



กรอบพัฒนาการจัดกลุ่มข้อความทวิตเตอร์ที่กล่าวถึงมหาวิทยาลัยไทย :
กรณีศึกษามหาวิทยาลัยธรรมศาสตร์

โดย

มนต์ชัย ป่านขาว

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
วิทยาศาสตรมหาบัณฑิต (วิทยาการคอมพิวเตอร์)
สาขาวิชาวิทยาการคอมพิวเตอร์
คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์
ปีการศึกษา 2565

A FRAMEWORK FOR CLUSTERING TWITTER MESSAGES REFERRING
TO THAI UNIVERSITIES : A CASE STUDY OF THAMMASAT
UNIVERSITY

BY

MONCHAI PARNKOW

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF MASTER OF SCIENCE (COMPUTER SCIENCE)
DEPARTMENT OF COMPUTER SCIENCE
FACULTY OF SCIENCE AND TECHNOLOGY
THAMMASAT UNIVERSITY
ACADEMIC YEAR 2022

มหาวิทยาลัยธรรมศาสตร์
คณะวิทยาศาสตร์และเทคโนโลยี

วิทยานิพนธ์

ของ

มนต์ชัย ป่านขาว

เรื่อง

กรอบพัฒนาการจัดกลุ่มข้อความทวิตเตอร์ที่กล่าวถึงมหาวิทยาลัยไทย : กรณีศึกษา
มหาวิทยาลัยธรรมศาสตร์

ได้รับการตรวจสอบและอนุมัติ ให้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร
วิทยาศาสตรมหาบัณฑิต (วิทยาการคอมพิวเตอร์)

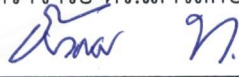
เมื่อ วันที่ 13 มิถุนายน พ.ศ. 2566

ประธานกรรมการสอบวิทยานิพนธ์




(ผู้ช่วยศาสตราจารย์ ดร.เสาวลักษณ์ วรรณานา)

กรรมการและอาจารย์ที่ปรึกษาวิทยานิพนธ์



(ผู้ช่วยศาสตราจารย์ ดร.ธนาธร ทยานานทอง)

กรรมการสอบวิทยานิพนธ์



(ผู้ช่วยศาสตราจารย์ ดร.ปกป้อง ส่องเมือง)

กรรมการสอบวิทยานิพนธ์



(รองศาสตราจารย์ ดร.ทศนัย ชุ่มวัฒนะ)

คณบดี



(รองศาสตราจารย์ ดร.สุเพชร จิระจรกุล)

| | |
|-----------------------------|---|
| หัวข้อวิทยานิพนธ์ | กรอบพัฒนาการจัดกลุ่มข้อความทวิตเตอร์ที่กล่าวถึง |
| ชื่อผู้เขียน | มหาวิทยาลัยไทย : กรณีศึกษามหาวิทยาลัยธรรมศาสตร์ |
| ชื่อปริญญา | มนต์ชัย ป่านขาว |
| สาขาวิชา/คณะ/มหาวิทยาลัย | วิทยาศาสตร์มหาบัณฑิต (วิทยาการคอมพิวเตอร์) |
| | สาขาวิชาวิทยาการคอมพิวเตอร์ |
| | คณะวิทยาศาสตร์และเทคโนโลยี |
| | มหาวิทยาลัยธรรมศาสตร์ |
| อาจารย์ที่ปรึกษาวิทยานิพนธ์ | ผู้ช่วยศาสตราจารย์ ดร.ธนาธร ทะนานทอง |
| ปีการศึกษา | 2565 |

บทคัดย่อ

งานวิจัยนี้นำเสนอกรอบพัฒนาการจัดกลุ่มข้อความภาษาไทยบนสื่อสังคมออนไลน์ ทวิตเตอร์ที่มีการกล่าวถึงมหาวิทยาลัยธรรมศาสตร์ โดยผู้วิจัยได้ใช้อัลกอริทึม K-Means สำหรับการ จัดกลุ่มข้อความและทำการเปรียบเทียบวิธีการคิดค่าความคล้ายคลึง (Similarity Measures) ที่ แตกต่างกันออกไป ได้แก่ Euclidean Distance, Cosine Similarity, Jaccard Similarity และ Pearson Similarity สำหรับการปรับปรุงประสิทธิภาพของการจัดกลุ่มข้อความ ในงานวิจัยนี้ ผู้วิจัย ทำการทดลองโดยใช้ชื่อเฉพาะ (Named Entities) และคำที่มีความหมายคล้ายคลึงกัน (Synonyms) จากผลการทดลองพบว่าการจัดกลุ่มข้อความโดยอัลกอริทึม K-Means ที่มีวิธีการคิดค่าความ คล้ายคลึงแบบ Cosine Similarity ให้ผลลัพธ์สูงที่สุด และพบว่าการใช้ชื่อเฉพาะ (Named Entity) และคำที่มีความหมายคล้ายคลึงกัน (Synonym) สามารถช่วยปรับปรุง ประสิทธิภาพของการสรุปหัวข้อกลุ่มของข้อความ

คำสำคัญ: การจัดกลุ่มข้อความภาษาไทย ชื่อเฉพาะ คำที่คล้ายคลึงกัน วิธีวัดความคล้าย

| | |
|-------------------------------|---|
| Thesis Title | A FRAMEWORK FOR CLUSTERING TWITTER MESSAGES REFERRING TO THAI UNIVERSITIES : A CASE STUDY OF THAMMASAT UNIVERSITY |
| Author | Monchai Parnkow |
| Degree | Master of Science (Computer Science) |
| Department/Faculty/University | Computer Science Faculty of Science and Technology Thammasat University |
| Thesis Advisor | Assistant Professor Tanatorn Tanantong, Ph.D. |
| Academic Year | 2022 |

ABSTRACT

This research presents a clustering framework of Thai messages on Twitter related to Thammasat University. The researcher used the K-Means algorithm for clustering the messages and compared different similarity measures, i.e., Euclidean Distance, Cosine Similarity, Jaccard Similarity, and Pearson Similarity. For improving the efficiency of message clustering, in this study, the researchers did experiments using Named Entities and Synonyms. From the obtained experimental results, text clustering using K-Means with Cosine Similarity yielded the best performance. In addition, the adoption of name entities and synonyms can help to improve the performance of concluding topics of the message group.

Keywords: Thai Text Clustering, Named Entity, Synonyms, Similarity Measures

กิตติกรรมประกาศ

วิทยานิพนธ์เล่มนี้ จะไม่สามารถสำเร็จได้ ถ้าไม่ได้รับความเมตตาช่วยเหลือจาก ผู้ช่วยศาสตราจารย์ ดร.ธนาธร ทะนานทอง ที่คอยสั่งสอนและให้คำแนะนำเสมอมา รวมถึงหน่วยวิจัยด้านนวัตกรรมข้อมูลและปัญญาประดิษฐ์แห่งมหาวิทยาลัยธรรมศาสตร์ ที่คอยรับฟังและให้แนวทางสำหรับดำเนินงานวิจัยนี้ ตลอดจนคณะกรรมการสอบวิทยานิพนธ์อันได้แก่ 1.ผู้ช่วยศาสตราจารย์ ดร.เสาวลักษณ์ วรรณภา ประธานกรรมการสอบวิทยานิพนธ์ 2. ผู้ช่วยศาสตราจารย์ ดร.ธนาธร ทะนานทอง กรรมการและอาจารย์ที่ปรึกษาวิทยานิพนธ์ 3.ผู้ช่วยศาสตราจารย์ ดร.ปกป้อง ส่องเมือง กรรมการสอบวิทยานิพนธ์ 4.รองศาสตราจารย์ ดร.ทศนีย์ ชุ่มวัฒนะ กรรมการสอบวิทยานิพนธ์ ที่สละเวลามาเป็นกรรมคุมสอบและเสนอข้อเสนอแนะอันทรงคุณค่าต่องานวิจัยนี้

ขอขอบคุณครอบครัวที่คอยสนับสนุนให้กำลังใจ อีกทั้งปัจจัยทางด้านทุนทรัพย์ที่ทำให้วิจัยนี้สามารถเกิดขึ้นได้

มนต์ชัย ป่านขาว

(4)

สารบัญ

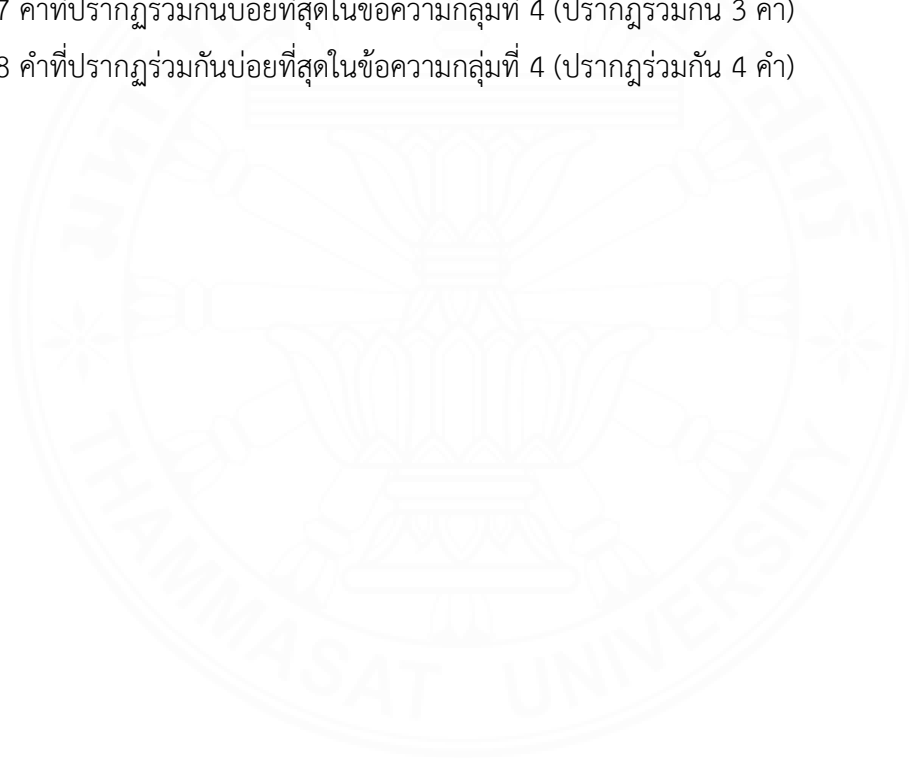
| | |
|---|-----|
| บทคัดย่อภาษาไทย | (1) |
| บทคัดย่อภาษาอังกฤษ | (2) |
| กิตติกรรมประกาศ | (3) |
| สารบัญ | (4) |
| สารบัญตาราง | (6) |
| สารบัญภาพ | (8) |
| บทที่ 1 | 1 |
| 1.1 วัตถุประสงค์ของงานวิจัย | 2 |
| 1.2 ขอบเขตของงานวิจัย | 3 |
| 1.3 ประโยชน์ที่คาดว่าจะได้รับ | 3 |
| 1.4 นิยามศัพท์ที่เกี่ยวข้อง | 3 |
| บทที่ 2 | 5 |
| 2.1 การจัดกลุ่มข้อความ | 5 |
| 2.2 การเตรียมข้อมูล | 5 |
| 2.3 การแปลงชื่อเฉพาะเป็นกลุ่มของชื่อเฉพาะ | 8 |
| 2.4 การแปลงข้อความที่ความหมายคล้ายคลึงกัน | 9 |
| 2.5 การสกัดคุณลักษณะ | 9 |
| 2.6 การจัดกลุ่มข้อมูลแบบเค-มีน | 10 |
| 2.7 Sum Squared Error | 12 |
| 2.8 Elbow method | 12 |

| | |
|--|-----|
| | (5) |
| 2.9 งานวิจัยที่เกี่ยวข้องกับการจัดกลุ่มข้อความ | 13 |
| บทที่ 3 | 16 |
| 3.1 การเก็บรวบรวมข้อมูล | 17 |
| 3.2 การสร้างพจนานุกรมการเตรียมและจัดการข้อมูล | 18 |
| 3.2 การจัดกลุ่มข้อความและสรุปหัวเรื่องข้อความ | 22 |
| บทที่ 4 | 26 |
| 4.1 ผลลัพธ์การเปรียบเทียบค่าความคล้ายคลึงที่เหมาะสม | 26 |
| 4.2 ผลลัพธ์การกำหนดจำนวนกลุ่ม(k)ที่เหมาะสมสำหรับการจัดกลุ่มข้อความ | 28 |
| 4.3 ผลลัพธ์การจัดกลุ่มข้อความที่กล่าวถึงมหาวิทยาลัยธรรมศาสตร์ | 31 |
| 4.4 ผลลัพธ์การสรุปหัวข้อของกลุ่มข้อความ | 53 |
| บทที่ 5 | 87 |
| 5.1 สรุปผล | 87 |
| 5.2 ข้อเสนอแนะสำหรับงานวิจัยครั้งต่อไป | 88 |
| รายการอ้างอิง | 89 |
| ประวัติผู้เขียน | 97 |

สารบัญตาราง

| ตารางที่ | หน้า |
|--|------|
| 2.1 ตัวอย่างการแปลงข้อความที่ความหมายคล้ายคลึงกัน | 9 |
| 3.1 ตัวอย่างคำในพจนานุกรมชื่อเฉพาะที่สร้างขึ้นจากคลังข้อมูลทวิตเตอร์ | 18 |
| 3.2 ตัวอย่างคำในพจนานุกรมคำพ้องความหมายที่สร้างขึ้นจากคลังข้อมูลทวิตเตอร์ | 19 |
| 3.3 ตัวอย่างผลลัพธ์ของการสกัดชื่อเฉพาะ | 19 |
| 3.4 ตัวอย่างการแปลงคำที่มีความหมายคล้ายคลึงกัน | 20 |
| 3.5 ตัวอย่างการสกัดคุณลักษณะโดยใช้ TFIDF | 21 |
| 3.6 ตัวอย่างคำที่มีความถี่สูงสุด 100 อันดับแรกของข้อความในกลุ่มที่ 1 | 24 |
| 4.1 ค่า Sum of Square Error เปรียบเทียบตามค่าความคล้ายคลึง สำหรับการจัดกลุ่มข้อความ | 26 |
| 4.2 ค่า Sum of Square Error เปรียบเทียบตามค่าความคล้ายคลึง สำหรับการจัดกลุ่มข้อความที่ใช้การแปลงกลุ่มของชื่อเฉพาะ | 27 |
| 4.3 Sum of Square Error และผลต่างของ Sum of Square Error สำหรับการจัดกลุ่มข้อความ | 29 |
| 4.4 Sum of Square Error และผลต่างของ Sum of Square Error สำหรับการจัดกลุ่มข้อความที่ใช้การแปลงกลุ่มของชื่อเฉพาะ | 30 |
| 4.5 ตัวอย่างข้อความกลุ่มที่ 1 ของการจัดกลุ่มข้อความที่ไม่ได้ใช้การแปลงกลุ่มของชื่อเฉพาะ | 33 |
| 4.6 ตัวอย่างข้อความกลุ่มที่ 2 ของการจัดกลุ่มข้อความที่ไม่ได้ใช้การแปลงกลุ่มของชื่อเฉพาะ | 36 |
| 4.7 ตัวอย่างข้อความกลุ่มที่ 3 ของการจัดกลุ่มข้อความที่ไม่ได้ใช้การแปลงกลุ่มของชื่อเฉพาะ | 41 |
| 4.8 ตัวอย่างข้อความกลุ่มที่ 4 ของการจัดกลุ่มข้อความที่ไม่ได้ใช้การแปลงกลุ่มของชื่อเฉพาะ | 44 |
| 4.9 ตัวอย่างข้อความกลุ่มที่ 1 ของการจัดกลุ่มข้อความที่ใช้การแปลงกลุ่มของชื่อเฉพาะ | 49 |
| 4.10 ตัวอย่างข้อความกลุ่มที่ 2 ของการจัดกลุ่มข้อความที่ใช้การแปลงกลุ่มของชื่อเฉพาะ | 51 |
| 4.11 ตัวอย่างข้อความกลุ่มที่ 3 ของการจัดกลุ่มข้อความที่ใช้การแปลงกลุ่มของชื่อเฉพาะ | 52 |
| 4.12 ตัวอย่างข้อความกลุ่มที่ 4 ของการจัดกลุ่มข้อความที่ใช้การแปลงกลุ่มของชื่อเฉพาะ | 55 |
| 4.13 คำที่มีความถี่สูงสุด 100 อันดับแรกในข้อความกลุ่มที่ 1 | 57 |
| 4.14 คำที่ปรากฏร่วมกันบ่อยที่สุดในข้อความกลุ่มที่ 1 (ปรากฏร่วมกัน 2 คำ) | 58 |
| 4.15 คำที่ปรากฏร่วมกันบ่อยที่สุดในข้อความกลุ่มที่ 1 (ปรากฏร่วมกัน 3 คำ) | 60 |
| 4.16 คำที่ปรากฏร่วมกันบ่อยที่สุดในข้อความกลุ่มที่ 1 (ปรากฏร่วมกัน 4 คำ) | 62 |
| 4.17 คำที่มีความถี่สูงสุด 100 อันดับแรกในข้อความกลุ่มที่ 2 | 64 |

| | |
|---|----|
| 4.18 คำที่ปรากฏร่วมกันบ่อยที่สุดในข้อความกลุ่มที่ 2 (ปรากฏร่วมกัน 2 คำ) | 65 |
| 4.19 คำที่ปรากฏร่วมกันบ่อยที่สุดในข้อความกลุ่มที่ 2 (ปรากฏร่วมกัน 3 คำ) | 67 |
| 4.20 คำที่ปรากฏร่วมกันบ่อยที่สุดในข้อความกลุ่มที่ 2 (ปรากฏร่วมกัน 4 คำ) | 69 |
| 4.21 คำที่มีความถี่สูงที่สุด 100 อันดับแรกในข้อความกลุ่มที่ 3 | 71 |
| 4.22 คำที่ปรากฏร่วมกันบ่อยที่สุดในข้อความกลุ่มที่ 3 (ปรากฏร่วมกัน 2 คำ) | 72 |
| 4.23 คำที่ปรากฏร่วมกันบ่อยที่สุดในข้อความกลุ่มที่ 3 (ปรากฏร่วมกัน 3 คำ) | 74 |
| 4.24 คำที่ปรากฏร่วมกันบ่อยที่สุดในข้อความกลุ่มที่ 3 (ปรากฏร่วมกัน 4 คำ) | 76 |
| 4.25 คำที่มีความถี่สูงที่สุด 100 อันดับแรกในข้อความกลุ่มที่ 4 | 79 |
| 4.26 คำที่ปรากฏร่วมกันบ่อยที่สุดในข้อความกลุ่มที่ 4 (ปรากฏร่วมกัน 2 คำ) | 80 |
| 4.27 คำที่ปรากฏร่วมกันบ่อยที่สุดในข้อความกลุ่มที่ 4 (ปรากฏร่วมกัน 3 คำ) | 82 |
| 4.28 คำที่ปรากฏร่วมกันบ่อยที่สุดในข้อความกลุ่มที่ 4 (ปรากฏร่วมกัน 4 คำ) | 84 |



สารบัญภาพ

| ภาพที่ | หน้า |
|--|------|
| 2.1 ตัวอย่างการตัดคำ | 6 |
| 2.2 ตัวอย่างการตัดคำและกำจัดคำหยุด | 8 |
| 2.3 กราฟระหว่างจำนวนกลุ่ม (k) และ Sum Squared Error (SSE) | 13 |
| 3.1 กรอบพัฒนาการจัดกลุ่มข้อความบนสื่อสังคมออนไลน์ | 16 |
| 3.2 การแปลงชื่อเฉพาะเป็นกลุ่มของชื่อเฉพาะ | 21 |
| 3.3 กราฟระหว่าง จำนวนกลุ่ม (k) และค่า SSE | 23 |
| 4.1 Elbow method สำหรับการจัดกลุ่มข้อความ | 29 |
| 4.2 Elbow method สำหรับการจัดกลุ่มข้อความที่ใช้การแปลงกลุ่มของชื่อเฉพาะ | 31 |
| 4.3 จำนวนข้อมูลของการจัดกลุ่มข้อความที่ไม่ได้ใช้การแปลงกลุ่มของชื่อเฉพาะ และวัดค่าความคล้ายคลึงแบบ Cosine Similarity | 32 |
| 4.4 Word Cloud ข้อความกลุ่มที่ 1 ของการจัดกลุ่มข้อความที่ไม่ได้ใช้การแปลงกลุ่ม ของชื่อเฉพาะ และวัดค่าความคล้ายคลึงแบบ Cosine Similarity | 32 |
| 4.5 Word Cloud ข้อความกลุ่มที่ 2 ของการจัดกลุ่มข้อความที่ไม่ได้ใช้การแปลงกลุ่ม ของชื่อเฉพาะ และวัดค่าความคล้ายคลึงแบบ Cosine Similarity | 36 |
| 4.6 Word Cloud ข้อความกลุ่มที่ 3 ของการจัดกลุ่มข้อความที่ไม่ได้ใช้การแปลงกลุ่ม ของชื่อเฉพาะ และวัดค่าความคล้ายคลึงแบบ Cosine Similarity | 40 |
| 4.7 Word Cloud ข้อความกลุ่มที่ 4 ของการจัดกลุ่มข้อความที่ไม่ได้ใช้การแปลงกลุ่ม ของชื่อเฉพาะ และวัดค่าความคล้ายคลึงแบบ Cosine Similarity | 43 |
| 4.8 จำนวนข้อมูลของการจัดกลุ่มข้อความที่ไม่ได้ใช้การแปลงกลุ่มของชื่อเฉพาะ และวัดค่าความคล้ายคลึงแบบ Cosine Similarity | 48 |
| 4.9 Word Cloud ข้อความกลุ่มที่ 1 ของการจัดกลุ่มข้อความที่ใช้การแปลงกลุ่ม ของชื่อเฉพาะ และวัดค่าความคล้ายคลึงแบบ Cosine Similarity | 48 |

| | |
|---|----|
| 4.10 Word Cloud ข้อความกลุ่มที่ 2 ของการจัดกลุ่มข้อความที่ใช้การแปลงกลุ่ม ของชื่อเฉพาะ และวัดค่าความคล้ายคลึงแบบ Cosine Similarity | 50 |
| 4.11 Word Cloud ข้อความกลุ่มที่ 3 ของการจัดกลุ่มข้อความที่ใช้การแปลงกลุ่ม ของชื่อเฉพาะ และวัดค่าความคล้ายคลึงแบบ Cosine Similarity | 52 |
| 4.12 Word Cloud ข้อความกลุ่มที่ 4 ของการจัดกลุ่มข้อความที่ใช้การแปลงกลุ่ม ของชื่อเฉพาะ และวัดค่าความคล้ายคลึงแบบ Cosine Similarity | 54 |



บทที่ 1

บทนำ

1.1 ที่มาและความสำคัญของปัญหา

สื่อสังคมออนไลน์ (Social Media) เป็นช่องทางในการติดต่อสื่อสารหลักระหว่างผู้คนบนโลก เนื่องจากสื่อสังคมออนไลน์ มีบทบาทที่สามารถตอบสนองความต้องการของผู้ใช้งานอย่างครบถ้วน ครบถ้วน ครบถ้วน (Wichitboonyarak, 2011) เช่น สามารถแสดงความคิดเห็นและแลกเปลี่ยนข้อมูล สามารถสื่อสารระหว่างผู้ใช้งานอื่น ผ่านทางข้อความ เสียง และวิดีโอ เป็นต้น นอกจากนี้ผู้ใช้งานยังสามารถเข้าถึงเครือข่ายสังคมออนไลน์ (Social Network) ("Social Networks: An Introduction," 2009) ได้ง่ายผ่านทาง Facebook, Twitter, Instagram ด้วยอุปกรณ์ที่เชื่อมต่อเครือข่ายอินเทอร์เน็ตที่มีอยู่ (จำปีหอม, 2018) นอกจากสื่อสังคมออนไลน์ ถูกนำมาใช้เป็นเครื่องมือที่ใช้ในการติดต่อสื่อสารแล้ว สื่อสังคมออนไลน์ ยังเป็นแหล่งแสดงความรู้สึกความคิดเห็นของผู้ใช้ คนส่วนใหญ่มักแสดงความคิดเห็นหรือความรู้สึกผ่านโพสต์หรือคอมเมนต์บนสื่อสังคมออนไลน์ ซึ่งโพสต์หรือคอมเมนต์ เป็นข้อความสั้นๆแต่แสดงออกถึงความรู้สึกหรือความคิดเห็นของผู้ใช้ (Charu C. Aggarwal, 2012) จากสถิติปี 2563 พบว่าคนไทย 75% ของจำนวนประชากรในประเทศใช้สื่อสังคมออนไลน์ (Social Media) เฉลี่ยวันละ 3 ชั่วโมงต่อคน (DataReportal, 2020) ประเทศไทยมีผู้ใช้งาน Facebook มากที่สุดเป็นอันดับ 8 ของโลก อยู่ที่ประมาณ 41 ล้านบัญชี (Positioning, 2020a) รวมไปถึง Twitter พบว่าประเทศไทยมี ผู้ใช้งาน 5.7 ล้านคน อยู่อันดับที่ 15 ของโลก (Positioning, 2020b) จากสถิติพบว่าการใช้โซเชียลมีเดียบนสื่อสังคมออนไลน์ (Social Media) ในประเทศไทยผ่านทางช่องทางต่างๆ บนโลกออนไลน์ ได้แก่ ทวิตเตอร์ (Twitter: TW) เฟซบุ๊ก (Facebook: FB) อินสตาแกรม (Instagram: IG) พันทิป (Pantip) ยูทูบ (YouTube) และเว็บไซต์ข่าวต่าง ๆ มีอัตราเติบโตสูงขึ้น เช่นบน Twitter ทวิตเตอร์ มีการโพสต์ข้อความเฉลี่ยอยู่ที่ 500 ล้านข้อความ/วัน (Ahlgren, 2020) ทำให้ข้อมูลที่เกิดจากเครือข่ายสังคมออนไลน์มีอัตราการเพิ่มขึ้นอย่างรวดเร็ว (M. IKONOMAKIS, KOTSIANTIS, & TAMPAKAS, 2005)

ข้อมูลที่เกิดจากสื่อสังคมออนไลน์ (Social media) มีอัตราการเพิ่มขึ้นอย่างรวดเร็วจึงทำให้เกิดการวิเคราะห์สื่อสังคมออนไลน์ (Social Media Analytics) (Susan Etlinger, 2011) ซึ่งเป็นการหาสกัดความรู้ที่ซ่อนอยู่ภายใต้ข้อมูลจำนวนมากมายมหาศาล เพื่อให้สามารถดำเนินการตัดสินใจได้อย่างมีประสิทธิภาพบนพื้นฐานข้อเท็จจริงของข้อมูล การวิเคราะห์สื่อสังคมออนไลน์

(Social Media Analytics จากการเติบโตของจำนวนผู้ใช้งานและปริมาณข้อมูลบนสื่อสังคมออนไลน์ การวิเคราะห์ข้อความบนสื่อสังคมออนไลน์เป็นประโยชน์ในหลายๆ ด้านโดยเฉพาะด้านธุรกิจ ทำให้ทราบความรู้สึกของผู้บริโภค(Phraewattanakun & Pramote, 2013)

การวิเคราะห์ข้อความบนสื่อสังคมออนไลน์มีความท้าทาย(Group, 2020) สำหรับการวิเคราะห์ข้อความคือ ข้อความส่วนใหญ่ที่อยู่บนอินเทอร์เน็ตนิยมใช้ภาษาที่มีโครงสร้างประโยคที่ไม่แน่นอน (Unstructured Data) และไม่ถูกต้องตามหลักไวยากรณ์ทางภาษา ทำให้ยากต่อการวิเคราะห์งานวิจัยด้านการวิเคราะห์ข้อความ จึงได้นำเทคนิคการ วิเคราะห์เหมืองข้อความ (Text mining) และการประมวลภาษาธรรมชาติ (Natural Language Processing: NLP) ตัวอย่างเช่น การตัดคำ การกำจัดคำหยุด เป็นต้น มาประยุกต์ใช้เพื่อวิเคราะห์ข้อความ(Masawee Masdisornchote, 2014)

การจัดกลุ่มข้อความ (Text Clustering) เป็นหนึ่งในเทคนิคการวิเคราะห์ข้อความ การจัดกลุ่มข้อความ คือ การจัดกลุ่มตามเนื้อหาของเอกสาร โดยที่ไม่จำเป็นต้องมีการ กำหนดกลุ่มหรือหมวดหมู่ของเอกสาร ด้วยวิธีการทางการเรียนรู้ด้วยคอมพิวเตอร์ (Machine Learning) ทำให้สามารถจัดกลุ่มของข้อความตามลักษณะและความคล้ายคลึงกันของข้อความได้ซึ่งเป็นวิธีการเอกสารอัตโนมัติ สามารถช่วยประหยัดแรงงานมนุษย์เป็นอย่างมากเพราะไม่ต้องใช้มนุษย์ในการวิเคราะห์และจัดกลุ่มข้อความด้วยตนเอง (Sebastiani, 2002)

งานวิจัยนี้ได้มุ่งเน้นกรอบพัฒนาการจัดกลุ่มข้อความบนสื่อสังคมออนไลน์ที่เป็นภาษาไทยโดยปรับปรุงวิธีการเตรียมข้อมูลโดยประยุกต์ใช้การรู้จำชื่อเฉพาะและการแปลงคำที่มีความหมายคล้ายคลึงกันเพื่อเพิ่มจัดกลุ่มและสรุปหัวข้อของข้อความที่กล่าวถึงมหาวิทยาลัยธรรมศาสตร์บนสื่อสังคมออนไลน์ทวิตเตอร์ และสามารถนำการจัดกลุ่มข้อความที่กล่าวถึงมหาวิทยาลัยธรรมศาสตร์ ไปใช้สำหรับการวิเคราะห์ความคิดเห็นของผู้ใช้สื่อสังคมออนไลน์ที่มีต่อมหาวิทยาลัย ในด้านต่างๆ เป็นต้น

1.2 วัตถุประสงค์ของงานวิจัย

งานวิจัยฉบับนี้ผู้วิจัยได้กำหนดวัตถุประสงค์สำหรับการจัดทำงานวิจัยไว้ ดังนี้

1.1.1 เพื่อศึกษาและพัฒนากรอบวิธีการจัดกลุ่มข้อความบนสื่อสังคมออนไลน์โดย k-means

1.1.2 เพื่อนำชื่อเฉพาะ และ คำพ้องความหมายมาปรับปรุงขั้นตอนการเตรียมข้อความสำหรับเพิ่มประสิทธิภาพของการจัดกลุ่มข้อความ

1.1.3 เพื่อกำหนดจำนวนกลุ่มของข้อความและสรุปความหมายหัวข้อของข้อความที่กล่าวถึงมหาวิทยาลัยธรรมศาสตร์

1.2 ขอบเขตของงานวิจัย

ผู้วิจัยได้กำหนดขอบเขตการวิจัยไว้ ดังนี้

1.2.1 ผู้วิจัยใช้ข้อความที่กล่าวถึงมหาวิทยาลัยธรรมศาสตร์ซึ่งทำการเก็บรวบรวมในเดือนเมษายน พ.ศ.2563

1.2.2 การจัดกลุ่มข้อความใช้อัลกอริทึม k-means และกำหนดพารามิเตอร์ค่าความคล้ายคลึงของข้อมูลได้แก่ Euclidean Distance, Cosine Similarity, Jaccard Similarity และ Pearson Similarity

1.2.3 ผู้วิจัยใช้ ภาษา Python สำหรับการเตรียมข้อมูลและสกัดคุณลักษณะ และใช้โปรแกรม Rapidminer สำหรับการจัดกลุ่มและการประเมินประสิทธิภาพข้อความ

1.3 ประโยชน์ที่คาดว่าจะได้รับ

1.3.1 ได้ศึกษาและพัฒนาวิธีการในการเตรียมข้อความที่เหมาะสมสำหรับข้อความที่กล่าวถึงมหาวิทยาลัยบนสื่อสังคมออนไลน์

1.3.2 ได้ศึกษาและพัฒนากรอบการจัดกลุ่มข้อความที่กล่าวถึงมหาวิทยาลัยต่างๆ ด้วยวิธีอัตโนมัติ

1.4 นิยามศัพท์ที่เกี่ยวข้อง

1.4.1 มหาวิทยาลัย(University) หมายถึง สถาบันอุดมศึกษาที่มีวัตถุประสงค์ให้การศึกษาด้านวิชาการและวิชาชีพชั้นสูงหลายสาขาวิชา หรือหลายกลุ่มสาขาวิชา เพื่อให้ประกาศนียบัตร อนุปริญญา ปริญญา และประกาศนียบัตรบัณฑิตแก่ผู้สำเร็จการศึกษา รวมทั้งดำเนินการวิจัยและให้บริการทางวิชาการแก่สังคม และทำนุบำรุงศิลปะและวัฒนธรรมของชาติ

1.4.2 สื่อสังคมออนไลน์(Social Media) หมายถึงสื่ออิเล็กทรอนิกส์ ซึ่งเป็นสื่อกลางที่ให้บุคคลทั่วไปมีส่วนร่วมสร้างและแลกเปลี่ยนความคิดเห็นต่างๆ ผ่านอินเทอร์เน็ตได้

1.4.3 การจัดกลุ่มข้อความ(Text Clustering) หมายถึง ประมวลผลภาษาธรรมชาติซึ่งทำการจัดกลุ่มข้อความโดยรวมข้อความที่มีลักษณะเนื้อหาและความหมายที่คล้ายคลึงกันเข้าไว้เป็นกลุ่มเดียวกัน

1.4.4 การจัดกลุ่มแบบเค-มีน(k-means) หมายถึง อัลกอริทึมการเรียนรู้แบบไม่ต้องสอน (Supervised Learning) สำหรับจัดกลุ่มข้อมูลตามค่าความคล้ายคลึงของข้อมูล

1.4.5 ชื่อเฉพาะ(Named Entity) หมายถึง วิสามานยนาม หรือเรียกว่า คำนามเฉพาะ คือ คำนามที่ใช้เรียกชื่อเฉพาะของคน สัตว์ สิ่งของ หรือสถานที่ เป็นคำเรียกเฉพาะเจาะจงลงไปว่าเป็นใครหรือเป็นอะไร

1.4.6 คำพ้อง(Synonym) หมายถึง คำหรือวลีที่เขียนต่างกันแต่มีความหมายเหมือนกันหรือใกล้เคียงกันมาก



บทที่ 2

วรรณกรรมและงานวิจัยที่เกี่ยวข้อง

จากการศึกษาค้นคว้าวรรณกรรมและงานวิจัยที่เกี่ยวข้อง ผู้วิจัยนำเสนอ โดยแบ่งเป็นรายละเอียด ดังต่อไปนี้ 2.1 การจัดกลุ่มข้อความ (Text Clustering) 2.2 การเตรียมข้อมูล (Data Preparation) 2.3 การแปลงชื่อเฉพาะเป็นกลุ่มของชื่อเฉพาะ 2.4 การแปลงข้อความที่ความหมายคล้ายคลึงกัน (Synonyms Substitution) 2.5 การสกัดคุณลักษณะ (Feature Extraction) 2.6 การจัดกลุ่มข้อมูลแบบเค-มีน (K-Means Clustering) 2.7 Sum Squared Error 2.8 Elbow method และ 2.9 งานวิจัยที่เกี่ยวข้องกับการจัดกลุ่มข้อความ

2.1 การจัดกลุ่มข้อความ (Text Clustering)

การจัดกลุ่มข้อความ คือวิธีการตัดแยกข้อความออกเป็นกลุ่ม ซึ่งถือว่าเป็นการเรียนรู้โดยไม่มีผู้สอน (Unsupervised Learning) เป็นการเรียนรู้ที่ให้เครื่องจักรนั้นสามารถเรียนรู้ได้ด้วยตนเอง โดยไม่ต้องเฉลยหรือระบุประเภทของกลุ่มแต่ละกลุ่ม โดยวิธีการตัดแยกข้อความออกเป็นกลุ่มอาศัยความเหมือนกันของข้อความ (Ahadi, Singh, Bower, & Garrett, 2022)

2.2 การเตรียมข้อมูล (Data Preparation)

การเตรียมข้อมูล (Data Preparation) หมายถึงกระบวนการ ที่จัดการกับข้อมูลดิบ (Raw Data) ที่ได้รับมา เพื่อปรับเปลี่ยนข้อมูลให้อยู่ในรูปแบบที่เหมาะสม ที่จะนำไปโหลดเข้าฐานข้อมูล หรือนำไปวิเคราะห์ต่อไป การเตรียมข้อมูล อาจจะแตกต่างกันไปสำหรับแต่ละชุดข้อมูล โดยขึ้นอยู่กับประเภทของข้อมูล และวัตถุประสงค์ของการวิเคราะห์ข้อมูล (Jotikabukkana, Sornlertlamvanich, & Manabu, 2016) สำหรับกระบวนการเตรียมข้อมูลที่อยู่ในรูปแบบข้อความ มีขั้นตอนเบื้องต้นดังนี้

2.2.1 การตัดคำ (Word Segmentation)

การตัดคำ คือ การแยกคำที่มีความหมายในแต่ละคำออกจากกัน ดังภาพที่ 1 ซึ่งเป็นกระบวนการพื้นฐานสำหรับงานทางด้านการวิเคราะห์ข้อความที่จะทำให้แต่ละเอกสาร แบ่งออกเป็นคำย่อย การตัดคำสามารถทำได้ 3 เทคนิค ได้แก่ การตัดคำโดยใช้กฎ การตัดคำโดยใช้พจนานุกรม และ การตัดคำโดยใช้คลังข้อมูล สำหรับการตัดข้อความภาษาไทยปัจจุบันสามารถทำได้โดยใช้เครื่องมือหรือโปรแกรมเพื่อช่วยอำนวยความสะดวกในการตัดคำได้หลากหลายเครื่องมือ

การตัดคำเป็นกระบวนการพื้นฐานสำหรับงานทางด้านการวิเคราะห์ข้อความที่จะทำให้แต่ละเอกสาร แบ่งออกเป็นคำย่อยดังภาพที่ 2.1 เพื่อนำคำเหล่านั้นไปสกัดคุณลักษณะ (Feature Extraction) หรือประมวลผลอื่นๆ การตัดคำสามารถทำได้โดยแยกคำที่มีความหมายในแต่ละคำออกจากกัน (Runnaphongsa & Urathamkun, 2006)

ภาพที่ 2.1

ตัวอย่างผลลัพธ์การตัดคำ

ประโยค : “ผมรอให้คุณสารภาพรักผมไม่ไหวแล้วคุณเขตคิดแต่เรื่องดีๆจริงจริงงคุณเขตสายมโนใครพินขอให้ยกมือขึ้น เจ้าสาวจำยอมทุกวันศุกร์2005นเสาร์อาทิตย์ 2015 นช่อง 7hd กด 35 liveสดfanpagech 7 hd แอปพลิเคชัน ch7hd และเว็บไซต์ละครย้อนหลังรับชมได้ก่อนใครch7hdไฮไลท์ละคร”

ผลลัพธ์การตัดคำ : | ผม | รอ | ให้ | คุณสารภาพ | รัก | ผม | ไม่ | ไหว | แล้ว | คุณเขต | คิด | แต่ | เรื่อง | ดี | ๆ | จริง | จริง | ง | คุณเขตสายมโน | ใคร | พิน | ขอ | ให้ | ยก | มือ | ขึ้น | เจ้า | สาว | จำยอม | ทุก | วัน | ศุกร์ | 2005 | นเสาร์ | อาทิตย์ | 2015 | นช่อง | 7 | hdกด | 35 | live | สดfanpagech | 7 | hd | แอปพลิเคชัน | ch | 7 | hd | และ | เว็บไซต์ | ละคร | ย้อนหลัง | รับ | ชม | ได้ | ก่อน | ใคร | ch | 7 | hdไฮไลท์ | ละคร |

หมายเหตุ: ผลลัพธ์การตัดคำเกิดจาก Deep Cut

การตัดคำสามารถทำได้ 3 เทคนิคได้แก่ การตัดคำโดยใช้กฎ (Rule-Based Approach) เป็นการตัดคำตามกฎเกณฑ์ที่กำหนดขึ้นโดยใช้หลักไวยากรณ์ภาษาไทยหรือพิจารณาจากพยัญชนะ สระ วรรณยุกต์ การันต์ การขึ้นย่อหน้าใหม่ เป็นต้น (Mahatthanachai, 2017) วิธีนี้จะทำให้ประหยัดเวลาในการตัดคำ ใช้พื้นที่ประมวลผลน้อย และไม่จำเป็นต้องใช้คลังคำหรือพจนานุกรม ในการตัดแต่ทำให้การตัดคำมีถูกต้องแม่นยำน้อย และไม่สามารถจัดการคำที่อยู่นอกเหนือจากกฎที่ตั้งขึ้นได้ การตัดคำโดยใช้พจนานุกรม (Dictionary-Based Approach) เป็นการเปรียบเทียบคำที่เก็บ

อยู่ในพจนานุกรมกับเอกสารที่ต้องการตัดคำ การตัดคำแบบนี้จะมีผลลัพธ์ที่ไม่เหมือนกันขึ้นอยู่กับคำที่เก็บอยู่ในแต่ละพจนานุกรมด้วย การตัดคำโดยใช้พจนานุกรมมีความแม่นยำสูง สามารถเพิ่มคำลงในพจนานุกรมได้ แต่ต้องใช้พื้นที่สำหรับเก็บคำในพจนานุกรมมาก การตัดคำโดยใช้คลังข้อมูล(Corpus-Based Approach) เป็นการตัดคำที่นำหลักการทางสถิติ ความน่าจะเป็น และกลไกการเรียนรู้เพื่อประมวลผลทางภาษา (Haruechaiyasak, Kongyoung, & Damrongrat, 2008) โดยอิงผลเฉลยการตัดคำจากคลังข้อมูลที่มีการตัดคำที่ถูกต้อง การตัดคำโดยใช้คลังข้อมูล มีความแม่นยำสูงมากเนื่องจากจะตัดคำโดยอิงจากบริบทรอบข้างด้วย แต่ต้องใช้เวลาและพื้นที่ในการประมวลผลสูง และความแม่นยำขึ้นอยู่กับข้อกำหนดผลเฉลยการตัดคำของคลังข้อมูลการตัดคำสามารถทำได้โดยใช้เครื่องมือหรือโปรแกรมเพื่อช่วยอำนวยความสะดวกในการตัดคำได้ หลากหลายเครื่องมือ ตัวอย่างเช่น Lexto, Nectec Lextoplus, Kucut และ PYThaiNLP เป็นต้น

2.2.2 การกำจัดคำหยุด (Stop Word)

คำหยุด หมายถึง คำที่ใช้กันโดยทั่วไปพบได้บ่อยในแต่ละเอกสาร คำหยุดเป็นคำที่ไม่มีนัยสำคัญต่อการจำแนกกลุ่มของข้อความ การกำจัดคำหยุดหรือการตัดคำหยุดคือการนำคำหยุดออกจากข้อความ การกำจัดคำหยุดถือว่าเป็นขั้นตอนที่สำคัญขั้นตอนหนึ่งในการลดขนาดคำภายในเอกสารและลดขนาดของการสร้างคุณลักษณะซึ่งจะส่งผลให้การประมวลผลเอกสารในขั้นตอนต่อไป คำหยุด (Stop word) หมายถึงคำที่ใช้กันโดยทั่วไปพบได้บ่อยในแต่ละเอกสาร คำหยุดเป็นคำที่

ไม่มีนัยสำคัญต่อการจำแนกกลุ่มของข้อความ หากทำการตัดคำเหล่านั้นออกจากข้อความก็ไม่ส่งผลให้ข้อความนั้นมีความหมายที่เปลี่ยนแปลง นอกจากนั้นแล้ว การกำจัดคำหยุดหรือการตัดคำหยุดคือการนำคำหยุดออกจากข้อความดังภาพที่ 2.2 การกำจัดคำยังส่งผลทำให้ขนาดของเอกสารและการประมวลผลลดลง ส่งผลให้เวลาที่ใช้ในการประมวลผลเร็วขึ้น (Choonui, 2012) โดยคำหยุดสามารถแบ่งออกได้ดังนี้

(1) สัญลักษณ์พิเศษ (Special character) หมายถึง อักขระที่เป็นเครื่องหมายต่าง ๆ เช่น + - ✓ / \$ = , " หรืออักขระที่ไม่มีอยู่บนแป้นพิมพ์ ตัวอย่างเช่น α, β, π เป็นต้น (Charoensuk & Sornil, 2018; Masawee Masdisornchote, 2016; Sirihattasak, 2019; Tapang, 2016)

(2) ตัวเลข หมายถึง สัญลักษณ์ที่ใช้แทนจำนวนจริง เช่น ๑๒ 53 VIII โดยปกติกำหนดไว้ให้มี 10 ตัว คือ 0 1 2 3 4 5 6 7 8 และ 9

(3) คำบุพบทคือ คำที่ใช้หน้าคำนาม คำสรรพนาม คำกริยา หรือคำวิเศษณ์ เพื่อบอกสถานภาพของคำเหล่านั้น หรือเพื่อแสดงความสัมพันธ์ระหว่างคำ

หรือประโยคคำสรรพนาม ตัวอย่างเช่น กับ ใน ของ ด้วย โดย เป็นต้น (Chumwatana & Wongkolkitsilp, 2019; Songram, Choompol, & Thipsanthia, 2016; Vateekul & Koomsubha, 2016)

(4) คำสรรพนาม หมายถึง คำที่ใช้แทนคำนามเพื่อไม่ต้องกล่าวคำนามนั้นซ้ำ เช่น เธอ เขา มัน เป็นต้น (Songram et al., 2016; Thetmueang & Chirawichitchai, 2017)

(5) คำวิเศษณ์หมายถึง คำที่ใช้ประกอบหรือขยายคำนาม สรรพนาม คำกริยา หรือคำวิเศษณ์ เพื่อให้ได้ใจความชัดเจนและละเอียดมากขึ้น เช่น ดี ชั่ว ใหญ่ ขาว ร้อน เย็น เป็นต้น (Thetmueang & Chirawichitchai, 2017)

ภาพที่ 2.2

ตัวอย่างการตัดคำและกำจัดคำหยุด

ข้อความ : “ไม่เห็นได้ผลเลยใช้มาจะ_6_ขวดแล้วอะ”

การตัดคำ : | ไม่ | เห็น | ได้ | ผล | เลย | ใช้ | มา | จะ | _ | 6 | _ | ขวด | แล้ว | อะ |

การกำจัดคำหยุด : | ไม่ | เห็น | ได้ | ผล | เลย | ใช้ | มา | จะ | แล้ว |

ที่มา: (Charoensuk & Sornil, 2018)

2.3 การแปลงชื่อเฉพาะเป็นกลุ่มของชื่อเฉพาะ

การรู้จำชื่อเฉพาะ (Named Entity) คือ กระบวนการในการระบุคำนาม ที่ทำหน้าที่ชี้เฉพาะถึงสิ่งใดสิ่งหนึ่ง เช่น ชื่อบุคคล ชื่อองค์กร ชื่อสถานที่ รวมไปถึงข้อความที่แสดงเกี่ยวกับวันเวลา เงิน และเปอร์เซ็นต์ ซึ่งเป็นงานที่สำคัญงานหนึ่งของการประมวลผลภาษาธรรมชาติ ตัวอย่างการรู้จำชื่อเฉพาะ ข้อความ “ฉันอยู่โรงพยาบาลสนาม” หลังจากตัดคำจะได้ผลลัพธ์เป็น “ฉัน | อยู่ | โรงพยาบาล | สนาม” จากผลลัพธ์จะเห็นได้ว่าคำว่า “โรงพยาบาลสนาม” ไม่ควรถูกตัดคำแยกกัน เนื่องจากคำว่า “โรงพยาบาลสนาม” ดังนั้นจึงต้องผ่านการรู้จำชื่อเฉพาะเพื่อให้ได้คำที่มีความหมายไม่คลาดเคลื่อน

2.4 การแปลงข้อความที่ความหมายคล้ายคลึงกัน (Synonyms Substitution)

การแปลงข้อความที่ความหมายคล้ายคลึงกัน (Synonyms Substitution) คือการแปลงคำที่มีความหมายเดียวกันให้อยู่ในรูปคำเดียวกัน เพื่อลดมิติของคำ และช่วยให้กระบวนการประมวลผลได้รวดเร็วขึ้น ตัวอย่างการแปลงข้อความที่ความหมายคล้ายคลึงกัน ดังตารางที่ 2.1

ตารางที่ 2.1

ตัวอย่างการแปลงข้อความที่ความหมายคล้ายคลึงกัน

| ข้อความ | ข้อความที่มีการแปลงความหมาย |
|---------------------------|-----------------------------|
| บ้านฉันมีหมาตัวใหญ่ | บ้านฉันมีสุนัขตัวใหญ่ |
| ฉันกำลังจะไปกินข้าว | ฉันกำลังจะไปรับประทานอาหาร |
| หมาที่บ้านเธอจ้าม่าจ้งเลย | สุนัขที่บ้านเธออ้วนจ้งเลย |

2.5 การสกัดคุณลักษณะ (Feature Extraction)

เนื่องจากคอมพิวเตอร์ไม่สามารถจำแนกหมวดหมู่ของเอกสารซึ่งเป็นภาษาธรรมชาติโดยตรงได้ ดังนั้นจึงต้องแปลงเอกสารให้อยู่ในรูปแบบที่คอมพิวเตอร์สามารถใช้ในการเรียนรู้ได้ซึ่งขั้นตอนนี้เราเรียกว่าการสกัดคุณลักษณะ (Feature Extraction) วัตถุประสงค์ที่สำคัญของขั้นตอนการสกัดคุณลักษณะเอกสารคือการดึงคุณลักษณะ (Feature) ของเอกสารออกมา ซึ่งคุณลักษณะดังกล่าวมีหลากหลายรูปแบบเช่น คำ (Single Word) วลี (Phrase) ชุดของคำ (N-Gram) ความถี่ของคำ TF-IDF (Songram, 2016; Songram et al., 2016) เป็นต้น คุณลักษณะสามารถอยู่ในรูปแบบทางคณิตศาสตร์ได้เช่น ตัวเลข เวกเตอร์ เมทริกซ์ เป็นต้น คุณลักษณะแต่ละรูปแบบเหมาะสมกับการประมวลผลที่แตกต่างกันไป ดังต่อไปนี้

ความถี่คำ – ความถี่ผกผันของคำในเอกสาร (Term frequency-inverse document frequency : TF-IDF) เป็นหนึ่งในคุณลักษณะ (Feature) ที่ได้รับความนิยมอย่างแพร่หลาย ซึ่ง TF-IDF เป็นวิธีการทางสถิติที่ใช้ในการประเมินความสำคัญของคำต่อเอกสาร ความสำคัญของคำเป็นสัดส่วนโดยตรงกับจำนวนครั้งที่คำคำนั้นปรากฏในเอกสารนั้น ซึ่งเป็นวิธี การคิดค่าคุณลักษณะที่ใช้ค่า TF weighting ร่วมกับค่าความสำคัญของแต่ละคำ (inverse document frequency) เป็นวิธีการ

ถ่วงน้ำหนักที่นิยมกันมากในการทำงานด้านการสืบค้นสารสนเทศ และ เหมืองข้อความ เป็นวิธีการทางสถิติที่ใช้ในการประเมินความสำคัญของคำต่อเอกสาร ความสำคัญของคำเป็นสัดส่วนโดยตรงกับจำนวนครั้งที่คำคำนั้นปรากฏในเอกสารนั้น แต่ถูกลดความสำคัญโดยความถี่ของคำนั้นในกลุ่มเอกสารทั้งหมด (Thanasopon, Buranapanitkij, & Netisopakul, 2019)

$$tf - idf = tf * \log \left(\frac{N}{df} \right)$$

N คือ จำนวนเอกสารทั้งหมด ทั้งหมด

Df (t) คือ จำนวนเอกสาร ที่มี term t อยู่

2.6 การจัดกลุ่มข้อมูลแบบเค-มีน (K-Means Clustering)

เทคนิคการจัดกลุ่ม K-Means เป็นวิธีการเรียนรู้โดยไม่มีผู้สอนซึ่งเป็นวิธีที่ง่ายที่สุดสำหรับการจัดกลุ่มข้อมูล โดยทำการจัดกลุ่มข้อมูล (partition) ออกเป็นกลุ่มจำนวน K กลุ่ม (cluster) ตามที่ผู้ใช้งานกำหนด ในแต่ละกลุ่มจะมีจุดศูนย์กลางประจำกลุ่ม (centroid) ซึ่งได้จากการคำนวณค่าเฉลี่ยของกลุ่ม และการพิจารณาว่าข้อมูลหนึ่ง ๆ ควรอยู่กลุ่มใด จะพิจารณาจากระยะห่างระหว่างข้อมูลกับจุดศูนย์กลางประจำกลุ่มแต่ละกลุ่ม ข้อมูลที่มีระยะห่างจากจุดศูนย์กลางประจำกลุ่มใดน้อยที่สุดจะถูกจัดอยู่ในกลุ่มนั้น เมื่อข้อมูลทุกตัวถูกพิจารณาจัดกลุ่มแล้ว จะทำการคำนวณหาค่าเฉลี่ยจุดศูนย์กลางประจำกลุ่มใหม่ หากจุดศูนย์กลางของแต่ละกลุ่มมีการเปลี่ยนตำแหน่ง จะทำการวนซ้ำเพื่อจัดกลุ่มข้อมูลตามจุดศูนย์กลางที่เปลี่ยนไป และจะหยุดกระบวนการเมื่อจุดศูนย์กลางประจำกลุ่มของทุกกลุ่มไม่มีการเปลี่ยนแปลง สุดท้ายจะได้ข้อมูลที่มีลักษณะเหมือนหรือคล้ายกันอยู่ในกลุ่มเดียวกัน และข้อมูลที่แตกต่างกันอยู่คนละกลุ่มกัน (Salih, 2022)

การคำนวณระยะห่างระหว่างข้อมูลหรือการคิดค่าความคล้ายคลึงระหว่างข้อมูลสามารถทำได้หลายวิธี เช่น

- ค่า Euclidean Distance เป็นการวัดค่าความระยะห่างระหว่างข้อมูล 2 ข้อมูล ในระบบพิกัดเชิงเวกเตอร์ ที่มาจากทฤษฎีบทพีทาโกรัส ซึ่งถ้าข้อมูล 2 ตัวมีความคล้ายกันมาก แสดงว่าข้อมูลแต่ละตัว จะอยู่ใกล้กันมาก จะทำให้ค่า Euclidean Distance มีค่าน้อย ๆ เข้าใกล้ศูนย์ ซึ่งมีสมการดังต่อไปนี้

$$\text{Euclidean Distance} = \sqrt{(\text{doc}_{11} - \text{doc}_{12})^2 + (\text{doc}_{21} - \text{doc}_{22})^2 + \dots + (\text{doc}_{ij} - \text{doc}_{ij})^2}$$

เมื่อ doc_{ij} แทนค่าของคุณลักษณะที่ i ในข้อความที่ j

- ค่า Cosine Similarity เป็นการวัดความคล้ายคลึงของข้อมูลจำนวน 2 ข้อมูลที่สอดคล้องกับความสัมพันธ์ระหว่างเวกเตอร์ โดยแต่ละข้อมูลจะถูกแทนด้วยเวกเตอร์ขนาด N ซึ่งการเปรียบเทียบจะแสดงได้จากมุมโคไซน์ของมุมระหว่าง 2 เวกเตอร์ของข้อมูล หากข้อมูลมีความคล้ายคลึงกันมากค่ามุม Cosine Similarity จะมีค่าน้อย โดยค่าจะอยู่ในช่วง 0 ถึง 1

$$\text{Cosine Similarity} = \frac{W_i \cdot W_j}{\|W_i\| \cdot \|W_j\|}$$

เมื่อ W_i แทนเวกเตอร์ของข้อมูลที่ i และ j

- ค่า Jaccard Similarity เป็นสถิติที่ประยุกต์แนวคิดในทฤษฎีเซตเพื่อนำมาใช้เปรียบเทียบความคล้ายของกุ่มตัวอย่าง แนวคิดของค่า Jaccard Similarity คือการวัดค่าความคล้ายระหว่างกลุ่มประชากร 2 กลุ่ม โดยคำนวณจากขนาดของประชากรที่ทั้งสองกลุ่มตัวอย่างมีร่วมกัน (อินเตอร์เซกชันในทฤษฎีเซต)หารด้วยขนาดของประชากรทั้งหมดจากทั้งสองกลุ่มตัวอย่าง (ยูเนียนในทฤษฎีเซต) สำหรับข้อมูลที่เป็นจำนวนจริง ค่า Jaccard Similarity จะเป็นไปดังสมการ

$$\text{Jaccard Similarity} = \frac{W_i \cdot W_j}{\|W_i\|^2 + \|W_j\|^2 - W_i \cdot W_j}$$

เมื่อ W_i แทนเวกเตอร์ของข้อมูลที่ i และ j

- ค่า Pearson Similarity เป็นเทคนิคทางสถิติที่ใช้หาขนาดและทิศทางของความสัมพันธ์เชิงเส้นตรงระหว่างข้อมูลที่มีมาตราวัดแบบช่วง (Interval scale) หรืออัตราส่วน (Ratio scale) สองตัว (metrics variables) ซึ่งสัมพันธ์แบบเพียร์สัน ให้ชื่อตาม คาร์ล เพียร์สัน (Karl Pearson) ผู้คิดสถิตินี้ขึ้น โดยใช้ชื่อว่า product-moment coefficient of correlation

$$\text{Pearson Similarity} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}}$$

| | | |
|-------|-----------|-----------------------------------|
| เมื่อ | x_i | แทนคุณลักษณะที่ i ของข้อมูล x |
| | y_i | แทนคุณลักษณะที่ i ของข้อมูล y |
| | \bar{x} | แทนค่าเฉลี่ยของข้อมูล x |
| | \bar{y} | แทนค่าเฉลี่ยของข้อมูล y |

2.7 Sum Squared Error

Sum Squared Error (SSE) เป็นวิธีการวัดผลทางสถิติวิธีหนึ่งที่ใช้ในการวัดผลรวมความผิดพลาดของข้อมูล สำหรับการจัดกลุ่มข้อมูล Sum Squared Error (SSE) เป็นการวัดผลรวมของระยะห่างระหว่างข้อมูลกับจุดศูนย์กลางประจำกลุ่ม (centroid) ซึ่งมีสมการดังต่อไปนี้ (Nainggolan, Perangin-angin, Simarmata, & Tarigan, 2019) :

$$SSE = \sum_{i=1}^n d$$

เมื่อ

n คือ จำนวนข้อมูล

d คือ ระยะห่างระหว่างข้อมูลกับจุดศูนย์กลางประจำกลุ่ม (centroid)

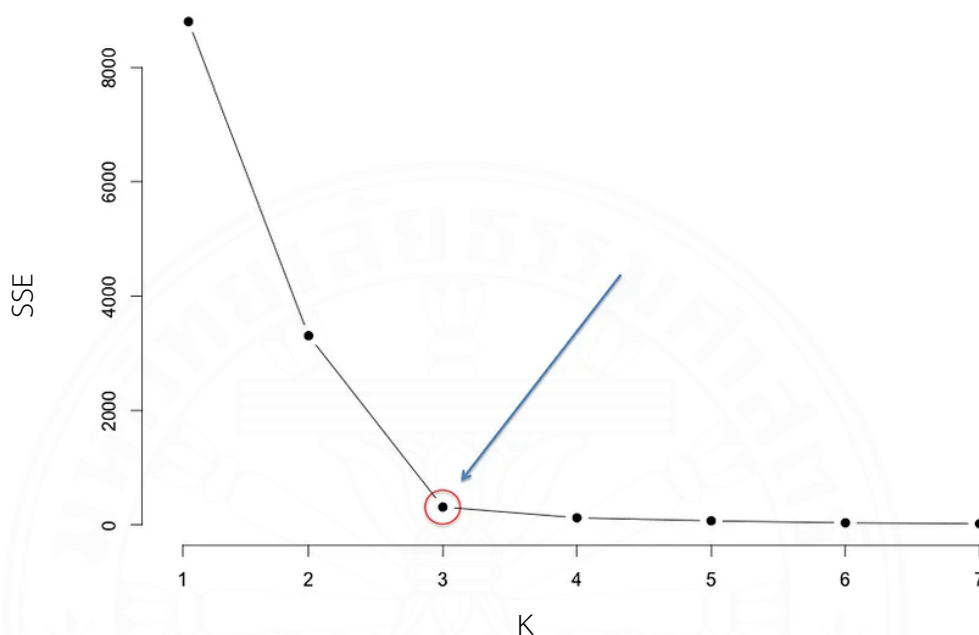
โดยหากค่า SSE มีค่าสูงแสดงให้เห็นว่าข้อมูลนั้นมีการกระจายตัวมาก ในทางตรงกันข้ามหากค่า SSE มีค่ามีค่าที่ต่ำแสดงให้เห็นว่าข้อมูลนั้นมีการเกาะกลุ่มกัน (วีระพันธ์ยานนท์ & มีสัจ, 2019)

2.8 Elbow method

Elbow method เป็นวิธีการเลือกค่า k เหมาะสม สำหรับการจัดกลุ่มข้อมูลด้วยอัลกอริทึมคือ k -means แนวคิดพื้นฐานของ Elbow method คือการวัดผลรวมของระยะห่างระหว่างข้อมูลกับจุดศูนย์กลางประจำกลุ่ม (centroid) ในแต่ละชุดข้อมูลที่มีค่า k แตกต่างกันไป จากนั้นใช้ SSE เป็นตัวบ่งชี้ประสิทธิภาพ ของข้อมูลที่ถูกแบ่ง K กลุ่ม หากจำนวนกลุ่มข้อมูล (k) เหมาะสม SSE จะแสดงมีค่าแตกต่างจาก SSE ของกลุ่มที่แล้วอย่างเห็นได้ชัดเจน หรือเมื่อนำค่า SSE และ จำนวนกลุ่มข้อมูล (k) มาสร้างเป็นกราฟดังภาพที่ 2.3 จุดที่เหมาะสมของจำนวนจำนวนกลุ่มข้อมูล (k) คือจุดที่กราฟมีลักษณะ “หัก” ที่สุดมากที่สุด (Larasati, Maren, & Wulandari, 2021)

ภาพที่ 2.3

กราฟระหว่างจำนวนกลุ่ม (k) และ Sum Squared Error (SSE)



2.9 งานวิจัยที่เกี่ยวข้องกับการจัดกลุ่มข้อความ

Thienrawish Pitchayaviwat (2016) เก็บรวบรวมความคิดเห็นและข้อเสนอแนะจากลูกค้าบนสื่อสังคมออนไลน์ที่เกี่ยวข้องกับบริการประกันภัย ได้ทำการจำแนกข้อความออกเป็น 3 กลุ่มได้แก่ กลุ่มที่ 1 ข้อความที่เกี่ยวข้องกับประกันรถยนต์ กลุ่มที่ 2 คำติชมของลูกค้าในการรับบริการเคลมประกัน กลุ่มที่ 3 ข้อความทั่วไป โดยใช้อัลกอริทึม k-means ร่วมกับ Self-Organization Map model (SOM) โดยใช้ TF-IDF เป็นคุณลักษณะพบว่า การใช้ k-means เพียงอย่างเดียวให้ประสิทธิภาพที่ดีกว่าการใช้ k-means ร่วมกับ Self-Organization Map model (SOM)

Shivani Vishwakarma et al. (2017) ทำการเปรียบเทียบประสิทธิภาพของอัลกอริทึม K-means กับ k-medoid สำหรับแบ่งกลุ่มข้อความ โดยทำการทดลองกับข้อความบนสื่อสังคมออนไลน์ โดยวัดผลจากการกระจายตัวของข้อมูลภายในกลุ่ม (Inter Cluster Distance) ซึ่งพบว่า K-means มีการกระจายตัวของข้อมูลภายในกลุ่มน้อยกว่า k-medoid แต่แลกมากด้วยการที่ K-means ใช้พื้นที่และเวลาสำหรับการประมวลผลที่มากกว่า k-medoid

Dhendra Marutho et al. (2018) นำข่าวต่างๆบนสื่อสังคมออนไลน์ มาเตรียมข้อมูล จากนั้นจัดกลุ่มข้อความด้วย K-means ซึ่งอาศัย Sum of Square Error (SSE) ร่วมกับเทคนิค Elbow method สำหรับการกำหนดจำนวนกลุ่ม(k)ที่เหมาะสมสำหรับการจัดกลุ่มข้อความ โดยสร้างกราฟความสัมพันธ์ระหว่าง Sum of Square Error (SSE) กับจำนวนกลุ่มต่างๆ (k) ซึ่งพบว่า การแบ่งข้อความออกเป็น 8 กลุ่มเหมาะสมที่สุด

สุภาพร วีระพันธ์ยานนท์ และ พยุง มีสัจ (2019) นำเสนอการประยุกต์ใช้เทคนิคเหมืองข้อความ (Text mining) สำหรับการจัดกลุ่มข้อมูลกรณีศึกษามะเร็ง โดยใช้ชุดข้อมูลทดสอบจากสื่อสังคมออนไลน์โดยรวบรวมข้อความจากเว็บไซต์ต่างๆ ที่เกี่ยวข้องกับ โรคมะเร็งการผู้วิจัยได้เสนอการเปรียบเทียบประสิทธิภาพของแต่ละคุณลักษณะได้แก่ TFIDF, WTFIDF และ FTFIDF ร่วมกับการทดลองการจัดกลุ่มข้อความแบบ Hierarchical ด้วยวิธี Single link วิธี Average link และวิธี Complete link ผลการทดลองแสดงให้เห็นว่าการใช้ คุณลักษณะแบบ WTFIDF อัลกอริทึมการจัดกลุ่มแบบ Complete link มีค่าความถูกต้อง (SSE) ในการจัดกลุ่มข้อความดีกว่า เมื่อเปรียบเทียบกับ อัลกอริทึมอื่น ๆ

Steinbach Michael et al. (2000) ได้ศึกษาเปรียบเทียบเทคนิคการจัดกลุ่มในการจัดกลุ่มเอกสารสามวิธี คือ การจัดกลุ่มแบบโครงสร้างลำดับชั้น (Agglomerative Hierarchical Clustering) ขั้นตอนวิธีการจัดกลุ่มแบบเค-มีน (K-Means) และ ขั้นตอนวิธีการจัดกลุ่มแบบไบเซ็คติงเค-มีน ซึ่งข้อมูลที่ใช้ในการทดลองเป็นเอกสารจำนวน 8 ชุด ซึ่งผ่านขั้นตอนการวิเคราะห์ด้วยวิธี Vector Space Model และหาความคล้ายคลึงระหว่างเอกสารแบบโคไซน์ก่อนเข้าสู่ขั้นตอนการจัดกลุ่ม โดยใช้ค่า Entropy, F-Measure และ Overall similarity ในการวัดประสิทธิภาพความถูกต้องของการจัดกลุ่ม ผลการทดลองพบว่า ขั้นตอน วิธีการจัดกลุ่มแบบไบเซ็คติงเค-มีน สามารถจัดกลุ่มเอกสารได้มีประสิทธิภาพกว่าการจัดกลุ่มแบบโครงสร้างลำดับชั้นและแบบเค-มีน

Karwan Jacksi et al. (2020) ได้เปรียบเทียบ ประสิทธิภาพอัลกอริทึมจัดกลุ่มข้อความระหว่าง อัลกอริทึม k-means และ การจัดกลุ่มแบบโครงสร้างลำดับชั้น (Hierarchical clustering : HAC) ใช้ข้อมูล 3 ชุดได้แก่ 1. 100 Movies (IMDB & Wiki) คือข้อความที่เกี่ยวข้องกับภาพยนตร์จาก IMDB & Wiki 2.Txt_Sentoken คือข้อความที่เกี่ยวข้องกับภาพยนตร์ และ 3. 20_Newsgrpap ข้อความที่เกี่ยวข้องกับข่าว โดยสกัดคุณลักษณะแบบ TF-IDF พบว่าจากจัดกลุ่มข้อความโดยอัลกอริทึม k-means มีประสิทธิภาพมากกว่าการจัดกลุ่มแบบโครงสร้างลำดับชั้น แต่การจัดกลุ่มแบบโครงสร้างลำดับชั้น มีการใช้เวลาในการประมวลผลที่น้อยกว่า

Aisyah Larasati et al. (2021) ใช้ข้อความที่กล่าวถึงสินค้าอิเล็กทรอนิกส์บนสื่อสังคมออนไลน์ทวิตเตอร์ หลังจากเตรียมข้อมูลได้ทำการสกัดคุณลักษณะแบบ TF-IDF จากนั้น จัดกลุ่ม

ข้อความโดย K-means และใช้ Sum of Square Error (SSE) ในการกำหนดจำนวนกลุ่มที่เหมาะสม (k) จากข้อมูลชุดนี้ พบว่า การเป็นข้อความเป็น 3 กลุ่ม เหมาะสมที่สุด จากนั้นทำการสร้าง Word Cloud สำหรับแสดงคำที่พบบ่อยในกลุ่มต่างๆ

Chotanansub Sophaken et al. (2022) ทำการดึงข้อมูลข่าวต่างๆ ที่อยู่ในสื่อสังคมออนไลน์ มาแบ่งกลุ่มของข้อความโดยสกัดคุณลักษณะแบบ TF-IDF และใช้ Latent Dirichlet Allocation (LDA) ในการจัดกลุ่มของข้อความและ สรุปหัวข้อของข้อความ ในแต่ละกลุ่ม โดยแสดงผลว่าข้อความแต่ละกลุ่ม มีคำใดปรากฏบ่อยที่สุด ซึ่งสามารถแบ่งได้ 8 กลุ่ม ดังนี้ กลุ่มที่ 1 พบคำว่า วัคซีน, ผู้เสียชีวิต, ติดเชื้อ, โควิด, อ่าง กลุ่มที่ 2 พบคำว่า หมู, เกษตรกร, ราคา, เนื้อ, อาหาร กลุ่มที่ 3 พบคำว่าพระ, บรม, เสด็จ, ถวาย, ราช กลุ่มที่ 4 พบคำว่า ติดเชื้อ, ดิน, สถานประกอบการ, เตือนภัย, ใช้จ่าย กลุ่มที่ 5 พบคำว่า กีฬา, การแข่งขัน, เกม, นักกีฬา, ฟุตบอล, กลุ่มที่ 6 พบคำว่า น้ำ, วิทยาศาสตร์, นวัตกรรม, คลัง, คาดการณ์ กลุ่มที่ 7 พบคำว่า ฉีดวัคซีน, ทางหลวง, ผู้โดยสาร, ประชากร, ขั้บรถ กลุ่มที่ 8 พบคำว่า การเลือกตั้ง, สมาชิกสภา, การคัดเลือก, วัง, ผู้แทนราษฎร

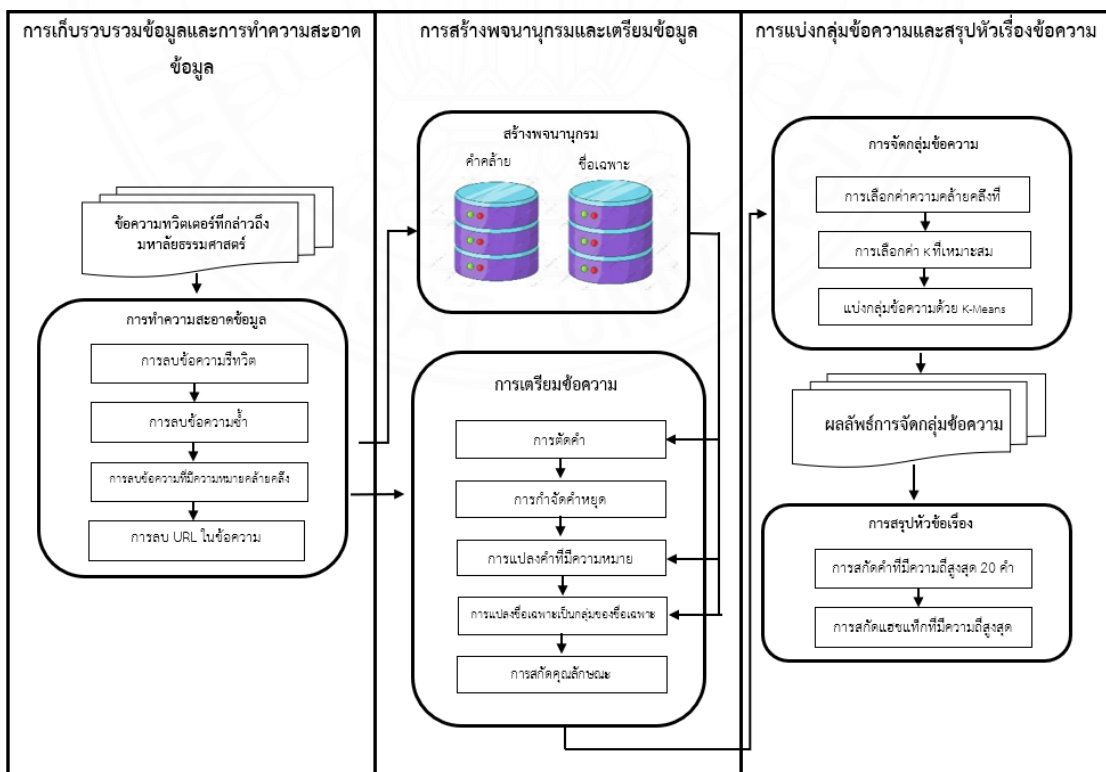


บทที่ 3 วิธีการวิจัย

งานวิจัยนี้มุ่งเน้นสร้างกรอบพัฒนาการจำแนกข้อความที่กล่าวถึงมหาวิทยาลัย ต่างๆบนสื่อสังคมออนไลน์บนสื่อสังคมออนไลน์ ดังแสดงรายละเอียดในภาพที่ 3.1 โดยทำการทดลองเพื่อสร้างกรอบพัฒนาการจำแนกข้อความซึ่งขั้นตอนการทดลอง มีขั้นตอนดังต่อไปนี้ 1.การเก็บรวบรวมข้อมูลและการกำหนดผลเฉลย เป็นขั้นตอนการรวบรวมข้อมูลจากสื่อสังคมออนไลน์และจัดการข้อมูลให้พร้อมสำหรับการจัดกลุ่มข้อความ 2.การเตรียมข้อมูลและการสกัดคุณลักษณะ เป็นขั้นตอนการทำให้ข้อความอยู่ในรูปแบบที่พร้อมสำหรับประมวลผลในอัลกอริทึม k-means ซึ่งขั้นตอนนี้ผู้วิจัยได้ทำการพัฒนาและเปรียบเทียบวิธีการเตรียมข้อมูลที่แตกต่างกันออกไป และ 3.การจัดกลุ่มข้อความและการประเมินประสิทธิภาพ เป็นขั้นตอนสุดท้าย ขั้นตอนนี้ผู้วิจัยได้ใช้ Rapidminer ในการจัดกลุ่มและประเมินประสิทธิภาพข้อความ

ภาพที่ 3.1

กรอบพัฒนาการจัดกลุ่มข้อความบนสื่อสังคมออนไลน์



3.1 การเก็บรวบรวมข้อมูล และการทำความสะอาดข้อมูล

3.1.1 วิธีการเก็บรวบรวมข้อมูล (Data Collection)

ผู้วิจัยได้ใช้ข้อความที่กล่าวถึงมหาวิทยาลัยวิทยาลัยธรรมศาสตร์ ในช่วง 1-30 เมษายน พ.ศ. 2563 จากทวิตเตอร์ (Twitter) โดยใช้ Twitter API ซึ่งเป็นเครื่องมือที่ได้รับการพัฒนาสำหรับการรวบรวมข้อมูลบน Twitter โดยอาศัยการค้นหาข้อความที่มีคำสำคัญต่างๆที่เกี่ยวข้องกับมหาวิทยาลัย เช่น “มหาวิทยาลัยธรรมศาสตร์” “มธ.” “ลูกแม่โดม” และ “TU” เป็นต้น ซึ่งสามารถรวบรวมข้อความได้ทั้งหมด 9,429 ข้อความ (ทะนานทอง, เทศขวัญ, & เกรียงไกรวิช, 2019)

3.1.2 การทำความสะอาดข้อมูล (Data Cleaning)

- การลบข้อความที่รีทวีต เนื่องจากข้อความที่รีทวีตนั้นแต่ละข้อความเขียนมีความหมายที่เหมือนกัน และเขียนเหมือนกันเมื่อนำไปประมวลผลจะทำให้ข้อมูลเกิดความซ้ำซ้อนกันและอาจให้ผลลัพธ์ที่ผิดพลาด โดยผู้วิจัยทำการรวมข้อความให้เหลือเพียง 1 ข้อความ
- การลบข้อความซ้ำ เป็นวิธีการเดียวกับ การลบข้อความที่รีทวีต เพียงแต่ข้อความที่ซ้ำกันนี้ไม่ใช่ข้อความรีทวีต แต่เป็นข้อความที่เขียนเหมือนกัน
- การลบข้อความที่มีความหมายคล้ายคลึงกัน เนื่องจากมีข้อความบางข้อความเป็นข้อความที่มีความหมายเดียวกัน แต่มีข้อความบางส่วนเขียนไม่เหมือนกัน ผู้วิจัยทำการรวมข้อความให้เหลือเพียง 1 ข้อความ โดยคำนึงจากการเหมือนกันของข้อความ ตัวอย่างเช่น “รับสอนพิเศษภาษาอังกฤษ เริ่มต้นที่ ชม.ละ 200 - 270 (ตัวต่อตัว) ไม่ว่าจะ เป็น ปูพื้นฐาน Grammar Gat Eng 9 สามัญ ผู้สอนเรียนอยู่ ม.ธรรมศาสตร์ มีปสภ.สอน4ปี (ตั้งแต่ปี60) #สอนพิเศษ #ติวเตอร์ #เรียนพิเศษ #หาติวเตอร์ #GATPAT #Dek65 #สอนพิเศษออนไลน์ #GatEng #เด็กชิว” กับข้อความ “รับสอนพิเศษภาษาอังกฤษ เริ่มต้นที่ ชม.ละ 200 - 300 (ตัวต่อตัว) ไม่ว่าจะ เป็น ปูพื้นฐาน Grammar Gat Eng 9 สามัญ ผู้สอนเรียนอยู่ ม.ธรรมศาสตร์ มีปสภ.สอน4ปี (ตั้งแต่ปี60) #สอนพิเศษ #ติวเตอร์ #เรียนพิเศษ #หาติวเตอร์ #GATPAT #Dek65 #สอนพิเศษออนไลน์ #GatEng #เด็กชิว” เห็นได้ว่า ข้อความทั้งสองข้อความมีความคล้ายคลึงกันจนเกือบจะเป็นข้อความเดียวกัน ต่างการเพียงข้อความที่ 2 นั้นมีการปรับเพิ่มของราคา

- การนำลิงก์ (URL) ที่อยู่ภายในข้อความ เช่น “ <https://t.co/NhmXx> ” และ “ <https://t.co/UoX> ” ลบออกจากข้อความ เป็นต้น

จากขั้นตอนข้างต้นนี้ทำให้ข้อความที่กล่าวถึงมหาวิทยาลัยทำธุรกรรมศาสตร์จำนวน 9,429 ข้อความ ลดลงเหลือ 2,401 ข้อความ

3.2 การสร้างพจนานุกรมการเตรียมและจัดการข้อมูล

3.2.1 การสร้างพจนานุกรมชื่อเฉพาะและพจนานุกรมคำพ้องความหมาย

ผู้วิจัยได้สร้างพจนานุกรมชื่อเฉพาะและพจนานุกรมคำพ้องความหมายจากการพิจารณาข้อความที่กล่าวถึงมหาวิทยาลัยต่างๆ และทำการเก็บรวบรวมคำเหล่านั้นมาสร้างเป็นพจนานุกรมขึ้นมาโดยนำพจนานุกรมนี้ไปใช้ในการพัฒนาการเตรียมข้อความ เพื่อเพิ่มประสิทธิภาพการจัดกลุ่มข้อความที่กล่าวถึงมหาวิทยาลัย โดยมีตัวอย่างพจนานุกรมดังนี้ ตารางที่ 3.1 และตารางที่ 3.2

ตารางที่ 3.1

ตัวอย่างคำในพจนานุกรมชื่อเฉพาะที่สร้างขึ้นจากคลังข้อมูลทวิตเตอร์

| ประเภทของชื่อเฉพาะ | ชื่อเฉพาะ |
|--------------------------|--|
| ชื่อเฉพาะของคณะ/สาขา | บริหารธุรกิจ สถาปัตย์ สนิสาดมธ และสังคมสงเคราะห์ เป็นต้น |
| ชื่อพื้นที่ต่างๆ | รังสิต เมืองทอง ปทุมธานี ดอนเมือง และฟิวเจอร์ เป็นต้น |
| ชื่อเฉพาะของดารา/นักแสดง | อุซอก ชินวอน ยูโตะ TillyBirds และtmd เป็นต้น |
| ชื่อเฉพาะข้อสอบ | ONET โอนีต 9วิชา #gat และ #pat1 เป็นต้น |
| ชื่อเฉพาะหอพัก | ดีคอนโดแคมป์รังสิต ดีคอนโด และ The kith เป็นต้น |

ตารางที่ 3.2

ตัวอย่างคำในพจนานุกรมคำพ้องความหมายที่สร้างขึ้นจากคลังข้อมูลทวิตเตอร์

| ความหมาย | คำที่มีความหมายเดียวกัน |
|------------------------|--|
| โรคโควิด-19 | โคจิด โควิด19 โควิด-19 Covid-19 COVID และโคโรนาไวรัส เป็นต้น |
| คณะสถาปัตยกรรมศาสตร์ | สถาปัต สถาปัตยกรรม และ สถาปัตยกรรม เป็นต้น |
| หน้ากากอนามัย | หน้ากาก แมส และ Mask เป็นต้น |
| NGV (รถโดยสารภายใน มธ) | NGV รถเอนจี้ รถNGV และ เอนจี้ เป็นต้น |
| โรงพยาบาลธรรมศาสตร์ | รพ.ธรรมศาสตร์และ โรงพยาบาลธรรมศาสตร์ เป็นต้น |
| Direct message | เดมมา ดีเอ็ม DM dm และ ดีเอ็ม เป็นต้น |

3.2.2 การตัดคำ (Word Segmentation)

การตัดคำ ผู้วิจัยเลือกใช้เครื่องมือ PyThaiNLP ซึ่งเป็นไลบรารีแพ็คเกจของภาษา Python เครื่องมือตัดคำนี้อาศัยการเทียบคำจากพจนานุกรมในการตัดคำ การตัดคำจะทำให้เกิดข้อความที่มีการแบ่งคำแยกออกมา ผู้วิจัยได้ปรับปรุงวิธีการตัดคำให้เครื่องมือตัดคำของ PyThaiNLP สามารถตัดคำได้ถูกต้องมากยิ่งขึ้นโดยใช้พจนานุกรมพจนานุกรมชื่อเฉพาะและพจนานุกรมคำพ้องความหมาย ในการเปรียบเทียบกับคำที่ถูกตัดและทำการปรับปรุงให้ตัดคำได้ถูกต้องมากยิ่งขึ้น ดังตารางที่ 3.3

ตารางที่ 3.3

ตัวอย่างผลลัพธ์ของการสกัดชื่อเฉพาะ

| ข้อความที่ผ่านการตัดคำ | ผลลัพธ์การสกัดชื่อเฉพาะ |
|--|--|
| “ท่า พระจันทร์” | “ท่าพระจันทร์” |
| “โรงพยาบาล สนาม” | “โรงพยาบาลสนาม” |
| “โรงพยาบาล ธรรมศาสตร์ เฉลิมพระเกียรติ” | “โรงพยาบาลธรรมศาสตร์ เฉลิมพระเกียรติ” |
| “นิธิ เอียว ศรี วงศ์ ” | “นิธิ เอียวศรีวงศ์ ” |

3.2.3 การกำจัดคำหยุด (Stop Word Removal)

การกำจัดคำหยุดในงานนี้ผู้วิจัยได้ใช้คำหยุดจาก PyThaiNLP ที่มีการสร้างคลังคำหยุดสำหรับภาษาไทย เช่น คำว่า “ที่” “ถึง” และ “ฉัน” เป็นต้น อีกทั้งผู้วิจัยกำหนดให้คำที่ปรากฏในข้อความทั้งหมดน้อยกว่า 5 คำเป็นคำหยุดที่ตัดออกด้วย รวมทั้งและคำที่มีอักษรตัวเดียว เช่น “ก” “ข” และ “ต” เป็นต้น

3.2.4 การแปลงคำที่มีความหมายคล้ายคลึงกัน

ขั้นตอนการแปลงคำที่มีความหมายคล้ายคลึงกัน คือการแปลงคำที่มีความหมายเดียวกันแต่เขียนไม่เหมือนกันให้เขียนในรูแบบเดียวกันตัวอย่างคำที่มีความหมายเดียวกันแต่เขียนไม่เหมือนกัน เช่น “กิน” และ “รับประทาน” ผู้วิจัยทำการแปลงคำที่มีความหมายคล้ายคลึงกันเพื่อเพิ่มอิทธิพลของคำนั้น ๆ ดังตารางที่ 3.4.

ตารางที่ 3.4

ตัวอย่างการแปลงคำที่มีความหมายคล้ายคลึงกัน

| ข้อความ | ข้อความที่ได้รับการแปลงคำที่มีความหมายคล้ายคลึงกัน |
|---|---|
| EXIM BANK บริจาคของใช้จำเป็นให้แก่กรุงเทพมหานครและ รพ.ธรรมศาสตร์เฉลิมพระเกียรติ เพื่อร่วมสนับสนุนให้พร้อมปฏิบัติงานดูแลรักษาผู้ป่วยโควิด-19 ประกอบด้วย พัดลม ชุด PPE แมส เครื่องทำน้ำร้อน/เย็น และน้ำดื่ม มูลค่ารวม 200,000 บ. ภายใต้โครงการ CSR “EXIM เพื่อสุขภาพชุมชน | EXIM BANK บริจาคของใช้จำเป็นให้แก่กรุงเทพมหานครและ รพ.ธรรมศาสตร์เฉลิมพระเกียรติ เพื่อร่วมสนับสนุนให้พร้อมปฏิบัติงานดูแลรักษาผู้ป่วยโควิด-19 ประกอบด้วย พัดลม ชุด PPE หน้ากากอนามัย เครื่องทำน้ำร้อน/เย็น และน้ำดื่ม มูลค่ารวม 200,000 บ. ภายใต้โครงการ CSR “EXIM เพื่อสุขภาพชุมชน |
| มูลนิธิชัยพัฒนา"โดย"กองทุนชัยพัฒนาสู้ภัยโควิด19 (และโรคระบาดต่างๆ)"มอบชุด PPE หน้ากากN95 และสิ่งของจำเป็นแก่โรงพยาบาลสนาม จังหวัดนนทบุรี และโรงพยาบาลสนาม มหาวิทยาลัยธรรมศาสตร์ ศูนย์รังสิต จังหวัดปทุมธานี | มูลนิธิชัยพัฒนา"โดย"กองทุนชัยพัฒนาสู้ภัยโควิด19 (และโรคระบาดต่างๆ)"มอบชุด PPE หน้ากากอนามัย และสิ่งของจำเป็นแก่โรงพยาบาลสนาม จังหวัดนนทบุรี และโรงพยาบาลสนาม มหาวิทยาลัยธรรมศาสตร์ ศูนย์รังสิต จังหวัดปทุมธานี |

3.2.5 การแปลงชื่อเฉพาะเป็นกลุ่มของชื่อเฉพาะ

ผู้วิจัยได้ทำการเปลี่ยนคำที่เป็นชื่อเฉพาะที่อยู่ในประเภทเดียวกันให้อยู่ในรูปประเภทเดียวกัน เพื่อเพิ่มสำคัญของชื่อเฉพาะในกลุ่มนั้นๆ โดยผู้วิจัยได้แปลงชื่อเฉพาะเป็นกลุ่มของชื่อเฉพาะโดยอ้างอิงจาก พจนานุกรมที่ได้สร้างขึ้นมา ซึ่งมีตัวอย่างการแปลงดังภาพที่ 3.2

ภาพที่ 3.2

การแปลงชื่อเฉพาะเป็นกลุ่มของชื่อเฉพาะ

| | |
|---|--|
| ข้อความ : | หาเมทหาค่าหอพักเดอะชาน |
| ข้อความหลังผ่านการตัดคำ : | หา เมท หาค่า หอพัก เดอะชาน |
| ข้อความหลังผ่านการแปลงกลุ่มของชื่อเฉพาะ : | หา เมท หาค่า หอพัก ชื่อเฉพาะของหอพัก |

3.2.6 การสกัดคุณลักษณะ (Feature Extraction)

ขั้นตอนนี้เป็นส่วนขั้นตอนที่เปลี่ยนข้อความให้อยู่ในรูปแบบของชุดตัวเลข ในงานนี้ได้ใช้ ความถี่คำ – ความถี่ผกผันของคำในเอกสาร (Term frequency-inverse document frequency : TF-IDF) ในการสกัดคุณลักษณะ โดยใช้ เครื่องมือ scikit-learn ซึ่งเป็นไลบรารีของภาษา Python ตัวอย่างการสกัดคุณลักษณะโดยวิธี TF-IDF ดังตารางที่ 3.5

ตารางที่ 3.5

ตัวอย่างการสกัดคุณลักษณะโดยใช้ TFIDF

| ข้อความ | ฉันทน์ | รัก | ธรรมศาสตร์ | จุฬา |
|---------------------------|--------|--------|------------|--------|
| ฉันทน์ รัก ธรรมศาสตร์ | 0.5015 | 0.5015 | 0.7049 | 0 |
| ฉันทน์ รัก จุฬา | 0.5015 | 0.5015 | 0 | 0.7049 |

3.3 การจัดกลุ่มข้อความและสรุปหัวเรื่องข้อความ

สำหรับขั้นตอนการจัดกลุ่มข้อความและสรุปหัวเรื่องข้อความ เป็นกระบวนการที่นำข้อความที่ได้รับการเตรียมข้อความ และสกัดคุณลักษณะ เรียบร้อย มาแบ่งกลุ่ม และผู้วิจัยได้นำข้อความที่แบ่งกลุ่มสำเร็จในแต่ละกลุ่มมาสรุปหัวเรื่องข้อความกลุ่มนั้นๆ ซึ่งการจัดกลุ่มข้อความประกอบไปด้วย ขั้นตอนดังนี้ การเลือกค่าความคล้ายคลึงที่เหมาะสม, การเลือกค่า K ที่เหมาะสม, และการจัดกลุ่มข้อความด้วย K-Means สำหรับการสรุปหัวเรื่องข้อความประกอบไปด้วยขั้นตอนดังนี้ การสกัดคำที่มีความถี่สูงสุด 100 คำ, การสร้าง itemset ของคำ, สรุปหัวข้อของกลุ่มข้อความจาก GNE ซึ่งขั้นตอนต่างๆมีรายละเอียดดังนี้

3.3.1 การเลือกค่าความคล้ายคลึงที่เหมาะสม

ขั้นตอนนี้เป็นขั้นตอนที่หาค่าความคล้ายคลึงที่เหมาะสมในการจัดกลุ่มข้อความ ซึ่งผู้วิจัยได้เลือกใช้โปรแกรม RapidMiner (Mierswa & Leblanc, 2001) โดยการนำข้อความที่ผ่านการสกัดคุณลักษณะหาคำน้ของหนักคำโดย TF-IDF มาหาค่า SSE โดยผู้วิจัยได้ใช้โอเปอเรเตอร์ต่างๆดังต่อไปนี้

โอเปอเรเตอร์ Read CSV สำหรับนำเข้าไฟล์ข้อมูลที่ผ่านการเตรียมและจัดการแล้ว โดยข้อมูลจะอยู่ในรูปแบบ ชุดของตัวเลข TF-IDF

- โอเปอเรเตอร์ KMeans เป็นโอเปอเรเตอร์สำหรับแบ่งกลุ่มข้อความด้วยวิธี k-means ซึ่งโอเปอเรเตอร์สามารถกำหนดพารามิเตอร์ค่าความคล้ายคลึงได้
- โอเปอเรเตอร์ Item Distribution Performance เป็นโอเปอเรเตอร์ที่สามารถวัดค่าผลรวมการกระจายของข้อมูลที่ถูกรวบรวมด้วย K-Means (SSE)

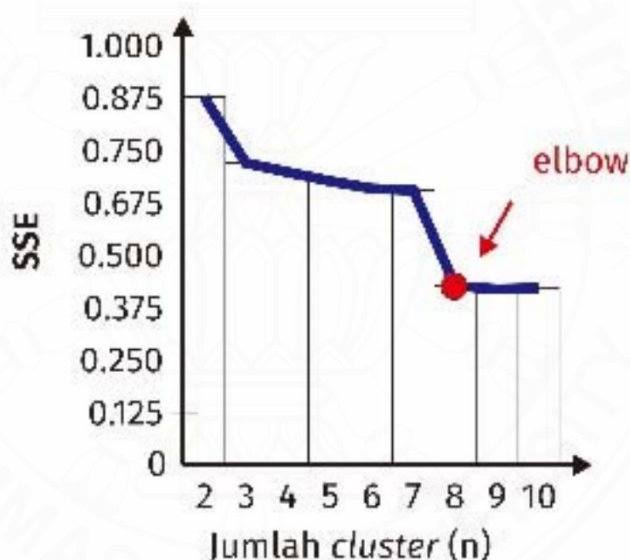
ผู้วิจัยได้ทำการทดลองแบ่งกลุ่มของข้อความด้วยอัลกอริทึม KMeans ซึ่งกำหนดค่า K ตั้งแต่ 2-20 กลุ่ม และทดลองใช้ค่าความคล้ายคลึงที่แตกต่างกันออกไป ได้แก่ Euclidean Distance, Cosine Similarity, Jaccard Similarity และ Pearson Similarity หลังจากนั้นผู้วิจัยวัดค่า SSE และเลือกค่าความคล้ายคลึงที่ให้ SSE น้อยที่สุด

3.3.2 การเลือกค่า K ที่เหมาะสม

ขั้นตอนนี้เป็นขั้นตอนที่หาค่า K ที่เหมาะสมในการจัดกลุ่มข้อความ ซึ่งมีวิธีที่คล้ายคลึงกับการเลือกค่าความคล้ายคลึงที่เหมาะสม เนื่องจากใช้ค่า SSE ที่วัดจากแบ่งกลุ่มข้อความด้วยอัลกอริทึม K-Means ที่กำหนดค่า K ตั้งแต่ 2-20 กลุ่มและนำ ค่า SSE มาสร้างเป็นกราฟระหว่างจำนวนกลุ่ม และค่า SSE ที่เกิดขึ้นในกลุ่มนั้นๆ ดังภาพที่ 3.3 เพื่อหาค่าความต่างระหว่าง SSE ที่เกิดขึ้นในแต่ละกลุ่ม และกำหนดให้กลุ่มที่มีค่าความต่างระหว่าง SSE มากที่สุดเป็นกลุ่มที่เหมาะสมที่สุดตามวิธีการของ Elbow Method

ภาพที่ 3.3

กราฟระหว่าง จำนวนกลุ่ม (k) และค่า SSE



ที่มา: (Mierswa & Leblanc, 2001)

3.3.3 การจัดกลุ่มข้อความด้วย K-Means

หลังจากผ่านกระบวนการเลือกค่าความคล้ายคลึงที่เหมาะสม และการเลือกค่า K ที่เหมาะสมผู้วิจัยได้แบ่งกลุ่มข้อความ ด้วยอัลกอริทึม K-Means ที่กำหนดค่าความคล้ายคลึงและจำนวนกลุ่มที่เหมาะสมดังข้างต้น โดยผู้วิจัยได้แบ่งเป็น 2 การทดลองคือ 1.การจัดกลุ่มข้อความ และ 2.การจัดกลุ่มข้อความที่ได้รับการแปลงชื่อเฉพาะเป็นกลุ่มของชื่อเฉพาะ

3.3.4 การสกัดคำที่มีความถี่สูงสุด 100 คำ

ผู้วิจัยนำข้อความที่แบ่งแล้วในแต่ละกลุ่มสกัดหาคำที่เกิดขึ้นบ่อยในกลุ่มนั้น ๆ 100 คำแรกดังตารางที่ 3.6 เพื่อนำคำเหล่านั้นไปสร้างเป็นไอเทมเซต สำหรับการกำหนดหัวเรื่องข้อความในกลุ่มนั้นๆ

ตารางที่ 3.6

ตัวอย่างคำที่มีความถี่สูงสุด 100 อันดับแรกของข้อความในกลุ่มที่ 1

| คำ | ความถี่ | คำ | ความถี่ | คำ | ความถี่ | คำ | ความถี่ |
|------------|---------|----------|---------|--------------------------|---------|---------------|---------|
| เรียน | 365 | ม. | 63 | วิชา | 36 | ชีวิต | 25 |
| พี่ | 154 | สอน | 63 | แบบนี้ | 35 | เดิน | 25 |
| สอบ | 152 | ชั้น | 61 | ชื่อพื้นที่ | 35 | เก่ง | 25 |
| ปี | 142 | ชอบ | 61 | ตึก | 35 | หน้า | 25 |
| ทำ | 136 | งง | 57 | ชื่อ | 33 | หนู | 25 |
| เพื่อน | 109 | กก | 57 | คุย | 33 | ตาย | 25 |
| อะ | 108 | นักศึกษา | 57 | นั่ง | 32 | ห้อง | 25 |
| รู้ | 104 | ตัวเอง | 57 | รอ | 31 | วะ | 25 |
| ชื่อมหาลัย | 102 | งาน | 56 | ชื่อสถานที่ในมหาวิทยาลัย | 30 | แถว | 25 |
| จบ | 100 | ตอนนี้ | 54 | ค่า | 30 | ค่า | 24 |
| คณะ | 95 | อ่าน | 51 | โท | 29 | ร้าน | 24 |
| ดี | 91 | บ้าน | 51 | กะ | 29 | กลัว | 24 |
| ชื่อคณะ | 88 | อี | 50 | อย่า | 29 | ออนไลน์ | 24 |
| ดู | 88 | ผม | 50 | นั่ง | 29 | เหมือนกัน | 24 |
| แม่ | 86 | ไหม | 48 | ตัว | 29 | รู้สึก | 24 |
| เรื่อง | 83 | เหมือน | 44 | ไทย | 28 | อิ | 24 |
| เด็ก | 81 | หา | 43 | โดน | 28 | ชื่อโรคติดต่อ | 23 |
| มหาลัย | 76 | บอ | 43 | เวลา | 28 | ดีกว่า | 23 |
| อะ | 76 | ตอบ | 42 | แล้วก็ | 27 | สังคม | 23 |
| น้อง | 75 | อาจารย์ | 40 | กิน | 27 | สนใจ | 23 |
| ติด | 75 | หรือ | 39 | หมอ | 27 | ขนาด | 23 |

ตารางที่ 3.6

ตัวอย่างคำที่มีความถี่สูงสุด 100 อันดับแรกของข้อความในกลุ่มที่ 1 (ต่อ)

| คำ | ความถี่ | คำ | ความถี่ | คำ | ความถี่ | คำ | ความถี่ |
|------|---------|-------|---------|------|---------|---------|---------|
| ถาม | 74 | พ่อ | 38 | ลูก | 26 | ทำงาน | 23 |
| เค้า | 71 | เลือก | 37 | ใจ | 26 | เล่น | 23 |
| ตอน | 69 | ใส่ | 37 | ชื่อ | 26 | ตอนนั้น | 22 |
| เจอ | 67 | ออ | 36 | อ. | 26 | หล่อ | 22 |

3.3.5 การสร้าง itemset

หลังจากที่สกัดคำที่มีความถี่สูงสุด 100 คำในแต่ละกลุ่มของข้อความ ผู้วิจัยได้นำคำที่มีความถี่สูงสุด 100 คำ มาสร้างเป็นไอเทมเซต สำหรับคำที่ปรากฏร่วมกันบ่อยที่สุดซึ่งได้กำหนดจำนวนคำที่ปรากฏร่วมกันจำนวน 2, 3 และ 4 คำ เพื่อหาชุดของคำที่ปรากฏร่วมกันในข้อความกลุ่มนั้นๆ โดยผู้วิจัยใช้เครื่องมือ Mlxtend.frequent_Patterns ซึ่งเป็นฟังก์ชัน ในไลบรารี Apriori ของภาษาไพธอน ในการสร้าง itemset

3.3.6 การสรุปหัวข้อของกลุ่มข้อความที่ได้รับการแปลงกลุ่มของชื่อเฉพาะ

สำหรับขั้นตอนสุดท้ายนี้ เป็นขั้นตอนการสรุปหัวข้อของกลุ่มข้อความที่ได้รับการแปลงชื่อเฉพาะเป็นกลุ่มของชื่อเฉพาะในแต่ละกลุ่ม โดยอ้างอิงจากคำที่มีความถี่สูงสุด 100 คำ ที่ปรากฏร่วมกันบ่อยที่สุด ซึ่งผู้วิจัยพิจารณาหัวข้อของกลุ่มข้อความจากคำต่างๆที่เกิดขึ้น

บทที่ 4

ผลการวิจัยและอภิปรายผล

การทดลองแบ่งกลุ่มข้อความที่กล่าวถึงมหาวิทยาลัยโดยใช้อัลกอริทึม k-means โดยใช้วิธีการคิดค่าความคล้ายคลึงแบบ Euclidean Distance, Cosine Similarity, Jaccard Similarity และ Pearson Similarity สามารถแบ่งการทดลองออกเป็น การทดลองการจัดกลุ่มข้อความที่ไม่ได้ใช้การแปลงกลุ่มของชื่อเฉพาะ และการทดลองการจัดกลุ่มข้อความที่ใช้การแปลงกลุ่มของชื่อเฉพาะ พบว่าได้ผลลัพธ์ดังนี้

4.1 ผลลัพธ์การเปรียบเทียบค่าความคล้ายคลึงที่เหมาะสม

สำหรับการเปรียบเทียบค่าความคล้ายคลึงที่เหมาะสมผู้วิจัย พิจารณาจากค่าเฉลี่ยของ SSE ที่เกิดขึ้นในแต่ละการจัดกลุ่มข้อความโดยเปรียบเทียบกันระหว่างค่าความคล้ายคลึงต่าง ๆ ได้แก่ Euclidean Distance, Cosine Similarity, Jaccard Similarity และ Pearson Similarity

4.1.1 ผลลัพธ์การเปรียบเทียบค่าความคล้ายคลึงที่เหมาะสมสำหรับการจัดกลุ่มข้อความที่ไม่ได้รับการแปลงกลุ่มของชื่อเฉพาะ

ตารางที่ 4.1

ค่าSum of Square Error เปรียบเทียบตามค่าความคล้ายคลึงสำหรับการจัดกลุ่มข้อความ

| k | ความคล้ายคลึง | | | |
|----|--------------------|-------------------|--------------------|--------------------|
| | Euclidean Distance | Cosine Similarity | Jaccard Similarity | Pearson Similarity |
| 2 | 0.7324 | 0.5125 | 0.5027 | 0.5761 |
| 3 | 0.6251 | 0.3957 | 0.3708 | 0.3865 |
| 4 | 0.4974 | 0.2623 | 0.2654 | 0.3221 |
| 5 | 0.4568 | 0.2217 | 0.2231 | 0.2557 |
| 6 | 0.3638 | 0.1870 | 0.1841 | 0.1832 |
| 7 | 0.3298 | 0.1729 | 0.1679 | 0.1646 |
| 8 | 0.2899 | 0.1319 | 0.1302 | 0.1386 |
| 9 | 0.2176 | 0.1181 | 0.1269 | 0.1429 |
| 10 | 0.2150 | 0.1200 | 0.1151 | 0.1063 |
| 11 | 0.2070 | 0.0989 | 0.1106 | 0.1066 |

ตารางที่ 4.1

ค่า Sum of Square Error เปรียบเทียบตามค่าความคล้ายคลึงสำหรับการจัดกลุ่มข้อความ(ต่อ)

| k | ความคล้ายคลึง | | | |
|-----------|--------------------|-------------------|--------------------|--------------------|
| | Euclidean Distance | Cosine Similarity | Jaccard Similarity | Pearson Similarity |
| 12 | 0.2004 | 0.0898 | 0.0983 | 0.0940 |
| 13 | 0.2266 | 0.0837 | 0.1008 | 0.0899 |
| 14 | 0.1446 | 0.0777 | 0.0903 | 0.0774 |
| 15 | 0.1608 | 0.0707 | 0.0771 | 0.0745 |
| 16 | 0.1596 | 0.0703 | 0.0784 | 0.0714 |
| 17 | 0.1319 | 0.0643 | 0.0668 | 0.0669 |
| 18 | 0.1396 | 0.0618 | 0.0673 | 0.0685 |
| 19 | 0.1262 | 0.0590 | 0.0639 | 0.0645 |
| 20 | 0.1084 | 0.0572 | 0.0600 | 0.0553 |
| ค่าเฉลี่ย | 0.2807 | 0.1503 | 0.1526 | 0.1603 |
| ผลรวม | 5.3328 | 2.8557 | 2.8998 | 3.0450 |

4.1.2 ผลลัพธ์การเปรียบเทียบค่าความคล้ายคลึงที่เหมาะสมสำหรับการจัดกลุ่มข้อความที่ใช้การแปลงกลุ่มของชื่อเฉพาะ

ตารางที่ 4.2

ค่า Sum of Square Error เปรียบเทียบตามค่าความคล้ายคลึงสำหรับการจัดกลุ่มข้อความที่ใช้การแปลงกลุ่มของชื่อเฉพาะ

| k | ความคล้ายคลึง | | | |
|---|--------------------|-------------------|--------------------|--------------------|
| | Euclidean Distance | Cosine Similarity | Jaccard Similarity | Pearson Similarity |
| 2 | 0.7081 | 0.5058 | 0.5202 | 0.5001 |
| 3 | 0.6173 | 0.3893 | 0.3850 | 0.3894 |
| 4 | 0.5067 | 0.2589 | 0.2839 | 0.3133 |
| 5 | 0.4475 | 0.2201 | 0.2344 | 0.2089 |
| 6 | 0.3845 | 0.1821 | 0.1800 | 0.1837 |
| 7 | 0.3350 | 0.1474 | 0.1633 | 0.1557 |
| 8 | 0.2680 | 0.1414 | 0.1447 | 0.1333 |
| 9 | 0.2368 | 0.1158 | 0.1451 | 0.1209 |

ตารางที่ 4.2

ค่า Sum of Square Error เปรียบเทียบตามค่าความคล้ายคลึงสำหรับการจัดกลุ่มข้อความที่ใช้การแปลงกลุ่มของชื่อเฉพาะ(ต่อ)

| k | ความคล้ายคลึง | | | |
|-----------|--------------------|-------------------|--------------------|--------------------|
| | Euclidean Distance | Cosine Similarity | Jaccard Similarity | Pearson Similarity |
| 10 | 0.2203 | 0.1045 | 0.1129 | 0.1171 |
| 11 | 0.2546 | 0.0996 | 0.1191 | 0.1107 |
| 12 | 0.2122 | 0.0862 | 0.0976 | 0.0965 |
| 13 | 0.1439 | 0.0842 | 0.0948 | 0.0805 |
| 14 | 0.1366 | 0.0749 | 0.0802 | 0.0770 |
| 15 | 0.1358 | 0.0694 | 0.0874 | 0.0753 |
| 16 | 0.1371 | 0.0654 | 0.0766 | 0.0733 |
| 17 | 0.1522 | 0.0646 | 0.0685 | 0.0646 |
| 18 | 0.0998 | 0.0603 | 0.0717 | 0.0599 |
| 19 | 0.1413 | 0.0582 | 0.0695 | 0.0627 |
| 20 | 0.1179 | 0.0573 | 0.0675 | 0.0617 |
| ค่าเฉลี่ย | 0.2766 | 0.1466 | 0.1580 | 0.1518 |
| ผลรวม | 5.2556 | 2.7853 | 3.0025 | 2.8847 |

จากตารางที่ 4.1 และ 4.2 พบว่าเมื่อทำการจัดกลุ่มข้อความโดยใช้วิธีการคิดค่าความคล้ายคลึงแบบ Cosine Similarity มีค่า SSE ต่ำที่สุดสำหรับการจัดกลุ่มข้อความและการจัดกลุ่มข้อความที่ใช้การแปลงกลุ่มของชื่อเฉพาะ

4.2 ผลลัพธ์การกำหนดจำนวนกลุ่ม(k)ที่เหมาะสมสำหรับการจัดกลุ่มข้อความ

การหาจำนวนกลุ่ม(k)ที่เหมาะสมสำหรับการจัดกลุ่มข้อความใช้วิธีหาผลต่างระหว่าง SSE กลุ่มก่อนหน้า (DIFF_SSE) ผู้วิจัยเลือกผลต่างระหว่าง SSE กลุ่มก่อนหน้า (DIFF_SSE) ที่มากที่สุดเป็นกลุ่มที่เหมาะสมที่สุด

4.2.1 การจัดกลุ่มข้อความ โดย อัลกอริทึม k-means พบว่าค่าจำนวนกลุ่ม (k) ที่เหมาะสมที่สุด คือ 4 กลุ่มโดยพิจารณาจากค่าของผลต่างSSE ดังตารางที่ 4.3 และภาพที่ 4.1

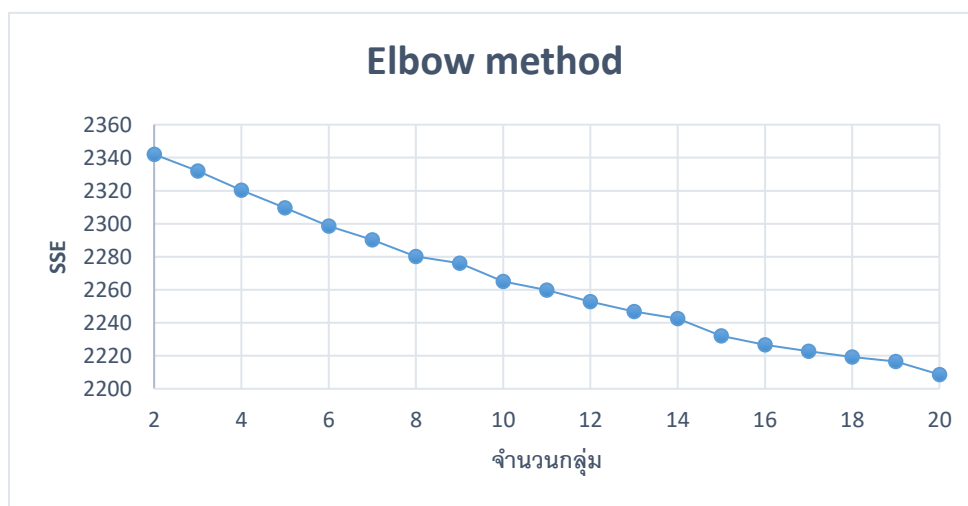
ตารางที่ 4.3

Sum of Square Error และผลต่างของ Sum of Square Error สำหรับการจัดกลุ่มข้อความ

| k | SSE | DIFF_SSE |
|----|----------|---------------|
| 2 | 2342.036 | - |
| 3 | 2331.879 | 10.158 |
| 4 | 2320.270 | 11.608 |
| 5 | 2309.676 | 10.595 |
| 6 | 2298.508 | 11.168 |
| 7 | 2290.232 | 8.276 |
| 8 | 2280.194 | 10.038 |
| 9 | 2276.090 | 4.104 |
| 10 | 2265.113 | 10.977 |
| 11 | 2259.870 | 5.243 |
| 12 | 2252.867 | 7.003 |
| 13 | 2246.844 | 6.023 |
| 14 | 2242.601 | 4.242 |
| 15 | 2231.973 | 10.628 |
| 16 | 2226.565 | 5.409 |
| 17 | 2222.715 | 3.849 |
| 18 | 2219.189 | 3.526 |
| 19 | 2216.583 | 2.606 |
| 20 | 2208.644 | 7.939 |

ภาพที่ 4.1

Elbow Method สำหรับการจัดกลุ่มข้อความ



4.1.2 การจัดกลุ่มข้อความที่ไม่ได้ใช้การแปลงกลุ่มของชื่อเฉพาะ โดย อัลกอริทึม k-means พบว่าค่าจำนวนกลุ่ม (k) ที่เหมาะสมที่สุด คือ 4 กลุ่มโดยพิจารณาจากค่าของผลต่าง SSE ดังตารางที่ 4.4 และ ภาพที่ 4.2

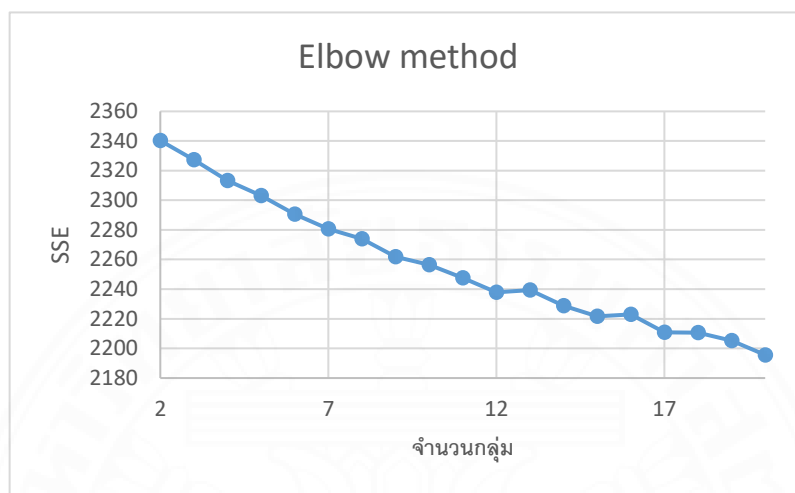
ตารางที่ 4.4

Sum of Square Error และผลต่างของ Sum of Square Error สำหรับการจัดกลุ่มข้อความที่ใช้การแปลงกลุ่มของชื่อเฉพาะ

| k | SSE | DIFF_SSE |
|----|----------|----------|
| 2 | 2340.270 | - |
| 3 | 2327.304 | 12.965 |
| 4 | 2313.294 | 14.010 |
| 5 | 2303.046 | 10.248 |
| 6 | 2290.496 | 12.550 |
| 7 | 2280.715 | 9.781 |
| 8 | 2273.965 | 6.749 |
| 9 | 2261.910 | 12.055 |
| 10 | 2256.442 | 5.468 |
| 11 | 2247.650 | 8.791 |
| 12 | 2237.748 | 9.902 |
| 13 | 2239.318 | -1.570 |
| 14 | 2228.708 | 10.610 |
| 15 | 2221.725 | 6.983 |
| 16 | 2222.904 | -1.179 |
| 17 | 2210.807 | 12.097 |
| 18 | 2210.545 | 0.262 |
| 19 | 2205.134 | 5.412 |
| 20 | 2195.602 | 9.532 |

ภาพที่ 4.2

Elbow method สำหรับการจัดกลุ่มข้อความที่ใช้การแปลงกลุ่มของชื่อเฉพาะ



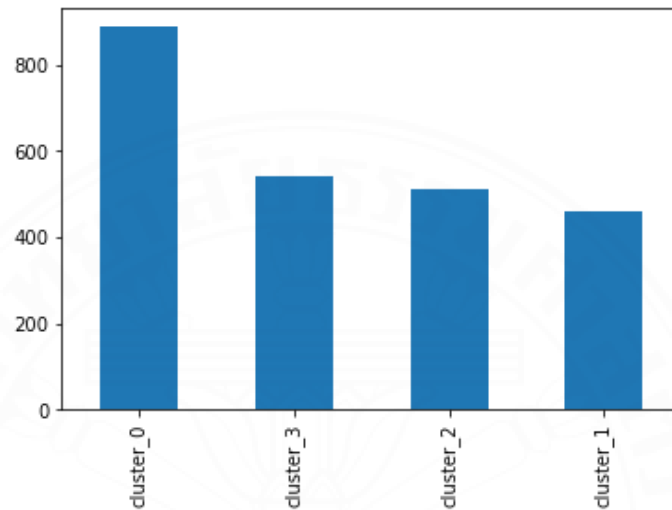
4.3 ผลลัพธ์การจัดกลุ่มข้อความบนสื่อสังคมออนไลน์ที่กล่าวถึงมหาวิทยาลัยธรรมศาสตร์

ผู้วิจัยจัดกลุ่มข้อความโดยอาศัยความคล้ายคลึงแบบ Cosine Similarity และใช้ จำนวนกลุ่ม 4 กลุ่ม ตามผลลัพธ์ การหาค่าความคล้ายคลึงและจำนวนกลุ่มที่เหมาะสม ดังการทดลองใน 4.1 และ 4.2

4.3.1 ผลลัพธ์การจัดกลุ่มข้อความที่โดยอัลกอริทึม k-means ซึ่งกำหนด $k = 4$ และ วัดค่าความคล้ายคลึงแบบ Cosine Similarity พบว่ากลุ่มข้อความที่ 1 มีข้อความ 888 ข้อความ กลุ่มข้อความที่ 2 มีข้อความ 460 ข้อความ กลุ่มข้อความที่ 3 มีข้อความ 512 ข้อความ และ กลุ่มข้อความที่ 4 มีข้อความ 541 ข้อความ ดังภาพที่ 4.3

ภาพที่ 4.3

จำนวนข้อมูลของการจัดกลุ่มข้อความที่ไม่ได้ใช้การแปลงกลุ่มของชื่อเฉพาะ และวัดค่าความคล้ายคลึงแบบ Cosine Similarity



จากข้อความกลุ่มที่ 1 จำนวน 888 ข้อความ พบคำที่มีความถี่สูงสุดได้แก่ “เรียน”, “พี่”, “ทำ”, “สอบ” “ปี”, “เพื่อน”, “รู้”, “คณะ”, “ดี” และ “จบ” ตามลำดับ ซึ่งนำคำที่มีความถี่สูงสุดมาแสดงได้ดัง ภาพที่ 4.4

ภาพที่ 4.4

Word Cloud ข้อความกลุ่มที่ 1 ของการจัดกลุ่มข้อความที่ไม่ได้ใช้การแปลงกลุ่มของชื่อเฉพาะ และวัดค่าความคล้ายคลึงแบบ Cosine Similarity



ตารางที่ 4.5

ตัวอย่างข้อความกลุ่มที่ 1 ของการจัดกลุ่มข้อความที่ไม่ได้ใช้การแปลงกลุ่มของชื่อเฉพาะ

| คำที่มี ความถี่สูง | ตัวอย่างข้อความ |
|-----------------------|---|
| “เรียน” | <ul style="list-style-type: none"> - สมัยเรียน ป. ตริ ที่ มธ. ผมใส่ตะแสมเสื้อยืดตลอด ถือเป็นอุดมการณ์ความเรียบง่ายและไม่ตัดสินคนด้วยเสื้อผ้า มีวิชาหนึ่งและเป็นวิชาเดียวที่เจอ ผมเข้าเรียนวันแรกอาจารย์ดูว่า ‘แต่งตัวแบบนี้มาเรียนถือว่าไม่เคารพอาจารย์ไม่เคารพห้องเรียน’ ผมเดินออกแล้วไปครี้อบวิชาที่นั่นทันที - นี่ตอนเรียนมธ. ก็คือตะแสมตลอด แทบจะชุดนอนไปเรียน ก็ไม่มีใคร judge กัน อจ. (ขณะตัวเอง) ไม่เคยว่าา ดีใจที่คนรอบตัวไม่ยุ่งกับเรื่องการแต่งตัวคนอื่น - เราว่าเรื่องการแต่งกาย ไม่ใช่จะมาบอกว่าต้องให้เกียรติสถานที่ซึ่งนั่นยังงี้เลย ให้มองเป็นเรื่องกาลเทศะดีกว่า สถานที่บางแห่งถูกคนยกย่องบูชาว่าต้องให้เกียรติ ทั้งที่เจ้าของสถานที่เขาก็ไม่ได้ต้องการอะไรขนาดนั้น ดูห้องเรียน ม.ธรรมศาสตร์ เป็นตัวอย่างเลย |
| “พี” | <ul style="list-style-type: none"> - มัธยมเป็นโปรเจกต์เดียวกันได้ไหมหละ ยังไง หรือธรรมศาสตร์ให้ไม่ได้ แต่พีก็ออล์พีให้ได้ เห้ยบ้าๆ - คือพณง.คูเรียร์มีรถก็เรทแบบนี้ แต่ฉันไปสาขาใน มธ อะ พีพณงเจ้าหน้าที่พัสดุต้องออกไปรับของที่กอฟวิว เห้ยมันคุ่มที่ทำเหนื่อยปาว - แม้นัดบั้งงไม่ไหวว จบนิตี มธ.เกียรตินิยมอันดับ2 ใครจะไฟท์กับแม่ คิดดีๆนะน้องนะ ถึงพีโน้ตจะไม่ได้ทำงานตามสายที่จบมา แต่พีโน้ตดูมีความสุขกับงานแต่งหน้าที่ทำมาเลย แต่งหน้าให้เจ้าสาวแต่ละคนสวยสวย แต่งหน้าพลุค้ก็สวย เวลาแต่งหน้าพลุค้ทอนหวานๆ อ้อหือ เหมือนพลุค้จะออกเรือน |
| “ทำ” | <ul style="list-style-type: none"> - การโคจรของ พีช ที่เคยฝากผลงาน มิว จากรักแห่งสยามและ มิน หรือ ดิณณ์ จาก บังเอิญรัก ทั้งคู่เคยผ่านผลงานความรักชายชาย และทั้งคู่เรียนฟิล์มที่ธรรมศาสตร์เหมือนกัน ตำแหน่ง Exclusive Producer คือเขาต้องควบคุมการผลิตงานและเขาต้องดูแลตั้งแต่กระบวนการคัดนักแสดง บท จนถ่ายทำ #หนังสือรุ่น - มีใครเคยไป U-Square บ้างไหมค้าบที่อยู่ติดกับมหาลัยผมเลย(มธ.รังสิต)ถ้าเคยไปฝากทำแบบสอบถามนี้หน่อยนะค้าบ - จำวันนี้ไว้จนตายเลยอะ วันที่ต้องไปเข้าโรงแรมอยู่เพราะหอมันจะปิด แล้วต้องออกไปฝึกงานไปเรียนไปสอบ ถ้าลูกอยากเข้าจุฬาๆจะขอให้ไปธรรมศาสตร์ เขาไม่มีวันทำกับหนูแบบนี้ ผู้บริหารต่าง mindsetต่าง #ไม่เอาคุณทอโน #ทอโนจุฬา #จุฬา #chula |

ตารางที่ 4.5

ตัวอย่างข้อความกลุ่มที่ 1 ของการจัดกลุ่มข้อความที่ไม่ได้ใช้การแปลงกลุ่มของชื่อเฉพาะ(ต่อ)

| คำที่มี ความถี่สูง | ตัวอย่างข้อความ |
|-----------------------|---|
| “สอบ” | <p>- อีกเรื่องหนึ่งที่ซื้อม ถ้าชั้นติดเตรียมอุดมศิลป์ภาษา มัณฑารัตนติดอักษรจุฬา นิเทศจุฬา ศิลปศาสตร์ มธ ไปเกินครึ่งละอะ แต่น้องแพทริคลาออกจากเตรียมอุดมตอน มอ4 ไปสอบ GED เพื่อมาแข่ง ชาวใจเด็ดซิบหายเลยชั้นนับถือมาก 55555555555555 #CHUANG2021 #PatrickRank13th</p> <p>- ทุกคนค้าบใครเห็นโพสนี้อย่าเพิ่งเลื่อนผ่านน้ำ เค้ากำลังตามหานักศึกษามาทำแบบสอบถามให้เค้าหน่อยคะ เป็นงานปลายภาคของบบเอง คนไหนเรียนธรรมศาสตร์หรือมีเพื่อนที่รู้จักเรียนอยู่รบกวนส่งต่อแบบสอบถามนี้ให้หน่อยนะคะ ฝากกรีทวิตด้วยนะคะ</p> <p>- เรื่องเนได้ประธานนักเรียนอะ ไม่จำบกล้วยเท่าวันที่เนเน บอกว่าสอบไม่ติด รัฐศาสตร์ มธ (น้องบอกว่าตอนนั้นจะไปต่ออินเดีย เราอะโคตรให้กำลังใจเลย) แล้วสลัมแซะกันเยอะมาก แล้วสรุปคือน้องแอดติด รัฐศาสตร์จุฬา กูอะโคตรสะใจ!!!!</p> |
| “ปี” | <p>- 6 ปีที่แล้ววันนี้ โกลด์สกรานต์ 2015 โปรดสังเกต อ.ยิม ยังอยู่กับปัญญาชนประชาธิปไตย ทว่า ถูกรองอธิการคนหนึ่งซบไล่ไสส่ง ไม่ให้จัดงานภายในรั้ว มธ ทำพระจันทร์ ทั้งทั้งที่อาจารย์สัญญา ธรรมศักดิ์ เคยกล่าวไว้ว่า "มีเสรีภาพทุกตารางนิ้ว" เมื่อปี 1973 (ภาพ:สงวน คุ่มรุ่งโรจน์)</p> <p>- เมื่อมีใครผ่านมานะคะ เราขอฝากแบบสอบถามการสำรวจการวางแผนชีวิตหลังจบการศึกษาของนักศึกษาชั้นปีสุดท้าย มธ. ฝากทำ ฝากกรีทวิต ขอให้ถูกหวงหวดหน้า รอดพ้นจากโควิดไปด้วยกัน</p> |
| “เพื่อน” | <p>- ปักหมุด แนวร่วมธรรมศาสตร์ศ กลับมาเล่น Instagram แล้ว! . เรากลับมาเปิด instagram แล้วเพื่อให้เพื่อน ๆ แชรข่าวสารได้ง่ายและรวดเร็วยิ่งขึ้น!</p> <p>- สังคมที่นั่นมันบ่มเพาะให้เด็กเตรียมตัวแต่เนิ่นๆด้วย ทั้งนี้อยู่กับตัวเด็กว่าจะปรับตัวได้มากแค่ไหน แต่เท่าที่รู้จักเพื่อนมา ศิลป์ภาษาเตรียม & อักษร พ /ศิลปศาสตร์ นิเทศพ วารสารมธ เยอะอยู่ -มธ ก็ไม่ต่างอะ ควรจะตรวจทั้งมธ คนไปก้ำติดโควิดแล้วคุณที่ดูคนไปก้ำวันนั้นสิคะ เทปซีคอกันกินเยอะกว่าคนไปงานรับเพื่อนใหม่อีกมั้ง</p> |
| “รู้” | <p>- สังคมที่นั่นมันบ่มเพาะให้เด็กเตรียมตัวแต่เนิ่นๆด้วย ทั้งนี้อยู่กับตัวเด็กว่าจะปรับตัวได้มากแค่ไหน แต่เท่าที่รู้จักเพื่อนมา ศิลป์ภาษาเตรียม & อักษรพ/ศิลปศาสตร์ นิเทศพ วารสาร มธ เยอะอยู่</p> <p>- กระจุกขยเกาะนั่นก็คือการ เมคฟัน กับความแตกต่างทางการเงินแบบนึ่ง คือรู้ว่าตัวเองมีทรัพย์สินที่คนอื่นไม่มีถึง ก็เลยมายืมพื้นที่ของคนประสบปัญหา มาเมคฟัน พอโดนคนจะรวมตัวกันเพื่อเอาชนะแบบจริงจัง ก็หอบแตก ตลก มธ สอนได้แค่นี้เหอ อายแทนอาจารย์บ๊วย</p> <p>- สัมผัสได้ตั้งแต่งานจับมือที่ไปครั้งแรกตอนนั้นไม่ได้รู้จักใครเลยนอกจากคุณกับตัน แต่รู้ว่่าน้ำหนึ่งเป็นรุ่นน้องที่ มธ. น่าจะพอมีเรื่องคุย และรู้ว่าต้องเรียนหนัก เลยไปให้กำลังใจ แต่กลับรู้สึกว่าได้พลังบวกกลับมาเยอะมาก โลกน่าอยู่ขึ้นเลย เธอน่ารักมาก ๆ รู้ตัวมั้ยคะ #มิลินเป็นคนน่ารัก</p> |

ตารางที่ 4.5

ตัวอย่างข้อความกลุ่มที่ 1 ของการจัดกลุ่มข้อความที่ไม่ได้ใช้การแปลงกลุ่มของชื่อเฉพาะ(ต่อ)

| คำที่มี ความถี่สูง | ตัวอย่างข้อความ |
|-----------------------|--|
| “คณะ” | <p>- นี่ตอนเรียนมธ. ก็คือเตะตลอด แทบจะชุนตอนไปเรียน ก็ไม่มีใคร judge กัน อจ. (คณะตัวเอง) ไม่เคยว่าาา ดีใจที่คนรอบตัวไม่ยุ่งกับเรื่องการแต่งตัวคนอื่น</p> <p>- ผมเรียนที่คณะศิลปศาสตร์ มธ ครับ เป็นอาจารย์จากสาขาสเปนและลาตินอเมริกาศึกษา เสียตายที่ผมแก่เกินไป เลยไปลงเรียน ป ตรี สาขานี้อีกใบไม่ได้แล้ว</p> <p>- และขอฝากสุดหล่อของเราหน่อยนะคะ น้ำหนึ่ง #Namneungbnk48 จบคณะวิทย์-สวล มธ. เป็นเซ็นเตอร์ #โตตติ่ง #Hightension กำลังจะมีหนังเรื่อง #ผ้าฝีบอก และซีรี่ ญ-ญ #GirlfriendPlannerTheseries มีหลายร่างให้เลือกหวิด ฝากน้ำหนึ่งไว้ในอ้อมอกอ้อมใจด้วยนะคะ</p> |
| “ดี” | <p>- เราว่าเรื่องการแต่งกาย ไม่ใช่จะมาบอกว่าต้องให้เกียรติสถานที่ที่ยังนั่งยังงี้เลย ให้มองเป็นเรื่องกาลเทศะดีกว่า สถานที่บางแห่งถูกคนยกย่องบูชาว่าต้องให้เกียรติ ทั้งที่เจ้าของสถานที่เขาก็ไม่ได้ต้องการอะไรขนาดนั้น ดูห้องเรียน ม.ธรรมศาสตร์ เป็นตัวอย่างเลย</p> <p>- องค์กรต่างๆควรตรวจเช็คให้ดีกว่าก่อนรับสมัครงาน ว่านศที่จบใหม่มาจาก จุฬา มธ นอนเคาฟงชาติรีไม่ เพราะคนพวกนี้ ไม่รักแม่แต่ชาติที่ตัวเองซุกหัวนอน ก็อย่าหวังว่าเค้าจะรักองค์กรของคุณ</p> <p>- เพื่อนทำไอโฟน8ตกสระแบบกลางสระเลยที่ตึก sc มธ. ตกไป3ชม.มั้งมีที่กู้ยัดน้ำเอาขึ้นมาให้ได้ ละดำน้าเจอเพราะว่ามีข้อความแดงๆมาทำให้เห็นแสง พอเอาขึ้นมาได้เหมือนมีสัญญาณข้อความก็แดงรัวๆเล่นได้ปกติเหมือนไม่เคยตกน้ำมาก่อน ดีที่ไม่ถอดใจเพราะตอนแรกคิดว่าพังแน่ๆ</p> |
| “จบ” | <p>- พี่มิวคะ จบรายการแล้วมาเฉลยหนูหน่อย - ตกลงพี่คุยกับใคร5555555555 - ตกลงสถานที่นั้นคือที่ไหน ไซ้ธรรมศาสตร์ไหม55555555555555 #HollywoodGNxMEW</p> <p>- เมื่อมีใครผ่านมานะคะ เราขอฝากแบบสอบถามการสำรวจการวางแผนชีวิตหลังจบการศึกษาของนักศึกษาชั้นปีสุดท้าย มธ. ฝากทำ ฝากรีทวีต ขอให้ถูกหวยงวดหน้า รอดพ้นจากโควิดไปด้วยกันนะคะ</p> |

ตารางที่ 4.6

ตัวอย่างข้อความกลุ่มที่ 2 ของการจัดกลุ่มข้อความที่ไม่ได้ใช้การแปลงกลุ่มของชื่อเฉพาะ(ต่อ)

| คำที่มี ความถี่สูง | ตัวอย่างข้อความ |
|-----------------------|--|
| “ขาย” | <p>- กลุ่มธรรมศาสตร์และการฝากบ้าน มีคนโพสต์ขายที่ในเกาะหวาย 350 ล้าน แล้วก็มีคนโพสต์หาหุ้นซื้อ 1 แแสนคน คนละ 3,500 เตรียมสร้างงานสร้างอาชีพให้นักศึกษา ปันผลปีละ 350 บาท 55555555555555555555 601</p> <p>- กรู้อัปเดตธรรมศาสตร์ฝากบ้านล่าสุดขายเกาะทั้งเกาะ 350 ล้าน ละมีคนเอาเครื่องบินไปบินสำรวจภาพมาให้ สุดเซดเสลดเปิดมาก 55555555 209</p> <p>- มีโพสต์ “ขายเกาะราคา 350 ล้านบาท” ในกลุ่ม มธ ฝากบ้าน งานนี้วงแชร์ต้องมา! มีคนโหวตไปบรจิตเชิญชวนลงขันคนละ ๒3,500 กัน 99,999 คน ชื่อเกาะให้เป็นทรัพย์สินมหา’ลัย เป็นสวัสดิการชาว มธ แต่เก็บเงินคนอื่นที่มาใช้บริการ 39</p> |
| #Dek64” | <p>- พุทโท ธรรมศาสตร์ ัฒโม ธรรมศาสตร์ พุทโท ธรรมศาสตร์ ัฒโม ธรรมศาสตร์ พุทโท ธรรมศาสตร์ ัฒโม ธรรมศาสตร์ พุทโท ธรรมศาสตร์ ัฒโม ธรรมศาสตร์ พุทโท ธรรมศาสตร์ ัฒโม ธรรมศาสตร์ พุทโท ธรรมศาสตร์ ัฒโม ธรรมศาสตร์ พุทโท ธรรมศาสตร์ ัฒโม ธรรมศาสตร์ 99 สาธุขอให้ทุกคนอยู่ในเกณฑ์ปลอดภัย ตัวฉันก็เช่นกัน โชคดีค่าาาา #dek64</p> <p>- มาแล้วววิวซีทฟริตามสัญญาา ซีทเกร็ดความรู้มหาลัยธรรมศาสตร์และคณะรัฐศาสตร์จากพี่KittyCat ไปโหดได้เลยยย และยังมีซีทอื่นๆที่น่าสนใจอีกมากมาย ถ้าสนใจก็กดซื้อทางเว็บได้เลยน่าา #สิงห์แดง #รัฐศาสตร์มธ #รัฐศาสตร์ #วิชาเฉพาะมธ #ซีทสรุป #ธรรมศาสตร์ #Dek64</p> <p>- บนค้ะ ติดหมอ ศรร. รามา รามาจุฬารณั ฬ มธ. มอ. มข. มช. มบ. วชิระ วพม. ้วยยยย ติดสายสุขภาพอะค่า อะไรก็ได้ จะกลับไปสอนน้องๆที่ร.ที่ไม่ได้โอกาสในการเรียนพิเศษ เพื่อลดความเหลื่อมล้ำในสังคม ฟรีค่าาา #dek64</p> |
| “บาท” | <p>- มีพร้อมส่ง 200 แพ็ค • แพ็คละ 30 บาทค่า ขึ้นต่ำ 3 แพ็คคะ นั้รับได้ ธรรมศาสตร์-รังสิต พิพเจอร์ คลอง1-6 นวนคร พระอินทร์ • ส่งตจว 5 แพ็คขึ้นไป ค่าส่ง 70 ค่ะ สนใจเดมมาได้เลยนะคะ //รับซื้อบิลด้วยนะคะ ให้65% เลยค่า #BKBeatตัวท็อป #BKxBamBam #BKBeatYourBe #bambam</p> <p>- ดิคอนโด แคมปัส รีสอร์ท รังสิต ธรรมศาสตร์ รังสิต ปทุมธานี คลองหลวง ขนาดห้อง 30.00 ตารางเมตร 1ห้องนอน 1ห้องน้ำ ชั้น8 เฟส 1. ตึกB ค่าส่วน 3เดือน 4040 บาท ราคา 2.1 ล้าน ทักเดม</p> <p>- งงไหม ไม่น่างนะ นีก็ตั้งใจเขียนกติกาอย่างละเอียดแล้ว ถ้ามีคนร่วมกิจกรรมนี้กับเราครบ 100 คน นั้นเท่ากับโรงพยาบาลสนามจะมีค่าอาหารเพิ่มอีก 1 มื้อ // ไซ้คะ ตอนนีโรงพยาบาลสนาม ธรรมศาสตร์ตอนนี้ใช้เงินค่าอาหารมื้อละประมาณ 20,000 บาท</p> |

ตารางที่ 4.6

ตัวอย่างข้อความกลุ่มที่ 2 ของการจัดกลุ่มข้อความที่ไม่ได้ใช้การแปลงกลุ่มของชื่อเฉพาะ(ต่อ)

| คำที่มี ความถี่สูง | ตัวอย่างข้อความ |
|-----------------------|---|
| “ราคา” | <p>- [ฟรี] CanCam Magazine ราคา นิตยสาร+การ์ด 370 (มัดจำได้) นิตยสารอย่างเดียว 200 เลือกการ์ดได้ตามลำดับการโอน รายละเอียด/สั่งซื้อ รู้การ์ดทันทีที่ของถึงบ้านญี่ปุ่น รับจำนวนจำกัด 40/60 (เล่มถัดไป +10) มีบัตรรับสยาม/มธ. #ตลาดนัดสวนแครอท</p> <p>- ขายสัญญาหอเพลินอพาร์ทเมนต์ ในซอยเจปาร์ค มีพิตเนสฟรี สระว่ายน้ำ มีสนามแบดมินตันราคาสัญญา 11,000 บาท ห้องพักเดือนละ 5,500 มีรถตู้รับส่งเข้ามา มีตู้เย็น ทีวี พร้อมออกค่าเปลี่ยนชื่อสัญญาให้ #หอนอกมธ #หอมธ #มธ #ขายสัญญาหอ สนใจทักdm มาได้เลยค่ะ พร้อมเข้าอยู่ได้เดือนพฤษา ไม่ค่าส่วนกลาง</p> <p>- ขายค่ะ จะเอาเงินไปจ่ายค่าสอบ Konica c35 ef ใช้งานได้ดีมีระบบ แฟลชติด วัตแสงได้ ห้องฟิล์มสะอาด 2500 รวมส่ง น้ดรับแถวพิวเจอร์ มธ. ตามแนวbts(ได้ถึงแค่btsหมอชิต) #ขายกล้องฟิล์ม #กล้องฟิล์ม #กล้องฟิล์มมือสองราคาถูก #กล้องฟิล์มมือสอง #ส่งต่อ</p> |
| “สนใจ” | <p>- มีพร้อมส่ง 200 แพ็ค • แพ็คละ 30 บาทค่า ชั้นต่ำ 3 แพ็คค่ะ น้ดรับได้ ธรรมศาสตร์-รังสิต พิวเจอร์ คลอง1-6 นวนคร พระอินทร์ • ส่งตจว 5 แพ็คขึ้นไป ค่าส่ง 70 ค่ะ สนใจเดมมาได้เลยนะคะ //รับซื้อบิลด้วยนะคะ ให้65% เลยค่า #BKBeatตัวท็อป #BKBeatYourBe #bambam</p> <p>- มาแล้ววววชิตพีรตามสัญญาา ชิตเกร็ดความรู้มหาลัยธรรมศาสตร์และคณะรัฐศาสตร์จากพี่KittyCat ไปโหลดได้เลยยย และยังมีชิตอื่นๆที่น่าสนใจอีกมากมาย ถ้าสนใจก็กดซื้อทางเว็บได้เลยน้าา #สิงห์แดง #รัฐศาสตร์มธ #วิชาเฉพาะมธ #ชิตสรุป #ธรรมศาสตร์ #Dek64</p> <p>- เมื่อก่อนเห็นโพสขายที่ดินบนเกาะหวาย จ.ตราด 350ลบ. ในกรู๊ปขายของ มธ. ก็ไม่ได้สนใจอะไร วันนี้พบว่าชาวมธ.จัดระดมทุนกันซื้อที่ดินแถบนี้ได้เงินจะครบแล้วด้วย ยังไม่พอมีคนเอาเครื่องบินไปบินดูเกาะมาให้จริงๆ และมีคนเสนอให้จัดขายเป็น digital token อีก</p> |
| “ห้อง” | <p>- ดิคอนโด แคมปัส รีสอร์ท รังสิต ธรรมศาสตร์ รังสิต ปทุมธานี ขนาดห้อง 30.00 ตารางเมตร 1 ห้องนอน 1ห้องน้ำ ชั้น8 เฟส 1. ดีกบ ค่าส่วน 3เดือน 4040 บาท ราคา 2.1 ล้าน ทักเดม</p> <p>- หาเมทผู้หญิง หารห้องหอ 18-22 หอสร้างใหม่ฝั่งตรงข้ามมธค้บ ห้อง studio plus 8,500 หารตกคนละ 4,250 บาทค่า เริ่มเข้าอยู่เดือนสิงหาค่ะ เปิดเทอม1 หรือใครสะดวกหอรห้องตั้งแต่ช่วงsummerก็ได้ ทักมาคุยกันก่อนได้เลยนะคะ เราอยู่ป้2ค้บ #หอนอกมธ #หอพักมธ #Dek64 #ทีมมธ #หาเมทมธ #มธรังสิต</p> <p>- ขายสัญญาหอเพลินอพาร์ทเมนต์ ในซอยเจปาร์ค มีพิตเนสฟรี สระว่ายน้ำ มีสนามแบดมินตันราคาสัญญา 11,000 บาท ห้องพักเดือนละ 5,500 มีรถตู้รับส่งเข้ามา มีตู้เย็น ทีวี พร้อมออกค่าเปลี่ยนชื่อสัญญาให้ #หอนอกมธ #หอมธ #มธ #ขายสัญญาหอ สนใจทักdm มาได้เลยค่ะ พร้อมเข้าอยู่ได้เดือนพฤษา ไม่ค่าส่วนกลาง 7</p> |

ตารางที่ 4.6

ตัวอย่างข้อความกลุ่มที่ 2 ของการจัดกลุ่มข้อความที่ไม่ได้ใช้การแปลงกลุ่มของชื่อเฉพาะ(ต่อ)

| คำที่มี ความถี่สูง | ตัวอย่างข้อความ |
|-----------------------|--|
| “#หอนอกมธ” | <p>- ขายสัญญา Jpark ตึกD (ตึกใหม่) ค่าเช่า ตอนนี้ลด 20% ทักเดมมาสอบถามเพิ่มเติมได้น้า #หอพักมธ #หอนอกมธ #หอมธ #ขายสัญญาหอมธ #หาหอมธ #มธ #ธรรมศาสตร์ #หอพักมธกรุง</p> <p>- ขายสัญญาห้องloft loft 11,000บาท อยู่ครบถึงปีได้ค่าสัญญาคืน ค่าห้องเดือนละ 5,500 บาท ช่วงโควิดค่าห้องลดเหลือ3,500 ค่าไฟยูนิตละ6บาท ค่าน้ำยูนิตละ35บาท เข้าอยู่ได้เดือนพ.ค.เลย #Dek64 #หอนอกมธ #หอมธ #ทิมมธ</p> <p>- ขายสัญญาหอCity Parkฝั่งหญิง ชั้น9 ระเบียบหันเข้าอินเตอร์ อยู่ตรงช่องลม ห้องไม่ร้อน ห้อง 7000/เดือน (ช่วงโควิดเหลือ5500) ค่าสัญญา16000 (ได้คืนตอนออก) ค่าเปลี่ยนชื่อสัญญา 1000 เข้าอยู่ได้เลย !ซื้อภายในเดือนนี้เดือนเมษาอยู่ฟรี! สนใจdm/line #หอมธ #หอนอกมธ #สัญญาหอมธ #cityparkมธ #ทิมมธ</p> |
| “ทัก” | <p>- ใครเรียน มธ. แล้วได้ค่าเทอม 10% คืนแล้วบ้างคะพอดีเราต้องการข้อมูลไปทำรายงานคะ รบกวนทักเดมหรือเมนขึ้นมาหน่อยนะคะ #มธ #ธรรมศาสตร์</p> <p>- หาเมทหออินเตอร์ปาร์ค ฝั่งหญิงค่า ตอนนี้มีห้องแล้ว ขอคนที่รับได้กับการตื่นเช้าและนอนเช้าในบางช่วง ปล. ทักมาคุยกรูกริบกันก่อนได้นะ ทาง DM หรือไลน์ madelineemito #หาเมทมธ #หอนอกมธ #ทิมมธ #หอพักมธ</p> <p>- ขายสัญญาหอ The Charn ค่ะ สนใจทักเดมหรือแอดไลน์นะคะ #หอนอกมธ #หอในมธ #หอนอกมธรังสิต #ทิมมธ #มธรังสิต #Dek64 3</p> |
| “หา” | <p>- ไหนๆ ก็พูดถึงแม่มดแล้ว nerd ประวัติศาสตร์ ขอวนไปหาหนังสือ nonfiction ของสำนักพิมพ์ ยิปซีหน่อย - แม่มด ประวัติศาสตร์แห่งสายเวทย์ 230฿ - เทพกรนัมแห่งฟินแลนด์ 260฿ - กองกำลังเอสเอสหน่วยพิฆาตแห่งนาซี 265฿ - หญิงร้าย หนังสือรางวัลวิทยานิพนธ์ตีมาจาก ธรรมศาสตร์ ปี2559 265฿</p> <p>- ข้อสอบจำลองสำหรับฝึกทำโค้งสุดท้ายก่อนสอบข้อสอบรวม3สาขาพร้อมข่าว เก่งทั้งเนื้อหาและข่าว เสมือนข้อสอบจริงพร้อมกระดาษคำตอบ ไม่ซื้อไม่ได้แล้วชาวว่าที่สิงห์แดง #รัฐศาสตร์มธ #รัฐศาสตร์ #รัฐศาสตร์จุฬา #Dek64 #TCAS64 #ธรรมศาสตร์ #9วิชาสามัญ64 ชื่อซีทเลย</p> <p>- สวัสดีครับ ขออนุญาตทุกคนนะครับ พวกเราจาก “Review หอนอกมธ” ตอนนี้พวกเรากำลังรวบรวม หาข้อมูลเกี่ยวกับหอพักรอบมหาวิทยาลัยธรรมศาสตร์ เพื่อที่จะทำเป็นอีกหนึ่งช่องทางในการรีวิวหอพักรอบมหาวิทยาลัยของเรา โดยเป็นส่วนหนึ่งของรายวิชา TU100 #มธ #หอมธ #หอในมธ</p> |

ตารางที่ 4.7

ตัวอย่างข้อความกลุ่มที่ 3 ของการจัดกลุ่มข้อความที่ไม่ได้ใช้การแปลงกลุ่มของชื่อเฉพาะ(ต่อ)

| คำที่มี ความถี่สูง | ตัวอย่างข้อความ |
|-----------------------|--|
| “แจก” | <p>- บนค้ำบรีได้เลยค้ำบ GAT 150+แจก 50 บาท PAT1 เกิน 90 แจก 50 บาท PAT2 125+แจก 50 บาท ติดเทคนิคการแพทย์ มหิดล แจก 100บาท ติดสทเวช มธ แจก 100 บาท ติดเทคนิคการแพทย์ มโหนกก็ได้ 50 บาท #Dek64 รีได้เลยค้ำบ</p> <p>- บน ถ้าติด นิติ มธ.รังสิต รอบ3 แจก 1000 ถ้าติด นิติ มธ.ลำปาง รอบ3 แจก 500 ถ้าติด นิติ มข. รอบ3 แจก 100 ถ้าติด นิติ มธ. ไม่ว่าที่ไหนจะโอนเงินให้มูลนิธิเกี่ยวกับน้องหมาเดือนละ 100 นาน 1ปี ร้อย่างเดียว #Dek64</p> <p>- หนูบน!!!! ถ้าสัมภาษณ์ ณ์ศิลปศาสตร์ มธ.ผ่าน หลังประกาศผล จะแจก 100 1 คน หนูงบน้อย แต่อยากแจก แงง ขอให้ติดๆๆ เพ็ยง!!!!!! #Dek64</p> |
| “ยื่น” | <p>- คนที่จะยื่นเทคนิคการแพทย์ มธ รอบ3.2 คิดคะแนนแล้วได้เท่าไรกันบ้างคะ เต็ม21,000นะคะ เนื่องจากonetยังไม่ออก เดียวจะแปะวิธีคิดไว้ในเม้นนะคะ #dek64</p> <p>- ใครจะยื่นบัญชี มธ รอบadmission2 ได้คะแนน gat pat1 เท่าไรกันบ้างคะ #dek64 #บัญชีมธ</p> <p>- รวมเกณฑ์คะแนนขั้นต่ำวิชาสามัญของจุฬาฯ / มธ. / มหิดล TCAS64 รอบที่ 3 (Admission 1) น้องๆ ทีม #dek64 ตามไปดูข้อมูลเกณฑ์คะแนนขั้นต่ำวิชาสามัญเพื่อใช้ยื่นใน TCAS รอบที่ 3 กัน ในบทความนี้เลย #TCAS64 #dek64</p> |
| “สอบ” | <p>- บนครับ ถ้า #สอบตรงนิติมธ ติดรอบ 2 ที่ธรรมศาสตร์ลำปาง แจก 500 บาท ร้อย่างเดียวครับ เดี่ยววันประกาศผลจะมาแจ้งอีกทีครับ #DEK64 #TCAS64 #ธรรมศาสตร์</p> <p>- *บน* ติดรัฐศาสตร์ มธ. การเมืองการปกครอง(รอบ3.1 สอบวิชาเฉพาะ) แจกบรวานี้ 1ปอนด์ครับ</p> <p>- โค้งสุดท้าย! อยากสรุปเป็นประเด็นสำคัญในรอบปีที่ผ่านมา ว่าเรื่องไหนมีแนวโน้มออกข้อสอบสูง โดยเฉพาะเรียงความ ลองเช็กกันครับ ว่ายังมีประเด็นไหนที่อาจจะยังไม่เข้าใจ ควรรีบทบทวนเลย อดทนอีกไม่กี่วันแล้วนะ ลู้ #รัฐศาสตร์ธรรมศาสตร์ #รัฐศาสตร์ #รัฐศาสตร์มธ #มธ #สิงห์แดง #สอบตรงรัฐศาสตร์</p> |
| “ปี” | <p>- บนจ้า ถ้าติด มธ. • แจกเงิน 100 บาท 1 คน • ดกกินของทอด 1 ปี ร้อย่างเดียว ไม่ต้องพอลจ้า</p> <p>- #dek64 แจกเพิ่มคะถ้าติดรัฐศาสตร์ มธ. สาขาบริหารรัฐกิจภายในปีนี้ แจก200 จำนวน2 คนคะ</p> <p>- ขออนุญาตบนคะ ดินนิติศาสตร์ ธรรมศาสตร์รังสิต แจก300 รีเรยๆ ไม่ตัดผม1ปี ไปสักลายมีเสื้อที่ ออกแบบเอง</p> |
| “Pat” | <p>- บน ติด บัญชี มธ.(ทำพระจันทร์) แจก500 1คน ติด บัญชี เกษตร(บางเขน) แจก300 1คน #Dek64 #gatpat64 #9วิชาสามัญ64</p> <p>- บน ถ้าได้คะแนน รีไว้เรย -gat 200+ -pat 7.4 200+ -9วิ รวมกัน3 วิชามากกว่า 180onet ทุกวิชา(ยกเว้น เลข วิท)) ได้มากกว่า 60 -ได้โควต้าเกษตรรอบ2แจกเพิ่มอีก50 -รอบ3.1 ได้เกษตร มธ จะเลี้ยงซาบู่1คน299 จากรีทวิตถ้าสู้ไม่ได้ละไม่ยากกิน จาให้เงินแทนค้ำ 50 #Dek64</p> |

ตารางที่ 4.8

ตัวอย่างข้อความกลุ่มที่ 4 ของการจัดกลุ่มข้อความที่ไม่ได้ใช้การแปลงกลุ่มของชื่อเฉพาะ

| คำที่มี ความถี่สูง | ตัวอย่างข้อความ |
|-----------------------|--|
| Covid-19 | <p>- ใครที่เข้าสอบกสพท ในวันที่10 มธ. รังสิต ห้อง sc3053 ชั้น 3 กักตัวด้วยนะคะ เราสอบห้องนั้น จิตตกมากๆ จะติดโควิดก่อนมหาลัยคือเรื่องจริง อีสัส นร.บอกคอกจะแตกให้เลื่อนสอบ ไม่เลื่อนให้ ทิ้งในศักดิ์ศรีเป็นไงละ อีเมลแจ้งพวกกูมาแค่นี้ #กสพท64 #dek64</p> <p>- สวัสดิ์คะเราได้ไปตรวจโควิดมาเมื่อวาน และวันนี้ทางรพ.แจ้งมาว่าผลเราเป็นบวกคะและก่อนหน้า นี้เราไม่ได้มีอาการอะไรเลย ซึ่งเมื่อวันที่ 9 เม.ย. เวลา 10.00-12.00 เราได้ไปสอบbir ที่ติกรัฐศาสตร์ (ห้อง404) มธ.ท่าพระจันทร์ (ไปจากอยุธยา) #โควิดวันนี้ #รัฐศาสตร์มธ</p> <p>- คณะรัฐศาสตร์ ม.ธรรมศาสตร์ แจ้ง มีผู้ติด Covid-19 จำนวน 1 ราย เข้าสอบ #วิชาเฉพาะ เมื่อวันที่ 9 และ 10 เม.ย. ที่ผ่านมา #dek64</p> |
| รังสิต | <p>- เอาจริงตอนตั้งศูนย์อุกเลนสินามิ ที่ธรรมศาสตร์ รังสิต คือทุกอย่างสามารถเสกออกมาได้ภายใน 3 ชั่วโมงเท่านั้น มีระบบรับลงทะเบียน ระบบจัดการห้องพัก ระบบรับบริจาค ระบบสวัสดิการทั้ง อาสาสมัครและผู้ประสบภัย + ประสานงานหน่วยงานราชการต่างๆ แล้วทุกเย็นก็มาประเมินกันว่า พรุ่งนี้ต้องปรับอะไรอีก</p> <p>- เพื่อนนักศึกษาไม่ได้รับความเป็นธรรม มวลชนที่นั่งอดอาหารหน้าศาลถูกจับ พระที่ออกมาเรียกร้องถูก คุกคาม . “ ก้าวอะไรแค้กับคนที่ออกมายื่นเฉยๆ!? ถึงเวลาแล้วที่เราจะออกมาแสดงพลัง! ขอเชิญ ร่วมกันยื่น หยุด ชัง 112 นาที ณ เชียงรอก 1 ธรรมศาสตร์ ศูนย์รังสิต เวลา 17:00 น. เป็นต้นไป</p> <p>- เริ่มแล้ว! กิจกรรมยื่น หยุด ชัง ณ บริเวณประตูเชียงรอก1 ธรรมศาสตร์ ศูนย์รังสิต สามารถเข้าร่วม ได้ตลอดเวลากิจกรรม โดยกิจกรรมจะจบลงเวลา 18:12 น. มาร่วมเรียกร้องความยุติธรรมคืนสู่ ประเทศได้แล้ว ณ ตอนนี! #ยื่นหยุดชัง #ปล่อยเพื่อนเรา</p> |
| เมษายน | <p>- แจ้งข่าวด่วน- มีคนโดน 112 อีกแล้ว พรุ่งนี้ 28 เม.ย. 64 ลัลนาจะต้องเข้าฟังการพิจารณาสั่งฟ้อง คดีตามม.112 กรณีปลดรูปหน้า มธ.ลำปาง โดยหลักฐานเป็นคลิปที่ไม่สามารถยืนยันตัวผู้กระทำได้ และเบื้องต้นลัลนาปฏิเสธทั้งหมด แน่นอนว่า 112 เสี่ยงฝากขังโดยไม่ให้ประกันสูง อย่าให้เจียบ . #Saveลัลนา</p> <p>- 6 เมษายนนี้ เปิดการรณรงค์เข้าชื่อทั่วประเทศ เริ่มลงชื่อครั้งแรกที่ ห้อง LT1 คณะนิติศาสตร์ ม. ธรรมศาสตร์ ท่าพระจันทร์ ขอเชิญทุกคนมาร่วมลงชื่อด้วยกัน ใช้เพียงบัตรประชาชนเท่านั้น . ยังมี รายชื่อมากเท่าไร่ ประตุของความเป็นไปได้ย่อมเปิดกว้างมากขึ้นเท่านั้น #ขอคนละชื่อรื้อระบอบ ประยุทธ์</p> <p>- พรุ่งนี้ (2 เม.ย.) มีเปิดตัว กลุ่ม #OCTDEM นำทีมโดยอดีตคนเดือนตุลา อาทิ -เกรียงกมล เลหา ไพโรจน์ -จาตุรนต์ ฉายแสง -ธเนศ อาภรณ์สุวรรณ -สมศักดิ์ ปริศนานันทนทุกุล -นพ.สุรพงษ์ สืบวงศ์ ลี ยื่นจม.ประธานศาลฎีกา เรียกร้องให้ประกันผู้ต้องหาการเมือง 13.00น. ลานประติมากรรม มธ. ท่าพระจันทร์</p> |

ตารางที่ 4.8

ตัวอย่างข้อความกลุ่มที่ 4 ของการจัดกลุ่มข้อความที่ไม่ได้ใช้การแปลงกลุ่มของชื่อเฉพาะ(ต่อ)

| คำที่มี ความถี่สูง | ตัวอย่างข้อความ |
|-----------------------|---|
| โรงพยาบาล สนาม | <p>- รพ.ธรรมศาสตร์ ฝากประชาสัมพันธ์ครับ ตอนนี้ทาง รพ.กำลังเตรียม รพ.สนามเพื่อรองรับผู้ป่วยโควิด ใครที่ประสงค์จะบริจาคของ ตอนนี้ทาง รพ.ต้องการสบู่ล้างมือ ไม้กวาด ยาแก้นุง ไม้ช้อนตุงครับ สามารถส่งทางไปรษณีย์ หรือไปส่งที่อาคาร มรว. สุวพรรณ (P2 ในแผนที่) ขอบคุณครับ โรงพยาบาลสนาม</p> <p>- ขอเชิญ #nontfam #nonttanont ร่วมสมทบทุนซื้อวัตถุดิบและเครื่องต้ม ให้กับศูนย์ฝึกประสบการณ์วิชาชีพ มรภ.วไลยอลงกรณ์ฯ เพื่อประกอบอาหารส่งมอบให้กับโรงพยาบาลสนาม ม.ธรรมศาสตร์ ที่บัญชี ธ.กสิกรไทย 744-2-73572-1 (ณรี สอนสา) 24 เม.ย.-6 พ.ค. 64</p> <p>- รพ.สนาม มธ ได้ใช้แล้วอะ คือเปิดมาตั้งแต่ระลอกแรกเพิ่งจะได้ใช้วันนี้วันแรก เนี่ยที่บอกว่ารอบนี้ น่ากลัว</p> |
| นักศึกษา | <p>- ข่าวด่วน! นศ.ธรรมศาสตร์ปี 1 - นิเว สิริชัย- ถูกอัยการสั่งฟ้อง ม.112 อย่างเร่งด่วนเช้านี้ จากการพนันสื่อความยกเลิก 112 หากศาลไม่ให้ประกัน นิเวจะถูกส่งเข้าเรือนจำทันที - ขณะนี้ทนายและอาจารย์กำลังยื่นประกันตัวอยู่ที่ศาลอัยการ #ยกเลิก112</p> <p>- เมษา 64 - อัยการสั่งฟ้องคดี 112 - พ.ร.บ.คอมฯ ‘นิเว’ สิริชัย นศ.มธ. #แนวร่วมธรรมศาสตร์และการชุมนุม เหตุพนันสื่อปราทวงค์ #ยกเลิก112 - ไม่ให้รหัสเครื่องมือสื่อสารกับ จนท. บก.ปอท. ก่อนได้ประกัน วางเงินแสนห้าสำหรับคดี 112 และห้าหมื่นสำหรับ พรบ.คอมฯ อ่านคำฟ้อง</p> <p>- วันนี้ที่สนง.อัยการจังหวัดธัญบุรี ญัฐชนน ไพโรจน์ นศ. มธ. ที่ถูกแจ้งข้อกล่าวหา #ม112 จากกรณีหนังสือ #ปกแดง “ปรากฏการณ์สะพานฟ้า” มาส่งตัวให้พนักงานอัยการพร้อมสำนวนคดี โดยอัยการได้นัดฟังคำสั่งฟ้องในวันที่ 30 เม.ย. 64 เวลา 10.30 น.</p> |
| ตรวจ | <p>- ใครอยู่แถวเชียงราก มธ.รังสิตเพื่อนเราติดโควิดจากพี่สาวแต่ไม่แสดงอาการอะไรเลยไปตรวจถึงรู้ว่า เป็น ได้ยินข่าวมาว่าวันที่20นี้จะมีการตรวจเชิงรุกที่ตึกsc มีคนสนใจไปตรวจด้วยกันไหมคะ;—;</p> <p>- ใครอยู่โซนรังสิต สามารถไปตรวจโควิดฟรีได้ วันที่ 19 เมษายน เวลา 8.00-15.00 ที่ plum condo ม.กรุงเทพ นะคะ ตรวจแบบ swab หรือถ้าเป็นที่โรงพยาบาลธรรมศาสตร์สามารถรอกฟอร์มได้นะคะ ถ้ามีอาการหรือเสี่ยงมากๆ เจ้าหน้าที่จะติดต่อมา เป็นตรวจแบบ swab เหมือนกันค่ะ #โควิด #โควิด19 #ตรวจโควิดฟรี</p> <p>- ฝากรื้อทวิตหน่อยนะคะ สำหรับชาวรังสิตหรือคนพื้นที่ใกล้เคียง จะมีการตรวจโควิดเชิงรุกฟรีนะคะ วันที่19เมษายน ตั้งแต่ 08:00น. เป็นต้นไป มีการตรวจโควิดฟรีที่ พลัมคอนโด รังสิต (ข้างม.กรุงเทพ) วันที่20เมษายน ตั้งแต่08:00น. ที่มหาวิทยาลัยธรรมศาสตร์ ศูนย์รังสิต</p> |

ตารางที่ 4.8

ตัวอย่างข้อความกลุ่มที่ 4 ของการจัดกลุ่มข้อความที่ไม่ได้ใช้การแปลงกลุ่มของชื่อเฉพาะ(ต่อ)

| คำที่มี ความถี่สูง | ตัวอย่างข้อความ |
|-----------------------|--|
| เตียง | <p>- มธ.พนัก รพ.ธรรมศาสตร์เฉลิมพระเกียรติ ประกาศเปิด รพ.สนาม รอบใหม่เพิ่มความจุจาก 308 เตียงเป็น 470 เตียง พร้อมรับคนไข้ 11 เม.ย.นี้ #ThaiPBSnews #COVID19 #มธ #โรงพยาบาลสนาม</p> <p>- โรงพยาบาลสนามธรรมศาสตร์ 470 เตียง พร้อมแล้วในการรับส่งต่อผู้ป่วยโควิดจากทุกโรงพยาบาลในเช้าวันอาทิตย์ที่ 11 เมษายน 64</p> <p>- ฟากเรื่องเตียงด้วยนะคะทุกคน มธยังพอรับได้แต่คิดว่าเหลือไม่มากเพราะทางรพต่างๆเริ่มจองส่งต่อผู้ป่วยมาที่นี้แล้วนะคะ ติดต่อเพจ Thammasat well being centerนะคะ #ClubLOVEIS มธ. เตรียมเปิด โรงพยาบาลสนาม 470 เตียง เจาะลึกทั่วไทย 09 เม.ย. 64 via</p> |
| ศูนย์ | <p>- “เมื่อกฎหมายกลายเป็นเครื่องมือปิดปาก พวกเราจึงต้องลุกขึ้นมา ไม่ยอมรับมัน!” . กิจกรรม ยืนหยุด ชั่ง 1 ชั่วโมง 12 นาที ในวันนี้เริ่มขึ้นแล้ว! โดยเราจะทำกิจกรรมถึงเวลา 18:12 น. . ท่านใดที่สนใจยังสามารถมาเข้าร่วมได้ทันที ณ บริเวณเชียงราก1 ธรรมศาสตร์ ศูนย์รังสิต . #ยืนหยุดชั่ง</p> <p>- “พอรัฐบาลสั่งปิดร้านเพราะโควิดโดยไม่มีแผนบริหารจัดการ อาชีพช่างทำผมไม่มีเงินกินข้าวเลย” มานพ วิเศษโวหาร ตัวแทนสมาคมช่างผม-เสริมสวยแห่งประเทศไทย จากเสวนา "ผลกระทบของโควิด-19 ต่อธุรกิจภาคบริการ” 2 เม.ย.64 โดยศูนย์ศึกษาสังคมและวัฒนธรรมร่วมสมัย คณะสังคมวิทยาและมานุษยวิทยา มธ.</p> <p>- อัยการยื่นยันนัดส่งฟ้อง 2 นศ. มธ.ศูนย์ลำปาง ลัลนา และ “เบนซ์” (นามสมมติ) คดี #ม112 “ปลดรูป” ในการชุมนุมหน้ามหาลัยของนศ.ช่วงเดือนตุลาคม 63 จากก่อนนี้แจ้งเพียงข้อหา #ม360 แม้ทางทนายความจะได้แจ้งเรื่องสถานการณ์โควิด-19 อาจสร้างความลำบากไปแล้วก่อนหน้านี้</p> |
| เวลา | <p>- [อัปเดต]เพนกวินอดอาหารเป็นเวลานานกว่า 21 วันแล้ว ซึ่งคานธีก็อดอาหารประท้วงต่อความอยุติธรรมในสังคมเป็นเวลา 21 วันเช่นเดียวกัน เราขอฝากบทความเรื่อง "สันติวิธี อดอาหารประท้วง" เขียนโดย ชัยวัฒน์ สถาอานันท์ อาจารย์คณะรัฐศาสตร์ ธรรมศาสตร์ และเพนกวินยังฝากแถลงการณ์อดอาหารออกมาด้วย</p> <p>- [อัปเดต] รุ่งอดอาหารเป็นเวลานาน 8 วันแล้ว เราขอฝากบทความเรื่อง "สันติวิธี อดอาหารประท้วง" เขียนโดย ชัยวัฒน์ สถาอานันท์ อาจารย์คณะรัฐศาสตร์ มหาวิทยาลัยธรรมศาสตร์ เว็บ : #อดอาหารประท้วง</p> <p>- [กรุณาแชร์โพสนี้ต่อ] . จัดไปอย่าหยุด! ขอเชิญเพื่อน ๆ ร่วม ยืน หยุด ชั่ง 1 ชั่วโมง 12 นาที กับพวกเรา ณ เชียงราก 1 ธรรมศาสตร์ รังสิต เวลา 16:30 น. เป็นต้นไป! . *กรุณาเว้นระยะห่างและสวมหน้ากากอนามัยตลอดเวลาทำกิจกรรมเพื่อป้องกันการแพร่ระบาดของโรค* #ยืนหยุดชั่ง #ปล่อยเพื่อนเรา</p> |

ตารางที่ 4.8

ตัวอย่างข้อความกลุ่มที่ 4 ของการจัดกลุ่มข้อความที่ไม่ได้ใช้การแปลงกลุ่มของชื่อเฉพาะ(ต่อ)

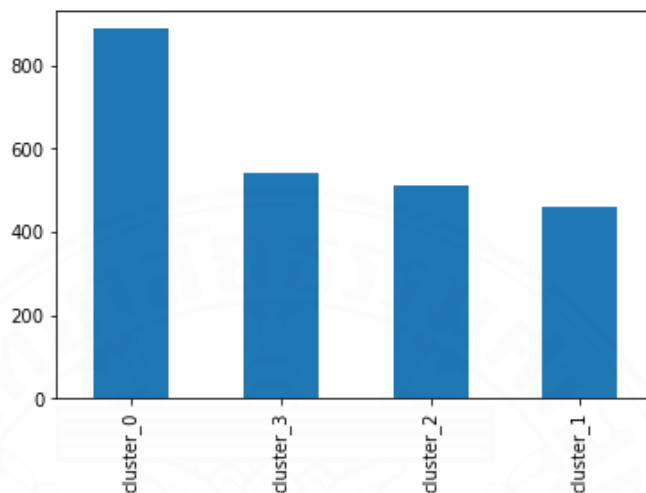
| คำที่มี ความถี่สูง | ตัวอย่างข้อความ |
|-----------------------|--|
| สอบ | <p>- ต่อมาวันที่ 10 เม.ย. เราก็ไปสอบข้อสอบเฉพาะรัฐศาสตร์ เวลา 8.30-11.30 ที่ มธ. รังสิต ตึกบร.1 ห้อง201 เราต้องขอโทษทุกคนด้วยจริงๆนะคะ ขอโทษที่ทำให้เดือดร้อน ขอโทษจริงๆค่ะ</p> <p>- คณะรัฐศาสตร์ ม.ธรรมศาสตร์ เน้นย้ำ จัดสอบ #วิชาเฉพาะ คณะรัฐศาสตร์ วันที่ 10 เม.ย. 08.30-11.30 น. - อาคารวิศิษฐ์อักษร (SC2) - อาคารบรรยายรวม 1 (บร.1) - อาคารกลุ่มสังคมศาสตร์ (คณะรัฐศาสตร์/คณะนิติศาสตร์/คณะสังคมสงเคราะห์ศาสตร์) อย่าลืม! บัตรประชาชน/บัตรนักเรียน + สวมหน้ากากอนามัย</p> <p>- คณะรัฐศาสตร์ ม.ธรรมศาสตร์ ขอความร่วมมือ ผู้มีสิทธิ์สอบ #วิชาเฉพาะ คณะรัฐศาสตร์(วันที่ 10 เม.ย.) ทั้ง 1,819 คน กรอกข้อมูลใน"แบบฟอร์มคัดกรองผู้เข้าสอบ" สแกน QR CODE หรือ ภายในวันที่ 8 เม.ย. เวลา 12.00 น. #dek64</p> |

จากภาพที่ 4.5, 4.6, 4.7 และ 4.8 แสดงให้เห็นถึงคำที่ปรากฏบ่อยในกลุ่มข้อความ ตัวอย่างเช่น ข้อความกลุ่มที่ 1 คำที่ปรากฏบ่อยคือ “เรียน”, “พี่”, “ทำ”, “สอบ” “ปี”, “เพื่อน”, “รู้”, “คณะ”, “ดี” และ “จบ” ข้อความกลุ่มที่ 2 คำที่ปรากฏบ่อยคือ “ค่า”, “ขาย”, “#Dek64”, “บาท”, “ราคา”, “สนใจ”, “ห้อง”, “#หอนอกมธ”, “หัก” และ “หา” ข้อความกลุ่มที่ 3 คำที่ปรากฏบ่อยคือ “ติด”, “#Dek64”, “คะแนน”, “รอบ”, “แจก”, “ยื่น”, “สอบ”, “ปี”, “Pat” และ “ไหม” และ ข้อความกลุ่มที่ 4 คำที่ปรากฏบ่อยคือ “Covid-19”, “รังสิต”, “เมษายน”, “โรงพยาบาลสนาม”, “นักศึกษา”, “ตรวจ”, “เตียง”, “ศูนย์”, “เวลา” และ “สอบ” เมื่อนำข้อความตัวอย่างมาพิจารณาแล้วพบว่าข้อความในแต่ละกลุ่มมีความหมายไปในทิศทางเดียวกัน แต่ยังมีคำบางคำที่ไม่สื่อถึงข้อความในกลุ่มนั้นๆ ได้

4.3.2 ผลลัพธ์การจัดกลุ่มข้อความที่ใช้การแปลงกลุ่มของชื่อเฉพาะ โดยอัลกอริทึม k-means ซึ่งกำหนด $k = 4$ และวัดค่าความคล้ายคลึงแบบ Cosine Similarity พบว่ากลุ่มข้อความที่ 1 มีข้อความ 848 ข้อความ กลุ่มข้อความที่ 2 มีข้อความ 555 ข้อความ กลุ่มข้อความที่ 3 มีข้อความ 536 ข้อความ และ กลุ่มข้อความที่ 4 มีข้อความ 462 ข้อความ ดังภาพที่ 4.8

ภาพที่ 4.8

จำนวนข้อมูลของการจัดกลุ่มข้อความที่ไม่ได้ใช้การแปลงกลุ่มของชื่อเฉพาะ และวัดค่าความคล้ายคลึงแบบ Cosine Similarity



จากข้อความกลุ่มที่ 1 จำนวน 848 ข้อความ พบคำที่มีความถี่สูงสุดได้แก่ “เรียน”, “พี่”, “ทำ”, “สอบ” “ปี”, “เพื่อน”, “รู้”, “ซุ่มทาลัย”, “ชื่อคณะ” และ “จบ” ตามลำดับ ซึ่งนำคำที่มีความถี่สูงที่สุดมาแสดงได้ดัง ภาพที่ 4.9

ภาพที่ 4.9

Word Cloud ข้อความกลุ่มที่ 1 ของการจัดกลุ่มข้อความที่ใช้การแปลงกลุ่มของชื่อเฉพาะ และวัดค่าความคล้ายคลึงแบบ Cosine Similarity



ตารางที่ 4.9

ตัวอย่างข้อความกลุ่มที่ 1 ของการจัดกลุ่มข้อความที่ใช้การแปลงกลุ่มของชื่อเฉพาะ

| คำที่มีความถี่สูง | ตัวอย่างข้อความ |
|----------------------------|--|
| “ชื่อมหาลัย” | <p>- มธ ก็ไม่ต่างค่ะ ควรจะตรวจทั้งมอ คนไปก้าติดโควิดแล้วคุณที่ดูคนไปก้าวันนั้นสิคะ แทบชื่ออ กันกิน เยอะกว่าคนไปงานรับเพื่อนใหม่อีกมั้ง</p> <p>- จำวันนี้ไว้จนตายเลยอะ วันที่ต้องไปเข้าโรงแรมอยู่เพราะหอนมันจะปิด แล้วต้องออกไปฝึกงาน ไปเรียนไปสอบ ถ้าลูกอยากเข้าจุฬาจะขอให้ไปธรรมศาสตร์ เขาไม่มีวันทำกับหนูแบบนี้ ผู้บริหารต่าง mindset ต่าง #ไม่เอาคุณหอน #หอนในจุฬา #จุฬา #chula</p> <p>- องค์กรต่างๆควรตรวจเช็คให้ดีกว่าก่อนรับสมัครงาน ว่าคนที่จบใหม่มาจาก จุฬา มธ นอนเคารพธงชาติรีไม่ เพราะคนพวกนี้ ไม่รักแม่แต่ชาติที่ตัวเองซุกหัวนอน ก็อย่าหวังว่าเค้าจะรักองค์กรของคุณ</p> |
| “ชื่อคณะ” | <p>- อีกเรื่องหนึ่งที่ซ้อค ถ้าชั้นติดเตรียมอุดมศิลป์ภาษา มัณฑารัตติอักษรจุฬา นิเทศจุฬา ศิลปศาสตร์ มธ ไปเกินครึ่งละอะ แต่น้องแพทริคลาออกจากเตรียมอุดมตอน มอ4 ไปสอบ GED เพื่อมาแข่ง ชวว ใจเด็ดซิบหายเลยชั้นนับถือมาก 555555555555 #CHUANG2021 #PatrickRank13th #แพทริค</p> <p>- แม่นิดปังงไม่ไหวว จบนิดิ มธ.เกียรตินิยมอันดับ2 ใครจะไฟท์กับแม่ คิดดีๆนะน้องนะ ถึงพี่ โน้ตจะไม่ได้ทำงานตามสายที่จบมา แต่พี่โน้ตมีความสุขกับงานแต่งหน้าที่ทำมาหลาย แต่งหน้าให้เจ้าสาวแต่ละคนสวยสวย แต่งหน้าฟลุ้คก็สวย เวลาแต่งหน้าฟลุ้คโหนหวานๆ อ้อหือ เหมือนฟลุ้คจะออกเรื่อน</p> <p>- ผมเรียนที่คณะศิลปศาสตร์ มธ ครับ เป็นอาจารย์จากสาขาสเปนและลาตินอเมริกันศึกษา เสียชีวิตที่ผมแก่เกินไป เลยไปลงเรียน ป ตรี สาขานี้ก็ไปไม่ได้แล้ว</p> |
| “ชื่อสถานที่ในมหาวิทยาลัย” | <p>- 6 ปีที่แล้ววันนี้ ไกล่สงกรานต์ 2015 โปรดสังเกต อ.ยิ้ม ยังอยู่กับปัญญาชนประชาธิปไตย ทว่าถูกรองอธิการเฮงชวยคนหนึ่งขับไล่ไสส่ง ไม่ให้จัดงานภายในรั้ว มธ ทำพระจันทร์ ทั้งทั้งที่อาจารย์สัญญา ธรรมศักดิ์ เคยกล่าวไว้ว่า "มีเสรีภาพทุกรางนี้" เมื่อปี 1973 (ภาพ:สงวนคุ้มครองโรจน์)</p> <p>- อ้อ ใ้ทำพระอาทิตย์55555 คือผมขึ้นต้นมาด้วยธรรมศาสตร์ก็ถูกทางแล้วแหละ พวดอบทำพระจันทร์ก็ถูกทาง แต่พอผมบอก ไม่ใช่อีกอันนึ่ง พี่ว่าบอก รังสิต อ้ออ ก็ถูกแต่มีคนละเรื่องเดียวกัน ตลึกกกกก55555 #HollywoodGNxMEW</p> <p>- ร้านเปิดทำพระจันทร์ที่ทุกคนรู้จักนั้น ตามตำนานเล่าขานว่า เด็กธรรมศาสตร์คนไหนไม่ได้กินคือไม่จบนะคะ คิดว่าพี่เจ๋งคงไปกินมาหลายจานแล้ว #เจ๋งพิท พล.เมื่อก่อนขอหาของกินและซ้อปแถวนั้นนะ มีหมอดูแม่นด้วย555</p> |

ตารางที่ 4.11

ตัวอย่างข้อความกลุ่มที่ 3 ของการจัดกลุ่มข้อความที่ใช้การแปลงกลุ่มของชื่อเฉพาะ(ต่อ)

| คำที่มีความถี่สูง | ตัวอย่างข้อความ |
|--|--|
| “ชื่อพื้นที่” | <ul style="list-style-type: none"> - เอาจริงตอนตั้งศูนย์ฉุกเฉินสึนามิ ที่ธรรมศาสตร์ รังสิต คือทุกอย่างสามารถเสกออกมาได้ ภายใน 3 ชั่วโมงเท่านั้น มีระบบรับลงทะเบียน ระบบจัดการห้องพัก ระบบรับบริจาค ระบบสวัสดิการทั้งอาสาสมัครและผู้ประสบภัย + ประสานงานหน่วยงานราชการต่างๆ แล้วทุกเย็นก็มาประเมินกันว่าพรุ่งนี้ต้องปรับอะไรอีก - เพื่อนนักสู้ไม่ได้รับความเป็นธรรม มวลชนที่นั่งอดอาหารหน้าศาลถูกจับ พระที่ออกมาเรียกร้องถูกคุกคาม . “ กว่าจะไรแค้นกับคนที่ออกมายื่นเฉยๆ !? “. ถึงเวลาแล้วที่เราจะออกมาแสดงพลัง! ขอเชิญร่วมกันยื่น หยุด ชัง 112 นาที ณ เชียงราก 1 ธรรมศาสตร์ ศูนย์ - เริ่มแล้ว! กิจกรรมยื่น หยุด ชัง ณ บริเวณประตูเชียงราก1 ธรรมศาสตร์ ศูนย์รังสิต สามารถเข้าร่วมได้ตลอดเวลากิจกรรม โดยกิจกรรมจะจบลงเวลา 18:12 น. มาร่วมเรียกร้องความยุติธรรมคืนสู่ประเทศได้แล้ว ณ ตอนนี! #ยื่นหยุดชัง #ปล่อยเพื่อนเรา |
| “ชื่อบุคคล/กลุ่มบุคคลที่เกี่ยวข้องกับการเมือง” | <ul style="list-style-type: none"> - เชิญไปกันเยอะๆ ให้ประยุทธ์หนาว ประชาเราอุ่นใจค่ะ ที่นิติศาสตร์ มธ. ทำพระจันทร์ งานนี้ดูจากคนพูดแล้ว ไฟลุกแน่นอน ติดตามรายละเอียดงานได้ที่ นะคะ - พรุ่งนี้ (2 เม.ย.) มีเปิดตัว กลุ่ม #OCTDEM นำทีมโดยอดีตคนเดือนตุลา อาทิ -เกรียงกมล เลขาไฟโรจน์ -จาตุรนต์ ฉายแสง -ธนศ อภรณ์สุวรรณ -สมศักดิ์ ปริศนานันทกุล -นพ.สุรพงษ์ สืบวงศ์ลี ยื่นจม.ประธานศาลฎีกา เรียกร้องให้ประกันผู้ต้องหาการเมือง - แนวร่วมธรรมศาสตร์ "นอน กิน เรียนหน้าศาลฎีกา" |
| “ชื่อโรงพยาบาล” | <ul style="list-style-type: none"> - รพ.ธรรมศาสตร์ ฝากประชาสัมพันธ์ครับ ตอนนีทาง รพ.กำลังเตรียม รพ.สนามเพื่อรองรับผู้ป่วยโควิด ใครที่ประสงค์จะบริจาคของ ตอนนีทาง รพ.ต้องการสบู่ล้างมือ ไม้กวาด ยาแก้นิยุง ไม้ขีดตุงครับ สามารถส่งทางไปรษณีย์ หรือไปส่งที่อาคาร มรว. สุพรรณณ (P2 ในแผนที่) - วิกฤตหนัก!! โรงพยาบาลธรรมศาสตร์เฉลิมพระเกียรติ เกิดสถานการณ์วิกฤต “โลหิตขาดแคลนทุกรูป” ขอเชิญทุกท่านร่วมบริจาคโลหิต ได้ที่ หน่วยรับบริจาคโลหิต ณ อาคารพลังใจ - #วิกฤตโลหิตขาดแคลนทุกรูป ขณะนี้โรงพยาบาลธรรมศาสตร์เฉลิมพระเกียรติ เกิดสถานการณ์วิกฤตโลหิตขาดแคลนทุกรูป จึงขอเชิญร่วมบริจาคโลหิตเพื่อต่อชีวิตและลมหายใจให้กับผู้ป่วย ทุกวันจันทร์-อาทิตย์ เวลา 08.30-15.30 (ไม่ปิดพักเที่ยง) ณ ชั้น 2 อาคารพลังใจ |
| “ชื่อสถานที่ในมหาวิทยาลัย” | <ul style="list-style-type: none"> - 6 เมษายนนี้ เปิดการณรงค์เข้าชื่อทั่วประเทศ เริ่มลงชื่อครั้งที่ ห้อง LT1 คณะนิติศาสตร์ ม.ธรรมศาสตร์ ทำพระจันทร์ ขอเชิญทุกคนมาร่วมลงชื่อด้วยกัน ใช้เพียงบัตรประชาชนเท่านั้น . ยังมีรายชื่อมากเท่าไร ประตุของความเป็นไปได้ย่อมเปิดกว้างมากขึ้นเท่านั้น - มาช่วยรื้อระบอบประยุทธ์ แก้อัฐธรรมนูญ ล้ม ส.ว. เลิกยุทธศาสตร์ชาติ ที่ธรรมศาสตร์ทำพระจันทร์ครับ #ปล่อยเพื่อนเรา - วิธีการเดินทางมา #รื้อระบอบประยุทธ์ พรุ่งนี้ 6 เมษายน 15.00 น. เป็นต้นไป มาร่วมลงชื่อ กันครั้งแรก! @ คณะนิติ มธ. รถมอเตอร์ เรือ "ทำพระจันทร์" BTS สะพานตากสิน แล้วเดินทางต่อมาโดยเรือด่วนเจ้าพระยา MRT สนามไชย ต่อวิน ตึกตุ๊ก หรือเดิน อยาลืมพกบัตร ปชช. มา |

ตารางที่ 4.12

ตัวอย่างข้อความกลุ่มที่ 4 ของการจัดกลุ่มข้อความที่ใช้การแปลงกลุ่มของชื่อเฉพาะ

| คำที่มีความถี่สูง | ตัวอย่างข้อความ |
|----------------------------|--|
| “ชื่อพื้นที่” | <p>- เด็กหอ-ชาวธรรมศาสตร์ รังสิตทุกคน ไปลงชื่อ #ขอคนละชื่อหรือระบอบบอประยุทธ์ พร้อมแฉะ จิบกาแพอร่อยๆ กันได้แล้วที่ ร้านกาแฟ URBIE ใครใจกล้า-มีสถานที่-พร้อมช่วยเปิดโต๊ะกับเรา แจ้งมาที่</p> <p>- นิตยสาร GQ วางที่ร้านกลมเกลียว บร.1 มธ รังสิต (เมื่อก่อนวางที่ศูนย์หนังสือ) ไปรับน้องกัน กันได้น้ำา ปล. เจ้าของร้านฝากมา</p> <p>- #รีวิว Tora Cha #รังสิต คาเฟ่โลโก้เจ้าเหมียวหน้าเสื่อคอนเซ็ปท์ว่าไรตี้ริงค์สุดคุลสาขาใหม่ ใน ม.ธรรมศาสตร์ รังสิต ที่รายล้อมด้วยพื้นที่สีเขียวร่มรื่นน่านั่งชิลล์ไปกับหลากหลายเมนูที่จัดเต็มทั้งความพรีเมียมของวัตถุดิบและความอร่อยแบบสบายกระเป๋าชั่งเริ่มต้นที่ 25.- เท่านั้น</p> |
| “ชื่อมหาลัย” | <p>- ข้อสอบจำลองสำหรับฝึกทำโค้งสุดท้ายก่อนสอบข้อสอบรวม3สาขาพร้อมข่าว เก็งทั้งเนื้อหาและข่าว เสมือนข้อสอบจริงพร้อมกระดาศคำตอบ ไม่ซื้อไม่ได้แล้วชาวว่าที่สิงห์แดง #รัฐศาสตร์ มธ #รัฐศาสตร์ #รัฐศาสตร์จุฬา #Dek64 #TCAS64 #ธรรมศาสตร์ #9วิชาสามัญ64 ชื่อซีทเลย</p> <p>- ขายสัญญาหอเพลินอพาร์ทเมนต์ ในซอยเจปาร์ค มีพิตเนสฟรี สระว่ายน้ำ มีสนามแบดมินตัน ราคาสัญญา 11,000 บาท ห้องพักเดือนละ 5,500 มีรถตู้รับส่งเข้ามา มีตู้เย็น ทีวี พร้อมออกค่าเปลี่ยนชื่อสัญญาให้ #หอนอกมธ #หอนมธ #มธ #ขายสัญญาหอ สนใจทักdm มาได้เลยค่ะ พร้อมเข้าอยู่ได้เดือนพฤษา ไม่ค่าส่วนกลาง</p> <p>- ถ้าสอบติดธรรมศาสตร์ หรือ จุฬา อยากส่งต่อหนังสือให้กับน้องๆ บางเล่มก็ขายถูกๆ บางเล่มก็ให้ฟรีๆคิดแค่ค่าส่งเลยค่า เห็นรุ่นน้องบางคนโพสแล้วเข้าใจฟิลตัวเองเลย #dek64</p> |
| “ชื่อสถานที่ในมหาวิทยาลัย” | <p>- ขออนุญาตฝากเมนูกุ้งหิมะ จากร้านป่าด้อยท่าพระจันทร์ สาย1 ร้านในตำนานเด็ก มธ.รุ่นท่าพระจันทร์ สาวๆ หนุ่มๆ ม.ธ ท่าพระจันทร์ ยุคสตรีทฟู้ด รอบดึก ต้องรู้จัก ตอนนีร้านอยู่ ถ.สาย 1 ค่ะ ราคาเมนูนี้ 99 บาท สั่งทาง online มีส่วนลด 10% นะคะ lineman ฟรีค่าส่ง 3 กม.</p> <p>- มีคนมาเฉลยด้วยน้ำว่า "ท่าพระอาทิตย์" ก็ใกล้ๆ ธรรมศาสตร์ (ท่าพระจันทร์) จริงๆ หนั้นละ</p> <p>- ฝากทำแบบสอบถามของเพื่อนเราหน่อยนะคะ ทำได้ทุกคะนะ ทุกชั้นปีเลยค่า ;_ ; ทั้งรังสิต ท่าพระจันทร์ พัทยาเลยค่า ฝากหน่อยนะคะ -ส่ม100บาท 2 คน ค่า - แคร์ทวิตและทำให้หน่อยนะคะ (เราจะส่มเมื่อยอดครบนะคะ โดยที่เราส่มแล้วให้ส่งหลักฐานการทำมาทางdmค่า)</p> |
| “ชื่อนักแสดง” | <p>- แลก เรา : ดงพโย คุณ : ลีอัน นั้รับฟิวเจอร์รังสิต มธ.รังสิต หรือปณ.ก็ได้ค่ะ *เราได้น้อง เพาะซ่า #ตลาดนัดmirae #แลกการ์ดmirae</p> <p>- คิวบ่มั่นเปงหนึ่คะทุกคน เปิดหารการ์ด JMUZE -ยวอน ไบละ 120 ค่าส่ง 30/50 (นั้รับฟิวเจอร์พาร์ค มธ.รังสิตได้นะคะ) #ตลาดนัดเพนทาگون #ตลาดนัดpentagon #ตลาดนัดยูนิ</p> <p>- เหมือนชีวิตมีหน้าที่ตื่นมาหารการ์ด เปิดหารการ์ด DMC2 -ชินวอน -ยวอน ไบละ 120 ค่าส่ง 30/50 (นั้รับฟิวเจอร์พาร์ค มธ.รังสิตได้นะคะ) #ตลาดนัดเพนทาگون #ตลาดนัดpentagon #ตลาดนัดยูนิ</p> |

จากภาพที่ 4.9, 4.10, 4.11 และ 4.12 แสดงตัวอย่างของข้อความที่มีกลุ่มของชื่อเฉพาะอยู่ในข้อความและแสดงถึงกลุ่มของชื่อเฉพาะที่ปรากฏบ่อยในกลุ่มข้อความนั้นๆ เมื่อนำข้อความตัวอย่างมาพิจารณาแล้วพบว่าข้อความในแต่ละกลุ่มมีความหมายไปในทิศทางเดียวกันมากขึ้นกว่าการใช้คำเพียงอย่างเดียว

4.4 ผลลัพธ์การสรุปหัวข้อของกลุ่มข้อความ

การสรุปหัวข้อของกลุ่มข้อความอาศัยกลุ่มของคำที่มีความถี่สูงสุด 100 อันดับแรกในกลุ่มข้อความนั้นๆ และพิจารณาจากการปรากฏร่วมกันของคำเหล่านั้น (Itemsets) โดยใช้ Itemsets ที่เกิดขึ้นในกลุ่มข้อความนั้น บ่อยมากำหนดหัวข้อของข้อความนั้นๆ

4.4.1 การสรุปหัวข้อของกลุ่มข้อความ กลุ่มที่ 1

การสรุปหัวข้อของกลุ่มข้อความ กลุ่มที่ 1 โดยการพิจารณาจากคำที่ปรากฏร่วมกันของคำที่มีความถี่สูงสุดข้อความกลุ่มที่ 1 ที่ปรากฏ 100 อันดับแรก ดังตารางที่ 4.13

ตารางที่ 4.13

คำที่มีความถี่สูงสุด 100 อันดับแรกในข้อความกลุ่มที่ 1

| ลำดับ | คำ | ความถี่ | ลำดับ | คำ | ความถี่ | ลำดับ | คำ | ความถี่ | ลำดับ | คำ | ความถี่ |
|-------|----------------|---------|-------|----------|---------|-------|------------------------------|---------|-------|-------------------|---------|
| 1 | เรียน | 365 | 26 | ม. | 63 | 51 | วิชา | 36 | 76 | ชีวิต | 25 |
| 2 | พี่ | 154 | 27 | สอน | 63 | 52 | แบบนี้ | 35 | 77 | เดิน | 25 |
| 3 | สอบ | 152 | 28 | ชั้น | 61 | 53 | ชื่อพื้นที่ | 35 | 78 | เก่ง | 25 |
| 4 | ปี | 142 | 29 | ชอบ | 61 | 54 | ตึก | 35 | 79 | หน้า | 25 |
| 5 | ทำ | 136 | 30 | งง | 57 | 55 | ชื่อ | 33 | 80 | หนู | 25 |
| 6 | เพื่อน | 109 | 31 | กก | 57 | 56 | คุย | 33 | 81 | ตาย | 25 |
| 7 | อะ | 108 | 32 | นักศึกษา | 57 | 57 | นั่ง | 32 | 82 | ห้อง | 25 |
| 8 | รู้ | 104 | 33 | ตัวเอง | 57 | 58 | รอ | 31 | 83 | วะ | 25 |
| 9 | ชื่อมหา ลัย | 102 | 34 | งาน | 56 | 59 | ชื่อสถานที่ใน มหาวิทยาลัย | 30 | 84 | แถว | 25 |
| 10 | จบ | 100 | 35 | ตอนนี้ | 54 | 60 | ดำ | 30 | 85 | ค่า | 24 |
| 11 | คณะ | 95 | 36 | อ่าน | 51 | 61 | โท | 29 | 86 | ร้าน | 24 |
| 12 | ดี | 91 | 37 | บ้าน | 51 | 62 | กะ | 29 | 87 | กลัว | 24 |
| 13 | ชื่อคณะ | 88 | 38 | อี | 50 | 63 | อย่า | 29 | 88 | ออนไลน์ | 24 |
| 14 | ดู | 88 | 39 | ผม | 50 | 64 | นั่ง | 29 | 89 | เหมือนกัน | 24 |
| 15 | แม่ | 86 | 40 | ใหม่ | 48 | 65 | ตัว | 29 | 90 | รู้สึก | 24 |
| 16 | เรื่อง | 83 | 41 | เหมือน | 44 | 66 | ไทย | 28 | 91 | อิ | 24 |
| 17 | เด็ก | 81 | 42 | หา | 43 | 67 | โดน | 28 | 92 | ชื่อ โรคติดต่อ | 23 |
| 18 | มหาลัย | 76 | 43 | บอ | 43 | 68 | เวลา | 28 | 93 | ดีกว่า | 23 |
| 19 | อะ | 76 | 44 | ตอบ | 42 | 69 | แล้วก็ | 27 | 94 | สังคม | 23 |
| 20 | น้อง | 75 | 45 | อาจารย์ | 40 | 70 | กิน | 27 | 95 | สนใจ | 23 |
| 21 | ติด | 75 | 46 | หรือ | 39 | 71 | หมอ | 27 | 96 | ขนาด | 23 |
| 22 | ถาม | 74 | 47 | พ่อ | 38 | 72 | ลูก | 26 | 97 | ทำงาน | 23 |
| 23 | เค้า | 71 | 48 | เลือก | 37 | 73 | ใจ | 26 | 98 | เล่น | 23 |
| 24 | ตอน | 69 | 49 | ใส่ | 37 | 74 | ชื่อ | 26 | 99 | ตอนนั้น | 22 |
| 25 | เจอ | 67 | 50 | ออ | 36 | 75 | อ. | 26 | 100 | หล่อ | 22 |

ตารางที่ 4.14

คำที่ปรากฏร่วมกันบ่อยที่สุดในข้อความกลุ่มที่ 1 (ปรากฏร่วมกัน 2 คำ)

| คำ, hashtag และกลุ่มของชื่อเฉพาะที่มี การปรากฏร่วมกัน | ตัวอย่างข้อความ |
|--|---|
| คำ | |
| ['เรียน', 'ปี'] | <ul style="list-style-type: none"> - [ฝากรี] สวัสดิ์ค่ะ ตอนนี้เราเรียนวิชาวิจัยอยู่ และได้ทำการวิจัยเรื่องการวางแผนชีวิตหลังจบการศึกษาของนักศึกษาชั้นปีที่4 มธ. รบกวณพีๆปี4 มธ. ที่เห็นทวิตนี้ช่วยเข้าไปทำแบบสอบถามให้หน่อยนะคะ ต้องการคนทำแบบสอบถามจำนวนมากคะ ขอขอบคุณมากๆคะ #ทิมมธ #thammasat - แล้วพอมานี้เราหมดความอยากเรียนหมอไปเยอะคือเห็นเพื่อนที่เรียนบัญชีอยู่มาเล่าให้ฟังเราก็อากเรียนมากขึ้นๆ แล้วก็รู้ว่าบัญชีจบไปไม่ตงงาน เลยตั้งใจว่าสอบรอบนี้เราจะเข้าบัญชีอันเดียวกับที่ติดเมื่อปีที่แล้วแต่ไม่ได้เลือก(มธ ตรีควบโท) แต่เราเฟลกับเลขปีนี้มากแล้วไม่รู้ว่าจะได้ถึงมั๊ย - ขายของที่ไมใช่การ์ดได้มั๊ยคะ พอดีเค้าเรียน coaching ปี3 คณะวิทยาการเรีนรู้และศึกษาศาสตร์ มธ ค่ะ ต้องเก็บชั่วโมงในการโค้ช เลยอยากหาคนมารับการโค้ชใครที่มีเรื่องอะไรในใจ ที่อยากได้รับการโค้ช อาจจะเป็นเรื่องเล็กๆน้อยก็ได้นะคะ สามารถ DM เค้ามาได้เลยนะคะ #ตลาดนัดnct |
| ['สอบ', 'เรียน'] | <ul style="list-style-type: none"> - มีใครเรียนธรรมศาสตร์ หรือเพื่อน พี่น้อง แฟน คนรู้จัก เรียนบ้างมั๊ยค่า ทุกคนเลยค่าฝากทำแบบสอบถามวิจัยเราหน่อยน้ำ ;_; - ถ้าบอกว่าเรียนที่ไหนก็เหมือนกันเด็กจะไม่ต้องเครียดว่าจะสอบเข้าเตรียมอุดม เด็กจะไม่ต้องเครียดว่าต้องเตรียมสอบเข้าธรรมศาสตร์ จุฬา เกษตร มหิดล แต่ที่เด็กเครียดเพราะอีกค่าว่าเรียนที่ไหนก็เหมือนกันมันเป็นแค่ประโยคเพื่อฝันไป - ใส่ชุดเฟรชเม่นมา แล้วทำไมคะแล้ววันนี้ชั้นก็ใส่ชุดเฟรชเม่นมาเรียน ถ้าหล่อนสอบติตรธรรมศาสตร์หล่อนคงตกตายที่เห็นเขาไม่ใส่ชุดนศ.กัน หล่อนใส่ชุดไรมาชั้นไม่รู้แต่ชั้นก็ไม่ได้ไปชี้หน้าถามคนอื่นว่าทำไมใส่ชุดนศ.มา #มทส |
| Hashtag | |
| ไม่พบปรากฏเนื่องจากไม่มี hashtag ในคำที่มีความถี่สูงสุด 100 คำ | |

ตารางที่ 4.14

คำที่ปรากฏร่วมกันบ่อยที่สุดในข้อความกลุ่มที่ 1 (ปรากฏร่วมกัน 2 คำ) (ต่อ)

| กลุ่มของชื่อเฉพาะ | |
|-----------------------|---|
| [ชื่อคณะ, 'เรียน'] | <ul style="list-style-type: none"> - ผมเรียนที่คณะศิลปศาสตร์ มธ ครับ เป็นอาจารย์จากสาขาสเปนและลาตินอเมริกัน ศึกษา เสียหายที่ผมแก่เกินไป เลยไปลงเรียน ป ตรี สาขานี้อีกใบไม่ได้แล้ว - สวัสดิ์คับ มีใครเรียนคณะ อักษรศาสตร์ จุฬา - ศิลปศาสตร์ มธ - มนุษยศาสตร์ บ้าง ไหมครับ ผมอยากทราบเกี่ยวกับหลักสูตร การเรียน การเตรียมตัวอะไรแบบนี้ อยากได้คำแนะนำและแนวทางครับ ผมเคว้ง ยพท ลพตท - ชื่อ จิม พิสิฐพล นิสัยน่ารักมากๆ อ่อนโยนแล้วก็น่ารักกับแฟนคลับทุกคนเสมอ เรียนอยู่ศิลปศาสตร์ มธ เป็น TUFolksong เล่นกีตาร์ได้ ร้องเพลงเก่ง เล่นกีฬาก็เก่ง ตอนนี้ใครอยากติดตามไอจีชื่อ pisitpon_j นะคะ กำลังมีงานถ่าย mv เรื่อยๆตอนนี้ ค่ะ #pisitpon_j |
| [เรียน, 'ชื่อมหาลัย'] | <ul style="list-style-type: none"> - กุจำวันนี้ไว้จนตายเลยอะ วันที่ต้องไปเข้าโรงแรมอยู่เพราะหอมันจะปิด แล้วก็ต้องออกไปฝึกงานไปเรียนไปสอบ ถ้าลูกอยากเข้าจุฬาจะขอให้ไปธรรมศาสตร์ เขาไม่มีวันทำกับหนูแบบนี้ ผู้บริหารต่าง mindsetต่าง #ไม่เอาคูกุหอใน #หอในจุฬา #จุฬา #chula -แล้วเมมตันนี้ก็บอกว่า... "ตอนนี้ลูกเรากำลังสอบติดจุฬา,มธ. แต่ไม่ได้เรียน ตอนนี้เราต้องจ่ายค่าเทอมให้ลูกได้เข้าเรียน ไม่มีตังค์ก็เข้าไม่ได้" แล้วเจ้าของโพสต์ก็ตอบว่า โอเคเข้าใจแล้ว แต่ดิฉันงงคะ ดิฉันโง่ลูกลมะ #นายท่านกรชิต -เออเกลียดคำว่า เป็นถึงเด็กจุฬา เด็กธรรมศาสตร์ แล้วไงคะ กูแค่สอบเข้าได้ มีเงินเรียน กูก็เป็นมนุษย์คนหนึ่งที่จ่ายเงินเรียน มีความรู้ไปสอบ คาดหวังให้หน้าสังคมแบบ ไหนอะ |

ตารางที่ 4.15

คำที่ปรากฏร่วมกันบ่อยที่สุดในข้อความกลุ่มที่ 1 (ปรากฏร่วมกัน 3 คำ)

| คำ, hashtag และกลุ่มของชื่อเฉพาะที่มีการปรากฏร่วมกัน | ตัวอย่างข้อความ |
|--|---|
| คำ | |
| ['เรียน', 'จบ', 'ปี'] | <p>- [ฝากรี] สวัสดิ์ค่ะ ตอนนี้เราเรียนวิชาวิจัยอยู่ และได้ทำการวิจัยเรื่องการวางแผนชีวิต หลังจบการศึกษาของนักศึกษาชั้นปีที่4 มธ. ครอบคลุมพีๆปี4 มธ. ที่เห็นทวิตนี้ช่วยเข้าไปทำแบบสอบถามให้หน่อยนะคะ ต้องการคนทำแบบสอบถามจำนวนมากคะ ขอคุณมากๆคะ #ทีมมธ #thammasat</p> <p>- แล้วพอมานี้เราหมดความอยากเรียนหมอไปเยอะคือเห็นเพื่อนที่เรียนบัญชีอยู่มาเล่าให้ฟังเราก็อยากเรียนมากขึ้นๆ แล้วก็รู้ว่าบัญชีจบไปไม่ตงงาน เลยตั้งใจว่าสอบรอบนี้เราจะเข้าบัญชีอันเดียวกับที่ติดเมื่อปีที่แล้วแต่ไม่ได้เลือก(มธ ตรีควบโท) แต่เราเพลกับเลขปีนี้มากแล้วไม่รู้ว่าจะได้ถึงมั้ย</p> <p>- ที่รับสอนคณิตศาสตร์, PAT1, 9วิชา, เพิ่มเกรดนะคะ มีประสบการณ์สอน 5 ปี จบวิศวะไฟฟ้า มหาวิทยาลัยธรรมศาสตร์ ขอดูประวัติที่ก่อนได้นำ ราคาตัวเองคะ ฟรีซีทเรียนตลอดการสอนเลย</p> |
| ['เรียน', 'พี', 'ปี'] | <p>- หลังจากเรียนมหาลัยมา2ปีตอนนี้ก็จะขึ้นปี3แล้ว เป็น2ปีที่ลำบากมากกก เรียนยากสุดไรสุด การแข่งขันในคณะก็สูง งานก็เยอะ สอบก็ยาก ในที่สุดวันนี้ก็รู้ตัวแล้วว่าตัวเองอยากจะทำอะไรจริงๆ ขอคุณมหาลัยธรรมศาสตร์ที่ทำให้เราได้อ่าน.. เราอยากเป็นแฟนที่เคนมะคะ</p> <p>- พีๆคนไหน เรียนอยู่ data sci มธ หรือลาดกระบัง ช่วยรีวิวให้หน่อยได้ไหมคะว่าแต่ละปีเรียนอะไรบ้าง #Dek64 #dek64</p> <p>- ประวัติคะแนนพีนะค้ำบ (คะแนนปี61) ส่วนของปี64 รออออก แล้วจะมาอัปเดตน้ำ พีเรียนอยู่ม.ธรรมศาสตร์ปี 3 (ชีว) มีประสบการณ์สอนGat Eng ตั้งแต่ปี 2560 สอนน้องมาไม่ต่ำกว่า 20คน #สอนพิเศษ #ติวเตอร์ #เรียนพิเศษ #หาติวเตอร์ #GATPAT #Dek65 #สอนพิเศษออนไลน์ #GatEng #เด็กชีว</p> |
| Hashtag | |
| ไม่พบปรากฏเนื่องจากไม่มี hashtag ในคำที่มีความถี่สูงสุด 100 คำ | |

ตารางที่ 4.15

คำที่ปรากฏร่วมกันบ่อยที่สุดในข้อความกลุ่มที่ 1 (ปรากฏร่วมกัน 3 คำ)(ต่อ)

| กลุ่มของชื่อเฉพาะ | |
|------------------------------------|---|
| <p>['ชื่อคณะ', 'เรียน', 'จบ']</p> | <ul style="list-style-type: none"> - แล้วพอมาปีนี้เราหมดความอยากเรียนหมอไปเยอะคือเห็นเพื่อนที่เรียนบัญชีอยู่มาเล่าให้ฟังเราก็อยากเรียนมากขึ้นๆ แล้วก็รู้ว่าบัญชีจบไปไม่ตกงาน เลยตั้งใจว่าสอบรอบนี้เราจะเข้าบัญชีอันเดียวกับที่ติดเมื่อปีที่แล้วแต่ไม่ได้เลือก(มธ ตรีควบโท) แต่เราเพลกับเลขปีนี้มากแล้วไม่รู้ว่าจะได้ถึงมั๊ย - สิ่งที่น่าสนใจอีกอย่างคือน้องเรียนจบแล้วจากคณะ เศรษฐศาสตร์ภาคอินเตอร์ มหาวิทยาลัยธรรมศาสตร์และถึงแม้ว่าจะไม่ชอบเลยแต่น้องก็ตั้งใจเรียนจนจบและทำมันออกมาได้ดีด้วย #WhatWinDone - ใค้จะเรียนนิติมข.บอกตรงๆนะหนีไป ยุคกูเรียนมีครูพิเศษที่สอนธรรมศาสตร์ แต่หลังจบมาแล้ว รุ่นน้องเรียนแต่กับครุคณะ ครุคณะที่ก๊วยกมที่คุณภาพที่สุดมีคนเดียวอ.วนิดานอกนั้นบ๊วยย ถ้าจะเรียนเอาต่อเนติ หนีไปซะ อย่ามาเรียนนิติมข.จากใจคนเคยเรียน มิ่งเรียนรามยังได้อ.ที่สอนคุณภาพกว่ากูพูดตรงๆ |
| <p>['ชื่อคณะ', 'เรียน', 'ชอบ']</p> | <ul style="list-style-type: none"> - น้องๆ #dek65 ที่สนใจสอบเข้า นิติศาสตร์ ธรรมศาสตร์ #นิติมธ หรือว่ายังค้นหาตัวเองอยู่ สามารถเข้าไปทดลองเรียนฟรีได้ตามคลิปนี้เลยนะคะ เนื้อหาการสอนเป็นการเช็คว่าคุณชอบคณะนี้จริงหรือไม่ และมีส่วนของข้อสอบวิชาเฉพาะที่ต้องใช้ในการสอบเข้าด้วยคะ - รัฐศาสตร์ กูอยากเรียนเพราะมันน่าเรียน ขอเรียนสิ่งที่ชอบบ้างเถอะซักอย่าง มธ.ไร้งี้ -จะมีใครมีข้อมูลเรื่องเรียนต่อภาคบัณฑิตที่ไม่ใช่ของมอรามมั้ยนะ อยากรู้ว่านอกจากนิติแล้ว มีคณะไหนอีกบ้าง ตอนนี้อยู่นิติ มธ. แต่ก็กลัวจะไม่ชอบจริงๆ ยังไม่ได้จะเรียนเร็วๆนี้แต่คิดว่าจะเรียนต่ออีกใ้บแน่ๆตอนนี้ขอเก็บตั้งก่อน |

ตารางที่ 4.16

คำที่ปรากฏร่วมกันบ่อยที่สุดในข้อความกลุ่มที่ 1 (ปรากฏร่วมกัน 4 คำ)

| คำ, hashtag และ กลุ่มของชื่อเฉพาะที่มีการปรากฏร่วมกัน | ตัวอย่างข้อความ |
|--|---|
| คำ | |
| ['จบ', 'สอน', 'ออนไลน์', 'ปี'] | <ul style="list-style-type: none"> - รับสอนคณิตศาสตร์ เด็กทุน จภ จบ เศรษฐศาสตร์ มธ ประสบการณ์ 3 ปี #Pat1/9 วิชา/ม.ต้น-ม.ปลาย ทิวสอบเข้า มธ/พิวเจอร์/อยุธยา มีวีดีโอตัวอย่างการสอนให้ดู r #สอนพิเศษ #หาติวเตอร์ #เรียนพิเศษ #เรียนออนไลน์ #dek65 #สสวท #9วิชาสามัญ - รับสอนเคมี ฟิสิกส์ ชีวะ คณิต วิทยาศาสตร์ Genchem ตั้งแต่อนุบาลถึงมหา |
| ['จบ', 'มหาลัย', 'สอน', 'ปี'] | <ul style="list-style-type: none"> ลัย ผู้สอนจบเคมี มธ ประสบการณ์สอน 3 ปี รับสอนเชียงใหม่ ออนไลน์ ก็รับ Dm ได้เลย #หาติวเตอร์ #เรียนพิเศษ #สอนพิเศษ #สอนพิเศษ เชียงใหม่ #เรียนพิเศษเชียงใหม่ #ติวเตอร์เชียงใหม่ #เรียนเคมี - รับสอนคณิตศาสตร์ เด็กทุน จภ จบ เศรษฐศาสตร์ มธ ประสบการณ์ 3 ปี #Pat1/9 วิชา/ม.ต้น-ม.ปลาย ทิวสอบเข้า มธ/พิวเจอร์/อยุธยา มีวีดีโอตัวอย่างการสอนให้ดู ออนไลน์/สอนสด 180-300 บาท/hr #สอนพิเศษ #หาติวเตอร์ #เรียนพิเศษ #เรียนออนไลน์ #dek65 |
| Hashtag | |
| ไม่พบปรากฏเนื่องจากไม่มี hashtag ในคำที่มีความถี่สูงสุด 100 คำ | |
| กลุ่มของชื่อเฉพาะ | |
| [ชื่อคณะ, 'เรียน', 'จบ', 'ปี'] | <ul style="list-style-type: none"> - แล้วพอมาปีนี้เราหมดความอยากเรียนหมอไปเยอะคือเห็นเพื่อนที่เรียนบัญชีอยู่มาเล่าให้ฟังเราก็อยากเรียนมากขึ้นๆ แล้วก็รู้ว่าบัญชีไปไม่ตกงาน เลยตั้งใจว่าสอบรอบนี้เราจะเข้าบัญชีอันเดียวกับที่ติดเมื่อปีที่แล้วแต่ไม่ได้เลือกแต่เราเพลกับเลขปีนี้มากแล้วไม่รู้ว่าจะได้ถึงมั๊ย - ขออนุญาตสอบถมนะคะ ที่บอกว่าได้งานตั้งแต่ปี 4 เทอมแรกคือสามารถสมัครงานที่ Big4 ได้ตั้งแต่ยังเรียนไม่จบ(แต่กำลังจะจบในเทอมนั้น)ใช่มั๊ยคะ จริงๆแล้วหนูเรียนนิติ มธ ค่ะ สนใจงาน legal ใน Big4 อยากจะสมัครงานไว้ตั้งแต่เนิ่นๆก่อนที่จะเรียนจบ แต่ก็ไม่รู้จะเริ่มสมัครตั้งแต่ยังไม่จบได้มั๊ย - ทุกคนครับคือตอนนี้เราเรียน ปวช.1 แผนกการตัดการโลจิสติกส์ กำลังจะขึ้น ปี2 แต่เราอยากเรียนนิติศาสตร์มากๆเลยครับ เราอยากถามว่าถ้าเราจบ ปวช.3 เราจะสามารถสอบเข้าคณะนิติศาสตร์ มหาลัยไหนได้บ้างครับ.....แอบหวัง มธ. ไว้เหมือนกัน #นิติมธ #นิติศาสตร์ ขออนุญาตใช้แท็กนะครับ |

ตารางที่ 4.16

คำที่ปรากฏร่วมกันบ่อยที่สุดในข้อความกลุ่มที่ 1 (ปรากฏร่วมกัน 4 คำ)(ต่อ)

| กลุ่มของชื่อเฉพาะ(ต่อ) | |
|-----------------------------------|---|
| ['ชื่อคณะ', 'เรียน', 'จบ', 'ชอบ'] | <ul style="list-style-type: none"> - สิ่งที่มีใจอีกอย่างคือน้องเรียนจบแล้วจากคณะ เศรษฐศาสตร์ภาคอินเตอร์มหาวิทยาลัยธรรมศาสตร์และถึงแม้ว่าจะไม่ชอบเลยแต่น้องก็ตั้งใจเรียนจนจบและทำมันออกมาได้ดีด้วย #WhatWinDone - 16. สิ่งที่ชอบของนิติ มธ. คือ การปล่อยให้นัก.รับผิดชอบตัวเองในการเรียน มันทำให้มีความเป็น liberal กว่าสมัยเรียนม.ปลายมาก แต่ก็มีสิ่งที่ไม่ชอบ คือรู้สึกว่าระบบเรียนแบบ sec ยังไม่สามารถหาจุดวางให้สมดุลได้ (แต่คงเป็นปัญหา common) กับอ.สายศาลเน้นการตอบแบบอิงธง เหตุผลดีแต่ผิดตรงคือจบ - ไม่ได้เรียน ม.4 5 6 ตอนจบมาเรียน ปวชัญชี ชอบงานด้านนิเทศ ไปเรียนสื่อสารมวลชนจบใน 3 ปี แล้วต่อโทธรรมศาสตร์ ทุกคนมีดีในตัวจริงๆ |

จากตารางที่ 4.14, 4.15 และ 4.16 พบคำที่ปรากฏร่วมกันได้แก่ ['เรียน', 'ปี'], ['สอบ', 'เรียน'], ['ชื่อคณะ', 'เรียน'], ['เรียน', 'ชื่อมหาลัย'], ['เรียน', 'จบ', 'ปี'], ['เรียน', 'พี', 'ปี'], ['ชื่อคณะ', 'เรียน', 'จบ'], ['ชื่อคณะ', 'เรียน', 'ชอบ'], ['จบ', 'สอน', 'ออนไลน์', 'ปี'], ['จบ', 'มหาลัย', 'สอน', 'ปี'], ['ชื่อคณะ', 'เรียน', 'จบ', 'ปี'] และ ['ชื่อคณะ', 'เรียน', 'จบ', 'ชอบ'] จากคำที่ปรากฏร่วมกัน และตัวอย่างของข้อความสามารถสรุปหัวข้อของข้อความ กลุ่มที่ 1 พบว่า เกี่ยวข้องกับการเรียนการสอน การศึกษาเป็นต้น

4.4.2 การสรุปหัวข้อของกลุ่มข้อความ กลุ่มที่ 2

การสรุปหัวข้อของกลุ่มข้อความ กลุ่มที่ 2 โดยการพิจารณาจากคำที่ปรากฏ
ร่วมกันของคำที่มีความถี่สูงสุดข้อความกลุ่มที่ 2 ที่ปรากฏ 100 อันดับแรก ดังตารางที่ 4.17

ตารางที่ 4.17

คำที่มีความถี่สูงสุด 100 อันดับแรกในข้อความกลุ่มที่ 2

| ลำดับ | คำ | ความถี่ | ลำดับ | คำ | ความถี่ | ลำดับ | คำ | ความถี่ | ลำดับ | คำ | ความถี่ |
|-------|----------------|---------|-------|-------------|---------|-------|-----------------|---------|-------|----------------|---------|
| 1 | ชื่อคณะ | 367 | 26 | เรียน | 32 | 51 | เพื่อน | 22 | 76 | ตอน | 15 |
| 2 | ติด | 332 | 27 | เงิน | 32 | 52 | อะ | 21 | 77 | สู้ | 15 |
| 3 | #Dek64 | 275 | 28 | ดู | 32 | 53 | มหิดล | 21 | 78 | เด็ก | 14 |
| 4 | คะแนน | 214 | 29 | สอบถาม | 31 | 54 | กิน | 21 | 79 | ไม่ต้อง | 14 |
| 5 | รอบ | 157 | 30 | #dek65 | 31 | 55 | หนังสือ | 21 | 80 | คอม | 14 |
| 6 | แจก | 129 | 31 | น้ำ | 31 | 56 | คุย | 20 | 81 | gat | 14 |
| 7 | สอบ | 118 | 32 | เอก | 30 | 57 | ลอง | 20 | 82 | วิทย์ | 14 |
| 8 | ยื่น | 117 | 33 | เล่น | 30 | 58 | Gt | 20 | 83 | สำหรับ | 14 |
| 9 | ปี | 95 | 34 | เลือก | 29 | 59 | สาขา | 20 | 84 | น้อง | 13 |
| 10 | ชื่อ ข้อสอบ | 94 | 35 | เดือน | 29 | 60 | ทวิ | 20 | 85 | น้อง ๆ | 13 |
| 11 | ชื่อมหาลัย | 67 | 36 | เค้า | 29 | 61 | รบกวน | 20 | 86 | หา | 13 |
| 12 | วิชา | 62 | 37 | เจ | 28 | 62 | งง | 19 | 87 | รุ่นพี่ | 13 |
| 13 | คณะ | 60 | 38 | ดี | 28 | 63 | เกณฑ์ | 19 | 88 | ศูนย์ | 13 |
| 14 | ใหม่ | 58 | 39 | มหาลัย | 26 | 64 | เตรียมตัว | 18 | 89 | สิน | 13 |
| 15 | บาท | 56 | 40 | กด | 26 | 65 | รีวิว | 18 | 90 | สอน | 13 |
| 16 | #นิติมธ | 49 | 41 | ชื่อพื้นที่ | 25 | 66 | บริจาค | 17 | 91 | งด | 13 |
| 17 | ค่า | 49 | 42 | รู้ | 25 | 67 | ข้อสอบ | 17 | 92 | หรือ | 13 |
| 18 | รี | 47 | 43 | สู่ม | 25 | 68 | แนน | 16 | 93 | นิง | 12 |
| 19 | คับ | 42 | 44 | หนู | 24 | 69 | สารู | 16 | 94 | จับ | 12 |
| 20 | บริหาร | 37 | 45 | ad | 23 | 70 | ตอนนี้ | 16 | 95 | สัก | 12 |
| 21 | ที่จะ | 37 | 46 | ค้าบ | 23 | 71 | สาด | 16 | 96 | แง | 12 |
| 22 | #Onet64 | 35 | 47 | น้องๆ | 23 | 72 | กรอก | 15 | 97 | อึ่ง | 12 |
| 23 | พี่ | 34 | 48 | ภาษา | 23 | 73 | อย่า | 15 | 98 | ตัวเอง | 12 |
| 24 | ขอให้ | 33 | 49 | ทำ | 23 | 74 | #ศิลป ศาสตร์ | 15 | 99 | # รัฐศาสตร์ | 12 |
| 25 | อันดับ | 32 | 50 | ประกาศ | 22 | 75 | มีโอกาส | 15 | 100 | มก | 12 |

ตารางที่ 4.18

คำที่ปรากฏร่วมกันบ่อยที่สุดในข้อความกลุ่มที่ 2 (ปรากฏร่วมกัน 2 คำ)

| คำ, hashtag และ กลุ่ม ของชื่อเฉพาะที่มีการ ปรากฏร่วมกัน | ตัวอย่างข้อความ |
|---|---|
| คำ | |
| ['แจก', 'ติด'] | <ul style="list-style-type: none"> - บ่นค่ะ! ถ้าปีนี้สอบติดนิติศาสตร์ ธรรมศาสตร์รังสิต หรือ นิติศาสตร์ จุฬาลงกรณ์ แจก 100 บาท 3 คนค่ะ (รีเลย์คั๊บ ไม่ต้องพอลก็ได้จ้า) #dek64 - RT บ่นค่า +°★ สอบติดเภสัช มธ. แจกเลยการ์ตจีของผ้าอ้อม 1 ใบ ปล.ใครไม่ตาม nct เปลี่ยนเป็นเงิน 500 บาท ได้นะคะ #Dek64 #ตลาดนัดอชทช - บ่นค่ะ ถ้าติด นิติ มธ. แจกเงิน 219 บาทค่ะ และจะมาเขียนรีวิวกการเตรียมตัวทั้งหมดของเราให้ด้วยค่ะ #dek64 #นิติมธ |
| ['ติด', 'รอบ'] | <ul style="list-style-type: none"> - บ่น+รีทวิต ถ้าสอบติด รัฐศาสตร์IR มธ. หรือ มช. รอบ2-3 จะมาแจกหนังสือมือ2 ฟรี!! 1คน (เอาไปทั้งหมดเลย) แจกหลังจากประกาศผลสอบนะคะทุกคน #ของสะสมเต๋ยเต๋ย #ส่งต่อหนังสือเตรียมสอบ #หนังสือมือสอง #dek65 #dek66 #วิชาสามัญ #gatpat #Onet - บ่น ถ้าติด นิติ มธ.รังสิต รอบ3 แจก 1000 ถ้าติด นิติ มธ.ลำปาง รอบ3 แจก 500 ถ้าติด นิติ มช. รอบ3 แจก 100 ถ้าติด นิติ มธ. ไม่ว่าจะที่ไหนจะโอนเงินให้ มูลนิธิเกี่ยวกับน้องหมาเดือนละ 100 นาน 1ปี ร้อยอย่างเดียว #Dek64 - *****บ่น***** ถ้าติดนิติศาสตร์ มธ.รอบ3 แจก250฿ (वल) ร้อยอย่างเดียวค่า ไม่ต้องพอลก็ได้คั๊บ สุ่มตอนประกาศผล #Dek64 |
| Hashtag | |
| ['#Dek64', 'คะแนน'] | <ul style="list-style-type: none"> - บ่น!!!! แจก 200 บาท 1คน -ติดบัญชี มธ. -คะแนนทุกอย่างผ่านเกินครึ่ง ร้อยอย่างเดียวไม่ต้องพอลก็ได้คะแนน นู๋อยากมีที่เรียนแนว ใครผ่านไปผ่านมาส่งกำลังจัยมาให้หนูด้วยฮรูกกก ขอให้มีที่เรียนกันทุกคนนะคะ #Onet64 #9วิชาสามัญ #Dek64 - ถ้าได้คะแนนสามัญ ไทย 60+ สังคม 60+ ฟิสิกส์ 40+ ชีวะ 60+ เคมี 50+ คณิต 60+ Eng 70+ สอบติดแพทย์ มธ แจก100บาท พร้อมแจกหนังสือ สารุ #Dek64 #9วิชาสามัญ64 #9วิชาสามัญ - บ่นเพิ่ม คะแนนเท่ากับแพทย์ปู่นไม่ถึง 200 555555 ถ้าคะแนนวิชาเฉพาะ นิติศาสตร์ธรรมศาสตร์ได้เกิน 55 คะแนน แจก 300 บาท 1 คน 200 บาท 1 คน 55 บาท 1 คนค่ะ ไม่ไหวจริง จะเรียนคณะอะไรดีคั๊บ 55555 #Dek64 |

ตารางที่ 4.18

คำที่ปรากฏร่วมกันบ่อยที่สุดในข้อความกลุ่มที่ 2 (ปรากฏร่วมกัน 2 คำ)(ต่อ)

| Hashtag(ต่อ) | |
|-----------------------|---|
| ['#Dek64', 'ติด'] | <ul style="list-style-type: none"> - บ่นค่ะ! ถ้าปีนี้สอบติดนิติศาสตร์ ธรรมศาสตร์รังสิต หรือ นิติศาสตร์ จุฬาลงกรณ์ แจก 100 บาท 3 คนค่ะ (รีเลยคับ ไม่ต้องพอลก็ได้้น้ำ) #dek64 - ถ้าติดศิลปศาสตร์ มธ แจก100 ถ้าติดศิลปศาสตร์ มธ แจก100 ถ้าติดศิลปศาสตร์ มธ แจก100 ถ้าติดศิลปศาสตร์ มธ แจก100 ถ้าติดศิลปศาสตร์ มธ แจก100 #Dek64 - บ่นไปแล้วแต่ขอบนอีกรอบละกัน GAT 250 PAT1 130 PAT2 150 สามัญคณิต 65 ไทย 70 ฟิสิก 50 เคมี 50 ชีวะ 70 อังกฤษ70 สังคม 50 ติดหมอหรือทันตฯ มน.หรือ มธ. ติด com-sci มน. ถ้าได้ทั้งหมดนี้สู่มแจก 2คน คนละ 300 บาท จะมาสู่มเมื่อผลประกาศเป็นที่เรียบร้อยแล้ว #Dek64 |
| กลุ่มของชื่อเฉพาะ | |
| ['#Dek64', 'ชื่อคณะ'] | <ul style="list-style-type: none"> - บ่นค่ะ! ถ้าปีนี้สอบติดนิติศาสตร์ ธรรมศาสตร์รังสิต หรือ นิติศาสตร์ จุฬาลงกรณ์ แจก 100 บาท 3 คนค่ะ (รีเลยคับ ไม่ต้องพอลก็ได้้น้ำ) #dek64 - บ่นค่ะ ถ้าติด นิติ มธ. แจกเงิน 219 บาทค่ะ และจะมาเขียนรีวิวกการเตรียมตัวทั้งหมดของเราให้ด้วยค่ะ #dek64 #นิติมธ |
| ['ชื่อคณะ', 'ติด'] | <ul style="list-style-type: none"> - ถ้าติดศิลปศาสตร์ มธ แจก100 ถ้าติดศิลปศาสตร์ มธ แจก100 ถ้าติดศิลปศาสตร์ มธ แจก100 ถ้าติดศิลปศาสตร์ มธ แจก100 ถ้าติดศิลปศาสตร์ มธ แจก100 #Dek64 |

ตารางที่ 4.19

คำที่ปรากฏร่วมกันบ่อยที่สุดในข้อความกลุ่มที่ 2 (ปรากฏร่วมกัน 3 คำ)

| คำ, hashtag และ กลุ่ม ของชื่อเฉพาะที่มีการ ปรากฏร่วมกัน | ตัวอย่างข้อความ |
|---|---|
| คำ | |
| ['แจก', 'บาท', 'ติด'] | <ul style="list-style-type: none"> - บ่นค่ะ! ถ้าปีนี้สอบติดนิติศาสตร์ ธรรมศาสตร์รังสิต หรือ นิติศาสตร์ จุฬาลงกรณ์ แจก 100 บาท 3 คนค่ะ (รีเลยคับ ไม่ต้องพอลก็ได้น้า) #dek64 - RT บ่นค่า +* สอบติดเกสซ์ มธ. แจกเลย์การ์ดจีของผ้าอ้อม 1 ใบ ปล.ใครไม่ตาม nct เปลี่ยนเป็นเงิน 500 บาท ได้นะคะ #Dek64 #ตลาดนัดอชทช - บ่นค่ะ ถ้าติด นิติ มธ. แจกเงิน 219 บาทค่ะ และจะมาเขียนรีวิวกการเตรียมตัว ทั้งหมดของเราให้ด้วยค่ะ #dek64 #นิติมธ |
| ['แจก', 'รี', 'ติด'] | <ul style="list-style-type: none"> - บ่นค่ะ! ถ้าปีนี้สอบติดนิติศาสตร์ ธรรมศาสตร์รังสิต หรือ นิติศาสตร์ จุฬาลงกรณ์ แจก 100 บาท 3 คนค่ะ (รีเลยคับ ไม่ต้องพอลก็ได้น้า) #dek64 - บ่น+รีทวิต ถ้าสอบติด รัฐศาสตร์IR มธ. หรือ มช. รอบ2-3 จะมาแจกหนังสือมือ2 ฟรี!! 1คน (เอาไปทั้งหมดเลย) แจกหลังจากประกาศผลสอบนะคะทุกคน #ของสะสมแต่ยแต่ย #ส่งต่อหนังสือเตรียมสอบ #หนังสือมือสอง #dek65 #dek66 #วิชาสามัญ #gatpat #Onet - บ่นค่า- ถ้าติดบัญชี มธ แจก 500 3คน บริหาร มธ แจก 300 3คน บัญชี มก แจก 200 4 คน ร้อย่างเดียวพอค่า #Dek64 #9วิชาสามัญ |
| Hashtag | |
| ['แจก', '#Dek64', 'ติด'] | <ul style="list-style-type: none"> - บ่นค่ะ! ถ้าปีนี้สอบติดนิติศาสตร์ ธรรมศาสตร์รังสิต หรือ นิติศาสตร์ จุฬาลงกรณ์ แจก 100 บาท 3 คนค่ะ (รีเลยคับ ไม่ต้องพอลก็ได้น้า) #dek64 - บ่นค่ะ ถ้าเพื่อนติดธรรมศาสตร์แจก 200฿ & & ร้อย่างเดียวค่ะ & & ; ขออนุญาตติดแท็กคับ #ตลาดนัดnct - บ่น+รีทวิต ถ้าสอบติด รัฐศาสตร์IR มธ. หรือ มช. รอบ2-3 จะมาแจกหนังสือมือ2 ฟรี!! 1คน (เอาไปทั้งหมดเลย) แจกหลังจากประกาศผลสอบนะคะทุกคน #ของสะสมแต่ยแต่ย #ส่งต่อหนังสือเตรียมสอบ #หนังสือมือสอง #dek65 #dek66 #วิชาสามัญ #gatpat #Onet |

ตารางที่ 4.20

คำที่ปรากฏร่วมกันบ่อยที่สุดในข้อความกลุ่มที่ 2 (ปรากฏร่วมกัน 4 คำ)

| คำที่มีการปรากฏร่วมกัน | ตัวอย่างข้อความ |
|---------------------------------|--|
| | คำ |
| ['ติด', 'บาท', 'แจก', 'รี'] | - บ่นค่ะ! ถ้าปีนี้สอบติดนิติศาสตร์ ธรรมศาสตร์รังสิต หรือ นิติศาสตร์ จุฬาลงกรณ์ แจก 100 บาท 3 คนค่ะ (รีเลยคับ ไม่ต้องพอลก็ได้จ้า) #dek64 - บ่นค่ะ ถ้าติด นิติ มธ. แจกเงิน 219 บาทค่ะ และจะมาเขียนรีวิวการเตรียมตัว ทั้งหมดของเราให้ด้วยค่ะ #dek64 #นิติมธ |
| ['แจก', 'เงิน', 'บาท', 'ติด'] | - บ่นจ้า ถ้าติด มธ. • แจกเงิน 100 บาท 1 คน • งดกินของทอด 1 ปี รืออย่าง เดียว ไม่ต้องพอลจ้า |
| | Hashtag |
| ['แจก', '#Dek64', 'บาท', 'ติด'] | - บ่นค่ะ! ถ้าปีนี้สอบติดนิติศาสตร์ ธรรมศาสตร์รังสิต หรือ นิติศาสตร์ จุฬาลงกรณ์ แจก 100 บาท 3 คนค่ะ (รีเลยคับ ไม่ต้องพอลก็ได้จ้า) #dek64 - RT บ่นค่า +°★ สอบติดเกสซ์ มธ. แจกเลย์การ์ดจีของผ้าอ้อม 1 ใบ ปล.ใครไม่ตาม nct เปลี่ยนเป็นเงิน 500 บาท ได้นะคะ #Dek64 #ตลาดนัดอชทช - บ่นค่ะ ถ้าติด นิติ มธ. แจกเงิน 219 บาทค่ะ และจะมาเขียนรีวิวการเตรียมตัว ทั้งหมดของเราให้ด้วยค่ะ #dek64 #นิติมธ |
| ['แจก', '#Dek64', 'รี', 'ติด'] | - บ่นค่ะ! ถ้าปีนี้สอบติดนิติศาสตร์ ธรรมศาสตร์รังสิต หรือ นิติศาสตร์ จุฬาลงกรณ์ แจก 100 บาท 3 คนค่ะ (รีเลยคับ ไม่ต้องพอลก็ได้จ้า) #dek64 - บ่นค่า- ถ้าติดบัญชี มธ แจก 500 3คน บริหาร มธ แจก 300 3คน บัญชี มก แจก 200 4 คน รืออย่างเดียวพอลค่า #Dek64 #9วิชาสามัญ - บ่นค่าบริได้เลยคับ GAT 150+แจก 50 บาท PAT1 เกิน 90 แจก 50 บาท PAT2 125+แจก 50 บาท ติดเทคนิคการแพทย์ มหิดล แจก 100บาท ติดสหเวช มธ แจก 100 บาท ติดเทคนิคการแพทย์ มโหนก็ได้ 50 บาท #Dek64 รีได้เลยคับ |

ตารางที่ 4.20

คำที่ปรากฏร่วมกันบ่อยที่สุดในข้อความกลุ่มที่ 2 (ปรากฏร่วมกัน 4 คำ)(ต่อ)

| กลุ่มของชื่อเฉพาะ | |
|--|---|
| ['แจก', '#Dek64', 'ชื่อคณะ', 'ติด'] | <ul style="list-style-type: none"> - บ่นค่ะ! ถ้าปีนี้สอบติดนิติศาสตร์ ธรรมศาสตร์รังสิต หรือ นิติศาสตร์ จุฬาลงกรณ์ แจก 100 บาท 3 คนค่ะ (รีเลยคับ ไม่ต้องพอลก็ได้จ้า) #dek64 - บ่นค่ะ ถ้าติด นิติ มธ. แจกเงิน 219 บาทค่ะ และจะมาเขียนรีวิวกการเตรียมตัว ทั้งหมดของเราให้ด้วยค่ะ #dek64 #นิติมธ - บ่นค่า- ถ้าติดบัญชี มธ แจก 500 3คน บริหาร มธ แจก 300 3คน บัญชี มก แจก 200 4 คน ร้อย่างเดียวพอค่า #Dek64 #9วิชาสามัญ |
| ['ชื่อคณะ', 'ยื่น', 'คะแนน', '#Dek64'] | <ul style="list-style-type: none"> - ใครจะยื่นบัญชี-บัญชี มธ รอบadmission2 ได้คะแนน gat pat1 เท่าไหร่กันบ้างคะ (โพล pat1 อยู่ข้างล่างคับ) #dek64 #บัญชีมธ - คนที่จะยื่นเทคนิคการแพทย์ มธ รอบ3.2 คิดคะแนนแล้วได้เท่าไหรกันบ้างคะ เต็ม21,000นะคะ เนื่องจากonetยังไม่ออก เดียวจะแปะวิธีคิดไว้ในเม้นนะคะ #dek64 - มีใครจะยื่นวารสาร มธ แบบคะแนน gat รวมมั้ยคะ ได้กันเท่าไรบ้างแงงๆ #dek64 #วารสารมธ |

จากตารางที่ 4.18, 4.19 และ 4.20 พบคำที่ปรากฏร่วมกันได้แก่ ['แจก', 'ติด'], ['ติด', 'รอบ'], ['#Dek64', 'คะแนน'], ['#Dek64', 'ติด'], ['#Dek64', 'ชื่อคณะ'], ['ชื่อคณะ', 'ติด'], ['แจก', 'บาท', 'ติด'], ['แจก', 'รี', 'ติด'], ['แจก', '#Dek64', 'ติด'], ['#Dek64', 'คะแนน', 'ยื่น'], ['#Dek64', 'ชื่อคณะ', 'ติด'], ['#Dek64', 'ชื่อคณะ', 'คะแนน'], ['ติด', 'บาท', 'แจก', 'รี'], ['แจก', 'เงิน', 'บาท', 'ติด'], ['แจก', '#Dek64', 'บาท', 'ติด'], ['แจก', '#Dek64', 'รี', 'ติด'], ['แจก', '#Dek64', 'ชื่อคณะ', 'ติด'], ['ชื่อคณะ', 'ยื่น', 'คะแนน', '#Dek64'] จากคำที่ปรากฏร่วมกัน และตัวอย่างของข้อความ สามารถสรุปหัวข้อของข้อความ กลุ่มที่ 2 พบว่าข้อความกลุ่มนี้ เกี่ยวข้องกับการสอบเข้า, ความต้องการที่จะศึกษาต่อที่มหาวิทยาลัยธรรมศาสตร์

4.4.3 การสรุปหัวข้อของกลุ่มข้อความ กลุ่มที่ 3

การสรุปหัวข้อของกลุ่มข้อความ กลุ่มที่ 3 โดยการพิจารณาจากคำที่ปรากฏ
ร่วมกันของคำที่มีความถี่สูงสุดข้อความกลุ่มที่ 3 ที่ปรากฏ 100 อันดับแรก ดังตารางที่ 4.21

ตารางที่ 4.21

คำที่มีความถี่สูงสุด 100 อันดับแรกในข้อความกลุ่มที่ 3

| ลำดับ | คำ | ความถี่ | ลำดับ | คำ | ความถี่ | ลำดับ | คำ | ความถี่ | ลำดับ | คำ | ความถี่ |
|-------|---|---------|-------|--------------|---------|-------|---------------------|---------|-------|-------------|---------|
| 1 | ชื่อ โรงพยาบาล | 168 | 26 | เอฟ | 33 | 51 | แจ้ง | 21 | 76 | บุคลากร | 17 |
| 2 | ชื่อโรคติดต่อ | 153 | 27 | #โควิด19 | 32 | 52 | ฟรี | 20 | 77 | ระบอบ | 17 |
| 3 | เมษายน | 109 | 28 | เรื่อง | 31 | 53 | น้อง | 20 | 78 | ชื่อ | 17 |
| 4 | ชื่อพื้นที่ | 107 | 29 | ห้อง | 28 | 54 | บริการ | 20 | 79 | งาน | 17 |
| 5 | นักศึกษา | 90 | 30 | อาจารย์ | 28 | 55 | ชลบุรี | 20 | 80 | ออนไลน์ | 16 |
| 6 | ตรวจ | 88 | 31 | ซึ่ง | 28 | 56 | สถานการณ์ | 20 | 81 | ราก | 16 |
| 7 | ชื่อบุคคลที่ เกี่ยวข้องกับ การเมือง | 86 | 32 | อย่า | 28 | 57 | เชียงใหม่ | 19 | 82 | เจ้าหน้าที่ | 16 |
| 8 | โรงพยาบาล | 79 | 33 | เพื่อน | 27 | 58 | จี | 19 | 83 | ฟ้า | 16 |
| 9 | เตียง | 67 | 34 | บริจาค | 27 | 59 | ลงชื่อ | 19 | 84 | สถานที่ | 16 |
| 10 | เวลา | 63 | 35 | คณะ | 27 | 60 | ข่าว | 19 | 85 | ดี | 16 |
| 11 | ศูนย์ | 58 | 36 | ยื่น | 25 | 61 | เชิญ | 19 | 86 | ทีม | 16 |
| 12 | ประกาศ | 55 | 37 | ชื่อจังหวัด | 25 | 62 | กักตัว | 19 | 87 | ดู | 16 |
| 13 | ชื่อสถานที่ใน มหาวิทยาลัย | 55 | 38 | กิจกรรม | 24 | 63 | อาคาร | 19 | 88 | เพจ | 16 |
| 14 | ผู้ป่วย | 49 | 39 | พรรค | 24 | 64 | #ปล่อยเพื่อน เรา | 19 | 89 | เอ | 16 |
| 15 | รอบ | 49 | 40 | ศาลฎีกา | 24 | 65 | ประกันตัว | 19 | 90 | ศาล | 16 |
| 16 | ชื่อคณะ | 49 | 41 | ติดเชื้อ | 24 | 66 | ข้อมูล | 19 | 91 | ชุมนุม | 16 |
| 17 | หน้า | 46 | 42 | ปี | 23 | 67 | มหาวิทยาลัย | 19 | 92 | ข้าง | 16 |
| 18 | สอบ | 44 | 43 | เพนกวิ้น | 23 | 68 | โลหิต | 19 | 93 | ศึกษา | 16 |
| 19 | สนาม | 39 | 44 | ออกมา | 23 | 69 | มือบ | 18 | 94 | บ้าน | 16 |
| 20 | ติด | 39 | 45 | คณะรัฐศาสตร์ | 23 | 70 | เชียงใหม่ | 18 | 95 | กล่อง | 15 |
| 21 | ตอนนี้ | 38 | 46 | รพ | 23 | 71 | เดือน | 18 | 96 | เรือ | 15 |
| 22 | หยุด | 36 | 47 | แนวร่วม | 22 | 72 | ชุด | 18 | 97 | แพทย์ | 15 |
| 23 | เฉลิมพระ เกียรติ | 34 | 48 | สู้ | 22 | 73 | ลิ้ม | 18 | 98 | เผย | 15 |
| 24 | ไทย | 34 | 49 | อาการ | 22 | 74 | รอ | 18 | 99 | เรียกร้อง | 15 |
| 25 | เอฟ | 33 | 50 | อ่าน | 21 | 75 | ประชาชน | 18 | 100 | รายละเอียด | 15 |

ตารางที่ 4.22

คำที่ปรากฏร่วมกันบ่อยที่สุดในข้อความกลุ่มที่ 3 (ปรากฏร่วมกัน 2 คำ)

| คำ, hashtag และ กลุ่ม ของชื่อเฉพาะที่มีการ ปรากฏร่วมกัน | ตัวอย่างข้อความ |
|---|---|
| คำ | |
| ['เตียง', 'โรงพยาบาลสนาม'] | - มธ.พนัก รพ.ธรรมศาสตร์เฉลิมพระเกียรติ ประกาศเปิด รพ.สนาม รอบใหม่ เพิ่มความจุจาก 308 เตียงเป็น 470 เตียง พร้อมรับคนไข้ 11 เม.ย.นี้ #ThaiPBSnews #COVID19 #มธ #โรงพยาบาลสนาม |
| ['ผู้ป่วย', 'โรงพยาบาลสนาม'] | - โรงพยาบาลสนามธรรมศาสตร์ 470 เตียง พร้อมแล้วในการรับส่งต่อผู้ป่วยโค วิดจากทุกโรงพยาบาลในเช้าวันอาทิตย์ที่ 11 เมษายน 64 - ผากเรื่องเตียงด้วยนะคะทุกคน มธยังพอรับได้แต่คาดว่าเหลือไม่มากเพราะทาง รพต่างๆเริ่มจองส่งต่อผู้ป่วยมาที่นี่แล้วนะคะ ติดต่อเพจ Thammasat well being center #ClubLOVEiS มธ.เตรียมเปิด โรงพยาบาลสนาม 470 เตียง |
| Hashtag | |
| ['#โควิด19', 'เมษายน'] | - BREAKING : เจ้าหน้าที่ไทยลีก ติดเชื้อ #โควิด19 ได้ปฏิบัติหน้าที่ในการ แข่งขันฟุตบอลช้าง เอเฟอ คัพ นัดชิงชนะเลิศ ที่สนาม ม.ธรรมศาสตร์ รังสิต เมื่อ 11 เม.ย. แต่ไม่ได้ติดต่อกับนักกีฬา จนท. หรือบุคคลภายนอก ทำในส่วนธุรการ อยู่ในห้อง จนท. จัดการแข่งขัน - สัญญาพุดิ เกิดบัณฑิต สมาคมธุรกิจท่องเที่ยวภายในประเทศ . จากเสวนา หัวข้อ "ผลกระทบของโควิด-19 ต่อธุรกิจภาคบริการ" เมื่อ 2 เม.ย. 2564 โดย ศูนย์ศึกษาสังคมและวัฒนธรรมร่วมสมัย คณะสังคมวิทยาและมานุษยวิทยา ม. ธรรมศาสตร์ . #โควิด19 #แรงงานใกล้ขีด #ธรรมศาสตร์ #ThelsaanRecord - 18.50 “ธรรมศาสตร์” ประกาศเปิด รพ.สนาม รอบใหม่เพิ่มความจุเป็น 470 เตียง พร้อมรับคนไข้ 11 เม.ย.นี้ #โควิดวันนี้ #โควิด19 #COVID19 |
| ['ตรวจ', '#โควิด19'] | - ใครอยู่โซนรังสิต สามารถไปตรวจโควิดฟรีได้ วันที่ 19 เมษายน เวลา 8.00- 15.00 ที่ plum condo ม.กรุงเทพ นะคะ ตรวจแบบ swab หรือถ้าเป็นที่ โรงพยาบาลธรรมศาสตร์สามารถกรอกฟอร์มได้นะคะ ถ้ามีอาการหรือเสี่ยง มากๆ เจ้าหน้าที่จะติดต่อมา เป็นตรวจแบบ swab เหมือนกันคะ #โควิด - ไทยพีบีเอสพบผู้ผลิตรายการสัญญาจ้าง ติดเชื้อ COVID-19 จำนวน 1 คน โดย ได้รับการตรวจและยืนยันจากแพทย์ว่าพบเชื้อ #โควิด19 และเข้ารับการรักษา ตัวใน รพ.ธรรมศาสตร์เฉลิมพระเกียรติ เมื่อวันที่ 6 เม.ย. ที่ผ่านมา อ่าน เพิ่มเติม: #ThaiPBS #COVID19 #โควิด |

ตารางที่ 4.22

คำที่ปรากฏร่วมกันบ่อยที่สุดในข้อความกลุ่มที่ 3 (ปรากฏร่วมกัน 2 คำ)(ต่อ)

| กลุ่มของชื่อเฉพาะ | |
|------------------------------------|---|
| ['ศุณย์', 'ชื่อพื้นที่'] | <p>- เอาจริงตอนตั้งศุณย์ฉุกเฉินสีนามิ ที่ธรรมศาสตร์ รังสิต คือทุกอย่างสามารถเสกออกมาได้ภายใน 3 ชั่วโมงเท่านั้น มีระบบรับลงทะเบียน ระบบจัดการห้องพัก ระบบรับบริจาค ระบบสวัสดิการทั้งอาสาสมัครและผู้ประสบภัย + ประสานงานหน่วยงานราชการต่างๆ แล้วทุกเย็นก็มาประเมินกันว่าพรุ่งนี้ต้องปรับอะไรอีก</p> <p>- เพื่อนักสู้ไม่ได้รับความเป็นธรรม มวลชนที่นั่งอดอาหารหน้าศาลถูกจับ พระที่ออกมาเรียกร้องถูกคุกคาม . “ กลัวอะไรแค่กับคนที่ออกมาเย็นเฉยๆ!? “ . ถึงเวลาแล้วที่เราจะออกมาแสดงพลัง! ขอเชิญร่วมกันยืน หยุด ชั่ง 112 นาที ณ เชียงราก 1 ธรรมศาสตร์ ศุณย์รังสิต เวลา 17:00 น. เป็นต้นไป</p> |
| ['ชื่อโรคติดต่อ', 'โรงพยาบาลสนาม'] | <p>- รพ.ธรรมศาสตร์ ผ่ากประชาสัมพันธ์รับ ตอนนี้อยู่ที่ รพ.กำลังเตรียม รพ.สนามเพื่อรองรับผู้ป่วยโควิด ใครที่ประสงค์จะบริจาคของ ตอนนี้อยู่ที่ รพ.ต้องการสบู่อล้างมือ ไม้กวาด ยากันยุง ไม้ช้อนตักช้อน สามารถส่งทางไปรษณีย์ หรือไปส่งที่อาคาร มรว. สุวพรรณ (P2 ในแผนที่) ขอขอบคุณครับ</p> <p>- เสริมจริงๆแล้วโรงพยาบาลสนามธรรมศาสตร์ ไม่ใช่ตึกในส่วนของรพ.มธ. แต่เป็นโรงแรมของธรรมศาสตร์ หาเงินเข้ามาหลาย แต่ปัจจุบันได้เปลี่ยนจากโรงแรมมาเป็นรพ.สนามรองรับผู้ป่วยโควิด แม้ขาดรายได้แต่มธ. ก็ไม่ทอดทิ้งประชาชนตามคติของมธ. ฉันรักธรรมศาสตร์ เพราะธรรมศาสตร์สอนให้ฉันประชาชน</p> <p>- โรงพยาบาลธรรมศาสตร์เฉลิมพระเกียรติ ขอเชิญร่วมบริจาคเพื่อต้านภัยโควิด สมทบทุนซื้ออุปกรณ์เครื่องมือทางการแพทย์ หรือ สามารถบริจาคสิ่งของ รับมือกับสถานการณ์COVID-19 ณ โรงพยาบาลสนามธรรมศาสตร์ เมื่อสแกนบริจาคผ่านQR CODEระบบจะส่งข้อมูลไปที่กรมสรรพากร ใช้ในการลดหย่อนภาษีได้ 2 เท่า</p> |

ตารางที่ 4.23

คำที่ปรากฏร่วมกันบ่อยที่สุดในข้อความกลุ่มที่ 3 (ปรากฏร่วมกัน 3 คำ)

| คำ, hashtag และ กลุ่มของ ชื่อเฉพาะที่มีการปรากฏ ร่วมกัน | ตัวอย่างข้อความ |
|---|---|
| คำ | |
| ['รพ.', 'ผู้ป่วย', 'โรงพยาบาล สนาม'] | <ul style="list-style-type: none"> - รพ.ธรรมศาสตร์ ฝากประชาสัมพันธ์ครับ ตอนนี้ทาง รพ.กำลังเตรียม รพ.สนาม เพื่อรองรับผู้ป่วยโควิด ใครที่ประสงค์จะบริจาคของ ตอนนี้ทาง รพ.ต้องการสบู่อล้างมือ ไม้กวาด ยาแก้นุง ไม้ช้อนตักครับ สามารถส่งทางไปรษณีย์ หรือไปส่งที่ อาคาร มรว. สุพรรณ (P2 ในแผนที่) ขอขอบคุณครับ - เสริมจริงๆแล้วโรงพยาบาลสนามธรรมศาสตร์ ไม่ใช่ตี๊กในส่วนของรพ.มธ. แต่เป็นโรงแรมของธรรมศาสตร์ หาเงินเข้ามาหลาย แต่ปัจจุบันได้เปลี่ยนจากโรงแรมมาเป็นรพ.สนามรองรับผู้ป่วยโควิด แม้ขาดรายได้แต่มธ. ก็ไม่ทอดทิ้งประชาชน ตามคติของมธ. ฉันทักธรรมศาสตร์ เพราะธรรมศาสตร์สอนให้ฉันทักประชาชน - อธิบดีกรมแพทย์ เผย ข้อมูลการครองเตียงผู้ป่วยโควิด เติงไม่เต็มตามที่รพ. เอกชน. อ้าง ชี้เหลือถึง 300 เติง และยังหาเตียงเสริมทั้งรพ.สนามจากรพ.มธ. |
| ['โล หิต ', 'อา คาร ', 'บริจาค'] | <ul style="list-style-type: none"> - วิกฤตหนัก!! โรงพยาบาลธรรมศาสตร์เฉลิมพระเกียรติ เกิดสถานการณ์วิกฤต “โลหิตขาดแคลนทุกกรุป” ขอเชิญทุกท่านร่วมบริจาคโลหิต ได้ที่ หน่วยรับบริจาคโลหิต ณ อาคารพลังใจ เปิดรับบริจาคทุกวัน (ยกเว้นวันหยุดนักขัตฤกษ์) เวลา 8.30-15.30 น. ค่ะ #Thammasat #ThammasatNews - #วิกฤตโลหิตขาดแคลนทุกกรุป ขณะนี้โรงพยาบาลธรรมศาสตร์เฉลิมพระเกียรติ เกิดสถานการณ์วิกฤตโลหิตขาดแคลนทุกกรุป จึงขอเชิญร่วมบริจาคโลหิต เพื่อต่อชีวิตและลมหายใจให้กับผู้ป่วย ทุกวันจันทร์-อาทิตย์ เวลา 08.30-15.30 น. (ไม่ปิดพักเที่ยง) ณ ชั้น 2 อาคารพลังใจ #Thammasat - เนื่องจากสถานการณ์ COVID-19 ทำให้ผู้บริจาคโลหิตลดลงและปริมาณโลหิตคงคลังก็ลดน้อยลงจนถึงขั้นขาดแคลน บริจาคโลหิต ณ อาคารพลังใจ โรงพยาบาลธรรมศาสตร์เฉลิมพระเกียรติ ทุกวัน เวลา 8.30-15.30 น. |
| Hashtag | |
| ['ติด', 'ชื่อโรคติดต่อ', '#โค วิด19'] | <ul style="list-style-type: none"> - ใครอยู่โซนรังสิต สามารถไปตรวจโควิดฟรีได้ วันที่ 19 เมษายน เวลา 8.00-15.00 ที่ plum condo ม.กรุงเทพ นะคะ ตรวจแบบ swab หรือถ้าเป็นที่โรงพยาบาลธรรมศาสตร์สามารถกรอกฟอร์มได้นะคะ ถ้ามีอาการหรือเสี่ยงมากๆ เจ้าหน้าที่จะติดต่อมา #โควิด #โควิด19 #ตรวจโควิดฟรี - ที่คณะชั้น(รัตสาดพ)มีคนติดโควิด นางไปก้า มธ.มาเมื่อวันที่ 26 มี.ค. จากนั้นไปอีกมากมายหลายที่ ใครผ่านมาเห็นลองเช็คตามไลน์กันดูน้า |

ตารางที่ 4.23

คำที่ปรากฏร่วมกันบ่อยที่สุดในข้อความกลุ่มที่ 3 (ปรากฏร่วมกัน 3 คำ)(ต่อ)

| Hashtag(ต่อ) | |
|--|--|
| <p>['ซัง', 'หยุด', '#ปล่อยเพื่อนเรา']</p> | <ul style="list-style-type: none"> - เริ่มแล้ว! กิจกรรมยื่น หยุด ซัง ณ บริเวณประตูเชียงใหม่ 1 ธรรมศาสตร์ ศูนย์รังสิต สามารถเข้าร่วมได้ตลอดเวลากิจกรรม โดยกิจกรรมจะจบลงเวลา 18:12 น. มาร่วมเรียกร้องความยุติธรรมคืนสู่ประเทศได้แล้ว ณ ตอนนี! - จัดไปอย่าหยุด! ขอเชิญเพื่อน ๆ ร่วม ยื่น หยุด ซัง 1 ชั่วโมง 12 นาที กับพวกเรา ณ เชียงราก 1 ธรรมศาสตร์ รังสิต เวลา 16:30 น. เป็นต้นไป! *กรุณาเว้นระยะห่างและสวมหน้ากากอนามัยตลอดเวลาทำกิจกรรมเพื่อป้องกันการแพร่ระบาดของโรค* #ยื่นหยุดซัง #ปล่อยเพื่อนเรา - กลับมาแล้วกับ ยื่น หยุด ซัง DAY3 ขอเชิญเพื่อน ๆ มาร่วมยื่นเพื่อแสดงออกเชิงสัญลักษณ์ว่า "เราจะไม่ยอมแพ้ต่อความไม่ยุติธรรม" ณ บริเวณเชียงใหม่ 1 ธรรมศาสตร์ รังสิต เวลา 17:00 น. เป็นต้นไป! #ยื่นหยุดซัง #ปล่อยเพื่อนเรา |
| กลุ่มของชื่อเฉพาะ | |
| <p>['ผู้ป่วย', 'ชื่อโรคติดต่อ', 'โรงพยาบาลสนาม']</p> | <ul style="list-style-type: none"> - รพ.ธรรมศาสตร์ ฝากประชาสัมพันธ์ครับ ตอนนี้ทาง รพ.กำลังเตรียม รพ.สนามเพื่อรองรับผู้ป่วยโควิด ใครที่ประสงค์จะบริจาคของ ตอนนี้ทาง รพ.ต้องการสบู่ล้างมือ ไม้กวาด ยาแก้นุง ไม้ช้อนตักครับ สามารถส่งทางไปรษณีย์ หรือไปส่งที่อาคาร มรว. สุวพรรณ (P2 ในแผนที่) ขอขอบคุณครับ - เสริมจริงๆแล้วโรงพยาบาลสนามธรรมศาสตร์ ไม่ใช่ตึกในส่วนของรพ.มธ. แต่เป็นโรงแรมของธรรมศาสตร์ หาเงินเข้ามาหลาย แต่ปัจจุบันได้เปลี่ยนจากโรงแรมมาเป็นรพ.สนามรองรับผู้ป่วยโควิด แม้ขาดรายได้แต่มธ. ก็ไม่ทอดทิ้งประชาชนตามคติของมธ. ฉันรักธรรมศาสตร์ เพราะธรรมศาสตร์สอนให้ฉันประชาชน - โรงพยาบาลสนามธรรมศาสตร์ 470 เต็มพร้อมแล้วในการรับส่งต่อผู้ป่วยโควิดจากทุกโรงพยาบาลในเช้าวันอาทิตย์ที่ 11 เมษายน 64 |
| <p>['เตียง', 'ชื่อโรคติดต่อ', 'โรงพยาบาลสนาม']</p> | <ul style="list-style-type: none"> - โรงพยาบาลสนามธรรมศาสตร์ 470 เต็มพร้อมแล้วในการรับส่งต่อผู้ป่วยโควิดจากทุกโรงพยาบาลในเช้าวันอาทิตย์ที่ 11 เมษายน 64 - อธิบดีกรมแพทย์ เผย ข้อมูลการครองเตียงผู้ป่วยโควิด เตียงไม่เต็มตามที่รพ.เอกชน. อ้าง ชี้เหลือถึง 300 เตียง และยังหาเตียงเสริมทั้งรพ.สนามจากรพ.มธ. และhospital เพิ่ม. เชื้อเพียงพอรับมือได้ - ขอเล่าจริงๆแล้ว รพ.สนามมธ. ไม่ใช่เตียงในรพ.มธ. แต่เป็นหอพักของนศ.และบุคลากรค่ะ เวลาสร้างรพ.สนามมธ. นศ.กับบุคลากร จะย้ายออกไปในที่ ที่มหาลัยรองรับใหม่ค่ะ เพื่อให้หอพักที่โรงแรมนี้ว่าง และสามารถรับรองคนไข้โควิดได้ตามแบบของชาวมธ. ฉันรักธรรมศาสตร์เพราะธรรมศาสตร์สอนให้ฉันรักประชาชน |

ตารางที่ 4.24

คำที่ปรากฏร่วมกันบ่อยที่สุดในข้อความกลุ่มที่ 3 (ปรากฏร่วมกัน 4 คำ)

| คำ, hashtag และ กลุ่มของ ชื่อเฉพาะที่มีการปรากฏ ร่วมกัน | ตัวอย่างข้อความ |
|---|---|
| คำ | |
| [‘ยืน’, ‘ซัง’, ‘หน้า’, ‘หยุด’] | <p>- เริ่มแล้ว! กิจกรรมยืน หยุด ซัง ณ บริเวณประตูเชียงราก1 ธรรมศาสตร์ ศูนย์รังสิต สามารถเข้าร่วมได้ตลอดเวลากิจกรรม โดยกิจกรรมจะจบลงเวลา 18:12 น. มาร่วมเรียกร้องความยุติธรรมคืนสู่ประเทศได้แล้ว ณ ตอนนี! #ยืนหยุด</p> <p>- [กรุณาแชร์โพสต์นี้ต่อ] . จัดไปอย่าหยุด! ขอเชิญเพื่อน ๆ ร่วม ยืน หยุด ซัง 1 ชั่วโมง 12 นาที กับพวกเรา ณ เชียงราก 1 ธรรมศาสตร์ รังสิต เวลา 16:30 น. เป็นต้นไป! . *กรุณาเว้นระยะห่างและสวมหน้ากากอนามัยตลอดเวลาทำกิจกรรมเพื่อป้องกันการแพร่ระบาดของโรค* #ยืนหยุดซัง #ปล่อยเพื่อนเรา</p> <p>- “เมื่อกฎหมายกลายเป็นเครื่องมือปิดปาก พวกเราจึงต้องลุกขึ้นมา ไม่ยอมรับมัน!” . กิจกรรม ยืน หยุด ซัง 1 ชั่วโมง 12 นาที ในวันนี้เริ่มขึ้นแล้ว! โดยเราจะทำกิจกรรมถึงเวลา 18:12 น. . ท่านใดที่สนใจยังสามารถมาเข้าร่วมได้ทันที ณ บริเวณเชียงราก1 ธรรมศาสตร์ ศูนย์รังสิต . #ยืนหยุดซัง</p> |
| [‘โลहित’, ‘อาคาร’, ‘เวลา’, ‘บริจาค’] | <p>- วิกฤตหนัก!! โรงพยาบาลธรรมศาสตร์เฉลิมพระเกียรติ เกิดสถานการณ์วิกฤต “โลहितขาดแคลนทุกรูป” ขอเชิญทุกท่านร่วมบริจาคโลหิต ได้ที่ หน่วยรับบริจาคโลหิต ณ อาคารพลังใจ เปิดรับบริจาคทุกวัน (ยกเว้นวันหยุดนักขัตฤกษ์) เวลา 8.30-15.30 น. ค่ะ #Thammasat #ThammasatNews</p> <p>- #วิกฤตโลหิตขาดแคลนทุกรูป ขณะนี้โรงพยาบาลธรรมศาสตร์เฉลิมพระเกียรติ เกิดสถานการณ์วิกฤตโลหิตขาดแคลนทุกรูป จึงขอเชิญร่วมบริจาคโลหิตเพื่อต่อชีวิตและลมหายใจให้กับผู้ป่วย ทุกวันจันทร์-อาทิตย์ เวลา 08.30-15.30 น. (ไม่ปิดพักเที่ยง) ณ ชั้น 2 อาคารพลังใจ #Thammasat #ThammasatNews</p> <p>- เนื่องจากสถานการณ์ COVID-19 ทำให้ผู้บริจาคโลหิตลดลงและปริมาณโลหิตคงคลังก็ลดน้อยลงจนถึงขั้นขาดแคลน บริจาคโลหิต ณ อาคารพลังใจ โรงพยาบาลธรรมศาสตร์เฉลิมพระเกียรติ ทุกวัน เวลา 8.30-15.30 น. ไม่หยุดเสาร์และอาทิตย์ เว้นวันหยุดนักขัตฤกษ์ รายละเอียดเพิ่มเติม:</p> |

ตารางที่ 4.24

คำที่ปรากฏร่วมกันบ่อยที่สุดในข้อความกลุ่มที่ 3 (ปรากฏร่วมกัน 4 คำ)(ต่อ)

| Hashtag | |
|---|---|
| <p>['ยืน', 'ซัง', 'หยุด', '#ปล่อยเพื่อนเรา']</p> | <p>- เริ่มแล้ว! กิจกรรมยืน หยุด ซัง ณ บริเวณประตูเชียงใหม่ 1 ธรรมศาสตร์ ศูนย์รังสิต สามารถเข้าร่วมได้ตลอดเวลา กิจกรรม โดยกิจกรรมจะจบลงเวลา 18:12 น. มาร่วมเรียกร้องความยุติธรรมคืนสู่ประเทศได้แล้ว ณ ตอนนี! #ยืนหยุดซัง #ปล่อยเพื่อนเรา</p> <p>- [กรุณาแชร์โพสต์นี้ต่อ] . จัดไปอย่าหยุด! ขอเชิญเพื่อน ๆ ร่วม ยืน หยุด ซัง 1 ชั่วโมง 12 นาที กับพวกเรา ณ เชียงราก 1 ธรรมศาสตร์ รังสิต เวลา 16:30 น. เป็นต้นไป! . *กรุณาเว้นระยะห่างและสวมหน้ากากอนามัยตลอดเวลาทำกิจกรรมเพื่อป้องกันการแพร่ระบาดของโรค* #ยืนหยุดซัง #ปล่อยเพื่อนเรา</p> <p>- กลับมาแล้วกับ ยืน หยุด ซัง DAY3 ขอเชิญเพื่อน ๆ มาร่วมยืนเพื่อแสดงออกเชิงสัญลักษณ์ว่า "เราจะไม่ยอมแพ้ต่อความไม่ยุติธรรม" ณ บริเวณเชียงใหม่ 1 ธรรมศาสตร์ รังสิต เวลา 17:00 น. เป็นต้นไป! #ยืนหยุดซัง #ปล่อยเพื่อนเรา</p> |
| <p>['ชื่อพื้นที่', 'ชื่อโรคติดต่อ', '#โควิด19', 'เมษายน']</p> | <p>- ใครอยู่โซนรังสิต สามารถไปตรวจโควิดฟรีได้ วันที่ 19 เมษายน เวลา 8.00-15.00 ที่ plum condo ม.กรุงเทพ นะคะ ตรวจแบบ swab หรือถ้าเป็นที่โรงพยาบาลธรรมศาสตร์สามารถกรอกฟอร์มได้้นะคะ ถ้ามีอาการหรือเสี่ยงมากๆ เจ้าหน้าที่จะติดต่อมา เป็นตรวจแบบ swab เหมือนกันค่ะ #โควิด #โควิด19 #ตรวจโควิดฟรี</p> <p>- คณะรัฐศาสตร์ มธ. ยืนยัน นร.สอบ TCAS ติดโควิด-19 แจ้งผู้เข้าสอบ คนคุมสอบ สังเกตอาการ เจ้าตัวยืนยันติดโควิด-19 พร้อมแจ้ง Timeline เข้าสนามสอบ มธ. ทำพระจันทร์-รังสิต วันที่ 9-10 เม.ย. #TCAS64 #โควิด19</p> <p>- เทรดรวมสถานที่ตรวจโควิดฟรี (3) 5. ม.ธรรมศาสตร์ รังสิต ตรวจเชิงรุกวันที่ 20 เมษา เริ่มเวลา 9 โมง #โควิด19 #โควิด19วันนี้ #โควิดวันนี้</p> |
| กลุ่มของชื่อเฉพาะ | |
| <p>['ระบอบ', 'ชื่อบุคคล/กลุ่มบุคคลที่เกี่ยวข้องกับการเมือง', 'ชื่อสถานที่ในมหาวิทยาลัย', 'ชื่อคณะ']</p> | <p>- พี่เจี๊ยบ ก็มาร่วมเป็นหนึ่งในชื่อ รื้อระบอบประยุทธ์ หยุดยั้งเวลาเผด็จการ! มาลงชื่อกันได้อย่างต่อเนื่อง @ คณะนิติศาสตร์ มธ. ทำพระจันทร์ ดุรายละเอียดและโหลดแผ่นพับที่นี่ :</p> <p>- มาร่วมเป็นหนึ่งในชื่อรื้อระบอบประยุทธ์กันตอนนี้! ก็เป็นหนึ่งในชื่อแล้ว มาลงชื่อกัน พบกันที่ คณะนิติศาสตร์ มธ. ทำพระจันทร์</p> <p>- ได้สถานที่ใหม่แล้ว! หลังโดนหอศิลป์เท @ ห้อง LT1 ชั้น 1 คณะนิติศาสตร์ มหาวิทยาลัยธรรมศาสตร์ ทำพระจันทร์ เปิดตัวแคมเปญใหม่สร้างแรงกระเพื่อมให้การเมืองไทย เดินหน้าเข้าชื่อทั่วประเทศแก้ไขรัฐธรรมนูญ เตรียมบัตรประชาชนให้พร้อม มาร่วมลงชื่อพร้อมกันครั้งแรก รื้อระบอบประยุทธ์</p> |

ตารางที่ 4.24

คำที่ปรากฏร่วมกันบ่อยที่สุดในข้อความกลุ่มที่ 3 (ปรากฏร่วมกัน 4 คำ)(ต่อ)

| กลุ่มของชื่อเฉพาะ(ต่อ) | |
|--|---|
| ['ผู้ป่วย', 'เตียง', 'ชื่อโรคติดต่อ', 'โรงพยาบาล'] | <ul style="list-style-type: none"> - โรงพยาบาลสนามธรรมศาสตร์ 470 เตียง พร้อมแล้วในการรับส่งต่อผู้ป่วยโควิดจากทุกโรงพยาบาลในเช้าวันอาทิตย์ที่ 11 เมษายน 64 - อธิบดีกรมแพทย์ เผย ข้อมูลการครองเตียงผู้ป่วยโควิด เตียงไม่เต็มตามที่รพ. เอกชน. อ้าง ชี้เหลือถึง 300 เตียง และยังหาเตียงเสริมทั้งรพ.สนามจากรพ.มธ. และhospital เพิ่ม. เชื้อเพียงพอรับมือได้ - 11 เม.ย 64 รพ.สนามธรรมศาสตร์ รับผู้ป่วยโควิด-19 วันแรก ขยายศักยภาพเพิ่มเป็น 470 เตียง รองรับผู้ป่วยทั้ง ปทุมธานี พื้นที่ใกล้เคียง และ กทม. |

จากตารางที่ 4.22, 4.23 และ 4.24 พบคำที่ปรากฏร่วมกันได้แก่ ['ศูนย์', 'ชื่อพื้นที่'], ['ชื่อโรคติดต่อ', 'โรงพยาบาลสนาม'], ['รพ.', 'ผู้ป่วย', 'โรงพยาบาลสนาม'], ['โลหิต', 'อาคาร', 'บริจาค'], ['ติด', 'ชื่อโรคติดต่อ', '#โควิด19'], ['ซัง', 'หยุด', '#ปล่อยเพื่อนเรา'], ['ผู้ป่วย', 'ชื่อโรคติดต่อ', 'โรงพยาบาลสนาม'], ['เตียง', 'ชื่อโรคติดต่อ', 'โรงพยาบาลสนาม'], ['ยื่น', 'ซัง', 'หน้า', 'หยุด'], ['โลหิต', 'อาคาร', 'เวลา', 'บริจาค'], ['ยื่น', 'ซัง', 'หยุด', '#ปล่อยเพื่อนเรา'], ['ชื่อพื้นที่', 'ชื่อโรคติดต่อ', '#โควิด19', 'เมษายน'], ['ระบอบ', 'ชื่อบุคคล/กลุ่มบุคคลที่เกี่ยวข้องกับการเมือง', 'ชื่อสถานที่ในมหาวิทยาลัย', 'ชื่อคณะ'], ['ผู้ป่วย', 'เตียง', 'ชื่อโรคติดต่อ', 'โรงพยาบาลสนาม'] จากคำที่ปรากฏร่วมกัน และตัวอย่างของข้อความ สามารถสรุปหัวข้อของข้อความ กลุ่มที่ 3 พบว่า ข้อเกี่ยวข้องกับโรคโคโรนาไวรัส 2019 และยังพบอีกว่าข้อความในกลุ่มนี้บางส่วนเป็นข้อความที่เกี่ยวข้องกับการเมือง

4.4.4 การสรุปหัวข้อของกลุ่มข้อความ กลุ่มที่ 4

การสรุปหัวข้อของกลุ่มข้อความ กลุ่มที่ 4 โดยการพิจารณาจากคำที่ปรากฏร่วมกันของคำที่มีความถี่สูงสุดข้อความกลุ่มที่ 4 ที่ปรากฏ 100 อันดับแรก ดังตารางที่ 4.25

ตารางที่ 4.25

คำที่มีความถี่สูงสุด 100 อันดับแรกในข้อความกลุ่มที่ 4

| ลำดับ | คำ | ความถี่ | ลำดับ | คำ | ความถี่ | ลำดับ | คำ | ความถี่ | ลำดับ | คำ | ความถี่ |
|-------|----------------|---------|-------|-------------------|---------|-------|------------------|---------|-------|----------------|---------|
| 1 | คำ | 240 | 26 | เดือน | 41 | 51 | ใบ | 23 | 76 | ออนไลน์ | 18 |
| 2 | ชื่อพื้นที่ | 156 | 27 | สอบถาม | 39 | 52 | ชื่อนักแสดง | 23 | 77 | รอบ | 17 |
| 3 | ขาย | 117 | 28 | Line | 36 | 53 | คอร์ส | 23 | 78 | คับ | 17 |
| 4 | บาท | 102 | 29 | #dek66 | 35 | 54 | ชื่อสถานที่สำคัญ | 23 | 79 | ง่าย | 17 |
| 5 | #Dek64 | 97 | 30 | มือสอง | 35 | 55 | ส่งต่อ | 22 | 80 | รูป | 17 |
| 6 | ราคา | 93 | 31 | ดู | 35 | 56 | หน้า | 20 | 81 | ไทย | 17 |
| 7 | สนใจ | 82 | 32 | เมท | 35 | 57 | อ่าน | 20 | 82 | สอน | 16 |
| 8 | ห้อง | 77 | 33 | #หอมธ | 32 | 58 | หาร | 20 | 83 | เรื่อง | 16 |
| 9 | #dek65 | 70 | 34 | ค้าบ | 32 | 59 | มหาลัย | 20 | 84 | เตรียม | 16 |
| 10 | #ทอนอมธ | 70 | 35 | หนังสือ | 31 | 60 | เริ่มต้น | 20 | 85 | คิว | 16 |
| 11 | หอ | 69 | 36 | ร้าน | 29 | 61 | รบกวน | 20 | 86 | น้องๆ | 16 |
| 12 | ทัก | 68 | 37 | เทอม | 29 | 62 | ตอนนี้ | 19 | 87 | เค้า | 16 |
| 13 | หา | 63 | 38 | #หอพักมธ | 29 | 63 | #หนังสือมือสอง | 19 | 88 | แลก | 16 |
| 14 | สัญญา | 59 | 39 | สภาพ | 28 | 64 | ติด | 19 | 89 | รายละเอียด | 16 |
| 15 | ชื่อมหาลัย | 58 | 40 | เรียน | 27 | 65 | ปล่อย | 19 | 90 | การ์ด | 16 |
| 16 | งาน | 58 | 41 | นักศึกษา | 27 | 66 | #รัฐศาสตร์มธ | 19 | 91 | #ขายสัญญาหอมธ | 16 |
| 17 | นัด | 57 | 42 | #หอโนมธ | 26 | 67 | #GATPAT | 19 | 92 | เมษายน | 15 |
| 18 | ชื่อ | 57 | 43 | ปี | 25 | 68 | ชื่อคณะ | 19 | 93 | #จุฬา | 15 |
| 19 | น้ำ | 56 | 44 | #หาเมทมธ | 25 | 69 | #มธรังสิต | 19 | 94 | หูน | 15 |
| 20 | Direct message | 54 | 45 | ดี | 25 | 70 | #นิติมธ | 18 | 95 | #มอกรุงเทพ | 15 |
| 21 | เล่ม | 50 | 46 | #หนังสือเตรียมสอบ | 25 | 71 | สวัสดี | 18 | 96 | #สอนพิเศษ | 15 |
| 22 | เกาะ | 49 | 47 | #หนังสือเตรียมสอบ | 24 | 72 | ลด | 18 | 97 | #ตลาดนัดnct | 15 |
| 23 | ทำ | 43 | 48 | คุย | 24 | 73 | ล้าน | 18 | 98 | #ส่งต่อหนังสือ | 15 |
| 24 | ฟรี | 42 | 49 | ฝั่ง | 24 | 74 | หญิง | 18 | 99 | พ.ค. | 15 |
| 25 | สอบ | 41 | 50 | พี่ | 23 | 75 | คนละ | 18 | 100 | บ้าน | 15 |

ตารางที่ 4.26

คำที่ปรากฏร่วมกันบ่อยที่สุดในข้อความกลุ่มที่ 4 (ปรากฏร่วมกัน 2 คำ)

| คำ, hashtag และ กลุ่มของ ชื่อเฉพาะที่มีการปรากฏ ร่วมกัน | ตัวอย่างข้อความ |
|---|---|
| คำ | |
| ['ค่า', 'สนใจ'] | <p>- มีพร้อมส่ง 200 แพ็ค • แพ็คละ 30 บาทค่า ขึ้นต่ำ 3 แพ็คคะ นั้รับได้ ธรรมชาติ-รังสิต พิวเจอร์ คลอง1-6 นวนคร พระอินทร์ • ส่งตจว 5 แพ็คขึ้น ไป ค่าส่ง 70 ค่ะ สนใจเดมมาได้เลยนะคะ //รับซื้อบิลด้วยนะคะ ให้65% เลย ค่า #BKBeatตัวท็อป #BKxBamBam #BKBeatYourBe #bambam</p> <p>- ฝากหน่อยค่า ตามหานักแสดงหนังสั้นนักศึกษา มธ. *ค่าตัว 4k-5k Shooting date : 29 พ.ค. - มิ.ย. (ถ่าย 2 วัน) Casting date 10 เม.ย. @ ธรรมชาติ รังสิต 24 - 25 เม.ย. @ กทม (แจ้งสถานที่อีกครั้ง) สนใจติดต่อ Line : fern- fernny</p> <p>- ขายสัญญาหอเพลินอพาร์ตเมนต์ ในซอยเจปาร์ค มีพิตเนสฟรี สระว่ายน้ำ มี สนามแบดมินตันราคาสัญญา 11,000 บาท ห้องพักเดือนละ 5,500 มีรถตู้รับส่ง เข้ามอ มีตู้เย็น ทีวี พร้อมออกค่าเปลี่ยนชื่อสัญญาให้ #หอนอกมธ #หอมธ #มธ #ขายสัญญาหอ สนใจทักdm มาได้เลยคะ พร้อมเข้าอยู่ได้เดือนพฤษา ไม่ค่า ส่วนกลาง</p> |
| [‘ขาย’, ‘ค่า’] | <p>- ขายสัญญาหอเพลินอพาร์ตเมนต์ ในซอยเจปาร์ค มีพิตเนสฟรี สระว่ายน้ำ มี สนามแบดมินตันราคาสัญญา 11,000 บาท ห้องพักเดือนละ 5,500 มีรถตู้รับส่ง เข้ามอ มีตู้เย็น ทีวี พร้อมออกค่าเปลี่ยนชื่อสัญญาให้ #หอนอกมธ #หอมธ #มธ #ขายสัญญาหอ สนใจทักdm มาได้เลยคะ พร้อมเข้าอยู่ได้เดือนพฤษา ไม่ค่า ส่วนกลาง</p> <p>- ถ้าสอบติดธรรมชาติ หรือ จุฬา อยากส่งต่อหนังสือให้กับน้องๆ บางเล่มก็ ขายถูกๆ บางเล่มก็ให้ฟรีๆคิดแค่ค่าส่งเลยค่า เห็นรุ่นน้องบางคนโพสแล้วเข้าใจ ฟิลตัวเองเลย #dek64</p> <p>- ขายสัญญาหอCity Parkฝั่งหญิง ชั้น9 ระเบียบหันเข้าอินเตอร์ อยู่ตรงช่องลม ห้องไม่ร้อน ห้อง7000/เดือน (ช่วงโควิดเหลือ5500) ค่าสัญญา16000 (ได้คืน ตอนออก) ค่าเปลี่ยนชื่อสัญญา 1000 เข้าอยู่ได้เลย !ซื้อภายในเดือนนี้เดือนเมษา อยู่ฟรี! สนใจdm/line #หอมธ #หอนอกมธ #สัญญาหอมธ #cityparkมธ # ทิมมธ</p> |

ตารางที่ 4.26

คำที่ปรากฏร่วมกันบ่อยที่สุดในข้อความกลุ่มที่ 4 (ปรากฏร่วมกัน 2 คำ)(ต่อ)

| hashtag | |
|------------------------|--|
| ['ค่า', '#หอนอกมธ'] | - หามทผู้หญิง หารห้องหอ 18-22 หอสร้างใหม่ฝั่งตรงข้ามมธค่าบ ห้อง studio plus 8,500 หารตกคนละ 4,250 บาทค่า เริ่มเข้าอยู่เดือนสิงหาคะ เปิดเทอม1 หรือใครสะดวกหอรห้องตั้งแต่ช่วงsummerก็ได้ ทักมาคุยกันก่อนได้เลยนะคะ เรออยู่ปี2ค่าบ #หอนอกมธ #หอพักมธ #Dek64 #ทีมมธ #หามทมธ #มธรังสิต |
| ['หอ', '#หอนอกมธ'] | - ขายสัญญาหอเพลินอพาร์ทเมนต์ ในซอยเจปาร์ค มีพิตเนสฟรี สระว่ายน้ำ มีสนามแบดมินตันราคาสัญญา 11,000 บาท ห้องพักเดือนละ 5,500 มีรถตู้รับส่ง เข้ามอ มีตู้เย็น ทีวี พร้อมออกค่าเปลี่ยนชื่อสัญญาให้ #หอนอกมธ #หอมธ #มธ #ขายสัญญาหอ สนใจทักdm มาได้เลยคะ พร้อมเข้าอยู่ได้เดือนพฤษา ไม่ค่า ส่วนกลาง - ขายสัญญาหอCity Parkฝั่งหญิง ชั้น9 ระเบียบหันเข้าอินเตอร์ อยู่ตรงช่องลม ห้องไม่ร้อน ห้อง7000/เดือน (ช่วงโควิดเหลือ5500) ค่าสัญญา16000 (ได้คืน ตอนออก) ค่าเปลี่ยนชื่อสัญญา 1000 เข้าอยู่ได้เลย !ซื้อภายในเดือนนี้เดือนเมษา อยู่ฟรี! สนใจdm/line #หอมธ #หอนอกมธ #สัญญาหอมธ #cityparkมธ #ทีมมธ |
| กลุ่มของชื่อเฉพาะ | |
| ['ค่า', 'ชื่อพื้นที่'] | - มีพร้อมส่ง 200 แพ็ค • แพ็คละ 30 บาทค่า ชั้นต่ำ 3 แพ็คคะ นั้ดรับได้ ธรรมชาติสร-รังสิต พิวจอร์ คลอง1-6 นวนคร พระอินทร์ • ส่งตจว 5 แพ็คขึ้น ไป ค่าส่ง 70 ค่ะ สนใจเดมมาได้เลยนะคะ //รับซื้อบิลด้วยนะคะ ให้65% เลย ค่า #BKBeatตัวท็อป #BKxBamBam #BKBeatYourBe #bambam |
| ['นัด', 'ชื่อพื้นที่'] | - พร้อมส่ง Lenticular pt.2 แจฮยอน 350฿ มาร์ค regular 250฿ แทง fire truck 250฿ ID เฮนเดอร์รี่ 180฿ ส่งไปร 30/50 ,นัดรับ มธ รังสิต สนใจ dm ค่าา #ตลาดนัดnct #แลกการ์ดnct - ขายเคส 11 pro max ค่า มีทั้ง2สีคะ กดไปแล้ว แต่จะเปลี่ยนทศ ราคา 1150 ค่าส่ง50/70(+10) นัดรับสยาม,เซนทรัลพระราม2,มธ.รังสิตได้คะ จ่ายเราชนะ/เรารักกัน(ม.33) ได้ค่า #REPRESENTXC3 #XC3 #representxmarktuan #ตลาดกัซ #ตลาดนัดอากาเซ #ตลาดนัดอากาเซ #ตลาดนัดอากาเซgot7 #ตลาดนัดgot7 |

ตารางที่ 4.27

คำที่ปรากฏร่วมกันบ่อยที่สุดในข้อความกลุ่มที่ 4 (ปรากฏร่วมกัน 3 คำ)

| คำ, hashtag และ กลุ่มของ ชื่อเฉพาะที่มีการปรากฏ ร่วมกัน | ตัวอย่างข้อความ |
|---|---|
| [คำ, 'สนใจ', 'Direct message'] | <p style="text-align: center;">คำ</p> <p>- มีพร้อมส่ง 200 แพ็ค • แพ็คละ 30 บาทค่า ชั้นต่ำ 3 แพ็คคะ นั้รับได้ ธรรมศาสตร์-รังสิต พิวเจอร์ คลอง1-6 นวนคร พระอินทร์ • ส่งตจว 5 แพ็คขึ้น ไป ค่าส่ง 70 ค่ะ สนใจเดมมาได้เลยนะคะ //รับซื้อบิลด้วยนะคะ ให้65% เลย ค่า #BKBeatตัวท็อป #BKxBamBam #BKBeatYourBe #bambam</p> <p>- ขายสัญญาหอเพลินอพาร์ทเมนต์ ในซอยเจปาร์ค มีพิตเนสฟรี สระว่ายน้ำ มี สนามแบดมินตันราคาสัญญา 11,000 บาท ห้องพักเดือนละ 5,500 มีรถตู้รับส่ง เข้ามา มีตู้เย็น ทีวี พร้อมออกค่าเปลี่ยนชื่อสัญญาให้ #หอนอกมธ #หอมธ #มธ #ขายสัญญาหอ สนใจทักdm มาได้เลยคะ พร้อมเข้าอยู่ได้เดือนพฤษา ไม่ค่า ส่วนกลาง</p> <p>- ขายสัญญาหอCity Parkฝั่งหญิง ชั้น9 ระเบียบหันเข้าอินเตอร์ อยู่ตรงช่องลม ห้องไม่ร้อน ห้อง7000/เดือน (ช่วงโควิดเหลือ5500) ค่าสัญญา16000 (ได้คืน ตอนออก) ค่าเปลี่ยนชื่อสัญญา 1000 เข้าอยู่ได้เลย !ชื่อภายในเดือนนี้เดือนเมษา อยู่ฟรี! สนใจdm/line #หอมธ #หอนอกมธ #สัญญาหอมธ #cityparkมธ # ทีมมธ</p> |
| [ขาย, 'ค่า', 'สัญญา'] | <p>- ขายสัญญาหอเพลินอพาร์ทเมนต์ ในซอยเจปาร์ค มีพิตเนสฟรี สระว่ายน้ำ มี สนามแบดมินตันราคาสัญญา 11,000 บาท ห้องพักเดือนละ 5,500 มีรถตู้รับส่ง เข้ามา มีตู้เย็น ทีวี พร้อมออกค่าเปลี่ยนชื่อสัญญาให้ #หอนอกมธ #หอมธ #มธ #ขายสัญญาหอ สนใจทักdm มาได้เลยคะ พร้อมเข้าอยู่ได้เดือนพฤษา ไม่ค่า ส่วนกลาง</p> <p>- ขายสัญญาห้องloft loft 11,000บาท อยู่ครบถึงปีได้ค่าสัญญาคืน ค่าห้องเดือน ละ 5,500 บาท ช่วงโควิดค่าห้องลดเหลือ3,500 ค่าไฟยูนิตละ6บาท ค่าน้ำยูนิ ตละ35บาท เข้าอยู่ได้เดือนพ.ค.เลย #Dek64 #หอนอกมธ #หอมธ #ทีมมธ</p> <p>- ขายสัญญาหอCity Parkฝั่งหญิง ชั้น9 ระเบียบหันเข้าอินเตอร์ อยู่ตรงช่องลม ห้องไม่ร้อน ห้อง7000/เดือน (ช่วงโควิดเหลือ5500) ค่าสัญญา16000 (ได้คืน ตอนออก) ค่าเปลี่ยนชื่อสัญญา 1000 เข้าอยู่ได้เลย !ชื่อภายในเดือนนี้เดือนเมษา อยู่ฟรี! สนใจdm/line #หอมธ #หอนอกมธ #สัญญาหอมธ #cityparkมธ # ทีมมธ</p> |

ตารางที่ 4.27

คำที่ปรากฏร่วมกันบ่อยที่สุดในข้อความกลุ่มที่ 4 (ปรากฏร่วมกัน 3 คำ)(ต่อ)

| hashtag | |
|--|--|
| ['ขาย', '#หอนอกมธ', 'สัญญา'] | - ขายสัญญาห้องloft loft 11,000บาท อยู่ครบถึงปีได้ค่าสัญญาคืน ค่าห้องเดือนละ 5,500 บาท ช่วงโควิดค่าห้องลดเหลือ3,500 ค่าไฟยูนิตละ6บาท ค่าน้ำยูนิตละ35บาท เข้าอยู่ได้เดือนพ.ค.เลย #Dek64 #หอนอกมธ #หอมธ #ทิมมธ |
| ['ค่า', '#หอนอกมธ', 'สัญญา'] | - ขายสัญญาหอเพลินอพาร์ตเมนต์ ในซอยเจปาร์ค มีพิตเนสฟรี สระว่ายน้ำ มีสนามแบดมินตันราคาสัญญา 11,000 บาท ห้องพักเดือนละ 5,500 มีรถตู้รับส่งเข้าชม มีตู้เย็น ทวีวี พร้อมออกค่าเปลี่ยนชื่อสัญญาให้ #หอนอกมธ #หอมธ #มธ #ขายสัญญาหอ สนใจทักdm มาได้เลยค่ะ พร้อมเข้าอยู่ได้เดือนพฤษา - ขายสัญญาหอCity Parkฝั่งหญิง ชั้น9 ระเบียงหันเข้าอินเตอร์ อยู่ตรงช่องลม ห้องไม่ร้อน ห้อง7000/เดือน (ช่วงโควิดเหลือ5500) ค่าสัญญา16000 (ได้คินตอนออก) ค่าเปลี่ยนชื่อสัญญา 1000 เข้าอยู่ได้เลย !ซื้อภายในเดือนนี้เดือนเมษาอยู่ฟรี! สนใจdm/line #หอมธ #หอนอกมธ #สัญญาหอมธ #cityparkมธ |
| กลุ่มของชื่อเฉพาะ | |
| ['#หอนอกมธ', '#มธรังสิต', 'ชื่อพื้นที่'] | - ขายสัญญาหอ The Charn ค่ะ สนใจทักเดมหรือแอดไลน์นะคะ #หอนอกมธ #หอโนมธ #หอนอกมธรังสิต #ทิมมธ #มธรังสิต #Dek64 - ขายสัญญาหอ twintown ฝั่งรวม - ค่าสัญญา 15,300 บาท (ได้คินตอนออก) - ค่าเช่าเดือนละ 5,300 บาท - มีตู้เย็นขาย 2,500 บาท (ไม่บังคับ) - เข้าอยู่ได้เลย ตั้งแต่ พ.ค.64 - Line: campyfyj - โทร: 0837325501 #หอนอกมธ #หอนอกมธรังสิต #หาเมทมธ #มธรังสิต #ธรรมศาสตร์ #ขายสัญญาหอมธ #หอพักมธ - หาดcon เฟส 1 ดึก c,d / เฟส 2 ราคาไม่เกิน 8000 ค่า #หาหอมธ #หอนอกมธ #หอนอกมธรังสิต #หาเมทมธ #มธรังสิต #ธรรมศาสตร์ |
| ['ค่า', 'ราคา', 'ชื่อพื้นที่'] | - ดิคอนโด แคมปัส รีสอร์ท รังสิต ธรรมศาสตร์ รังสิต ปทุมธานี คลองหลวง ขนาดห้อง 30.00 ตารางเมตร 1ห้องนอน 1ห้องน้ำ ชั้น8 เฟส 1. ดึกB ค่าส่วน 3 เดือน 4040 บาท ราคา 2.1 ล้าน ทักเดม - ขายเคส 11 pro max ค่า มีทั้ง2สีคะ กดไปแล้ว แต่จะเปลี่ยนทศ ราคา 1150 ค่าส่ง50/70(+10) นักรับสยาม,เซนทรัลพระราม2,มธ.รังสิตได้คะ จ่ายเราชนะ/เรารักกัน(ม.33) ได้ค่า #REPRESENTXC3 #XC3 #representxmarktuan #ตลาดกัซ #ตลาดนัดอากาเซ #ตลาดนัดอากาเซ #ตลาดนัดอากาเซgot7 - มึงงงงงงงมันแซ่บบบบบบมากเว่อ โปรหอยแครงลวกพร้อมน้ำจิ้มแซ่บนัว ราคาพิเศษไซส์ M ได้หอย 20-25 ตัว เลือกความสุกได้ 3 ระดับ ราคา 100 รวมส่ง!!! ในเขตกรุงเทพ - นน - รังสิต (มธ,มกรุง) เหลือไม่กี่ออเดอร์แล้วค่า ต้องจัดแล้ว #อรอยไปแดก #อรอยบอกต่อ #หอยแครง #รีวิวของกิน |

ตารางที่ 4.28

คำที่ปรากฏร่วมกันบ่อยที่สุดในข้อความกลุ่มที่ 4 (ปรากฏร่วมกัน 4 คำ)

| คำ, hashtag และ กลุ่มของ ชื่อเฉพาะที่มีการปรากฏ ร่วมกัน | ตัวอย่างข้อความ |
|---|--|
| คำ | |
| ['ขาย', 'ค่า', 'สนใจ', 'สัญญา'] | <p>- ขายสัญญาหอCity Parkฝั่งหญิง ชั้น9 ระเบียงหันเข้าอินเตอร์ อยู่ตรงช่องลม ห่องไม่ร้อน ห่อง7000/เดือน (ช่วงโควิดเหลือ5500) ค่าสัญญา16000 (ได้คืน ตอนออก) ค่าเปลี่ยนชื่อสัญญา 1000 เข้าอยู่ได้เลย !ซื้อภายในเดือนนี้เดือนเมษา อยู่ฟรี! สนใจdm/line #หอมธ #หอนอกมธ #สัญญาหอมธ #cityparkมธ #ทิมมธ</p> |
| ['ขาย', 'ค่า', 'สัญญา', 'ห้อง'] | <p>- ขายสัญญา Twintown ฝั่งหญิงคะ ค่าเปลี่ยนสัญญาออกให้คะ ห้องอยู่ชั้น 7 ไกลลิฟท์ ห้องหันเข้ามอ เข้าอยู่ได้ประมาณพฤษภาเป็นต้นไป สนใจ/สอบถาม DMมาได้เลย #หอนอกมธ #ขายสัญญาหอมธ #หอมธ #ทิมมธ</p> <p>- ขายสัญญาหอ Loft Loft ตึก B ชั้น 7 ค่าห้อง 5,500 (โควิดลดเหลือ 3,500) ค่าสัญญา 10,000 สามารถเข้าอยู่ได้เลย! ทักมาสอบถามได้คับ สนใจทัก DM หรือ Line: Kamonthipwong #ขายสัญญาหอมธ #หอนอกมธ #ทิมมธ #สัญญาหอมธ #หอมธ #ซื้อสัญญาหอมธ #หอพักมธ #มธ #หาหอมธ #หาสัญญาหอมธ #ธรรมศาสตร์</p> |
| hashtag | |
| ['ขาย', 'ค่า', '#หอนอกมธ', 'สัญญา'] | <p>- ขายสัญญาห้องloft loft 11,000บาท อยู่ครบถึงปีได้ค่าสัญญาคืน ค่าห้องเดือนละ 5,500 บาท ช่วงโควิดค่าห้องลดเหลือ3,500 ค่าไฟยูนิตละ6บาท ค่าน้ำยูนิตละ35บาท เข้าอยู่ได้เดือนพ.ค.เลย #Dek64 #หอนอกมธ #หอมธ #ทิมมธ</p> <p>- ขายสัญญาหอเพลินอพาร์ทเมนต์ ในซอยเจปาร์ค มีพิตเนสฟรี สระว่ายน้ำ มีสนามแบดมินตันราคาสัญญา 11,000 บาท ห้องพักเดือนละ 5,500 มีรถตู้รับส่งเข้ามอ มีตู้เย็น ทีวี พร้อมออกค่าเปลี่ยนชื่อสัญญาให้ #หอนอกมธ #หอมธ #มธ #ขายสัญญาหอ สนใจทักdm มาได้เลยคะ พร้อมเข้าอยู่ได้เดือนพฤษา ไม่ค่าส่วนกลาง</p> <p>- ขายสัญญาหอCity Parkฝั่งหญิง ชั้น9 ระเบียงหันเข้าอินเตอร์ อยู่ตรงช่องลม ห่องไม่ร้อน ห่อง7000/เดือน (ช่วงโควิดเหลือ5500) ค่าสัญญา16000 (ได้คืน ตอนออก) ค่าเปลี่ยนชื่อสัญญา 1000 เข้าอยู่ได้เลย !ซื้อภายในเดือนนี้เดือนเมษา อยู่ฟรี! สนใจdm/line #หอมธ #หอนอกมธ #สัญญาหอมธ #cityparkมธ #ทิมมธ</p> |

ตารางที่ 4.28

คำที่ปรากฏร่วมกันบ่อยที่สุดในข้อความกลุ่มที่ 4 (ปรากฏร่วมกัน 4 คำ)(ต่อ)

| Hashtag(ต่อ) | |
|---|---|
| [เมท, '#หอนอกมธ', 'หา', '# หาเมทมธ'] | <p>- หาเมทผู้หญิง หารห้องหอ 18-22 หอสร้างใหม่ฝั่งตรงข้ามมธค้ำบ ห้อง studio plus 8,500 หารตกคนละ 4,250 บาทค่า เริ่มเข้าอยู่เดือนสิงหาคม เปิดเทอม1 หรือใครสะดวกหอรห้องตั้งแต่ช่วงsummerก็ได้ ทักมาคุยกันก่อนได้เลยนะคะ เราอยู่ปี2ค้ำบ #หอนอกมธ #หอพักมธ #Dek64 #ทีมมธ #หาเมทมธ #มธรังสิต</p> <p>- หาเมทหออินเตอร์ปาร์ค ฝั่งหญิงค่า ตอนนี้มีห้องแล้ว ขอคนที่รับได้กับการตื่นเช้าและนอนเช้าในบางช่วง ปล. ทักมาคุยกรูกริบกันก่อนได้นะ ทาง DM หรือไลน์ madelineemito #หาเมทมธ #หอนอกมธ #ทีมมธ #หอพักมธ</p> <p>- หาเมทค่ะ หอแฉนเดือนพ.ค.นี้เลยค่า ขอเป็นผู้หญิงนะคะ ค่าสัญญาหอรกันคนละ 4600 ค่าหออีกคนละ 2250 รวม 6850 เข้าอยู่ได้เลยค่า ผู้ชายขึ้นได้ หอถูกหอใกล้มาก ทักมาคุยสอบถามก่อนได้น้า line : maiathicha #หอนอกมธ</p> |
| กลุ่มของชื่อเฉพาะ | |
| [ชาย', 'ค่า', 'สัญญา', 'ชื่อ พื้นที่'] | <p>- ชายสัญญาหอพัก SUNTA ค่า เข้าอยู่ได้ต้นเดือน พ.ค. หรือคุยกันได้เลยคะ สัญญาหมด 31/07/64 สนใจดูภาพเพิ่มเติมได้เลยคะ #Dek64 #dek65 #TCAS64 #ทีมมธ #มธ #ธรรมศาสตร์ #หอนอกมธ #หอนอกมธรังสิต</p> <p>- ชายสัญญาหอ twintown ฝั่งรวม - ค่าสัญญา 15,300 บาท (ได้คืนตอนออก) - ค่าเช่าเดือนละ 5,300 บาท - มีตู้เย็นชาย 2,500 บาท (ไม่บังคับ) - เข้าอยู่ได้เลย ตั้งแต่ พ.ค.64 - Line: campyfyj - โทร: 0837325501 #หอนอกมธ #หอนอกมธรังสิต #หาเมทมธ #มธรังสิต #ธรรมศาสตร์ #ชายสัญญาหอพัก #หอพักมธ</p> <p>- ชายสัญญาหอZoom place ค่า ห้อง4,500นะคะ #ทีมมธ #หอนอกมธรังสิต #Dek64 (เช่าทิ้งไว้ ไม่ได้อยู่ห้องเลยจับ แล้วคิดว่าคงไม่ได้กลับไปยาวๆเลย)</p> |
| [#หอพักมธ', '#หอนอกมธ', '#มธรังสิต', 'ชื่อพื้นที่'] | <p>- ชายสัญญาหอ twintown ฝั่งรวม - ค่าสัญญา 15,300 บาท (ได้คืนตอนออก) - ค่าเช่าเดือนละ 5,300 บาท - มีตู้เย็นชาย 2,500 บาท (ไม่บังคับ) - เข้าอยู่ได้เลย ตั้งแต่ พ.ค.64 - Line: campyfyj - โทร: 0837325501 #หอนอกมธ #หอนอกมธรังสิต #หาเมทมธ #มธรังสิต #ธรรมศาสตร์ #ชายสัญญาหอพัก #หอพักมธ</p> <p>- หาดcon เฟส 1 ดีก c,d / เฟส 2 ราคาไม่เกิน 8000 ค่า #หอนอกมธ #หอนอกมธรังสิต #หาเมทมธ #มธรังสิต #ธรรมศาสตร์ #ชายสัญญาหอพัก</p> <p>- ตอนนี้มีเพจ Facebook : RESTTWIN แล้วน้า ฝากทุกคนเข้าไปกดติดตาม กดไลค์กดแชร์ให้หน่อยนะคะ ทางเพจจะคอยอัปเดตข่าวสารรอบมธ รีวิวหอพัก ตอนนี้ทางเพจจะมีกิจกรรมให้เล่นชิงเงินรางวัลสนุกๆด้วยแหละไปไลค์เพจกันเลยยย #หอนอกมธ #หอนอกมธรังสิต #หาเมทมธ #มธรังสิต #ธรรมศาสตร์ #หอพักมธ #dek64</p> |

จากตารางที่ 4.26, 4.27 และ 4.28 พบคำที่ปรากฏร่วมกันได้แก่ ['ค่า', 'สนใจ'], ['ขาย', 'ค่า'], ['ค่า', '#หอนอกมธ'], ['หอ', '#หอนอกมธ'], ['ค่า', 'ชื่อพื้นที่'], ['นัด', 'ชื่อพื้นที่'], ['ค่า', 'สนใจ', 'Direct message'], ['ขาย', 'ค่า', 'สัญญา'], ['ขาย', '#หอนอกมธ'], ['ค่า', '#หอนอกมธ', 'สัญญา'], ['#หอนอกมธ', '#มธรังสิต', 'ชื่อพื้นที่'], ['ค่า', 'ราคา', 'ชื่อพื้นที่'], ['ขาย', 'ค่า', 'สนใจ', 'สัญญา'], ['ขาย', 'ค่า', 'สัญญา', 'ห้อง'], ['ขาย', 'ค่า', '#หอนอกมธ', 'สัญญา'], ['เมท', '#หอนอกมธ', 'หา', '#หาเมทมธ'], ['ขาย', 'ค่า', 'สัญญา', 'ชื่อพื้นที่'], ['#หอพักมธ', '#หอนอกมธ', '#มธรังสิต', 'ชื่อพื้นที่'] จากคำที่ปรากฏร่วมกัน และตัวอย่างของข้อความ สามารถ สรุปหัวข้อของข้อความ กลุ่มที่ 4 พบว่า ข้อความในกลุ่มนี้เกี่ยวข้องกับ หอพัก และการซื้อ/ขายสินค้า



บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

ในบทนี้เป็นการสรุปผลการจัดกลุ่มข้อความทวิตเตอร์ที่กล่าวถึงมหาวิทยาลัยธรรมศาสตร์โดยใช้อัลกอริทึม k-means และข้อเสนอสำหรับงานวิจัยในครั้งต่อไป โดยมีเนื้อหา ดังนี้

5.1 สรุปผล

การจัดกลุ่มข้อความที่กล่าวถึงมหาวิทยาลัยธรรมศาสตร์โดยใช้อัลกอริทึม k-means พบว่า ความคล้ายคลึงที่เหมาะสมสำหรับการจัดกลุ่มคือ Cosine Similarity สำหรับการจัดกลุ่มข้อความที่ได้รับการเตรียมข้อมูลด้วยการแปลงชื่อเฉพาะเป็นกลุ่มของชื่อเฉพาะ และการจัดกลุ่มข้อความที่ไม่ได้รับการเตรียมข้อมูลด้วยการแปลงชื่อเฉพาะเป็นกลุ่มของชื่อเฉพาะ มีจำนวนกลุ่มที่เหมาะสม 4 กลุ่มเท่ากัน เป็นการจัดกลุ่มที่ไปในทิศทางเดียวกัน และเมื่อแปลงชื่อเฉพาะเป็นกลุ่มของชื่อเฉพาะ พบว่าข้อมูลในแต่ละกลุ่มมีความชัดเจนมากยิ่งขึ้น

ความถี่สูงสุดของคำในแต่ละกลุ่มข้อความนั้นยังไม่สามารถนำมาสรุปหัวข้อของกลุ่มนั้นๆได้ เนื่องจากคำที่มีความถี่สูงสุดในแต่ละกลุ่มนั้นมีบางคำที่ไม่สื่อถึงข้อความในกลุ่มนั้นๆ การสร้างชุดของคำ, hashtag และ กลุ่มของชื่อเฉพาะที่มีการปรากฏร่วมกัน เป็นวิธีที่มาช่วยในการกำหนดหัวข้อของกลุ่มนั้นๆได้

การจัดกลุ่มข้อความที่กล่าวถึงมหาวิทยาลัยธรรมศาสตร์ที่ได้รับการเตรียมข้อมูลด้วยการแปลงชื่อเฉพาะเป็นกลุ่มของชื่อเฉพาะ สามารถจัดกลุ่มได้ 4 กลุ่ม โดยกลุ่มที่ 1 เกี่ยวข้องกับด้านกรเรียนการศึกษา กลุ่มที่ 2 เกี่ยวข้องกับการสอบเข้าหรือความสนใจที่จะเข้าศึกษาต่อในมหาวิทยาลัย กลุ่มที่ 3 เกี่ยวข้องกับการซื้อ/ขาย/เช่า หอพักรวมไปถึง สินค้าต่างๆ และกลุ่มที่ 4 เกี่ยวข้องกับโรคโคโรนาไวรัส 2019 และ การเมืองภายในประเทศไทย จะเห็นได้ว่าการจัดกลุ่มที่ 4 ไม่ชัดเจนเท่ากับการจัดกลุ่มอื่นๆ เนื่องจาก กลุ่มที่ 4 เป็นข้อความที่เกี่ยวข้องกับโรคโคโรนาไวรัส 2019 และ การเมืองภายในประเทศไทย ซึ่งเป็นสองหัวข้อที่ไม่ได้ไปในทิศทางเดียวกัน

5.2 ข้อเสนอแนะสำหรับงานวิจัยครั้งต่อไป

5.2.1 การสร้างพจนานุกรมชื่อเฉพาะในงานวิจัยนี้ยังไม่อยู่ในระบบอัตโนมัติหากใช้วิธีที่เป็นอัตโนมัติจะช่วยลดเวลา และขั้นตอนได้

5.2.2 การสกัดคุณลักษณะของข้อความปัจจุบันมีหลากหลายวิธีควรวิเคราะห์ปัจจัยที่ส่งผลต่อการจัดกลุ่มข้อความของคุณลักษณะต่างๆ

5.2.3 ขอบเขตของข้อมูลอยู่เพียงมหาวิทยาลัยธรรมศาสตร์หากเพิ่มมหาวิทยาลัยอื่นๆ จะสามารถนำมาเปรียบเทียบและวิเคราะห์ได้แม่นยำขึ้น



รายการอ้างอิง

หนังสือและบทความในหนังสือ

Charu C. Aggarwal, ChengXiang Zhai (2012). *Mining Text Data*. Springer New York, Published: 03 February 2012

บทความวารสารและงานประชุมวิชาการ

Charoensuk, J., & Sornil, O. (2018). A Hierarchical Emotion Classification Technique for Thai Reviews. *Journal of Information and Communication Technology ITB* 12, 280-296.

Chumwatana, T., & Wongkolkitsilp, K. (2019). *Using Classification Technique for Customer Relationship Management based on Thai Social Media Data. Proceeding of the International Conference on Computer and Automation Engineering*

Haruechaiyasak, C., Kongyoung, S., & Damrongrat, C. (2008). LearnLexTo: A Machine-Learning Based Word Segmentation for Indexing Thai Texts. *Proceeding of the 2nd ACM workshop on Improving Non English Web Searching, iNEWS*.

Jotikabukkana, P., Sornlertlamvanich, V., & Manabu, O. (2016). Social Media Text Classification by Enhancing Well-Formed Text Trained Model. *Institut Teknologi Bandung Journal*, 10, 177-196

Larasati, A., Maren, R., & Wulandari, R. (2021). Utilizing Elbow Method for Text Clustering Optimization in Analyzing Social Media Marketing Content of Indonesian e-Commerce. *Jurnal Teknik Industri*, 23, 111-120.

M. Ikonomakis, kotsiantis, s., & tampakas, V. (2005). Text Classification Using Machine Learning Techniques. *WSEAS TRANSACTIONS on COMPUTERS*, 4(8), 966-974.

Masdisornchote, M. (2016). A sentiment analysis framework in implicit opinions for Thai language. *Annual. Proceeding of the IEEE Industrial Electronics Society (IES)*.

Nainggolan, R., Perangin-angin, R., Simarmata, E., & Tarigan, F. A. (2019). Improved the Performance of the K-Means Cluster Using the Sum of Squared Error (SSE) optimized by using the Elbow Method. *Journal of Physics: Conference Series*.

- Phraewattanakun, K., & Pramote. (2013). Opinion Mining From Online Social Networks. *Modern Management Journal*, 11, 11-20.
- Salih, N. (2022). Semantic-Based K-Means Clustering for IMDB Top 100 Movies. *Journal of Applied Science and Technology Trends*, 3(2), 58-61.
- Social Networks: An Introduction. (2009). *Development and Learning in Organizations: An International Journal*, 23(4).
- Songram, p. (2016). Opinion Classification of Politics on Social Network using Associative Classification. *Science and Technology RMUTT Journal*, 6, 83-93.
- Thanasopon, B., Buranapanitkij, J., & Netisopakul, P. (2019). mining social media crowd trends from thai text posts and comments. *Malaysian Journal of Computer Science. Information Technology and Electrical Engineering Special Issue*, 96-108.
- Thetmueang, R., & Chirawichitchai, N. (2017). Thai Sentiment Analysis of Product Review Online Using Support Vector Machine. *Engineering Journal of Siam University*, 18(1).
- Vateekul, P., & Koomsubha, T. (2016). *A Study of Sentiment Analysis Using Deep Learning Techniques on Thai Twitter Data. Proceeding of the International Joint Conference on Computer Science and Software Engineering (JCSSE)*.
- ธนาธร ทยานานทอง และคณะ (2019). ระบบเฝ้าติดตามความคิดเห็นบนสื่อสังคมออนไลน์สำหรับมหาวิทยาลัยไทย. *TNI Journal of Engineering and Technology*, 7.
- พยุ่ง มีสังข์และคณะ (2019). Application of Text Mining for Data Clustering: A Case Study for Cancer. *Science and Technology RMUTT Journal*, 9, 71-79.
- Wichitboonyarak, P. (2011). Social Media: Future Media. *Executive Journal*, 31(4), 99-103.
- Masdisornchote, M. (2014). Constructing Thai Opinion Mining. *Proceedings of the 8th Workshop on Asian Language Resources*, 7, 64-71.

สื่ออิเล็กทรอนิกส์

- Ahlgren, M. (2020). 50 + สถิติ TWITTER & ข้อเท็จจริงสำหรับ 2020. Retrieved from <https://www.websitehostingrating.com/th/twitter-statistics/>. สืบค้นเมื่อวันที่ 18 มิถุนายน 2565,
- DataReportal. (2020). *DIGITAL 2020: THAILAND*. Retrieved from <https://datareportal.com/reports/digital-2020-thailand> สืบค้นเมื่อวันที่ 18 มิถุนายน 2565,
- Group, A. R. (2020). Social Media Analytics คืออะไรและทำไมจึงมีความสำคัญ. Retrieved from <https://www.ar.co.th/kp/th/575> สืบค้นเมื่อวันที่ 17 มิถุนายน 2565,
- Positioning. (2020a). คนไทยติดหนึบ ‘โซเชียล มีเดีย’ ยอดเฟซบุ๊กทะลุ 41 ล้านราย ติดอันดับ 8 ของโลก. Retrieved from <https://positioningmag.com/1092090> สืบค้นเมื่อวันที่ 21 มิถุนายน 2565,
- Positioning. (2020b). คนไทยเทใจใช้ “ทวิตเตอร์” โด่งดัง สานทางเฟซบุ๊ก โนสนโนแคร์ Retrieved from <https://positioningmag.com/1159478> สืบค้นเมื่อวันที่ 21 มิถุนายน 2565,
- จำปี หอม, ส. (2018). สรุปความรู้เบื้องต้นเกี่ยวกับสังคมออนไลน์. Retrieved from <https://sites.google.com/site/suwinaycapihxm/bi-ngan-srup-khwam-ru-beuxng-tn-keiyw-kab-sangkhn-xxnlin> สืบค้นเมื่อวันที่ 11 สิงหาคม 2565,

วิทยานิพนธ์

- Choonui, W. (2012). *Feature Reduction using FCA for Web Page Classification*. (Master of Science). Prince of Songkla University,
- Sirihattasak, S. (2019). *Annotation and Classification of Toxicity for Thai Twitter*. (Master degree). Tokyo Metropolitan University, Miyazaki, Japan.
- Tapang, K. (2016). *Sarcasm classification in twitter using probability of tweets*. (Master of Science). Chulalongkorn University,



ภาคผนวก ก
พจนานุกรมชื่อเฉพาะ

| พจนานุกรมชื่อเฉพาะ | |
|--|--|
| กลุ่มของชื่อเฉพาะ | ชื่อเฉพาะ |
| ชื่อข้อสอบ | Pat2 9 วิชาสามัญ 9 วิชา TCAS64 TCAS65 TCAS 9สามัญ TU GET PAT1 9วิชา TU100 Onet ความถนัดแพทย์ getpat ONET Pat pat2 |
| ชื่อคณะ | สถาปัต บริหารธุรกิจ คณะสาธารณสุขศาสตร์ JC คณะศิลปศาสตร์ คณะแพทยศาสตร์ สถาปัต วิทยาการเรีนรู้อู วารสาร เศรษฐศาสตร์ เทคนิคการแพทย์ คณะทางวิทย์ เสดสาด รัฐศาสตร์ BMCI วิศวกรรมไฟฟ้า BEC BAS ศิลปศาสตร์ อี้จรรณกรรม คณะนิติศาสตร์ อักษรฯ วิศวกรรม บัญชี นิติ นิติศาสตร์ รัฐศาสตร์มธ ศิลปศาสตร์มธ คณะสถาปัตยกรรม คณะวิศวกรรมศาสตร์ |
| ชื่อจังหวัด | ปทุมธานี ออยุธยา สุพรรณบุรี ราชบุรี |
| ชื่อนักแสดง | ฮุย อุซอก ชินวอน ยอวอน ยูโตะ ดงพโย ลีอัน tmd tilly three man down |
| ชื่อบุคคล/กลุ่มบุคคลที่ เกี่ยวข้องกับการเมือง | คสช. ไชยันต์ ประติมานุสรณ์ 6 ตุลา กลุ่มคนตุลาเพื่อประชาธิปไตย OctDem สลิ้ม พรรคธรรมเพื่อโดม แนวร่วมธรรมศาสตร์ ลานประติมากรรม ชาญวิทย์ เกษตรศิริ พันัส ทศนียานนท์ #โตโต้ปิยะรัฐ แนวร่วมธรรมศาสตร์และการชุมนุม โตโต้ ปิยะรัฐ แนวร่วมธรรมศาสตร์ฯ ชูเวช ปิยะบุตร ไอติม ประยุทธ์ เกศินี วิฑูรชาติ ขวัญชนก โชคสกุล |
| ชื่อพื้นที่ | รังสิต เมืองเอก เมืองทอง นวนคร ปากเกร็ด สายไหม ดอนเมือง เชียงราก2 พระราม 7 |
| ชื่อมหาลัย | ม.กรุงเทพ ม.รังสิต จุฬา มช มอชอ |
| ชื่อโรคติดต่อ | โคโรนา โควิด Covid19 Covid-19 |
| ชื่อโรงพยาบาล | รพ.ธรรมศาสตร์ โรงพยาบาลธรรมศาสตร์ รพ มธ คลินิกพิเศษธรรมศาสตร์ โรงพยาบาลสนาม รพ.ภูมิพล |

| พจนานุกรมชื่อเฉพาะ | |
|--------------------------|--|
| กลุ่มของชื่อเฉพาะ | ชื่อเฉพาะ |
| ชื่อสถานที่ในมหาวิทยาลัย | ท่าพระจันทร์ ห้องกรีก อาคารบรรยายรวม หอป่วย |
| ชื่อสถานที่สำคัญ | เวสเกต ทะเลทรายซาฮารา ฟิวเจอร์ ศ.สิริกิติ์ ศาลแขวงดุสิต ฟิวเจอร์พาร์ค |
| ชื่อสิ่งของในมหาวิทยาลัย | NGV รถเอนจี้ |
| ชื่อหอพัก | Zommlace เดอะชาญ หอวินทาวน์ ดีคอนโด The kith หอมธ city park หอโนมธ หอนอกมธ หอโนมธรังสิต หอนอกมธรังสิต |
| อื่นๆ | พ่อปรีดี เมษา Youtube April fools day กองทัพบก MK Media ไทย ญี่ปุ่น มี.ค. เม.ย. ไวอากร้า อังกฤษ ตุลา โรงพยาบาล สจล. คาเฟ่โลโลโดม เรือนจำพิเศษกรุงเทพฯ สน.ประชาชื่น สังคม คณิต ฟิสิกส์ ชิวะ เคมี ธรรมศาสตร์และการขายของ เรือนจำพิเศษกรุงเทพ พฤษภาคม LT1 งานบอลประเพณี Genchem gpax gpa สเปน |

ภาคผนวก ข

พจนานุกรมคำพ้องความหมาย

| กลุ่มของคำที่มีความหมายคล้ายกัน | |
|---------------------------------|---|
| คำตัวแทนคำพ้อง ความหมาย | รายการคำพ้องความหมาย |
| Covid-19 | Covid-19, โควิด19, โควิด-19, Covid, Covid19, โคจิด, COVID, COVID19, COVID-19, โคโรนาไวรัส, โคโรนา, Coronaviruses, โควิด |
| คณะสถาปัตยกรรม | คณะสถาปัตยกรรม, สถาปัต, คณะสถาปัตยกรรม, สถาปัตยกรรม, สถาปัตยกรรม, สถาปัต |
| น้อยใจ | น้อยด์, น้อยใจ, น้อยจ้ย |
| หมดกำลังใจ | ดาวน, เศร้า, หมดกำลังใจ |
| เมษายน | เมษายน, เม.ย., เมย., เมษายน, April |
| คณะรักษาความสงบ แห่งชาติ | คณะรักษาความสงบแห่งชาติ, คสช., คสช |
| บริหารธุรกิจ | บริหารธุรกิจ, คณะบริหาร |
| ติวเตอร์ | ติวเตอร์, สอนพิเศษ, Tutor |
| Passion | Passion, แพชชั่น, แพทชั่น |
| รถNGV | รถNGV, รถเอนจี้, NEV, NGV, ngv, Ngv, เอนจี้ |
| คณะสาธารณสุข ศาสตร์ | คณะสาธารณสุขศาสตร์, สาธา, สาธารณสุขศาสตร์, สาธารณสุข |
| กองทัพบก | กองทัพบก, ทหาร, ทบ. |
| หน้ากากอนามัย | หน้ากากอนามัย, แมส, หน้ากาก, mask Mask |
| zommplice | zommplice, หอชุม, หอzoom |
| ประเทศไทย | ไทย, TH, ประเทศไทย |
| ประเทศญี่ปุ่น | ญี่ปุ่น, JP, ประเทศญี่ปุ่น |
| นักเรียน | นักเรียน, นร, นร. |
| นักศึกษา | นักศึกษา, นศ, นศ. |
| ป.โท | ป.โท, บัณฑิตศึกษา |

| กลุ่มของคำที่มีความหมายคล้ายกัน | |
|---------------------------------|---|
| คำตัวแทนคำพ้อง ความหมาย | รายการคำพ้องความหมาย |
| Line | Line, ไลน์, line |
| จังหวัดปทุมธานี | ปทุมธานี, ปทุม, จังหวัดปทุมธานี, จังหวัดปทุม |
| 9 วิชาสามัญ | 9 วิชาสามัญ, 9 วิชา, ข้อสอบ9วิชา |
| Direct message | Direct message, DM, dm, Dm, ดีเอ็ม, เดมมา |
| อธิการบดี | อธิการบดี , อธิการ |
| คณะ วิศวกรรมศาสตร์ | คณะวิศวกรรมศาสตร์, วิศวกรรม, วิศวกรรมศาสตร์, วิศวกรรม |
| ประชาธิปไตย | ประชาธิปไตย, ปะชาธิปไตย, ประชาธิปไตย |
| PAT | Pat, pat, PAT, แพท |
| #Dek64 | #Dek64, #dek64, #DEK64 |
| โรงพยาบาลสนาม | โรงพยาบาลสนาม, ร.พ.สนาม โรงบาลสนาม, รพ.สนาม |
| โรงพยาบาล | โรงพยาบาล, รพ., โรงบาล |



ประวัติผู้เขียน

| | |
|-----------------|--|
| ชื่อ | มนต์ชัย ป่านขาว |
| วันเดือนปีเกิด | 30 เมษายน พ.ศ. 2539 |
| วุฒิการศึกษา | ปีการศึกษา 2557 : วิทยาศาสตร์บัณฑิต มหาวิทยาลัยธรรมศาสตร์ |
| ผลงานทางวิชาการ | |

มนต์ชัย ป่านขาว, และ ธนาธร ทะนานทอง. (2565) การแบ่งกลุ่มข้อความภาษาไทยบนทวิตเตอร์ที่กล่าวถึงมหาวิทยาลัยโดยใช้เทคนิค K-Means. การประชุมวิชาการระดับประเทศทางด้านเทคโนโลยีสารสนเทศครั้งที่ 14 (NCIT2022), 58-63

