การพัฒนาระบบการเทียบโอนผลการเรียนอัตโนมัติด้วยเทคโนโลยีโอซีอาร์ และการประมวลผลภาษาธรรมชาติ กรณีศึกษามหาวิทยาลัยธนบุรี

Development of an Automated Credit Transfer Prototype System Using OCR and Natural Language Processing Technologies: A Case Study of Thonburi University

ปฐมพงษ์ ฤกษ์สมุทร[1], จิรพันธุ์ ศรีสมพันธุ์[2], ธัญญรัตน์ น้อมพลกรัง[3]

[1]ภาควิชาคอมพิวเตอร์ศึกษา, มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, S6402042910029@email.kmutnb.ac.th

[2]ภาควิชาคอมพิวเตอร์ศึกษา, มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, jiraphan.s@fte.kmutnb.ac.th

[3]ภาควิชาคอมพิวเตอร์ศึกษา, มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, thanyarat.n@fte.kmutnb.ac.th

บทคัดย่อ

การเทียบโอนหน่วยกิต เป็นกลไกสำคัญในการส่งเสริมความคล่องตัวทางการศึกษา โดยเฉพาะในบริบทของการเชื่อมโยงระหว่างระดับประกาศนียบัตรวิชาชีพชั้นสูง (ปวส.) สู่ระดับปริญญาตรี อย่างไรก็ตาม กระบวนการตรวจสอบเอกสารระเบียนผลการเรียนในปัจจุบันยังคงดำเนินการด้วยมือ ซึ่งก่อให้เกิดความล่าช้าและเสี่ยงต่อ ความผิดพลาด งานวิจัยนี้จึงมีวัตถุประสงค์เพื่อพัฒนาระบบเว็บแอปพลิเคชันสำหรับการเทียบโอนหน่วยกิตอัตโนมัติ โดยการบูรณาการเทคโนโลยีการรู้จำอักขระด้วยแสงผ่าน Tesseract Engine เพื่อสกัดข้อมูล จากภาพถ่ายระเบียนผลการเรียน และประยุกต์ใช้เทคนิคการประมวลผลภาษาธรรมชาติ ในการวิเคราะห์ความคล้ายคลึงของรายวิชา ระบบที่พัฒนาขึ้นด้วยภาษา Python และเฟรมเวิร์ก Streamlit สามารถแสดงผล การวิเคราะห์และสถานะการอนุมัติเบื้องต้นได้อย่างมีประสิทธิภาพ ผลการทดสอบกับกลุ่มตัวอย่างระเบียนผลการเรียน จำนวน 100 ฉบับ ของนักศึกษาสาขาวิชาเทคโนโลยีดิจิทัลมีเดีย คณะวิทยาศาสตร์และเทคโนโลยี แสดงให้เห็นว่าระบบสามารถสกัดข้อมูลรหัสวิชาและเกรดเฉลี่ยจากระเบียนผลการเรียนได้อย่างถูกต้องแม่นยำ ร้อยละ 100 และสามารถวิเคราะห์ความคล้ายคลึงของรายวิชาเพื่อแนะนำผลการเทียบโอนได้อย่างมีประสิทธิภาพ โดยรายวิชาที่มีความคล้ายคลึงกันสูง เช่น วิชาเทคโนโลยีดิจิทัลเพื่อการจัดการ มีค่าความเหมือนร้อยละ 92.5 และระบบสามารถแนะนำให้เทียบโอนได้อย่างถูกต้อง ระบบนี้สามารถลดภาระงานของเจ้าหน้าที่ทะเบียนและเพิ่มมาตรฐานในการพิจารณาเทียบโอนได้ ซึ่งเป็นพื้นฐานสำคัญในการพัฒนาสู่ระบบที่มีความปลอดภัยสูง ด้วยเทคโนโลยีบล็อกเชนในอนาคต

คำหลัก: การเทียบโอนหน่วยกิต, การรู้จำอักขระด้วยแสง, การประมวลผลภาษาธรรมชาติ, ระบบอัตโนมัติ

Abstract

Credit transfer is a crucial process that facilitates student mobility, especially in assisting students as they transition from the high vocational certificate level to the bachelor's degree level. Currently, the verification of academic transcripts is mainly conducted manually, resulting in operational delays and a higher risk of human error. This study aims to develop a prototype for an automated web application designed to simplify the credit transfer process. The system utilizes the Tesseract Engine, which is part of optical character recognition (OCR) technology, to extract information from images of academic transcripts. Additionally, to assess the similarity between various courses, the application incorporates Natural Language Processing (NLP) techniques. Developed using Python and the Streamlit framework, the application effectively presents the analysis results alongside the initial approval statuses. Experimental findings from a sample of 100 academic transcripts of students in the Digital Media Technology program, Faculty of Science and Technology, show that the system achieves 100% accuracy in extracting course codes and grade point averages (GPA). Moreover, it accurately analyzes course similarity to recommend credit transfers; for example, courses with substantial content overlap, such as "Digital Technology for Management," received a similarity score of 92.5%, resulting in a valid recommendation for transfer approval. This system significantly decreases the workload of registrar staff and standardizes the credit transfer process, marking an important step toward the future development of a highly secure system based on blockchain technology.

Keywords: Credit Transfer, Optical Character Recognition (OCR), Natural Language Processing (NLP), Automation

ความเป็นมาและความสำคัญของปัญหา

ในยุคดิจิทัลสถาบันอุดมศึกษาจำเป็นต้องปรับเปลี่ยนกระบวนการทำงานเพื่อให้สามารถตอบสนองต่อการเปลี่ยนแปลงได้ดียิ่งขึ้น โดยเฉพาะ "การเทียบโอนหน่วยกิต" ซึ่งเป็นกระบวนการที่สำคัญมากที่เชื่อมโยงหลักสูตรระดับอนุปริญญาและปริญญาตรี รวมถึงเพิ่มความยืดหยุ่นทางการศึกษาให้กับนักศึกษา (Heppner et al., 2019) อย่างไรก็ตาม กระบวนการตรวจสอบเอกสารผลการเรียน (Transcript) ของมหาวิทยาลัยธนบุรียังคงดำเนินการตรวจสอบด้วยตนเอง (Manual Verification) ซึ่งเป็นขั้นตอนที่ใช้

เวลานานและมีความเสี่ยงต่อความผิดพลาดจากมนุษย์อันเนื่องมาจากข้อมูลจำนวนมากที่ต้องตรวจสอบ (Jayoma et al., 2020; Kamisetty & Jeyanthi, 2022) ส่งผลให้เกิดความล่าช้าและการทำงานซ้ำซ้อนในระหว่างช่วงเวลาการยื่นคำขอในอนาคต

ดังนั้น ผู้วิจัยจึงตระหนักว่าจำเป็นต้องบูรณาการเทคโนโลยีการรู้จำอักษรด้วยแสง (OCR) ซึ่งมีประสิทธิภาพสูงในการแปลงภาพเอกสารเป็นข้อความดิจิทัล (Kirana et al., 2025) โดยเน้นเป็นพิเศษที่การใช้ Tesseract Engine ในเอกสารภาษาไทยและภาษาอังกฤษ (Joshi & Arolkar, 2024) และปัญหาของอักษรไทยในเอกสาร ทางวิชาการ (Mapato et al., 2025) เข้ากับการประมวลผลภาษาธรรมชาติ (NLP) โดยมีจุดประสงค์เพื่อสร้างระบบอัตโนมัติที่สามารถวิเคราะห์ความคล้ายคลึงกันของรายวิชาได้อย่างแม่นยำ ช่วยลดภาระงานของเจ้าหน้าที่ และเพิ่มความเร็วและความสม่ำเสมอในการจัดการการลงทะเบียน (Heppner et al., 2019)

วัตถุประสงค์ของการวิจัย

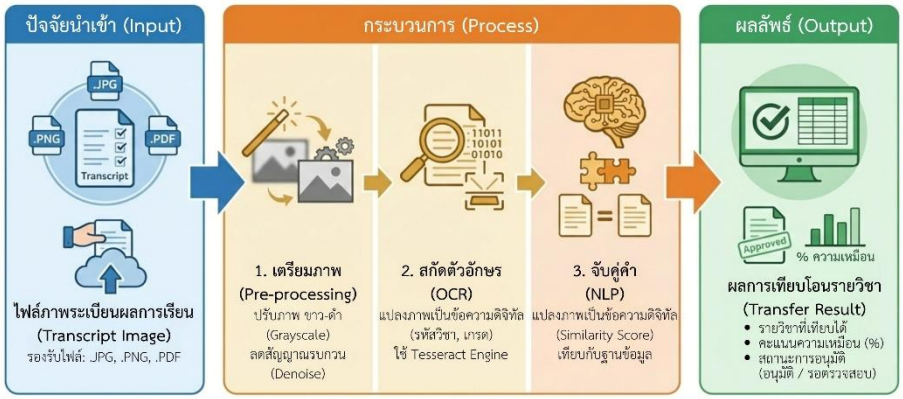
1. เพื่อพัฒนาระบบที่สามารถสกัดข้อมูลจากภาพถ่ายเอกสารผลการเรียนได้อย่างถูกต้อง
2. เพื่อประยุกต์ใช้อัลกอริทึมในการวิเคราะห์ความเหมือนของรายวิชาและแนะนำผลการเทียบโอน
3. เพื่อลดระยะเวลาและขั้นตอนการทำงานของเจ้าหน้าที่ในการบันทึกข้อมูลเข้าสู่ระบบ

ประโยชน์ที่คาดว่าจะได้รับ

1. ได้ระบบที่ช่วยลดความผิดพลาดและลดเวลาในการตรวจสอบเอกสารเทียบโอนหน่วยกิต
2. เจ้าหน้าที่ทะเบียนสามารถปฏิบัติงานได้อย่างมีประสิทธิภาพและมีมาตรฐานการเทียบเคียงรายวิชา
3. นักศึกษาได้รับผลการพิจารณาเทียบโอนที่รวดเร็วขึ้น ส่งเสริมความคล่องตัวทางการศึกษา
4. เป็นแนวทางพื้นฐานในการนำเทคโนโลยี AI มาประยุกต์ใช้ในงานทะเบียนและต่อยอดสู่ระบบบล็อกเชน

กรอบแนวคิด

การวิจัยนี้ใช้กระบวนการ Input-Process-Output โดยนำเข้าไฟล์ภาพระเบียนผลการเรียนผ่านกระบวนการประมวลผลภาพ (Pre-processing) การสกัดตัวอักษร (OCR) และการจับคู่คำ (NLP) เพื่อให้ได้ผลลัพธ์เป็นการเทียบโอนรายวิชา



ภาพ 1 กรอบแนวคิดระบบเทียบโอนหน่วยกิตอัตโนมัติด้วย OCR และ NLP

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

เทคโนโลยีการรู้จำอักขระด้วยแสง (OCR) เป็นเทคโนโลยีหลักในการแปลงเอกสารภาพ Mapato et al. (2025) ได้ศึกษาประสิทธิภาพของ OCR สำหรับเอกสารวิชาการภาษาไทย และพบว่าความท้าทายหลัก อยู่ที่วรรณยุกต์ลอยและสระซ้อน ในขณะที่ Joshi and Arolkar (2024) ยืนยันว่า Tesseract OCR เวอร์ชัน 5 ซึ่งใช้สถาปัตยกรรม LSTM (Long Short-Term Memory) มีประสิทธิภาพสูงในการอ่านเอกสาร หลายภาษา (Multilingual) และเป็นซอฟต์แวร์โอเพนซอร์สที่เหมาะสมสำหรับการพัฒนาระบบ

การประมวลผลภาษาธรรมชาติ (NLP) การเทียบโอนรายวิชาจำเป็นต้องอาศัยการเปรียบเทียบเนื้อหา Heppner et al. (2019) เสนอการใช้ NLP เพื่อวัดความคล้ายคลึงทางความหมาย (Semantic Similarity) ของคำอธิบายรายวิชา ซึ่งช่วยให้ระบบสามารถจับคู่รายวิชาที่มีชื่อเรียกต่างกันแต่มีเนื้อหาใกล้เคียงกันได้แม่นยำกว่าการเปรียบเทียบคำตรง (Keyword Matching)

วิธีดำเนินการวิจัย

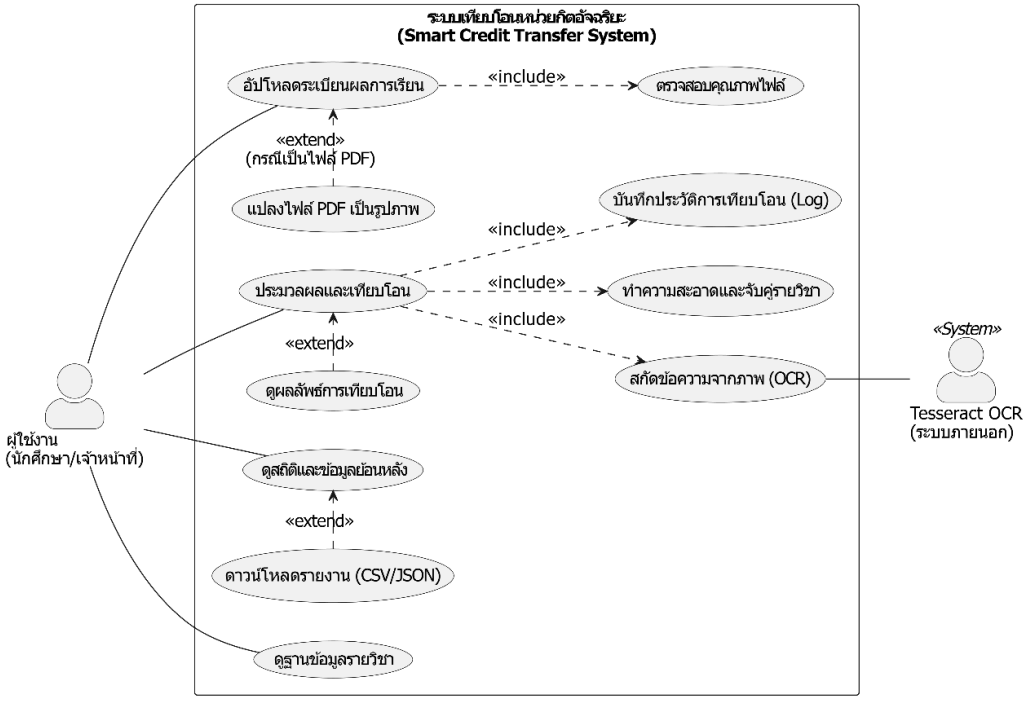
การวิจัยครั้งนี้ ผู้วิจัยได้กำหนดกลุ่มตัวอย่างสำหรับการทดสอบระบบ คือ ภาพระเบียนผลการเรียน (Transcript) ของนักศึกษาระดับประกาศนียบัตรวิชาชีพชั้นสูง (ปวส.) ที่ยื่นความจำนงขอเทียบโอนผลการเรียนเข้าสู่ระดับปริญญาตรี สาขาวิชาเทคโนโลยีดิจิทัลมีเดีย คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธนบุรี จำนวน 100 ฉบับ โดยใช้วิธีการเลือกกลุ่มตัวอย่างแบบเจาะจง เพื่อนำมาประเมินประสิทธิภาพความแม่นยำ ของการสกัดข้อความด้วยเทคโนโลยี OCR และทดสอบการจับคู่ความคล้ายคลึงของรายวิชาด้วยอัลกอริทึม NLP

ในการออกแบบโครงสร้างการทำงานของระบบ ผู้วิจัยได้กำหนดสถาปัตยกรรมแบบโมดูลาร์ (Modular Architecture) เพื่อให้ระบบมีความยืดหยุ่นและง่ายต่อการบำรุงรักษา โดยแบ่งการทำงานออกเป็น 2 ส่วนหลัก ได้แก่ ส่วนการนำเข้าและเตรียมความพร้อมข้อมูล (Input & Pre-processing)

ทำหน้าที่รับภาพระเบียนผลการเรียนผ่านส่วนต่อประสานผู้ใช้งานที่พัฒนาด้วย Streamlit และดำเนินการปรับปรุงคุณภาพภาพด้วยไลบรารี Pillow เพื่อเตรียมความพร้อมของข้อมูล และส่วนการประมวลผลอัจฉริยะและการตัดสินใจ (AI Processing & Decision) ซึ่งทำหน้าที่สกัดข้อความสำคัญด้วย Tesseract-OCR บนพื้นฐานโครงข่ายประสาทเทียมแบบ LSTM และนำข้อมูลที่ได้มาวิเคราะห์ความคล้ายคลึงผ่านโมดูลตรรกะการตัดสินใจ เพื่อประมวลผลสถานะการเทียบโอนว่าเป็น "อนุมัติ" หรือ "รอการตรวจสอบ" ดังแสดงรายละเอียดสถาปัตยกรรมในภาพ 2

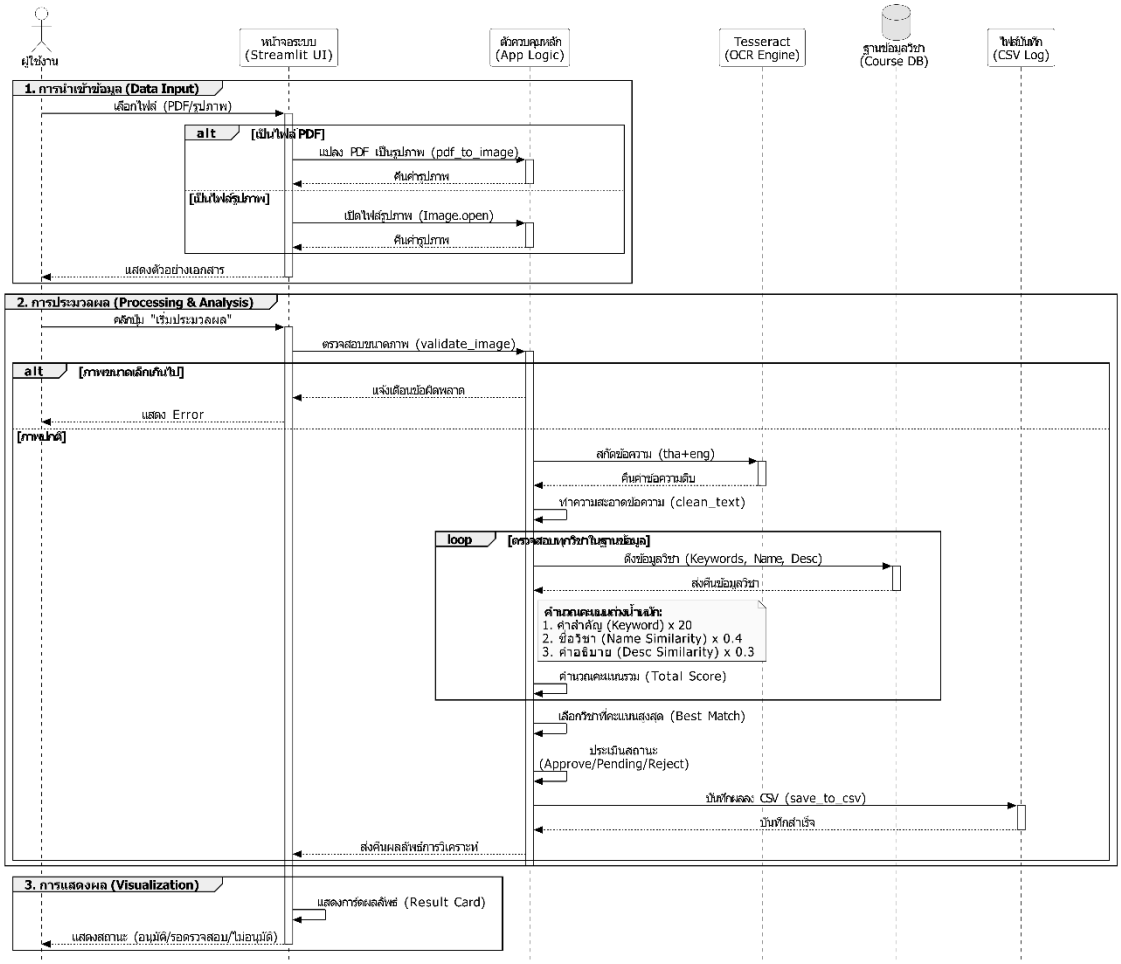
ภาพ 2 สถาปัตยกรรมของระบบ

เพื่อให้เห็นภาพรวมหน้าที่และการทำงานของระบบที่ชัดเจน ผู้วิจัยได้วิเคราะห์และออกแบบระบบ ผ่านแผนภาพยูสเคส (Use Case Diagram) ซึ่งแสดงให้เห็นถึงปฏิสัมพันธ์ระหว่างผู้ใช้งานกับระบบ โดยผู้ใช้งานสามารถเข้าถึงฟังก์ชันหลักต่าง ๆ ได้แก่ การอัปโหลดระเบียนผลการเรียน การดูผลลัพธ์การเทียบโอน การดูสถิติและข้อมูลย้อนหลัง และการดูฐานข้อมูลรายวิชา ทั้งนี้ระบบยังมีฟังก์ชันเบื้องหลังที่สำคัญ เช่น การแปลงไฟล์ PDF เป็นรูปภาพ การตรวจสอบคุณภาพไฟล์ การสกัดข้อความจากภาพ (OCR) การทำความสะอาดและจับคู่รายวิชา และการบันทึกประวัติการเทียบโอน (Log) ซึ่งทำงานร่วมกันเพื่อสนับสนุนกระบวนการเทียบโอนให้สมบูรณ์ ดังแสดงในภาพ 3



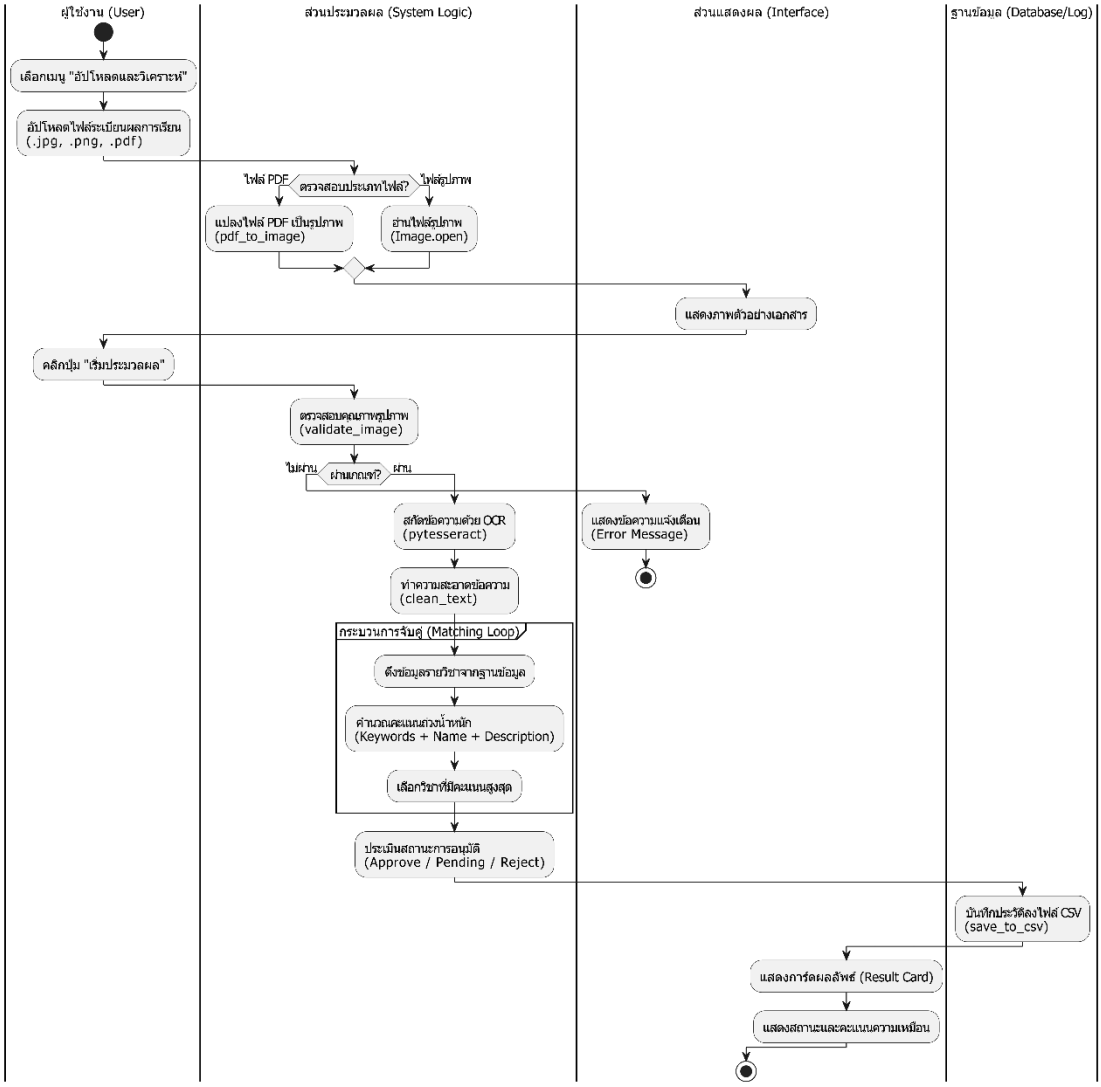
ภาพ 3 แผนภาพยูสเคสของระบบ

สำหรับการอธิบายขั้นตอนการทำงานภายในระบบอย่างละเอียด ผู้วิจัยได้จำลองลำดับการประมวลผลผ่านแผนภาพลำดับ (Sequence Diagram) เริ่มตั้งแต่ผู้ใช้งานทำการเลือกไฟล์ข้อมูล ซึ่งระบบจะตรวจสอบว่าเป็นไฟล์ PDF หรือไฟล์รูปภาพ หากเป็น PDF ระบบจะทำการแปลงไฟล์เป็นรูปภาพก่อน จากนั้นเมื่อผู้ใช้งาน สั่งประมวลผล ระบบจะตรวจสอบขนาดภาพ หากภาพผ่านเกณฑ์จะเข้าสู่กระบวนการสกัดข้อความด้วย OCR การทำความสะอาดข้อความ และเข้าสู่กระบวนการตรวจสอบรายวิชาในฐานข้อมูล โดยมีการคำนวณคะแนน แบบถ่วงน้ำหนักจากคำสำคัญ ชื่อวิชา และคำอธิบายรายวิชา เพื่อเลือกวิชาที่มีคะแนนสูงสุดและประเมินสถานะการเทียบโอน ก่อนจะบันทึกผลลงในไฟล์ CSV และส่งคืนผลลัพธ์กลับมาแสดงผลให้ผู้ใช้งานทราบ ดังแสดงขั้นตอนในภาพ 4



ภาพ 4 แผนภาพลำดับการทำงานของระบบ

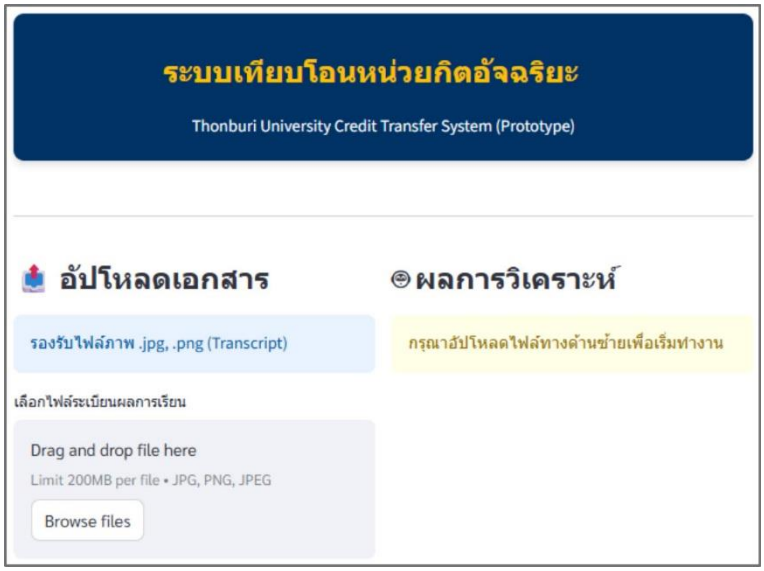
นอกจากนี้ เพื่อให้เห็นภาพรวมของกระบวนการทำงานตั้งแต่ต้นจนจบอย่างเป็นระบบ ผู้วิจัยได้จัดทำแผนภาพกิจกรรม (Workflow Diagram) แสดงเส้นทางการทำงานของผู้ใช้งาน ส่วนประมวลผล ส่วนแสดงผล และฐานข้อมูล โดยเริ่มจากการเลือกเมนูและอัปโหลดไฟล์ การตรวจสอบประเภทและคุณภาพไฟล์ การสกัดและทำความสะอาดข้อความ การเข้าสู่กระบวนการจับคู่รายวิชา การประเมินสถานะการอนุมัติ การบันทึกประวัติ ลงไฟล์ และสิ้นสุดที่การแสดงผลลัพธ์การ์ดผลลัพธ์พร้อมสถานะและคะแนนความเหมือนแก่ผู้ใช้งาน ซึ่งแสดงถึงความเชื่อมโยงของทุกส่วนงานอย่างชัดเจน ดังแสดงในภาพ 5



ภาพ 5 แผนภาพกระบวนการทำงานของระบบ

ผลการดำเนินงาน

จากการพัฒนาและทดสอบระบบการเทียบโอนผลการเรียนอัตโนมัติ ผู้วิจัยได้ดำเนินการประเมินประสิทธิภาพของระบบใน 3 ด้านหลัก ได้แก่ ส่วนติดต่อผู้ใช้ ความถูกต้องของการสกัดข้อมูล และผลลัพธ์การเทียบโอนรายวิชา โดยเริ่มจากการพัฒนาส่วนติดต่อผู้ใช้ให้มีความทันสมัยและใช้งานง่ายตามอัตลักษณ์ของมหาวิทยาลัยธนบุรี ซึ่งผู้ใช้งานสามารถอัปโหลดไฟล์ภาพระเบียนผลการเรียน ผ่านระบบลากและวางเพื่อเข้าสู่กระบวนการประมวลผลได้ทันที ดังแสดงในภาพ 6



ภาพ 6 หน้าจอหลักของระบบเทียบโอนหน่วยกิตอัจฉริยะ

ในส่วนของการทดสอบความถูกต้องของการสกัดข้อมูล ผู้วิจัยได้นำภาพระเบียนผลการเรียนของกลุ่มตัวอย่าง จำนวน 100 ฉบับ จากสาขาวิชาเทคโนโลยีดิจิทัลมีเดีย เข้าสู่กระบวนการสกัดข้อความและเปรียบเทียบผลลัพธ์กับข้อมูลจริงบนเอกสาร พบว่าระบบมีความแม่นยำสูงมาก ในการอ่านข้อมูลประเภทตัวเลข เช่น รหัสประจำตัว รหัสวิชา และเกรดเฉลี่ย โดยมีความถูกต้องถึง 100% เนื่องจากรูปแบบตัวอักษรมีความชัดเจน ในขณะที่ข้อมูลส่วนหัวเอกสารและชื่อรายวิชาภาษาไทยยังพบ ความคลาดเคลื่อนบ้างในกรณีที่ฟอนต์มีรูปแบบพิเศษหรือไม่เป็นมาตรฐาน ส่งผลให้อ่านเป็นขยะอักขระในบางส่วน ดังรายละเอียดในตาราง 1

ตาราง 1 การเปรียบเทียบข้อมูลดิบจาก OCR กับข้อมูลจริงบนเอกสาร

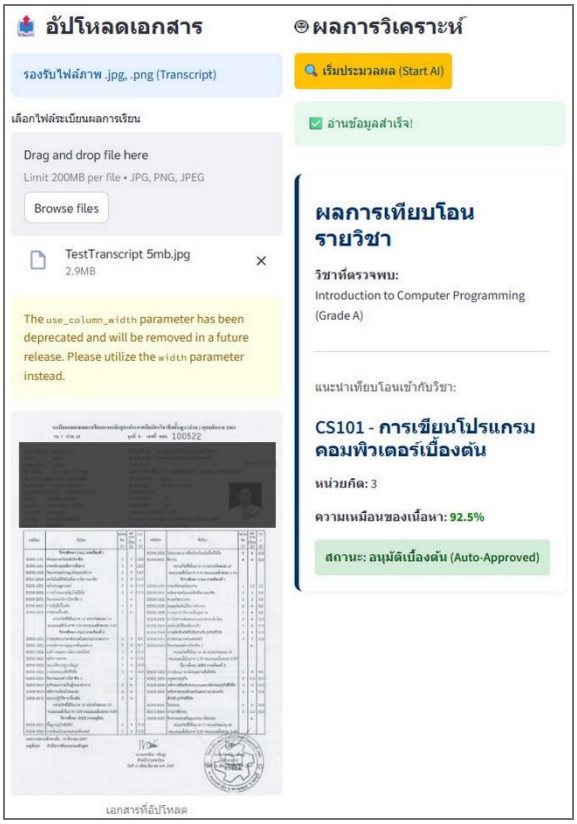
ส่วนของเอกสาร	ข้อมูลดิบที่ระบบอ่านได้	การวิเคราะห์ผล
ส่วนหัวเอกสาร	siDsuMansnamasanmunanq assemansiass...	**ความแม่นยำต่ำ** เกิดความผิดพลาดเนื่องจากฟอนต์หัวเรื่องมีความหนาและรูปแบบพิเศษ (Decorative Font) ทำให้ระบบอ่านเป็นขยะอักขระ (Garbage Text)
รหัสประจำตัว	1320100114	**ความแม่นยำสูง (100%)** ตัวเลขมีความชัดเจน ระบบสามารถระบุแพทเทิร์นได้ถูกต้อง
รหัสวิชา	30000-1101, 30204-2006, 30001-2003	**ความแม่นยำสูง (100%) ระบบสามารถจับแพทเทิร์น** xxxxx-xxxx ได้อย่างแม่นยำแม้จะมีสัญญาณรบกวนรอบข้าง
เกรดเฉลี่ย	4.00, 3.88, 3.90	**ความแม่นยำสูง** จุดทศนิยมและตัวเลขชัดเจน
ชื่อรายวิชา	Anwisangquriionisaea1s	**ความแม่นยำปานกลาง-ต่ำ** การอ่านภาษาไทยในเอกสารเก่าหรือฟอนต์ที่ไม่ใช่มาตรฐานยังมีความคลาดเคลื่อนสูง

สำหรับผลการวิเคราะห์และเทียบโอนรายวิชา ระบบใช้อัลกอริทึมการประมวลผลภาษาธรรมชาติ (NLP) เข้ามาช่วยวิเคราะห์ความคล้ายคลึงของเนื้อหารายวิชาเพื่อชดเชยความคลาดเคลื่อนจากการอ่านชื่อวิชาภาษาไทย โดยผลการทดสอบแสดงให้เห็นว่าระบบสามารถจับคู่รายวิชาต้นทางกับวิชาเป้าหมายได้อย่างมีประสิทธิภาพ พร้อมคำนวณค่าความเหมือน และระบุสถานะการอนุมัติได้อย่างถูกต้อง เช่น วิชาเทคโนโลยีดิจิทัลเพื่อการจัดการ มีความเหมือน 92.5% ได้รับสถานะแนะนำให้เทียบโอน ในขณะที่วิชาภาษาอังกฤษเพื่อการสื่อสาร มีความเหมือน 78.2% ซึ่งต่ำกว่าเกณฑ์ จึงได้รับสถานะรอการตรวจสอบ ดังแสดงตัวอย่างผลลัพธ์ในตาราง 2

ตาราง 2 ผลลัพธ์การเทียบโอนรายวิชาหลังผ่านกระบวนการประมวลผล

รหัสวิชาต้นทาง	วิชาเป้าหมาย	หน่วยกิต	ความเหมือน	สถานะการอนุมัติ
30001-2003	CS103 เทคโนโลยีดิจิทัลเพื่อการจัดการ	3	92.5%	แนะนำให้เทียบโอน
30204-2001	CS101 ระบบปฏิบัติการเบื้องต้น	3	85.0%	แนะนำให้เทียบโอน
30000-1201	GE101 ภาษาอังกฤษเพื่อการสื่อสาร	3	78.2%	รอการตรวจสอบ

สุดท้ายเมื่อกระบวนการวิเคราะห์เสร็จสิ้น ระบบจะแสดงผลลัพธ์ในรูปแบบการ์ดข้อมูลที่เข้าใจง่าย โดยระบุชื่อวิชาที่ตรวจพบ เกรดที่ได้ วิชาที่แนะนำให้เทียบโอน หน่วยกิต ค่าความเหมือน และสถานะการอนุมัติเบื้องต้น เพื่อให้เจ้าหน้าที่หรือนักศึกษาสามารถตรวจสอบผลการพิจารณาได้อย่างรวดเร็วและสะดวกสบาย ดังแสดงตัวอย่างหน้าจอผลลัพธ์ในภาพ 7



ภาพ 7 หน้าจอแสดงผลลัพธ์การวิเคราะห์และสถานะการอนุมัติ

ในด้านการประเมินระยะเวลาการปฏิบัติงานเพื่อตอบสนองวัตถุประสงค์ในการลดขั้นตอนการทำงาน ผู้วิจัยได้ทำการเปรียบเทียบระยะเวลาที่ใช้ในการพิจารณาเทียบโอนต่อ 1 กรณีศึกษา พบว่า กระบวนการตรวจสอบและเทียบโอนด้วยมือแบบเดิม เจ้าหน้าที่ต้องใช้เวลาในการตรวจสอบเอกสารและค้นหาเทียบเคียงรายวิชาเฉลี่ยประมาณ 15 นาทีต่อ 1 ระเบียนผลการเรียน ในขณะที่ระบบเทียบโอนอัตโนมัติที่พัฒนาขึ้น ใช้เวลาในกระบวนการอัปโหลดไฟล์ สกัดข้อความ และประมวลผลจับคู่รายวิชาด้วย NLP จนถึงขั้นตอนแสดงสถานะการอนุมัติเบื้องต้น เฉลี่ยเพียง 1-2 นาทีต่อ 1 ระเบียนผลการเรียน ผลการทดสอบเชิงปริมาณนี้แสดงให้เห็นว่าระบบสามารถลดระยะเวลาการทำงานของเจ้าหน้าที่ลงได้กว่า ร้อยละ 86 ส่งผลให้นักศึกษาได้รับทราบผลการพิจารณาเบื้องต้นที่รวดเร็วขึ้นอย่างมาก

สรุปผลการวิจัย

การพัฒนาระบบการเทียบโอนผลการเรียนอัตโนมัติด้วยเทคโนโลยีโอซีอาร์และการประมวลผลภาษาธรรมชาติ กรณีศึกษามหาวิทยาลัยธนบุรี ประสบความสำเร็จตามวัตถุประสงค์ที่ตั้งไว้ โดยระบบที่พัฒนาขึ้นในรูปแบบเว็บแอปพลิเคชันด้วยภาษา Python และเฟรมเวิร์ก Streamlit สามารถทำงานได้จริงบนสภาพแวดล้อมเครือข่ายอินเทอร์เน็ต มีความสามารถในการนำเข้าไฟล์ภาพระเบียนผลการเรียน และผ่านกระบวนการเตรียมภาพ เพื่อให้ Tesseract OCR ทำการสกัดข้อมูลตัวอักษร จากผลการทดสอบกับ

กลุ่มตัวอย่างระเบียนผลการเรียน จำนวน 100 ฉบับ ของนักศึกษาสาขาวิชาเทคโนโลยีดิจิทัลมีเดีย คณะวิทยาศาสตร์และเทคโนโลยี พบว่าระบบมีประสิทธิภาพสูงมากในการอ่านข้อมูลโครงสร้างหลักที่เป็นภาษาอังกฤษและตัวเลข ได้แก่ รหัสวิชา และเกรดเฉลี่ย โดยมีความถูกต้องถึงร้อยละ 100 ในขณะที่การอ่านชื่อรายวิชาภาษาไทยยังมีความคลาดเคลื่อนในส่วนของวรรณยุกต์และสระลอย อย่างไรก็ตาม เมื่อนำข้อมูลที่สกัดได้ไปผ่านกระบวนการวิเคราะห์ความคล้ายคลึง ด้วยอัลกอริทึมการประมวลผลภาษาธรรมชาติ (NLP) พบว่าระบบสามารถจับคู่รายวิชาต้นทางกับรายวิชาเป้าหมายได้อย่างแม่นยำ และสามารถแสดงสถานะการอนุมัติเบื้องต้นได้ถูกต้องตามเกณฑ์ที่กำหนด ช่วยลดระยะเวลาและขั้นตอนการทำงานของเจ้าหน้าที่ทะเบียนจากการตรวจสอบด้วยมือลงได้อย่างมีนัยสำคัญ โดยสามารถลดระยะเวลาการประมวลผลจากเดิมที่ใช้เวลาเฉลี่ย 15 นาทีต่อเอกสาร เหลือเพียง 1-2 นาทีต่อเอกสาร หรือลดลงกว่าร้อยละ 86 ซึ่งช่วยเพิ่มประสิทธิภาพในการให้บริการนักศึกษาได้ตามวัตถุประสงค์ที่ตั้งไว้อย่างเป็นรูปธรรม

อภิปรายผล

จากผลการวิจัย สามารถอภิปรายผลได้ว่า ประสิทธิภาพของเทคโนโลยี Tesseract OCR เวอร์ชัน 5 ซึ่งใช้สถาปัตยกรรมโครงข่ายประสาทเทียมแบบ LSTM มีความแม่นยำสูงในการจัดการกับเอกสารภาษาอังกฤษและตัวเลข ซึ่งสอดคล้องกับงานวิจัยของ Joshi & Arolkar (2024) ที่ระบุว่า Tesseract OCR มีประสิทธิภาพสูงในการอ่านเอกสารหลายภาษาและเหมาะสมสำหรับการพัฒนาระบบ แต่ในบริบทของเอกสารภาษาไทย ระบบยังพบข้อจำกัดในการอ่านอักขระที่มีความซับซ้อน เช่น สระซ้อนและวรรณยุกต์ลอย ซึ่งสอดคล้องกับผลการศึกษาของ Mapato et al. (2025) ที่พบความท้าทายในลักษณะเดียวกัน อย่างไรก็ตาม งานวิจัยนี้ได้พิสูจน์ให้เห็นว่าการบูรณาการเทคนิคการประมวลผลภาษาธรรมชาติ (NLP) เข้ามาช่วยในการวัดความคล้ายคลึงทางความหมาย (Semantic Similarity) สามารถชดเชยข้อผิดพลาดจากการอ่านตัวอักษรของ OCR ได้อย่างมีประสิทธิภาพ ส่งผลให้ระบบสามารถจับคู่รายวิชาได้ถูกต้องแม้ข้อมูลนำเข้าจะไม่สมบูรณ์ ซึ่งสอดคล้องกับแนวคิดของ Heppner et al. (2019) ที่เสนอให้ใช้ NLP ในกระบวนการเทียบโอนหน่วยกิตเพื่อแก้ปัญหาความล่าช้าและลดภาระงานของบุคลากร ดังที่ Jayoma et al. (2020) ได้กล่าวถึงปัญหาคอขวดในกระบวนการจัดการเอกสารด้วยมือ

ข้อเสนอแนะ

จากการวิจัยและพัฒนาระบบ ผู้วิจัยมีข้อเสนอแนะเพื่อการนำไปใช้และการพัฒนาต่อยอดในอนาคต ดังนี้

- ข้อเสนอแนะในการนำผลการวิจัยไปใช้ ผู้ใช้งานควรตรวจสอบคุณภาพของไฟล์ภาพต้นฉบับให้มีความคมชัดและปราศจากเงามืดเพื่อให้ OCR ทำงานได้เต็มประสิทธิภาพ และแม้ระบบจะมีความ

แม่นยำในการคัดกรองเบื้องต้น แต่ในขั้นตอนการอนุมัติผลอย่างเป็นทางการ ควรให้เจ้าหน้าที่ผู้เชี่ยวชาญเป็นผู้ตรวจสอบและยืนยันความถูกต้องครั้งสุดท้าย ตามระเบียบข้อบังคับของมหาวิทยาลัย

- ข้อเสนอแนะสำหรับการวิจัยครั้งต่อไป ควรมีการศึกษาและพัฒนาต่อยอดโดยการนำเทคโนโลยีบล็อกเชน โดยเฉพาะ Hyperledger Fabric มาประยุกต์ใช้ในการจัดเก็บและยืนยันผลการเทียบโอน เพื่อป้องกันการปลอมแปลงเอกสารและสร้างความโปร่งใส นอกจากนี้ ควรทดลองใช้โมเดล Deep Learning ขั้นสูงสำหรับ การประมวลผลเอกสาร เพื่อเพิ่มขีดความสามารถในการอ่านเอกสารที่มีรูปแบบซับซ้อนและลดความคลาดเคลื่อนในการอ่านภาษาไทยให้ดียิ่งขึ้น

เอกสารอ้างอิง

Heppner, A., et al. (2019). Automating Articulation: Applying Natural Language Processing to Post-Secondary Credit Transfer. *IEEE Access,* 7.

Jayoma, J., et al. (2020). OCR Based Document Archiving and Indexing Using PyTesseract: A Record Management System for DSWD Caraga, Philippines. *2020 IEEE 12th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM).*

Joshi, K., & Arolkar, H. (2024). Comparative Analysis of Outcomes of Tesseract OCR for Different Languages. *2024 5th International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV).*

Kamisetty, V. N. S. R., & Jeyanthi, P. (2022). Digitization of Data from Invoice using OCR. *2022 6th International Conference on Computing Methodologies and Communication (ICCMC).*

Kirana, K. C., et al. (2025). Comparison of Tesseract OCR, Easy OCR, and Transformer OCR on Handwritten Image. *2025 9th International Conference on Electrical, Electronics and Information Engineering (ICEEIE).*

Mapato, M., Sornlertlamvanich, M., & Charoenporn, T. (2025). Evaluating OCR-Based Retrieval Performance in Thai Academic Literature. *2025 17th International Conference on Human System Interaction (HSI).*