

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ก
กิตติกรรมประกาศ.....	ข
สารบัญตาราง.....	ง
สารบัญภาพ.....	จ
บทที่ 1 บทนำ.....	12
1.1 หลักการและเหตุผล.....	12
1.2 วัตถุประสงค์ของโครงการ.....	12
1.3 ขอบเขตของโครงการ.....	12
1.4 ภาพรวมของระบบ.....	13
1.5 ประโยชน์ที่คาดว่าจะได้รับ.....	14
1.6 อุปกรณ์และเครื่องมือที่ใช้ในการดำเนินงาน.....	15
1.7 แผนการดำเนินงาน.....	15
1.8 ตัวอย่างโปรแกรม และ/หรือ ฮาร์ดแวร์.....	16
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง.....	18
2.1 ทฤษฎีที่เกี่ยวข้อง.....	18
2.2 งานวิจัยที่เกี่ยวข้อง/ระบบงานที่เกี่ยวข้อง (Related Word).....	32
บทที่ 3 วิธีดำเนินงานวิจัย.....	34
3.1 กรอบการดำเนินงาน.....	34
3.2 การตรวจจับข้อความ (Text Detection).....	36
3.3 ขั้นตอนการเรียกใช้งาน Tesseract OCR.....	50
3.4 การรู้จำข้อความ (Text Recognition).....	51
3.5 การเรียนรู้เชิงลึก (Deep Learning).....	52
3.6 การวัดประสิทธิภาพ.....	62
3.7 ชุดข้อมูลที่ใช้ในการทดลอง.....	63
3.8 ผลการออกแบบและพัฒนาระบบ.....	65
บทที่ 4 ผลการทดลอง.....	68
4.1 ผลการทดลองขั้นตอนการ Pre-process.....	68
4.2 การตัดภาพส่วนที่มีข้อความ.....	76
4.3 การเปรียบเทียบประสิทธิภาพการรู้จำโดยการปรับค่า Optimizers.....	86
4.4 ผลลัพธ์ประสิทธิภาพการรู้จำจาก Model Optimizer Adam.....	91
4.5 เปรียบเทียบประสิทธิภาพการรู้จำใบเสร็จจาก Tesseract.....	105

สารบัญ (ต่อ)

	หน้า
4.6 ตัวอย่างการทำนายผลบน Mobile Application	111
4.7 ตัวอย่างการทำนายผลบน Website	111
4.8 สรุปและวิเคราะห์ผลการทดลอง	112
บทที่ 5 สรุปและอภิปรายผลการทดลอง	113
5.1 สรุปผล และอภิปรายผล	113
5.2 ปัญหา และอุปสรรคในการดำเนินงาน	113
5.3 ข้อเสนอแนะ	113
5.1 ปัญหา และอุปสรรคในการดำเนินงาน	133
5.2 ข้อเสนอแนะ	133
เอกสารอ้างอิง	113
ภาคผนวก	115
ภาคผนวก ก คู่มือการใช้งาน Mobile Application	117
ภาคผนวก ข คู่มือการใช้งาน Website Application	120
บทความวิจัย	123
โปสเตอร์โครงงาน	134
ประวัติผู้จัดทำโครงงาน	136

สารบัญตาราง

	หน้า
ตารางที่ 1.1 แผนการดำเนินงาน	16
ตารางที่ 3.1 ตารางการสร้าง One-Hot Encode.....	58
ตารางที่ 3.2 ตารางแสดง External Entity Description	65
ตารางที่ 3.3 ตารางแสดง Data Store Description.....	65
ตารางที่ 3.4 ตารางแสดง Data Structure Description.....	66
ตารางที่ 3.5 ตารางแสดงข้อมูลการทำนายผล (Predict).....	67
ตารางที่ 4.1 ตัวอย่างการทำงานของขั้นตอนการ Pre-process ภาพที่ 1	68
ตารางที่ 4.2 ตัวอย่างการทำงานของขั้นตอนการ Pre-process ภาพที่ 2	70
ตารางที่ 4.3 ตัวอย่างการทำงานของขั้นตอนการ Pre-process ภาพที่ 3	71
ตารางที่ 4.4 ตัวอย่างการทำงานของขั้นตอนการ Pre-process ภาพที่ 4	73
ตารางที่ 4.5 ตัวอย่างการทำงานของขั้นตอนการ Pre-process ภาพที่ 5	74
ตารางที่ 4.6 ตัวอย่างผลลัพธ์ของการตัดภาพส่วนที่มีข้อความ ภาพที่ 1	76
ตารางที่ 4.7 ตัวอย่างผลลัพธ์ของการตัดภาพส่วนที่มีข้อความ ภาพที่ 2	78
ตารางที่ 4.8 ตัวอย่างผลลัพธ์ของการตัดภาพส่วนที่มีข้อความ ภาพที่ 3	79
ตารางที่ 4.9 ตัวอย่างผลลัพธ์ของการตัดภาพส่วนที่มีข้อความ ภาพที่ 4	80
ตารางที่ 4.10 ตัวอย่างผลลัพธ์ของการตัดภาพส่วนที่มีข้อความ ภาพที่ 5	82
ตารางที่ 4.11 ตัวอย่างผลลัพธ์ของการตัดภาพส่วนที่มีข้อความ ภาพที่ 6	84
ตารางที่ 4.12 ตัวอย่างผลลัพธ์ของการตัดภาพส่วนที่มีข้อความ ภาพที่ 7	85
ตารางที่ 4.13 รูปแบบ Model ที่ทำการทดลอง ทั้ง 2 Model	86
ตารางที่ 4.14 ผลการประเมินประสิทธิภาพค่า CER ที่ได้จากการเรียนรู้ Optimizer ทั้ง 3 ประเภท	87
ตารางที่ 4.15 ตัวอย่างผลลัพธ์การทำนายผลจาก Optimizer ทั้ง 3 ประเภท ภาพที่ 1	88
ตารางที่ 4.16 ตัวอย่างผลลัพธ์การทำนายผลจาก Optimizer ทั้ง 3 ประเภท ภาพที่ 2	88
ตารางที่ 4.17 ตัวอย่างผลลัพธ์การทำนายผลจาก Optimizer ทั้ง 3 ประเภท ภาพที่ 3	88
ตารางที่ 4.18 ตัวอย่างผลลัพธ์การทำนายผลจาก Optimizer ทั้ง 3 ประเภท ภาพที่ 4	88
ตารางที่ 4.19 ตัวอย่างผลลัพธ์การทำนายผลจาก Optimizer ทั้ง 3 ประเภท ภาพที่ 5	89
ตารางที่ 4.20 ตัวอย่างผลลัพธ์การทำนายผลจาก Optimizer ทั้ง 3 ประเภท ภาพที่ 6	89
ตารางที่ 4.21 ตัวอย่างผลลัพธ์การทำนายผลจาก Optimizer ทั้ง 3 ประเภท ภาพที่ 7	89
ตารางที่ 4.22 ตัวอย่างผลลัพธ์การทำนายผลจาก Optimizer ทั้ง 3 ประเภท ภาพที่ 8	89
ตารางที่ 4.23 ตัวอย่างผลลัพธ์การทำนายผลจาก Optimizer ทั้ง 3 ประเภท ภาพที่ 9	90
ตารางที่ 4.24 ตัวอย่างผลลัพธ์การทำนายผลจาก Optimizer ทั้ง 3 ประเภท ภาพที่ 10	90
ตารางที่ 4.25 ตัวอย่างผลลัพธ์การทำนายผลจาก Optimizer ทั้ง 3 ประเภท ภาพที่ 11	90
ตารางที่ 4.26 ตัวอย่างผลลัพธ์การทำนายผลจาก Optimizer ทั้ง 3 ประเภท ภาพที่ 12	90
ตารางที่ 4.27 ตัวอย่างผลลัพธ์การทำนายผลจาก Optimizer ทั้ง 3 ประเภท ภาพที่ 13	91

สารบัญตาราง (ต่อ)

	หน้า
ตารางที่ 4.61 ตัวอย่างการทำนายผลลัพธ์ระหว่างวิธีที่ใช้พัฒนากับการใช้ Tesseract ภาพที่ 11.....	108
ตารางที่ 4.62 ตัวอย่างการทำนายผลลัพธ์ระหว่างวิธีที่ใช้พัฒนากับการใช้ Tesseract ภาพที่ 12.....	108
ตารางที่ 4.63 ตัวอย่างการทำนายผลลัพธ์ระหว่างวิธีที่ใช้พัฒนากับการใช้ Tesseract ภาพที่ 13.....	108
ตารางที่ 4.64 ตัวอย่างการทำนายผลลัพธ์ระหว่างวิธีที่ใช้พัฒนากับการใช้ Tesseract ภาพที่ 14.....	109
ตารางที่ 4.65 ตัวอย่างการทำนายผลลัพธ์ระหว่างวิธีที่ใช้พัฒนากับการใช้ Tesseract ภาพที่ 15.....	109
ตารางที่ 4.66 ตัวอย่างการทำนายผลลัพธ์ระหว่างวิธีที่ใช้พัฒนากับการใช้ Tesseract ภาพที่ 16.....	109
ตารางที่ 4.67 ตัวอย่างการทำนายผลลัพธ์ระหว่างวิธีที่ใช้พัฒนากับการใช้ Tesseract ภาพที่ 17.....	109
ตารางที่ 4.68 ตัวอย่างการทำนายผลลัพธ์ระหว่างวิธีที่ใช้พัฒนากับการใช้ Tesseract ภาพที่ 18.....	110
ตารางที่ 4.69 ตัวอย่างการทำนายผลลัพธ์ระหว่างวิธีที่ใช้พัฒนากับการใช้ Tesseract ภาพที่ 19.....	110
ตารางที่ 4.70 ตัวอย่างการทำนายผลลัพธ์ระหว่างวิธีที่ใช้พัฒนากับการใช้ Tesseract ภาพที่ 20.....	110

สารบัญญภาพประกอบ

	หน้า
ภาพประกอบที่ 1.1 ตัวอย่างไบเซิร์จ.....	13
ภาพประกอบที่ 1.2 ภาพรวมของระบบ.....	14
ภาพประกอบที่ 1.3 User Interface ในส่วนของ Application Mobile	16
ภาพประกอบที่ 1.4 User Interface ในส่วนของ Website	17
ภาพประกอบที่ 2.1 ตัวอย่างภาพก่อนและหลังการปรับปรุงคุณภาพด้วย Median Blur.....	19
ภาพประกอบที่ 2.2 ตัวอย่างการแปลงภาพสีเป็นภาพระดับเทา	20
ภาพประกอบที่ 2.3 แสดงจุดตัด (Threshold) จากฮิสโตแกรมของภาพด้วย Otsu Algorithm.....	21
ภาพประกอบที่ 2.4 ตัวอย่างการหาค่า Threshold แบบ Adaptive local thresholding	22
ภาพประกอบที่ 2.5 ตัวอย่าง Structuring Element.....	22
ภาพประกอบที่ 2.6 ตัวอย่างก่อนทำไดเลชัน (Dilation).....	23
ภาพประกอบที่ 2.7 ตัวอย่างหลังทำไดเลชัน (Dilation).....	23
ภาพประกอบที่ 2.8 ภาพตัวอย่างการทำไดเลชัน (Dilation).....	23
ภาพประกอบที่ 2.9 ตัวอย่างก่อนทำอีโรชัน (Erosion).....	24
ภาพประกอบที่ 2.10 ตัวอย่างหลังทำอีโรชัน (Erosion).....	24
ภาพประกอบที่ 2.11 ภาพตัวอย่างการทำอีโรชัน (Erosion).....	24
ภาพประกอบที่ 2.12 การ Contour รูปภาพ.....	25
ภาพประกอบที่ 2.13 ตัวอย่างก่อนและหลังการทำ Contour.....	25
ภาพประกอบที่ 2.14 กระบวนการทำงานของ CNN.....	27
ภาพประกอบที่ 2.15 ฝั่งคุณลักษณะ (Feature map).....	27
ภาพประกอบที่ 2.16 แสดง filter ที่ใช้ในการสร้าง feature maps ด้วย CNN.....	27
ภาพประกอบที่ 2.17 แสดงการทำงานของ Max Pooling	28
ภาพประกอบที่ 2.18 แสดงการทำงานของ Average pooling	28
ภาพประกอบที่ 2.19 แสดงการทำงานของ Fully connected layer	29
ภาพประกอบที่ 2.20 กระบวนการทำงานของ RNN.....	29
ภาพประกอบที่ 2.21 กระบวนการทำงานของ LSTM.....	30
ภาพประกอบที่ 2.22 Forget Gate LSTM.....	31
ภาพประกอบที่ 2.23 Input Gate LSTM.....	31
ภาพประกอบที่ 2.24 Output Gate LSTM.....	31
ภาพประกอบที่ 2.25 กระบวนการทำงานของ GRU	32
ภาพประกอบที่ 3.1 กรอบการดำเนินงานของระบบ	34
ภาพประกอบที่ 3.2 ขั้นตอนการตรวจจับข้อความ	35
ภาพประกอบที่ 3.3 ตัวอย่างภาพนำเข้า	36
ภาพประกอบที่ 3.4 ตัวอย่างข้อมูลภาพก่อนทำ Median ที่ตำแหน่ง (3,3).....	37
ภาพประกอบที่ 3.5 ตัวอย่างเรียงพิกเซลภาพจากค่าน้อยไปหาค่ามาก ที่ตำแหน่ง (3,3).....	37

สารบัญภาพประกอบ (ต่อ)

	หน้า
ภาพประกอบที่ 3.6 ตัวอย่างหลังทำ Median ที่ตำแหน่ง (3,3).....	37
ภาพประกอบที่ 3.7 ตัวอย่างข้อมูลภาพก่อนทำ Median ที่ตำแหน่ง (3,4).....	37
ภาพประกอบที่ 3.8 ตัวอย่างเรียงพิกเซลภาพจากค่าน้อยไปหาค่ามาก ที่ตำแหน่ง (3,4).....	38
ภาพประกอบที่ 3.9 ตัวอย่างหลังทำ Median ที่ตำแหน่ง (3,4).....	38
ภาพประกอบที่ 3.10 ภาพก่อนละหลังการทำ Median Blur.....	38
ภาพประกอบที่ 3.11 ตารางค่าสี R G B ตามลำดับ	39
ภาพประกอบที่ 3.12 การแปลงภาพสี RGB ไปเป็นภาพระดับเทา	39
ภาพประกอบที่ 3.13 ตัวอย่างหลังจากการแปลงเป็นภาพระดับเทา.....	39
ภาพประกอบที่ 3.14 ตัวอย่างภาพต้นฉบับก่อนการแปลงเป็นภาพระดับเทา.....	40
ภาพประกอบที่ 3.15 ตัวอย่างภาพหลังการแปลงเป็นภาพระดับเทา.....	40
ภาพประกอบที่ 3.16 ตัวอย่างการแปลงภาพระดับเทา เป็นภาพ Binary.....	42
ภาพประกอบที่ 3.17 ส่วนประกอบโครงสร้าง 4 ทิศทาง	43
ภาพประกอบที่ 3.18 ภาพต้นฉบับก่อนการทำ Dilation	43
ภาพประกอบที่ 3.19 ขั้นตอนการนำภาพต้นฉบับและส่วนประกอบโครงสร้างมายูเนียนกัน	43
ภาพประกอบที่ 3.20 ภาพหลังการยูเนียนกันของ ภาพประกอบที่ 3.19.....	44
ภาพประกอบที่ 3.21 ภาพผลลัพธ์การทำ Dilation	44
ภาพประกอบที่ 3.22 ตัวอย่างการสแกนรูปภาพจากซ้ายไปขวา.....	45
ภาพประกอบที่ 3.23 ตัวอย่างการ Contour ขั้นตอนที่ 2	45
ภาพประกอบที่ 3.24 ตัวอย่างการหาเส้นขอบตามเข็มนาฬิกา	46
ภาพประกอบที่ 3.25 ตัวอย่างการกำหนดค่าให้ (i_2, j_2) เท่ากับ (i_1, j_1) และให้ (i_3, j_3) เท่ากับ (i, j)	46
ภาพประกอบที่ 3.26 ตัวอย่างการหาเส้นขอบแบบทวนเข็มนาฬิกา	47
ภาพประกอบที่ 3.27 ตัวอย่างการเปลี่ยนค่าพิกเซลตามเงื่อนไขที่ 3	47
ภาพประกอบที่ 3.28 ภาพผลลัพธ์หลังการทำ Contour จนครบ.....	48
ภาพประกอบที่ 3.29 ตัวอย่างก่อนการทำ Contour	48
ภาพประกอบที่ 3.30 ตัวอย่างภาพผลลัพธ์หลังการทำ Contour.....	48
ภาพประกอบที่ 3.31 ขั้นตอนการหาค่าความกว้างและความสูง	48
ภาพประกอบที่ 3.32 ขั้นตอนการนำค่าความสูงมาหารสอง.....	49
ภาพประกอบที่ 3.33 ตัวอย่างภาพใบเสร็จที่เป็น Input	49
ภาพประกอบที่ 3.34 ตัวอย่างภาพใบเสร็จหลังจากนำพิกัดมาตัดภาพเอาเฉพาะส่วนที่ต้องการ.....	50
ภาพประกอบที่ 3.35 ตัวอย่างภาพใบเสร็จ	50
ภาพประกอบที่ 3.36 ตัวอย่าง Code และ การสกัดตัวอักษรภาษาไทยและภาษาอังกฤษ	51
ภาพประกอบที่ 3.37 ผลลัพธ์หลังจากทดลองใช้งาน Tesseract	51
ภาพประกอบที่ 3.38 ขั้นตอนการทำงานของกระบวนการรู้จำตัวอักษร.....	51
ภาพประกอบที่ 3.39 ภาพต้นฉบับก่อนทำ Convolution ขนาด 5×5	52

สารบัญภาพประกอบ (ต่อ)

	หน้า
ภาพประกอบที่ 3.40 Filter สำหรับการทำ Convolution ขนาด 3 x 3.....	52
ภาพประกอบที่ 3.41 การทำ Convolution ในรอบแรก.....	53
ภาพประกอบที่ 3.42 ค่าใหม่หลังจากการ Stride ในรอบที่ 1.....	53
ภาพประกอบที่ 3.43 การทำ Convolution ในรอบที่ 2.....	54
ภาพประกอบที่ 3.44 ค่าใหม่หลังจากการ Stride ในรอบที่ 2.....	54
ภาพประกอบที่ 3.45 ผลลัพธ์จากการทำ Convolution จนครบทุกช่อง.....	55
ภาพประกอบที่ 3.46 ภาพก่อนการทำ Max Pooling.....	55
ภาพประกอบที่ 3.47 ภาพหลังการทำ Max Pooling.....	56
ภาพประกอบที่ 3.48 ตัวอย่าง Time-Step ของข้อมูล.....	56
ภาพประกอบที่ 3.49 ตัวอย่างการ Feed ข้อมูลเข้าที่ละ Time Step.....	57
ภาพประกอบที่ 3.50 กราฟแสดงความน่าจะเป็นของอักขระที่ละ Time-Step.....	61
ภาพประกอบที่ 3.51 ตัวอย่างการทำงานของ Levenshtein Edit Distance ของคำว่า เต่าปิ้งย่าง	62
ภาพประกอบที่ 3.52 ตัวอย่างการทำงานของ Levenshtein Edit Distance ของคำว่า จานแรกหมู..	63
ภาพประกอบที่ 3.53 ตัวอย่างข้อมูลใบเสร็จของร้านชาบู อู๊ดเป็นต่อ (ชุดข้อมูล 1 ภาพ).....	64
ภาพประกอบที่ 3.54 ตัวอย่างชุดข้อมูลใบเสร็จของร้านชาบู อู๊ด เป็นต่อ (ชุดข้อมูล 1 บรรทัด).....	64
ภาพประกอบที่ 3.55 ตัวอย่างชุดข้อมูลใบเสร็จของร้านชาบู อู๊ด เป็นต่อ (ชุดข้อมูล 1 คำ).....	64
ภาพประกอบที่ 3.56 Context Diagram.....	65
ภาพประกอบที่ 3.57 ภาพแผนภาพกระแสข้อมูล ระดับที่1 (Data Flow Diagram Level).....	65
ภาพประกอบที่ 3.58 แผนผัง ER Diagram.....	66
ภาพประกอบที่ 4.1 การตรวจจับข้อความผิดพลาด ภาพที่ 1.....	77
ภาพประกอบที่ 4.2 การตรวจจับข้อความไม่ได้ ภาพที่ 1.....	77
ภาพประกอบที่ 4.3 ตรวจจับข้อความไม่ได้ ภาพที่ 2.....	79
ภาพประกอบที่ 4.4 ตรวจจับข้อความไม่ได้ ภาพที่ 3.....	80
ภาพประกอบที่ 4.5 การตรวจจับข้อความผิดพลาด ภาพที่ 4.....	81
ภาพประกอบที่ 4.6 การตรวจจับข้อความผิดพลาด ภาพที่ 5.....	83
ภาพประกอบที่ 4.7 ตรวจจับข้อความไม่ได้ ภาพที่ 5.....	83
ภาพประกอบที่ 4.8 ตรวจจับข้อความไม่ได้ ภาพที่ 6.....	85
ภาพประกอบที่ 4.9 กราฟแสดงค่า Loss ที่ได้จากการเรียนรู้ด้วย Optimizers 3 ประเภท.....	87
ภาพประกอบที่ 4.10 ภาพกราฟแสดงค่า Loss ที่ได้จาก Model.....	91
ภาพประกอบที่ 4.11 ตัวอย่างการทำนายผลบน Mobile Application.....	111
ภาพประกอบที่ 4.12 ตัวอย่างการทำนายผลบน Website.....	111
ภาพประกอบที่ ก - 1 หน้าแรกของ Application.....	118
ภาพประกอบที่ ก - 2 หน้าเลือกรูปภาพที่จะตรวจสอบ.....	118
ภาพประกอบที่ ก - 3 หน้ากดฟังก์ชัน process.....	119

สารบัญภาพประกอบ (ต่อ)

	หน้า
ภาพประกอบที่ ก - 4 หน้าแสดงผลทำนาย	119
ภาพประกอบที่ ข - 1 หน้าแรกของ Application	121
ภาพประกอบที่ ข - 2 หน้าเลือกรูปภาพนำเข้า	121
ภาพประกอบที่ ข - 3 เมื่อเลือกรูปภาพแล้วให้กด Submit.....	122
ภาพประกอบที่ ข - 4 โปรแกรมแสดงผลทำนาย	122