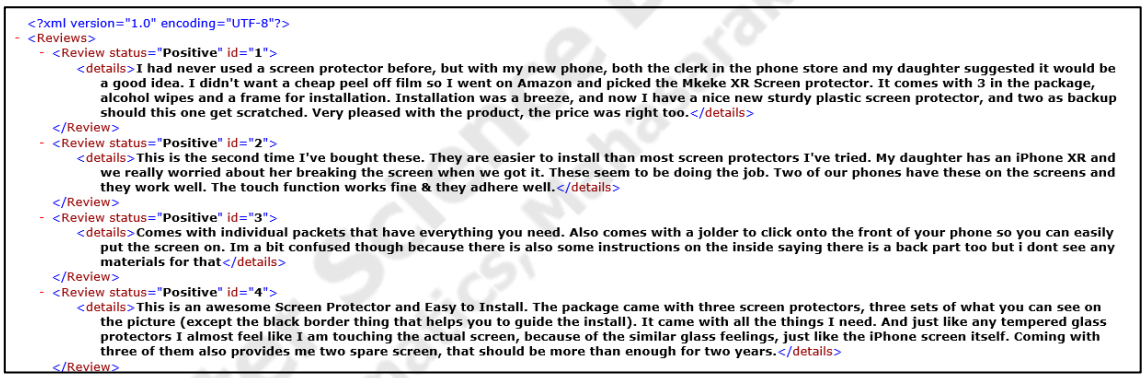
บทที่ 4

ผลการทดลอง

ในบทนี้จะกล่าวถึงการทดลองและผลการทดลอง ในการนำตัวจำแนกบทวิจารณ์สินค้าอิเล็กทรอนิกส์ ที่ได้จากขั้นตอนการดำเนินงาน มาทำการทดลองเพื่อจำแนกบทวิจารณ์สินค้าอิเล็กทรอนิกส์ที่ต้องการตรวจสอบ

4.1 ข้อมูลที่ใช้ในการทดสอบ

ข้อมูลที่ใช้การในทดสอบการสร้างโมเดลสำหรับการจำแนกบทวิจารณ์อิเล็กทรอนิกส์นั้น จะเป็นชุดข้อมูลทดสอบ (Test set) ที่ได้ทำการคัดเลือกไว้แล้วในขั้นตอนข้างต้น ที่เก็บอยู่ในรูปแบบของ XML ดังภาพประกอบที่ 4.1



ภาพประกอบที่ 4.1 ตัวอย่างบทวิจารณ์สินค้าอิเล็กทรอนิกส์ที่ใช้ในการทดสอบ

4.2 Algorithm Setup

ในหัวข้อนี้จะกล่าวถึงการตั้งค่าในอัลกอริทึมที่ใช้ในปริญญานิพนธ์นี้ ซึ่งประกอบไปด้วย 3 อัลกอริทึมดังนี้

4.2.1 KNN Setup

อัลกอริทึม *KNN* นั้นมีการกำหนดค่า *k* โดยค่า *k* ที่ใช้ในงานปริญญานิพนธ์นี้คือ 7 โดยการที่ได้ค่า *k* นั้นมาจากการทำการทดสอบค่า *k* ทั้งหมด 4 ค่า คือ 5, 7, 11 และ 15 กับทุกสัดส่วนที่ใช้ในการสร้างโมเดลแล้วนำมาเฉลี่ยหาค่าความระลึก ค่าความแม่นยำ และค่าเฉลี่ย *F-measure* โดยเราจะทำการนำค่า *k* ไปทดสอบกับทุกการให้น้ำหนักคำกับทุกสัดส่วนในการสร้างโมเดล

ตารางที่ 4.1 ตารางการทดสอบประสิทธิภาพของค่า *k*

ค่า *k*	ค่าความระลึก	ค่าความแม่นยำ	ค่าเฉลี่ย F-measure
5	0.6547	0.6615	0.6580
7	**0.6834**	**0.6713**	**0.6772**
11	0.6458	0.6450	0.6454
15	0.6232	0.6220	0.6226

ดังนั้นจากตารางที่ 4.1 เห็นได้ว่าค่า *k* ที่มาค่าเฉลี่ยมากที่สุด คือ *k*=7 เนื่องจากข้อมูลที่แล้วรองลงมาคือ *k*=5 เนื่องจาก ข้อมูลที่ใช้ในการสร้างโมเดลนั้น เป็นชุดข้อมูลที่ไม่มีความสมดุล ทำให้การที่ค่า *k* เยอะมีประสิทธิภาพที่ต่ำนั้นเป็นเรื่องที่เห็นได้เป็นปกติ ดังนั้น เราจึงได้ทำการเลือกใช้ค่า *k*=7 ในงานปริญญานิพนธ์นี้

4.2.2 Naïve Bayes

สำหรับอัลกอรึทึม Naïve Bayes นั้นได้ทำการใช้ Multinomial Naïve Bayes (MNB) ในการสร้างและทดสอบโมเดล เนื่องจาก MNB นั้นถูกสร้างขึ้นมาเพื่อใช้ในการจำแนกเอกสาร โดยมีการคำนวณสัดส่วนเอกสาร ซึ่ง MNB คือ ตัวทำนายที่ใช้โดยลักษณนามคือความถี่ของคำที่มีอยู่ในเอกสารมาใช้ให้เกิดประโยชน์มากที่สุด เนื่องจาก Naïve Bayes อื่น นั้นไม่เหมาะสมกับการนำมาจำแนกข้อมูลที่ไม่สมดุลในการสร้างโมเดลมากนัก

4.2.3 CNN Setup

ในส่วนของอัลกอริทึม *CNN* นั้นจะมีการเซ็ตค่าในการสร้างโมเดลของอัลกอริทึมโดยในแต่ละส่วนของการตั้งค่าได้มีการทดสอบประสิทธิภาพในการตั้งค่าเสมอ จึงจะนำการตั้งค่านั้นไปใช้ในงานจริง โดยการทดสอบในงานปริญญานิพนธ์นี้ได้ใช้ตัว Conv1D ซึ่งเป็นตัวที่ถูกใช้สำหรับ NLP มากที่สุด โดยเราได้กำหนด kernel_size = 4 เนื่องจากมีประสิทธิภาพที่ดีและใช้เวลาสั้นในการสร้างโมเดล

ตารางที่ 4.2 ค่าเฉลี่ยในการทดลองค่า input ในการทดสอบกับอัลกอริทึม *CNN*

filters	ค่าความระลึก	ค่าความแม่นยำ	ค่าเฉลี่ย F-measure
20	0.3104	0.3641	0.3322
30	0.4323	0.4752	0.4511
50	**0.6475**	**0.6654**	**0.6534**

ตารางที่ 4.2 ค่าเฉลี่ยในการทดลองค่า input ในการทดสอบกับอัลกอริทึม *CNN* (ต่อ)

filters	ค่าความระลึก	ค่าความแม่นยำ	ค่าเฉลี่ย F-measure
60	0.5497	0.5293	0.5395
70	0.3764	0.3354	0.3549

จากตารางที่ 4.2 จะเห็นได้ว่าเมื่อค่า filters อยู่ในระดับ 50 มีค่าเฉลี่ยสูงที่สุดในการทดสอบ เนื่องจากข้อมูลที่ใช้ในการสร้างโมเดลนั้น มีข้อมูลอยู่ในระดับกลาง ทำให้การที่ค่า filters น้อยหรือมากจนเกินไปจะทำให้ประสิทธิภาพของข้อมูลลดลง ดังนั้นในปริญญานิพนธ์นี้จึงเลือกค่า filters = 50

4.3 ผลการทดลอง (Results)

ในหัวข้อนี้จะกล่าวถึงผลการทดลองทั้งหมดในระบบการจำแนกบทวิจารณ์อิเล็กทรอนิกส์ ซึ่งได้มีการสร้างโมเดลโดยใช้อัลกอริทึมนาอีฟเบย์ (Naïve Bayes) อัลกอริทึมการหาเพื่อนบ้านใกล้ที่สุด (K-Nearest Neighbor: KNN) และอัลกอริทึมโครงข่ายประสาทแบบคอนโวลูชัน (Convolutional Neural Network: CNN) ดังหัวข้อต่อไปนี้

4.3.1 การทดสอบโมเดลในการจำแนกบทวิจารณ์โดยอัลกอริทึม KNN

สำหรับโมเดลการจำแนกบทวิจารณ์สินค้าอิเล็กทรอนิกส์แบบ 2 กลุ่ม ได้แก่ Positive class และ Negative class ซึ่งจะใช้เอกสารในการสร้างโมเดลตามสัดส่วนของเอกสารที่ไม่สมดุลกัน โดยจะให้ Positive class เป็นคลาสหลัก ที่มีเอกสาร 500 เอกสาร และให้ Negative class เป็นคลาสรองที่มีสัดส่วนเอกสารเป็นร้อยละ 10 20 และ 30 ของคลาสหลัก โดยใช้อัลกอริทึม *KNN* ในการทำนายเอกสารที่มีการให้น้ำหนักคำ

ในขั้นตอนการทดสอบโมเดลการจำแนกบทวิจารณ์สินค้าอิเล็กทรอนิกส์แบบ 2 กลุ่ม จะใช้เอกสารในการทดสอบจำนวน 1000 เอกสาร ซึ่งแบ่งออกเป็น 2 กลุ่ม จำนวนกลุ่มละ 500 เอกสาร เพื่อหาค่าความระลึก ค่าความแม่นยำ และค่า *F-measure* ในการประเมินความถูกต้องในการวิเคราะห์

ตารางที่ 4.3 ผลการทดสอบด้วยอัลกอริทึม *KNN*

การให้น้ำหนักคำ	สัดส่วนเอกสารที่ใช้ในการสร้างโมเดล (ร้อยละ)	จำนวนFeature ที่ใช้ในการสร้างโมเดล	เวลาที่ใช้ในการสร้างโมเดล (นาที)	เวลาที่ใช้ในการทดสอบโมเดล (นาที)	ค่าความระลึก	ค่าความแม่นยำ	ค่าเฉลี่ย F-measure
TF-IDF	100:10	1924	1.44	0.11	0.5162	0.5021	0.5074
	100:20	2081	1.54	0.12	0.5447	0.5246	0.5342
	100:30	2119	2.04	0.14	0.5941	0.5702	0.5801
	ค่าเฉลี่ย				**0.5462**	**0.5346**	**0.5403**
Delta TF-IDF	100:10	1924	1.38	0.10	0.5562	0.5544	0.5546
	100:20	2081	1.49	0.15	0.5714	0.5804	0.5766
	100:30	2119	2.14	0.14	0.5922	0.5752	0.5812
	ค่าเฉลี่ย				**0.5566**	**0.5550**	**0.5574**
TF-ICF-IDF	100:10	1924	1.40	0.12	0.5610	0.5532	0.5564
	100:20	2081	1.58	0.15	0.6012	0.5830	0.5912
	100:30	2119	2.10	0.14	0.6332	0.6242	0.6262
	ค่าเฉลี่ย				**0.5967**	**0.5834**	**0.5866**

ตารางที่ 4.3 ผลการทดสอบด้วยอัลกอริทึม KNN (ต่อ)

การให้น้ำหนักคำ	สัดส่วนเอกสารที่ใช้ในการสร้างโมเดล (ร้อยละ)	จำนวนFeature ที่ใช้ในการสร้างโมเดล	เวลาที่ใช้ในการสร้างโมเดล (นาที)	เวลาที่ใช้ในการทดสอบโมเดล (นาที)	ค่าความระลึก	ค่าความแม่นยำ	ค่าเฉลี่ย F-measure
TF-RF	100:10	1924	1.37	0.14	0.6401	0.6410	0.6403
	100:20	2081	1.52	0.13	0.6711	0.6862	0.6812
	100:30	2119	2.07	0.15	0.7035	0.7046	0.7062
	ค่าเฉลี่ย				**0.6763**	**0.6734**	**0.6724**
TF- IGM	100:10	1924	1.39	0.13	0.6456	0.6684	0.6594
	100:20	2081	1.50	0.13	0.6803	0.6614	0.6703
	100:30	2119	2.02	0.15	0.7045	0.7164	0.7021
	ค่าเฉลี่ย				**0.6734**	**0.6794**	**0.6782**

จากผลการทดสอบโมเดลการจำแนกบทวิจารณ์สินค้าอิเล็กทรอนิกส์ที่มีข้อมูลไม่สมดุลแบบ 2 กลุ่ม โดยใช้อัลกอริทึม *KNN* ดังตารางที่ 4.3 จะเห็นว่าการให้น้ำหนักคำ *TF-IGM* มีค่า *F-measure* สูงสุดในทุกสัดส่วนเอกสารที่ใช้ในการสร้างโมเดลด้วยอัลกอริทึม *KNN* โดยมีค่าเฉลี่ย *F-measure* อยู่ที่ 0.7052

4.3.2 การทดสอบโมเดลในการจำแนกบทวิจารณ์โดยอัลกอริทึม Naïve Bayes

สำหรับโมเดลการจำแนกบทวิจารณ์สินค้าอิเล็กทรอนิกส์แบบ 2 กลุ่ม ได้แก่ Positive class และ Negative class ซึ่งจะใช้เอกสารในการสร้างโมเดลตามสัดส่วนของเอกสารที่ไม่สมดุลกัน โดยจะให้ Positive class เป็นคลาสหลัก ที่มีเอกสาร 500 เอกสาร และให้ Negative class เป็นคลาสรองที่มีสัดส่วนเอกสารเป็นร้อยละ 10 20 และ 30 ของคลาสหลัก โดยใช้อัลกอริทึม *Naïve Bayes* ในการทำนายเอกสารที่มีการให้น้ำหนักคำ

ในขั้นตอนการทดสอบโมเดลการจำแนกบทวิจารณ์สินค้าอิเล็กทรอนิกส์แบบ 2 กลุ่ม จะใช้เอกสารในการทดสอบจำนวน 1000 เอกสาร ซึ่งแบ่งออกเป็น 2 กลุ่ม จำนวนกลุ่มละ 500 เอกสาร เพื่อหาค่าความระลึก ค่าความแม่นยำ และค่า *F-measure* ในการประเมินความถูกต้องในการวิเคราะห์

ตารางที่ 4.4 ผลการทดสอบด้วยอัลกอริทึม *Naïve Bayes*

การให้น้ำหนักคำ	สัดส่วนเอกสารที่ใช้ในการสร้างโมเดล (ร้อยละ)	จำนวนFeature ที่ใช้ในการสร้างโมเดล	เวลาที่ใช้ในการสร้างโมเดล (นาที)	เวลาที่ใช้ในการทดสอบโมเดล (นาที)	ค่าความระลึก	ค่าความแม่นยำ	ค่าเฉลี่ย F-measure
TF-IDF	100:10	1924	1.08	0.08	0.5546	0.5350	0.5421
	100:20	2081	1.38	0.09	0.5747	0.5542	0.5632
	100:30	2119	1.45	0.09	0.6304	0.6346	0.6323
	ค่าเฉลี่ย				**0.5767**	**0.5764**	**0.5737**
Delta TF-IDF	100:10	1924	1.07	0.08	0.5431	0.5294	0.5374
	100:20	2081	1.32	0.09	0.6466	0.6546	0.6575
	100:30	2119	1.44	0.08	0.7045	0.6978	0.7068
	ค่าเฉลี่ย				**0.6369**	**0.6268**	**0.6274**
TF-ICF-IDF	100:10	1924	1.10	0.09	0.6143	0.6233	0.6176
	100:20	2081	1.29	0.08	0.6436	0.6312	0.6366
	100:30	2119	1.39	0.08	0.6744	0.6561	0.6674
	ค่าเฉลี่ย				**0.6424**	**0.6337**	**0.6339**

ตารางที่ 4.4 ผลการทดสอบด้วยอัลกอริทึม *Naïve Bayes* (ต่อ)

การให้น้ำหนักคำ	สัดส่วนเอกสารที่ใช้ในการสร้างโมเดล (ร้อยละ)	จำนวนFeature ที่ใช้ในการสร้างโมเดล	เวลาที่ใช้ในการสร้างโมเดล (นาที)	เวลาที่ใช้ในการทดสอบโมเดล (นาที)	ค่าความระลึก	ค่าความแม่นยำ	ค่าเฉลี่ย F-measure
TF-RF	100:10	1924	1.07	0.08	0.6332	0.6242	0.6262
	100:20	2081	1.34	0.08	0.6811	0.6862	0.6812
	100:30	2119	1.47	0.08	0.7135	0.7146	0.7122
	ค่าเฉลี่ย				**0.6763**	**0.6734**	**0.6732**
TF- IGM	100:10	1924	1.39	0.09	0.6477	0.6345	0.6412
	100:20	2081	1.48	0.09	0.6716	0.6764	0.6735
	100:30	2119	2.05	0.10	0.7548	0.7068	0.7282
	ค่าเฉลี่ย				**0.6913**	**0.6725**	**0.6809**

จากผลการทดสอบโมเดลการจำแนกบทวิจารณ์สินค้าอิเล็กทรอนิกส์ที่มีข้อมูลไม่สมดุลแบบ 2 กลุ่ม โดยใช้อัลกอริทึม *Naïve Bayes* ดังตารางที่ 4.4 จะเห็นว่าการให้น้ำหนักคำ *TF-IGM* มีค่าเฉลี่ย *F-measure* สูงที่สุดอยู่ที่ 0.6809

4.3.3 การทดสอบโมเดลในการจำแนกบทวิจารณ์โดยอัลกอริทึม CNN

สำหรับโมเดลการจำแนกบทวิจารณ์สินค้าอิเล็กทรอนิกส์แบบ 2 กลุ่ม ได้แก่ Positive class และ Negative class ซึ่งจะใช้เอกสารในการสร้างโมเดลตามสัดส่วนของเอกสารที่ไม่สมดุลกัน โดยจะให้ Positive class เป็นคลาสหลัก ที่มีเอกสาร 500 เอกสาร และให้ Negative class เป็นคลาสรองที่มีสัดส่วนเอกสารเป็นร้อยละ 10 20 และ 30 ของคลาสหลัก โดยใช้อัลกอริทึม *CNN* ในการทำนายเอกสารที่มีการให้น้ำหนักคำ

ในขั้นตอนการทดสอบโมเดลการจำแนกบทวิจารณ์สินค้าอิเล็กทรอนิกส์แบบ 2 กลุ่ม จะใช้เอกสารในการทดสอบจำนวน 1000 เอกสาร ซึ่งแบ่งออกเป็น 2 กลุ่ม จำนวนกลุ่มละ 500 เอกสาร เพื่อหาค่าความระลึก ค่าความแม่นยำ และค่า *F-measure* ในการประเมินความถูกต้องในการวิเคราะห์

ตารางที่ 4.5 ผลการทดสอบด้วยอัลกอริทึม *CNN*

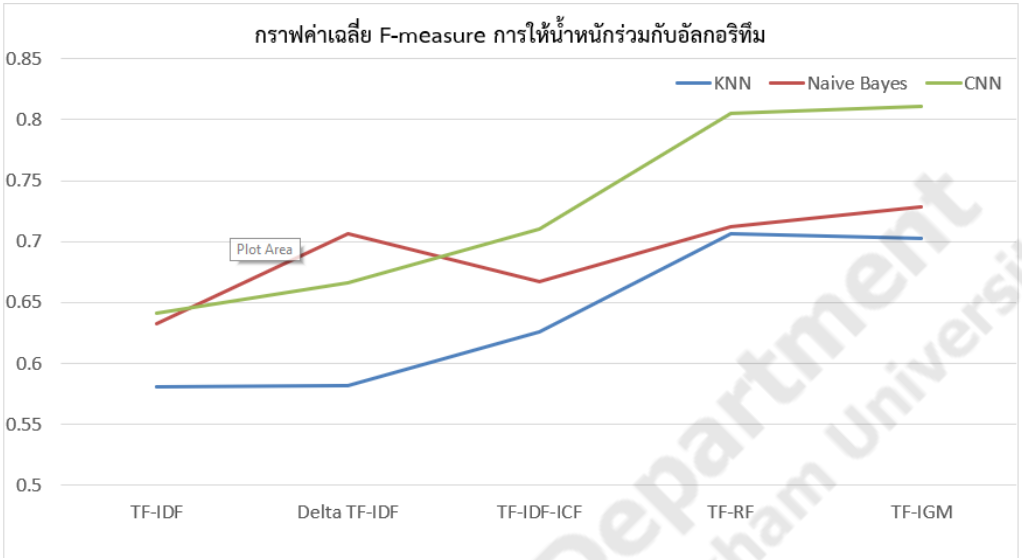
การให้น้ำหนักคำ	สัดส่วนเอกสารที่ใช้ในการสร้างโมเดล (ร้อยละ)	จำนวนFeature ที่ใช้ในการสร้างโมเดล	เวลาที่ใช้ในการสร้างโมเดล (นาที)	เวลาที่ใช้ในการทดสอบโมเดล (นาที)	ค่าความระลึก	ค่าความแม่นยำ	ค่าเฉลี่ย F-measure
TF-IDF	100:10	1924	2.08	1.05	0.5562	0.5644	0.5546
	100:20	2081	2.58	1.06	0.5914	0.6304	0.6066
	100:30	2119	3.45	1.05	0.6398	0.6202	0.6412
	ค่าเฉลี่ย				**0.5966**	**0.6150**	**0.6074**
Delta TF-IDF	100:10	1924	2.42	1.05	0.5610	0.5832	0.5764
	100:20	2081	3.44	1.05	0.6112	0.6530	0.6412
	100:30	2119	3.55	1.06	0.6632	0.6742	0.6662
	ค่าเฉลี่ย				**0.6267**	**0.6334**	**0.6266**
TF-ICF-IDF	100:10	1924	2.42	1.06	0.6342	0.6384	0.6362
	100:20	2081	3.48	1.07	0.6871	0.6872	0.6852
	100:30	2119	4.25	1.07	0.7135	0.7066	0.7102
	ค่าเฉลี่ย				**0.6782**	**0.6774**	**0.6761**

ตารางที่ 4.5 ผลการทดสอบด้วยอัลกอริทึม *CNN* (ต่อ)

การให้น้ำหนักคำ	สัดส่วนเอกสารที่ใช้ในการสร้างโมเดล (ร้อยละ)	จำนวนFeature ที่ใช้ในการสร้างโมเดล	เวลาที่ใช้ในการสร้างโมเดล (นาที)	เวลาที่ใช้ในการทดสอบโมเดล (นาที)	ค่าความระลึก	ค่าความแม่นยำ	ค่าเฉลี่ย F-measure
TF-RF	100:10	1924	2.22	1.06	0.6221	0.6512	0.6354
	100:20	2081	3.38	1.05	0.7398	0.7130	0.7212
	100:30	2119	4.05	1.05	0.8132	0.7954	0.8054
	ค่าเฉลี่ย				**0.7257**	**0.7198**	**0.7282**
TF- IGM	100:10	1924	2.52	1.06	0.6552	0.6752	0.6652
	100:20	2081	3.48	1.06	0.7598	0.7430	0.7512
	100:30	2119	4.35	1.06	0.8142	0.8214	0.8112
	ค่าเฉลี่ย				**0.7430**	**0.7438**	**0.7431**

จากผลการทดสอบโมเดลการจำแนกบทวิจารณ์สินค้าอิเล็กทรอนิกส์ที่มีข้อมูลไม่สมดุลแบบ 2 กลุ่ม โดยใช้อัลกอริทึม *CNN* ดังตารางที่ 4.5 จะเห็นว่าการให้น้ำหนักคำ *TF-IGM* มีค่า *F-measure* สูงสุดในทุกสัดส่วนเอกสารที่ใช้ในการสร้างโมเดลด้วยอัลกอริทึม *CNN* โดยมีค่าเฉลี่ย *F-measure* อยู่ที่ 0.7431

4.3.4 ภาพรวมผลการทดลอง



ภาพประกอบที่ 4.2 กราฟค่าเฉลี่ย *F-measure* การให้น้ำหนักร่วมกับอัลกอริทึม

จากภาพประกอบที่ 4.2 และตารางที่ 4.6 จะเห็นได้ว่าการให้น้ำหนักในรูปแบบต่างๆ มีประสิทธิภาพที่ดีในแต่ละอัลกอริทึมที่แตกต่างกัน ยกเว้นการให้น้ำหนัก *TF-IGM* ที่มีประสิทธิภาพที่ดีในทุกอัลกอรทึม

ตารางที่ 4.6 ตารางค่าเฉลี่ย *F-measure* การให้น้ำหนักร่วมกับอัลกอริทึม

	KNN	*Naive Bayes*	*CNN*
TF-IDF	0.5801	0.6323	0.6412
Delta TF-IDF	0.5812	0.7068	0.6662
TF-IDF-ICF	0.6262	0.6674	0.7102
TF-RF	0.7062	0.7122	0.8054
TF-IGM	0.7021	0.7282	0.8112

4.4 การทดสอบการจำแนกบทวิจารณ์ที่มีข้อมูลที่ต่างกัน 3 ชุดข้อมูลในทุกสัดส่วน

เนื่องจากอัลกอริทึมที่ใช้ในงานปริญญานิพนธ์นี้นั้น มีการกล่าวถึงการเพิ่มประสิทธิภาพของข้อมูลหากมีข้อมูลในการสร้างโมเดลที่มากขึ้น ดังนั้นในหัวข้อนี้จะทำการทดสอบการสร้างโมเดลที่มีข้อมูลในแต่ละสัดส่วนต่างกันดังต่อไปนี้

4.4.1 ทดสอบโมเดลกับ 3 สัดส่วนด้วยข้อมูล 3 ชุดที่ต่างกันกับอัลกอริทึม KNN

สำหรับการทดสอบโมเดลที่มีสัดส่วน 100 : 10, 100 : 20 และ 100 : 30 กับชุดข้อมูล 3 ชุด โดยข้อมูลกลุ่มหลักใช้ 100, 250 และ 500 เอกสาร และข้อมูลกลุ่มรองใช้ 10, 25 และ 50 เอกสาร ซึ่งจะสร้างโมเดลและทดสอบในอัลกอริทึมเพื่อนบ้านที่ใกล้ที่สุด (KNN)

โดยในขั้นตอนการทดสอบโมเดลการจำแนกบทวิจารณ์สินค้าอิเล็กทรอนิกส์แบบ 2 กลุ่ม จะใช้เอกสารในการทดสอบจำนวน 1000 เอกสาร ซึ่งแบ่งออกเป็น 2 กลุ่ม จำนวนกลุ่มละ 500 เอกสาร เพื่อหาค่าความระลึก ค่าความแม่นยำ และค่า *F-measure* ในการประเมินความถูกต้องในการวิเคราะห์

ตารางที่ 4.7 ทดสอบโมเดลที่มีสัดส่วน 100 : 10 กับข้อมูล 3 ชุดที่ต่างกันกับทุกอัลกอริทึม *KNN*

P : N	Recall					Precision					F-measure				
	TF-IDF	Delta TF-IDF	TF-IDF-ICF	TF-RF	TF-IGM	TF-IDF	Delta TF-IDF	TF-IDF-ICF	TF-RF	TF-IGM	TF-IDF	Delta TF-IDF	TF-IDF-ICF	TF-RF	TF-IGM
100 : 10	0.30	0.33	0.40	0.44	0.32	0.50	0.48	0.44	0.38	0.50	0.40	0.41	0.42	0.41	0.38
250 : 25	0.48	0.50	0.46	0.46	0.54	0.52	0.49	0.47	0.48	0.51	0.48	0.51	0.46	0.47	0.52
500 : 50	0.51	0.55	0.56	0.64	0.64	0.52	0.55	0.55	0.64	0.66	0.51	0.55	0.55	0.64	0.65

ตารางที่ 4.8 ทดสอบโมเดลที่มีสัดส่วน 100 : 20 กับข้อมูล 3 ชุดที่ต่างกันกับทุกอัลกอริทึม *KNN*

P : N	Recall					Precision					F-measure				
	TF-IDF	Delta TF-IDF	TF-IDF-ICF	TF-RF	TF-IGM	TF-IDF	Delta TF-IDF	TF-IDF-ICF	TF-RF	TF-IGM	TF-IDF	Delta TF-IDF	TF-IDF-ICF	TF-RF	TF-IGM
100 : 10	0.45	0.52	0.47	0.49	0.52	0.47	0.50	0.48	0.48	0.48	0.46	0.51	0.47	0.48	0.50
250 : 25	0.50	0.54	0.51	0.54	0.55	0.49	0.54	0.50	0.55	0.52	0.49	0.54	0.50	0.54	0.53
500 : 50	0.54	0.57	0.60	0.67	0.68	0.52	0.58	0.58	0.68	0.66	0.53	0.57	0.59	0.68	0.67

ตารางที่ 4.9 ทดสอบโมเดลที่มีสัดส่วน 100 : 30 กับข้อมูล 3 ชุดที่ต่างกันกับทุกอัลกอริทึม *KNN*

P : N	Recall					Precision					F-measure				
	TF-IDF	Delta TF-IDF	TF-IDF-ICF	TF-RF	TF-IGM	TF-IDF	Delta TF-IDF	TF-IDF-ICF	TF-RF	TF-IGM	TF-IDF	Delta TF-IDF	TF-IDF-ICF	TF-RF	TF-IGM
100 : 10	0.52	0.57	0.49	0.60	0.62	0.51	0.56	0.52	0.57	0.54	0.52	0.56	0.51	0.58	0.58
250 : 25	0.54	0.60	0.51	0.63	0.65	0.56	0.58	0.57	0.59	0.62	0.55	0.59	0.54	0.61	0.63
500 : 50	0.59	0.65	0.63	0.70	0.70	0.57	0.64	0.62	0.70	0.71	0.58	0.64	0.62	0.70	0.70

ในการทดสอบกับอัลกอริทึม *KNN* ดังตารางที่ 4.7 - ตารางที่ 4.9 นั้นเห็นได้ชัดว่าหากข้อมูลที่ใช้ในการสร้างโมเดลมีน้อยจะทำให้ประสิทธิภาพการทำงานของอัลกอริทึมลดลง และในขณะที่ข้อมูลในการสร้างโมเดลมีมากขึ้นประสิทธิภาพในการสร้างโมเดลก็ยิ่งเพิ่มขึ้นเช่นกัน โดยในการทดสอบกับโมเดลที่มีสัดส่วน 100:10, 100:20 และ 100:30 นั้นในชุดข้อมูลกลุ่มหลักที่มีขนาด 100 และ 250 เอกสาร นั้นการให้น้ำหนักคำแบบ *Delta TF-IDF* สามารถดึงประสิทธิภาพออกมากได้มากขึ้นในขณะที่การให้น้ำหนักแบบ *TF-IGM* ไม่สามารถดึงประสิทธิภาพออกมาได้เท่าที่ควร แต่ในชุดข้อมูลกลุ่มหลักที่มีขนาด 500 เอกสาร การให้น้ำหนัก TF-RF และ TF-IGM กับให้ประสิทธิภาพที่ดีในทุกสัดส่วน

4.4.2 ทดสอบโมเดลกับ 3 สัดส่วนด้วยข้อมูล 3 ชุดที่ต่างกันกับอัลกอริทึมนาอีฟเบย์

สำหรับการทดสอบโมเดลที่มีสัดส่วน 100 : 10, 100 : 20 และ 100 : 30 กับชุดข้อมูล 3 ชุด โดยข้อมูลกลุ่มหลักใช้ 100, 250 และ 500 เอกสาร และข้อมูลกลุ่มรองใช้ 10, 25 และ 50 เอกสาร ซึ่งจะสร้างโมเดลและทดสอบในอัลกอริทึมนาอีฟเบย์

โดยในขั้นตอนการทดสอบโมเดลการจำแนกบทวิจารณ์สินค้าอิเล็กทรอนิกส์แบบ 2 กลุ่ม จะใช้เอกสารในการทดสอบจำนวน 1000 เอกสาร ซึ่งแบ่งออกเป็น 2 กลุ่ม จำนวนกลุ่มละ 500 เอกสาร เพื่อหาค่าความระลึก ค่าความแม่นยำ และค่า *F-measure* ในการประเมินความถูกต้องในการวิเคราะห์

ตารางที่ 4.10 ทดสอบโมเดลที่มีสัดส่วน 100 : 10 กับข้อมูล 3 ชุดที่ต่างกันกับทุกอัลกอริทึมนาอีฟเบย์

P : N	Recall					Precision					F-measure				
	TF-IDF	Delta TF-IDF	TF-IDF-ICF	TF-RF	TF-IGM	TF-IDF	Delta TF-IDF	TF-IDF-ICF	TF-RF	TF-IGM	TF-IDF	Delta TF-IDF	TF-IDF-ICF	TF-RF	TF-IGM
100 : 10	0.43	0.45	0.48	0.44	0.46	0.43	0.42	0.48	0.47	0.46	0.43	0.42	0.48	0.45	0.46
250 : 25	0.47	0.50	0.52	0.51	0.49	0.47	0.51	0.52	0.51	0.50	0.47	0.50	0.52	0.51	0.49
500 : 50	0.55	0.54	0.61	0.63	0.64	0.53	0.55	0.62	0.62	0.63	0.54	0.54	0.61	0.62	0.64

ตารางที่ 4.11 ทดสอบโมเดลที่มีสัดส่วน 100 : 20 กับข้อมูล 3 ชุดที่ต่างกันกับทุกอัลกอริทึมนาอีฟเบย์

P : N	Recall					Precision					F-measure				
	TF-IDF	Delta TF-IDF	TF-IDF-ICF	TF-RF	TF-IGM	TF-IDF	Delta TF-IDF	TF-IDF-ICF	TF-RF	TF-IGM	TF-IDF	Delta TF-IDF	TF-IDF-ICF	TF-RF	TF-IGM
100 : 10	0.46	0.48	0.51	0.48	0.50	0.47	0.48	0.51	0.49	0.49	0.46	0.48	0.51	0.48	0.49
250 : 25	0.48	0.51	0.54	0.52	0.57	0.49	0.52	0.56	0.54	0.56	0.48	0.51	0.55	0.53	0.56
500 : 50	0.57	0.64	0.64	0.68	0.67	0.55	0.65	0.63	0.68	0.67	0.56	0.65	0.63	0.68	0.67

ตารางที่ 4.12 ทดสอบโมเดลที่มีสัดส่วน 100 : 30 กับข้อมูล 3 ชุดที่ต่างกันกับทุกอัลกอริทึมนาอีฟเบย์

P : N	Recall					Precision					F-measure				
	TF-IDF	Delta TF-IDF	TF-IDF-ICF	TF-RF	TF-IGM	TF-IDF	Delta TF-IDF	TF-IDF-ICF	TF-RF	TF-IGM	TF-IDF	Delta TF-IDF	TF-IDF-ICF	TF-RF	TF-IGM
100 : 10	0.48	0.51	0.51	0.50	0.47	0.48	0.51	0.53	0.49	0.47	0.48	0.51	0.52	0.49	0.47
250 : 25	0.53	0.58	0.57	0.60	0.64	0.51	0.57	0.55	0.57	0.63	0.52	0.57	0.56	0.58	0.63
500 : 50	0.63	0.64	0.67	0.71	0.75	0.63	0.69	0.65	0.71	0.70	0.63	0.66	0.66	0.71	0.72

ในการทดสอบกับอัลกอริทึมนาอีฟเบย์ ดังตารางที่ 4.10 - ตารางที่ 4.12 นั้นเห็นได้ชัดว่าหากข้อมูลที่ใช้ในการสร้างโมเดลหากมีน้อยจะทำให้ประสิทธิภาพการทำงานของอัลกอริทึมลดลง และในขณะที่ข้อมูลในการสร้างโมเดลมีมากขึ้นประสิทธิภาพในการสร้างโมเดลก็ยิ่งเพิ่มขึ้นเช่นกัน โดยในการทดสอบกับโมเดลที่มีสัดส่วน 100:10, 100:20 และ 100:30 นั้นในชุดข้อมูลกลุ่มหลักที่มีขนาด 100 นั้นการให้น้ำหนักคำแบบ *Delta TF-IDF* และ *TF-IDF-ICF* สามารถดึงประสิทธิภาพออกมากได้มากขึ้นในขณะที่การให้น้ำหนักแบบอื่น ไม่สามารถดึงประสิทธิภาพออกมาได้เท่าที่ควร แต่ในชุดข้อมูลกลุ่มหลักที่มีขนาด 250 และ 500 เอกสาร การให้น้ำหนัก TF-RF และ TF-IGM กับให้ประสิทธิภาพที่ดีในทุกสัดส่วน และการให้น้ำหนัก *TF-IDF* ก็ยังคงให้ประสิทธิภาพการทำงานร่วมกับอัลกอริทึมน้อยเช่นเดิม

4.4.3 ทดสอบโมเดลกับ 3 สัดส่วนด้วยข้อมูล 3 ชุดที่ต่างกันกับ CNN

สำหรับการทดสอบโมเดลที่มีสัดส่วน 100 : 10, 100 : 20 และ 100 : 30 กับชุดข้อมูล 3 ชุด โดยข้อมูลกลุ่มหลักใช้ 100, 250 และ 500 เอกสาร และข้อมูลกลุ่มรองใช้ 10, 25 และ 50 เอกสาร ซึ่งจะสร้างโมเดลและทดสอบในอัลกอริทึม CNN

โดยในขั้นตอนการทดสอบโมเดลการจำแนกบทวิจารณ์สินค้าอิเล็กทรอนิกส์แบบ 2 กลุ่ม จะใช้เอกสารในการทดสอบจำนวน 1000 เอกสาร ซึ่งแบ่งออกเป็น 2 กลุ่ม จำนวนกลุ่มละ 500 เอกสาร เพื่อหาค่าความระลึก ค่าความแม่นยำ และค่า *F-measure* ในการประเมินความถูกต้องในการวิเคราะห์

ตารางที่ 4.13 ทดสอบโมเดลที่มีสัดส่วน 100 : 10 กับข้อมูล 3 ชุดที่ต่างกันกับทุกอัลกอริทึม *CNN*

P : N	Recall					Precision					F-measure				
	TF-IDF	Delta TF-IDF	TF-IDF-ICF	TF-RF	TF-IGM	TF-IDF	Delta TF-IDF	TF-IDF-ICF	TF-RF	TF-IGM	TF-IDF	Delta TF-IDF	TF-IDF-ICF	TF-RF	TF-IGM
100 : 10	0.43	0.45	0.48	0.44	0.46	0.43	0.42	0.48	0.47	0.46	0.43	0.42	0.48	0.45	0.46
250 : 25	0.47	0.50	0.52	0.51	0.49	0.47	0.51	0.52	0.51	0.50	0.47	0.50	0.52	0.51	0.49
500 : 50	0.55	0.54	0.61	0.63	0.64	0.53	0.55	0.62	0.62	0.63	0.54	0.54	0.61	0.62	0.64

ตารางที่ 4.14 ทดสอบโมเดลที่มีสัดส่วน 100 : 20 กับข้อมูล 3 ชุดที่ต่างกันกับทุกอัลกอริทึม *CNN*

P : N	Recall					Precision					F-measure				
	TF-IDF	Delta TF-IDF	TF-IDF-ICF	TF-RF	TF-IGM	TF-IDF	Delta TF-IDF	TF-IDF-ICF	TF-RF	TF-IGM	TF-IDF	Delta TF-IDF	TF-IDF-ICF	TF-RF	TF-IGM
100 : 10	0.46	0.48	0.51	0.48	0.50	0.47	0.48	0.51	0.49	0.49	0.46	0.48	0.51	0.48	0.49
250 : 25	0.48	0.51	0.54	0.52	0.57	0.49	0.52	0.56	0.54	0.56	0.48	0.51	0.55	0.53	0.56
500 : 50	0.57	0.64	0.64	0.68	0.67	0.55	0.65	0.63	0.68	0.67	0.56	0.65	0.63	0.68	0.67

ตารางที่ 4.15 ทดสอบโมเดลที่มีสัดส่วน 100 : 30 กับข้อมูล 3 ชุดที่ต่างกันกับทุกอัลกอริทึม *CNN*

P : N	Recall					Precision					F-measure				
	TF-IDF	Delta TF-IDF	TF-IDF-ICF	TF-RF	TF-IGM	TF-IDF	Delta TF-IDF	TF-IDF-ICF	TF-RF	TF-IGM	TF-IDF	Delta TF-IDF	TF-IDF-ICF	TF-RF	TF-IGM
100 : 10	0.48	0.51	0.51	0.50	0.47	0.48	0.51	0.53	0.49	0.47	0.48	0.51	0.52	0.49	0.47
250 : 25	0.53	0.58	0.57	0.60	0.64	0.51	0.57	0.55	0.57	0.63	0.52	0.57	0.56	0.58	0.63
500 : 50	0.63	0.64	0.67	0.71	0.75	0.63	0.69	0.65	0.71	0.70	0.63	0.66	0.66	0.71	0.72

ในการทดสอบกับอัลกอริทึม *CNN* ดังตารางที่ 4.13 - ตารางที่ 4.15 นั้นเห็นได้ชัดว่าหากข้อมูลที่ใช้ในการสร้างโมเดลมีน้อยจะทำให้ประสิทธิภาพการทำงานของอัลกอริทึมลดลง และในขณะที่ข้อมูลในการสร้างโมเดลมีมากขึ้นประสิทธิภาพในการสร้างโมเดลก็ยิ่งเพิ่มขึ้นเช่นกัน โดยในการทดสอบกับโมเดลที่มีสัดส่วน 100:10, 100:20 และ 100:30 นั้นในชุดข้อมูลกลุ่มหลักที่มีขนาด 100, 250 และ 500 เอกสาร นั้นการให้น้ำหนักคำแบบ *Delta TF-IDF*, *TF-IDF-ICF, TF-RF* และ *TF-IGM* ให้ประสิทธิภาพที่ดีในทุกสัดส่วน

4.5 การวิเคราะห์ผล

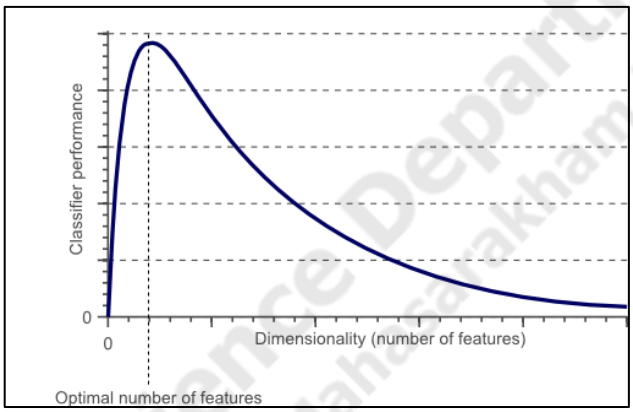
จากผลการทดสอบจะเห็นได้ว่ากรณีที่อัตราส่วนของข้อมูลที่สูงขึ้นนั้นส่งผลให้ประสิทธิภาพของอัลกอริทึมนั้นดีขึ้น และถ้าหากข้อมูลที่ใช้ในการสร้างมีมากขึ้นก็ส่งผลให้มีประสิทธิภาพที่ดีขึ้นเช่นกันต่อให้มีการไม่สมดุลของข้อมูลมากก็ตาม ซึ่งในการวิเคราะห์ผลนั้นประกอบไปด้วย

1) วิเคราะห์เกี่ยวกับวิธีการให้น้ำหนักคำ

- สำหรับรูปแบบการให้น้ำหนักคำทั้ง 5 รูปแบบจะเห็นได้ชัดว่ารูปแบบการให้น้ำหนักคำ *TF-IGM* มีค่าเฉลี่ยสูงสุดในทุกอัลกอริทึมเนื่องจากรูปแบบการให้น้ำหนักคำแบบ *TF-IGM* นั้น ถูกนำเสนอให้วัดความไม่สม่ำเสมอหรือความเข้มข้นของการแจกแจงคำศัพท์ระหว่างคลาสซึ่งสะท้อนให้เห็นถึงอำนาจการจำแนกชั้นข้อตกลง จึงทำให้เห็นความชัดเจนของการแยกข้อมูลในแต่ละคลาสเป็นอย่างดี ซึ่งเมื่อนำรูปแบบการให้น้ำหนักคำไปใช้กับอัลกอริทึม *CNN* แล้วทำให้เห็นว่าหากเอกสารมีข้อมูลไม่สมดุลทำให้การให้น้ำหนักคำแบบ *TF-IGM* ที่ใช้ร่วมกับอัลกอริทึม *CNN* สามารถแก้ปัญหาได้ดีที่สุด เมื่อเอกสารมีสัดส่วนที่ 100: 10 โดยมีค่าเฉลี่ยอยู่ที่ 0.6652 เมื่อเทียบกับรูปแบบอื่นๆ
- รองลงมาคือรูปแบบการให้น้ำหนักคำแบบ *TF-ICF-IDF* ที่มีค่าเฉลี่ยอยู่ที่ 0.6362 และรูปแบบที่มีค่าเฉลี่ยต่ำสุดคือ *TF-IDF* ที่มีค่าเฉลี่ยอยู่ที่ 0.5546
- สำหรับรูปแบบการให้น้ำหนักที่มีค่าเฉลี่ยมากสุดที่ทดสอบกับชุดข้อมูลมีสัดส่วน 100:20 และ 100:30 คือการให้น้ำหนักคำ *TF-IGM* ที่ทดสอบร่วมกับอัลกอริทึม *CNN* เช่นเดียวกับสัดส่วน 100:10 โดยมีค่า *F-measure* อยู่ที่ 0.7512 และ 0.8112 ตามลำดับ ซึ่งสัดส่วน 100:30 เป็นค่าที่สูงที่สุดในการทดสอบรูปแบบการให้ทั้งหมด และเห็นได้ชัดว่าหากข้อมูลมีค่าความไม่สมดุลต่างกันน้อยจะให้การจำแนกข้อมูลมีประสิทธิภาพมาก
- ส่วนการให้น้ำหนัก *TF-IDF* นั้นมีประสิทธิภาพต่ำที่สุดในทุกอัลกอริทึม เนื่องจากการให้น้ำหนัก *TF-IDF* นั้นเป็นการให้น้ำหนักคำที่คิดจากความถี่ของคำทั้งหมดของคลังเอกสาร นั้นทำให้การให้น้ำหนักในรูปแบบนี้มีประสิทธิภาพที่ต่ำกว่ารูปแบบอื่นที่คิดจากความถี่ของคำในแต่ละคลาส
- ดังนั้นสรุปได้ว่ารูปแบบการให้น้ำหนัก *UTW (TF-IDF)* นั้นให้ประสิทธิภาพต่ำกว่ารูปแบบการให้น้ำหนัก *STW (Delta TF-IDF, TF-IDF-ICF, TF-RF และ TF-IGM)* เนื่องจากการให้น้ำหนักคำรูปแบบ UTW เป็นการให้น้ำหนักที่คิดจากข้อมูลทั้งหมดในคลังข้อมูลแต่ในส่วนของ STW เป็นการให้น้ำหนักคำที่คิดจากกลุ่มของเอกสารเป็นหลัก ซึ่งเหมาะสมกับการใช้ในการแก้ปัญหาในงานปริญญานิพนธ์นี้

2) วิเคราะห์เกี่ยวกับอัลกอริทึม

- สำหรับอัลกอริทึมเพื่อนบ้านที่ใกล้ที่สุด จะมีประสิทธิภาพต่ำเมื่อมีคุณลักษณะจำนวนมาก เพราะถูกรบกวนจากคุณลักษณะที่ไม่เกี่ยวข้องได้ง่าย แต่เมื่อนำมาใช้ร่วมกับการให้น้ำหนักคำแบบ STW จะมีประสิทธิภาพดีขึ้น ตามที่ได้กล่าวไว้ในข้างต้นเพราะว่าอัลกอริทึมเพื่อนบ้านที่ใกล้ที่สุดได้รับผลกระทบจากคุณลักษณะที่ไม่เกี่ยวข้อง (Irrelevant feature) ต่อการวัดระยะทาง หรือการเกิดปัญหาของมิติข้อมูล (Curse of Dimensionality) อีกทั้งอัลกอริทึมเพื่อนบ้านที่ใกล้ที่สุดเหมาะกับชุดข้อมูลสอนที่มีปริมาณมาก แต่มีตัวอย่างคุณลักษณะจำนวนน้อย ดังภาพประกอบที่ 4.3



ภาพประกอบที่ 4.3 Curse of Dimensionality

- และอีกข้อที่สำคัญคือ การเลือกค่า *k* เพราะถ้าหากใช้ค่า *k* น้อยเกินไปอาจจะทำให้ไวต่อสัญญาณรบกวนได้ และถ้าหากเลือกค่า *k* มากเกินไปก็อาจจะทำให้มีกลุ่มข้อมูลอื่นๆ มาปะปนกับข้อมูลที่กำลังสนใจได้เช่นกัน

- ต่อมาสำหรับอัลกอริทึมนาอีฟเบย์ จะใช้งานได้ดีเมื่อมีคุณลักษณะจำนวนมาก และคุณลักษณะเป็นอิสระต่อกัน สังเกตได้จากตารางที่ 4.4 จะเห็นว่าอัลกอริทึมนาอีฟเบย์จะมีประสิทธิภาพดีที่สุด เมื่อใช้ร่วมกับการให้น้ำหนักคำแบบ *STW* ที่เป็นความถี่ของคำที่เกิดในแต่ละกลุ่มเท่านั้น แต่ถ้าหากใช้ร่วมกับการให้น้ำหนักคำที่มีการนำ *global weight* มาคำนวณร่วมด้วย อาจจะทำให้ประสิทธิภาพโมเดลลดลง

- สำหรับอัลกอริทึมโครงข่ายประสาทแบบคอนโวลูชัน เมื่อทำการทดสอบกับการให้น้ำหนักคำแบบ *STW* แล้วทำให้การจำแนกข้อมูลที่ไม่สมดุลมีประสิทธิภาพที่มากขึ้นเนื่องจาก *CNN* นั้นมีประสิทธิภาพในการจำแนกข้อมูลที่ไม่สมดุลอยู่แล้ว และหากต้องการให้ *CNN* มีประสิทธิภาพมากขึ้น

ควรใช้ชุดข้อมูลมีขนาดใหญ่ เนื่องจาก *CNN* นั้นถูกสร้างมาเพื่อทดสอบกับชุดข้อมูลชุดสอนที่มีขนาดใหญ่

3) วิเคราะห์การใช้งานการให้น้ำหนักร่วมกับอัลกอริทึม

- สำหรับการใช้งานการให้น้ำหนักร่วมกับอัลกอริทึมที่มีค่าเฉลี่ยมากที่สุดคือ การให้น้ำหนัก *TF-IGM* ร่วมกับอัลกอริทึม *CNN* ที่มีค่าเฉลี่ย *F-measure* สูงที่สุดอยู่ที่ 0.8112 ซึ่งมีประสิทธิภาพที่สุดที่สุด ซึ่งได้อธิบายไปข้างต้นแล้ว
- รองลงมาคือ การให้น้ำหนัก *TF-RF* ร่วมกับอัลกอริทึม *CNN* เช่นกันกับการให้น้ำหนัก *TF-IGM* เนื่องจากการให้น้ำหนักของทั้งสองรูปแบบนั้นมีการคำนวณการให้น้ำหนักคำที่คล้ายคลึงกันจึงทำให้มีประสิทธิภาพที่ใกล้เคียงกันมากที่สุด

4) วิเคราะห์เกี่ยวกับเวลาที่ใช้ในการสร้างและทดสอบ

สำหรับเวลาที่ใช้ในการประมวลผลจะขึ้นอยู่กับปัจจัย ดังนี้

1.จำนวนคุณลักษณะ (Feature)

ถ้าหากมีคุณลักษณะจำนวนมากเวลาที่ใช้ในการประมวลผลก็จะมากขึ้นตามไปด้วย เนื่องจากระบบต้องนำคุณลักษณะเหล่านั้นมาประมวลผล ดังนั้นโครงงานนี้จึงได้มีการลดจำนวนคุณลักษณะด้วยการใช้ *information gain* และการคัดเลือกคำด้วยพจนานุกรม ซึ่งการลดคุณลักษณะเหล่านี้ ไม่ส่งผลต่อความถูกต้องของการจัดกลุ่มเอกสาร เนื่องจากคุณลักษณะที่ถูกคัดออกไปไม่มีความสำคัญต่อการจัดกลุ่มเอกสาร แต่เป็นข้อมูลจริง ตัวอย่างคำที่ถูกคัดออก

ภาพประกอบที่ 4.4 คุณลักษณะที่ไม่ส่งผลต่อการจัดกลุ่ม

2. อัลกอริทึมที่ใช้ในการจัดกลุ่มเอกสาร (Algorithm)

สำหรับอัลกอริทึมนาอีฟเบย์นั้น ในการสร้างและทดสอบโมเดลจะใช้เวลาในการประมวลผลค่อนข้างเร็วเนื่องจากการคำนวณไม่ซับซ้อน ซึ่งแตกต่างจากอัลกอริทึมเพื่อนบ้านที่ใกล้ที่สุดที่ใช้ในการสร้างโมเดลจะมีความรวดเร็ว แต่จะใช้เวลาค่อนข้างนานในการทดสอบโมเดล เนื่องจากอัลกอริทึมเพื่อนบ้านที่ใกล้ที่สุดจะเป็นการนำเอกสารที่เข้ามาใหม่ไปคำนวณน้ำหนัก แล้ววัดระยะทางระหว่างเอกสารที่เข้ามาใหม่กับทุกเอกสารในโมเดล แล้วนำมาเรียงลำดับทำให้ตอนทดสอบโมเดลใช้เวลามากกว่าตอนสร้างโมเดลนั่นเอง สุดท้ายอันกอริทึมโครงข่ายประสาทคอนโวลูชันใช้เวลาในการสร้างโมเดลค่อนข้างนานเนื่องจากมีความซับซ้อนในการคำนวณอัลกอริทึม แต่เวลาที่ใช้ในการทดสอบโมเดลมีความเร็วใกล้เคียงกับทั้งสองอัลกอริทึม