

สารบัญ

หน้า

บทคัดย่อ	ก
กิตติกรรมประกาศ.....	ข
สารบัญ.....	ค
สารบัญตาราง.....	จ
สารบัญภาพประกอบ.....	ฉ
บทที่ 1 บทนำ	1
1.1 หลักการและเหตุผล.....	1
1.2 วัตถุประสงค์ของโครงการ.....	2
1.3 ขอบเขตของโครงการ.....	2
1.4 ประโยชน์ที่คาดว่าจะได้รับ.....	3
1.5 อุปกรณ์และเครื่องมือที่ใช้ในการดำเนินงาน.....	3
1.6 แผนการดำเนินงาน.....	4
1.7 กรอบการดำเนินงาน.....	4
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง.....	7
2.1 ความเป็นมาของโมเดลแบบทรานส์ฟอร์มเมอร์ (History of Transformer-based Model).....	7
2.2 ประเภทของโมเดลทรานส์ฟอร์มเมอร์	12
2.3 ความเป็นมาของ BERT.....	13
2.4 เทคนิคสำหรับการประเมินโมเดล	24
2.5 Google Colaboratory (Google Colab).....	26
2.6 งานวิจัยที่เกี่ยวข้อง (Related Works)	26
บทที่ 3 วิธีดำเนินงานวิจัย.....	29
3.1 ชุดข้อมูล (Dataset).....	29
3.2 กรอบการดำเนินงาน.....	30
3.3 Pre-training of BERT	30
3.4 Fine-tuning BERT	55
3.5 ตัวอย่างการคำนวณการประเมินในแบบ Multi-class.....	56

สารบัญ(ต่อ)

	หน้า
บทที่ 4 ผลการทดลอง.....	62
4.1 ข้อมูลที่ใช้ในการทดสอบ.....	62
4.2 ผลการทดลองของโมเดลที่ศึกษา (Result).....	62
4.3 การทดสอบเปรียบเทียบกับโมเดลการจำแนกความรู้สึกอื่นๆ	64
บทที่ 5 สรุปและอภิปรายผลการทดลอง.....	72
5.1 สรุปผลและอภิปรายผล	72
5.2 ปัญหาและอุปสรรคในการดำเนินงาน	74
5.3 ข้อเสนอแนะ.....	75
เอกสารอ้างอิง	77
ภาคผนวก.....	80
ภาคผนวก ก คู่มือการติดตั้ง.....	81
ภาคผนวก ข คู่มือการใช้งาน.....	90
บทความวิจัย	93
โปสเตอร์โครงงาน.....	107
ประวัติย่อผู้จัดทำโครงงาน.....	109

สารบัญตาราง

หน้า

ตารางที่ 1.1	แผนการดำเนินงาน	4
ตารางที่ 2.1	เปรียบเทียบประสิทธิภาพของโมเดล Transformer กับโมเดล state-of-the-art.....	11
ตารางที่ 3.1	การตัดคำด้วย WordPiece ที่ได้รูปแบบของคำเต็ม	33
ตารางที่ 3.2	การตัดคำด้วย WordPiece ที่ได้รูปแบบของชิ้นส่วนของคำ.....	33
ตารางที่ 3.3	เมทริกซ์แสดงเวกเตอร์ของแต่ละคำในข้อมูลเอกสาร	36
ตารางที่ 3.4	ตัวอย่างการสร้าง Segment Embeddings.....	37
ตารางที่ 3.5	ค่า Positional encoding ของแต่ละเวกเตอร์ในข้อมูลเอกสาร	38
ตารางที่ 3.6	ผลรวมของค่า Positional encoding, Token Embedding, Segment Embedding	42
ตารางที่ 3.7	ผลลัพธ์ของผลรวมค่า Positional encoding และ Word embedding	43
ตารางที่ 3.8	ข้อมูลของเวกเตอร์ “love”	48
ตารางที่ 3.9	ข้อมูลของเวกเตอร์ “love” หลังจากปรับด้วยค่ามาตรฐาน.....	50
ตารางที่ 3.10	ข้อมูลของเวกเตอร์ “love” หลังจากปรับด้วยค่าสเกลและชิฟ.....	51
ตารางที่ 3.11	ตัวอย่าง คอนฟิวชั่นเมทริกซ์แบบ 3 class	57
ตารางที่ 4.1	คอนฟิวชั่นเมทริกซ์ของโมเดลการจำแนกความรู้สึกจากโมเดล BERT-base	63
ตารางที่ 4.2	แสดงการวัดประสิทธิภาพของโมเดลการจำแนกความรู้สึกจากโมเดล BERT-base.....	63
ตารางที่ 4.3	คอนฟิวชั่นเมทริกซ์ของโมเดลการจำแนกความรู้สึกจากอัลกอริทึม KNN.....	64
ตารางที่ 4.4	แสดงการวัดประสิทธิภาพของโมเดลการจำแนกความรู้สึกจากอัลกอริทึม KNN	65
ตารางที่ 4.5	คอนฟิวชั่นเมทริกซ์ของโมเดลการจำแนกความรู้สึกจากอัลกอริทึม SVM.....	66
ตารางที่ 4.6	แสดงการวัดประสิทธิภาพของโมเดลการจำแนกความรู้สึกจากอัลกอริทึม SVM	66
ตารางที่ 4.7	คอนฟิวชั่นเมทริกซ์ของโมเดลการจำแนกความรู้สึกจากอัลกอริทึม RF	67
ตารางที่ 4.8	แสดงการวัดประสิทธิภาพของโมเดลการจำแนกความรู้สึกจากอัลกอริทึม RF.....	67
ตารางที่ 4.9	คอนฟิวชั่นเมทริกซ์ของโมเดลการจำแนกความรู้สึกจากโมเดล roBERTa-base	68
ตารางที่ 4.10	แสดงการวัดประสิทธิภาพของโมเดลการจำแนกความรู้สึกจากโมเดล roBERTa-base..	69
ตารางที่ 4.11	สรุปผลการทดสอบโมเดลและเปรียบเทียบประสิทธิภาพของโมเดลแบบต่างๆ	70

สารบัญภาพประกอบ

หน้า

ภาพประกอบที่ 1.1	กรอบการดำเนินงานของโมเดลทรานสฟอร์มเมอร์สำหรับการจำแนกความรู้สึก.....	4
ภาพประกอบที่ 1.2	Fine-tuning โมเดล BERT สำหรับ Sentiment Analysis ด้วย Google Colab..	5
ภาพประกอบที่ 1.3	การใช้โมเดลทรานสฟอร์มเมอร์สำหรับการจำแนกความรู้สึกด้วย Test Set.....	6
ภาพประกอบที่ 1.4	การใช้โมเดลทรานสฟอร์มเมอร์สำหรับการจำแนกความรู้สึกด้วยการเขียนรีวิว	6
ภาพประกอบที่ 2.1	ตัวอย่างการทำ Machine Translation เพื่อแปลภาษาจีนเป็นภาษาไทย	7
ภาพประกอบที่ 2.2	แผนภาพของโมเดล seq2seq	7
ภาพประกอบที่ 2.3	การทำ Attention สำหรับโมเดล seq2seq	8
ภาพประกอบที่ 2.4	ตัวอย่างการจัดตำแหน่งของคำที่ได้จาก Attention	9
ภาพประกอบที่ 2.5	ภาพรวมของสถาปัตยกรรมแบบทรานสฟอร์มเมอร์.....	10
ภาพประกอบที่ 2.6	ตัวอย่างการใช้งานโมเดลแบบทรานสฟอร์มเมอร์ใน Machine Translator	10
ภาพประกอบที่ 2.7	ตัวอย่างการใช้งานโมเดลแบบทรานสฟอร์มเมอร์ใน Google Translate	11
ภาพประกอบที่ 2.8	แสดงตัวอย่าง One Hot Vector ของประโยค “I really love my dog.”	13
ภาพประกอบที่ 2.9	ตัวอย่าง BOW	14
ภาพประกอบที่ 2.10	ตัวอย่างโมเดล tf-idf.....	15
ภาพประกอบที่ 2.11	แสดงตัวอย่าง Word Embedding ของประโยค “I really love my dog.”	16
ภาพประกอบที่ 2.12	แสดงการใช้งาน Word Embedding	16
ภาพประกอบที่ 2.13	ตัวอย่าง Sparse vector และ Dense Vector	16
ภาพประกอบที่ 2.14	การพิจารณาค่าความเหมือนของคำด้วย Word2Vec ที่ใช้ window-size = 2.....	17
ภาพประกอบที่ 2.15	โมเดลแบบ CBOW	18
ภาพประกอบที่ 2.16	ตัวอย่างการทำนาย “คำ” จากบริบทรอบๆ ด้วยโมเดลแบบ CBOW	18
ภาพประกอบที่ 2.17	โมเดลแบบ Skip-gram.....	18
ภาพประกอบที่ 2.18	ตัวอย่างการทำนายบริบทที่อยู่รอบๆ “คำ” ด้วยโมเดลแบบ Skip-gram	19
ภาพประกอบที่ 2.19	กรอบการดำเนินงานใน BERT	19
ภาพประกอบที่ 2.20	Pre-training และ Fine-tuning ของ BERT	20
ภาพประกอบที่ 2.21	กรอบการดำเนินงานในส่วนของ Masked Language Model	21
ภาพประกอบที่ 2.22	ตัวอย่างการคำนวณค่าความน่าจะเป็นด้วยฟังก์ชัน SoftMax	22
ภาพประกอบที่ 2.23	กรอบการทำงานใน Next Sentence Prediction	23
ภาพประกอบที่ 2.24	ตารางคอนฟิวชั่นเมตริกซ์	25
ภาพประกอบที่ 3.1	ชุดข้อมูลบทวิจารณ์โรงแรมจาก TripAdvisor.....	29

สารบัญภาพประกอบ(ต่อ)

หน้า

ภาพประกอบที่ 3.2	กรอบการดำเนินงานของโมเดลทรานสฟอร์มเมอร์สำหรับการจำแนกความรู้สึก...	30
ภาพประกอบที่ 3.3	Language แบบ BERT.....	31
ภาพประกอบที่ 3.4	ตัวอย่างข้อมูลเอกสารจากเว็บ Tripadvisor	31
ภาพประกอบที่ 3.5	ตัวอย่างโปรแกรมสำหรับตัดคำ	32
ภาพประกอบที่ 3.6	ตัวอย่างคำใน “คำ” ในคลังคำศัพท์	34
ภาพประกอบที่ 3.7	ขั้นตอนการเตรียมแต่ละโทเค็นในข้อมูลเอกสารก่อนเข้าสู่โมเดลแบบ BERT	34
ภาพประกอบที่ 3.8	การเพิ่มโทเค็นพิเศษ [CLS] และ [SEP].....	35
ภาพประกอบที่ 3.9	การเปลี่ยนคำให้เป็นตัวเลขสำหรับบอกที่อยู่ของเวกเตอร์ใน lookup table	35
ภาพประกอบที่ 3.10	การดึงเวกเตอร์ขนาด 786 มิติที่ของแต่ละคำในข้อมูลเอกสาร	36
ภาพประกอบที่ 3.11	สร้างเวกเตอร์ Query (Q), Key (K) และ Value (V)	44
ภาพประกอบที่ 3.12	การแบ่งเวกเตอร์ Q, K และ V ที่มีขนาดเท่ากันให้กับแต่ละ Attention Head..	44
ภาพประกอบที่ 3.13	Attention Head แต่ละตัวที่ประกอบด้วยส่วนของเวกเตอร์ Q, K และ V.....	45
ภาพประกอบที่ 3.14	แสดงการคำนวณ Attention Score ของแต่ละ Attention Head	46
ภาพประกอบที่ 3.15	การคูณเวกเตอร์ V และการสร้างเวกเตอร์ r.....	46
ภาพประกอบที่ 3.16	Concatenation เวกเตอร์ r ทุก Attention Head เข้าด้วยกัน	47
ภาพประกอบที่ 3.17	แสดงการสร้างเวกเตอร์ r-Projected	47
ภาพประกอบที่ 3.18	กระบวนการทำ Residual Connections	48
ภาพประกอบที่ 3.19	การทำงานของ Feed-forward neural networks	52
ภาพประกอบที่ 3.20	ชั้นของ Encoder ภายในโมเดลแบบ BERT.....	52
ภาพประกอบที่ 3.21	การสุ่มเปลี่ยนคำในข้อมูลเอกสารจำนวน 15% เป็นโทเค็น [MASK].....	53
ภาพประกอบที่ 3.22	ขั้นตอนการทำงานในส่วนของ Mask Language Model.....	54
ภาพประกอบที่ 3.23	ขั้นตอนการทำงานในส่วนของ Next Sentient Prediction	55
ภาพประกอบที่ 3.24	ขั้นตอนการ Fine-tuning BERT สำหรับงานจำแนกความรู้สึก	56
ภาพประกอบที่ 4.1	ตัวอย่างบทวิจารณ์โรงแรมที่ใช้ในการทดสอบ	62
ภาพประกอบที่ 5.1	ตัวอย่างบทวิจารณ์ที่มีการใช้โครงสร้างทางภาษาที่ผิด.....	74
ภาพประกอบที่ 5.2	การใช้ทรัพยากรณ์ในการสร้างโมเดลด้วย BERT-Large	75
ภาพประกอบที่ 5.3	ข้อผิดพลาดด้านทรัพยากรณ์ไม่เพียงพอในการสร้างโมเดลด้วย BERT-Large.....	75
ภาพประกอบที่ ก-1	หน้าเว็บสำหรับเลือกดาวน์โหลด Node.js.....	82
ภาพประกอบที่ ก-2	ไฟล์สำหรับติดตั้ง Node.js	82

สารบัญภาพประกอบ(ต่อ)

	หน้า
ภาพประกอบที่ ก-3 แสดงการติดตั้ง Node.js ขั้นตอนที่ 1.....	83
ภาพประกอบที่ ก-4 แสดงการติดตั้ง Node.js ขั้นตอนที่ 2.....	83
ภาพประกอบที่ ก-5 แสดงการติดตั้ง Node.js ขั้นตอนที่ 3.....	84
ภาพประกอบที่ ก-6 แสดงการติดตั้ง Node.js ขั้นตอนที่ 4.....	84
ภาพประกอบที่ ก-7 แสดงการติดตั้ง Node.js ขั้นตอนที่ 5.....	85
ภาพประกอบที่ ก-8 แสดงการติดตั้ง Node.js ขั้นตอนที่ 6.....	85
ภาพประกอบที่ ก-9 หน้าสำหรับเลือกดาวน์โหลดโปรแกรม Visual Studio Code.....	86
ภาพประกอบที่ ก-10 หน้าที่อยู่ของโปรแกรมสำหรับติดตั้ง Visual Studio Code.....	86
ภาพประกอบที่ ก-11 แสดงการติดตั้ง Visual Studio Code ขั้นตอนที่ 1	87
ภาพประกอบที่ ก-12 แสดงการติดตั้ง Visual Studio Code ขั้นตอนที่ 2	87
ภาพประกอบที่ ก-13 แสดงการติดตั้ง Visual Studio Code ขั้นตอนที่ 3	88
ภาพประกอบที่ ก-14 แสดงการติดตั้ง Visual Studio Code ขั้นตอนที่ 4	88
ภาพประกอบที่ ก-15 แสดงการติดตั้ง Visual Studio Code ขั้นตอนที่ 5	89
ภาพประกอบที่ ก-16 แสดงการติดตั้ง Visual Studio Code ขั้นตอนที่ 6	89
ภาพประกอบที่ ข-1 หน้าเว็บแอปพลิเคชัน	91
ภาพประกอบที่ ข-2 แสดงส่วนการทำงานต่างๆของการทำงานแบบกรอกบทวิจารณ์	91
ภาพประกอบที่ ข-3 แสดงส่วนการทำงานต่างๆของการทำงานแบบใส่ชุดข้อมูล	92