# รายงานความก้าวหน้าวิชา CE Project

**ครั้งที่ 5**

**ระหว่างวันที่ 24 ก.ย. 65 ถึงวันที่ 21 ต.ค. 65**

1. ชื่อโครงงาน (อังกฤษ) Analytics and Prediction System for CE Curriculum administrators

## การดำเนินงานมีความก้าวหน้า 43 % (ใช้ค่า **% Complete** จาก MS Project)

มีความก้าวหน้าเพิ่มขึ้นจากรายงานความก้าวหน้า ครั้งก่อน 8 %

□ เร็วกว่าแผน วัน □ ช้ากว่าแผน 34 วัน

## **รายละเอียดความก้าวหน้า**

**นัดประชุมกับที่ปรึกษาจำนวน 2 ครั้ง**

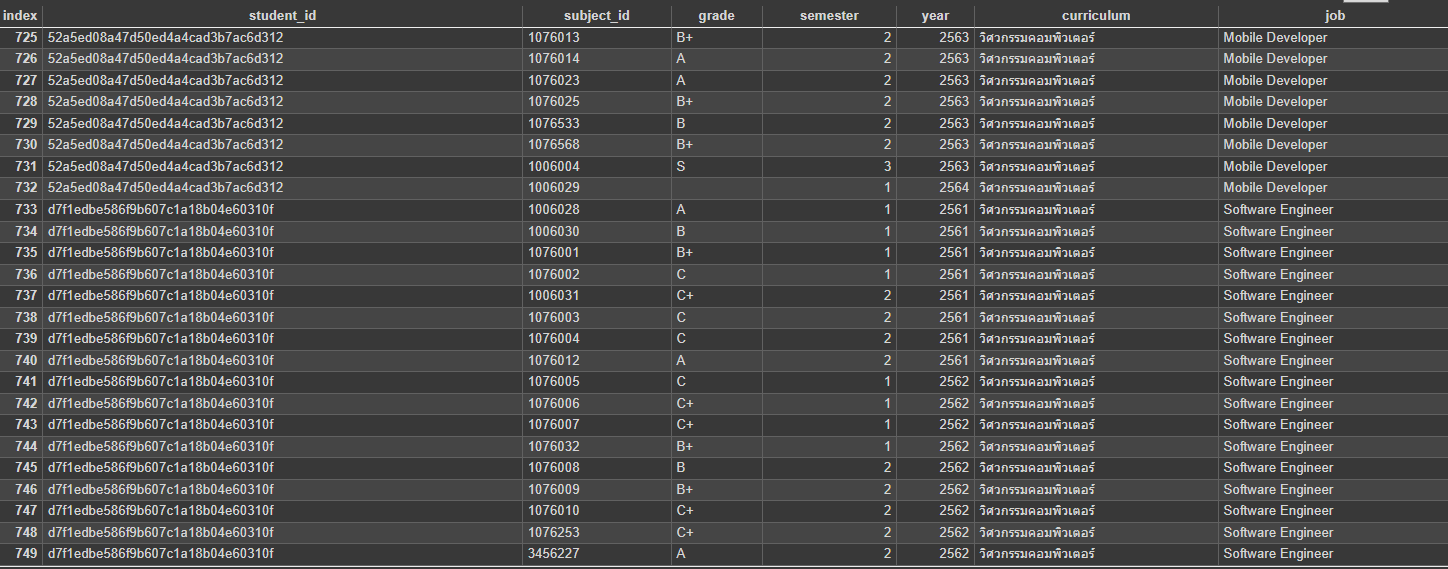
ครั้งที่ 1 : รายละเอียดประชุม ที่ปรึกษาได้ตาม progress ของโครงงานในช่วงที่ผ่านมา ผู้จัดทำได้แจ้งรายละเอียดโดย 1 แจ้ง progress ปัจจุบัน โดยผู้จัดทำเน้นการทำโครงงานในช่วง progress นี้เป็นการนำ function ที่ทดลองใน google collab มา implement ลง Django Web เพื่อทำ demo และเพื่อให้ผู้จัดทำใช้งานทดลอง function ได้สะดวกยิ่งขึ้น 2 ได้ทำการแสดงผลลัพท์การทดลองตัว Model Job Classification ที่ได้ทำการทดลองกับชุดข้อมูลที่ได้ทำการ survey มา 3 ที่ปรึกษาเสนอแนะแนวทางแก้ไข้และปรับเปลี่ยน 1 การเก็บ survey 2 การ Generate File CSV 3 การ Update Career ลง Database

**หัวข้อการพัฒนาโครงงานตาม Gantt Chart**

**ศึกษาทฤษฎีที่เกี่ยวข้อง Complete 75 % (remaining 15 Hr)** หมายเหตุ พักไว้สำหรับเพื่อในอณนาคตมีเรื่องที่จะต้องศึกษาเพิ่มเติม

**เตรียม Server Complete 50 % (remaining 7 Hr)** หมายเหตุ จากที่ได้ Discuss กับที่ปรึกษาทำให้ต้องพักส่วนนี้ไปก่อนแล้วไปมุ่งเน้นกับการพัฒนาตัว Model ทั้งหมด และ แสดงผลลัพท์

**เตรียม Data สำหรับการพัฒนา Complete 89 % (remaining 0 Hr)** หมายเหตุเนื่องด้วย Data test สำหรับ Model ทั้งหมดคือ Data ของนักศึกษาปี 1-4 ปัจจุบัน ซึ่งที่ปรึกษาจะเตรียม Data ชุดนี้ให้ในเทอมการศึกษาหน้า

**พัฒนา Job Classification Model Complete 90 %** หลังจากที่ได้ข้อมูล survey ที่ผ่านการ Hash รหัสนศจากที่ปรึกษามาแล้วได้นำข้อมูลมา join กับข้อมูล grade หลังจากนั้นได้ทำการ transpose dataframe ให้อยู่ในรูปแบบ ที่จะนำไปใช้งานกับ Model KNN, Decision tree ของ Scikit learn ทั้งนี้ผู้จัดทำได้เลือก Model Decision tree Classification เนื่องด้วยตัวข้อมูล label อยู่ในรูปแบบของ Multiclass ทำให้การทำงานเหมาะสมมากกว่า หลังจากทำการทดลองพบว่าข้อมูลที่มีนั้นมีน้อยเกินไปทำให้ค่า accuracy ของผลลัพท์ ไม่ถึง 20%

A picture containing text

Description automatically generatedText

Description automatically generated*รูปที่ 1 นำ data ที่ survey มา join กับ data ที่มีอยู่*

*รูปที่ 2 data ที่ survey มา* *รูปที่ 3 นำ data ของแต่ละ user มา transpose*

**หัวข้อการพัฒนาโครงงานเพิ่มเติม**

**1 Update Schema Student :** เพิ่ม status และ career ลงไปใน schema เพื่อให้ง่ายต่อการใช้งานและไม่เกิดความซ้ำซ้อน

**2 Function Update Student Status :** พัฒนา Function ที่จะ Update Status ทั้งหมดของข้อมูลนักศึกษาที่ใช้ในการ train model ในระบบ

**3 Function Update Student Career :** พัฒนา Function ที่จะ Update Career ทั้งหมดของข้อมูลนักศึกษาที่ใช้ในการ train model ในระบบ ทั้งในรูปแบบของการ upload file เข้ามาแล้ว update ทั้งหมด หรือ update ที่ละคนตาม request body

**4 Function Update Subject Class :** พัฒนา Function ที่จะ Update Subject Class ได้แบบ auto เมื่อมีการ upload file subject เข้ามาใหม่จะทำการใช้ NLP หา class และ update class ให้กับวิชาที่มีอยู่ในฐานข้อมูลอยู่แล้ว

Text

Description automatically generatedGraphical user interface

Description automatically generated

*รูปที่ 4 schema ที่ redesign มาใหม่เพิ่ม status และ Career**รูปที่ 5 Update function ให้รองรับการเพิ่มของวิชาใหม่*

## **ปัญหาที่เกิดขั้นและแนวทางการแก้ไข**

ปัญหาจากครั้งก่อน

1. ข้อมูลในการสร้างโมเดลมีไม่เพียงพอ ทำให้การทำงานในส่วนของ job prediction ยังคงต้องรอข้อมูลจากอาจารย์ที่ปรึกษาก่อน จากที่ปรึกษากับที่ปรึกษาเรื่องปัญหานี้แล้วทางผู้จัดทำได้ข้อสรุปว่าทางผู้จัดทำจะทำการเก็บ survey อีกครั้งตอนที่นักศึกษาที่จบไปแล้วกลับมารับปริญญา และทางที่ปรึกษาได้แนะนำให้เตรียม function สำหรับการเพิ่มข้อมูลเข้า database และทางผู้จัดทำได้ทำการพัฒนาส่วนของ function เรียบร้อยแล้ว

2. การเทรนข้อมูลมีการเทรนใหม่ทุกรอบ ทำให้ผลลัพธ์ออกมามีค่าที่ไม่นิ่ง จากปัญหานี้ทางผู้จัดทำได้ทำการสำรวจแล้วพบว่าปัญหาเกิดจากการที่ผู้จัดทำได้ทำการ train model ใหม่ทุกรอบที่มีการร้องข้อการ prediction ทำให้ค่าผลลัพท์ออกมาไม่เหมือนเดิม ดังนั้นผู้จัดทำจึงแก้ไขปัญหาโดยการนำ model ที่ train เสร็จแล้วลงฐานข้อมูลในรูปแบบของ Pickled File และเพิ่ม Table ในการเก็บ Model และเมื่อมีการเรียกใช้จะทำการเรียก Model มาใช้งานจากฐานข้อมูลซึ่งในรูปแบบ Demo ผู้ใช้งานสามารถเลือกใช้ Model ที่ต่างกันได้ในฐานข้อมูล

ปัญหา ณ ปัจจุบัน

1. การสร้าง file csv ผ่านภาษา Python เนื่องด้วยผู้จัดทำได้ทำการทดลองใช้งานแล้วพบปัญหาว่าเมื่อใช้ Program Microsoft Excel ในการ edit หรือ view ตัว file นั้นจะทำให้ภาษาไทยใน column curriculum นั้นกลายเป็นภาษาประหลาด

## **สิ่งที่จะดำเนินการต่อไป**

1. เขียนบทที่ 3

2. ปรับแก้ไข้ตัว generate csv file ให้เป็นปกติ โดยวิธีคร่าวๆคือ เมื่อ generate file จะทำการบอก encoding rule ให้ด้วยเพื่อให้ Microsoft Excel เปิดออกมาแล้วภาษาไทยเป็นปกติ

3 จัดระเบียบ Function ใน Demo App ให้เป็นไปตาม Usecase Diagram และ Sequence Diagram มากขึ้นเนื่องด้วยปัจจุบันมี Function ที่เกิดขึ้นมานอกเหนือจาก Usecase และ Sequence Diagram มากมาย จะจัดระเบียบโดยการไล่ดูที่ละ Usecase และ Sequence ตัด Function ที่ไม่จำเป็นออกรวมทั้งแก้ไขชื่อให้สอดคล้อง

4 เพิ่ม log การทำงานของตัวระบบ เนื่องด้วยบาง request ใช้เวลาในการ process ค่อนข้างนานผู้จัดทำเล็งเห็นถึงความไม่สะดวกในการ Monitoring Process จึงจะเพิ่ม print log ลงไปในแต่ละ Function ว่าทำงานถึง process ไหนแล้ว