การเปรียบเทียบการเรียนรู้ของเครื่องและโครงข่ายประสาทเทียมสำหรับการตรวจสอบคุณภาพงานเขียนอัตโนมัติ

น.ส.ณิชาพรรณ น้อยอยู่

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาวิทยาศาสตร์คอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2565
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

A COMPARISON OF MACHINE LEARNING AND NEURAL NETWORK ALGORITHMS FOR AN AUTOMATED THAI ESSAY QUALITY CHECKING

Miss Nichaphan Noiyoo

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Computer Science
Department of Computer Engineering
FACULTY OF ENGINEERING
Chulalongkorn University
Academic Year 2022

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	การเปรียบเทียบการเรียนรู้ของเครื่องและโครงข่ายประสาทเทียมสำหรับการตรวจสอบคุณภาพงานเขียนอัตโนมัติ
โดย	น.ส.ณิชาพรรณ น้อยอยู่
สาขาวิชา	วิทยาศาสตร์คอมพิวเตอร์
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	อาจารย์ ดร.เจษฎา ธัชแก้วกรพินธุ์

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

..... คณบดีคณะวิศวกรรมศาสตร์
(ศาสตราจารย์ ดร.สุพจน์ เตชวรสินสกุล)

คณะกรรมการสอบวิทยานิพนธ์

..... ประธานกรรมการ
(รองศาสตราจารย์ ดร.พีรพล เวทีกูล)

..... อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(อาจารย์ ดร.เจษฎา ธัชแก้วกรพินธุ์)

..... กรรมการภายนอกมหาวิทยาลัย
(ผู้ช่วยศาสตราจารย์ ดร.ภิญโญ แท้ประสาทสิทธิ์)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ณิชาพรรณ น้อยอยู่ : การเปรียบเทียบการเรียนรู้ของเครื่องและโครงข่ายประสาทเทียมสำหรับการตรวจสอบคุณภาพงานเขียนอัตโนมัติ. (A COMPARISON OF MACHINE LEARNING AND NEURAL NETWORK ALGORITHMS FOR AN AUTOMATED THAI ESSAY QUALITY CHECKING) อ.ที่ปรึกษาหลัก : อ. ดร.เจษฎา ธัชแก้วกรพินธุ์

การตรวจสอบคุณภาพงานเขียนเรียงความในภาษาไทยยังคงเป็นงานที่ยุ่งยาก เนื่องจากเป็นภาษาที่ซับซ้อนมากทั้งในด้านเครื่องหมายวรรคตอน โครงสร้างประโยค การซ้ำคำ การสะกดคำ การแสดงความคิดเห็น และการให้เหตุผล ดังนั้นการตรวจสอบคุณภาพงานเขียนเรียงความภาษาไทยจึงเป็นงานที่ต้องอาศัยทักษะของผู้ตรวจทั้งด้านการอ่านและการตีความ ทำให้ใช้เวลาในการตรวจคุณภาพงานเขียนมาก นอกจากนี้หากมีผู้ตรวจมากกว่า 1 คน อาจส่งผลต่อมาตรฐานที่ใช้ในการตรวจสอบคุณภาพงานเขียนที่แตกต่างกัน งานวิจัยนี้ได้รวบรวมข้อมูลเรียงความภาษาไทยที่เขียนโดยนิสิตที่ลงทะเบียนเรียนในหลักสูตรการเขียนย่อหน้า จากสถาบันภาษาไทยสิรินธรแห่งจุฬาลงกรณ์มหาวิทยาลัย และนำแบบจำลองหน่วยความจำระยะสั้นแบบยาว (LSTM) แบบจำลองโครงข่ายประสาทเทียมแบบสังวัตนาการ (CNN) แบบจำลองเบิร์ต (BERT) และแบบจำลองวังจันทร์เบอร์ต้า (WangchanBERTa) มาพัฒนาระบบสำหรับการตรวจสอบคุณภาพงานเขียนอัตโนมัติ เพื่อเปรียบเทียบประสิทธิภาพการตรวจสอบคุณภาพงานเขียนเรียงความในภาษาไทย จากผลการทดลองแบบจำลอง WangchanBERTa ที่ประมวลผลด้วย Adam Optimizer และใช้ Binary Crossentropy เป็น Loss Function เหมาะกับปัญหาประเภทการจัดกลุ่ม โดยมีประสิทธิภาพในการทำนายคุณภาพการเขียนเรียงความในภาษาไทยสูง และมีค่าความถูกต้องสูงกว่า 90% ส่วนแบบจำลอง CNN มีค่าความถูกต้องสูงกว่า 87% ในขณะที่เมื่อประมวลผลด้วย RMSprop Optimizer และใช้ Mean Squared Error เป็น Loss Function เหมาะกับปัญหาประเภทการถดถอยและมีความถูกต้องอยู่ในช่วง 90% - 98% จึงสรุปได้ว่าแบบจำลอง WangchanBERTa เหมาะสำหรับปัญหาการจำแนกประเภท และแบบจำลอง CNN เหมาะสำหรับปัญหาการถดถอย เพื่อให้สามารถทำนายคุณภาพงานเขียนเรียงความในความภาษาไทยได้อย่างมีประสิทธิภาพสูงสุด

สาขาวิชา วิทยาศาสตร์คอมพิวเตอร์ ลายมือชื่อนิสิต .....

ปีการศึกษา 2565 ลายมือชื่อ อ.ที่ปรึกษาหลัก .....

6470182621 : MAJOR COMPUTER SCIENCE
KEYWORD:

Nichaphan Noiyoo : A COMPARISON OF MACHINE LEARNING AND NEURAL NETWORK ALGORITHMS FOR AN AUTOMATED THAI ESSAY QUALITY CHECKING. Advisor: Prof. JESSADA THUTKAWKORAPIN

Checking the quality of essay writing in Thai language is still a complicated task because Thai language is very complex language in terms of punctuation, sentence structure, word repetition, spelling, commenting, and reasoning in content. Therefore, checking the quality of an essay and scoring require the reviewer's skills in reading and interpreting that make long time to review. In addition, if in reviewing process using more than one reviewer, it might affect different quality checking standards. We collected essay in Thai language which is written by student who registered paragraph writing course from The Sirindhorn Thai Language Institute of Chulalongkorn University. This work implemented LSTM model, CNN model, BERT model and WangchanBERTa model to compare the effectiveness of checking the quality of Thai essay writing. Our experimental result shows that classification analysis compiled with WangchanBERTa can achieve high accuracy up to 90%. However, CNN model compiled with classification analysis can achieve high accuracy up to 87% while compiled with regression analysis can achieve high accuracy in the range 90%. In conclusion, the system that we proposed can predict the quality of Thai essays with high accuracy. Therefore, we recommended Wangchanberta model for classification problem and CNN model for regression problem.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

Field of Study: Computer Science Student's Signature .....
Academic Year: 2022 Advisor's Signature .....

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยดีจากแรงสนับสนุน คำแนะนำ ความช่วยเหลือ และกำลังใจจากบุคคลหลายฝ่าย ผู้วิจัยจึงใคร่ขอใช้เนื้อหาในส่วนของกิตติกรรมประกาศเพื่อขอขอบพระคุณเป็นอย่างสูงมา ณ ที่นี้

ขอขอบพระคุณ อาจารย์ เจษฎา ธัชแก้วกรพินธุ์ อาจารย์ที่ปรึกษาวิทยานิพนธ์ ที่ได้เสียสละเวลาให้ความอนุเคราะห์และเอาใจใส่ ให้คำปรึกษาแนะนำ แก้ไขปรับปรุงข้อบกพร่อง รวมทั้งให้คำแนะนำองค์ความรู้ แนวทางในการศึกษาค้นคว้าเป็นอย่างดียิ่งมาโดยตลอด ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้ ดิฉันรู้สึกเป็นเกียรติอย่างสูงที่ได้เป็นลูกศิษย์ของอาจารย์

ขอขอบพระคุณ คณะกรรมการสอบวิทยานิพนธ์ผู้ช่วยศาสตราจารย์ ดร.พีรพล เวทีกูล ประธานกรรมการวิทยานิพนธ์ และผู้ช่วยศาสตราจารย์ ดร.ภิญโญ แท้ประสาทสิทธิ์ กรรมการภายนอกมหาวิทยาลัย ที่ได้สละเวลาในการให้ข้อเสนอแนะเพื่อให้วิทยานิพนธ์นี้เป็นประโยชน์ต่อการศึกษาอย่างแท้จริง

ขอขอบพระคุณสถาบันภาษาไทยสิรินธร จุฬาลงกรณ์มหาวิทยาลัย ที่ให้ความอนุเคราะห์ในการเก็บข้อมูลเรียงความและตรวจสอบคุณภาพงานเขียนแต่ละฉบับเพื่อใช้ในการทำวิจัย จนทำให้วิทยานิพนธ์ฉบับนี้สำเร็จลงด้วยดี

สุดท้ายนี้ขอขอบพระคุณ คุณพ่อคุณแม่และครอบครัว รวมถึงพี่อนๆ ที่เรียนปริญญาโท คณะวิศวกรรมคอมพิวเตอร์ด้วยกัน ที่สนับสนุนและให้กำลังใจกันตลอดจนวิทยานิพนธ์เล่มนี้เสร็จสมบูรณ์

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ณิชาพรรณ น้อยอยู่

สารบัญ

หน้า

..... ค

บทคัดย่อภาษาไทย ..... ค

..... ง

บทคัดย่อภาษาอังกฤษ ..... ง

กิตติกรรมประกาศ ..... จ

สารบัญ ..... ฉ

สารบัญตาราง ..... ฌ

สารบัญรูปภาพ ..... ฎ

บทที่ 1 บทนำ 1

1.1 ที่มาและความสำคัญของปัญหา 1

1.2 วัตถุประสงค์ 2

1.3 ขอบเขตของงานวิจัย 2

1.4 ประโยชน์ที่คาดว่าจะได้รับ 3

1.5 วิธีการดำเนินการ 3

บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง 5

2.1 ทฤษฎีที่เกี่ยวข้อง 5

การประมวลผลภาษาธรรมชาติ (Natural language processing : NLP) 5

การเรียนรู้ของเครื่อง (Machine Learning : ML) 5

โครงข่ายประสาทเทียม (Artificial Neural Network : ANN) 6

โครงข่ายประสาทเทียมแบบสังวัตนาการ (Convolutional Neural Networks : CNN) 7

หน่วยความจำระยะสั้นแบบยาว (Long Short-Term Memory : LSTM) 7

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ทรานฟอร์เมอร์ (Transformers) 8
เบิร์ต (Bidirectional Encoder Representations from Transformers : BERT) 9
การเข้ารหัสวันฮอต (One-hot Encoding) 10
เซนเทนซ์พีส (Sentencepiece) 11
การฝังคำ (Word Embedding) 11
คลังคำศัพท์ (Bag of Words) 11
การวัดประสิทธิภาพและตัวชี้วัด 11
เมทริกซ์ความสับสน (Confusion Matrix) 11
ค่าความถูกต้อง (Accuracy) 12
ค่าความแม่นยำ (Precision) 12
ค่าความครบถ้วน (Recall) 12
เอฟวัน สกอร์ (F1-Score) 13
2.2 งานวิจัยที่เกี่ยวข้อง 13
บทที่ 3 แนวทางการดำเนินงาน 17
3.1 เครื่องมือ 17
3.2 แนวคิดการประมวลผลการจำแนกคุณภาพของงานเขียน 17
3.3 ข้อมูลที่ใช้ในการทดลอง 19
3.4 การเตรียมชุดข้อมูล 20
3.4.1 เทคนิคการสุ่มตัวอย่างเพิ่ม (Over-sampling) 20
3.4.2 การตัดคำภาษาไทย 22
3.5 การประมวลผลข้อมูล 24
3.6 ตัวปรับพารามิเตอร์ที่ใช้ในการจำแนกคุณภาพของงานเขียน 25
3.6.1 Optimizer 25
3.6.1.1 Adaptive Moment Estimation (Adam) 25

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

3.6.1.2 Root Mean Square Propagation (RMSprop) 26
3.6.2 Loss Function 27
3.7 การพัฒนาแบบจำลอง 27
3.8 เกณฑ์การวัดประสิทธิภาพของแบบจำลอง (Evaluation Metrics) 28
บทที่ 4 ผลการทดลอง 29
4.1 การจำแนกคุณภาพการเกริ่นเข้าเรื่องของบทนำ 29
4.2 การจำแนกคุณภาพความน่าสนใจของบทนำ 30
4.3 การจำแนกคุณภาพการเรียงลำดับความของเนื้อเรื่อง 32
4.4 การจำแนกคุณภาพการแสดงความคิดเห็นของเนื้อเรื่อง 33
4.5 การจำแนกคุณภาพการขมวดความของสรุปความ 34
4.6 การจำแนกคุณภาพความน่าสนใจของสรุปความ 35
4.7 การจำแนกส่วนของเรียงความ 36
4.8 ความสัมพันธ์ระหว่างกับความยาวของประโยคกับประสิทธิภาพของแบบจำลอง 37
บทที่ 5 สรุปผลและอภิปรายผลการทดลอง 41
5.1 สรุปผลและอภิปรายผลการทดลอง 41
5.2 ปัญหาและอุปสรรค 42
5.3 แนวทางการวิจัยในขั้นถัดไป 43
5.4 ผลงานวิจัยที่ได้รับการตีพิมพ์ 44
บรรณานุกรม 45
ประวัติผู้เขียน 46

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

สารบัญตาราง

หน้า

ตารางที่ 1 : แสดงการเข้ารหัสวันฮอต 11

ตารางที่ 2 : แสดงเมทริกซ์ความสับสน 11

ตารางที่ 3 : แสดงหัวข้อที่เป็นเกณฑ์การพิจารณาคุณภาพของงานเขียนของเรียงความ 21

ตารางที่ 4 : แสดงจำนวนเรียงความของแต่ละหัวข้อหลังทำหลังทำการสุ่มตัวอย่างเพิ่ม 22

ตารางที่ 5 : แสดงตัวอย่างข้อมูลเรียงความและการตัดคำแสดง 23

ตารางที่ 6 : แสดงค่า Accuracy ผลการทดลองการจำแนกคุณภาพการเกริ่นเข้าเรื่องของบทนำ 30

ตารางที่ 7 : แสดงค่า Precision ผลการทดลองการจำแนกคุณภาพการเกริ่นเข้าเรื่องของบทนำ 30

ตารางที่ 8 : แสดงค่า Recall ผลการทดลองการจำแนกคุณภาพการเกริ่นเข้าเรื่องของบทนำ..... 30

ตารางที่ 9 : แสดงค่า F1-Score ผลการทดลองการจำแนกคุณภาพการเกริ่นเข้าเรื่องของบทนำ..... 30

ตารางที่ 10 : แสดงค่า Accuracy ผลการทดลองการจำแนกคุณภาพความน่าสนใจของบทนำ 31

ตารางที่ 11 : แสดงค่า Precision ผลการทดลองการจำแนกคุณภาพความน่าสนใจของบทนำ..... 31

ตารางที่ 12 : แสดงค่า Recall ผลการทดลองการจำแนกคุณภาพความน่าสนใจของบทนำ..... 31

ตารางที่ 13 : แสดงค่า F1-Score ผลการทดลองการจำแนกคุณภาพความน่าสนใจของบทนำ..... 31

ตารางที่ 14 : แสดงค่า Accuracy ผลการทดลองการจำแนกการเรียงลำดับความของเนื้อเรื่อง..... 32

ตารางที่ 15 : แสดงค่า Precision ผลการทดลองการจำแนกการเรียงลำดับความของเนื้อเรื่อง..... 32

ตารางที่ 16 : แสดงค่า Recall ผลการทดลองการจำแนกการเรียงลำดับความของเนื้อเรื่อง 32

ตารางที่ 17 : แสดงค่า F1-Score ผลการทดลองการจำแนกการเรียงลำดับความของเนื้อเรื่อง 33

ตารางที่ 18 : แสดงค่า Accuracy ผลการทดลองการจำแนกการแสดงความคิดเห็นของเนื้อเรื่อง ... 33

ตารางที่ 19 : แสดงค่า Precision ผลการทดลองการจำแนกการแสดงความคิดเห็นของเนื้อเรื่อง.... 33

ตารางที่ 20 : แสดงค่า Recall ผลการทดลองการจำแนกการแสดงความคิดเห็นของเนื้อเรื่อง..... 34

ตารางที่ 21 : แสดงค่า F1-Score ผลการทดลองการจำแนกการแสดงความคิดเห็นของเนื้อเรื่อง 34

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตารางที่ 22 : แสดงค่า Accuracy ผลการทดลองการจำแนกการขมวดความของสรุปความ 34
ตารางที่ 23 : แสดงค่า Precision ผลการทดลองการจำแนกการขมวดความของสรุปความ 35
ตารางที่ 24 : แสดงค่า Recall ผลการทดลองการจำแนกการขมวดความของสรุปความ 35
ตารางที่ 25 : แสดงค่า F1-Score ผลการทดลองการจำแนกการขมวดความของสรุปความ..... 35
ตารางที่ 26 : แสดงค่า Accuracy ผลการทดลองการจำแนกความน่าสนใจของสรุปความ 36
ตารางที่ 27 : แสดงค่า Precision ผลการทดลองการจำแนกความน่าสนใจของสรุปความ 36
ตารางที่ 28 : แสดงค่า Recall ผลการทดลองการจำแนกความน่าสนใจของสรุปความ..... 36
ตารางที่ 29 : แสดงค่า F1-Score ผลการทดลองการจำแนกความน่าสนใจของสรุปความ..... 36
ตารางที่ 30 : แสดงค่า Accuracy ผลการทดลองการจำแนกส่วนของเรียงความ..... 37
ตารางที่ 31 : แสดงค่า Precision ผลการทดลองการจำแนกส่วนของเรียงความ..... 37
ตารางที่ 32 : แสดงค่า Recall ผลการทดลองการจำแนกส่วนของเรียงความ..... 37
ตารางที่ 33 : แสดงค่า F1-Score ผลการทดลองการจำแนกส่วนของเรียงความ 37
ตารางที่ 34 : แสดงค่า Accuracy การจำแนกการเรียงลำดับความของเนื้อเรื่องที่ความยาวต่างกัน. 38
ตารางที่ 35 : แสดงค่า Accuracy ผลการจำแนกคุณภาพของแบบจำลองที่มีประสิทธิภาพที่สุด 43

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

สารบัญรูปภาพ

หน้า

รูปที่ 1 : โครงข่ายประสาทเทียม 6

รูปที่ 2 : โครงข่ายประสาทเทียมแบบสังวัตนาการ (CNN) 7

รูปที่ 3 : หน่วยความจำระยะสั้นแบบยาว (LSTM) 8

รูปที่ 4 : สถาปัตยกรรมทรานฟอร์เมอร์ 9

รูปที่ 5 : สถาปัตยกรรมเบิร์ต 10

รูปที่ 6 : ขั้นตอนการดำเนินการ 18

รูปที่ 7 : กราฟแท่งแสดงความผันแปรของจำนวนคำและจำนวนเรียงความของบทนำ 19

รูปที่ 8 : กราฟแท่งแสดงความผันแปรของจำนวนคำและจำนวนเรียงความของเนื้อเรื่อง 20

รูปที่ 9 : กราฟแท่งแสดงความผันแปรของจำนวนคำและจำนวนเรียงความของสรุปความ 20

รูปที่ 10 : ตัวอย่างข้อมูลเรียงความหนึ่งชุดในรูปแบบ JSON 24

รูปที่ 11 : ตัวอย่างข้อมูลเรียงความหนึ่งชุดในรูปแบบ Dataframe 25

รูปที่ 12 : สูตรการคำนวณของตัวปรับพารามิเตอร์ Adam 26

รูปที่ 13 : สูตรการคำนวณของตัวปรับพารามิเตอร์ RMSprop 26

รูปที่ 14 : Confusion Matrix จำแนกการเกริ่นนำเข้าเรื่องของ LSTM กับ Adam 39

รูปที่ 15 : Confusion Matrix จำแนกการเกริ่นนำเข้าเรื่องของ CNN กับ Adam 39

รูปที่ 16 : Confusion Matrix จำแนกการเกริ่นนำเข้าเรื่องของ LSTM กับ RMSprop 40

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บทที่ 1
บทนำ

1.1 ที่มาและความสำคัญของปัญหา

การตรวจสอบคุณภาพของงานเขียนเรียงความ เป็นงานอย่างหนึ่งที่ใช้เวลาเป็นอย่างมากต่อเรียงความ 1 ฉบับ จึงมีผู้พัฒนาระบบสำหรับการตรวจสอบคุณภาพงานเขียนและประเมินคะแนนเรียงความโดยอัตโนมัติออกมาเป็นจำนวนมาก แต่อย่างไรก็ตามยังไม่มีผู้ที่พัฒนาระบบอัตโนมัติสำหรับการตรวจสอบคุณภาพงานเขียนเรียงความที่มีเนื้อความเป็นภาษาไทยออกมาอย่างเป็นทางการ งานวิจัยนี้จึงจัดทำขึ้นเพื่อพัฒนาระบบอัตโนมัติสำหรับการตรวจสอบคุณภาพงานเขียนเรียงความที่มีเนื้อความเป็นภาษาไทย เพื่อตรวจสอบคุณภาพงานเขียนของเนื้อหาส่วนต่าง ๆ ในเนื้อความได้ ทำให้ผู้เขียนสามารถทราบถึงคุณภาพงานเขียนของตนเอง และนำไปปรับปรุงคุณภาพงานเขียนได้ในอนาคต อีกทั้งยังช่วยลดเวลาและเพิ่มประสิทธิภาพในการตรวจให้กับผู้ตรวจเรียงความอีกด้วย งานวิจัยนี้ได้รับความอนุเคราะห์จากสถาบันภาษาไทยสิรินธร จุฬาลงกรณ์มหาวิทยาลัย ซึ่งเป็นสถาบันจัดการเรียนการสอนเพื่อการพัฒนาความสามารถในการใช้ภาษาไทยแก่ผู้เรียนทุกระดับการศึกษาตลอดจนดำเนินการวิจัย และเผยแพร่ความรู้เกี่ยวกับการใช้ภาษาไทย ที่ให้ข้อมูลเรียงความสำหรับใช้ในการฝึกสอนและทดสอบในการพัฒนาแบบจำลองการเรียนรู้ของเครื่องและโครงข่ายประสาทเทียม

ในงานวิจัยนี้ได้นำเทคนิคการเรียนรู้ของเครื่อง (Machine Learning) และโครงข่ายประสาทเทียม (Artificial Neural Network) อย่างแบบจำลองหน่วยความจำระยะสั้นแบบยาว (LSTM) แบบจำลองโครงข่ายประสาทเทียมแบบสังวัตนาการ (CNN) แบบจำลองเบิร์ต (BERT) และแบบจำลองวังจันทร์เบอร์ต้า (WangchanBERTa) มาพัฒนาแบบจำลองทางภาษาโดยใช้ Optimizer 2 ประเภท ได้แก่ Adaptive Moment Estimation (Adam) และ Root Mean Square Propagation (RMSProp) เพื่อเปรียบเทียบว่าแบบจำลองใดมีประสิทธิภาพในการสามารถทำนายคุณภาพของงานเขียนของเนื้อหาส่วนต่าง ๆ อาทิ เช่น บทนำ เนื้อเรื่อง สรุปความ ในเรียงความภาษาไทยที่ดีและเหมาะสมสำหรับข้อมูลเรียงความที่สุด เนื่องจากในงานวิจัย [3] ได้นำเทคนิคโครงข่ายประสาทเทียมอย่าง LSTM มาใช้ในการตรวจให้คะแนนเรียงความอัตโนมัติ ได้แบบจำลองที่มีคุณภาพ มีค่า Quadratic weighted kappa (QWK) สูงถึง 0.70 และในงานวิจัย [4] ได้นำแบบจำลองทางภาษาซึ่งพัฒนาจากสถาปัตยกรรมทรานฟอร์เมอร์อย่าง BERT มาทำการทดลองร่วมกับการทำ Features ด้วย ในการพัฒนาแบบจำลองสำหรับการตรวจให้คะแนนเรียงความ

อัตโนมัติ ซึ่งมีค่า QWK สูงถึง 0.80 นอกเหนือจากนี้ [5] ได้นำแบบจำลองทางภาษาที่พัฒนาจากสถาปัตยกรรมทรานฟอร์เมอร์อย่างแบบจำลอง WangchanBERTa ซึ่งเป็นแบบจำลองทางภาษาที่พัฒนาจากชุดข้อมูลภาษาไทย มาใช้ในการวิเคราะห์ความคิดเห็นเกี่ยวกับหุ้นไทย ซึ่งแบบจำลองในงานวิจัยดังกล่าวมีค่า Accuracy สูงถึงร้อยละ 92.52 และ [6] ได้นำแบบจำลอง WangchanBERTa มาพัฒนาแบบจำลองทางภาษาที่ใช้ในการจัดประเภทของบทความเกี่ยวกับการตลาดในภาษาไทย เพื่อเปรียบเทียบกับแบบจำลองอื่น ๆ อย่าง แบบจำลอง CNN Conv1D, CNN Skip-gram, CNN FastText Skip-gram, BERT และ RoBERTa เมื่อเปรียบเทียบของแต่ละแบบจำลองจะพบว่าการใช้แบบจำลอง WangchanBERTa มีประสิทธิภาพในการจัดประเภทของบทความเกี่ยวกับการตลาดในภาษาไทยดีที่สุดในทุกๆด้านและมีผลลัพธ์ค่า Accuracy ที่ดีที่สุดอยู่ที่ร้อยละ 93.00 โดยแบบจำลอง WangchanBERTa เองนั้นก็ได้รับขนานนามว่าเป็นการเปลี่ยนแปลงครั้งยิ่งใหญ่ในงานด้านภาษาธรรมชาติในภาษาไทย ถือเป็นแบบจำลองทางภาษาที่มีความถูกต้องและแม่นยำมากในข้อมูลภาษาไทยโดยเฉพาะ

1.2 วัตถุประสงค์

1. เพื่อพัฒนาแบบจำลองการเรียนรู้ของเครื่องและโครงข่ายประสาทเทียม ที่มีประสิทธิภาพในการทำนายคุณภาพของเนื้อความส่วนต่าง ๆ อาทิ เช่น บทนำ เนื้อเรื่อง สรุปความ ในเรียงความภาษาไทย ได้ตรงหรือใกล้เคียงกับความเป็นจริงมากที่สุด โดยอ้างอิงจากค่าความถูกต้อง
2. เพื่อเพิ่มประสิทธิภาพในการตรวจสอบคุณภาพงานเขียนเรียงความให้กับผู้ตรวจเรียงความ
3. เพื่อลดเวลาที่ใช้ในการตรวจสอบคุณภาพงานเขียนเรียงความให้กับผู้ตรวจเรียงความ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

1.3 ขอบเขตของงานวิจัย

ข้อมูลที่ใช้ในการวิจัยเป็นข้อมูลเรียงความที่ประกอบด้วยเนื้อหาส่วนต่าง ๆ ดังนี้ หัวข้อเรียงความที่ได้รับ, ความครอบคลุมของหัวข้อ, ความน่าสนใจของหัวข้อ, การเกริ่นเข้าเรื่องของบทนำ, ความน่าสนใจของบทนำ, การเรียงลำดับความของเนื้อเรื่อง, การแสดงความคิดเห็นของเนื้อเรื่อง, การขมวดความของสรุปความ, ความน่าสนใจของสรุปความ, การแบ่งย่อหน้าของเรียงความ, การใช้ภาษาของเรียงความ, คะแนนรวมที่ได้ของเรียงความ, และส่วนเนื้อความของเรียงความ จากนักเรียนของ

สถาบันภาษาไทยสิรินธร จุฬาลงกรณ์มหาวิทยาลัย ที่อยู่ในรูปแบบ JSON โดยที่ไม่มีข้อมูลส่วนบุคคลของผู้เขียนประกอบอยู่ด้วย โดยแต่ละเรียงความจะได้รับการตรวจสอบคุณภาพงานเขียนของเนื้อความทั้งจากบทนำ, เนื้อเรื่อง, และสรุป ตามหมวดหมู่ และการประเมินคะแนนของเนื้อความตามเกณฑ์ที่กำหนดไว้ ซึ่งผู้เขียนจะได้รับหัวข้อในการเขียนเรียงความ และผู้เขียนจะมีอิสระในการเขียนเพราะแต่ละหัวข้อจะไม่มีคำตอบที่ตายตัว แต่คะแนนการเขียนจะขึ้นอยู่กับความสามารถในการอธิบาย, การใช้คำ, การแบ่งวรรคตอน, ความครอบคลุมของเนื้อหากับหัวข้อ และการนำเสนอตามหัวข้อที่ได้รับ โดยจำนวนเรียงความที่นำมาใช้ในงานวิจัยจะถูกเขียนด้วยภาษาไทยและมีจำนวน 576 เรียงความ

1.4 ประโยชน์ที่คาดว่าจะได้รับ

1. แบบจำลองการเรียนรู้ของเครื่องและโครงข่ายประสาทเทียมสามารถทำนายคุณภาพงานเขียนของเนื้อหาส่วนต่าง ๆ อาทิ เช่น บทนำ เนื้อเรื่อง สรุปความ ในเรียงความภาษาไทยที่ดีและเหมาะสมสำหรับข้อมูลเรียงความที่สุด

2. แบบจำลองการเรียนรู้ของเครื่องและโครงข่ายประสาทเทียมสามารถช่วยผู้ตรวจในด้านการเพิ่มประสิทธิภาพของการตรวจสอบคุณภาพงานเขียนเรียงความ

3. แบบจำลองการเรียนรู้ของเครื่องและโครงข่ายประสาทเทียมสามารถช่วยผู้ตรวจในด้านการลดเวลาที่ใช้ในการตรวจสอบคุณภาพงานเขียนเรียงความ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

1.5 วิธีการดำเนินการ

1. ศึกษาทฤษฎีและงานวิจัยที่เกี่ยวข้องกับการตรวจสอบคุณภาพงานเขียนของเนื้อหาในเรียงความและการให้คะแนนเรียงความอัตโนมัติ

2. แปลงข้อมูลในรูปแบบ JSON ให้อยู่ในรูปแบบ Dataframe

3. ทำ Feature Extraction ข้อมูลที่ต้องใช้ในขั้นตอนการสร้างแบบจำลองการเรียนรู้ของเครื่องและโครงข่ายประสาทเทียม

4. ทดลองสร้างแบบจำลองการเรียนรู้ของเครื่องและโครงข่ายประสาทเทียมที่เหมาะสมกับการตรวจสอบคุณภาพงานเขียนของเนื้อหาในเรียงความ

5. วัดประสิทธิภาพของแต่ละแบบจำลอง

6. ปรับพารามิเตอร์ให้แต่ละแบบจำลองมีประสิทธิภาพในการทำนายคุณภาพงานเขียนของเนื้อหาในเรียงความได้ดีที่สุด

7. เปรียบเทียบประสิทธิภาพของทุกแบบจำลอง

8. ตีพิมพ์ผลงานทางวิชาการ

9. สรุปผลการดำเนินงาน

10. จัดทำเล่มวิทยานิพนธ์และสอบวิทยานิพนธ์

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

บทที่ 2
ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ทฤษฎีที่เกี่ยวข้อง

การประมวลผลภาษาธรรมชาติ (Natural language processing : NLP)

การประมวลผลภาษาธรรมชาติ คือ วิทยาการแขนงหนึ่งในเทคโนโลยีปัญญาประดิษฐ์ (Artificial Intelligence) ทำงานในเชิงโปรแกรมคอมพิวเตอร์ที่สามารถดำเนินการวิเคราะห์และแปลงข้อมูลภาษาธรรมชาติได้ ช่วยให้มนุษย์ปฏิสัมพันธ์กับคอมพิวเตอร์ด้วยภาษาธรรมชาติ และช่วยให้คอมพิวเตอร์สามารถเข้าใจภาษาธรรมชาติของมนุษย์ใช้สนการสื่อสาร โดยในการทำงานที่เกี่ยวกับ NLP จะต้องใช้ความรู้ในด้านหลายด้าน เช่น ด้านวิทยาการคอมพิวเตอร์ (Computer science) ด้านภาษาศาสตร์ (Linguistics) ด้านคณิตศาสตร์ (Mathematics) ด้านสถิติศาสตร์ (Statistics) ด้านจิตวิทยา (Psychology) และด้านปรัชญา (Philosophy) เพื่อปิดช่องว่างทางการสื่อสารระหว่างมนุษย์และคอมพิวเตอร์และสามารถทำงานในด้านภาษาได้ เช่น การย่อความ (Summarizing) การจำแนกข้อความ (Text Classification) การแปลภาษา (Translation) การตอบคำถาม (Questions and Answers) การวิเคราะห์อารมณ์ความรู้สึก (Sentiment Analysis)

ในงานวิจัยนี้จะใช้แบบจำลองทางคอมพิวเตอร์ ได้แก่ การเรียนรู้ของเครื่อง โครงข่ายประสาทเทียม และนำเทคนิคในด้านการประมวลภาษาธรรมชาติเพื่อให้คอมพิวเตอร์ เรียนรู้และสามารถตรวจสอบคุณภาพงานเขียนของเนื้อหาในเรียงความภาษาไทยได้

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

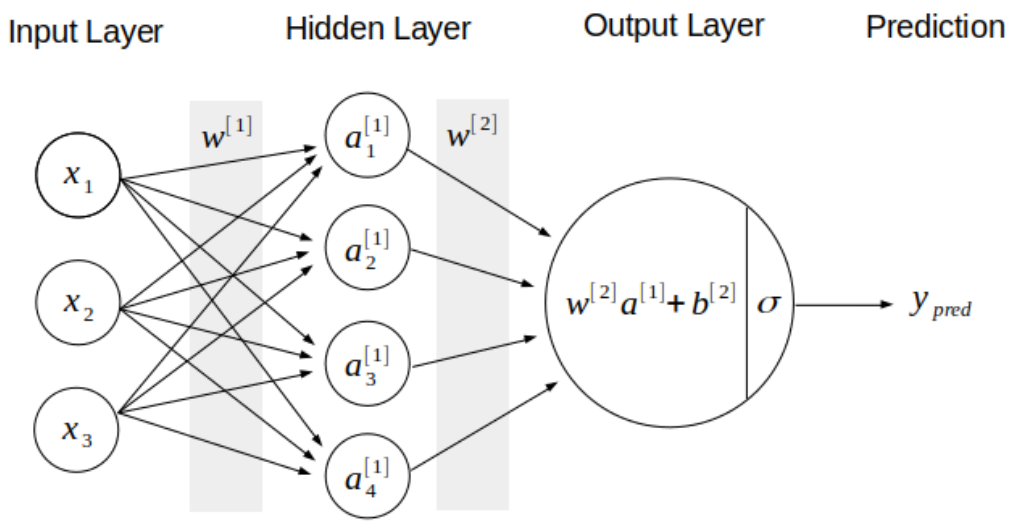
การเรียนรู้ของเครื่อง (Machine Learning : ML)

การเรียนรู้ของเครื่อง คือ วิทยาการแขนงหนึ่งในเทคโนโลยีปัญญาประดิษฐ์ (Artificial Intelligence) ทำงานในเชิงโปรแกรมคอมพิวเตอร์ที่สามารถดำเนินการวิเคราะห์ข้อมูล ที่ดำเนินการวิเคราะห์ด้วยแบบจำลองโดยอัตโนมัติ ภายใต้แนวคิดที่ว่า ระบบต่าง ๆ สามารถที่เรียนรู้และมีปฏิสัมพันธ์กับชุดข้อมูลต่าง ๆ รวมถึงสามารถระบุ และทราบรูปแบบต่าง ๆ ที่เกิดขึ้นภายในชุดข้อมูล และนำไปสู่การตัดสินใจคำตอบได้เองโดยไม่จำเป็นต้องพึ่งพามนุษย์ การเรียนรู้ของเครื่องสามารถแบ่งออกเป็น 3 ประเภท ดังนี้ 1. การเรียนรู้แบบผู้สอน (Supervised Learning) 2. การเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) 3. การเรียนรู้แบบเสริมกำลัง (Reinforcement Learning)

โดยในงานวิจัยนี้จะใช้แบบจำลองประเภทการเรียนรู้ของเครื่องแบบมีผู้สอน ซึ่งข้อมูลแต่ละเรียงความสามารถจำแนกคุณภาพของงานเขียนได้ด้วยคุณภาพงานเขียนของเนื้อหาในเรียงความส่วนต่าง ๆ ที่ได้รับการตรวจโดยผู้เชี่ยวชาญแล้ว

โครงข่ายประสาทเทียม (Artificial Neural Network : ANN)

โครงข่ายประสาทเทียม หรือ ข่ายงานประสาทเทียม คือ แบบจำลองทางคณิตศาสตร์ที่พัฒนาขึ้นเพื่อจำลองเอาวิธีการทำงานของสมองมนุษย์ หรือการทำให้คอมพิวเตอร์รู้จักคิดและจดจำเหมือนกับโครงข่ายประสาทของสมองมนุษย์ เพื่อช่วยให้คอมพิวเตอร์สามารถเข้าใจภาษามนุษย์ได้ โดยที่โครงข่ายประสาทของมนุษย์ประกอบด้วยเซลล์ประสาท (Neuron) หลายพันล้านเซลล์ เชื่อมต่อระหว่างเซลล์ดเข้าด้วยกันจนเกิดโครงข่ายประสาท การเชื่อมต่อของระบบประสาทนี้ยังถูกนำมาประยุกต์ใช้เป็นค่าถ่วงน้ำหนัก (Weight) ในการคำนวณค่าผลลัพธ์ของฟังก์ชันกระตุ้น (Activation function) และส่งค่าผลลัพธ์ไปยังชั้น (Layer) ถัดไป เพื่อปรับค่าถ่วงน้ำหนักให้ทำนายผลลัพธ์ให้มีความคลาดเคลื่อนน้อยที่สุด

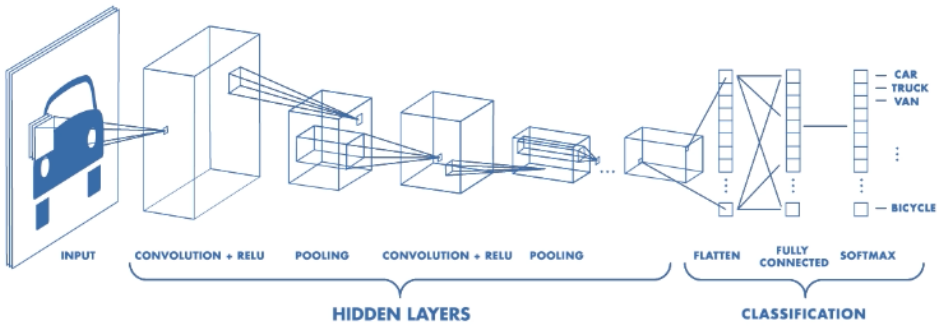


รูปที่ 1 : โครงข่ายประสาทเทียม

ที่มา : https://guopai.github.io/ml-blog14.html

โครงข่ายประสาทเทียมแบบสังวัตนาการ (Convolutional Neural Networks : CNN)

โครงข่ายประสาทเทียมแบบสังวัตนาการ คือ โครงข่ายประสาทเทียมประเภทหนึ่งในกลุ่ม bio-inspired โดยจำลองการมองเห็นของมนุษย์ที่มองพื้นที่เป็นที่ย่อย ๆ และนำกลุ่มของพื้นที่ย่อย ๆ มาผสานกัน เพื่อดูว่าสิ่งที่เห็นอยู่คืออะไร มักนิยมนำไปใช้กับข้อมูลที่เป็นรูปภาพ โดยโครงข่ายประสาทเทียมแบบสังวัตนาการมีความสามารถในการแยกคุณลักษณะ (feature) ของข้อมูลออกมาเป็นลักษณะย่อย ๆ โดยใช้การคำนวณตัวกรอง (filter) และเคอร์เนล (kernel) ที่จะดึงคุณลักษณะที่ใช้ในการรู้จำวัตถุออก โดยหนึ่งตัวกรองหรือหนึ่งเคอร์เนลจะดึงคุณลักษณะที่สนใจออกมาได้หนึ่งอย่าง ดังนั้นจึงต้องใช้หลายตัวกรองและหลายเคอร์เนลทำงานร่วมกันเพื่อหาคุณลักษณะหลาย ๆ อย่างประกอบกัน

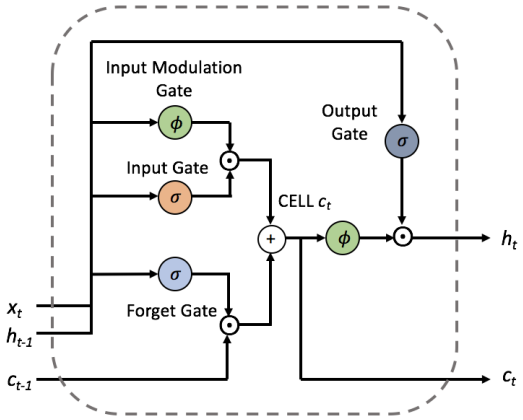


รูปที่ 2 : โครงข่ายประสาทเทียมแบบสังวัตนาการ (CNN)
ที่มา : https://medium.com/@natthawatphongchit/

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

หน่วยความจำระยะสั้นแบบยาว (Long Short-Term Memory : LSTM)

หน่วยความจำระยะสั้นแบบยาว คือ โครงข่ายประสาทเทียมแบบหนึ่งซึ่งถูกพัฒนามาจากโครงข่ายประสาทเทียมแบบวนซ้ำ (Recurrent Neural Network : RNN) เพื่อให้มีความเสถียรและมีประสิทธิภาพมากขึ้น หน่วยความจำระยะสั้นแบบยาวสร้างขึ้นเพื่อจำลองรูปแบบความจำของคนที่มีความจุของความทรงจำอยู่จำกัด โดยมีหลักการทำงาน คือ เซลล์สถานะจะเก็บสถานะ (State) หรือข้อมูลของแต่ละโหนด (Node) เอาไว้เพื่อที่เวลาย้อนกลับไปดูจะได้ได้ทราบถึงที่มาของข้อมูลค่าดังกล่าวค่าเดิมเป็นค่าอะไร และมีฟังก์ชันพิเศษที่ทำหน้าที่เหมือนประตู (Gate) ที่คอยควบคุมข้อมูลที่จะเข้ามาในแต่ละโหนดว่าควรจะถูกเก็บไว้หรือควรจะละทิ้งไป ซึ่งประกอบไปด้วย Forget gate layer, Input gate layer และ Output gate layer



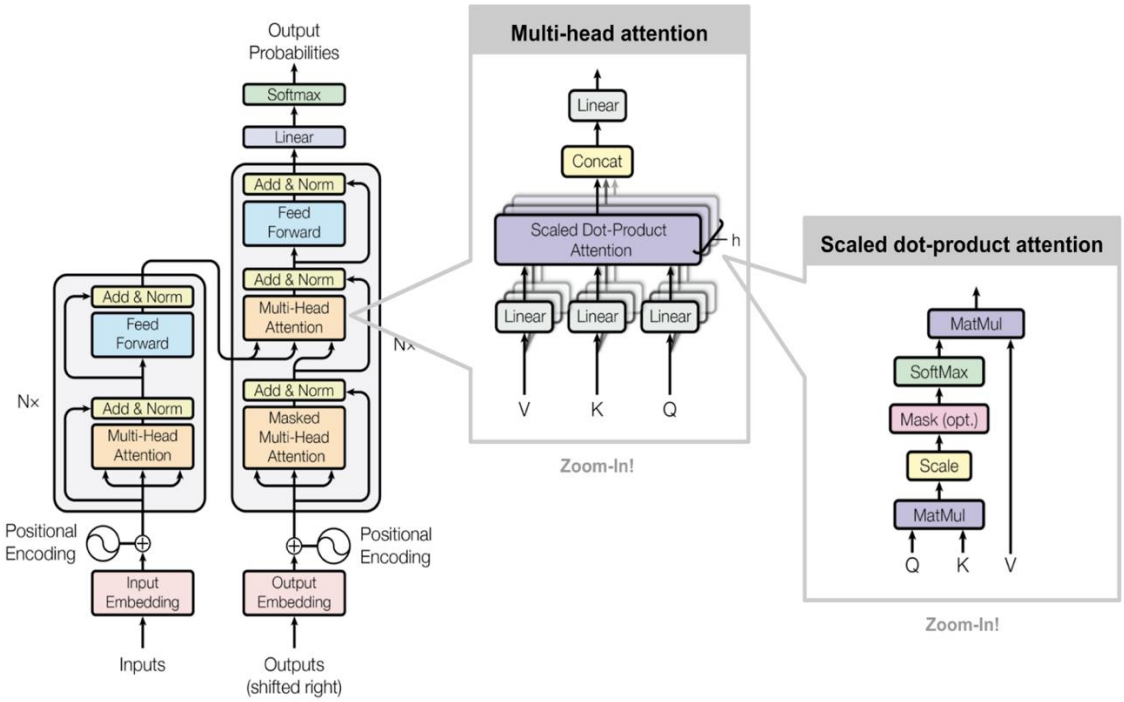
รูปที่ 3 : หน่วยความจำระยะสั้นแบบยาว (LSTM)

ที่มา : https://medium.com/@sinart.t/long-short-term-memory-lstm

ทรานฟอร์เมอร์ (Transformers)

ทรานฟอร์เมอร์ถูกนำเสนอขึ้นเมื่อปี 2017 ในงานวิจัยที่มีชื่อว่า “Attention Is All You Need” [1] โดยเป้าหมายในการสร้างแบบจำลองคือเพื่องานในการแปลภาษา หลังจากนั้นก็มีแบบจำลองที่ได้รับแรงบันดาลใจต่อ ๆ มา เช่น BERT, GPT, GPT-2, BART, GPT-3, หรือ Transformer XL หัวใจหลักของทรานฟอร์เมอร์ คือ กระบวนการ Self-Attention ที่แสดงถึงความข้องเกี่ยวกันของคำต่างๆ ในข้อความ ทำให้สามารถแก้ไขปัญหาการหานิพจน์ที่สามารถอ้างถึงสิ่ง ๆ เดียวกันในข้อความที่กำหนด (Coreference Resolution) ซึ่งมีความสำคัญอย่างมากต่องานกระบวนการที่คอมพิวเตอร์ทำความเข้าใจภาษาธรรมชาติ (Natural Language Processing : NLP) หลายประเภท โดยแนวคิดของ Self-Attention คือ การนำคำทุกคำในประโยคมาเปรียบเทียบกันให้แบบจำลองเรียนรู้และเลือกว่าจะสนใจคำไหน เมื่อไร ด้วยการแปลง Input เป็น 3 เวกเตอร์ คือ Q Query, K Key ซึ่งเอาไว้ Match กัน และ V Value ซึ่งค่าที่จะนำไป Attend จากสัดส่วนคะแนนผลลัพธ์การ Match ของ เวกเตอร์ Q กับ เวกเตอร์ K โดยการคำนวณทั้ง 3 เวกเตอร์ จะคำนวนจากการคูณ Input เข้ากับ Weight Matrix WQ, WK, WV ซึ่งก็คือ Linear Layer โดย Weight เหล่านั้นก็จะเป็น Learned Parameter ที่แบบจำลองจะเรียนรู้ขึ้นมาเองจากการฝึกสอนด้วย Gradient Descent ตามปกติ ข้อดีของทรานฟอร์เมอร์ คือ สามารถรองรับการเรียนรู้บนข้อมูลมหาศาลได้

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

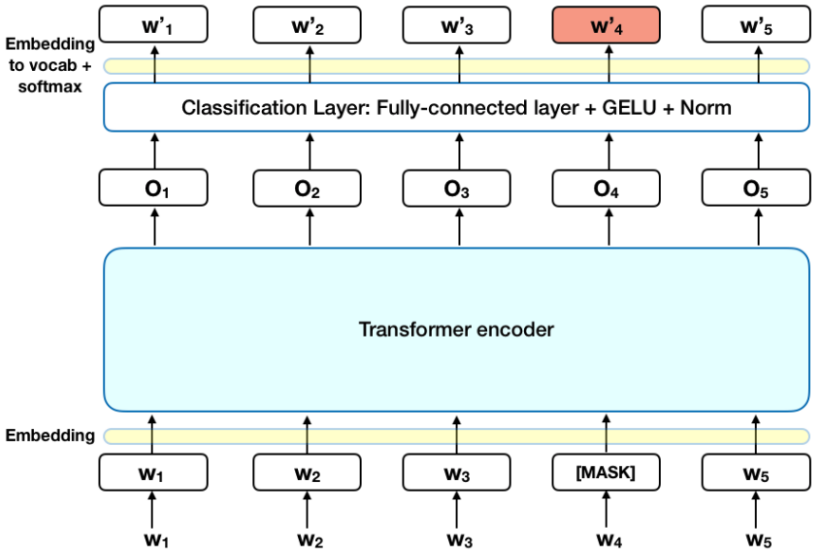


รูปที่ 4 : สถาปัตยกรรมทรานฟอร์เมอร์

ที่มา : https://deepfrench.gitlab.io/deep-learning-project/

เบิร์ต (Bidirectional Encoder Representations from Transformers : BERT)

เบิร์ต [2] คือ แบบจำลองที่ถูกพัฒนาต่อยอดจากสถาปัตยกรรมทรานฟอร์เมอร์ที่มีความสามารถในการประมวลผลแบบขนานทำให้สามารถเรียนรู้จากข้อมูลที่มีมหาศาลได้ คล้ายๆกับสถาปัตยกรรม GPT โดยแบบจำลองเบิร์ตประกอบด้วย 2 ส่วนหลัก ๆ คือ ส่วนเตรียมการเรียนรู้ (Pre-Training) และส่วนการปรับตั้งละเอียด (Fine-Tuning) และออกแบบให้เลือกใช้เฉพาะส่วนที่เป็นตัวเข้ารหัส (Encoder) จากสถาปัตยกรรมทรานฟอร์เมอร์ ซึ่งจะทำหน้าที่แปลงคำต่าง ๆ ในประโยคให้เปลี่ยนไปเป็นเวกเตอร์ โดยเบิร์ตได้เพิ่มแบบจำลองอีก 1 ตัว ต่อจากตัวเข้ารหัสที่มีอยู่เดิมเพื่อทำหน้าที่เป็นส่วนจำแนก (Classifier) เวกเตอร์ที่ได้จากตัวเข้ารหัสไปคำนวณต่อและประยุกต์ใช้ในงานด้านการประมวลผลทางภาษาประเภทต่าง ๆ ได้ เช่น การสรุปข้อความ การแปลภาษา การจำแนกประเภทของข้อความ การตอบคำถาม โดยในงานวิจัยนี้จะใช้แบบจำลองเบิร์ต เบส (BERT based) ซึ่งผ่านการเตรียมการเรียนรู้ภาษาไทยมาแล้วในการดำเนินงานสร้างแบบจำลองการเรียนรู้ที่สามารถจำแนกคุณภาพของงานเขียนของเนื้อหาในเรียงความส่วนต่าง ๆ



รูปที่ 5 : สถาปัตยกรรมเบิร์ต

ที่มา : https://towardsdatascience.com/bert-explained-state-of-the-art-language-model

วังจันทร์เบอร์ต้า (WangchanBERTa)

วังจันทร์เบอร์ต้า คือ แบบจำลองที่ใช้ในการฝึกฝนแบบจำลองภาษาไทย ที่พัฒนามาจากสถาปัตยกรรม RoBERTa-base ด้วยข้อมูลภาษาไทยขนาดใหญ่ที่มีความหลากหลายและมีการจัดการให้ข้อมูลมีความสะอาดที่สุดเท่าที่จะทำได้ โดยใช้กฎการจัดการข้อมูลที่สร้างขึ้นเพื่อภาษาไทยโดยเฉพาะ โดยวังจันทร์เบอร์ต้าถูกฝึกสอนบนชุดข้อมูลในภาษาไทยที่ได้จากแหล่งต่าง ๆ โดยขนาดของข้อมูลที่ใช้ในการฝึกสอนรวม 78.5 กิกะไบต์ (GB) และได้วัดประสิทธิภาพของแบบจำลองภาษาที่ทำการปรับตั้งละเอียดแล้ว ได้ผลคะแนนเฉลี่ยของ F1 Score สูงที่สุดบน 5 ชุดข้อมูล จากข้อมูลทั้งหมด 6 ชุดข้อมูล เนื่องจากวังจันทร์เบอร์ต้านั้นได้ใช้ข้อมูลจากโซเชียลมีเดียในการฝึกสอนด้วยทำให้สามารถเข้าใจภาษาไทยได้ดีทั้งแบบที่เป็นภาษาทางการและภาษาที่ใช้ในการแสดงความคิดเห็นต่าง ๆ อีกทั้งสถาปัตยกรรมวังจันทร์เบอร์ต้ายังได้ใช้แบบจำลองสำหรับตัดคำย่อยอย่าง SentencePiece ในการแบ่งคำด้วย ทำให้สถาปัตยกรรมวังจันทร์เป็นตัวเลือกแรก ๆ สำหรับงานด้านภาษาธรรมชาติที่เป็นภาษาไทยในปัจจุบัน

การเข้ารหัสวันฮอต (One-hot Encoding)

การเข้ารหัสวันฮอต คือ การแปลงข้อมูลที่ถูกเก็บในลักษณะข้อมูลที่จัดเป็นกลุ่ม (Categorical) ให้อยู่ในรูปแบบของระบบตัวเลขฐานสองที่มีค่าเป็น 0 หรือ 1 เท่านั้น ดังตารางที่ 1

	การขมวดความของสรุปความ	การเข้ารหัสวันฮอต
0	ขมวดความคิดได้	[1, 0, 0, 0]
1	ขมวดความคิดได้ดี	[0, 1, 0, 0]
2	ขมวดความคิดได้บ้าง	[0, 0, 1, 0]
3	ขมวดความคิดไม่ได้	[0, 0, 0, 1]

ตารางที่ 1 : แสดงการเข้ารหัสวันฮอต

เซนเทนซ์พีส (Sentencepiece)

เซนเทนซ์พีส คือ แบบจำลองการเรียนรู้ของเครื่องแบบไม่มีผู้สอนที่ใช้ในการแบ่งคำย่อย (Text Tokenizer) และ Detokenizer ซึ่งส่วนใหญ่ใช้สำหรับระบบการสร้างข้อความบนเครือข่ายประสาทเทียม โดยจะมีการกำหนดขนาดของคำศัพท์ไว้ก่อนการฝึกแบบจำลองประสาท

การฝังคำ (Word Embedding)

การฝังคำ คือ การแปลงคำให้อยู่ในรูปเวกเตอร์คุณลักษณะ ซึ่งในการแปลงชุดของคำจะนำเสนอในรูปของเวกเตอร์ โดยแปลงคำให้เป็นตัวเลขไม่ซ้ำกัน โดยการคำนวณความน่าจะเป็นของคำและบริบทภายในประโยค เวกเตอร์คุณลักษณะที่ได้สามารถนำไปคำนวณความคล้ายคลึงกับคำอื่น ๆ ในบริบทของคำที่แตกต่างออกไปได้

คลังคำศัพท์ (Bag of Words)

คลังคำศัพท์ คือ เทคนิคสำหรับการสร้าง Vector ของคำเพื่อแปลงคำแต่ละคำทั้งหมดที่ใช้ในการฝึกสอนแบบจำลองการเรียนรู้ของเครื่อง ซึ่งเป็นคำที่ไม่ซ้ำกัน ให้เป็นไอดี (ID) ของคำนั้น ๆ โดยยิ่งจำนวนของชุดข้อมูลเยอะ จำนวนคุณลักษณะก็ยิ่งมากขึ้นตาม

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

การวัดประสิทธิภาพและตัวชี้วัด

เมทริกซ์ความสับสน (Confusion Matrix)

เมทริกซ์ความสับสน ถือเป็นเครื่องมือสำคัญในการประเมินประสิทธิภาพของอัลกอริทึมในการทำนายผลลัพธ์ที่ได้จากแบบจำลองที่เราสร้างขึ้นในเครื่องมือการเรียนรู้ ซึ่งจะนำเสนอผลลัพธ์จากคลาสจริงและผลลัพธ์จากคลาสที่ถูกทำนาย

	คลาสที่ทำนายเป็นบวก	คลาสที่ทำนายเป็นลบ
คลาสจริงที่เป็นบวก	TP	FN
คลาสจริงที่เป็นลบ	FP	TN

ตารางที่ 2 : แสดงเมทริกซ์ความสับสน

True Positive (TP) – ผลลัพธ์จากคลาสที่ถูกทำนายเป็น “จริง” และผลลัพธ์จากคลาสจริงเป็น “จริง”

True Negative (TN) – ผลลัพธ์จากคลาสที่ถูกทำนายเป็น “ไม่จริง” และผลลัพธ์จากคลาสจริงเป็น “ไม่จริง”

False Positive (FP) - ผลลัพธ์จากคลาสที่ถูกทำนายเป็น “จริง” และผลลัพธ์จากคลาสจริงเป็น “ไม่จริง”

False Negative (FN) - ผลลัพธ์จากคลาสที่ถูกทำนายเป็น “ไม่จริง” และผลลัพธ์จากคลาสจริงเป็น “จริง”

ค่าความถูกต้อง (Accuracy)

ค่าความถูกต้องมักจะถูกใช้ในการเปรียบเทียบประสิทธิภาพของแบบจำลองการเรียนรู้ของเครื่อง ซึ่งค่าความถูกต้องจะพิจารณาจากค่าในคลาสทั้งหมดจากเมทริกซ์ความสับสน โดยการคำนวณหาอัตราส่วนของจำนวนคำตอบที่ถูกต้องเทียบกับจำนวนคำตอบทั้งหมด ดังสมการนี้

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

ค่าความแม่นยำ (Precision)

ค่าความแม่นยำจะใช้เพื่อวัดประสิทธิภาพของแบบจำลองการเรียนรู้ของเครื่อง ในการจำกัดผลลัพธ์ที่เป็นคลาสบวกเท็จ โดยการคำนวณหาอัตราส่วนของค่าที่แบบจำลองการเรียนรู้ของเครื่องทำนายเป็นคลาสบวกจริงกับค่าที่แบบจำลองการเรียนรู้ของเครื่องทำนายเป็นทั้งคลาสบวกจริงและคลาสเท็จจริง ดังสมการนี้

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

ค่าความครบถ้วน (Recall)

ค่าความครบถ้วนจะใช้เพื่อวัดประสิทธิภาพของแบบจำลองการเรียนรู้ของเครื่อง ในการทำนายผลลัพธ์ที่เป็นคลาสบวกจริง โดยการคำนวณหาอัตราส่วนของค่าที่แบบจำลองการเรียนรู้ของเครื่องทำนายเป็นคลาสบวกจริงกับค่าเหตุการณ์จริงเป็นทั้งคลาสบวกจริงและคลาสเท็จจริง ดังสมการนี้

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ค่าความครบถ้วนมักจะถูกนำมาพิจารณาในชุดข้อมูลที่ต้องการหลีกเลี่ยงการทำนายที่ได้ผลลัพธ์เป็นคลาสลบเท็จให้มากที่สุด

เอฟวัน สกอร์ (F1-Score)

เอฟวัน สกอร์จะใช้เพื่อวัดประสิทธิภาพของแบบจำลองการเรียนรู้ของเครื่อง เพื่อเป็นสรุปผลโดยรวมของแบบจำลองการเรียนรู้ โดยการคำนวณหาค่าเฉลี่ยแบบฮาร์โมนิคระหว่างค่าความแม่นยำกับค่าความครบถ้วน ดังสมการนี้

$$\mathrm{F1-Score} = 2 \times \frac{\mathrm{Precision} \times \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}}$$

2.2 งานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวขhองกับงานวิจัยนี้ คือ งานวิจัยที่ใชhแบบจำลองการเรียนรู้ของเครื่อง โครงข่ายประสาทเทียมรวมถึงการเรียนรู้เชิงลึกเพื่อทำนายคะแนนเรียงความ ซึ่งประเภทของงานวิจัยที่เกี่ยวข้องสามารถแบ่งออกเป็น 2 กลุ่ม ได้แก่ งานวิจัยที่ใช้ฟีเจอร์ประเภทการฝังคำและงานวิจัยที่ใช้สถาปัตยกรรมทรานฟอร์เมอร์

การใช้แบบจำลองที่เป็นโครงข่ายประสาทเทียมเพื่อเปรียบเทียบประสิทธิภาพในการให้คะแนนเรียงความอัตโนมัติ [3]

งานวิจัยนี้ได้นำเทคนิค LSTM, GRU และ BiLSTM ร่วมกับเทคนิค GloVe embedding มาใช้ในการพัฒนาแบบจำลองทางภาษาเพื่อเปรียบเทียบประสิทธิภาพในการทำนายคะแนนเรียงความ ชุด Automated Student Assessment Prize (ASAP) เป็นชุดข้อมูลสำหรับ Automated Essay Scoring (AES) บน Kaggle สำหรับการฝึกสอนและใช้ค่า QWK เป็นเกณฑ์ในการวัดผล เมื่อเปรียบเทียบผลลัพธ์กัน พบว่าการใช้แบบจำลอง LSTM ที่มี embedding dimensions เท่ากับ 200 ได้ผลลัพธ์ที่ดีที่สุดที่ค่าเฉลี่ย QWK = 0.68

การใช้แบบจำลองที่เป็นสถาปัตยกรรมทรานส์ฟอร์เมอร์กับแบบจำลองที่เป็นโครงข่ายประสาทเทียมเพื่อเปรียบเทียบประสิทธิภาพในการให้คะแนนเรียงความอัตโนมัติ [4]

งานวิจัยนี้ได้นำเทคนิค LSTM, LSTM+CNN, BERT(base) และ BERT + Features มาใช้ในการพัฒนาแบบจำลองทางภาษาเพื่อเปรียบเทียบประสิทธิภาพในการทำนายคะแนนเรียงความ โดยมีชุดข้อมูล ASAP จำนวน 8 ชุด สำหรับการฝึกสอนและใช้ค่า QWK เป็นเกณฑ์ในการวัดผล เมื่อเปรียบเทียบผลลัพธ์กัน พบว่าการใช้ BERT รวมกับการทำ Features ได้ผลลัพธ์ที่ดีที่สุดที่ค่าเฉลี่ย QWK = 0.801

การใช้แบบจำลองที่เป็นสถาปัตยกรรมทรานส์ฟอร์เมอร์เพื่อเปรียบเทียบประสิทธิภาพในการวิเคราะห์ความคิดเห็นเกี่ยวกับหุ้นไทย [5]

งานวิจัยนี้ได้นำเทคนิค TF-IDF, BERT และ WangchanBERTa มาใช้ในการพัฒนาแบบจำลองทางภาษาเพื่อเปรียบเทียบประสิทธิภาพในการวิเคราะห์ความคิดเห็นเกี่ยวกับหุ้นไทย โดยแบ่งชุดข้อมูลสำหรับการฝึกสอนออกเป็น 3 กลุ่ม ได้แก่ Buy (Positive polarity) จำนวน 165 รายการ Hold (Neutral polarity) จำนวน 165 รายการ Sell (Negative polarity) จำนวน 80 รายการ และแบ่งออกเป็นชุดฝึกสอนร้อยละ 70 ชุดทดสอบร้อยละ 30 และใช้ค่า Accuracy, Precision, Recall, and F1-score เป็นเกณฑ์ในการวัดผล เมื่อเปรียบเทียบผลลัพธ์กัน พบว่าการใช้ WangchanBERTa มีประสิทธิภาพในการวิเคราะห์ความคิดเห็นเกี่ยวกับหุ้นไทยดีกว่า BERT ในทุกๆ ด้าน โดยมีค่า Accuracy ที่ดีที่สุดอยู่ที่ร้อยละ 92.52 ในขณะที่ Accuracy ของ TF-IDF และ BERT มีค่าเท่ากับ ร้อยละ 85.03 และร้อยละ 89.12 ตามลำดับ

การใช้แบบจำลอง CNN และแบบจำลองที่เป็นสถาปัตยกรรมทรานส์ฟอร์เมอร์เพื่อเปรียบเทียบประสิทธิภาพในการจัดประเภทของบทความเกี่ยวกับการตลาดในภาษาไทย [6]

งานวิจัยนี้ได้นำเทคนิค CNN Conv1D, CNN Skip-gram, CNN FastText Skip-gram, BERT, RoBERTa และ WangchanBERTa มาใช้ในการพัฒนาแบบจำลองทางภาษาเพื่อเปรียบเทียบประสิทธิภาพในการจัดประเภทของบทความเกี่ยวกับการตลาดในภาษาไทย โดยชุดข้อมูลสำหรับการฝึกสอนแบ่งออกเป็น 2 กลุ่ม ได้แก่ บทความประเภท Timely และบทความประเภท Timeless โดยแบ่งออกเป็นชุดฝึกสอนร้อยละ 56 ชุดการตรวจสอบร้อยละ 24 และชุดทดสอบร้อยละ 20 ในการวัดผลใช้ค่า Accuracy, Precision, Recall, and F1-score เป็นเกณฑ์ เมื่อเปรียบเทียบผลลัพธ์กัน พบว่าการใช้แบบจำลอง WangchanBERTa มีประสิทธิภาพในการจัดประเภทของบทความเกี่ยวกับการตลาดในภาษาไทยดีที่สุดในทุกๆด้าน โดยมีค่า Accuracy ที่ดีที่สุดอยู่ที่ร้อยละ 93.00 ในขณะที่แบบจำลอง CNN FastText Skip-gram และแบบจำลอง RoBERTa มีค่า Accuracy เท่ากับร้อยละ 90.00 ซึ่งถือเป็นลำดับที่สองรองจากแบบจำลอง WangchanBERTa

แบบจำลอง SentencePiece ที่ใช้สำหรับการทำ Tokenizer และ Detokenizer ของคำย่อยในงานภาษาธรรมชาติ [7]

งานวิจัยนี้ได้นำเสนอ SentencePiece ซึ่งเป็นเครื่องมือ open-source ที่เปิดให้ใช้งานภายใต้ Apache 2 พัฒนาด้วยภาษา C++ และ Python สำหรับงานการแบ่งหน่วยคำย่อย โดยมีแรงจูงใจจากเครื่องมือแบ่งกลุ่มคำย่อยที่มีอยู่นั้นจะรับอินพุตที่ถูกแปลงเป็น tokenize มาแล้วล่วงหน้าก่อนจะแปลงเป็นลำดับของคำ (word sequences) แต่ SentencePiece นั้นสามารถฝึก

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

แบบจำลองการแบ่งคำย่อยได้โดยตรงจากประโยคดิบที่ยังไม่ผ่านการทำ tokenize มาก่อนเลย ทำให้สามารถสร้างระบบแบบ end-to-end และภาษาที่เป็นอิสระอย่างแท้จริงได้ ซึ่งในงานวิจัยนี้ได้ทำการทดลองตรวจสอบความถูกต้องของ NMT เกี่ยวกับการแปลภาษาระหว่างภาษาอังกฤษและภาษาญี่ปุ่น และจากค่า accuracy ที่เป็นไปได้ในการทดลองพบว่า SentencePiece มีความประสิทธิภาพในการจัดการคำย่อยที่ถูกต้องแม้ฝึกอบรมคำย่อยโดยตรงจากประโยคดิบเท่านั้น นอกจากนี้ยังเปรียบเทียบประสิทธิภาพของการฝึกคำย่อยและการแบ่งกลุ่มกับการกำหนดค่าต่างๆ โดย SentencePiece เปิดให้ใช้งานภายใต้ Apache 2

แบบจำลองทางภาษาไทย WangchanBERTa ซึ่งเป็น Pre-Training ที่พัฒนาจากสถาปัตยกรรมทรานส์ฟอร์เมอร์ [8]

งานวิจัยนี้ได้นำเสนอแบบจำลองทางภาษาไทยซึ่งเป็น Pre-Training ที่มีชื่อว่า WangchanBERTa พัฒนาจากแบบจำลองทางภาษาที่ใช้สถาปัตยกรรม Transformer เนื่องจากมีประสิทธิภาพสูงในงานที่มีดาวน์สตรีมจำนวนมาก และเหมาะสมกับการพัฒนาสำหรับภาษาที่มีทรัพยากรค่อนข้างต่ำ อย่างเช่น ภาษาไทย ซึ่งในงานวิจัยนี้ได้ใช้แบบจำลองทางภาษา RoBERTa-base สำหรับชุดข้อมูลขนาดใหญ่ที่ขจัดข้อมูลซ้ำซ้อนที่ใช้ในการฝึกสอน ขนาดรวม 78 กิกะไบต์ (GB) ซึ่งรวบรวมจากโซเชียลมีเดีย บทความข่าว และชุดข้อมูลอื่น ๆ ที่เปิดเผยต่อสาธารณะ ได้แก่ Wisesight, Wongnai, Prachathai และ Generated Reviews (EN–TH) ในงานวิจัยนี้ได้ใช้กฎการประมวลผลข้อความเฉพาะสำหรับภาษาไทย ซึ่งจุดที่สำคัญที่สุดคือการรักษาช่องว่าง ซึ่งเป็นขอบเขตสำคัญในการแบ่งประโยคในภาษาไทยก่อนการแปลง tokenize ของคำย่อย โดยมีการทดลองทั้งระดับคำ ระดับพยางค์ และการสร้าง tokenize ของ SentencePiece ด้วยชุดข้อมูลที่เล็กลงเพื่อสำรวจผลกระทบต่อการสร้าง tokenize ต่อประสิทธิภาพดาวน์สตรีม ผลลัพธ์ที่ได้คือแบบจำลอง Wangchanberta-base-att-spm-uncased มีประสิทธิภาพเหนือกว่า baselines อย่าง NBSVM, CRF และ ULMFit และแบบจำลองทางภาษาหลายภาษาอย่าง XLMR และ mBERT ทั้งด้านการจำแนกลำดับและงานการจัดหมวดหมู่ tokenize ในคำอธิบายประกอบโดยมนุษย์

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

แบบจำลองทางภาษา RoBERTa ซึ่งเป็น Pre-Training ที่พัฒนาจากแบบจำลองเบิร์ต [9]

งานวิจัยนี้ได้นำเสนอแบบจำลองทางภาษาที่มีชื่อว่า Robustly Optimized BERT Pretraining Approach (RoBERTa) ซึ่งเป็นตัว Pre-Training ที่ได้พัฒนาจากสถาปัตยกรรม Transformer และพัฒนาต่อยอดจากแบบจำลองทางภาษาอย่างแบบจำลองเบิร์ต (BERT) เนื่องจากมีโครงสร้างสถาปัตยกรรม Transformer ส่วนของ Encoder เหมือนกัน แต่มีขั้นตอนการฝึกสอนและ

ตัวตัดคำ (Tokenizer) ที่แตกต่างกัน โดยแบบจำลอง RoBERTa มีขั้นตอนการฝึกสอน ดังนี้ การเปลี่ยน Masking ทุกรอบของการฝึกสอน การเลือกจะไม่ใช้การทำนายประโยคถัดไป และใช้ตัวตัดคำแบบ Byte-Pair-Encoding (BPE)

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

บทที่ 3
แนวทางการดำเนินงาน

ในบทนี้จะกล่าวถึงแนวทางการดำเนินงาน โดยนำทฤษฎีและงานวิจัยที่เกี่ยวข้องในบทที่ 2 มาประยุกต์ใช้กับการทำนายคุณภาพงานเขียนของเนื้อหาส่วนต่าง ๆ เช่น บทนำ เนื้อเรื่อง สรุปความ ในเรียงความภาษาไทย ได้แก่ เครื่องมือ แนวคิดการประมวลผลการจำแนกคุณภาพของงานเขียน ข้อมูลที่ใช้ในการลอง การประมวลผลข้อมูล ตัวปรับพารามิเตอร์ที่ใช้ในการจำแนกคุณภาพของงานเขียน การพัฒนาแบบจำลอง และเกณฑ์การวัดประสิทธิภาพของแบบจำลอง

3.1 เครื่องมือ

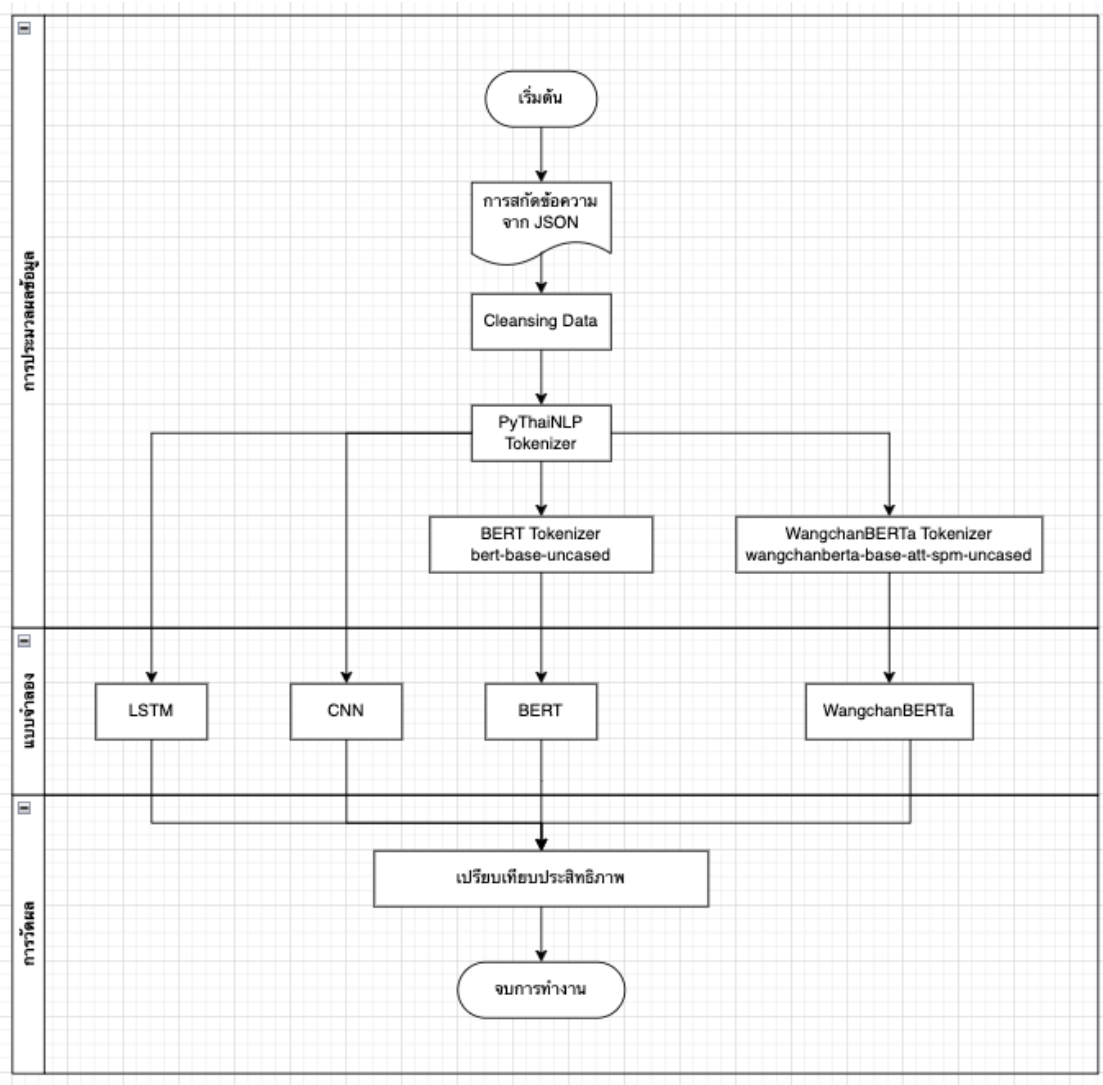
- ภาษาไพทอน (Python)
- เครื่องมือกูเกิล โคแล็บ (Google Colab)
- ไลบรารี TensorFlow
- ไลบรารี Pytorch
- ไลบรารี PythaiNLP

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

3.2 แนวคิดการประมวลผลการจำแนกคุณภาพของงานเขียน

งานวิจัยนี้จัดทำขึ้นเพื่อเปรียบเทียบความประสิทธิภาพการทำนายคุณภาพงานเขียนของเนื้อหาเรียงความในภาษาไทย จากแบบจำลองการเรียนรู้ของเครื่องและแบบจำลองโครงข่ายประสาทเทียมประเภทต่าง ๆ โดยการวัดผลจะเปรียบเทียบประสิทธิภาพของแบบจำลองแต่ละประเภทจากเทคนิคที่แตกต่างกันใช้ค่าความถูกต้อง (Accuracy), ค่าความแม่นยำ (Precision), ค่าความครบถ้วน (Recall), และเอฟวัน สกอร์ (F1-Score) เป็นเกณฑ์ในการเปรียบเทียบประสิทธิภาพ ในงานวิจัยนี้ได้สร้างแบบจำลองทั้งหมด 4 ประเภท ได้แก่ LSTM, CNN, BERT และ WangchangBERTa เพื่อทำนายคุณภาพงานเขียนของเนื้อหาส่วนต่าง ๆ อาทิ เช่น บทนำ, เนื้อเรื่อง, สรุปความ ในเรียงความภาษาไทยที่ดีและเหมาะสมสำหรับข้อมูลเรียงความที่สุด ในการสร้างแบบจำลองจะมีการกำหนดค่า 2 ชุดสำหรับการฝึกสอนแบบจำลอง โดยการกำหนดค่าชุดที่ 1 แบบจำลองจะถูกประมวลผลโดยใช้ตัวปรับพารามิเตอร์ Adam และ ใช้ Binary Crossentropy เป็น Loss Function การกำหนดค่านี้

ออกแบบมาเพื่อดูประสิทธิภาพของแบบจำลองจากมุมมองการจัดหมวดหมู่ และการกำหนดค่าชุดที่ 2 แบบจำลองจะถูกประมวลผลโดยใช้ตัวปรับพารามิเตอร์ RMSprop และ ใช้ Mean Squared Error เป็น Loss Function ซึ่งการกำหนดค่าเช่นนี้ออกแบบมาเพื่อดูประสิทธิภาพของแบบจำลองจากมุมมองการวิเคราะห์การถดถอยของปัญหา โดยในการพัฒนาแบบจำลองการเรียนรู้ของเครื่องและแบบจำลองโครงข่ายประสาทเทียมเพื่อทำนายคุณภาพงานเขียนของเนื้อหาเรียงความในภาษาไทยมีขั้นตอนการดำเนินการดังรูปที่ 6



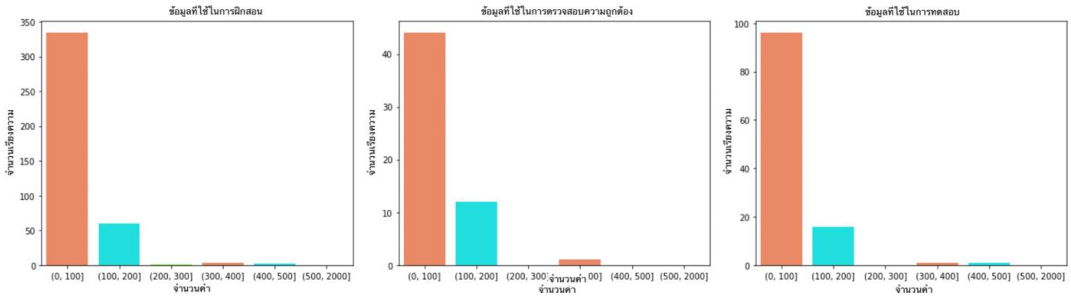
รูปที่ 6 : ขั้นตอนการดำเนินการ

3.3 ข้อมูลที่ใช้ในการทดลอง

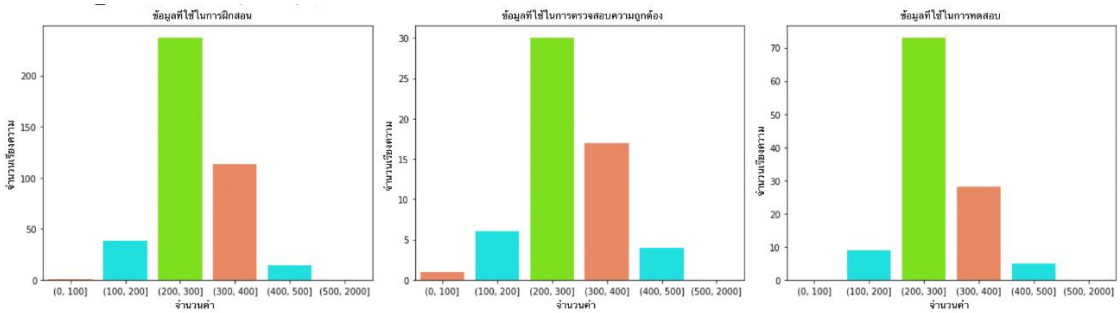
ข้อมูลที่นำมาใช้ในงานวิจัยนี้เป็นข้อมูลเรียงความของนิสิตที่ลงทะเบียนเรียนในหลักสูตรการเขียนย่อหน้าจากสถาบันภาษาไทยสิรินธร จุฬาลงกรณ์มหาวิทยาลัย ซึ่งเป็นสถาบันจัดการเรียนการสอนเพื่อการพัฒนาความสามารถในการใช้ภาษาไทยแก่ผู้เรียนทุกระดับการศึกษา ซึ่งประกอบด้วยจำนวนเรียงความทั้งหมด 576 เรียงความ ซึ่งประกอบหัวข้อเรียงความ จำนวน 3 หัวข้อ ได้แก่ งมงาย , บูชาคนเก่ง และก้าวร้าว โดยข้อมูลที่ได้นำมาใช้ในงานวิจัยนี้จะไม่มีส่วนที่ระบุตัวตนของผู้เขียนอยู่ด้วย โดยเรียงความจะมีผู้เชี่ยวชาญตรวจประเมินคุณภาพของงานเขียนตามแต่ละส่วนได้แก่ บทนำ, เนื้อเรื่อง และสรุปความ ซึ่งผู้ตรวจจะประเมินคุณภาพของงานเขียนตามหัวข้อต่าง ๆ ดังนี้ หัวข้อเรียงความที่ได้รับ, ความครอบคลุมของหัวข้อ, ความน่าสนใจของหัวข้อ, การเกริ่นเข้าเรื่องของบทนำ, ความน่าสนใจของบทนำ, การเรียงลำดับความของเนื้อเรื่อง, การแสดงความคิดเห็นของเนื้อเรื่อง, การขมวดความของสรุปความ, และความน่าสนใจของสรุปความ ดังตารางที่ 3

3.1.1 ความผันแปรของข้อมูล

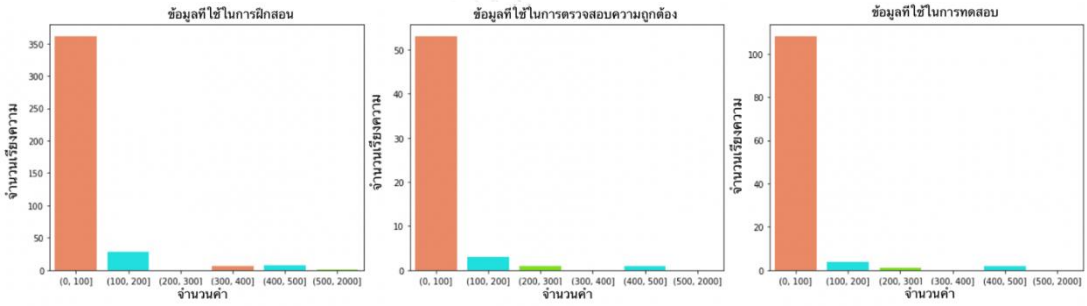
เมื่อนำข้อมูลมาแสดงผลในรูปแบบกราฟแท่งที่แสดงความผันแปรของข้อมูลเรียงความโดยแบ่งออกเป็นส่วน ได้แก่ บทนำ, เนื้อเรื่อง และสรุปความ และแบ่งข้อมูลเป็น 3 กลุ่ม กลุ่มแรกเป็นข้อมูลสำหรับฝึกสอนแบบจำลองคิดเป็นร้อยละ 70 กลุ่มที่สองเป็นข้อมูลสำหรับการตรวจสอบความถูกต้องของแบบจำลองคิดเป็นร้อยละ 10 และกลุ่มที่สามเป็นข้อมูลสำหรับการทดสอบแบบจำลองคิดเป็นร้อยละ 20 พบว่าส่วนใหญ่เป็นเรียงความที่มีขนาดสั้น บทนำจะมีจำนวนคำอยู่ในช่วง 0 ถึง 300 คำ ดังรูปที่ 7 ส่วนเนื้อเรื่องจะมีจำนวนคำอยู่ในช่วง 0 ถึง 500 คำ ดังรูปที่ 8 และสรุปความจะมีจำนวนคำอยู่ในช่วง 0 ถึง 100 คำ ดังรูปที่ 9



รูปที่ 7 : กราฟแท่งแสดงความผันแปรของจำนวนคำและจำนวนเรียงความของบทนำ



รูปที่ 8 : กราฟแท่งแสดงความผันแปรของจำนวนคำและจำนวนเรียงความของเนื้อเรื่อง



รูปที่ 9 : กราฟแท่งแสดงความผันแปรของจำนวนคำและจำนวนเรียงความของสรุปความ

เนื่องจากความผันแปรของจำนวนคำที่ไม่เท่ากันในแต่ละส่วนของเรียงความ ดังนั้นผู้วิจัยจึงมีการตัดความยาวของแต่ละส่วนประกอบเรียงความตามข้อมูลส่วนมาก โดยส่วนของบทนำจะตัดให้มีจำนวนคำ 300 คำ ส่วนของเนื้อเรื่องจะตัดให้มีจำนวนคำ 500 คำ แล ส่วนของสรุปความจะตัดให้มีจำนวนคำ 100 คำ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

3.4 การเตรียมชุดข้อมูล

การเตรียมชุดข้อมูลสำหรับการทำนายคุณภาพงานเขียนของเนื้อหาส่วนต่าง ๆ เช่น บทนำ เนื้อเรื่อง สรุปความ ในเรียงความภาษาไทย ได้แก่ เทคนิคการสุ่มตัวอย่างเพิ่ม (Over-sampling) การตัดคำภาษาไทย การกำจัดอักขระพิเศษ

3.4.1 เทคนิคการสุ่มตัวอย่างเพิ่ม (Over-sampling)

จากข้อมูลเรียงความของนักเรียนจากสถาบันภาษาไทยสิรินธร จุฬาลงกรณ์มหาวิทยาลัยที่ถูกรวบรวมเพื่อใช้เป็นชุดข้อมูลในการวิจัยเกิดปัญหาข้อมูลสมดุล ผู้วิจัยจึง

เลือกใช้เทคนิคการสุ่มตัวอย่างเพิ่ม (Over-sampling) ทำให้ตัวอย่างในแต่ละคลาสที่ใช้ในการฝึกสอนแบบจำลองประเภทต่าง ๆ มีจำนวนเท่ากันดังตารางที่ 4 เพื่อให้ชุดข้อมูลมีความสมดุลและไม่เกิดการเรียนรู้ของแบบจำลองที่ Overfitting เกินไป

ส่วน	หัวข้อที่พิจารณา	เกณฑ์คุณภาพของงานเขียน	จำนวนข้อมูล
บทนำ	การเกริ่นนำเข้าเรื่อง	เกริ่นเข้าเรื่องได้ดี	246
		เกริ่นเข้าเรื่องได้	287
		เกริ่นเข้าเรื่องไม่ดี	43
	ความน่าสนใจ	น่าสนใจ : ยกตัวอย่าง	183
		น่าสนใจ : ให้ข้อมูลเป็นเบื้องต้น	137
		ไม่น่าสนใจ	108
		น่าสนใจ : อ้างเหตุผล	81
		น่าสนใจ : อื่น ๆ	67
เนื้อเรื่อง	การเรียงลำดับความ	เนื้อเรื่องมีความราบรื่น	334
		ไม่ราบรื่น สับสน 1 แห่ง	193
		ไม่ราบรื่น สับสน 2 แห่งขึ้นไป	49
	การแสดงความคิดเห็น	ชัดเจนน่าเชื่อถือ	169
		บกพร่อง 1 แห่ง	256
		บกพร่อง 2 แห่ง	123
		บกพร่อง 3 แห่งขึ้นไป	28
สรุปความ	การขมวดความคิด	ขมวดความคิดได้ดี	242
		ขมวดความคิดได้	265
		ขมวดความคิดได้บ้าง	50
		ขมวดความคิดไม่ได้	15
	ความน่าสนใจ	ยกคำคม สุภาษิต	281
		ถามให้คิด	37
		โยงไปเรื่องอื่น	55
		ไม่น่าสนใจ	17
		อื่น ๆ	182

ตารางที่ 3 : แสดงหัวข้อที่เป็นเกณฑ์การพิจารณาคุณภาพของงานเขียนของเรียงความ

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

ส่วน	หัวข้อที่พิจารณา	เกณฑ์คุณภาพของงานเขียน	จำนวนข้อมูล
บทนำ	การเกริ่นนำเข้าเรื่อง	เกริ่นเข้าเรื่องได้ดี	287
		เกริ่นเข้าเรื่องได้	287
		เกริ่นเข้าเรื่องไม่ดี	287
	ความน่าสนใจ	น่าสนใจ : ยกตัวอย่าง	183
		น่าสนใจ : ให้ข้อมูลเป็นเบื้องต้น	183
		ไม่น่าสนใจ	183
		น่าสนใจ : อ้างเหตุผล	183
		น่าสนใจ : อื่น ๆ	183
เนื้อเรื่อง	การเรียงลำดับความ	เนื้อเรื่องมีความราบรื่น	334
		ไม่ราบรื่น สับสน 1 แห่ง	334
		ไม่ราบรื่น สับสน 2 แห่งขึ้นไป	334
	การแสดงความคิดเห็น	ชัดเจนน่าเชื่อถือ	256
		บกพร่อง 1 แห่ง	256
		บกพร่อง 2 แห่ง	256
		บกพร่อง 3 แห่งขึ้นไป	256
สรุปความ	การขมวดความคิด	ขมวดความคิดได้ดี	265
		ขมวดความคิดได้	265
		ขมวดความคิดได้บ้าง	265
		ขมวดความคิดไม่ได้	265
	ความน่าสนใจ	ยกคำคม สุภาษิต	281
		ถามให้คิด	281
		โยงไปเรื่องอื่น	281
		ไม่น่าสนใจ	281
		อื่น ๆ	281

ตารางที่ 4 : แสดงจำนวนเรียงความของแต่ละหัวข้อหลังทำหลังทำการสุ่มตัวอย่างเพิ่ม

3.4.2 การตัดคำภาษาไทย

จากชุดข้อมูลเรียงความภาษาไทยที่รวบรวมโดยสถาบันภาษาไทยสิรินธร จุฬาลงกรณ์มหาวิทยาลัยเพื่อใช้เป็นชุดข้อมูลสำหรับการฝึกสอนแบบจำลองต่าง ๆ มีรูปแบบ

การเขียนที่แตกต่างกัน เนื่องจากชุดข้อมูลมีหัวข้อเรียงความทั้งหมด 3 หัวข้อ ผู้จัดทำวิจัยจึงได้สร้างคุณลักษณะการตัดคำภาษาไทยโดยเลือกใช้ไลบรารี PyThaiNLP (Engine : NEWMM) และไลบรารี Sentencepiece เพื่อเรียกใช้แบบจำลอง Bert-base-uncase ในการแบ่งคำย่อยข้อมูลเรียงความสำหรับการฝึกสอนแบบจำลองเบิร์ต และใช้แบบจำลอง Wangchanberta-base-att-spm-uncased ในการแบ่งคำย่อยข้อมูลเรียงความสำหรับการฝึกสอนแบบจำลองวังจันทร์เบอร์ต้า (Wangchanberta) ตัวอย่างข้อมูลเรียงความและการตัดคำแสดงในตารางที่ 4

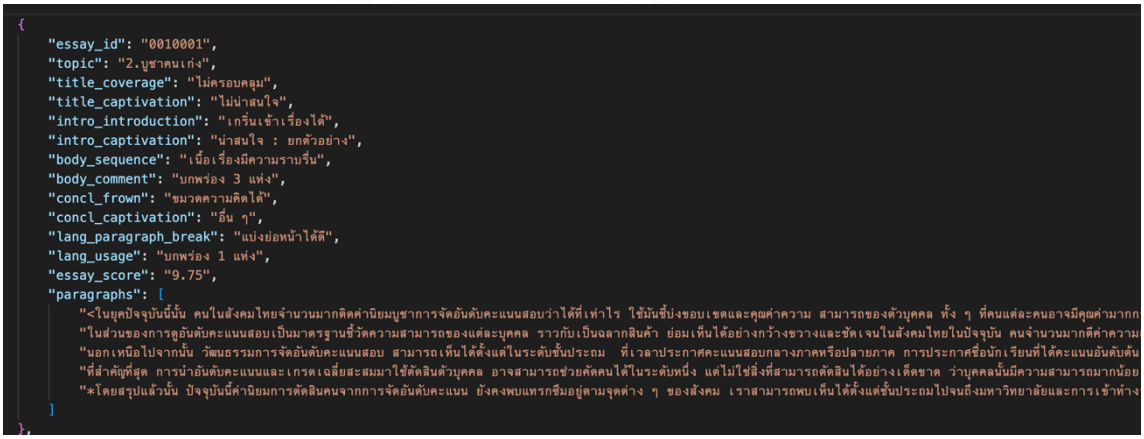
ส่วน	ข้อมูลเรียงความก่อนตัดคำ	ข้อมูลเรียงความหลังตัดคำ
บทนำ	ระบบการศึกษาถือเป็นสิ่งพื้นฐานที่คนไทยทุกคนคู่ควรที่จะได้รับ คนส่วนมากจึงผ่านการใช้ชีวิตของกรอบสังคมที่มีชื่อว่า โรงเรียน ถึงแม้ในความเป็นจริงแล้วหน้าที่ของโรงเรียนเป็นเพียงแค่สถานที่ที่ใช้ในการเรียนการสอนสำหรับนักเรียนแต่ด้วยบริบทการที่ต้องเข้ามาใช้ชีวิตภายในสถานที่นี้ตั้งแต่เช้าจรดเย็น ตั้งแต่วันจันทร์ถึงวันศุกร์ โรงเรียนจึงกลับกลายมาเป็นสถานที่ที่หล่อหลอมค่านิยมและบ่มเพาะนิสัยของคน หนึ่งไปโดยปริยาย จนทำให้เกิดข้อสงสัยว่าแล้วเราได้ถูกปลูกฝังค่านิยมอะไรจากการไปโรงเรียน	ระบบ การศึกษา ถือเป็น สิ่ง พื้นฐาน ที่ คนไทย ทุกคน คู่ควร ที่จะ ได้รับ คน ส่วนมาก จึง ผ่าน การ ใช้ชีวิต ของ กรอบ สังคม ที่ มี ชื่อว่า โรงเรียน ถึงแม้ ใน ความเป็นจริง แล้ว หน้าที่ ของ โรงเรียน เป็น เพียงแค่ สถานที่ ที่ ใช้ ใน การเรียน การสอน สำหรับ นักเรียน แต่ ด้วย บริบท การ ที่ ต้อง เข้ามา ใช้ชีวิต ภายใน สถานที่ นี้ ตั้ง แต่เช้า จรด เย็น ตั้ง แต่วัน จันทร์ ถึง วัน ศุกร์ โรงเรียน จึง กลับ กลาย มา เป็น สถานที่ ที่ หล่อหลอม ค่านิยม และ บ่ม เพาะ นิสัย ของ คน หนึ่ง ไป โดยปริยาย จน ทำให้เกิด ข้อ สงสัย ว่าแล้ว เรา ได้ ถูก ปลูกฝัง ค่านิยม อะไร จาก การ ไป โรงเรียน

ตารางที่ 5 : แสดงตัวอย่างข้อมูลเรียงความและการตัดคำแสดง

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

3.5 การประมวลผลข้อมูล

เนื่องจากข้อมูลดั้งเดิมเป็นข้อมูลที่นำออกมาจากฐานข้อมูลในรูปแบบ JSON ซึ่งประกอบด้วยข้อมูล ไอดีของเรียงความ, หัวข้อเรียงความที่ได้รับ, ความครอบคลุมของหัวข้อ, ความน่าสนใจของหัวข้อ, การเกริ่นเข้าเรื่องของบทนำ, ความน่าสนใจของบทนำ, การเรียงลำดับความของเนื้อเรื่อง, การแสดงความคิดเห็นของเนื้อเรื่อง, การขมวดความของสรุปความ, ความน่าสนใจของสรุปความ, การแบ่งย่อหน้าของเรียงความ, การใช้ภาษาของเรียงความ, คะแนนรวมที่ได้ของเรียงความ, และส่วนเนื้อความของเรียงความ ดังรูปที่ 11 จากนั้นจึงแปลงให้ข้อมูลของเรียงความที่อยู่ในรูปแบบ JSON ให้อยู่ในรูปแบบของ Dataframe ดังรูปที่ 12 ขั้นตอนถัดมาคือการทำความสะอาดข้อความ เช่น ลบอักขระพิเศษ เพื่อให้ข้อความอยู่ในรูปแบบที่เหมาะสมก่อนทำการตัดคำ และการสกัดฟีเจอร์จากข้อความ โดยสำหรับแบบจำลองหน่วยความจำระยะสั้นแบบยาว (LSTM) และแบบจำลองโครงข่ายประสาทเทียมแบบสังวัตนาการ (CNN) จะใช้เพียงโมดูล PyThaiNLP ในการตัดคำ แต่ในส่วนของแบบจำลองเบิร์ต (BERT) และแบบจำลองวังจันทร์เบอร์ต้า (WangchanBERTa) จะมีการเพิ่มขั้นตอนการตัดคำที่แตกต่างจากแบบจำลองอื่น ๆ เพื่อเพิ่มความถูกต้อง โดยใช้แบบจำลองการตัดคำอย่าง Sentencepiece มาทำการตัดคำอีกครั้งหลังจากใช้โมดูล PyThaiNLP ในการตัดคำมาแล้ว โดยจะใช้แบบจำลองการตัดคำ bert-base-uncased กับแบบจำลอง BERT และใช้แบบจำลองการตัดคำ wangchanberta-base-att-spm-uncased กับแบบจำลอง WangchanBERTa



รูปที่ 10 : ตัวอย่างข้อมูลเรียงความหนึ่งชุดในรูปแบบ JSON

intro_introduction	intro_captivation	body_sequence	body_comment	concl_frown	concl_captivation	...	intro	body	conclusion
เกริ่นเข้าเรื่องได้ดี	ไม่น่าสนใจ	เนื้อเรื่องมีความราบรื่น	บกพร่อง 3 แห่ง	None	None	...	ในปัจจุบันสังคมในประเทศไทยให้ความสำคัญกับคะแนน...	ถ้าหากมองอีกมุมหนึ่งคนที่สอบได้คะแนนไม่ดีหรือส...	ซึ่งการยกย่องผู้คนที่คะแนนสอบนั้นไม่ได้เป็นเรื่...
เกริ่นเข้าเรื่องได้ดี	ไม่น่าสนใจ	เนื้อเรื่องมีความราบรื่น	บกพร่อง 1 แห่ง	None	None	...	เราคงปฏิเสธไม่ได้ว่าคนไทยนั้นให้ความสำคัญด้านก...	ตั้งแต่เล็กจนโตเราคงจะได้ยินคนพูดกันเรื่องคะแน...	'เราคงปฏิเสธไม่ได้ว่าคนไทยนั้นให้ความสำคัญด้า...
เกริ่นเข้าเรื่องได้ดี	น่าสนใจ : ให้ข้อมูลเป็นเบื้องต้น	ไม่ราบรื่น สับสน 2 แห่ง	บกพร่อง 2 แห่ง	None	None	...	วัยรุ่น เป็นวัยที่เด็กเริ่มมีความคิดเห็นเป็นของ...	ในปัจจุบันเริ่มมีการเข้ามาของเด็ก generation ...	'วัยรุ่น เป็นวัยที่เด็กเริ่มมีความคิดเห็นเป็น...
เกริ่นเข้าเรื่องไม่ดี	น่าสนใจ : ให้ข้อมูลเป็นเบื้องต้น	ไม่ราบรื่น สับสน 2 แห่ง	บกพร่อง 4 แห่ง	None	None	...	ปัจจุบันการแข่งขันในระบบการศึกษาสูงมาก จนบางที...	จะเป็นไปไม่ได้เลยหรอที่ความเก่งของเด็กคนหนึ่งไ...	'ปัจจุบันการแข่งขันในระบบการศึกษาสูงมากจนบาง...

รูปที่ 11 : ตัวอย่างข้อมูลเรียงความหนึ่งชุดในรูปแบบ Dataframe

3.6 ตัวปรับพารามิเตอร์ที่ใช้ในการจำแนกคุณภาพของงานเขียน

จากแนวคิดการประมวลผลการจำแนกคุณภาพของงานเขียนได้กล่าวในหัวข้อที่ 3.1 ผู้วิจัยได้นำแนวคิดมาออกแบบตัวปรับพารามิเตอร์เพื่อใช้ในการจำแนกคุณภาพของงานเขียนดังนี้

3.6.1 Optimizer

Optimizer ใช้เพื่อปรับพารามิเตอร์ของแบบจำลองเพื่อลด loss function นอกจากนี้ optimizer สามารถทำการ fine-tuned ได้ด้วยการตั้งค่า Hyperparameter ที่แตกต่างกันได้ และ optimizer สามารถทำงานร่วมกับเทคนิคอื่นๆเพื่อปรับปรุงประสิทธิภาพของโมเดลให้ดียิ่งขึ้นไปอีก

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

3.6.1.1 Adaptive Moment Estimation (Adam)

เป็นวิธีการปรับพารามิเตอร์ที่ได้รับความนิยม เนื่องจากเป็นวิธีที่รวมข้อดีและลดข้อเสียของตัวปรับพารามิเตอร์ AdaGrad และ RMSprop ออกไป นอกจากนี้ยังสามารถประมวลผลได้เร็วและลดปัญหาการแกว่งของพารามิเตอร์ได้อีกด้วย โดยตัวปรับพารามิเตอร์ Adam มีสูตรในการคำนวณดังรูปที่ 12 ดังนั้นจะเห็นได้ว่าตัวปรับพารามิเตอร์ Adam มีการคำนวณโดยใช้ค่าเกรเดียน และค่าเกรเดียนยกกำลังสอง ทำให้มีความแม่นยำและเป็นที่นิยมมากที่สุด

$$\nu_t = \rho\nu_{t-1} + (1-\rho) * g_t^2$$
$$\Delta\omega_t = -\frac{\eta}{\sqrt{\nu_t + \epsilon}} * g_t$$
$$\omega_{t+1} = \omega_t + \Delta\omega_t$$

$\eta : Initial\ Learning\ rate$
$\nu_t : Exponential\ Average\ of\ squares\ of\ gradients$
$g_t : Gradient\ at\ time\ t\ along\ \omega^j$

รูปที่ 12 : สูตรการคำนวณของตัวปรับพารามิเตอร์ Adam

3.6.1.2 Root Mean Square Propagation (RMSprop)

เป็นวิธีการปรับพารามิเตอร์ที่ทำให้อัตราการเรียนรู้ถูกปรับให้เหมาะสมกับพารามิเตอร์แต่ละตัว โดยจะเก็บค่าเกรเดียน (Gradient) ของการเรียนรู้ครั้งก่อนหน้ามาใช้ในการเรียนรู้ในครั้งต่อไป โดยใช้ค่า Mean Square Error ในการเก็บค่าเกรเดียน ตัวปรับพารามิเตอร์ RMSprop มีสูตรในการคำนวณดังรูปที่ 13 ดังนั้นจะเห็นได้ว่าตัวปรับพารามิเตอร์ RMSprop จะใช้คำนวณค่าเกรเดียนโดยใช้การยกกำลังสอง

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

$$\nu_t = \beta_1 * \nu_{t-1} - (1-\beta_1) * g_t$$
$$s_t = \beta_2 * s_{t-1} - (1-\beta_2) * g_t^2$$
$$\Delta\omega_t = -\eta\frac{\nu_t}{\sqrt{s_t + \epsilon}} * g_t$$
$$\omega_{t+1} = \omega_t + \Delta\omega_t$$

$\eta : Initial\ Learning\ rate$
$g_t : Gradient\ at\ time\ t\ along\ \omega^j$
$\nu_t : Exponential\ Average\ of\ gradients\ along\ \omega_j$
$s_t : Exponential\ Average\ of\ squares\ of\ gradients\ along\ \omega_j$
$\beta_1, \beta_2 : Hyperparameters$

รูปที่ 13 : สูตรการคำนวณของตัวปรับพารามิเตอร์ RMSprop

3.6.2 Loss Function

ใช้เพื่อวัดความแม่นยำของการทำนายผลลัพธ์ของแบบจำลอง โดยจะคำนวณความแตกต่างของแต่ละตัวอย่างการฝึกสอนระหว่างผลลัพธ์ที่ทำนายไว้และผลลัพธ์จริง ซึ่งเป้าหมายของแบบจำลองคือการลด loss function ให้เหลือน้อยที่สุด โดยการหาพารามิเตอร์ที่ดีที่สุดที่จะทำให้แบบจำลองมีประสิทธิภาพมากที่สุกที่สามารถทำนายผลลัพธ์ได้อย่างแม่นยำที่สุด

ในงานวิจัยนี้ ผู้วิจัยได้ตั้งสมมติฐานว่าตัวปรับพารามิเตอร์ Adam จะเหมาะกับการจำแนกคุณภาพของงานเขียนตามส่วนต่าง ๆ ของเรียงความในหัวข้อที่เกี่ยวกับการจำแนกประเภทการจัดหมวดหมู่ทั้งหมด 3 หัวข้อ ได้แก่ ความน่าสนใจของบทนำ, ความน่าสนใจของสรุปความ และองค์ประกอบของเรียงความ และตัวปรับพารามิเตอร์ RMSprop จะเหมาะกับการจำแนกคุณภาพของงานเขียนตามส่วนต่าง ๆ ของเรียงความในหัวข้อที่เกี่ยวการจำแนกประเภทถดถอยทั้งหมด 4 หัวข้อ ได้แก่ การเกริ่นนำเข้าเรื่องของบทนำ การเรียงลำดับความของเนื้อเรื่อง การแสดงความคิดเห็นของเนื้อเรื่อง และการขมวดความคิดของสรุปความ

3.7 การพัฒนาแบบจำลอง

แบบจำลองของงานวิจัยนี้จะแบ่งออกเป็น 4 แบบจำลอง โดยแบบจำลองแรก คือ แบบจำลองหน่วยความจำระยะสั้นแบบยาว (LSTM) ที่ใช้ฟีเจอร์ทางภาษาและถุงคำในการฝึกสอนแบบจำลอง แบบจำลองที่สอง คือ แบบจำลองโครงข่ายประสาทเทียมแบบสังวัตนาการ (CNN) ซึ่งจะใช้ฟีเจอร์ทางภาษาและถุงคำในการฝึกสอนแบบจำลองเช่นเดียวกับแบบจำลองหน่วยความจำระยะสั้นแบบยาว (LSTM) แบบจำลองที่สาม คือ แบบจำลองเบิร์ต (BERT) ซึ่งจะใช้ฟีเจอร์การฝังคำด้วยแบบจำลอง Sentencepiece อย่าง bert-base-uncased ซึ่งถูกพัฒนาบนสถาปัตยกรรมแบบทรานส์ฟอร์เมอร์ และแบบจำลองสุดท้าย คือ แบบจำลองวังจันทร์เบอร์ต้า (WangchanBERTa) ซึ่งจะใช้ฟีเจอร์การฝังคำของสถาปัตยกรรม WangchanBERTa เองซึ่งถูกพัฒนาบนสถาปัตยกรรมแบบทรานส์ฟอร์เมอร์เช่นเดียวกับแบบจำลองเบิร์ต (BERT) แต่ถูกฝึกสอนด้วยชุดข้อมูลภาษาไทยและได้วัดประสิทธิภาพของแบบจำลองภาษาที่ทำการปรับตั้งละเอียดแล้ว

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

3.8 เกณฑ์การวัดประสิทธิภาพของแบบจำลอง (Evaluation Metrics)

ในการวิจัยนี้จะเปรียบเทียบผลความถูกต้องของค่าเฉลี่ยความถูกต้อง และค่า Macro F1-Score จากการทำนายคุณภาพของงานเขียนของเนื้อหาส่วนต่าง ๆ อาทิ เช่น บทนำ เนื้อเรื่อง สรุปความ ในเรียงความภาษาไทย เพื่อวัดประสิทธิภาพของแบบจำลองแต่ละแบบและนำมาเปรียบเทียบกัน

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บทที่ 4
ผลการทดลอง

บทนี้จะกล่าวถึงผลการทดลอง โดยได้นำแนวคิดและแนวคิดการดำเนินงานในบทที่ 3 มาทดลอง งานวิจัยนี้ได้นำแบบจำลองหลายประเภท ได้แก่ แบบจำลองหน่วยความจำระยะสั้นแบบยาว (LSTM) แบบจำลองโครงข่ายประสาทเทียมแบบสังวัตนาการ (CNN) แบบจำลองเบิร์ต (BERT) และแบบจำลองวังจันทร์เบอร์ต้า (WangchanBERTa) มาทดลองเพื่อเปรียบเทียบประสิทธิภาพของแบบจำลองในการทำนายคุณภาพของงานเขียนของเนื้อหาส่วนต่าง ๆ อาทิ เช่น บทนำ เนื้อเรื่อง สรุปความ ในเรียงความภาษาไทย ซึ่งได้แบ่งผลการทดลองการทำนายคุณภาพของงานเขียนเรียงความภาษาไทยออกเป็น 8 หัวข้อการทดลอง ได้แก่ การจำแนกคุณภาพการเกริ่นเข้าเรื่องของบทนำ การจำแนกคุณภาพความน่าสนใจของบทนำ การจำแนกคุณภาพการเรียงลำดับความของเนื้อเรื่อง การจำแนกคุณภาพการแสดงความคิดเห็นของเนื้อเรื่อง การจำแนกคุณภาพการขมวดความของสรุปความ การจำแนกคุณภาพความน่าสนใจของสรุปความ การจำแนกส่วนของเรียงความ และการวิเคราะห์ โดยมีการกำหนดค่า 2 ชุดสำหรับการฝึกสอนแบบจำลอง โดยการกำหนดค่าชุดที่ 1 แบบจำลองจะถูกประมวลผลโดยใช้ตัวปรับพารามิเตอร์ Adam และ ใช้ Binary Crossentropy เป็น Loss Function และการกำหนดค่าชุดที่ 2 แบบจำลองจะถูกประมวลผลโดยใช้ตัวปรับพารามิเตอร์ RMSprop และ ใช้ Mean Squared Error เป็น Loss Function

4.1 การจำแนกคุณภาพการเกริ่นเข้าเรื่องของบทนำ

แบบจำลองการจำแนกคุณภาพการเกริ่นเข้าเรื่องของบทนำ แบ่งเป็น 3 คลาส ได้แก่ เรียงความที่มีการเกริ่นเข้าเรื่องได้ดี จำนวน 246 เรียงความ เรียงความที่มีการเกริ่นเข้าเรื่องได้ จำนวน 287 เรียงความ และเรียงความที่มีการเกริ่นเข้าเรื่องไม่ดี จำนวน 43 เรียงความ ดังตารางที่ 3 โดยผลการทดลองการจำแนกคุณภาพการเกริ่นเข้าเรื่องของบทนำ พบว่าแบบจำลองโครงข่ายประสาทเทียมแบบสังวัตนาการ (CNN) ที่ประมวลผลด้วยตัวปรับพารามิเตอร์ Adam และตัวปรับพารามิเตอร์ RMSprop และแบบจำลองเบิร์ต (BERT) ที่ประมวลผลด้วยตัวปรับพารามิเตอร์ Adam มีประสิทธิภาพที่ดีที่สุด โดยมีค่าความถูกต้องเท่ากันที่ 92.44% รองลงมาเป็นแบบจำลองวังจันทร์เบอร์ต้า (WangchanBERTa) ที่ประมวลผลด้วยตัวปรับพารามิเตอร์ Adam มีค่าความถูกต้องเท่ากับ 91.92% ผลลัพธ์ค่าความถูกต้อง ผลลัพธ์ค่าความแม่นยำ ผลลัพธ์ค่าความครบถ้วน และผลลัพธ์ค่าเอฟวันสกอร์ในการจำแนกคุณภาพการเกริ่นเข้าเรื่องของบทนำแสดงดังตารางที่ 6 - 9 ตามลำดับ

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

แบบจำลอง	LSTM	CNN	BERT	WangchanBERTa
Adam	38.37	**92.44**	**92.44**	91.28
RMSprop	59.88	**92.44**	36.63	36.63

ตารางที่ 6 : แสดงค่า Accuracy ผลการทดลองการจำแนกคุณภาพการเกริ่นเข้าเรื่องของบทนำ

แบบจำลอง	LSTM	CNN	BERT	WangchanBERTa
Adam	45.76	**92.69**	92.48	91.48
RMSprop	74.10	**92.56**	12.21	12.21

ตารางที่ 7 : แสดงค่า Precision ผลการทดลองการจำแนกคุณภาพการเกริ่นเข้าเรื่องของบทนำ

แบบจำลอง	LSTM	CNN	BERT	WangchanBERTa
Adam	35.19	**92.08**	92.59	91.62
RMSprop	57.91	**92.47**	33.33	33.33

ตารางที่ 8 : แสดงค่า Recall ผลการทดลองการจำแนกคุณภาพการเกริ่นเข้าเรื่องของบทนำ

แบบจำลอง	LSTM	CNN	BERT	WangchanBERTa
Adam	21.61	**92.26**	92.53	91.44
RMSprop	57.14	**92.51**	17.87	17.87

ตารางที่ 9 : แสดงค่า F1-Score ผลการทดลองการจำแนกคุณภาพการเกริ่นเข้าเรื่องของบทนำ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

4.2 การจำแนกคุณภาพความน่าสนใจของบทนำ

แบบจำลองการจำแนกคุณภาพความน่าสนใจของบทนำ แบ่งเป็น 5 คลาส ได้แก่ เรียงความที่มีบทนำน่าสนใจโดยการยกตัวอย่าง จำนวน 183 เรียงความ เรียงความที่มีบทนำน่าสนใจโดยการให้ข้อมูลเป็นเบื้องต้น จำนวน 137 เรียงความที่มีบทนำไม่มีความน่าสนใจ จำนวน 108 เรียงความ เรียงความที่มีบทนำน่าสนใจโดยการอ้างเหตุผล จำนวน 81 เรียงความ และเรียงความที่มีบทนำน่าสนใจด้วยเหตุผลอื่น ๆ จำนวน 67 เรียงความ ดังตารางที่ 3 โดยผลการทดลองการจำแนกคุณภาพความน่าสนใจของบทนำ พบว่าแบบจำลองโครงข่ายประสาทเทียมแบบสังวัตนาการ (CNN) ที่ประมวลผลด้วยตัวปรับพารามิเตอร์ RMSprop มีประสิทธิภาพที่ดีที่สุด โดยมีค่าความถูกต้องเท่ากับ 90.71% รองลงมาเป็นแบบจำลองวังจันทร์เบอร์ต้า (WangchanBERTa) ที่ประมวลผลด้วยตัวปรับ

พารามิเตอร์ Adam มีค่าความถูกต้องเท่ากับ 90.16% แบบจำลองโครงข่ายประสาทเทียมแบบสังวัตนาการ (CNN) ที่ประมวลผลด้วยตัวปรับพารามิเตอร์ Adam มีค่าความถูกต้องเท่ากับ 87.98% และแบบจำลองวังจันทร์เบอร์ต้า (WangchanBERTa) ที่ประมวลผลด้วยตัวปรับพารามิเตอร์ RMSprop มีค่าความถูกต้องเท่ากับ 86.89% ตามลำดับ ผลลัพธ์ค่าความถูกต้อง ผลลัพธ์ค่าความแม่นยำ ผลลัพธ์ค่าความครบถ้วน และผลลัพธ์ค่าเอฟวันสกอร์ในการจำแนกคุณภาพความน่าสนใจของบทนำแสดงดังตารางที่ 10 - 13 ตามลำดับ

แบบจำลอง	**LSTM**	**CNN**	**BERT**	**WangchanBERTa**
Adam	20.77	87.98	77.05	**90.16**
RMSprop	22.95	**90.71**	54.10	86.89

ตารางที่ 10 : แสดงค่า Accuracy ผลการทดลองการจำแนกคุณภาพความน่าสนใจของบทนำ

แบบจำลอง	**LSTM**	**CNN**	**BERT**	**WangchanBERTa**
Adam	4.15	88.73	77.52	**90.10**
RMSprop	14.53	**90.19**	62.90	87.81

ตารางที่ 11 : แสดงค่า Precision ผลการทดลองการจำแนกคุณภาพความน่าสนใจของบทนำ

แบบจำลอง	**LSTM**	**CNN**	**BERT**	**WangchanBERTa**
Adam	20.00	87.93	76.64	**90.15**
RMSprop	20.50	**90.01**	54.15	86.92

ตารางที่ 12 : แสดงค่า Recall ผลการทดลองการจำแนกคุณภาพความน่าสนใจของบทนำ

แบบจำลอง	**LSTM**	**CNN**	**BERT**	**WangchanBERTa**
Adam	8.34	88.03	75.27	**90.06**
RMSprop	8.34	**90.02**	52.33	87.01

ตารางที่ 13 : แสดงค่า F1-Score ผลการทดลองการจำแนกคุณภาพความน่าสนใจของบทนำ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

4.3 การจำแนกคุณภาพการเรียงลำดับความของเนื้อเรื่อง

แบบจำลองจำแนกการเรียงลำดับความของเนื้อเรื่อง แบ่งเป็น 3 คลาส ได้แก่ เรียงความที่เนื้อเรื่องมีความราบรื่น จำนวน 334เรียงความ เรียงความเนื้อเรื่องไม่ราบรื่นมีความสับสน 1 แห่ง จำนวน 193 เรียงความ และเรียงความเนื้อเรื่องไม่ราบรื่นมีความสับสน 2 แห่งขึ้นไป จำนวน 49 เรียงความ ดังตารางที่ 3 โดยผลการทดลองการจำแนกคุณภาพการเรียงลำดับความของเนื้อเรื่อง พบว่าแบบจำลองวังจันทร์เบอร์ต้า (WangchanBERTa) ที่ประมวลผลด้วยตัวปรับพารามิเตอร์ Adam มีค่าความถูกต้องเท่ากับ 97.00% ซึ่งเป็นแบบจำลองที่มีประสิทธิภาพดีที่สุด รองลงมาเป็นแบบจำลองโครงข่ายประสาทเทียมแบบสังวัตนาการ (CNN) ที่ประมวลผลด้วยตัวปรับพารามิเตอร์ Adam มีค่าความถูกต้องเท่ากับ 96.50% และประมวลผลด้วยตัวปรับพารามิเตอร์ RMSprop มีค่าความถูกต้องเท่ากับ 93.00% ตามลำดับ ผลลัพธ์ค่าความถูกต้อง ผลลัพธ์ค่าความแม่นยำ ผลลัพธ์ค่าความครบถ้วน และผลลัพธ์ค่าเอฟวันสกอร์ในการจำแนกคุณภาพการเรียงลำดับความของเนื้อเรื่อง แสดงดังตารางที่ 14 - 17 ตามลำดับ

แบบจำลอง	**LSTM**	**CNN**	**BERT**	**WangchanBERTa**
Adam	88.50	96.50	64.50	**97.00**
RMSprop	34.00	**93.00**	79.00	89.00

ตารางที่ 14 : แสดงค่า Accuracy ผลการทดลองการจำแนกการเรียงลำดับความของเนื้อเรื่อง

แบบจำลอง	**LSTM**	**CNN**	**BERT**	**WangchanBERTa**
Adam	88.42	96.56	64.49	**97.45**
RMSprop	11.33	**93.52**	80.66	88.74

ตารางที่ 15 : แสดงค่า Precision ผลการทดลองการจำแนกการเรียงลำดับความของเนื้อเรื่อง

แบบจำลอง	**LSTM**	**CNN**	**BERT**	**WangchanBERTa**
Adam	86.54	96.26	63.44	**96.42**
RMSprop	33.33	**93.71**	80.76	88.32

ตารางที่ 16 : แสดงค่า Recall ผลการทดลองการจำแนกการเรียงลำดับความของเนื้อเรื่อง

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

แบบจำลอง	LSTM	CNN	BERT	WangchanBERTa
Adam	87.12	96.31	63.81	**96.83**
RMSprop	16.92	**93.51**	78.18	88.50

ตารางที่ 17 : แสดงค่า F1-Score ผลการทดลองการจำแนกการเรียงลำดับความของเนื้อเรื่อง

4.4 การจำแนกคุณภาพการแสดงความคิดเห็นของเนื้อเรื่อง

แบบจำลองการจำแนกคุณภาพการแสดงความคิดเห็นของเนื้อเรื่อง แบ่งเป็น 4 คลาส ได้แก่ เรียงความที่มีการแสดงความคิดเห็นชัดเจนน่าเชื่อถือ จำนวน 169 เรียงความ เรียงความที่มีการแสดงความคิดเห็นบกพร่อง 1 แห่ง จำนวน 256 เรียงความ เรียงความที่มีการแสดงความคิดเห็นบกพร่อง 2 แห่ง จำนวน 123 เรียงความ และเรียงความที่มีการแสดงความคิดเห็นบกพร่อง 3 แห่งขึ้นไป จำนวน 28 เรียงความ ดังตารางที่ 3 โดยผลการทดลองการจำแนกคุณภาพการแสดงความคิดเห็นของเนื้อเรื่อง พบว่าแบบจำลองโครงข่ายประสาทเทียมแบบสังวัตนาการ (CNN) ที่ประมวลผลด้วยตัวปรับพารามิเตอร์ Adam มีประสิทธิภาพที่ดีที่สุด โดยมีค่าความถูกต้องเท่ากับ 95.61% รองลงมาเป็นแบบจำลองโครงข่ายประสาทเทียมแบบสังวัตนาการ (CNN) ที่ประมวลผลด้วยตัวปรับพารามิเตอร์ RMSprop ค่าความถูกต้องเท่ากับ 91.71% และแบบจำลองวังจันทร์เบอร์ต้า (WangchanBERTa) ที่ประมวลผลด้วยตัวปรับพารามิเตอร์ Adam มีค่าความถูกต้องเท่ากับ 90.24% ผลลัพธ์ค่าความถูกต้อง ผลลัพธ์ค่าความแม่นยำ ผลลัพธ์ค่าความครบถ้วน และผลลัพธ์ค่าเอฟวันสกอร์ในการจำแนกคุณภาพการแสดงความคิดเห็นของเนื้อเรื่องแสดงดังตารางที่ 18 - 21 ตามลำดับ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

แบบจำลอง	LSTM	CNN	BERT	WangchanBERTa
Adam	44.39	**95.61**	71.71	90.24
RMSprop	26.83	**91.71**	81.46	51.22

ตารางที่ 18 : แสดงค่า Accuracy ผลการทดลองการจำแนกการแสดงความคิดเห็นของเนื้อเรื่อง

แบบจำลอง	LSTM	CNN	BERT	WangchanBERTa
Adam	46.97	**96.37**	72.71	90.71
RMSprop	6.71	**91.66**	84.60	53.63

ตารางที่ 19 : แสดงค่า Precision ผลการทดลองการจำแนกการแสดงความคิดเห็นของเนื้อเรื่อง

แบบจำลอง	LSTM	CNN	BERT	WangchanBERTa
Adam	42.65	**95.05**	73.00	90.87
RMSprop	25.00	**91.45**	82.57	53.00

ตารางที่ 20 : แสดงค่า Recall ผลการทดลองการจำแนกการแสดงความคิดเห็นของเนื้อเรื่อง

แบบจำลอง	LSTM	CNN	BERT	WangchanBERTa
Adam	29.85	**95.38**	71.87	90.55
RMSprop	10.58	**91.54**	82.24	49.81

ตารางที่ 21 : แสดงค่า F1-Score ผลการทดลองการจำแนกการแสดงความคิดเห็นของเนื้อเรื่อง

4.5 การจำแนกคุณภาพการขมวดความของสรุปความ

แบบจำลองการจำแนกคุณภาพการขมวดความของสรุปความ แบ่งเป็น 4 คลาส ได้แก่ เรียงความที่มีการขมวดความคิดได้ดี จำนวน 242 เรียงความ เรียงความที่มีการขมวดความคิดได้ จำนวน 265 เรียงความ เรียงความที่มีการขมวดความคิดได้บ้าง จำนวน 50 เรียงความ และเรียงความที่มีขมวดความคิดไม่ได้ จำนวน 15 เรียงความ ดังตารางที่ 3 โดยผลการทดลองการจำแนกคุณภาพการขมวดความของสรุปความ พบว่าแบบจำลองวังจันทร์เบอร์ต้า (WangchanBERTa) ที่ประมวลผลด้วยตัวปรับพารามิเตอร์ Adam และแบบจำลองโครงข่ายประสาทเทียมแบบสังวัตนาการ (CNN) ที่ประมวลผลด้วยตัวปรับพารามิเตอร์ RMSprop มีประสิทธิภาพที่ดีที่สุด โดยมีค่าความถูกต้องเท่ากันคือ 93.40% รองลงมาคือแบบจำลองโครงข่ายประสาทเทียมแบบสังวัตนาการ (CNN) แบบจำลองหน่วยความจำระยะสั้นแบบยาว (LSTM) และแบบจำลองเบิร์ต (BERT) ที่ประมวลผลด้วยตัวปรับพารามิเตอร์ Adam โดยมีค่าความถูกต้องเท่ากับ 92.45%, 91.98% และ 91.04% ตามลำดับ ผลลัพธ์ค่าความแม่นยำ ผลลัพธ์ค่าความครบถ้วน และผลลัพธ์ค่าเอฟวันสกอร์ในการจำแนกคุณภาพการขมวดความของสรุปความแสดงดังตารางที่ 22 - 25 ตามลำดับ

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

แบบจำลอง	LSTM	CNN	BERT	WangchanBERTa
Adam	91.98	92.45	91.04	**93.40**
RMSprop	84.43	**93.40**	44.81	31.60

ตารางที่ 22 : แสดงค่า Accuracy ผลการทดลองการจำแนกการขมวดความของสรุปความ

แบบจำลอง	LSTM	CNN	BERT	WangchanBERTa
Adam	92.33	92.83	91.19	**93.61**
RMSprop	85.13	**93.08**	35.49	58.90

ตารางที่ 23 : แสดงค่า Precision ผลการทดลองการจำแนกการขมวดความของสรุปความ

แบบจำลอง	LSTM	CNN	BERT	WangchanBERTa
Adam	91.97	92.99	91.28	**93.57**
RMSprop	84.91	**93.06**	44.86	31.25

ตารางที่ 24 : แสดงค่า Recall ผลการทดลองการจำแนกการขมวดความของสรุปความ

แบบจำลอง	LSTM	CNN	BERT	WangchanBERTa
Adam	92.01	92.84	91.22	**93.57**
RMSprop	84.17	**93.04**	36.90	22.42

ตารางที่ 25 : แสดงค่า F1-Score ผลการทดลองการจำแนกการขมวดความของสรุปความ

4.6 การจำแนกคุณภาพความน่าสนใจของสรุปความ

แบบจำลองการจำแนกคุณภาพความน่าสนใจของสรุปความ แบ่งเป็น 5 คลาส ได้แก่ เรียงความที่มีความน่าสนใจของสรุปความโดยการยกคำคมสุภาษิต จำนวน 281 เรียงความ เรียงความที่มีความน่าสนใจของสรุปความโดยการถามให้คิด จำนวน 37เรียงความ เรียงความที่มีความน่าสนใจของสรุปความโดยการโยงไปเรื่องอื่น จำนวน 55 เรียงความ เรียงความที่ไม่มีความน่าสนใจของสรุปความ จำนวน 17 เรียงความ และเรียงความที่มีความน่าสนใจของสรุปความด้วยเหตุผลอื่น ๆ จำนวน 182 เรียงความ ดังตารางที่ 3 โดยผลการทดลองการจำแนกคุณภาพความน่าสนใจของสรุปความ พบว่าแบบจำลองวังจันทร์เบอร์ต้า (WangchanBERTa) ที่ประมวลผลด้วยตัวปรับพารามิเตอร์ Adam และแบบจำลองโครงข่ายประสาทเทียมแบบสังวัตนาการ (CNN) ที่ประมวลผลด้วยตัวปรับพารามิเตอร์ RMSprop มีประสิทธิภาพที่ดีที่สุด โดยมีค่าความถูกต้องเท่ากัน คือ 97.15% รองลงมาคือแบบจำลองโครงข่ายประสาทเทียมแบบสังวัตนาการ (CNN) และแบบจำลองหน่วยความจำระยะสั้นแบบยาว (LSTM) ที่ประมวลผลด้วยตัวปรับพารามิเตอร์ Adam โดยมีค่าความถูกต้องเท่ากับ 96.80% และ 95.37% ตามลำดับ ผลลัพธ์ค่าความแม่นยำ ผลลัพธ์ค่าความครบถ้วน และผลลัพธ์ค่าเอฟวันสกอร์ในการจำแนกคุณภาพความน่าสนใจของสรุปความแสดงดังตารางที่ 26 - 29 ตามลำดับ

แบบจำลอง	LSTM	CNN	BERT	WangchanBERTa
Adam	95.37	96.80	59.43	**97.15**
RMSprop	90.39	**97.15**	57.30	77.94

ตารางที่ 26 : แสดงค่า Accuracy ผลการทดลองการจำแนกความน่าสนใจของสรุปความ

แบบจำลอง	LSTM	CNN	BERT	WangchanBERTa
Adam	95.39	96.81	62.30	**97.02**
RMSprop	90.55	**97.26**	73.57	78.22

ตารางที่ 27 : แสดงค่า Precision ผลการทดลองการจำแนกความน่าสนใจของสรุปความ

แบบจำลอง	LSTM	CNN	BERT	WangchanBERTa
Adam	95.46	96.82	59.23	**96.89**
RMSprop	90.78	**97.38**	57.33	76.54

ตารางที่ 28 : แสดงค่า Recall ผลการทดลองการจำแนกความน่าสนใจของสรุปความ

แบบจำลอง	LSTM	CNN	BERT	WangchanBERTa
Adam	95.39	96.81	59.82	**96.95**
RMSprop	90.39	**97.31**	57.59	75.68

ตารางที่ 29 : แสดงค่า F1-Score ผลการทดลองการจำแนกความน่าสนใจของสรุปความ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

4.7 การจำแนกส่วนของเรียงความ

แบบจำลองการจำแนกส่วนของเรียงความ แบ่งเป็น 3 คลาส ได้แก่ บทนำ เนื้อเรื่อง และสรุปความ จำนวน 576 ชุดข้อมูล โดยผลการทดลองการจำแนกส่วนของเรียงความ พบว่าแบบจำลองโครงข่ายประสาทเทียมแบบสังวัตนาการ (CNN) ที่ประมวลผลด้วยตัวปรับพารามิเตอร์ Adam โดยมีค่าความถูกต้องเท่ากันคือ 98.55% ซึ่งเป็นแบบจำลองที่มีประสิทธิภาพดีที่สุด รองลงมาคือแบบจำลองโครงข่ายประสาทเทียมแบบสังวัตนาการ (CNN) ที่ประมวลผลด้วยตัวปรับพารามิเตอร์ RMSprop โดยมีค่าความถูกต้องเท่ากับ 98.27% และแบบจำลองวังจันทร์เบอร์ต้า (WangchanBERTa) ที่ประมวลผลด้วยตัวปรับพารามิเตอร์ Adam มีค่าความถูกต้องเท่ากับ 97.98%

ตามลำดับ ผลลัพธ์ค่าความแม่นยำ ผลลัพธ์ค่าความครบถ้วน และผลลัพธ์ค่าเอฟวันสกอร์ในการจำแนกส่วนของเรียงความแสดงดังตารางที่ 30 - 33 ตามลำดับ

แบบจำลอง	**LSTM**	**CNN**	**BERT**	**WangchanBERTa**
Adam	55.78	**98.55**	93.06	97.98
RMSprop	34.39	**98.27**	33.24	33.24

ตารางที่ 30 : แสดงค่า Accuracy ผลการทดลองการจำแนกส่วนของเรียงความ

แบบจำลอง	**LSTM**	**CNN**	**BERT**	**WangchanBERTa**
Adam	37.26	**98.65**	93.06	97.98
RMSprop	11.46	**98.24**	11.08	11.08

ตารางที่ 31 : แสดงค่า Precision ผลการทดลองการจำแนกส่วนของเรียงความ

แบบจำลอง	**LSTM**	**CNN**	**BERT**	**WangchanBERTa**
Adam	35.19	**92.08**	92.59	91.62
RMSprop	33.33	**98.26**	33.33	33.33

ตารางที่ 32 : แสดงค่า Recall ผลการทดลองการจำแนกส่วนของเรียงความ

แบบจำลอง	**LSTM**	**CNN**	**BERT**	**WangchanBERTa**
Adam	44.72	**98.61**	93.07	97.98
RMSprop	17.06	**98.25**	16.63	16.63

ตารางที่ 33 : แสดงค่า F1-Score ผลการทดลองการจำแนกส่วนของเรียงความ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

4.8 ความสัมพันธ์ระหว่างกับความยาวของประโยคกับประสิทธิภาพของแบบจำลอง

เพื่อวิเคราะห์ประสิทธิภาพในการทำนายคุณภาพงานเขียนว่ามีความสัมพันธ์กับความยาวของประโยคที่ใช้ในการฝึกสอนแบบจำลองอย่างไร ผู้วิจัยจึงได้เลือกหัวข้อ การจำแนกการเรียงลำดับความของเนื้อเรื่องมาใช้ในการทดสอบ และกำหนดความยาวของประโยตในการฝึกสอนให้แบบจำลองเป็น 100 คำ 300 คำ 500 คำ และ 600 คำ แต่แบบจำลองเบิร์ต (BERT) และแบบจำลองวังจันทร์เบอร์ต้า (WangchanBERTa) ไม่สามารถทดลองการประมวลผลข้อมูลฝึกสอนที่มีความยาว 600 คำ

ได้ เนื่องจากมีข้อจำกัดในการประมวลผลข้อมูลอยู่ที่ 512 คำเท่านั้น จากผลการทดลองแบบจำลองหน่วยความจำระยะสั้นแบบยาว (LSTM) ที่ใช้ข้อมูลในการฝึกสอน 300 คำ มีประสิทธิภาพการจำแนกการเรียงลำดับความของเนื้อเรื่องสูงที่สุด โดยมีค่า Accuracy เท่ากับร้อยละ 94.00 แต่เมื่อใช้ข้อมูลในการฝึกสอน 600 คำ จะมีค่า Accuracy ลงลดเหลือเพียงร้อยละ 35.50 ในขณะที่ความยาวของข้อมูลในการฝึกสอนไม่ค่อยมีผลกระทบต่อประสิทธิภาพการจำแนกการเรียงลำดับความของเนื้อเรื่องของแบบจำลองอื่นๆ ได้แก่ แบบจำลองโครงข่ายประสาทเทียมแบบสังวัตนาการ (CNN) แบบจำลองเบิร์ต (BERT) และแบบจำลองวังจันทร์เบอร์ต้า (WangchanBERTa) จึงสรุปได้ว่าความยาวของประโยคที่ใช้ในการฝึกสอนแบบจำลองหน่วยความจำระยะสั้นแบบยาว (LSTM) ส่งผลต่อประสิทธิภาพการทำนายคุณภาพงานเขียนมาก ยิ่งใช้ความยาวของข้อมูลในการฝึกสอนแบบจำลองมาก ประสิทธิภาพยิ่งลดลง โดยผลการทดลองแสดงดังตารางที่ 34

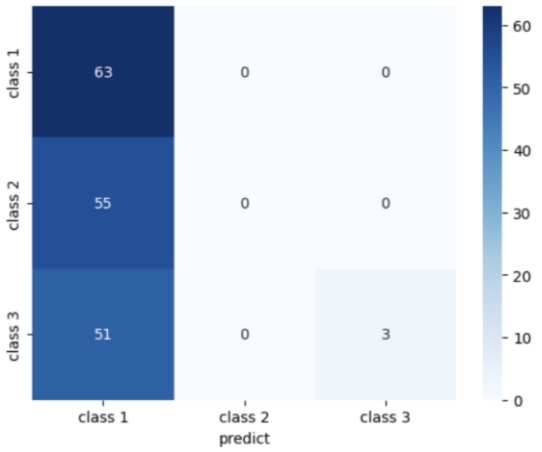
แบบจำลอง	**ความยาว 100 คำ**	**ความยาว 300 คำ**	**ความยาว 500 คำ**	**ความยาว 600 คำ**
LSTM	92.00	**94.00**	88.50	35.50
CNN	96.00	**96.50**	**96.50**	95.50
BERT	**76.00**	71.00	64.50	-
WangchanBERTa	94.00	**97.00**	**97.00**	-

ตารางที่ 34 : แสดงค่า Accuracy การจำแนกการเรียงลำดับความของเนื้อเรื่องที่ความยาวต่างกัน

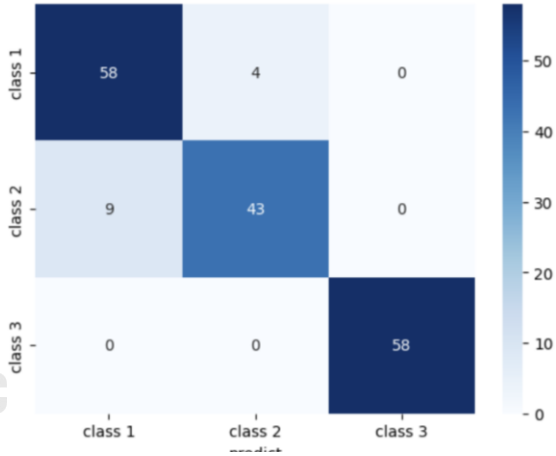
จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

จากตารางที่ 14 - 17 ผลการทดลองใช้แบบจำลองหน่วยความจำระยะสั้นแบบยาว (LSTM) ประมวลผลด้วย Adam Optimizer ในการจำแนกคุณภาพการเกริ่นเข้าเรื่องของบทนำ มีค่า Accuracy เท่ากับร้อยละ 38.37 ค่า Precision เท่ากับร้อยละ 45.76 ค่า Recall เท่ากับร้อยละ 35.19 และมีค่า F1-Score เท่ากับร้อยละ 21.61 โดยผลการทำนายการเกริ่นนำเข้าเรื่องส่วนใหญ่จะเป็นคลาสที่ 1 หรือเรียงความที่มีการเกริ่นเข้าเรื่องได้ดี เนื่องจากเป็นชุดข้อมูลเรียงความที่มีจำนวนต่อคลาสมากที่สุดในการฝึกสอนแบบจำลอง ดังรูปที่ 14 ในขณะที่แบบจำลองโครงข่ายประสาทเทียมแบบสังวัตนาการ (CNN) ซึ่งเป็นแบบจำลองที่มีประสิทธิภาพที่สุดในการจำแนกคุณภาพการเกริ่นเข้าเรื่องของบทนำ มีค่า Accuracy เท่ากับร้อยละ 92.44 ค่า Precision เท่ากับร้อยละ 92.69 ค่า Recall เท่ากับร้อยละ 92.08 และมีค่า F1-Score เท่ากับร้อยละ 92.26 โดย Confusion Matrix แสดงดังรูปที่ 15 จะเห็นได้ว่าแบบจำลองสามารถทำนายคุณภาพการเกริ่นเข้าเรื่องของบทนำได้ตรง

กับคลาสจริงเป็นส่วนมาก แต่จะมีเรียงความส่วนน้อยที่แบบจำลองทำนายผิดเป็นคลาสที่ 1 ที่เรียงความมีการเกริ่นเข้าเรื่องได้ดี



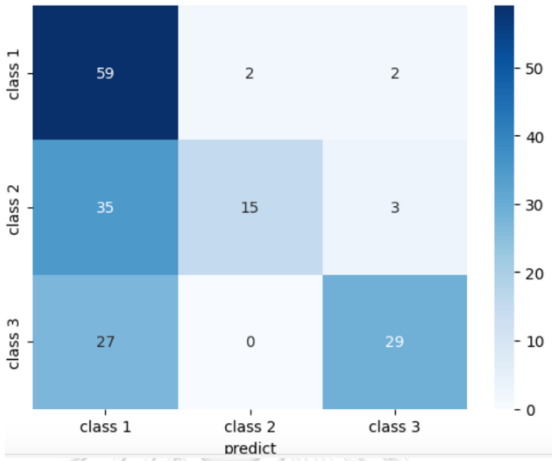
รูปที่ 14 : Confusion Matrix จำแนกการเกริ่นนำเข้าเรื่องของ LSTM กับ Adam



รูปที่ 15 : Confusion Matrix จำแนกการเกริ่นนำเข้าเรื่องของ CNN กับ Adam

นอกจากนี้ผู้วิจัยได้ทดลองใช้แบบจำลองหน่วยความจำระยะสั้นแบบยาว (LSTM) ประมวลผลด้วย RMSprop Optimizer ในการจำแนกคุณภาพการเกริ่นเข้าเรื่องของบทนำ โดยมีผลการทดลอง ดังนี้ ค่า Accuracy เท่ากับร้อยละ 59.88 ค่า Precision เท่ากับร้อยละ 74.10 ค่า Recall เท่ากับร้อยละ 57.91 และมีค่า F1-Score เท่ากับร้อยละ 57.14 ดังตารางที่ 14 – 17 และเมื่อพิจารณาจาก Confusion Matrix ดังรูปที่ 16 จะเห็นว่าในตัวอย่างข้อมูลเรียงความที่แบบจำลองยังทำนายผิดนั้น ส่วนใหญ่จะเป็นการทำนายว่าเรียงความดังกล่าวเป็นเรียงความมีการเกริ่นเข้าเรื่องได้

ดี ซึ่งเป็นคลาสที่มีตัวอย่างข้อมูลมากที่สุด แต่ในความเป็นจริงข้อมูลเรียงความเหล่านั้นเป็นชุดข้อมูลเรียงความมีการเกริ่นเข้าเรื่องได้หรือชุดข้อมูลเรียงความมีการเกริ่นเข้าเรื่องไม่ได้ แบบจำลองนี้จึงยังไม่มีประสิทธิภาพเหมาะสมสำหรับการไปใช้งานจริง



รูปที่ 16 : Confusion Matrix จำแนกการเกริ่นนำเข้าเรื่องของ LSTM กับ RMSprop

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บทที่ 5
สรุปผลและอภิปรายผลการทดลอง

ในบทนี้จะกล่าวถึงการสรุปผลการทดลองการทำนายคุณภาพงานเขียนของเนื้อหาส่วนต่าง ๆ เช่น บทนำ เนื้อเรื่อง สรุปความ ในเรียงความภาษาไทยด้วยแบบจำลองแบบต่าง ๆ ในบทที่ 4 และนำมาอภิปรายผลการทดลอง รวมถึงวิเคราะห์ปัญหาและอุปสรรคที่เกิดระหว่างทำการวิจัย แนวทางการวิจัยในขั้นถัดไป และผลงานวิจัยที่ได้รับการตีพิมพ์

5.1 สรุปผลและอภิปรายผลการทดลอง

งานวิจัยนี้มีจุดประสงค์เพื่อคุณภาพงานเขียนของเนื้อหาส่วนต่าง ๆ เช่น บทนำ เนื้อเรื่อง สรุปความ ในเรียงความภาษาไทย เพื่อช่วยเพิ่มประสิทธิภาพและลดเวลาที่ใช้ในการตรวจสอบคุณภาพงานเขียนเรียงความให้กับผู้ตรวจเรียงความ โดยข้อมูลเรียงความภาษาไทยที่ใช้ในงานวิจัยนี้สามารถแบ่งหัวข้อเกณฑ์การพิจารณาคุณภาพของงานเขียนของเรียงความออกเป็น 7 หัวข้อ โดยเป็นแบ่งออกเป็นหัวข้อการจำแนกประเภทการจัดหมวดหมู่ 3 หัวข้อ ได้แก่ ความน่าสนใจของบทนำ ความน่าสนใจของสรุปความ และองค์ประกอบของความเรียง และหัวข้อการจำแนกประเภทถดถอยทั้งหมด 4 หัวข้อ ได้แก่ การเกริ่นนำเข้าเรื่องของบทนำ การเรียงลำดับความของเนื้อเรื่อง การแสดงความคิดเห็นของเนื้อเรื่อง และการขมวดความคิดของสรุปความ

งานวิจัยนี้แสดงให้เห็นว่าแบบจำลองใดมีประสิทธิภาพสูงที่สุดในการทำนายคุณภาพงานเขียนในเรียงความภาษาไทย โดยได้ทำการเสนอแบบจำลองทั้งหมด 4 ชนิด ได้แก่ แบบจำลองหน่วยความจำระยะสั้นแบบยาว (LSTM) แบบจำลองโครงข่ายประสาทเทียมแบบสังวัตนาการ (CNN) แบบจำลองเบิร์ต (BERT) และแบบจำลองวังจันทร์เบอร์ต้า (WangchanBERTa) ร่วมกับตัวปรับพารามิเตอร์ทั้งหมด 2 ประเภท ได้แก่ Adaptive Moment Estimation (Adam) และ Root Mean Square Propagation (RMSProp) เพื่อใช้ประมวลผลแบบจำลองการจำแนกคุณภาพของงานเขียน

ผลการทดลองแสดงให้เห็นว่าแบบจำลองวังจันทร์เบอร์ต้า (WangchanBERTa) ที่ประมวลผลโดยใช้ตัวปรับพารามิเตอร์ Adam และ ใช้ Binary Crossentropy เป็น Loss Function เป็นแบบจำลองที่มีประสิทธิภาพที่สุดเมื่อนำไปใช้จำแนกคุณภาพงานเขียนในหัวข้อ ดังนี้ การจำแนกคุณภาพความน่าสนใจของบทนำ การจำแนกคุณภาพการเรียงลำดับความของเนื้อเรื่อง การจำแนก

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

คุณภาพการขมวดความของสรุปความ และการจำแนกคุณภาพความน่าสนใจของสรุปความโดยมีค่าความถูกต้องเป็น 90.16%, 97.00%, 93.40% และ 97.15% ตามลำดับ ส่วนแบบจำลองโครงข่ายประสาทเทียมแบบสังวัตนาการ (CNN) ที่ประมวลผลโดยใช้ตัวปรับพารามิเตอร์ Adam และ ใช้ Binary Crossentropy เป็น Loss Function เป็นแบบจำลองที่มีประสิทธิภาพที่สุดเมื่อนำไปใช้จำแนกคุณภาพงานเขียนในหัวข้อ ดังนี้ การจำแนกคุณภาพการเกริ่นเข้าเรื่องของบทนำ การจำแนกคุณภาพการแสดงความคิดเห็นของเนื้อเรื่อง และการจำแนกส่วนของเรียงความ โดยมีค่าความถูกต้องเป็น 92.44%, 95.61%, และ 98.55% ตามลำดับ นอกจากนี้แบบจำลองโครงข่ายประสาทเทียมแบบสังวัตนาการ (CNN) ที่ประมวลผลโดยใช้ตัวปรับพารามิเตอร์ RMSprop และ ใช้ Mean Squared Error เป็น Loss Function เป็นแบบจำลองที่มีประสิทธิภาพที่สุดเมื่อนำไปใช้จำแนกคุณภาพงานเขียนทั้ง 7 หัวข้อ โดยมีค่าความถูกต้องเป็น 92.44%, 90.71%, 93.00%, 91.71%, 93.40%, 97.15% และ 98.27% ตามลำดับ โดยตารางที่ 35 แสดงค่า Accuracy ของแบบจำลองที่ใช้จำแนกคุณภาพงานเขียนหัวข้อต่าง ๆ ที่มีประสิทธิภาพที่สุดของเมื่อประมวลผลด้วยตัวปรับพารามิเตอร์ทั้ง 2 ประเภท

จึงสรุปได้ว่าแบบจำลองโครงข่ายประสาทเทียมแบบสังวัตนาการ (CNN) มีประสิทธิภาพสูงในการตรวจสอบคุณภาพของเรียงความภาษาไทย เมื่อประมวลผลด้วยตัวปรับพารามิเตอร์แบบ Adam และ RMSprop โดยสามารถแก้ไขปัญหาได้ดีทั้งการจำแนกประเภทการจัดหมวดหมู่และการจำแนกประเภทถดถอย ในทางกลับกัน แบบจำลองวังจันทร์เบอร์ต้า (WangchanBERTa) มีประสิทธิภาพสูงในการตรวจสอบคุณภาพของเรียงความภาษาไทยเมื่อประมวลผลด้วยตัวปรับพารามิเตอร์แบบ Adam แต่ไม่เหมาะสมกับการประมวลผลด้วยตัวปรับพารามิเตอร์ RMSprop เนื่องจากทำให้มีประสิทธิภาพในการจำแนกคุณภาพของเรียงความภาษาไทยต่ำ

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

5.2 ปัญหาและอุปสรรค

ข้อมูลที่นำมาใช้ในงานวิจัยนี้เป็นข้อมูลเรียงความของนิสิตที่ลงทะเบียนเรียนในหลักสูตรการเขียนย่อหน้าจากสถาบันภาษาไทยสิรินธร จุฬาลงกรณ์มหาวิทยาลัย ซึ่งมีระยะเวลาในการเก็บรวมรวบเพียง 1 ภาคการศึกษา ทำให้มีจำนวนข้อมูลเรียงความน้อยและมีความไม่สมดุลของชุดข้อมูลสูง ดังนั้นต้องมีการจัดการกับข้อมูลที่ไม่สมดุลก่อนจะส่งให้แบบจำลองประมวลผลต่อไปและควรมีการเก็บข้อมูลเพิ่มขึ้นในภาคการศึกษาต่อไป

หัวข้อ	Adam	RMSprop
การจำแนกคุณภาพการเกริ่นเข้าเรื่องของบทนำ	**CNN / BERT 92.44%**	**CNN 92.44%**
การจำแนกคุณภาพความน่าสนใจของบทนำ	WangchanBERTa 90.16%	**CNN 90.71%**
การจำแนกคุณภาพการเรียงลำดับความของเนื้อเรื่อง	**WangchanBERTa 97.00%**	CNN 93.00%
การจำแนกคุณภาพการแสดงความคิดเห็นของเนื้อเรื่อง	**CNN 95.61%**	CNN 91.71%
การจำแนกคุณภาพการขมวดความของสรุปความ	**WangchanBERTa 93.40%**	**CNN 93.40%**
การจำแนกคุณภาพความน่าสนใจของสรุปความ	**WangchanBERTa 97.15%**	**CNN 97.15%**
การจำแนกส่วนของเรียงความ	**CNN 98.55%**	CNN 98.27%

ตารางที่ 35 : แสดงค่า Accuracy ผลการจำแนกคุณภาพของแบบจำลองที่มีประสิทธิภาพที่สุด

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

5.3 แนวทางการวิจัยในขั้นถัดไป

1. นำแนวคิดการทำนายคุณภาพงานเขียนของเนื้อหาส่วนต่าง ๆ เช่น บทนำ เนื้อเรื่อง สรุปความ ในเรียงความภาษาไทยที่นำเสนอไปประยุกต์ใช้กับชุดข้อมูลเรียงความในหัวข้อเรียงความอื่น ๆ

2. นำแนวคิดการทำนายคุณภาพงานเขียนของเนื้อหาส่วนต่าง ๆ เช่น บทนำ เนื้อเรื่อง สรุปความ ในเรียงความภาษาไทยไปพัฒนาแบบจำลองประเภทอื่น ๆ ต่อไป

3. จากการศึกษาทฤษฎีและงานวิจัยที่เกี่ยวข้องพบว่า การตัดคำในภาษาไทยมีหลายวิธี โดยในงานวิจัยนี้เลือกใช้ไลบรารี่ PyThaiNLP สำหรับแบบจำลอง LSTM และแบบจำลอง CNN

ร่วมกับการใช้แบบจำลองการแบ่งคำอย่าง Sentencepiece สำหรับแบบจำลอง BERT และแบบจำลอง WangchanBERTa แต่ยังมีวิธีการตัดคำอื่น ๆ เช่น การตัดคำโดยใช้กฎไวยากรณ์ทางภาษา (Rule base) และการตัดคำแบบตรงมากที่สุด (Maximum Matching) ที่สามารถนำมาประยุกต์ใช้เพื่อปรับปรุงความถูกต้องของการตัดคำในภาษาไทยได้เช่นกัน ซึ่งอาจส่งผลให้การทำนายคุณภาพงานเขียนของเนื้อหาส่วนต่าง ๆ เช่น บทนำ เนื้อเรื่อง สรุปความ ในเรียงความภาษาไทยมีประสิทธิภาพเพิ่มขึ้นได้

4. นำแนวคิดการพัฒนาแบบจำลองที่มีประสิทธิภาพสูงสุดในการทำนายคุณภาพงานเขียนของเนื้อหาส่วนต่าง ๆ เช่น บทนำ เนื้อเรื่อง สรุปความ ในเรียงความภาษาไทยที่นำเสนอไปประยุกต์ใช้และพัฒนาเป็นเว็บแอปพลิเคชันที่สามารถกรอกข้อมูลเรียงความผ่านทางหน้าจอและตรวจสอบคุณภาพงานเขียนได้ทันทีเมื่อกดส่งเรียงความ

5.4 ผลงานวิจัยที่ได้รับการตีพิมพ์

ผลงานวิจัยนี้ได้รับการคัดเลือกและตีพิมพ์เป็นบทความวิชาการเรื่อง “A Comparison of Machine Learning and Neural Network Algorithms for an Automated Thai Essay Quality Checking” โดยมี นางสาวณิชาพรรณ น้อยอยู่ และ ดร.เจษฎา ธัชแก้วกรพินธุ์ เป็นผู้เขียน และได้นำผลงานไปเสนอในงานประชุมวิชาการ “The 20th International Joint Conference on Computer Science and Software Engineering (JCSSE 2023)” ซึ่งจัดขึ้นระหว่างวันที่ 28 มิถุนายน – 1 กรกฎาคม พ.ศ. 2566 ที่มหาวิทยาลัยนเรศวร จังหวัดพิษณุโลก ประเทศไทย

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บรรณานุกรม

[1] A. Vaswani *et al.*, "Attention is all you need," *Advances in neural information processing systems,* vol. 30, 2017.

[2] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," *arXiv preprint arXiv:1810.04805,* 2018.

[3] W. Zhu and Y. Sun, "Automated essay scoring system using multi-model machine learning," *MLNLP, BDIOT, ITCCMA, CSITY, DTMN, AIFZ, SIGPRO,* 2020.

[4] C. Ormerod, A. Malhotra, and A. Jafari, *Automated essay scoring using efficient transformer-based language models*. 2021.

[5] P. Harnmetta and T. Samanchuen, "Sentiment Analysis of Thai Stock Reviews Using Transformer Models," in 2022 19th International Joint Conference on Computer Science and Software Engineering (JCSSE), 2022: IEEE, pp. 1-6.

[6] P. Gatchalee, S. Waijanya and P. Promrit, "Thai text classification experiment using cnn and transformer models for timely-timeless content marketing," ICIC Express Letters, vol. 19, pp. 91-101, 2023.

[7] T. Kudo and J. Richardson, "Sentencepiece: A simple and language independent subword tokenizer and detokenizer for neural text processing," *arXiv preprint arXiv:1808.06226,* 2018.

[8] L. Lowphansirikul, C. Polpanumas, N. Jantrakulchai, and S. Nutanong, "Wangchanberta: Pretraining transformer-based thai language models," *arXiv preprint arXiv:2101.09635,* 2021.

[9] Yinhan Liu, Myle Ott, Naman Goyal, et al., “Roberta: a robustly optimized bert pretraining approach,” *arXiv preprint arXiv: 1907.11692*, 2019.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ประวัติผู้เขียน

ชื่อ-สกุล	นางสาว ณิชาพรรณ น้อยอยู่
วัน เดือน ปี เกิด	12 มีนาคม พ.ศ. 2541
สถานที่เกิด	กรุงเทพมหานคร, ประเทศไทย
วุฒิการศึกษา	วิศวกรรมศาสตรบัณฑิต (วิศวกรรมอิเล็กทรอนิกส์), มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, ประเทศไทย
ที่อยู่ปัจจุบัน	99/7 หมู่.5 ตำบลเสาธงหิน, อำเภอบางใหญ่, จังหวัดนนทบุรี 11140
ผลงานตีพิมพ์	The 20th International Joint Conference on Computer Science and Software Engineering

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University