

Recommendation System for CE Curriculum administrators

ระบบแนะนำสำหรับการบริหารหลักสูตร
วิศวกรรมคอมพิวเตอร์



The Team



62010299

ณิชาภัณต์ สุขุมจิตพิทโยทัย



The Team



62010453

มนทพันธุ์ รุจิรภา



The Team



62010465
ນຣົງຊຸນ ອຢູ່ບັ້ວ



Topic

→ **Motivation**

ที่มาและความสำคัญ

→ **Objective**

วัตถุประสงค์

→ **Theoretical Background**

ทฤษฎีที่เกี่ยวข้อง

→ **Related Work**

งานวิจัยที่เกี่ยวข้อง

→ **Scope**

ขอบเขตของโครงการ

→ **Project Development**

การพัฒนาโครงการ

→ **Experiment**

ผลการศึกษาเทคโนโลยีที่ใช้พัฒนา

→ **Expected Benefit**

ประโยชน์ที่คาดว่าจะได้รับ





Motivation

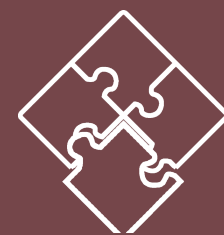
ที่มาและความสำคัญ



ปัญหาที่พวกเราสังเกตเห็น



การนำข้อมูลผลการเรียนของนักศึกษาในอดีตมาใช้ประโยชน์ในการบริหารหลักสูตร และนำมาวิเคราะห์ผล เพื่อช่วยในการวางแผนการเรียนของนักศึกษา จะสามารถช่วยอาจารย์และบุคลากรที่เกี่ยวข้องกับการ ศึกษาในด้านของการบริหารหลักสูตร เพื่อวางแผนการเพิ่มหรือลดจำนวนผู้เรียนในรายวิชาต่าง ๆ ซึ่งส่งผลต่อการผลิตบัณฑิตด้านต่าง ๆ ได้



Solution !!!

ดังนั้นผู้จัดทำจึงได้เห็นถึงความสำคัญการประเมินสถานะของหลักสูตร ของระบบแนะนำการวางแผน การคาดการณ์ จากการใช้ความรู้ทางด้าน **Data Analytics, Prediction** และระบบ **Data Visualization** โดยใช้ข้อมูลผลการเรียนของนักศึกษาในอดีต เพื่อพัฒนาระบบช่วยเหลือ และตอบ โจทย์ให้แก่นักศึกษาและบุคลากรทางการศึกษาหรือบุคคลที่เกี่ยวข้องได้





Objective

วัตถุประสงค์



Objective



1. เพื่อนำข้อมูลของผลการเรียนของนักศึกษาในอดีต และข้อมูลจากแบบสำรวจการปฏิบัติงานทำของบัณฑิตมาใช้ในการวางแผนการเรียนหรือประเมินอาชีพในอนาคตของนักศึกษาได้



2. ประมวลผลข้อมูลผลการเรียนของนักศึกษาในอดีต และข้อมูลจากแบบสำรวจการปฏิบัติงานทำของบัณฑิต และทำแผนภาพกราฟิกเพื่อนำเสนอข้อมูลอำนวยความสะดวกในการบริหารหลักสูตร เช่น ถ้าต้องการผลิตบัณฑิตจำนวน X คน จะต้องเปิดวิชาอะไร จำนวนเท่าไร



Objective



3. เพื่อนำข้อมูลผลการเรียนของนักศึกษาในอดีตมาพัฒนาเป็นระบบแนะนำและวางแผนการเรียนของนักศึกษาได้

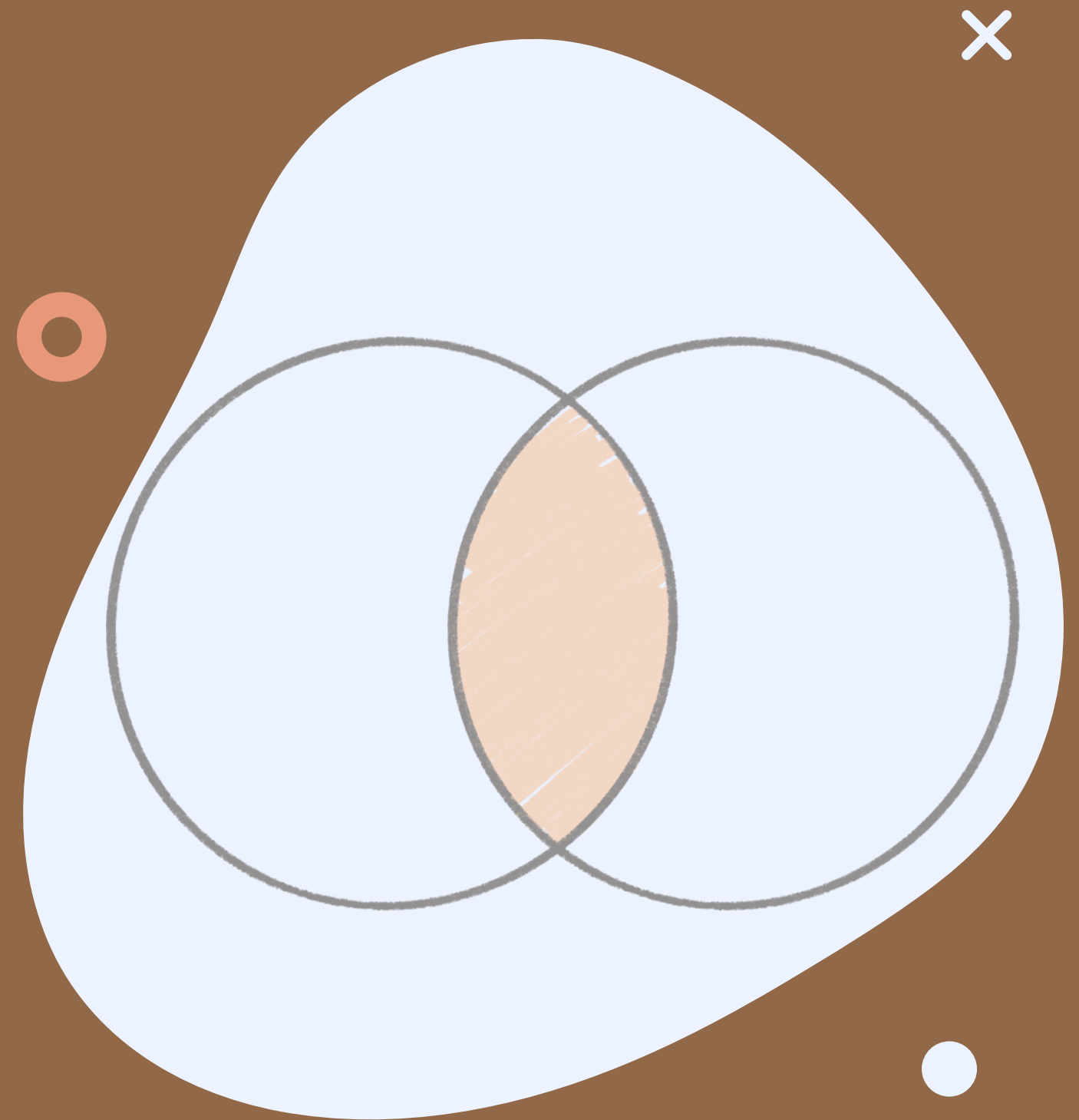


4. เพื่อนำข้อมูลการพยากรณ์อาชีพในอนาคตของนักศึกษาในสถาบันมาแสดงเป็นแผนภาพกราฟิกในการอำนวยความสะดวกให้หน่วยงานภายนอกได้รับทราบว่าหลักสูตรปัจจุบันของสถาบันสามารถผลิตบุคลากรที่มีความชำนาญด้านใดได้บ้าง



Theoretical Background

កុមារភូមិភាគខាង



Recommendation System

○ What is it ?

ระบบที่จะแนะนำสิ่ง (item) ที่ “เหมาะสม” ให้กับผู้ใช้ โดย item เป็นได้ตั้งแต่ ข่าว, เนื้อหา, เพลง, course เรียน, โดยระบบสามารถให้คำแนะนำที่ต่างกันตามความชอบของแต่ละผู้ใช้ได้

○ ประเภท

Content-based Filtering
Collaborative Filtering
Hybrid system



Classification and Prediction

○ What is it ?

เทคนิคการจำแนกประเภทข้อมูล โดยสร้างโมเดลจัดการข้อมูลให้อยู่ในกลุ่มที่กำหนดมาให้ จากกลุ่มตัวอย่างข้อมูลที่เรียกว่าข้อมูลสอนระบบ ที่แต่ละแถวของข้อมูลประกอบด้วยฟิลด์ หรือแอตทริบิวต์ จำนวนมาก

○ ประเภท

Binary classification
Multi-Class Classification
Multi-Label Classification
Imbalanced Classification



Decision Tree

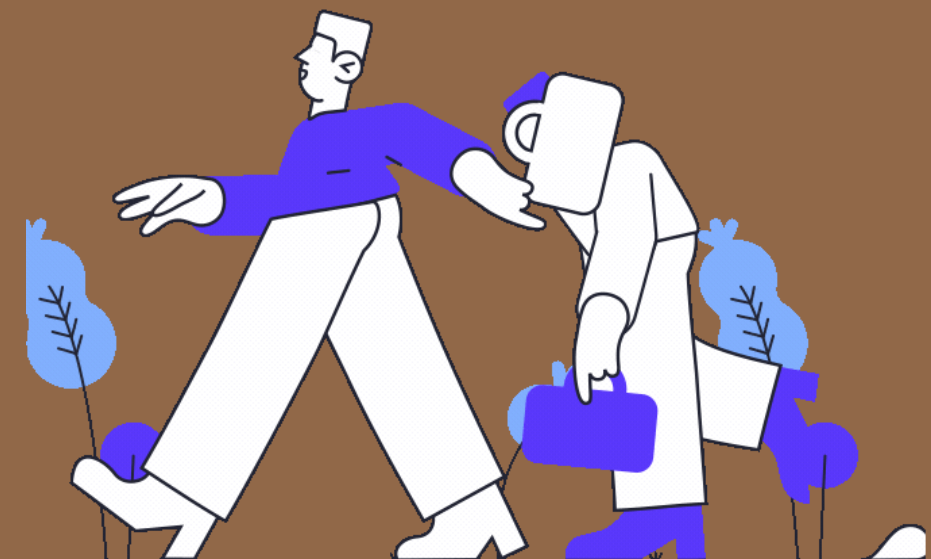
○ What is it ?

เป็นหนึ่งในวิธีการเรียนรู้ซึ่งใช้ในสถิติ, การเรียนรู้ของเครื่อง และการทำเหมืองข้อมูล โดยพิจารณาการสังเกตการแบ่งแยกข้อมูล โดยพิจารณาจากลักษณะของวัตถุ

○ ประเภท

Classification trees

Regression trees



K-Nearest Neighbors (KNN)



○ What is it ?

เป็นวิธีที่ใช้ในการจัดแบ่งคลาส โดยเทคนิคนี้จะตัดสินใจว่า คลาสใดที่จะแทนเงื่อนไขหรือกรณีใหม่ๆ ได้บ้าง โดยการตรวจสอบจำนวนบางจำนวน ของกรณีหรือเงื่อนไขที่เหมือนกัน หรือ ใกล้เคียงกันมากที่สุด

โดยจะหาผลรวม (Count Up) ของจำนวนเงื่อนไข หรือกรณีต่างๆ สำหรับแต่ละคลาส และกำหนดเงื่อนไขใหม่ๆ ให้คลาสที่เหมือนกันกับคลาสที่ใกล้เคียงกันมากที่สุด













Memory-based

○ What is it ?

memory-based นั้น จะใช้หลักการของ nearest neighbor คือ เก็บข้อมูลการชอบของผู้ใช้ทั้งหมดมา แล้วดูว่ามีผู้ใช้คนใด หรือ สินค้าใด ที่มีลักษณะคล้ายกันบ้างจากข้อมูลนั้น

○ Example

ตัวอย่างของการใช้ memory-based เช่น การใช้ nearest neighbor คือการคาดเดา rating จาก การดูจากผู้ใช้ (user-based) หรือ สินค้า (item-based) ที่ใกล้เคียงกัน

		1	2	3	4	5
						
A		7.00	6.00	6.80		8.00
B		8.30	6.00	8.00	7.40	7.00
C			7.00		8.00	9.50
D		8.00	6.00	7.75	?	7.50
E			6.50	5.50	7.75	



Naive Bayesian Classifier



○ What is it ?

คือโมเดลการจำแนกประเภทข้อมูลที่ใช้หลักความน่าจะเป็นซึ่งอยู่บนพื้นฐานของ Bayes' Theorem และแนวคิดของ Bayes' Theorem นั้นเราสามารถทำนายเหตุการณ์ที่พิจารณาได้จากการเกิดของเหตุการณ์ต่างๆ ได้ดังสมการ

$$P(h|D) = [P(D|h) * P(h)]/P(D)$$



Visualization

○ What is it ?

การวิเคราะห์ด้วยภาพ ซึ่งว่าด้วยการมองดูข้อมูลในเชิงปริมาณหรือปริมาตรแล้วสามารถสกัดเอาความรู้หรือข้อสรุปออกมาได้ ซึ่งมี Interactive Visual Interfaces เป็นเครื่องมือและสื่อกลาง

○ Advantage

- ลดการทำงานในส่วนความจำ โดยใช้ภาพที่มองเห็นด้วยตาเปล่าเป็นสื่อรับรู้
- ลดการค้นหากที่หนักและซับซ้อน โดยการนำเสนอข้อมูลที่มีขนาดใหญ่ในพื้นที่แสดงขนาดเล็ก
- ใช้รูปแบบเฉพาะเพื่อให้เราสามารถเข้าใจความหมายได้อย่างรวดเร็ว
- ช่วยในการอนุมานความสัมพันธ์ให้ง่ายขึ้น
- ติดตามเหตุการณ์หรือปรากฏการณ์ในข้อมูลหลายเรื่อง ๆ ให้สามารถมองเห็นได้ด้วยตาเปล่า
- มีเครื่องมือเพื่อให้จัดการเปลี่ยนแปลงค่าพารามิเตอร์ต่าง ๆ



Scikit-learn

○ What is it ?

เป็นโมดูลหนึ่งของภาษา Python เป็นแพ็คเกจที่รวบรวม Library ด้าน การเรียนรู้ของเครื่อง (Machine Learning) เอาไว้ และถูกออกแบบมาให้ทำงานร่วมกับ Library ของภาษา Python อย่าง NumPy และ SciPy ได้ดี

Tools

1. Classification
2. Regression
3. Clustering
4. Model selection
5. Pre-processing





Related Work

งานวิจัยที่เกี่ยวข้อง



Development and Analysis of an Online Computer Books Recommendation System Using Content-Based Filtering

วิเคราะห์และพัฒนาระบบแนะนำหนังสือคอมพิวเตอร์แบบออนไลน์โดยใช้เทคนิคการกรองแบบอิงเนื้อหา

งานวิจัยนี้คือการพัฒนาาระบบแนะนำหนังสือคอมพิวเตอร์แบบออนไลน์โดยใช้เทคนิคการกรองแบบอิงเนื้อหาเพื่อประเมินความพึงพอใจของการพัฒนาระบบแนะนำหนังสือคอมพิวเตอร์แบบออนไลน์ที่พัฒนาขึ้น



A Prediction System for Undergraduate Student Dropout at Faculty of Science, Buriram Rajabhat University using Data Mining Techniques

ระบบทำนายการพ้นสภาพของนักศึกษาระดับปริญญาตรี คณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏบุรีรัมย์ ด้วยเทคนิคการทำเหมืองข้อมูล

งานวิจัยนี้คือการ สร้างแบบจำลองการทำนายการพ้นสภาพของนักศึกษา และ พัฒนาระบบทำนายการพ้นสภาพด้วยเทคนิคเหมืองข้อมูล โดย ใช้คลังข้อมูล จากฐานข้อมูลนักศึกษาคณะที่มีอัตราการพ้นสภาพมากที่สุด นำมาเข้ากระบวนการเตรียมข้อมูล และจำแนกประเภทโดย อัลกอริทึม J48 และวัดประสิทธิภาพโดยวิธีการตรวจสอบแบบไขว้



The Comparison of Performance Models for Predicting Students Success in Vocation Education

การเปรียบเทียบประสิทธิภาพของแบบจำลองในการพยากรณ์ ความสำเร็จการศึกษาของนักเรียนระดับประกาศนียบัตรวิชาชีพ

- งานวิจัยนี้คือการ ศึกษาปัจจัยที่มีผลต่อการ ความสำเร็จการศึกษาในสถานศึกษา และ พัฒนา แบบจำลองที่มีประสิทธิภาพในการพยากรณ์ เพื่อ อำนวยความสะดวกให้ครูแนะแนวและผู้บริหารในการวางแผนการแนะแนว



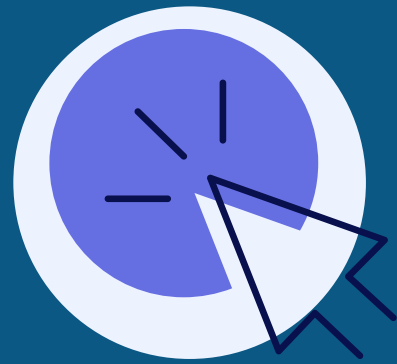


Scope

ขอบเขตของโครงการ



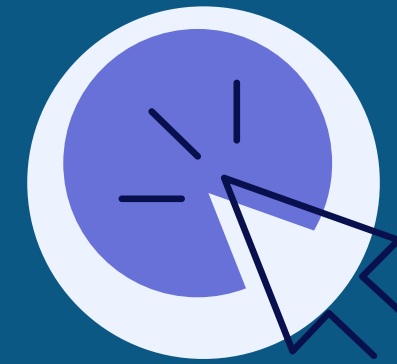
Scope



ข้อมูลผลการเรียนในอดีตจะได้จากสำนัก
ทะเบียนและประมวลผล โดยกรรมการ
หลักสูตรจะเป็นผู้ร้องขอข้อมูลดังกล่าวและ
นำเข้าระบบ

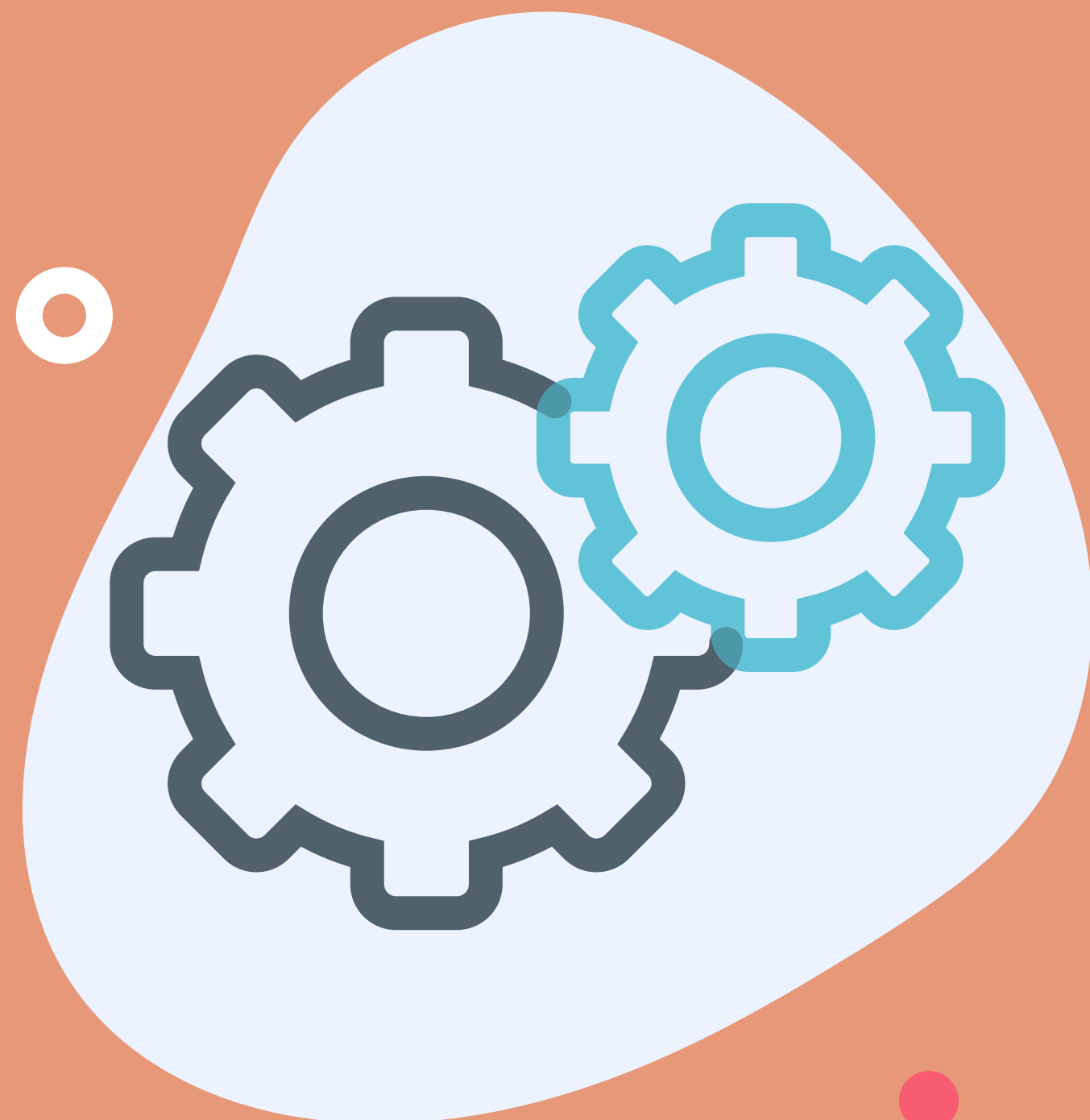


การดำเนินงานต่าง ๆ จะใช้ข้อมูลเพียง 2 แหล่ง
คือข้อมูลผลการเรียนของนักศึกษาจาก
สำนักทะเบียนและประมวลผล และ
แบบสอบถามการปฏิบัติงานทำของบัณฑิต
เท่านั้น



Algorithm ในระบบทั้งหมดจะใช้
Algorithm พื้นฐานตาม library ของ
scikit-learn เท่านั้น ไม่มีการประยุกต์ใช้
จาก research paper อื่น ๆ





Project Development

การพัฒนาโครงการ

Development

การพัฒนาโครงการ

Step 1

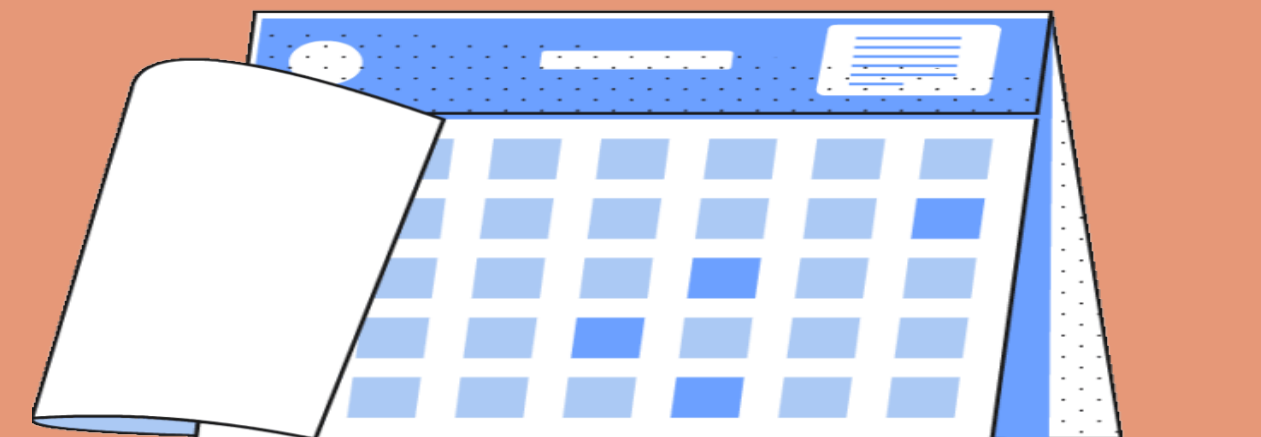
กำหนดขอบเขต เป้าหมาย และ
วัตถุประสงค์ในการทำโครงการ

Step 2

กำหนดขอบเขต เป้าหมาย
และวัตถุประสงค์ในการทำ
โครงการ

Step 3

วางแผนการพัฒนาระบบ



Development

การพัฒนาโครงงาน

ออกแบบระบบที่จะพัฒนา

Step 4

ทดสอบการทำงานในแต่ละส่วน

Step 5

รวมแต่ละส่วนของระบบเข้าด้วยกัน

Step 6

ทดสอบการทำงานร่วมกันของระบบ

Step 7



ส่วนการทำงานของระบบ

- ส่วนของการรับข้อมูลจาก User
- ส่วนของเซิร์ฟเวอร์สำหรับการประมวลผล แบบจำลองพยากรณ์
- ส่วนของเซิร์ฟเวอร์สำหรับดูแลรับส่งข้อมูลของ ผู้ใช้งาน
- ส่วนของเซิร์ฟเวอร์สำหรับส่วนแสดงผลของ Web Application
- ส่วนของฐานข้อมูลสำหรับจัดเก็บ แบบจำลองพยากรณ์ ที่ผ่านการ Trained
- ส่วนของฐานข้อมูลสำหรับจัดเก็บข้อมูลที่ใช้ Train แบบจำลองพยากรณ์

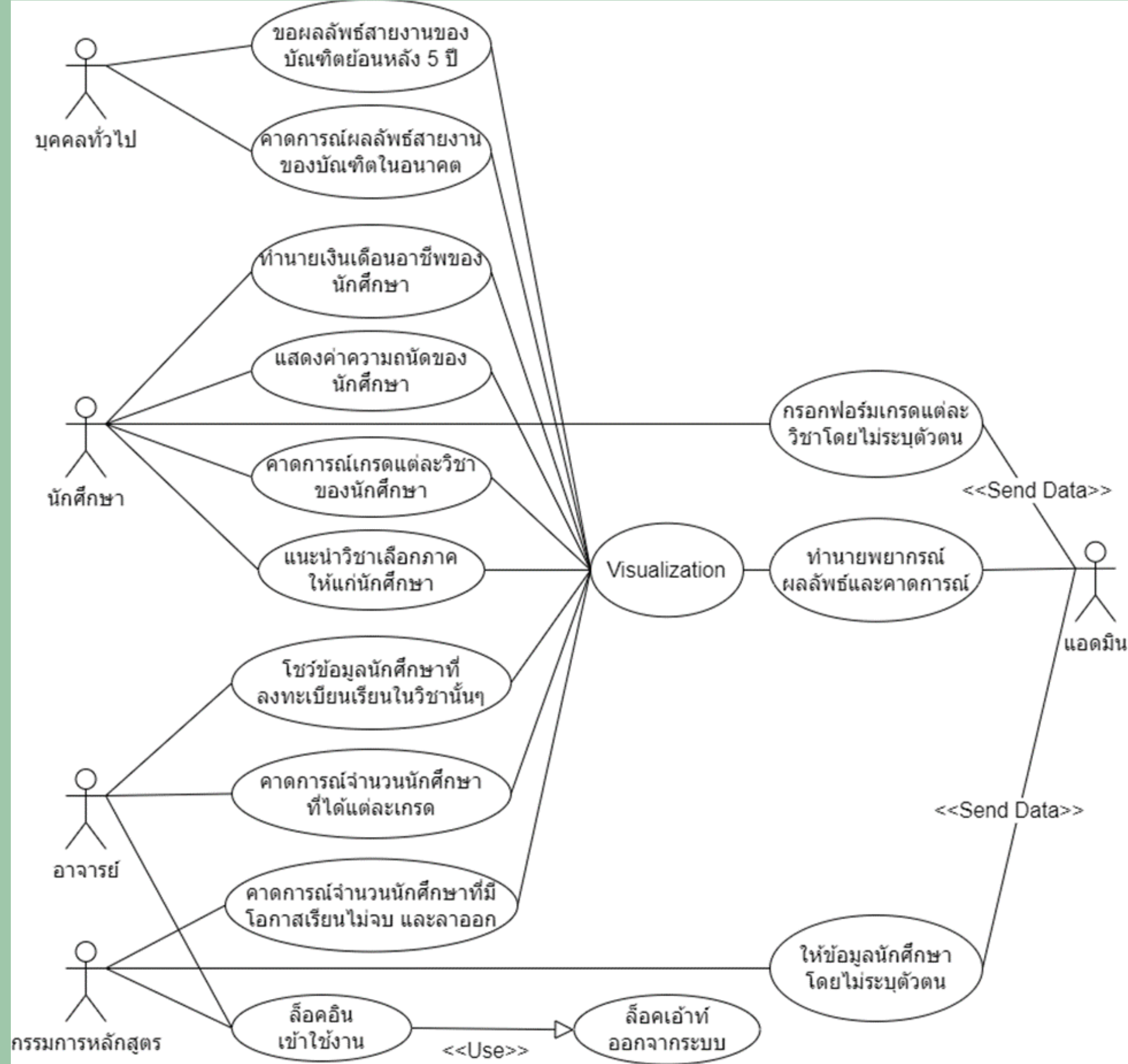


Design

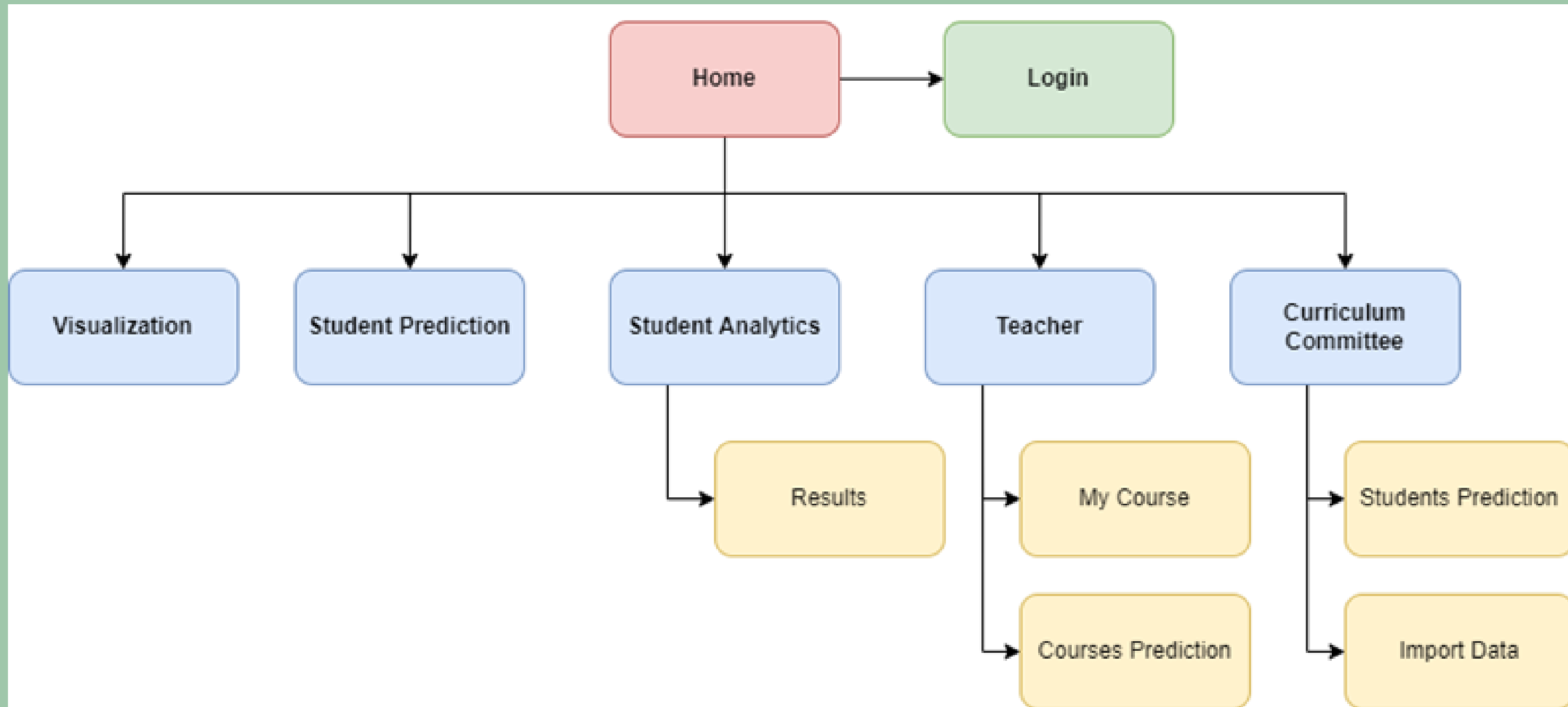
การออกแบบ



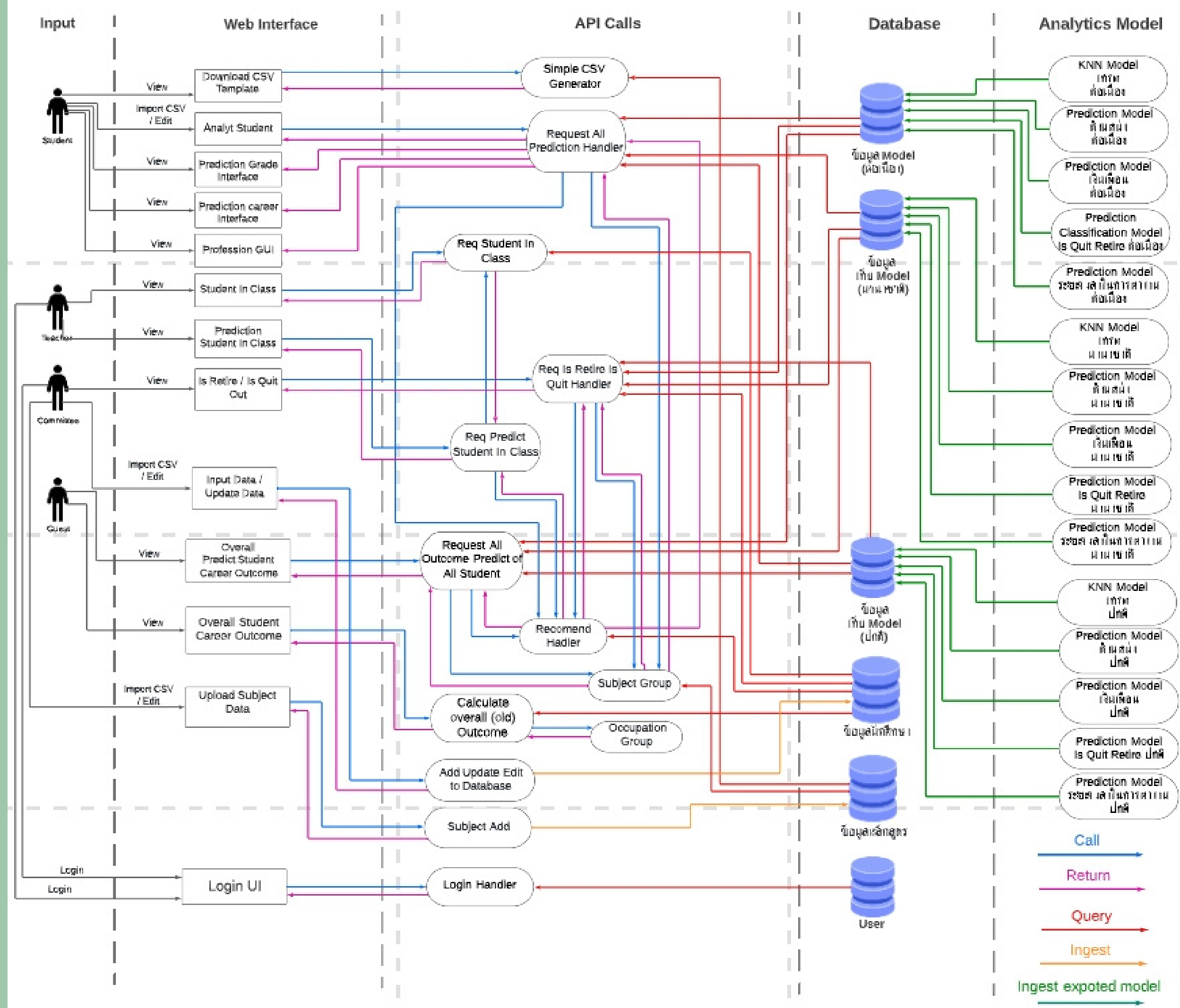
Use Case Diagram



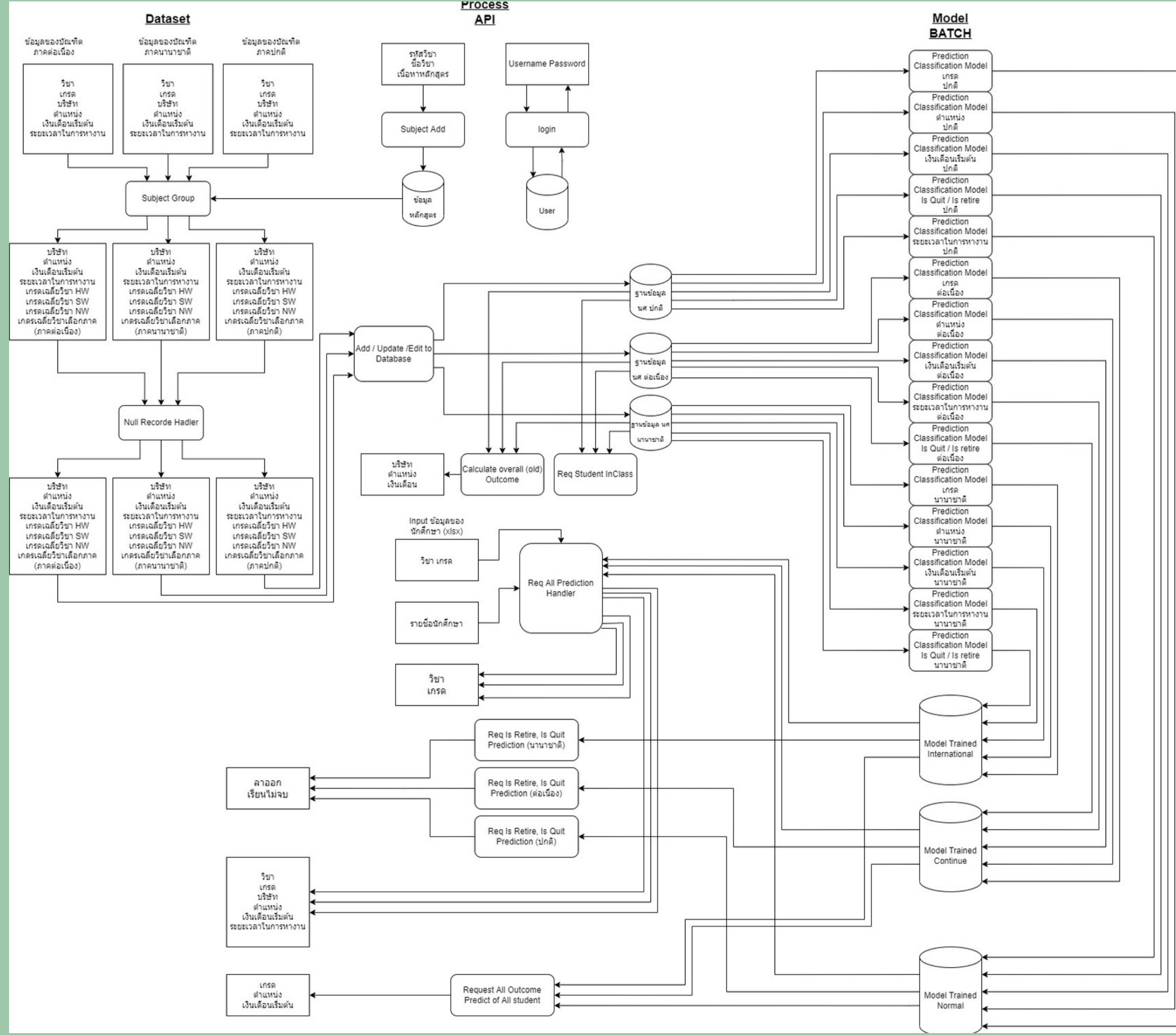
Site Map



System Design



Data Flow Diagram



Test and Performance Evaluation Approaches

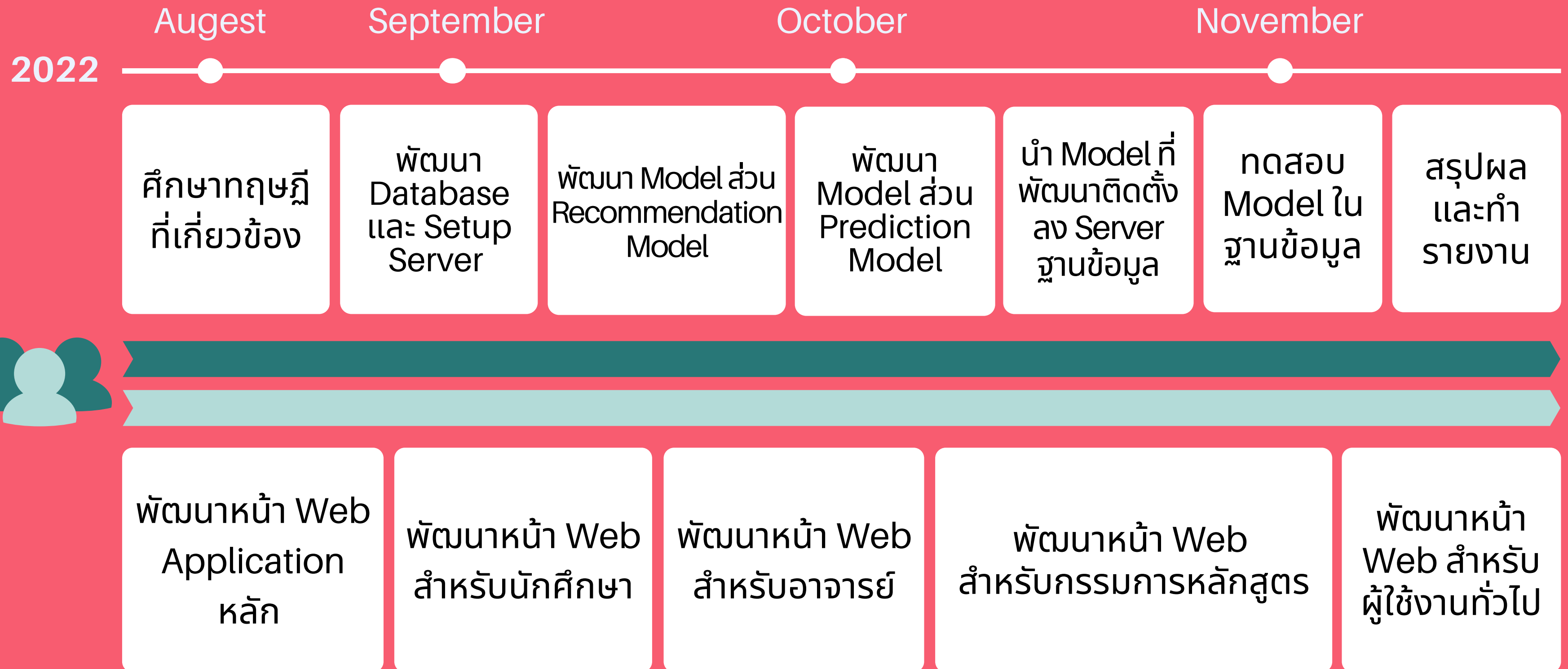
แนวทางการทดสอบและการวัดประสิทธิภาพ

- การทดสอบด้วย Confusion Matrix วัดประสิทธิภาพความถูกต้องของ Model โดยใช้สูตร Accuracy (สูตร $(TP+TN) / (TP+TN+FP+FN)$)
- การทดสอบการวัดด้วย Classification accuracy หาสัดส่วนของจำนวนความถูกต้องในการ Predict ต่อจำนวนที่นำมา Predict ทั้งหมด
- การทดสอบการวัดด้วย Precision หาค่าการทำนายผิด positives (ทำนายผิดแต่ไม่เป็นผลเสีย)
- การทดสอบการวัดด้วย Recall หาค่าการทำนายผิด negative แล้วเสียหาย (ทำนายผิดและมีผลเสีย)
- การทดสอบการวัดด้วย F1-score โดยหาค่าเฉลี่ยระหว่าง Precision กับ Recall



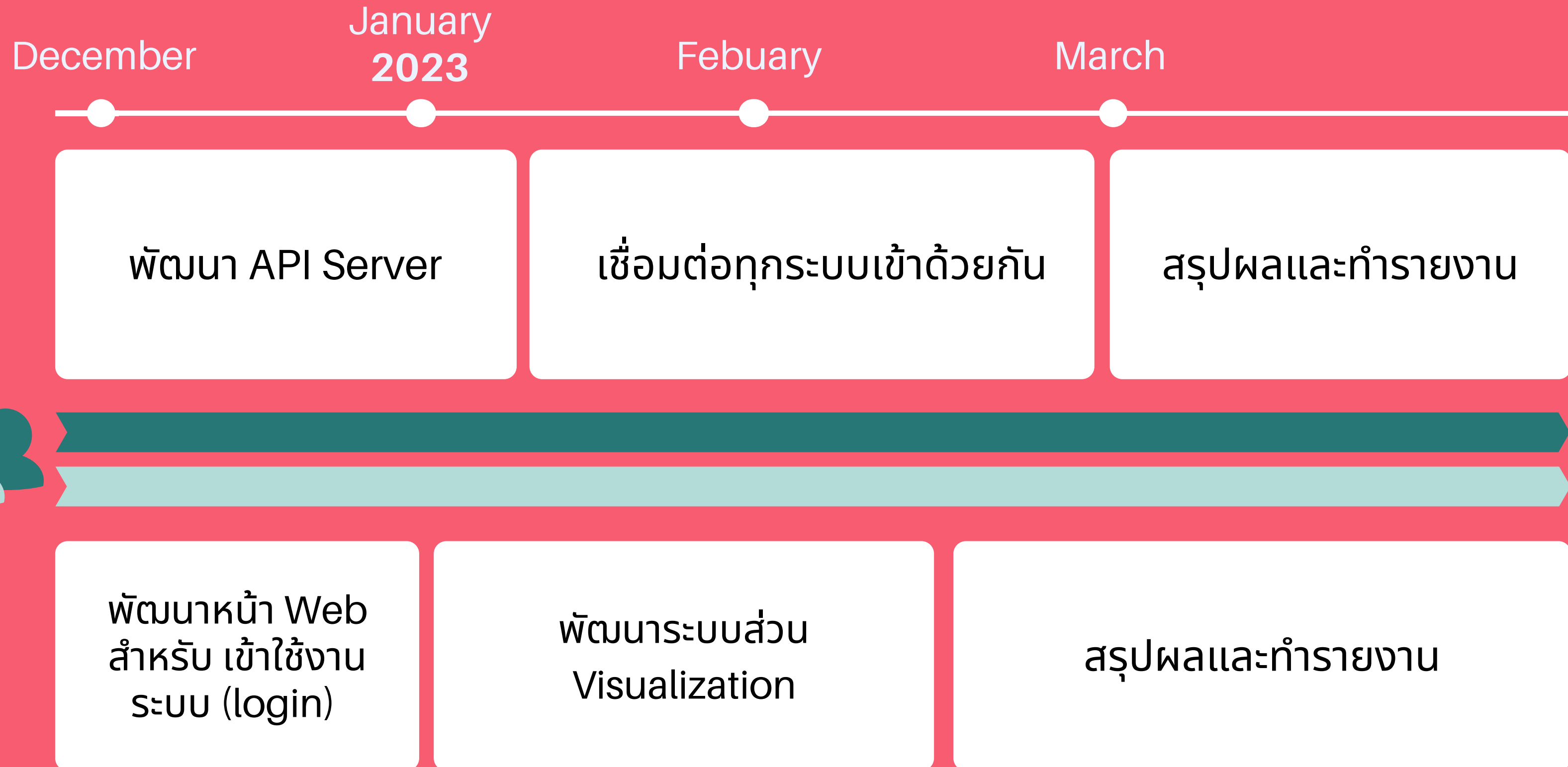
Work Plan

Project 1



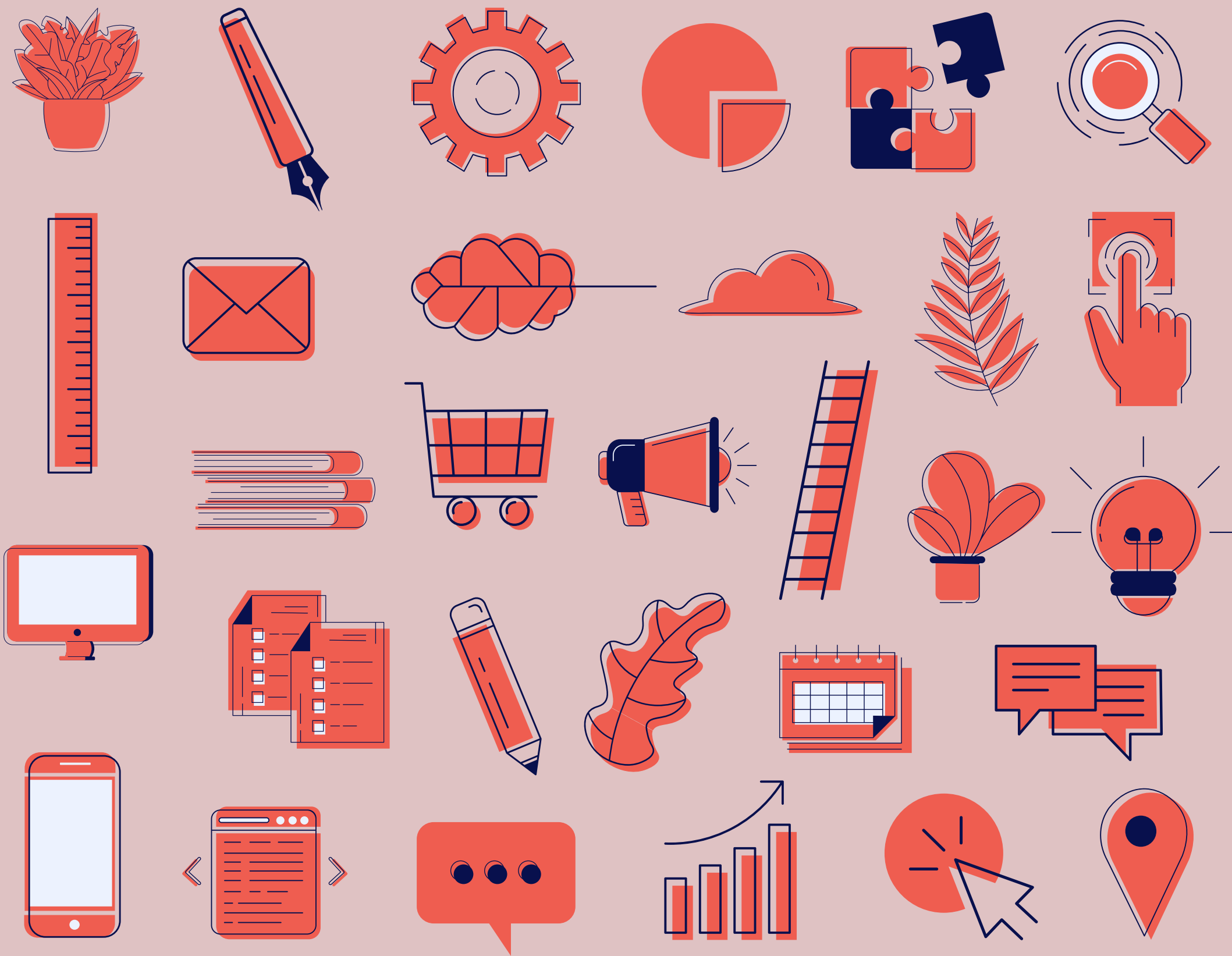
Work Plan

Project 2



Experiment

ผลการศึกษาเทคโนโลยีที่ใช้พัฒนา



Experiment

ผลการศึกษาเทคโนโลยีที่ใช้พัฒนา



What we do and how is it ?

Pandas

Scikit learn

Django Framework

MongoDB Atlas



Expected Benefit



Expected Benefit

ประโยชน์ที่คาดว่าจะได้รับ

1. ได้ระบบรวบรวมข้อมูลผลการเรียนของนักศึกษาและข้อมูลแบบสำรวจการทำงาน
ของบัณฑิต แล้วนำมาวิเคราะห์และนำเสนอข้อมูลที่เป็นประโยชน์ในการบริหาร
หลักสูตรของกรรมการหลักสูตร
2. มีระบบที่สามารถแนะนำ วางแผน และพยากรณ์อาชีพในอนาคตจากผลการเรียน
ของนักศึกษา





**THANK
YOU**

