

เว็บแอปพลิเคชันการสกัดข้อมูลจากรูปภาพสลิปเงินเดือนหลาย
รูปแบบ

WEB APPLICATION MULTIMODAL INFORMATION
EXTRACTION FOR PAYSリップ IMAGES



ปริญญาานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญา

วิศวกรรมศาสตรบัณฑิต

สาขาวิชาวิศวกรรมสารสนเทศ

คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้เพื่อใช้ในการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ปีการศึกษา 2563

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

WEB APPLICATION MULTIMODAL INFORMATION
EXTRACTION FOR PAYSLIP IMAGES



THE THESIS IS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
BACHELOR OF ENGINEERING IN INFORMATION ENGINEERING
FACULTY OF ENGINEERING

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG
ACADEMIC YEAR 2020

เอกสารนี้เป็นเอกสารที่จัดทำขึ้นเพื่อใช้ในการศึกษาเท่านั้น ไม่สามารถนำออกจำหน่ายหรือทำซ้ำโดยไม่ได้รับอนุญาต
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

หัวข้อปริญญานิพนธ์	เว็บแอปพลิเคชันการสกัดข้อมูลจากรูปภาพเอกสารแบบหลายรูปแบบ
ชื่อนักศึกษา	นาย ขจรยศ บุญเกตุ รหัสนักศึกษา 60010105
ปริญญา	วิศวกรรมศาสตรบัณฑิต
สาขาวิชา	วิศวกรรมสารสนเทศ
พ.ศ.	2563
อาจารย์ที่ปรึกษา	ผศ.ดร. เกล็ดดาว สัตย์เจริญ

ปริญญานิพนธ์ฉบับนี้ได้รับการอนุมัติให้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
วิศวกรรมศาสตรบัณฑิต คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหาร
ลาดกระบัง



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

หัวข้อปริญญานิพนธ์	เว็บแอปพลิเคชันการสกัดข้อมูลจากรูปภาพสลิปเงินเดือนหลายรูปแบบ
ชื่อนักศึกษา	นาย ขจรยศ บุญเกตุ รหัสนักศึกษา 60010105
ปริญญา	วิศวกรรมศาสตรบัณฑิต
สาขาวิชา	วิศวกรรมสารสนเทศ
พ.ศ.	2563
อาจารย์ที่ปรึกษาปริญญานิพนธ์	ผศ.ดร. เกียรติดาว สัตย์เจริญ

บทคัดย่อ

จากในปัจจุบันนี้โลกกำลังเข้าสู่ยุคข้อมูลดิจิทัลมากขึ้นเรื่อย ๆ แต่การขั้นตอนหลายอย่างในทางธุรกิจนั้นยังคงมีการใช้กระดาษกันอย่างมาก เช่นเดียวกันกับในระบบการดำเนินการในสถาบันทางการเงินที่จะมีการรับมือกับการประมวลผลเอกสารที่เป็นกระดาษอย่างมาก เช่น Pay Slip เว็บแอปพลิเคชันที่ผู้จัดทำทำขึ้นนั้นมีความสามารถในการดึงข้อมูลที่จำเป็นหรือส่วนที่ต้องการจะนำไปใช้ในรูปภาพสลิปเงินเดือนที่มีหลายรูปแบบได้เพื่อให้ง่ายต่อการนำข้อมูลจากรูปภาพสลิปเงินเดือนไปใช้ต่อในการประมวลผลในขั้นตอนอื่น ๆ ได้อย่างรวดเร็วการใช้งานเว็บแอปพลิเคชันสามารถทำได้โดยทำการอัปโหลดรูปภาพสลิปเงินเดือนเข้าไปในตัวระบบหลังจากนั้นระบบจะทำการประมวลผลและทำการแปลงรูปภาพเอกสารให้ออกมาเป็นข้อความและจะทำการทำนายฟิลด์ที่ผู้ใช้ต้องการออกมาเช่นฟิลด์ของ ชื่อ-นามสกุล หรือ ฟิลด์ของชื่อบริษัท เป็นต้น และสามารถที่จะแก้ไขในกรณีที่มีการทำนายผิดพลาด และสามารถเพิ่มฟิลด์ที่ต้องการเพิ่มได้ หลังจากนั้นจะสามารถเลือกส่งออกข้อมูลให้ไปอยู่ในรูปแบบของไฟล์เอกสารหรือ ส่งไปทางอีเมลก็ได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

Thesis Title	WEB APPLICATION MULTIMODAL INFORMATION EXTRACTION FOR PAYSLIP IMAGE
Student	Mr.Khajohnyos Boonkate Student ID 60010105
Degree	Bachelor of Engineering
Program	Information Engineering
Year	2020
Thesis Advisor	Asst.Prof. Dr. Kleddao Satcharoen

ABSTRACT

Nowadays, the world is entering more and more digital Information age, but many of the business processes are still heavily used in paper. The same is true for financial institutions to deal with the processing of paper documents such as statements or payslip. The authors developed web application which extract the necessary information or the desired part to be used in the images, documents that are multi-format, so that it is easy to use the information from the document for further processing. Web application can be done by uploading images and documents into the system. After that, the system will process and convert the payslip image to text and will predict the desired field for the user such as name field or company name field etc. and can be edited in case of wrong prediction and can add the fields you want to add after that you can choose to export the data to the form of a document file or send it by email.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

||

Forbidden to modify the content, and cite the document when use.

กิตติกรรมประกาศ

ปริญญานิพนธ์ฉบับนี้สำเร็จขึ้นได้นั้นต้องขอขอบคุณ ผศ.ดร. เกียรติดาว สัตย์เจริญ ที่ให้คำปรึกษาและตรวจสอบความถูกต้องของเล่มปริญญานิพนธ์ และขอขอบคุณพี่ ๆ ที่ได้ทำสื่อความรู้ คำแนะนำและวิธีการประยุกต์ใช้งานต่าง ๆ ในบนอินเทอร์เน็ต

ขอขอบพระคุณอาจารย์ทุกท่าน ที่ได้ให้ความรู้และคำแนะนำต่าง ๆ มากมายตลอดทั้งการศึกษา รวมถึงเพื่อน ๆ พี่ ๆ และน้อง ๆ ที่คอยให้การสนับสนุนช่วยเหลือ ให้คำปรึกษาและเป็นการกำลังใจ

ท้ายที่สุดผู้จัดทำขอขอบพระคุณ บิดา มารดา ซึ่งเป็นบุคคลที่มีส่วนสำคัญในการสนับสนุนในทุก ๆ ด้าน ทางผู้จัดทำหวังเป็นอย่างยิ่งว่าปริญญานิพนธ์ฉบับนี้จะเป็นประโยชน์ให้กับผู้ที่สนใจเป็นอย่างสูงสุด



ขจรยศ บุญเกตุ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

สารบัญ

หน้า

บทคัดย่อภาษาไทย.....	I
บทคัดย่อภาษาอังกฤษ.....	II
กิตติกรรมประกาศ.....	III
สารบัญ.....	IV
สารบัญตาราง.....	IX
สารบัญรูปภาพ.....	X
บทที่ 1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของโครงการ.....	1
1.2 วัตถุประสงค์ของโครงการ.....	1
1.3 ขอบเขตของโครงการ.....	1
1.4 อุปกรณ์ที่ใช้ในโครงการ.....	2
1.4.1 ซอฟต์แวร์ (Software).....	2
1.4.2 ฮาร์ดแวร์ (Hardware).....	2
1.4.3 อัลกอริทึม (Algorithms).....	2
1.5 ประโยชน์ที่คาดว่าจะได้รับ.....	2
1.6 ตารางเวลาที่ใช้ในการดำเนินโครงการ.....	3
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง.....	4
2.1 ทฤษฎีที่เกี่ยวข้องกับการทำงานของโครงการ.....	4
2.1.1 Machine Learning.....	4
2.1.1.1 ประเภทของ Machine Learning.....	4
2.1.1.2 Artificial Neural Network.....	5
2.1.1.3 Activation Function.....	7
2.1.1.4 Cost Function.....	7

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นิยมนำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมีเหตุใดเปลี่ยนแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

สารบัญ (ต่อ)

2.1.1.5 Forward Propagation.....	8
2.1.1.6 Backward Propagation	10
2.1.1.7 Under Fitting / Good Fitting / Over Fitting	11
2.1.1.8 Optimization	11
2.1.1.9 การวัดผลของโมเดลในการทำ Classification.....	12
2.1.2 Graph Convolutional Network.....	13
2.1.3 Computer Vision.....	14
2.1.3.1 Optical Character Recognition (OCR).....	14
2.1.4 Image Processing.....	15
2.1.4.1 Sauvola Local Thresholding	15
2.1.4.2 Morphology.....	16
2.1.5 Natural Language Processing.....	18
2.1.5.1 Cosine Distance.....	18
2.1.6 RESTful API.....	19
2.1.7 HTTP Protocol.....	20
2.1.7.1 HTTP Request.....	20
2.1.7.2 HTTP Response.....	21
2.2 ภาษาและเฟรมเวิร์กที่ใช้ในการพัฒนาในโครงการ	23
2.2.1 ภาษา.....	23
2.2.1.1 Python.....	23
2.2.1.2 JavaScript	24
2.2.1.3 SQL	24
2.2.2 เฟรมเวิร์ก.....	25
2.2.2.1 Django & Django Rest Framework.....	25

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

สารบัญ (ต่อ)

2.2.2.2 ReactJS.....	26
2.3 โปรแกรมที่ใช้ในการพัฒนาโครงการ.....	27
2.3.1 Visual Studio Code.....	27
2.3.2 Docker.....	28
2.3.3 Git.....	29
บทที่ 3 วิธีการดำเนินงานโครงการ.....	30
3.1 ภาพรวมของโครงการ.....	30
3.2 การออกแบบหน้าเว็บแอปพลิเคชันและเซิร์ฟเวอร์.....	31
3.2.1 การออกแบบหน้าเว็บแอปพลิเคชัน.....	31
3.2.2 การออกแบบเว็บเซิร์ฟเวอร์.....	36
3.3 ขั้นตอนการ Deployment เว็บแอปพลิเคชันและเซิร์ฟเวอร์ขึ้นบนคลาวด์เซิร์ฟเวอร์.....	41
3.3.1 การ Deploy เว็บแอปพลิเคชันขึ้นบน Heroku Cloud Platform.....	41
3.3.2 การ Deploy เซิร์ฟเวอร์ขึ้นบน Google Cloud Platform.....	42
3.4 ขั้นตอนการพัฒนากระบวนการทำนายข้อมูลบนรูปภาพสลิปเงินเดือน.....	43
3.4.1 การสร้างกราฟของข้อมูล.....	43
3.4.1.1 การปรับ Thresholding และการลบเส้นตารางออกจากรูปภาพสลิปเงินเดือน.....	44
3.4.2 การเตรียมชุดข้อมูลฝึกและข้อมูลทดสอบ.....	45
3.4.2.1 การ Label ข้อมูลในแต่ละรูปภาพ.....	45
3.4.3 การเลือกคุณสมบัติและการแปลงข้อมูล.....	45
3.4.3.1 การเลือกคุณสมบัติ.....	45
3.4.3.2 การแปลงข้อมูลเพื่อเข้าสู่การฝึกฝนและทดสอบ.....	46
3.4.4 การฝึกฝนโมเดล.....	47
3.5 ขั้นตอนการพัฒนากระบวนการแก้ค่าและประโยคผิดโดยใช้ความใกล้เคียงกันของคำ.....	48
3.5.1 การเตรียมชุดค่าและประโยค.....	48

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับอาจารย์ใช้เท่านั้น เพื่อการสื่อสารเท่านั้น ไม่สามารถนำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

สารบัญ (ต่อ)

3.5.2 การแก้คำและประโยคที่ผิด.....	48
บทที่ 4 ผลการดำเนินงาน	49
4.1 ผลการทดสอบการทำงานของหน้าเว็บแอปพลิเคชัน.....	49
4.1.1 หน้าเข้าสู่ระบบ.....	49
4.1.2 หน้าลงทะเบียน.....	50
4.1.3 หน้าหลัก.....	51
4.1.4 หน้าอัปโหลดรูปภาพสลิปเงินเดือน	52
4.1.5 หน้าแสดงผลและแก้ไขข้อมูล.....	53
4.2 ผลการทดสอบการทำงานของเซอร์วิส.....	54
4.2.1 การตรวจสอบและยืนยันตัวตนการเข้าถึงหน้าเว็บแอปพลิเคชัน	54
4.2.2 การลงทะเบียน	55
4.2.3 การอัปโหลดรูปภาพและการดึงข้อมูลฟิลด์โดยใช้โมเดลการทำนายผล	56
4.2.4 แสดงข้อมูลรายการที่เรียกใช้การอัปโหลดรูปภาพและการดึงข้อมูลฟิลด์	57
4.2.5 แสดงผลลัพธ์การดึงข้อมูลฟิลด์โดยใช้โมเดลการทำนายผลและแก้ไขข้อมูล	58
4.3 ผลการแปลงข้อมูลเพื่อเข้าสู่การฝึกฝนและทดสอบ	59
4.4 ผลการฝึกฝนและทดสอบโมเดลทำนายผล.....	59
4.4.1 ผลเปรียบเทียบการเลือกคุณสมบัติของข้อมูล.....	61
4.5 ผลการเรียกใช้การดึงข้อมูลด้วยโมเดลการทำนายผลผลการปรับ Thresholding และการลบ เส้นตารางในรูปภาพสลิปเงินเดือน	63
4.6 ผลการเรียกใช้ OCR และการจัดรูปแบบ	64
4.7 ผลการเรียกใช้การแก้ประโยคและคำผิด.....	65
4.8 ผลการเรียกใช้การดึงข้อมูลด้วยโมเดลการทำนายผล.....	66
บทที่ 5 สรุปผลการดำเนินงาน	68
5.1 สรุปผลการทดลอง.....	68

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับใช้ในงานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น ลีโอนี่ หวังเป็ ให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

5.2 ปัญหาและอุปสรรคของการดำเนินงาน.....	68
5.3 ข้อเสนอแนะและแนวทางการแก้ไข.....	68
บรรณานุกรม.....	69
ประวัติผู้จัดทำ.....	74



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

สารบัญตาราง

ตารางที่	หน้า
2.1 แสดงตัวอย่างของ Activation Function.....	7
3.1 แสดง HTTP Status Code ของเซอร์วิส Authentication.....	37
3.2 แสดง HTTP Status Code ของเซอร์วิสการลงทะเบียน.....	37
3.3 แสดง HTTP Status Code ของเซอร์วิสการอัปโหลดรูปภาพและการดึงข้อมูล.....	38
3.4 แสดง HTTP Status Code ของการแสดงผลการดึงข้อมูลไฟล์โดยใช้โมเดลการทำนายผล.....	40
3.5 แสดง HTTP Status Code ของการแก้ไขข้อมูล.....	40
3.1 ตารางแสดงค่า Hyperparameter ที่ใช้ในการตั้งค่า model.....	47
4.1 ตารางแสดง Confusion Matrix ของการฝึกฝนโมเดล.....	60
4.2 ตารางแสดง Confusion Matrix ของการทดสอบโมเดล.....	61
4.3 ตารางเปรียบเทียบค่า Precision, Recall และ F1-Score ของการฝึกโมเดล.....	61
4.4 ตารางเปรียบเทียบค่า Precision, Recall และ F1-Score ของการทดสอบโมเดล.....	62

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

สารบัญรูปภาพ

รูปภาพที่	หน้า
2.1 ภาพแสดงโครงข่ายประสาทของมนุษย์.....	5
2.2 ภาพแสดงโครงข่ายประสาทเทียม.....	6
2.3 ภาพแสดง Forward และ Backward Propagation.....	8
2.4 ภาพแสดงโครงข่ายประสาทเทียมที่มี 1 hidden layer.....	9
2.5 ภาพแสดงการทำ Backward Propagation เบื้องต้น.....	10
2.6 ภาพแสดงความแตกต่างของ Underfitting, Good Fit และ Overfitting.....	11
2.8 ภาพแสดงตัวอย่างคอนเซปของ Graph Convolutional Network.....	13
2.9 ภาพแสดงตัวอย่าง GCN 2-layer.....	14
2.10 ภาพแสดงขั้นตอนการทำ OCR.....	15
2.11 รูปภาพก่อนที่จะถูกดำเนินการ.....	16
2.12 รูปภาพหลังทำการ Erosion ด้วยเคอร์เนลขนาด 5x5.....	17
2.12 รูปภาพหลังทำการ Dilation ด้วยเคอร์เนลขนาด 5x5.....	17
2.14 รูปโลโก้ของภาษาไพทอน.....	24
2.15 ตัวอย่างไวยากรณ์ของภาษาไพทอน.....	24
2.15 ตัวอย่างไวยากรณ์ของภาษา JavaScript.....	24
2.16 ตัวอย่างการใช้งานภาษา SQL.....	25
2.17 ภาพการทำงานคร่าว ๆ ของ Django Framework.....	26
2.18 โลโก้ของ ReactJS.....	27
2.19 ตัวอย่างการใช้ Extension Github บน VSCode.....	27
2.20 การทำงานของ Docker ในแต่ละแอปพลิเคชัน.....	28
2.21 โลโก้ของ Git.....	29

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้เผยแพร่ไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดก็ตาม ห้ามทำซ้ำหรือดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

X

Forbidden to modify the content, and cite the document when use.

3.3 ภาพแสดง Flow ของหน้าเข้าสู่ระบบ	32
3.4 ภาพแสดง Flow ของหน้าลงทะเบียน.....	33
3.5 ภาพแสดง Flow ของหน้าหลัก.....	34
3.6 ภาพแสดง Flow ของหน้าอัปโหลดรูปภาพสลิปเงินเดือน	34
3.7 ภาพแสดง Flow ของหน้าแสดงผลและแก้ไขข้อมูล	35
3.8 ภาพแสดง Sequence Diagram ของเซอร์วิสตรวจสอบยืนยันตัวตน	36
3.9 ภาพแสดง Sequence Diagram ของเซอร์วิสการลงทะเบียน	37
3.10 ภาพแสดง Sequence Diagram ของเซอร์วิสการอัปโหลดรูปภาพและการดึงข้อมูล.....	38
3.11 ภาพแสดง Sequence Diagram ของเซอร์วิสแสดงรายการ.....	39
3.12 ภาพแสดง Sequence Diagram ของเซอร์วิสแสดงผลลัพธ์	39
3.13 ภาพแสดง Sequence Diagram ของเซอร์วิสการแก้ไขผลลัพธ์	40
3.14 แสดงการสร้าง Heroku app	41
3.15 แสดงการเชื่อมต่อ Heroku app กับ Github.....	41
3.16 แสดงการสร้าง VM instance บน Google Cloud Platform.....	42
3.17 แสดงการ git clone โปรเจกต์ลงบน VM instance.....	42
3.18 แสดงการรันคำสั่งเพื่อให้เซอร์วิสทำงาน.....	43
3.19 แสดงภาพรวมของระบบการถ่ายโอนข้อมูลบนรูปภาพสลิปเงินเดือน	43
3.20 แสดงตัวอย่างการสร้างกราฟ	44
3.21 แสดงตัวอย่างของโปรแกรม Label	45
3.22 ภาพแสดงตัวอย่างการตัดคำ.....	46
3.23 ภาพแสดงตัวอย่างการแปลงข้อมูล.....	46
3.24 ภาพแสดง Layer และ Activation Function.....	47
3.25 แสดงตัวอย่างของชุดค่าและประโยคที่ใช้ในการแก้คำผิด (a) คำ (b) ประโยค.....	48
3.26 แสดงการแปลงคำหรือประโยคสำหรับการแก้ไข.....	48

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมีเหตุดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

4.3	รูปแสดงหน้าลงทะเบียน.....	50
4.4	รูปแสดงหน้าลงทะเบียน.....	51
4.5	รูปแสดงหน้าหลัก.....	51
4.6	รูปแสดงหน้าอัปโหลดรูปภาพ.....	52
4.7	รูปแสดงหน้าอัปโหลดรูปภาพ.....	53
4.8	รูปแสดงหน้าแสดงผลลัพธ์.....	53
4.9	รูปแสดงหน้าแสดงผลลัพธ์.....	54
4.10	รูปภาพแสดงข้อมูลที่เซอร์วิสการยืนยันตัวตนต้องการ.....	55
4.11	รูปภาพตัวอย่างการตอบกลับของเซอร์วิสการยืนยันตัวตน.....	55
4.12	รูปภาพตัวอย่างการตอบกลับของเซอร์วิสการยืนยันตัวตน.....	55
4.13	รูปภาพแสดงข้อมูลที่เซอร์วิสการลงทะเบียนต้องการ.....	56
4.14	รูปภาพแสดงการตอบกลับของเซอร์วิสลงทะเบียนไม่สำเร็จ.....	56
4.15	รูปภาพแสดงการตอบกลับของเซอร์วิสลงทะเบียนสำเร็จ.....	56
4.16	รูปภาพแสดงการตอบกลับของเซอร์วิสการอัปโหลดรูปภาพ.....	57
4.17	รูปภาพแสดงการตอบกลับของเซอร์วิสการอัปโหลดรูปภาพ.....	57
4.18	รูปภาพแสดงการตอบกลับของเซอร์วิสการอัปโหลดรูปภาพ.....	57
4.19	รูปภาพแสดงการตอบกลับของเซอร์วิสผลลัพธ์การดึงข้อมูล.....	58
4.20	รูปภาพแสดงการตอบกลับของเซอร์วิสผลลัพธ์การดึงข้อมูล.....	58
4.21	ตัวอย่างการแปลงข้อมูลให้เป็นไฟล์ “.cites”.....	59
4.22	ตัวอย่างการแปลงข้อมูลให้เป็นไฟล์ “.content”.....	59
4.23	รูปภาพแสดงค่า precision, recall and f1-score ของการฝึกฝนในแต่ละคลาส.....	60
4.24	รูปภาพแสดงค่า precision, recall and f1-score ของการทดสอบในแต่ละคลาส.....	60
4.22	รูปภาพแสดงตัวอย่างรูปภาพก่อนทำการปรับ Threshold.....	63
4.23	รูปภาพแสดงการปรับ Threshold ของรูปภาพ.....	64
4.24	รูปภาพแสดงการลบเส้นตารางออกจากในรูปภาพ.....	64

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นอนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมีเหตุดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

4.25	รูปภาพแสดงผลลัพธ์การเรียกใช้ OCR.....	65
4.26	รูปภาพแสดงการจัดรูปแบบใหม่.....	65
4.27	รูปภาพแสดงการแก้ประโยค.....	66
4.28	รูปภาพแสดงการแก้คำ.....	66
4.29	รูปภาพแสดงการผลของการใช้โมเดลทำนายผล.....	66
4.30	รูปภาพแสดงการผลของการใช้โมเดลทำนายผล.....	67



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของโครงการ

เนื่องจากในปัจจุบันการดำเนินการประมวลผลในระบบสถาบันทางการเงินในบางกระบวนการจะต้องรับมือกับการรับมือกับการประมวลผลเอกสารที่เป็นกระดาษอย่างมาก ซึ่งทำให้เกิดการเสียเวลาในการนั่งพิมพ์หรือการกรอกข้อมูลจึงทำให้เกิดเทคโนโลยี OCR หรือ Optical Character Recognition เป็นเทคโนโลยีที่ช่วยในการแปลงตัวอักษรในรูปภาพให้ออกมาเป็นข้อความได้แต่การใช้เพียงแค่การทำ OCR นั้นก็ยังไม่สามารถช่วยตอบโจทย์ปัญหานี้ได้เพราะข้อมูลที่ได้ออกมานั้นในการที่จะเอาไปใช้ต่อเลยนั้นทำไม่ได้ เว็บแอปพลิเคชันการสกัดข้อมูลจากรูปภาพสลิปเงินเดือนหลายรูปแบบนั้นจะช่วยให้การประมวลผลเอกสารที่เป็นกระดาษนั้นได้ง่ายขึ้น และสามารถที่จะสกัดข้อมูลที่ต้องการหรือที่จำเป็นต้องใช้ และตรวจสอบความถูกต้องของข้อมูลและแก้ไขได้อย่างรวดเร็วเพื่อให้พร้อมสำหรับการนำไปใช้งานต่อในงานส่วนอื่น ๆ ได้อย่างง่ายโดยสามารถส่งออกมาเป็นในรูปแบบของไฟล์เอกสารหรือจะเป็นการส่งอีเมลให้พนักงานท่านอื่นก็ทำได้อย่างง่ายดาย ซึ่งจะช่วยให้ลดภาระงานที่ไม่จำเป็นและยังช่วยประหยัดเวลาไปทำงานในส่วนที่สำคัญกว่าได้

1.2 วัตถุประสงค์ของโครงการ

- 1.2.1 เพื่อให้สามารถทำการดึงข้อมูลจากรูปภาพสลิปเงินเดือนได้
- 1.2.2 เพื่อดึงข้อมูลที่ต้องการแล้วสามารถนำไปใช้ต่อในกระบวนการอื่น ๆ ได้ง่าย
- 1.2.3 เพื่อให้สามารถทำการแก้ไขข้อมูลในส่วนที่มีการดึงข้อมูลผิดพลาดได้ง่าย

1.3 ขอบเขตของโครงการ

สามารถทำการดึงข้อมูลที่ต้องการในรูปภาพสลิปเงินเดือนที่มีหลากหลายรูปแบบตามแต่ละบริษัทแล้วสามารถนำไปใช้ต่อในรูปแบบของ key-value เพื่อให้ง่ายต่อการนำไปใช้ต่อในกระบวนการอื่น ๆ ได้และสามารถแก้ไขในส่วนที่มีการทำนายผิดได้อย่างง่ายและสะดวก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

1.4 อุปกรณ์ที่ใช้ในโครงการ

1.4.1 ซอฟต์แวร์ (Software)

- Python
- JavaScript
- Google Cloud Platform
- ReactJS
- Django & Django Rest Framework
- Database SQL
- HTML

1.4.2 ฮาร์ดแวร์ (Hardware)

- คอมพิวเตอร์ หรือ แล็ปทอป ที่สามารถใช้เขียนโปรแกรมได้

1.4.3 อัลกอริทึม (Algorithms)

- Graph Convolutional Network
- Graph Theory
- Natural Language Processing
- Cosine Distance

1.5 ประโยชน์ที่คาดว่าจะได้รับ

- 1.5.1 ได้รับความรู้ในเรื่องของการใช้ Django and Django Rest Framework ในการทำ Backend
- 1.5.2 ได้รับความรู้ในเรื่องของการใช้ Django and Django Rest Framework ในการทำ Backend
- 1.5.3 ได้รับความรู้ในเรื่องของการใช้ React.js Framework ในส่วนของ Frontend
- 1.5.4 ได้รับความรู้ในเรื่องของการใช้ Google Cloud Platform
- 1.5.5 ได้รับความรู้ในเรื่องของการใช้ของการเขียน Restful API และ HTTP Protocol
- 1.5.6 ได้รับความรู้ในเรื่องของการใช้ Semi - Graph Convolutional Network
- 1.5.7 ได้รับความรู้ในเรื่องของ Machine Learning and Neural Network

เอกสารนี้เป็นเอกสารที่ 1.5.8 ทำให้สามารถดึงข้อมูลจากรูปภาพสลิปเงินเดือนได้ง่ายขึ้นให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น 1.5.9 ได้รับความรู้เรื่องของ Image Processing อิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

1.6 ตารางเวลาที่ใช้ในการดำเนินโครงการ

หัวข้อ	2020					2021				
	8	9	10	11	12	1	2	3	4	5
1.วางแผนและหา แนวทางการศึกษา	■									
2.ศึกษาเรื่องของเ ทคโนโลยีที่จำเป็น		■	■	■						
3.ศึกษาการทำ API				■	■					
4.ศึกษาการทำแล ะออกแบบ UX/UI				■	■	■	■			
5.ศึกษาเรื่องของ GCN					■	■				
6.เขียนการตรวจสอบ การทำงาน OCR						■				
7. ทำการเชื่อมต่อ API กับ Front end								■		
8.ทำการปรับปรุง GCN							■	■	■	
9.ทำการทดสอบร บบและปรับปรุง								■		
10. จัดทำรูปเล่มปริณ ญาบัตร					■				■	■

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ทฤษฎีที่เกี่ยวข้องกับการทำงานของโครงการ

2.1.1 Machine Learning

Machine Learning หรือ การเรียนรู้ของเครื่องคอมพิวเตอร์นั้นเป็นหนึ่งในสาขาวิทยาศาสตร์คอมพิวเตอร์ที่เติบโตเร็วที่สุด เริ่มต้นจากปี ค.ศ. 1959 อาร์เธอร์ ซามูเอล นักวิทยาศาสตร์คอมพิวเตอร์ผู้บุกเบิกการศึกษาปัญญาประดิษฐ์กล่าวว่า การเรียนรู้ของเครื่องคอมพิวเตอร์เป็นการศึกษาที่ช่วยให้คอมพิวเตอร์สามารถเรียนรู้ได้โดยไม่ต้องตั้งโปรแกรมไว้อย่างชัดเจน เอกสารสรุปข้อมูลของ Alan Turing (Turing, 1950) ได้นำเสนอมาตรฐานสำหรับการแสดงความฉลาดของเครื่องจักรเช่นว่า เครื่องจักรจะต้องมีความชาญฉลาดและตอบสนองในลักษณะที่ไม่สามารถแยกความแตกต่างจากของมนุษย์ได้ คำจำกัดความทางเทคนิคเพิ่มเติมจาก Tom M. Mitchell's (1997) บอกเอาไว้ว่า “โปรแกรมคอมพิวเตอร์ถูกกล่าวว่าเป็นการเรียนรู้จากประสบการณ์ E เกี่ยวกับงานบางประเภท T และการวัดประสิทธิภาพ P หากประสิทธิภาพของงานใน T ตามที่ P วัดปรับปรุงด้วยประสบการณ์ E”

2.1.1.1 ประเภทของ Machine Learning

1. Supervised Learning เป็นการเรียนรู้โดยที่จำเป็นต้องมีข้อมูลหรือชุดข้อมูลฝึกฝนพร้อมกับการ Labels หรือเป้าหมายสำหรับแต่ละข้อมูลไว้ด้วย โดยที่ Labels ในแต่ละข้อมูลนั้นจะช่วยอัลกอริทึมในการเชื่อมโยงคุณลักษณะต่าง ๆ โดยที่ Supervised Learning จะมีสองวิธีที่พบเจอได้บ่อย คือแบบ Classification จะเป็นการจำแนกประเภทของข้อมูลที่เป็นรูปแบบของค่าที่ไม่ต่อเนื่อง (discrete) ซึ่งก็คือเครื่องคอมพิวเตอร์จะทำการคาดคะเนหรือทำนาย (Prediction) ว่าจะมีโอกาสเป็นคลาสหรือหมวดหมู่ไหนมากที่สุด ในการประยุกต์ใช้ Classification เราจะเห็นตัวอย่างได้เช่น การทำโมเดล classification ในการทำนายรูปภาพว่าจะเป็นรูปภาพนั้นเป็นสุนัขหรือแมว ซึ่งตัวโมเดลนี้จะต้องใช้เรื่องของ Computer Vision เข้ามาช่วยด้วย แบบที่สองก็คือ Regression จะถูกใช้สำหรับการทำนายค่าของข้อมูลที่มีความต่อเนื่องกัน (continuous) ตัวอย่างได้แก่ การทำนายยอดขายสำหรับในแต่ละเดือนโดย

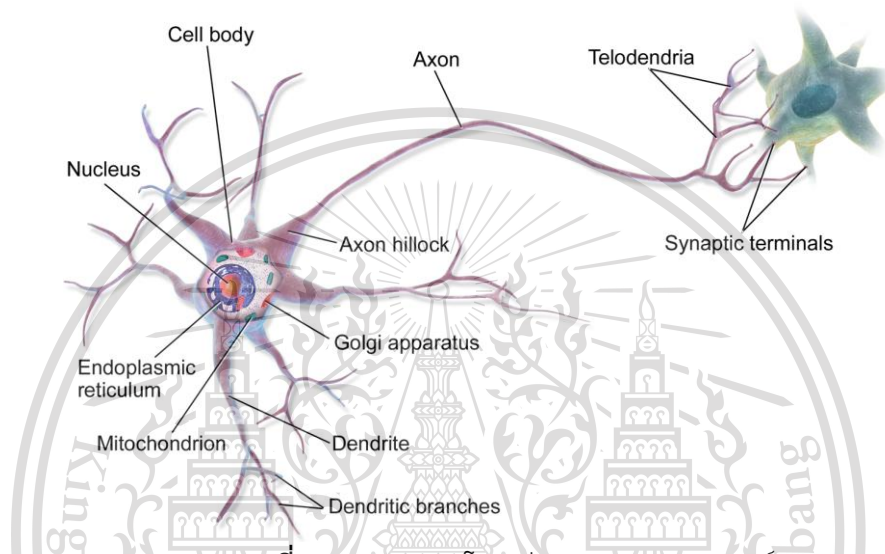
2. Unsupervised Learning เป็นการเรียนรู้ที่แตกต่างจาก Supervised Learning คือไม่จำเป็นที่จะมีชุดข้อมูลฝึกฝนมาให้แต่จะสามารถเรียนรู้ได้ด้วยการเรียนรู้จากรูปแบบของข้อมูล เอกสารนี้เป็นเอกสารที่ห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

ซึ่งงานส่วนใหญ่ นั่นคือ การจัดหมวดหมู่หรือจัดกลุ่มของข้อมูล (Clustering) โดยใช้อัลกอริทึมเช่น K Nearest Neighbors, K Mean Clustering เป็นต้น

3. Reinforcement Learning เป็นการเรียนรู้ที่มีเป้าหมายมุ่งเน้นชัดเจนซึ่งจะเรียนรู้วิธีการเพื่อจะบรรลุเป้าหมายหรือวัตถุประสงค์วิธีนี้จะช่วยให้คอมพิวเตอร์สามารถเรียนรู้พฤติกรรมที่เหมาะสมโดยอัตโนมัติภายในบริบทเฉพาะโดยจะใช้หลักการของการให้รางวัลหรือลบคะแนน เช่น การเพิ่มคะแนนเมื่อทำเป้าหมายได้สำเร็จและทำการลบคะแนนลงเมื่อทำเป้าหมายได้ไม่ถูกต้อง

2.1.1.2 Artificial Neural Network



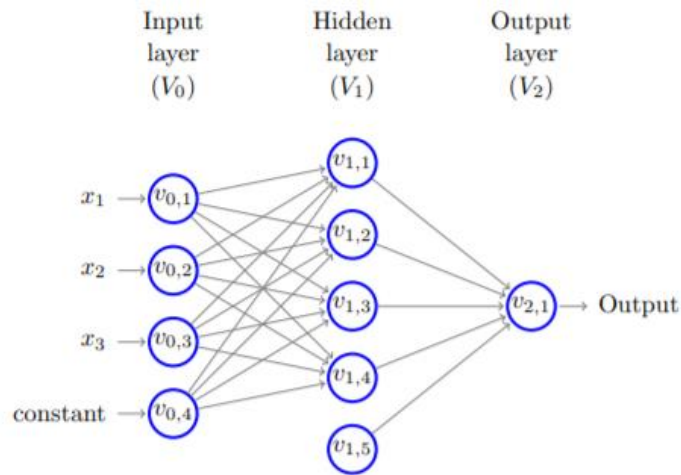
รูปภาพที่ 2.1 ภาพแสดงโครงข่ายประสาทของมนุษย์

อ้างอิง <https://www.oreilly.com/library/view/neural-networks-and/9781492037354/ch01.html>

Artificial Neural Network (Neural Network) หรือโครงข่ายประสาทเทียมนั้น เป็นการจำลองของการคำนวณที่ได้รับแรงบันดาลใจจากโครงข่ายประสาทในสมองของมนุษย์ ซึ่งในแบบโครงข่ายประสาทเทียมนั้นจะประกอบไปด้วยเซลล์ประสาท (Neurons) ที่เชื่อมต่อกันในโครงข่ายประสาทที่ซับซ้อนซึ่งสามารถทำการคำนวณที่ซับซ้อนสูงได้ โครงข่ายประสาทเทียมสามารถอธิบายได้ว่าเป็นกราฟที่มีโหนด (Node) ที่เปรียบเสมือนเป็นเซลล์ประสาทและมีการเชื่อมโยงระหว่างกันโดยใช้ผลรวมการถ่วงน้ำหนัก (Weight) ของผลลัพธ์ของเซลล์ประสาทที่เชื่อมต่อกัน โดยโครงข่ายประสาทเทียมจะประกอบไปด้วย 5 ส่วนหลักได้แก่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.



รูปภาพที่ 2.2 ภาพแสดงโครงข่ายประสาทเทียม

อ้างอิง Understanding Machine Learning From Theory to Algorithms: 2014 by Shai Shalev-Shwartz and Shai Ben-David

1. Neurons หรือจะถูกเรียกว่า Node ซึ่งเป็นหน่วยของข้อมูลซึ่งจะแตกต่างกันไปตามเลเยอร์ ที่มันอยู่เช่นถ้าอยู่ใน Input layer นั้นเจ้าตัว Node ก็จะเป็นหน่วยข้อมูลของข้อมูลที่รับเข้ามา ถ้าอยู่ใน Hidden layer ก็จะมีการคำนวณทางคณิตศาสตร์เข้ามาช่วยเพื่อให้ปรับข้อมูลใน Node นั้น

2. Input layer เป็นเลเยอร์ที่มีหน้าที่ได้การรับข้อมูล (Input) เข้ามาและส่งต่อไปยัง layer ต่อไป (Hidden layer) ในโครงข่ายประสาทเทียมโดย Input layer นั้นจะมีเพียงแค่ชั้นเดียวเท่านั้นและจำนวนของ Neurons หรือ Node นั้นจะขึ้นอยู่กับจำนวนของข้อมูลที่นำเข้ามา

3. Hidden layer เป็นเลเยอร์ที่จะทำการคำนวณทางคณิตศาสตร์ต่าง ๆ โดยจะใช้ค่า Weight และ Bias เข้ามาคำนวณด้วยโดย Hidden layer นั้นจะสามารถมีได้มากกว่า 1 เลเยอร์ โดยจะขึ้นอยู่กับปรับให้เหมาะสมในแต่ละโมเดลการใช้งานการเพิ่มเลเยอร์ในชั้นนี้นั้นอาจจะทำให้โครงข่ายประสาทเทียมนั้นมีความฉลาดขึ้นหรือไม่ก็ได้

4. Output layer เป็นเลเยอร์สุดท้ายของโครงข่ายประสาทเทียมโดยจำนวน Neurons หรือ Node นั้นจะขึ้นอยู่กับจำนวนของคลาสหรือจำนวนของหมวดหมู่ที่เราต้องการโดยค่าในแต่ละ Node ของเลเยอร์ นั้นนั้นจะมีค่าได้มาจาก Hidden layer โดยที่จะมีค่าไม่เท่ากัน

5. Weight and Bias จะเป็นส่วนที่เชื่อมต่อ Hidden layer กับ Input layer และ Output layer หรือเชื่อมต่อกับ Hidden layer เองโดยที่ Bias จะมีไว้เพื่อให้การคำนวณทางคณิตศาสตร์ต่าง ๆ ให้มีความเท่าเทียมมากขึ้น ส่วนค่า Weight นั้นจะเป็นค่าที่ทำหน้าที่เข้าไปถ่วงน้ำหนักให้กับแต่ละ Node เพื่อที่จะกำหนดความสำคัญโดยที่ Node ที่มี Weight มากนั้นจะช่วยให้

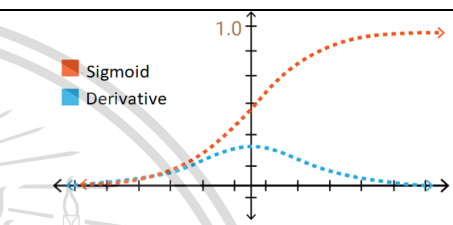
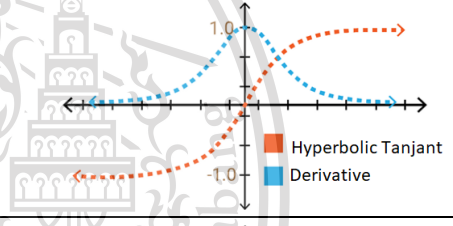
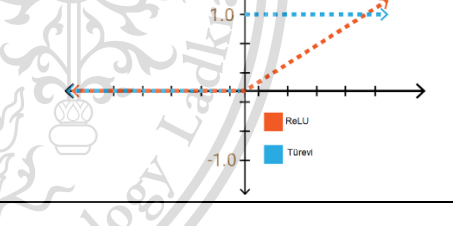
Output ที่ส่งออกไปมีความสำคัญมากขึ้นเมื่อเทียบกับตัวอื่น โดยค่า Weight เริ่มต้นนั้นส่วนใหญ่จะ
ไม่ว่าการคิดใช้การสุ่มที่มีค่าน้อย ๆ เพื่อให้โครงข่ายประสาทเทียมสามารถเรียนรู้ได้ง่ายขึ้น

2.1.1.3 Activation Function

Activation Function เป็นส่วนที่สำคัญในโครงข่ายประสาทเทียมโดยส่วนใหญ่จะถูกใช้ใน Hidden layer โดยตัว Activation Function นั้นอาจจะถูกเรียกว่า Transfer Function เป็นฟังก์ชันที่ใช้สำหรับการหาเอาต์พุตของแต่ละโหนดโดยใช้การดำเนินการทางคณิตศาสตร์ สามารถแบ่งออกได้เป็น 2 แบบคือ

1. ฟังก์ชันแบบเชิงเส้น
2. ฟังก์ชันแบบไม่เชิงเส้น

ตารางที่ 2.1 แสดงตัวอย่างของ Activation Function

ฟังก์ชัน	สมการ	อนุพันธ์ของสมการ	กราฟ
Sigmoid	$S(x) = \frac{1}{1 + e^{-x}}$	$S'(x) = S(x)(1 - S(x))$	
Tanh	$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$	$\tanh'(x) = 1 - (\tanh(x))^2$	
ReLU	$f(x) = \begin{cases} 0, & x \leq 0 \\ x, & x > 0 \end{cases}$	ไม่มี	
Softmax	$f_i(x) = \frac{e^{x_i}}{\sum_{j=0}^J e^{x_j}}$	ไม่มี	ไม่มี

2.1.1.4 Cost Function

Cost Function, Objective Function หรือ Loss Function เป็นฟังก์ชันทางคณิตศาสตร์ที่จะในการดูระยะห่างระหว่างเอาต์พุตปัจจุบันกับเอาต์พุตที่ตั้งเอาไว้ เป็นวิธีการประเมินว่าอัลกอริทึมที่ใช้ในโครงข่ายประสาทเทียมนั้นมีผลเป็นอย่างไร โดยจะแบ่งออกเป็นสองกลุ่มคือ

1. การทำ Classification สำหรับข้อมูลไม่ต่อเนื่อง (Discrete)

2. การทำ Regression สำหรับข้อมูลที่มีความต่อเนื่อง (Continuous)

This material is reserved for educational use only, not allowed for commercial use.

Categorical Cross Entropy Loss

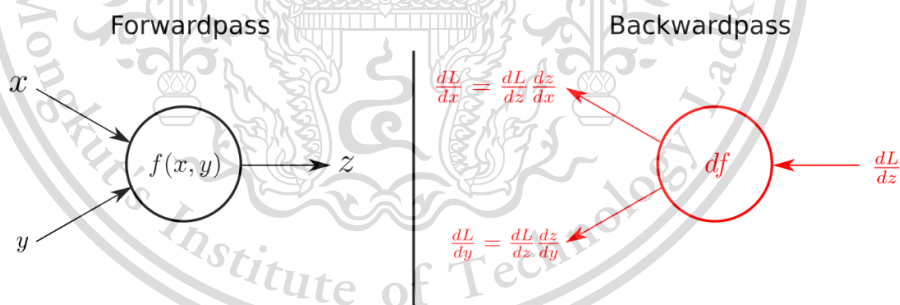
เป็น Loss Function ที่ใช้สำหรับการทำ Classification โดยจะใช้สำหรับข้อมูลที่มีมากกว่า 2 คลาสขึ้นไปซึ่งจะใช้งานร่วมกับ Activation Softmax โดยวางไว้ในเลเยอร์สุดท้ายของโครงข่ายประสาทเทียมเพื่อที่จะให้เอาต์พุตออกเป็นความน่าจะเป็นซึ่งจะได้ผลรวมของเอาต์พุตที่ได้ออกมาเท่ากับ 1

$$L = -\left(\frac{1}{m}\right) \sum_{i=1}^m \sum_{j=1}^k y_{i,j} \cdot \log (\hat{y}_{i,j})$$

สมการที่ (1) สมการของ Categorical Cross Entropy Loss

จากสมการที่ 1 จะให้ $y_{i,j}$ คือผลของเอาต์พุตตัวที่ i คลาสที่ j และ $\hat{y}_{i,j}$ คือผลของเอาต์พุตที่ได้มาจากการทำนายของโครงข่ายประสาทเทียมตัวที่ i คลาสที่ j โดยที่ $i = 1, 2, 3, \dots, m$ โดยที่ m คือจำนวนของเอาต์พุตและ $j = 1, 2, 3, \dots, k$ โดยที่ k คือจำนวนของคลาสทั้งหมด จากที่ Cross Entropy Loss จะได้ออกมาเป็นความน่าจะเป็นก็จะหมายความว่าถ้าตัวโครงข่ายประสาทเทียมนั้นทำการทายถูกแต่ไม่มั่นใจ loss ที่ได้นั้นก็จะมาก หรือ ถ้าทายผิดแต่ไม่มั่นใจมาก loss ก็จะมีค่ามากตามไปด้วยโดยการมั่นใจนั้นจะดูได้จากค่าของความน่ามาทำการคำนวณ

2.1.1.5 Forward Propagation

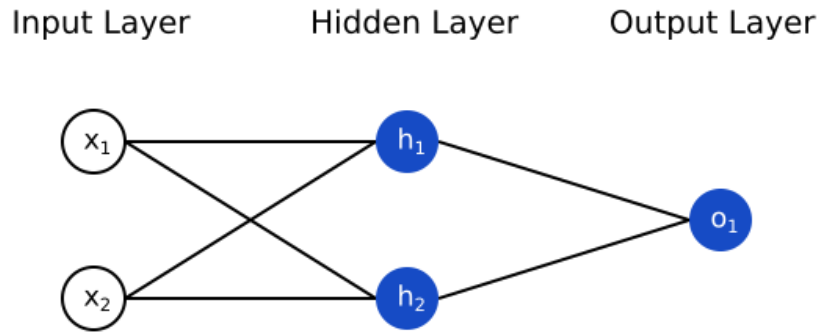


รูปภาพที่ 2.3 ภาพแสดง Forward และ Backward Propagation

อ้างอิง medium deep-learning-แบบฉบับคนสามัญชน: 2019 by Mr.P L

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.



รูปภาพที่ 2.4 ภาพแสดงโครงข่ายประสาทเทียมที่มี 1 hidden layer

เป็นการคำนวณและการจัดเก็บข้อมูลรวมถึงเอาต์พุต สำหรับโครงข่ายประสาทเทียมตามลำดับจากชั้นของอินพุตไปถึงชั้นของเอาต์พุต โดยจะทำงานทีละชั้นตอนผ่านกลไกของโครงข่ายประสาทเทียม โดยสามารถอธิบายได้ตามสมการที่ (2) โดยที่กำหนดให้ X เป็นอินพุตโดยที่ $X \in R^n$ และ $n \in I^+$ คือจำนวนของอินพุต และมีค่า bias คือ b และ W เป็นค่าน้ำหนักจะได้สมการในการทำ Forward Propagation ดังนี้

$$Z^{[l]} = W^{[l]T} X + b \quad (2)$$

โดยที่ $W \in R^{h \times n}$ โดยที่ $h, n, l \in I^+$ โดยจากรูปภาพที่ 2.3 เมื่อเราพิจารณาที่ hidden layer 1 จะได้ว่า h คือจำนวน Node ซึ่งจะมีค่าเท่ากับ 3 และ n คือจำนวนของอินพุตจะมีค่าเท่ากับ 2 และ l คือเลเยอร์ที่ l ซึ่งในที่นี้จะมีค่าเท่ากับ 1

$$A^{[l]} = \phi(Z^{[l]}) \quad (3)$$

จากสมการที่ (3) A คือผลลัพธ์หรือเอาต์พุตที่ได้ออกมาจาก hidden layer ที่ l และ ϕ เป็นสัญลักษณ์แทนการใช้ Activation Function และจะทำการส่งต่อไปที่เลเยอร์ถัดไปเรื่อยๆ โดยที่เอาต์พุตของเลเยอร์ก่อนหน้าจะเป็นอินพุตของเลเยอร์ถัดไปจนไปถึงเอาต์พุตเลเยอร์ และหลังจากนั้นจะทำการคำนวณ Cost Function หรือ Loss Function โดยจะทำการเทียบเอาต์พุตที่ได้จากเอาต์พุตเลเยอร์กับเอาต์พุตที่คาดหวังไว้หรือทำการตั้งเอาไว้

$$loss = L(A, y) \quad (4)$$

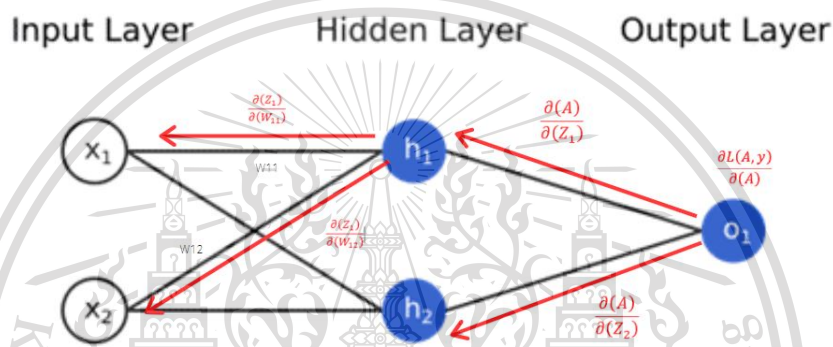
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับกิจกรรมเชิงงานเพื่อการศึกษานั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

จากสมการที่ (4) L จะเป็น Cost Function หรือ Loss Function โดยการเลือกใช้นั้นจะแตกต่างกันไปตามความเหมาะสม และ A คือผลลัพธ์หรือเอาต์พุตที่ได้ในเอาต์พุตเลเยอร์ y คือผลลัพธ์ที่เราต้องการหรือคาดหวัง

2.1.1.6 Backward Propagation

Backward propagation เป็นวิธีที่ถูกใช้ในการทำให้โครงข่ายประสาทเทียมเกิดการเรียนรู้ขึ้น โดยแนวคิดของ backward propagation นั้นจะใช้ผลของ Cost Function หรือ Error ที่ได้จากเอาต์พุตเลเยอร์มาเทียบกับ Weight และค่า Bias โดยจะทำย้อนกลับไปจากเอาต์พุตเลเยอร์จนถึงอินพุตเลเยอร์



รูปภาพที่ 2.5 ภาพแสดงการทำ Backward Propagation เบื้องต้น

ซึ่งการคำนวณ Backward propagation นั้นจะมีความซับซ้อนและการคำนวณ Cost หรือ Error ที่เกิดขึ้นนั้นเกิดขึ้นจากหลายตัวแปรดังนั้นจะใช้ Chain Rule และ Partial Differential เข้ามาใช้ในการคำนวณหา โดยสิ่งเราต้องการคือ $\frac{\partial L(A,y)}{\partial (W)}$ คือค่า Weight นั้นมีส่งผลกับ Error แค่นั้นซึ่งจะได้สมการเป็นดังสมการที่ (5)

$$\frac{\partial L(A,y)}{\partial (W_{11})} = \frac{\partial L(A,y)}{\partial (A)} \cdot \frac{\partial (A)}{\partial (Z_1)} \cdot \frac{\partial (Z_1)}{\partial (W_{11})}$$

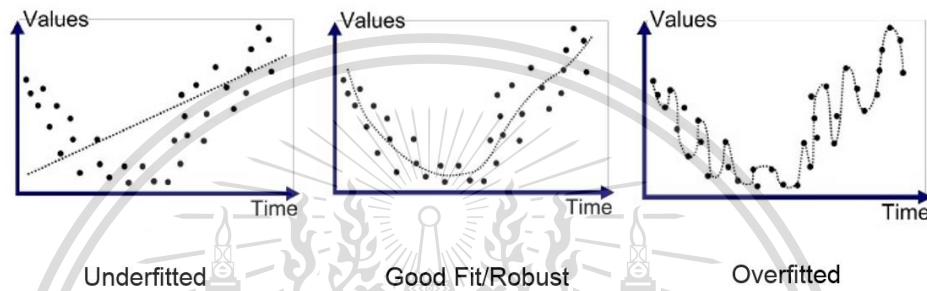
สมการที่ (5)

โดยสมการที่ 5 โดยที่ $L(A,y)$ คือ Cost Function เอาต์พุตที่ได้เทียบกับเอาต์พุตที่ควรจะเป็น A คือเอาต์พุตที่ได้จากโครงข่ายประสาทเทียม Z คือเอาต์พุตที่ได้จาก hidden layer ก่อนหน้า โดยสมการที่ (5) นั้นจะเป็นสมการตัวอย่างในการหาค่าของ Weight ที่ส่งผลกระทบต่อ Error ที่เกิดขึ้นของ h_1 เทียบกับเอาต์พุตที่ได้ในเอาต์พุตเลเยอร์โดยอ้างอิงจากรูปภาพที่ 2.2 ซึ่งการไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

คำนวณสำหรับ Weight อื่น ๆ ก็จะใช้วิธีเดียวกันโดยยิ่งถ้ามี hidden layer มากขึ้นการคำนวณก็ยิ่งซับซ้อนขึ้นไป

2.1.1.7 Under Fitting / Good Fitting / Over Fitting

ในการทำ Machine Learning แบบ Supervised Learning นั้นโดยปกติแล้วจะทำการสอนให้เจ้าตัว Model นั้นด้วยการทำการแบ่งชุดข้อมูลออกเป็น 2 อย่างคือ ชุดข้อมูลฝึกฝน กับ ชุดข้อมูลทดสอบ ซึ่งการชุดข้อมูลทดสอบนั้นจะเป็นชุดข้อมูลที่ต้องไม่มีอยู่ในชุดข้อมูลฝึกฝนหรือตัว Model นั้นจะต้องไม่เคยเห็นมาก่อนเพื่อใช้ในการวัดความแม่นยำของ Model



รูปภาพที่ 2.6 ภาพแสดงความแตกต่างของ Underfitting, Good Fit และ Overfitting

อ้างอิง Under Fitting / Over Fitting ปัญหาที่มองไม่เห็นแต่สัมผัสได้ว่ามี..... : Machine Learning 101 2018 by Mr.P L

1. Under Fitting คือการที่ Model นั้นเมื่อนำไปทดสอบกับชุดข้อมูลทดสอบหรือนำไปใช้งานจริงแล้ว การทำนายหรือทำการแบ่งกลุ่ม ผลลัพธ์ไม่ถูกต้องหรือไม่ใกล้เคียงกับผลลัพธ์ที่ต้องเป็น ซึ่งสามารถเกิดขึ้นได้ในกรณีที่ใช้ชุดข้อมูลฝึกฝนน้อยเกินไปจนทำให้ Model ไม่เกิดการเรียนรู้
2. Good Fit/Robust คือการที่ Model นั้นเมื่อนำไปทดสอบกับชุดข้อมูลทดสอบหรือนำไปใช้งานจริงแล้ว สามารถการทำนายได้ถูกต้องและมีความแม่นยำได้ระดับที่ยอมรับได้
3. Over Fitting คือการที่ Model เมื่อนำไปทดสอบกับชุดข้อมูลทดสอบหรือนำไปใช้งานจริงแล้วไม่สามารถที่จะการทำนายหรือทำการแบ่งกลุ่มของข้อมูลได้ถูกต้อง แต่ในการทดสอบกับชุดข้อมูลฝึกฝนนั้นกลับการทำนายได้อย่างถูกต้องแม่นยำ

2.1.1.8 Optimization

การทำ optimization เป็นวิธีที่ใช้หลัก ๆ ในการปรับค่า Weight และ Bias โดยจะทำการหาว่าจะต้องปรับเท่าไรแล้วให้เกิดค่า Loss หรือ Error ที่เกิดจาก Cost Function ได้น้อยที่สุด

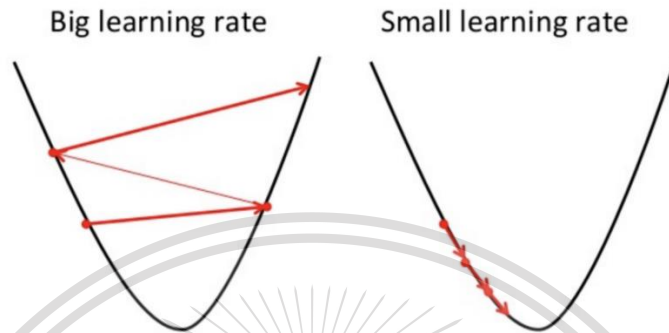
1. Gradient descent เป็นอัลกอริทึมที่ถูกเริ่มใช้ในการปรับค่าพารามิเตอร์ต่าง ๆ โดยจะมี

สมการเป็น เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลง
$$W = W - \alpha \frac{\partial L(A,y)}{\partial (w)} \quad (6)$$
 เอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

W คือ ค่า Weight α คือ Learning rate และ $L(A, y)$ คือ Loss Function ของการเปรียบเทียบเอาต์พุตของโครงข่ายประสาทเทียมกับเอาต์พุตที่ควรจะเป็นโดยจากสมการที่ 6 นั้นจะเป็นการใช้ Gradient descent เข้ามาทำการปรับค่า Weight ที่จะทำให้เกิดค่า Loss น้อยที่สุด โดยที่จะทำการปรับตาม Learning rate (α) ที่ตั้งเอาไว้เพราะฉะนั้น Learning rate จึงเป็นสิ่งสำคัญ เพราะถ้าเกิดใช้มากไปก็จะทำให้เลยจุดต่ำสุด แต่ถ้าปรับน้อยไปก็จะทำให้ถึงจุดต่ำสุดได้ช้าลง



รูปภาพที่ 2.7 ภาพแสดงความแตกต่างของการตั้งค่า Learning Rate ระหว่าง Big และ Small อ้างอิง Gradient Descent คีย์เวิร์ดที่ Data Science ต้องรู้จัก ! : Machine Learning 101: Machine Learning 101 2018 by Mr.P L

2. Adam หรือ Adaptive Moment Estimation เป็น optimizer ที่สามารถทำการปรับค่าการเรียนรู้ สำหรับพารามิเตอร์ในแต่ละครั้งได้ Adam เป็น optimizer ที่นิยมมากที่สุด เพราะสามารถแก้ปัญหาต่าง ๆ ที่เกิดขึ้นได้และยังให้ความเร็วมากกว่า Gradient descent และลดปัญหาการแกว่งของพารามิเตอร์ได้อีกด้วย.

2.1.1.9 การวัดผลของโมเดลในการทำ Classification

1. Accuracy เป็น Matric ที่นิยมใช้งานกันมากที่สุด โดยจะใช้ในการบอกว่าโมเดลนั้นทำการทำนายถูกทั้งหมดกี่เปอร์เซ็นต์ โดยจะใช้สูตรตามสมการที่ (7)

$$acc = \frac{True\ Positive}{True\ Positive + True\ Negative + False\ Positive + False\ Negative} \quad (7)$$

2. Precision เป็น Matric ที่ใช้วัดว่าโมเดลนั้นมีความแม่นยำมากแค่ไหนโดยจะดูจากการทำนายที่ถูกเทียบกับจำนวนข้อมูลของคลาสนั้น โดยจะได้ตามสมการที่ (8)

$$precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \quad (8)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

3. Recall เป็น Matric ที่ใช้วัดว่าโมเดลนั้นสามารถทำนายคลาสนั้นได้ถูกต้องเท่าไรเมื่อเทียบกับการทำนายที่ได้ออกเป็นคลาสนั้นทั้งหมด โดยจะสามารถเขียนเป็นสมการได้ตามสมการที่ (9)

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negative} \quad (9)$$

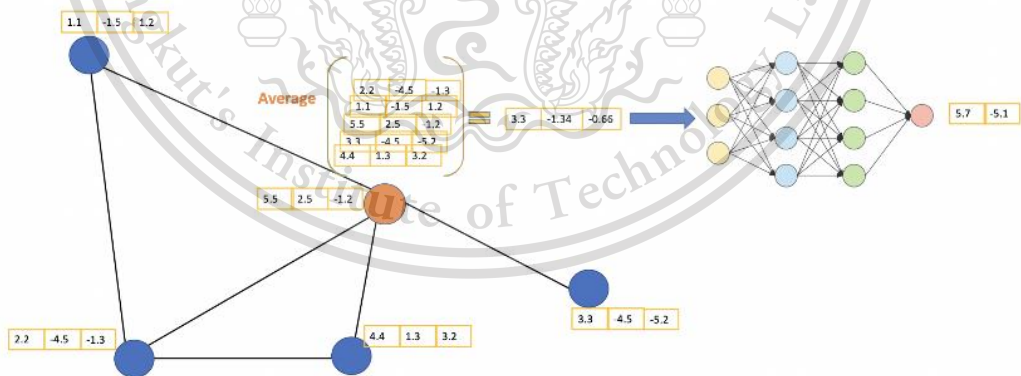
4. F1-Score เป็น Matric ที่จะได้ผลออกมาเป็น harmonic mean ระหว่าง precision และ recall โดยจะเหมาะสมในการใช้วัดโมเดลที่มีข้อมูลในรูปแบบของ Imbalance Classes โดยจะได้ตามสมการที่ (10)

$$F1 - Score = 2 * \left(\frac{precision * recall}{precision + recall} \right) \quad (10)$$

2.1.2 Graph Convolutional Network

Graph Convolutional Network หรือ GCN เป็นสถาปัตยกรรมของโครงข่ายประสาทเทียมที่มีประสิทธิภาพสำหรับการเรียนรู้ในรูปแบบของกราฟ โดยจะประกอบด้วย node, adjacent node และ edge ซึ่งจะใช้การคำนวณในรูปแบบของ Matrix

GCN จะใช้การทำ “Convolution” เข้ามาใช้ในการเรียนรู้คุณลักษณะของ node และ adjacent node เพื่อใช้เป็นข้อมูลส่งเข้าไปในโครงข่ายประสาทเทียม



รูปภาพที่ 2.8 ภาพแสดงตัวอย่างคอนเซ็ปของ Graph Convolutional Network

อ้างอิง Graph Convolutional Networks (GCN) 2020 by Chau Pham

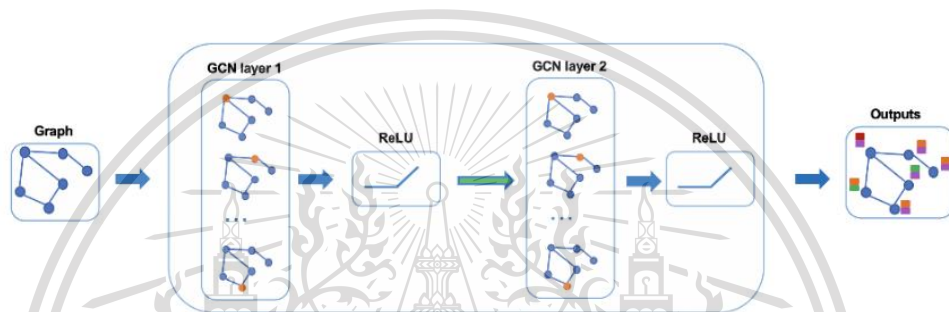
โดย GCN คือโครงข่ายประสาทเทียมที่ทำในรูปแบบของกราฟ (G) โดยที่ $G = (V, E)$ (11) เมื่อ V คือโหนดของข้อมูลโดยที่ $V = \{v_1, v_2, v_3, \dots, v_n\}$ และ E คือเส้นเชื่อมการคำนวณ ไม่ว่ากรณีใดการต่อกันของแต่ละโหนดและอินพุตในการส่งเข้าไปในโครงข่ายประสาทเทียมจะประกอบไปด้วย

This material is reserved for educational use only, not allowed for commercial use.

Feature Matrix และ Adjacent Matrix โดยที่ X คือ Feature Matrix ซึ่งจะมีขนาดเท่ากับ $N \times X^0$ โดยที่ N คือจำนวนของโหนดทั้งหมดและ X^0 คือจำนวนของคุณสมบัติในแต่ละโหนด และ A คือ Adjacent Matrix โดยที่จะมีขนาดเท่ากับ $N \times N$

ใน GCN สามารถเขียนการทำ propagation ได้รูปแบบของสมการที่ (9) โดยที่ W^i คือ weight matrix ของ เลเยอร์ ที่ i และ σ คือ non-linear activation function เช่น ReLU Function โดยที่ weight matrix จะมีขนาดเท่ากับ $X^i \times X^{i+1}$ และ $H^0 = X$

$$f(H^i, A) = \sigma(AH^iW^i) \quad (12)$$



รูปภาพที่ 2.9 ภาพแสดงตัวอย่าง GCN 2-layer

อ้างอิง Graph Convolutional Networks (GCN) 2020 by Chau Pham

2.1.3 Computer Vision

Computer Vision หรือ คอมพิวเตอร์วิทัศน์เป็นศาสตร์ที่จะทำให้คอมพิวเตอร์สามารถเข้ารูปภาพหรือวิดีโอได้ในรูปแบบเดียวกันกับที่มนุษย์มองเห็น โดยจะทำการประมวลผลภาพด้วย machine learning ในตัวอย่างการใช้งานของคอมพิวเตอร์วิทัศน์ที่เห็นได้อย่างกว้างขวางเช่น

- การตรวจจับ หรือการระบุตำแหน่ง และรู้จำของวัตถุที่ต้องการในภาพเช่น หน้าคน หรือ ตัวหนังสือ
- การติดตาม วัตถุที่ต้องการในภาพ

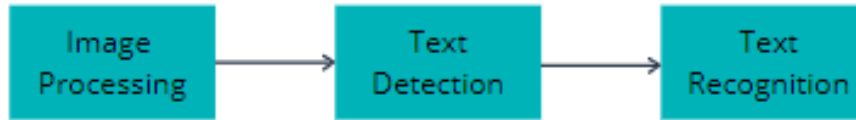
2.1.3.1 Optical Character Recognition (OCR)

OCR เป็นเครื่องมือที่ใช้ในการทำให้รูปภาพหรือเอกสารที่สแกนแปลงให้เป็นในรูปแบบข้อความ ที่ถูกใช้อย่างแพร่หลายในรูปแบบของการป้อนข้อมูลจากกระดาษเอกสารต่าง ๆ ไม่ว่าจะเป็น หนังสือเดินทาง, ใบเสร็จ หรือสลิปเงินเดือน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

OCR เป็นหนึ่งในการทำ Computer Vision ในด้านการจดจำรูปแบบและการตรวจจับโดยที่ในการทำงานของการแปลงรูปภาพมาเป็นข้อความนั้นจะทำการตรวจจับตำแหน่งของข้อความในรูปภาพ (Text Detection) ก่อนและจึงทำการจดจำข้อความ (Text Recognition)



รูปภาพที่ 2.10 ภาพแสดงขั้นตอนการทำ OCR

แต่การใช้งานนั้นจะต้องอาศัยรูปภาพเอกสารต่าง ๆ ที่มีความคมชัดทั้งภาพและเนื้อหาเพื่อที่จะทำให้การแปลงออกมาเป็นข้อความนั้นทำได้อย่างมีประสิทธิภาพมากที่สุด

2.1.4 Image Processing

Image Processing หรือ การประมวลผลภาพเป็นวิธีการดำเนินการบางอย่างกับรูปภาพเพื่อทำการปรับปรุงรูปภาพให้เป็นไปอย่างที่ต้องการหรือทำการดึงข้อมูลที่เป็นประโยชน์ออกมา ซึ่งเป็นวิธีการประมวลผลสัญญาณประเภทหนึ่งซึ่งอินพุตเป็นรูปภาพและเอาต์พุตอาจจะเป็นรูปภาพหรือลักษณะ / คุณสมบัติที่เกี่ยวข้องกับรูปภาพนั้น

2.1.4.1 Sauvola Local Thresholding

Thresholding คือการปรับหรือแบ่งส่วนภาพประเภทหนึ่ง (Segmentation) โดยทำการเปลี่ยนพิกเซลของภาพซึ่งจะทำแปลงภาพจากสีเทาหรือ Gray Scale ให้เป็นไบนารี (Binary Image) ซึ่งจะทำการปรับตามเกณฑ์ความเข้มของพิกเซลตามที่กำหนดโดยจะสามารถอธิบายเป็นสมการได้ดังนี้อธิบายเป็นสมการได้ดังนี้ กำหนดให้เกณฑ์ที่ตั้งคือ $T(x, y)$ และ $B(x, y)$ คือ binarized image และ $I(x, y)$ คือความเข้มของพิกเซลที่ตำแหน่ง x, y ของรูปภาพโดยที่ $I(x, y) \in [0, 1]$ โดยจะได้สมการในการปรับ Threshold ดังนี้

$$B(x, y) = \begin{cases} 0, & I(x, y) < T(x, y) \\ 1, & \text{otherwise} \end{cases} \quad (13)$$

ในการปรับ threshold แบบ local adaptive นั้นจะคำนวณในแต่ละพิกเซลโดยจะดูจากสถิติเช่น ช่วงของความแปรปรวน หรือค่าพิกเซลของพื้นที่ใกล้เคียง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

Sauvola Local Threshold จะใช้เทคนิค local variance โดยวิธีนี้จะใช้เกณฑ์จากการคำนวณค่าเฉลี่ย (local mean) $m(x, y)$ และค่าส่วนเบี่ยงเบนมาตรฐาน $\delta(x, y)$ ภายในหน้าต่างขนาด $w \times w$ โดยจะได้สมการได้การตั้งเกณฑ์ดังนี้

$$T(x, y) = m(x, y) \left[1 + k \left(\frac{\delta(x, y)}{R} - 1 \right) \right] \quad (14)$$

โดยที่ R คือค่าส่วนเบี่ยงเบนมาตรฐานที่มากที่สุด ($R = 128$ สำหรับรูปภาพแบบ gray scale) และ k คือค่า bias ซึ่งจะมีค่าอยู่ในช่วงของ $[0.2, 0.5]$ ค่าเฉลี่ยหรือ local mean และค่าส่วนเบี่ยงเบนมาตรฐานจะปรับค่าเกณฑ์ตามความคมชัด (Contrast) ในพื้นที่ใกล้เคียงของพิกเซล โดยที่เมื่อค่าความคมชัดสูงการเปรียบเทียบในบางจุดบางพื้นที่ของภาพจะมีค่า local sum หรือเทียบเท่ากับ R โดจะได้ผลลัพธ์เป็น $T(x, y)$ ซึ่งจะเทียบเท่ากับ $m(x, y)$ แต่ถ้าเมื่อมีความคมชัดในพื้นที่ใกล้เคียงต่ำจะเป็นในกรณีของการที่ค่าของ $T(x, y)$ น้อยกว่าค่า $m(x, y)$ ดังนั้นจึงสามารถทำการลบพื้นที่ที่มีความคมชัดต่ำหรือค่อนข้างมืดได้ดี และค่า k จะคอยควบคุมค่าเกณฑ์ในพื้นที่ท้องถิ่น (Local window) โดยที่ถ้าค่า k ยิ่งสูงจะยิ่งทำให้ค่าเกณฑ์หรือ $T(x, y)$ ต่ำ โดยที่ใน Sauvola จะใช้ค่า k อยู่ที่ 0.5

อย่างไรก็ตามลำดับในการคำนวณค่าเกณฑ์หรือ $T(x, y)$ ค่าของค่าเฉลี่ย และค่าส่วนเบี่ยงเบนมาตรฐานจะต้องคำนวณสำหรับแต่ละพิกเซล ซึ่งการคำนวณของ $m(x, y)$ และ $\delta(x, y)$ เป็นการคำนวณที่ไม่มีประสิทธิภาพซึ่งจะได้ผลของ complexity ในการคำนวณในสำหรับรูปภาพหนึ่งจะเท่ากับ $O(n^2w^2)$ เมื่อรูปภาพมีขนาดเท่ากับ $n \times n$

2.1.4.2 Morphology

เป็นการดำเนินการแบบเรียบง่ายโดยปกติจะดำเนินการกับภาพรูปแบบไบนารี โดยการดำเนินการจะประกอบไปด้วยรูปภาพที่เป็นอินพุตและเคอร์เนลซึ่งเป็นตัวกำหนดลักษณะของการดำเนินการ โดยการดำเนินการนั้นจะเริ่มจากพื้นฐานสองอย่างคือ การทำ Erosion และ การทำ Dilation



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาค้นคว้าเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
รูปภาพที่ 2.11 รูปภาพก่อนที่จะถูกดำเนินการ
 ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

1. Erosion

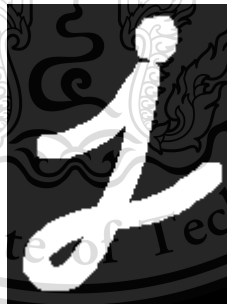
จะเป็นวิธีการตรวจสอบและลดขนาดรูปร่างที่มีอยู่ในรูปภาพอินพุตจะทำหน้าที่คล้ายกับเป็น local minimum filter โดยจะทำการเลื่อนเคอร์เนลเลื่อนผ่านรูปภาพโดยที่ถ้าเคอร์เนลเลื่อนผ่านจะมีค่าเป็นเท่ากับ 1 เมื่อพิกเซลทั้งหมดในเคอร์เนลเป็น 1 ถ้าไม่จะมีค่าเป็น 0 เพราะฉะนั้นสิ่งที่เกิดขึ้นคือพิกเซลที่อยู่ใกล้กับขอบจะถูกทำให้เป็น 0 ดังนั้นความหนาหรือขนาดของวัตถุในรูปจะมีขนาดเล็กกลงโดยจะขึ้นอยู่กับขนาดของเคอร์เนล



รูปภาพที่ 2.12 รูปภาพหลังทำการ Erosion ด้วยเคอร์เนลขนาด 5x5

2. Dilation

จะเป็นทำสิ่งที่ตรงกันข้ามกับการทำ Erosion โดยที่ในการเลื่อนผ่านของเคอร์เนลถ้ามีอย่างน้อยหนึ่งพิกเซลภายในเคอร์เนลมีค่าเท่ากับ 1 จะทำการเพิ่มพื้นที่สีขาวหรือเปลี่ยนจากค่า 0 เป็น 1 ดังนั้นจึงจะทำให้ขนาดของวัตถุในรูปมีขนาดใหญ่ขึ้นตามขนาดของเคอร์เนล ซึ่งจะมีประโยชน์ในการเชื่อมชิ้นส่วนระหว่างวัตถุในรูปภาพ



รูปภาพที่ 2.12 รูปภาพหลังทำการ Dilation ด้วยเคอร์เนลขนาด 5x5

3. Opening

Opening ก็เป็นแค่ชื่อเรียกของการทำ erosion และตามด้วยการทำ Dilation โดยจะช่วยในการลบ Noise ในรูปภาพดังรูปตัวอย่างที่ 2.13

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.



รูปภาพที่ 2.12 รูปภาพหลังแสดงการทำ opening ด้วยคอร์เนลขนาด 5x5

2.1.5 Natural Language Processing

Natural Language Processing หรือ NLP เป็นการประมวลผลภาษาธรรมชาติ หรือก็คือภาษาของมนุษย์โดยที่ NLP นั้นก็เป็นสาขาหนึ่งของ Artificial intelligence โดยเป้าหมายหลักของ NLP นั้นคือการทำให้ทำให้คอมพิวเตอร์สามารถที่จะ อ่านและเข้าใจภาษาของมนุษย์ได้ โดยมีตัวอย่างการนำไปใช้งานที่เห็นได้อย่างกว้างเช่น การทำ Google Translate หรือ Siri เป็นต้น.

โดยจะขั้นตอนในการเตรียมพร้อมในการทำ NLP หลัก ๆ คือ

1. Tokenization จะเป็นการทำให้ประโยคแยกออกเป็นคำหรือโทเคนโดย เช่น ไปกินข้าว จะถูกการทำ Tokenization เป็น 'ไป', 'กิน', 'ข้าว'
2. Stemming and Lemmatization จะเป็นการเปลี่ยนแปลงรูปของคำให้ไปเป็นรูปแบบของรากศัพท์นั้น ๆ โดยจะพบเห็นได้บ่อยได้การทำ NLP กับภาษาอังกฤษ เช่น seen จะถูกเปลี่ยนเป็น see

2.1.5.1 Cosine Distance

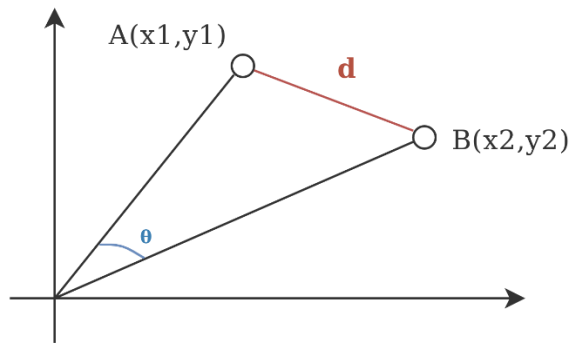
Cosine Distance เป็นหนึ่งในวิธีการที่นิยมใช้ในการทำ NLP ถูกนำมาใช้เพื่อวัดความใกล้เคียงกันของคำสองคำ ซึ่งที่จะมีความสัมพันธ์กับ Cosine Similarity หรือก็คือความเหมือนกันโดยที่จะสามารถหาได้จากสมการที่ด้านล่าง

$$\text{similarity} = \cos(\theta) = \frac{A \cdot B}{\|A\| \cdot \|B\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}}$$

โดยที่ถ้าสองเวกเตอร์นั้นมีความเหมือนกันมากเท่าไรค่าของ cosine similarity จะมีค่าเข้าใกล้ 0 โดยที่ถ้าเหมือนกันเลยจะมีค่าเท่ากับ 0 แต่ถ้าต่างกันค่าจะเข้าใกล้ 1 โดยที่ถ้าแตกต่างกันโดยที่ไม่มีอะไรเหมือนกันเลยจะมีค่าเท่ากับ 1

เอกสารนี้เป็นเอกสารที่สงวนไว้ใช้สำหรับอาจารย์และบุคลากรในมหาวิทยาลัยเท่านั้น ไม่สามารถนำออกนอกมหาวิทยาลัยได้
โดยค่าของ cosine distance นั้นจะมีค่าเท่ากับ $1 - \cos(\theta)$ โดยจะมีค่าตรงกันไม่ว่ากรณีใดซ้ำกับ cosine similarity

This material is reserved for educational use only, not allowed for commercial use.



รูปภาพที่ 2.13 รูปภาพตัวอย่าง cosine distance

2.1.6 RESTful API

RESTful คือ Web Services ที่จะอยู่บนพื้นฐานของ REST ซึ่งย่อมาจากคำว่า Representational State Transfer ซึ่งเป็น Architectural Style ซึ่งก็เป็นเพียงสถาปัตยกรรมรวมกับการกำหนดข้อจำกัดซึ่งการที่จะเป็น REST นั้นจะต้องมีข้อจำกัดต่อไปนี้

1. Performance คือประสิทธิภาพในการโต้ตอบกลับ
2. Scalability คือสามารถทำการปรับขยายเพื่อรองรับส่วนประกอบต่าง ๆ จำนวนมากได้
3. Simplicity คือต้องมีความเรียบง่ายในส่วนของการเชื่อมต่อและมีความเหมือนกัน
4. Modifiability คือสามารถทำการปรับเปลี่ยนส่วนประกอบต่าง ๆ เพื่อให้ตอบสนองกับการใช้งาน
5. Visibility คือสามารถมองเห็นการเชื่อมต่อหรือการสื่อสารระหว่างส่วนประกอบได้
6. Portability คือสามารถทำการปรับเปลี่ยนและโยกย้ายในส่วนของข้อมูลกับโปรแกรมได้ง่าย
7. Reliability คือมีระบบมีความน่าเชื่อถือระบบไม่ล่มง่าย

ซึ่ง RESTful นั้นจะมีระบบที่สามารถเข้าถึงทรัพยากรหรือการ Request ผ่านเว็บและจะทำการตอบโต้กับโดยจะใช้ HTTP Protocol หรือเรียกว่า Response ซึ่ง RESTful นั้นจะประกอบไปด้วย 6 ข้อกำหนดดังนี้

1. Client-Server Architecture คือ การแยกกันระหว่าง Client กับ Server โดยที่ Client จะไม่มีส่วนและหน้าที่โดยตรงกับฝั่งของ Server และ Server นั้นมีหน้าที่แค่บริการทรัพยากรให้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

2. Statelessness คือ การติดต่อระหว่าง Client และ Server นั้นทางฝั่งของ Client จะทำหน้าที่แค่ส่ง Request ไปและ Server จะทำหน้าที่แค่ Response กลับมาให้ก็จะถือว่าสิ้นสุดการติดต่อกัน
3. Cache คือ สามารถที่จะ Request และได้รับ Response กลับมาแม้จะไม่ได้ทำการ Request ไปที่ Server
4. Layered System คือ ในระบบ Layered นั้นจะสามารถใช้ตัวกลางหรือ Proxy ระหว่าง Client กับ Server ได้ซึ่งจะสามารถช่วยในการทำ Load Balance ได้
5. Uniform Interface คือการแสดงให้เห็นถึงวิธีที่จะติดต่อกับฝั่งของ Server โดยที่จะต้อง มีหน้าตาการติดต่อสื่อสารเหมือนกันแม้จะเชื่อมต่อกันคนละอุปกรณ์
 - 5.1 Identification of Resources เป็นการกำหนดชื่อให้กับข้อมูลที่จะถูกส่งออกมาเช่น รูปภาพหรือเอกสาร
 - 5.2 Manipulation of Resources through representations คือรูปแบบของการตอบกลับเช่น JSON, XML
 - 5.3 Self-description message คือข้อความที่จะทำการส่ง Request และ Response กลับนั้นจะต้องถูกอธิบายได้อย่างเข้าใจ เช่น HTTP Status
 - 5.4 Hypermedia as the engine of application state (HATEOAS) การที่จะเข้าถึงทรัพยากรนั้นจะเข้าถึงได้ด้วย URL
6. Code on demand (Optional) เป็นทางเลือกและอนุญาตให้ Client ดาวน์โหลด โปรแกรมสำหรับการดำเนินการฝั่ง Client เช่น JavaScript, jQuery

2.1.7 HTTP Protocol

HTTP หรือ Hypertext Transfer Protocol คือโพรโทคอลที่ใช้ในการติดต่อสื่อสารระหว่างฝั่ง Client กับฝั่งของ Server ใน Application layer ของ TCP/IP model ผ่านทาง URL หรือ PATH โดยการที่ Client นั้นจะติดต่อไปยัง Server จะเรียกว่า Request และทาง Server จะทำการตอบกลับหรือที่เรียกว่า Response กลับไปยังทางฝั่งของ Client

2.1.7.1 HTTP Request

ในการ Request ไปในแต่ละครั้งนั้นจะต้องประกอบไปด้วยส่วนต่าง ๆ ดังนี้

1. Method เป็นส่วนที่จะต้องกำหนดไว้บอกทางฝั่งของ Server รู้ว่าทาง Client นั้นต้องการจะดำเนินการอะไร โดยจะประกอบไปด้วย method หลัก ๆ ดังนี้
 - ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

- GET เป็น method ที่ใช้เมื่อต้องการเรียกรับข้อมูลจากทางฝั่งของ Server โดยผ่านทาง URL ที่กำหนด

- POST เป็น method ที่ใช้ในกรณีที่ต้องการ Request การสร้างข้อมูลใหม่เข้าไปในทางฝั่งของ Server โดยจะทำการส่ง

- PUT เป็น method ที่ใช้ในการอัปเดตข้อมูลที่มีอยู่แล้วโดยจะแทนที่ด้วยข้อมูลใหม่ทั้งข้อมูล

- PATCH เป็น method ที่คล้ายกับ PUT method แต่จะทำการแก้ไขข้อมูลเฉพาะส่วนที่ถูกส่งไปเท่านั้น

- DELETE เป็น method ที่ใช้ในการลบข้อมูลที่มีอยู่แล้ว

2. Headers จะเป็นส่วนที่ผู้ใช้จะกำหนดข้อมูลเพิ่มเติมที่เกี่ยวกับ Request เช่น ข้อมูลที่จะส่งไปนั้นเป็นชนิดไหน โดยจะกำหนดในส่วนของ Content-Type หรือส่วนของการ Authorization

3. URL หรือ PATH เป็นส่วนที่กำหนดปลายทางหรือ end point ที่ต้องการ Request ไป

4. Version เป็นส่วนที่ใช้กำหนด version ของ http ที่ทางฝั่งของ Client นั้นใช้งาน

5. Body จะเป็นส่วนที่ใส่ข้อมูลสำหรับการส่งไปให้กับฝั่งของ Server หรือจะถูกเรียกว่า payload ซึ่งจะถูกใช้ในบาง method เช่น POST, PUT, PATCH ซึ่งอาจจะใส่หรือไม่ใส่ก็ได้แล้วแต่การกำหนดใน Request นั้น ๆ

2.1.7.2 HTTP Response

การตอบกลับหรือ Response จากทาง Server จะประกอบไปด้วย 3 ส่วนหลักได้แก่

1. Status Code & Status message โดยจะถูกแบ่งชนิดตามเลขที่อยู่ตัวหน้าสุด โดยแบ่งได้ดังนี้

- 1XX Informational Request จะเป็นตัวบอกสถานะว่า Server กำลังประมวลผลข้อมูล และจะแจ้งเตือนให้ผู้ใช้งานรอการ Response

- 100 (Continue) Server ได้รับส่วนของ Header ของการ Request

แล้ว Client สามารถส่งส่วน message body ตามมาได้

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

- 101 (Switching Protocol) ผู้ส่ง Request ให้ฝั่ง Server เปลี่ยน protocol แล้ว Server ทำการตกลง

- 102 (Processing) เมื่อ Server ได้รับ Request ที่มีการเกี่ยวข้องกับไฟล์ เป็นจำนวนมากและใช้เวลานาน การตอบกลับด้วย status code นี้จะเป็นการบอกว่า Server ในได้รับ แล้วและกำลังประมวลผลอยู่

- 2XX Success จะเป็นรหัสที่ทำการบอกว่าการ Request นั้นได้รับแล้ว และทาง ฝั่งของ Server ได้ทำการดำเนินตามคำขอนั้นเสร็จเรียบร้อยแล้ว

- 200 (OK) เป็นมาตรฐานของ HTTP Response โดยจะบ่งบอกว่าการ ติดต่อระหว่าง Client กับ Server นั้นสำเร็จและเป็นไปตามข้อตกลงและไม่มีการเกิดข้อผิดพลาด ไต ๆ เกิดขึ้น

- 201 (Created) เป็นการบอกว่าได้มีการสร้างข้อมูลใหม่เกิดขึ้นและทำ สำเร็จแล้วโดยจะใช้สำหรับ POST Method

- 204 (No Content) เป็นการบ่งบอกว่า Request ที่ได้รับนั้นดำเนินการ สำเร็จแล้วแต่ไม่มีการ Response Message กลับไปให้ทางฝั่งของ Client

- 3XX Redirection เป็นการระบุว่าทางฝั่งของ Client นั้นจำเป็นที่จะต้องทำการ เปลี่ยน PATH หรือ URL เพื่อที่จะได้ดำเนินการต่อตาม Request

- 300 (Multiple Choices) มีการระบุ PATH หรือ URL ให้ทางฝั่งของ Client เลือก

- 301 (Moved Permanently) Request ในการเข้าถึงทรัพยากรเดิมนั้น ได้ถูกทำการเปลี่ยนไปยังที่ทรัพยากรใหม่อย่างถาวรแล้ว

- 302 (Found) เหมือนกับ status code 301 แต่จะเป็นการเปลี่ยนแค่ ชั่วคราวเท่านั้น

- 4XX Client Errors จะเป็นการระบุข้อผิดพลาดที่เกิดขึ้นที่ฝั่งของ Client

- 400 (Bad Request) จะเป็นการบอกว่าทาง Server ไม่สามารถที่จะ ประมวลผลตาม Request ที่ทางฝั่งของ Client นั้นส่งมาได้เนื่องจาก Client อาจจะทำการส่งผิด รูปแบบ

- 401 (Unauthorized) เป็นการบ่งบอกว่าการ Request ที่จะทำการ เข้าถึงนั้นจะ ต้องได้มีการอนุญาต หรือมีสิทธิในการเข้าถึงก่อน

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้เพื่อใช้ในการเรียนการสอนเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

- 404 (Not Found) เป็นการบ่งบอกว่าการ Request นั้นไม่เจอสิ่งที่ทาง Client ร้องขอมาและทางฝั่ง Server ไม่สามารถส่งข้อมูลใด ๆ คืนได้

- 5XX Server Errors เป็นการแสดงข้อผิดพลาดที่เกิดขึ้นจากทางฝั่งของ Server

- 501 (Internal Server Error) เป็นการบอกว่า Request นั้นถูกต้องแต่ทางฝั่งของ Server นั้นไม่สามารถที่จะทำการตอบกลับได้

- 502 (Bad Gateway) เป็นการบอกว่าทางฝั่งของ Server ที่ทำหน้าที่เป็น Gateway นั้นไม่สามารถทำการติดต่อไปยังปลายทางที่ทางฝั่งของ Client ร้องขอได้

2. Version เหมือนกับในการทำ HTTP Request จะเป็นการบอกเวอร์ชันของ HTTP ที่ใช้ในทางฝั่งของ Server

3. Headers จะเป็นส่วนที่มีข้อมูลเพิ่มเติมจาก response body เช่น content-length และ content-type ที่ทำการตอบกลับไปยังฝั่งของ Client

4. Body เป็นส่วนที่สำคัญของ HTTP Response โดยที่ในส่วนของ Body นั้นจะมีข้อมูลที่ทางฝั่งของ Client ต้องการจากการ Request มา

2.2 ภาษาและเฟรมเวิร์กที่ใช้ในการพัฒนาในโครงการ

2.2.1 ภาษา

2.2.1.1 Python

Python เป็นภาษาโปรแกรมระดับสูงหรือ High Level ภาษาไพทอนนั้นถูกออกแบบเพื่อที่จะให้ผู้ใช้สามารถอ่านและเข้าใจชุดคำสั่งได้ง่าย โดยจะตัดความซับซ้อนของไวยากรณ์ (Syntax) ของตัวภาษาออกไปโดยจะใช้เวลาว่างจำนวนมากแทน ไพทอนยังมีการประยุกต์ใช้แนวคิดการเขียนโปรแกรมเชิงวัตถุหรือ Object Oriented Programming (OOP) ซึ่งสามารถเขียนและทำการบำรุงรักษาตัวโปรแกรมได้อย่างง่ายอีกด้วย

การแปลงชุดคำสั่งของภาษาไพทอนนั้นจะทำการแปลงให้เป็นภาษาเครื่องแบบ Interpreter หรือก็คือการที่จะทำการแปลงชุดคำสั่งในไพทอนให้เป็นภาษาเครื่องที่ละบรรทัดเพื่อป้อนเข้าสู่หน่วยประมวลผลให้คอมพิวเตอร์ทำงานอย่างที่ต้องการ ภาษาไพทอนนั้นถูกออกแบบให้ใช้งานสำหรับงานทั่วไป (General-Purpose Language) ซึ่งสามารถที่จะนำไปใช้งานได้หลากหลายประเภทโดยไม่จำกัดอยู่เฉพาะงานใดงานหนึ่ง จึงทำให้เป็นที่นิยมใช้กันอย่างแพร่หลายในหลายองค์กรเช่น Google, Youtube เป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.



รูปภาพที่ 2.14 รูปโลโก้ของภาษาไพทอน

```
print("Hello World")
for index in range(10):
    if index%2 == 0:
        print("Even")
    else:
        print("Odd")
```

รูปภาพที่ 2.15 ตัวอย่างไวยากรณ์ของภาษาไพทอน

2.2.1.2 JavaScript

JavaScript เป็นภาษาโปรแกรมมิ่งส่วนใหญ่ใช้สำหรับการพัฒนาเว็บไซต์ โดยที่จะช่วยให้สามารถจัดการคุณสมบัติที่ซับซ้อนบนหน้าเว็บแอปพลิเคชัน เพื่อให้เว็บไซต์มีความสวยงามและมีการเคลื่อนไหว สามารถส่งค่าไปมาระหว่าง Object ได้และมีการตอบสนองกับผู้ใช้งานได้ทันที โดยภาษา JavaScript นั้นจะถูกใช้ในส่วนของผู้ Client

```
export const checkIntersec = (regionProps, listRegion) => {
  const x1 = regionProps.x
  const y1 = regionProps.y
  const w1 = regionProps.width
  const h1 = regionProps.height

  let arr = []
  listRegion.map((val, key) => {
    const x2 = parseFloat(val[0])
    const y2 = parseFloat(val[1])
    const w2 = parseFloat(val[2])
    const h2 = parseFloat(val[3])
    const start = [Math.max(x1, x2), Math.max(y1, y2)]
    const end = [Math.min(x1 + w1, x2 + w2), Math.min(y1 + h1, y2 + h2)]
    const dx = end[0] - start[0]
    const dy = end[1] - start[1]

    if (dx > 0 && dy > 0) {
      arr = [...arr, val[4]]
    }
  })
  return arr
}
```

รูปภาพที่ 2.15 ตัวอย่างไวยากรณ์ของภาษา JavaScript

2.2.1.3 SQL

SQL หรือ Structured Query Language เป็นภาษาโปรแกรมมิ่งที่ใช้ในการจัดการกับฐานข้อมูลที่มีความสัมพันธ์ (Relational Databases) โดยที่ภาษา SQL นั้นจะช่วยในการเอ็กสแตร์นี้เป็นเอ็กสแตร์ที่ส่งงานให้หรือการเขียนเพื่อการศึกษาเท่านั้น เมื่อนักศึกษาไปใช้ประโยชน์ด้านการค้าเข้าถึงข้อมูล ดึงข้อมูล หรือจัดการกับข้อมูลที่ถูกรวบรวมไว้ในฐานข้อมูลได้ ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งยังมีให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอ็กสแตร์ทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

ภาษา SQL จะใช้งานโดยการเขียนเป็นชุดคำสั่งโดยมักจะถูกเรียกว่าคิวรี โดยจะมีคำสั่งที่จะเจอได้บ่อยสุดดังนี้

- SELECT ใช้ในการดึงข้อมูลจากฐานข้อมูล
- UPDATE ใช้ในการอัปเดตหรือเปลี่ยนแปลงข้อมูลที่มีอยู่ในฐานข้อมูล
- DELETE ใช้ในการลบข้อมูลที่มีอยู่ในฐานข้อมูล
- INSERT INTO ใช้ในการเพิ่มข้อมูลใหม่ลงในฐานข้อมูล
- CREATE DATABASE ใช้ในการสร้างฐานข้อมูลใหม่
- CREATE TABLE ใช้ในการสร้างตารางข้อมูลภายในฐานข้อมูล

โดยจะมีตัวอย่างการใช้งานชุดคำสั่งในการเลือกหรือทำการคิวรีข้อมูลทุกข้อมูลจากตารางข้อมูลที่มีชื่อว่า Student ดังรูปภาพที่ 2.16



```
SELECT * FROM Student;
```

รูปภาพที่ 2.16 ตัวอย่างการใช้งานภาษา SQL

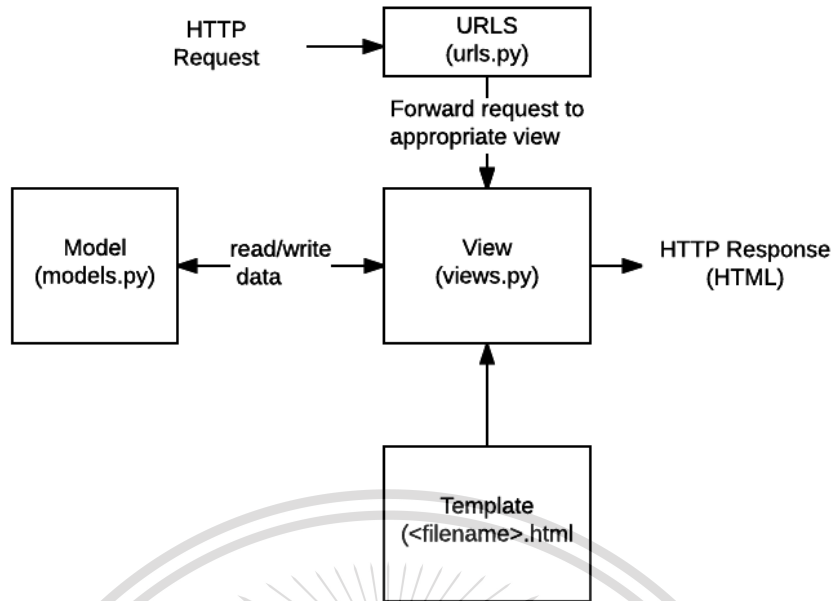
2.2.2 เฟรมเวิร์ก

2.2.2.1 Django & Django Rest Framework

Django เป็นเว็บเฟรมเวิร์กระดับสูงที่พัฒนาด้วยภาษาไพทอน ที่มีการสนับสนุนและการพัฒนาอย่างรวดเร็ว Django ออกแบบมาให้สามารถเข้าใจได้ง่ายและสามารถใช้งานได้จริง โดยจะใช้หลักการทำงานแบบ MVC หรือ Model-View-Control แต่ในเจ้าตัว Django เฟรมเวิร์กนั้นจะถูกเรียก MTV (Model-Template-View)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.



รูปภาพที่ 2.17 ภาพการทำงานคร่าว ๆ ของ Django Framework

โดยจากรูปภาพที่ 2.17 จะสามารถอธิบายส่วนประกอบและการทำงานของแต่ละไฟล์หลัก ๆ ใน Django Framework ได้ดังนี้

- URLs เป็นไฟล์ที่ใช้ในการกำหนด URL หรือ PATH ที่ผู้ใช้งานทำการ Request เข้ามา โดยจะทำการแมพเข้ากับการเรียกใช้ฟังก์ชันต่าง ๆ ใน View เพื่อให้สามารถเข้าใจและพัฒนาต่อไปได้ง่าย

- View เป็นไฟล์ที่ใช้ในการจัดการกับ Request ที่เข้ามาผ่านทาง URLs โดยจะทำงานตามฟังก์ชันที่ได้ทำการเขียนไว้แล้วทำการ Return หรือ Render ไปที่ไฟล์ Template หรืออาจจะใช้งานร่วมกับ Django Rest Framework แล้วทำการ Response กลับไปโดยใช้ HTTP Protocol เป็น Json, หรือ XML

- Model เป็นไฟล์ที่ใช้ในการกำหนดโครงสร้างของข้อมูลโดยจะใช้หลักการของ Object Oriented Programming หรือ OOP เพื่อทำการติดต่อกับฐานข้อมูลในการที่จะเพิ่ม ลบ ค้นหาหรือแก้ไขข้อมูล

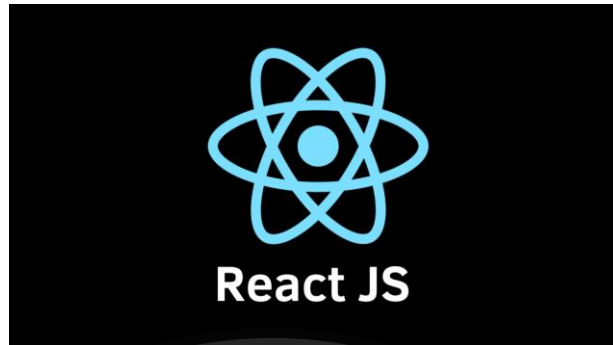
- Template เป็นไฟล์ที่ใช้ในการแสดงผลลัพธ์จากไฟล์ View โดยส่วนใหญ่ในไฟล์นี้จะทำการแสดงผลในรูปแบบของภาษา HTML หรืออาจจะใช้ประเภทอื่นก็ได้

2.2.2.2 ReactJS

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับเอาไว้ใช้ในการศึกษาเพื่อการศึกษาเท่านั้น ไม่สามารถนำไปใช้เชิงพาณิชย์ได้ React เป็น JavaScript ไลบรารีสำหรับการสร้าง UX/UI บนเว็บไซต์ ซึ่ง React จะใช้แนวคิดที่ง่ายต่อการเข้าใจในการสร้างเว็บไซต์ และมีเป้าหมายที่จะเพิ่มประสิทธิภาพและ

This material is reserved for educational use only, not allowed for commercial use.

ความยืดหยุ่นต่อการใช้งาน ReactJS เป็นที่นิยมใช้กันอย่างแพร่หลายและมีคอมมูนิตี้ที่ดีโดยถูกพัฒนาจากมาที่ทีมงานของ Facebook เจ้าตัว React นั้นจะทำการสร้าง UI แบบอินเทอร์แอคทีฟและมีการใช้ state ที่เรียบง่ายโดยจะทำให้สามารถแก้ไขหรือปรับปรุงจุดบกพร่องได้ง่ายขึ้น.



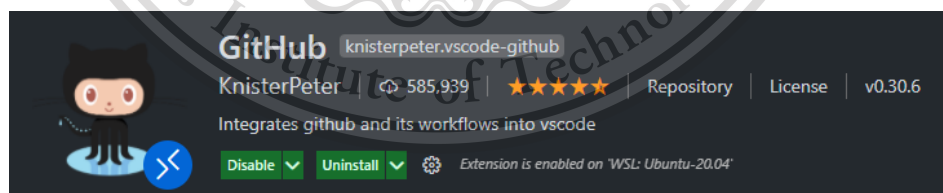
รูปภาพที่ 2.18 โลโก้ของ ReactJS

2.3 โปรแกรมที่ใช้ในการพัฒนาโครงการงาน

2.3.1 Visual Studio Code

Visual Studio Code หรือ VSCode นั้นเป็น Text editor ที่เป็นที่นิยมใช้กันอย่างแพร่หลายโดยมีการรองรับการใช้งานบนหลายระบบปฏิบัติการไม่ว่าจะเป็น Windows, Linux และ macOS โดย VSCode นั้นจะถูกออกแบบมาบนความเรียบง่ายและยังมี Extension เสริมต่าง ๆ มากมายที่ช่วยให้การพัฒนาโปรแกรมต่าง ๆ หรือทำการดีบั๊กโปรแกรมได้ง่ายขึ้น

VScode นั้นเป็น Text Editor ที่ถือว่าใช้ทรัพยากรและสภาพแวดล้อมของเครื่องไม่มากซึ่งทำให้การใช้งานนั้นสิ้นเปลืองและใช้เวลาน้อยในการรันโปรแกรมต่าง ๆ



รูปภาพที่ 2.19 ตัวอย่างการใช้ Extension Github บน VScode

จากรูปภาพที่ 2.19 ในการใช้ Extension Github นั้นจะช่วยให้สามารถทำการเชื่อมต่อส่วนของโค้ดบนเครื่องกับโค้ดบน Github Server และยังสามาถใช้งานการ commit, push หรือ pull นั้นง่ายมากขึ้นโดยไม่จำเป็นต้องพิมพ์คำสั่งใด ๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

2.3.2 Docker

Docker เป็น open-source platform ที่ใช้ในการพัฒนาและรันแอปพลิเคชันต่าง ๆ โดย Docker นั้นถูกออกแบบบนแนวความคิดของการที่สามารถแยกทรัพยากรได้เพื่อให้ง่ายต่อการใช้งานโดยที่จะสามารถทำการติดตั้งและรันได้ทุกที่ที่ต้องการโดยใช้หลักการของ Container

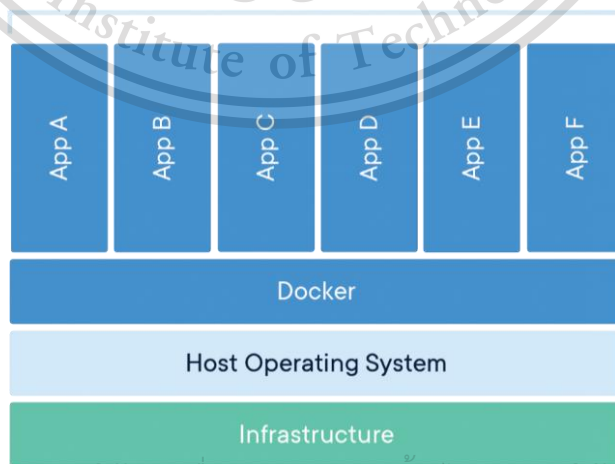
Container คือมาตรฐานหน่วยของซอฟต์แวร์ที่เก็บรวบรวมชุดคำสั่งหรือโค้ดและการอ้างอิงทั้งหมดเพื่อให้แอปพลิเคชันทำงานได้อย่างรวดเร็วและเชื่อถือได้จากสภาพแวดล้อมการประมวลผลหนึ่งไปยังอีกการประมวลผลหนึ่ง

โดยในการใช้ Docker นั้นจะทำการสร้างวัตถุหลัก ๆ ได้แก่ images, container โดยที่จะสามารถอธิบายได้ดังนี้

- Images เป็นรูปแบบหรือเทมเพลตแบบที่จะสามารถทำการอ่านได้อย่างเดียวพร้อมกันกับสิ่งต่าง ๆ ในการสร้าง โดย Images นั้นอาจจะใช้งานของคนอื่นที่มีอยู่แล้วหรือสามารถทำการสร้างขึ้นได้เองโดยในการสร้าง images เองนั้นจะทำการสร้าง Dockerfile สำหรับการกำหนดขั้นตอนหรือคำสั่งที่จำเป็นหรือที่ต้องการในการสร้าง images โดยแต่คำสั่งใน Dockerfile จะทำการสร้างเลเยอร์ใน images เมื่อมีการเปลี่ยนแปลงของ Dockerfile และทำการสร้าง images ใหม่ นั้นระบบจะทำการสร้างเฉพาะส่วนของเลเยอร์ที่เปลี่ยนแปลงไป

- Container เป็นอินสแตนซ์ที่ได้มาจาก Docker Images ด้วยจะสามารถทำการสร้างสิ่งให้รัน หยุด หรือลบออกได้โดยการใช้ Docker API or CLI หรือชุดคำสั่งของ Docker นั้นเอง และยังสามารถทำการเชื่อมต่อระหว่าง Container โดยใช้งาน Docker Compose ซึ่งจะช่วยทำการรวบรวม container ต่าง ๆ ที่มีการใช้งานรวมกันเข้าด้วยกันได้

Containerized Applications



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้าม **รูปภาพที่ 2.20** การทำงานของ Docker ในแต่ละแอปพลิเคชัน ทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

2.3.3 Git

Git เป็น open-source ที่ถูกใช้ในการทำ Version Control System (VCS) โดยจะมีจุดประสงค์เพื่อใช้ในการติดตาม (Track) ไฟล์ในโครงการหรือโปรเจกต์ต่าง ๆ ซึ่งอาจจะมีการเปลี่ยนแปลงตลอดเวลาโดยเมื่อมีการปรับเปลี่ยนเกิดขึ้นจากผู้ใช้งาน Git จะจัดเก็บข้อมูลของการเปลี่ยนแปลงนั้นบนสิ่งที่เรียกว่า Repository และจะเรียกการเก็บของการเปลี่ยนแปลงเหล่านั้นว่าการ commits โดยในการอ้างอิงของแต่ละ commits จะถูกเรียกว่า HEAD และข้อมูลทั้งหมดจะถูกเก็บไว้ในโฟลเดอร์เดียวกับโปรเจกต์ซึ่งส่วนใหญ่จะถูกซ่อนแล้ว โดยที่สามารถการอัปเดตชุดของโค้ดหรือไฟล์ต่าง ๆ ในโปรเจกต์ไปได้ทั้งหมด Github หรือ Gitlab แล้วแต่การเลือกใช้งาน



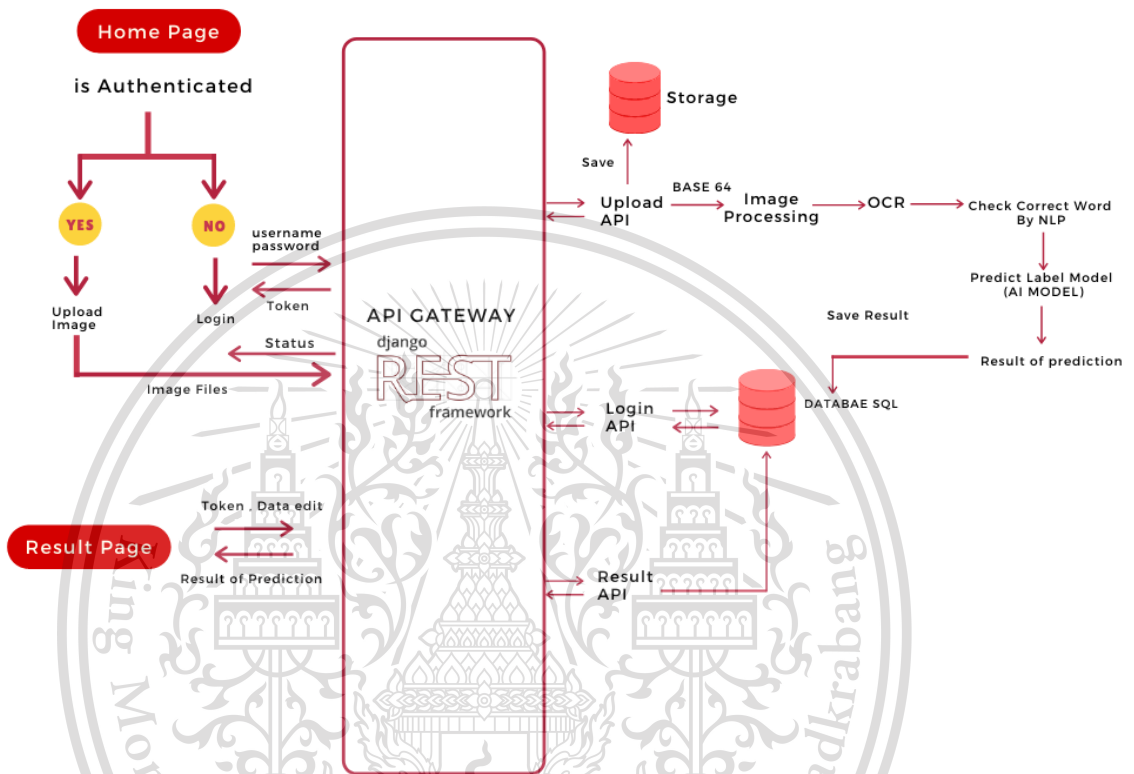
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

บทที่ 3

วิธีการดำเนินงานโครงการ

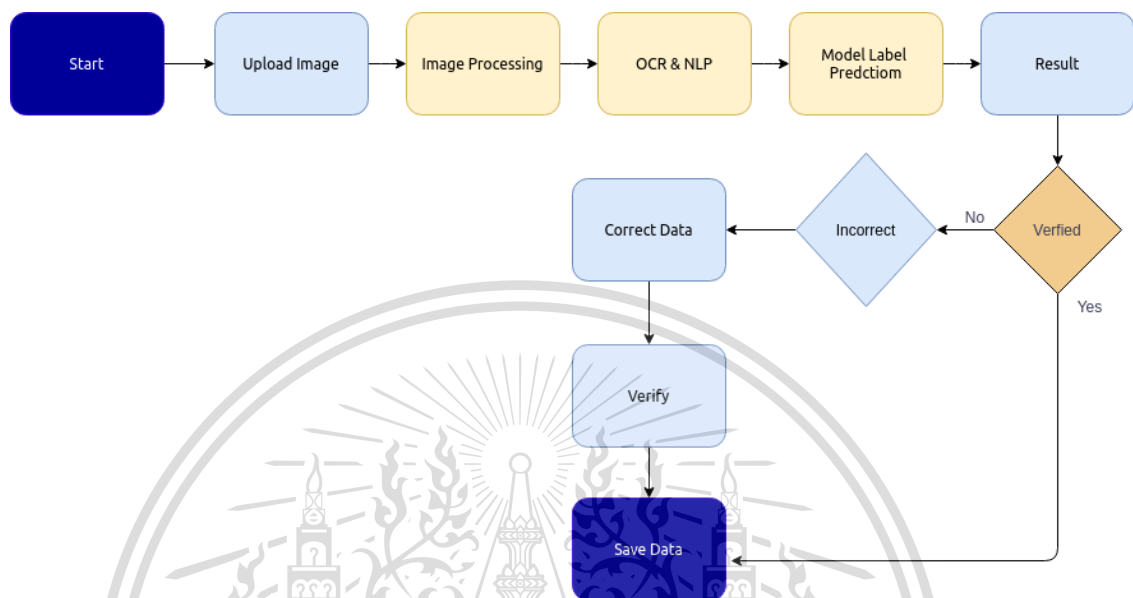
3.1 ภาพรวมของโครงการ



รูปภาพที่ 3.1 ภาพรวมของเว็บแอปพลิเคชัน

จากรูปที่ 3.1 ภาพรวมเว็บแอปพลิเคชัน จะแสดงโครงสร้างการติดต่อสื่อสารของกายระบบ โดยที่การทำงานหลักของระบบนั้นจะประกอบด้วยหน้า Home Page จะเป็นหน้าที่แสดงประวัติของการเรียนการใช้ Prediction Model โดยจะแสดงสถานะว่าถูก Verified แล้วหรือยัง หน้า Upload Page จะเป็นหน้าที่เอาไว้สำหรับการอัปโหลดรูปภาพที่ต้องการจะดึงข้อมูลออกมาโดยจะส่งรูปภาพเข้าไปเรียกการทำ OCR หลังจากนั้นจะทำการเช็คค่าที่ได้มาจากการ OCR แล้วสุดท้ายจึงไปเรียก Prediction Model หลังจากนั้นจะบันทึกผลลัพธ์ที่ได้ลงใน Database และหน้าสุดท้ายคือหน้า Result Page จะแสดงข้อมูลที่ถูกบันทึกลงใน Database ไว้แล้วมาแสดงให้กับผู้ใช้งานว่าผลลัพธ์ที่ได้เป็นอย่างไร โดยผลลัพธ์นั้นจะแสดงในรูปแบบ Key-Value พร้อมทั้งกับ ผู้ใช้สามารถที่จะแก้ไขในเอกสารนี้เป็นการ Prediction ผิด หรือการทำ OCR แล้วค่าไม่ถูกต้องได้ และสามารถที่จะเพิ่มฟิลด์ที่ต้องการเข้าไปเพิ่มได้ โดยที่การติดต่อจากฝั่งหน้าบ้านหรือ Client ต่างๆ นั้น จะติดต่อกับฝั่งหลัง

บ้านหรือ Server โดยผ่าน API โดยใช้ Django Rest Framework มาติดต่อกับ Django ในการประมวลผลต่าง ๆ เช่น การบันทึกข้อมูลลง Database หรือ การประมวลผลภาพและ Model Prediction โดยการติดต่อกันทั้งหมดจะต้องมีการ Authentication โดยใช้ Token เพื่อเข้าใช้งานระบบทั้งหมด



รูปภาพที่ 3.2 ภาพแสดงขั้นตอนการทำงานของระบบ

3.2 การออกแบบหน้าเว็บแอปพลิเคชันและเซอร์วิส

3.2.1 การออกแบบหน้าเว็บแอปพลิเคชัน

การออกแบบหน้าเว็บแอปพลิเคชันนั้นจะคำนึงการทำอะไรให้ผู้ใช้สามารถใช้งานและสามารถที่จะทำการแก้ไขข้อมูลที่ผิดพลาดได้อย่างรวดเร็วและง่ายสะดวกสบายในการใช้งานหน้าเว็บแอปพลิเคชันและใช้งานได้อย่างมีประสิทธิภาพโดยได้มีการออกแบบหน้าผู้ใช้งานไว้ดังนี้

- | | |
|-----------------------------------|--|
| 1. หน้าเข้าสู่ระบบ | สำหรับการยืนยันสิทธิการใช้งาน |
| 2. หน้าลงทะเบียน | สำหรับการลงทะเบียนเพื่อขอสิทธิการใช้งาน |
| 3. หน้าหลัก | สำหรับแสดงรายการที่เรียกใช้การดึงข้อมูล |
| 4. หน้าอัปโหลดรูปภาพสลิปเงินเดือน | สำหรับอัปโหลดรูปภาพเพื่อเรียกใช้การดึงข้อมูล |
| 5. หน้าแสดงผลและแก้ไขข้อมูล | สำหรับแสดงผลของการดึงข้อมูลและแก้ไขให้ |

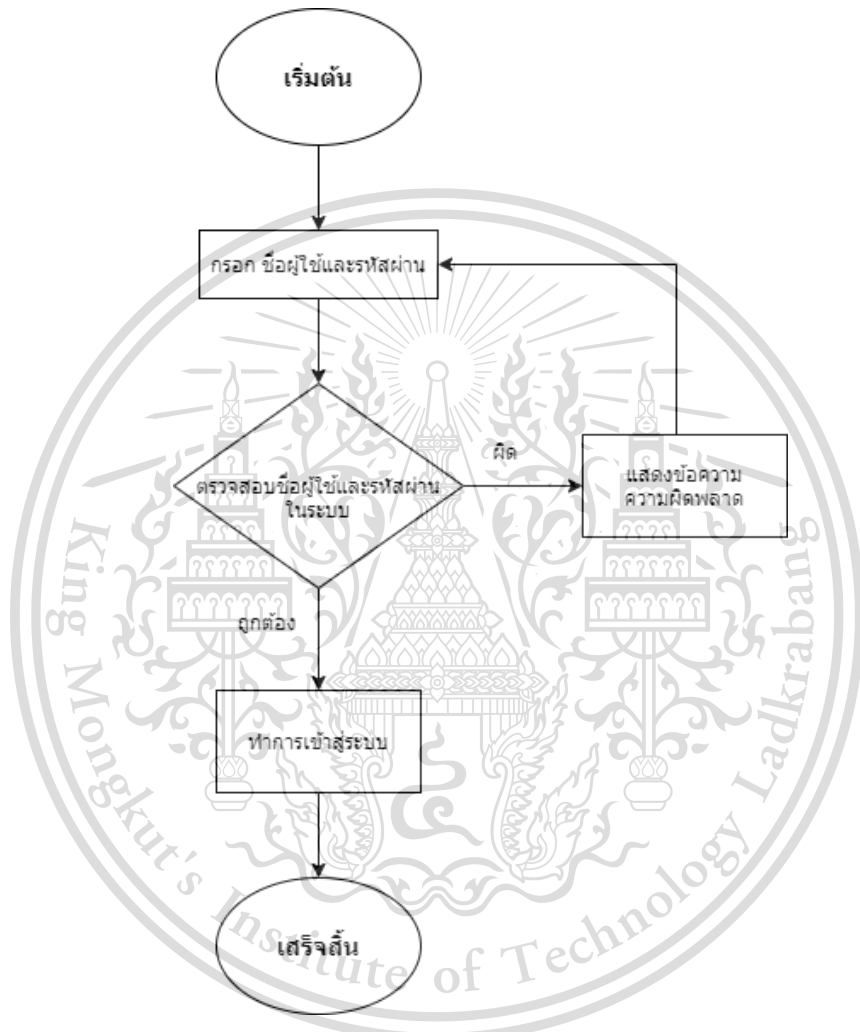
ถูกต้อง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

1. หน้าเข้าสู่ระบบ

เป็นหน้าที่ผู้ใช้งานนั้นจะต้องทำการกรอก ชื่อผู้ใช้งาน และ รหัสผ่าน ตามที่ผู้ดูแลระบบได้ทำการสร้างไว้ให้ เพื่อทำการยืนยันสิทธิให้สามารถทำการใช้งานในฟังก์ชันต่าง ๆ ของเว็บแอปพลิเคชันได้ โดยจะนำชื่อผู้ใช้งาน และ รหัสผ่านไปตรวจสอบกับฐานข้อมูลว่าถูกต้องหรือไม่



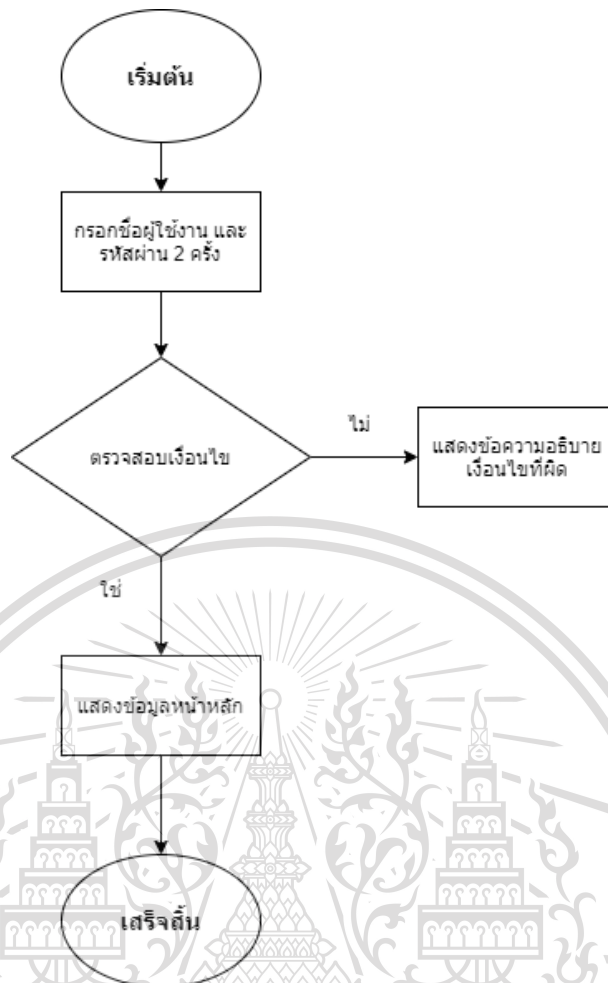
รูปภาพที่ 3.3 ภาพแสดง Flow ของหน้าเข้าสู่ระบบ

2. หน้าลงทะเบียน

เป็นหน้าที่ผู้ใช้งานโดยจะต้องทำการกดปุ่ม Register ที่หน้าเข้าสู่ระบบโดยผู้ใช้งานจะต้องทำการกรอกชื่อผู้ใช้งานและรหัสผ่าน โดยจะต้องทำการยืนยันรหัสผ่านสองครั้งให้และต้องให้ตรงกันซึ่งการตั้งรหัสผ่านนั้นจะมีเงื่อนไขในการตั้งเพื่อไม่ให้ง่ายเกินไป และไม่สามารถใช้ชื่อผู้ใช้งานซ้ำได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.



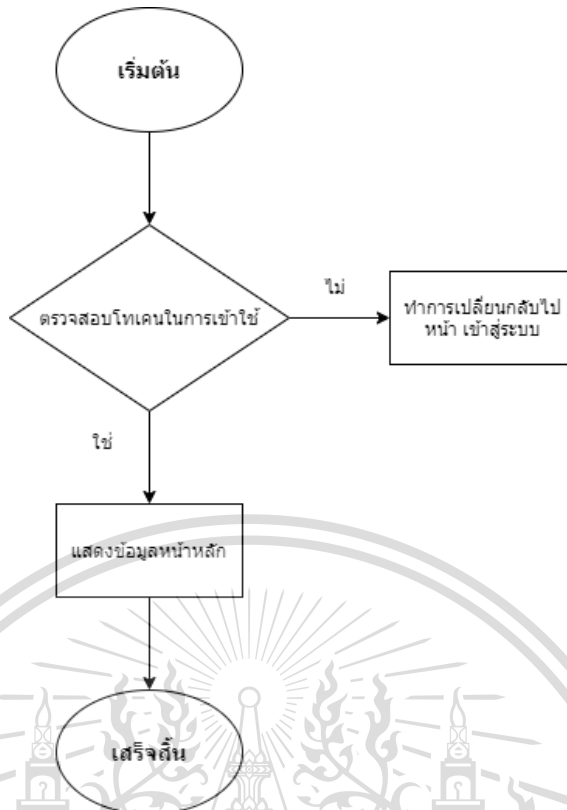
รูปภาพที่ 3.4 ภาพแสดง Flow ของหน้าลงทะเบียน

3. หน้าหลัก

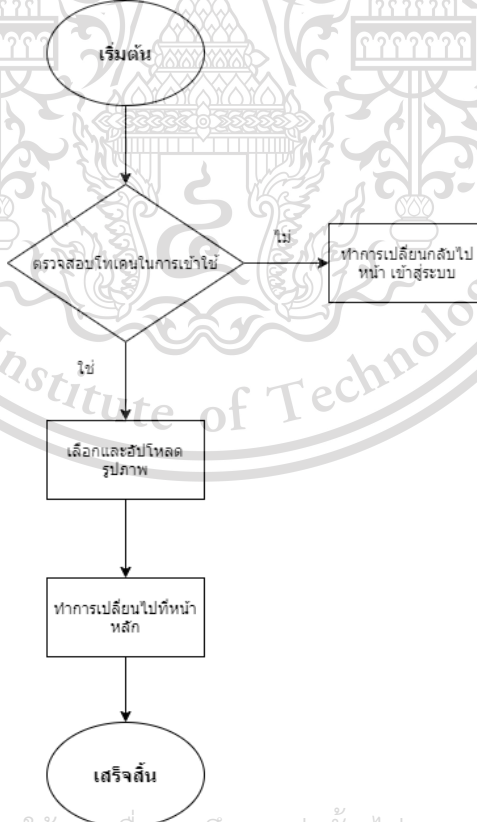
เป็นหน้าแสดงรายการที่ได้มีการอัปโหลดรูปภาพสลิปเงินเดือนที่ผ่านการดึงข้อมูลฟิลด์โดยใช้โมเดลการทำนายผลเรียบร้อยแล้ว โดยแต่ละรายการจะแสดงผลโดยมีสถานะเป็น unverified ไว้เพื่อที่จะให้ผู้ใช้งาน เข้าไปดูหรือตรวจสอบและแก้ไขให้ถูกต้องที่หน้าของแสดงผลและแก้ไขข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.



รูปภาพที่ 3.5 ภาพแสดง Flow ของหน้าหลัก



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 รูปภาพที่ 3.6 ภาพแสดง Flow ของหน้าอัปเดตรูปภาพสลิปเงินเดือน
 ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมีเหตุดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

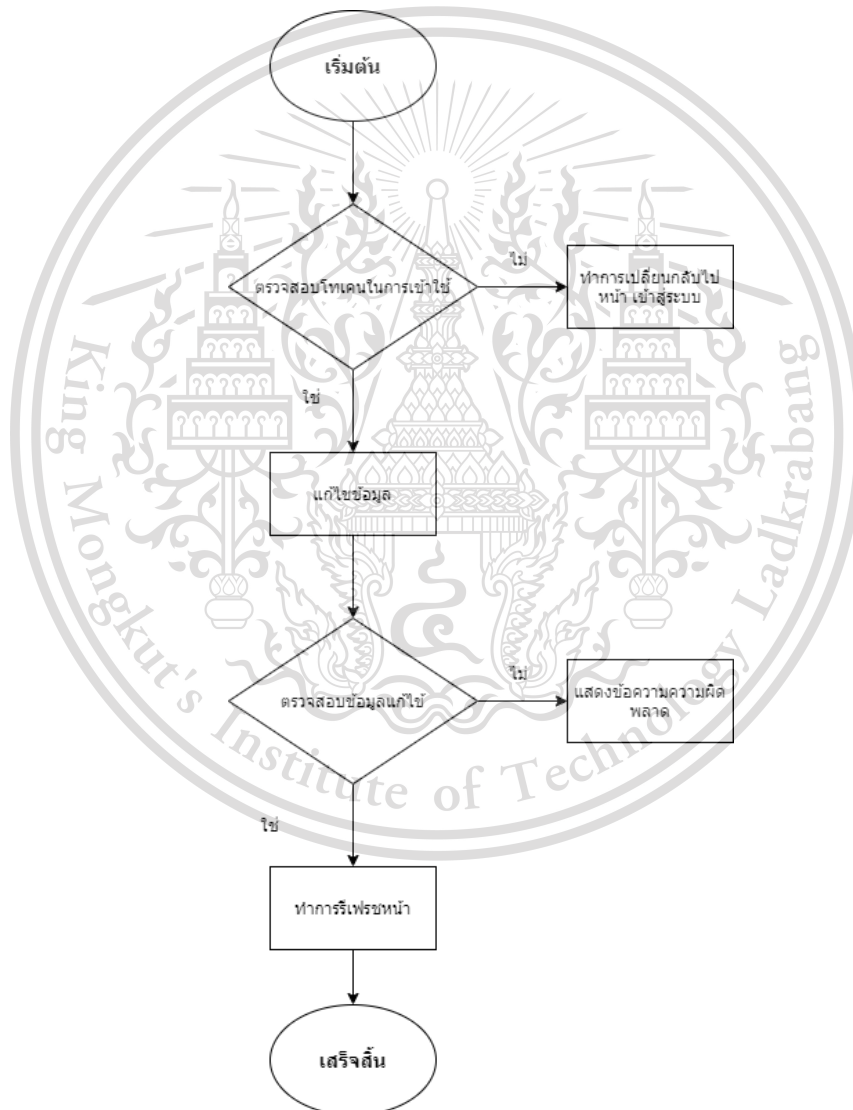
This material is reserved for educational use only, not allowed for commercial use.

4. หน้าอัปโหลดรูปภาพสลิบเงินเดือน

เป็นหน้าที่เอาไว้สำหรับให้ผู้ใช้งานนั้นทำการเลือกรูปภาพสลิบเงินเดือนที่ต้องและอัปโหลดเข้าสู่ตัวระบบเพื่อที่จะทำการดึงข้อมูลฟิลต์โดยใช้โมเดลการทำนายผลและเมื่อทำการเสร็จสิ้นรายการที่อัปโหลดจะถูกแสดงผลได้หน้าหลัก

5. หน้าแสดงผลและแก้ไขข้อมูล

เป็นหน้าที่แสดงผลลัพธ์ของการเรียกใช้ระบบการดึงข้อมูลฟิลต์โดยใช้โมเดลการทำนายผลว่าผลลัพธ์ที่ได้แต่ละฟิลต์นั้นมีค่าอะไรบ้าง และสามารถที่จะแก้ไขข้อมูลที่ทำการทำนายผิดพลาดได้อย่างง่ายและสะดวกสบายมากที่สุด



รูปภาพที่ 3.7 ภาพแสดง Flow ของหน้าแสดงผลและแก้ไขข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

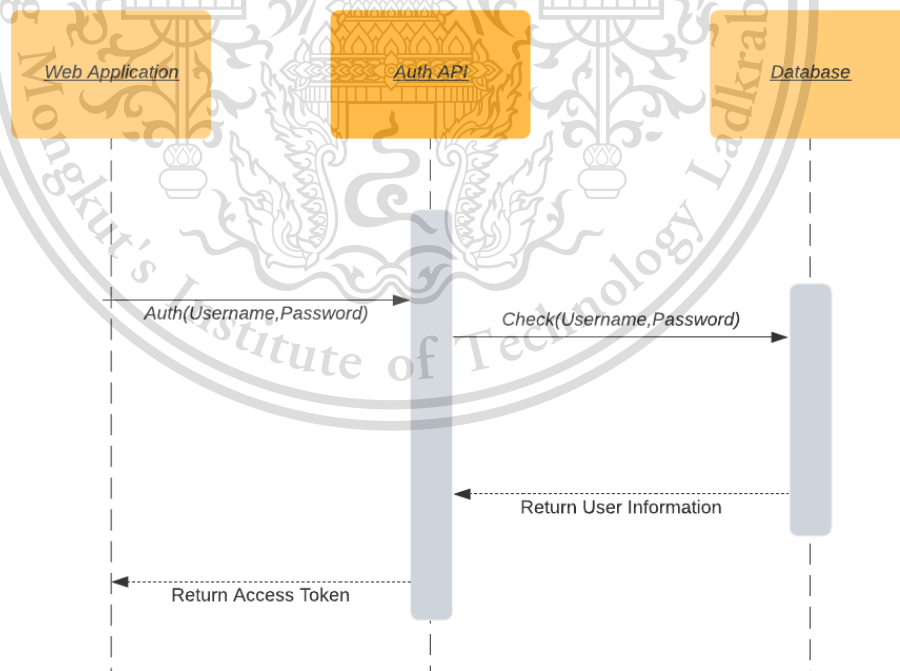
This material is reserved for educational use only, not allowed for commercial use.

3.2.2 การออกแบบเว็บเซอร์วิส

การออกแบบเซอร์วิสทั้งหมดที่มีให้บริการนั้นจะใช้ Django เป็นตัวหลักและใช้ Django Rest Framework เป็นตัวให้บริการ API ซึ่งจะให้บริการเซอร์วิสต่าง ๆ โดยใช้ RESTful API ซึ่ง API ทั้งหมดนั้นจะต้องทำการยืนยันตัวตนและสิทธิการเข้าถึงก่อนโดยจะใช้โทเคนในการยืนยันตัวตนและจะใช้รูปแบบของ JSON เป็น Response และ Request Body ในการรับ-ส่งข้อมูลต่าง ๆ ภายในแต่ละเซอร์วิสซึ่งจะประกอบไปด้วยดังนี้

1. การตรวจสอบและยืนยันตัวตนการเข้าถึงหน้าเว็บแอปพลิเคชัน

เป็นเซอร์วิสที่ใช้ในการตรวจสอบและยืนยันตัวตนและสิทธิในการสามารถที่จะเข้าถึงหรือเข้าใช้หน้าเว็บแอปพลิเคชันและฟังก์ชันต่าง ๆ โดยจะทำการตรวจสอบชื่อผู้ใช้งานและรหัสผ่านที่ผู้ใช้งานได้ทำการกรอกข้อมูลที่หน้าเข้าสู่ระบบ และระบบจะทำการตรวจสอบในฐานข้อมูลว่ามีชื่อผู้ใช้งานและรหัสผ่านนี้หรือไม่ถ้ามีระบบจะทำการตอบกลับไปโดยจะมี HTTP Response Status Code 200 (OK) และ HTTP Response Body เป็นโทเคนของผู้ใช้งานเพื่อใช้ในการยืนยันสิทธิในการเข้าใช้งานฟังก์ชันต่าง ๆ ในระบบต่อไปแต่ถ้าไม่ระบบจะทำการตอบกลับโดยจะมี HTTP Response Status Code 401 (Unauthorized) และ HTTP Response Body เป็นข้อความบอกว่า ชื่อผู้ใช้งานและ/หรือรหัสผ่านไม่ถูกต้อง



รูปภาพที่ 3.8 ภาพแสดง Sequence Diagram ของเซอร์วิสตรวจสอบยืนยันตัวตน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

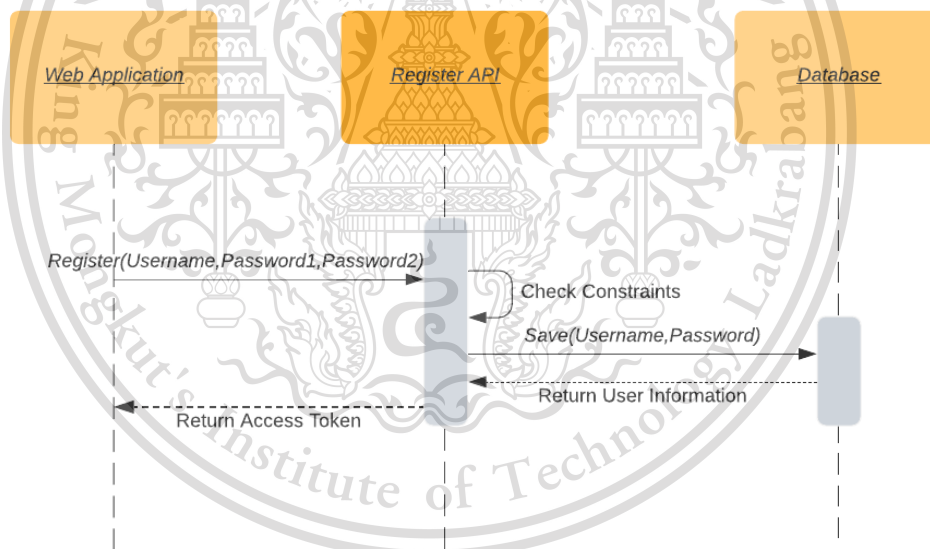
This material is reserved for educational use only, not allowed for commercial use.

ตารางที่ 3.1 แสดง HTTP Status Code ของเซอร์วิส Authentication

HTTP Status Code	Description
200 (OK)	ชื่อผู้ใช้งานและรหัสผ่านถูกต้อง
401 (Unauthorized)	ชื่อผู้ใช้งานหรือรหัสผ่านไม่ถูกต้อง
400 (Bad Request)	รูปแบบการส่งข้อมูลผิดพลาด

2. การลงทะเบียน

เป็นเซอร์วิสที่จะใช้ในการลงทะเบียนเพื่อที่จะสามารถใช้ชื่อผู้ใช้งานและรหัสผ่านเข้าสู่ระบบ เพื่อเข้าใช้งานตัวเซอร์วิสจะรับชื่อผู้ใช้งานและรหัสผ่านเข้ามา โดยทำการตรวจสอบชื่อผู้ใช้งานก่อนว่ามีชื่อผู้ใช้งานนี้แล้วหรือไม่และทำการตรวจสอบรหัสผ่านว่ารหัสผ่านทั้งสองครั้งตรงกันหรือไม่แล้วมีความง่ายเกินไปหรือไม่โดยจะทำการกำหนดเกณฑ์การตรวจสอบโดยจะต้องมีความยาวไม่น้อยกว่า 8 อักขรและจะต้องประกอบไปด้วยตัวเลขและตัวอักษรภาษาอังกฤษผสมกัน



รูปภาพที่ 3.9 ภาพแสดง Sequence Diagram ของเซอร์วิสการลงทะเบียน

ตารางที่ 3.2 แสดง HTTP Status Code ของเซอร์วิสการลงทะเบียน

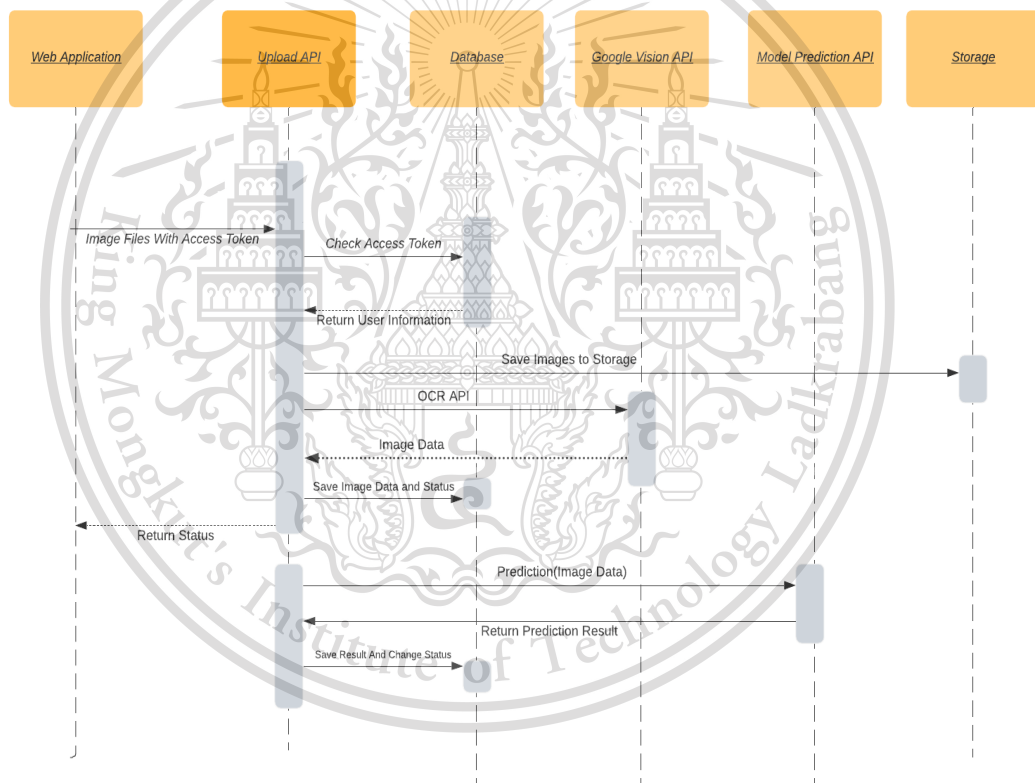
HTTP Status Code	Description
200 (OK)	ระบบทำการบันทึกสมัครสมาชิกสำเร็จ
400 (Bad Request)	เงื่อนไขของ Username หรือ Password ไม่ถูกต้อง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

3. การอัปโหลดรูปภาพและการดึงข้อมูลฟิลต์โดยใช้โมเดลการทำนายผล

เป็นเซอวีริสที่ใช้ในการบันทึกรูปภาพสลิปเงินเดือนและทำการดึงข้อมูลฟิลต์ที่ต้องการโดยใช้โมเดลการทำนายผลโดยเมื่อผู้ใช้งานทำการอัปโหลดภาพเข้าสู่ระบบแล้ว ระบบจะทำบันทึกรูปภาพลงในพื้นที่จัดเก็บข้อมูล(Storage) และบันทึกข้อมูลรายการลงในฐานข้อมูลแต่จะมีสถานะของรายการนี้ว่าเป็น Waiting หลังจากนั้นจะทำการเรียกใช้ฟังก์ชันการดึงข้อมูลและทำนายผลโดยฟังก์ชันจะทำการ Image Thresholding และทำการลบเส้นตารางในรูปภาพและหลังจากนั้นจะทำการเรียกใช้ Google Vision API เพื่อ OCR หรือแปลงรูปภาพให้ออกเป็นข้อความแล้วทำการนำข้อมูลที่ทำการแปลงรูปแบบหลังจากทำ OCR แล้วเข้าสู่การดึงข้อมูลฟิลต์ที่ต้องการโดยใช้โมเดลการทำนายผล เมื่อทำเสร็จสิ้นการดึงข้อมูลแล้วจะทำการเปลี่ยนสถานะของรายการเป็น Success แต่ถ้าเกิดมีข้อผิดพลาดเกิดขึ้นระหว่างการดึงข้อมูลขึ้นสถานะของรายการนั้นจะเปลี่ยนเป็น Error



รูปภาพที่ 3.10 ภาพแสดง Sequence Diagram ของเซอวีริสการอัปโหลดรูปภาพและการดึงข้อมูล
ตารางที่ 3.3 แสดง HTTP Status Code ของเซอวีริสการอัปโหลดรูปภาพและการดึงข้อมูล

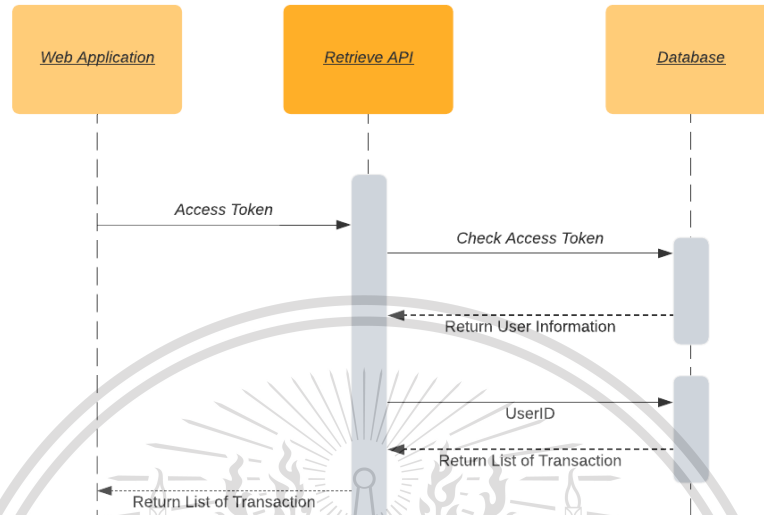
HTTP Status Code	Description
200 (OK)	ระบบทำการบันทึกข้อมูลสำเร็จไม่เกิดข้อผิดพลาด
401 (Unauthorized)	ไม่มีโทเคนในการยืนยันหรือโทเคนไม่ถูกต้อง
400 (Bad Request)	รูปแบบการส่งข้อมูลผิดพลาด

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์การใช้งานเพื่อการศึกษาค้นคว้าเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น ยกเว้นที่ให้มีเหตุที่เปลี่ยนแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

4. แสดงข้อมูลรายการที่เรียกใช้การอัปโหลดรูปภาพและการดึงข้อมูลฟิลด์

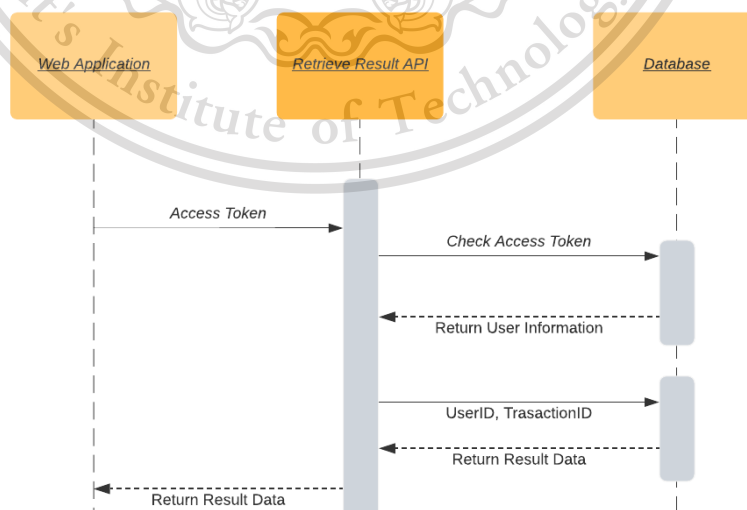
เป็นเซอว์ริสที่ใช้ในการเรียกแสดงลิสต์รายการประวัติที่เคยเรียกใช้การอัปโหลดภาพและการดึงข้อมูลฟิลด์ของผู้ใช้งานนั้น ๆ โดยจะไปทำการคิวรีรายการของข้อมูลจากฐานข้อมูลที่มีการบันทึกไว้



รูปภาพที่ 3.11 ภาพแสดง Sequence Diagram ของเซอว์ริสแสดงรายการ

5. แสดงผลลัพธ์การดึงข้อมูลฟิลด์โดยใช้โมเดลการทำนายผล

เป็นเซอว์ริสที่ใช้ในการแสดงผลลัพธ์ของที่ได้จากการทำการดึงข้อมูลฟิลด์โดยใช้โมเดลการทำนายผลซึ่งจะส่งข้อมูลของแต่ละฟิลด์ว่ามีค่าอะไรบ้าง



รูปภาพที่ 3.12 ภาพแสดง Sequence Diagram ของเซอว์ริสแสดงผลลัพธ์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามแก้ไขตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

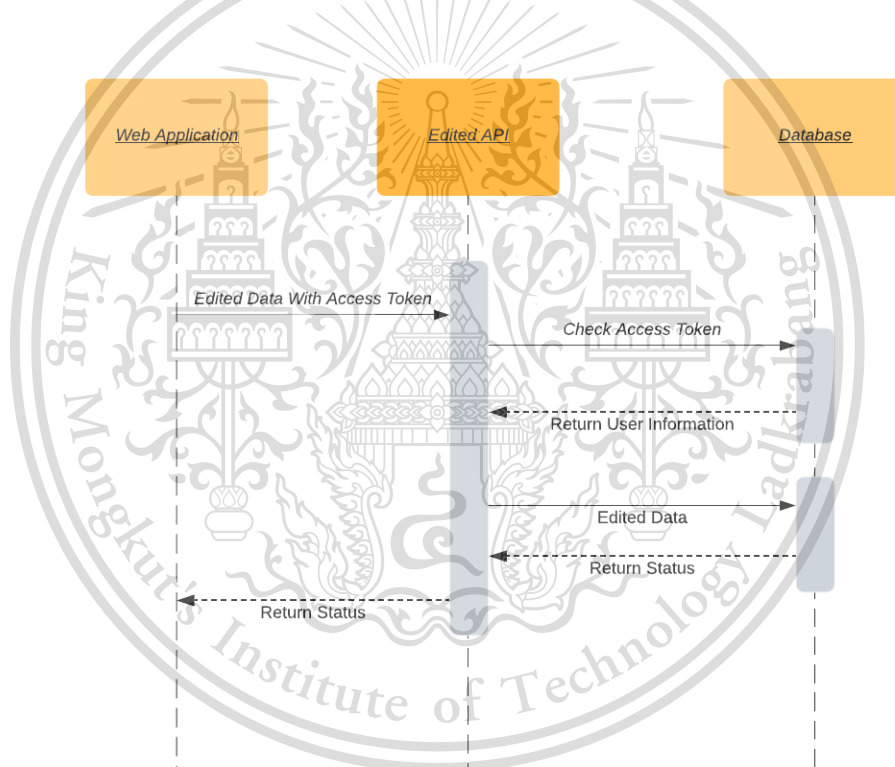
This material is reserved for educational use only, not allowed for commercial use.

ตารางที่ 3.4 แสดง HTTP Status Code ของการแสดงผลการดึงข้อมูลฟิลด์โดยใช้โมเดล
การทำงานายผล

HTTP Status Code	Description
200 (OK)	ระบบทำการดึงข้อมูลได้สำเร็จ
401 (Unauthorized)	ไม่มีโทเคนในการยืนยันหรือโทเคนไม่ถูกต้อง
400 (Bad Request)	รูปแบบการส่งข้อมูลผิดพลาด

6. การแก้ไขข้อมูล

เป็นเซอวิริสที่ใช้เมื่อมีการทำการแก้ไขข้อมูลต่าง ๆ โดยจะทำการส่งข้อมูลที่มีการแก้ไขนั้น
กลับไปและทำการแก้ไขข้อมูลภายในฐานข้อมูล



รูปภาพที่ 3.13 ภาพแสดง Sequence Diagram ของเซอวิริสการแก้ไขผลลัพธ์

ตารางที่ 3.5 แสดง HTTP Status Code ของการแก้ไขข้อมูล

HTTP Status Code	Description
200 (OK)	ระบบทำการแก้ไขข้อมูลได้สำเร็จ
401 (Unauthorized)	ไม่มีโทเคนในการยืนยันหรือโทเคนไม่ถูกต้อง
400 (Bad Request)	รูปแบบการส่งข้อมูลผิดพลาด

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์การใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

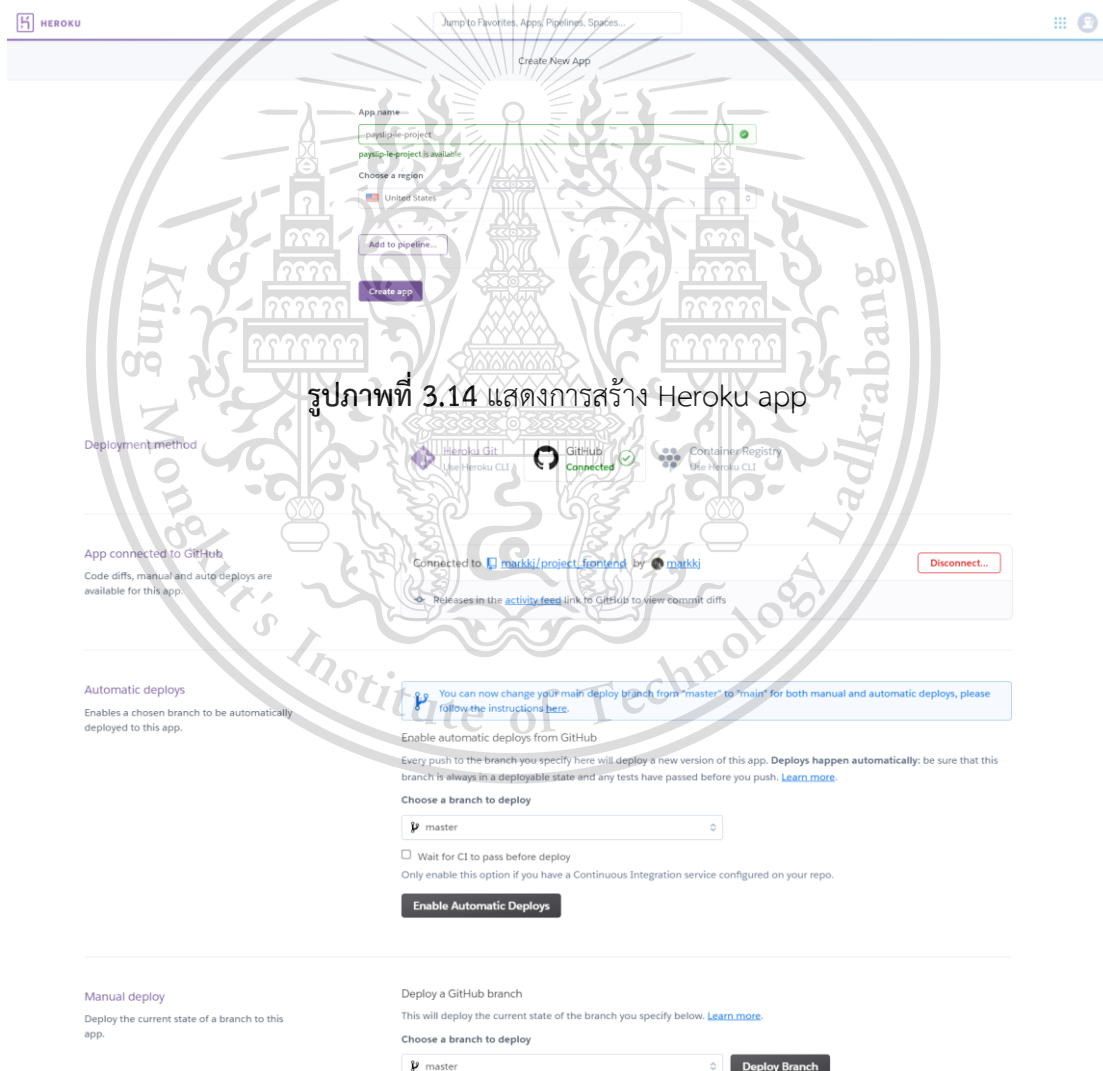
This material is reserved for educational use only, not allowed for commercial use.

3.3 ขั้นตอนการ Deployment เว็บแอปพลิเคชันและเซอร์วิสขึ้นบนคลาวด์เซอร์วิส

ในขั้นตอนนี้จะเป็นการนำเว็บแอปพลิเคชันและเซอร์วิสที่ได้ทำการออกแบบไว้ขึ้นไปบนคลาวด์ โดยจะทำการแบ่งการ Deployment ของเว็บแอปพลิเคชันและเซอร์วิสนั้นจะแตกต่างกันซึ่งจะทำได้ดังนี้

3.3.1 การ Deploy เว็บแอปพลิเคชันขึ้นบน Heroku Cloud Platform

ในส่วนของเว็บแอปพลิเคชันนั้นจะทำการ Deploy ขึ้นบน Heroku app engine โดยจะทำการสร้างโปรเจกต์และ deploy ผ่านหน้าเว็บของทาง Heroku ได้เลย โดยจะทำการเชื่อมต่อ Heroku app ที่สร้างขึ้นกับ Github Repository ของไฟล์โปรเจกต์ โดยจะทำการสร้างตามรูปภาพที่ 3.14



รูปภาพที่ 3.14 แสดงการสร้าง Heroku app

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับใช้ภายในเท่านั้น ไม่สามารถนำออกเผยแพร่ไปภายนอกได้โดยไม่ใช้ประโยชน์ด้านการค้า

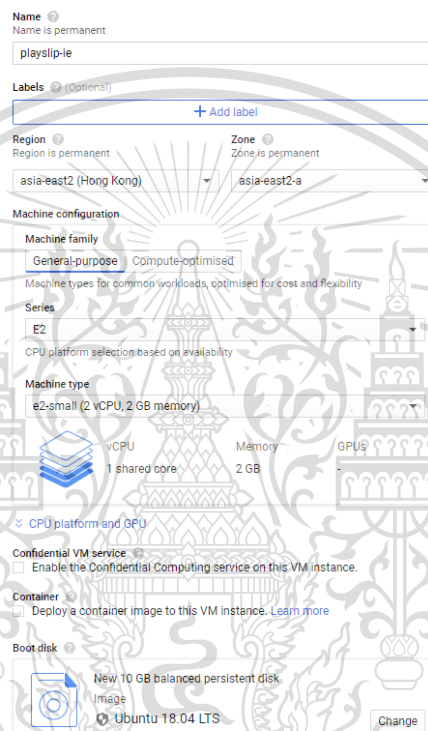
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

จากรูปที่ 3.15 เมื่อทำการเชื่อมต่อ Heroku app กับ Github เรียบร้อยแล้ว สามารถทำการกดปุ่ม “Deploy Branch” เพื่อทำการ Deploy ขึ้นบน Heroku app ได้เลย

3.3.2 การ Deploy เซอร์วิสขึ้นบน Google Cloud Platform

ในการส่วนของเซอร์วิสนั้นจะทำการ Deploy จะใช้ Google Compute Engine โดยจะต้องทำการสร้าง VM instance หรือก็คือการสร้างเครื่อง Virtual Server Machine โดยที่สามารถทำการกำหนดฮาร์ดแวร์และซอฟต์แวร์ของเครื่องที่ต้องการจะสร้างเพื่อให้เหมาะสมพอดีกับสิ่งที่ต้องการได้แต่ก็มีราคาที่แตกต่างกัน



รูปภาพที่ 3.16 แสดงการสร้าง VM instance บน Google Cloud Platform

เมื่อทำการสร้าง VM instance เรียบร้อยแล้วจะเข้าใช้งานผ่าน Secure Shell หรือ SSH ซึ่งสามารถกดเข้าใช้งานผ่านหน้าเว็บ Google Cloud ได้เลย หลังจากนั้นจะทำการ git clone โปรเจกต์ที่ได้ทำการสร้างไว้ลงมาที่เครื่อง VM instance ที่ได้ทำการสร้างไว้ และทำการรันคำสั่งให้ เซอร์วิสต่าง ๆ ทำงานโดยจะได้ผลลัพธ์ดังรูปภาพที่ 3.17 และ 3.18

```
johnyosbk@payslip-ie:~$ git clone https://github.com/markkj/project_backend.git
Cloning into 'project_backend'...
Username for 'https://github.com':
Password for 'https://github.com':
remote: Enumerating objects: 287, done.
remote: Counting objects: 100% (203/203), done.
remote: Compressing objects: 100% (123/123), done.
remote: Total 287 (delta 96), reused 169 (delta 70), pack-reused 84
Receiving objects: 100% (287/287), 116.63 MiB | 16.27 MiB/s, done.
Resolving deltas: 100% (109/109), done.
johnyosbk@payslip-ie:~$
```

เอกสารนี้เป็น

ค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งรูปภาพที่ 3.17 แสดงการ git clone โปรเจกต์ลงบน VM instance ครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

Forbidden to modify the content, and cite the document when use.

```

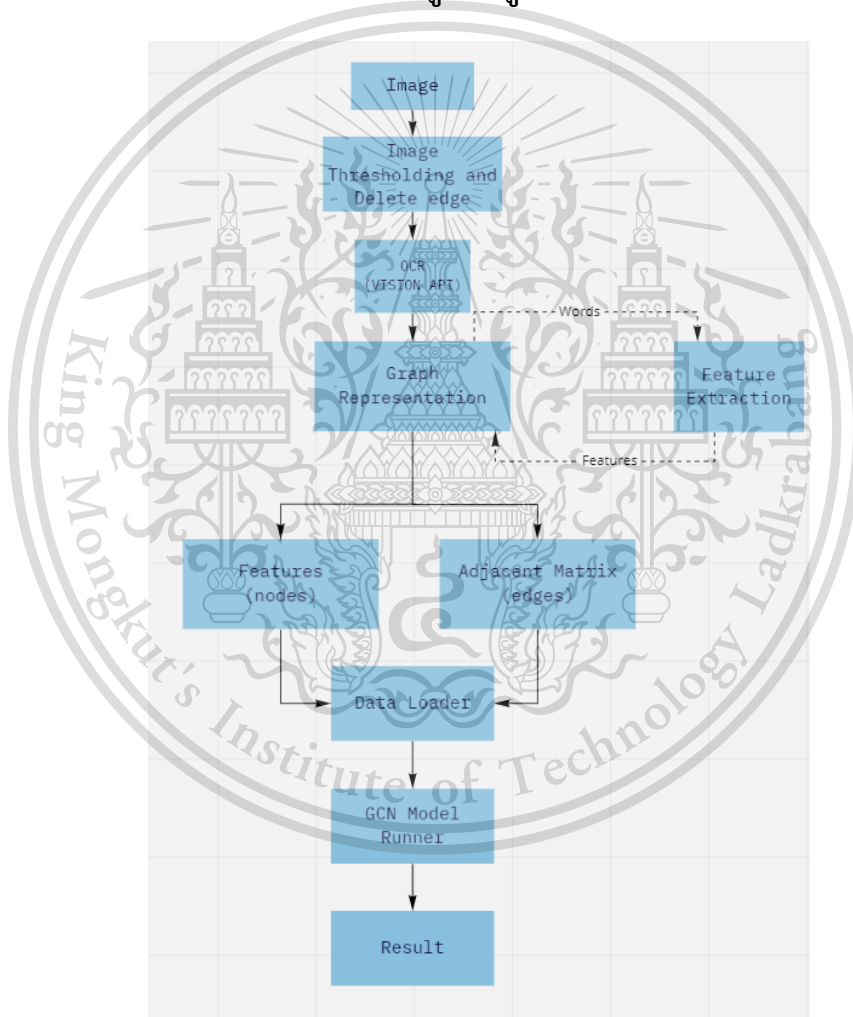
khajohnyosbk@payslip-ie:~/project_backend$ sudo docker-compose up -d
Starting celery_worker ... done
Starting backend_db ... done
Starting rd01 ... done
Starting backend_api ... done
khajohnyosbk@payslip-ie:~/project_backend$

```

รูปภาพที่ 3.18 แสดงการรันคำสั่งเพื่อให้เซอวิสทำงาน

โดยจากรูปที่ 3.18 จะทำการรันให้เซอวิสทำงานผ่าน Docker โดยจะใช้ของ docker-compose ในการสั่งการทำงานของเซอวิสทั้ง

3.4 ขั้นตอนการพัฒนากระบวนการทำนายข้อมูลบนรูปภาพสลิปเงินเดือน



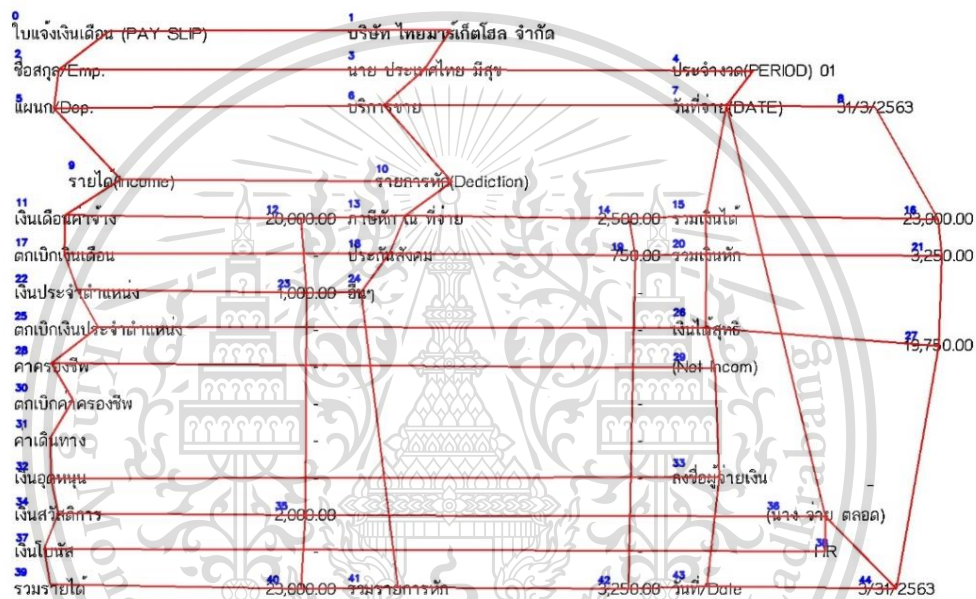
รูปภาพที่ 3.19 แสดงภาพรวมของระบบการทำนายข้อมูลบนรูปภาพสลิปเงินเดือน

3.4.1 การสร้างกราฟของข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้รูปภาพของสลิปเงินเดือนจะถูกทำการปรับ Thresholding และทำการลบเส้นด้านการค้า
ไม่ว่ากรณีใดก็ตามจะทำการ OCR ผ่าน Google Vision API โดยที่จะได้ข้อมูล

ออกมาในรูปแบบของ word boxes หลังจากนั้นก็จะทำการรวม word boxes ที่น่าจะเป็น segment เดียวกันโดยดูจากระยะห่างแต่ละ word boxes โดยที่ถ้ามีระยะห่างน้อยกว่าค่าควอไรรี่ที่ 0.15 จากระยะห่างทั้งหมดของ word boxes ภายใน 1 image ก็จะทำให้การรวมเป็น text areas เดียวกันโดยที่จะเรียกแต่ละ text areas ว่า โหนด

หลังจากทำการจัด text area เรียบร้อยแล้วก็จะทำการสร้าง Graph โดยดูจากความสัมพันธ์ของ แต่ละโหนดโดยที่แต่ละโหนดนั้นจะมีได้สูงสุด 4 edges สำหรับแต่ละ node edge จะเชื่อมต่อกับ node อื่นที่มีระยะใกล้มากที่สุดในแต่ละทิศทาง คือ บน ล่าง ซ้าย และ ขวา โดยจะเป็นดังรูปภาพที่ 3.20



รูปภาพที่ 3.20 แสดงตัวอย่างการสร้างกราฟ

3.4.1.1 การปรับ Thresholding และการลบเส้นตารางออกจากรูปภาพสลิป

เงินเดือน

ในการทำการปรับ Threshold ของรูปภาพที่เข้ามานั้นจะใช้อัลกอริทึม “Sauvola local threshold” ซึ่งจะทำให้รับมือกับรูปภาพที่มีแสงสะท้อนเข้ามาได้ดี โดยหลังจากทำการปรับ Threshold เรียบร้อยแล้วจะนำไปทำการลบเส้นตารางในรูปภาพออกโดยใช้การทำ Morphological ซึ่งจะใช้ไลบรารีของไพทอนที่ชื่อว่า open-cv เข้ามาช่วย

โดยจะใช้การทำ opening ตามแนวนอนและแนวตั้งของรูปภาพซึ่งจะใช้ kernel ที่มีความแตกต่างกันใน แนวนอนนั้นจะใช้ kernel ที่มีขนาดเท่ากับ 1xความกว้างของรูปภาพ/100 และในแนวตั้งจะใช้ขนาดเท่ากับ 1xความสูงของรูปภาพ/100 ซึ่งหลังจากนั้นจะทำการลบเส้นทั้งแนวตั้งและแนวนอนด้วยการจับ Contour

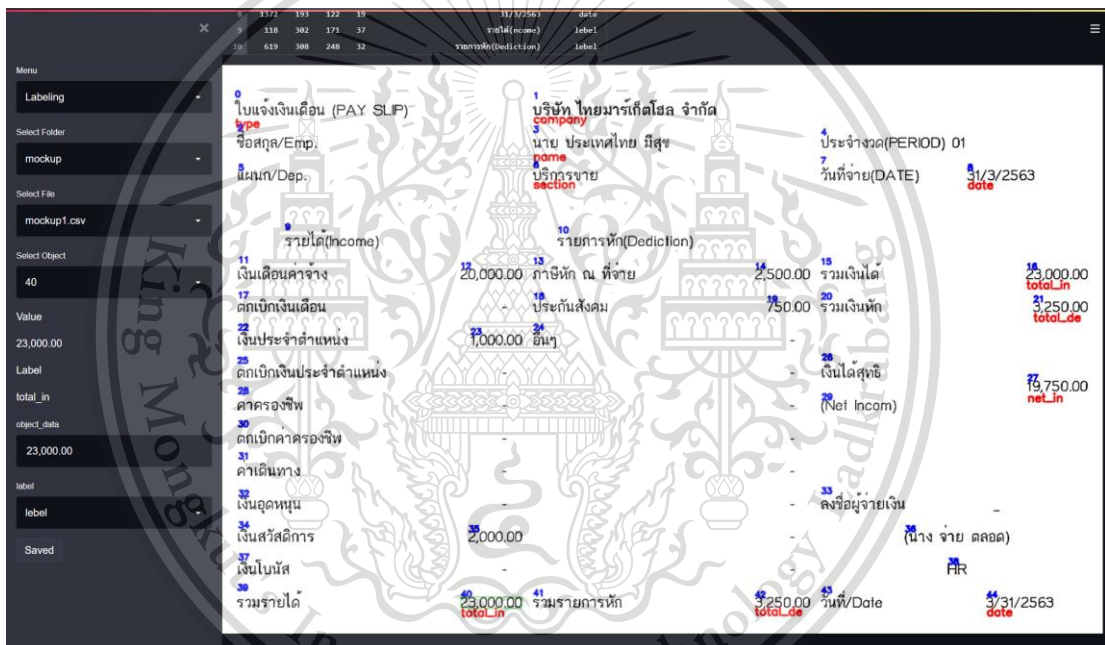
เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้เพื่อการเรียนการสอนเท่านั้น มิใช่เพื่อเผยแพร่โดยไม่ได้รับอนุญาต การค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมีเหตุดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.4.2 การเตรียมชุดข้อมูลฝึกและข้อมูลทดสอบ

การเตรียมชุดข้อมูลฝึกและข้อมูลทดสอบนั้นจะใช้รูปภาพของสลิปเงินเดือนทั้งหมด 119 รูปภาพ โดยประกอบไปด้วยชุดข้อมูลฝึกเป็น 80 รูปภาพ และชุดข้อมูลทดสอบ 39 รูปภาพ โดยที่ข้อมูลที่ต้องการดึงออกมาจากสลิปเงินเดือนนั้นมีทั้งหมด 9 tags ได้แก่ ชื่อ-นามสกุล ประเภทแผนก ชื่อบริษัท วันที่ จำนวนเงินได้ทั้งหมด จำนวนเงินหักทั้งหมด เงินได้สุทธิ และ null โดยที่ null นั้นจะเป็นคลาสหรือ tag ที่เราไม่สนใจ

3.4.2.1 การ Label ข้อมูลในแต่ละรูปภาพ

โดยจะทำการเขียนฟังก์ชันในการทำ Label ข้อมูลโดยใช้ Library ของ Streamlit โดยทำการบันทึกข้อมูลในแต่ละ text boxes นั้นเป็น Label อะไร เช่น ในรูปภาพที่ 3.21 นั้นได้มีการ Label ไว้ว่า text boxes ที่ 1 นั้นเป็น tag ของ ชื่อบริษัท



รูปภาพที่ 3.21 แสดงตัวอย่างของโปรแกรม Label

3.4.3 การเลือกคุณสมบัติและการแปลงข้อมูล

3.4.3.1 การเลือกคุณสมบัติ

ในการเลือกคุณสมบัติของข้อมูลนั้นเนื่องจากข้อมูลที่ใช้การฝึกฝนและทดสอบนั้นมีไม่มากจึงจะต้องมีการเก็บกลุ่มของประโยคและคำที่มีความหมายเหมือนกันแต่รูปแบบการเขียนนั้นแตกต่างกันเช่น “รวมรายการหัก” กับ “รวมรายจ่าย” ให้เป็นรูปแบบประโยคเดียวกันคือ “รวมรายการหัก” โดยจะทำการเก็บประโยคที่มีความเหมือนกันไว้ในรูปแบบ data type แบบ dictionary ของไพทอนโดยจะทำการตรวจสอบและแก้ไขประโยคก่อนแล้วจากนั้นจะทำการตรวจสอบกับประโยคที่เข้ามาว่ามีอยู่ในข้อมูลที่เก็บไว้ข้างต้นหรือไม่หลังจากนั้นจะทำการตัดคำโดยใช้

library ของ pythainlp engine newmm เป็นตัว tokenization และทำการนำตัวอักษรพิเศษต่าง ๆ ออกจากชุดข้อมูลเช่น ! \ ? และจากนั้นทำการตรวจสอบและแก้ไขคำอีกครั้งหนึ่งจากชุดข้อมูลทั้งหมด 119 ข้อมูล และทำการเลือกคำที่จะมาใช้เป็นคุณสมบัติของแต่ละโหนดโดยจะนำคำทั้งหมดที่มีในข้อมูลยกเว้นคำที่เป็นเอกลักษณ์เช่น ชื่อของคน หรือชื่อบริษัท มาใช้ซึ่งจะได้คำศัพท์ขนาด 913 คำที่ไม่ซ้ำกันทั้งหมดมาเป็นคุณสมบัติของข้อมูลทั้งหมดและจะเพิ่มอีก 2 คุณลักษณะเข้าไปสำหรับตรวจสอบว่าเป็นค่าจำนวนตัวเลขที่มี ลูกน้ำ หรือ เป็นค่าจำนวนตัวเลขที่มี / อยู่ด้วย เพื่อใช้ในการตรวจสอบว่าเป็นจำนวนเงินหรือเป็นวันที่

เงินได้สะสม
['เงินได้', 'สะสม']
ภาษีสะสม
['ภาษี', 'สะสม']
ประกันสังคม
['ประกันสังคม']
สำรองเลี้ยงชีพ
['สำรอง', 'เลี้ยงชีพ']

รูปภาพที่ 3.22 ภาพแสดงตัวอย่างการตัดคำ

3.4.3.2 การแปลงข้อมูลเพื่อเข้าสู่การฝึกฝนและทดสอบ

หลังจากทำการสร้างกราฟแบบในหัวข้อที่ 3.4.1 และนั่นจะได้ความสัมพันธ์ของแต่ละโหนดในรูปแบบของกราฟและการแปลงข้อมูลจะทำการแปลงออกเป็นสองไฟล์โดยที่ไฟล์หนึ่งจะทำการเก็บข้อมูลโหนดที่ติดกันในข้อมูลซึ่งจะโดยจะให้มีนามสกุลไฟล์เป็น “.cites” โดยที่จะทำการเก็บเป็นคู่ของหมายเลขโหนดที่ติดกัน และอีกไฟล์จะทำการเก็บข้อมูลคุณสมบัติของแต่ละโหนดโดยจะมีนามสกุลไฟล์เป็น “.content” โดยจะทำการแปลงข้อมูลด้วยวิธีการ One-Hot Encoding คือการทำข้อมูลภายในแต่ละโหนดจากข้อมูลประเภทข้อความในรูปแบบไบนารีหรือมีค่าแค่ 0 กับ 1 โดยจะดูจากคุณสมบัติที่ได้ทำการคัดเลือกตาม 3.3.3.1 ซึ่งจะได้มาเป็นคุณสมบัติโดยจะทำการแปลงโหนดที่มีคุณสมบัติตรงกับที่ได้บอกไว้ใน 3.3.3.1 นั้นจะให้ค่าเท่ากับหนึ่ง แต่ถ้าไม่ตรงก็จะให้คุณสมบัตินั้นมีค่าเท่ากับ 0

feature words = ['รายรับ', 'รายจ่าย', 'ค่าใช้จ่าย', 'รวม']

ประโยค One-Hot Encoding
รวมรายรับ = [1,0,0,1]
ค่าใช้จ่าย = [0,0,1,0]

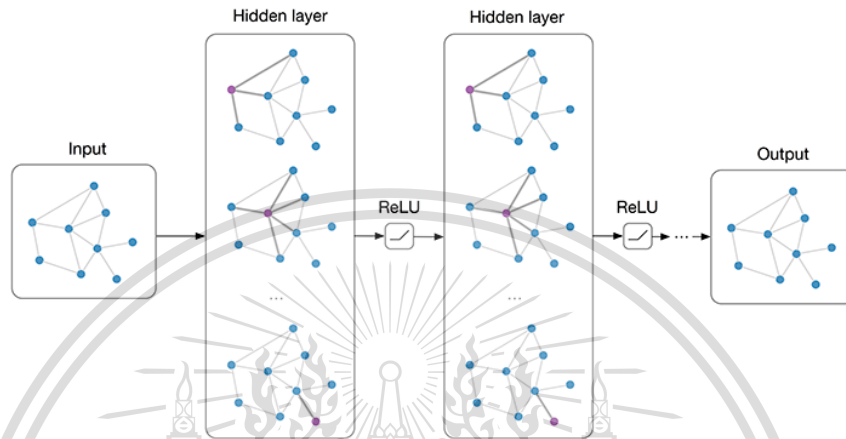
รูปภาพที่ 3.23 ภาพแสดงตัวอย่างการแปลงข้อมูล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

3.4.4 การฝึกฝนโมเดล

ทำการฝึกฝนโมเดลโดยจะใช้ Graph Convolutional Network (node classification) ซึ่งเป็น Semi-Supervised Learning รูปแบบหนึ่ง โดยมีเลเยอร์และ activation function ต่าง ๆ ตามรูปภาพที่ 3.24



รูปภาพที่ 3.24 ภาพแสดง Layer และ Activation Function
ที่มา: tkipf/pygcn (2017)

โดยที่ Input layer นั้นจะเป็นประกอบด้วย feature ของแต่ละ node และ adjacent node ที่อยู่ใกล้กันตามหัวข้อที่ 3.3.1 ในส่วนของ hidden layer นั้นจะใช้จำนวน 2 hidden layer ประกอบไปด้วยเลเยอร์ละ 16 node และใช้ ReLU เป็น activation function และในส่วนของ Output Layer นั้นจะใช้ softmax เป็น activation function โดยจะได้ออกมาเป็นในรูปแบบของความน่าจะเป็น โดยจะมีการปรับค่า hyperparameter ต่างๆ ดังตารางที่ 3.1

ตารางที่ 3.1 ตารางแสดงค่า Hyperparameter ที่ใช้ในการตั้งค่า model

Parameter	Value
Learning Rate	0.01
Epochs	400
Weight decay	5e-05
Dropout	0.25

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการฝึกฝนข้อมูลนั้นจะใช้ Adam Optimizer และใช้ Cross-Entropy เป็น Loss Function และทำการฝึกชุดข้อมูลตาม epochs ที่ได้ทำการตั้งค่าไว้ตาม hyperparameter

This material is reserved for educational use only, not allowed for commercial use.

บทที่ 4

ผลการวิจัย

4.1 ผลการทดสอบการทำงานของหน้าเว็บแอปพลิเคชัน

ในส่วนการของการทดสอบการทำงานของหน้าเว็บแอปพลิเคชันจะเป็นการทดสอบการทำงานต่าง ๆ ที่ได้ถูกออกแบบไว้สามารถทำงานได้ โดยการทดสอบจะมีหน้าผู้ใช้งานดังนี้

4.1.1 หน้าเข้าสู่ระบบ

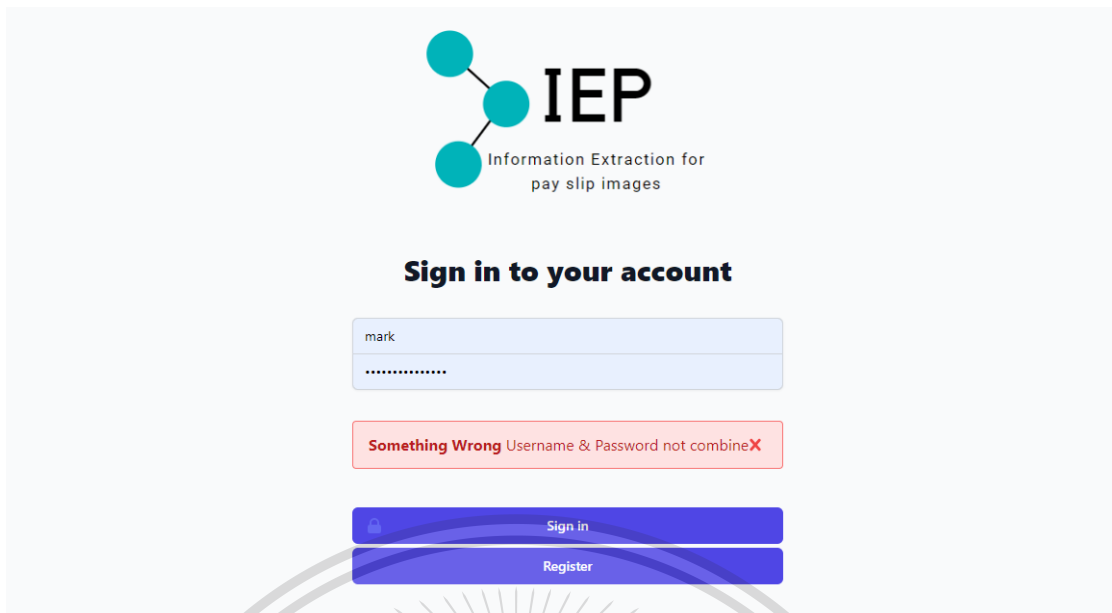


รูปภาพที่ 4.1 รูปแสดงหน้าเข้าสู่ระบบ

จากรูปที่ 4.1 เป็นหน้าที่ให้ผู้ใช้กรอกชื่อผู้ใช้งาน และรหัสผ่านเพื่อทำการยืนยันตัวตนและสิทธิในการเข้าใช้หน้าเว็บแอปพลิเคชันในส่วนถัดไปได้ โดยถ้าผู้ใช้ไม่กรอกชื่อผู้ใช้งาน และรหัสผ่าน หรือทำการกรอกไม่ถูกต้อง ระบบจะทำการแสดงข้อความแจ้งเตือนโดยจะเป็นดังรูปที่ 4.2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.



รูปภาพที่ 4.2 รูปแสดงการตรวจสอบของหน้าเข้าสู่ระบบ

4.1.2 หน้าลงทะเบียน



รูปภาพที่ 4.3 รูปแสดงหน้าลงทะเบียน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.



Register in to your account

Username
Password
Password

username This field may not be blank.

password This field may not be blank.

password2 This field may not be blank.

Register

รูปภาพที่ 4.4 รูปแสดงหน้าลงทะเบียน

จากรูปภาพที่ 4.3 จะเป็นหน้าที่ผู้ใช้งานจะต้องทำการกรอกข้อมูลชื่อผู้ใช้งานและรหัสผ่านทั้งสองครั้งให้ตรงกันโดยจะเมื่อทำการกดปุ่ม “Register” แล้วระบบจะทำการเรียกใช้งานเซิร์ฟเวอร์แล้วทำการตรวจสอบโดยจะแสดงข้อความที่ไม่ผ่านเงื่อนไขต่าง ๆ ดังรูปภาพที่ 4.4 จะเป็นรูปภาพตัวอย่างที่แสดงข้อความเมื่อผู้ใช้งานทำการกรอกข้อมูลแล้วไม่ได้ทำการกรอกข้อมูลช่องใดช่องหนึ่งโดยจะบอกให้ผู้ใช้งานรู้ว่าจำเป็นต้องกรอกช่องใดบ้าง

4.1.3 หน้าหลัก



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และเปลี่ยนแปลงของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

จากรูปภาพ 4.5 จะเป็นรูปแสดงหน้าหลักหลังจากทำการเข้าสู่ระบบเรียบร้อยแล้ว โดยที่หน้านี้จะแสดงข้อมูลรายการที่ผู้ใช้งานเคยทำการอัปโหลดภาพสลิปเงินเดือนไว้ โดยแต่ละรายการที่แสดง จะประกอบไปด้วย “Status” คือสถานะของการดึงข้อมูลโดยใช้โมเดลการทำนายผล ว่ามีการตรวจสอบข้อมูลแล้วหรือยังถ้ายังจะขึ้นสถานะว่า Unverified แต่ถ้าทำการตรวจสอบและยืนยันแล้วจะขึ้นว่า Verified และ “Task Status” คือการแสดงสถานะของการใช้งานการดึงข้อมูลว่าสำเร็จหรือไม่ “Create at” คือแสดงวันที่ทำการอัปโหลดรูปภาพ และ “Export To” ซึ่งผู้ใช้งานสามารถทำการส่งออกข้อมูลที่ทำการดึงข้อมูลด้วยโมเดลทำนายผลแล้วออกเป็นไฟล์เอกสารในรูปแบบของ CSV ได้ และผู้ใช้งานสามารถกดที่รายแต่ละรายการเพื่อเข้าไปดูผลลัพธ์ของแต่ละรายการได้

4.1.4 หน้าอัปโหลดรูปภาพสลิปเงินเดือน

ในหน้าอัปโหลดรูปภาพสลิปเงินเดือนนั้นเป็นหน้าที่ให้ผู้ใช้งานทำการเลือกรูปภาพสลิปเงินเดือนที่ต้องการทำการดึงข้อมูลโดยทำการกดที่ “SELECT A FILE” โดยผู้ใช้งานจะสามารถทำการอัปโหลดไฟล์ได้เฉพาะไฟล์รูปภาพเท่านั้น โดยจะเป็นดังรูปที่ 4.6



รูปภาพที่ 4.6 รูปแสดงหน้าอัปโหลดรูปภาพ

ซึ่งผู้ใช้งานสามารถทำการเลือกรูปภาพสลิปเงินเดือนได้หลาย ๆ รูปภายในครั้งเดียว โดยจะได้เป็นดังรูปที่ 4.7 และเมื่อผู้ใช้งานทำการเลือกรูปภาพเสร็จแล้วสามารถทำการกด “Upload” เพื่อเรียกใช้งานเซิร์ฟเวอร์ดึงข้อมูลได้หรือถ้าผู้ใช้งานต้องการทำการลบรูปภาพที่เลือกเอาไว้สามารถทำการกดปุ่ม “Clear All” เพื่อทำการลบรูปภาพที่เลือกเอาไว้ทั้งหมดได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

ใบแจ้งรายได้ PAY SLIP

บริษัท ทรานส์ เทคโนโลยี จำกัด
81/2 เซ็นทรัลพลาซ่า ชั้น 10 อาคาร ทรานส์เทคโนโลยี 10230
Tel: 0-2456-8010 Fax: 02-0951-6981
เลขประจำตัวผู้เสียภาษี 3152361100000

นาย พงษ์ ใจดีสร้าง
ตำแหน่ง: Project Manager
แผนก: Management
วันที่: 31/01/2025
บัญชีเงินฝาก: 120-4156-11

รายการเงินได้ INCOME	จำนวนเงิน AMOUNT	รายการเงินหัก DEDUCTION	จำนวนเงิน AMOUNT
เงินเดือน/ค่าจ้าง Salary/Wage	37,000.00	ภาษี / Tax ประกันสังคม / Social Security	3,400.00 1,500.00
รวมเงินได้ Total Income		รวมเงินหัก Total Deduction	
37,000.00		3,400.00	
เงินได้สุทธิ Net Income		33,600.00	

Label verify all

- company
- date: 31/01/2025
- name: นาย พงษ์ ใจดีสร้าง
- net_in: 32,100.00
- section: Management
- total_de: 4,900.00
- total_lin: 37,000.00
- type: ใบแจ้งรายได้ PAY SLIP

Save

รูปภาพที่ 4.9 รูปแสดงหน้าแสดงผลลัพธ์

เมื่อผู้ใช้งานทำการแก้ไขข้อมูลในแต่ละช่องเรียบร้อยแล้วสามารถที่จะเลือกช่องที่ทำการแก้ไขหรือตรวจสอบแล้วข้อมูลนั้นถูกต้องว่า Verified แล้ว หรือสามารถทำการเลือกทั้งหมดพร้อมกันที่เดียวก็ได้ ซึ่งเมื่อต้องการที่จะบันทึกผลที่ทำการแก้ไขแล้วนั้นผู้ใช้งานสามารถทำการกดปุ่ม “Save” เพื่อทำการบันทึกได้

4.2 ผลการทดสอบการทำงานของเซอร์วิส

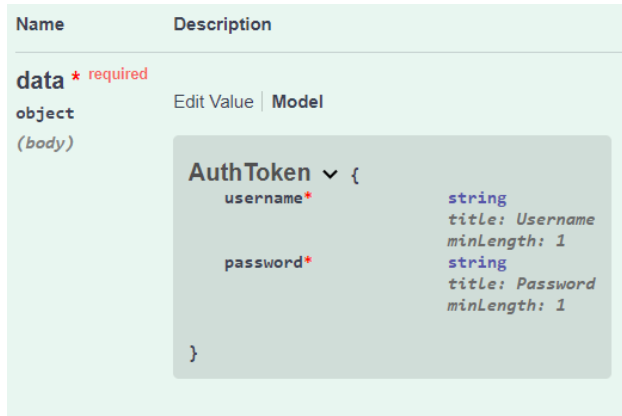
ในส่วนการทดสอบเซอร์วิสนั้นจะเซอร์วิสต่าง ๆ ที่ทดสอบนั้นจะทำการทดสอบด้วยการใช้ Swagger โดยจะใช้งานผ่านเว็บเบราว์เซอร์ได้เลย โดย Swagger ที่ผู้จัดทำได้ทำขึ้นนั้นจะเป็น API Document ที่สามารถทำการทดลองการใช้เซอร์วิสได้เลยโดยจะมีรายละเอียดต่าง ๆ ของแต่ละเซอร์วิสบอกไว้อย่างละเอียดว่าแต่ละเซอร์วิสนั้นจะต้องใช้ข้อมูลอะไรและรูปแบบใด โดยจะมีการทดสอบเซอร์วิสดังต่อไปนี้

4.2.1 การตรวจสอบและยืนยันตัวตนการเข้าถึงหน้าเว็บแอปพลิเคชัน

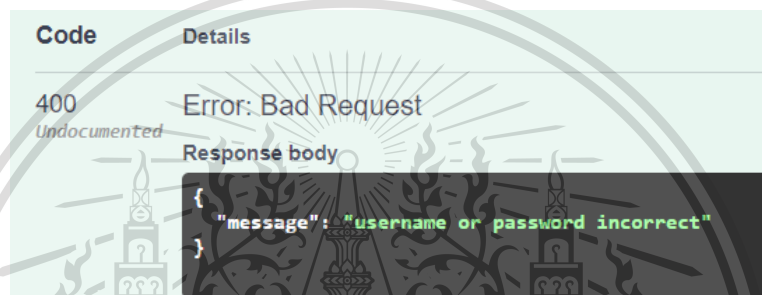
ในเซอร์วิสนี้จะทำการรับค่าชื่อผู้ใช้และรหัสผ่านจากทางหน้าเว็บแอปพลิเคชันที่ทำการกรอกข้อมูลในหน้าเข้าสู่ระบบโดยข้อมูลและรูปแบบข้อมูลที่เซอร์วิสที่ต้องการจะเป็นดังรูปภาพที่ 4.10

โดยเมื่อมีการเรียกใช้งานเซอร์วิสและผู้ใช้งานทำการกรอก ชื่อผู้ใช้และรหัสผ่านไม่ถูกต้องนั้นเซอร์วิสจะทำการตอบกลับแบบดังรูปภาพที่ 4.11 แต่ถ้าทำการกรอกข้อมูลได้ถูกต้อง เซอร์วิสจะทำการตอบกลับพร้อมกับ Access Token กลับไปเพื่อใช้ในการยืนยันเข้าใช้งานส่วนต่าง ๆ ของเว็บแอปพลิเคชันต่อไปได้ โดยจะเป็นดังรูปภาพที่ 4.12

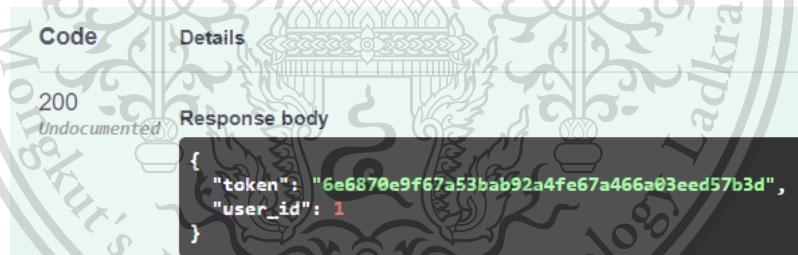
เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้สำหรับการใช้งานเพื่อการศึกษาค้นคว้าเท่านั้น เมื่อผู้ผู้เห็นไปใช้ประโยชน์ทางการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมีเหตุดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปภาพที่ 4.10 รูปภาพแสดงข้อมูลที่เซอริสการยืนยันตัวตนต้องการ



รูปภาพที่ 4.11 รูปภาพตัวอย่างการตอบกลับของเซอริสการยืนยันตัวตน



รูปภาพที่ 4.12 รูปภาพตัวอย่างการตอบกลับของเซอริสการยืนยันตัวตน

4.2.2 การลงทะเบียน

เซอริสนี้จะทำการรับข้อมูลจากหน้าลงทะเบียนในเว็บแอปพลิเคชันโดยจะรับ ชื่อ ผู้ใช้งาน รหัสผ่านและรหัสผ่านที่ยืนยันโดยจะต้องการข้อมูลดังรูปภาพที่ 4.13

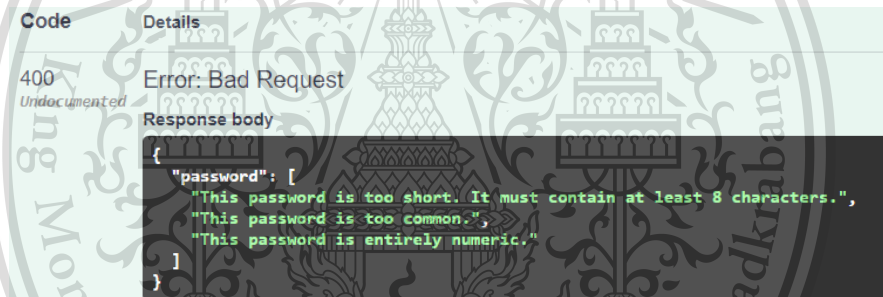
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

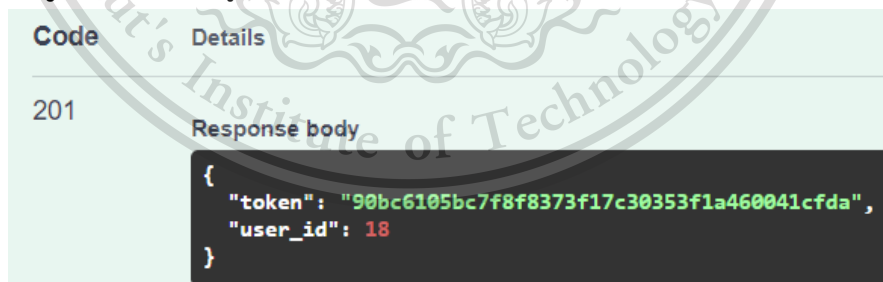


รูปภาพที่ 4.13 รูปภาพแสดงข้อมูลที่เซอวีส์การลงทะเบียนต้องการ

เมื่อถูกเรียกใช้ตัวเซอวีส์จะทำการตรวจสอบข้อมูลที่รับเข้ามาโดยถ้าชื่อผู้ใช้งานหรือรหัสผ่านที่ ผู้ใช้งานทำการกรอกผ่านหน้าลงทะเบียนนั้นไม่ถูกต้องตามเงื่อนไขที่กำหนด เซอวีส์จะมีการตอบกลับไปว่าผู้ใช้งานทำการกรอกข้อมูลใดไม่ถูกต้อง โดยจากรูปภาพที่ 4.14 นั้นจะเป็นตัวอย่างการตอบกลับของเซอวีส์ในกรณีที่ผู้ใช้งานกรอกรหัสผ่านที่ง่ายเกินไป แต่ถ้าผู้ใช้งานทำการกรอกข้อมูลได้ถูกต้องเซอวีส์จะทำการตอบกลับดังรูปภาพที่ 4.15



รูปภาพที่ 4.14 รูปภาพแสดงการตอบกลับของเซอวีส์ลงทะเบียนไม่สำเร็จ



รูปภาพที่ 4.15 รูปภาพแสดงการตอบกลับของเซอวีส์ลงทะเบียนสำเร็จ

4.2.3 การอัปเดตรูปภาพและการดึงข้อมูลฟิลต์โดยใช้โมเดลการทำนายผล

เซอวีส์นี้จะรับรูปภาพสลิปเงินเดือนที่ผู้ใช้งานอัปเดตผ่านทางหน้าเว็บแอปพลิเคชัน

ขั้นแล้วทำการเรียกใช้การดึงข้อมูลด้วยโมเดลทำนายผลโดยเซอวีส์จะทำการถ้าผู้ใช้งานไม่ได้ทำการยืนยันตัวด้วย Access Token ที่จะได้รับมาจากการเรียกใช้เซอวีส์เข้าสู่ระบบเซอวีส์จะตอบกลับดังรูปภาพที่ 4.16

This material is reserved for educational use only, not allowed for commercial use.

Code	Details
401 <i>Undocumented</i>	Error: Unauthorized Response body <pre>{ "detail": "Authentication credentials were not provided." }</pre>

รูปภาพที่ 4.16 รูปภาพแสดงการตอบกลับของเซอรัวิสการอัปโหลดรูปภาพ

และถ้าสามารถทำการอัปโหลดและเรียกใช้งานการดึงข้อมูลได้สำเร็จเซอรัวิสจะทำการตอบกลับดังรูปภาพที่ 4.17

Code	Details
200 <i>Undocumented</i>	Response body <pre>{ "message": "success" }</pre>

รูปภาพที่ 4.17 รูปภาพแสดงการตอบกลับของเซอรัวิสการอัปโหลดรูปภาพ

4.2.4 แสดงข้อมูลรายการที่เรียกใช้การอัปโหลดรูปภาพและการดึงข้อมูลฟิลต์

เซอรัวิสที่จะทำการคิวรีข้อมูลรายการที่ผู้ใช้งานได้เคยทำการเรียกใช้เซอรัวิสการอัปโหลดรูปภาพสลิปเงินและการดึงข้อมูลไว้ในของแต่ละผู้ใช้งาน ดังรูปภาพที่ 4.18

Code	Details
200	Response body <pre>{ "count": 4, "next": "http://127.0.0.1:8000/api/list/?page=2", "previous": null, "results": [{ "id": 41, "image": "storage/images/mockup8_I0#xenL.png", "status": "unverified", "status_task": "SC", "create_at": "05 May, 2021 08:51:23" }, { "id": 7, "image": "storage/images/mockup6_5PwVYXJ.png", "status": "unverified", "status_task": "SC", "create_at": "05 May, 2021 08:51:44" }, { "id": 39, "image": "storage/images/mockup5_CoauQi7.png", "status": "unverified", "status_task": "SC", "create_at": "05 May, 2021 08:52:06" }] }</pre>

รูปภาพที่ 4.18 รูปภาพแสดงการตอบกลับของเซอรัวิสการอัปโหลดรูปภาพ

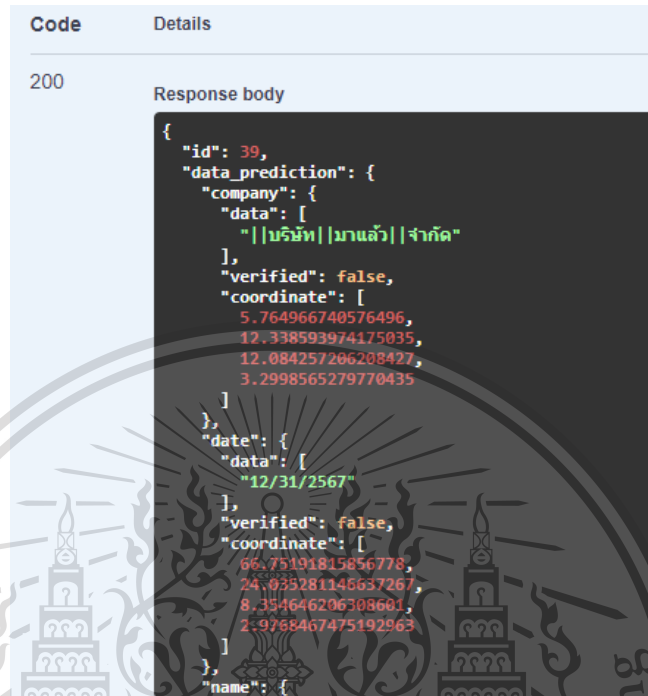
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งาน ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

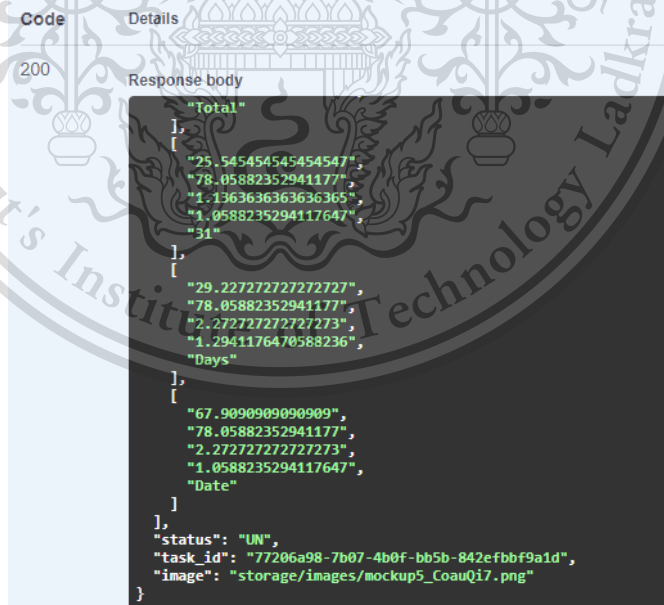
4.2.5 แสดงผลลัพธ์การดึงข้อมูลฟิลต์โดยใช้โมเดลการทำนายผลและแก้ไขข้อมูล

ในส่วนของการเรียกเซิร์ฟวิสนี้จะอ้างอิงจากเลข id ของรายการอัปโหลดรูปภาพสลิปเงินที่ผู้ใช้งานได้ทำการเลือกในหน้าหลัก ดังรูปภาพที่ 4.19 และ 4.20



```
Code    Details
200     Response body
{
  "id": 39,
  "data_prediction": {
    "company": {
      "data": [
        "|บริษัท|มาแล้ว|จำกัด"
      ],
      "verified": false,
      "coordinate": [
        5.764966740576496,
        12.338593974175035,
        12.084257266208427,
        3.2998565279770435
      ]
    },
    "date": {
      "data": [
        "12/31/2567"
      ],
      "verified": false,
      "coordinate": [
        66.75191815856778,
        24.035281146637267,
        8.354646206308601,
        2.9768467475192963
      ]
    }
  },
  "name": {}
}
```

รูปภาพที่ 4.19 รูปภาพแสดงการตอบกลับของเซิร์ฟวิสนี้ผลลัพธ์การดึงข้อมูล



```
Code    Details
200     Response body
{
  "Total": [
    [
      "25.545454545454547",
      "78.05882352941177",
      "1.1363636363636365",
      "1.0588235294117647",
      "31"
    ],
    [
      "29.227272727272727",
      "78.05882352941177",
      "2.272727272727273",
      "1.2941176470588236",
      "Days"
    ],
    [
      "67.909090909090909",
      "78.05882352941177",
      "2.272727272727273",
      "1.0588235294117647",
      "Date"
    ]
  ],
  "status": "UN",
  "task_id": "77206a98-7b07-4b0f-bb5b-842efbbf9a1d",
  "image": "storage/images/mockup5_CoauQi7.png"
}
```

รูปภาพที่ 4.20 รูปภาพแสดงการตอบกลับของเซิร์ฟวิสนี้ผลลัพธ์การดึงข้อมูล

โดยจะประกอบไปด้วยผลข้อมูลที่ทำกรเรียกใช้โมเดลทำนายผลออกมาได้ทั้ง 8 tags และข้อมูลที่ได้ออกมาจากการทำ OCR โดยในเซิร์ฟวิสนี้จะสามารถทำการ PATCH เพื่อที่จะทำการแก้ไขข้อมูลได้อีกด้วย

	precision	recall	f1-score	support
company	0.87	0.95	0.91	142
date	0.92	0.87	0.90	140
label	0.96	0.99	0.98	8005
name	0.93	0.72	0.81	143
net_in	0.78	0.56	0.65	123
section	0.93	0.54	0.68	117
total_de	0.90	0.58	0.70	123
total_in	0.88	0.40	0.55	123
type	0.96	0.93	0.94	99
accuracy			0.96	9015
macro avg	0.90	0.73	0.79	9015
weighted avg	0.96	0.96	0.95	9015

รูปภาพที่ 4.23 รูปภาพแสดงค่า precision, recall and f1-score ของการฝึกฝนในแต่ละคลาส

	precision	recall	f1-score	support
company	0.86	0.96	0.91	45
date	0.89	0.87	0.88	47
label	0.96	0.99	0.97	2608
name	0.89	0.63	0.74	49
net_in	0.76	0.47	0.58	40
section	0.87	0.51	0.65	39
total_de	0.83	0.50	0.62	40
total_in	0.75	0.23	0.35	40
type	1.00	0.91	0.96	35
accuracy			0.95	2943
macro avg	0.87	0.68	0.74	2943
weighted avg	0.95	0.95	0.94	2943

รูปภาพที่ 4.24 รูปภาพแสดงค่า precision, recall and f1-score ของการทดสอบในแต่ละคลาส

ตารางที่ 4.1 ตารางแสดง Confusion Matrix ของการฝึกฝนโมเดล

	company	date	label	name	net_in	section	total_de	total_in	type	total
company	135	0	5	0	0	0	0	0	2	142
date	0	122	18	0	0	0	0	0	0	140
label	20	9	7937	8	11	4	8	6	2	8005
net_in	0	0	39	103	0	1	0	0	0	143
section	0	0	54	0	69	0	0	0	0	123
total_de	0	1	53	0	0	63	0	0	0	117
total_in	0	0	43	0	8	0	71	1	0	123
type	0	0	74	0	0	0	0	49	0	123
total	0	0	7	0	0	0	0	0	92	99

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับใช้ภายในเท่านั้น ไม่อนุญาตให้ไปใช้ประโยชน์ด้วยวิธีการ

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

ตารางที่ 4.2 ตารางแสดง Confusion Matrix ของการทดสอบโมเดล

	company	date	lebel	name	net_in	section	total_de	total_in	type	total
company	43	0	2	0	0	0	0	0	0	45
date	0	41	6	0	0	0	0	0	0	47
lebel	7	4	2580	4	3	3	4	3	0	2608
net_in	0	0	18	31	0	0	0	0	0	49
section	0	0	21	0	19	0	0	0	0	40
total_de	0	1	18	0	0	20	0	0	0	39
total_in	0	0	17	0	3	0	20	0	0	40
type	0	0	31	0	0	0	0	9	0	40
total	0	0	3	0	0	0	0	0	32	35

4.4.1 ผลเปรียบเทียบการเลือกคุณสมบัติของข้อมูล

ในการทำการเลือกค่าที่จะเป็นคุณสมบัติของข้อมูลนั้นได้ทำการเปรียบเทียบการเลือกค่าโดยจะทำการเลือกใช้ทุกค่าที่มีในข้อมูลกับใช้ค่าที่มีความถี่หรือค่าซ้ำกันมากกว่าเท่ากับ 2 และ 3 โดยจะทำการเปรียบเทียบทั้งในการฝึกฝนและทดสอบโมเดล โดยจะทำการกำหนดตัวแปรดังต่อไปนี้

- เลือกใช้ทุกค่าที่มีในข้อมูลยกเว้นค่าที่มีความเป็นเอกลักษณ์เช่นชื่อหรือชื่อบริษัท เป็นตัวแปร "A"
- เลือกใช้ค่าซ้ำที่มากกว่าเท่ากับ 2 เป็นตัวแปร "B"

ตารางที่ 4.3 ตารางเปรียบเทียบค่า Precision, Recall และ F1-Score ของการฝึกโมเดล

Class	Precision		Recall		F1-Score		Support
	A	B	A	B	A	B	
Company	0.87	0.85	0.95	0.91	0.91	0.88	142
Date	0.92	0.88	0.87	0.79	0.90	0.83	140
Name	0.93	0.94	0.72	0.75	0.91	0.83	143
Net Income	0.78	0.75	0.56	0.50	0.65	0.60	123

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษานั้น ไม่อนุญาตให้ไปใช้ ระเบียบด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

Section	0.93	0.99	0.54	0.59	0.68	0.74	117
Total Deduction	0.90	0.86	0.58	0.50	0.70	0.64	123
Total Income	0.88	0.75	0.40	0.20	0.55	0.31	123
Type	0.96	0.99	0.93	0.93	0.94	0.90	99
Label (Null)	0.96	0.96	0.99	0.99	0.98	0.97	8005
accuracy					0.96	0.95	9015
macro average	0.90	0.87	0.73	0.68	0.79	0.75	9015
weighted average	0.96	0.95	0.96	0.95	0.95	0.95	9015

ตารางที่ 4.4 ตารางเปรียบเทียบค่า Precision, Recall และ F1-Score ของการทดสอบโมเดล

Class	Precision		Recall		F1-Score		Support
	A	B	A	B	A	B	
Company	0.86	0.78	0.96	0.89	0.91	0.83	45
Date	0.89	0.82	0.87	0.70	0.88	0.76	48
Name	0.89	0.92	0.63	0.71	0.74	0.80	49
Net Income	0.76	0.76	0.47	0.40	0.58	0.52	40
Section	0.87	1.00	0.51	0.49	0.65	0.66	39
Total Deduction	0.83	0.76	0.50	0.47	0.62	0.58	40
Total Income	0.75	0.60	0.23	0.15	0.35	0.24	40
Type	1.00	0.89	0.91	0.91	0.96	0.90	35
Label (Null)	0.96	0.95	0.99	0.99	0.97	0.97	2608
accuracy					0.95	0.94	2943
macro average	0.87	0.83	0.68	0.64	0.74	0.70	2943
weighted average	0.95	0.94	0.95	0.94	0.94	0.93	2943

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่ออนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

จากตารางที่ 4.1 และ 4.2 จะสังเกตได้ว่าข้อมูลนั้นมีลักษณะที่เป็น Imbalance Classes หรือการที่มีคลาสในคลาสหนึ่งมากกว่าคลาสอื่นมาก ๆ และคลาสที่เราสนใจนั้นมีน้อยดังนั้น จึงจะใช้ค่าของ F1-Score แบบ Marco Average นั้นจะสะท้อนถึงความแม่นยำของโมเดลได้ดีที่สุด เพราะจะ Marco นั้นจะสะท้อนความแม่นยำของคลาสที่มีจำนวนน้อย ๆ หรือก็คือคลาสที่เราสนใจ

ซึ่งเมื่อทำการเปรียบเทียบ F1-Score Marco Average ของตัวแปร A, B นั้นจะได้สรุปได้ว่าตัวแปร A หรือก็คือค่าที่มีในข้อมูลยกเว้นค่าที่มีความเป็นเอกลักษณ์เช่นชื่อหรือชื่อบริษัท นั้นจะมีความแม่นยำมากที่สุดซึ่งจะมีค่าเท่ากับ 0.74

4.5 ผลการเรียกใช้การดึงข้อมูลด้วยโมเดลการทำนายผลผลการปรับ Thresholding และการลบเส้นตารางในรูปภาพสลิปเงินเดือน

ในส่วนของการทำ Thresholding และลบเส้นตารางในของรูปภาพนั้นจะถูกเรียกใช้งานก่อนที่จะนำรูปภาพสลิปเงินเดือนเข้าสู่การดึงข้อมูลด้วยโมเดลการทำนายผลซึ่งผลลัพธ์ของการปรับ Threshold จะเป็นดังรูปที่ 4.23 และหลังจากทำการปรับ Threshold เรียบร้อยแล้วก็จะทำการลบเส้นตารางในรูปภาพออกซึ่งจะได้ผลลัพธ์ตามรูปที่ 4.24 โดยรูปภาพตัวอย่างที่นำเข้าจะเป็นดังรูปภาพที่ 4.22

รายได้ Income	รายหัก Deduction	ชม. ล่วงเวลา Over Time Hours
เงินเดือน 35,000	ภาษี 3,000	ชม./นาที ต่อชม. จำนวนเงิน
ค่าล่วงเวลา 2,000	OT 750	
เบี้ยขยันโอที -		ชม.การลา Over Time Hours
ค่าอาหาร -		การลา ชม./นาที ต่อชม. จำนวนเงิน
เบี้ยขยัน -		
ค่าตำแหน่ง -		
ค่ารักษา -		
ค่าจ้างงาน -		
ค่าปิ่นรถ -		
อื่นๆ -		
รวมเงินได้ Total Income 37,000	รวมรายหัก Total Deduction 3,750	เงินได้สุทธิ Net Income 33,250

รูปภาพที่ 4.22 รูปภาพแสดงตัวอย่างรูปภาพก่อนทำการปรับ Threshold

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

บริษัท มาแล้ว จำกัด
 81/2 เลี้ยวไทย เขต ดินนาถาว แขวง ดินนาถาว กรุงเทพฯ 10230
 Tel 0-2456-9010 Fax. 02-0951-6901
 เลขประจำตัวผู้เสียภาษี 3152361100000

ใบแจ้งรับเงินเดือน
 Payroll Slip

วันที่จ่ายเงิน Pay Date 12/31/2567
 งวดที่ Period No. ประจำเดือน ธันวาคม

รหัสพนักงาน Code 125-C12 ชื่อ Name นาย สมพร กว้างใหญ่
 แผนก Section ช่างซ่อมบำรุง
 ประเภทพนักงาน ประจำ

รายได้ Income		รายหัก Deduction		ชม.ล่วงเวลา Over Time Hours			
เงินเดือน	35,000	ภาษี	3,000	OT	ชม./นาที	ต่อชม.	จำนวนเงิน
ค่าล่วงเวลา	2,000	ประกันสังคม	750				
เบี้ยขยันโอที	-						
ค่าอาหาร	-						
ค่าประกัน	-						
เบี้ยขยัน	-						
ค่าค่าแห่ง	-						
ค่าวิชา	-						
ค่าอายุงาน	-						
ค่าบำนาญ	-						
อื่นๆ	-						
รวมเงินได้ Total Income	37,000	รวมรายหัก Total Deduction	3,750	เงินได้สุทธิ Net Income			33,250

จำนวนวันที่คิดเงิน 31 วัน ลงชื่อ ผู้รับเงิน วันที่ 31 Days Total Date

รูปภาพที่ 4.23 รูปภาพแสดงการปรับ Threshold ของรูปภาพ

บริษัท มาแล้ว จำกัด
 81/2 เลี้ยวไทย เขต ดินนาถาว แขวง ดินนาถาว กรุงเทพฯ 10230
 Tel 0-2456-9010 Fax. 02-0951-6901
 เลขประจำตัวผู้เสียภาษี 3152361100000

ใบแจ้งรับเงินเดือน
 Payroll Slip

วันที่จ่ายเงิน Pay Date 12/31/2567
 งวดที่ Period No. ประจำเดือน ธันวาคม

รหัสพนักงาน Code 125-C12 ชื่อ Name นาย สมพร กว้างใหญ่
 แผนก Section ช่างซ่อมบำรุง
 ประเภทพนักงาน ประจำ

รายได้ Income		รายหัก Deduction		ชม.ล่วงเวลา Over Time Hours			
เงินเดือน	35,000	ภาษี	3,000	OT	ชม./นาที	ต่อชม.	จำนวนเงิน
ค่าล่วงเวลา	2,000	ประกันสังคม	750				
เบี้ยขยันโอที	-						
ค่าอาหาร	-						
ค่าประกัน	-						
เบี้ยขยัน	-						
ค่าค่าแห่ง	-						
ค่าวิชา	-						
ค่าอายุงาน	-						
ค่าบำนาญ	-						
อื่นๆ	-						
รวมเงินได้ Total Income	37,000	รวมรายหัก Total Deduction	3,750	เงินได้สุทธิ Net Income			33,250

จำนวนวันที่คิดเงิน 31 วัน ลงชื่อ ผู้รับเงิน วันที่ 31 Days Total Date

รูปภาพที่ 4.24 รูปภาพแสดงการลบเส้นตารางออกจากในรูปภาพ

4.6 ผลการเรียกใช้ OCR และการจัดรูปแบบ

ในการทำ OCR นั้นจะทำการใช้บริการของ Google Vision API โดยทำการแปลงรูปภาพเป็นรูปแบบของ Bytes Array ก่อนแล้วจึงทำการส่งไปที่ Vision API โดยจะได้ผลลัพธ์ดังรูปภาพที่

4.25

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

บริษัท อนาคต จำกัด
 81/2 เลี้ยวใหม่ เขต ดินนายาว แขวง ดินนายาว กรุงเทพฯ 10230
 โทร 02-456-9010 โทร 02-0951-6901
 เลขประจำตัวผู้เสียภาษี 3152361100000

ในแจ้งรับเงินเดือน
 Payroll Slip

วันที่จ่ายเงิน Pay Date 12/31/2567
 งวดที่ Period No. ประจำเดือน ธันวาคม

รหัสพนักงาน Code 125-C12 ชื่อ Name นาย สมุทร กว้างใหญ่
 แผนก Section 11 เลขชอมนำร่อง 12
 ประเภทพนักงาน 10 มร.จ่า

19 รายได้ Income	20 รายหัก Deduction	21 ชม.ล่วงเวลา Over Time Hours	22 ชม./นาที	23 ต่อชม.	24 จำนวนเงิน
เงินเดือน 85,000	ภาษี 3,000				
ค่าล่วงเวลา 2,000	ประกันสังคม 750				
เบี้ยขยัน 0					
ค่าอาหาร 0					
ค่าชดเชย 0					
เบี้ยขยัน 0					
ค่าตำแหน่ง 0					
ค่าเช่า 0					
ค่าจ่าย 0					
ค่าประกัน 0					
อื่นๆ 0					
รวมรายได้ Total Income 87,000	รวมรายหัก Total Deduction 3,750	เงินได้สุทธิ Net Income 83,250			

จำนวนวันที่คิดเงิน Total 31 Days
 วันที่ 31 ธันวาคม

รูปภาพที่ 4.25 รูปภาพแสดงผลลัพธ์การเรียกใช้ OCR

ซึ่งผลลัพธ์ที่ได้ดังนั้นยังไม่สามารถนำไปใช้งานได้จึงต้องทำการจัดรูปแบบใหม่ก่อนโดยจะทำการรวม word boxes ที่อยู่ใกล้กันให้เป็น text area เดียวกันแต่ถ้าเป็นในตัวเลขและมีค่าต่อด้านหลังจะไม่ทำการรวมโดยจะได้ดังรูปภาพที่ 4.26

บริษัท อนาคต จำกัด
 81/2 เลี้ยวใหม่ เขต ดินนายาว แขวง ดินนายาว กรุงเทพฯ 10230
 โทร 02-456-9010 โทร 02-0951-6901
 เลขประจำตัวผู้เสียภาษี 3152361100000

ในแจ้งรับเงินเดือน
 Payroll Slip

วันที่จ่ายเงิน Pay Date 12/31/2567
 งวดที่ Period No. ประจำเดือน ธันวาคม

รหัสพนักงาน Code 125-C12 ชื่อ Name นาย สมุทร กว้างใหญ่
 แผนก Section 11 เลขชอมนำร่อง 12
 ประเภทพนักงาน 10 มร.จ่า

19 รายได้ Income	20 รายหัก Deduction	21 ชม.ล่วงเวลา Over Time Hours	22 ชม./นาที	23 ต่อชม.	24 จำนวนเงิน
เงินเดือน 85,000	ภาษี 3,000				
ค่าล่วงเวลา 2,000	ประกันสังคม 750				
เบี้ยขยัน 0					
ค่าอาหาร 0					
ค่าชดเชย 0					
เบี้ยขยัน 0					
ค่าตำแหน่ง 0					
ค่าเช่า 0					
ค่าจ่าย 0					
ค่าประกัน 0					
อื่นๆ 0					
รวมรายได้ Total Income 87,000	รวมรายหัก Total Deduction 3,750	เงินได้สุทธิ Net Income 83,250			

จำนวนวันที่คิดเงิน Total 31 Days
 วันที่ 31 ธันวาคม

รูปภาพที่ 4.26 รูปภาพแสดงการจัดรูปแบบใหม่

4.7 ผลการเรียกใช้การแก้ประโยคและคำผิด

โดยจะถูกเรียกใช้ในตอนที่จะต้องทำการแปลงข้อมูลให้อยู่รูปแบบของ One-Hot Encoding ซึ่งจะทำให้การเช็คข้อมูลที่เป็นประโยคก่อนแล้วทำการเช็คคำ โดยจะได้ดังรูปภาพที่ 4.27 และ 4.28 เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

และประจำตัวผู้เสียภาษี --> เลขประจำตัวผู้เสียภาษี
รายหัก --> รายการหัก
ค่าแรง --> ค่าแรง
ชั่วโมงงานที่ --> ชั่วโมงงานที่

รูปภาพที่ 4.27 รูปภาพแสดงการแก้ประโยค

bongkok --> bangkok
เบีย --> เบี้ย

รูปภาพที่ 4.28 รูปภาพแสดงการแก้คำ

4.8 ผลการเรียกใช้การดึงข้อมูลด้วยโมเดลการทำนายผล

ในส่วนของการเรียกใช้การดึงข้อมูลด้วยโมเดลการทำนายผลนั้น จะทำการเรียกไฟล์ของโมเดลที่ได้เคยทำการฝึกฝนไว้ซึ่งในไฟล์ของโมเดลนั้นจะมีการเก็บค่าพารามิเตอร์ต่าง ๆ ไว้ โดยจะทำการนำรูปภาพสลิปเงินเดือนไปทำการปรับ Threshold และลบเส้นตารางก่อนโดยผลลัพธ์ของการเรียกใช้จะได้ข้อมูลของแต่ละ tags โดยจะประกอบด้วยตำแหน่งของข้อมูล tag นั้นในรูปภาพ และข้อมูลที่เป็นข้อความ โดยจะเป็นดังรูปภาพที่ 4.25 และ 4.6

```
ผลการใช้โมเดลทำนายผล
{
  "company": {
    "coordinate": [
      "",
      "",
      ""
    ],
    "data": "",
    "verified": false
  },
  "date": {
    "coordinate": [
      "6.681818181818182",
      "75.41176470588235",
      "9.090909090909092",
      "1.5294117647058822"
    ],
    "data": "จำนวนที่คิดเงิน",
    "verified": false
  },
  "name": {
    "coordinate": [
      "43.09090909090909",
      "34.23529411764706",
      "10.727272727272727",
      "1.6470588235294117"
    ],
    "data": " นาย สมพร กว้างใหญ่",
    "verified": false
  },
  "net_in": {
    "coordinate": [
      "",
      "",
      ""
    ],
    "data": "",
    "verified": false
  }
}
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งาน... อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปภาพที่ 4.29 รูปภาพแสดงการแสดงผลของการใช้โมเดลทำนายผล

This material is reserved for educational use only, not allowed for commercial use.

```
    "section": {
      "coordinate": [
        "",
        "",
        "",
        ""
      ],
      "data": "",
      "verified": false
    },
    "total_de": {
      "coordinate": [
        "57.227272727273",
        "70.3529411764706",
        "2.81818181818183",
        "1.2941176470588236"
      ],
      "data": "3,750",
      "verified": false
    },
    "total_in": {
      "coordinate": [
        "6.6363636363637",
        "71.47058823529412",
        "6.6363636363637",
        "1.1176470588235294"
      ],
      "data": " Total Income",
      "verified": false
    },
    "type": {
      "coordinate": [
        "6.72727272727275",
        "27.941176470588236",
        "5.72727272727275",
        "1.3529411764705883"
      ],
      "data": " Payroll Slip",
      "verified": false
    }
  }
}
```

รูปภาพที่ 4.30 รูปภาพแสดงการผลของการใช้โมเดลทำนายผล



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

บทที่ 5

สรุปผลการดำเนินงาน

5.1 สรุปผลการทดลอง

วิทยานิพนธ์นี้มีวัตถุประสงค์ เพื่อช่วยในการดึงข้อมูลออกจากรูปภาพสลิปเงินเดือนนั้นมีความง่ายขึ้นต่อการที่จะนำไปใช้งานในหรือนำไปประมวลผลในขั้นตอนต่อ ๆ และยังมีส่วนของหน้าเว็บแอปพลิเคชันหรือทาง front-end นั้นจะใช้ ReactJS และ Tailwindcss ในการพัฒนาเพื่อที่จะให้ผู้ใช้งานสามารถใช้งานได้สะดวกสบายอีกทั้งยังสามารถที่จะทำการแก้ไขข้อมูลที่มีความผิดพลาดได้อย่างง่ายดายอีกด้วย โดยจะมีการพัฒนาในส่วนของ back-end เพื่อที่จะใช้ในการเชื่อมต่อระหว่างเว็บแอปพลิเคชันกับฐานข้อมูลหรือการเรียกใช้งานการดึงข้อมูลด้วยโมเดลทำนายผลหรือการเรียกใช้งานเซอร์วิสต่าง ๆ ผ่าน API ซึ่งทำการพัฒนาโดยใช้ Django & Django Rest Framework และใช้ PostgreSQL สำหรับฐานข้อมูล โดยในส่วนของโมเดลทำนายผลนั้นใช้ Graph Convolutional Network ซึ่งเป็นโครงข่ายประสาทเทียมที่ใช้ข้อมูลที่มีความสัมพันธ์กันในรูปแบบของกราฟซึ่งจะใช้ Macro F1-Score เป็นค่าวัดที่ผลความสามารถของโมเดลโดยในการฝึกฝนของโมเดลนั้นจะได้ค่าอยู่ที่ 0.79 และในการทดสอบโมเดลนั้นจะได้ค่าอยู่ที่ 0.74

5.2 ปัญหาและอุปสรรคของการดำเนินงาน

5.2.1 รูปภาพของสลิปเงินเดือนที่ใช้ในการฝึกฝนและทดสอบนั้นมีจำนวนและความหลากหลายที่ไม่มากพอทำให้การเรียนรู้ของโมเดลนั้นยังไม่ฉลาดมากพอ

5.2.2 ในการแก้ไขคำหรือประโยคด้วยการใช้ Cosine Distance นั้นยังไม่ตอบโจทย์ในกรณีที่คำนั้นมีลักษณะคล้ายกันแต่มีความหมายต่างกันเช่น “นาง” กับ “งาน”

5.2.3 การทำ OCR ของรูปภาพสลิปเงินเดือนบ้างคำหรือประโยคนั้นไม่ถูกต้องจึงทำให้ความหมายนั้นเปลี่ยนไปได้

5.3 ข้อเสนอแนะและแนวทางการแก้ไข

5.2.1 ในการฝึกฝนโมเดลนั้นควรใช้รูปภาพสลิปเงินเดือนที่มีความหลากหลายเพื่อให้โมเดลเรียนรู้ได้ฉลาดมากขึ้น

5.2.2 ในการแก้ไขคำในอนาคตอาจจะต้องใช้อัลกอริทึมใหม่เข้ามาช่วยเช่น LSTM

5.2.3 ในการเรียกใช้ OCR นั้นรูปภาพสลิปเงินเดือนที่ควรที่จะมีความคมชัดและสามารถอ่านออกได้เพื่อให้การทำ OCR นั้นมีประสิทธิภาพมากที่สุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น อนุญาตให้นำไปเผยแพร่บนสื่อออนไลน์ได้เพื่อการทำ OCR นั้นมีประสิทธิภาพมากที่สุด
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

บรรณานุกรม

[1] Chau Pham. 2020. Graph Convolutional Networks. [Online].

Available: <https://ai.plainenglish.io/graph-convolutional-networks-gcn-baf337d5cb6b>

[2] Thomas N Kipf and Max Welling. 2016. Semi-supervised classification with graph convolutional networks. [Online].

Available: <https://arxiv.org/pdf/1609.02907.pdf>

[3] d2lai. Forward Propagation, Backward Propagation, and Computational Graphs. [Online].

Available: https://d2lai.com/chapter_multilayer-perceptrons/backprop.html

[4] HD COE. 2019. Cost, Deep learning พื้นฐาน (Back propagation). [Online].

Available: <https://medium.com/@hadee2531earvesdrop/deep-learning-พื้นฐาน-back-propagation-10eead7496fe>

[5] Z2 Little. 2020. Activation Functions (Linear/Non-linear) in Deep Learning. [Online].

Available: <https://xzz201920.medium.com/activation-functions-linear-non-linear-in-deep-learning-relu-sigmoid-softmax-swish-leaky-relu-a6333be712ea>

[6] Mr.P L. 2019. Deep Learning แบบฉบับสามัญชน EP 2 Optimization & Activation Function เรียนกันสบายๆสไตล์ชิลๆ. [Online].

Available: <https://medium.com/mmp-li/deep-learning-แบบฉบับสามัญชน-ep-2-optimization-activation-function-เรียนกันสบายๆสไตล์ชิลๆ-9feb5a87e3b2>

[7] Javaid Nabi. 2018. Machine Learning —Fundamentals. [Online].

Available: <https://towardsdatascience.com/machine-learning-basics-part-1-a36d38c7916>

เอกสารนี้เป็นเอกสารที่จัดทำขึ้นไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

บรรณานุกรม(ต่อ)

[8] Victor Zhou. 2019. Machine Learning for Beginners: An Introduction to Neural Networks. [Online].

Available: <https://towardsdatascience.com/machine-learning-for-beginners-an-introduction-to-neural-networks-d49f22d238f9>

[9] XuanKhanh Nguyen. 2020. Minimizing the cost function: Gradient descent. [Online].

Available: <https://towardsdatascience.com/minimizing-the-cost-function-gradient-descent-a5dd6b5350e1>

[10] Mr.P L. 2018. Gradient Descent คีย์เวิร์ดที่ Data Science ต้องรู้จัก ! : Machine Learning 101. [Online].

Available: <https://medium.com/mmp-li/gradient-descent-คีย์เวิร์ดที่-data-science-ต้องรู้จัก-machine-learning-101-b5c78c1787a0>

[11] Mr.P L. 2018. Under Fitting / Over Fitting ปัญหาที่มองไม่เห็นแต่สัมผัสได้ว่ามี..... : Machine Learning 101. [Online].

Available: <https://medium.com/mmp-li/under-fitting-over-fitting-ปัญหาที่มองไม่เห็นแต่สัมผัสได้ว่ามี-4c694bffc06>

[12] Engkarat Techapanurak. 2018. Overfitting คำคำตอบของข้อสอบอย่างเดียว ไม่สามารถเข้าใจความรู้ที่แท้จริงได้ — ปัญหาที่ AI เจอเหมือนมนุษย์. [Online].

Available: <https://medium.com/@dopplerz/overfitting-คำตอบของข้อสอบอย่างเดียว-ไม่สามารถเข้าใจความรู้ที่แท้จริงได้-ปัญหาที่-ai-5bf27c266462>

[13] Engkarat Techapanurak. 2018. Overfitting คำคำตอบของข้อสอบอย่างเดียว ไม่สามารถเข้าใจความรู้ที่แท้จริงได้ — ปัญหาที่ AI เจอเหมือนมนุษย์. [Online].

Available: <https://medium.com/@dopplerz/overfitting-คำตอบของข้อสอบอย่างเดียว-ไม่สามารถเข้าใจความรู้ที่แท้จริงได้-ปัญหาที่-ai-5bf27c266462>

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

บรรณานุกรม(ต่อ)

- [14] T.Romen Singh, Sudipta Roy, O.Imocha Singh, Tejmani Sinam , Kh.Manglem Singh. 2011. A New Local Adaptive Thresholding Technique in Binarization. [Online]. Available: <https://arxiv.org/ftp/arxiv/papers/1201/1201.5227.pdf>
- [15] Opencv. Morphological Transformations. [Online]. Available: <https://shorturl.at/dkCET>
- [16] Nickson Joram. 2020. Morphological Operations in Image Processing. [Online]. Available: <https://himnickson.medium.com/morphological-operations-in-image-processing-cb8045b98fcc>
- [17] Microsoft. 2021. Visual Studio Code. [Online]. Available: <https://code.visualstudio.com/>
- [18] Mozilla. 2021. Django Introduction. [Online]. Available: <https://developer.mozilla.org/en-US/docs/Learn/Server-side/Django/Introduction>
- [19] Django Rest Frame. 2021. Django Rest Framework. [Online]. Available: <https://www.django-rest-framework.org/>
- [20] Docker Docs. 2021. Docker overview. [Online]. Available: <https://docs.docker.com/get-started/overview/>

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

Poster

Department of Computer Engineering (Information Engineering)

Web Application Information Extraction For Pay Slip Images

Author Khajohnyos Boonkate Advisor Asst. Prof. Dr. Kleddao Satcharoen

Abstract

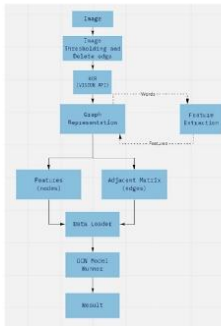
Payslip documents come in a variety of formats, making it difficult to retrieve the required information. In this project, we will present the required extraction of the document using Graph Convolution Network (GCN) using node classification. It uses node and adjacent node relationships by looking at the four directions - top, bottom, left and right - they are all deployed via web applications for ease of use and ease of editing.

Introduction

Nowadays, the world is entering more and more digital Information age, but many of the business processes are still heavily used in paper. The same is true for financial institutions to deal with the processing of paper documents such as statements or payslip. The authors developed web application which can extract the necessary information or the desired part to be used in the images, documents that are multi-format, so that it is easy to use the information from the document for further processing. Web application can be done by uploading images and documents into the system. After that, the system will process and convert the payslip image to text and will predict the desired field for the user such as name field or company name field etc. and can be edited in case of wrong prediction and can add the fields you want to add after that you can choose to export the data to the form of a document file or send it by email.

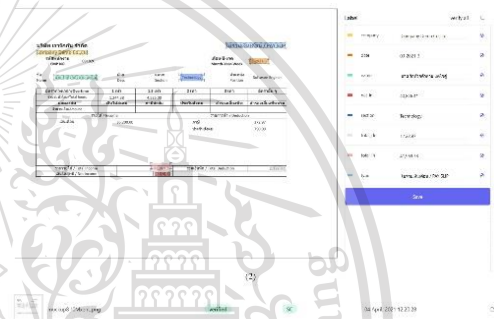
Methodology

This project consists of a frontend backend and model prediction to retrieve required fields. In frontend was built to make it easier for users to use and edit data, developed using React JS which is written in JavaScript Language. In backend was built to connect to database and to connect between user in frontend and model prediction, using Django & Django Rest Framework for RESTful API and using PostgreSQL for database. In model prediction can be separated into 4 main part.



- 1. Image Processing for thresholding image to reduce noise and delete grid line in image.
2. OCR for converting image data to text data for use in Natural Language Processing in next step.
3. Natural Language Processing for word tokenize and correct the sentences or words and use to feature extraction for each node.
4. Graph Convolution Network for using node classification to classify which class a node should be. It use the relationship in the form of a graph. The Graph consists of nodes and adjacent nodes with edges connected in 4 directions: top, bottom, left and right. The node represents the text area in document.

Results



User can easily to information extraction from payslip image via web application. Which can be easily edited and exported as a document (CSV) file.

Conclusion

This project makes it easier for users to retrieve data from payslip images, which can save time in this task. But the model of prediction is not very accurate. As there are not many data sets used in train and test in figure (4) showing model's performance by F1-Score, Precision and Recall.

Table with 5 columns: precision, recall, f1-score, support. Rows include company, date, label, name, net_in, section, total_in, total_in, type, accuracy, macro avg, weighted avg.

References

Thomas N Kipf and Max Welling. 2016. Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:1609.02907.
Xiaojing Liu, Feiyu Gao, Qiong Zhang, Huasha Zhao. 2019 Graph Convolution for Multimodal Information Extraction for Visually Rich Documents. arXiv preprint arXiv:1903.11279v1.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.

ประวัติผู้จัดทำ



หัวข้อโครงการ - เว็บแอปพลิเคชันการสกัดข้อมูลจากรูปภาพสลิปเงินเดือนหลายรูปแบบ

ชื่อ - สกุล นาย ขจรยศ บุญเกิด

รหัสนักศึกษา 60010105

คณะ วิศวกรรมศาสตร์

ภาควิชา วิศวกรรมคอมพิวเตอร์

สาขาวิชา วิศวกรรมสารสนเทศ

ประวัติการศึกษา

2553 - 2556 จบการศึกษาจากโรงเรียนนวมินทราชูทิศ กรุงเทพมหานคร (มัธยมต้น)

2557 - 2560 จบการศึกษาจากโรงเรียนนวมินทราชูทิศ กรุงเทพมหานคร (มัธยมปลาย)

2560 - 2564 สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง (ระดับปริญญาตรี)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

This material is reserved for educational use only, not allowed for commercial use.