# STUDENT PERFORMANCE PREDICTION IN AN ONLINE LEARNING MANAGEMENT SYSTEM USING MACHINE LEARNING

ปริยานุช ประเสริฐสิริกุล

บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ

2564

# การทำนายผลการเรียนของนิสิตที่ใช้งานระบบการจัดการเรียนรู้ออนไลน์ โดยใช้การเรียนรู้ของเครื่อง

ปริยานุช ประเสริฐสิริกุล

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
วิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาการข้อมูล
คณะวิทยาศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ
ปีการศึกษา 2564
ลิขสิทธิ์ของมหาวิทยาลัยศรีนครินทรวิโรฒ

# STUDENT PERFORMANCE PREDICTION IN AN ONLINE LEARNING MANAGEMENT SYSTEM USING MACHINE LEARNING

PARIYANUCH PRASERTSIRIKUL

A Thesis Submitted in Partial Fulfillment of the Requirements

for the Degree of MASTER OF SCIENCE

(Data Science)

Faculty of Science, Srinakharinwirot University

2021

Copyright of Srinakharinwirot University

ปริญญานิพนธ์

เรื่อง

การทำนายผลการเรียนของนิสิตที่ใช้งานระบบการจัดการเรียนรู้ออนไลน์

โดยใช้การเรียนรู้ของเครื่อง

ของ

ปริยานุช ประเสริฐสิริกุล

ได้รับอนุมัติจากบัณฑิตวิทยาลัยให้นับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาการข้อมูล

ของมหาวิทยาลัยศรีนครินทรวิโรฒ

---

(รองศาสตราจารย์ นายแพทย์ฉัตรชัย เอกปัญญาสกุล)

คณบดีบัณฑิตวิทยาลัย

---

คณะกรรมการสอบปากเปล่าปริญญานิพนธ์

..... ที่ปรึกษาหลัก

(อาจารย์ ดร.ศิริสรรพ เหล่าหะเกียรติ)

..... ประธาน

(อาจารย์ ดร.รัตน์ชัยนันท์ ธรรมสุจริต)

..... กรรมการ

(ผู้ช่วยศาสตราจารย์ ดร.นุวีย์ วิวัฒนวัฒนา)

| | |
|---|---|
| ชื่อเรื่อง | การทำนายผลการเรียนของนิสิตที่ใช้งานระบบการจัดการเรียนรู้ออนไลน์ โดยใช้การเรียนรู้ของเครื่อง |
| ผู้วิจัย | ปริยานุช ประเสริฐสิริกุล |
| ปริญญา | วิทยาศาสตรมหาบัณฑิต |
| ปีการศึกษา | 2564 |
| อาจารย์ที่ปรึกษา | อาจารย์ ดร. ศิริสรรพ เหล่าหะเกียรติ |

งานวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาการทำนายผลการเรียนของนิสิตที่ใช้งานระบบการจัดการเรียนรู้ออนไลน์ โดยใช้การเรียนรู้ของเครื่อง เพื่อหาแบบจำลองการจำแนกประเภทสำหรับการทำนายผลการเรียนของนิสิตที่ดีที่สุด โดยใช้ชุดข้อมูลการเรียนในช่วงครึ่งแรกของภาคการศึกษา ตั้งแต่วันที่ 7 มกราคม ถึง 5 มีนาคม 2564 ของนิสิตที่ลงทะเบียนเรียนรายวิชา SWU 141 ชีวิตในโลกดิจิทัล ภาคการศึกษา 2 ปีการศึกษา 2563 จำนวน 405 คน ผู้วิจัยได้ศึกษาเพื่อเปรียบเทียบประสิทธิภาพของแบบจำลอง โดยใช้ชุดข้อมูลที่แตกต่างกันตามประเภทของคุณลักษณะ ได้แก่ คุณลักษณะข้อมูลทั่วไป คุณลักษณะกลุ่มพฤติกรรมการใช้งานระบบ คุณลักษณะกลุ่มคะแนน โดยเริ่มจากการนำข้อมูลการเรียนผ่านทางระบบจัดการเรียนรู้ออนไลน์ของนิสิตจำนวน 99,925 รายการ มาวิเคราะห์ เพื่อสร้างคุณลักษณะกลุ่มพฤติกรรมการใช้งานระบบ โดยใช้เทคนิคการทำวิศวกรรมคุณลักษณะ (Feature Engineering) แล้วนำคุณลักษณะกลุ่มพฤติกรรมการใช้งานระบบที่สร้างขึ้นมาร่วมกับคุณลักษณะกลุ่มคะแนน และคุณลักษณะข้อมูลทั่วไป ไปสร้างแบบจำลองที่ได้รับความนิยมในปัจจุบันหลากหลายแบบจำลอง ได้แก่ Logistic Regression, Naïve Bayes, K-Nearest Neighbor, Support Vector Machines, Decision Tree, Random Forest และ XGBoost ผู้วิจัยได้เปรียบเทียบและวัดประสิทธิภาพของแบบจำลองเหล่านี้ร่วมกับการใช้เทคนิคการจัดการความไม่สมดุลของข้อมูลด้วยวิธีการสุ่มตัวอย่างเพิ่มกลุ่มน้อยด้วยการสร้างตัวอย่างสังเคราะห์ (Synthetic Minority Oversampling Technique: SMOTE) รวมทั้งการคัดเลือกคุณลักษณะ (Feature Selection) จากการทดลองพบว่า แบบจำลอง XGBoost ให้ประสิทธิภาพในการทำนายสูงสุด มีความแม่นยำถึง 83.95% นอกจากนี้ยังพบว่า คุณลักษณะกลุ่มคะแนนให้ผลการทำนายที่ดีกว่าการใช้คุณลักษณะกลุ่มพฤติกรรมการใช้งานระบบ แต่การใช้คุณลักษณะทั้งสองด้านร่วมกัน ให้ผลการทำนายดีที่สุด และการคัดเลือกคุณลักษณะช่วยเพิ่มประสิทธิภาพในแบบจำลองที่มีความซับซ้อนต่ำ แต่ในแบบจำลองที่มีความซับซ้อนสูง การคัดเลือกคุณลักษณะไม่ได้ช่วยเพิ่มประสิทธิภาพการทำนาย อีกทั้งการศึกษานี้ ยังแสดงคุณลักษณะที่สำคัญ เพื่อเป็นภาพสะท้อนให้เห็นคุณลักษณะที่มีบทบาทสำคัญในการทำนายผลการเรียนของนิสิต โดยมีกลุ่มของพฤติกรรมการใช้งานระบบของนิสิตมากถึง 19 คุณลักษณะ ในขณะที่คุณลักษณะสำคัญในกลุ่มของคะแนนผลการประเมินการเรียนรู้มีเพียง 10 คุณลักษณะ และคุณลักษณะข้อมูลทั่วไปจะปรากฏเพียงคุณลักษณะเดียว คือ ข้อมูลคณะของนิสิต

คำสำคัญ : ระบบการจัดการเรียนรู้ออนไลน์, การทำนาย, ผลการเรียน, การเรียนรู้ของเครื่อง

| | |
|---|---|
| Title | STUDENT PERFORMANCE PREDICTION IN AN ONLINE LEARNING MANAGEMENT SYSTEM USING MACHINE LEARNING |
| Author | PARIYANUCH PRASERTSIRIKUL |
| Degree | MASTER OF SCIENCE |
| Academic Year | 2021 |
| Thesis Advisor | Dr. Sirisup Laohakiat |

This research aims to study student performance prediction in an online learning management system using a machine learning model by using the learning dataset from Moodle, an online learning platform. In order to study the impact of the features on prediction performance, the features were divided into two groups based on their characteristics, i.e., the features related to the system usage behavior of the students and the features related to the scores obtained by the students. Then, four sets of datasets were created using the combination of these two groups of features to study the impact of the features on the predictive performance. Several well-known machine learning models were experimented with including Logistic Regression, Naïve Bayes, K-Nearest Neighbor, Support Vector Machines, Decision Tree, Random Forest and XGBoost. The imbalanced datasets were handled by the synthetic minority oversampling technique (SMOTE). The results showed that the XGBoost model yielded the highest prediction accuracy at 83.95%. The models trained by features related to score alone yielded better performance than those trained by the features related to system usage behavior, but using the features from both groups together yielded the best performance. Moreover, the researcher found that feature selection can improve the performance of low-complexity models, while it impaired the performance of high-complexity models.

Keyword : Online Learning Management System, Prediction, Student Performance, Machine Learning

# กิตติกรรมประกาศ

ปริญญานิพนธ์นี้สำเร็จลุล่วงไปด้วยความกรุณาของ อาจารย์ ดร.ศิริสรรพ เหล่าหะเกียรติ อาจารย์ที่ปรึกษาหลัก ขอกราบขอบพระคุณท่านที่ได้เสียสละเวลาอันมีค่าในการให้ข้อเสนอแนะ ให้คำปรึกษา และคำแนะนำที่เป็นประโยชน์อย่างยิ่ง ทำให้ผู้วิจัยสามารถดำเนินการวิจัยจนสำเร็จสมบูรณ์

ขอกราบขอบพระคุณ คณะกรรมการสอบปริญญานิพนธ์ ที่ได้กรุณาให้คำแนะนำ ข้อเสนอแนะ และการตรวจสอบ ทำให้การจัดทำปริญญานิพนธ์นี้มีความสมบูรณ์ยิ่งขึ้น

ขอกราบขอบพระคุณ บัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ สำหรับทุนสนับสนุนการเข้าร่วมประชุมและการนำเสนอผลงานของนิสิตบัณฑิตศึกษา ทำให้มีโอกาสเผยแพร่การนำเสนอผลงานวิจัย และได้เข้าร่วมฟังผลงานของนักวิจัยในสาขาที่ศึกษา

ขอกราบขอบพระคุณ สำนักนวัตกรรมการเรียนรู้ มหาวิทยาลัยศรีนครินทรวิโรฒ ที่ได้กรุณาให้ความอนุเคราะห์แก่ผู้วิจัยได้จัดเก็บข้อมูลสำหรับการทำปริญญานิพนธ์นี้ ตลอดจนสนับสนุนข้อมูลการเรียนการสอนของนิสิต

ปริยานุช ประเสริฐสิริกุล

# สารบัญ

หน้า


บทคัดย่อภาษาไทย ..... ง

บทคัดย่อภาษาอังกฤษ ..... จ

กิตติกรรมประกาศ ..... ฉ

สารบัญ ..... ช

สารบัญตาราง ..... ฎ

สารบัญรูปภาพ ..... ฐ

บทที่ 1 บทนำ ..... 1

ความเป็นมาและความสำคัญของปัญหา ..... 1

คำถามงานวิจัย ..... 2

วัตถุประสงค์ของการวิจัย ..... 3

ขอบเขตการวิจัย ..... 3

ประชากรที่ใช้ในการวิจัย ..... 3

กลุ่มตัวอย่างที่ใช้ในการวิจัย ..... 3

ตัวแปรที่ศึกษา ..... 3

กรอบแนวคิดในการวิจัย ..... 5

สมมติฐานการวิจัย ..... 5

ข้อจำกัดในการวิจัย ..... 5

ประโยชน์ที่คาดว่าจะได้รับจากการวิจัย ..... 6

บทที่ 2 ทฤษฎี และงานวิจัยที่เกี่ยวข้อง ..... 7

ระบบการจัดการเรียนรู้ (Learning Management System: LMS) ..... 7

พฤติกรรมการเรียนรู้ออนไลน์ (Online Learning Behavior) ..... 8

การทำเหมืองข้อมูล (Data Mining) ..... 9
การเรียนรู้ของเครื่อง (Machine Learning) ..... 11
การทำวิศวกรรมคุณลักษณะ (Feature Engineering) ..... 14
การจัดการความไม่สมดุลของข้อมูล (Handling Imbalanced Data) ..... 16
การสร้างแบบจำลองการจำแนกประเภท (Classification Modeling Techniques) ..... 17
1. แบบจำลอง Logistic Regression ..... 17
2. แบบจำลอง Naïve Bayes ..... 19
3. แบบจำลอง Support Vector Machine ..... 19
4. แบบจำลอง K-Nearest Neighbor ..... 23
5. แบบจำลอง Decision Tree ..... 24
6. แบบจำลอง Random Forest ..... 26
7. แบบจำลอง XGBoost ..... 28
การคัดเลือกคุณลักษณะ (Feature Selection) ..... 29
การวัดประสิทธิภาพของแบบจำลอง (Model Evaluation) ..... 30
งานวิจัยที่เกี่ยวข้อง (Literature Review) ..... 32
บทที่ 3 วิธีการดำเนินงานวิจัย ..... 42
แผนการดำเนินงานวิจัย ..... 42
ภาพรวมของกระบวนการทำงาน ..... 44
การทำความเข้าใจในธุรกิจ ..... 46
การทำความเข้าใจข้อมูล ..... 54
การเก็บรวบรวมข้อมูลและการเตรียมข้อมูล ..... 56
การทำวิศวกรรมคุณลักษณะ ..... 56
การสำรวจข้อมูล ..... 66

การจัดการความไม่สมดุลของข้อมูล ..... 79
การสร้างแบบจำลองการจำแนกประเภท ..... 80
การคัดเลือกคุณลักษณะ ..... 80
การวัดประสิทธิภาพของแบบจำลอง ..... 89
บทที่ 4 ผลการดำเนินงานวิจัย ..... 90
ผลลัพธ์จากการเปรียบเทียบระหว่างแบบจำลองการทำนายผลการเรียนที่ใช้ชุดข้อมูลที่ 1 ข้อมูลบันทึกการใช้งาน คะแนนกิจกรรมในห้องเรียนออนไลน์ และคะแนนกลางภาค โดยใช้ข้อมูลคุณลักษณะทั้งหมด และแบบจำลองที่มีการคัดเลือกคุณลักษณะ ..... 91
ผลลัพธ์จากการเปรียบเทียบระหว่างแบบจำลองการทำนายผลการเรียนที่ใช้ชุดข้อมูลที่ 2 ข้อมูลบันทึกการใช้งาน และคะแนนกิจกรรมในห้องเรียนออนไลน์ โดยใช้ข้อมูลคุณลักษณะทั้งหมด และแบบจำลองที่มีการคัดเลือกคุณลักษณะ ..... 92
ผลลัพธ์จากการเปรียบเทียบระหว่างแบบจำลองการทำนายผลการเรียนที่ใช้ชุดข้อมูลที่ 3 ข้อมูลบันทึกการใช้งาน โดยใช้ข้อมูลคุณลักษณะทั้งหมด และแบบจำลองที่มีการคัดเลือกคุณลักษณะ ..... 93
ผลลัพธ์จากการเปรียบเทียบระหว่างแบบจำลองการทำนายผลการเรียนที่ใช้ชุดข้อมูลที่ 4 คะแนนกิจกรรมในห้องเรียนออนไลน์ และคะแนนกลางภาค โดยใช้ข้อมูลคุณลักษณะทั้งหมด และแบบจำลองที่มีการคัดเลือกคุณลักษณะ ..... 94
ผลลัพธ์จากการเปรียบเทียบระหว่างแบบจำลองการทำนายผลการเรียนที่ใช้ชุดข้อมูล 4 ชุด โดยใช้ข้อมูลคุณลักษณะทั้งหมด และแบบจำลองที่มีการคัดเลือกคุณลักษณะ ..... 95
บทที่ 5 สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ ..... 97
สรุปผลการวิจัย ..... 97
อภิปรายผลการวิจัย ..... 98
ข้อเสนอแนะ ..... 103
บรรณานุกรม ..... 104
ภาคผนวก ..... 110

ประวัติผู้เขียน..... 124

# สารบัญตาราง

**หน้า**

ตาราง 1 แสดงสรุปงานวิจัยที่เกี่ยวข้องทั้งหมดจำนวน 10 เรื่อง ..... 37

ตาราง 2 แสดงแผนการดำเนินงานวิจัย ..... 42

ตาราง 3 แสดงรายละเอียดเงื่อนไขของกิจกรรมการเรียนการสอนในระบบ Moodle ..... 46

ตาราง 4 แสดงเกณฑ์การพิจารณาตัดเกรด ..... 52

ตาราง 5 แสดงเกณฑ์การประเมินผลการเรียนรู้ ..... 52

ตาราง 6 การจำแนกคุณลักษณะที่ใช้ในการวิจัย ..... 54

ตาราง 7 แสดงรายละเอียดคำอธิบายของชุดข้อมูล ..... 63

ตาราง 8 แสดงการจำแนกกลุ่มเกรดตามระดับผลการเรียน ..... 67

ตาราง 9 แสดงการเรียงลำดับคะแนนความสำคัญของชุดข้อมูลสำหรับกรณีที่ 1: ข้อมูลบันทึกการใช้งานระบบ, คะแนนกิจกรรมในห้องเรียนออนไลน์ (30%) และคะแนนกลางภาค (30%) ที่คำนวณจากแบบจำลอง XGBoost ..... 81

ตาราง 10 แสดงการเรียงลำดับคะแนนความสำคัญของชุดข้อมูลสำหรับกรณีที่ 2: ข้อมูลบันทึกการใช้งานระบบ และคะแนนกิจกรรมในห้องเรียนออนไลน์ (30%) ที่คำนวณจากแบบจำลอง XGBoost ..... 83

ตาราง 11 แสดงการเรียงลำดับคะแนนความสำคัญของชุดข้อมูลในกรณีที่ 3: ข้อมูลบันทึกการใช้งานระบบ ที่คำนวณจากแบบจำลอง XGBoost ..... 85

ตาราง 12 แสดงการเรียงลำดับคะแนนความสำคัญของชุดข้อมูลในกรณีที่ 4: คะแนนกิจกรรมในห้องเรียนออนไลน์ และคะแนนสอบกลางภาค ที่คำนวณจากแบบจำลอง XGBoost ..... 87

ตาราง 13 แสดงการเปรียบเทียบความแม่นยำสำหรับชุดข้อมูลที่ตกต่างกัน 4 ชุด ของแต่ละแบบจำลอง ..... 96

ตาราง 14 แสดงค่า Prediction และค่า Recall ของแบบจำลอง XGBoost ของข้อมูลชุดที่ 1 ข้อมูลคุณลักษณะกลุ่มคะแนน และคุณลักษณะกลุ่มพฤติกรรมการใช้งานระบบ โดยใช้คุณลักษณะครบถ้วน ..... 98

ตาราง 15 แสดงจำนวนคุณลักษณะก่อนและหลังจากทำการคัดเลือกคุณลักษณะ..... 101

ตาราง 16 แสดงคุณลักษณะที่สำคัญ 30 อันดับแรกที่ได้จากแบบจำลอง XGBoost ที่ใช้ชุดข้อมูลที่ 1 โดยใช้คุณลักษณะครบถ้วน ..... 101

# สารบัญรูปภาพ

**หน้า**

ภาพประกอบ 1 แสดงการเปรียบเทียบแพลตฟอร์มการเรียนรู้ที่แตกต่างกันในระบบการจัดการเรียนรู้ออนไลน์ โดยวัดคะแนนโหวต (แกน x) และคะแนนความนิยม (แกน y) ..... 8

ภาพประกอบ 2 แสดงกระบวนการวิเคราะห์ข้อมูลด้วย CRISP-DM ..... 10

ภาพประกอบ 3 แสดงตัวอย่างการทำงานของการเรียนรู้ของเครื่องแบบมีผู้สอน ..... 12

ภาพประกอบ 4 แสดงตัวอย่างการทำงานของการเรียนรู้ของเครื่องแบบไม่มีผู้สอน ..... 12

ภาพประกอบ 5 แสดงตัวอย่างการทำงานของการเรียนรู้ของเครื่องแบบกึ่งมีผู้สอน โดยนำรูปวงกลมสีฟ้าและสามเหลี่ยมสีแดงเป็นข้อมูลที่มีป้ายกำกับ 2 คลาส และวงกลมสีเขียวเป็นข้อมูลที่ไม่มีป้ายกำกับ ให้กับเครื่องเรียนรู้ว่า วงกลมสีเขียวอยู่คลาสใด ..... 13

ภาพประกอบ 6 แสดงตัวอย่างการทำงานของการเรียนรู้ของเครื่องแบบเสริมกำลัง โดยสร้างเขาวงกตให้หนูวิ่งเข้าออกเขาวงกตจากการลองผิดลองถูก ..... 14

ภาพประกอบ 7 แสดงกระบวนการทำวิศวกรรมคุณลักษณะ โดยนำข้อมูลหลายแหล่งเข้าด้วยกัน แล้วทำความสะอาดและเปลี่ยนแปลงโครงสร้างให้เป็นคุณลักษณะที่พร้อมเข้าสู่แบบจำลองการทำนาย ..... 14

ภาพประกอบ 8 แสดงตัวอย่างการทำงานของการจัดการข้อมูลที่ไม่สมดุลด้วยวิธี SMOTE โดยนำข้อมูลคลาสบวกที่เป็นสี่เหลี่ยมและคลาสลบที่เป็นวงกลมไปจัดการความสมดุล แล้วจะได้ข้อมูลเทียมที่เป็นรูปสามเหลี่ยมเพิ่มขึ้นมา ..... 17

ภาพประกอบ 9 แสดงการเปรียบเทียบก่อนและหลังการทำงานของแบบจำลอง Logistic Regression ..... 18

ภาพประกอบ 10 แสดงการทำงานของแบบจำลอง Support Vector Machine สำหรับการแบ่งข้อมูล 2 คลาส ได้แก่ รูปสี่เหลี่ยม (คลาสที่ 1) และรูปวงกลม (คลาสที่ 2) ..... 20

ภาพประกอบ 11 แสดงการเปรียบเทียบการทำงานของแบบจำลอง Support Vector Machine โดยใช้ข้อมูลที่สามารถแบ่ง 2 คลาสด้วยเส้นตรง และข้อมูลที่ไม่สามารถแบ่ง 2 คลาสด้วยเส้นตรง จากการกำหนดคลาสที่เป็นรูปวงกลมสีม่วง (คลาสที่ 1) และรูปวงกลมสีฟ้า (คลาสที่ 2) ..... 21

ภาพประกอบ 12 แสดงการเปรียบเทียบการทำงานของแบบจำลอง Support Vector Machine โดยใช้ข้อมูลที่ถูกแบ่งคลาสแบบ One versus all และข้อมูลที่ถูกแบ่งคลาสแบบ One versus one จากการกำหนดคลาสที่เป็นรูปวงกลมสีม่วง (คลาสที่ 1) รูปวงกลมสีฟ้า (คลาสที่ 2) และรูปวงกลมสีเหลือง (คลาสที่ 3) ..... 22

ภาพประกอบ 13 แสดงตัวอย่างการทำงานของแบบจำลอง K-Nearest Neighbor จากการกำหนดคลาสที่เป็นรูปสี่เหลี่ยม (คลาสที่ 1) และคลาสที่เป็นรูปสามเหลี่ยม (คลาสที่ 2) ..... 23

ภาพประกอบ 14 แสดงการทำงานของแบบจำลอง Decision Tree ในโครงสร้างของผังงาน ..... 24

ภาพประกอบ 15 แสดงการทำงานของแบบจำลอง Random Forest สำหรับการสร้างชุดข้อมูลด้วยเทคนิค Bootstrapping แต่ละต้น ..... 26

ภาพประกอบ 16 แสดงการทำงานของแบบจำลอง Random Forest สำหรับการสุ่มเลือกคุณลักษณะของชุดข้อมูลที่ได้จากการใช้วิธี Bootstrapping ..... 27

ภาพประกอบ 17 แสดงการทำงานของแบบจำลอง Random Forest สำหรับการทำงานร่วมกันของแบบจำลอง Decision Tree หลายต้น ..... 27

ภาพประกอบ 18 แสดงตัวอย่างการทำงานของแบบจำลอง XGBoost จากการเรียนรู้ข้อมูล 3 ครั้ง โดยแบ่งข้อมูลเป็น 2 คลาส ได้แก่ คลาสบวก และคลาสลบ ..... 29

ภาพประกอบ 19 แสดงกระบวนการคัดเลือกคุณลักษณะด้วย Wrapped Method ..... 30

ภาพประกอบ 20 แสดงตัวอย่างการวัดประสิทธิภาพของแบบจำลองการทำนายอีเมลว่าเป็นสแปมหรือไม่ โดยใช้ Confusion Matrix สำหรับข้อมูลที่มี 2 คลาส ..... 30

ภาพประกอบ 21 แสดงตัวอย่างการวัดประสิทธิภาพของแบบจำลองการทำนายอีเมลว่าอยู่ประเภทใด โดยใช้ Confusion Matrix สำหรับข้อมูลที่มี 3 คลาส ..... 32

ภาพประกอบ 22 แสดงแผนผังกระบวนการทำงาน ..... 45

ภาพประกอบ 23 แสดงหน้าแรกการเข้าสู่ระบบ Moodle รายวิชา SWU 141 ชีวิตในโลกดิจิทัล บนเว็บไซต์ https://course.ilc.swu.ac.th ..... 48

ภาพประกอบ 24 แสดงการเข้าถึงสื่อการเรียนรู้ และแบบทดสอบ จากระบบ Moodle ..... 49

ภาพประกอบ 25 แสดงการเข้าถึงงานมอบหมายจากระบบ Moodle ..... 49

ภาพประกอบ 26 แสดงรายการบันทึกกิจกรรมการใช้งานของนิสิตจากระบบ Moodle ..... 50

ภาพประกอบ 27 แสดงคะแนนแบบทดสอบของนิสิตจากระบบ Moodle ..... 51

ภาพประกอบ 28 แสดงไฟล์ข้อมูลบันทึกกิจกรรมการใช้งานระบบ Moodle รายวิชา SWU 141 ชีวิตในโลกดิจิทัล จากเว็บไซต์ https://course.ilc.swu.ac.th ..... 54

ภาพประกอบ 29 แสดงตัวอย่างบันทึกการใช้งานระบบของนิสิต ..... 55

ภาพประกอบ 30 แสดงจำนวนครั้งที่คลิกเข้าสู่ระบบ ..... 56

ภาพประกอบ 31 แสดงจำนวนครั้งที่คลิกเข้าถึงสื่อการเรียนรู้ ..... 57

ภาพประกอบ 32 แสดงจำนวนวันที่ชมสื่อการเรียนรู้ ..... 57

ภาพประกอบ 33 แสดงจำนวนวันระหว่างวันแรกที่เรียน และวันแรกที่ชมสื่อการเรียนรู้ ..... 57

ภาพประกอบ 34 แสดงจำนวนเรื่องที่ชมสื่อการเรียนรู้ ..... 57

ภาพประกอบ 35 แสดงจำนวนครั้งที่คลิกเข้าถึงแบบทดสอบ ..... 58

ภาพประกอบ 36 แสดงจำนวนวันระหว่างวันที่ทำแบบทดสอบครั้งแรก และวันสุดท้ายการทำแบบทดสอบ ..... 58

ภาพประกอบ 37 แสดงจำนวนครั้งที่ทำแบบทดสอบซ้ำ ..... 58

ภาพประกอบ 38 แสดงจำนวนเรื่องที่ทำแบบทดสอบ ..... 59

ภาพประกอบ 39 แสดงระยะเวลาการทำแบบทดสอบ ครั้งที่ 1 ถึง ครั้งที่ 3 ..... 59

ภาพประกอบ 40 แสดงระยะเวลารวมการทำแบบทดสอบรวมทุกครั้ง ..... 59

ภาพประกอบ 41 แสดงระยะเวลาเฉลี่ยการทำแบบทดสอบรวมทุกครั้ง ..... 60

ภาพประกอบ 42 แสดงจำนวนครั้งที่คลิกเข้าถึงงานมอบหมาย ..... 60

ภาพประกอบ 43 แสดงจำนวนวันระหว่างวันแรกที่ส่งงานมอบหมาย และวันสุดท้ายการส่งงานมอบหมาย ..... 60

ภาพประกอบ 44 แสดงคะแนนกิจกรรมในห้องเรียนออนไลน์ ..... 61

ภาพประกอบ 45 แสดงคะแนนการเข้าเรียน ..... 61

ภาพประกอบ 46 แสดงคะแนนแบบทดสอบ ..... 61

ภาพประกอบ 47 แสดงคะแนนการสอบกลางภาค ..... 62

ภาพประกอบ 48 แสดงคะแนนงานมอบหมาย ..... 62
ภาพประกอบ 49 แสดงจำนวนนิสิตจำแนกเกรด ..... 67
ภาพประกอบ 50 แสดงจำนวนนิสิตจำแนกเพศ ..... 68
ภาพประกอบ 51 แสดงจำนวนนิสิตจำแนกคณะ ..... 68
ภาพประกอบ 52 แสดงความสัมพันธ์ระหว่างจำนวนครั้งที่คลิกเข้าสู่ระบบกับกลุ่มเกรด ..... 69
ภาพประกอบ 53 แสดงความสัมพันธ์ระหว่างจำนวนครั้งที่คลิกเข้าถึงสื่อการเรียนรู้รวมทุกเรื่องกับกลุ่มเกรด ..... 69
ภาพประกอบ 54 แสดงความสัมพันธ์ระหว่างจำนวนวันที่ชมสื่อการเรียนรู้รวมทุกเรื่องกับกลุ่มเกรด ..... 70
ภาพประกอบ 55 แสดงความสัมพันธ์ระหว่างจำนวนวันระหว่างวันแรกที่เรียนและวันแรกที่ชมสื่อการเรียนรู้กับกลุ่มเกรด ..... 70
ภาพประกอบ 56 แสดงความสัมพันธ์ระหว่างจำนวนเรื่องที่ชมสื่อการเรียนรู้กับกลุ่มเกรด ..... 71
ภาพประกอบ 57 แสดงความสัมพันธ์ระหว่างจำนวนครั้งที่คลิกเข้าถึงแบบทดสอบรวมทุกเรื่องกับกลุ่มเกรด ..... 72
ภาพประกอบ 58 แสดงความสัมพันธ์ระหว่างจำนวนวันระหว่างวันที่ทำแบบทดสอบครั้งแรกและวันสุดท้ายการทำแบบทดสอบกับกลุ่มเกรด ..... 72
ภาพประกอบ 59 แสดงความสัมพันธ์ระหว่างจำนวนครั้งที่ทำแบบทดสอบซ้ำกับกลุ่มเกรด ..... 73
ภาพประกอบ 60 แสดงความสัมพันธ์ระหว่างจำนวนเรื่องที่ทำแบบทดสอบกับกลุ่มเกรด ..... 73
ภาพประกอบ 61 แสดงความสัมพันธ์ระหว่างระยะเวลารวมการทำแบบทดสอบรวมทุกครั้ง (นาที) และทุกเรื่อง กับกลุ่มเกรด ..... 74
ภาพประกอบ 62 แสดงความสัมพันธ์ระหว่างระยะเวลาเฉลี่ยการทำแบบทดสอบรวมทุกครั้งและทุกเรื่อง (นาที) กับกลุ่มเกรด ..... 74
ภาพประกอบ 63 แสดงความสัมพันธ์ระหว่างจำนวนครั้งที่คลิกเข้าถึงงานมอบหมายกับกลุ่มเกรด ..... 75

ภาพประกอบ 64 แสดงความสัมพันธ์ระหว่างจำนวนวันระหว่างวันแรกที่ส่งงานมอบหมายและวันสุดท้ายการส่งงานมอบหมายกับกลุ่มเกรด ..... 75

ภาพประกอบ 65 แสดงความสัมพันธ์ระหว่างคะแนนกิจกรรมในห้องเรียนออนไลน์รวมทุกเรื่องกับกลุ่มเกรด ..... 76

ภาพประกอบ 66 แสดงความสัมพันธ์ระหว่างคะแนนการเข้าเรียนกับกลุ่มเกรด ..... 77

ภาพประกอบ 67 แสดงความสัมพันธ์ระหว่างคะแนนแบบทดสอบรวมทุกเรื่องกับกลุ่มเกรด ..... 77

ภาพประกอบ 68 แสดงความสัมพันธ์ระหว่างคะแนนสอบกลางภาคกับกลุ่มเกรด ..... 78

ภาพประกอบ 69 แสดงความสัมพันธ์ระหว่างคะแนนงานมอบหมายรวมทุกส่วนกับกลุ่มเกรด ... 78

ภาพประกอบ 70 แสดงจำนวนนิสิตแต่ละระดับผลการเรียน ก่อนการทำ SMOTE ..... 79

ภาพประกอบ 71 แสดงจำนวนนิสิตแต่ละระดับผลการเรียน หลังการทำ SMOTE ..... 79

ภาพประกอบ 72 แสดงความสัมพันธ์ระหว่างชุดข้อมูลสำหรับกรณีที่ 1 และระดับผลการเรียน โดยวัดคะแนนความสำคัญจากแบบจำลอง XGBoost ..... 82

ภาพประกอบ 73 แสดงความสัมพันธ์ระหว่างชุดข้อมูลสำหรับกรณีที่ 2 และระดับผลการเรียน โดยวัดคะแนนความสำคัญจากแบบจำลอง XGBoost ..... 84

ภาพประกอบ 74 แสดงความสัมพันธ์ระหว่างชุดข้อมูลสำหรับกรณีที่ 3 และระดับผลการเรียน โดยวัดคะแนนความสำคัญจากแบบจำลอง XGBoost ..... 86

ภาพประกอบ 75 แสดงความสัมพันธ์ระหว่างชุดข้อมูลสำหรับกรณีที่ 4 และระดับผลการเรียน โดยวัดคะแนนความสำคัญจากแบบจำลอง XGBoost ..... 88

ภาพประกอบ 76 แสดงการเปรียบเทียบความแม่นยำของแบบจำลองการทำนายผลการเรียน โดยใช้ชุดข้อมูลที่ 1 ข้อมูลบันทึกการใช้งาน คะแนนกิจกรรมในห้องเรียนออนไลน์ และคะแนนกลางภาค โดยใช้ข้อมูลคุณลักษณะทั้งหมด และแบบจำลองที่มีการคัดเลือกคุณลักษณะ ..... 92

ภาพประกอบ 77 แสดงการเปรียบเทียบความแม่นยำของแบบจำลองการทำนายผลการเรียน โดยใช้ชุดข้อมูลที่ 2 ข้อมูลบันทึกการใช้งาน และคะแนนกิจกรรมในห้องเรียนออนไลน์ โดยใช้ข้อมูลคุณลักษณะทั้งหมด และแบบจำลองที่มีการคัดเลือกคุณลักษณะ ..... 93

ภาพประกอบ 78 แสดงการเปรียบเทียบความแม่นยำของแบบจำลองการทำนายผลการเรียน โดยใช้ชุดข้อมูลที่ 3 ข้อมูลบันทึกการใช้งาน โดยใช้ข้อมูลคุณลักษณะทั้งหมด และแบบจำลองที่มีการคัดเลือกคุณลักษณะ ..... 94

ภาพประกอบ 79 แสดงการเปรียบเทียบความแม่นยำของแบบจำลองการทำนายผลการเรียน โดยใช้ชุดข้อมูลที่ 4 คะแนนกิจกรรมในห้องเรียนออนไลน์ และคะแนนกลางภาค โดยใช้ข้อมูลคุณลักษณะทั้งหมด และแบบจำลองที่มีการคัดเลือกคุณลักษณะ ..... 95

# บทที่ 1
# บทนำ

## ความเป็นมาและความสำคัญของปัญหา

ปัจจุบันสถาบันการศึกษาหลายแห่งได้นำระบบการจัดการเรียนรู้ (Learning Management System: LMS) มาใช้ในการจัดการเรียนการสอน เนื่องด้วยระบบดังกล่าวนี้มีเทคโนโลยีที่ช่วยอำนวยความสะดวกในการจัดการสอนให้แก่ผู้สอน และช่วยสนับสนุนการเรียนรู้ของผู้เรียนได้เป็นอย่างดี สร้างบรรยากาศเหมือนห้องเรียนเสมือน (Virtual Classroom) ที่ผู้สอนและผู้เรียนสามารถมีปฏิสัมพันธ์กันได้ ระบบการจัดการเรียนรู้นี้ โดยทั่วไปแล้วมีลักษณะเป็นแอปพลิเคชัน (Application) ที่นำเสนอชุดเครื่องมือและโปรแกรมการเรียนรู้ทางอินเทอร์เน็ต มีฟังก์ชันการทำงานที่สนับสนุนการสอนและการเรียนรู้ของผู้เรียนเป็นจำนวนมาก โดยเปรียบเสมือนสะพานที่เชื่อมระหว่างผู้เรียนและผู้สอน ประโยชน์ที่ได้รับจากการนำระบบการจัดการเรียนรู้มาใช้นั้นสามารถกล่าวได้ดังนี้ 1) อาจารย์สามารถเข้าถึงสื่อการเรียนรู้ได้โดยง่าย สามารถตรวจสอบผลคะแนนจากการทำแบบทดสอบของผู้เรียนได้ สามารถติดตามความก้าวหน้าในการเรียนของผู้เรียนได้เป็นรายบุคคล 2) ผู้เรียนสามารถเรียกใช้งานระบบได้ทุกที่ทุกเวลา สามารถส่งงานมอบหมาย หรือทำแบบทดสอบได้แบบออนไลน์ และทราบผลคะแนนได้โดยทันที

รูปแบบการจัดการเรียนการสอนในระบบการจัดการเรียนรู้ เป็นการเรียนการสอนแบบออนไลน์ (Online Learning) ซึ่งไม่ได้เป็นการเรียนแบบปกติที่เผชิญหน้ากันระหว่างผู้สอนและผู้เรียน (Face-to-Face Learning) อาจทำให้เกิดการปฏิสัมพันธ์ที่น้อยลงระหว่างผู้สอนและผู้เรียน ตลอดจนนำไปสู่การติดตามความก้าวหน้าการเรียนรู้ของผู้เรียนโดยผู้สอนได้ยาก และผลการเรียนของผู้เรียนมีประสิทธิภาพน้อยลง ในด้านการใช้ประโยชน์เพื่อวิเคราะห์ข้อมูลทางการศึกษา ข้อมูลที่จัดเก็บอยู่ในระบบการจัดการเรียนรู้นี้ เช่น สถิติการเข้าใช้งานระบบ คะแนนงานมอบหมาย คะแนนแบบทดสอบ เป็นต้น ข้อมูลบันทึกการใช้งานกิจกรรมที่หลากหลายของผู้ใช้งาน รวมทั้งผู้ดูแลระบบ ผู้สอน และผู้เรียน จะถูกจัดเก็บในระบบเป็นจำนวนมาก และมีปริมาณมากขึ้นในทุกวันจนกลายเป็นข้อมูลขนาดใหญ่ (Big data) สามารถนำมาใช้เพื่อวิเคราะห์พฤติกรรมของผู้เรียนในมิติต่าง ๆ ได้ โดยการนำข้อมูลเหล่านี้มาวิเคราะห์เพื่อทำนายผลการเรียนของผู้เรียนจากข้อมูลพฤติกรรมการเรียนรู้ล่วงหน้า เป็นแนวทางหนึ่งที่จะช่วยนำเสนอข้อมูลที่เป็นประโยชน์ให้แก่ผู้สอน เพื่อให้ผู้สอนสามารถช่วยเหลือผู้เรียนได้อย่างเจาะจง และทำให้การเรียนการสอนเกิดประสิทธิภาพสูงขึ้นได้

จากแนวทางการวิเคราะห์ข้อมูลดังกล่าว ต้องอาศัยเทคนิคขั้นสูงในการวิเคราะห์ข้อมูลขนาดใหญ่ที่มีความซับซ้อน เปรียบเสมือนการขุดหาสินแร่ที่อยู่ในเหมือง ซึ่งเรียกได้ว่า เป็นการทำเหมืองข้อมูล (Data Mining) นั่นเอง การทำเหมืองข้อมูลนี้เป็นกระบวนการดึงรูปแบบที่ไม่เคยเห็นมาก่อนอยู่ในชุดข้อมูลขนาดใหญ่ สามารถนำเทคนิคการทำเหมืองข้อมูลมาประยุกต์ใช้ในด้านการศึกษา วิธีการนี้เรียกว่า การทำเหมืองข้อมูลเพื่อการศึกษา (Educational Data Mining) ช่วยให้ผู้สอนรู้จักตัวผู้เรียนและพฤติกรรมของผู้เรียนได้เป็นอย่างดี เทคนิคหนึ่งที่กล่าวถึงกันมากในการทำเหมืองข้อมูล เรียกว่า “การเรียนรู้ของเครื่อง (Machine Learning)” ผู้วิจัยอื่นได้ศึกษาวิจัยที่เกี่ยวข้องการทำเหมืองข้อมูลการศึกษา โดยใช้เทคนิคการเรียนรู้ของเครื่อง ได้แก่ การทำนายผลการเรียนจากพฤติกรรมการส่งงานมอบหมายของผู้เรียนในระบบการจัดการเรียนรู้ (Hooshyar และ Pedaste, 2019), การทำนายการลาออกกลางคันของผู้เรียนจากข้อมูลทั่วไป และผลการเรียนของผู้เรียน (Tan และ Shao, 2015) และการทำนายผลการเรียนเฉลี่ยสะสม (GPA) จากผลการเรียนของรายวิชาและข้อมูลทั่วไป (Almasri, Alkhawaldeh, และ Çelebi, 2020)

ทั้งนี้ผู้วิจัยจึงสนใจศึกษาและวิเคราะห์พฤติกรรมการใช้งานระบบการจัดการเรียนรู้ออนไลน์ของนิสิต เพื่อทำนายผลการเรียนของนิสิต โดยใช้แบบจำลองการจำแนกประเภท (Classification) ในเทคนิคการเรียนรู้ของเครื่อง (Machine Learning) ได้แก่ Logistic Regression, Naïve Bayes, Support Vector Machine, K-Nearest Neighbor, Decision Tree, Random Forest, Extreme Gradient Boosting (XGBoost) เป็นต้น และมุ่งเน้นเทคนิคการทำวิศวกรรมคุณลักษณะ (Feature Engineering) การจัดการความไม่สมดุลของข้อมูล (Handling Imbalanced Data) ด้วยวิธีการสุ่มตัวอย่างเพิ่มกลุ่มน้อยด้วยการสร้างตัวอย่างสังเคราะห์ (Synthetic Minority Oversampling Technique: SMOTE) และการคัดเลือกคุณลักษณะ (Feature Selection) การวัดประสิทธิภาพของแบบจำลองการจำแนกประเภท โดยพิจารณาค่าวัดต่าง ๆ ในเมทริกซ์ความสับสน (Confustion Matrix) เพื่อหาแบบจำลองการทำนายผลการเรียนของนิสิตที่ดีที่สุด

**คำถามงานวิจัย**

การสร้างแบบจำลองการจำแนกประเภทสำหรับการทำนายผลการเรียนของนิสิต จากพฤติกรรมการใช้งานระบบของนิสิต และคะแนนของนิสิต ในช่วงเวลาการเรียนของครึ่งภาคการศึกษา มีประสิทธิภาพการทำนายผลการเรียนที่ดีกว่าข้อมูลที่ใช้เพียงคะแนนหรือไม่ และแบบจำลองที่สร้างมีประสิทธิภาพการทำนายมากน้อยเพียงใด

**วัตถุประสงค์ของการวิจัย**

การวิจัยนี้ผู้วิจัยได้กำหนดเป้าหมายของการวิจัยไว้ ดังนี้

1. เพื่อสร้างแบบจำลองการจำแนกประเภทสำหรับการทำนายผลการเรียนของนิสิตจากพฤติกรรมการใช้งานระบบของนิสิต และคะแนนของนิสิต

2. เพื่อศึกษาและวิเคราะห์พฤติกรรมการใช้งานระบบของนิสิต

3. เพื่อจัดการความไม่สมดุลของข้อมูล โดยใช้เทคนิคการสุ่มตัวอย่างเพิ่มกลุ่มน้อยด้วยการสร้างตัวอย่างสังเคราะห์

4. เพื่อหาคุณลักษณะที่ส่งผลต่อการทำนายผลการเรียนของนิสิต โดยใช้เทคนิคการทำวิศวกรรมคุณลักษณะ และเทคนิคการคัดเลือกคุณลักษณะ

5. เพื่อหาแบบจำลองการจำแนกประเภทสำหรับการทำนายผลการเรียนของนิสิตที่ดีที่สุดจากการเปรียบเทียบประสิทธิภาพของแบบจำลองต่าง ๆ

6. เพื่อให้นิสิตมีผลการเรียนที่ดีขึ้น โดยนำผลการทำนายล่วงหน้าในระยะเวลาครึ่งแรกของภาคการศึกษาให้อาจารย์ช่วยแนะแนวให้นิสิตที่มีโอกาสได้เกรดไม่ค่อยดีนัก และทำให้นิสิตมีเวลาปรับปรุงการเรียนรู้ของตนเองมากขึ้น

**ขอบเขตการวิจัย**

**ประชากรที่ใช้ในการวิจัย**

ประชากรที่ใช้ในการวิจัยคือ นิสิตระดับปริญญาตรี ชั้นปีที่ 1 มหาวิทยาลัยศรีนครินทรวิโรฒ ที่ลงทะเบียนเรียนรายวิชา SWU 141 ชีวิตโลกดิจิทัล ภาคการศึกษาที่ 2 ปีการศึกษา 2563 มีจำนวนนิสิตที่ลงทะเบียน 2,894 คน โดยมาจากจำนวนห้องเรียน 29 ห้อง

**กลุ่มตัวอย่างที่ใช้ในการวิจัย**

กลุ่มตัวอย่างที่ใช้ในการวิจัยจำนวน 405 คน มาจากจำนวนห้องเรียน 3 ห้อง เป็นกลุ่มของนิสิตในอาจารย์ผู้สอน 2 ท่าน ที่ได้ให้ความอนุเคราะห์ข้อมูลต่อการวิจัยชุดนี้

**ตัวแปรที่ศึกษา**

งานวิจัยครั้งนี้จะทำนายผลการเรียนของนิสิตที่ใช้งานระบบการจัดการเรียนรู้ออนไลน์ การกำหนดตัวแปรที่ใช้ในการวิจัยนี้ ผู้วิจัยจึงศึกษางานวิจัยที่เกี่ยวข้อง รายละเอียดกิจกรรมการเรียนการสอนที่กำหนดในระบบ Moodle รายวิชา SWU 141 ชีวิตในโลกดิจิทัล ภาคการศึกษา 2 ปีการศึกษา 2563 และเกณฑ์การประเมินผลการเรียนรู้สำหรับรายวิชานี้ และแบ่งตัวแปรที่ใช้ในการวิจัยออกเป็น 2 ประเภท คือ ตัวแปรอิสระ และตัวแปรตาม ดังนี้

1. ตัวแปรอิสระ ประกอบด้วย

1.1 ตัวแปรเชิงคุณภาพ (Qualitative Variable)

1.1.1 ข้อมูลทั่วไป (Almasri และคนอื่น ๆ, 2020; Amrieh, Hamtini, และ Aljarah, 2016; Ashfaq, Poolan Marikannan, และ Raheem, 2020; Ramaswami, Susnjak, Mathrani, และ Umer, 2020)

1.1.1.1 เพศ

1.1.1.2 คณะ

1.2 ตัวแปรเชิงปริมาณ (Quantitative Variable)

1.2.1 พฤติกรรมการใช้งานระบบของนิสิต (Ahmad, Anuar, และ Hassan, 2020; Amrieh และคนอื่น ๆ, 2016; Ashfaq และคนอื่น ๆ, 2020; Hooshyar และ Pedaste, 2019; Mueen, Zafar, และ Manzoor, 2016; Quinn และ Gray, 2019; Ramaswami และคนอื่น ๆ, 2020; Riestra-González, Paule-Ruíz, และ Ortin, 2021; Shrestha และ Pokharel, 2021)

1.2.1.1 จำนวนครั้งที่คลิกเข้าสู่ระบบ

1.2.1.2 จำนวนครั้งที่คลิกเข้าถึงสื่อการเรียนรู้

1.2.1.3 จำนวนวันที่ชมสื่อการเรียนรู้

1.2.1.4 จำนวนวันระหว่างวันแรกที่เรียนและวันแรกที่ชมสื่อการเรียนรู้

1.2.1.5 จำนวนเรื่องที่ชมสื่อการเรียนรู้

1.2.1.6 จำนวนครั้งที่คลิกเข้าถึงแบบทดสอบ

1.2.1.7 จำนวนวันระหว่างวันที่ทำแบบทดสอบครั้งแรกและวันสุดท้ายการทำแบบทดสอบ

1.2.1.8 จำนวนครั้งที่ทำแบบทดสอบซ้ำ

1.2.1.9 จำนวนเรื่องที่ทำแบบทดสอบ

1.2.1.10 ระยะเวลาการทำแบบทดสอบ

1.2.1.11 จำนวนครั้งที่คลิกเข้าถึงงานมอบหมาย

1.2.1.12 จำนวนวันระหว่างวันแรกที่ส่งงานมอบหมายและวันสุดท้ายการส่งงานมอบหมาย

1.2.2 คะแนนของนิสิต (Mueen และคนอื่น ๆ, 2016; Ramaswami และคนอื่น ๆ, 2020; Riestra-González และคนอื่น ๆ, 2021; Shrestha และ Pokharel, 2021)

1.2.2.1 คะแนนกิจกรรมในห้องเรียนออนไลน์

1.2.2.2 คะแนนการเข้าเรียน

1.2.2.3 คะแนนแบบทดสอบ

1.2.2.4 คะแนนกลางภาค

1.2.2.5 คะแนนงานมอบหมาย

2. ตัวแปรตาม คือ ระดับผลการเรียนของนิสิต

**กรอบแนวคิดในการวิจัย**

งานวิจัยนี้ศึกษาการวิเคราะห์พฤติกรรมการใช้งานระบบการจัดการเรียนรู้ออนไลน์ของนิสิต ที่ใช้ Moodle (Modular Object Oriented Dynamic Learning Environment) เป็นแพลตฟอร์มการเรียนรู้ออนไลน์ รายวิชา SWU 141 ชีวิตในโลกดิจิทัล จากสำนักนวัตกรรมการเรียนรู้ มหาวิทยาลัยศรีนครินทรวิโรฒ โดยใช้ข้อมูลการบันทึกการใช้งานระบบของนิสิตในช่วงครึ่งแรกของภาคการศึกษา ตั้งแต่วันที่ 7 มกราคม ถึง 5 มีนาคม 2564 คะแนนต่าง ๆ และข้อมูลทั่วไป เพื่อสร้างแบบจำลองการจำแนกประเภทสำหรับการทำนายผลการเรียนของนิสิต โดยใช้เทคนิคการเรียนรู้ของเครื่อง ได้แก่ Logistic Regression, Naïve Bayes, Support Vector Machine, K-Nearest Neighbor, Decision Tree, Random Forest, Extreme Gradient Boosting (XGBoost) และใช้เทคนิคการจัดข้อมูลด้วยการทำวิศวกรรมคุณลักษณะ การจัดการความไม่สมดุลของข้อมูลด้วยวิธีการสุ่มตัวอย่างเพิ่มกลุ่มน้อยด้วยการสร้างตัวอย่างสังเคราะห์ และการคัดเลือกคุณลักษณะ

**สมมติฐานการวิจัย**

1. แบบจำลองที่ใช้ข้อมูลพฤติกรรมการใช้งานระบบของนิสิต และคะแนน มีประสิทธิภาพที่ดีกว่า เมื่อเทียบกับแบบจำลองที่ใช้เพียงคะแนน

2. สามารถเลือกแบบจำลองที่มีประสิทธิภาพ โดยดูจากผลการประเมินด้วยข้อมูลสำหรับตรวจสอบ (test set)

**ข้อจำกัดในการวิจัย**

สำนักนวัตกรรมการเรียนรู้ มศว มีการนำสื่อการเรียนรู้ รายวิชา SWU 141 ชีวิตในโลกดิจิทัล ภาคการศึกษา 2 ปีการศึกษา 2563 เข้าสู่ระบบ Moodle ซึ่งสื่อการเรียนรู้นี้ไม่ได้เชื่อมต่อกับระบบโดยตรง เนื่องจากรูปแบบของสื่อการเรียนรู้เป็นลิงก์ที่อยู่ภายนอกระบบ เช่น YouTube, Microsoft OneDrive, E-book แต่ระบบมีการจัดเก็บข้อมูลการคลิกหัวข้อสื่อบนหน้าเว็บไซต์

ดังนั้น ยังคงมีความน่าเชื่อถือพอที่จะสามารถใช้ข้อมูลเหล่านี้ในการวิเคราะห์ข้อมูลเพื่อทำนายผลการเรียนของนิสิต

**ประโยชน์ที่คาดว่าจะได้รับจากการวิจัย**

1. แบบจำลองที่สร้างจากพฤติกรรมการใช้งานระบบของนิสิต สามารถทำนายผลการเรียนของนิสิต เพื่อเป็นประโยชน์ต่อการเรียนการสอนให้แก่อาจารย์และนิสิต ได้แก่ อาจารย์ให้คำแนะนำและช่วยเหลือนิสิตได้อย่างเจาะจง สามารถติดตามความก้าวหน้าในการเรียนของผู้เรียนได้เป็นรายบุคคล และวางแผนการสอนให้เหมาะสมกับนิสิต

2. สามารถวิเคราะห์พฤติกรรมการใช้งานระบบของนิสิต และนำผลการวิเคราะห์นี้ไปสร้างคุณลักษณะสำหรับการเข้าสู่แบบจำลองการทำนายผลการเรียนของนิสิต

3. สามารถจัดการความไม่สมดุลของข้อมูล โดยใช้เทคนิคการสุ่มตัวอย่างเพิ่มกลุ่มน้อยด้วยการสร้างตัวอย่างสังเคราะห์

4. ได้คุณลักษณะจำนวนน้อยที่ส่งผลต่อการทำนายผลการเรียนของนิสิต โดยใช้การทำวิศวกรรมคุณลักษณะ และการคัดเลือกคุณลักษณะ

5. ได้แบบจำลองการจำแนกประเภทสำหรับการทำนายผลการเรียนของนิสิตดีที่สุด โดยใช้เทคนิคการเรียนรู้ของเครื่อง จากการตัดสินใจเลือกแบบจำลองอย่างเหมาะสม และสามารถนำแบบจำลองนี้ไปใช้กับพฤติกรรมการใช้งานระบบของนิสิตทั้งหมดที่ลงทะเบียนรายวิชา SWU 141 ชีวิตในโลกดิจิทัล

6. สถาบันการศึกษาหลายแห่งที่มีการจัดการเรียนการสอนออนไลน์ โดยใช้งานระบบการจัดการเรียนรู้ สามารถนำแนวทางการวิจัยนี้ โดยวิเคราะห์พฤติกรรมการใช้งานของผู้เรียนที่ได้จากระบบ ได้แก่ จำนวนวันระหว่างวันแรกที่เรียนและวันแรกที่ชมสื่อการเรียนรู้ จำนวนวันระหว่างวันแรกที่ส่งงานมอบหมายและวันสุดท้ายการส่งงานมอบหมาย จำนวนครั้งที่ทำแบบทดสอบซ้ำ เพื่อนำพฤติกรรมเหล่านี้ร่วมกับคะแนนไปทำนายผลการเรียนที่แม่นยำสูงขึ้น

# บทที่ 2
# ทฤษฎี และงานวิจัยที่เกี่ยวข้อง

การศึกษาวิจัยนี้ ผู้วิจัยได้ศึกษาทฤษฎี และงานวิจัยที่เกี่ยวข้อง โดยมีการนำเสนอหัวข้อดังนี้

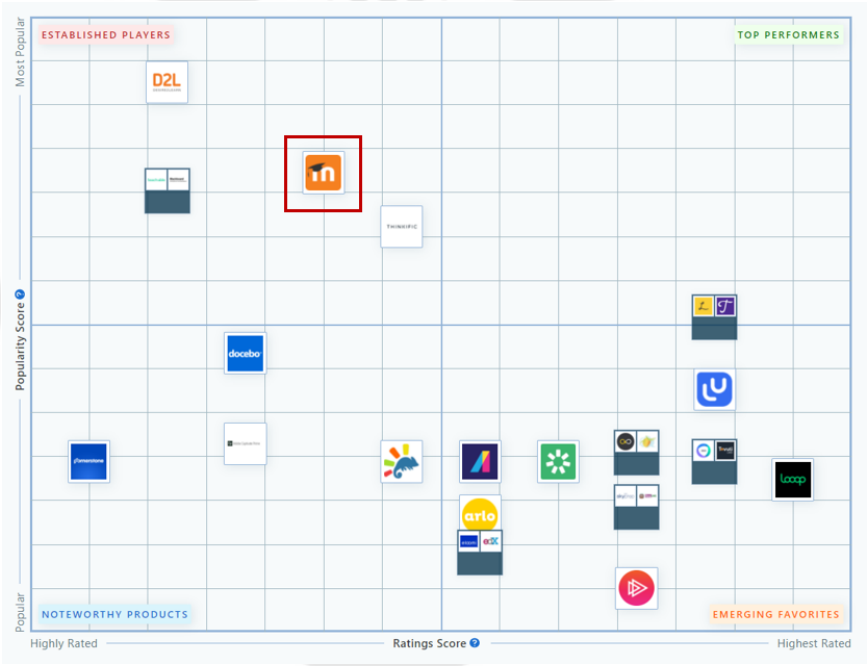
1. ระบบการจัดการเรียนรู้ (Learning Management System: LMS)
2. พฤติกรรมการเรียนรู้ออนไลน์ (Online Learning Behavior)
3. การทำเหมืองข้อมูล (Data Mining)
4. การเรียนรู้ของเครื่อง (Machine Learning)
5. การทำวิศวกรรมคุณลักษณะ (Feature Engineering)
6. การจัดการความไม่สมดุลของข้อมูล (Handling Imbalanced Data)
7. การสร้างแบบจำลองการจำแนกประเภทข้อมูล (Classification Modeling Techniques)
8. การคัดเลือกคุณลักษณะ (Feature Selection)
9. การวัดประสิทธิภาพของแบบจำลอง (Model Evaluation)
10. งานวิจัยที่เกี่ยวข้อง (Literature Review)

**ระบบการจัดการเรียนรู้ (Learning Management System: LMS)**

การเรียนการสอนแบบออนไลน์ (E-Learning) เป็นรูปแบบการเรียนผ่านทางเครือข่ายอินเทอร์เน็ต โดยใช้เนื้อหาบทเรียนร่วมกับระบบบริหารจัดการเรียนการสอนที่เรียกว่า “Learning Management System: LMS” ซึ่งมีหน้าที่จัดการข้อมูลของผู้เรียน ผู้สอน เนื้อหารายวิชา แบบทดสอบ รวมถึงการประเมินผลและติดตามความก้าวหน้าการทำกิจกรรมการเรียนการสอนของผู้เรียน และช่องทางการติดต่อสื่อสารระหว่างผู้สอนและผู้เรียน โดยมีสถานะของผู้ใช้งานตามหน้าที่การใช้งานระบบ ได้แก่ ผู้ดูแลระบบ ผู้สอน และผู้เรียน ผู้ดูแลระบบมีหน้าที่ติดตั้งระบบ ดูแลบำรุงรักษาระบบ กำหนดค่าเริ่มต้นของระบบ กำหนดสิทธิ์ของผู้ใช้งานให้แก่ผู้สอนสร้างรายวิชา โดยให้ผู้สอนสามารถสร้างและจัดการเนื้อหาในรูปแบบเอกสารและสื่อการเรียนรู้ สร้างแบบทดสอบและกำหนดระยะเวลาการทำแบบทดสอบ ดึงข้อมูลบันทึกการใช้งานระบบของผู้เรียน เพื่อให้ผู้สอนสามารถนำไปวิเคราะห์ข้อมูลไปประเมินผลการเรียนการสอนของผู้เรียนแต่ละคนได้ และผู้เรียนสามารถศึกษาเนื้อหารายวิชาและทำกิจกรรมต่าง ๆ ที่ผู้สอนได้จัดเตรียมให้ นอกจากนี้

ระบบมีการรองรับการใช้งานได้หลากหลายอุปกรณ์ เช่น คอมพิวเตอร์ แท็บเล็ต โทรศัพท์มือถือ และสามารถใช้งานได้ทุกที่ตลอดเวลาตามความต้องการของผู้ใช้งาน (อาณัติ, 2553)

มีหนึ่งในระบบการจัดการเรียนรู้ที่นิยมใช้งานกันอย่างกว้างขวาง เรียกว่า “Moodle (Modular Object Oriented Dynamic Learning Environment” เป็นระบบที่ได้รับการพัฒนาโดย Martin Dougiamas และเป็นแพลตฟอร์มการเรียนรู้ออนไลน์ฟรีแบบโอเพ่นซอร์ส (Open Source) ที่ผู้ใช้งานในหลายประเทศทั่วโลกสามารถเข้าถึง โดยมีการแปลภาษากว่า 100 ภาษา (Wikipedia, 2021b) ทาง Capterra ได้ศึกษาการเก็บรวบรวมข้อมูล LMS กว่า 1,000 ระบบ จากบทความวิจัย ข้อมูลสาธารณะ และข้อมูลจากแพลตฟอร์มการเรียนรู้ที่แตกต่างกัน พบว่าระบบ Moodle อยู่ใน 20 อันดับความนิยมและคะแนนโหวต (Capterra, 2021) ดังภาพประกอบ 1




ภาพประกอบ 1 แสดงการเปรียบเทียบแพลตฟอร์มการเรียนรู้ที่แตกต่างกันในระบบการจัดการเรียนรู้ออนไลน์ โดยวัดคะแนนโหวต (แกน x) และคะแนนความนิยม (แกน y)

ที่มา: (Capterra, 2021)

**พฤติกรรมการเรียนรู้ออนไลน์** (Online Learning Behavior)

พฤติกรรมการเรียนรู้ออนไลน์ (Online Learning Behavior) เป็นพฤติกรรมการเรียนรู้ประเภทหนึ่งที่ผ่านระบบเครือข่ายอินเทอร์เน็ต และเป็นคุณลักษณะของผู้เรียนที่สังเกตได้จากข้อมูลการคลิกและการดาวน์โหลด เช่น คลิกเข้าหลักสูตรรายวิชา เรียกใช้งานเว็บไซต์ ชมคลิป

วิดีโอ และดาวน์โหลดเอกสาร ข้อมูลพฤติกรรมการเรียนรู้เหล่านี้มีสามารถวิเคราะห์ทางสถิติในเรื่องประสิทธิภาพการเรียนรู้ออนไลน์ รวมทั้งระยะเวลาการเรียนรู้ กิจกรรมการเรียนรู้ แหล่งการเรียนรู้ และผลสอบ (Liang และคนอื่น ๆ, 2017) ปัจจัยสำคัญที่ส่งผลต่อการเรียนรู้ออนไลน์ของผู้เรียน โดยแบ่งออกเป็น 3 ด้าน ได้แก่

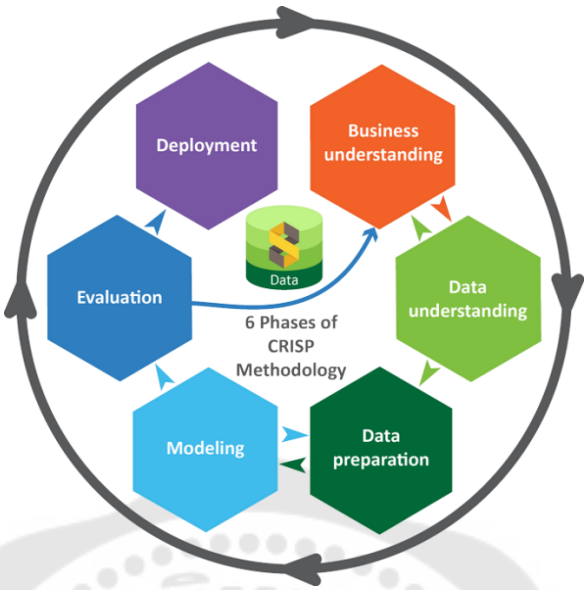
1. ด้านเนื้อหา: เนื้อหาที่ผู้สอนจัดเตรียมให้ที่มีความกระชับส่งผลต่อการเรียนรู้ของผู้เรียนเป็นอย่างยิ่ง โดยจัดส่งเนื้อหาเข้าสู่ระบบ เพื่อให้ผู้เรียนเข้าใจง่าย สามารถอ่านล่วงหน้า และทบทวนย้อนหลังผ่านระบบออนไลน์ได้

2. ด้านผู้สอน: ผู้สอนออกแบบสื่อการเรียนรู้ให้ดึงดูดความสนใจของผู้เรียน มอบหมายงานให้ผู้เรียนชมสื่อการเรียนรู้และทำแบบทดสอบ ติดตามความก้าวหน้าการเรียนของผู้เรียน และเปิดช่องทางการสื่อสารให้ผู้เรียนสามารถขอคำปรึกษา ซักสอบถามได้ เพื่อส่งเสริมให้มีปฏิสัมพันธ์ระหว่างผู้สอนและผู้เรียน และกระตุ้นให้ผู้เรียนเกิดความกระตือรือร้น

3. ด้านผู้เรียน: ผู้เรียนจะต้องศึกษาการใช้ระบบผ่านเครือข่ายอินเทอร์เน็ต เพื่อให้มีความพร้อมการใช้งานระบบก่อนเรียนจริง โดยต้องอาศัยความรู้พื้นฐานการใช้คอมพิวเตอร์ในการเข้าถึงสื่อการเรียนรู้ การทำกิจกรรมต่าง ๆ ในระบบ และการเชื่อมต่ออินเทอร์เน็ต เพื่อสามารถใช้งานระบบได้อย่างมีประสิทธิภาพ

**การทำเหมืองข้อมูล** (Data Mining)

การทำเหมืองข้อมูล เปรียบได้กับการขุดเจาะเข้าไปในข้อมูลลงสู่ระดับลึก เพื่อค้นหาสิ่งที่มีคุณค่า การทำเหมืองข้อมูลเป็นเทคนิคหนึ่งที่กล่าวถึงมากในการวิเคราะห์ข้อมูลขนาดใหญ่ (Big Data) ซึ่งเป็นสิ่งที่สำคัญที่จะทำให้องค์กรได้รับประโยชน์จากข้อมูลมหาศาล มีกรณีศึกษาเกี่ยวกับการทำเหมืองข้อมูล สามารถนำไปใช้ประโยชน์ในหลายด้าน เช่น 1) ด้านธุรกิจ: บริษัท Amazon นำข้อมูลขนาดใหญ่มาวิเคราะห์เพื่อเลือกนำเสนอขายสินค้าที่ตรงตามความต้องการของลูกค้ามากที่สุด 2) ด้านการศึกษา: ข้อมูลการเรียนผ่านระบบออนไลน์ ผลประเมินการเรียน ช่วยให้ผู้สอนรู้จักตัวเรียนมากขึ้น 3) ด้านระบบการจราจรและการขนส่ง: ข้อมูลจากอุปกรณ์เซ็นเซอร์ตรวจจับความเร็ว 4) ด้านสาธารณสุข: การหากลุ่มคนที่มีความเสี่ยงต่อการติดโรคเพื่อวางแผนการป้องกันโรคจากข้อมูลทางด้านประชากรศาสตร์ ข้อมูลสุขภาพของคนในครอบครัว (ชนัตถ์, 2562) การทำเหมืองข้อมูลมีหนึ่งในกระบวนการตอบโจทย์ธุรกิจที่สำคัญ เรียกว่า “Cross-Industry Standard Process for Data Mining (CRISP-DM)” ประกอบด้วย 6 ขั้นตอน ดังภาพประกอบ 2 (อสมา, 2561)




ภาพประกอบ 2 แสดงกระบวนการวิเคราะห์ข้อมูลด้วย CRISP-DM

ที่มา: (Sailotech, 2017)

1. **การทำความเข้าใจในธุรกิจ** (Business Understanding)

เป็นขั้นตอนที่สำคัญที่สุดในการทำความเข้าใจปัญหาและความต้องการขององค์กร เพื่อกำหนดเป้าหมายให้ชัดเจน รวมถึงแนวทางการแก้ปัญหาด้วย หากผู้วิจัยต้องการสร้างผลลัพธ์ที่ดี จะต้องอาศัยผู้ที่มีความเข้าใจธุรกิจและนักวิทยาศาสตร์ข้อมูล (Data Scientist) เพื่อให้ได้ผลลัพธ์อย่างที่ควรเป็น

2. **การทำความเข้าใจข้อมูล** (Data Understanding)

ก่อนอื่นทำการศึกษาข้อมูลว่า ต้องการข้อมูลส่วนใดบ้าง ข้อมูลควรมีคุณสมบัติอย่างไร ต้องการให้ข้อมูลมีความสัมพันธ์กันอย่างไร

3. **การเตรียมข้อมูล** (Data Preprocessing)

เป็นขั้นตอนที่ใช้เวลานานที่สุดในการทำโครงการ ยังเป็นขั้นตอนที่มีความละเอียดสูงในการจัดการข้อมูล ประกอบด้วย

3.1 การเลือกข้อมูล (Data Selection) ในที่นี้ไม่ได้หมายถึงช่วงเวลาเพียงอย่างเดียว รวมถึงตัวแปรที่เกี่ยวข้องกับข้อมูลที่ส่งผลต่อการทำนาย

3.2 การทำความสะอาดข้อมูล (Data Cleaning) ข้อมูลต่าง ๆ ที่ได้มาจากหลายแหล่งบางประเภทยังไม่ได้ถูกจัดโครงสร้าง จึงมีการทำความสะอาดข้อมูลก่อน

3.3 การรวบรวมข้อมูล (Data Collection) หากข้อมูลมาจากหลายแหล่ง ทำให้การดึงข้อมูลแต่ละครั้งใช้เวลาค่อนข้างนาน ดังนั้น ต้องนำข้อมูลหลายแหล่งมารวมไว้ที่เดียวกัน และนำข้อมูลที่ซ้ำซ้อนมาไว้ในตำแหน่งเดียวกันด้วย

3.4 การจัดรูปแบบของข้อมูลให้อยู่ในรูปแบบพร้อมสร้างแบบจำลอง (Data Transformation) ชุดข้อมูลจะต้องมีโครงสร้างที่ชัดเจน เช่น จำนวนรายการเท่าไร และตัวแปรมีความพร้อมหรือไม่ มีข้อมูลสูญหาย (Missing data) หรือข้อมูลผิดปกติ (Outlier data) หรือไม่

4. **การสร้างแบบจำลอง** (Data Modeling)

ในการหาความสัมพันธ์เชิงลึกจากข้อมูลชุดใดชุดหนึ่ง จึงมีความจำเป็นที่ต้องสร้างแบบจำลอง ซึ่งแตกต่างจากการวิเคราะห์ข้อมูลแบบทั่วไปที่อาจไม่สามารถมองหาความแม่นยำ หรือผลลัพธ์เชิงลึกได้

5. **การวัดประสิทธิภาพของแบบจำลอง** (Model Evaluation)

ต้องการทราบว่าแบบจำลองมีผลลัพธ์ที่ดีหรือไม่ สามารถพิจารณาความแม่นยำของผลลัพธ์ที่ได้จากการทำนาย หากแบบจำลองนี้ผ่านการประเมินผลลัพธ์ จะเข้าสู่ขั้นตอนการนำแบบจำลองไปใช้งาน แต่หากแบบจำลองนี้ไม่ผ่านการประเมินผลลัพธ์ อาจจะต้องย้อนกลับไปที่ขั้นตอนการทำความเข้าใจข้อมูลใหม่ หรือการสร้างแบบจำลองใหม่

6. **การนำไปใช้จริง** (Implementation)

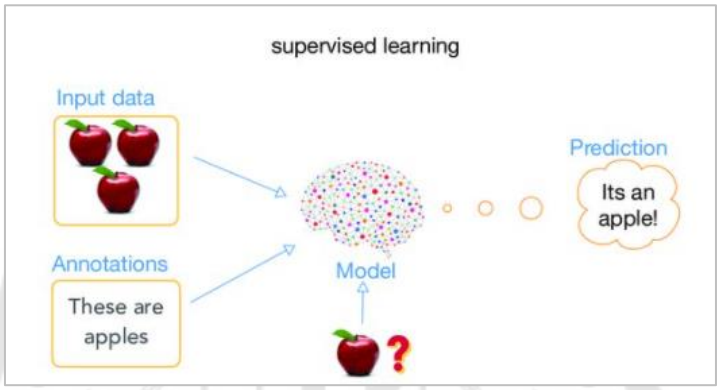
ก่อนจะใช้งานจริงทั้งระบบ จะต้องมีการทดสอบระบบก่อน โดยแบ่งออกเป็นการทดสอบระบบเชื่อมโยงข้อมูล การทดสอบสคริปต์ การทดสอบการประมวลผล รวมถึงการทดสอบแบบจำลอง

**การเรียนรู้ของเครื่อง** (Machine Learning)

การที่จะพัฒนาให้เครื่องคอมพิวเตอร์มีความฉลาดและความสามารถเหมือนมนุษย์ เครื่องจำเป็นต้องมีการเรียนรู้สถานการณ์ เพื่อวางแผนและตัดสินใจแก้ปัญหาได้ ซึ่งเรียกกระบวนการนี้ว่า “การเรียนรู้ของเครื่อง (Machine Learning)” การเรียนรู้ของเครื่องเป็นสาขาหนึ่งของปัญญาประดิษฐ์ (Artificial Intelligence) โดยมุ่งเน้นศึกษาการทำให้คอมพิวเตอร์มีความสามารถในการเรียนรู้ได้ โดยไม่ต้องเขียนโปรแกรมกำกับไว้อย่างชัดเจน กระบวนการเรียนรู้ของเครื่องมีการแบ่งชุดข้อมูลออกเป็น 2 ส่วน คือ ชุดข้อมูลการเรียนรู้ (Training Set) เป็นข้อมูลที่ถูกนำไปเรียนรู้ให้กับเครื่อง และชุดข้อมูลทดสอบ (Test Set) เป็นข้อมูลที่ใช้สำหรับการประเมินประสิทธิภาพการทำนาย การเรียนรู้ของเครื่องนั้นมีการเรียนรู้หลายแบบที่ขึ้นอยู่กับลักษณะของข้อมูล สามารถแบ่งออกเป็น 4 ประเภท (Awad และ Khanna, 2015) ได้แก่

1. **การเรียนรู้แบบมีผู้สอน** (Supervised Learning)

การที่เรียนรู้ของเครื่องนั้น จำเป็นต้องให้ผู้ใช้สอนการเรียนรู้ก่อน โดยนำข้อมูลขาเข้า (Input) หรืออีกชื่อหนึ่งเรียกว่า ข้อมูลการเรียนรู้ (Training Set) ให้เครื่องเรียนรู้จนได้ข้อมูลขาออก (Output) หรือป้ายกำกับ (Label หรือ Class) โดยแบ่งออกเป็น 2 ประเภทตามลักษณะของ ป้ายกำกับ ได้แก่ การจำแนกประเภทของข้อมูล (Classification) และการวิเคราะห์การถดถอย (Regression) ดังภาพประกอบ 3

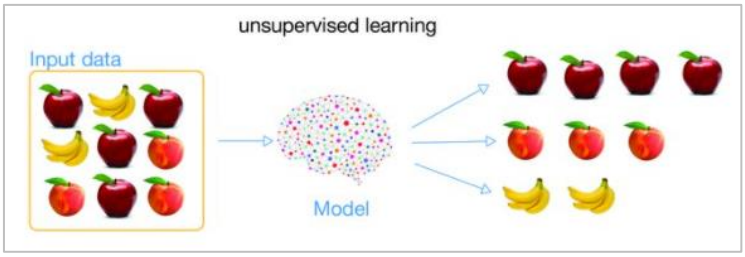



ภาพประกอบ 3 แสดงตัวอย่างการทำงานของการเรียนรู้ของเครื่องแบบมีผู้สอน

ที่มา: (M. Yan และคนอื่น ๆ, 2018)

2. **การเรียนรู้แบบไม่มีผู้สอน** (Unsupervised Learning)

การเรียนรู้แบบนี้เป็นสิ่งที่อยู่ตรงกันข้ามกับการเรียนรู้แบบมีผู้สอน ผู้ใช้ไม่จำเป็นต้องสอนเครื่องจากการป้อนข้อมูลเข้าสู่เครื่อง เครื่องสามารถเรียนรู้ข้อมูลได้เอง มีแบบจำลองสำหรับการเรียนรู้แบบไม่มีผู้สอน ได้แก่ การจัดกลุ่ม (Clustering), การลดมิติ (Dimensionality Reduction) ดังภาพประกอบ 4

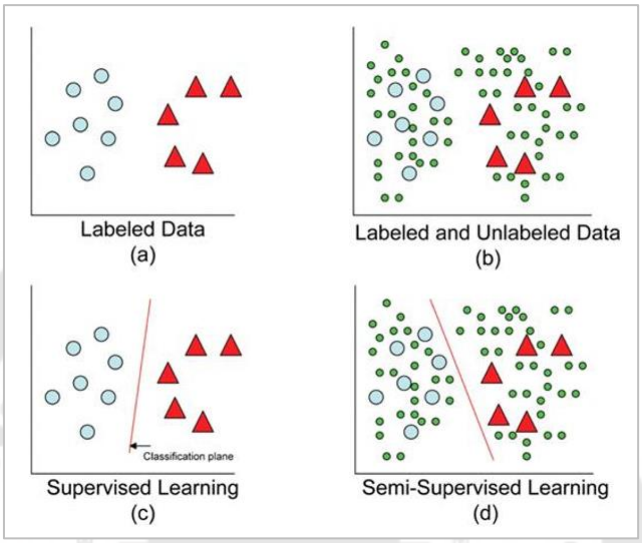



ภาพประกอบ 4 แสดงตัวอย่างการทำงานของการเรียนรู้ของเครื่องแบบไม่มีผู้สอน

ที่มา: (M. Yan และคนอื่น ๆ, 2018)

3. **การเรียนรู้ของเครื่องแบบกึ่งมีผู้สอน** (Semi-Supervised Machine Learning)

เป็นการผสมระหว่างการเรียนรู้แบบมีผู้สอนและไม่มีผู้สอน วิธีการเรียนรู้ของเครื่องในลักษณะนี้จะใช้สำหรับข้อมูลการเรียนรู้ที่มีป้ายกำกับบางส่วนและไม่มีป้ายกำกับบางส่วนของข้อมูลทั้งหมด ดังนั้น การที่จะติดป้ายกำกับให้กับข้อมูลที่ไม่มีป้ายกำกับจะใช้เวลานาน เนื่องจากต้องอาศัยความสามารถการเรียนรู้โดยผู้เชี่ยวชาญ ดังภาพประกอบ 5

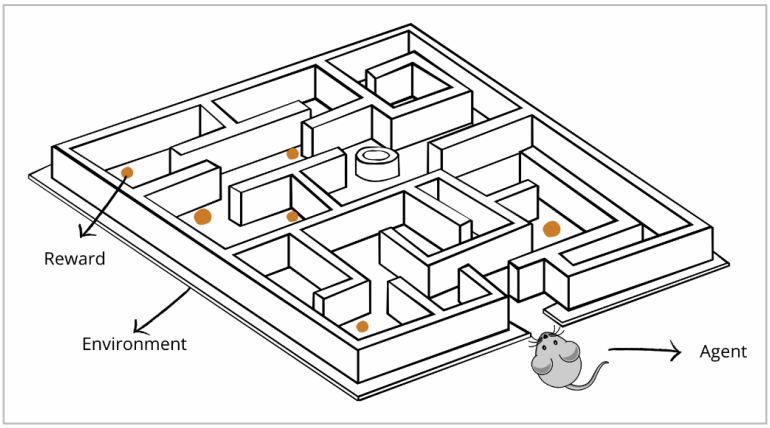



ภาพประกอบ 5 แสดงตัวอย่างการทำงานของการเรียนรู้ของเครื่องแบบกึ่งมีผู้สอน โดยนำรูปวงกลมสีฟ้าและสามเหลี่ยมสีแดงเป็นข้อมูลที่มีป้ายกำกับ 2 คลาส และวงกลมสีเขียวเป็นข้อมูลที่ไม่มีป้ายกำกับ ให้กับเครื่องเรียนรู้ว่า วงกลมสีเขียวอยู่คลาสใด

ที่มา: (Rodriguez, 2021)

4. **การเรียนรู้ของเครื่องแบบเสริมกำลัง** (Reinforcement Learning)

การที่เครื่องเรียนรู้ที่เกิดจากปฏิสัมพันธ์การลองผิดลองถูกระหว่างผู้เรียนรู้ (Agent) และสิ่งแวดล้อม (Environment) โดยผู้เรียนรู้สามารถเรียนรู้เองได้จากการเลือกการกระทำใด (Action) ในสิ่งแวดล้อม เพื่อให้ได้รางวัลสูงที่สุด เปรียบเสมือนกับการเล่นเกมส์หาทางออกจากเขาวงกตโดยเร็วที่สุด ดังภาพประกอบ 6

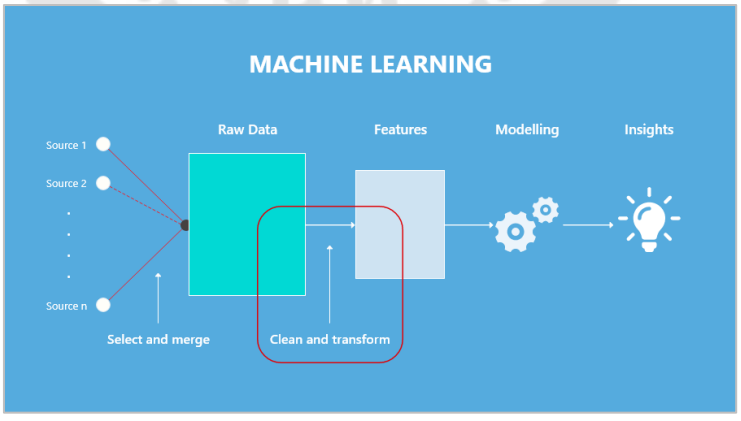



ภาพประกอบ 6 แสดงตัวอย่างการทำงานของการเรียนรู้ของเครื่องแบบเสริมกำลัง โดยสร้างเขาวงกตให้หนูวิ่งเข้าออกเขาวงกตจากการลองผิดลองถูก

ที่มา: (Tas, 2021)

**การทำวิศวกรรมคุณลักษณะ (Feature Engineering)**

วิศวกรรมคุณลักษณะ (Feature Engineering) เป็นกระบวนที่ใช้ความรู้เฉพาะทาง (Domain Knowledge) ในการสร้างคุณลักษณะที่จะทำความสะอาด และแปลงข้อมูลดิบในรูปแบบที่สามารถเรียนรู้ด้วยแบบจำลอง ทั้งนี้ช่วยลดระยะเวลาการเรียนรู้ข้อมูลจากการลดจำนวนคุณลักษณะ และส่งผลต่อประสิทธิภาพของแบบจำลองดีขึ้น (Mohamad, Ahmad, Jawawi, และ Mohd Hashim, 2020) ดังภาพประกอบ 7




ภาพประกอบ 7 แสดงกระบวนการทำวิศวกรรมคุณลักษณะ โดยนำข้อมูลหลายแหล่งเข้าด้วยกันแล้วทำความสะอาดและเปลี่ยนแปลงโครงสร้างให้เป็นคุณลักษณะที่พร้อมเข้าสู่แบบจำลองการทำนาย

ที่มา: (MasterSoftwareSolutions, 2021)

เทคนิคการทำวิศวกรรมคุณลักษณะในการสร้างคุณลักษณะใหม่ หรือปรับรูปแบบข้อมูลของคุณลักษณะ ซึ่งมีหลายวิธี (Brownlee, 2020b; Datacamp, 2021) ดังนี้

1. **การแทนค่าที่ขาดหาย** (Imputation)

ส่วนใหญ่พบบ่อยในแบบสอบถามที่ระบุข้อมูลไม่ครบ สามารถเติมค่าที่ขาดหายด้วยค่าหนึ่ง เช่น ค่าเฉลี่ย ฐานนิยม มัธยฐาน

2. **การจัดการข้อมูลผิดปกติ** (Handling Outliers)

อาจจะเกิดจากข้อผิดพลาดหลายสาเหตุ เช่น การป้อนข้อมูล เครื่องการวัด การประมวลผลข้อมูล การทดลอง

3. **การคำนวณข้อมูลทางสถิติต่าง ๆ** (Data Aggregation)

นำข้อมูลไปวิเคราะห์และสรุปข้อมูลทางสถิติ เช่น ผลรวม ค่าเฉลี่ย ระยะเวลา

4. **การจัดการวันที่และเวลา** (Dealing with DateTime Features)

ทำการแยกวันที่และเวลาออกจากกัน โดยมาจากข้อมูลชนิด DateTime

5. **การลดขนาดของข้อมูล** (Scaling Data)

ปกติชุดข้อมูลมีขนาดข้อมูลที่แตกต่างกันแต่ละคุณลักษณะ จึงปรับช่วงของข้อมูลให้เหมาะสมก่อนนำข้อมูลไปเรียนรู้ต่อไป การปรับช่วงของข้อมูลสามารถทำได้ด้วยหลายวิธี เช่น Standardization, Mean Normalization, Min-Max Normalization

6. **การแปลงตัวอักษรเป็นข้อมูลตัวเลข** (Categorical Data Encoding)

เช่น การแบ่งกลุ่มผู้เรียนตามเพศเป็นชายและหญิง เปลี่ยนเป็นค่า 1 และ 2 ตามลำดับ สามารถแปลงได้โดยใช้ 2 วิธี ดังนี้

6.1 Label Encoding

มีหน้าที่แปลงข้อมูลตัวอักษรเป็นตัวเลขแต่ละค่าต่าง ๆ โดยใช้คอลัมน์เดิม

6.2 One-hot Encoding

มีหน้าที่แยกคอลัมน์ใหม่จากคอลัมน์เดิม โดยมีจำนวนคอลัมน์เท่ากับจำนวนประเภทของข้อมูลตัวอักษร หากรายการใดอยู่ประเภทใด จะกำหนดค่าเท่ากับ 1 ในคอลัมน์สำหรับประเภทนี้ ส่วนรายการใดไม่ได้อยู่ประเภท จะแสดงค่าศูนย์ในคอลัมน์ประเภทนี้

กระบวนการวิศวกรรมคุณลักษณะส่วนใหญ่ถูกนำไปใช้กับข้อมูลในรูปแบบการบันทึกการใช้งาน (Log File) เนื่องจากข้อมูลบันทึกการใช้งานเป็นข้อมูลดิบที่มาจากแพลตฟอร์มออนไลน์ (Mohamad และคนอื่น ๆ, 2020) เช่น การเข้าถึงสื่อการเรียนการสอนใน Massive Open Online Course (MOOC), การเล่นผ่านแอพ Facebook, การสั่งซื้อสินค้าบนแอพ Shopee

**การจัดการความไม่สมดุลของข้อมูล** (Handling Imbalanced Data)

ในการสร้างแบบจำลองสำหรับการจำแนกประเภท มีสิ่งหนึ่งที่พบปัญหาบ่อยมากคือ ข้อมูลแต่ละคลาสมีจำนวนไม่เท่ากัน เรียกว่า “ข้อมูลที่ไม่สมดุล (Imbalanced Data)” ซึ่งเป็นปัญหาของการจำแนกประเภทเท่านั้น หมายความว่าเมื่อทำนายข้อมูลที่มีจำนวนแตกต่างกันมากแต่ละคลาส โดยที่ไม่ได้จัดการข้อมูลที่ไม่สมดุลก่อน ผลลัพธ์การทำนายที่ได้เป็นคลาสที่มีข้อมูลจำนวนมาก (Major class) ข้อมูลในลักษณะนี้จะต้องถูกปรับความสมดุลก่อนเข้าสู่การทำนายข้อมูลด้วยแบบจำลองการจำแนกประเภท โดยสามารถใช้เทคนิคหนึ่งในวิธีการจัดการความไม่สมดุลของข้อมูลสำหรับงานวิจัยนี้ เทคนิคนี้เรียกว่า “การสุ่มตัวอย่างเพิ่มกลุ่มน้อยด้วยการสร้างตัวอย่างสังเคราะห์ (Synthetic Minority Oversampling Technique: SMOTE)” เป็นการสุ่มเพิ่มจำนวนข้อมูลในคลาสที่มีข้อมูลอยู่น้อยให้มากขึ้น โดยให้แต่ละคลาสมีจำนวนเท่า ๆ กัน เพื่อให้ข้อมูลมีความสมดุลมากขึ้น โดยมีหลักการทำงานดังภาพประกอบ 8 และมีขั้นตอนการทำงานดังนี้ (Sun, Lang, Fujita, และ Li, 2018)

1. สุ่มค่าข้อมูลที่อยู่ในคลาสบวกที่มีอยู่น้อยขึ้นมาจำนวน 1 ค่า จากจำนวนตัวอย่างในคลาสบวกทั้งหมดเท่ากับ $S_{Pos}$ โดยกำหนดตัวแปรของจุดที่สุ่มในคลาสบวกเป็น $x^{Pos}$

2. กำหนดจำนวนเพื่อนบ้านในคลาสบวกที่ใกล้เคียงค่าที่สุ่มที่สุดจำนวน $K$ โดยกำหนดตัวแปรของเพื่อนบ้านเป็น $x^{Pos-Near}$

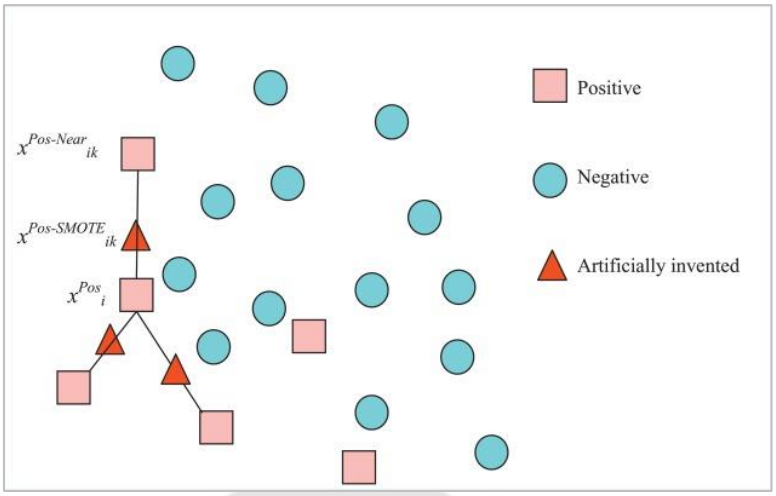
3. แล้วลากเส้นจากจุดที่สุ่มไปยังเพื่อนบ้านที่อยู่ใกล้เคียงที่มีจำนวน $K$

4. สุ่มเลือกจุดที่อยู่บนเส้นนี้ในการสร้างข้อมูลเทียมที่อยู่ระหว่างจุดที่สุ่มกับเพื่อนบ้านที่อยู่ใกล้เคียงจุดที่สุ่ม โดยกำหนดตัวแปรของจุดเทียมเป็น $x^{Pos-SMOTE}$

การคำนวณการสร้างจุดเทียมขึ้นระหว่างจุดที่อยู่คลาสบวกที่มีจำนวนน้อยกว่าคลาสลบ 2 จุด โดยใช้สมการ (1) ดังนี้

$$x^{Pos-SMOTE}{}_{ij} = x^{Pos}{}_{i} + rand(0,1) \times \left(x^{Pos-Near}{}_{ij} - x^{Pos}{}_{i}\right) \quad (1)$$

เมื่อ $i \in [1, 2, \dots, S_{Pos}]$, $j = 1, 2, \dots, K$ และ $rand(0,1)$ เป็นค่าสุ่มแบบต่อเนื่องระหว่าง 0 ถึง 1




ภาพประกอบ 8 แสดงตัวอย่างการทำงานของการจัดการข้อมูลที่ไม่สมดุลด้วยวิธี SMOTE โดยนำข้อมูลคลาสบวกที่เป็นสี่เหลี่ยมและคลาสลบที่เป็นวงกลมไปจัดการความสมดุล แล้วจะได้ข้อมูลเทียมที่เป็นรูปสามเหลี่ยมเพิ่มขึ้นมา

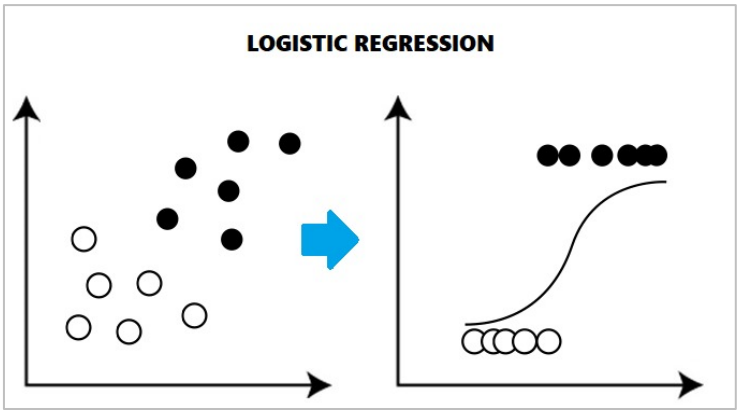
ที่มา: (Sun และคนอื่น ๆ, 2018)

**การสร้างแบบจำลองการจำแนกประเภท (Classification Modeling Techniques)**

แบบจำลองการจำแนกประเภทที่ใช้ในการวิจัยครั้งนี้ จำนวน 7 แบบ ได้แก่ Logistic Regression, Naïve Bayes, Support Vector Machine, K-Nearest Neighbor, Decision Tree, Random Forest และ Extreme Gradient Boosting (XGBoost) แบบจำลองเหล่านี้มีการใช้งานอย่างแพร่หลายในการนำมาทำนายผลการเรียนของนิสิต ซึ่งมีความหมายของแบบจำลอง ดังนี้

**1. แบบจำลอง Logistic Regression**

Logistic Regression เป็นแบบจำลองการจำแนกประเภทในการวิเคราะห์ข้อมูลที่อาศัยความน่าจะเป็นของผลลัพธ์การทำนายที่จะเกิดขึ้นหรือไม่ การทำนายจะขึ้นอยู่กับปัจจัยที่ส่งผลต่อการทำนาย ซึ่งเป็นตัวแปรอิสระที่สามารถเป็นทั้งแบบไม่ต่อเนื่องและแบบต่อเนื่อง (Wikipedia, 2021a) สมการของ Logistic Regression เป็นฟังก์ชัน Sigmoid ที่มีลักษณะเป็นเส้นโค้งรูป S โดยมีค่าระหว่าง 0 ถึง 1 ฟังก์ชันนี้ถูกแปลงมาจากฟังก์ชัน Logit ซึ่งมีค่าเป็น Infinity เนื่องจากค่าความน่าจะเป็นอยู่ระหว่าง 0 ถึง 1 (Brixius, 2016) แสดงภาพประกอบ 9




ภาพประกอบ 9 แสดงการเปรียบเทียบก่อนและหลังการทำงานของแบบจำลอง Logistic Regression

ที่มา: (Equiskill, 2018)

ดังนั้น สมการที่ได้อยู่ในรูปสูตร ดังสมการ (2)

$$P(class) = \frac{e^{(\beta_0+\beta_1 X_1+\beta_2 X_2+\cdots+\beta_n X_n)}}{1+e^{(\beta_0+\beta_1 X_1+\beta_2 X_2+\cdots+\beta_n X_n)}} \quad (2)$$

โดยที่ $X_i$ คือ คุณลักษณะตัวที่ $i$ เมื่อ $i = 1, 2, 3, \ldots, N$

$\beta_0$ คือ ค่าไบแอส (Bias)

$\beta_i$ คือ ค่าสัมประสิทธิ์ของ $X_i$ เมื่อ $i = 1, 2, 3, \ldots, N$

$e$ คือ ค่าคงที่ทางคณิตศาสตร์ที่เป็นฐานของลอการิทึมธรรมชาติ มีค่าโดยประมาณ 2.71828

แบบจำลอง Logistic Regression ไม่ได้ใช้พลังการประมวลผลในการคำนวณข้อมูลสูง ตีความได้ง่าย แต่อาจจะเกิด Overfitting และแบ่งออกเป็น 3 ประเภท (Wikipedia, 2021a) ดังนี้

1. **Binary Logistic Regression** เป็น Logistic Regression ที่วิเคราะห์ข้อมูลที่มี 2 คลาส เช่น ผู้ใช้งานจะคลิกลิงก์ชมสื่อการเรียนรู้หรือไม่ (คลิก, ไม่ได้คลิก)

2. **Multinomial Logistic Regression** เป็น Logistic Regression ที่วิเคราะห์ข้อมูลที่มี 3 คลาสขึ้นไป และไม่สามารถนำข้อมูลมาจัดลำดับ เช่น ผู้เรียนจะอ่านหนังสือก่อนสอบหรือไม่ (อ่านเป็นประจำ, อ่านนานๆ ครั้ง, ไม่อ่าน)

**3. Ordinal Logistic Regression** เป็น Logistic Regression ที่วิเคราะห์ข้อมูลที่สามารถนำมาจัดลำดับได้ เช่น ผู้เรียนมีโอกาสได้ระดับผลการเรียนในระดับใด (ดีเยี่ยม, ปานกลาง, เป็นที่น่าพอใจ)

**2. แบบจำลอง Naïve Bayes**

Naïve Bayes เป็นแบบจำลองที่อาศัยหลักการความน่าจะเป็นตามทฤษฎีของเบย์ (Bayes Theorem) ที่บ่งบอกถึงความสัมพันธ์ระหว่างสิ่งที่กำลังเกิดขึ้นและสิ่งที่เกิดขึ้นแล้ว เพื่อทำนายคลาสจากคุณลักษณะที่เป็นอิสระต่อกัน (Amra และ Maghari, 2017) โดยกำหนดตัวแปรต่าง ๆ และสามารถเขียนดังสมการ (3)

$$P(c|x) = \frac{P(x|C)P(c)}{P(x)} \quad (3)$$

โดยกำหนดตัวแปร ดังนี้

$P(c|x)$ เป็นความน่าจะเป็นที่ได้คลาสใดคลาสหนึ่ง เมื่อใช้คุณลักษณะแล้ว เรียกว่า “Posterior Probability”

$P(c)$ เป็นความน่าจะเป็นที่ได้คลาสใดคลาสหนึ่ง “Prior Probability”

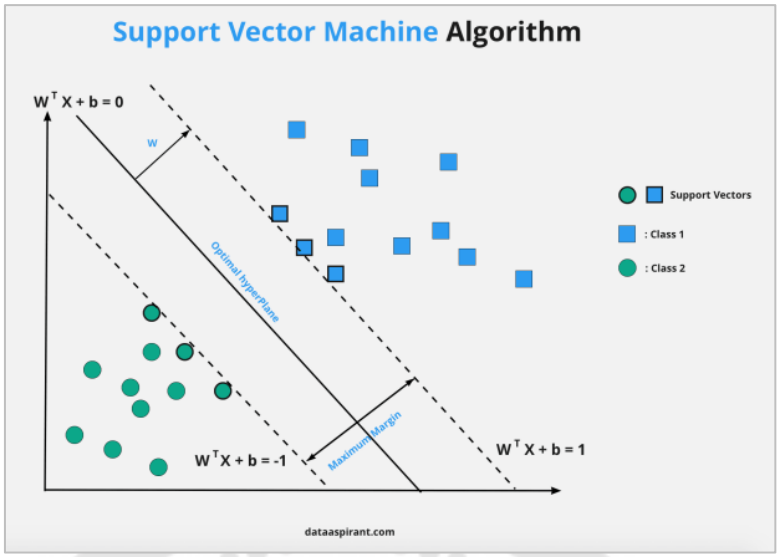
$P(x|c)$ เป็นความน่าจะเป็นที่ได้คุณลักษณะ เมื่อเลือกคลาสใดคลาสหนึ่งแล้ว เรียกว่า “Likelihood”

$P(x)$ เป็นความน่าจะเป็นที่ได้คุณลักษณะ เรียกว่า “Prior Probability”

Naïve Bayes ทำนายข้อมูลได้ง่ายและเร็ว สามารถใช้ในการจำแนกคลาสจำนวนมาก (Multiple Classification) แต่หากไม่พบข้อมูลทดสอบในชุดข้อมูลการเรียนรู้ จะได้จำนวนข้อมูลที่เกิดขึ้นเป็นศูนย์ ทำให้ไม่มีค่าความน่าจะเป็น จึงใช้วิธีการประมาณค่า Laplace (Laplace estimation) เข้ามาช่วยในการแก้ปัญหานี้ (Shamsi และ Lakshmi, 2016)

**3. แบบจำลอง Support Vector Machine**

Support Vector Machine เป็นแบบจำลองที่ใช้ในการจำแนกประเภทข้อมูลท โดยสร้างเส้นที่แบ่งคลาสออกจากกันที่เรียกว่า “Hyperplane” ที่ทำให้มีระยะขอบของ 2 คลาสให้มากที่สุด (Maximum Margin) โดยให้เส้นขอบออกไปสัมผัสจุดที่เป็นสมาชิกของข้อมูลการเรียนรู้ให้ใกล้ที่สุดที่เรียกว่า “Support Vector” ซึ่งมีอย่างน้อย 1 จุด แบบจำลองนี้มีความทนทานต่อข้อมูลรบกวน (Noisy data) ได้ดี เนื่องจากมี Maximum Margin สามารถจำแนกข้อมูลที่แบ่งแยกออกจากกันได้ทั้งแบบเชิงเส้น (Linear SVM) และแบบไม่เชิงเส้น (Non-Linear SVM) (Jakkula, 2011) และมีหลักการทำงานของ SVM ดังภาพประกอบ 10




ภาพประกอบ 10 แสดงการทำงานของแบบจำลอง Support Vector Machine สำหรับการแบ่งข้อมูล 2 คลาส ได้แก่ รูปสี่เหลี่ยม (คลาสที่ 1) และรูปวงกลม (คลาสที่ 2)

ที่มา: (Awasthi, 2020)

การคำนวณค่าของเส้นของทั้ง 2 คลาส ในการจำแนกประเภทข้อมูล โดยใช้สมการที่กำหนดตัวแปรต่าง ๆ ได้แก่ $w^T$ คือ ค่าน้ำหนัก, $x_i$ คือ คุณลักษณะของรายการที่ $i$, $b$ คือ ค่าไบแอส (Bias) และ $i = 1, 2, \dots, N$ เป็นของรายการที่ $i$ ของชุดข้อมูล (Wikipedia, 2021d) ดังสมการ (4) – (6)

1. สมการ Hyperplane คือ $w^T x_i + b = 0$ (4)
2. หาก $y_i = \ class\ 1$ จะได้ว่า $w^T x_i + b \geq 1$ (5)
3. หาก $y_i = \ class\ 2$ จะได้ว่า $w^T x_i + b \leq -1$ (6)

หากไม่สามารถแบ่งข้อมูลออกเป็น 2 คลาสด้วยเส้นตรง Hyperplane จึงแก้ปัญหาด้วยการใช้สมการ Kernel Function แทน เพื่อให้ข้อมูลสามารถแบ่งออกจากกันได้ ดังสมการ Kernel Function 4 สมการที่นิยม (Abuzalata, 2019) ดังสมการ (7) – (10)

1. Linear:

$$K(x_i, x_j) = {x_i}^T x_j \quad (7)$$

2. Polynomial:

$$K(x_i, x_j) = (\gamma {x_i}^T x_j + r)^j, \gamma > 0 \quad (8)$$

3. Radial Basis Function (RBF):

$$K(x_i, x_j) = \exp\left(-\gamma \left\| x_i - x_j \right\|^2\right), \gamma > 0 \quad (9)$$

4. Sigmoid

$$K(x_i, x_j) = \tanh\left(\gamma {x_i}^T x_j + r\right) \quad (10)$$

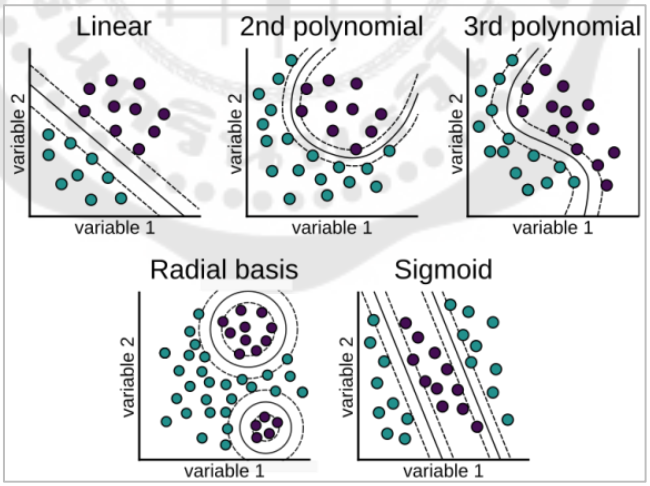
เมื่อ $x_i$ หมายถึง คุณลักษณะที่ 1 ของรายการที่ $i$ เมื่อ $i = 1, 2, \ldots, N$

$x_j$ หมายถึง คุณลักษณะที่ 2 ของรายการที่ $j$ เมื่อ $j = 1, 2, \ldots, N$

$\gamma$ หมายถึง ค่าความแคบของส่วนโค้ง

$r$ หมายถึง ค่าพารามิเตอร์ของสมการที่เป็นสัมประสิทธิ์ของค่าไบแอส

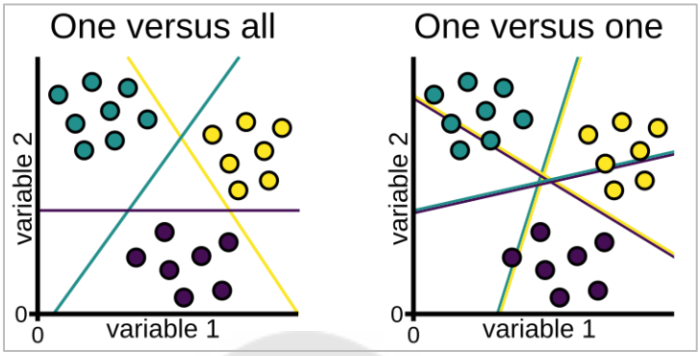
และแสดงกราฟของสมการ Kernel Function ดังภาพประกอบ 11




ภาพประกอบ 11 แสดงการเปรียบเทียบการทำงานของแบบจำลอง Support Vector Machine โดยใช้ข้อมูลที่สามารถแบ่ง 2 คลาสด้วยเส้นตรง และข้อมูลที่ไม่สามารถแบ่ง 2 คลาสด้วยเส้นตรง จากการกำหนดคลาสที่เป็นรูปวงกลมสีม่วง (คลาสที่ 1) และรูปวงกลมสีฟ้า (คลาสที่ 2)

ที่มา: (hefinioanrhys, 2019)

ต่อไปกล่าวถึงการจำแนกประเภทข้อมูลอย่างน้อย 3 คลาสขึ้นไป สามารถใช้เทคนิค On- versus-All และ One-versus-One ดังภาพประกอบ 12




ภาพประกอบ 12 แสดงการเปรียบเทียบการทำงานของแบบจำลอง Support Vector Machine โดยใช้ข้อมูลที่ถูกแบ่งคลาสแบบ One versus all และข้อมูลที่ถูกแบ่งคลาสแบบ One versus one จากการกำหนดคลาสที่เป็นรูปวงกลมสีม่วง (คลาสที่ 1) รูปวงกลมสีฟ้า (คลาสที่ 2) และรูปวงกลมสีเหลือง (คลาสที่ 3)

ที่มา: (hefinioanrhys, 2019)

โดยสมมติจำนวนคลาสเท่ากับ 4 จะได้ว่า

1. One-versus-All ทำหน้าที่แบ่งชุดข้อมูลการจำแนกประเภทจำนวน 3 ชุด ดังนี้ (หากกำหนดจำนวน $k$ คลาส จะได้ชุดข้อมูลการจำแนกประเภททั้งหมด $k$ ชุด)

1.1 ชุดที่ 1: คลาสที่ 1 กับ [คลาสที่ 2, คลาสที่ 3, คลาสที่ 4]

1.2 ชุดที่ 2: คลาสที่ 2 กับ [คลาสที่ 1, คลาสที่ 3, คลาสที่ 1]

1.3 ชุดที่ 3: คลาสที่ 3 กับ [คลาสที่ 1, คลาสที่ 2, คลาสที่ 4]

2. One-versus-One ทำหน้าที่แบ่งชุดข้อมูลการจำแนกประเภทจำนวน 6 ชุด ดังนี้ (หากกำหนดจำนวน $k$ คลาส จะได้ชุดข้อมูลการจำแนกประเภททั้งหมด $\frac{k \times (k-1)}{2}$ ชุด)

2.1 ชุดที่ 1: คลาสที่ 1 กับ คลาสที่ 2

2.2 ชุดที่ 2: คลาสที่ 1 กับ คลาสที่ 3

2.3 ชุดที่ 3: คลาสที่ 1 กับ คลาสที่ 4

2.4 ชุดที่ 4: คลาสที่ 2 กับ คลาสที่ 3

2.5 ชุดที่ 5: คลาสที่ 2 กับ คลาสที่ 4

2.6 ชุดที่ 6: คลาสที่ 3 กับ คลาสที่ 4

**4. แบบจำลอง K-Nearest Neighbor**

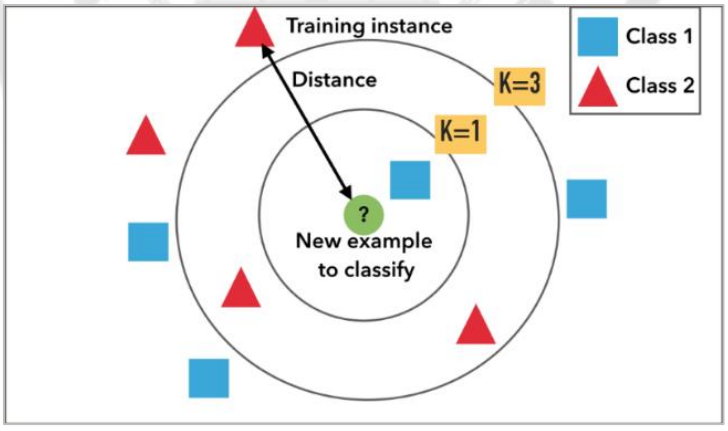
K-Nearest Neighbor เป็นแบบจำลองการจำแนกประเภทข้อมูล โดยพิจารณาเพื่อนบ้านใกล้เคียงที่สุดจำนวน K ตำแหน่ง จากระยะห่างระหว่างแต่ละคุณลักษณะในข้อมูลตัวอย่างและข้อมูลที่สนใจว่าอยู่คลาสใด โดยเลือกระยะทางน้อยที่สุด เป็นวิธีการที่เข้าใจง่าย เรียนรู้ข้อมูลได้เร็ว มีความทนทานต่อข้อมูลรบกวน (Noisy data) ได้ดี แต่ค่อนข้างประมวลผลช้า มีความไวต่อโครงสร้างข้อมูล (Data Structure) แบบจำลองนี้มีการทำงานดังภาพประกอบ 13 (Amra และ Maghari, 2017) โดยมีขั้นตอนการทำงาน ดังนี้

1. ต้องการทราบว่าจุดสีเขียวอยู่คลาสใด จากชุดข้อมูลการเรียนรู้ (Training Set) ที่มี 2 คลาส ได้แก่ คลาสที่ 1 (รูปสี่เหลี่ยม) และคลาสที่ 2 (รูปสามเหลี่ยม)

2. กำหนดจำนวนจุดที่อยู่ใกล้เคียงกับจุดสีเขียวที่สุดเป็น K จุด (จำนวนคี่)

3. สมมติกำหนด K เท่ากับ 3 แล้วคำนวณระยะทาง (Distance) ระหว่างจุดที่ใกล้เคียงจุดสีเขียวที่สุดจำนวน 3 จุด และจุดสีเขียว พบว่ามีจุดที่ใกล้เคียงจุดสีเขียวที่สุดจำนวน 3 จุด ได้แก่ รูปสี่เหลี่ยม 1 รูป และรูปสามเหลี่ยม 2 รูป

4. ต่อมาใช้หลักการ Majority Vote ในการเลือกคลาสส่วนใหญ่ที่อยู่ใกล้เคียงจุดสีเขียวที่สุดจำนวน 3 จุด แสดงว่ามีรูปสามเหลี่ยม 2 รูป มากกว่ารูปสี่เหลี่ยม 1 รูป นั่นคือจุดสีเขียวอยู่คลาสที่ 1 ซึ่งเป็นรูปสามเหลี่ยม




ภาพประกอบ 13 แสดงตัวอย่างการทำงานของแบบจำลอง K-Nearest Neighbor จากการกำหนดคลาสที่เป็นรูปสี่เหลี่ยม (คลาสที่ 1) และคลาสที่เป็นรูปสามเหลี่ยม (คลาสที่ 2)

ที่มา: (A. Ali, 2018)

**5. แบบจำลอง Decision Tree**

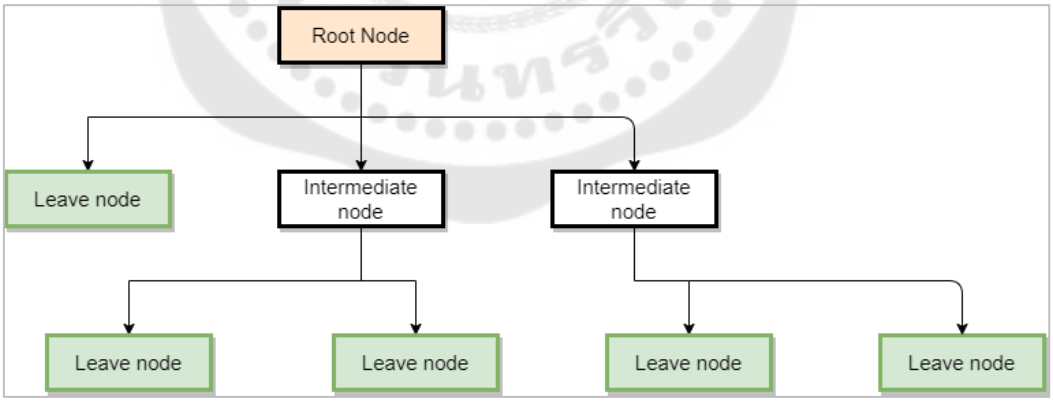
Decision Tree เป็นหนึ่งในแบบจำลองทั้งการจำแนกประเภทและการวิเคราะห์การถดถอย แบบจำลองนี้ตีความได้ง่าย เนื่องจากต้นไม้ตัดสินใจเป็นการเลียนแบบการตัดสินใจของมนุษย์ ทำงานกับคุณลักษณะได้ทุกชนิด สามารถจัดการข้อมูลสูญหาย แต่ค่อนข้างไม่ยืดหยุ่น หากเปลี่ยนแปลงข้อมูล ทำให้แผนภาพ Decision Tree เปลี่ยนได้ อาจทำให้เกิด Overfitting กับข้อมูลได้ง่าย ทำให้ส่งผลต่อการทำนายข้อมูลได้ไม่ค่อยดีนัก Decision Tree มีลักษณะคล้ายต้นไม้กลับหัว (Patel และ Prajapati, 2018) ประกอบด้วย

1. รากอยู่ด้านบนที่เรียกว่า “Root node” เป็นจุดเริ่มต้นการตัดสินใจของคุณลักษณะ

2. ใบไม้ที่อยู่ระหว่างรากและใบไม้ที่อยู่ด้านล่างสุด เรียกว่า “Intermediate node” โดยแต่ละ Node แสดงเป็นคุณลักษณะที่ไม่ใช่ Root node และทำหน้าที่แบ่งแยกข้อมูลเช่นเดียวกับ Root node

3. ใบไม้ที่อยู่ด้านล่างสุด เรียกว่า “Leaf node” แสดงเป็นตัวกำหนดคลาสของข้อมูลตัวอย่าง (Sample) ที่มีลักษณะเป็นค่าแบบไม่ต่อเนื่อง (Discrete) หรือค่าแบบต่อเนื่อง (Continuous)

โดยแสดงแผนผังต้นไม้การตัดสินใจ ดังภาพประกอบ 14




ภาพประกอบ 14 แสดงการทำงานของแบบจำลอง Decision Tree ในโครงสร้างของผังงาน

ที่มา: (z_ai, 2020a)

Decision Tree มีพารามิเตอร์การแบ่งแยกข้อมูลได้ดีคือ เอนโทรปี (Entropy) เป็นตัวชี้วัดที่ดีของข้อมูลที่มีหลายเหตุการณ์ โดยสมมติตัวแปรต่าง ๆ ดังนี้ (Guleria, Thakur, และ Sood, 2014)

1. จำนวนเหตุการณ์ทั้งหมด $n$ เหตุการณ์
2. กลุ่มตัวอย่างในชุดข้อมูลการเรียนรู้ (Training set) แสดงด้วย $S$
3. เอนโทรปีของชุดกลุ่มตัวอย่าง $S$ แสดงโดย $H(S)$
4. ความน่าจะเป็นของเหตุการณ์ถูกกำหนดด้วย $P_i$
5. จำนวนค่าที่เป็นเป้าหมายเท่ากับ $\mathrm{b}$

ดังนั้น Entropy สามารถคำนวณด้วยสูตรดังสมการ (11)

$$H(S) = \sum_{i=1}^{n} -P_i log_b P_i \quad (11)$$

เมื่อ $\sum_{i=1}^{n} P_i = 1$

นำค่า Entropy นี้ไปคำนวณค่า Information Gain ซึ่งเป็นการวัดที่บ่งบอกว่าคุณลักษณะแบ่งแยกข้อมูลได้ดีเพียงใด หากคุณลักษณะให้ค่า Information Gain มากที่สุด ต้นไม้จะเลือกคุณลักษณะนี้เป็นราก $A_i$ และแตกกิ่งก้านออก โดยแบ่งข้อมูลการเรียนรู้เรื่อย ๆ จนได้จำนวนข้อมูลแต่ละคลาส เปรียบเสมือนการแตกใบไม้ที่มีลักษณะตามอายุของใบไม้ การคำนวณค่า Information Gain ดังสมการ (12)

$$IG(S, A_i) = H(S) - \sum_{v \in Values(A_i)}^{n} P(A_i = v) H(S_v) \quad (12)$$

เมื่อ $v$ หมายถึง ค่าของลักษณะประจำ $(Value(A_i))$

$S_v = \{s | s \in S_v, Value(s, a) = v\}$ โดยที่ $a \in A$

การศึกษาวิจัยนี้มีคุณลักษณะที่มีค่าจำนวนมากที่ไม่ซ้ำกัน จึงใช้ค่า Gain Ratio (GR) แทน เนื่องจากค่า Information Gain ของคุณลักษณะนี้จะมีค่าจำนวนมาก ทำให้ถูกเลือกเป็นรากและแบ่งแยกข้อมูลก่อน ทำให้มีจำนวนกิ่งมาก ค่า Gain Ratio จะมาก เมื่อข้อมูลถูกแบ่งเท่า ๆ กัน และค่า Gain Ratio น้อย เมื่อข้อมูลอยู่คลาสเดียว (Gulati, Sharma, และ Gupta, 2016) โดยคำนวณค่า Gain Ratio ทำได้ ดังสมการ (13)

$$GR(S, A_i) = \frac{IG(S,A_i)}{SI(S,A_i)} \quad (13)$$

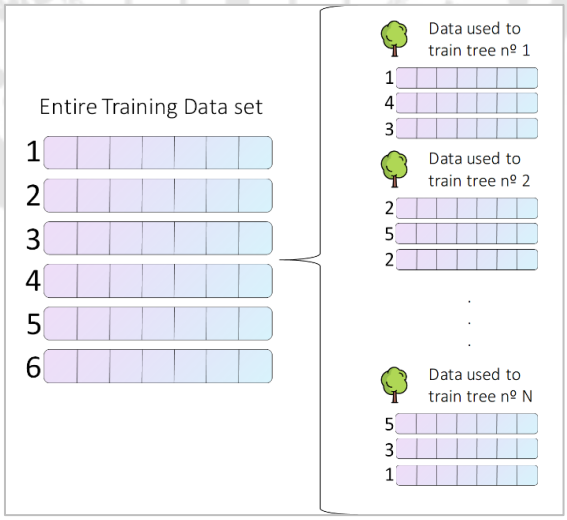
เมื่อ $SI$ เป็น Split Information โดยคำนวณจากสมการ (14)

$$SI(S, A_i) = -\sum_{v \in Values(A_i)}^{n} \frac{|S_v|}{|S|} log_b \frac{|S_v|}{|S|} \quad (14)$$

**6. แบบจำลอง Random Forest**

Random Forest เป็นแบบจำลองทั้งการจำแนกประเภทและการวิเคราะห์การถดถอย เช่นเดียวกับ Decision Tree มีหลักการใช้วิธี Ensemble Learning โดยนำ Decision Tree หลายต้นที่ไม่ซ้ำกันมาทำงานร่วมกัน สามารถแก้ปัญหาความไม่ยืดหยุ่นของต้นไม้ตัดสินใจและการเกิด Overfitting ง่าย และเป็นวิธีที่ช่วยเพิ่มประสิทธิภาพการทำนายของข้อมูลทดลอง (Test set) ด้วยแบบจำลองต้นไม้ตัดสินใจแต่ละตัว Random Forest เป็นส่วนหนึ่งของเทคนิค Bagging (J. Ali, Khan, Ahmad, และ Maqsood, 2012; Jayaprakash, Krishnan, และ Jaiganesh, 2020) โดยมีหลักการทำงาน 2 วิธี ดังนี้

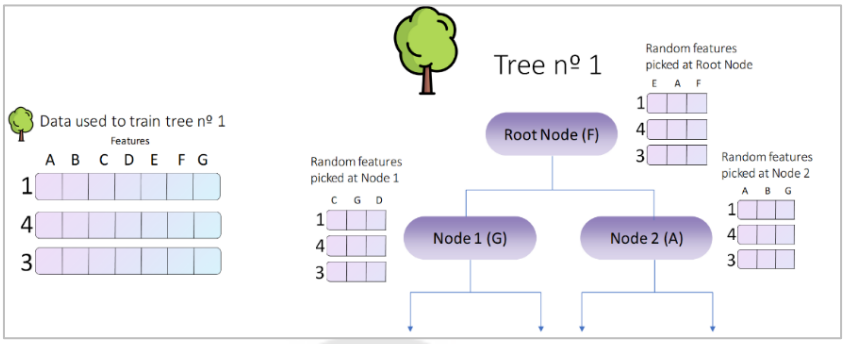
1. Bootstrapping เป็นเทคนิคที่เพิ่มปริมาณข้อมูล และไม่ได้ทำให้ข้อมูลมีความหลากหลายเพิ่ม โดยจัดการข้อมูลตามแนวตั้ง จะได้ชุดข้อมูลที่ต่างกับชุดข้อมูลดั้งเดิมหลาย ๆ ชุด ดังภาพประกอบ 15




ภาพประกอบ 15 แสดงการทำงานของแบบจำลอง Random Forest สำหรับการสร้างชุดข้อมูลด้วยเทคนิค Bootstrapping แต่ละต้น

ที่มา: (z_ai, 2020b)

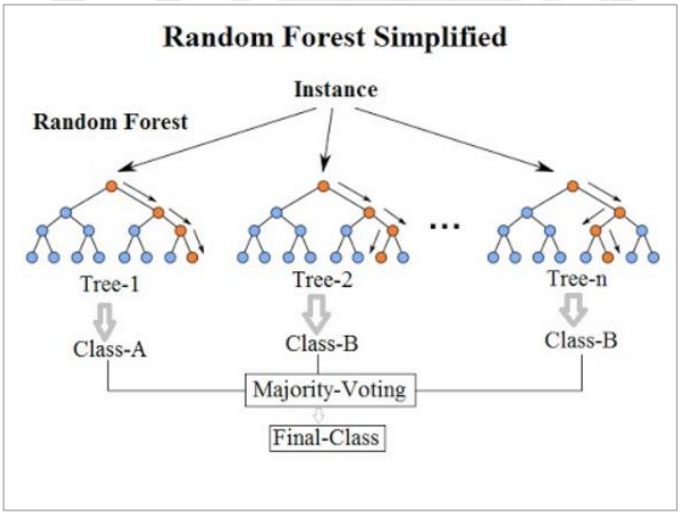
2. Random Feature Projection เป็นเทคนิคที่สุ่มเลือกคุณลักษณะแต่ละแบบจำลอง โดยจัดการข้อมูลตามแนวนอน ทำให้ได้ชุดข้อมูลที่ต่างกันหลายชุด ดังภาพประกอบ 16




ภาพประกอบ 16 แสดงการทำงานของแบบจำลอง Random Forest สำหรับการสุ่มเลือกคุณลักษณะของชุดข้อมูลที่ได้จากการใช้วิธี Bootstrapping

ที่มา: (z_ai, 2020b)

นำเทคนิค Bootstrapping และเทคนิค Random Feature Projection มาทำงานร่วมกันหลายแบบจำลอง จะได้คลาสที่ได้จากการทำนายแต่ละแบบจำลอง แล้วเลือกคลาสส่วนใหญ่ (Majority Vote) ซึ่งก็คือ ผลสรุปการทำนายของทุกแบบจำลองนั่นเอง ดังภาพประกอบ 17




ภาพประกอบ 17 แสดงการทำงานของแบบจำลอง Random Forest สำหรับการทำงานร่วมกันของแบบจำลอง Decision Tree หลายต้น

ที่มา: (Wikipedia, 2021c)

**7. แบบจำลอง XGBoost**

XGBoost ย่อมาจาก Extreme Gradient Boosting เป็นแบบจำลองทั้งการจำแนกประเภทและการวิเคราะห์การถดถอยที่อยู่ในตระกูล Decision Tree ที่นิยมใช้ในการแข่งขัน Kaggle สามารถถูกนำไปใช้ประยุกต์ในหลาย ๆ ด้านสำหรับการทำนาย ช่วยเพิ่มประสิทธิภาพการทำนายในด้านความแม่นยำ (Accuracy) ความไว (Sensitivity) และความจำเพาะ (Specificity) ของข้อมูลการทดสอบ (K. Yan, 2021) ช่วยลดการเกิด Overfitting สามารถจัดการข้อมูลที่ขาดหาย นอกจากนี้ยังสามารถแก้ปัญหาได้จริงโดยใช้ทรัพยากรการประมวลผลเพียงเล็กน้อย (Chen และ Guestrin, 2016) และเป็นเทคนิค Boosting ซึ่งเป็นวิธีการทั่วไปที่สร้างตัวแยกประเภทที่แข็งแกร่ง (Strong Classifier) จากตัวแยกประเภทที่อ่อนแอ (Weak Classifier) โดยเรียนรู้ข้อมูลดั้งเดิมด้วยแบบจำลอง Decision Tree แล้วสร้างแบบจำลองใหม่ เพื่อแก้ไขการเกิดความผิดพลาดการทำนายจากแบบจำลองครั้งแรก (Brownlee, 2016) การทำงานของ Boosting เป็นการเรียนรู้ต่อเนื่องกันแบบ Sequential fashion โดยมีหลักการทำงานของ XGBoost (Saraswat, 2016) ตามภาพประกอบ 18 ดังนี้

1. นำชุดข้อมูลดั้งเดิม (ตัวลบสีน้ำเงิน และตัวบวกสีแดง) เข้าสู่แบบจำลองที่ 1 โดยมี Weight ของข้อมูลแต่ละตัวเท่ากัน

2. เมื่อได้ผลการทำนายด้วยแบบจำลองที่ 1 ได้ผล ดังนี้

2.1 ตัวที่ทำนายผิดพลาด (ตัวลบสีน้ำเงิน 2 ตัว ตกในช่องสีแดง ทั้ง ๆ ที่ต้องอยู่ในช่องสีน้ำเงิน และตัวบวกสีแดง 1 ตัว ตกในช่องสีน้ำเงิน ทั้ง ๆ ที่ต้องอยู่ในช่องสีน้ำแดง)

2.2 ตัวที่ทำนายถูกต้อง (ตัวลบสีน้ำเงิน 3 ตัว อยู่ในช่องสีน้ำเงิน และตัวบวกสีแดง 4 ตัว อยู่ในช่องสีแดง อย่างถูกต้อง)

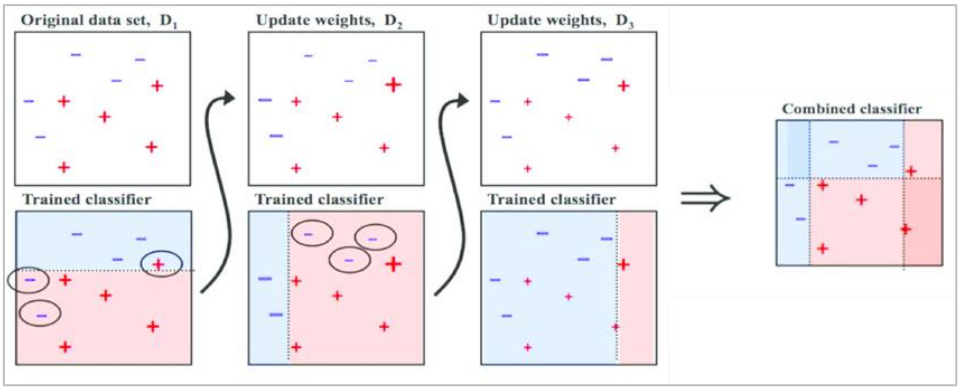
3. เพิ่ม Weight ให้กับตัวที่ทำนายผิดพลาด และลด Weight ให้กับตัวที่ทายถูกต้อง โดยใช้วิธี Gradient Descent ในการหา Weight ที่ทำให้เกิดค่าความผิดพลาดน้อยที่สุด

4. ต่อไปนำข้อมูลที่เพิ่มและลด Weight แล้วทำนายด้วยแบบจำลองที่ 2 ต่อไป โดยแบบจำลองที่ 2 ต้องเป็นอิสระต่อแบบจำลองที่ 1

5. ทำงานตามข้อ 2) - ข้อ 4) เรื่อย ๆ จนกว่าจะได้จำนวนแบบจำลองตามที่กำหนด

6. ทำการรวมผลการทำนายแต่ละแบบจำลองที่ได้ โดยใช้ Weighted Average ที่ได้เรียนรู้จากความผิดพลาดของแบบจำลองครั้งก่อน

การทำงานของ XGBoost เหล่านี้ เพื่อช่วยลดการทำนายข้อมูลผิดพลาด โดยเรียนรู้จากแบบจำลองก่อนหน้านี้เรื่อย ๆ เป็น Sequential fashion

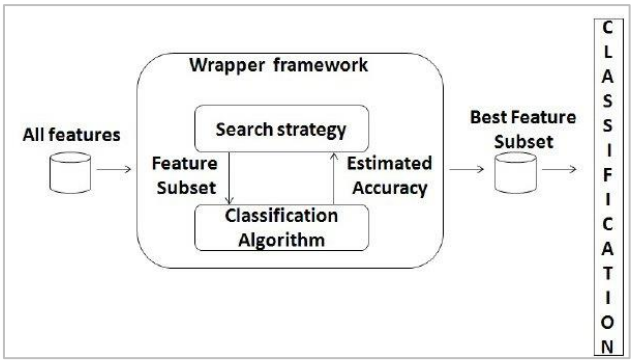



ภาพประกอบ 18 แสดงตัวอย่างการทำงานของแบบจำลอง XGBoost จากการเรียนรู้ข้อมูล 3 ครั้ง โดยแบ่งข้อมูลเป็น 2 คลาส ได้แก่ คลาสบวก และคลาสลบ

ที่มา: (Alto, 2020)

**การคัดเลือกคุณลักษณะ** (Feature Selection)

การคัดเลือกคุณลักษณะ มีผลต่อประสิทธิภาพของแบบจำลองเป็นอย่างมาก มีหนึ่งในเทคนิคการคัดเลือกคุณลักษณะที่เรียกว่า “Feature Importance” ซึ่งบ่งบอกว่าคุณลักษณะใดมีความสำคัญกับแบบจำลองมากน้อยเพียงใด เมื่อเทียบกับคุณลักษณะอื่น ๆ โดยพิจารณาจากคะแนนความสำคัญของคุณลักษณะที่ได้มาจากแบบจำลองที่คำนวนคะแนนความสำคัญให้ เพื่อจัดอันดับความสำคัญของแต่ละคุณลักษณะ และเลือกคุณลักษณะที่ส่งผลต่อแบบจำลองการทำนายมากที่สุดด้วยจำนวนคุณลักษณะที่กำหนด หากคุณลักษณะใดที่ได้คะแนนความสำคัญมาก แสดงว่ามีค่าผิดพลาด (Loss) มากจากการทำนาย คุณลักษณะนี้จะถูกจัดเป็นอันดับแรกก่อน สามารถใช้กับคุณลักษณะทั้งแบบค่าต่อเนื่องและแบบค่าไม่ต่อเนื่อง วิธีการนี้เป็นวิธีที่เรียกว่า “Wrapper Method” ซึ่งเป็นวิธีการอย่างง่าย โดยคัดเลือกกลุ่มคุณลักษณะย่อยต่าง ๆ จากคุณลักษณะทั้งหมดไปเรียนรู้กับชุดข้อมูลการเรียนรู้ ต่อมาวัดประสิทธิภาพของแบบจำลองที่สนใจสำหรับแต่ละกลุ่มคุณลักษณะย่อย จากนั้นนำกลุ่มคุณลักษณะย่อยที่ให้ผลลัพธ์ดีที่ที่สุดไปเรียนรู้ด้วยแบบจำลองที่นำไปใช้จริง (Brownlee, 2020a) ดังภาพประกอบ 19




ภาพประกอบ 19 แสดงกระบวนการคัดเลือกคุณลักษณะด้วย Wrapped Method

ที่มา: (Bouaguel, 2015)

**การวัดประสิทธิภาพของแบบจำลอง** (Model Evaluation)

หลังจากสร้างแบบจำลองประเภทการจำแนกประเภทแล้ว ต้องการประเมินแบบจำลองว่าแบบจำลองมีประสิทธิภาพในระดับดีมากหรือน้อยแค่ไหน เครื่องมือการวัดประสิทธิภาพของแบบจำลองการจำแนกประเภท เรียกว่า “เมทริกซ์ความสับสัน (Confusion Matrix)” ซึ่งอยู่ในรูปของเมทริกซ์ที่เป็นตารางแสดงแถวและคอลัมน์ ซึ่งมีลักษณะเป็น 2 มิติ มีหลักการทำงานโดยนำตัวเลขต่าง ๆ ในเมทริกซ์ความสับสน ไปวัดความถูกต้อง (Accuracy), ความแม่นยำ (Precision), ความระลึก (Recall) และค่าวัดทั้งค่า Precision และค่า Recall พร้อมกัน (F1-score) (Almasri และคนอื่น ๆ, 2020) ดังภาพประกอบ 20

| | Predicted class POSITIVE (spam ✉) | Predicted class NEGATIVE (normal 🕵) | |
|---|---|---|---|
| Actual class POSITIVE (spam ✉) | TRUE POSITIVE (TP) ✉✉ 320 | FALSE NEGATIVE (FN) ✉🕵 43 | $Recall = \frac{TP}{TP+FN} = \frac{320}{320+43} = 0.882$ |
| Actual class NEGATIVE (normal ✉) | FALSE POSITIVE (FP) ✉✉ 20 | TRUE NEGATIVE (TN) ✉🕵 538 | |
| | $Precision = \frac{TP}{TP+FP} = \frac{320}{320+20} = 0.941$ | | |

ภาพประกอบ 20 แสดงตัวอย่างการวัดประสิทธิภาพของแบบจำลองการทำนายอีเมลว่าเป็นสแปมหรือไม่ โดยใช้ Confusion Matrix สำหรับข้อมูลที่มี 2 คลาส

ที่มา: (Widmann, 2019)

จากภาพประกอบ 20 เป็นการทำนายข้อมูลที่มีจำนวน 2 คลาส โดยมีการคำนวณสูตรการวัดค่าต่าง ๆ ในเมทริกซ์ความสับสน ดังสมการ (15) – (18)

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (15)$$

$$Precision = \frac{TP}{TP+FP} \quad (16)$$

$$Recall = \frac{TP}{TP+FN} \quad (17)$$

$$F1 - score = 2 \times \left(\frac{Recall \times Precision}{Recall+Precision}\right) \quad (18)$$

โดยตัวแปรต่าง ๆ จากสมการ (15) – (18) มีความหมาย ดังนี้

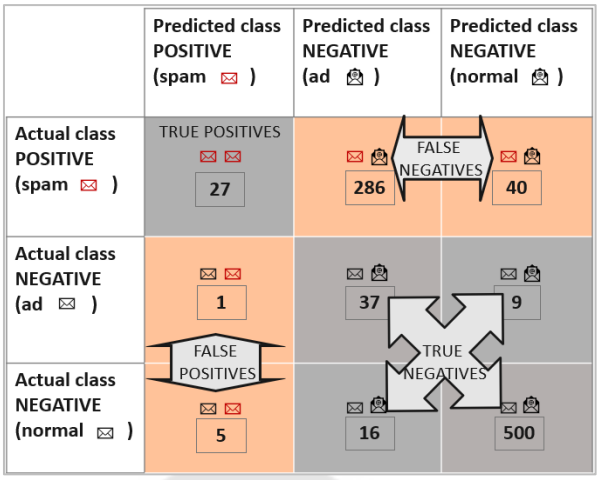
1. True Positive (TP) คือ ผลการทำนายว่าเป็นคลาสที่สนใจ เช่นเดียวกับข้อมูลแท้จริงเป็นคลาสที่สนใจ มีความหมายว่าเป็นการทำนายถูกต้อง

2. True Negative (TN) คือ ผลการทำนายว่าไม่ได้เป็นคลาสที่สนใจ เช่นเดียวกับข้อมูลแท้จริงไม่ได้เป็นคลาสที่สนใจ มีความหมายว่าเป็นการทำนายถูกต้อง

3. False Positive (FP) คือ ผลการทำนายว่าเป็นคลาสที่สนใจ แต่ข้อมูลแท้จริงไม่ได้เป็นคลาสที่สนใจ มีความหมายว่าเป็นการทำนายผิดพลาด

4. False Negative (FN) คือ ผลการทำนายว่าไม่ได้เป็นคลาสที่สนใจ แต่ข้อมูลแท้จริงเป็นคลาสที่สนใจ มีความหมายว่าเป็นการทำนายผิดพลาด

อีกกรณีหนึ่งสำหรับข้อมูลที่มีจำนวน 3 คลาสขึ้นไป สามารถคำนวณค่าวัดค่าต่าง ๆ ในเมทริกซ์ความสับสนเช่นเดียวกัน ดังภาพประกอบ 21




ภาพประกอบ 21 แสดงตัวอย่างการวัดประสิทธิภาพของแบบจำลองการทำนายอีเมลว่าอยู่ประเภทใด โดยใช้ Confusion Matrix สำหรับข้อมูลที่มี 3 คลาส

ที่มา: (Widmann, 2019)

**งานวิจัยที่เกี่ยวข้อง** (Literature Review)

1. **บทความวิจัย เรื่อง** Improved students' performance prediction for multi-class imbalanced problems using hybrid and ensemble approach in educational data mining (Ahmad, Anuar, & Hassan, 2020)

ผู้วิจัยได้ศึกษาพฤติกรรมการใช้งานระบบการจัดการเรียนรู้ของผู้เรียน และข้อมูลทั่วไป เพื่อทำนายระดับผลการเรียนเฉลี่ยของผู้เรียน โดยใช้ชุดข้อมูล 2 ชุด ที่มาจากระบบจัดเก็บข้อมูลผู้เรียน และระบบการเรียนรู้ออนไลน์ของผู้เรียนจากมหาวิทยาลัย Malaysia ในประเทศมาเลเซีย การนำข้อมูลเหล่านี้ไปจัดการความไม่สมดุลของข้อมูลหลายวิธี ทั้งเทคนิคการสุ่มเพิ่มข้อมูล 3 แบบ ได้แก่ Synthetic Minority Oversampling Technique: SMOTE, Random Oversampling: ROS และ Adaptive Synthetic Sampling: ADASYN เทคนิคการสุ่มลดข้อมูล 2 แบบ ได้แก่ Random Undersampling: RUS และ Near Miss Undersampling: NearMiss ส่วนเทคนิคการสุ่มทั้งการเพิ่มและลดข้อมูล 2 แบบ ได้แก่ SMOTE with Edited Nearest Neighbors: SMOTE-ENN, SMOTE-TL แล้วทำชุดคุณลักษณะที่แตกต่างกันคือ ชุดข้อมูลที่ 1 ชุดข้อมูลที่ 2 และชุดข้อมูล 2 ชุด พบว่าแบบจำลอง Bagging ที่ใช้คุณลักษณะของชุดข้อมูล 2 ชุด มีค่า F1-score มากที่สุด จากนั้นนำชุดข้อมูล 2 ชุดนี้ ไปใช้เทคนิคการคัดเลือกคุณลักษณะด้วยวิธี Filter Method นอกจากนี้ยังมีการเปรียบเทียบประสิทธิภาพการทำนายของแบบจำลองการจำแนกประเภท ได้แก่ Random Forest,

Bagging, AdaBoost, Gradient Boosting, XGBoost พบว่าแบบจำลอง AdaBoost ที่มีการสุ่มเพิ่มข้อมูลแบบ ROS มีความแม่นยำมากที่สุดเท่ากับ 92% แต่เนื่องจากหลังจากการจัดข้อมูลที่ไม่สมดุลด้วยการสุ่มเพิ่มข้อมูลแบบ SMOTE-ENN สรุปได้ว่าแบบจำลองส่วนใหญ่ให้ผลลัพธ์การทำนายที่ดีที่สุดจากการสุ่มตัวอย่างเพิ่มนี้ ดังนั้น SMOTE-ENN เป็นเทคนิคที่เหมาะสมที่สุดในการปรับความไม่สมดุลของข้อมูล ทั้งนี้ช่วยให้ติดตามความก้าวหน้าของผู้เรียน และสามารถบ่งบอกได้ว่าผู้เรียนมีความเสี่ยงที่จะได้ผลการเรียนไม่ผ่านในช่วงระยะแรก

2. **บทความวิจัย เรื่อง** Managing Student Performance: A Predictive Analytics using Imbalanced Data (Ashfaq **และคนอื่น ๆ**, 2020)

ผู้วิจัยได้ศึกษาพฤติกรรมการใช้งานระบบการจัดการเรียนรู้ของผู้เรียน และข้อมูลทั่วไปเพื่อทำนายระดับผลการเรียน โดยใช้ชุดข้อมูลของผู้เรียน 480 คน จากมหาวิทยาลัย Jordan ในประเทศจอร์แดน ก่อนอื่นนำข้อมูลไปทำนายโดยใช้เทคนิคการปรับความไม่สมดุลของข้อมูล 2 วิธี คือ Synthetic Minority Over-sampling Technique (SMOTE) และ Adaptive Synthetic Sampling (ADASYN) แล้วคัดเลือกคุณลักษณะโดยใช้เทคนิค 2 วิธี คือ Fast Correlation Based Feature selection (FCBF) และ Recursive Feature Elimination (RFE) จากนั้นเข้าสู่แบบจำลอง Classification 3 แบบ ได้แก่ Random Forest, Support Vector Machine และ Artificial Neural Network พบว่าแบบจำลอง Random Forest ที่คัดเลือกคุณลักษณะ มีความแม่นยำที่สุดเท่ากับ 86.74% หลังจากปรับข้อมูลให้มีความสมดุลด้วยวิธี ADASYN

3. **บทความวิจัย เรื่อง** Modeling and Predicting Students' Academic Performance Using Data Mining Techniques (Mueen **และคนอื่น ๆ**, 2016)

ผู้วิจัยได้ศึกษาพฤติกรรมการใช้งานระบบการจัดการเรียนรู้ของผู้เรียน และคะแนนเก็บเพื่อทำนายผลการเรียนของผู้เรียนว่าผ่านหรือไม่ผ่าน โดยใช้ชุดข้อมูลของผู้เรียนหลักสูตรการเขียนโปรแกรมระดับพื้นฐานและหลักสูตรการเขียนโปรแกรมระดับขั้นสูงในระบบการจัดการเรียนรู้จำนวน 60 คน แล้วนำข้อมูลเหล่านี้ไปจัดการความไม่สมดุลของข้อมูลด้วยวิธี Synthetic Minority Over-sampling Technique (SMOTE) ต่อมาคัดเลือกคุณลักษณะด้วย Filter Method จากนั้นเข้าสู่แบบจำลองการจำแนกประเภท ได้แก่ Naïve Bayes, Neural Network และ Decision Tree พบว่าแบบจำลอง Naïve Bayes ที่ใช้คุณลักษณะทั้งหมดและจัดการข้อมูลที่ไม่สมดุลด้วยวิธี SMOTE มีความแม่นยำที่สุดเท่ากับ 86% การทำนายนี้ช่วยให้อาจารย์ทราบล่วงหน้าว่าผู้เรียนมีความเสี่ยงที่จะล้มเหลวในการเรียนหรือไม่ และช่วยเหลือผู้เรียนได้อย่างเจาะจง

4. **บทความวิจัย เรื่อง** Mining Educational Data to Predict Student's academic Performance using Ensemble Methods (Amrieh **และคนอื่น ๆ**, 2016)

ผู้วิจัยได้ศึกษาพฤติกรรมการใช้งานระบบการจัดการเรียนรู้ของผู้เรียน ข้อมูลทั่วไป และผลตอบประเมินความพึงพอใจของผู้ปกครองต่อการจัดการเรียนการสอน เพื่อทำนายระดับผลการเรียน โดยใช้ชุดข้อมูลของผู้เรียนจำนวน 500 คน ต่อมาทำชุดคุณลักษณะที่แตกต่างกันจำนวน 2 ชุด ได้แก่ ชุดที่ 1 คุณลักษณะทั้งหมด และชุดที่ 2 คุณลักษณะที่ไม่ได้ใช้ข้อมูลพฤติกรรมการใช้งาน จากนั้นคัดเลือกคุณลักษณะด้วยการใช้ Filter Method ต่อไปเข้าสู่แบบจำลอง Artificial Neural Network, Naïve Bayes และ Decision Tree ร่วมกับการทำ Bagging, Boosting และ Random Forest พบว่าแบบจำลองสำหรับคุณลักษณะที่ใช้พฤติกรรมของผู้เรียนมีประสิทธิภาพที่ดีกว่า เมื่อเทียบกับคุณลักษณะที่ไม่ได้ใช้พฤติกรรมของผู้เรียน โดยมีความแม่นยำกว่า 22.1% ส่วนแบบจำลองที่ร่วมการทำ Ensemble Method 3 วิธี ได้แก่ Bagging, Boosting และ Random Forest มีประสิทธิภาพที่ดีกว่าแบบจำลองสำหรับคุณลักษณะที่ไม่ได้ร่วมการทำ Ensemble Method โดยมีความแม่นยำกว่า 25.8% พบว่าแบบจำลองทุกแบบที่ร่วมการทำ Boosting ให้ผลลัพธ์การทำนายได้ดีที่สุดเป็นส่วนใหญ่ สรุปได้ว่าแบบจำลอง Decision Tree ที่ร่วมกับการทำ Boosting ที่มีการคัดเลือกคุณลักษณะ มีความแม่นยำที่สุดเท่ากับ 85%

5. **บทความวิจัย เรื่อง** Predicting Students Final Academic Performance using Feature Selection Approaches (Ramaswami **และคนอื่น ๆ**, 2020)

ผู้วิจัยได้ศึกษาพฤติกรรมการใช้งานระบบการจัดการเรียนรู้ของผู้เรียน คะแนนเก็บ และข้อมูลทั่วไป เพื่อทำนายผลการเรียนของผู้เรียนว่าผ่านหรือไม่ผ่าน โดยใช้ขุดข้อมูลของผู้เรียนจำนวน 112 คน จากมหาวิทยาลัยใน Australasian แล้วไปคัดเลือกคุณลักษณะด้วยวิธี Embedded Method ต่อมาทำนายด้วยแบบจำลองการจำแนกประเภท ได้แก่ Naïve Bayes, Random Forest, Logistic Regression และ K-Nearest Neighbor พบว่าแบบจำลอง Logistic Regression ที่มีการคัดเลือกคุณลักษณะ มีค่า F1-score สูงถึง 85.29% และทุกแบบจำลองที่มีการคัดเลือกคุณลักษณะ มีค่า F1-score สูงกว่า เมื่อเทียบกับการที่ไม่มีการคัดเลือกคุณลักษณะ การทำนายความเสี่ยงที่จะล้มเหลวในการเรียนหลักสูตรนี้มีประโยชน์ในการเข้าช่วยเหลือผู้เรียนในช่วงระยะแรก เพื่อช่วยปรับปรุงประสิทธิภาพการเรียนการสอนตลอดหลักสูตร

6. **บทความวิจัย เรื่อง** Educational data mining in moodle data (Shrestha **และ** Pokharel, 2021)

ผู้วิจัยได้ศึกษาพฤติกรรมการใช้งานในระบบการจัดการเรียนรู้ของผู้เรียน คะแนนเก็บ และข้อมูลทั่วไป โดยใช้ข้อมูลของผู้เรียนจำนวน 128 คน จากมหาวิทยาลัย Kathmandu ในประเทศ

เนปาล เพื่อทำนายระดับคะแนนแบบทดสอบของผู้เรียน นำข้อมูลไปคัดเลือกคุณลักษณะด้วย Wrapper Method แล้วทำนายด้วยแบบจำลองการจำแนกประเภทต่าง ๆ ได้แก่ K-Nearest Neighbor, Naïve Bayes, Support Vector Machine, Random Forest และ CART Decision Tree พบว่า แบบจำลอง Support Vector Machine ที่คัดเลือกคุณลักษณะ มีความแม่นยำสูงที่สุดเท่ากับ 93.94% ส่วนแบบจำลอง Support Vector Machine และ CART Decision Tree ที่ใช้เพียง 2 คุณลักษณะ มีความแม่นยำเท่ากับการใช้คุณลักษณะทั้งหมด 8 คุณลักษณะ ทั้งนี้สามารถปรับปรุงการเรียนการสอนให้คำปรึกษาแก่ผู้เรียน ลดการลาออกกลางคันของผู้เรียน

7. **บทความวิจัย เรื่อง** Massive LMS log data analysis for the early prediction of course-agnostic student performance (Riestra-González **และคนอื่น ๆ**, 2021)

ผู้วิจัยได้ศึกษาพฤติกรรมการใช้งานระบบการจัดการเรียนรู้ของผู้เรียน และคะแนนเก็บเพื่อทำนายผลการเรียนของผู้เรียนว่าผ่านหรือไม่ผ่าน โดยใช้ชุดข้อมูลของผู้เรียนจำนวน 29,602 คน จากมหาวิทยาลัย Oviedo ในประเทศสเปน ซึ่งมีหลักสูตรทั้งหมด 5,112 มีการทำคุณลักษณะที่แตกต่างกัน โดยใช้ข้อมูลการเรียนหลักสูตรเพียง 10%, 25%, 33% และ 50% เริ่มจากการแบ่งกลุ่มตามรูปแบบปฏิสัมพันธ์การเรียนรู้ในระบบของผู้เรียนด้วยวิธี K-means Clustering จะได้จำนวน 4 กลุ่มที่มีความสัมพันธ์กับผลการเรียนของผู้เรียนจากจำนวน 6 กลุ่มทั้งหมด จากนั้นคัดเลือกคุณลักษณะโดยใช้วิธี Recursive Feature Elimination and Cross-Validation Selection (RFECV) และใช้แบบจำลองการจำแนกประเภทต่าง ๆ ได้แก่ Decision Tree, Naïve Bayes, Logistic Regression, Multilayer Perceptron และ Support Vector Machine พบว่าแบบจำลอง Multiple perceptron มีประสิทธิภาพการทำนายดีที่สุด มีความแม่นยำเท่ากับ 80.1% โดยใช้ข้อมูลการเรียนหลักสูตรเพียง 10% เท่านั้น

8. **บทความวิจัย เรื่อง** Prediction of student academic performance using Moodle data from a Further Education setting (Quinn **และ** Gray, 2019)

ผู้วิจัยได้ศึกษาพฤติกรรมการใช้งานระบบการจัดการเรียนรู้ของผู้เรียน เพื่อทำนายระดับผลการเรียนของผู้เรียน โดยใช้ข้อมูลการใช้งานระบบทั้งหมดตลอดหลักสูตร และทำนายผลการเรียนของผู้เรียนว่าผ่านหรือไม่ผ่าน โดยใช้คุณลักษณะที่แตกต่างกัน 2 ชุด คือ ข้อมูลการเรียน 6 สัปดาห์แรก และ 10 สัปดาห์แรก ข้อมูลเหล่านี้มาจากชุดข้อมูลของผู้เรียน 690 คน จากสถาบันการศึกษา Limerick and Clare Education and Training Board ในประเทศไอร์แลนด์ นำข้อมูลเข้าสู่แบบจำลอง Random Forest, Gradient Boosting, K-Nearest Neighbor และ Linear Discriminant Analysis พบว่าความแม่นยำการทำนายระดับผลการเรียนมากที่สุดเท่ากับ 60.5% ด้วยแบบจำลอง Random Forest และความแม่นยำการทำนายผลการเรียนของผู้เรียนผ่านหรือไม่ผ่าน โดย

มีค่ามากที่สุดเท่ากับ 82.18% ด้วยแบบจำลอง Random Forest เช่นเดียวกัน โดยใช้ข้อมูลการเรียน 10 สัปดาห์แรก ทั้งนี้ ช่วยให้แจ้งเตือนผู้เรียนล่วงหน้า

9. **บทความวิจัย เรื่อง** Mining Educational Data to Predict Students' Performance through Procrastination Behavior (Hooshyar **และ** Pedaste, 2019)

ผู้วิจัยได้ศึกษาพฤติกรรมการส่งงานมอบหมายในระบบการจัดการเรียนรู้ของผู้เรียน และประเภทของสื่อการเรียนรู้ เพื่อทำนายระดับผลการเรียนของผู้เรียน โดยใช้ข้อมูลของผู้เรียน 242 คน จากมหาวิทยาลัย Tartu ในประเทศ Estonia เพื่อทำนายระดับผลการเรียนด้วยเทคนิค การจัดกลุ่ม และการจำแนกประเภท เริ่มจากการสร้างสถานะการส่งงานมอบหมายเป็นคุณลักษณะ จากวันที่ระบบเปิดการส่งงานมอบหมาย วันแรกที่ชมงานมอบหมาย และวันสุดท้ายการส่งงานมอบหมาย จากนั้นเข้าสู่แบบจำลองการแบ่งกลุ่มด้วยวิธี K-means Clustering และแบบจำลองการจำแนกประเภท ได้แก่ Support Vector Machine, Gaussian processes, Decision Tree, Random Forest, Neural Network, AdaBoost และ Naïve Bayes พบว่าสามารถแบ่งกลุ่มพฤติกรรมการส่งงานมอบหมายตามประเภทการผัดกันประกันพรุ่งล่วงหน้า และทำนายระดับผลการเรียนจากพฤติกรรมของผู้เรียนที่แบ่งกลุ่มนั้นด้วยความแม่นยำที่สุดเท่ากับ 96% จากการใช้แบบจำลอง Support Vector Machine สำหรับข้อมูลที่เป็นตัวเลข (Numerical feature) ร่วมกับแบบจำลอง Neural Network สำหรับข้อมูลที่เป็นหมวดหมู่ (Categorical feature)

10. **บทความวิจัย เรื่อง** Clustering-Based EMT Model for Predicting Student Performance (Almasri **และคนอื่น ๆ**, 2020)

ผู้วิจัยได้ศึกษาผลการเรียนแต่ละวิชาของผู้เรียน และข้อมูลทั่วไป เพื่อทำนายระดับผลการเรียนเฉลี่ยสะสม โดยใช้ชุดข้อมูลของผู้เรียน 1,062 คน จากมหาวิทยาลัย Balqa Applied ในประเทศจอร์แดน เริ่มจากการกำหนดจำนวนกลุ่มที่เหมาะสมด้วยวิธี Canopy Clustering จากนั้นแบ่งกลุ่มด้วยวิธี K-means Clustering แล้วทำการคัดเลือกคุณลักษณะด้วย Wrapper Method ต่อไปเข้าสู่แบบจำลองการจำแนกประเภทต่าง ๆ เช่น Multilayer Perceptron, J48 Decision Trees, Naïve Bayes และ Ensemble Meta-based Tree พบว่าแบบจำลองที่ร่วมการแบ่งกลุ่มและการจำแนกประเภทได้ผลดีกว่า เมื่อเทียบกับแบบจำลองการจำแนกประเภทเท่านั้น สรุปได้ว่าแบบจำลองการจัดกลุ่มร่วมกับแบบจำลอง Ensemble Meta-based Tree ที่คัดเลือกคุณลักษณะเพียง 50% มีความแม่นยำที่สุดเท่ากับ 96.96%

จากการศึกษาบทความวิจัยที่เกี่ยวข้องทั้งหมดจำนวน 10 เรื่อง ผู้วิจัยจึงรวบรวมบทความวิจัยเหล่านี้ไปสรุปประเภทของชุดข้อมูลที่ใช้ในการวิจัย การเปรียบเทียบวิธีการดำเนินงานวิจัยแต่ละเรื่อง และผลลัพธ์การทำนายที่ดีที่สุด ดังตาราง 1

ตาราง 1 แสดงสรุปงานวิจัยที่เกี่ยวข้องทั้งหมดจำนวน 10 เรื่อง

| ที่ | ผู้วิจัย | ชุดข้อมูลแต่ละด้าน | | | วิธีการดำเนินงานวิจัย (ตัวพิมพ์หนา หมายถึง วิธีที่ให้ผลลัพธ์การทำนายที่ดีที่สุด) | | | | ผลลัพธ์การทำนาย |
|---|---|---|---|---|---|---|---|---|---|
| | | **พฤติกรรมการใช้ระบบ** | **คะแนน** | **ข้อมูลทั่วไป** | **ทำวิศวกรรมคุณลักษณะ** | **จัดการความไม่สมดุลของข้อมูล** | **คัดเลือกคุณลักษณะ** | **แบบจำลอง** | **ประสิทธิภาพของแบบจำลองที่ดีที่สุด** |
| 1 | Ahmad, Anuar, และ Hassan (2020) | ✓ | | | ลำดับ 2<br>**ชุดที่ 1 คุณลักษณะที่ใช้ข้อมูลระบบจัดเก็บข้อมูลผู้เรียน**<br>**ชุดที่ 2 คุณลักษณะที่ใช้ข้อมูลระบบการเรียนรู้ออนไลน์ของผู้เรียน** | ลำดับ 1<br>Synthetic Minority Oversampling Technique (SMOTE), **Random Oversampling: ROS**, Adaptive Synthetic Sampling (ADASYN), Random Undersampling (RUS), Near Miss Undersampling (NearMiss), SMOTE with Edited Nearest Neighbors (SMOTE-ENN), SMOTE-TL | ลำดับ 3<br>**Filter Method** | Random Forest, Bagging, **AdaBoost**, Gradient Boosting, XGBoost | ความแม่นยำการทำนายระดับผลการเรียนเฉลี่ยที่สุด 92% โดยใช้ชุดข้อมูล 2 ชุด |

ตาราง 1 (ต่อ)

| ที่ | ผู้วิจัย | ชุดข้อมูลแต่ละด้าน | | | วิธีการดำเนินงานวิจัย (ตัวพิมพ์หนา หมายถึง วิธีที่ให้ผลลัพธ์การทำนายที่ดีที่สุด) | | | ผลลัพธ์การทำนาย | |
|---|---|---|---|---|---|---|---|---|---|
| | | พฤติกรรมการใช้ระบบ | คะแนน | ข้อมูลทั่วไป | ทำวิศวกรรมคุณลักษณะ | จัดการความไม่สมดุลของข้อมูล | คัดเลือกคุณลักษณะ | แบบจำลอง | ประสิทธิภาพของแบบจำลองที่ดีที่สุด |
| 2 | Ashfaq, Poolan Marikannan, และ Raheem (2020) | ✓ | | ✓ | - | ลำดับ 1 Synthetic Minority Over-sampling Technique (SMOTE), **Adaptive Synthetic Sampling (ADASYN)** | ลำดับที่ 2 Fast Correlation Based Feature selection (FCBF), **Recursive Feature Elimination (RFE)** | **Random Forest**, Support Vector Machine และ Artificial Neural Network | ความแม่นยำการทำนายระดับผลการเรียน 86.74% |
| 3 | Mueen, Zafar, และ Manzoor (2016) | ✓ | ✓ | | - | ลำดับที่ 2 **Synthetic Minority Over-sampling Technique (SMOTE)** | ลำดับที่ 1 Filter Method, **All Feature** | **Naïve Bayes**, Neural Network และ Decision Tree | ความแม่นยำการทำนายผลการเรียนว่าผ่านหรือไม่ผ่าน 86% |
| 4 | Amrieh, Hamtini, และ Aljarah (2016) | ✓ | | ✓ | ลำดับที่ 1 **ชุดที่ 1 คุณลักษณะที่ใช้ข้อมูลพฤติกรรมการใช้งาน** ชุดที่ 2 คุณลักษณะที่ไม่ได้ใช้ข้อมูลพฤติกรรมการใช้งาน | - | ลำดับที่ 2 **Filter Method** | Artificial Neural Network, Naïve Bayes และ **Decision Tree ร่วมกับการทำ** Bagging, **Boosting** และ Random Forest | ความแม่นยำการทำนายระดับผลการเรียน 85% โดยใช้ข้อมูลพฤติกรรมการใช้งาน |

ตาราง 1 (ต่อ)

| ที่ | ผู้วิจัย | ชุดข้อมูลแต่ละด้าน | | | วิธีการดำเนินงานวิจัย (ตัวพิมพ์หนา หมายถึง วิธีที่ให้ผลลัพธ์การทำนายที่ดีที่สุด) | | | ผลลัพธ์การทำนาย | |
|---|---|---|---|---|---|---|---|---|---|
| | | พฤติกรรมการใช้ระบบ | คะแนน | ข้อมูลทั่วไป | ทำวิศวกรรมคุณลักษณะ | จัดการความไม่สมดุลของข้อมูล | คัดเลือกคุณลักษณะ | แบบจำลอง | ประสิทธิภาพของแบบจำลองที่ดีที่สุด |
| 5 | Ramaswami, Susnjak, Mathrani, และ Umer (2020) | ✓ | ✓ | ✓ | - | - | ลำดับที่ 1 **Embedded Method** | Naïve Bayes, Random Forest, **Logistic Regression** และ K-Nearest Neighbor | ค่า F1-score การทำนายทำนายผลการเรียนว่าผ่านหรือไม่ผ่าน 85.29% |
| 6 | Shrestha และ Pokharel (2021) | ✓ | ✓ | | - | - | ลำดับที่ 1 **Wrapper Method** | K-Nearest Neighbor, Naïve Bayes, **Support Vector Machine**, Random Forest และ CART Decision Tree | ความแม่นยำการทำนายระดับคะแนนแบบทดสอบ 93.94% |
| 7 | Riestra-González, Paule-Ruíz, และ Ortin (2021) | ✓ | ✓ | | ลำดับที่ 1 ชุดคุณลักษณะ 4 ชุด ที่ใช้ **ข้อมูลการเรียน** 10%, 25%, 33% และ 50% ของหลักสูตร | - | ลำดับที่ 2 **Recursive Feature Elimination and Cross-Validation Selection (RFECV)** | **K-means Clustering ร่วมกับ** Decision Tree, Naïve Bayes, Logistic Regression, **Multilayer Perceptron** และ Support Vector Machine | ความแม่นยำการทำนายผลการเรียนว่าผ่านหรือไม่ผ่าน 80.1% โดยใช้ข้อมูลการเรียนเพียง 10% ของหลักสูตร |

ตาราง 1 (ต่อ)

| ที่ | ผู้วิจัย | ชุดข้อมูลแต่ละด้าน | | | วิธีการดำเนินงานวิจัย (ตัวพิมพ์หนา หมายถึง วิธีที่ให้ผลลัพธ์การทำนายที่ดีที่สุด) | | | ผลลัพธ์การทำนาย | |
|---|---|---|---|---|---|---|---|---|---|
| | | พฤติกรรมการใช้ระบบ | คะแนน | ข้อมูลทั่วไป | ทำวิศวกรรมคุณลักษณะ | จัดการความไม่สมดุลของข้อมูล | คัดเลือกคุณลักษณะ | แบบจำลอง | ประสิทธิภาพของแบบจำลองที่ดีที่สุด |
| 8 | Quinn และ Gray (2019) | ✓ | | | ลำดับที่ 1 **ชุดคุณลักษณะที่ทำนายระดับผลการเรียน โดยใช้ข้อมูลการเรียนตลอดหลักสูตร** **ชุดคุณลักษณะที่ทำนายผลการเรียนว่าผ่านหรือไม่ผ่าน โดยใช้ข้อมูลการเรียน** 6 สัปดาห์แรก และ **10 สัปดาห์แรก** | - | - | **Random Forest**, Gradient Boosting, K-Nearest Neighbor และ Linear Discriminant Analysis | - ความแม่นยำการทำนายระดับผลการเรียน 60.5% โดยใช้ข้อมูลการเรียนตลอดหลักสูตร<br>- ความแม่นยำการทำนายผลการเรียนว่าผ่านหรือไม่ผ่านที่สุด 82.18% โดยใช้ข้อมูลการเรียน 6 สัปดาห์ |
| 9 | Hooshyar และ Pedaste (2019) | ✓ | | | ลำดับที่ 1 **สร้างคุณลักษณะสถานะการส่งงานมอบหมาย**จากวันที่ระบบเปิดการส่งงานมอบหมาย วันแรกที่ชมงานมอบหมาย และวันสุดท้ายการส่งงานมอบหมาย | - | - | **K-means Clustering ร่วม Support Vector Machine (For Numerical Feature)**, Gaussian processes, Decision Tree, Random Forest, **Neural Network (For Categorical Feature)**, AdaBoost และ Naïve Bayes | ความแม่นยำการทำนายระดับผลการเรียน 96% |

ตาราง 1 (ต่อ)

| ที่ | ผู้วิจัย | ชุดข้อมูลแต่ละด้าน | | | วิธีการดำเนินงานวิจัย (ตัวพิมพ์หนา หมายถึง วิธีที่ให้ผลลัพธ์การทำนายที่ดีที่สุด) | | | | ผลลัพธ์การทำนาย |
|---|---|---|---|---|---|---|---|---|---|
| | | พฤติกรรมการใช้ระบบ | คะแนน | ข้อมูลทั่วไป | ทำวิศวกรรมคุณลักษณะ | จัดการความไม่สมดุลของข้อมูล | คัดเลือกคุณลักษณะ | แบบจำลอง | ประสิทธิภาพของแบบจำลองที่ดีที่สุด |
| 10 | Almasri, Alkhawaldeh, และ Çelebi (2020) | | | ✓ | - | - | ลำดับที่ 1 **Wrapper Method** | **K-means Clustering ร่วมกับ** Multilayer Perceptron, J48 Decision Trees, Naïve Bayes และ **Ensemble Meta-based Tree** | ความแม่นยำการทำนายระดับผลการเรียนเฉลี่ยสะสม 96.96% |
| งานวิจัยที่สนใจทำ | | ✓ | ✓ | ✓ | ลำดับที่ 1 **ชุดข้อมูลที่ 1 คุณลักษณะบันทึกการใช้งาน คะแนนทั้งหมด** ชุดข้อมูลที่ 2 คุณลักษณะบันทึกการใช้งาน และคะแนนกิจกรรม ชุดข้อมูลที่ 3 คุณลักษณะบันทึกการใช้งาน ชุดข้อมูลที่ 4 คุณลักษณะคะแนนกิจกรรม และคะแนนกลางภาค | ลำดับที่ 2 **Synthetic Minority Oversampling Technique: SMOTE** | ลำดับที่ 3 **Wrapper Method** | Logistic Regression, Naïve Bayes, Support Vector Machine, K-Nearest Neighbor, Decision Tree, Random Forest และ **XGBoost** | ผลที่คาดหวังจะได้รับแบบจำลองที่ดีที่สุด คือ XGBoost ทำนายระดับผลการเรียน โดยใช้ชุดข้อมูลที่ 1 คุณลักษณะบันทึกการใช้งาน คะแนนกิจกรรม และคะแนนกลางภาค |

# บทที่ 3
# วิธีการดำเนินงานวิจัย

ผู้วิจัยได้วางแผนการดำเนินงานวิจัย และกำหนดรายละเอียดการดำเนินงานวิจัยดังตาราง 2 โดยมีหัวข้อขั้นตอนการดำเนินงานวิจัย ดังนี้

1. การทำความเข้าใจในธุรกิจ
2. การทำความเข้าใจข้อมูล
3. การเก็บรวบรวมข้อมูลและการเตรียมข้อมูล
4. การสำรวจข้อมูล
5. การทำวิศวกรรมคุณลักษณะ
6. การจัดการความไม่สมดุลของข้อมูล
7. การสร้างแบบจำลอง
8. การคัดเลือกคุณลักษณะ
9. การวัดประสิทธิภาพของแบบจำลอง

**แผนการดำเนินงานวิจัย**

ตาราง 2 แสดงแผนการดำเนินงานวิจัย

| กิจกรรม | เดือน | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| 1. การศึกษาทฤษฎี และงานวิจัยที่เกี่ยวข้อง | ✓ | ✓ | | | | | | | | | | |
| 2. การกำหนดหัวข้องานวิจัย และกรอบแนวคิดในการวิจัย | | | ✓ | | | | | | | | | |
| 3. การวางแผนการดำเนินงานวิจัย | | | ✓ | | | | | | | | | |
| 4. การทำความเข้าใจในธุรกิจ<br>- กิจกรรมการเรียนการสอนที่กำหนดในระบบ Moodle รายวิชา SWU 141 ชีวิตในโลกดิจิทัล ภาคการศึกษา 2/2563<br>- วิธีการใช้งานในระบบ Moodle บนเว็บไซต์<br>- เกณฑ์การประเมินผลการเรียนรู้ | | | | ✓ | ✓ | | | | | | | |

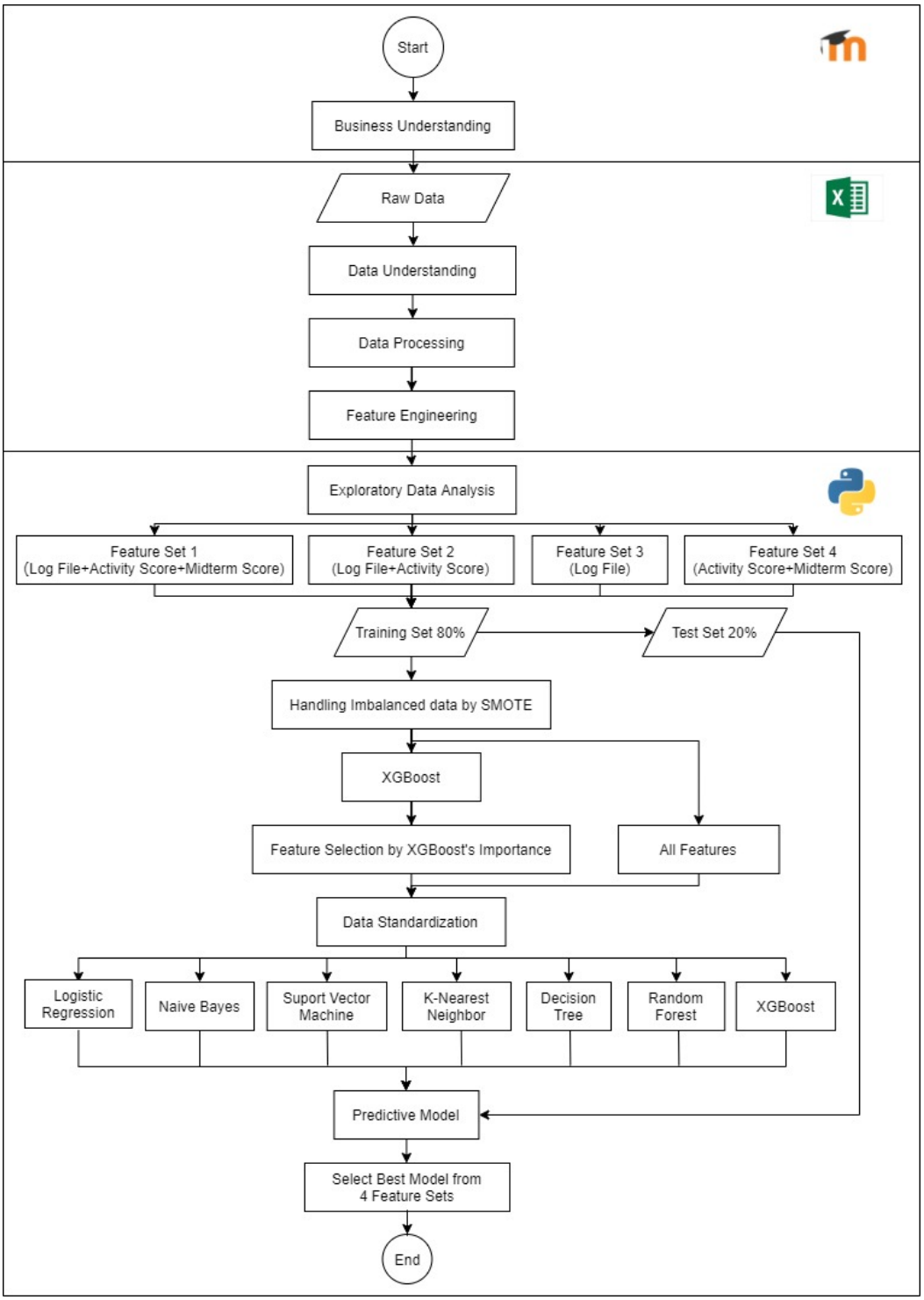
ตาราง 2 (ต่อ)

| กิจกรรม | เดือน | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| 6. การเก็บรวมรวมข้อมูลและการเตรียมข้อมูล<br>- เลือกใช้ข้อมูลของนิสิตตามที่กำหนดกลุ่มตัวอย่าง ได้แก่ ข้อมูลทั่วไป ข้อมูลบันทึกการใช้งานระบบ คะแนน และผลการเรียน<br>- วิเคราะห์ข้อมูลบันทึกการใช้งานระบบของนิสิตเพื่อสร้างคุณลักษณะในการสรุปสถิติการใช้งานระบบ<br>- รวบรวมข้อมูลทั้งหมดของนิสิตเข้าด้วยกัน<br>- ทำความสะอาดข้อมูล<br>- ทำชุดข้อมูลที่ใช้คุณลักษณะที่แตกต่างกัน 4 ชุด ที่แบ่งตามกรณี | | | | | | ✓ | ✓ | | | | |
| 7. การสำรวจข้อมูล<br>- ตัวแปรของผลลัพธ์การทำนาย<br>- ตัวแปรของคุณลักษณะ | | | | | | | | ✓ | | | | |
| 8. การสร้างแบบจำลอง<br>- สร้างแบบจำลองการจำแนกประเภทจำนวน 7 แบบจำลอง<br>- ใช้เทคนิคการจัดการข้อมูล ได้แก่ การทำวิศวกรรมคุณลักษณะ การจัดการความไม่สมดุลของข้อมูล และการคัดเลือกคุณลักษณะ | | | | | | | | ✓ | ✓ | | | |
| 9. การวัดประสิทธิภาพของแบบจำลอง<br>- เปรียบเทียบประสิทธิภาพของแบบจำลองที่ใช้คุณลักษณะที่แตกต่างกัน และเทคนิคการจัดข้อมูลที่แตกต่างกัน<br>- เลือกแบบจำลองที่ให้ประสิทธิภาพดีที่สุด | | | | | | | | | ✓ | | | |
| 10. สรุปผลการวิจัย | | | | | | | | | | ✓ | ✓ | ✓ |

**ภาพรวมของกระบวนการทำงาน**

งานวิจัยนี้จะใช้เทคนิคการเรียนรู้ของเรื่องในการวิเคราะห์พฤติกรรมการใช้งานระบบของนิสิต เพื่อทำนายผลการเรียนของนิสิต รายวิชา SWU 141 ชีวิตในโลกดิจิทัล ภาคการศึกษา 2 ปีการศึกษา 2563 จำนวน 405 คน จากสำนักนวัตกรรมการเรียนรู้ มหาวิทยาลัยศรีนครินทรวิโรฒ สำนักฯ มีการนำระบบการจัดการเรียนรู้ออนไลน์มาใช้ในการจัดการเรียนการสอน ซึ่งใช้ระบบ Moodle เป็นแพลตฟอร์มการเรียนรู้ออนไลน์ ในที่นี้จะอธิบายถึงภาพรวมของกระบวนการทำงาน โดยเริ่มจากการทำความเข้าใจในธุรกิจ ได้แก่ 1) กิจกรรมต่าง ๆ ที่อาจารย์มอบหมายให้นิสิตทำในระบบ Moodle และการกำหนดระยะเวลาของกิจกรรม 2) วิธีการใช้งานในระบบ Moodle บนเว็บไซต์ 3) เกณฑ์การประเมินผลการเรียนรู้ หลังจากนิสิตใช้งานระบบแล้ว ระบบจะจัดเก็บบันทึกการใช้งานระบบตลอดเวลาในรูปไฟล์ Excel จากนั้นดึงไฟล์ออกจากระบบไปทำความเข้าใจพฤติกรรมการใช้งานระบบ โดยสุ่มตัวอย่างการใช้งานของนิสิตคนหนึ่ง แล้วเข้าสู่กระบวนการเตรียมข้อมูล โดยผ่านการทำความสะอาดข้อมูล และการแปลงข้อมูลให้เหมาะสม จากนั้นสร้างคุณลักษณะจากข้อมูลบันทึกการใช้งานระบบในการวิเคราะห์พฤติกรรมการใช้งานระบบ โดยใช้เทคนิคการทำวิศวกรรมคุณลักษณะ เพื่อสรุปสถิติการใช้งานของนิสิตของแต่ละกิจกรรมเป็นรายบุคคลจากการใช้เครื่องมือต่าง ๆ ในโปรแกรม Excel Workbook ต่อไปนำคุณลักษณะสถิติการใช้งานรวมกับคะแนน ข้อมูลทั่วไป และเกรดเป็นชุดข้อมูล ต่อมาเขียนภาษา Python ในการนำเข้าชุดข้อมูลนี้ แล้วเข้าสู่กระบวนการสำรวจข้อมูล จากนั้นแบ่งชุดข้อมูลออกเป็น 4 กรณี ตามพฤติกรรมการใช้งานระบบ และเกณฑ์การประเมินผลการเรียนรู้ ได้แก่ ชุดที่ 1 ข้อมูลบันทึกการใช้งาน คะแนนกิจกรรมในห้องเรียนออนไลน์ และคะแนนกลางภาค, ชุดที่ 2 ข้อมูลบันทึกการใช้งาน และคะแนนกิจกรรมในห้องเรียนออนไลน์, ชุดที่ 3 ข้อมูลบันทึกการใช้งาน และชุดที่ 4 คะแนนกิจกรรมในห้องเรียนออนไลน์ และคะแนนกลางภาค ต่อมาแยกชุดข้อมูลเป็นชุดข้อมูลการเรียนรู้จำนวน 80% และชุดข้อมูลการทดลองจำนวน 20% แล้วนำชุดข้อมูลการเรียนรู้ไปปรับความไม่สมดุลของข้อมูลด้วยวิธี SMOTE ของแต่ละชุดข้อมูล แล้วปรับขนาดของข้อมูลแบบ Standardization แล้วเข้าสู่แบบจำลอง XGBoost เพื่อใช้แบบจำลอง XGBoost ในการคัดเลือกกลุ่มคุณลักษณะที่มีความสำคัญต่อผลการเรียน จากนั้นนำชุดคุณลักษณะนี้ไปทำนายข้อมูลด้วยแบบจำลอง 7 แบบ ได้แก่ Logistic Regression, Naïve Bayes, Support Vector Machine, K-Nearest Neighbor, Decision Tree, Random Forest และ Extreme Gradient Boosting (XGBoost) แล้วเลือกแบบจำลองที่ดีที่สุดไปปรับค่าพารามิเตอร์ จากนั้นนำชุดข้อมูลการทดสอบ 4 ชุด เข้าสู่แบบจำลองนี้ และเปรียบเทียบประสิทธิภาพของแบบจำลองของชุดข้อมูลที่ใช้

คุณลักษณะทั้งหมด และชุดข้อมูลที่คัดเลือกคุณลักษณะ แล้วเลือกแบบจำลองที่ให้ประสิทธิภาพที่ดีที่สุด ดังภาพประกอบ 22




ภาพประกอบ 22 แสดงแผนผังกระบวนการทำงาน

**การทำความเข้าใจในธุรกิจ**

ก่อนเปิดภาคการศึกษา 2 ปีการศึกษา 2563 ผู้ดูแลระบบมีการติดตั้งระบบ Moodle มาใช้การสนับสนุนการเรียนการสอนของสำนักนวัตกรรมการเรียนรู้ มศว จากนั้นเปิดรายวิชา SWU 141 ชีวิตในโลกดิจิทัล ร่วมกับการสร้างสื่อการเรียนรู้ แบบทดสอบ งานมอบหมายล่วงหน้า โดยกำหนดเงื่อนไขของกิจกรรมการเรียนการสอนที่แตกต่างกันในการตั้งค่าระบบ ซึ่งมีเนื้อหาบทเรียนรายวิชา SWU 141 ชีวิตในโลกดิจิทัล จำนวน 11 เรื่อง ดังตาราง 3

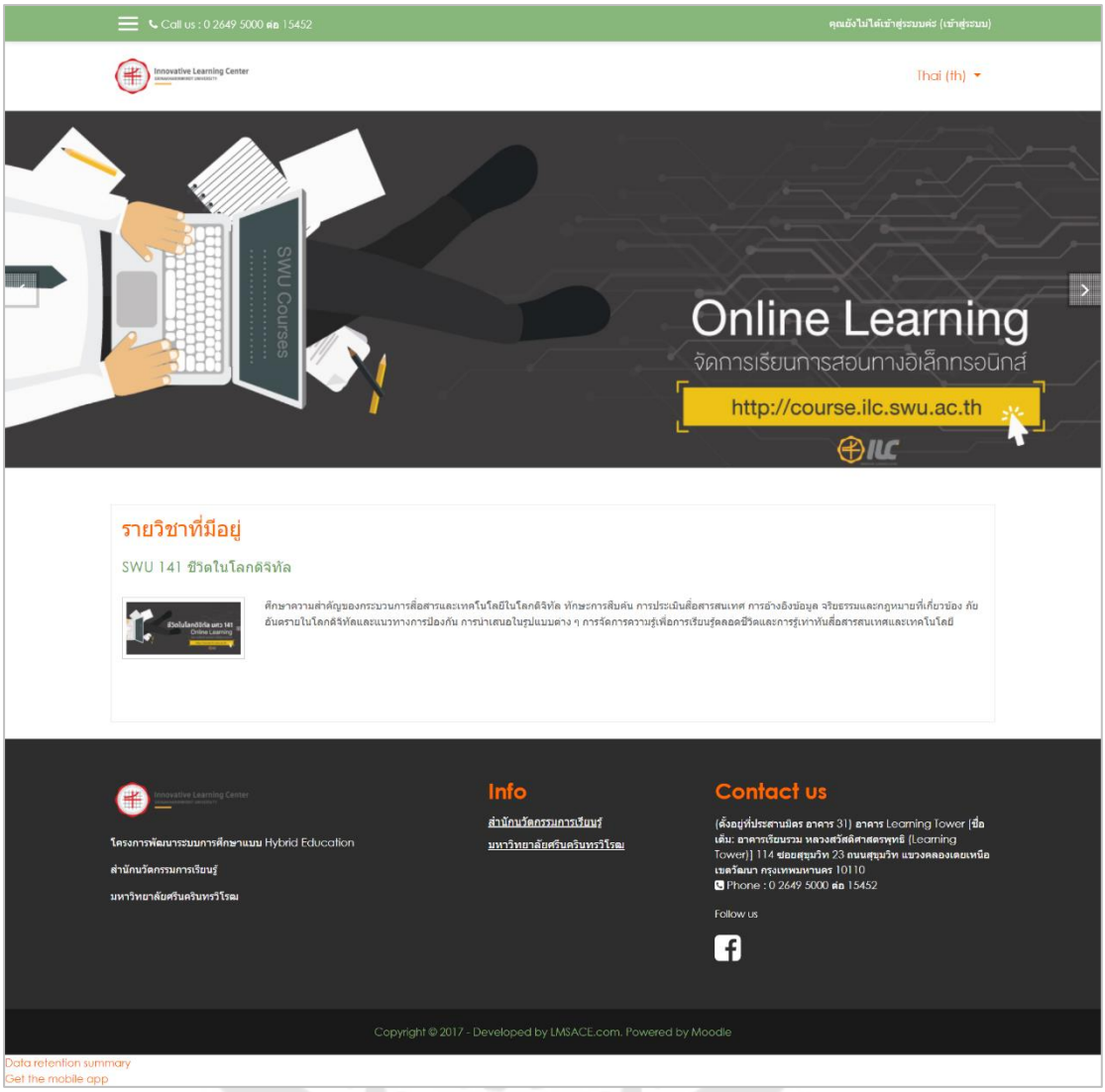
ตาราง 3 แสดงรายละเอียดเงื่อนไขของกิจกรรมการเรียนการสอนในระบบ Moodle

| **เรื่องของสื่อการเรียนรู้ (สามารถชมสื่อการเรียนรู้ล่วงหน้าทุกเรื่องตลอดเวลา)** | **แบบทดสอบ (ทำได้เพียง 3 ครั้งเท่านั้น)** | | | **งานมอบหมาย** | |
|---|---|---|---|---|---|
| | **ช่วงเวลาที่สามารถเข้าทำแบบทดสอบได้** | **ระยะเวลาการทำแบบทดสอบ (นาที)** | **คะแนนเต็ม** | **ช่วงเวลาการส่งงานมอบหมาย** | **คะแนนเต็ม** |
| **ครึ่งแรกของภาคการศึกษา (ตั้งแต่วันแรกของการเรียนรายวิชานี้ วันที่ 7 มกราคม 2564 ถึงวันสุดท้ายของสัปดาห์การสอบกลางภาค วันที่ 5 มีนาคม 2564)** | | | | | |
| 1. พลเมืองดิจิทัลที่รู้เท่าทันสื่อสารสนเทศและเทคโนโลยี | 4 ม.ค. -28 ก.พ. 64 | 10 | 10 | | |
| 2. แหล่งสารสนเทศและการสืบค้น | 4 ม.ค. - 28 ก.พ. 64 | 10 | 10 | | |
| 3. การประเมินสารสนเทศ | 4 ม.ค. - 28 ก.พ. 64 | 10 | 10 | | |
| 4. การวิเคราะห์สังเคราะห์สารสนเทศ | | | | | |
| 5. จริยธรรมในการใช้เทคโนโลยีสารสนเทศและการสื่อสาร | 4 ม.ค. - 28 ก.พ. 64 | 20 | 10 | | |
| 6. การลอกเลียนวรรณกรรมและแนวทางการหลีกเลี่ยงการลอกเลียนวรรณกรรมด้วยการเขียนอ้างอิงและบรรณานุกรม | 4 ม.ค. - 28 ก.พ. 64 | 30 | 15 | | |
| 7. เทคโนโลยีสารสนเทศและการสื่อสารที่น่าสนใจในปัจจุบันและแนวโน้มในอนาคต | | | | 8 - 20 ก.พ. 64 | 100 |

ตาราง 3 (ต่อ)

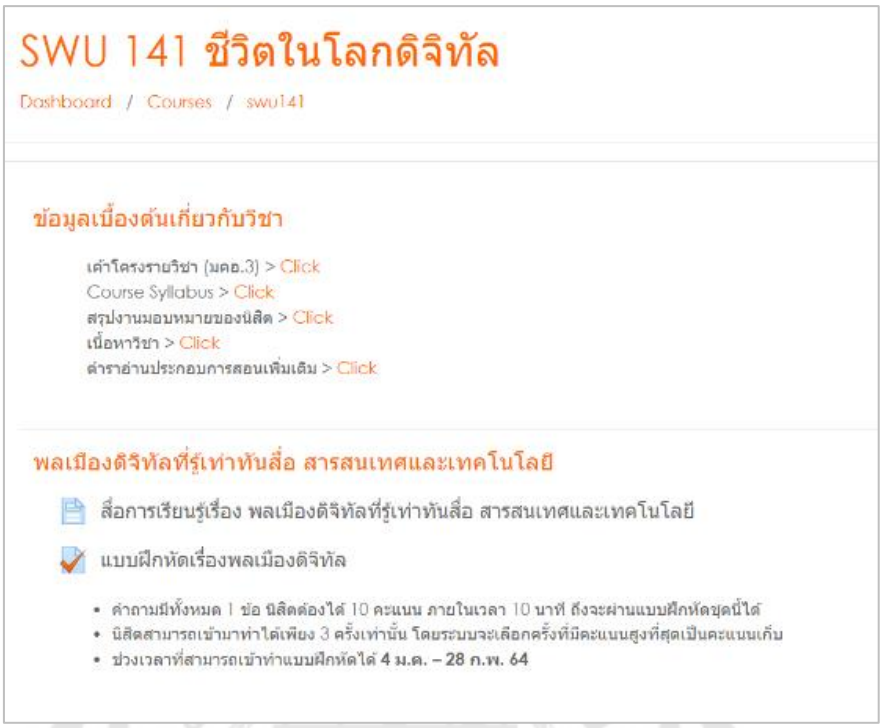
| เรื่องของสื่อการเรียนรู้ (สามารถชมสื่อการเรียนรู้ล่วงหน้าทุกเรื่องตลอดเวลา) | แบบทดสอบ (ทำได้เพียง 3 ครั้งเท่านั้น) | | | งานมอบหมาย | |
|---|---|---|---|---|---|
| | ช่วงเวลาที่สามารถเข้าทำแบบทดสอบได้ | ระยะเวลาการทำแบบทดสอบ (นาที) | คะแนนเต็ม | ช่วงเวลาการส่งงานมอบหมาย | คะแนนเต็ม |
| **ครึ่งหลังของภาคการศึกษา (ตั้งแต่หลังวันสุดท้ายของสัปดาห์การสอบกลางภาค วันที่ 5 มีนาคม 2564 ถึงวันสุดท้ายการเรียน วันที่ 30 เมษายน 2564)** | | | | | |
| 8. การประยุกต์ใช้ทักษะการรู้สารสนเทศในการเขียนบทความวิชาการ | | | | 8 - 22 มี.ค. 64 | 100 |
| 9. ภัยอันตรายในโลกยุคดิจิทัลและแนวทางการป้องกัน | 4 ม.ค. - 28 ก.พ. 64 | 10 | 10 | | |
| 10. การนำเสนอสารสนเทศในรูปแบบต่าง ๆ | | | | | |
| 11. เทคนิคการจัดการข้อมูลอย่างมีระบบในรูปแบบรายงานด้วยโปรแกรม MS Word และ MS Excel | | | | 8 มี.ค. - 2 เม.ย. 64 | 10 |

ในระหว่างเปิดภาคการศึกษาอาจารย์ได้มอบหมายงานให้นิสิตดำเนินกิจกรรมในระบบ Moodle ได้แก่ การชมสื่อการเรียนรู้ล่วงหน้า การทำแบบทดสอบภายในระยะเวลาที่กำหนด โดยสามารถทำแบบทดสอบซ้ำได้ 3 ครั้ง และการส่งงานมอบหมายภายในระยะเวลาที่กำหนด ระบบจะจัดเก็บข้อมูลการบันทึกกิจกรรมของนิสิตเหล่านี้ตลอดเวลา และจัดเก็บคะแนนแบบทดสอบมากที่สุดเป็นคะแนนเก็บจากการทำแบบทดสอบซ้ำ 3 ครั้ง ดังภาพประกอบ 23 – 27



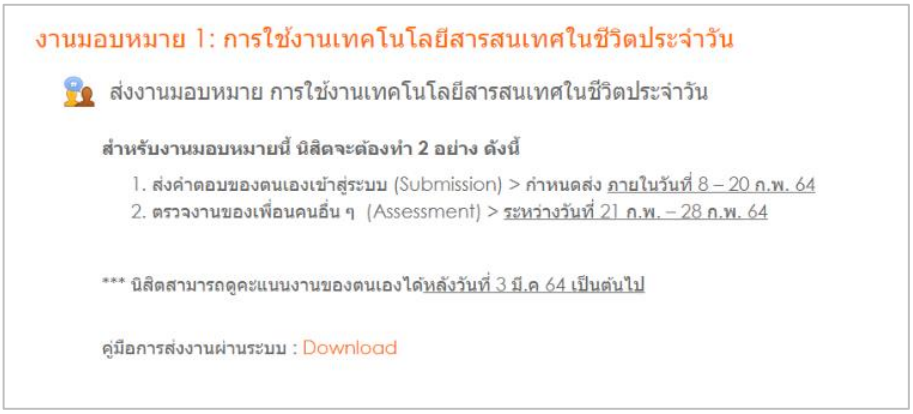

ภาพประกอบ 23 แสดงหน้าแรกการเข้าสู่ระบบ Moodle รายวิชา SWU 141 ชีวิตในโลกดิจิทัล บนเว็บไซต์ https://course.ilc.swu.ac.th

ที่มา: (InnovativeLearningCenterSrinakharinwirotUniversity, 2017)



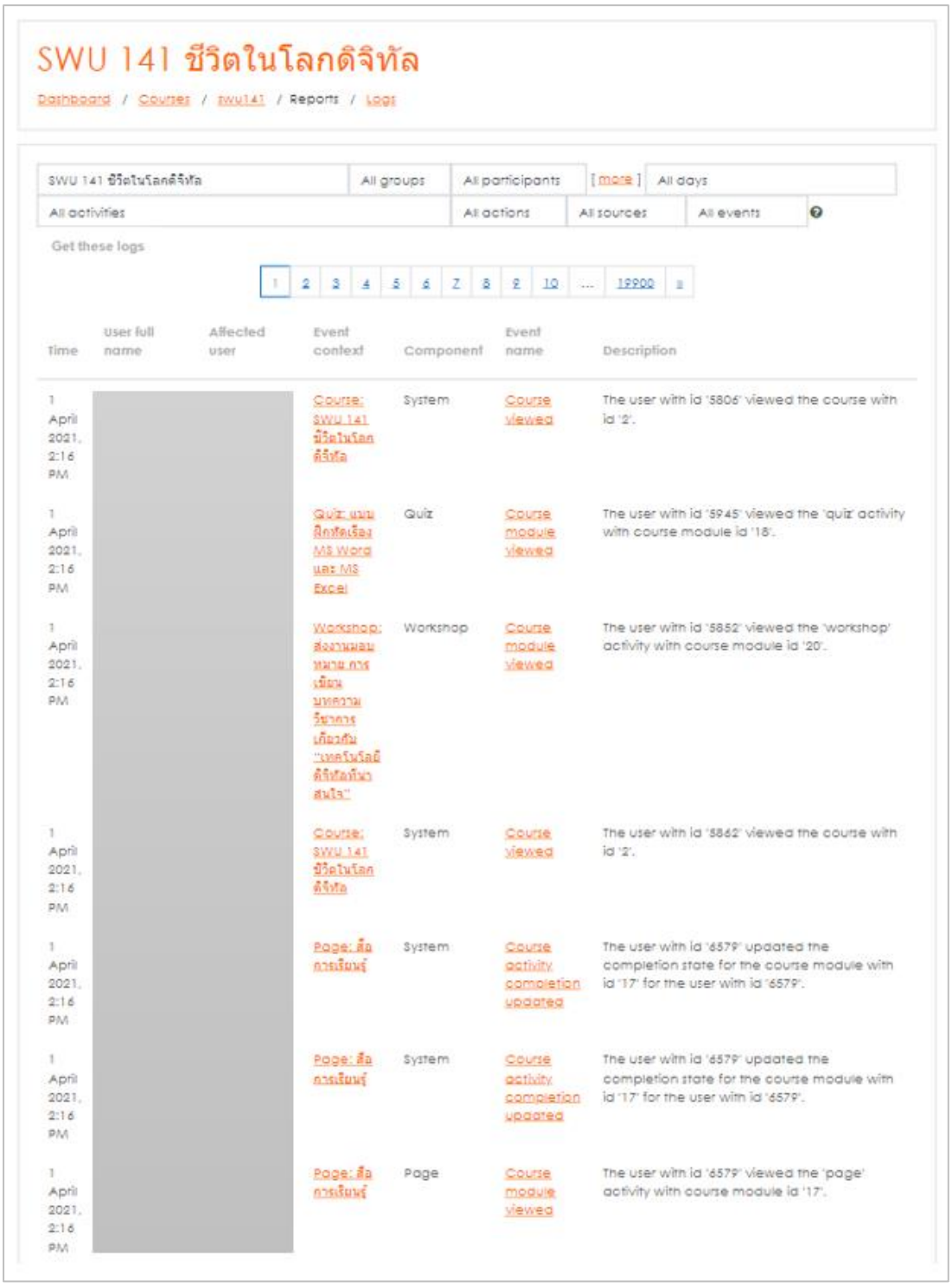

ภาพประกอบ 24 แสดงการเข้าถึงสื่อการเรียนรู้ และแบบทดสอบ จากระบบ Moodle

ที่มา: (InnovativeLearningCenterSrinakharinwirotUniversity, 2017)




ภาพประกอบ 25 แสดงการเข้าถึงงานมอบหมายจากระบบ Moodle

ที่มา: (InnovativeLearningCenterSrinakharinwirotUniversity, 2017)



# SWU 141 ชีวิตในโลกดิจิทัล

Dashboard / Courses / swu141 / Reports / Logs

SWU 141 ชีวิตในโลกดิจิทัล | All groups | All participants | [ more ] | All days

All activities | All actions | All sources | All events

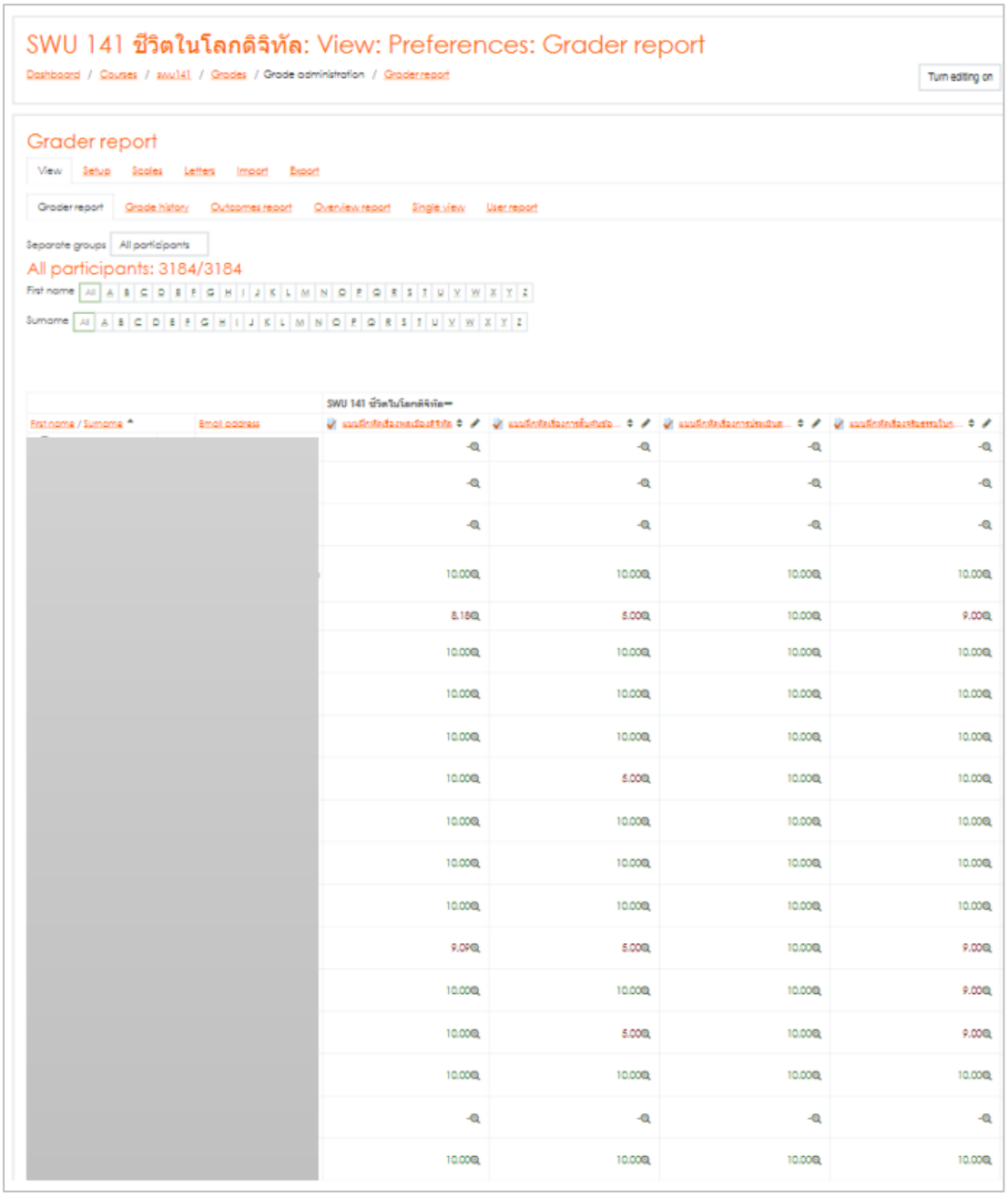
Get these logs

1 2 3 4 5 6 7 8 9 10 ... 19900 »

| Time | User full name | Affected user | Event context | Component | Event name | Description |
|---|---|---|---|---|---|---|
| 1 April 2021, 2:16 PM | | | Course: SWU 141 ชีวิตในโลกดิจิทัล | System | Course viewed | The user with id '5806' viewed the course with id '2'. |
| 1 April 2021, 2:16 PM | | | Quiz: แบบฝึกหัดเรื่อง MS Word และ MS Excel | Quiz | Course module viewed | The user with id '5945' viewed the 'quiz' activity with course module id '18'. |
| 1 April 2021, 2:16 PM | | | Workshop: ส่งงานมอบหมาย การเขียนบทความวิชาการเกี่ยวกับ "เทคโนโลยีดิจิทัลที่น่าสนใจ" | Workshop | Course module viewed | The user with id '5852' viewed the 'workshop' activity with course module id '20'. |
| 1 April 2021, 2:16 PM | | | Course: SWU 141 ชีวิตในโลกดิจิทัล | System | Course viewed | The user with id '5862' viewed the course with id '2'. |
| 1 April 2021, 2:16 PM | | | Page: สื่อการเรียนรู้ | System | Course activity completion updated | The user with id '6579' updated the completion state for the course module with id '17' for the user with id '6579'. |
| 1 April 2021, 2:16 PM | | | Page: สื่อการเรียนรู้ | System | Course activity completion updated | The user with id '6579' updated the completion state for the course module with id '17' for the user with id '6579'. |
| 1 April 2021, 2:16 PM | | | Page: สื่อการเรียนรู้ | Page | Course module viewed | The user with id '6579' viewed the 'page' activity with course module id '17'. |

ภาพประกอบ 26 แสดงรายการบันทึกกิจกรรมการใช้งานของนิสิตจากระบบ Moodle

ที่มา: (InnovativeLearningCenterSrinakharinwirotUniversity, 2017)




ภาพประกอบ 27 แสดงคะแนนแบบทดสอบของนิสิตจากระบบ Moodle

ที่มา: (InnovativeLearningCenterSrinakharinwirotUniversity, 2017)

นอกจากนี้สำนักนวัตกรรมการเรียนรู้ มศว ได้กำหนดเกณฑ์การพิจารณาตัดเกรด ดังตาราง 4 และเกณฑ์การประเมินผลการเรียนรู้ ดังตาราง 5

ตาราง 4 แสดงเกณฑ์การพิจารณาตัดเกรด

| การพิจารณาตัดเกรด | สัดส่วนการประเมิน |
|---|---|
| A | มากกว่าหรือเท่ากับ 90 % |
| B+ | มากกว่าหรือเท่ากับ 85% และน้อยกว่า 90% |
| B | มากกว่าหรือเท่ากับ 80% และน้อยกว่า 95% |
| C+ | มากกว่าหรือเท่ากับ 75% และน้อยกว่า 80% |
| C | มากกว่าหรือเท่ากับ 70% และน้อยกว่า 75% |
| D+ | มากกว่าหรือเท่ากับ 65% และน้อยกว่า 70% |
| D | มากกว่าหรือเท่ากับ 60% และน้อยกว่า 65% |
| E | น้อยกว่า 60% |

ตาราง 5 แสดงเกณฑ์การประเมินผลการเรียนรู้

| คะแนนการประเมินผลการเรียนรู้ | คะแนนภายในระบบ | คะแนนภายนอกระบบ | สัดส่วนการประเมิน |
|---|---|---|---|
| **1. ครึ่งแรกของภาคการศึกษา** | | | |
| **1.1 ด้านคะแนนกิจกรรมในห้องเรียน** | | | |
| - คะแนนกิจกรรมในห้องเรียนออนไลน์ | | ✓ | 30% |
| - คะแนนการเข้าเรียน | ✓ | ✓ | |
| - คะแนนแบบทดสอบ | | | |
| **1.2 ด้านคะแนนกลางภาค** | | | |
| - คะแนนสอบกลางภาค | | ✓ | 30% |
| - คะแนนงานมอบหมาย เรื่องเทคโนโลยีสารสนเทศและการสื่อสารที่น่าสนใจในปัจจุบันและแนวโน้มในอนาคต | ✓ | | |
| **รวม** | | | **60%** |

ตาราง 5 (ต่อ)

| คะแนนการประเมินผลการเรียนรู้ | คะแนนภายในระบบ | คะแนนภายนอกระบบ | สัดส่วนการประเมิน |
|---|---|---|---|
| **2. ครึ่งหลังของภาคการศึกษา** | | | |
| **2.1 ด้านคะแนนโครงงาน** | | | |
| - คะแนนงานมอบหมาย เรื่อง การประยุกต์ใช้ทักษะการรู้สารสนเทศในการเขียนบทความวิชาการ (งานเดี่ยว) | ✓ | | 30% |
| - คะแนนงานมอบหมาย เรื่อง การประยุกต์ใช้ทักษะการรู้สารสนเทศในการเขียนบทความวิชาการ (งานกลุ่ม) | | ✓ | |
| - คะแนนการนำเสนอ เรื่อง การประยุกต์ใช้ทักษะการรู้สารสนเทศในการเขียนบทความวิชาการ (งานกลุ่ม) | | ✓ | |
| **2.2 ด้านคะแนนพฤติกรรมการมีส่วนร่วม** | | | |
| - คะแนนจิตพิสัย | | ✓ | 10% |
| - คะแนนการประเมินโดยเพื่อน (Peer Assessment) | | ✓ | |
| **รวม** | | | **40%** |
| **รวมทั้งหมด** | | | **100%** |

งานวิจัยครั้งนี้จะศึกษาข้อมูลผลการเรียนรู้ของนิสิตในช่วงครึ่งแรกของภาคการศึกษา ได้แก่ ข้อมูลบันทึกการใช้งาน คะแนนกิจกรรมในห้องเรียน และคะแนนกลางภาค เพื่อเปรียบเทียบประสิทธิภาพของแบบจำลองการทำนายผลการเรียนของนิสิต จึงสร้างชุดคุณลักษณะที่ใช้การวิจัย จำนวน 4 ชุด โดยแบ่งกรณีตามประเภทของผลการเรียนรู้สำหรับนิสิต ดังตาราง 6

ตาราง 6 การจำแนกคุณลักษณะที่ใช้ในการวิจัย

| ชุดของคุณลักษณะ | คุณลักษณะข้อมูลบันทึกการใช้งาน ระหว่างวันที่ 7 มกราคม ถึง 5 มีนาคม 2564 | คุณลักษณะคะแนนกิจกรรมในห้องเรียน (30%) | คุณลักษณะคะแนนกลางภาค (30%) |
|---|---|---|---|
| กรณีที่ 1 | ✓ | ✓ | ✓ |
| กรณีที่ 2 | ✓ | ✓ | |
| กรณีที่ 3 | ✓ | | |
| กรณีที่ 4 | | ✓ | ✓ |

**การทำความเข้าใจข้อมูล**

ในระหว่างการใช้งานระบบของผู้ใช้งาน ระบบจะบันทึกข้อมูลการใช้งานแบบเฉพาะเจาะจงของผู้ใช้งานทั้งหมด รวมทั้งผู้ดูแลระบบ อาจารย์ และนิสิต ตั้งแต่การเข้าสู่ระบบ การเข้าถึงกิจกรรมในบทเรียน ตลอดจนการจัดเก็บคะแนนกิจกรรม รวมถึงวันที่และเวลาการใช้งานด้วย โดยข้อมูลเหล่านี้จะถูกจัดเก็บเป็นไฟล์ Excel ในระบบ ดังภาพประกอบ 28

| DateTime | IdStudent | User | AffectedUser | EventContext | Component | EventName | Description | Origin | IPAddress |
|---|---|---|---|---|---|---|---|---|---|
| 1/02/21, 01:20 | | | | Course: SWU 141 ชีวิตในโลกดิจิทัล | System | Course viewed | The user with id '6243' viewed the course with id '2'. | web | 184.22.205.95 |
| 1/02/21, 01:21 | | | | Page: สื่อการเรียนรู้เรื่อง การประเมินสารสนเทศ | System | Course activity completion updated | The user with id '6243' updated the completion state for the course module wit | web | 184.22.205.95 |
| 1/02/21, 01:21 | | | | Page: สื่อการเรียนรู้เรื่อง การประเมินสารสนเทศ | System | Course activity completion updated | The user with id '6243' updated the completion state for the course module wit | web | 184.22.205.95 |
| 1/02/21, 01:21 | | | | Page: สื่อการเรียนรู้เรื่อง การประเมินสารสนเทศ | Page | Course module viewed | The user with id '6243' viewed the 'page' activity with course module id '6'. | web | 184.22.205.95 |
| 1/02/21, 11:46 | | | | Quiz: แบบฝึกหัดเรื่องพลเมืองดิจิทัล | Quiz | Quiz attempt reviewed | The user with id '6216' has had their attempt with id '34713' reviewed by the us | web | 171.97.73.170 |
| 1/02/21, 11:46 | | | | Quiz: แบบฝึกหัดเรื่องพลเมืองดิจิทัล | Quiz | Course module viewed | The user with id '6216' viewed the 'quiz' activity with course module id '3'. | web | 171.97.73.170 |
| 1/02/21, 11:46 | | | | Course: SWU 141 ชีวิตในโลกดิจิทัล | System | Course viewed | The user with id '6216' viewed the course with id '2'. | web | 171.97.73.170 |
| 1/02/21, 11:52 | | | | Quiz: แบบฝึกหัดเรื่องพลเมืองดิจิทัล | Quiz | Quiz attempt viewed | The user with id '6216' has viewed the attempt with id '43597' belonging to the | web | 171.97.73.170 |
| 1/02/21, 11:52 | | | | Quiz: แบบฝึกหัดเรื่องพลเมืองดิจิทัล | Quiz | Quiz attempt started | The user with id '6216' has started the attempt with id '43597' for the quiz with | web | 171.97.73.170 |
| 1/02/21, 11:54 | | | | Quiz: แบบฝึกหัดเรื่องพลเมืองดิจิทัล | Quiz | Course module viewed | The user with id '6216' viewed the 'quiz' activity with course module id '3'. | web | 171.97.73.170 |
| 1/02/21, 11:54 | | | | Quiz: แบบฝึกหัดเรื่องพลเมืองดิจิทัล | Quiz | Quiz attempt reviewed | The user with id '6216' has had their attempt with id '43597' reviewed by the us | web | 171.97.73.170 |
| 1/02/21, 11:54 | | | | Quiz: แบบฝึกหัดเรื่องพลเมืองดิจิทัล | Quiz | Quiz attempt submitted | The user with id '6216' has submitted the attempt with id '43597' for the quiz v | web | 171.97.73.170 |
| 1/02/21, 11:54 | | | | Course: SWU 141 ชีวิตในโลกดิจิทัล | System | User graded | The user with id '6216' updated the grade with id '49654' for the user with id '6 | web | 171.97.73.170 |
| 1/02/21, 11:54 | | | | Quiz: แบบฝึกหัดเรื่องพลเมืองดิจิทัล | System | Course activity completion updated | The user with id '6216' updated the completion state for the course module wit | web | 171.97.73.170 |
| 1/02/21, 11:54 | | | | Quiz: แบบฝึกหัดเรื่องพลเมืองดิจิทัล | Quiz | Quiz attempt summary viewed | The user with id '6216' has viewed the summary for the attempt with id '43597 | web | 171.97.73.170 |
| 1/02/21, 11:55 | | | | Quiz: แบบฝึกหัดเรื่องการสืบค้นข้อมูลด้วย Google | Quiz | Quiz attempt viewed | The user with id '6216' has viewed the attempt with id '43598' belonging to the | web | 171.97.73.170 |
| 1/02/21, 11:55 | | | | Quiz: แบบฝึกหัดเรื่องการสืบค้นข้อมูลด้วย Google | Quiz | Quiz attempt started | The user with id '6216' has started the attempt with id '43598' for the quiz with | web | 171.97.73.170 |
| 1/02/21, 11:55 | | | | Quiz: แบบฝึกหัดเรื่องการสืบค้นข้อมูลด้วย Google | System | Course activity completion updated | The user with id '6216' updated the completion state for the course module wit | web | 171.97.73.170 |
| 1/02/21, 11:55 | | | | Quiz: แบบฝึกหัดเรื่องการสืบค้นข้อมูลด้วย Google | Quiz | Course module viewed | The user with id '6216' viewed the 'quiz' activity with course module id '5'. | web | 171.97.73.170 |
| 1/02/21, 11:55 | | | | Course: SWU 141 ชีวิตในโลกดิจิทัล | System | Course viewed | The user with id '6216' viewed the course with id '2'. | web | 171.97.73.170 |
| 1/02/21, 11:58 | | | | Course: SWU 141 ชีวิตในโลกดิจิทัล | System | Course viewed | The user with id '6216' viewed the course with id '2'. | web | 171.97.73.170 |
| 1/02/21, 11:58 | | | | Quiz: แบบฝึกหัดเรื่องการสืบค้นข้อมูลด้วย Google | Quiz | Course module viewed | The user with id '6216' viewed the 'quiz' activity with course module id '5'. | web | 171.97.73.170 |
| 1/02/21, 11:58 | | | | Quiz: แบบฝึกหัดเรื่องการสืบค้นข้อมูลด้วย Google | Quiz | Quiz attempt reviewed | The user with id '6216' has had their attempt with id '43598' reviewed by the us | web | 171.97.73.170 |
| 1/02/21, 11:58 | | | | Quiz: แบบฝึกหัดเรื่องการสืบค้นข้อมูลด้วย Google | Quiz | Quiz attempt submitted | The user with id '6216' has submitted the attempt with id '43598' for the quiz v | web | 171.97.73.170 |
| 1/02/21, 11:58 | | | | Course: SWU 141 ชีวิตในโลกดิจิทัล | System | User graded | The user with id '6216' updated the grade with id '49656' for the user with id '6 | web | 171.97.73.170 |
| 1/02/21, 11:58 | | | | Quiz: แบบฝึกหัดเรื่องการสืบค้นข้อมูลด้วย Google | System | Course activity completion updated | The user with id '6216' updated the completion state for the course module wit | web | 171.97.73.170 |
| 1/02/21, 11:58 | | | | Quiz: แบบฝึกหัดเรื่องการสืบค้นข้อมูลด้วย Google | Quiz | Quiz attempt summary viewed | The user with id '6216' has viewed the summary for the attempt with id '43598 | web | 171.97.73.170 |
| 1/02/21, 12:00 | | | | Page: สื่อการเรียนรู้เรื่อง การออกเลียนวรรณกรรม แ | System | Course activity completion updated | The user with id '6173' updated the completion state for the course module wit | web | 115.87.82.254 |
| 1/02/21, 12:00 | | | | Page: สื่อการเรียนรู้เรื่อง การออกเลียนวรรณกรรม แ | System | Course activity completion updated | The user with id '6173' updated the completion state for the course module wit | web | 115.87.82.254 |
| 1/02/21, 12:00 | | | | Page: สื่อการเรียนรู้เรื่อง การออกเลียนวรรณกรรม แ | Page | Course module viewed | The user with id '6173' viewed the 'page' activity with course module id '10'. | web | 115.87.82.254 |
| 1/02/21, 12:00 | | | | Course: SWU 141 ชีวิตในโลกดิจิทัล | System | Course viewed | The user with id '6173' viewed the course with id '2'. | web | 115.87.82.254 |
| 1/02/21, 12:01 | | | | Quiz: แบบฝึกหัดเรื่องจริยธรรมในการใช้เทคโนโลย | Quiz | Quiz attempt viewed | The user with id '6216' has viewed the attempt with id '43601' belonging to the | web | 171.97.73.170 |
| 1/02/21, 12:01 | | | | Quiz: แบบฝึกหัดเรื่องจริยธรรมในการใช้เทคโนโลย | Quiz | Quiz attempt started | The user with id '6216' has started the attempt with id '43601' for the quiz with | web | 171.97.73.170 |
| 1/02/21, 12:01 | | | | Quiz: แบบฝึกหัดเรื่องจริยธรรมในการใช้เทคโนโลย | System | Course activity completion updated | The user with id '6216' updated the completion state for the course module wit | web | 171.97.73.170 |
| 1/02/21, 12:01 | | | | Quiz: แบบฝึกหัดเรื่องจริยธรรมในการใช้เทคโนโลย | Quiz | Course module viewed | The user with id '6216' viewed the 'quiz' activity with course module id '9'. | web | 171.97.73.170 |
| 1/02/21, 12:04 | | | | Quiz: แบบฝึกหัดเรื่องจริยธรรมในการใช้เทคโนโลย | Quiz | Quiz attempt viewed | The user with id '6216' has viewed the attempt with id '43601' belonging to the | web | 171.97.73.170 |

ภาพประกอบ 28 แสดงไฟล์ข้อมูลบันทึกกิจกรรมการใช้งานระบบ Moodle รายวิชา SWU 141 ชีวิตในโลกดิจิทัล จากเว็บไซต์ https://course.ilc.swu.ac.th

เพื่อให้มีความเข้าใจข้อมูลการใช้งานระบบของนิสิตมากขึ้น จึงยกตัวอย่างสำหรับข้อมูลบันทึกการใช้งานระบบของนิสิตคนหนึ่ง ดังภาพประกอบ 29

| IdStudent | DateTime | EventContext | EventName |
|---|---|---|---|
| | **14/01/21, 13:14** | Course: SWU 141 **ชีวิตในโลกดิจิทัล** | Course viewed |
| | **14/01/21, 13:15** | Quiz1 | Course module viewed |
| | **14/01/21, 13:16** | Course: SWU 141 **ชีวิตในโลกดิจิทัล** | Course viewed |
| | **14/01/21, 13:17** | Course: SWU 141 **ชีวิตในโลกดิจิทัล** | Course viewed |
| | **14/01/21, 13:40** | Quiz1 | Course module viewed |
| | | | Quiz attempt started |
| | | | Quiz attempt viewed |
| | **14/01/21, 13:47** | Course: SWU 141 **ชีวิตในโลกดิจิทัล** | User graded |
| | | Quiz1 | Course module viewed |
| | | | Quiz attempt reviewed |
| | | | Quiz attempt submitted |
| | | | Quiz attempt summary viewed |
| | **14/01/21, 13:48** | Course: SWU 141 **ชีวิตในโลกดิจิทัล** | Course viewed |
| | | Page1 | Course module viewed |
| | **14/01/21, 13:49** | Course: SWU 141 **ชีวิตในโลกดิจิทัล** | Course viewed |
| | | Quiz1 | Course module viewed |
| | | | Quiz attempt started |
| | | | Quiz attempt viewed |
| | **14/01/21, 13:56** | Course: SWU 141 **ชีวิตในโลกดิจิทัล** | User graded |
| | | Quiz1 | Quiz attempt reviewed |
| | | | Quiz attempt submitted |
| | | | Quiz attempt summary viewed |

ภาพประกอบ 29 แสดงตัวอย่างบันทึกการใช้งานระบบของนิสิต

จากภาพประกอบ 29 สรุปว่า นิสิตได้ดำเนินกิจกรรมต่าง ๆ ในระบบ ดังนี้

1. เมื่อวันที่ 14 ม.ค. 64 เวลา 13.14 น. นิสิตได้คลิกเข้าสู่ระบบ รายวิชา SWU 141 ชีวิตในโลกดิจิทัล ในเว็บไซต์

2. คลิกลิงก์เพื่อเข้าหน้าเว็บไซต์การทำแบบทดสอบ เมื่อเวลา 13.15 น.

3. กลับไปสู่หน้าแรกเว็บไซต์ เมื่อเวลา 13.16 น.

4. คลิกทำแบบทดสอบครั้งที่ 1 เมื่อเวลา 13.40 น. แล้วทบทวนคำตอบแบบทดสอบและส่งแบบทดสอบ เมื่อเวลา 13.47 น. ต่อไปเข้าสู่หน้าเว็บไซต์แสดงผลคะแนนแบบทดสอบ โดยมีระยะเวลาการทำแบบทดสอบทั้งหมด 7 นาที

5. ต่อมากลับเข้าสู่หน้าแรกอีกครั้ง เพื่อชมสื่อการเรียนรู้ เมื่อเวลา 13.48 น.

6. ทำแบบทดสอบซ้ำในครั้งที่ 2 ตั้งแต่เวลา 13.49 น. จนถึง 13.56 น. เป็นระยะเวลา 7 นาที

**การเก็บรวบรวมข้อมูลและการเตรียมข้อมูล**

การศึกษาวิจัยนี้เลือกใช้ข้อมูลบันทึกการใช้งานระบบ Moodle ของนิสิตระดับปริญญาตรี ชั้นปีที่ 1 มหาวิทยาลัยศรีนครินทรวิโรฒ จำนวน 405 คน ที่ลงทะเบียนรายวิชา SWU 141 ชีวิตในโลกดิจิทัล ภาคการศึกษาที่ 2 ปีการศึกษา 2563 ตั้งแต่วันที่ 7 มกราคม ถึง 5 มีนาคม 2564 ในช่วงการเรียนของครึ่งภาคการศึกษาเป็นระยะเวลา 9 สัปดาห์ ข้อมูลบันทึกการใช้งานระบบเหล่านี้ที่ดึงมาจากระบบ เป็นข้อมูลการเข้าถึงกิจกรรมการใช้งานที่หลากหลายของผู้ดูแลระบบ อาจารย์ และนิสิต เป็นรายบุคคลในแต่ละวันที่และเวลา จึงต้องมีการทำความสะอาดข้อมูล และปรับเปลี่ยนข้อมูลให้อยู่ในรูปแบบที่เหมาะสม โดยมีขั้นตอนการดำเนินงาน โดยใช้เครื่องมือต่าง ๆ ในโปรแกรม Microsoft Excel ดังนี้

1. เลือกข้อมูลเฉพาะข้อมูลการใช้งานของนิสิต ตั้งแต่วันที่ 7 มกราคม ถึง 5 มีนาคม 2564 จำนวน 99,925 รายการ

2. ปรับรูปแบบวันที่และเวลาการใช้งานให้แยกวันที่และเวลาออกจากกัน

**การทำวิศวกรรมคุณลักษณะ**

หลังจากทำการเก็บรวบรวมและเตรียมข้อมูลบันทึกการใช้งานระบบแล้ว จะสร้างคุณลักษณะที่เกี่ยวข้องกับพฤติกรรมของนิสิตในระบบ จึงใช้การทำวิศวกรรมคุณลักษณะ เพื่อสร้างคุณลักษณะสำหรับสรุปสถิติการใช้งานระบบที่มี 4 ด้าน ได้แก่ ด้านการเข้าสู่ระบบ, ด้านการชมสื่อการเรียนรู้, ด้านการทำแบบทดสอบ และด้านการทำงานมอบหมาย โดย ทั้งนี้ใช้เครื่องมือต่าง ๆ ในโปรแกรม Microsoft Excel ในการสรุปข้อมูล ได้แก่ Pivot Table เป็นเครื่องมือที่นำข้อมูลไปวิเคราะห์และสรุปผลข้อมูลตามเงื่อนไขที่กำหนดได้อย่างง่ายดายในรูปแบบของตาราง, ฟังก์ชัน Vlookup เป็นฟังก์ชันที่ช่วยค้นหาข้อมูลที่ต้องการ เพื่ออ้างอิงข้อมูล การทำวิศวกรรมคุณลักษณะมีขั้นตอนการดำเนินงาน ดังนี้

1. วิเคราะห์และสรุปข้อมูลพฤติกรรมการใช้งานระบบจากข้อมูลบันทึกการใช้งานระบบ โดยใช้เครื่องมือ Pivot Table ฟังก์ชัน Vookup และสูตรการคำนวณทางสถิติ ดังภาพประกอบ 30 - 48

| IdStudent | ClickLogin |
|---|---|
| | 51 |
| | 39 |
| | 57 |
| | 60 |
| | 104 |
| | 14 |
| | 45 |
| | 47 |
| | 23 |
| | 33 |

ภาพประกอบ 30 แสดงจำนวนครั้งที่คลิกเข้าสู่ระบบ

| IdStudent | ClickPage1 | ClickPage2 | ClickPage3 | ClickPage4 | ClickPage5 | ClickPage6 | ClickPage7 | ClickPage8 | ClickPage9 | ClickPage10 | ClickPage11 | Total_ClickPage |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 5 | 3 | 3 | 3 | 4 | 3 | 4 | 1 | | | | 26 |
| | 1 | 2 | 1 | 2 | 1 | 5 | 3 | 2 | | | 1 | 18 |
| | 4 | 5 | 4 | 2 | 4 | 5 | 3 | 3 | 2 | 1 | | 33 |
| | 4 | 3 | 2 | 4 | 5 | 3 | 4 | 1 | 3 | 2 | | 31 |
| | 6 | 5 | 4 | 2 | 6 | 4 | 4 | 1 | 2 | | | 34 |
| | 2 | 2 | 3 | 2 | 2 | 4 | 1 | | 1 | | | 17 |
| | 3 | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 3 | 1 | | 20 |
| | 2 | 1 | 2 | 2 | 1 | 4 | 1 | 1 | 2 | | | 16 |
| | 2 | 3 | 1 | 2 | 3 | 5 | 1 | 1 | 2 | | | 20 |
| | 2 | 2 | 1 | 1 | 1 | 2 | 2 | 1 | | | | 12 |

ภาพประกอบ 31 แสดงจำนวนครั้งที่คลิกเข้าถึงสื่อการเรียนรู้

| IdStuden | Page1 | Page10 | Page11 | Page2 | Page3 | Page4 | Page5 | Page6 | Page7 | Page8 | Page9 | NumDaysPageViewed |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 3 | | | 2 | 2 | 2 | 2 | 3 | 3 | 1 | | 8 |
| | 1 | | 1 | 2 | 1 | 2 | 1 | 2 | 2 | 2 | | 3 |
| | 3 | 1 | | 3 | 3 | 2 | 2 | 3 | 2 | 2 | 2 | 7 |
| | 2 | 2 | | 3 | 2 | 3 | 3 | 3 | 3 | 1 | 2 | 5 |
| | 3 | | | 3 | 2 | 1 | 3 | 3 | 2 | 1 | 2 | 7 |
| | 2 | | | 2 | 2 | 2 | 2 | 3 | 1 | | 1 | 5 |
| | 2 | 1 | | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 3 |
| | 2 | | | 1 | 2 | 2 | 1 | 4 | 1 | 1 | 1 | 8 |
| | 2 | | | 2 | 1 | 2 | 2 | 3 | 1 | 1 | 1 | 4 |
| | 2 | | | 1 | 1 | 1 | 1 | 1 | 1 | 1 | | 5 |

ภาพประกอบ 32 แสดงจำนวนวันที่ชมสื่อการเรียนรู้

| Sum of Duration | Colum | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| IdStudent | Page1 | Page10 | Page11 | Page2 | Page3 | Page4 | Page5 | Page6 | Page7 | Page8 | Page9 | TotalDuration (Days) |
| | 1 | | | 8 | 15 | 52 | 22 | 29 | 43 | 52 | | 222 |
| | 29 | | 53 | 29 | 29 | 29 | 29 | 29 | 29 | 46 | | 302 |
| | 22 | 52 | | 8 | 8 | 22 | 22 | 29 | 52 | 52 | 52 | 319 |
| | 42 | 42 | | 8 | 42 | 42 | 42 | 29 | 42 | 42 | 42 | 373 |
| | 7 | | | 8 | 31 | 55 | 31 | 29 | 43 | 55 | 53 | 312 |
| | 1 | | | 8 | 51 | 51 | 51 | 29 | 55 | | 51 | 297 |
| | 52 | 52 | | 52 | 22 | 22 | 22 | 52 | 52 | 52 | 52 | 430 |
| | 8 | | | 44 | 22 | 44 | 44 | 29 | 44 | 44 | 52 | 331 |
| | 53 | | | 53 | 53 | 53 | 53 | 29 | 43 | 53 | 53 | 443 |
| | 1 | | | 8 | 29 | 29 | 29 | 29 | 43 | 45 | | 213 |

ภาพประกอบ 33 แสดงจำนวนวันระหว่างวันแรกที่เรียน และวันแรกที่ชมสื่อการเรียนรู้

| IdStudent | NumTitlePageView |
|---|---|
| | 8 |
| | 8 |
| | 9 |
| | 9 |
| | 9 |
| | 8 |
| | 9 |
| | 9 |
| | 9 |
| | 8 |

ภาพประกอบ 34 แสดงจำนวนเรื่องที่ชมสื่อการเรียนรู้

| IdStudent | ClickQuiz1 | ClickQuiz2 | ClickQuiz3 | ClickQuiz4 | ClickQuiz5 | ClickQuiz6 | Total_ClickQuiz |
|---|---|---|---|---|---|---|---|
| | 7 | 2 | 1 | 5 | 3 | | 18 |
| | 9 | 5 | 4 | 9 | 9 | | 36 |
| | 1 | 1 | 2 | 5 | 3 | 2 | 14 |
| | 3 | 2 | 3 | 4 | 3 | 2 | 17 |
| | 14 | 8 | 12 | 13 | 8 | 5 | 60 |
| | 2 | 8 | 5 | 2 | 2 | 1 | 20 |
| | 5 | 7 | 4 | 2 | 2 | 2 | 22 |
| | 6 | 7 | 3 | 3 | 7 | 5 | 31 |
| | 4 | 4 | 4 | 2 | 2 | 3 | 19 |
| | 2 | 2 | 1 | 4 | 5 | | 14 |

ภาพประกอบ 35 แสดงจำนวนครั้งที่คลิกเข้าถึงแบบทดสอบ

| EventName | Quiz attempt submitted |
|---|---|

| IdStudent | EventContext | Month | Dat | Deadline | Duration (days) |
|---|---|---|---|---|---|
| | Quiz1 | ม.ค. | 14-ม.ค. | 28/2/2021 | 46 |
| | Quiz2 | ม.ค. | 14-ม.ค. | 28/2/2021 | 46 |
| | Quiz3 | ม.ค. | 28-ม.ค. | 28/2/2021 | 32 |
| | Quiz4 | ม.ค. | 28-ม.ค. | 28/2/2021 | 32 |
| | | ก.พ. | 27-ก.พ. | 28/2/2021 | 2 |
| | Quiz5 | ก.พ. | 27-ก.พ. | 28/2/2021 | 2 |
| | Quiz1 | ก.พ. | 28-ก.พ. | 28/2/2021 | 1 |
| | Quiz2 | ก.พ. | 28-ก.พ. | 28/2/2021 | 1 |
| | Quiz3 | ก.พ. | 28-ก.พ. | 28/2/2021 | 1 |
| | Quiz4 | ก.พ. | 28-ก.พ. | 28/2/2021 | 1 |
| | Quiz5 | ก.พ. | 4-ก.พ. | 28/2/2021 | 25 |
| | | | 28-ก.พ. | 28/2/2021 | 1 |

ภาพประกอบ 36 แสดงจำนวนวันระหว่างวันที่ทำแบบทดสอบครั้งแรก และวันสุดท้ายการทำแบบทดสอบ

| EventName | Quiz attempt submitted |
|---|---|

| Count of EventName | EventContext | | | | | | |
|---|---|---|---|---|---|---|---|
| IdStudent | Quiz1 | Quiz2 | Quiz3 | Quiz4 | Quiz5 | Quiz6 | TotalRepeat (times) |
| | 3 | 1 | 1 | 3 | 2 | | 10 |
| | 3 | 3 | 3 | 3 | 3 | | 15 |
| | 1 | 1 | 1 | 2 | 1 | 1 | 7 |
| | 2 | 1 | 1 | 2 | 3 | 1 | 10 |
| | 3 | 1 | 3 | 3 | 3 | 3 | 16 |
| | 1 | 3 | 3 | 2 | 1 | 1 | 11 |
| | 3 | 3 | 2 | 2 | 1 | 1 | 12 |
| | 2 | 3 | 1 | 2 | 3 | 3 | 14 |
| | 2 | 3 | 2 | 2 | 1 | 3 | 13 |
| | 1 | 1 | 1 | 3 | 3 | | 9 |

ภาพประกอบ 37 แสดงจำนวนครั้งที่ทำแบบทดสอบซ้ำ

| IdStudent | NumTitleQuiz |
|---|---|
| | 5 |
| | 5 |
| | 6 |
| | 6 |
| | 6 |
| | 6 |
| | 6 |
| | 6 |
| | 6 |
| | 5 |

ภาพประกอบ 38 แสดงจำนวนเรื่องที่ทำแบบทดสอบ

| | Quiz1 | | | Quiz2 | | | Quiz3 | | | Quiz4 | | | Quiz5 | | | Quiz6 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| IdStudent | Duration1 | Duration2 | Duration3 | Duration1 | Duration2 | Duration3 | Duration1 | Duration2 | Duration3 | Duration1 | Duration2 | Duration3 | Duration1 | Duration2 | Duration3 | Duration1 | Duration2 | Duration3 |
| | 7 | 7 | 3 | 2 | 0 | 0 | 3 | 0 | 0 | 7 | 2 | 1 | 3 | 2 | 0 | 0 | 0 | 0 |
| | 7 | 2 | 5 | 2 | 2 | 5 | 1 | 2 | 1 | 5 | 1 | 2 | 30 | 27 | 5 | 0 | 0 | 0 |
| | 2 | 0 | 0 | 2 | 0 | 0 | 3 | 0 | 0 | 8 | 1 | 0 | 3 | 0 | 0 | 2 | 0 | 0 |
| | 5 | 2 | 0 | 4 | 0 | 0 | 3 | 0 | 0 | 6 | 2 | 0 | 15 | 1 | 1 | 4 | 0 | 0 |
| | 6 | 5 | 5 | 2 | 0 | 0 | 3 | 3 | 3 | 20 | 17 | 2 | 5 | 1 | 1 | 4 | 2 | 2 |
| | 1 | 0 | 0 | 4 | 0 | 1 | 3 | 3 | 2 | 5 | 1 | 0 | 17 | 0 | 0 | 5 | 0 | 0 |
| | 10 | 3 | 2 | 1 | 2 | 10 | 4 | 1 | 0 | 2 | 1 | 0 | 9 | 0 | 0 | 4 | 0 | 0 |
| | 6 | 7 | 0 | 2 | 2 | 0 | 2 | 0 | 0 | 7 | 1 | 0 | 18 | 9 | 30 | 2 | 2 | 1 |
| | 3 | 3 | 0 | 2 | 2 | 2 | 2 | 1 | 0 | 3 | 2 | 0 | 23 | 0 | 0 | 2 | 2 | 1 |
| | 3 | 0 | 0 | 3 | 0 | 0 | 2 | 0 | 0 | 2 | 2 | 2 | 19 | 2 | 1 | 0 | 0 | 0 |

ภาพประกอบ 39 แสดงระยะเวลาการทำแบบทดสอบ ครั้งที่ 1 ถึง ครั้งที่ 3

| IdStudent | Quiz1 | Quiz2 | Quiz3 | Quiz4 | Quiz5 | Quiz6 | SumDuration (minutes) |
|---|---|---|---|---|---|---|---|
| | 17 | 2 | 3 | 10 | 5 | 0 | 37 |
| | 14 | 9 | 4 | 8 | 62 | 0 | 97 |
| | 2 | 2 | 3 | 9 | 3 | 2 | 21 |
| | 7 | 4 | 3 | 8 | 17 | 4 | 43 |
| | 16 | 2 | 9 | 39 | 7 | 8 | 81 |
| | 1 | 5 | 8 | 6 | 17 | 5 | 42 |
| | 15 | 13 | 5 | 3 | 9 | 4 | 49 |
| | 13 | 4 | 2 | 8 | 57 | 5 | 89 |
| | 6 | 6 | 3 | 5 | 23 | 5 | 48 |
| | 3 | 3 | 2 | 6 | 22 | 0 | 36 |

ภาพประกอบ 40 แสดงระยะเวลารวมการทำแบบทดสอบรวมทุกครั้ง

| IdStudent | Quiz1 | Quiz2 | Quiz3 | Quiz4 | Quiz5 | Quiz6 | AverageDuration (minutes) |
|---|---|---|---|---|---|---|---|
| | 5 | 2 | 3 | 3 | 2 | 0 | 2 |
| | 4 | 3 | 1 | 2 | 20 | 0 | 5 |
| | 2 | 2 | 3 | 4 | 3 | 2 | 2 |
| | 3 | 4 | 3 | 4 | 5 | 4 | 3 |
| | 5 | 2 | 3 | 13 | 2 | 2 | 4 |
| | 1 | 1 | 2 | 3 | 17 | 5 | 4 |
| | 5 | 4 | 2 | 1 | 9 | 4 | 4 |
| | 6 | 1 | 2 | 4 | 19 | 1 | 5 |
| | 3 | 2 | 1 | 2 | 23 | 1 | 5 |
| | 3 | 3 | 2 | 2 | 7 | 0 | 2 |

ภาพประกอบ 41 แสดงระยะเวลาเฉลี่ยการทำแบบทดสอบรวมทุกครั้ง

| IdStudent | ClickAssignment1 | ClickAssignment2 | ClickAssignment3 | Total_ClickAssignment |
|---|---|---|---|---|
| | 11 | 3 | | 14 |
| | 9 | | | 9 |
| | 8 | | | 8 |
| | 7 | | | 7 |
| | 14 | 1 | | 15 |
| | 4 | | | 4 |
| | 9 | | | 9 |
| | 13 | 1 | | 14 |
| | 6 | 1 | | 7 |
| | 5 | 1 | | 6 |

ภาพประกอบ 42 แสดงจำนวนครั้งที่คลิกเข้าถึงงานมอบหมาย

| EventName | A submission has been uploaded | | | |
|---|---|---|---|---|
| EventContext | Assignment | | | |
| | | | | |
| IdStudent | Months | Date | Deadline | Duration (days) |
| | ก.พ. | 18-ก.พ. | 20/2/2021 | 3 |
| | ก.พ. | 20-ก.พ. | 20/2/2021 | 1 |
| | ก.พ. | 20-ก.พ. | 20/2/2021 | 1 |
| | ก.พ. | 19-ก.พ. | 20/2/2021 | 2 |
| | ก.พ. | 20-ก.พ. | 20/2/2021 | 1 |
| | ก.พ. | 20-ก.พ. | 20/2/2021 | 1 |
| | ก.พ. | 20-ก.พ. | 20/2/2021 | 1 |
| | ก.พ. | 20-ก.พ. | 20/2/2021 | 1 |
| | ก.พ. | 20-ก.พ. | 20/2/2021 | 1 |
| | ก.พ. | 20-ก.พ. | 20/2/2021 | 1 |

ภาพประกอบ 43 แสดงจำนวนวันระหว่างวันแรกที่ส่งงานมอบหมาย และวันสุดท้ายการส่งงานมอบหมาย

| IdStudent | ScoreOnlineMeeting1 | ScoreOnlineMeeting2 | ScoreOnlineMeeting3 | TotalScoreOnlineMeeting |
|---|---|---|---|---|
| | 3 | 3 | 1.8 | 7.8 |
| | 1.31 | 1 | 0 | 2.31 |
| | 3 | 3 | 1.95 | 7.95 |
| | 2.38 | 3 | 1.95 | 7.33 |
| | 2.69 | 1.5 | 2.4 | 6.59 |
| | 2.44 | 2.5 | 1.5 | 6.44 |
| | 3 | 3 | 0 | 6 |
| | 3 | 3 | 1.2 | 7.2 |
| | 0.5 | 2 | 1.95 | 4.45 |
| | 0 | 0 | 0 | 0 |

ภาพประกอบ 44 แสดงคะแนนกิจกรรมในห้องเรียนออนไลน์

| IdStudent | ScoreAttendance |
|---|---|
| | 10.00 |
| | 10.00 |
| | 0.00 |
| | 0.00 |
| | 10.00 |
| | 10.00 |
| | 10.00 |
| | 10.00 |
| | 10.00 |
| | 9.00 |

ภาพประกอบ 45 แสดงคะแนนการเข้าเรียน

| IdStudent | ScoreQuiz1 | ScoreQuiz2 | ScoreQuiz3 | ScoreQuiz4 | ScoreQuiz5 | ScoreQuiz6 | Total_ScoreQuiz |
|---|---|---|---|---|---|---|---|
| | 8.18 | 10 | 10 | 10 | 15 | | 53.18 |
| | 9.09 | 5 | 2 | 9 | 0 | | 25.09 |
| | 10 | 10 | 10 | 10 | 15 | 10 | 65.00 |
| | 10 | 10 | 10 | 10 | 15 | 10 | 65.00 |
| | 9.09 | 10 | 6 | 9 | 10 | 10 | 54.09 |
| | 10 | 5 | 10 | 10 | 5 | 10 | 50.00 |
| | 10 | 5 | 10 | 10 | 15 | 10 | 60.00 |
| | 10 | 10 | 10 | 10 | 10 | 10 | 60.00 |
| | 10 | 5 | 10 | 10 | 0 | 10 | 45.00 |
| | 10 | 10 | 10 | 10 | 10 | | 50.00 |

ภาพประกอบ 46 แสดงคะแนนแบบทดสอบ

| IdStudent | ScoreMidterm |
|---|---|
| | 15.00 |
| | 19.50 |
| | 14.00 |
| | 14.50 |
| | 21.00 |
| | 14.50 |
| | 14.00 |
| | 14.00 |
| | 14.50 |
| | 16.50 |

ภาพประกอบ 47 แสดงคะแนนการสอบกลางภาค

| IdStudent | ScoreAssignment |
|---|---|
| | 55.00 |
| | 0.00 |
| | 60.00 |
| | 60.00 |
| | 46.67 |
| | 0.00 |
| | 50.00 |
| | 60.00 |
| | 40.00 |

ภาพประกอบ 48 แสดงคะแนนงานมอบหมาย

2. นำคุณลักษณะพฤติกรรมการใช้งานระบบเหล่านี้ไปรวมกับคะแนน ข้อมูลทั่วไป และผลการเรียนของนิสิต โดยใช้ฟังก์ชัน Vlookup

3. ตรวจสอบความผิดปกติของข้อมูล พบว่ามีระยะเวลาการทำแบบทดสอบที่สูงกว่าที่กำหนดเท่านั้น เนื่องจากการทำแบบทดสอบซ้ำได้ 3 ครั้ง อาจจะทำให้นิสิตใช้เวลาในการคิดวิเคราะห์ในการหาคำตอบมากขึ้น หรือไม่ได้ออกจากระบบจนกว่าสัญญาณอินเทอร์เน็ตหลุด ทั้งนี้อาจจะเกิดจากการที่ไม่ได้ตั้งเวลาเปิดปิดการทำแบบทดสอบแบบอัตโนมัติ

4. ทำความสะอาดข้อมูล โดยการแทนระยะเวลาการทำแบบทดสอบที่ผิดปกติด้วยระยะเวลาที่กำหนด และแทนข้อมูลที่ไม่มีการใช้งานที่เป็นค่าว่างด้วยค่าศูนย์

5. จะได้ชุดข้อมูลของนิสิตจำนวน 405 คน ที่พร้อมเข้าสู่แบบจำลองสำหรับการทำนายผลการเรียนของนิสิตจำนวน 4 ชุด ดังตาราง 7 ที่แบ่งตามกรณีของชุดคุณลักษณะจากตาราง 6

ตาราง 7 แสดงรายละเอียดคำอธิบายของชุดข้อมูล

| ตัวแปร | คำอธิบายของตัวแปร | ชุดข้อมูลที่ 1 จากกรณีที่ 1 (116 คอลัมน์) | ชุดข้อมูลที่ 2 จากกรณีที่ 2 (112 คอลัมน์) | ชุดข้อมูลที่ 3 จากกรณีที่ 3 (100 คอลัมน์) | ชุดข้อมูลที่ 4 จากกรณีที่ 4 (19 คอลัมน์) |
|---|---|---|---|---|---|
| **1. ผลลัพธ์การทำนาย (Target หรือ Class)** | | | | | |
| **ประเภทหมวดหมู่** | | | | | |
| - เกรด (Grade) | A, B+, B, C+, C, D+, D, E, W | ✓ | ✓ | ✓ | ✓ |
| **2. ตัวทำนาย (Predictor หรือ Feature)** | | | | | |
| **2.1 ประเภทหมวดหมู่** | | | | | |
| - คณะที่นิสิตเรียน (Faculty) | มนุษยศาสตร์, ศึกษาศาสตร์, สังคมศาสตร์ | ✓ | ✓ | ✓ | ✓ |
| - เพศ (Gender) | ชาย, หญิง | ✓ | ✓ | ✓ | ✓ |
| **2.2 ประเภทตัวเลข** | | | | | |
| **2.2.1 ด้านพฤติกรรมการใช้งานระบบของนิสิต** | | | | | |
| **การเข้าสู่ระบบ** | | | | | |
| - จำนวนครั้งที่คลิกเข้าสู่ระบบ (ClickLogin) | จำนวนครั้งที่คลิกเข้าสู่ระบบ | ✓ | ✓ | ✓ | |
| **การชมสื่อการเรียนรู้ จำนวน 10 เรื่อง** | | | | | |
| - จำนวนครั้งที่คลิกเข้าถึงสื่อการเรียนรู้ (ClickPage) | จำนวนครั้งที่คลิกเข้าถึงสื่อการเรียนรู้แต่ละเรื่อง และรวมทุกเรื่อง | ✓ | ✓ | ✓ | |
| - จำนวนวันที่ชมสื่อการเรียนรู้ (DayPageView) | จำนวนวันที่ชมสื่อการเรียนรู้แต่ละเรื่อง และรวมทุกเรื่อง | ✓ | ✓ | ✓ | |
| - จำนวนวันระหว่างวันแรกที่เรียน และวันแรกที่ชมสื่อการเรียนรู้ (DayFirstPageView) | จำนวนวันระหว่างวันแรกที่เรียน และวันแรกที่ชมสื่อการเรียนรู้แต่ละเรื่อง และรวมทุกเรื่อง | ✓ | ✓ | ✓ | |

ตาราง 7 (ต่อ)

| ตัวแปร | คำอธิบายของตัวแปร | ชุดข้อมูลที่ 1 จากกรณีที่ 1 (116 คอลัมน์) | ชุดข้อมูลที่ 2 จากกรณีที่ 2 (112 คอลัมน์) | ชุดข้อมูลที่ 3 จากกรณีที่ 3 (100 คอลัมน์) | ชุดข้อมูลที่ 4 จากกรณีที่ 4 (19 คอลัมน์) |
|---|---|---|---|---|---|
| - จำนวนเรื่องที่ชมสื่อการเรียนรู้ (NumTitlePageView) | จำนวน 1 ถึง 10 เรื่อง | ✓ | ✓ | ✓ | |
| **การทำแบบทดสอบ จำนวน 6 เรื่อง โดยสามารถทำซ้ำได้ 3 ครั้ง** | | | | | |
| - จำนวนครั้งที่คลิกเข้าถึงแบบทดสอบ (ClickQuiz) | จำนวนครั้งที่คลิกเข้าถึงแบบทดสอบแต่ละเรื่อง และรวมทุกเรื่อง | ✓ | ✓ | ✓ | |
| - จำนวนวันระหว่างวันที่ทำแบบทดสอบครั้งแรกและวันสุดท้ายการทำแบบทดสอบ (DayQuizBefore) | จำนวนวันระหว่างวันที่ทำแบบทดสอบครั้งแรก และวันสุดท้ายการทำแบบทดสอบแต่ละเรื่อง และรวมทุกเรื่อง | ✓ | ✓ | ✓ | |
| - จำนวนครั้งที่ทำแบบทดสอบซ้ำ (RepeatQuiz) | จำนวน 1 ถึง 3 ครั้ง | ✓ | ✓ | ✓ | |
| - จำนวนเรื่องที่ทำแบบทดสอบ (NumTitleQuiz) | จำนวน 1 ถึง 6 เรื่อง | ✓ | ✓ | ✓ | |
| - ระยะเวลาการทำแบบทดสอบ ครั้งที่ 1 (DurationQuiz_1) | ระยะเวลาการทำแบบทดสอบแต่ละเรื่อง (นาที) | ✓ | ✓ | ✓ | |
| - ระยะเวลาการทำแบบทดสอบ ครั้งที่ 2 (DurationQuiz_2) | ระยะเวลาการทำแบบทดสอบแต่ละเรื่อง (นาที) | ✓ | ✓ | ✓ | |

ตาราง 7 (ต่อ)

| ตัวแปร | คำอธิบายของตัวแปร | ชุดข้อมูลที่ 1 จากกรณีที่ 1 (116 คอลัมน์) | ชุดข้อมูลที่ 2 จากกรณีที่ 2 (112 คอลัมน์) | ชุดข้อมูลที่ 3 จากกรณีที่ 3 (100 คอลัมน์) | ชุดข้อมูลที่ 4 จากกรณีที่ 4 (19 คอลัมน์) |
|---|---|---|---|---|---|
| - ระยะเวลาการทำแบบทดสอบ ครั้งที่ 3 (DurationQuiz_3) | ระยะเวลาการทำแบบทดสอบแต่ละเรื่อง (นาที) | ✓ | ✓ | ✓ | |
| - ระยะเวลารวมการทำแบบทดสอบรวมทุกครั้ง (SumMinuteQuiz) | ระยะเวลารวมการทำแบบทดสอบรวมทุกครั้งแต่ละเรื่อง (นาที) | ✓ | ✓ | ✓ | |
| - ระยะเวลาเฉลี่ยการทำแบบทดสอบรวมทุกครั้ง (AvgMinuteQuiz) | ระยะเวลาเฉลี่ยการทำแบบทดสอบรวมทุกครั้งแต่ละเรื่อง (นาที) | ✓ | ✓ | ✓ | |
| **การส่งงานมอบหมาย** | | | | | |
| - จำนวนครั้งที่คลิกเข้าถึงงานมอบหมาย (ClickAssignment) | จำนวนครั้งที่คลิกเข้าถึงงานมอบหมายแต่ละเรื่อง และรวมทุกเรื่อง | ✓ | ✓ | ✓ | |
| - จำนวนวันระหว่างวันแรกที่ส่งงานมอบหมายและวันสุดท้ายการส่งงานมอบหมาย (DayAssignment Before) | จำนวนวันระหว่างวันแรกที่ส่งงานมอบหมายและวันสุดท้ายการส่งงานมอบหมาย | ✓ | ✓ | ✓ | |
| **2.2.2 ด้านคะแนนการประเมินผลการเรียนรู้** | | | | | |
| **คะแนนกิจกรรมในห้องเรียน (30%) จำนวน 3 เรื่อง** | | | | | |
| - คะแนนกิจกรรมในห้องเรียนออนไลน์ (ScoreOnline Meeting) | คะแนนกิจกรรมในห้องเรียนออนไลน์แต่ละเรื่อง และรวมทุกเรื่อง | ✓ | ✓ | | ✓ |

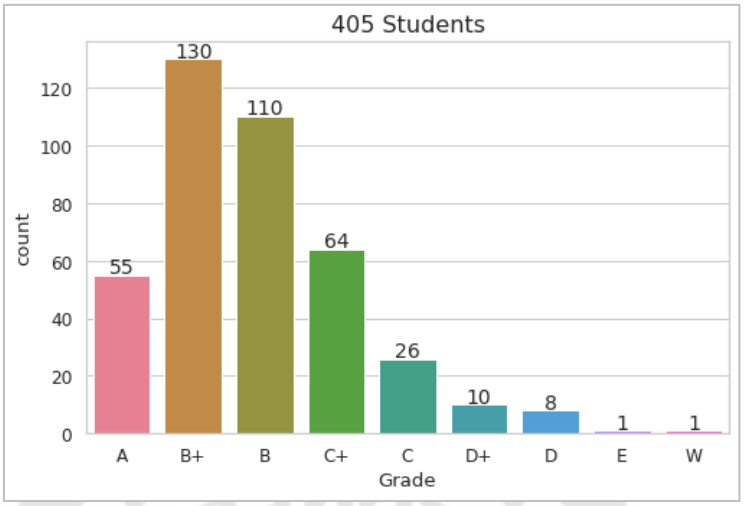
ตาราง 7 (ต่อ)

| ตัวแปร | คำอธิบายของตัวแปร | ชุดข้อมูลที่ 1 จากกรณีที่ 1 (116 คอลัมน์) | ชุดข้อมูลที่ 2 จากกรณีที่ 2 (112 คอลัมน์) | ชุดข้อมูลที่ 3 จากกรณีที่ 3 (100 คอลัมน์) | ชุดข้อมูลที่ 4 จากกรณีที่ 4 (19 คอลัมน์) |
|---|---|---|---|---|---|
| - คะแนนการเข้าเรียน (ScoreAttendance) | คะแนนเฉลี่ยการเข้าเรียน | ✓ | ✓ | | ✓ |
| - คะแนนแบบทดสอบ (ScoreQuiz) | คะแนนแบบทดสอบแต่ละเรื่อง และรวมทุกเรื่อง | ✓ | ✓ | | ✓ |
| **คะแนนกลางภาค (30%)** | | | | | |
| - คะแนนสอบกลางภาค (ScoreMidterm) | คะแนนสอบกลางภาค | ✓ | | | ✓ |
| - คะแนนงานมอบหมาย (ScoreAssignment) | คะแนนงานมอบหมาย เรื่องเทคโนโลยีสารสนเทศและการสื่อสารที่น่าสนใจในปัจจุบันและแนวโน้มในอนาคต แต่ละส่วน และรวมทุกส่วน | ✓ | | | ✓ |

**การสำรวจข้อมูล**

จากการสำรวจข้อมูล โดยใช้ภาษา Python พบว่าไม่มีข้อมูลที่ขาดหายไปในคุณลักษณะใด ๆ และนำคุณลักษณะเหล่านี้ไปศึกษาความสัมพันธ์กับระดับผลการเรียนด้วยวิธีการกระจายของข้อมูลแบบฮิสโตแกรม (Histogram) ที่แสดงความถี่ของข้อมูล และวิธีการกระจายของข้อมูลแบบการประมาณความหนาแน่นเชิงพื้นที่แบบเคอร์เนล (Kernel Density Estimation) ที่แสดงเส้นกราฟให้มีความเรียบและมีความชัดเจนขึ้น การสำรวจข้อมูลนี้แบ่งตามผลลัพธ์การทำนายและประเภทของคุณลักษณะ

**1. ผลลัพธ์การทำนาย: ผลการเรียนของนิสิต**



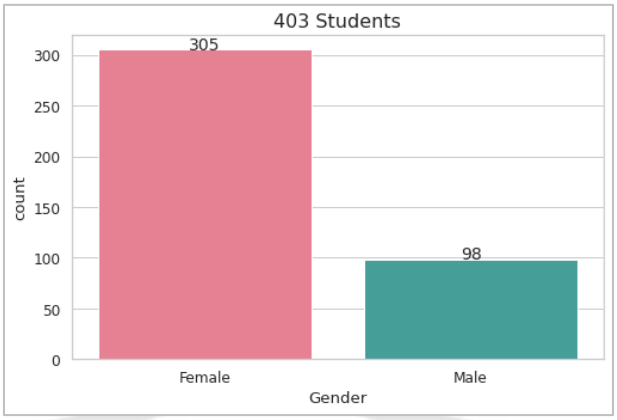

ภาพประกอบ 49 แสดงจำนวนนิสิตจำแนกเกรด

จากภาพประกอบ 49 พบว่ามีจำนวนนิสิตที่ได้เกรด E และเกรด W น้อยมาก และจะส่งผลต่อการกระจายตัวของข้อมูลพฤติกรรมการเรียนรู้เป็นอย่างมาก เพื่อให้มีความแตกต่างกันของตัวจำแนกข้อมูลน้อยลง ทำให้ทำนายง่ายขึ้น จึงทำการคัดข้อมูลของนิสิตที่ได้เกรด E และ W ออก จะได้จำนวนนิสิตที่เหลือ 403 คน และทำการรวมกลุ่มเกรดแต่ละระดับผลการเรียน ดังตาราง 8

ตาราง 8 แสดงการจำแนกกลุ่มเกรดตามระดับผลการเรียน

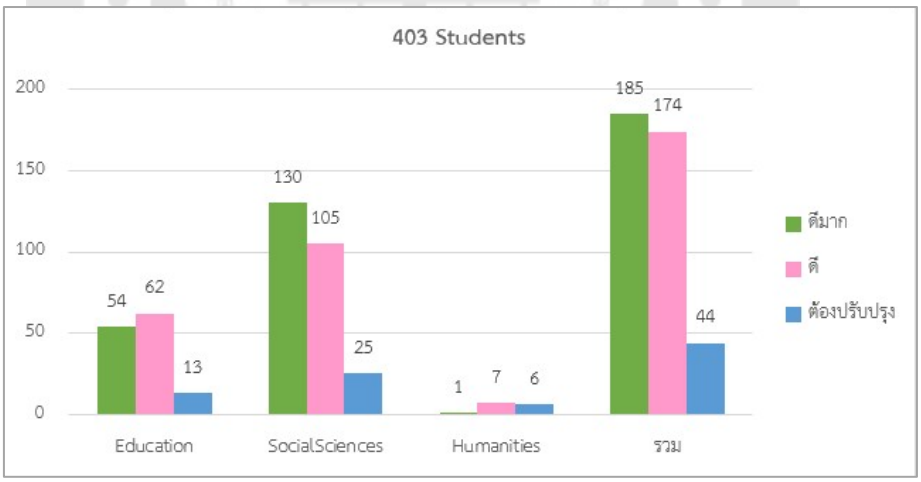
| **เกรด** | **ระดับผลการเรียนจากการรวมเกรด** | **จำนวนนิสิต (คน)** |
|---|---|---|
| A และ B+ | ดีมาก | 185 |
| B และ C+ | ดี | 174 |
| C, D+ และ D | ต้องปรับปรุง | 44 |

**2. คุณลักษณะที่เกี่ยวข้องกับข้อมูลทั่วไป**




ภาพประกอบ 50 แสดงจำนวนนิสิตจำแนกเพศ

จากภาพประกอบ 50 พบว่านิสิตเพศหญิง จำนวน 305 คน คิดเป็นร้อยละ 75.68 โดยประมาณสามเท่าของนิสิตเพศชาย จำนวน 98 คน คิดเป็นร้อยละ 24.32



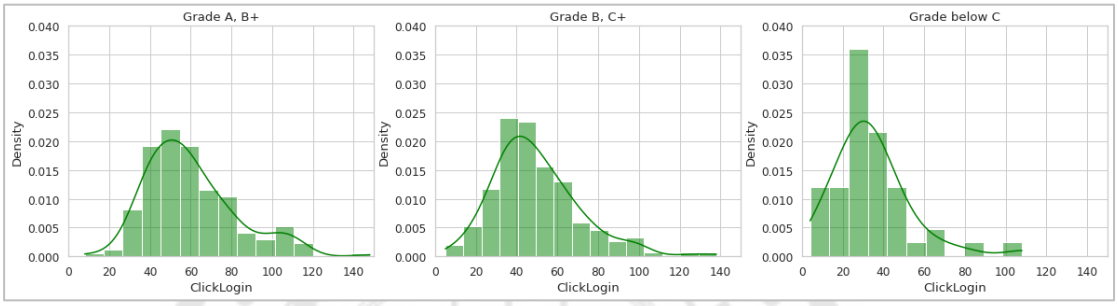

ภาพประกอบ 51 แสดงจำนวนนิสิตจำแนกคณะ

จากภาพประกอบ 51 พบว่านิสิตคณะสังคมศาสตร์ จำนวน 260 คน มาก รองลงมาเป็นนิสิตคณะศึกษาศาสตร์ จำนวน 129 คน นิสิตคณะมนุษยศาสตร์ จำนวน 14 คน ตามลำดับ และนิสิตคณะศึกษาศาสตร์ส่วนใหญ่ ได้รับผลการเรียนดีเป็นจำนวน 62 คน รองลงมากลุ่มผลการเรียนดีมากจำนวน 54 คน กลุ่มผลการเรียนต้องปรับปรุงจำนวน 13 คน ส่วนนิสิตคณะสังคมศาสตร์ส่วนใหญ่ได้รับผลการเรียนดีมากเป็นจำนวน 130 คน รองลงมากลุ่มผลการเรียนดีจำนวน 105 คน กลุ่มผลการ

เรียนต้องปรับปรุงจำนวน 25 คน และนิสิตคณะมนุษยศาสตร์ส่วนใหญ่ ได้รับผลการเรียนดีเป็นจำนวน 7 คน รองลงมากลุ่มผลการเรียนต้องปรับปรุงจำนวน 6 คน กลุ่มผลการเรียนดีมากจำนวน 1 คน โดยรวมสรุปว่านิสิตรวมทุกคณะส่วนใหญ่ได้รับผลการเรียนดีมากเป็นจำนวน 185 คน รองลงมากลุ่มผลการเรียนดีจำนวน 174 คน กลุ่มผลการเรียนต้องปรับปรุงจำนวน 44 คน

### 3. คุณลักษณะที่เกี่ยวข้องกับพฤติกรรมการใช้งานระบบของนิสิต

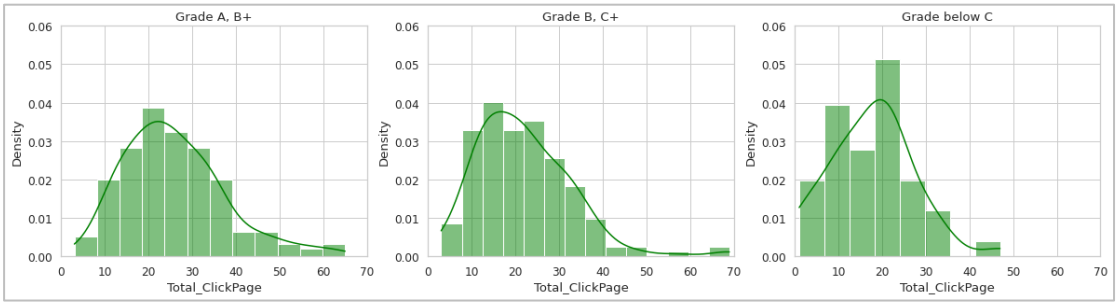
**- ด้านการเข้าสู่ระบบ**




ภาพประกอบ 52 แสดงความสัมพันธ์ระหว่างจำนวนครั้งที่คลิกเข้าสู่ระบบกับกลุ่มเกรด

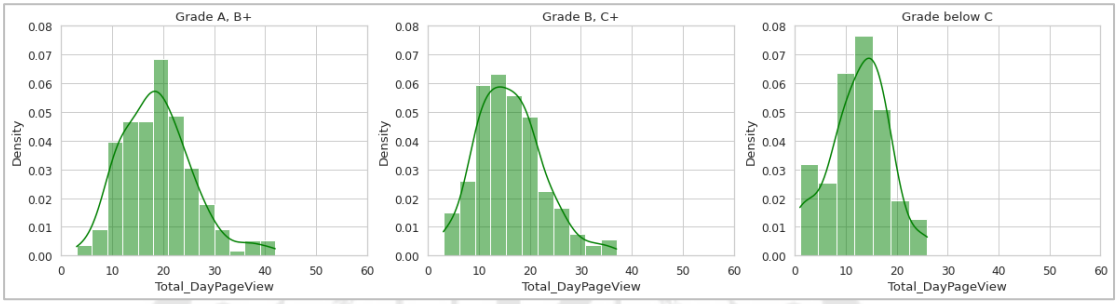
จากภาพประกอบ 52 แสดงให้เห็นว่า จำนวนครั้งที่คลิกเข้าสู่ระบบมากที่สุดในอันดับแรก คือ นิสิตกลุ่มผลการเรียนดีมาก จำนวนครั้งโดยประมาณ 150 ครั้ง ต่อมานิสิตกลุ่มผลการเรียนดี จำนวนครั้งโดยประมาณ 140 ครั้ง สุดท้ายนิสิตกลุ่มผลการเรียนต้องปรับปรุง จำนวนครั้งโดยประมาณ 110 ครั้ง ส่วนจำนวนนิสิตที่คลิกเข้าสู่ระบบน้อยที่สุดเป็นส่วนใหญ่ในอันดับแรก คือ นิสิตกลุ่มผลการเรียนต้องปรับปรุง รองลงมานิสิตกลุ่มผลการเรียนดี ดีมาก ตามลำดับ จะเห็นได้ว่า จำนวนครั้งที่คลิกเข้าสู่ระบบ มีความสัมพันธ์กับระดับผลการเรียนแต่ละกลุ่มได้ดี

**- ด้านการชมสื่อการเรียนรู้ จำนวน 10 เรื่อง**



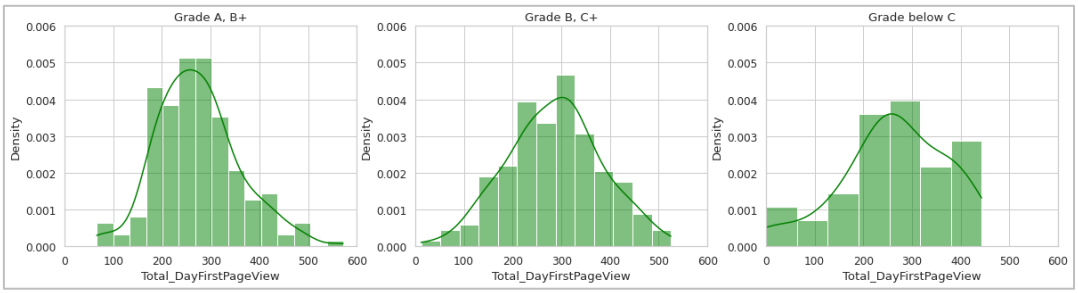

ภาพประกอบ 53 แสดงความสัมพันธ์ระหว่างจำนวนครั้งที่คลิกเข้าถึงสื่อการเรียนรู้รวมทุกเรื่องกับกลุ่มเกรด

จากภาพประกอบ 53 แสดงให้เห็นว่า จำนวนครั้งที่คลิกเข้าถึงสื่อการเรียนรู้รวมทุกเรื่องมากที่สุดในอันดับแรก คือ นิสิตกลุ่มผลการเรียนดี จำนวนครั้งโดยประมาณ 70 ครั้ง ต่อมานิสิตกลุ่มผลการเรียนดีมาก จำนวนครั้งโดยประมาณ 65 ครั้ง สุดท้ายนิสิตกลุ่มผลการเรียนต้องปรับปรุง จำนวนครั้งโดยประมาณ 45 ครั้ง ส่วนจำนวนนิสิตที่คลิกเข้าถึงสื่อการเรียนรู้รวมทุกเรื่องน้อยที่สุดเป็นส่วนใหญ่ในอันดับแรกคือ นิสิตกลุ่มผลการเรียนต้องปรับปรุง รองลงมานิสิตกลุ่มผลการเรียนดี ดีมาก ตามลำดับ จะเห็นได้ว่า จำนวนครั้งที่คลิกเข้าถึงสื่อการเรียนรู้รวมทุกเรื่อง มีความสัมพันธ์กับระดับผลการเรียนแต่ละกลุ่มได้ดี



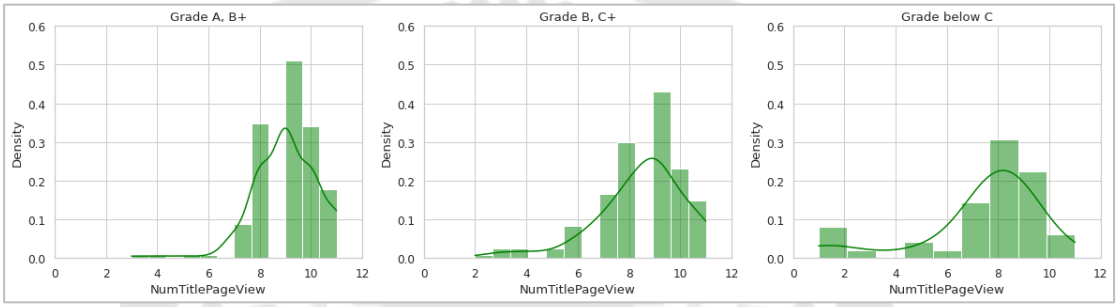

ภาพประกอบ 54 แสดงความสัมพันธ์ระหว่างจำนวนวันที่ชมสื่อการเรียนรู้รวมทุกเรื่องกับกลุ่มเกรด

จากภาพประกอบ 54 แสดงให้เห็นว่า จำนวนวันที่ชมสื่อการเรียนรู้รวมทุกเรื่องมากที่สุดในอันดับแรก คือ นิสิตกลุ่มผลการเรียนดีมาก จำนวนวันโดยประมาณ 42 วัน ต่อมานิสิตกลุ่มผลการเรียนดี จำนวนวันโดยประมาณ 35 วัน สุดท้ายนิสิตกลุ่มผลการเรียนต้องปรับปรุง จำนวนวันโดยประมาณ 25 วัน ในขณะที่จำนวนนิสิตที่ชมสื่อการเรียนรู้รวมทุกเรื่องน้อยที่สุด มีนิสิตเป็นส่วนใหญ่ในอันดับแรกคือ นิสิตกลุ่มผลการเรียนต้องปรับปรุง รองลงมานิสิตกลุ่มผลการเรียนดี ดีมาก ตามลำดับ จะเห็นได้ว่า จำนวนวันที่ชมสื่อการเรียนรู้รวมทุกเรื่อง มีความสัมพันธ์กับระดับผลการเรียนแต่ละกลุ่มได้ดี




ภาพประกอบ 55 แสดงความสัมพันธ์ระหว่างจำนวนวันระหว่างวันแรกที่เรียนและวันแรกที่ชมสื่อการเรียนรู้กับกลุ่มเกรด

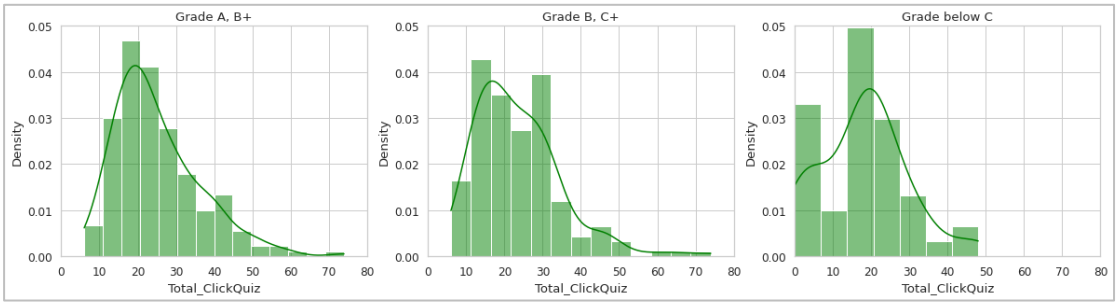
จากภาพประกอบ 55 แสดงให้เห็นว่า จำนวนวันระหว่างวันแรกที่เรียนและวันแรกที่ชมสื่อการเรียนรู้มากที่สุดในอันดับแรก คือ นิสิตกลุ่มผลการเรียนดีมาก จำนวนวันโดยประมาณ 550 วัน ต่อมานิสิตกลุ่มผลการเรียนดี จำนวนวันโดยประมาณ 520 วัน สุดท้ายนิสิตกลุ่มผลการเรียนต้องปรับปรุง จำนวนวันโดยประมาณ 450 วัน ในขณะที่จำนวนวันระหว่างวันแรกที่เรียนและวันแรกที่ชมสื่อการเรียนรู้น้อยที่สุด มีนิสิตเป็นส่วนใหญ่ในอันดับแรกคือ นิสิตกลุ่มผลการเรียนต้องปรับปรุง รองลงมานิสิตกลุ่มผลการเรียนดี ดีมาก ตามลำดับ จะเห็นได้ว่า จำนวนวันระหว่างวันแรกที่เรียนและวันแรกที่ชมสื่อการเรียนรู้ มีความสัมพันธ์กับระดับผลการเรียนแต่ละกลุ่มได้ดี




ภาพประกอบ 56 แสดงความสัมพันธ์ระหว่างจำนวนเรื่องที่ชมสื่อการเรียนรู้กับกลุ่มเกรด

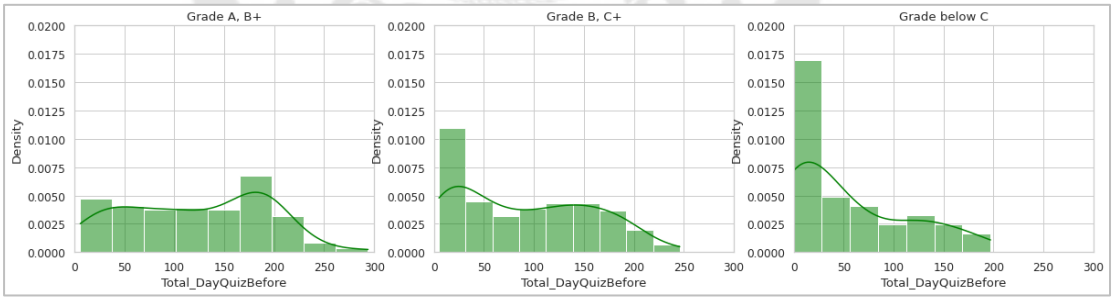
จากภาพประกอบ 56 แสดงให้เห็นว่า จำนวนเรื่องที่ชมสื่อการเรียนรู้มากที่สุดคือนิสิตกลุ่มผลการเรียนดีมาก ดี ต้องปรับปรุง จำนวนเรื่องโดยประมาณ 11 เรื่อง ใกล้เคียงกัน แต่ในขณะที่จำนวนเรื่องที่ชมสื่อการเรียนรู้มากที่สุด มีนิสิตเป็นส่วนใหญ่ในอันดับแรกคือ นิสิตกลุ่มผลการเรียนต้องดีมาก รองลงมานิสิตกลุ่มผลการเรียนดี ต้องปรับปรุง ตามลำดับ ส่วนจำนวนเรื่องที่ชมสื่อการเรียนรู้น้อยที่สุดในอันดับแรก คือ นิสิตกลุ่มผลการเรียนปรับปรุง จำนวนเรื่องโดยประมาณ 1 เรื่อง ต่อมานิสิตกลุ่มผลการเรียนดี จำนวนเรื่องโดยประมาณ 2 เรื่อง สุดท้ายนิสิตกลุ่มผลการเรียนดีมากจำนวนเรื่องโดยประมาณ 3 เรื่อง จะเห็นได้ว่า จำนวนเรื่องที่ชมสื่อการเรียนรู้ มีความสัมพันธ์กับระดับผลการเรียนแต่ละกลุ่มได้ดี

**- ด้านการทำแบบทดสอบ จำนวน 6 เรื่อง**




ภาพประกอบ 57 แสดงความสัมพันธ์ระหว่างจำนวนครั้งที่คลิกเข้าถึงแบบทดสอบรวมทุกเรื่องกับกลุ่มเกรด

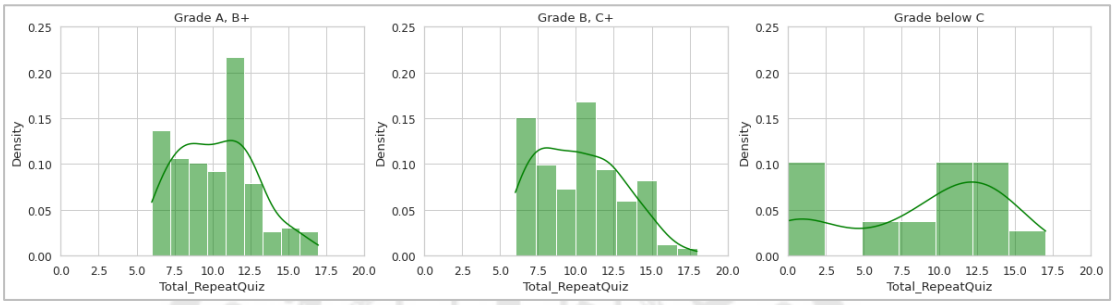
จากภาพประกอบ 57 แสดงให้เห็นว่า จำนวนครั้งที่คลิกเข้าถึงแบบทดสอบรวมทุกเรื่องมากที่สุดในอันดับแรก คือ นิสิตกลุ่มผลการเรียนดีมาก ดี จำนวนครั้งโดยประมาณ 75 ครั้ง ใกล้เคียงกัน ต่อมานิสิตกลุ่มผลการเรียนต้องปรับปรุง จำนวนครั้งโดยประมาณ 48 ครั้ง ส่วนจำนวนครั้งที่คลิกเข้าถึงสื่อการเรียนรู้รวมทุกเรื่องน้อยที่สุดเป็นส่วนใหญ่ในอันดับแรกคือ นิสิตกลุ่มผลการเรียนต้องปรับปรุง รองลงมานิสิตกลุ่มผลการเรียนดี ดีมาก ตามลำดับ จะเห็นได้ว่า จำนวนครั้งที่คลิกเข้าถึงแบบทดสอบรวมทุกเรื่อง มีความสัมพันธ์กับระดับผลการเรียนแต่ละกลุ่มได้ดี




ภาพประกอบ 58 แสดงความสัมพันธ์ระหว่างจำนวนวันระหว่างวันที่ทำแบบทดสอบครั้งแรกและวันสุดท้ายการทำแบบทดสอบกับกลุ่มเกรด

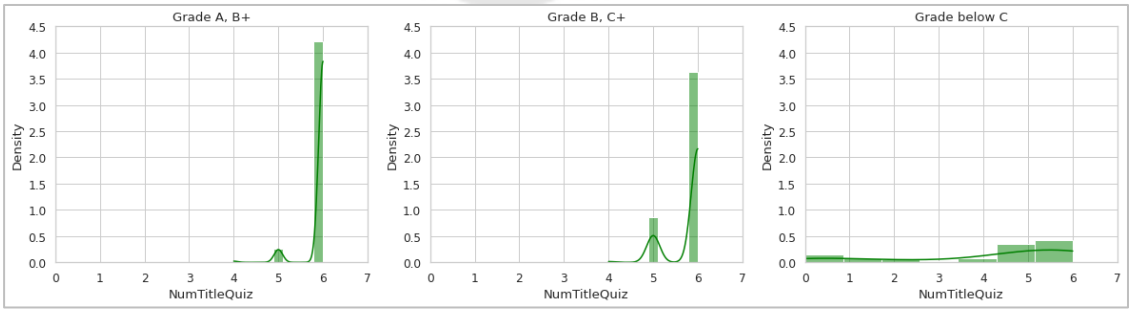
จากภาพประกอบ 58 แสดงให้เห็นว่า จำนวนวันระหว่างวันที่ทำแบบทดสอบครั้งแรกและวันสุดท้ายการทำแบบทดสอบมากที่สุดในอันดับแรก คือ นิสิตกลุ่มผลการเรียนดีมาก จำนวนวันโดยประมาณ 300 วัน รองลงมานิสิตกลุ่มผลการเรียนดี จำนวนวันโดยประมาณ 250 วัน นิสิตกลุ่มผลการเรียนต้องปรับปรุง จำนวนวันโดยประมาณ 200 วัน ตามลำดับ ในขณะที่

จำนวนวันระหว่างวันที่ทำแบบทดสอบครั้งแรกและวันสุดท้ายการทำแบบทดสอบน้อยที่สุด มีนิสิตเป็นส่วนใหญ่ในอันดับแรกคือ นิสิตกลุ่มผลการเรียนต้องปรับปรุง รองลงมานิสิตกลุ่มผลการเรียนดี ดีมาก ตามลำดับ จะเห็นได้ว่า จำนวนวันระหว่างวันที่ทำแบบทดสอบครั้งแรกและวันสุดท้ายการทำแบบทดสอบ มีความสัมพันธ์กับระดับผลการเรียนแต่ละกลุ่มได้ดี จะเห็นได้ว่า จำนวนวันระหว่างวันที่ทำแบบทดสอบครั้งแรกและวันสุดท้ายการทำแบบทดสอบ มีความสัมพันธ์กับระดับผลการเรียนแต่ละกลุ่มได้ดี



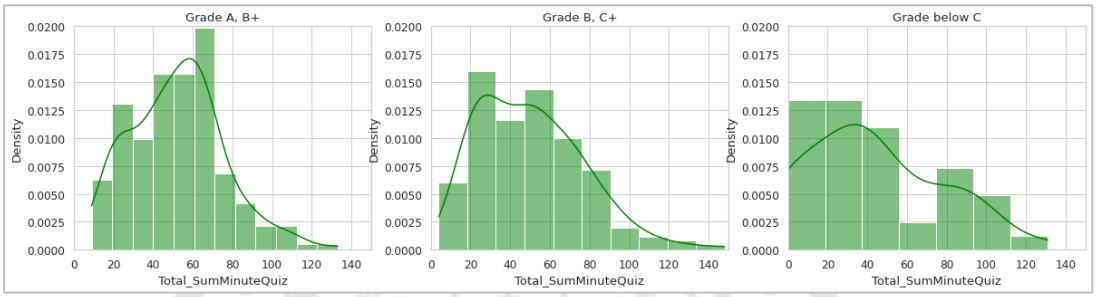

ภาพประกอบ 59 แสดงความสัมพันธ์ระหว่างจำนวนครั้งที่ทำแบบทดสอบซ้ำกับกลุ่มเกรด

จากภาพประกอบ 59 แสดงให้เห็นว่า จำนวนครั้งที่ทำแบบทดสอบซ้ำมากที่สุดในแต่ละกลุ่มผลการเรียนใกล้เคียงกัน คือ นิสิตทุกกลุ่มผลการเรียนดีมาก จำนวนครั้งโดยประมาณ 17 ครั้ง ในขณะที่จำนวนครั้งที่ทำแบบทดสอบซ้ำน้อยที่สุด มีนิสิตเป็นส่วนใหญ่ในอันดับแรกคือ นิสิตกลุ่มผลการเรียนต้องปรับปรุง รองลงมานิสิตกลุ่มผลการเรียนดี ดีมาก ตามลำดับ จะเห็นได้ว่า จำนวนครั้งที่ทำแบบทดสอบซ้ำ มีความสัมพันธ์กับระดับผลการเรียนแต่ละกลุ่มได้ดี



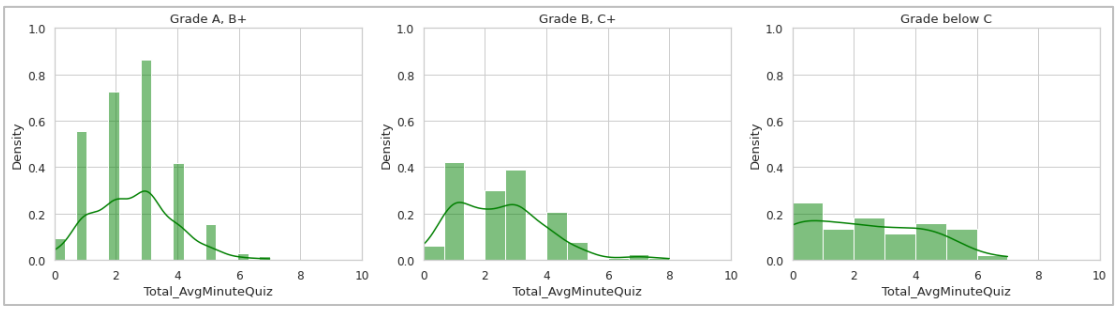

ภาพประกอบ 60 แสดงความสัมพันธ์ระหว่างจำนวนเรื่องที่ทำแบบทดสอบกับกลุ่มเกรด

จากภาพประกอบ 60 แสดงให้เห็นว่า จำนวนเรื่องที่ทำแบบทดสอบมากที่สุดในแต่ละกลุ่มผลการเรียนใกล้เคียงกัน นิสิตทุกกลุ่มผลการเรียน จำนวนเรื่องโดยประมาณ 6 เรื่อง ในขณะที่จำนวนเรื่องที่ทำแบบทดสอบมากที่สุด มีนิสิตเป็นส่วนใหญ่ในอันดับแรกคือ นิสิตกลุ่มผลการเรียนดีมาก รองลงมานิสิตกลุ่มผลการเรียนดี ปรับปรุง ตามลำดับ และพบว่ามีนิสิตกลุ่มผลการเรียนต้องปรับปรุงไม่ได้ทำแบบทดสอบใด ๆ เท่านั้น จะเห็นได้ว่า จำนวนเรื่องที่ทำแบบทดสอบ มีความสัมพันธ์กับระดับผลการเรียนแต่ละกลุ่มได้ดี




ภาพประกอบ 61 แสดงความสัมพันธ์ระหว่างระยะเวลารวมการทำแบบทดสอบรวมทุกครั้ง (นาที) และทุกเรื่อง กับกลุ่มเกรด

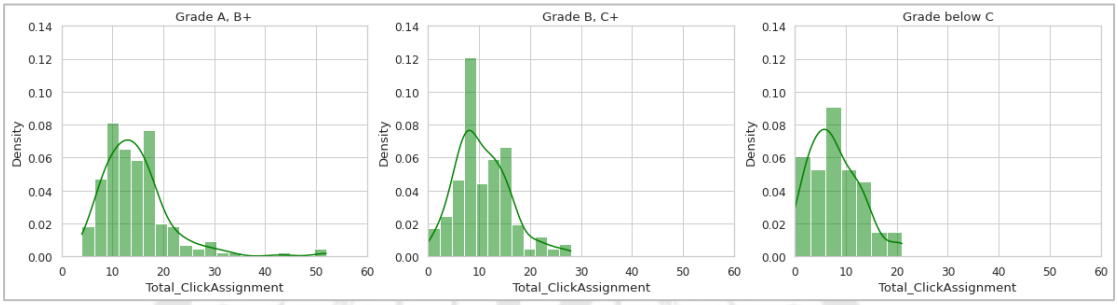
จากภาพประกอบ 61 แสดงให้เห็นว่า ระยะเวลารวมการทำแบบทดสอบรวมทุกครั้งทุกเรื่องมากที่สุดในแต่ละกลุ่มผลการเรียนใกล้เคียงกัน โดยใช้เวลาประมาณ 140 นาที ในขณะที่ระยะเวลารวมการทำแบบทดสอบรวมทุกครั้งทุกเรื่องน้อยที่สุด มีนิสิตเป็นส่วนใหญ่ในอันดับแรกคือ นิสิตกลุ่มผลการเรียนต้องปรับปรุง รองลงมานิสิตกลุ่มผลการเรียนดี ดีมาก ตามลำดับ จะเห็นได้ว่า ระยะเวลารวมการทำแบบทดสอบรวมทุกครั้งทุกเรื่อง มีความสัมพันธ์กับระดับผลการเรียนแต่ละกลุ่มได้ดี




ภาพประกอบ 62 แสดงความสัมพันธ์ระหว่างระยะเวลาเฉลี่ยการทำแบบทดสอบรวมทุกครั้งและทุกเรื่อง (นาที) กับกลุ่มเกรด

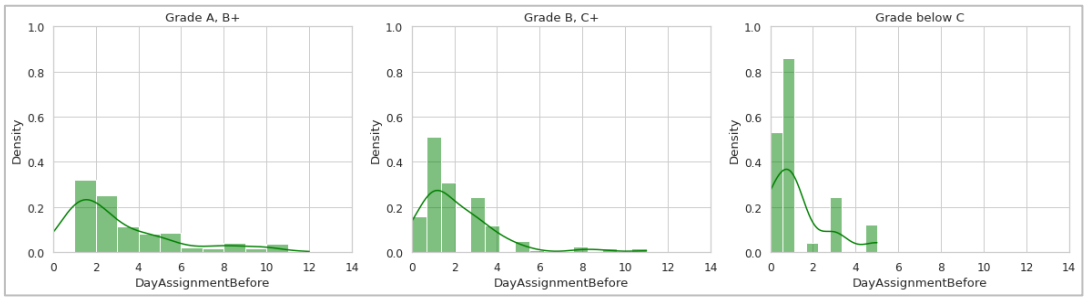
จากภาพประกอบ 62 แสดงให้เห็นว่า ระยะเวลาเฉลี่ยการทำแบบทดสอบรวมทุกครั้งทุกเรื่องมากที่สุดในแต่ละกลุ่มผลการเรียนใกล้เคียงกัน โดยใช้เวลาประมาณ 8 นาที ในขณะที่ระยะเวลาเฉลี่ยการทำแบบทดสอบรวมทุกครั้งทุกเรื่องน้อยที่สุด มีนิสิตเป็นส่วนใหญ่ในอันดับแรก คือ นิสิตกลุ่มผลการเรียนต้องปรับปรุง รองลงมานิสิตกลุ่มผลการเรียนดี ดีมาก ตามลำดับ จะเห็นได้ว่า ระยะเวลาเฉลี่ยการทำแบบทดสอบรวมทุกครั้งทุกเรื่อง มีความสัมพันธ์กับระดับผลการเรียนแต่ละกลุ่มได้ดี

**- ด้านการส่งงานมอบหมาย จำนวน 1 เรื่อง**




ภาพประกอบ 63 แสดงความสัมพันธ์ระหว่างจำนวนครั้งที่คลิกเข้าถึงงานมอบหมายกับกลุ่มเกรด

จากภาพประกอบ 63 แสดงให้เห็นว่า จำนวนครั้งที่คลิกเข้าถึงงานมอบหมายมากที่สุดในอันดับแรก คือ นิสิตกลุ่มผลการเรียนดีมาก จำนวนครั้งโดยประมาณ 50 ครั้ง รองลงมานิสิตกลุ่มผลการเรียนดี จำนวนครั้งโดยประมาณ 30 ครั้ง นิสิตกลุ่มผลการเรียนต้องปรับปรุง จำนวนครั้งโดยประมาณ 20 ครั้ง ตามลำดับ ในขณะที่จำนวนที่คลิกเข้าสู่ระบบน้อยที่สุดเป็นส่วนใหญ่ในอันดับแรกคือ นิสิตกลุ่มผลการเรียนต้องปรับปรุง รองลงมานิสิตกลุ่มผลการเรียนดี ดีมาก ตามลำดับ จะเห็นได้ว่า จำนวนครั้งที่คลิกเข้าถึงงานมอบหมาย มีความสัมพันธ์กับระดับผลการเรียนแต่ละกลุ่มได้ดี

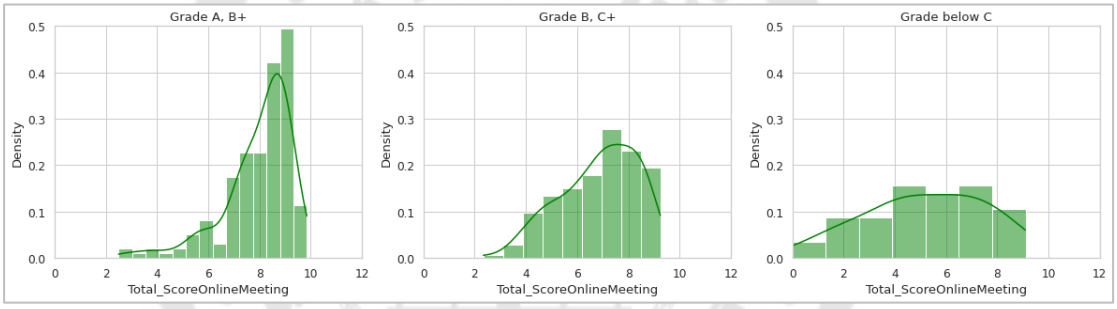



ภาพประกอบ 64 แสดงความสัมพันธ์ระหว่างจำนวนวันระหว่างวันแรกที่ส่งงานมอบหมายและวันสุดท้ายการส่งงานมอบหมายกับกลุ่มเกรด

จากภาพประกอบ 64 แสดงให้เห็นภาพรวมว่า จำนวนวันระหว่างวันแรกที่ส่งงานมอบหมายและวันสุดท้ายการส่งงานมอบหมายมากที่สุดเป็นอันดับแรก คือ นิสิตกลุ่มผลการเรียนดีมาก ส่งก่อนวันสุดท้ายที่ต้องส่งเร็วที่สุดโดยประมาณ 12 วัน รองลงมาเป็นนิสิตกลุ่มผลการเรียนดี ส่งก่อนวันสุดท้ายที่ต้องส่งเร็วที่สุดโดยประมาณ 11 วัน นิสิตกลุ่มผลการเรียนต้องปรับปรุง ส่งก่อนวันสุดท้ายที่ต้องส่งเร็วที่สุดโดยประมาณ 5 วัน ในขณะที่นิสิตส่งงานในวันสุดท้ายที่ต้องส่งช้าที่สุดเป็นอันดับแรก คือ นิสิตกลุ่มผลการเรียนต้องปรับปรุง รองลงมานิสิตกลุ่มผลการเรียนดี ดีมาก ตามลำดับ จะเห็นได้ว่า จำนวนวันระหว่างวันแรกที่ส่งงานมอบหมายและวันสุดท้ายการส่งงานมอบหมาย มีความสัมพันธ์กับระดับผลการเรียนแต่ละกลุ่มได้ดี

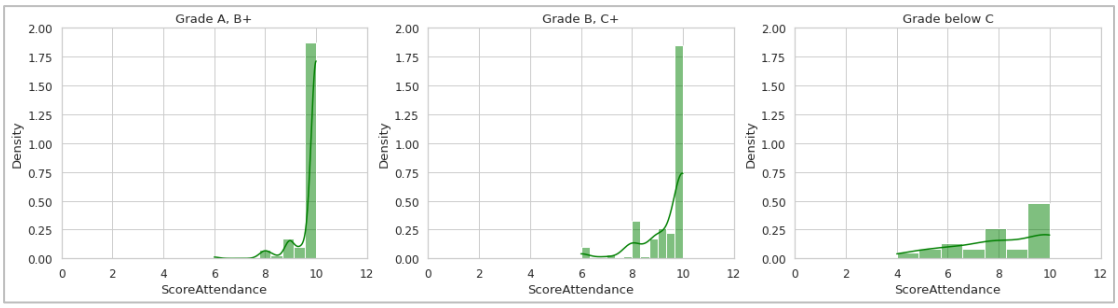
**4. คุณลักษณะที่เกี่ยวข้องกับคะแนนการประเมินผลการเรียนรู้**

**- ด้านคะแนนกิจกรรมในห้องเรียน (30%) จำนวน 3 เรื่อง**



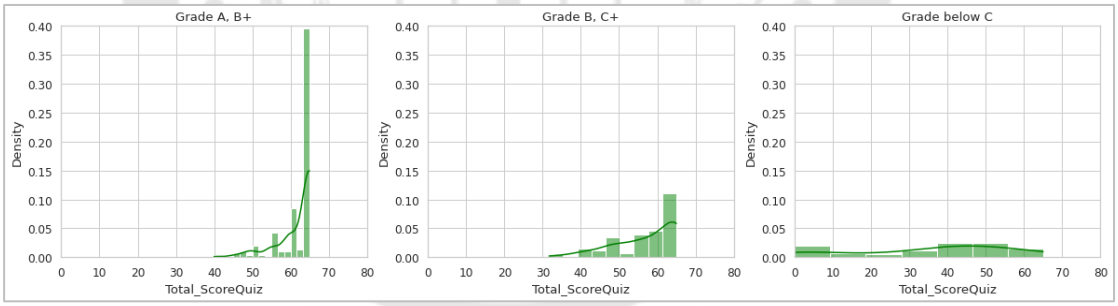

ภาพประกอบ 65 แสดงความสัมพันธ์ระหว่างคะแนนกิจกรรมในห้องเรียนออนไลน์รวมทุกเรื่องกับกลุ่มเกรด

จากภาพประกอบ 65 แสดงให้เห็นว่า คะแนนกิจกรรมในห้องเรียนออนไลน์รวมทุกเรื่องมากที่สุดของแต่ละกลุ่มผลการเรียนใกล้เคียงกัน คือ นิสิตทุกกลุ่มผลการเรียนได้ประมาณ 10 คะแนน ในขณะที่นิสิตได้คะแนนมากที่สุดเป็นส่วนใหญ่คือ นิสิตกลุ่มผลการเรียนดีมาก รองลงมาผลการเรียนดี ต้องปรับปรุง ส่วนนิสิตที่ได้คะแนนน้อยที่สุดเป็นส่วนใหญ่ในอันดับแรกคือ นิสิตกลุ่มผลการเรียนต้องปรับปรุง รองลงมานิสิตกลุ่มผลการเรียนดี ดีมาก ตามลำดับ จะเห็นได้ว่า คะแนนกิจกรรมในห้องเรียนออนไลน์รวมทุกเรื่อง มีความสัมพันธ์กับระดับผลการเรียนแต่ละกลุ่มได้ดี




ภาพประกอบ 66 แสดงความสัมพันธ์ระหว่างคะแนนการเข้าเรียนกับกลุ่มเกรด

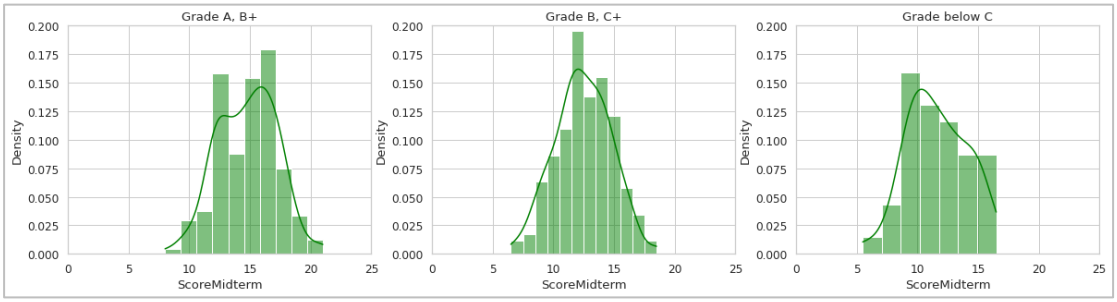
จากภาพประกอบ 66 แสดงให้เห็นว่า คะแนนการเข้าเรียนมากที่สุดของแต่ละกลุ่มผลการเรียนใกล้เคียงกัน คือ นิสิตทุกกลุ่มผลการเรียนได้ประมาณ 10 คะแนน ในขณะที่นิสิตได้คะแนนมากที่สุดเป็นส่วนใหญ่คือ นิสิตกลุ่มผลการเรียนดีมาก รองลงมาผลการเรียนดี ต้องปรับปรุง ส่วนนิสิตที่ได้คะแนนน้อยที่สุดเป็นส่วนใหญ่ในอันดับแรกคือ นิสิตกลุ่มผลการเรียนต้องปรับปรุง รองลงมานิสิตกลุ่มผลการเรียนดี ดีมาก ตามลำดับ จะเห็นได้ว่า คะแนนการเข้าเรียน มีความสัมพันธ์กับระดับผลการเรียนแต่ละกลุ่มได้ดี




ภาพประกอบ 67 แสดงความสัมพันธ์ระหว่างคะแนนแบบทดสอบรวมทุกเรื่องกับกลุ่มเกรด

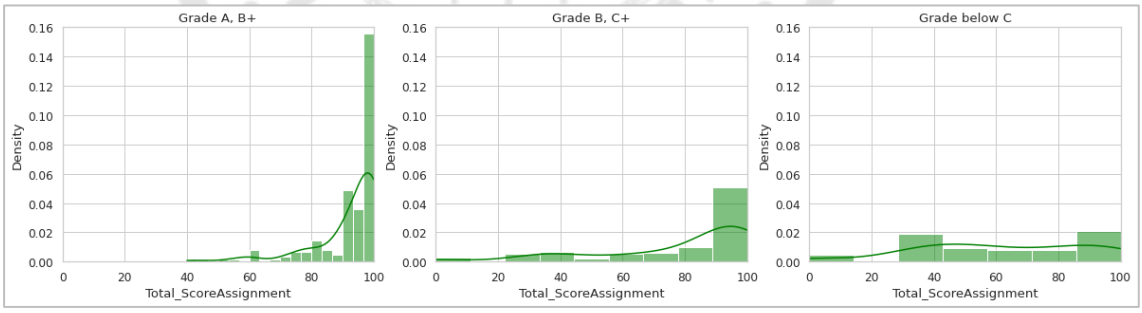
จากภาพประกอบ 67 แสดงให้เห็นว่า คะแนนแบบทดสอบรวมทุกเรื่องมากที่สุดของแต่ละกลุ่มผลการเรียนใกล้เคียงกัน คือ นิสิตทุกกลุ่มผลการเรียนได้ประมาณ 65 คะแนน ในขณะที่นิสิตได้คะแนนมากที่สุดเป็นส่วนใหญ่คือ นิสิตกลุ่มผลการเรียนดีมาก รองลงมาผลการเรียนดี ต้องปรับปรุง ส่วนนิสิตที่ได้คะแนนน้อยที่สุดเป็นส่วนใหญ่ในอันดับแรกคือ นิสิตกลุ่มผลการเรียนต้องปรับปรุง รองลงมานิสิตกลุ่มผลการเรียนดี ดีมาก ตามลำดับ จะเห็นได้ว่า คะแนนแบบทดสอบรวมทุกเรื่อง มีความสัมพันธ์กับระดับผลการเรียนแต่ละกลุ่มได้ดี

- **ด้านคะแนนกลางภาค** (30%)




ภาพประกอบ 68 แสดงความสัมพันธ์ระหว่างคะแนนสอบกลางภาคกับกลุ่มเกรด

จากภาพประกอบ 68 แสดงให้เห็นว่า คะแนนกลางภาคมากที่สุดในอันดับแรก คือ นิสิตกลุ่มผลการเรียนดีมากได้ประมาณ 20 คะแนน รองลงมาเป็นนิสิตผลการเรียนดีได้ประมาณ 18 คะแนน ผลการเรียนต้องปรับปรุงได้ประมาณ 16 คะแนน ตามลำดับ ในขณะที่นิสิตได้คะแนนน้อยที่สุดเป็นส่วนใหญ่ในอันดับแรกคือ นิสิตกลุ่มผลการเรียนต้องปรับปรุง รองลงมา นิสิตกลุ่มผลการเรียนดี ดีมาก ตามลำดับ จะเห็นได้ว่า คะแนนกลางภาค มีความสัมพันธ์กับระดับผลการเรียนแต่ละกลุ่มได้ดี




ภาพประกอบ 69 แสดงความสัมพันธ์ระหว่างคะแนนงานมอบหมายรวมทุกส่วนกับกลุ่มเกรด

จากภาพประกอบ 69 แสดงให้เห็นว่า คะแนนงานมอบหมายรวมทุกส่วนมากที่สุดของแต่ละกลุ่มผลการเรียนใกล้เคียงกัน คือ นิสิตทุกกลุ่มผลการเรียนได้ประมาณ 100 คะแนน ในขณะที่นิสิตได้คะแนนมากที่สุดเป็นส่วนใหญ่คือ นิสิตกลุ่มผลการเรียนดีมาก รองลงมาผลการเรียนดี ต้องปรับปรุง ตามลำดับ ส่วนนิสิตที่ได้คะแนนน้อยที่สุดเป็นส่วนใหญ่ในอันดับแรกคือ นิสิต

กลุ่มผลการเรียนต้องปรับปรุง รองลงมานิสิตกลุ่มผลการเรียนดี ดีมาก ตามลำดับ จะเห็นได้ว่าคะแนนงานมอบหมายรวมทุกส่วน มีความสัมพันธ์กับระดับผลการเรียนแต่ละกลุ่มได้ดี

**การจัดการความไม่สมดุลของข้อมูล**

ก่อนจัดการข้อมูลที่สมดุลให้ทำการแบ่งชุดข้อมูลออกเป็น 2 ส่วน คือ ข้อมูลการเรียนรู้ (80%) จำนวน 322 คน และข้อมูลการทดสอบ (20%) จำนวน 81 คน นำข้อมูลการเรียนรู้ที่นิสิตได้ผลการเรียนดีมาก 147 คน, ผลการเรียนดี 141 คน และผลการเรียนต้องปรับปรุง 34 คน จะเห็นได้ว่าชุดข้อมูลนี้มีลักษณะของข้อมูลที่ไม่สมดุลกัน อาจจะส่งผลต่อประสิทธิภาพการทำนายได้ไม่ค่อยดี เมื่อนำข้อมูลใหม่เข้ามา จะทำนายได้คลาสที่มีจำนวนมากคือ ระดับผลการเรียนดีมากและดี ดังนั้น ต้องปรับความสมดุลของข้อมูล โดยให้ระดับผลการเรียนทุกระดับมีจำนวนนิสิตเท่า ๆ กัน งานวิจัยนี้ใช้เทคนิคการสุ่มตัวอย่างเพิ่มกลุ่มน้อยด้วยการสร้างตัวอย่างสังเคราะห์ (Synthetic Minority Oversampling Technique: SMOTE) ในการจัดความไม่สมดุลของข้อมูล ดังภาพประกอบ 70 – 71

ภาพประกอบ 70 แสดงจำนวนนิสิตแต่ละระดับผลการเรียน ก่อนการทำ SMOTE

ภาพประกอบ 71 แสดงจำนวนนิสิตแต่ละระดับผลการเรียน หลังการทำ SMOTE

**การสร้างแบบจำลองการจำแนกประเภท**

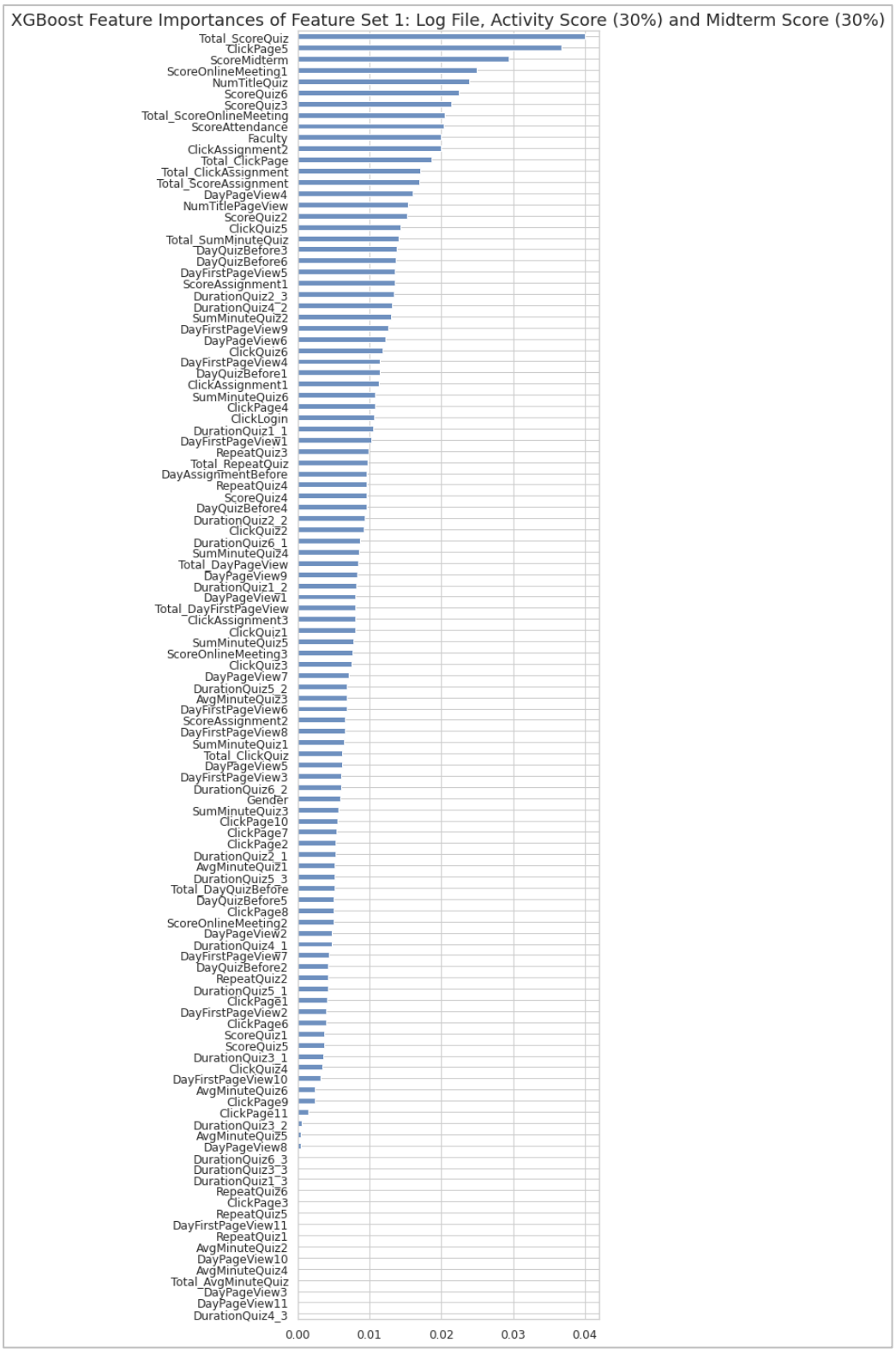
การจะสร้างแบบจำลองการทำนายระดับผลการเรียน โดยใช้ชุดข้อมูลที่มีตัวแปรเชิงหมวดหมู่และตัวแปรเชิงตัวเลข จึงทำการแปลงข้อมูลข้อความเป็นข้อมูลตัวเลข (Label Encoding) และปรับขนาดข้อมูลให้เป็นมาตรฐาน โดยใช้วิธี Standardization ในการปรับช่วงข้อมูลให้มีค่าเฉลี่ยเท่ากับศูนย์ และค่าเบี่ยงเบนมาตรฐานเท่ากับ 1 แล้ว นำชุดข้อมูลการเรียนรู้จำนวน 4 ชุด ที่ได้จากการทำวิศวกรรมคุณลักษณะ ไปเรียนรู้ข้อมูลแต่ละชุด จากนั้นทำนายระดับผลการเรียนด้วยแบบจำลองการจำแนกประเภท 7 แบบ ได้แก่ Logistic Regression, GaussianNB, K-Nearest Neighbor, Support Vector Machines, Decision Tree, Random Forest และ XGBoost

**การคัดเลือกคุณลักษณะ**

สำหรับงานวิจัยนี้ใช้แบบจำลอง XGBoost ในการวัดคะแนนความสำคัญของคุณลักษณะ (Feature Importance) เพื่อคัดเลือกกลุ่มคุณลักษณะที่มีความสัมพันธ์กับระดับผลการเรียนด้วยการเรียนรู้ โดยมีหลักการนำกลุ่มคุณลักษณะที่แตกต่างกันไปเรียนรู้และทำนายข้อมูลด้วยแบบจำลอง XGBoost หากเกิดความผิดพลาดการทำนาย แบบจำลอง XGBoost จะนำความผิดพลาดที่เกิดขึ้นไปเรียนรู้ข้อมูลใหม่เรื่อย ๆ จนกว่าจะไม่เกิดความผิดพลาด แล้วคัดเลือกกลุ่มคุณลักษณะที่ทำให้ได้ผลลัพธ์การทำนายดีที่สุด ยังส่งผลต่อให้จำนวนคุณลักษณะลดลงตลอดจนนำไปสู่การลดระยะเวลาการเรียนรู้ข้อมูล และช่วยเพิ่มประสิทธิภาพการทำนาย ทั้งนี้ คะแนนความสำคัญของคุณลักษณะมากที่สุดจำนวน 10 คุณลักษณะ โดยเรียงลำดับคะแนนความสำคัญของชุดคุณลักษณะที่แตกต่างกันจากชุดข้อมูล 4 ชุด ดังตาราง 9 - 12 และแสดงความสัมพันธ์ระหว่างชุดคุณลักษณะ 4 ชุด และระดับผลการเรียน ดังภาพประกอบ 72 - 75

ตาราง 9 แสดงการเรียงลำดับคะแนนความสำคัญของชุดข้อมูลสำหรับกรณีที่ 1: ข้อมูลบันทึกการใช้งานระบบ, คะแนนกิจกรรมในห้องเรียนออนไลน์ (30%) และคะแนนกลางภาค (30%) ที่คำนวณจากแบบจำลอง XGBoost

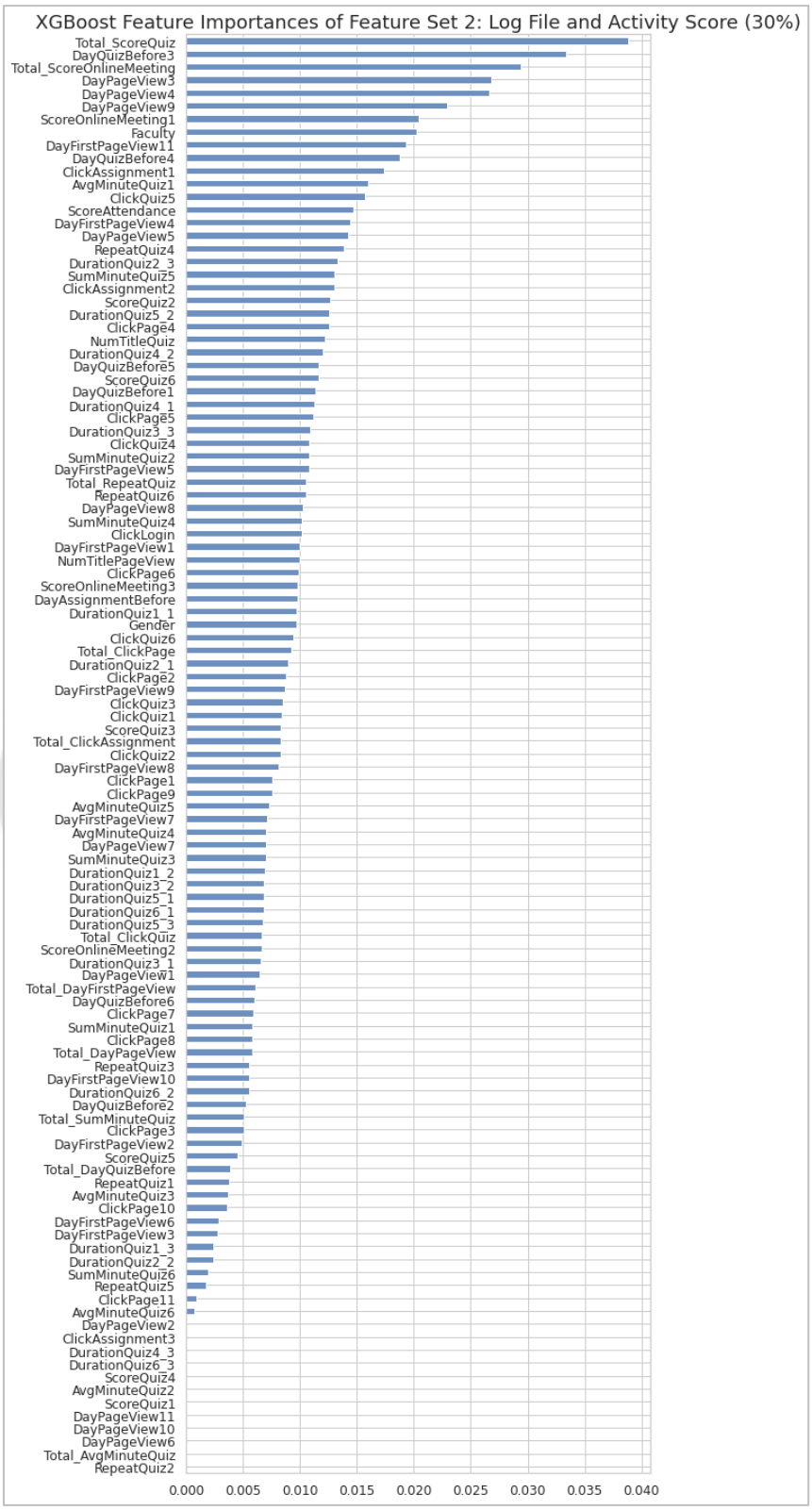
| อันดับ | คุณลักษณะ | ความหมายของคุณลักษณะ | คะแนนความสำคัญ |
|---|---|---|---|
| 1 | Total_ScoreQuiz | คะแนนแบบทดสอบรวมทุกเรื่อง | 0.0400 |
| 2 | ClickPage5 | จำนวนครั้งที่คลิกเข้าถึงสื่อการเรียนรู้ เรื่องที่ 5 จริยธรรมในการใช้เทคโนโลยีสารสนเทศและการสื่อสาร | 0.0368 |
| 3 | ScoreMidterm | คะแนนการสอบกลางภาค | 0.0295 |
| 4 | ScoreOnlineMeeting1 | คะแนนกิจกรรมในห้องเรียนออนไลน์ เรื่องที่ 1 แหล่งสารสนเทศและการสืบค้น | 0.0249 |
| 5 | NumTitleQuiz | จำนวนเรื่องที่ทำแบบทดสอบ | 0.0239 |
| 6 | ScoreQuiz6 | คะแนนแบบทดสอบ เรื่องที่ 6 ภัยอันตรายในโลกยุคดิจิทัลและแนวทางการป้องกัน | 0.0225 |
| 7 | ScoreQuiz3 | คะแนนแบบทดสอบ เรื่องที่ 3 การประเมินสารสนเทศ | 0.0214 |
| 8 | Total_ScoreOnlineMeeting | คะแนนกิจกรรมในห้องเรียนออนไลน์รวมทุกเรื่อง | 0.0205 |
| 9 | ScoreAttendance | คะแนนการเข้าเรียน | 0.0204 |
| 10 | Faculty | คณะที่นิสิตเรียน | 0.0200 |




ภาพประกอบ 72 แสดงความสัมพันธ์ระหว่างชุดข้อมูลสำหรับกรณีที่ 1 และระดับผลการเรียน โดยวัดคะแนนความสำคัญจากแบบจำลอง XGBoost

ตาราง 10 แสดงการเรียงลำดับคะแนนความสำคัญของชุดข้อมูลสำหรับกรณีที่ 2: ข้อมูลบันทึกการใช้งานระบบ และคะแนนกิจกรรมในห้องเรียนออนไลน์ (30%) ที่คำนวณจากแบบจำลอง XGBoost

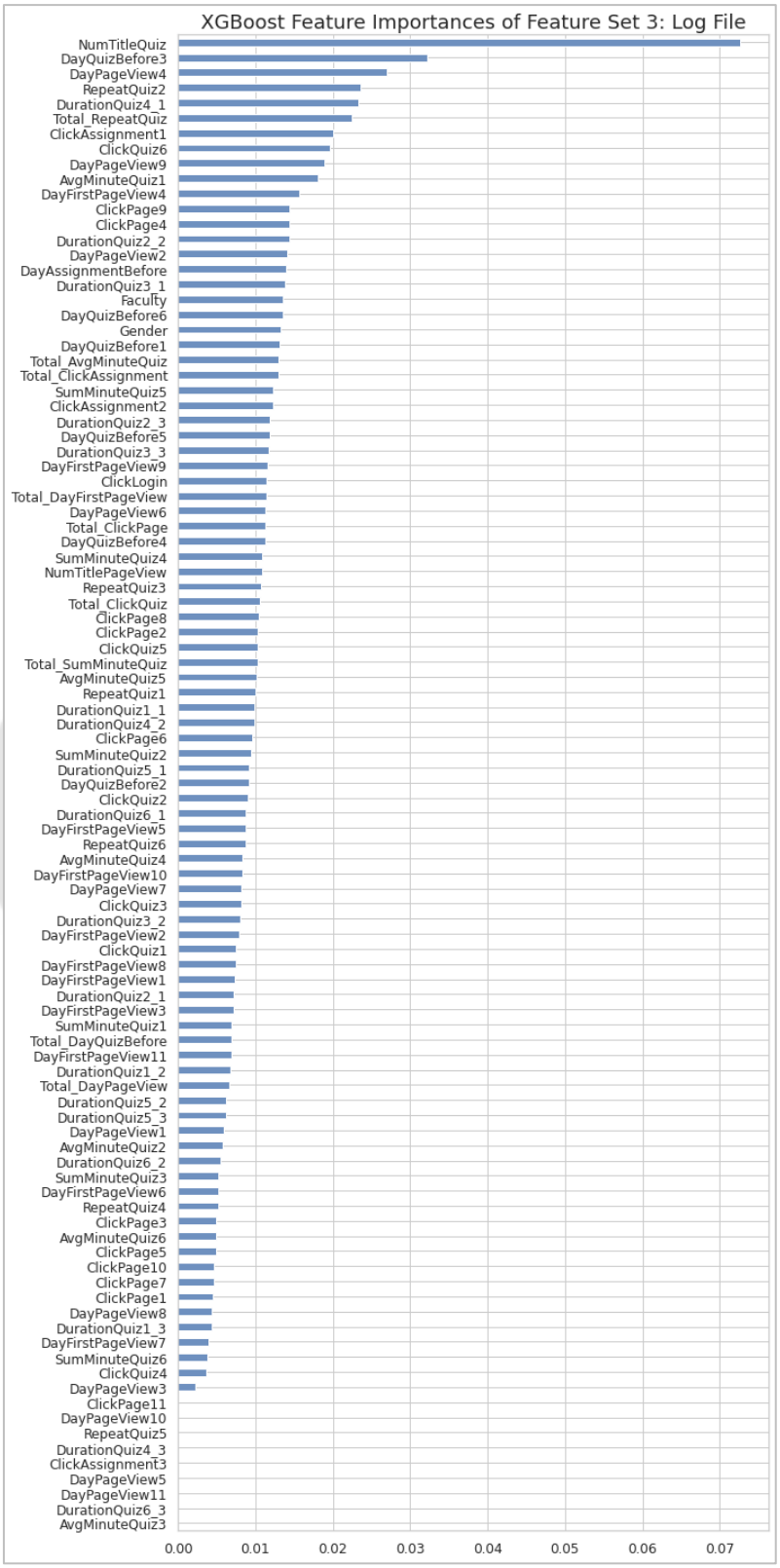
| อันดับ | คุณลักษณะ | ความหมายของคุณลักษณะ | คะแนนความสำคัญ |
|---|---|---|---|
| 1 | Total_ScoreQuiz | คะแนนแบบทดสอบรวมทุกเรื่อง | 0.0388 |
| 2 | DayQuizBefore3 | จำนวนวันระหว่างวันที่ทำแบบทดสอบครั้งแรกเรื่องที่ 3 การประเมินสารสนเทศ และวันสุดท้ายการทำแบบทดสอบ | 0.0334 |
| 3 | Total_ScoreOnlineMeeting | คะแนนกิจกรรมในห้องเรียนออนไลน์รวมทุกเรื่อง | 0.0294 |
| 4 | DayPageView3 | จำนวนวันที่ชมสื่อการเรียนรู้ เรื่องที่ 3 การประเมินสารสนเทศ | 0.0268 |
| 5 | DayPageView4 | จำนวนวันที่ชมสื่อการเรียนรู้ เรื่องที่ 4 การวิเคราะห์สังเคราะห์สารสนเทศ | 0.0266 |
| 6 | DayPageView9 | จำนวนวันที่ชมสื่อการเรียนรู้ เรื่องที่ 9 ภัยอันตรายในโลกยุคดิจิทัลและแนวทางการป้องกัน | 0.0229 |
| 7 | ScoreOnlineMeeting1 | คะแนนกิจกรรมในห้องเรียนออนไลน์ เรื่องที่ 1 แหล่งสารสนเทศและการสืบค้น | 0.0204 |
| 8 | Faculty | คณะที่นิสิตเรียน | 0.0203 |
| 9 | DayFirstPageView11 | จำนวนวันระหว่างวันแรกที่ที่เรียน และวันแรกที่ชมสื่อการเรียนรู้ เรื่องที่ 11 เทคนิคการจัดการข้อมูลอย่างมีระบบในรูปแบบรายงานด้วยโปรแกรม MS Word และ MS Excel | 0.0193 |
| 10 | DayQuizBefore4 | จำนวนวันระหว่างวันที่ทำแบบทดสอบครั้งแรกเรื่องที่ 4 จริยธรรมในการใช้เทคโนโลยีสารสนเทศและการสื่อสาร และวันสุดท้ายการทำแบบทดสอบ | 0.0188 |




ภาพประกอบ 73 แสดงความสัมพันธ์ระหว่างชุดข้อมูลสำหรับกรณีที่ 2 และระดับผลการเรียน โดยวัดคะแนนความสำคัญจากแบบจำลอง XGBoost

ตาราง 11 แสดงการเรียงลำดับคะแนนความสำคัญของชุดข้อมูลในกรณีที่ 3: ข้อมูลบันทึกการใช้งานระบบ ที่คำนวณจากแบบจำลอง XGBoost

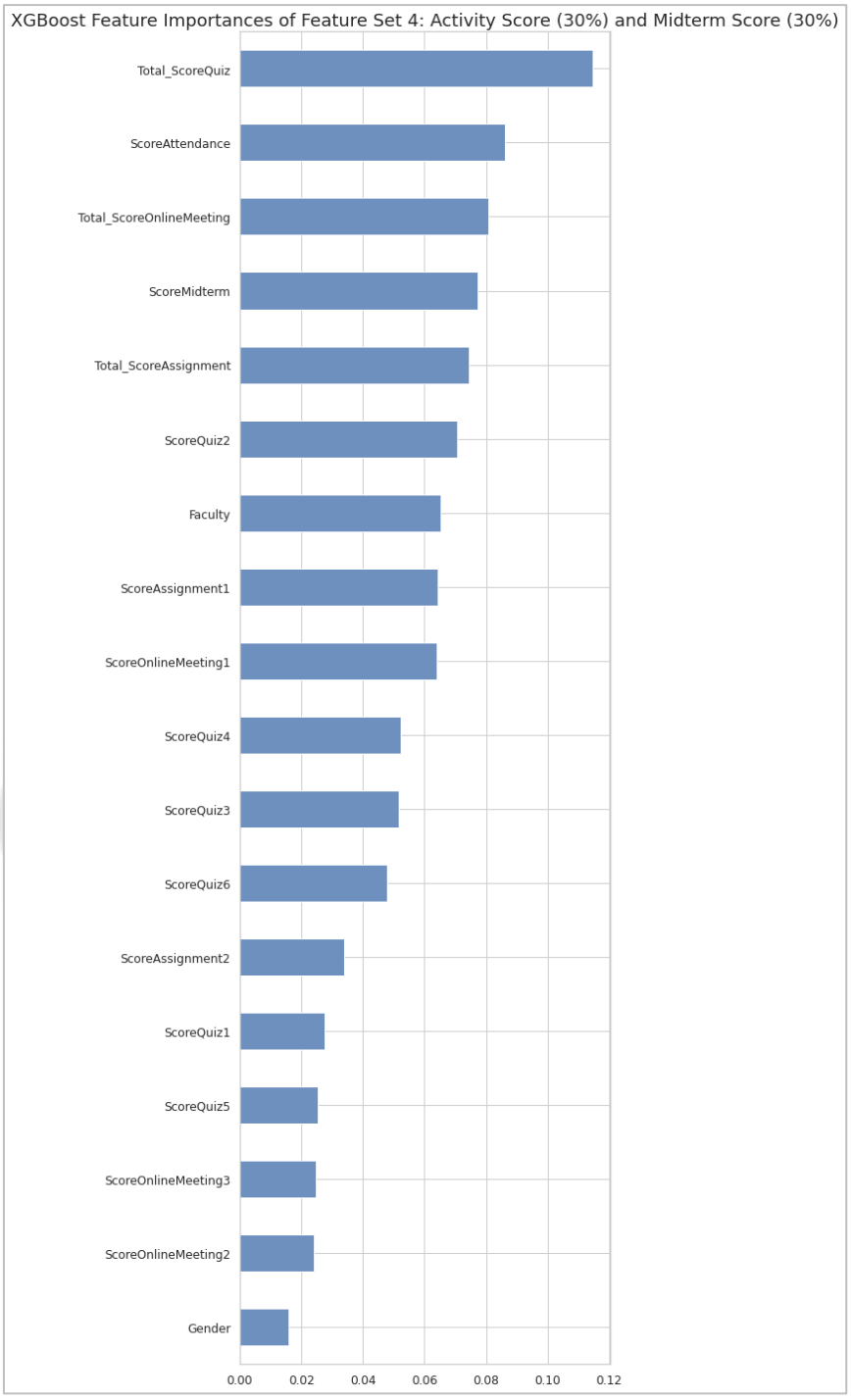
| อันดับ | คุณลักษณะ | ความหมายของคุณลักษณะ | คะแนนความสำคัญ |
|---|---|---|---|
| 1 | NumTitleQuiz | จำนวนเรื่องที่ทำแบบทดสอบ | 0.0726 |
| 2 | DayQuizBefore3 | จำนวนวันระหว่างวันที่ทำแบบทดสอบครั้งแรก เรื่องที่ 3 การประเมินสารสนเทศ และวันสุดท้ายการทำแบบทดสอบ | 0.0322 |
| 3 | DayPageView4 | จำนวนวันที่ชมสื่อการเรียนรู้ เรื่องที่ 4 การวิเคราะห์สังเคราะห์สารสนเทศ | 0.0269 |
| 4 | RepeatQuiz2 | จำนวนครั้งที่ทำแบบทดสอบซ้ำ เรื่องที่ 2 แหล่งสารสนเทศและการสืบค้น | 0.0237 |
| 5 | DurationQuiz4_1 | ระยะเวลาการทำแบบทดสอบครั้งแรก เรื่องที่ 4 จริยธรรมในการใช้เทคโนโลยีสารสนเทศและการสื่อสาร (นาที) | 0.0232 |
| 6 | Total_RepeatQuiz | จำนวนครั้งที่ทำแบบทดสอบซ้ำรวมทุกเรื่อง | 0.0225 |
| 7 | ClickAssignment1 | จำนวนครั้งที่คลิกเข้าถึงสื่อการเรียนรู้ เรื่องที่ 1 เทคโนโลยีสารสนเทศและการสื่อสารที่น่าสนใจในปัจจุบันและแนวโน้มในอนาคต | 0.0201 |
| 8 | ClickQuiz6 | จำนวนครั้งที่คลิกเข้าถึงแบบทดสอบ เรื่องที่ 6 ภัยอันตรายในโลกยุคดิจิทัลและแนวทางการป้องกัน | 0.0196 |
| 9 | DayPageView9 | จำนวนวันที่ชมสื่อการเรียนรู้ เรื่องที่ 9 ภัยอันตรายในโลกยุคดิจิทัลและแนวทางการป้องกัน | 0.0189 |
| 10 | AvgMinuteQuiz1 | ระยะเวลาเฉลี่ยการทำแบบทดสอบรวมทุกครั้ง เรื่องที่ 1 พลเมืองดิจิทัลที่รู้เท่าทันสื่อ สารสนเทศและเทคโนโลยี (นาที) | 0.0181 |




ภาพประกอบ 74 แสดงความสัมพันธ์ระหว่างชุดข้อมูลสำหรับกรณีที่ 3 และระดับผลการเรียน โดยวัดคะแนนความสำคัญจากแบบจำลอง XGBoost

ตาราง 12 แสดงการเรียงลำดับคะแนนความสำคัญของชุดข้อมูลในกรณีที่ 4: คะแนนกิจกรรมในห้องเรียนออนไลน์ และคะแนนสอบกลางภาค ที่คำนวณจากแบบจำลอง XGBoost

| อันดับ | คุณลักษณะ | ความหมายของคุณลักษณะ | คะแนนความสำคัญ |
|---|---|---|---|
| 1 | Total_ScoreQuiz | คะแนนแบบทดสอบรวมทุกเรื่อง | 0.1146 |
| 2 | ScoreAttendance | คะแนนการเข้าเรียน | 0.0861 |
| 3 | Total_ScoreOnlineMeeting | คะแนนกิจกรรมในห้องเรียนออนไลน์รวมทุกเรื่อง | 0.0808 |
| 4 | ScoreMidterm | คะแนนสอบกลางภาค | 0.0772 |
| 5 | Total_ScoreAssignment | คะแนนงานมอบหมาย เรื่อง เทคโนโลยีสารสนเทศและการสื่อสารที่น่าสนใจในปัจจุบันและแนวโน้มในอนาคต รวมทุกส่วน | 0.0744 |
| 6 | ScoreQuiz2 | คะแนนแบบทดสอบ เรื่องที่ 2 แหล่งสารสนเทศและการสืบค้น | 0.0707 |
| 7 | Faculty | คณะที่นิสิตเรียน | 0.0652 |
| 8 | ScoreAssignment1 | คะแนนงานมอบหมาย เรื่อง เทคโนโลยีสารสนเทศและการสื่อสารที่น่าสนใจในปัจจุบันและแนวโน้มในอนาคต ส่วนที่ 1 | 0.0642 |
| 9 | ScoreOnlineMeeting1 | คะแนนกิจกรรมในห้องเรียนออนไลน์ เรื่องที่ 1 แหล่งสารสนเทศและการสืบค้น | 0.0640 |
| 10 | ScoreQuiz4 | คะแนนแบบทดสอบ เรื่องที่ 4 จริยธรรมในการใช้เทคโนโลยีสารสนเทศและการสื่อสาร | 0.0522 |




ภาพประกอบ 75 แสดงความสัมพันธ์ระหว่างชุดข้อมูลสำหรับกรณีที่ 4 และระดับผลการเรียน โดยวัดคะแนนความสำคัญจากแบบจำลอง XGBoost

**การวัดประสิทธิภาพของแบบจำลอง**

นำชุดทดสอบจำนวน 4 ชุด ที่ได้ปรับความไม่สมดุลของข้อมูลด้วยวิธี SMOTE แล้ว คัดเลือกคุณลักษณะที่ใช้แบบจำลอง XGBoost ในการวัดคะแนนความสำคัญของคุณลักษณะ แล้วเข้าสู่แบบจำลองการจำแนกประเภท 7 แบบ ที่ผ่านการเรียนรู้ของชุดข้อมูลการเรียนรู้จำนวน 4 ชุด ได้แก่ Logistic Regression, GaussianNB, K-Nearest Neighbor, Support Vector Machines, Decision Tree, Random Forest และ XGBoost ไปเปรียบเทียบประสิทธิภาพของแบบจำลองของชุดข้อมูลที่ใช้คุณลักษณะทั้งหมด และชุดข้อมูลที่คัดเลือกคุณลักษณะ โดยพิจารณาค่าความถูกต้อง (Accuracy), ความแม่นยำ (Precision), ความระลึก (Recall) จากนั้นคัดเลือกแบบจำลองการทำนายผลการเรียนที่ให้ประสิทธิภาพของแบบจำลองที่ดีที่สุดจากชุดข้อมูลทดสอบจำนวน 4 ชุด ที่ใช้คุณลักษณะทั้งหมด และคัดเลือกคุณลักษณะ

# บทที่ 4
# ผลการดำเนินงานวิจัย

ในการวิจัยเพื่อศึกษาเกี่ยวกับการทำนายผลการเรียนของนิสิตที่ใช้งานระบบการจัดการเรียนรู้ออนไลน์ โดยใช้การเรียนรู้ของเครื่อง ผู้วิจัยได้ดำเนินการวิจัย ตั้งแต่การทบทวนวรรณกรรมที่เกี่ยวข้อง การทำความเข้าใจในธุรกิจ การทำความเข้าใจข้อมูล การเตรียมข้อมูล การสำรวจข้อมูล การสร้างแบบจำลอง ตลอดจนถึงการวัดประสิทธิภาพของแบบจำลอง เพื่อให้บรรลุผลสำเร็จตามวัตถุประสงค์ที่ได้กำหนดไว้ ดังนี้

1. ผลลัพธ์จากการเปรียบเทียบระหว่างแบบจำลองการทำนายผลการเรียนที่ใช้ชุดข้อมูลที่ 1 ข้อมูลบันทึกการใช้งาน คะแนนกิจกรรมในห้องเรียนออนไลน์ และคะแนนกลางภาค โดยใช้ข้อมูลคุณลักษณะทั้งหมด และแบบจำลองที่มีการคัดเลือกคุณลักษณะ
2. ผลลัพธ์จากการเปรียบเทียบระหว่างแบบจำลองการทำนายผลการเรียนที่ใช้ชุดข้อมูลที่ 2 ข้อมูลบันทึกการใช้งาน และคะแนนกิจกรรมในห้องเรียนออนไลน์ โดยใช้ข้อมูลคุณลักษณะทั้งหมด และแบบจำลองที่มีการคัดเลือกคุณลักษณะ
3. ผลลัพธ์จากการเปรียบเทียบระหว่างแบบจำลองการทำนายผลการเรียนที่ใช้ชุดข้อมูลที่ 3 ข้อมูลบันทึกการใช้งาน โดยใช้ข้อมูลคุณลักษณะทั้งหมด และแบบจำลองที่มีการคัดเลือกคุณลักษณะ
4. ผลลัพธ์จากการเปรียบเทียบระหว่างแบบจำลองการทำนายผลการเรียนที่ใช้ชุดข้อมูลที่ 4 คะแนนกิจกรรมในห้องเรียนออนไลน์ และคะแนนกลางภาค โดยใช้ข้อมูลคุณลักษณะทั้งหมด และแบบจำลองที่มีการคัดเลือกคุณลักษณะ
5. ผลลัพธ์จากการเปรียบเทียบระหว่างแบบจำลองการทำนายผลการเรียนที่ใช้ชุดข้อมูล 4 ชุด โดยใช้ข้อมูลคุณลักษณะทั้งหมด และแบบจำลองที่มีการคัดเลือกคุณลักษณะ

ผลลัพธ์ที่ได้ดังข้อ 1 ถึงข้อ 5 มาจากการสร้างแบบจำลองที่เริ่มจากการปรับความสมดุลด้วยการทำ SMOTE โดยใช้ SMOTE(k_neighbors=13, random_state=42) จากการที่เลือกใช้ k_neighbors=13 นี้ เนื่องจากทดลองค่า k_neighbors= 3, 5, 7, 9, 11 , 13, 15, 17, 19, 21, 23 พบว่า k_neighbors=13 มีค่าความแม่นยำที่สุด ต่อมาสร้างแบบจำลองที่มีการใช้ pipeline โดยปรับขนาดของข้อมูลโดยใช้ StandardScaler() และแบบจำลองการจำแนกประเภทต่าง ๆ ที่มีการกำหนดพารามิเตอร์ ดังนี้

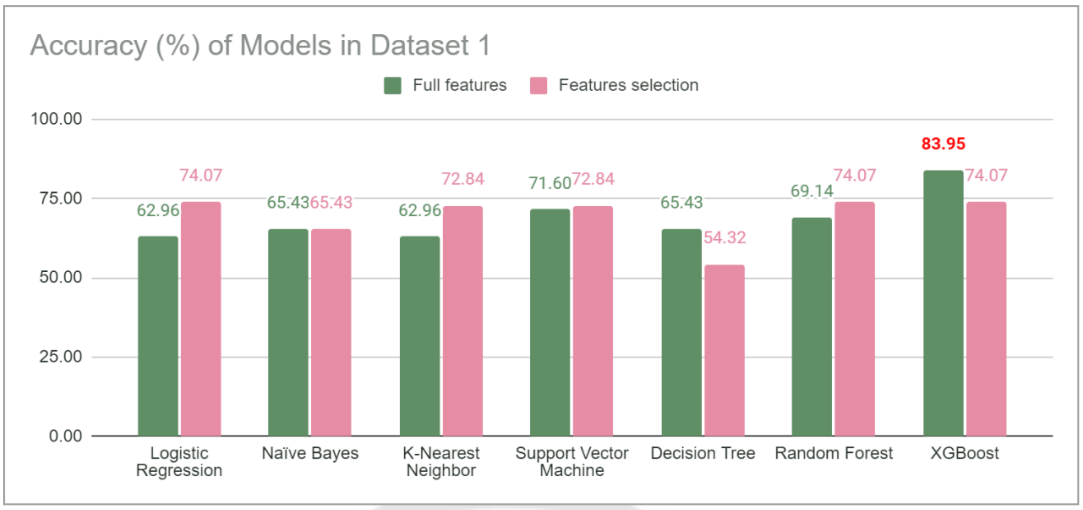
1. Logistic Regression โดยใช้ LogisticRegression(random_state=42)
2. Naïve Bayes โดยใช้ GaussianNB()
3. K-Nearest Neighbor โดยใช้ KNeighborsClassifier()
4. Support Vector Machines โดยใช้ SVC(kernel='linear')
5. Decision Tree โดยใช้ DecisionTreeClassifier(random_state=42)
6. Random Forest โดยใช้ RandomForestClassifier(random_state=42)
7. XGBoost โดยใช้ XGBClassifier(randomm_state=42)

จากการเลือกแบบจำลองดังกล่าวที่ไม่มีการปรับค่าพารามิเตอร์ เนื่องจากเริ่มจากการสร้างแบบจำลองเบื้องต้น โดยที่ยังไม่ได้ปรับค่าพารามิเตอร์ พบว่า มีค่าความแม่นยำสูงกว่าแบบจำลองที่มีการปรับค่าพารามิเตอร์ จึงเลือกแบบจำลองเบื้องต้นแทน จากนั้นได้เปรียบเทียบประสิทธิภาพระหว่างแบบจำลองที่ใช้คุณลักษณะของข้อมูลในแต่ละชุดครบถ้วน กับแบบจำลองที่ใช้ข้อมูลที่ผ่านการเลือกคุณลักษณะ โดยใช้แบบจำลอง XGBoost เป็นตัวคัดเลือกคุณลักษณะ ซึ่งใช้ SelectFromModel(XGBClassifier(random_state=42))

ส่วนสำหรับคุณลักษณะของคณะนั้น มี 3 คณะได้แก่ คณะศึกษาศาสตร์ คณะมนุษย์ศาสตร์ คณะสังคมศาสตร์ ใช้วิธีการทำ Ordinal Encoding แทนการทำ One Hot Encoding เนื่องจากทดลองการทำคุณลักษณะนี้ทั้ง 2 แบบ คือ One Hot Encoding และ Ordinal Encoding แล้ว พบว่าการใช้คุณลักษณะด้วยวิธี Ordinal Encoding ให้ค่าความแม่นยำที่สูงกว่าวิธี One Hot Encoding แต่อาจจะไม่เหมาะสม เนื่องจากทำให้แบบจำลองแบ่งค่าคุณลักษณะ ซึ่งไม่สอดคล้องกับความเป็นจริงได้ในกรณีแบบจำลองที่เป็นพวกตระกูลต้นไม้ตัดสินใจ ได้แก่ Decision Tree, Random Tree, XGBoost ซึ่งโดยปกติทำงานได้ดี เมื่อคุณลักษณะที่มี 2 ค่า

**ผลลัพธ์จากการเปรียบเทียบระหว่างแบบจำลองการทำนายผลการเรียนที่ใช้ชุดข้อมูลที่ 1 ข้อมูลบันทึกการใช้งาน คะแนนกิจกรรมในห้องเรียนออนไลน์ และคะแนนกลางภาค โดยใช้ข้อมูลคุณลักษณะทั้งหมด และแบบจำลองที่มีการคัดเลือกคุณลักษณะ**

จากการสร้างแบบจำลองการทำนายผลการเรียนที่ใช้ชุดข้อมูลที่ 1 ข้อมูลบันทึกการใช้งาน คะแนนกิจกรรมในห้องเรียนออนไลน์ และคะแนนกลางภาค เพื่อเปรียบเทียบประสิทธิภาพในการทำนายของแบบจำลองของชุดนี้ โดยใช้ข้อมูลคุณลักษณะทั้งหมด และแบบจำลองที่มีการคัดเลือกคุณลักษณะ ดังภาพประกอบ 76

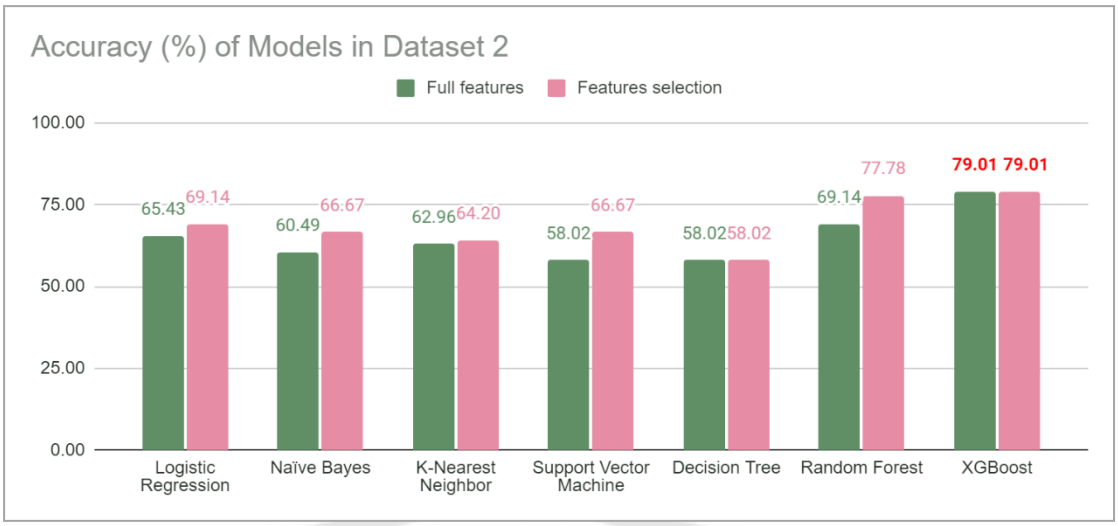



ภาพประกอบ 76 แสดงการเปรียบเทียบความแม่นยำของแบบจำลองการทำนายผลการเรียน โดยใช้ชุดข้อมูลที่ 1 ข้อมูลบันทึกการใช้งาน คะแนนกิจกรรมในห้องเรียนออนไลน์ และคะแนนกลางภาค โดยใช้ข้อมูลคุณลักษณะทั้งหมด และแบบจำลองที่มีการคัดเลือกคุณลักษณะ

จากภาพประกอบ 76 พบว่า ประสิทธิภาพของแบบจำลองของชุดข้อมูลที่ 1 คุณลักษณะกลุ่มคะแนน และคุณลักษณะกลุ่มพฤติกรรมการเข้าใช้งานระบบ โดยใช้คุณลักษณะครบถ้วน และการคัดเลือกคุณลักษณะ จะเห็นได้ว่า แบบจำลอง XGBoost ที่ใช้คุณลักษณะทั้งหมดให้ผลการทำนายที่แม่นยำที่สุด ซึ่งมีค่าเท่ากับ 83.95% รองลงมาเป็นแบบจำลอง Logistic Regression, Random Forest และ XGBoost ที่มีการคัดเลือกคุณลักษณะ ให้ค่าความแม่นยำเท่ากัน โดยมีค่าเท่ากับ 74.07% ส่วนแบบจำลองที่ให้ค่าความแม่นยำเป็นอันดับ 3 คือ แบบจำลอง K-Nearest Neighbor และ Support Vector Machine ที่มีการคัดเลือกคุณลักษณะ ให้ค่าเท่ากับ 72.84%

**ผลลัพธ์จากการเปรียบเทียบระหว่างแบบจำลองการทำนายผลการเรียนที่ใช้ชุดข้อมูลที่ 2 ข้อมูลบันทึกการใช้งาน และคะแนนกิจกรรมในห้องเรียนออนไลน์ โดยใช้ข้อมูลคุณลักษณะทั้งหมด และแบบจำลองที่มีการคัดเลือกคุณลักษณะ**

จากการสร้างแบบจำลองการทำนายผลการเรียนที่ใช้ชุดข้อมูลที่ 2 จะไม่ใช้คะแนนสอบกลางภาค เหลือเพียงคุณลักษณะกลุ่มคะแนนในส่วนของกิจกรรมในห้องเรียนออนไลน์ และกลุ่มพฤติกรรมการเข้าใช้งานระบบเพื่อเปรียบเทียบประสิทธิภาพในการทำนายของแบบจำลองของชุดนี้ โดยใช้ข้อมูลคุณลักษณะทั้งหมด และแบบจำลองที่มีการคัดเลือกคุณลักษณะ ดังภาพประกอบ 77

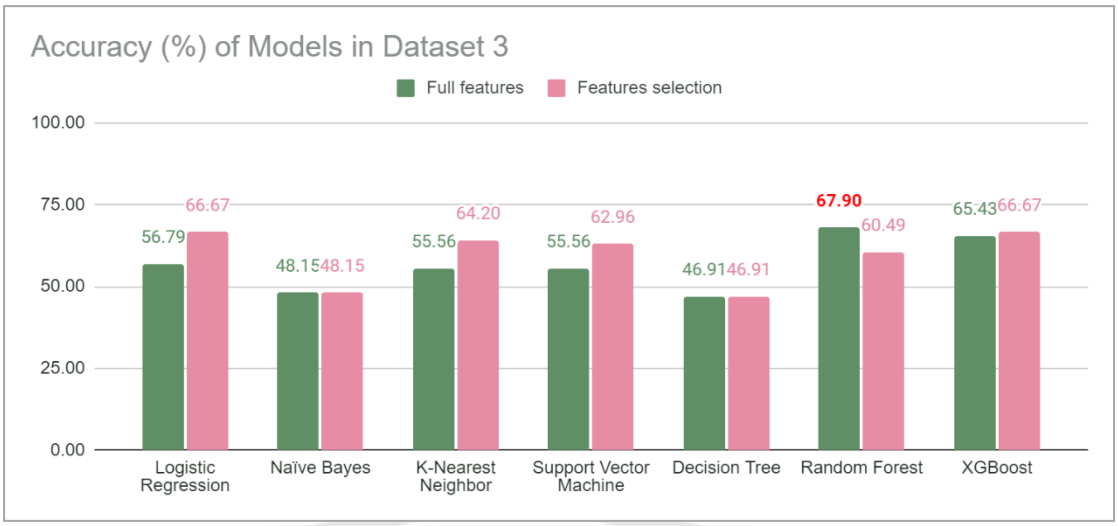



ภาพประกอบ 77 แสดงการเปรียบเทียบความแม่นยำของแบบจำลองการทำนายผลการเรียน โดยใช้ชุดข้อมูลที่ 2 ข้อมูลบันทึกการใช้งาน และคะแนนกิจกรรมในห้องเรียนออนไลน์ โดยใช้ข้อมูลคุณลักษณะทั้งหมด และแบบจำลองที่มีการคัดเลือกคุณลักษณะ

จากภาพประกอบ 77 พบว่า ประสิทธิภาพของแบบจำลองของชุดข้อมูลที่ 2 โดยใช้คุณลักษณะทั้งหมด และการคัดเลือกคุณลักษณะ จะเห็นได้ว่า แบบจำลอง XGBoost ที่ใช้คุณลักษณะทั้งหมด และแบบจำลอง XGBoost ที่มีการคัดเลือกคุณลักษณะ ให้ผลการทำนายที่แม่นยำที่สุด ซึ่งมีค่าเท่ากับ 79.01% รองลงมาเป็นแบบจำลอง Random Forest ที่มีการคัดเลือกคุณลักษณะ ให้ค่าเท่ากับ 77.78% ส่วนแบบจำลองที่ให้ค่าความแม่นยำเป็นอันดับ 3 คือ แบบจำลอง Random Forest ที่ใช้คุณลักษณะทั้งหมด และแบบจำลอง Logistic Regression ที่มีการคัดเลือกคุณลักษณะ ให้ค่าความแม่นยำเท่ากัน โดยมีค่าเท่ากับ 69.14%

**ผลลัพธ์จากการเปรียบเทียบระหว่างแบบจำลองการทำนายผลการเรียนที่ใช้ชุดข้อมูลที่ 3 ข้อมูลบันทึกการใช้งาน โดยใช้ข้อมูลคุณลักษณะทั้งหมด และแบบจำลองที่มีการคัดเลือกคุณลักษณะ**

จากการสร้างแบบจำลองการทำนายผลการเรียนที่ใช้ชุดข้อมูลที่ 3 จะใช้คุณลักษณะด้านพฤติกรรมการเข้าใช้งานระบบเท่านั้น เพื่อเปรียบเทียบประสิทธิภาพในการทำนายของแบบจำลองของชุดนี้ โดยใช้ข้อมูลคุณลักษณะทั้งหมด และแบบจำลองที่มีการคัดเลือกคุณลักษณะ ดังภาพประกอบ 78

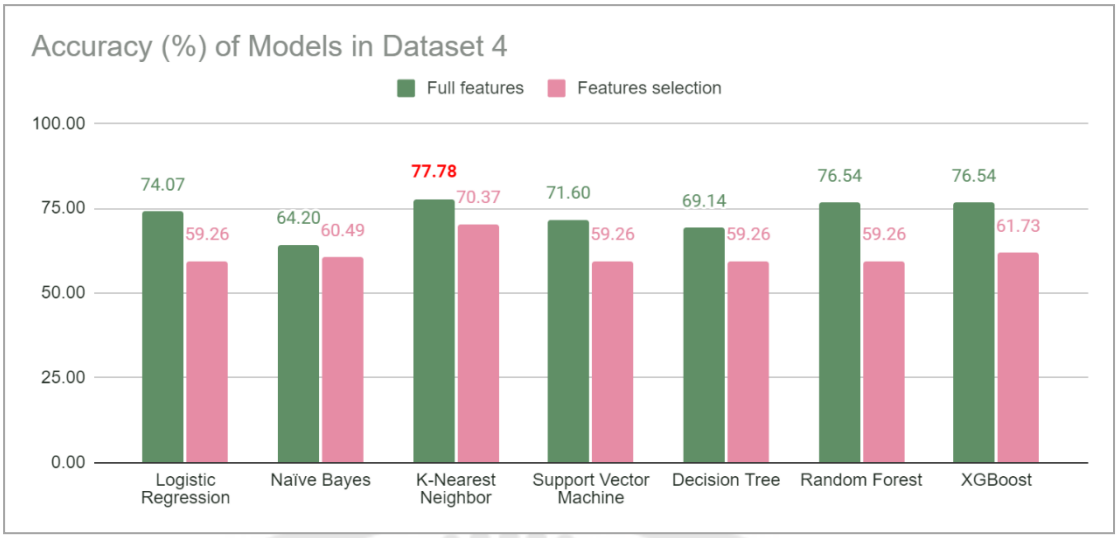



ภาพประกอบ 78 แสดงการเปรียบเทียบความแม่นยำของแบบจำลองการทำนายผลการเรียน โดยใช้ชุดข้อมูลที่ 3 ข้อมูลบันทึกการใช้งาน โดยใช้ข้อมูลคุณลักษณะทั้งหมด และแบบจำลองที่มีการคัดเลือกคุณลักษณะ

จากภาพประกอบ 78 พบว่า ประสิทธิภาพของแบบจำลองของชุดข้อมูลที่ 3 โดยใช้คุณลักษณะทั้งหมด และการคัดเลือกคุณลักษณะ จะเห็นได้ว่า แบบจำลอง Random Forest ที่ใช้คุณลักษณะทั้งหมด ให้ผลการทำนายที่แม่นยำที่สุด ซึ่งมีค่าเท่ากับ 67.90% รองลงมาเป็นแบบจำลอง Logistic Regression ที่มีการคัดเลือกคุณลักษณะ และแบบจำลอง XGBoost ที่มีการคัดเลือกคุณลักษณะ ให้ค่าเท่ากับ 66.67% ส่วนแบบจำลองที่ให้ค่าความแม่นยำเป็นอันดับ 3 คือ แบบจำลอง XGBoost ที่ใช้คุณลักษณะทั้งหมด ให้ค่าความแม่นยำเท่ากับ 65.43%

**ผลลัพธ์จากการเปรียบเทียบระหว่างแบบจำลองการทำนายผลการเรียนที่ใช้ชุดข้อมูลที่ 4 คะแนนกิจกรรมในห้องเรียนออนไลน์ และคะแนนกลางภาค โดยใช้ข้อมูลคุณลักษณะทั้งหมด และแบบจำลองที่มีการคัดเลือกคุณลักษณะ**

จากการสร้างแบบจำลองการทำนายผลการเรียนที่ใช้ชุดข้อมูลที่ 4 จะใช้คุณลักษณะด้านคะแนนเท่านั้น เพื่อเปรียบเทียบประสิทธิภาพในการทำนายของแบบจำลองของชุดนี้ โดยใช้ข้อมูลคุณลักษณะทั้งหมด และแบบจำลองที่มีการคัดเลือกคุณลักษณะ ดังภาพประกอบ 79




ภาพประกอบ 79 แสดงการเปรียบเทียบความแม่นยำของแบบจำลองการทำนายผลการเรียน โดยใช้ชุดข้อมูลที่ 4 คะแนนกิจกรรมในห้องเรียนออนไลน์ และคะแนนกลางภาค โดยใช้ข้อมูลคุณลักษณะทั้งหมด และแบบจำลองที่มีการคัดเลือกคุณลักษณะ

จากภาพประกอบ 79 พบว่า ประสิทธิภาพของแบบจำลองของชุดข้อมูลที่ 4 โดยใช้คุณลักษณะทั้งหมด และการคัดเลือกคุณลักษณะ จะเห็นได้ว่า แบบจำลอง K-Nearest Neighbor ที่ใช้คุณลักษณะทั้งหมด ให้ผลการทำนายที่แม่นยำที่สุด ซึ่งมีค่าเท่ากับ 77.78% รองลงมาเป็นแบบจำลอง Random Forest ที่ใช้คุณลักษณะทั้งหมด และแบบจำลอง XGBoost ที่ใช้คุณลักษณะทั้งหมด ให้ค่าเท่ากับ 76.54% ส่วนแบบจำลองที่ให้ค่าความแม่นยำเป็นอันดับ 3 คือ แบบจำลอง Logistic Regression ที่ใช้คุณลักษณะทั้งหมด ให้ค่าความแม่นยำเท่ากัน โดยมีค่าเท่ากับ 74.07%

**ผลลัพธ์จากการเปรียบเทียบระหว่างแบบจำลองการทำนายผลการเรียนที่ใช้ชุดข้อมูล 4 ชุด โดยใช้ข้อมูลคุณลักษณะทั้งหมด และแบบจำลองที่มีการคัดเลือกคุณลักษณะ**

เมื่อพิจารณาค่าความแม่นยำในการเปรียบเทียบประสิทธิภาพของแบบจำลองการทำนายผลการเรียนของชุดข้อมูล 4 ชุด ที่ใช้คุณลักษณะครบถ้วน และการคัดเลือกคุณลักษณะ โดยแบบจำลองในแต่ละชุดที่มีประสิทธิภาพดีที่สุด จะแสดงด้วยตัวเลขเข้ม (Bold Font) ดังตาราง 13

ตาราง 13 แสดงการเปรียบเทียบความแม่นยำสำหรับชุดข้อมูลที่ตกต่างกัน 4 ชุด ของแต่ละแบบจำลอง

| แบบจำลอง | | ชุดที่ 1 ข้อมูลบันทึกการใช้งาน คะแนนกิจกรรมในห้องเรียนออนไลน์ และคะแนนกลางภาค | | ชุดที่ 2 ข้อมูลบันทึกการใช้งานและคะแนนกิจกรรมในห้องเรียนออนไลน์ | | ชุดที่ 3 ข้อมูลบันทึกการใช้งาน | | ชุดที่ 4 คะแนนกิจกรรมในห้องเรียนออนไลน์และคะแนนกลางภาค | |
|---|---|---|---|---|---|---|---|---|---|
| | | ความแม่นยำ (%) | เปลี่ยนแปลง* | ความแม่นยำ (%) | เปลี่ยนแปลง* | ความแม่นยำ (%) | เปลี่ยนแปลง* | ความแม่นยำ (%) | เปลี่ยนแปลง* |
| 1.Logistic Regression | ใช้คุณลักษณะทั้งหมด | 62.96 | 11.11 | 65.43 | 3.71 | 56.79 | 9.88 | 74.07 | -14.81 |
| | คัดเลือกคุณลักษณะ | 74.07 | | 69.14 | | 66.67 | | 59.26 | |
| 2.Naïve Bayes | ใช้คุณลักษณะทั้งหมด | 65.43 | - | 60.49 | 6.18 | 48.15 | - | 64.20 | -3.71 |
| | คัดเลือกคุณลักษณะ | 65.43 | | 66.67 | | 48.15 | | 60.49 | |
| 3.K-Nearest Neighbor | ใช้คุณลักษณะทั้งหมด | 62.96 | 9.88 | 62.96 | 1.24 | 55.56 | 8.64 | **77.78** | -7.41 |
| | คัดเลือกคุณลักษณะ | 72.84 | | 64.20 | | 64.20 | | 70.37 | |
| 4.Support Vector Machine | ใช้คุณลักษณะทั้งหมด | 71.60 | 1.24 | 58.02 | 8.65 | 55.56 | 7.40 | 71.60 | -12.34 |
| | คัดเลือกคุณลักษณะ | 72.84 | | 66.67 | | 62.96 | | 59.26 | |
| 5.Decision Tree | ใช้คุณลักษณะทั้งหมด | 65.43 | -11.11 | 58.02 | - | 46.91 | - | 69.14 | -9.88 |
| | คัดเลือกคุณลักษณะ | 54.32 | | 58.02 | | 46.91 | | 59.26 | |
| 6.Random Forest | ใช้คุณลักษณะทั้งหมด | 69.14 | 4.93 | 69.14 | 8.64 | **67.90** | -7.41 | 76.54 | -17.28 |
| | คัดเลือกคุณลักษณะ | 74.07 | | 77.78 | | 60.49 | | 59.26 | |
| 7.XGBoost | ใช้คุณลักษณะทั้งหมด | **83.95** | -9.88 | **79.01** | - | 65.43 | 1.24 | 76.54 | -14.81 |
| | คัดเลือกคุณลักษณะ | 74.07 | | **79.01** | | 66.67 | | 61.73 | |

หมายเหตุ เปลี่ยนแปลง* หมายถึง ร้อยละการเปลี่ยนแปลงของความแม่นยำ หลังจากผ่านการคัดเลือกคุณลักษณะ

จากตาราง 13 พบว่า แบบจำลองการทำนายผลการเรียนที่ไม่นำคะแนนสอบกลางภาคในชุดข้อมูลที่ 2 นั้น โดยรวม มีความแม่นยำลดลงจากแบบจำลองในชุดข้อมูลที่ 1 โดยเฉลี่ย 4.06% หากจะไม่พิจารณาคุณลักษณะกลุ่มคะแนนทั้งหมดไปทำนาย จะเห็นได้ว่ามีความแม่นยำลดลงอย่างเห็นได้ชัด โดยเฉลี่ย 12.17% ส่วนหากจะพิจารณาคุณลักษณะกลุ่มคะแนนเท่านั้น จะมีความแม่นยำเพิ่มขึ้นไม่มาก โดยเฉลี่ย 4.06% แต่แบบจำลอง XGBoost ในชุดที่ 1 คุณลักษณะกลุ่มพฤติกรรมการใช้งานระบบร่วมกับคุณลักษณะกลุ่มคะแนน มีความแม่นยำสูงกว่าแบบจำลองนี้ในชุดที่ 4 ที่ใช้เพียงคุณลักษณะคะแนน มากถึง 7.41%

# บทที่ 5
# สรุปผลการวิจัย อภิปรายผล และข้อเสนอแนะ

ในการวิจัยเพื่อศึกษาเกี่ยวกับการทำนายผลการเรียนของนิสิตที่ใช้งานระบบการจัดการเรียนรู้ออนไลน์ โดยใช้การเรียนรู้ของเครื่อง ผู้วิจัยได้วัดประสิทธิภาพของแบบจำลองต่าง ๆ ที่สร้างขึ้นมาในแต่ละชุดข้อมูล เพื่อนำผลที่ได้มาเปรียบเทียบและสรุปผล โดยสามารถแบ่งหัวข้อดังนี้

1. สรุปผลการวิจัย
2. อภิปรายผลการวิจัย
3. ข้อเสนอแนะ

## สรุปผลการวิจัย

งานวิจัยนี้สร้างแบบจำลองการทำนายผลการเรียนของนิสิตที่ใช้งานระบบการจัดการเรียนรู้ออนไลน์ โดยใช้การเรียนรู้ของเครื่อง แบบจำลองเหล่านี้มีการร่วมการปรับความสมดุลของข้อมูลด้วยเทคนิค SMOTE และการคัดเลือกคุณลักษณะ ในผลการดำเนินการวิจัยดังตาราง 13 จะเห็นได้ว่า เมื่อเปรียบเทียบประสิทธิภาพของแบบจำลองต่าง ๆ ที่ใช้ชุดข้อมูลที่แตกต่างกัน ซึ่งถูกแบ่งตามช่วงเวลาการเรียนและประเภทของคุณลักษณะ พบว่า แบบจำลอง XGBoost ของชุดข้อมูลที่ 1 ที่ใช้ข้อมูลคุณลักษณะกลุ่มพฤติกรรมการใช้งานระบบร่วมกับคุณลักษณะกลุ่มคะแนน โดยใช้คุณลักษณะครบถ้วน มีประสิทธิภาพสูงสุด โดยมีความแม่นยำที่สุดเท่ากับ 83.95% ส่วนแบบจำลองที่ทำนายแม่นยำที่สุดสำหรับชุดข้อมูลที่ 2 ที่ไม่นำคะแนนสอบกลางภาคมาทำนายผลการเรียน คือ แบบจำลอง XGBoost ที่ใช้คุณลักษณะทั้งหมด ให้ค่าเท่ากับ 79.01% ส่วนชุดข้อมูลที่ 3 ใช้เพียงคุณลักษณะกลุ่มพฤติกรรมการใช้งานระบบเท่านั้น แบบจำลองที่ให้ประสิทธิภาพสูงสุดคือ แบบจำลอง Random Forest มีค่าความแม่นยำเท่ากับ 67.90% รองลงมาเป็นแบบจำลอง XGBoost มีค่าความแม่นยำเท่ากับ 65.43% สุดท้ายชุดข้อมูลที่ 4 ใช้เพียงคุณลักษณะกลุ่มคะแนนเท่านั้น แบบจำลอง XGBoost ให้ค่าความแม่นยำเท่ากับ 76.54% ซึ่งเป็นอันดับสอง ขณะที่แบบจำลอง K-Nearest Neighbor มีค่าความแม่นยำสูงสุดเท่ากับ 77.78% สรุปโดยรวมแล้ว แบบจำลอง XGBoost ให้ประสิทธิภาพโดยรวมสูงที่สุดอย่างสม่ำเสมอในชุดข้อมูลทุกประเภท แบบจำลองนี้มีศักยภาพสูงในการนำมาใช้ในการทำนายผลการเรียนของนิสิตได้อย่างมีประสิทธิภาพ ดังนั้น แบบจำลอง XGBoost ที่ใช้ชุดข้อมูลที่ 1 เป็นแบบจำลองที่เหมาะกับการนำข้อมูลชุดนี้มาใช้ในงานมากที่สุด จะเห็นได้ว่า แบบจำลองที่ใช้คุณลักษณะกลุ่มพฤติกรรม

การใช้งานระบบร่วมกับคุณลักษณะกลุ่มคะแนน มีประสิทธิภาพที่ดีกว่าแบบจำลองที่เพียงใช้คุณลักษณะกลุ่มคะแนน ซึ่งตรงกับผลการวิจัยของ Amrieh และคนอื่น ๆ (2016) ที่ได้ข้อสรุปว่าแบบจำลองสำหรับคุณลักษณะที่มีกลุ่มพฤติกรรมการใช้งานระบบมีประสิทธิภาพที่ดีกว่า เมื่อเทียบกับคุณลักษณะที่ไม่มีกลุ่มพฤติกรรมกรรมการใช้งานระบบ และการคัดเลือกคุณลักษณะช่วยให้แบบจำลองที่มีความซับซ้อนไม่มาก เช่น Logistic Regression หรือ Naïve Bayes มีประสิทธิภาพดีขึ้น แต่ไม่ได้ช่วยแบบจำลองที่มีความซับซ้อนสูงอยู่แล้ว เช่น XGBoost ซึ่งมีความสอดคล้องกับบทความวิจัยของ Ramaswami และคนอื่น ๆ (2020) ในส่วนที่แบบจำลอง Naïve Bayes และ Logistic Regression มีประสิทธิภาพที่สูงขึ้น หลังจากทำการคัดเลือกคุณลักษณะ นอกจากนี้ยังพบว่า คุณลักษณะที่สำคัญในการทำนายผลการเรียนของนิสิต คือ พฤติกรรมการใช้งานระบบของนิสิต มีหลายพฤติกรรมถึง 19 คุณลักษณะ คะแนนผลการประเมินการเรียนรู้ มีเพียง 10 คุณลักษณะ และข้อมูลคณะของนิสิต ดังนั้น สามารถนำแบบจำลอง XGBoost ไปทำนายผลการเรียนของนิสิต โดยใช้คุณลักษณะของพฤติกรรมการใช้งานระบบร่วมกับคะแนนและคณะของนิสิต เพื่อนำผลการทำนายนี้ไปคำแนะนำและช่วยเหลือนิสิตได้อย่างเจาะจง สามารถติดตามความก้าวหน้าในการเรียนของนิสิตได้เป็นรายบุคคล หรือวางแผนการสอนให้เหมาะสมกับนิสิตและนิสิตมีเวลาพัฒนาทักษะการเรียนรู้มากขึ้น

**อภิปรายผลการวิจัย**

เมื่อนำแบบจำลองที่มีความแม่นยำที่สุดจากการเปรียบเทียบประสิทธิภาพของแบบจำลองต่าง ๆ โดยใช้ชุดข้อมูลที่แตกต่างกัน ไปพิจารณาค่า Prediction และค่า Recall ของแบบจำลองนี้ ดังตาราง 14

ตาราง 14 แสดงค่า Prediction และค่า Recall ของแบบจำลอง XGBoost ของข้อมูลชุดที่ 1 ข้อมูลคุณลักษณะกลุ่มคะแนน และคุณลักษณะกลุ่มพฤติกรรมการใช้งานระบบ โดยใช้คุณลักษณะครบถ้วน

| ระดับผลการเรียน | ค่า Precision | ค่า Recall |
|---|---|---|
| ดีมาก | 0.93 | 0.86 |
| ดี | 0.76 | 0.81 |
| ต้องปรับปรุง | 0.73 | 0.80 |

จากตาราง 14 พบว่า แบบจำลอง XGBoost ของชุดข้อมูลที่ 1 ที่ใช้ข้อมูลคุณลักษณะกลุ่มคะแนน และคุณลักษณะกลุ่มพฤติกรรมการใช้งานระบบ โดยใช้คุณลักษณะครบถ้วน สำหรับกลุ่มผลการเรียนต้องปรับปรุงมีค่า Recall เท่ากับ 0.80 ถึงแม้ว่าในกลุ่มผลการเรียนนี้มีจำนวนที่น้อยกว่ากลุ่มอื่น ๆ แต่ยังมีค่า Recall ใกล้เคียงกับกลุ่มผลการเรียนดีมาก และดี แสดงว่าแบบจำลองนี้ยังมีความสามารถในการค้นหานิสิตกลุ่มผลการเรียนต้องปรับปรุงออกมาได้ เพื่อให้อาจารย์ผู้สอนให้ความเอาใจใส่กลุ่มนิสิตที่ได้ผลการเรียนนี้ให้เป็นพิเศษ และช่วยให้นิสิตมีโอกาสปรับปรุงผลการเรียนที่ดีขึ้นได้

ในส่วนของการทำนายผลการเรียนโดยไม่ใช้คะแนนสอบกลางภาคในชุดที่ 2 นั้น จะเห็นได้ว่า โดยรวมแล้ว ประสิทธิภาพของแบบจำลองในข้อมูลชุดที่ 1 สูงกว่าแบบจำลองในข้อมูลชุดที่ 2 โดยเฉลี่ย 4.06% แสดงให้เห็นว่า การทำนายล่วงหน้าก่อนที่นิสิตจะได้ผลสอบกลางภาคออกมานั้น มีความแม่นยำลดลงไม่มากนัก นอกจากนี้แม้ว่าค่าความแม่นยำของแบบจำลอง XGBoost จะตกลงบ้าง แต่ค่า Recall ในกรณีข้อมูลชุดที่ 2 ในกลุ่มนิสิตผลการเรียนต้องปรับปรุงก็ยังมีค่าอยู่ที่ 0.80 เช่นเดิม

ทว่า หากเราไม่พิจารณาคะแนน ดังในข้อมูลชุดที่ 3 จะเห็นว่า ประสิทธิภาพของแบบจำลองทุกชนิด ลดลงอย่างเห็นได้ชัดเจน โดยเฉลี่ยแล้ว ลดลงถึง 12.17% เมื่อเทียบกับประสิทธิภาพของแบบจำลองในข้อมูลชุดที่ 1 ทว่า ในข้อมูลชุดนี้ แบบจำลอง XGBoost ก็ยังมีประสิทธิภาพสูงอยู่ เมื่อเทียบกับแบบจำลองชนิดอื่น อีกทั้ง ในข้อมูลชุดที่ 3 นี้ ค่า Recall ในกลุ่มผลการเรียนต้องปรับปรุงของทุกแบบจำลอง จะมีค่าต่ำกว่าหรือเท่ากับ 0.60 ทั้งสิ้น (โดยแบบจำลอง XGBoost มีค่า Recall เท่ากับ 0.50) ดังนั้น ในการทำนายผลการเรียน โดยอาศัยแต่เพียงข้อมูลพฤติกรรมการใช้งานระบบของนิสิตอย่างเดียว โดยไม่อาศัยคะแนนในชั้นเรียน อาจไม่เพียงพอต่อการทำนายผลการเรียนของนิสิตได้อย่างแม่นยำ

ในส่วนของประสิทธิภาพแบบจำลองที่ใช้ข้อมูลชุดที่ 4 จะมีความแม่นยำเพิ่มขึ้นไม่มาก โดยเฉลี่ยถึง 4.06% เมื่อเปรียบเทียบกับความแม่นยำของข้อมูลชุดที่ 1 แต่แบบจำลอง XGBoost ในชุดที่ 1 มีค่าความแม่นยำสูงกว่าแบบจำลองนี้ในชุดที่ 4 ถึง 7.41% แสดงให้เห็นว่า แบบจำลอง XGBoost ที่ใช้แต่เพียงคะแนนที่ได้อย่างเดียว โดยไม่พิจารณาข้อมูลพฤติกรรมการใช้งานระบบของนิสิต ให้ผลการทำนายแม่นยำที่น้อยกว่า เมื่อเทียบกับแบบจำลองนี้ที่ใช้ข้อมูลพฤติกรรมการใช้งานระบบของนิสิตร่วมกับคะแนน

จากการศึกษาในส่วนนี้ จะเห็นว่า การใช้คุณลักษณะครบถ้วนทั้งสองกลุ่มคือ คุณลักษณะด้านคะแนน และ คุณลักษณะด้านพฤติกรรมการเข้าใช้ระบบร่วมกัน ในชุดข้อมูลที่ 1 จะให้

ประสิทธิภาพของทุกแบบจำลองสูงสุด ในขณะที่ในชุดข้อมูลที่ 3 เราใช้คุณลักษณะด้านพฤติกรรมการเข้าใช้ระบบแต่เพียงอย่างเดียว จะให้ผลการทำนายที่ต่ำกว่าชุดข้อมูลที่ 4 ที่ใช้แต่เพียงคุณลักษณะด้านคะแนนอย่างเดียว จึงทำให้เห็นได้ว่า อย่างไรก็ตาม คุณลักษณะด้านคะแนน ช่วยให้ทำนายผลการเรียนได้แม่นยำกว่า คุณลักษณะด้านพฤติกรรมแต่เพียงอย่างเดียว แต่ทั้งนี้แบบจำลองจะให้ความแม่นยำสูงสุด เมื่อใช้คุณลักษณะทั้งสองกลุ่มนี้ร่วมกัน

ในส่วนของประสิทธิภาพแบบจำลองชุดข้อมูลแต่ละชุดแล้ว จะเห็นว่า แบบจำลอง XGBoost ให้ประสิทธิภาพการทำนายสูงสุด ทั้งในข้อมูลชุดที่ 1 และ 2 ซึ่งเป็นข้อมูลส่วนที่ใช้คุณลักษณะมากที่สุด และมีผลการทำนายที่แม่นยำที่สุด ส่วนในชุดข้อมูลที่ 3 แบบจำลองที่ให้ประสิทธิภาพดีที่สุด จะเป็น Random forest ได้ค่า 67.9 ส่วนแบบจำลอง XGBoost จะให้ค่าความแม่นยำเป็นอันดับสอง ที่ค่า 65.43 ส่วนในชุดข้อมูลที่ 4 แบบจำลอง XGBoost จะให้ความแม่นยำ 76.54 ซึ่งมีค่าเป็นอันดับสอง ในขณะที่แบบจำลอง K-Nearest Neighbor ให้ค่าความแม่นยำสูงสุดที่ 77.78 จะเห็นได้ว่า โดยรวมแล้ว แบบจำลอง XGBoost สามารถทำนายข้อมูลชุดต่างๆ ได้อย่างแม่นยำอย่างสม่ำเสมอ แม้ว่าในบางชุดข้อมูล แบบจำลอง XGBoost อาจมิได้ให้ค่าประสิทธิภาพที่สูงที่สุด แต่ค่าความแม่นยำที่ได้ ก็มิได้แย่กว่าแบบจำลองที่ดีที่สุดมากนัก ทั้งในชุดข้อมูลที่ 3 และ 4 แบบจำลอง XGBoost ยังให้ความแม่นยำเป็นอันดับสอง จึงสรุปได้ว่าแบบจำลอง XGBoost เป็นแบบจำลองที่เหมาะกับการนำมาใช้งานกับในข้อมูลชุดนี้มากที่สุด เนื่องจากให้ประสิทธิภาพโดยรวมสูงที่สุดอย่างสม่ำเสมอในชุดข้อมูลทุกประเภท

ในส่วนของการศึกษาชุดข้อมูลที่ 3 และ 4 นั้น เราต้องการศึกษาประสิทธิภาพการทำนายของการใช้คุณลักษณะกลุ่มของพฤติกรรมการเข้าใช้งานระบบ เทียบกับการใช้คุณลักษณะกลุ่มของคะแนนสอบ ซึ่งในการศึกษานี้ เราพบว่า ในการสร้างแบบจำลองที่โดยรวมแล้ว การทำนายผลการเรียนด้วยคุณลักษณะกลุ่มของคะแนนสอบแต่เพียงอย่างเดียว จะให้ความแม่นยำสูงกว่า การใช้คุณลักษณะกลุ่มของพฤติกรรมการเข้าใช้งานระบบแต่เพียงอย่างเดียว

ดังนั้น ในการสร้างแบบจำลองเพื่อใช้งานจริง เราจะสร้างแบบจำลองที่ใช้ชุดข้อมูลที่ 1 เป็นหลัก จึงได้ข้อสรุปว่า แบบจำลอง XGBoost เป็นแบบจำลองที่เหมาะสมกับข้อมูลชุดนี้ที่สุด

ต่อไปจะเปรียบเทียบแบบจำลองการทำนายผลการเรียนที่ใช้คุณลักษณะทั้งหมด และแบบจำลองที่มีการคัดเลือกคุณลักษณะการคัดเลือกคุณลักษณะ ดังตาราง 15

ตาราง 15 แสดงจำนวนคุณลักษณะก่อนและหลังจากทำการคัดเลือกคุณลักษณะ

| การคัดเลือกคุณลักษณะ | ชุดที่ 1 | ชุดที่ 2 | ชุดที่ 3 | ชุดที่ 4 |
|---|---|---|---|---|
| ก่อนการคัดเลือกคุณลักษณะ | 115 | 111 | 99 | 18 |
| หลังการคัดเลือกคุณลักษณะ | 36 | 33 | 41 | 5 |

โดยรวมแล้ว การคัดเลือกคุณลักษณะ ไม่ค่อยได้ช่วยให้แบบจำลองที่มีความแม่นยำค่อนข้างสูงนัก เช่น พวก XGBoost เป็นต้น สำหรับแบบจำลอง Naïve Bayes หรือ Logistic Regression จะให้ประสิทธิภาพที่ดีขึ้น เนื่องจากอาจจะเป็นแบบจำลองที่มีความซับซ้อนไม่มากนัก เมื่อนำข้อมูลที่มีจำนวนคุณลักษณะมาก ๆ มาทำนายผลการเรียน จะทำให้สามารถประมวลผลข้อมูลได้ดีไม่พอ หากลดจำนวนคุณลักษณะลง แล้วทำให้แบบจำลองเหล่านี้มีความแม่นยำมากขึ้น ส่วนเมื่อนำแบบจำลองที่มีความซับซ้อนสูง เช่น แบบจำลอง XGBoost จะทำให้สามารถประมวลผลข้อมูลที่มีความซ้อนมาก เมื่อนำแบบจำลองนี้ไปทำการคัดเลือกคุณลักษณะ ทำให้จำเป็นต้องคัดข้อมูลบางส่วนออก มีความแม่นยำน้อยลงจากแบบจำลองที่ใช้คุณลักษณะครบถ้วน

ต่อไปพิจารณาคุณลักษณะที่สำคัญของแบบจำลองที่มีความแม่นยำที่สุดในการทำนายผลการเรียน ดังตาราง 16 จะเห็นว่า คุณลักษณะสำคัญจะอยู่ในกลุ่มของพฤติกรรมการใช้งานระบบของนิสิตมากถึง 19 คุณลักษณะ ในขณะที่คุณลักษณะสำคัญในกลุ่มของคะแนนผลการประเมินการเรียนรู้จะมีเพียง 10 คุณลักษณะ และคุณลักษณะข้อมูลทั่วไปจะปรากฏเพียงคุณลักษณะเดียว คือ ข้อมูลคณะของนิสิต

ตาราง 16 แสดงคุณลักษณะที่สำคัญ 30 อันดับแรกที่ได้จากแบบจำลอง XGBoost ที่ใช้ชุดข้อมูลที่ 1 โดยใช้คุณลักษณะครบถ้วน

| อันดับความสำคัญ | คุณลักษณะ | ความหมาย | จัดเป็นข้อมูลในกลุ่ม |
|---|---|---|---|
| 1 | NumTitleQuiz | จำนวนเรื่องที่ทำแบบทดสอบ | ด้านคะแนนผลการประเมินการเรียนรู้ |
| 2 | Total_ScoreQuiz | คะแนนแบบทดสอบรวมทุกเรื่อง | ด้านคะแนนผลการประเมินการเรียนรู้ |
| 3 | ScoreQuiz2 | คะแนนแบบทดสอบ เรื่องที่ 2 | ด้านคะแนนผลการประเมินการเรียนรู้ |
| 4 | DayPageView4 | จำนวนวันที่ชมสื่อการเรียนรู้ เรื่องที่ 4 | ด้านพฤติกรรมการใช้งานระบบของนิสิต |
| 5 | Total_ScoreOnlineMeeting | คะแนนกิจกรรมในห้องเรียนออนไลน์รวมทุกเรื่อง | ด้านคะแนนผลการประเมินการเรียนรู้ |
| 6 | DayAssignmentBefore | จำนวนวันระหว่างวันแรกที่ส่งงานมอบหมาย และวันสุดท้ายการส่งงานมอบหมาย | ด้านพฤติกรรมการใช้งานระบบของนิสิต |

ตาราง 16 (ต่อ)

| อันดับความสำคัญ | คุณลักษณะ | ความหมาย | จัดเป็นข้อมูลในกลุ่ม |
|---|---|---|---|
| 7 | DayQuizBefore3 | จำนวนวันระหว่างวันที่ทำแบบทดสอบครั้งแรกและวันสุดท้ายการทำแบบทดสอบ เรื่องที่ 3 | ด้านพฤติกรรมการใช้งานระบบของนิสิต |
| 8 | ScoreAttendance | คะแนนการเข้าเรียน | ด้านคะแนนผลการประเมินการเรียนรู้ |
| 9 | ScoreQuiz5 | คะแนนแบบทดสอบ เรื่องที่ 5 | ด้านคะแนนผลการประเมินการเรียนรู้ |
| 10 | ScoreMidterm | คะแนนสอบกลางภาค | ด้านคะแนนผลการประเมินการเรียนรู้ |
| 11 | ScoreOnlineMeeting1 | คะแนนกิจกรรมในห้องเรียนออนไลน์ เรื่องที่ 1 | ด้านคะแนนผลการประเมินการเรียนรู้ |
| 12 | ScoreQuiz3 | คะแนนแบบทดสอบ เรื่องที่ 3 | ด้านคะแนนผลการประเมินการเรียนรู้ |
| 13 | RepeatQuiz3 | จำนวนครั้งที่ทำแบบทดสอบซ้ำ | ด้านพฤติกรรมการใช้งานระบบของนิสิต |
| 14 | DurationQuiz3_3 | ระยะเวลาการทำแบบทดสอบ เรื่องที่ 3 ในครั้งที่ 3 | ด้านพฤติกรรมการใช้งานระบบของนิสิต |
| 15 | ScoreAssignment1 | คะแนนงานมอบหมาย ส่วนที่ 1 | ด้านคะแนนผลการประเมินการเรียนรู้ |
| 16 | Faculty | คณะของนิสิต | ข้อมูลทั่วไป |
| 17 | DayPageView3 | จำนวนวันที่ชมสื่อการเรียนรู้ เรื่องที่ 3 | ด้านพฤติกรรมการใช้งานระบบของนิสิต |
| 18 | ClickPage9 | จำนวนครั้งที่คลิกเข้าถึงสื่อการเรียนรู้ เรื่องที่ 9 | ด้านพฤติกรรมการใช้งานระบบของนิสิต |
| 19 | DurationQuiz2_3 | ระยะเวลาการทำแบบทดสอบ เรื่องที่ 2 ในครั้งที่ 3 | ด้านพฤติกรรมการใช้งานระบบของนิสิต |
| 20 | DayQuizBefore4 | จำนวนวันระหว่างวันที่ทำแบบทดสอบครั้งแรกและวันสุดท้ายการทำแบบทดสอบ เรื่องที่ 4 | ด้านพฤติกรรมการใช้งานระบบของนิสิต |
| 21 | ClickAssignment1 | จำนวนครั้งที่คลิกเข้าถึงงานมอบหมาย เรื่องที่ 1 | ด้านพฤติกรรมการใช้งานระบบของนิสิต |
| 22 | ClickPage4 | จำนวนครั้งที่คลิกเข้าถึงสื่อการเรียนรู้ เรื่องที่ 4 | ด้านพฤติกรรมการใช้งานระบบของนิสิต |
| 23 | Total_ScoreAssignment | คะแนนงานมอบหมายรวมทุกส่วน | ด้านคะแนนผลการประเมินการเรียนรู้ |
| 24 | AvgMinuteQuiz4 | ระยะเวลาเฉลี่ยการทำแบบทดสอบรวมทุกครั้ง เรื่องที่ 4 | ด้านพฤติกรรมการใช้งานระบบของนิสิต |
| 25 | DayQuizBefore1 | จำนวนวันระหว่างวันที่ทำแบบทดสอบครั้งแรกและวันสุดท้ายการทำแบบทดสอบ เรื่องที่ 1 | ด้านพฤติกรรมการใช้งานระบบของนิสิต |
| 26 | ClickPage8 | จำนวนครั้งที่คลิกเข้าถึงสื่อการเรียนรู้ เรื่องที่ 8 | ด้านพฤติกรรมการใช้งานระบบของนิสิต |
| 27 | DayQuizBefore5 | จำนวนวันระหว่างวันที่ทำแบบทดสอบครั้งแรกและวันสุดท้ายการทำแบบทดสอบ เรื่องที่ 5 | ด้านพฤติกรรมการใช้งานระบบของนิสิต |
| 28 | DayFirstPageView5 | จำนวนวันระหว่างวันแรกที่เรียน และวันแรกที่ชมสื่อการเรียนรู้ เรื่องที่ 5 | ด้านคะแนนผลการประเมินการเรียนรู้ |
| 29 | DayFirstPageView9 | จำนวนวันระหว่างวันแรกที่เรียน และวันแรกที่ชมสื่อการเรียนรู้ เรื่องที่ 9 | ด้านพฤติกรรมการใช้งานระบบของนิสิต |
| 30 | Total_ClickAssignment | จำนวนครั้งที่คลิกเข้าถึงงานมอบหมายรวมทุกเรื่อง | ด้านคะแนนผลการประเมินการเรียนรู้ |

**ข้อเสนอแนะ**

1. การวิเคราะห์สถิติการใช้งานระบบในส่วนการประมวลผลระยะเวลาการทำแบบทดสอบ โดยใช้โปรแกรม Microsoft Excel เพื่อสร้างคุณลักษณะกลุ่มพฤติกรรมการใช้งานระบบสำหรับการทำนายผลการเรียนนั้น สำหรับวิธีการประมวลผลนั้นค่อนข้างใช้เวลานาน เนื่องจากข้อมูลที่ดึงมาจากระบบการจัดการเรียนรู้ออนไลน์อยู่ในรูปแบบ Timeline ดังนั้น อาจใช้เทคนิคการใช้งาน Microsoft Excel ที่เชี่ยวชาญมากขึ้น หรือตั้งค่ารูปแบบการแสดงข้อมูลสรุประยะเวลาการทำแบบทดสอบในไฟล์ที่ดึงออกจากระบบว่าสามารถทำได้หรือไม่

2. มีบางรายการที่ไม่มีข้อมูลเวลาเริ่มต้น หรือเวลาสิ้นสุดในการทำแบบทดสอบของนิสิต ทำให้ลบข้อมูลที่สำคัญออกไป ดังนั้น อาจจะทำให้ระบบการจัดการเรียนรู้ออนไลน์มีความเสถียรภาพมากขึ้น จะได้ข้อมูลที่เป็นประโยชน์และถูกต้องมากขึ้น ตลอดจนนำไปสู่การเพิ่มประสิทธิภาพในการทำนายผลการเรียน

3. งานวิจัยนี้สร้างแบบจำลองการทำนายผลการเรียนของนิสิตที่ใช้งานระบบการจัดการเรียนรู้ออนไลน์ โดยใช้ข้อมูลของนิสิตที่มีจำนวนไม่มากนัก ในอนาคตอาจจะเพิ่มจำนวนนิสิต เพื่อให้แบบจำลองนี้มีประสิทธิภาพการทำนายที่สูงขึ้น

4. เนื่องจากงานวิจัยนี้ออกแบบแบบจำลองการทำนายผลการเรียนสำหรับ Boosting เพียงชนิดเดียวคือ XGBoost แต่แบบจำลองพวก Boosting มีหลากหลายแบบจำลอง นอกจาก XGBoost ที่ทดลองสร้างแล้ว ได้แก่ AdaBoost, Gradient Boosting, Light GBM และ CatBoost เพื่อช่วยเพิ่มประสิทธิการทำนายมากยิ่งขึ้น

5. หากภายในระบบการจัดการเรียนรู้ออนไลน์สามารถทำงานร่วมกับการสร้างแบบจำลองการทำนายผลการเรียน ทำให้ได้ผลการทำนายเป็น Real Time ที่รวดเร็วขึ้น และถูกต้องมากขึ้น

# บรรณานุกรม


Abuzalata, A. (2019). *Academic Performance Prediction for Engineering Students using RBF-SVM Classification Model. Case Study: PPU.* (Master's thesis, Palestine Polytechnic University, Hebron). http://scholar.ppu.edu/bitstream/handle/123456789/1733/ThesisAfterModifications.pdf?sequence=1&isAllowed=y

Ahmad, N., Anuar, S., และ Hassan, N. (2020). Improved students' performance prediction for multi-class imbalanced problems using hybrid and ensemble approach in educational data mining. *Journal of Physics Conference Series, 1877*, 52041.

Ali, A. (2018, Jul 21, 2018). K-Nearest Neighbor with Practical Implementation. https://medium.com/machine-learning-researcher/k-nearest-neighbors-in-machine-learning-e794014abd2a

Ali, J., Khan, R., Ahmad, N., และ Maqsood, I. (2012). Random Forests and Decision Trees. *International Journal of Computer Science Issues(IJCSI), 9*.

Almasri, A., Alkhawaldeh, R. S., และ Çelebi, E. (2020). Clustering-Based EMT Model for Predicting Student Performance. *Arabian Journal for Science and Engineering, 45*(12), 10067-10078.

Alto, V. (2020, Jan 11, 2020). Understanding AdaBoost for Decision Tree. https://towardsdatascience.com/understanding-adaboost-for-decision-tree-ff8f07d2851

Amra, I. A. A., และ Maghari, A. Y. A. (2017, 17-18 May 2017). *Students performance prediction using KNN and Naïve Bayesian.* Paper presented at the 2017 8th International Conference on Information Technology (ICIT).

Amrieh, E., Hamtini, T., และ Aljarah, I. (2016). Mining Educational Data to Predict Student's academic Performance using Ensemble Methods. *International Journal of*

*Database Theory and Application, 9*, 119-136.

Ashfaq, U., Poolan Marikannan, B., และ Raheem, M. (2020). Managing Student Performance: A Predictive Analytics using Imbalanced Data. *International Journal of Recent Technology and Engineering, 8*, 2277-2283.

Awad, M., และ Khanna, R. (2015). Machine Learning. https://link.springer.com/chapter/10.1007/978-1-4302-5990-9_1

Awasthi, S. (2020, December 17, 2020). SEVEN MOST POPULAR SVM KERNELS. https://dataaspirant.com/svm-kernels/

Bouaguel, W. (2015). *On Feature Selection for Credit Scoring*.

Brixius, N. (2016, June 4, 2016). The Logit and Sigmoid Functions. https://nathanbrixius.wordpress.com/2016/06/04/functions-i-have-known-logit-and-sigmoid/

Brownlee, J. (2016, April 25, 2016). Boosting and AdaBoost for Machine Learning. https://machinelearningmastery.com/boosting-and-adaboost-for-machine-learning/

Brownlee, J. (2020a, August 20, 2020). How to Calculate Feature Importance With Python. https://machinelearningmastery.com/calculate-feature-importance-with-python/

Brownlee, J. (2020b, June 19, 2020). Tour of Data Preparation Techniques for Machine Learning. https://machinelearningmastery.com/data-preparation-techniques-for-machine-learning/

Capterra. (2021). The 2021 Capterra Shortlist: Learning Management System Software. https://www.capterra.com/learning-management-system-software#shortlist

Chen, T., และ Guestrin, C. (2016). *XGBoost: A Scalable Tree Boosting System*. Paper presented at the Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.

Datacamp. (2021). Preprocessing for Machine Learning in Python.

https://www.datacamp.com/courses/preprocessing-for-machine-learning-in-python

Equiskill. (2018, July 5, 2018). Understanding Logistic Regression. https://www.equiskill.com/understanding-logistic-regression/

Gulati, P., Sharma, A., และ Gupta, M. (2016). Theoretical Study of Decision Tree Algorithms to Identify Pivotal Factors for Performance Improvement: A Review. *International Journal of Computer Applications, 141*, 19-25.

Guleria, P., Thakur, N., และ Sood, M. (2014, 11-13 Dec. 2014). *Predicting student performance using decision tree classifiers and information gain.* Paper presented at the 2014 International Conference on Parallel, Distributed and Grid Computing.

hefinioanrhys. (2019, October 10, 2019). Support Vector Machines with the mlr package. https://www.r-bloggers.com/2019/10/support-vector-machines-with-the-mlr-package/

Hooshyar, D., และ Pedaste, M. (2019). Mining Educational Data to Predict Students' Performance through Procrastination Behavior. *Entropy, 22*, 12.

InnovativeLearningCenterSrinakharinwirotUniversity. (2017). SWU 141 ชีวิตในโลกดิจิทัล. https://course.ilc.swu.ac.th/

Jakkula, V. (2011). *Tutorial on Support Vector Machine (SVM)*.

Jayaprakash, S., Krishnan, S., และ Jaiganesh, V. (2020, 12-14 March 2020). *Predicting Students Academic Performance using an Improved Random Forest Classifier.* Paper presented at the 2020 International Conference on Emerging Smart Computing and Informatics (ESCI).

Liang, K., Zhang, Y., He, Y., Zhou, Y., Tan, W., และ Li, X. (2017). Online Behavior Analysis-Based Student Profile for Intelligent E-Learning. *Journal of Electrical and Computer Engineering, 2017*, 9720396.

MasterSoftwareSolutions. (2021). Machine Learning Solutions - Key Features to Look Into.

https://www.mastersoftwaresolutions.com/feature-for-machine-learning/

Mohamad, N., Ahmad, N., Jawawi, D., และ Mohd Hashim, S. (2020). Feature Engineering for Predicting MOOC Performance. *IOP Conference Series: Materials Science and Engineering, 884*, 012070.

Mueen, A., Zafar, B., และ Manzoor, U. (2016). Modeling and Predicting Students' Academic Performance Using Data Mining Techniques. *International Journal of Modern Education and Computer Science, 11*, 36-42.

Patel, H., และ Prajapati, P. (2018). Study and Analysis of Decision Tree Based Classification Algorithms. *International Journal of Computer Sciences and Engineering, 6*, 74-78.

Quinn, R., และ Gray, G. (2019). Prediction of student academic performance using Moodle data from a Further Education setting. *Irish Journal of Technology Enhanced Learning, 5*.

Ramaswami, G., Susnjak, T., Mathrani, A., และ Umer, R. (2020). Predicting Students Final Academic Performance using Feature Selection Approaches. *2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE)*, 1-5.

Riestra-González, M., Paule-Ruíz, M. d. P., และ Ortin, F. (2021). Massive LMS log data analysis for the early prediction of course-agnostic student performance. *Computers & Education, 163*, 104108.

Rodriguez, J. (2021). Two Minutes of Semi-Supervised Learning.

Sailotech. (2017, May 30, 2017). CRISP-DM methodology is used to develop predictive analytical models. https://twitter.com/sailotech/status/869440807794647040

Saraswat, M. (2016, December 20, 2016). Beginners Tutorial on XGBoost and Parameter Tuning in R. https://www.hackerearth.com/blog/developers/beginners-tutorial-on-xgboost-parameter-tuning-r/

Shamsi, M., และ Lakshmi, J. (2016). Student performance prediction using classification

data mining techniques.

Shrestha, S., และ Pokharel, M. (2021). Educational data mining in moodle data. *International Journal of Informatics and Communication Technology (IJ-ICT), 10*, 9.

Sun, J., Lang, J., Fujita, H., และ Li, H. (2018). Imbalanced enterprise credit evaluation with DTE-SBD: Decision tree ensemble based on SMOTE and bagging with differentiated sampling rates. *Information Sciences, 425*, 76-91.

Tan, M., และ Shao, P. (2015). Prediction of Student Dropout in E-Learning Program Through the Use of Machine Learning Method. *International Journal of Emerging Technologies in Learning (iJET), 10*.

Tas, S. (2021). Reinforcement Learning For Mice.

Widmann, M. (2019, May 27, 2019). From Modeling to Scoring: Confusion Matrix and Class Statistics. https://www.knime.com/blog/from-modeling-to-scoring-confusion-matrix-and-class-statistics

Wikipedia. (2021a, July 31, 2021). Logistic regression. https://en.wikipedia.org/wiki/Logistic_regression

Wikipedia. (2021b, August 2, 2021). Moodle. https://en.wikipedia.org/wiki/Moodle

Wikipedia. (2021c, July 12, 2021). Random forest. https://en.wikipedia.org/wiki/Random_forest

Wikipedia. (2021d, July 12, 2021). Support-vector machine. https://en.wikipedia.org/wiki/Support-vector_machine

Yan, K. (2021, 28-29 Jan. 2021). *Student Performance Prediction Using XGBoost Method from A Macro Perspective.* Paper presented at the 2021 2nd International Conference on Computing and Data Science (CDS).

Yan, M., Liu, K., Guan, Z., Xinkai, X., Qian, X., และ Bao, H. (2018). Background Augmentation Generative Adversarial Networks (BAGANs): Effective Data

Generation Based on GAN-Augmented 3D Synthesizing. *Symmetry, 10*, 734.

z_ai. (2020a, Mar 9, 2020). Decision Trees Explained. https://towardsdatascience.com/decision-trees-explained-3ec41632ceb6

z_ai. (2020b, Sep 17, 2020). Random Forest Explained. {z_ai, 2020 #82}

ชนัตถ์ พูนเดช. (2562). Life in a Digital World ชีวิตในโลกดิจิทัล. กรุงเทพฯ: สำนักนวัตกรรมการเรียนรู้ มหาวิทยาลัยศรีนครินทรวิโรฒ.

อสมา กุลวานิชไชยนันท์. (2561). Big Data Series II: Think Like a Data Scientist คิดแบบนักวิทยาศาสตร์ข้อมูล (พิมพ์ครั้งที่ 2). นนทบุรี: สำนักพิมพ์มหาวิทยาลัยสุโขทัยธรรมาธิราช.

อาณัติ รัตนถิรกุล. (2553). สร้างระบบ e-Learning ด้วย moodle ฉบับสมบูรณ์. กรุงเทพฯ: ซีเอ็ดยูเคชั่น, บมจ.

# ภาคผนวก

# มคอ 3

## รายละเอียดของรายวิชา ชีวิตในโลกดิจิทัล (มศว 141)

## หมวดวิชา ศึกษาทั่วไป

## สำนักนวัตกรรมการเรียนรู้ มหาวิทยาลัยศรีนครินทรวิโรฒ

## ภาคเรียนที่ 2 ประจำปีการศึกษา 2563

### หมวดที่ 1 ข้อมูลทั่วไป

**1. รหัสและชื่อรายวิชา**

มศว 141 ชีวิตในโลกดิจิทัล

SWU 141 Life in a digital world

**2. จำนวนหน่วยกิต**

3(3-0-6)

**3. หลักสูตรและประเภทของรายวิชา**

หลักสูตร หมวดวิชาศึกษาทั่วไป

ประเภทของ รายวิชาในกลุ่มวิชาบูรณาการ (วิทยาศาสตร์ คณิตศาสตร์ และเทคโนโลยี)

**4. อาจารย์ผู้รับผิดชอบรายวิชาและอาจารย์ผู้สอน**

อาจารย์ผู้รับผิดชอบรายวิชา อ.ดร.สุนทรี สกุลพราหมณ์

| | อาจารย์ผู้สอน | สถานที่ติดต่อ | Email |
|---|---|---|---|
| B01 | อ.วัชรพงศ์ แสงอ่อน | สำนักนวัตกรรมการเรียนรู้ | watcharapongs@g.swu.ac.th |
| B02 | ผศ.ดร.อนอมศักดิ์ ศรีจันทรา | สำนักนวัตกรรมการเรียนรู้ | tnomsak@g.swu.ac.th |
| B03 | อ.จาริณี มาเวหา | สำนักนวัตกรรมการเรียนรู้ | jariseem@g.swu.ac.th |
| B04 | ผศ.ดร.อนอมศักดิ์ ศรีจันทรา | สำนักนวัตกรรมการเรียนรู้ | tnomsak@g.swu.ac.th |
| B05 | อ.ภานุวัฒน์ บุตรเรียง | สำนักนวัตกรรมการเรียนรู้ | panuwat@g.swu.ac.th |
| B06 | อ.ภานุวัฒน์ บุตรเรียง | สำนักนวัตกรรมการเรียนรู้ | panuwat@g.swu.ac.th |
| B07 | รศ.ดร.ธนิตา เลิศพรกุลรัตน์ | สำนักนวัตกรรมการเรียนรู้ | thanital@g.swu.ac.th |
| B08 | อ.ดร.อรุโณทัย พยัคฆพงษ์ | สำนักนวัตกรรมการเรียนรู้ | aroonothai@g.swu.ac.th |
| B09 | อ.ดร.สุนทรี สกุลพราหมณ์ | สำนักนวัตกรรมการเรียนรู้ | soontaree@g.swu.ac.th |
| B10 | อ.ดร.อรุโณทัย พยัคฆพงษ์ | สำนักนวัตกรรมการเรียนรู้ | aroonothai@g.swu.ac.th |
| B11 | อ.สิริพันธุ์ พงศ์ประยูร | สำนักนวัตกรรมการเรียนรู้ | siriphan@g.swu.ac.th |
| B12 | อ.ดร.สราวุฒิ ตรีศรี | สำนักนวัตกรรมการเรียนรู้ | sarawut.trisri@g.swu.ac.th |
| B13 | อ.ภานุวัฒน์ บุตรเรียง | สำนักนวัตกรรมการเรียนรู้ | panuwat@g.swu.ac.th |
| B14 | อ.วัชรพงศ์ แสงอ่อน | สำนักนวัตกรรมการเรียนรู้ | watcharapongs@g.swu.ac.th |
| B15 | อ.สุเมธ โอสถานนท์ | สำนักนวัตกรรมการเรียนรู้ | sumateo@g.swu.ac.th |
| B16 | อ.ดร.อรุโณทัย พยัคฆพงษ์ | สำนักนวัตกรรมการเรียนรู้ | aroonothai@g.swu.ac.th |

| | อาจารย์ผู้สอน | สถานที่ติดต่อ | Email |
|---|---|---|---|
| B17 | อ.อธิปัตย์ สมิทธองค์ | สำนักนวัตกรรมการเรียนรู้ | athipat@g.swu.ac.th |
| B18 | ผศ.ดร.อนอมศักดิ์ ศรีจันทรา | สำนักนวัตกรรมการเรียนรู้ | tnomsak@g.swu.ac.th |
| B19 | อ.ภานุวัฒน์ บุตรเรียง | สำนักนวัตกรรมการเรียนรู้ | panuwat@g.swu.ac.th |
| B20 | อ.วัธรพงศ์ แสงอ่อน | สำนักนวัตกรรมการเรียนรู้ | watcharapongs@g.swu.ac.th |
| B21 | อ.ดร.เพดิจ เพชรออด | คณะแพทยศาสตร์ | Phadet2520@gmail.com |
| B22 | รศ.ดร.ชนัตย์ พูนเดช | สำนักนวัตกรรมการเรียนรู้ | chanutp@g.swu.ac.th |
| B24 | อ.ดร.สุนทรี สกุลพราหมณ์ | สำนักนวัตกรรมการเรียนรู้ | soontaree@g.swu.ac.th |
| B26 | รศ.ดร.ธนิตา เลิศพรกุลรัตน์ | สำนักนวัตกรรมการเรียนรู้ | thanital@g.swu.ac.th |
| B27 | อ.ดร.วุฒิพร เสือเมือง | สำนักนวัตกรรมการเรียนรู้ | S.wuttiporn@gmail.com |
| B29 | อ.ดร.สุนทรี สกุลพราหมณ์ | สำนักนวัตกรรมการเรียนรู้ | soontaree@g.swu.ac.th |
| B30 | อ.ดร.วุฒิพร เสือเมือง | สำนักนวัตกรรมการเรียนรู้ | S.wuttiporn@gmail.com |
| B31 | อ.สุเมธ โอสถานนท์ | สำนักนวัตกรรมการเรียนรู้ | sumateo@g.swu.ac.th |
| B32 | รศ.ดร.ชนัตย์ พูนเดช<br>รศ.ดร.ธนิตา เลิศพรกุลรัตน์ | สำนักนวัตกรรมการเรียนรู้ | chanutp@g.swu.ac.th<br>thanital@g.swu.ac.th |

**5. ภาคการศึกษา / ชั้นปีที่เรียน**

2/2563 ชั้นปีที่ 1

**6. รายวิชาที่ต้องเรียนมาก่อน (Pre-requisite) (ถ้ามี)**

- ไม่มี -

**7. รายวิชาที่ต้องเรียนพร้อมกัน (Co-requisite) (ถ้ามี)**

- ไม่มี -

**8. สถานที่เรียน**

มหาวิทยาลัยศรีนครินทรวิโรฒ ประสานมิตร และ มหาวิทยาลัยศรีนครินทรวิโรฒ องครักษ์
ระบบการจัดการเรียนการสอนแบบออนไลน์

**9. วันที่จัดทำหรือปรับปรุงรายละเอียดของรายวิชาครั้งล่าสุด**

14 ธันวาคม 2563

## หมวดที่ 2 จุดมุ่งหมายและวัตถุประสงค์

**1. วัตถุประสงค์ (หรือจุดมุ่งหมาย) ของรายวิชา**

1.1 เพื่อให้นิสิตมีความรู้และความเข้าใจพื้นฐานทางด้านกระบวนการสื่อสารและเทคโนโลยีในโลกดิจิทัล เพื่อเป็นพื้นฐานการเรียนรู้ และการแสวงหาความรู้ได้ด้วยตนเอง

1.2 เพื่อให้นิสิตมีทักษะขั้นพื้นฐานในการสืบค้นและประเมินสื่อสารสนเทศ การอ้างอิงข้อมูล และการนำเสนอสารสนเทศในรูปแบบต่าง ๆ

1.3 เพื่อให้นิสิตตระหนักถึงจริยธรรมและกฎหมายที่เกี่ยวข้องในการใช้เทคโนโลยีสารสนเทศและการสื่อสาร รวมทั้งภัยอันตรายในโลกดิจิทัลที่ส่งผลกระทบต่อบุคคลและสังคม พร้อมทั้งแนวทางการป้องกัน

1.4 เพื่อให้นิสิตเป็นพลเมืองดิจิทัลที่รู้เท่าทันสื่อ สารสนเทศ และเทคโนโลยี

**2. วัตถุประสงค์ในการพัฒนา/ ปรับปรุงรายวิชา**

2.1 เพื่อพัฒนาการจัดการเรียนรู้ในรูปแบบ Active Learning ที่เน้นนิสิตเป็นสำคัญ

2.2 เพื่อพัฒนานิสิตให้เป็นไปตามกรอบมาตรฐาน TQF

2.3 เพื่อส่งเสริมอัตลักษณ์นิสิตมหาวิทยาลัยศรีนครินทรวิโรฒ

## หมวดที่ 3 ลักษณะและการดำเนินการ

**1. คำอธิบายรายวิชา**

ศึกษาความสำคัญของกระบวนการสื่อสารและเทคโนโลยีในโลกดิจิทัล ทักษะการสืบค้น การประเมินสื่อสารสนเทศ การอ้างอิงข้อมูล จริยธรรมและกฎหมายที่เกี่ยวข้อง ภัยอันตรายในโลกดิจิทัลและแนวทางการป้องกัน การนำเสนอในรูปแบบต่าง ๆ การจัดการความรู้เพื่อการเรียนรู้ตลอดชีวิต และการรู้เท่าทันสื่อสารสนเทศและเทคโนโลยี

**2. จำนวนชั่วโมงที่ใช้ต่อภาคการศึกษา**

| บรรยาย | การเรียนรู้จากการฝึกปฏิบัติ | โครงงาน/กิจกรรมมอบหมาย/เรียนรู้ด้วยตนเอง |
|---|---|---|
| 45 | 0 | 90 |

**3. ความรับผิดชอบหลัก/ความรับผิดชอบรอง**

| รายวิชาศึกษาทั่วไป | ด้านที่ 1 คุณธรรมและจริยธรรม | | | ด้านที่ 2 ความรู้ | ด้านที่ 3 ทักษะทางปัญญา | | | | ด้านที่ 4 ทักษะความสัมพันธ์ระหว่างบุคคลและความรับผิดชอบ | | | ด้านที่ 5 การวิเคราะห์เชิงตัวเลข การสื่อสารและการใช้เทคโนโลยีสารสนเทศ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1.1 | 1.2 | 1.3 | 2.1 | 3.1 | 3.2 | 3.3 | 3.4 | 4.1 | 4.2 | 4.3 | 5.1 | 5.2 | 5.3 |
| มคว 141 | ● | ● | | ● | ● | ● | ● | ○ | ● | ● | ○ | ● | ● | ● |

**4. จำนวนชั่วโมงต่อสัปดาห์ที่อาจารย์ให้คำปรึกษาและแนะนำทางวิชาการแก่นักศึกษาเป็นรายบุคคล**

อาจารย์จัดเวลาให้คำปรึกษาและแนะนำทางวิชาการแก่นิสิตนอกชั้นเรียนเป็นรายบุคคล (เฉพาะรายที่ต้องการ) โดยการนัดหมายทางโทรศัพท์ สื่อสังคมออนไลน์ หรือ e-mail

มคอ. 3

## หมวดที่ 4 การพัฒนาผลการเรียนรู้ของนิสิต

| 1. คุณธรรม จริยธรรม : มีคุณธรรม จริยธรรมในการดำเนินชีวิต พร้อมให้ความช่วยเหลือเพื่อนมนุษย์ มีจิตสำนึกสาธารณะ และเป็นพลเมืองที่มีคุณค่าของสังคม | | |
|---|---|---|
| **ผลการเรียนรู้** | **วิธีการสอน** | **วิธีการประเมินผล** |
| 1.1 มีความซื่อสัตย์ มีวินัย ตรงต่อเวลา | - ฝึกปฏิบัติกิจกรรม เพื่อเสริมสร้างความซื่อสัตย์ มีวินัย และตรงต่อเวลา<br>- สอดแทรกจรรยาบรรณทางวิชาการและวัฒนธรรมองค์กร เพื่อให้นิสิตมีค่านิยมพื้นฐานที่ถูกต้อง<br>- ผู้สอนประพฤติตนเป็นแบบอย่างที่ดี<br>- ฝึกปฏิบัติทักษะการใช้สื่อออนไลน์บนพื้นฐานของมิติด้านจริยธรรม สิทธิและความรับผิดชอบในฐานะพลเมืองดิจิทัล (Digital citizenship) | สังเกตพฤติกรรม<br>- ความซื่อสัตย์ จากการสอบ การทำโครงงานเดี่ยวและโครงงานกลุ่ม การอ้างอิงแหล่งข้อมูลตามหลักและจรรยาบรรณทางวิชาการ<br>- การปฏิบัติตามระเบียบของมหาวิทยาลัยและข้อตกลงในชั้นเรียน<br>- การเข้าชั้นเรียนและการส่งงานตรงเวลา<br>- การใช้งานระบบจัดการเรียนการสอนของรายวิชาตามข้อตกลงของรายวิชา<br>(มีการกำหนดเกณฑ์มาตรฐาน Rubrics ในการประเมิน) |
| 1.2 มีจิตสาธารณะ เสียสละเพื่อส่วนรวม | - สอดแทรกกรณีศึกษาบุคคลตัวอย่าง เพื่อสร้างแรงบันดาลใจ เรื่อง การเสียสละ และจิตสาธารณะ<br>- ฝึกทำกิจกรรมที่มีส่วนร่วมรับผิดชอบ และเสียสละเพื่อส่วนรวม | - ประเมินจากการแสดงความคิดเห็นและสะท้อนคิดสิ่งที่ได้รับจากประสบการณ์ในการเรียนรู้<br>- ประเมินจากกิจกรรมและพฤติกรรมการมีส่วนร่วมในการทำงานกลุ่ม |
| 1.3 กระพือในคุณค่าของศิลปะและวัฒนธรรม | - | - |
| 2. ความรู้ : มีความรอบรู้อย่างกว้างขวาง เข้าใจ และเห็นคุณค่าของตนเองผู้อื่น สังคม และธรรมชาติ บนพื้นฐานปรัชญาเศรษฐกิจพอเพียง | | |
| **ผลการเรียนรู้** | **วิธีการสอน** | **วิธีการประเมินผล** |
| 2.1 มีความรู้พื้นฐานศึกษาทั่วไป | - อธิบายแนวคิดในการพัฒนามนุษย์ทั้งด้านพฤติกรรม จิตใจ ปัญญา<br>- จัดการเรียนรู้แบบ Active Learning<br>- จัดกิจกรรมเพื่อให้เข้าใจและเห็นคุณค่าของตนเอง ผู้อื่น และสังคม พร้อมตั้งเป้าหมายในการพัฒนาตนเอง สังคม และสิ่งแวดล้อม | - ประเมินจากการแสดงความคิดเห็นและสะท้อนคิดสิ่งที่ได้รับจากประสบการณ์ในการเรียนรู้<br>- ประเมินจากคุณภาพของโครงงาน<br>- ประเมินจากการสอบข้อเขียน<br>(มีการกำหนดเกณฑ์มาตรฐาน Rubrics ในการประเมิน) |

มคอ. 3

| 3. ทักษะทางปัญญา : ตระหนักถึงความเปลี่ยนแปลงของสรรพสิ่ง เรียนรู้ตลอดชีวิต สามารถบูรณาการความรู้สู่การคิดริเริ่มสร้างสรรค์และพัฒนาตนเองอย่างต่อเนื่อง | | |
|---|---|---|
| ผลการเรียนรู้ | วิธีการสอน | วิธีการประเมินผล |
| 3.1 สามารถประยุกต์ความรู้ให้เกิดประโยชน์<br>3.2 สามารถคิดอย่างเป็นระบบ และแก้ไขปัญหาได้<br>3.3 สามารถประเมินวิพากษ์ สถานการณ์ต่าง ๆ โดยใช้ความรู้เป็นฐาน | - มอบหมายกิจกรรมให้นิสิตฝึกประยุกต์ความรู้ให้เกิดประโยชน์<br>- ใช้กรณีศึกษาให้นิสิตฝึกคิดวิเคราะห์ วิพากษ์ และนำความรู้มาใช้ในการแก้ไขปัญหาอย่างเป็นระบบ<br>- นำเสนอ อภิปรายแลกเปลี่ยนความคิดเห็น และสะท้อนคิดในชั้นเรียน | - ประเมินจากการแสดงความคิดเห็นและสะท้อนคิดสิ่งที่ได้รับจากประสบการณ์ในการเรียนรู้<br>- ประเมินจากคุณภาพของโครงงาน<br>- ประเมินจากการสอบข้อเขียน<br>(มีการกำหนดเกณฑ์มาตรฐาน Rubrics ในการประเมิน) |
| 3.4 มีความคิดริเริ่มสร้างสรรค์นวัตกรรม | | |
| 4. ทักษะความสัมพันธ์ระหว่างบุคคลและความรับผิดชอบ : สามารถทำงานร่วมกับผู้อื่น มีความรับผิดชอบต่อตนเอง สังคม และสิ่งแวดล้อม | | |
| ผลการเรียนรู้ | วิธีการสอน | วิธีการประเมินผล |
| 4.1 สามารถทำงานร่วมกับผู้อื่นทั้งในฐานะผู้นำและผู้ร่วมงานได้<br>4.2 มีความรับผิดชอบต่อตนเอง สังคม และสิ่งแวดล้อม<br>4.3 สามารถปรับตัวให้เข้ากับสถานการณ์ต่าง ๆ | - ฝึกปฏิบัติทักษะการใช้ข้อมูลสารสนเทศออนไลน์ บนพื้นฐานของสิทธิและความรับผิดชอบของตนเองในฐานะพลเมืองดิจิทัล (Digital citizenship)<br>- มอบหมายกิจกรรมกลุ่ม ให้นิสิตฝึกทำงานร่วมกับผู้อื่น ฝึกรับผิดชอบต่อตนเอง สังคม สิ่งแวดล้อม และปรับตัวในสถานการณ์ต่าง ๆ<br>- ใช้กรณีศึกษาให้นิสิตเรียนรู้ และตระหนักถึงผลกระทบที่เกิดขึ้นจากการใช้เทคโนโลยีต่อตนเอง สังคม และสิ่งแวดล้อม พร้อมทั้งกระตุ้นให้คิดหาวิธีที่จะมีส่วนร่วมรับผิดชอบในการแก้ปัญหาโดยเริ่มต้นจากตัวนิสิตเอง<br>- นำเสนอ อภิปรายแลกเปลี่ยนความคิดเห็น และสะท้อนคิดในชั้นเรียน | - ประเมินจากคุณภาพของโครงงานและสื่อนำเสนอของนิสิต<br>- สังเกตจากพฤติกรรมการทำกิจกรรมกลุ่ม เช่น ภาวะผู้นำ/ผู้ร่วมงาน ความรับผิดชอบ การแสดงจุดยืนของตนเอง การรับฟังความคิดเห็นของเพื่อนร่วมกลุ่ม และค้นหาทางออกร่วมกันได้<br>- ประเมินจากการแสดงความคิดเห็นและสะท้อนคิดสิ่งที่ได้รับจากประสบการณ์ในการเรียนรู้<br>- ประเมินจากการปรับตัวให้เข้ากับสถานการณ์ต่าง ๆ ตามกาลเทศะ<br>- ให้นิสิต/ผู้ที่เกี่ยวข้องมีส่วนร่วมในการประเมินผล<br>- ประเมินจากคุณภาพของโครงงาน<br>- ประเมินจากการสอบข้อเขียน<br>(มีการกำหนดเกณฑ์มาตรฐาน Rubrics ในการประเมิน |

มคอ. 3

| 5. ทักษะการวิเคราะห์เชิงตัวเลข การสื่อสาร และการใช้เทคโนโลยีสารสนเทศ : สามารถวิเคราะห์ข้อมูลเชิงตัวเลข มีทักษะสื่อสาร พร้อมทั้งสามารถเลือกใช้เทคโนโลยีสารสนเทศอย่างเหมาะสม | | |
|---|---|---|
| ผลการเรียนรู้ | วิธีการสอน | วิธีการประเมินผล |
| 5.1 สามารถ วิเคราะห์ข้อมูลเชิงตัวเลข เพื่อให้เข้าใจองค์ความรู้หรือประเด็นปัญหา | - มอบหมายกิจกรรมให้นิสิตฝึกฝนการคิดวิเคราะห์ข้อมูลเชิงตัวเลขในการทำแบบฝึกหัดและโครงงาน<br>- ฝึกใช้ข้อมูลทางสถิติประกอบการวิเคราะห์ในการทำโครงงาน | - ประเมินจากการแสดงความคิดเห็นและสะท้อนคิดสิ่งที่ได้รับจากประสบการณ์ในการเรียนรู้<br>- ประเมินจากคุณภาพของโครงงาน<br>- ประเมินจากการสอบข้อเขียน (มีการกำหนดเกณฑ์มาตรฐาน Rubrics ในการประเมิน) |
| 5.2 สามารถสื่อสารได้อย่างมีประสิทธิภาพ | ฝึกการใช้ทักษะสื่อสารในรายวิชาทั้งการฟัง การอ่าน การพูด และการเขียน | - ประเมินทักษะการฟังและการอ่านจากการตั้งคำถามและตอบคำถาม<br>-ประเมินทักษะการพูดโดยพิจารณาจากการนำเสนอเป็นลำดับขั้นตอน พูดชัดเจน กระชับ ตรงประเด็น เข้าใจง่าย และรักษาเวลา<br>- ประเมินทักษะการเขียนจากคุณภาพโครงงานที่มีการเขียนเป็นลำดับขั้นตอน ชัดเจน ตรงประเด็น เข้าใจง่าย<br>- ประเมินจากการสอบข้อเขียน |
| 5.3 สามารถใช้เทคโนโลยีในการสืบค้น เก็บรวบรวมข้อมูล และเลือกใช้ได้เหมาะสมลักษณะการณ์อย่างมีประสิทธิภาพ | - ฝึกใช้เทคโนโลยีในการสืบค้นข้อมูลจากแหล่งที่เชื่อถือได้<br>- ฝึกเลือกใช้ข้อมูลสารสนเทศที่เหมาะสมประกอบการนำเสนองาน<br>- สอดแทรกหลักการเลือกแหล่งข้อมูลที่น่าเชื่อถือและวิธีการอ้างอิงแหล่งข้อมูลอย่างถูกต้องตามหลักและจรรยาบรรณทางวิชาการ | - ประเมินจากคุณภาพโครงงาน<br>- ประเมินจากการเลือกแหล่งข้อมูลได้เหมาะสมน่าเชื่อถือ และอ้างอิงแหล่งข้อมูลอย่างถูกต้องตามหลักและจรรยาบรรณทางวิชาการ |

มคอ. 3

## หมวดที่ 5 แผนการสอนและการประเมินผล

### 1. แผนการสอน

| สัปดาห์ | หัวข้อ/รายละเอียด | กิจกรรมการเรียนการสอน/สื่อที่ใช้ | งานมอบหมาย |
|---|---|---|---|
| 1<br>(4-8 ม.ค. 64) | 1. ปฐมนิเทศรายวิชา ชี้แจงรายละเอียดการเรียน<br>2. แนะนำการใช้ระบบการจัดการเรียนการสอนออนไลน์ในการเรียนรายวิชา ประกอบด้วย<br>- SWU Moodle<br>- ILC Moodle<br>- Microsoft TEAMS | จัดการเรียนการสอนแบบออนไลน์โดยมีรายละเอียด ดังนี้<br>1. ปฐมนิเทศ แนะนำรายวิชา ชี้แจงการเรียน วัตถุประสงค์ กิจกรรม และเกณฑ์การประเมินผลในรายวิชา<br>2. แนะนำการใช้ระบบเรียนรู้ออนไลน์ Microsoft TEAMS ประจำ Section ของตนเอง<br>3. แนะนำการใช้ระบบการจัดการเรียนการสอนของรายวิชา (LMS) http://course.ilc.swu.ac.th | 1. สมัครใช้บริการระบบจัดการเรียนการสอนของรายวิชา<br>2. เรียนรู้ด้วยตนเองและทำแบบฝึกหัดในหัวข้อ “พลเมืองดิจิทัลที่รู้เท่าทันสื่อสารสนเทศและเทคโนโลยี” และ “แหล่งสารสนเทศและการสืบค้น” จากเว็บไซต์ http://course.ilc.swu.ac.th |
| 2<br>(11-15 ม.ค. 64) | 1. พลเมืองดิจิทัลที่รู้เท่าทันสื่อ สารสนเทศ และเทคโนโลยี<br>2. แหล่งสารสนเทศและการสืบค้น | จัดการเรียนการสอนแบบออนไลน์โดยมีกิจกรรม Brainstorming/Student's reflection: อภิปรายแลกเปลี่ยนเรียนรู้ประสบการณ์ร่วมกัน ในหัวข้อ “พลเมืองดิจิทัลที่รู้เท่าทันสื่อ สารสนเทศ และเทคโนโลยี” และ “แหล่งสารสนเทศและการสืบค้น”<br>**วิดิทัศน์ประกอบการบรรยาย**<br>- Living in a Digital World<br>- การค้นหาข้อมูลบนอินเทอร์เน็ต<br>- 9 เครื่องมือลับ ที่ซ่อนอยู่ใน Google | กิจกรรมบนเว็บไซต์ http://course.ilc.swu.ac.th<br>1. ทำแบบฝึกหัดหัวข้อ “พลเมืองดิจิทัลที่รู้เท่าทันสื่อสารสนเทศและเทคโนโลยี” และ “แหล่งสารสนเทศและการสืบค้น” จากเว็บไซต์<br>2. เรียนรู้ด้วยตนเองในหัวข้อ<br>- การประเมินสารสนเทศ<br>- การวิเคราะห์และสังเคราะห์สารสนเทศ<br>- จริยธรรมในการใช้เทคโนโลยีสารสนเทศและการสื่อสาร |
| 3<br>(18-22 ม.ค. 64) | 1. การประเมินสารสนเทศ<br>2. การวิเคราะห์และสังเคราะห์สารสนเทศ | จัดการเรียนการสอนแบบออนไลน์โดยมีกิจกรรม Interactive Lecture/Student's reflection: อภิปรายแลกเปลี่ยนเรียนรู้ประสบการณ์ร่วมกัน ในหัวข้อ “การประเมินสารสนเทศ” และ “การวิเคราะห์และสังเคราะห์สารสนเทศ”<br>**วิดิทัศน์ประกอบการบรรยาย**<br>- เว็บปลอม เพจปลอม ข่าวปลอม ดูยังไง | กิจกรรมบนเว็บไซต์ http://course.ilc.swu.ac.th<br>1. ทำแบบฝึกหัดหัวข้อ<br>- การประเมินสารสนเทศ<br>- การวิเคราะห์และสังเคราะห์สารสนเทศ<br>2. เรียนรู้ด้วยตนเองในหัวข้อ “จริยธรรมในการใช้เทคโนโลยีสารสนเทศและการสื่อสาร” |

มคอ. 3

| สัปดาห์ | หัวข้อ/รายละเอียด | กิจกรรมการเรียนการสอน/สื่อที่ใช้ | งานมอบหมาย |
|---|---|---|---|
| 4<br>(28-29 ม.ค. 64) | จริยธรรมในการใช้เทคโนโลยีสารสนเทศและการสื่อสาร | จัดการเรียนการสอนแบบออนไลน์โดยมีกิจกรรม Interactive Lecture/Student's reflection: อภิปรายแลกเปลี่ยนเรียนรู้ประสบการณ์ร่วมกัน ในหัวข้อ "จริยธรรมในการใช้เทคโนโลยีสารสนเทศและการสื่อสาร"<br>**วิธีทัศน์ประกอบการบรรยาย**<br>- คดไลค์ คดแชร์ คิดกฎหมาย<br>- เตือนภัยชาวเน็ต การกระทำคิด พรบ. คอมพิวเตอร์ "โพสต์ต้องคิด คลิกเสี่ยงคุก" โดย บก.ปอท.<br>- พ.ร.บ.ลิขสิทธิ์ฉบับใหม่ ชาวเน็ตทำอะไรได้บ้าง<br>- รู้จักกับ Creative Common | กิจกรรมบนเว็บไซต์ http://course.ilc.swu.ac.th<br>1. ทำแบบฝึกหัดหัวข้อ "จริยธรรมในการใช้เทคโนโลยีสารสนเทศและการสื่อสาร"<br>2. เรียนรู้ด้วยตนเองในหัวข้อ "การลอกเลียนวรรณกรรมและแนวทางการหลีกเลี่ยงการลอกเลียนวรรณกรรมด้วยการเขียนอ้างอิงและบรรณานุกรม" |
| 5<br>(1-5 ก.พ. 64) | การลอกเลียนวรรณกรรมและแนวทางการหลีกเลี่ยงการลอกเลียนวรรณกรรมด้วยการเขียนอ้างอิงและบรรณานุกรม | จัดการเรียนการสอนแบบออนไลน์โดยมีกิจกรรม Interactive Lecture: อภิปรายแลกเปลี่ยนเรียนรู้ประสบการณ์ร่วมกัน ในหัวข้อ "การลอกเลียนวรรณกรรม และแนวทางการหลีกเลี่ยงการลอกเลียนวรรณกรรมด้วยการเขียนอ้างอิงและบรรณานุกรม"<br>**วิธีทัศน์ประกอบการบรรยาย**<br>- การหลีกเลี่ยงการลอกเลียนวรรณกรรม | กิจกรรมบนเว็บไซต์ http://course.ilc.swu.ac.th<br>1. ทำแบบฝึกหัดหัวข้อ "การลอกเลียนวรรณกรรมและแนวทางการหลีกเลี่ยงการลอกเลียนวรรณกรรมด้วยการเขียนอ้างอิงและบรรณานุกรม"<br>2. เรียนรู้ด้วยตนเองในหัวข้อ<br>- เทคโนโลยีสารสนเทศและการสื่อสารที่น่าสนใจในปัจจุบันและแนวโน้มในอนาคต<br>- ภัยอันตรายในโลกยุคดิจิทัล และแนวทางการป้องกัน |
| 6<br>(8-12 ก.พ. 64)<br><br>*12 ก.พ. 64 วันหยุดเนื่องในวันตรุษจีน | 1. เทคโนโลยีสารสนเทศและการสื่อสารที่น่าสนใจในปัจจุบันและแนวโน้มในอนาคต<br>2. ภัยอันตรายในโลกยุคดิจิทัล และแนวทางการป้องกัน | จัดการเรียนการสอนแบบออนไลน์โดยมีกิจกรรม Brainstorming/Student's reflection: อภิปรายแลกเปลี่ยนเรียนรู้ประสบการณ์ร่วมกัน ในหัวข้อ "เทคโนโลยีสารสนเทศและการสื่อสารที่น่าสนใจในปัจจุบันและแนวโน้มในอนาคต" และ "ภัยอันตรายในโลกยุคดิจิทัลและแนวทางการป้องกัน"<br>**วิธีทัศน์ประกอบการบรรยาย**<br>- อุทาหรณ์ โลกออนไลน์ หญิงสาวถูกหลอกทางเฟซบุ๊คให้โอนเงิน<br>- แฮกเกอร์ขโมยเงินจากธนาคาร<br>- จอมโจรในโลกไซเบอร์<br>- ภัยคุกคามทางอินเทอร์เน็ต<br>- PayPal ปลอม e-mail หัวขโมย!! | กิจกรรมบนเว็บไซต์ http://course.ilc.swu.ac.th<br>1. ทำแบบฝึกหัดหัวข้อ "ภัยอันตรายในโลกยุคดิจิทัล และแนวทางการป้องกัน"<br>2. ทำงานมอบหมายในหัวข้อ "การใช้งานเทคโนโลยีสารสนเทศในชีวิตประจำวัน" |

มคอ. 3

| สัปดาห์ | หัวข้อ/รายละเอียด | กิจกรรมการเรียนการสอน/สื่อที่ใช้ | งานมอบหมาย |
|---|---|---|---|
| | | - การดักจับ Packet บนเครือข่ายคอมพิวเตอร์<br>- Denial of Service หรือ DoS คืออะไร<br>- การโจมตีแบบ DDoS Attack<br>- คิดก่อนคลิก<br>- วิธีการป้องกันความปลอดภัยใน Facebook | |
| 7<br>(15-19 ก.พ. 64) | การประยุกต์ใช้เทคโนโลยีสารสนเทศในการศึกษา | จัดการเรียนการสอนแบบออนไลน์โดยมีกิจกรรม ดังนี้<br>1. Brainstorming/Student's reflection: อภิปรายแลกเปลี่ยนเรียนรู้ประสบการณ์ร่วมกัน ในหัวข้อ "การประยุกต์ใช้เทคโนโลยีสารสนเทศในการศึกษา"<br>2. อธิบายงานมอบหมาย เรื่อง การเขียนบทความ<br>**วีดิทัศน์ประกอบการบรรยาย**<br>- การเขียนบทความ | กิจกรรมบนเว็บไซต์ http://course.ilc.swu.ac.th<br>- ทำแบบฝึกหัดในหัวข้อ "การประเมินบทความ" |
| 8<br>(22-26 ก.พ. 64)<br>*26 ก.พ. 64 วันพุธเฉลิมวันมาฆบูชา | ประมวลความรู้ทั้งหมด | จัดการเรียนการสอนแบบออนไลน์โดยมีกิจกรรม Brainstorming/Student's reflection/World cafe: ประมวลความรู้ | กิจกรรมบนเว็บไซต์ http://course.ilc.swu.ac.th<br>ทำกิจกรรม Peer Assessment ในหัวข้อ "การใช้งานเทคโนโลยีสารสนเทศในชีวิตประจำวัน" |
| สอบกลางภาค มศว 141 วันอังคารที่ 2 มีนาคม 2564 เวลา 13.00 – 15.00 น. | | | |
| 9<br>(8-12 มี.ค. 64) | การประยุกต์ใช้เทคโนโลยีสารสนเทศในการศึกษา (ต่อ) | จัดการเรียนการสอนแบบออนไลน์โดยมีกิจกรรม Task-based Learning: บทความวิชาการ เรื่อง เทคโนโลยีดิจิทัลที่น่าสนใจ<br>**วีดิทัศน์ประกอบการบรรยาย**<br>- การตรวจบทความ | กิจกรรมบนเว็บไซต์ http://course.ilc.swu.ac.th<br>ส่งงานส่วนบุคคล "บทความวิชาการ เรื่อง เทคโนโลยีดิจิทัลที่น่าสนใจ" |
| 10<br>(15-19 มี.ค. 64) | การนำเสนอสารสนเทศในรูปแบบต่าง ๆ | จัดการเรียนการสอนแบบออนไลน์โดยมีกิจกรรม ดังนี้<br>1. Interactive Lecture: อภิปรายแลกเปลี่ยนเรียนรู้ประสบการณ์ร่วมกัน ในหัวข้อ "การนำเสนอสารสนเทศในรูปแบบต่าง ๆ"<br>2. นิสิตฝึกปฏิบัติการนำเสนอสารสนเทศ | กิจกรรมบนเว็บไซต์ http://course.ilc.swu.ac.th<br>- ทำแบบฝึกหัดในหัวข้อ MS Word และ MS Excel |

มคอ. 3

| สัปดาห์ | หัวข้อ/รายละเอียด | กิจกรรมการเรียนการสอน/สื่อที่ใช้ | งานมอบหมาย |
|---|---|---|---|
| 11<br>(22-26 มี.ค. 64) | นิสิตนำเสนอหัวข้อโครงงานและวิธีการดำเนินการ | จัดการเรียนการสอนแบบออนไลน์โดยมีกิจกรรม ดังนี้<br>1. Student's reflection: อภิปราย และซักถามข้อสงสัย<br>2. นิสิตทุกกลุ่มนำเสนอหัวข้อ และวิธีการดำเนินการโครงงาน<br>3. อธิบายกิจกรรม Peer Assessment ในหัวข้อ "การประเมินบทความวิชาการ" | ปรับปรุงโครงงาน และประเมินด้วยกระบวนการสะท้อนคิด (Reflection Process) |
| 12<br>(29 มี.ค. - 2 เม.ย. 64) | 1. การติดตามโครงงานและวิธีการดำเนินการ<br>2. นิสิตศึกษาปฏิบัติการประเมินโครงงานบทความวิชาการ | นิสิตทำกิจกรรม Peer Assessment ในหัวข้อ "การประเมินบทความวิชาการ" | กิจกรรม Peer Assessment ในหัวข้อ "การใช้งานเทคโนโลยีสารสนเทศในชีวิตประจำวัน" การประเมินบทความตาม Rubrics ที่กำหนดจากเว็บไซต์ http://course.ilc.swu.ac.th |
| 13<br>(5-9 เม.ย. 64)<br>*6 เม.ย. 64 วันหยุดเนื่องในวันจักรี | การติดตามโครงงานกลุ่ม | กิจกรรม Student's reflection: อภิปรายและซักถามข้อสงสัยเกี่ยวกับโครงงานกลุ่ม (บทความวิชาการและสื่อนำเสนอ) | ปรับปรุงโครงงานตามข้อเสนอแนะ<br>*หมายเหตุ การดำเนินกิจกรรมจัดขึ้นในชั้นเรียนหรือออนไลน์ (ขึ้นอยู่กับสถานการณ์) |
| สัปดาห์วันหยุดสงกรานต์ | | | |
| 14<br>(19-23 เม.ย. 64) | นิสิตนำเสนอโครงงานกลุ่ม | จัดการเรียนการสอนแบบออนไลน์โดยมีกิจกรรม Task-based Learning: การนำเสนอบทความวิชาการ | กิจกรรม Peer Assessment ในหัวข้อ "การประเมินโดยเพื่อน"<br>*หมายเหตุ การนำเสนอผลงานจัดขึ้นในชั้นเรียนหรือออนไลน์ (ขึ้นอยู่กับสถานการณ์) |
| 15<br>(26-30 เม.ย. 64) | นิสิตนำเสนอโครงงานกลุ่ม | จัดการเรียนการสอนแบบออนไลน์โดยมีกิจกรรม ดังนี้<br>1. Task-based Learning: การนำเสนอบทความวิชาการ<br>2. นิสิตทำแบบประเมิน ILC001-003 | กิจกรรม Peer Assessment ในหัวข้อ "การประเมินโดยเพื่อน"<br>*หมายเหตุ การนำเสนอผลงานจัดขึ้นในชั้นเรียนหรือออนไลน์ (ขึ้นอยู่กับสถานการณ์) |

มคอ. 3

2. แผนการประเมินผลการเรียนรู้

| สาระการประเมิน | การเรียนรู้ | วิธีการประเมิน | สัดส่วนของการประเมิน |
|---|---|---|---|
| ภาคเนื้อหา | 2.1 มีความรู้พื้นฐานศึกษาทั่วไป<br>3.3 สามารถประเมิน วิพากษ์ สถานการณ์ต่าง ๆ โดยใช้ความรู้เป็นฐาน<br>5.1 สามารถวิเคราะห์ข้อมูลเชิงตัวเลข เพื่อให้เข้าใจองค์ความรู้หรือประเด็นปัญหา | 1. การประเมินงานมอบหมาย<br>- แบบฝึกหัดออนไลน์<br>- กิจกรรมแต่ละบทเรียน<br>2. การสอบกลางภาค | <br>15%<br>15%<br>30% |
| ภาคกิจกรรม<br>- โครงงาน | 1.1 ซื่อสัตย์ มีวินัย ตรงต่อเวลา<br>1.2 มีจิตสำนึกสาธารณะ เสียสละเพื่อส่วนรวม<br>2.1 มีความรู้พื้นฐานศึกษาทั่วไป<br>3.1 สามารถประยุกต์ความรู้ให้เกิดประโยชน์<br>3.2 คิดวิเคราะห์อย่างเป็นระบบและแก้ไขปัญหาได้<br>3.4 มีความคิดริเริ่ม สร้างสรรค์นวัตกรรม<br>4.1 ทำงานร่วมกับผู้อื่นในฐานะผู้นำและผู้ร่วมงานได้<br>4.2 มีความรับผิดชอบต่อตนเอง สังคมและสิ่งแวดล้อม<br>4.3 สามารถปรับตัวให้เข้ากับสถานการณ์ต่าง ๆ<br>5.2 สื่อสารได้อย่างมีประสิทธิภาพ<br>5.3 ใช้เทคโนโลยีในการสืบค้น เก็บรวบรวมข้อมูล และเลือกใช้ได้เหมาะสมกับสถานการณ์ | 1. การประเมินโครงงานด้วยการสะท้อนคิด<br>(Reflection process)<br>- ครั้งที่ 1<br>บทความวิชาการเดี่ยว<br>- ครั้งที่ 2<br>บทความวิชาการกลุ่ม<br>2. การนำเสนอโครงงาน<br>- สื่อและการนำเสนอ | <br><br><br><br><br><br><br><br><br>10%<br><br>12%<br><br>8% |
| ภาคพฤติกรรมการมีส่วนร่วม | 1.1 ซื่อสัตย์ มีวินัย ตรงต่อเวลา<br>1.2 มีจิตสำนึกสาธารณะ เสียสละเพื่อส่วนรวม<br>4.1 ทำงานร่วมกับผู้อื่นในฐานะผู้นำและผู้ร่วมงานได้<br>4.2 มีความรับผิดชอบต่อตนเอง สังคมและสิ่งแวดล้อม | 1. จิตพิสัย<br>2. การประเมินโดยเพื่อน<br>(Peer Assessment) | 5%<br>5% |
| | | รวม | 100% |

## หมวดที่ 6 ทรัพยากรประกอบการเรียนการสอน

### 1. ตำราและเอกสารหลัก

ชนัดดา พูนเดช. (2561). *ชีวิตในโลกดิจิทัล*. กรุงเทพฯ: สำนักนวัตกรรมการเรียนรู้ มหาวิทยาลัยศรีนครินทรวิโรฒ.

ธนิดา เลิศพรกุลรัตน์. (2561). *การประยุกต์ใช้ทักษะการรู้สารสนเทศในการเขียนบทความวิชาการ*. กรุงเทพฯ: สำนักนวัตกรรมการเรียนรู้ มหาวิทยาลัยศรีนครินทรวิโรฒ.

Alvermann, D. E. (2002). *Adolescents and Literacies in a Digital World*: P. Lang.

Jacob, H. H. (2014). Mastering Digital Literacy. Indiana: Solution Tree Press.

Scheyen, C. (2015). *Beyond Media Literacy*: New Paradigms in Media Education: Five Rivers Chapmanry.

Shelly, G., & Vermaat, M. (2011). *Discovering Computers, Introductory: Your Interactive Guide to the Digital World*: Cengage Learning.

Solomon, A., Wilson, G., & Taylor, T. (2011). *100% Information Literacy Success*: Cengage Learning.

White, J. (2015). Digital Literacy Skills for FE Teachers. London: Learning Matters.

### 2. เอกสารและข้อมูลสำคัญ

- ไม่มี -

### 3. เอกสารและข้อมูลแนะนำ

- ไม่มี -

## หมวดที่ 7 การประเมินและปรับปรุงการดำเนินการของรายวิชา

**1. กลยุทธ์การประเมินประสิทธิผลของรายวิชาโดยนิสิต**

1.1 การสนทนากลุ่มระหว่างผู้สอนและผู้เรียน

1.2 การสะท้อนคิดจากพฤติกรรมของผู้เรียน

1.3 การประเมินจิตพิสัย แบบฝึกหัด และงานมอบหมาย

1.4 การประเมินโครงงาน

**2. กลยุทธ์การประเมินการสอน**

2.1 การสังเกตพฤติกรรมของผู้เรียน

2.2 ผลสัมฤทธิ์ทางการเรียน

2.3 แบบประเมินผู้สอนและแบบประเมินความพึงพอใจต่อรายวิชา

**3. การปรับปรุงการสอน**

3.1 นำผลการประเมินรายวิชาโดยนิสิต และผลการประชุมระหว่างภาคและปลายภาคของผู้สอน มาพิจารณาและหาแนวทางปรับปรุงวิธีการสอนและการเรียนรู้ของนิสิต

3.2 จัดประชุมระหว่างอาจารย์ผู้สอนเพื่อพัฒนารูปแบบการจัดการเรียนรู้ของรายวิชา

**4. การทวนสอบมาตรฐานผลสัมฤทธิ์ของนิสิตในรายวิชา**

4.1 ประชุมพิจารณาการออกข้อสอบให้เป็นไปตามวัตถุประสงค์การเรียนรู้

4.2 ประชุมติดตามผลการจัดการเรียนการสอนระหว่างภาคให้เป็นไปตาม มคอ. 3

4.3 ประชุมพิจารณาการให้คะแนนข้อสอบของอาจารย์แต่ละท่าน ให้เป็นไปตามเกณฑ์การประเมิน (Rubrics) เพื่อให้มีมาตรฐานการให้คะแนนใกล้เคียงกัน

**5. การดำเนินการทบทวนและการวางแผนปรับปรุงประสิทธิผลของรายวิชา**

5.1 การปรับปรุงตามข้อเสนอแนะและการวิพากษ์ของคณาจารย์เพื่อปรับปรุงประสิทธิผลของรายวิชา

5.2 การปรับปรุงตามข้อเสนอแนะของนิสิตเพื่อปรับปรุงประสิทธิผลของรายวิชา

# ประวัติผู้เขียน

| | |
|---|---|
| **ชื่อ-สกุล** | ปริยานุช ประเสริฐสิริกุล |
| **วัน เดือน ปี เกิด** | 21 เมษายน 2535 |
| **สถานที่เกิด** | กรุงเทพมหานคร |
| **วุฒิการศึกษา** | พ.ศ.2556 วิทยาศาสตรบัณฑิต (วท.บ.) เกียรตินิยมอันดับ 1<br>สาขาวิชาคณิตศาสตร์ จากมหาวิทยาลัยศรีนครินทรวิโรฒ |
| **ที่อยู่ปัจจุบัน** | 211/9 หมู่บ้านนนทรีแกลลอรี่ ถนนนครอินทร์ ตำบลบางไผ่<br>อำเภอเมืองนนทบุรี จังหวัดนนทบุรี 11000 |
| **ผลงานตีพิมพ์** | ปริยานุช ประเสริฐสิริกุล. (2020). การพัฒนาระบบสนับสนุนการตัดสินใจสำหรับนิสิตเพื่อวางแผนการเรียนในหมวดวิชาศึกษาทั่วไป. วารสาร Mahidol R2R e-Journal, 7(2), 102-114. |