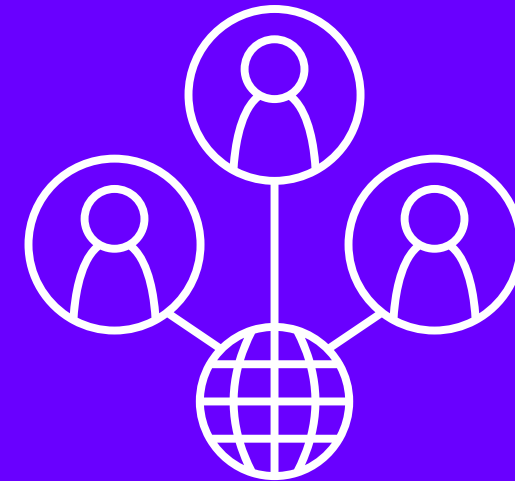
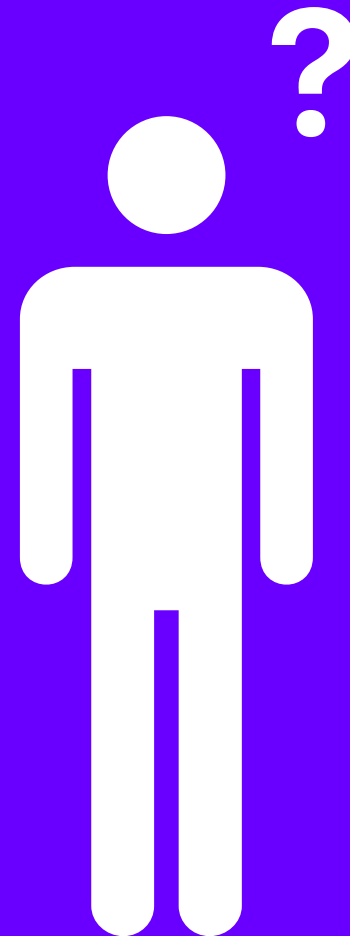
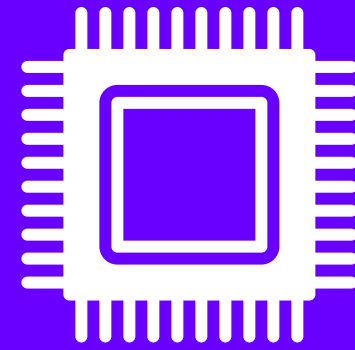
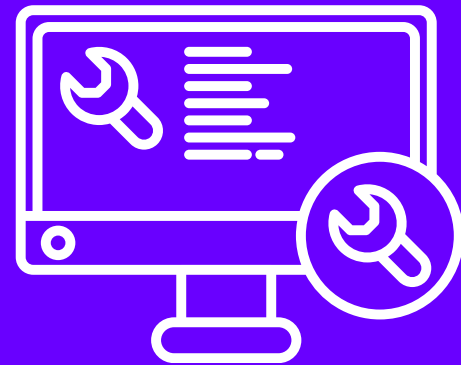


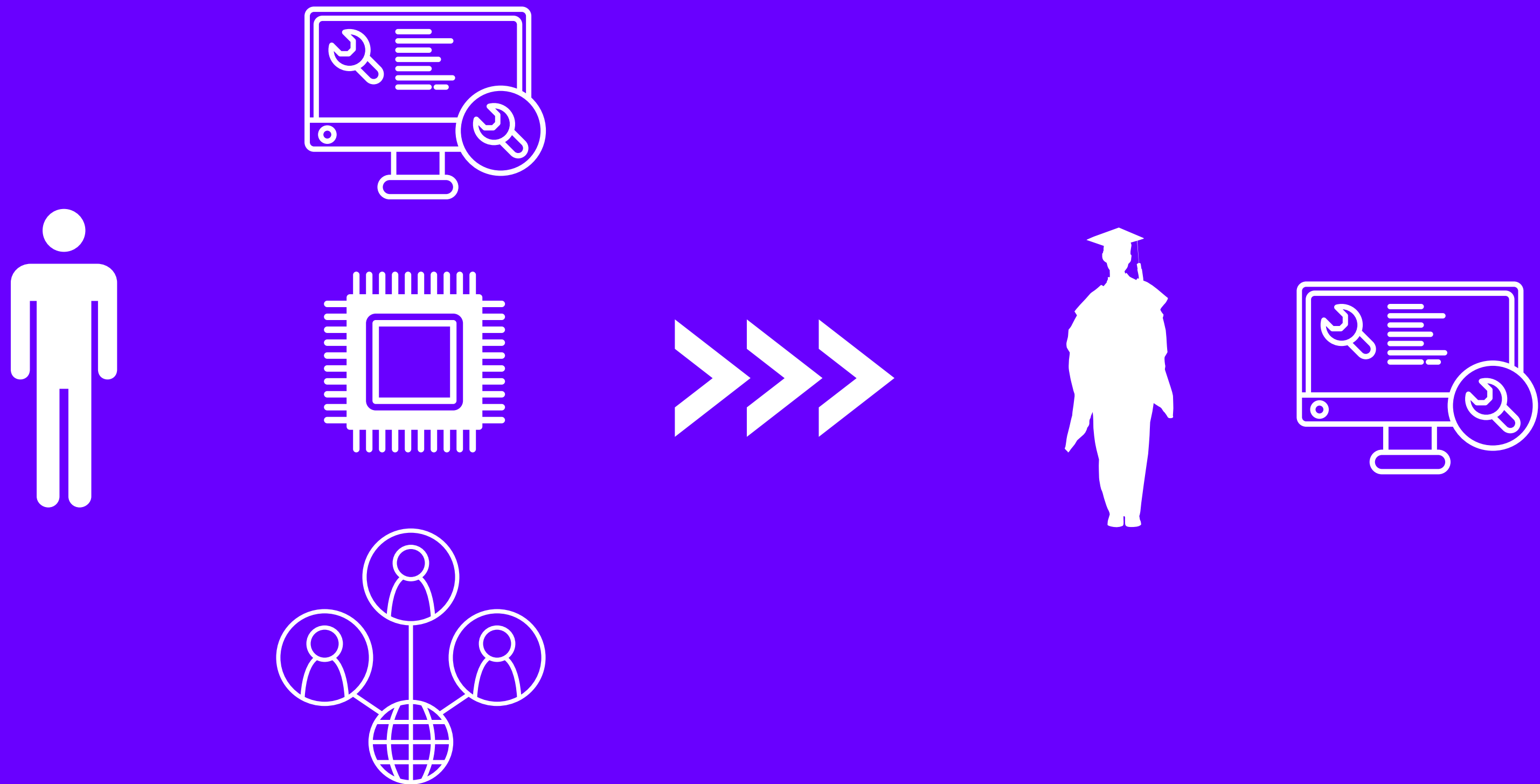
Curriculum Output Prediction From Student Academic Data

อาจารย์ที่ปรึกษา
ผศ. ดร. รณัญชัย ตรีภาค

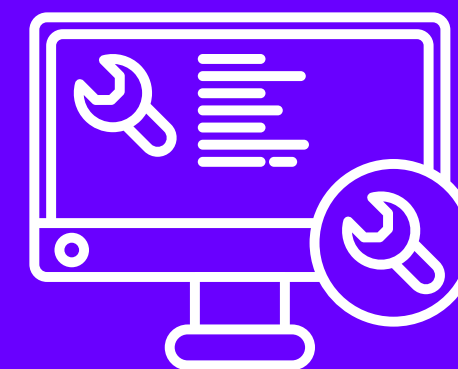
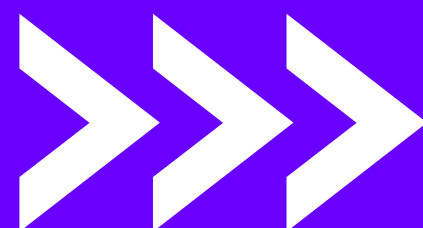
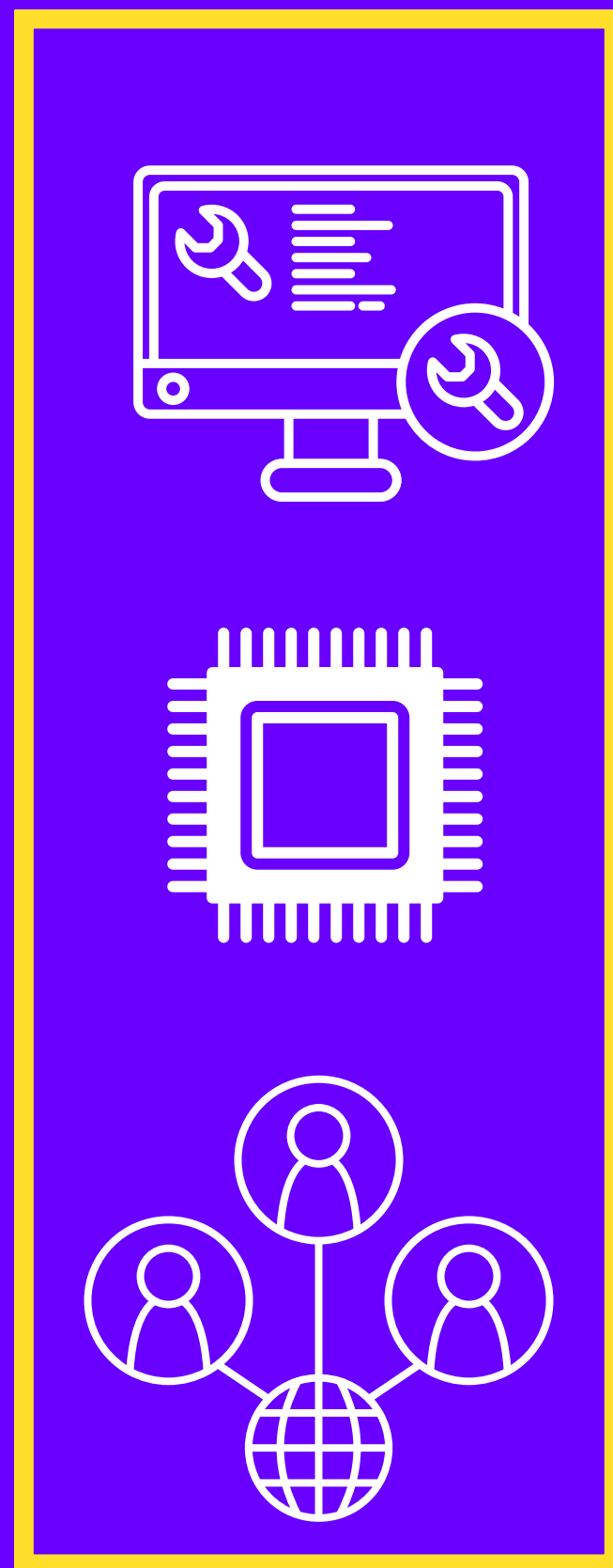
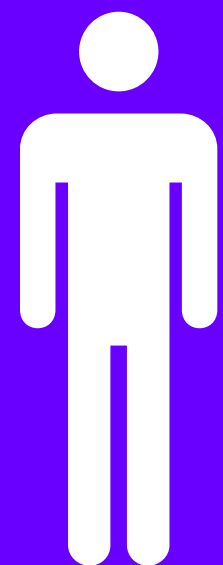
What is the Problem?



What is the Problem?



What is the Problem?

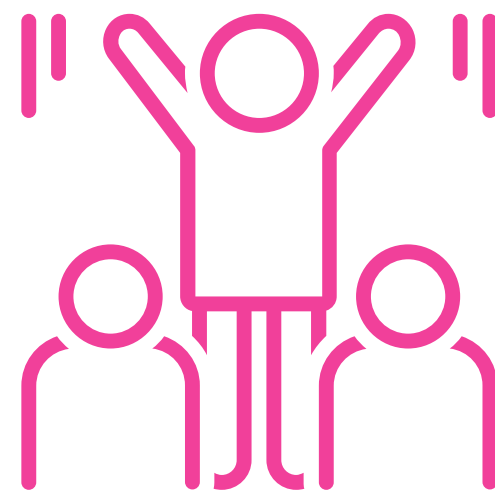
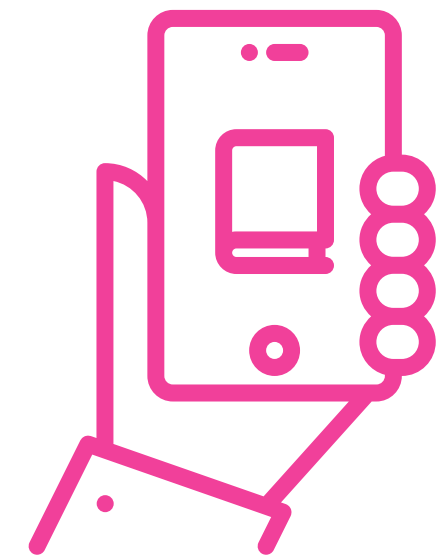


What is the Problem?

**หน่วยงานภายนอกต้องการทราบผลลัพธ์
สายงานอาชีพของนักศึกษาจากหลักสูตร
ปัจจุบัน**

**๘ นักศึกษาในปัจจุบันไม่รู้ว่าเกรดที่ตัวเองมีนั้น มีความสามารถไป
สายงานไหนได้บ้าง**

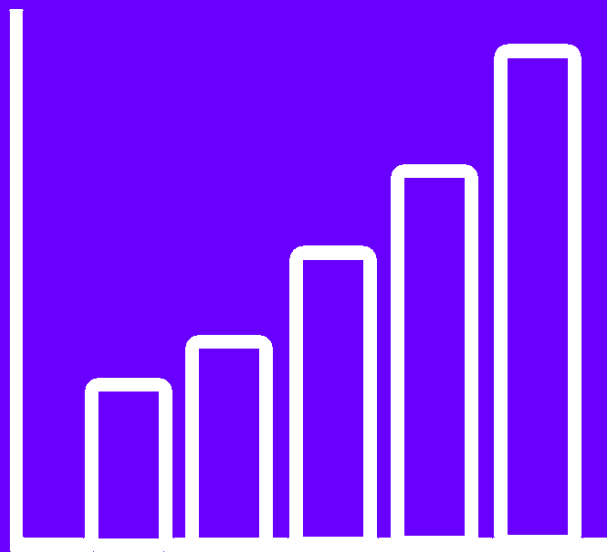
OBJECTIVE



Our Solution

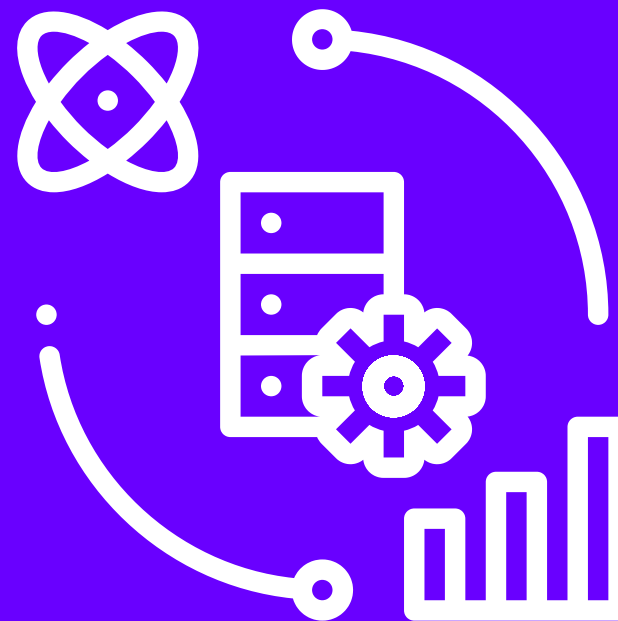
01

การวิเคราะห์ข้อมูลที่มีอยู่ตั้งแต่อดีตจนถึงปัจจุบัน



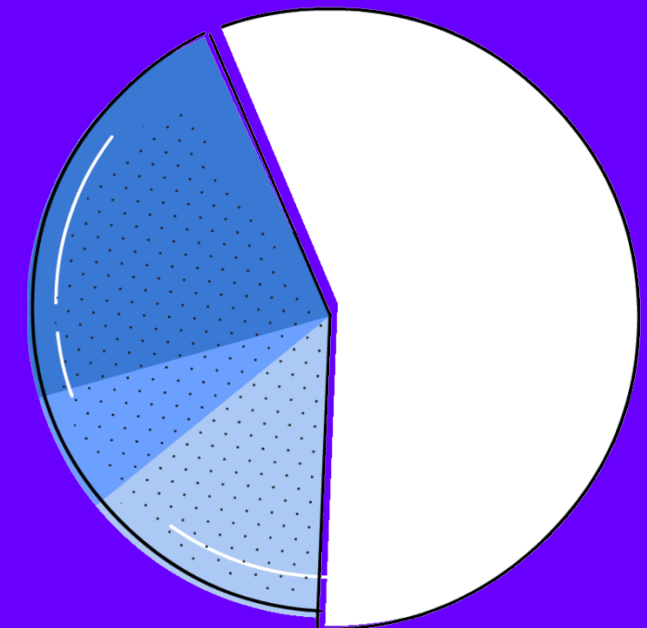
02

การนำข้อมูลผลการเรียนของนักศึกษาในอดีตมาใช้



03

Data Analytics, Prediction, Visualization



SCOPE



01

SCI-KIT Learn base
Model

02

Data from registry office



03

Data on past academic
only past 2 years



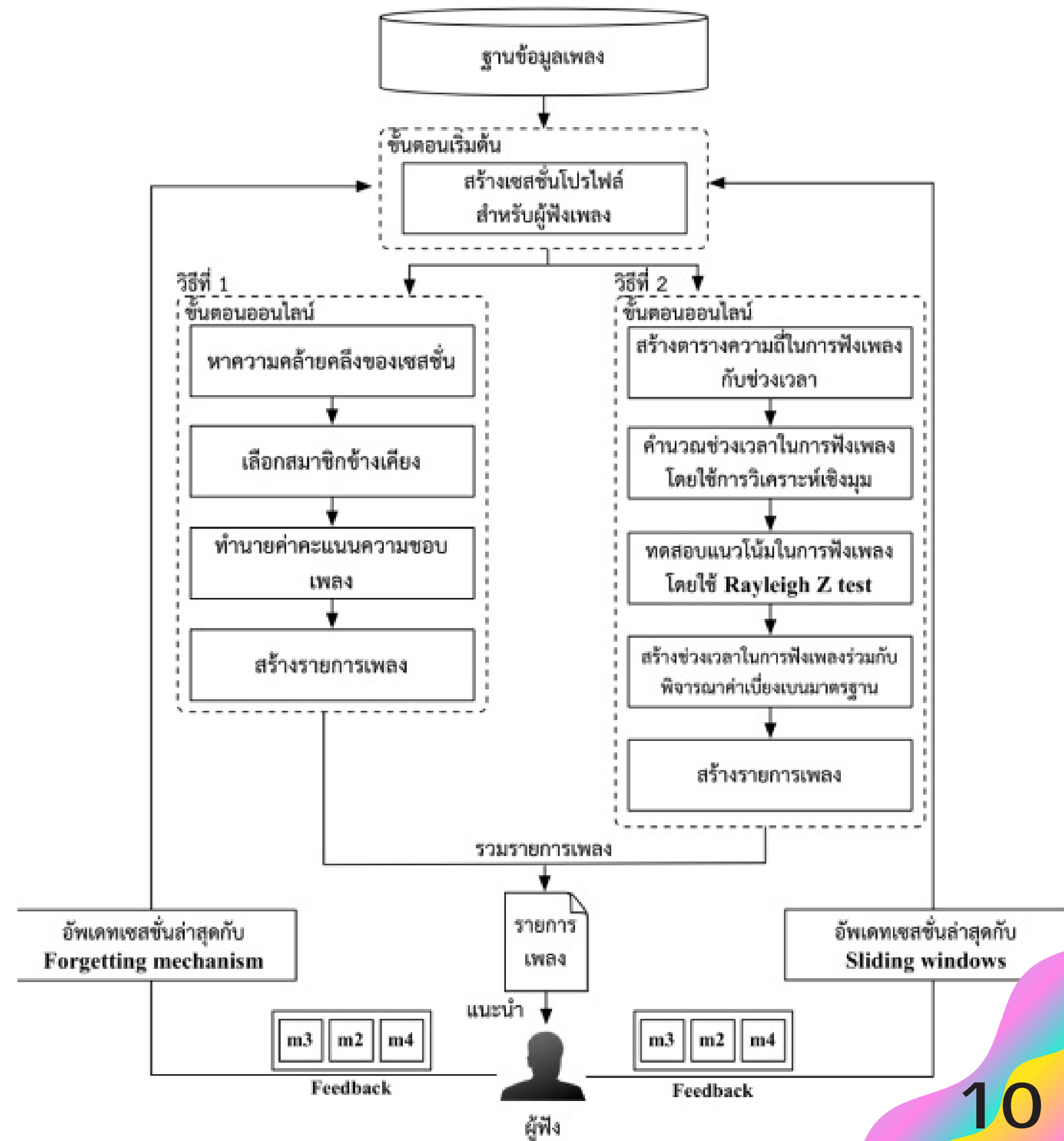
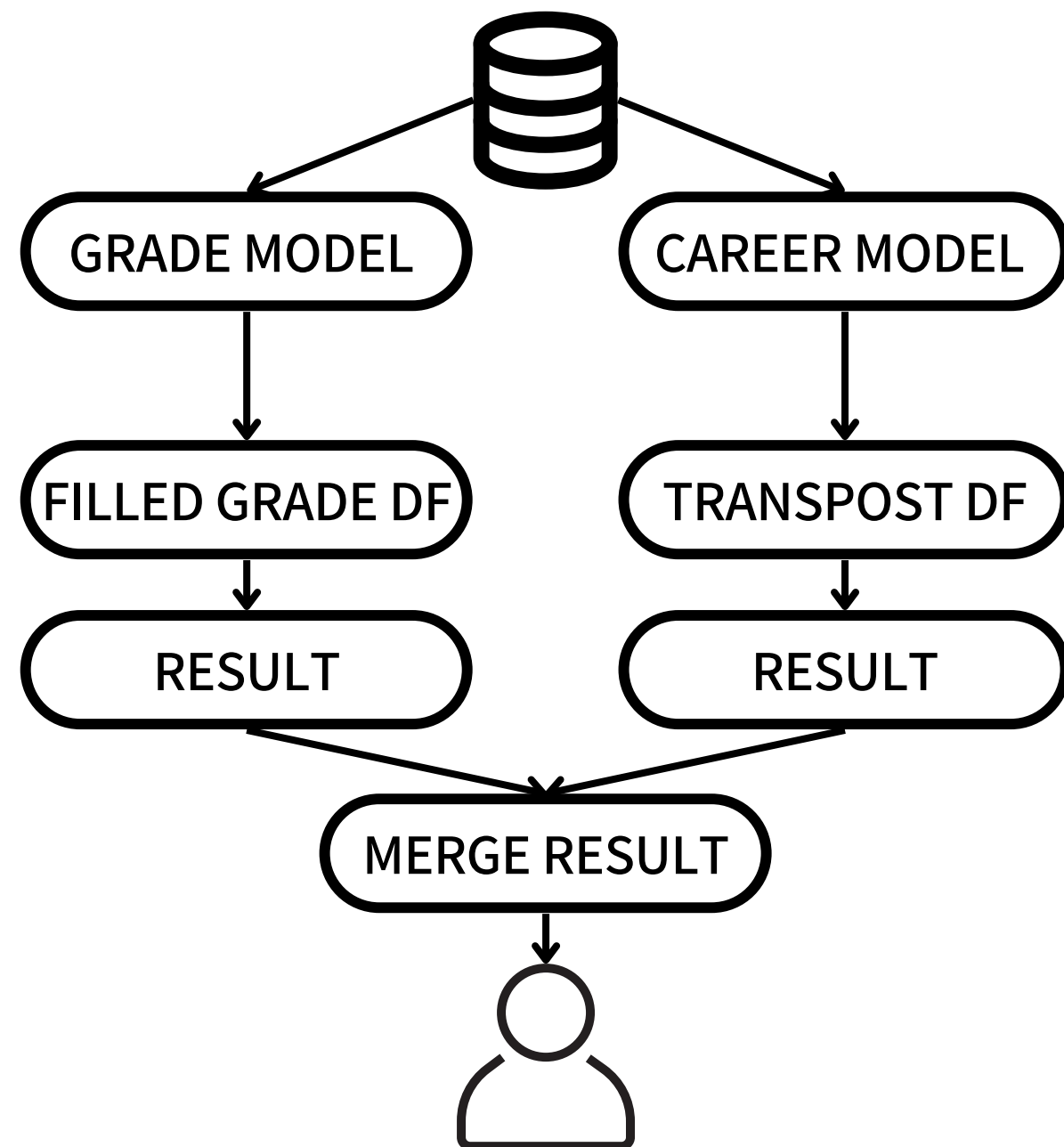
Related Research

Research that made our project possible

- การสร้างรายการเพลงโดยใช้การกรอกร่วมแบบเซสชันที่เพิ่มขึ้นด้วยกลไกการลืมหและการวิเคราะห์สถิติเชิงมุม
- ระบบแนะนำสินค้าอาหารโดยใช้ระบบแนะนำแบบผสมผสาน

การสร้างรายการเพลงโดยใช้การกรองร่วมแบบเซสชันที่เพิ่มขึ้นด้วยกลไกการลืมและการวิเคราะห์สถิติเชิงมุม

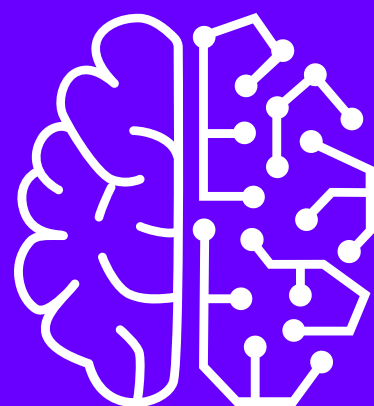
เราได้นำวิธีการการนำเข้าและส่งออกข้อมูลของงานวิจัยนี้มาใช้งาน



ระบบแนะนำสินค้าอาหารโดยใช้ระบบแนะนำแบบผสมผสาน

01

นำผลลัพธ์การทดลองของงานวิจัยมาเป็น
รากฐานในการตัดสินใจ ในการใช้ Model

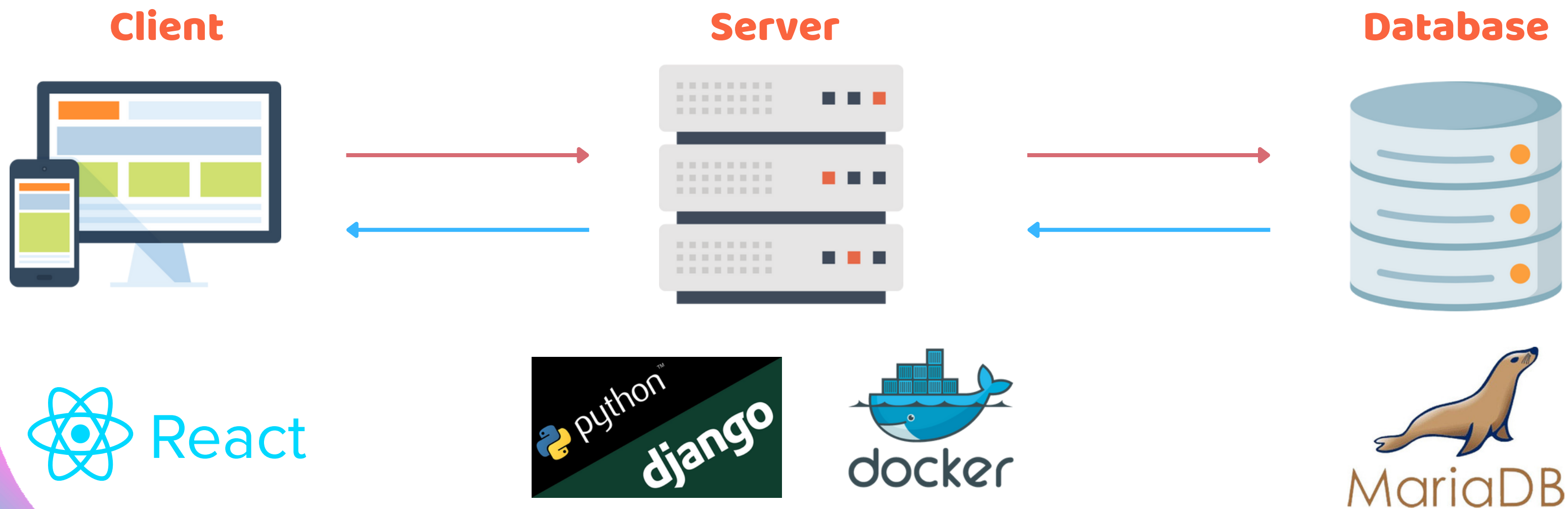


02

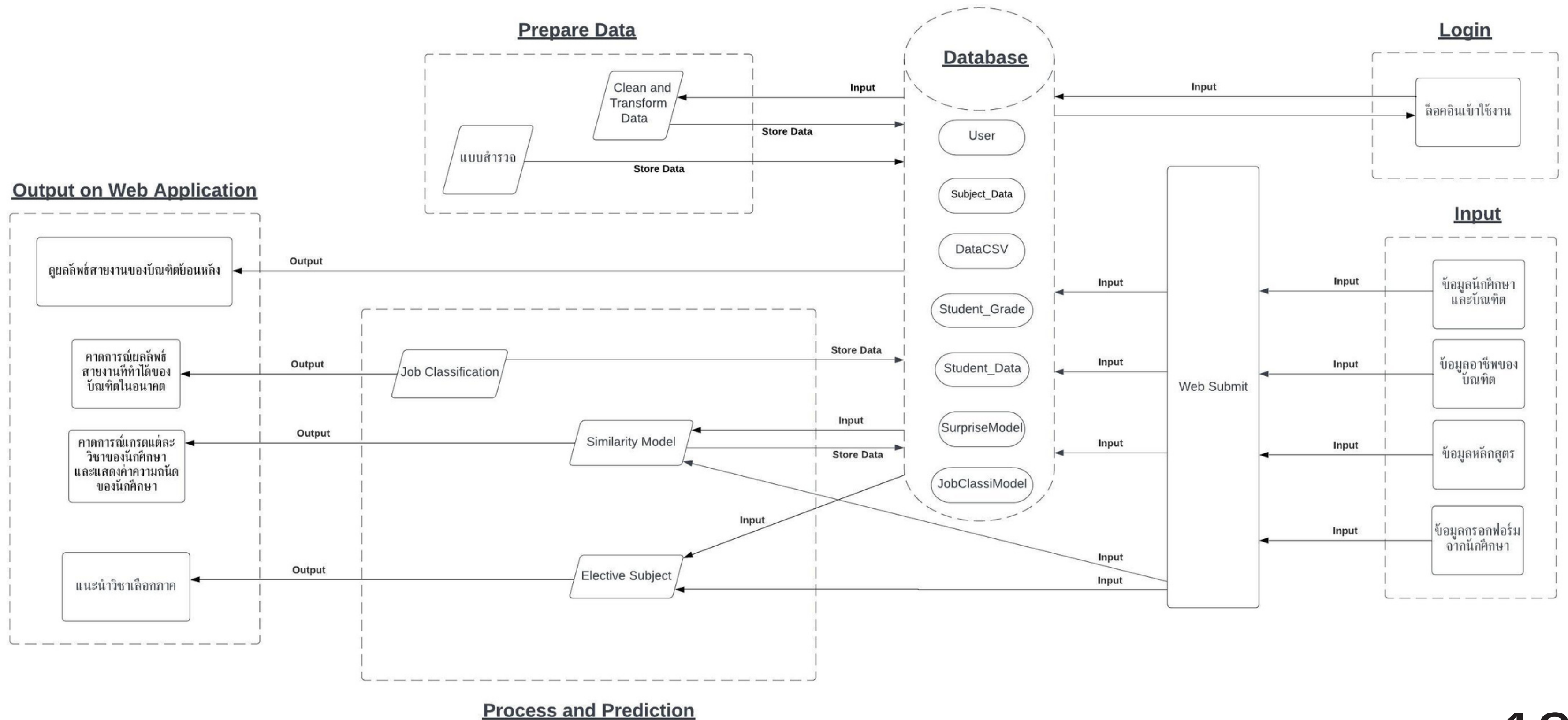
นำวิธีการวัดประสิทธิภาพของแบบ
จำลองจากงานวิจัยมาปรับใช้



โครงสร้างระบบ



แผนการทำงาน



การเตรียมข้อมูล

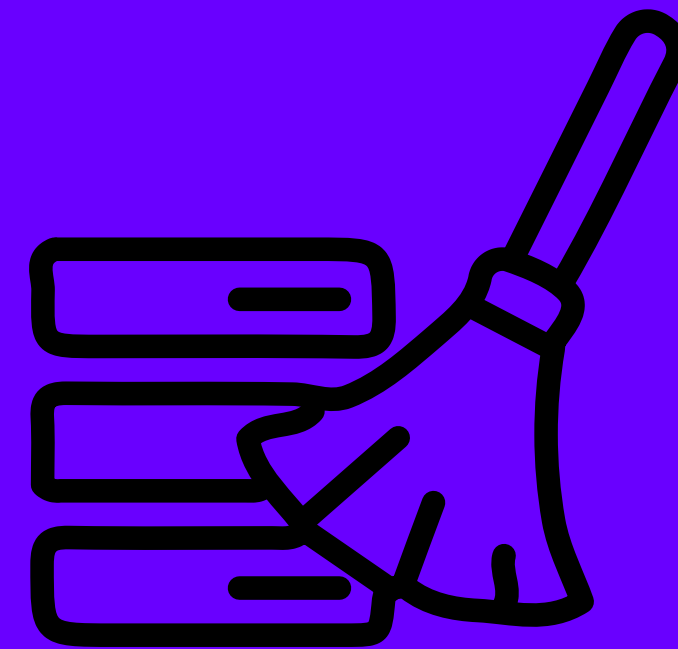
01

สำรวจแบบสอบถามอาชีพบัณฑิต



02

Clean , Transform and Ingestion process



สำรวจแบบสอบถามอาชีพบัณฑิต

01

สร้างแบบสำรวจการมีงานทำ



02

กระจายแบบสำรวจผ่านช่องทาง online



03

ลงพื้นที่เก็บแบบสำรวจด้วยตัวเอง



04

นำชุดข้อมูลที่ได้มา Encrypt



Clean , Transform and Ingestion process

ข้อมูลจากสำนักทะเบียน

- Encrypt รหัสนักศึกษา
- แยกชุดข้อมูลออกเป็น 2 ส่วน
 - 1) Student Data
 - 2) Student Grade
- นำเอาข้อมูลสายงานอาชีพมา Join
- นำข้อมูลทั้ง 2 ส่วนเก็บลงในฐานข้อมูล

ข้อมูลจากแบบสำรวจ

- Encrypt รหัสนักศึกษา
- clean data ที่ไม่จำเป็นออก
- ทำการปรับหมวดหมู่ของอาชีพ
- จัดการกับอาชีพที่มีสมาชิกน้อย
- นำข้อมูลไป join กับ student data

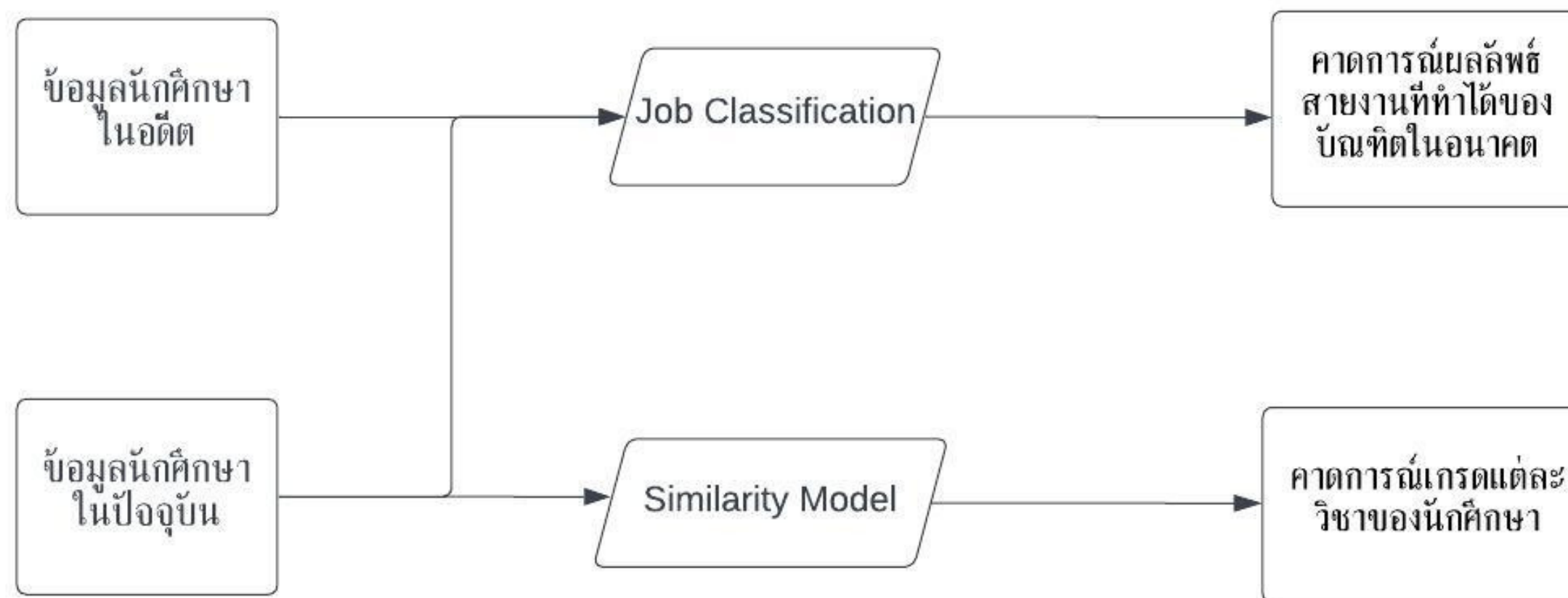
พัฒนา Model

01

Similarity Model

02

Job Classification Model



Similarity Model

01

แปลง input grade ของ user เป็น dataframe

02

สร้าง dataset จากการ query ข้อมูลจากฐานข้อมูล

03

ทำการสร้างแบบจำลองจากชุดข้อมูล

04

ทำการ Hyperparameter Tuning

05

ทำการวัดประสิทธิภาพด้วยค่า RMSE

Job Classification Model

01

กำหนดปีของหลักสูตรที่จะนำแบบจำลองไปทำนาย

02

สร้าง dataset จากการ query ข้อมูลจากฐานข้อมูล

03

ทำการสร้างแบบจำลองจากชุดข้อมูล

04

ทำการ Hyperparameter Tuning

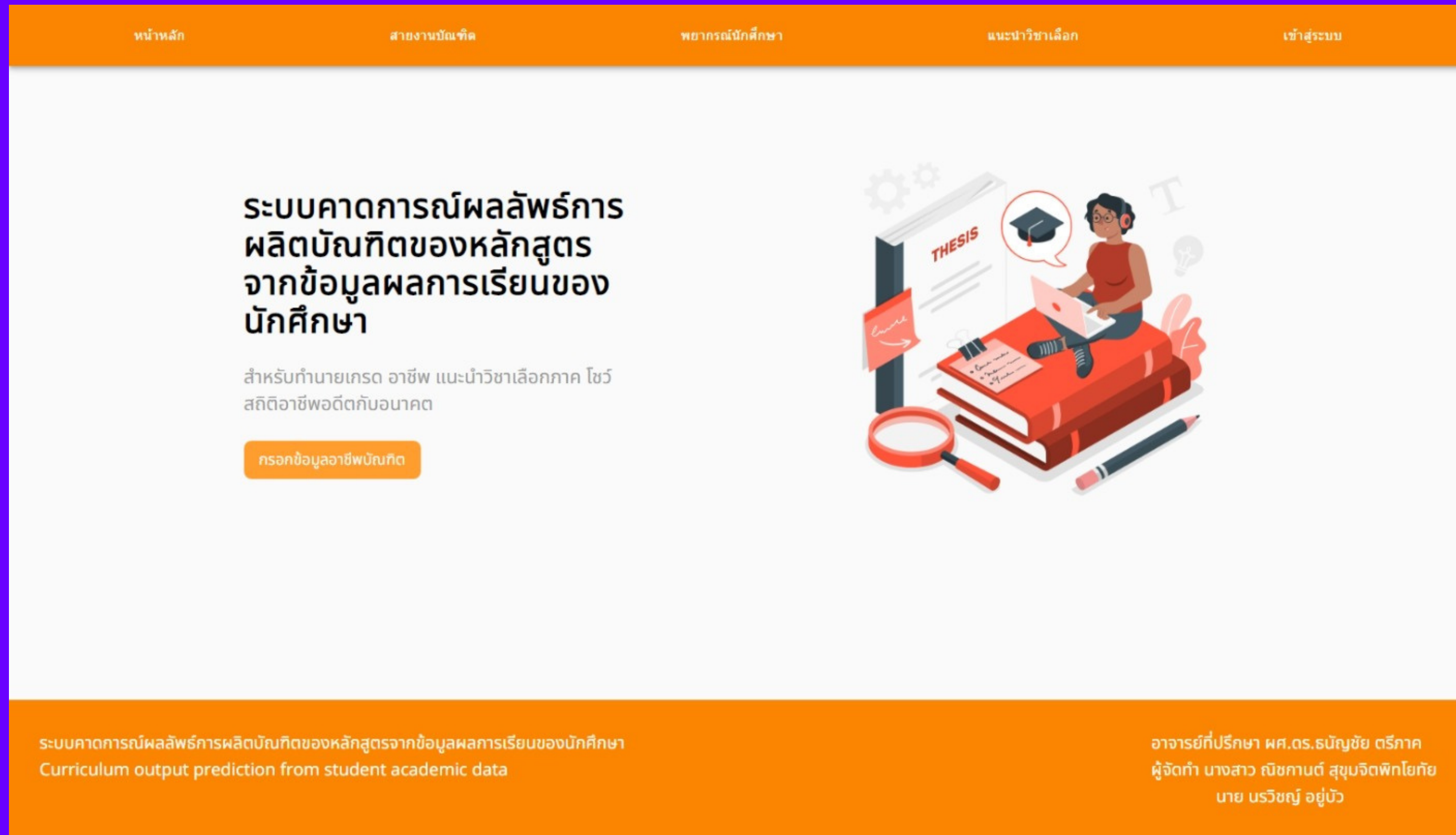
05

ทำการวัดประสิทธิภาพด้วยค่า Accuracy

06

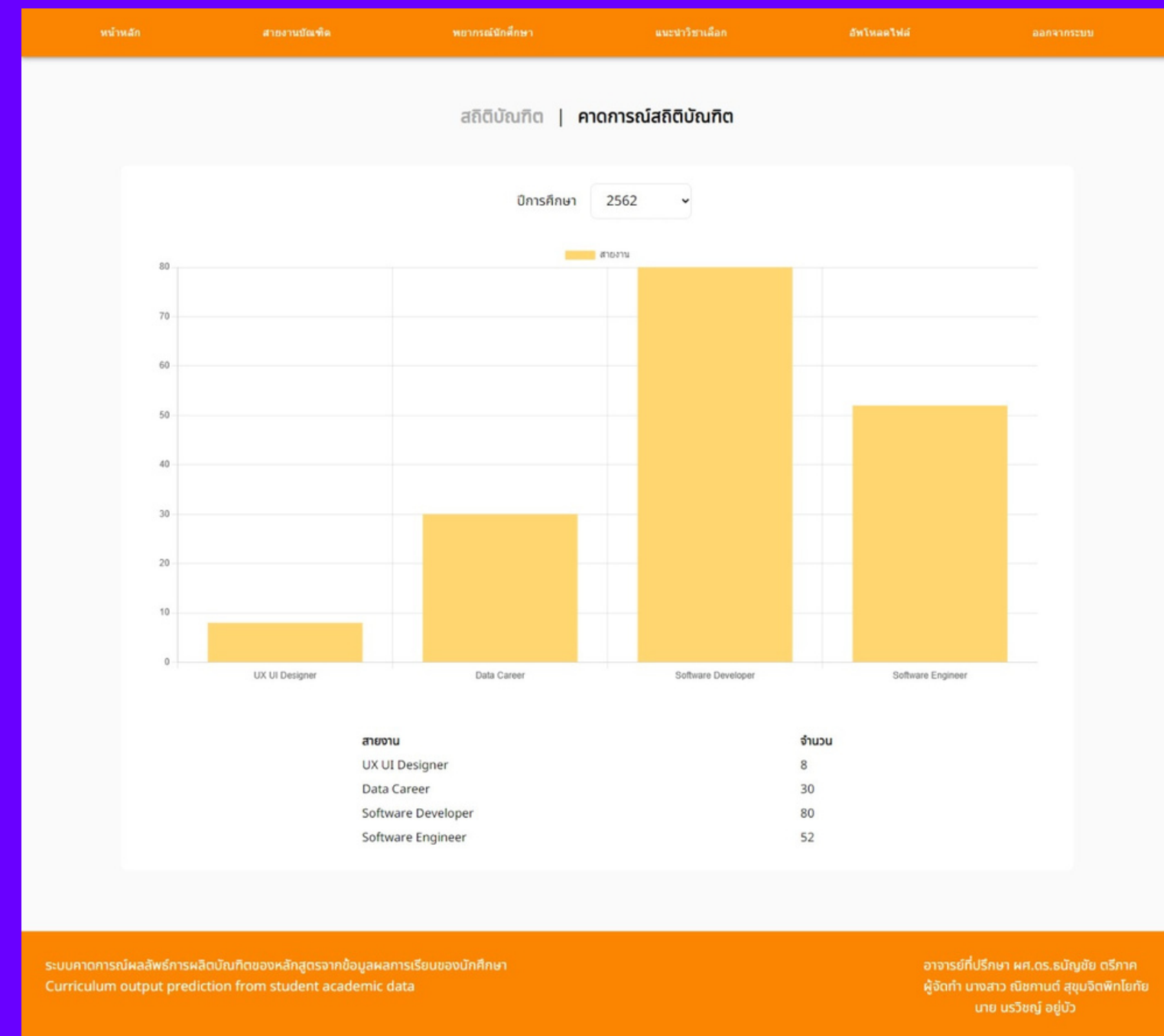
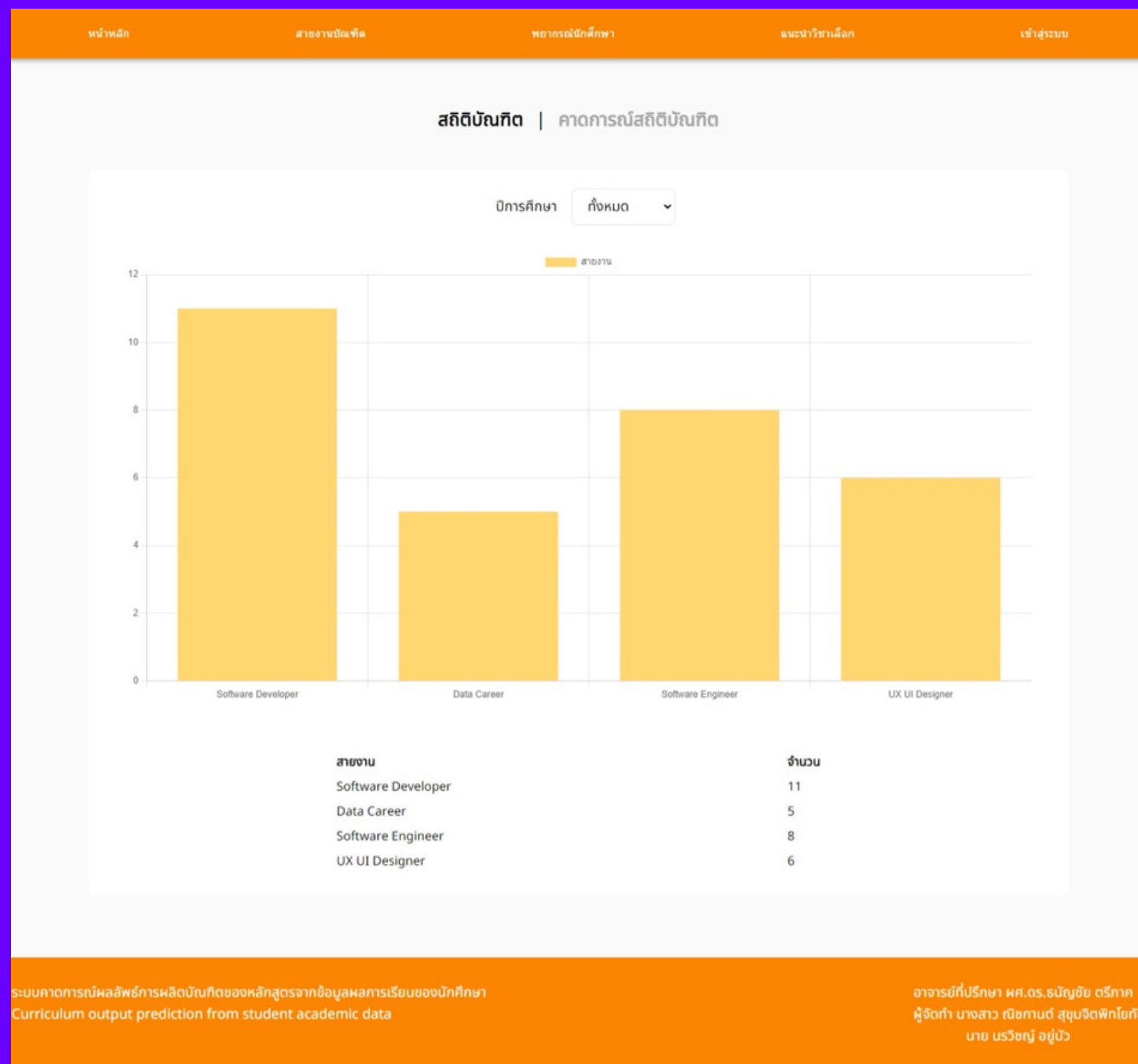
แปลงแบบจำลองเป็น Binary แล้วเก็บลงฐานข้อมูล

WEB APPLICATION



ผลลัพธ์

ค่า Accuracy ของอัลกอริทึม Random Forest Classification ร่วมกับการปรับลดจำนวนของกลุ่มผลลัพธ์นั้นออกมาอยู่ที่ 0.75 หรือคิดเป็น 75%



ผลลัพธ์

หน้าหลัก

สายงานบัณฑิต

พยากรณ์นักศึกษา

แนะนำวิชาเลือก

เข้าสู่ระบบ

ผลการวิเคราะห์

คณะวิศวกรรมศาสตร์คอมพิวเตอร์ หลักสูตรปี 2560

ประเมินเกรดรายวิชา

วิชา	เกรด
01076005 OVERSEA TRAINING	D+
01076007 DATA COMMUNICATIONS	D+
01076014 COMPUTER ENGINEERING PROJECT PREPARATION	B
01076021 COMPUTER ARCHITECTURE	C
01076024 SOFTWARE ARCHITECTURE AND DESIGN	C
01076029 CO-OPERATIVE EDUCATION	C+
01076120 WEB APPLICATION DEVELOPMENT PROJECT	C
01076041 INTERNETWORKING STANDARDS AND TECHNOLOGIES IN PRACTICE	C
01076044 INTRODUCTION TO CLOUD ARCHITECTURE IN PRACTICE	C
01076414 INTRODUCTION TO AUTONOMOUS VEHICLE	C
01076559 SOFTWARE ARCHITECTURE AND DESIGN	C
01076635 IP SWITCHED NETWORKS	C

อาชีพที่เป็นไปได้

UX UI Designer

หลักการทำงาน

1. การทำนายเกรดของแต่ละวิชา

โดยใช้หลักการเปรียบเทียบความคล้ายคลึงของพฤติกรรมของนักศึกษา มาเป็นชุดข้อมูลพื้นฐานในการวิเคราะห์เชิงสถิติผ่าน library surprise scikit-learn ของ Python แล้ววัดประสิทธิภาพด้วยการทำ K-Fold Cross validation แล้วอ้างอิงประสิทธิภาพผ่านค่า RMSE

2. การทำนายอาชีพ

โดยใช้หลักการการจับกลุ่มเชิงสถิติผ่านกระบวนการ Random Forest Classification ด้วย library scikit-learn ของ Python แล้ววัด ประสิทธิภาพด้วยการใช้ผลลัพธ์การทำนายต่อผลลัพธ์จริงออกมาเป็นค่า Accuracy

ทั้งนี้การทำนายดังกล่าวทั้งหมดมาจากการวิเคราะห์ข้อมูลด้วย กระบวนการเชิงสถิติทั้งหมด ซึ่งในผลลัพธ์ในความเป็นจริงอาจมีตัวแปรอื่น ที่ทางผู้จัดทำไม่ได้นำมาใช้วิเคราะห์ซึ่งอาจทำให้ผลลัพธ์มีความคลาดเคลื่อน เป็นบางส่วน

กลับ

ระบบคาดการณ์ผลลัพธ์การผลิตบัณฑิตของหลักสูตรจากข้อมูลผลการเรียนของนักศึกษา

Curriculum output prediction from student academic data

อาจารย์ที่ปรึกษา ผศ.ดร.ธนัญชัย ตรีภาค

ผู้จัดทำ นางสาว ณิษกานต์ สุขุมจิตพิทยโยทัย

นาย นรวิชญ์ อยู่บัว

ชุดข้อมูลในส่วนของหลักสูตรวิศวกรรมคอมพิวเตอร์มีค่า RMSE อยู่ที่ 0.5751 และในส่วนของหลักสูตรวิศวกรรมคอมพิวเตอร์(ต่อเนื่อง)มีค่า RMSE อยู่ที่ 0.7351 ซึ่งเมื่อวัดจากช่วงของ Rating หรือ Grade ของชุดข้อมูลที่อยู่ตั้งแต่ 0 ถึง 4

แนวทางพัฒนาต่อ

01

พัฒนาค่าผลลัพธ์ Accuracy ของ Model Job Classification ให้ถึง 0.9

02

สำรวจข้อมูลและใช้ข้อมูลของนักศึกษามากขึ้น



THANK YOU

FEEL FREE TO ASK A QUESTION

นางสาว ณัชกานต์ สุขุมจิตพิทย 62010299

นาย นรวิชญ์ อยู่บัว 62010465