



ความคิดเห็นที่เป็นตัวแทนของผลิตภัณฑ์

มหาวิทยาลัยศิลปากร สวทช. นวัตกรรม

โดย

นางล ภาณุพร เทศศรีเมือง

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาเทคโนโลยีสารสนเทศ

ภาควิชาคอมพิวเตอร์

บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร

ปีการศึกษา 2559

ลิขสิทธิ์ของบัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร

ความคิดเห็นที่เป็นตัวแทนของผลิตภัณฑ์

โดย

นางสาวนิภาพร เทศศรีเมือง

มหาวิทยาลัยศิลปากร สภานิติศาสตร์

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาดำเนินการตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาเทคโนโลยีสารสนเทศ

ภาควิชาคอมพิวเตอร์

บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร

ปีการศึกษา 2559

ลิขสิทธิ์ของบัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร

OPINION REPRESENTATIVE OF PRODUCTS

By

Miss Nipaporn Thessrimuang

มหาวิทยาลัยศิลปากร สภานิติศาสตร์

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree

Master of Science Program in Information Technology

Department of Computing

Graduate School, Silpakorn University

Academic Year 2016

Copyright of Graduate School, Silpakorn University

บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร อนุมัติให้วิทยานิพนธ์เรื่อง “ ความคิดเห็นที่เป็น
ตัวแทนของผลิตภัณฑ์ ” เสนอโดย นางสาวนิภาพร เทศศรีเมือง เป็นส่วนหนึ่งของการศึกษาตาม
หลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ

.....
(รองศาสตราจารย์ ดร.ปานใจ ธารทัศนวงศ์)

คณบดีบัณฑิตวิทยาลัย

วันที่ 1 เดือน มิ.ย. พ.ศ. 2560

อาจารย์ที่ปรึกษาวิทยานิพนธ์

อาจารย์ ดร.อรวรรณ เขาวลิต

คณะกรรมการตรวจสอบวิทยานิพนธ์

..... *ทัศนวงศ์* ประธานกรรมการ *สงวนลิขสิทธิ์*

(อาจารย์ ดร.ทัศนวงศ์ ศูนย์กลาง)

26 / 6 / 2560

..... *เทพชัย ทรัพย์นันทิ* กรรมการ

(ดร.เทพชัย ทรัพย์นันทิ)

25 / 6 / 2560

..... *สุนีย์ พงษ์พินิจบุญโญ* กรรมการ

(อาจารย์ ดร.สุนีย์ พงษ์พินิจบุญโญ)

26 / 6 / 2560

..... *อรวรรณ* กรรมการ

(อาจารย์ ดร.อรวรรณ เขาวลิต)

26 / 6 / 2560

57309203 : สาขาวิชาเทคโนโลยีสารสนเทศ

คำสำคัญ : การสรุปใจความสำคัญ/การวิเคราะห์ความรู้สึก/ความคล้ายคลึงแบบโคซายน์/อัลกอริทึม
นาอ็ฟเบย์/ อัลกอริทึมเพจเรงก์/ข่ายงานฮอปฟิลด์

นิภาพร เทศศรีเมือง : ความคิดเห็นที่เป็นตัวแทนของผลิตภัณฑ์. อาจารย์ที่ปรึกษา
วิทยานิพนธ์ : อ.ดร.อรวรรณ เชาวลิต. 113 หน้า.

ในปัจจุบันเครื่องสำอางเป็นสินค้าที่มีผู้คนให้ความสนใจมาก จึงทำให้มีการแข่งขันในการขายสินค้านี้ระหว่างผู้ประกอบการ รวมทั้งผู้บริโภคได้ศึกษาข้อมูลและรายละเอียดจากเว็บไซต์ต่างๆ ที่มีความคิดเห็นเกี่ยวกับผลิตภัณฑ์ก่อนตัดสินใจซื้อสินค้า ซึ่งเว็บไซต์ส่วนใหญ่ได้เปิดช่องทางให้ผู้บริโภคได้แสดงความคิดเห็นมากขึ้น ทำให้บทวิจารณ์ที่มีต่อสินค้าและบริการมีจำนวนมาก ทำให้เสียเวลาในการอ่านบทวิจารณ์นั้นได้อย่างครบถ้วน งานวิจัยนี้มีวัตถุประสงค์ในการสรุปความคิดเห็น (Text Summarization) จากบทวิจารณ์สินค้าของภาษาไทย และใช้เทคนิควิเคราะห์ความรู้สึก (Sentiment Analysis) ร่วมด้วย ซึ่งทำให้มองเห็นความรู้สึกที่มีต่อสินค้าและผลิตภัณฑ์ได้อย่างชัดเจน ทั้งความคิดเห็นเชิงบวกและเชิงลบ เพื่อช่วยให้ง่ายในการอ่านสรุปการแสดงความคิดเห็นของสินค้าประเภทเครื่องสำอาง การนำไปใช้ประโยชน์ ในด้านผู้บริโภคช่วยในการเลือกซื้อสินค้าได้ตรงตามความต้องการ ส่วนด้านผู้ประกอบการช่วยในการพัฒนาสินค้าให้ดีขึ้นตรงตามความต้องการของผู้บริโภค และใช้ในการวิจัยทางการตลาดเพื่อผลิตสินค้าชนิดใหม่ออกมา

งานวิจัยนี้มีขั้นตอนการทำงาน 3 ขั้นตอนหลัก (1) ทำการเก็บรวบรวมข้อมูล โดยข้อมูลที่ใช้ในการทดลอง คือบทวิจารณ์สินค้าประเภทเครื่องสำอางที่เป็นภาษาไทย (2) ใช้อัลกอริทึมนาอ็ฟเบย์ (Naive Bayes) ในการจัดกลุ่มเพื่อจำแนกประเภทของผลิตภัณฑ์ จากนั้นสร้างกลุ่มคำที่มีความคิดเห็นเชิงบวกและเชิงลบ เพื่อใช้วิเคราะห์ความรู้สึกของประโยค (3) หาความคล้ายคลึงของเอกสาร (Cosine similarity) จากนั้นใช้ PageRank algorithm, Hopfield Network algorithm และ Term Frequency Inverted Document Frequency หาประโยคที่มีความสำคัญที่สุด การวัดประสิทธิภาพของโปรแกรมโดยใช้ ROUGE มีค่าความแม่นยำสูงสุดที่ 90.46200% เป็นผลลัพธ์ที่ได้จาก Hopfield Network algorithm

ภาควิชาคอมพิวเตอร์

บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร

ลายมือชื่อนักศึกษา*นิภาพร*.....

ปีการศึกษา 2559

ลายมือชื่ออาจารย์ที่ปรึกษาวิทยานิพนธ์*อ.ร. /x*.....

57309203 : MAJOR : INFORMATION TECHNOLOGY

KEY WORD : TEXT SUMMARIZATION/SENTIMENT ANALYSIS/NAIVE BAYES/COSINE
SIMILARITY/PAGERANK ALGORITHM/HOPFIELD NETWORK
ALGORITHM/TERM FREQUENCY INVERTED DOCUMENT
FREQUENCY

NIPAPORN THESSRIMUANG : OPINION REPRESENTATIVE OF PRODUCTS.
THESIS ADVISOR : ORAWAN CHAOWALIT,Ph.D. 113 pp.

Nowadays people are more interested in cosmetic products; therefore cosmetic companies compete to be able to sell more. Customers can also easily search for make-up information on the internet. More websites has a feature for consumers to review the products so the consumers can see the reviews before making a purchase; however, reading long product reviews are very time consuming. Although some websites have a rating system, oftentimes the score and the opinions are not related. The purpose of this research is to summarize customer reviews of cosmetic products that were written in Thai by using sentiment analysis technics. The researcher expects to conclude the product reviews that can help the company and consumers. The company can use the product reviews to improve their products and service while consumers can find the right products that meets their needs and requirements. The consumers can read product reviews before purchasing the products or services. The consumers' opinions of the products are also highly influential in the decision making process to purchase the product.

This research summarizes the reviews by: 1) we used the customer reviews from popular website and store them in Text File. 2) Use Naïve Bayes to categorize the reviews into specific categories and analyze sentiment phrases of each type of products and group them into positive or negative sentiments 3) Use Cosine similarity to find similar documents and Choose important phrases by using PageRank algorithm, Hopfield Network algorithm and Term Frequency Inverted Document Frequency to conclude the most important phrase. To verify results, ROUGE technic was utilized. The accuracy of our system is 90.46200% on maximum which measured from Hopfield Network algorithm.

Department of Computing
Student's signature
Thesis Advisor's signature

Graduate School, Silpakorn University
Academic Year 2016

กิตติกรรมประกาศ

ผู้วิจัยขอกราบขอบพระคุณอาจารย์ ดร. อรวรรณ เซาวลิต อาจารย์ที่ปรึกษาวิทยานิพนธ์ที่ได้สละเวลา ในการชี้แนะแนวทางและให้ความรู้ต่างๆ ที่เป็นประโยชน์ ช่วยเหลือ ให้คำปรึกษา แก้ไขส่วนที่บกพร่อง และดูแลเอาใจใส่ทั้งเรื่องการทำงานวิจัยและการดำเนินชีวิตของผู้วิจัยเสมอมา ทำให้วิทยานิพนธ์นี้สำเร็จลุล่วงไปได้ด้วยดี ขอกราบขอบพระคุณอาจารย์ ดร. สุนีย์ พงษ์พิณีจัญญ์ อาจารย์ ดร. ทศนวรรณ ศูนย์กลาง และดร. เทพชัย ทรัพย์นิธิ ที่ได้ให้ความรู้ คำปรึกษาแนะนำ และชี้แนะแนวทาง แก้ไขส่วนที่บกพร่อง ด้วยดีตลอดมา

ผู้วิจัยขอกราบขอบพระคุณบิดามารดา และครอบครัว ที่คอยสนับสนุน ช่วยเหลือ และเป็นกำลังใจในยามที่รู้สึกท้อแท้ และได้มอบโอกาสในการศึกษาจนประสบความสำเร็จ ขอขอบคุณนาย ศุภนันท์ ไกรตะนะ นายณัฐพงษ์ แดงจ้อย และพี่ๆ เพื่อนๆ ทุกคนที่คอยช่วยเหลือและให้กำลังใจมาตลอดจนทำให้งานวิจัยนี้สำเร็จลงได้ ขอกราบขอบพระคุณเจ้าของหนังสือ วารสาร เอกสาร วิทยานิพนธ์ทุกเล่ม ที่ช่วยให้วิทยานิพนธ์นี้มีความสมบูรณ์

ท้ายนี้ผู้วิจัยหวังเป็นอย่างยิ่งว่างานวิทยานิพนธ์นี้จะเป็นประโยชน์สูงสุดต่อผู้ที่สนใจ สามารถนำไปใช้ประโยชน์ได้อย่างมีประสิทธิภาพ และเกิดประสิทธิผลมากที่สุด หากมีข้อผิดพลาดประการใด ผู้วิจัยขออภัยมา ณ ที่นี้ด้วย

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ง
บทคัดย่อภาษาอังกฤษ.....	จ
กิตติกรรมประกาศ.....	ฉ
สารบัญตาราง.....	ณ
สารบัญภาพ	ญ
บทที่	
1 บทนำ	1
ความเป็นมาและความสำคัญของปัญหา	1
วัตถุประสงค์การวิจัย	2
ขอบเขตการศึกษา	2
ขั้นตอนการศึกษา	2
เครื่องมือและอุปกรณ์	4
ประโยชน์ที่คาดว่าจะได้รับ.....	4
2 วรรณกรรมที่เกี่ยวข้อง	5
Query-Based Summarization : A survey	5
An Automated System for Summarizing Structured Product Reviews	8
Opinion Strength Identification in Customer Review Summarizing System Using Association Rule Technique.....	10
Mining and Summarizing Customer Reviews.....	13
Emotion Categorization System for the Thai Texts.....	16
Automatically Rating of Hotel Service Aspects from Textual Customer Reviews	17
Automatic Feeling Analysis from Opinion Text	19
Thai News Text Summarization and Its Application	21
Ranking Explanatory Sentences for Opinion Summarization.....	23
3 ทฤษฎีที่เกี่ยวข้อง.....	26
ทฤษฎีของเบย์ส์ (Bayes Theorem).....	26
การหาความคล้ายคลึงของเอกสาร (Cosine similarity)	26

บทที่	หน้า
อัลกอริทึม PageRank.....	28
อัลกอริทึม Hopfield networks.....	29
การหาค่าความถี่ของเอกสาร (TF-IDF).....	30
การวัดประสิทธิภาพ (ROUGE)	31
ค่าความถูกต้อง, ค่าความครบถ้วน, ค่าเอฟเมชเชอร์	37
4 วิธีการดำเนินงานวิจัย	38
ข้อมูลที่ใช้ในงานวิจัย	38
วิธีการดำเนินงานวิจัย	38
ขั้นตอนการเตรียมข้อมูล.....	41
ขั้นตอนการจำแนกข้อมูลตามกลุ่มของความคิดเห็น.....	43
ขั้นตอนการประมวลผลข้อมูล.....	49
การเลือกประโยคที่มีคะแนนเท่ากัน	59
แผนการดำเนินงาน.....	61
5 ผลการดำเนินงานวิจัย.....	62
ผลการดำเนินงานวิจัย.....	62
การวัดประสิทธิภาพ (ROUGE).....	64
6 สรุปผลการทดลอง และข้อเสนอแนะ.....	71
สรุปผลการทดลอง.....	71
ข้อเสนอแนะและงานที่จะศึกษาต่อ.....	72
รายการอ้างอิง.....	74
ภาคผนวก.....	78
ภาคผนวก ก ผลลัพธ์จากการดำเนินงานวิจัย.....	79
ภาคผนวก ข คลังข้อมูล.....	106
ภาคผนวก ค คู่มือการใช้งานโปรแกรม (User' s Document).....	110
ประวัติผู้วิจัย.....	113

สารบัญตาราง

ตารางที่		หน้า
1	ตัวอย่างรายการคำแสดงความคิดเห็นเชิงบวกจากงานวิจัยของ (Trithep Thumrongluck, 2010).....	9
2	ตัวอย่างรายการคำแสดงความคิดเห็นเชิงลบจากงานวิจัยของ (Trithep Thumrongluck, 2010).....	10
3	ความถี่ของคำในประโยค.....	27
4	ตัวอย่างการคำนวณหาค่าถูกต้อง, ค่าความครบถ้วน, ค่าเอฟเมชเชอร์.....	37
5	ตัวอย่างความถี่ของคำในประโยค.....	49
6	ค่าความคล้ายคลึงกันของประโยค.....	52
7	ตัวอย่างการคำนวณหาค่า PageRank.....	55
8	ตัวอย่างการคำนวณหาค่า Hopfield.....	57
9	ความถี่ของคำในประโยค (tf).....	60
10	ค่าของ idf.....	60
11	ค่าของ tf-idf.....	60
12	ขั้นตอนและแผนการดำเนินงาน.....	61
13	ค่าการวัดประสิทธิภาพผลิตภัณ์ของ ROUGE-1 (Recall).....	64
14	ค่าการวัดประสิทธิภาพผลิตภัณ์ของ ROUGE-1 (Precision).....	66
15	ค่าการวัดประสิทธิภาพผลิตภัณ์ของ ROUGE-1 (F-measure).....	68
16	ค่าการวัดประสิทธิภาพของผลิตภัณ์ประเภทที่ปิดแก้ม M.A.C-PowderBlush ที่ได้จากอัลกอริทึม PageRank.....	79
17	ค่าการวัดประสิทธิภาพของผลิตภัณ์ประเภทที่ปิดแก้ม M.A.C-PowderBlush ที่ได้จากอัลกอริทึม Hopfield.....	80
18	ค่าการวัดประสิทธิภาพของผลิตภัณ์ประเภทที่ปิดแก้ม M.A.C-PowderBlush ที่ได้จากอัลกอริทึม TF-IDF.....	81
19	ค่าการวัดประสิทธิภาพของผลิตภัณ์ประเภทที่ปิดแก้ม NARS-Blush ที่ได้จากอัลกอริทึม PageRank.....	82
20	ค่าการวัดประสิทธิภาพของผลิตภัณ์ประเภทที่ปิดแก้ม NARS-Blush ที่ได้จากอัลกอริทึม Hopfield.....	83

ตารางที่	หน้า
21	ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภทที่ปิดแก้ม NARS-Blush ที่ได้จากอัลกอริทึม TF-IDF 84
22	ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภทที่ปิดแก้ม OP-ReadyToWearAutomatic ที่ได้จากอัลกอริทึม PageRank 85
23	ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภทที่ปิดแก้ม OP-ReadyToWearAutomatic ที่ได้จากอัลกอริทึม Hopfield 86
24	ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภทที่ปิดแก้ม OP-ReadyToWearAutomatic ที่ได้จากอัลกอริทึม TF-IDF 87
25	ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภททำความสะอาด BIODERMA-SensibioH2O ที่ได้จากอัลกอริทึม PageRank 88
26	ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภททำความสะอาด BIODERMA-SensibioH2O ที่ได้จากอัลกอริทึม Hopfield 89
27	ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภททำความสะอาด BIODERMA-SensibioH2O ที่ได้จากอัลกอริทึม TF-IDF 90
28	ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภททำความสะอาด Cetaphil-GentleSkin- Cleanser ที่ได้จากอัลกอริทึม PageRank 91
29	ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภททำความสะอาด Cetaphil-GentleSkin- Cleanser ที่ได้จากอัลกอริทึม Hopfield 92
30	ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภททำความสะอาด Cetaphil-GentleSkin- Cleanser ที่ได้จากอัลกอริทึม TF-IDF 93
31	ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภททำความสะอาด DHC-DeepCleansingOil ที่ได้จากอัลกอริทึม PageRank 94
32	ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภททำความสะอาด DHC-DeepCleansingOil ที่ได้จากอัลกอริทึม Hopfield 95
33	ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภททำความสะอาด DHC-DeepCleansingOil ที่ได้จากอัลกอริทึม TF-IDF 96
34	ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภทครีมรองพื้น CathyDoll-CCCream- SPF50PA+++ ที่ได้จากอัลกอริทึม PageRank 97
35	ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภทครีมรองพื้น CathyDoll-CCCream- SPF50PA+++ ที่ได้จากอัลกอริทึม Hopfield..... 98

ตารางที่	หน้า
36	ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภทครีมรองพื้น CathyDoll-CCCream-SPF50PA+++ ที่ได้จากอัลกอริทึม TF-IDF 99
37	ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภทครีมรองพื้น Revlon-ColorstayMakeup ที่ได้จากอัลกอริทึม PageRank 100
38	ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภทครีมรองพื้น Revlon-ColorstayMakeup ที่ได้จากอัลกอริทึม Hopfield 101
39	ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภทครีมรองพื้น Revlon-ColorstayMakeup ที่ได้จากอัลกอริทึม TF-IDF 102
40	ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภทครีมรองพื้น Revlon-PhotoReady-Foundation ที่ได้จากอัลกอริทึม PageRank 103
41	ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภทครีมรองพื้น Revlon-PhotoReady-Foundation ที่ได้จากอัลกอริทึม Hopfield 104
42	ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภทครีมรองพื้น Revlon-PhotoReady-Foundation ที่ได้จากอัลกอริทึม TF-IDF 105
43	รายการคำแสดงความคิดเห็นเชิงบวกจากงานวิจัยของ (Trithep Thumrongluck, 2010)..... 106
44	รายการคำแสดงความคิดเห็นเชิงบวกจากงานวิจัยนี้ 107
45	รายการคำแสดงความคิดเห็นเชิงลบจากงานวิจัยของ (Trithep Thumrongluck, 2010)..... 108
46	รายการคำแสดงความคิดเห็นเชิงลบจากงานวิจัยนี้ 109

สารบัญภาพ

ภาพที่		หน้า
1	สัญลักษณ์และความหมายที่ใช้ในการให้ความหนักแน่นของคำ.....	10
2	ตัวอย่างการระบุทิศทางของความคิดเห็น	11
3	ตัวอย่างระดับความหนักแน่นของความคิดเห็น	11
4	ตัวอย่างที่ได้จากกฎความสัมพันธ์	12
5	ภาพรวมการทำงานของ Mining and Summarizing Customer Reviews.....	13
6	ตัวอย่างการระบุหน้าที่ (POS tags).....	14
7	ตัวอย่างการจัดกลุ่มคำที่มีความหมายเหมือนกันและตรงข้ามกัน	14
8	ตัวอย่างประโยคสรุปของความคิดเห็น	15
9	แผนภาพกิจกรรมระบบจำแนกระดับอารมณ์สำหรับข้อความภาษาไทย	16
10	กระบวนการจัดระดับความคิดเห็นเกี่ยวกับบริการโรงแรมแบบอัตโนมัติ.....	18
11	กระบวนการวิเคราะห์ความรู้สึกของแบบอัตโนมัติ.....	19
12	การกำหนดโครงสร้างของประโยคภาษาไทย	21
13	ภาพรวมการทำงานของระบบสรุปข่าว	22
14	ภาพรวมการทำงานของสรุปความคิดเห็นด้วยการจัดลำดับ (Ranking).....	24
15	การเปรียบเทียบคะแนน	25
16	ตัวอย่างผลลัพธ์ที่ได้จากการทำงานของ SumWordLR และ LexRank.....	25
17	ตัวอย่างการเชื่อมโยงแต่ละประโยค.....	29
18	ภาพรวมการทำงานของระบบสรุปทวิจรรย์สินค้าแบบที่ 1 และแบบที่ 2	39
19	ภาพรวมการทำงานของระบบสรุปทวิจรรย์สินค้าแบบที่ 3	40
20	ตัวอย่างความคิดเห็นของผู้ใช้จากเว็บไซต์.....	41
21	ตัวอย่างความคิดเห็นเก็บในรูปแบบของ Text File.....	42
22	ตัวอย่างที่ได้จากการตัดคำ.....	42
23	ตัวอย่างความคิดเห็นที่เป็นบวก	43
24	ตัวอย่างความคิดเห็นที่เป็นลบ	43
25	ข้อมูล (Training) JSON files.....	45
26	ไฟล์ข้อมูลที่ใช้ในการจำแนกกลุ่มของความคิดเห็น	45
27	ตัวอย่างโค้ด (Code) การจัดกลุ่มโดยใช้อัลกอริทึมนาอ็ฟเบย์ (Naive Bayes).....	46

ภาพที่	หน้า
28	ตัวอย่างผลลัพธ์ที่ได้จากการจัดกลุ่ม 46
29	ความสัมพันธ์ของแต่ละประโยค 53
30	สรุปการวัดประสิทธิภาพจากอัลกอริทึม PageRank ของ ROUGE-1 (Recall)..... 65
31	สรุปการวัดประสิทธิภาพจากอัลกอริทึม Hopfield ของ ROUGE-1 (Recall) 65
32	สรุปการวัดประสิทธิภาพจากอัลกอริทึม TF-IDF ของ ROUGE-1 (Recall)..... 66
33	สรุปการวัดประสิทธิภาพจากอัลกอริทึม PageRank ของ ROUGE-1 (Precision)..... 67
34	สรุปการวัดประสิทธิภาพจากอัลกอริทึม Hopfield ของ ROUGE-1 (Precision) 67
35	สรุปการวัดประสิทธิภาพจากอัลกอริทึม TF-IDF ของ ROUGE-1 (Precision)..... 68
36	สรุปการวัดประสิทธิภาพจากอัลกอริทึม PageRank ของ ROUGE-1 (F-measure)..... 69
37	สรุปการวัดประสิทธิภาพจากอัลกอริทึม Hopfield ของ ROUGE-1 (F-measure) 69
38	สรุปการวัดประสิทธิภาพจากอัลกอริทึม TF-IDF ของ ROUGE-1 (F-measure)..... 70
39	ตัวอย่างโครงสร้าง Deep learning..... 73
40	หน้าแรกของระบบ 110
41	เลือกผลิตภัณฑ์เครื่องสำอางที่ผู้ใช้ต้องการ 111
42	เลือกยี่ห้อของผลิตภัณฑ์เครื่องสำอางที่ผู้ใช้ต้องการ..... 111
43	กดปุ่ม Submit..... 112
44	ระบบแสดงผลลัพธ์ของความคิดเห็นที่ผู้ใช้ต้องการ 112

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

จากการที่อินเทอร์เน็ตมีความจำเป็นมากในการดำรงชีวิตในปัจจุบัน ทั้งเทคโนโลยีทางด้านต่างๆ มีความเจริญก้าวหน้า พัฒนาไปอย่างรวดเร็ว ดังนั้นจะเห็นได้ว่าในปัจจุบันแนวโน้มของจำนวนผู้ใช้อินเทอร์เน็ตเพิ่มขึ้นอย่างต่อเนื่อง (Thailand web directory and Advance Web Statistics, 2016) จะสังเกตได้จากการใช้งาน Social Network ในการติดต่อสื่อสาร จึงทำให้เกิดช่องทางในการทำธุรกิจซื้อขายสินค้าระหว่างผู้ประกอบการกับผู้บริโภค และมีการแข่งขันในการซื้อขายสินค้ากันอย่างมาก ในปัจจุบันเครื่องสำอางเป็นสินค้าที่มีผู้คนที่ให้ความสนใจมาก และผลิตภัณฑ์เกี่ยวกับเครื่องสำอางก็ได้รับความนิยมในกลุ่มผู้หญิงและมีข้อมูลผลิตภัณฑ์สินค้าแพร่หลายในอินเทอร์เน็ต

นอกจากการซื้อขายสินค้าผ่านบริการออนไลน์แล้ว ผู้ประกอบการและผู้บริโภค ยังมีความสนใจในการแสดงความคิดเห็นต่อสินค้าเป็นอย่างมาก ซึ่งการแสดงความคิดเห็นของผู้บริโภคที่มีต่อสินค้า ยังมีอิทธิพลอย่างมากในการตัดสินใจซื้อสินค้าด้วย เพราะผู้บริโภคจะหาข้อมูลบทวิจารณ์เกี่ยวกับสินค้านั้นก่อนตัดสินใจซื้อสินค้า ซึ่งเว็บไซต์ส่วนใหญ่ได้เปิดช่องทางให้ผู้บริโภคได้แสดงความคิดเห็นมากขึ้น สำหรับผู้บริโภคก่อนตัดสินใจซื้อสินค้าได้ศึกษาข้อมูลและรายละเอียดจากเว็บไซต์ต่างๆที่มีความคิดเห็นเกี่ยวกับสินค้านั้น ตัวอย่างเว็บไซต์เครื่องสำอาง เช่น www.jeban.com, www.vanilla.in.th, www.cosmenet.in.th เป็นต้น ทำให้บทวิจารณ์ที่มีต่อสินค้าและบริการมีจำนวนมาก และมีการแสดงความคิดเห็นที่หลากหลาย ทำให้เสียเวลาในอ่านบทวิจารณ์นั้นได้อย่างครบถ้วน ถึงแม้จะมีการกำหนดการให้คะแนนต่อสินค้าแล้ว แต่ในหลายครั้งคะแนนที่ได้กับความคิดเห็นก็ไม่สัมพันธ์กัน ผู้วิจัยคาดว่าในการทำสรุปความคิดเห็นต่อสินค้าจะมีประโยชน์ทั้งทางด้านผู้ประกอบการ และผู้บริโภค ทางด้านผู้ประกอบการนำบทวิจารณ์สินค้าไปปรับปรุงแก้ไขสินค้าและบริการให้ดีขึ้น และเพิ่มประสิทธิภาพสินค้าให้ตรงตามความต้องการของลูกค้ามากขึ้น ส่วนทางด้านผู้บริโภคการอ่านบทวิจารณ์สินค้า เพื่อช่วยในการตัดสินใจซื้อสินค้า และบริการได้ง่าย สะดวก และตรงตามความต้องการมากขึ้น

ดังนั้นงานวิจัยนี้จึงได้จัดทำระบบสรุปความคิดเห็น โดยมีวิธีการคือ (1) ทำการเก็บรวบรวมข้อมูล โดยข้อมูลที่ใช้ในการทดลอง คือบทวิจารณ์สินค้าประเภทเครื่องสำอางที่เป็นภาษาไทย ใช้บทวิจารณ์ประเภทเครื่องสำอางจากมุมมองของผู้บริโภค โดยบทวิจารณ์ทั้งหมดนำมาจากเว็บไซต์ ที่มี

ชื่อเสียงด้านการแต่งหน้า (2) ใช้อัลกอริทึมนาอิวเบย์ (Naive Bayes) ในการจัดกลุ่มเพื่อจำแนกคุณสมบัติที่สนใจตามประเภทของผลิตภัณฑ์ จากนั้นสร้างกลุ่มประโยคที่มีความคิดเห็นเชิงบวก และเชิงลบ เพื่อวิเคราะห์ความรู้สึกของประโยค (3) หาความคล้ายคลึงของประโยค (Cosine similarity) จากนั้นใช้ PageRank algorithm, Hopfield Network algorithm และ Term Frequency Inverted Document Frequency เพื่อทำการสรุปประโยคที่เป็นตัวแทนของความคิดเห็นของผลิตภัณฑ์ ตามคุณสมบัติที่สนใจเป็นบทสรุปความรู้สึกต่อสินค้านั้นๆ

1.2 วัตถุประสงค์การวิจัย

1. เพื่อพัฒนาระบบสรุปบทวิจารณ์สินค้า
2. เพื่อสรุปบทวิจารณ์สินค้าสำหรับผู้บริโภค และผู้ประกอบการ
3. เพื่อศึกษาวิธีการระบุระดับความหนักแน่นของความคิดเห็นสำหรับนำไปใช้ในระบบสรุปบทวิจารณ์สินค้าที่เป็นภาษาไทย

มหาวิทยาลัยศิลปากร สวทช.วิจัย

1.3 ขอบเขตการศึกษา

1. เว็บไซต์ที่ใช้ในการทดลองเป็นเว็บไซต์เกี่ยวกับเครื่องสำอางที่เป็นภาษาไทยเท่านั้น
2. เก็บข้อมูลบทวิจารณ์จากเว็บไซต์ www.jeban.com เท่านั้น โดยใช้ข้อมูลบทวิจารณ์จากผลิตภัณฑ์เครื่องสำอาง 3 ผลิตภัณฑ์
3. งานวิจัยนี้จะศึกษาวิธีการระบุระดับความหนักแน่นของความคิดเห็นในบทวิจารณ์สินค้าของผู้บริโภคและผู้ประกอบการที่เป็นข้อความภาษาไทยเท่านั้น
4. เนื่องจากงานวิจัยใช้ข้อความแสดงความคิดเห็นมีลักษณะเป็นข้อความภาษาพูด ในบางครั้งอาจไม่ถูกต้องตามหลักไวยากรณ์ภาษาไทย เช่น การใช้ศัพท์แสลง ซึ่งข้อความนี้ส่งผลต่อความถูกต้องและแม่นยำของระบบที่ทำงานบนหลักไวยากรณ์ที่ถูกต้องเป็นหลัก

1.4 ขั้นตอนการศึกษา

1. ศึกษาวิธีการตัดคำ และกำหนดทิศทางความคิดเห็น
2. ศึกษาวิธีการสรุปบทวิจารณ์สินค้า

3. เก็บข้อมูลบทวิจารณ์จากเว็บไซต์ www.jeban.com แล้วเก็บไว้ในรูปของ Text File
4. เมื่อได้ข้อมูลที่ต้องการแล้ว ใช้โปรแกรมตัดคำจากข้อมูลที่เป็นบทวิจารณ์ จากนั้นทำการสร้างคลังข้อมูลระบุทิศทางของความคิดเห็น (Corpus Positive word, Corpus Negative word) เพื่อนำข้อมูลนี้ไปสร้างกลุ่มคำเพื่อระบุทิศทางของความคิดเห็น (Polarity) ซึ่งแบ่งเป็น 2 ทิศทาง คือทางบวกและทางลบ
5. จัดกลุ่มข้อมูลโดยใช้อัลกอริทึมนาอิวเบย์ (Naive Bayes) เพื่อสร้างกลุ่มของผลิตภัณฑ์แต่ละประเภทด้วยการนำไฟล์ประโยคที่ได้จากกระบวนการข้างต้น โดยในการจัดกลุ่มจะให้ผู้ที่เคยใช้ผลิตภัณฑ์เครื่องสำอางแต่ละประเภทจำนวน 10 คน เป็นผู้จัดกลุ่มข้อมูล โดยการจัดกลุ่มข้อมูลผลลัพธ์ที่ได้มีทั้งหมดดังนี้

ที่ปิดแก้ม (Blush On) จัดกลุ่มได้ทั้งหมด 3 กลุ่ม ได้แก่

“ติดทน”

“สี/เมื่อดสี”

“แพคเกจ”

ผลิตภัณฑ์ทำความสะอาดผิวหน้า (Cleansing) จัดกลุ่มได้ทั้งหมด 4 กลุ่ม ได้แก่

“กลิ่น”

“ความสะอาด”

“แพ็/อ่อนโยน”

“ราคา”

ครีมรองพื้น (Foundation) จัดกลุ่มได้ทั้งหมด 5 กลุ่ม ได้แก่

“เกลี่ยง่าย”

“ปกปิด”

“ควบคุมความมัน”

“กันน้ำ”

“ติดทน”

6. สร้างกลุ่มความรู้สึกของประโยค (Opinion word list)
7. หาค่าความคล้ายคลึง (Similarity)

8. ใช้อัลกอริทึม PageRank และ อัลกอริทึม Hopfield ในการสรุปและเลือกประโยคที่ดีที่สุด ในกรณีที่คะแนนเท่ากันจะใช้ การหาค่าความถี่ของประโยค (TF-IDF)
9. ออกแบบและพัฒนาระบบที่ใช้ในการสรุปทวิจรรย์สินค้าโดยผู้บริโภค

1.5 เครื่องมือและอุปกรณ์

1. หน่วยประมวลผลข้อมูล 2.6 GHz Intel Core i5
2. ภาษาที่ใช้ในการพัฒนาระบบ: Python
3. ภาษาที่ใช้ในการพัฒนาโปรแกรม: PHP Script Language Version 6.0.0-dev
4. ฮาร์ดดิสก์ ขนาด 1 TB

1.6 ประโยชน์ที่คาดว่าจะได้รับ

1. สามารถนำไปพัฒนาต่อยอดสำหรับระบบสรุปทวิจรรย์สินค้าและบริการให้มีประสิทธิภาพดียิ่งขึ้น
2. ผู้บริโภค และผู้ประกอบการได้ข้อมูลสรุปทวิจรรย์สินค้าที่ละเอียดแม่นยำมากขึ้น สำหรับผู้บริโภคนำไปใช้ประกอบการตัดสินใจซื้อสินค้าและบริการ และผู้ประกอบการสามารถใช้เป็นข้อมูลประกอบการวางแผนพัฒนากลยุทธ์การตลาดให้ดียิ่งขึ้น

บทที่ 2 วรรณกรรมที่เกี่ยวข้อง

2.1 บทนำ

ในบทนี้เป็นการนำเสนอวรรณกรรมในอดีตที่เกี่ยวข้องกับการพัฒนาระบบสรุปบทวิจารณ์สินค้าและบริการ วิธีการระบุความหนักแน่นของความคิดเห็นที่เกี่ยวข้อง ได้แก่ การสรุปความคิดเห็นของผู้บริโภค การระบุระดับความหนักแน่นของความคิดเห็น และการวัดประสิทธิภาพของระบบสรุปบทวิจารณ์สินค้า

2.2 การสรุปความคิดเห็นของผู้บริโภค

2.2.1 Query-Based Summarization : A survey

จากงานวิจัยของ Mariana Damova and Ivan Koychev (2010: 142-146) ได้ทำการสรุปเกี่ยวกับงานวิจัยทางการสรุปข้อความ โดยวิธีที่ใช้ในการสร้างการสรุปผล มีอยู่ 2 วิธี

1. Abstract summaries เป็นการสร้างการสรุปผลจากสิ่งสำคัญที่เรารวบรวม และประมวลผลแล้ว นำมาทำการสรุปผลสั้นๆจากความคิดเห็นของเราเอง (ความคิด ความเห็น ที่บ่งบอกความเป็นตัวเอง)

2. Extractive summaries เป็นการระบุสิ่งสำคัญ และดึงส่วนที่สำคัญออกมา เพื่อใช้ในการสรุปโดยเลือกสิ่งสำคัญบางส่วนในเอกสารโดยที่ดึงข้อมูลที่อยู่ในเอกสารนั้นมาสรุปผล

แหล่งที่มาของเอกสารที่เรานำมาใช้ในการสรุปผล แบ่งออกเป็น 2 แบบ คือ

1. Single document การสรุปจากแหล่งที่มาจากเอกสารฉบับเดียว
2. Multi-document การสรุปจากแหล่งที่มาจากเอกสารหลายฉบับ

ปัญหาที่พบคือ

1. การยอมรับความซ้ำซ้อนของเอกสาร
2. การระบุความสำคัญที่แตกต่างกันของแต่ละเอกสาร
3. สร้างความเชื่อมั่นจากการสรุปที่สอดคล้อง และเชื่อมโยงกัน

วิธีการที่ใช้สรุปผลแบบ query-based มี 3 วิธี ดังนี้

1. Approaches based on Document Graphs

จากงานวิจัย Ahmed A. Mohamed and Sanguthevar Rajasekaran (2006) เสนอการแยกเอกสารที่มีหลายเอกสารด้วยวิธีการสรุป เพื่อแสดงให้เห็นถึงผลลัพธ์ของ document graph ที่ได้มาจากข้อความในเอกสาร โดยขั้นตอนมีดังนี้

1. เริ่มจากการตัดคำ (Token) แล้วแยกแต่ละคำเป็น NPs (Net Promoter Score) เป็นตัววัดผลตัวหนึ่งซึ่งเป็นที่นิยมในการวัดความพึงพอใจและความซื่อสัตย์ของลูกค้าได้แก่ ตัววัดที่เรียกชื่อว่า เปรียบเทียบ Centric Graph ของเอกสารจากแนวคิดเห็น

2. สร้างความสัมพันธ์ตาม Heuristic Rules (กฎการแก้ไขปัญหา)

สำหรับวิธีการที่จะทำเพื่อให้ได้มาซึ่งผลลัพธ์มี 3 วิธี

1. การเปรียบเทียบ Centric Graph กับ เอกสารคำถาม (นำเอกสารมาเทียบกับกราฟ)

2. การสร้างกราฟของเอกสาร และกราฟของคำถามขึ้นมา และประมาณค่าของประโยคคำถามเพื่อให้คล้ายคลึงกันแต่ละประโยค และนำกราฟของเอกสาร และกราฟของคำถามมาเปรียบเทียบกัน

3. เทคนิคการปรับเปลี่ยนคำถาม โดยใช้กราฟในการเลือกประโยคจากกราฟคำถาม สำหรับวิธีที่ดีที่สุด คือ วิธีที่ 2 และวิธีการหาคำตอบ จะใช้ RST (Rhetorical Structure Theory) โดยการสร้างกราฟ และหาความสัมพันธ์กันระหว่างประโยค และเป็นการให้ค่าน้ำหนัก เพื่อหาค่าน้ำหนักของ 2 ประโยค แล้วนำมาดูถึงความแตกต่างของน้ำหนักระหว่างประโยคคำถามและประโยคคำตอบ เพื่อหาค่าน้ำหนักที่น้อยที่สุดเป็นผลลัพธ์สุดท้าย

2. Approaches using linguistics

1. John M. Conroy et al. (2005) ใช้วิธีการ HMM (Hidden Markov Model) ในการเลือกประโยค แล้วใช้อัลกอริทึมในการสร้างสรุปผลจากเอกสารหลายๆเอกสาร กระบวนการสร้างชุดข้อมูลของเอกสารซึ่งมีความหลากหลายของกลุ่ม เช่น สถานที่ คน วันที่ ฯลฯ การประมวลผลทางภาษาจะใช้การสร้าง Query Term โดยจะใช้โมเดลของ Hidden Markov Model ในการให้คะแนนของแต่ละประโยค รวมทั้งการแก้ปัญหาและปรับปรุงโครงสร้างให้ดีขึ้นรวมถึงความสามารถในการนำไปประยุกต์ใช้

2. เป็นวิธีการสรุป Multi-Document โดยใช้การแปลความหมายของแบบสอบถาม และวิเคราะห์เพื่อหารายละเอียด ของแต่ละหัวข้อ แล้วทำการจัดกลุ่มเอกสาร และสร้างการสรุปผล

3. Machine-learning approaches

1. Frank Schilder and Ravikumar Kondadadi (2008) เป็นวิธีที่ได้จากการค้นหาข้อมูล ร่วมกับเทคนิคในการสรุปผลลัพธ์ที่ได้จากการแยกข้อความ หรือการสกัดสิ่งที่สำคัญ วิธีนี้ประกอบด้วย แนวความคิดใหม่ของข้อความที่ไม่ขึ้นอยู่กับคำถามที่อยู่ในผลลัพธ์สุดท้าย เป็นการให้คะแนนประโยค จะใช้ set ของคุณสมบัติจากหลายๆประโยค เช่น การให้คะแนนมากที่สุด (เป็นการคำนวณน้ำหนักจาก คะแนนของแต่ละประโยค) โดยเลือกจากประโยคที่มีคะแนนสูงสุดมาเป็นผลลัพธ์สุดท้าย

การกำหนดค่าคุณสมบัติของประโยคใหม่ มี 2 วิธีคือ

1. คำถามที่ขึ้นอยู่กับการจัดลำดับของ เอกสาร/ประโยค
 2. จัดตามความสำคัญที่ชัดเจนของ เอกสาร/ประโยค
2. เป็นวิธีที่ใช้ความถี่ของคำที่ปรากฏในเอกสารเพื่อจัดกลุ่ม โดยมีกระบวนการดังนี้
 1. ทำการจัดกลุ่มคำโดยใช้ Ontology ของ Cluster ในการจัดกลุ่ม
 2. ทำการเรียงลำดับความสำคัญของเอกสาร/ประโยค โดยใช้ SVM (Support Vector Machine)

การใช้ application ในการปรับปรุงระบบ มีอยู่ 2 ระบบ ดังนี้

1. Subject domain ontology based approach

Query-based Medical Information Summarization System จะใช้ Ontology Knowledge ทางด้านแพทย์ในการเก็บรวบรวมข้อมูล และ Ping Chen, Rakesh Verma ได้เสนอ เทคนิคของ UMLS และ Ontology จาก National Library of Medicine และกระบวนการสรุป คือ การแก้ไขคำถามด้วย UMLP Ontology Knowledge

คำนวณช่วงหรือระยะห่างของประโยคในเอกสาร final query

คำนวณการจับคู่ระหว่างประโยคที่อยู่ในหมู่เดียวกัน โดยจะแบ่งออกเป็นแต่ละกลุ่ม และ เลือกการจัดลำดับที่อยู่สูงสุดของแต่ละประโยค

2. Opinion Summarization

ระบบนี้สามารถแยกแยะ แนวโน้มกระแสข่าวที่เพิ่มมากขึ้น และสิ่งที่สำคัญมาก คือ เวลา และระบบนี้สามารถรายงานเป็นการสรุปจากความรู้สึกโดยอัตโนมัติและข้อมูลที่ป้อนเข้าไปจะเป็นการสรุปมาจากงานที่มีความคิดเห็นที่สัมพันธ์กันกับคำถาม และเป้าหมายของผลลัพธ์คือคำตอบตรงกับคำถาม โดยจะใช้การบ่งบอกความรู้สึกทางบวกและทางลบ (Positive and Negative Polarity) ในการสร้างการสรุปผล แล้วเรียงลำดับของประโยคที่สร้างขึ้นมา

อย่างไรก็ตามการวิเคราะห์ความคิดเห็นบนเครือข่ายสังคมออนไลน์นั้น ไม่สามารถมองเป็นสองแบบนี้ได้ชัดเจนขึ้นอยู่กับเป้าหมายที่สนใจ สำหรับการวิเคราะห์ความคิดเห็นในงานวิจัยนี้มองว่าข้อความแสดงความคิดเห็นส่วนใหญ่ที่อยู่บนอินเทอร์เน็ตนิยมใช้ภาษาที่มีโครงสร้างประโยคที่ไม่แน่นอน (Unstructured data) หรือเป็นภาษาธรรมชาติ (Natural language) ไม่ถูกต้องตามหลักไวยากรณ์ทางภาษา ทำให้ยากต่อการวิเคราะห์งานวิจัยด้านการวิเคราะห์ความคิดเห็น จึงได้นำเทคนิคการวิเคราะห์เหมืองข้อความ (Text mining) และการประมวลผลภาษาธรรมชาติ (Natural Language Processing: NLP) มาประยุกต์ใช้ เพื่อวิเคราะห์ความคิดเห็นของผู้บริโภค ซึ่งเรียกว่าการวิเคราะห์เหมืองความคิดเห็น (opinion mining) หรือการวิเคราะห์ความรู้สึก (Sentiment analysis) ซึ่งงานวิจัยต่าง ๆ ได้นำเสนอการวิเคราะห์ความพึงพอใจของผู้บริโภคในหัวข้อเรื่องที่สนใจ อาทิ สินค้า บริการ ข่าว และ การเมือง ซึ่งบทความนี้จะอธิบายแนวคิดทฤษฎี เทคนิคต่าง ๆ ที่เกี่ยวข้อง รวมถึงกระบวนการในการวิเคราะห์การสร้างแบบจำลอง เช่น ในงานวิจัยของ Trithep Thumrongluck (2010) ได้นำเสนอการวิเคราะห์ความคิดเห็นในระดับเอกสารโดยใช้คำวิเศษณ์ หรือวลีวิเศษณ์ คำคุณศัพท์ คำกริยาวิเศษณ์ และคำนามหรือนามวลีในการบ่งชี้หัวข้อที่สนใจ และสรุปความคิดเห็นของเอกสารเป็นความรู้สึกเชิงบวก (positive) หรือเชิงลบ (negative) ด้วยวิธีการนับจำนวนคำที่แสดงความคิดเห็นเชิงบวกและเชิงลบ หากความคิดเห็นใดมีจำนวนมากกว่า จะสรุปเป็นความคิดเห็นของเอกสารนั้น และงานวิจัยของ Parnicha Apisuwankun and Janjao Mongkolnavin (2013) นำเสนอการวิเคราะห์เหมืองความคิดเห็นในระดับคุณลักษณะ (feature level) ซึ่งทำให้ผู้อ่านทราบว่าผู้แสดงความคิดเห็นมีทัศนคติ เชิงบวก เชิงลบ หรือเป็นกลางต่อคุณลักษณะของหัวข้อที่สนใจ และยังมีงานวิจัยอีกมากมายได้นำแนวคิดทฤษฎี เทคนิคต่าง ๆ ที่เกี่ยวข้อง รวมถึงกระบวนการในการวิเคราะห์การสร้างแบบจำลอง ซึ่งมารายละเอียดดังนี้

2.2.2 An Automated System for Summarizing Structured Product Reviews

จากวิทยานิพนธ์ของ Trithep Thumrongluck (2010) มีวิธีสร้างรายการคำแสดงคุณลักษณะของสินค้าจะทำโดยสกัดคุณลักษณะของสินค้า (Feature Extraction) ส่วนการค้นหาคำแสดงคุณลักษณะจากบทวิจารณ์จะใช้คำที่มีหน้าที่คำเป็นคำนาม (Noun) และคำกริยา (Verb) การสร้างรายการคำแสดงคุณลักษณะในภาษาไทย เนื่องจากภาษาไทยเป็นภาษาที่ไม่สามารถหาจุดสิ้นสุดของประโยคได้อย่างชัดเจน จึงไม่สามารถสกัดคุณลักษณะของสินค้าโดยวิธีการหาคุณลักษณะของสินค้าที่มีการประมวลผลแบบประโยคต่อประโยค จึงได้การสกัดคุณลักษณะของสินค้าจากบทวิจารณ์โดยใช้วิธีการดึงคำสำคัญ (Key Word) ซึ่งเป็นวิธีการหาความถี่หรือคะแนนของคำนามและคำกริยาที่ปรากฏในบทวิจารณ์สินค้าของแต่ละสินค้า โดยนำค่าที่เอฟไอดีเอฟ TF-IDF (Term Frequency-

Inverse Document Frequency) ซึ่งเป็นค่าความถี่ของคำที่ปรากฏในเอกสารมาปรับใช้กับคำที่ปรากฏในบทวิจารณ์ของสินค้าที่มีมากกว่า 1 สินค้า เรียกว่า มาตรฐานค่าคะแนนทีเอฟไอซีเอฟ TFICF (Term Frequency Inverse Class Frequency) โดยนำค่าทีเอฟไอซีเอฟของคำในบทวิจารณ์ของแต่ละสินค้ามาเปรียบเทียบกับกัน ถ้าคำในสินค้าใดมีค่าคะแนนมากกว่าคำเดียวกันในสินค้าอื่น คำนั้นจะถูกจำแนกออกมาเป็นคุณลักษณะของสินค้านั้น

การระบุทิศทางของคำแสดงความคิดเห็น เริ่มจากการสร้างรายการต้นกำเนิดซึ่งเป็นรายการเริ่มต้นสำหรับใช้ระบุทิศทางให้กับคำใหม่ที่จะเพิ่มเข้ามาในรายการคำแสดงความคิดเห็น โดยใช้เวิร์ดเน็ต (Wordnet) ที่มีลักษณะเป็นโครงสร้างคำคุณศัพท์สองขั้ว (Bipolar Adjective Structure) ในการค้นหาคำที่มีความหมายคล้ายคลึง (Synonym) หรือคำที่มีความหมายตรงข้าม (Antonym) เพื่อให้สามารถระบุทิศทางของคำแสดงความคิดเห็นได้ โดยหากคำใหม่มีความสัมพันธ์แบบพ้องความหมาย (Synonym) กับคำในรายการต้นกำเนิด คำใหม่นั้นจะถูกระบุทิศทางให้เป็นทางเดียวกับคำในรายการต้นกำเนิด แต่หากคำใหม่มีความหมายตรงข้ามกับคำในรายการต้นกำเนิด คำใหม่นั้นจะถูกระบุทิศทางเป็นทางตรงข้ามกับทิศทางของคำนั้นในรายการต้นกำเนิด ตัวอย่างรายการคำแสดงความคิดเห็น แสดงในตารางที่ 1 และ ตารางที่ 2

มหาวิทยาลัยศิลปากร สมุทรปราการ

ตารางที่ 1 ตัวอย่างรายการคำแสดงความคิดเห็นเชิงบวกจากงานวิจัยของ (Trithep Thumrongluck, 2010)

กลุ่มบรัชออนทาแก้ม	กลุ่มโลชั่นทำความสะอาดหน้า	กลุ่มครีมรองพื้น
สวย	ถูก	ถูก
ดีดทน	หอม	บางเบา
ถูก	สะอาด	หอม
นาน	สวย	เนียน
ละเอียด	ง่าย	เรียบเนียน
มาก	ดี	สวย
ดี	เบาบาง	ดี
ง่าย	เนียนนุ่ม	เหลว

ตารางที่ 2 ตัวอย่างรายการคำแสดงความคิดเห็นเชิงลบจากงานวิจัยของ (Trithep Thumrongluck, 2010)

กลุ่มบริษัชอณาแก้ม	กลุ่มโลชั่นทำความสะอาดหน้า	กลุ่มครีมรองพื้น
แพง	แพง	แพง
ง่าย	เหม็น	เหนียว
น้อย	เหนียว	เหนอะหนะ
หยาบ	ฉุน	น้อย
ยาก	น้อย	เหม็น
เร็ว	ยาก	ยาก
เปราะบาง	ง่าย	ช้า
ยาก	แพ้	ฉุน

2.2.3 Opinion Strength Identification in Customer Review Summarizing System Using Association Rule Technique

จากงานวิจัยของ Parnicha Apisuwankun and Janjao Mongkolnavin (2013) ได้เห็นถึงความสำคัญของการสรุปความคิดเห็นของลูกค้าที่มีต่อสินค้าและบริการ จึงได้ทำระบบสรุปความคิดเห็นเกี่ยวกับเครื่องสำอางโดยใช้กฎของความสัมพันธ์ (Association Rule Technique) ในการระบุระดับความหนักแน่นของความคิดเห็นมาใช้ในการสรุปความคิดเห็น โดยใช้ข้อมูลการระบุความคิดเห็นของคำจาก (Trithep Thumrongluck, 2010) ตัวอย่างสัญลักษณ์การระบุความหนักแน่นของความคิดเห็นแสดงดังรูปที่ 1

Symbols	Meaning
+	"Slightly positive opinion"
++	"positive opinion"
+++	"very strong positive opinion"
-	"Slightly negative opinion"
--	"negative opinion"
---	"very strong negative opinion"

รูปที่ 1 สัญลักษณ์และความหมายที่ใช้ในการให้ความหนักแน่นของคำ

ขั้นตอนการวิจัย

1. Data Collection

สร้างแบบสอบถามที่ใช้ในการเก็บรวบรวมข้อมูลคำและคะแนนความหนักแน่นของคำ จากบทวิจารณ์สินค้าที่ได้เลือกใช้เพื่อรวบรวมรายการคำ และคะแนนความหนักแน่นของคำ จำนวน 180 บทวิจารณ์ จะนำมาสร้างแบบสอบถามสำหรับเก็บข้อมูลจำนวน 9 ชุด โดยแบ่งตามกลุ่มเครื่องสำอาง และยี่ห้อสินค้า 3 กลุ่ม กลุ่มละ 3 ยี่ห้อ ยี่ห้อละ 20 บทวิจารณ์

ข้อ	ข้อความแสดงความคิดเห็น
	<p style="text-align: center;">++ + +++</p> <p style="text-align: center;">ชอบกลิ่นมาก หอมอ่อนๆ ถูกใจมาก</p>
	<p style="text-align: center;">-</p> <p style="text-align: center;">แป้งสีไม่ขาวเท่าไหร่ (เลือกสีขาวสุดแล้วนะ)</p>
	<p style="text-align: center;">- ++ +++ --</p> <p style="text-align: center;">ไม่ถึงกะหมอง ดูเป็นธรรมชาติ เนียนมาก แต่ไม่ปกปิด</p>

รูปที่ 2 ตัวอย่างการระบุทิศทางของความคิดเห็น

2. Data Preparation

นำข้อมูลที่เก็บรวบรวมมาจากตัวอย่างนำมาระบุความหนักแน่นของความคิดเห็นดังรูปที่ 2 แล้วนำความคิดเห็นทั้งหมดมาทำการตัดคำ และกำหนดหน้าที่ของคำ โดยเป็นการนำข้อความบทวิจารณ์สินค้าจากเว็บตัวอย่างเครื่องสำอางมาตัดคำ ตัดคำโดยใช้โปรแกรม Lexto ดังตัวอย่างรูปที่ 3 และกำหนดหน้าที่คำโดยใช้ฐานข้อมูล Lexitron

ID	Word	Strength
1	ชอบ กลิ่น มาก	2
2	หอม	1
3	ถูกใจ มาก	3
4	ไม่ ขาว เท่าไหร่	-1
5	ธรรมชาติ	2

รูปที่ 3 ตัวอย่างระดับความหนักแน่นของความคิดเห็น

3. Data Mining Process

นำข้อมูลที่จัดเตรียมไว้ซึ่งประกอบไปด้วยค่าและคะแนนความหนักแน่นของค่า มาสร้างกฎในการระบุความหนักแน่นความคิดเห็น โดยใช้เทคนิคกฎความสัมพันธ์ (Association Rule) ในการหาความสัมพันธ์ของค่ากับคะแนนความหนักแน่นของค่า ดังตัวอย่างรูปที่ 4

No.	Confident	Suport	Lift	Count	Rule
1	30.36	4.92	2.09	575	ไม่ ==> -1
2	35.14	4.38	1.29	512	มาก ==> 2
3	50.26	4.18	1.61	488	ดี ==> 1
4	34.21	3.12	1.25	364	ชอบ ==> 2
5	50.74	2.34	3.50	273	ไม่ & ค่อย ==> -1
6	44.93	2.20	1.44	257	สวย ==> 1
7	46.64	2.02	1.49	236	สะอาด ==> 1

รูปที่ 4 ตัวอย่างที่ได้จากกฎความสัมพันธ์

ข้อจำกัด และส่วนที่สามารถพัฒนาต่อได้

จากผลการวิจัยและข้อจำกัดที่สรุปได้ในงานวิจัยนี้ สามารถเป็นแนวทางในการศึกษาต่อเนื่องกับผู้สนใจศึกษาพัฒนาต่อยอดให้ดียิ่งขึ้น ดังนี้

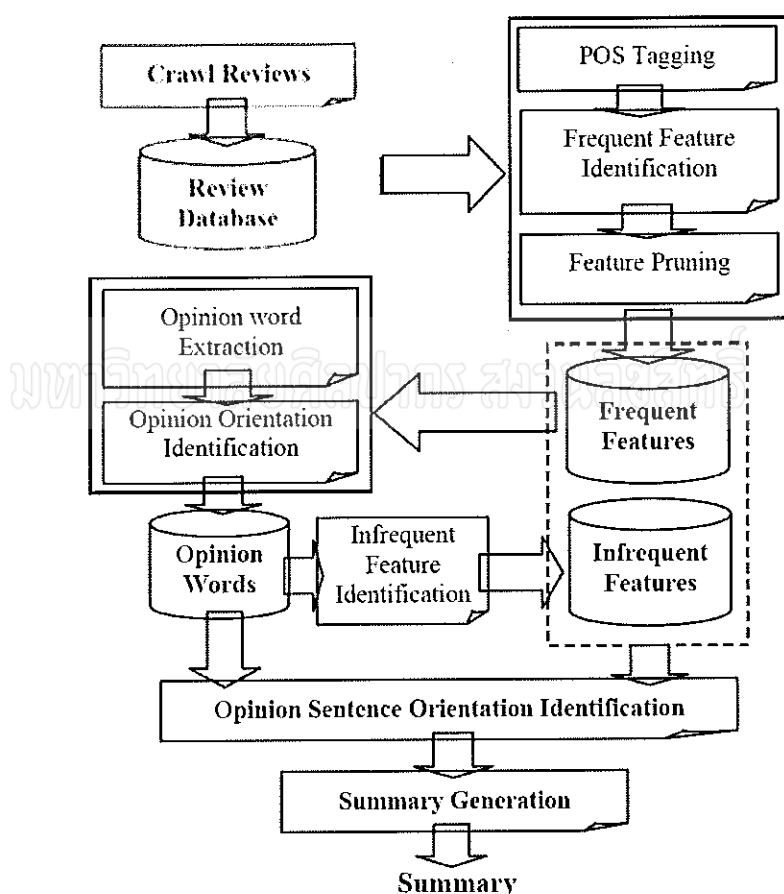
1. พัฒนาพจนานุกรมคำศัพท์ที่ใช้ในระบบสรุปบทวิจารณ์สินค้าสำหรับใช้เฉพาะในการสรุปบทวิจารณ์สินค้าโดยตรง เนื่องจากบางคำมีความหมายและหน้าที่คำมากกว่า 1 หน้าที่ และอาจเป็นคำที่ไม่เกี่ยวกับบทวิจารณ์สินค้า รวมถึงศึกษาหาคำที่ไม่เป็นทางการที่มักใช้ในการแสดงความคิดเห็น และเพิ่มค่าน้ำลงในฐานข้อมูล

2. พัฒนาวิธีการค้นหาคุณลักษณะของสินค้าให้สามารถดึงคุณลักษณะของสินค้าที่เป็นคุณลักษณะทั่วไปหรือเป็นคุณลักษณะของสินค้ามากกว่า 1 กลุ่มสินค้าได้ เช่น ราคา กลิ่น สี เป็นต้น ซึ่งจะทำให้ระบบสามารถดึงคุณลักษณะของสินค้าออกมาได้ใกล้เคียงกับมนุษย์มากขึ้น อาจเพิ่มจำนวนบทวิจารณ์ตัวอย่างในการนำขั้นตอนวิธีการสร้างกฎการระบุความหนักแน่นของความคิดเห็นในงานวิจัยนี้ไปประยุกต์ใช้ ซึ่งอาจทำให้พบค่าบอกระดับความหนักแน่นของความคิดเห็นมากขึ้น

2.2.4 Mining and Summarizing Customer Reviews

จากงานวิจัยของ Minqing Hu and Bing Liu (2004: 168-177) ได้พัฒนาระบบสรุปความคิดเห็น เพื่อผู้บริโภคได้เข้าถึงความคิดเห็นที่มีทั้งข้อดี ข้อเสีย ในแง่มุมต่างๆ สำหรับนำมาใช้ในการตัดสินใจเลือกซื้อสินค้า ส่วนผู้บริโภคก็จะเห็นถึงข้อดี ข้อเสียของผลิตภัณฑ์นั้นๆ จากความคิดเห็นของผู้ใช้ และในด้านของผู้ผลิตก็จะได้นำข้อเสียมาปรับปรุงสินค้าและบริการต่างๆ ได้

ขั้นตอนการทำงาน



รูปที่ 5 ภาพรวมการทำงานของ Mining and Summarizing Customer Reviews

1. POS (Part-of-Speech Tagging) เป็นการแยกกลุ่มคำ โดยใช้ NLPProcessor linguistic parser โดยจะแบ่งออกเป็น Noun, Verb, Adjective ดังตัวอย่างในรูปที่ 6

```

<S> <NG><W C='PRP' L='SS' T='w' S='Y'> I </W> </NG>
<VG> <W C='VBP'> am </W><W C='RB'> absolutely
</W></VG> <W C='IN'> in </W> <NG> <W C='NN'> awe
</W> </NG> <W C='IN'> of </W> <NG> <W C='DT'> this
</W> <W C='NN'> camera </W></NG><W C='.'> .
</W></S>

```

รูปที่ 6 ตัวอย่างการระบุหน้าที่ (POS tags)

2. Frequent Features Identification เป็นการระบุความถี่ของคุณสมบัติ เช่น

Picture ----> The pictures are very clear.

Size ----> While light, it will not easily fit in pockets.

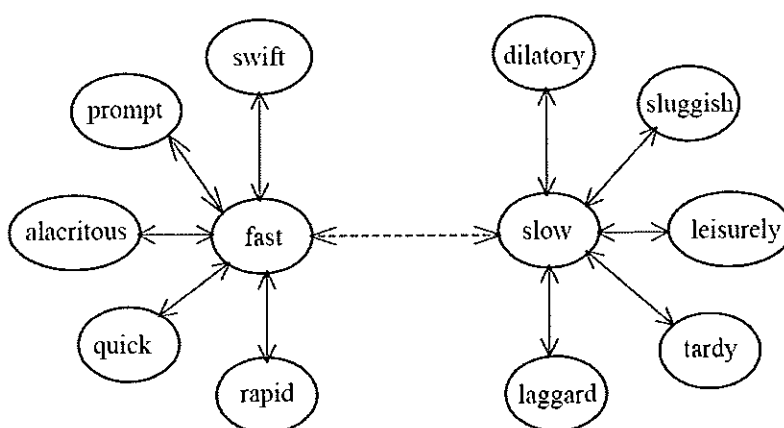
3. Opinion Words Extraction And Orientation Identification for Opinion Words

เป็นการสกัดคำจากความคิดเห็น แล้วนำมาระบุความเห็นของคำโดยการกำหนดความหมายของคำแต่ละคำว่าเป็น Positive หรือ Negative เช่น

Positive ----> adjectives: great, fantastic, nice, cool.

Negative ----> adjectives: bad, dull.

เมื่อทำการกำหนดคำแล้วจะนำ WordNet มาช่วยในการจัดกลุ่มคำที่มีความหมายเหมือนกันและตรงข้ามกัน (Similarity, Antonym) ดังรูปที่ 7



รูปที่ 7 ตัวอย่างการจัดกลุ่มคำที่มีความหมายเหมือนกันและตรงข้ามกัน

4. Predicting the Orientations of Opinion Sentences เมื่อได้กลุ่มคำที่แบ่งเป็นกลุ่ม Positive และ Negative แล้วนำคำมาเทียบกับประโยคเพื่อหาความคิดเห็นของประโยคนั้นว่ามีความคิดเห็นต่อสินค้าในทางบวกหรือทางลบ

5. Summary Generation ผลสรุปจากระบบแสดงความคิดเห็นนี้จะได้ความคิดเห็นที่เป็นบวกและความคิดเห็นที่เป็นลบ ของผู้บริโภคที่มีต่อสินค้านั้น ตัวอย่างประโยคสรุปของความคิดเห็นแสดงดังรูปที่ 8

Feature: picture

Positive: 12

- Overall this is a good camera with a really good picture clarity.
- The pictures are absolutely amazing - the camera captures the minutest of details.
- After nearly 800 pictures I have found that this camera takes incredible pictures.

...

Negative: 2

- The pictures come out hazy if your hands shake even for a moment during the entire process of taking a picture.
- Focusing on a display rack about 20 feet away in a brightly lit room during day time, pictures produced by this camera were blurry and in a shade of orange.

รูปที่ 8 ตัวอย่างประโยคสรุปของความคิดเห็น

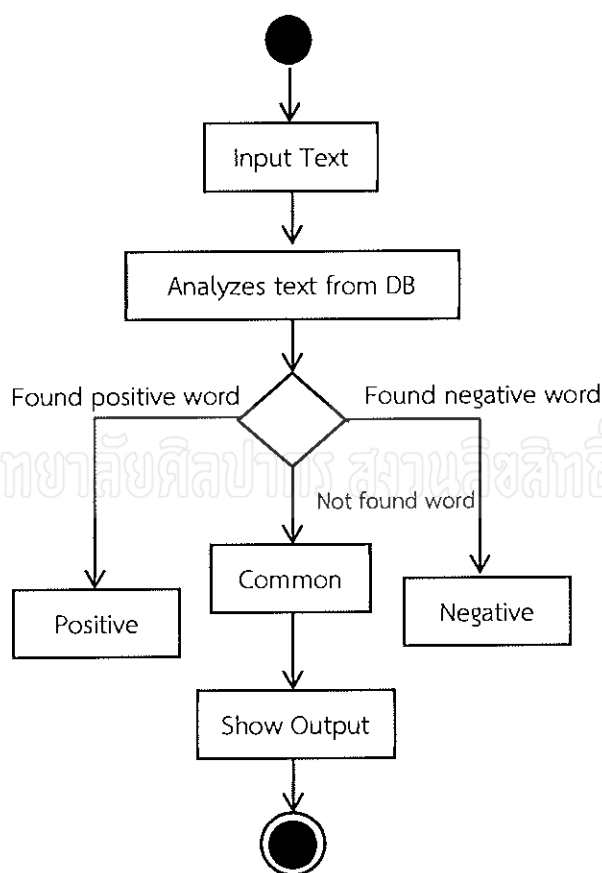
สรุป

ในงานวิจัยนี้ เสนอเทคนิคของ Mining and Summarizing Product Reviews ด้วยวิธีการของ Mining and NLP (Natural Language Processing) เพื่อสรุปความคิดเห็นให้ลูกค้าตัดสินใจง่ายขึ้น งานในอนาคตวางแผนที่จะปรับปรุง และปรับแต่งเทคนิคและการจัดการกับปัญหาให้ดีขึ้น ระบุคำสรรพนาม กำหนดความหนักแน่นของความคิดเห็น และตรวจสอบคำวิเศษณ์คำกริยาและคำนาม

2.2.5 Emotion Categorization System for the Thai Texts

จากงานวิจัยของ Lalita Wonghaeon et al. (2016) มีวัตถุประสงค์เพื่อพัฒนาระบบสำหรับวิเคราะห์อารมณ์จากข้อความตัวอักษรภาษาไทยที่ปรากฏบนสื่อสังคมออนไลน์ ระบบจะวิเคราะห์ข้อความภาษาไทย และสามารถพิจารณาแนวโน้มของอารมณ์ในเชิงบวกหรือเชิงลบ เพื่อช่วยวิเคราะห์หา ข้อความนั้นสื่ออารมณ์ออกมาในลักษณะใด

วิธีการดำเนินการวิจัย



รูปที่ 9 แผนภาพกิจกรรมระบบจำแนกระดับอารมณ์สำหรับข้อความภาษาไทย

เก็บรวบรวมสถานะจาก Facebook กระทั่งจากเว็บไซต์ Pantip และจากแหล่งอื่น ๆ และนำมาคัดแยกออกเป็น Emoticon เชิงบวกและเชิงลบ จากนั้นแล้วให้กรรมการ จำนวนสามคนพิจารณาเพื่อให้คะแนนโดย จะมีเกณฑ์การกำหนดคะแนนที่จะนำมาคำนวณต่อได้ คือ หากพิจารณาแล้วมีความเห็นว่าสถานะ เป็นเชิงบวก จะต้องให้คะแนนเท่ากับ 2 หากเป็นเชิงลบค่าคะแนนจะเป็น 2 แต่หากพิจารณาแล้วไม่มีแนวโน้มไปทั้งเชิงบวกและเชิงลบ จะให้ค่าคะแนนเท่ากับ 1 เมื่อให้คะแนนเสร็จเรียบร้อยแล้วจะต้อง นำคำที่พิจารณาว่าเป็นความคิดเห็นเชิงบวกหรือเชิงลบ มาแสดงไว้ประกอบเป็นเหตุผลด้วย เพื่อนำมาสรุปเป็นสถิติและนำไปเก็บรวบรวมเพื่อการวิเคราะห์ระบบต่อไป หลังจาก

ขั้นตอนการเก็บข้อมูล และวิเคราะห์ระบบแล้ว ก็มาถึงขั้นตอนของการพัฒนาตัวเว็บแอปพลิเคชัน เพื่อให้ระบบสามารถใช้งานได้ ซึ่งมีตัวช่วยในการ พัฒนาคือ Bootstrap Framework ที่ช่วยให้การพัฒนาที่มีความง่าย และสมบูรณ์มากขึ้น

2.2.6 Automatically Rating of Hotel Service Aspects from Textual Customer Reviews

จากงานวิจัยของ Nattakit Srikarnjanapert and Jantima Polpinij (2015) นำเสนอ กระบวนการของการจัดระดับความคิดเห็นเกี่ยวกับบริการของโรงแรมแบบอัตโนมัติจากบทวิจารณ์ ของลูกค้า กระบวนการที่นำเสนอการประยุกต์ใช้อัลกอริทึมนาอูฟเวียร์ร่วมกับลาพราซสมูทติ้ง และการใช้คุณลักษณะร่วมกับ tf-idf สำหรับการจัดระดับความคิดเห็นต่อบริการของโรงแรมแบบอัตโนมัติ จากบทวิจารณ์ของลูกค้า

การวิเคราะห์ความรู้สึก (Sentiment Analysis) หรือ การวิเคราะห์ความคิดเห็น (Opinion Analysis) เป็นการวิเคราะห์ และทำความเข้าใจความรู้สึก เกี่ยวกับสิ่งที่คนเหล่านั้นกำลังสนใจ โดยเฉพาะเรื่องของสินค้าและบริการ การท่องเที่ยว ภาพยนตร์ ข่าว หรืออื่นๆ โดยกระบวนการ ทำงานเริ่มจากการเก็บรวบรวมข้อมูลแสดงความคิดเห็นโดยทั่วไปการวิเคราะห์ความรู้สึกมักจะ เกี่ยวข้องกับการวิเคราะห์ข้อความ (Text) ซึ่งจัดได้ว่าเป็นส่วนหนึ่งของงานด้านการประมวลผล ภาษาธรรมชาติ (Natural Language Processing: NLP)

คลังคำแสดงความรู้สึก (Sentiment Word Polarity Corpora)

ในงานวิจัยนี้ได้เลือกใช้คลังคำ แสดงความรู้สึกในภาษาอังกฤษที่มีชื่อว่า SentiWordNet (Stefano Baccianella et al, 2010) แต่สำหรับภาษาไทย โดยได้พัฒนาคลังคำแสดงความรู้สึกด้วย การใช้เทคนิคของค่าความน่าจะเป็น และใช้ค่าสถิติ Cohen's Kappa

วิธีการดำเนินงานวิจัย (Research Methodology)

ในส่วนนี้จะอธิบายขั้นตอนการดำเนินงานดังรูปที่ 10

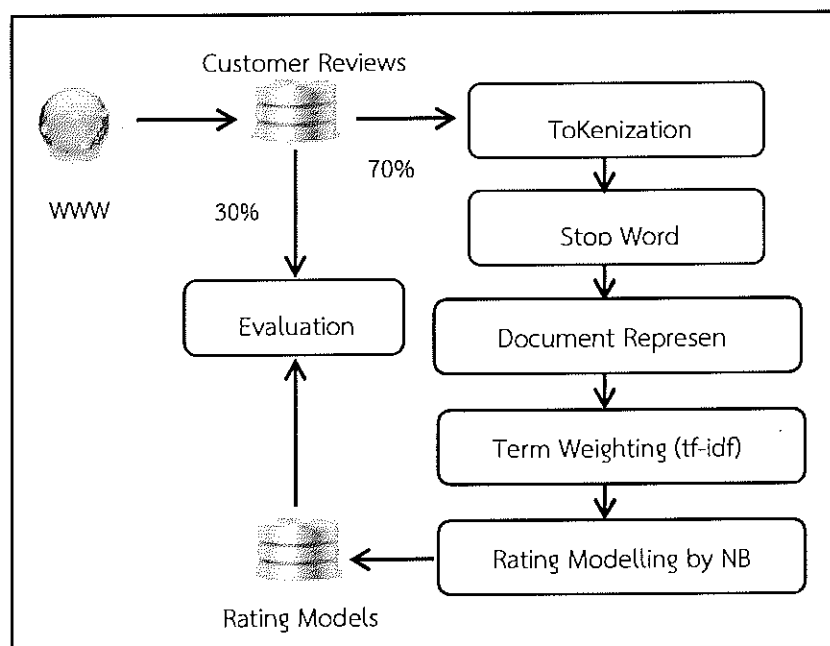
1. การรวบรวมข้อมูล (Data Collection)

รวบรวมข้อมูลบทวิจารณ์มาจากโรงแรมต่างๆ โดยจะเก็บทั้งเอกสารภาษาไทยและภาษาอังกฤษ ในรูปแบบของ Text File งานวิจัยนี้จะทำการเก็บข้อมูลจำนวน 500 เอกสารต่อกลุ่ม คือ

กลุ่ม “0” ความรู้สึกเชิงลบ (Negative Feeling)

กลุ่ม “1” ความรู้สึกเป็นกลาง (Neutral Feeling)

กลุ่ม “2” ความรู้สึกเชิงบวก (Positive Feeling)



รูปที่ 10 กระบวนการจัดระดับความคิดเห็นเกี่ยวกับบริการโรงแรมแบบอัตโนมัติ

2. การเตรียมข้อมูล (Data Preparation)

ในขั้นตอนนี้จะประกอบไปด้วยการทำงานหลักๆ ทั้งหมด 4 ขั้นตอนดังนี้

1. ขั้นตอนของการตัดคำ (Tokenization)
 2. ขั้นตอนของการตัดคำหยุด (Stop-word Removal)
 3. ขั้นตอนการคัดเลือกคำที่เป็นคุณลักษณะ (Feature Selection)
 4. ตัวแทนเนื้อหาของเอกสารและการให้น้ำหนักคำ
3. สร้างโมเดลในการจำแนกระดับความรู้สึกใช้อัลกอริธึมนาอิวเบย์ (Naive Bayes) เนื่องจากนาอิวเบย์เป็นอัลกอริธึมที่ง่าย ไม่ซับซ้อนแต่ให้ประสิทธิภาพในการจำแนกเอกสารได้ดี

สรุป

การวิเคราะห์ความรู้สึกเป็นงานวิจัยที่อยู่ในกลุ่มของการประมวลผลภาษาธรรมชาติ มีกระบวนการมุ่งเน้นการวิเคราะห์และตรวจสอบความรู้สึกของผู้คน จากงานวิจัยฉบับนี้เสนอการจัดระดับความคิดเห็นต่อบริการของโรงแรมแบบอัตโนมัติจากบทวิจารณ์ของลูกค้าที่เป็นข้อความ โดยเบื้องต้นจะจำแนกรูปร่างความรู้สึกออกเป็น 3 ระดับคือ “0” ความรู้สึกเชิงลบ “1” ความรู้สึกเป็นกลาง และ “2” ความรู้สึกเชิงบวก งานวิจัยที่น่าเสนอนี้จะสามารถสร้างโมเดลใช้ในการวิเคราะห์ความรู้สึกได้ทั้งภาษาไทยและภาษาอังกฤษ ซึ่งจากผลการทดสอบโมเดลดังกล่าวได้ผลลัพธ์ที่น่าพอใจคือ 74%

2.2.7 Automatic Feeling Analysis from Opinion Text

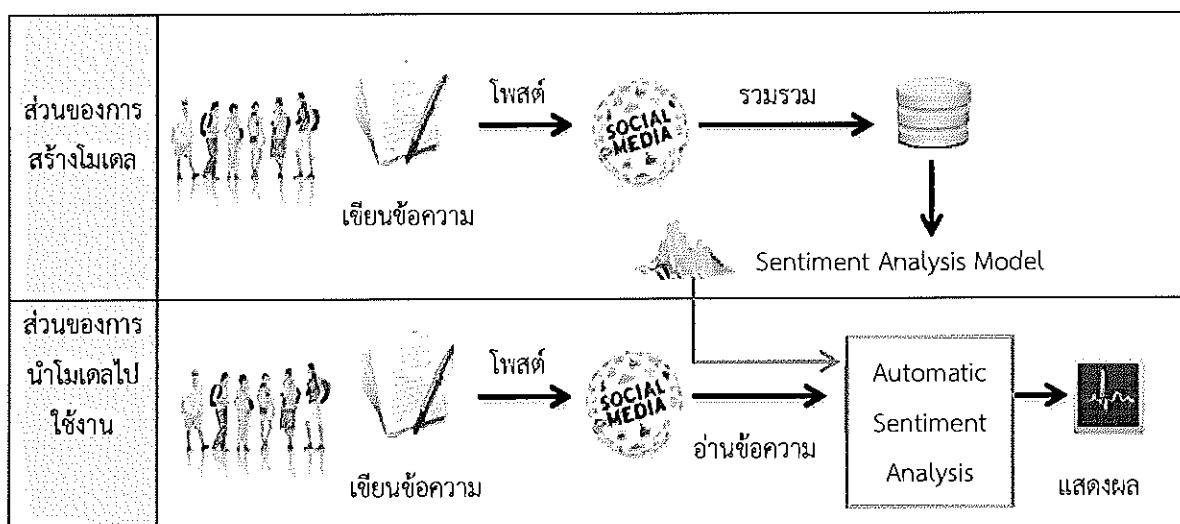
จากงานวิจัยของ Chetarin Wongsin et al. (2015) นำเสนอกระบวนการของการวิเคราะห์ความรู้สึก เพื่อวิเคราะห์ ความรู้สึกของนักท่องเที่ยวภายหลังจากที่ได้เยี่ยมชมสถานที่ท่องเที่ยวในภาคตะวันออกเฉียงเหนือวัตถุประสงค์ของงานวิจัยฉบับนี้ เพื่อวิเคราะห์ความรู้สึกของนักท่องเที่ยวที่มีการเขียนในรูปแบบของข้อความบนโซเชียลมีเดียเป็น 2 กลุ่มความรู้สึก คือความรู้สึกที่ดี และความรู้สึกที่ไม่ดีต่อสถานที่ท่องเที่ยวที่ได้ไปเยี่ยมชมโดยงานวิจัยฉบับนี้ได้พัฒนาโมเดลเพื่อการวิเคราะห์ความรู้สึกด้วยตัวจัดกลุ่มเอกสารแบบนาอูฟเบย์ หลังจากทดสอบด้วยการวัดค่าเอฟ

การวิเคราะห์ความรู้สึก (Sentiment Analysis)

การวิเคราะห์ความรู้สึก (Sentiment Analysis) เป็นงานวิจัยที่อยู่ในกลุ่มของการประมวลผลภาษาธรรมชาติ(Natural Language Processing: NLP) ที่มีกระบวนการมุ่งเน้นการวิเคราะห์และตรวจสอบความรู้สึก (Opinion) ของผู้คนที่จากข้อความ (Text) ที่คนเหล่านั้นเขียนหรือโพสต์เอาไว้ เพื่อบ่งบอกความรู้สึกของตนเองที่มีต่อบางสิ่งบางอย่าง เช่น ความรู้สึกดี (Positive หรือ Good) หรือความรู้สึกที่ไม่ดีหรือไม่ชอบ (Negative หรือ Bad)

กระบวนการวิจัย (Research Methodology)

ในส่วนนี้จะแสดงภาพรวมของการดำเนินการวิจัยดังรูปที่ 11



รูปที่ 11 กระบวนการวิเคราะห์ความรู้สึกของแบบอัตโนมัติ

1. ส่วนของการสร้างโมเดล

การสร้างโมเดลเพื่อการวิเคราะห์ความรู้สึก (Sentiment Classifier Modelling)

ขั้นตอนในการสร้างโมเดลเพื่อวิเคราะห์

ขั้นตอนที่ 1: การเตรียมเอกสาร (Document Preprocessing) ในการเตรียมเอกสารจะเริ่มจากการตัดคำ (Tokenization)

ขั้นตอนที่ 2: การนำเสนอเอกสาร (Document Representation) ในการนำเสนอเอกสาร (Document Representation) คือการแสดงความสัมพันธ์ระหว่าง “คำ ” และ “เอกสาร” ที่มีอยู่ในรูปแบบของเวกเตอร์

ขั้นตอนที่ 3: การสร้างโมเดลเพื่อการวิเคราะห์ความรู้สึกด้วย Naïve Bayes และนาอิวเบย์ (Naïve Bayes) เป็นอัลกอริทึมที่ถูกลำเอียงมาใช้อย่างแพร่หลายในงานจำแนก เอกสาร และให้ผลที่ดีในการสร้างตัวจำแนกแบบอัตโนมัติด้วยนาอิวเบย์

2. ส่วนของการนำโมเดลไปใช้ (Sentiment Model Usage)

เป็นขั้นตอนของการนำเอาโมเดลเพื่อการวิเคราะห์ความรู้สึกจากข้อความมาใช้งาน โดยเมื่อมีการอ่าน “ข้อความ” เข้ามา โมเดลดังกล่าวจะวิเคราะห์ให้ได้ว่า ความรู้สึกของนักท่องเที่ยวต่อสถานที่ท่องเที่ยวต่างๆ เป็น Positive หรือ Negative

สรุปงานวิจัย

งานวิจัยนี้เป็นการวิเคราะห์ ข้อความที่แสดงความรู้สึกแบบอัตโนมัติของนักท่องเที่ยวที่มีต่อสถานที่ท่องเที่ยวในประเทศไทย ที่มีการแสดงออกไว้ในข้อความแสดงความรู้สึกผ่าน Social Media ต่างๆ โดยในงานวิจัยนี้ จะวิเคราะห์ความรู้สึกแบบ 2 กลุ่ม คือ ความรู้สึกที่ดีต่อสถานที่ท่องเที่ยว (Positive) และความรู้สึกที่ไม่ดีต่อสถานที่ท่องเที่ยว(Negative) ซึ่งอัลกอริทึม Naïve Bayes ถูกใช้เป็นอัลกอริทึมหลักในการสร้างโมเดลสำหรับวิเคราะห์ความรู้สึก

งานวิจัยที่ผ่านมาที่ทำการสรุปความคิดเห็นจากบทวิจารณ์ภาษาไทยประกอบด้วย (Naris Prombut, 2007) และ (Choochart Haruechaiyasak et al., 2010: 64–71) ส่วนในงานวิจัย (Nongnuch Ketui et al., 2014: 215-220) เป็นการสรุปข่าวที่เป็นภาษาไทยซึ่งแนวทางการแก้ปัญหา คือ การสรุปความของบทความแต่ละบทความเพื่อหาประโยคที่จะใช้แทนบทความนั้น ใช้วิธี Thai Elementary Discourse Unit (TEDU) ซึ่งเป็นวิธีที่ใช้ในการกำหนดโครงสร้างของประโยคภาษาไทยตามรูปแบบ(Type) ที่ได้กำหนดไว้ทั้งหมด 12 รูปแบบ และ Common phrases (COMP) ส่วนในงานวิจัยภาษาไทยที่ใช้การสร้างรายการคำแสดงคุณลักษณะของสินค้า (Feature Extraction) โดยระบุทิศทางของความคิดเห็นด้วยแนวคิดแบบเวิร์ดเน็ต (Wordnet) และงานวิจัยที่ (Aekkasit Chongsuntornsri and Ohm Sornil, 2006) ทำการสรุปใจความสำคัญ ที่มีการทำกับเอกสารวิชาการ หรืองานวารสารที่ใช้คำถูกต้อง โดยให้ความสำคัญกับคำ ตำแหน่งของประโยค โดยใช้ อัลกอริทึม Hopfield และอัลกอริทึม PageRank มาใช้ในการหาประโยคสรุปที่สำคัญที่สุด ซึ่งมีรายละเอียดงานของวิจัยต่างๆดังนี้

2.2.8 Thai News Text Summarization and Its Application

จากงานวิจัยของ Nongnuch Ketui et al. (2014: 215-220) ได้พัฒนาวิธีการสรุปข่าวที่เป็นภาษาไทย เนื่องจากในปัจจุบันบทความข่าวมีจำนวนมากมาย สถานการณ์เช่นนี้ ทำให้เกิดสารสนเทศหรือข้อมูลจำนวนมากเกินไป ทำให้เกิดความยุ่งยากและลำบากในการสรุปบทความของข่าว และข่าวต่างๆ ก็มีมาจากหลายแหล่งข่าว แต่ละบทความก็จะมีคุณค่าคล้ายคลึงกัน จึงได้ศึกษาวิธีการในการสรุปข่าวจากบทความข่าวที่มีมากมายให้มีประสิทธิภาพและประสิทธิผล

วิธีการดำเนินงาน

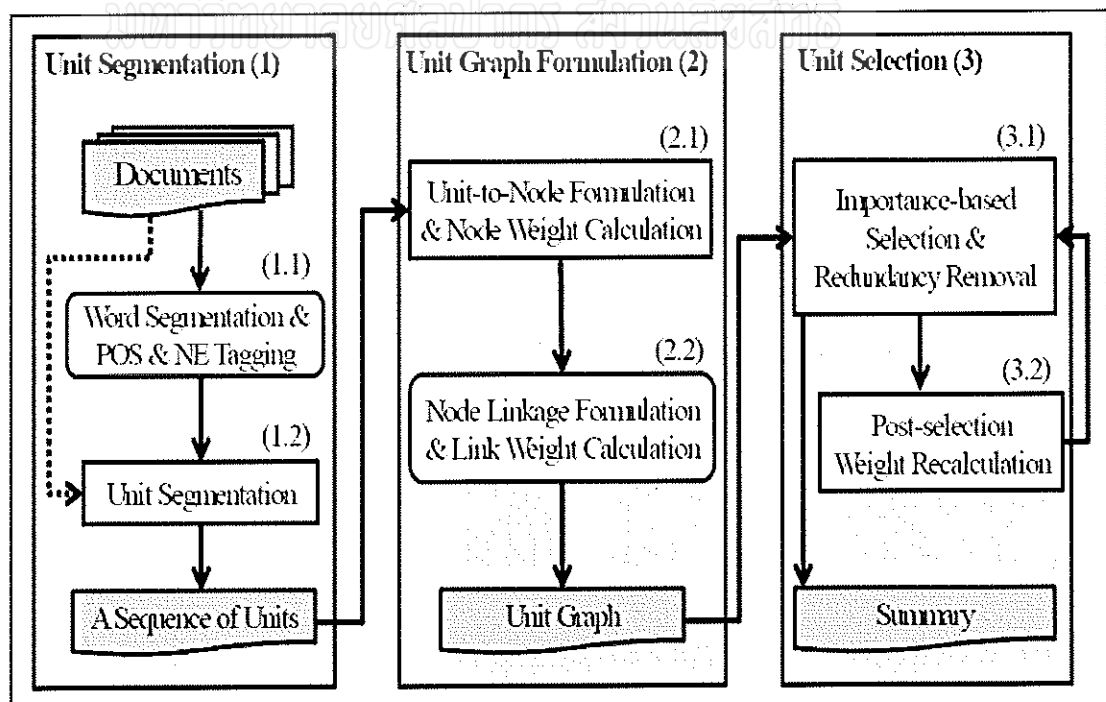
Simple clauses	Temporal phrases
Subject zero-anaphora clauses	Spatial phrases
Clauses with attribution verb	Conjunction phrases
Comparative clauses	Embedded phrases
Question clauses	Clausal subjects/objects
Embedded conjunction clauses	Synthetic nominal compounds

รูปที่ 12 การกำหนดโครงสร้างของประโยคภาษาไทย

แนวทางการแก้ปัญหา คือ การสรุปความของบทความแต่ละบทความ เพื่อหาประโยคที่จะใช้แทนบทความนั้น โดยงานวิจัยนี้เสนอวิธี Thai Elementary Discourse Unit (TEDU) ซึ่งเป็นวิธีที่ใช้ในการกำหนดโครงสร้างของประโยคภาษาไทยตามรูปแบบ (Type) ที่ได้กำหนดไว้ทั้งหมด 12 รูปแบบ และ Common phrases (COMP)

ขั้นตอนการดำเนินงานวิจัย

1. **Unit Segmentation** เริ่มจากนำบทความข่าวภาษาไทยที่เป็นเอกสาร นำมาตัดออกเป็นประโยค เมื่อได้เป็นประโยคแล้วก็นำมาตัดออกเป็นคำ โดยในแต่ละคำเราจะกำหนดหน้าที่ของคำว่าเป็นประเภทไหน (เช่น “กิน” เป็นคำ กริยา)
2. **Unit Graph Formulation** เป็นการกำหนดน้ำหนักให้กับคำแต่ละคำว่ามีน้ำหนักเท่าไร โดยการหาค่าน้ำหนักคำจะหาได้จากวิธี (TF-IDF) และในการหาความสัมพันธ์ระหว่างประโยคจะใช้วิธีของ (Cosine similarity)
3. **Unit Selection** ในขั้นตอนการเลือกประโยคที่สำคัญ ที่จะนำมาเป็นประโยคสรุป จะเลือกได้จากประโยคที่มีความสำคัญมากที่สุด โดยหาได้จาก ประโยคที่มีน้ำหนักสูงที่สุด



รูปที่ 13 ภาพรวมการทำงานของระบบสรุปข่าว

ผลสรุป

ในการสรุปผลใช้วิธีการวัดประสิทธิภาพด้วย ROUGE-1, ROUGE-2 และ ROUGE-SU4 ในช่วงระหว่าง (0.1 - 0.5) เพื่อหาวิธีที่ดีและมีประสิทธิภาพมากที่สุด คือ ROUGE-1 โดยที่วิธีการ Combined Thai elementary discourse units (CTEDU) ได้ประสิทธิภาพดีกว่า Thai elementary discourse units + Common phrases (COMP) (TEDU+COMP)

ข้อดีของงานวิจัย

งานวิจัยนี้อธิบายวิธีการและขั้นตอนเข้าใจได้ง่าย และสามารถนำงานวิจัยนี้มาพัฒนาต่อได้

ข้อเสียของงานวิจัย

ในงานวิจัยมีการอ้างอิงถึงวิธีการทำจากงานวิจัยอื่นเยอะเกินไป ทำให้เราต้องตามหาวิธีการจากงานวิจัยอื่นเป็นส่วนมาก จึงทำให้เราเข้าใจยาก และตารางในการสรุปผลอ่านเข้าใจยาก เนื่องจากมีการใช้ตัวย่อมากเกินไปและอธิบายแต่ละอย่างไม่ละเอียด

2.2.9 Ranking Explanatory Sentences for Opinion Summarization

จากงานวิจัยของ Hyun Duk Kim et al. (2013) ได้พัฒนาการสรุปความคิดเห็น ด้วยการวิเคราะห์ความรู้สึก (ความคิดเห็นบวก ลบ) ถึงแม้ว่าเทคนิคต่างๆ ที่มีอยู่สามารถเห็นการกระจายตัวของความคิดเห็นทั่วไป เช่น มีความคิดเห็นเป็นบวก 70% และลบ 30% แต่ก็ให้ผลลัพธ์ไม่เพียงพอในการให้ผลลัพธ์ที่ดีและมีประสิทธิภาพ สำหรับผู้ใช้งาน ดังนั้นในงานวิจัยนี้จึงได้นำเทคนิคต่างๆ มาช่วยเพื่อใช้ในการจัดเรียงประโยค เพื่อให้เราสามารถนำมาใช้งานได้ง่าย เป็นประโยชน์และมีประสิทธิภาพ สำหรับคนที่ต้องการนำข้อมูลเหล่านี้ไปใช้งาน กระบวนการทำงานของการสรุปความคิดเห็นแสดงดังรูปที่ 14

ขั้นตอนการทำงาน

1. เตรียมข้อมูล 2 ชุด คือ Product review และ Hotel review
2. นำมาทำการ classify ประโยคตามหัวข้อ (89 หัวข้อ 3799 ประโยค)

การทำ Input data set โดยทำทั้งหมด 4 ขั้นตอน

2.1 แบ่งออกตามหัวข้อ (T = topic (iPhone))

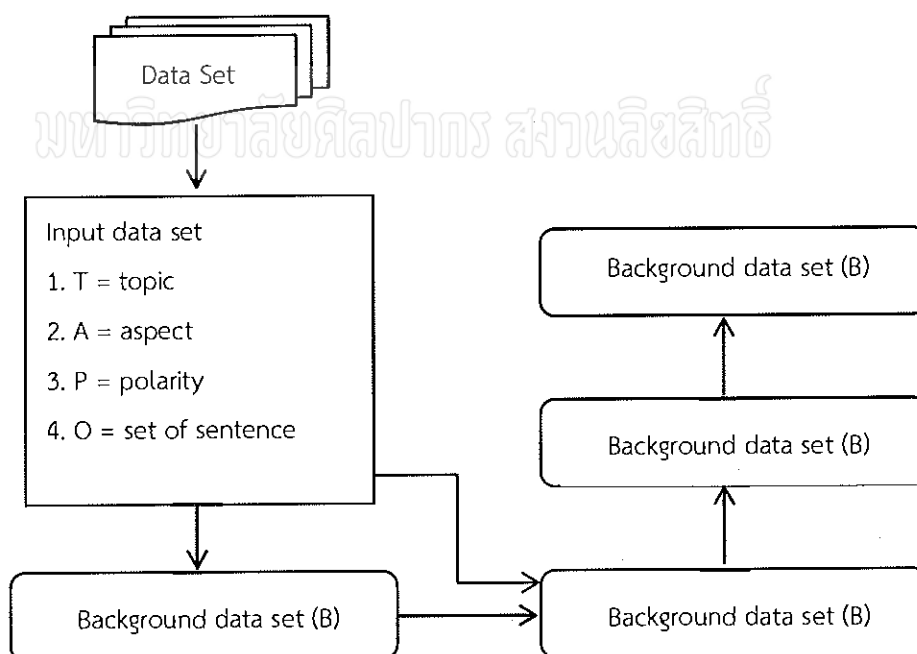
เมื่อได้หัวข้อแล้วทำการแบ่งตามเรื่องในหัวข้อนั้นๆ เป็นประโยค (A = aspect ('screen' of iPhone))

2.2 ขั้นตอนนี้จะทำการระบุความคิดเห็น (Tag) ของประโยคโดยใช้ผู้เชี่ยวชาญทั้งหมด 3 คน ('positive', 'negative')

2.3 นำประโยคทั้งหมดที่ทำการให้ความคิดเห็นแล้ว มารวมเป็น set ของข้อมูล ($O = \{S_1, \dots, S_n\}$)

3. เป็นการเตรียมหาค่าคะแนน (explanatoriness scoring) เพื่อใช้ในการหาความสำคัญของประโยคเพื่อเรียงลำดับประโยคที่มีความสำคัญที่สุด (Ranking sentence) โดยนำ three heuristics มาช่วยในการออกแบบ เพื่อสร้างฟังก์ชันในการให้คะแนนประโยค โดยวิเคราะห์พฤติกรรม 3 แบบ คือ ความยาวของประโยคสร้าง (Sentence length) ความนิยมและตัวแทนของประโยค (Popularity and representativeness) สิ่งที่ได้จากการเปรียบเทียบต้นฉบับ (Discriminateness relative to background)

4. ได้ประโยคสรุปที่ได้จากการเรียงลำดับความสำคัญ



รูปที่ 14 ภาพรวมการทำงานของ การสรุปความคิดเห็นด้วยการจัดลำดับ (Ranking)

ผลสรุป

Scoring Method	Product	Hotel
LexRank	0.4612	0.5869
BM25	0.7498*	0.6060
SumWordLR	0.7730*	0.6143

รูปที่ 15 การเปรียบเทียบคะแนน

จากรูปที่ 15 จะเห็นได้ว่าการวัดค่าความแม่นยำเฉลี่ย (wMAP) ของทั้งสามวิธี คือ โดยการเปรียบเทียบข้อมูล 2 ชุด คือ Product review และ Hotel review จะเห็นได้ว่าค่าของ Product review มีค่ามากกว่า เพราะว่า ข้อมูลที่นำมาใช้ในการทดลอง Hotel review มีทั้งหมด 23 หัวข้อ แต่ข้อมูลของ Product review มีทั้งหมด 66 หัวข้อ และอีกเหตุผลหนึ่งคือ ข้อมูลของ Hotel review อาจจะมี ความซับซ้อน หรือไม่ชัดเจนของข้อมูลด้วย ดังนั้นในการเปรียบเทียบทั้ง 3 วิธีนี้ จะเห็นว่า วิธีของ SumWordLR ได้ผลลัพธ์ที่ดีที่สุดดังรูปที่ 16

EXPLANATORY SUMMARY (SUMWORDLR)
when i played back a symphony orchestra you would swear that you were seated dead center in front of the orchestra - it 's that good ! 1) great sound (> 98db signal-to-noise ratio beats ipod 's " unspecified " ratio) and good power output allow the zen xtra to drive large head-phones as well as external speakers . once again , the sound is awesome , the battery life is only 6-8 hours , and that is because all my music is 320 kbps which does affect the battery life . the sound is awesome ,
BASELINE SUMMARY (LEXRANK)
the sound from the player is ok . - best in class sound the sound is excellent as one would suspect from a creative product . the sound is great even with the supplied earbuds - but i find earbuds uncomfortable so i use different headphones . plusses are the easy to remove battery and the terrific sound produced by the nomad . on the positive side , the sound of the player is pretty good , once you have everything configured . the sound is great , and the volume is more than satisfactory for

รูปที่ 16 ตัวอย่างผลลัพธ์ที่ได้จากการทำงานของ SumWordLR และ LexRank

บทที่ 3 ทฤษฎีที่เกี่ยวข้อง

3.1 ทฤษฎีของเบย์ส (Bayes Theorem)

Naive Bayes (Andrew McCallum and Kamal Nigam, 2003: 1265–1287) เป็นโมเดลการจำแนกกลุ่มที่ใช้หลักความน่าจะเป็น (Probability) ซึ่งอยู่บนพื้นฐานของ Bayes' s Theorem ซึ่งเป็นอัลกอริทึมที่ไม่ซับซ้อน เป็นขั้นตอนวิธีในการจำแนกข้อมูลโดยการเรียนรู้ปัญหาที่เกิดขึ้น เพื่อสร้างกลุ่มเงื่อนไขในการจำแนกข้อมูลใหม่ โดยใช้หลักความน่าจะเป็นในการจำแนกข้อมูล หลักการของ Naive Bayes (David McG.Squire, 2004) ใช้การคำนวณหาความน่าจะเป็นซึ่งถูกใช้ในการทำนายผล เป็นเทคนิคในการแก้ปัญหาแบบการจำแนกประเภทที่สามารถคาดการณ์ผลลัพธ์ได้ มันจะทำการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรเพื่อใช้ในการสร้างเงื่อนไขความน่าจะเป็นสำหรับแต่ละความสัมพันธ์ ตามทฤษฎีของ Naive Bayes ดังสมการที่ 1

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)} \quad (1)$$

โดยที่ $P(A|B)$ คือ ความน่าจะเป็นของ A เมื่อรู้ว่ามีเหตุการณ์ B
 $P(B|A)$ คือ ความน่าจะเป็นของ B เมื่อรู้ว่ามีเหตุการณ์ A
 $P(A)$ คือ ความน่าจะเป็นของเหตุการณ์ A
 $P(B)$ คือ ความน่าจะเป็นของเหตุการณ์ B

3.2 การหาความคล้ายคลึงของเอกสาร (Cosine similarity)

ในงานวิจัยนี้ใช้การวัดความคล้ายคลึงของประโยคด้วย (Cosine Similarity) ซึ่งเป็นวิธีการเปรียบเทียบความคล้ายคลึงของ 2 ประโยค โดยแต่ละประโยคจะถูกแทนด้วยเวกเตอร์ขนาดเอ็น (N-Dimensional Vector) ซึ่งเก็บค่าน้ำหนักคำแต่ละคำในประโยคนั้น (N-Dimensional Vector in Term Space) การเปรียบเทียบความคล้ายคลึงของประโยคจะเปรียบเทียบโดยดูจากมุมโคไซน์ของมุมระหว่าง 2 เวกเตอร์ของประโยค หากประโยคทั้งสองคล้ายคลึงกันมาก เวกเตอร์ของประโยคทั้ง 2

จะทับกันเกือบสนิทมุมจึงมีค่าน้อยและค่าโคไซน์จะมีค่ามาก ตัวอย่างในการวัดค่าความคล้ายคลึง
ระหว่างเวกเตอร์ s_1 และเวกเตอร์ s_2 จะมีค่าดังสมการที่ 2

$$\text{sim}(s_1, s_2) = \frac{\sum_{i=1}^n s_{1i} \times s_{2i}}{\sqrt{\sum_{i=1}^n s_{1i}^2} \times \sqrt{\sum_{i=1}^n s_{2i}^2}} \quad (2)$$

เมื่อ $\text{sim}(s_1, s_2)$ คือ ความคล้ายคลึงของประโยค s_1 และ s_2
 s_1 และ s_2 คือ ค่าเวกเตอร์ของประโยค
 n คือ จำนวนค่าของประโยค

ตัวอย่าง

ประโยคที่ 1 ความคล้ายคลึงของเอกสาร

เมื่อตัดคำจะได้ว่า ความคล้ายคลึง|ของ|เอกสาร|

ประโยคที่ 2 การหาความคล้ายคลึงของเอกสาร

เมื่อตัดคำจะได้ว่า การ|หา|ความคล้ายคลึง|ของ|เอกสาร|

ลำดับชั้นของคำศัพท์ได้ดังนี้

ความคล้ายคลึง|ของ|เอกสาร|การ|หา|

หาค่าความถี่ของคำในประโยค

ตารางที่ 3 ความถี่ของคำในประโยค

ประโยค \ คำ	คำ	ความ คล้ายคลึง	ของ	เอกสาร	การ	หา
ประโยคที่ 1		1	1	1	0	0
ประโยคที่ 2		1	1	1	1	1

นำความถี่ของประโยคจากรายที่ 3.1 มาเขียนไว้ในรูปของเวกเตอร์ดังนี้

ประโยคที่ 1	ความคล้ายคลึงของเอกสาร
เขียนในรูปของเวกเตอร์	(1, 1, 1, 0, 0)
ประโยคที่ 2	การหาความคล้ายคลึงของเอกสาร
เขียนในรูปของเวกเตอร์	(1, 1, 1, 1, 1)

หาค่าความคล้ายคลึงของประโยคที่ 1 และประโยคที่ 2 จากสมการ 2 ได้ดังนี้

$$\begin{aligned}
 &= \frac{(1 \times 1) + (1 \times 1) + (1 \times 1) + (0 \times 1) + (0 \times 1)}{\sqrt{(1^2 + 1^2 + 1^2 + 0^2 + 0^2) \times (1^2 + 1^2 + 1^2 + 1^2 + 1^2)}} \\
 &= \frac{3}{\sqrt{15}} \\
 &= 0.77461
 \end{aligned}$$

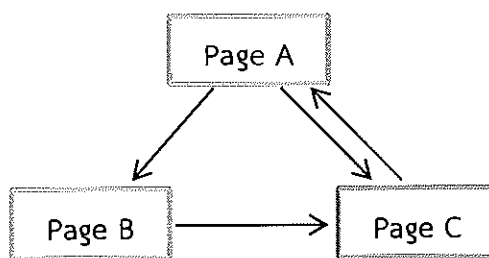
ดังนั้น ค่าความคล้ายคลึงของประโยคที่ 1 และประโยคที่ 2 เท่ากับ 0.775

3.3 อัลกอริทึม PageRank

PageRank (Aekkasit Chongsuntornsri and Ohm Sornil, 2006: 547–552) เป็นหลักการของ back link คือถ้าหน้าไหนถูกอ้างอิงมาก หรือมี link มาถึงหน้านั้นมาก แสดงว่าหน้านั้นมีความสำคัญ ซึ่งเป็นการบ่งบอกถึงความสำคัญของ website และได้นำมาประยุกต์ใช้ในงานวิจัยเพื่อนำมาใช้ในการจัดอันดับถึงความสำคัญของประโยคในแต่ละกลุ่ม เพื่อหาประโยคที่มีความสำคัญมากที่สุด สูตรในการคำนวณค่า PageRank ดังสมการที่ 3

$$PR(A) = \left(\frac{1 - D}{N} \right) + D \left(\frac{PR(T_1)}{C(T_1)} + \dots + \frac{PR(T_n)}{C(T_n)} \right) \quad (3)$$

$PR(A)$	คือ ค่าอันดับของประโยค A
D	คือ ค่า Damping Factor มีค่าอยู่ระหว่าง 0 – 1
$PR(T_i)$	คือ ค่าอันดับของประโยค T_i ที่เชื่อมโยงไปยังประโยค A
$C(T_i)$	คือ จำนวนประโยคที่ Link ออกจากประโยค T_i
N	คือ จำนวนประโยคทั้งหมด



รูปที่ 17 ตัวอย่างการเชื่อมโยงแต่ละประโยค

ตัวอย่างการคำนวณค่า PageRank โดยสมมติค่าให้ $D = 0.5$

$$\text{จากสมการ (3)} \quad PR(A) = \frac{(1-D)}{N} + D \times \left(\frac{PR(T1)}{C(T1)} + \dots + \frac{PR(TN)}{C(TN)} \right)$$

แทนค่าในสมการ

$$PR(A) = \frac{(1-0.5)}{3} + 0.5 \times PR(C)$$

$$PR(B) = \frac{(1-0.5)}{3} + 0.5 \times \left(\frac{PR(A)}{2} \right)$$

$$PR(C) = \frac{(1-0.5)}{3} + 0.5 \times \left(\frac{PR(A)}{2} + PR(B) \right)$$

เมื่อดูจากการแทนค่าในสูตรจะเห็นว่า

ค่า $PR(Ti)$ ของเพจ A คือ เพจ C เพราะมีเพียงเพจ C เท่านั้นที่ลิงก์มา เช่นเดียวกับเพจ B ที่มีเพจ A เท่านั้นที่ลิงก์มา ส่วนเพจ C มีเพจ A และเพจ B จำนวนสองเพจที่ลิงก์มา ส่วนค่า $C(Ti)$ คือ จำนวนเพจที่ถูกลิงก์ ในตัวอย่างข้างต้น จะมีเพียงเพจ C ที่ถูกลิงก์โดยสองเพจ ส่วนเพจ A และ B ถูกลิงก์เพียงเพจเดียวเท่านั้น เมื่อคำนวณหาค่า PageRank มาแล้ว ค่าของเพจไหนมีค่ามากที่สุด จะทำให้เพจนั้นถูกจัดอันดับไว้เพจแรกและเรียงลำดับไล่ลงมา

3.4 อัลกอริทึม Hopfield networks

อัลกอริทึม Hopfield networks (Hsinchun Chen et al., 1998), (Hsinchun Chen and Tobundorbin Ng, 2003) เป็นโครงข่ายประสาทเทียม (Neural network) (ธนาวุฒิ ประกอบผล, 2552) เป็นแขนงหนึ่งทางด้านปัญญาประดิษฐ์ (Artificial Intelligence: AI) ที่สามารถนำมา

ประยุกต์ใช้ทางการจำแนกรูปแบบการหาความเหมาะสมและการจัดกลุ่ม (Aekkasit Chongsuntornsri and Ohm Sornil, 2006: 547–552) โครงข่ายประสาทเทียมประกอบด้วย 5 องค์ประกอบดังนี้

- 1) ข้อมูลป้อนเข้า (input) (r_i) ใช้ค่าเดียวกับค่าความคล้ายคลึงของอัลกอริทึม PageRank
- 2) ข้อมูลส่งออก (output) ผลลัพธ์จากกระบวนการเรียนรู้ของโครงข่ายประสาทเทียม
- 3) ค่าน้ำหนัก (weight) w_i ค่าน้ำหนักเริ่มต้นที่มาจากการสุ่ม โดยกำหนดค่าน้ำหนักเริ่มต้นให้ทุกประโยคมีค่าเท่ากัน
- 4) ฟังก์ชันผลรวม (summation function) เป็นผลรวมของข้อมูลที่ป้อนเข้า (r_i) กับค่าน้ำหนัก (w_i)

$$\text{Summation } f^n = \sum_{i=1}^n r_i w_{ij} \quad (4)$$

- 5) ฟังก์ชันการแปลง (Transfer function) แบ่งได้ 2 ประเภท คือ

5.1) ฟังก์ชันการแปลงเชิงเส้น (Liner Transfer Function)

5.2) ฟังก์ชันการแปลงไม่เชิงเส้น (Nonlinear Transfer Function)

โดยฟังก์ชันการแปลงที่ใช้ในงานวิจัยนี้คือ ฟังก์ชันการแปลงไม่เชิงเส้น (Nonlinear Transfer Function) ชนิดที่ใช้คือ ฟังก์ชันการแปลง ซิกมอยด์ฟังก์ชัน (Sigmoid Function)

$$f_s(\text{net}_j) = \frac{1}{1 + \exp\left[\frac{\theta_j - \text{net}_j}{\theta_0}\right]} \quad (5)$$

$$\text{net}_j = \sum_{i=1}^n r_i w_{ij} \quad (6)$$

θ_j คือ ค่าความโน้มเอียง (bias)

θ_0 คือ ค่าที่ใช้ในการปรับเปลี่ยนฟังก์ชันซิกมอยด์

3.5 การหาค่าความถี่ของเอกสาร (TF-IDF)

การหาค่าความถี่ของเอกสาร TF-IDF (term frequency-inverse document frequency) เป็นการหาค่าโดยนำคำในเอกสารมาเขียนให้อยู่ในรูปของเวกเตอร์ จากนั้นจะทำการหาค่าส่วนกลับความถี่ของคำในเอกสาร (idf) เพื่อบ่งบอกถึงความสำคัญของคำ (Vanessa Graham Murdock, 2006) สามารถคำนวณได้ดังสมการที่ 7 และ 8

$$idf = \log \frac{N}{df} \quad (7)$$

$$tf - idf = tf \times idf \quad (8)$$

โดย	tf	คือ จากค่าความถี่ของคำ
	idf	คือ ค่าส่วนกลับความถี่ของเอกสาร
	N	คือ จำนวนเอกสารในชุดเอกสารทั้งหมด
	df	คือ ค่าความถี่ของเอกสารของแต่ละคำ

3.6 การวัดประสิทธิภาพ (ROUGE)

ในการวัดประสิทธิภาพนั้นจะใช้การวัดประสิทธิภาพแบบ ROUGE (Chin-Yew Lin, 2004) โดย ROUGE คือ ตัววัดปัญหาการสรุปข้อความ โดยใช้ n-gram และคำที่ซ้ำกัน เป็นมาตรฐานหลักที่ใช้วัดผลระหว่างผลลัพธ์ที่ได้จากอัลกอริทึมกับผลลัพธ์ที่ได้จากผู้ทดสอบ เราใช้ ROUGE-1, ROUGE-2, ROUGE-3, ROUGE-4 และ ROUGE-SU4 ในการวัดประสิทธิภาพ ROUGE โดยวัดผลลัพธ์ที่ได้จากอัลกอริทึม PageRank อัลกอริทึม Hopfield และ TF-IDF กับผลลัพธ์ที่ได้จากผู้ทดสอบในการสรุปประโยค ซึ่งผู้ทดสอบที่ใช้ คือผู้เชี่ยวชาญทางด้านเครื่องสำอางทั้งหมด 5 คน สามารถคำนวณได้จากสมการที่ 9

$$ROUGE_n(x) = \frac{\sum_{j=1}^h \sum_{i \in N_n} \min(x_n(i), M_n(i, j))}{\sum_{j=1}^h \sum_{i \in N_n} M_n(i, j)} \quad (9)$$

โดย	N_n	คือ เซตของคำทั้งหมดในประโยคที่ได้จากอัลกอริทึมที่ต้องการนำมาเปรียบเทียบ
	h	คือ จำนวนผู้ทดสอบที่เป็นผู้สรุปผลทั้งหมด h คน
	i	คือ จำนวนคำใน N_n
	j	คือ ลำดับของผู้ทดสอบ
	x	คือ ประโยคที่ได้จากอัลกอริทึมที่ต้องการนำมาเปรียบเทียบ
	$x_n(i)$	คือ คำที่ i ในประโยคที่ได้จากอัลกอริทึม ที่ต้องการนำมาเปรียบเทียบ
	$M_n(i, j)$	คือ คำที่ i, j ในประโยคที่ได้จากผู้ทดสอบ ลำดับที่ j
	n	คือ คำในประโยคที่จะนำมาเปรียบเทียบเป็น n-gram

$$ROUGE_n(x) = \frac{\text{ผลรวมของคำที่ } i \text{ ในประโยคที่ได้จากอัลกอริทึมเมื่อเทียบกับคำที่ } i \text{ ในประโยคที่ได้จากผู้ทดสอบลำดับที่ } j}{\text{ผลรวมของคำที่ } i \text{ ในประโยคที่ได้จากผู้ทดสอบลำดับที่ } j}$$

ตัวอย่างการคิดค่า ROUGE จะคิดจากการเปรียบเทียบคำในประโยคคำตอบที่ได้จากผู้ทดสอบกับคำในประโยคคำตอบที่ได้จากอัลกอริทึม PageRank และ Hopfield ดังตัวอย่าง ให้ประโยคคำตอบที่ได้จากผู้ทดสอบ 5 คน

ประโยคคำตอบที่ได้จากอัลกอริทึม : เนื้อ|ครีม|เกลี่ย|ง่าย|เรียบ|เนียน|ดี|

ประโยคคำตอบที่ได้จากผู้ทดสอบคนที่ 1 : เกลี่ย|ง่าย|มาก|

ประโยคคำตอบที่ได้จากผู้ทดสอบคนที่ 2 : เกลี่ย|ง่าย|และ|เนียน|

ประโยคคำตอบที่ได้จากผู้ทดสอบคนที่ 3 : เกลี่ย|ง่าย|ดี|ชอบ|มาก|

ประโยคคำตอบที่ได้จากผู้ทดสอบคนที่ 4 : เรียบ|เนียน|ดี|มาก|เลย|

ประโยคคำตอบที่ได้จากผู้ทดสอบคนที่ 5 : ครีม|เกลี่ย|ง่าย|เรียบ|เนียน|ดี|ชอบ|มาก|

การคำนวณค่า ROUGE-1 จากตัวอย่างประโยคคำตอบที่ได้จากอัลกอริทึม : เนื้อ|ครีม|เกลี่ย|ง่าย|เรียบ|เนียน|ดี| มีทั้งหมด 7 คำ คือคำว่า “เนื้อ”, “ครีม”, “เกลี่ย”, “ง่าย”, “เรียบ”, “เนียน”, “ดี” นำคำในประโยคคำตอบที่ได้จากอัลกอริทึมทีละคำตามลำดับ มาเปรียบเทียบกับคำในประโยคคำตอบที่ได้จากผู้ทดสอบคนที่ n สามารถอธิบายได้ดังนี้

1. คำแรกในประโยคคำตอบที่ได้จากอัลกอริทึมคือ คำว่า “เนื้อ” นำไปเปรียบเทียบกับคำในประโยคที่ได้จากผู้ทดสอบตั้งแต่คนที่ 1 ถึงคนที่ n ผลปรากฏว่าไม่มีคำไหนในประโยคผู้ทดสอบตรงกับคำว่า “เนื้อ” เลย
2. เลื่อนคำในประโยคคำตอบที่ได้จากอัลกอริทึมเป็นคำถัดไปเรื่อยๆ จนเจอคำที่เหมือนกันคือ คำว่า “ครีม” นำไปเปรียบเทียบกับคำในประโยคที่ได้จากผู้ทดสอบตั้งแต่คนที่ 1 ถึงคนที่ h ผลปรากฏว่า พบคำว่า “ครีม” 1 ครั้ง
3. เลื่อนคำในประโยคคำตอบที่ได้จากอัลกอริทึมเป็นคำถัดไปเรื่อยๆ จนเจอคำที่เหมือนกันคือ คำว่า “เกลี่ย” นำไปเปรียบเทียบกับคำในประโยคที่ได้จากผู้ทดสอบตั้งแต่คนที่ 1 ถึงคนที่ n ผลปรากฏว่า จะได้จำนวนคำที่เกี่ยวข้องที่ค้นเจอ คือคำว่า “เกลี่ย” ทั้งหมด 4 ครั้ง ซึ่งปรากฏในประโยคที่ 1, 2, 3 และ 5 ตามลำดับ
4. เลื่อนคำและทำการเปรียบเทียบระหว่างคำในประโยคคำตอบที่ได้จากอัลกอริทึมกับคำในประโยคคำตอบที่ได้จากผู้ทดสอบ ทำซ้ำแบบนี้ไปเรื่อยๆ จนครบ

ผลปรากฏว่า มีคำในประโยคคำตอบที่ได้จากผู้ทดสอบทั้ง 5 คน ตรงกับคำในประโยคคำตอบที่ได้จากอัลกอริทึม

จำนวน 17 คำ คือ คำว่า “ครีม”, “เกลี่ย”, “ง่าย”, “เรียบ”, “เนียน”, “ดี” โดยแต่ละคำมีจำนวน 1, 4, 4, 2, 3 และ 3 คำตามลำดับ

ส่วนจำนวนคำทั้งหมดให้นับจากคำทุกคำที่อยู่ในประโยคคำตอบที่ได้จากผู้ทดสอบ ตั้งแต่คนที่ 1 ถึงคนที่ n จะได้จำนวนคำทั้งหมด 25 คำ

จะได้ว่า $ROUGE-1 = 17/25 = 0.68000$

การคำนวณค่า ROUGE-2 จะมีวิธีการคำนวณคล้ายๆ กันกับค่า ROUGE-1 เพียงแต่คำที่จะนำมาเปรียบเทียบจากหนึ่งคำเป็นสองคำที่ติดกันในประโยค (bigram) จากตัวอย่างประโยคคำตอบที่ได้จากอัลกอริทึม : เนื้อ|ครีม|เกลี่ย|ง่าย|เรียบ|เนียน|ดี| จะแบ่งเป็นคำเพื่อใช้ในการเปรียบเทียบดังนี้คือ “เนื้อ|ครีม”, “ครีม|เกลี่ย”, “เกลี่ย|ง่าย”, “ง่าย|เรียบ”, “เรียบ|เนียน”, “เนียน|ดี|” นำคำในประโยคคำตอบที่ได้จากอัลกอริทึมตามลำดับ มาเปรียบเทียบกับคำในประโยคคำตอบที่ได้จากผู้ทดสอบคนที่ n สามารถอธิบายได้ดังนี้

1. คำสองคำที่อยู่ติดกันในประโยคคำตอบที่ได้จากอัลกอริทึม คือ คำว่า “เนื้อ|ครีม” นำไปเปรียบเทียบกับคำสองคำที่อยู่ติดกันในประโยคที่ได้จากผู้ทดสอบตั้งแต่คนที่ 1 ถึงคนที่ n ผลปรากฏว่าไม่มีคำไหนในประโยคผู้ทดสอบตรงกับคำว่า “เนื้อ|ครีม” เลย
2. เลื่อนคำในประโยคคำตอบที่ได้จากอัลกอริทึมเป็นคำถัดไปสองคำที่อยู่ติดกัน คือ คำว่า “ครีม|เกลี่ย” นำไปเปรียบเทียบกับคำสองคำที่อยู่ติดกันในประโยคที่ได้จากผู้ทดสอบ ตั้งแต่คนที่ 1 ถึงคนที่ n ผลปรากฏว่า พบคำว่า “ครีม|เกลี่ย” จำนวน 1 ครั้ง
3. เลื่อนคำในประโยคคำตอบที่ได้จากอัลกอริทึมเป็นคำถัดไปสองคำที่อยู่ติดกัน คือ คำว่า “เกลี่ย|ง่าย” นำไปเปรียบเทียบกับคำสองคำที่อยู่ติดกันในประโยคที่ได้จากผู้ทดสอบ ตั้งแต่คนที่ 1 ถึงคนที่ n จะได้จำนวนคำที่เกี่ยวข้องที่ค้นเจอ คือคำว่า “เกลี่ย|ง่าย|” ทั้งหมด 4 ครั้ง ซึ่งปรากฏในประโยคที่ 1, 2, 3 และ 5 ตามลำดับ
4. เลื่อนคำและทำการเปรียบเทียบระหว่างคำสองคำที่อยู่ติดกันในประโยคคำตอบที่ได้จากอัลกอริทึมกับคำสองคำที่อยู่ติดกันในประโยคคำตอบที่ได้จากผู้ทดสอบ ทำซ้ำแบบนี้ไปเรื่อยๆ จนครบ

ผลปรากฏว่า มีคำสองคำที่อยู่ติดกันในประโยคคำตอบที่ได้จากผู้ทดสอบทั้ง 5 คน ตรงกับคำสองคำที่อยู่ติดกันในประโยคคำตอบที่ได้จากอัลกอริทึม จำนวน 4 คำ คือ คำว่า “ครีม|เกลี่ย|”, “เกลี่ย|ง่าย|”, “ง่าย|เรียบ|”, “เรียบ|เนียน|”, “เนียน|ดี|” โดยมีจำนวนทั้งหมด 1, 4, 1, 2 และ 2 คำตามลำดับ

ส่วนจำนวนคำทั้งหมดให้นับจากคำสองคำที่อยู่ติดกันในประโยคคำตอบที่ได้จากผู้ทดสอบ ตั้งแต่คนที่ 1 ถึงคนที่ n จะได้จำนวนคำทั้งหมด 20

$$\text{จะได้ว่า} \quad \text{ROUGE-2} = 10/20 = 0.50000$$

การคำนวณค่า ROUGE-3 มีวิธีคำนวณคล้าย ๆ กันกับค่า ROUGE-2 เพียงแต่คำที่จะนำมาเปรียบเทียบจากสองคำที่ติดกันเป็นสามคำที่ติดกันในประโยค (trigram) คำตอบที่ได้จากอัลกอริทึม จากตัวอย่างประโยคคำตอบที่ได้จากอัลกอริทึม : เนื้อ|ครีม|เกลี่ย|ง่าย|เรียบ|เนียน|ดี| จะแบ่งเป็นคำเพื่อใช้ในการเปรียบเทียบดังนี้คือ “เนื้อ|ครีม|เกลี่ย|”, “ครีม|เกลี่ย|ง่าย|”, “เกลี่ย|ง่าย|เรียบ|”, “ง่าย|เรียบ|เนียน|”, “เรียบ|เนียน|ดี|” นำคำในประโยคคำตอบที่ได้จากอัลกอริทึมตามลำดับ มาเปรียบเทียบกับคำในประโยคคำตอบที่ได้จากผู้ทดสอบคนที่ n สามารถอธิบายได้ดังนี้

1. คำสามคำที่อยู่ติดกันในประโยคคำตอบที่ได้จากอัลกอริทึม คือ คำว่า “เนื้อ|ครีม|เกลี่ย” นำไปเปรียบเทียบกับคำสามคำที่อยู่ติดกันในประโยคที่ได้จากผู้ทดสอบตั้งแต่คนที่ 1 ถึงคนที่ n ผลปรากฏว่าไม่มีคำไหนในประโยคผู้ทดสอบตรงกับคำว่า “เนื้อ|ครีม|เกลี่ย” เลย
2. เลื่อนคำในประโยคคำตอบที่ได้จากอัลกอริทึมเป็นคำถัดไปสามคำที่อยู่ติดกัน คือ คำว่า “ครีม|เกลี่ย|ง่าย|” นำไปเปรียบเทียบกับคำสามคำที่อยู่ติดกันในประโยคที่ได้จากผู้ทดสอบตั้งแต่คนที่ 1 ถึงคนที่ n ผลปรากฏว่า พบคำว่า “ครีม|เกลี่ย|ง่าย|” จำนวน 1 ครั้ง
3. เลื่อนคำและทำการเปรียบเทียบระหว่างคำสามคำที่อยู่ติดกันในประโยคคำตอบที่ได้จากอัลกอริทึมกับคำสามคำที่อยู่ติดกันในประโยคคำตอบที่ได้จากผู้ทดสอบ ทำซ้ำแบบนี้ไปเรื่อยๆ จนครบ

ผลปรากฏว่า มีคำสามคำที่อยู่ติดกันในประโยคคำตอบที่ได้จากผู้ทดสอบทั้ง 5 คน ตรงกับคำสามคำที่อยู่ติดกันในประโยคคำตอบที่ได้จากอัลกอริทึม จำนวน 4 คำ คือ คำว่า “ครีม|เกลี่ย|ง่าย|”,

“เกลี่ย|ง่าย|เรียบ|”, “ง่าย|เรียบ|เนียน|”, “เรียบ|เนียน|ดี|” โดยมีจำนวนทั้งหมด 1, 1, 1 และ 2 คำตามลำดับ

ส่วนจำนวนคำทั้งหมดให้นับจากคำสามคำที่อยู่ติดกันในประโยคคำตอบที่ได้จากผู้ทดสอบ ตั้งแต่คนที่ 1 ถึงคนที่ n จะได้จำนวนคำทั้งหมด 20

$$\text{จะได้ว่า} \quad \text{ROUGE-3} = 5/15 = 0.33333$$

การคำนวณค่า ROUGE-4 มีวิธีคำนวณคล้าย ๆ กันกับค่า ROUGE-3 เพียงแต่คำที่จะนำมาเปรียบเทียบจากสามคำที่ติดกันเป็นสี่คำที่ติดกันในประโยค (4-gram) คำตอบที่ได้จากอัลกอริทึม จากตัวอย่างประโยคคำตอบที่ได้จากอัลกอริทึม : เนื้อ|ครีม|เกลี่ย|ง่าย|เรียบ|เนียน|ดี| จะแบ่งเป็นคำเพื่อใช้ในการเปรียบเทียบดังนี้คือ “เนื้อ|ครีม|เกลี่ย|ง่าย|”, “ครีม|เกลี่ย|ง่าย|เรียบ|”, “เกลี่ย|ง่าย|เรียบ|เนียน|”, “ง่าย|เรียบ|เนียน|ดี|” นำคำในประโยคคำตอบที่ได้จากอัลกอริทึมตามลำดับ มาเปรียบเทียบกับคำในประโยคคำตอบที่ได้จากผู้ทดสอบคนที่ n สามารถอธิบายได้ดังนี้

1. คำสี่คำที่อยู่ติดกันในประโยคคำตอบที่ได้จากอัลกอริทึม คือ คำว่า “เนื้อ|ครีม|เกลี่ย|ง่าย|” นำไปเปรียบเทียบกับคำสี่คำที่อยู่ติดกันในประโยคที่ได้จากผู้ทดสอบตั้งแต่คนที่ 1 ถึงคนที่ n ผลปรากฏว่าไม่มีคำไหนในประโยคผู้ทดสอบตรงกับคำว่า “เนื้อ|ครีม|เกลี่ย|ง่าย|” เลย
2. เลื่อนคำในประโยคคำตอบที่ได้จากอัลกอริทึมเป็นคำถัดไปสี่คำที่อยู่ติดกัน คือ คำว่า “ครีม|เกลี่ย|ง่าย|เรียบ|” นำไปเปรียบเทียบกับคำสี่คำที่อยู่ติดกันในประโยคที่ได้จากผู้ทดสอบตั้งแต่คนที่ 1 ถึงคนที่ n ผลปรากฏว่า พบคำว่า “ครีม|เกลี่ย|ง่าย|เรียบ|” จำนวน 1 ครั้ง
3. เลื่อนคำและทำการเปรียบเทียบระหว่างคำสี่คำที่อยู่ติดกันในประโยคคำตอบที่ได้จากอัลกอริทึมกับคำสี่คำที่อยู่ติดกันในประโยคคำตอบที่ได้จากผู้ทดสอบ ทำซ้ำแบบนี้ไปเรื่อยๆ จนครบ

ผลปรากฏว่า มีคำสี่คำที่อยู่ติดกันในประโยคคำตอบที่ได้จากผู้ทดสอบทั้ง 5 คน ตรงกับคำสี่คำที่อยู่ติดกันในประโยคคำตอบที่ได้จากอัลกอริทึม จำนวน 3 คำ คือ คำว่า “ครีม|เกลี่ย|ง่าย|เรียบ|”, “เกลี่ย|ง่าย|เรียบ|เนียน|”, “ง่าย|เรียบ|เนียน|ดี|” โดยมีจำนวนทั้งหมดอย่างละ 1 คำตามลำดับ

ส่วนจำนวนคำทั้งหมดให้นับจากคำสี่คำที่อยู่ติดกันในประโยคคำตอบที่ได้จากผู้ทดสอบ ตั้งแต่คนที่ 1 ถึงคนที่ n จะได้จำนวนคำทั้งหมด 10

$$\text{จะได้ว่า} \quad \text{ROUGE-4} = 3/10 = 0.30000$$

การคำนวณค่า ROUGE-SU4 มีวิธีคำนวณคล้าย ๆ กันกับค่า ROUGE อื่นๆ ข้างต้น โดยจะทำการจับคู่ของคำเพื่อนำมาเปรียบเทียบ ซึ่งมีข้อแม้ว่าคำในประโยคคำตอบที่ได้จากอัลกอริทึมที่จะนำมาจับคู่กันนั้นต้องอยู่ห่างกันไม่เกิน 4 คำ

จากตัวอย่างเมื่อทำการจับคู่แล้วจะได้คำที่จับคู่กัน 20 คำ จากคำในประโยคคำตอบที่ได้จากอัลกอริทึม ซึ่งประกอบด้วย

“เนื้อ ครีม ”	“เนื้อ เกลี่ย ”	“เนื้อ ง่าย ”	“เนื้อ เรียบ ”	“เนื้อ เนียน ”
“ครีม เกลี่ย ”	“ครีม ง่าย ”	“ครีม เรียบ ”	“ครีม เนียน ”	“ครีม ดี ”
“เกลี่ย ง่าย ”	“เกลี่ย เรียบ ”	“เกลี่ย เนียน ”	“เกลี่ย ดี ”	“ง่าย เรียบ ”
“ง่าย เนียน ”	“ง่าย ดี ”	“เรียบ เนียน ”	“เรียบ ดี ”	“เนียน ดี ”

เมื่อได้คำที่ทำการจับคู่กันแล้วจากคำในประโยคคำตอบที่ได้จากอัลกอริทึม ก็นำมาทำการเปรียบเทียบกับคำในประโยคคำตอบที่ได้จากผู้ทดสอบทั้ง 5 คน ซึ่งเป็นวิธีการเดียวกันกับการเปรียบเทียบแบบค่า ROUGE ต่างๆ ที่ผ่านมา ผลปรากฏว่า

จำนวนคำที่จับคู่กันโดยห่างไม่เกิน 4 คำที่เกี่ยวข้องที่ค้นเจอถูกต้อง ซึ่งประกอบด้วย 15 คำ คือ คำว่า

“ครีม เกลี่ย ”	“ครีม ง่าย ”	“ครีม เรียบ ”	“ครีม เนียน ”	“ครีม ดี ”
“เกลี่ย ง่าย ”	“เกลี่ย เรียบ ”	“เกลี่ย เนียน ”	“เกลี่ย ดี ”	“ง่าย เรียบ ”
“ง่าย เนียน ”	“ง่าย ดี ”	“เรียบ เนียน ”	“เรียบ ดี ”	“เนียน ดี ”

ซึ่งแต่ละคำจะมีจำนวนคำเท่ากับ 1, 1, 1, 1, 1, 4, 1, 2, 2, 1, 2, 2, 2, 2 และ 2 คำตามลำดับ

ส่วนจำนวนคำทั้งหมดให้นับจากคำที่จับคู่กันโดยห่างไม่เกิน 4 คำในประโยคคำตอบที่ได้จากผู้ทดสอบ ตั้งแต่คนที่ 1 ถึงคนที่ n จะได้จำนวนคำทั้งหมด 54 คำ ได้แก่

$$\text{จะได้ว่า } \text{ROUGE-SU4} = 25/54 = 0.46296$$

3.7 ค่าความถูกต้อง, ค่าความครบถ้วน, ค่าเอฟเมซเซอร์

การทดสอบเพื่อประเมินประสิทธิภาพในงานวิจัย จะทำการวัดผลออกมาเป็นค่าความถูกต้อง (Precision) ค่าความครบถ้วน (Recall) และค่าเอฟเมซเซอร์ (F-measure) โดยวิธีวัดผลมีดังนี้

1. ค่า Recall

$$\text{Recall} = \frac{\text{จำนวนคำที่เกี่ยวข้องที่ค้นเจอ}}{\text{จำนวนคำที่เกี่ยวข้องทั้งหมด}}$$

2. ค่า Precision

$$\text{Precision} = \frac{\text{จำนวนคำที่เกี่ยวข้องที่ค้นเจอ}}{\text{จำนวนคำที่ค้นเจอทั้งหมด}}$$

3. ค่า F- measure

$$\text{F - measure} = \frac{2 \times (\text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})}$$

ตารางที่ 4 ตัวอย่างการคำนวณหาค่าถูกต้อง, ค่าความครบถ้วน, ค่าเอฟเมซเซอร์

ประโยคคำตอบ	Recall	Precision	F- measure
ผู้ทดสอบคนที่ 1	2/3 = 0.66667	2/7 = 0.28571	0.39494
ผู้ทดสอบคนที่ 2	3/4 = 0.75000	3/7 = 0.42857	0.54661
ผู้ทดสอบคนที่ 3	3/5 = 0.60000	3/7 = 0.42857	0.50097
ผู้ทดสอบคนที่ 4	3/5 = 0.60000	3/7 = 0.42857	0.50097
ผู้ทดสอบคนที่ 5	6/8 = 0.75000	6/7 = 0.85714	0.79687
ROUGE-1	0.68000	0.48571	0.54807

บทที่ 4

วิธีการดำเนินงานวิจัย

ในบทนี้จะกล่าวถึงแนวทางในการทำวิจัยเพื่อตอบวัตถุประสงค์ของงานวิจัยนี้ซึ่งประกอบไปด้วย สมมติฐานงานวิจัย แบบแผนการทดลอง เครื่องมือที่ใช้ในงานวิจัย แนวทางการทำวิจัย ข้อมูลที่ใช้ในการศึกษาและการเก็บรวบรวมข้อมูล การศึกษาเบื้องต้น ภาพรวมการทำงานของระบบสรุปบทวิจารณ์สินค้า ที่มีการวิเคราะห์ระดับความหนักแน่นของความคิดเห็น ซึ่งมีรายละเอียดดังรูปที่ 18

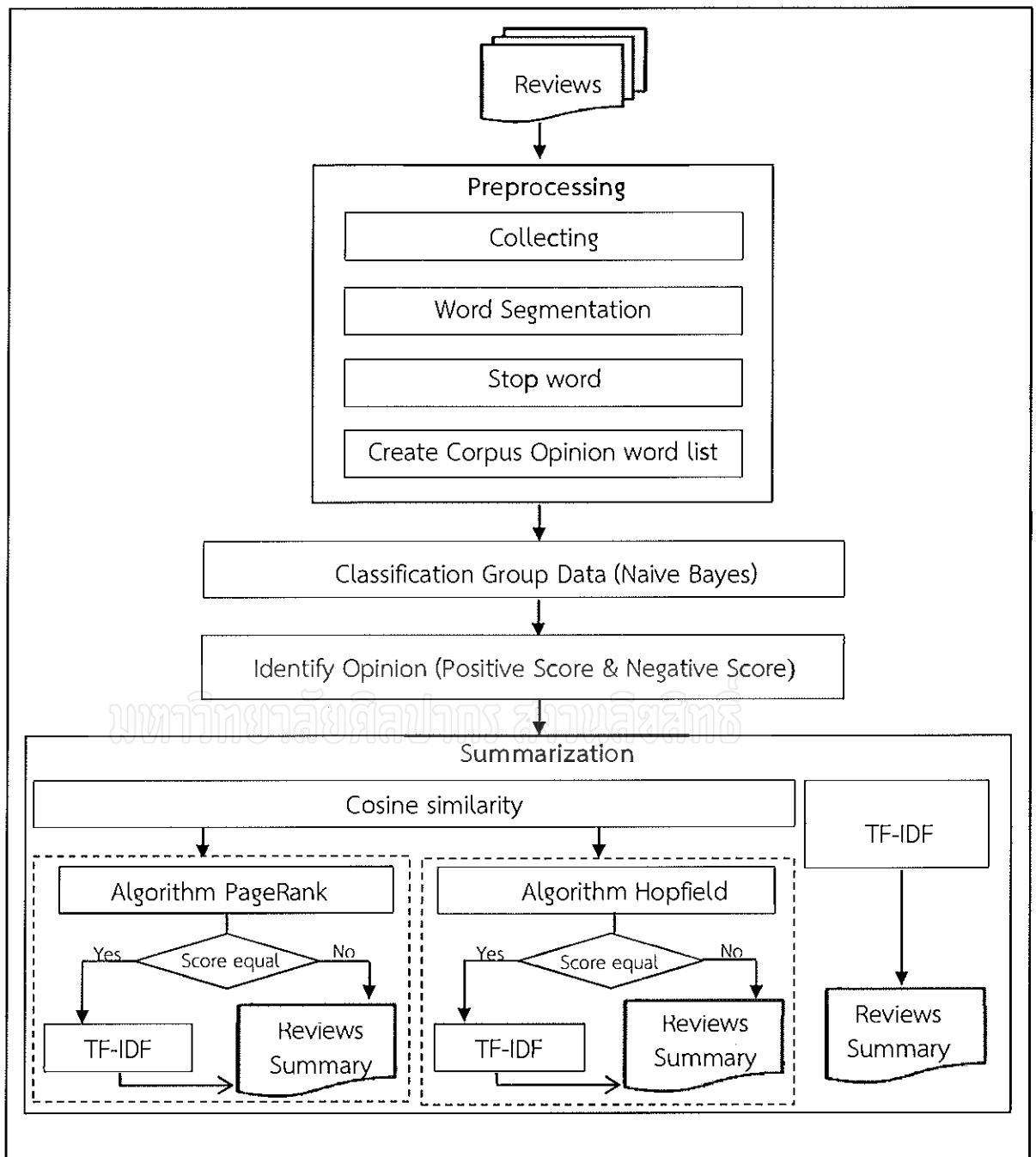
4.1 ข้อมูลที่ใช้ในงานวิจัย

โปรแกรมจะมีกระบวนการทำงานโดยจะดึงความคิดเห็นเกี่ยวกับผลิตภัณฑ์เครื่องสำอางที่ได้จากเว็บไซต์ www.Jeban.com มาจำนวน 3 ประเภท มาเก็บในรูปของ Text File ซึ่งในแต่ละประเภทจะประกอบด้วย 3 ผลิตภัณฑ์ ดังนี้

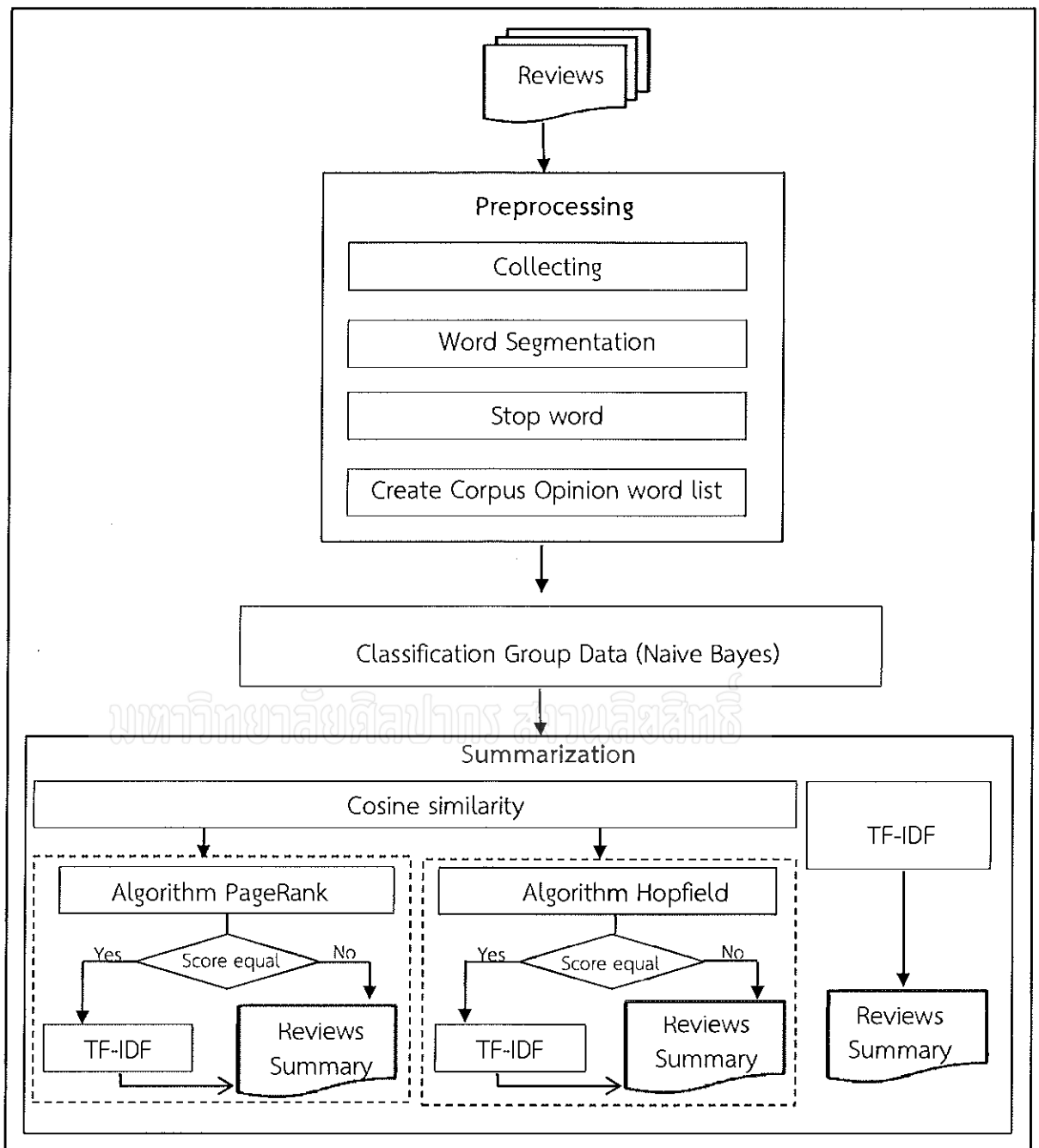
1. ประเภทครีมรองพื้น (Foundation)
CathyDoll-CCCreamSPF50PA+++
Revlon-ColorstayMakeupWithSoftflexforCombinationOilySkin
Revlon-PhotoReadyFoundation
2. ประเภทที่ปิดแก้ม (Blush on)
M.A.C-PowderBlush
NARS-Blush
Illamasqua - PowderBlusher
3. ประเภทผลิตภัณฑ์ทำความสะอาดผิวหน้า (Facial cleanser)
BIODERMA-SensibioH2O
Cetaphil-GentleSkinCleanser
DHC-DeepCleansingOil

4.2 วิธีการดำเนินงานวิจัย

โดยจะแบ่งเป็น 3 ขั้นตอนหลักคือ ขั้นตอนการเตรียมข้อมูล ขั้นตอนการสร้างกลุ่มข้อมูล และขั้นตอนการประมวลผลข้อมูล



รูปที่ 18 ภาพรวมการทำงานของระบบสรุปบทวิจารณ์สินค้าแบบที่ 1 และแบบที่ 2



รูปที่ 19 ภาพรวมการทำงานของระบบสรุปทวิจรรย์สินค้าแบบที่ 3

4.2.1 ขั้นตอนการเตรียมข้อมูล

4.2.1.1 เก็บข้อมูลความคิดเห็นจากเว็บไซต์

ในการทดลองได้ใช้ข้อมูลจากเว็บไซต์ที่มีชื่อเสียงด้านการแต่งหน้า และเครื่องสำอาง ดังรูปที่ 19 เก็บไว้ในรูปของ Text File โดยนำจากข้อมูลความคิดเห็นเกี่ยวกับผลิตภัณฑ์ของเครื่องสำอาง 3 ประเภท ได้แก่ ที่ปิดแก้ม (Blush On) ผลิตภัณฑ์ทำความสะอาดผิวหน้า (Cleanser) และครีมรองพื้น (Foundation) ซึ่งในแต่ละประเภทจะประกอบด้วย 3 ผลิตภัณฑ์

Figure 19: Screenshot of user reviews for beauty products. The image shows two review snippets. The first review is for 'MayBeauty' with a 4.0 rating, dated 28-04-16, and a 3-week-old review. The second review is for 'BeautyBee' with a 5.0 rating, dated 18-25-16, and a 3-week-old review. Both reviews include product names and user comments in Thai.

Figure 20: ตัวอย่างความคิดเห็นของผู้ใช้จากเว็บไซต์

ที่มา: User Reviews, accessed August 20, 2016, available from
<http://www.jeban.com/reviews.php>

แพคเกจโอเค
 ดิดทนทั้งวัน
 สีสวยมากเลย
 ราคาถูก
 ดิดทนดีมาก
 ราคาไม่แพงเลย
 แพคเกจไม่โอเคเลย
 ราคาเหมาะสม
 มีสีให้เลือกเยอะ
 สีไม่สวย
 ไม่ทน

รูปที่ 21 ตัวอย่างความคิดเห็นเก็บในรูปแบบของ Text File

4.2.1.2 ตัดคำ (Token)

เลือกใช้โปรแกรมตัดในการตัดคำภาษาไทย (Thai word breaker for PHP, 2016) เป็นโปรแกรมตัดคำแบบใช้ PHP โดยใช้ฐานข้อมูลภาษาไทย และเพิ่มคำศัพท์สำคัญที่เกี่ยวข้องกับเครื่องสำอาง เพื่อให้การตัดคำตรงตามข้อมูลที่ทำในงานวิจัยนี้

| แพคเกจ | โอเค |
 | ดิดทน | ทั้งวัน |
 | สี | สวย | มาก | เลย
 | ราคา | ถูก
 | ดิดทน | ดี | มาก
 | ราคา | ไม่ | แพง | เลย
 | แพคเกจ | ไม่ | โอเค | เลย
 | ราคา | เหมาะสม
 | มี | สี | ให้ | เลือก | เยอะ
 | สี | ไม่ | สวย
 | ไม่ | ทน

รูปที่ 22 ตัวอย่างที่ได้จากการตัดคำ

4.2.1.3 ตัดคำที่ไม่สำคัญ

คำที่ไม่สำคัญ (Stop Word) คือคำที่เป็นส่วนประกอบหรือขยายคำอื่นๆ แต่คำนั้นจะไม่มี ความหมายในตัว สามารถตัดทิ้งได้ และไม่ทำให้ความหมายของเอกสารนั้นเปลี่ยนไป ทำการตัดคำที่ ไม่สำคัญ คำที่ไม่มีความหมาย หรือคำที่ไม่เกี่ยวข้องออกไป

4.2.1.4 สร้างคลังข้อมูลระบุทิศทางของความคิดเห็น (Corpus Positive word, Corpus Negative word)

เนื่องจากงานวิจัยนี้ใช้ความคิดเห็นจากสินค้าชนิดเดียวกับงานวิจัยของ (Trithep Thumrongluck, 2010) จึงใช้คำแสดงความคิดเห็นจาก (Trithep Thumrongluck, 2010) เป็นฐานข้อมูล เริ่มต้น และเพิ่มคำแสดงความคิดเห็นที่ได้จากงานวิจัยใหม่เข้าไปด้วย ทำให้ได้ข้อมูลที่มีความละเอียด และถูกต้องมากขึ้น เพื่อนำข้อมูลนี้ไปสร้างกลุ่มคำเพื่อระบุทิศทางของความคิดเห็น (Polarity) ซึ่งมีการระบุทิศทางของความคิดเห็นแบ่งเป็น 2 ทิศทาง คือทางบวกและทางลบ

สดใส	น้อย
น่ารัก	หลุด
มาก	ไม่
เนียน	เพราะบาง
อ่อนโยน	เหม็น
สวย	แข็ง
ติดทน	หม่น
ทน	จืดเหี่ยว
ชัดเจน	บาง
ละเอียด	ต่ำ
แข็งแรง	ยาก
ชัดเจน	แย่

รูปที่ 23 ตัวอย่างความคิดเห็นที่เป็นบวก

รูปที่ 24 ตัวอย่างความคิดเห็นที่เป็นลบ

4.2.2 ขั้นตอนการจำแนกข้อมูลตามกลุ่มของความคิดเห็น

จำแนกกลุ่มข้อมูลโดยใช้อัลกอริทึมเบย์ (Naive Bayes) โดยสร้างกลุ่มของผลิตภัณฑ์ แต่ละประเภทด้วยการนำไฟล์ประโยคที่ได้จากกระบวนการข้างต้น มาทำการจัดกลุ่มประโยคผลลัพธ์ ที่ได้จากการจัดกลุ่มจะได้กลุ่มผลิตภัณฑ์รวมทั้งสิ้น 15 กลุ่ม ดังนี้

1. ประเภทที่ปิดแก้ม (Blush on) 3 กลุ่ม ได้แก่
 - กลุ่ม “ติดทน”
 - กลุ่ม “สี/เมื่อดสี”
 - กลุ่ม “แพคเกจ”
2. ประเภทผลิตภัณฑ์ทำความสะอาดผิวหน้า (Cleansing) 4 กลุ่ม ได้แก่
 - กลุ่ม “กลิ่น”
 - กลุ่ม “ความสะอาด”
 - กลุ่ม “แพ้/อ่อนโยน”
 - กลุ่ม “ราคา”
3. ประเภทครีมรองพื้น (Foundation) 5 กลุ่ม ได้แก่
 - กลุ่ม “เกลี่ยง่าย”
 - กลุ่ม “ปกปิด”
 - กลุ่ม “ควบคุมความมัน”
 - กลุ่ม “กันน้ำ”
 - กลุ่ม “ติดทน”

มหาวิทยาลัยศิลปากร สวทช.ลิขสิทธิ์

ตัวอย่างการใช้อัลกอริทึมเบย์ (Naive Bayes) ในการจำแนกกลุ่มของความคิดเห็น

ที่มา: (Tutorial: Building a Text Classification System, accessed April 5 2016, available from <http://textblob.readthedocs.io/en/dev/classifiers.html>)

1. ทำการเตรียมข้อมูลสำหรับไฟล์เพื่อให้โปรแกรมใช้เรียนรู้ (Training) และใช้ทดสอบ (Test) เพื่อหาค่าความถูกต้องในการจัดกลุ่มข้อมูล โดยแบ่งข้อมูลในการ (Training) 70% และแบ่งข้อมูลในการ (Test) 30% โดยที่รูปแบบของไฟล์ต้องเป็นรูปแบบ (Format) JSON files ดังรูปภาพที่ 25

JSON files ประกอบไปด้วย 2 ส่วน คือส่วนของ “text” กับ “label”

“text” คือ ประโยค

“label” คือ กลุ่มของของผลิตภัณฑ์

เช่น “A” กลุ่ม “ติดทน”

“B” กลุ่ม “สี”

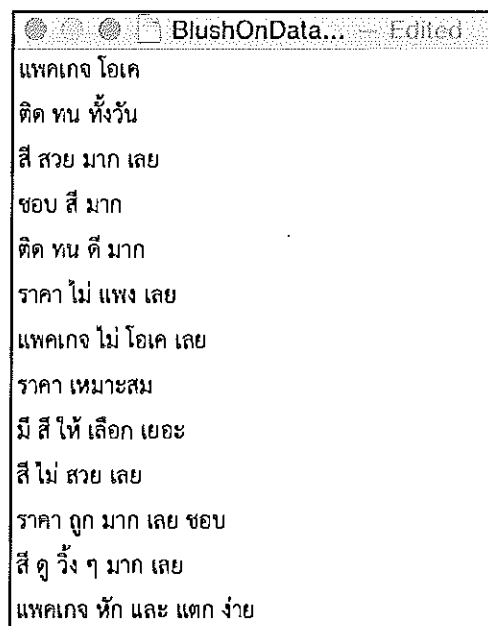
“C” กลุ่ม “แพคเกจ”

“X” คือ “ไม่มีกลุ่ม”



รูปที่ 25 ข้อมูล (Training) JSON files

- สร้างไฟล์ข้อมูล(Raw Data) เพื่อใช้ในการจำแนกกลุ่มของความคิดเห็นเป็นไฟล์ .txt โดยนำข้อมูลของประโยคความคิดเห็นเกี่ยวกับผลิตภัณฑ์ของเครื่องสำอาง 3 ประเภท ได้แก่ ที่ปิดแก้ม (Blush On) ผลิตภัณฑ์ทำความสะอาดผิวหน้า (Cleanser) และครีมรองพื้น (Foundation) ประเภทละ 3 ผลิตภัณฑ์ (จำนวนทั้งหมด 9,206 ประโยค) ไฟล์ข้อมูลที่ใช้ในการจำแนกกลุ่มของผลิตภัณฑ์แสดงดังตัวอย่างในรูปที่ 26



รูปที่ 26 ไฟล์ข้อมูลที่ใช้ในการจำแนกกลุ่มของความคิดเห็น

หลังจากจำแนกข้อมูลโดยใช้อัลกอริทึมนาอีฟเบย์ (Naive Bayes) เสร็จเรียบร้อยแล้ว จึงทำการจำแนกกลุ่มความรู้สึกของประโยค (Opinion word list) ของผลิตภัณฑ์แต่ละประเภทที่ผ่านการจัดกลุ่มตามขั้นตอนข้างต้น ประโยคของความคิดเห็นเหล่านี้จะถูกรวบรวม และเพื่อแสดงให้เห็นถึงความรู้สึกทั้งบวกหรือลบแล้วเก็บไว้ในคลังความคิดเห็น นอกจากนี้คำที่เป็นตัวแทนของความรู้สึกเชิงบวกหรือเชิงลบจะถูกดึงมาจาก Corpus Positive word , Corpus Negative word (Trithep Thumrongluck, 2010) ตัวอย่างในการคำนวณ

ในงานวิจัยนี้แบ่งการจัดกลุ่มออกเป็น 3 แบบ ดังนี้

แบบที่ 1 จัดกลุ่มความคิดเห็นของคำที่มีค่าน้ำหนักดังนี้

กลุ่ม “0” ความคิดเห็นเป็นกลาง (Neutral word)

กลุ่ม “+1” ความคิดเห็นเชิงบวก (Positive word)

กลุ่ม “-1” ความคิดเห็นเชิงลบ (Negative word)

ตัวอย่างเช่น

GROUP [เกลี้ยง่าย]

Example 1 : “เนื้อเบาบางเกลี้ยง่ายดี”

ความคิดเห็นของคำ “เนื้อ” → (Polarity of word “เนื้อ” คือ 0)

ความคิดเห็นของคำ “เบาบาง” → (Polarity of word “เบาบาง” คือ +1)

ความคิดเห็นของคำ “เกลี้ย” → (Polarity of word “เกลี้ย” คือ 0)

ความคิดเห็นของคำ “ง่าย” → (Polarity of word “ง่าย” คือ +1)

ความคิดเห็นของคำ “ดี” → (Polarity of word “ดี” คือ +1)

คำนวณคะแนนความคิดเห็นได้ $(0+1+0+1+1= +3)$ ดังนั้น “เนื้อเบาบางเกลี้ยง่ายดี” มีความคิดเห็นเป็นบวก

Example 2 : “เนื้อหยาบเกลี้ยยาก”

ความคิดเห็นของคำ “เนื้อ” → (Polarity of word “เนื้อ” คือ 0)

ความคิดเห็นของคำ “หยาบ” → (Polarity of word “หยาบ” คือ -1)

ความคิดเห็นของคำ “เกลี้ย” → (Polarity of word “เกลี้ย” คือ 0)

ความคิดเห็นของคำ “ยาก” → (Polarity of word “ยาก” คือ -1)

คำนวณคะแนนความคิดเห็นได้ $(0-1-0-1= -2)$ ดังนั้น “เนื้อหยาบเกลี้ยยาก” มีความคิดเห็นเป็นลบ

ตัวอย่างประโยค กลุ่ม “เกลี้ยง่าย” ความรู้สึกทางบวก

1. เนื้อเบาบางเกลี้ยง่ายดี
2. เกลี้ยง่ายมาก
3. เนื้อเกลี้ยง่ายดี

ตัวอย่างประโยค กลุ่ม “เกลี้ยง่าย” ความรู้สึกทางลบ

1. เนื้อหยาบเกลี้ยยาก
2. เกลี้ยลำบากมาก
3. เกลี้ยยากไม่ตีเลย

แบบที่ 2 จัดกลุ่มความคิดเห็นของคำที่มีค่าน้ำหนักดังนี้

กลุ่ม “0” ความคิดเห็นเป็นกลาง (Neutral word)

กลุ่ม “+1” ความคิดเห็นเชิงบวก (Positive word)

กลุ่ม “-1” ความคิดเห็นเชิงลบ (Negative word)

กลุ่ม “-2” (Very Negative word) ตัวอย่างเช่น คำว่า “ไม่” มีค่าเท่ากับ -2

มหาวิทยาลัยศิลปากร สวทช. สวทช. สวทช.

ตัวอย่างเช่น

GROUP [ติดทน]

Example 1 : “ติดทนดีมากเลย”

ความคิดเห็นของคำ “ติดทน” → (Polarity of word “เนื้อ” คือ +1)

ความคิดเห็นของคำ “ดี” → (Polarity of word “บางเบา” คือ +1)

ความคิดเห็นของคำ “มาก” → (Polarity of word “เกลี้ย” คือ 0)

ความคิดเห็นของคำ “เลย” → (Polarity of word “ง่าย” คือ 0)

คำนวณคะแนนความคิดเห็นได้ $(1+1+0+0= +2)$ ดังนั้น “ติดทนดีมากเลย” มีความคิดเห็นเป็นบวก

Example 2 : “ไม่ติดทนเลย”

ความคิดเห็นของคำ “ไม่” → (Polarity of word “เนื้อ” คือ -2)

ความคิดเห็นของคำ “ติดทน” → (Polarity of word “บางเบา” คือ +1)

ความคิดเห็นของคำ “เลย” → (Polarity of word “ง่าย” คือ 0)

คำนวณคะแนนความคิดเห็นได้ $((-2)+1+0= -1)$ ดังนั้น “ไม่ติดทนเลย” มีความคิดเห็นเป็นลบ

แบบที่ 3 จัดกลุ่มความคิดเห็นของประโยคด้วยการแยกกลุ่มประโยคในการฝึกสอน (Train) ออกเป็นกลุ่มของประโยคที่เป็นบวกและกลุ่มประโยคที่เป็นลบ เช่น สีน้าประเพทที่ปิดแก้ม (Blush on) แบ่งออกเป็น 3 กลุ่ม กลุ่มละ 2 แบบ คือ ได้แก่

- กลุ่ม “ติดทน” ทางบวก
- กลุ่ม “ติดทน” ทางลบ
- กลุ่ม “สี/เมืดสี” ทางบวก
- กลุ่ม “สี/เมืดสี” ทางลบ
- กลุ่ม “แพคเกจ” ทางบวก
- กลุ่ม “แพคเกจ” ทางลบ

4.2.3 ขั้นตอนการประมวลผลข้อมูล

4.2.3.1 การหาค่าความคล้ายคลึง (Similarity)

จากขั้นตอนที่สองนำข้อมูลแต่ละกลุ่มมาหาค่าความคล้ายคลึง (Similarity) อันดับแรกต้องนำประโยคที่ได้จากการจัดกลุ่มมาหาดัชนี (index) ของคำศัพท์ทั้งหมด จากนั้นหาความถี่ของคำในประโยคมาสร้าง Term Document Matrix ตัวอย่างประโยค กลุ่ม “ความสะอาด” ความรู้สึกทางบวก

1. ล้าง|ออก|ง่าย|สะอาด|ดี|
2. สะอาด|ดี|มาก|
3. ล้าง|ออก|ง่าย|
4. ล้าง|ง่าย|

ดัชนีคำศัพท์ของคำศัพท์ คือ

ล้าง|ออก|ง่าย|สะอาด|ดี|มาก|

ตารางที่ 5 ตัวอย่างความถี่ของคำในประโยค

	ล้าง	ออก	ง่าย	สะอาด	ดี	มาก
ประโยคที่ 1	1	1	1	1	1	0
ประโยคที่ 2	0	0	0	1	1	1
ประโยคที่ 3	1	1	1	0	0	0
ประโยคที่ 4	1	0	1	0	0	0

นำความถี่ของประโยคจากตารางที่ 5 มาเขียนไว้ในรูปของเวกเตอร์ดังนี้

ประโยคที่ 1

ล้างออกง่ายสะอาดดี

เขียนในรูปของเวกเตอร์	(1, 1, 1, 1, 1, 0)
ประโยคที่ 2	สะอาดดีมาก
เขียนในรูปของเวกเตอร์	(0, 0, 0, 1, 1, 1)
ประโยคที่ 3	ล้างออกง่าย
เขียนในรูปของเวกเตอร์	(1, 1, 1, 0, 0, 0)
ประโยคที่ 4	ล้างง่าย
เขียนในรูปของเวกเตอร์	(1, 0, 1, 0, 0, 0)

หาค่าความคล้ายคลึงของแต่ละประโยคโดยเปรียบเทียบทุกๆ ประโยคที่มีโอกาสเปรียบเทียบกันได้ โดยการแทนค่าค่าความคล้ายคลึงในสมการที่ 2 ยกเว้นประโยคตัวมันเองจะกำหนดค่าความคล้ายคลึงเมื่อเทียบกับตัวมันเองให้เท่ากับ 0 โดยถ้าประโยคไหนเปรียบเทียบกันแล้วมีค่าความคล้ายคลึงมาก แสดงว่าประโยคนั้นมีความคล้ายคลึงกันของประโยคมาก และหากเปรียบเทียบกันแล้วค่าความคล้ายคลึงน้อย ก็แสดงว่าความคล้ายคลึงกันของประโยคนั้นน้อยเช่นกัน

หาค่าความคล้ายคลึงกันระหว่างแต่ละประโยคโดยการแทนค่าในสมการที่ 2 ที่ทำการเปรียบเทียบกัน ได้ดังนี้

มหาวิทยาลัยศิลปากร สมานลียสิทธิ์

ค่าความคล้ายคลึงกันระหว่างประโยคที่ 1 และประโยคที่ 2

$$\begin{aligned} \text{Sim}(s_1, s_2) &= \frac{\sum_{i=1}^n s_1 \times s_2}{\sqrt{(\sum_{i=1}^n s_1^2 \times \sum_{i=1}^n s_2^2)}} \\ &= \frac{(1 \times 0) + (1 \times 0) + (1 \times 0) + (1 \times 1) + (1 \times 1) + (0 \times 1)}{\sqrt{(1^2 + 1^2 + 1^2 + 1^2 + 1^2 + 0^2) \times (0^2 + 0^2 + 0^2 + 1^2 + 1^2 + 1^2)}} \\ &= \frac{2}{\sqrt{15}} \\ &= 0.51640 \end{aligned}$$

ค่าความคล้ายคลึงกันระหว่างประโยคที่ 1 และประโยคที่ 3

$$\begin{aligned} \text{Sim}(s_1, s_3) &= \frac{\sum_{i=1}^n s_1 \times s_3}{\sqrt{(\sum_{i=1}^n s_1^2 \times \sum_{i=1}^n s_3^2)}} \\ &= \frac{(1 \times 1) + (1 \times 1) + (1 \times 1) + (1 \times 0) + (1 \times 0) + (0 \times 0)}{\sqrt{(1^2 + 1^2 + 1^2 + 1^2 + 1^2 + 0^2) \times (1^2 + 1^2 + 1^2 + 0^2 + 0^2 + 0^2)}} \end{aligned}$$

$$= \frac{3}{\sqrt{15}}$$

$$= 0.77461$$

ค่าความคล้ายคลึงกันระหว่างประโยคที่ 1 และประโยคที่ 4

$$\begin{aligned} \text{Sim}(s_1, s_4) &= \frac{\sum_{i=1}^n s_1 \times s_4}{\sqrt{(\sum_{i=1}^n s_1^2 \times \sum_{i=1}^n s_4^2)}} \\ &= \frac{(1 \times 1) + (1 \times 0) + (1 \times 1) + (1 \times 0) + (1 \times 0) + (0 \times 0)}{\sqrt{(1^2 + 1^2 + 1^2 + 1^2 + 1^2 + 0^2) \times (1^2 + 0^2 + 1^2 + 0^2 + 0^2 + 0^2)}} \\ &= \frac{2}{\sqrt{10}} \end{aligned}$$

$$= 0.63247$$

ค่าความคล้ายคลึงกันระหว่างประโยคที่ 2 และประโยคที่ 3

$$\begin{aligned} \text{Sim}(s_2, s_3) &= \frac{\sum_{i=1}^n s_2 \times s_3}{\sqrt{(\sum_{i=1}^n s_2^2 \times \sum_{i=1}^n s_3^2)}} \\ &= \frac{(0 \times 1) + (0 \times 1) + (0 \times 1) + (1 \times 0) + (1 \times 0) + (1 \times 0)}{\sqrt{(0^2 + 0^2 + 0^2 + 1^2 + 1^2 + 1^2) \times (1^2 + 1^2 + 1^2 + 0^2 + 0^2 + 0^2)}} \\ &= \frac{0}{\sqrt{9}} \end{aligned}$$

$$= 0$$

ค่าความคล้ายคลึงกันระหว่างประโยคที่ 2 และประโยคที่ 4

$$\begin{aligned} \text{Sim}(s_2, s_4) &= \frac{\sum_{i=1}^n s_2 \times s_4}{\sqrt{(\sum_{i=1}^n s_2^2 \times \sum_{i=1}^n s_4^2)}} \\ &= \frac{(0 \times 1) + (0 \times 0) + (0 \times 1) + (1 \times 0) + (1 \times 0) + (1 \times 0)}{\sqrt{(0^2 + 0^2 + 0^2 + 1^2 + 1^2 + 1^2) \times (1^2 + 0^2 + 1^2 + 0^2 + 0^2 + 0^2)}} \end{aligned}$$

$$= \frac{0}{\sqrt{6}}$$

$$= 0$$

ค่าความคล้ายคลึงกันระหว่างประโยคที่ 3 และประโยคที่ 4

$$\begin{aligned} \text{Sim}(s_3, s_4) &= \frac{\sum_{i=1}^n s_3 \times s_4}{\sqrt{(\sum_{i=1}^n s_3^2 \times \sum_{i=1}^n s_4^2)}} \\ &= \frac{(1 \times 1) + (1 \times 0) + (1 \times 1) + (0 \times 0) + (0 \times 0) + (0 \times 0)}{\sqrt{(1^2 + 1^2 + 1^2 + 0^2 + 0^2 + 0^2) \times (1^2 + 0^2 + 1^2 + 0^2 + 0^2 + 0^2)}} \\ &= \frac{2}{\sqrt{6}} \\ &= 0.81649 \end{aligned}$$

เมื่อได้ค่าความคล้ายคลึงจากการเปรียบเทียบทุกประโยคในกลุ่ม สามารถนำมาเขียนในรูปของตารางได้ดังตารางที่ 6

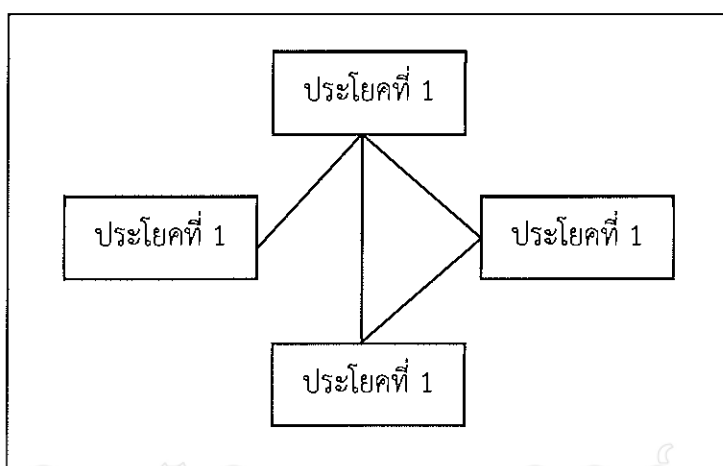
ตารางที่ 6 ค่าความคล้ายคลึงของแต่ละประโยค

	ประโยคที่ 1	ประโยคที่ 2	ประโยคที่ 3	ประโยคที่ 4
ประโยคที่ 1	0	0.51640	0.77461	0.63247
ประโยคที่ 2	0.51640	0	0	0
ประโยคที่ 3	0.77461	0	0	0.81649
ประโยคที่ 4	0.63247	0	0.81649	0

จากค่าความคล้ายคลึงที่ได้จากขั้นตอนนี้นำมาใช้เป็นข้อมูลนำเข้าคำนวณโดยใช้ อัลกอริทึม PageRank และ อัลกอริทึม Hopefield

4.2.3.2 การคำนวณโดยใช้อัลกอริทึม PageRank

จากตารางแสดงค่าความคล้ายคลึงกันของแต่ละประโยคในแต่ละกลุ่ม โดยประโยคที่มีค่าความคล้ายคลึงกันเท่ากับ 0 แสดงว่าประโยคที่นำมาเปรียบเทียบกับนั้นไม่มีความคล้ายคลึงกันของประโยคเลย หมายความว่าประโยคที่ถูกเปรียบเทียบกับนั้นไม่มีความสัมพันธ์กัน ซึ่งสามารถนำมาเขียนความสัมพันธ์ของแต่ละประโยคได้ดังรูปที่ 29



รูปที่ 29 ความสัมพันธ์ของแต่ละประโยค

ตัวอย่างการคำนวณโดยใช้อัลกอริทึม PageRank จากสมการที่ 2

$$PR(A) = \frac{(1-D)}{N} + D \times \left(\frac{PR(T1)}{C(T1)} + \dots + \frac{PR(TN)}{C(TN)} \right)$$

จากประโยคตัวอย่างเดิม คือ

1. ล้าง|ออก|ง่าย|สะอาด|ดี|
2. สะอาด|ดี|มาก|
3. ล้าง|ออก|ง่าย|
4. ล้าง|ง่าย|

มีทั้งหมด 4 ประโยค ดังนั้น ค่า $N = 4$ และค่า $D = 0.85000$ ค่า $C = 3, 1, 2, 2$ ตามลำดับ สามารถแทนค่าในสมการข้างต้นได้ดังนี้

$$PR(\text{ประโยชน์ที่ 1}) = \frac{(1-0.85000)}{4} + 0.85000 \times \left(\frac{PR(\text{ประโยชน์ที่ 2})}{1} + \frac{PR(\text{ประโยชน์ที่ 3})}{2} + \frac{PR(\text{ประโยชน์ที่ 4})}{2} \right)$$

$$PR(\text{ประโยชน์ที่ 2}) = \frac{(1-0.85000)}{4} + 0.85000 \times \left(\frac{PR(\text{ประโยชน์ที่ 1})}{3} \right)$$

$$PR(\text{ประโยชน์ที่ 3}) = \frac{(1-0.85000)}{4} + 0.85000 \times \left(\frac{PR(\text{ประโยชน์ที่ 1})}{3} + \frac{PR(\text{ประโยชน์ที่ 4})}{2} \right)$$

$$PR(\text{ประโยชน์ที่ 4}) = \frac{(1-0.85000)}{4} + 0.85000 \times \left(\frac{PR(\text{ประโยชน์ที่ 1})}{3} + \frac{PR(\text{ประโยชน์ที่ 3})}{2} \right)$$

การคำนวณหาค่า PageRank ของประโยชน์ทั้ง 4 ประโยค สามารถแสดงการคำนวณได้ดังตารางที่ 7 โดยการคำนวณหาค่า PageRank ของประโยชน์จะทำการคำนวณรอบต่อไปเรื่อยๆ จนกว่าค่ามีการเปลี่ยนแปลงน้อยที่สุด แล้วจึงทำการเลือกประโยชน์ที่มีค่า PageRank มากที่สุด จากตารางที่ 7 แสดงการคำนวณหาค่า PageRank ผลลัพธ์ของแต่ละประโยชน์ได้ค่าดังนี้

- | | | |
|-----------------------|----------------------|---------|
| 1. ล้างออกง่ายสะอาดดี | ค่า PageRank เท่ากับ | 0.40636 |
| 2. สะอาดดีมาก | ค่า PageRank เท่ากับ | 0.15267 |
| 3. ล้างออกง่าย | ค่า PageRank เท่ากับ | 0.26964 |
| 4. ล้างง่าย | ค่า PageRank เท่ากับ | 0.26723 |

เมื่อเรียงประโยชน์ตามค่าของ PageRank จะได้ประโยชน์ดังต่อไปนี้

ล้างออกง่ายสะอาดดี > ล้างออกง่าย, ล้างง่าย > สะอาดดีมาก

เพราะฉะนั้นประโยชน์ที่มีค่ามากที่สุดเป็นประโยชน์สรุปของอัลกอริทึม PageRank โดยการใช้การหาค่าความถี่ของเอกสารมาช่วยในการเลือกประโยชน์ คือ ประโยค “ล้างออกง่ายสะอาดดี”

ตารางที่ 7 ตัวอย่างการคำนวณหาค่า PageRank

รอบ	การคำนวณหา PageRank ของประวัติที่ 1	ค่า PageRank	การคำนวณค่า PageRank ของประวัติที่ 2	ค่า PageRank	การคำนวณหา PageRank ของประวัติที่ 3	ค่า PageRank	การคำนวณค่า PageRank ของประวัติที่ 4	ค่า PageRank
1	$(1-0.85000)/4+(0.85000*((0.51640/1)+(0.37461/2)+(0.62047/2)))$	1.07493	$(1-0.85000)/4+(0.85000*(1.07493/3))$	0.94212	$(1-0.85000)/4+(0.85000*((1.37493/3)+(0.81649/2)))$	0.68912	$(1-0.85000)/4+(0.85000*((0.68912/2)+(0.07493/3)+(0.68912/2)))$	0.63494
2	$(1-0.85000)/4+(0.85000*((0.34212/1)+(0.68912/2)+(0.63494/2)))$	0.89092	$(1-0.85000)/4+(0.85000*(0.89092/3))$	0.28894	$(1-0.85000)/4+(0.85000*((0.89092/3)+(0.63494/2)))$	0.55985	$(1-0.85000)/4+(0.85000*((0.55985/2)+(0.89092/3)+(0.55985/2)))$	0.52787
3	$(1-0.85000)/4+(0.85000*((0.28894/1)+(0.55985/2)+(0.52787/2)))$	0.74625	$(1-0.85000)/4+(0.85000*(0.74625/3))$	0.24893	$(1-0.85000)/4+(0.85000*((0.74625/3)+(0.52787/2)))$	0.47323	$(1-0.85000)/4+(0.85000*((0.47323/2)+(0.74625/3)+(0.47323/2)))$	0.45000
4	$(1-0.85000)/4+(0.85000*((0.24893/1)+(0.47323/2)+(0.45000/2)))$	0.64154	$(1-0.85000)/4+(0.85000*(0.64154/3))$	0.21934	$(1-0.85000)/4+(0.85000*((0.64154/3)+(0.45000/2)))$	0.41055	$(1-0.85000)/4+(0.85000*((0.41055/2)+(0.64154/3)+(0.41055/2)))$	0.39376
5	$(1-0.85000)/4+(0.85000*((0.21934/1)+(0.41055/2)+(0.39376/2)))$	0.56572	$(1-0.85000)/4+(0.85000*(0.56572/3))$	0.19782	$(1-0.85000)/4+(0.85000*((0.56572/3)+(0.39376/2)))$	0.36514	$(1-0.85000)/4+(0.85000*((0.36514/2)+(0.56572/3)+(0.36514/2)))$	0.33293
6	$(1-0.85000)/4+(0.85000*((0.19782/1)+(0.36514/2)+(0.35293/2)))$	0.51082	$(1-0.85000)/4+(0.85000*(0.51082/3))$	0.18223	$(1-0.85000)/4+(0.85000*((0.51082/3)+(0.35293/2)))$	0.33022	$(1-0.85000)/4+(0.85000*((0.33022/2)+(0.51082/3)+(0.33022/2)))$	0.30945
7	$(1-0.85000)/4+(0.85000*((0.18223/1)+(0.33022/2)+(0.30945/2)))$	0.47104	$(1-0.85000)/4+(0.85000*(0.47104/3))$	0.17102	$(1-0.85000)/4+(0.85000*((0.47104/3)+(0.30945/2)))$	0.30644	$(1-0.85000)/4+(0.85000*((0.30644/2)+(0.47104/3)+(0.30644/2)))$	0.30005
8	$(1-0.85000)/4+(0.85000*((0.17102/1)+(0.30944/2)+(0.30205/2)))$	0.44239	$(1-0.85000)/4+(0.85000*(0.44239/3))$	0.16287	$(1-0.85000)/4+(0.85000*((0.44239/3)+(0.30205/2)))$	0.29125	$(1-0.85000)/4+(0.85000*((0.29125/2)+(0.44239/3)+(0.29125/2)))$	0.28663
9	$(1-0.85000)/4+(0.85000*((0.16287/1)+(0.29125/2)+(0.27876/2)))$	0.42146	$(1-0.85000)/4+(0.85000*(0.42146/3))$	0.15694	$(1-0.85000)/4+(0.85000*((0.42146/3)+(0.27876/2)))$	0.27976	$(1-0.85000)/4+(0.85000*((0.27976/2)+(0.42146/3)+(0.27976/2)))$	0.27531
10	$(1-0.85000)/4+(0.85000*((0.15694/1)+(0.27876/2)+(0.27531/2)))$	0.40625	$(1-0.85000)/4+(0.85000*(0.40625/3))$	0.15267	$(1-0.85000)/4+(0.85000*((0.40625/3)+(0.27531/2)))$	0.26964	$(1-0.85000)/4+(0.85000*((0.26964/2)+(0.40625/3)+(0.26964/2)))$	0.26723

4.2.3.3 การคำนวณโดยใช้อัลกอริทึม Hopfield

การคำนวณโดยใช้อัลกอริทึม Hopfield กำหนดค่าน้ำหนักเริ่มต้นของทุกประโยคเท่ากับ 1 และกำหนดให้ค่าความโน้มเอียง θ_j (ค่าที่ได้จากการทดลอง) ค่าที่ใช้ปรับฟังก์ชันซิกมอยด์ $\theta_0 = 0.10000$ (ค่าที่ได้จากการทดลอง) ค่าฟังก์ชันเปลี่ยนแปลงของแต่ละประโยค คือ นำค่าผลลัพธ์ไปใช้ในการคำนวณรอบต่อไปเรื่อยๆ จนกว่าค่ามีการเปลี่ยนแปลงน้อยที่สุดแล้วจึงนำค่าสุดท้ายที่ได้มาหาค่าที่มากที่สุด แล้วจึงทำการเลือกประโยคที่มีค่า Hopfield มากที่สุดมาเป็นประโยคสรุปว่ามีใจความสำคัญที่สุดในแต่ละกลุ่ม ตัวอย่างการคำนวณโดยใช้อัลกอริทึม Hopfield

1. ล้าง|ออก|ง่าย|สะอาด|ดี|
2. สะอาด|ดี|มาก|
3. ล้าง|ออก|ง่าย|
4. ล้าง|ง่าย|

กำหนดค่าน้ำหนักเริ่มต้น ประโยคที่ 1 = 1, ประโยคที่ 2 = 1, ประโยคที่ 3 = 1, ประโยคที่ 4 = 1 และกำหนดให้ค่า $\theta_j = 0.85000$, $\theta_0 = 0.10000$

มหาวิทยาลัยศิลปากร สภานิติศาสตร์

ตัวอย่างการคำนวณของประโยคที่ 1

$$\text{ฟังก์ชันผลรวม (ประโยคที่ 1)} = ((1 * \text{ค่าความคล้ายคลึงประโยค 1}) + (1 * \text{ค่าความคล้ายคลึงประโยค 2}) + (1 * \text{ค่าความคล้ายคลึงประโยค 3}) + (1 * \text{ค่าความคล้ายคลึงประโยค 4}))$$

$$\text{ฟังก์ชันการเปลี่ยนแปลง (ประโยคที่ 1)} = 1 / (1 + \exp((- (\text{ค่าฟังก์ชันผลรวม(ประโยคที่ 1)} + 0.85000) / 0.10000)))$$

ค่าฟังก์ชันการเปลี่ยนแปลงของแต่ละประโยค คือ ค่าผลลัพธ์ที่นำไปใช้ในรอบต่อ ๆ ไป โดยจะทำการเปลี่ยนแปลงค่าน้ำหนักเป็นค่าผลลัพธ์ที่ได้ในแต่ละรอบ

การคำนวณค่า Hopfield ของประโยคทั้ง 4 ประโยค สามารถแสดงการคำนวณได้ดังตารางที่ 8

ตารางที่ 8 ตัวอย่างการคำนวณหาค่า Hopfield

รอบ	การคำนวณค่า Hopfield ของประยงค์ที่ 1	ค่า Hopfield	การคำนวณค่า Hopfield ของประยงค์ที่ 2	ค่า Hopfield	การคำนวณค่า Hopfield ของประยงค์ที่ 3	ค่า Hopfield	การคำนวณค่า Hopfield ของประยงค์ที่ 4	ค่า Hopfield
1	$((1^0) \cdot (1^0 \cdot 0.51640) + (1^0 \cdot 0.77461) + (1^0 \cdot 0.63247))$	1.92552	$((1^0 \cdot 0.51640) + (1^0) \cdot (1^0) + (1^0) \cdot (1^0))$	0.51640	$((1^0 \cdot 0.77461) + (1^0) \cdot (1^0) + (1^0) \cdot (1^0 \cdot 0.91649))$	1.59112	$((1^0 \cdot 0.63247) + (1^0) \cdot (1^0) + (1^0 \cdot 0.91649) + (1^0))$	1.44502
	$1/(1 + \exp(-1.92552 + 0.85000/0.10000))$	1.00000	$1/(1 + \exp(-0.51640 + 0.85000/0.10000))$	0.05445	$1/(1 + \exp(-1.59112 + 0.85000/0.10000))$	0.99945	$1/(1 + \exp(-1.44502 + 0.85000/0.10000))$	0.99755
2	$((1^0) \cdot (0.03445 \cdot 0.51640) + (0.99943 \cdot 0.77461) + (0.99753 \cdot 0.63247))$	1.42285	$((1^0 \cdot 0.51640) + (0.05445 \cdot 0) + (0.99943 \cdot 0) + (0.99753 \cdot 0))$	0.51640	$((1^0 \cdot 0.77461) + (0.03445 \cdot 0) + (0.99943 \cdot 0) + (0.99753 \cdot 0.91649))$	1.58905	$((1^0 \cdot 0.63247) + (0.03445 \cdot 0) + (0.99943 \cdot 0.91649) + (0.99753 \cdot 0))$	1.44856
	$1/(1 + \exp(-1.42285 + 0.85) / 0.100000))$	0.99685	$1/(1 + \exp(-0.51640 + 0.85000/0.10000))$	0.05442	$1/(1 + \exp(-1.58905 + 0.85000/0.10000))$	0.99945	$1/(1 + \exp(-1.44856 + 0.85000/0.10000))$	0.99754
3	$((0.99685 \cdot 0) + (0.05442 \cdot 0.51640) + (0.99943 \cdot 0.77461) + (0.99753 \cdot 0.63247))$	1.42285	$((0.99685 \cdot 0.51640) + (0.05442 \cdot 0) + (0.99943 \cdot 0) + (0.99753 \cdot 0))$	0.51475	$((0.99685 \cdot 0.77461) + (0.05442 \cdot 0) + (0.99943 \cdot 0) + (0.99685 \cdot 0.91649))$	1.58653	$((0.99685 \cdot 0.63247) + (0.05442 \cdot 0) + (0.99943 \cdot 0.91649) + (0.99685 \cdot 0))$	1.44643
	$1/(1 + \exp(-1.42285 + 0.85000/0.100000))$	0.99686	$1/(1 + \exp(-0.51475 + 0.85000/0.10000))$	0.05394	$1/(1 + \exp(-1.58653 + 0.85000/0.10000))$	0.99945	$1/(1 + \exp(-1.44643 + 0.85000/0.10000))$	0.99748
4	$((0.99686 \cdot 0) + (0.05394 \cdot 0.51640) + (0.99945 \cdot 0.74615) + (0.99748 \cdot 0.63247))$	1.42244	$((0.99748 \cdot 0.51640) + (0.05394 \cdot 0) + (0.99945 \cdot 0) + (0.99748 \cdot 0))$	0.51476	$((0.99748 \cdot 0.77461) + (0.05394 \cdot 0) + (0.99945 \cdot 0) + (0.99748 \cdot 0.91649))$	1.5865	$((0.99748 \cdot 0.63247) + (0.05394 \cdot 0) + (0.99945 \cdot 0.91649) + (0.99748 \cdot 0))$	1.44645
	$1/(1 + \exp(-1.42244 + 0.85000/0.10000))$	0.99675	$1/(1 + \exp(-0.51476 + 0.85000/0.10000))$	0.05392	$1/(1 + \exp(-1.58653 + 0.85000/0.10000))$	0.99948	$1/(1 + \exp(-1.44645 + 0.85000/0.10000))$	0.99749

ตารางที่ 8 ตัวอย่างการคำนวณทาค่า Hopfield (ต่อ)

5	$((0.99749^0) - (0.03382^0.51640)) + (0.99949^0) + (0.99749^0.63247)$	1.42247	$((0.99749^0.51640) + (0.03382^0.51640)) + (0.99949^0) + (0.99749^0)$	0.51476	$((0.99749^0.77461) + (0.03382^0.91649)) + (0.99949^0) + (0.99749^0)$	1.58655	$((0.99749^0.63247) - (0.03382^0.91649)) + (0.99949^0.91649) + (0.99749^0)$	1.44646
	$1/(1 + \exp(-1.42247 - 0.85000/0.10000))$	0.99677	$1/(1 + \exp(-0.51476 + 0.85000/0.10000))$	0.03383	$1/(1 + \exp(-1.58655 + 0.85000/0.10000))$	0.99948	$1/(1 + \exp(-1.44646 + 0.85000/0.10000))$	0.99740
6	$((0.99677^0) + (0.03383^0.51640)) + (0.99949^0.63247) + (0.99740^0.63247)$	1.42244	$((0.03383^0.51640) + (0.99949^0.63247)) + (0.99677^0.51640) + (0.99740^0)$	0.51476	$((0.99740^0.77461) + (0.03383^0.91649)) + (0.99949^0) + (0.99740^0.91649)$	1.58654	$((0.99740^0.63247) - (0.03383^0.91649)) + (0.99949^0.91649) + (0.99740^0)$	1.44640
	$1/(1 + \exp(-1.42244 - 0.85000/0.10000))$	0.99670	$1/(1 + \exp(-0.51476 + 0.85000/0.10000))$	0.03385	$1/(1 + \exp(-1.58654 + 0.85000/0.10000))$	0.99949	$1/(1 + \exp(-1.44640 + 0.85000/0.10000))$	0.99742
7	$((0.99670^0) + (0.03385^0.51640)) + (0.99949^0.63247) + (0.99949^0.63247)$	1.42245	$((0.03385^0.51640) + (0.99949^0.63247)) + (0.99742^0.51640) + (0.99742^0)$	0.51476	$((0.99742^0.77461) + (0.03385^0.91649)) + (0.99949^0) + (0.99742^0.91649)$	1.58657	$((0.99742^0.63247) - (0.03385^0.91649)) + (0.99949^0.91649) + (0.99742^0)$	1.44646
	$1/(1 + \exp(-1.42245 - 0.85000/0.10000))$	0.99675	$1/(1 + \exp(-0.51476 + 0.85000/0.10000))$	0.03386	$1/(1 + \exp(-1.58657 + 0.85000/0.10000))$	0.99947	$1/(1 + \exp(-1.44646 + 0.85000/0.10000))$	0.99749

จากตารางที่ 8 เป็นตารางแสดงการคำนวณการหาค่า Hopfield ของประโยคในแต่ละกลุ่ม เพื่อค้นหาประโยคที่ดีที่สุด โดยจะทำการคำนวณรอบต่อไปเรื่อยๆ จนกว่าค่ามีการเปลี่ยนแปลงน้อยที่สุด แล้วจึงนำค่าสุดท้ายที่ได้มาหาค่าที่มากที่สุด แล้วจึงทำการเลือกประโยคที่มีค่า Hopfield มากที่สุด มาเป็นประโยคสรุปว่ามีใจความสำคัญที่สุดในแต่ละกลุ่ม

ผลลัพธ์จากการคำนวณได้ดังนี้

1. ล้างออกง่ายสะอาดดี	ค่า Hopfield เท่ากับ	0.99675
2. สะอาดดีมาก	ค่า Hopfield เท่ากับ	0.03386
3. ล้างออกง่าย	ค่า Hopfield เท่ากับ	0.99947
4. ล้างง่าย	ค่า Hopfield เท่ากับ	0.99749

เมื่อเรียงประโยคตามค่าของ Hopfield จะได้ประโยคดังต่อไปนี้

ล้างออกง่าย > ล้างง่าย > ล้างออกง่ายสะอาดดี > สะอาดดีมาก

เพราะฉะนั้น ประโยคที่มีค่ามากที่สุดเป็นประโยคสรุปของอัลกอริทึม Hopfield คือ ประโยค “ล้างออกง่าย”

มหาวิทยาลัยศิลปากร สวทช. สวทช. สวทช.

การคำนวณโดยใช้ TF-ID

เป็นการหาค่าโดยนำค่าในเอกสารมาเขียนให้อยู่ในรูปของเวกเตอร์ จากนั้นจะทำการหาค่าส่วนกลับความถี่ของคำในเอกสาร (idf) เพื่อบ่งบอกถึงความสำคัญของคำ จะได้ว่าประโยคที่มีค่ามากที่สุดเป็นประโยคสรุปของ TF-IDF

4.2.4 การเลือกประโยคที่มีคะแนนเท่ากัน

จากการประมวลผลที่ได้จากขั้นตอนข้างต้น จะได้คะแนนมาเรียงลำดับ โดยคะแนนที่มากที่สุดถือว่าเป็นประโยคที่ถูกเลือกจากกลุ่มความคิดเห็นนั้นว่าเป็นประโยคที่มีความสำคัญที่สุด ในกรณีที่คะแนนเท่ากันจะใช้การหาค่าความถี่ของเอกสาร (TF-IDF) มาใช้ในการเลือกประโยค ถ้าประโยคใดมีค่า (TF-IDF) มากที่สุดจะเลือกประโยคนั้น

ตัวอย่างเช่น ประโยคที่ 1 และประโยคที่ 2 มีค่าเท่ากัน

ประโยคที่ 1 ปรับสีผิวหน้าให้ดูสว่าง

ประโยคที่ 2 ปรับสีผิวให้สว่างขึ้น

4.3 แผนการดำเนินงาน

ตารางที่ 12 ขั้นตอนและแผนการดำเนินงาน

การดำเนินงาน	เดือน												
	มิ.ย. 59	ก.ค. 59	ส.ค. 59	ก.ย. 59	ต.ค. 59	พ.ย. 59	ธ.ค. 59	ม.ค. 60	ก.พ. 60	มี.ค. 60	เม.ย. 60	พ.ค. 60	
1. ศึกษางานวิจัย ก่อนหน้า	←—————→												
2. ศึกษาเครื่องมือ และทฤษฎีที่ใช้ใน งานวิจัย	←—————→												
3. เสนอหัวข้อ งานวิจัย				←————→									
4. วิเคราะห์และ ออกแบบระบบ				←—————→									
5. พัฒนาระบบ			←—————→										
6. ทดสอบและ ประเมินระบบ								←—————→					
7. สรุปงานวิจัย และจัดทำเอกสาร								←—————→					

บทที่ 5

ผลการดำเนินงานวิจัย

การวัดประสิทธิภาพ (ROUGE)

ในการวัดประสิทธิภาพนั้นจะใช้การวัดประสิทธิภาพแบบ ROUGE (Chin-Yew Lin, 2004) โดย ROUGE คือ ตัววัดปัญหาการสรุปข้อความ โดยใช้ n-gram และคำที่ซ้ำกัน เป็นมาตรฐานหลักที่ใช้วัดผลระหว่างผลลัพธ์ที่ได้จากอัลกอริทึมกับผลลัพธ์ที่ได้จากผู้ทดสอบ เราใช้ ROUGE-1, ROUGE-2, ROUGE-3, ROUGE-4 และ ROUGE-SU4 ในการวัดประสิทธิภาพ ROUGE โดยวัดผลลัพธ์ที่ได้จากอัลกอริทึมPageRank อัลกอริทึมHopfield และ TF-IDF กับผลลัพธ์ที่ได้จากผู้ทดสอบในการสรุปประโยค ซึ่งผู้ทดสอบที่ใช้ คือผู้เชี่ยวชาญทางด้านเครื่องสำอางทั้งหมด 5 คน สามารถคำนวณได้จากสมการที่ 9

5.1 ผลการดำเนินงานวิจัย

ผลการทดลองแบบที่ 1 คือ

1. ใช้อัลกอริทึมนาอิวเบย์ (Naive Bayes) ในการจัดกลุ่มเพื่อจำแนกคุณสมบัติที่สนใจตามประเภทของผลิตภัณฑ์
2. จัดกลุ่มความคิดเห็นเชิงบวก เชิงลบของคำโดยมีค่าน้ำหนักดังนี้
 - กลุ่ม “0” ความคิดเห็นเป็นกลาง (Neutral word)
 - กลุ่ม “+1” ความคิดเห็นเชิงบวก (Positive word)
 - กลุ่ม “-1” ความคิดเห็นเชิงลบ (Negative word)
3. หาความคล้ายคลึงของเอกสาร (Cosine similarity)
4. ใช้ PageRank algorithm, Hopfield Network algorithm และ Term Frequency Inverted Document Frequency เพื่อทำการสรุปประโยคที่เป็นตัวแทนของความคิดเห็นของผลิตภัณฑ์

ผลการทดลองแบบที่ 2 คือ

1. ใช้อัลกอริทึมนาอิวเบย์ (Naive Bayes) ในการจัดกลุ่มเพื่อจำแนกคุณสมบัติที่สนใจตามประเภทของผลิตภัณฑ์
2. กลุ่มความคิดเห็นเชิงบวก เชิงลบของคำโดยมีค่าน้ำหนักดังนี้
 - กลุ่ม “0” ความคิดเห็นเป็นกลาง (Neutral word)
 - กลุ่ม “+1” ความคิดเห็นเชิงบวก (Positive word)
 - กลุ่ม “-1” ความคิดเห็นเชิงลบ (Negative word)
 - กลุ่ม “-2” (Very Negative word)
3. หาความคล้ายคลึงของเอกสาร (Cosine similarity)
4. ใช้ PageRank algorithm, Hopfield Network algorithm และ Term Frequency Inverted Document Frequency เพื่อทำการสรุปประโยคที่เป็นตัวแทนของความคิดเห็นของผลิตภัณฑ์

ผลการทดลองแบบที่ 3 คือ

1. ใช้อัลกอริทึมนาอิวเบย์ (Naive Bayes) จัดกลุ่มความคิดเห็นของคำด้วยการแยกกลุ่มคำในการฝึกสอน (Train) ออกเป็นกลุ่มของคำที่เป็นบวกและกลุ่มคำที่เป็นลบ เช่น สีน้าประเภทที่ปิดแก้ม (Blush on) แบ่งออกเป็น 4 กลุ่ม กลุ่มละ 2 แบบ
2. หาความคล้ายคลึงของเอกสาร (Cosine similarity)
3. ใช้ PageRank algorithm, Hopfield Network algorithm และ Term Frequency Inverted Document Frequency เพื่อทำการสรุปประโยคที่เป็นตัวแทนของความคิดเห็นของผลิตภัณฑ์

ผลิตภัณฑ์เครื่องสำอาง 3 ผลิตภัณฑ์ ผลิตภัณฑ์ละ 3 ยี่ห้อ มีทั้งหมด 9 แบบ (Product) ดังนี้

ประเภทที่ปิดแก้ม (Blush on)

1. M.A.C-PowderBlush
2. NARS-Blush
3. OP-ReadyToWearAutomaticFace

ประเภทผลิตภัณฑ์ทำความสะอาดผิวหน้า (Facial cleanser)

4. BIODERMA-SensibioH2O
5. Cetaphil-GentleSkinCleanser

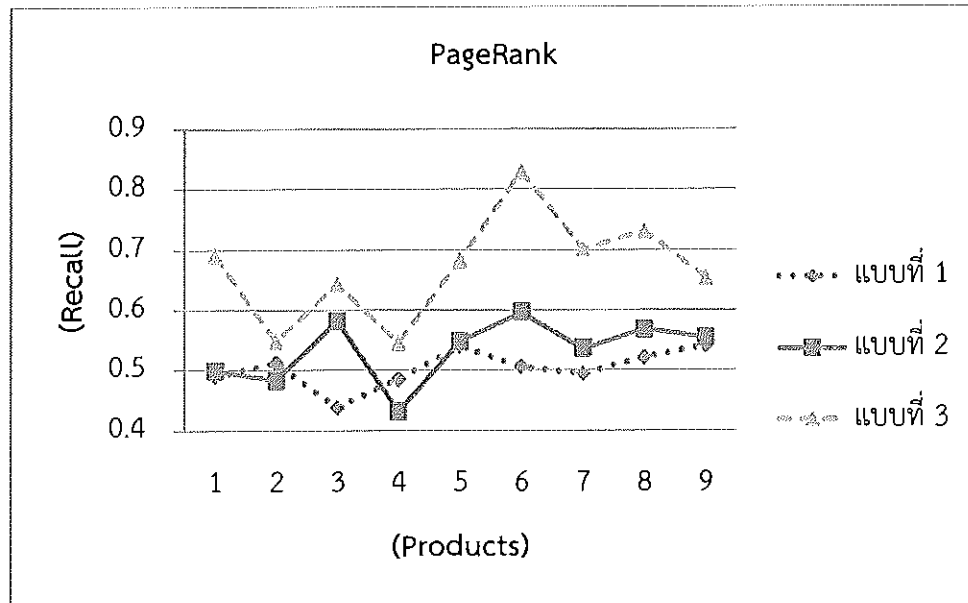
6. DHC-DeepCleansingOil
ประเภทครีมรองพื้น (Foundation)
7. CathyDoll-CCCreamSPF50PA+++
8. Revlon-ColorstayMakeupWithSoftflexforCombinationOilySkin
9. Revlon-PhotoReadyFoundation

5.2 การวัดประสิทธิภาพ (ROUGE)

การวัดประสิทธิภาพใช้โปรแกรม ROUGE (Recall Oriented Understudy for Gisting Evaluation) ผลลัพธ์ที่ได้จากการวัดประสิทธิภาพข้อมูลสามารถแสดงค่าได้ดังตารางที่ 13, 14 และ 15

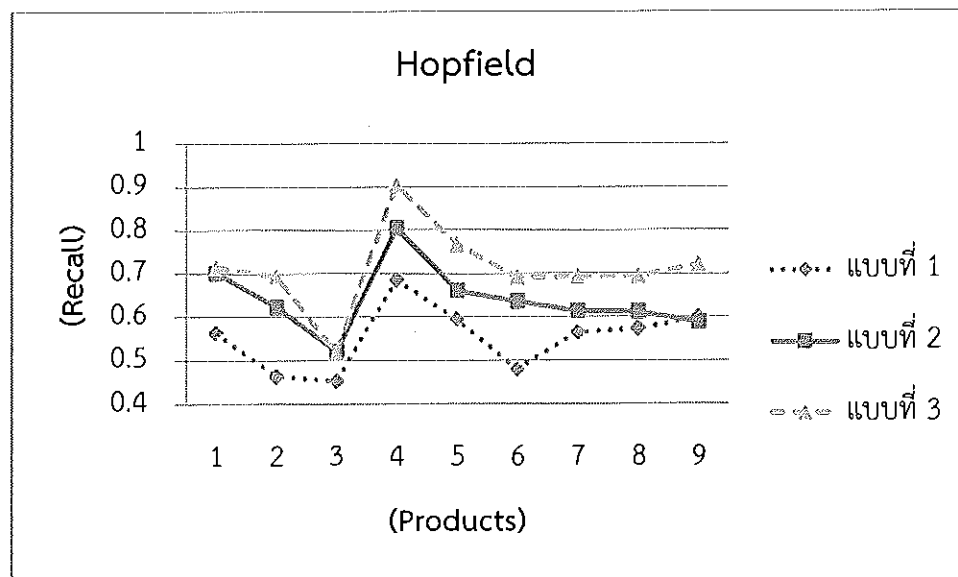
ตารางที่ 13 ค่าการวัดประสิทธิภาพผลิตภัณฑ์ของ ROUGE-1 (Recall)

Product	PageRank			Hopfield			TF-IDF		
	แบบที่ 1	แบบที่ 2	แบบที่ 3	แบบที่ 1	แบบที่ 2	แบบที่ 3	แบบที่ 1	แบบที่ 2	แบบที่ 3
1	0.48999	0.49957	0.69119	0.56355	0.70133	0.71304	0.46406	0.62016	0.70791
2	0.51228	0.48355	0.54785	0.63770	0.62385	0.69449	0.49553	0.66398	0.68778
3	0.43766	0.58182	0.64322	0.35261	0.52216	0.51757	0.66111	0.66038	0.67199
4	0.48501	0.43148	0.54534	0.68478	0.80480	0.90462	0.39322	0.48784	0.55293
5	0.54045	0.54804	0.68230	0.59435	0.66113	0.76707	0.49748	0.57272	0.67032
6	0.50556	0.59622	0.79984	0.47919	0.63618	0.69096	0.49416	0.56859	0.72674
7	0.49375	0.53596	0.70085	0.56441	0.61361	0.69439	0.43762	0.47518	0.69185
8	0.52131	0.56857	0.70053	0.57275	0.61299	0.69444	0.52206	0.55694	0.60782
9	0.54154	0.55496	0.65320	0.59959	0.58912	0.72120	0.53663	0.57216	0.67048

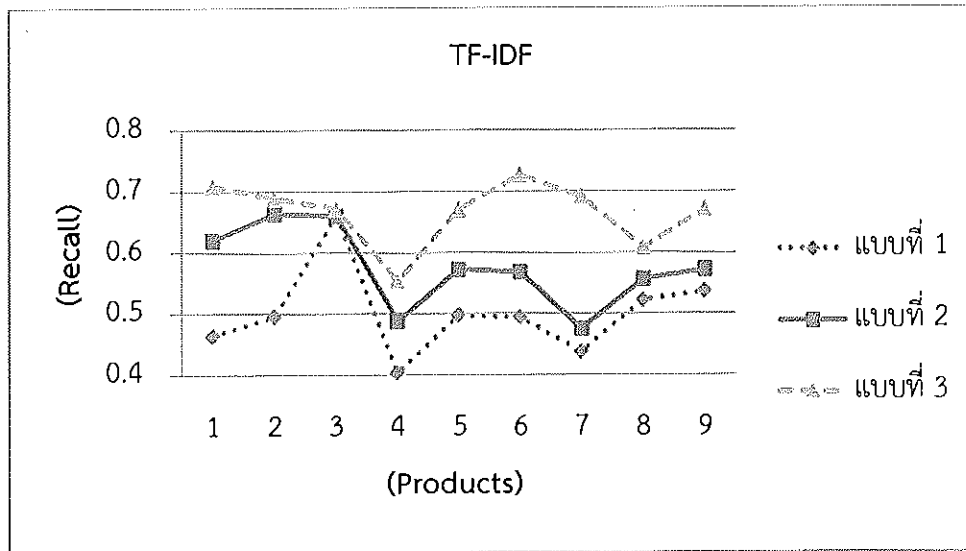


รูปที่ 30 สรุปการวัดประสิทธิภาพจากอัลกอริทึม PageRank ของ ROUGE-1 (Recall)

มหาวิทยาลัยศิลปากร สวทช. สวทช.



รูปที่ 31 สรุปการวัดประสิทธิภาพจากอัลกอริทึม Hopfield ของ ROUGE-1 (Recall)

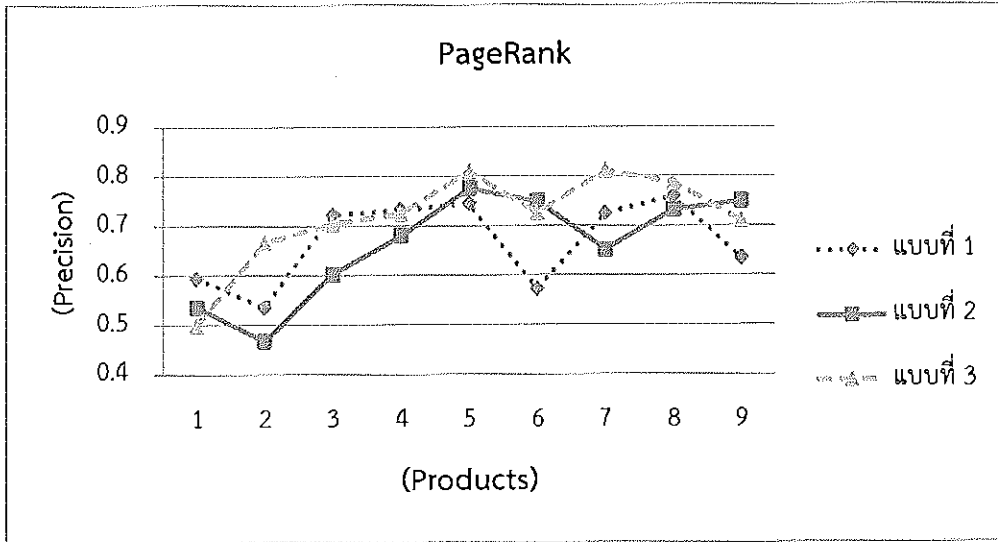


รูปที่ 32 สรุปการวัดประสิทธิภาพจากอัลกอริทึม TF-IDF ของ ROUGE-1 (Recall)

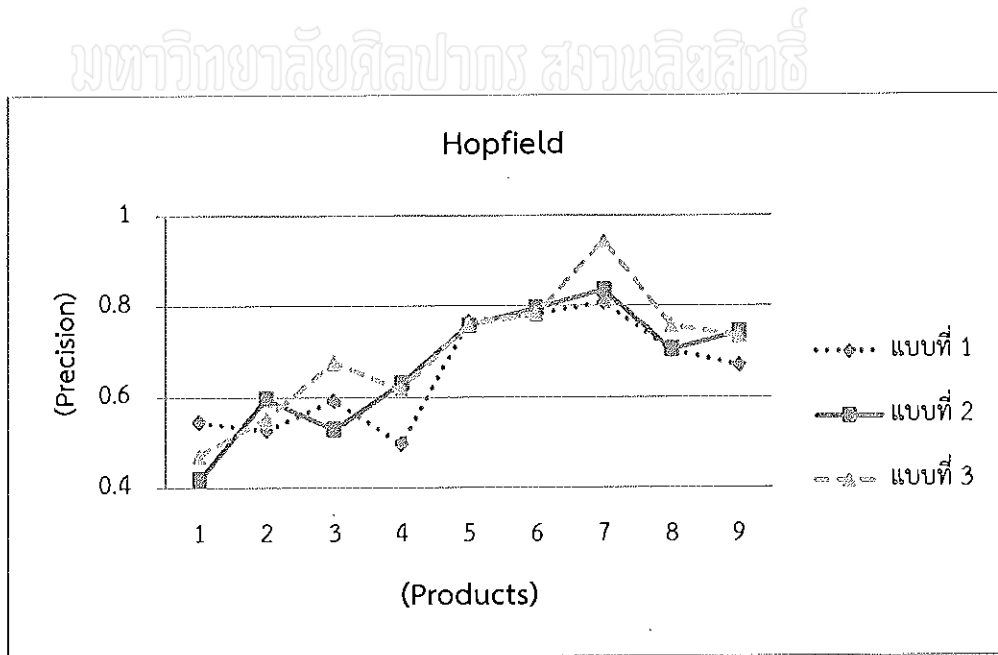
จากรูปที่ 30, 31 และ 32 ผลลัพธ์จากวิธีการแบบที่ 3 เป็นวิธีที่ดีที่สุดสำหรับการเลือกประโยคที่เป็นตัวแทนของผลิตภัณฑ์

ตารางที่ 14 ค่าการวัดประสิทธิภาพผลิตภัณฑ์ของ ROUGE-1 (Precision)

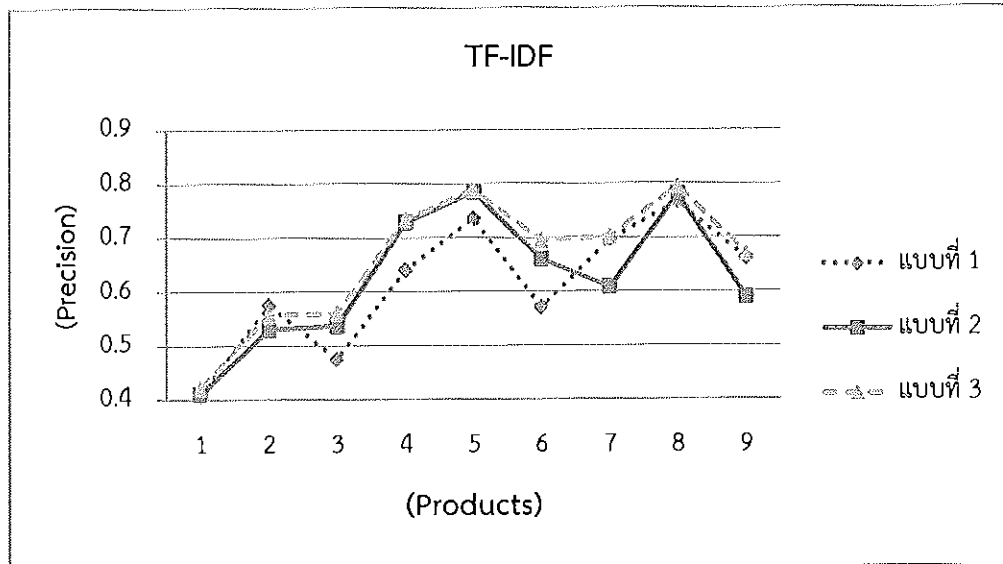
Product	PageRank			Hopfield			TF-IDF		
	แบบที่ 1	แบบที่ 2	แบบที่ 3	แบบที่ 1	แบบที่ 2	แบบที่ 3	แบบที่ 1	แบบที่ 2	แบบที่ 3
1	0.59342	0.53614	0.49948	0.54379	0.31848	0.47074	0.36788	0.37079	0.42214
2	0.53530	0.46793	0.66583	0.52640	0.59625	0.55044	0.57509	0.43017	0.45631
3	0.72222	0.60272	0.58405	0.59167	0.52897	0.67611	0.47440	0.53746	0.46062
4	0.73274	0.68006	0.65499	0.49660	0.63087	0.61530	0.63982	0.76886	0.71308
5	0.74464	0.79560	0.50970	0.76546	0.75813	0.44126	0.73598	0.78356	0.48887
6	0.57227	0.74958	0.45601	0.78162	0.79631	0.78247	0.57052	0.65992	0.59428
7	0.72473	0.65229	0.81127	0.80700	0.83589	0.94460	0.69578	0.60867	0.70138
8	0.75907	0.73217	0.78726	0.70261	0.70461	0.75647	0.76602	0.77968	0.79481
9	0.63508	0.74974	0.71079	0.66845	0.74295	0.67393	0.66163	0.59008	0.67020



รูปที่ 33 สรุปการวัดประสิทธิภาพจากอัลกอริทึม PageRank ของ ROUGE-1 (Precision)



รูปที่ 34 สรุปการวัดประสิทธิภาพจากอัลกอริทึม Hopfield ของ ROUGE-1 (Precision)



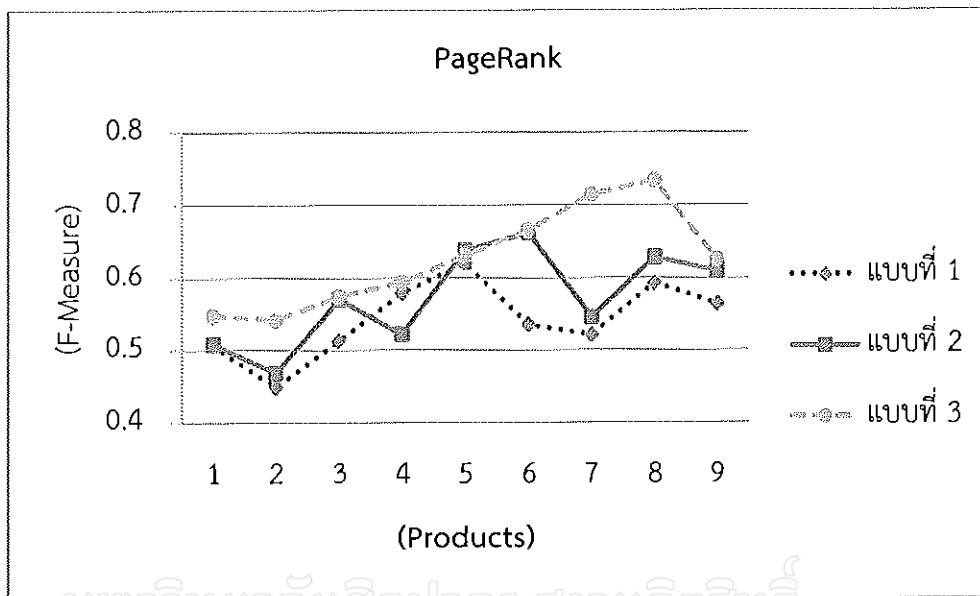
รูปที่ 35 สรุปการวัดประสิทธิภาพจากอัลกอริทึม TF-IDF ของ ROUGE-1 (Precision)

จากรูปที่ 33, 34 และ 35 ผลลัพธ์จากวิธีการแบบที่ 3 เป็นวิธีที่ดีที่สุดสำหรับการเลือกประโยคที่เป็นตัวแทนของผลิตภัณฑ์

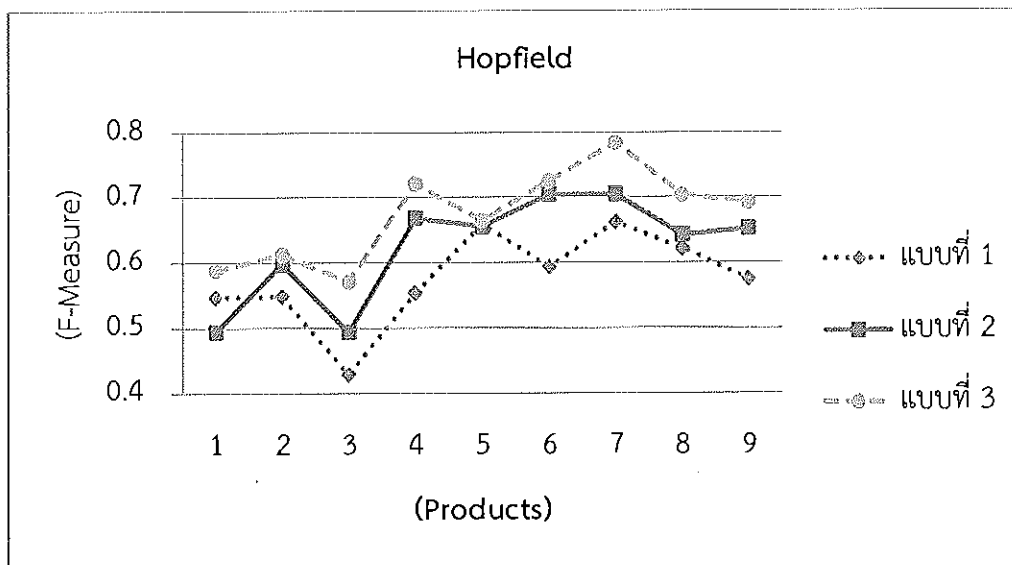
ตารางที่ 15 ค่าการวัดประสิทธิภาพผลิตภัณฑ์ของ ROUGE-1 (F-measure)

Product	PageRank			Hopfield			TF-IDF		
	แบบที่ 1	แบบที่ 2	แบบที่ 3	แบบที่ 1	แบบที่ 2	แบบที่ 3	แบบที่ 1	แบบที่ 2	แบบที่ 3
1	0.50601	0.50837	0.54705	0.54642	0.49381	0.58817	0.40780	0.45253	0.51599
2	0.44840	0.46879	0.54158	0.54767	0.59791	0.61165	0.53221	0.51882	0.52699
3	0.51247	0.57072	0.57427	0.42788	0.49359	0.57051	0.51477	0.57820	0.52403
4	0.57848	0.52227	0.59348	0.55289	0.66810	0.72090	0.48463	0.59409	0.62067
5	0.61983	0.63826	0.63108	0.65934	0.65576	0.66396	0.53194	0.58402	0.60267
6	0.53467	0.66153	0.66556	0.59268	0.70457	0.72451	0.52698	0.60862	0.63019
7	0.52123	0.54594	0.71453	0.66167	0.70500	0.78338	0.47247	0.44171	0.67881
8	0.59198	0.62832	0.73401	0.62099	0.64258	0.70290	0.59568	0.63861	0.67049
9	0.56344	0.60914	0.62443	0.57410	0.65257	0.69217	0.54474	0.52870	0.61886

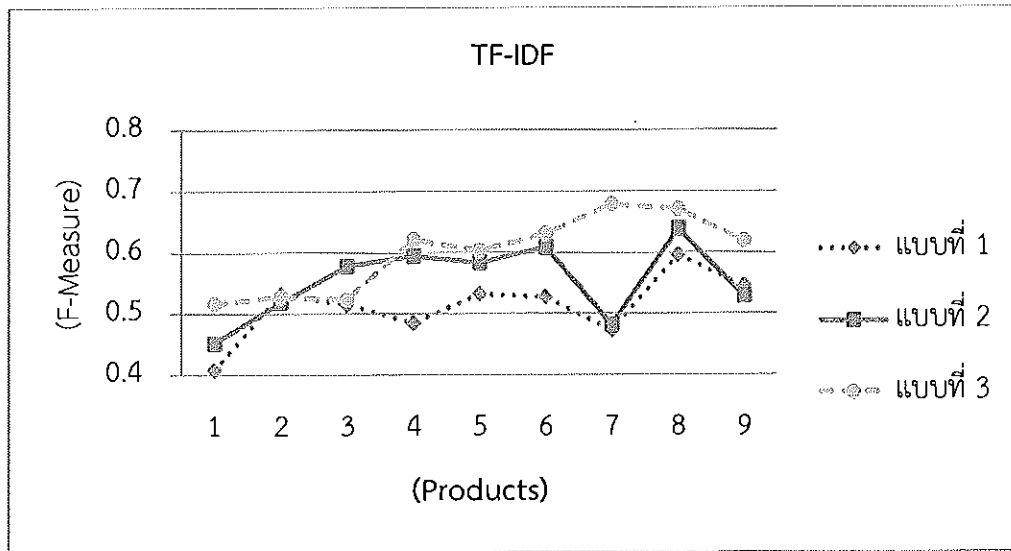
จากตารางที่ 15 สรุปได้ว่าค่าที่ได้จากการวัดประสิทธิภาพของ ROUGE-1 (F-measure) โดยใช้ อัลกอริทึม Hopfield วิธีการแบบที่ 3 ได้ประสิทธิภาพดีที่สุดสำหรับการเลือกประโยคที่เป็นตัวแทนของผลิตภัณฑ์



รูปที่ 36 สรุปการวัดประสิทธิภาพจากอัลกอริทึม PageRank ของ ROUGE-1 (F-measure)



รูปที่ 37 สรุปการวัดประสิทธิภาพจากอัลกอริทึม Hopfield ของ ROUGE-1 (F-measure)



รูปที่ 38 สรุปการวัดประสิทธิภาพจากอัลกอริทึม TF-IDF ของ ROUGE-1 (F-measure)

จากรูปที่ 36, 37 และ 38 ผลลัพธ์จากวิธีการแบบที่ 3 เป็นวิธีที่ดีที่สุดสำหรับการเลือกประโยคที่เป็นตัวแทนของผลิตภัณฑ์

บทที่ 6

สรุปผลการทดลอง และข้อเสนอแนะ

6.1 สรุปผลการทดลอง

งานวิจัยนี้เป็นการวิเคราะห์ความรู้สึกเกี่ยวกับผลิตภัณฑ์ประเภทเครื่องสำอาง และหาค่าความสำคัญของประโยคจากบทวิจารณ์ เพื่อสรุปว่าผลิตภัณฑ์ประเภทเครื่องสำอางมีข้อดีและข้อเสียในด้านใดบ้าง โดยจะนำไปใช้ประกอบการตัดสินใจในการเลือกซื้อหรือเลือกใช้ผลิตภัณฑ์ประเภทเครื่องสำอาง ในการทดลองประกอบไปด้วย 3 ขั้นตอนหลัก คือ ขั้นตอนการเตรียมข้อมูล ขั้นตอนการจัดกลุ่ม และขั้นตอนการประมวลผลข้อมูล โดยสามารถสรุปผลการทดสอบระบบสรุปความคิดเห็นผลิตภัณฑ์เครื่องสำอาง และผลการประเมินประสิทธิภาพโดยใช้ ROUGE สรุปได้ว่าจากข้อมูลที่ทำทดสอบนั้นผลลัพธ์ที่ได้จากอัลกอริทึม Hopfield แบบที่ 3 มีค่าความแม่นยำสูงที่สุด

เพราะแบบที่ 3 มีการจัดกลุ่มความคิดเห็นโดยใช้อัลกอริทึมนาอิวเบย์ (Naive Bayes) จัดกลุ่มความคิดเห็นของประโยคด้วยการแยกกลุ่มประโยคในการฝึกสอน (Train) ออกเป็นกลุ่มของคำที่เป็นความคิดเห็นบวกและกลุ่มคำที่เป็นความคิดเห็นเชิงลบ เช่น สินค้าประเภทที่ปิดแก้ม (Blush on) ออกเป็น 3 กลุ่ม กลุ่มละ 2 แบบ ได้แก่

กลุ่ม “ติดทน” ทางบวก

กลุ่ม “ติดทน” ทางลบ

กลุ่ม “สี/เมื่อดสี” ทางบวก

กลุ่ม “สี/เมื่อดสี” ทางลบ

กลุ่ม “แพกเกจ” ทางบวก

กลุ่ม “แพกเกจ” ทางลบ

จึงทำให้แบบที่ 3 ได้ประโยคความคิดเห็นที่ถูกต้องและแม่นยำมากกว่าวิธีการแบบที่ 1 และแบบที่ 2 เพราะว่าแบบที่ 1 และ แบบที่ 2 มีการจัดกลุ่มความคิดเห็นโดยใช้อัลกอริทึมนาอิวเบย์ (Naive Bayes) จัดกลุ่มความคิดเห็นของประโยค ด้วยการแยกกลุ่มของประโยคในการฝึกสอน (Train) ออกเป็นกลุ่มที่เป็นตัวแทนของสินค้าผลิตภัณฑ์นั้นๆ เช่น สินค้าประเภทที่ปิดแก้ม (Blush on) ออกเป็น 3 กลุ่ม ได้แก่

กลุ่ม “ติดหน”

กลุ่ม “สี่/เม็ดสี่”

กลุ่ม “แพกเก็จ”

เมื่อจัดกลุ่มเสร็จ นำประโยคที่ได้จากการจัดกลุ่มมาคำนวณคะแนนความคิดเห็นเพื่อหาประโยคความคิดเห็นเชิงบวก เชิงลบ ความผิดพลาดของแบบที่ 1 และ แบบที่ 2 เกิดขึ้นจาก คำว่า “ไม่” เป็นคำในในคลังข้อมูลความคิดเห็นเชิงลบ ดังนั้นจึงทำให้ความคิดเห็นของประโยค “ใช้แล้วไม่แพ้อยู่” อยู่ในกลุ่มประโยคของความคิดเห็นเชิงลบ แต่ความเป็นจริงแล้ว ประโยค “ใช้แล้วไม่แพ้อยู่” เป็นประโยคของความคิดเห็นเชิงบวก

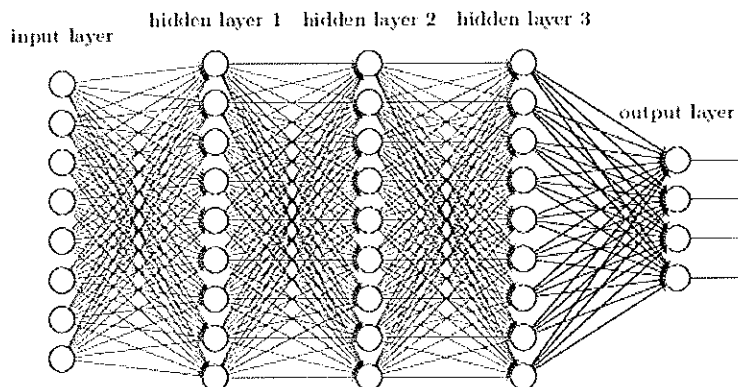
เนื่องจากการทำวิจัยเกี่ยวกับข้อความภาษามีข้อจำกัดของการตัดคำ และตัดประโยคของภาษาไทยที่ประสิทธิภาพยังไม่ดีมากพอ และด้วยการที่ประโยค ข้อความแสดงความคิดเห็นส่วนใหญ่ที่อยู่บนอินเทอร์เน็ตนิยมใช้ภาษาที่มีโครงสร้างประโยคที่ไม่แน่นอน (unstructured data) หรือเป็นภาษาธรรมชาติ (natural language) ไม่ถูกต้องตามหลักไวยากรณ์ทางภาษานอกจากนั้นแล้วยังไม่มีการจัดการเรื่องคำพ้องความหมายทำให้ยากต่อการวิเคราะห์ จึงทำให้ผลการทดลองมีค่าความแม่นยำไม่สูงมาก

6.2. ข้อเสนอแนะและงานที่จะศึกษาต่อ

1. สามารถนำไปพัฒนาให้ใช้งานกับเว็บไซต์ได้หลากหลายมากขึ้น โดยสามารถนำงานวิจัยนี้ไปพัฒนาเพื่อให้ได้การเลือกประโยคที่ดีขึ้น เช่น ใช้ฐานข้อมูลที่มีขนาดใหญ่ขึ้น ใช้ชุดข้อมูลที่หลากหลายชุด

2. สามารถนำไปพัฒนาต่อโดยใช้ Deep learning ในการเลือกประโยคที่ดีที่สุด

Deep Learning (Deep Learning, 2017) คือรูปแบบหนึ่งของ Machine Learning ซึ่งไอดียั้งตั้งต้นของ Deep Learning มาจากแบบจำลอง Machine Learning ชนิดโครงข่ายประสาทเทียม (Artificial Neural Network) ที่เลียนแบบการทำงานของโครงข่ายเซลล์สมองของคนเรา ซึ่งแบบจำลองที่วันนี้ได้รับการพิสูจน์ทางทฤษฎีมาในอดีตว่า ในกรณีที่โครงข่ายภายในลึก (Deep) และมีจำนวนโหนดหรือเซลล์มากพอ ระดับความซับซ้อนของแบบจำลองจะเพียงพอสำหรับแก้ปัญหาที่เราโยนเข้าไปได้ทุกชนิด ดังตัวอย่างในรูปที่ 29



รูปที่ 39 ตัวอย่างโครงสร้าง Deep learning

ที่มา: Deep Learning, accessed March 2, 2017, available from
<http://www.manager.co.th/daily/ViewNews.aspx?NewsID=9590000091323>

3. ทำการจัดกลุ่มความคิดเห็นผลิตภัณฑ์ให้มากขึ้น ครอบคลุมคุณสมบัติต่างๆของผลิตภัณฑ์ เครื่องสำอางแต่ละประเภทแต่ละยี่ห้อ เช่น ประเภทที่ปิดแก้ม (Blush on) เพิ่มเป็น 5 กลุ่ม ได้แก่

- กลุ่ม “ติดทน”
- กลุ่ม “สี/เม็ดสี”
- กลุ่ม “แพคเกจ”
- กลุ่ม “ราคา”
- กลุ่ม “ซิมเมอร์”

4. นำบทวิจารณ์ความคิดเห็นที่เป็นภาษาไทยจากเว็บไซต์เครื่องสำอางที่มีชื่อเสียง จากหลากหลายเว็บไซต์มาใช้ในการสรุปผลเพื่อหาตัวแทนของผลิตภัณฑ์ที่ดีที่สุด

รายการอ้างอิง

- ชนาวุฒิ ประกอบผล. (2552). “โครงข่ายประสาทเทียม.” วารสาร มฉก. วิชาการ 73, ปีที่ 12, ฉบับที่ 24, ม.ค.-มิ.ย.
- Deep Learning. Accessed March 2, 2017. Available from <http://www.manager.co.th/daily/ViewNews.aspx?NewsID=9590000091323>
- PageRank. Accessed July 20, 2016. Available from <https://en.wikipedia.org/wiki/PageRank>
- Tutorial: Building a Text Classification System. Accessed April 5, 2016. Available from <http://textblob.readthedocs.io/en/dev/classifiers.html>
- Thailand web directory and Advance Web Statistics. Accessed January 12, 2016. Available from <http://truehits.net/>
- Thai word breaker for PHP. Accessed May 3, 2016. Available from <https://github.com/veer66/PhlongTalam>
- User Reviews. Accessed August 20, 2016. Available from <http://www.jeban.com/reviews.php>
- Aekkasit Chongsuntornsri and Ohm Sornil. (2006). “An automatic thai text summarization using topic sensitive pagerank.” In Proc. of International Symposium on Communications and Information Technologies (ISCIT '06): 547–552.
- Ahmed A. Mohamed and Sanguthevar Rajasekaran. (2006). “Query-Based Summarization Based on Document Graphs.” In Proceedings of IEEE International Symposium on Signal Processing and Information Technology Vancouver, Canada: 408-410.
- Andrew McCallum and Kamal Nigam. (2003). “A comparison of event models for naïve Bayes text classification.” Journal of Machine Learning Research, Vol. 3, 2003: 1265–1287.
- Chetarin Wongsin, Nattakit Srikarnjanapert and Jantima Polpinij. (2015). “Automatic Feeling Analysis from Opinion Text.” In Proceedings of the Eleventh

National Conference on Computing and Information Technology (NCCIT2015), 2-3 July 2015.

Chin-Yew Lin. (2004). "ROUGE:A Package for Automatic Evaluation of Summaries." In *Proceedings of Workshop on Text Summarization of AC*.

Choochart Haruechaiyasak, Alisa Kongthon, Pornpimon Palingoon and Chatchawal Sangkeettrakarn. (2010). "Constructing Thai Opinion Mining Resource: A Case Study on Hotel Reviews." *Proceedings of the 8th Workshop on Asian Language Resources*: 64–71.

David McG.Squire. (2004). "CSE5230 Tutorial:The Naïve Bays Classifier." *Faculty of Information Technology*, August 19, 2004.

Frank Schilder and Ravikumar Kondadadi. (2008). "FastSum: Fast and accurate query-based multi-document summarization." In *Proceedings of the 46th meeting of the Association for Computational Linguistics*, Columbus, Ohio.

Hsinchun Chen and Tobundorbin Ng. (2003). "An Algorithmic Approach to Concept Exploration in a Large Knowledge Network (automatic Thesaurus Consultation): Symbolic Branch-and-Bound Search vs. Connectionist Hopfield Net Activation." *Journal of the American Society for Information Science* 46.

Hsinchun Chen, Yin Zhang and Andrea L. Houston. (1998). "Semantic indexing and searching using a Hopfield net." *Journal of information Science*, Vol. 24, No. 1: 3-18.

Hyun Duk Kim, Malu G Castellanos, Meichun Hsu, ChengXiang Zhai, Umeshwar Dayal and Riddhiman Ghosh. (2013). "Ranking Explanatory Sentences for Opinion Summarization." In *Proceedings of (SIGIR'13)*, July 28-August 1, Dublin, Ireland.

John M. Conroy, Judith D. Schlesinger and Jade Goldstein Stewart. (2005). "CLASSY Query Based Multi-Document Summarization." In *DUC 05 Conference Proceedings*, Boston, USA.

- Lalita Wonghaeon, Natwara Keereedet, Simaporn Punpauk and Nattapong Tongteng. (2016). "Emotion Categorization System for the Thai Texts." In Proceedings of the twelfth National Conference on Computing and Information Technology (NCCIT2016), 7-8 July 2016.
- Mariana Damova and Ivan Koychev. (2010). "Query-based summarization: a survey." In International Conference on Software, Services and Semantic Technologies (S3T '2010) Varna, Bulgaria: 142-146.
- Minqing Hu and Bing Liu. (2004). "Mining and Summarizing Customer Reviews." In Proceedings of the tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining: 168-177.
- Naris Prombut. (2007). "Mining Opinion in Product Reviews : A Case Study of Mobile Phone Reviews." Master's Project, Department of Computer Science Faculty of Applied Sciences King Mongkut's University of Technology North Bangkok.
- Nattakit Srikarnjanapert and Jantima Polpinij. (2015). "Automatically Rating of Hotel Service Aspects from Textual Customer Reviews." In Proceedings of the Eleventh National Conference on Computing and Information Technology (NCCIT2015), 2-3 July 2015.
- Nongnuch Ketui, Thanaruk Theeramunkong and Chutamane Onsuwan. (2013). "Thai News Text Summarization and Its Application." In Proceedings of the Tenth Symposium on Natural Language Processing (SNLP 2013), 28-30 October, Phuket Thailand: 215-220.
- Ohm Sornil and Kornnika Gree-ut. (2006). "An Automatic Text Summarization Approach Combining Content-Based and Graph-Based Characteristics." In Proceedings of IEEE CIS 2006, Bangkok, Thailand.
- Parnicha Apisuwankun and Janjao Mongkolnavin. (2013). "Opinion Strength Identification in Customer Review Summarization System Using Association Rule Technique." The International Conference on E-Technologies and Business on the Web (EBW2013).

- Stefano Baccianella, Andrea Esuli, and Fabrizio Sebastiani. (2010). "SentiWordNet 3.0: sentiment analysis and opinion mining." In: LREC 2010 Seventh conference on International Language Resources and Evaluation Proceedings: 2200-2204
- Trithep Thumrongluck. (2010). "An Automated System for Summarizing Structured Product Reviews." Master's Thesis, Department of Statistics Faculty of Commerce and Accountancy Chulalongkorn University.
- Vanessa Graham Murdock. (2006). "ASPECTS OF SENTENCE RETRIEVAL." Ph.D. dissertation, University of Massachusetts Amherst, aAI3242373.

มหาวิทยาลัยศิลปากร สงวนลิขสิทธิ์

ภาคผนวก

มหาวิทยาลัยศิลปากร สภานิติศาสตร์

ภาคผนวก ก ผลลัพธ์จากการดำเนินงานวิจัย

ตารางที่ 16 ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภทที่ปิดแก้ม M.A.C-PowderBlush อัลกอริทึม PageRank

Group	ROUGE-1			ROUGE-2			ROUGE-3			ROUGE-SU4		
	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure
แบบที่ 1												
ดีดทม	0.35616	0.52000	0.40276	0.27941	0.42222	0.33628	0.23810	0.37500	0.29127	0.12314	0.33704	0.18038
สี	0.60833	0.76842	0.63907	0.53043	0.67778	0.59512	0.48182	0.62353	0.54359	0.45714	0.60000	0.51892
แม่คเณจ	0.50547	0.49185	0.47622	0.16832	0.13077	0.14719	0.14583	0.11200	0.12670	0.29398	0.19416	0.23386
แบบที่ 2												
ดีดทม	0.49315	0.31304	0.39298	0.30882	0.19091	0.63636	0.25397	0.15238	0.19048	0.23410	0.12582	0.16367
สี	0.60000	0.75789	0.66977	0.52174	0.66667	0.58537	0.43636	0.56471	0.49231	0.35582	0.63069	0.45496
แม่คเณจ	0.40566	0.53750	0.46237	0.23762	0.32000	0.27272	0.20833	0.28571	0.24096	0.19598	0.36148	0.25416
แบบที่ 3												
ดีดทม	0.76712	0.29474	0.42586	0.39706	0.14595	0.21344	0.33333	0.11667	0.17284	0.46008	0.09714	0.15318
สี	0.60833	0.76842	0.67907	0.53043	0.67778	0.59512	0.48182	0.62353	0.54359	0.39224	0.69524	0.50153
แม่คเณจ	0.69811	0.43529	0.53623	0.35644	0.21818	0.27068	0.31250	0.18750	0.23438	0.43373	0.18182	0.25623

ตารางที่ 17 ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภทที่ปิดแก้ม M.A.C-PowderBlush ที่ได้จากอัลกอริทึม Hopfield

Group	ROUGE-1				ROUGE-2				ROUGE-3				ROUGE-SU4			
	Recall	Precision	F-measure		Recall	Precision	F-measure		Recall	Precision	F-measure		Recall	Precision	F-measure	
แบบที่ 1																
ติดทน	0.50685	0.41111	0.45399	0.26471	0.21176	0.23529	0.22222	0.17500	0.19580	0.20027	0.17412	0.18628				
สี	0.60833	0.76842	0.67907	0.53043	0.67778	0.59512	0.4818	0.62353	0.54359	0.39224	0.69524	0.50153				
แก้สีแดง	0.57547	0.45185	0.50622	0.16832	0.13077	0.14719	0.14583	0.11200	0.12670	0.29398	0.19416	0.23386				
แบบที่ 2																
ติดทน	0.69863	0.29143	0.51129	0.26471	0.10588	0.15126	0.19048	0.07273	0.10527	0.37754	0.09871	0.14366				
สี	0.71667	0.33220	0.47229	0.18261	0.07241	0.10370	0.01818	0.00702	0.01013	0.60776	0.11509	0.19353				
แก้สีแดง	0.68868	0.33182	0.49785	0.26733	0.12558	0.17089	0.18750	0.08571	0.11764	0.41205	0.1037	0.16575				
แบบที่ 3																
ติดทน	0.76712	0.39706	0.53333	0.29474	0.14595	0.11667	0.42586	0.21344	0.17284	0.46008	0.09189	0.15318				
สี	0.68333	0.68333	0.68333	0.58261	0.58261	0.58261	0.50000	0.50000	0.50000	0.48478	0.54314	0.51230				
แก้สีแดง	0.68868	0.33182	0.54785	0.26733	0.12558	0.17089	0.18750	0.08571	0.11764	0.41205	0.10374	0.16575				

ตารางที่ 18 ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภทที่ปิดแก้ม M.A.C-PowderBlush ที่ได้จากอัลกอริทึม TF-IDF

Group	ROUGE-1			ROUGE-2			ROUGE-3			ROUGE-SU4		
	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure
แบบที่ 1												
ดีดทม	0.67123	0.44545	0.53551	0.50000	0.32381	0.39306	0.44444	0.28000	0.34356	0.47767	0.28016	0.35318
สี	0.48333	0.42963	0.45490	0.09565	0.08462	0.08980	0.01818	0.01600	0.01702	0.20955	0.18621	0.19719
แพ็คแดง	0.23762	0.22857	0.23301	0.17708	0.17000	0.17347	0.17708	0.17000	0.17347	0.23775	0.23492	0.23633
แบบที่ 2												
ดีดทม	0.73973	0.30857	0.43548	0.39706	0.15882	0.22689	0.33333	0.12727	0.18421	0.41543	0.09762	0.15809
สี	0.80000	0.48000	0.60000	0.53043	0.31282	0.39355	0.44545	0.25789	0.32666	0.57970	0.23712	0.33657
แพ็คแดง	0.32075	0.32381	0.32227	0.02970	0.03000	0.02985	0.01042	0.01053	0.01047	0.08675	0.09391	0.09019
แบบที่ 3												
ดีดทม	0.76712	0.29474	0.42586	0.39706	0.14595	0.21344	0.33333	0.11667	0.17284	0.46008	0.09189	0.15318
สี	0.80000	0.48000	0.60000	0.53043	0.31282	0.39355	0.44545	0.25789	0.32666	0.57970	0.23712	0.33657
แพ็คแดง	0.55660	0.49167	0.52212	0.24752	0.21739	0.23148	0.17708	0.15455	0.16505	0.29157	0.24281	0.26497

ตารางที่ 19 ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภทที่ปิดแก้ม NARS-Blush อัลกอริทึม PageRank

Group	ROUGE-1			ROUGE-2			ROUGE-3			ROUGE-SU4		
	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure
แบบที่ 1												
ดีดทนม	0.56604	0.41379	0.47809	0.18812	0.13571	0.15767	0.06250	0.04444	0.05195	0.20515	0.12857	0.15807
สี	0.40515	0.44545	0.32343	0.22826	0.20000	0.2132	0.11494	0.10000	0.10695	0.25955	0.21032	0.23236
แม่สีแดง	0.56566	0.74667	0.54368	0.40426	0.54286	0.46342	0.34831	0.47692	0.40259	0.26267	0.48381	0.34048
แบบที่ 2												
ดีดทนม	0.56604	0.41379	0.47809	0.18812	0.13571	0.15767	0.06250	0.04444	0.05195	0.17951	0.11133	0.13743
สี	0.28866	0.40000	0.33533	0.15217	0.21538	0.17834	0.08046	0.11667	0.09524	0.06674	0.13626	0.08960
แม่สีแดง	0.59596	0.59000	0.59297	0.40426	0.40000	0.40212	0.31461	0.31111	0.31285	0.2864	0.29158	0.28899
แบบที่ 3												
ดีดทนม	0.81132	0.53750	0.64662	0.58416	0.38065	0.46094	0.52083	0.33333	0.40650	0.46154	0.28966	0.35594
สี	0.19588	0.76000	0.31148	0.14130	0.6500	0.23214	0.08046	0.46667	0.13726	0.04799	0.70000	0.08982
แม่สีแดง	0.63636	0.70000	0.66666	0.40426	0.44706	0.42458	0.33708	0.37500	0.35503	0.33648	0.42000	0.37363

ตารางที่ 20 ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภทที่ปิดแก้ม NARS-Blush ที่ได้จากอัลกอริทึม Hopfield

Group	ROUGE-1			ROUGE-2			ROUGE-3			ROUGE-SU4		
	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure
แบบที่ 1												
ดีดทน	0.40566	0.66154	0.50292	0.25743	0.43333	0.32298	0.18750	0.32727	0.23841	0.12647	0.38222	0.19005
สี	0.61856	0.40000	0.48583	0.19565	0.12414	0.15190	0.10345	0.06429	0.07930	0.29709	0.12690	0.17784
แพ็คเกจ	0.88889	0.51765	0.65428	0.64894	0.36970	0.47105	0.56180	0.31250	0.40161	0.47619	0.25806	0.33472
แบบที่ 2												
ดีดทน	0.70755	0.46875	0.56391	0.43564	0.28387	0.34375	0.33333	0.21333	0.26016	0.44485	0.22960	0.30288
สี	0.61856	0.60000	0.60914	0.44565	0.43158	0.43850	0.36782	0.35556	0.36159	0.39079	0.38182	0.38625
แพ็คเกจ	0.54545	0.72000	0.62069	0.47872	0.64286	0.54878	0.43820	0.60000	0.50649	0.31291	0.55798	0.40096
แบบที่ 3												
ดีดทน	0.70755	0.46875	0.56391	0.43564	0.28387	0.34375	0.33333	0.21333	0.26016	0.44485	0.22960	0.30288
สี	0.60825	0.49167	0.54378	0.42391	0.33913	0.37681	0.35632	0.28182	0.31472	0.38492	0.26288	0.31240

ตารางที่ 21 ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภทที่ปิดแก้ม NARS-Blush ที่ได้จากอัลกอริทึม TF-IDF

Group	ROUGE-1			ROUGE-2			ROUGE-3			ROUGE-SU4		
	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure
แบบที่ 1												
ติดทน	0.51887	0.57895	0.54727	0.24752	0.27778	0.26178	0.19792	0.22353	0.20995	0.15385	0.25503	0.20911
สี	0.40206	0.48750	0.44068	0.15217	0.18667	0.16766	0.08046	0.10000	0.08917	0.12243	0.18519	0.14741
เนื้อผิว	0.56566	0.65882	0.60870	0.38298	0.45000	0.41379	0.30337	0.36000	0.32927	0.27804	0.38816	0.32400
แบบที่ 2												
ติดทน	0.70755	0.36585	0.48231	0.33663	0.17000	0.22591	0.23958	0.11795	0.15808	0.42059	0.13302	0.20212
สี	0.57732	0.38621	0.46281	0.10870	0.07143	0.08621	0.01149	0.00741	0.00901	0.30264	0.14240	0.19367
เนื้อผิว	0.70707	0.53846	0.61135	0.53191	0.40000	0.45662	0.46067	0.34167	0.39235	0.49859	0.30229	0.37638
แบบที่ 3												
ติดทน	0.73585	0.35455	0.47853	0.36634	0.17209	0.23418	0.27083	0.12381	0.16993	0.46985	0.12922	0.20269
สี	0.71134	0.33659	0.45696	0.18478	0.08500	0.11644	0.05747	0.02564	0.03546	0.43585	0.09293	0.15153
เนื้อผิว	0.61616	0.67778	0.64550	0.48936	0.54118	0.51397	0.42697	0.47500	0.44971	0.36194	0.45752	0.40416

ตารางที่ 22 ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภทที่ปิดแถม OP-ReadyToWearAutomaticFace อัลกอริทึม PageRank

Group	ROUGE-1			ROUGE-2			ROUGE-3			ROUGE-SU4		
	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure
แบบที่ 1												
ดีดกน	0.55263	0.70000	0.61765	0.40845	0.52727	0.46032	0.34848	0.46000	0.39655	0.27869	0.37778	0.32076
ลี	0.54000	0.60000	0.56842	0.37895	0.42353	0.40000	0.26667	0.30000	0.28235	0.28863	0.38235	0.32894
แพ็ดแกจ	0.22034	0.86667	0.35135	0.14815	0.80000	0.25000	0.08163	0.80000	0.14814	0.05115	0.80000	0.09615
แบบที่ 2												
ดีดกน	0.56579	0.50588	0.53416	0.40845	0.36250	0.38411	0.33333	0.29333	0.31205	0.32589	0.28816	0.30587
ลี	0.40000	0.72727	0.51613	0.25263	0.48000	0.33103	0.14444	0.28889	0.19259	0.07059	0.15000	0.09600
แพ็ดแกจ	0.77966	0.57500	0.66187	0.64815	0.46667	0.54264	0.55102	0.38571	0.45378	0.56777	0.32889	0.41651
แบบที่ 3												
ดีดกน	0.75000	0.36774	0.49350	0.36620	0.17333	0.23529	0.28788	0.13103	0.18009	0.51488	0.13980	0.21989
ลี	0.40000	0.72727	0.51613	0.25263	0.48000	0.33103	0.14444	0.28889	0.19259	0.07059	0.15000	0.09600
แพ็ดแกจ	0.77966	0.65714	0.71318	0.59259	0.49231	0.53762	0.42857	0.35000	0.38532	0.58568	0.44038	0.50274

ตารางที่ 23 ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภทที่ปิดแถม OP-ReadyToWearAutomaticFace ที่ได้จากอัลกอริทึม

Group	ROUGE-1			ROUGE-2			ROUGE-3			ROUGE-SU4		
	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure
แบบที่ 1												
ดีดทนม	0.42105	0.80000	0.55172	0.35211	0.71429	0.47170	0.28788	0.63333	0.39583	0.18006	0.69143	0.28572
สี่	0.23000	0.57500	0.32857	0.06316	0.17143	0.09231	0.18788	0.25335	0.21584	0.05329	0.34286	0.09224
แม่คเภา	0.40678	0.40000	0.40336	0.31481	0.30909	0.31192	0.22449	0.22000	0.22222	0.19437	0.19740	0.19587
แบบที่ 2												
ดีดทนม	0.42105	0.80000	0.55172	0.35211	0.71429	0.47170	0.28788	0.63333	0.39583	0.18006	0.69143	0.28572
สี่	0.62000	0.47692	0.53913	0.38947	0.29600	0.33636	0.26667	0.20000	0.22857	0.37567	0.24171	0.29416
แม่คเภา	0.52542	0.31000	0.38994	0.44444	0.25263	0.32215	0.36735	0.20000	0.20000	0.37084	0.13876	0.20195
แบบที่ 3												
ดีดทนม	0.64474	0.65333	0.64901	0.46479	0.47143	0.46809	0.37879	0.38462	0.38168	0.38839	0.43866	0.41200
สี่	0.23000	0.57500	0.32857	0.06316	0.17143	0.09231	0.22000	0.35000	0.21429	0.05329	0.34286	0.09224
แม่คเภา	0.67797	0.80000	0.73395	0.61111	0.73333	0.66666	0.55102	0.67500	0.60674	0.49105	0.71111	0.58094

ตารางที่ 24 ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภทที่ปิดแถม OP-ReadyToWearAutomaticFace ที่ได้จากอัลกอริทึม TF-IDF

Group	ROUGE-1			ROUGE-2			ROUGE-3			ROUGE-SU4		
	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure
แบบที่ 1												
ดีดทนม	0.72368	0.32353	0.44715	0.25352	0.10909	0.15254	0.09091	0.03750	0.05310	0.44196	0.1000	0.16310
สี่	0.48000	0.32000	0.38400	0.15789	0.10345	0.12500	0.03333	0.02143	0.02609	0.20959	0.10172	0.13697
แหม้คแกจ	0.77966	0.77966	0.71318	0.59259	0.59259	0.53782	0.42857	0.35000	0.38532	0.58568	0.58568	0.50274
แบบที่ 2												
ดีดทนม	0.76316	0.46400	0.57712	0.52113	0.30833	0.38743	0.39394	0.22609	0.28730	0.47917	0.19877	0.28098
สี่	0.54000	0.34839	0.42353	0.20000	0.12667	0.15510	0.03333	0.02069	0.02553	0.25755	0.11717	0.16106
แหม้คแกจ	0.67797	0.80000	0.73395	0.61111	0.73333	0.66666	0.55102	0.67500	0.60674	0.49105	0.71111	0.58094
แบบที่ 3												
ดีดทนม	0.77632	0.36875	0.50000	0.47887	0.21935	0.30088	0.33333	0.14667	0.20371	0.50744	0.12941	0.20623
สี่	0.46000	0.57500	0.51111	0.22105	0.28000	0.24706	0.12222	0.15714	0.13750	0.21581	0.36000	0.26985
แหม้คแกจ	0.77966	0.43810	0.56098	0.59259	0.32000	0.41558	0.44898	0.23158	0.30556	0.62404	0.21217	0.31667

ตารางที่ 25 ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภททำความสะอาด BIODERMA-SensibioH2O อัลตราไวท์ม PageRank

Group	ROUGE-1			ROUGE-2			ROUGE-3			ROUGE-SU4		
	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure
แบบที่ 1												
กลิ่น	0.36957	0.56667	0.44737	0.21951	0.36000	0.27273	0.11111	0.20000	0.14286	0.13136	0.31000	0.18453
ความสะอาด	0.40000	0.74286	0.52000	0.30000	0.60000	0.40000	0.23636	0.52000	0.32500	0.14706	0.55556	0.23256
แพ้/อ่อนโยน	0.58427	0.65000	0.61538	0.47619	0.53333	0.50314	0.46835	0.52857	0.49664	0.37309	0.54222	0.44203
ราคา	0.58621	0.97143	0.73119	0.50450	0.86154	0.63636	0.47170	0.83333	0.60241	0.27091	0.94038	0.42064
แบบที่ 2												
กลิ่น	0.48404	0.70000	0.57233	0.32240	0.47200	0.38311	0.27528	0.40833	0.32886	0.22787	0.47657	0.30832
ความสะอาด	0.42857	0.84000	0.56757	0.37634	0.77778	0.50724	0.32955	0.72500	0.45313	0.10792	0.75185	0.18875
แพ้/อ่อนโยน	0.45946	0.52308	0.48921	0.27711	0.65714	0.38983	0.18861	0.80513	0.30562	0.21455	0.45908	0.29243
ราคา	0.35385	0.65714	0.46000	0.30435	0.46667	0.36842	0.16000	0.40000	0.22857	0.12000	0.51429	0.19459
แบบที่ 3												
กลิ่น	0.48404	0.70000	0.57233	0.32240	0.47200	0.38311	0.27528	0.40833	0.32886	0.22787	0.47657	0.30832
ความสะอาด	0.63218	0.73333	0.67901	0.49704	0.57931	0.53503	0.44512	0.52143	0.48026	0.41160	0.56293	0.47552
แพ้/อ่อนโยน	0.48235	0.53478	0.50721	0.32000	0.35556	0.33684	0.26531	0.29545	0.27957	0.24562	0.30648	0.27270
ราคา	0.58278	0.65185	0.61538	0.35616	0.40000	0.37681	0.24823	0.28000	0.26316	0.32251	0.40584	0.35941

ตารางที่ 26 ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภททำความสะอาด BIODERMA-SensibioH2O ได้จากอัลกอริทึม Hopfield

Group	ROUGE-1			ROUGE-2			ROUGE-3			ROUGE-SU4		
	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure
แบบที่ 1												
กลิ่น	0.76712	0.29474	0.42586	0.39706	0.14595	0.21344	0.33333	0.11667	0.17284	0.46008	0.09189	0.15318
ความสะอาด	0.80000	0.48000	0.60000	0.53043	0.31282	0.39355	0.393550	0.25789	0.32666	0.57970	0.23712	0.33657
แพ้ง่าย/อ่อนโยน	0.55660	0.49167	0.52212	0.24752	0.21739	0.23148	0.17708	0.15455	0.16505	0.29157	0.24281	0.26497
ราคา	0.61538	0.72000	0.66359	0.43750	0.51579	0.47343	0.33645	0.40000	0.36548	0.36911	0.51005	0.42828
แบบที่ 2												
กลิ่น	0.82447	0.54386	0.65539	0.65574	0.42857	0.51836	0.60674	0.39273	0.47682	0.57225	0.36667	0.44695
ความสะอาด	0.71233	0.94545	0.81250	0.41176	0.56000	0.47457	0.28571	0.40000	0.33333	0.48333	0.89231	0.62702
แพ้ง่าย/อ่อนโยน	0.95918	0.32982	0.49086	0.81720	0.27143	0.40751	0.76136	0.24364	0.36915	0.97820	0.22276	0.36288
ราคา	0.72321	0.70435	0.71366	0.59817	0.58222	0.59009	0.53271	0.51818	0.52534	0.53223	0.50611	0.51884
แบบที่ 3												
กลิ่น	0.91304	0.84000	0.87500	0.85366	0.77778	0.81396	0.83333	0.75000	0.78947	0.85593	0.74815	0.79842
ความสะอาด	0.90769	0.69412	0.78667	0.86667	0.65000	0.74286	0.85455	0.62667	0.72308	0.92157	0.61842	0.74016
แพ้ง่าย/อ่อนโยน	0.79775	0.44375	0.57028	0.70238	0.38065	0.49373	0.64557	0.34000	0.44541	0.78186	0.29108	0.42422
ราคา	1.00000	0.48333	0.65168	0.89189	0.42128	0.57226	0.84906	0.39130	0.53571	0.98116	0.30145	0.46120

ตารางที่ 27 ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภทความสะอาด BIODERMA-SensibioH2O ที่ได้จากอัลกอริทึม TF-IDF

Group	ROUGE-1			ROUGE-2			ROUGE-3			ROUGE-SU4		
	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure
แบบที่ 1												
กลิ่น	0.36957	0.56667	0.44737	0.21951	0.36000	0.27273	0.11111	0.20000	0.14286	0.13136	0.31000	0.18453
ความสะอาด	0.40000	0.74286	0.52000	0.30000	0.60000	0.40000	0.23636	0.52000	0.32500	0.14706	0.55556	0.23256
แห้ง/อ่อนโยน	0.44944	0.59259	0.51118	0.30058	0.40000	0.34324	0.22024	0.29600	0.25256	0.23327	0.40318	0.29554
ราคา	0.35385	0.65714	0.46000	0.30435	0.46667	0.36842	0.16000	0.40000	0.22857	0.12000	0.51429	0.19459
แบบที่ 2												
กลิ่น	0.48404	0.70000	0.57233	0.32240	0.47200	0.38311	0.27528	0.40833	0.32886	0.22787	0.47657	0.30832
ความสะอาด	0.42857	0.84000	0.56757	0.37634	0.77778	0.50724	0.32955	0.72500	0.45313	0.10792	0.75185	0.18875
แห้ง/อ่อนโยน	0.44944	0.59259	0.51118	0.30058	0.40000	0.34324	0.22024	0.29600	0.25256	0.23327	0.40318	0.29554
ราคา	0.58929	0.94286	0.72528	0.49533	0.81538	0.61628	0.47059	0.80000	0.59259	0.34493	0.89615	0.49813
แบบที่ 3												
กลิ่น	0.48404	0.70000	0.57233	0.32240	0.47200	0.38311	0.27528	0.40833	0.32886	0.22787	0.47657	0.30832
ความสะอาด	0.63218	0.73333	0.67901	0.49704	0.57931	0.53503	0.44512	0.52143	0.48026	0.41160	0.56293	0.47552
แห้ง/อ่อนโยน	0.48235	0.53478	0.50721	0.32000	0.35556	0.33684	0.26531	0.29545	0.27957	0.24562	0.30648	0.27270
ราคา	0.61314	0.88421	0.72414	0.43939	0.64444	0.52252	0.38583	0.57647	0.46227	0.37908	0.78624	0.51153

ตารางที่ 28 ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภททำความสะอาด Cetaphil-GentleSkinCleanser อัลกอริทึม PageRank

Group	ROUGE-1			ROUGE-2			ROUGE-3			ROUGE-SU4		
	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure
แบบที่ 1												
กลิ่น	0.58278	0.65185	0.61538	0.35616	0.40000	0.37681	0.24823	0.28000	0.26316	0.32251	0.40584	0.35941
ความสะอาด	0.58427	0.65000	0.61538	0.47619	0.53333	0.50314	0.46835	0.52857	0.49664	0.37309	0.54222	0.44203
แพ้ง่าย/อ่อนโยน	0.40854	0.70526	0.51738	0.26415	0.46667	0.33735	0.18831	0.34118	0.24268	0.16344	0.49206	0.24538
ราคา	0.58621	0.97143	0.73119	0.50450	0.86154	0.63636	0.47170	0.83333	0.60241	0.27091	0.94038	0.42064
แบบที่ 2												
กลิ่น	0.48404	0.70000	0.57233	0.32240	0.47200	0.38311	0.27528	0.40833	0.32886	0.22787	0.47657	0.30832
ความสะอาด	0.42857	0.84000	0.56757	0.37634	0.77778	0.50724	0.32955	0.72500	0.45313	0.10792	0.75185	0.18875
แพ้ง่าย/อ่อนโยน	0.69333	0.67097	0.68197	0.52414	0.50667	0.51526	0.43571	0.42069	0.42807	0.51613	0.48485	0.50000
ราคา	0.58621	0.97143	0.73119	0.50450	0.86154	0.63636	0.47170	0.83333	0.60241	0.27091	0.94038	0.42064
แบบที่ 3												
กลิ่น	0.77033	0.29009	0.42147	0.6876	0.25701	0.37417	0.63971	0.23727	0.34615	0.42043	0.28290	0.33822
ความสะอาด	0.63218	0.73333	0.67901	0.49704	0.57931	0.53503	0.44512	0.52143	0.48026	0.41160	0.56293	0.47552
แพ้ง่าย/อ่อนโยน	0.68382	0.60000	0.63917	0.44275	0.38667	0.41281	0.33333	0.28966	0.30996	0.46927	0.36404	0.41001
ราคา	0.64286	0.41538	0.50467	0.47107	0.30000	0.36656	0.42241	0.26486	0.32558	0.39242	0.17279	0.23993

ตารางที่ 29 ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภทฟาคาความสะอาด Cetaphil-GentleSkinCleanser ได้จากอัลกอริทึม Hopfield

Group	ROUGE-1			ROUGE-2			ROUGE-3			ROUGE-SU4		
	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure
แบบที่ 1												
กลิ่น	0.59036	0.61250	0.60123	0.38462	0.40000	0.39216	0.38356	0.40000	0.39098	0.41247	0.46074	0.43527
ความสะอาด	0.78750	0.90000	0.84000	0.73333	0.84615	0.78571	0.68571	0.80000	0.73846	0.61461	0.82500	0.70443
แพ้ย่อมโยม	0.43286	0.86933	0.57795	0.35245	0.69881	0.46857	0.2942	0.54149	0.38132	0.39201	0.44167	0.41536
ราคา	0.56667	0.68000	0.61818	0.48000	0.60000	0.53333	0.4666	0.60000	0.52500	0.46939	0.68148	0.55589
แบบที่ 2												
กลิ่น	0.59036	0.61250	0.60123	0.38462	0.40000	0.39216	0.38356	0.40000	0.39098	0.41247	0.46074	0.43527
ความสะอาด	0.78750	0.90000	0.84000	0.73333	0.84615	0.78571	0.68571	0.80000	0.73846	0.61461	0.82500	0.70443
แพ้ย่อมโยม	0.70000	0.84000	0.56364	0.49091	0.60000	0.54000	0.48000	0.60000	0.53333	0.46667	0.60000	0.52500
ราคา	0.56667	0.68000	0.61818	0.48000	0.60000	0.53333	0.4666	0.60000	0.52500	0.46939	0.68148	0.55589
แบบที่ 3												
กลิ่น	0.73494	0.24400	0.36637	0.32051	0.10204	0.15480	0.28767	0.08750	0.13419	0.54111	0.06405	0.11454
ความสะอาด	1.00000	0.13333	0.93529	0.66667	0.08403	0.14925	0.60000	0.07119	0.12728	0.92693	0.01783	0.03499
แพ้ย่อมโยม	0.76667	0.70769	0.73600	0.43636	0.40000	0.41739	0.44000	0.40000	0.41905	0.57398	0.50000	0.53444
ราคา	0.56667	0.68000	0.61818	0.48000	0.60000	0.53333	0.4666	0.60000	0.52500	0.46939	0.68148	0.55589

ตารางที่ 30 ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภททำความสะอาด Cetaphil-GentleSkinCleanser ที่ได้จากอัลกอริทึม TF-IDF

Group	ROUGE-1			ROUGE-2			ROUGE-3			ROUGE-SU4		
	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure
แบบที่ 1												
กลิ่น	0.58278	0.65185	0.61538	0.35616	0.40000	0.37681	0.24823	0.28000	0.26316	0.32251	0.40584	0.35941
ความสะอาด	0.41237	0.61538	0.49382	0.23913	0.36667	0.28947	0.09195	0.14545	0.11267	0.18146	0.40444	0.25052
แพ้ง่าย/อ่อนโยน	0.40854	0.70526	0.51738	0.26415	0.46667	0.33735	0.18831	0.34118	0.24268	0.16344	0.49206	0.24538
ราคา	0.58621	0.97143	0.50119	0.50450	0.86154	0.63636	0.47170	0.83333	0.60241	0.27091	0.94038	0.42064
แบบที่ 2												
กลิ่น	0.58278	0.65185	0.61538	0.35616	0.40000	0.37681	0.24823	0.28000	0.26316	0.32251	0.40584	0.35941
ความสะอาด	0.42857	0.84000	0.56757	0.37634	0.77778	0.50724	0.32955	0.72500	0.45313	0.10792	0.75185	0.18875
แพ้ง่าย/อ่อนโยน	0.69333	0.67097	0.65197	0.52414	0.50667	0.51526	0.43571	0.42069	0.42807	0.51613	0.48485	0.50000
ราคา	0.58621	0.97143	0.50119	0.50450	0.86154	0.63636	0.47170	0.83333	0.60241	0.27091	0.94038	0.42064
แบบที่ 3												
กลิ่น	0.77033	0.29009	0.42147	0.6876	0.25701	0.37417	0.63971	0.23727	0.34615	0.42043	0.28290	0.33822
ความสะอาด	0.58427	0.65000	0.61538	0.47619	0.53333	0.50314	0.46835	0.52857	0.49664	0.37309	0.54222	0.44203
แพ้ง่าย/อ่อนโยน	0.68382	0.60000	0.63917	0.44275	0.38667	0.41281	0.33333	0.28966	0.30996	0.46927	0.36404	0.41001
ราคา	0.64286	0.41538	0.73467	0.47107	0.30000	0.36656	0.42241	0.26486	0.32558	0.39242	0.17279	0.23993

ตารางที่ 31 ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภททำความสะอาด DHC-DeepCleansingOil อัลกอริทึม PageRank

Group	ROUGE-1			ROUGE-2			ROUGE-3			ROUGE-SU4		
	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure
แบบที่ 1												
กลิ่น	0.50267	0.64828	0.56626	0.31319	0.40714	0.35404	0.20904	0.27407	0.23718	0.26384	0.43917	0.32964
ความสะอาด	0.48235	0.53478	0.50721	0.32000	0.35556	0.33684	0.2653	0.29545	0.27957	0.24562	0.30648	0.27270
แห้ง/อ่อนโยน	0.56818	0.68182	0.61983	0.37008	0.44762	0.40517	0.27049	0.33000	0.29730	0.31702	0.47302	0.37962
ราคา	0.46903	0.42400	0.44538	0.17756	0.14667	0.16064	0.01852	0.01667	0.01755	0.20267	0.16852	0.18402
แบบที่ 2												
กลิ่น	0.55738	0.85000	0.67327	0.43128	0.66182	0.52224	0.36691	0.31553	0.49057	0.27386	0.74194	0.40005
ความสะอาด	0.63218	0.73333	0.67901	0.49704	0.57931	0.53503	0.44512	0.52143	0.48026	0.41160	0.56293	0.47552
แห้ง/อ่อนโยน	0.52258	0.67500	0.58909	0.30000	0.39130	0.33962	0.23448	0.30909	0.26666	0.26790	0.44548	0.33459
ราคา	0.67273	0.74000	0.70476	0.34000	0.37778	0.35790	0.22222	0.25000	0.23329	0.40909	0.50000	0.45000
แบบที่ 3												
กลิ่น	0.65333	0.75385	0.70000	0.51429	0.60000	0.55385	0.47692	0.56364	0.51667	0.43333	0.52000	0.47273
ความสะอาด	0.72222	0.53529	0.61486	0.69421	0.50909	0.58741	0.68103	0.49375	0.57246	0.46628	0.43300	0.44902
แห้ง/อ่อนโยน	0.94382	0.42000	0.58131	0.66667	0.28718	0.40143	0.64557	0.26842	0.37918	0.92428	0.25934	0.40503
ราคา	1.00000	0.11489	0.40610	0.93878	0.09892	0.17898	0.90909	0.08696	0.15874	0.96875	0.01389	0.02739

ตารางที่ 32 ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภททำความสะอาด DHC-DeepCleansingOil ได้จากอัลกอริทึม Hopfield

Group	ROUGE-1				ROUGE-2				ROUGE-3				ROUGE-SU4			
	Recall	Precision	F-measure		Recall	Precision	F-measure		Recall	Precision	F-measure		Recall	Precision	F-measure	
แบบที่ 1																
กลิ่น	0.55738	0.85000	0.67327	0.43128	0.66182	0.52224	0.36691	0.56667	0.44542	0.27386	0.74194	0.40005				
ความสะอาด	0.44578	0.74000	0.55639	0.32051	0.55556	0.40650	0.23288	0.42500	0.30089	0.16667	0.5222	0.25269				
แพ้ง่าย/อ่อนโยน	0.49367	0.93134	0.64529	0.38278	0.72727	0.50157	0.31350	0.60000	0.41182	0.25608	0.49375	0.33725				
ราคา	0.41993	0.60513	0.49580	0.38744	0.49978	0.43650	0.36605	0.75889	0.49388	0.14193	0.48444	0.21954				
แบบที่ 2																
กลิ่น	0.62162	0.87619	0.72727	0.50515	0.71707	0.59274	0.45455	0.65000	0.53498	0.41993	0.60513	0.49580				
ความสะอาด	0.69277	0.79310	0.73955	0.53416	0.61429	0.57143	0.42308	0.48889	0.45361	0.42570	0.56636	0.48606				
แพ้ง่าย/อ่อนโยน	0.62222	0.65882	0.64000	0.42353	0.45000	0.43636	0.31250	0.33333	0.32258	0.39977	0.44868	0.42282				
ราคา	0.60811	0.85714	0.71146	0.36605	0.75889	0.4938	0.34962	0.57824	0.43576	0.37370	0.42206	0.39641				
แบบที่ 3																
กลิ่น	0.74667	0.62222	0.67879	0.65714	0.54118	0.59355	0.63077	0.51250	0.56552	0.65246	0.46824	0.54521				
ความสะอาด	0.69277	0.79310	0.73955	0.53416	0.61429	0.57143	0.42308	0.48889	0.45361	0.42570	0.56636	0.48606				
แพ้ง่าย/อ่อนโยน	0.52809	0.85455	0.65278	0.39286	0.39286	0.4925	0.36709	0.64444	0.46774	0.22019	0.77846	0.34328				
ราคา	0.79630	0.86000	0.82693	0.73469	0.80000	0.76596	0.72727	0.80000	0.76190	0.69063	0.81852	0.74916				

ตารางที่ 33 ค่าการวัดประสิทธิภาพของผลผลิตประเภททำความสะอาด DHC-DeepCleansingOil ที่ได้จากอัลกอริทึม TF-IDF

Group	ROUGE-1				ROUGE-2				ROUGE-3				ROUGE-SU4			
	Recall	Precision	F-measure	F-measure	Recall	Precision	F-measure	F-measure	Recall	Precision	F-measure	F-measure	Recall	Precision	F-measure	F-measure
แบบที่ 1																
กลิ่น	0.50267	0.64828	0.56626	0.31319	0.40714	0.35404	0.20904	0.27407	0.23718	0.26384	0.43917	0.32964				
ความสะอาด	0.48235	0.53478	0.50721	0.32000	0.35556	0.33684	0.2653	0.29545	0.27957	0.24562	0.30648	0.27270				
แพ้ง่าย/อ่อนโยน	0.52258	0.67500	0.58909	0.30000	0.39130	0.33962	0.23448	0.30909	0.26666	0.26790	0.44548	0.33459				
ราคา	0.46903	0.42400	0.44538	0.17756	0.14667	0.16064	0.01852	0.01667	0.01755	0.20267	0.16852	0.18402				
แบบที่ 2																
กลิ่น	0.50267	0.64828	0.56626	0.31319	0.40714	0.35404	0.20904	0.27407	0.23718	0.26384	0.43917	0.32964				
ความสะอาด	0.57639	0.57639	0.57439	0.39568	0.39286	0.39426	0.29104	0.28889	0.28996	0.35907	0.35576	0.35741				
แพ้ง่าย/อ่อนโยน	0.52258	0.67500	0.58909	0.30000	0.39130	0.33962	0.23448	0.30909	0.26666	0.26790	0.44548	0.33459				
ราคา	0.67273	0.74000	0.70476	0.34000	0.37778	0.35790	0.22222	0.25000	0.23529	0.40909	0.50000	0.45000				
แบบที่ 3																
กลิ่น	0.56818	0.68182	0.61983	0.37008	0.44762	0.40517	0.27049	0.33000	0.29730	0.31702	0.47302	0.37962				
ความสะอาด	0.72222	0.53529	0.61486	0.69421	0.50909	0.58741	0.68103	0.49375	0.57246	0.46628	0.43300	0.44902				
แพ้ง่าย/อ่อนโยน	0.94382	0.42000	0.58131	0.66667	0.28718	0.40143	0.64557	0.26842	0.37918	0.92428	0.25934	0.40503				
ราคา	0.67273	0.74000	0.70476	0.34000	0.37778	0.35790	0.22222	0.25000	0.23529	0.40909	0.50000	0.45000				

ตารางที่ 34 ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภทครีมรองพื้น CathyDoll-CCreamSPF50PA+++ อัลกอริทึม PageRank

Group	ROUGE-1			ROUGE-2			ROUGE-3			ROUGE-SU4		
	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure
แบบที่ 1												
เกลือง่าย	0.55932	0.47143	0.51163	0.37037	0.30769	0.33613	0.30612	0.25000	0.27523	0.31397	0.24908	0.27712
ปกปิด	0.69919	0.95000	0.76119	0.63769	0.90000	0.70968	0.62500	0.90000	0.70175	0.51236	0.80779	0.62702
คุมมัน	0.47897	0.75556	0.58621	0.33333	0.55000	0.41509	0.26230	0.45714	0.33334	0.21541	0.62273	0.32010
กันน้ำ	0.31894	0.99000	0.46808	0.27069	0.90000	0.40450	0.25000	0.90000	0.39095	0.09601	0.93333	0.15593
ดีทอน	0.42254	0.66667	0.27907	0.21368	0.60300	0.33212	0.23997	0.67702	0.35434	0.19123	0.60556	0.29067
แบบที่ 2												
เกลือง่าย	0.55932	0.47143	0.51163	0.37037	0.30769	0.33613	0.30612	0.25000	0.27523	0.31397	0.24908	0.27712
ปกปิด	0.69919	0.95000	0.76119	0.63769	0.90000	0.70968	0.62500	0.90000	0.70175	0.51236	0.80779	0.62702
คุมมัน	0.66667	0.32000	0.43243	0.43143	0.09972	0.16067	0.34082	0.21756	0.26559	0.23997	0.67702	0.55434
กันน้ำ	0.31894	0.99000	0.46808	0.27069	0.90000	0.40450	0.25000	0.90000	0.39095	0.09601	0.93333	0.15593
ดีทอน	0.44578	0.74000	0.55639	0.32051	0.53356	0.40650	0.23298	0.42500	0.30099	0.16667	0.5222	0.25269
แบบที่ 3												
เกลือง่าย	0.97950	0.99524	0.99679	0.77451	0.79800	0.78218	0.72165	0.73694	0.72917	0.76230	0.79479	0.77920
ปกปิด	0.71429	1.00000	0.93334	0.66667	1.00000	0.90000	0.60000	1.00000	0.75000	0.49611	1.00000	0.65420
คุมมัน	0.91937	0.95745	0.94713	0.97179	0.97143	0.91892	0.82353	0.93333	0.97300	0.92707	คุมมัน	0.91937
กันน้ำ	0.31894	0.99000	0.46808	0.27069	0.90000	0.40450	0.25000	0.90000	0.39095	0.09601	กันน้ำ	0.31894
ดีทอน	0.67424	0.32364	0.43735	0.37795	0.1777	0.24182	0.27049	0.12453	0.17054	0.44962	ดีทอน	0.67424

ตารางที่ 35 ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภทครีมรองพื้น CathyDoll-CCCreamSPF50PA+++ ได้จากอัลกอริทึม Hopfield

Group	ROUGE-1			ROUGE-2			ROUGE-3			ROUGE-SU4		
	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure
แบบที่ 1												
เกลือง่าย	0.61290	0.76000	0.67857	0.53846	0.70000	0.60869	0.42857	0.60000	0.50000	0.35294	0.60000	0.44444
ปกปิด	0.60526	0.76667	0.67647	0.49256	0.63636	0.55556	0.42424	0.56000	0.48276	0.30517	0.62957	0.41087
คุมมัน	0.47887	0.75556	0.58621	0.33333	0.53000	0.41509	0.26230	0.43714	0.33334	0.21541	0.62273	0.32010
กันน้ำ	0.57143	0.97778	0.72131	0.51389	0.92500	0.66072	0.46269	0.88571	0.60795	0.20753	0.97727	0.34436
ดีดทน	0.55357	0.77500	0.64383	0.35294	0.60000	0.44444	0.31213	0.70393	0.45238	0.29630	0.64000	0.40507
แบบที่ 2												
เกลือง่าย	0.61290	0.76000	0.67857	0.53846	0.70000	0.60869	0.42857	0.60000	0.50000	0.35294	0.60000	0.44444
ปกปิด	0.60526	0.76667	0.67647	0.49256	0.63636	0.55556	0.42424	0.56000	0.48276	0.30517	0.62957	0.41087
คุมมัน	0.58929	0.82500	0.68750	0.49533	0.70667	0.58242	0.47039	0.68571	0.58914	0.33105	0.71556	0.45267
กันน้ำ	0.57143	0.97778	0.72131	0.51389	0.92500	0.66072	0.46269	0.88571	0.60795	0.20753	0.97727	0.34436
ดีดทน	0.68919	0.85000	0.76119	0.63768	0.80000	0.70968	0.62500	0.80000	0.70175	0.51236	0.80779	0.62702
แบบที่ 3												
เกลือง่าย	0.87850	0.89524	0.88679	0.77451	0.73000	0.78218	0.72165	0.73684	0.72917	0.66304	0.67778	0.67033
ปกปิด	0.42373	1.00000	0.59524	0.37037	1.00000	0.54054	0.30612	1.00000	0.46875	0.17052	1.00000	0.29107
คุมมัน	0.90909	1.00000	0.95238	0.87179	0.97143	0.91892	0.82353	0.93333	0.87500	0.82707	1.00000	0.90535
กันน้ำ	0.57143	0.97778	0.72131	0.51389	0.92500	0.66072	0.46269	0.88571	0.60795	0.20753	0.97727	0.34436
ดีดทน	0.68919	0.85000	0.76119	0.63768	0.80000	0.70968	0.62500	0.80000	0.70175	0.51236	0.80779	0.62702

ตารางที่ 36 ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภทครีมรองพื้น CathyDoll-CCreamSPF50PA+++ ที่ได้จากอัลกอริทึม TF-IDF

Group	ROUGE-1			ROUGE-2			ROUGE-3			ROUGE-4		
	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure
แบบที่ 1												
เกลือง่าย	0.55992	0.47143	0.51163	0.37037	0.30769	0.33613	0.30612	0.25000	0.27523	0.31387	0.24808	0.27712
ปกปิด	0.40854	0.70526	0.51739	0.26415	0.46667	0.33735	0.18931	0.34119	0.24269	0.14094	0.26250	0.18341
คุมมัน	0.47887	0.73356	0.58621	0.33333	0.55800	0.41509	0.26230	0.45714	0.33334	0.21541	0.62273	0.32010
กันน้ำ	0.31894	0.89000	0.46908	0.27068	0.80000	0.40450	0.23000	0.80000	0.38095	0.08401	0.93333	0.15393
ดีทน	0.42254	0.66667	0.27907	0.21368	0.60500	0.35212	0.23997	0.67702	0.35434	0.19123	0.60556	0.29067
แบบที่ 2												
เกลือง่าย	0.55992	0.47143	0.51163	0.37037	0.30769	0.33613	0.30612	0.25000	0.27523	0.31387	0.24808	0.27712
ปกปิด	0.40854	0.70526	0.51739	0.26415	0.46667	0.33735	0.18931	0.34119	0.24269	0.14094	0.26250	0.18341
คุมมัน	0.66667	0.32800	0.43243	0.43145	0.09872	0.16067	0.34082	0.21736	0.26559	0.23997	0.67702	0.35434
กันน้ำ	0.31894	0.89000	0.46908	0.27068	0.80000	0.40450	0.23000	0.80000	0.38095	0.08401	0.93333	0.15393
ดีทน	0.42254	0.66667	0.27907	0.21368	0.60500	0.35212	0.23997	0.67702	0.35434	0.19123	0.60556	0.29067
แบบที่ 3												
เกลือง่าย	0.74251	0.67027	0.70454	0.61728	0.55556	0.58480	0.54593	0.44872	0.49257	0.37705	0.51111	0.43396
ปกปิด	0.71429	1.00000	0.83334	0.63667	1.00000	0.80000	0.60000	1.00000	0.75000	0.49611	1.00000	0.65420
คุมมัน	0.91837	0.95745	0.94713	0.87179	0.97143	0.91892	0.82353	0.93333	0.87500	0.82707	0.81081	0.90335
กันน้ำ	0.40984	0.53356	0.47170	0.16071	0.22500	0.18750	0.03922	0.05714	0.04651	0.17250	0.31364	0.22258
ดีทน	0.67424	0.92364	0.43735	0.37795	0.1777	0.24192	0.27049	0.12453	0.17054	0.44862	0.10669	0.17246

ตารางที่ 37 ค่าการวัดประสิทธิภาพของผลิตภัณฑ์กันตบประเภทครีมรองพื้น Revlon-ColorstayMakeup อัลกอริทึม PageRank

Group	ROUGE-1			ROUGE-2			ROUGE-3			ROUGE-SU4		
	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure
แบบที่ 1												
เกลี้ยงง่าย	0.50267	0.64928	0.56626	0.51319	0.40714	0.35404	0.20904	0.27407	0.23718	0.26394	0.43917	0.32964
ปกปิด	0.52809	0.85455	0.65278	0.39286	0.39286	0.49250	0.36709	0.64444	0.46774	0.22019	0.77846	0.34328
คุมมัน	0.22034	0.86667	0.35135	0.14815	0.80000	0.25000	0.08163	0.80000	0.14914	0.05115	0.80000	0.09615
กันน้ำ	0.60185	0.62397	0.61339	0.59748	0.62397	0.61101	0.59114	0.62397	0.60751	0.57925	0.62397	0.60075
ติดทน	0.75362	0.80000	0.77612	0.67188	0.71667	0.69355	0.62712	0.67273	0.64912	0.59259	0.64000	0.61339
แบบที่ 2												
เกลี้ยงง่าย	0.55263	0.70000	0.61765	0.40845	0.52727	0.46052	0.34849	0.46000	0.39655	0.30357	0.52987	0.38600
ปกปิด	0.54000	0.60000	0.56842	0.37895	0.42353	0.40000	0.26667	0.30000	0.28235	0.28863	0.38235	0.32994
คุมมัน	0.39474	0.93496	0.56604	0.33211	0.34146	0.52083	0.27273	0.90000	0.41861	0.14993	0.93496	0.34146
กันน้ำ	0.60185	0.62397	0.61339	0.59748	0.62397	0.61101	0.59114	0.62397	0.60751	0.57925	0.62397	0.60075
ติดทน	0.75362	0.80000	0.77612	0.67188	0.71667	0.69355	0.62712	0.67273	0.64912	0.59259	0.64000	0.61339
แบบที่ 3												
เกลี้ยงง่าย	0.75862	0.69474	0.72528	0.67073	0.61111	0.63953	0.64933	0.58824	0.61729	0.63956	0.55767	0.59581
ปกปิด	0.76364	0.84000	0.80000	0.62000	0.69899	0.65263	0.60000	0.67500	0.63529	0.61493	0.76296	0.68099
คุมมัน	0.33971	0.84000	0.48276	0.39070	0.80000	0.41558	0.23077	0.80000	0.35821	0.18800	0.82426	0.30617
กันน้ำ	0.87500	0.86154	0.86822	0.84746	0.83333	0.84034	0.83333	0.81918	0.82569	0.81633	0.76296	0.68099
ติดทน	0.91667	0.70000	0.79382	0.89608	0.66667	0.76087	0.87839	0.65000	0.74713	0.61493	0.77556	0.76195

ตารางที่ 38 ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภทเครื่องสำอาง Revlon-ColorstayMakeup ได้จากอัลกอริทึม Hopfield

Group	ROUGE-1			ROUGE-2			ROUGE-3			ROUGE-SU4		
	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure
แบบที่ 1												
เกลี้ยงง่าย	0.60556	0.66061	0.63189	0.43429	0.47500	0.45373	0.36471	0.40000	0.39154	0.39052	0.47071	0.42698
ปกปิด	0.23000	0.57500	0.32957	0.06316	0.17143	0.09231	0.10276	0.47754	0.16913	0.22000	0.55000	0.31429
คุมมัน	0.67797	0.80000	0.73395	0.61111	0.73333	0.66666	0.55102	0.67500	0.60674	0.46598	0.69779	0.55972
กันน้ำ	0.59659	0.67742	0.63444	0.45029	0.51333	0.47975	0.40361	0.46207	0.43087	0.36479	0.47636	0.41318
ดีทน	0.75362	0.90000	0.77612	0.67189	0.71667	0.69355	0.62712	0.67273	0.64912	0.50626	0.62989	0.56095
แบบที่ 2												
เกลี้ยงง่าย	0.64474	0.65333	0.64901	0.46479	0.47143	0.46909	0.37879	0.39462	0.39169	0.39939	0.43966	0.41200
ปกปิด	0.23000	0.57500	0.32957	0.06316	0.17143	0.09231	0.10276	0.47754	0.16913	0.22000	0.55000	0.31429
คุมมัน	0.67797	0.80000	0.73395	0.61111	0.73333	0.66666	0.55102	0.67500	0.60674	0.46598	0.69779	0.55972
กันน้ำ	0.75962	0.69474	0.72528	0.67073	0.61111	0.63953	0.64935	0.5892	0.61729	0.63956	0.55767	0.59581
ดีทน	0.75362	0.90000	0.77612	0.67189	0.71667	0.69355	0.62712	0.67273	0.64912	0.50626	0.62989	0.56095
แบบที่ 3												
เกลี้ยงง่าย	0.64474	0.65333	0.64901	0.46479	0.47143	0.46909	0.37879	0.39462	0.39168	0.39839	0.43966	0.41200
ปกปิด	0.33971	0.84000	0.49276	0.29070	0.80000	0.41558	0.23077	0.80000	0.35821	0.10959	0.80000	0.19277
คุมมัน	0.85513	0.75276	0.79923	0.64837	0.55478	0.59793	0.56674	0.49421	0.52223	0.75453	0.55789	0.64147
กันน้ำ	0.75862	0.69474	0.72528	0.67073	0.61111	0.63953	0.64935	0.5892	0.61729	0.63956	0.55767	0.59581
ดีทน	0.97500	0.96154	0.96822	0.84746	0.83333	0.84034	0.83333	0.91819	0.82569	0.76369	0.77556	0.76957

ตารางที่ 39 ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ที่ประเภทเครื่องสำอาง Revlon-ColorstayMakeup ที่ได้จากอัลกอริทึม TF-IDF

Group	ROUGE-1			ROUGE-2			ROUGE-3			ROUGE-SU4		
	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure
แบบที่ 1												
เสียง่าย	0.50267	0.64929	0.56626	0.31919	0.40714	0.35404	0.20904	0.27407	0.23719	0.26394	0.43917	0.32264
ปกปิด	0.52809	0.85455	0.65279	0.39296	0.39296	0.4925	0.36709	0.64444	0.46774	0.22019	0.77946	0.34329
คุมมัน	0.22034	0.96667	0.35135	0.14815	0.80000	0.25000	0.09163	0.80000	0.14914	0.05115	0.80000	0.09615
กันน้ำ	0.60556	0.66061	0.63189	0.43429	0.43429	0.45373	0.36471	0.40000	0.38154	0.39032	0.47071	0.42689
ดีดทน	0.75362	0.80000	0.77612	0.67189	0.71667	0.69355	0.62712	0.67273	0.64912	0.59259	0.64000	0.61539
แบบที่ 2												
เสียง่าย	0.50267	0.64929	0.56626	0.31919	0.40714	0.35404	0.20904	0.27407	0.23719	0.26394	0.43917	0.32264
ปกปิด	0.52809	0.85455	0.65279	0.39296	0.39296	0.4925	0.36709	0.64444	0.46774	0.22019	0.77946	0.34329
คุมมัน	0.39474	0.93496	0.56604	0.35211	0.34146	0.52083	0.27273	0.90000	0.41861	0.14993	0.93496	0.34146
กันน้ำ	0.60556	0.66061	0.63189	0.43429	0.43429	0.45373	0.36471	0.40000	0.38154	0.39032	0.47071	0.42689
ดีดทน	0.75362	0.80000	0.77612	0.67189	0.71667	0.69355	0.62712	0.67273	0.64912	0.59259	0.64000	0.61539
แบบที่ 3												
เสียง่าย	0.55263	0.70000	0.61765	0.40845	0.52727	0.46032	0.34848	0.46000	0.39635	0.30357	0.52997	0.39600
ปกปิด	0.52809	0.85455	0.65279	0.39296	0.39296	0.4925	0.36709	0.64444	0.46774	0.22019	0.77946	0.34329
คุมมัน	0.33971	0.84000	0.48276	0.29070	0.80000	0.41559	0.23077	0.80000	0.35821	0.18900	0.92426	0.30617
กันน้ำ	0.97500	0.86154	0.96822	0.84746	0.83333	0.94034	0.83333	0.81918	0.82569	0.81633	0.76296	0.68099
ดีดทน	0.74468	0.71795	0.73107	0.63389	0.61053	0.62199	0.56180	0.54054	0.48333	0.53321	0.50090	0.51635

ตารางที่ 40 ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภทเครื่องสำอาง Revlon-PhotoReadyFoundation อัลกอริทึม PageRank

Group	ROUGE-1			ROUGE-2			ROUGE-3			ROUGE-SU4		
	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure
แบบที่ 1												
เกลือง่าย	0.45652	0.60000	0.51852	0.31034	0.41338	0.35526	0.29268	0.40000	0.33803	0.22283	0.47308	0.30296
ปกปิด	0.41237	0.61538	0.49382	0.23913	0.36667	0.29947	0.09195	0.14545	0.11267	0.18146	0.40444	0.25052
คุมมัน	0.56338	0.89899	0.69966	0.51515	0.85000	0.64151	0.45902	0.80000	0.58334	0.31449	0.80909	0.45293
กันน้ำ	0.50435	0.64444	0.56585	0.38364	0.467059	0.41026	0.33333	0.43730	0.37838	0.22690	0.39294	0.28768
สีชมพู	0.77108	0.42667	0.54936	0.57692	0.31034	0.40358	0.52055	0.27143	0.35681	0.74468	0.27155	0.39798
แบบที่ 2												
เกลือง่าย	0.45652	0.84000	0.59155	0.39024	0.80000	0.52459	0.33333	0.90000	0.47059	0.19113	0.80000	0.30854
ปกปิด	0.41237	0.61538	0.49382	0.23913	0.36667	0.29947	0.09195	0.14545	0.11267	0.18146	0.40444	0.25052
คุมมัน	0.56338	0.89899	0.69966	0.51515	0.85000	0.64151	0.45902	0.80000	0.58334	0.31449	0.80909	0.45293
กันน้ำ	0.57143	0.97778	0.72131	0.51389	0.92500	0.66072	0.46269	0.88571	0.60785	0.20753	0.97727	0.34236
สีชมพู	0.77108	0.42667	0.54936	0.57692	0.31034	0.40358	0.52055	0.27143	0.35681	0.74468	0.27155	0.39798
แบบที่ 3												
เกลือง่าย	0.61290	0.76000	0.67857	0.53946	0.70000	0.60869	0.42857	0.60000	0.50000	0.35294	0.60000	0.44444
ปกปิด	0.60526	0.76667	0.67647	0.49296	0.65636	0.55556	0.24242	0.56000	0.48276	0.30517	0.62957	0.41087
คุมมัน	0.59929	0.82500	0.68750	0.49533	0.70667	0.58242	0.47059	0.88571	0.55814	0.44330	0.66134	0.33087
กันน้ำ	0.57143	0.97778	0.72131	0.51389	0.92500	0.66072	0.46269	0.88571	0.60785	0.20753	0.97727	0.34236
สีชมพู	0.89710	0.22449	0.35831	0.61404	0.14593	0.25569	0.55769	0.12340	0.20208	0.68388	0.03408	0.10023

ตารางที่ 4.1 ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ที่ประเภทเครื่องมือ Revlon-PhotoReadyFoundation ได้จากอัลกอริทึม Hopfield

Group	ROUGE-1			ROUGE-2			ROUGE-3			ROUGE-5U4		
	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure
แบบที่ 1												
เกลือง่าย	0.45652	0.60000	0.51952	0.31034	0.41538	0.35526	0.29269	0.40000	0.33903	0.22293	0.47309	0.30296
ปกปิด	0.56339	0.99899	0.68966	0.51515	0.95000	0.64151	0.45902	0.90000	0.59334	0.39286	0.73333	0.51165
คุมน้ำมัน	0.44378	0.74000	0.55639	0.30051	0.55556	0.40630	0.23299	0.42500	0.30899	0.16667	0.52222	0.25269
กันน้ำ	0.64516	0.99999	0.74766	0.59649	0.95000	0.70103	0.55769	0.92957	0.66666	0.31064	0.90000	0.62339
ติดทน	0.99710	0.22449	0.53931	0.61404	0.14593	0.23569	0.55769	0.12340	0.20209	0.68389	0.05409	0.10923
แบบที่ 2												
เกลือง่าย	0.45652	0.60000	0.51952	0.31034	0.41539	0.35526	0.29268	0.40000	0.33903	0.22293	0.47309	0.30296
ปกปิด	0.60195	0.62597	0.61339	0.59749	0.62597	0.61101	0.59114	0.62597	0.60751	0.59117	0.62597	0.60195
คุมน้ำมัน	0.44379	0.74000	0.55639	0.30051	0.55556	0.40650	0.23299	0.42500	0.30899	0.16667	0.52222	0.25269
กันน้ำ	0.64516	0.99999	0.74766	0.59649	0.95000	0.70103	0.55769	0.92957	0.66666	0.31064	0.90000	0.62339
ติดทน	0.79330	0.96000	0.82693	0.73469	0.90000	0.76596	0.72727	0.90000	0.76190	0.69063	0.91952	0.74916
แบบที่ 3												
เกลือง่าย	0.74467	0.62222	0.67879	0.65714	0.54118	0.59555	0.63077	0.51250	0.56552	0.65246	0.46924	0.54521
ปกปิด	0.60195	0.62597	0.61339	0.59749	0.62597	0.61101	0.59114	0.62597	0.60751	0.59117	0.62597	0.60195
คุมน้ำมัน	0.61739	0.41765	0.49925	0.39182	0.25455	0.30546	0.33239	0.23125	0.27924	0.51902	0.25724	0.34399
กันน้ำ	0.69271	0.76000	0.72490	0.59802	0.59892	0.53221	0.42308	0.46667	0.44381	0.45772	0.56789	0.39911
ติดทน	0.94737	0.94390	0.94363	0.91379	0.91043	0.91211	0.99715	0.99593	0.99549	0.99090	0.99453	0.99765

ตารางที่ 42 ค่าการวัดประสิทธิภาพของผลิตภัณฑ์ประเภทครีมรองพื้น Revlon-PhotoReadyFoundation ที่ได้จากอัลกอริทึม TF-IDF

Group	ROUGE-1			ROUGE-2			ROUGE-3			ROUGE-SU4		
	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure	Recall	Precision	F-measure
แบบที่ 1												
เฉลี่ยง่าย	0.45652	0.60000	0.51852	0.51034	0.41538	0.55526	0.29268	0.40000	0.33803	0.22283	0.47308	0.30296
ปกติ	0.44944	0.59239	0.51118	0.30058	0.40000	0.34324	0.22024	0.29600	0.25256	0.23327	0.44554	0.29554
คุณมัน	0.56338	0.88889	0.69966	0.51515	0.95000	0.64151	0.45902	0.80000	0.58334	0.31449	0.80909	0.45293
กันน้ำ	0.44275	0.38667	0.41281	0.36018	0.4197	0.38769	0.24793	0.21429	0.22989	0.45333	0.34882	0.39427
เฉลี่ยง่าย	0.77108	0.84000	0.59155	0.39024	0.80000	0.52459	0.33333	0.80000	0.47058	0.19113	0.80000	0.30854
แบบที่ 2												
เฉลี่ยง่าย	0.45652	0.60000	0.51852	0.51034	0.41538	0.55526	0.29268	0.40000	0.33803	0.22283	0.47308	0.30296
ปกติ	0.44944	0.59239	0.51118	0.30058	0.40000	0.34324	0.22024	0.29600	0.25256	0.23327	0.44554	0.29554
คุณมัน	0.56338	0.88889	0.69966	0.51515	0.95000	0.64151	0.45902	0.80000	0.58334	0.31449	0.80909	0.45293
กันน้ำ	0.50435	0.64444	0.56585	0.36364	0.47059	0.41026	0.33333	0.43750	0.37538	0.22680	0.39294	0.28788
ดีดทน	0.99710	0.22449	0.35831	0.61404	0.14583	0.23569	0.55769	0.12340	0.20208	0.68388	0.03408	0.10023
แบบที่ 3												
เฉลี่ยง่าย	0.69333	0.67097	0.69197	0.52414	0.50667	0.51526	0.43371	0.42069	0.42807	0.51613	0.48485	0.50000
ปกติ	0.60526	0.76667	0.67647	0.49296	0.63636	0.55556	0.42424	0.56000	0.48376	0.30517	0.62857	0.41087
คุณมัน	0.56338	0.88889	0.69966	0.51515	0.95000	0.64151	0.45902	0.80000	0.58334	0.31449	0.80909	0.45293
กันน้ำ	0.60335	0.80000	0.68790	0.48851	0.65385	0.53921	0.40828	0.55200	0.46939	0.56022	0.63501	0.45968
ดีดทน	0.99710	0.22449	0.35831	0.61404	0.14583	0.23569	0.55769	0.12340	0.20208	0.68388	0.03408	0.10023

ภาคผนวก ข คลังข้อมูล

สร้างคลังข้อมูลระบุทิศทางของความคิดเห็น (Corpus Positive word, Corpus Negative word)

ตารางที่ 43 รายการคำแสดงความคิดเห็นเชิงบวกจากงานวิจัยของ (Trithep Thumrongluck, 2010)

กลุ่มบริษัชอันทาแก้ม	กลุ่มโลชั่นทำความสะอาดหน้า	กลุ่มครีมรองพื้น
สวย	ถูก	ถูก
ดีดทน	หอม	บางเบา
ถูก	สะอาด	หอม
นาน	สวย	เนียน
ละเอียด	ง่าย	เรียบเนียน
มาก	ดี	สวย
ดี	เบาบาง	ดี
ง่าย	เนียนนุ่ม	เหลว
สดใส	เหมาะสม	นาน
หอม	มาก	รู้จัก
เหมาะสม	หมดจด	ง่าย
เนียน	น่าเชื่อถือ	สมมุเสมอ
สะดวก	ทนทาน	มาก
ทนทาน	เร็ว	น่าเชื่อถือ
เข้ม	อ่อนโยน	สว่าง
อ่อน	สะดวก	เหมาะสม
น่ารัก	รู้จัก	สมเหตุสมผล
อ่อนโยน	ละเอียด	ละเอียด
เยอะ	คุ้มค่า	เข้ม
อ่อนนุ่ม	สิ้น	ทนทาน
รู้จัก	สวย	เร็ว

ชัดเจน	เหลว	น้อย
หรุหระ	ไว	อ่อนโยน
สมเหตุสมผล	น่าใช้	คุ้นเคย
ทันสมัย	สมเหตุสมผล	อ่อน
น่าเชื่อถือ	ใหญ่	
น่าใช้	เยอะ	
	ช้า	
	ทันสมัย	
	กระต๊าด	
	สูง	
	ใส	
	หรุหระ	
	น้อย	
	เหนียว	
	ขึ้น	

ตารางที่ 44 รายการค่าแสดงความคิดเห็นเชิงบวกจากงานวิจัยนี้

กลุ่มบรรชอนทาแก้ม	กลุ่มโลชันทำความสะอาดหน้า	กลุ่มครีมรองพื้น
ใช้ดี	เกลี้ยง	ติดทน
มีสีสัน	ไม่แพ้	เรียบ
สีสวย	ใช้ง่าย	ปกปิด
เลิศ	นุ่ม	ผองใส
เยี่ยม	นุ่มนวล	เยอะ
	ไม่แพ้	เนียนนุ่ม
	คุ้มค่า	คุ้มค่า
	สดชื่น	น่าใช้
	ประทับใจ	

ตารางที่ 45 รายการคำแสดงความคิดเห็นเชิงลบจากงานวิจัยของ (Trithep Thumrongluck, 2010)

กลุ่มบริษัชอันทากัม	กลุ่มโลชันทำความสะอาดหน้า	กลุ่มครีมรองพื้น
แพง	แพง	แพง
ง่าย	เหม็น	เหนียว
น้อย	เหนียว	เหนอะหนะ
หยาบ	ฉุน	น้อย
ยาก	น้อย	เหม็น
เร็ว	ยาก	ยาก
เปลือง	ง่าย	ช้า
ยาก	แพ้	ฉุน
เร็ว	หยาบ	ง่าย
เปลือง	เหนอะหนะ	ใหญ่
เหม็น	สกปรก	เปลือง
น้ำเกลียด	เปลือง	แฉะ
ธรรมดา	ช้า	ลำบาก
เซย	แฉะ	หยาบ
เข้มน	ระคายเคือง	หนา
จืด	ธรรมดา	เหลว
น้ำเกลียด	น้ำเกลียด	เข้มน
แฉะ	ซีแห้ง	ซีแห้ง
แข็ง	ตกค้าง	ต่ำ
อ่อน	หนัก	คราบ
ดูฉาด	ต่ำ	หนัก
หม่น	คราบ	เลอะเทอะ
ซีด	ชุ่มฉะ	เทอะทะ
ซีแห้ง	เซย	เบาบาง
ต่ำ	หนืด	ลื่น
มาก	แฉะ	จืด

ใหญ่	เยอะ	เซย
กลุ่มบริษัทอันทาแก้ม	กลุ่มโลชั่นทำความสะอาดหน้า	กลุ่มครีมรองพื้น
แข่งกระด้าง	เทอะทะ	น่าเกลียด
บาง	เหลว	หนืด
	แสบ	ชุ่มมัว
	คัน	เส็ก
	เส็ก	ขึ้น
	ใหญ่	หมอง
	เฉยๆ	มาก
	เร็ว	เร็ว

ตารางที่ 46 รายการคำแสดงความคิดเห็นเชิงลบจากงานวิจัยนี้

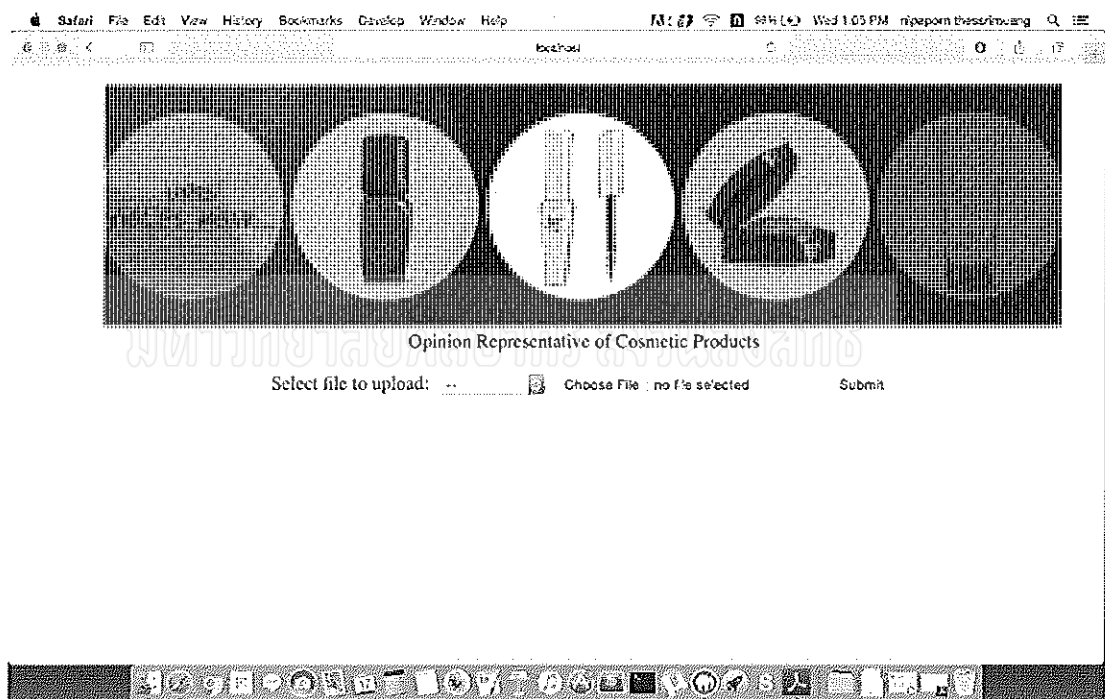
กลุ่มบริษัทอันทาแก้ม	กลุ่มโลชั่นทำความสะอาดหน้า	กลุ่มครีมรองพื้น
น้อย	ไม่สะอาด	ระคายเคือง
แสบ	ติด	หม่น
	ค้ำ	เหนอะ
	คราบ	ไม่ติด
		ก้อน
		ขุย
		คราบ
		แพ้

ภาคผนวก ค คู่มือการใช้งานโปรแกรม (User' s Document)

ภาคผนวกนี้ จะกล่าวถึงการใช้งานโปรแกรมตามลำดับขั้นตอนอย่างละเอียด เพื่อให้ผู้ใช้สามารถใช้งานได้อย่างถูกต้องและเข้าใจการทำงานของโปรแกรมมากขึ้น

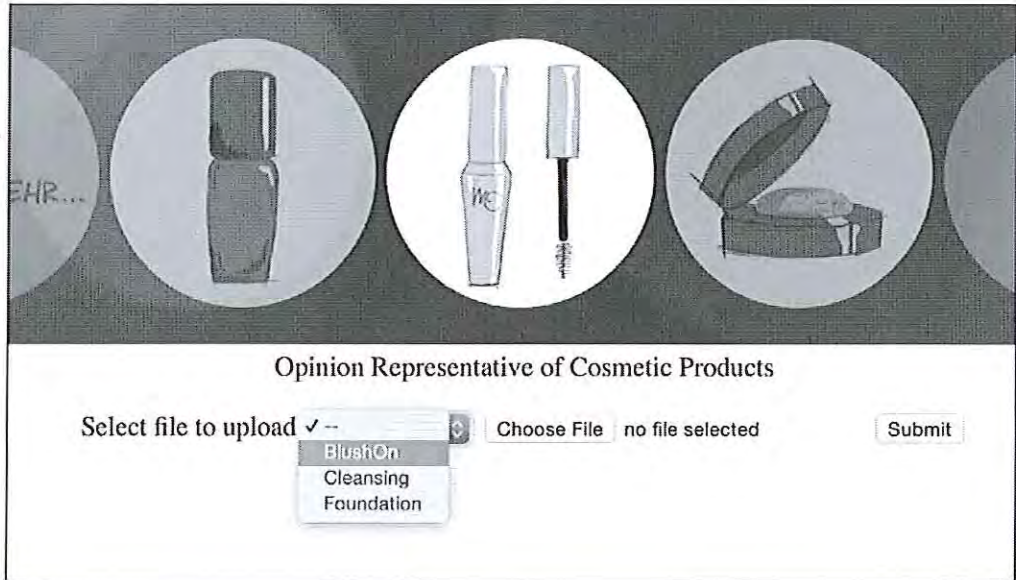
การใช้งานระบบ

แสดงหน้าแรกของระบบความคิดเห็นที่เป็นตัวแทนของผลิตภัณฑ์



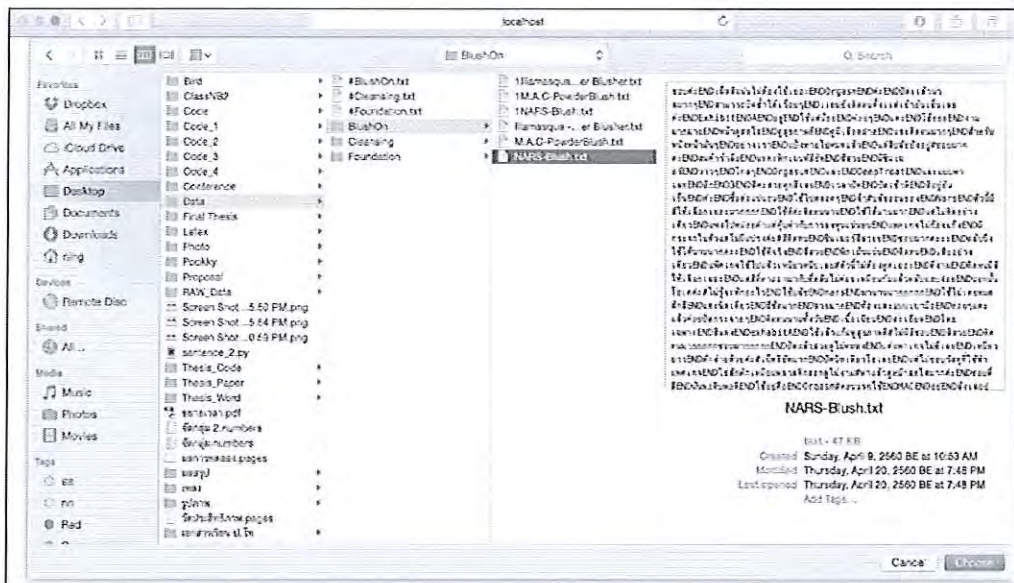
รูปที่ 40 หน้าแรกของระบบ

เลือกผลิตภัณฑ์เครื่องสำอางที่ผู้ใช้งานต้องการ



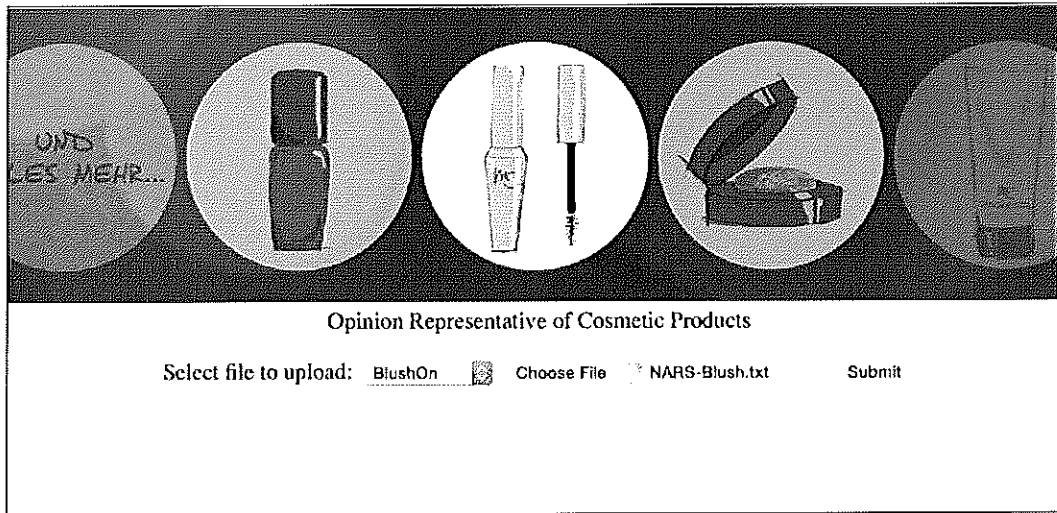
รูปที่ 41 เลือกผลิตภัณฑ์เครื่องสำอางที่ผู้ใช้งาน

กดปุ่มเลือกไฟล์ (Choose File) เพื่อเลือกชื่อของผลิตภัณฑ์เครื่องสำอางที่ผู้ใช้งาน



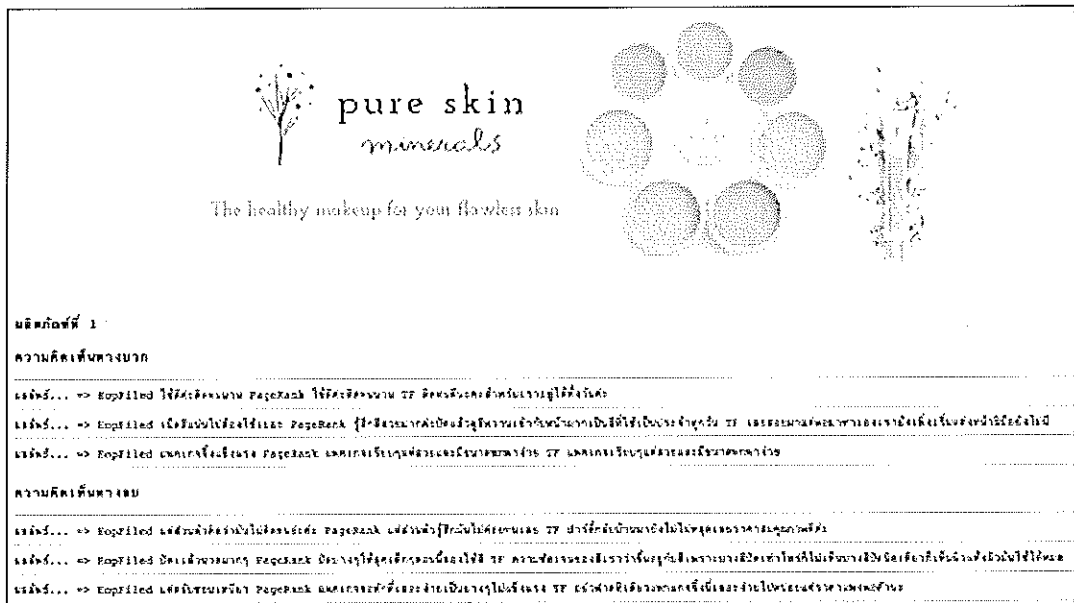
รูปที่ 42 เลือกชื่อของผลิตภัณฑ์เครื่องสำอางที่ผู้ใช้งาน

กดปุ่มตกลง (Submit) เพื่อให้ระบบทำการเลือกประโยคความคิดเห็นที่เป็นตัวแทนของ
ผลิตภัณฑ์



มหาวิทยาลัยศิลปากร
รูปที่ 43 กดปุ่มตกลง (Submit)

ระบบแสดงความคิดเห็นที่เป็นตัวแทนของผลิตภัณฑ์ที่ผู้ใช้ต้องการ



รูปที่ 44 ระบบแสดงความคิดเห็นที่เป็นตัวแทนของผลิตภัณฑ์ที่ผู้ใช้ต้องการ

ประวัติผู้วิจัย

ชื่อ – สกุล

นิภาพร เทศศรีเมือง

ที่อยู่

8 หมู่ 6 ตำบล หุ่นหลวง อำเภอ โพนพิสัย จังหวัด หนองคาย 43120

ประวัติการศึกษา

พ.ศ. 2557

สำเร็จการศึกษาปริญญาวิทยาศาสตรบัณฑิต

สาขาวิชาเทคโนโลยีสารสนเทศ

มหาวิทยาลัยศิลปากร

อาจารย์ที่ปรึกษาวิทยานิพนธ์ อาจารย์ ดร. อรวรรณ เชาวลิต

GPA : 2.73

พ.ศ. 2557

ศึกษาต่อระดับปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาเทคโนโลยีสารสนเทศ

มหาวิทยาลัยศิลปากร

อาจารย์ที่ปรึกษาวิทยานิพนธ์ อาจารย์ ดร. อรวรรณ เชาวลิต

GPA : 3.45

มหาวิทยาลัยศิลปากร สมานลิขสิทธิ์