รายงานวิจัยฉบับสมบูรณ์

สารสนเทศเชิงประยุกต์เพื่อนวัตกรรมสารสนเทศทางการแพทย์

Applied Informatics for Innovation of Medical Informatics

ปรัชญา นวนแก้ว และ คณะ

ได้รับทุนอุดหนุนการวิจัย ประจำปี พ.ศ. 2568

มหาวิทยาลัยพะเยา

รายงานวิจัยฉบับสมบูรณ์

สารสนเทศเชิงประยุกต์เพื่อนวัตกรรมสารสนเทศทางการแพทย์

Applied Informatics for Innovation of Medical Informatics

คณะผู้วิจัย

หัวหน้าโครงการ

ปรัชญา นวนแก้ว

คณะเทคโนโลยีสารสนเทศและการสื่อสาร

ได้รับทุนอุดหนุนการวิจัย ประจำปี พ.ศ. 2568

มหาวิทยาลัยพะเยา

1. บทสรุปผู้บริหาร (Executive Summary)

1.1 หลักการและเหตุผล

ระบบบริการสาธารณสุขไทยในปัจจุบันได้ขยายความครอบคลุมสถานบริการสาธารณสุขในทุกจังหวัดและมีหลักประกันสุขภาพถ้วนหน้าแก่ประชาชนไทย เพื่อให้เกิดความเป็นธรรมในการเข้าถึงบริการสาธารณสุขที่จำเป็น ทว่ายังพบความไม่เป็นธรรมในการเข้าถึงบริการสาธารณสุข เนื่องจากการขาดการกระจายเทคโนโลยีสารสนเทศพื้นฐาน การกระจายบุคลากรสาธารณสุขที่มีความรู้ความเข้าใจเทคโนโลยีและส่งเสริมเทคโนโลยีทางการแพทย์ที่เหมาะสม

เทคโนโลยีสารสนเทศเชิงประยุกต์และนวัตกรรมปัญญาประดิษฐ์ รวมถึงเทคโนโลยีทางการแพทย์ที่มีความก้าวหน้าอย่างรวดเร็วทำให้การดูแลรักษาผู้ป่วยมีประสิทธิภาพและสัมฤทธิ์ผลมากขึ้น ทว่าในขณะเดียวกันองค์กรและประชาชนต่างมีภาระค่าใช้จ่ายที่เพิ่มมากขึ้น ประกอบกับความเหลื่อมล้ำทางรายได้และโอกาสของประชาชนในท้องถิ่นชนบท จึงเป็นความท้าทายในการเพิ่มขีดความสามารถในการแข่งขันและพัฒนาเทคโนโลยีเพื่อสนับสนุนและให้บริการด้านสาธารณสุขแต่ประชาชนในพื้นที่ชนบท

การขับเคลื่อนดำเนินงานในระบบบริการสุขภาพต่อประชาชนให้มีประสิทธิภาพจำเป็นต้องใช้เทคโนโลยีสารสนเทศและการสื่อสาร ในการบริหารจัดการและสนับสนุนการดำเนินงาน องค์กรสำคัญต่างๆ ของประเทศต่างให้ความสำคัญต่อการพัฒนาเทคโนโลยีสารสนเทศ ได้แก่ กระทรวงสาธารณสุขได้กำหนดแผนแม่บทเทคโนโลยีสารสนเทศของประเทศไทยเพื่อพัฒนาระบบเทคโนโลยีสารนเทศและการสื่อสาร ได้กำหนดร่างยุทธศาสตร์เทคโนโลยีสารสนเทศสุขภาพ กระทรวงสาธารณสุข (255-2563) เพื่อกำหนดแนวปฏิบัติที่ดีและมุ่งประเด็นไปที่การพัฒนาระบบเทคโนโลยีสารสนเทศและการสื่อสารเพื่อสุขภาพประชาชน ตามแนวทาง eHealth ขององค์การอนามัยโลก (World Health Organization: WHO) และสหภาพโทรคมนาคมนานาชาติ (International Telecommunication Union: ITU) ที่หลากหลายประเทศได้นำมาเป็นกรอบแนวทางในการดำเนินงาน ซึ่งช่วยตอบโจทย์การทำงานที่ตรองต่อภาระงานด้านสุขภาพของประชากรโลกได้มากมาย โดยนัยสำคัญของการดำเนินงานตามเป้าหมายสำคัญของ WHO และ ITU คือมุ้งเป้าไปที่การส่งเสริมและดูแลสุขภาพของประชาชนเป็นหลัก โดยการนำระบบเทคโนโลยีสารสนเทศและการสื่อสาร (Information and Communication Technology: ICT) เพื่อแก้ปัญหาและสนับสนุนให้ประชาชนได้รับบริการด้านสุขภาพอย่างมีประสิทธิภาพ ทั่วถึง เป็นธรรมและปลอดภัย

ร่างยุทธศาสตร์เทคโนโลยีสารสนเทศสุขภาพ กระทรวงสาธารณสุข (2559-2563) ของประเทศไทยได้จัดทำการวิเคราะห์กรอบนโยบาย ยุทธศาสตร์ แผนงานด้านเทคโนโลยีสารสนเทศและการสื่อสารได้อย่างละเอียดโดยมีกรอบเป้าหมายการดำเนินงานได้แก่ การกำหนดกรอบนโยบายเทคโนโลยีสารสนเทศและการสื่อสารของประเทญไทย พ.ศ. 2554-2563 (National ICT Policy Framework 2011-2020: ICT 2020) การกำหนดแผนแม่บทด้านเทคโนโลยีสารสนเทศและการสื่อสารของประชาคมอาเซียน (ASEAN ICT Master Plan) ยุทธศาสตร์การบูรณาการรัฐบาลอิเล็กทรอนิกส์ (eGovernment) การขับเคลื่อนเศรษฐกิจและสังคมดิจิทัล ระยะเร่งด่วนปี 2558 และแนวคิด Digital Health จากประเทศต่างๆ ประกอบกับแผนและสภาพแวดล้อมของ eHealth ในประเทศไทยตามองค์ประกอบของ eHealth (Regional Strategy for Strengthening eHealth in South East Asia Region 2558-2564)

นอกจากนี้ เมื่อพิจารณาจากสถานการณ์ปัญหาสถานะสุขภาพของประชาชน ยุทธศาสตร์และเป้าหมายในแผนพัฒนาสุขภาพแห่งชาติ ฉบับที่ 12 (พ.ศ. 2560-2564) โดยได้กำหนดยุทธศาสตร์การพัฒนาสุขภาพ กำหนดไว้ 4 ยุทธศาสตร์ แต่ละยุทธศาสตร์ได้กำหนดวัตถุประสงค์ ตัวชี้วัด มาตรการและแนวทางการพัฒนาไว้อย่างละเอียด

โดยยุทธศาสตร์ที่ 1 คือการเร่งการเสริมสร้างสุขภาพคนไทยเชิงรุก ประกอบด้วยวัตถุประสงค์สำคัญ 3 ประการ ได้แก่ 1) เพื่อสร้างความเข้มแข็งของบุคคล ชุมชน ประชาชน องค์กรปกครองส่วนท้องถิ่น ภาคีเครือข่ายภาคประชาชนและภาคประชาสังคมด้านสุขภาพ ให้มีศักยภาพ มีความรู้และทัศนคติที่ถูกต้องด้านพฤติกรรมสุขภาพ มีการเรียนรู้ มีพฤติกรรมเสี่ยงทางสุขภาพลดลง สามารถช่วยเหลือ ดูแลตนเองและชุมชน ตลอดจนมีส่วนร่วมในการสร้างและจัดการระบบสุขภาพ 2) เพื่อสร้างระบบสุขภาพเชิงรุก ที่มุ่งสร้างเสริมให้คนไทยทุกช่วงวัยมีสุขภาพดี มีระบบการป้องกันควบคุมโรคและปัจจัยเสี่ยงด้านสุขภาพ มีการคุ้มครองผู้บริโภคด้านสุขภาพที่ดี มีสภาพแวดล้อมที่เอื้อต่อการมีสุขภาพดี 3) เพื่อพัฒนาระบบการดูแลผู้สูงอายุระยะยาวที่เกิดจากความร่วมมือของครอบครัว ชุมชน และสถานพยาบาล ให้มีความพอเพียงและเหมาะสมต่อการเข้าถึงบริการของผู้สูงอายุ ส่งผลให้ผู้สูงอายุมีคุณภาพชีวิตที่ดี เข้าสู่สังคมสูงวัยได้อย่างมีความสุข

ยุทธศาสตร์ที่ 2: สร้างความเป็นธรรม ลดความเหลื่อมล้ำในระบบบริการสุขภาพ ประกอบด้วยวัตถุประสงค์ 4 ข้อ ดังนี้ 1) เพื่อสร้างและพัฒนาระบบบริการปฐมภูมิที่มีแพทย์เวชศาสตร์ครอบครัวประจำให้ครบทุกแห่งเพื่อเพิ่มความเป็นธรรมในการเข้าถึงบริการและคุณภาพในการดูแลประชาชน 2) เพื่อยกระดับขีดความสามารถของหน่วยบริการทุกระดับให้มีคุณภาพ ตามมาตรฐาน มีความเป็นธรรมตอบสนองต่อความต้องการของประชาชนและสามารถเข้าถึงบริการได้อย่างทั่วถึง ทัดเทียมกัน 3) เพื่อ

สร้างความเข้มแข็งและสามารถในการแข่งขันให้กับระบบสุขภาพของประเทศ 4) เพื่อลดความเหลื่อมล้ำในการได้รับบริการด้านการแพทย์และสาธารณสุข

ยุทธศาสตร์ที่ 3: พัฒนาและสร้างกลไกเพื่อเพิ่มประสิทธิภาพการบริหารจัดการกำลังคนด้านสุขภาพ โดยมีวัตถุประสงค์ 4 ข้อ ได้แก่ 1) เพื่อวางแผนกำลังคนด้านสุขภาพที่สอดรับกับการออกแบบระบบสุขภาพและความจำเป็นด้านสุขภาพของประชาชนแต่ละพื้นที่และทิศทางระบบสุขภาพของประเทศ 2) เพื่อบูรณาการระบบการผลิตกำลังคนด้านสุขภาพของประเทศให้ตั้งอยู่บนฐานของความร่วมมือระหว่างผู้ผลิต และผู้ใช้กำลังคนด้านสุขภาพทั้งภาครัฐและเอกชน 3) เพื่อสร้างกลไกและระบบการบริหารจัดการก ำลังคนด้านสุขภาพ รวมทั้งระบบการติดตามและประเมินผลการบริหารจัดการกำลังคนด้านสุขภาพในทุกระดับ 4) เพื่อสร้างเครือข่ายกำลังคนด้านสุขภาพ ที่ประกอบด้วยภาครัฐ เอกชน องค์กรปกครองส่วนท้องถิ่น และภาคประชาชน

ยุทธศาสตร์ที่ 4: พัฒนาและสร้างความเข้มแข็งในการอภิบาลระบบสุขภาพ ประกอบด้วย 2 วัตถุประสงค์ ได้แก่ 1) เพื่ออภิบาลระบบสุขภาพอย่างมีธรรมาภิบาล เป็นเอกภาพ อันจะส่งผลให้มีความมั่นคง ยั่งยืนของระบบสุขภาพ 2) เพื่อพัฒนาระบบสนับสนุนการบริการสุขภาพ ระบบข้อมูลข่าวสารด้านสุขภาพ การเงินการคลังด้านสุขภาพ รวมถึงยาและเทคโนโลยีด้านสุขภาพ

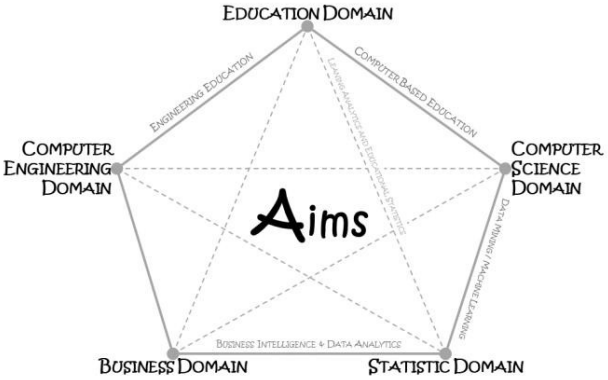
ด้วยความสำคัญ ที่มาของปัญหา และเหตุผลต่าง ๆ เหล่านี้ ทำให้นักวิจัยเชื่อว่า โครงการวิจัยเรื่อง สารสนเทศเชิงประยุกต์เพื่อนวัตกรรมสารสนเทศทางการแพทย์ (Applied Informatics for Innovation of Medical Informatics) จะสามารถตอบโจทย์ปัญหาและตอบสนองต่อยุทธศาสตร์เทคโนโลยีสารสนเทศสุขภาพ กระทรวงสาธารณสุข ตอบโจทย์แนวปฏิบัติที่ดีและมุ่งประเด็นไปที่การพัฒนาระบบเทคโนโลยีสารสนเทศและการสื่อสารเพื่อสุขภาพประชาชน ตามแนวทาง eHealth ขององค์การอนามัยโลก (World Health Organization: WHO) และตอบโจทย์ยุทธศาสตร์และเป้าหมายตามแผนพัฒนาสุขภาพแห่งชาติ ฉบับที่ 12 (พ.ศ. 2560-2564) ทั้ง 4 ยุทธศาสตร์ โดยทีมนักวิจัยมีกรอบการดำเนินงานในการนำเทคโนโลยีสารสนเทศเชิงประยุกต์มาปรับใช้ต่อการพัฒนานวัตกรรมสารสนเทศทางการแพทย์ โดยการทำเทคโนโลยีปัญญาประดิษฐ์ เทคโนโลยีสารสนเทศและการสื่อสาร เทคโนโลยีและนวัตกรรมอื่นๆ ที่เกี่ยวข้องมาส่งเสริมและพัฒนาตามที่ได้กำหนดแนวทางไว้ และจากการสืบค้นเบื้องต้นยังไม่พบว่ามีงานวิจัยใดที่ทำการวิจัยเรื่องนี้ ผู้วิจัยจึงมีความสนใจที่จะศึกษาวิจัย อันจะส่งผลให้เกิดประโยชน์และพัฒนาผู้เรียนอย่างมีประสิทธิภาพ ต่อไป

1.2 ปัญหา โจทย์การวิจัย และโจทย์การพัฒนานวัตกรรม

ในยุคที่เทคโนโลยีมีอิทธิพลต่อการดำรงชีวิตมนุษย์ ความก้าวหน้าของเทคโนโลยีส่งผลให้เกิดการเปลี่ยนแปลงอย่างต่อเนื่องและรุนแรง ความพร้อมเพื่อที่จะสร้างโอกาสบนเส้นทางสายต่างๆ ไม่ว่าจะเป็นการใช้ชีวิตประจำวัน การประกอบอาชีพ การศึกษารูปแบบใหม่ ธุรกิจสมัยใหม่ เกษตรอัจฉริยะ วิศวกรรมคอมพิวเตอร์ ปัญญาประดิษฐ์ ไปจนถึงเศรษฐกิจและนโยบายระดับชาติ ต่างมีความสัมพันธ์และได้รับผลกระทบจากการเปลี่ยนแปลงและเทคโนโลยีทั้งสิ้น เหล่านี้ทำให้ผู้คนหันมาเปลี่ยนแปลงและปรับเปลี่ยนเพื่อที่จะได้สามารถปรับตัวให้อยู่รอดในยุคที่ Disruptive Technology หรือ หนึ่งในการเปลี่ยนแปลงจนทำให้เกิดผลกระทบอย่างรอบด้าน Disruptive Technology คือ เทคโนโลยี หรือนวัตกรรมที่เข้ามาสร้างองค์ความรู้ ความรู้ใหม่ วิถีใหม่ สร้างตลาดและมูลค่าให้กับผลิตภัณฑ์ที่ใช้เทคโนโลยี จนทำให้เกิดผลกระทบอย่างรุนแรงและทำให้ผลิตภัณฑ์เดิม หรือองค์ความรู้เดิมถูก Disrupt ไป โดยที่มีการเปลี่ยนแปลงทั้งด้านกระบวนการ ด้านคุณภาพ ด้านประสิทธิภาพ ด้านการผลิตและการขาย หรือแม้กระทั่งการปรับเปลี่ยนกระบวนการสร้างองค์ความรู้ สร้างผลิตภัณฑ์ใหม่ที่แตกต่างจากการผลิตรูปแบบเดิม ส่งผลให้มีการเปลี่ยนแปลงอย่างรุนแรงและมากยิ่งขึ้น ได้แก่ การนำระบบปัญญาประดิษฐ์ (Artificial Intelligence: AI) มาประยุกต์ใช้เพื่อการพยากรณ์ การคาดการณ์ และการตั้งสมมุติฐาน หรือการนำระบบอัตโนมัติเข้ามาใช้ในกระบวนการผลิตสินค้า ตรวจนับสินค้า การเก็บข้อมูลเอาไว้ในระบบคลาวด์ (Cloud Storage) แทนการเก็บเป็นเอกสาร เป็นต้น

เมื่อกล่าวถึงเทคโนโลยีที่จะเข้ามามีอิทธิพลในปัจจุบันและอนาคตนั้น หนีไม่พ้นกับการเปลี่ยนแปลงด้านเทคโนโลยีอินเทอร์เน็ตไร้สายและอินเทอร์เน็ตความเร็วสูง (Wireless Broadband and High Speed Technologies) เทคโนโลยีการวิเคราะห์ข้อมูลขนาดใหญ่ (Big Data Analytics Technologies) เทคโนโลยีหุ่นยนต์ (Robotics Technology) เทคโนโลยีชีวภาพ (Biotechnology) เทคโนโลยีการพิมพ์ภาพสามมิติ (3D Printing Technology) เทคโนโลยีการเก็บกักพลังงานและพลังงานทดแทน (Energy Storage Technology and Renewable Energy) และอื่น ๆ มากมาย เห็นได้ว่า Disruptive Technology คือ กระบวนการที่ทำให้เกิดการสร้างอาชีพใหม่ (New Areas) สร้างองค์ความรู้ใหม่ และสร้างวิถีใหม่ (New Normal) ดังนั้น องค์กรต่างๆ จะเอาตัวรอดจาก Disruptive Technology ได้ องค์กรจำเป็นต้องมีการศึกษาและสังเกต การเปลี่ยนแปลงอย่างต่อเนื่อง เพื่อปรับตัวให้เข้ากับเทคโนโลยีที่เกี่ยวข้อง และตอบสนองต่อการเปลี่ยนแปลงของเทคโนโลยีอย่างเหมาะสม สิ่งสำคัญสำหรับองค์กร คือ การเตรียมตัวและปรับตัวให้ดียิ่งขึ้น

โดยโครงการวิจัยนี้ ได้เสนอกรอบเทคโนโลยีการศึกษาใหม่ ดังแสดงกรอบความเชื่อมโยงความสัมพันธ์ และความเกี่ยวข้องเทคโนโลยีและศาสตร์ต่าง ๆ ในการเข้าสู่ยุค Disruptive Technology ดังแสดงในภาพ



จากกรอบเทคโนโลยีการศึกษาใหม่ (New Educational Technology Domains Era) แสดงให้เห็นองค์ประกอบ 5 ศาสตร์สำคัญ (Domains) และ 5 ศาสตร์ย่อย (Sub-Domains) ประกอบด้วย ศาสตร์ด้านการศึกษา (Education Domain) ศาสตร์ด้านวิศวกรรมคอมพิวเตอร์ (Computer Engineering Domain) ศาสตร์ด้านวิทยาการคอมพิวเตอร์ (Computer Science Domain) ศาสตร์ด้านสถิติ (Statistical Domain) ศาสตร์ด้านธุรกิจ (Business Domain) ศาสตร์ด้านการประยุกต์คอมพิวเตอร์เพื่อการศึกษา (Computer Based Education Domain) ศาสตร์ด้านเหมืองข้อมูลและการเรียนรู้ด้วยเครื่อง (Data Mining and Machine Learning Domain) ศาสตร์ด้านวิศวกรรมการศึกษา (Engineering Education) ศาสตร์ด้านคอมพิวเตอร์ธุรกิจ (Business Computer Domain) และศาสตร์ด้านสถิติการประยุกต์เพื่อการศึกษา (Leaning Analytics and Educational Statistics Domain) ทั้งหมด คือ องค์ประกอบและส่วนสำคัญ เพื่อการมุ่งสู่เป้าหมายของโครงการวิจัยสารสนเทศเชิงประยุกต์เพื่อนวัตกรรมสารสนเทศทางการแพทย์ คณะเทคโนโลยีสารสนเทศและการสื่อสาร มหาวิทยาลัยพะเยา

เป้าหมาย :

เป้าหมายสำคัญของโครงการวิจัยสารสนเทศเชิงประยุกต์เพื่อนวัตกรรมสารสนเทศทางการแพทย์ของคณะเทคโนโลยีสารสนเทศและการสื่อสาร มหาวิทยาลัยพะเยา ประกอบด้วย 2 เป้าหมายหลัก ได้แก่ เป้าหมายที่ 1 การสร้างความเป็นเลิศด้านวิชาการ ด้านการวิจัย ตามยุทธศาสตร์การพัฒนา มหาวิทยาลัยพะเยา (2562-2565) เป้าหมายที่ 2 การบรรลุเป้าหมายของโครงการวิจัยสารสนเทศเชิงประยุกต์เพื่อ

นวัตกรรมสารสนเทศทางการแพทย์ คณะเทคโนโลยีสารสนเทศและการสื่อสาร มหาวิทยาลัยพะเยา โดยมีรายละเอียด ดังนี้

เป้าหมายที่ 1 การสร้างความเป็นเลิศด้านวิชาการ ด้านการวิจัย ตามยุทธศาสตร์การพัฒนามหาวิทยาลัยพะเยา (2562-2565) ประกอบด้วย 6 ยุทธศาสตร์ ดังนี้

ยุทธศาสตร์ที่ 1 การเตรียมคนและเสริมสร้างศักยภาพคน

ยุทธศาสตร์ที่ 2 การสร้างงานวิจัยและนวัตกรรมและการเป็นผู้นำทางด้านวิชาการ

ยุทธศาสตร์ที่ 3 การบริการวิชาการเพื่อพัฒนาชุมชน สังคม และประเทศ

ยุทธศาสตร์ที่ 4 การเสริมสร้างและส่งเสริมการทำนุบำรุงศิลปะและวัฒนธรรมให้ยั่งยืน

ยุทธศาสตร์ที่ 5 การเสริมสร้างและพัฒนาความเป็นสากลหรือนานาชาติ

ยุทธศาสตร์ที่ 6 การบริหารที่มีประสิทธิภาพและโปร่งใส

เป้าหมายที่ 2 การบรรลุเป้าหมายของโครงการวิจัยสารสนเทศเชิงประยุกต์เพื่อนวัตกรรมสารสนเทศทางการแพทย์ คณะเทคโนโลยีสารสนเทศและการสื่อสาร มหาวิทยาลัยพะเยา ประกอบด้วย 5 องค์ประกอบ ดังนี้

องค์ประกอบที่ 1 การนำเสนอผลงานตีพิมพ์และเผยแพร่ผลงานวิจัยในระดับนานาชาติ

องค์ประกอบที่ 2 การใช้ประโยชน์ผลงานวิจัย

องค์ประกอบที่ 3 การแสวงหาแหล่งทุนวิจัยภายนอก

องค์ประกอบที่ 4 การสร้างนักวิจัยรุ่นใหม่

องค์ประกอบที่ 5 การสนับสนุนการผลิตบัณฑิต หรือการบูรณาการกับการเรียนการสอน

1.3 วัตถุประสงค์

1.3.1 เพื่อส่งเสริมและสร้างเครือข่ายนักวิจัยด้านสารสนเทศเชิงประยุกต์เพื่อนวัตกรรมสารสนเทศทางการแพทย์

1.3.2 เพื่อสร้างแหล่งการเรียนรู้และสนับสนุนนักวิจัย นักวิชาการ ด้านสารสนเทศเชิงประยุกต์เพื่อนวัตกรรมสารสนเทศทางการแพทย์

1.3.3 เพื่อเผยแพร่นวัตกรรม เทคโนโลยี องค์ความรู้ และผลงานวิจัยในระดับชาติ ระดับนานาชาติด้านสารสนเทศเชิงประยุกต์เพื่อนวัตกรรมสารสนเทศทางการแพทย์

1.4 กรอบการวิจัย และพัฒนา

กรอบการวิจัยแสดงเป้าหมายและตัวชี้วัดโครงการวิจัย ซึ่งมีการแสดงความเชื่อมโยงโครงการย่อยเพื่อตอบเป้าหมายและวัตถุประสงค์โครงการ ดังแสดงในภาพ

กรอบการวิจัย:
โครงการวิจัยสารสนเทศเชิงประยุกต์เพื่อนวัตกรรมสารสนเทศทางการแพทย์

กรอบการวิจัย โครงการวิจัยสารสนเทศเชิงประยุกต์เพื่อนวัตกรรมสารสนเทศทางการแพทย์ ของคณะเทคโนโลยีสารสนเทศและการสื่อสาร มหาวิทยาลัยพะเยา มีกรอบการดำเนินงานสำคัญ สำหรับปีงบประมาณ 2568 โดยกรอบเป้าหมายการวิจัยที่ 1 เพื่อมุ่งเป้าตามวัตถุประสงค์ของโครงการ ซึ่งประกอบด้วย 3 ข้อสำคัญ ได้แก่ 1) เพื่อส่งเสริมและสร้างเครือข่ายนักวิจัยด้านสารสนเทศเชิงประยุกต์เพื่อนวัตกรรมสารสนเทศทางการแพทย์ 2) เพื่อสร้างแหล่งการเรียนรู้และสนับสนุนนักวิจัย นักวิชาการ ด้านสารสนเทศเชิงประยุกต์เพื่อนวัตกรรมสารสนเทศทางการแพทย์ 3) เพื่อเผยแพร่นวัตกรรม เทคโนโลยี องค์ความรู้ และผลงานวิจัยในระดับชาติ ระดับนานาชาติด้านสารสนเทศเชิงประยุกต์เพื่อนวัตกรรมสารสนเทศทางการแพทย์ โดยกรอบเป้าหมายการวิจัยที่ 1 นี้ มีความสำคัญต่อการขยายผลการดำเนินงานของกรอบการวิจัยที่ 2 โดยกรอบการวิจัยที่ 2 มุ่งเป้าเพื่อการพัฒนากิจกรรมย่อย ซึ่งประกอบด้วย 2 กิจกรรมย่อย ได้แก่ กิจกรรมย่อยที่ 1 ระบบสารสนเทศเพื่อการบริหารการจัดการการดูแลสุขภาวะในชุมชนพื้นที่ห่างไกล ด้วยสารสนเทศเชิงประยุกต์และเทคโนโลยีปัญญาประดิษฐ์ กิจกรรมย่อยที่ 2 การพัฒนานักพัฒนานวัตกรรมสารสนเทศทางการแพทย์ ผลลัพธ์ที่ได้จากกรอบการวิจัยที่ 2 คือ การสร้างเครือข่ายความร่วมมือ นักวิจัย และสถาบันการศึกษาต่างๆ และกรอบเป้าหมายการวิจัยที่ 3 คือ การเผยแพร่ผลงานการวิจัยที่ได้ค้นพบ โดยการนำเสนอในวารสารที่ได้รับการยอมรับ ระดับชาติ หรือ ระดับนานาชาติต่อไป

1.5 เป้าหมายผลงานที่วางไว้ (Committed Targets/Outputs)

ตารางที่ 1 เป้าหมายผลงานที่วางไว้ (Committed Targets/Outputs)

ประเภทของค่าเป้าหมาย		**เป้าหมายที่โครงการกำหนด (Premier League)**
1. ผลงานตีพิมพ์/เผยแพร่	1.1 ผลงานตีพิมพ์ระดับนานาชาติ (ISI/Scopus: Impact factor /Quartile)	- มีผลงานตีพิมพ์ระดับนานาชาติ 5 ผลงาน
	1.2 ผลงานตีพิมพ์ระดับนานาชาติ ที่ไม่มีค่า Impact factor หรือไม่อยู่ในฐานข้อมูล ISI/Scopus /Quartile	-
	1.3 ผลงานตีพิมพ์ระดับชาติ (TCI กลุ่ม 1 และ 2)	-
	1.4 ผลงานตีพิมพ์/เผยแพร่ในรูปแบบ Proceedings	-
2. การใช้ประโยชน์ผลงานวิจัย	2.1 มีการรับรองการใช้ประโยชน์ผลงานวิจัย	-
3. การแสวงหาแหล่งทุนวิจัยภายนอก	3.1 มีการแสวงหาแหล่งทุนวิจัยภายนอก (เพิ่มเติมจาก ทุนวิจัย UoE)	-
4. การสร้างนักวิจัยรุ่นใหม่	4.1 มีสร้างนักวิจัยรุ่นใหม่	1 คน
5. การสนับสนุนการผลิตบัณฑิต หรือการบูรณาการกับการเรียนการสอน	5.1 มีส่วนร่วมในการผลิตบัณฑิต หรือการบูรณาการกับการเรียนการสอน	-

1.6 สรุปผลการดำเนินงานโดยภาพรวม

โครงการ “สารสนเทศเชิงประยุกต์เพื่อนวัตกรรมสารสนเทศทางการแพทย์” ได้ดำเนินโครงการเสร็จสิ้นแล้ว โดยมีรายงานการสรุป 3 ประเด็นหลัก ได้แก่ การประเมินผลการดำเนินโครงการตาม KPI สรุปข้อมูลการดำเนินกิจกรรม และ สรุปข้อมูลการใช้จ่ายงบประมาณ โดยมีรายละเอียด ดังนี้

1.6.1 สรุปผลการประเมินผลการดำเนินโครงการตาม KPI

ตารางที่ 2 สรุปผลการประเมินผลการดำเนินโครงการตาม KPI

ประเด็น	ค่าเป้าหมาย	ผลลัพธ์	KPI %
- กำลังคน หรือ หน่วยงานที่ได้รับการพัฒนาทักษะ - นักวิจัยหน่วยงานรัฐ	1 คน	1 คน	100%
- ต้นฉบับบทความวิจัย (Manuscript)	5 เรื่อง	5 เรื่อง	100%

ผลการดำเนินโครงการ “สารสนเทศเชิงประยุกต์เพื่อนวัตกรรมสารสนเทศทางการแพทย์” โครงการมีค่าเฉลี่ยผลลัพธ์ KPI รวม 100% โดยโครงการฯ สามารถบรรลุทุกวัตถุประสงค์ที่ตั้งไว้

1.6.2 สรุปข้อมูลการดำเนินกิจกรรม

ตารางที่ 3 สรุปข้อมูลการดำเนินกิจกรรม

ชื่อกิจกรรม	ระยะเวลาดำเนินการ	ผลดำเนินการ
กิจกรรมที่ 1: ศึกษาและสำรวจความเป็นไปได้	10/2567 – 12/2567	เสร็จสิ้น
กิจกรรมที่ 2: สร้างเครือข่ายความร่วมมือ เครือข่ายการวิจัย ทั้งภายในและภายนอกมหาวิทยาลัยพะเยา	12/2567 – 6/2568	เสร็จสิ้น
กิจกรรมที่ 3: จัดตั้งแหล่งเรียนรู้และสนับสนุนนักวิจัย นักวิชาการ	2/2568 – 8/2568	เสร็จสิ้น
กิจกรรมที่ 4: เผยแพร่เทคโนโลยีใหม่ องค์ความรู้ใหม่ และผลงานวิจัยในระดับชาติ ระดับนานาชาติ	2/2568 – 8/2568	เสร็จสิ้น
กิจกรรมที่ 5: ติดตามประเมินการดำเนินงานของหน่วยวิจัยตามแผน	12/2568 – 10/2568	เสร็จสิ้น
กิจกรรมที่ 6: สรุปผลและจัดทำรายงาน	3/2568, 6/2568, 9/2568	เสร็จสิ้น

ภาพรวมแผนการดำเนินงานของโครงการ “ สารสนเทศเชิงประยุกต์เพื่อนวัตกรรมสารสนเทศทางการแพทย์” บรรลุวัตถุประสงค์ทุกข้อ

1.6.3 สรุปข้อมูลการใช้จ่ายงบประมาณ

ตารางที่ 4 สรุปข้อมูลการใช้จ่ายงบประมาณ

รายการ	ยอดจัดสรร	ใช้ไป	คงเหลือ
1. งบบุคลากร	0.00	0.00	0.00
2. งบดำเนินงาน			
2.1 ค่าจ้าง	340,000	450,000	0.00
2.2 ค่าวัสดุ	160,000	50,000	0.00
2.3 ค่าใช้สอย	0.00	0.00	0.00
3. งบลงทุน	0.00	0.00	0.00
รวมทั้งสิ้น	**500,000.00**	**500,000.00**	**0.00**

ภาพรวมของข้อมูลการใช้จ่ายงบประมาณของการดำเนินโครงการ “ สารสนเทศเชิงประยุกต์เพื่อนวัตกรรมสารสนเทศทางการแพทย์” บรรลุทุกวัตถุประสงค์

****

2. ผลงานวิจัยโดยสรุป

ผลงานวิจัยของโครงการ “ สารสนเทศเชิงประยุกต์เพื่อนวัตกรรมสารสนเทศทางการแพทย์” ได้บรรลุตามวัตถุประสงค์ ทั้ง 3 ข้อ โดย โครงการสารสนเทศเชิงประยุกต์เพื่อนวัตกรรมสารสนเทศทางการแพทย์ คณะเทคโนโลยีสารสนเทศและการสื่อสาร มหาวิทยาลัยพะเยา ได้ส่งเสริมและสร้างเครือข่ายความร่วมมือ เครือข่ายการวิจัย ทั้งภายในและภายนอกมหาวิทยาลัยพะเยา ในศาสตร์ที่เกี่ยวข้อง โดยมีรายชื่อเครือข่าย ดังนี้

- ดร.วงษ์ปัญญา นวนแก้ว มหาวิทยาลัยพะเยา ผู้ร่วมวิจัย
- รศ.ดร.พรรณฤมล เต็มดี มหาวิทยาลัยแม่ฟ้าหลวง ที่ปรึกษาโครงการ
- รศ.ดร.สิทธิชัย บุษหมั่น มหาวิทยาลัยราชภัฏมหาสารคาม ที่ปรึกษาโครงการ
- รศ.ดร.ทิพรัตน์ สิทธิวงศ์ มหาวิทยาลัยนเรศวร ที่ปรึกษาโครงการ
- ผศ.ดร.ฆณการ ภัณณิพงส์ มหาวิทยาลัยเทคโนโลยีราชมงคลตะวันออก ที่ปรึกษาโครงการ
- ผศ.ดร.ทิพวิมล ชมภูคำ มหาวิทยาลัยราชภัฏมหาสารคาม ที่ปรึกษาโครงการ
- ผศ.ดร.ณิฏฐะญาร์ บรรเทา มหาวิทยาลัยราชภัฏมหาสารคาม ที่ปรึกษาโครงการ
- ดร.พัชระ นาเสงี่ยม มหาวิทยาลัยราชภัฏมหาสารคาม ที่ปรึกษาโครงการ
- ดร.อุมาพร ไชยสูง มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน ที่ปรึกษาโครงการ

นอกจากนั้น โครงการ” สารสนเทศเชิงประยุกต์เพื่อนวัตกรรมสารสนเทศทางการแพทย์” ได้ตั้งแหล่งเรียนรู้และสนับสนุนนักวิจัย นักวิชาการ ในศาสตร์ที่เกี่ยวข้อง เผยแพร่เทคโนโลยีใหม่ องค์ความรู้ใหม่ และผลงานวิจัยในระดับชาติ ระดับนานาชาติ ในศาสตร์ที่เกี่ยวข้อง ผ่านทางเว็บไซต์ http://www.aims.or.th โดยโครงการได้สรุปผลงานที่ได้รับการเผยแพร่ทั้ง 5 ผลงาน ซึ่งทั้ง 5 ผลงานประกอบด้วย

[FF68 – UoE – 5028 – P05] “Classification Model for Screening Knee Osteoarthritis Patients in Northern Thailand Using Data Analytics” to The 5th International Conference on Data Engineering and Communication Technology (ICDECT) at https://icdect.com/

[FF68 – UoE – 5028 – P04] “Sign Language Detection Mobile Application for Thai Patients Using Medical Image Processing to Support Medical Consultations” to The 3rd World Conference on Information Systems for Business Management (ISBM’ 2024) at https://isbm.ict4sd.org

[FF68 – UoE – 5028 – P03] “AI for Healthcare: A Classification Model for Personalized Premenstrual Symptoms and Depressive Crisis Risk Tracking Using Data Analytics and Machine Learning” to The 10th International Conference on Digital Arts, Media and Technology (DAMT) and 8th ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (NCON): ECTI DAMT & NCON 2025 at https://www.icdamt.org/

[FF68 – UoE – 5028 – P02] Buntao, N., Somkhuean, R., Nuankaew, P., & Nuankaew S., W. (2025). Evolutionary Feature Selection to Classify Elderly Diseases from Dietary and Exercise Habits and Emotions. International Journal of Engineering Trends and Technology – IJETT, 73(1), 166–176. https://doi.org/10.14445/22315381/IJETT-V73I1P114

[FF68 – UoE – 5028 – P01] Chompookham, T., Nuankaew, W. S., & Nuankaew, P. (2024). Two-Factor Authentication Application Using Artificial Intelligence to Support Academic Information Systems. International Journal of Engineering Trends and Technology – IJETT, 72(12), 14–29. https://doi.org/10.14445/22315381/IJETT-V72I12P102

ดังที่รายงานข้อมูลมาข้างต้นแสดงให้เห็นว่า โครงการ ” สารสนเทศเชิงประยุกต์เพื่อนวัตกรรมสารสนเทศทางการแพทย์” ได้บรรลุวัตถุประสงค์ทุกประการ

3. ตารางเปรียบเทียบผลงานที่ทำได้จริง กับผลงานตามเป้าหมายที่วางไว้ (Committed Targets/Outputs) โดยสรุป

โครงการ “ สารสนเทศเชิงประยุกต์เพื่อนวัตกรรมสารสนเทศทางการแพทย์” ได้ดำเนินงานตามเป้าหมายที่วางไว้ ดังแสดงในตารางที่ 1 เป้าหมายผลงานที่วางไว้ (Committed Targets/Outputs) โดยผลสรุปของการดำเนินโครงการถูกสรุป ดังแสดงในตารางที่ 2 สรุปผลการประเมินผลการดำเนินโครงการตาม KPI และการเปรียบเทียบของการดำเนินงานของโครงการ ได้แสดงในตารางที่ 5 ตารางเปรียบเทียบผลงานที่ทำได้จริง กับผลงานตามเป้าหมายที่วางไว้ (Committed Targets/Outputs) ดังนี้

ตารางที่ 5 ตารางเปรียบเทียบผลงานที่ทำได้จริง กับผลงานตามเป้าหมายที่วางไว้ (Committed Targets/Outputs)

ประเภทของค่าเป้าหมาย		เป้าหมาย	ผลลัพธ์
1. ผลงานตีพิมพ์/เผยแพร่	1.1 ผลงานตีพิมพ์ระดับนานาชาติ (ISI/Scopus: Impact factor /Quartile)	- มีผลงานตีพิมพ์ระดับนานาชาติ 5 ผลงาน	- บรรลุเป้าหมายโดยมีผลงานตีพิมพ์ระดับนานาชาติ ถ ผลงาน ดังนี้ [FF68 – UoE – 5028 – P05] “Classification Model for Screening Knee Osteoarthritis Patients in Northern Thailand Using Data Analytics” to The 5th International Conference on Data Engineering and Communication Technology (ICDECT) at https://icdect.com/ [FF68 – UoE – 5028 – P04] “Sign Language Detection Mobile Application for Thai Patients Using Medical Image Processing to Support Medical Consultations” to The 3rd World Conference on Information Systems for Business Management (ISBM’ 2024) at https://isbm.ict4sd.org

ประเภทของค่าเป้าหมาย		เป้าหมาย	ผลลัพธ์
			[FF68 – UoE – 5028 – P03] “AI for Healthcare: A Classification Model for Personalized Premenstrual Symptoms and Depressive Crisis Risk Tracking Using Data Analytics and Machine Learning” to The 10th International Conference on Digital Arts, Media and Technology (DAMT) and 8th ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (NCON): ECTI DAMT & NCON 2025 at https://www.icdamt.org/ [FF68 – UoE – 5028 – P02] Buntao, N., Somkhuean, R., Nuankaew, P., & Nuankaew S., W. (2025). Evolutionary Feature Selection to Classify Elderly Diseases from Dietary and Exercise Habits and Emotions. International Journal of Engineering Trends and

ประเภทของค่าเป้าหมาย		เป้าหมาย	ผลลัพธ์
			Technology – IJETT, 73(1), 166–176. https://doi.org/10.14445/22315381/IJETT-V73I1P114 [FF68 – UoE – 5028 – P01] Chompookham, T., Nuankaew, W. S., & Nuankaew, P. (2024). Two-Factor Authentication Application Using Artificial Intelligence to Support Academic Information Systems. International Journal of Engineering Trends and Technology – IJETT, 72(12), 14–29. https://doi.org/10.14445/22315381/IJETT-V72I12P102
2. การใช้ประโยชน์ผลงานวิจัย	2.1 มีการรับรองการใช้ประโยชน์ผลงานวิจัย	- มีผลงานวิจัยที่ได้รับการอ้างอิง (Citation)	- บรรลุเป้าหมายโดยมีผลงานวิจัยที่ได้รับการอ้างอิง (Citation) ดังแสดงในเอกสารประกอบและเอกสารภาคผนวก (อื่นๆ)
3. การแสวงหาแหล่งทุนวิจัยภายนอก	3.1 มีการแสวงหาแหล่งทุนวิจัยภายนอก (เพิ่มเติมจากทุนวิจัย UoE)	- จำนวนโครงร่างการวิจัยที่ได้เสนอขอรับทุนในระดับชาติ หรือระดับนานาชาติ 1 โครงการ	- บรรลุเป้าหมายโดยมีโครงร่างการวิจัยที่ได้เสนอขอรับทุนในระดับชาติ หรือระดับนานาชาติ 1 โครงการ และได้รับจัดสรรทุนวิจัย 1 โครงการ ดังนี้

ประเภทของค่าเป้าหมาย		เป้าหมาย	ผลลัพธ์
			• โครงการ “สารสนเทศเชิงประยุกต์และนวัตกรรมสารสนเทศทางการแพทย์เพื่อการดูแลสุขภาพผู้สูงอายุ” งบประมาณด้าน ววน. ประเภท Fundamental Fund ประจำปีงบประมาณ 2569 (ผ่านหน่วยงาน) ดังแสดงในเอกสารประกอบและเอกสารภาคผนวก (อื่นๆ)
4. การสร้างนักวิจัยรุ่นใหม่	4.1 มีสร้างนักวิจัยรุ่นใหม่	-	-
5. การสนับสนุนการผลิตบัณฑิต หรือการบูรณาการกับการเรียนการสอน	5.1 มีส่วนร่วมในการผลิตบัณฑิต หรือการบูรณาการกับการเรียนการสอน	- มีนิสิตระดับปริญญาตรี หรือนิสิตระดับบัณฑิตศึกษา ตีพิมพ์และเผยแพร่ผลงานวิจัยในระดับชาติ หรือระดับนานาชาติ 1 ผลงาน	- บรรลุเป้าหมาย โดยมีนิสิตในระดับปริญญาตรี จำนวน 8 คนได้รับการสนับสนุนและส่งเสริมให้เข้าร่วมโครงการวิจัย นำโจทย์และปัญหาที่โครงการวิจัยได้ออกแบบไปดำเนินการประยุกต์ในโครงงานนิสิต จำนวน 4 โครงงาน ระหว่างปีการศึกษา 2567-2568

4. ผลงานอื่นๆ / เครือข่ายงานวิจัย โดยสรุป

โครงการ “ สารสนเทศเชิงประยุกต์เพื่อนวัตกรรมสารสนเทศทางการแพทย์” ได้สร้างเครือข่ายงานวิจัย ดังมีรายชื่อนักวิจัยที่เข้าร่วม ดังนี้

- รศ.ดร.พรรณฤมล เต็มดี มหาวิทยาลัยแม่ฟ้าหลวง
- รศ.ดร.สิทธิชัย บุษหมั่น มหาวิทยาลัยราชภัฏมหาสารคาม
- รศ.ดร.ทิพรัตน์ สิทธิวงศ์ มหาวิทยาลัยนเรศวร
- ผศ.ดร.ฆณการ ภัณณิพงส์ มหาวิทยาลัยเทคโนโลยีราชมงคลตะวันออก
- ผศ.ดร.ทิพวิมล ชมภูคำ มหาวิทยาลัยราชภัฏมหาสารคาม
- ผศ.ดร.ณิฏะญาร์ บรรเทา มหาวิทยาลัยราชภัฏมหาสารคาม
- ดร.พัชระ นาเสงี่ยม มหาวิทยาลัยราชภัฏมหาสารคาม
- ดร.อุมาพร ไชยสูง มหาวิทยาลัยเทคโนโลยีราชมงคลอีสาน วิทยาเขตสุรินทร์

จากความร่วมมือในเครือข่ายนักวิจัยดังกล่าว ส่งผลให้มีผลงานวิจัยที่ได้รับการวิจัยและตีพิมพ์เผยแพร่จำนวน 10 ผลงาน ดังนี้

[FF68 – UoE – 5028 – P05] “Classification Model for Screening Knee Osteoarthritis Patients in Northern Thailand Using Data Analytics” to The 5th International Conference on Data Engineering and Communication Technology (ICDECT) at https://icdect.com/

[FF68 – UoE – 5028 – P04] “Sign Language Detection Mobile Application for Thai Patients Using Medical Image Processing to Support Medical Consultations” to The 3rd World Conference on Information Systems for Business Management (ISBM’ 2024) at https://isbm.ict4sd.org

[FF68 – UoE – 5028 – P03] “AI for Healthcare: A Classification Model for Personalized Premenstrual Symptoms and Depressive Crisis Risk Tracking Using Data Analytics and Machine Learning” to The 10th International Conference on Digital Arts, Media and Technology (DAMT) and 8th ECTI Northern Section Conference on Electrical,

Electronics, Computer and Telecommunications Engineering (NCON): ECTI DAMT & NCON 2025 at https://www.icdamt.org/

[FF68 – UoE – 5028 – P02] Buntao, N., Somkhuean, R., Nuankaew, P., & Nuankaew S., W. (2025). Evolutionary Feature Selection to Classify Elderly Diseases from Dietary and Exercise Habits and Emotions. International Journal of Engineering Trends and Technology – IJETT, 73(1), 166–176. https://doi.org/10.14445/22315381/IJETT-V73I1P114

[FF68 – UoE – 5028 – P01] Chompookham, T., Nuankaew, W. S., & Nuankaew, P. (2024). Two-Factor Authentication Application Using Artificial Intelligence to Support Academic Information Systems. International Journal of Engineering Trends and Technology – IJETT, 72(12), 14–29. https://doi.org/10.14445/22315381/IJETT-V72I12P102

จากผลงานที่ปรากฏแสดงให้เห็นถึงความร่วมมือและเครือข่ายนักวิจัย ซึ่งโครงการ "สารสนเทศเชิงประยุกต์เพื่อนวัตกรรมสารสนเทศทางการแพทย์" หวังเป็นอย่างยิ่งที่จะต่อยอดและพัฒนาส่งเสริมงานวิจัยต่อไป

5. แผนงานในอนาคต/ ข้อเสนอแนะ โดยสรุป

โครงการ “ สารสนเทศเชิงประยุกต์เพื่อนวัตกรรมสารสนเทศทางการแพทย์” มีแผนการดำเนินงานในอนาคต โดยได้ต่อยอดการเสนอขอรับทุนส่งเสริมการวิจัยจาก กองทุนส่งเสริมวิทยาศาสตร์ วิจัย และนวัตกรรม (กสว.) : งบประมาณด้าน ววน. ประเภท Fundamental Fund ประจำปีงบประมาณ 2569 และ ปีงบประมาณ 2570 โดยทีมนักวิจัยได้รับการส่งเสริมจากศักยภาพของนักวิจัยที่ได้ปรากฏ ดังนั้น จึงสรุปเบื้องต้นได้ว่า หน่วยงาน และองค์ในระดับมหาวิทยาลัย และระดับชาติ ต่างให้ความสำคัญและเล็งเห็นศักยภาพของนักวิจัย ซึ่งทีมนักวิจัยต่างมุ่งมั่นที่จะพัฒนางานวิจัย เพื่อรับใช้สังคมและประเทศชาติต่อไป

ข้อเสนอแนะ จากโครงการวิจัย โครงการ “ สารสนเทศเชิงประยุกต์เพื่อนวัตกรรมสารสนเทศทางการแพทย์” ประกอบด้วยประเด็นเล็กน้อยได้แก่ ทีมผู้วิจัยยังเป็นนักวิจัยหน้าใหม่ ที่ยังมีประสบการณ์น้อย จึงทำให้ไม่สามารถวางแผนและควบคุมการบริหารจัดการโครงการได้ดีเท่าที่ควร ตัวอย่างเช่น การตีพิมพ์ผลงานวิจัยนอกกรอบระยะเวลา อย่างไรก็ตามปัญหานี้ เป็นปัญหาที่เกิดขึ้นจากวารสารที่ผู้วิจัยส่งไปตีพิมพ์ ดังนั้น ในอนาคตทีมนักวิจัยจำเป็นต้องการมีวางแผนที่ครอบคลุมและกำหนดแผนสำรองไว้ด้วย

6. ภาคผนวก

6.1 เอกสารประกอบ และเอกสารอื่นๆ

เอกสารประกอบและรายละเอียดการเผยแพร่ต่างๆ ถูกสรุปในตอนท้ายของเอกสาร

International Journal of Engineering Trends and Technology
ISSN: 2231–5381 / https://doi.org/10.14445/22315381/IJETT-V72I12P102

Volume 72 Issue 12, 14-29, December 2024

Original Article

Two-Factor Authentication Application Using Artificial Intelligence to Support Academic Information Systems

Thipwimon Chompookham[1], Wongpanya S. Nuankaew[2], Pratya Nuankaew[3]

[1]*Department of Computer Technology and Digital, Faculty of Information Technology, Rajabhat Maha Sarakham University, Maha Sarakham, Thailand.*
[2]*Department of Computer Science, School of Information and Communication Technology,University of Phayao, Phayao, Thailand.*
[3]*Department of Digital Business, School of Information and Communication Technology, University of Phayao, Phayao, Thailand.*

[3]*Corresponding Author : pratya.nu@up.ac.th*

Received: 14 August 2024 Revised: 14 November 2024 Accepted: 18 November 2024 Published: 25 December 2024

Abstract - *Integrating facial recognition technology is crucial for enhancing the security of online learning and consultation platforms, particularly in computer technology and engineering education. This project aims to synthesize components and technologies to develop a prototype of a two-factor authentication system that leverages artificial intelligence technology. The goal is to enhance the efficiency and capacity of identification verification. Moreover, a comprehensive evaluation will be carried out to determine the prototype's performance and level of acceptance. A purposive random sampling strategy was employed to select 40 participants from the Computer Technology and Digital Program at Rajabhat Maha Sarakham University, including students and faculty members. The research instruments comprised a thorough questionnaire encompassing several facets of the academic information system's development, a prototype of a two-factor authentication system, a quality assessment form, and a questionnaire to gauge the information system's acceptance. The performance criteria comprised accuracy, precision, recall, f1-score, average, and standard deviation. The findings indicated that the authorized prototype comprised four distinct modules: an authentication module, a member module, an information module, and a management module. The evaluation results for identifying faces, using CNN Face Detector, VGG-Face net, and classification by Logistic Regression, attained a remarkable accuracy of 83.54%. The precision and recall values were 0.84 and 0.88, respectively. The evaluation findings indicate that the overall quality is quite acceptable, with a mean score of 4.44 and a standard deviation of 0.55. Similarly, the user satisfaction with the 2FA system prototype is high, as indicated by a mean score of 4.54 and a standard deviation of 0.51.*

Keywords - *AI in Education, Facial recognition technology, Mobile learning support, Pedagogical prototype, Two-factor authentication.*

1. Introduction

In the present era, characterized by the dominance of the digital sphere, human existence is primarily influenced by online activities, which are directly connected to daily behaviors. These activities include various behaviors, such as doing online transactions, making purchases, paying for products and services, communicating with colleagues, documenting client information, and monitoring account balances via online banking. The Internet has had a significant influence on human existence and has witnessed a worrisome surge in patterns. However, the Internet is commonly recognized as an extensive repository of information, providing online venues for users to engage in conversation, exchange their experiences, and avail themselves of a diverse array of convenient services. Due to these concerns, data security is imperative for individuals and companies. There exist specific crucial pieces of knowledge that are not suitable for public disclosure. To access computer systems, computer system security must diligently identify and address potential vulnerabilities that may compromise particular data, such as ID card information, credit card information, usernames, and passwords. Therefore, authentication is crucial and essential when entering any system to confirm the identity of the data owner. Authentication is the procedure of verifying and validating the accuracy and legitimacy of an individual's information. This can be achieved by directly communicating with the data owner or employing computer systems or artificial intelligence to analyse the data. Authentication is a crucial and necessary element of security protocols. During the authentication process, the user provides evidence of authorization to access the system or data. Presently, there are multiple forms and protocols for authentication, such as user

IDs, usernames, passwords, and biometric authentication. Each technique has its advantages and disadvantages. The predominant method of authentication in contemporary society is the utilization of usernames and passwords, which has been broadly embraced and firmly established as a customary procedure for a substantial duration. The main concern with username and password authentication is the vulnerability of user information to interception.

Furthermore, individuals are unaware that their data has been taken, and service providers are unsure about the user's ownership of the data. When the system has a significant impact, relying solely on username and password authentication is deemed inadequate and insufficient in terms of security, as it only involves a single authentication step to access the system. In order to resolve the problem mentioned before, researchers utilize a two-step verification method called Two-Factor Authentication (2FA) when users log in. 2FA is a security measure that requires two keys to access a system, much like having two doors. This signifies that the system has enforced a prerequisite for verifying at least two criteria to access the data or system. Usually, individuals are required to verify their identity by inputting their user ID and password to gain access to the system.

Subsequently, the system proceeds to authenticate the user's identity at the second tier, utilizing diverse techniques that align with the capabilities of the created system. Current examples of two-step verification methods include the transmission of a One-Time Password (OTP) by SMS, the encryption of code through emails, the use of a second verified device, the utilization of fingerprints, and other comparable approaches. Due to the improved efficiency of contemporary communication technologies, users now possess the capacity to transmit greater quantities of information across networks. Consequently, two-step verification has progressed to include facial recognition as a security feature. Facial recognition is an artificial intelligence approach that can assess and make comparisons between the facial features of individuals.

Facial recognition technology can swiftly and precisely recognize persons in photos or videos containing their faces. Afterwards, it compares the faces in the database to determine the identification of the detected face and then analyzes its unique facial features. Moreover, using facial recognition technology for unlocking or logging in reduces the user's need for physical contact with objects or surfaces that may transmit diseases like COVID-19. Facial recognition technology has become widely popular in several businesses and people's everyday routines. It is utilized to verify identities to gain access to highly secure places, monitor staff entry and exit times, record entry permissions for telephone systems, and validate users for online banking security systems. Facial recognition systems offer a straightforward and fast method of verifying someone's identity by quickly differentiating their face. This approach is more convenient and efficient than using cards and fingerprints for identification. To improve student achievement, educational institutions and organizations should prioritise creating information systems for education, even though there may be limited use of information technology and a lack of student engagement since many instructors heavily rely on cognitive teaching methods. Maximizing the capabilities of driving pupils is a step-by-step procedure, particularly given the transition from conventional classroom environments to online platforms, while encouraging sustainable learning. Academic instructors and educational institutions require the development of academic expert systems and educational technologies that improve the advancement of student learning. Currently, learners necessitate an information system that functions as a learning tool and is compatible with their learning behavior. Blended online learning and online mentoring are crucial elements in fulfilling the requirements of students. Moreover, concerns regarding the system's trustworthiness generate apprehension among children, discouraging them from accessing the system to seek help. Based on the observations reported earlier, the researchers proposed using facial recognition technology to improve instructional technologies. They accomplished this by creating a prototype for identification purposes, aiming to enhance the security of information systems that facilitate learning and provide online guidance for counselors and students. The researchers conducted a study and created a preliminary version of a two-factor authentication system. They employed two approaches to authenticate individuals, utilizing artificial intelligence technology incorporating facial recognition into the login process.

The prototype may verify users' identity by requiring the input of a username and password during the initial stage. The system will capture a facial image from a camera or webcam, which will be utilized for identity authentication in the subsequent stage to enhance security during system usage, as depicted in Figure 1 of the study framework. Figure 1 illustrates a study framework that uses two-factor authentication (2FA) to improve the functionality of academic information systems (AIS). The framework consists of two distinct study eras. The first phase entails constructing the two-factor authentication (2FA) login prototype. There are two elements, as depicted in Figure 2 and Figure 3.

The last phase entails implementing the two-factor authentication (2FA) login prototype to streamline access to academic information systems (AIS). This research proposal exclusively addresses the preliminary phase of the investigation, while the research has a wide-ranging reach. Assessing the efficacy of academic information systems requires conducting tests with actual users, who are the primary target audience based on research data. Figure 2 comprehensively describes the initial two-factor authentication (2FA) element that aids academic information systems (AIS).

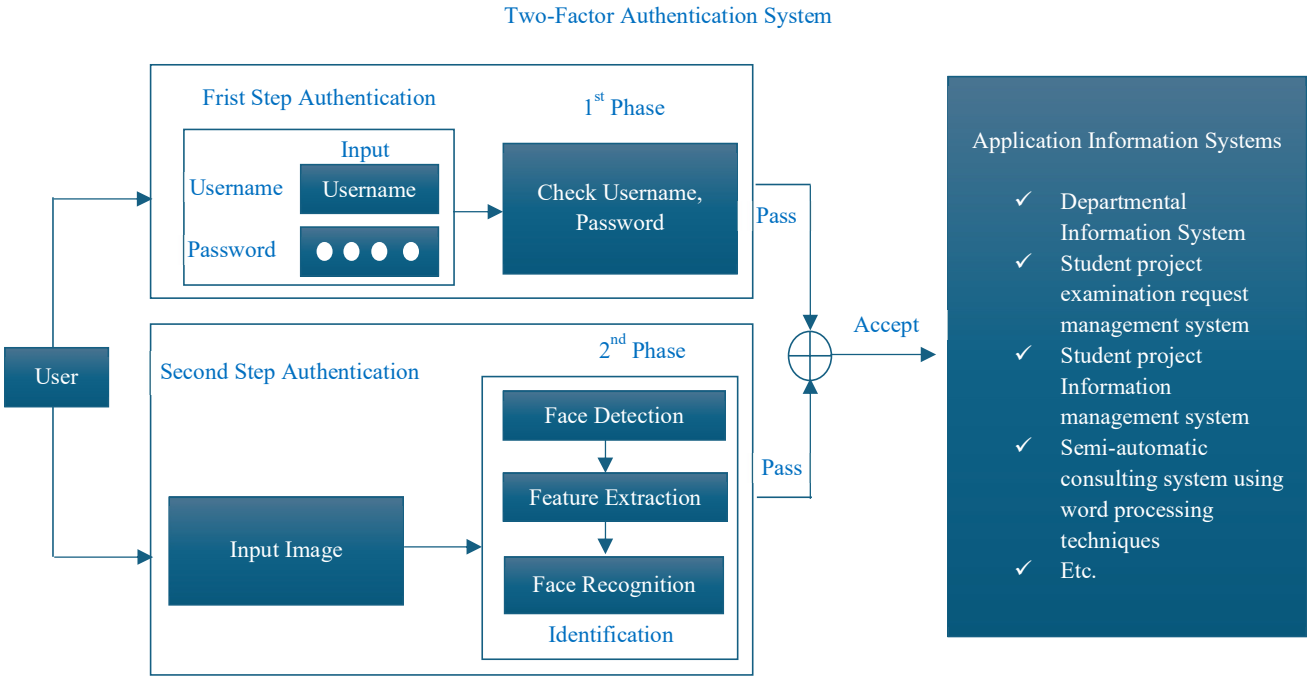


Fig. 1 The research framework

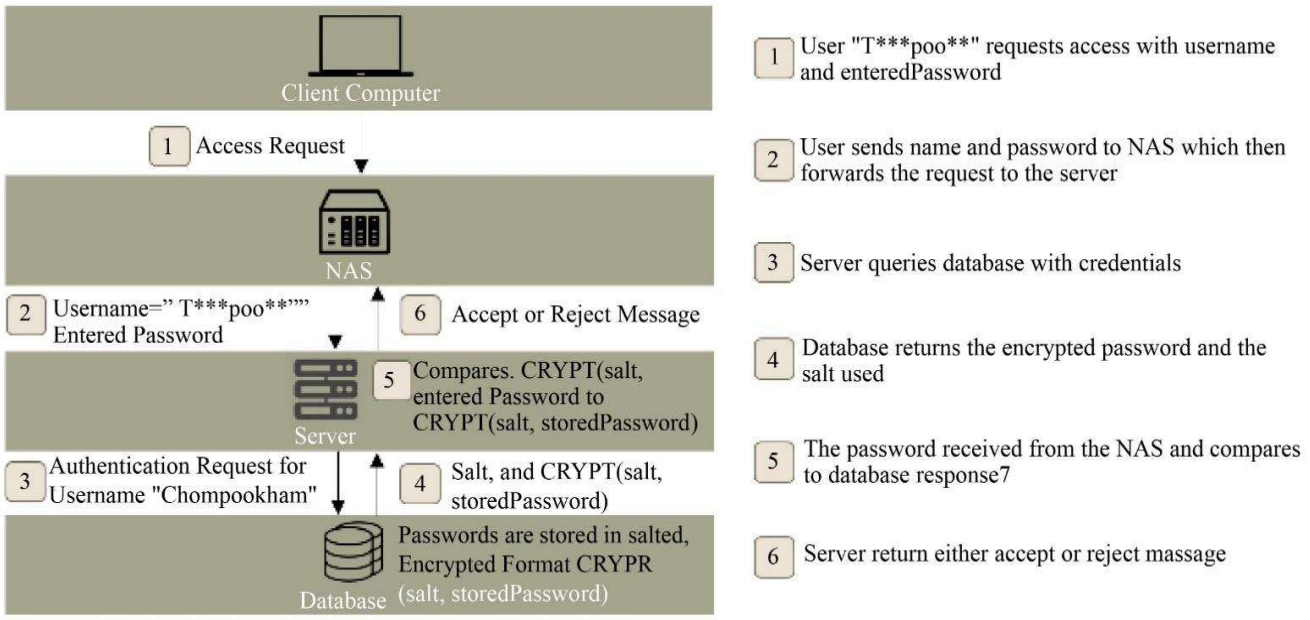


Fig. 2 The initial component of the two-factor authentication (2FA) system

The primary objective of this component is to create a robust authentication system that requires a username and password for secure login. The process comprises six sequential stages, elaborated in the materials and methods part and Pseudo Code, as specified in Table 4. Figure 3 provides a complete description of the second component of two-factor authentication (2FA), specifically designed to improve the security of academic information systems (AIS).

The main objective of the second phase is to incorporate facial recognition into the prototype of the two-factor authentication system with artificial intelligence. This integration will enhance the functionality of academic information systems, as described in the materials and techniques section, and will adhere to the Pseudo Code provided in Table 5. Significantly, it defined the study objectives, encompassing three main aims.

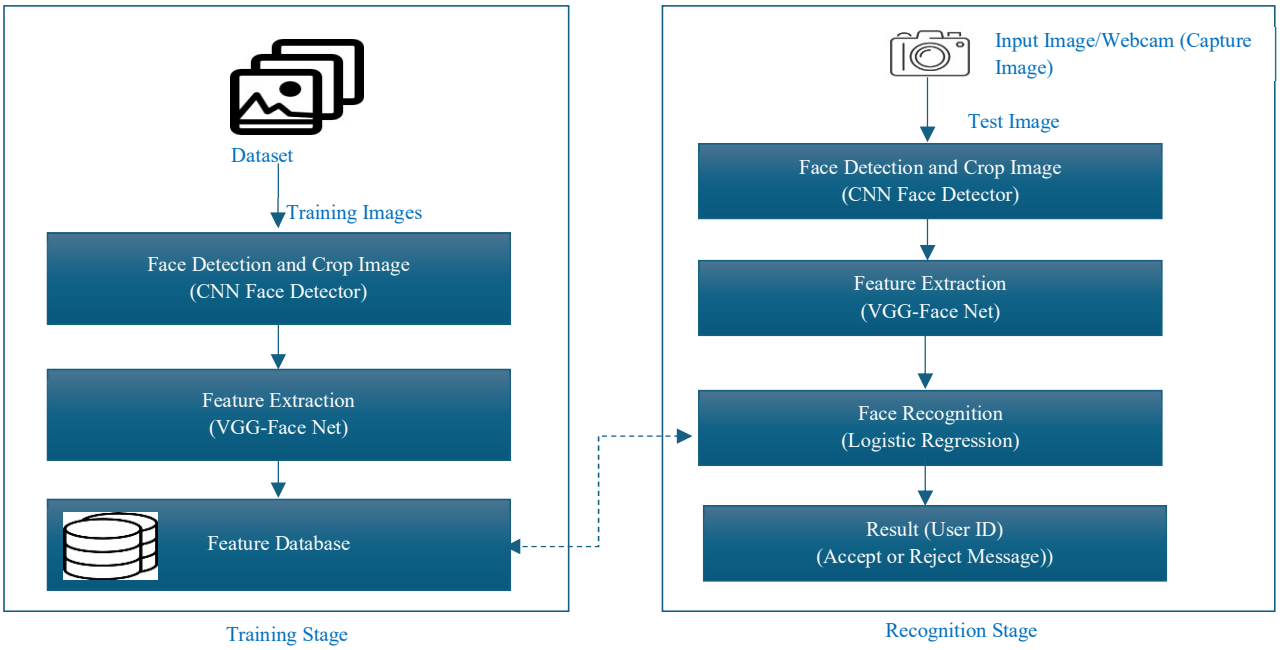


Fig. 3 The second component of the two-factor authentication (2FA) system

The study aimed to integrate the essential elements and technologies to prototype a two-factor authentication system that utilizes artificial intelligence technology. The second study objective was to prototype a two-factor authentication system using artificial intelligence technologies. This system would optimize the effectiveness and capability of identity verification while integrating robust security measures for system access. The third research objective involved evaluating the efficacy and studying the user response to the prototype of the two-factor authentication system utilizing artificial intelligence technology. The researchers are confident that this discovery will significantly benefit education, computer science, computer engineering, and related academic subjects in the future. Their unwavering will and diligent efforts will guarantee a substantial and enduring influence.

2. Literature Review and Related Works

2.1. Authentication System

Authentication systems must meet three fundamental criteria [1]: The primary criterion is to possess ample system resources for authorized users and the capability to deliver uninterrupted services to users with dependability. The second requirement is that the system must be able to preserve the integrity of the data, thereby preventing any unauthorized individuals from modifying it and safeguarding against unauthorized infiltration and manipulation of the data. The last determinant is confidentiality, which concerns the capacity to maintain the concealment and privacy of users. The system should incorporate safeguards to deter illegal access to data. Data encryption technology ensures the secrecy and accuracy of data within the system [2]. Adhering to these criteria is crucial when creating, implementing, and maintaining an authentication system to ensure reliability. The level of protection against unauthorized use and data access is determined by requirements, resources, priorities, and environment, which are established by suitable authentication [3].

2.2. Multi-Factor Authentication

Multi-Factor Authentication (MFA) is an authentication principle that is applied to the login process by providing an additional layer of authentication to the system or a means of verifying at least two factors to increase the security of the system even more [4], [5], [6] There are three authentication factors.

- What you know, what you have, and who you are. What you know may be using passwords [7], [8], digital signatures [9], or personal data.
- What you have might be a smartcard [10], [11], Radio-Frequency Identification (RFID) [12], and Near-Field Communication (NFC) [13]
- Who you are is biometric identity [14], [15], [16] such as fingerprint, face, hand geometry, voice, and retina. Two-factor authentication means using an authentication method organized into two different authentication factors.

2.3. Face Recognition for User Authentication

Face recognition technology for authenticating users increases system security and prevents potential vulnerabilities [17]. Some features of facial authentication systems are that they utilize different people's facial

characteristics to create unique identities, increasing security and reducing the risk of unauthorized access through facial recognition technology. Facial recognition has the advantage of being easy to use and convenient, not disturbing users, and having no hardware requirements [18], [19]. Facial recognition requires image processing techniques and machine learning algorithms for face detection and identification. Rameswari et al. [20] present a method to encode faces in the input image using the Histogram Object Gradient (HOG) method and detecting faces using the Facenet algorithm. Then, the Support Vector Machine (SVM) method compares and classifies faces. Chen et al. [21] presented an efficient convolutional neural network model called Mobile Facenets, which uses less than 1 million parameters and can efficiently detect real-time faces on mobile devices. Sagar and Narasimha [22] developed a thoughtful and robust face detection-based lock system using facial detection and recognition algorithms, including the application of Principal Component Analysis(PCA), Histogram Equalization, and Linear Discriminant Analysis (LDA) where the system decides which algorithm to use to detect and recognize faces based on the current light intensity. Kim et al. [23] present a practical algorithm to improve recognition accuracy using a hierarchical deep neural network structure by extracting features from the appearance of a feature-based network fused with the geometric feature in a hierarchical structure.

2.4. Academic Information System (AIS)

The swift advancement of information technology presents a notable difficulty in implementing information systems [24]. It is necessary to carefully assess and implement suitable measures to guarantee results that align with the requirements and goals of technology users. Data processing in information technology includes acquiring, processing, storing, and managing data in diverse formats to derive high-quality information. An Academic Information System (AIS) is a data storage and management system that offers academic information services within educational institutions. Typically, it encompasses student details such as personal information, academic background, grades, and other information about the student's education at that institution [25]. Indrayani's study [26] defines AIS as a collection of methods and activities employed in higher education institutions to arrange, handle, and utilize information. AIS serves multiple purposes, including generating reports on activity performance, addressing hypothetical scenarios, facilitating decision-making, and assessing results in institutional development. AIS must address the requirements of various users, such as students, teachers, administrative personnel, and executives.

3. Materials and Methods

The prototype two-factor authentication system was developed by combining artificial intelligence, face recognition technology, and the conventional technique of validating identity using a login and password.

3.1. Population and Sampling

The study focuses on the attributes of the sample population and the process employed to choose participants. The study analysed the demographic characteristics of students, teachers, and staff associated with the Department of Computer Technology and Animation at the Faculty of Information Technology, Rajabhat Maha Sarakham University.

The study utilized purposive sampling methods, explicitly targeting a cohort of 40 individuals comprising students, lecturers, and staff affiliated with the Department of Computer Technology and Digital at the Faculty of Information Technology, Rajabhat Maha Sarakham University, for the academic year 2022. Participants granted their informed consent to utilize their data for research purposes. The researchers will abstain from disseminating or disclosing the acquired data without getting agreement from the individuals.

3.2. Data Collection

The data underwent three distinct phases. The initial stage involved gathering data to analyse and combine the many components of the academic information system. The second stage involved gathering data to create and construct a prototype two-factor authentication system utilizing artificial intelligence technologies. The third step involved collecting data to assess the system's quality and studying user approval of the prototype.

3.2.1. Academic Information Systems Synthesis Data

The purpose of the Data collection for the synthesis of academic information systems was to identify critical factors for creating a prototype of a two-factor authentication system utilizing artificial intelligence to enhance academic information systems.
This technique consisted of four steps.

- Step 1: The design of the research framework for developing a two-factor authentication system was a collaborative effort. In consultation with experts from various fields, including system development, context analysis, documentation, and research, researchers analyzed the system's boundaries, constraints, and issues.

- Step 2: We synthesized suitable components for a prototype of a two-factor authentication system that utilizes artificial intelligence to enhance academic information systems.

- Step 3: The synthesized elements were distributed to five experts to solicit their arguments and acceptance of the cumulative generalization of the elements for a prototype of a two-factor authentication system that utilizes artificial intelligence to enhance academic information systems. This process is elaborated in Table 1.

- Step 4: They analysed and summarised the results of synthesizing suitable components for the prototype of a two-factor authentication system using artificial intelligence. These findings are presented in Table 6.

3.2.2. Data for Designing and Development

The data-collecting process for prototype design and development was conducted by analysing the results and conclusions derived from synthesizing the relevant system components, with input from five experts. The results of the design and development of the prototype system were separated into two components. The initial element consisted of a sign-in system, as depicted in the operational architecture illustrated in Figure 2. The second element consisted of a facial recognition two-factor authentication system prototype, as described in the functional framework illustrated in Figure 3. The research tools section provides a comprehensive overview of the development process for each framework.

3.2.3. Data for Prototype Evaluation

The purpose of gathering data for testing and evaluating the prototype's effectiveness was to examine the level of satisfaction and acceptability toward the two-factor authentication system prototype, which utilizes artificial intelligence to assist academic information systems. Our team of dedicated researchers undertook this task with five meticulous steps, outlined below.

- Step 1: Our team of researchers meticulously developed a comprehensive and robust prototype for a two-factor authentication system trial, ensuring its quality and reliability.

- Step 2: The researchers created detailed questionnaires to evaluate the system's quality, as described in Table 2, and to provide a comprehensive analysis of user approval, as described in Table 3, of the prototype for the two-factor authentication system.

- Step 3: Researchers organized and requested collaboration from sample groups and experimented on the system prototype. During this phase, the researchers exercised control over, explained, clarified, and assisted the samples throughout the testing process.

- Step 4: Questionnaires were gathered and assessed for acceptability based on their sample questionnaires, and a prototype system was developed for testing activities with the samples.

- Step 5: The researchers condensed the questionnaire responses, summarized the reaction results, analysed the level of acceptability among the sample groups, assessed the performance of the system prototype, and deliberated on the findings, as documented in the assessment reports presented in Tables 10 and 11.

3.3. Research Tools

Four instruments were used for the research. The first instrument used was a questionnaire designed to assess the effectiveness of the prototype components of a two-factor authentication system that utilizes artificial intelligence technology to enhance academic information systems.

The initial questionnaire was created for assessment by specialists, as outlined in Table 1, and the results are presented in Table 7. The second instrument refers to the prototype of a two-factor authentication system that utilizes artificial intelligence to enhance academic information systems. This prototype is described and outlined in the part dedicated to its production.

The third instrument employed was a questionnaire designed to assess the efficacy of the prototype two-factor authentication system, which utilizes artificial intelligence to enhance academic information systems. This questionnaire was arranged for specialist assessment, as specified in Table 2, and the result is displayed in Table 10.

The fourth instrument was a questionnaire designed to assess the acceptance of the prototype two-factor authentication system, which incorporates artificial intelligence to enhance academic information systems. This survey was conducted to investigate the level of acceptance among a selected group, as outlined in Table 3, and yielded Table 11.

3.4. Prototype Construction and Quality of Tools

The conceptual notion of developing a prototype for a two-factor authentication system employing artificial intelligence to assist academic information systems is depicted in the study framework, shown in Figure 1. The structure and components comprised two primary elements. The primary element is creating and manufacturing a secure login system that requires a username and password, as depicted in the workflow illustrated in Figure 2.

The second element involves integrating a prototype of a two-factor authentication system that utilizes artificial intelligence to enhance the functionality of academic information systems, as described in Figure 3 of the framework. This section offers a comprehensive explanation of both frameworks.

3.4.1. Development of the 2FA System Prototype

The development of the 2FA system prototype commenced by incorporating a login mechanism that required the user to enter both a username and password.

The researchers developed an initial iteration of a two-factor authentication (2FA) system utilizing the Systems Development Life Cycle (SDLC) framework, which encompasses five distinct stages.

- Step 1: Planning

The system's development has been methodically and extensively planned. An extensive investigation was conducted to collect various user requirements, meticulously evaluate the details, and pinpoint the underlying reasons for the problems. Subsequently, these discoveries were crucial in creating and implementing the login system.

- Step 2: Performing system analysis

The data collected in Step 1 was analyzed to incorporate user requirements by developing a framework that addresses the varied system needs.

- Step 3: The procedure of system design.

The data acquired from the system analysis process was employed to build the database system and its associated procedures, influencing its design.

- Step 4: System development.

A fresh technique was developed by utilizing appropriate technologies. Information was acquired from the first round of data gathering, which involved the participation of five specialists. The precise research inquiries may be located in Table 1, while the accompanying analytical discoveries are displayed in Table 7.

The data in this section was employed to evaluate and select the most appropriate technology. Afterwards, the system underwent thorough testing, which included developing its installation, assessing its compatibility with the current system, and providing training to equip system users with the required information.

The login system is activated by entering a username and password. Figure 2 illustrates the arrangement, while Table 4 displays the Pseudo code.

- Step 5: System Maintenance

The researchers were following the completion of tests using the prototype of the 2FA system. Multiple modifications were implemented to complete and assemble a guidebook to improve system efficiency.

Table 1. Questionnaire to evaluate the suitability of 2FA prototype components

Stage	Details
CP1. Suitability of system components	
CP1.1	Suitability of members
CP1.2	Suitability of information
CP1.3	Suitability of user management
CP1.4	Suitability of two-factor authentication
CP1.5	Suitability of user permission
CP2. Suitability of technological components in system development	
CP2.1	Suitability of hardware
CP2.2	Suitability of code and programming
CP2.3	Suitability of database management system
CP2.4	Suitability of software for two-factor authentication
CP2.5	Suitability of reports via browser and platform

Table 2. Questionnaire for quality assessment

Stage	Details
AC1. Function test	
AC1.1	Completeness of prototype components
AC1.2	Completeness of prototype data presentation
AC1.3	Completeness of the prototype's data retrieval
AC1.4	Capabilities of signing in to the system
AC1.5	Two-factor authentication capabilities
AC2. Result test	
AC 2.1	Correctness of the system in adding data
AC 2.2	Correctness of the system in editing data
AC 2.3	Correctness of the system in deleting data
AC 2.4	Correctness of the system in two-factor authentication
AC 2.5	Correctness according to the framework of the system
AC3. Usability test	
AC3.1	Usability to access the data
AC3.2	Arrangement of system complexity
AC3.3	Overview of usability testing with the system

AC4. Security test	
AC4.1	Appropriateness to assign permissions to access data
AC4.2	Security in accessing data
AC4.3	Appropriateness of the overall security system

Table 3. Questionnaire for acceptance

Stage	Details
AC1. Perceived usefulness	
AC1.1	The system can effectively verify the identity of factor 1
AC1.2	The system can effectively verify the identity of Factor 2
AC1.3	The system can link information efficiently.
AC1.4	The system processes overall identity verification information correctly.
AC1.5	The identity verification system helps secure the work.
AC2. Perceived ease of use	
AC 2.1	Manual for recommending use
AC 2.2	The usage process is flexible and not complicated.
AC 2.3	Information systems can quickly access information.
AC 2.4	The information system uses visual explanations, content, and appropriate organization of elements to make it easy to understand
AC 2.5	The information system can be used anywhere and anytime via web browsers from many devices and all platforms.
AC3. Attitude towards use	
AC3.1	The user is satisfied with the system overall.
AC3.2	The user is satisfied with the system's process.
AC3.3	The user is satisfied with the stability of the system.
AC3.4	The user is satisfied with the manual that guides them using the system.
AC3.5	The user is satisfied with the manual that guides them using the system.

Table 4. Pseudo code for the first 2FA component

Pseudocode for the initial step of the authentication system
Input Username EnteredPassword
Process Authenticate by providing a Username and the corresponding EnteredPassword. The user sends their username and entered password to the Network-Attached Storage (NAS), which relays the request to the server. The server queries data from the database using authentication credentials. The database returns the encrypted password and the corresponding permissions. If the encrypted password in the database response matches the encrypted EnteredPassword returned from NAS, Then Server returns the accept Proceed to the second step authentication **Else** Server returns the reject Username or Password may be invalid or wrong
Try again

3.4.2. Second Stage of Development 2FA System Prototype

The second stage of creating the 2FA system prototype entailed integrating facial recognition technology into the existing 2FA system prototype.

A facial verification system was implemented by combining a human image recognition model into the 2FA system prototype. The system follows a five-step process, as depicted in Figure 4.

- Step 1: Data Preparation

The stage obtained image data of the provided sample groups. The dataset comprised 400 photographs, with ten images for each of the 40 students and teachers affiliated with the Faculty of Information Technology at Rajabhat Maha Sarakham University.

These photographs depicted the unique ambiance and overall surroundings.The data owner has explicitly granted permission to collect and utilize the data for research purposes.

The data was partitioned into two segments, with 80% allocated for evaluating the prototype system's performance and 20% designated for split testing.

Thus, 80% of the resources were spent on developing the prototype, while the remaining 20% were dedicated to testing its performance.

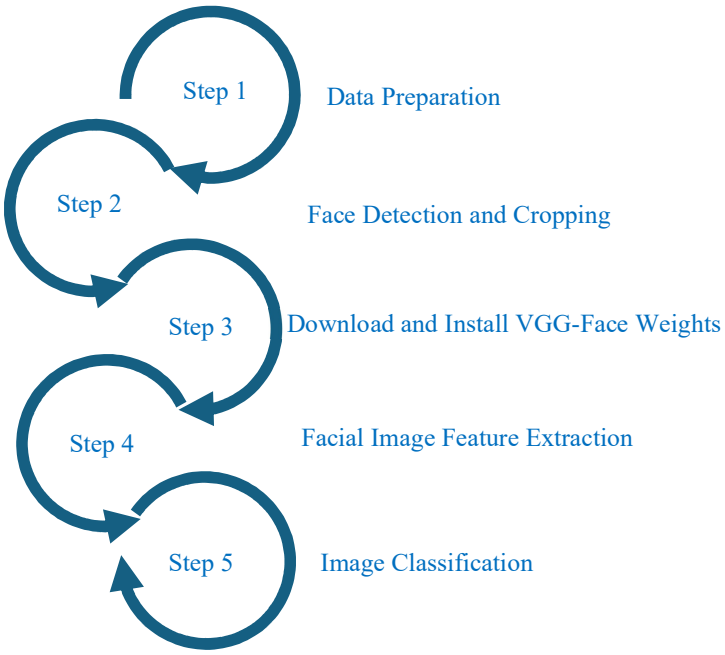


Fig. 4 Constructing a facial recognition module

- Step 2: Face Detection and Cropping.

At this stage, facial pictures were collected and analyzed. Researchers utilized the CNN-Face Detector, a powerful tool that accurately detects and locates faces in photographs and then returns this information to the system. Afterwards, the researchers cropped the facial photos for facial recognition.

- Step 3: Download and Install VGG-Face Weights.

It acquired pre-trained neural networks. The experiment utilized VGG-Face Net, a neural network that has undergone rigorous training on a vast dataset of facial photos to recognize human faces precisely.

- Step 4: Facial Image Feature Extraction.

The VGG-Face Net model extracted picture features from the generated pages. Subsequently, each image was assigned the name of the data owner.

- Step 5: Image Classsification.

In this step, a machine learning algorithm was employed to generate and categorize photographs by randomly splitting the experimental data into two distinct datasets. The datasets were later used to create and evaluate a "Split Test" model, with 80% of the data used for development and 20% used for testing. 80% of the initial dataset was utilized to build a model, with the data being divided into two smaller segments. During the initial stage, 70% of the resources were allocated towards developing the prototype model.

The subsequent portion assigned 10% of the resources to validate the model. An equivalent of 20% of the remaining data was allocated to evaluate the prototype's efficiency. The data was classified, and then a machine learning model was created using four classifiers: K-Nearest Neighbors (K-NN), Logistic Regression (LR), Multi-Layer Perceptron (MLP), and Support Vector Machine (SVM). The results of the prototype development, utilizing the four strategies, are outlined in the part dedicated to the prototype development report. The second element of the 2FA system prototype demonstrates the functioning of the process in the Pseudo code, as depicted in Table 5.

Table 5. Pseudo code for the second 2FA

Pseudo code the Second Step Authentication System
Input Username Image (from computer) or capture image (from webcam)
Process Face detection and crop image by CNN face detector Face image feature extraction will use the VGG-Face-Net model. Classify image using a trained Logistic Regression model. **If** the model return classification results (Username) and Username input are the same, Then Server return accept Enter academic information system: AIS **Else** Server return reject
Try again

Table 6. The Confusion matrix

	Actually Positive	Actually Negative
Predicted Positive	True Positive (TP)	False Positive (FP)
Predicted Negative	False Negative (FN)	True Negative (TN)

After developing the two-component subsystems, a prototype system was created for testing. The test target group consisted of two parts: 5 experts and 40 students and instructors. Tables 10 and 11 display the test results and satisfaction study data.

3.5. Model Performance Assessment

The confusion matrix technique and four indicators—accuracy, precision, recall, and F1-Score—were used to evaluate the machine learning models' performance with four classifiers. From the confusion matrix shown in Table 5, the four indicators used to determine the model performance are displayed in Equation (1) to (4).

$$Accuracy = \frac{TP+}{TP+FP+FN+TN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$F1-Score = 2 * (\frac{Precision*Recall}{Precision+Recal}) \quad (4)$$

Where, True Positive (TP) and True Negative (TN) mean what was predicted matches what happened. In contrast, False Positive (FP) and False Negative (FN) mean what was predicted miss-matches what happened.

3.6. Research Analysis and Interpretation

The Statistics used in data analysis. The collected data was analysed, and results were summarized as mean and Standard Deviation (S.D.). A rating scale used in the study of acceptance and satisfaction consisted of 5 levels. A score of 5 = very satisfied; 4 = somewhat satisfied; 3 = neither satisfied nor dissatisfied; 2 = slightly dissatisfied; 1 = very dissatisfied. The rating scale was used with experts and sample groups, as shown in the questionnaire details in Table 1 and Table 2.The acceptance and satisfaction levels data were interpreted using five interval levels. A mean result between 1.00 and 1.80 at level 1 indicates a highly undesirable state. A mean value falling from 1.81 to 2.60 is deemed unsatisfactory at level 2. If the average number falls between 2.61 and 3.40 at level 3, it suggests a neither acceptable nor unacceptable state. A mean value falling within the range of 3.41 to 4.20 was deemed sufficient for level 4. A mean number between 4.21 and 5.00 at level 5 signifies a high level of acceptability. The data analysis and interpretation of acceptance and satisfaction levels were condensed and shown in Tables 7, 10, and 11.

4. Results

The investigation and construction of a prototype 2FA system utilizing artificial intelligence to enhance academic information systems yielded three key findings, documented in sections 4.1, 4.2, and 4.3. The first aspect pertains to the study's findings and the synthesis of components inside the information system. The second aspect pertains to the results derived from designing and creating the prototype for the 2FA system. The third aspect involves the experts' and sample groups' acceptance of the 2FA system prototype.

4.1. Synthesis of the 2FA System Prototype Components

The experts' input on the components necessary for constructing the 2FA system prototype was summarized and utilized to design the system process, as depicted in Figure 5. The findings of their evaluation of the acceptability of the 2FA system prototype module were reported and summarized in Table 7. Figure 5 illustrates that the 2FA system prototype comprised four modules. The Authentication Module is a component that offers authentication functionality for system access. Users are required to complete a two-step login process. Initially, the user must input a username and password.

As part of the second phase, users must authenticate their identity by capturing a photo using a camera. To log in, users must complete both stages. The Member Module is a component designed to store and manage data related to members. The recorded data includes information about both students and instructors. The individuals capable of documenting information are students and instructors. The Information Module is designed to contain many components of information from the Academic Information System (AIS). Only instructors are allowed to record information. The purpose of the Management Module is to streamline the process of managing member data access permissions. According to Table 7, the experts expressed that the components utilized in developing the 2FA prototype system were highly acceptable.

4.2. Results of the 2FA Prototype Systems' Development

The development results of the 2FA prototype system, which utilizes artificial intelligence to enhance academic information systems, are categorized into two sections. The initial section comprehensively analyses the progress in creating the facial recognition model. The second section entails a comprehensive report detailing the execution of the prototype system.

4.2.1. Facial Recognition Model

The development results of the prototype model of the human face verification system using the facial recognition module were tested for model performance with the steps shown in Figure 4. To test the performance of the facial recognition module, it was run on Google Colab. The experiment used A CNN-Face Detector to detect faces and

crop facial images. Following that, VGG-Face Net was used to extract image features and assign labels to the images. Once the facial image data was prepared, the data was divided into two parts for testing: a training set with 80% of the data and a testing set with 20% of the data.

The researchers used the first 80% of the data to develop a model using four classifiers: K-Nearest Neighbours (K-NN), Logistic Regression (LR), Multi-Layer Perceptron (MLP), and Supports Vector Machine (SVM), as detailed parameters in the experiment as shown in Table 8. Once the models from each technique were obtained, the remaining 20% of the data was tested with a confusion metric technique and four indicators. The test results for each classifier are reported in Table 9. The data in Table 9 demonstrates that the categorization efficiency test produced results with an accuracy surpassing 80 percent. The Logistic Regression (LR) classifier achieved an accuracy of 83.54%, while the Support Vector Machine (SVM) classifier achieved an accuracy of 81.01%. Therefore, the Logistic Regression (LR) model was chosen to be integrated into the login system in the facial recognition module because of its exceptional precision.

4.2.2. Implementing the Prototype System

Facial recognition technology in the 2FA prototype system has effectively streamlined user authentication in deploying Academic Information Systems (AIS) that require a login and password.

AIS development involves utilizing code programs written in PHP, Python, HTML5, CSS3, JavaScript, and a MySQL database system. To evaluate the effectiveness of the facial recognition authentication system, researchers utilized OKER-A229 cameras with a resolution of 1920×1080 pixels (Full HD-1080p). The system categorizes users into two distinct groups: students and teachers. Each user group possesses distinct permissions. Figures 6 and 7 depict the whole two-part login system. Upon accessing the system, users are granted access to all academic information systems within the Department of Computer Technology and Digital, Faculty of Information Technology, Rajabhat Maha Sarakham University.

4.3. Performance and Acceptance of the 2FA System

After the two systems were integrated, the system underwent testing, and arrangements were established to ssess the quality of the 2FA system prototype, which was separated into two groups.Group one comprised five experts, whereas group two was a specifically chosen sample of 40 students and instructors, as documented in Tables 10 and 11.

Table 10 indicates that the experts expressed high satisfaction and rated the overall quality of the 2FA system prototype as strongly acceptable. Upon careful evaluation, it was determined that the security and functionality testing components of the 2FA system prototype were highly satisfactory. The test results on the performance and ease of use of the 2FA system prototype met the acceptable criteria. Table 11 shows that the samples' satisfaction and acceptance of the overall quality of the 2FA system prototype is at a strongly acceptable level. When considering each aspect, it was found that the perceived usefulness, ease of use aspects, and attitude towards using the 2FA system prototype are strongly acceptable.

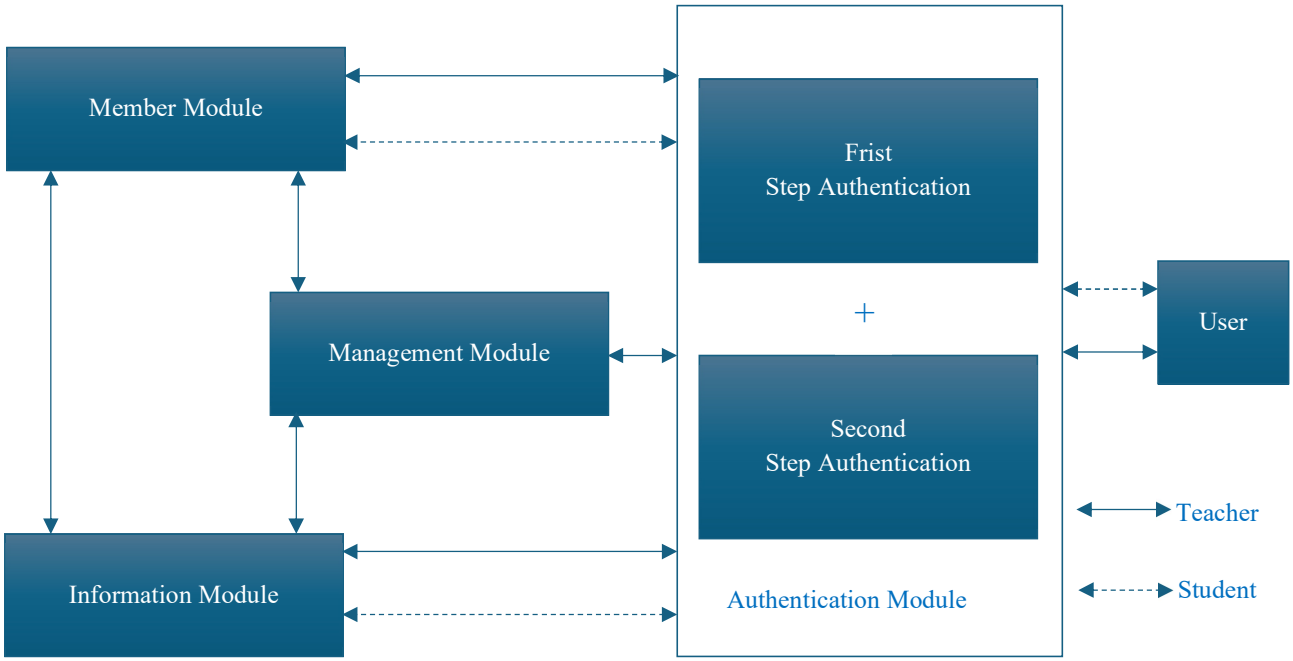


Fig. 5 The components of the 2FA system prototype

Table 7. Summary of the Experts' acceptance of the 2FA prototype components

Stage	Mean	S.D.	Interpretation
CP1. Suitability of system components			
CP1.1	4.40	0.55	Strongly Acceptable
CP1.2	4.60	0.55	Strongly Acceptable
CP1.3	4.40	0.55	Strongly Acceptable
CP1.4	4.80	0.45	Strongly Acceptable
CP1.5	4.60	0.55	Strongly Acceptable
Average:	4.56	0.51	Strongly Acceptable
CP2. Suitability of technological components in system development			
CP2.1	4.20	0.45	Acceptable
CP2.2	4.40	0.55	Strongly Acceptable
CP2.3	4.60	0.55	Strongly Acceptable
CP2.4	4.60	0.55	Strongly Acceptable
CP2.5	4.60	0.55	Strongly Acceptable
Average:	4.48	0.51	Strongly Acceptable
Total Average:	4.52	0.50	Strongly Acceptable

Table 8. Parameter in the experiment

Classifiers	Hyperparameters	Search Space	Bast Parameters
K-Nearest Neighbors (K-NN)	n-neighbors	[1, 15]	1
Logistic Regression (LR)	solvers penalty c_values max_iters	['lbfgs', 'newton-cg', 'sag', 'saga', 'liblinear'] ['l2', 'l1', 'none'] [100, 10, 1.0, 0.1, 0.01] 100 - 2000	Sag none 1.0 500
Multi-Layer Perceptron (MLP)	max_iters beta alpha activation solver	100 - 2000 0.1 - 0.9 0.0001 - 0.1 ['sigmoid', 'relu', 'Tanh'] ['lbfgs', 'sgd', 'adam', default='adam']	200 0.9 0.0001 relu adam
Supports Vector Machine (SVM)	c gamma kernel	[0.1 - 50] [0.0001 - 1] ['liner', 'rbf']	20 0.001 rbf

Fig. 6 Login with a username and password

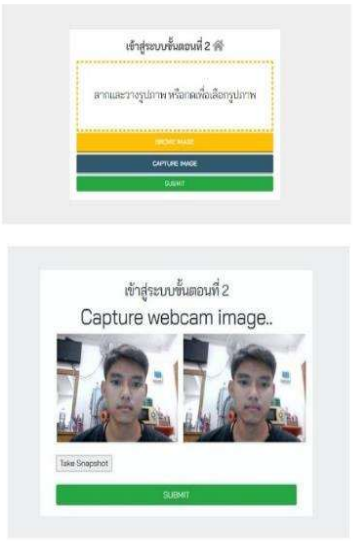


Fig. 7 Login with facial recognition

Table 9. Performance for each classifier

Classifiers	Accuracy (%)	Precision	Recall	F1-Score
K-NN	75.94	0.76	0.74	0.73
LR	83.54	0.84	0.88	0.83
MLP	78.48	0.78	0.81	0.76
SVM	81.01	0.81	0.84	0.80

Table 10. Acceptance from experts

Stage	Mean	S.D.	Interpretation
AC1. Function test			
AC1.1	4.60	0.55	Strongly Acceptable
AC1.2	4.40	0.55	Strongly Acceptable
AC1.3	4.60	0.55	Strongly Acceptable
AC1.4	4.60	0.55	Strongly Acceptable
AC1.5	4.40	0.55	Strongly Acceptable
Average:	4.52	0.51	Strongly Acceptable
AC2. Result test			
AC 2.1	4.20	0.84	Acceptable
AC 2.2	4.60	0.55	Strongly Acceptable
AC 2.3	4.20	0.45	Acceptable
AC 2.4	4.60	0.55	Strongly Acceptable
AC 2.5	4.40	0.55	Strongly Acceptable
Average:	4.40	0.58	Strongly Acceptable
AC3. Usability test			
AC3.1	4.00	0.71	Acceptable
AC3.2	4.40	0.55	Strongly Acceptable
AC3.3	4.40	0.55	Strongly Acceptable
Average:	4.27	0.59	Acceptable
AC4. Security test			
AC4.1	4.40	0.55	Strongly Acceptable
AC4.2	4.80	0.45	Strongly Acceptable
AC4.3	4.40	0.55	Strongly Acceptable
Average:	4.53	0.52	Strongly Acceptable
Total Average:	4.44	0.55	Strongly Acceptable

Table 11. Acceptance from sample groups

Stage	Mean	S.D.	Interpretation
AC1. Perceived usefulness			
AC1.1	4.50	0.51	Strongly Acceptable
AC1.2	4.40	0.50	Strongly Acceptable
AC1.3	4.55	0.50	Strongly Acceptable
AC1.4	4.55	0.55	Strongly Acceptable
AC1.5	4.65	0.48	Strongly Acceptable
Average:	4.53	0.51	Strongly Acceptable
AC2. Perceived ease of use			
AC 2.1	4.58	0.55	Strongly Acceptable
AC 2.2	4.45	0.55	Strongly Acceptable
AC 2.3	4.50	0.55	Strongly Acceptable
AC 2.4	4.50	0.51	Strongly Acceptable
AC 2.5	4.68	0.47	Strongly Acceptable
Average:	4.54	0.53	Strongly Acceptable
AC3. Attitude towards use			
AC3.1	4.65	0.48	Strongly Acceptable
AC3.2	4.53	0.51	Strongly Acceptable
AC3.3	4.50	0.51	Strongly Acceptable

AC3.4	4.58	0.50	Strongly Acceptable
AC3.5	4.65	0.48	Strongly Acceptable
Average:	4.56	0.50	Strongly Acceptable
Total Average:	4.54	0.51	Strongly Acceptable

5. Discussion

Three issues emerge from this work and require discussion.

- Evaluation of the component quality of the 2FA system prototype, which incorporates artificial intelligence technology specifically developed to improve academic information systems (AIS). Table 7 indicates that experts agree that the components and technologies used in the development process are very suited.

- The assessment of the efficiency of constructing a facial recognition model, as indicated in Table 9, resulted in the maximum degree of accuracy, reaching 83.54%. The error in evaluating efficiency can be ascribed to the limited availability of image datasets and the utilization of particular images in the experiment. The photos had diminutive dimensions and exhibited subpar resolution, leading to inaccuracies in identifying human faces and compromising the efficacy of image recognition. It is recommended that high-resolution pictures be used for future investigations and to explore the application of deep learning or other more efficient approaches.

- The data summary from the quality evaluation and acceptance study of the 2FA system prototype, which uses artificial intelligence technology created for academic information systems, is shown in Tables 10 and 11. Experts highly praised the system's overall quality, and consumers likewise strongly accepted the system's general use. The researcher has examined the factors contributing to identity verification challenges to integrate them into the development and design of information systems. The result is the creation of a prototype for a 2FA system that employs artificial intelligence technologies specifically tailored to assist academic information systems. By using the SDLC method, artificial intelligence technology can be used for two-factor authentication, improving the efficiency of identity verification and guaranteeing secure access to academic information systems. This academic information system efficiently manages, analyzes, and supervises activity data and offers responses to inquiries within the Department of Computer Technology and Digital, Faculty of Information Technology, Rajabhat Maha Sarakham University.

6. Conclusion

This article describes the research and development of the 2FA system prototype, which uses artificial intelligence technology developed to support academic information systems. The prototype consists of three main parts.

- Part 1. Studying the components of the 2FA system prototype to support academic information systems development, which consists of 4 modules: 1) Authentication Module, 2) Member Module, 3) Information Module, 4) Management Module, and experts agreed that the components and technology used in system development were overall strongly acceptable level.

- Part 2. The design and development of the 2FA system prototype, which uses artificial intelligence technology to emphasize accuracy, completeness, and security in verifying identity before accessing academic information systems, was divided into two phases. The first phase was developing a face recognition model consisting of a CNN Face Detector, VGG-Face net, and classification with Logistic Regression (LR), which achieved the highest accuracy of 83.54% compared to other methods. The second phase developed an Academic Information System (AIS) for username and password login combined with a facial recognition model to authenticate users through two-factor authentication. It was developed using PHP, Python, HTML5, CSS3, JavaScript, and a database system with MySQL. The developed system consisted of 2 types of users: teachers and students.
- Part 3. Evaluating system quality and inquiring about user acceptance of the 2FA system prototype to support academic information systems was divided into two components. - 1) Evaluation of the quality of the system by experts found that the opinions of the experts on the overall quality of the system were at the strongly acceptable level, and 2) Acceptance of the use of the 2FA system prototype to support academic information systems overall in a strongly acceptable level.

7. Limitations and Suggestions

7.1. Suggestions for Applying the Research Results

This research has produced a 2FA system prototype that used artificial intelligence technology and emphasized accuracy, completeness, and security in verifying identity before accessing academic information systems. However, suppose the results of this research are to be used practically. In that case, there is a need to collect more data used in facial recognition and training the facial image recognition module to recognize individuals.

7.2. Suggestions for Future Research

Future research should study the nature of identity verification by additional methods such as fingerprint scanning or personal voice recognition or increase recognition efficiency and combine each classification result obtained (ensemble method).

Acknowledgments

This research project was supported by the Rajabhat Maha Sarakham University, the Thailand Science Research and Innovation Fund, and the University of Phayao. In addition, this research was supported by many advisors, academics, researchers, staff, and students. The authors would like to thank all of them for their support and collaboration in making this research possible.

References

[1] Yasser M. Hausawi, William H. Allen, and Gisela Susanne Bahr, "Choice-Based Authentication: A Usable-Security Approach," *In Universal Access in Human-Computer Interaction, Design and Development Methods for Universal Access*, Heraklion, Crete, Greece, pp. 114-124, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[2] Vashek Matyáš, and Zdeněk Říha, "Security of Biometric Authentication Systems," *In 2010 International Conference on Computer Information Systems and Industrial Management Applications (CISIM)*, Krakow, Poland, pp. 19-28, 2010. [CrossRef] [Google Scholar] [Publisher Link]

[3] Sarah Abdulkader, Ayman Atia, and Mostafa-Sami Mostafa, "Authentication Systems: Principles and Threats," *Computer and Information Science*, vol. 8, no. 3, pp. 155-179, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[4] Tance Suleski et al., "A Review of Multi-Factor Authentication in The Internet of Healthcare Things," *Digital Health*, vol. 9, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[5] Zigang Chen et al., "FSMFA: Efficient Firmware-Secure Multi-Factor Authentication Protocol for IoT Devices," *Internet Things*, vol. 21, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[6] Tok Yen Xin, Norliza Katuk, and Ahmad Suki Che Mohamed Arif, "Smart Home Multi-Factor Authentication Using Face Recognition and One-Time Password on Smartphone," *International Journal of Interactive Mobile Technologies (IJIM)*, vol. 15, no. 24, pp. 32-48, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[7] Mariam. M. Taha et al., "On Password Strength Measurements: Password Entropy and Password Quality," *In 2013 International Conference on Computing, Electrical and Electronics Engineering (ICCEEE)*, Khartoum, Sudan, pp. 497-501, 2013. [CrossRef] [Google Scholar] [Publisher Link]

[8] Eliana Stavrou, "Enhancing Cyber Situational Awareness: A New Perspective of Password Auditing Tools," *In 2018 International Conference on Cyber Situational Awareness, Data Analytics and Assessment (Cyber SA)*, Glasgow, UK, pp. 1-4, 2018.

[9] Zahoor Ahmed Alizai, Noquia Fatima Tareen, and Iqra Jadoon, "Improved IoT Device Authentication Scheme Using Device Capability and Digital Signatures," *In 2018 International Conference on Applied and Engineering Mathematics (ICAEM)*, Taxila, Pakistan, pp. 1-5, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[10] Franck Favier, "Smart Cards and Healthcare," *Card Technology Today*, vol. 19, no. 11-12, 2007. [CrossRef] [Google Scholar] [Publisher Link]

[11] T. İ. Gündem, and Ö. Armağan, "Efficient Storage of Healthcare Data in XML-Based Smart Cards," *Computer Methods and Programs in Biomedicine*, vol. 81, no. 1, pp. 26-40, 2006. [CrossRef] [Google Scholar] [Publisher Link]

[12] Junbin Kang et al., "An Ultra-Light Weight and Secure RFID Batch Authentication Scheme for IoMT," *Computer Communications*, vol. 167, pp. 48-54, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[13] Sakhaa B. Al-Saedi, and Mohamed Mostafa A. Azim, "Radio Frequency Near Communication (RFNC) Technology: An Integrated RFID-NFC System for Objects' Localization," *In 2017 9th IEEE-GCC Conference and Exhibition (GCCCE)*, Manama, Bahrain, pp. 1-5, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[14] Anu Rathi et al., "Improvement of Existing Security System by Using Elliptic Curve and Biometric Cryptography," *In International Conference on Computing, Communication & Automation*, Greater Noida, India, pp. 994-998, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[15] Philip Black, "Building A Trusted Guardian for Our Biometric Identities," *Biometric Technology Today*, vol. 2021, no. 6, pp. 7-10, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[16] Chunhua Jin et al., "EBIAC: Efficient Biometric Identity-Based Access Control for Wireless Body Area Networks," *Journal of Systems Architecture*, vol. 121, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[17] Yasmin Makki Mohialden, Nadia Mahmood Hussien, and Doaa Muhsin Abd Ali, "Enhancing User Authentication with Facial Recognition and Feature-Based Credentials," *Journal La Multiapp*, vol. 4, no. 6, pp. 243-252, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[18] Mohammed Abdulhakim Al-Absi et al., "Real-Time Access Control System Method Using Face Recognition," *In Proceedings of International Conference on Smart Computing and Cyber Security*, Springer, Singapore, pp. 101-108, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[19] Gabit Tolendiyev, Hyotaek Lim, and Byung-Gook Lee, "A Margin-based Face Liveness Detection with Behavioral Confirmation," *International Journal of Internet, Broadcasting and Communication*, vol. 13, no. 2, pp. 187-194, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[20] R. Rameswari, "Automated Access Control System Using Face Recognition," *Materials Today: Proceedings*, vol. 45, pp. 1251-1256, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[21] Sheng Chen et al., "MobileFaceNets: Efficient CNNs for Accurate Real-Time Face Verification on Mobile Devices," *In Biometric Recognition, In Lecture Notes in Computer Science*, Springer, Cham, pp. 428-438, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[22] D. Sagar, and Murthy K. R. Narasimha, "Development and Simulation Analysis of a Robust Face Recognition Based Smart Locking System," *In Innovations in Electronics and Communication Engineering, in Lecture Notes in Networks and Systems*, Springer, Singapore, pp. 3-14, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[23] Ji-Hae Kim et al., "Efficient Facial Expression Recognition Algorithm Based on Hierarchical Deep Neural Network Structure," *IEEE Access*, vol. 7, pp. 41273-41285, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[24] Melda Agnes Manuhutu, Lulu Jola Uktolseja, and Sherly Gaspersz, "Academic Information System for Student (Case Study: Victory University of Sorong)," *International Journal of Computer Applications*, vol. 180, no. 43, pp. 26-33, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[25] H. P. Utomo, A. T. Bon, and M. Hendayun, "Academic Information System Support in the Era of Education 3.0," *International Research and Innovation Summit (IRIS2017)*, Melaka, Malaysia, vol. 226, no. 1, pp. 1-9, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[26] Etin Indrayani, "Management of Academic Information System (AIS) at Higher Education in the City of Bandung," *Procedia - Social and Behavioral Sciences*, vol. 103, pp. 628-636, 2013. [CrossRef] [Google Scholar] [Publisher Link]

International Journal of Engineering Trends and Technology
ISSN: 2231–5381 / https://doi.org/10.14445/22315381/IJETT-V73I1P114

Volume 73 Issue 1, 166-176, January 2025

Original Article

Evolutionary Feature Selection to Classify Elderly Diseases from Dietary and Exercise Habits and Emotions

Nitaya Buntao[1], Rada Somkhuean[2], Wongpanya S. Nuankaew[3], Pratya Nuankaew[4]

[1]*Department of Applied Statistics, Faculty of Science and Technology, Rajabhat Maha Sarakham University, Maha Sarakham, Thailand.*
[2]*Department of Mathematics, Faculty of Science and Agricultural Technology, Rajamangala University of Technology Lanna Chiang Mai, Thailand.*
[3]*Department of Computer Science, School of Information and Communication Technology, University of Phayao, Phayao, Thailand.*
[4]*Department of Digital Business, School of Information and Communication Technology, University of Phayao, Phayao, Thailand.*

[4]*Corresponding Author : pratya.nu@up.ac.th*

Received: 02 August 2024 Revised: 11 December 2024 Accepted: 17 December 2024 Published: 31 January 2025

Abstract - *The research investigates the effectiveness of various feature selection methods in enhancing disease classification models for elderly populations based on dietary habits, physical activity, and emotional well-being. It is conducted in Maha Sarakham Province, Thailand, and addresses critical healthcare challenges specific to this demographic. Traditional greedy algorithms (Forward Selection, Backward Elimination) are contrasted with metaheuristic approaches like evolutionary feature selection, evaluating their impact on accuracy and model robustness across classification algorithms (Deep Learning with H2O, Naïve Bayes, Gradient Boosted Trees, KNN, Decision Trees, Generalized Linear Models). Results show that evolutionary feature selection consistently outperforms traditional methods, achieving an average accuracy of 79.69% with Logistic Regression and Generalized Linear Models and demonstrating a superior balance between precision and recall. Deep Learning with H2O performs strongly across all methods, while Naïve Bayes benefits from Backward Elimination. The findings highlight the potential of evolutionary feature selection to enhance disease classification accuracy and model reliability, emphasizing the need for personalized healthcare strategies tailored to individual profiles in older adults.*

Keywords - *Evolutionary feature selection, Meta heuristic approaches, Personalized healthcare strategies, Elderly diseases, Healthcare challenges, Classification algorithms.*

1. Introduction

The world's senior population is booming, with the World Health Organization (WHO) predicting a doubling by 2050 [1]. This translates to a rise in chronic illnesses among older adults, which affect roughly 80% of this demographic [2]. Chronic conditions like cardiovascular diseases, diabetes, chronic respiratory disorders, and dementia take a toll on both quality of life and healthcare systems, and It is evident that innovative approaches to elder care and disease management are needed. The key to tackling this challenge lies in understanding how daily habits. People already know regular diet, exercise, and emotional well-being impact health. Research shows regular exercise can significantly reduce chronic disease risk, improve mental health, and boost physical function in older adults [3]). Similarly, studies highlight the positive effects of balanced nutrition and exercise on managing conditions like hypertension [4]. Beyond physical health, emotional well-being plays a crucial role. Research suggests managing negative emotions might be more effective than simply boosting positive ones to reduce inflammation in older adults [5]. It opens doors for exploring how emotional regulation can contribute to overall health. Previous research has explored the application of machine learning (ML) algorithms for early diagnosis of chronic diseases, emphasizing their potential to improve patient outcomes and treatment strategies. Feature selection is critical in developing accurate machine-learning models [6, 7]. The main feature selection methods include filtering, wrapping, and Embedding [8]. Advanced techniques like Genetic Algorithms (GAs) optimize feature subsets for complex datasets. Feature selection enhances model performance, reduces computational costs, and improves interpretability. The choice of method depends on the dataset size, interpretability requirements, and computational capacity [9]. This study identifies crucial features for accurate, efficient machine learning models in elderly disease classification. It

aims to enable proactive health management, potentially enhancing the aging population's quality of life. Accurately predicting diseases like mild cognitive impairment (MCI) in older adults is crucial. This research tackles this challenge by focusing on feature selection, a critical step in building robust machine-learning models. Finding the most important data: Forward Selection, Backward Elimination, Stepwise Selection, and Features Finally Used. Evaluate models directly to identify the most impactful features, leading to the best predictions [10]. Effective feature selection, like focusing on diet and exercise, helps build efficient and accurate models for predicting diseases in older adults. Studies show the positive impact of healthy habits. Research by Nitschke et al. (2022) analysed 82 studies and found that interventions promoting nutrition and physical activity improved weight, blood pressure, and blood sugar - all factors crucial for reducing chronic disease risk [11].

It highlights the importance of lifestyle interventions for overall health. For doctors to use these models effectively, clear explanations are essential. Abbas Saad Alatrany et al. (2024) proposed an approach for Alzheimer's disease (AD) classification that achieves high accuracy and provides clear explanations. It allows doctors to understand better the factors influencing the diagnosis [12]. In conclusion, effective feature selection and explainable machine-learning approaches are essential for improving disease prediction and management. These methods and lifestyle interventions play a critical role in promoting overall health and preventing chronic diseases. Advanced computational techniques, particularly evolutionary feature selection, effectively analyze large datasets to classify diseases among the elderly. These methods use diet, physical activity, and emotional state data to predict health outcomes accurately. Khanna et al. (2024) introduced a computer-aided diagnosis system for breast cancer classification using Teaching Learning-Based Optimization and Elephant Herding Optimization.

These methods improved classification accuracy and reduced unnecessary features [13]. Rashid et al. (2022) developed an AI-based method for chronic disease prediction, integrating Artificial Neural Networks with Particle Swarm Optimization. This approach focused on diseases like breast cancer, diabetes, and heart attack, outperforming traditional methods [14]. De Lacy et al. (2022) explored integrated evolutionary learning for complex medical datasets, automating features, and hyperparameter selection. These techniques aim to enhance the accuracy, reliability, and interpretability of disease diagnosis models for the elderly, potentially revolutionizing healthcare strategies for this growing population [15]. This research explores machine learning for identifying disease risks in older adults from Maha Sarakham, Thailand. It focuses on feature selection to pick crucial data (diet, activity, emotions) for accurate models. The research compares greedy algorithms (e.g., Forward Selection, Backward Elimination) with genetic algorithms (Evolutionary Method) to find the most effective approach for disease classification, aiming to improve early detection and health outcomes.

- Identify key factors contributing to disease prediction across different approaches for the elderly.
- Conduct a comparative analysis of these algorithms in elderly disease classification.
- Explore method synergies to enhance model robustness and accuracy.

This research contributes to geriatric health informatics by developing interpretable models. By explaining their reasoning, these models can inform personalized healthcare strategies in Thailand and globally.

2. Literature Review

Disease classification in elderly populations is vital for enhancing healthcare outcomes. Traditional feature selection methods, such as greedy algorithms (e.g., forward selection, backward elimination), sequentially select features based on individual contributions to model performance. However, these methods often lack feature interactions and struggle with large datasets. Advanced feature selection methods, like evolutionary algorithms, offer a more sophisticated approach. Inspired by natural selection, they explore and select features by considering their interactions and impact on model performance. This approach improves model accuracy by reducing dimensionality and identifying complex feature interactions. Effective feature selection is essential for developing accurate and interpretable machine learning models. Identifying key predictors allows researchers to tailor healthcare strategies for the elderly. The literature reveals a growing trend of using machine learning and Artificial Intelligence for disease classification and prediction in elderly populations, particularly for chronic conditions like Parkinson's, diabetes, and heart disease. Future directions could include integrating domain knowledge with evolutionary algorithms or developing more advanced feature selection techniques to enhance model performance further and advance personalized medicine for the elderly.

2.1. Disease Classification in Elderly Populations

Recent studies highlight a shift towards advanced computational methods for disease classification in geriatric medicine. Khera and Kumar (2020) proposed an ensemble learning classifier with optimal feature selection for Parkinson's disease, showcasing sophisticated algorithmic approaches in geriatric medicine [16]. Similarly, Qin et al. (2022) developed machine learning models for predicting diabetes based on lifestyle factors, emphasizing the importance of data-driven methods in managing chronic diseases in older adults [17]. Ali et al. (2023) also investigated Parkinson's disease detection using filter feature selection and genetic algorithms combined with ensemble learning, highlighting advanced computational methods in neurological

disorder diagnosis [18]. Further, Chawla et al. (2024) and Bhakar et al. (2024) focused on Parkinson's disease classification, employing nature-inspired feature selection methods and hybrid models, respectively. These studies demonstrate ongoing refinement in classification techniques for age-related neurological disorders [19, 20]. Collectively, these studies indicate a shift towards personalized and precise disease classification methods for elderly patients, leveraging machine learning to enhance diagnostic accuracy and inform tailored treatment strategies.

2.2. Existing Approaches to Feature Selection

2.2.1. Traditional Methods (Greedy Algorithms)

Feature selection is vital in developing predictive models for medical conditions like Chronic Kidney Disease (CKD) and Mild Cognitive Impairment (MCI). Traditional methods, including Filter, Wrapper, and Embedded approaches, provide foundational techniques in this area. These methods, such as forward selection, backward elimination, and stepwise regression, often serve as baseline comparisons in research studies. Recent investigations have demonstrated the significant impact of feature selection on model accuracy. For example, Zeynu and Patil (2018) showed that feature selection techniques substantially improved the precision of CKD prediction models. Their research used both Filter and Wrapper methods to refine the dataset and identify key attributes. Additionally, they implemented an ensemble model integrating multiple classifiers through a voting mechanism, enhancing prediction performance [21].

Similarly, Lim S-J et al. (2021) explored feature selection in predicting MCI using medical records. Their approach incorporated both Filter and Wrapper methods. The Filter method assessed individual features based on relevance, while the Wrapper method employed recursive elimination to identify optimal feature subsets. These strategies aimed to boost prediction accuracy by focusing on essential attributes and reducing dimensionality. The study also compared various classifiers to evaluate the impact of feature selection on model performance [10]. Additionally, Purwaningsih (2022) utilized forward selection to predict CKD, a technique that iteratively adds features to a Support Vector Machine (SVM) model based on their impact on performance. This approach seeks to identify the most relevant features for CKD detection, thereby enhancing the SVM's effectiveness. Despite its benefits, forward selection has limitations, including the potential for reduced generalizability due to the small dataset size and the possibility of overlooking important feature interactions. Broader feature selection methods or additional validation techniques could address these issues, potentially improving model robustness and applicability across various datasets [22]. More recently, K Hema et al. (2024) investigated feature selection techniques for early CKD prediction, employing Filter, Wrapper, and Embedded methods. Their study demonstrated that advanced feature selection methods improved prediction accuracy. The Filter method evaluated individual features' relevance, while the Wrapper method refined feature subsets through iterative approaches. Embedded methods optimized feature selection during model training. However, limitations included a lack of exploration of diverse techniques and potential overfitting due to dataset constraints. Future research should address these limitations by expanding feature selection techniques and testing on more varied datasets to enhance model robustness and generalizability [23]. In conclusion, while traditional feature selection methods have shown promise, there is significant room for advancement. By addressing current limitations and exploring more advanced techniques, future studies can contribute to developing even more accurate and reliable predictive models in healthcare.

2.2.2. Advanced Methods (Evolutionary Algorithms)

Recent research demonstrates the growing prominence of advanced, often nature-inspired or evolutionary approaches: The prominence of advanced feature selection algorithms in medical diagnostics is evident in recent research. Abdollahi and Nouri-Moghaddam (2021) evaluated a hybrid Stacked-Genetic approach for heart disease diagnosis, illustrating the integration of evolutionary algorithms in medical feature selection [24]. This trend highlights the effectiveness of evolutionary and nature-inspired algorithms in medical diagnosis and prediction tasks, often outperforming traditional greedy algorithms in accuracy and robustness for elderly disease classification.

Moreover, de Lacy et al. (2022) introduced an integrated evolutionary learning approach that simultaneously optimizes both feature selection and model parameters, showcasing a sophisticated method for medical diagnostics [15]. Ali et al. (2023) combined genetic algorithms with filter feature selection, demonstrating a hybrid approach that blends traditional and evolutionary methods [18]. Similarly, Chawla et al. (2023) employed nature-inspired feature selection techniques, moving towards bio-inspired optimization in healthcare data analysis [19]. Bhakar et al. (2024) also proposed a hybrid model incorporating random classification and feature selection, indicating a trend of combining multiple advanced techniques for enhanced performance [20]. Collectively, these studies reflect a shift towards more personalized and precise disease classification methods for elderly patients, leveraging machine learning to improve diagnostic accuracy and inform tailored treatment strategies.

2.3. Studies on the Impact of Diet, Exercise, and Emotional Health on Disease Outcomes in the Elderly

While most studies focus on computational methods, some address the impact of lifestyle factors on health outcomes: Studies on the Impact of Diet, Exercise, and Emotional Health on Disease Outcomes in the Elderly: While most of the provided papers focus on computational methods for disease classification, some address the impact of lifestyle factors on health outcomes in elderly populations. For

instance, Qin et al. (2022) developed machine learning models for diabetes prediction based on lifestyle types, implicitly considering factors such as diet and exercise in their analysis [17]. Additionally, Rashid et al. (2022) proposed an augmented artificial intelligence approach for chronic disease prediction, likely incorporating lifestyle factors as part of its predictive model [14].

Furthermore, while focusing on breast cancer, Khanna et al. (2024) developed an enhanced approach for chronic human disease prediction that could potentially be applied to lifestyle-related conditions in the elderly [13]. While traditional feature selection methods have proven valuable, several areas warrant further exploration. First, studies on more diverse elderly populations are needed to ensure the generalizability of findings. Second, incorporating environmental and lifestyle factors, such as diet, exercise, and sleep quality, could offer a more holistic understanding of disease risk in this population. Feature selection techniques have already shown promise in identifying the most impactful factors within these domains (e.g., Khanna et al., 2024). Finally, exploring more advanced feature selection techniques beyond traditional methods holds the potential to further improve model accuracy and robustness.

3. Methodology

3.1. Feature Selection Algorithm

In the realm of machine learning, feature selection stands as a cornerstone in model development, significantly enhancing efficiency and mitigating complexity. The research methodology often incorporates three principal approaches: Evolutionary Algorithm (EA), Forward Selection (FS), and Backward Elimination (BE). Evolutionary feature selection improves classification by identifying multiple optimal feature subsets through complex interactions using heuristic search methods. Techniques include multimodal optimization, differential evolution, duplication analysis, niching-based, binary differential evolution, and feature clustering-assisted selection. These methods select smaller feature subsets while maintaining accuracy, generating diverse non-dominated solutions, and reducing redundancy. Solutions with high diversity scores enhance population diversity. This approach excels in navigating intricate search spaces, making it effective for handling complex datasets [25, 26]. Forward feature selection begins with an empty feature set and adds features incrementally based on their contribution to model performance. Initially, it selects the feature that improves performance the most. The algorithm then evaluates combinations of the selected and remaining features, adding the feature with the highest performance boost. This process continues until a stopping criterion is met. It aims to maximize classification accuracy or minimize error rates while being computationally efficient and suitable for large datasets. However, its greedy approach may not always yield the optimal subset, and its success depends on the chosen selection criteria [27]. In contrast, Backward Elimination is a feature selection method that improves predictive models by removing the least significant features based on their statistical impact. By retaining only the most relevant variables, this technique enhances model accuracy and generalization. It simplifies the model, making it easier to interpret and less resource-intensive. Additionally, it speeds up the training process by reducing the number of features, leading to greater efficiency [28]. While these methodologies offer versatility across various data types, their efficacy is inherently tied to the specific research context. Astute researchers must carefully weigh factors such as dataset dimensions, problem intricacy, and available computational resources when selecting the most appropriate feature selection technique for their unique challenges.

3.2. Classification Algorithm

In academic research, various classification algorithms are employed to analyze and interpret complex datasets. These algorithms range from simple, intuitive methods to sophisticated machine-learning techniques, each with its own strengths and limitations. Deep Learning algorithms, particularly those implemented using platforms like H2O, represent the cutting edge of machine learning. These algorithms utilize multi-layered neural networks to extract features and learn from data, making them exceptionally adept at handling complex, high-dimensional datasets. They excel in tasks such as image classification and natural language processing. However, their power comes at a cost: they typically require large datasets, involve time-consuming training processes, and can produce results that are challenging to interpret [29].

On the other end of the spectrum, algorithms like Naïve Bayes operate on simpler principles. Naïve Bayes employs Bayes' probability theorem, assuming independence between features. This approach is user-friendly, computationally efficient, and effective for small to medium-sized datasets, making it particularly useful for tasks like text classification and spam detection. However, its underlying assumption of feature independence may not always hold in real-world scenarios [30]. Gradient Boosted Trees offer a middle ground, combining multiple decision trees to create powerful predictive models. This method effectively manages complex and imbalanced datasets, making it suitable for data with non-linear relationships and numerous features.

However, careful parameter tuning is required to avoid overfitting, which can involve lengthy training periods [31]. The K-Nearest Neighbors (KNN) algorithm provides an intuitive approach to classification, basing its decisions on the K nearest data points in the training set. While it's easy to implement and makes no assumptions about data distribution, its performance can degrade with high-dimensional data, and prediction times increase for large datasets [32]. Decision Trees offer a highly interpretable model, constructing a tree-like structure where each node represents a feature-based

decision. This approach is particularly valuable when model explainability is crucial. However, Decision Trees are prone to overfitting, especially when allowed to grow too deep [33].

For data that follows specific probability distributions, Generalized Linear Models (GLMs) extend the concepts of linear regression beyond normal distributions. GLMs are flexible and capable of elucidating variable relationships, but they may struggle with highly complex, non-linear relationships and require statistical expertise for proper interpretation [34]. Logistic regression is a tool for binary classification, like predicting heart disease. It evaluates how risk factors (such as high cholesterol and smoking) relate to the likelihood of developing cardiovascular disease.

The model computes probabilities based on input features and classifies individuals accordingly. Stored with the Pickle library for convenient deployment and reuse, it uses a logistic function to convert features into probabilities for binary predictions [35]. When selecting an appropriate algorithm for classification tasks, it is crucial to consider various factors, including the characteristics and size of the dataset, the complexity of the problem at hand, requirements for result interpretation, and available computational resources.

Experimenting with multiple algorithms and comparing their performance using metrics such as Accuracy, F1-score, or Area Under the ROC Curve (AUC-ROC) is often beneficial. This empirical approach allows the identification of the most suitable model for specific research endeavors, balancing predictive power with interpretability and computational efficiency.

3.3. Preprocessing

3.3.1. Data Collection

This research collected data from 215 elderly individuals aged 60 and above residing in the Kang Leung Chan Sub-district, Mueang District of Maha Sarakham Province. These participants were selected from a total population of 1,505 elderly individuals in Maha Sarakham Province, Thailand, between 2021 and 2022. The data collection tool was a researcher-adapted questionnaire consisting of two sections:

Section 1: General Information Questionnaire

This section includes questions on gender, age, weight, height, marital status, educational level, occupation, income, source of income, marital status, living conditions, household status, caregiver, history of alcohol consumption, smoking history, and chronic diseases. Respondents are asked to fill in or select the information that corresponds to their own.

Section 2: Questionnaire on Eating Habits, Exercise, and Mood

This section features 19 questions where respondents select answers by marking the appropriate box. It uses a rating scale to classify behaviors into three levels: regular (5 - 7 days/week), occasional (1 - 4 days/week), and never. The assessed behaviors include smoking, sleeping, and drinking water, as outlined in Table 1. The researchers conducted the data collection process by explaining the purpose of the data collection and describing the nature of the questionnaire, including how to respond. The researchers personally gathered the data through interviews, allowing participants to complete the questionnaire themselves. The researchers then verified the accuracy and completeness of the questionnaires, recorded the data, and documented the process with photographs as evidence.

Table 1. Questionnaire on Eating Habits, Exercise, and Mood

Question	Feature name	Choice
Do you smoke?	Smoking	Yes No
How many hours did you sleep per night on average in the past week?	Sleep_per_night	Less than 5 hours/night 5 – 6 hours/night 7 – 8 hours/night
How often do you drink at least 8 glasses of water per day in a week?	Drink_water_per_day	1 – 3 days/a week 4 – 6 days/a week 7 days /a week
Eating habits, exercise, and mood behaviors		
Do you consume a balanced diet consisting of all five food groups (meat-dairy-eggs, grains, vegetables, fruits, and oils)?	Feature_Q1	Regular (5-7 days/week) Occasional (1-4 days/week) Never
Do you have breakfast as your main meal?	Feature_Q2	
Do you eat at least six servings of vegetables per day?	Feature_Q3	
Do you eat 4-5 servings of fruit per day (one serving equals 6-8 bites)?	Feature_Q4	
Do you eat fish at least once a day?	Feature_Q5	
Do you eat lean meat 2-3 times per week?	Feature_Q6	
Do you drink plain milk, low-fat milk, skim milk, or unsweetened soy milk with black sesame once or twice a day?	Feature_Q7	
Do you eat dinner at least 4 hours before bedtime?	Feature_Q8	

Do you consume foods that are boiled, steamed, blanched, baked, or grilled?	Feature_Q9	
Do you avoid high-fat foods?	Feature_Q10	
Do you avoid drinks, desserts, and snacks high in flour and sugar or very sweet?	Feature_Q11	
Do you eat bland food?	Feature_Q12	
Do you choose to drink water instead of soda or sweetened beverages?	Feature_Q13	
Do you avoid alcoholic beverages?	Feature_Q14	
Do you maintain a good mood and avoid stress?	Feature_Q15	
Do you sleep at least 7-8 hours per night?	Feature_Q16	
Do you exercise 5 days a week or 5 times a week?	Feature_Q17	
Do you exercise for at least 30 minutes a day?	Feature_Q18	
During exercise, do you breathe faster than usual and break a sweat?	Feature_Q19	

Table 2. Data Preparation

Category	Male	Female
Total Number	82	129
Average Age (years)	69.39	68.62
Average weight (kg)	56.16	57.21
Average height (cm)	160.95	158.36
Has Chronic Illness	43	83
Has More Than 1 Chronic Illness	39	32
No Chronic Illness	18	46
Exercises Regularly	62	112
Smokes	12	9
Sleeps More Than 5 Hours/Day	74	121
Drinks At Least 8 Glasses of Water/Day/Week	81	128

3.3.2. Data Preparation

After collecting data from the questionnaire, researchers proceeded with preparing the data for use in model building, which involved the following steps:

- Data Cleaning:
  Addressing errors and inconsistencies like missing values, duplicates, and outliers.
- Data Transformation:
  Converting data into a suitable format for analysis, including normalization, scaling, encoding categorical variables, and aggregating data.
- Data Integration:
  Consolidating data from different sources into a cohesive dataset, maintaining consistency and integrity.
- Data Reduction:
  Simplifying the dataset by selecting relevant features, aggregating data, and removing redundant or irrelevant information.
- Data Validation:
  Ensuring data accuracy and quality through consistency checks and verification against established benchmarks.
- Data Formatting:
  Structuring data for analysis or modelling, organizing it into tables with appropriate headers and ensuring consistent data types.
- Data Splitting:
  Dividing data into training and testing sets, 70:30 was used to evaluate model and algorithm performance.

These steps ensure the data is accurate, consistent, and ready for analysis, leading to more reliable and meaningful results. The 22-question questionnaire on eating habits, exercise, and mood was used for data modelling, with chronic diseases (Yes/No) as the class label derived from the general information questionnaire. Following data preparation, the dataset comprised 211 elderly individuals, as detailed in Table 2.

3.4. Modelling

After Data Preparation, the modelling process involve comparing feature selection methods: Evolutionary Algorithm (EA), Forward Selection (FS), and Backward Elimination (BE). These assess which variables best enhance model performance. Additionally, various classification algorithms are compared: Deep Learning (H2O), Naïve Bayes, Gradient Boosted Trees, K-Nearest Neighbours (KNN), Decision Trees, Generalized Linear Models, and Logistic Regression. Evaluation criteria include accuracy, precision, recall, F1-score, and computational efficiency to identify the optimal approach for the dataset's needs.

3.4.1. First Objective

Conduct a comparative analysis of greedy algorithms (e.g., Forward Selection, Backward Elimination) versus metaheuristic algorithms (e.g., evolutionary methods) to enhance disease classification accuracy for the elderly. Identify key factors by analysing the overlap and uniqueness of selected variables and assessing the impact of dietary

habits, physical activity, and emotional well-being on disease prediction.

3.4.2. Second Objective

Conducting a comparative analysis of classification models is essential to evaluate the performance of greedy feature selection algorithms versus evolutionary algorithms. This study employs various models, including Deep Learning (H2O), Naïve Bayes, Gradient Boosted Trees, K-Nearest Neighbours (KNN), Decision Trees, Generalized Linear Models, and Logistic Regression, to improve disease classification accuracy for elderly populations. Greedy feature selection algorithms, such as Forward Selection and Backward Elimination, incrementally add or remove features based on immediate performance impact. These methods are efficient but may overlook complex feature interactions. In contrast, evolutionary algorithms use metaheuristic techniques inspired by natural evolution, such as mutation, crossover, and selection, to explore a broader search space. These methods handle complex, high-dimensional datasets effectively, uncovering intricate feature interactions. The analysis aims to identify the most effective feature selection method for improving disease prediction accuracy in elderly populations. It considers dietary habits, physical activity levels, and emotional well-being, comprehensively evaluating factors contributing to disease outcomes. This study offers insights into the strengths and limitations of greedy and evolutionary feature selection methods, guiding the choice of techniques to enhance disease classification accuracy and leading to better health outcomes and targeted interventions. The modeling process in this research is outlined in Algorithm 1.

Algorithm1: Framework process

Input: Training set, Testing set

1. Read the Training Set
2. Define the range of training data (i to j) and attributes (m to n)
3. Define classifiers c (c_1, c_2, ..., c_k)
4. Define attribute weights w (w_1, w_2, ..., w_z)
5. Compute attribute weights w_1 to w_z using Forward Selection, Backward Elimination, and Evolutionary Methods
6. Rank attributes by weight for each feature selection method
7. Define rankings r (r_1, r_2, ..., r_z) from highest to lowest weight for each method
8. Select attributes from ranked list r (1 to z) for the best classification using c_1 to c_k
9. Build classification models using the selected attributes from each feature selection method
10. Read the Testing Set
11. Evaluate the models

Output:

1. Attribute weight values
2. Accuracy, Precision, Recall, and F1 Score values

3.5. Evaluation

The evaluation employs statistical methods to assess the significance of differences in classification performance between greedy and evolutionary algorithms. Model quality was assessed using a 70:30 split of training and testing sets. Efficiency was measured using the following metrics: accuracy, precision, recall, and F1 score, as defined by the equations below [23]:

$$\text{Accuracy}=\frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$\text{Precision}=\frac{TP}{TP+FP} \quad (2)$$

$$\text{Recall}=\frac{TP}{TP+FN} \quad (3)$$

$$\text{F1}=\frac{2 * (Precision * Recall)}{(Precision + Recall)} \quad (4)$$

In these equations, TP, FP, TN, and FN refer to true positive, false positive, true negative, and false negative counts, respectively.

4. Results and Discussion

4.1. Research Results

This study explores feature selection methods for classifying diseases in elderly populations, emphasizing dietary habits, exercise, and emotional well-being as predictors. It compares Evolutionary Algorithms (EA), Forward Selection (FS), and Backward Elimination (BE) to enhance predictive accuracy and guide targeted healthcare interventions for the second objective, as detailed in Table 3. For the first objective, the research focuses on features with a weight value greater than 0.05, as listed below. The Evolutionary Algorithm identified nine key features: Smoking, Feature_Q1, Feature_Q2, Feature_Q3, Feature_Q10, Feature_Q11, Feature_Q13, Feature_Q15, Feature_Q17, and Feature_Q18. This focus underscores the significance of dietary habits and lifestyle in maintaining overall well-being. For dietary habits, the algorithm highlighted essential features such as consuming a balanced diet (Feature_Q1), eating breakfast regularly (Feature_Q2), consuming at least six servings of vegetables daily (Feature_Q3), avoiding high-fat foods (Feature_Q10), and

eating bland food (Feature_Q11). These aspects are crucial for health promotion and disease prevention. In terms of lifestyle, the selected features include smoking, choosing water over sugary drinks (Feature_Q13), getting adequate sleep (Feature_Q15), and engaging in regular exercise (Feature_Q17 and Feature_Q18). These behaviors are vital components of a healthy lifestyle and significantly impact overall health.

The Forward Algorithm selected thirteen key features: Smoking, Sleep_per_night, Drink_water_per_day, Feature_Q1, Feature_Q2, Feature_Q3, Feature_Q4, Feature_Q10, Feature_Q12, Feature_Q11, Feature_Q13, Feature_Q19, and Feature_Q9. This algorithm identifies essential features for evaluating health-related behaviors, focusing on smoking, sleep, hydration, diet, and exercise. It assesses smoking habits, average sleep duration (Sleep_per_night), and daily water intake (Drink_water_per_day), which are crucial for hydration. Dietary aspects include whether the individual consumes a balanced diet (Feature_Q1), eats breakfast regularly (Feature_Q2), and their daily intake of vegetables (Feature_Q3) and fruits (Feature_Q4). The algorithm also evaluates avoidance of high-fat foods (Feature_Q10), preference for bland food (Feature_Q12), avoidance of sugary foods (Feature_Q11), and choosing water over sugary drinks (Feature_Q13). Additionally, it measures exercise intensity (Feature_Q19) and preference for boiled, steamed, or grilled foods (Feature_Q9).

These features provide a comprehensive view of lifestyle factors impacting overall health and well-being, incorporating specific measures such as average sleep hours, daily water consumption, and dietary and exercise-related factors. The Backward Elimination algorithm selected thirteen features: Smoking, Sleep_per_night, Feature_Q2, Feature_Q3, Feature_Q4, Feature_Q7, Feature_Q9, Feature_Q10, Feature_Q11, Feature_Q12, Feature_Q13, Feature_Q14, and Feature_Q16. It has identified key features for assessing health-related behaviors. These include smoking (whether the individual smokes) and sleep_per_night (average hours of sleep). Dietary habits are evaluated through breakfast frequency (Feature_Q2), daily vegetable intake (Feature_Q3), and daily fruit intake (Feature_Q4). Additional factors include milk or soymilk consumption (Feature_Q7), preference for boiled, steamed, or grilled foods (Feature_Q9), and avoidance of high-fat foods (Feature_Q10), sugary foods (Feature_Q11), and alcohol (Feature_Q14) indicates a preference for bland food (Feature_Q12), measures choosing water over sugary drinks (Feature_Q13) and checks if the individual gets at least 7-8 hours of sleep (Feature_Q16).

The Backward Elimination algorithm introduced unique elements like milk or soymilk consumption and alcohol avoidance, which were not featured in the other algorithms, offering additional insights into health-related behaviors. Based on a comprehensive analysis of various algorithms and feature selection methods, evolutionary feature selection consistently achieved the highest accuracy, notably with Logistic Regression and Generalized Linear Models averaging 79.69%. Precision and recall metrics varied across methods, with evolutionary approaches demonstrating superior balance compared to Forward and Backward elimination techniques.

F1 scores, reflecting the harmonic mean of precision and recall, also favored evolutionary methods across diverse algorithms. Deep learning using H2O showed consistently strong performance across all feature selection methods, maintaining high accuracy, precision, recall, and F1 scores with minimal variation. Naïve Bayes performed well in precision and recall, especially enhanced by Backward Elimination. Gradient Boosted Trees, KNN, Decision Trees, and Generalized Linear Models exhibited mixed performance across different feature selection techniques, with evolutionary methods generally providing more stable outcomes.

Overall, Evolutionary Feature Selection emerged as the preferred method due to its superior performance in accuracy, precision, recall, and F1 scores across various classification algorithms. This underscores its potential for optimizing disease classification models in elderly populations based on dietary habits, physical activity, and emotional well-being. Leveraging evolutionary methods, particularly with Logistic Regression and Deep Learning using H2O, is recommended for enhancing model robustness and predictive accuracy in healthcare applications.

Third Objective: Explore synergies between methods to enhance model robustness and accuracy. To improve model performance, especially for elderly populations, integrating Evolutionary Feature Selection with algorithms like Logistic Regression and Deep Learning is proposed.

Table 3. Research Results

	Accuracy	Precision		Recall		F1-Score	
		Yes	No	Yes	No	Yes	No
Forward Selection							
Deep Learning algorithm using H2O	75.00	72.00	85.71	94.74	46.15	81.82	60.00
Naïve Bayes	71.88	69.23	83.33	94.74	38.46	80.00	52.63
Gradient Boosted Trees	70.31	75.68	62.96	73.68	65.38	74.67	64.15

KNN	70.31	70.17	65.22	78.95	57.69	75.95	61.22
Decision Tree	70.31	68.63	76.92	92.11	38.46	78.65	51.28
Generalized Linear Model	71.88	69.23	83.33	94.74	38.46	80.00	52.63
Logistic Regression	73.44	78.38	66.67	76.23	69.23	77.34	67.92
Backward Elimination							
Deep Learning algorithm using H2O	76.56	81.08	70.37	78.95	73.08	80.00	71.70
Naïve Bayes	75.00	84.38	65.62	71.05	80.77	77.14	72.41
Gradient Boosted Trees	71.88	73.81	68.18	81.58	57.69	77.75	62.49
KNN	71.88	70.83	75.00	89.47	46.15	79.07	57.14
Decision Tree	71.88	72.73	70.00	84.21	53.85	78.05	60.87
Generalized Linear Model	75.00	80.56	67.86	76.32	73.08	78.37	70.37
Logistic Regression	76.56	84.85	67.74	73.68	80.77	78.87	73.68
Evolutionary							
Deep Learning algorithm using H2O	78.12	81.58	73.08	81.58	73.08	81.59	73.08
Naïve Bayes	76.56	82.86	68.97	76.32	76.92	79.45	72.72
Gradient Boosted Trees	73.44	72.34	76.47	89.47	50.00	80.00	60.47
KNN	73.44	76.92	68.00	78.95	65.38	77.92	66.67
Decision Tree	78.12	83.33	71.43	78.95	76.92	81.08	74.07
Generalized Linear Model	79.69	80.49	78.26	86.84	69.23	83.54	73.47
Logistic Regression	79.69	83.78	74.07	81.58	76.92	82.67	75.47

This combined approach leverages the strengths of each method to optimize disease classification models by prioritizing factors such as diet, physical activity, and emotional well-being.

4.2. Discussion

Investigating feature selection methods for disease classification among elderly populations based on dietary habits, exercise routines, and emotional well-being provides crucial insights for advancing healthcare practices. Quantitative analysis highlights evolutionary approaches as particularly effective, with Logistic Regression and Generalized Linear Models achieving notable average accuracies of 79.69%. Evolutionary methods excel in balancing precision and recall metrics compared to traditional Forward and Backward elimination methods, underscoring their superiority. These methods consistently identify predictive factors like dietary habits, exercise routines, and emotional well-being indicators, offering nuanced insights into health outcomes among older adults.

They outperform traditional approaches across various classification algorithms by managing complex feature interactions, thereby enhancing model robustness and predictive accuracy. Conversely, traditional methods often struggle to maintain this balance, potentially overlooking critical relationships between dietary, exercise, and emotional variables. Understanding these synergistic relationships is pivotal for effective disease classification in elderly populations. Evolutionary feature selection effectively captures these dynamics, revealing how specific dietary patterns and exercise frequencies influence both emotional well-being and physical health outcomes. This comprehensive understanding informs tailored healthcare interventions integrating dietary modifications, personalized exercise regimens, and emotional support strategies to improve disease prevention and management among older adults.

5. Conclusion

In conclusion, this study rigorously evaluates various feature selection methods to enhance disease classification models for elderly populations, specifically focusing on dietary, exercise, and emotional factors. Evolutionary Algorithms (EA) are highlighted for consistently achieving superior predictive accuracy, precision, recall, and balanced F1 scores across diverse algorithms, effectively identifying critical predictive features and revealing nuanced relationships between lifestyle factors and health outcomes among older adults.

In contrast, traditional methods like Forward Selection (FS) and Backward Elimination (BE) show variable performance, often grappling with precision-recall trade-offs and occasionally missing subtle yet significant feature interactions. While each method offers valuable insights into feature relevance, Evolutionary Algorithms emerge as the optimal choice for enhancing model robustness and accuracy in complex healthcare scenarios. Moving forward, further research should extend algorithm comparisons to include longitudinal studies and considerations of ethical implications. Addressing study limitations such as sample size constraints and data quality issues will be pivotal in enhancing the generalizability and applicability of predictive models across diverse healthcare settings. In summary, the integration of evolutionary feature selection methods marks a crucial advancement in geriatric healthcare, fostering more precise disease classification models that cater to the evolving needs of aging populations worldwide. These advancements hold promise for shaping future healthcare strategies, ultimately

enhancing the quality of life and health outcomes for elderly individuals globally.

Limitation

Collecting data from elderly individuals can be challenging due to age-related factors such as vision problems, speech difficulties, or cognitive impairments like dementia. These challenges necessitate using clear language, extra support during interviews or surveys, and adaptive methods to ensure elderly participants can comfortably engage and provide accurate information. Respecting their abilities and ensuring their comfort during the data collection process is crucial.

Acknowledgements

This research was supported by the Thailand Science Research and Innovation (TSRI) through a grant fund for Rajabhat Maha Sarakham University. Numerous advisors and researchers from the University of Phayao also provided partial support. The authors extend their gratitude to all for their support and cooperation in completing this research, facilitated by the Research and Development Institute (RDI) of Rajabhat Maha Sarakham University. Special thanks are extended for the research data support from the population in Kang Leung Chan Sub-district, Mueang District of Maha Sarakham Province.

References

[1] Ageing and Health, World Health Organization, 2024. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health

[2] Efraim Jaul, and Jeremy Barron, "Age-Related Diseases and Clinical and Public Health Implications for the 85 Years Old and Over Population," *Frontiers in Public Health*, vol. 5, pp. 1-7, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[3] Pawel Posadzki et al., "Exercise/Physical Activity and Health Outcomes: An Overview of Cochrane Systematic Reviews," *BMC Public Health*, vol. 20, pp. 1-12, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[4] Leonardo Santos Lopes da Silva et al., "Nutritional Status, Health Risk Behaviors, and Eating Habits are Correlated with Physical Activity and Exercise of Brazilian Older Hypertensive Adults: A Cross-Sectional Study," *BMC Public Health*, vol. 22, pp. 1-12, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[5] Jennifer E. Graham-Engeland et al., "Negative and Positive Affect as Predictors of Inflammation: Timing Matters," *Brain, Behavior, and Immunity*, vol. 74, pp. 222-230, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[6] Wongpanya Nuankaew, and Jaree Thongkam, "Improving Student Academic Performance Prediction Models Using Feature Selection," *2020 17th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, Phuket, Thailand, pp. 392-395, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[7] Rakibul Islam, Azrin Sultana, and Mohammad Rashedul Islam, "A Comprehensive Review for Chronic Disease Prediction Using Machine Learning Algorithms," *Journal of Electrical Systems and Information Technology*, vol. 11, no. 1, pp. 1-28, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[8] Girish Chandrashekar, and Ferat Sahin, "A Survey on Feature Selection Methods," *Computers and Electrical Engineering*, vol. 40, no. 1, pp. 16-28, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[9] Miguel García-Torres, Roberto Ruiz, and Federico Divina, "Evolutionary Feature Selection on High Dimensional Data Using a Search Space Reduction Approach," *Engineering Applications of Artificial Intelligence*, vol. 117, pp. 1-15, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[10] Soo-Jin Lim et al., "Medical Health Records-Based Mild Cognitive Impairment (MCI) Prediction for Effective Dementia Care," *International Journal of Environmental Research and Public Health*, vol. 18, no. 17, pp. 1-15, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[11] Erin Nitschke et al., "Impact of Nutrition and Physical Activity Interventions Provided by Nutrition and Exercise Practitioners for the Adult General Population: A Systematic Review and Meta-Analysis," *Nutrients*, vol. 14, no. 9, pp. 1-33, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[12] Abbas Saad Alatrany et al., "An Explainable Machine Learning Approach for Alzheimer's Disease Classification," *Scientific Reports*, vol. 14, no. 1, pp. 1-18, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[13] Munish Khanna et al., "An Enhanced and Efficient Approach for Feature Selection for Chronic Human Disease Prediction: A Breast Cancer Study," *Heliyon*, vol. 10, no. 5, pp. 1-21, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[14] Junaid Rashid et al., "An Augmented Artificial Intelligence Approach for Chronic Diseases Prediction," *Frontiers in Public Health*, vol. 10, pp. 1-20, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[15] Nina de Lacy, Michael J. Ramshaw, and J. Nathan Kutz, "Integrated Evolutionary Learning: An Artificial Intelligence Approach to Joint Learning of Features and Hyperparameters for Optimized, Explainable Machine Learning," *Frontiers in Artificial Intelligence*, vol. 5, pp. 1-16, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[16] Preeti Khera, and Neelesh Kumar, "Ensemble Learning Classifier with Optimal Feature Selection for Parkinson's Disease," *2020 International Conference on Advances in Computing, Communication & Materials*, Dehradun, India, pp. 427-431, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[17] Yifan Qin et al., "Machine Learning Models for Data-Driven Prediction of Diabetes by Lifestyle Type," *International Journal of Environmental Research and Public Health*, vol. 19, no. 22, pp. 1-16, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[18] Abdullah Marish Ali, Farsana Salim, and Faisal Saeed, "Parkinson's Disease Detection Using Filter Feature Selection and a Genetic Algorithm with Ensemble Learning," *Diagnostics*, vol. 13, no. 17, pp. 1-14, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[19] Prabhleen Kaur Chawla et al., "Parkinson's Disease Classification Using Nature Inspired Feature Selection and Recursive Feature Elimination," *Multimedia Tools and Applications*, vol. 83, pp. 35197-35220, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[20] Suman Bhakar et al., "A Hybrid Model: Random Classification and Feature Selection Approach for Diagnosis of the Parkinson Syndrome," *Scalable Computing: Practice and Experience*, vol. 25, no. 1, pp. 167-176, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[21] Sirage Zeynu, and Shruti Patil, "Prediction of Chronic Kidney Disease Using Data Mining Feature Selection and Ensemble Method," *WSEAS Transactions on Information Science and Applications*, vol. 15, pp. 168-176, 2018. [Google Scholar] [Publisher Link]

[22] Esty Purwaningsih, "Improving the Performance of Support Vector Machine with Forward Selection for Prediction of Chronic Kidney Disease," *Journal of Computer Science and Technology*, vol. 8, no. 1, pp. 18-24, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[23] K. Hema, K. Meena, and Ramaraj Pandian, "Analyze the Impact of Feature Selection Techniques in the Early Prediction of CKD," *International Journal of Cognitive Computing in Engineering*, vol. 5, pp. 66-77, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[24] Jafar Abdollahi, and Babak Nouri-Moghaddam, "Feature Selection for Medical Diagnosis: Evaluation for Using a Hybrid STACKED-Genetic Approach in the Diagnosis of Heart Disease," *Arxiv*, pp. 1-11, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[25] Peng Wang et al., "Multiobjective Differential Evolution for Feature Selection in Classification," *IEEE Transactions on Cybernetics*, vol. 53, no. 7, pp. 4579-4593, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[26] Wongpanya S. Nuankaew, Sittichai Bussaman, and Pratya Nuankaew, "Evolutionary Feature Weighting Optimization and Majority Voting Ensemble Learning for Curriculum Recommendation in the Higher Education," *15th International Conference: Multi-Disciplinary Trends in Artificial Intelligence*, Virtual Event, pp. 14-25, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[27] Afnan M. Alhassan, and Wan Mohd Nazmee Wan Zainon, "Review of Feature Selection, Dimensionality Reduction and Classification for Chronic Disease Diagnosis," *IEEE Access*, vol. 9, pp. 87310-87317, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[28] F. Maulidina et al., "Feature Optimization Using Backward Elimination and Support Vector Machines (SVM) Algorithm for Diabetes Classification," *Journal of Physics: Conference Series*, *International Conference on Mathematics: Pure, Applied and Computation*, Surabaya, Indonesia (virtual), vol. 1821, pp. 1-8, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[29] Erin LeDell, and Sebastien Poirier, "H2O AutoML: Scalable Automatic Machine Learning," *7th ICML Workshop on Automated Machine Learning*, pp. 1-16, 2020. [Google Scholar] [Publisher Link]

[30] Harry Zhang, and Jiang Su, "Naive Bayes for Optimal Ranking," *Journal of Experimental and Theoretical Artificial Intelligence*, vol. 20, no. 2, pp. 79-93, 2008. [CrossRef] [Google Scholar] [Publisher Link]

[31] Tianqi Chen, and Carlos Guestrin, "XGBoost: A Scalable Tree Boosting System," *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, California, USA, pp. 785-794, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[32] T. Cover, and P. Hart, "Nearest Neighbor Pattern Classification," *IEEE Transactions on Information Theory*, vol. 13, no. 1, pp. 21-27, 1967. [CrossRef] [Google Scholar] [Publisher Link]

[33] J.R. Quinlan, "Induction of Decision Trees," *Machine Learning*, vol. 1, pp. 81-106, 1986. [CrossRef] [Google Scholar] [Publisher Link]

[34] J.A. Nelder, and R.W.M. Wedderburn, "Generalized Linear Models," *Journal of the Royal Statistical Society. Series A (General)*, vol. 135, no. 3, pp. 370-384, 1972. [CrossRef] [Google Scholar] [Publisher Link]

[35] Faris Hrvat, Lemana Spahić, and Amina Aleta, "Heart Disease Prediction Using Logistic Regression Machine Learning Model," *Proceedings of the Mediterranean Conference on Medical and Biological Engineering and Computing (MEDICON) and International Conference on Medical and Biological Engineering (CMBEBIH)*, Sarajevo, Bosnia and Herzegovina, vol. 1, pp. 654-662, 2024. [CrossRef] [Google Scholar] [Publisher Link]

AI for Healthcare: A Classification Model for Personalized Premenstrual Symptoms and Depressive Crisis Risk Tracking Using Data Analytics and Machine Learning

Pratya Nuankaew[1], Jidapa Sorat[2], Jindaporn Intajak[3], Jirapron Inta[4], Wongpan S. Nuankaew[5,*]
[1, 2, 3, 4, 5] School of Information and Communicatin Technology, University of Phayao, Phayao 56000 Thailand
pratya.nu@up.ac.th[1], 65020788@up.ac.th[2], 65020799@up.ac.th[3], 65025389@up.ac.th[4],
*Corresponding Author: wongpanya.nu@up.ac.th[5]

***Abstract* — A significant obstacle to women's lives is menstruation, which, in some cases, leads to depression, anxiety, and stress. This research, therefore, aims to solve this problem with three objectives: to study the context of female adolescents who experience premenstrual syndrome (PMS) problems and depression crises, to develop a classification model for personalized premenstrual symptoms and depressive crisis risk tracking using data analytics and machine learning techniques, and to evaluate the performance of all developed models to select suitable models for future application. The research data included 282 volunteers who gave consent from the University of Phayao, with random sampling. The machine learning used in this research consists of six techniques: decision trees, k-nearest neighbors, logistic regression, naïve bayes, random forests, and support vector machines. The model performance is tested using 10-fold cross-validation and evaluated using five metrics: accuracy, precision, recall, f1-score, and time. The results of the construction and evaluation of all models are at a high level, which is encouraging. The researchers expect that this research can be further developed for future mobile applications.**

Keywords* — *AI for Healthcare, Depressive Crisis Risk Tracking, Medical Data Analytics, Medical Informatics, Personalized Premenstrual Symptom

I. INTRODUCTION

Premenstrual Syndrome (PMS) is a significant health issue affecting women of reproductive age worldwide. Studies show that about 47.80% of women experience PMS, with 3-8% having severe symptoms [1]. Women with PMS often experience both physical and mental symptoms, such as stomach pain, breast tenderness, irritability, and depression. These symptoms are closely linked to hormone changes, significant changes in estrogen and allopregnanolone, and imbalances in brain chemicals like GABA.

Current treatments for PMS and its more severe form, Premenstrual Dysphoric Disorder (PMDD), usually involve medications like antidepressants (SSRIs) and hormone therapy [1]. However, these methods often don't fully help with all symptoms because PMS involves complex interactions between hormones, the nervous system, and behavior. It shows women need more personalized treatment approaches based on individual patient data. The severity of PMS symptoms poses a significant challenge for women, particularly students, in obtaining help. These symptoms can interfere with their academic performance and daily routines, influencing their physical and mental health [2], [3]. Furthermore, younger individuals were found to experience greater severity of PMS symptoms.

Recent advances in Artificial Intelligence (AI) and Machine Learning are creating new opportunities in healthcare, especially for conditions [4] that need personalized risk assessment and treatment. Using AI-based classification models, researchers can combine different types of data—including hormone levels, lifestyle factors, and symptom history—to predict and track each person's risk levels. This approach not only helps make diagnosis more accurate but also allows for better management of PMS and related depression.

This research proposes to apply Machine Learning using a classification model to track and predict PMS symptoms and depression risks. Using effective data analysis and machine learning techniques, this model will help overcome the limitations of traditional treatment methods. It will offer a personalized solution that fits each person's specific needs, ultimately aiming to improve the quality of life for women worldwide through precise and timely PMS management. The main ideas that do this research significantly are: PMS and PMDD is a widespread problem affecting many women.

Current treatments don't always work well enough. Artificial Intelligence technology can help create better, more personalized solutions. This research could help improve how we treat PMS and depression. The research goal is to help women have a better quality of life through better PMS management.

Research objective: The objectives of this research are threefold. The first objective is to study the context of female adolescents who experience premenstrual syndrome (PMS) problems and depression crises. The second objective is to develop a classification model for personalized premenstrual symptoms and depressive crisis risk tracking using data analytics and machine learning techniques. The third objective is to evaluate the performance of all developed models to select suitable models for future application.

II. LITERATURE REVIEWS

Effectively addressing premenstrual syndrome (PMS) and depressive crises necessitates the use of advanced machine learning (ML) and data analytics. Research highlights the influence of PMS on professional performance and underscores ML's capability for precise symptom prediction. While challenges such as data imbalance remain, ML algorithms significantly improve accuracy and scalability.

This study seeks to create a tailored classification model for monitoring PMS and depressive risks, enhancing health outcomes for women.

This research examined the impact of premenstrual syndrome (PMS) symptoms on clinical performance among female nurses in Iranian public hospitals. Given the high occupational stress experienced by nurses, the study highlighted the importance of understanding how PMS affects their ability to deliver quality patient care. Three hundred eighteen female nurses from three hospitals participated, with data collected using validated tools such as the PSST, PHQ-9, SPST-20, and WHOQOL-BREF-THAI. Statistical analyses identified a significant negative correlation between PMS symptoms and clinical performance, with PMS explaining 26.5% of the variance in performance scores. Although limited to one region, the findings stress the need to address PMS symptoms to improve nursing performance and healthcare quality [5]. In addition, the study recruited participants from 46 public hospitals across Shandong Province, including 22,924 dedicated nurses who completed a detailed questionnaire. It aimed to develop a predictive model to assess the risk of premenstrual syndrome (PMS) in nurses, a group with unique occupational stressors. Key risk factors for PMS were identified, including tea or coffee consumption, sleep quality, menstrual cycle regularity, dysmenorrhea severity, workplace bullying, and levels of anxiety and stress. A predictive model was created using stepwise multivariate logistic regression and the LASSO method, achieving strong performance (AUC of 0.765 for the training set and 0.769 for the test set). The model proved effective for early detection and intervention, improving nurses' well-being [6].

Furthermore, many women faced mental and emotional challenges before menstruation, especially with severe symptoms. This study developed a machine learning model to predict the severity of Premenstrual Syndrome (PMS) based on various factors. In addition, data from the Kelly Wallace questionnaire were collected from 317 women aged 13-19, covering PMS symptoms such as anxiety, appetite changes, depression, and fluid retention. To optimize the model, the Gradient Boost Regressor (GBR) model, combined with Grid Search Optimization (GSO), was used to determine the optimal parameters. The results showed that the GBR-GSO model demonstrated high accuracy, with 99.99% for PMS-A, 99.93% for PMS-C, 99.87% for PMS-D, 99.92% for PMS-H, and 99.97% for other symptoms. It indicated that the model showed promise in diagnosing and managing PMS, especially in adolescents, by improving early detection and treatment. Overall, the study's main strength was the application of machine learning for accurate predictions. At the same time, limitations included sample diversity and data imbalance across PMS categories [7].

Integrating DASS-21 with Machine Learning (ML) and Deep Learning (DL) methods has proven transformative in the context of advancements in mental health assessment. Utilizing 39,775 responses collected between 2017 and 2019, clustering techniques like K-Modes and Hierarchical Clustering identified psychological subgroups. Predictive models such as SVM, Naive Bayes, Decision Trees, and ensemble methods (e.g., Random Forest, Adaboost) achieved remarkable accuracy, with SVM and Adaboost reaching up to 96%. Deep Learning models like MLP and ANN offered further processing speed and efficiency advantages for large-scale assessments. While limited by sample diversity, the findings highlight that combining DASS-21 with ML and DL enhances accuracy and scalability, supporting personalized mental health interventions [8].

This study created a tool to assess the risk of self-reported anxiety, depression, and stress using Machine Learning (ML) and Explainable AI (XAI). The dataset from COVICAT included 9,291 participants from Northern Spain, focusing on mental health after COVID-19. Data was gathered through an online survey, and various ML models (Random Forest, SVM, and GBR) were used for prediction. Additionally, SHAP (SHapley Additive exPlanations) and UMAP (Uniform Manifold Approximation and Projection) techniques helped interpret the results. As a result, the model produced AUROC values of 0.77 for depression, 0.72 for anxiety, and 0.73 for stress. Furthermore, the main risk factors identified included poor overall health and low social support, crucial for identifying high-risk groups. While the study's strengths lie in its prediction accuracy and potential for targeted interventions, the lack of diversity in the sample limits its broader applicability [9].

Finally, this investigation focused on identifying effective methods for mental health detection using machine learning models (K-NN, MLP, SVM) in recovered COVID-19 patients. Drawing from data on depression, anxiety, and stress collected from 549 participants, the study employed feature selection techniques (Recursive Feature Elimination and Extra Trees) and hyper-parameter tuning to enhance model accuracy. The results demonstrated that SVM provided the highest accuracy ($\geq$ 0.984) for predicting mental health conditions, while Extra Trees outperformed RFE in selecting features for anxiety and stress datasets. These findings contribute to advancements in predictive modeling for early and precise mental health detection [10].

III. Materials and Methods

A. Research Scope

The scope of this research is designed to create a risk prediction model for menstruating women who are more likely to be affected by depression, stress, and anxiety. The population in this research was defined as undergraduate students of the University of Phayao, located in Mueang District, Phayao Province, Northern Thailand. The sample group that provided information totaled 282 undergraduate female students who consented to the data for the research, as detailed in Table I. In addition, the conceptual framework is shown in Figure 1.

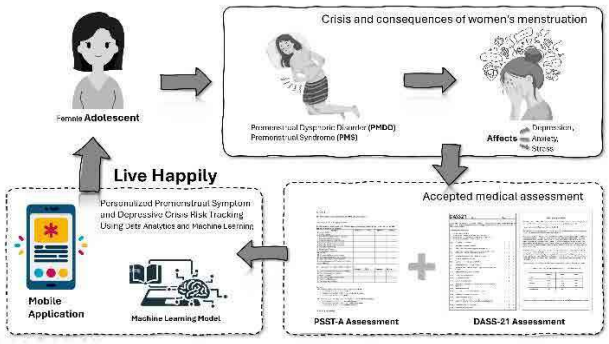


Fig. 1. Conceptual Frameworks

Figure 1 demonstrates the research framework in which researchers discovered a significant problem for adolescent

women: menstruation often has severe effects. It can result in varying degrees of depression, anxiety, and stress. This research, therefore, aims to apply data analytics and machine learning to develop a classification model for personalized premenstrual symptoms and depressive crisis risk tracking. The ultimate goal of the research is to construct a mobile application that can support and care for adolescent females so they can be delighted and live safe lives.

B. Data Collection

The data collected for this research has been obtained with the consent of the informants. It is kept confidential and cannot be traced back to them.

Data was collected through random sampling from 282 students across the first and second semesters of the 2024 academic year, using the PMDD and DASS-21 forms. Slovin's formula (T. Ryan, *Sample Size Determination and Power*, John Wiley & Sons, 2013.), based on a population of 3,026 (https://up.ac.th) and a margin of error of e=0.05, suggested a sample size of 353.33, but the actual sample size was limited to 282 due to resource constraints. This may impact the reliability of the findings on menstruation among female students. The data were obtained from an initial sample of 360 students, but some were excluded due to incomplete information. The final data were obtained from students across eighteen educational institutions at the University of Phayao, including the School of Agriculture and Natural Resources, the School of Allied Health Sciences, the School of Architecture and Fine Arts, the School of Business and Communication Arts, the School of Dentistry, the School of Education, the School of Energy and Environment, the School of Engineering, the School of Information and Communication Technology, the School of Law, the School of Liberal Arts, the School of Medical Sciences, the School of Medicine, the School of Nursing, the School of Pharmaceutical Sciences, the School of Political and Social Science, the School of Public Health, and the School of Science. The insights of the collected data are presented in Table I.

TABLE I. THE DATA COLLECTION

Issues	Statistics	
	Count	*Percentage*
Level Year		
• First Year	80	28.37%
• Second Year	61	21.63%
• Third Year	127	45.04%
• Fourth Year	14	4.96%
Age		
• Lower than 19 years old	35	12.41%
• 19 years old	56	19.86%
• 20 years old	88	31.21%
• 21 years old	91	32.27%
• 22 years old	10	3.55%
• More than 22 years old	2	0.70%
At what age did you have your first menstruation?		
• At the age of lower than 10 years	8	2.84%
• At the age of 10 years	14	4.96%
• At the age of 11 years	40	14.18%
• At the age of 12 years	107	37.94%
• At the age of 13 years	57	20.21%
• At the age of 14 years	31	10.99%
• At the age of 15 years	25	8.87%
Weight		
• Between 31 - 40 Kilograms	21	7.45%
• Between 41 - 50 Kilograms	84	29.79%
• Between 51 - 60 Kilograms	97	34.40%
• Between 61 - 70 Kilograms	45	15.96%
• Between 71 - 80 Kilograms	20	7.09%
• More than 80 Kilograms	15	5.32%
Hight		
• Between 141 - 150 Centimeters	22	7.80%
• Between 151 - 160 Centimeters	145	51.42%
• Between 161 - 170 Centimeters	104	36.88%
• More than 170 Centimeters	11	3.90%
Have you ever had sexual intercourse before?		
• Yes	131	46.45%
• No	151	53.55%
Total	282	100%

Table I shows the in-depth details of the informants. The total sample size was 282 representatives, divided into 80 first-year students (28.37%), 61 second-year students (21.63%), 127 third-year students (45.04%), and 14 fourth-year students (4.96%). Other summary information includes an average age of 19.97 years, an average weight of 56.61 kilograms, and an average height of 160.06 centimeters.

The researchers also found that, on average, respondents had their first menstruation at age 12.36 years and that 131 respondents (46.45 percent) had had sexual intercourse.

C. Research Tools

The research tools are divided into two tools. The first tool is the instrument used to collect data; the researcher uses a three-part questionnaire: the first part is to collect general information regarding the respondents, the second part is the menstruation status assessment form called "Premenstrual Symptoms Screening Tool for Adolescents (PSST-A)" [11], [12], as detailed assessment form in https://shorturl.at/cK9si, and the third part is the depression, stress, and anxiety assessment form called "Depression Anxiety Stress Scales (DASS-21)" [13], [14], as detailed assessment form in https://shorturl.at/hJJOs. Stress, anxiety, and depression are classified into five distinct levels: Normal, Mild, Moderate, Severe, and Extremely Severe [15]**.** These class labels provide a structured framework for assessing the severity of symptoms and tailoring interventions accordingly. The features of the data used in this study are of categorical type, all features are represented in Table II.

In developing a machine learning model for tracking premenstrual symptoms and depressive crises, the process starts by importing a reliable dataset with relevant features, ensuring data integrity through platforms like Google Colab. Non-essential columns are eliminated to simplify the model and reduce the risk of overfitting, enabling a focus on significant patterns for accurate predictions. String-based data is then transformed into numeric format, and categorical variables are encoded using methods such as label encoding. Labels like "Mild" or "Severe" are converted using the EncodeLabels function to improve computational efficiency and compatibility, as shown in Figure 2.

TABLE II. DASS-21 QUESTIONS

Questions	Features
I found it hard to wind down.	Stress
I was aware of dryness of my mouth.	Anxiety
I couldn't seem to experience any positive feeling at all.	Depression
I experienced breathing difficulty.	Anxiety
I found it difficult to work up the initiative to do things.	Depression
I tended to over-react to situations.	Stress
I experienced trembling.	Anxiety
I felt that I was using a lot of nervous energy.	Stress

Questions	Features
I was worried about situations in which I might panic and make a fool of myself.	Anxiety
I felt that I had nothing to look forward to.	Depression
I found myself getting agitated.	Stress
I found it difficult to relax.	Stress
I felt down-hearted and blue.	
I was intolerant of anything that kept me from getting	Stress
on with what I was doing.	Anxiety
I felt I was close to panic.	Depression
I was unable to become enthusiastic about anything.	Depression
I felt I wasn't worth much as a person.	Stress
I felt I was rather touchy.	Anxiety
I was aware of the action of my heart in the absence of physical exertion.	Anxiety
I felt scared without any good reason.	Depression

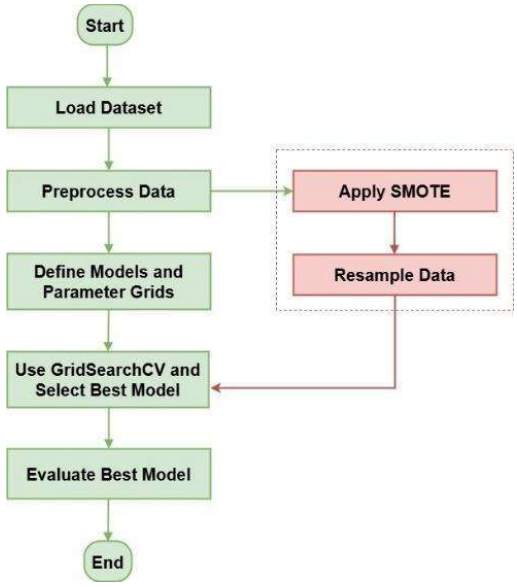


Fig. 2. Creating A Machine Learning Model

To address class imbalance, the SMOTE (Synthetic Minority Oversampling Technique) method is used to generate synthetic samples for the minority classes. This approach enhances model accuracy and robustness, especially when identifying high-risk cases [16].

Hyperparameter tuning is performed by defining a parameter grid for each model, optimizing performance with Grid Search or Random Search to determine the best combination of parameters [16].

The second tool is the tool used to develop the model. The researcher uses six machine learning techniques: Decision Trees (DT), K-nearest neighbors (K-NN), Logistic Regression (LR), Naïve Bayes (NB), Random Forests (RF), and Support Vector Machines (SVM). Details of each technique are explained in the modeling section.

D. Modeling

Researchers used the CRISP-DM technique [17], [18], [19], which consists of six steps, in the model development process. In the first step, the business understanding, researchers applied two accepted questionnaires PSST-A [20] and DASS-21 [21] to predict depression in adolescents who may be affected by irregular menstruation. In the second step, the data understanding, researchers found that the context of mass questionnaires was tedious and often yielded inaccurate information, so they applied a curated questionnaire to create an effective depression prediction model. In the third step, the data preparation, the researchers collected the data thoroughly, as summarized in Table I.

In the fourth step, modeling, the researchers used machine learning techniques to create a predictive model of risk and potential impact on adolescents during menstruation. The fifth step is the evaluation, in which the researchers use data partitioning techniques to test and summarize the assessment, as detailed in the model performance testing section. The final step is the deployment, where researchers plan to develop a mobile application prototype using data analytics and machine learning to personalized monitor premenstrual symptoms and the risk of depressive crises.

E. Model Performance Testing

To test the model's performance, a 10-fold cross-validation technique [22], [23] was used. The data is divided into ten equal parts, and the model is run for ten iterations. Each iteration extracts nine parts of the data to generate the model and uses the remaining one part to test it. The tools used to measure performance consist of five indicators: accuracy, precision, recall, f1-score, and time (seconds).

IV. Results

The research report is divided into two parts: the first part is a summary of the context of the collected data and the second part is a report of the development and performance analysis of the model.

A. Context of the Collected Data

The results of the contextual analysis of the data are reported in two ways: The first is the risk analysis of menstrual severity (PSST-A assessment), as shown in Table III, and the second is the analysis of depression, anxiety, and stress severity (DASS-21 evaluation), as shown in Table IV.

TABLE III. PSST-A Assessment Results

Severity	PSST-A Assessment Results	
	Yes	*No*
Premenstrual Dysphoric Disorder (PMDD)	29 (10.28%)	253 (89.72%)
Premenstrual Syndrome (PMS)	35 (12.41%)	247 (87.59%)

Table III illustrates the analysis results of the collected data and the Premenstrual Symptoms Screening Tool for Adolescents (PSST-A assessment). The number of students who were analyzed as having the premenstrual dysphoric disorder (PMDD) symptom (10.28%) was 29 samples (12.41%), and the premenstrual syndrome (PMS) was 35 samples (12.41%). Overall, the majority of the sample was not affected by these two symptoms.

TABLE IV. DASS-21 Evaluation Results

Issues	DASS-21 Assessment Results				
	Normal	*Mild*	*Moderate*	*Severe*	*Extremely Severe*
Depression	151 (53.55%)	55 (19.50%)	42 (14.89%)	19 (6.74%)	15 (5.32%)
Anxiety	110 (39.01%)	49 (17.38%)	48 (17.02%)	49 (17.38%)	26 (9.22%)
Stress	161 (57.09%)	42 (14.89%)	34 (12.06%)	11 (3.90%)	34 (12.06%)

Table IV shows the analysis results of the collected data and the Depression Anxiety Stress Scales (DASS-21) assessment. It was found that the sample group from which data were collected had symptoms of extremely severe

depression in 15 cases (5.32%), extremely severe anxiety disorder in 26 cases (9.22%), and extremely severe stress disorder in 34 cases (12.06%). Overall, it can be concluded that students of the University of Phayao have low levels of depression, anxiety, and stress.

B. Model Results and Model Performance

Model development and performance testing results were reported in three parts based on risk assessment for depression, anxiety, and stress severity (DASS-21 evaluation), as indicated in Tables V to VII.

TABLE V. EVALUATION OF THE DEPRESSION SEVERITY MODEL

Model	Evaluation of the Depression Severity Model				
	Accuracy	*Precision*	*Recall*	*F1-Score*	*Time (s.)*
RF	0.8874	0.8866	0.8874	0.8865	98.8590
Best Parameters: {'max_depth': 10, 'min_samples_split': 2, 'n_estimators': 50}					
LR	0.6690	0.6594	0.6689	0.6619	2.6361
Best Parameters: {'C': 1, 'solver': 'lbfgs'}					
SVM	0.9006	0.8994	0.9007	0.8994	22.7895
Best Parameters: {'C': 10, 'kernel': 'rbf'}					
DT	0.8025	0.7998	0.8026	0.7995	1.7152
Best Parameters: {'max_depth': None, 'min_samples_split': 2}					
K-NN	0.8780	0.8804	0.8781	0.8745	2.0236
Best Parameters: {'n_neighbors': 3, 'weights': 'distance'}					
NB	0.6027	0.5914	0.6026	0.5904	0.4682
Best Parameters: {}					

Table V shows the evaluation results of the depression severity model using six techniques for model construction and performance analysis. It was found that the support vector machine (SVM) technique had the highest accuracy value, with 90.06%, precision of 89.94%, recall of 90.07%, and f1-score of 89.94%.

TABLE VI. EVALUATION OF THE ANXIETY SEVERITY MODEL

Model	Evaluation of the Anxiety Severity Model				
	Accuracy	*Precision*	*Recall*	*F1-Score*	*Time (s.)*
RF	0.8018	0.8029	0.8018	0.8022	130.1979
Best Parameters: {'max_depth': None, 'min_samples_split': 2, 'n_estimators': 200}					
LR	0.4655	0.4615	0.4655	0.4604	5.2877
Best Parameters: {'C': 1, 'solver': 'liblinear'}					
SVM	0.7964	0.7981	0.7964	0.7968	40.0226
Best Parameters: {'C': 10, 'kernel': 'rbf'}					
DT	0.6273	0.6278	0.6273	0.6263	1.2792
Best Parameters: {'max_depth': None, 'min_samples_split': 2}					
K-NN	0.8109	0.8143	0.8109	0.8086	1.0536
Best Parameters: {'n_neighbors': 3, 'weights': 'distance'}					
NB	0.4273	0.4337	0.4273	0.3926	0.2983
Best Parameters: {}					

Table VI shows the evaluation results of the anxiety severity model using six techniques for model construction and performance analysis. It was found that the k-nearest neighbor (K-NN) technique had the highest accuracy value, with 81.09%, precision of 81.43%, recall of 81.09%, and f1-score of 80.86%.

TABLE VII. EVALUATION OF THE STRESS SEVERITY MODEL

Model	Evaluation of the Stress Severity Model				
	Accuracy	*Precision*	*Recall*	*F1-Score*	*Time (s.)*
RF	0.8980	0.8975	0.8981	0.8973	107.3233
Best Parameters: {'max_depth': 10, 'min_samples_split': 2, 'n_estimators': 200}					
LR	0.6236	0.6100	0.6236	0.6095	2.5072
Best Parameters: {'C': 0.01, 'solver': 'lbfgs'}					
SVM	0.9093	0.9087	0.9093	0.9086	30.4058
Best Parameters: {'C': 10, 'kernel': 'rbf'}					
DT	0.7988	0.8004	0.7988	0.7984	1.2964
Best Parameters: {'max_depth': 10, 'min_samples_split': 2}					
K-NN	0.9018	0.9080	0.9019	0.8996	1.1513
Best Parameters: {'n_neighbors': 3, 'weights': 'distance'}					
NB	0.5639	0.5576	0.5640	0.5485	0.3691
Best Parameters: {}					

Table VII shows the evaluation results of the stress severity model using six techniques for model construction and performance analysis. It was found that the support vector machine (SVM) technique had the highest accuracy value, with 90.93%, precision of 90.87%, recall of 90.93%, and f1-score of 90.86%.

V. DISCUSSION

The results show that the SVM model achieved the highest performance in terms of accuracy and other metrics for both depression and stress severity prediction, demonstrating its strength in managing complex data. In contrast, the K-NN model performed best for anxiety severity, suggesting it is better suited for this task. Meanwhile, NB and LR showed relatively weaker performance, indicating that they might not be as effective for handling the intricacies of these psychological conditions. Depression and stress data are more concentrated in the "Normal" category with fewer cases at severe levels, providing clearer class boundaries that SVM can efficiently separate. Anxiety data, however, is more evenly distributed across severity levels, forming localized clusters that align well with K-NN's proximity-based classification approach.

The use of SMOTE (Synthetic Minority Over-sampling Technique) and Grid Search played a crucial role in improving the model results, as they helped to balance class distribution and fine-tune hyperparameters for each model. These findings emphasize the importance of choosing the right model based on the dataset and the specific psychological condition being addressed. Additionally, the results point to the potential benefit of further refining model tuning and feature engineering, especially when simpler models fall short.

VI. CONCLUSION

This research has developed and applied AI for healthcare, with a classification model for personalized premenstrual symptoms and depressive crisis risk tracking constructed using data analytics and machine learning. The objectives of this research are threefold. The first objective is to study the context of female adolescents who experience premenstrual syndrome (PMS) problems and depression crises. The second objective is to develop a classification model for personalized premenstrual symptoms and depressive crisis risk tracking using data analytics and machine learning techniques. The third objective is to evaluate the performance of all developed models to select suitable models for future application. The research data were collected from 282 students at the University of Phayao during the first and second semesters of the academic year 2024. All students who provided data gave their consent for the use of their data in the research.

The research instruments were divided into two parts: a questionnaire for data collection and a machine learning technique for model development, which consisted of six techniques: decision trees, k-nearest neighbors, logistic regression, naïve bayes, random forests, and support vector machines. The model performance is tested using 10-fold cross-validation and evaluated using five metrics: accuracy, precision, recall, f1-score, and time.

The research results found that most students at the University of Phayao did not have problems with premenstrual syndrome (PMS) and depressive crisis. In addition, Tables V to VII also show the developed models classified by the characteristics of depression, anxiety, and

stress. It shows that the researchers successfully created efficient models with high accuracy in all techniques. Three models should be further developed in mobile applications. The first model was built using the Support Vector Machine (SVM) technique, which is suitable for predicting depression with an accuracy of 90.06%. The second model was developed using the K-Nearest Neighbor (K-NN) technique. This model is ideal for predicting anxiety, with an accuracy rate of 81.09%. The third model was developed using the Support Vector Machine (SVM) technique. This model is suitable for predicting stress with an accuracy rate of 90.93%. SVM is more effective for depression and stress data due to its ability to handle distinct patterns, while K-NN is better suited for the more evenly spread anxiety data, as it captures localized relationships. This emphasizes the need to select the appropriate algorithm based on the data's structure for the best results. Finally, the researchers hope this research will be of public benefit for other researchers to develop further.

ACKNOWLEDGMENTS

This research project was supported by the Thailand Science Research and Innovation Fund and the University of Phayao. It also received support from many advisors, academics, researchers, students, and staff. The authors thank everyone for their support and cooperation in completing this research. Moreover, thanks to Claude, ChatGPT, Sci Space, Gemini, and Perplexity Generative AI for their invaluable assistance in gathering information, reading, and providing.

REFERENCES

[1] S. Modzelewski, A. Oracz, X. Żukow, K. Iłendo, Z. Śledzikowka, and N. Waszkiewicz, "Premenstrual syndrome: new insights into etiology and review of treatment methods," *Front. Psychiatry*, vol. 15, Apr. 2024, doi: 10.3389/fpsyt.2024.1363875.

[2] M. Upadhyay, A. Mahishale, and A. Kari, "Prevalence of premenstrual syndrome in college going girls - A cross sectional study," *Clin. Epidemiol. Glob. Health*, vol. 20, Mar. 2023, doi: 10.1016/j.cegh.2023.101234.

[3] T. Ratchatasitthigul and C. Roomruangwong, "Prevalence and associated factors of premenstrual syndrome and premenstrual dysphoric disorder among physical therapy internship students in Bangkok and surrounding areas," *J. Med. Biosci.*, vol. 6, no. 2, Art. no. 2, Jun. 2024, doi: 10.14456/jmbs.2024.18.

[4] W. Abbaoui, S. Retal, B. El Bhiri, N. Kharmoum, and S. Ziti, "Towards revolutionizing precision healthcare: A systematic literature review of artificial intelligence methods in precision medicine," *Inform. Med. Unlocked*, vol. 46, p. 101475, Jan. 2024, doi: 10.1016/j.imu.2024.101475.

[5] P. Khazdoozi, S. Javanmardifard, and Z. Keshtkaran, "The predictive role of common symptoms of premenstrual syndrome in the clinical practice of nurses: a cross-sectional study," *BMC Nurs.*, vol. 23, no. 1, p. 603, Aug. 2024, doi: 10.1186/s12912-024-02280-6.

[6] L. Li, X. Lv, Y. Li, X. Zhang, M. Li, and Y. Cao, "Development and validation of risk prediction model for premenstrual syndrome in nurses: results from the nurses-based the TARGET cohort study," *Front. Public Health*, vol. 11, Oct. 2023, doi: 10.3389/fpubh.2023.1203280.

[7] R. Joshi, P. Mathur, A. K. Gupta, S. Singh, V. Paliwal, and S. Nayar, "Mathematical modeling of intelligent system for predicting effectiveness of premenstrual syndrome," *J. Interdiscip. Math.*, vol. 26, no. 3, pp. 551–562, 2023, doi: 10.47974/JIM-1681.

[8] J. Ghorpade-Aher, A. Memon, S. Chugh, A. Chebolu, P. Chaudhari, and J. Chavan, "DASS-21 Based Psychometric Prediction Using Advanced Machine Learning Techniques," *J. Adv. Inf. Technol.*, vol. 14, no. 3, pp. 571–580, 2023, doi: 10.12720/jait.14.3.571-580.

[9] G. V. Benito *et al.*, "Machine learning for anxiety and depression profiling and risk assessment in the aftermath of an emergency," *Artif. Intell. Med.*, vol. 157, p. 102991, Nov. 2024, doi: 10.1016/j.artmed.2024.102991.

[10] T. A. Tuan, L. T. T. Trang, T. D. An, N. H. Nghia, and D. T. T. Loan, "An Analysis of Machine Learning for Detecting Depression, Anxiety, and Stress of Recovered COVID-19 Patients," *J. Hum. Earth Future*, vol. 5, no. 1, Art. no. 1, Mar. 2024, doi: 10.28991/HEF-2024-05-01-01.

[11] M. Steiner, M. Peer, E. Palova, E. W. Freeman, M. Macdougall, and C. N. Soares, "The premenstrual symptoms screening tool revised for adolescents (PSST-A): prevalence of severe PMS and premenstrual dysphoric disorder in adolescents.," *Arch. Womens Ment. Health*, vol. 14, no. 1, pp. 77–82, Feb. 2011.

[12] M. Steiner, M. Macdougall, and E. Brown, "The premenstrual symptoms screening tool (PSST) for clinicians," *Arch. Women's Ment. Health*, vol. 6, no. 3, pp. 203–209, Aug. 2003, doi: 10.1007/s00737-003-0018-4.

[13] F. R. Ronk, J. R. Korman, G. R. Hooke, and A. C. Page, "Assessing clinical significance of treatment outcomes using the DASS-21," *Psychol. Assess.*, vol. 25, no. 4, pp. 1103–1110, 2013, doi: 10.1037/a0033100.

[14] A. Osman, J. L. Wong, C. L. Bagge, S. Freedenthal, P. M. Gutierrez, and G. Lozano, "The Depression Anxiety Stress Scales—21 (DASS-21): Further Examination of Dimensions, Scale Reliability, and Correlates," *J. Clin. Psychol.*, vol. 68, no. 12, pp. 1322–1338, 2012, doi: 10.1002/jclp.21908.

[15] S. H. Lovibond and P. F. Lovibond, "Depression Anxiety Stress Scales." Sep. 12, 2011. doi: 10.1037/t01004-000.

[16] E. Sara, C. Laila, and I. Ali, "The Impact of SMOTE and Grid Search on Maintainability Prediction Models," in *2019 IEEE/ACS 16th International Conference on Computer Systems and Applications (AICCSA)*, Nov. 2019, pp. 1–8. doi: 10.1109/AICCSA47632.2019.9035342.

[17] S. Tripathi, D. Muhr, M. Brunner, H. Jodlbauer, M. Dehmer, and F. Emmert-Streib, "Ensuring the Robustness and Reliability of Data-Driven Knowledge Discovery Models in Production and Manufacturing," *Front. Artif. Intell.*, vol. 4, Jun. 2021, doi: 10.3389/frai.2021.576892.

[18] P. Nuankaew and P. Temdee, "Matching of Compatible Different Attributes for Compatibility of Members and Groups," *Int. J. Mob. Learn. Organ.*, vol. 13, no. 1, pp. 4–29, Jan. 2019, doi: 10.1504/IJMLO.2019.096469.

[19] P. Nuankaew and P. Temdee, "Online Mentoring Model by Using Compatible Different Attributes," *Wirel. Pers. Commun.*, vol. 85, no. 2, pp. 565–584, Nov. 2015, doi: 10.1007/s11277-015-2755-x.

[20] M. Steiner, M. Peer, E. Palova, E. W. Freeman, M. Macdougall, and C. N. Soares, "The premenstrual symptoms screening tool revised for adolescents (PSST-A): prevalence of severe PMS and premenstrual dysphoric disorder in adolescents," *Arch. Womens Ment. Health*, vol. 14, no. 1, pp. 77–81, Feb. 2011, doi: 10.1007/s00737-010-0202-2.

[21] C.-H. Cao, X.-L. Liao, X.-Y. Jiang, X.-D. Li, I.-H. Chen, and C.-Y. Lin, "Psychometric evaluation of the depression, anxiety, and stress scale-21 (DASS-21) among Chinese primary and middle school teachers," *BMC Psychol.*, vol. 11, no. 1, p. 209, Jul. 2023, doi: 10.1186/s40359-023-01242-y.

[22] P. Nuankaew, P. Nasa-Ngium, and W. S. Nuankaew, "Self-Regulated Learning Styles in Hybrid Learning Using Educational Data Mining Analysis," in *2022 26th International Computer Science and Engineering Conference (ICSEC)*, Dec. 2022, pp. 208–212. doi: 10.1109/ICSEC56337.2022.10049322.

[23] S. Bussaman, P. Nasa-Ngium, T. Sararat, W. S. Nuankaew, and P. Nuankaew, "Influence Analytics Model of the General Education Courses Toward the Academic Achievement of Rajabhat University Students Using Data Mining Techniques," in *Information Systems for Intelligent Systems*, C. So In, N. D. Londhe, N. Bhatt, and M. Kitsing, Eds., in Smart Innovation, Systems and Technologies. Singapore: Springer Nature, Feb. 2024, pp. 117–129. doi: 10.1007/978-981-99-8612-5_10.

Gmail Pratya Nuankaew <nuankaew.p@gmail.com>

ECTI DAMT and NCON 2025 submission 4164

1 message

ECTI DAMT and NCON 2025 <ectidamtandncon2025@easychair.org> Thu, Dec 12, 2024 at 3:02 PM
To: Pratya Nuankaew <pratya.nu@up.ac.th>

Dear authors,

We received your submission to ECTI DAMT and NCON 2025 (2025 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunication Engineering):

Authors : Pratya Nuankaew, Jidapa Sorat, Jindaporn Intajak, Jirapron Inta and Wongpanya Nuankaew
Title : AI for Healthcare: A Classification Model for Personalized Premenstrual Symptoms and Depressive Crisis Risk Tracking Using Data Analytics and Machine Learning
Number : 4164
Track : Knowledge and Innovative Management

The submission was uploaded by Pratya Nuankaew <pratya.nu@up.ac.th>.
You can access it via the ECTI DAMT and NCON 2025 EasyChair Web page

https://easychair.org/conferences/?conf=ectidamtandncon2025

Thank you for submitting to ECTI DAMT and NCON 2025.

Best regards,
EasyChair for ECTI DAMT and NCON 2025.

Gmail Pratya Nuankaew <nuankaew.p@gmail.com>

Notification of Acceptance of Your Paper ID 4164 for the ECTI DAMT and NCON 2025 International Conference

1 message

ECTI DAMT and NCON 2025 <ectidamtandncon2025@easychair.org> Mon, Dec 30, 2024 at 11:36 AM

To: Pratya Nuankaew <pratya.nu@up.ac.th>

Dear Pratya Nuankaew,

The decision for the paper ID 4164 entitled AI for Healthcare: A Classification Model for Personalized Premenstrual Symptoms and Depressive Crisis Risk Tracking Using Data Analytics and Machine Learning is "Accept with minor revision" and review results are shown below. Please revise and update your camera-ready to the EasyChair system by 15 January 2025. The format of the conference both on Microsoft Word (A4) and Latex is available at https://www.icdamt.org/submission/.

For further preparation, please carefully check the conference's important date at https://www.icdamt.org/call-for-paper/. In addition, it is a condition of paper acceptance that you or the nominated presenting co-author must register for the conference by the registration deadline of 13 January 2025 otherwise, the papers will be removed from the program.

The instructions and registration site are available at https://www.openbadgepay.com/.

For more information, please visit www.icdamt.org. If you have any inquiries, please feel free to contact us.

Best Regards,
ECTI DAMT and NCON 2025 Committee

SUBMISSION: 4164
TITLE: AI for Healthcare: A Classification Model for Personalized Premenstrual Symptoms and Depressive Crisis Risk Tracking Using Data Analytics and Machine Learning

---------- REVIEW 1 ----------
SUBMISSION: 4164
TITLE: AI for Healthcare: A Classification Model for Personalized Premenstrual Symptoms and Depressive Crisis Risk Tracking Using Data Analytics and Machine Learning
AUTHORS: Pratya Nuankaew, Jidapa Sorat, Jindaporn Intajak, Jirapron Inta and Wongpanya Nuankaew

---------- Overall evaluation ----------
SCORE: 0 (major revision)
----- TEXT:
The authors developed the classified model for stress, anxiety, and depression effects from survey data using several ML techniques.

The submission requires revision to align with established academic standards.

Data Collection:
This section should provide a detailed explanation of how the data was collected, including the sampling design, tools used (e.g., survey forms), and the type of data gathered (continuous or categorical).

- Please clarify the sampling design implemented in this study.
- What is the total population size (N)?
- Based on the "random sampling" method referenced by the authors, how many female students were registered during the 2024 academic year?
- How was the sample size (n) determined from the population?

The authors should relocate Table I to the results section, possibly under the subsection "Context of the Collected Data."

Additionally, the descriptive statistics of the sample data should be presented, specifically highlighting the representation across the 18 educational institutions at the university. Are all faculties within the university included?

Providing this information would substantiate the predictive model's applicability for future analysis involving female students at the University of Phayao.

- Were some data points in Table I initially continuous and subsequently converted into categorical data?
- Were all features utilized in the ML model exclusively as categorical variables? If the data was regrouped, the authors should cite relevant references to justify this categorization.

Research Tools:
Consider renaming this section to "Data Preprocessing" or "Data Manipulation" to reflect its purpose better.

This section should comprehensively outline the dataset preparation steps for ML model training. The authors should include:
- A description of the raw data format.
- Data cleaning and transformation processes (if applicable).
- Any feature engineering conducted.

Furthermore, the final dataset should be described in detail, including:
- The final sample size.
- The outcome variables (labels) and their data types.
- The predictor variables (features) and their data types.

SMOTE should be properly cited, as should any techniques such as Grid Search or Random Search. Note that the last paragraph in this section is redundant and may be omitted.

Modeling:
The CRISP-DM methodology should be cited and introduced at the beginning of the "Materials and Methods" section, possibly after the research scope. As this methodology is central to the study, a brief explanation of its six steps would provide valuable context.

If feasible, include a short description or key characteristics of the ML models utilized.

Model Results and Model Performance:
Tables IV to VI currently share the same title, "Evaluation of the Anxiety Severity Model," which should be corrected for clarity.

Discussion:
This section repeats content from the results section and should instead provide a more comprehensive analysis of the findings. For example:

Why does the SVM model perform best in classifying depression and stress severity, while K-NN outperforms in anxiety severity classification?

Is this outcome consistent with the distribution of samples across the five severity levels presented in Table III?

Additional aspects that merit discussion include the impact of sample size and resampling techniques (e.g., oversampling and undersampling). The use of SMOTE, which creates synthetic samples by interpolating between points and their K-nearest neighbors, may increase the risk of overfitting and should be evaluated in this context.

Questions:

- Given that SMOTE addresses class imbalance, why are additional metrics beyond accuracy necessary, particularly when there is no inherent bias toward majority or minority classes?

Suggestions:

- Use abbreviations for healthcare terms and ML model names after their first mention in the text.
- The authors should perform preliminary data analysis or exploratory data analysis (EDA) to derive insights and obtain critical findings that could support or lead to the study's results and discussion.

---------- REVIEW 2 ----------
SUBMISSION: 4164
TITLE: AI for Healthcare: A Classification Model for Personalized Premenstrual Symptoms and Depressive Crisis Risk Tracking Using Data Analytics and Machine Learning
AUTHORS: Pratya Nuankaew, Jidapa Sorat, Jindaporn Intajak, Jirapron Inta and Wongpanya Nuankaew

---------- Overall evaluation ----------

SCORE: 1 (minor revision)
----- TEXT:
The research work is interesting and could provide a strong impact. However, the classification is unclear. There should be details regarding the features (how many and what are they?) and the labels (how to get the labels?) used in classification.

---------- REVIEW 3 ----------
SUBMISSION: 4164
TITLE: AI for Healthcare: A Classification Model for Personalized Premenstrual Symptoms and Depressive Crisis Risk Tracking Using Data Analytics and Machine Learning
AUTHORS: Pratya Nuankaew, Jidapa Sorat, Jindaporn Intajak, Jirapron Inta and Wongpanya Nuankaew

---------- Overall evaluation ----------
SCORE: 2 (accept)
----- TEXT:
Contributions
The paper proposes a machine learning-based classification model aimed at tracking and predicting premenstrual symptoms and depressive crises. Utilizing six machine learning techniques (Decision Trees, K-Nearest Neighbors, Logistic Regression, Naïve Bayes, Random Forests, and Support Vector Machines), the study evaluates their performance using a dataset from 282 female students. It aims to address the critical need for personalized healthcare solutions in managing Premenstrual Syndrome (PMS) and related mental health issues. The research contributes to healthcare by identifying effective predictive models and providing actionable insights for future applications, such as mobile health tools.

Strengths
Focus on Personalized Healthcare: The study addresses a relevant issue in women's health, emphasizing the need for personalized risk assessment and early intervention for PMS and depressive crises.
Comprehensive Methodology: The use of multiple machine learning techniques with detailed evaluation metrics (accuracy, precision, recall, F1-score, and time) strengthens the reliability of the findings.
Real-World Application: The proposed models are well-suited for deployment in mobile health applications, providing a pathway for practical use in healthcare.
Effective Handling of Imbalanced Data: The application of SMOTE (Synthetic Minority Oversampling Technique) to address class imbalances enhances the robustness of the models.
Weaknesses
Limited Dataset: The dataset is relatively small (282 participants) and geographically restricted, which may limit the generalizability of the findings.
Model Scope: While the study explores six models, it omits other advanced techniques such as ensemble methods or deep learning, which could potentially improve performance.
Lack of External Validation: The models are not validated using external datasets, raising questions about their robustness in diverse contexts.
Insufficient Analysis of Psychological Context: The study does not delve deeply into the psychological and behavioral factors influencing PMS and depression, which could add depth to the findings.
Suggestions for Improvement
Expand Dataset: Include a larger and more diverse sample to improve the generalizability of the findings and ensure the robustness of the models across different populations.
Explore Advanced Techniques: Investigate additional machine learning techniques, such as ensemble methods or deep learning models, to potentially enhance prediction accuracy.
Conduct External Validation: Test the developed models on external datasets to evaluate their robustness and applicability in different settings.
Integrate Behavioral Insights: Incorporate a deeper analysis of psychological and behavioral factors influencing PMS and depressive crises to provide a more comprehensive perspective.
Assess Model Scalability: Include a discussion on the scalability and computational feasibility of the models in real-world applications, especially in resource-constrained settings.

Sign Language Detection Mobile Application for Thai Patients Using Medical Image Processing to Support Medical Consultations

Wongpanya S. Nuankaew[1[0000-0003-3805-9529]], Natthida Nuttaphum[2], Thapanapong Sararat[3[0000-0001-8504-4026]], Phanombongkot Banyaem[4], and Pratya Nuankaew[5[0000-0002-3297-4198]]

[1, 2, 3, 3, 4, 5] School of Information and Communication Technology, University of Phayao, Phayao 56000, Thailand
pratya.nu@up.ac.th

Abstract. Effective communication between patients and healthcare providers is pivotal for accurate diagnosis, appropriate treatment, and optimal patient care. However, individuals with hearing impairments often encounter significant communication barriers when accessing medical services, hindering their ability to convey health-related information effectively. A sign language detection mobile application tailored for Thai patients has been developed to address this challenge. The primary objectives of this initiative are to construct the most effective sign language detection model for Thai patients, to create a mobile application for Thai patients with a sign language detection model to support medical consultations, and to study satisfaction with the sign language detection mobile application. Leveraging cutting-edge technologies such as the Flutter and Dart framework, TensorFlow Lite, and Teachable Machine for model training, the application aims to bridge the communication gap between deaf and hard-of-hearing patients and healthcare professionals. The dataset utilized in this experiment comprised video recordings of 60 sign language vocabulary and phrases related to health and wellness.

Keywords: Clinical Informatics, Medical Application, Medical Image Processing, Sign Language Detection, Support Medical Consultations.

1 Introduction

Medical personnel whose essential duties are to provide knowledge, consider illnesses, and treat disease, which requires assessment of understanding, knowledge, and primary treatment skills of patients to ensure that patients can care for themselves. Patients, therefore, need to communicate with their doctors or medical personnel regarding their illnesses. However, the treatment or disease information is complex and may not be understood by some patients. Thus, receiving appropriate treatment that matches the disease heavily depends on the patient's communication, behavior, and knowledge. For example, patients may have language barriers during

transmission, and doctors may have difficulty communicating due to factors from the patient themselves, such as hearing loss or deafness. As a result, communication between medical staff and patients has more obstacles and errors.

Statistics in 2023 Thailand has 405,920 people with hearing disabilities, accounting for 18.62 percent of all people with disabilities. For people who are hard of hearing, the most important thing is daily life and communication with others. Especially when receiving health services, such as communication between doctors and nurses at the hospital, because if the transmission is not understood, it will lead to incorrect treatment. The most popular solution is the annotation method. This method is effective for certain patients, including those who are hard of hearing but can communicate understandably through reading and writing or patients who can read the lips of their conversation partners. These patients often have no problems receiving health services at hospitals [1, 2]. This group of patients will prepare before seeing the doctor by writing a report of their symptoms or questions they want to know to save communication time and prevent misunderstandings. However, using this method for deaf people with limited reading and writing skills may result in the patient not always understanding what the doctor wants to communicate through writing. Researchers have therefore studied and developed an application for deaf and hard of hearing people to use in health communication with doctors through sign language that patients are familiar with, using a camera to capture the sign language gestures of the patient and translate sign language into text, and convert text to sign language for medical communication.

The purpose of this research consists of three main parts. The first objective is to construct a mobile application for Thai patients with a sign language detection to support medical consultations. The second objective is to develop a consultation chatbot for Thai patients to inquire about basic medical treatment information. Finally, the third objective is to study the satisfaction with the sign language detection mobile application. The scope and framework of the research study are two components. The first component is a communication channel for Thai patients consisting of converting sign language into text to communicate with healthcare professionals. The second component is a communication channel for chatbots and medical personnel to provide information and inquire about the primary symptoms of Thai patients, as indicated in the research framework in Figure 1.

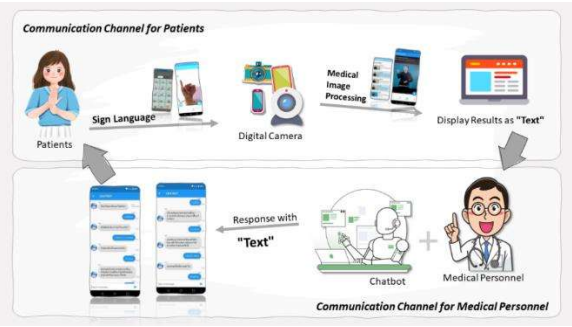


Fig. 1. The scope and framework of the research

Overview of system and application functionality: this application is designed to support use for the deaf and hard of hearing who want to communicate with medical personnel to inquire about the patient's essential condition. The application concept begins with capturing the user's sign language gestures through a webcam or mobile device. The system then receives sign language images, records new images, and compares them with existing images in the database to convert them into text for communication with medical personnel. Feedback communication for patients consists of two states: answering questions by the chatbot and medical personnel. The nature of communication is through online conversations. One last thing, benefits expected from the application include facilitating communication problems between medical personnel who do not have sign language expertise and the deaf and hard of hearing, being a medical innovation that promotes and encourages medical services for patients with communication problems.

2 Related works

This stage reviews recent literature on sign language systems to enable an understanding of designed and developed sign language systems and the prediction mechanisms of deep learning models. The details are as follows: Dony Novaliendry et al. (2023) have designed and developed a specialized sign language learning application to boost students' sign language proficiency and encourage greater participation [3]. The system utilizes the Media Development Life Cycle (MDLC) methodology and the Flutter framework. It targets first-grade students who are deaf.

Anusorn Chaikaew et al. (2018) developed a mobile application to facilitate Thai sign language communication and serve the sizeable deaf population [4]. The mobile application provides ease of sign language learning and is used as an animated dictionary with vocabulary classifications that are easy to search and learn. It focuses on everyday terms and phrases to expand communication access and skills for people who are deaf or hard of hearing. In their 2021 study, Anusorn Chaikaew et al. demonstrate how developing human-computer interaction technology could enhance the quality of life for people with disabilities [5]. Their research aims to develop a Thai sign language recognition application to create a real-time sign language translation mobile app. Jakkrapan Sudthipadh and Suree Pumrin (2022) have developed an innovative system designed to interpret Thai Sign Language, facilitating communication both within the hearing-impaired community and between individuals with and without hearing impairments. This cutting-edge technology leverages camera or webcam input to detect and capture hand gestures.

Kun Xia et al. (2022) applied a deep learning model called MobileNet-YOLOv3 to recognize Chinese sign language gestures used by deaf-mute patients when communicating medical needs to doctors. Their model was trained and test on a dataset of 4000 sign language images. When tested, it achieved a robust accuracy rate of 90.77% in correctly classifying and interpreting the signs. This technology aims to act as an automated interpreter, bridging more seamless communication between deaf patients and medical professionals [1].

Candy Obdulia Sosa-Jiménez et al. (2022) proposed a prototype system for recognizing and synthesizing Mexican sign language in the context of primary healthcare services. Using 82 distinct sign gestures, their model attained mean accuracy and F1 scores of 99% and 88% - also representing an effective level of sign language translation capability [2]. K. J. Kevin Feng and David W. Mcdonald (2023) [9] talked about the for training Machine learning model using Teachable machine. It has been used in various applications. In this research, the researchers have developed a machine learning model utilizing Teachable Machine and TensorFlow lite, a tool that facilitates the creation of models for data learning. This tool is renowned for its speed, efficiency, and universal accessibility. It enables training computers to recognize and memorize various inputs, including images, sounds, and motion gestures. The resulting models can then be integrated into web applications, applications, and other platforms, offering a wide range of practical applications.

To conclude the related work, deep learning has shown promising accuracy in modeling and recognizing sign language from images and video input. In addition, mobile applications are making sign language more accessible and beneficial. These studies demonstrate the effective model and implementation of sign language systems based on development application systems. Undoubtedly, technology enhances communication and quality of life for the deaf community.

3 Research Methodology

The scope of the research has been designed and implemented by the research and development (R&D) process, which has defined the population and sample groups, has a system design and development process with a software development life cycle (SDLC), and has tested the developed system to study satisfaction and efficiency of the application with expertise and Thai patients.

3.1 Population and Sample

The researcher determined the research population and the sample as two target groups. The first target group is the general patients and the patients who are deaf or hard of hearing in Mueang District, Phayao Province, Thailand, who received medical services at the University of Phayao Hospital, Mueang District, Phayao Province, Thailand, between 2022 and 2023, totaling 30 people. The second target group is three medical personnel working at the University of Phayao Hospital with at least five years of experience working in a hospital and expertise in communicating with deaf patients. Both sample groups were assigned to evaluate their satisfaction with the sign language detection mobile application in studying their satisfaction and attitude towards using medical technology and innovation to support work efficiency at the University of Phayao Hospital.

3.2 Research Instruments

The research instruments are composed of two tools. The first tool is a sign language detection mobile application. The second tool is a questionnaire on attitudes and satisfaction with using a sign language detection mobile application, which ICO according to the effective questionnaire development process.

3.3 Construction of Research Instruments

Mobile application development for sign language detection.

Data analysis.

Data analysis is the analysis of words in sign language, considering basic communication sentences, asking about patient symptoms, medical diseases, and specific terms related to patient illnesses. To collect data aligned with the research objectives, the words contained sixty healthcare sign language vocabulary and phrases, as shown in Table 1.

Data collection.

The collection data used in this experiment was stored in video format, extension MP4, and images from shot-to-shot photography from video, extension JPG, in the RGB system, with the cropped image size being 512x512 pixels. The researchers used 60 healthcare sign language gestures published on the Thai Telecommunication Relay Service (TTRS) website: https://www.ttrs.or.th and the Office of Educational Technology, Sukhothai Thammathirat Open University (STOU) website: https://oet.stou.ac.th/edtech, as detailed in Table 1.

Table 1. Healthcare sign languages in research

60 Healthcare sign languages in research					
Abortion	AIDS	Air allergies	Allergic reactions	Allergy	Anorexia
Anus	Arm pain	Astigmatism	Bladder	Blisters	Bloating
Blurred vision	Blurry vision	Can't eat.	Can't sleep.	Can't swallow.	Constipation
Cough	Cramp	Diarrhea	Dizziness	Drug allergy	Epilepsy
Eyes pain	Feeling feverish	Flu	Food allergy	Gradually invisible	Headache
Heartbeat	High blood pressure	High fever	Hurt	Itchy	Jaw pain
Liquid medicine	Nausea	Neuralgia	Panting	Pinching	Pregnant
Pulse	Red Eyes	Sick	Sneezing	Snot	Stinging nose
Stomach	Stomach pain	Stressed	Swollen	Tightness in the chest	Tired
Tired easily	Tooth filling	Trembling	Vomiting	Watery eyes	Worry

System design analysis.

The researchers analyzed the system design in two parts. The first part is an analysis of the traditional system. The researchers found the communication system with deaf and hard-of-hearing patients was complicated. Doctors and nurses must use assistants to communicate with patients. Moreover, some patients cannot read and write and do not have any relatives, which is a significant obstacle to the work of medical personnel. Therefore, health organization personnel need technology to support and solve such problems.

The second part is analyzing the new system to support and solve the discovered problems. The new scheme needed to solve communication barriers between hard-of-hearing patients and medical personnel, and the researchers designed the outline as shown in Figure 1. The main components of the designed new system comprise two parts: a communication system for patients that translates sign language into text to communicate with medical personnel and a communication system for medical personnel with two subsystems: feedback communication by Chatbot and doctors and nurses. The new system has been designed and developed following the software development life cycle (SDLC) principles [10], which are presented in the following sections.

Questionnaire development.

The research questionnaire is divided into two parts. The first is used to study the acceptance and satisfaction of application users. The second analyzes the test results of comparing sign language gestures with three experts, as shown in Tables 3 and 4, respectively.

3.4 Research Interpretation

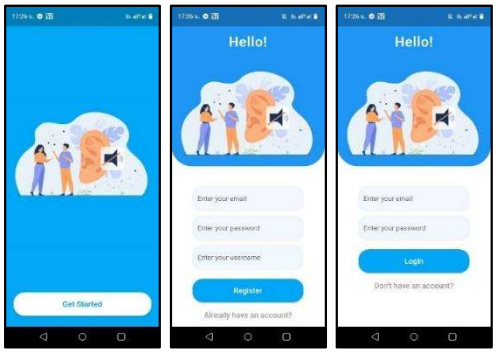
The interpretation of the research was tested on the user satisfaction assessment. The questionnaire used by Google Forms to complete the questionnaire and explore the results of acceptance and satisfaction. The researchers restricted the level of satisfaction with the application as follows: The first level is satisfaction and acceptance at the highest level, with a score equal to 5. The second level is satisfaction and acceptance at the high level, with a score equal to 4. The third level is satisfaction and acceptance at the accepted level, with a score equal to 3. The fourth level is satisfaction and acceptance at an unacceptable level, with a score equal to 2. The fifth level is satisfaction and acceptance at an extremely unacceptable level, with a score equal to 1.

After collecting the acceptance and satisfaction levels from the questionnaire, the values are processed to demonstrate the average score, and the results are interpreted according to the class interval, as specified in five classes as follows: The first-class average score is 4.21-5.00, acceptance at the highest level. The second-class average score is 3.41-4.20, with support at a high level. The third-class average score is 2.61-3.40, with approval at the accepted level. The fourth-class average score is 1.81-2.60, with consent at the accepted level. The fifth-class average score is 1.00-1.80, with permission at the accepted level.

4 Research Results

4.1 Sign language detection mobile application

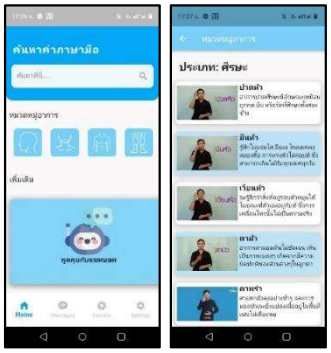
The application development results completed following all activity diagrams according to the established research framework are illustrated in Figures 2 to 5.



(a) (b) (c)

Fig. 2. Sign language detection mobile application

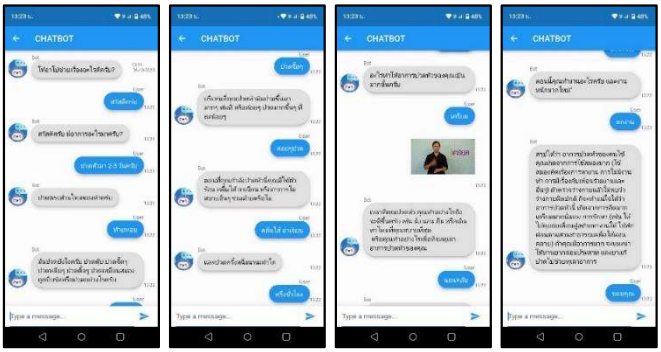
Figure 2(a) shows entering the application, Figure 2(b) shows registration, and Figure 2(c) shows the application login. The researchers designed the application page to be user-friendly and minimally complex.



(a) (b)

Fig. 3. Sign language detection mobile application

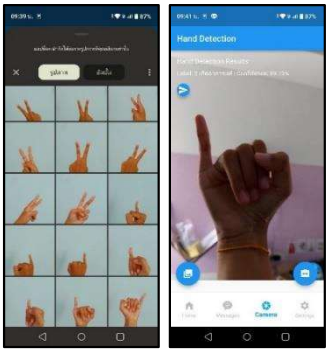
After entering the application, the researchers designed four main sections: the function of searching in the sign language vocabulary and phrases, symptom categories, talking with the chatbot, and related menus such as the home, messages, camera, and settings. Figure 3(a) shows the application page elements, and Figure 3(b) shows examples of sign language with descriptions.



(a) (b) (c) (d)

Fig. 4. Sign language detection mobile application

Figures 4(a) to 4(d) show an example of a conversation between a patient and a chatbot.



(a) (b)

Fig. 5. Sign language detection mobile application

Figures 5(a) and 5(b) illustrate an example of using a camera to capture patient images and translate them into text for communication in the system.

After developing the application according to the designed components, the researchers tested and evaluated its effectiveness. The test results and satisfaction studies will be presented in the next section.

4.2 Acceptance and satisfaction levels

The analysis results of the acceptance and satisfaction levels with the sign language detection mobile application for Thai patients are presented in Table 3, with 30 respondents, as detailed in Table 2.

Table 2. Respondent details

Details	Respondents	Percentages
Gender		
• Male	9	30.00%
• Female	21	70.00%
Age		
• 20 – 30 years old	22	73.33%
• 31 – 40 years old	5	16.67%
• 41 – 50 years old	2	6.67%
• More than 50 years old	1	3.33%
Occupation		
• Employee	11	36.67%
• General employee	8	26.33%
• Government officer	10	33.33%
• Student	1	3.33%
Total	*30*	*100%*

Table 2 presents respondents' satisfaction and acceptance of the application, classified by gender, age, and occupation. The total number of respondents is 30 samples. There were 21 females, representing 70.00 %, while there were nine males, representing 30.00 percent, with the most respondents being in the age range of 20-30 years old, 22 instances, representing 73.33 %. In addition, most of the respondents' occupations were employees, a total of 11 samples, accounting for 36.67 %.

Table 3. Acceptance and satisfaction level toward the mobile application

Stage	Mean	S.D.	Interpretation
Application performance / Usefulness			
• The design of the application has reasonable design elements that fulfill its purpose.	4.45	0.66	Highest accepted
• The sign language search function is quick, accurate, and straightforward.	4.36	0.71	Highest accepted
• Mobile camera capture can accurately grasp the user's sign language gestures and communication preferences.	4.32	0.76	Highest accepted
• The chatbot's functionality can answer questions quickly and directly to the communication point between the patient and the system.	4.59	0.58	Highest accepted
Average	*4.43*	*0.68*	Highest accepted

Table 3 indicates the results of the analysis of the acceptance level from the sign language detection mobile application testing. It was found that respondents assigned the highest level of acceptance of the application in all areas assessed. It has an overall average of 4.43 and a standard deviation of 0.68. It can, therefore, be concluded that the developed application is practical.

5 Research discussion

The results of the study showed that deaf and hard-of-hearing people were able to communicate with medical personnel with the highest level of satisfaction and acceptance. Medical personnel can understand sign language gestures by sharing through sign language detection mobile applications for counseling. In addition, the expert testing results with sign language accuracy values at 1,253 times, equal to 69.61 percent, show that the developed application can communicate with medical personnel. Moreover, it can also be applied to share with other people. Ordinary people interested in sign language can also use this application to study or learn sign language vocabulary and phrases related to health as an alternative way to learn sign language.

From the summary of research results, the researchers have issues that need to be discussed, including 1) user satisfaction and acceptance of the application and 2) the comparative analysis of sign language gestures with the experts. A study of satisfaction and acceptance of the application found that it has important implications for communication problems among medical personnel and people who are deaf or hard of hearing. The analysis results in Table 3 show that the application responds to the need for technology to solve communication problems in healthcare organizations. It has an overall acceptance of 4.43 and a standard deviation of 0.68. It can, therefore, be concluded that the developed application is practical.

The second point is the results of the comparative analysis of sign language gestures with the application classified by sign language vocabulary and phrases found that the testing had an accuracy of 1,253 times, equal to 69.61 percent, and the incorrect predictions were 547 times, equivalent to 30.39 percent, as detailed in Table 4. The error was caused by the images used to train the model at a poor resolution and insufficient clarity to enable high-performance image processing. It, therefore, causes and results in errors in vocabulary and phrase prediction. However, the researcher has tried to study and learn more from other research to solve the mistakes and make the system as accurate as possible. The researchers found that property and data recording storage were the main limitations. Finally, this research is an undergraduate study, so it is a study to develop a prototype to be presented to organizations that want to expand the research results further.

6 Research limitations and suggestions

Suggestions that should be improved and adapted in further research, including categorizing vocabulary and phrases according to patient symptoms, require sub-categorization to expand research results and usefulness to users. This is significant because computer models can learn and make more accurate predictions. In addition, statistics on vocabulary and phrases that are likely to be frequently used and essential for communication between patients and medical personnel need to be accelerated. The reason is that space and resource constraints can be controlled by the terms necessary for communication.

Finally, an influential recommendation is that public health organizations must continue investing in information and communication technology to solve critical and necessary problems for patients in healthcare facilities.

7 Conclusion

In this study, researchers selected 60 basic Thai sign language gestures commonly used for medical visits for describing health symptoms. These gestures were utilized to train a machine learning model on the Teachable Machine platform, which employs deep learning algorithms for model creation and training. The model was integrated into a mobile application to detect and interpret signing gestures from video feeds. Following software engineering best practices, the application was designed and developed together with an Artificial intelligence (AI) chatbot to enable doctors and deaf/hard-of-hearing patients to communicate, using a structured software development life cycle (SDLC) approach. When tested, the model achieved 69.61% accuracy in recognizing the 60 signs. The application received a 4.43 out of 5 user satisfaction rating. However, 18.33% of signs language had under 50% recognition accuracy. These problematic signs included blurred vision, swallowing issues, constipation, epilepsy, food allergy, headache, liquid medicine, stress, swelling, chest tightness, and worry. As such, exploring machine learning approaches and deep learning approaches may improve accuracy for intricate signs language in the future.

In summary, this prototype mobile application demonstrates capabilities in translating basic medical sign language, although performance for some gestures was disappointing. The application shows potential to facilitate communication between deaf individuals and healthcare providers regarding basic symptoms and ailments. However, future work may require exploring machine learning approaches to improve model accuracy across a wider range of intricate signs. While still in early stages, this research indicates feasibility of using technology to interpret healthcare signing language. This foundation could ultimately help bridge communication barriers and enhance inclusivity in healthcare services.

8 Conflicts of interest

The authors declare no conflicts of interest.

9 Acknowledgment

This research project was supported by the Thailand Science Research and Innovation Fund and the University of Phayao. In addition, this research was supported by many advisors, academics, researchers, staff, and students. The authors would like to thank all of them for their support and collaboration in making this research possible.

References

1. Xia, K., Lu, W., Fan, H., Zhao, Q.: A Sign Language Recognition System Applied to Deaf-Mute Medical Consultation. Sensors. 22, 9107 (2022). https://doi.org/10.3390/s22239107.
2. Sosa-Jiménez, C.O., Ríos-Figueroa, H.V., Solís-González-Cosío, A.L.: A Prototype for Mexican Sign Language Recognition and Synthesis in Support of a Primary Care Physician. IEEE Access. 10, 127620–127635 (2022). https://doi.org/10.1109/ACCESS.2022.3226696.
3. Novaliendry, D., Pratama, M.F.P., Budayawan, K., Huda, Y., Rahiman, W.M.Y.: Design and Development of Sign Language Learning Application for Special Needs Students Based on Android Using Flutter. Int. J. Online Biomed. Eng. IJOE. 19, 76–92 (2023). https://doi.org/10.3991/ijoe.v19i16.44669.
4. Chaikaew, A., Somkuan, K., Sarapee, P.: Mobile Application for Thai Sign language. In: 2018 22nd International Computer Science and Engineering Conference (ICSEC). pp. 1–4 (2018). https://doi.org/10.1109/ICSEC.2018.8712709.
5. Chaikaew, A., Somkuan, K., Yuyen, T.: Thai Sign Language Recognition: an Application of Deep Neural Network. In: 2021 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunication Engineering. pp. 128–131 (2021). https://doi.org/10.1109/ECTIDAMTNCON51128.2021.9425711.
6. Sudthipadh, J., Pumrin, S.: Development of Thai Sign Language Interpretation with MediaPipe. In: 2022 37th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC). pp. 841–843 (2022). https://doi.org/10.1109/ITC-CSCC55581.2022.9894962.
7. Vijitkunsawat, W., Racharak, T., Nguyen, C., Minh, N.: Video-Based Sign Language Digit Recognition for the Thai Language: A New Dataset and Method Comparisons. In: Proceedings of the 12th International Conference on Pattern Recognition Applications and Methods - Volume 1: ICPRAM,. pp. 775–782. SciTePress (2023). https://doi.org/10.5220/0011643700003411.
8. Kothadiya, D.R., Bhatt, C.M., Rehman, A., Alamri, F.S., Saba, T.: SignExplainer: An Explainable AI-Enabled Framework for Sign Language Recognition With Ensemble Learning. IEEE Access. 11, 47410–47419 (2023). https://doi.org/10.1109/ACCESS.2023.3274851.
9. Feng, K.J.K., Mcdonald, D.W.: Addressing UX Practitioners' Challenges in Designing ML Applications: an Interactive Machine Learning Approach. In: Proceedings of the 28th International Conference on Intelligent User Interfaces. pp. 337–352. Association for Computing Machinery, New York, NY, USA (2023). https://doi.org/10.1145/3581641.3584064.
10. Zhang, P., Carey, J., Te'eni, D., Tremaine, M.: Integrating Human-Computer Interaction Development into the Systems Development Life Cycle: A Methodology. Commun. Assoc. Inf. Syst. 15, (2005). https://doi.org/10.17705/1CAIS.01529.

Gmail Pratya Nuankaew <nuankaew.p@gmail.com>

ISBM 2024 submission 174

1 message

ISBM 2024 <isbm2024@easychair.org> Sun, Jun 23, 2024 at 3:56 PM
To: Pratya Nuankaew <pratya.nu@up.ac.th>

Dear authors,

We received your submission to ISBM 2024 (World Conference on Information Systems for Business Management):

Authors : Wongpanya Nuankaew, Natthida Nuttaphum, Thapanapong Sararat, Phanombongkot Banyaem and Pratya Nuankaew
Title : Sign Language Detection Mobile Application for Thai Patients Using Medical Image Processing to Support Medical Consultations
Number : 174

The submission was uploaded by Pratya Nuankaew <pratya.nu@up.ac.th>.
You can access it via the ISBM 2024 EasyChair Web page

https://easychair.org/conferences/?conf=isbm2024

Thank you for submitting to ISBM 2024.

Best regards,
EasyChair for ISBM 2024.

Gmail Pratya Nuankaew <nuankaew.p@gmail.com>

ISBM 2024 notification for paper 174

1 message

ISBM 2024 <isbm2024@easychair.org> Mon, Jul 1, 2024 at 2:07 PM
To: Pratya Nuankaew <pratya.nu@up.ac.th>

Dear Pratya Nuankaew,

Paper ID : 174

Title : Sign Language Detection Mobile Application for Thai Patients Using Medical Image Processing to Support Medical Consultations

Congratulations! On behalf of the Program Committee of ISBM 2024 - Bangkok, Thailand, I am happy to inform you that your above-mentioned paper has been ACCEPTED for oral presentation in ISBM 2024 and publication in Springer SIST series subject to fulfillment of Guidelines by Springer. An accepted paper will be published in the Springer proceedings (SIST) only if the final version is accompanied by the payment information (i.e transaction reference number) subject to quality check as per Springer Guidelines.

Kindly follow the below-mentioned guidelines (strictly), related to the preparation of the final manuscript, copyright transfer form, payment, and final submission. The procedure has been detailed as a five-step process (I)-(III):

(I) Preparation of CRC (Final Manuscript for Inclusion in Springer Proceedings Book-SIST Series)

You are requested to give strict attention to the below-mentioned points during the preparation of the final manuscript to avoid any last-minute revert backs from the publisher (Springer).

(a) Please format your paper on the Springer template downloaded from our conference website: https://isbm.ict4sd.org//cfp.php#important_downloads

(b) Authors should not mention designation etc. in the name and affiliation part below the paper title. You are requested to mention only the name, college/organization name, city name, and mail ids of all the authors. However, they may mention acknowledgment if necessary (at the end of the paper-prior to references). Kindly mention the full name of your institution and city name.

(c) Authors should not give any photos and biography at the end of the paper.

(d) Please address the comments raised by the reviewer(s) as far as practicable. The responsibility lies with the author to incorporate these comments in preparation for your final Camera Ready Copy (CRC) for submission. However, it is worth noting that your paper will not be reviewed again and the submitted CRC (after payment) will be directed to Springer for publication.

(e) In case there is any content matching issue that is communicated to you by review comments, kindly see that you rewrite the paper to avoid any plagiarism issue in the future, in no case organizer will be responsible for dropping off the paper from Springer Publication at a later stage under such circumstances.

(f) The sole responsibility for the correctness of the paper lies with the Author(s).

(g) It is mandatory that the submitted manuscript is solely from the authors' own work and not from the work of others (unless explicit permission for the same has been granted). This includes text, figures, and tables. Information from published articles must always be cited explicitly. Proper citation is to give the credit to the work that is originally published (not to follow-up work or reviews).

(h) Citations should be given close to the information within the sentences or at the end of the sentence, not after several sentences or near the end of the paragraph. Even when citations are given, the exact copying of a whole sentence or paragraphs should be indicated by quotation marks. Furthermore, the reuse of part of a published figure or table requires copyright permission from the publishers that hold the rights. All re-published figures and tables should explicitly indicate the original source.

(i) You have to doubly ensure the format of the reference section as well as the mode of citations. All references which are cited must be complete in terms of details like Year of publication, Vol No, Issue No, pages, name of conference/journal, etc. wherever required. [Refer to detailed guidelines given in the template itself] All references listed at the end (under reference section) must be cited in the body of the text.

(j) Prepare your CRC (manuscript) using Springer template without containing any page numbers or any special headers or footers. From the final CRC formatted, authors should generate the electronic version of CRC in PDF format as well.

(k) You can add or delete the author's name in your final paper. Once sent no further alteration will be possible.

(l) By submitting the final paper it is understood that all co-authors have agreed to the submission of the paper.

(II) Author(s) Registration

(a) Each paper must be registered for the inclusion of paper in Springer SIST.

(b) Extra page length charges would be strictly applicable beyond 10 pages (after converting to Springer format only). You have to make a combined payment in such a case (i.e. Normal Registration + Extra Page Charges). For paper registration and extra page charges kindly check the conf website for details. If by chance extra page charges are not included, the paper may not be forwarded to Springer for publication. (Extra Pages can be allowed per paper with an Extra payment charge per page)

(c) One author of the paper gets automatically registered along with the paper, other authors if wish to participate in the conference have to separately register them under the participant category.

(d) You are requested to attend the conference and present your paper under the oral presentation session. If the registered/corresponding author is unable to attend the conference, the same can be attended and presented on his/her behalf by the co-authors. However, if none of the authors are able to attend the conference, then it is mandatory that you should give us prior information and seek permission. Your registered paper will be forwarded for publication in Springer. But, it may be noted that the presentation certificate will not be issued in any case.

(e) The paper will not be uploaded to Springer unless we receive the complete payment including Extra Pages.

(III) Payments

The Complete payment and registration process is online on the registration page of our conference website: https://isbm.ict4sd.org//registration.php

(IV) Final Submission

Your final submission is the last step of the registration/submission process.

You are requested to submit the following files after renaming those with your paper id.

(a) Final camera-ready paper in Springer format: PDF file.

(b) Final camera-ready Springer format source paper: Either in WORD file or in Latex. (Mandatory).

(c) Kindly rename all files with your PAPER ID (for example 164.pdf, 164.docx etc.) .

(d) Store all files in a folder again named as your paper id and convert the same to zipped form.

(e) Online transfer Reference number to be added in the mail.

(f) Forward this zipped folder to us as an attachment in ONE mail only to: support@isbm.ict4sd.org

With your paper id in the subject column of the email (as: Final Submission ISBM 2024- PAPER ID)

Once we get the confirmation of the receipt of the payment and receiving of Final Submission, authors will be directed for the signing of Consent to publish.

(V) Completion/Filling of Copyright Form(a) Conference team will fill the CTP form and share it with you.

(b) Take a printout of this Copyright form and kindly mention/enter the following:

Title of Book or Conference Name: World Conference on Information Systems for Business Management (ISBM 2024) Author(s): Mention the names of all the author(s) in the same order as mentioned in your paper.

Corresponding Author’s Name, Address, Affiliation, and Email: The author signing the copyright form should mention his/her details: Full name, affiliation, complete postal address, and email id under this head.Title: Mention your complete paper title

(c) After making the above entries in the consent to publish, it can be signed by any author (on behalf of all the authors), i.e. the corresponding author and mention the date.

(d) Scan this completed and signed copyright form and convert it to pdf. (you can also use the facility of e-signature available in PDF)

**The consent to publish form will be shared with the authors after the registration.

Please complete the above process (Registration) by 9th July 2024. (Strict deadline).

We would like to further extend our congratulations to you and we are looking forward to meeting you in Bangkok, India.

On behalf of the program committee and team ISBM 2024

Program Secretary

Email - support@isbm.ict4sd.org

Once again, I thank you on behalf of the organizing committee for your interest in ISBM 2024. Please treat this letter as an Official document for all conference-related activities & quote the Paper No. & Name for future correspondence.

Last Date for Registration and Payments - 9th July 2024

With regards and best wishes.

(Program Secretary, ISBM 2024)

SUBMISSION: 174
TITLE: Sign Language Detection Mobile Application for Thai Patients Using Medical Image Processing to Support Medical Consultations

---------- REVIEW 1 ----------
SUBMISSION: 174
TITLE: Sign Language Detection Mobile Application for Thai Patients Using Medical Image Processing to Support Medical Consultations
AUTHORS: Wongpanya Nuankaew, Natthida Nuttaphum, Thapanapong Sararat, Phanombongkot Banyaem and Pratya Nuankaew

---------- Overall evaluation ----------
SCORE: 2 (accept)
----- TEXT:
- The study addresses a significant issue of communication barriers between healthcare providers and patients with hearing impairments.

- The use of machine learning and mobile technology to develop a sign language detection application is innovative and timely.

- How do you plan to address the accuracy challenges for signs with lower recognition rates?

- Have you considered integrating feedback mechanisms from users to continuously improve the application?

---------- REVIEW 2 ----------
SUBMISSION: 174
TITLE: Sign Language Detection Mobile Application for Thai Patients Using Medical Image Processing to Support Medical Consultations
AUTHORS: Wongpanya Nuankaew, Natthida Nuttaphum, Thapanapong Sararat, Phanombongkot Banyaem and Pratya Nuankaew

---------- Overall evaluation ----------
SCORE: 2 (accept)
----- TEXT:
- The integration of an AI chatbot enhances the functionality of the application, allowing for more interactive and efficient communication.

- The thorough methodology, including data collection, model training, and application development, is well-

documented and follows best practices.

- What are the next steps in terms of scaling this application for broader deployment in healthcare facilities?

- Expand the application's capabilities to include real-time translation and additional functionalities that cater to specific medical scenarios.

Classification Model for Screening Knee Osteoarthritis Patients in Northern Thailand Using Data Analytics

Pratya Nuankaew[1[0000-0002-3297-4198]], Prachayakorn Chanthong[2[0009-0000-5799-1309]], Tanaporn Somphuak[3[0009-0008-0682-7304]], Thapanapong Sararat[3[0000-0001-8504-4026]], and Wongpanya S. Nuankaew[5[0000-0003-3805-9529]]

[1, 2, 3, 4,5] School of Information and Communication Technology, University of Phayao, Phayao 56000, Thailand
{ pratya.nu, 64020114, 64020079, thapanapong.sa, wongpanya.nu}@up.ac.th

Abstract. Thailand faces a continuously increasing number of older adults, which impacts future healthcare arrangements for older people. This research aims to achieve three objectives: to study the context and risks of knee osteoarthritis based on Thai people's lifestyle and food consumption in the northern region of Thailand, to construct the classification model for screening knee osteoarthritis patients in northern Thailand using data analytics, and to evaluate the classification model for screening knee osteoarthritis patients in northern Thailand using data analytics. The dataset is a sample of 360 people from eight villages in the Sop Prap Subdistrict, Sop Prap District, Lampang Province, Northern Thailand. The research tools are classified into two main areas: descriptive statistical analysis and data mining analysis, which has eight techniques: Naïve Bayes, Decision Tree, Generalized Linear Model, Logistic Regression, Deep Learning, Random Forest, Gradient Boosted Trees, and Support Vector Machine. Model performance testing uses cross-validation techniques and confusion matrix performance, which has three indicators: accuracy, precision, and recall. The results showed that the majority of patients with osteoarthritis were female, aged 51 years and over, and whose main occupation was agriculture. In addition, the sample group also has significant awareness of health care and food consumption control. At the same time, the overall developed model had high accuracy values for all techniques, with the GBT technique having the highest accuracy value of 0.827 and S.D. equal to 0.24. Therefore, the results of this research are a report on the context of care and self-protection of the elderly in the area. This mathematical model can be applied in planning care for the Thai elderly society in the future.

Keywords: Applied Informatics, Medical Data Mining, Medical Informatics, Medical Innovations, Screening Osteoarthritis.

1 Introduction

According to the World Health Organization, Long et al. (2022) [1], and Cieza et al. (2021) [2], there is an increase in patients with bones and joints, especially osteoarthritis, which has an increasing rate of this disease every year. Osteoarthritis usually occurs in elderly people aged 60 years and over and affects up to 10 percent

of patients. In addition, wrists are the part where degeneration symptoms are most common and more commonly found in females than males. Meanwhile, older adults 65 years and older will most likely experience deterioration in their knees, hips, wrists, spine, and ankles.

For Thailand, knee osteoarthritis is a significant public health problem [3, 4]. It is commonly found among middle-aged and elderly people and tends to increase continuously in patients with severe knee osteoarthritis. In patients who do not receive proper treatment, osteoarthritis progresses to a condition where articular cartilage deteriorates. It is destroyed, along with a decrease in synovial fluid, which causes the ends of joint bones to develop. The knees rub against each other directly, causing bone growth around the periphery of the knee joint to degenerate, resulting in pain, inflammation, and loss of joint movement. A survey in Thailand found that in 2013, 72.90% of elderly people had never been examined for osteoarthritis. The group that was discussed and diagnosed with this disease was only 10.60 percent, and it was found that 10.60 percent began to have abnormal symptoms. 12.80 had moderately severe symptoms, 16.20 percent and 4.40 percent had severe symptoms, and 56.40 percent did not experience any abnormal symptoms.

Currently, knee osteoarthritis patients can be screened in various ways, such as questioning the patient's history, using an osteoarthritis screening assessment (the Oxford Knee Score [5–7]), physical examination, laboratory testing, etc. History questionnaires and physical examinations are simple and convenient methods for screening patients with osteoarthritis, but they may not be highly accurate. The researchers, therefore, used data mining techniques to find hidden knowledge in big data, using it as a supplementary tool for screening patients with osteoarthritis. Data mining techniques can solve many problems, including prediction, trend analysis, clustering, and relationship mining. Using data mining techniques to develop a screening model for osteoarthritis patients has many advantages, such as analyzing large amounts of data efficiently and accurately identifying risk factors for osteoarthritis. In addition, models can be developed and adapted to suit various target groups.

From the foregoing, the researchers are confident and believe that the classification model for screening knee osteoarthritis patients in Northern Thailand using data analytics will help detect osteoarthritis patients at an early stage, which will help start treatment sooner and more effectively. In addition, this model can be applied to develop guidelines for preventing osteoarthritis.

Research Objectives.

1. To study the context and risks of knee osteoarthritis based on Thai people's lifestyle and food consumption in the northern region of Thailand.
2. To construct the classification model for screening knee osteoarthritis patients in northern Thailand using data analytics.
3. To evaluate the classification model for screening knee osteoarthritis patients in northern Thailand using data analytics.

2 Materials & Methods

2.1 Population and Sampling

The population and sample selection were determined in the northern region of Thailand by selecting eight villages from Sop Prap Subdistrict, Sop Prap District, Lampang Province, Thailand. The sample was determined using the Krejcie and Morgan method and the equal proportion method. The population and sample used in the research are shown in Table 1.

Table 1. Population and Sampling.

Village Name	Population	Samples with Equal Proportion	Osteoarthritis Patients	General Public	Actual Samples
Baan Phae	593	47	24	24	48
Baan Thung	801	63	32	32	64
Baan Wattana	449	35	18	18	36
Baan Thung Phatthana	927	73	37	37	74
Baan Mai Watthana	264	21	11	11	22
Baan Thung Charoen	549	43	22	22	44
Baan Thung Ruang Thong	590	46	23	23	46
Baan Hong Pu Samakkhi	328	26	13	13	26
Total:	4,501	354	180	180	360

Table 1 presents the research population and sample. The researchers divided the sample into two parts: patients with osteoarthritis and the general public in the target area. The latter will be used to develop a classification model and forecast the risk of osteoarthritis among people in northern Thailand.

2.2 Research Tools

Research tools were divided into two main parts: data collection tools, where researchers used questionnaires to collect data, and tools for building the classification model for screening knee osteoarthritis patients in northern Thailand, where researchers used data mining techniques.

Questionnaire.

The questionnaires used in the research were all communicated and conducted in Thai so that the data could be given truthfully. The questionnaire in this research was divided into three main parts: the first part collected general information from the informants, the second part assessed the risk of osteoarthritis using the Oxford Knee Score (OKS) for the Thai version, which is accepted in Thailand [5, 8], and the third part surveyed the sample group's food consumption behavior. In addition, this questionnaire has been evaluated by seven experts, including five public health experts and two academics from educational institutions. The details of the three sections are presented in Table 2 to Table 5.

Table 2. General Information Query.

Questions and Options	
Consent to provide your data for this research.	
❒ Consent,	❒ Not consent
Gender	
❒ Male	❒ Female
Year of Birth	
Weight ..... Kilograms	Height ..... Centimeters
Occupation	
❒ Agriculturist	❒ Enterprise employee
❒ General employee	❒ Government official
❒ Housekeeper	❒ Merchant/Own business
❒ Student	❒ Unemployed
Education level	
❒ Primary level	❒ Secondary level
❒ Diploma level	❒ Bachelor level
❒ Higher than a bachelor's degree	
Number of family members	
Congenital disease (as diagnosed by a doctor)	
❒ No congenital disease	
❒ Blood pressure disease	
❒ Cardiovascular disease	
❒ Diabetes	
❒ Gout	
❒ Hyperlipidemia disease	
❒ Kidney disease	
❒ Osteoarthritis	
❒ Rheumatoid disease	

Table 3. The Oxford Knee Score (OKS) English version.

Questions and Options		
During the past 4 weeks.....		
1. How would you describe the pain you <u>usually</u> have from your knee?		
❒ None (4)	❒ Very mild (3)	❒ Mild (2)
❒ Moderate (1)	❒ Severe (0)	
2. Have you had any trouble with washing and drying yourself (all over) <u>because of your knee</u>?		
❒ No trouble at all (4)	❒ Very little trouble (3)	❒ Moderate trouble (2)
❒ Extreme difficulty (1)	❒ Impossible to do (0)	
3. Have you had any trouble getting in and out of a car or using public transport <u>because of your knee</u>? (whichever you would tend to use)		
❒ No trouble at all (4)	❒ Very little trouble (3)	❒ Moderate trouble (2)
❒ Extreme difficulty (1)	❒ Impossible to do (0)	
4. For how long have you been able to walk before <u>pain from your knee</u> becomes severe? (with or without a stick)		
❒ No pain/More than 30 minutes (4)		❒ 16 to 30 minutes (3)
❒ 5 to 15 minutes (2)		❒ Around the house <u>only</u> (1)
❒ Not at all - pain severe when walking (0)		

Questions and Options
5. After a meal (sat at a table), how painful has it been for you to stand up from a chair because of your knee? ❒ Not at all painful (4) ❒ Slightly painful (3) ❒ Moderately painful (2) ❒ Very painful (1) ❒ Unbearable (0)
6. Have you been limping when walking, because of your knee? ❒ Rarely/never (4) ❒ Sometimes, or just at first (3) ❒ Often, not just at first (2) ❒ Most of the time (1) ❒ All of the time (0)
7. **Could** you kneel down and get up again afterwards? ❒ Yes, Easily (4) ❒ With little difficulty (3) ❒ With moderate difficulty (2) ❒ With extreme difficulty (1) ❒ No, Impossible (0)
8. Have you been troubled by pain from your knee in bed at night? ❒ No nights (4) ❒ Only 1 or 2 nights (3) ❒ Some nights (2) ❒ Most nights (1) ❒ Every night (0)
9. How much has pain from your knee interfered with your usual work (including housework)? ❒ Not at all (4) ❒ A little bit (3) ❒ Moderately (2) ❒ Greatly (1) ❒ Totally (0)
10. Have you felt that your knee might suddenly 'give way' or let you down? ❒ Rarely/never (4) ❒ Sometimes, or just at first (3) ❒ Often, not just at first (2) ❒ Most of the time(1) ❒ All of the time (0)
11. **Could** you do the household shopping on your own? ❒ Yes, Easily (4) ❒ With little difficulty (3) ❒ With moderate difficulty (2) ❒ With extreme difficulty (1) ❒ No, Impossible (0)
12. **Could** you walk down one flight of stairs? ❒ Yes, Easily (4) ❒ With little difficulty (3) ❒ With moderate difficulty (2) ❒ With extreme difficulty (1) ❒ No, Impossible (0)

Table 3 shows the English version of the Oxford Knee Score (OKS). Each question has five options, and the value of each option is displayed. After respondents provided their information, researchers summed and calculated their risk of osteoarthritis using criteria as shown in Table 4.

Table 4. The Oxford Knee Score (OKS) Criteria.

Score range	Score grading	Severity level
0 – 19	May indicate severe knee osteoarthritis.	Level 4
20 – 29	May indicate moderate to severe knee osteoarthritis.	Level 3
30 – 39	May indicate mild to moderate knee osteoarthritis.	Level 2
40 – 48	May indicate satisfactory joint function.	Level 1

Table 5. Food consumption behavior questionnaire.

Stage	Questions for food consumption behavior
Q1.	Consistency in consuming high-fiber foods such as eggplant, passion fruit, guava, etc.
Q2.	Consistency in consuming low-fiber foods such as cucumber, lychee, longan, etc.
Q3.	Consistency in fast food consumption.
Q4.	Consistency in consuming fried foods.
Q5.	Consistency in consuming 6 – 8 glasses of water per day.
Q6.	Consistency in consuming soft drinks and syrups.
Q7.	Consistency in consuming tea and coffee.
Q8.	Consistency in consuming alcohol.
Q9.	Consistency in consuming milk.
Q10.	Consistency in consuming instant noodles and canned foods.
Q11.	Consistency in consuming snacks.
Q12.	Consistency in consuming fermented foods.
Q13.	Consistency in consuming spicy foods.

Table 5 shows questions about food consumption behavior consisting of 13 questions with five levels (scores) of information criteria as follows: Level 1 contains a value of 0 points, which means it has never been consumed. Level 2 includes a value of 1 point, which means it has been consumed 1 – 2 times per week. Level 3 contains a value of 2 points, meaning it has been consumed 3 – 4 times weekly. Level 4 comprises a value of 3 points, which means it has been consumed 5 – 6 times per week. Level 5 retains a value of 4 points, meaning it has been consumed regularly daily.

Cross Industry Standard Process for Data Mining (CRISP-DM).

The researchers used the CRISP-DM [9–11] method to develop a model to classify groups at risk for osteoarthritis. This method consists of six steps: business understanding, data understanding, data preparation, modeling, evaluation, and deployment, as detailed below.

Business Understanding.

Business understanding involves analyzing the research problem to determine the research direction. The researchers found that a large number of people in the target area at Sop Prap Subdistrict, Sop Prap District, Lampang Province, Thailand, had issues with osteoarthritis. Thus, researchers intend to study various behaviors that may influence the result of osteoarthritis.

Data Understanding.

Data understanding concerns understanding the source of data to align with research objectives. The research population consists of people in rural areas, which the Thai government defines as poor. The majority of people work in agriculture. Therefore, daily life and food consumption in the target areas are distinctive and different from those in other regions of Thailand.

To understand this data, therefore, the researchers focuses on studying population behavior that reflects the possible causes of osteoarthritis. However, to balance the data in developing the classification model, the researchers used equal parts data collection between people with osteoarthritis and healthy people in the target area, as shown in the data collection in Table 1.

Data Preparation.

The objective of data preparation is to obtain data ready to create a classification model for patients at risk of developing knee osteoarthritis. This step is very consistent with the data collection process, in which the researchers used face-to-face, individual data collection.

Before starting to collect data, the research team asked the sample subjects whether they consented to providing personal information for the research. If the sample group does not consent to the data, the researchers do not collect it. After the researchers have collected the data according to the set goals, they proceed with data preparation according to CRISP-DM principles [11], which consist of selecting data (no discrepancies), cleaning data (limiting the data to the specified scope), constructing data (defining a data structure based on attributes), integrating data, and formatting data.

Modeling.

In this phase of modeling, many techniques are available, depending on the data collected and the research goals, to achieve the most optimal model. Modeling is closely linked to data preparation. Therefore, researchers need to be aware of data problems that may arise during data collection and affect them while building models.

Techniques used to develop this model include Naïve Bayes, Decision Tree, Generalized Linear Model , Logistic Regression, Deep Learning, Random Forest, Gradient Boosted Trees, and Support Vector Machine. The researchers used all machine learning techniques to classify and predict osteoarthritis risk.

Evaluation.

The purpose of the evaluation is to obtain the most efficient model, and the evaluation method depends on the techniques used to develop the model.

The evaluation techniques in this research include cross-validation technique and confusion matrix performance. The cross-validation technique is a method of dividing data to build and test a model by dividing the data into equal parts called k-fold. Examples of testing data: The data is set to have ten parts. Modeling uses nine pieces of data to create a model called training. The remaining part is used to test, called testing. Each model run is tested using three metrics: accuracy, precision, and recall, known as the “confusion matrix performance”.

Accuracy is calculated as the total number of correctly predicted data divided by the total number. Precision is calculated as correctly forecasted data in each class divided by the total data. The recall value is calculated as the actual data for each class divided by the total data.

Deployment.

The deployment goal is to implement the results of selecting the best or most efficient model. Deployment can be done in many ways. In this research, researchers have planned to choose the most effective model to develop an application to facilitate the recruitment of osteoarthritis patients in Northern Thailand, which is the following research area.

2.3 Data Collection

The data collected this time followed the research objectives and the designed population goals, as summarized in Table 6 to Table 8. The collected data will be prepared and organized to create a patient classification model using the CRISP-DM method and process as presented previously.

Table 6. The collected samples were classified according to gender and status.

Gender	Status		Total
	Osteoarthritis Patients	General Public	
Male	83 (23.06%)	68 (18.89%)	151 (41.94%)
Female	100 (27.78%)	109 (30.28%)	209 (58.06%)
Total:	183 (50.83%)	177 (49.17%)	360 (100.00%)

Table 6 shows the data collection of the sample group by gender and status. The number of females outnumbered males, 209 (58.06%) per 151 (41.94%). When considering patients with osteoarthritis, females also had more patients than males, 100 (27.78%) per 83 (23.06%).

Table 7. The collected samples were classified according to age and status.

Age	Status		Total
	Osteoarthritis Patients	General Public	
Age less than 30 years	0 (0.00%)	0 (0.00%)	0 (0.00%)
Age 30 – 40 years	0 (0.00%)	13 (3.61%)	13 (3.61%)
Age 41 – 50 years	0 (0.00%)	35 (9.72%)	35 (9.72%)
Age 51 – 60 years	61 (16.94%)	37 (10.28%)	98 (27.22%)
Age 61 – 70 years	81 (22.50%)	77 (21.39%)	158 (43.89%)
Age 71 – 80 years	28 (7.78%)	13 (3.61%)	41 (11.39%)
Age 81 – 90 years	11 (3.06%)	2 (0.56%)	13 (3.61%)
Age over 90 years	2 (0.56%)	0 (0.00%)	2 (0.56%)
Total:	183 (50.83%)	177 (49.17%)	360 (100.00%)

Table 7 shows the data collection of the sample group classified by age and status. Osteoarthritis patients were aged 51 years and up, with the age range of most patients being between 61 and 70, numbering 81 cases (22.50%).

Table 8. The collected samples were classified according to occupation and status.

Occupation	Status		Total
	Osteoarthritis Patients	General Public	
Agriculturist	105 (29.17%)	104 (28.89%)	209 (58.06%)
Enterprise employee	16 (4.44%)	15 (4.17%)	31 (8.61%)
General employee	35 (9.72%)	32 (8.89%)	67 (18.61%)
Government official	0 (0.00%)	2 (0.56%)	2 (0.56%)
Housekeeper	2 (0.56%)	2 (0.56%)	4 (1.11%)
Merchant/Own business	16 (4.44%)	22 (6.11%)	38 (10.56%)
Unemployed	9 (2.50%)	0 (0.00%)	9 (2.50%)
Total:	183 (50.83%)	177 (49.17%)	360 (100.00%)

Table 8 shows the data collection of the sample grouped by occupation and status. It was found that the sample had a primary occupation in agriculture (209 samples, 58.06%), classified as 105 patients with osteoarthritis (29.17%) and 104 general farmers (28.89%).

2.4 Analysis and Interpretation

The analysis and interpretation of the results were divided into two main directions. The first direction was the sample's analysis and feasibility assessed according to the Oxford Knee Score Criteria, as presented in Table 4. It is composed of four levels of knee osteoarthritis symptoms. The first level has a total score between 40 and 49 points, meaning the test respondent does not have knee osteoarthritis symptoms. The second level has a total score between 30 and 39 points, meaning the test respondent is at risk of mild to moderate knee osteoarthritis symptoms. The third level has a total score between 20 and 29 points, meaning the test respondent is at risk of moderate to severe knee osteoarthritis symptoms. The fourth level has a total score between 0 and 19 points, meaning the test respondent is at risk of severe knee osteoarthritis symptoms. The sample's descriptive analysis includes statistics such as count, mean, mode, median, percentage, and standard deviation (S.D.), as presented in Table 9 and Table 10.

The second direction was model performance, as demonstrated in the evaluation section of the CRISP-DM processes. It consists of partitioning the test data using cross-validation techniques and three confusion metric performance indicators: accuracy, precision, and recall, as presented in Tables 11 and 12.

3 Results

The research report is divided into three parts according to the two directions of interpretation analysis. The first part summarizes the results of the sample group's risk assessment for knee osteoarthritis, as detailed in Table 9. The second part summarizes the sample behavior data using descriptive statistics techniques, as detailed in Table 9. The third part reports on the model development and selection of practical models, as detailed in Tables 11 and 12.

3.1 Results of Knee Osteoarthritis Risk Assessment

Table 9. The results of knee osteoarthritis risk assessment.

Stage of Knee Osteoarthritis	Score grading	Instances
Level 4	May indicate severe knee osteoarthritis.	19 (5.28%)
Level 3	May indicate moderate to severe knee osteoarthritis.	149 (41.39%)
Level 2	May indicate mild to moderate knee osteoarthritis.	19 (5.28%)
Level 1	May indicate satisfactory joint function.	173 (48.06%)
Total:		360 (100.00%)

Table 9 shows the results of the analysis of the severity of osteoarthritis. The majority were people with standard bodies. However, most respondents had level 3 osteoarthritis, which was at a moderate to critical level, with 149 cases, accounting for 41.39 percent.

3.2 Consumption Behavior Analysis Results

Table 10. The results of consumption behavior analysis.

Stage	Quantity (Percentage)					Average	S.D.
	Score 4	Score 3	Score 2	Score 1	Score 0		
Q1.	7 (1.94%)	6 (1.67%)	89 (24.72%)	200 (55.56%)	58 (16.11%)	1.18	0.79
Q2.	3 (0.83%)	1 (0.28%)	135 (37.50%)	165 (45.83%)	56 (15.56%)	1.25	0.74
Q3.	1 (0.28%)	15 (4.17%)	54 (15.00%)	107 (29.72%)	183 (50.83%)	0.73	0.88
Q4.	1 (0.28%)	28 (7.78%)	87 (24.17%)	133 (36.94%)	111 (30.83%)	1.1	0.94
Q5.	356 (98.89%)	3 (0.83%)	1 (0.28%)	0 (0.00%)	0 (0.00%)	3.99	0.14
Q6.	2 (0.56%)	1 (0.28%)	22 (6.11%)	130 (36.11%)	205 (56.94%)	0.51	0.67
Q7.	36 (10.00%)	8 (2.22%)	27 (7.50%)	77 (21.39%)	212 (58.89%)	0.83	1.27
Q8.	3 (0.83%)	3 (0.83%)	20 (5.56%)	120 (33.33%)	214 (59.44%)	0.5	0.71
Q9.	151 (41.94%)	27 (7.50%)	44 (12.22%)	93 (25.83%)	45 (12.50%)	2.41	1.53
Q10.	0 (0.00%)	1 (0.28%)	11 (3.06%)	136 (37.78%)	212 (58.89%)	0.45	0.57
Q11.	2 (0.56%)	0 (0.00%)	44 (12.22%)	156 (43.33%)	158 (43.89%)	0.7	0.72
Q12.	0 (0.00%)	4 (1.11%)	51 (14.17%)	153 (42.50%)	152 (42.22%)	0.74	0.74
Q13.	12 (3.33%)	59 (16.39%)	119 (33.06%)	126 (35.00%)	44 (12.22%)	1.64	0.99

Table 10 presents information for all 360 sample subjects, which can be analyzed to show that the sample groups take good care of themselves. For example, question number 5 asks about drinking eight glasses of water per day, which the majority of the sample group perceives that they should do, with a frequency equal to 356 (98.89%) of the total sample group. However, this information only describes the overall context of the population and sample. It cannot be clarified and needs to be developed into a mathematical model, as presented in Tables 11 and 12.

3.3 Results of Development and Selection of Practical Models

Table 11. The results of the model construction and model performance.

Model	Accuracy	S.D.	Gains	Training Time*
Naïve Bayes	0.748	0.36	56	256 ms
Decision Tree	0.769	0.23	62	108 ms
Generalized Linear Model	0.770	0.49	60	747 ms
Logistic Regression	0.743	0.64	56	550 ms
Deep Learning	0.738	0.29	54	472 ms
Random Forest	0.787	0.54	64	150 ms
Gradient Boosted Trees	0.827	0.24	72	306 ms
Support Vector Machine	0.815	0.67	70	186 ms

* Training Time (1,000 Rows)

Table 11 shows the results of model construction and model performance testing. It was found that the model built with Gradient Boosted Trees (GBT) technique had the highest accuracy with a value of 0.827 and an S.D. of 0.24. The details of testing with Confusion Matrix as shown in Table 12.

Table 12. The Gradient Boosted Trees (GBT) model performance.

	True Level 4	True Level 3	True Level 2	True Level 1	Class Precision
Pred. Level 4	2	2	0	0	50.00%
Pred. Level 3	2	36	1	2	87.80%
Pred. Level 2	0	0	0	0	0.00%
Pred. Level 1	4	3	4	48	81.36%
Class Recall	25.00%	87.80%	0.00%	96.00%	

4 Discussions

All discussions are based on hypotheses and set research objectives. The first issue is to discuss the context and risks of knee osteoarthritis based on Thai people's lifestyle and food consumption in the northern region of Thailand. The researchers found that most of the sample had knee osteoarthritis at level 3, with 149 instances (41.39%), as shown in Table 9. Level 3 means the patient's symptoms may indicate moderate to severe knee osteoarthritis. The sample's osteoarthritis may be due to the nature of their occupations, most of which are agriculture, with 105 farmers (29.17%), as shown in Table 8, which requires using the body for occupation.

For this reason, it is necessary to study certain behaviors that may reflect the causes of disease. Researchers are interested in consumption characteristics. The researchers found that the samples were aware of their consumption and self-care, reinforcing their past lifestyles or occupations, as detailed in Table 10. Moreover, an in-depth study in which researchers use data mining and machine learning techniques to create mathematical models, as presented in the second issue, is necessary.

The researchers found that even though the sample took good care of themselves regarding food consumption, factors may play a role in osteoarthritis, which researchers believe uses mathematical models and data analytics to reveal more complex issues. The second issue concerns constructing the classification model for screening knee osteoarthritis patients in northern Thailand. Researchers used eight machine-learning techniques to devise an osteoarthritis risk classification model. These include Naïve Bayes, Decision Tree, Generalized Linear Model, Logistic Regression, Deep Learning, Random Forest, Gradient Boosted Trees, and Support Vector Machine, as reported in Table 11. Overall, all models performed well, with the model developed using the Gradient Boosted Trees (GBT) technique having the highest accuracy, with 0.827 and S.D. equal to 0.24. Therefore, it was selected to further test its performance, which is presented in the third issue.

The third issue aims to evaluate the classification model for screening knee osteoarthritis patients in northern Thailand using data analytics. Although the overall performance of the model tests was high, the prediction of level 2 illnesses still had problems, with the model's prediction precision and recall equal to 0.00%, which are essential problems that need to be addressed.

In conclusion, this research achieved the results as set out. The researchers will use the discussion to improve in the following research.

5 Conclusion

As Thailand enters an aging society, where the majority of citizens are over 60, this will affect the preparation and care of the people. Illness is inevitable, but preparation for dealing with it can be done. Therefore, this research has three objectives: to study the context and risks of knee osteoarthritis based on Thai people's lifestyle and food consumption in the northern region of Thailand, to construct the classification model for screening knee osteoarthritis patients in northern Thailand using data analytics, and to evaluate the classification model for screening knee osteoarthritis patients in northern Thailand using data analytics.

The context in which the study was conducted found that mostly females had a higher number of people with osteoarthritis than males. In addition, patients with osteoarthritis will be aged 51 years and up, consistent with Tangtrakulwanich and Suwanno's research [12]. In addition, Fransen et al., [13] also suggests that osteoarthritis is influenced by occupation.

The researchers, therefore, developed a risk classification model for osteoarthritis. All the models developed had high overall performance, with the Gradient Boosted Trees (GBT) technique, as shown in Table 11, having the best performance, with 0.827 and S.D. equal to 0.24. It was selected to test its performance, as detailed in Table 12. It was found that although the overall model performs well and efficiently,

there are still some classes where the model does not predict correctly, which requires further improvement and research. However, this research is an undergraduate project that aims to serve as a guideline for compiling the knowledge that students have studied in the Bachelor of Business Administration program in Business Computer at the School of Information and Communication Technology, University of Phayao. So that students can apply the knowledge they have gained to use in the future. Researchers, therefore, expect to promote and support research for the further development of humanity.

6 Limitations

A notable limitation of this research is that it is an undergraduate student project that encourages students to learn and explore research experiences. Therefore, some steps have slight glitches. However, all researchers tried their best to achieve all the objectives of this research.

7 Conflict of Interest

The researchers declare that there is no conflict of interest for this research.

8 Acknowledgement

This research project was supported by the Thailand Science Research and Innovation Fund and the University of Phayao. It also received support from many advisors, academics, researchers, students, and staff. The authors thank everyone for their support and cooperation in completing this research.

References

1. Long, H., Liu, Q., Yin, H., Wang, K., Diao, N., Zhang, Y., Lin, J., Guo, A.: Prevalence Trends of Site-Specific Osteoarthritis From 1990 to 2019: Findings From the Global Burden of Disease Study 2019. Arthritis Rheumatol. Hoboken NJ. 74, 1172–1183 (2022). https://doi.org/10.1002/art.42089.
2. Cieza, A., Causey, K., Kamenov, K., Hanson, S.W., Chatterji, S., Vos, T.: Global estimates of the need for rehabilitation based on the Global Burden of Disease study 2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Lond. Engl. 396, 2006–2017 (2021). https://doi.org/10.1016/S0140-6736(20)32340-0.
3. Roopsawang, I., Aree-Ue, S.: Knee osteoarthritis in adult and older Thais living in rural and urban areas: A comparative study. Pac. Rim Int. J. Nurs. Res. 19, 187–201 (2015).
4. Luksameesate, P., Tanavalee, A., Taychakhoonavudh, S.: An economic evaluation of knee osteoarthritis treatments in Thailand. Front. Pharmacol. 13, (2022). https://doi.org/10.3389/fphar.2022.926431.

5. Charoencholvanich, K., Pongcharoen, B.: Oxford knee score and SF-36: translation & reliability for use with total knee arthroscopy patients in Thailand. J. Med. Assoc. Thail. Chotmaihet Thangphaet. 88, 1194–1202 (2005).
6. Whitehouse, S.L., Blom, A.W., Taylor, A.H., Pattison, G.T.R., Bannister, G.C.: The Oxford Knee Score; problems and pitfalls. The Knee. 12, 287–291 (2005). https://doi.org/10.1016/j.knee.2004.11.005.
7. Jenny, J.-Y., Diesinger, Y.: The Oxford Knee Score: Compared performance before and after knee replacement. Orthop. Traumatol. Surg. Res. 98, 409–412 (2012). https://doi.org/10.1016/j.otsr.2012.03.004.
8. Khonglormyati, J., Khunpra, R.: Effectiveness of Derris Scandens (Roxb.) Benth. (Thaowan Priang) in Elderly Patients with Knee Osteoarthritis. J. Prev. Med. Assoc. Thail. 9, 304–312 (2019).
9. Schröer, C., Kruse, F., Gómez, J.M.: A Systematic Literature Review on Applying CRISP-DM Process Model. Procedia Comput. Sci. 181, 526–534 (2021). https://doi.org/10.1016/j.procs.2021.01.199.
10. Martínez-Plumed, F., Contreras-Ochando, L., Ferri, C., Hernández-Orallo, J., Kull, M., Lachiche, N., Ramírez-Quintana, M.J., Flach, P.: CRISP-DM Twenty Years Later: From Data Mining Processes to Data Science Trajectories. IEEE Trans. Knowl. Data Eng. 33, 3048–3061 (2021). https://doi.org/10.1109/TKDE.2019.2962680.
11. Wirth, R., Hipp, J.: CRISP-DM: Towards a standard process model for data mining. In: Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining. pp. 29–39. Manchester (2000).
12. Tangtrakulwanich, B., Suwanno, P.: Epidemiology and risk factors of patellofemoral osteoarthritis in adults: a population-based study in southern Thailand. J. Med. Assoc. Thai. 95, 1048 (2012).
13. Fransen, M., Bridgett, L., March, L., Hoy, D., Penserga, E., Brooks, P.: The epidemiology of osteoarthritis in Asia. Int. J. Rheum. Dis. 14, 113–121 (2011). https://doi.org/10.1111/j.1756-185X.2011.01608.x.

Gmail Pratya Nuankaew <nuankaew.p@gmail.com>

5th International Conference on Data Engineering and Communication Technology (ICDECT) : Submission (93) has been created.

1 message

Microsoft CMT <email@msr-cmt.org> Tue, Jul 30, 2024 at 9:20 AM
Reply-To: Microsoft CMT - Do Not Reply <noreply@msr-cmt.org>
To: pratya.nu@up.ac.th

Hello,

The following submission has been created.

Track Name: ICDECT2024

Paper ID: 93

Paper Title: Classification Model for Screening Knee Osteoarthritis Patients in Northern Thailand Using Data Analytics

Abstract:
Thailand faces a continuously increasing number of older adults, which impacts future healthcare arrangements for older people. This research aims to achieve three objectives: to study the context and risks of knee osteoarthritis based on Thai people's lifestyle and food consumption in the northern region of Thailand, to construct the classification model for screening knee osteoarthritis patients in northern Thailand using data analytics, and to evaluate the classification model for screening knee osteoarthritis patients in northern Thailand using data analytics. The dataset is a sample of 360 people from eight villages in the Sop Prap Subdistrict, Sop Prap District, Lampang Province, Northern Thailand. The research tools are classified into two main areas: descriptive statistical analysis and data mining analysis, which has eight techniques: Naïve Bayes, Decision Tree, Generalized Linear Model, Logistic Regression, Deep Learning, Random Forest, Gradient Boosted Trees, and Support Vector Machine. Model performance testing uses cross-validation techniques and confusion matrix performance, which has three indicators: accuracy, precision, and recall. The results showed that the majority of patients with osteoarthritis were female, aged 51 years and over, and whose main occupation was agriculture. In addition, the sample group also has significant awareness of health care and food consumption control. At the same time, the overall developed model had high accuracy values for all techniques, with the GBT technique having the highest accuracy value of 0.827 and S.D. equal to 0.24. Therefore, the results of this research are a report on the context of care and self-protection of the elderly in the area. This mathematical model can be applied in planning care for the Thai elderly society in the future.

Created on: Tue, 30 Jul 2024 02:19:56 GMT

Last Modified: Tue, 30 Jul 2024 02:19:56 GMT

Authors:
- pratya.nu@up.ac.th (Primary)
- 64020114@up.ac.th
- 64020079@up.ac.th
- thapanapong.sa@up.ac.th
- wongpanya.nu@up.ac.th

Secondary Subject Areas: Not Entered
Submission Files: ICDECT_2024@icdect.com (03) - 1stDraft [2024-07-29].pdf (226 Kb, Tue, 30 Jul 2024 02:17:28 GMT)

Submission Questions Response: Not Entered

Thanks,
CMT team.

Gmail Pratya Nuankaew <nuankaew.p@gmail.com>

ICDECT-2024:-Notification of Acceptance for Paper-ID [*93*] and Guidelines for Registration & CRC Submission

1 message

Editor ICDECT <editor.icdect@gmail.com> Sun, Aug 18, 2024 at 12:04 PM
To: pratya.nu@up.ac.th, 64020114@up.ac.th, 64020079@up.ac.th, thapanapong.sa@up.ac.th, wongpanya.nu@up.ac.th

ICDECT-2024:-Notification of Acceptance for Paper-ID [*93*] and Guidelines for Registration & CRC Submission

Dear Authors,

Paper-ID: [*93*]
Title: [*Classification Model for Screening Knee Osteoarthritis Patients in Northern Thailand Using Data Analytics*]

On behalf of editorial board of ICDECT-2024, we are happy to inform you that your above referenced paper has been ACCEPTED (based on the review reports) for presentation in ICDECT-2024 conference in hybrid mode to be organized by Asia Pacific University of Technology & Innovation (APU), Technology Park, Kuala Lumpur, Malaysia during 28-29 September 2024.

Please refer to the reviewers' comments attached below which are intended to help you to improve your paper for final publication. You are advised to address the comments of the reviewers in your final paper suitably. **Kindly note that the conference committee follows a rigorous double-blind peer-review system for screening and reviewing papers for the conference proceedings.**

An accepted and presented paper will be published in the conference proceedings (Springer LNNS Series-indexed in SCOPUS, INSPEC, WTI Frankfurt eG, zbMATH, SCImago and Web of Science) only if the final version is accompanied by the registration fee payment information for at least one of the authors.

Instructions for Author(s) Registration & Payment of Registration Fee
---
(a) The details of registration fee as well as guidelines may be referred to at:
https://icdect.com/?page_id=235
(b) Kindly pay the conference registration fee via the following registration form and payment link(s):
https://buy.stripe.com/9AQ3ecaXJh0s6FG7sv (For payment in MYR)
https://buy.stripe.com/00g6qo0j511u9RSfZ2 (For payment in USD)
https://rzp.io/l/7neoIJw (For payment in INR).
(c) After completion of payment through the above link--you are requested to take a snapshot of this payment confirmation, rename it with your paper id and preserve it as [*PAPER-ID*]_PaymentProof.

Mode of Paper Presentation

The conference will be conducted in hybrid mode.
In Physical Mode: September 28, 2024 (Saturday)
Venue: Asia Pacific University of Technology & Innovation (APU), Technology Park Malaysia
In Online Mode: September 29, 2024 (Sunday)-On MS Teams/Google Meet Platform

As the conference is being conducted in hybrid mode; the mode of presentation can be decided by the author individually. The payment of registration shall be made accordingly as there are different

registration fee for physical (offline) and virtual (online) mode.

Guidelines for Preparation of Camera-Ready Copy (CRC) for Publication in Springer Proc. of ICDECT-2024

(a) CRC for ICDECT-2024 should be prepared as per the guidelines mentioned at: https://icdect.com/?page_id=526
(b) Template for preparation of CRC is available at the following link: https://icdect.com/?page_id=164
The process of filling the "Springer Copyright Form/Publishing Agreement Form" by authors will be initiated after the completion of the conference.

Upload of CRC, Payment Proof and Deadline

You are requested to finalize the following files after renaming those with your paper id.
(a) Final paper (based on review comments) in pdf file.
(b) Final paper source file: Either in WORD file or in LaTex (For latex, all the associated files should be given as zip file).
(c) Payment proof of your online registration payment (received on your email after payment)
(d) Kindly rename all files with your PAPER ID (for example: [*PAPER-ID*]__CRC.pdf, [*PAPER-ID*]_CRC.docx, [*PAPER-ID*]_PaymentProof.pdf).
(e) Kindly fill the following submission form to upload the above documents through the given link: https://forms.gle/jyh7TbkZa39Xv6NSA

The above process should be completed by or before: **30th August 2024**

For any queries related to the registration visit the conference website or email the undersigned. Kindly mention your Paper-ID (as mentioned above) in each communication for easy tracking at our end and serving you better in answering your queries.

Thanking You

Warm Regards
Program Chair
ICDECT-2024,
Asia Pacific University of Technology & Innovation (APU),
Kuala Lumpur, Malaysia
Conference Website: https://icdect.com/

*******REVIEW#1*******

This paper provides a valuable contribution to the field of public health and data analytics by developing and evaluating a classification model for knee osteoarthritis screening in Northern Thailand. The study is well-structured, with a clear methodology and robust analysis. However, it could be improved by offering more details on the sample, feature selection, and cross-validation techniques, as well as a deeper discussion of the results and their broader implications.

1. Sample Size Justification ?
2. The paper does not detail the specific features used in the classification models.

The paper is well written and organized .

*******EDITORIAL COMMENTS*******

Authors are advised to strictly incorporate the following points in the final paper:

1. The content matching must be within 20%. Single source matching not to exceed 4%.
2. All authors name, affiliation and mail ids must be there and the corresponding author to be marked. This is very important.
3. No biography or photos of authors should be there.
4. Authors must check the English language carefully.
5. Paper should be at least 10 pages in length on the Springer template. Beyond 12 pages, extra page charges apply.
6. The work must be original and should not have been submitted or published in any other conf or journal. At any point of time, if it is being traced that the paper/content is AI-generated then the paper will be rejected without refund of registration fee. Authors take their own responsibility to ensure that it is not AI generated.
7. Use uniform style for headings and subheading as per Springer template. Do not use IEEE Template.
8. The figures given in the papers should be 800dpi so that the picture clarity can be maintained.
9. All figures, equations and tables must be numbered properly and should be cited within the text.
10. Add missing publication details in references. All the references must be cited within the text.

ที่ อว ๗๓๑๙/ว๐๔๕๗

คณะเทคโนโลยีสารสนเทศและการสื่อสาร
มหาวิทยาลัยพะเยา ตำบลแม่กา
อำเภอเมือง จังหวัดพะเยา ๕๖๐๐๐

๑๗ ธันวาคม ๒๕๖๗

เรื่อง ขอความอนุเคราะห์บุคลากรในสังกัดของท่านเป็นที่ปรึกษาโครงการวิจัย

เรียน คณบดีสำนักวิชาเทคโนโลยีสารสนเทศ มหาวิทยาลัยแม่ฟ้าหลวง

สิ่งที่ส่งมาด้วย แบบตอบรับที่ปรึกษาโครงการวิจัย จำนวน ๑ ฉบับ

ตามที่ มหาวิทยาลัยพะเยา ได้อนุมัติทุนอุดหนุนโครงการวิจัย ประจำปีงบประมาณ พ.ศ. ๒๕๖๘ ให้กับ ผู้ช่วยศาสตราจารย์ ดร. ปรัชญา นวนแก้ว สังกัด สาขาวิชาธุรกิจดิจิทัล คณะเทคโนโลยีสารสนเทศและการสื่อสาร เป็นหัวหน้าโครงการวิจัย "สารสนเทศเชิงประยุกต์เพื่อนวัตกรรมสารสนเทศทางการแพทย์" ตามคำสั่งมหาวิทยาลัยพะเยา เลขที่ ๕๐๒๘/๒๕๖๗ ลงวันที่ ๔ ตุลาคม ๒๕๖๗ ระยะเวลาดำเนินโครงการ ตั้งแต่วันที่ ๑ ตุลาคม ๒๕๖๗ – ๓๐ กันยายน ๒๕๖๘ นั้น

เพื่อให้การดำเนินโครงการวิจัยเป็นไปด้วยความเรียบร้อยและบรรลุตามวัตถุประสงค์ คณะเทคโนโลยีสารสนเทศและการสื่อสาร มหาวิทยาลัยพะเยา พิจารณาแล้วเห็นว่าบุคลากรในสังกัดของท่าน คือ รองศาสตราจารย์ ดร. พรรณฤมล เต็มดี เป็นผู้มีความรู้ความสามารถในงานวิจัยดังกล่าวเป็นอย่างดี จึงขอความอนุเคราะห์บุคลากรในสังกัดของท่านเป็นที่ปรึกษาโครงการวิจัย ทั้งนี้ หากท่านมีข้อสงสัย สามารถติดต่อสอบถามได้ที่ ผู้ช่วยศาสตราจารย์ ดร.ปรัชญา นวนแก้ว เบอร์โทรติดต่อ ๐๘๙-๙๖๑-๔๘๓๒, ๐๘๙-๔๑๙-๗๑๒๙ อีเมลล์ pratya.nu@up.ac.th .

คณะเทคโนโลยีสารสนเทศและการสื่อสาร มหาวิทยาลัยพะเยา หวังเป็นอย่างยิ่งว่าคงได้รับความอนุเคราะห์จากท่านและขอขอบคุณมา ณ โอกาสนี้

จึงเรียนมาเพื่อโปรดพิจารณาให้ความอนุเคราะห์ จะขอบคุณยิ่ง

ขอแสดงความนับถือ

(ผู้ช่วยศาสตราจารย์ ดร.พรเทพ โรจนวสุ)
คณบดีคณะเทคโนโลยีสารสนเทศและการสื่อสาร

คณะเทคโนโลยีสารสนเทศและการสื่อสาร มหาวิทยาลัยพะเยา
โทรศัพท์ ๐-๕๔๔๖-๖๖๖๖ ต่อ ๒๓๒๐

ที่ อว ๗๓๑๙/ว๐๔๕๗

คณะเทคโนโลยีสารสนเทศและการสื่อสาร
มหาวิทยาลัยพะเยา ตำบลแม่กา
อำเภอเมือง จังหวัดพะเยา ๕๖๐๐๐

๑๗ ธันวาคม ๒๕๖๗

เรื่อง ขอความอนุเคราะห์บุคลากรในสังกัดของท่านเป็นที่ปรึกษาโครงการวิจัย

เรียน คณบดีคณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏมหาสารคาม

สิ่งที่ส่งมาด้วย แบบตอบรับที่ปรึกษาโครงการวิจัย จำนวน ๑ ฉบับ

ตามที่ มหาวิทยาลัยพะเยา ได้อนุมัติทุนอุดหนุนโครงการวิจัย ประจำปีงบประมาณ พ.ศ. ๒๕๖๘ ให้กับ ผู้ช่วยศาสตราจารย์ ดร. ปรัชญา นวนแก้ว สังกัด สาขาวิชาธุรกิจดิจิทัล คณะเทคโนโลยีสารสนเทศและการสื่อสาร เป็นหัวหน้าโครงการวิจัย "สารสนเทศเชิงประยุกต์เพื่อนวัตกรรมสารสนเทศทางการแพทย์" ตามคำสั่งมหาวิทยาลัยพะเยา เลขที่ ๕๐๒๘/๒๕๖๗ ลงวันที่ ๔ ตุลาคม ๒๕๖๗ ระยะเวลาดำเนินโครงการ ตั้งแต่วันที่ ๑ ตุลาคม ๒๕๖๗ – ๓๐ กันยายน ๒๕๖๘ นั้น

เพื่อให้การดำเนินโครงการวิจัยเป็นไปด้วยความเรียบร้อยและบรรลุตามวัตถุประสงค์ คณะเทคโนโลยีสารสนเทศและการสื่อสาร มหาวิทยาลัยพะเยา พิจารณาแล้วเห็นว่าบุคลากรในสังกัดของท่าน คือ รองศาสตราจารย์ ดร. สิทธิชัย บุษหมั่น เป็นผู้มีความรู้ความสามารถในงานวิจัยดังกล่าวเป็นอย่างดี จึงขอความอนุเคราะห์บุคลากรในสังกัดของท่านเป็นที่ปรึกษาโครงการวิจัย ทั้งนี้ หากท่านมีข้อสงสัย สามารถติดต่อสอบถามได้ที่ ผู้ช่วยศาสตราจารย์ ดร.ปรัชญา นวนแก้ว เบอร์โทรติดต่อ ๐๘๙-๙๖๑-๔๘๓๒, ๐๘๙-๔๑๙-๗๑๒๙ อีเมลล์ pratya.nu@up.ac.th

คณะเทคโนโลยีสารสนเทศและการสื่อสาร มหาวิทยาลัยพะเยา หวังเป็นอย่างยิ่งว่าคงได้รับความอนุเคราะห์จากท่านและขอขอบคุณมา ณ โอกาสนี้

จึงเรียนมาเพื่อโปรดพิจารณาให้ความอนุเคราะห์ จะขอบคุณยิ่ง

ขอแสดงความนับถือ

(ผู้ช่วยศาสตราจารย์ ดร.พรเทพ โรจนวสุ)
คณบดีคณะเทคโนโลยีสารสนเทศและการสื่อสาร

คณะเทคโนโลยีสารสนเทศและการสื่อสาร มหาวิทยาลัยพะเยา
โทรศัพท์ ๐-๕๔๔๖-๖๖๖๖ ต่อ ๒๓๒๐

ที่ อว ๗๓๑๙/ว๐๔๕๗

คณะเทคโนโลยีสารสนเทศและการสื่อสาร
มหาวิทยาลัยพะเยา ตำบลแม่กา
อำเภอเมือง จังหวัดพะเยา ๕๖๐๐๐

๑๗ ธันวาคม ๒๕๖๗

เรื่อง ขอความอนุเคราะห์บุคลากรในสังกัดของท่านเป็นที่ปรึกษาโครงการวิจัย

เรียน คณบดีคณะบริหารธุรกิจและเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีราชมงคลตะวันออก

สิ่งที่ส่งมาด้วย แบบตอบรับที่ปรึกษาโครงการวิจัย จำนวน ๑ ฉบับ

ตามที่ มหาวิทยาลัยพะเยา ได้อนุมัติทุนอุดหนุนโครงการวิจัย ประจำปีงบประมาณ พ.ศ. ๒๕๖๘ ให้กับ ผู้ช่วยศาสตราจารย์ ดร. ปรัชญา นวนแก้ว สังกัด สาขาวิชาธุรกิจดิจิทัล คณะเทคโนโลยีสารสนเทศและการสื่อสาร เป็นหัวหน้าโครงการวิจัย "สารสนเทศเชิงประยุกต์เพื่อนวัตกรรมสารสนเทศทางการแพทย์" ตามคำสั่งมหาวิทยาลัยพะเยา เลขที่ ๕๐๒๘/๒๕๖๗ ลงวันที่ ๔ ตุลาคม ๒๕๖๗ ระยะเวลาดำเนินโครงการ ตั้งแต่วันที่ ๑ ตุลาคม ๒๕๖๗ – ๓๐ กันยายน ๒๕๖๘ นั้น

เพื่อให้การดำเนินโครงการวิจัยเป็นไปด้วยความเรียบร้อยและบรรลุตามวัตถุประสงค์ คณะเทคโนโลยีสารสนเทศและการสื่อสาร มหาวิทยาลัยพะเยา พิจารณาแล้วเห็นว่าบุคลากรในสังกัดของท่าน คือ ผู้ช่วยศาสตราจารย์ ดร. ฆณการ ภัณณิพงส์ เป็นผู้มีความรู้ความสามารถในงานวิจัยดังกล่าวเป็นอย่างดี จึงขอความอนุเคราะห์บุคลากรในสังกัดของท่านเป็นที่ปรึกษาโครงการวิจัย ทั้งนี้ หากท่านมีข้อสงสัย สามารถติดต่อสอบถามได้ที่ ผู้ช่วยศาสตราจารย์ ดร.ปรัชญา นวนแก้ว เบอร์โทรติดต่อ ๐๘๙-๙๖๑-๔๘๓๒, ๐๘๙-๔๑๙-๗๑๒๙ อีเมลล์ pratya.nu@up.ac.th .

คณะเทคโนโลยีสารสนเทศและการสื่อสาร มหาวิทยาลัยพะเยา หวังเป็นอย่างยิ่งว่าคงได้รับความอนุเคราะห์จากท่านและขอขอบคุณมา ณ โอกาสนี้

จึงเรียนมาเพื่อโปรดพิจารณาให้ความอนุเคราะห์ จะขอบคุณยิ่ง

ขอแสดงความนับถือ

(ผู้ช่วยศาสตราจารย์ ดร.พรเทพ โรจนวสุ)
คณบดีคณะเทคโนโลยีสารสนเทศและการสื่อสาร

คณะเทคโนโลยีสารสนเทศและการสื่อสาร มหาวิทยาลัยพะเยา
โทรศัพท์ ๐-๕๔๔๖-๖๖๖๖ ต่อ ๒๓๒๐

ที่ อว ๗๓๑๙/ว๐๔๕๗

คณะเทคโนโลยีสารสนเทศและการสื่อสาร
มหาวิทยาลัยพะเยา ตำบลแม่กา
อำเภอเมือง จังหวัดพะเยา ๕๖๐๐๐

๑๗ ธันวาคม ๒๕๖๗

เรื่อง ขอความอนุเคราะห์บุคลากรในสังกัดของท่านเป็นที่ปรึกษาโครงการวิจัย

เรียน คณบดีคณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยราชภัฏมหาสารคาม

สิ่งที่ส่งมาด้วย แบบตอบรับที่ปรึกษาโครงการวิจัย จำนวน ๑ ฉบับ

ตามที่ มหาวิทยาลัยพะเยา ได้อนุมัติทุนอุดหนุนโครงการวิจัย ประจำปีงบประมาณ พ.ศ. ๒๕๖๘ ให้กับ ผู้ช่วยศาสตราจารย์ ดร. ปรัชญา นวนแก้ว สังกัด สาขาวิชาธุรกิจดิจิทัล คณะเทคโนโลยีสารสนเทศและการสื่อสาร เป็นหัวหน้าโครงการวิจัย "สารสนเทศเชิงประยุกต์เพื่อนวัตกรรมสารสนเทศทางการแพทย์" ตามคำสั่งมหาวิทยาลัยพะเยา เลขที่ ๕๐๒๘/๒๕๖๗ ลงวันที่ ๔ ตุลาคม ๒๕๖๗ ระยะเวลาดำเนินโครงการ ตั้งแต่วันที่ ๑ ตุลาคม ๒๕๖๗ – ๓๐ กันยายน ๒๕๖๘ นั้น

เพื่อให้การดำเนินโครงการวิจัยเป็นไปด้วยความเรียบร้อยและบรรลุตามวัตถุประสงค์ คณะเทคโนโลยีสารสนเทศและการสื่อสาร มหาวิทยาลัยพะเยา พิจารณาแล้วเห็นว่าบุคลากรในสังกัดของท่าน คือ ผู้ช่วยศาสตราจารย์ ดร. ทิพวิมล ชมภูคำ เป็นผู้มีความรู้ความสามารถในงานวิจัยดังกล่าวเป็นอย่างดี จึงขอความอนุเคราะห์บุคลากรในสังกัดของท่านเป็นที่ปรึกษาโครงการวิจัย ทั้งนี้ หากท่านมีข้อสงสัย สามารถติดต่อสอบถามได้ที่ ผู้ช่วยศาสตราจารย์ ดร.ปรัชญา นวนแก้ว เบอร์โทรติดต่อ ๐๘๙-๙๖๑-๔๘๓๒, ๐๘๙-๔๑๙-๗๑๒๙ อีเมลล์ pratya.nu@up.ac.th

คณะเทคโนโลยีสารสนเทศและการสื่อสาร มหาวิทยาลัยพะเยา หวังเป็นอย่างยิ่งว่าคงได้รับความอนุเคราะห์จากท่านและขอขอบคุณมา ณ โอกาสนี้

จึงเรียนมาเพื่อโปรดพิจารณาให้ความอนุเคราะห์ จะขอบคุณยิ่ง

ขอแสดงความนับถือ

(ผู้ช่วยศาสตราจารย์ ดร.พรเทพ โรจนวสุ)
คณบดีคณะเทคโนโลยีสารสนเทศและการสื่อสาร

คณะเทคโนโลยีสารสนเทศและการสื่อสาร มหาวิทยาลัยพะเยา
โทรศัพท์ ๐-๕๔๔๖-๖๖๖๖ ต่อ ๒๓๒๐

ที่ อว ๗๓๑๙/ว๐๔๕๗

คณะเทคโนโลยีสารสนเทศและการสื่อสาร
มหาวิทยาลัยพะเยา ตำบลแม่กา
อำเภอเมือง จังหวัดพะเยา ๕๖๐๐๐

๑๗ ธันวาคม ๒๕๖๗

เรื่อง ขอความอนุเคราะห์บุคลากรในสังกัดของท่านเป็นที่ปรึกษาโครงการวิจัย

เรียน คณบดีคณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏมหาสารคาม

สิ่งที่ส่งมาด้วย แบบตอบรับที่ปรึกษาโครงการวิจัย จำนวน ๑ ฉบับ

ตามที่ มหาวิทยาลัยพะเยา ได้อนุมัติทุนอุดหนุนโครงการวิจัย ประจำปีงบประมาณ พ.ศ. ๒๕๖๘ ให้กับ ผู้ช่วยศาสตราจารย์ ดร. ปรัชญา นวนแก้ว สังกัด สาขาวิชาธุรกิจดิจิทัล คณะเทคโนโลยีสารสนเทศและการสื่อสาร เป็นหัวหน้าโครงการวิจัย "สารสนเทศเชิงประยุกต์เพื่อนวัตกรรมสารสนเทศทางการแพทย์" ตามคำสั่งมหาวิทยาลัยพะเยา เลขที่ ๕๐๒๘/๒๕๖๗ ลงวันที่ ๔ ตุลาคม ๒๕๖๗ ระยะเวลาดำเนินโครงการ ตั้งแต่วันที่ ๑ ตุลาคม ๒๕๖๗ – ๓๐ กันยายน ๒๕๖๘ นั้น

เพื่อให้การดำเนินโครงการวิจัยเป็นไปด้วยความเรียบร้อยและบรรลุตามวัตถุประสงค์ คณะเทคโนโลยีสารสนเทศและการสื่อสาร มหาวิทยาลัยพะเยา พิจารณาแล้วเห็นว่าบุคลากรในสังกัดของท่าน คือ ดร. พัชระ นาเสงี่ยม เป็นผู้มีความรู้ความสามารถในงานวิจัยดังกล่าวเป็นอย่างดี จึงขอความอนุเคราะห์บุคลากรในสังกัดของท่านเป็นที่ปรึกษาโครงการวิจัย ทั้งนี้ หากท่านมีข้อสงสัยสามารถติดต่อสอบถามได้ที่ ผู้ช่วยศาสตราจารย์ ดร.ปรัชญา นวนแก้ว เบอร์โทรติดต่อ ๐๘๙-๙๖๑-๔๘๓๒, ๐๘๙-๔๑๙-๗๑๒๙ อีเมลล์ pratya.nu@up.ac.th

คณะเทคโนโลยีสารสนเทศและการสื่อสาร มหาวิทยาลัยพะเยา หวังเป็นอย่างยิ่งว่าคงได้รับความอนุเคราะห์จากท่านและขอขอบคุณมา ณ โอกาสนี้

จึงเรียนมาเพื่อโปรดพิจารณาให้ความอนุเคราะห์ จะขอบคุณยิ่ง

ขอแสดงความนับถือ

(ผู้ช่วยศาสตราจารย์ ดร.พรเทพ โรจนวสุ)
คณบดีคณะเทคโนโลยีสารสนเทศและการสื่อสาร

คณะเทคโนโลยีสารสนเทศและการสื่อสาร มหาวิทยาลัยพะเยา
โทรศัพท์ ๐-๕๔๔๖-๖๖๖๖ ต่อ ๒๓๒๐