บทความวิชาการ (Academic Article)

การพยากรณ์ผลสัมฤทธิ์ทางการเรียนของผู้เรียนชั้นประถมศึกษา โดยใช้อัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน

Prediction of Achievement in Primary School Students Using Support Vector Machine Algorithm

จุฑามาศ ตันยะลา[1*] ชัยมงคล ปินะสา[1] และ อนุภูมิ คำยัง[1]

Jutamart Tanyala[1*], Chaimongkhon Pinasa[1] and Anuphum Kumyoung[1]

[1] สาขาวิจัยนวัตกรรมทางการศึกษา คณะครุศาสตร์ มหาวิทยาลัยราชภัฏเลย จังหวัดเลย 42000

[1] Educational Innovation Research, Faculty of Education, Loei Rajabhat University, Loei 42000

* Corresponding author: sm677911101@lru.ac.th

Received: December 22, 2024; **Revised:** February 14, 2025; **Accepted:** March 13, 2025

บทคัดย่อ

บทความทางวิชาการนี้มุ่งเน้นการใช้อัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีนพยากรณ์ผลสัมฤทธิ์ทาง การเรียนของนักเรียนชั้นประถมศึกษาปีที่ 6 ปีการศึกษา 2567 โดยใช้เทคโนโลยีปัญญาประดิษฐ์ ChatGPT เพื่อจำลองชุดข้อมูลของนักเรียน จำนวน 1,000 ชุด โดยอ้างอิงข้อมูลจริงจากโรงเรียน 3 แห่งในจังหวัดเลย ข้อมูลที่ศึกษา ได้แก่ เพศ, เกรดเฉลี่ย, เกรดเฉลี่ยสะสม, จำนวนของชั่วโมงที่เรียน, และสถานะทางการเรียน จากนั้นนำข้อมูลเข้าสู่อัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีนโดยแบ่งข้อมูลเป็น 2 กลุ่ม คือ กลุ่มฝึกสอนโมเดล 70% และกลุ่มทดสอบโมเดล 30% เพื่อประเมินประสิทธิภาพของโมเดล ซึ่งขั้นตอนการสร้างโมเดลมี 5 ขั้นตอน ดังนี้ 1) ขั้นรวบรวมข้อมูล 2) ขั้นเตรียมข้อมูล 3) ขั้นสร้างโมเดล 4) ขั้นทดสอบและประเมินประสิทธิภาพของโมเดลด้วยวิธีการ 10–Fold cross–validation และ 5) ขั้นการนำแบบโมเดลไปใช้งานจริงโดยผสานร่วมกับโปรแกรม Altair AI Studio Educational 2025.0.0 ผลการศึกษาพบว่า โมเดลมีประสิทธิภาพสูงทั้งในด้านของค่าความถูกต้อง ความแม่นยำ และค่าความระลึก ทั้งนี้ได้ผสานการทำงานเข้ากับแพลตฟอร์ม One Compiler ซึ่งเป็นสภาพแวดล้อมออนไลน์สำหรับการรันโค้ดและออกแบบส่วนติดต่อผู้ใช้ด้วยภาษา HTML ซึ่งสามารถป้อนข้อมูลนักเรียนในการพยากรณ์ผลสัมฤทธิ์ทางการเรียนล่วงหน้าได้อย่างสะดวก ทั้งนี้เครื่องมือนี้จึงอำนวยความสะดวกแก่ผู้บริหารสถานศึกษา ครู และผู้ที่สนใจ ในการพัฒนาการจัดการเรียนการสอนให้มีประสิทธิภาพมากยิ่งขึ้น

คำสำคัญ: การพยากรณ์ ผลสัมฤทธิ์ทางการเรียน แพลตฟอร์ม อัลกอริทึม ซัพพอร์ตเวกเตอร์แมชชีน

Abstract

This academic article focuses on the application of the Support Vector Machine (SVM) algorithm to predict the academic performance of Grade 6 students in the 2024 academic year. Artificial intelligence technology, specifically ChatGPT, was employed to simulate a dataset of 1,000 student records, based on real data collected from three schools in Loei Province, Thailand. The dataset included variables such as gender, grade point average (GPA), cumulative GPA, study hours, and academic status, while personal identifiers were excluded to ensure privacy. The data were divided into two sets: 70% for model training and 30% for testing to evaluate the model's performance. The model development process consisted of five main stages: (1) data collection, (2) data preparation,

(3) model construction, (4) performance evaluation using 10-fold cross-validation, and (5) implementation using Altair AI Studio Educational 2025.0.0. The results revealed that the SVM model demonstrated high levels of performance in terms of accuracy, precision, and recall, confirming its potential as a reliable tool for academic performance prediction. Additionally, the model was integrated into the One Compiler application, an online coding platform that supports HTML-based user interface design. This enabled users to input student data and conveniently receive predictive outcomes. The tool serves as a practical resource for school administrators, teachers, and stakeholders to enhance decision-making and improve the effectiveness of teaching and learning management.

Keywords: Prediction, Achievement, Platform, Algorithm, Support vector machine

บทนำ

การจัดการศึกษา ทั้งการศึกษาในระบบ การศึกษานอกระบบ และการศึกษาตามอัธยาศัยต้องเน้นความสำคัญทั้งความรู้คุณธรรม กระบวนการเรียนรู้และบูรณาการตามความเหมาะสมของ แต่ละระดับการศึกษา (สำนักนายกรัฐมนตรี, 2545) การศึกษาแต่ละช่วงวัยในหลักสูตรได้กำหนดตัวชี้วัด และการประเมินผลสัมฤทธิ์ทางการเรียนของผู้เรียน ซึ่งเป็นผลลัพธ์ที่สำคัญที่บ่งชี้ถึงประสิทธิภาพของบุคคล และเป็นปัจจัยสำคัญที่ช่วยให้ครูผู้สอน ผู้บริหาร และนักวิชาการศึกษาสามารถวิเคราะห์แนวโน้มการเรียนรู้และพัฒนากลยุทธ์การสอนได้อย่างมีประสิทธิภาพ การพยากรณ์ผลการเรียนของผู้เรียนมีความสำคัญเป็นอย่างมากต่อสถาบันการศึกษา รัฐบาล ผู้ปกครอง และผู้มีส่วนได้ส่วนเสีย เพื่อนำมาช่วยส่งเสริมการพัฒนาหลักสูตร (Rastrollo-Guerrero, Gómez-Pulido & Durán-Domínguez, 2020)

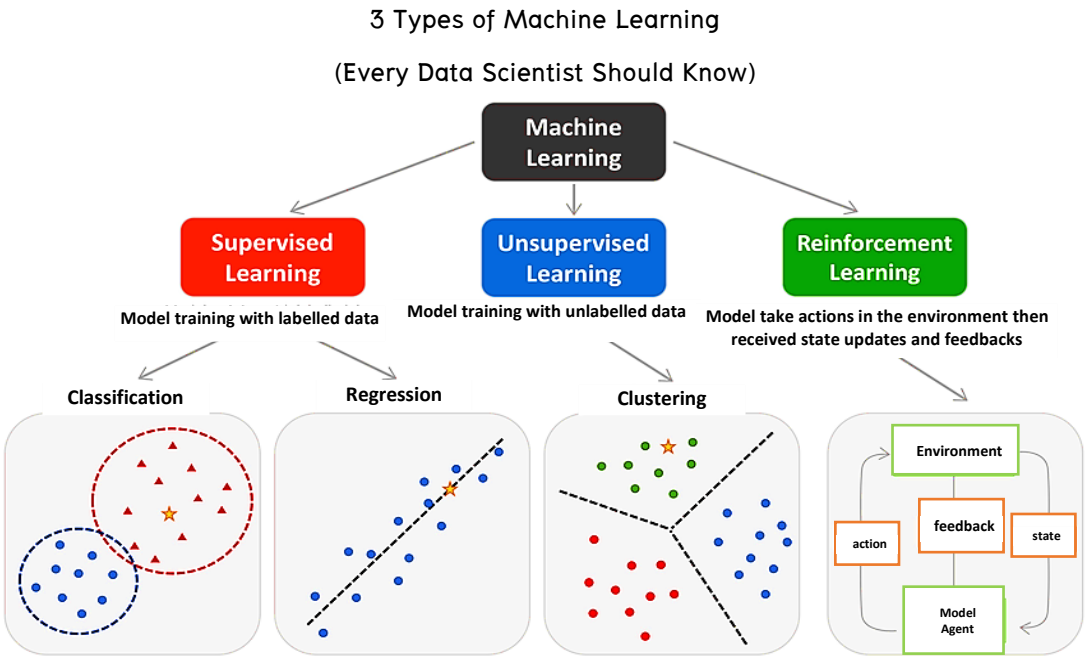
การทำเหมืองข้อมูลหรือดาต้าไมน์นิง (Education Data Mining) คือ วิธีการค้นพบ และการกลั่นกรองความรู้จากผลลัพธ์การวิเคราะห์โอกาสหรือความเป็นไปได้ในทางใดทางหนึ่งจากฐานข้อมูลที่มีขนาดใหญ่ (Larose & Larose, 2014; Peanpedlerd, 2023; Ramageri, 2010) โดยประโยชน์ของการทำเหมืองข้อมูลทางการศึกษาเพื่อนำมาใช้หาความสัมพันธ์ของชุดข้อมูลนั้น ๆ ได้แก่ การตัดสินใจ การวางแผน การทำนายหรือพยากรณ์แนวโน้มกับสิ่งที่จะเกิดขึ้นในอนาคตจากการจำแนก (Classification) เป็นกระบวนการนำเข้าของข้อมูลมาสร้างเป็นโมเดลแล้ววิเคราะห์ข้อมูล และส่งออกเป็นข้อมูลใหม่เพื่อทำนายอนาคต โดยอัลกอริทึมที่นิยมนำมาใช้วิเคราะห์ข้อมูล เช่น ต้นไม้ตัดสินใจ (Decision Tree) แบบจำลองเบย์สอย่างง่าย (Naïve Bayesian) โครงข่ายประสาทเทียม (Artificial Neural Networks: ANN) และซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine: SVM) (นคร ละลอกน้ำ, 2561) สอดคล้องแนวคิดของ Alamri, Almuslim, Alotibi, Alkadi, UllahKhan, Aslam (2020) ที่ได้ทำนายผลการเรียนของนักศึกษาด้วยชุดข้อมูล 369 คน และ 649 คน ตามลำดับ โดยใช้อัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน และป่าสุ่ม (Random Forest: RF) เพื่อหาแนวทางแก้ไขปัญหาของนักศึกษาที่มีผลการเรียนต่ำโดยการจำแนกประเภทแบบไบนารีและการถดถอยไปใช้กับชุดข้อมูลทั้งสองชุด พบว่า ความแม่นยำในกรณีของการจำแนกแบบไบนารีให้ผลทำนายที่แม่นยำถึง 93% ขณะที่ค่าความคลาดเคลื่อนในการทำนายอยู่ที่ 1.13

ด้วยการประยุกต์ใช้เทคนิคเหมืองข้อมูลทางการศึกษา ถือเป็นเครื่องมือสำคัญที่สถาบันหรือโรงเรียนนำมาใช้ยกระดับคุณภาพการเรียนการสอน (Bisri, Heryatun & Navira, 2025) ปัจจุบันการเรียนรู้ของเครื่อง (Machine Learning) โดยใช้อัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีนได้ถูกนำมาใช้ในการจำแนกนักเรียนที่มีผลการเรียนต่ำกว่าเกณฑ์ ซึ่งเป็นกลุ่มเสี่ยงที่อาจสำเร็จการศึกษาล่าช้า บทความทางวิชาการนี้มุ่งนำเสนอการสร้างโมเดลในการพยากรณ์ผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้นประถมศึกษาปีที่ 6 เพื่อนำข้อมูลที่ได้ไปใช้ในการวางแผน พัฒนา และปรับปรุงคุณภาพการศึกษาอย่างเหมาะสม

แนวคิด และทฤษฎีที่เกี่ยวข้อง

1. เหมืองข้อมูลทางการศึกษา (Education Data Mining)

ปัจจุบันนี้การนำเหมืองข้อมูลที่เกี่ยวข้องกับความสำเร็จในด้านการพัฒนาธุรกิจอีคอมเมิร์ซ และการปรับประยุกต์ใช้ด้านการศึกษาเพิ่มจำนวนมากขึ้น การทำเหมืองข้อมูลด้านการศึกษาจะซ่อนอยู่ภายใต้ชุดข้อมูลที่มีขนาดใหญ่ ช่วยให้การตัดสินใจมีประสิทธิภาพมากยิ่งขึ้น ด้วยการวิเคราะห์ข้อมูลที่แตกต่างกันไป เช่น การทำนาย และการจำแนกประเภท (Adekitan & Salau, 2019; Batool, Rashid, Nisar, Kim, Kwon & Hussain, 2023) ประโยชน์ที่ได้จากการทำเหมืองข้อมูลโดยใช้อัลกอริทึมเหล่านี้ถูกนำมาเพื่อออกแบบสร้างโมเดลในการทำนายผลการเรียนรู้ของเครื่อง (Machine Learning Models : ML) ซึ่งประเภทของเครื่องจักรอัลกอริทึมที่นำมาใช้ในการทำเหมืองข้อมูลทางการศึกษาได้แบ่งออกเป็น 3 ประเภท ได้แก่ 1) ประเภทการเรียนรู้แบบมีผู้สอน (Supervised Learning) เพื่อทำนายผลลัพธ์เป็นหมวดหมู่ 2) ประเภทการเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) เพื่อประมวลผลและเรียนรู้โครงสร้างของข้อมูลฝึกอบรมที่แฝงอยู่ภายใน และ 3) ประเภทการเรียนรู้แบบเสริมแรง (Reinforcement Learning) เป็นการปรับตัวตามสภาพแวดล้อมทั้งเชิงบวกและเชิงลบล้วนส่งผลต่อประสิทธิภาพการทำงานของโมเดล Peng, Jury, Dönnes & Ciurtin (2021) ปัจจุบันมีสถาบันการศึกษาจำนวนมากได้นำชุดข้อมูลจากเหมืองข้อมูลทางการศึกษาไปวิเคราะห์หาสาเหตุ และปัจจัยสำคัญที่มีผลต่อด้านการยกระดับคุณภาพการศึกษา โดยนำข้อมูลเหล่านี้ได้ถูกนำไปใช้ในการประเมินความสำเร็จ ความต้องการหรือแนวโน้มด้านการศึกษา รวมถึงพฤติกรรมการเรียนรู้ของผู้เรียน (Adekitan & Noma-Osaghae, 2019; Batool et al., 2023) ดังภาพที่ 1



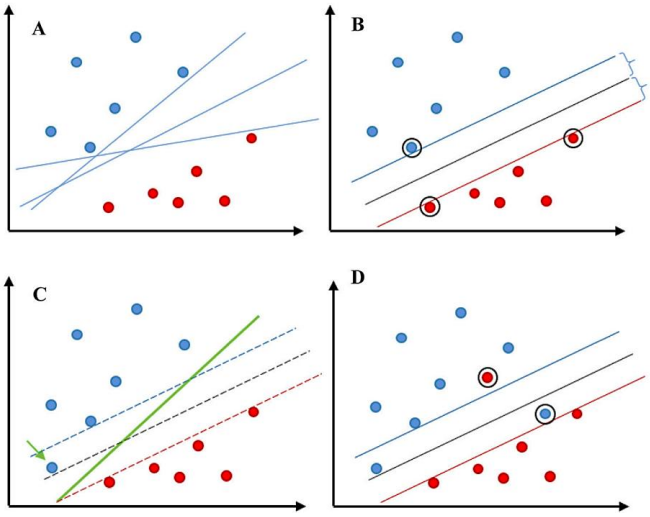
ภาพที่ 1 ประเภทของการเรียนรู้ของเครื่อง

ที่มา: Peng et al. (2021)

2. อัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine: SVM)

SVM เป็นอัลกอริทึมของการเรียนรู้ของเครื่องประเภทหนึ่งโดยการศึกษาของ Chen and Zhai (2023) ได้กล่าวถึงอัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน ถูกสร้างขึ้นเป็นครั้งแรกและได้นำไปใช้แก้ปัญหาการจำแนกประเภทต่าง ๆ เป็นผลสำเร็จ โดย Vapnik (Cortes & Vapnik, 1995) ซึ่งแปลงข้อมูลนำเข้า (Input) ของความสัมพันธ์ระหว่างตัวแปรที่ไม่ได้อยู่

ในรูปแบบเส้นตรง และมีฟังก์ชันเคอร์เนล (Kernel) ที่แปลงข้อมูลจากมิติเดิมไปสู่มิติที่สูงขึ้น ส่งผลให้เกิดความแม่นยำในมิติต่าง ๆ (Tree, 2020) โดยการจำแนกประเภทของอัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีนเป็นวิธีการนำเข้าข้อมูล (Input) ของตัวแปรที่มีความสัมพันธ์ทั้งแบบเชิงเส้นและไม่เชิงเส้นโดยอาศัยการแบ่งแยกข้อมูลด้วยเส้นตรงหรือไฮเปอร์เพลนที่เหมาะสมเพื่อเป็นขอบเขตในการตัดสินใจ ลดข้อผิดพลาดของผลลัพท์ (Output) ที่ได้จากการจำแนกประเภทของชุดข้อมูลที่ต้องการพยากรณ์ (Widaningsih, 2019) ตัวอย่างที่แสดงข้อมูลของการจำแนกประเภทแบบไบนารีโดยใช้โมเดลอัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีนอย่างง่ายที่มีขอบเขตกว้างที่สุด (Maximum-margin classifier) ในภาพ A และภาพ B ขณะเดียวกัน ภาพ C และภาพ D แสดงถึงข้อมูลที่ไม่สามารถแยกออกจากกันได้ด้วยไฮเปอร์เพลนเชิงเส้น (Valkenborg et al., 2023) ดังแสดงในภาพที่ 2



ภาพที่ 2 การจัดจำแนกประเภทของอัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน

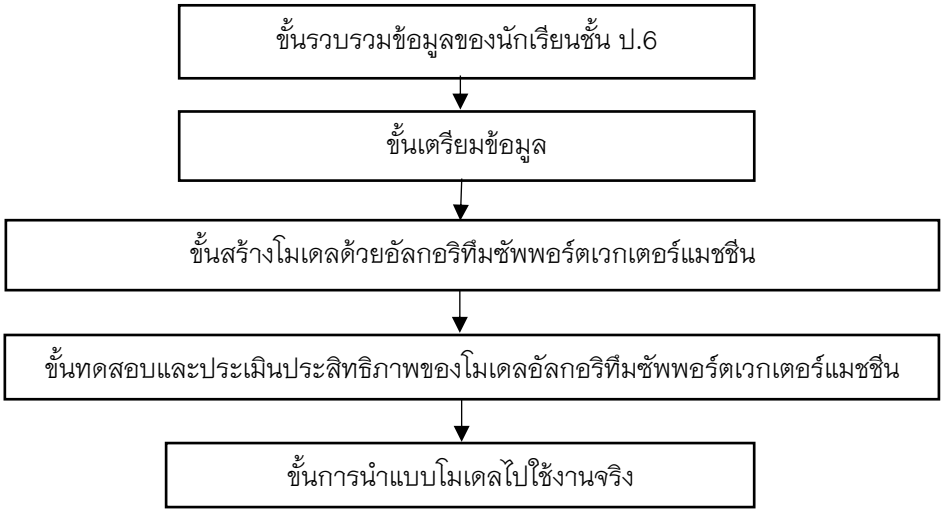
ที่มา: Valkenborg et al. (2023)

วีระพันธ์ พานิชย์ (2564) ได้กล่าวถึงหลักการทำงานของอัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน เป็นการสร้างและเลือกเส้นแบ่งเขตข้อมูลที่เหมาะสมที่สุดในการจำแนกประเภทโดยเฉพาะการจำแนกนักเรียนเป็น 2 กลุ่ม เช่น “ผ่าน/ไม่ผ่าน” โดยระยะห่างระหว่างขอบเขตของ 2 กลุ่มให้มีระยะห่างมากที่สุด ถ้าระยะห่างยิ่งมากความผิดพลาดจะน้อยลง และได้นำเสนอโมเดลเพื่อทำนายผลการเรียนวิชา Web Database ของนักศึกษาสาขาเทคโนโลยีการศึกษา คณะศึกษาศาตร์ มหาวิทยาลัยบูรพาโดยใช้การเรียนรู้ของเครื่อง 4 วิธีด้วยโปรแกรม RapidMiner Studio การศึกษานี้ใช้ข้อมูลจำนวน 271 คน กับข้อมูลการวิเคราะห์ความถดถอยพหุคูณแบบลำดับขั้นด้วยการ Atribute จำนวน 6 Atribute แบ่งเป็น 2 กลุ่ม ได้แก่ กลุ่มเรียนรู้ (Train Set) 80% และกลุ่มทดสอบ (Test set) 20% ผลการศึกษานี้ชี้ให้เห็นว่าประสิทธิภาพการทำนายด้วยวิธี Neural Network มีค่าความถูกต้อง (Accuracy) มากที่สุด 85.89 %รองลงมาวิธี Support Vector Machine 85.19 % ลำดับที่สามวิธี Decision Tree 79.63 % ลำดับสุดท้ายวิธี Naive Bayes 74.07 % ซึ่งสอดคล้องกับ Yağcı (2022) ได้ศึกษาการทำเหมืองข้อมูลทางด้านการศึกษาเกี่ยวกับความสัมพันธ์ของเครื่องมือที่มีประสิทธิภาพในการพยากรณ์ผลสัมฤทธิ์ทางการเรียนของนักศึกษา จำนวน 1,854 คน ที่เรียนวิชาภาษาตุรกีของมหาวิทยาลัยแห่งหนึ่ง ในปี ค.ศ. 2019-2020 ด้วยการเปรียบเทียบประสิทธิภาพของอัลกอริทึมป่าสุ่มและอัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน โดยมีตัวแปรที่ใช้ศึกษาได้แก่ คะแนนสอบกลางภาค ข้อมูลรหัสวิชา และภาควิชา ผลการศึกษาพบว่า การวิเคราะห์ผลการเรียนรู้ของนักศึกษาที่คาดว่าจะไม่จบการศึกษาในการจำแนกประเภทมีความแม่นยำ 70-75% เช่นเดียวกับการทำนายความสำเร็จทางการศึกษาของ Al-Hidayah เมือง Bogor ที่มีข้อมูลของผลการศึกษา 97 รายการ พบว่า การทำเหมืองข้อมูลโดยใช้

อัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีนสามารถทำนายความสำเร็จทางการศึกษาของผู้เรียนที่จบการศึกษาได้ตรงตามเวลาอย่างแม่นยำถึง 85% (Mailana, Putra, Hidayat & Wibowo, 2021)

3. ขั้นตอนการสร้างโมเดล (Modeling) ด้วยอัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน

บทความทางวิชาการนี้ได้เสนอการพยากรณ์ผลสัมฤทธิ์ทางการเรียนโดยใช้อัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีนของนักเรียนชั้นประถมศึกษาปีที่ 6 ด้วย ในบทความนี้ได้ศึกษาเป็นขั้นตอนการสร้างโมเดล ประกอบด้วย 5 ขั้น (ชัยมงคล ปินะสา, สำราญ มีแจ้ง และ น้ำทิพย์ องอาจวาณิชย์, 2567) ดังนี้ 1) ขั้นรวบรวมข้อมูล (Dataset) 2) ขั้นเตรียมข้อมูล (Data Preparation) 3) ขั้นสร้างโมเดล (Modeling) ด้วยอัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน 4) การทดสอบและประเมินประสิทธิภาพของโมเดล (Evaluation) และ 5) ขั้นการนำแบบโมเดลไปใช้งานจริง (Deployment) ดังภาพที่ 3



ภาพที่ 3 ขั้นตอนการสร้างโมเดลอัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน

3.1 การรวบรวมข้อมูล

การเก็บรวบรวมข้อมูลทางการศึกษาในครั้งนี้ได้รวบรวมข้อมูลผลการเรียนของผู้เรียนชั้นประถมศึกษาปีที่ 6 ในปีการศึกษา 2567 ได้แก่ ผู้เรียนโรงเรียนบ้านหนองดอกบัว โรงเรียนบ้านท่าสวรรค์ และโรงเรียนบ้านป่าม่วง จังหวัดเลย จำนวนทั้งหมด 18 คน ด้วยการเลือกประชากรทั้งหมดที่ศึกษา เนื่องจากชุดข้อมูลมีขนาดเล็กจึงได้จำลองข้อมูล (Simulate Data) ซึ่งอ้างอิงจากฐานข้อมูลสารสนเทศของทั้ง 3 โรงเรียน แล้วนำเข้าไฟล์ข้อมูลให้เทคโนโลยีปัญญาประดิษฐ์ ChatGPT ซึ่งข้อดีของ ChatGPT มีหลายประการ ได้แก่ การประหยัดเวลาและทรัพยากร มีความยืดหยุ่นในกำหนดเงื่อนไขต่าง ๆ ของข้อมูลจำลอง และหลีกเลี่ยงความเสี่ยงในการละเมิดข้อมูลส่วนบุคคล ดังนั้นผู้เขียนจึงได้จำลองชุดข้อมูล จำนวน 1,000 ชุดโดยใช้ ChatGPT ได้แก่ ชื่อนามสกุล เพศ เกรดเฉลี่ย จำนวนชั่วโมงที่เรียน เกรดเฉลี่ยสะสม และสถานะทางการเรียน โดยการเขียน Prompt “ช่วยจำลองชุดข้อมูล ได้แก่ ชื่อ นามสกุล เพศ เกรดเฉลี่ย จำนวนชั่วโมงที่เรียน เกรดเฉลี่ยสะสม และสถานะทางการเรียน จากไฟล์ที่แนบ จำนวน 1000 ชุดข้อมูล” ทั้งนี้ ผู้เขียนได้พิจารณาอย่างรอบคอบตามหลักจริยธรรมการวิจัยในมนุษย์โดยไม่นำข้อมูลส่วนบุคคลมาใช้โดยตรงเพื่อป้องกันการละเมิดสิทธิส่วนบุคคล แต่อย่างไรก็ตามชุดข้อมูลที่ได้จำลองนั้นมีลักษณะใกล้เคียงกับข้อมูลจริง เมื่อได้ชุดข้อมูลจะนำมากำหนดค่าตัวแปรตามชนิดของข้อมูล เช่น ชื่อ–นามสกุล ชนิดของข้อมูลเป็น text ถ้าเป็น เพศ ชนิดของข้อมูลเป็น Nominal ส่วนเกรดเฉลี่ยสะสม หรือ GPA เป็นผลการเรียนรายปีการศึกษา ชนิดของข้อมูลเป็น Real โดยผู้เรียนที่ได้เกรด 0 หมายถึง จะมีคะแนนต่ำสุด และผู้เรียนที่ได้เกรด 4 หมายถึง จะมีคะแนนสูงสุด อีกทั้ง Wiyono, Abidin, Wibowo, Hidayatullah & Dairoh (2019) ได้กล่าวถึง Satuan Kredit

Semester: SKS เป็นหน่วยกิตที่ผู้เรียนได้ลงทะเบียนเรียนในแต่ละปีการศึกษา และบ่งบอกถึงสถานะทางการเรียน ซึ่งเป็นข้อมูลที่นำมาใช้ในการทำนายความแม่นยำหรือความล่าช้าของความสำเร็จในการเรียน เมื่อการเก็บรวบรวมข้อมูลของการสร้างโมเดลเสร็จสิ้นแล้ว จากนั้นบันทึกข้อมูลเป็นไฟล์นามสกุล CSV (Comma Separated Values) เพื่อนำไปสู่ขั้นการเตรียมข้อมูล (Novianto & Suhirman, 2024)

3.2 ขั้นการเตรียมข้อมูลเพื่อประมวลผลของชุดข้อมูล ดังนี้

ขั้นตอนที่ 1 ขั้นจัดรูปแบบของข้อมูล (Formatting) เพื่อหลีกเลี่ยงและลดสัญญาณรบกวนของข้อมูลที่ไม่ครบถ้วนสมบูรณ์ (Switrayana, Ashadi, Hairani, & Aminuddin, 2023) ได้แก่ การปรับสเกลของข้อมูล การเปลี่ยนข้อความ (Text) แทนด้วยตัวเลข (Numeric) ได้แก่ เพศชาย แทน 0 เพศหญิง แทน 1 เป็นต้น

ขั้นตอนที่ 2 ขั้นทำความสะอาดของชุดข้อมูล เป็นขั้นตอนในการกำหนดลักษณะของข้อมูลที่ไม่สอดคล้องกัน (Putri & Waspada, 2018)

ตัวอย่างการเตรียมข้อมูลเพื่อป้อนเข้าภายหลังจากการกำหนดค่าตัวแปร ได้แก่ 1) ชื่อ –นามสุกล เลือกชนิดของข้อมูลเป็น Nominal 2) เพศ เลือกชนิดของข้อมูลเป็น Text 3) เกรดเฉลี่ย เลือกชนิดของข้อมูลเป็น Real 4) หน่วยกิต เลือกชนิดของข้อมูล Integer และ 5) สถานะทางการเรียน เลือกชนิดของข้อมูลเป็น Nominal และ Label

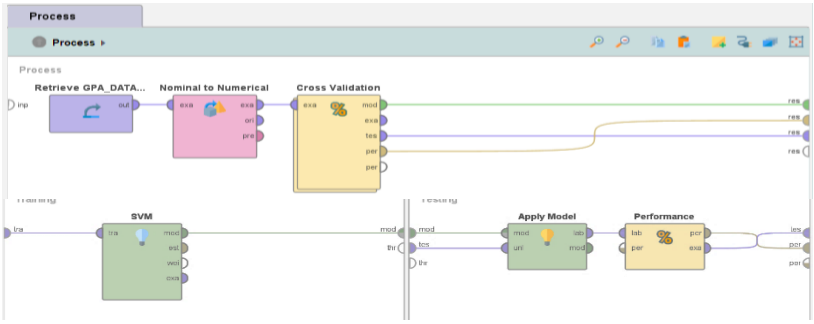
3.3 ขั้นการสร้างโมเดลด้วยอัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน มีขั้นตอนดังนี้

1. สร้างโมเดล (Modeling) โดยการกำหนดโครงสร้างโมเดลด้วยอัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน ดังนี้

1.1 เลือก Kernel Function ที่เหมาะสมกับลักษณะของข้อมูล เส้นตรงสามารถแยกข้อมูลออกเป็นสองกลุ่มได้ชัดเจน โมเดลที่ใช้ Linear Kernel มีความเรียบง่าย ใช้เวลาฝึกน้อย และเหมาะกับชุดข้อมูลขนาดใหญ่

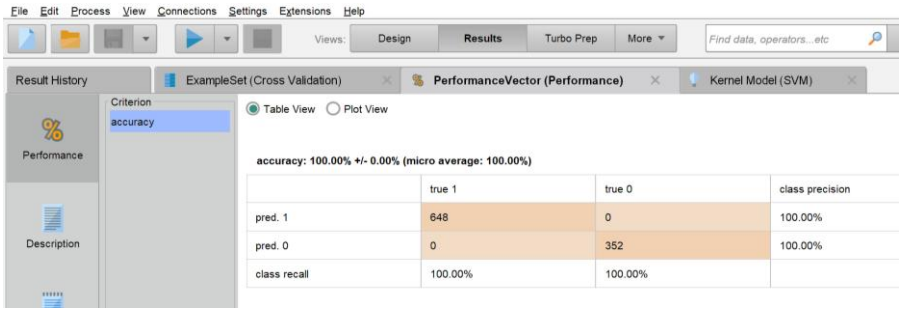
1.2 การตั้งค่าพารามิเตอร์ C ที่ใช้ควบคุมความสมดุลระหว่างการจำแนกข้อมูลให้ถูกต้องข้อผิดพลาดบางกรณีในชุดข้อมูลฝึก (Training Data) เพื่อหลีกเลี่ยงการฝึกโมเดลจนจำรายละเอียดมากเกินไป (Overfitting) และ Gamma ที่ควบคุมขอบเขตของอิทธิพลที่ข้อมูลแต่ละจุดของโมเดล อย่างรอบคอบ

2. ขั้นตอนการทำเหมืองข้อมูลทางการศึกษา เป็นขั้นตอนภายหลังจากการเก็บรวบรวม การจัดรูปแบบและการทำความสะอาดของชุดข้อมูล เพื่อวิเคราะห์ข้อมูลเชิงปริมาณและเชิงคุณภาพ โดยใช้โปรแกรม Altair AI Studio Educational 2025.0.0 (Uska, Wirasasmita, Usuluddin & Arianti, 2020) ซึ่งเป็นแพลตฟอร์มสำหรับข้อมูลขนาดใหญ่ มีข้อดี คือ เป็นแพลตฟอร์มที่นักพัฒนาอนุญาตให้ใช้งานได้ฟรี อีกทั้งยังสามารถขอมีลิขสิทธิ์ในเชิงพาณิชย์ได้อีกด้วย (Madyatmadja et al., 2021) ขั้นตอนแรกในการใช้งานของ FETATNIA (2024) ดังนี้ 1) เปิดโปรแกรม Altair AI Studio Educational 2025.0.0 แล้วคลิกมุมบนซ้ายคำว่า "Repository" เปิดไฟล์ที่จัดเก็บข้อมูล 2) คลิกขวาเพื่อเลือก "Import Data" นำเข้าไฟล์ CSV 3)เลือกไฟล์ CSV จากนั้นคลิก "Open" 3) ตรวจสอบคุณสมบัติประเภทของข้อมูล และ 4) คลิก "Finish" เพื่อนำเข้าข้อมูลไฟล์นามสกุล CSV เข้าสู่โปรแกรม Altair AI Studio Educational 2025.0.0 จากนั้นแบ่งชุดข้อมูลออกเป็น 2 กลุ่ม คือ กลุ่มฝึกสอนโมเดล (Training Set) 70% และกลุ่มทดสอบโมเดล (Testing Set) 30% ดังภาพที่ 4



ภาพที่ 4 ตัวอย่างการสร้างโมเดล (Modeling) ด้วยอัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน

3.4 ขั้นทดสอบและประเมินประสิทธิภาพของโมเดลอัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีนด้วยวิธีการ 10-Fold cross-validation ซึ่งเป็นวิธีการที่ได้รับการยอมรับและนำมาใช้กันอย่างแพร่หลายในงานวิจัยด้านการเรียนรู้ของเครื่อง เนื่องจากสามารถสร้างความสมดุลระหว่างความแม่นยำของการประเมินและภาระในการคำนวณ (Kohavi, 1995) โดยแบ่งชุดข้อมูลแต่ละรอบออกเป็น 10 ส่วน โดย 9 ส่วนจะเป็นกลุ่มฝึกสอนโมเดล 70% และอีก 1 ส่วนจะเป็นกลุ่มทดสอบโมเดล 30% (Testing Set) แล้วสลับกันฝึกและทดสอบ จนครบทั้ง 10 ส่วน เมื่อกระบวนการครบถ้วนแล้ว จะนำผลลัพธ์จากแต่ละรอบมาคำนวณค่าเฉลี่ยเพื่อประเมินประสิทธิภาพของภาพรวมโมเดล จากนั้นจึงนำโมเดลที่ได้ไปรันเพื่อประเมินผลลัพธ์จริงออกมา ซึ่งการวิเคราะห์ผลลัพธ์ที่ได้จากตารางประเมินผล (Confusion Matrix) เป็นการพยากรณ์ของอัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีนเทียบกับค่าจริงของข้อมูล (Gu, Zhu & Cai, 2009; Valero-Carreras, Alcaraz & Landete, 2023) ดังภาพที่ 5 ได้แสดงค่าประสิทธิภาพหลักของโมเดล ได้แก่ ค่าความถูกต้อง (Accuracy) ที่บ่งชี้ระดับความถูกต้องของการจำแนกประเภทของโมเดล ค่าความแม่นยำ (Precision) โดยการวัดความแม่นยำของการพยากรณ์กับตัวอย่างที่โมเดลพยากรณ์ และค่าความระลึก (Recall) จะวัดความสามารถของโมเดลในการตรวจจับตัวอย่าง (Ali & Saleem, 2022; Valero-Carreras et al., 2023)



ภาพที่ 5 ค่าประสิทธิภาพการพยากรณ์ของโมเดล

หากนำผลลัพธ์ (Output) จากทดสอบประสิทธิภาพของโมเดล (Testing) มาเขียนสมการแปลงค่าผลลัพธ์การพยากรณ์ผลสัมฤทธิ์ทางการเรียนด้วยสูตรของ Putri et al. (2023) จะได้ค่าการคำนวณประสิทธิภาพการพยากรณ์ของโมเดล ดังนี้

$$\text{ค่าความถูกต้อง accuracy} = \frac{TP+TN}{TP+TN+FP+FN} = \frac{648+352}{648+352+0+0} = \frac{1000}{1000} = 1.0 = 100\%$$

$$\text{ค่าความแม่นยำ precision} = \frac{TP}{TP+FP} = \frac{648}{648+0} = \frac{648}{648} = 1.0 = 100\%$$

$$\text{ค่าความระลึก recall} = \frac{TP}{TP+FN} = \frac{352}{352+0} = \frac{352}{352} = 1.0 = 100\%$$

ผลลัพธ์ของการศึกษาครั้งนี้สอดคล้องการศึกษาของ Novianto and Suhirman (2024) จากการทบทวนวรรณกรรม(Hairani, 2021; Haryatmi & Hervianti, 2021)ที่ได้เปรียบเทียบประสิทธิภาพความแม่นยำขออัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน ไว้อย่างน่าสนใจ แสดงในตารางที่ 1

ตารางที่ 1 ตารางการเปรียบเทียบผลการวิจัยเกี่ยวกับอัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน

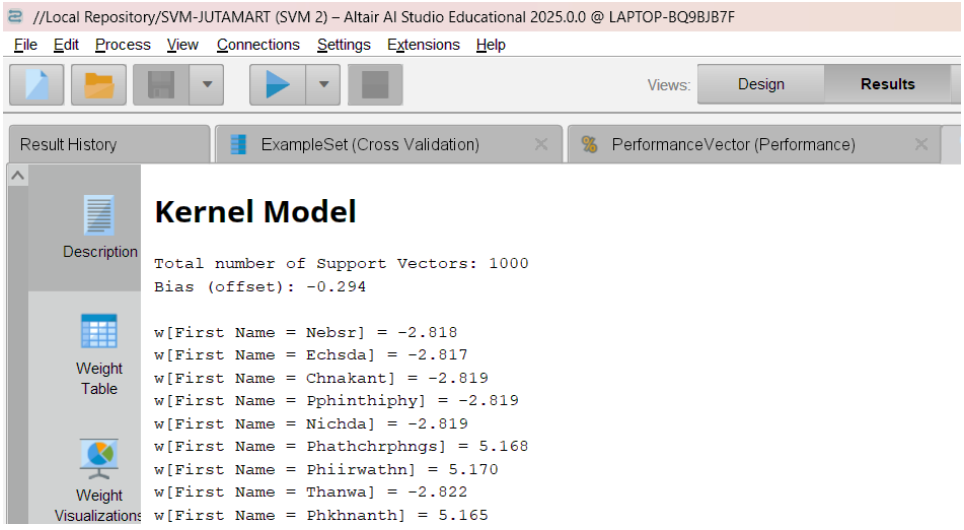
ผู้วิจัย (Author)	**ค่าความถูกต้อง** (Accuracy)
Hairani, 2021	SVM: 83.9%
Haryatmi & Hervianti, 2021	SVM: 94.4%
Novianto and Suhirman, 2024	K-NN: 98.45% และ SVM: 96.90%
บทความที่นำเสนอ	SVM: 100%

ที่มา: Novianto and Suhirman (2024)

3.5 ขั้นการนำโมเดลอัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีนไปใช้งานจริง

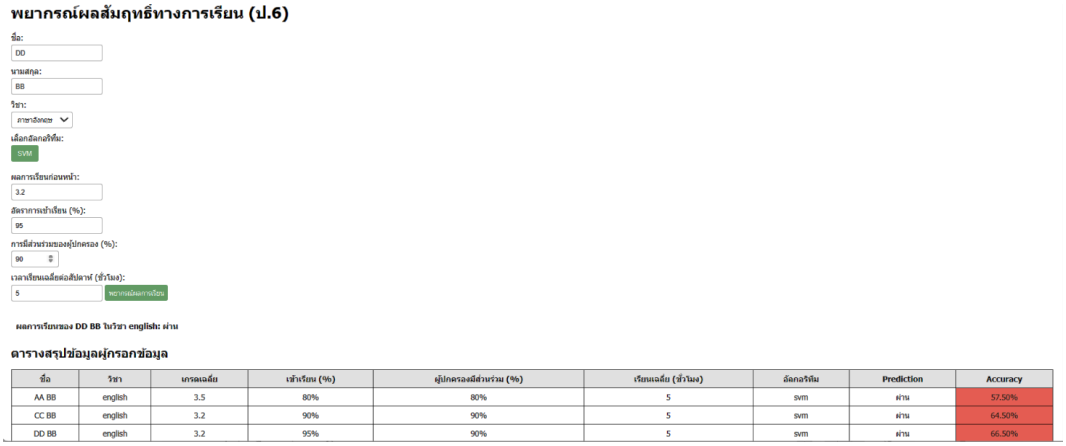
หลังจากทดสอบและประเมินประสิทธิภาพของโมเดลด้วยอัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีนแล้ว ขั้นตอนต่อไปคือการนำโมเดลที่ผ่านการฝึกฝนด้วยโปรแกรม Altair AI Studio Educational 2025.0.0 แล้วแบ่งกลุ่มข้อมูลกลุ่มฝึกสอน 70% และข้อมูลกลุ่มทดสอบ 30% จากนั้นให้เลือก ตัวดำเนินการ (Operators) ดังนี้ 1) Retrieve Student Data เรียกใช้งานในรูปแบบไฟล์ เช่น .csv 2) Select Attributes เพื่อเลือกคุณลักษะณะของชุดข้อมูล 3) Set Role ให้กำหนดบทบาทของตัวแปรในชุดข้อมูล 4) Cross Validation ดำเนินการตรวจสอบประสิทธิภาพความแม่นยำของโมเดลด้วยวิธีการ 10-Fold cross-validation 5) เลือก Support Vector Machine ปรับตั้งค่าพารามิเตอร์และ Kernel Function 6) Apply Model นำโมเดลที่ฝึกฝนแล้วไปประยุกต์ใช้กับข้อมูล และ 7) Performance ประเมินประสิทธิภาพของโมเดลด้วยค่าความถูกต้อง ค่าความแม่นยำ และค่าความระลึก

จะเห็นได้ว่าผลลัพธ์ของโมเดลมีค่าเฉลี่ยรวมจากทุกชั้นเท่ากับ 100 % ซึ่งโมเดลที่ถูกฝึกด้วยเวกเตอร์สนับสนุนจะมีทั้งหมด 1,000 รายการ โดยมีค่าน้ำหนักเฉพาะตัวที่สะท้อนถึง เส้นแบ่งการจำแนกประเภท การมีค่าบวกและลบในน้ำหนักชี้ให้เห็นถึงความสามารถของโมเดลในการจำแนกกลุ่มข้อมูลออกจากกันในลักษณะของ Binary Classification นอกจากนี้ โมเดลยังแสดงค่าไบแอส (Bias) เท่ากับ -0.294 เป็นค่าคงที่ที่ใช้ในสมการของเส้นแบ่งการตัดสินใจ ดังแสดงในภาพที่ 6



ภาพที่ 6 แสดงค่าฟังก์ชันเคอร์เนลของโมเดล

จากนั้นส่งออกโมเดลในส่วนต่อประสานระหว่างผู้ใช้ (User) กับระบบคอมพิวเตอร์หรือแพลตฟอร์มซึ่งยังไม่ได้เชื่อมต่อกับโมเดลจากอัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน โดยตรงจากโปรแกรม Altair AI Studio Educational 2025.0.0 ในรูปแบบ JavaScript ซึ่งเป็นมาตรฐานกลางสำหรับการนำโมเดลที่ฝึกแล้วไปใช้งานจริง หรือเชื่อมต่อเข้ากับระบบออนไลน์ของเครื่องโดยบูรณาการร่วมกับแพลตฟอร์ม One Compiler เพื่อสาธิตรูปแบบการใช้งาน ซึ่งเป็นสภาพแวดล้อมออนไลน์สำหรับการรันโค้ดและออกแบบส่วนติดต่อผู้ใช้ด้วยภาษา HTML ที่สามารถป้อนข้อมูลนักเรียนในการพยากรณ์ผลสัมฤทธิ์ทางการเรียนล่วงหน้าได้อย่างสะดวก ได้แก่ ชื่อ–สกุล รายวิชา ผลการเรียนก่อนหน้า เปอร์เซ็นต์ของอัตราการเข้าเรียน การมีส่วนร่วมของผู้ปกครอง รวมถึงเวลาเรียนเฉลี่ยต่อสัปดาห์ เมื่อมีการกรอกข้อมูลที่ครบถ้วน จะทำการเลือกพยากรณ์ผลการเรียน แล้วระบบจะทำหน้าที่วิเคราะห์และประเมินผลสัมฤทธิ์ทางการเรียนที่ได้จากการพยากรณ์ไว้ล่วงหน้าได้อย่างเป็นระบบโดยใช้เกณฑ์ตัดสินผลการเรียนผ่านหรือไม่ผ่าน สะดวกในการนำไปใช้งาน ซึ่งเป็นประโยชน์แก่ครู ผู้บริหารสถานศึกษา และผู้ปกครอง ในการกรอกข้อมูลต่าง ๆ พยากรณ์ผลการเรียนผ่านระบบออนไลน์ ดังภาพที่ 10



ภาพที่ 10 แพลตฟอร์ม One Compiler เพื่อการพยากรณ์ผลสัมฤทธิ์ทางการเรียน

ที่มา: https://drive.google.com/drive/folders/1_wWlu0Bou_1GsjaTQv62daSolsgkQQD2?usp=sharing

บทสรุป

บทความทางวิชาการนี้มุ่งนำเสนอผลการศึกษามุ่งนำเสนอการสร้างโมเดลในการพยากรณ์ผลสัมฤทธิ์ทางการเรียนของนักเรียนชั้นประถมศึกษาปีที่ 6 เพื่อนำข้อมูลที่ได้ไปใช้ในการวางแผน พัฒนา และปรับปรุงคุณภาพการศึกษาบนพื้นฐานแนวคิดพื้นฐานของการทำเหมืองข้อมูลทางการศึกษาอัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีน และการใช้โปรแกรม Altair AI Studio ในการพัฒนาโมเดลเชิงพยากรณ์ ซึ่งมี 5 ขั้นตอนหลัก ได้แก่ การรวบรวมข้อมูล การเตรียมข้อมูล การสร้างโมเดล การทดสอบและประเมินผลประสิทธิภาพโมเดล และการนำไปใช้งานจริงโดยจำลองชุดข้อมูล จำนวน 1,000 คน จาก ChatGPT ที่อ้างอิงจากฐานข้อมูลจริง ประกอบด้วย ชื่อ–นามสกุล เพศ เกรดเฉลี่ย จำนวนชั่วโมงที่เรียน เกรดเฉลี่ยสะสม และสถานะทางการเรียน ผ่านการวิเคราะห์ด้วยโปรแกรม Altair AI Studio Educational 2025.0.0 ด้วยวิธีการ 10–Fold Cross–Validation เพื่อเพิ่มความแม่นยำในการ พยากรณ์ผลสัมฤทธิ์ทางการเรียน

ผลการศึกษาพบว่า การสร้างโมเดลโดยใช้อัลกอริทึมซัพพอร์ตเวกเตอร์แมชชีนมีค่าความแม่นยำของค่าความถูกต้อง ค่าความแม่นยำ และค่าความระลึกสูงถึง 100% บ่งบอกถึงประสิทธิภาพของโมเดลที่นำมาบูรณาการร่วมกับแพลตฟอร์ม One Compiler เพื่อสาธิตรูปแบบการใช้งาน ซึ่งถูกออกแบบให้สามารถเชื่อมต่อกับส่วนติดต่อผู้ใช้ในรูปแบบ HTML เพื่อเอื้อประโยชน์แก่ครู ผู้บริหารสถานศึกษา และผู้ปกครอง ในการกรอกข้อมูลต่าง ๆ พยากรณ์ผลการเรียนผ่านระบบออนไลน์ ทั้งนี้ในการสร้างชุดข้อมูลจำลอง จำนวน 1,000 ชุดข้อมูลด้วย ChatGPT ที่อ้างอิงจาก

ฐานข้อมูลจริง ถือเป็นแนวทางใหม่ในการเข้าถึงข้อมูลผู้เรียนเกี่ยวกับข้อจำกัดด้านจริยธรรมและความเป็นส่วนตัว ช่วยให้การสร้างโมเดลเบื้องต้นมีคุณภาพในการนำมาใช้งาน

แม้โมเดล SVM นี้จะถูกฝึกด้วยข้อมูลจากโรงเรียนต้นแบบ และยังไม่มีการตรวจสอบความแม่นยำในสภาพแวดล้อมจริงแต่สามารถทำนายผลสัมฤทธิ์ทางการเรียนได้อย่างมีประสิทธิภาพ สะท้อนถึงความสามารถของโมเดลในการนำไปใช้งานจริงกับข้อมูลจากโรงเรียนอื่นที่ไม่อยู่ในชุดข้อมูลฝึกหัด อย่างไรก็ตามหากบริบทของโรงเรียนมีความแตกต่างกันอย่างชัดเจน เช่น ลักษณะของผู้เรียน แนวทางการพัฒนาโมเดลในอนาคตอาจจำเป็นต้องมีการปรับปรุงหรือฝึกฝนโมเดลเพิ่มเติม โดยนำข้อมูลจริงจากหลายพื้นที่หลากหลายเพื่อเพิ่มความสามารถของการปรับตัว และลดความเสี่ยงจาก Overfitting จะช่วยให้ได้ผลลัพธ์ที่แม่นยำมากยิ่งขึ้น นอกจากนี้ยังมีข้อจำกัดในด้านชุดข้อมูลจริงที่เป็นเพียงกลุ่มตัวอย่างจึงทำให้เกิดความลำเอียงของชุดข้อมูลและไม่อาจสะท้อนการพยากรณ์ผลสัมฤทธิ์ทางการเรียนของนักเรียนในพื้นที่อื่นได้

เอกสารอ้างอิง

นคร ละลอกน้ำ. (2561). การใช้เทคนิคการขุดข้อมูลทางการศึกษาเพื่อการศึกษา. **วารสารการศึกษาและการพัฒนาสังคม, 14**(2), 1–16.

ชัยมงคล ปินะสา, สำราญ มีแจ้ง และ น้ำทิพย์ องอาจวาณิชย์. (2567). การประยุกต์ใช้โครงข่ายประสาทเทียมในการคัดเลือกข้อสอบข้อถัดไป. **วารสารวัดผล, 30**(1), 1–15.

วีระพันธ์ พาณิชย์. (2564). **การประยุกต์ใช้ Machine Learing ทำนายผลการเรียนวิชา Web Database ของนิสิตสาขาเทคโนโลยีการศึกษา คณะศึกษาศาตร์ มหาวิทยาลัยบูรพา**. (วิทยานิพนธ์การศึกษาตามหลักสูตรวิศวกรรมศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมคอมพิวเตอร์และโทรคมนาคม). กรุงเทพฯ: วิทยาลัยนวัตกรรมด้านเทคโนโลยีและวิศวกรรมศาสตร์มหาวิทยาลัยธุรกิจบัณฑิตย์.

สำนักนายกรัฐมนตรี. (2545). **พระราชบัญญัติการศึกษาแห่งชาติ พ.ศ.2542 และที่แก้ไขเพิ่มเติม (ฉบับที่ 2) พ.ศ.2545 สำนักนายกรัฐมนตรีประกาศในราชกิจจานเบกษา เล่มที่ 119 ตอนที่ 123ก วันที่ 19 ธันวาคม 2545**. กรุงเทพฯ: พริกหวานกราฟฟิค.

Adekitan, A. I., & Noma-Osaghae, E. (2019). Data mining approach to predicting the performance of first year student in a university using the admission requirements. **Education and Information Technologies, 24**, 1527–1543. Form https://doi.org/10.1007/s10639-018-9839-7.

Adekitan, A. I., & Salau, O. (2019). The impact of engineering students' performance in the first three years on their graduation result using educational data mining. **Heliyon, 5**(2). Form https://doi.org/10.1016/j.heliyon.2019. e01250.

Ali, A., & Saleem, N. N. (2022). Classification of Software Systems attributes based on quality factors using linguistic knowledge and machine learning: A review. **Journal of Education and Science, 31**(3), 66.60-90.60. Form https://doi.org/10.33899/edusj.2022.134024.1245.

Batool, S., Rashid, J., Nisar, M. W., Kim, J., Kwon, H.-Y., & Hussain, A. (2023). Educational data mining to predict students' academic performance: A survey study. **Education and Information Technologies, 28**(1), 905–971. Form https://doi.org/10.1007/s10639-022-11152-y.

Bisri, A., Heryatun, Y., & Navira, A. (2025). Educational Data Mining Model Using Support Vector Machine for Student Academic Performance Evaluation. **Journal of Education and Learning (EduLearn), 19**(1), 478–486. Form https://doi.org/10.11591/edulearn.v19i1.21609.

Chen, Y., & Zhai, L. (2023). A comparative study on student performance prediction using machine learning. **Education and Information Technologies, 28**(9), 12039–12057. Form https://doi.org/10.1007/s10639-023-11672-1.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. **Machine learning, 20**, 273–297.

Fetatnia, A. (2024). **A Comparative Study of Data Mining Tools**. Form http://dspace.univ-guelma.dz/jspui/handle/123456789/16489.

Gu, Q., Zhu, L., & Cai, Z. (2009). Evaluation measures of the classification performance of imbalanced data sets. **2009 International Symposium on Intelligence Computation and Applications, Huangshi** (461–471). Form https://doi.org/10.1007/978-3-642-04962-0_53.

H. Alamri, L., S. Almuslim, R., S. Alotibi, M. ,K. Alkadi, D., UllahKhan, I., Aslam,N., (2020). Predicting student academic performance using support vector machine and random forest. **2020 3rd International Conferenceon Education Technology Management** (100–107). Form https://doi.org/10.1145/3446590.3446607.

Haryatmi, E., & Hervianti, S. P. (2021). Penerapan algoritma support vector machine untuk model prediksi kelulusan mahasiswa tepat waktu. **Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi), 5**(2), 386–392. Form https://doi.org/10.29207/resti.v5i2.3007.

Kohavi, R. (1995). A Study of Cross-Validation and bootstrap for accuracy estimation and model selection. **International Joint Conference on Artificial Intelligence (IJCAI)** (1137–1143). Morgan Kaufmann. Form https://www.researchgate.net/publication/2352264.

Larose, D. T., & Larose, C. D. (2014). **Discovering knowledge in data: an introduction to data mining**. John Wiley & Sons.

Madyatmadja, E. D., Jordan, S. I., & Andry, J. F. (2021). Big data analysis using rapidminer studio to predict suicide rate in several countries. **ICIC Express Letters, Part B: Applications, 12**(8), 757–764. Form https://doi.org/10.24507/icicelb.12.08.757.

Mailana, A., Putra, A. A., Hidayat, S., & Wibowo, A. (2021). Comparison of C4. 5 algorithm and support vector machine in predicting the student graduation timeliness. **Jurnal Online Informatika, 6**(1), 11–16. Form https://doi.org/10.15575/join.v6i1.608.

Novianto, E., & Suhirman, S. (2024). Comparison of K-nearest neighbor classification methods and support vector machine in predicting students' Sstudy period. **Journal of Education and Science, 33**(1), 32.30–45.30. Form https://doi.org/10.33899/edusj.2023.144865.1408.

Peng, J., Jury, E. C., Dönnes, P., & Ciurtin, C. (2021). Machine learning techniques for personalised medicine approaches in immune-mediated chronic inflammatory diseases: applications and challenges. **Frontiers in pharmacology, 12**, 720694. Form https://doi.org/10.3389/fphar.2021.720694.

Putri, R. P. S., & Waspada, I. (2018). Penerapan algoritma C4. 5 pada aplikasi prediksi kelulusan mahasiswa prodi informatika. **Khazanah Informatika: Jurnal Ilmu Komputer dan Informatika, 4**(1), 1–7. Form https://doi.org/10.23917/khif.v4i1.5975.

Ramageri, B. M. (2010). Data mining techniques and applications. **Indian journal of computer science and engineering, 1**(4), 301–305.

Rastrollo–Guerrero, J. L., Gómez–Pulido, J. A., & Durán–Domínguez, A. (2020). Analyzing and predicting students' performance by means of machine learning: A review. **Applied sciences, 10**(3), 1042. Form https://doi.org/10.3390/app10031042.

Switrayana, I. N., Ashadi, D., Hairani, H., & Aminuddin, A. (2023). Sentiment analysis and topic modeling of Kitabisa applications using Support Vector Machine (SVM) and Smote–Tomek links methods. **International Journal of Engineering and Computer Science Applications (IJECSA), 2**(2), 87–98. Form https://doi.org/10.30812/ijecsa.v2i2.3406.

Tree, D. (2020). Klasifikasi mahasiswa her berbasis algoritma svm dan decision tree. **Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK), 7**(6). Form https://doi.org/10.25126/jtiik.202073080.

Uska, M. Z., Wirasasmita, R. H., Usuluddin, U., & Arianti, B. D. D. (2020). Evaluation of rapidminer–aplication in data mining learning using persiva model. **Edumatic: Jurnal Pendidikan Informatika, 4**(2), 164–171. Form https://doi.org/10.29408/edumatic.v4i2.2688.

Valero–Carreras, D., Alcaraz, J., & Landete, M. (2023). Comparing two SVM models through different metrics based on the confusion matrix. **Computers & Operations Research, 152**, 106131. Form https://doi.org/10.1016/j.cor.2022.106131.

Valkenborg, D., Rousseau, A.–J., Geubbelmans, M., & Burzykowski, T. (2023). Support vector machines. **American Journal of Orthodontics and Dentofacial Orthopedics, 164**(5), 754–757. Form https://doi.org/10.1016/j.ajodo.2023.08.003.

Widaningsih, S. (2019). Perbandingan metode data mining untuk prediksi nilai dan waktu kelulusan mahasiswa prodi teknik informatika dengan algoritma C4, 5, Naïve Bayes, Knn Dan Svm. **Jurnal Tekno Insentif, 13**(1), 16–25. Form https://doi.org/10.36787/jti.v13i1.78.

Wiyono, S., Abidin, T., Wibowo, D., Hidayatullah, M., & Dairoh, D. (2019). Comparative study of machine learning knn, svm, and decision tree algorithm to predict students performance. **International Journal of Research–Granthaalayah, 7**(1), 190–196. Form https://doi.org/10.29121/granthaalayah.v7.i1.2019.1048.

Yağcı, M. (2022). Educational data mining: prediction of students' academic performance using machine learning algorithms. **Smart Learning Environments, 9**(1), 11. Form https://doi.org/10.1186/s40561–022–00192–z.