บทที่ 4

ผลการทดลอง

ในบทนี้จะกล่าวถึงการทดลองและผลการทดลอง ในการนำโมเดลสำหรับการจำแนกความรู้สึก (Sentiment Classification Model) ที่ได้จากขั้นตอนการดำเนินงาน มาทำการทดลองเพื่อจำแนกบทวิจารณ์โรงแรมที่ต้องการตรวจสอบ

4.1 ข้อมูลที่ใช้ในการทดสอบ

ข้อมูลที่ใช้การในทดสอบการสร้างโมเดลสำหรับการจำแนกความรู้สึกบทวิจารณ์โรงแรมนั้น จะเป็นชุดข้อมูลทดสอบ (Test set) ที่มีจำนวนทั้งสิ้น 9024 ข้อความ ที่เก็บอยู่ในรูปแบบของไฟล์ .csv ดังภาพประกอบที่ 4.1

	Text	RateType
0	The story started after we ate at Baoshuan and headed downstairs to exit the hotel with allsmiles. Then, though our littl	1
1	This is worst and pathetic experience with The Oberoi New Delhi Zahir Hussain marg. Staffs doesn't know how to speak	1
2	Very bad Experience with property ever in my life and there is two Staff Member Vikrant and Jatin they miss behave..Do	1
3	This is very bad property and pathatic services. Staff doesn't have manner to talk with customers. There is two staff Ja	1
4	On our second night, as we were about to enter the elevator, a guy in black from the front desk came rushing to us sa	1
5	I checked into this hotel on Saturday and given below has been experienced Good experience The driver who picked m	1
6	Coming from Europe I chose the Oberoi due to its excellent reputation... Before departure getting an exrtra bed for my	1
7	Seems like I over rated / over expected from this big old brand an otherwise excellent engagement with the sales, until	1
8	Wanted to stay at the hotel and after speaking to the Sales Director, a room was booked for me. On the day of arrival	1
9	One of the worse dining experience as we booked for lunch buffet. The service was poor, empty dishes and even after	1
10	The hotel was oldfashioned and outdated in a bad way. Even the superior rooms were small and desperately in need of	1
11	First off we have stayed at the Oberoi on two different occasions while in Delhi and it truly is a beautiful hotel and lovely	1
12	After being left at the airport for over an hour only to get lost for two hours in a local taxi, the oberoi started their resolu	1
13	I'm in Delhi every month for work but this was the first time I have stayed at the Oberoi which on this occasion was for	1
14	I had money 180 stolen from my room on the last day of our holiday. Luckily I had gone to put the AED in it to chsnge	1
15	Worst experience I've ever had in a hotel in Dubai I will avoid staying at this hotel from the staff, services, and outdatec	1
16	Service level was clearly shown before arrival when I tried to contact hotel and Accor customer service to update our res	1
17	I frequently travel, 2 weeks every month and in different hotels experiencing different levels of services and that is why	1
18	On checking in the lovely male receptionist offered us an upgrade for 1000 dirims. In total After two nights I had a bad	1
19	We were looking forward to our stay at the Fairmont hotel also due to its pool, beach and garden which can be seen on	1
20	Bad experience, the air conditioning is very, very, very bad, and there is a female employee in the reception who treats	1
21	My family and I checked in yesterday to the fairmont on the palm. Quite frankly, check in process was unprofessional a	1
22	We booked Gold rooms and expected to be on levels 8 or 9 as per the website. We were allocated rooms on levels 3 a	1
23	READ BEFORE YOU TAKE YOUR FAMILY!Don't recommend this hotel for the price. I took my family of 5 here with a 1	1
24	Probably the worst hotel Ive ever stayed in in Dubai. Arrived at 9am had to wait until 3pm for a room, I booked a sea	1
25	We have been regularly visiting Little Miss India, and the experience have always been nice. This time, it was my wife's	1
26	This is the 56th time I have stayed at the Fairmont Palm in the last 8 years. There was been significant deterioration ov	1
27	My stay at The Fairmont Hotel The Palm was a disaster. The hotel was very extremely dirty, bed sheets smells bad an	1
28	The hotel is nice, but there is CONSTRUCTION works going all nigh just in front the building beach side. It is impossible t	1
29	Lovely hotel haveztayed in better, but there are building works on wither side of the hotel as well as further down on We	1

ภาพประกอบที่ 4.1 ตัวอย่างบทวิจารณ์โรงแรมที่ใช้ในการทดสอบ

4.2 ผลการทดลองของโมเดลที่ศึกษา (Result)

ในหัวข้อนี้จะกล่าวถึงผลการทดลองทั้งหมดในระบบการให้ระดับคะแนนของบทวิจารณ์โรงแรม ซึ่งได้มีการสร้างโมเดลจากโมเดลแบบทรานสฟอร์มเมอร์แบบ BERT-base และเข้าสู่ขั้นตอน Fine-Tuned ด้วยชุดข้อมูลชุดสอน (Training set) โดยผลการทดลองของโมเดลจากถูกประเมินประสิทธิภาพ

ด้วยตารางคอนฟิวชั่นเมทริกซ์ (ดังตารางที่ 4.1) และค่าความระลึก ค่าความแม่นยำ และ F1-score (ดังตารางที่ 4.2)

ตารางที่ 4.1 คอนฟิวชั่นเมทริกซ์ของโมเดลการจำแนกความรู้สึกจากโมเดล BERT-base

		Actual Results		
		Positive	Neural	Negative
Prediction Results	Positive	1918	66	23
	Neural	63	1716	243
	Negative	11	164	1825

ตารางที่ 4.2 แสดงการวัดประสิทธิภาพของโมเดลการจำแนกความรู้สึกจากโมเดล BERT-base

Class	Recall	Precision	F1-score	Accuracy
Negative	0.9125	0.8728	0.8922	
Neural	0.8487	0.8818	0.8649	
Positive	0.9557	0.9629	0.9593	
Average Scores	0.9035	0.9041	0.9038	0.9055

จากผลการทดลองดังที่แสดงในตารางที่ 4.2 แสดงให้เป็นว่าการสร้างโมเดลเพื่อการจำแนกความรู้สึกจากบทวิจารณ์โรงแรมนั้นให้ผลลัพธ์ในมุมมองของค่าความระลึก ค่าความแม่นยำ และค่าความถูกต้องได้อย่างน่าพอใจ เนื่องจาก

ในมุมมองด้านค่า Recall นั้น BERT-base มีการเรียนรู้แบบต่อเนื่อง (Sequential Learning) ซึ่งช่วยให้โมเดลสามารถเข้าใจความหมายของประโยคหรือข้อความที่ซับซ้อนได้เป็นอย่างดี การเรียนรู้แบบต่อเนื่องช่วยให้ BERT-base สามารถจดจำและระบุลักษณะที่เกี่ยวข้องกับข้อความที่ต้องการจำแนกความรู้สึกได้อย่างแม่นยำ ซึ่งส่งผลให้มีค่า recall สูง เพราะโมเดลสามารถตรวจจับข้อความที่เกี่ยวข้องกับความรู้สึก (Sentiment) ได้หลายกรณี

ในมุมมองด้านค่า Precision นั้น เนื่องจาก BERT-base มีความสามารถในการเข้าใจและระบุลักษณะที่เกี่ยวข้องกับประโยคหรือข้อความที่ต้องการจำแนกความรู้สึก โมเดลสามารถสร้างข้อความที่ถูกต้องเกี่ยวกับความรู้สึกออกมาได้เป็นอย่างดี ซึ่งนำไปสู่การมีค่า precision สูง

Computer Science Department
Faculty of Informatics, Mahasarakham University

ในมุมมองด้านค่า Accuracy นั้น BERT-base เป็นโมเดลที่มีความซับซ้อนและถูกฝึกสอนด้วยข้อมูลจำนวนมาก ซึ่งช่วยให้สามารถจำแนกข้อมูลความรู้สึกได้อย่างถูกต้องแม่นยำ ดังนั้นจึงส่งผลให้มีค่า accuracy สูง

4.3 การทดสอบเปรียบเทียบกับโมเดลการจำแนกความรู้สึกอื่นๆ

ในหัวข้อนี้จะนำโมเดลการจำแนกความรู้สึกที่สร้างจากอัลกอริทึมอื่นๆ มาเปรียบเทียบกับโมเดลการจำแนกความรู้สึกจากโมเดลแบบทรานสฟอร์มเมอร์แบบ BERT-base โดยโมเดลการจำแนกความรู้สึกที่สร้างจากอัลกอริทึมอื่นๆ นั้นจะสร้างขึ้นจากโมเดลชุดสอน และทดสอบประสิทธิภาพโมเดลด้วยข้อมูลชุดทดสอบเดียวกันกับโมเดลการจำแนกความรู้สึกจากโมเดลแบบทรานสฟอร์มเมอร์แบบ BERT-base และวัดประสิทธิภาพด้วยเกณฑ์เดียวกัน โดยโมเดลการจำแนกความรู้สึกที่นำมาทดสอบมีทั้งหมด 4 รูปแบบดังนี้

แบบที่ 1 โมเดลการจำแนกความรู้สึกจากเพื่อนบ้านใกล้เคียงที่สุด (K-Nearest Neighbor: KNN)

แบบที่ 2 โมเดลการจำแนกความรู้สึกจากซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine: SVM)

แบบที่ 3 โมเดลการจำแนกความรู้สึกจากป่าสุ่ม (Random Forest: RF)

แบบที่ 4 โมเดลการจำแนกความรู้สึกจากโมเดลแบบทรานสฟอร์มเมอร์แบบ roBERTa-base

4.3.1 ผลการทดลองกับโมเดลการจำแนกความรู้สึกจากเพื่อนบ้านใกล้เคียงที่สุด (K-Nearest Neighbor: KNN)

ตารางที่ 4.3 คอนฟิวชั่นเมทริกซ์ของโมเดลการจำแนกความรู้สึกจากอัลกอริทึม KNN

		Actual Results		
		Positive	Neural	Negative
Prediction Results	Positive	2865	61	78
	Neural	1407	1167	429
	Negative	1321	475	1221

Computer Science Department
Faculty of Informatics, Mahasarakham University

ตารางที่ 4.4 แสดงการวัดประสิทธิภาพของโมเดลการจำแนกความรู้สึกจากอัลกอริทึม KNN

Class	Recall	Precision	F1-score	Accuracy
Negative	0.7066	0.4047	0.5146	
Neural	0.6853	0.3886	0.4960	
Positive	0.5122	0.9537	0.6665	
Average Scores	0.6347	0.5823	0.5590	0.5821

ในโครงงานปริญญานิพนธ์นี้ใน KNN กำหนดให้ค่า K มีค่าเท่ากับ 3 หากเปรียบเทียบโมเดลของการวิเคราะห์ความรู้สึกที่สร้างจาก KNN และ BERT-base จะพบว่า KNN จะให้ประสิทธิภาพต่ำกว่า BERT-base สามารถมองในแต่ละมิติได้ดังนี้

1. การเรียนรู้แบบแบ่งกลุ่ม: KNN ใช้วิธีการเรียนรู้แบบแบ่งกลุ่ม (instance-based learning) ซึ่งจะใช้ข้อมูลการเรียนรู้เพื่อจำแนกคำในประโยคหรือข้อความที่ต้องการวิเคราะห์ โดยการหาคำที่ใกล้เคียงกันที่สุดจากชุดข้อมูลการเรียนรู้ (training dataset) อยู่ในตำแหน่งที่เหมือนกัน แต่ถึงแม้ว่า KNN จะสามารถจำแนกคำในประโยคหรือข้อความที่ใกล้เคียงให้ได้ แต่การที่จะเปรียบเทียบประสิทธิภาพแบบต่อเนื่องและสอดคล้องกับความซับซ้อนของประโยคหรือข้อความที่ต้องการวิเคราะห์จะทำให้ KNN มีประสิทธิภาพต่ำกว่า BERT-base ที่เรียนรู้และรู้เข้าใจความหมายของประโยคและข้อความอย่างละเอียด

2. การแปลงคำ (word embeddings): BERT-base มีการฝึกสอนด้วยการแปลงคำ (word embeddings) ที่มีความซับซ้อนและสามารถเข้าใจบริบทและความสัมพันธ์ระหว่างคำได้อย่างถูกต้อง ทำให้ BERT-base สามารถเข้าใจและรู้เข้าใจความหมายของประโยคหรือข้อความที่ต้องการจำแนก sentiment ได้เป็นอย่างดี ในทางตรงกันข้าม KNN ไม่ได้มีการเรียนรู้ความหมายของคำหรือประโยค แต่จะพยายามจำแนกข้อมูลตามความคล้ายคลึงกันของคำในชุดข้อมูลการเรียนรู้ ซึ่งอาจทำให้มีความคลาดเคลื่อนในการจำแนก sentiment ได้

จากเหตุผลข้างบน เมื่อเปรียบเทียบประสิทธิภาพระหว่าง KNN และ BERT-base ใน sentiment classification จะพบว่า BERT-base มักจะให้ผลลัพธ์ที่ดีกว่า KNN โดยทั่วไป โดยเฉพาะเมื่อมีความซับซ้อนและความหมายที่ซับซ้อนในประโยคหรือข้อความที่ต้องการจำแนกความรู้สึก

Computer Science Department
Faculty of Informatics, Mahasarakham University

4.3.2 โมเดลการจำแนกความรู้สึกจากซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine: SVM)

ตารางที่ 4.5 คอนฟิวชั่นเมทริกซ์ของโมเดลการจำแนกความรู้สึกจากอัลกอริทึม SVM

		Actual Results		
		Positive	Neural	Negative
Prediction Results	Positive	2711	141	150
	Neural	202	1513	1352
	Negative	98	44	2813

ตารางที่ 4.6 แสดงการวัดประสิทธิภาพของโมเดลการจำแนกความรู้สึกจากอัลกอริทึม SVM

Class	Recall	Precision	F1-score	Accuracy
Negative	0.6519	0.9519	0.7738	
Neural	0.8910	0.4933	0.6350	
Positive	0.9004	0.9031	0.9017	
Average Scores	0.8144	0.7828	0.7702	0.7798

ในโครงงานปริญญานิพนธ์นี้ใน SVM ร่วมกับฟังก์ชันฐานเชิงรัศมี (Radial Basis Function: RBF) โดย SVM เป็นอัลกอริทึมในการจำแนกข้อมูลที่ใช้เวกเตอร์เป็นพื้นฐานในการสร้างเส้นแบ่งแยกข้อมูลระหว่างกลุ่ม แม้ว่า SVM สามารถให้ผลลัพธ์ที่ดีในหลายงานที่ซับซ้อนได้ แต่ใน sentiment classification บางครั้ง SVM อาจให้ประสิทธิภาพที่ต่ำกว่า BERT-base ด้วยเหตุผลต่อไปนี้:

1. การจำแนกแบบเชิงเส้น: SVM ทำการจำแนกโดยการสร้างเส้นแบ่งแยกระหว่างกลุ่มข้อมูลในเชิงเส้น (Linear Separation) ซึ่งอาจทำให้ยากที่จะจำแนกคำหรือข้อความที่มีความซับซ้อนหรือไม่สามารถแยกได้ด้วยเส้นตรงเท่านั้น เช่น กรณีที่มีคำที่มีความหมายเปลี่ยนแปลงขึ้นอยู่กับบริบทหรือประโยคที่เกี่ยวข้อง
2. การจัดเรียงแบบเชิงเส้นเดียว: SVM ใช้การจัดเรียงแบบเชิงเส้นเดียว (Linear Ordering) ในการจำแนกกลุ่มข้อมูล ซึ่งไม่สามารถตรวจจับความซับซ้อนและความสัมพันธ์ที่ซับซ้อนระหว่างคำหรือ

Computer Science Department
Faculty of Informatics, Mahasarakham University

ประโยคได้อย่างละเอียดเหมือน BERT-base ที่เรียนรู้และจัดเรียงข้อมูลในรูปแบบที่ซับซ้อนและครอบคลุม

3. การจัดเรียงเส้นแบ่งที่ไม่เหมาะสม: SVM อาจพบข้อจำกัดในการจัดเรียงเส้นแบ่งที่ไม่เหมาะสมสำหรับข้อมูลที่ซับซ้อนหรือมีความหลากหลายมาก เช่น กรณีที่ข้อมูลมีการกระจายตัวแบบไม่สม่ำเสมอหรือข้อมูลที่แยกกันอยู่ในระดับสูง

4. ความซับซ้อนของการประมวลผล: SVM อาจมีความซับซ้อนในการประมวลผลเมื่อมีข้อมูลขนาดใหญ่ โดยที่ต้องคำนึงถึงการเลือกและปรับพารามิเตอร์ของ SVM ให้เหมาะสม เพื่อให้ได้ผลลัพธ์ที่ดี

จากข้างต้น จึงสามารถสรุปได้ว่า SVM อาจให้ประสิทธิภาพต่ำกว่า BERT-base ใน sentiment classification เนื่องจาก SVM มีข้อจำกัดในการจำแนกข้อมูลที่มีความซับซ้อนหรือไม่สามารถแยกแยะได้ด้วยเส้นเชิงเส้นอย่างเดียว ในขณะที่ BERT-base มีความสามารถในการรู้เข้าใจความหมายและความสัมพันธ์ของประโยคหรือข้อความที่ซับซ้อนมากกว่า

4.3.3 โมเดลการจำแนกความรู้สึกจากป่าสุ่ม (Random Forest: RF)

ตารางที่ 4.7 คอนฟิวชั่นเมทริกซ์ของโมเดลการจำแนกความรู้สึกจากอัลกอริทึม RF

		Actual Results		
		Positive	Neural	Negative
Prediction Results	Positive	2611	246	74
	Neural	1862	853	238
	Negative	1622	651	801

ตารางที่ 4.8 แสดงการวัดประสิทธิภาพของโมเดลการจำแนกความรู้สึกจากอัลกอริทึม RF

Class	Recall	Precision	F1-score	Accuracy
Negative	0.7197	0.2606	0.3826	
Neural	0.4874	0.2889	0.3628	
Positive	0.4345	0.8932	0.5846	
Average Scores	0.5472	0.4809	0.4433	0.4799

Computer Science Department
Faculty of Informatics, Mahasarakham University

Random Forest เป็นอัลกอริทึมที่ใช้หลักการของการสร้างต้นไม้การตัดสินใจ (Decision Tree) หลายต้นมาช่วยในการจำแนกข้อมูล แม้ว่า Random Forest สามารถให้ผลลัพธ์ที่ดีในหลายงานได้ แต่ใน sentiment classification บางครั้ง Random Forest อาจให้ประสิทธิภาพที่ต่ำกว่า BERT-base ด้วยเหตุผลต่อไปนี้:

1. ข้อจำกัดของต้นไม้การตัดสินใจ: แม้ว่า Random Forest จะใช้หลายต้นไม้ในการจำแนกข้อมูล แต่แต่ละต้นไม้ยังคงมีข้อจำกัดในการจัดเรียงและจำแนกข้อมูล ซึ่งอาจทำให้ยากต่อการจำแนกคำหรือข้อความที่มีความซับซ้อนหรือประโยคที่ต้องการวิเคราะห์อย่างละเอียด
2. การตัดสินใจที่ไม่เหมาะสม: ต้นไม้การตัดสินใจใน Random Forest อาจตัดสินใจที่ไม่เหมาะสมในบางกรณี โดยเฉพาะเมื่อมีข้อมูลที่มีความซับซ้อนหรือความสัมพันธ์ที่ซับซ้อนระหว่างคำหรือประโยค ซึ่ง BERT-base ที่มีการฝึกสอนและเรียนรู้ความหมายของประโยคและข้อความเป็นอย่างละเอียดสามารถจัดเรียงและตัดสินใจได้อย่างถูกต้องมากกว่า
3. ความซับซ้อนในการประมวลผล: Random Forest อาจมีความซับซ้อนในการประมวลผลเมื่อมีข้อมูลขนาดใหญ่ โดยที่ต้องคำนึงถึงการเลือกและปรับพารามิเตอร์ของ Random Forest ให้เหมาะสม เพื่อให้ได้ผลลัพธ์ที่ดี

สรุปแล้ว Random Forest อาจให้ประสิทธิภาพต่ำกว่า BERT-base ใน sentiment classification เนื่องจาก Random Forest อาจมีข้อจำกัดในการจำแนกข้อมูลที่มีความซับซ้อนหรือไม่สามารถแยกแยะได้ด้วยต้นไม้การตัดสินใจอย่างเดียว ในขณะที่ BERT-base มีความสามารถในการรู้เข้าใจความหมายและความสัมพันธ์ของประโยคหรือข้อความที่ซับซ้อนมากกว่า

4.3.4 โมเดลการจำแนกความรู้สึกจากโมเดลแบบทรานสฟอร์มเมอร์แบบ roBERTa-base

ตารางที่ 4.9 คอนฟิวชั่นเมทริกซ์ของโมเดลการจำแนกความรู้สึกจากโมเดล roBERTa-base

		Actual Results		
		Positive	Neural	Negative
Prediction Results	Positive	3001	44	14
	Neural	57	2721	216
	Negative	11	217	2673

Computer Science Department
Faculty of Informatics Mahasarakham University

ตารางที่ 4.10 แสดงการวัดประสิทธิภาพของโมเดลการจำแนกความรู้สึกจากโมเดล roBERTa-base

Class	Recall	Precision	F1-score	Accuracy
Negative	0.9238	0.9232	0.9235	
Neural	0.9125	0.9088	0.9106	
Positive	0.9557	0.9629	0.9593	
Average Scores	0.9307	0.9316	0.9311	0.9382

roBERTa-base เป็นรุ่นที่ปรับปรุงของ BERT base โดยใช้วิธีการที่ซับซ้อนขึ้นในการฝึกสอน และปรับปรุงโมเดลภายใน BERT architecture ซึ่งเป็นเหตุผลที่ทำให้ roBERTa-base ให้ประสิทธิภาพที่ดีกว่า BERT base ใน sentiment classification ดังนี้:

1. การเพิ่มขนาดของข้อมูลการฝึกสอน: roBERTa-base มีข้อมูลการฝึกสอนมากกว่า BERT base ซึ่งช่วยให้โมเดลสามารถเรียนรู้ความหมายและความสัมพันธ์ของคำและประโยคได้มากขึ้น
2. การเพิ่มจำนวนการฝึกสอน: roBERTa-base ถูกฝึกสอนในระยะเวลาที่นานขึ้นและด้วยข้อมูลมากขึ้น ทำให้โมเดลสามารถเรียนรู้และตีความข้อมูลในระดับสูงมากขึ้น
3. การเพิ่มขนาดและความลึกของโมเดล: roBERTa-base มีจำนวนของเลเยอร์ transformer มากกว่า BERT-base และถูกฝึกสอนในระยะเวลาที่นานขึ้น ซึ่งช่วยให้โมเดลสามารถเรียนรู้และจัดเรียงข้อมูลในลักษณะที่ซับซ้อนและครอบคลุมมากขึ้น
4. การกำจัดปัญหาโอเวอร์ฟิตติ้ง (overfitting): roBERTa-base ใช้วิธีการเพิ่มขนาดของข้อมูลการฝึกสอนและการฝึกสอนในระยะเวลาที่นานขึ้น ซึ่งช่วยลดปัญหาโอเวอร์ฟิตติ้งที่อาจเกิดขึ้นใน BERT-base
5. การปรับปรุงการฝึกสอน: roBERTa-base ได้รับการปรับปรุงในกระบวนการฝึกสอนและการตัดสินใจเพื่อให้ได้ผลลัพธ์ที่ดีขึ้นในการจำแนก sentiment

สรุปแล้ว roBERTa-base ให้ประสิทธิภาพที่ดีกว่า BERT-base ใน sentiment classification เนื่องจากการปรับปรุงในกระบวนการฝึกสอนและโครงสร้างภายในที่ช่วยให้โมเดลมีความสามารถในการเรียนรู้และตีความข้อมูลได้ดีกว่า

จากตารางการวัดประสิทธิภาพของโมเดลอื่นๆ สามารถสรุปและเปรียบเทียบประสิทธิภาพของโมเดลได้ดังนี้

Computer Science Department
Faculty of Informatics, Mahasarakham University

ตารางที่ 4.11 สรุปผลการทดสอบโมเดลและเปรียบเทียบประสิทธิภาพของโมเดลแบบต่างๆ

โมเดล	Accuracy	ค่าความระลึก	ค่าความแม่นยำ	ค่า F1-score
BERT-base	0.9055	0.9125	0.8728	0.8922
		0.8487	0.8818	0.8649
		0.9557	0.9629	0.9593
	ค่าเฉลี่ย	0.9035	0.9041	0.9038
roBERTa-base	0.9382	0.9238	0.9232	0.9235
		0.9125	0.9088	0.9106
		0.9557	0.9629	0.9593
	ค่าเฉลี่ย	0.9307	0.9316	0.9311
KNN	0.5821	0.7066	0.4047	0.5146
		0.6853	0.3886	0.4960
		0.5122	0.9537	0.6665
	ค่าเฉลี่ย	0.6347	0.5823	0.5590
SVM	0.7789	0.6519	0.9519	0.7738
		0.8910	0.4933	0.6350
		0.9004	0.9031	0.9017
	ค่าเฉลี่ย	0.8144	0.7828	0.7702
RF	0.4799	0.7197	0.2606	0.3826
		0.4874	0.2889	0.3628
		0.4345	0.8932	0.5846
	ค่าเฉลี่ย	0.5472	0.4809	0.4433

จากตารางที่ 4.11 จะเห็นว่า roBERTa-base ให้ประสิทธิภาพดีกว่า BERT base, KNN, SVM และ Random Forest ใน sentiment classification เนื่องจากการฝึกสอนและการปรับปรุงที่ละเอียดและซับซ้อนของโมเดลใน roBERTa-base ช่วยให้โมเดลเรียนรู้และตีความข้อมูลได้ดียิ่งขึ้นและมีความสามารถในการจัดเรียงและจำแนกข้อมูลที่เหมาะสมมากยิ่งขึ้น เนื่องจาก

1. Pre-training อย่างละเอียด: roBERTa-base ได้รับการฝึกสอนในระดับคำศัพท์โดยรวม (word level) และระดับส่วนของคำ (subword level) โดยใช้ข้อมูลจำนวนมาก ซึ่งช่วยให้ระบบเรียนรู้และเข้าใจคำศัพท์และความหมายในประโยคได้ลึกซึ้งมากขึ้น

Computer Science Department
Faculty of Informatics, Mahasarakham University

2. Masked Language Model (MLM): roBERTa-base มีการใช้ MLM ในการฝึกสอนโมเดล โดยการฝึกสอนโมเดลให้เดาคำที่ถูกซ่อนไว้ในประโยค ซึ่งช่วยให้โมเดลเรียนรู้ความสัมพันธ์ระหว่างคำและประโยคได้ดีกว่า

3. Dynamic Masking: roBERTa-base มีการใช้วิธี Dynamic Masking ในการฝึกสอน ซึ่งทำให้คำที่ถูกซ่อนไว้ในประโยคเปลี่ยนไปในแต่ละรอบการฝึกสอน ทำให้โมเดลมีการเรียนรู้ความสัมพันธ์ของคำในทางที่หลากหลายมากขึ้น

4. การทำซ้ำแบบสุ่ม (Random Iteration): roBERTa-base ได้ใช้การทำซ้ำแบบสุ่มในการฝึกสอน โดยการใช้ข้อมูลจากอินเทอร์เน็ตมากขึ้นและใช้กระบวนการสุ่มเพื่อสร้างตัวอย่างการฝึกสอนที่หลากหลายมากขึ้น

5. การปรับปรุงโครงสร้าง: roBERTa-base ได้รับการปรับปรุงในโครงสร้างที่ซับซ้อนขึ้น เช่น การเพิ่มจำนวนเลเยอร์ transformer, การปรับค่าพารามิเตอร์ต่างๆ ซึ่งช่วยให้โมเดลมีความสามารถในการจัดเรียงและจำแนกข้อมูลที่ดีกว่า

Computer Science Department
Faculty of Informatics, Mahasarakham University