การตรวจสอบสลิปการเงินด้วยวิธีการเรียนรู้เชิงลึก

FINANCIAL DATA EXTRACTION FROM PAPER SLIP WITH DEEP LEARNING

โดย

นายกฤษฎา ใจดี

นายศรัณญ์ สุวรรณพล

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต

ภาควิชาวิศวกรรมโทรคมนาคม

คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2565

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การตรวจสอบสลิปการเงินด้วยวิธีการเรียนรู้เชิงลึก

FINANCIAL DATA EXTRACTION FROM PAPER SLIP WITH DEEP LEARNING

โดย

นายกฤษฎา	ใจดี	62010032
นายศรัณญ์	สุวรรณพล	62010846

อาจารย์ที่ปรึกษา

ดร.สมปอง	วิเศษพานิชกิจ
ผศ. ดร.ธเนศ	พัฒนธาดาพงษ์

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต

ภาควิชาวิศวกรรมโทรคมนาคม

คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2565

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ปริญญานิพนธ์ปีการศึกษา 2565

ภาควิชาวิศวกรรมโทรคมนาคม

คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เรื่อง การตรวจสอบสลิปการเงินด้วยวิธีการเรียนรู้เชิงลึก

FINANCIAL DATA EXTRACTION FROM PAPER SLIP WITH DEEP LEARNING

ผู้จัดทำ

1. นายกฤษฎา ใจดี 62010032
2. นายศรัณญ์ สุวรรณพล 62010846

..... **อาจารย์ที่ปรึกษา**

(ดร.สมปอง วิเศษพานิชกิจ)

..... **อาจารย์ที่ปรึกษาร่วม**

(ผศ. ดร.ธเนศ พัฒนธาดาพงษ์)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

การดำเนินปริญญานิพนธ์เรื่อง “การตรวจสอบสลิปการเงินด้วยวิธีการเรียนรู้เชิงลึก” จะไม่สามารถสำเร็จลุล่วงไปได้ หากไม่ได้รับความอนุเคราะห์อย่างยิ่งจากอาจารย์ที่ปรึกษาโครงงาน คือ ดร.สมปอง วิเศษพานิชกิจ, ผศ. ดร.ธเนศ พัฒนธาดาพงษ์, ผศ. ดร.นภัทร สระเอี่ยม และ ผศ. ดร.สมเกียรติ ฤกษ์วรัญญู ที่กรุณาให้คำสั่งสอน และแนวทางการแก้ไขปัญหาที่เป็นประโยชน์ต่อการศึกษา ตลอดระยะเวลาในการจัดทำปริญญานิพนธ์นี้ รวมทั้งสนับสนุนสถานที่ เครื่องมือ และอุปกรณ์ต่าง ๆ ที่จำเป็นต้องใช้ในระหว่างการจัดทำปริญญานิพนธ์ ขอขอบพระคุณท่านในความห่วงใย และความปรารถนาดีที่ให้แก่คณะผู้จัดทำเป็นอย่างยิ่ง

ขอขอบคุณคณาจารย์และเจ้าหน้าที่ประจำภาควิชาวิศวกรรมโทรคมนาคม คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบังทุกท่าน ที่ได้อบรมสั่งสอน ประสิทธิ์ประสาทวิชา ความรู้ และประสบการณ์ให้แก่คณะผู้จัดทำ

ขอขอบคุณผู้มีส่วนเกี่ยวข้องทุกท่านอาทิ เช่น บิดา มารดา และเพื่อนนักศึกษา ที่คอยสนับสนุน แนะนำช่วยเหลือ และให้กำลังใจแก่คณะผู้จัดทำเสมอมา จนกระทั่งโครงงานสำเร็จลุล่วงไปได้ด้วยดี

นายกฤษฎา ใจดี
นายศรัณญ์ สุวรรณพล
ผู้จัดทำ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การตรวจสอบสลิปการเงินด้วยวิธีการเรียนรู้เชิงลึก
FINANCIAL DATA EXTRACTION FROM PAPER SLIP
WINTH DEPP LEARNING

โดย นายกฤษฎา ใจดี 62010032
นายศรัณญ์ สุวรรณพล 62010846

อาจารย์ที่ปรึกษา ดร. สมปอง วิเศษพานิชกิจ
อาจารย์ที่ปรึกษาร่วม ผศ. ดร.ธเนศ พัฒนธาดาพงษ์

บทคัดย่อ

การตรวจสอบข้อมูลรูปภาพจากรูปภาพสลิปการเงินจำเป็นต้องใช้เวลาในการดำเนินการที่ยาวนานโดยเฉพาะการหารูปภาพสลิปการเงินจากรูปภาพหลากหลายชนิดในปริมาณที่มาก และการระบุหาข้อมูลที่สำคัญจากรูปภาพสลิปการเงิน จุดประสงค์ของปริญญานิพนธ์นี้ต้องการศึกษา และใช้งานเทคโนโลยีที่สามารถช่วยเพิ่มประสิทธิภาพหรือความรวดเร็วในการตรวจสอบข้อมูลรูปภาพสลิปการเงิน โดยเริ่มจากการแยกรูปภาพสลิปการเงินออกจากรูปภาพชนิดอื่น และแบ่งประเภทสลิปการตามธนาคารจะมุ่งเน้นไปที่การใช้งานการเรียนรู้เชิงลึก (Deep Learning) สำหรับการฝึกสอน และทดสอบประสิทธิภาพของโมเดลการเรียนรู้เชิงลึกจะใช้ชุดข้อมูลรูปภาพสลิปการเงินที่ออกโดยธนาคารพาณิชย์ในไทย และสถาบันการเงินเฉพาะกิจ ขั้นตอนถัดมา คือการเก็บข้อมูล (Data Extraction) โดยในการเก็บข้อมูลใช้เทคโนโลยี Optical Character Recognition เพื่อนำข้อมูลที่ได้ไปทำการประมวลผล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ABSTRACT

The checking of image data from financial slip images requires a long processing time, especially searching for financial slip images from various types of images in large quantities and identifying important information from pictures of financial slips. The purpose of this thesis is to study and use technology that can increase efficiency or quickness in checking financial slip image data. It begins by separating financial slip images from other types of images and categorizing the bank slips using deep learning methods. For training and testing the effectiveness of the deep learning, this thesis uses the financial slip image dataset from Thai commercial banks and specialized financial institutions. The next step is data collection by using OCR technology to extract data to use the information obtained for processing.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ

		หน้า
กิตติกรรมประกาศ		I
บทคัดย่อ		II
สารบัญ		IV
สารบัญรูป		VI
สารบัญตาราง		XII
บทที่ 1	**บทนำ**	**1**
	1.1 ความเป็นมาและความสำคัญของปัญหา	1
	1.2 วัตถุประสงค์	1
	1.3 ขอบเขตของโครงงาน	2
บทที่ 2	**ทฤษฎีและหลักการที่เกี่ยวข้อง**	**4**
	2.1 Financial Data Extraction	4
	2.2 การเรียนรู้เชิงลึก (Deep Learning)	11
	2.3 Confusion Matrix	19
	2.4 การเพิ่มประสิทธิภาพของโครงข่ายประสาทเทียมแบบคอนโวลูชัน	21
	2.5 RGB	21
	2.6 การแปลงภาพสีเป็นภาพระดับสีเทา (Color to Grayscale Image)	22
	2.7 ภาษา Python	23
	2.8 Visual Studio Code	24
	2.9 Jupyter Notebook	25
	2.10 User Interface	25
	2.11 Data Visualization	25

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ (ต่อ)

		หน้า
บทที่ 3	**การออกแบบและการจัดทำปริญญานิพนธ์**	**26**
	3.1 การออกแบบ	26
	3.2 เครื่องมือที่ใช้ในการทดลอง	57
	3.3 การจัดเก็บผลการทดลอง	57
บทที่ 4	**ผลการทดลอง**	**59**
	4.1 ทดสอบการปรับขนาดรูปภาพเพื่อให้ทราบค่าขนาดที่เหมาะสมกับโมเดลที่ออกแบบ	59
	4.2 ทดสอบการจำแนกรูปภาพสลิปการเงินออกจากรูปภาพชนิดอื่น	66
	4.3 ทดสอบการเปลี่ยนภาพระดับสีเทา	67
	4.4 ทดสอบการแยกชนิดของสลิปการเงิน	68
	4.5 ทดสอบการแบ่งประเภทสลิปตามธนาคารของสลิปการเงินแบบอิเล็กทรอนิกส์	69
	4.6 ทดสอบการแบ่งประเภทสลิปตามธนาคารของสลิปการเงินกระดาษ	71
	4.7 ทดสอบการเก็บข้อมูลสลิปการเงิน (Financial Data Extraction)	73
	4.8 ทดสอบการแสดงผลข้อมูล	80
บทที่ 5	**สรุปผลและข้อเสนอแนะ**	**85**
	5.1 สรุปผล	85
	5.2 ข้อเสนอแนะ	85
บรรณานุกรม		**86**
ภาคผนวก ก	โปรแกรมสำหรับโมเดลการเรียนรู้เชิงลึก	89
ภาคผนวก ข	โปรแกรมแบ่งประเภทสลิปตามธนาคาร	100
ภาคผนวก ค	โปรแกรมดึงข้อมูลจากรูปภาพสลิปการเงิน	126

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป

หน้า

รูปที่

1.1 ตัวอย่างของสลิปการเงินแบบอิเล็กทรอนิกส์ 2
1.2 ตัวอย่างของสลิปการเงินแบบกระดาษ 2
2.1 โครงสร้างของ GPT-2 5
2.2 กราฟความสัมพันธ์ของเอนทิตี 6
2.3 ตัวอย่างการทำ NER 6
2.4 เทคโนโลยี Computer Vision 8
2.5 เทคโนโลยี NLP 8
2.6 เทคโนโลยี Fuzzy Logic 9
2.7 เทคโนโลยี Machine Learning 9
2.8 การทำงานของ OCR 10
2.9 เปรียบเทียบฟังก์ชัน OCR และ IDP 11
2.10 ตัวอย่างโครงสร้างของ CNN 12
2.11 ตัวอย่างของการทำการคอนโวลูชัน 13
2.12 การทำ Feature map จากการคอนโวลูชัน 13
2.13 ตัวกรองขนาด 3X3 13
2.14 ตัวอย่างในการเลื่อนตัวกรอง (Filter) เมื่อกำหนดให้ Stride เท่ากับ 1 14
2.15 ตัวอย่างของการทำ Padding โดยการตั้งค่าให้ padding = 1 14
2.16 ตัวอย่างการทำ Max pooling และ Average pooling 15
2.17 ตัวอย่างการใช้งาน Flatten layer 15
2.18 การเชื่อมกันของแต่ละชั้นอย่างสมบูรณ์ใน Fully connected layer 16
2.19 กราฟการคำนวณสมการ Gradient 17
2.20 การลู่เข้าของ Gradient Descent ทั้ง 3 ประเภท 17
2.21 Confusion Matrix 19
2.22 Multiclass Confusion Matrix 20
2.23 รูปภาพก่อน และหลังการแปลงเฉดสีเป็นภาพระดับสีเทา 22

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

หน้า

รูปที่

2.24 ข้อมูลที่ถูกนำเสนอแบบ Data Visualization 25
3.1 บล็อกไดอะแกรมของการตรวจสอบสลิปการเงินด้วยวิธีการเรียนรู้เชิงลึก 26
3.2 แผนผังการแยกรูปภาพสลิปการเงินออกจากรูปภาพชนิดอื่น 27
3.3 Layer ของ CNN Model สำหรับแยกรูปภาพสลิปการเงินและรูปภาพชนิดอื่น 30
3.4 พารามิเตอร์ในแต่ละ Layer ของโมเดลที่ใช้ในการแยกรูปภาพสลิปการเงิน 31
3.5 แผนผังการแยกชนิดของสลิปการเงิน 32
3.6 Layer ของ CNN Model สำหรับแยกชนิดของสลิปการเงิน 34
3.7 พารามิเตอร์ในแต่ละ Layer ของโมเดลที่ใช้แยกชนิดของสลิปการเงิน 35
3.8 แผนผังการแบ่งประเภทของสลิปการเงินตามธนาคารของสลิปการเงินแบบอิเล็กทรอนิกส์ 36
3.9 Layer ของ CNN Model สำหรับแบ่งประเภทตามธนาคารของสลิปการเงินแบบอิเล็กทรอนิกส์ 41
3.10 พารามิเตอร์ในแต่ละ Layer ของโมเดลที่ใช้สำหรับแบ่งประเภทสลิปตามธนาคารของสลิปการเงินแบบอิเล็กทรอนิกส์ 42
3.11 แผนผังการแบ่งประเภทของสลิปการเงินตามธนาคารของสลิปการเงินแบบกระดาษ 43
3.12 Layer ของ CNN Model สำหรับแบ่งประเภทตามธนาคารของสลิปการเงินแบบกระดาษ 47
3.13 พารามิเตอร์ในแต่ละ Layer ของโมเดลที่ใช้สำหรับแบ่งประเภทสลิปตามธนาคารของสลิปการเงินแบบกระดาษ 48
3.14 แผนผังการดึงข้อมูลออกจากรูปภาพสลิปการเงิน 49
3.15 แผนผังการแสดงผลข้อมูล 50
3.16 เว็บแอปพลิเคชันที่ได้ออกแบบ 51
3.17 เว็บแอปพลิเคชันที่ได้ออกแบบ 51

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

หน้า

รูปที่

3.18 แนะนำการอัปโหลดข้อมูล 52
3.19 คู่มือการใช้งานกราฟแสดงวันเวลาที่ทำธุรกรรม 52
3.20 แนะนำ Filter 53
3.21 Slip Recognition 53
3.22 Financial Data Extraction 54
3.23 Data Visualization 54
3.24 Upload file 55
3.25 กราฟวันเวลาที่ทำธุรกรรม 55
3.26 แผนภูมิข้อมูลจำนวนครั้งในการโอน และจำนวนเงินทั้งหมดในการทำธุรกรรม 55
3.27 Option 56
3.28 Name Filter 56
3.29 ผลลัพธ์การ Filter 57
4.1 ผลการฝึกสอนโมเดลที่มีขนาดอินพุต 8x8 pixel 59
4.2 กราฟค่า Loss, Accuracy ที่มีขนาดอินพุต 8x8 pixel 60
4.3 Confusion Matrix ของโมเดลที่มีขนาดอินพุต 8x8 pixel 60
4.4 ผลการฝึกสอนโมเดลที่มีขนาดอินพุต 16x16 pixel 60
4.5 กราฟค่า Loss, Accuracy ของโมเดลที่มีขนาดอินพุต 16x16 pixel 61
4.6 Confusion Matrix ของโมเดลที่มีขนาดอินพุต 16x16 pixel 61
4.7 ผลการฝึกสอนโมเดลที่มีขนาดอินพุต 28x28 pixel 61
4.8 กราฟค่า Loss, Accuracy ของโมเดลที่มีขนาดอินพุต 28x28 pixel 62
4.9 Confusion Matrix ของโมเดลที่มีขนาดอินพุต 28x28 pixel 62
4.10 ผลการฝึกสอนโมเดลที่มีขนาดอินพุต 32x32 pixel 62
4.11 กราฟค่า Loss, Accuracy ของโมเดลที่มีขนาดอินพุต 32x32 pixel 63

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

หน้า

รูปที่

4.12 Confusion Matrix ที่มีขนาดอินพุต 32x32 pixel 63
4.13 ผลการฝึกสอนโมเดลที่มีขนาดอินพุต 64x64 pixel 63
4.14 กราฟค่า Loss, Accuracy ของโมเดลที่มีขนาดอินพุต 64x64 pixel 64
4.15 Confusion Matrix ของโมเดลที่มีขนาดอินพุต 64x64 pixel 64
4.16 ผลการฝึกสอนโมเดลที่มีขนาดอินพุต 128x128 pixel 64
4.17 กราฟค่า Loss, Accuracy ของโมเดลที่มีขนาดอินพุต 128x128 pixel 65
4.18 Confusion Matrix ของโมเดลที่มีขนาดอินพุต 128x128 Pixel 65
4.19 ผลการฝึกสอนโมเดลโมเดลจำแนกรูปภาพสลิปการเงิน 66
4.20 กราฟค่า Loss, Accuracy ของโมเดลจำแนกรูปภาพสลิปการเงิน 66
4.21 ผลการทดสอบโมเดลจำแนกรูปภาพสลิปการเงิน 66
4.22 Confusion Matrix ของโมเดลจำแนกรูปภาพสลิปการเงิน 67
4.23 รูปภาพก่อน และหลังการแปลงเฉดสีเป็นภาพระดับสีเทา 67
4.24 ผลการฝึกสอนโมเดลแยกชนิดของสลิปการเงิน 68
4.25 กราฟค่า Loss, Accuracy ของโมเดลแยกชนิดของสลิปการเงิน 68
4.26 ผลการทดสอบโมเดลแยกชนิดของสลิปการเงิน 68
4.27 Confusion Matrix ของโมเดลแยกชนิดของสลิปการเงิน 69
4.28 ผลการฝึกสอนโมเดลแบ่งประเภทสลิปอิเล็กทรอนิกส์ 70
4.29 กราฟค่า Loss, Accuracy ของโมเดลแบ่งประเภทสลิปอิเล็กทรอนิกส์ 70
4.30 ผลการทดสอบโมเดลแบ่งประเภทสลิปอิเล็กทรอนิกส์ 70
4.31 Confusion Matrix ของโมเดลแบ่งประเภทสลิปอิเล็กทรอนิกส์ 71
4.32 ผลการทดสอบโมเดลแบ่งประเภทสลิปกระดาษ 72
4.33 กราฟค่า Loss, Accuracy ของโมเดลแบ่งประเภทสลิปกระดาษ 72
4.34 ผลการทดสอบโมเดลแบ่งประเภทสลิปกระดาษ 72
4.35 Confusion Matrix ของโมเดลแบ่งประเภทสลิปกระดาษ 73
4.36 สลิปการเงินแบบอิเล็กทรอนิกส์ก่อนเข้าสู่กระบวนการ OCR 73

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

หน้า

รูปที่

4.37 ข้อความที่ได้จากการทำ OCR ของรูปภาพสลิปการเงินแบบอิเล็กทรอนิกส์ 74
4.38 สลิปการเงินแบบกระดาษก่อนเข้าสู่กระบวนการ OCR 74
4.39 ข้อความที่ได้จากการทำ OCR ของรูปภาพสลิปการเงินแบบกระดาษ 75
4.40 Data Frame ของธนาคาร SCB 75
4.41 Data Frame ของธนาคาร TTB 76
4.42 Data Frame ของธนาคาร BBL 76
4.43 Data Frame ของธนาคาร KTB 76
4.44 Data Frame ของธนาคาร Kbank 77
4.45 Data Frame ของธนาคาร BAY 77
4.46 .xlsx ของธนาคาร SCB 78
4.47 xlsx ของธนาคาร TTB 78
4.48 xlsx ของธนาคาร BBL 78
4.49 xlsx ของธนาคาร KTB 79
4.50 xlsx ของธนาคาร Kbank 79
4.51 xlsx ของธนาคาร BAY 79
4.52 วันเวลาที่ทำธุรกรรมของธนาคาร SCB 80
4.53 วันเวลาที่ทำธุรกรรมของธนาคาร TTB 80
4.54 วันเวลาที่ทำธุรกรรมของธนาคาร BBL 81
4.55 วันเวลาที่ทำธุรกรรมของธนาคาร KTB 81
4.56 วันเวลาที่ทำธุรกรรมของธนาคาร Kbank 81
4.57 วันเวลาที่ทำธุรกรรมของธนาคาร BAY 81
4.58 แผนภูมิแท่งของจำนวนครั้งในการโอนเงิน และจำนวนเงินทั้งหมดของแต่ละธนาคาร 82
4.59 ข้อมูลที่ถูกแสดงก่อนผ่าน Filter 83
4.60 ข้อมูลหลังผ่าน Filter ธนาคาร 83

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป (ต่อ)

หน้า

รูปที่

4.61 ข้อมูลหลังผ่าน Filter เวลา 84
4.62 ข้อมูลหลังผ่าน Advance Option 84

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่		หน้า
1.1	ธนาคารพาณิชย์ในไทย	3
1.2	สถาบันการเงินเฉพาะกิจ	3
3.1	ชุดข้อมูลสำหรับใช้ฝึกสอน และทดสอบโมเดล สำหรับการจำแนกรูปภาพสลิปการเงินออกจากรูปภาพชนิดอื่น	28
3.2	ชุดข้อมูลสำหรับใช้ฝึกสอน และทดสอบโมเดล สำหรับแยกชนิดของรูปภาพสลิปการเงิน	33
3.3	ชุดข้อมูลสำหรับใช้ฝึกสอน และทดสอบโมเดล สำหรับแบ่งประเภทของรูปภาพสลิปการเงินตามธนาคารของสลิปการเงินแบบอิเล็กทรอนิกส์	37
3.4	ชุดข้อมูลสำหรับใช้ฝึกสอน และทดสอบโมเดล สำหรับแบ่งประเภทของรูปภาพสลิปการเงินตามธนาคารของสลิปการเงินแบบกระดาษ	44
4.1	สรุปผลการทดสอบขนาดโมเดล	65

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

เนื่องจากการตรวจสอบข้อมูลรูปภาพ (Images Data) เช่น การตรวจสอบหาข้อมูลที่สนใจ (Keyword) ในข้อมูลรูปภาพโดยใช้เทคโนโลยีเพียงชนิดเดียว หรือการตรวจสอบโดยใช้สายตามนุษย์นั้นสามารถทำได้อย่างมีข้อจำกัด และปัญหาที่เกิดขึ้นคือข้อมูลรูปภาพมีมากกว่าขอบเขต และข้อจำกัดเกินกว่าที่สายตาของมนุษย์หรือเทคโนโลยีเพียงชนิดเดียวจะทำได้ ยกตัวอย่างในหน่วยความจำรอง (Second Memory) เช่น แฟรชไดร์ฟ ฮาร์ดดิสไดรฟ์ โซลิดสเตตไดรฟ์ และเมมโมรี่การ์ด โดยในหน่วยความจำที่กล่าวมาข้างต้นสามารถเก็บข้อมูลรูปภาพได้เป็นจำนวนมาก ทำให้ในการตรวจสอบแต่ละครั้งจำเป็นต้องทำการจำแนกข้อมูลรูปภาพที่สนใจออกมาจากข้อมูลรูปภาพที่ไม่สนใจก่อน ซึ่งในกระบวนการนี้ทำให้ใช้เวลาในการดำเนินการที่ยาวนาน และยังส่งผลต่อประสิทธิภาพในการทำงานที่มีผลให้เกิดความผิดพลาดได้ โดยทางคณะผู้จัดทำได้เห็นถึงความสำคัญของปัญหาที่เกิดขึ้น ปริญญานิพนธ์ฉบับนี้จึงมีจุดประสงค์ที่ต้องการเพิ่มประสิทธิภาพ และความรวดเร็วในการตรวจสอบข้อมูลรูปภาพ โดยใช้เทคโนโลยีที่กำลังเป็นที่สนใจในปัจจุบัน และถูกใช้กันอย่างแพร่หลายอย่างการเรียนรู้เชิงลึก (Deep Learning) และเทคโนโลยี Optical Character Recognition (OCR) ในการประมวลผลข้อมูล

1.2 วัตถุประสงค์

1) เพื่อศึกษาและใช้งานการเรียนรู้เชิงลึก
2) เพื่อวิเคราะห์และนำเสนอข้อมูลของสลิปการเงิน
3) เพื่อเพิ่มประสิทธิภาพและความรวดเร็วในการตรวจสอบสลิปการเงิน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.3 ขอบเขตของปริญญานิพนธ์

แยกรูปภาพสลิปการเงินทั้ง 2 ชนิด คือ รูปภาพสลิปการเงินแบบอิเล็กทรอนิกส์ที่เป็นหลักฐานที่แอปพลิเคชันของธนาคารสร้างขึ้นจากการทำธุรกรรมการเงิน แสดงตัวอย่างดังรูป 1.1 และรูปภาพสลิปการเงินแบบกระดาษที่ได้จากตู้ ATM แสดงตัวอย่างดังรูปที่ 1.2 โดยทั้ง 2 ชนิด ถูกออกโดยธนาคารพาณิชย์ในไทย 13 แห่ง [1] ที่มีจำนวนสาขาในกรุงเทพฯ มากที่สุด [2] ดังตารางที่ 1.1 และรูปภาพสลิปการเงินที่ออกโดยสถาบันการเงินเฉพาะกิจ 4 แห่ง [3] แสดงดังตารางที่ 1.2 โดยใช้การเรียนรู้เชิงลึกในการแยกรูปภาพสลิปการเงินออกจากรูปภาพชนิดอื่นเพื่อนำมาแบ่งประเภทสลิปการเงินตามธนาคาร รวมถึงการดึงข้อมูลจากรูปภาพสลิปการเงินให้อยู่ในรูปแบบของข้อความโดยใช้เทคนิคการอ่านอักขระด้วยแสงโดยนำข้อมูลที่ได้ไปประมวลผล และแสดงผลในรูปแบบ Data Visualization ผ่านเว็บแอปพลิเคชัน

รูปที่ 1.1 สลิปการเงินแบบอิเล็กทรอนิกส์

รูปที่ 1.2 สลิปการเงินแบบกระดาษ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 1.1 ธนาคารพาณิชย์ในไทย

ลำดับ	ชื่อธนาคาร
1	ธนาคารกรุงเทพ จำกัด (มหาชน)
2	ธนาคารกรุงไทย จำกัด (มหาชน)
3	ธนาคารกรุงศรีอยุธยา จำกัด (มหาชน)
4	ธนาคารกสิกรไทย จำกัด (มหาชน)
5	ธนาคารทหารไทยธนชาต จำกัด (มหาชน)
6	ธนาคารไทยพาณิชย์ จำกัด (มหาชน)
7	ธนาคารซีไอเอ็มบี ไทย จำกัด (มหาชน)
8	ธนาคารสแตนดาร์ดชาร์เตอร์ด (ไทย) จำกัด (มหาชน
9	ธนาคารยูโอบี จำกัด (มหาชน)
10	ธนาคารทิสโก้ จำกัด (มหาชน)
11	ธนาคารเกียรตินาคินภัทร จำกัด (มหาชน)
12	ธนาคารแลนด์ แอนด์ เฮ้าส์ จำกัด (มหาชน)
13	ธนาคารไอซีบีซี (ไทย) จำกัด (มหาชน)

ตารางที่ 1.2 สถาบันการเงินเฉพาะกิจ

ลำดับ	ชื่อธนาคาร
1	ธนาคารเพื่อการเกษตรและสหกรณ์การเกษตร
2	ธนาคารอาคารสงเคราะห์
3	ธนาคารออมสิน
4	ธนาคารอิสลามแห่งประเทศไทย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2

ทฤษฎีและหลักการที่เกี่ยวข้อง

ปริญญานิพนธ์เรื่องการตรวจสอบสลิปการเงินด้วยวิธีการเรียนรู้เชิงลึกได้ทำการออกแบบ และพัฒนา โดยมีทฤษฎีและหลักการที่เกี่ยวข้องดังต่อไปนี้

2.1 Financial Data Extraction

Data Extraction [4] หรือการดึงข้อมูล คือการดึงข้อมูลจากเอกสาร และแหล่งข้อมูลอื่น ๆ โดย Wikipedia ได้ให้คำนิยามที่เป็นทางการว่าการดึงข้อมูล คือการกระทำหรือกระบวนการดึงข้อมูลจากแหล่งข้อมูลซึ่งโดยปกติเป็นข้อมูลที่ไม่มีโครงสร้างให้มีโครงสร้างเพื่อใช้ในกระบวนการประมวลผลข้อมูลหรือการจัดเก็บข้อมูล

Financial Statements [4] หรืองบการเงิน คือรายงานทางการที่นำเสนอเกี่ยวกับข้อมูลต่าง ๆ ของกิจการทั้งฐานะทางการเงิน ผลการดำเนินงาน และกระแสเงินสด ในรอบหนึ่งงวดบัญชี โดยความน่าเชื่อถือของงบการเงินขึ้นอยู่กับการปฏิบัติตามมาตรฐานการบัญชี โดยในปริญญานิพนธ์ฉบับนี้ให้คำนิยามคำว่า Financial เป็นสลิปการเงินที่ออกโดยธนาคารพาณิชย์ และสถาบันการเงินเฉพาะกิจที่นำเสนอข้อมูลต่าง ๆ เช่น บัญชีต้นทาง บัญชีปลายทาง ชื่อเจ้าของบัญชีต้นทาง ชื่อเจ้าของบัญชีปลายทาง วันเวลาที่ทำรายการ และจำนวนเงินที่ทำธุรกรรม

จากการทบทวนวรรณกรรม ทำให้ทราบว่าในการทำ Data Extraction เช่น Financial Data Extraction หรือ Extracting Data from Financial PDFs มีเทคโนโลยี และวิธีการอยู่หลากหลายวิธี จากการนำความสามารถ Artificial Intelligence (AI) และ Machine Learning (ML) หลายแขนงเข้าด้วยกัน ซึ่งวิธีที่ได้จากการทบทวนวรรณกรรมแสดงดังต่อไปนี้ [5]

2.1.1 Methodology with Machine Learning

2.1.1.1 Text Extraction tools

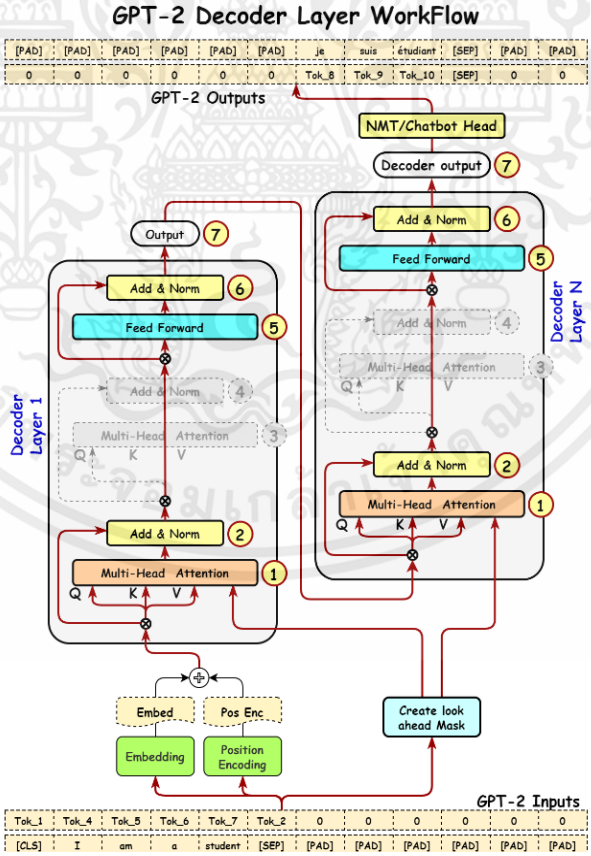
เป็นการใช้เครื่องในการสกัดข้อมูล เช่น Tabula ที่ใช้งานในข้อมูลที่อยู่ในรูปแบบของตาราง และ PDFMiner กับ Tesseract สำหรับข้อความในรูปแบบปกติสามารถใช้งานร่วมกับไฟล์ PDF และรูปภาพได้ โดยการทำงานของเครื่องมือเหล่านี้จะทำการจำแนกส่วนของข้อความบนสื่อที่แสดงผล ไม่ว่าจะเป็นรูปภาพหรือเอกสาร PDF ให้เป็นข้อมูลที่เป็นของความ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

แต่ข้อเสียที่พบของ Tabula คือเหมาะสำหรับการทำงานบนข้อมูลที่เป็นตารางเท่านั้น ซึ่งสลิปการเงินไม่ได้แสดงอยู่ในรูปแบบของตารางเสมอไป ในขณะที่เครื่องมืออย่าง PDFMiner และ Tesseract ก็พบปัญหาของข้อความดิบที่ส่งออกไปนั้นไม่มีโครงสร้างอย่างสมบูรณ์ทำให้มนุษย์ไม่สามารถอ่านได้

2.1.1.2 GPT-2

GPT-2 เป็นหนึ่งในโมเดลของ Transformers มีความสามารถในการสร้างเรื่อง ซึ่งการทำงานของ GPT-2 เมื่อใส่ประโยคเริ่มต้น ตัวของ GPT-2 จะทำการแต่งประโยคต่อขึ้นมาเอง โดย GPT-2 มีพารามิเตอร์ขนาด 1.5 พันล้านตัวที่พัฒนาจาก OpenAI แต่สำหรับกรณีของการนำไปใช้งานที่เกี่ยวข้องกับสลิปการเงินนั้นต้องทำการประมวลผลล่วงหน้าเพื่อแยกข้อมูลที่สนใจออกมาก่อนที่จะเข้าสู่กลไก GPT-2 แต่ผลที่ออกมานั้นยังไม่มีประสิทธิภาพมากพอเพราะยังไม่สามารถสรุปข้อมูลที่ไม่มีโครงสร้างได้ และมีความเป็นไปได้ว่า GPT-2 จะลบข้อมูลที่สำคัญออกจากข้อความ ซึ่งต่อมาถูกพัฒนาเป็น GPT-3 แสดงตัวอย่างโครงสร้างของ GPT-2 ดังรูปที่ 2.1 [6]

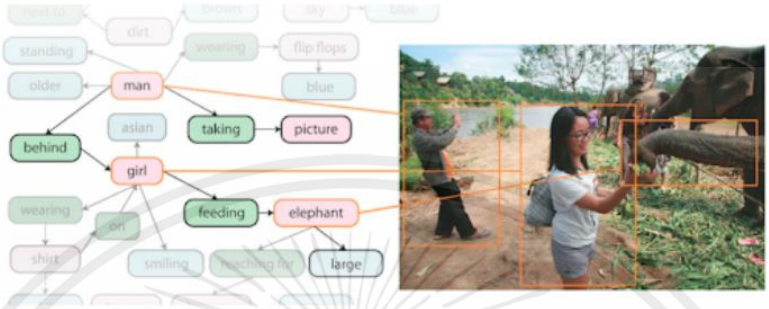


รูปที่ 2.1 โครงสร้างของ GPT-2 [6]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.1.1.3 Bounding Boxes Relationship-OpenCV

Bounding Boxes Relationship เป็นวิธีการสร้างกราฟความสัมพันธ์ระหว่างวัตถุที่กำลังถูกตรวจจับจากกรอบสัญญาณ แสดงตัวอย่างดังรูป 2.2 [5]

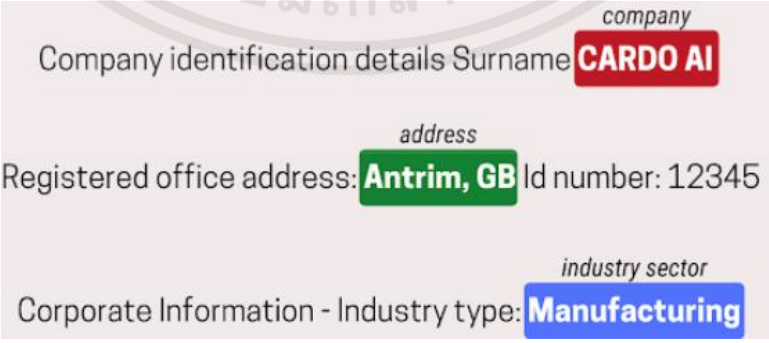


รูปที่ 2.2 กราฟความสัมพันธ์ของเอนทิตี [5]

เมื่อนำไปใช้งานกับเอกสาร หรือสลิปการเงินพบว่าต้องทำป้ายกำกับ และสร้างกรอบสัญญาณได้แค่จุดที่สามารถระบุตัวตนได้ ทำให้ต้องพัฒนา AI มาเพื่อใช้คำนวณเมทริกซ์เพื่อแสดงค่าความสัมพันธ์ระหว่างกรอบสัญญาณตามตำแหน่งเชิงพื้นที่เพื่อให้ AI สามารถตัดสินใจตำแหน่งของเอนทิตีถัดไปตามลำดับ แต่ปัญหาที่เกิดเพราะเป็นเรื่องยากที่จะระบุตำแหน่งกรอบสัญญาณถัดไป

2.1.1.4 Named Entity Recognition (NER)

Named Entity Recognition (NER) เป็นการจัดหาตำแหน่ง และทำการจัดหมวดหมู่ของกลุ่มคำที่อยู่ในเอกสาร เช่น ชื่อองค์กร ชื่อคน สถานที่ จำนวนเงิน และอื่น ๆ โดยข้อดีของ NER คือไม่จำเป็นต้องกำหนดตำแหน่งของข้อมูลไว้ล่วงหน้า โดยตัวอย่างของการทำ NER แสดงดังรูป 2.3 [5]



รูปที่ 2.3 ตัวอย่างการทำ NER [5]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.1.1.5 NLP model-Spacy

Natural Language Processing (NLP) หรือการประมวลผลภาษาธรรมชาติ เป็นการใช้ปัญญาประดิษฐ์ร่วมกับศาสตร์ของภาษาคอมพิวเตอร์ที่ทำให้คอมพิวเตอร์สามารถสื่อสารโดยใช้ภาษาของมนุษย์ และทำให้คอมพิวเตอร์เข้าใจเกี่ยวกับภาษามนุษย์มากขึ้น ยกตัวอย่าง Siri, Google Assistant และ Alexa

โดยวิธีการของ Machine Learning ที่กล่าวมาข้างต้น รูปแบบที่เหมาะสำหรับงานประเภทประมวลสลิปการเงินคือ NLP แต่ผลลัพธ์ของการทำ NLP ก็ยังไม่ใช่ผลลัพธ์ที่ดีที่สุดเพราะ Machine Learning ยังสามารถพัฒนาเพื่อแก้ไขช่องโหว่ที่เกิดขึ้นได้

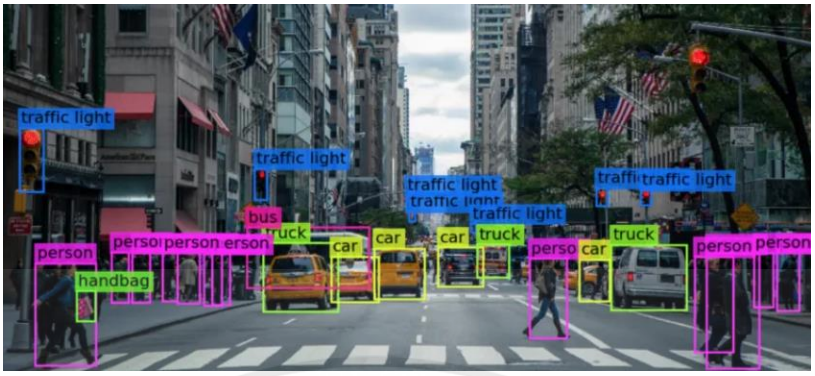
2.1.2 Intelligent Document Processing (IDP)

Intelligent Document Processing หรือการประมวลผลเอกสารอัจฉริยะ [7] เป็นการแก้ปัญหาการประมวลผลเอกสารในรูปแบบเดิม จากการที่ต้องกำหนดรูปแบบของเอกสารไว้ก่อนที่จะทำการประมวลผล เนื่องจากการทำงานในรูปแบบเดิมคือสามารถดึงข้อมูลที่อยู่ในรูปแบบเดียวกันได้เท่านั้น ซึ่งจะไม่สามารถใช้งานได้เลยหากต้องการที่จะดึงข้อมูลจากเอกสารในรูปแบบอื่น ๆ ในขณะที่เทคโนโลยีอย่าง IDP สามารถทำได้โดยการใช้ AI หลายชนิดอยู่เบื้องหลัง โดย IDP ได้เอาจุดเด่นของเทคโนโลยีอย่าง OCR กับจุดแข็งของ AI มารวมกัน ทำให้แก้ไขเรื่องการจัดรูปแบบของเอกสารไว้ล่วงหน้า ดังนั้นประสิทธิภาพของ IDP คือสามารถทำงานได้อย่างแม่นยำ และลดเวลาในการกระบวนการที่ไม่จำเป็นออก โดยมีเทคโนโลยี AI คือผู้อยู่เบื้องหลังการทำงานที่สำคัญ ซึ่งเป็นการรวมตัวของเทคโนโลยีหลากหลายแขนง เพื่อที่จะดึงข้อมูลที่เกี่ยวข้องจากเอกสาร และรูปภาพออกมาได้ดังนี้

2.1.2.1 เทคโนโลยี Computer Vision

นำมาใช้ในการจำแนกเอกลักษณ์ของข้อมูลแต่ละอย่าง และสามารถกำหนดขอบเขตพื้นที่ที่สนใจ ช่วยให้ลดขั้นตอนในการทำงานลงเพราะไม่ต้องจัดทำรูปแบบของเอกสารไว้ล่วงหน้า แสดงเทคโนโลยี Computer Vision ดังรูปที่ 2.4 [8]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.4 เทคโนโลยี Computer Vision [8]

2.1.2.2 เทคโนโลยี NLP (Natural language processing)

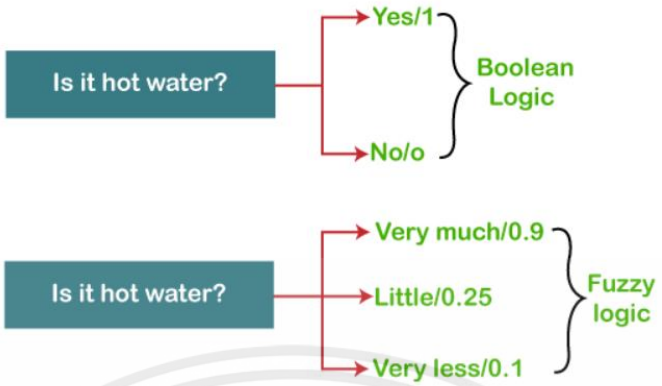
โดยเทคโนโลยีอย่าง NLP รองรับหลากหลายภาษา ทำให้การประมวลผลข้อมูลหรือข้อมูลที่ออกมานั้นสามารถเข้าถึงเนื้อหาใจความได้ เพราะเอกสารหรือข้อมูลที่ถูกดึงออกมาได้เข้าสู่การเปรียบเทียบพจนานุกรม แสดงดังรูปที่ 2.5 [9]

รูปที่ 2.5 เทคโนโลยี NLP [9]

2.1.2.3 Fuzzy Logic

เป็นเทคโนโลยีที่ช่วยในการตัดสินใจให้เหมือนกับการใช้คนทำ เพียงแต่มีความรวดเร็วมากกว่าคนได้หลายเท่า ซึ่งเมื่อนำมารวมกับ NLP ทำให้เพิ่มประสิทธิภาพของระบบได้มากยิ่งขึ้น แสดงดังรูปที่ 2.6 [10]

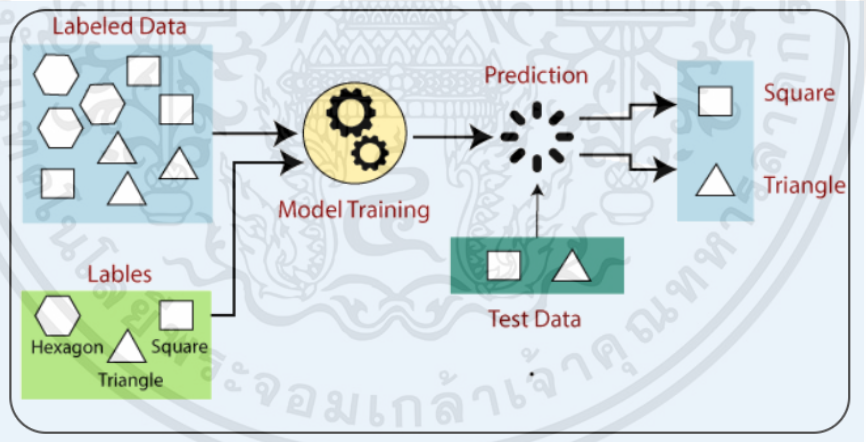
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.6 เทคโนโลยี Fuzzy Logic [10]

2.1.2.4 เทคโนโลยี Machine Learning

ทำหน้าที่ในการหาความผิดปกติจากข้อมูลที่ถูกดึงออกมาจากระบบ หรือตัวเอกสาร ทำให้สามารถตรวจสอบ แก้ไข เพื่อเพิ่มความแม่นยำ และความถูกต้องได้มากขึ้น แสดงตัวอย่างของเทคโนโลยี Machine Learning ดังรูปที่ 2.7 [11]



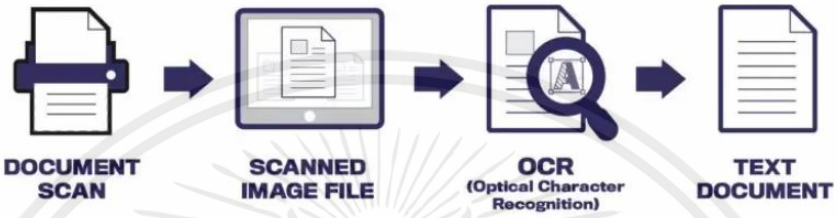
รูปที่ 2.7 เทคโนโลยี Machine Learning [11]

2.1.3 การอ่านอักระด้วยแสง (Optical Character Recognition)

OCR [12] ย่อมาจาก Optical Character Recognition (การอ่านอักขระด้วยแสง) โดย OCR เป็นเทคโนโลยีที่ใช้เพื่อ "รู้จำ" หรือ "อ่าน" ข้อความที่ปรากฏในไฟล์ภาพ OCR จะแปลงภาพที่มีข้อความซึ่งพิมพ์ด้วยมือ เขียน หรือพิมพ์ด้วยเครื่องพิมพ์จากไฟล์ โดยรวมถึงเอกสารที่ผ่าน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การสแกน รูปภาพเอกสาร รูปภาพที่มีสัญลักษณ์หรือฉากที่มีเครื่องหมาย หรือคำบรรยายใต้ภาพใน ภาพ ให้กลายเป็นข้อความซึ่งเครื่องจักรสามารถประมวลผล และค้นหาได้ในหลาย ๆ ธุรกิจมีปัญหาเรื่องการจัดการงบการเงินเพราะเป็นเอกสารที่ต้องแยกอย่างระมัดระวังเพียงแค่เกิดข้อผิดพลาดเดียวก็มีผลที่ทำให้เกิดปัญหาใหญ่ที่ตามมาได้ โดยเครื่องมือที่ช่วยลดข้อผิดพลาด และมีข้อดีอันดับต้น ๆ ก็คือ OCR โดยการทำงานของ OCR แสดงดังรูปที่ 2.8 [13]



รูปที่ 2.8 การทำงานของ OCR [13]

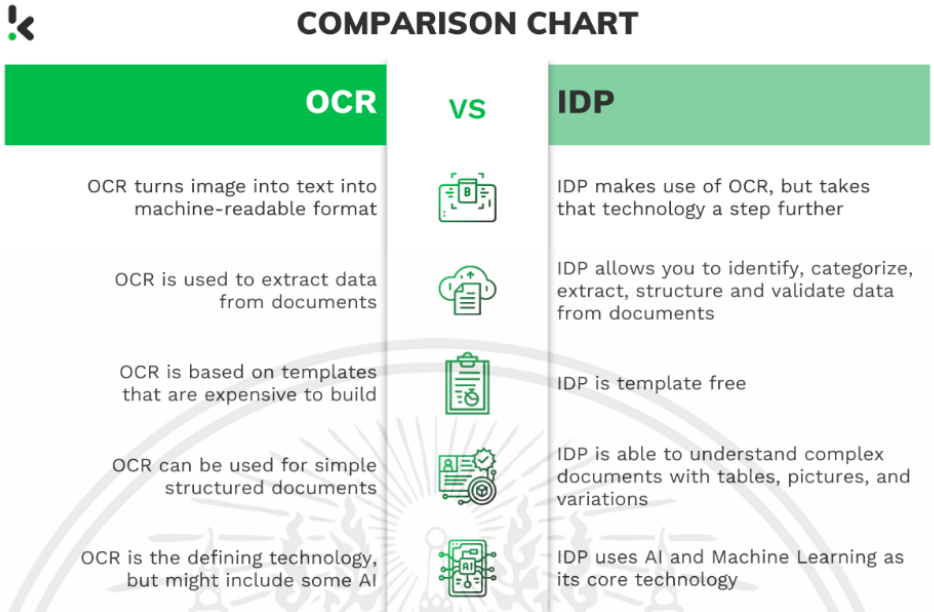
โดยในปัจจุบันเทคโนโลยีอย่าง OCR ถูกนำไปใช้งานอย่างแพร่หลาย และหนึ่งในนั้นคือการนำประมวลผลงานเรื่องสลิปการเงิน และมีการนำไปพัฒนาสู่เทคนิคต่าง ๆ โดยมีพื้นฐานของเทคโนโลยี OCR

2.1.4 การรู้จำเครื่องหมายด้วยแสง (Optical Mark Recognition)

การรู้จำเครื่องหมายด้วยแสง (OMR) [14] เป็นเครื่องมือหรือวิธีการทางอิเล็กทรอนิกส์ที่ใช้ในการสแกน และอ่านค่าตำแหน่งในการจัดการรวบรวมข้อมูลโดยต้องอาศัยเครื่องสแกนในการระบุตำแหน่งที่ต้องการสแกน โดย OMR ที่รู้จักอย่างแพร่หลาย คือเครื่องตรวจข้อสอบ เพราะเทคโนโลยีนี้จำเป็นต้องใช้การระบุไปที่ตำแหน่งสแกนอย่างดินสอ 2B

ข้อสรุปที่ได้จากการทบทวนวรรณกรรม ทำให้ทราบว่าเทคโนโลยีที่เหมาะกับการทำ Data Extraction คือ OCR และ IDP โดยที่ IDP ยังคงได้เปรียบ OCR ในทางเทคนิคเพราะเกิดจากการนำเทคโนโลยีอย่าง OCR ไปพัฒนาต่อ และรวมเข้ากับเทคโนโลยีอื่น ๆ ทำให้มีประสิทธิภาพและความรวดเร็วที่มากกว่า แต่ปริญญานิพนธ์ฉบับนี้มุ่งเน้นไปที่ข้อดีของ OCR ที่ถูกแพร่หลายมากกว่า IDP และถูกพัฒนาจนสามารถใช้งานได้สะดวก ตอบโจทย์กับการทำงานบางชนิดมากกว่า รวมไปถึงการใช้งานรวมกับภาษา Python ที่มี Library ที่สามารถเข้าถึงการใช้งานของ OCR โดยแสดงรูปการเปรียบเทียบฟังก์ชันของ OCR และ IDP ดังรูปที่ 2.9 [14]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.9 เปรียบเทียบฟังก์ชันของ OCR และ IDP [14]

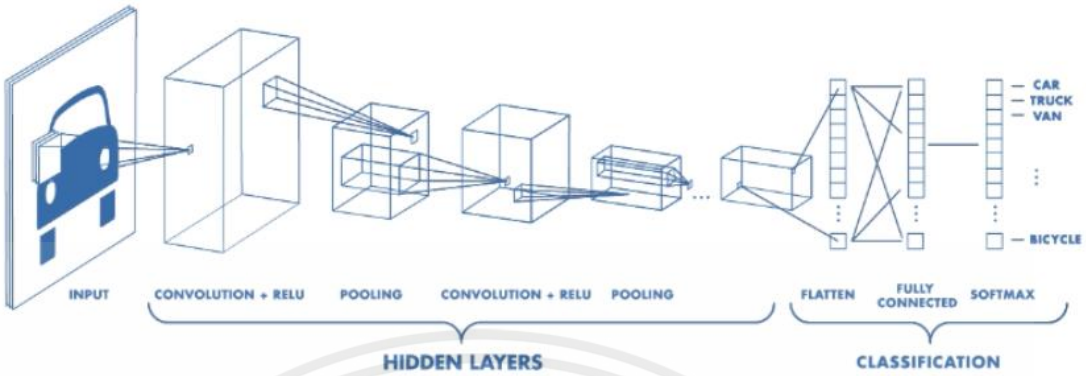
2.2 การเรียนรู้เชิงลึก (Deep Learning)

การเรียนรู้เชิงลึก [15] คือวิธีการเรียนรู้แบบอัตโนมัติด้วยการเลียนแบบการทำงานของโครงข่ายประสาทของมนุษย์ (Neurons) โดยนำระบบโครงข่ายประสาท (Neural network) มาซ้อนกันหลายชั้น (Layer) และทำการเรียนรู้ข้อมูลตัวอย่าง ซึ่งข้อมูลดังกล่าวจะถูกนำไปใช้ในการตรวจจับรูปแบบ (Pattern) หรือจัดหมวดหมู่ข้อมูล (Classify the data)

2.2.1 โครงข่ายประสาทเทียมแบบคอนโวลูชัน (Convolution Neural Networks: CNN)

Convolutional Neural Network (CNN) [16] หรือโครงข่ายประสาทแบบคอนโวลูชัน แสดงดังรูปที่ 2.10 [16] เป็นโครงข่ายประสาทเทียมที่จำลองการมองเห็นของมนุษย์ที่มองพื้นที่เป็นส่วนย่อย ๆ และนำกลุ่มของพื้นที่ย่อย ๆ เช่น ลายเส้น และการตัดกันของสี ซึ่งมนุษย์รู้ว่าพื้นที่ตรงนี้เป็นเส้นตรงหรือสีตัดกันเพราะมนุษย์ดูทั้งจุดที่สนใจและบริเวณรอบ ๆ ประกอบกัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.10 ตัวอย่างโครงสร้างของ CNN [16]

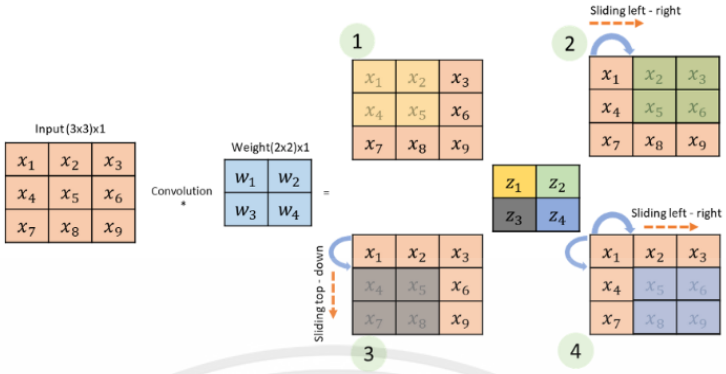
การทำงานของ CNN จะถูกแบ่งออกเป็น 2 ส่วน [17] ซึ่งในส่วนของ Hidden layer หรือ Feature extraction จะมี Convolution layer เพิ่มขึ้นมาเป็นส่วนที่ใช้ในการฟิลเตอร์ภาพ โดยมี Kernel layer เพื่อใช้แปลงคุณลักษณะและแยกองค์ประกอบออกมา เช่น ขอบรูป สี ลักษณะรูปทรง เป็นต้น ก่อนจะนำมาผ่าน Activation function ที่ช่วยแปลงค่าให้อยู่ในผลลัพธ์ที่ดียิ่งขึ้น รวมถึงการทำพูลลิง (Pooling) ที่ช่วยปรับขนาดของข้อมูลให้เล็กลง โดยที่ยังคงรักษารายละเอียดสำคัญของข้อมูลเดิมไว้ และส่วนสุดท้ายจะเป็น Fully connected layer ที่เชื่อมต่อในแต่ละชั้นเข้าด้วยกันเพื่อที่จะนำเข้าสู่กระบวนการที่เรียกว่า Flatten เข้ามาช่วยในการแปลงข้อมูลให้กลายเป็น 1 มิติ

2.2.1.1 การดึงคุณลักษณะ (Feature extraction)

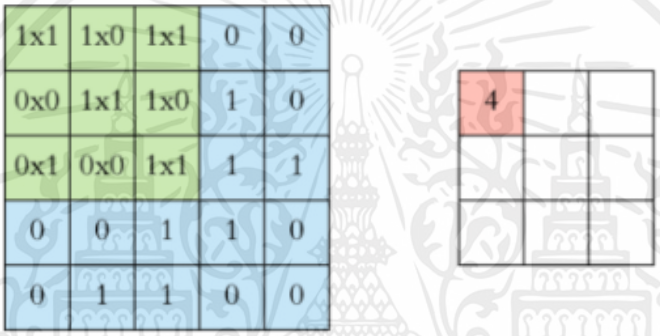
การดึงคุณลักษณะเป็นการเปลี่ยนแปลงขนาดของข้อมูลให้มีขนาดที่เล็กลง และยังคงคุณลักษณะเด่นของข้อมูลเดิมไว้ ซึ่งต้องทำการใช้ตัวกรอง (Filter) หรือเคอร์เนล (Kernel) เพื่อช่วยในการดึงคุณลักษณะที่สนใจออกมา โดยจะถูกแบ่งเป็นส่วนย่อยที่ทำงานร่วมกันอีก 3 ส่วน [17] คือ

1) Convolution layer เป็นกระบวนการที่ใช้สกัดเอาคุณลักษณะเด่นที่สำคัญของรูปภาพออกมา โดยยังคงความสัมพันธ์ของพิกเซลที่อยู่บริเวณพื้นที่ใกล้เคียงเอาไว้ [18] โดยการนำรูปภาพอินพุตมาคูณกับตัวกรองหรือเคอร์เนลที่ทำหน้าที่ดึงคุณลักษณะออกมา ดังรูปที่ 2.11 [19] โดยจะสกัดคุณลักษณะของอินพุตที่สนใจออกมาหรือที่เรียกว่า Feature map ดังตัวอย่างในรูปที่ 2.12 [16]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.11 ตัวอย่างของการทำการคอนโวลูชัน [19]



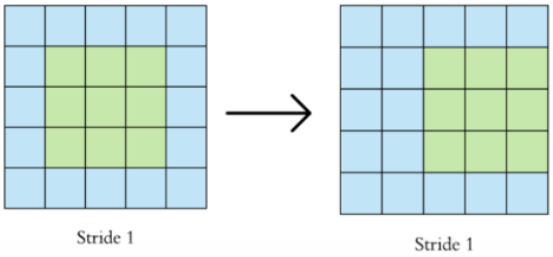
รูปที่ 2.12 การทำ Feature map จากการคอนโวลูชัน [16]

- ลักษณะของตัวกรอง (Filter) เป็นตัวกรองที่ใช้ในการดึงคุณลักษณะที่สนใจของข้อมูล ตัวอย่างของตัวกรองขนาด 3x3 แสดงดังรูปที่ 2.13 [14]

รูปที่ 2.13 ตัวกรองขนาด 3x3 [20]

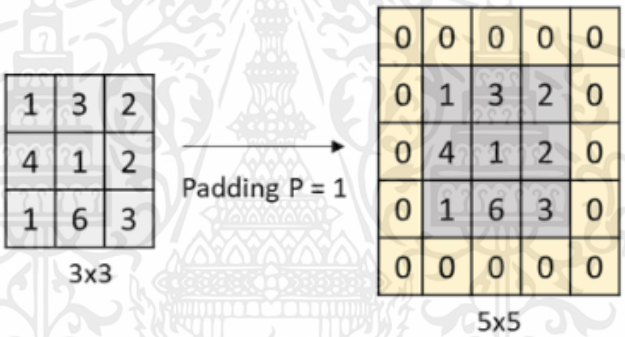
- Stride เป็นค่าที่กำหนดในการเลื่อนตัวกรอง (Filter) ว่าต้องการให้ตัวกรองถูกเลื่อนไปด้วย Step เท่าไร แสดงดังตัวอย่างในรูปที่ 2.14 [16] เมื่อกำหนดให้ Stride เท่ากับ 1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.14 ตัวอย่างในการเลื่อนตัวกรอง (Filter) เมื่อกำหนดให้ Stride เท่ากับ 1 [16]

- Padding เป็นส่วนที่ถูกเติมเข้าไป โดยทั่วไปจะเติมค่าที่เป็น 0 เพื่อให้การทำคอนโวลูชันแล้วมีผลลัพธ์ของ Feature map ยังคงมีขนาดเท่ากับอินพุต ดังรูปที่ 2.15 [17]



รูปที่ 2.15 ตัวอย่างของการทำ Padding โดยการตั้งค่าให้ Padding = 1 [17]

2) ฟังก์ชันเรคติไฟด์เชิงเส้น (Rectified Linear Unit: ReLU) เป็นฟังก์ชันเส้นตรงที่ถูกปรับแก้ไข โดยจะกำหนดให้ max(0, x) คือฟังก์ชันที่ส่งออกค่าอินพุตโดยตรงหากค่าเป็นบวก และค่าที่ต่ำกว่า 0 จะถูกทำให้เป็น 0 ส่งผลให้โมเดลมีประสิทธิภาพและความรวดเร็วมากขึ้น โดยสมการของ ReLU function [21] แสดงดังสมการที่ (2.1)

$$f(x) = \max(0, x) = \begin{cases} 0 & for \quad x \leq 0 \\ x & for \quad x > 0 \end{cases} \tag{2.1}$$

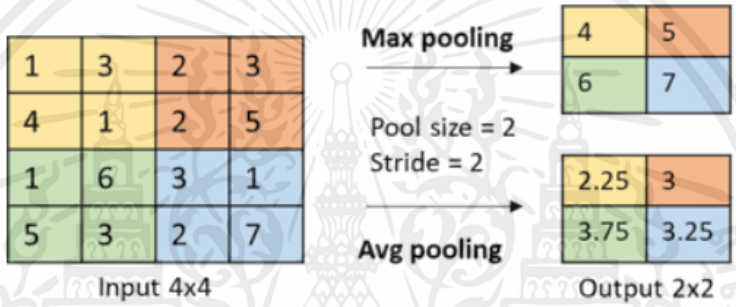
และสมการ Derivative ของ ReLU function ที่ ReLU function = 1 แสดงดังสมการที่ (2.2)

$$f'(x) = \begin{cases} 0 & for \quad x \leq 0 \\ 1 & for \quad x > 0 \end{cases} \tag{2.2}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3) Pooling layer เป็นการช่วยลดมิติของ Feature map ให้มีขนาดเล็กลงแต่ยังคงรักษาข้อมูลที่สำคัญไว้ โดยทั่วไปการ Pooling มีสองแบบคือ Pooling ด้วยค่าสูงสุด (Max pooling) และ Pooling ด้วยค่าเฉลี่ย (Average pooling) นอกจากนี้การทำ Pooling ยังช่วยลดจำนวนของพารามิเตอร์และการคำนวณที่เกินจำเป็นในโครงข่าย [19]

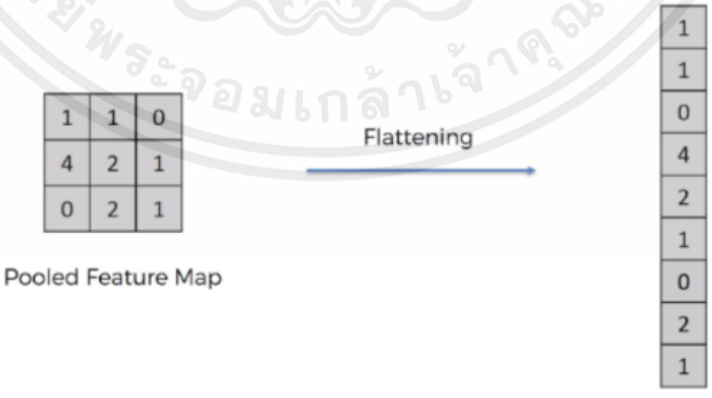
ในกรณีที่ต้องการ Pooling ด้วยค่าสูงสุด (Max pooling) ตัวอย่างการกำหนดให้ตัวกรองมีขนาด 2x2 พิกเซล และทำการเลื่อนที่ละ 2 พิกเซล โดยจะเลือกค่าที่สูงที่สุดในแต่ละ Pool ที่ถูกทาบด้วยตัวกรอง และการทำ Pooling ด้วยค่าเฉลี่ย (Average pooling) ดังรูปที่ 2.16 [19] โดยจะทำการหาค่าเฉลี่ยในแต่ละ Pool ที่ถูกทาบด้วยตัวกรอง



รูปที่ 2.16 ตัวอย่างการทำ Max pooling และ Average pooling [19]

2.2.1.2 การจำแนกประเภท (Classification)

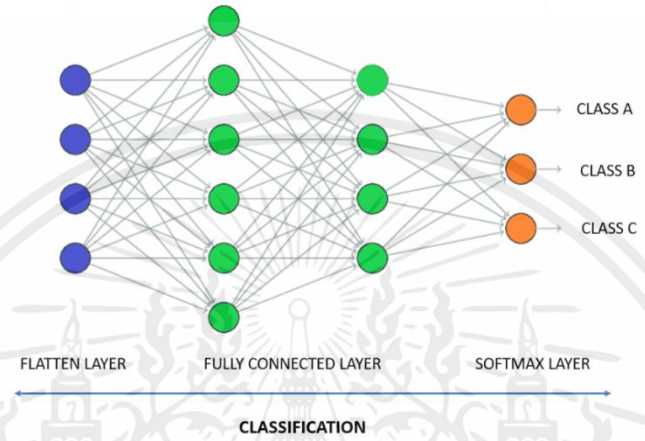
1) Flatten layer เป็นชั้นที่ทำการแปลงข้อมูลจากเมทริกซ์จากหลายมิติ ให้กลายเป็นเวกเตอร์ 1 มิติ ดังตัวอย่างในรูป 2.17 [22]



รูปที่ 2.17 ตัวอย่างการใช้งาน Flatten layer [22]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2) Fully connected layer เป็นขั้นตอนสุดท้ายที่ทำให้ทุกชั้นมีการเชื่อมต่อกันอย่างสมบูรณ์ ผลลัพธ์จากการคอนโวลูชันและ Pooling นั้นให้ลักษณะเด่นของรูปที่รับเข้ามา และในชั้นสุดท้ายนี้จะนำลักษณะเด่นไปทำการคัดกรองรูปที่รับเข้ามาให้อยู่ในรูปของ Classes ดังรูปที่ 2.18 [23]



รูปที่ 2.18 การเชื่อมกันของแต่ละชั้นอย่างสมบูรณ์ใน Fully connected layer [23]

3) ฟังก์ชันซอฟต์แมกซ์ (Soft Max) เป็นฟังก์ชันที่รับอินพุตเป็นเวกเตอร์ของโดเมลโลจิสติคจำนวนจริงแล้ว Normalize ออกมาเป็นความน่าจะเป็นที่ผลรวมเท่ากับ 1 เพื่อแสดงการกระจายของความน่าจะเป็นแบบเลขชี้กำลังระหว่างคลาสต่าง ๆ ใน Output layer [24] โดยสมการของ Soft Max function แสดงดังสมการที่ (2.3)

$$soft\max(\vec{x})_i = \frac{e^{x_i}}{\sum_{j=1}^{n} e^{x_j}} \tag{2.3}$$

เมื่อ $\vec{x}$ คือ เวกเตอร์อินพุต

x_i คือ เวกเตอร์อินพุตตัวที่ i โดยที่ $i = 1 \rightarrow n$

$\sum_{j=1}^{n} e^{x_j}$ คือ ผลรวมของค่าเวกเตอร์อินพุตที่ใช้ในการทำ Normalization เพื่อให้ค่าเอาต์พุตของฟังก์ชันที่ได้มีผลรวมเท่ากับ 1 โดย $x_i = x_j$ เมื่อ $i = j$

n คือ จำนวนผลลัพธ์ที่เป็นไปได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2.2 Training Method

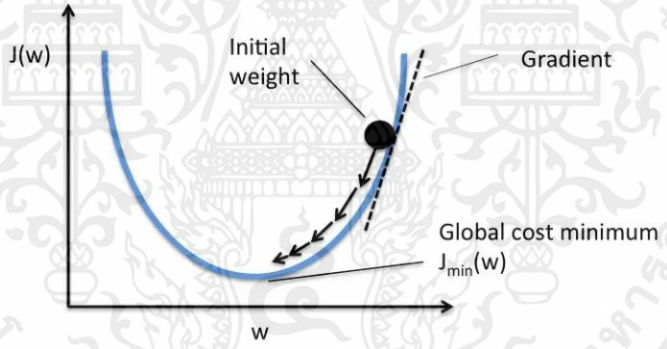
Gradient Descent [25] เป็นอัลกอริธึมที่อัปเดตพารามิเตอร์ w_t โดยใช้การวนซ้ำเพื่อหาค่าที่ต่ำสุดจากการคำนวณจากความชันที่จุดที่เราอยู่ แล้วพยายามเดินทางไปทางตรงข้ามกับความชันที่คำนวณขึ้นมา โดยมีสมการของ Gradient Descent แสดงดังสมการที่ (2.4) และกราฟการคำนวณสมการ Gradient แสดงดังรูปที่ 2.19 [26]

$$w_{t+1} = w_t - \alpha \frac{\partial L}{\partial w_t} \tag{2.4}$$

เมื่อ w_t คือ ค่าของพารามิเตอร์ ณ เวลาปัจจุบัน

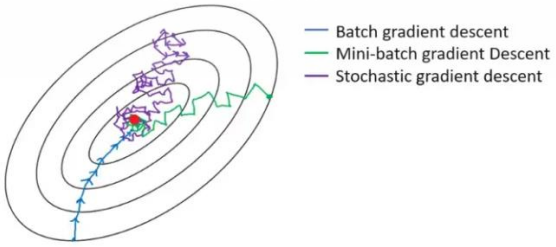
$\frac{\partial L}{\partial w_t}$ คือ ค่า Gradient หรือความชันที่เกิดจากการเปลี่ยนแปลง Loss เมื่อ w_t เปลี่ยน

α คือ Learning rate



รูปที่ 2.19 กราฟการคำนวณสมการ Gradient [26]

Gradient Descent มี 3 ประเภท แสดงดังรูปที่ 2.20 [25]



รูปที่ 2.20 การลู่เข้าของ Gradient Descent ทั้ง 3 ประเภท [25]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2.2.1 Batch Gradient Descent (BGD)

จะใช้ชุดข้อมูลทั้งหมดในการคำนวณเพื่ออัปเดตค่าพารามิเตอร์น้ำหนักตามค่า Gradient ที่จุดปัจจุบัน

2.2.2.2 Stochastic Gradient Descent (SGD)

จะสุ่มตัวอย่าง 1 ตัวจากชุดข้อมูลทั้งหมดมาใช้ในการคำนวณเพื่ออัปเดตค่าพารามิเตอร์น้ำหนักตามค่า Gradient ที่จุดปัจจุบัน

2.2.2.3 Mini-Batch Gradient Descent (MBGD)

เป็นการนำข้อดีของ BGD และ SGD มารวมกัน โดยจะสุ่มตัวอย่าง n ตัวจากชุดข้อมูลทั้งหมดมาใช้ในการคำนวณเพื่ออัปเดตค่าพารามิเตอร์น้ำหนักตามค่า Gradient ที่จุดปัจจุบัน

2.2.3 Loss Function

Loss Function [27] หรือฟังก์ชันการสูญเสียที่ใช้ในการปรับค่าถ่วงน้ำหนักของโมเดลระหว่างการฝึก โดยมีเป้าหมายเพื่อให้โมเดลมีค่า Loss เหลือน้อยที่สุด ซึ่ง Loss Function มีหลายประเภทให้เลือกใช้ตามความเหมาะสมกับปัญหา เช่น Cross-entropy และ MSE

ครอส-เอนโทรปี (Cross-entropy) เป็นฟังก์ชันการสูญเสียที่ใช้บอกประสิทธิภาพของแบบจำลองที่ใช้ในการจำแนก โดยมีสมการของ Cross-entropy สำหรับ 2 Class แสดงดังสมการที่ (2.5) และแบบ Multi-class แสดงดังสมการที่ (2.6)

$$L = -\frac{1}{N}\left[\sum_{j=1}^{N}[t_j \log(p_j) + (1-t_j)\log(1-p_j)]\right] \tag{2.5}$$

$$L(\hat{y}, y) = -\sum_{k}^{K} y^{(k)} \log \hat{y}^{(k)} \tag{2.6}$$

เมื่อ $\hat{y}$ คือ ค่าน้ำหนักในการระบุ Class มีค่า 0 ถึง 1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.2.4 Hyperparameter Tunning

Hyperparameter Tunning [28] เป็นตัวแปรที่มีค่าการเปลี่ยนแปลงตามขนาดของข้อมูล หรือโครงสร้างของ Network

2.2.4.1 Grid Search

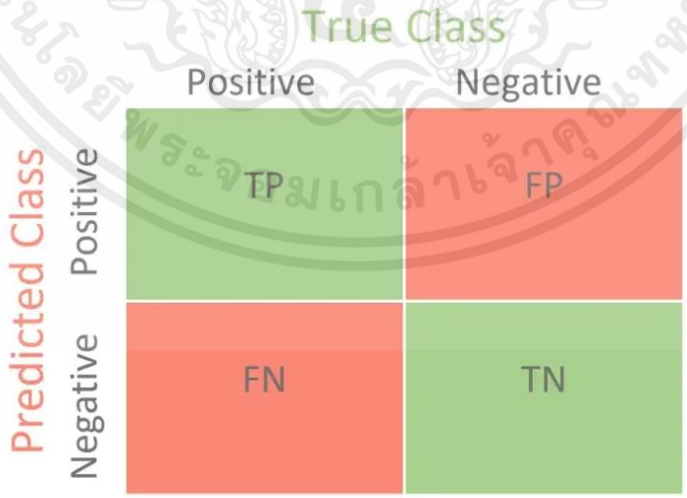
ทดสอบค่า Hyperparameter ของ Input size เพื่อหาค่าขนาดที่เหมาะกับโมเดล และมีประสิทธิภาพที่ได้จากการทดลองปรับค่าเป็นจำนวนมาก ๆ ออกมาเป็นตาราง โดยวิธีการนี้จะเหมาะกับโมเดลขนาดเล็ก

2.2.4.2 Random Search

เป็นการทดสอบโดยการสุ่มค่าของพารามิเตอร์ออกมาจากการกำหนดขอบเขต ทำให้ใช้เวลาน้อยกว่าแบบ Grid Search และยังได้ค่าที่กระจายกว่า

2.3 Confusion Matrix

Confusion Matrix [29] คือ เมทริกซ์ที่เกิดจากการเปรียบเทียบค่าระหว่างค่าที่ใช้ในการทดสอบ และค่าที่คาดการณ์จากโมเดล โดยการนำ Confusion Matrix มาใช้เพื่ออธิบายประสิทธิภาพของโมเดล โดยองค์ประกอบของ Confusion Matrix แสดงดังรูปที่ 2.21 [30] และแสดงตัวอย่าง Multiclass Confusion Matrix ดังรูปที่ 2.22 [31]

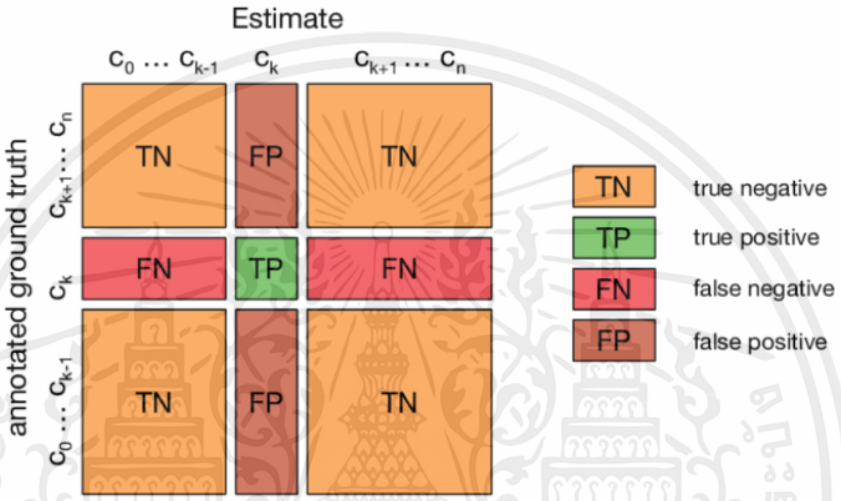


รูปที่ 2.21 Confusion Matrix [30]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูปที่ 2.21 พารามิเตอร์ที่ได้จากการทำ Confusion Matrix คือ

1. True Positive (TP) คือ สิ่งที่โมเดลทำนายว่า จริง และมีค่าเป็น จริง
2. False Negative (FN) คือ สิ่งที่โมเดลทำนายว่า ไม่จริง และมีค่าเป็น จริง
3. False Positive (FP) คือ สิ่งที่โมเดลทำนายว่า จริง และมีค่าเป็น ไม่จริง
4. True Negative (TN) คือ สิ่งที่โมเดลทำนายว่า ไม่จริง แต่มีค่าเป็น ไม่จริง



รูปที่ 2.22 Multiclass Confusion Matrix [31]

F1 Score คือ เป็นค่าเฉลี่ยถ่วงน้ำหนักของ Recall และ Precision ให้อยู่ในสมการเดียวกันเพื่อแก้ปัญหาในการแยกแยะโมเดลที่มีค่า Recall ต่ำและ Precision สูง โดย F1 Score มีค่าสูงสุดที่ 1 และต่ำสุดคือ 0 โดยสมการของ F1 Score ที่เป็นแบบ Average เป็นดังสมการ (2.7)

$$F1Score = \frac{TP}{TP + \frac{1}{2}(FP + FN)} \tag{2.7}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.4 การเพิ่มประสิทธิภาพของโครงข่ายประสาทเทียมแบบคอนโวลูชัน

การเพิ่มประสิทธิภาพของโครงข่ายประสาทเทียมแบบคอนโวลูชัน หรือการ Optimizer [32] เป็น Algorithm ที่ใช้สำหรับลด Loss หรือทำให้โมเดลมีประสิทธิภาพที่ดีขึ้น โดย Optimizer จะไปทำการปรับพารามิเตอร์ที่เรียนรู้ได้ของโมเดล เช่น ค่าถ่วงน้ำหนัก ซึ่งช่วยให้การอัปเดตโมเดลเข้าสู่การลู่เข้าได้เร็วขึ้น โดยปัจจุบันมี Optimizer อยู่หลากหลายให้เลือกใช้ เช่น Momentum, Adagrad, RMSProb, Adam และอื่น ๆ

Adaptive Moment Estimation หรือ Adam [33] เป็น Optimizer ที่สามารถปรับ Learning rates สำหรับพารามิเตอร์ในแต่ละครั้งที่ฝึกสอน ซึ่ง Adam นำข้อดีของ RMSProp และ Momentum มารวมไว้ด้วยกัน โดยสามารถแก้ปัญหา Decaying learning rate ช่วยให้พารามิเตอร์มีการปรับจูนตลอดเวลา และลดปัญหาการแกว่งของพารามิเตอร์ โดยมีสมการของเวกเตอร์ m และ v เพื่อแสดงถึงพารามิเตอร์ที่เก็บ First order gradient และ กำลังสองของ First order gradient ตามลำดับ ซึ่งมีสมการดังสมการที่ (2.7) และ (2.8)

$$m_t = \beta_1 m_{t-1} + (1-\beta_1) g_t \tag{2.7}$$

$$v_t = \beta_2 v_{t-1} + (1-\beta_2) g_t^2 \tag{2.8}$$

เมื่อ m คือค่าเฉลี่ยเลขชี้กำลังของการไล่ระดับสี

v คือค่าเฉลี่ยเลขชี้กำลังยกกำลังสองของการไล่ระดับสี

g คือ Gradient

β คือไฮเปอร์พารามิเตอร์ โดยมีค่า default ที่ 0.9 และ 0.999 ตามลำดับ

2.5 RGB

RGB [34] ย่อมาจาก Red, Green และ Blue คือระบบสีของแสงที่เกิดจากการหักเหของแสงกลายเป็นสีรุ้งด้วยกัน 7 สี โดยมีแสงสีม่วงที่มีความถี่สูงสุดเรียกว่า Ultraviolet และแสงสีแดงที่มีความถี่ต่ำสุดถูกเรียกว่า Infrared ซึ่งคลื่นแสงที่มีความถี่สูงกว่าสีม่วง และต่ำกว่าสีแดงนั้นสายตามนุษย์ไม่สามารถรับได้ทำ แสงสีทั้งหมดเกิดจากแสงสี 3 สีคือ สีแดง สีน้ำเงิน และสีเขียว ที่ถูกเรียกว่าแม่สีของแสง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.6 การแปลงภาพสีเป็นภาพระดับสีเทา (Color to Grayscale image)

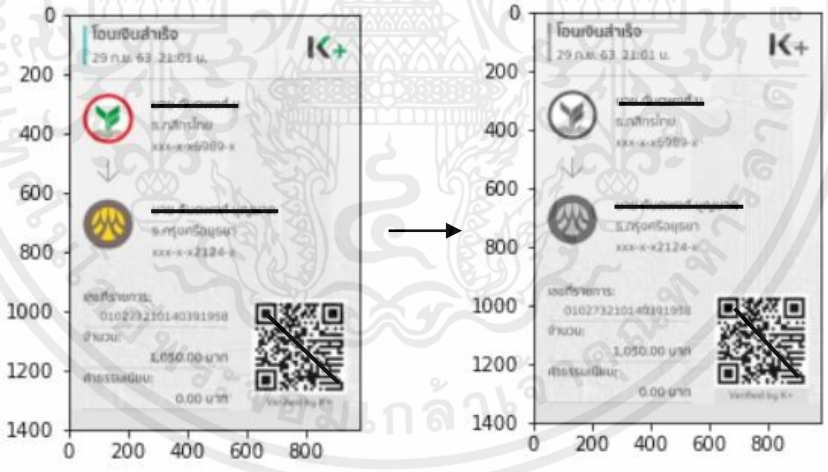
การแปลงภาพสีเป็นภาพระดับสีเทา หรือ Color to Grayscale เป็นการแปลงภาพสีให้เป็นภาพระดับสีเทา (Grayscale) [35] ที่ทำให้ภาพที่มีเฉดสี RGB กลายเป็นภาพระดับสีเทา ซึ่งช่วยให้ลดเวลาในการประมวลผลภาพ เนื่องจากข้อมูลมีขนาดเล็กลงมาเพราะเฉดสีของข้อมูลมีตั้งแต่สีขาวไปจนถึงดำ จากเดิมที่เป็นภาพ RGB ซึ่งช่วยให้สามารถดำเนินการประมวลผลที่ซับซ้อนได้มากขึ้นในเวลาที่สั้นลง โดยมีสมการการแปลงภาพระดับสีเทาดังสมการที่ (2.9) โดยตัวอย่างหลังการแปลงภาพระดับสีเทาแสดงดังรูปที่ 2.23

$$Gray = 0.299R + 0.587G + 0.114B \tag{2.9}$$

เมื่อ R คือค่าความเข้มของสีแดง มีค่า 0-255

G คือค่าความเข้มของสีเขียว มีค่า 0-255

B คือค่าความเข้มของสีน้ำเงิน มีค่า 0-255



รูปที่ 2.23 รูปภาพก่อน และหลังการแปลงเฉดสีเป็นภาพระดับสีเทา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.7 ภาษา Python

ภาษาโปรแกรม Python [36] เป็นภาษาโปรแกรมคอมพิวเตอร์ระดับสูง โดยถูกออกแบบมาให้เป็นภาษาสคริปต์ที่อ่านง่าย โดยตัดความซับซ้อนของโครงสร้าง และไวยกรณ์ของภาษาออกไป ในส่วนของการแปลงชุดคำสั่งที่เราเขียนให้เป็นภาษาเครื่อง Python มีการทำงานแบบ Interpreter คือเป็นการแปลชุดคำสั่งทีละบรรทัด เพื่อป้อนเข้าสู่หน่วยประมวลผลให้คอมพิวเตอร์ทำงานตามที่เราต้องการ นอกจากนั้นภาษาโปรแกรม Python ยังสามารถนำไปใช้ในการเขียนโปรแกรมได้หลากหลายประเภท โดยไม่ได้จำกัดอยู่ที่งานเฉพาะทางใดทางหนึ่ง (General-purpose language) จึงทำให้มีการนำไปใช้กันแพร่หลายในหลายองค์กรใหญ่ระดับโลก เช่น Google, YouTube, Instagram, Dropbox และ NASA เป็นต้น

2.7.1 OpenCV-Python

OpenCV ย่อมาจาก Opensource Computer Vision หรือก็คือ Computer Vision Library ที่สามารถใช้งานได้ฟรี โดยความสามารถในการทำงานของไลบรารีคือสามารถช่วยจัดการกับวีดีโอหรือรูปภาพ เพื่อให้เกิดการเปลี่ยนแปลงกับรูปภาพ อย่างเช่น ทำให้ภาพชัดขึ้น ทำให้ภาพเบลอ รวมถึงช่วยในการลดสัญญาณรบกวน (Noise) ในรูปภาพ และไม่เพียงแค่นั้น ยังมีเทคนิคที่ใช้ในการประมวลผลภาพที่เรียกว่าการแบ่งประเภทของวัตถุอีกด้วย

2.7.2 Keras

Keras เป็นไลบรารีที่เกิดจากการวิจัยของ Google Lab ที่ได้คิดค้น และพัฒนา Platform ใหม่ขึ้นมาเพื่อเรียกใช้งาน TensorFlow ในเบื้องหลังแทน อันเนื่องมาจากความซับซ้อนในการใช้งาน TensorFlow ทำให้สามารถเขียนโปรแกรมผ่าน Keras ได้โดยตรง โดย Keras เป็นไลบรารีที่เหมาะสำหรับการทำ Machine Learning และ Deep Learning ที่ใช้งานได้ง่าย แต่มีประสิทธิภาพสูงในการประมวลผลโมเดล เนื่องจาก Back-end ของ Keras มีทั้ง TensorFlow และ Theano ที่เป็น Deep Learning Library ที่มีประสิทธิภาพสูงทั้งคู่ ทำให้ Keras ของ Python นั้นรองรับการประมวลข้อมูล และความยืดหยุ่นในการใช้งานที่สูงมากกว่า

2.7.3 Pytesseract

เป็นไลบรารี่บน Python ที่เรียกใช้งานเครื่องมือที่เรียกใช้งาน Tesseract OCR Engine ใช้ในการรู้จำตัวอักษรหรือที่รู้จักกัน คือ Optical Character Recognition หรือ OCR ซึ่ง Tesseract มีความสามารถในการเรียนรู้จดจำได้มากกว่า 100 ภาษา และยังเป็น Open Source ที่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เปิดโอกาสให้ดาวโหลด แก้ไข ดัดแปลง และเผยแพร่ได้อย่างอิสระ โดยในปัจจุบันทาง Google ได้ใช้ Tesseract ในการทำงาน Text Detection กับวีดีโอ และตรวจจับสแปมรูปภาพทาง Gmail และยังสามารถสแกนเพื่อแปลงไฟล์รูปภาพเป็นไฟล์เอกสารหรือข้อความ

2.7.4 Pandas

เป็นไลบรารี่ที่มีเครื่องมือในการจัดการ และวิเคราะห์ข้อมูลที่มีประสิทธิภาพสูงเหมาะกับจัดการข้อมูลที่มีรูปแบบหลากหลาย และสามารถจัดการกับข้อมูลที่เป็นตาราง (Data Frame) ได้อย่างมีประสิทธิภาพ

2.7.5 Streamlit

Streamlit เป็นไลบรารี่ที่ถูกนำมาใช้งานในการพัฒนาเว็บแอปพลิเคชัน และมีฟังก์ชันที่เหมาะสมในการใช้งานทางด้านแสดงผลในรูปแบบของ Data Visualization และเหมาะกับการนำมาใช้งานร่วมกับ Machine Learning Model

2.7.6 Plotly

เป็นไลบรารี่ในการแสดงผลข้อมูลกราฟในรูปแบบต่าง ๆ เป็นโมดูลสร้างกราฟรายงานข้อมูลทั้งแบบ 2 มิติ และแบบ 3 มิติ

2.7.7 Datetime

เป็นไลบรารี่ในการนำเข้าข้อมูลเวลา และจัดการกับข้อมูลที่อยู่ในรูปของเวลา เช่นการแปลงหน่วยเวลา

2.8 Visual Studio Code

Visual Studio Code (VS Code) [37] เป็นโปรแกรมแก้ไขซอร์สโค้ดที่สามารถทำงานบนเดสก์ท็อป และใช้สำหรับ Windows, macOS และ Linux ที่สนับสนุนภาษาโปรแกรมหลากหลายภาษา เช่น C++, C#, Java, Python, PHP และ Go โดยในปริญญานิพนธ์ฉบับนี้ได้นำ VS Code มาเพื่อใช้งานร่วมกับภาษา Python.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.9 Jupyter Notebook

Jupyter Notebook [38] คือ เครื่องมือที่ใช้เขียน Code ในการสร้าง Machine Learning Model และยังเป็นที่นิยมสำหรับสายงานทางด้าน Data Science ที่เหมาะกับการทำงานที่เกี่ยวกับการจัดเก็บข้อมูลที่มีปริมาณมาก

2.10 User Interface

User Interface หรือ UI [39] คือส่วนที่เชื่อมต่อกับผู้ใช้งาน (User) เพื่อให้ผู้ใช้งานสามารถเข้าถึงระบบถูกได้ถูกพัฒนาขึ้น โดยมุ่งเน้นไปที่การออกแบบหน้าตาของเว็บไซต์ให้มีความสวยงาม หรือออกแบบให้ดึงดูดใจ และใช้งานได้ง่าย เช่น ขนาดตัวอักษร ปุ่มกด การวางภาพ เป็นต้น โดยในปัจจุบันการออกแบบ UI ต้องคำนึงถึง User Experience หรือ UX ที่เป็นหลักการที่ทำให้ผู้ใช้งานสามารถใช้งานเว็บไซต์ได้ง่าย ไม่สับสน

2.11 Data Visualization

Data Visualization หรือการสรุปข้อมูล และแสดงออกมาเป็นแผนภาพ กราฟ ที่สามารถอธิบายถึงข้อมูลที่ต้องการจะแสดง [40] ถูกนำมาใช้ร่วมกับการประมวลผล วิเคราะห์ข้อมูล หรือ Data Analysis เพื่อให้ข้อมูลสามารถจับต้องได้ง่ายขึ้น หรือเห็นภาพได้ชัดเจน และเข้าใจง่ายกว่าเดิม โดยการนำเสนอในรูปของ Data Visualization สามารถช่วยจัดการเรื่องความซับซ้อนและยุ่งยากของข้อมูลให้มีประสิทธิภาพ ตัวอย่างของการแสดงผลแบบ Data Visualization แสดงดังรูปที่ 2.24 [40]

รูปที่ 2.24 ข้อมูลที่ถูกนำเสนอแบบ Data Visualization [40]

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

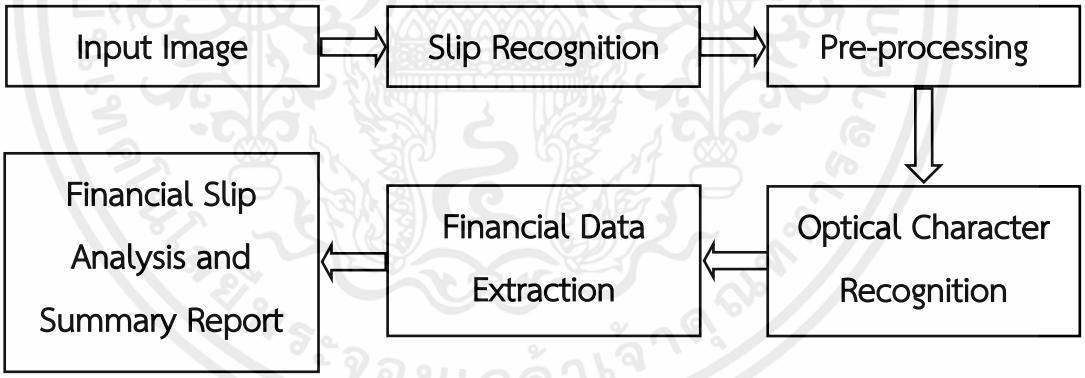
บทที่ 3

การออกแบบและการจัดทำปริญญานิพนธ์

3.1 การออกแบบ

3.1.1 การออกแบบระแบบ

บล็อกไดอะแกรมของการเรียนรู้เชิงลึกสำหรับการตรวจสอบสลิปการเงินแสดงได้ดังรูปที่ 3.1 โดยชุดข้อมูลที่ได้จากการเก็บรวบรวมรูปภาพสลิปการเงิน และรูปภาพชนิดอื่นจะถูกจำแนกโดยใช้วิธีการเรียนรู้เชิงลึกเพื่อคัดแยกรูปภาพสลิปการเงินทั้งหมดออกมาจากรูปภาพชนิดอื่น หลังจากนั้นรูปภาพสลิปการเงินจะถูกนำเข้าสู่กระบวนการเตรียมข้อมูลเพื่อแบ่งประเภทสลิปการเงินตามธนาคาร และแบ่งชนิดของรูปภาพสลิปการเงินโดยผลลัพธ์ที่ได้จะถูกนำไปเข้าสู่กระบวนการอ่านอักขระด้วยแสง OCR เพื่อดึงข้อมูลที่สนใจออกมาจากรูปภาพสลิปการเงิน โดยนำข้อมูลที่ได้ไปประมวลผลเพื่อจัดหมวดหมู่ข้อมูลในรูปแบบ Data Frame และแสดงผลข้อมูลในรูปแบบ Data Visualization ผ่านเว็บแอปพลิเคชัน

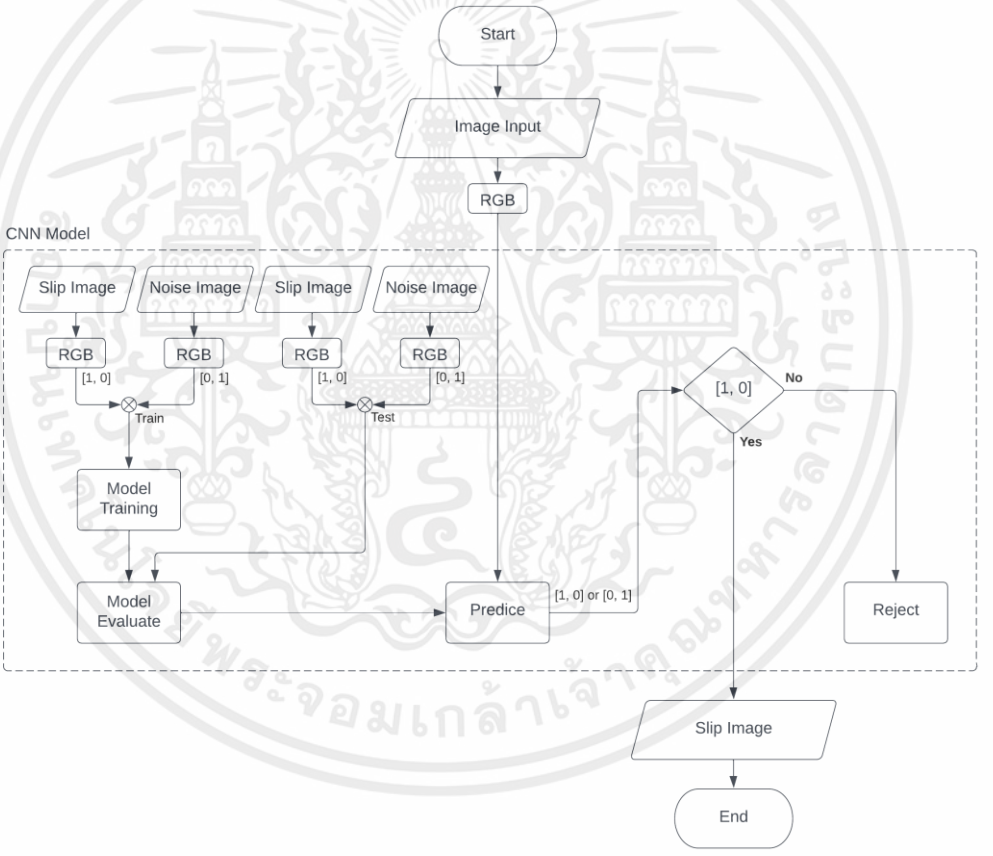


รูปที่ 3.1 บล็อกไดอะแกรมของการตรวจสอบสลิปการเงินด้วยวิธีการเรียนรู้เชิงลึก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.1.2 การออกแบบโมเดลแยกรูปภาพสลิปการเงินออกจากรูปภาพชนิดอื่น

ขั้นตอนแรก คือการทดสอบหาค่า Hyperparameter ของ Input size เพื่อหาค่า Input size ที่มีประสิทธิภาพในการเรียนรู้ของการเรียนรู้เชิงลึก โดยวิธีการ Grid search โดยกำหนดขนาดเป็น 8x8 pixels, 16x16 pixels, 28x28 pixel, 32x32 pixels, 64x64 pixel และ 128x128 pixel โดยขั้นตอนต่อไปเป็นการนำรูปภาพทั้งหมดเข้าสู่ CNN Model ที่ผ่านการฝึกสอนและทดสอบประสิทธิภาพ เพื่อแยกรูปภาพที่ไม่สนใจออกจากรูปภาพที่สนใจมีเอาต์พุตที่ได้คือรูปภาพสลิปการเงิน โดยแผนผังการแยกรูปภาพสลิปการเงินออกจากรูปภาพชนิดอื่นแสดงได้ดังรูปที่ 3.2



รูปที่ 3.2 แผนผังการแยกรูปภาพสลิปการเงินออกจากรูปภาพชนิดอื่น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.1.2.1 การกำหนดข้อมูลสำหรับฝึกสอน และทดสอบโมเดล

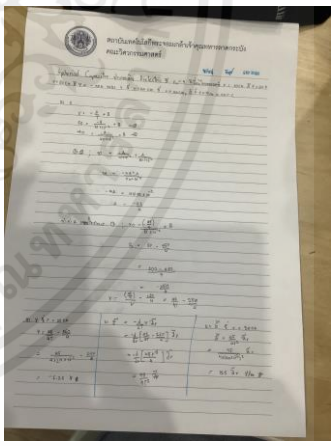
ในขั้นตอนการฝึกสอน และทดสอบโมเดลจำเป็นต้องทำการแบ่งชุดข้อมูลสำหรับใช้การฝึกสอน และทดสอบโมเดลเพื่อใช้ในการทดสอบความแม่นยำของโมเดล โดยทำการแบ่งชุดข้อมูลออกเป็น 3 ส่วน [41] คือ Training set, Validation set, Test set โดยมีอัตราส่วนของ Training set เป็น 80% และ Test set 20% ในส่วนของ Validation set จะทำการแบ่งมาจาก Training set 20% โดยชุดข้อมูลสำหรับฝึกสอน และทดสอบโมเดลแสดงได้ดังตารางที่ 3.1

ตารางที่ 3.1 ชุดข้อมูลสำหรับใช้ฝึกสอน และทดสอบโมเดล สำหรับการจำแนกรูปภาพสลิปการเงินออกจากรูปภาพชนิดอื่น

Data set		Training set 80 % (image)	Test set 20 % (image)	Sample picture
สลิปการเงิน	อิเล็กทรอนิกส์	7,200	1,800	
	กระดาษ	1,200	300	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

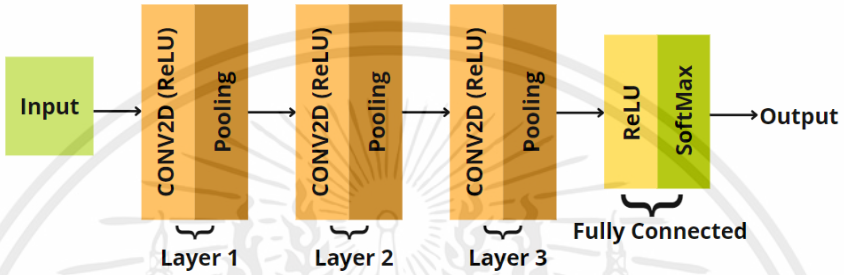
ตารางที่ 3.1 ชุดข้อมูลสำหรับใช้ฝึกสอน และทดสอบโมเดล สำหรับการจำแนกรูปภาพสลิปการเงินออกจากรูปภาพชนิดอื่น (ต่อ)

Data set		Training set 80 % (image)	Test set 20 % (image)	Sample picture
รูปภาพชนิดอื่น	คล้ายสลิป	4,000	1,000	
	ไม่คล้ายสลิป	4,000	1,000	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.1.2.2 CNN Model สำหรับแยกรูปภาพสลิปการเงินออกจากรูปภาพชนิดอื่น

ในการแยกรูปภาพสลิปการเงินออกจากรูปภาพชนิดอื่น ได้ทำการออกแบบ CNN Model เป็นจำนวน 5 layer ประกอบไปด้วยส่วนที่เป็น Convolution มี ReLU 3 layer และส่วนที่เป็น Fully Connected มี ReLU กับ SoftMax อย่างละ 1 layer แสดงดังรูปที่ 3.3



รูปที่ 3.3 Layer ของ CNN Model สำหรับแยกรูปภาพสลิปการเงินและรูปภาพชนิดอื่น

จากรูป 3.3 แต่ละ Layer ของ CNN ที่ออกแบบมีคุณสมบัติดังแสดงดังต่อไปนี้ โดยมีตัวอย่างของพารามิเตอร์ที่กำหนดให้ขนาดของอินพุตเข้ามีขนาด 28x28 ดังรูปที่ 3.4

1) Input size: ข้อมูลรูปภาพสลิปการเงินที่ได้จากการทดลองปรับขนาด 8x8 pixels, 16x16 pixels, 28x28 pixel, 32x32 pixels, 64x64 pixel, 128x128 pixel ที่เหมาะสมที่สุด

2) Layer 1-3: เป็นส่วนของ Convolution ที่ประกอบไปด้วย Kernel size ขนาด 3x3 pixels ที่มี Step ในการเลือกคือ 1 pixels โดยการเติม Padding 0 เข้าไป และมีใช้ฟังก์ชันของ ReLU ในการคำนวณก่อนส่งไปยังการทำ Max pooling เพื่อลดจำนวนพารามิเตอร์ที่ใช้ ซึ่งออกแบบขนาดการ Max pooling ให้มีขนาด 2x2 pixels และเลื่อนที่ละ 2-pixel แบบไม่เติม Padding ทำให้มีขนาดลง 4 เท่า

3) Fully Connected: เริ่มจากการทำ Flatten เพื่อจัดเรียงข้อมูลให้เป็นเวกเตอร์ 1 มิติ หลังจากนั้นนำเข้าสู่ฟังก์ชัน ReLU เพื่อคำนวณ และส่งไปยังฟังก์ชันของ SoftMax เพื่อทำการคำนวณหาค่าในการทำนาย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4) Output: จะแสดงค่า 0-1 โดยผลรวมของความน่าจะเป็นของคลาสทั้งหมดต้องเท่ากับ ซึ่งค่าที่มากที่สุดจะถูกนำมาเป็นการระบุคลาสของข้อมูลที่ใช้ทำนาย

5) การเพิ่มประสิทธิภาพประกอบด้วย Optimizer ใช้ Adam ในส่วนของ Loss Function ใช้ Cross entropy

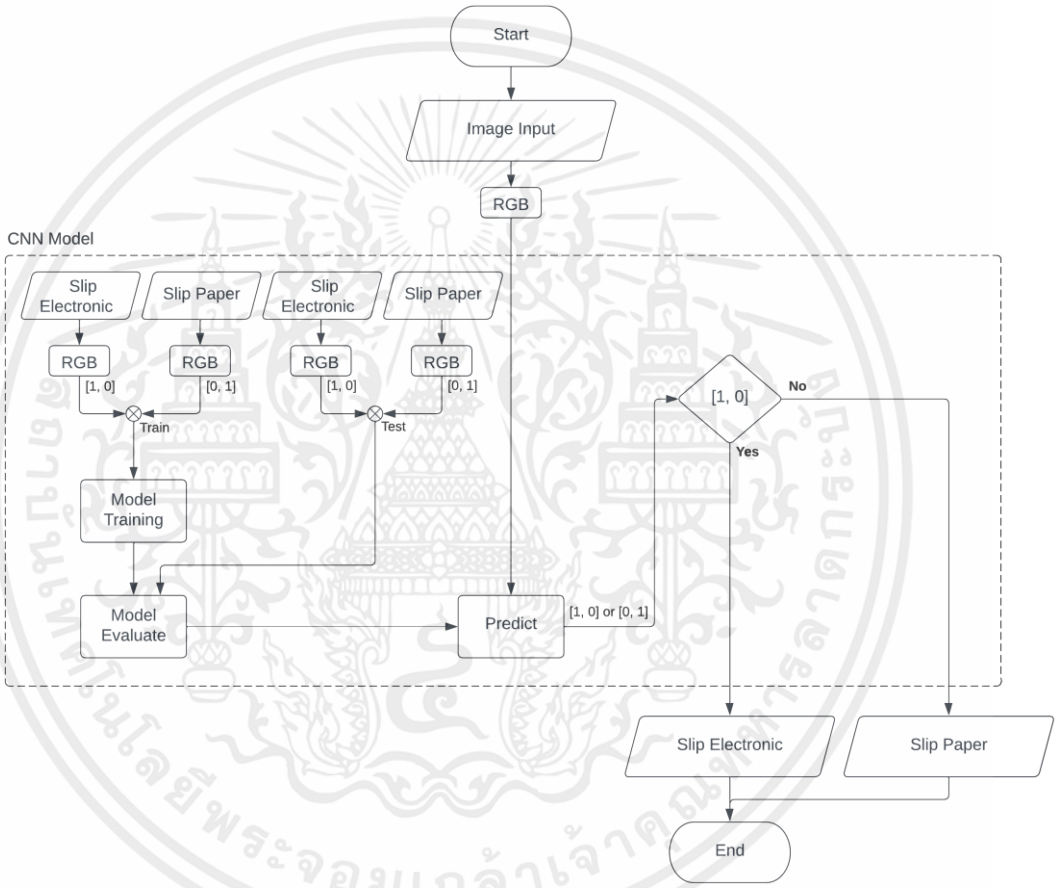
```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d_14 (Conv2D)           (None, 28, 28, 32)        896

max_pooling2d_14 (MaxPoolin  (None, 14, 14, 32)        0
g2D)

conv2d_15 (Conv2D)           (None, 14, 14, 128)       36992

max_pooling2d_15 (MaxPoolin  (None, 7, 7, 128)         0
g2D)

conv2d_16 (Conv2D)           (None, 7, 7, 256)         295168

max_pooling2d_16 (MaxPoolin  (None, 3, 3, 256)         0
g2D)

dropout_8 (Dropout)          (None, 3, 3, 256)         0

flatten_4 (Flatten)          (None, 2304)              0

dense_8 (Dense)              (None, 128)               295040

dropout_9 (Dropout)          (None, 128)               0

dense_9 (Dense)              (None, 2)                 258

=================================================================
Total params: 628,354
Trainable params: 628,354
Non-trainable params: 0
_________________________________________________________________
```

รูปที่ 3.4 พารามิเตอร์ในแต่ละ Layer ของโมเดลที่ใช้ในการแยกรูปภาพสลิปการเงิน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.1.3 การออกแบบโมเดลแยกชนิดของสลิปการเงิน

ในขั้นตอนทำการแบ่งแยกชนิดของสลิปการเงินตามธนาคารโดยใช้ CNN Model ที่ผ่านการฝึกสอน และทดสอบประสิทธิภาพ โดยนำอินพุตรูปภาพสลิปการเงินเข้าสู่ CNN Model ได้ผลลัพธ์คือรูปภาพสลิปการเงินที่ถูกแบ่งแยกชนิดของสลิปการเงินเพื่อง่ายต่อการกำหนดรูปแบบการดึงข้อมูล แสดงแผนผังการทำงานของโมเดลแยกชนิดของสลิปการเงินได้ดังรูปที่ 3.5



รูปที่ 3.5 แผนผังการแยกชนิดของสลิปการเงิน

3.1.3.1 การกำหนดข้อมูลสำหรับฝึกสอน และทดสอบโมเดล

ในขั้นตอนนี้ใช้มีอัตราส่วนของ Training set เป็น 80% และ Test set 20% ในส่วนของ Validation set จะทำการแบ่งมาจาก Training set 20% โดยชุดข้อมูลสำหรับฝึกสอน และทดสอบโมเดลแสดงได้ดังตารางที่ 3.2

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

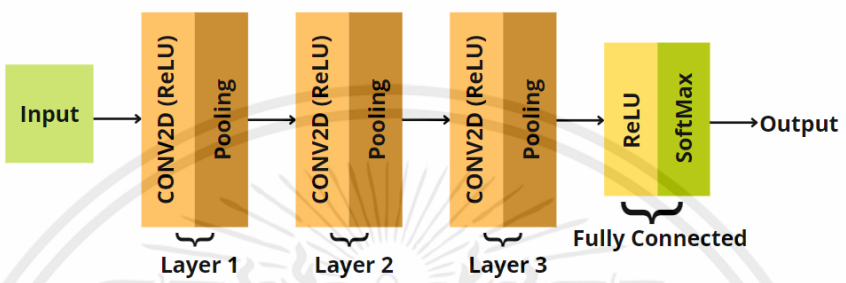
ตารางที่ 3.2 ชุดข้อมูลสำหรับใช้ฝึกสอน และทดสอบโมเดล สำหรับแยกชนิดของรูปภาพสลิปการเงิน

Data set	Training set 80 % (image)	Test set 20 % (image)	Sample picture
สลิปการเงินแบบอิเล็กทรอนิกส์	1,200	300	
สลิปการเงินแบบกระดาษ	1,200	300	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.1.3.2 CNN Model สำหรับแยกชนิดของสลิปการเงิน

ในการแยกชนิดของสลิปการเงิน ได้ทำการออกแบบ CNN Model เป็นจำนวน 5 layer ประกอบไปด้วยส่วนที่เป็น Convolution มี ReLU 3 layer และส่วนที่เป็น Fully Connected มี ReLU กับ SoftMax อย่างละ 1 layer แสดงดังรูปที่ 3.6



รูปที่ 3.6 Layer ของ CNN Model สำหรับแยกชนิดของสลิปการเงิน

จากรูป 3.6 แต่ละ Layer ของ CNN ที่ออกแบบมีคุณสมบัติดังแสดงดังต่อไปนี้ โดยมีตัวอย่างของพารามิเตอร์ที่กำหนดให้ขนาดของอินพุตเข้ามีขนาด 28x28 ดังรูปที่ 3.7

1) Input: ข้อมูลรูปภาพสลิปการเงินที่ได้จากการทดลองปรับขนาด 8x8 pixels, 16x16 pixels, 28x28 pixel, 32x32 pixels, 64x64 pixel, 128x128 pixel หน่วยเป็น pixels ที่เหมาะสมที่สุด

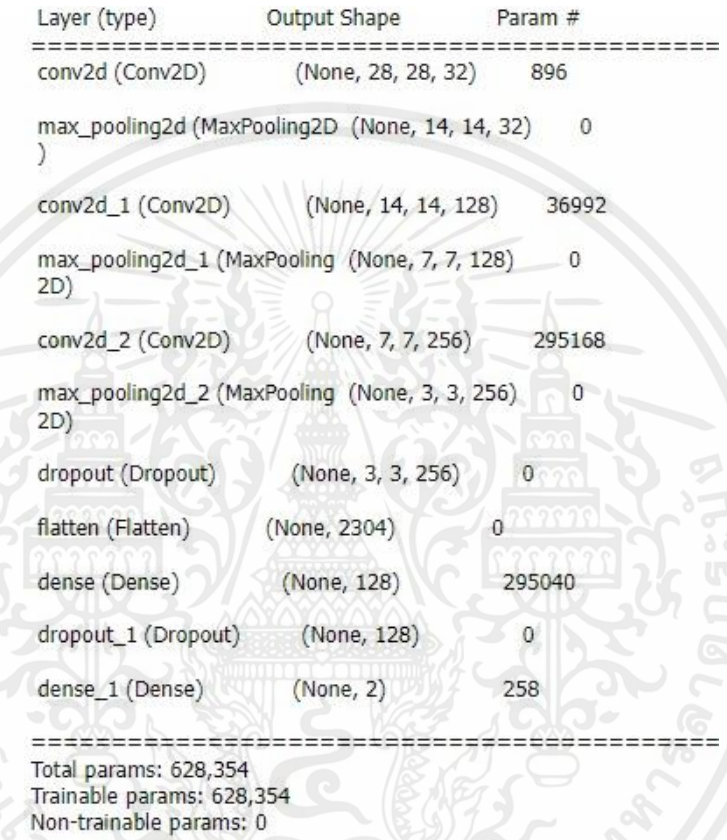
2) Layer 1-3: เป็นส่วนของ Convolution ที่ประกอบไปด้วย Kernel size ขนาด 3x3 pixels ที่มี Step ในการเลือกคือ 1 pixels โดยการเติม Padding 0 เข้าไป และมีใช้ฟังก์ชันของ ReLU ในการคำนวณก่อนส่งไปยังการทำ Max pooling เพื่อลดจำนวนพารามิเตอร์ที่ใช้ ซึ่งออกแบบขนาดการ Max pooling ให้มีขนาด 2x2 pixels และเลื่อนที่ละ 2-pixel แบบไม่เติม Padding ทำให้มีขนาดลง 4 เท่า

3) Fully Connected: เริ่มจากการทำ Flatten เพื่อจัดเรียงข้อมูลให้เป็นเวกเตอร์ 1 มิติ หลังจากนั้นนำเข้าสู่ฟังก์ชัน ReLU เพื่อคำนวณ และส่งไปยังฟังก์ชันของ SoftMax เพื่อทำการคำนวณหาค่าในการทำนาย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4) Output: จะแสดงค่า 0-1 โดยผลรวมของความน่าจะเป็นของคลาสทั้งหมดต้องเท่ากับ ซึ่งค่าที่มากที่สุดจะถูกนำมาเป็นการระบุคลาสของข้อมูลที่ใช้ทำนาย

5) การเพิ่มประสิทธิภาพประกอบด้วย Optimizer ใช้ Adam ในส่วนของ Loss Function ใช้ Cross entropy



Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 28, 28, 32)	896
max_pooling2d (MaxPooling2D)	(None, 14, 14, 32)	0
conv2d_1 (Conv2D)	(None, 14, 14, 128)	36992
max_pooling2d_1 (MaxPooling2D)	(None, 7, 7, 128)	0
conv2d_2 (Conv2D)	(None, 7, 7, 256)	295168
max_pooling2d_2 (MaxPooling2D)	(None, 3, 3, 256)	0
dropout (Dropout)	(None, 3, 3, 256)	0
flatten (Flatten)	(None, 2304)	0
dense (Dense)	(None, 128)	295040
dropout_1 (Dropout)	(None, 128)	0
dense_1 (Dense)	(None, 2)	258

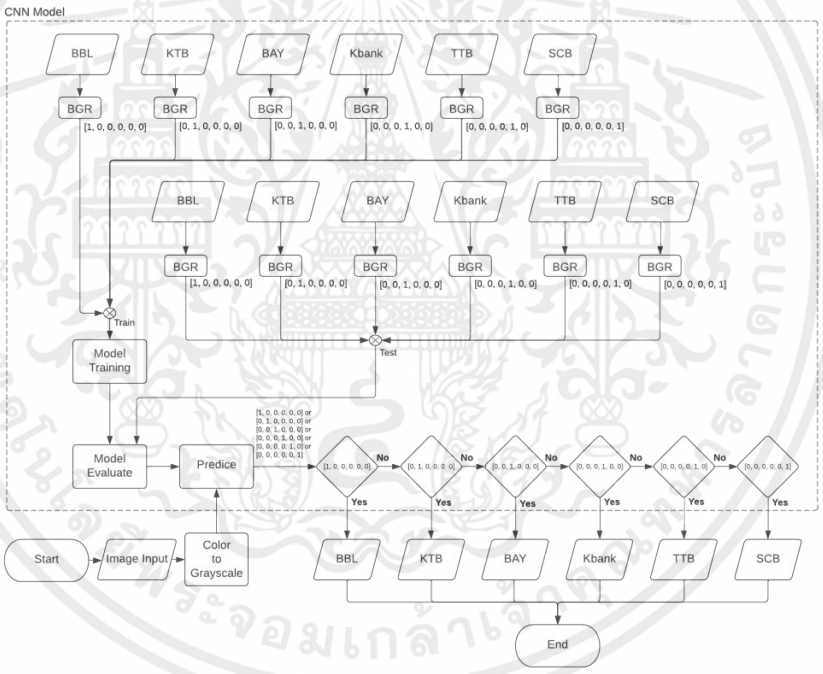
Total params: 628,354
Trainable params: 628,354
Non-trainable params: 0

รูปที่ 3.7 พารามิเตอร์ในแต่ละ Layer ของโมเดลที่ใช้แยกชนิดของสลิปการเงิน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.1.4 การออกแบบโมเดลแบ่งประเภทของสลิปการเงินตามธนาคารของสลิปการเงินแบบอิเล็กทรอนิกส์

ในขั้นตอนนี้ทำการแบ่งประเภทของสลิปการเงินตามธนาคารของสลิปการเงินแบบอิเล็กทรอนิกส์โดยใช้ CNN Model ที่ผ่านการฝึกสอน และทดสอบประสิทธิภาพ โดยขั้นตอนแรกจะนำอินพุตรูปภาพสลิปการเงินมาแปลงเป็นรูปภาพระดับสีเทา เพื่อเพิ่มความรวดเร็วในการประมวลผลข้อมูล และนำผลลัพธ์ที่ได้เข้าสู่ CNN Model เพื่อแบ่งประเภทของสลิปการเงินตามธนาคารแบบอิเล็กทรอนิกส์ โดยแผนผังการแบ่งประเภทของสลิปการเงินตามธนาคารแสดงได้ดังรูปที่ 3.8



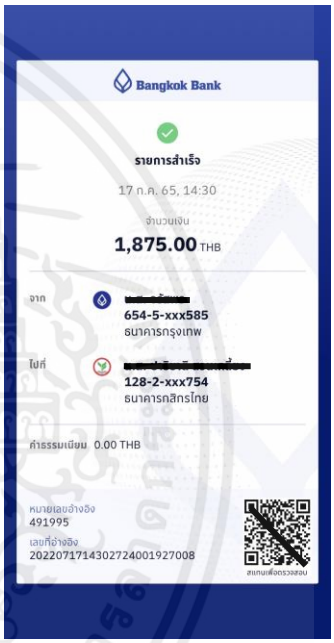
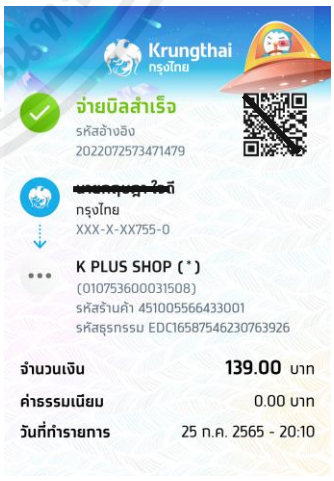
รูปที่ 3.8 แผนผังการแบ่งประเภทของสลิปการเงินตามธนาคารของสลิปการเงินแบบอิเล็กทรอนิกส์

3.1.4.1 การกำหนดข้อมูลสำหรับฝึกสอน และทดสอบโมเดล

ในขั้นตอนนี้ใช้มีอัตราส่วนของ Training set เป็น 80% และ Test set 20% ในส่วนของ Validation set จะทำการแบ่งมาจาก Training set 20% โดยชุดข้อมูลสำหรับฝึกสอน และทดสอบโมเดลแสดงได้ดังตารางที่ 3.3

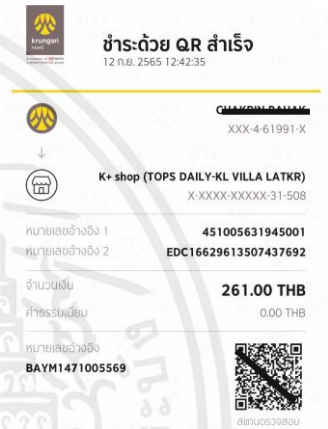
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.3 ชุดข้อมูลสำหรับใช้ฝึกสอน และทดสอบโมเดล สำหรับแบ่งประเภทของรูปภาพสลิปการเงินตามธนาคารของสลิปการเงินแบบอิเล็กทรอนิกส์

Data set		Training set 80 % (image)	Test set 20 % (image)	Sample picture
ธนาคารกรุงเทพ	อิเล็กทรอนิกส์	1,200	300	
ธนาคารกรุงไทย	อิเล็กทรอนิกส์	1,200	300	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.3 ชุดข้อมูลสำหรับใช้ฝึกสอน และทดสอบโมเดล สำหรับแบ่งประเภทของรูปภาพสลิปการเงินตามธนาคารของสลิปการเงินแบบอิเล็กทรอนิกส์ (ต่อ)

Data set		Training set 80 % (image)	Test set 20 % (image)	Sample picture
ธนาคารกรุงศรีอยุธยา	อิเล็กทรอนิกส์	1,200	300	
ธนาคารกสิกรไทย	อิเล็กทรอนิกส์	1,200	300	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.3 ชุดข้อมูลสำหรับใช้ฝึกสอน และทดสอบโมเดล สำหรับแบ่งประเภทของรูปภาพสลิปการเงินตามธนาคารของสลิปการเงินแบบอิเล็กทรอนิกส์ (ต่อ)

Data set		Training set 80 % (image)	Test set 20 % (image)	Sample picture
ธนาคารไทยธนชาติ	อิเล็กทรอนิกส์	1,200	300	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

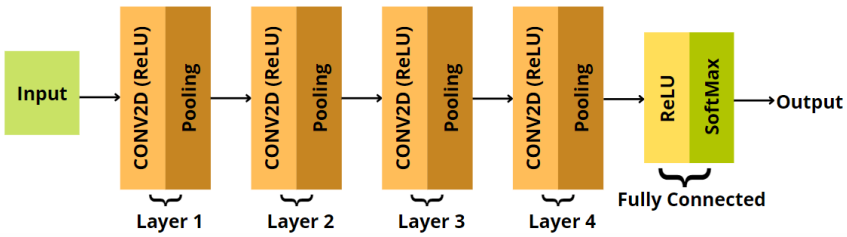
ตารางที่ 3.3 ชุดข้อมูลสำหรับใช้ฝึกสอน และทดสอบโมเดล สำหรับแบ่งประเภทของรูปภาพสลิปการเงินตามธนาคารของสลิปการเงินแบบอิเล็กทรอนิกส์ (ต่อ)

Data set		Training set 80 % (image)	Test set 20 % (image)	Sample picture
ธนาคารไทยธนชาติ	อิเล็กทรอนิกส์	1,200	300	

3.1.4.2 CNN Model สำหรับแบ่งประเภทสลิปตามธนาคารของสลิปการเงินแบบอิเล็กทรอนิกส์

ในการแบ่งประเภทสลิปตามธนาคารได้ทำการออกแบบ CNN Model เป็นจำนวน 6 layer ประกอบไปด้วยส่วนที่เป็น Convolution มี ReLU จำนวน 4 layer และส่วนที่เป็น Fully Connected มี ReLU กับ SoftMax อย่างละ 1 layer แสดงดังรูปที่ 3.9

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.9 Layer ของ CNN Model สำหรับแบ่งประเภทสลิปตามธนาคารของสลิปการเงินแบบอิเล็กทรอนิกส์

จากรูป 3.9 แต่ละ Layer ของ CNN ที่ออกแบบมีคุณสมบัติดังแสดงดังต่อไปนี้ โดยมีตัวอย่างของพารามิเตอร์ที่กำหนดให้ขนาดของอินพุตเข้ามีขนาด 28x28 ดังรูปที่ 3.10

1) Input: ข้อมูลรูปภาพสลิปการเงินที่ได้จากการทดลองปรับขนาด 8x8 pixels, 16x16 pixels, 28x28 pixel, 32x32 pixels, 64x64 pixel, 128x128 pixel หน่วยเป็น pixels ที่เหมาะสมที่สุด

2) Layer 1-4: เป็นส่วนของ Convolution ที่ประกอบไปด้วย Kernel size ขนาด 3x3 pixels ที่มี Step ในการเลือกคือ 1 pixels โดยการเติม Padding 0 เข้าไป และมีใช้ฟังก์ชันของ ReLU ในการคำนวณก่อนส่งไปยังการทำ Max pooling เพื่อลดจำนวนพารามิเตอร์ที่ใช้ ซึ่งออกแบบขนาดการ Max pooling ให้มีขนาด 2x2 pixels และเลื่อนที่ละ 2-pixel แบบไม่เติม Padding ทำให้มีขนาดลง 4 เท่า

3) Fully Connected: เริ่มจากการทำ Flatten เพื่อจัดเรียงข้อมูลให้เป็นเวกเตอร์ 1 มิติ หลังจากนั้นนำเข้าสู่ฟังก์ชัน ReLU เพื่อคำนวณ และส่งไปยังฟังก์ชันของ SoftMax เพื่อทำการคำนวณหาค่าในการทำนาย

4) Output: จะแสดงค่า 0-1 โดยผลรวมของความน่าจะเป็นของคลาสทั้งหมดต้องเท่ากับ ซึ่งค่าที่มากที่สุดจะถูกนำมาเป็นการระบุคลาสของข้อมูลที่ใช้ทำนาย

5) การเพิ่มประสิทธิภาพประกอบด้วย Optimizer ใช้ Adam ในส่วนของ Loss Function ใช้ Cross entropy

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Layer (type)	Output Shape	Param #
conv2d_43 (Conv2D)	(None, 28, 28, 32)	896
max_pooling2d_43 (MaxPooling2D)	(None, 14, 14, 32)	0
dropout_31 (Dropout)	(None, 14, 14, 32)	0
conv2d_44 (Conv2D)	(None, 14, 14, 128)	36992
max_pooling2d_44 (MaxPooling2D)	(None, 7, 7, 128)	0
dropout_32 (Dropout)	(None, 7, 7, 128)	0
conv2d_45 (Conv2D)	(None, 7, 7, 256)	295168
max_pooling2d_45 (MaxPooling2D)	(None, 3, 3, 256)	0
dropout_33 (Dropout)	(None, 3, 3, 256)	0
conv2d_46 (Conv2D)	(None, 3, 3, 512)	1180160
max_pooling2d_46 (MaxPooling2D)	(None, 1, 1, 512)	0
dropout_34 (Dropout)	(None, 1, 1, 512)	0
flatten_12 (Flatten)	(None, 512)	0
dense_24 (Dense)	(None, 128)	65664
dropout_35 (Dropout)	(None, 128)	0
dense_25 (Dense)	(None, 6)	774

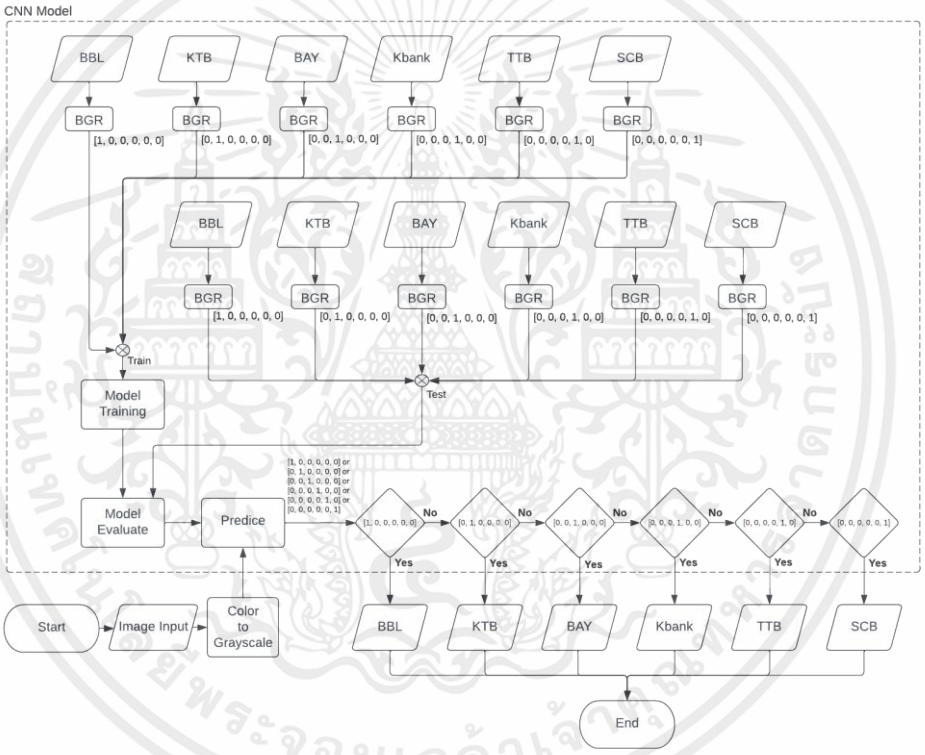
Total params: 1,579,654
Trainable params: 1,579,654
Non-trainable params: 0

รูปที่ 3.10 พารามิเตอร์ในแต่ละ Layer ของโมเดลที่ใช้สำหรับแบ่งประเภทสลิปตามธนาคารของสลิปการเงินแบบอิเล็กทรอนิกส์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.1.5 การออกแบบโมเดลแบ่งประเภทของสลิปการเงินตามธนาคารของสลิปการเงินแบบกระดาษ

ในขั้นตอนนี้จะทำการแบ่งประเภทของสลิปการเงินตามธนาคารโดยใช้ CNN Model ที่ผ่านการฝึกสอน และทดสอบประสิทธิภาพ โดยขั้นตอนแรกจะนำอินพุตรูปภาพสลิปการเงินแบบกระดาษมาแปลงเป็นรูปภาพระดับสีเทา เพื่อเพิ่มความรวดเร็วในการประมวลผลข้อมูล และนำผลลัพธ์ที่ได้เข้าสู่ CNN Model เพื่อแบ่งประเภทของสลิปการเงินตามธนาคารของสลิปการเงินแบบกระดาษ โดยแผนผังการแบ่งประเภทของสลิปการเงินตามธนาคารแสดงได้ดังรูปที่ 3.11



รูปที่ 3.11 แผนผังการแบ่งประเภทของสลิปการเงินตามธนาคารของสลิปการเงินแบบกระดาษ

3.1.5.1 การกำหนดข้อมูลสำหรับฝึกสอน และทดสอบโมเดล

ในขั้นตอนนี้ใช้มีอัตราส่วนของ Training set เป็น 80% และ Test set 20% ในส่วนของ Validation set จะทำการแบ่งมาจาก Training set 20% โดยชุดข้อมูลสำหรับฝึกสอน และทดสอบโมเดลแสดงได้ดังตารางที่ 3.4

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.4 ชุดข้อมูลสำหรับใช้ฝึกสอน และทดสอบโมเดล สำหรับแบ่งประเภทของรูปภาพสลิปการเงินตามธนาคารของสลิปการเงินแบบกระดาษ

Data set		Training set 80 % (image)	Test set 20 % (image)	Sample picture
ธนาคารกรุงเทพ	กระดาษ	200	50	
ธนาคารกรุงไทย	กระดาษ	200	50	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.4 ชุดข้อมูลสำหรับใช้ฝึกสอน และทดสอบโมเดล สำหรับแบ่งประเภทของรูปภาพสลิปการเงินตามธนาคารของสลิปการเงินแบบกระดาษ (ต่อ)

Data set		Training set 80 % (image)	Test set 20 % (image)	Sample picture
ธนาคารกรุงศรีอยุธยา	กระดาษ	200	50	
ธนาคารกสิกรไทย	กระดาษ	200	50	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

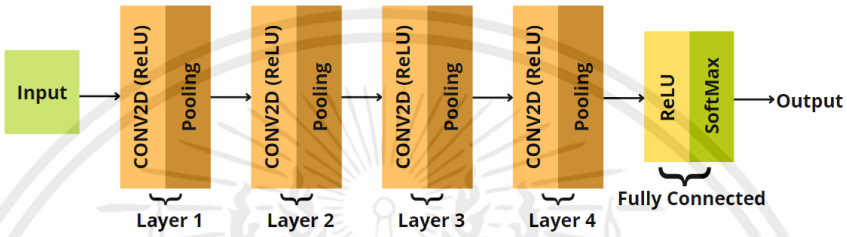
ตารางที่ 3.4 ชุดข้อมูลสำหรับใช้ฝึกสอน และทดสอบโมเดล สำหรับแบ่งประเภทของรูปภาพสลิปการเงินตามธนาคารของสลิปการเงินแบบกระดาษ (ต่อ)

Data set		Training set 80 % (image)	Test set 20 % (image)	Sample picture
ธนาคารไทยธนชาติ	กระดาษ	200	50	
ธนาคารไทยพาณิชย์	กระดาษ	200	50	

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.1.5.2 CNN Model สำหรับแบ่งประเภทสลิปตามธนาคารของสลิปการเงินแบบกระดาษ

ในการแบ่งประเภทสลิปตามธนาคารได้ทำการออกแบบ CNN Model เป็นจำนวน 6 layer ประกอบไปด้วยส่วนที่เป็น Convolution มี ReLU จำนวน 4 layer และส่วนที่เป็น Fully Connected มี ReLU กับ SoftMax อย่างละ 1 layer แสดงดังรูปที่ 3.12



รูปที่ 3.12 Layer ของ CNN Model สำหรับแบ่งประเภทสลิปตามธนาคารของสลิปการเงินแบบกระดาษ

จากรูป 3.12 แต่ละ Layer ของ CNN ที่ออกแบบมีคุณสมบัติดังแสดงดังต่อไปนี้ โดยมีตัวอย่างของพารามิเตอร์ที่กำหนดให้ขนาดของอินพุตเข้ามีขนาด 28x28 ดังรูปที่ 3.13

1) Input: ข้อมูลรูปภาพสลิปการเงินที่ได้จากการทดลองปรับขนาด 8x8 pixels, 16x16 pixels, 28x28 pixel, 32x32 pixels, 64x64 pixel, 128x128 pixel หน่วยเป็น pixels ที่เหมาะสมที่สุด

2) Layer 1-4: เป็นส่วนของ Convolution ที่ประกอบไปด้วย Kernel size ขนาด 3x3 pixels ที่มี Step ในการเลือกคือ 1 pixels โดยการเติม Padding 0 เข้าไป และมีใช้ฟังก์ชันของ ReLU ในการคำนวณก่อนส่งไปยังการทำ Max pooling เพื่อลดจำนวนพารามิเตอร์ที่ใช้ ซึ่งออกแบบขนาดการ Max pooling ให้มีขนาด 2x2 pixels และเลื่อนที่ละ 2-pixel แบบไม่เติม Padding ทำให้มีขนาดลง 4 เท่า

3) Fully Connected: เริ่มจากการทำ Flatten เพื่อจัดเรียงข้อมูลให้เป็นเวกเตอร์ 1 มิติ หลังจากนั้นนำเข้าสู่ฟังก์ชัน ReLU เพื่อคำนวณ และส่งไปยังฟังก์ชันของ SoftMax เพื่อทำการคำนวณหาค่าในการทำนาย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4) Output: จะแสดงค่า 0-1 โดยผลรวมของความน่าจะเป็นของคลาสทั้งหมดต้องเท่ากับ ซึ่งค่าที่มากที่สุดจะถูกนำมาเป็นการระบุคลาสของข้อมูลที่ใช้ทำนาย

5) การเพิ่มประสิทธิภาพประกอบด้วย Optimizer ใช้ Adam ในส่วนของ Loss Function ใช้ Cross entropy

Layer (type)	Output Shape	Param #
conv2d_43 (Conv2D)	(None, 28, 28, 32)	896
max_pooling2d_43 (MaxPooling2D)	(None, 14, 14, 32)	0
dropout_31 (Dropout)	(None, 14, 14, 32)	0
conv2d_44 (Conv2D)	(None, 14, 14, 128)	36992
max_pooling2d_44 (MaxPooling2D)	(None, 7, 7, 128)	0
dropout_32 (Dropout)	(None, 7, 7, 128)	0
conv2d_45 (Conv2D)	(None, 7, 7, 256)	295168
max_pooling2d_45 (MaxPooling2D)	(None, 3, 3, 256)	0
dropout_33 (Dropout)	(None, 3, 3, 256)	0
conv2d_46 (Conv2D)	(None, 3, 3, 512)	1180160
max_pooling2d_46 (MaxPooling2D)	(None, 1, 1, 512)	0
dropout_34 (Dropout)	(None, 1, 1, 512)	0
flatten_12 (Flatten)	(None, 512)	0
dense_24 (Dense)	(None, 128)	65664
dropout_35 (Dropout)	(None, 128)	0
dense_25 (Dense)	(None, 6)	774

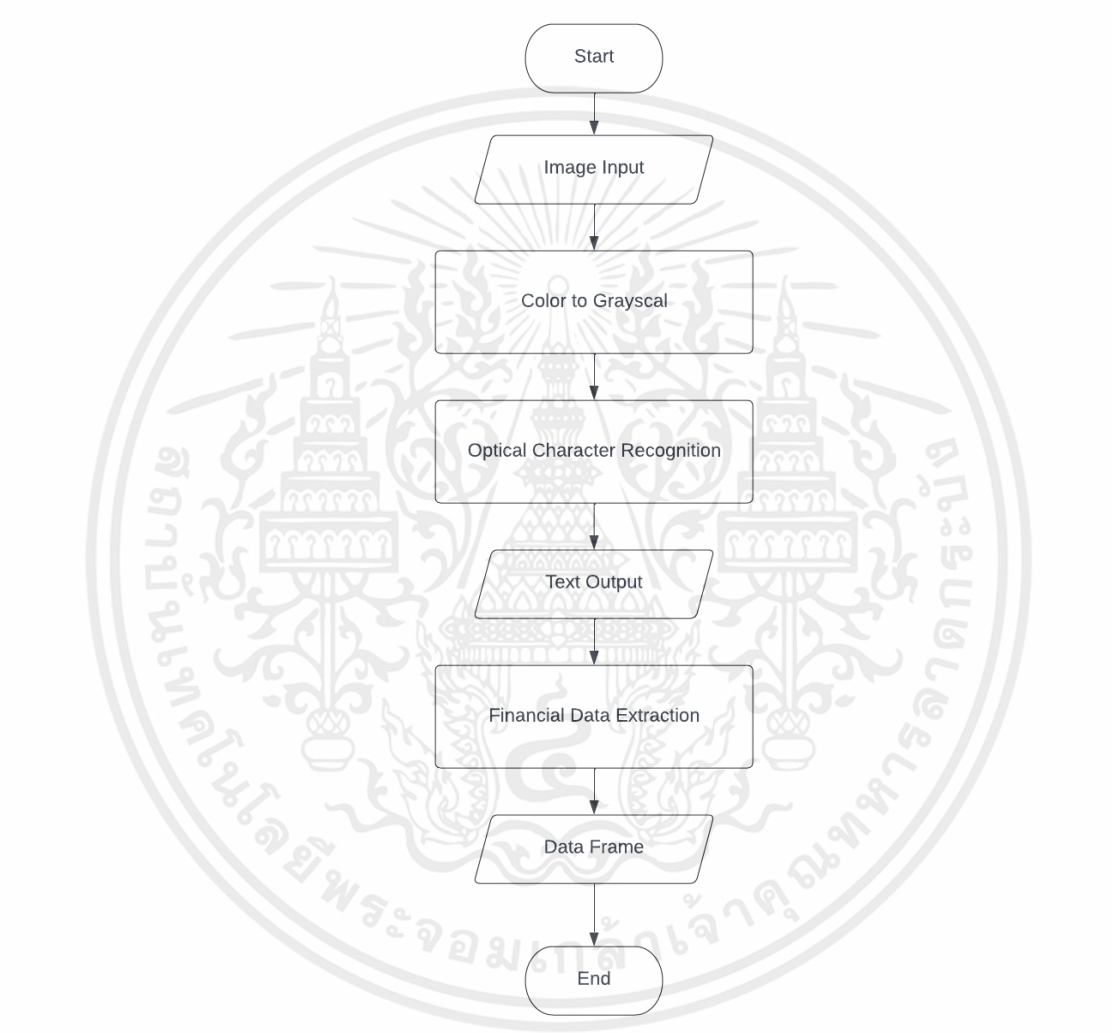
Total params: 1,579,654
Trainable params: 1,579,654
Non-trainable params: 0

รูปที่ 3.13 พารามิเตอร์ในแต่ละ Layer ของโมเดลที่ใช้ในการแยกรูปภาพสลิปการเงินของสลิปการเงินแบบกระดาษ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.1.6 การออกแบบการดึงข้อมูลออกจากรูปภาพสลิปการเงิน

ในการดึงข้อมูลออกจากรูปภาพสลิปการเงินจะใช้วิธีการอ่านอักขระด้วยแสง OCR โดยมีอินพุตเป็นรูปภาพสลิปการเงิน และเมื่อผ่านกระบวนการนี้จะได้เอาต์พุตเป็นข้อความหรือ Text โดยจัดเก็บข้อมูลในรูปของ Data Frame โดยแผนผังการดึงข้อมูลแสดงได้ดังรูปที่ 3.14

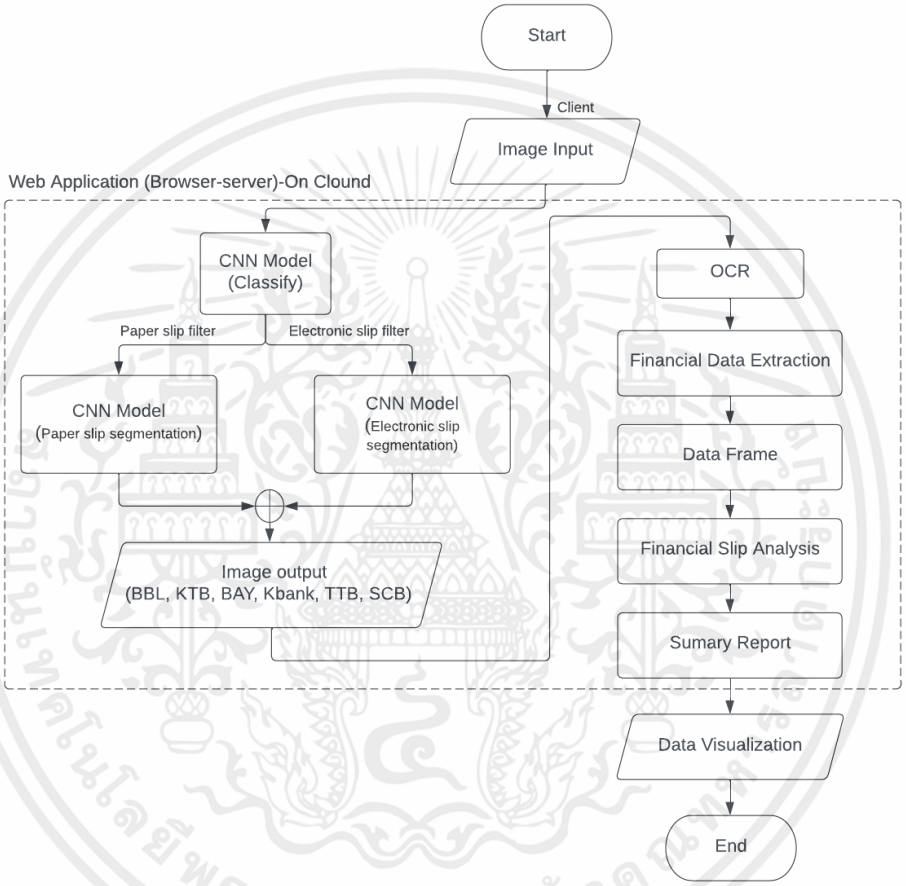


รูปที่ 3.14 แผนผังการดึงข้อมูลออกจากรูปภาพสลิปการเงิน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.1.7 การออกแบบการแสดงผลข้อมูล

ในขั้นตอนการแสดงผลข้อมูล คือ ข้อมูลจะถูกวิเคราะห์ข้อมูล และสรุปผลโดยจะแสดงข้อมูลดังกล่าวผ่านเว็บแอปพลิเคชันในรูปแบบของ Data Visualization แสดงแผนผังการแสดงผลข้อมูลได้ดังรูปที่ 3.15

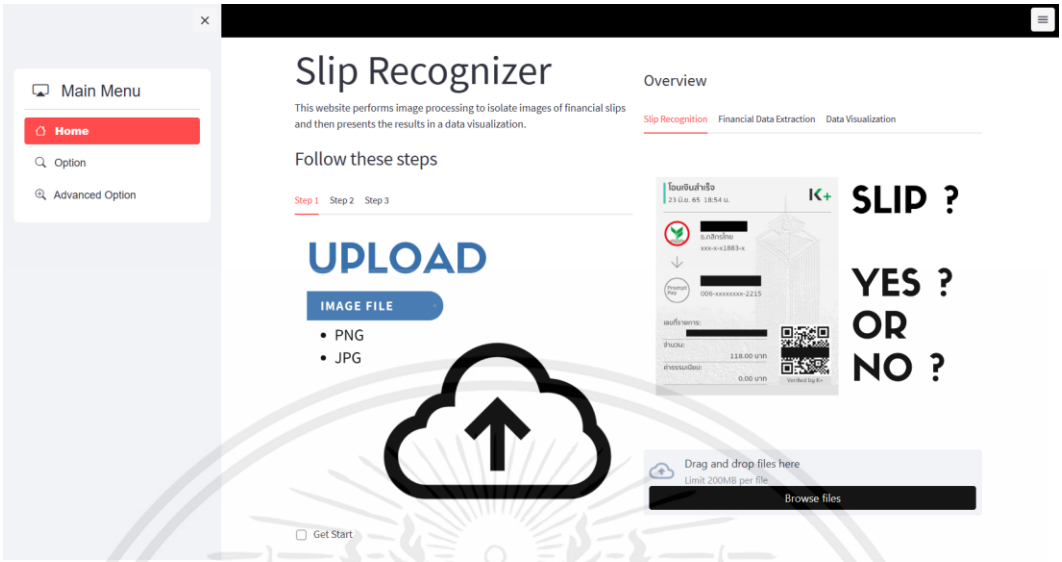


รูปที่ 3.15 แผนผังการแสดงผลข้อมูล

3.1.7.1 ออกแบบเว็บแอปพลิเคชันที่ใช้แสดงผล

ในขั้นตอนนี้ใช้ Python ในการสร้างเว็บแอปพลิเคชันเพื่อแสดงผลข้อมูล แสดงได้ดังรูปที่ 3.16 และ 3.17 โดยออกแบบให้มีองค์ประกอบทั้งหมด 5 ส่วน คือ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



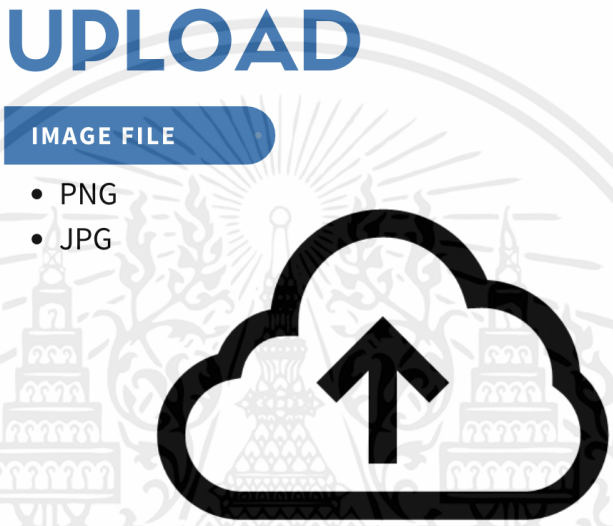
รูปที่ 3.16 เว็บแอปพลิเคชันที่ได้ออกแบบ

รูปที่ 3.17 เว็บแอปพลิเคชันที่ได้ออกแบบ

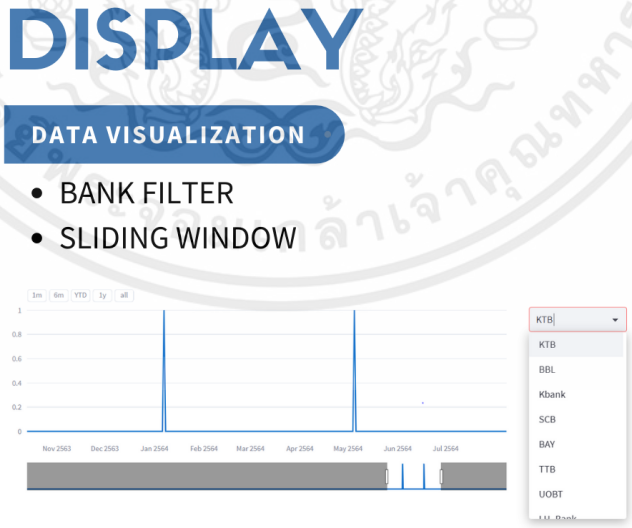
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1) สอนการใช้งานเว็บแอปพลิเคชัน

ในส่วนของการสอนใช้งานเว็บแอปพลิเคชันได้สอนการใช้งาน Feature ต่าง ๆ โดยแนะนำการอัปโหลดข้อมูลแสดงดังรูปที่ 3.18 คู่มือการใช้งานกราฟแสดงวันเวลาที่ทำธุรกรรมพร้อมกับ Feature แสดงดังรูปที่ 3.19 และส่วนที่แนะนำ Filter ให้ User ได้รู้จักตัว Filter แสดงดังรูปที่ 3.20

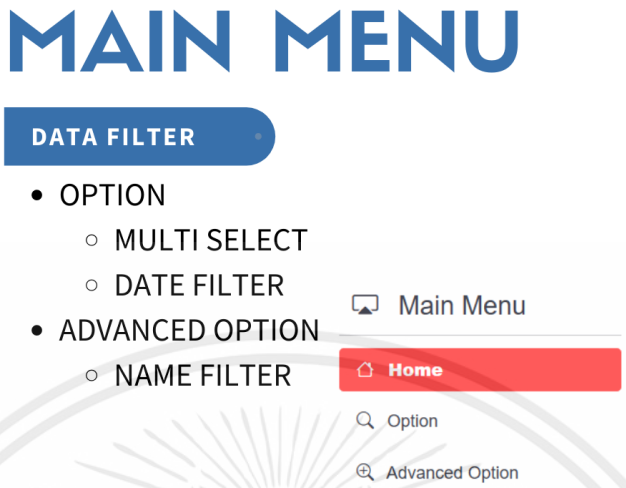


รูปที่ 3.18 แนะนำการอัปโหลดข้อมูล



รูปที่ 3.19 คู่มือการใช้งานกราฟแสดงวันเวลาที่ทำธุรกรรม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



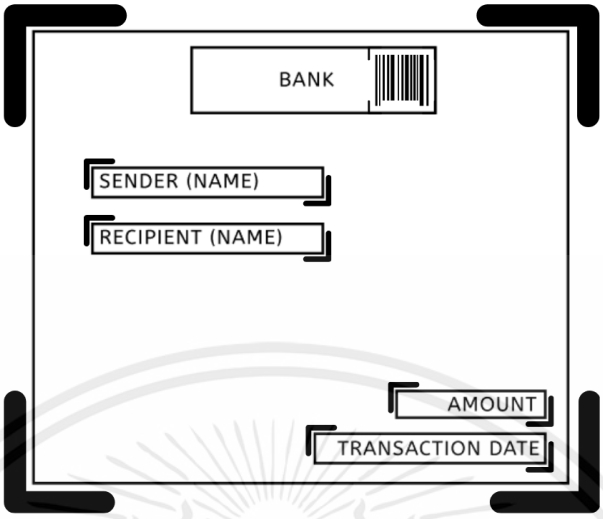
รูปที่ 3.20 แนะนำ Filter

2) อธิบายภาพรวมของปริญญานิพนธ์

ในส่วนของการออกแบบอธิบายภาพรวมการทำงานทั้งหมดเป็น 3 ส่วน คือ Slip Recognition แสดงดังรูปที่ 3.21 ส่วนของ Financial Data Extraction แสดงดังรูปที่ 3.22 และส่วนการนำเสนอข้อมูลที่ถูกนำเสนอด้วยการทำ Data Visualization แสดงดังรูปที่ 3.23

รูปที่ 3.21 Slip Recognition

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



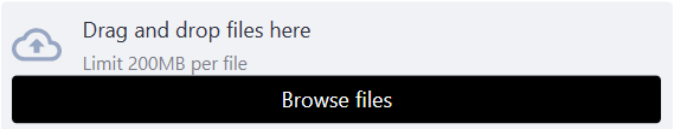
รูปที่ 3.22 Financial Data Extraction

รูปที่ 3.23 Data Visualization

3) Upload file

ส่วนที่ใช้ในการอัปโหลดไฟล์รูปภาพที่เป็นอินพุตของระบบแสดง ดังรูปที่ 3.24

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

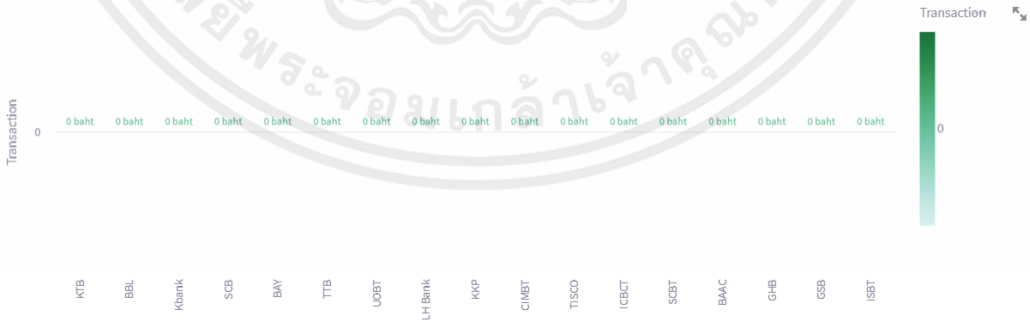


รูปที่ 3.24 Upload file

4) แสดงผลของข้อมูลสลิปการเงิน (Data Visualization)

ในส่วนของการแสดงข้อมูลได้ถูกแบ่งออกเป็น 2 ส่วน คือ ส่วนที่เป็นกราฟแสดงวันเวลาที่ทำธุรกรรม โดยออกแบบ Feature สำหรับ Filter ธนาคาร เพื่อสามารถดูกราฟในแต่ละธนาคารแสดงดังรูปที่ 3.25 และส่วนที่ออกแบบเป็นแผนภูมิแท่งแสดงข้อมูลจำนวนครั้งในการโอน และจำนวนเงินทั้งหมดที่ใช้ในแต่ละธนาคารแสดงดังรูปที่ 3.26

รูปที่ 3.25 กราฟวันเวลาที่ทำธุรกรรม

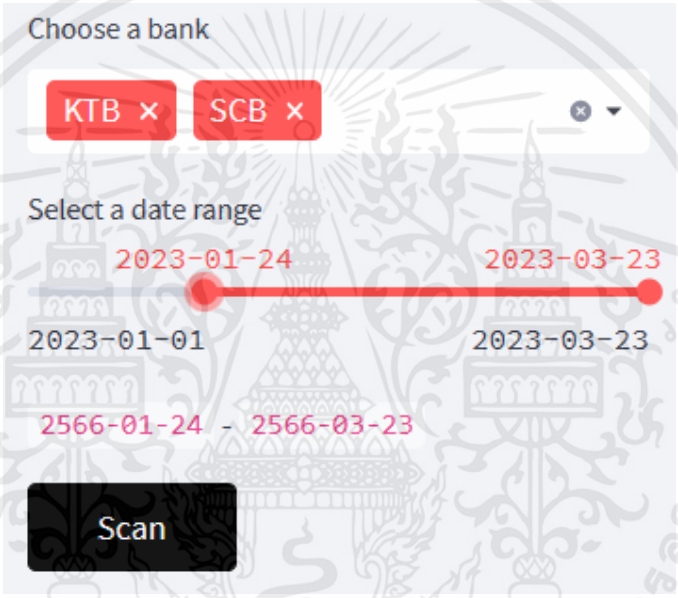


รูปที่ 3.26 แผนภูมิข้อมูลจำนวนครั้งในการโอน และจำนวนเงินทั้งหมดในการทำธุรกรรม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5) Filter

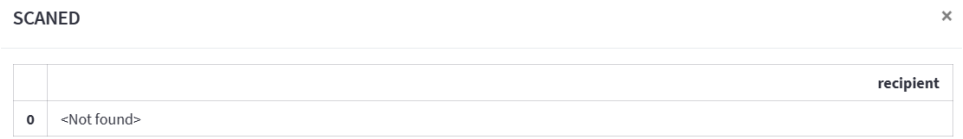
Filter คือ Feature ที่ออกแบบมาเพื่อเพิ่มประสิทธิภาพในการใช้งานของเว็บแอปพลิเคชันได้มากขึ้นในการค้นหาข้อมูลที่สนใจ โดยถูกแบ่งออกเป็น 2 ส่วน คือ ส่วนที่เป็น Option ที่มีความสามารถในการเลือกข้อมูลธนาคารแบบ multi-select พร้อมกับ Feature ที่สามารถเลือกช่วงเวลาในการทำธุรกรรมที่สนใจแสดงดังรูปที่ 3.27 และ Name Filters ที่มี Feature ในการพิมพ์ค้นหาชื่อผู้รับที่สนใจแสดงดังรูปที่ 3.28 โดยออกแบบผลลัพธ์ให้แสดงข้อมูลในการแสดง Filter ตามที่กำหนดใน Data Frame แสดงตัวอย่างดังรูปที่ 3.29



รูปที่ 3.27 Option

รูปที่ 3.28 Name Filters

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้


SCANED

	recipient
0	<Not found>

รูปที่ 3.29 ผลลัพธ์การ Filter

3.2 เครื่องมือที่ใช้ในการทดลอง

ในปริญญานิพนธ์นี้ มีอุปกรณ์และเครื่องมือที่ใช้ในการทดลอง ดังนี้

3.2.1 Jupyter Notebook

ใช้สำหรับการเขียนโปรแกรมโดยใช้ภาษา Python ในการสร้าง CNN Model และการทำ OCR เพื่อจัดการข้อมูลให้อยู่ในรูปของ Data Frame

3.2.2 Visual Studio Code

ใช้สำหรับการเขียนโปรแกรมโดยใช้ภาษา Python ในการทำเว็บแอปพลิเคชันเพื่อแสดงผลข้อมูลในรูปแบบ Data Visualization

3.2.3 คอมพิวเตอร์ ที่มีคุณสมบัติดังนี้

- หน่วยประมลผล NVIDIA GeForce GTX 1650
- หน่วยความจำ 8GB DDR4 2666 MHz

3.3 การจัดเก็บผลการทดลอง

3.3.1 ทดสอบการปรับขนาดรูปภาพเพื่อให้ทราบค่าขนาดที่เหมาะสมกับโมเดลที่ออกแบบ

3.3.1.1 ปรับขนาดอินพุตของโมเดลเป็น 8x8 pixel

3.3.1.2 ปรับขนาดอินพุตของโมเดลเป็น 16x16 pixel

3.3.1.3 ปรับขนาดอินพุตของโมเดลเป็น 28x28 pixel

3.3.1.4 ปรับขนาดอินพุตของโมเดลเป็น 32x32 pixel

3.3.1.5 ปรับขนาดอินพุตของโมเดลเป็น 64x64 pixel

3.3.1.6 ปรับขนาดอินพุตของโมเดลเป็น 128x128 pixel

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3.2 ทดสอบการจำแนกรูปภาพสลิปการเงินออกจากรูปภาพชนิดอื่น

3.3.2.1 ทดสอบด้วยชุดข้อมูลทดสอบ

3.3.2.2 ประเมินประสิทธิภาพของโมเดลด้วย Confusion Matrix

3.3.3 ทดสอบการเปลี่ยนภาพระดับสีเทา

3.3.4 ทดสอบการแยกชนิดของสลิปการเงิน

3.3.4.1 ทดสอบด้วยชุดข้อมูลทดสอบ

3.3.4.2 ประเมินประสิทธิภาพของโมเดลด้วย Confusion Matrix

3.3.5 ทดสอบการแบ่งประเภทสลิปตามธนาคารของสลิปการเงินแบบอิเล็กทรอนิกส์

3.3.5.1 ทดสอบด้วยชุดข้อมูลทดสอบ

3.3.5.2 ประเมินประสิทธิภาพของโมเดลด้วย Confusion Matrix

3.3.6 ทดสอบการแบ่งประเภทสลิปตามธนาคารของสลิปการเงินแบบกระดาษ

3.3.6.1 ทดสอบด้วยชุดข้อมูลทดสอบ

3.3.6.2 ประเมินประสิทธิภาพของโมเดลด้วย Confusion Matrix

3.3.7 ทดสอบการเก็บข้อมูลสลิปการเงิน (Financial Data Extraction)

3.3.7.1 ทดสอบการดึงข้อมูลจากรูปภาพสลิปการเงิน

3.3.7.2 ทดสอบการทำ Data Frame

3.3.7.3 ทดสอบการเก็บข้อมูลในรูปแบบไฟล์ .xlsx

3.3.8 ทดสอบการแสดงผลข้อมูล

3.3.8.1 ทดสอบกราฟวันเวลาที่ทำธุรกรรม

3.3.8.2 ทดสอบแผนภูมิแท่งของข้อมูลจำนวนครั้งในการโอน และจำนวนเงินทั้งหมดของแต่ละธนาคาร

3.3.8.3 ทดสอบ Option

3.3.8.4 ทดสอบ Advanced Option

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

ผลการทดลอง

ผู้จัดทำได้ทำการเก็บผลการทำงานของระบบโดยแบ่งการทดสอล และจัดเก็บผลการทดสอลเป็นส่วนต่าง ๆ ดังต่อไปนี้

4.1 ทดสอบการปรับขนาดรูปภาพเพื่อให้ทราบค่าขนาดที่เหมาะสมกับโมเดลที่ออกแบบ

ในการทดสอบปรับขนาดได้ประเมินประสิทธิภาพของโมเดลด้วยการทำ Confusion Matrix โดยทดสอบกับโมเดลในการจำแนกประเภทรูปภาพสลิปการเงินออกจากรูปภาพชนิดอื่น ทำให้ Confusion Matrix มีขนาดเป็น 2x2 เนื่องจาก Class ของข้อมูลที่มี 2 Class คือรูปภาพสลิปการเงิน และรูปภาพที่ไม่ใช่สลิปการเงิน นิยามให้เป็น Noise และในการทดสอบได้ใช้รูปภาพสลิปการเงิน 2,400 รูป และรูปภาพที่ไม่ใช่สลิปการเงิน 2,000 รูป ที่นอกเหนือจากรูปภาพที่ใช้ฝึกสอนโมเดล ทำให้ค่า TP, TN, FN และ FN บอกจำนวนของรูปภาพที่เกิดจากการประเมินประสิทธิภาพ

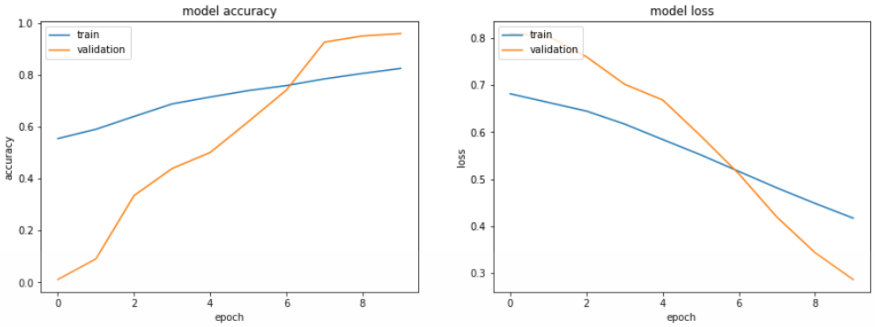
4.1.1 ปรับขนาดอินพุตของโมเดลเป็น 8x8 pixel

จากการปรับขนาดรูปภาพเป็น 8x8 pixel ได้ผลลัพธ์ของการฝึกสอนโมเดลดังรูปที่ 4.1 แสดงกราฟของผลการฝึกสอนโมเดลได้ดังรูปที่ 4.2, Confusion Matrix แสดงดังรูปที่ 4.3

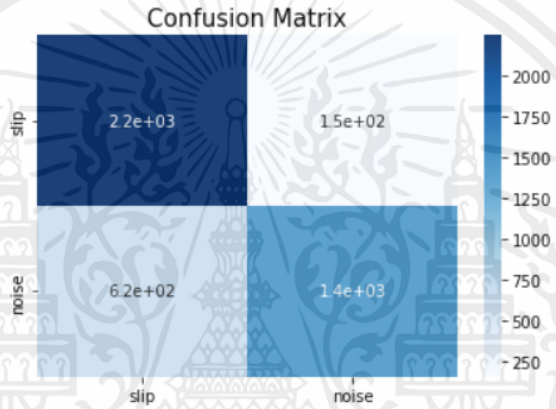
```
Epoch 1/10
116/116 [==============================] - 4s 22ms/step - loss: 0.6814 - accuracy: 0.5553 - val_loss: 0.8096 - val_accuracy: 0.0110
Epoch 2/10
116/116 [==============================] - 2s 20ms/step - loss: 0.6629 - accuracy: 0.5913 - val_loss: 0.8040 - val_accuracy: 0.0909
Epoch 3/10
116/116 [==============================] - 2s 20ms/step - loss: 0.6446 - accuracy: 0.6408 - val_loss: 0.7600 - val_accuracy: 0.3354
Epoch 4/10
116/116 [==============================] - 2s 20ms/step - loss: 0.6170 - accuracy: 0.6892 - val_loss: 0.7017 - val_accuracy: 0.4396
Epoch 5/10
116/116 [==============================] - 2s 20ms/step - loss: 0.5841 - accuracy: 0.7159 - val_loss: 0.6683 - val_accuracy: 0.5018
Epoch 6/10
116/116 [==============================] - 2s 20ms/step - loss: 0.5515 - accuracy: 0.7407 - val_loss: 0.5917 - val_accuracy: 0.6207
Epoch 7/10
116/116 [==============================] - 2s 20ms/step - loss: 0.5163 - accuracy: 0.7601 - val_loss: 0.5111 - val_accuracy: 0.7427
Epoch 8/10
116/116 [==============================] - 2s 20ms/step - loss: 0.4812 - accuracy: 0.7859 - val_loss: 0.4190 - val_accuracy: 0.9280
Epoch 9/10
116/116 [==============================] - 2s 20ms/step - loss: 0.4483 - accuracy: 0.8070 - val_loss: 0.3435 - val_accuracy: 0.9518
Epoch 10/10
116/116 [==============================] - 2s 20ms/step - loss: 0.4169 - accuracy: 0.8264 - val_loss: 0.2863 - val_accuracy: 0.9610
```

รูปที่ 4.1 ผลการฝึกสอนโมเดลที่มีขนาดอินพุต 8x8 pixel

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.2 กราฟค่า Loss, Accuracy ของโมเดลที่มีขนาดอินพุต 8x8 pixel



รูปที่ 4.3 Confusion Matrix ของโมเดลที่มีขนาดอินพุต 8x8 pixel

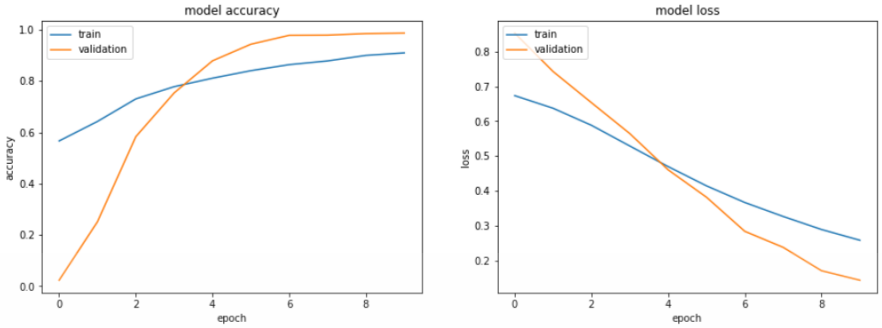
4.1.2 ปรับขนาดอินพุตของโมเดลเป็น 16x16 pixel

จากการปรับขนาดรูปภาพเป็น 16x16 pixel ได้ผลลัพธ์ของการฝึกสอนโมเดลดังรูปที่ 4.4 แสดงกราฟของผลการฝึกสอนโมเดลได้ดังรูปที่ 4.5 และ Confusion Matrix แสดงดังรูปที่ 4.6

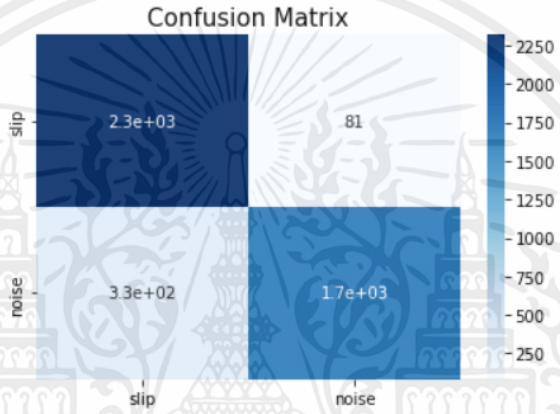
```
Epoch 1/10
116/116 [==============================] - 8s 63ms/step - loss: 0.6736 - accuracy: 0.5663 - val_loss: 0.8533 - val_accuracy: 0.0232
Epoch 2/10
116/116 [==============================] - 7s 60ms/step - loss: 0.6374 - accuracy: 0.6423 - val_loss: 0.7430 - val_accuracy: 0.2512
Epoch 3/10
116/116 [==============================] - 7s 60ms/step - loss: 0.5884 - accuracy: 0.7301 - val_loss: 0.6534 - val_accuracy: 0.5829
Epoch 4/10
116/116 [==============================] - 7s 60ms/step - loss: 0.5289 - accuracy: 0.7778 - val_loss: 0.5646 - val_accuracy: 0.7537
Epoch 5/10
116/116 [==============================] - 7s 62ms/step - loss: 0.4697 - accuracy: 0.8109 - val_loss: 0.4604 - val_accuracy: 0.8787
Epoch 6/10
116/116 [==============================] - 7s 61ms/step - loss: 0.4143 - accuracy: 0.8402 - val_loss: 0.3817 - val_accuracy: 0.9433
Epoch 7/10
116/116 [==============================] - 7s 61ms/step - loss: 0.3665 - accuracy: 0.8637 - val_loss: 0.2836 - val_accuracy: 0.9780
Epoch 8/10
116/116 [==============================] - 7s 61ms/step - loss: 0.3263 - accuracy: 0.8782 - val_loss: 0.2376 - val_accuracy: 0.9787
Epoch 9/10
116/116 [==============================] - 7s 61ms/step - loss: 0.2888 - accuracy: 0.8998 - val_loss: 0.1704 - val_accuracy: 0.9848
Epoch 10/10
116/116 [==============================] - 7s 61ms/step - loss: 0.2580 - accuracy: 0.9096 - val_loss: 0.1431 - val_accuracy: 0.9866
```

รูปที่ 4.4 ผลการฝึกสอนโมเดลที่มีขนาดอินพุต 16x16 pixel

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.5 กราฟค่า Loss, Accuracy ของโมเดลที่มีขนาดอินพุต 16x16 pixel



รูปที่ 4.6 Confusion Matrix ของโมเดลที่มีขนาดอินพุต 16x16 pixel

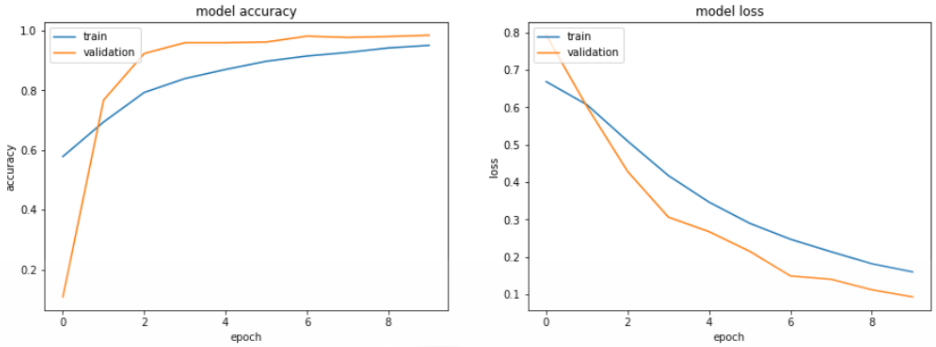
4.1.3 ปรับขนาดอินพุตของโมเดลเป็น 28x28 pixel

จากการปรับขนาดรูปภาพเป็น 28x28 pixel ได้ผลลัพธ์ของการฝึกสอนโมเดลดังรูปที่ 4.7 แสดงกราฟของผลการฝึกสอนโมเดลได้ดังรูปที่ 4.8 และ Confusion Matrix แสดงดังรูปที่ 4.9

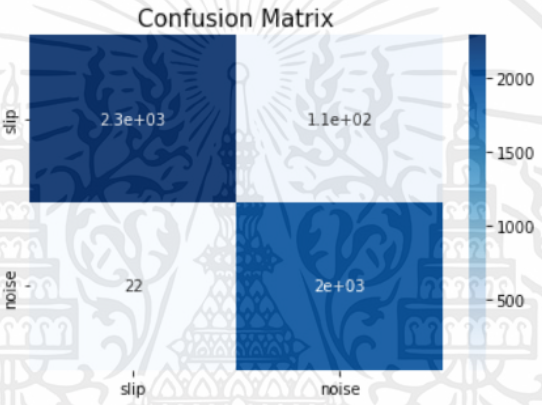
```
Epoch 1/10
116/116 [==============================] - 20s 161ms/step - loss: 0.6687 - accuracy: 0.5783 - val_loss: 0.7929 - val_accuracy: 0.1085
Epoch 2/10
116/116 [==============================] - 19s 160ms/step - loss: 0.6069 - accuracy: 0.6938 - val_loss: 0.6016 - val_accuracy: 0.7671
Epoch 3/10
116/116 [==============================] - 19s 161ms/step - loss: 0.5095 - accuracy: 0.7928 - val_loss: 0.4282 - val_accuracy: 0.9232
Epoch 4/10
116/116 [==============================] - 19s 163ms/step - loss: 0.4175 - accuracy: 0.8392 - val_loss: 0.3063 - val_accuracy: 0.9591
Epoch 5/10
116/116 [==============================] - 19s 162ms/step - loss: 0.3463 - accuracy: 0.8697 - val_loss: 0.2674 - val_accuracy: 0.9591
Epoch 6/10
116/116 [==============================] - 19s 162ms/step - loss: 0.2897 - accuracy: 0.8971 - val_loss: 0.2147 - val_accuracy: 0.9616
Epoch 7/10
116/116 [==============================] - 20s 171ms/step - loss: 0.2472 - accuracy: 0.9147 - val_loss: 0.1489 - val_accuracy: 0.9811
Epoch 8/10
116/116 [==============================] - 19s 165ms/step - loss: 0.2136 - accuracy: 0.9266 - val_loss: 0.1399 - val_accuracy: 0.9768
Epoch 9/10
116/116 [==============================] - 19s 163ms/step - loss: 0.1815 - accuracy: 0.9417 - val_loss: 0.1119 - val_accuracy: 0.9799
Epoch 10/10
116/116 [==============================] - 19s 164ms/step - loss: 0.1598 - accuracy: 0.9502 - val_loss: 0.0929 - val_accuracy: 0.9841
```

รูปที่ 4.7 ผลการฝึกสอนโมเดลที่มีขนาดอินพุต 28x28 pixel

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.8 กราฟค่า Loss, Accuracy ของโมเดลที่มีขนาดอินพุต 28x28 pixel



รูปที่ 4.9 Confusion Matrix ของโมเดลที่มีขนาดอินพุต 28x28 pixel

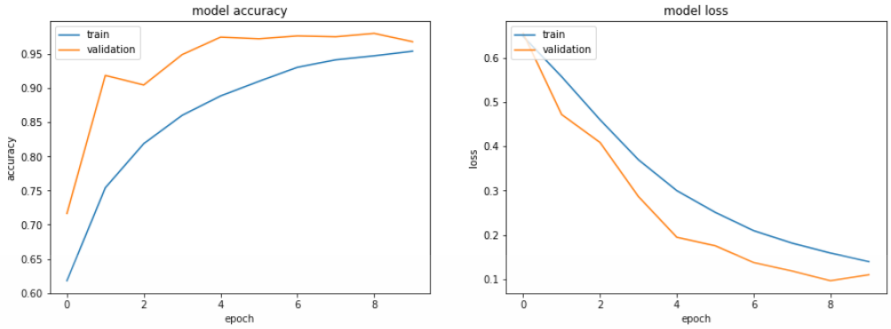
4.1.4 ปรับขนาดอินพุตของโมเดลเป็น 32x32 pixel

จากการปรับขนาดรูปภาพเป็น 32x32 pixel ได้ผลลัพธ์ของการฝึกสอนโมเดลดังรูปที่ 4.10 แสดงกราฟของผลการฝึกสอนโมเดลได้ดังรูปที่ 4.11 และ Confusion Matrix ดังรูปที่ 4.12

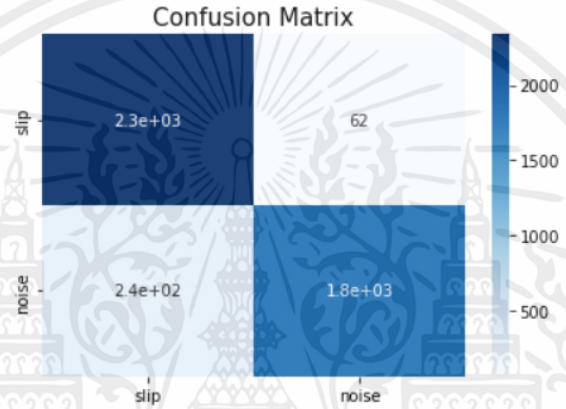
```
Epoch 1/10
116/116 [==============================] - 26s 211ms/step - loss: 0.6480 - accuracy: 0.6179 - val_loss: 0.6549 - val_accuracy: 0.7165
Epoch 2/10
116/116 [==============================] - 25s 212ms/step - loss: 0.5573 - accuracy: 0.7542 - val_loss: 0.4719 - val_accuracy: 0.9189
Epoch 3/10
116/116 [==============================] - 25s 215ms/step - loss: 0.4601 - accuracy: 0.8188 - val_loss: 0.4088 - val_accuracy: 0.9049
Epoch 4/10
116/116 [==============================] - 25s 214ms/step - loss: 0.3697 - accuracy: 0.8604 - val_loss: 0.2867 - val_accuracy: 0.9494
Epoch 5/10
116/116 [==============================] - 25s 214ms/step - loss: 0.3000 - accuracy: 0.8888 - val_loss: 0.1947 - val_accuracy: 0.9750
Epoch 6/10
116/116 [==============================] - 25s 214ms/step - loss: 0.2510 - accuracy: 0.9102 - val_loss: 0.1756 - val_accuracy: 0.9726
Epoch 7/10
116/116 [==============================] - 25s 215ms/step - loss: 0.2096 - accuracy: 0.9308 - val_loss: 0.1377 - val_accuracy: 0.9768
Epoch 8/10
116/116 [==============================] - 25s 214ms/step - loss: 0.1816 - accuracy: 0.9418 - val_loss: 0.1186 - val_accuracy: 0.9756
Epoch 9/10
116/116 [==============================] - 25s 215ms/step - loss: 0.1591 - accuracy: 0.9476 - val_loss: 0.0967 - val_accuracy: 0.9805
Epoch 10/10
116/116 [==============================] - 25s 214ms/step - loss: 0.1397 - accuracy: 0.9544 - val_loss: 0.1102 - val_accuracy: 0.9683
```

รูปที่ 4.10 ผลการฝึกสอนโมเดลที่มีขนาดอินพุต 32x32 pixel

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.11 กราฟแสดงค่า Loss, Accuracy ของโมเดลที่มีขนาดอินพุต 32x32 pixel



รูปที่ 4.12 Confusion Matrix ของโมเดลที่มีขนาดอินพุต 32x32 pixel

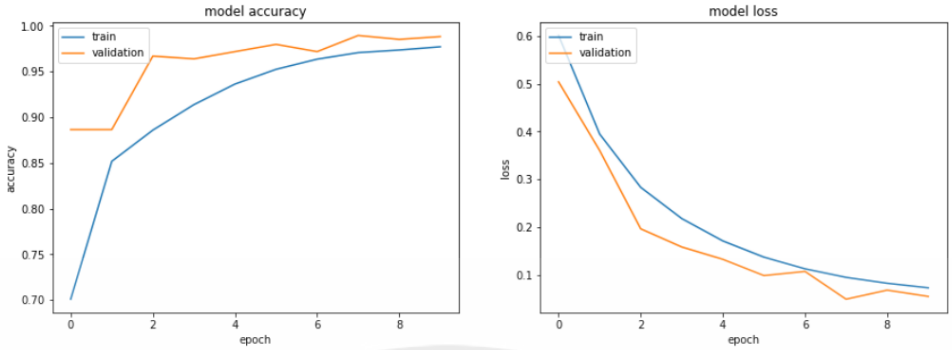
4.1.5 ปรับขนาดอินพุตของโมเดลเป็น 64x64 pixel

จากการปรับขนาดรูปภาพเป็น 64x64 pixel ได้ผลลัพธ์ของการฝึกสอนโมเดลดังรูปที่ 4.13 กราฟของผลการฝึกสอนโมเดลได้ดังรูปที่ 4.14 และ Confusion Matrix แสดงดังรูปที่ 4.15

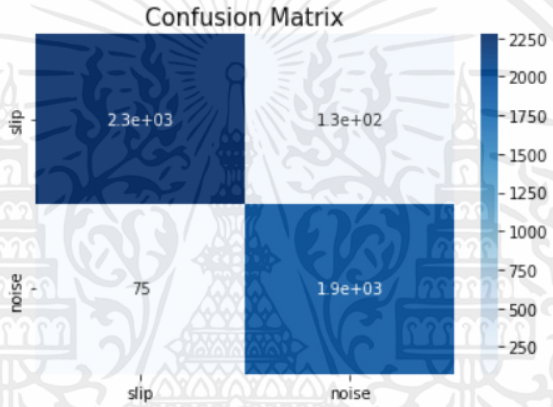
```
Epoch 1/10
116/116 [==============================] - 103s 877ms/step - loss: 0.6008 - accuracy: 0.7008 - val_loss: 0.5039 - val_accuracy: 0.8866
Epoch 2/10
116/116 [==============================] - 103s 884ms/step - loss: 0.3948 - accuracy: 0.8518 - val_loss: 0.3604 - val_accuracy: 0.8866
Epoch 3/10
116/116 [==============================] - 102s 881ms/step - loss: 0.2834 - accuracy: 0.8860 - val_loss: 0.1967 - val_accuracy: 0.9671
Epoch 4/10
116/116 [==============================] - 103s 884ms/step - loss: 0.2182 - accuracy: 0.9139 - val_loss: 0.1588 - val_accuracy: 0.9640
Epoch 5/10
116/116 [==============================] - 104s 894ms/step - loss: 0.1714 - accuracy: 0.9364 - val_loss: 0.1330 - val_accuracy: 0.9720
Epoch 6/10
116/116 [==============================] - 103s 887ms/step - loss: 0.1375 - accuracy: 0.9526 - val_loss: 0.0991 - val_accuracy: 0.9799
Epoch 7/10
116/116 [==============================] - 103s 890ms/step - loss: 0.1132 - accuracy: 0.9637 - val_loss: 0.1077 - val_accuracy: 0.9720
Epoch 8/10
116/116 [==============================] - 103s 885ms/step - loss: 0.0953 - accuracy: 0.9709 - val_loss: 0.0497 - val_accuracy: 0.9896
Epoch 9/10
116/116 [==============================] - 103s 884ms/step - loss: 0.0830 - accuracy: 0.9737 - val_loss: 0.0686 - val_accuracy: 0.9854
Epoch 10/10
116/116 [==============================] - 104s 898ms/step - loss: 0.0735 - accuracy: 0.9772 - val_loss: 0.0555 - val_accuracy: 0.9884
```

รูปที่ 4.13 ผลการฝึกสอนโมเดลที่มีขนาดอินพุต 64x64 pixel

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.14 กราฟค่า Loss, Accuracy ของโมเดลที่มีขนาดอินพุต 64x64 pixel



รูปที่ 4.15 Confusion Matrix ของโมเดลที่มีขนาดอินพุต 64x64 pixel

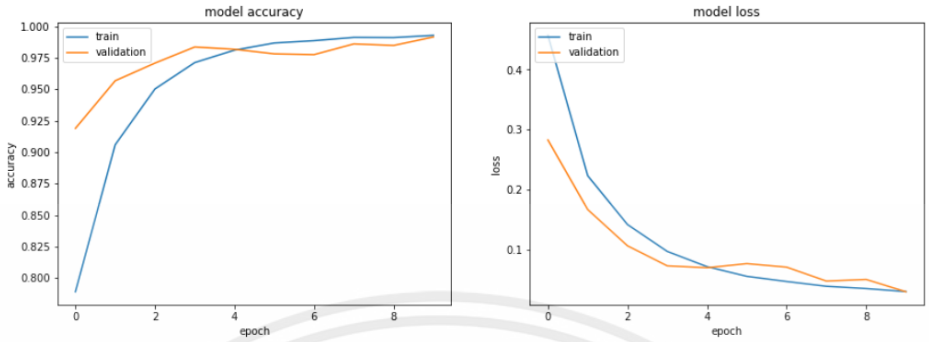
4.1.6 ปรับขนาดอินพุตของโมเดลเป็น 128x128 pixel

จากการปรับขนาดรูปภาพเป็น 128x128 pixel ได้ผลลัพธ์ของการฝึกสอนโมเดลดังรูปที่ 4.16 กราฟของผลการฝึกสอนโมเดลได้ดังรูปที่ 4.17 และ Confusion Matrix แสดงดังรูปที่ 4.18

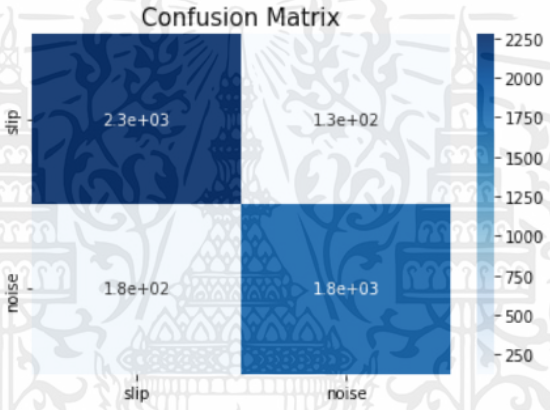
```
Epoch 1/10
116/116 [==============================] - 399s 3s/step - loss: 0.4571 - accuracy: 0.7892 - val_loss: 0.2827 - val_accuracy: 0.9189
Epoch 2/10
116/116 [==============================] - 393s 3s/step - loss: 0.2231 - accuracy: 0.9059 - val_loss: 0.1667 - val_accuracy: 0.9567
Epoch 3/10
116/116 [==============================] - 389s 3s/step - loss: 0.1417 - accuracy: 0.9502 - val_loss: 0.1063 - val_accuracy: 0.9707
Epoch 4/10
116/116 [==============================] - 388s 3s/step - loss: 0.0969 - accuracy: 0.9711 - val_loss: 0.0729 - val_accuracy: 0.9835
Epoch 5/10
116/116 [==============================] - 388s 3s/step - loss: 0.0716 - accuracy: 0.9810 - val_loss: 0.0697 - val_accuracy: 0.9817
Epoch 6/10
116/116 [==============================] - 387s 3s/step - loss: 0.0554 - accuracy: 0.9867 - val_loss: 0.0766 - val_accuracy: 0.9780
Epoch 7/10
116/116 [==============================] - 387s 3s/step - loss: 0.0468 - accuracy: 0.9886 - val_loss: 0.0707 - val_accuracy: 0.9774
Epoch 8/10
116/116 [==============================] - 388s 3s/step - loss: 0.0388 - accuracy: 0.9911 - val_loss: 0.0475 - val_accuracy: 0.9860
Epoch 9/10
116/116 [==============================] - 388s 3s/step - loss: 0.0350 - accuracy: 0.9910 - val_loss: 0.0502 - val_accuracy: 0.9848
Epoch 10/10
116/116 [==============================] - 387s 3s/step - loss: 0.0299 - accuracy: 0.9928 - val_loss: 0.0297 - val_accuracy: 0.9915
```

รูปที่ 4.16 ผลการฝึกสอนโมเดลที่มีขนาดอินพุต 128x128 pixel

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.17 กราฟค่า Loss, Accuracy ของโมเดลที่มีขนาดอินพุต 128x128 pixel



รูปที่ 4.18 Confusion Matrix ของโมเดลที่มีขนาดอินพุต 128x128 pixel

4.1.7 สรุปผลการทดสอบขนาดโมเดล

ตารางที่ 4.1 แสดงผลลัพธ์ และประเมินประสิทธิภาพด้วยวิธีการของ F1 Score โดยทางผู้จัดทำได้เลือกขนาดอินพุตที่มีค่า F1 Score สูงที่สุด

Size (pixel)	Accuracy	F1 score
8x8	0.82	0.82
16x16	0.91	0.90
28x28	0.97	0.97
32x32	0.93	0.93
64x64	0.95	0.95
128x128	0.93	0.93

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2 ทดสอบการจำแนกรูปภาพสลิปการเงินออกจากรูปภาพชนิดอื่น

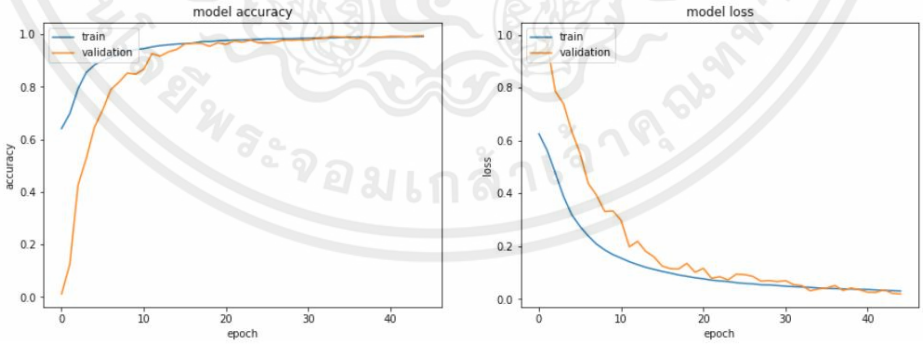
ในการทดสอบประสิทธิภาพได้ประเมินประสิทธิภาพของโมเดลการจำแนกประเภทรูปภาพสลิปการเงินออกจากรูปภาพชนิดอื่นด้วยการทำ Confusion Matrix และกำหนดให้ Confusion Matrix มีขนาดเป็น 2x2 เนื่องจาก Class ของข้อมูลที่มี 2 Class คือรูปภาพสลิปการเงิน และรูปภาพที่ไม่ใช่สลิปการเงิน นิยามให้เป็น Noise และค่าของ TP, TN, FN และ FN บอกจำนวนของรูปภาพที่เกิดจากการประเมินประสิทธิภาพ

4.2.1 ทดสอบด้วยชุดข้อมูล

ได้ทำการฝึกสอนโมเดลเป็นจำนวน Epoch = 45 รอบ โดยมีผลการฝึกสอนโมเดล กราฟค่า Loss, Accuracy กับจำนวนรอบที่ฝึกสอน และผลการทดสอบโมเดล แสดงได้ดังรูปที่ 4.19, 4.20 และ 4.21 ตามลำดับ

```
103/103 [==============================] - 17s 166ms/step - loss: 0.0389 - accuracy: 0.9894 - val_loss: 0.0328 - val_accuracy: 0.9918
Epoch 39/45
103/103 [==============================] - 17s 166ms/step - loss: 0.0384 - accuracy: 0.9894 - val_loss: 0.0422 - val_accuracy: 0.9878
Epoch 40/45
103/103 [==============================] - 17s 166ms/step - loss: 0.0371 - accuracy: 0.9894 - val_loss: 0.0353 - val_accuracy: 0.9899
Epoch 41/45
103/103 [==============================] - 17s 166ms/step - loss: 0.0366 - accuracy: 0.9896 - val_loss: 0.0265 - val_accuracy: 0.9924
Epoch 42/45
103/103 [==============================] - 17s 166ms/step - loss: 0.0347 - accuracy: 0.9900 - val_loss: 0.0257 - val_accuracy: 0.9921
Epoch 43/45
103/103 [==============================] - 17s 166ms/step - loss: 0.0334 - accuracy: 0.9909 - val_loss: 0.0355 - val_accuracy: 0.9896
Epoch 44/45
103/103 [==============================] - 17s 168ms/step - loss: 0.0321 - accuracy: 0.9905 - val_loss: 0.0220 - val_accuracy: 0.9939
Epoch 45/45
103/103 [==============================] - 17s 167ms/step - loss: 0.0310 - accuracy: 0.9914 - val_loss: 0.0197 - val_accuracy: 0.9951
```

รูปที่ 4.19 ผลการฝึกสอนโมเดลจำแนกรูปภาพสลิปการเงิน



รูปที่ 4.20 กราฟค่า Loss, Accuracy ของโมเดลจำแนกรูปภาพสลิปการเงิน

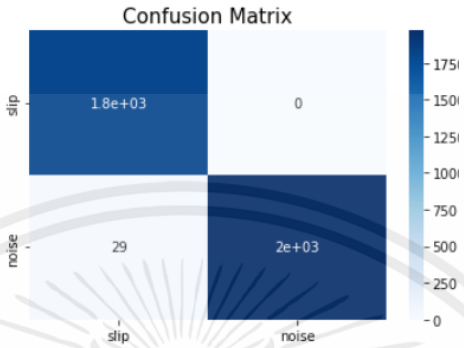
```
119/119 [==============================] - 1s 9ms/step - loss: 0.0274 - accuracy: 0.9924
Test Accuracy 0.99
```

รูปที่ 4.21 ผลการทดสอบโมเดลจำแนกรูปภาพสลิปการเงิน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.2.2 ประเมินประสิทธิภาพด้วย Confusion Matrix

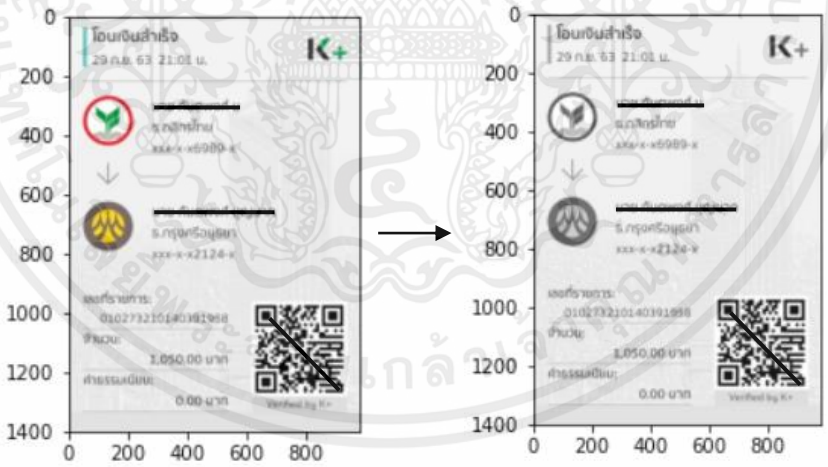
ประสิทธิภาพของโมเดลถูกประเมิณจากการทำ Confusion Matrix โดยมีผลลัพธ์ที่ได้แสดงได้ดังรูปที่ 4.22 โดยมีคะแนนของ F1 Score คือ 0.99



รูปที่ 4.22 Confusion Matrix ของโมเดลจำแนกรูปภาพสลิปการเงิน

4.3 ทดสอบการเปลี่ยนภาพระดับสีเทา

จากการเปลี่ยนภาพระดับสีเทาทำให้ได้ผลลัพธ์แสดงดังรูป 4.23 ซึ่งเปลี่ยนโดยใช้งานไลบรารี่ของ OpenCV



รูปที่ 4.23 รูปภาพก่อน และหลังการแปลงเฉดสีเป็นภาพระดับสีเทา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.4 ทดสอบการแยกชนิดของสลิปการเงิน

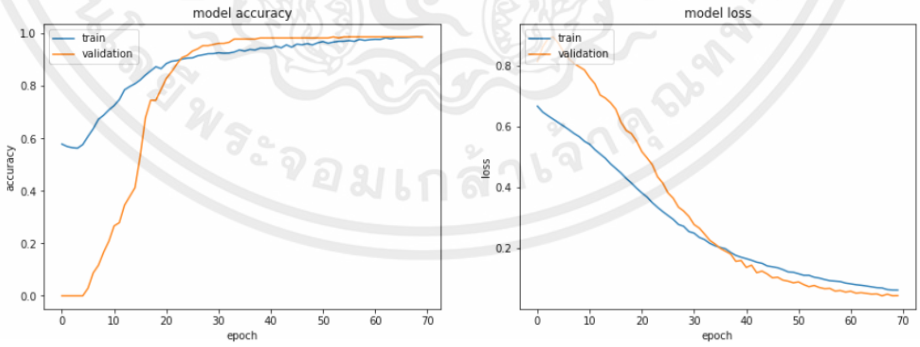
ในการทดสอบประสิทธิภาพได้ประเมินประสิทธิภาพของโมเดลแยกชนิดของสลิปการเงินด้วยการทำ Confusion Matrix และกำหนดให้ Confusion Matrix มีขนาดเป็น 2x2 เนื่องจาก Class ของข้อมูลที่มี 2 Class คือรูปภาพสลิปการเงินแบบอิเล็กทรอนิกส์กับรูปภาพสลิปการเงินแบบกระดาษ และค่าของ TP, TN, FN และ FN บอกจำนวนของรูปภาพที่เกิดจากการประเมินประสิทธิภาพ

4.4.1 ทดสอบด้วยชุดข้อมูล

ได้ทำการฝึกสอนโมเดลเป็นจำนวน Epoch = 70 รอบ โดยมีผลการฝึกสอนโมเดล กราฟค่า Loss, Accuracy กับจำนวนรอบที่ฝึกสอน และผลการทดสอบโมเดล แสดงได้ดังรูปที่ 4.24, 4.25 และ 4.26 ตามลำดับ

```
51/51 [==============================] - 15s 300ms/step - loss: 0.0113 - accuracy: 0.9978 - val_loss: 0.0289 - val_accuracy: 0.9979
Epoch 62/70
51/51 [==============================] - 15s 300ms/step - loss: 0.0112 - accuracy: 0.9981 - val_loss: 0.0167 - val_accuracy: 0.9993
Epoch 63/70
51/51 [==============================] - 15s 302ms/step - loss: 0.0112 - accuracy: 0.9975 - val_loss: 0.0156 - val_accuracy: 0.9993
Epoch 64/70
51/51 [==============================] - 16s 310ms/step - loss: 0.0106 - accuracy: 0.9977 - val_loss: 0.0243 - val_accuracy: 0.9979
Epoch 65/70
51/51 [==============================] - 16s 309ms/step - loss: 0.0096 - accuracy: 0.9981 - val_loss: 0.0182 - val_accuracy: 0.9986
Epoch 66/70
51/51 [==============================] - 16s 311ms/step - loss: 0.0107 - accuracy: 0.9978 - val_loss: 0.0287 - val_accuracy: 0.9979
Epoch 67/70
51/51 [==============================] - 16s 311ms/step - loss: 0.0088 - accuracy: 0.9986 - val_loss: 0.0139 - val_accuracy: 0.9993
Epoch 68/70
51/51 [==============================] - 16s 309ms/step - loss: 0.0089 - accuracy: 0.9986 - val_loss: 0.0199 - val_accuracy: 0.9986
Epoch 69/70
51/51 [==============================] - 16s 310ms/step - loss: 0.0091 - accuracy: 0.9984 - val_loss: 0.0133 - val_accuracy: 0.9993
Epoch 70/70
51/51 [==============================] - 16s 310ms/step - loss: 0.0082 - accuracy: 0.9986 - val_loss: 0.0269 - val_accuracy: 0.9972
```

รูปที่ 4.24 ผลการฝึกสอนโมเดลแยกชนิดของสลิปการเงิน



รูปที่ 4.25 กราฟค่า Loss, Accuracy ของโมเดลแยกชนิดของสลิปการเงิน

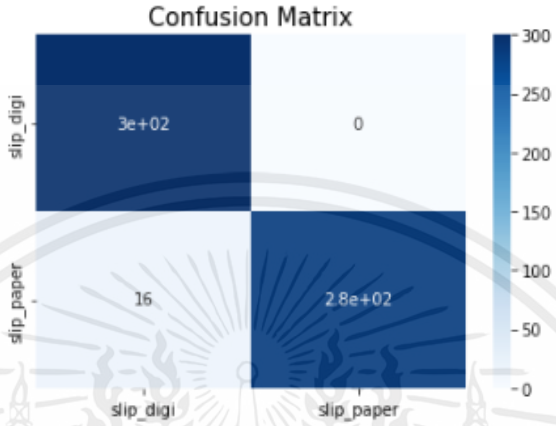
```
19/19 [==============================] - 0s 8ms/step - loss: 0.0857 - accuracy: 0.9767
Test Accuracy 0.98
```

รูปที่ 4.26 ผลการทดสอบโมเดลแยกชนิดของสลิปการเงิน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.4.2 ประเมินประสิทธิภาพด้วย Confusion Matrix

ประสิทธิภาพของโมเดลถูกประเมิณจากการทำ Confusion Matrix โดยมีผลลัพธ์ที่ได้แสดงได้ดังรูปที่ 4.27 โดยมีคะแนน F1 Score คือ 0.97



รูปที่ 4.27 Confusion Matrix ของโมเดลแยกชนิดของสลิปการเงิน

4.5 ทดสอบการแบ่งประเภทสลิปตามธนาคารของสลิปการเงินแบบอิเล็กทรอนิกส์

ในการทดสอบประสิทธิภาพได้ประเมินประสิทธิภาพของโมเดลแบ่งประเภทสลิปตามธนาคารของสลิปการเงินแบบอิเล็กทรอนิกส์ด้วยการทำ Confusion Matrix และกำหนดให้ Confusion Matrix มีขนาดเป็น 6x6 เนื่องจาก Class ของข้อมูลที่มี 6 Class คือรูปภาพสลิปการเงินแบบอิเล็กทรอนิกส์ที่ประกอบไปด้วยธนาคาร BBL, KTB, BAY, Kbank, TTB และ SCB โดยค่าของ TP, TN, FN และ FN บอกจำนวนของรูปภาพที่เกิดจากการประเมินประสิทธิภาพ

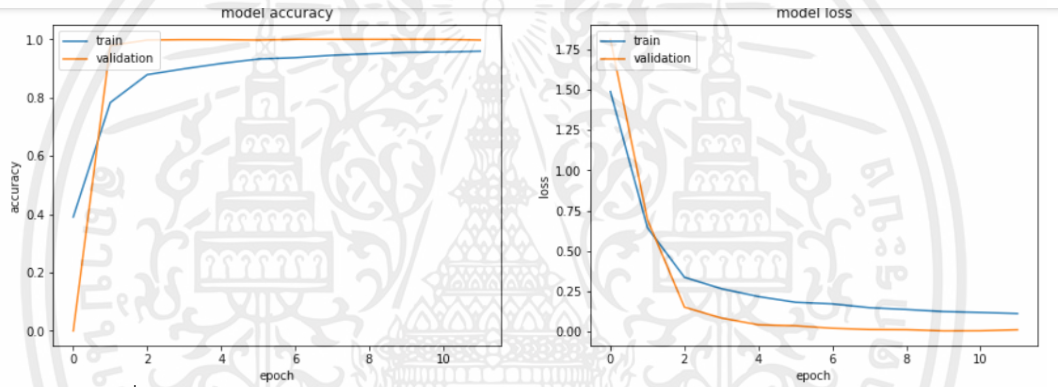
4.5.1 ทดสอบด้วยชุดข้อมูล

ได้ทำการฝึกสอนโมเดลเป็นจำนวน Epoch = 12 รอบ โดยมีผลการฝึกสอนโมเดลกราฟค่า Loss, Accuracy กับจำนวนรอบที่ฝึกสอน และผลการทดสอบโมเดล แสดงได้ดังรูปที่ 4.28, 4.29 และ 4.30 ตามลำดับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
Epoch 1/12
119/119 [==============================] - 33s 273ms/step - loss: 1.4887 - accuracy: 0.3909 - val_loss: 1.8137 - val_accuracy: 0.0000e+
00
Epoch 2/12
119/119 [==============================] - 33s 277ms/step - loss: 0.6439 - accuracy: 0.7832 - val_loss: 0.6934 - val_accuracy: 0.9786
Epoch 3/12
119/119 [==============================] - 33s 279ms/step - loss: 0.3359 - accuracy: 0.8790 - val_loss: 0.1504 - val_accuracy: 0.9976
Epoch 4/12
119/119 [==============================] - 35s 294ms/step - loss: 0.2649 - accuracy: 0.8991 - val_loss: 0.0828 - val_accuracy: 0.9988
Epoch 5/12
119/119 [==============================] - 33s 278ms/step - loss: 0.2162 - accuracy: 0.9172 - val_loss: 0.0408 - val_accuracy: 0.9988
Epoch 6/12
119/119 [==============================] - 33s 280ms/step - loss: 0.1808 - accuracy: 0.9325 - val_loss: 0.0347 - val_accuracy: 0.9976
Epoch 7/12
119/119 [==============================] - 36s 301ms/step - loss: 0.1707 - accuracy: 0.9372 - val_loss: 0.0193 - val_accuracy: 1.0000
Epoch 8/12
119/119 [==============================] - 34s 287ms/step - loss: 0.1471 - accuracy: 0.9454 - val_loss: 0.0116 - val_accuracy: 1.0000
Epoch 9/12
119/119 [==============================] - 33s 281ms/step - loss: 0.1358 - accuracy: 0.9504 - val_loss: 0.0105 - val_accuracy: 1.0000
Epoch 10/12
119/119 [==============================] - 34s 282ms/step - loss: 0.1229 - accuracy: 0.9552 - val_loss: 0.0030 - val_accuracy: 1.0000
Epoch 11/12
119/119 [==============================] - 33s 280ms/step - loss: 0.1175 - accuracy: 0.9565 - val_loss: 0.0037 - val_accuracy: 1.0000
Epoch 12/12
119/119 [==============================] - 33s 280ms/step - loss: 0.1107 - accuracy: 0.9593 - val_loss: 0.0098 - val_accuracy: 0.9976
```

รูปที่ 4.28 ผลการฝึกสอนโมเดลแบ่งประเภทสลิปอิเล็กทรอนิกส์



รูปที่ 4.29 กราฟค่า Loss, Accuracy ของโมเดลแบ่งประเภทสลิปอิเล็กทรอนิกส์

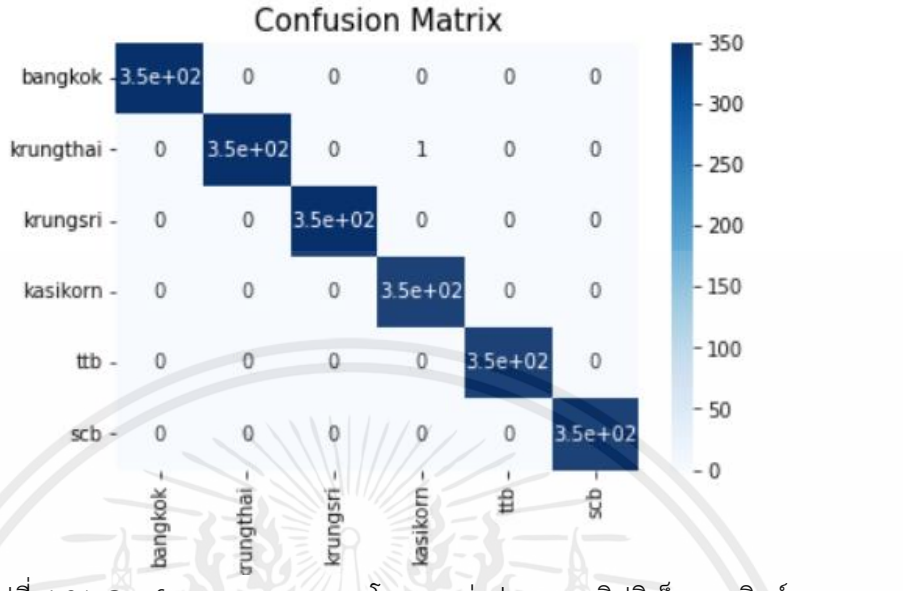
```
66/66 [==============================] - 1s 13ms/step - loss: 0.0013 - accuracy: 0.9995
Test Accuracy 1.00
```

รูปที่ 4.30 ผลการทดสอบโมเดลแบ่งประเภทสลิปอิเล็กทรอนิกส์

4.5.2 ประเมินประสิทธิภาพด้วย Confusion Matrix

ประสิทธิภาพของโมเดลถูกประเมิณจากการทำ Confusion Matrix โดยมีผลลัพธ์ที่ได้แสดงได้ดังรูปที่ 4.31 โดยมีคะแนน F1 Score คือ 0.99

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.31 Confusion Matrix ของโมเดลแบ่งประเภทสลิปอิเล็กทรอนิกส์

4.6 ทดสอบการแบ่งประเภทสลิปตามธนาคารของสลิปการเงินแบบกระดาษ

ในการทดสอบประสิทธิภาพได้ประเมินประสิทธิภาพของโมเดลแบ่งประเภทสลิปตามธนาคารของสลิปการเงินแบบกระดาษด้วยการทำ Confusion Matrix และกำหนดให้ Confusion Matrix มีขนาดเป็น 6x6 เนื่องจาก Class ของข้อมูลที่มี 6 Class คือรูปภาพสลิปการเงินแบบกระดาษที่ประกอบไปด้วยธนาคาร BBL, KTB, BAY, Kbank, TTB และ SCB โดยค่าของ TP, TN, FN และ FN บอกจำนวนของรูปภาพที่เกิดจากการประเมินประสิทธิภาพ

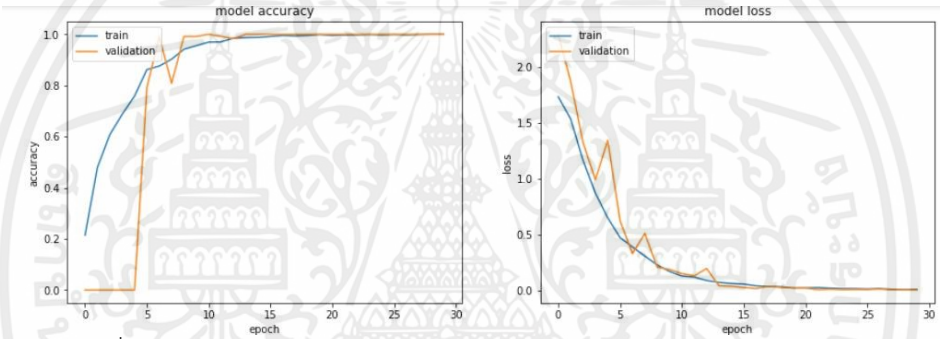
4.6.1 ทดสอบด้วยชุดข้อมูล

ได้ทำการฝึกสอนโมเดลเป็นจำนวน Epoch = 30 รอบ โดยมีผลการฝึกสอนโมเดล กราฟค่า Loss, Accuracy กับจำนวนรอบที่ฝึกสอน และผลการทดสอบโมเดล แสดงได้ดังรูปที่ 4.32, 4.33 และ 4.34 ตามลำดับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
Epoch 18/30
34/34 [==============================] - 6s 164ms/step - loss: 0.0349 - accuracy: 0.9935 - val_loss: 0.0366 - val_accuracy: 1.0000
Epoch 19/30
34/34 [==============================] - 6s 163ms/step - loss: 0.0335 - accuracy: 0.9954 - val_loss: 0.0264 - val_accuracy: 1.0000
Epoch 20/30
34/34 [==============================] - 5s 161ms/step - loss: 0.0214 - accuracy: 0.9991 - val_loss: 0.0228 - val_accuracy: 1.0000
Epoch 21/30
34/34 [==============================] - 5s 160ms/step - loss: 0.0230 - accuracy: 0.9954 - val_loss: 0.0213 - val_accuracy: 1.0000
Epoch 22/30
34/34 [==============================] - 5s 162ms/step - loss: 0.0256 - accuracy: 0.9972 - val_loss: 0.0073 - val_accuracy: 1.0000
Epoch 23/30
34/34 [==============================] - 6s 162ms/step - loss: 0.0204 - accuracy: 0.9972 - val_loss: 0.0113 - val_accuracy: 1.0000
Epoch 24/30
34/34 [==============================] - 6s 162ms/step - loss: 0.0169 - accuracy: 0.9991 - val_loss: 0.0070 - val_accuracy: 1.0000
Epoch 25/30
34/34 [==============================] - 6s 162ms/step - loss: 0.0160 - accuracy: 0.9972 - val_loss: 0.0083 - val_accuracy: 1.0000
Epoch 26/30
34/34 [==============================] - 6s 162ms/step - loss: 0.0131 - accuracy: 0.9991 - val_loss: 0.0083 - val_accuracy: 1.0000
Epoch 27/30
34/34 [==============================] - 6s 163ms/step - loss: 0.0168 - accuracy: 0.9972 - val_loss: 0.0124 - val_accuracy: 1.0000
Epoch 28/30
34/34 [==============================] - 6s 162ms/step - loss: 0.0097 - accuracy: 0.9991 - val_loss: 0.0053 - val_accuracy: 1.0000
Epoch 29/30
34/34 [==============================] - 6s 162ms/step - loss: 0.0073 - accuracy: 1.0000 - val_loss: 0.0046 - val_accuracy: 1.0000
Epoch 30/30
34/34 [==============================] - 6s 164ms/step - loss: 0.0096 - accuracy: 1.0000 - val_loss: 0.0036 - val_accuracy: 1.0000
```

รูปที่ 4.32 ผลการฝึกสอนโมเดลแบ่งประเภทสลิปกระดาษ



รูปที่ 4.33 กราฟค่า Loss, Accuracy ของโมเดลแบ่งประเภทสลิปกระดาษ

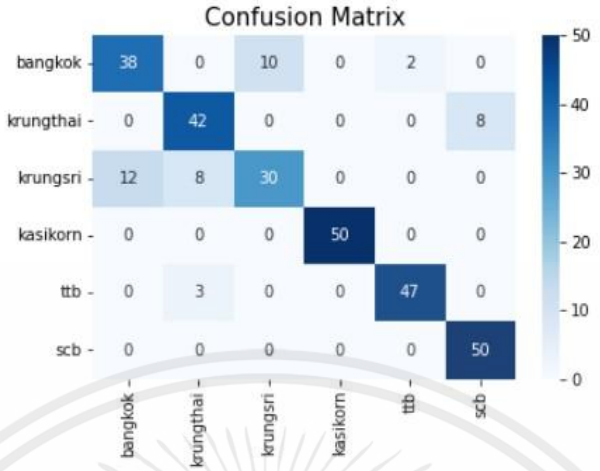
```
10/10 [==============================] - 0s 12ms/step - loss: 0.7110 - accuracy: 0.8567
Test Accuracy 0.86
```

รูปที่ 4.34 ผลการทดสอบโมเดลแบ่งประเภทสลิปกระดาษ

4.6.2 ประเมินประสิทธิภาพด้วย Confusion Matrix

ประสิทธิภาพของโมเดลถูกประเมิณจากการทำ Confusion Matrix โดยมีผลลัพธ์ที่ได้แสดงได้ดังรูปที่ 4.35 โดยมีคะแนน F1 Score คือ 0.86

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.35 Confusion Matrix ของโมเดลแบ่งประเภทสลิปกระดาษ

4.7 ทดสอบการเก็บข้อมูลสลิปการเงิน (Financial Data Extraction)

4.7.1 ทดสอบการดึงข้อมูลจากรูปภาพสลิปการเงิน

จากการทดสอบการดึงข้อมูลจากรูปภาพสลิปการเงินโดยมีอินพุตเป็นรูปภาพสลิปการเงินแบบอิเล็กทรอนิกส์แสดงดังรูป 4.36 และมีผลลัพธ์ที่ได้จากการทำ OCR แสดงดังรูป 4.37 และรูปภาพสลิปการเงินแบบกระดาษดังรูปที่ 4.38 และผลลัพธ์ที่ได้จากการทำ OCR แสดงดังรูปที่ 4.39

รูปที่ 4.36 สลิปการเงินแบบอิเล็กทรอนิกส์ก่อนเข้าสู่กระบวนการ OCR

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

@ เติมเงินสำเร็จ

19 พ.ย. 2563 - 15:05
รหัสอ้างอิง: 202011199aZ8LAat0O9TnvZ02x

จาก @ นส. ปัณฑารีย์ Wan
XXX-XXxX440-7

ไปยัง ® เติมเงินพร้อมเพย์
004999045919148

จำนวนเงิน 3.00

ข้อมูลเพิ่มเติมจากผู้ให้บริการ

น.ส. ~~อภิชญา วงศ์สมวงศ์~~ (K Plus W)

ผู้รับเงินสามารถสแกนคิวอาร์โค้ดนี้เพื่อ
ตรวจสอบสถานะการเติมเงิน

รูปที่ 4.37 ข้อความที่ได้จากการทำ OCR ของรูปภาพสลิปการเงินแบบอิเล็กทรอนิกส์

รูปที่ 4.38 สลิปการเงินแบบกระดาษก่อนเข้าสู่กระบวนการ OCR

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
[') Krungthai',
 'กรุงไทย',
 'ใบบันทึกรายการ',
 'TRANSACTION RECORD',
 'Sui        เวลา สถานที่        ธนาคารผู้ออกบัตร',
 'Date    Time    Location        Issuer Bank',
 '16/12/65 21:25 K30186              KTBA',
 'จากบัเปิเลขที่       เข้าบัญชีเถยกี่       ประเภทรายการ',
 'Flom 4/2 No.         ToA'E No        1 5',
 'จ9จ8จ9จ9450     AUTO DEP',
 'To A/C Name :นายศักดิ์สยรา แดงชาติ',
 'จำแวนเงิน                  จว่านวนอนบัตร',
 'Amount Bart                Derominston',
 '300.00',
 'ล้าดับก์          คำอรรมเนียบ',
 'Seq No              (uM Balt',
 '2217, =A       6.60',
 'eee',
 'สอบถาม/จายัดบัตรกรุงโทน',
 'Krung hat Contact Center 0-20-1111 หรือ wew krungthai com']
```

รูปที่ 4.39 ข้อความที่ได้จากการทำ OCR ของรูปภาพสลิปการเงินแบบกระดาษ

4.7.2 ทดสอบการทำ Data Frame

ทดสอบการทำ Data Frame โดยการดึงข้อมูลที่สนใจหรือ Financial Data Extraction มีเป้าหมายคือ Date, Bank, Sender, Recipient และ Amount โดยมีตัวอย่างผลลัพธ์การทำ Data Frame ของธนาคาร SCB, TTB, BBL, KTB, Kbank และ BAY แสดงตัวอย่างดังรูปที่ 4.40, 4.41, 4.42, 4.43, 4.44 และ 4.45 ตามลำดับ

	date	bank	sender	recipient	amount
0	19 พ.ย. 2563	SCB	ปัณฑาริย์ พิลึก XXX-XXX440-7	เติมเงินพร้อมเพย์ 004999045919148	100
1	19 พ.ย. 2563	SCB	ปัณฑาริย์ พิลึก XXX-XXX440-7	WATSON CO,LTD Biller ID : 010753600037401 หมาย...	156
2	19 พ.ย. 2563	SCB	ปัณฑาริย์ พิลึก XXX-XXX440-7	® เติมเงินพร้อมเพย์ 004999045919148	3
3	11 ส.ค. 2565	SCB	ชยวิชญ์ บุญรุ่ง XXX-XXX373-3	เดชา ศรธินอ XXX-XXX-8727	130
4	11 ส.ค. 2565	SCB	ชยวิชญ์ บุญรุ่ง XXX-XXX373-3	เติมเงินพร้อมเพย์ 140000934474262	25
...	...	...	...	...	...
195	09 เม.ย. 2565	SCB	ปัณฑาธริย์ พิลึก XXX-XXX440-7	อิทธิชัย สถิตย์ปัญญาพันธุ์ XXX-XXX-8600	63
196	10 เม.ย. 2565	SCB	ปัณฑาริย์ พิลึก XXX-XXX440-7	BY BOONTERM) Biller ID : 010753600031508 รหัสร...	50
197	10 เม.ย. 2565	SCB	ปัณฑาธิย์ พิลึก XXX-XXX440-7	(อุมาอิ ซานูซาน) Biller ID : 010753600031508 ...	299
198	11 เม.ย. 2565	SCB	ปัณฑาธิย์ พิลึก XXX-XXX440-7	ธิรแมธ พรหมเรือง X-9261	83
199	11 เม.ย. 2565	SCB	ปัณฑาธิย์ พิลึก XXX-XXX440-7	กาญจนา อ่อนแก้ว XXX-XXX-5854	85

รูปที่ 4.40 ตัวอย่าง Data Frame ของธนาคาร SCB

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

	date	bank	sender	recipient	amount
0	15 พ.ย. 65, 09:03 น.	TTB	ภัครพล บัวเทศ XXX-X-XX042-2 ttb	พยุงศักดิ์ สะราคำ XXX-X-XX309-6 ttb	30
1	12 พ.ย. 65, 17:18 น.	TTB	ศุภณัฐ พัสสาริกรณ์ - - XXX-X-XX927-1 ttb	รณกร ภู่ระยม้า XXX-X-XX515-8 KBANK	330
2	15 พ.ย. 65, 09:12 น.	TTB	ภัครพล บัวเทศ XXX-X-XX042-2 ttb	นางสาว ไพลินมอพิมพ์ XXX-X-XX441-3 ttb	35000
3	25 พ.ย. 65, 17:00 น.	TTB	ภัครพล บัวเทศ XXX-X-XX042-2 ttb	อาริรัตน์ ลาโพธิ์ XXX-XXX-5319 พร้อมเพย์	110
4	27 พ.ค. 65, 17:47 น.	TTB	ภัครพล บัวเทศ XXX-X-XX042-2 ttb	เจษฎา สีดำนิล XXX-X-XX299-1 ttb	700
5	5 ส.ค. 65, 09:29 น.	TTB	ภัครพล บัวเทศ XXX-X-XX042-2 ttb	เจษฎา สีดำนิล XXX-X-XX299-1 ttb	200
6	13 ก.ย. 65, 11:32 น.	TTB	ภัครพล บัวเทศ XXX-X-XX042-2 ttb	พยุงศักดิ์ สะราคำ XXX-X-XX309-6 ttb	50
7	31 ต.ค. 65, 20:52 น.	TTB	เพ็ญพิชชา แย้มงาม XXX-X-XX710-2 ttb	ภาวัญชน์ ธนฐฉไพศาล XXX-XXX-6514 พร้อมเพย์	100
8	29 พ.ค. 65, 09:53 น.	TTB	ธิติภาส คงไทย XXX-X-XX212-7 ttb	ภรคัตสรร คอร์เปอร์เธชั่ XXX-X-XX283-3 SCB	5644
9	14 ต.ค. 65, 20:23 น.	TTB	ธิติภาส คงไทย วงะห ม 2197 ttb	ชิษณุพงศ์ จันทรัตนันท์ # XXX-=X-XX732-1 KTB	1000

รูปที่ 4.41 Data Frame ของธนาคาร TTB

	date	bank	sender	recipient	amount
0	11 ธ.ค. 65, 21:09	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	ศรัณญ์ สุวรรณพล 044-3-XXX065 ธนาคารกสิกรไทย	20
1	16 ส.ค. 65, 08:36	BBL	จรัสพร 654-5-XXX585 ธนาคารกรุงเทพ	ชิดชนก จำรัส 096-XXX-6898 พร้อมเพย์	124
2	11 ธ.ค. 65, 21:09	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	ศรัณญ์ สุวรรณพล 044-3-XXX065 ธนาคารกสิกรไทย	20
3	16 ส.ค. 65, 10:33	BBL	จรัสพร 654-5-XXX585 ธนาคารกรุงเทพ	อาริยา ป้อมแก้ว เติมเงินพร้อมเพย์ / G-Wallet K...	19
4	11 ธ.ค. 65, 21:09	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	ศรัณญ์ สุวรรณพล 044-3-XXX065 ธนาคารกสิกรไทย	20
...	...	...	...	...	...
195	08 มี.ค. 65, 11:38	BBL	จรัสพร 654-5-XXX585 ธนาคารกรุงเทพ	ไวร์เลส เน็ทเวอร์ ค จำกัด Service Code:AIS12C...	50
196	08 มี.ค. 65, 10:34	BBL	จรัสพร 654-5-XXX585 ธนาคารกรุงเทพ	osisnv อุดรวงศ์ เติมเงินพร้อมเพย์ / G-Wallet K...	20
197	08 มี.ค. 65, 19:31	BBL	จรัสพร 654-5-XXX585 ธนาคารกรุงเทพ	มยุรา โพธิ์ศิร 064-XXX-5980 พร้อมเพย์	10
198	02 ส.ค. 65, 10:34	BBL	จรัสพร 654-5-XXX585 ธนาคารกรุงเทพ	จราพร เติมโภค เติมเงินพร้อมเพย์ / G-Wallet ทรู...	40
199	11 ธ.ค. 65, 20:47	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	ศรัณญ์ สุวรรณพล 044-3-XXX065 ธนาคารกสิกรไทย	20

รูปที่ 4.42 Data Frame ของธนาคาร BBL

	date	bank	sender	recipient	amount
0	15 ม.ค. 2565 - 11:24	KTB	นสุภัญญดา กำลังเจริญ กรุงไทย ว XXX-X-XX568-6	Jud วอลเล็ท (1022 หมายเลขโทรศัพท์มือถือ 10หลัก...	100
1	15 ม.ค. 2565 - 13:21	KTB	[ต] X นสภัญญดา กำลังเจริญ กรุงไทย	: XXX-X-XX568-6 Fe] กฤษณา ชากระโทก พร้อมเพย์ ...	35
2	15 ม.ค. 2565 - 13:25	KTB	นสภัญญดา กำลังเจริญ กรุงไทย XXX-X-XX568-6	ร้านถุงเงิน (010753700088205) หมายเลขประจำตัว...	45
3	15 ม.ค. 2565 - 17:24	KTB	นสภัญญดา กำลังเจริญ กรุงไทย XXX-X-XX690-0	/ จ่ายบิลข้อปปี้ (24184) หมายเลขอ้างอิง = 2022...	1213
4	15 ม.ค. 2565 - 17:28	KTB	นสุภัญญดา กำลังเจริญ กรุงไทย : XXX-X-XX690-0	/ จ่ายบิลข้อปปี้ (24184) หมายเลขอ้างอิง = 2022...	200
...	...	...	...	...	...
195	22 มี.ค. 2565 - 21:30	KTB	นสภัญญดา กำลังเจริญ กรุงไทย XXX-X-XX568-6	รัชฎาภรณ์ พลพันธ์ ไทยพาณิชย์ XXX-X-XX861-1	86
196	23 มี.ค. 2565 - 12:40	KTB	นสภัญญดา กำลังเจริญ กรุงไทย XXX-X-XX568-6	พิมพ์จันทร์ ยินเนหันหา พร้อมเพย์ XXX XXX 6638	40
197	24 มี.ค. 2565 - 12:29	KTB	202203248591579 นสภัญญดา กำลังเจริญ กรุงไทย	XXX-X-XX568-6 SCB มณี Shop (*) (01075360001028...	20
198	24 มี.ค. 2565 - 16:53	KTB	นสภัญญดา กำลังเจริญ กรุงไทย XXX-X-XX568-6	มัทนา อวนศรี พร้อมเพย์ XXX XXX 4594	1
199	24 มี.ค. 2565 - 17:56	KTB	นสภัญญดา กำลังเจริญ กรุงไทย XXX-X-XX568-6	สรศักดิ์ ธนกมลนันท์ พร้อมเพย์ XXX XXX 9491	189

รูปที่ 4.43 Data Frame ของธนาคาร KTB

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

	date	bank	sender	recipient	amount
0	29 ก.ย. 63 21:01 น.	Kbank	กันตพงศ์ บ ธ.กสิกรไทย XXX-X-X6989-X	กันตพงศ์ บุญมาก ธ.กรุงศรีอยุธยา XXX-X-X2124-X	1050
1	30 ก.ย 63 17:34น.	Kbank	กันตพงศ์ บ ธ.กสิกรไทย XXX-XX-X6989-X	ราชศักดิ์ เล็ศคปัญญา Prompt	89
2	30 ก.พ. 63 18:56 น.	Kbank	ศรัณญ์ a ธ.กสิกรไทย XXX-X-X9706-X	อัครีน DUUV Prompt ง ม _ Pay รฟ...	211
3	1 ต.ค. 63 02:40 น.	Kbank	กันตพงศ์ บ ธ.กสิกรไทย XXX-X-X6989-X	พัฒน์ภูมิ หาแก้ว ธ.กสิกรไทย XXX-X-X2758-X	25
4	1 ต.ค. 63 18:13 น.	Kbank	ศรัณญ์ a ธ.กสิกรไทย XXX-X-X9706-X	ศรัณย์วิทย์ สันติสุขคง ธ.กสิกรไทย 2XX-X-X2597-X	240
...	...	...	...	...	...
195	14 Sep 22 7:27 น.	Kbank	Kuntapong B KBank XXX-XX-X6989-X	PIRAMAI KRUEANUAN KBank XXX-X-X6683-X	115
196	15 Sep 22 1:17 น.	Kbank	Kuntapong B KBank XXX-XX-X6989-X	SOK ROEUN KBank XXX-X-X43 59-X	146
197	15 ก.ย. 65 11:18 น.	Kbank	อนุชา โ ธ.กสิกรไทย XXX-X-X1883-X	TrueMoney Wallet 0617279429	100
198	15 Sep 22 1:32 น.	Kbank	Kuntapong B KBank XXX-XX-X6989-X	KITSADA JAIDE Prompt Pay PromptPay ...	50
199	15 Sep 22 3:10 น.	Kbank	Kuntapong B KBank XXX-XX-X6989-X	KUNLARAT SAIBUA KBank XXX-X-X3567-X	15

รูปที่ 4.44 Data Frame ของธนาคาร Kbank

	date	bank	sender	recipient	amount	referral code
0	19 ส.ก. 2565	BAY	AVIRUT CHANTH XXX-4-61542-X	คณากร คำชาติ XXX-XXX-9975	500	BAYM1419942310 een a H
1	01 nov 2020	BAY	PITIPAT SANPR XXX-1-31228-X	NETNAPAR KASI XXX-XXX-6790	40	BAYM423056890
2	27 ct 2020	BAY	PITIPAT SANPR XXX-1-31228-X	KTB G-WALLET 006-990-2-XXXX597-6	100	BAYM417049833
3	26 ct 2020	BAY	PITIPAT SANPR XXX-1-31228-X	NULL NULL 140-000-9-XXXX249-0	155	BAYM416408018
4	26 ct 2020	BAY	PITIPAT SANPR XXX-1-31228-X	MR. SONGWUT AREEROB 004-000-0-XXXX307-7	73	BAYM416236092
...	...	...	...	...	...	...
195	17 ส.ค. 2565	BAY	KITSADA JAIDE XXX-4-62360-X	กนกพร ฟักเขียว เธค ...	20	BAYM1415501275 Peet nH
196	17 ส.ค. 2565	BAY	KITSADA JAIDE XXX-4-62360-X	เสาวลักษณ์ เครือนวล ...	60	BAYM1416062964 SE Ea
197	18 Gก. 2565	BAY	KITSADA JAIDE XXX-4-62360-X	ปิติพัฒน์ สารปรัง ...	55	BAYM1417878358 ...
198	18 ส.ค. 2565	BAY	KITSADA JAIDE XXX-4-62360-X	PAVISA NILPEC XXX-4-05338-X จำนวนเงิน	98	BAYM1418106165 ร เล
199	06 ก.ย. 2565	BAY	KITSADA JAIDE XXX-4-62360-X	กันตพงศ์ บุญมาก ...	20	BAYM1459632467 r=" 'a Th

รูปที่ 4.45 Data Frame ของธนาคาร BAY

4.7.3 ทดสอบการเก็บข้อมูลในรูปแบบไฟล์ .xlsx

ทดสอบการทำไฟล์ .xlsx โดยมีผลลัพธ์การทำของธนาคาร SCB, TTB, BBL, KTB, Kbank และ BAY แสดงตัวอย่างดังรูปที่ 4.46, 4.47, 4.48, 4.49, 4.50 และ 4.51 ตามลำดับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

667	15 ก.พ. 2565	SCB	ปัณฑาธรีย์ พิลึก XXX-XX9X440-7	พาณญิภัคร ค้าบอง XXX-XXX-6847	124
668	15 ม.ค. 2564	SCB	ปัณฑาธรีย์ Wan XXX-XXX440-7	สยาม สงวนรัมย์ XXX-XXX176-3	285
669	15 ก.พ. 2565	SCB	ปัณฑารีย์ wan XXX-XXX440-7	® เติมเงินพร้อมเพย์ 006990306915703	150
670	16 ก.พ. 2565	SCB	ปัณฑารีย์ wan XXX-XXX440-7	CinepleX Group Biller ID : 010555206135	200
671	16 ก.พ. 2565	SCB	ปัณฑารีย์ wan XXX-XXX440-7	® เติมเงินพร้อมเพย์ 004999001476307	400
672	17 ก.พ. 2565	SCB	ปัณฑาธรีย์ Wan XXX-XX9X440-7	พาณญิภัคร ค้าบอง XXX-XXX-6847	715
673	17 ก.พ. 2565	SCB	ปัณฑารีย์ Wan XXX-XXX44.0-7	shop (เทคโนอินเตอร์) Biller ID : 0107536	40
674	17 ก.พ. 2565	SCB	ปัณฑารีย์ wan XXX-XXX440-7	® เติมเงินพร้อมเพย์ 004999006397912	50
675	22 ก.พ. 2565	SCB	ปัณฑารีย์ Wan XXX-XXX440-7	shop (กะเพราปิลเลี่ยนกรุป์) Biller ID : 01	40
676	22 ก.พ. 2565	SCB	ปัณฑารีย์ Wan XXX-XXX440-7	© เติมเงินพร้อมเพย์ 006990306915703	150
677	24 ก.พ. 2565	SCB	ปัณฑาธรีย์ Wan XXX-XX9X440-7	พาณญิภัคร ค้าบอง XXX-XXX-6847	200
678	24 ก.พ. 2565	SCB	ปัณฑาธีย์ Wan XXX-XXX440-7	ปัณฑาธีย์ Wan X-2751	1000
679	26 ก.พ. 2565	SCB	ปัณฑารีย์ Wan XXX-XX9X440-7	© เติมเงินพร้อมเพย์ 006990306915703	100
680	26 ก.พ. 2565	SCB	ปัณฑาธีย์ Wan XXX-XXX440-7	พ.ส.วริษฐา เจริญสุน X-5663	55
681	26 ก.พ. 2565	SCB	ปัณฑาธีย์ Wan XXX-XXX440-7	ไพเกุล แสงสุง 3-4005-0057X-XX-X	62
682	27 ก.พ. 2565	SCB	ปัณฑาธีย์ Wan XXX-XXX440-7	วรกร สุงรัง X-4342	300
683	27 ก.พ. 2565	SCB	ปัณฑารีย์ wan XXX-XXX440-7	® เติมเงินพร้อมเพย์ 004999028650991	60
684	28 ก.พ. 2565	SCB	ปัณฑาธรีย์ Wan XXX-XX9X440-7	ปุญกวัฒน์ บุษดี XXX-XXX-3438	59
685	01 มี.ค. 2565	SCB	ปัณฑารีย์ Wan XXX-XXX440-7	© เติมเงินพร้อมเพย์ 006990306915703	100
686	01 มี.ค. 2565	SCB	ปัณฑาธรีย์ Wan XXX-XXX440-7	ธนวรรณ ถนอมศรี XXX-XXX-6568	75
687	02 มี.ค. 2565	SCB	ปัณฑารีย์ wan XXX-XXX440-7	® เติมเงินพร้อมเพย์ 004999051026077	30

รูปที่ 4.46 .xlsx ของธนาคาร SCB

505	18 ต.ค. 2565	TTB	ศุภณัฐ พัสสาริกรณ์ XXX-X-XX929-7 ttb	พลอยนวล ชนะบุญ XXX-XXX-6643	104
506	18 ต.ค. 2565	TTB	ศุภณัฐ พัสสาริกรณ์ XXX-X-XX929-7 ttb	® เติมเงินพร้อมเพย์ 004999035584969	20
507	18 ต.ค. 2565	TTB	ศุภณัฐ พัสสาริกรณ์ XXX-X-XX929-7 ttb	ร้านถุงเงิน (ข้าวกล่องบ้านซี่ด) Biller ID : 0	20
508	18 ต.ค. 2565	TTB	ศุภณัฐ พัสสาริกรณ์ XXX-X-XX929-7 ttb	ร้านถุงเงิน (ร้านวิไลวรรณลูกชุป) Biller ID	20
509	18 ต.ค. 2565	TTB	ศุภณัฐ พัสสาริกรณ์ XXX-X-XX929-7 ttb	® เติมเงินพร้อมเพย์ 004999067921840	25
510	19 ต.ค. 2565	TTB	ศุภณัฐ พัสสาริกรณ์ XXX-X-XX929-7 ttb	พลอยนวล ชนะบุญ XXX-XXX-6643	20
511	19 ต.ค. 2565	TTB	ศุภณัฐ พัสสาริกรณ์ XXX-X-XX929-7 ttb	ช้อป (เดอะบูนคาเฟ่) Biller ID : 010753600	39
512	19 ต.ค. 2565	TTB	ศุภณัฐ พัสสาริกรณ์ XXX-X-XX929-7 ttb	น.ส.รัชนีกร โมอ่อน XXX-XXXX-0690	300
513	20 ต.ค. 2565	TTB	ศุภณัฐ พัสสาริกรณ์ XXX-X-XX929-7 ttb	ทอดหอย ป้าพัช) ชื่อบัญชี: อารีย์ รุ่งสง่า E	55
514	21 มี.ค. 2565	TTB	ศุภณัฐ พัสสาริกรณ์ XXX-X-XX929-7 ttb	MONEY CO.,LTD. Biller ID : 0105546149	200
515	22 มี.ค. 2565	TTB	ศุภณัฐ พัสสาริกรณ์ XXX-X-XX929-7 ttb	shop (BCR-INTHANIN CHALONG KRUN) I	125
516	22 มี.ค. 2565	TTB	ศุภณัฐ พัสสาริกรณ์ XXX-X-XX929-7 ttb	shop (BCR-INTHANIN CHALONG KRUN) I	55
517	22 มี.ค. 2565	TTB	ศุภณัฐ พัสสาริกรณ์ XXX-X-XX929-7 ttb	ทอดหอย ป้าพัช) ชื่อบัญชี: อารีย์ รุ่งสง่า E	105
518	22 มี.ค. 2565	TTB	ศุภณัฐ พัสสาริกรณ์ XXX-X-XX929-7 ttb	SHOP (ซัมฟรีด) ชื่อบัญชี: เกษมศักดิ์ ฉันท	35
519	23 มี.ค. 2565	TTB	ศุภณัฐ พัสสาริกรณ์ XXX-X-XX929-7 ttb	เติมเงินพร้อมเพย์ 004999065269176	50
520	23 มี.ค. 2565	TTB	ศุภณัฐ พัสสาริกรณ์ XXX-X-XX929-7 ttb	ชัยยา สมสา X-1805	75
521	23 มี.ค. 2565	TTB	ศุภณัฐ พัสสาริกรณ์ XXX-X-XX929-7 ttb	shop (BCR-INTHANIN CHALONG KRUN) I	50
522	24 มี.ค. 2565	TTB	ศุภณัฐ พัสสาริกรณ์ XXX-X-XX929-7 ttb	shop (BCR-INTHANIN CHALONG KRUN) I	130
523	24 มี.ค. 2565	TTB	ศุภณัฐ พัสสาริกรณ์ XXX-X-XX929-7 ttb	MONEY CO.,LTD. Biller ID : 0105546149	500
524	25 มี.ค. 2565	TTB	ศุภณัฐ พัสสาริกรณ์ XXX-X-XX929-7 ttb	ไปยัง siranee Biller ID : 0107536000374	100
525	25 มี.ค. 2565	TTB	ศุภณัฐ พัสสาริกรณ์ XXX-X-XX929-7 ttb	shop (BCR-INTHANIN CHALONG KRUN) I	135

รูปที่ 4.47 xlsx ของธนาคาร TTB

203	07 ธ.ค. 2564	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	® เติมเงินพร้อมเพย์ 006990306915703	130
204	14 พ.ย. 2565	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	เพ็ญพิชชา แย้มงาม XXX-XXXX-1072	40
205	07 ธ.ค. 2564	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	ทอดหอย ป้าพัช) Biller ID : 01075360001	50
206	07 ธ.ค. 2564	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	กาญจนา อ่อนแก้ว XXX-XXX-5854	55
207	07 ธ.ค. 2564	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	พาณญิภัคร ค้าบอง XXX-XXX-6847	20
208	18 ต.ค. 2563	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	พรนภัส งามพาณิชยกิจ XXX-XXX-9875	103
209	08 ธ.ค. 2564	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	ปัณฑาธีย์ wan 1-6099-0055X-XX-X	400
210	08 ธ.ค. 2564	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	ไปยัง TUNGNGERN Biller ID : 010753700	95
211	16 พ.ย. 2565	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	C RATCH) Biller ID : 010753600031508 ร	1191
212	21 ต.ค. 2563	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	พาณญิภัคร Arve" XXX-XXX-6847	84
213	25 มิ.ย. 2565	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	shop (RAYONGPHUEANPHE) Biller ID : 01	70
214	25 มิ.ย. 2565	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	(COFF CIRCLE) Biller ID : 0107536000315	140
215	25 มิ.ย. 2565	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	AND INSPI-DEPT) Biller ID : 0107536000	159
216	25 มิ.ย. 2565	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	AND INSPI-DEPT) Biller ID : 0107536000	189
217	26 มิ.ย. 2565	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	น.ส.รธรสา สุพรรณกนก XXX-XXX-7722	30
218	26 มิ.ย. 2565	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	shop (บ้านพัฒก์) Biller ID : 01075360003	145
219	26 มิ.ย. 2565	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	เติมเงินพร้อมเพย์ 004999085471372	105
220	26 มิ.ย. 2565	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	(TOPS CARE- CHAENGWATTANA) Biller I	134
221	27 มิ.ย. 2565	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	ไปยัง TUNGNGERN Biller ID : 010753700	35
222	27 มิ.ย. 2565	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	แพรพฟันวา เทศประสิทธิ์ XXX-XXX720-9	30

รูปที่ 4.48 xlsx ของธนาคาร BBL

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

304	15 ก.ค. 2565	KTB	นสภิญญดา กำลังเจริญ กรุงไทย ร	XXX-X-X	shop (TOPS-CHAENGWATTANA) Biller ID	67
305	16 ก.ค. 2565	KTB	นสภิญญดา กำลังเจริญ กรุงไทย ร	XXX-X-X	© เติมเงินพร้อมเพย์ 004999057919432	100
306	16 ก.ค. 2565	KTB	นสภิญญดา กำลังเจริญ กรุงไทย ร	XXX-X-X	โรสท์ รันเนอร์) Biller ID : 010753600031	130
307	03 ม.ค. 2565	KTB	นสภิญญดา กำลังเจริญ กรุงไทย ร	XXX-X-X	ปัณเฑาธิย์ wan 1-6099-0055X-XX-X	150
308	16 ก.ค. 2565	KTB	นสภิญญดา กำลังเจริญ กรุงไทย ร	XXX-X-X	เติมเงินพร้อมเพย์ 004999059857804	329
309	16 ก.ค. 2565	KTB	นสภิญญดา กำลังเจริญ กรุงไทย ร	XXX-X-X	(ESPRESSOMAN SUPPLY) Biller ID : 01075	100
310	16 ก.ค. 2565	KTB	นสภิญญดา กำลังเจริญ กรุงไทย ร	XXX-X-X	สุภาพร สยามนิกร 3-8005-0021X-XX-X	315
311	17 ก.ค. 2565	KTB	นสภิญญดา กำลังเจริญ กรุงไทย ร	XXX-X-X	นรินทร์ โสภาบุตร X-6679	120
312	17 ก.ค. 2565	KTB	นสภิญญดา กำลังเจริญ กรุงไทย ร	XXX-X-X	AT BTS) Biller ID : 010753600031508 รหัส	26
313	17 ก.ค. 2565	KTB	นสภิญญดา กำลังเจริญ กรุงไทย ร	XXX-X-X	AT BTS Biller ID : 010555201212600 ชื่อล	52
314	17 ก.ค. 2565	KTB	นสภิญญดา กำลังเจริญ กรุงไทย ร	XXX-X-X	(พีค กะบุ) ชื่อบัญชี: จิรทีปต์ เลิศอุฬาร Bill	1067
315	05 ม.ค. 2565	KTB	นสภิญญดา กำลังเจริญ กรุงไทย ร	XXX-X-X	von. นิปเปอร์อิสตาเบสท์ X-9144	730
316	18 ก.ค. 2565	KTB	นสภิญญดา กำลังเจริญ กรุงไทย ร	XXX-X-X	shop (TOPS-CHAENGWATTANA) Biller ID	167
317	19 ก.ค. 2565	KTB	นสภิญญดา กำลังเจริญ กรุงไทย ร	XXX-X-X	ไปยัง TUNGNGERN Biller ID : 010753700	35
318	19 ก.ค. 2565	KTB	นสภิญญดา กำลังเจริญ กรุงไทย ร	XXX-X-X	เบญจวีณ์ ศรีศุภรหัศน์ XXX-XXX-9609	60
319	06 ม.ค. 2565	KTB	นสภิญญดา กำลังเจริญ กรุงไทย ร	XXX-X-X	ปัณเฑาธีย์ wan 1-6099-0055X-XX-X	100
320	20 ก.ค. 2565	KTB	นสภิญญดา กำลังเจริญ กรุงไทย ร	XXX-X-X	© เติมเงินพร้อมเพย์ 004999070025613	45
321	20 ก.ค. 2565	KTB	นสภิญญดา กำลังเจริญ กรุงไทย ร	XXX-X-X	shop (TOPS-CHAENGWATTANA) Biller ID	34
322	20 ก.ค. 2565	KTB	นสภิญญดา กำลังเจริญ กรุงไทย ร	XXX-X-X	ไปยัง TUNGNGERN Biller ID : 010753700	40

รูปที่ 4.49 xlsx ของธนาคาร KTB

1	date	bank	sender	recipient	amount
2	14 ก.ย. 2565	KBANK	Kuntapong B KBank XXX-XX-X6989-X	สัมฤทธิ์ ริตนทิพย์ XXX-XXX-6145	80
3	24 ต.ค. 2565	KBANK	Kuntapong B KBank XXX-XX-X6989-X	® เติมเงินพร้อมเพย์ 004999094360988	29
4	04 ส.ค. 2563	KBANK	Kuntapong B KBank XXX-XX-X6989-X	พาณิภัคร ค้าของ XXX-XXX255-1	36
5	24 ต.ค. 2565	KBANK	Kuntapong B KBank XXX-XX-X6989-X	® เติมเงินพร้อมเพย์ 006990306915703	70
6	25 ต.ค. 2565	KBANK	Kuntapong B KBank XXX-XX-X6989-X	® เติมเงินพร้อมเพย์ 004999043976447	45
7	25 ต.ค. 2565	KBANK	Kuntapong B KBank XXX-XX-X6989-X	gua ลีลา XXX-XXX-5952	40
8	25 ต.ค. 2565	KBANK	Kuntapong B KBank XXX-XX-X6989-X	พลอยนวล ชนะบุญ XXX-XXX-6643	275
9	25 ต.ค. 2565	KBANK	Kuntapong B KBank XXX-XX-X6989-X	ร้านถุงเงิน (สุชาดาช่าเก็ด) Biller ID : 0107	275
10	26 ต.ค. 2565	KBANK	Kuntapong B KBank XXX-XX-X6989-X	นนทกานต์ ชูสถาน X-9142	1100
11	28 ต.ค. 2565	KBANK	Kuntapong B KBank XXX-XX-X6989-X	พาณิภัคร Arve" XXX-XXX-6847	382
12	28 ต.ค. 2565	KBANK	Kuntapong B KBank XXX-XX-X6989-X	พลอยนวล ชนะบุญ XXX-XXX-6643	150
13	29 ต.ค. 2565	KBANK	Kuntapong B KBank XXX-XX-X6989-X	พาณิภัคร ค้าบอง XXX-XXX-6847	53
14	15 ก.ย. 2565	KBANK	Kuntapong B KBank XXX-XX-X6989-X	ณัชธิดา วิเศษจิรโชค 1-5603-0008X-XX-X	40
15	29 ต.ค. 2565	KBANK	Kuntapong B KBank XXX-XX-X6989-X	ไพเทูล แสงสูง 3-4005-00057X-XX-X	55
16	29 ต.ค. 2565	KBANK	Kuntapong B KBank XXX-XX-X6989-X	มานี หารมาก XXX-XXX-6490	60
17	30 ต.ค. 2565	KBANK	Kuntapong B KBank XXX-XX-X6989-X	SHOP (ร้านไข่ปลาทอดทะเลน้อย) ชื่อบัญ	20
18	30 ต.ค. 2565	KBANK	Kuntapong B KBank XXX-XX-X6989-X	ไปยัง เติมเงินพร้อมเพย์ 0049991052182	40
19	30 ต.ค. 2565	KBANK	Kuntapong B KBank XXX-XX-X6989-X	@ หมูอบบอล Biller ID : 01055610620760	60
20	31 ต.ค. 2565	KBANK	Kuntapong B KBank XXX-XX-X6989-X	ไปยัง เติมเงินพร้อมเพย์ 1400008943995	24

รูปที่ 4.50 xlsx ของธนาคาร Kbank

403	10 ก.ย. 2565	BAY	KITSADA JAIDE XXX-4-62360-X	พลอยนวล ชนะบุญ XXX-XXX-6643	20
404	10 ก.ย. 2565	BAY	KITSADA JAIDE XXX-4-62360-X	(สามโทน เดียวเนื้อน่าตก) Biller ID : 0107	120
405	11 ก.ย. 2565	BAY	KITSADA JAIDE XXX-4-62360-X	จิรภาส อมาตยกุล X-5234	114
406	11 ก.ย. 2565	BAY	KITSADA JAIDE XXX-4-62360-X	น.ส.ชวณัชฐ ทิพวัลย์ X-9940	15
407	12 ก.ย. 2565	BAY	KITSADA JAIDE XXX-4-62360-X	พาณิภัคร ค้าบอง XXX-XXX-6847	100
408	12 ก.ย. 2565	BAY	KITSADA JAIDE XXX-4-62360-X	(470 WATSONS MEGA BANGNA) Biller ID	221
409	12 ก.ย. 2565	BAY	KITSADA JAIDE XXX-4-62360-X	ชัชชญา wan XXX-XXX-4450	50
410	13 ก.ย. 2565	BAY	KITSADA JAIDE XXX-4-62360-X	พาณิภัคร Arve" XXX-XXX-6847	105
411	14 ก.ย. 2565	BAY	KITSADA JAIDE XXX-4-62360-X	พลอยนวล ชนะบุญ XXX-XXX-6643	104
412	14 ก.ย. 2565	BAY	KITSADA JAIDE XXX-4-62360-X	มาโนช สมหวัง X-6426	132
413	15 1.8. 2565	BAY	KITSADA JAIDE XXX-4-62360-X	® เติมเงินพร้อมเพย์ 004999006069719	65
414	16 ก.ย. 2565	BAY	KITSADA JAIDE XXX-4-62360-X	MONEY CO.,LTD. Biller ID : 0105546149	200
415	16 ก.ย. 2565	BAY	KITSADA JAIDE XXX-4-62360-X	ปัณเฑาธีย์ Wan X-2751	200
416	16 1.8). 2565	BAY	KITSADA JAIDE XXX-4-62360-X	พลอยนวล ชนะบุญ XXX-XXX-6643	25
417	17 ก.ย. 2565	BAY	KITSADA JAIDE XXX-4-62360-X	(PORJAI UNIFORM) Biller ID : 010753600	340
418	17 ก.ย. 2565	BAY	KITSADA JAIDE XXX-4-62360-X	พลอยนวล ชนะบุญ XXX-XXX-6643	110
419	17 ก.ย. 2565	BAY	KITSADA JAIDE XXX-4-62360-X	วิษณุ กองศรีมา 1-4506-0002X-XX-X	50
420	17 ก.ย. 2565	BAY	KITSADA JAIDE XXX-4-62360-X	วิษณุ กองศรีมา 1-4506-0002X-XX-X	10
421	18 1.2. 2565	BAY	KITSADA JAIDE XXX-4-62360-X	สาวิตรี แก้วขุมเหล็ก XXX-XXX-4942	370
422	18 ก.ย. 2565	BAY	KITSADA JAIDE XXX-4-62360-X	shop (TOPS-LADKRABANG) Biller ID : 010	249

รูปที่ 4.51 xlsx ของธนาคาร BAY

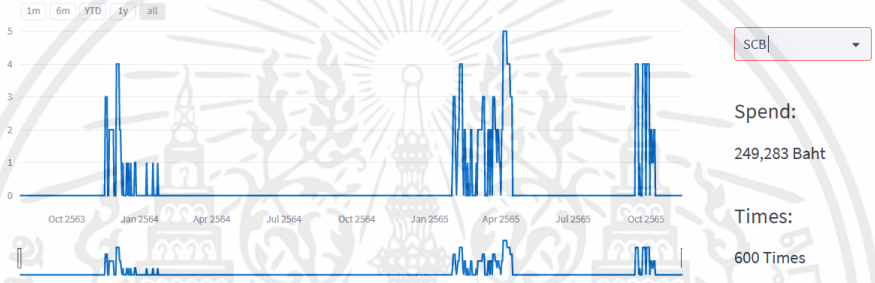
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.8 ทดสอบการแสดงผลข้อมูล

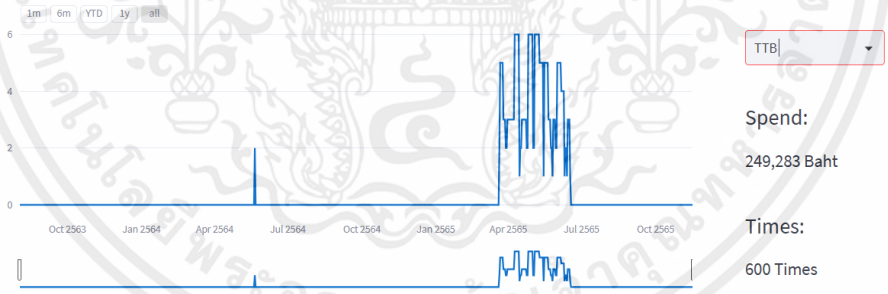
จากการทดสอบการแสดงผลข้อมูลโดยใช้ข้อมูลของธนาคารกรุงเทพ ธนาคารกรุงไทย ธนาคารกสิกรไทย ธนาคารไทยพาณิชย์ ธนาคารกรุงศรีอยุธยา และธนาคารทหารไทยธน จากรูปที่ 4.46, 4.47, 4.48, 4.49, 4.50 และ 4.51

4.8.1 ทดสอบกราฟวันเวลาที่ทำธุรกรรม

ทดสอบการเลือกธนาคารในการแสดงผลกราฟวันเวลาที่ทำธุรกรรมโดยมีผลลัพธ์ของธนาคาร SCB, TTB, BBL, KTB, Kbank และ BAY แสดงดังรูปที่ 4.52, 4.53, 4.54, 4.55, 4.56 และ 4.57 ตามลำดับ

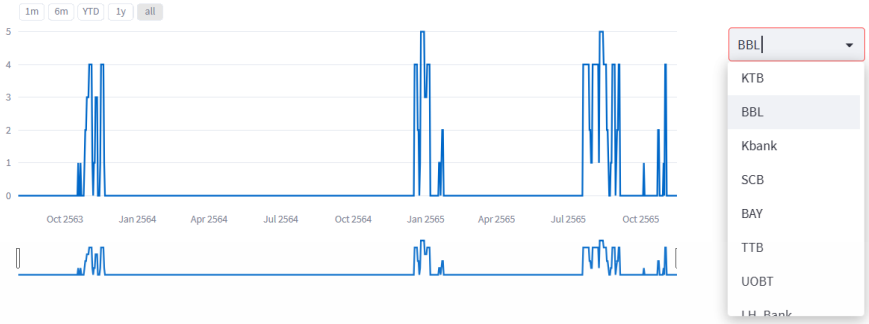


รูปที่ 4.52 วันเวลาที่ทำธุรกรรมของธนาคาร SCB

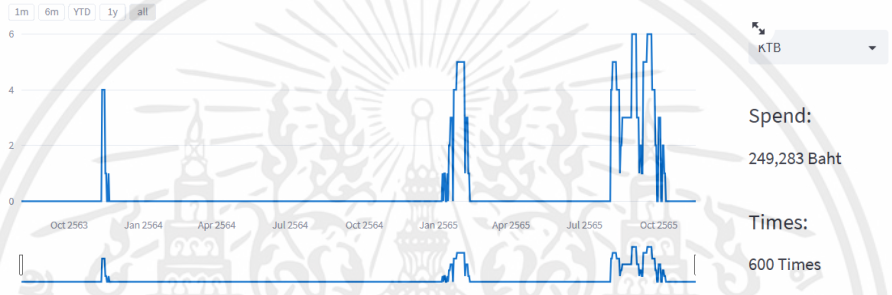


รูปที่ 4.53 วันเวลาที่ทำธุรกรรมของธนาคาร TTB

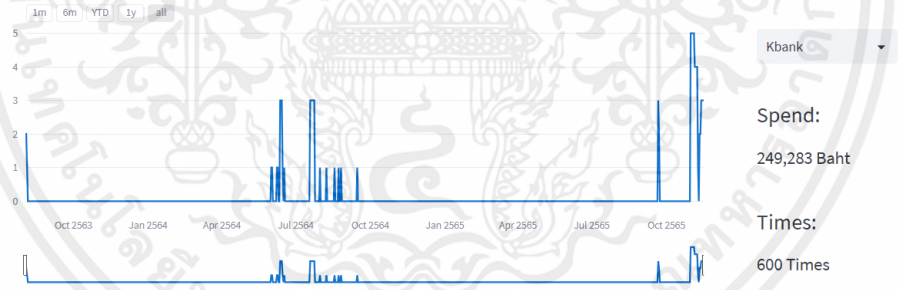
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.54 วันเวลาที่ทำธุรกรรมของธนาคาร BBL



รูปที่ 4.55 วันเวลาที่ทำธุรกรรมของธนาคาร KTB



รูปที่ 4.56 วันเวลาที่ทำธุรกรรมของธนาคาร Kbank

รูปที่ 4.57 วันเวลาที่ทำธุรกรรมของธนาคาร BAY

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.8.2 ทดสอบแผนภูมิแท่งของจำนวนครั้งในการโอนเงิน และจำนวนเงินทั้งหมดของแต่ละธนาคาร

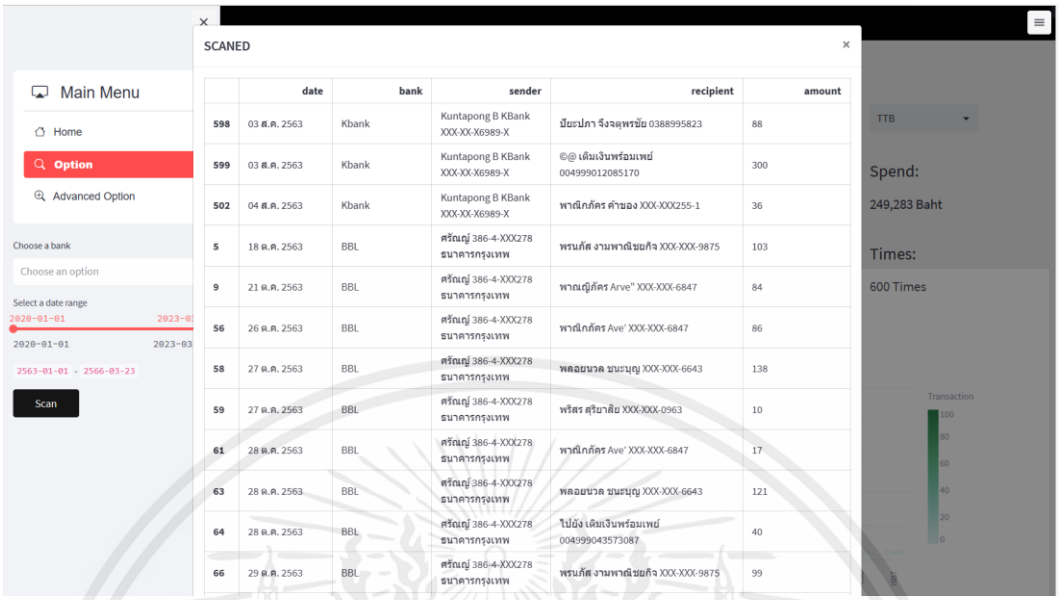
ทดสอบการเลือกธนาคารในการแสดงแผนภูมิแท่งของจำนวนครั้งในการโอนเงิน และจำนวนเงินทั้งหมดของแต่ละธนาคารที่ทำธุรกรรม โดยมีผลลัพธ์แสดงดังรูปที่ 4.58

รูปที่ 4.58 แผนภูมิแท่งของของจำนวนครั้งในการโอนเงิน และจำนวนเงินทั้งหมดของแต่ละธนาคาร

4.8.3 ทดสอบ Option

ในการทดสอบ Option ทำการทดสอบโดยการเลือก Filter ไปที่ธนาคาร SCB และ BAY โดยแสดงข้อมูลก่อนใช้ Filter ดังรูปที่ 4.59 และข้อมูลหลัง Filter แสดงดังรูปที่ 4.60 และ Filter ข้อมูลจากเวลามีผลลัพธ์ดังรูปที่ 4.61

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



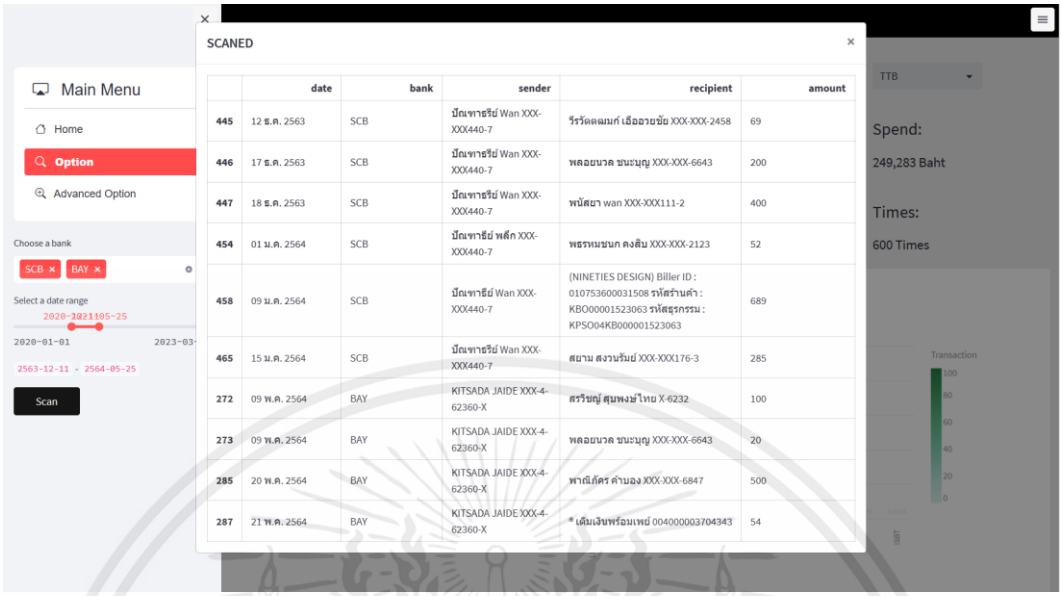
	date	bank	sender	recipient	amount
598	03 ส.ค. 2563	Kbank	Kuntapong B KBank XXX-XX-X6989-X	ปิยะปภา จึงจตุพรชัย 0388995823	88
599	03 ส.ค. 2563	Kbank	Kuntapong B KBank XXX-XX-X6989-X	©@ เติมเงินพร้อมเพย์ 004999012085170	300
502	04 ส.ค. 2563	Kbank	Kuntapong B KBank XXX-XX-X6989-X	พาณิกภัคร คำของ XXX-XXX255-1	36
5	18 ส.ค. 2563	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	พรนภัส งามพาณิชยกิจ XXX-XXX-9875	103
9	21 ส.ค. 2563	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	พาณภัคร Arve" XXX-XXX-6847	84
56	26 ส.ค. 2563	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	พาณิกภัคร Ave' XXX-XXX-6847	86
58	27 ส.ค. 2563	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	พลอยนวล ขนะบุญ XXX-XXX-6643	138
59	27 ส.ค. 2563	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	พริสร สุริยาลัย XXX-XXX-0963	10
61	28 ส.ค. 2563	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	พาณิกภัคร Ave' XXX-XXX-6847	17
63	28 ส.ค. 2563	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	พลอยนวล ขนะบุญ XXX-XXX-6643	121
64	28 ส.ค. 2563	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	ไปยัง เติมเงินพร้อมเพย์ 004999043573087	40
66	29 ส.ค. 2563	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	พรนภัส งามพาณิชยกิจ XXX-XXX-9875	99

4.59 ข้อมูลที่ถูกแสดงก่อนผ่าน Filter

444	09 ธ.ค. 2563	SCB	ปัณฑาธรีย์ Wan XXX-XXX440-7	พรนภัส งามพาณิชยกิจ XXX-XXX-9875	45
445	12 ธ.ค. 2563	SCB	ปัณฑาธรีย์ Wan XXX-XXX440-7	วีรวัตตฒมก์ เอืออวยชัย XXX-XXX-2458	69
446	17 ธ.ค. 2563	SCB	ปัณฑาธรีย์ Wan XXX-XXX440-7	พลอยนวล ขนะบุญ XXX-XXX-6643	200
447	18 ธ.ค. 2563	SCB	ปัณฑาธรีย์ Wan XXX-XXX440-7	พนัสยา wan XXX-XXX111-2	400
454	01 ม.ค. 2564	SCB	ปัณฑาธีย์ พลึก XXX-XXX440-7	พรรหมชนก คงลิม XXX-XXX-2123	52
458	09 ม.ค. 2564	SCB	ปัณฑาธีย์ Wan XXX-XXX440-7	(NINETIES DESIGN) Biller ID : 010753600031508 รหัสร้านค้า : KBO00001523063 รหัสธุรกรรม : KPSO04KB000001523063	689
465	15 ม.ค. 2564	SCB	ปัณฑาธรีย์ Wan XXX-XXX440-7	สยาม สงวนริมย์ XXX-XXX176-3	285
272	09 พ.ค. 2564	BAY	KITSADA JAIDE XXX-4-62360-X	สรวิชญ์ สุบพงษ์ไทย X-6232	100
273	09 พ.ค. 2564	BAY	KITSADA JAIDE XXX-4-62360-X	พลอยนวล ขนะบุญ XXX-XXX-6643	20
285	20 พ.ค. 2564	BAY	KITSADA JAIDE XXX-4-62360-X	พาณิภัคร คำของ XXX-XXX-6847	500
287	21 พ.ค. 2564	BAY	KITSADA JAIDE XXX-4-62360-X	* เติมเงินพร้อมเพย์ 004000003704343	54
422	21 ม.ค. 2565	SCB	ปัณฑาริย์ Wan XXX-XXX440-7	© เติมเงินพร้อมเพย์ 004999028650991	30
423	22 ม.ค. 2565	SCB	ปัณฑาธรีย์ Wan XXX-XXX440-7	von. จี.อาร์.โอ.กรุป์ X-3564	1540
424	22 ม.ค. 2565	SCB	ปัณฑาริย์ wan XXX-XXX440-7	* เติมเงินพร้อมเพย์ 004999071551435	145

4.60 ข้อมูลหลังผ่าน Filter ธนาคาร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



	date	bank	sender	recipient	amount
445	12 ธ.ค. 2563	SCB	ปัณเทาธรีย์ Wan XXX-XXX440-7	วีรวัตผลมก์ เอืออวยชัย XXX-XXX-2458	69
446	17 ธ.ค. 2563	SCB	ปัณเทาธรีย์ Wan XXX-XXX440-7	พลอยนวล ชนะบุญ XXX-XXX-6643	200
447	18 ธ.ค. 2563	SCB	ปัณเทาธรีย์ Wan XXX-XXX440-7	พนัสยา wan XXX-XXX111-2	400
454	01 ม.ค. 2564	SCB	ปัณเทาธีย์ พลัก XXX-XXX440-7	พธรหมชนก คงสิม XXX-XXX-2123	52
458	09 ม.ค. 2564	SCB	ปัณเทาธีย์ Wan XXX-XXX440-7	(NINETIES DESIGN) Biller ID : 010753600031508 รหัสร้านค้า : KBO00001523063 รหัสธุรกรรม : KPSO04KB000001523063	689
465	15 ม.ค. 2564	SCB	ปัณเทาธรีย์ Wan XXX-XXX440-7	สยาม สงวนรัมย์ XXX-XXX176-3	285
272	09 พ.ค. 2564	BAY	KITSADA JAIDE XXX-4-62360-X	สรวิชญ์ สุนพงษ์ไทย X-6232	100
273	09 พ.ค. 2564	BAY	KITSADA JAIDE XXX-4-62360-X	พลอยนวล ชนะบุญ XXX-XXX-6643	20
285	20 พ.ค. 2564	BAY	KITSADA JAIDE XXX-4-62360-X	พาณิภัคร คำนอง XXX-XXX-6847	500
287	21 พ.ค. 2564	BAY	KITSADA JAIDE XXX-4-62360-X	* เติมเงินพร้อมเพย์ 0040000037043 43	54

4.61 ข้อมูลหลังผ่าน Filter เวลา

4.8.4 ทดสอบ Advance Option

ในการทดสอบ Advance Option ทำการทดสอบโดยการพิมพ์ Keyword ที่สนใจ แสดงตัวอย่างของผลลัพธ์การ Filter ดังรูปที่ 4.62

	date	bank	sender	recipient	amount
6	08 ธ.ค. 2564	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	ปัณเทาธีย์ wan 1-6099-0055X-XX-X	400
21	11 ธ.ค. 2564	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	ปัณเทิดา โสดาราม XXX-XXX-9288	50
80	02 พ.ย. 2563	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	ปัณเทาธีย์ Wan X-2751	3500
86	25 ธ.ค. 2564	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	ปัณเทาธีย์ wan 1-6099-0055X-XX-X	100
87	26 ธ.ค. 2564	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	ปัณเทาธีย์ wan 1-6099-0055X-XX-X	100
95	29 ธ.ค. 2564	BBL	ศรัณญ์ 386-4-XXX278 ธนาคารกรุงเทพ	ปัณเทาธีย์ wan 1-6099-0055X-XX-X	500
104	03 ม.ค. 2565	KTB	นลภัญญดา กำลังเจริญ กรุงไทย ร XXX-X-XX568-6	ปัณเทาธีย์ wan 1-6099-0055X-XX-X	150

4.62 ข้อมูลหลังผ่าน Advance Option

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5

สรุปผลและข้อเสนอแนะ

5.1 สรุปผล

ปริญญานิพนธ์นี้นำเสนอการศึกษา และใช้งานการเรียนรู้เชิงลึกร่วมกับเทคโนโลยีการสกัดข้อมูลอย่าง OCR เข้ามาแก้ไขปัญหาของสังคมที่เกิดขึ้น โดยปริญญานิพนธ์ฉบับนี้มีกระบวนการจำแนกรูปภาพสลิปการเงินออกจากรูปภาพชนิดอื่น แบ่งแยกประเภทสลิปตามธนาคาร และแบ่งประเภทของสลิปการเงิน เพื่อกำหนดรูปแบบการเก็บข้อมูลให้มีความรวดเร็วมากขึ้นในขั้นตอน OCR เพื่อนำข้อมูลที่ได้ไปทำการวิเคราะห์ และแสดงผลในรูปแบบ Data Visualization ผ่านเว็บแอปพลิเคชัน

จากการทดสอบเพื่อหาขนาดที่เหมาะสมการวิธีการ Grid search พบว่าให้ขนาดอินพุตของโมเดลมีขนาด 28x28 มีประสิทธิภาพที่ดีที่สุดโดยวัดผลจาก Confusion Matrix และประสิทธิภาพของโมเดลที่ใช้ในการจำแนกรูปภาพสลิปการเงินมีคะแนน Accuracy = 0.9924, Loss = 0.0274 ต่อมาทดสอบประสิทธิภาพของโมเดลที่ใช้ในการแบ่งชนิดสลิปการเงินมีคะแนน Accuracy = 0.9767, Loss = 0.0857 จากนั้นทดสอบประสิทธิภาพของโมเดลที่ใช้แบ่งประเภทของสลิปการเงินตามธนาคารของสลิปการเงินแบบอิเล็กทรอนิกส์มีคะแนน Accuracy = 0.9624, Loss = 0.1252 และทดสอบประสิทธิภาพของโมเดลที่ใช้แบ่งประเภทของสลิปการเงินตามธนาคารของสลิปการเงินแบบกระดาษมีคะแนน Accuracy = 0.8567, Loss = 0.7110 และ Web Application สามารถนำเข้ารูปภาพ และมีผลลัพธ์คือ Data Visualization ของข้อมูลสลิปการเงิน

5.2 ข้อเสนอแนะ

ประสิทธิภาพของ CNN Model สามารถปรับปรุงประสิทธิภาพให้ดียิ่งขึ้น โดยการเพิ่มชุดข้อมูลของรูปภาพสลิปการเงินและความหลากหลายของชุดข้อมูล และในการใช้ OCR จาก Tesseract OCR ใน Python มีข้อจำกัดในการสกัดข้อมูลจากรูปภาพสลิปการเงิน ดังนั้นควรจะลองใช้ OCR ที่เป็น API ที่นอกเหนือจาก Library ของ Python เพื่อทดสอบประสิทธิภาพเพื่อให้ได้ผลที่ดีที่สุด เช่น Google Vision API, AWS Textract และ Microsoft Azure OCR เป็นต้น และในการประมวลผลถ้าใช้ Data base เข้าร่วมด้วยอาจจะช่วยเพิ่มประสิทธิภาพในการประมวลผล และดึงข้อมูลมาใช้เพื่อเพิ่มความรวดเร็ว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บรรณานุกรม

[1] Deposit Protection Agency. “รายชื่อสถาบันการเงินที่ได้รับการคุ้มครอง.” https://www.dpa.or.th/articles/view/list-of-insured-financial-institutions.

[2] ธนาคารแห่งประเทศไทย. “สรุปจำนวนสาขาและจุดให้บริการของธนาคารพาณิชย์.” https://www.bot.or.th/App/BTWS_STAT/statistics/BOTWEBSTAT.aspx?reportID=904&language=th.

[3] ศูนย์คุ้มครองผู้ใช้บริการทางการเงิน. “Specialized Financial Institutions.” https://www.1213.or.th/th/aboutfcc/finservices/Pages/SFIs.aspx.

[4] Docextractor. “Data Extraction from Financial Statements using OCR and IDP.” https://docextractor.com/data-extraction-from-financial-statements-using-ocr-and-idp/.

[5] CARDOAI. “Financial Data Extraction from Statements with Machine Learning.” https://cardoai.com/data-extraction-from-financial-statement-with-ml/?fbclid=IwAR2UbWcZBzletl8QzJVRFoy77h1da_ZToSr-g2Mzt2urJ5k79dWXh_FEFK8.

[6] Wikidocs. “GPT-2.” https://wikidocs.net/181226.

[7] Aigen. “AI.” https://aigencorp.com/intelligent-document-processing/.

[8] Ilija Mihajlovie. “Computer Vision.” https://towardsdatascience.com/everything-you-ever-wanted-to-know-about-computer-vision-heres-a-look-why-it-s-so-awesome-e8a58dfb641e.

[9] Revolveai. “NLP.” https://revolveai.com/natural-language-processing-techniques/.

[10] Javatpoint. “Fuzzy.” https://www.javatpoint.com/fuzzy-logic.

[11] Dataset.org.in. “Machine Learning.” https://dataset.org.in/MACHINE-LEARNING.html.

[12] Iconext. “OCR.” https://iconext.co.th/th/2021/10/11/.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

[13] Sasiwut Chaiyadecha. “Optical Character Recognition with TensorFlow and Keras.” https://lengyi.medium.com/part-1-optical-character-recognition-with-tensorflow-and-keras-ocr-94c271c0a0a3.

[14] Klippa. “IDP.” https://www.klippa.com/en/blog/information/idp-buyers-guide/.

[15] Divya Sheel. “Deep Learning.” https://new.abb.com/news/detail/58004.

[16] Natthawat Phongchit. “Convolutional Neural Network (CNN).” https://medium.com/@natthawatphongchit.

[17] Patompat Kamtar. “Image Recognition with Convolutional Neural Networks (CNN).” https://patompatkamtar.medium.com.

[18] Surapong Kanoktipsatharporn. “Convolutional Neural Network.” https://www.bualabs.co m/archives/2461.

[19] HD COE. “Convolutional Neural Network.” https://medium.com/@hadee2531earvesdrop.

[20] Austin G. Waltes. “Convolutional Neural Networks (CNN) to Classify Sentences.” https://a ustingwalters.com/convolutional-neural-networks-cnn-to-classify-sentences/.

[21] Surapong Kanoktipsatharporn. “ReLU Function.” https://www.bualabs.com/archives/1355.

[22] SuperDataScience Team. “Convolutional Neural Networks (CNN): Step 3 – Flattening.” https://www.superdatascience.com/blogs/convolutional-neural-networks-cnn-step-3-flattening.

[23] IndianTechWarror. “Fully Connected Layers in Convolutional Neural Networks.” https://indiantechwarrior.com/fully-connected-layers-in-convolutional-neural-networks/.

[24] Thomas Wood. “What is the Softmax Function.” https://deepai.org/machine-learning-glossary-and-terms/softmax-layer.

[25] Pakawat Nakwijit. “Optimizer.” https://medium.com/@chameleontk/.

[26] A P. “Gradient.” https://medium.com/swlh/strengths-and-weaknesses-of-optimization-algorithms-used-for-machine-learning-58926b1d69dd

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

[27] Kiprono Elijah Koech. “Cross-Entropy Loss Function.” https://towardsdatascience.com/cross-entropy-loss-function-f38c4ec8643e.

[28] Kizito Nyuytiymbit. “Hyperparameter.” https://towardsdatascience.com/.

[29] ScienceDirect. “Confusion Matrix.” https://www.sciencedirect.com/topics/engineering/confusion-matrix.

[30] Joydwip Mohajon. “Confusion Matrix.” https://towardsdatascience.com/confusion-matrix-for-your-multi-class-machine-learning-model-ff9aa3bf7826.

[31] Frank Kruger. “Confusion matrix.” https://www.researchgate.net/figure/.

[32] Mindphp. “Optimizer.” https://www.mindphp.com/.

[33] Vitaly Bushaev. “Adam.” https://towardsdatascience.com/adam-latest-trends-in-deep-learning-optimization-6be9a291375c.

[34] iop. “RGB.” https://www.iop.co.th/content/6222/.

[35] Had to know. “RGB to Grayscale.” https://www.had2know.org/technology/rgb-to-gray-scale-converter.

[36] Sarayut Nonsiri. “Python.” https://www.9experttraining.com/articles/.

[37] Microsoft. “VS Code.” https://learn.microsoft.com/th-th/power-apps/maker/portals/vs-code-extension.

[38] Sirirat Kantanat. “Jupyter Notebook.” https://blog.ragnar.co.th/blog/what-is-jupyter-notebook.

[39] Worrapob Parboonpasent. “UI.” https://grappik.com/user-interface-design.

[40] Boostlabs. “Data Visualization.” https://boostlabs.com/10-types-of-data-visualization-tools/.

[41] Surapong Kanoktipsatharporn. “Training/Test Set Split.” https://www.bualabs.com/archiv es/532/.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ก

โปรแกรมสำหรับโมเดลการเรียนรู้เชิงลึก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
width = 28
trainpath = 'data_train_part2.2y/'
testpath = 'data_test_part2.2x/'
trainImg = [trainpath+f for f in listdir(trainpath)]
testImg = [testpath+f for f in listdir(testpath)]
def img2data(path):
    rawImgs = []
    labels = []
    h=0
    width = 28
    for imagePath in (path):
        img_train = list(paths.list_images(imagePath))
        for i in img_train:
            img= cv2.imread(img_train[h])
            img = cv2.cvtColor(img, cv2.COLOR_BGR2RGB)
            img = cv2.resize(img ,(width,width))
            rawImgs.append(img)
            h+=1
            l = imagePath.split('/')[1]
            if l == '1_slip_digi':
                labels.append([1,0])
            elif l == '2_slip_paper':
                labels.append([0,1])
        print(str(h)+"/"+str(len(img_train)))
        print(imagePath)
        h=0
    return rawImgs, labels
x_train = []
y_train = []
x_train, y_train = img2data(trainImg)
print(len(x_train))
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
x_test = []
y_test = []
x_test, y_test = img2data(testImg)
x_train = np.array(x_train)
y_train = np.array(y_train)
x_test = np.array(x_test)
y_test = np.array(y_test)
x_train = x_train.astype('float32')
x_test = x_test.astype('float32')
x_train = x_train/255
x_test= x_test/255
model=keras.Sequential()
model.add(keras.layers.Conv2D(filters=32,kernel_size = 3,strides =
(1,1),padding = 'same', activation='relu',input_shape = (width,width,3)))
model.add(keras.layers.MaxPool2D(pool_size=(2,2), strides = (2,2),
padding = 'valid'))
model.add(keras.layers.Conv2D(filters=128,kernel_size = 3,strides =
(1,1),padding = 'same', activation='relu'))
model.add(keras.layers.MaxPool2D(pool_size=(2,2), strides = (2,2),
padding = 'valid'))
model.add(keras.layers.Conv2D(filters=256,kernel_size = 3,strides =
(1,1),padding = 'same', activation='relu'))
model.add(keras.layers.MaxPool2D(pool_size=(2,2), strides = (2,2),
padding = 'valid'))
model.add(keras.layers.Dropout(0.2))
model.add(keras.layers.Flatten())
model.add(keras.layers.Dense(units=128, activation='relu'))
model.add(keras.layers.Dropout(0.2))
model.add(keras.layers.Dense(2,activation='softmax'))
model.summary()
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
model.compile(optimizer=tf.keras.optimizers.Adam(learning_rate=0.00001
),loss="categorical_crossentropy",metrics=['accuracy'])
history=model.fit(x=x_train,y=y_train,epochs=70,batch_size=256,validatio
n_split=0.1)
model.save('./model/CNN_part3_x_model.h5')
```

```
width = 28
trainpath = 'data_train_part1x/'
testpath = 'data_test_part1x/'
trainImg = [trainpath+f for f in listdir(trainpath)]
testImg = [testpath+f for f in listdir(testpath)]
def img2data(path):
    rawImgs = []
    labels = []
    h=0
    width = 28
    for imagePath in (path):
        img_train = list(paths.list_images(imagePath))
        for i in img_train:
                img= cv2.imread(img_train[h])
                img = cv2.cvtColor(img, cv2.COLOR_BGR2RGB)
                img = cv2.resize(img ,(width,width))
                rawImgs.append(img)
                h+=1
                l = imagePath.split('/')[1]
                if l == '1_slip':
                    labels.append([1,0])
                elif l == '2_noise':
                    labels.append([0,1])
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
            print(str(h)+"/"+str(len(img_train)))
            print(imagePath)
            h=0
        return rawImgs, labels

x_train = []
y_train = []
x_train, y_train = img2data(trainImg)
print(len(x_train))
print(len(y_train))
x_test = []
y_test = []
x_test, y_test = img2data(testImg)
x_train = np.array(x_train)
y_train = np.array(y_train)
x_test = np.array(x_test)
y_test = np.array(y_test)
x_train = x_train.astype('float32')
x_test = x_test.astype('float32')
x_train = x_train/255
x_test= x_test/255
model=keras.Sequential()
model.add(keras.layers.Conv2D(filters=32,kernel_size = 3,strides =
(1,1),padding = 'same', activation='relu',input_shape = (width,width,3)))
model.add(keras.layers.MaxPool2D(pool_size=(2,2), strides = (2,2),
padding = 'valid'))
model.add(keras.layers.Conv2D(filters=128,kernel_size = 3,strides =
(1,1),padding = 'same', activation='relu'))
model.add(keras.layers.MaxPool2D(pool_size=(2,2), strides = (2,2),
padding = 'valid'))
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
model.add(keras.layers.Conv2D(filters=256,kernel_size = 3,strides =
(1,1),padding = 'same', activation='relu'))
model.add(keras.layers.MaxPool2D(pool_size=(2,2), strides = (2,2),
padding = 'valid'))
model.add(keras.layers.Dropout(0.2))
model.add(keras.layers.Flatten())
model.add(keras.layers.Dense(units=128, activation='relu'))
model.add(keras.layers.Dropout(0.2))
model.add(keras.layers.Dense(2,activation='softmax'))
model.summary()
model.compile(optimizer=tf.keras.optimizers.Adam(learning_rate=0.00001
),loss="categorical_crossentropy",metrics=['accuracy'])
history=model.fit(x=x_train,y=y_train,epochs=45,batch_size=128,validatio
n_split=0.2)
model.save('./model/CNN_part1_x_model.h5')

width = 28
trainpath = 'data_train_part2x/'
testpath = 'data_test_part2x/'
trainImg = [trainpath+f for f in listdir(trainpath)]
testImg = [testpath+f for f in listdir(testpath)]
trainImg,testImg
def img2data(path):
    rawImgs = []
    labels = []
    h=0
    width = 28
    for imagePath in (path):
        img_train = list(paths.list_images(imagePath)
        for i in img_train:
                img= cv2.imread(img_train[h])
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
                img = cv2.cvtColor(img, cv2.COLOR_BGR2RGB)
                img = cv2.resize(img ,(width,width))
                rawImgs.append(img)
                h+=1
                l = imagePath.split('/')[1]
                if l == '1_bangkok':
                    labels.append([1,0,0,0,0,0])
                elif l == '2_krungthai':
                    labels.append([0,1,0,0,0,0])
                elif l == '3_krungsri':
                    labels.append([0,0,1,0,0,0])
                elif l == '4_kasikorn':
                    labels.append([0,0,0,1,0,0])
                elif l == '5_ttb':
                    labels.append([0,0,0,0,1,0])
                elif l == '6_scb':
                    labels.append([0,0,0,0,0,1])
            print(imagePath)
            print(str(h)+"/"+str(len(img_train)))
            h=0
        return rawImgs, labels
x_train = []
y_train = []
x_train, y_train = img2data(trainImg)
x_test = []
y_test = []
x_test, y_test = img2data(testImg)
model=keras.Sequential()
model.add(keras.layers.Conv2D(filters=32,kernel_size = 3,strides =
(1,1),padding = 'same', activation='relu',input_shape = (width,width,3)))
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
model.add(keras.layers.MaxPool2D(pool_size=(2,2), strides = (2,2),
padding = 'valid'))
model.add(keras.layers.Dropout(0.2))
model.add(keras.layers.Conv2D(filters=128,kernel_size = 3,strides =
(1,1),padding = 'same', activation='relu'))
#model.add(keras.layers.Dropout(0.2))
model.add(keras.layers.MaxPool2D(pool_size=(2,2), strides = (2,2),
padding = 'valid'))
model.add(keras.layers.Dropout(0.2))
model.add(keras.layers.Conv2D(filters=256,kernel_size = 3,strides =
(1,1),padding = 'same', activation='relu'))
#model.add(keras.layers.Dropout(0.2))
model.add(keras.layers.MaxPool2D(pool_size=(2,2), strides = (2,2),
padding = 'valid'))
model.add(keras.layers.Dropout(0.2))
model.add(keras.layers.Conv2D(filters=512,kernel_size = 3,strides =
(1,1),padding = 'same', activation='relu'))
#model.add(keras.layers.Dropout(0.2))
model.add(keras.layers.MaxPool2D(pool_size=(2,2), strides = (2,2),
padding = 'valid'))
model.add(keras.layers.Dropout(0.2))
model.add(keras.layers.Flatten())
model.add(keras.layers.Dense(units=128, activation='relu'))
model.add(keras.layers.Dropout(0.2))
model.add(keras.layers.Dense(6,activation='softmax'))
model.summary()
model.compile(optimizer=tf.keras.optimizers.Adam(learning_rate=0.0001),
loss="categorical_crossentropy",metrics=['accuracy'])
history=model.fit(x=x_train,y=y_train,epochs=12,batch_size=64,validation
_split=0.1)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
model.save('./model/CNN_part2_last_model.h5')


width = 28
trainpath = 'data_train_part2z/'
testpath = 'data_test_part2z/'
trainImg = [trainpath+f for f in listdir(trainpath)]
testImg = [testpath+f for f in listdir(testpath)]
trainImg,testImg
def img2data(path):
    rawImgs = []
    labels = []
    h=0
    width = 28
    for imagePath in (path):
        img_train = list(paths.list_images(imagePath))
        for i in img_train:
            img= cv2.imread(img_train[h])
            img = cv2.cvtColor(img, cv2.COLOR_BGR2RGB)
            img = cv2.resize(img ,(width,width))
            rawImgs.append(img)
            h+=1
            l = imagePath.split('/')[1]
            if l == '1_bangkok':
                labels.append([1,0,0,0,0,0])
            elif l == '2_krungthai':
                labels.append([0,1,0,0,0,0])
            elif l == '3_krungsri':
                labels.append([0,0,1,0,0,0])
            elif l == '4_kasikorn':
                labels.append([0,0,0,1,0,0])
            elif l == '5_ttb':
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
                labels.append([0,0,0,0,1,0])
            elif l == '6_scb':
                labels.append([0,0,0,0,0,1])
        print(imagePath)
        print(str(h)+"/"+str(len(img_train)))
        h=0
    return rawImgs, labels
x_train = []
y_train = []
x_train, y_train = img2data(trainImg)
print(len(x_train))
print(len(y_train))
x_test = []
y_test = []
x_test, y_test = img2data(testImg)
print(len(x_test))
print(len(y_test))
x_train = np.array(x_train)
y_train = np.array(y_train)
x_test = np.array(x_test)
y_test = np.array(y_test)
x_train = x_train.astype('float32')
x_test = x_test.astype('float32')
x_train = x_train/255
x_test= x_test/255
model.add(keras.layers.Conv2D(filters=32,kernel_size = 3,strides =
(1,1),padding = 'same', activation='relu',input_shape = (width,width,3)))
model.add(keras.layers.MaxPool2D(pool_size=(2,2), strides = (2,2),
padding = 'valid'))
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
#model.add(keras.layers.Dropout(0.2))
model.add(keras.layers.Conv2D(filters=128,kernel_size = 3,strides =
(1,1),padding = 'same', activation='relu'))
#model.add(keras.layers.Dropout(0.2))
model.add(keras.layers.MaxPool2D(pool_size=(2,2), strides = (2,2),
padding = 'valid'))

#model.add(keras.layers.Dropout(0.2))
model.add(keras.layers.Conv2D(filters=256,kernel_size = 3,strides =
(1,1),padding = 'same', activation='relu'))
#model.add(keras.layers.Dropout(0.2))
model.add(keras.layers.MaxPool2D(pool_size=(2,2), strides = (2,2),
padding = 'valid'))

#model.add(keras.layers.Dropout(0.2))
model.add(keras.layers.Conv2D(filters=512,kernel_size = 3,strides =
(1,1),padding = 'same', activation='relu'))
#model.add(keras.layers.Dropout(0.2))
model.add(keras.layers.MaxPool2D(pool_size=(2,2), strides = (2,2),
padding = 'valid'))
#model.add(keras.layers.Dropout(0.2))
model.add(keras.layers.Flatten())
model.add(keras.layers.Dense(units=128, activation='relu'))
model.add(keras.layers.Dropout(0.2))
model.add(keras.layers.Dense(6,activation='softmax'))
model.summary()
model.compile(optimizer=tf.keras.optimizers.Adam(learning_rate=0.0001),
loss="categorical_crossentropy",metrics=['accuracy'])
history=model.fit(x=x_train,y=y_train,epochs=10,batch_size=32,validation
_split=0.1)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ข

โปรแกรมนำเข้า Model และ OCR

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
def run_ai():
    path = os.path.dirname(os.path.realpath(__file__))
    new_model_part1 = load_model(path+"\\CNN_part1_x_model.h5")
    new_model_part2 = load_model(path+"\\CNN_part2_last_model.h5")
    new_model_part3 = load_model(path+"\\CNN_part3_x_model.h5")
    new_model_part2_paper =
load_model(path+"\\CNN_part2_paper_model.h5")

    def delete_dot(z):
        c=""
        c=z
        count_dot = c.count('.')
        while count_dot  !=1 :
            inddateL = c.find(".")
            c = c[:inddateL] + c[inddateL+1:]
            count_dot = c.count('.')
        return c
    def cor_data_acc(z):
        c=""
        c=z
        c = c.replace("%", "X")
        c = c.replace("<", "X")
        c = c.replace("o", "X")
        c = c.replace("x", "X")
        c = c.replace("(@)", "")
        return c

    def cor_date(z):
        c=""
        c=z
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
    c = c.replace( "PM", "น.")
    c = c.replace( "AM", "น.")
    c = c.replace("ุ", "")
    c = c.replace( "Jan", "ม.ค.")
    c = c.replace( "..", ".")
    return c

def cor_year(z):
    c=z
    if c.find("66") != -1:
        c = c.replace( "66", "2566")
    elif c.find("65") != -1:
        c = c.replace( "65", "2565")
    elif c.find("64") != -1:
        c = c.replace( "64", "2564")
    elif c.find("63") != -1:
        c = c.replace( "63", "2563")
    elif c.find("62") != -1:
        c = c.replace( "62", "2562")
    elif c.find("61") != -1:
        c = c.replace( "61", "2561")
    elif c.find("60") != -1:
        c = c.replace( "60", "2560")
    elif c.find("59") != -1:
        c = c.replace( "59", "2559")
    elif c.find("58") != -1:
        c = c.replace( "58", "2558")
    elif c.find("57") != -1:
        c = c.replace( "57", "2557")
    elif c.find("56") != -1:
        c = c.replace( "56", "2556")
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
    elif c.find("55") != -1:
        c = c.replace( "55", "2555")
    elif c.find("54") != -1:
        c = c.replace( "54", "2554")
    elif c.find("53") != -1:
        c = c.replace( "53", "2553")
    else:
        c="-"
    return c
def cor_kbang(z):
    try:
        z = z.split(" ")
        c=z[2]
        if c=="66":
            c = c.replace( "66", "2566")
        elif c=="65":
            c = c.replace( "65", "2565")
        elif c=="64":
            c = c.replace( "64", "2564")
        elif c=="63":
            c = c.replace( "63", "2563")
        elif c=="62":
            c = c.replace( "62", "2562")
        elif c=="61":
            c = c.replace( "61", "2561")
        elif c=="60":
            c = c.replace( "60", "2560")
        elif c=="59":
            c = c.replace( "59", "2559")
        elif c=="58":
            c = c.replace( "58", "2558")
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
elif c=="57":
    c = c.replace( "57", "2557")
elif c=="56":
    c = c.replace( "56", "2556")
elif c=="55":
    c = c.replace( "55", "2555")
elif c=="54":
    c = c.replace( "54", "2554")
elif c=="53":
    c = c.replace( "53", "2553")
elif c=="23":
    c = c.replace( "23", "2566")
elif c=="22":
    c = c.replace( "22", "2565")
elif c=="21":
    c = c.replace( "21", "2564")
elif c=="20":
    c = c.replace( "20", "2563")
elif c=="19":
    c = c.replace( "19", "2562")
elif c=="18":
    c = c.replace( "18", "2561")
elif c=="17":
    c = c.replace( "17", "2560")
elif c=="16":
    c = c.replace( "16", "2559")
elif c=="15":
    c = c.replace( "15", "2558")
elif c=="14":
    c = c.replace( "14", "2557")
elif c=="13":
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
            c = c.replace( "13", "2556")
        elif c=="12":
            c = c.replace( "12", "2555")
        elif c=="11":
            c = c.replace( "11", "2554")
        elif c=="10":
            c = c.replace( "10", "2553")
        else:
            c="-"
        o = z[0]+" "+z[1]+" "+c
        return o
    except:
        return "-"

def add_dot(z):
    if z[-1]!=".":
        z = z+"."
    return


def cor_year_eng(z):
    c=z
    c = c.replace( "2023", "2566")
    c = c.replace( "2022", "2565")
    c = c.replace( "2021", "2564")
    c = c.replace( "2020", "2563")
    c = c.replace( "2019", "2562")
    c = c.replace( "2018", "2561")
    c = c.replace( "2017", "2560")
    c = c.replace( "2016", "2559")
    c = c.replace( "2015", "2558")
    c = c.replace( "2014", "2557")
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
    c = c.replace( "2013", "2556")
    c = c.replace( "2012", "2555")
    c = c.replace( "2011", "2554")
    c = c.replace( "2010", "2553")
    return c

def seperate(txt):
    pattern = r"(\d{1,2})(.*\D)(\d{2,4})"
    lit = re.findall(pattern, txt)[0]
    return lit

def edit_day(d):
    day = '0'+d
    return day[-2:]
def edit_mouth(m):
    format_mouth = ['ม.ค.', 'ก.พ.', 'มี.ค.', 'เม.ย.', 'พ.ค.','มิ.ย.', 'ก.ค.', 'ส.ค.',
'ก.ย.', 'ต.ค.', 'พ.ย.', 'ธ.ค.',]
    mouth = m.strip()
    if mouth in format_mouth:
        return mouth
    else:
        num = random.randint(0, 11)
        return format_mouth[num]

def detect_error(z):
    try:
        dect_1=0
        dect_2=0
        dect_0=0
        a=[]
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
        format_mouth = ['0','ม.ค.', 'ก.พ.', 'มี.ค.', 'เม.ย.', 'พ.ค.','มิ.ย.', 'ก.ค.',
'ส.ค.', 'ก.ย.', 'ต.ค.', 'พ.ย.', 'ธ.ค.',]
        a=z
        a = a.split(" ")
        c_0=a[0]
        c_1=a[1]
        c_2=a[2]

        if c_1 in format_mouth:
            dect_1=0
        else:
            dect_1=1

        if len(c_2) ==4:
            dect_2=0
        else:
            dect_2=1

        if int(c_0)<=31 and int(c_0)>0 :
            dect_0=0
        else:
            dect_0=1
        if dect_0==1 or dect_1==1 or dect_2==1:
            o="-"
        else:
            o = z
    except:
        o = "-"
    return o

def len_error(z):
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
        try:
            if len(z)==0:
                o="-"
            else:
                o=z
        except:
            o = "-"

        return o
    #digi examp
    def ex_bank_KTB(img_slip):
        custom_config = r'-l tha+eng --oem 3 --psm 6 -c
language_model_ngram_space_delimited_language=2'
        bank="KTB"
        n_date = 0
        n_date_2 = 0
        n_code = 0
        n_sen = 0
        n_re = 0
        n_amount = 0
        x_code= []
        x_amount= []
        x_sen =""
        x_re =""
        x_amount =""

        T_date = ""
        T_code = ""
        T_sen = ""
        T_re = ""
        T_amount =0
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
img= img_slip
img = cv2.cvtColor(img, cv2.COLOR_BGR2RGB)
data = pytesseract.image_to_string(img, lang='Tha+Eng')
x = data.split("\n")
while '' in x:
    x.remove('')

x = [s for s in x if len(s) >= 3]

for i, s in enumerate(x):
    if 'สำ' in s or 'เร็จ' in s:
        n_sen = i+2
    if  'วันที่ทำรายการ' in s or 'รายการ' in s or 'วันที่' in s:
        n_date = i
    if 'จำนวนเงิน' in s or 'จำนวน' in s or 'นวน' in s:
        n_amount = i
n_code=1
x_code= re.sub(r'[^\d.]', '', x[n_code])
while len(x_code)  < 10:
    n_code+=1
    x_code = re.sub(r'[^\d.]', '', x[n_code])
if n_sen ==0:
    n_sen = n_code+1

n_re = n_sen+3

x_sen = x[n_sen].split(" ")
while '' in x_sen:
    x_sen.remove('')
try:
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
    x[n_sen] = (x_sen[-2]+" "+x_sen[-1])
except:
    x[n_sen] = (x_sen[-1])



x_re = x[n_re].split(" ")
while '' in x_re:
    x_re.remove('')
try:
    x[n_re] = (x_re[-2]+" "+x_re[-1])
except:
    x[n_re] = (x_re[-1])

if n_date == 0:
    try:
        n_date=-1
        inddateL = x[n_date].find(":")
        n_dateL = x[n_date][:inddateL+3]
        while len(n_dateL)  < 10:
            n_date -=1
            inddateL = x[n_date].find(":")
            n_dateL = x[n_date][:inddateL+3]
    except:
        n_date=-1

inddateL = x[n_date].find(":")
x[n_date] = x[n_date][inddateL-20 :inddateL+3]
x[n_date] = cor_date(x[n_date])
inddateL =x[n_date].find("-")
x[n_date] = x[n_date][:inddateL-1]
x[n_date] =edit_all(x[n_date])
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
T_date = x[n_date]

N_sen = n_sen+1
while N_sen  < n_re:
    x[n_sen] = x[n_sen] +" "+x[N_sen]
    N_sen +=1

x[n_sen] = cor_data_acc(x[n_sen])
T_sen = x[n_sen]


N_re = n_re+1
while N_re  < n_amount:
    x[n_re] = x[n_re] +" "+x[N_re]
    N_re +=1

x[n_re] = cor_data_acc(x[n_re])
T_re =x[n_re]

try:
    x[n_amount] = re.sub(r'[^\d.]', '', x[n_amount])
    while len(x[n_amount])  < 1:
        n_amount+=1
        x[n_amount] = re.sub(r'[^\d.]', '', x[n_amount])

    x[n_amount] = re.sub(r'[^\d.]', '', x[n_amount])
    count_dot =  x[n_amount].count('.')
    if count_dot > 1:
        x[n_amount] = delete_dot(x[n_amount])
    if count_dot ==0:
        x_not_dot = x[n_amount]
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
            x[n_amount] = x_not_dot[:-2]

        T_amount = int(float(x[n_amount]))
    except:
        T_amount = 0

    if n_code == 0:
        T_code ="-"
    else:
        x[n_code] = re.sub(r'[^\d.]', '', x[n_code])
        T_code = x[n_code]

    return     T_date,bank,T_sen,T_re,T_amount,T_code

#paper exam
def ex_bank_BAY_paper(img_slip):
    custom_config = r'-l tha+eng --oem 3 --psm 6 -c
language_model_ngram_space_delimited_language=2'
    bank_paper="BAY"
    n_date = 0
    n_date_2 = 0
    n_code = 0
    n_sen = 0
    n_re = 0
    n_amount = 0
    x_code= []
    x_amount= []
    x_sen =""
    x_re =""
    x_amount =""
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
T_date_paper = ""
T_code_paper = ""
T_sen_paper = ""
T_re_paper = ""
T_amount_paper =0

img= img_slip
img = cv2.cvtColor(img, cv2.COLOR_BGR2RGB)
#  data  =  pytesseract.image_to_string(img,config=custom_config,
lang='Tha+Eng')
data = pytesseract.image_to_string(img, lang='Tha+Eng')
x = data.split("\n")
while '' in x:
    x.remove('')
x = [s for s in x if len(s) >= 4]
for i, s in enumerate(x):
    if 'BAY' in s:
        n_date = i
    if 'TRANSFER'  in s:
        n_re = i+1
n_amount = n_re+1

try:
    x_re = x[n_re].split(" ")
    x[n_re] = (x_re[-1])
    x[n_sen] = (x_re[0])
except:
    x[n_re] = "-"
    x[n_sen] = "-"
try:
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
            x[n_date] = cor_date(x[n_date])
            ind = x[n_date].find(":")
            x[n_date] = x[n_date][:ind-2]
            x[n_date] =cor_date_paper(x[n_date])
            x[n_date] =edit_all(x[n_date])
            x[n_date]=cor_kbang(x[n_date])
            T_date_paper = x[n_date]
        except:
            T_date_paper ='-'

        T_sen_paper = x[n_sen]

        T_re_paper =x[n_re]

        try:
            x_amount= x[n_amount].split(" ")
            x[n_amount] = x_amount[0]
            x[n_amount] = re.sub(r'[^\d.]', '', x[n_amount])
            while len(x[n_amount])  < 1:
                x[n_amount] = re.sub(r'[^\d.]', '', x[n_amount])
                n_amount+=1
            x[n_amount] = re.sub(r'[^\d.]', '', x[n_amount])
            T_amount_paper = int(float(x[n_amount]))
        except:
            T_amount_paper =0
        return
T_date_paper,bank_paper,T_sen_paper,T_re_paper,T_amount_paper,T_c
ode_paper
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
#slip/noise
noise =[]
slip =[]
x_pre =  []
n_pre = 0
no_noise = 1
no_slip = 1
width = 28
path = os.path.dirname(os.path.realpath(__file__))
img_pre= list(paths.list_images(path+"\\save"))
while n_pre  < len(img_pre):
    img= cv2.imread(img_pre[n_pre])
    show_img_pre = cv2.cvtColor(img , cv2.COLOR_BGR2RGB)
    img_re = cv2.resize(show_img_pre ,(width,width))
    x_pre = np.array(img_re)
    x_pre = x_pre.astype('float32')
    x_pre = x_pre/255
    x_pre = np.reshape(x_pre ,(1,width,width,3))
    predict = new_model_part1.predict(x_pre)
    label = ['slip','noise']
    result = label[np.argmax(predict)]
    print('number:'+str(n_pre+1))
    print(predict)
    print('predict:'+str(result))
    if result == 'noise':
        noise.append(img_pre[n_pre])
        #cv2.imwrite('./predict_noise/noise_'+str(no_noise)+'.jpg', img)
        no_noise += 1
    if result == 'slip':
        slip.append(img_pre[n_pre])
        #cv2.imwrite('./predict_slip/slip_'+str(no_slip)+'.jpg', img)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
            no_slip += 1
        n_pre =n_pre + 1


    #digi/paper
    slip_digi =[]
    slip_paper =[]
    x_pre =  []
    n_pre = 0
    no_slip_digi = 1
    no_slip_pape = 1
    width = 28
    while n_pre  < len(slip):
        img= cv2.imread(slip[n_pre])
        show_img_pre = cv2.cvtColor(img , cv2.COLOR_BGR2RGB)
        img_re = cv2.resize(show_img_pre ,(width,width))
        x_pre = np.array(img_re)
        x_pre = x_pre.astype('float32')
        x_pre = x_pre/255
        x_pre = np.reshape(x_pre ,(1,width,width,3))
        predict = new_model_part3.predict(x_pre)
        label = ['slip_digi','slip_paper']
        result = label[np.argmax(predict)]
        print('number:'+str(n_pre+1))
        print(predict)
        print('predict:'+str(result))
        if result == 'slip_digi':
            slip_digi.append(slip[n_pre])
            #cv2.imwrite('./table/slip_digi_'+str(no_slip_digi)+'.jpg', img)
            no_slip_digi += 1
        if result == 'slip_paper':
            slip_paper.append(slip[n_pre])
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
        no_slip_pape += 1
    n_pre =n_pre + 1

#6 bank digi
slip_bangkok =[]
slip_krungthai =[]
slip_krungsri =[]
slip_kasikorn =[]
slip_ttb =[]
slip_scb =[]

lable_slip_digi=[]

x_pre =  []
n_pre = 0
no_noise = 1
no_slip = 1
width = 28
while n_pre  < len(slip_digi):
    img= cv2.imread(slip_digi[n_pre])
    show_img_pre = cv2.cvtColor(img , cv2.COLOR_BGR2RGB)
    img_re = cv2.resize(show_img_pre ,(width,width))
    x_pre = np.array(img_re)
    x_pre = x_pre.astype('float32')
    x_pre = x_pre/255
    x_pre = np.reshape(x_pre ,(1,width,width,3))
    predict = new_model_part2.predict(x_pre)
    label = ['bangkok','krungthai','krungsri','kasikorn','ttb','scb']
    result = label[np.argmax(predict)]
    print('number:'+str(n_pre+1))
    print(predict)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
        print('predict:'+str(result))

        if result == 'bangkok':
            lable_slip_digi.append(result)
            slip_bangkok.append(slip_digi[n_pre])
        if result == 'krungthai':
            lable_slip_digi.append(result)
            slip_krungthai.append(slip_digi[n_pre])
        if result == 'krungsri':
            lable_slip_digi.append(result)
            slip_krungsri.append(slip_digi[n_pre])
        if result == 'kasikorn':
            lable_slip_digi.append(result)
            slip_kasikorn.append(slip_digi[n_pre])
        if result == 'ttb':
            lable_slip_digi.append(result)
            slip_ttb.append(slip_digi[n_pre])
        if result == 'scb':
            lable_slip_digi.append(result)
            slip_scb.append(slip_digi[n_pre])

        n_pre =n_pre + 1

#6 bank paper
slip_bangkok_paper =[]
slip_krungthai_paper =[]
slip_krungsri_paper =[]
slip_kasikorn_paper =[]
slip_ttb_paper =[]
slip_scb_paper =[]
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
lable_slip_paper=[]

x_pre =  []
n_pre = 0
no_noise = 1
no_slip = 1
width = 28
while n_pre  < len(slip_paper):
    img= cv2.imread(slip_paper[n_pre])
    show_img_pre = cv2.cvtColor(img , cv2.COLOR_BGR2RGB)
    img_re = cv2.resize(show_img_pre ,(width,width))
    x_pre = np.array(img_re)
    x_pre = x_pre.astype('float32')
    x_pre = x_pre/255
    x_pre = np.reshape(x_pre ,(1,width,width,3))
    predict = new_model_part2_paper.predict(x_pre)
    label = ['bangkok','krungthai','krungsri','kasikorn','ttb','scb']
    result = label[np.argmax(predict)]
    print('number:'+str(n_pre+1))
    print(predict)
    print('predict:'+str(result))

    if result == 'bangkok':
        lable_slip_paper.append(result)
        slip_bangkok_paper.append(slip_paper[n_pre])

    if result == 'krungthai':
        lable_slip_paper.append(result)
        slip_krungthai_paper.append(slip_paper[n_pre])

    if result == 'krungsri':
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
            lable_slip_paper.append(result)
            slip_krungsri_paper.append(slip_paper[n_pre])

        if result == 'kasikorn':
            lable_slip_paper.append(result)
            slip_kasikorn_paper.append(slip_paper[n_pre])

        if result == 'ttb':
            lable_slip_paper.append(result)
            slip_ttb_paper.append(slip_paper[n_pre])

        if result == 'scb':
            lable_slip_paper.append(result)
            slip_scb_paper.append(slip_paper[n_pre])

        n_pre =n_pre + 1

    #all pedict
    date = []
    sender = []
    recipient = []
    amount =[]
    referral_code = []
    bank = []

    date_digi = []
    sender_digi = []
    recipient_digi = []
    amount_digi =[]
    referral_code_digi = []
    bank_digi = []
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
    date_paper = []
    sender_paper = []
    recipient_paper = []
    amount_paper =[]
    referral_code_paper = []
    bank_paper = []

    slip_error = []
    slip_error_name = []

    TE_date = ""
    TE_sen = ""
    TE_re = ""
    TE_amount =0
    TE_code = ""

    TE_date_paper = ""
    TE_sen_paper = ""
    TE_re_paper = ""
    TE_amount_paper =0
    TE_code_paper = ""
    tn_digi = range(len(slip_digi))
    tn_paper = range(len(slip_paper))
    #tn = range(10)
    for n in tqdm(tn_digi):
        try:
            img= cv2.imread(slip_digi[n])
            if lable_slip_digi[n] =="bangkok":
                TE_date,TE_bank,TE_sen,TE_re,TE_amount,TE_code            =
ex_bank_BBL(img)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
            if lable_slip_digi[n] =="krungthai":
                TE_date,TE_bank,TE_sen,TE_re,TE_amount,TE_code          =
ex_bank_KTB(img)
            if lable_slip_digi[n] =="krungsri":
                TE_date,TE_bank,TE_sen,TE_re,TE_amount,TE_code          =
ex_bank_BAY(img)
            if lable_slip_digi[n] =="kasikorn":
                TE_date,TE_bank,TE_sen,TE_re,TE_amount,TE_code          =
ex_bank_KBANK(img)
            if lable_slip_digi[n] =="ttb":
                TE_date,TE_bank,TE_sen,TE_re,TE_amount,TE_code          =
ex_bank_TTB(img)
            if lable_slip_digi[n] =="scb":
                TE_date,TE_bank,TE_sen,TE_re,TE_amount,TE_code          =
ex_bank_SCB(img)

            TE_date= detect_error(TE_date)
        except:
            TE_date = "-"
            TE_bank = "-"
            TE_sen = "-"
            TE_re = "-"
            TE_amount = 0

        if TE_date !="-" and TE_amount !=0 and TE_amount <1000000:
            date_digi.append(TE_date)
            bank_digi.append(TE_bank)
            sender_digi.append(TE_sen)
            recipient_digi.append(TE_re)
            amount_digi.append(TE_amount)
            referral_code_digi.append(TE_code)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
        else:
            slip_error.append(slip_digi[n])

    for n in tqdm(tn_paper):
        try:
            img_paper= cv2.imread(slip_paper[n])
            if lable_slip_paper[n] =="bangkok":

TE_date_paper,TE_bank_paper,TE_sen_paper,TE_re_paper,TE_amount_p
aper,TE_code_paper = ex_bank_BBL_paper(img_paper)
            if lable_slip_paper[n] =="krungthai":

TE_date_paper,TE_bank_paper,TE_sen_paper,TE_re_paper,TE_amount_p
aper,TE_code_paper = ex_bank_KTB_paper(img_paper)
            if lable_slip_paper[n] =="krungsri":

TE_date_paper,TE_bank_paper,TE_sen_paper,TE_re_paper,TE_amount_p
aper,TE_code_paper = ex_bank_BAY_paper(img_paper)
            if lable_slip_paper[n] =="kasikorn":

TE_date_paper,TE_bank_paper,TE_sen_paper,TE_re_paper,TE_amount_p
aper,TE_code_paper = ex_bank_KBANK_paper(img_paper)
            if lable_slip_paper[n] =="ttb":

TE_date_paper,TE_bank_paper,TE_sen_paper,TE_re_paper,TE_amount_p
aper,TE_code_paper = ex_bank_TTB_paper(img_paper)
            if lable_slip_paper[n] =="scb":

TE_date_paper,TE_bank_paper,TE_sen_paper,TE_re_paper,TE_amount_p
aper,TE_code_paper = ex_bank_SCB_paper(img_paper)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
            TE_date_paper= detect_error(TE_date_paper)
        except:
            TE_date_paper = "-"
            TE_bank_paper = "-"
            TE_sen_paper = "-"
            TE_re_paper = "-"
            TE_amount_paper = 0

        if TE_date_paper !="-" and TE_amount_paper !=0 and
TE_amount_paper <1000000:
            date_paper.append(TE_date_paper)
            bank_paper.append(TE_bank_paper)
            sender_paper.append(TE_sen_paper)
            recipient_paper.append(TE_re_paper)
            amount_paper.append(TE_amount_paper)
            referral_code_paper.append(TE_code_paper)
        else:
            slip_error.append(slip_paper[n])


    date = date_digi+date_paper
    sender = sender_digi+sender_paper
    recipient = recipient_digi+recipient_paper
    amount =amount_digi+amount_paper
    referral_code = referral_code_digi+referral_code_paper
    bank = bank_digi+bank_paper


    slip_error_name = len_error(slip_error)
    print(slip_error_name)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
    import pandas as pd
    df=pd.DataFrame({'date' : date, 'bank' : bank, 'sender' : sender,
'recipient' : recipient, 'amount' : amount})

    for i in range(len(img_pre)):
        os.remove(img_pre[i])

    def parse_csv(df):
        res = df.to_json(orient="records")
        parsed = json.loads(res)
        return parsed
    return df
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ค

โปรแกรมหน้าแสดงผล (Web Application)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
# Button
path = os.path.dirname(os.path.realpath(__file__))
data = {'recipient':'Initial', 'bank':[None], 'amount':[0], 'date':['1 ม.ค. 2566']}
df = pd.DataFrame(data)
output_path = os.path.join(path, 'Data_set.xlsx')
df.to_excel(output_path, index=False)
with st.container():
    button_state = st.checkbox(label="Get Start")
    if button_state:
        path = os.path.dirname(os.path.realpath(__file__))
        if uploaded_files is not None :
            for uploaded_file in uploaded_files:
                bytes_data = io.BytesIO(uploaded_file.read())
                image = Image.open(bytes_data)
                if image.mode == 'RGBA':
                    image = image.convert('RGB')
                path = os.path.dirname(os.path.realpath(__file__))
                path = os.path.join(path, "save")
                path = os.path.join(path, uploaded_file.name)
                image.save(path, format='JPEG')
            data = run_ai()
            print(data)
            df = pd.DataFrame(data)
            path = os.path.dirname(os.path.realpath(__file__))
            output_path = os.path.join(path, 'Data_set.xlsx')
            df.to_excel(output_path, index=False)
        uploaded_files = None
path = os.path.dirname(os.path.realpath(__file__))
Data_set = os.path.join(path, 'Data_set.xlsx')
Data_set = pd.read_excel(str(Data_set))
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
Data_set = pd.DataFrame(Data_set)
if int(len(Data_set['bank']))>0:
    data = Data_set
    df = pd.DataFrame(data)
else:
    data = {'recipient':'Initial', 'bank':[None], 'amount':[0], 'date':['1 ม.ค. 2566']}
    df = pd.DataFrame(data)


# Main menu
with st.container():
    add_selectbox = st.multiselect("Choose a bank ",('KTB', 'BBL', 'Kbank', 'SCB', 'BAY', 'TTB'
                                                     , 'UOBT', 'LH Bank', 'KKP', 'CIMBT' , 'TISCO'
                                                     , 'ICBCT', 'SCBT', 'BAAC', 'GHB', 'GSB', 'ISBT'))
    days = [] ; months = [] ; years = []
    for i in range(len(df)):
        month_names = {
                        'ม.ค.': 1,
                        'ก.พ.': 2,
                        'มี.ค.': 3,
                        'เม.ย.': 4,
                        'พ.ค.': 5,
                        'มิ.ย.': 6,
                        'ก.ค.': 7,
                        'ส.ค.': 8,
                        'ก.ย.': 9,
                        'ต.ค.': 10,
                        'พ.ย.': 11,
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
            'ธ.ค.': 12
        }
    date_string = str(df['date'][i])
    day_str, month_str, year_str = date_string.split()
    month_num = month_names[month_str]
    day = int(day_str) ; days.append(day)
    month = int(month_num) ; months.append(month)
    year = int(year_str) ; years.append(year)

nowtime = str(datetime.datetime.now())[:10]
day_start = int(nowtime[8:])
month_start = int(nowtime[5:7])
year_start = int(nowtime[0:4])

for i in range(len(df)):
    if year_start == years[i]-543 :
        year_start = year_start
        if month_start == months[i]:
            month_start = months[i]
            if day_start == days[i]:
                day_start = days[i]
            elif day_start < days[i]:
                day_start = day_start
            elif day_start > days[i]:
                day_start = days[i]
        elif month_start < months[i] :
            month_start = month_start
            if day_start == days[i]:
                day_start = days[i]
            elif day_start < days[i]:
                day_start = day_start
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
            elif day_start > days[i]:
                day_start = days[i]
        elif month_start > months[i] :
            month_start = months[i]
            if day_start == days[i]:
                day_start = days[i]
            elif day_start < days[i]:
                day_start = day_start
            elif day_start > days[i]:
                day_start = days[i]

    elif year_start < years[i]-543 :
        year_start = year_start
        if month_start == months[i]:
            month_start = months[i]
            if day_start == days[i]:
                day_start = days[i]
            elif day_start < days[i]:
                day_start = day_start
            elif day_start > days[i]:
                day_start = days[i]
        elif month_start < months[i] :
            month_start = month_start
            if day_start == days[i]:
                day_start = days[i]
            elif day_start < days[i]:
                day_start = day_start
            elif day_start > days[i]:
                day_start = days[i]
        elif month_start > months[i] :
            month_start = months[i]
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
            if day_start == days[i]:
                day_start = days[i]
            elif day_start < days[i]:
                day_start = day_start
            elif day_start > days[i]:
                day_start = days[i]

        elif year_start > years[i]-543 :
            year_start = years[i]-543
            if month_start == months[i]:
                month_start = months[i]
                if day_start == days[i]:
                    day_start = days[i]
                elif day_start < days[i]:
                    day_start = day_start
                elif day_start > days[i]:
                    day_start = days[i]
            elif month_start < months[i] :
                month_start = month_start
                if day_start == days[i]:
                    day_start = days[i]
                elif day_start < days[i]:
                    day_start = day_start
                elif day_start > days[i]:
                    day_start = days[i]
            elif month_start > months[i] :
                month_start = months[i]
                if day_start == days[i]:
                    day_start = days[i]
                elif day_start < days[i]:
                    day_start = day_start
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
            elif day_start > days[i]:
                day_start = days[i]

    min_date = datetime.date(year_start, month_start, day_start)
    max_date = datetime.datetime.now().date()

    start_date, end_date = st.slider("Select a date range",
min_value=min_date, max_value=max_date, value=(min_date,
max_date))

    year_s = int(str(start_date)[:4])+543
    month_s = int(str(start_date)[5:7])
    day_s = int(str(start_date)[8:])
    start_time = datetime.date(year_s, month_s, day_s)
    year_e = int(str(end_date)[:4])+543
    month_e = int(str(end_date)[5:7])
    day_e = int(str(end_date)[8:])
    end_time = datetime.date(year_e, month_e , day_e)
    st.write(start_time, "-", end_time)
    time_se = np.arange(start_time, end_time,
dtype='datetime64[D]').astype('str')

    # Filter
    month_names = {
                'ม.ค.': 1,
                'ก.พ.': 2,
                'มี.ค.': 3,
                'เม.ย.': 4,
                'พ.ค.': 5,
                'มิ.ย.': 6,
                'ก.ค.': 7,
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
                'ส.ค.': 8,
                'ก.ย.': 9,
                'ต.ค.': 10,
                'พ.ย.': 11,
                'ธ.ค.': 12
            }
    date_time_se = []
    for i in range(len(df)) :
        input_date_str = str(df['date'][i])
        day_str, month_str, year_str = input_date_str.split()
        month_num = month_names[month_str]
        day = int(day_str)
        month = int(month_num)
        year = int(year_str)
        output_date_se  =  datetime.datetime(year=year,  month=month,
day=day)
        output_date_str_se = output_date_se.strftime("%Y-%m-%d")
        date_time_se.append(output_date_str_se)

    if data is not None:
        df = pd.DataFrame(data)
        t = []
        for i in range(len(time_se)):
            for j in range(len(date_time_se)):
                if date_time_se[j]==time_se[i]:
                    t.append(j)
        k = []
        for i in range(len(add_selectbox)):
            for j in range(len(df)):
                if df['bank'][j]==add_selectbox[i]:
                    k.append(j)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
        pointer = []
        for i in range(len(t)):
            for j in range(len(k)):
                if k[j]==t[i]:
                    pointer.append(k[j])

        if len(add_selectbox )>=1 and len(time_se)>=1 and df['recipient'][0]
!= 'Initial' :
            bank_df = df.iloc[pointer]
        elif len(add_selectbox)>=1 and len(time_se)==0 and
df['recipient'][0] != 'Initial' :
            bank_df = df.iloc[k]
        elif len(add_selectbox)==0 and len(time_se)>=1 and
df['recipient'][0] != 'Initial' :
            bank_df = df.iloc[t]
        elif len(add_selectbox)==0 and len(time_se)==0 and
df['recipient'][0] != 'Initial':
            bank_df = df
        elif df['recipient'][0]=='Initial':
            bank_df = pd.DataFrame(data='<Not found>', index=[0],
columns=['recipient'])

    # Run Modal
    outer_modal = Modal("Financial Slip Data", key="outer_modal")
    outer_button = st.button("Data Search")
    if outer_button:
        outer_modal.open()

    if outer_modal.is_open():
        with outer_modal.container():
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
            dfModal = bank_df
            col1, col2 = st.columns((1,1))
            with col1:
                name = st.text_input('Recipient Search')

            if data['recipient'][0] == 'Initial':
                dfModal = pd.DataFrame(data='<Not found>', index=[0],
columns=['recipient'])

            else :
                if len(name) >= 1 :
                    matches = dfModal[dfModal['recipient'].str.contains(name)]
                    pointer_name = []
                    if not matches.empty:
                        for i, row in matches.iterrows():
                            pointer_name.append(i)
                        dfModal = df.iloc[pointer_name]
                    else :
                        dfModal = pd.DataFrame(data='<Not found>', index=[0],
columns=['recipient'])
                else:
                    dfModal = dfModal
            with col2:
                image_name = st.text_input("View Image:")

            path = os.path.abspath(os.path.join(os.path.dirname(__file__),
"..", "static"))
            if image_name is None:
                filepath =
os.path.abspath(os.path.join(os.path.dirname(__file__), "..", "1.png"))
            else:
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
            path = os.path.abspath(os.path.join(os.path.dirname(__file__),
"..", "static"))
            filepath = os.path.join(path, image_name + ".jpg")
            try:
                with open(filepath, "rb") as f:
                    image_data = f.read()
                    image_base64                                    =
base64.b64encode(image_data).decode()
            except FileNotFoundError:
                filepath                                            =
os.path.abspath(os.path.join(os.path.dirname(__file__), "..", "1.png"))
                with open(filepath, "rb") as f:
                    image_data = f.read()
                    image_base64                                    =
base64.b64encode(image_data).decode()
        modal_image_code = """
        <div>
            <button type="button" class="btn btn-primary" data-
toggle="modal" data-target="#exampleModal1">Show Image</button>
            <div class="modal fade" id="exampleModal1" tabindex="-1"
role="dialog" aria-labelledby="exampleModalLabel1" aria-hidden="true">
                <div class="modal-dialog" role="document">
                    <div class="modal-content">
                        <div class="modal-header">
                            <h2                                 class="modal-title"
id="exampleModalLabel1">Image</h2>
                            <button type="button" class="close" data-
dismiss="modal" aria-label="Close">
                                <span aria-hidden="true">&times;</span>
                            </button>
                        </div>
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
                    <div class="modal-body">
                        <img  src="{}"  style="max-
width:100%;height:auto;">
                    </div></div></div>  </div>  </div>
        """.format(image_base64)
        st.markdown(modal_image_code, unsafe_allow_html=True)
        st.markdown("<br/>", unsafe_allow_html=True)
        st.dataframe(dfModal)

# Data Visualization with datetime
with st.container():
    st.markdown("<br/><br/><br/>", unsafe_allow_html=True)
    col1, col2 = st.columns((5,1))
    # Summary times and Amount
    with col2:
        option = st.selectbox('',('KTB', 'BBL', 'Kbank', 'SCB', 'BAY',
        'TTB', 'UOBT', 'LH_Bank', 'KKP', 'CIMBT', 'TISCO', 'ICBCT', 'SCBT',
        'BAAC', 'GHB', 'GSB', 'ISBT'))
        spend = 0
        for i in range(len(df)):
            spend = spend + df['amount'][i]
        formatted_spend = '{:,}'.format(spend)
        st.markdown("<br/>", unsafe_allow_html=True)
        st.markdown(f"<p                    style='font-size:25px'>Spend:</p>",
unsafe_allow_html=True)
        text_spend = '{:,}'.format(spend)+" Baht"
        st.markdown(f"<p              style='font-size:20px'>{text_spend}</p>",
unsafe_allow_html=True)
        st.markdown("<br/>", unsafe_allow_html=True)
        if df['amount'][0] == 0:
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
            st.markdown(f"<p                style='font-size:25px'>Times:</p>",
unsafe_allow_html=True)
            st.markdown(f"<p        style='font-size:20px'>0        Times</p>",
unsafe_allow_html=True)
        elif df['amount'][0] > 0:
            st.markdown(f"<p                style='font-size:25px'>Times:</p>",
unsafe_allow_html=True)
            text_times = '{:,}'.format(len(df))+" Times"
            st.markdown(f"<p          style='font-size:20px'>{text_times}</p>",
unsafe_allow_html=True)

    # Data Visualization with datetime
    with col1:
        month_names = {
            'ม.ค.': 1,
            'ก.พ.': 2,
            'มี.ค.': 3,
            'เม.ย.': 4,
            'พ.ค.': 5,
            'มิ.ย.': 6,
            'ก.ค.': 7,
            'ส.ค.': 8,
            'ก.ย.': 9,
            'ต.ค.': 10,
            'พ.ย.': 11,
            'ธ.ค.': 12
        }
        date_time = []
        for i in range(len(df)) :
            input_date_str = str(df['date'][i])
            day_str, month_str, year_str = input_date_str.split()
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
            month_num = month_names[month_str]
            day = int(day_str)
            month = int(month_num)
            year = int(year_str)
            output_date = datetime.datetime(year=year, month=month,
day=day)
            output_date_str = output_date.strftime("%Y-%m-%d")
            date_time.append(output_date_str)
        date_time.sort()

        new_values = np.arange(date_time[0], date_time[-1],
dtype='datetime64[D]').astype('str')
        my_date_time = np.unique(np.concatenate((date_time,
new_values)), axis=0)
        count_dict = Counter(date_time)
        x_values = np.array([*count_dict.items()])[:,0]
        y_values = map(int,np.array([*count_dict.items()])[:,1])
        y_val = np.array([*count_dict.items()])[:,1]

        data_times = []
        for i in range(len(my_date_time)):
            count = 0
            for j in range(len(x_values)):
                if my_date_time[i]==x_values[j]:
                    data_times.append(int(y_val[j]))
                    count = count+0
                elif count+1 != len(x_values):
                    count = count+1
                else :
                    data_times.append(0)
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
        if option == 'KTB':
            input_KTB = []
            for i in range(len(df)) :
                if str(df['bank'][i])=='KTB':
                    input =  str(df['date'][i])
                    day_str, month_str, year_str = input.split()
                    month_num = month_names[month_str]
                    day = int(day_str); month = int(month_num); year =
int(year_str)
                    output = datetime.datetime(year=year, month=month,
day=day)
                    output = output.strftime("%Y-%m-%d")
                    input_KTB.append(output)
            input_KTB.sort()
            data_KTB = []
            for i in range(len(my_date_time)):
              count = 0
              for j in range(len(input_KTB)):
                if my_date_time[i]==input_KTB[j]:
                    data_KTB.append(int(data_times[i]))
                    count = count+0
                elif count+1 != len(input_KTB):
                    count = count+1
                else :
                    data_KTB.append(0)
            date_position = data_KTB

        elif option == 'BBL':
            input_BBL = []
            for i in range(len(df)) :
                if str(df['bank'][i])=='BBL':
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
                input =  str(df['date'][i])
                day_str, month_str, year_str = input.split()
                month_num = month_names[month_str]
                day = int(day_str); month = int(month_num); year =
int(year_str)
                output = datetime.datetime(year=year, month=month,
day=day)
                output = output.strftime("%Y-%m-%d")
                input_BBL.append(output)
            input_BBL.sort()
            data_BBL = []
            for i in range(len(my_date_time)):
                count = 0
                for j in range(len(input_BBL)):
                if my_date_time[i]==input_BBL[j]:
                    data_BBL.append(int(data_times[i]))
                    count = count+0
                elif count+1 != len(input_BBL):
                    count = count+1
                else :
                    data_BBL.append(0)
            date_position = data_BBL

        elif option == 'Kbank':
            input_Kbank = []
            for i in range(len(df)) :
                if str(df['bank'][i])=='Kbank':
                    input =  str(df['date'][i])
                    day_str, month_str, year_str = input.split()
                    month_num = month_names[month_str]
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
                day = int(day_str); month = int(month_num); year =
int(year_str)
                output = datetime.datetime(year=year, month=month,
day=day)
                output = output.strftime("%Y-%m-%d")
                input_Kbank.append(output)
        input_Kbank.sort()
        data_Kbank = []
        for i in range(len(my_date_time)):
            count = 0
            for j in range(len(input_Kbank)):
                if my_date_time[i]==input_Kbank[j]:
                    data_Kbank.append(int(data_times[i]))
                    count = count+0
                elif count+1 != len(input_Kbank):
                    count = count+1
                else :
                    data_Kbank.append(0)
        date_position = data_Kbank

    fig = go.Figure(layout_height=400, layout_width=800)
    fig.add_trace(go.Scatter(x=my_date_time, y=date_position))
    fig.update_layout(
        xaxis=dict(
            rangeselector=dict(
                buttons=list([
                    dict(count=1,          label="1m",          step="month",
stepmode="backward"),
                    dict(count=6,          label="6m",          step="month",
stepmode="backward"),
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
                dict(count=1, label="YTD", step="year",
stepmode="todate"),
                dict(count=1, label="1y", step="year",
stepmode="backward"),
                dict(step="all")
            ])
        ),
        rangeslider=dict(visible=True),
        type="date"
    ),
    margin=dict(b=0)
)
st.plotly_chart(fig)

# Bank chart
with st.container():
    KTB = 0; BBL=0; Kbank=0; SCB=0; BAY=0; TTB=0;
    for i in range(len(df)):
        if data['bank'][i]=='KTB':
            KTB=KTB+1
        elif data['bank'][i]=='BBL':
            BBL=BBL+1
        elif data['bank'][i]=='Kbank':
            Kbank=Kbank+1
        elif data['bank'][i]=='SCB':
            SCB=SCB+1
        elif data['bank'][i]=='BAY':
            BAY=BAY+1
        elif data['bank'][i]=='TTB':
            TTB=TTB+1
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
    x_values = ['KTB', 'BBL', 'Kbank', 'SCB', 'BAY', 'TTB', 'UOBT', 'LH Bank',
'KKP', 'CIMBT' , 'TISCO' , 'ICBCT', 'SCBT', 'BAAC', 'GHB', 'GSB', 'ISBT']
    y_values = [int(KTB), int(BBL), int(Kbank), int(SCB), int(BAY), int(TTB),
int(UOBT), int(LH_Bank), int(KKP), int(CIMBT), int(TISCO), int(ICBCT),
int(SCBT), int(BAAC), int(GHB), int(GSB), int(ISBT)]
    bank = []; amount = []
    for i in range(len(df)):
        x = str(df['bank'][i])
        y = str(df['amount'][i])
        bank.append(x)
        amount.append(y)

    KTB_m = 0; BBL_m=0; Kbank_m=0; SCB_m=0; BAY_m=0; TTB_m=0;
UOBT_m=0;   LH_Bank_m=0;   KKP_m=0;   CIMBT_m=0;   TISCO_m=0;
ICBCT_m=0; SCBT_m=0; BAAC_m=0; GHB_m=0; GSB_m=0; ISBT_m=0

    for i in range(len(bank)):
        if bank[i]=='KTB' :
            KTB_m = KTB_m+int(amount[i])
        elif bank[i]=='BBL':
            BBL_m=BBL_m+int(amount[i])
        elif bank[i]=='Kbank':
            Kbank_m=Kbank_m+int(amount[i])
        elif bank[i]=='SCB':
            SCB_m = SCB_m+int(amount[i])
        elif bank[i]=='BAY':
            BAY_m = BAY_m+int(amount[i])
        elif bank[i]=='TTB':
            TTB_m = TTB_m+int(amount[i])
    sum_values = ['{:,}'.format(KTB_m)+" baht", '{:,}'.format(BBL_m)+" baht",
'{:,}'.format(Kbank_m)+" baht", '{:,}'.format(SCB_m)+" baht",
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
        '{:,}'.format(BAY_m)+" baht", '{:,}'.format(TTB_m)+" baht",
'{:,}'.format(UOBT_m)+" baht", '{:,}'.format(LH_Bank_m)+" baht",
        '{:,}'.format(KKP_m)+" baht", '{:,}'.format(CIMBT_m)+" baht",
'{:,}'.format(TISCO_m)+" baht", '{:,}'.format(ICBCT_m)+" baht",
        '{:,}'.format(SCBT_m)+" baht", '{:,}'.format(BAAC_m)+"
baht",'{:,}'.format(GHB_m)+" baht", '{:,}'.format(GSB_m)+" baht",
'{:,}'.format(ISBT_m)+" baht"]
    data = {'x': x_values,'Transaction': y_values,'label': sum_values}
    chart_data = pd.DataFrame(data, columns=['x', 'Transaction', 'label'])
    bars = alt.Chart(chart_data).mark_bar().encode(
        x=alt.X('x', title=None, sort=None),
        y='Transaction',
        color=alt.Color('Transaction', scale=alt.Scale(scheme='bluegreen')),
        tooltip=['label']
    )
    text = bars.mark_text(
        align='center',
        baseline='bottom',
        dy=-5,
    ).encode(
        text='label'
    )
    st.altair_chart(bars + text, use_container_width=True
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหาและต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้