



การทำแบบจำลองหัวข้อด้วยเทคนิค LDA, Top2vec, BERTopic  
กับข้อความรีวิวการทำงานภายในบริษัทต่างๆ บนชุดข้อมูล Glassdoor

โดย  
นายสินธุ์ ใจติกเสถียร  
นายนิติภูมิ นุชน้อย

โครงการพิเศษนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร  
วิทยาศาสตรบัณฑิต  
สาขาวิชาวิทยาการคอมพิวเตอร์  
คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์  
ปีการศึกษา 2566  
ลิขสิทธิ์ของมหาวิทยาลัยธรรมศาสตร์

การทำแบบจำลองหัวข้อด้วยเทคนิค LDA, Top2vec, BERTopic  
กับข้อความรีวิวการทำงานภายในบริษัทต่างๆบนชุดข้อมูล Glassdoor

โดย  
นายสินธุ์ โชคิกเสถียร  
นายนิติภูมิ นุชน้อย

โครงการพิเศษนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร  
วิทยาศาสตรบัณฑิต  
สาขาวิชาวิทยาการคอมพิวเตอร์  
คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์  
ปีการศึกษา 2566  
ลิขสิทธิ์ของมหาวิทยาลัยธรรมศาสตร์

TOPIC MODELING WITH LDA, TOP2VEC, BERTOPIC ON JOB  
REVIEWS IN GLASSDOOR DATASET

BY

MISTER SINCHOO CHOTIKASATHIAN  
MISTER NITIPOOM NUCHNOI

A FINAL-YEAR PROJECT REPORT SUBMITTED IN PARTIAL  
FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF  
BACHELOR OF SCIENCE  
COMPUTER SCIENCE  
FACULTY OF SCIENCE AND TECHNOLOGY  
THAMMASAT UNIVERSITY  
ACADEMIC YEAR 2023  
COPYRIGHT OF THAMMASAT UNIVERSITY

มหาวิทยาลัยธรรมศาสตร์  
คณะวิทยาศาสตร์และเทคโนโลยี

รายงานโครงการพิเศษ

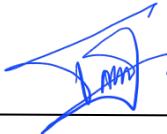
ของ  
นายสินชู ใจติกเสถียร  
นายนิติภูมิ นุชน้อย

เรื่อง

การทำแบบจำลองหัวข้อด้วยเทคนิค LDA, Top2vec, BERTopic กับข้อความรีวิวการทำงานภายใน  
บริษัทต่างๆบนชุดข้อมูล Glassdoor

ได้รับการตรวจสอบและอนุมัติ ให้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร  
หลักสูตรวิทยาศาสตรบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์  
เมื่อ วันที่ 28 ธันวาคม พ.ศ. 2566

อาจารย์ที่ปรึกษา




---

(ผศ.ดร. ฐานะปนา บุญชู)

กรรมการสอบโครงการพิเศษ

---

(ผศ.ดร. เสาวลักษณ์ วรรธนาภา)

กรรมการสอบโครงการพิเศษ

---

(อ.ดร. ศานนาภิ กิจศิรานุวัตร)

กรรมการสอบโครงการพิเศษ

---

(อ.ดร. นุชจรินทร์ อินตัชหล้า)

มหาวิทยาลัยธรรมศาสตร์  
คณะวิทยาศาสตร์และเทคโนโลยี

รายงานโครงการพิเศษ

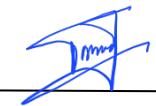
ของ  
นายสินชู ใจติกเสถียร  
นายนิติภูมิ นุชน้อย

เรื่อง

การทำแบบจำลองหัวข้อด้วยเทคนิค LDA, Top2vec, BERTopic กับข้อความรีวิวการทำงานภายใน  
บริษัทต่างๆบนชุดข้อมูล Glassdoor

ได้รับการตรวจสอบและอนุมัติ ให้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร  
หลักสูตรวิทยาศาสตรบัณฑิต สาขาวิชาการคอมพิวเตอร์  
เมื่อ วันที่ 28 ธันวาคม พ.ศ. 2566

อาจารย์ที่ปรึกษา



(ผศ.ดร. ฐานา บุญชู)

กรรมการสอบโครงการพิเศษ

(ผศ.ดร. เสาวลักษณ์ วรรณนาภา)

กรรมการสอบโครงการพิเศษ

(อ.ดร. ศานนา กิจศิรานุวัตร)

กรรมการสอบโครงการพิเศษ

(อ.ดร. นุชจรินทร์ อินตัชหล้า)

หัวข้อโครงการพิเศษ	การสร้างแบบจำลองหัวข้อด้วยเทคนิค LDA, Top2vec, BERTopic กับข้อความรีวิวการทำงานภายในบริษัทต่างๆ บนชุดข้อมูล Glassdoor
ชื่อผู้เขียน	นายสินธุ์ ใจศิริ
ชื่อผู้เขียน	นายนิติภูมิ นุชน้อย
ชื่อปริญญา	วิทยาศาสตรบัณฑิต สาขาวิชาภาษาคอมพิวเตอร์
สาขาวิชา/คณะ/มหาวิทยาลัย	สาขาวิชาภาษาคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์
อาจารย์ที่ปรึกษาโครงการพิเศษ	อ.ดร. ฐานา บุญชู
ปีการศึกษา	2566

### บทคัดย่อ

การจัดทำโครงการพิเศษนี้มีเป้าหมายคือ ทดลองศึกษาการสร้างแบบจำลองหัวข้อกับข้อความรีวิวการทำงานภายในบริษัทต่างๆ บนชุดข้อมูล Glassdoor ผ่านการวิเคราะห์หาคำที่มีนัยสำคัญและการจำแนกหัวข้อในข้อความที่มีความคล้ายคลึงกันโดยอาศัยประโยชน์จากการทำแบบจำลองหัวข้อ (Topic Modeling) เป็นการทำประดิ่นสำคัญที่ช่วยอุปกรณ์ในการทำงานรูปแบบของหัวข้อจากชุดเอกสารที่มีจำนวนข้อมูลมากๆ โดยใช้เทคนิคในการสร้างแบบจำลองหัวข้อ 3 เทคนิคได้แก่ 1. การจัดสรรเดรีเคลฟัง (Latent Dirichlet Allocation) 2. Top2vec และ 3. BERTopic

จากการศึกษาพบว่าเทคนิค LDA ทำประสิทธิภาพในเชิงปริมาณและคุณภาพเป็นอันดับหนึ่งรองลงมาเป็น BERTopic และ Top2vec ตามลำดับโดยผู้วิจัยมีความเห็นว่าเทคนิค LDA อาจจะมีความเหมาะสมกับข้อความรีวิวการทำงานภายในบริษัทต่างๆ ของชุดข้อมูล Glassdoor นี้ ดังนั้นผู้วิจัยมีความเห็นว่าในการศึกษาเกี่ยวกับการสร้างแบบจำลองหัวข้อกับข้อความรีวิวการทำงานภายในบริษัทการใช้เทคนิค LDA จึงควรพิจารณาเป็นอันดับแรกๆ

**คำสำคัญ:** Glassdoor, LDA, Top2vec, BERTopic

Thesis Title	Topic modeling with LDA, Top2vec, BERTopic on job reviews in Glassdoor dataset
Author	Sinchoo Chotikasathian
Author	Nitipoom Nuchnoi
Degree	Bachelor of Science
Major Field/Faculty/University	Computer Science Faculty of Science and Technology Thammasat University
Project Advisor	Academic Title Firstname Surname
Academic Years	2023

## ABSTRACT

The goal of this special project is Experimental study of topic modeling with reviews of work within companies on the Glassdoor dataset through analysis of significant words and classification of similar topics in the text using Topic Modeling. Topic Modeling is the process of finding important hidden points in the form of topics from a set of documents with a large amount of information. It uses three topic modeling techniques: 1) Latent Dirichlet Allocation, 2) Top2vec, and 3) BERTopic.

From the results of this study, it was found that the LDA technique performed first in terms of quantity and quality, followed by BERTopic and Top2vec, respectively. The researcher have an opinion that the LDA technique may be appropriate for text reviews of work within a company in Glassdoor dataset. Therefore, the researcher have an opinion that in order to studies of topic modeling with text reviews of work within a company, LDA techniques should be considered first.

**Keywords:** Glassdoor, LDA, Top2vec, BERTopic

## กิตติกรรมประกาศ

โครงการนี้สำเร็จลุล่วงได้ด้วยความกรุณาจากอาจารย์ดร.ฐานะปนา บุญชู ที่ให้คำแนะนำ และให้คำปรึกษาในการทำโครงการตั้งแต่เริ่มต้นทำโครงการ อีกทั้งยังช่วยเหลือในการแก้ไขข้อบกพร่องต่างๆที่เกิดขึ้นในโครงการสำเร็จลุล่วงไปได้ด้วยดี

ขอขอบคุณเพื่อนร่วมชั้นภายในสาขาวิชาการคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์ ศูนย์รังสิต ที่ให้ความช่วยเหลือในการทำโครงการ

สุดท้ายนี้ขอขอบคุณทุกคนที่ช่วยให้คำแนะนำและให้กำลังใจในการทำโครงการนี้จนสามารถทำสำเร็จลุล่วง

นายสินชู ใจติกเสถียร

นายนิติภูมิ นุชน้อย

สารบัญ	หน้า
บทคัดย่อ	๑
ABSTRACT	๒
กิตติกรรมประกาศ	๓
สารบัญ	๔
สารบัญตาราง	๕
สารบัญภาพ	๖
รายการสัญลักษณ์และคำย่อ	๗
บทที่ 1 บทนำ	๑
1.1 ความเป็นมาและความสำคัญของโครงการ	๑
1.2 วัตถุประสงค์	๒
1.3 ขอบเขตของโครงการ	๒
1.4 ประโยชน์ของโครงการ	๒
บทที่ 2 วรรณกรรมและงานวิจัยที่เกี่ยวข้อง	๓
2.1 แนวคิดทฤษฎีที่เกี่ยวข้อง	๓
2.1.1 การประมวลผลภาษาธรรมชาติ (Natural language processing: NLP)	๓
2.1.2 ค่า Topic Coherence Cv	๖
2.2 งานวิจัยที่เกี่ยวข้อง	๘
2.3 เทคโนโลยีที่เกี่ยวข้อง	๘
2.3.1 ภาษาโปรแกรมไพธอน (Python Programming Language)	๘
2.3.2 เครื่องมือ (Tools)	๘

บทที่ ๓ วิธีการวิจัย	9
3.1 ภาพรวมของโครงงาน	9
3.2 ขั้นตอนการดำเนินงานวิจัย	9
3.2.1 การนำเข้าและคัดแยกข้อมูล (Import and Extracting Data)	9
3.2.2 การเตรียมข้อมูลและทำความสะอาดข้อมูล (Data Preprocessing)	11
3.2.3 ตัวแบบหัวข้อ	14
3.2.4 ประเมินประสิทธิภาพ (Evaluation)	16
3.2.5 การสรุปผลลัพธ์การทดลอง (Conclusion)	16
บทที่ ๔ ผลการดำเนินวิจัย	17
บทที่ ๕ สรุป	139
5.1 สรุปผลการทดลอง	139
5.1.1 สรุปกลุ่มหัวข้อที่คล้ายคลึงในแต่ละเทคนิคแบ่งตาม บริษัทและประเภทข้อมูล	139
5.1.2 สรุปผลการประเมินประสิทธิภาพเชิงคุณภาพของ แบบจำลองหัวข้อในแต่ละเทคนิค	142
รายการอ้างอิง	143

## สารบัญตาราง

หน้า

ตารางที่ 4.1 ผลการวิเคราะห์หัวคำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล headline บริษัท Apple	17
ตารางที่ 4.2 กลุ่มสมมุติฐานหัวข้อความหมายที่คล้ายคลึง headline ใน Apple	19
ตารางที่ 4.3 ผลการวิเคราะห์หัวคำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล รีวิว ประโยชน์ที่ได้ บริษัท apple	20
ตารางที่ 4.4 กลุ่มสมมุติฐานหัวข้อความหมายที่คล้ายคลึง ประโยชน์ที่ได้ ใน Apple	22
ตารางที่ 4.5 ผลการวิเคราะห์หัวคำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล รีวิว ผลเสียที่ได้ บริษัท apple	22
ตารางที่ 4.6 กลุ่มสมมุติฐานหัวข้อความหมายที่คล้ายคลึง ผลเสียที่ได้ ใน Apple	24
ตารางที่ 4.7 ผลการวิเคราะห์หัวคำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล รีวิว headline บริษัท Google	25
ตารางที่ 4.8 กลุ่มสมมุติฐานหัวข้อความหมายที่คล้ายคลึง headline ใน Google	27
ตารางที่ 4.9 ผลการวิเคราะห์หัวคำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล รีวิว ผลประโยชน์ที่ได้ บริษัท Google	27
ตารางที่ 4.10 กลุ่มสมมุติฐานหัวข้อความหมายที่คล้ายคลึง ประโยชน์ที่ได้ ใน Google	29
ตารางที่ 4.11 ผลการวิเคราะห์หัวคำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล รีวิวผลเสียที่ได้ บริษัท Google	30
ตารางที่ 4.12 กลุ่มสมมุติฐานหัวข้อความหมายที่คล้ายคลึง ผลเสียที่ได้ ใน Google	32
ตารางที่ 4.13 ผลการวิเคราะห์หัวคำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล headline บริษัท Microsoft	32
ตารางที่ 4.14 กลุ่มสมมุติฐานหัวข้อความหมายที่คล้ายคลึง headline ใน Microsoft	34
ตารางที่ 4.15 ผลการวิเคราะห์หัวคำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล รีวิวประโยชน์ที่ได้ บริษัท Microsoft	35
ตารางที่ 4.16 กลุ่มสมมุติฐานหัวข้อความหมายที่คล้ายคลึง ประโยชน์ที่ได้ ใน Microsoft	37

ตารางที่ 4.17 ผลการวิเคราะห์หาคำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูลรีวิวผลเสียที่ได้ บริษัท Microsoft	38
ตารางที่ 4.18 กลุ่มสมมุติฐานหัวข้อความหมายที่คล้ายคลึง ผลเสียที่ได้ ใน Microsoft	39
ตารางที่ 4.19 ตารางสมมุติฐานกลุ่มหัวข้อ headline ของ Apple	125
ตารางที่ 4.20 ตารางสมมุติฐานกลุ่มหัวข้อ ผลประโยชน์ที่ได้ ของ Apple	126
ตารางที่ 4.21 ตารางสมมุติฐานกลุ่มหัวข้อ ผลเสียที่ได้ ของ Apple	127
ตารางที่ 4.22 ตารางสมมุติฐานกลุ่มหัวข้อ headline ของ Google	129
ตารางที่ 4.23 ตารางสมมุติฐานกลุ่มหัวข้อ ประโยชน์ที่ได้ ของ Google	130
ตารางที่ 4.24 ตารางสมมุติฐานกลุ่มหัวข้อ ผลเสียที่ได้ ของ Google	132
ตารางที่ 4.25 ตารางสมมุติฐานกลุ่มหัวข้อ headline ของ Microsoft	133
ตารางที่ 4.26 ตารางสมมุติฐานกลุ่มหัวข้อ ผลประโยชน์ที่ได้ของ Microsoft	135
ตารางที่ 4.27 ตารางสมมุติฐานกลุ่มหัวข้อ ผลเสียที่ได้ของ Microsoft	136
ตารางที่ 5.1 กลุ่มหัวเรื่องที่คล้ายคลึงในแต่ละเทคนิคของ Apple	139
ตารางที่ 5.2 กลุ่มประโยชน์ที่คล้ายคลึงในแต่ละเทคนิคของ Apple	139
ตารางที่ 5.3 กลุ่มผลเสียที่คล้ายคลึงในแต่ละเทคนิคของ Apple	139
ตารางที่ 5.4 กลุ่มหัวข้อหัวเรื่องที่คล้ายคลึงในแต่ละเทคนิคของ Google	140
ตารางที่ 5.5 กลุ่มประโยชน์ที่คล้ายคลึงในแต่ละเทคนิคของ Google	140
ตารางที่ 5.6 กลุ่มผลเสียที่คล้ายคลึงในแต่ละเทคนิคของ Google	140
ตารางที่ 5.7 กลุ่มหัวข้อหัวเรื่องที่คล้ายคลึงในแต่ละเทคนิคของ Microsoft	140
ตารางที่ 5.8 กลุ่มประโยชน์ที่คล้ายคลึงในแต่ละเทคนิคของ Microsoft	140
ตารางที่ 5.9 กลุ่มผลเสียที่คล้ายคลึงในแต่ละเทคนิคของ Microsoft	141

## สารบัญภาพ

	หน้า
ภาพที่ 3.1 แผนผังวิธีการดำเนินงานวิจัย	9
ภาพที่ 3.2 การนำชุดข้อมูล import เข้า Google Colab	10
ภาพที่ 3.3 การกรองข้อมูลเฉพาะชื่อบริษัทที่เลือก	10
ภาพที่ 3.4 ลบรายการคอลัมน์ที่ไม่ได้นำมาวิเคราะห์	10
ภาพที่ 3.5 การเปลี่ยนจาก DataFrame เป็น List	10
ภาพที่ 3.6 การเตรียมข้อมูลและทำความสะอาดข้อมูล	11
ภาพที่ 3.7 การปรับรูปแบบของคำให้เหมือนกัน	11
ภาพที่ 3.8 การตรวจสอบข้อมูลตัวอักษรที่ซ้ำกัน	11
ภาพที่ 3.9 ตัวอย่างการลบตัวอักษรซ้ำ	12
ภาพที่ 3.10 การลบตัวเลขออกจากข้อความและลบอักขระหรือสัญลักษณ์ต่างๆ	12
ภาพที่ 3.11 ตัวอย่างการลบตัวเลขออกจากข้อความและลบอักขระหรือสัญลักษณ์ต่างๆ	12
ภาพที่ 3.12 กระบวนการตัดคำ	12
ภาพที่ 3.13 การลบ Stop words	12
ภาพที่ 3.14 กระบวนการ lemmatization	13
ภาพที่ 3.15 ตัวอย่างการ lemmatization	13
ภาพที่ 3.16 ตัวอย่างการ lemmatization	13
ภาพที่ 3.17 คัดแยกข้อมูล	14
ภาพที่ 3.18 library ที่นำมาใช้ ของ LDA	14
ภาพที่ 3.19 การกำหนดจำนวนหัวข้อ LDA	15
ภาพที่ 3.20 library ที่นำมาใช้ ของ Top2Vec	15
ภาพที่ 3.21 การกำหนดสร้างตัวแบบหัวข้อของเทคนิค Top2vec	15
ภาพที่ 3.22 library ที่นำมาใช้ ของ BERTopic	15
ภาพที่ 3.23 การกำหนดสร้างตัวแบบหัวข้อของเทคนิค BERTopic	16
ภาพที่ 4.1 แสดงการจัดกลุ่มผลลัพธ์หัวข้อของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท apple	41
ภาพที่ 4.2 แสดงผลลัพธ์แต่ละหัวข้อในกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท apple	41



ภาพที่ 4.18 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท apple	49
ภาพที่ 4.19 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท headline ของบริษัท apple	50
ภาพที่ 4.20 แสดงการจัดกลุ่มผลลัพธ์หัวข้อของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท apple	50
ภาพที่ 4.21 แสดงผลลัพธ์หัวข้อที่ 0 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท apple	50
ภาพที่ 4.22 แสดงผลลัพธ์การนับความถี่ของคำในหัวข้อที่ 0 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท apple	51
ภาพที่ 4.23 word cloud แสดงคำทุกคำที่ปรากฏในหัวข้อที่ 0 ในข้อมูลประเภท pros ของบริษัท apple	52
ภาพที่ 4.24 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อที่ 0 ของเทคนิค Top2vec ของข้อมูลประเภท pros ของบริษัท apple	52
ภาพที่ 4.25 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในหัวข้อที่ 0 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท apple	52
ภาพที่ 4.26 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในหัวข้อที่ 0 ในข้อมูลประเภท pros ของบริษัท apple	53
ภาพที่ 4.27 แสดงผลลัพธ์หัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท apple	53
ภาพที่ 4.28 แสดงผลลัพธ์การนับความถี่ของคำในหัวข้อที่ 0 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท apple	54
ภาพที่ 4.29 word cloud แสดงคำทุกคำที่ปรากฏในหัวข้อที่ 1 ในข้อมูลประเภท pros ของบริษัท apple	54
ภาพที่ 4.30 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อที่ 1 ของเทคนิค Top2vec ของข้อมูลประเภท headline ของบริษัท apple	54
ภาพที่ 4.31 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท apple	55
ภาพที่ 4.32 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในหัวข้อที่ 1 ในข้อมูลประเภท pros ของบริษัท apple	55

ภาพที่ 4.33 แสดงการจัดกลุ่มผลลัพธ์หัวข้อของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท apple	56
ภาพที่ 4.34 แสดงผลลัพธ์หัวข้อที่ 0 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท apple	56
ภาพที่ 4.35 แสดงผลลัพธ์การนับความถี่ของคำในหัวข้อที่ 0 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท apple	56
ภาพที่ 4.36 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 0 ในข้อมูลประเภท cons ของบริษัท apple	57
ภาพที่ 4.37 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อที่ 0 ของเทคนิค Top2vec ของข้อมูลประเภท cons ของบริษัท apple	57
ภาพที่ 4.38 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในหัวข้อที่ 0 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท apple	57
ภาพที่ 4.39 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในหัวข้อที่ 0 ในข้อมูลประเภท cons ของบริษัท apple	58
ภาพที่ 4.40 แสดงผลลัพธ์หัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท apple	58
ภาพที่ 4.41 แสดงผลลัพธ์การนับความถี่ของคำในหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท apple	59
ภาพที่ 4.42 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท cons ของบริษัท apple	59
ภาพที่ 4.43 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อที่ 1 ของเทคนิค Top2vec ของข้อมูลประเภท cons ของบริษัท apple	59
ภาพที่ 4.44 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท apple	60
ภาพที่ 4.45 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในหัวข้อที่ 1 ในข้อมูลประเภท cons ของบริษัท apple	60
ภาพที่ 4.46 แสดงผลลัพธ์หัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท apple	61
ภาพที่ 4.47 แสดงผลลัพธ์การนับความถี่ของคำในหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท apple	61

ภาพที่ 4.48 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท cons ของบริษัท apple	62
ภาพที่ 4.49 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อที่ 2 ของเทคนิค Top2vec ของข้อมูลประเภท cons ของบริษัท apple	62
ภาพที่ 4.50 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท apple	62
ภาพที่ 4.51 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในหัวข้อที่ 2 ในข้อมูลประเภท cons ของบริษัท apple	63
ภาพที่ 4.52 แสดงการจัดกลุ่มผลลัพธ์หัวข้อของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท google	63
ภาพที่ 4.53 แสดงผลลัพธ์แต่ละหัวข้อในกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท google	64
ภาพที่ 4.54 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท google	64
ภาพที่ 4.55 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท headline ของบริษัท google	65
ภาพที่ 4.56 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ	65
ในกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ของข้อมูลประเภท headline ของบริษัท google	
ภาพที่ 4.57 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1	66
ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท google	
ภาพที่ 4.58 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท headline ของบริษัท google	66
ภาพที่ 4.59 แสดงผลลัพธ์แต่ละหัวข้อในกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท google	67
ภาพที่ 4.60 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท google	67
ภาพที่ 4.61 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท headline ของบริษัท google	68
ภาพที่ 4.62 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ	68
ในกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ของข้อมูลประเภท headline ของบริษัท google	

ภาพที่ 4.63 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท google	69
ภาพที่ 4.64 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท headline ของบริษัท google	69
ภาพที่ 4.65 แสดงผลลัพธ์แต่ละหัวข้อในกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท google	70
ภาพที่ 4.66 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท google	70
ภาพที่ 4.67 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ใน ข้อมูลประเภท headline ของบริษัท google	71
ภาพที่ 4.68 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ ในกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ของข้อมูลประเภท headline ของบริษัท google	71
ภาพที่ 4.69 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท google	71
ภาพที่ 4.70 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท headline ของบริษัท google	72
ภาพที่ 4.71 แสดงการจัดกลุ่มผลลัพธ์หัวข้อของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท google	72
ภาพที่ 4.72 แสดงผลลัพธ์แต่ละหัวข้อในกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท google	73
ภาพที่ 4.73 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท google	74
ภาพที่ 4.74 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ใน ข้อมูลประเภท pros ของบริษัท google	74
ภาพที่ 4.75 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ ในกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ของข้อมูลประเภท pros ของบริษัท google	75
ภาพที่ 4.76 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท google	75

ภาพที่ 4.77 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท pros ของบริษัท google	76
ภาพที่ 4.78 แสดงผลลัพธ์แต่ละหัวข้อในกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท google	77
ภาพที่ 4.79 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท google	78
ภาพที่ 4.80 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท pros ของบริษัท google	78
ภาพที่ 4.81 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อในกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท google	79
ภาพที่ 4.82 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท google	79
ภาพที่ 4.83 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท pros ของบริษัท google	80
ภาพที่ 4.84 แสดงผลลัพธ์แต่ละหัวข้อในกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท google	80
ภาพที่ 4.85 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท google	81
ภาพที่ 4.86 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท pros ของบริษัท google	82
ภาพที่ 4.87 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อในกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท google	82
ภาพที่ 4.88 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท google	82
ภาพที่ 4.89 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท pros ของบริษัท google	83
ภาพที่ 4.90 แสดงการจัดกลุ่มผลลัพธ์หัวข้อของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท google	83

ภาพที่ 4.91 แสดงผลลัพธ์แต่ละหัวข้อในกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท google	84
ภาพที่ 4.92 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท google	85
ภาพที่ 4.93 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท cons ของบริษัท google	85
ภาพที่ 4.94 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อในกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท google	86
ภาพที่ 4.95 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท google	86
ภาพที่ 4.96 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท cons ของบริษัท google	87
ภาพที่ 4.97 แสดงผลลัพธ์แต่ละหัวข้อในกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท google	87
ภาพที่ 4.98 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท google	88
ภาพที่ 4.99 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท cons ของบริษัท google	89
ภาพที่ 4.100 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อในกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท google	89
ภาพที่ 4.101 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท google	90
ภาพที่ 4.102 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท cons ของบริษัท google	90
ภาพที่ 4.103 แสดงผลลัพธ์แต่ละหัวข้อในกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท google	91
ภาพที่ 4.104 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท google	92

ภาพที่ 4.105 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท cons ของบริษัท google	92
ภาพที่ 4.106 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อในกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ของข้อมูลประเภท cons ของบริษัท google	93
ภาพที่ 4.107 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท google	93
ภาพที่ 4.108 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท cons ของบริษัท google	94
ภาพที่ 4.109 แสดงการจัดกลุ่มผลลัพธ์หัวข้อของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท microsoft	94
ภาพที่ 4.110 แสดงผลลัพธ์แต่ละหัวข้อในกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท microsoft	95
ภาพที่ 4.111 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท Microsoft	95
ภาพที่ 4.112 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท headline ของบริษัท microsoft	96
ภาพที่ 4.113 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อในกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ของข้อมูลประเภท headline ของบริษัท microsoft	96
ภาพที่ 4.114 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท microsoft	97
ภาพที่ 4.115 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท headline ของบริษัท Microsoft	97
ภาพที่ 4.116 แสดงผลลัพธ์แต่ละหัวข้อในกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท microsoft	98
ภาพที่ 4.117 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท microsoft	98
ภาพที่ 4.118 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท headline ของบริษัท microsoft	99

ภาพที่ 4.119 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ ในกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ของข้อมูลประเภท headline ของบริษัท microsoft	99
ภาพที่ 4.120 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท Microsoft	100
ภาพที่ 4.121 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท headline ของบริษัท microsoft	100
ภาพที่ 4.122 แสดงผลลัพธ์แต่ละหัวข้อในกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท microsoft	101
ภาพที่ 4.123 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท microsoft	101
ภาพที่ 4.124 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ใน ข้อมูลประเภท headline ของบริษัท microsoft	102
ภาพที่ 4.125 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ ในกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ของข้อมูลประเภท headline ของบริษัท microsoft	102
ภาพที่ 4.126 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท microsoft	103
ภาพที่ 4.127 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท headline ของบริษัท microsoft	103
ภาพที่ 4.128 แสดงการจัดกลุ่มผลลัพธ์หัวข้อของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท Microsoft	104
ภาพที่ 4.129 แสดงผลลัพธ์แต่ละหัวข้อในกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท microsoft	104
ภาพที่ 4.130 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท microsoft	105
ภาพที่ 4.131 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อที่ 1 ของเทคนิค Top2vec ของข้อมูลประเภท cons ของบริษัท apple	105
ภาพที่ 4.132 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ	106

ในกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ของข้อมูลประเภท pros ของบริษัท microsoft ภาพที่ 4.133 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1	106
ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท microsoft ภาพที่ 4.134 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท microsoft	107
ภาพที่ 4.135 แสดงผลลัพธ์หัวข้อในกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท microsoft	107
ภาพที่ 4.136 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท microsoft	108
ภาพที่ 4.137 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ใน ข้อมูลประเภท pros ของบริษัท Microsoft	108
ภาพที่ 4.138 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ ในกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ของข้อมูลประเภท pros ของบริษัท microsoft	109
ภาพที่ 4.139 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2	109
ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท microsoft ภาพที่ 4.140 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท pros ของบริษัท microsoft	109
ภาพที่ 4.141 แสดงผลลัพธ์แต่ละหัวข้อในกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท microsoft	110
ภาพที่ 4.142 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท microsoft	111
ภาพที่ 4.143 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ใน ข้อมูลประเภท pros ของบริษัท Microsoft	111
ภาพที่ 4.144 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ ในกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ของข้อมูลประเภท pros ของบริษัท microsoft	112
ภาพที่ 4.145 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3	112
ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท microsoft	

ภาพที่ 4.146 word cloud และคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท pros ของบริษัท microsoft	112
ภาพที่ 4.147 แสดงการจัดกลุ่มผลลัพธ์หัวข้อของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท microsoft	113
ภาพที่ 4.148 แสดงผลลัพธ์แต่ละหัวข้อในกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท microsoft	113
ภาพที่ 4.149 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท microsoft	114
ภาพที่ 4.150 word cloud และคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท cons ของบริษัท microsoft	114
ภาพที่ 4.151 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อในกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ของข้อมูลประเภท cons ของบริษัท microsoft	115
ภาพที่ 4.152 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1	115
ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท microsoft	
ภาพที่ 4.153 word cloud และคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท pros ของบริษัท Microsoft	116
ภาพที่ 4.154 แสดงผลลัพธ์แต่ละหัวข้อในกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท microsoft	117
ภาพที่ 4.155 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท microsoft	118
ภาพที่ 4.156 word cloud และคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท cons ของบริษัท microsoft	118
ภาพที่ 4.157 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อในกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ของข้อมูลประเภท cons ของบริษัท microsoft	119
ภาพที่ 4.158 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2	119
ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท microsoft	
ภาพที่ 4.159 word cloud และคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท cons ของบริษัท microsoft	120

ภาพที่ 4.160 แสดงผลลัพธ์แต่ละหัวข้อในกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท microsoft	120
ภาพที่ 4.161 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท microsoft	121
ภาพที่ 4.162 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ใน ข้อมูลประเภท cons ของบริษัท microsoft	121
ภาพที่ 4.163 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ ในกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ของข้อมูลประเภท cons ของบริษัท microsoft	122
ภาพที่ 4.164 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3	122
ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท microsoft	
ภาพที่ 4.165 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท cons ของบริษัท microsoft	123
ภาพที่ 4.166 ผลการวิเคราะห์คำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล headline บริษัท apple	124
ภาพที่ 4.167 ผลการวิเคราะห์คำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล รีวิว ประโยชน์ที่ได้ บริษัท apple	126
ภาพที่ 4.168 ผลการวิเคราะห์คำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล รีวิว ผลเสียที่ได้ บริษัท apple	127
ภาพที่ 4.169 ผลการวิเคราะห์คำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล headline บริษัท Google	128
ภาพที่ 4.170 ผลการวิเคราะห์คำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล รีวิว ประโยชน์ที่ได้ บริษัท Google	130
ภาพที่ 4.171 ผลการวิเคราะห์คำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล รีวิว ผลเสียที่ได้ บริษัท Google	131
ภาพที่ 4.172 ผลการวิเคราะห์คำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล headline บริษัท Microsoft	133
ภาพที่ 4.173 ผลการวิเคราะห์คำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล รีวิว ประโยชน์ที่ได้ บริษัท Microsoft	134
ภาพที่ 4.174 ผลการวิเคราะห์คำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล รีวิว ผลเสียที่ได้ บริษัท Microsoft	136

ภาพที่ 4.175 ตารางผลลัพธ์ของการประเมินประสิทธิภาพเชิงปริมาณ

138

## รายการสัญลักษณ์และคำย่อ

### สัญลักษณ์/คำย่อ

NLP

PoS

LDA

### คำเต็ม/คำจำกัดความ

Natural language processing

Part of Speech Tagging

Latent Dirichlet Allocation

## บทที่ 1 บทนำ

### 1.1 ความเป็นมาและความสำคัญของโครงการ

ผู้วิจัยมีความสนใจในเรื่องรีวิวการทำงาน (Job review) บนแพลตฟอร์มออนไลน์ โดยผู้ให้บริการจัดหางานในที่นี้คือเว็บไซต์ Glassdoor.com เป็นเว็บไซต์หางานยอดนิยมในสหรัฐอเมริกา และเป็นพื้นที่ให้รีวิวการทำงานได้อย่างอิสระ อีกทั้งยังมีความแพร่หลายในการใช้งานที่อีกมาก many และคนไทยใช้งานด้วยเช่นกันดังนั้นสิ่งที่ผู้วิจัยจะศึกษาดังนี้คือการสำรวจข้อความรีวิวการทำงาน (Job review) บนเว็บไซต์ Glassdoor.com ว่าข้อความรีวิวล่าwiększเกี่ยวกับงานและบริษัทอย่างไร แต่การที่จะทราบความเห็นจากข้อความรีวิวที่มีจำนวนมากกว่า 800,000 ข้อความในชุดข้อมูล Glassdoor นั้นเป็นเรื่องที่ท้าทายเช่นกัน ผู้วิจัยจึงอาศัยประโยชน์จากการทำ Topic Modeling แบบจำลองหัวข้อ (Topic Modeling) เป็นการทำประดิณสำคัญที่ช่อนอยู่อุปแบบของหัวข้อ จากชุดเอกสารที่มีจำนวนข้อมูลมากๆ โดยใช้เทคนิคในการสร้างแบบจำลองหัวข้อ 3 เทคนิคได้แก่ 1) การจัดสรรดีรีเคลล์เพง(Latent Dirichlet Allocation), 2) Top2vec, 3) BERTopic ดังนั้นเพื่อที่จะทราบสิ่งที่รีวิวโดยรวมกล่าวถึงงานและบริษัทผ่านผลลัพธ์ที่อยู่ในรูปแบบหัวข้อและคาดหวังว่าจะให้ผลลัพธ์ที่สามารถอธิบายได้ว่ารีวิวโดยรวมกล่าวถึงงานและบริษัทอย่างไร

เพื่อที่จะทราบว่ารีวิวโดยรวมกล่าวถึงการทำงานและบริษัทอย่างไรนั้นผู้วิจัยจะใช้การสร้างแบบจำลองหัวข้อ (Topic Modeling) เพื่อจัดกลุ่มข้อความในที่นี้คือ ข้อความที่เป็นการรีวิวการทำงาน (Job review) ทั้งหมด แต่ก่อนการสร้างแบบจำลองหัวข้อนั้นผู้วิจัยจะทำการ Filter ข้อมูลในชุดข้อมูล Glassdoor โดยแบ่งตามบริษัทซึ่งเลือกมา 3 บริษัทได้แก่ Apple, Google, Microsoft หลังจากนั้น ผู้วิจัยจะทำการคัดแยกข้อมูลอุปมา Extracting Data จะได้ชุดข้อมูลที่มีข้อมูลเฉพาะที่เลือกมา 3 บริษัท โดยในโครงการนี้ผู้วิจัยใช้เทคนิคในการสร้างแบบจำลองหัวข้อ 3 เทคนิคได้แก่ 1) การจัดสรรดีรีเคลล์เพง (Latent Dirichlet Allocation), 2) Top2vec, 3) BERTopic เพื่อหาประเด็นสำคัญที่อยู่ในรูปแบบหัวข้อ (หัวข้อประกอบไปด้วยคำหลักคำที่แสดงความเป็นหัวข้อนั้นๆ) และทำการวิเคราะห์และตีความคำที่อยู่ภายในหัวข้อเหล่านั้นเพื่อที่จะสามารถเข้าใจว่ารีวิวโดยรวมกล่าวเกี่ยวกับงานและบริษัทอย่างไรผ่านผลลัพธ์ที่อยู่ในรูปแบบหัวข้อ ซึ่งอ้างอิงจากงานวิจัย [1] วรอัปสร บุบพาเพื่อที่จะเข้าใจถึงบริบทของการสนทนารึใช้การสรุปแบบอุปนัย กล่าวคือ ตีความสร้างข้อสรุปข้อมูลจากปракติกการณ์ที่มองเห็น หมายความว่า ผู้วิจัยจะวิเคราะห์ตีความจากคำที่แสดงในแต่ละผลลัพธ์หัวข้อนั้นๆและตั้งสมมติฐานว่า หัวข้อนั้นๆควรจะสื่อถึงอะไร โดยชุดข้อมูลที่นำมาศึกษานั้นมีข้อมูลที่เป็นข้อความประเภทรีวิวการทำงาน (Job review) ซึ่งมีอยู่ 3 ส่วนในชุดข้อมูล Glassdoor นี้คือ headline (ชื่อรีวิว), pros (รีวิวประโยชน์ที่ได้), cons (รีวิวผลเสียที่ได้) โดยจะสร้างแบบจำลองหัวข้อกับข้อความรีวิวทั้ง 3 ส่วนนี้เพื่อที่จะทราบสิ่งที่รีวิวโดยรวมในแต่ละส่วนนั้นกล่าวถึงงานและบริษัททั้ง 3 อย่างไร

ส่วนการประเมินประสิทธิภาพของแบบจำลองหัวข้อนั้นจะพิจารณาใน 2 ด้านคือ เชิงคุณภาพ และเชิงปริมาณ โดยในด้านเชิงคุณภาพผู้วิจัยจะพิจารณาจากผลลัพธ์หัวข้อทั้ง 3 เทคนิค ซึ่งในแต่ละเทคนิคความยากง่ายในการวิเคราะห์ตีความผลลัพธ์หัวข้อและตั้งสมมติฐานอย่างไร ในด้านเชิง

ปริมาณผู้วิจัยจะประเมินจากค่า Coherence score ค่าความสอดคล้อง (Coherence score) คือ วิธีประเมินประสิทธิภาพของตัวแบบหัวข้อโดยให้ผลลัพธ์เป็นค่าความน่าจะเป็น (Probability) ค่าที่ยิงเข้าใกล้ 1 แสดงถึงความสอดคล้องของคำในแต่ละผลลัพธ์หัวข้อนั้นๆ และมีความหมายอีกนัยว่ามนุษย์สามารถตีความหัวข้อเหล่านี้ได้มากน้อยเพียงใดตามค่า score ที่แสดงเป็นเชิง probability

ในการวิเคราะห์ผู้วิจัยคาดหวังผลลัพธ์ที่สามารถเป็นประโยชน์ให้กับผู้ที่กำลังสนใจเรื่อง การวิเคราะห์ผลลัพธ์จากการทำแบบจำลองหัวข้อ (Topic Modeling) กับข้อความประเภทรีวิวการทำงาน (Job Review) ได้และผู้ที่สนใจสามารถนำไปเป็นแนวทางในการศึกษาและปรับปรุงต่ออยู่ได้

## 1.2 วัตถุประสงค์

1. เพื่อศึกษาการสร้างแบบจำลองหัวข้อ (Topic modeling) ที่ใช้เทคนิค LDA, Top2vec, BERTopic
2. เพื่อวิเคราะห์ข้อความรีวิวนะเด็ข้อมูล Glassdoor โดยการวิเคราะห์จากหัวข้อที่ถูกค้นพบ โดยการทำแบบจำลองหัวข้อ (Topic modeling)
3. เพื่อประเมินประสิทธิภาพของเทคนิคการสร้างแบบจำลองหัวข้อ (Topic modeling) สำหรับหัวข้อที่ค้นพบจากข้อความรีวิวการทำงาน

## 1.3 ขอบเขตโครงการ

1. โครงการนี้ศึกษาเรื่องการทำงานในบริษัทต่างๆ บนเว็บไซต์ Glassdoor.com โดยชุดข้อมูลมีช่วงเวลาระหว่างปี 2010 ถึง 2021 โดยเลือกใช้ข้อมูลรีวิวเฉพาะของบริษัท apple, google, microsoft
2. โครงการนี้ประเมินประสิทธิภาพตัวแบบจำลองหัวข้อ (Topic modeling) ทั้ง 3 เทคนิค ได้แก่ LDA, Top2vec, BERTopic

## 1.4 ประโยชน์ที่คาดหวัง

1. Topic modeling ที่ใช้เทคนิค LDA, Top2vec, BERTopic ในเชิงลึกมากขึ้น
2. คาดหวังว่าสามารถประเมินประสิทธิภาพของตัวแบบจำลองหัวข้อ (Topic modeling) และวิเคราะห์ทั้งในเชิงคุณภาพและเชิงปริมาณได้บนชุดข้อมูล Glassdoor ด้วยชุด dataset ของ apple, google, microsoft
3. คาดหวังว่าสามารถช่วยวิเคราะห์สภาพการทำงานภายในของบริษัทต่างๆ ผ่านผลลัพธ์ของการสร้างแบบจำลองหัวข้อในรูปแบบหัวข้อ

## บทที่ 2

### วรรณกรรมและงานวิจัยที่เกี่ยวข้อง

#### 2.1 แนวคิดทฤษฎีที่เกี่ยวข้อง

##### 2.1.1 การประมวลผลภาษาธรรมชาติ (Natural language processing: NLP)

การประมวลผลภาษาธรรมชาติเป็นการที่ระบบคอมพิวเตอร์เข้าใจภาษาตามนุxyz โดยมีการจัดการกับข้อมูลที่คล้ายกับภาษาตามนุxyzในลักษณะที่แบบทางการ เขียนและพูด รวมถึงการรับรู้ถึงความหมายโดยนัย ความรู้สึกของผู้เขียน ความแตกต่างทางบริบทของภาษาสามารถทำการวิเคราะห์ในรูปแบบต่าง ๆ

เนื่องจากข้อมูลในรูปแบบของข้อความสามารถทำการวิเคราะห์ได้ในหลากหลายมุมมอง เครื่องมือสำหรับงาน NLP จึงมีความหลากหลายเช่นเดียวกัน โดยในการทำโครงงานอาจมีความจำเป็นที่ต้องใช้เครื่องมือมากกว่าหนึ่งรายการ ซึ่งในที่นี้ จะมีการแนะนำเครื่องมือในโครงงานนี้

###### 2.1.1.1 การตัดคำ (Tokenization)

กระบวนการตัดคำ Tokenization คือ การตัดคำออกจากข้อความยาว ๆ ซึ่งแต่ละคำที่ถูกตัดออกมาจะถูกเรียกว่า Token โดยกระบวนการนี้มักจะถูกทำเป็นกระบวนการนี้แรก ๆ และ Token ที่ได้มาจะถูกนำไปใช้ในเครื่องมือ NLP อื่น ๆ ต่อไป

###### 2.1.1.2 การระบุประเภทของคำ (Part of Speech Tagging :PoS Tagging)

กระบวนการในการระบุประเภทของคำ Part of Speech Tagging (PoS Tagging) คือ แต่ละคำตามหน้าที่ในประโยค อาทิ คำนาม (Nouns), คำสรรพนาม (Pronouns), คำคุณศัพท์ (Adjectives), คำกริยา (Verbs), คำกริยาวิเศษณ์ (Adverbs), คำบุพบท (Prepositions), คำสันฐาน (Conjunctions), และ คำอุทาน (Interjections)

###### 2.1.1.3 การแปลงคำให้อยู่รูปพื้นฐานของคำ (Lemmatization)

เป็นกระบวนการที่ใช้คำศัพท์และการวิเคราะห์ทางสัณฐานวิทยา (morphological analysis) ของคำเพื่อลบจุดสิ้นสุดที่ผันกลับมาเพื่อให้ได้รูปแบบพื้นฐาน (รูปแบบพจนานุกรม) ซึ่งรู้จักกันในชื่อ บทแทรก (lemma) มันเป็นกระบวนการที่ซับซ้อนซึ่งต้องการความเข้าใจในบริบทที่คำต่าง ๆ pragmatics เพื่อตัดสินใจเกี่ยวกับความหมายดังนั้น มันจึงใช้คำศัพท์เพื่อให้ได้รูปแบบราก (root form)

###### 2.1.1.4 แบบจำลองหัวข้อ (Topic Modeling)

แบบจำลองหัวข้อ (Topic Modeling) เป็นเทคนิคการเรียนรู้ของเครื่อง (Machine Learning) เป็นการเรียนรู้ของเครื่องแบบ Unsupervised ใช้ในการวิเคราะห์ข้อมูลตัวอักษรเพื่อกำหนดกลุ่มคำสำหรับชุดเอกสาร (Document) เป็นการจำลองหัวข้อแบบการกระจายตัวของข้อมูล เพื่อนำไปใช้ในการจัดกลุ่มข้อมูล โดยเอกสารแต่ละเอกสารจะประกอบด้วยการรวมกัน หลายหัวข้อ แต่ละหัวข้อมีการแจกแจงความน่าจะเป็นของคำนั้นๆที่เกิดขึ้นในแต่ละหัวข้อ

แบบจำลองหัวข้อสามารถทำได้หลากหลายวิธี ซึ่งจำเป็นที่จะต้องเลือกวิธีที่ดีและเหมาะสมกับข้อมูล เพื่อให้ไม่เดลอุ่นมา มีคุณภาพมากที่สุด ซึ่งแต่ละรูปแบบก็จะมีข้อดี ข้อเสียแตกต่างกันออกไป ดังรายละเอียดต่อไปนี้

### ก. การจัดสรรดีรีเคลแฟง (Latent Dirichlet Allocation หรือ LDA)

การจัดสรรดีรีเคลแฟง (Latent Dirichlet Allocation หรือ LDA) เป็นแบบจำลองความน่าจะเป็นสำหรับการทำแบบจำลองหัวข้อ (Topic Modeling) เป็นการหาหัวข้อ (topic) ที่ซ่อนอยู่ ในเอกสาร โดยมีหัวข้อที่แฟงอยู่ในเอกสารถูกสร้างจากแนวคิดที่ว่า ในเอกสารนั้นจะประกอบไปด้วยหัวข้อที่ถูกรวบกันแบบไม่มีรูปแบบและกระจายในกลุ่มของคำศัพท์ของเอกสารนั้น โดย การทำ LDA จะทำการหาค่าความน่าจะเป็นของความสัมพันธ์ของหัวข้อของแต่ละประโยชน์และค่าความน่าจะเป็นของคำในแต่ละหัวข้อ ลักษณะของเทคนิคนี้เป็นแบบการเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) และใช้หลักความน่าจะเป็นในการหาพารามิเตอร์ (Probabilistic Model) หลักการทำงานของ LDA ได้แก่

- (1) เลือกพารามิเตอร์ โดยกำหนดจำนวนหัวข้อตัวอย่างเช่น Clustering algorithms เช่น K-mean ที่เลือกจำนวน Cluster หรือค่า k
- (2) LDA จะทำการอ่านทุกคำของแต่ละเอกสารและกำหนดหัวข้อ โดยผลลัพธ์ที่ได้คือ กลุ่มคำของหัวข้อ (Topic Representation) ที่อยู่รวมกันใน ถุงคำ (bag of words) และหัวข้อของแต่ละเอกสาร

จำนวนหัวข้อที่ดีที่สุดมี 2 วิธี คือการหา Perplexity และ Log-Likelihood function

วิธีแรก Perplexity หาจำนวน k ที่เหมาะสมและประเมินประสิทธิภาพของโมเดลในการอธิบายข้อมูล ค่า Perplexity ต่ำมากก็หมายความว่าโมเดลนั้นมีประสิทธิภาพดีมาก วิธีการหาค่า นี้เมื่อมีการกำหนดค่า k จากการ LDA ได้กลุ่มของคำ นำคำเหล่านี้ที่ได้กำหนดความเชื่อมโยงกับแต่ละหัวข้อเปรียบเทียบกับการกระจายของคำในเอกสาร เพื่อหา cluster ของหัวข้อของเอกสาร นั้นๆ สามารถอธิบายเป็นสมการได้ดังนี้

$$\text{perplexity}(D_{test}) = \exp\left(-\frac{\sum_{d=1}^m \log[p(w_d)]}{\sum_{d=1}^m N_d}\right)$$

วิธีที่สอง Log-Likelihood function การหาค่าความใกล้เคียงโดยผลลัพธ์ที่ได้หากมีค่า สูง แสดงว่ามีความใกล้เคียงกันมาก

$$L(\theta, m | D) = p_{\theta, m}(d) = P_{\theta, m}(D = d)$$

$$L(\theta, m | D) = p(\theta, m | D)$$

## ข. Top2Vec

Top2Vec คือ อัลกอริทึมสำหรับการสร้างแบบจำลองหัวข้อและการค้นหาความหมาย ตรวจจับหัวข้อที่มีอยู่ในข้อความโดยอัตโนมัติและสร้างหัวข้อเอกสารและเวกเตอร์คำที่ผังร่วมกัน อัลกอริทึมจะดำเนินการตามขั้นตอนต่อไปนี้เพื่อค้นหาหัวข้อในรายการเอกสาร

(1) สร้างเวกเตอร์ผังสำหรับเอกสารและคำ: การแปลงเอกสารทั้งหมดให้เป็นตัวเลขจะถูกแปลงรวมเป็นเวกเตอร์ผัง โดยเวกเตอร์ผังจะช่วยสามารถแสดงเอกสารคำหรือข้อความในพื้นที่หลายมิติ แนวคิดเบื้องหลังการผังเวกเตอร์คือคำหรือเอกสารข้อความ ที่คล้ายกันจะมีเวกเตอร์ที่คล้ายกัน การสร้างเวกเตอร์ผังสำหรับแต่ละเอกสารทำให้ถือว่าแต่ละเอกสารเป็นจุดในพื้นที่หลายมิติ Top2Vec ยังสร้างเวกเตอร์คำที่ผังร่วมกัน ซึ่งช่วยให้สามารถกำหนดคำหลักของหัวข้อได้ในภายหลัง

(2) การลดขนาดเวกเตอร์โดยใช้อัลกอริทึม: หลังจากที่มีเวกเตอร์แต่ละเอกสารแล้ว ขั้นตอนต่อไปคือการแบ่งออกเป็นกลุ่มโดยใช้อัลกอริทึมการจัดกลุ่มนี้เนื่องจากเวกเตอร์ที่สร้างขึ้นในขั้นตอนแรกมีมิติจำนวนมากจากการผังเอกสารโดยขึ้นอยู่กับรูปแบบการผังที่ใช้ ด้วยเหตุผลนี้ การทำอัลกอริทึมที่ลดขนาดบางประเภทเพื่อลดจำนวนมิติในข้อมูลจึงเป็นเรื่องสมเหตุผล Top2Vec ใช้อัลกอริทึมที่เรียกว่า UMAP ซึ่งจะต้องถูกบีบให้เรียบ จนนั้นจึงสามารถสร้างกราฟเป็นสองมิติหรือสามมิติให้เข้าใจได้ง่าย

(3) จัดกลุ่มเวกเตอร์โดยใช้อัลกอริทึมการจัดกลุ่ม: เมื่อเอกสารทั้งหมดถูกจัดลงในมิติแล้วจะค้นหาพื้นที่ที่มีความเข้มข้นหรือบริเวณที่เป็นกระจุกจะบ่งบอกถึงความคล้ายคลึงกันถึงความหมายที่มีอยู่

(4) กำหนดหัวข้อให้กับแต่ละคลัสเตอร์: เอกสารที่มีความคล้ายคลึงกันจะมีหัวข้อที่แฟรงอยู่ การค้นหาหัวข้อที่แฟรงในความคล้ายคลึงของกลุ่มเอกสารที่หนาแน่นจะระบุว่าคำที่ดึงดูดเอกสารเหล่านั้นเข้าด้วยกันแต่ละคำที่ดึงดูดเอกสารไปยังพื้นที่หนาแน่นเป็นหัวข้อ

## ค. BERTopic

โดยสามองค์ประกอบแรกเหมือนกับ Top2Vec องค์ประกอบที่แตกต่างหลังมีดังนี้

(1) cTF-IDF: เป็นขั้นตอน TF-IDF เทคนิคการคัดแยกคำตามความสำคัญโดยการให้น้ำหนักคำในแต่ละคลัสเตอร์

โดยใช้ 2 ปัจจัยคือ TF และ IDF

TF มาจาก Term Frequency หรือความถี่ของคำศัพท์โดยจุดประสงค์ของ TF นั้น เพื่อคุณว่าคำแต่ละคำนั้นปรากฏบ่อยแค่ไหนในแต่ละ Document

การหาค่า TF หาได้จากสูตร

$$TF(t,d) = \frac{\text{จำนวนคำที่ปรากฏใน Document}}{\text{จำนวนคำทั้งหมดใน Document}}$$

IDF ย่อมาจาก Inverse Document Frequency หรือ การผกผันในความถี่ของเอกสาร ซึ่งมีที่มาจากการผกผันของเอกสารที่มีค่า TF ไม่ตอบว่าเป็นคำที่สำคัญ เพราะ TF จะ focus ในแต่ละ Document แต่ IDF จะ เป็นค่าที่คำนวณจาก Document ทั้งหมด จึงมีความเป็น Generalize มากกว่าและใช้ปัจจัยความเป็น General ได้ในระดับหนึ่ง ค่า IDF สามารถคำนวณได้จาก

$$IDF(t) = \log \left[ \frac{(1 + n)}{(1 + DF(t))} \right] + 1$$

จากนั้น นำค่า TF และ IDF มาคูณกันเพื่อให้ weight ของ TF และ weight ของ IDF ที่ซึ่งมีค่าตรงข้ามกันนั้นเป็นนำหนักที่สามารถแยกคำสำคัญออกจากมาได้จริง

### 2.1.2 ค่า Topic Coherence Cv

Topic Coherence Cv เป็นเครื่องมือที่ใช้ในการวัดคุณภาพของหัวข้อที่ถูกสร้างขึ้นจากระบบคอมพิวเตอร์หรือโมเดลที่ทำงานเกี่ยวกับการจัดหัวข้อ และมีหลายวิธีในการวัดความสัมพันธ์ระหว่างคำศัพท์ในหัวข้อนั้น ๆ ซึ่ง Normalized Pointwise Mutual Information (NPMI) เป็นหนึ่งในวิธีที่นิยม

Normalized Pointwise Mutual Information (NPMI) เป็นหนึ่งในวิธีการวัดความสัมพันธ์ หรือความเชื่อมโยงระหว่างคำศัพท์ในหัวข้อกับคลังคำศัพท์ในเอกสารถูกนำมาใช้ในการคำนวณ Topic Coherence Cv

NPMI ใช้เพื่อวัดความสัมพันธ์ระหว่างคำศัพท์ในเอกสารและคำในหัวข้อ ทำให้สามารถประเมินความเข้าใจและความสัมพันธ์ที่มีความหมายระหว่างคำศัพท์ได้

NPMI สามารถคำนวณสูตรได้ดังนี้ :

$$NPMI(w_1, w_2) = \frac{\log(\frac{P(w_1, w_2)}{P(w_1) P(w_2)})}{-\log(P(w_1, w_2))}$$

โดยที่:

- $w_1$  และ  $w_2$  คือคำศัพท์ที่ต้องการวัดความสัมพันธ์
- $P(w_1, w_2)$  คือความน่าจะเป็นที่คำศัพท์  $w_1$  และ  $w_2$  จะปรากฏพร้อมกัน
- $P(w_1)$  และ  $P(w_2)$  คือความน่าจะเป็นที่คำศัพท์  $w_1$  และ  $w_2$  จะปรากฏคนละครั้ง
- ฐาน log ของส่วนหารในเศษส่วนนั้นจะทำให้ NPMI มีค่าอยู่ระหว่าง -1 ถึง 1

การคำนวณ Topic Coherence Cv คือรูปแบบหนึ่งของ NPMI ในการคำนวณ co-occurrence (จะ

สื่อถึงความถี่ของคำสองคำที่จะปรากฏอยู่ด้วยกันในเอกสารหรือชุดเอกสาร) ของคำในหัวข้อ (Topic) เทียบกับคำทั้งหมดในหัวข้อ (Topic) เดียวกัน โดยจะอธิบายการคำนวณจากตัวอย่างดังต่อไปนี้

สมมติให้ topic  $t_1 = \{roses, red, tulip, blue\}$

สามารถสร้าง vector ของคำว่า roses, red ตามสูตรดังนี้

$$v_{roses,red} = NPMI_{roses,red} = \frac{\log\left(\frac{P(roses,red)}{P(roses) \cdot P(red)}\right)}{-\log(P(roses,red))}$$

ในการคำนวณ Topic Coherence  $C_V$  จะเป็นต้องสร้าง vector ของคำทุกคำภายใน topic  $t_1$  โดยจะสร้าง vector ของคำว่า roses เป็นตัวอย่างตามสูตรดังนี้

$$v_{roses} = \begin{cases} NPMI(roses, roses), NPMI(roses, red), \\ NPMI(roses, tulip), NPMI(roses, blue) \end{cases}$$

จากนั้นจะวัดระยะของ vector โดยใช้ cosine similarity (เป็นการหาความคล้ายคลึงของ vector ด้วยองศา) ก่อนจะเริ่มคำนวณ cosine similarity จะเป็นต้องสร้าง vector ที่เป็นขนาดรวมของ vector ทุกคำภายใน topic  $t_1$  ซึ่งจะแทนด้วย  $v_c$  ตามสูตรดังนี้

$$v_c = v_{roses} + v_{red} + v_{tulip} + v_{blue}$$

สุดท้ายจะคำนวณ Topic Coherence  $C_V$  ของ topic  $t_1$  โดยใช้ cosine similarity ของ vector แต่ละคำใน topic  $t_1$  ซึ่งแสดงตามสูตรดังนี้

$$C_V = \frac{1}{4} \cdot (\cos(v_{roses}, v_c) + \cos(v_{red}, v_c) + \cos(v_{tulip}, v_c) + \cos(v_{blue}, v_c))$$

โดยที่:

$\frac{1}{4}$  มาจาก  $\frac{1}{N}$  โดย  $N$  คือจำนวนคำทั้งหมดใน topic  $t_1$

$\cos(v_{roses}, v_c)$  คือการหา cosine similarity ระหว่าง  $v_{roses}$ ,  $v_c$

จากการคำนวณของ Topic Coherence  $C_V$  หรือ  $C_V$  โดยค่าที่คำนวณออกมาได้ถ้าค่าอยู่สูงหมายถึง ประสิทธิภาพที่ดีมากตาม หมายความว่าคำหลายคำที่ปรากฏภายในหัวข้อนั้นมีความสัมพันธ์อยู่จริงในเอกสารหรือชุดเอกสารซึ่งค่า  $C_V$  โดยทั่วไปจะอยู่ในช่วง  $[0,1]$

## 2.2 งานวิจัยที่เกี่ยวข้อง

### 2.2.1 งานวิจัยของ วรอปสร บุบพา และ เอกรัฐ รัฐกาญจน

[1] ได้ทำการศึกษาเรื่องการวิเคราะห์ความรู้สึกแต่ละแบบของผู้ใช้บริการศูนย์ซื้อขายสินทรัพย์ดิจิทัล ในโซเชียลมีเดีย โดยใช้ข้อมูลความคิดเห็นของผู้ใช้บริการในเพจเฟซบุ๊กศูนย์ซื้อขายสินทรัพย์ดิจิทัลในการหา Topic Modeling เพื่อจำแนกແนียงปัญหาแล้ววิเคราะห์ความรู้สึกตัวแบบสำหรับการระบุความรู้สึกซึ้งตัวแบบ Random Forest และ Logistic Regression ให้ประสิทธิภาพดีต่างกัน โดยภาพรวมและเฉพาะความคิดเห็นในเชิงลบตามลำดับ

### 2.2.2 งานวิจัยของ ณิชาภา รัตนปรีชาเวช

[2] ได้ศึกษาเรื่องความสัมพันธ์ระหว่างประเด็นต่าง ๆ ที่ถูกพูดถึงใน โซเชียลมีเดียกรณีศึกษาเว็บบอร์ดกลุ่มแม่และเด็ก โดยใช้ข้อมูลความคิดเห็นทั้งหมดที่เกิดขึ้นในเว็บบอร์ด Pantip.com ห้อง chan เรื่อง ผลการวิจัยพบว่า ผู้บริโภคกลุ่มแม่และเด็กส่วนใหญ่พูดคุยกันในประเด็นเรื่องค่าใช้จ่ายของโรงเรียน ความกังวลต่อการตั้งครรภ์และค่าใช้จ่ายของโรงพยาบาล ที่ Jaccard Similarity 0.2, 0.19, 0.19 ตามลำดับ อีกทั้งยังมีการแสดงความเครียดและกังวลกับทุกเรื่องที่ระดับ Normalized Degree Centrality สูงสุด 0.648 โดยเฉพาะ สถานการณ์ท้องไม่มีพร้อม

### 2.2.3 งานวิจัยของ ปิยวรรณ ทองพโลย

[3] วิเคราะห์ข้อความภาษาไทยเกี่ยวกับการตั้งครรภ์ด้วยวิธีการสร้างแบบจำลองหัวข้อให้ผลการทดลองคืองานวิจัยนี้เป็นการศึกษาวิธีการวิเคราะห์หากำที่มีนัยสำคัญและหัวข้อในข้อความและทำการจัดกลุ่มข้อความที่มีลักษณะคล้ายคลึงกันอยู่ในกลุ่มเดียวกันด้วยเทคนิคการจัดตีรีเคลฟแห่ง (Latent Dirichlet Allocation) ซึ่งให้เห็นว่าผู้ใช้งานอินเตอร์เน็ตจำนวนมากที่เข้ามาอ่านปัญหาเกี่ยวกับแพทย์ผ่านชุมชนออนไลน์นั้น ส่วนมากมีความกังวลในเรื่องเกี่ยวกับ การผิดปกติในการตั้งครรภ์ ความผิดปกติของประจำเดือนและการคุมกำเนิด โดยงานวิจัยนี้ได้ทำการแบ่งกลุ่มข้อมูลโดยใช้ K-means ในการวัดประสิทธิภาพในการจัดกลุ่มของเทคนิคการจัดสรรคีรีเคลฟแห่ง (Latent Dirichlet Allocation) โดยค่า Silhouette Coefficient มีการจัดกลุ่มข้อมูลที่ 4 กลุ่ม มีค่าอยู่ที่ 0.5397 ซึ่งมีค่าที่สูงและข้อมูลมีการจับกลุ่มได้ดีเมื่อเทียบกับการจัดกลุ่มด้วยจำนวนอื่นๆ

## 2.3 เทคโนโลยีที่เกี่ยวข้อง

### 2.3.1 ภาษาโปรแกรมพythoN (Python Programming Language)

โปรแกรมภาษา Python เป็นภาษาทางคอมพิวเตอร์ชนิดหนึ่งที่ใช้ในการเขียนโปรแกรม โดย Python ถูกออกแบบมาให้เน้นการอ่าน code ได้ง่ายและให้มีความคล้ายคลึงกับภาษาอังกฤษ มีรูปแบบการเขียน Syntax ที่ช่วยให้นักพัฒนาสามารถเขียนโปรแกรมได้ง่ายด้วยจำนวนบรรทัด การเขียนน้อยกว่าภาษาทางคอมพิวเตอร์ภาษาอื่นๆ และทำงานในรูปแบบ object-oriented

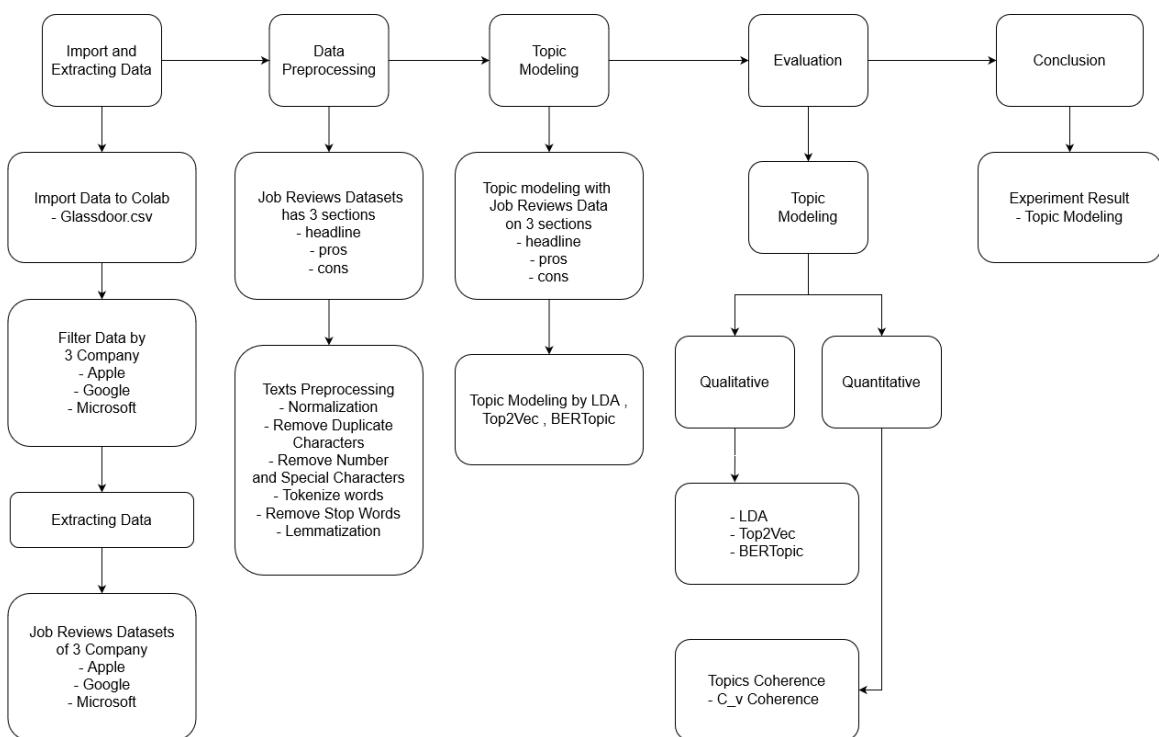
### 2.3.2 เครื่องมือ (Tools)

2.3.2.1 Google Colab (Google Colaboratory) เป็นบริการ Software as a service โดยเป็นโปรแกรม Jupyter Notebook บน Cloud จาก Google ไม่จำเป็นต้องติดตั้งโปรแกรมอื่นๆ

## บทที่ 3 วิธีการวิจัย

บทนี้จะกล่าวถึงภาพรวมของโครงการซึ่งจะอธิบายถึงการศึกษาของแต่ละส่วนอย่างละเอียด

### 3.1 ภาพรวมของโครงการ



ภาพที่ 3.1 แผนผังวิธีการดำเนินงานวิจัย

### 3.2 ขั้นตอนการดำเนินงานวิจัย

ขั้นตอนการดำเนินงานวิจัยแบ่งเป็น 5 ขั้นตอน ได้แก่ การนำเข้าและคัดแยกข้อมูล (Import and Extracting Data), การทำความสะอาดข้อมูล (Data Preprocessing), แบบจำลองหัวข้อ (Topic Modeling), ประเมินประสิทธิภาพ (Evaluation), สรุปผลลัพธ์การทดลอง (Conclusion) โดยมีรายละเอียดดังต่อไปนี้

#### 3.2.1 การนำเข้าและคัดแยกข้อมูล (Import and Extracting Data)

ข้อมูลที่นำมาใช้ในงานวิจัยนี้ จะเป็นข้อมูลเกี่ยวกับรีวิวอาชีพบนเว็บไซต์ Glassdoor.com ซึ่งผู้วิจัยได้ชุดข้อมูลของ Glassdoor มาจากเว็บไซต์ฟรีเดต้าอย่าง Kaggle.com ซึ่งเป็นเว็บไซต์ยอดนิยมสำหรับค้นหาเดต้าเซ็ทต่าง ๆ ได้โดยไม่มีค่าใช้จ่าย โดยชุดข้อมูลเป็นไฟล์สกุล CSV มีข้อมูล 838,566 แถวและมี 18 หลัก ซึ่งข้อมูลที่จะนำมาใช้วิเคราะห์ได้แก่ firm (ชื่อบริษัท), headline (ชื่อรีวิว), pros (รีวิวประโภชน์ที่ได้), cons (รีวิวผลเสียที่ได้)

นำชุดข้อมูล import เข้า Google Colab โดยเรียกใช้ผ่าน drive.mount('/content/drive') ซึ่งผู้จัดใช้ Google drive ในการเก็บชุดข้อมูลแล้วจึงสร้าง path นำไปใช้งานใน Colab ตัวอย่างเช่น path = '/content/drive/MyDrive/glassdoor\_reviews.csv'

```
[ ] path = '/content/drive/MyDrive/403/glassdoor_reviews.csv'
df = pd.read_csv(path)
```

ภาพที่ 3.2 การนำชุดข้อมูล import เข้า Google Colab

Filter data by 3 company - Apple, Google, Microsoft

```
▶ dfapple = df.query("firm == 'Apple'")
dfgoogle = df.query("firm == 'Google'")
dfmicrosoft = df.query("firm == 'Microsoft'")
```

ภาพที่ 3.3 การกรองข้อมูลเฉพาะชื่อบริษัทที่เลือก

Extracting data

```
[ ] apple=dfapple.drop(['firm','date_review','job_title','current','location',
                      'overall_rating','work_life_balance','culture_values',
                      'diversity_inclusion','career_opp','comp_benefits',
                      'senior_mgmt','recommend','ceo_approv','outlook'],axis = 1)

google=dfgoogle.drop(['firm','date_review','job_title','current','location',
                      'overall_rating','work_life_balance','culture_values',
                      'diversity_inclusion','career_opp','comp_benefits',
                      'senior_mgmt','recommend','ceo_approv','outlook'],axis = 1)

microsoft=dfmicrosoft.drop(['firm','date_review','job_title','current','location',
                           'overall_rating','work_life_balance','culture_values',
                           'diversity_inclusion','career_opp','comp_benefits',
                           'senior_mgmt','recommend','ceo_approv','outlook'],axis = 1)
```

ภาพที่ 3.4 ลบรายการคอลัมน์ที่ไม่ได้นำมาวิเคราะห์

Job reviews Datasets of 3 company

```
[ ] applelist = apple.values.tolist()
googlelist = google.values.tolist()
microsoftlist = microsoft.values.tolist()
```

ภาพที่ 3.5 การเปลี่ยนจาก DataFrame เป็น List

### 3.2.2 การเตรียมข้อมูลและทำความสะอาดข้อมูล (Data Preprocessing)

การเตรียมข้อมูลและทำความสะอาดข้อมูลโดยจะทำทุกขั้นตอนกับข้อมูลประเภท headline (ชื่อรีวิว), pros (รีวิวประโยชน์ที่ได้), cons (รีวิวผลเสียที่ได้) ซึ่งทั้ง 3 ประเภทนี้คือ ข้อมูลที่เป็นข้อความรีวิวทั้งหมด

```
def clean_data_lemmatizer(row):
    clean_row = []
    for w in row:
        w = unicode_to_ascii(w)

        w = w.lower() # Lower casing

        w = re.sub(' +', ' ', w).strip(' ') # Remove multiple whitespaces, also leading and trailing whitespaces

        w = re.sub(r'[^W\S]', '', w) # Remove special characters and punctuation
        w = re.sub(r"[-_]", r" ", w) # Remove Numerical data

        w = re.sub("(.)\\1{2,}", "\\\\1", w) # Remove duplicate characters

        words = w.split() # Tokenization

        clean_words = [word for word in words if (word not in stopwords_list) and len(word) > 2]

        clean_words= [lemmatizer.lemmatize(w,get_wordnet_pos(w)) for w in clean_words] # For lemmatization
```

ภาพที่ 3.6 การเตรียมข้อมูลและทำความสะอาดข้อมูล

```
w = w.lower() # Lower casing
```

ภาพที่ 3.7 การปรับรูปแบบของคำให้เหมือนกัน

1) Normalization เป็นขั้นตอนการปรับรูปแบบของคำให้เหมือนกัน เช่น การแปลงให้เป็นตัวอักษรตัวเล็กทั้งหมด

```
w = re.sub("(.)\\1{2,}", "\\\\1", w) # Remove duplicate characters
```

ภาพที่ 3.8 การตรวจสอบข้อมูลตัวอักษรที่ซ้ำกัน

2) Remove Duplicate Characters เป็นการตรวจสอบข้อมูลตัวอักษรที่ซ้ำกันโดย จะลบตัวอักษรที่ซ้ำกันเกิน 3 ตัว ซึ่งใช้ “.” เพื่อรับตัวอักษรของคำแต่ละ 1 ตัว และมี “\\\\1” ในการ อ้างอิงถึงวงลีบว่าจะตรวจสอบแต่ละอักษรไม่ให้มีจำนวนมากกว่าเกิน 2 ขึ้นไป “{2, }” จึงถูกแทนที่ ด้วยตัวอักษรเดิมตัวเดียว

```
[ ] print(re.sub("(.)\\1{2,}", "\\\1", 'meee'))
    print(re.sub("(.)\\1{2,}", "\\\1", 'gooddd'))
```

me  
good

ภาพที่ 3.9 ตัวอย่างการลบตัวอักษรซ้ำ

```
w = re.sub(r'[^w\s]', '', w) # Remove special characters and punctuation
w = re.sub(r"([0-9])", r" ", w) # Remove Numerical data
```

ภาพที่ 3.10 การลบตัวเลขออกจากข้อความและลบอักษรหรือสัญลักษณ์ต่างๆ

```
w = re.sub(r'[^w\s]', '', "i've 33333 dogs")
print(w)
w = re.sub(r"([0-9])", r" ", w)
print(w)

ive 33333 dogs
ive  dogs
```

ภาพที่ 3.11 ตัวอย่างการลบตัวเลขออกจากข้อความและลบอักษรหรือสัญลักษณ์ต่างๆ

3) Remove Number and Special Characters เป็นการลบตัวเลขออกจากข้อความ รวมไปถึงลบอักษรหรือลบสัญลักษณ์ต่างๆที่ไม่เกี่ยวข้องออก

```
words = w.split() # Tokenization
```

ภาพที่ 3.12 กระบวนการตัดคำ

4) Tokenize words เป็นกระบวนการตัดคำหรือแบ่งข้อความออกให้เป็นส่วนที่เล็กลง เช่น แบ่งข้อความออกเป็นประโยค และแบ่งจากประโยคให้เป็นคำก่อนแล้วจึงทำการลบเครื่องหมายวรรคตอนออกทั้งหมด

```
clean_words = [word for word in words if (word not in stopwords_list) and len(word) > 2]
```

ภาพที่ 3.13 การลบ Stop words

5) Remove Stop Words ใน การลบ Stop Words จะเป็นคำที่ว่าไปที่พบบ่อยและไม่ได้สื่อความหมายพิเศษ พร้อมทั้งลบคำที่มีอักษรน้อยกว่า 3 ตัวอักษรออก เช่น a, an, the, is, am, are, in, on, at

```
clean_words= [lemmatizer.lemmatize(w,get_wordnet_pos(w)) for w in clean_words] # For lemmatization
```

ภาพที่ 3.14 กระบวนการ lemmatization

6) lemmatization เป็นกระบวนการที่ใช้คำศัพท์และการวิเคราะห์ทางสัณฐานวิทยา (morphological analysis) แล้วลบคำในจุดสิ้นสุดที่ผันกลับมาเพื่อให้ได้รูปแบบพื้นฐาน (รูปแบบพจนานุกรม) เป็นที่รู้จักกันในชื่อ บทแทรก (lemma) โดยเป็นกระบวนการที่ซับซ้อนต้องการความเข้าใจในบริบทที่คำต่าง ๆ จากการทำ PoS Tagging pragmatically เพื่อตัดสินใจเกี่ยวกับความหมาย ดังนั้นจึงสามารถเข้าใจและใช้คำศัพท์เพื่อให้ได้รูปแบบราก (root form) ใช้เวลานานกว่าการ stemming รวมไปถึงแนวโน้มที่จะให้ผลลัพธ์ที่แม่นยำ

```
[46] print(lemmatizer.lemmatize("fly"))
    print(lemmatizer.lemmatize("flies"))
    print(lemmatizer.lemmatize("flying"))
    print(lemmatizer.lemmatize("flew"))
    print(lemmatizer.lemmatize("flown"))

fly
fly
flying
flew
flown
```

ภาพที่ 3.15 ตัวอย่างการ lemmatization

```
[49] print(lemmatizer.lemmatize("foot"))
    print(lemmatizer.lemmatize("feet"))
    print(lemmatizer.lemmatize("foots"))
    print(lemmatizer.lemmatize('footing'))

foot
foot
foot
footing
```

ภาพที่ 3.16 ตัวอย่างการ lemmatization

### 3.2.3 ตัวแบบหัวข้อ (Topic modeling)

```

applehead = []
for x in range(len(ppcs_apple_data)):
    applehead.append(ppcs_apple_data[x][0])

applepros = []
for x in range(len(ppcs_apple_data)):
    applepros.append(ppcs_apple_data[x][1])

applecons = []
for x in range(len(ppcs_apple_data)):
    applecons.append(ppcs_apple_data[x][2])

googlehead = []
for x in range(len(ppcs_google_data)):
    googlehead.append(ppcs_google_data[x][0])

googlepros = []
for x in range(len(ppcs_google_data)):
    googlepros.append(ppcs_google_data[x][1])

googlecons = []
for x in range(len(ppcs_google_data)):
    googlecons.append(ppcs_google_data[x][2])

microsofthead = []
for x in range(len(ppcs_microsoft_data)):
    microsofthead.append(ppcs_microsoft_data[x][0])

microsoftpros = []
for x in range(len(ppcs_microsoft_data)):
    microsoftpros.append(ppcs_microsoft_data[x][1])

microsoftcons = []
for x in range(len(ppcs_microsoft_data)):
    microsoftcons.append(ppcs_microsoft_data[x][2])

```

ภาพที่ 3.17 คัดแยกข้อมูล

ก่อนการสร้างตัวแบบหัวข้อผู้วิจัยจะทำการ Filter ข้อมูลโดยแบ่งตามบริษัท ซึ่งเลือกมา 3 บริษัทได้แก่ Apple, Google, Microsoft หลังจากนั้นผู้วิจัยจะทำการคัดแยกข้อมูล ออกมานา (Extracting data) จะได้ชุดข้อมูลที่มีข้อมูลเฉพาะจากที่เลือกมา 3 บริษัท

ในกระบวนการนี้ใช้เฉพาะข้อความที่เป็นประเภทข้อความรีวิวการทำงาน (Job review) ซึ่งมีอยู่ 3 ส่วนในชุดข้อมูล Glassdoor นี้คือ headline (หัวข้อรีวิว), pros (รีวิวประโภชณ์ที่ได้), cons (รีวิวผลเสียที่ได้) เพื่อหาหัวข้อหรือประเด็นสำคัญที่พูดถึงในแต่ละส่วน

#### 3.2.3.1 Topic Modeling by Latent Dirichlet Allocation (LDA)

```

# Gensim libraries
import gensim

```

ภาพที่ 3.18 library ที่นำมาใช้ ของ LDA

การสร้างตัวแบบหัวข้อโดยการจัดสรรตีเร่เคลแฟง (Latent Dirichlet Allocation-LDA) โดยการสร้างตัวแบบหัวข้อนั้นได้มีการใช้ Library จาก gensim.models

```
# Build LDA model
lda_model1 = gensim.models.ldamodel.LdaModel(corpus=corpus,
                                              id2word=id2word,
                                              num_topics=10,
```

ภาพที่ 3.19 การกำหนดจำนวนหัวข้อ LDA

ในการสร้างตัวแบบของเทคนิค LDA ในที่นี้ได้กำหนดให้พารามิเตอร์ที่ต้องการ คือ n\_components (จำนวนของหัวข้อ) เท่ากับ 10

### 3.2.3.2 Top2Vec

การสร้างตัวแบบหัวข้อโดย Top2vec ได้มีการข้อมูลส่วนใช้ Library จาก Top2vec ในการสร้างตัวแบบ

```
%pip install -q top2vec[sentence_transformers]
from top2vec import Top2Vec
```

ภาพที่ 3.20 library ที่นำมาใช้ ของ Top2Vec

```
t2v_app_h_16v2_dp = Top2Vec(documents=applehead_1_docs,
                             embedding_model='all-MiniLM-L6-v2',
                             speed="deep-learn", workers=4)
```

ภาพที่ 3.21 การกำหนดสร้างตัวแบบหัวข้อของเทคนิค Top2vec

### 3.2.3.3 BERTopic

การสร้างตัวแบบหัวข้อโดย BERTTopic ได้มีการข้อมูลส่วนใช้ Library จาก BERTTopic ในการสร้างตัวแบบ

```
%pip install -q BERTopic
from bertopic import BERTopic
```

ภาพที่ 3.22 library ที่นำมาใช้ ของ BERTTopic

```

from bertopic import BERTopic

berttopic_model = BERTopic(language="english")

topics, probabilities = berttopic_model.fit_transform(applehead)

```

ภาพที่ 3.23 การกำหนดสร้างตัวแบบหัวข้อของเทคนิค BERTopic

### 3.2.4 ประเมินประสิทธิภาพ (Evaluation)

#### 3.2.4.1 การประเมินประสิทธิภาพเชิงปริมาณ

ค่าความสอดคล้องหัวข้อ (Topic coherence score) คือ วิธีประเมินประสิทธิภาพของตัวแบบหัวข้อโดยให้ผลลัพธ์เป็นค่าความน่าจะเป็น (Probability) ค่ายิ่งเข้าใกล้ 1 แสดงถึงความสอดคล้องของคำในแต่ละผลลัพธ์หัวข้อนั้นๆ และมีความหมายอีกนัยว่ามันนุชร์สามารถตีความหัวข้อเหล่านั้นได้มากน้อยเพียงใดตามค่า score ที่แสดงเป็นเชิง probability

#### 3.2.4.2 การประเมินประสิทธิภาพเชิงคุณภาพ

- 1) เทคนิค LDA ผู้วิจัยจะพิจารณาจากผลลัพธ์ของแบบจำลองหัวข้อและสมมติฐานหัวข้อโดยดูผลลัพธ์ 10 คำที่นำเสนอด้วยกันแล้วว่าอย่างไรได้บ้าง
- 2) เทคนิค Top2vec ผู้วิจัยจะพิจารณาจากคำที่พบในผลลัพธ์หัวข้อซึ่งแสดงผ่าน word cloud และสมมติฐานกลุ่มหัวข้อโดยจับกลุ่มหัวข้อจากการใช้ Hierarchical topic reduction ใน library ของ top2vec
- 3) เทคนิค BERTopic ผู้วิจัยจะพิจารณาจากผลลัพธ์ของแบบจำลองหัวข้อและสมมติฐานกลุ่มหัวข้อโดยจับกลุ่มจากราฟ Hierarchical Clustering

### 3.2.5 การสรุปผลลัพธ์การทดลอง (Conclusion)

สรุปผลการทดลองของตัวแบบหัวข้อ(Topic modeling) ผู้วิจัยจะสรุปจากผลลัพธ์ที่ได้จากการทดลอง ดังนั้นผลลัพธ์ของ Topic modeling นั้นน่าเชื่อถือมากแค่ไหนขึ้นอยู่กับการประเมินผลลัพธ์หัวข้อของผู้วิจัยโดยการวิเคราะห์แบบสรุปอุปนัย (Analytical Induction) และประเมินความเหมาะสมของค่า Coherence score ควบคู่ไปด้วยกันซึ่งอ้างอิงจากงานวิจัยของ [1] รวมถึง บุบพา เพื่อที่จะเข้าใจถึงบริบทของการสนทนาระหว่างหัวข้อ หมายความว่า ผู้วิจัยจะวิเคราะห์ตีความสร้างข้อสรุปข้อมูลจากปรากฏการณ์ที่มองเห็น หมายความว่า หัวข้อนั้นๆ ควรจะสื่อถึงอะไร

## บทที่ 4

### ผลการดำเนินวิจัย

งานวิจัยนี้ได้ทำการวิเคราะห์เชิงความรู้สึกจากการทำแบบจำลองหัวข้อของรีวิวอาชีพในเว็บไซต์ Glassdoor โดยการทำแบบจำลองหัวข้อของรีวิวมีอยู่ด้วยกันสามรูปแบบ คือ การจัดสรรดีรีเคลแฟง (Latent Dirichlet Allocation), Top2Vec และBERTopic พัฒนาด้วยภาษาไพธอน (Python) และ เครื่องมือ Google Colab มีการรวบรวมรีวิว 19,890 รายการ โดยมีผลการวิเคราะห์คำที่มีนัยสำคัญและหัวข้อในข้อความจากแบบจำลอง

1. ผลลัพธ์ของการจำแนกกลุ่มข้อความด้วยเทคนิคการจัดสรรดีรีเคลแฟง (Latent Dirichlet Allocation)
2. ผลลัพธ์ของการจำแนกกลุ่มข้อความด้วย Top2Vec
3. ผลลัพธ์ของการจำแนกกลุ่มข้อความด้วย BERTopic

**ผลลัพธ์ของการจำแนกกลุ่มข้อความด้วยเทคนิคการจัดสรรดีรีเคลแฟง (Latent Dirichlet Allocation)**

1.

ลำดับหัวข้อ	10 อันดับคำที่ปรากฏในหัวข้อ
1	('nice', 0.081231505), ('diverse', 0.06432118), ('fantastic', 0.061784383), ('environment', 0.059033286), ('inclusive', 0.05840986), ('interest', 0.053145494), ('enrich', 0.051216673), ('experienced', 0.051216673), ('office', 0.051209416), ('excellent', 0.035942655)
2	('apple', 0.3184953), ('review', 0.07095938), ('specialist', 0.035098612), ('love', 0.031434417), ('care', 0.027496103), ('depends', 0.025728896), ('technical', 0.024386218), ('home', 0.024350965), ('sale', 0.022070665), ('manager', 0.01966646)
3	('culture', 0.16058522), ('experience', 0.1524269), ('advisor', 0.051936403), ('pretty', 0.03666327), ('cut', 0.029101092), ('tech', 0.020919604), ('edge', 0.018151365), ('toxic', 0.017890636), ('grateful', 0.017294895), ('positive', 0.016624635)
4	('great', 0.47983047), ('company', 0.2965644), ('people', 0.026314445), ('management', 0.01873161), ('growth', 0.0078120884), ('take', 0.0073107244), ('cool', 0.006623101), ('value', 0.006075932), ('little', 0.005026706), ('meet', 0.0049300473)

		('benefit', 0.11808828), ('lot', 0.10501514), ('flexibility', 0.080769576),
5		('career', 0.040933605), ('pay', 0.037979357), ('team', 0.03381804),
		('like', 0.032031942), ('poor', 0.023267446), ('software', 0.020270169),
		('decent', 0.017961571)
6		('work', 0.4585783), ('place', 0.28225288), ('fun', 0.028005918), ('okay', 0.017666161), ('balance', 0.015452149), ('employee', 0.014684082),
		('worklife', 0.007957737), ('grow', 0.0067838375), ('money', 0.0061787786), ('term', 0.0055845226)
7		('awesome', 0.0921886), ('well', 0.05689393), ('expect', 0.043296643), ('development', 0.039446976), ('technology', 0.024259327), ('door', 0.024255909), ('make', 0.021295004), ('advance', 0.018764883), ('soul', 0.016710598), ('world', 0.016221123)
8		('job', 0.34392926), ('time', 0.04021697), ('opportunity', 0.03716692), ('student', 0.032439675), ('high', 0.027937073), ('college', 0.025382353), ('position', 0.024128217), ('part', 0.023039022), ('much', 0.021231828), ('entry', 0.01789921)
9		('amaze', 0.1807173), ('workplace', 0.13030538), ('far', 0.10865526), ('fav', 0.10081764), ('business', 0.018695176), ('reward', 0.016888954), ('youre', 0.0149744665), ('terrible', 0.014901055), ('important', 0.013848791), ('youll', 0.010746247)
10		('good', 0.34892622), ('best', 0.08106721), ('ive', 0.07972521), ('retail', 0.060635682), ('overall', 0.059743028), ('ever', 0.05450702), ('side', 0.03897589), ('view', 0.038239043), ('bad', 0.019829236), ('hard', 0.016606933)

ตารางที่ 4.1 ผลการวิเคราะห์คำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล headline บริษัท Apple

ผลการทำนายของแบบจำลองได้วิเคราะห์คำที่มีนัยสำคัญและหัวข้อ โดยในคอลัมน์ที่ 2 เป็น 10 อันดับของคำที่ปรากฏในหัวข้อในห้วนนั้นๆ ซึ่งแต่ละหัวข้อมีการวิเคราะห์ชุดของคำที่ประกอบในหัวข้อและคำหลักแต่ละคำมีการถ่วงน้ำหนัก (weightage) กำหนดให้กับแต่ละหัวข้อ ซึ่งเมื่อตั้งสมมุติฐาน และวิเคราะห์คำที่มีนัยสำคัญในแต่ละหัวข้อ พบร่วมกัน

หัวข้อที่ 1 ความน่าสนใจบริษัทของผู้รีวิว โดยมีคำที่เกี่ยวข้อง คือ nice, interest, fantastic, inclusive

หัวข้อที่ 2 ตำแหน่งอาชีพที่ผู้รีวิวเขียนหัวข้อรีวิวบริษัท Apple โดยมีคำที่เกี่ยวข้องคือ apple, review, technical, speacialist, sale, manager

หัวข้อที่ 3 วัฒนธรรมและประสบการณ์ โดยมีคำที่เกี่ยวข้องคือ culture, experience

หัวข้อที่ 4 บริษัทที่ดี โดยมีคำที่เกี่ยวข้องคือ great, company, people, management

หัวข้อที่ 5 สิ่งที่ได้รับ โดยมีคำที่เกี่ยวข้องคือ benefit, lot, flexibility, career, pay

หัวข้อที่ 6 ความน่าสนใจบริษัทของผู้รีวิว โดยมีคำที่เกี่ยวข้องคือ work, place, fun, okay, balance

หัวข้อที่ 7 แนะนำที่ทำงานที่ดี โดยมีคำที่เกี่ยวข้องคือ awesome, well, expect, development, technology, door, advance, make, soul, world

หัวข้อที่ 8 การเรียนรู้นักศึกษา โดยมีคำที่เกี่ยวข้องคือ job, time, opportunity, student, high, college, position, part, much, entry

หัวข้อที่ 9 ชื่นชอบบริษัท โดยมีคำที่เกี่ยวข้องคือ amaze, workplace, far, fav, business, reward, youre, terrible, important, youll

หัวข้อที่ 10 โดยรวมของบริษัท โดยมีคำที่เกี่ยวข้องคือ good, best, ive, retail, overall, ever, bad, side, view, hard

### มีกลุ่มหัวข้อความหมายที่คล้ายคลึงกันคือ

กลุ่มสมมติฐาน	หัวข้อที่เกี่ยวข้อง
กลุ่มหัวข้อ แนะนำที่ทำงานที่ดี	ความน่าสนใจบริษัทของผู้รีวิว, บริษัทที่ดี, ชื่นชอบบริษัท, แนะนำที่ทำงานที่ดี
กลุ่มหัวข้อ ประสบการณ์ของทำงานที่บริษัท	ตำแหน่งอาชีพที่ผู้รีวิวเขียนหัวข้อรีวิวบริษัท Apple, การเรียนรู้นักศึกษา, วัฒนธรรม ประสบการณ์
กลุ่มหัวข้อ ภาพรวมของบริษัท	สิ่งที่ได้รับ, โดยรวมของบริษัท

ตารางที่ 4.2 กลุ่มสมมติฐานหัวข้อความหมายที่คล้ายคลึง headline ใน Apple

2.

ลำดับหัวข้อ	10 อันดับคำที่ปรากฏในหัวข้อ
1	('environment', 0.32451585), ('fun', 0.10630905), ('coworkers', 0.09077361), ('decent', 0.04807017), ('lovely', 0.047786348), ('level', 0.01786657), ('strong', 0.015320367), ('location', 0.014302465), ('super', 0.014275585), ('leader', 0.012207556)
2	('get', 0.078425065), ('really', 0.07480343), ('training', 0.06964042), ('manager', 0.05039106), ('person', 0.050107293), ('well', 0.04921321), ('make', 0.045217924), ('compensation', 0.02619375), ('manage', 0.024660414), ('sign', 0.02461278)
3	('lot', 0.11195514), ('opportunity', 0.101053394), ('growth', 0.067416616), ('many', 0.065305464), ('learn', 0.058307707), ('ive', 0.0355976), ('always', 0.024632454), ('new', 0.02415832), ('skill', 0.023033144), ('supportive', 0.021745646)
4	('flexibility', 0.11657566), ('development', 0.09532892), ('focus', 0.06762919), ('personal', 0.060961295), ('right', 0.04600959), ('tuitionfitnesscommute', 0.044439442), ('diversity', 0.028168285), ('role', 0.02113696), ('open', 0.018602878), ('store', 0.017768612)
5	('work', 0.29613915), ('people', 0.14777546), ('apple', 0.07820268), ('culture', 0.06909745), ('balance', 0.03694486), ('life', 0.0361273), ('best', 0.030951599), ('everyone', 0.025328843), ('driven', 0.01883997), ('truly', 0.01745489)
6	('job', 0.19462305), ('time', 0.0824698), ('love', 0.072289646), ('part', 0.037988152), ('option', 0.034451995), ('colleague', 0.022771124), ('atmosphere', 0.022359962), ('bonus', 0.022221366), ('even', 0.020650385), ('diverse', 0.019310907)
7	('nice', 0.11425046), ('perk', 0.07737141), ('like', 0.073789135), ('office', 0.057532948), ('family', 0.054960486), ('feel', 0.052100934), ('view', 0.03721321), ('retail', 0.03404098), ('within', 0.029009972), ('inclusive', 0.019788785)
8	('product', 0.09409616), ('high', 0.08197604), ('salary', 0.053915527), ('management', 0.052313663), ('career', 0.050177865), ('friendly',

0.04595091), ('reimbursement', 0.03825637), ('sale', 0.03617079),  
 ('excellent', 0.035371266), ('customer', 0.03476079)

9 ('good', 0.30321214), ('pay', 0.11738916), ('stock', 0.047842175),  
 ('home', 0.032124806), ('positive', 0.027244871), ('flexible',  
 0.026413877), ('hour', 0.025348362), ('position', 0.023834538),  
 ('expectation', 0.020015208), ('snack', 0.018715106)

10 ('great', 0.2371239), ('benefit', 0.16898996), ('discount', 0.06089609),  
 ('company', 0.060856473), ('care', 0.04769242), ('team', 0.041543104),  
 ('employee', 0.041040685), ('amaze', 0.038188986), ('help',  
 0.026683096), ('experience', 0.0172435)

ตารางที่ 4.3 ผลการวิเคราะห์คำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล รีวิวประโยชน์ที่ได้  
 บริษัท apple

ผลการทำนายของแบบจำลองได้วิเคราะห์คำที่มีนัยสำคัญและหัวข้อ โดยในคอลัมน์ที่ 2 เป็น<sup>10</sup> อันดับของคำที่ปรากฏในหัวข้อในหัวนั้นๆ ซึ่งแต่ละหัวข้อมีการวิเคราะห์ชุดของคำที่ประกอบใน<sup>หัวข้อและคำหลักแต่ละคำมีการถ่วงน้ำหนัก (weightage) กำกับให้กับแต่ละหัวข้อ</sup> ซึ่งเมื่อตั้งสมมุติฐาน และวิเคราะห์คำที่มีนัยสำคัญในแต่ละหัวข้อ พบร่วม

หัวข้อที่ 1 สภาพรอบข้างในการทำงาน โดยมีคำที่เกี่ยวข้อง คือ environment, fun, coworkers, decent, lovely, level, strong, location, super, leader

หัวข้อที่ 2 การได้รับด้านบวก โดยมีคำที่เกี่ยวข้องคือ get, really, training, manager, person, well, make, compensation, manage, sign

หัวข้อที่ 3 พัฒนาการ โดยมีคำที่เกี่ยวข้องคือ lot, opportunity, growth, many, learn, ive, always, new, skill, supportive

หัวข้อที่ 4 หลากหลายวิธี โดยมีคำที่เกี่ยวข้องคือ flexibility, development, focus, personal, right, tuitionfitnesscommute, diversity, role, open, store

หัวข้อที่ 5 work life balance โดยมีคำที่เกี่ยวข้องคือ work, people, apple, culture, balance, life, best, everyone, driven, truly

หัวข้อที่ 6 การทำงาน โดยมีคำที่เกี่ยวข้องคือ job, time, love, part, option, colleague, atmosphere, bonus, even, diverse

หัวข้อที่ 7 สิทธิพิเศษ โดยมีคำที่เกี่ยวข้องคือ nice, perk, like, office, family, feel, view, retail, within, inclusive

หัวข้อที่ 8 โดยรวม โดยมีคำที่เกี่ยวข้องคือ product, high, salary, management, career, friendly, reimbursement, sale, excellent, customer

หัวข้อที่ 9 การเงิน โดยมีคำที่เกี่ยวข้องคือ good, pay, stock, home, positive, flexible, hour, position, expectation, snack

หัวข้อที่ 10 ประโยชน์ โดยมีคำที่เกี่ยวข้องคือ great, benefit, discount, company, care, team, employee, amaze, help, experience

มีกลุ่มหัวข้อความหมายที่คล้ายคลึงกันคือ

กลุ่มสมมติฐาน	หัวข้อที่เกี่ยวข้อง
กลุ่มหัวข้อ วัฒนธรรม	สภาพรอบข้างในการทำงาน
กลุ่มหัวข้อ เงินเดือน	การเงิน
กลุ่มหัวข้อ สวัสดิการ	การได้รับด้านบวก, สิทธิพิเศษ, ประโยชน์
กลุ่มหัวข้อ Worklifebalance	Work Life Balance, การทำงาน, โดยรวม
กลุ่มหัวข้อ การเรียนรู้	พัฒนาการ, หลากหลายวิธี

ตารางที่ 4.4 กลุ่มสมมติฐานหัวข้อความหมายที่คล้ายคลึง ประโยชน์ที่ได้ ใน Apple

3.

ลำดับหัวข้อ	10 อันดับคำที่ปรากฏในหัวข้อ
1	('long', 0.09603633), ('none', 0.0757127), ('day', 0.05971902), ('experience', 0.047701515), ('internally', 0.040811457), ('promote', 0.03008383), ('raise', 0.025399903), ('demand', 0.02314532), ('enough', 0.022030566), ('back', 0.01925868)
2	('work', 0.20670068), ('balance', 0.06270123), ('hour', 0.06254865), ('retail', 0.05943438), ('life', 0.051611807), ('weekend', 0.033937), ('even', 0.033842407), ('always', 0.023812477), ('one', 0.019148499), ('general', 0.018068098)
3	('customer', 0.09022938), ('lack', 0.075140454), ('would', 0.044138364), ('progression', 0.032763507), ('leadership', 0.027328936), ('everything', 0.0268542), ('politics', 0.024685884), ('didnt', 0.02141142), ('office', 0.015475036), ('diversity', 0.01445073)
4	('sometimes', 0.045863558), ('people', 0.042469643), ('little', 0.041782167), ('pay', 0.0388197), ('worklife', 0.034435585), ('culture', 0.031699233), ('high', 0.03003012), ('change', 0.026188826), ('many', 0.024214676), ('position', 0.023843765)

		('management', 0.099658884), ('time', 0.09642656), ('difficult', 0.04392553), ('well', 0.040163834), ('opportunity', 0.029605523), ('good', 0.025395578), ('job', 0.023440227), ('keep', 0.022027101), ('grow', 0.020987382), ('low', 0.018373787)
5		('apple', 0.07002246), ('con', 0.053475518), ('employee', 0.039417483), ('lot', 0.037374128), ('manager', 0.034454856), ('nothing', 0.028403852), ('poor', 0.027757077), ('youre', 0.026453847), ('think', 0.021541547), ('unless', 0.02042114)
6		('get', 0.070307806), ('hard', 0.05389529), ('company', 0.041100085), ('move', 0.033829313), ('much', 0.032738566), ('like', 0.024668746), ('role', 0.0230941), ('bad', 0.022855828), ('team', 0.018909141), ('dont', 0.0182741)
7		('busy', 0.14230382), ('process', 0.0635374), ('know', 0.042801756), ('full', 0.03243569), ('mean', 0.03232022), ('living', 0.02870112), ('stress', 0.028478611), ('cost', 0.024410646), ('technical', 0.01795003), ('value', 0.017015146)
8		('call', 0.06387678), ('environment', 0.054884475), ('bit', 0.04879768), ('great', 0.04593717), ('stressful', 0.04049729), ('customer', 0.036279064), ('week', 0.022174494), ('extremely', 0.018775221), ('competitive', 0.0183367), ('base', 0.015648626)
9		('schedule', 0.1856374), ('super', 0.06917135), ('less', 0.04109451), ('try', 0.03184149), ('flexible', 0.02642148), ('free', 0.025914162), ('availability', 0.023797186), ('food', 0.022962283), ('strict', 0.013685469), ('liberal', 0.010380007)
10		

ตารางที่ 4.5 ผลการวิเคราะห์คำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล รีวิวผลเสียที่ได้ บริษัท apple

ผลการทำนายของแบบจำลองได้วิเคราะห์คำที่มีนัยสำคัญและหัวข้อ โดยในคอลัมน์ที่ 2 เป็น 10 อันดับของคำที่ปรากฏในหัวข้อในหัวนั้นๆ ซึ่งแต่ละหัวข้อมีการวิเคราะห์ชุดของคำที่ประกอบในหัวข้อและคำหลักแต่ละคำมีการถ่วงน้ำหนัก (weightage) กำกับให้กับแต่ละหัวข้อ ซึ่งเมื่อตั้งสมมุติฐาน และวิเคราะห์คำที่มีนัยสำคัญในแต่ละหัวข้อ พบร่วม

หัวข้อที่ 1 ประสบการณ์ โดยมีคำที่เกี่ยวข้อง คือ lone, none, day, experience, internally, promote, raise, demand, enough, back

หัวข้อที่ 2 Work Life Balance โดยมีคำที่เกี่ยวข้องคือ work, balance, hour, retail, life, weekend, even, always, one, general

หัวข้อที่ 3 ความคืบหน้า โดยมีคำที่เกี่ยวข้องคือ customer, lack, would, progression, leadership, everything, politics, didnt, office, diversity

หัวข้อที่ 4 จ่ายเงินเดือนช้า โดยมีคำที่เกี่ยวข้องคือ sometimes, people, little, pay, worklife, culture, high, change, many, position

หัวข้อที่ 5 บริหารงาน โดยมีคำที่เกี่ยวข้องคือ management, time, difficult, well, opportunity, good, job, keep, grow, low

หัวข้อที่ 6 ไม่มีข้อเสีย โดยมีคำที่เกี่ยวข้องคือ apple, con, employee, lot, manager, nothing, poor, youre, think, unless

หัวข้อที่ 7 องค์กร โดยมีคำที่เกี่ยวข้องคือ get, hard, company, move, much, like, role, bad, team, dont

หัวข้อที่ 8 วุ่นวาย โดยมีคำที่เกี่ยวข้องคือ busy, process, know, full, mean, living, stress, cost, technical, value

หัวข้อที่ 9 มีตึงตลอดเวลา โดยมีคำที่เกี่ยวข้องคือ call, environment, bit, great, stressful, customer, week, extremely, competitive, base

หัวข้อที่ 10 ตารางเวลาไม่ยืดหยุ่น โดยมีคำที่เกี่ยวข้องคือ schedule, super, less, try, flexible, free, availability, food, strict, liberal

มีกลุ่มหัวข้อความหมายที่คล้ายคลึงกันคือ

กลุ่มสมมติฐาน	หัวข้อที่เกี่ยวข้อง
กลุ่มหัวข้อ องค์กร	ความคืบหน้า, องค์กร, บริหารงาน
กลุ่มหัวข้อ worklifebalance	ประสบการณ์, Work Life Balance
กลุ่มหัวข้อ เงินเดือน	จ่ายเงินเดือนช้า
กลุ่มหัวข้อ การจัดการเวลา	วุ่นวาย, มีตึงตลอดเวลา, ตารางเวลาไม่ยืดหยุ่น

ตารางที่ 4.6 กลุ่มสมมติฐานหัวข้อความหมายที่คล้ายคลึง ผลเสียที่ได้ใน Apple

4.

ลำดับหัวข้อ	10 อันดับคำที่ปรากฏในหัวข้อ
1	('company', 0.41025725), ('best', 0.15709655), ('still', 0.02861966), ('big', 0.025394589), ('ever', 0.02121909), ('top', 0.0154693285), ('fantastic', 0.015354643), ('interest', 0.01392948), ('startup', 0.01195821), ('challenge', 0.0116224745)
2	('good', 0.5937144), ('wlb', 0.040404372), ('bad', 0.028068949), ('high', 0.026158437), ('job', 0.025058389), ('far', 0.017202085), ('move', 0.016203033), ('view', 0.015390453), ('sale', 0.015045497), ('mountain', 0.014568276)
3	('marketing', 0.06327808), ('specialist', 0.057418782), ('huge', 0.036598407), ('performance', 0.034929384), ('feel', 0.034613553), ('like', 0.03204099), ('matter', 0.02974718), ('bureaucracy', 0.022522233), ('would', 0.02198459), ('process', 0.021192435)
4	('opportunity', 0.08199396), ('environment', 0.06698865), ('learn', 0.05859761), ('manager', 0.052609388), ('worklife', 0.041984014), ('large', 0.029972881), ('hard', 0.026782913), ('management', 0.026354678), ('technology', 0.021812737), ('assistant', 0.021352693)
5	('great', 0.38994265), ('culture', 0.10080822), ('people', 0.061495133), ('pay', 0.034289204), ('career', 0.03245656), ('benefit', 0.02827007), ('colleague', 0.025316242), ('low', 0.015320435), ('demand', 0.012058133), ('atmosphere', 0.011846879)
6	('google', 0.33814904), ('job', 0.07120074), ('love', 0.047516197), ('engineer', 0.04428056), ('software', 0.043343667), ('salary', 0.03864928), ('developer', 0.024499048), ('lottery', 0.021117646), ('satisfaction', 0.015767625), ('cool', 0.014794554)
7	('work', 0.29462108), ('place', 0.20287488), ('great', 0.17746232), ('nice', 0.040555526), ('balance', 0.025760453), ('life', 0.020553732), ('overall', 0.017091077), ('business', 0.010989836), ('fun', 0.0104539655), ('position', 0.009414041)
8	('review', 0.14652254), ('perk', 0.06832597), ('world', 0.06764116), ('one', 0.049749143), ('class', 0.049697816), ('always', 0.0261313),

		('talent', 0.025846463), ('talented', 0.021864561), ('upward', 0.020018945), ('trajectory', 0.018265164)
9		('amaze', 0.15720831), ('experience', 0.14756711), ('excellent', 0.124703564), ('grow', 0.07366033), ('awesome', 0.07037469), ('employer', 0.05508588), ('slow', 0.018435687), ('positive', 0.013502444), ('little', 0.010756162), ('simply', 0.009746744)
10		('well', 0.060608517), ('year', 0.045730975), ('employee', 0.04054204), ('school', 0.03796076), ('district', 0.032092527), ('solid', 0.028194478), ('product', 0.024208369), ('office', 0.02315762), ('quite', 0.019065933), ('operation', 0.016899707)

ตารางที่ 4.7 ผลการวิเคราะห์คำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล headline บริษัท Google

ผลการทำนายของแบบจำลองได้วิเคราะห์คำที่มีนัยสำคัญและหัวข้อ โดยในคอลัมน์ที่ 2 เป็น 10 อันดับของคำที่ปรากฏในหัวข้อในหัวนี้ๆ ซึ่งแต่ละหัวข้อมีการวิเคราะห์ชุดของคำที่ประกอบในหัวข้อและคำหลักแต่ละคำมีการถ่วงน้ำหนัก (weightage) กำกับให้กับแต่ละหัวข้อ ซึ่งเมื่อตั้งสมมุติฐาน และวิเคราะห์คำที่มีนัยสำคัญในแต่ละหัวข้อ พบร่วม

หัวข้อที่ 1 บริษัทที่ดี โดยมีคำที่เกี่ยวข้อง คือ company, best, still, big, ever, top, fantastic, interest, startup, challenge

หัวข้อที่ 2 ชีวิตกับงาน โดยมีคำที่เกี่ยวข้องคือ good, well, bad, high, job, far, move, view, sale, mountain

หัวข้อที่ 3 ตำแหน่ง โดยมีคำที่เกี่ยวข้องคือ marketing, specialist, huge, performance, feel, like, matter, bureaucracy, would, process

หัวข้อที่ 4 โอกาสพัฒนา โดยมีคำที่เกี่ยวข้องคือ opportunity, environment, learn, manager, worklife, large, hard, management, technology, assistant

หัวข้อที่ 5 วัฒนธรรม โดยมีคำที่เกี่ยวข้องคือ great, culture, people, pay, career, benefit, colleague, low, demand, atmosphere

หัวข้อที่ 6 งานภายใน Google โดยมีคำที่เกี่ยวข้องคือ google, job, love, engineer, software, salary, developer, lottery, satisfaction, cool

หัวข้อที่ 7 work life balance โดยมีคำที่เกี่ยวข้องคือ work, place, great, nice, balance, life, overall, business, fun, position

หัวข้อที่ 8 สิทธิพิเศษ โดยมีคำที่เกี่ยวข้องคือ review, perk, world, one, class, always, talent, talented, upward, trajectory

หัวข้อที่ 9 ประสบการณ์ โดยมีคำที่เกี่ยวข้องคือ amaze, experience, excellent, grow, awesome, employer, slow, positive, little, simply

หัวข้อที่ 10 บรรยายกาศ โดยมีคำที่เกี่ยวข้องคือ well, year, employee, school, district, solid, product, office, quite, operation

มีกลุ่มหัวข้อความหมายที่คล้ายคลึงกันคือ

กลุ่มสมมติฐาน	หัวข้อที่เกี่ยวข้อง
กลุ่มหัวข้อ แนะนำว่าเป็นที่ทำงานที่ดี	บริษัทที่ดี, Work Life Balance
กลุ่มหัวข้อ ประสบการณ์ของทำงานที่บริษัท	ตำแหน่ง, พัฒนา, ประสบการณ์, งานภายใน Google
กลุ่มหัวข้อ ภาพรวมของบริษัท	ชีวิตกับงาน, วัฒนธรรม, สิทธิพิเศษ, บรรยายกาศ

ตารางที่ 4.8 กลุ่มสมมติฐานหัวข้อความหมายที่คล้ายคลึง headline ใน Google

5.

ลำดับหัวข้อ	10 อันดับคำที่ปรากฏในหัวข้อ
1	('compensation', 0.115879856), ('excellent', 0.085575335), ('everything', 0.08041645), ('technology', 0.044902336), ('freedom', 0.026205571), ('bad', 0.025977004), ('billion', 0.022232873), ('design', 0.019192936), ('scope', 0.018803634), ('base', 0.01761841)
2	('well', 0.058910344), ('like', 0.05124543), ('flexible', 0.045571845), ('know', 0.034121387), ('ever', 0.03393055), ('supportive', 0.033287775), ('internal', 0.032758683), ('resource', 0.025759697), ('around', 0.024386393), ('make', 0.0234999)
3	('lot', 0.12216744), ('smart', 0.10392801), ('opportunity', 0.09369037), ('learn', 0.06812028), ('team', 0.054023143), ('interest', 0.041447107), ('growth', 0.036279436), ('impact', 0.03572008), ('product', 0.034781724), ('manager', 0.021802733)
4	('employee', 0.12832303), ('care', 0.064768024), ('day', 0.034875944), ('comp', 0.0345448), ('top', 0.03118008), ('super', 0.030272888), ('take', 0.023627946), ('market', 0.02311819), ('everyone', 0.023067657), ('perfect', 0.020415792)

('fun', 0.058582366), ('awesome', 0.05467439), ('experience',  
 5 ('experience', 0.04515904), ('thing', 0.042676236), ('wonderful', 0.042029053),  
 ('leader', 0.03371702), ('professional', 0.032748193), ('incredible',  
 6 ('incredible', 0.030746238), ('strong', 0.029841416), ('development', 0.022561349)  
 ('best', 0.11019115), ('google', 0.10610611), ('time', 0.04375616),  
 ('love', 0.040860508), ('challenge', 0.033805236), ('competitive',  
 7 ('competitive', 0.027891226), ('skill', 0.022781473), ('mobility', 0.01947215),  
 ('provide', 0.01943441), ('also', 0.018305728)  
 ('great', 0.18832788), ('people', 0.10580211), ('benefit', 0.08410904),  
 ('culture', 0.07091218), ('company', 0.06028098), ('environment',  
 8 ('environment', 0.05339887), ('perk', 0.0474506), ('amaze', 0.043180097), ('nice',  
 ('nice', 0.040538497), ('work', 0.024774395)  
 ('project', 0.07590246), ('coworkers', 0.07556684), ('get', 0.072618045),  
 ('many', 0.068657644), ('scale', 0.03561084), ('cool', 0.032961804),  
 ('support', 0.022758823), ('feel', 0.02268499), ('role', 0.02187086),  
 ('health', 0.018025449)  
 ('salary', 0.16700095), ('food', 0.13240647), ('free', 0.0984345), ('high',  
 9 ('high', 0.07345626), ('money', 0.03730468), ('new', 0.03179467), ('etc',  
 ('etc', 0.028517459), ('snack', 0.022112839), ('start', 0.017555924), ('name',  
 ('name', 0.015663493)  
 ('work', 0.2798341), ('good', 0.2579833), ('balance', 0.06732777), ('life',  
 10 ('life', 0.06274495), ('place', 0.056754593), ('pay', 0.048536558),  
 ('management', 0.016609427), ('pro', 0.014776091), ('flexibility',  
 ('flexibility', 0.0125407195), ('worklife', 0.012144486)

ตารางที่ 4.9 ผลการวิเคราะห์คำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล  
รีวิวผลประโยชน์ที่ได้ บริษัท Google

ผลการทำนายของแบบจำลองได้วิเคราะห์คำที่มีนัยสำคัญและหัวข้อ โดยในคอลัมน์ที่ 2 เป็น 10 อันดับของคำที่ปรากฏในหัวข้อในหัวนั้นๆ ซึ่งแต่ละหัวข้อมีการวิเคราะห์ชุดของคำที่ประกอบในหัวข้อและคำหลักแต่ละคำมีการถ่วงน้ำหนัก (weightage) กำกับให้กับแต่ละหัวข้อ ซึ่งเมื่อตั้งสมมุติฐาน และวิเคราะห์คำที่มีนัยสำคัญในแต่ละหัวข้อ พบร่วม

หัวข้อที่ 1 ภาพรวมของบริษัท โดยมีคำที่เกี่ยวข้อง คือ compensation, excellent, everything, technology, freedom, bad, billion, design, scope, base

หัวข้อที่ 2 สวัสดิการ โดยมีคำที่เกี่ยวข้องคือ well, like, flexible, know, ever, supportive, internal, resource, around, make

หัวข้อที่ 3 โอกาสการเรียนรู้ โดยมีคำที่เกี่ยวข้องคือ lot, smart, opportunity, learn, team, interest, growth, impact, product, manager

หัวข้อที่ 4 พนักงาน โดยมีคำที่เกี่ยวข้องคือ employee, care, day, comp, top, super, take, market, everyone, perfect

หัวข้อที่ 5 สนุกสนาน โดยมีคำที่เกี่ยวข้องคือ fun, awesome, experience, thing, wonderful, leader, professional, incredible, strong, development

หัวข้อที่ 6 ดีที่สุด โดยมีคำที่เกี่ยวข้องคือ best, google, time, love, challenge, competitive, skill, mobility, provide, also

หัวข้อที่ 7 สังคม โดยมีคำที่เกี่ยวข้องคือ great, people, benefit, culture, company, environment, perk, amaze, nice, work

หัวข้อที่ 8 โครงงาน โดยมีคำที่เกี่ยวข้องคือ project, coworkers, get, many, scale, cool, support, feel, role, health

หัวข้อที่ 9 เงินเดือน โดยมีคำที่เกี่ยวข้องคือ salary, food, free, high, money, new, etc, snack, start, name

หัวข้อที่ 10 worklifebalance โดยมีคำที่เกี่ยวข้องคือ work, good, balance, life, place, pay, management, pro, flexibility, worklife

#### มีกลุ่มหัวข้อความหมายที่คล้ายคลึงกันคือ

กลุ่มสมมติฐาน	หัวข้อที่เกี่ยวข้อง
กลุ่มหัวข้อ วัฒนธรรม	พนักงาน, สังคม
กลุ่มหัวข้อ เงินเดือน	เงินเดือน
กลุ่มหัวข้อ สวัสดิการ	สวัสดิการ
กลุ่มหัวข้อ Worklifebalance	ดีที่สุด, โครงงาน, worklifebalance
กลุ่มหัวข้อ การเรียนรู้	โอกาสการเรียนรู้, สนุกสนาน

ตารางที่ 4.10 กลุ่มสมมติฐานหัวข้อความหมายที่คล้ายคลึง ประยุกต์ได้ ใน Google

6.

ลำดับหัวข้อ	10 อันดับคำที่ปรากฏในหัวข้อ
1	('much', 0.11883626), ('promotion', 0.06603627), ('leadership', 0.039663393), ('challenge', 0.033155493), ('pro', 0.03147371), ('competitive', 0.027668666), ('red', 0.016492583), ('tape', 0.016492583), ('help', 0.013830043), ('frustrate', 0.012410829)
2	('project', 0.06841864), ('process', 0.055808958), ('bit', 0.038016155), ('performance', 0.033507816), ('focus', 0.02807748), ('difficult', 0.027822174), ('internal', 0.026163409), ('competition', 0.026001472), ('tech', 0.023045883), ('review', 0.02239658)
3	('get', 0.084664345), ('people', 0.06125273), ('bad', 0.043886665), ('like', 0.039799713), ('thing', 0.032064658), ('promote', 0.030351134), ('feel', 0.028497523), ('politics', 0.026532829), ('say', 0.02505275), ('environment', 0.021641945)
4	('management', 0.06034896), ('growth', 0.042468287), ('make', 0.035813563), ('team', 0.034727704), ('culture', 0.03273164), ('great', 0.031695545), ('need', 0.023895111), ('product', 0.021798242), ('lead', 0.0163848), ('even', 0.014809263)
5	('lot', 0.114770606), ('slow', 0.09911096), ('time', 0.08340311), ('career', 0.07637522), ('lack', 0.04190052), ('life', 0.0374972), ('bureaucracy', 0.03551027), ('move', 0.025691943), ('stress', 0.021480277), ('progression', 0.020747928)
6	('nothing', 0.09781243), ('job', 0.040446274), ('really', 0.039028786), ('come', 0.031121675), ('well', 0.03106159), ('office', 0.024910117), ('pay', 0.023335624), ('role', 0.022839425), ('one', 0.021004876), ('best', 0.019048484)
7	('high', 0.08018884), ('manager', 0.046702996), ('experience', 0.04346698), ('less', 0.034024395), ('little', 0.031772774), ('know', 0.025950143), ('learn', 0.022667442), ('far', 0.02199153), ('everyone', 0.018176189), ('area', 0.015143666)
8	('work', 0.2558811), ('hard', 0.0814586), ('sometimes', 0.053400498), ('hour', 0.05029933), ('long', 0.047327366), ('good', 0.043725424),

		('balance', 0.03060692), ('pressure', 0.02676331), ('everything', 0.02549554), ('worklife', 0.013089857)
9		('none', 0.1093949), ('large', 0.080056086), ('think', 0.0645943), ('impact', 0.03421421), ('level', 0.032761384), ('year', 0.028542195), ('limited', 0.025857873), ('mind', 0.023857584), ('pace', 0.022031022), ('every', 0.017533198)
10		('company', 0.174489), ('con', 0.13567801), ('google', 0.0761162), ('big', 0.074491665), ('dont', 0.03756826), ('many', 0.032490384), ('become', 0.022895804), ('senior', 0.020898344), ('political', 0.018344264), ('see', 0.01595975)

ตารางที่ 4.11 ผลการวิเคราะห์คำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล รีวิวผลเสียที่ได้  
บริษัท Googlee

ผลการทำนายของแบบจำลองได้วิเคราะห์คำที่มีนัยสำคัญและหัวข้อ โดยในคอลัมน์ที่ 2 เป็น 10 อันดับของคำที่ปรากฏในหัวข้อในหัวนี้ฯ ซึ่งแต่ละหัวข้อมีการวิเคราะห์ชุดของคำที่ประกอบในหัวข้อและคำหลักแต่ละคำมีการถ่วงน้ำหนัก (weightage) กำกับให้กับแต่ละหัวข้อ ซึ่งเมื่อตั้งสมมุติฐาน และวิเคราะห์คำที่มีนัยสำคัญในแต่ละหัวข้อ พบร่วม

หัวข้อที่ 1 การเดือนขึ้น โดยมีคำที่เกี่ยวข้อง คือ much, promotion, leadership, challenge, pro, competitive, red, tape, help, frustrate

หัวข้อที่ 2 โครงการยก โดยมีคำที่เกี่ยวข้องคือ project, process, bit, performance, focus, difficult, internal, competition, tech, review

หัวข้อที่ 3 ผู้คน โดยมีคำที่เกี่ยวข้องคือ get, people, bad, like, thing, promote, feel, politics, say, environment

หัวข้อที่ 4 บริหารคน โดยมีคำที่เกี่ยวข้องคือ management, growth, make, team, culture, great, need, product, lead, even

หัวข้อที่ 5 พนักงาน โดยมีคำที่เกี่ยวข้องคือ lot, slow, time, career, lack, life, bureaucracy, move, stress, progression

หัวข้อที่ 6 คิดว่าไม่มีผลเสีย โดยมีคำที่เกี่ยวข้องคือ nothing, job, really, come, well, office, pay, role, one, best

หัวข้อที่ 7 มีการจัดการ โดยมีคำที่เกี่ยวข้องคือ high, manager, experience, less, little, know, learn, far, everyone, area

หัวข้อที่ 8 งานหนัก โดยมีคำที่เกี่ยวข้องคือ work, hard, sometimes, hour, long, good, balance, pressure, everything, worklife

หัวข้อที่ 9 คิดว่าไม่มีผลเสีย โดยมีคำที่เกี่ยวข้องคือ none, large, think, impact, level, year, limited, mind, pace, every

หัวข้อที่ 10 บริษัท โดยมีคำที่เกี่ยวข้องคือ company, con, google, big, dont, many, become, senior, political, see

มีกลุ่มหัวข้อความหมายที่คล้ายคลึงกันคือ

กลุ่มสมนติฐาน	หัวข้อที่เกี่ยวข้อง
กลุ่มหัวข้อ องค์กร	โครงงาน, ผู้คน, พนักงาน, บริษัท, การเลื่อนขั้น, บริหารคน
กลุ่มหัวข้อ worklifebalance	งานหนัก, โครงงานยาก
กลุ่มหัวข้อ การจัดการเวลา	มีการจัดการ

ตารางที่ 4.12 กลุ่มสมนติฐานหัวข้อความหมายที่คล้ายคลึง ผลเสียที่ได้ใน Google

## 7.

ลำดับหัวข้อ	10 อันดับคำที่ปรากฏในหัวข้อ
1	('great', 0.3843641), ('company', 0.3322897), ('culture', 0.066611096), ('big', 0.011130648), ('tech', 0.010820952), ('get', 0.010582059), ('nothing', 0.009734177), ('product', 0.009204391), ('dont', 0.008766993), ('new', 0.0070804837)
2	('career', 0.16707672), ('start', 0.051737286), ('term', 0.0342029), (('manger', 0.03016672), ('early', 0.028677942), ('honestly', 0.028028952), ('leadership', 0.027304169), ('toss', 0.027015071), (('pandemic', 0.025053233), ('fast', 0.02342194))
3	('best', 0.23693906), ('awesome', 0.096228175), ('team', 0.087473), (('lot', 0.047300015), ('love', 0.042878695), ('depends', 0.034720004), (('people', 0.034046765), ('workplace', 0.02359238), ('pretty', 0.01819103), ('satisfaction', 0.017577887))
4	('work', 0.38895506), ('place', 0.23408256), ('balance', 0.06783618), (('learn', 0.05028052), ('benefit', 0.04093736), ('nice', 0.028465673), (('worklife', 0.024192153), ('well', 0.010730009), ('role', 0.008252), (('healthy', 0.0070746643))

('amaze', 0.088452354), ('opportunity', 0.074564256), ('review',  
 5 0.07177979), ('engineer', 0.057910386), ('employee', 0.050744943),  
 ('software', 0.037358955), ('grow', 0.037328564), ('project',  
 6 0.03368176), ('support', 0.031673416), ('cool', 0.021574847)  
 ('job', 0.12392486), ('stay', 0.050624993), ('ever', 0.050301146), ('away',  
 6 0.045277767), ('director', 0.039680853), ('business', 0.03905783),  
 ('limited', 0.039041102), ('please', 0.031334206), ('comfortable',  
 7 0.026345363), ('india', 0.014868906)  
 ('good', 0.5416413), ('growth', 0.053464185), ('experience',  
 7 0.040608875), ('stable', 0.029213646), ('positive', 0.016292304),  
 ('politics', 0.01627541), ('fantastic', 0.015277369), ('option',  
 8 0.011114875), ('perk', 0.010966837), ('interest', 0.010428503)  
 ('excellent', 0.13040769), ('environment', 0.10918209), ('manager',  
 8 0.09521654), ('one', 0.07342109), ('overall', 0.06689846), ('technical',  
 0.024608122), ('cloud', 0.023483207), ('still', 0.023301188), ('decent',  
 9 0.021851035), ('yes', 0.017813748)  
 ('microsoft', 0.17806716), ('life', 0.10189498), ('bad', 0.079163626),  
 ('salary', 0.061467078), ('employer', 0.04268622), ('pay', 0.041405823),  
 9 ('wlb', 0.03471366), ('low', 0.033035696), ('azure', 0.024354413),  
 ('slow', 0.019370856)  
 ('value', 0.08107382), ('fair', 0.04770416), ('colleague', 0.037102323),  
 10 ('day', 0.025427176), ('intentional', 0.015769562), ('every',  
 0.009394938), ('deal', 0.009270136), ('confuse', 0.009209552),  
 ('strives', 0.008159839), ('deliver', 0.008159839)

ตารางที่ 4.13 ผลการวิเคราะห์คำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล headline บริษัท Microsoft

ผลการทำนายของแบบจำลองได้วิเคราะห์คำที่มีนัยสำคัญและหัวข้อ โดยในคอลัมน์ที่ 2 เป็น 10 อันดับของคำที่ปรากฏในหัวข้อในหัวนั้นๆ ซึ่งแต่ละหัวข้อมีการวิเคราะห์ชุดของคำที่ประกอบในหัวข้อและคำหลักแต่ละคำมีการถ่วงน้ำหนัก (weightage) กำกับให้กับแต่ละหัวข้อ ซึ่งเมื่อตั้งสมมุติฐาน และวิเคราะห์คำที่มีนัยสำคัญในแต่ละหัวข้อ พบร่วมกัน

หัวข้อที่ 1 บริษัทใหญ่ โดยมีคำที่เกี่ยวข้อง คือ great, company, culture, big, tech, get, nothing, product, dont, new

หัวข้อที่ 2 เส้นสายอาชีพ Apple โดยมีคำที่เกี่ยวข้องคือ career, start, term, manager, early, honestly, leadership, toss, pandemic, fast

หัวข้อที่ 3 เพื่อนร่วมงานดี โดยมีคำที่เกี่ยวข้องคือ best, awesome, team, lot, love, depends, people, workplace, pretty, satisfaction

หัวข้อที่ 4 สวัสดิการ โดยมีคำที่เกี่ยวข้องคือ work, place, balance, learn, benefit, nice, worklife, well, role, healthy

หัวข้อที่ 5 โอกาสที่ดี โดยมีคำที่เกี่ยวข้องคือ amaze, opportunity, review, engineer, employee, software, grow, project, support, cool

หัวข้อที่ 6 งาน โดยมีคำที่เกี่ยวข้องคือ job, stay, ever, away, director, business, limited, please, comfortable, india

หัวข้อที่ 7 ประสบการณ์ โดยมีคำที่เกี่ยวข้องคือ good, growth, experience, stable, positive, politics, fantastic, option, perk, interest

หัวข้อที่ 8 สภาพแวดล้อม โดยมีคำที่เกี่ยวข้องคือ excellent, environment, manager, one, overall, technical, cloud, still, decent, yes

หัวข้อที่ 9 ภาพรวมบริษัท โดยมีคำที่เกี่ยวข้องคือ Microsoft, life, bad, salary, employer, pay, wlb, low, azure, slow

หัวข้อที่ 10 คุณค่า โดยมีคำที่เกี่ยวข้องคือ value, fair, colleague, day, intentional, every, deal, confuse, strives, deliver

#### มีกลุ่มหัวข้อความหมายที่คล้ายคลึงกันคือ

กลุ่มสมมติฐาน	หัวข้อที่เกี่ยวข้อง
กลุ่มหัวข้อ แนะนำว่าเป็นที่ทำงานที่ดี	บริษัทใหญ่, เพื่อนร่วมงานดี, สวัสดิการ, คุณค่า
กลุ่มหัวข้อ ประสบการณ์ของทำงานที่บริษัท	โอกาสที่ดี, ประสบการณ์
กลุ่มหัวข้อ ภาพรวมของบริษัท	เส้นสายอาชีพ Apple, งาน, สภาพแวดล้อม, ภาพรวมบริษัท

ตารางที่ 4.14 กลุ่มสมมติฐานหัวข้อความหมายที่คล้ายคลึง headline ใน Microsoft

8.

ลำดับหัวข้อ	10 อันดับคำที่ปรากฏในหัวข้อ
1	('amaze', 0.12265104), ('brand', 0.056031276), ('tech', 0.053601425), ('like', 0.040108778), ('business', 0.039807044), ('big', 0.02904672), ('etc', 0.024628194), ('feel', 0.024417004), ('development', 0.021561593), ('access', 0.021442004)
2	('strong', 0.044284273), ('fantastic', 0.04360175), ('area', 0.035477895), ('service', 0.03483233), ('campus', 0.03249539), ('within', 0.032280233), ('move', 0.029394077), ('cloud', 0.022610152), ('recognition', 0.022519398), ('long', 0.02210394)
3	('opportunity', 0.11822484), ('learn', 0.101727866), ('lot', 0.098521665), ('technology', 0.054966234), ('compensation', 0.030700698), ('product', 0.029174833), ('grow', 0.028932085), ('growth', 0.0280799), ('new', 0.025318269), ('diversity', 0.024726227)
4	('care', 0.062152576), ('customer', 0.054786768), ('supportive', 0.054390047), ('support', 0.05400828), ('much', 0.038320474), ('take', 0.036308855), ('always', 0.029206766), ('kind', 0.027195321), ('positive', 0.02521524), ('nothing', 0.023003297)
5	('good', 0.24213894), ('environment', 0.058700096), ('pay', 0.04495831), ('nice', 0.037084777), ('best', 0.030772917), ('growth', 0.028956315), ('microsoft', 0.026673477), ('management', 0.021120165), ('awesome', 0.017766232), ('really', 0.017584989)
6	('manager', 0.09055669), ('well', 0.08658008), ('edge', 0.049292307), ('oslo', 0.033286713), ('cut', 0.02896876), ('look', 0.022702148), ('paid', 0.020974686), ('vacation', 0.018413002), ('autonomy', 0.017898945), ('week', 0.017816562)
7	('work', 0.18413955), ('great', 0.14570768), ('benefit', 0.10683731), ('culture', 0.086669065), ('company', 0.07607604), ('people', 0.050915502), ('team', 0.0348101), ('salary', 0.02884012), ('career', 0.025104418), ('worklife', 0.021395732)
8	('employee', 0.10603538), ('flexibility', 0.048886128), ('focus', 0.041282255), ('health', 0.035866186), ('world', 0.034459304),

		('community', 0.026717292), ('love', 0.023429329), ('impact', 0.022845982), ('talented', 0.02095775), ('mission', 0.020141298)
9		('balance', 0.2132651), ('life', 0.16851772), ('work', 0.08388359), ('excellent', 0.038957357), ('smart', 0.03346808), ('colleague', 0.030893894), ('help', 0.02933367), ('job', 0.026289158), ('interest', 0.026020337), ('time', 0.024863565)
10		('perk', 0.10602594), ('flexible', 0.09832693), ('hour', 0.04769965), ('different', 0.028850578), ('industry', 0.028241199), ('condition', 0.028078875), ('late', 0.027015226), ('home', 0.025749331), ('politics', 0.024813866), ('responsibility', 0.023523148)

ตารางที่ 4.15 ผลการวิเคราะห์คำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล รีวิวประโยชน์ที่ได้ บริษัท Microsoft

ผลการทำนายของแบบจำลองได้วิเคราะห์คำที่มีนัยสำคัญและหัวข้อ โดยในคอลัมน์ที่ 2 เป็น 10 อันดับของคำที่ปรากฏในหัวข้อในหัวนั้นๆ ซึ่งแต่ละหัวข้อมีการวิเคราะห์ชุดของคำที่ประกอบในหัวข้อและคำหลักแต่ละคำมีการถ่วงน้ำหนัก (weightage) กำกับให้กับแต่ละหัวข้อ ซึ่งเมื่อตั้งสมมุติฐาน และวิเคราะห์คำที่มีนัยสำคัญในแต่ละหัวข้อ พบร่วมกัน

หัวข้อที่ 1 แบรนด์ที่ดี โดยมีคำที่เกี่ยวข้อง คือ amaze, brand, tech, like, business, big, etc, feel, development, access

หัวข้อที่ 2 สวัสดิการ โดยมีคำที่เกี่ยวข้องคือ strong, fantastic, area, service, campus, within, move, cloud, recognition, long

หัวข้อที่ 3 เรียนรู้ประสบการณ์ โดยมีคำที่เกี่ยวข้องคือ opportunity, learn, lot, technology, compensation, product, grow, growth, new, diversity

หัวข้อที่ 4 ดูแลซัพพอร์ต โดยมีคำที่เกี่ยวข้องคือ care, customer, supportive, support, much, take, always, kind, positive, nothing

หัวข้อที่ 5 โดยรวม โดยมีคำที่เกี่ยวข้องคือ good, environment, pay, nice, best, growth, Microsoft, management, awesome, really

หัวข้อที่ 6 บริหาร โดยมีคำที่เกี่ยวข้องคือ manager, well, edge, oslo, cut, look, paid, vacation, autonomy, week

หัวข้อที่ 7 งาน โดยมีคำที่เกี่ยวข้องคือ work, great, benefit, culture, company, people, team, salary, career, worklife

หัวข้อที่ 8 พนักงาน โดยมีคำที่เกี่ยวข้องคือ employee, flexibility, focus, health, world, community, love, impact, talented, mission

หัวข้อที่ 9 worklifebalance โดยมีคำที่เกี่ยวข้องคือ balance, life, work, excellent, smart, colleague, help, job, interest, time

หัวข้อที่ 10 วันหยุด โดยมีคำที่เกี่ยวข้องคือ perk, flexible, hour, different, industry, condition, late, home, politics, responsibility

มีกลุ่มหัวข้อความหมายที่คล้ายคลึงกันคือ

กลุ่มสมมติฐาน	หัวข้อที่เกี่ยวข้อง
กลุ่มหัวข้อ สวัสดิการ	สวัสดิการ, ดูแลซัพพอร์ต, วันหยุด
กลุ่มหัวข้อ Worklifebalance	งาน, Work Life Balance
กลุ่มหัวข้อ การเรียนรู้	แบรนด์ที่ดี, เรียนรู้ประสบการณ์, โดยรวม, พนักงาน

ตารางที่ 4.16 กลุ่มสมมติฐานหัวข้อความหมายที่คล้ายคลึง ประযุณ์ที่ได้ใน Microsoft

## 9.

ลำดับหัวข้อ	10 อันดับคำที่ปรากฏในหัวข้อ
1	('microsoft', 0.047384273), ('think', 0.04717649), ('role', 0.043536607), ('sometimes', 0.043325286), ('change', 0.042803325), ('none', 0.03796925), ('need', 0.030675553), ('internal', 0.025777478), ('organization', 0.023772726), ('limited', 0.02184511)
2	('team', 0.12038362), ('get', 0.07985656), ('old', 0.040517285), ('worklife', 0.034632556), ('base', 0.033526387), ('experience', 0.027911307), ('system', 0.024554657), ('still', 0.024523731), ('challenge', 0.022736534), ('might', 0.022063065)
3	('less', 0.0692977), ('manager', 0.06634921), ('people', 0.063065246), ('good', 0.05718305), ('year', 0.041015137), ('meeting', 0.034602363), ('hire', 0.028938679), ('structure', 0.021616453), ('join', 0.018119076), ('every', 0.01708912)
4	('con', 0.13469683), ('bad', 0.08395426), ('really', 0.06537687), ('high', 0.057482064), ('position', 0.030481188), ('market', 0.029961651), ('senior', 0.029307278), ('cant', 0.026355237), ('do', 0.023095246), ('use', 0.019465752)

	(‘salary’, 0.0872542), (‘culture’, 0.0485729), (‘dont’, 0.04597189), (‘job’, 0.043431398), (‘like’, 0.03730283), (‘large’, 0.03588117), (‘competitive’, 0.031162715), (‘complex’, 0.02452704), (‘problem’, 0.022848286), (‘everything’, 0.02015349)
5	(‘work’, 0.20746139), (‘balance’, 0.06662906), (‘much’, 0.06413646), (‘time’, 0.055083457), (‘life’, 0.04626595), (‘hard’, 0.037137307), (‘long’, 0.031320926), (‘difficult’, 0.027715867), (‘pressure’, 0.02742417), (‘hour’, 0.02667905)
6	(‘many’, 0.05333779), (‘process’, 0.0487853), (‘new’, 0.030620893), (‘environment’, 0.02921481), (‘place’, 0.028742654), (‘lack’, 0.026637893), (‘support’, 0.022669692), (‘technology’, 0.020132417), (‘project’, 0.017603504), (‘poor’, 0.016857423)
7	(‘employee’, 0.050841596), (‘move’, 0.044199392), (‘well’, 0.036285724), (‘make’, 0.036076944), (‘learn’, 0.033061694), (‘could’, 0.024311744), (‘product’, 0.022753682), (‘bureaucracy’, 0.022560673), (‘right’, 0.019484056), (‘huge’, 0.019181827)
8	(‘management’, 0.075997725), (‘great’, 0.05819132), (‘promotion’, 0.050309945), (‘thing’, 0.03330062), (‘customer’, 0.029314045), (‘may’, 0.02719774), (‘feel’, 0.023370652), (‘focus’, 0.020510545), (‘easy’, 0.018404141), (‘load’, 0.016519303)
9	(‘company’, 0.09195588), (‘lot’, 0.05782778), (‘slow’, 0.049554456), (‘pay’, 0.047038637), (‘growth’, 0.041580223), (‘big’, 0.041290104), (‘low’, 0.040635884), (‘politics’, 0.03784646), (‘nothing’, 0.033062905), (‘career’, 0.030558847)
10	ตารางที่ 4.17 ผลการวิเคราะห์คำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล รีวิวผลเสียที่ได้ บริษัท Microsoft

ผลการทำนายของแบบจำลองได้วิเคราะห์คำที่มีนัยสำคัญและหัวข้อ โดยในคอลัมน์ที่ 2 เป็น 10 อันดับของคำที่ปรากฏในหัวข้อในหัวนั้นๆ ซึ่งแต่ละหัวข้อมีการวิเคราะห์ชุดของคำที่ประกอบในหัวข้อและคำหลักแต่ละคำมีการถ่วงน้ำหนัก (weightage) กำกับให้กับแต่ละหัวข้อ ซึ่งเมื่อตั้งสมมุติฐาน และวิเคราะห์คำที่มีนัยสำคัญในแต่ละหัวข้อ พบร่วม

หัวข้อที่ 1 ระบุเบี้ยบบริหาร โดยมีคำที่เกี่ยวข้อง คือ microsoft, think, role, sometimes, change, none, need, internal, organization,

หัวข้อที่ 2 เพื่อนร่วมทีม โดยมีคำที่เกี่ยวข้องคือ team, get, old, worklife, base, experience, system, still, challenge, might

หัวข้อที่ 3 องค์กร โดยมีคำที่เกี่ยวข้องคือ less, manager, people, good, year, meeting, hire, structure, join, every

หัวข้อที่ 4 ผลเสีย โดยมีคำที่เกี่ยวข้องคือ con, bad, really, high, position, market, senior, cant, do, use

หัวข้อที่ 5 เงินเดือน โดยมีคำที่เกี่ยวข้องคือ salary, culture, dont, job, like, large, competitive, complex, problem, everything

หัวข้อที่ 6 worklifebalance โดยมีคำที่เกี่ยวข้องคือ work, balance, much, time, life, hard, long, difficult, pressure, hour

หัวข้อที่ 7 งานหนัก โดยมีคำที่เกี่ยวข้องคือ many, process, new, environment, place, lack, support, technology, project, poor

หัวข้อที่ 8 เปลี่ยนตำแหน่ง โดยมีคำที่เกี่ยวข้องคือ employee, move, well, make, learn, could, product, bureaucracy, right, huge

หัวข้อที่ 9 บริหาร โดยมีคำที่เกี่ยวข้องคือ management, great, promotion, thing, customer, may, feel, focus, easy, load

หัวข้อที่ 10 เติบโตชา โดยมีคำที่เกี่ยวข้องคือ company, lot, slow, pay, growth, big, low, politics, nothing, career

#### มีกลุ่มหัวข้อความหมายที่คล้ายคลึงกันคือ

กลุ่มสมมติฐาน

หัวข้อที่เกี่ยวข้อง

กลุ่มหัวข้อองค์กร

องค์กร, เพื่อนร่วมทีม, ระเบียบบริหาร, เปลี่ยนตำแหน่ง,  
บริหาร

กลุ่มหัวข้อworklifebalance

Work Life Balance,งานหนัก

กลุ่มหัวข้อแรงกดดัน

เติบโตชา

กลุ่มหัวข้อเงินเดือน

เงินเดือน

ตารางที่ 4.18 กลุ่มสมมติฐานหัวข้อความหมายที่คล้ายคลึง ผลเสียที่ได้ ใน Microsoft

## ผลลัพธ์ของการจำแนกกลุ่มข้อความด้วยเทคนิค Top2Vec

ก่อนผู้วิจัยจะนำเสนอด้วยกับผลลัพธ์หัวข้อของเทคนิค Top2Vec ผู้วิจัยมีเรื่องซึ่งต้องไปนี้

ผู้วิจัยมีแนวการวิเคราะห์ผลลัพธ์ของเทคนิค Top2vec โดยจะวิเคราะห์ตีความเป็นกลุ่มๆ หัวข้อเนื่องจากเทคนิค Top2vec จะมีการกำหนดจำนวนข้อโดยอัตโนมัติจากกลุ่มความหมายแน่นที่ถูกค้นพบใน data ที่ผ่านกระบวนการ Embedding และตัวแบบจะระบุหัวข้อตามจำนวนนั้นซึ่งหมายความว่าเทคนิค Top2vec ให้จำนวนผลลัพธ์หัวข้อในแต่ละ data แตกต่างกัน บาง data ให้จำนวนมาก บาง data ให้จำนวนน้อยจึงคาดเดาได้ยากไม่ตายตัวทำให้เกิดอุปสรรคต่อการวิเคราะห์ตีความในแต่ละผลลัพธ์ ผู้วิจัยมีแนวทางการวิเคราะห์ผลลัพธ์หัวข้อความไม่แน่นอนของจำนวนหัวข้อที่ระบุออกมากของเทคนิค Top2vec โดยผู้วิจัยจะจัดกลุ่มหัวข้อเป็นกลุ่มๆ หมายความว่าการจัดกลุ่มหัวข้อนั้น แสดงให้เห็นว่าหัวข้อนั้นมีความคล้ายคลึงกันและสื่อความหมายไปในทำนองเดียวกัน ดังนั้นผู้วิจัยจะวิเคราะห์ตีความจากคำที่พบในแต่ละหัวข้อจากกลุ่มหัวข้อนั้นๆ แล้วตั้งข้อสันนิษฐานเกี่ยวกับกลุ่มหัวข้อนั้นๆว่าควรจะสื่อถึงอะไร

ผู้วิจัยมีแนวการวิเคราะห์โดยจะวิเคราะห์จากคำที่พบในแต่ละผลลัพธ์หัวข้อ (Topic) ในกลุ่มหัวข้อโดยจะพิจารณาจากคำที่ปรากฏเป็น 5 อันดับแรกในแต่ละผลลัพธ์หัวข้อ (Topic) ซึ่งเหตุผลที่เลือกเป็น 5 อันดับ เพราะลำดับคำในผลลัพธ์หัวข้อ (Topic) มีความหมายว่า ยิ่งอันดับต้นๆ การถ่วงน้ำหนัก (weightage) ของคำจะยิ่งสูงแปลว่าคำนั้นจะสื่อถึงความเกี่ยวเนื่องของหัวข้อ (Topic) นั้นมากตาม ดังนั้นผู้วิจัยจึงพยายามตีกรอบคำให้เป็น 5 อันดับในแต่ละหัวข้อ (Topic) ในกลุ่มหัวข้อที่จัดได้เพื่อให้วิเคราะห์สังเกตได้สะดวกขึ้น

- ผลการวิเคราะห์หากคำที่มีนัยสำคัญจากการจัดกลุ่มหัวข้อซึ่งแสดงเป็น wordcloud ในข้อมูลรีวิวประเภท headline บริษัท apple

ผู้วิจัยจัดกลุ่มหัวข้อให้เป็น 3 กลุ่ม โดยเรียกใช้ .hierarchical\_topic\_reduction(num\_topics = 3) ผ่าน library ของ top2vec กับตัวแปร topic model ของเรา

```
[63] # reduce to group of 3
t2v_a_h_16v2_dp_load.hierarchical_topic_reduction(num_topics=3)

[[26,
 27,
 20,
 22,
 28,
 17,
 35,
 6,
 12,
 1,
 16,
 25,
 19,
 29,
 5,
 37,
 39,
 23,
 4,
 13,
 18,
 30,
 3],
 [32, 31, 36, 11, 24, 2, 7, 9, 15, 38, 10],
 [8, 14, 21, 33, 34, 0]]
```

ภาพที่ 4.1 แสดงการจัดกลุ่มผลลัพธ์หัวข้อของเทคโนโลยี Top2vec ในข้อมูลประเภท headline ของบริษัท apple

หลังจากที่จัดกลุ่มหัวข้อแล้วนี้คือผลลัพธ์ของแต่ละหัวข้อ (ผลลัพธ์หัวข้อจะประกอบไปด้วยคำหลายคำ อยู่ภายในและสื่อถึงความเป็นหัวข้อนั้นๆ) ในกลุ่มที่ 1 ในข้อมูลประเภท headline ของบริษัท apple ซึ่งมีอยู่ 23 หัวข้อ

```
[['apple' 'retail' 'company' 'job' 'experience' 'amazing' 'work' 'best'
'place' 'great' 'benefits' 'good' 'culture'],
['apple' 'work' 'job' 'company' 'retail' 'benefits' 'experience'
'amazing' 'best' 'place' 'great' 'good' 'culture'],
['apple' 'company' 'retail' 'best' 'amazing' 'experience' 'great'
'place' 'work' 'job' 'benefits' 'good' 'culture'],
['apple' 'retail' 'company' 'place' 'job' 'work' 'best' 'culture'
'experience' 'amazing' 'great' 'good' 'benefits'],
[job' 'company' 'experience' 'apple' 'work' 'retail' 'place' 'benefits'
'culture' 'best' 'amazing' 'great' 'good'],
[company' 'job' 'experience' 'retail' 'work' 'apple' 'amazing' 'best'
'place' 'benefits' 'great' 'culture' 'good'],
[company' 'job' 'culture' 'work' 'retail' 'best' 'experience' 'place'
'good' 'great' 'benefits' 'apple' 'amazing'],
[company' 'job' 'work' 'retail' 'benefits' 'experience' 'best' 'great'
'place' 'culture' 'good' 'amazing' 'apple'],
[place' 'great' 'best' 'job' 'experience' 'company' 'good' 'amazing'
'culture' 'work' 'retail' 'benefits' 'apple'],
[place' 'work' 'job' 'company' 'best' 'great' 'experience' 'amazing'
'good' 'retail' 'benefits' 'culture' 'apple'],
[job' 'work' 'benefits' 'experience' 'great' 'good' 'company' 'retail'
'amazing' 'best' 'place' 'culture' 'apple'],
[benefits' 'company' 'job' 'experience' 'work' 'great' 'retail' 'best'
'good' 'amazing' 'culture' 'place' 'apple'],
[benefits' 'experience' 'great' 'good' 'work' 'best' 'job' 'amazing'
'company' 'retail' 'culture' 'place' 'apple'],
[work' 'job' 'benefits' 'company' 'experience' 'retail' 'best'
'culture' 'place' 'amazing' 'great' 'good' 'apple'],
[work' 'job' 'company' 'experience' 'best' 'place' 'retail' 'amazing'
'benefits' 'great' 'good' 'culture' 'apple'],
[culture' 'experience' 'company' 'retail' 'benefits' 'job' 'place'
'work' 'apple' 'great' 'good' 'best' 'amazing'],
[culture' 'company' 'experience' 'retail' 'job' 'work' 'benefits'
'place' 'best' 'great' 'apple' 'amazing' 'good'],
[culture' 'experience' 'place' 'best' 'benefits' 'company' 'good' 'job'
'retail' 'great' 'work' 'amazing' 'apple'],
[job' 'job' 'company' 'experience' 'work' 'apple' 'best' 'place'
'benefits' 'great' 'culture' 'good' 'amazing'],
[job' 'work' 'company' 'amazing' 'great' 'best' 'experience' 'retail'
'place' 'benefits' 'good' 'culture' 'apple'],
[job' 'experience' 'work' 'place' 'company' 'benefits' 'best' 'amazing'
'great' 'culture' 'retail' 'good' 'apple'],
[experience' 'amazing' 'great' 'best' 'benefits' 'place' 'job' 'work'
'good' 'retail' 'culture' 'company' 'apple'],
[place' 'job' 'experience' 'work' 'company' 'best' 'retail' 'benefits'
'culture' 'great' 'amazing' 'good' 'apple']]
```

ภาพที่ 4.2 แสดงผลลัพธ์ทั้งหมดหัวข้อในกลุ่มหัวข้อที่ 1 ของเทคโนโลยี Top2vec ในข้อมูลประเภท headline ของบริษัท apple

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท headline ของบริษัท apple ซึ่งมีอยู่ 23 หัวข้อ

และนี่คือผลลัพธ์จากการนับความถี่ของคำทุกคำที่ปรากฏในแต่ละหัวข้อ

```
Word frequencies:
apple: 23
retail: 23
company: 23
job: 23
experience: 23
amazing: 23
work: 23
best: 23
place: 23
great: 23
benefits: 23
good: 23
culture: 23
```

ภาพที่ 4.3 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท apple

คำและความถี่ที่นับได้ในกลุ่มที่ 1 ในข้อมูลประเภท headline ของบริษัท apple ไปแสดงผลผ่าน word cloud

Group topics 1 – apple\_headline all words each topics



ภาพที่ 4.4 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท headline ของบริษัท apple

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท headline ของบริษัท apple ซึ่งมีอยู่ 23 หัวข้อแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
[['apple' 'retail' 'company' 'job' 'experience']
 ['apple' 'work' 'job' 'company' 'retail']
 ['apple' 'company' 'retail' 'best' 'amazing']
 ['apple' 'retail' 'company' 'place' 'job']
 ['job' 'company' 'experience' 'apple' 'work']
 ['company' 'job' 'experience' 'retail' 'work']
 ['company' 'job' 'culture' 'work' 'retail']
 ['company' 'job' 'work' 'retail' 'benefits']
 ['place' 'great' 'best' 'job' 'experience']
 ['place' 'work' 'job' 'company' 'best']
 ['job' 'work' 'benefits' 'experience' 'great']
 ['benefits' 'company' 'job' 'experience' 'work']
 ['benefits' 'experience' 'great' 'good' 'work']
 ['work' 'job' 'benefits' 'company' 'experience']
 ['work' 'job' 'company' 'experience' 'best']
 ['culture' 'experience' 'company' 'retail' 'benefits']
 ['culture' 'company' 'experience' 'retail' 'job']
 ['culture' 'experience' 'place' 'best' 'benefits']
 ['retail' 'job' 'company' 'experience' 'work']
 ['job' 'work' 'company' 'amazing' 'great']
 ['job' 'experience' 'work' 'place' 'company']
 ['experience' 'amazing' 'great' 'best' 'benefits']
 ['place' 'job' 'experience' 'work' 'company']]
```

ภาพที่ 4.5 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ  
ในกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท apple

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏซึ่งแสดงเป็นคำ 5 อันดับแรก  
ในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท headline ของบริษัท apple ซึ่งมีอยู่ 23 หัวข้อ  
และนี่คือผลลัพธ์จากการนับความถี่เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
Word frequencies:
apple: 5
retail: 10
company: 18
job: 18
experience: 16
work: 15
best: 6
amazing: 3
place: 6
culture: 4
benefits: 8
great: 5
good: 1
```

ภาพที่ 4.6 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1  
ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท apple

คำนับและความถี่ที่นับได้ในกลุ่มหัวข้อที่ 1 ซึ่งแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ  
ในข้อมูลประเภท headline ของบริษัท apple ไปแสดงผลผ่าน word cloud

### Group topics 1 – apple\_headline top 5 words each topics



ภาพที่ 4.7 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท headline ของบริษัท apple

ผู้วิจัยพิจารณาจาก word cloud แสดงเฉพาะคำ 5 อันดับแรกปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท headline ของบริษัท apple ผู้วิจัยมีข้อสันนิษฐานว่าในกลุ่มหัวข้อที่ 1 ของ apple ข้อความรีวิวประเภท headline

มีความเกี่ยวข้องกับ: ชื่อบริษัท ประสบการณ์เยี่ยม งานค้าปลีกเยี่ยม สถานที่เยี่ยม ผลประโยชน์

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท headline ของบริษัท apple ซึ่งมีอยู่ 11 หัวข้อ

```
[[['best' 'amazing' 'good' 'job' 'great' 'work' 'experience' 'company'
  'culture' 'place' 'retail' 'apple' 'benefits']
  ['work' 'job' 'experience' 'company' 'great' 'best' 'good' 'place'
  'retail' 'amazing' 'culture' 'benefits' 'apple']
  ['good' 'great' 'best' 'amazing' 'job' 'experience' 'work' 'culture'
  'place' 'company' 'retail' 'benefits' 'apple']
  ['best' 'good' 'great' 'amazing' 'job' 'work' 'experience' 'place'
  'culture' 'company' 'retail' 'apple' 'benefits']
  ['good' 'great' 'amazing' 'job' 'best' 'work' 'experience' 'retail'
  'company' 'place' 'apple' 'culture' 'benefits']
  ['good' 'great' 'job' 'amazing' 'work' 'best' 'company' 'experience'
  'place' 'retail' 'apple' 'culture' 'benefits']
  ['good' 'great' 'amazing' 'best' 'job' 'work' 'experience' 'company'
  'retail' 'benefits' 'culture' 'place' 'apple']
  ['amazing' 'great' 'job' 'work' 'good' 'best' 'company' 'experience'
  'place' 'retail' 'benefits' 'apple' 'culture']
  ['amazing' 'great' 'best' 'good' 'work' 'job' 'experience' 'company'
  'place' 'retail' 'apple' 'culture' 'benefits']
  ['great' 'amazing' 'good' 'best' 'work' 'job' 'experience' 'place'
  'company' 'retail' 'apple' 'benefits' 'culture']
  ['great' 'good' 'amazing' 'best' 'job' 'work' 'experience' 'company'
  'place' 'retail' 'apple' 'benefits' 'culture']]
```

ภาพที่ 4.8 แสดงผลลัพธ์แต่ละหัวข้อในกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท apple

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท headline ของบริษัท apple ซึ่งมีอยู่ 11 หัวข้อ

และนี่คือผลลัพธ์จากการนับความถี่ของคำทุกคำที่ปรากฏในแต่ละหัวข้อ

```
Word frequencies:
best: 11
amazing: 11
good: 11
job: 11
great: 11
work: 11
experience: 11
company: 11
culture: 11
place: 11
retail: 11
apple: 11
benefits: 11
```

ภาพที่ 4.9 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท apple

นำคำและความถี่ที่นับได้ในกลุ่มที่ 2 ในข้อมูลประเภท headline ของบริษัท apple ไปแสดงผลผ่าน word cloud

Group topics 2 – apple\_headline all words each topics



ภาพที่ 4.10 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท headline ของบริษัท apple

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท headline ของบริษัท apple ซึ่งมีอยู่ 11 หัวข้อแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
[['best' 'amazing' 'good' 'job' 'great']
 ['work' 'job' 'experience' 'company' 'great']
 ['good' 'great' 'best' 'amazing' 'job']
 ['best' 'good' 'great' 'amazing' 'job']
 ['good' 'great' 'amazing' 'job' 'best']
 ['good' 'great' 'job' 'amazing' 'work']
 ['good' 'great' 'amazing' 'best' 'job']
 ['amazing' 'great' 'job' 'work' 'good']
 ['amazing' 'great' 'best' 'good' 'work']
 ['great' 'amazing' 'good' 'best' 'work']
 ['great' 'good' 'amazing' 'best' 'job']]
```

ภาพที่ 4.11 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ  
ในกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท apple

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏชื่้นแสดงเป็นคำ 5 อันดับแรก  
ในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท headline ของบริษัท apple ซึ่งมีอยู่ 11 หัวข้อ  
และนี่คือผลลัพธ์จากการนับความถี่เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
Word frequencies:
best: 8
amazing: 10
good: 10
job: 9
great: 11
work: 5
experience: 1
company: 1
```

ภาพที่ 4.12 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2  
ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท apple

นำคำและความถี่ที่นับได้ในกลุ่มหัวข้อที่ 2 ซึ่งแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ  
ในข้อมูลประเภท headline ของบริษัท apple ไปแสดงผลผ่าน word cloud

Group topics 2 – apple\_headline top 5 words each topics



ภาพที่ 4.13 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท headline ของบริษัท apple

ผู้วิจัยพิจารณาจาก word cloud แสดงเฉพาะคำ 5 อันดับแรกปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท headline ของบริษัท apple ผู้วิจัยมีข้อสันนิษฐานว่าในกลุ่มหัวข้อที่ 2 ของ apple ข้อความรีวิวประเภท headline มีความเกี่ยวข้องกับ: ชื่อชุมบริษัท ประสบการณ์เยี่ยม

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท headline ของบริษัท apple ซึ่งมีอยู่ 6 หัวข้อ

```
[[['company' 'best' 'job' 'work' 'retail' 'apple' 'experience' 'place'
  'benefits' 'culture' 'amazing' 'great' 'good']
 ['company' 'job' 'great' 'best' 'experience' 'work' 'retail' 'benefits'
  'amazing' 'good' 'culture' 'place' 'apple']
 ['company' 'good' 'best' 'job' 'great' 'amazing' 'work' 'retail'
  'experience' 'place' 'benefits' 'apple' 'culture']
 ['company' 'best' 'job' 'great' 'work' 'experience' 'amazing' 'good'
  'retail' 'place' 'benefits' 'culture' 'apple']
 ['company' 'work' 'job' 'amazing' 'best' 'experience' 'great' 'retail'
  'benefits' 'apple' 'place' 'good' 'culture']
 ['company' 'great' 'job' 'best' 'amazing' 'work' 'experience' 'good'
  'retail' 'benefits' 'place' 'apple' 'culture']]
```

ภาพที่ 4.14 แสดงผลลัพธ์แต่ละหัวข้อในกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท apple

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท headline ของบริษัท apple ซึ่งมีอยู่ 6 หัวข้อ

และนี่คือผลลัพธ์จากการนับความถี่ของคำทุกคำที่ปรากฏในแต่ละหัวข้อ

```
Word frequencies:
company: 6
best: 6
job: 6
work: 6
retail: 6
apple: 6
experience: 6
place: 6
benefits: 6
culture: 6
amazing: 6
great: 6
good: 6
```

ภาพที่ 4.15 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท apple

นำคำและความถี่ที่นับได้ในกลุ่มที่ 3 ในข้อมูลประเภท headline ของบริษัท apple ไปแสดงผลผ่าน word cloud

Group topics 3 – apple\_headline all words each topics



ภาพที่ 4.16 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท headline ของบริษัท apple

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท headline ของบริษัท apple ซึ่งมีอยู่ 6 หัวข้อแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
[[ 'company' 'best' 'job' 'work' 'retail']
 ['company' 'job' 'great' 'best' 'experience']
 ['company' 'good' 'best' 'job' 'great']
 ['company' 'best' 'job' 'great' 'work']
 ['company' 'work' 'job' 'amazing' 'best']
 ['company' 'great' 'job' 'best' 'amazing']]
```

ภาพที่ 4.17 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ  
ในกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท apple

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏซึ่งแสดงเป็นคำ 5 อันดับแรก  
ในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท headline ของบริษัท apple ซึ่งมีอยู่ 6 หัวข้อ  
และนี่คือผลลัพธ์จากการนับความถี่เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
Word frequencies:
company: 6
best: 6
job: 6
work: 3
retail: 1
great: 4
experience: 1
good: 1
amazing: 2
```

ภาพที่ 4.18 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3  
ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท apple

นำคำและความถี่ที่นับได้ในกลุ่มหัวข้อที่ 3 ซึ่งแสดงเฉพาะคำ 5 อันดับแรกในหัวข้อ ในข้อมูลประเภท  
headline ของบริษัท apple ไปแสดงผลผ่าน word cloud

Group topics 3 – apple\_headline top 5 words each topics



ภาพที่ 4.19 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท headline ของบริษัท apple

ผู้วิจัยมีข้อสันนิษฐานว่าในกลุ่มหัวข้อที่ 3 ของ apple ข้อความรีวิวประเภท headline มีความเกี่ยวข้องกับ: ชื่อชุมชนบริษัท งานค้าปลีกคือ ประสบการณ์เยี่ยม

2. ผลการวิเคราะห์หาคำที่มีนัยสำคัญจากการจัดกลุ่มหัวข้อซึ่งแสดงเป็น wordcloud ในข้อมูลรีวิวประเภท pros บริษัท apple

การจัดกลุ่มที่เป็นไปได้เท่ากับ 1 กลุ่ม แต่เนื่องจากผลลัพธ์ที่ได้มีแค่ 2 หัวข้อจึงแสดง wordcloud ในแต่ละหัวข้อ

```
[78] # reduce to group of 1
      t2v_a_p_16v2_dp_load.hierarchical_topic_reduction(num_topics=1)
[[1, 0]]
```

ภาพที่ 4.20 แสดงการจัดกลุ่มผลลัพธ์หัวข้อของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท apple

ผลลัพธ์หัวข้อที่ 0 ในข้อมูลประเภท pros ของบริษัท apple

```
[['benefits' 'opportunities' 'employees' 'career' 'job' 'company' 'work'
 'employee' 'management' 'experience' 'coworkers' 'working' 'pay'
 'training' 'development' 'discounts' 'salary' 'health' 'decent'
 'friendly' 'perks' 'life' 'culture' 'products' 'environment' 'fun'
 'learn' 'growth' 'hours' 'best' 'excellent' 'place' 'balance' 'great'
 'team' 'care' 'good' 'amazing' 'get' 'people' 'lot' 'lots' 'really'
 'time' 'nice' 'flexible' 'apple' 'stock']]
```

ภาพที่ 4.21 แสดงผลลัพธ์หัวข้อที่ 0 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท apple

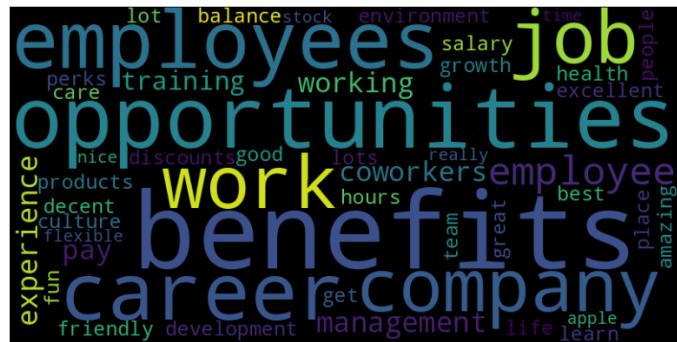
ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏในหัวข้อที่ 0 ในข้อมูลประเภท pros ของบริษัท apple และนี่คือผลลัพธ์จากการนับความถี่ของคำทุกคำที่ปรากฏในหัวข้อ

```
Word frequencies: environment: 1
benefits: 1           fun: 1
opportunities: 1      learn: 1
employees: 1           growth: 1
career: 1              hours: 1
job: 1                 best: 1
company: 1             excellent: 1
work: 1                place: 1
employee: 1            balance: 1
management: 1          great: 1
experience: 1          team: 1
coworkers: 1           care: 1
working: 1              good: 1
pay: 1                  amazing: 1
training: 1             get: 1
development: 1          people: 1
discounts: 1            lot: 1
salary: 1               lots: 1
health: 1               really: 1
decent: 1               time: 1
friendly: 1             nice: 1
perks: 1                flexible: 1
life: 1                 apple: 1
culture: 1              stock: 1
products: 1
```

ภาพที่ 4.22 แสดงผลลัพธ์การนับความถี่ของคำในหัวข้อที่ 0 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท apple

นำคำและความถี่ที่นับได้ในหัวข้อที่ 0 ในข้อมูลประเภท pros ของบริษัท apple ไปแสดงผลผ่าน word cloud

### Topic 0 – apple\_pros all words



ภาพที่ 4.23 word cloud แสดงคำทุกคำที่ปรากฏในหัวข้อที่ 0 ในข้อมูลประเภท pros ของบริษัท apple

ผลลัพธ์หัวข้อที่ 0 ในข้อมูลประเภท pros ของบริษัท apple แสดงเฉพาะคำ 5 อันดับแรกในหัวข้อ

```
[[ 'benefits' 'opportunities' 'employees' 'career' 'job']]
```

ภาพที่ 4.24 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ  
ที่ 0 ของเทคนิค Top2vec ของข้อมูลประเภท pros ของบริษัท apple

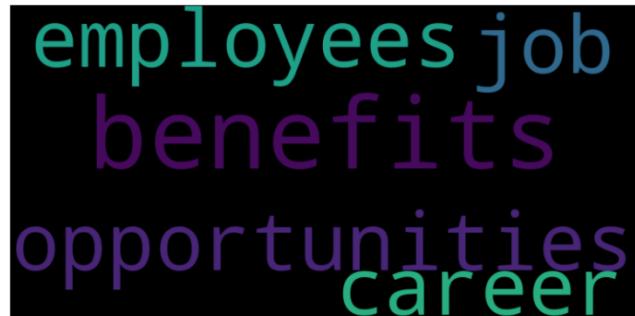
ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏในหัวข้อที่ 0 ในข้อมูลประเภท pros ของบริษัท apple และนี้คือผลลัพธ์จากการนับความถี่เฉพาะคำ 5 อันดับแรกในหัวข้อ

```
Word frequencies:  
benefits: 1  
opportunities: 1  
employees: 1  
career: 1  
job: 1
```

ภาพที่ 4.25 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในหัวข้อที่ 0  
ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท apple

นำคำและความถี่ที่นับได้ในหัวข้อที่ 0 ซึ่งแสดงเฉพาะคำ 5 อันดับแรกในหัวข้อของข้อมูลประเภท pros ของบริษัท apple ไปแสดงผลผ่าน word cloud

Topic 0 – apple\_pros top 5 words



ภาพที่ 4.26 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในหัวข้อที่ 0 ในข้อมูลประเภท pros ของบริษัท apple

ผู้วิจัยพิจารณาจาก word cloud แสดงเฉพาะคำ 5 อันดับแรกปรากฏในหัวข้อที่ 0 ผู้วิจัยมีข้อสันนิษฐานว่าในหัวข้อที่ 0 ของ apple ข้อความรีวิวประเภท pros มีความเกี่ยวข้องกับ: ผลประโยชน์ของพนักงาน โอกาสของพนักงาน

ผลลัพธ์หัวข้อที่ 1 ในข้อมูลประเภท pros ของบริษัท apple

```
[['apple' 'benefits' 'company' 'opportunities' 'employees' 'products'
'employee' 'work' 'coworkers' 'experience' 'management' 'job' 'working'
'career' 'discounts' 'perks' 'development' 'pay' 'health' 'training'
'learn' 'best' 'growth' 'environment' 'decent' 'care' 'salary'
'friendly' 'life' 'stock' 'balance' 'culture' 'place' 'amazing' 'get'
'hours' 'excellent' 'great' 'really' 'flexible' 'time' 'people' 'fun'
'lot' 'good' 'nice' 'team' 'lots']]
```

ภาพที่ 4.27 แสดงผลลัพธ์หัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท apple

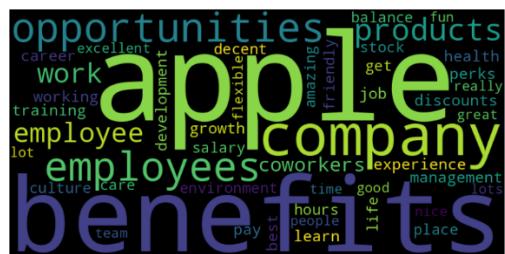
ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏในหัวข้อที่ 0 ในข้อมูลประเภท pros ของบริษัท apple และนี้คือผลลัพธ์จากการนับความถี่ของคำทุกคำที่ปรากฏในหัวข้อ

```
Word frequencies: decent: 1
apple: 1 care: 1
benefits: 1 salary: 1
company: 1 friendly: 1
opportunities: 1 life: 1
employees: 1 stock: 1
products: 1 balance: 1
employee: 1 culture: 1
work: 1 place: 1
coworkers: 1 amazing: 1
experience: 1 get: 1
management: 1 hours: 1
job: 1 excellent: 1
working: 1 great: 1
career: 1 really: 1
discounts: 1 flexible: 1
perks: 1 time: 1
development: 1 people: 1
pay: 1 fun: 1
health: 1 lot: 1
training: 1 good: 1
learn: 1 nice: 1
best: 1 team: 1
growth: 1 lots: 1
environment: 1
```

ภาพที่ 4.28 แสดงผลลัพธ์การนับความถี่ของคำในหัวข้อที่ 0 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท apple

นำคำและความคิดที่นับได้ในหัวข้อที่ 1 ในข้อมูลประเภท pros ของบริษัท apple ไปแสดงผลผ่าน word cloud

## Topic 1 – apple pros all words



ภาพที่ 4.29 word cloud แสดงคำทุกคำที่ปรากฏในหัวข้อที่ 1 ในข้อมูลประเภท pros ของบริษัท apple

ผลลัพธ์หัวข้อที่ 0 ในข้อมูลประเภท pros ของบริษัท apple แสดงเฉพาะคำ 5 อันดับแรกในหัวข้อ

```
[[ 'apple' 'benefits' 'company' 'opportunities' 'employees' ]]
```

ภาพที่ 4.30 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อที่ 1 ของเทคนิค Top2vec ของข้อมูลประเภท headline ของบริษัท apple

ในการวิเคราะห์ผลลัพธ์ทั่วข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏในหัวข้อที่ 1 ในข้อมูลประเภท pros ของบริษัท apple และนี่คือผลลัพธ์จากการนับความถี่เฉพาะคำ 5 อันดับแรกในหัวข้อ

```
Word frequencies:
apple: 1
benefits: 1
company: 1
opportunities: 1
employees: 1
```

ภาพที่ 4.31 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท apple

นำคำและความถี่ที่นับได้ในหัวข้อที่ 1 ซึ่งแสดงเฉพาะคำ 5 อันดับแรกในหัวข้อของข้อมูลประเภท pros ของบริษัท apple ไปแสดงผลผ่าน word cloud

Topic 1 – apple\_pros top 5 words



ภาพที่ 4.32 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในหัวข้อที่ 1 ในข้อมูลประเภท pros ของบริษัท apple

ผู้วิจัยพิจารณาจาก word cloud แสดงเฉพาะคำ 5 อันดับแรกปรากฏในหัวข้อที่ ผู้วิจัยมีข้อสันนิษฐานว่าในหัวข้อที่ 1 ของ apple ข้อความรีวิวประเภท pros มีความเกี่ยวข้องกับ: ผลประโยชน์ของพนักงาน โอกาสของพนักงาน

3. ผลการวิเคราะห์หาคำที่มีนัยสำคัญจากการจัดกลุ่มหัวข้อซึ่งแสดงเป็น wordcloud ในข้อมูลรีวิวประเภท cons บริษัท apple

การจัดกลุ่มที่เป็นไปได้แต่เนื่องจากผลลัพธ์ที่ได้มีแค่ 3 หัวข้อจึงแสดง wordcloud ในแต่ละหัวข้อ

```
[ ] # reduce to group of 2
t2v_a_c_16v2_f1_load.hierarchical_topic_reduction(num_topics=2)
[[[0], [2, 1]]]

[57] # reduce to group of 1
t2v_a_c_16v2_f1_load.hierarchical_topic_reduction(num_topics=1)
[[[2, 1, 0]]]
```

ภาพที่ 4.33 แสดงการจัดกลุ่มผลลัพธ์หัวข้อของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท apple

ผลลัพธ์หัวข้อที่ 0 ในข้อมูลประเภท cons ของบริษัท apple

```
[['management' 'opportunities' 'employees' 'managers' 'work' 'job'
'career' 'customers' 'working' 'company' 'retail' 'busy' 'schedule'
'experience' 'hours' 'life' 'times' 'growth' 'balance' 'pay'
'difficult' 'time' 'culture' 'move' 'environment' 'get' 'sometimes'
'really' 'hard' 'day' 'lot' 'team' 'think' 'like' 'bad' 'people'
'nothing' 'apple' 'high' 'always' 'dont' 'even' 'good' 'great' 'much'
'cons' 'many' 'could' 'none' 'long']]
```

ภาพที่ 4.34 แสดงผลลัพธ์หัวข้อที่ 0 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท apple

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏในหัวข้อที่ 0 ในข้อมูลประเภท cons ของบริษัท apple และนี่คือผลลัพธ์จากการนับความถี่ของคำที่ปรากฏในหัวข้อ

```
Word frequencies: culture: 1
management: 1 move: 1
opportunities: 1 environment: 1
employees: 1 get: 1
managers: 1 sometimes: 1
work: 1 really: 1
job: 1 hard: 1
career: 1 day: 1
customers: 1 lot: 1
working: 1 team: 1
company: 1 think: 1
retail: 1 like: 1
busy: 1 bad: 1
schedule: 1 people: 1
experience: 1 nothing: 1
hours: 1 apple: 1
life: 1 high: 1
times: 1 always: 1
growth: 1 dont: 1
balance: 1 even: 1
pay: 1 good: 1
difficult: 1 great: 1
time: 1 much: 1
cons: 1 many: 1
could: 1 none: 1
long: 1
```

ภาพที่ 4.35 แสดงผลลัพธ์การนับความถี่ของคำในหัวข้อที่ 0 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท apple

นำคำและความถี่ที่นับได้ในหัวข้อที่ 0 ในข้อมูลประเภท pros ของบริษัท apple ไปแสดงผลผ่าน word cloud

## Topic 0 – apple\_cons all words



ภาพที่ 4.36 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 0 ในข้อมูลประเภท cons ของบริษัท apple

ผลลัพธ์ทั้งหมดที่ 0 ในข้อมูลประเภท cons ของบริษัท apple แสดงเฉพาะคำ 5 อันดับแรกในหัวข้อ

```
[[ 'management' 'opportunities' 'employees' 'managers' 'work']]
```

ภาพที่ 4.37 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อที่ 0 ของเทคนิค Top2vec ของข้อมูลประเภท cons ของบริษัท apple

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏในหัวข้อที่ 0 ในข้อมูลประเภท cons ของบริษัท apple และนี่คือผลลัพธ์จากการนับความถี่เฉพาะคำ 5 อันดับแรกในหัวข้อ

```
Word frequencies:  
management: 1  
opportunities: 1  
employees: 1  
managers: 1  
work: 1
```

ภาพที่ 4.38 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในหัวข้อที่ 0 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท apple

นำคำและความถี่ที่นับได้ในหัวข้อที่ 0 ซึ่งแสดงเฉพาะคำ 5 อันดับแรกในหัวข้อของข้อมูลประเภท cons ของบริษัท apple ไปแสดงผลผ่าน word cloud

Topic 0 – apple\_cons top 5 words



ภาพที่ 4.39 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในหัวข้อที่ 0 ในข้อมูลประเภท cons ของบริษัท apple

ผู้วิจัยพิจารณาจาก word cloud แสดงเฉพาะคำ 5 อันดับแรกปรากฏในหัวข้อที่ 0 ผู้วิจัยมีข้อสันนิษฐานว่าในหัวข้อที่ 0 ของ apple ข้อความรีวิวประเภท cons มีความเกี่ยวข้องกับ: การจัดการพนักงาน โอกาสของพนักงาน ผู้จัดการ

ผลลัพธ์หัวข้อที่ 1 ในข้อมูลประเภท cons ของบริษัท apple

```
[['cons' 'bad' 'dont' 'always' 'really' 'like' 'think' 'time' 'good'
 'opportunities' 'much' 'sometimes' 'life' 'experience' 'long' 'times'
 'none' 'great' 'hard' 'nothing' 'job' 'difficult' 'could' 'hours'
 'many' 'pay' 'lot' 'culture' 'high' 'retail' 'work' 'get' 'environment'
 'team' 'working' 'even' 'career' 'day' 'balance' 'move' 'management'
 'schedule' 'busy' 'apple' 'company' 'customers' 'managers' 'people'
 'little' 'employees']]
```

ภาพที่ 4.40 แสดงผลลัพธ์หัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท apple

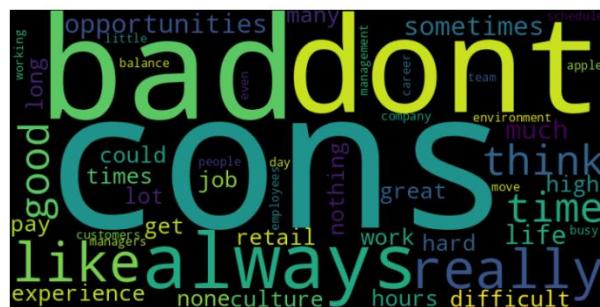
ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏในหัวข้อที่ 0 ในข้อมูลประเภท cons ของบริษัท apple และนี่คือผลลัพธ์จากการนับความถี่ของคำทุกคำที่ปรากฏในหัวข้อ

```
Word frequencies: many: 1  
cons: 1 pay: 1  
bad: 1 lot: 1  
dont: 1 culture: 1  
always: 1 high: 1  
really: 1 retail: 1  
like: 1 work: 1  
think: 1 get: 1  
time: 1 environment: 1  
good: 1 team: 1  
opportunities: 1 working: 1  
much: 1 even: 1  
sometimes: 1 career: 1  
life: 1 day: 1  
experience: 1 balance: 1  
long: 1 move: 1  
times: 1 management: 1  
none: 1 schedule: 1  
great: 1 busy: 1  
hard: 1 apple: 1  
nothing: 1 company: 1  
job: 1 customers: 1  
difficult: 1 managers: 1  
could: 1 people: 1  
hours: 1 little: 1  
employees: 1
```

ภาพที่ 4.41 แสดงผลลัพธ์การนับความถี่ของคำในหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท apple

นำคำและความถี่ที่นับได้ในหัวข้อที่ 0 ในข้อมูลประเภท pros ของบริษัท apple ไปแสดงผลผ่าน word cloud

## Topic 1 – apple cons all words



ภาพที่ 4.42 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท cons ของบริษัท apple

ผลลัพธ์ทั่วไปที่ 1 ในข้อมูลประเภท cons ของบริษัท apple แสดงเฉพาะคำ 5 อันดับแรกในทั้งหมด

```
[[ 'cons' 'bad' 'dont' 'always' 'really']]
```

ภาพที่ 4.43 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อที่ 1 ของเทคนิค Top2vec ของข้อมูลประเภท cons ของบริษัท apple

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏในหัวข้อที่ 1 ในข้อมูลประเภท cons ของบริษัท apple และนี่คือผลลัพธ์จากการนับความถี่เฉพาะคำ 5 อันดับแรกในหัวข้อ

```
Word frequencies:
cons: 1
bad: 1
dont: 1
always: 1
really: 1
```

ภาพที่ 4.44 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในหัวข้อที่ 1 ของเทคนิค Top2vec  
ในข้อมูลประเภท cons ของบริษัท apple

คำและความถี่ที่นับได้ในหัวข้อที่ 1 ซึ่งแสดงเฉพาะคำ 5 อันดับแรกในหัวข้อของข้อมูลประเภท pros ของบริษัท apple ไปแสดงผลผ่าน word cloud

Topic 1 – apple\_cons top 5 words



ภาพที่ 4.45 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในหัวข้อที่ 1 ในข้อมูลประเภท cons  
ของบริษัท apple

ผู้วิจัยพิจารณาจาก word cloud แสดงเฉพาะคำ 5 อันดับแรกปรากฏในหัวข้อที่ 1  
ผู้วิจัยมีข้อสันนิษฐานว่าในหัวข้อที่ 1 ของ apple ข้อความรีวิวประเภท cons  
มีความเกี่ยวข้องกับ: คิดว่าไม่นี๊อเสีย

ผลลัพธ์หัวข้อที่ 2 ในข้อมูลประเภท cons ของบริษัท apple

```
[['cons' 'company' 'job' 'employees' 'opportunities' 'management'
'managers' 'work' 'customers' 'career' 'retail' 'working' 'pay' 'dont'
'experience' 'hours' 'really' 'life' 'like' 'environment' 'bad' 'time'
'times' 'always' 'sometimes' 'think' 'team' 'good' 'balance' 'culture'
'great' 'apple' 'much' 'growth' 'schedule' 'could' 'get' 'difficult'
'many' 'hard' 'lot' 'busy' 'move' 'high' 'day' 'long' 'people' 'even'
'none' 'nothing']]
```

ภาพที่ 4.46 แสดงผลลัพธ์หัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท apple

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏในหัวข้อที่ 2 ในข้อมูลประเภท cons ของบริษัท apple และนี่คือผลลัพธ์จากการนับความถี่ของคำที่ปรากฏในหัวข้อ

```
Word frequencies: sometimes: 1
cons: 1           think: 1
company: 1        team: 1
job: 1            good: 1
employees: 1      balance: 1
opportunities: 1 culture: 1
management: 1     great: 1
managers: 1       apple: 1
work: 1            much: 1
customers: 1      growth: 1
career: 1          schedule: 1
retail: 1          could: 1
working: 1         get: 1
pay: 1             difficult: 1
dont: 1            many: 1
experience: 1     hard: 1
hours: 1           lot: 1
really: 1          busy: 1
life: 1            move: 1
like: 1             high: 1
environment: 1    day: 1
bad: 1              long: 1
time: 1             people: 1
times: 1            even: 1
always: 1          none: 1
never: 1           nothing: 1
```

ภาพที่ 4.47 แสดงผลลัพธ์การนับความถี่ของคำในหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท apple

นำคำและความถี่ที่นับได้ในหัวข้อที่ 2 ในข้อมูลประเภท cons ของบริษัท apple ไปแสดงผลผ่าน word cloud

### Topic 2 – apple\_cons all words



ภาพที่ 4.48 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูล  
ประเภท cons ของบริษัท apple

ผลลัพธ์หัวข้อที่ 2 ในข้อมูลประเภท cons ของบริษัท apple แสดงเฉพาะคำ 5 อันดับแรกในหัวข้อ

```
[['cons' 'company' 'job' 'employees' 'opportunities']]
```

ภาพที่ 4.49 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อที่ 2 ของเทคนิค Top2vec  
ของข้อมูลประเภท cons ของบริษัท apple

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏในหัวข้อที่ 2 ในข้อมูลประเภท cons ของบริษัท apple และนี่คือผลลัพธ์จากการนับความถี่เฉพาะคำ 5 อันดับแรกในหัวข้อ

```
Word frequencies:  
cons: 1  
company: 1  
job: 1  
employees: 1  
opportunities: 1
```

ภาพที่ 4.50 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในหัวข้อที่ 2 ของเทคนิค Top2vec  
ในข้อมูลประเภท cons ของบริษัท apple

นำคำและความถี่ที่นับได้ในหัวข้อที่ 2 ซึ่งแสดงเฉพาะคำ 5 อันดับแรกในหัวข้อของข้อมูลประเภท cons ของบริษัท apple ไปแสดงผลผ่าน word cloud

### Topic 2 – apple cons top 5 words



ภาพที่ 4.51 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในหัวข้อที่ 2 ในข้อมูลประเภท cons ของบริษัท apple

ผู้วิจัยพิจารณาจาก word cloud แสดงเฉพาะคำ 5 อันดับแรกปรากฏในหัวข้อที่ 2 ผู้วิจัยมีข้อสันนิษฐานว่าในหัวข้อที่ 2 ของ apple ข้อความรีวิวประเภท cons มีความเกี่ยวข้องกับ: งานมีข้อเสีย โอกาสของพนักงาน

4. ผลการวิเคราะห์หาคำที่มีนัยสำคัญจากการจัดกลุ่มหัวข้อซึ่งแสดงเป็น wordcloud ในข้อมูลรีวิวประเภท headline บริษัท google  
จัดกลุ่มหัวข้อให้เป็น 3 กลุ่ม

```
[83] # reduce to group of 3
      t2v_gg_h_16v2_dp_load.hierarchical_topic_reduction(num_topics=3)
[[[37, 20, 19, 0, 35, 11, 9, 13, 38, 8, 17, 10, 24, 30, 25, 14, 22, 6],
 [12, 7, 18, 21, 27, 3, 39, 1, 32, 23, 29, 31, 33, 36, 4],
 [16, 5, 28, 15, 34, 26, 2]]]
```

ภาพที่ 4.52 แสดงการจัดกลุ่มผลลัพธ์หัวข้อของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท google

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท headline ของบริษัท google ซึ่งมีอยู่ 18 หัวข้อ

```
[['place' 'best' 'good' 'great' 'company' 'experience' 'nice' 'work'
 'amazing' 'google' 'culture' 'people']
 ['place' 'great' 'best' 'amazing' 'experience' 'company' 'good' 'nice'
 'work' 'culture' 'google' 'people']
 ['place' 'work' 'best' 'company' 'experience' 'google' 'culture'
 'people' 'amazing' 'great' 'nice' 'good']
 ['place' 'work' 'company' 'best' 'great' 'amazing' 'experience' 'nice'
 'google' 'good' 'people' 'culture']
 ['google' 'place' 'company' 'great' 'best' 'work' 'amazing' 'experience'
 'people' 'good' 'nice' 'culture']
 ['google' 'company' 'place' 'work' 'people' 'best' 'experience'
 'culture' 'great' 'amazing' 'nice' 'good']
 ['google' 'company' 'work' 'experience' 'best' 'place' 'great' 'amazing'
 'nice' 'people' 'good' 'culture']
 ['company' 'google' 'experience' 'work' 'place' 'people' 'culture'
 'best' 'nice' 'amazing' 'good' 'great']
 ['experience' 'google' 'place' 'company' 'amazing' 'work' 'best' 'great'
 'culture' 'nice' 'good' 'people']
 ['company' 'experience' 'work' 'google' 'people' 'culture' 'place'
 'amazing' 'best' 'great' 'good' 'nice']
 ['culture' 'place' 'experience' 'best' 'company' 'great' 'people' 'good'
 'work' 'amazing' 'nice' 'google']
 ['culture' 'company' 'experience' 'work' 'best' 'good' 'great' 'place'
 'amazing' 'nice' 'people' 'google']
 ['great' 'experience' 'best' 'amazing' 'nice' 'good' 'work' 'company'
 'place' 'people' 'culture' 'google']
 ['experience' 'work' 'great' 'best' 'amazing' 'good' 'company' 'place'
 'nice' 'culture' 'google' 'people']
 ['experience' 'best' 'amazing' 'good' 'great' 'nice' 'place' 'work'
 'company' 'culture' 'google' 'people']
 ['experience' 'place' 'work' 'company' 'best' 'culture' 'amazing'
 'great' 'good' 'people' 'nice' 'google']
 ['work' 'company' 'best' 'experience' 'amazing' 'culture' 'place'
 'great' 'people' 'good' 'nice' 'google']
 ['place' 'company' 'experience' 'work' 'best' 'culture' 'great'
 'amazing' 'google' 'good' 'nice' 'people']]
```

ภาพที่ 4.53 แสดงผลลัพธ์แต่ละหัวข้อในกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท google

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท headline ของบริษัท google ซึ่งมีอยู่ 18 หัวข้อ

และนี่คือผลลัพธ์จากการนับความถี่ของคำทุกคำที่ปรากฏในแต่ละหัวข้อ

```
Word frequencies:
place: 18
best: 18
good: 18
great: 18
company: 18
experience: 18
nice: 18
work: 18
amazing: 18
google: 18
culture: 18
people: 18
```

ภาพที่ 4.54 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท google

คำและความถี่ที่นับได้ในกลุ่มที่ 1 ในข้อมูลประเภท headline ของบริษัท google ไปแสดงผลผ่าน word cloud

Group topics 1 – google\_headline all words each topics



ภาพที่ 4.55 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1  
ในข้อมูลประเภท headline ของบริษัท google

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 1 ของข้อมูลประเภท headline ของบริษัท google ซึ่งมีอยู่ 18 หัวข้อแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
[['place' 'best' 'good' 'great' 'company'],
 ['place' 'great' 'best' 'amazing' 'experience'],
 ['place' 'work' 'best' 'company' 'experience'],
 ['place' 'work' 'company' 'best' 'great'],
 ['google' 'place' 'company' 'great' 'best'],
 ['google' 'company' 'place' 'work' 'people'],
 ['google' 'company' 'work' 'experience' 'best'],
 ['company' 'google' 'experience' 'work' 'place'],
 ['experience' 'google' 'place' 'company' 'amazing'],
 ['company' 'experience' 'work' 'google' 'people'],
 ['culture' 'place' 'experience' 'best' 'company'],
 ['culture' 'company' 'experience' 'work' 'best'],
 ['great' 'experience' 'best' 'amazing' 'nice'],
 ['experience' 'work' 'great' 'best' 'amazing'],
 ['experience' 'best' 'amazing' 'good' 'great'],
 ['experience' 'place' 'work' 'company' 'best'],
 ['work' 'company' 'best' 'experience' 'amazing'],
 ['place' 'company' 'experience' 'work' 'best']]
```

ภาพที่ 4.56 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ  
ในกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ของข้อมูลประเภท headline ของบริษัท google

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏซึ่งแสดงเป็นคำ 5 อันดับแรก ในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท headline ของบริษัท google ซึ่งมีอยู่ 18 หัวข้อ และนี่คือผลลัพธ์จากการนับความถี่เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
Word frequencies:
place: 11
best: 14
good: 2
great: 7
company: 14
amazing: 6
experience: 14
work: 11
google: 6
people: 2
culture: 2
nice: 1
```

ภาพที่ 4.57 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท google

นำคำและความถี่ที่นับได้ในกลุ่มหัวข้อที่ 1 ซึ่งแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อของข้อมูลประเภท headline ของบริษัท google ไปแสดงผลผ่าน word cloud

Group topics 1 – google\_headline top 5 words each topics



ภาพที่ 4.58 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท headline ของบริษัท google

ผู้วิจัยพิจารณาจาก word cloud แสดงเฉพาะคำ 5 อันดับแรกปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ของข้อมูลประเภท headline ของบริษัท google ผู้วิจัยมีข้อสันนิษฐานว่าในกลุ่มหัวข้อที่ 1 ของ google ข้อความรีวิวประเภท headline

มีความเกี่ยวข้องกับ: ชื่อบริษัท ประสบการณ์เยี่ยม สถานที่ดี

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท headline ของบริษัท apple ซึ่งมีอยู่ 15 หัวข้อ

```
[['good' 'great' 'nice' 'amazing' 'work' 'best' 'experience' 'company'
'people' 'place' 'culture' 'google']
['company' 'good' 'best' 'nice' 'great' 'experience' 'work' 'amazing'
'place' 'culture' 'people' 'google']
['work' 'experience' 'company' 'best' 'amazing' 'place' 'great' 'good'
'people' 'culture' 'nice' 'google']
['google' 'work' 'experience' 'great' 'company' 'best' 'good' 'place'
'amazing' 'nice' 'culture' 'people']
['experience' 'best' 'good' 'great' 'work' 'place' 'amazing' 'nice'
'company' 'culture' 'people' 'google']
['good' 'great' 'best' 'amazing' 'nice' 'work' 'experience' 'company'
'place' 'people' 'culture' 'google']
['good' 'great' 'nice' 'amazing' 'best' 'work' 'experience' 'company'
'place' 'people' 'culture' 'google']
['good' 'great' 'nice' 'amazing' 'best' 'work' 'company' 'experience'
'people' 'culture' 'place' 'google']
['best' 'amazing' 'great' 'good' 'nice' 'work' 'experience' 'place'
'company' 'google' 'culture' 'people']
['amazing' 'great' 'best' 'nice' 'good' 'work' 'experience' 'company'
'place' 'google' 'culture' 'people']
['amazing' 'great' 'nice' 'good' 'work' 'best' 'company' 'experience'
'place' 'google' 'people' 'culture']
['nice' 'great' 'good' 'amazing' 'best' 'work' 'experience' 'place'
'company' 'people' 'google' 'culture']
['great' 'nice' 'amazing' 'good' 'best' 'work' 'experience' 'place'
'company' 'google' 'people' 'culture']
['great' 'good' 'amazing' 'nice' 'best' 'work' 'experience' 'company'
'place' 'people' 'google' 'culture']
['great' 'good' 'amazing' 'nice' 'best' 'work' 'experience' 'company'
'place' 'google' 'people' 'culture']]
```

ภาพที่ 4.59 แสดงผลลัพธ์แต่ละหัวข้อในกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท google

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท headline ของบริษัท google ซึ่งมีอยู่ 15 หัวข้อ

และนี่คือผลลัพธ์จากการนับความถี่ของคำทุกคำที่ปรากฏในแต่ละหัวข้อ

```
Word frequencies:
good: 15
great: 15
nice: 15
amazing: 15
work: 15
best: 15
experience: 15
company: 15
people: 15
place: 15
culture: 15
google: 15
```

ภาพที่ 4.60 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท google

คำและความถี่ที่นับได้ในกลุ่มที่ 2 ในข้อมูลประเภท headline ของบริษัท google ไปแสดงผลผ่าน word cloud

Group topics 2 – google\_headline all words each topics



ภาพที่ 4.61 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2  
ในข้อมูลประเภท headline ของบริษัท google

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 2 ของข้อมูลประเภท headline ของบริษัท google  
ซึ่งมีอยู่ 15 หัวข้อแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
[['good' 'great' 'nice' 'amazing' 'work']
['company' 'good' 'best' 'nice' 'great']
['work' 'experience' 'company' 'best' 'amazing']
['google' 'work' 'experience' 'great' 'company']
['experience' 'best' 'good' 'great' 'work']
['good' 'great' 'best' 'amazing' 'nice']
['good' 'great' 'nice' 'amazing' 'best']
['good' 'great' 'nice' 'amazing' 'best']
['best' 'amazing' 'great' 'good' 'nice']
['amazing' 'great' 'best' 'nice' 'good']
['amazing' 'great' 'nice' 'good' 'work']
['nice' 'great' 'good' 'amazing' 'best']
['great' 'nice' 'amazing' 'good' 'best']
['great' 'good' 'amazing' 'nice' 'best']
['great' 'good' 'amazing' 'nice' 'best']]
```

ภาพที่ 4.62 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ  
ในกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ของข้อมูลประเภท headline ของบริษัท google

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏซึ่งแสดงเป็นคำ 5 อันดับแรก  
ในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท headline ของบริษัท google ซึ่งมีอยู่ 15 หัวข้อ  
และนี่คือผลลัพธ์จากการนับความถี่เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
Word frequencies:
good: 13
great: 14
nice: 12
amazing: 12
work: 5
company: 3
best: 12
experience: 3
google: 1
```

ภาพที่ 4.63 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท google

นำคำและความถี่ที่นับได้ในกลุ่มหัวข้อที่ 2 ซึ่งแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อของข้อมูลประเภท headline ของบริษัท google ไปแสดงผลผ่าน word cloud

Group topics 2 – google\_headline top 5 words each topics



ภาพที่ 4.64 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท headline ของบริษัท google

ผู้วิจัยพิจารณาจาก word cloud แสดงเฉพาะคำ 5 อันดับแรกปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ของข้อมูลประเภท headline ของบริษัท google ผู้วิจัยมีข้อสันนิษฐานว่าในกลุ่มหัวข้อที่ 2 ของ google ข้อความรีวิวประเภท headline มีความเกี่ยวข้องกับ: ชื่อบริษัท ประสบการณ์เยี่ยม

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท headline ของบริษัท google ซึ่งมีอยู่ 7 หัวข้อ

```
[['company' 'best' 'place' 'experience' 'google' 'work' 'amazing' 'great'
 'culture' 'people' 'good' 'nice']
 ['company' 'work' 'best' 'experience' 'place' 'google' 'people'
 'culture' 'amazing' 'great' 'good' 'nice']
 ['company' 'work' 'best' 'good' 'great' 'experience' 'people' 'place'
 'amazing' 'google' 'culture' 'nice']
 ['company' 'good' 'best' 'great' 'amazing' 'work' 'nice' 'people'
 'experience' 'place' 'google' 'culture']
 ['company' 'work' 'experience' 'great' 'best' 'good' 'google' 'place'
 'amazing' 'culture' 'nice' 'people']
 ['company' 'work' 'best' 'great' 'people' 'place' 'experience' 'amazing'
 'culture' 'good' 'nice' 'google']
 ['company' 'great' 'amazing' 'best' 'work' 'experience' 'good' 'google'
 'nice' 'place' 'people' 'culture']]
```

ภาพที่ 4.65 แสดงผลลัพธ์แต่ละหัวข้อในกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท google

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท headline ของบริษัท google ซึ่งมีอยู่ 7 หัวข้อ

และนี่คือผลลัพธ์จากการนับความถี่ของคำที่ปรากฏในแต่ละหัวข้อ

```
Word frequencies:
company: 7
best: 7
place: 7
experience: 7
google: 7
work: 7
amazing: 7
great: 7
culture: 7
people: 7
good: 7
nice: 7
```

ภาพที่ 4.66 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท google

นำคำและความถี่ที่นับได้ในกลุ่มที่ 3 ในข้อมูลประเภท headline ของบริษัท google ไปแสดงผลผ่าน word cloud

Group topics 3 – google\_headline all words each topics



ภาพที่ 4.67 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท headline ของบริษัท google

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 3 ของข้อมูลประเภท headline ของบริษัท google ซึ่งมีอยู่ 7 หัวข้อแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
[[ 'company' 'best' 'place' 'experience' 'google']
 ['company' 'work' 'best' 'experience' 'place']
 ['company' 'work' 'best' 'good' 'great']
 ['company' 'good' 'best' 'great' 'amazing']
 ['company' 'work' 'experience' 'great' 'best']
 ['company' 'work' 'best' 'great' 'people']
 ['company' 'great' 'amazing' 'best' 'work']]
```

ภาพที่ 4.68 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อในกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ของข้อมูลประเภท headline ของบริษัท google

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏซึ่งแสดงเป็นคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท headline ของบริษัท google ซึ่งมีอยู่ 7 หัวข้อ และนี่คือผลลัพธ์จากการนับความถี่เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

Word frequencies:
company: 7
best: 7
place: 2
experience: 3
google: 1
work: 5
good: 2
great: 5
amazing: 2
people: 1

ภาพที่ 4.69 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท google

นำคำและความถี่ที่นับได้ในกลุ่มหัวข้อที่ 3 ซึ่งแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อของข้อมูลประเภท headline ของบริษัท google ไปแสดงผลผ่าน word cloud

Group topics 3 – google\_headline top 5 words each topics



ภาพที่ 4.70 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท headline ของบริษัท google

ผู้วิจัยพิจารณาจาก word cloud แสดงเฉพาะคำ 5 อันดับแรกปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ของข้อมูลประเภท headline ของบริษัท google ผู้วิจัยมีข้อสันนิษฐานว่าในกลุ่มหัวข้อที่ 3 ของ google ข้อความรีวิวประเภท headline มีความเกี่ยวข้องกับ: ชื่อบริษัท ประสบการณ์เยี่ยม สถานที่เยี่ยม

5. ผลการวิเคราะห์หาคำที่มีนัยสำคัญจากการจัดกลุ่มหัวข้อซึ่งแสดงเป็น wordcloud ในข้อมูลรีวิวประเภท pros บริษัท google

จัดกลุ่มหัวข้อให้เป็น 3 กลุ่ม

```
[ ] # reduce to group of 3
t2v_gg_p_16v2_dp_load.hierarchical_topic_reduction(num_topics=3)
[[[19, 2, 12, 4, 5, 8, 13, 0], [16, 9, 3, 7, 18, 10, 1], [6, 17, 11, 14, 15]]]
```

ภาพที่ 4.71 แสดงการจัดกลุ่มผลลัพธ์หัวข้อของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท google

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท pros ของบริษัท google ซึ่งมีอยู่ 8 หัวข้อ

```
[['culture' 'balance' 'career' 'opportunities' 'employees' 'coworkers'
 'life' 'work' 'colleagues' 'benefits' 'office' 'working' 'environment'
 'company' 'salary' 'projects' 'compensation' 'pay' 'growth' 'friendly'
 'fun' 'learn' 'team' 'place' 'best' 'smart' 'perks' 'amazing' 'food'
 'really' 'people' 'interesting' 'lot' 'high' 'excellent' 'great' 'lots'
 'nice' 'good' 'free' 'google' 'many']
['culture' 'benefits' 'opportunities' 'employees' 'coworkers'
 'colleagues' 'career' 'company' 'salary' 'compensation' 'environment'
 'pay' 'friendly' 'perks' 'work' 'growth' 'office' 'projects' 'learn'
 'working' 'life' 'place' 'fun' 'food' 'team' 'best' 'people' 'balance'
 'excellent' 'good' 'smart' 'really' 'lots' 'great' 'lot' 'amazing'
 'interesting' 'google' 'high' 'many' 'nice' 'free']
['culture' 'opportunities' 'employees' 'work' 'career' 'office'
 'coworkers' 'working' 'compensation' 'projects' 'colleagues' 'company'
 'salary' 'perks' 'pay' 'friendly' 'life' 'fun' 'environment' 'balance'
 'growth' 'smart' 'amazing' 'interesting' 'team' 'nice' 'excellent'
 'really' 'great' 'learn' 'free' 'culture' 'place' 'best' 'google'
 'lots' 'food' 'good' 'people' 'lot' 'many' 'high']
['balance' 'benefits' 'work' 'career' 'opportunities' 'compensation'
 'life' 'employees' 'working' 'office' 'coworkers' 'salary' 'pay'
 'company' 'colleagues' 'growth' 'projects' 'environment' 'fun' 'perks'
 'best' 'smart' 'friendly' 'amazing' 'lot' 'learn' 'culture' 'team'
 'place' 'great' 'excellent' 'high' 'food' 'really' 'good' 'free' 'nice'
 'lots' 'interesting' 'people' 'google' 'many']
['food' 'benefits' 'opportunities' 'employees' 'perks' 'company' 'place'
 'pay' 'coworkers' 'friendly' 'culture' 'projects' 'work' 'career'
 'free' 'life' 'office' 'environment' 'colleagues' 'working' 'growth'
 'fun' 'salary' 'people' 'compensation' 'learn' 'best' 'smart' 'great'
 'excellent' 'balance' 'lots' 'good' 'high' 'google' 'amazing' 'lot'
 'team' 'many' 'interesting' 'really' 'nice']
['perks' 'benefits' 'compensation' 'employees' 'opportunities' 'pay'
 'salary' 'career' 'work' 'coworkers' 'colleagues' 'projects' 'working'
 'company' 'office' 'great' 'friendly' 'smart' 'growth' 'best' 'learn'
 'excellent' 'people' 'nice' 'balance' 'good' 'life' 'lot' 'amazing'
 'free' 'really' 'environment' 'fun' 'place' 'lots' 'culture' 'food'
 'team' 'google' 'interesting' 'many' 'high']
['employees' 'colleagues' 'coworkers' 'company' 'benefits'
 'compensation' 'office' 'working' 'work' 'career' 'opportunities'
 'salary' 'pay' 'people' 'perks' 'growth' 'projects' 'balance' 'life'
 'smart' 'pest' 'friendly' 'culture' 'team' 'great' 'environment' 'fun'
 'food' 'amazing' 'learn' 'really' 'excellent' 'high' 'place' 'lots'
 'good' 'many' 'google' 'lot' 'interesting' 'nice' 'free']
['benefits' 'salary' 'compensation' 'opportunities' 'pay' 'career'
 'employees' 'work' 'company' 'working' 'coworkers' 'office' 'perks'
 'colleagues' 'projects' 'growth' 'friendly' 'life' 'learn' 'good' 'fun'
 'excellent' 'smart' 'environment' 'best' 'culture' 'food' 'lot' 'great'
 'balance' 'place' 'free' 'lots' 'nice' 'really' 'amazing' 'high'
 'interesting' 'team' 'people' 'google' 'many']]
```

ภาพที่ 4.72 แสดงผลลัพธ์แต่ละหัวข้อในกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท google

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ด้วยความจากคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท pros ของบริษัท google ซึ่งมีอยู่ 8 หัวข้อ

และนี่คือผลลัพธ์จากการนับความถี่ของคำทุกคำที่ปรากฏในแต่ละหัวข้อ

```
Word frequencies: smart: 8
culture: 8      perks: 8
balance: 8      amazing: 8
career: 8       food: 8
opportunities: 8 really: 8
employees: 8    people: 8
coworkers: 8    interesting: 8
life: 8          lot: 8
work: 8          high: 8
colleagues: 8   excellent: 8
benefits: 8     great: 8
office: 8        lots: 8
working: 8       nice: 8
environment: 8  good: 8
company: 8       free: 8
salary: 8        google: 8
projects: 8      many: 8
compensation: 8
pay: 8
growth: 8
friendly: 8
fun: 8
learn: 8
team: 8
place: 8
best: 8
```

ภาพที่ 4.73 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท google

คำและความถี่ที่นับได้ในกลุ่มที่ 2 ในข้อมูลประเภท pros ของบริษัท google ไปแสดงผลผ่าน word cloud

Group topics 1 – google\_pros all words each topics



ภาพที่ 4.74 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1  
ในข้อมูลประเภท pros ของบริษัท google

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 1 ของข้อมูลประเภท pros ของบริษัท google ซึ่งมีอยู่ 8 หัวข้อแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
[['culture' 'balance' 'career' 'opportunities' 'employees']
 ['culture' 'benefits' 'opportunities' 'employees' 'coworkers']
 ['benefits' 'opportunities' 'employees' 'work' 'career']
 ['balance' 'benefits' 'work' 'career' 'opportunities']
 ['food' 'benefits' 'opportunities' 'employees' 'perks']
 ['perks' 'benefits' 'compensation' 'employees' 'opportunities']
 ['employees' 'colleagues' 'coworkers' 'company' 'benefits']
 ['benefits' 'salary' 'compensation' 'opportunities' 'pay']]
```

ภาพที่ 4.75 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ  
ในกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ของข้อมูลประเภท pros ของบริษัท google

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏซึ่งแสดงเป็นคำ 5 อันดับแรกอันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท pros ของบริษัท google ซึ่งมีอยู่ 8 หัวข้อ และนี่คือผลลัพธ์จากการนับความถี่เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
Word frequencies:
culture: 2
balance: 2
career: 3
opportunities: 7
employees: 6
benefits: 7
coworkers: 2
work: 2
food: 1
perks: 2
compensation: 2
colleagues: 1
company: 1
salary: 1
pay: 1
```

ภาพที่ 4.76 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท google

นำคำและความถี่ที่นับได้ในกลุ่มหัวข้อที่ 1 ซึ่งแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อของข้อมูลประเภท pros ของบริษัท google ไปแสดงผลผ่าน word cloud

Group topics 1 – google\_pros top 5 words each topics



ภาพที่ 4.77 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูล  
ประเภท pros ของบริษัท google

ผู้วิจัยพิจารณาจาก word cloud แสดงเฉพาะคำ 5 อันดับแรกปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ของข้อมูลประเภท pros ของบริษัท google ผู้วิจัยมีข้อสันนิษฐานว่าในกลุ่มหัวข้อที่ 1 ของ google ข้อความรีวิวประเภท pros

มีความเกี่ยวข้องกับ: โอกาสของพนักงาน ผลประโยชน์สิทธิพิเศษ เพื่อร่วมงาน วัฒนธรรม  
ความสมดุล

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท pros ของบริษัท apple ซึ่งมีอยู่ 7 หัวข้อ

```
[['place' 'learn' 'opportunities' 'projects' 'environment' 'career'
'office' 'fun' 'company' 'friendly' 'coworkers' 'work' 'culture'
'working' 'best' 'life' 'benefits' 'google' 'growth' 'smart' 'great'
'lots' 'colleagues' 'perks' 'food' 'excellent' 'pay' 'employees'
'amazing' 'lot' 'good' 'many' 'high' 'free' 'team' 'interesting' 'nice'
'people' 'really' 'salary' 'compensation' 'balance']
['place' 'working' 'work' 'office' 'projects' 'company' 'coworkers'
'best' 'opportunities' 'employees' 'career' 'friendly' 'excellent'
'great' 'fun' 'colleagues' 'amazing' 'good' 'life' 'team' 'nice'
'environment' 'pay' 'google' 'benefits' 'lot' 'lots' 'learn' 'people'
'smart' 'perks' 'high' 'culture' 'compensation' 'food' 'many' 'free'
'really' 'interesting' 'growth' 'salary' 'balance']
['environment' 'employees' 'office' 'colleagues' 'coworkers'
'opportunities' 'projects' 'working' 'work' 'friendly' 'company'
'career' 'benefits' 'place' 'life' 'culture' 'people' 'best' 'fun'
'growth' 'smart' 'team' 'balance' 'learn' 'excellent' 'amazing' 'perks'
'compensation' 'food' 'great' 'good' 'pay' 'salary' 'high' 'free'
'nice' 'really' 'lot' 'google' 'lots' 'many' 'interesting']
['google' 'company' 'opportunities' 'projects' 'working' 'career' 'work'
'employees' 'benefits' 'smart' 'office' 'coworkers' 'colleagues'
'perks' 'learn' 'environment' 'growth' 'best' 'free' 'pay' 'place'
'life' 'lot' 'friendly' 'fun' 'great' 'compensation' 'amazing' 'really'
'many' 'people' 'excellent' 'salary' 'lots' 'interesting' 'balance'
'team' 'food' 'good' 'culture' 'nice' 'high']
['benefits' 'smart' 'compensation' 'opportunities' 'pay' 'employees'
'perks' 'salary' 'career' 'colleagues' 'office' 'work' 'coworkers'
'company' 'projects' 'working' 'free' 'friendly' 'growth' 'learn'
'life' 'people' 'great' 'balance' 'google' 'food' 'excellent' 'really'
'environment' 'interesting' 'culture' 'amazing' 'nice' 'best' 'fun'
'lot' 'good' 'lots' 'many' 'team' 'place' 'high']
['smart' 'colleagues' 'employees' 'coworkers' 'career' 'people'
'opportunities' 'work' 'office' 'projects' 'working' 'company' 'learn'
'google' 'life' 'pay' 'friendly' 'environment' 'perks' 'benefits'
'growth' 'culture' 'best' 'team' 'salary' 'food' 'really' 'many'
'balance' 'lot' 'lots' 'place' 'high' 'compensation' 'amazing' 'great'
'fun' 'excellent' 'interesting' 'free' 'good' 'nice']
['opportunities' 'career' 'projects' 'learn' 'work' 'employees'
'working' 'coworkers' 'company' 'colleagues' 'benefits' 'office'
'growth' 'life' 'environment' 'people' 'place' 'culture' 'lots' 'fun'
'lot' 'perks' 'smart' 'many' 'google' 'pay' 'excellent' 'team' 'food'
'friendly' 'best' 'amazing' 'compensation' 'great' 'really'
'interesting' 'good' 'balance' 'salary' 'high' 'free' 'nice']]
```

ภาพที่ 4.78 แสดงผลลัพธ์แต่ละหัวข้อในกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท google

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ต่อความจากคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท pros ของบริษัท google ซึ่งมีอยู่ 7 หัวข้อ

และนี่คือผลลัพธ์จากการนับความถี่ของคำทุกคำที่ปรากฏในแต่ละหัวข้อ

```
Word frequencies: food: 7
place: 7           excellent: 7
learn: 7            pay: 7
opportunities: 7   employees: 7
projects: 7          amazing: 7
environment: 7      lot: 7
career: 7            good: 7
office: 7            many: 7
fun: 7               high: 7
company: 7           free: 7
friendly: 7          team: 7
coworkers: 7         interesting: 7
work: 7               nice: 7
culture: 7            people: 7
working: 7            really: 7
best: 7               salary: 7
life: 7                compensation: 7
benefits: 7           balance: 7
google: 7
growth: 7
smart: 7
great: 7
lots: 7
colleagues: 7
perks: 7
```

ภาพที่ 4.79 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท google

นำคำและความถี่ที่นับได้ในกลุ่มที่ 2 ในข้อมูลประเภท pros ของบริษัท google ไปแสดงผลผ่าน word cloud

Group topics 2 – google\_pros all words each topics



ภาพที่ 4.80 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท pros ของบริษัท google

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 2 ของข้อมูลประเภท pros ของบริษัท google ซึ่งมีอยู่ 7 หัวข้อแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
[['place' 'learn' 'opportunities' 'projects' 'environment']
 ['place' 'working' 'work' 'office' 'projects']
 ['environment' 'employees' 'office' 'colleagues' 'coworkers']
 ['google' 'company' 'opportunities' 'projects' 'working']
 ['benefits' 'smart' 'compensation' 'opportunities' 'pay']
 ['smart' 'colleagues' 'employees' 'coworkers' 'career']
 ['opportunities' 'career' 'projects' 'learn' 'work']]
```

ภาพที่ 4.81 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อในกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ของข้อมูลประเภท pros ของบริษัท google

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏซึ่งแสดงเป็นคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท pros ของบริษัท google ซึ่งมีอยู่ 7 หัวข้อ และนี่คือผลลัพธ์จากการนับความถี่เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
Word frequencies:
place: 2
learn: 2
opportunities: 4
projects: 4
environment: 2
working: 2
work: 2
office: 2
employees: 2
colleagues: 2
coworkers: 2
google: 1
company: 1
benefits: 1
smart: 2
compensation: 1
pay: 1
career: 2
```

ภาพที่ 4.82 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท google

คำและความถี่ที่นับได้ในกลุ่มหัวข้อที่ 2 ซึ่งแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อของข้อมูลประเภท pros ของบริษัท google ไปแสดงผลผ่าน word cloud

### Group topics 2 – google\_pros top 5 words each topics



ภาพที่ 4.83 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท pros ของบริษัท google

ผู้วิจัยพิจารณาจาก word cloud แสดงเฉพาะคำ 5 อันดับแรกปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ของข้อมูลประเภท pros ของบริษัท Google ผู้วิจัยมีข้อสันนิษฐานว่าในกลุ่มหัวข้อที่ 2 ของ Google ข้อความรีวิวประเภท pros มีความเกี่ยวข้องกับ: โอกาสของพนักงาน สภาพแวดล้อม โครงการต่างๆ

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท pros ของบริษัท google ซึ่งมีอยู่ 5 หัวข้อ

```
[[['good', 'great', 'best', 'excellent', 'amazing', 'friendly', 'nice', 'lot',
  'fun', 'working', 'life', 'lots', 'work', 'high', 'smart', 'learn', 'many',
  'really', 'perks', 'career', 'place', 'food', 'opportunities', 'team',
  'interesting', 'environment', 'projects', 'free', 'company', 'culture', 'pay',
  'balance', 'benefits', 'growth', 'office', 'people', 'colleagues', 'google',
  'compensation', 'employees', 'coworkers', 'salary'],
 ['company', 'excellent', 'good', 'best', 'employees', 'working', 'great',
  'work', 'team', 'amazing', 'friendly', 'colleagues', 'office', 'fun', 'nice',
  'life', 'opportunities', 'coworkers', 'career', 'high', 'growth', 'smart',
  'projects', 'people', 'compensation', 'learn', 'benefits', 'really', 'place',
  'lot', 'pay', 'perks', 'many', 'google', 'culture', 'balance', 'environment',
  'free', 'food', 'lots', 'salary', 'interesting'],
 ['company', 'working', 'work', 'employees', 'office', 'career', 'best',
  'excellent', 'great', 'projects', 'colleagues', 'coworkers', 'amazing',
  'team', 'good', 'opportunities', 'compensation', 'benefits', 'fun', 'pay',
  'friendly', 'life', 'growth', 'nice', 'salary', 'smart', 'learn', 'lot',
  'place', 'perks', 'high', 'people', 'google', 'balance', 'really',
  'environment', 'free', 'culture', 'lots', 'food', 'many', 'interesting'],
 ['company', 'opportunities', 'growth', 'career', 'employees', 'office',
  'projects', 'benefits', 'learn', 'colleagues', 'compensation', 'environment',
  'work', 'working', 'life', 'salary', 'google', 'smart', 'coworkers', 'balance',
  'culture', 'team', 'perks', 'pay', 'excellent', 'great', 'place', 'best',
  'amazing', 'many', 'lot', 'good', 'fun', 'people', 'food', 'high', 'really',
  'lots', 'friendly', 'free', 'interesting', 'nice'],
 ['company', 'best', 'employees', 'office', 'working', 'opportunities',
  'career', 'work', 'colleagues', 'coworkers', 'projects', 'team', 'place',
  'life', 'google', 'growth', 'benefits', 'amazing', 'smart', 'environment',
  'culture', 'great', 'friendly', 'excellent', 'compensation', 'salary',
  'really', 'people', 'perks', 'pay', 'fun', 'learn', 'high', 'lot', 'balance',
  'good', 'food', 'nice', 'many', 'lots', 'free', 'interesting']]
```

ภาพที่ 4.84 แสดงผลลัพธ์แต่ละหัวข้อในกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท google

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท pros ของบริษัท google ซึ่งมีอยู่ 5 หัวข้อ

และนี่คือผลลัพธ์จากการนับความถี่ของคำทุกคำที่ปรากฏในแต่ละหัวข้อ

Word frequencies:	
team:	5
good:	5
great:	5
best:	5
excellent:	5
amazing:	5
friendly:	5
nice:	5
lot:	5
fun:	5
working:	5
life:	5
lots:	5
work:	5
high:	5
smart:	5
learn:	5
many:	5
really:	5
perks:	5
career:	5
place:	5
food:	5
opportunities:	5
interesting:	5
environment:	5
projects:	5
free:	5
company:	5
culture:	5
pay:	5
balance:	5
benefits:	5
growth:	5
office:	5
people:	5
colleagues:	5
google:	5
compensation:	5
employees:	5
coworkers:	5
salary:	5

ภาพที่ 4.85 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ของเทคนิค

Top2vec ในข้อมูลประเภท pros ของบริษัท google

นำคำและความถี่ที่นับได้ในกลุ่มที่ 3 ในข้อมูลประเภท pros ของบริษัท google ไปแสดงผลผ่าน word cloud

### Group topics 3 – google\_pros all words each topics



ภาพที่ 4.86 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท pros ของบริษัท Google

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 3 ของข้อมูลประเภท pros ของบริษัท Google ซึ่งมีอยู่ 5 หัวข้อแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
[[['good', 'great', 'best', 'excellent', 'amazing'],
  ['company', 'excellent', 'good', 'best', 'employees'],
  ['company', 'working', 'work', 'employees', 'office'],
  ['company', 'opportunities', 'growth', 'career', 'employees'],
  ['company', 'best', 'employees', 'office', 'working']]]
```

ภาพที่ 4.87 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ  
ในกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ของข้อมูลประเภท pros ของบริษัท Google

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏซึ่งแสดงเป็นคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท pros ของบริษัท Google ซึ่งมีอยู่ 5 หัวข้อ และนี่คือผลลัพธ์จากการนับความถี่เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
Word frequencies:
good: 2
great: 1
best: 3
excellent: 2
amazing: 1
company: 4
employees: 4
working: 2
work: 1
office: 2
opportunities: 1
growth: 1
career: 1
```

ภาพที่ 4.88 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท Google

นำคำและความถี่ที่นับได้ในกลุ่มหัวข้อที่ 3 ซึ่งแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อของข้อมูลประเภท pros ของบริษัท google ไปแสดงผลผ่าน word cloud

Group topics 3 – google\_pros top 5 words each topics



ภาพที่ 4.89 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท pros ของบริษัท google

ผู้วิจัยพิจารณาจาก word cloud แสดงเฉพาะคำ 5 อันดับแรกปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ของข้อมูลประเภท pros ของบริษัท google ผู้วิจัยมีข้อสันนิษฐานว่าในกลุ่มหัวข้อที่ 3 ของ google ข้อความรีวิวประเภท pros มีความเกี่ยวข้องกับ: ชื่อบริษัท พนักงานดี

6. ผลการวิเคราะห์หาคำที่มีนัยสำคัญจากการจัดกลุ่มหัวข้อซึ่งแสดงเป็น wordcloud ในข้อมูลรีวิวประเภท cons บริษัท google

จัดกลุ่มหัวข้อให้เป็น 3 กลุ่ม

```
[ ] # reduce to group of 3
t2v_gg_c_16v2_dp_load.hierarchical_topic_reduction(num_topics=3)

[[[18, 17, 13, 16, 1, 9, 6, 7, 12, 11, 3],
  [23, 2, 21, 29, 22, 25, 28, 0],
  [15, 5, 19, 24, 26, 27, 20, 8, 10, 14, 4]]]
```

ภาพที่ 4.90 แสดงการจัดกลุ่มผลลัพธ์หัวข้อของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท google

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท cons ของบริษัท google ซึ่งมีอยู่ 11 หัวข้อ

```
[['google' 'cons' 'job' 'company' 'working' 'work' 'much' 'bad' 'really'
'slow' 'like' 'nothing' 'great' 'sometimes' 'lot' 'dont' 'good'
'management' 'life' 'many' 'difficult' 'none' 'career' 'big' 'growth'
'politics' 'get' 'hard' 'time' 'hours' 'say' 'think' 'large' 'people'
'balance' 'long' 'team' 'process' 'culture' 'high']
['google' 'company' 'management' 'growth' 'job' 'career' 'work'
'working' 'process' 'difficult' 'politics' 'balance' 'slow' 'much'
'lot' 'culture' 'get' 'really' 'people' 'many' 'life' 'hard' 'like'
'hours' 'sometimes' 'great' 'big' 'nothing' 'none' 'dont' 'time'
'large' 'good' 'think' 'team' 'cons' 'say' 'long' 'high' 'bad']
['job' 'career' 'work' 'company' 'working' 'management' 'growth' 'much'
'balance' 'good' 'lot' 'like' 'get' 'really' 'none' 'say' 'nothing'
'bad' 'hours' 'cons' 'high' 'culture' 'life' 'think' 'sometimes'
'great' 'difficult' 'dont' 'politics' 'slow' 'people' 'time' 'google'
'hard' 'team' 'many' 'large' 'big' 'long' 'process']
['company' 'management' 'job' 'working' 'work' 'career' 'team' 'life'
'google' 'growth' 'bad' 'great' 'good' 'people' 'really' 'sometimes'
'hours' 'like' 'culture' 'say' 'nothing' 'lot' 'none' 'many' 'balance'
'slow' 'time' 'politics' 'get' 'big' 'difficult' 'dont' 'cons' 'think'
'process' 'hard' 'much' 'high' 'large' 'long']
['company' 'growth' 'management' 'large' 'big' 'job' 'career' 'work'
'working' 'team' 'people' 'balance' 'google' 'lot' 'culture' 'many'
'politics' 'slow' 'process' 'like' 'sometimes' 'difficult' 'life'
'long' 'hard' 'get' 'high' 'say' 'really' 'great' 'dont' 'none' 'time'
'think' 'nothing' 'hours' 'much' 'good' 'cons' 'bad']
['management' 'politics' 'growth' 'career' 'company' 'process' 'culture'
'job' 'difficult' 'work' 'slow' 'people' 'working' 'balance' 'like'
'hard' 'google' 'life' 'large' 'really' 'big' 'lot' 'sometimes' 'time'
'cons' 'team' 'good' 'many' 'think' 'bad' 'high' 'say' 'none' 'long'
'nothing' 'great' 'much' 'hours' 'get' 'dont']
['politics' 'management' 'career' 'company' 'culture' 'job' 'work'
'difficult' 'grow' 'working' 'really' 'people' 'process' 'hard'
'team' 'balance' 'life' 'sometimes' 'get' 'time' 'think' 'google'
'many' 'high' 'nothing' 'like' 'slow' 'cons' 'bad' 'good' 'lot' 'none'
'say' 'dont' 'hours' 'large' 'much' 'big' 'long' 'great']
['career' 'growth' 'management' 'job' 'work' 'company' 'working' 'life'
'politics' 'slow' 'difficult' 'hard' 'process' 'get' 'really' 'culture'
'time' 'hours' 'balance' 'think' 'google' 'lot' 'long' 'like' 'many'
'say' 'bad' 'high' 'much' 'nothing' 'team' 'good' 'sometimes' 'none'
'dont' 'large' 'great' 'cons' 'big' 'people']
['management' 'career' 'job' 'get' 'work' 'company' 'growth' 'working'
'slow' 'politics' 'difficult' 'process' 'hard' 'good' 'really' 'time'
'hours' 'high' 'like' 'culture' 'life' 'say' 'team' 'dont' 'bad'
'great' 'long' 'much' 'sometimes' 'nothing' 'think' 'people' 'balance'
'cons' 'lot' 'google' 'big' 'many' 'none' 'large']
['culture' 'management' 'politics' 'company' 'career' 'growth' 'people'
'difficult' 'balance' 'job' 'work' 'life' 'really' 'many' 'working'
'like' 'hard' 'team' 'sometimes' 'cons' 'think' 'process' 'get'
'nothing' 'good' 'bad' 'say' 'dont' 'google' 'none' 'lot' 'hours'
'time' 'great' 'much' 'high' 'slow' 'large' 'big' 'long']
['career' 'management' 'company' 'growth' 'job' 'work' 'politics'
'culture' 'working' 'life' 'balance' 'google' 'get' 'process' 'like'
'difficult' 'say' 'team' 'think' 'really' 'nothing' 'many' 'people'
'hard' 'none' 'hours' 'dont' 'time' 'lot' 'sometimes' 'much' 'good'
'high' 'cons' 'slow' 'bad' 'great' 'long' 'big' 'large']]
```

ภาพที่ 4.91 แสดงผลลัพธ์แต่ละหัวข้อในกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท google

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ดีความจากคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท cons ของบริษัท google ซึ่งมีอยู่ 11 หัวข้อ

และนี่คือผลลัพธ์จากการนับความถี่ของคำทุกคำที่ปรากฏในแต่ละหัวข้อ



ภาพที่ 4.92 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท Google

นำคำและความถี่ที่นับได้ในกลุ่มที่ 1 ในข้อมูลประเภท cons ของบริษัท google ไปแสดงผลผ่าน word cloud

Group topics 1 – google cons all words each topics



ภาพที่ 4.93 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท cons ของบริษัท google

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 1 ของข้อมูลประเภท cons ของบริษัท Google ซึ่งมีอยู่ 11 หัวข้อแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
[['google' 'cons' 'job' 'company' 'working']
 ['google' 'company' 'management' 'growth' 'job']
 ['job' 'career' 'work' 'company' 'working']
 ['company' 'management' 'job' 'working' 'work']
 ['company' 'growth' 'management' 'large' 'big']
 ['management' 'politics' 'growth' 'career' 'company']
 ['politics' 'management' 'career' 'company' 'culture']
 ['career' 'growth' 'management' 'job' 'work']
 ['management' 'career' 'job' 'get' 'work']
 ['culture' 'management' 'politics' 'company' 'career']
 ['career' 'management' 'company' 'growth' 'job']]
```

ภาพที่ 4.94 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ  
ในกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ของข้อมูลประเภท cons ของบริษัท google

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏซึ่งแสดงเป็นคำ 5 อันดับแรก  
ในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท cons ของบริษัท google ซึ่งมีอยู่ 11 หัวข้อ  
และนี่คือผลลัพธ์จากการนับความถี่เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
Word frequencies:
google: 2
cons: 1
job: 7
company: 9
working: 3
management: 9
growth: 5
career: 7
work: 4
large: 1
big: 1
politics: 3
culture: 2
get: 1
```

ภาพที่ 4.95 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1  
ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท google

นำคำและความถี่ที่นับได้ในกลุ่มหัวข้อที่ 1 ซึ่งแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อของข้อมูล  
ประเภท cons ของบริษัท google ไปแสดงผลผ่าน word cloud

Group topics 1 – google\_cons top 5 words each topics



ภาพที่ 4.96 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูล  
ประเภท cons ของบริษัท google

ผู้วิจัยพิจารณาจาก word cloud แสดงเฉพาะคำ 5 อันดับแรกปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ของข้อมูลประเภท cons ของบริษัท google ผู้วิจัยมีข้อสันนิษฐานว่าในกลุ่มหัวข้อที่ 1 ของ google ข้อความรีวิวประเภท cons

มีความเกี่ยวข้องกับ: เส้นทางอาชีพ การจัดการของบริษัท การเมืองภายใน

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท cons ของบริษัท google ซึ่งมีอยู่ 8 หัวข้อ

```
[[['cons', 'company', 'management', 'job', 'career', 'work', 'working', 'hours',
'dont', 'team', 'life', 'culture', 'really', 'like', 'growth', 'bad',
'politics', 'great', 'balance', 'time', 'good', 'sometimes', 'google', 'say',
'much', 'many', 'get', 'difficult', 'process', 'hard', 'big', 'people', 'lot',
'think', 'high', 'none', 'large', 'slow', 'long', 'nothing'],
['cons', 'bad', 'dont', 'really', 'like', 'say', 'think', 'time', 'good',
'sometimes', 'life', 'long', 'politics', 'much', 'hard', 'difficult', 'job',
'nothing', 'slow', 'hours', 'great', 'none', 'culture', 'many', 'lot', 'high',
'big', 'get', 'work', 'team', 'career', 'working', 'balance', 'large',
'management', 'company', 'people', 'google', 'growth', 'process'],
['think', 'say', 'dont', 'really', 'lot', 'good', 'get', 'like', 'much', 'great',
'time', 'hard', 'many', 'long', 'high', 'bad', 'big', 'life', 'sometimes',
'slow', 'job', 'team', 'none', 'nothing', 'work', 'difficult', 'working',
'large', 'career', 'people', 'cons', 'hours', 'politics', 'balance', 'culture',
'process', 'company', 'management', 'growth', 'google'],
['think', 'nothing', 'dont', 'none', 'bad', 'slow', 'say', 'big', 'life', 'lot',
'get', 'much', 'time', 'high', 'job', 'long', 'like', 'work', 'really', 'hard',
'good', 'career', 'sometimes', 'many', 'cons', 'working', 'great', 'team',
'large', 'difficult', 'politics', 'people', 'hours', 'balance', 'culture',
'process', 'google', 'company', 'management', 'growth'],
['think', 'none', 'many', 'nothing', 'lot', 'dont', 'get', 'much', 'good', 'bad',
'big', 'great', 'people', 'say', 'high', 'slow', 'really', 'large', 'like',
'life', 'hard', 'long', 'sometimes', 'team', 'job', 'time', 'cons', 'working',
'work', 'career', 'hours', 'difficult', 'politics', 'process', 'google',
'culture', 'growth', 'balance', 'company', 'management'],
[['cons', 'bad', 'really', 'nothing', 'say', 'life', 'much', 'politics', 'good',
'none', 'sometimes', 'dont', 'culture', 'many', 'difficult', 'balance',
'think', 'career', 'like', 'time', 'slow', 'great', 'get', 'management', 'job',
'company', 'growth', 'lot', 'work', 'team', 'long', 'working', 'big', 'google',
'hard', 'high', 'people', 'large', 'hours', 'process'],
[['life', 'nothing', 'bad', 'great', 'cons', 'good', 'work', 'none', 'job', 'dont',
'working', 'management', 'get', 'culture', 'time', 'company', 'people', 'team',
'balance', 'slow', 'high', 'hours', 'like', 'growth', 'difficult', 'think',
'politics', 'lot', 'career', 'much', 'really', 'say', 'big', 'many', 'google',
'hard', 'large', 'long', 'process', 'sometimes'],
[['nothing', 'none', 'bad', 'good', 'life', 'much', 'slow', 'lot', 'say', 'think',
'job', 'great', 'dont', 'time', 'big', 'get', 'many', 'work', 'working',
'really', 'like', 'career', 'sometimes', 'difficult', 'long', 'high',
'politics', 'hard', 'culture', 'hours', 'team', 'large', 'cons', 'balance',
'people', 'google', 'growth', 'company', 'management', 'process']]]]
```

ภาพที่ 4.97 แสดงผลลัพธ์แต่ละหัวข้อในกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท  
cons ของบริษัท google

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท cons ของบริษัท google ซึ่งมีอยู่ 8 หัวข้อ

และนี่คือผลลัพธ์จากการนับความถี่ของคำทุกคำที่ปรากฏในแต่ละหัวข้อ

```
Word frequencies: say: 8
cons: 8           much: 8
company: 8        many: 8
management: 8     get: 8
job: 8            difficult: 8
career: 8          process: 8
work: 8            hard: 8
working: 8         big: 8
hours: 8           people: 8
dont: 8            lot: 8
team: 8            think: 8
life: 8             high: 8
culture: 8          none: 8
really: 8           large: 8
like: 8             slow: 8
growth: 8           long: 8
bad: 8              nothing: 8
politics: 8
great: 8
balance: 8
time: 8
good: 8
sometimes: 8
google: 8
```

ภาพที่ 4.98 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท google

นำคำและความถี่ที่นับได้ในกลุ่มที่ 2 ในข้อมูลประเภท cons ของบริษัท google ไปแสดงผลผ่าน word cloud

Group topics 2 – google\_cons all words each topics



ภาพที่ 4.99 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูล  
ประเภท cons ของบริษัท google

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 2 ของข้อมูลประเภท cons ของบริษัท google ซึ่งมีอยู่ 8  
หัวข้อแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
[['cons', 'company', 'management', 'job', 'career'],
 ['cons', 'bad', 'dont', 'really', 'like'],
 ['think', 'say', 'dont', 'really', 'lot'],
 ['think', 'nothing', 'dont', 'none', 'bad'],
 ['think', 'none', 'many', 'nothing', 'lot'],
 ['cons', 'bad', 'really', 'nothing', 'say'],
 ['life', 'nothing', 'bad', 'great', 'cons'],
 ['nothing', 'none', 'bad', 'good', 'life']]
```

ภาพที่ 4.100 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ  
ในกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ของข้อมูลประเภท cons ของบริษัท google

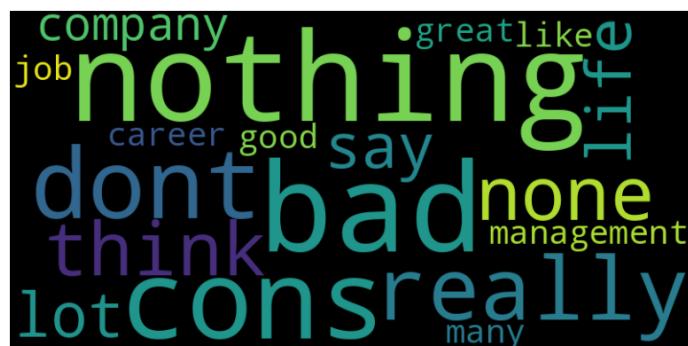
ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏซึ่งแสดงเป็นคำ 5 อันดับแรก  
ในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท cons ของบริษัท google ซึ่งมีอยู่ 8 หัวข้อ  
และนี้คือผลลัพธ์จากการนับความถี่เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
Word frequencies:
cons: 4
company: 1
management: 1
job: 1
career: 1
bad: 5
dont: 3
really: 3
like: 1
think: 3
say: 2
lot: 2
nothing: 5
none: 3
many: 1
life: 2
great: 1
good: 1
```

ภาพที่ 4.101 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท google

คำและความถี่ที่นับได้ในกลุ่มหัวข้อที่ 2 ซึ่งแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อของข้อมูลประเภท cons ของบริษัท google ไปแสดงผลผ่าน word cloud

Group topics 2 – google\_cons top 5 words each topics



ภาพที่ 4.102 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท cons ของบริษัท google

ผู้วิจัยพิจารณาจาก word cloud แสดงเฉพาะคำ 5 อันดับแรกปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ของข้อมูลประเภท cons ของบริษัท google ผู้วิจัยมีข้อสันนิษฐานว่าในกลุ่มหัวข้อที่ 2 ของ google ข้อความรีวิวประเภท cons มีความเกี่ยวข้องกับ: ไม่มีข้อเสีย ไม่คิดว่ามีข้อเสีย

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท cons ของบริษัท google ซึ่งมีอยู่ 11 หัวข้อ

```
[['balance' 'work' 'life' 'career' 'working' 'management' 'job' 'hours'
 'growth' 'company' 'time' 'politics' 'lot' 'team' 'culture' 'really'
 'get' 'process' 'high' 'sometimes' 'say' 'difficult' 'think' 'hard'
 'much' 'long' 'great' 'people' 'dont' 'nothing' 'like' 'bad' 'good'
 'cons' 'google' 'slow' 'none' 'many' 'big' 'large']
 ['hours' 'work' 'working' 'job' 'long' 'time' 'career' 'life' 'company'
 'management' 'lot' 'difficult' 'hard' 'bad' 'really' 'sometimes' 'slow'
 'get' 'great' 'cons' 'like' 'say' 'much' 'politics' 'growth' 'balance'
 'good' 'large' 'team' 'high' 'big' 'nothing' 'think' 'dont' 'many'
 'culture' 'people' 'process' 'none' 'google']
 ['slow' 'work' 'career' 'working' 'hours' 'difficult' 'management' 'job'
 'time' 'life' 'process' 'balance' 'hard' 'high' 'politics' 'people'
 'like' 'think' 'get' 'really' 'growth' 'say' 'nothing' 'sometimes'
 'lot' 'bad' 'team' 'long' 'company' 'culture' 'good' 'google' 'dont'
 'cons' 'great' 'none' 'many' 'large' 'much' 'big']
 ['work' 'high' 'job' 'working' 'process' 'management' 'hours' 'balance'
 'hard' 'company' 'difficult' 'lot' 'slow' 'life' 'growth' 'career'
 'get' 'team' 'people' 'cons' 'politics' 'nothing' 'time' 'much' 'dont'
 'sometimes' 'think' 'large' 'say' 'big' 'none' 'like' 'really' 'many'
 'culture' 'bad' 'great' 'long' 'good' 'google']
 ['management' 'career' 'job' 'work' 'company' 'team' 'working'
 'difficult' 'life' 'growth' 'balance' 'politics' 'culture' 'hard'
 'process' 'time' 'high' 'people' 'hours' 'sometimes' 'like' 'get' 'say'
 'slow' 'think' 'really' 'lot' 'good' 'dont' 'cons' 'google' 'many'
 'nothing' 'bad' 'large' 'much' 'great' 'long' 'big' 'none']
 ['time' 'hours' 'work' 'management' 'difficult' 'working' 'job' 'career'
 'hard' 'politics' 'slow' 'life' 'process' 'high' 'team' 'nothing'
 'company' 'really' 'get' 'balance' 'culture' 'cons' 'long' 'sometimes'
 'dont' 'bad' 'say' 'think' 'much' 'none' 'good' 'growth' 'people'
 'great' 'like' 'many' 'lot' 'large' 'google' 'big']
 ['work' 'working' 'job' 'life' 'career' 'hard' 'time' 'difficult'
 'management' 'hours' 'company' 'balance' 'high' 'say' 'cons' 'get'
 'sometimes' 'culture' 'politics' 'process' 'dont' 'nothing' 'team'
 'growth' 'people' 'like' 'think' 'bad' 'lot' 'slow' 'much' 'really'
 'many' 'good' 'great' 'none' 'long' 'google' 'large' 'big']
 ['slow' 'growth' 'management' 'work' 'process' 'working' 'difficult'
 'company' 'team' 'balance' 'job' 'people' 'time' 'hard' 'career'
 'sometimes' 'big' 'life' 'large' 'politics' 'nothing' 'get' 'think'
 'lot' 'long' 'really' 'bad' 'like' 'much' 'culture' 'hours' 'great'
 'say' 'good' 'google' 'many' 'dont' 'none' 'high' 'cons']
 ['job' 'working' 'work' 'company' 'management' 'career' 'cons' 'life'
 'culture' 'like' 'hours' 'great' 'difficult' 'bad' 'say' 'good'
 'people' 'hard' 'google' 'get' 'really' 'slow' 'lot' 'time' 'team'
 'sometimes' 'dont' 'politics' 'nothing' 'process' 'big' 'think' 'high'
 'growth' 'many' 'much' 'balance' 'none' 'large' 'long']
 ['work' 'difficult' 'working' 'career' 'job' 'hard' 'management'
 'google' 'politics' 'company' 'process' 'hours' 'life' 'time' 'really'
 'think' 'growth' 'team' 'get' 'like' 'lot' 'people' 'sometimes'
 'balance' 'slow' 'many' 'nothing' 'much' 'say' 'culture' 'cons' 'high'
 'none' 'great' 'good' 'large' 'long' 'bad' 'dont' 'big']
 ['work' 'working' 'job' 'career' 'hard' 'difficult' 'lot' 'hours' 'life'
 'management' 'company' 'sometimes' 'time' 'slow' 'much' 'get' 'really'
 'like' 'say' 'great' 'process' 'growth' 'think' 'team' 'many' 'nothing'
 'good' 'bad' 'balance' 'high' 'big' 'people' 'politics' 'long' 'dont'
 'google' 'large' 'none' 'cons' 'culture']]
```

ภาพที่ 4.103 แสดงผลลัพธ์แต่ละหัวข้อในกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท google

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ด้วยความจากคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท cons ของบริษัท google ซึ่งมีอยู่ 11 หัวข้อ

และนี่คือผลลัพธ์จากการนับความถี่ของคำทุกคำที่ปรากฏในแต่ละหัวข้อ

```

Word frequencies: think: 11
balance: 11      hard: 11
work: 11         much: 11
life: 11          long: 11
career: 11        great: 11
working: 11       people: 11
management: 11    dont: 11
job: 11           nothing: 11
hours: 11          like: 11
growth: 11         bad: 11
company: 11        good: 11
time: 11           cons: 11
politics: 11       google: 11
lot: 11            slow: 11
team: 11           none: 11
culture: 11        many: 11
really: 11          big: 11
get: 11            large: 11
process: 11
high: 11
sometimes: 11
say: 11
difficult: 11

```

ภาพที่ 4.104 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท google

นำคำและความถี่ที่นับได้ในกลุ่มที่ 3 ในข้อมูลประเภท cons ของบริษัท google ไปแสดงผลผ่าน word cloud

Group topics 3 – google\_cons all words each topics



ภาพที่ 4.105 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท cons ของบริษัท google

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 3 ของข้อมูลประเภท cons ของบริษัท google ซึ่งมีอยู่ 11 หัวข้อแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
[['balance' 'work' 'life' 'career' 'working']
['hours' 'work' 'working' 'job' 'long']
['slow' 'work' 'career' 'working' 'hours']
['work' 'high' 'job' 'working' 'process']
['management' 'career' 'job' 'work' 'company']
['time' 'hours' 'work' 'management' 'difficult']
['work' 'working' 'job' 'life' 'career']
['slow' 'growth' 'management' 'work' 'process']
['job' 'working' 'work' 'company' 'management']
['work' 'difficult' 'working' 'career' 'job']
['work' 'working' 'job' 'career' 'hard']]
```

ภาพที่ 4.106 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ  
ในกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ของข้อมูลประเภท cons ของบริษัท google

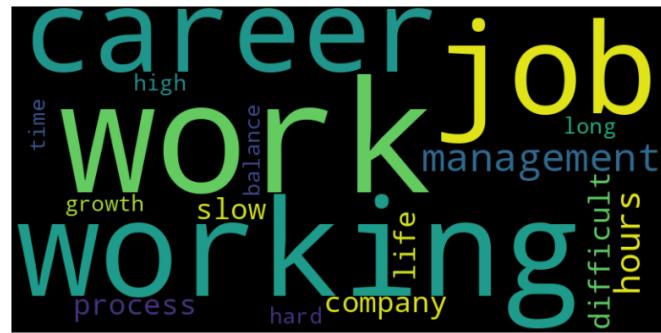
ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏซึ่งแสดงเป็นคำ 5 อันดับแรก  
ในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท cons ของบริษัท google ซึ่งมีอยู่ 11 หัวข้อ  
และนี่คือผลลัพธ์จากการนับความถี่เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
Word frequencies:
balance: 1
work: 11
life: 2
career: 6
working: 8
hours: 3
job: 7
long: 1
slow: 2
high: 1
process: 2
management: 4
company: 2
time: 1
difficult: 2
growth: 1
hard: 1
```

ภาพที่ 4.107 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3  
ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท google

นำคำและความถี่ที่นับได้ในกลุ่มหัวข้อที่ 3 ซึ่งแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อของข้อมูล  
ประเภท cons ของบริษัท google ไปแสดงผลผ่าน word cloud

Group topics 3 – google\_cons top 5 words each topics



ภาพที่ 4.108 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท cons ของบริษัท Google

ผู้วิจัยพิจารณาจาก word cloud แสดงเฉพาะคำ 5 อันดับแรกปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ของข้อมูลประเภท cons ของบริษัท Google ผู้วิจัยมีข้อสันนิษฐานว่าในกลุ่มหัวข้อที่ 3 ของ Google ข้อความรีวิวประเภท cons

มีความเกี่ยวข้องกับ: เส้นทางอาชีพ การทำงานชา งานยาก การจัดการ ชั่วโมงการทำงานสูง

7. ผลการวิเคราะห์หากำที่มีนัยสำคัญจากการจัดกลุ่มหัวข้อซึ่งแสดงเป็น wordcloud ในข้อมูลรีวิวประเภท headline บริษัท Microsoft

จัดกลุ่มหัวข้อให้เป็น 3 กลุ่ม

```
[ ] # reduce to group of 3
t2v_m_h_16v2_dp_load.hierarchical_topic_reduction(num_topics=3)

[[3, 27, 19, 32, 9, 10, 22, 17, 20, 12, 8, 15, 1],
 [24, 5, 31, 21, 2, 25, 30, 11, 18, 14, 28, 0],
 [34, 13, 33, 16, 6, 26, 7, 23, 29, 35, 4]]
```

ภาพที่ 4.109 แสดงการจัดกลุ่มผลลัพธ์หัวข้อของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท microsoft

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท headline ของบริษัท microsoft ซึ่งมีอยู่ 11 หัวข้อ

```
[['microsoft' 'company' 'experience' 'work' 'benefits' 'culture'
 'amazing' 'life' 'best' 'place' 'balance' 'excellent' 'great' 'awesome'
 'good']
[['microsoft' 'company' 'great' 'work' 'best' 'experience' 'benefits'
 'excellent' 'amazing' 'place' 'awesome' 'good' 'balance' 'culture'
 'life']]
[['microsoft' 'company' 'work' 'experience' 'benefits' 'great' 'amazing'
 'awesome' 'excellent' 'best' 'life' 'good' 'balance' 'culture' 'place']]
[['microsoft' 'company' 'work' 'experience' 'best' 'amazing' 'great'
 'place' 'balance' 'life' 'excellent' 'culture' 'good' 'awesome'
 'benefits']]
[['microsoft' 'company' 'experience' 'culture' 'best' 'great' 'life'
 'place' 'work' 'amazing' 'balance' 'excellent' 'benefits' 'awesome'
 'good']]
[['culture' 'company' 'experience' 'benefits' 'work' 'best' 'microsoft'
 'life' 'place' 'balance' 'excellent' 'great' 'amazing' 'good' 'awesome']]
[['culture' 'life' 'experience' 'benefits' 'place' 'balance' 'company'
 'best' 'good' 'work' 'microsoft' 'excellent' 'great' 'amazing'
 'awesome']]
[['culture' 'experience' 'best' 'life' 'place' 'benefits' 'great'
 'company' 'excellent' 'good' 'amazing' 'work' 'awesome' 'microsoft'
 'balance']]
[['benefits' 'experience' 'work' 'good' 'excellent' 'best' 'great' 'life'
 'awesome' 'amazing' 'company' 'balance' 'culture' 'microsoft' 'place']]
[['benefits' 'company' 'experience' 'work' 'good' 'excellent' 'great'
 'microsoft' 'best' 'life' 'amazing' 'balance' 'awesome' 'culture'
 'place']]
[['balance' 'work' 'life' 'benefits' 'company' 'best' 'experience'
 'amazing' 'microsoft' 'culture' 'great' 'awesome' 'place' 'excellent'
 'good']]
[['company' 'work' 'benefits' 'experience' 'best' 'excellent' 'great'
 'microsoft' 'good' 'place' 'amazing' 'life' 'culture' 'awesome'
 'balance']]
[['company' 'experience' 'microsoft' 'work' 'benefits' 'place' 'best'
 'life' 'culture' 'great' 'excellent' 'amazing' 'good' 'balance'
 'awesome']]
```

ภาพที่ 4.110 แสดงผลลัพธ์แต่ละหัวข้อในกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท microsoft

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท headline ของบริษัท microsoft ซึ่งมีอยู่ 13 หัวข้อ

และนี่คือผลลัพธ์จากการนับความถี่ของคำทุกคำที่ปรากฏในแต่ละหัวข้อ

```
Word frequencies:
microsoft: 13
company: 13
experience: 13
work: 13
benefits: 13
culture: 13
amazing: 13
life: 13
best: 13
place: 13
balance: 13
excellent: 13
great: 13
awesome: 13
good: 13
```

ภาพที่ 4.111 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท Microsoft

นำคำและความถี่ที่นับได้ในกลุ่มที่ 1 ในข้อมูลประเภท headline ของบริษัท microsoft ไปแสดงผลผ่าน word cloud

Group 1 – microsoft\_headline all words each topics



ภาพที่ 4.112 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท headline ของบริษัท Microsoft

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 1 ของข้อมูลประเภท headline ของบริษัท microsoft ซึ่งมีอยู่ 13 หัวข้อแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
[['microsoft' 'company' 'experience' 'work' 'benefits'],
 ['microsoft' 'company' 'great' 'work' 'best'],
 ['microsoft' 'company' 'work' 'experience' 'benefits'],
 ['microsoft' 'company' 'work' 'experience' 'best'],
 ['microsoft' 'company' 'experience' 'culture' 'best'],
 ['culture' 'company' 'experience' 'benefits' 'work'],
 ['culture' 'life' 'experience' 'benefits' 'place'],
 ['culture' 'experience' 'best' 'life' 'place'],
 ['benefits' 'experience' 'work' 'good' 'excellent'],
 ['benefits' 'company' 'experience' 'work' 'good'],
 ['balance' 'work' 'life' 'benefits' 'company'],
 ['company' 'work' 'benefits' 'experience' 'best'],
 ['company' 'experience' 'microsoft' 'work' 'benefits']]
```

ภาพที่ 4.113 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อในกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ของข้อมูลประเภท headline ของบริษัท microsoft

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏซึ่งแสดงเป็นคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท headline ของบริษัท microsoft ซึ่งมีอยู่ 13 หัวข้อ และนี้คือผลลัพธ์จากการนับความถี่เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
Word frequencies:
microsoft: 6
company: 10
experience: 11
work: 10
benefits: 9
great: 1
best: 5
culture: 4
life: 3
place: 2
good: 2
excellent: 1
balance: 1
```

ภาพที่ 4.114 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท microsoft

คำและความถี่ที่นับได้ในกลุ่มหัวข้อที่ 1 ซึ่งแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อของข้อมูลประเภท headline ของบริษัท microsoft ไปแสดงผลผ่าน word cloud

Group 1 – microsoft\_headline top 5 words each topics



ภาพที่ 4.115 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท headline ของบริษัท Microsoft

ผู้วิจัยพิจารณาจาก word cloud แสดงเฉพาะคำ 5 อันดับแรกปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ของข้อมูลประเภท headline ของบริษัท microsoft ผู้วิจัยมีข้อสันนิษฐานว่าในกลุ่มหัวข้อที่ 1 ของ microsoft ข้อความรีวิวประเภท headline มีความเกี่ยวข้องกับ: ผลประโยชน์ดีและประสบการณ์การทำงานเยี่ยม

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท headline ของบริษัท microsoft ซึ่งมีอยู่ 12 หัวข้อ

```
[['place' 'work' 'experience' 'company' 'life' 'benefits' 'culture'
'microsoft' 'best' 'amazing' 'excellent' 'balance' 'great' 'good'
'awesome']
['place' 'work' 'company' 'best' 'experience' 'life' 'excellent' 'great'
'amazing' 'awesome' 'benefits' 'good' 'culture' 'microsoft' 'balance']
['place' 'best' 'good' 'excellent' 'great' 'experience' 'company' 'life'
'work' 'amazing' 'awesome' 'culture' 'benefits' 'balance' 'microsoft']
['place' 'great' 'excellent' 'best' 'amazing' 'awesome' 'experience'
'company' 'good' 'life' 'work' 'culture' 'benefits' 'microsoft'
'balance']
['place' 'work' 'great' 'excellent' 'company' 'best' 'amazing' 'awesome'
'experience' 'good' 'life' 'benefits' 'microsoft' 'culture' 'balance']
['company' 'work' 'best' 'excellent' 'good' 'great' 'benefits' 'amazing'
'life' 'experience' 'microsoft' 'place' 'awesome' 'balance' 'culture']
['amazing' 'great' 'excellent' 'awesome' 'work' 'company' 'good' 'best'
'life' 'experience' 'place' 'benefits' 'microsoft' 'balance' 'culture']
['company' 'good' 'excellent' 'best' 'great' 'amazing' 'work' 'life'
'awesome' 'microsoft' 'experience' 'place' 'benefits' 'culture'
'balance']
['company' 'work' 'best' 'microsoft' 'experience' 'benefits' 'place'
'life' 'amazing' 'great' 'excellent' 'culture' 'awesome' 'balance'
'good']
['company' 'best' 'microsoft' 'place' 'experience' 'work' 'life'
'amazing' 'great' 'benefits' 'culture' 'excellent' 'awesome' 'good'
'balance']
['company' 'awesome' 'work' 'amazing' 'great' 'excellent' 'microsoft'
'best' 'experience' 'place' 'life' 'benefits' 'good' 'balance'
'culture']
['company' 'great' 'excellent' 'amazing' 'best' 'work' 'awesome'
'microsoft' 'experience' 'good' 'life' 'benefits' 'place' 'balance'
'culture']]
```

ภาพที่ 4.116 แสดงผลลัพธ์แต่ละหัวข้อในกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท microsoft

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ด้วยความจากคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท headline ของบริษัท microsoft ซึ่งมีอยู่ 12 หัวข้อ

และนี่คือผลลัพธ์จากการนับความถี่ของคำทุกคำที่ปรากฏในแต่ละหัวข้อ

```
Word frequencies:
place: 12
work: 12
experience: 12
company: 12
life: 12
benefits: 12
culture: 12
microsoft: 12
best: 12
amazing: 12
excellent: 12
balance: 12
great: 12
good: 12
awesome: 12
```

ภาพที่ 4.117 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท microsoft

นำคำและความถี่ที่นับได้ในกลุ่มที่ 2 ในข้อมูลประเภท headline ของบริษัท microsoft ไปแสดงผลผ่าน word cloud

Group topics 2 – microsoft\_headline all words each topics



ภาพที่ 4.118 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท headline ของบริษัท microsoft

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 2 ของข้อมูลประเภท headline ของบริษัท microsoft ซึ่งมีอยู่ 12 หัวข้อแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
[['place' 'work' 'experience' 'company' 'life'],
 ['place' 'work' 'company' 'best' 'experience'],
 ['place' 'best' 'good' 'excellent' 'great'],
 ['place' 'great' 'excellent' 'best' 'amazing'],
 ['place' 'work' 'great' 'excellent' 'company'],
 ['company' 'work' 'best' 'excellent' 'good'],
 ['amazing' 'great' 'excellent' 'awesome' 'work'],
 ['company' 'good' 'excellent' 'best' 'great'],
 ['company' 'work' 'best' 'microsoft' 'experience'],
 ['company' 'best' 'microsoft' 'place' 'experience'],
 ['company' 'awesome' 'work' 'amazing' 'great'],
 ['company' 'great' 'excellent' 'amazing' 'best']]
```

ภาพที่ 4.119 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ  
ในกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ของข้อมูลประเภท headline ของบริษัท microsoft

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏซึ่งแสดงเป็นคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท headline ของบริษัท microsoft ซึ่งมีอยู่ 12 หัวข้อและนี่คือผลลัพธ์จากการนับความถี่เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
Word frequencies:
place: 6
work: 7
experience: 4
company: 9
life: 1
best: 8
good: 3
excellent: 7
great: 7
amazing: 4
awesome: 2
microsoft: 2
```

ภาพที่ 4.120 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท Microsoft

นำคำและความถี่ที่นับได้ในกลุ่มหัวข้อที่ 2 ซึ่งแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อของข้อมูลประเภท headline ของบริษัท microsoft ไปแสดงผลผ่าน word cloud

Group topics 2 – microsoft\_headline top 5 words each topics



ภาพที่ 4.121 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท headline ของบริษัท microsoft

ผู้วิจัยพิจารณาจาก word cloud แสดงเฉพาะคำ 5 อันดับแรกปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ของข้อมูลประเภท headline ของบริษัท microsoft ผู้วิจัยมีข้อสันนิษฐานว่าในกลุ่มหัวข้อที่ 2 ของ microsoft ชื่อความรีวิวประเภท headline

มีความเกี่ยวข้องกับ : ชื่นชมบริษัท ประสบการการทำงานที่ดีเยี่ยม ชื่นชมสถานที่

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท headline ของบริษัท microsoft ซึ่งมีอยู่ 11 หัวข้อ

```
[[['excellent' 'great' 'good' 'amazing' 'awesome' 'best' 'experience'
  'work' 'company' 'place' 'life' 'balance' 'benefits' 'microsoft'
  'culture']
[['great' 'excellent' 'awesome' 'good' 'amazing' 'best' 'work'
  'experience' 'life' 'place' 'company' 'microsoft' 'balance' 'benefits'
  'culture']
[['good' 'great' 'excellent' 'awesome' 'amazing' 'best' 'work'
  'experience' 'company' 'life' 'place' 'microsoft' 'balance' 'benefits'
  'culture']
[['good' 'great' 'excellent' 'awesome' 'amazing' 'work' 'best' 'life'
  'company' 'experience' 'place' 'culture' 'microsoft' 'balance'
  'benefits']
[['good' 'excellent' 'great' 'amazing' 'awesome' 'best' 'work' 'life'
  'company' 'experience' 'benefits' 'culture' 'place' 'balance'
  'microsoft']
[['company' 'experience' 'work' 'life' 'place' 'culture' 'balance' 'best'
  'microsoft' 'good' 'benefits' 'excellent' 'great' 'awesome' 'amazing']
[['work' 'life' 'company' 'experience' 'place' 'best' 'balance'
  'benefits' 'good' 'culture' 'amazing' 'excellent' 'great' 'microsoft'
  'awesome']
[['experience' 'excellent' 'amazing' 'great' 'best' 'awesome' 'good'
  'place' 'benefits' 'work' 'life' 'culture' 'company' 'microsoft'
  'balance']
[['work' 'experience' 'excellent' 'company' 'best' 'great' 'life' 'good'
  'place' 'amazing' 'balance' 'culture' 'microsoft' 'benefits' 'awesome']
[['excellent' 'good' 'great' 'best' 'amazing' 'awesome' 'balance' 'work'
  'life' 'experience' 'benefits' 'company' 'microsoft' 'culture' 'place']
[['best' 'great' 'excellent' 'amazing' 'good' 'awesome' 'life' 'work'
  'experience' 'place' 'company' 'culture' 'balance' 'benefits'
  'microsoft']]]
```

ภาพที่ 4.122 แสดงผลลัพธ์แต่ละหัวข้อในกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท microsoft

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท headline ของบริษัท microsoft ซึ่งมีอยู่ 11 หัวข้อ

และนี่คือผลลัพธ์จากการนับความถี่ของคำทุกคำที่ปรากฏในแต่ละหัวข้อ

```
Word frequencies:
excellent: 11
great: 11
good: 11
amazing: 11
awesome: 11
best: 11
experience: 11
work: 11
company: 11
place: 11
life: 11
balance: 11
benefits: 11
microsoft: 11
culture: 11
```

ภาพที่ 4.123 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท microsoft

นำคำและความถี่ที่นับได้ในกลุ่มที่ 3 ในข้อมูลประเภท headline ของบริษัท microsoft ไปแสดงผลผ่าน word cloud

Group topics 3 – microsoft\_headline all words each topics



ภาพที่ 4.124 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท headline ของบริษัท microsoft

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 3 ของข้อมูลประเภท headline ของบริษัท microsoft ซึ่งมีอยู่ 11 หัวข้อแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
[[ 'excellent' 'great' 'good' 'amazing' 'awesome' ]
 ['great' 'excellent' 'awesome' 'good' 'amazing' ]
 ['good' 'great' 'excellent' 'awesome' 'amazing' ]
 ['good' 'great' 'excellent' 'awesome' 'amazing' ]
 ['good' 'excellent' 'great' 'amazing' 'awesome' ]
 ['company' 'experience' 'work' 'life' 'place' ]
 ['work' 'life' 'company' 'experience' 'place' ]
 ['experience' 'excellent' 'amazing' 'great' 'best' ]
 ['work' 'experience' 'excellent' 'company' 'best' ]
 ['excellent' 'good' 'great' 'best' 'amazing' ]
 ['best' 'great' 'excellent' 'amazing' 'good']]
```

ภาพที่ 4.125 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ  
ในกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ของข้อมูลประเภท headline ของบริษัท microsoft

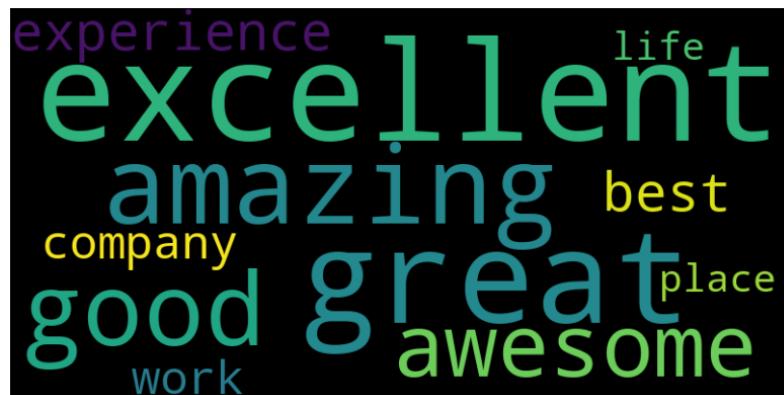
ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏซึ่งแสดงเป็นคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท headline ของบริษัท microsoft ซึ่งมีอยู่ 11 หัวข้อ และนี่คือผลลัพธ์จากการนับความถี่เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
Word frequencies:
excellent: 9
great: 8
good: 7
amazing: 8
awesome: 5
company: 3
experience: 4
work: 3
life: 2
place: 2
best: 4
```

ภาพที่ 4.126 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ในข้อมูลประเภท headline ของบริษัท microsoft

คำและความถี่ที่นับได้ในกลุ่มหัวข้อที่ 3 ซึ่งแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อของข้อมูลประเภท headline ของบริษัท microsoft ไปแสดงผลผ่าน word cloud

Group topics 3 – microsoft\_headline top 5 words each topics



ภาพที่ 4.127 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท headline ของบริษัท microsoft

ผู้วิจัยพิจารณาจาก word cloud แสดงเฉพาะคำ 5 อันดับแรกปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ของข้อมูลประเภท headline ของบริษัท microsoft ผู้วิจัยมีข้อสันนิษฐานว่าในกลุ่มหัวข้อที่ 3 ของ microsoft ข้อความรีวิวประเภท headline

มีความเกี่ยวข้องกับ: ชื่อบริษัท ประสบการณ์การทำงานเยี่ยม ชื่อสถานที่

8. ผลการวิเคราะห์หาคำที่มีนัยสำคัญจากการจัดกลุ่มหัวข้อซึ่งแสดงเป็น wordcloud ในข้อมูลรีวิวประเภท pros บริษัท Microsoft  
จัดกลุ่มหัวข้อให้เป็น 3 กลุ่ม

```
[ ] # reduce to group of 3
t2v_m_p_16v2_dp_load.hierarchical_topic_reduction(num_topics=3)

[[12, 7, 10, 17, 9, 16, 14, 15, 5, 3, 4, 13, 1], [0], [6, 8, 11, 2]]
```

ภาพที่ 4.128 แสดงการจัดกลุ่มผลลัพธ์หัวข้อของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท Microsoft

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท pros ของบริษัท microsoft ซึ่งมีอยู่ 13 หัวข้อ

```
[['culture' 'diversity' 'opportunities' 'technology' 'learn' 'experience' 'growth' 'grow' 'opportunity' 'management' 'worklife' 'environment' 'learning' 'world' 'benefits' 'leadership' 'career' 'life' 'flexibility' 'place' 'people' 'projects' 'best' 'products' 'company' 'smart' 'lots' 'really' 'perks' 'care' 'many' 'employees' 'balance' 'interesting' 'team' 'job' 'lot' 'pay' 'excellent' 'good' 'microsoft' 'great' 'work' 'amazing' 'flexible' 'salary' 'working' 'employee' 'get' 'awesome'],
 ['culture' 'company' 'worklife' 'employees' 'management' 'opportunities' 'employee' 'experience' 'diversity' 'career' 'leadership' 'technology' 'benefits' 'job' 'environment' 'microsoft' 'products' 'projects' 'place' 'work' 'best' 'opportunity' 'working' 'growth' 'world' 'team' 'life' 'salary' 'grow' 'people' 'learn' 'perks' 'excellent' 'great' 'pay' 'learning' 'compensation' 'smart' 'balance' 'amazing' 'care' 'flexibility' 'really' 'good' 'awesome' 'lots' 'flexible' 'lot' 'many' 'nice'],
 ['diversity' 'opportunities' 'culture' 'benefits' 'worklife' 'environment' 'experience' 'opportunity' 'employees' 'management' 'leadership' 'career' 'flexibility' 'projects' 'company' 'perks' 'technology' 'employee' 'balance' 'learn' 'grow' 'compensation' 'place' 'job' 'growth' 'work' 'microsoft' 'learning' 'life' 'people' 'flexible' 'world' 'products' 'team' 'working' 'pay' 'interesting' 'really' 'many' 'lots' 'awesome' 'salary' 'care' 'great' 'smart' 'excellent' 'get' 'nice' 'good' 'lot'],
 ['benefits' 'opportunities' 'products' 'culture' 'perks' 'worklife' 'pay' 'diversity' 'projects' 'management' 'employees' 'company' 'grow' 'life' 'opportunity' 'environment' 'experience' 'job' 'place' 'growth' 'balance' 'smart' 'work' 'learn' 'career' 'compensation' 'learning' 'working' 'people' 'technology' 'employee' 'flexibility' 'world' 'get' 'team' 'excellent' 'great' 'good' 'best' 'leadership' 'lots' 'care' 'microsoft' 'awesome' 'lot' 'salary' 'flexible' 'amazing' 'really' 'interesting'],
 ['benefits' 'culture' 'worklife' 'salary' 'opportunities' 'pay' 'compensation' 'employees' 'career' 'experience' 'management' 'diversity' 'employee' 'work' 'job' 'company' 'technology' 'growth' 'working' 'leadership' 'opportunity' 'learn' 'perks' 'projects' 'environment' 'flexibility' 'life' 'learning' 'grow' 'microsoft' 'place' 'balance' 'good' 'best' 'excellent' 'products' 'team' 'world' 'smart' 'really' 'care' 'get' 'lots' 'great' 'interesting' 'lot' 'awesome' 'people' 'flexible' 'amazing'],
 ['microsoft' 'worklife' 'company' 'management' 'employees' 'opportunities' 'technology' 'benefits' 'employee' 'experience' 'career' 'projects' 'products' 'job' 'growth' 'leadership' 'opportunity' 'work' 'working' 'environment' 'perks' 'smart' 'pay' 'compensation' 'learning' 'balance' 'best' 'salary' 'diversity' 'culture' 'really' 'great' 'world' 'amazing' 'grow' 'life' 'learn' 'flexibility' 'excellent' 'awesome' 'team' 'flexible' 'good' 'get' 'people' 'place' 'care' 'nice' 'lot' 'many'],
 ['smart' 'benefits' 'technology' 'opportunities' 'worklife' 'career' 'projects' 'employees' 'work' 'job' 'microsoft' 'working' 'employee' 'learning' 'products' 'experience' 'management' 'opportunity' 'company' 'people' 'learn' 'perks' 'pay' 'leadership' 'growth' 'environment' 'life' 'diversity' 'culture' 'flexibility' 'salary' 'compensation' 'world' 'really' 'care' 'get' 'best' 'lot' 'place' 'amazing' 'grow' 'awesome' 'great' 'interesting' 'balance' 'excellent' 'lots' 'many' 'flexible' 'team'],
 ['opportunities' 'learning' 'career' 'projects' 'opportunity' 'technology' 'management' 'leann' 'growth' 'experience' 'worklife' 'company' 'benefits' 'microsoft' 'products' 'environment' 'job' 'grow' 'employees' 'leadership' 'diversity' 'work' 'culture' 'flexibility' 'working' 'employee' 'smart' 'place' 'life' 'perks' 'world' 'pay' 'balance' 'compensation' 'lots' 'many' 'get' 'lot' 'flexible' 'salary' 'excellent' 'amazing' 'really' 'interesting' 'great' 'best' 'team' 'people' 'care' 'good']]
```

ภาพที่ 4.129 แสดงผลลัพธ์แต่ละหัวข้อในกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท microsoft

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท headline ของบริษัท microsoft ซึ่งมีอยู่ 13 หัวข้อ

และนี่คือผลลัพธ์จากการนับความถี่ของคำทุกคำที่ปรากฏในแต่ละหัวข้อ

```
Word frequencies: care: 13
culture: 13      many: 10
diversity: 13    employees: 13
opportunities: 13 balance: 13
technology: 13   interesting: 9
learn: 13        team: 13
experience: 13   job: 13
growth: 13       lot: 13
grow: 13         pay: 13
opportunity: 13  excellent: 13
management: 13   good: 12
worklife: 13     microsoft: 12
environment: 13  great: 13
learning: 13     work: 13
world: 13        amazing: 12
benefits: 13     flexible: 10
leadership: 13   salary: 12
career: 13        working: 13
life: 13          employee: 13
flexibility: 11  get: 12
place: 13         awesome: 12
people: 13        compensation: 12
projects: 13     nice: 8
best: 12
products: 13
company: 13
smart: 13
lots: 12
really: 13
perks: 13
```

ภาพที่ 4.130 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท microsoft

คำและความถี่ที่นับได้ในกลุ่มที่ 1 ในข้อมูลประเภท pros ของบริษัท microsoft ไปแสดงผ่าน word cloud

Group 1 – microsoft\_pros all words each topics



ภาพที่ 4.131 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท pros ของบริษัท Microsoft

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 1 ของข้อมูลประเภท pros ของบริษัท microsoft ซึ่งมีอยู่ 13 หัวข้อแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
[['culture' 'diversity' 'opportunities' 'technology' 'learn']
 ['culture' 'company' 'worklife' 'employees' 'management']
 ['diversity' 'opportunities' 'culture' 'benefits' 'worklife']
 ['benefits' 'opportunities' 'products' 'culture' 'perks']
 ['benefits' 'culture' 'worklife' 'salary' 'opportunities']
 ['good' 'great' 'excellent' 'many' 'lots']
 ['people' 'employees' 'job' 'working' 'work']
 ['place' 'work' 'working' 'job' 'worklife']
 ['company' 'working' 'job' 'work' 'employees']
 ['environment' 'worklife' 'employees' 'opportunities' 'management']
 ['microsoft' 'worklife' 'company' 'management' 'employees']
 ['smart' 'benefits' 'technology' 'opportunities' 'worklife']
 ['opportunities' 'learning' 'career' 'projects' 'opportunity']]
```

ภาพที่ 4.132 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ  
ในกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ของข้อมูลประเภท pros ของบริษัท microsoft

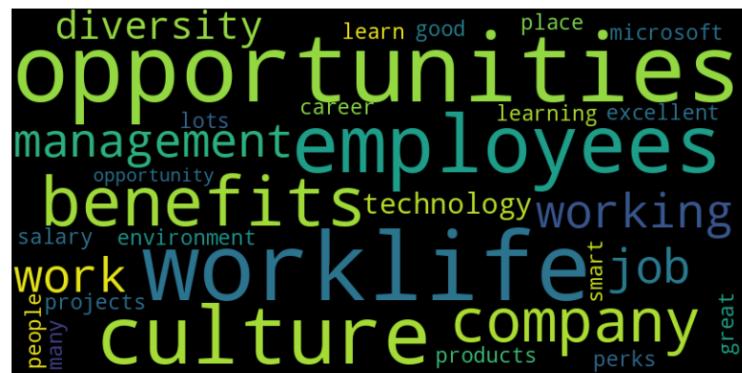
ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏชื่้นแสดงเป็นคำ 5 อันดับแรก  
ในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท pros ของบริษัท microsoft ซึ่งมีอยู่ 13 หัวข้อ  
และนี่คือผลลัพธ์จากการนับความถี่เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

Word frequencies:
culture: 5
diversity: 2
opportunities: 7
technology: 2
learn: 1
company: 3
worklife: 7
employees: 5
management: 3
benefits: 4
products: 1
perks: 1
salary: 1
good: 1
great: 1
excellent: 1
many: 1
lots: 1
people: 1
job: 3
working: 3
work: 3
place: 1
environment: 1
microsoft: 1
smart: 1
learning: 1
career: 1
projects: 1
opportunity: 1

ภาพที่ 4.133 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1  
ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท microsoft

นำคำและความถี่ที่นับได้ในกลุ่มหัวข้อที่ 1 ซึ่งแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อของข้อมูลประเภท pros ของบริษัท microsoft ไปแสดงผลผ่าน word cloud

Group 1 – microsoft\_pros top 5 words each topics



ภาพที่ 4.134 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท pros ของบริษัท microsoft

ผู้วิจัยพิจารณาจาก word cloud แสดงเฉพาะคำ 5 อันดับแรกปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ของข้อมูลประเภท pros ของบริษัท microsoft ผู้วิจัยมีข้อสันนิษฐานว่าในกลุ่มหัวข้อที่ 1 ของ microsoft ข้อความรีวิวประเภท pros

มีความเกี่ยวข้องกับ: โอกาสของพนักงานและการจัดการพนักงาน ชีวิตการทำงาน ผลประโยชน์ วัฒนธรรม ความหลากหลาย

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท pros ของบริษัท microsoft ซึ่งมีอยู่ 1 หัวข้อ

```
[[ 'benefits' 'compensation' 'salary' 'opportunities' 'pay' 'worklife'
  'career' 'employees' 'employee' 'job' 'work' 'experience' 'company'
  'perks' 'working' 'management' 'opportunity' 'growth' 'projects'
  'learn' 'learning' 'flexibility' 'good' 'technology' 'excellent' 'get'
  'life' 'products' 'microsoft' 'leadership' 'smart' 'diversity' 'great'
  'best' 'balance' 'lot' 'lots' 'grow' 'awesome' 'nice' 'environment'
  'culture' 'really' 'interesting' 'amazing' 'care' 'place' 'team' 'many'
  'flexible']]
```

ภาพที่ 4.135 แสดงผลลัพธ์หัวข้อในกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท microsoft

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท pros ของบริษัท microsoft ซึ่งมีอยู่ 1 หัวข้อ

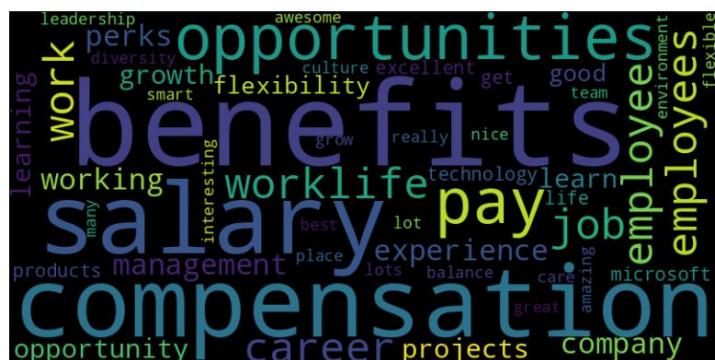
และนี่คือผลลัพธ์จากการนับความถี่ของคำทุกคำที่ปรากฏในแต่ละหัวข้อ

```
Word frequencies: get: 1
benefits: 1      life: 1
compensation: 1 products: 1
salary: 1        microsoft: 1
opportunities: 1 leadership: 1
pay: 1           smart: 1
worklife: 1      diversity: 1
career: 1         great: 1
employees: 1     best: 1
employee: 1      balance: 1
job: 1            lot: 1
work: 1           lots: 1
experience: 1    grow: 1
company: 1        awesome: 1
perks: 1          nice: 1
working: 1        environment: 1
management: 1    culture: 1
opportunity: 1   really: 1
growth: 1          interesting: 1
projects: 1        amazing: 1
learn: 1           care: 1
learning: 1        place: 1
flexibility: 1    team: 1
good: 1            many: 1
technology: 1     flexible: 1
excellent: 1
```

ภาพที่ 4.136 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท microsoft

คำและความถี่ที่นับได้ในกลุ่มที่ 2 ในข้อมูลประเภท pros ของบริษัท microsoft ไปแสดงผ่าน word cloud

Group topics 2 – microsoft\_pros all words



ภาพที่ 4.137 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท pros ของบริษัท Microsoft

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 2 ของข้อมูลประเภท pros ของบริษัท microsoft ซึ่งมีอยู่ 1 หัวข้อแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
[[ 'benefits' 'compensation' 'salary' 'opportunities' 'pay']]
```

ภาพที่ 4.138 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ  
ในกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ของข้อมูลประเภท pros ของบริษัท microsoft

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏชื่้นแสดงเป็นคำ 5 อันดับแรก  
ในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท pros ของบริษัท microsoft ซึ่งมีอยู่ 1 หัวข้อ  
และนี่คือผลลัพธ์จากการนับความถี่เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
Word frequencies:  
benefits: 1  
compensation: 1  
salary: 1  
opportunities: 1  
pay: 1
```

ภาพที่ 4.139 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2  
ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท microsoft

นำคำและความถี่ที่นับได้ในกลุ่มหัวข้อที่ 2 ซึ่งแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อของข้อมูล  
ประเภท pros ของบริษัท microsoft ไปแสดงผลผ่าน word cloud

Group topics 2 – microsoft\_pros top 5 words



ภาพที่ 4.140 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ใน  
ข้อมูลประเภท pros ของบริษัท microsoft

ผู้วิจัยพิจารณาจาก word cloud แสดงเฉพาะคำ 5 อันดับแรกปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ของข้อมูลประเภท pros ของบริษัท microsoft ผู้วิจัยมีข้อสันนิษฐานว่าในกลุ่มหัวข้อที่ 2 ของ microsoft ข้อความรีวิวประเภท pros มีความเกี่ยวข้องกับ: ผลประโยชน์และโอกาส, เงินชดเชยและเงินเดือน

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท pros ของบริษัท microsoft ซึ่งมีอยู่ 4 หัวข้อ

```
[['worklife' 'flexibility' 'flexible' 'opportunities' 'benefits' 'work'
 'career' 'employees' 'employee' 'working' 'compensation' 'job'
 'management' 'experience' 'projects' 'salary' 'company' 'opportunity'
 'pay' 'life' 'technology' 'balance' 'environment' 'diversity'
 'microsoft' 'perks' 'leadership' 'awesome' 'growth' 'smart' 'learning'
 'amazing' 'products' 'interesting' 'place' 'culture' 'get' 'excellent'
 'grow' 'learn' 'world' 'great' 'really' 'team' 'best' 'nice' 'lots'
 'good' 'lot' 'care']
['worklife' 'culture' 'balance' 'career' 'opportunities' 'employees'
 'management' 'work' 'benefits' 'life' 'employee' 'diversity' 'job'
 'working' 'salary' 'company' 'experience' 'compensation' 'opportunity'
 'leadership' 'pay' 'growth' 'technology' 'flexibility' 'environment'
 'projects' 'world' 'best' 'grow' 'learn' 'place' 'microsoft' 'team'
 'flexible' 'smart' 'care' 'perks' 'learning' 'products' 'really'
 'amazing' 'lot' 'people' 'excellent' 'awesome' 'great' 'get'
 'interesting' 'good' 'lots']
['worklife' 'balance' 'team' 'employees' 'management' 'career' 'work'
 'opportunities' 'leadership' 'life' 'working' 'projects' 'employee'
 'company' 'benefits' 'job' 'opportunity' 'flexibility' 'diversity'
 'experience' 'growth' 'compensation' 'best' 'culture' 'environment'
 'people' 'pay' 'care' 'microsoft' 'technology' 'perks' 'lot' 'learn'
 'flexible' 'world' 'products' 'smart' 'really' 'amazing' 'grow'
 'learning' 'excellent' 'place' 'great' 'awesome' 'salary' 'get' 'lots'
 'nice' 'good']
['worklife' 'balance' 'work' 'career' 'life' 'benefits' 'opportunities'
 'employees' 'working' 'employee' 'job' 'compensation' 'management'
 'company' 'opportunity' 'growth' 'salary' 'pay' 'projects'
 'flexibility' 'experience' 'leadership' 'diversity' 'environment'
 'best' 'smart' 'technology' 'care' 'amazing' 'perks' 'world'
 'microsoft' 'learn' 'lot' 'flexible' 'culture' 'products' 'team' 'grow'
 'great' 'excellent' 'place' 'awesome' 'get' 'learning' 'really' 'good'
 'nice' 'lots' 'interesting']]
```

ภาพที่ 4.141 แสดงผลลัพธ์แต่ละหัวข้อในกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท microsoft

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท pros ของบริษัท microsoft ซึ่งมีอยู่ 4 หัวข้อ

และนี่คือผลลัพธ์จากการนับความถี่ของคำทุกคำที่ปรากฏในแต่ละหัวข้อ

```
Word frequencies: microsoft: 4
worklife: 4 perks: 4
flexibility: 4 leadership: 4
flexible: 4 awesome: 4
opportunities: 4 growth: 4
benefits: 4 smart: 4
work: 4 learning: 4
career: 4 amazing: 4
employees: 4 products: 4
employee: 4 interesting: 3
working: 4 place: 4
compensation: 4 culture: 4
job: 4 get: 4
management: 4 excellent: 4
experience: 4 grow: 4
projects: 4 learn: 4
salary: 4 world: 4
company: 4 great: 4
opportunity: 4 really: 4
pay: 4 team: 4
life: 4 best: 4
technology: 4 nice: 3
balance: 4 lots: 4
environment: 4 good: 4
diversity: 4 lot: 4
people: 2 care: 4
```

ภาพที่ 4.142 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท microsoft

คำและความถี่ที่นับได้ในกลุ่มที่ 3 ในข้อมูลประเภท pros ของบริษัท microsoft ไปแสดงผลผ่าน word cloud

Group topics 3 – microsoft\_pros all words each topics



ภาพที่ 4.143 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท pros ของบริษัท microsoft

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 3 ของข้อมูลประเภท pros ของบริษัท microsoft ซึ่งมีอยู่ 4 หัวข้อแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
[['worklife' 'flexibility' 'flexible' 'opportunities' 'benefits']
 ['worklife' 'culture' 'balance' 'career' 'opportunities']
 ['worklife' 'balance' 'team' 'employees' 'management']
 ['worklife' 'balance' 'work' 'career' 'life']]
```

ภาพที่ 4.144 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ  
ในกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท microsoft

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏชื่้นแสดงเป็นคำ 5 อันดับแรก ในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท pros ของบริษัท microsoft ซึ่งมีอยู่ 4 หัวข้อ และนี่คือผลลัพธ์จากการนับความถี่เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
Word frequencies:
worklife: 4
flexibility: 1
flexible: 1
opportunities: 2
benefits: 1
culture: 1
balance: 3
career: 2
team: 1
employees: 1
management: 1
work: 1
life: 1
```

ภาพที่ 4.145 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3  
ของเทคนิค Top2vec ในข้อมูลประเภท pros ของบริษัท microsoft

นำคำและความถี่ที่นับได้ในกลุ่มหัวข้อที่ 3 ซึ่งแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อของข้อมูล  
ประเภท pros ของบริษัท microsoft ไปแสดงผลผ่าน word cloud

Group topics 3 – microsoft\_pros top 5 words each topics



ภาพที่ 4.146 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ใน  
ข้อมูลประเภท pros ของบริษัท Microsoft

ผู้วิจัยพิจารณาจาก word cloud แสดงเฉพาะคำ 5 อันดับแรกปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ของข้อมูลประเภท pros ของบริษัท microsoft ผู้วิจัยมีข้อสันนิษฐานว่าในกลุ่มหัวข้อที่ 3 ของ microsoft ข้อความรีวิวประเภท pros มีความเกี่ยวข้องกับ: ความสมดุลและโอกาส, ชีวิตการทำงานและความยืดหยุ่น

9. ผลการวิเคราะห์หาคำที่มีนัยสำคัญจากการจัดกลุ่มหัวข้อซึ่งแสดงเป็น wordcloud ในข้อมูลรีวิวประเภท cons บริษัท Microsoft

จัดกลุ่มหัวข้อให้เป็น 3 กลุ่ม

```
[ ] # reduce to group of 3
t2v_m_c_l6v2_dp_load.hierarchical_topic_reduction(num_topics=3)

[[25, 6, 21, 27, 0, 10, 28, 2, 14, 17, 7, 11, 18, 15, 13, 4],
 [19, 8, 5, 23, 16, 12, 22, 1],
 [3, 9, 24, 26, 28]]
```

ภาพที่ 4.147 แสดงการจัดกลุ่มผลลัพธ์หัวข้อของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท microsoft

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท cons ของบริษัท microsoft ซึ่งมีอยู่ 16 หัวข้อ

```
[[{'cons': 'companies', 'company': 'management', 'employees': 'managers', 'politics': 'team', 'culture': 'internal', 'hours': 'don\'t', 'change': 'compared', 'team': 'salary', 'really': 'politics', 'team': 'balance', 'life': 'good', 'like': 'great', 'time': 'long', 'culture': 'new', 'pressure': 'sometimes', 'like': 'much', 'hours': 'large', 'none': 'people'}, {'cons': 'bad', 'don\'t': 'compared', 'change': 'really', 'like': 'think', 'good': 'pay', 'team': 'internal', 'politics': 'hours', 'work': 'tech', 'team': 'compared', 'team': 'working', 'good': 'low', 'not': 'change', 'none': 'team', 'internal', 'politics': 'hours', 'com': 'pressure', 'team': 'difficult', 'get': 'many', 'lots': 'hard', 'high', 'lot': 'large', 'none': 'people'}, {'cons': 'bad', 'don\'t': 'compared', 'change': 'really', 'like': 'think', 'good': 'pay', 'team': 'internal', 'politics': 'hours', 'work': 'tech', 'team': 'compared', 'team': 'working', 'good': 'low', 'not': 'change', 'none': 'team', 'internal', 'politics': 'hours', 'com': 'pressure', 'team': 'difficult', 'get': 'many', 'lots': 'hard', 'nothing', 'dont': 'many'}, {'cons': 'pay', 'team': 'internal', 'politics': 'hours', 'work': 'tech', 'team': 'compared', 'team': 'working', 'good': 'low', 'not': 'change', 'none': 'team', 'internal', 'politics': 'hours', 'com': 'pressure', 'team': 'difficult', 'get': 'many', 'lots': 'hard', 'nothing', 'like': 'think', 'team': 'think', 'team': 'life', 'great', 'large', 'time': 'bad', 'lots': 'hard', 'nothing', 'dont': 'many'}, {'cons': 'pay', 'team': 'internal', 'politics': 'hours', 'work': 'tech', 'team': 'compared', 'team': 'working', 'good': 'low', 'not': 'change', 'none': 'team', 'internal', 'politics': 'hours', 'com': 'pressure', 'team': 'difficult', 'get': 'many', 'lots': 'hard', 'nothing', 'like': 'think', 'team': 'think', 'team': 'life', 'great', 'large', 'time': 'bad', 'lots': 'hard', 'nothing', 'dont': 'many'}, {'cons': 'pay', 'team': 'internal', 'politics': 'hours', 'work': 'tech', 'team': 'compared', 'team': 'working', 'good': 'low', 'not': 'change', 'none': 'team', 'internal', 'politics': 'hours', 'com': 'pressure', 'team': 'difficult', 'get': 'many', 'lots': 'hard', 'nothing', 'like': 'think', 'team': 'think', 'team': 'life', 'great', 'large', 'time': 'bad', 'lots': 'hard', 'nothing', 'dont': 'many'}, {'cons': 'pay', 'team': 'internal', 'politics': 'hours', 'work': 'tech', 'team': 'compared', 'team': 'working', 'good': 'low', 'not': 'change', 'none': 'team', 'internal', 'politics': 'hours', 'com': 'pressure', 'team': 'difficult', 'get': 'many', 'lots': 'hard', 'nothing', 'like': 'think', 'team': 'think', 'team': 'life', 'great', 'large', 'time': 'bad', 'lots': 'hard', 'nothing', 'dont': 'many'}, {'cons': 'pay', 'team': 'internal', 'politics': 'hours', 'work': 'tech', 'team': 'compared', 'team': 'working', 'good': 'low', 'not': 'change', 'none': 'team', 'internal', 'politics': 'hours', 'com': 'pressure', 'team': 'difficult', 'get': 'many', 'lots': 'hard', 'nothing', 'like': 'think', 'team': 'think', 'team': 'life', 'great', 'large', 'time': 'bad', 'lots': 'hard', 'nothing', 'dont': 'many'}, {'cons': 'pay', 'team': 'internal', 'politics': 'hours', 'work': 'tech', 'team': 'compared', 'team': 'working', 'good': 'low', 'not': 'change', 'none': 'team', 'internal', 'politics': 'hours', 'com': 'pressure', 'team': 'difficult', 'get': 'many', 'lots': 'hard', 'nothing', 'like': 'think', 'team': 'think', 'team': 'life', 'great', 'large', 'time': 'bad', 'lots': 'hard', 'nothing', 'dont': 'many'}, {'cons': 'pay', 'team': 'internal', 'politics': 'hours', 'work': 'tech', 'team': 'compared', 'team': 'working', 'good': 'low', 'not': 'change', 'none': 'team', 'internal', 'politics': 'hours', 'com': 'pressure', 'team': 'difficult', 'get': 'many', 'lots': 'hard', 'nothing', 'like': 'think', 'team': 'think', 'team': 'life', 'great', 'large', 'time': 'bad', 'lots': 'hard', 'nothing', 'dont': 'many'}, {'cons': 'pay', 'team': 'internal', 'politics': 'hours', 'work': 'tech', 'team': 'compared', 'team': 'working', 'good': 'low', 'not': 'change', 'none': 'team', 'internal', 'politics': 'hours', 'com': 'pressure', 'team': 'difficult', 'get': 'many', 'lots': 'hard', 'nothing', 'like': 'think', 'team': 'think', 'team': 'life', 'great', 'large', 'time': 'bad', 'lots': 'hard', 'nothing', 'dont': 'many'}, {'cons': 'pay', 'team': 'internal', 'politics': 'hours', 'work': 'tech', 'team': 'compared', 'team': 'working', 'good': 'low', 'not': 'change', 'none': 'team', 'internal', 'politics': 'hours', 'com': 'pressure', 'team': 'difficult', 'get': 'many', 'lots': 'hard', 'nothing', 'like': 'think', 'team': 'think', 'team': 'life', 'great', 'large', 'time': 'bad', 'lots': 'hard', 'nothing', 'dont': 'many'}, {'cons': 'pay', 'team': 'internal', 'politics': 'hours', 'work': 'tech', 'team': 'compared', 'team': 'working', 'good': 'low', 'not': 'change', 'none': 'team', 'internal', 'politics': 'hours', 'com': 'pressure', 'team': 'difficult', 'get': 'many', 'lots': 'hard', 'nothing', 'like': 'think', 'team': 'think', 'team': 'life', 'great', 'large', 'time': 'bad', 'lots': 'hard', 'nothing', 'dont': 'many'}, {'cons': 'pay', 'team': 'internal', 'politics': 'hours', 'work': 'tech', 'team': 'compared', 'team': 'working', 'good': 'low', 'not': 'change', 'none': 'team', 'internal', 'politics': 'hours', 'com': 'pressure', 'team': 'difficult', 'get': 'many', 'lots': 'hard', 'nothing', 'like': 'think', 'team': 'think', 'team': 'life', 'great', 'large', 'time': 'bad', 'lots': 'hard', 'nothing', 'dont': 'many'}, {'cons': 'pay', 'team': 'internal', 'politics': 'hours', 'work': 'tech', 'team': 'compared', 'team': 'working', 'good': 'low', 'not': 'change', 'none': 'team', 'internal', 'politics': 'hours', 'com': 'pressure', 'team': 'difficult', 'get': 'many', 'lots': 'hard', 'nothing', 'like': 'think', 'team': 'think', 'team': 'life', 'great', 'large', 'time': 'bad', 'lots': 'hard', 'nothing', 'dont': 'many'}, {'cons': 'pay', 'team': 'internal', 'politics': 'hours', 'work': 'tech', 'team': 'compared', 'team': 'working', 'good': 'low', 'not': 'change', 'none': 'team', 'internal', 'politics': 'hours', 'com': 'pressure', 'team': 'difficult', 'get': 'many', 'lots': 'hard', 'nothing', 'like': 'think', 'team': 'think', 'team': 'life', 'great', 'large', 'time': 'bad', 'lots': 'hard', 'nothing', 'dont': 'many'}, {'cons': 'pay', 'team': 'internal', 'politics': 'hours', 'work': 'tech', 'team': 'compared', 'team': 'working', 'good': 'low', 'not': 'change', 'none': 'team', 'internal', 'politics': 'hours', 'com': 'pressure', 'team': 'difficult', 'get': 'many', 'lots': 'hard', 'nothing', 'like': 'think', 'team': 'think', 'team': 'life', 'great', 'large', 'time': 'bad', 'lots': 'hard', 'nothing', 'dont': 'many'}, {'cons': 'pay', 'team': 'internal', 'politics': 'hours', 'work': 'tech', 'team': 'compared', 'team': 'working', 'good': 'low', 'not': 'change', 'none': 'team', 'internal', 'politics': 'hours', 'com': 'pressure', 'team': 'difficult', 'get': 'many', 'lots': 'hard', 'nothing', 'like': 'think', 'team': 'think', 'team': 'life', 'great', 'large', 'time': 'bad', 'lots': 'hard', 'nothing', 'dont': 'many'}, {'cons': 'pay', 'team': 'internal', 'politics': 'hours', 'work': 'tech', 'team': 'compared', 'team': 'working', 'good': 'low', 'not': 'change', 'none': 'team', 'internal', 'politics': 'hours', 'com': 'pressure', 'team': 'difficult', 'get': 'many', 'lots': 'hard', 'nothing', 'like': 'think', 'team': 'think', 'team': 'life', 'great', 'large', 'time': 'bad', 'lots': 'hard', 'nothing', 'dont': 'many'}]]
```

ภาพที่ 4.148 แสดงผลลัพธ์แต่ละหัวข้อในกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท microsoft

ในการวิเคราะห์ผลลัพธ์ทั่วไปผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท cons ของบริษัท microsoft ซึ่งมีอยู่ 16 หัวข้อ

และนี่คือผลลัพธ์จากการนับความถี่ของคำทุกคำที่ปรากฏในแต่ละหัวข้อ

Word frequencies:	balance: 16
cons:	life: 16
companies:	bad: 14
company:	growth: 15
management:	good: 15
employees:	like: 16
managers:	great: 16
career:	times: 15
work:	time: 15
compensation:	culture: 16
tech:	new: 14
working:	pressure: 14
microsoft:	sometimes: 15
pay:	low: 13
internal:	much: 14
hours:	think: 16
dont:	big: 13
change:	difficult: 16
compared:	get: 16
teams:	many: 15
salary:	lots: 16
really:	hard: 16
politics:	high: 13
team:	lot: 16
	large: 12
	none: 16
	people: 15
	long: 9
	nothing: 14
	slow: 14
	less: 12

ภาพที่ 4.149 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท microsoft

นำคำและความคิดที่นับได้ในกลุ่มที่ 1 ในข้อมูลประเภท cons ของบริษัท microsoft ไปแสดงผลผ่าน word cloud

Group topics 1 – microsoft cons all words each topics



ภาพที่ 4.150 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท cons ของบริษัท microsoft

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท cons ของบริษัท microsoft ซึ่งมีอยู่ 11 หัวข้อ

```
[[['cons' 'companies' 'company' 'management' 'employees'],
  ['cons' 'bad' 'dont' 'compared' 'change'],
  ['working' 'work' 'company' 'great' 'career'],
  ['none' 'think' 'many' 'lots' 'nothing'],
  ['nothing' 'none' 'think' 'bad' 'lot'],
  ['compensation' 'salary' 'pay' 'companies' 'employees'],
  ['pay' 'companies' 'company' 'employees' 'managers'],
  ['salary' 'compensation' 'pay' 'less' 'companies'],
  ['slow' 'change' 'time' 'growth' 'times'],
  ['tech' 'microsoft' 'companies' 'career' 'growth'],
  ['career' 'management' 'growth' 'managers' 'companies'],
  ['microsoft' 'companies' 'tech' 'management' 'cons'],
  ['employees' 'management' 'companies' 'company' 'managers'],
  ['growth' 'companies' 'company' 'slow' 'management'],
  ['management' 'companies' 'company' 'managers' 'employees'],
  ['companies' 'company' 'management' 'managers' 'employees']]]
```

ภาพที่ 4.151 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ  
ในกลุ่มหัวข้อที่ 1 ของเทคนิค Top2vec ของข้อมูลประเภท cons ของบริษัท microsoft

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏชื่้นแสดงเป็นคำ 5 อันดับแรก  
ในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท cons ของบริษัท microsoft ซึ่งมีอยู่ 16 หัวข้อ<sup>1</sup>  
และนี่คือผลลัพธ์จากการนับความถี่เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
Word frequencies:
cons: 3
companies: 11
company: 7
management: 7
employees: 6
bad: 2
dont: 1
compared: 1
change: 2
working: 1
work: 1
great: 1
career: 3
none: 2
think: 2
many: 1
lots: 1
nothing: 2
lot: 1
compensation: 2
salary: 2
pay: 3
managers: 5
less: 1
slow: 2
time: 1
growth: 4
times: 1
tech: 2
microsoft: 2
```

ภาพที่ 4.152 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1  
ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท microsoft

นำคำและความถี่ที่นับได้ในกลุ่มหัวข้อที่ 1 ซึ่งแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อของข้อมูลประเภท cons ของบริษัท microsoft ไปแสดงผลผ่าน word cloud

Group topics 1 – microsoft\_cons top 5 words each topics



ภาพที่ 4.153 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ในข้อมูลประเภท cons ของบริษัท microsoft

ผู้วิจัยพิจารณาจาก word cloud แสดงเฉพาะคำ 5 อันดับแรกปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 1 ของข้อมูลประเภท cons ของบริษัท microsoft ผู้วิจัยมีข้อสันนิษฐานว่าในกลุ่มหัวข้อที่ 1 ของ microsoft ข้อความรีวิวประเภท cons มีความเกี่ยวข้องกับ: การจัดการพนักงานของบริษัท, ผู้จัดการ(managers)

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท cons ของบริษัท microsoft ซึ่งมีอยู่ 8 หัวข้อ

```
[[{"management": "politics", "managers": "employees", "companies": "career", "company": "culture", "internal": "tech", "growth": "microsoft", "slow": "difficult", "work": "change", "people": "pay", "working": "compared", "compensation": "balance", "like": "times", "teams": "really", "salary": "life", "hard": "time", "sometimes": "lots", "team": "hours", "cons": "large", "big": "many", "lot": "long", "less": "think", "bad": "pressure", "good": "new", "none": "great", "much": "high"}, {"politics": "management", "managers": "internal", "companies": "culture", "company": "times", "employees": "difficult", "growth": "work", "change": "company", "really": "lots", "time": "hard", "people": "working", "pressure": "microsoft", "many": "life", "slow": "balance", "lot": "sometimes", "teams": "think", "pay": "get", "compared": "hours", "cons": "like", "bad": "less", "none": "large", "long": "good", "tech": "big", "team": "much", "nothing": "high", "compensation": "new"}, {"teams": "team", "managers": "employees", "management": "companies", "company": "culture", "internal": "tech", "work": "people", "working": "growth", "microsoft": "compared", "balance": "difficult", "change": "slow", "hard": "times", "internal": "culture", "politics": "time", "really": "think", "hours": "life", "big": "sometimes", "lots": "like", "large": "lot", "pressure": "none", "pay": "many", "get": "bad", "compensation": "cons", "new": "nothing", "dont": "long", "high": "good", "less"}, {"culture": "management", "managers": "employees", "growth": "internal", "company": "politics", "people": "team", "teams": "change", "culture": "companies", "tech": "life", "compared": "difficult", "balance": "really", "work": "think", "like": "times", "hard": "lots", "many": "bad", "cons": "sometimes", "salary": "nothing", "pay": "working", "get": "slow", "time": "microsoft", "dont": "lot", "new": "good", "high": "pressure", "hours": "compensation", "low": "none", "large": "much"}, {"employees": "management", "managers": "careers", "companies": "culture", "company": "politics", "tech": "growth", "change": "balance", "work": "microsoft", "compensation": "teams", "internal": "working", "people": "difficult", "times": "many", "team": "really", "salary": "like", "pay": "lots", "compared": "get", "think": "none", "less": "sometimes", "life": "hours", "nothing": "dont", "pressure": "hard", "cons": "low", "lot": "new", "high": "time", "good": "much", "bad": "great"}, {"managers": "management", "companies": "career", "growth": "company", "employees": "compensation", "work": "pay", "get": "teams", "times": "working", "politics": "really", "hours": "difficult", "high": "hard", "microsoft": "culture", "slow": "change", "time": "team", "tech": "good", "life": "less", "like": "compared", "salary": "low", "bad": "internal", "sometimes": "great", "dont": "new", "think": "balance", "cons": "people", "long": "nothing", "mush": "pressure", "many": "lots"}, {"management": "careers", "managers": "employees", "companies": "compensation", "tech": "change", "internal": "company", "work": "working", "growth": "balance", "teams": "salary", "difficult": "team", "hard": "microsoft", "pay": "politics", "hours": "time", "get": "times", "slow": "think", "pressure": "life", "really": "new", "like": "nothing", "cons": "people", "high": "lots", "many": "lot", "mush": "great", "low": "good", "compared": "long", "dont": "none"}, {"management": "managers", "employees": "companies", "career": "company", "politics": "growth", "internal": "culture", "microsoft": "work", "teams": "working", "change": "team", "balance": "compensation", "compared": "people", "times": "difficult", "salary": "tech", "high": "life", "really": "sometimes", "pay": "get", "hard": "like", "pressure": "bad", "time": "good", "low": "think", "many": "cons", "large": "slow", "less": "nothing", "lots": "lot", "hours": "none", "great": "dont"}]]
```

ภาพที่ 4.154 แสดงผลลัพธ์แต่ละหัวข้อในกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท microsoft

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท cons ของบริษัท microsoft ซึ่งมีอยู่ 8 หัวข้อ

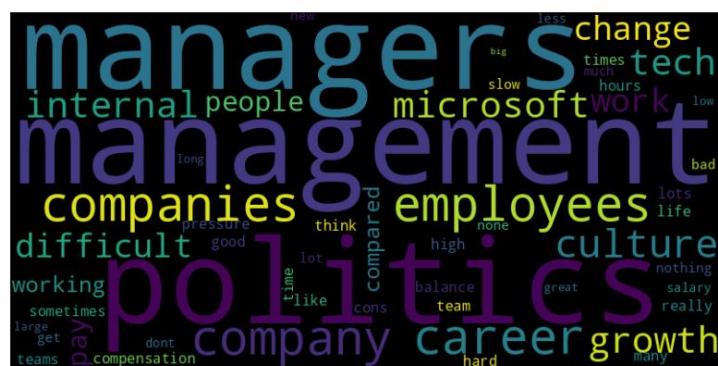
และนี่คือผลลัพธ์จากการนับความถี่ของคำทุกคำที่ปรากฏในแต่ละหัวข้อ

Word frequencies:	really: 8
management:	8
politics:	8
managers:	8
employees:	8
companies:	8
career:	8
company:	8
culture:	8
internal:	8
tech:	8
growth:	8
microsoft:	8
slow:	7
difficult:	8
work:	8
change:	8
people:	8
pay:	8
working:	8
compared:	8
compensation:	8
balance:	8
like:	8
times:	8
teams:	8
salary:	6
life:	8
hard:	8
time:	8
sometimes:	8
lots:	8
team:	8
hours:	8
cons:	8
large:	5
big:	3
many:	8
lot:	7
long:	5
less:	6
think:	8
bad:	7
pressure:	8
good:	8
new:	7
none:	7
great:	5
much:	6
high:	8
get:	7
nothing:	7
dont:	6
low:	5

ภาพที่ 4.155 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท microsoft

คำและความถี่ที่นับได้ในกลุ่มที่ 2 ในข้อมูลประเภท cons ของบริษัท microsoft ไปแสดงผลผ่าน word cloud

Group topics 2 – microsoft\_cons all words each topics



ภาพที่ 4.156 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท cons ของบริษัท microsoft

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 2 ของข้อมูลประเภท cons ของบริษัท microsoft ซึ่งมีอยู่ 8 หัวข้อแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
[['management' 'politics' 'managers' 'employees' 'companies']
 ['politics' 'management' 'managers' 'internal' 'companies']
 ['teams' 'team' 'managers' 'employees' 'management']
 ['culture' 'management' 'managers' 'employees' 'growth']
 ['employees' 'management' 'managers' 'career' 'companies']
 ['managers' 'management' 'companies' 'career' 'growth']
 ['management' 'career' 'managers' 'employees' 'companies']
 ['management' 'managers' 'employees' 'companies' 'career']]
```

ภาพที่ 4.157 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ  
ในกลุ่มหัวข้อที่ 2 ของเทคนิค Top2vec ของข้อมูลประเภท cons ของบริษัท microsoft

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏซึ่งแสดงเป็นคำ 5 อันดับแรก  
ในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท cons ของบริษัท microsoft ซึ่งมีอยู่ 8 หัวข้อ  
และนี่คือผลลัพธ์จากการนับความถี่เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
Word frequencies:
management: 8
politics: 2
managers: 8
employees: 6
companies: 6
internal: 1
teams: 1
team: 1
culture: 1
growth: 2
career: 4
```

ภาพที่ 4.158 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2  
ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท microsoft

คำและความถี่ที่นับได้ในกลุ่มหัวข้อที่ 2 ซึ่งแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อของข้อมูล  
ประเภท cons ของบริษัท microsoft ไปแสดงผลผ่าน word cloud

### Group topics 2 – microsoft\_cons top 5 words each topics



ภาพที่ 4.159 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ในข้อมูลประเภท cons ของบริษัท microsoft

ผู้วิจัยพิจารณาจาก word cloud แสดงเฉพาะคำ 5 อันดับแรกปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 2 ของข้อมูลประเภท cons ของบริษัท microsoft ผู้วิจัยมีข้อสันนิษฐานว่าในกลุ่มหัวข้อที่ 2 ของ microsoft ข้อความรีวิวประเภท cons

มีความเกี่ยวข้องกับ: การจัดการพนักงาน, การทำงานเป็นทีมในบริษัท, การเมืองภายใน

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท cons ของบริษัท microsoft ซึ่งมีอยู่ 5 หัวข้อ

```
[[ 'balance' 'work' 'career' 'life' 'management' 'working' 'employees'
  'compensation' 'managers' 'company' 'growth' 'companies' 'pay' 'salary'
  'tech' 'hours' 'times' 'microsoft' 'politics' 'culture' 'change' 'lot'
  'difficult' 'team' 'teams' 'get' 'time' 'really' 'internal' 'compared'
  'high' 'hard' 'sometimes' 'much' 'less' 'pressure' 'think' 'people'
  'low' 'great' 'lots' 'nothing' 'dont' 'like' 'bad' 'good' 'cons' 'none'
  'long' 'many']
 [ 'hours' 'work' 'times' 'working' 'time' 'long' 'career' 'employees'
  'life' 'pay' 'company' 'management' 'managers' 'compensation'
  'companies' 'lots' 'lot' 'sometimes' 'slow' 'difficult' 'really' 'tech'
  'balance' 'hard' 'bad' 'get' 'great' 'cons' 'much' 'growth' 'like'
  'salary' 'politics' 'less' 'microsoft' 'dont' 'think' 'compared' 'good'
  'high' 'team' 'internal' 'nothing' 'teams' 'many' 'change' 'culture'
  'large' 'big' 'none']
 [ 'pressure' 'work' 'employees' 'working' 'high' 'management' 'company'
  'managers' 'companies' 'low' 'hard' 'lot' 'balance' 'compensation'
  'tech' 'slow' 'difficult' 'growth' 'hours' 'life' 'people' 'get'
  'internal' 'microsoft' 'career' 'team' 'lots' 'politics' 'cons'
  'change' 'pay' 'time' 'sometimes' 'times' 'much' 'dont' 'less' 'teams'
  'nothing' 'large' 'big' 'think' 'many' 'new' 'like' 'really' 'salary'
  'none' 'culture' 'compared']
 [ 'work' 'working' 'employees' 'hours' 'slow' 'balance' 'managers'
  'pressure' 'management' 'lot' 'company' 'lots' 'companies' 'growth'
  'microsoft' 'pay' 'times' 'career' 'tech' 'high' 'compensation' 'hard'
  'difficult' 'less' 'internal' 'sometimes' 'much' 'life' 'people'
  'salary' 'politics' 'get' 'really' 'nothing' 'time' 'low' 'compared'
  'team' 'great' 'bad' 'teams' 'none' 'change' 'new' 'think' 'large'
  'many' 'cons' 'long' 'like']
 [ 'work' 'pressure' 'management' 'working' 'employees' 'managers' 'times'
  'time' 'difficult' 'life' 'career' 'hard' 'hours' 'internal' 'company'
  'balance' 'politics' 'high' 'sometimes' 'companies' 'compensation'
  'get' 'tech' 'culture' 'dont' 'cons' 'team' 'teams' 'nothing' 'bad'
  'pay' 'microsoft' 'slow' 'think' 'like' 'people' 'really' 'growth'
  'lots' 'low' 'change' 'lot' 'compared' 'good' 'much' 'great' 'many'
  'long' 'none' 'less']]
```

ภาพที่ 4.160 แสดงผลลัพธ์แต่ละหัวข้อในกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท microsoft

ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท cons ของบริษัท microsoft ซึ่งมีอยู่ 5 หัวข้อ

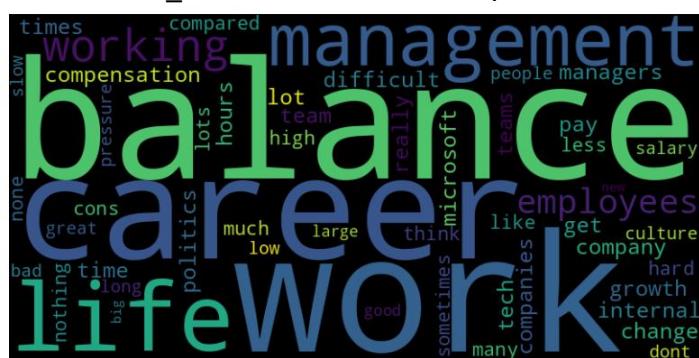
และนี่คือผลลัพธ์จากการนับความถี่ของคำทุกคำที่ปรากฏในแต่ละหัวข้อ

```
Word frequencies: get: 5
balance: 5      time: 5
work: 5         really: 5
career: 5        internal: 5
life: 5          compared: 5
management: 5   high: 5
working: 5       hard: 5
employees: 5    sometimes: 5
compensation: 5 much: 5
managers: 5     less: 5
company: 5       pressure: 4
growth: 5        think: 5
companies: 5    people: 4
pay: 5           low: 4
salary: 4        great: 4
tech: 5          lots: 5
hours: 5         nothing: 5
times: 5         dont: 4
microsoft: 5    like: 5
politics: 5     bad: 4
culture: 4      good: 3
change: 5        cons: 5
lot: 5           none: 5
difficult: 5    long: 4
team: 5          many: 5
teams: 5         slow: 4
                large: 3
                big: 2
                new: 2
```

ภาพที่ 4.161 แสดงผลลัพธ์การนับความถี่ของคำในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท microsoft

นำคำและความถี่ที่นับได้ในกลุ่มที่ 3 ในข้อมูลประเภท cons ของบริษัท microsoft ไปแสดงผลผ่าน word cloud

Group topics 3 – microsoft\_cons all words each topics



ภาพที่ 4.162 word cloud แสดงคำทุกคำที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท cons ของบริษัท microsoft

ผลลัพธ์ของแต่ละหัวข้อในกลุ่มหัวข้อที่ 3 ของข้อมูลประเภท cons ของบริษัท microsoft ซึ่งมีอยู่ 5 หัวข้อแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
[['balance' 'work' 'career' 'life' 'management']
 ['hours' 'work' 'times' 'working' 'time']
 ['pressure' 'work' 'employees' 'working' 'high']
 ['work' 'working' 'employees' 'hours' 'slow']
 ['work' 'pressure' 'management' 'working' 'employees']]
```

ภาพที่ 4.163 แสดงผลลัพธ์เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ  
ในกลุ่มหัวข้อที่ 3 ของเทคนิค Top2vec ของข้อมูลประเภท cons ของบริษัท microsoft

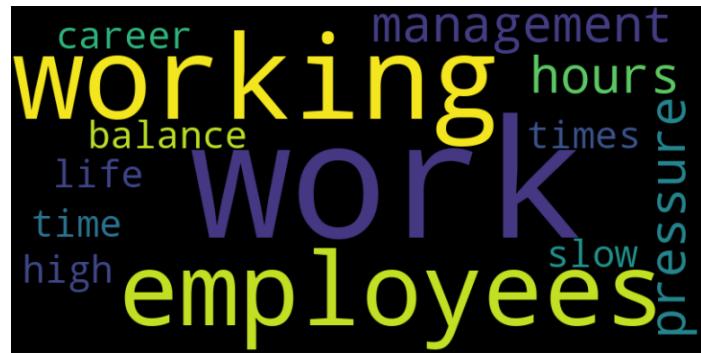
ในการวิเคราะห์ผลลัพธ์หัวข้อผู้วิจัยจะวิเคราะห์ตีความจากคำที่ปรากฏซึ่งแสดงเป็นคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท cons ของบริษัท microsoft ซึ่งมีอยู่ 5 หัวข้อ และนี่คือผลลัพธ์จากการนับความถี่เฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อ

```
Word frequencies:
balance: 1
work: 5
career: 1
life: 1
management: 2
hours: 2
times: 1
working: 4
time: 1
pressure: 2
employees: 3
high: 1
slow: 1
```

ภาพที่ 4.164 แสดงผลลัพธ์การนับความถี่ของคำ 5 อันดับแรกในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3  
ของเทคนิค Top2vec ในข้อมูลประเภท cons ของบริษัท microsoft

คำและความถี่ที่นับได้ในกลุ่มหัวข้อที่ 3 ซึ่งแสดงเฉพาะคำ 5 อันดับแรกในแต่ละหัวข้อของข้อมูลประเภท cons ของบริษัท microsoft ไปแสดงผลผ่าน word cloud

Group topics 3 – microsoft\_cons top 5 words each topics



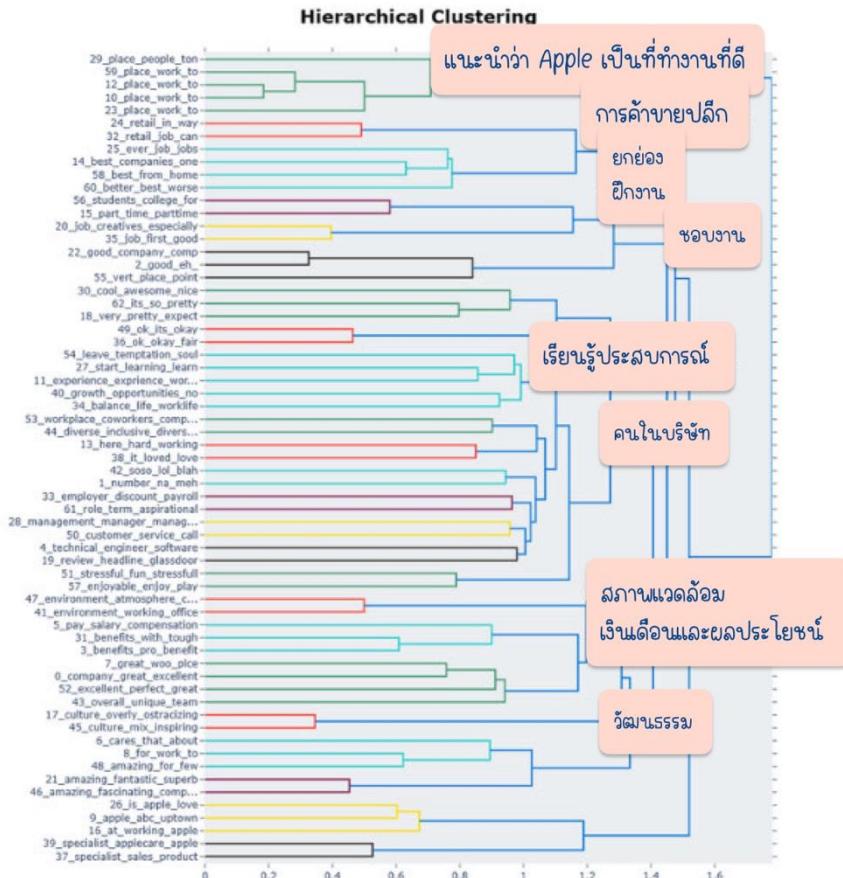
ภาพที่ 4.165 word cloud แสดงคำ 5 อันดับแรกที่ปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ในข้อมูลประเภท cons ของบริษัท microsoft

ผู้วิจัยพิจารณาจาก word cloud แสดงเฉพาะคำ 5 อันดับแรกปรากฏในแต่ละหัวข้อของกลุ่มหัวข้อที่ 3 ของข้อมูลประเภท cons ของบริษัท microsoft ผู้วิจัยมีข้อสันนิษฐานว่าในกลุ่มหัวข้อที่ 3 ของ microsoft ข้อความรีวิวประเภท cons

มีความเกี่ยวข้องกับ: การจัดการพนักงานในการทำงานและชีวิตการทำงาน, ความสมดุล, ความกดดัน, ช่วงเวลาการทำงานที่สูง

## ผลลัพธ์ของการจำแนกกลุ่มข้อความด้วย BERTopic

- ผลการวิเคราะห์หัวคำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล headline บริษัท apple



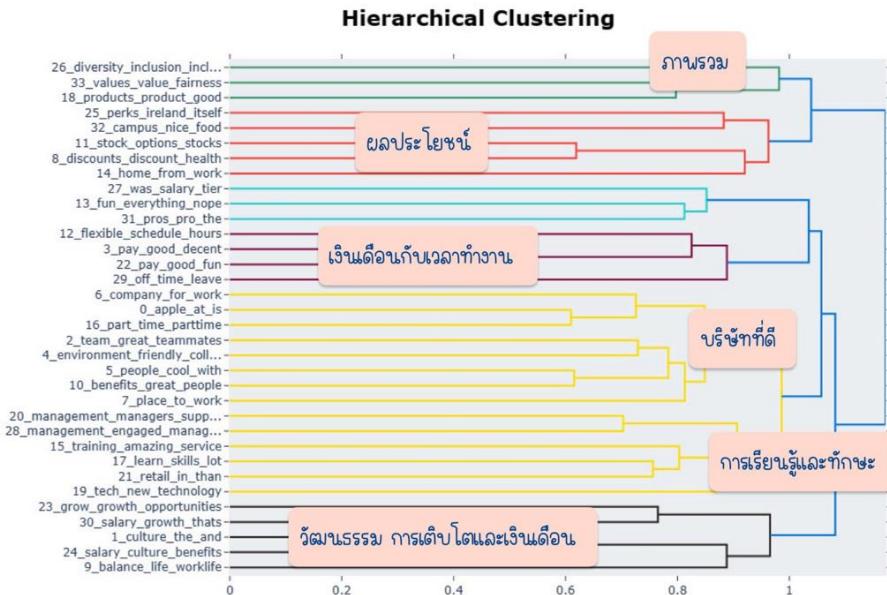
ภาพที่ 4.166 ผลการวิเคราะห์หัวคำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล headline บริษัท apple

ผลการทำนายของแบบจำลองได้วิเคราะห์คำที่มีนัยสำคัญและหัวข้อ โดยจากภาพ เมื่อตั้งสมมุติฐาน และวิเคราะห์คำที่มีนัยสำคัญในแต่ละหัวข้อ พบร่วม

สมมติฐานหัวข้อ	คำที่ปรากฏในหัวข้อ
แนะนำว่าเป็นที่ทำงานที่ดี	Place, work, to, facilities, great, nice, really, one, ton, people, miss, decline
เกี่ยวกับการค้าขายปลีก	Retail, in, way, job, can, as, store, you, best
ยกย่องบริษัท	Best, better, worse, companies, company, the, ever, one, from, home job, jobs, my
ประสบการณ์ฝึกงาน	Students, college, for, parttime, part, time, kids, student
ชوبงาน	Job, creatives, especially, first, good, aight
ได้รับการเรียนรู้ประสบการณ์	Start, learning, learn, experience, growth, opportunities
คนในบริษัท	Diverse, inclusive, diversity, action, lives, workplace, coworkers, competent, community
สภาพแวดล้อม	Environment, atmosphere, classy, supportive, pandemic, working, office, views, largely
เงินเดือนและผลประโยชน์	Pay, salary, benefits, pro, benefit, con, perks, with, company, tough, incredible
วัฒนธรรม	Culture, mix, inspiring, overly, ostracizing, pressure, europeanamerican

ตารางที่ 4.19 ตารางสมมติฐานกลุ่มหัวข้อ headline ของ Apple

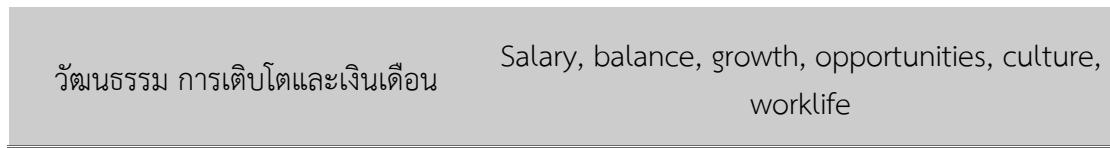
2. ผลการวิเคราะห์หาคำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล รีวิวประโยชน์ที่ได้ บริษัท apple



ภาพที่ 4.167 ผลการวิเคราะห์หาคำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล รีวิวประโยชน์ที่ได้ บริษัท apple

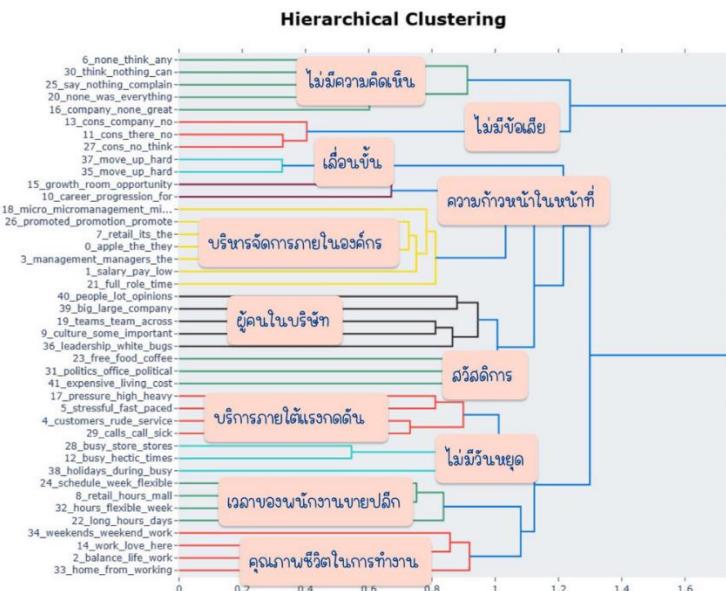
ผลการทำนายของแบบจำลองได้วิเคราะห์คำที่มีนัยสำคัญและหัวข้อ โดยจากภาพ เมื่อตั้งสมมติฐาน และวิเคราะห์คำที่มีนัยสำคัญในแต่ละหัวข้อ พบร่วมกัน

สมมติฐานหัวข้อ	คำที่ปรากฏในหัวข้อ
ภาพรวม	Diversity, inclusion, values, products, product, good
ผลประโยชน์	Perks, ireland, itself, campus, nice, food, stock, option, stocks, discounts, discount, health
เงินเดือนกับเวลาทำงาน	Flexible, schedule, hours, pay, good, fun, decent
เพื่อนร่วมงานที่ดี	Apple, place, to, work, company, cool, great, people, environment, friendly, teammate
การเรียนรู้ทักษะ	learn, skills, lot, management, manager, tech, new, technology, support



ตารางที่ 4.20 ตารางสมมติฐานกลุ่มหัวข้อ ผลประโยชน์ที่ได้ ของ Apple

3. ผลการวิเคราะห์หัวคำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล รีวิวผลเสียที่ได้ บริษัท apple



ภาพที่ 4.168 ผลการวิเคราะห์หัวคำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล รีวิวผลเสียที่ได้ บริษัท apple

ผลการทำนายของแบบจำลองได้วิเคราะห์คำที่มีนัยสำคัญและหัวข้อ โดยจากภาพ เมื่อตั้งสมมติฐาน และวิเคราะห์คำที่มีนัยสำคัญในแต่ละหัวข้อ พบร่วมๆ

สมมติฐานหัวข้อ	คำที่ปรากฏในหัวข้อ
คิดว่าไม่มีผลเสีย	None, nothing , say, complain, cons, think, any, no
เลื่อนขั้น	Move, up, hard, in, company, career, progression, growth opportunity

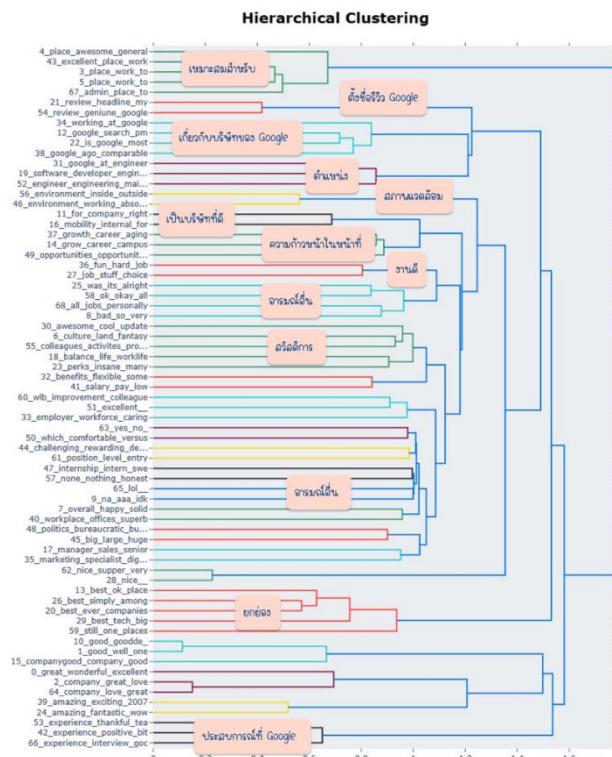
บริหารจัดการองค์กร

Micro, micromanagement, micromanaged, managing, promoted, promotion, promote



ตารางที่ 4.21 ตารางสมมติฐานกลุ่มหัวข้อ ผลเสียที่ได้ ของ Apple

4. ผลการวิเคราะห์หาคำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล headline บริษัท Google



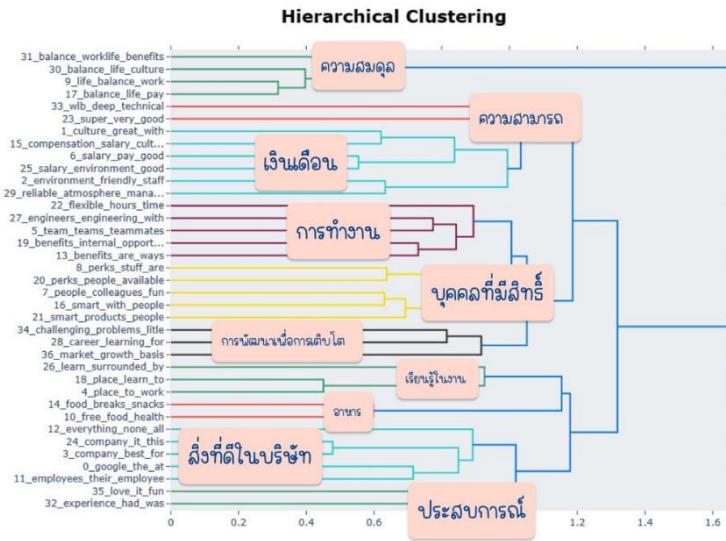
ภาพที่ 4.169 ผลการวิเคราะห์หาคำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล headline บริษัท Google

ผลการทำนายของแบบจำลองได้วิเคราะห์คำที่มีนัยสำคัญและหัวข้อ โดยจากภาพ เมื่อตั้งสมมุติฐาน และวิเคราะห์คำที่มีนัยสำคัญในแต่ละหัวข้อ พบร่วม

สมมติฐานหัวข้อ	คำที่ปรากฏในหัวข้อ
แนะนำว่าเป็นที่ทำงานที่ดี	Place, work, to, awesome, great, company ,general
ตำแหน่ง	Engineer, engineering, google, at software, developer, manager
ยกย่องบริษัท	Company, right, mobility, internal, for
ความก้าวหน้าในหน้าที่	Growth, grow, career, aging, campus, opportunities, opportunity
เกี่ยวกับงาน	Fun, hard, job, stuff, choice
สวัสดิการ	Awesome, cool, update, culture, land, fantasy, colleagues, activities, balance, life, worklife, perks, insane, many
ประสบการณ์ google	Experiences, positive, bit, interview, goc, them, working

ตารางที่ 4.22 ตารางสมมติฐานกลุ่มหัวข้อ headline ของ Google

5. ผลการวิเคราะห์หัวคำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล รีวิวประโยชน์ที่ได้ บริษัท Google



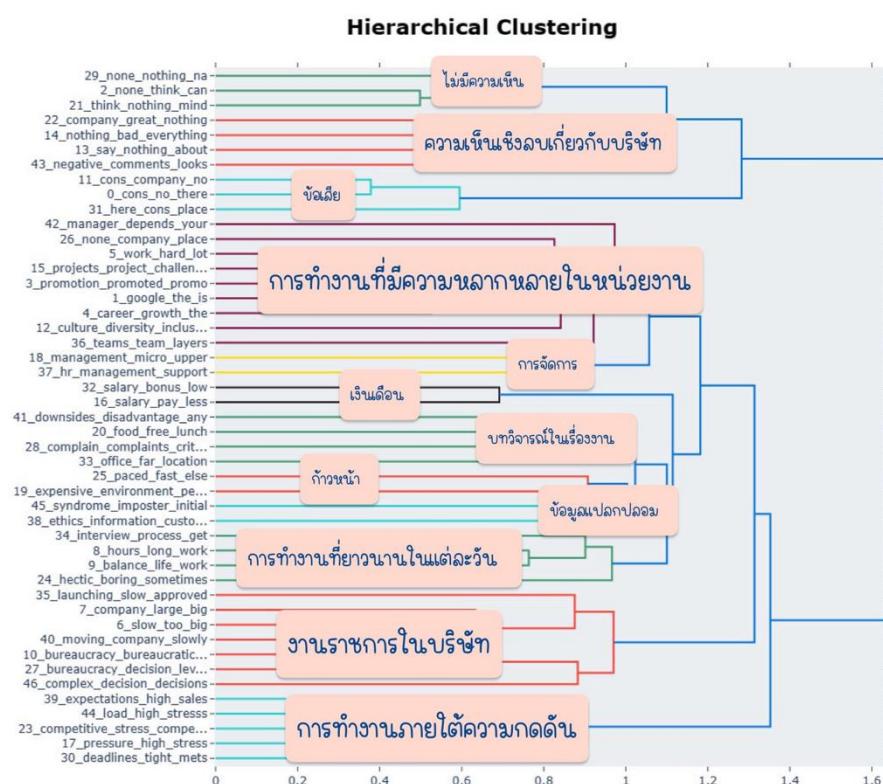
ภาพที่ 4.170 ผลการวิเคราะห์หัวคำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล รีวิวประโยชน์ที่ได้ บริษัท Google

ผลการทำนายของแบบจำลองได้วิเคราะห์คำที่มีนัยสำคัญและหัวข้อ โดยจากการ เมื่อตั้งสมมติฐาน และวิเคราะห์คำที่มีนัยสำคัญในแต่ละหัวข้อ พบว่า

สมมติฐานหัวข้อ	คำที่ปรากฏในหัวข้อ
ความสมดุล	Balance, worklife, benefits, life, culture, pay
ความสามารถ	Work life balance (wlb), deep, technical, super, very, good
เงินเดือน	Salary, culture, great, with, compensation, pay, good, environment, friendly, staff
การทำงาน	Flexible, hours, time, engineers, engineering, with, team, teams, teammates, benefits, internal
สิทธิพิเศษ	Perks, people, stuff, are, available, colleague, fun, smart, with, products

ได้รับการเรียนรู้ประสบการณ์	Challenging, learning, problems, career, for
อาหาร	Food, breaks, snacks, free, health
ตารางที่ 4.23 ตารางสมมติฐานกลุ่มหัวข้อ ประโยชน์ที่ได้ ของ Google	

6. ผลการวิเคราะห์หัวคำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล รีวิวผลเสียที่ได้ บริษัท Google



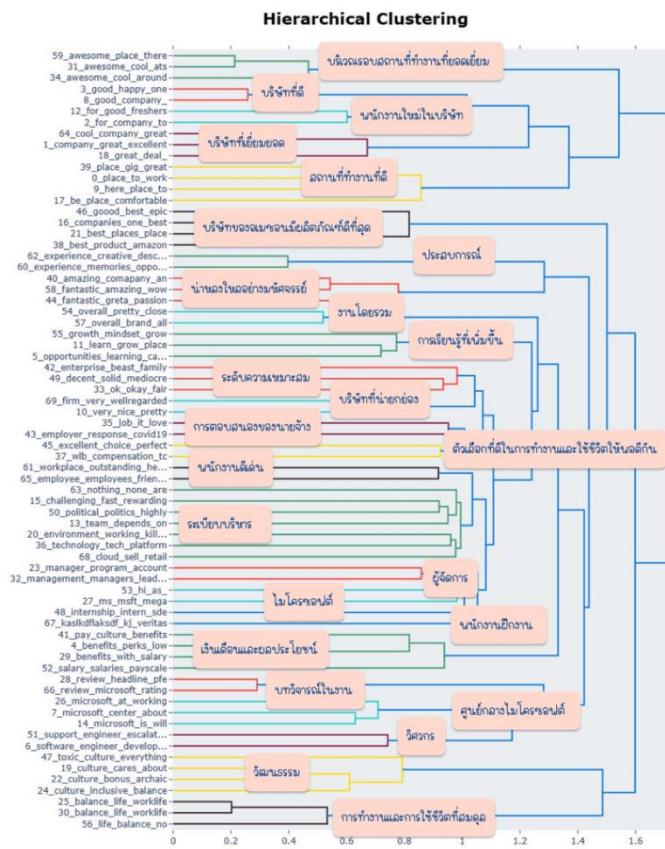
ภาพที่ 4.171 ผลการวิเคราะห์หัวคำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล รีวิวผลเสียที่ได้ บริษัท Google

ผลการทำนายของแบบจำลองได้วิเคราะห์คำที่มีนัยสำคัญและหัวข้อ โดยจากภาพ เมื่อตั้งสมมติฐาน และวิเคราะห์คำที่มีนัยสำคัญในแต่ละหัวข้อ พบร่วม

สมมติฐานหัวข้อ	คำที่ปรากฏในหัวข้อ
คิดว่าไม่มีผลเสีย	Cons, none, nothing, negative, comment, bad, everything, looks, company, think, place, say
โดยรวม	Career, promoted, promotion, promo, projects, project, work, hard, culture, diversity, inclusion
การจัดการ	Management, micro, upper, hr, support
เงินเดือน	Salary, bonus, low, pay, less
ทำงานหนัก	Work, interview, process, get, hours, long, balance, life
ระเบียบบริหาร	Bureaucracy, bureaucratic, decision, company, large, big, complex, decisions
สวัสดิการ	Food, free, lunch, complain, complains
มีความ เครียด กดดัน	Stress, expectations, high, sales, load, competitive, pressure, tight

ตารางที่ 4.24 ตารางสมมติฐานกลุ่มหัวข้อ ผลเสียที่ได้ ของ Google

## 7. ผลการวิเคราะห์หัวคำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล headline บริษัท Microsoft



ภาพที่ 4.172 ผลการวิเคราะห์หัวคำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล headline บริษัท Microsoft

ผลการทำนายของแบบจำลองได้วิเคราะห์คำที่มีนัยสำคัญและหัวข้อ โดยจากภาพ เมื่อตั้งสมมติฐาน และวิเคราะห์คำที่มีนัยสำคัญในแต่ละหัวข้อ พบว่า

สมมติฐานหัวข้อ	คำที่ปรากฏในหัวข้อ
แนะนำว่าเป็นที่ทำงานที่ดี	Awesome, place, there, cool, around, good, happy, one, company, great, excellent, deal, to, work, here, firm, very
ประสบการณ์ที่ทำงานกับ Microsoft	Microsoft, experience, creative, describe, word, horrible, fantastic, amazing

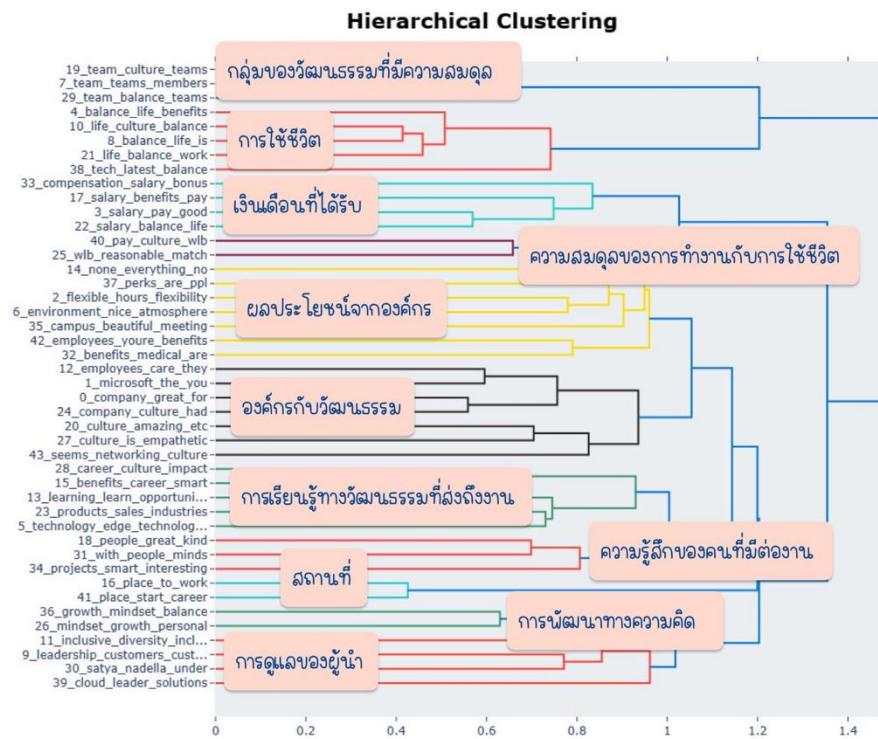
โดยรวม

Overall, brand, all, pretty, close, headstrong

เรียนรู้ เติบโต	Growth, grow, learn, mindset, place, opportunities, learning
ตำแหน่ง	Manager, program, account, management, internship, intern, engineer, support, software
เงินเดือนและสิทธิพิเศษ	Benefits, pay, culture, perks, low, with, salary
วัฒนธรรม	Culture, toxic, everything, cares, about, bonus, archaic, inclusive, balance

ตารางที่ 4.25 ตารางสมมุติฐานกลุ่มหัวข้อ headline ของ Microsoft

8. ผลการวิเคราะห์ห้าคำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล รีวิวประโยชน์ที่ได้ บริษัท Microsoft



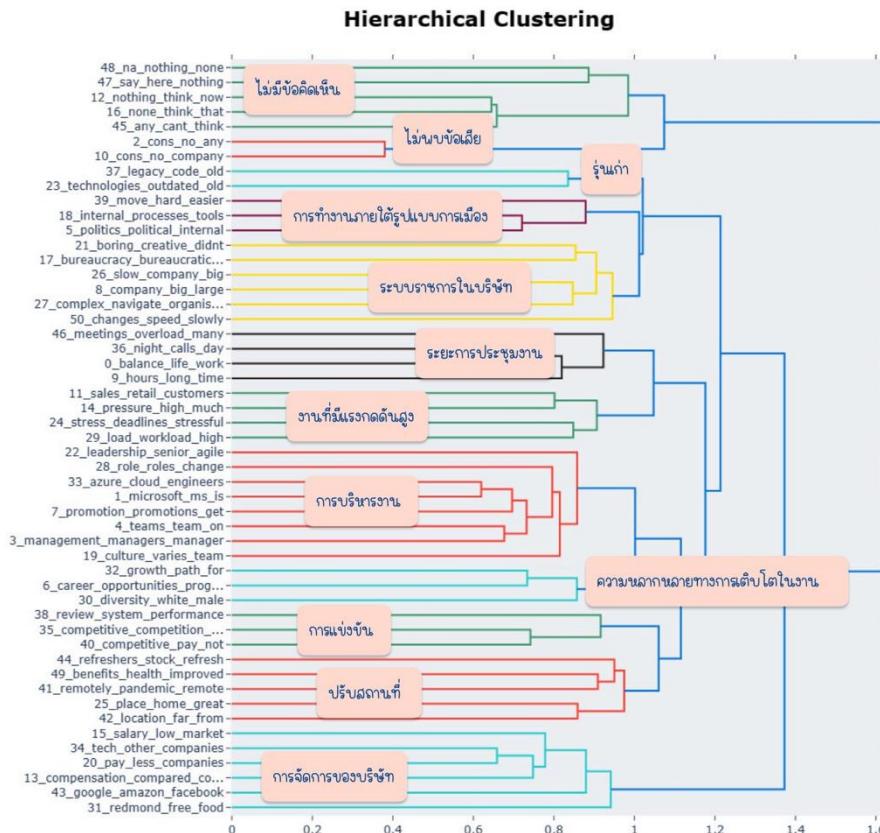
ภาพที่ 4.173 ผลการวิเคราะห์ห้าคำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล รีวิวประโยชน์ที่ได้ บริษัท Microsoft

ผลการทนายของแบบจำลองได้วิเคราะห์คำที่มีนัยสำคัญและหัวข้อ โดยจากภาพ เมื่อตั้งสมมุติฐาน และวิเคราะห์คำที่มีนัยสำคัญในแต่ละหัวข้อ พบร่วมๆ

สมมติฐานหัวข้อ	คำที่ปรากฏในหัวข้อ
เพื่อร่วมทีม	Team, teams, culture, members, balance
คุณภาพชีวิต	Balance, life, benefits, culture, is, work
เงินเดือน	Salary, compensation, bonus, benefits, pay, good, balance, life
ผลตอบแทนที่ดี	Benefit, career, learnning, learn, opportunity
องค์กร	Company, employees, care, they, Microsoft, the, you, great, for, culture, had, amazing, is, empathetic, seem, networking
การเรียนรู้	Career, culture, smart, impact, benefits, learning, learn, products, sales, industries, technology, edge
พัฒนาทางความคิด	Growth, mindset, balance, personal

ตารางที่ 4.26 ตารางสมมติฐานกลุ่มหัวข้อ ผลประโยชน์ที่ได้ของ Microsoft

9. ผลการวิเคราะห์หาคำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล รีวิวผลเสียที่ได้ บริษัท Microsoft



ภาพที่ 4.174 ผลการวิเคราะห์หาคำที่มีนัยสำคัญและหัวข้อของแบบจำลองข้อมูล รีวิวผลเสียที่ได้ บริษัท Microsoft

ผลการนำযາของแบบจำลองได้วิเคราะห์คำที่มีนัยสำคัญและหัวข้อ โดยจากภาพ เมื่อตั้งสมมติฐาน และวิเคราะห์คำที่มีนัยสำคัญในแต่ละหัวข้อ พบร่วมๆ

สมมติฐานหัวข้อ	คำที่ปรากฏในหัวข้อ
คิดว่าไม่มีผลเสีย	Nothing, na, none, say, here, think, now, that, cons, no, any, company
เทคโนโลยีล้าสมัย	Old, legacy, code, technologies, outdated
และเบี้ยงบบริหาร	Bureaucacy, bureaucratic, slow, company,big large, com plex, navigate, organisation

ประชุมตลอดเวลา	Meeting, overload, many, night, calls, day, balance, life, work, hours, long, time
รับแรงกดดัน	Pressure, stress, sales, retail, customers, high, much, deadlines, workload, stressful, load, high
การเลื่อนขั้น	Career, growth, path, promotion, promotions
การแข่งขัน	Competitive, review, system, performance, competition, pay

ตารางที่ 4.27 ตารางสมมติฐานกลุ่มหัวข้อ ผลเสียที่ได้ของ Microsoft

### ผลลัพธ์ของการประเมินประสิทธิภาพเชิงปริมาณ

Topic Coherence : Coherence score									
LDA	apple			google			microsoft		
	headline	pros	cons	headline	pros	cons	headline	pros	cons
LDA	0.603	0.441	0.433	0.603	0.445	0.422	0.596	0.413	0.423
BERTopic	0.375	0.406	0.371	0.394	0.377	0.346	0.364	0.377	0.368
Top2vec	0.166	0.237	0.282	0.272	0.220	0.349	0.322	0.221	0.368

ภาพที่ 4.175 ตารางผลลัพธ์ของการประเมินประสิทธิภาพเชิงปริมาณ

ตัววัดที่ใช้ประเมินคือค่า Topic Coherence Cv หรือ Cv ตามที่ได้อธิบายไปในบทที่ 2 โดยที่เป็นค่าที่คำนวนออกมาได้จะอยู่ในช่วง [0,1] ถ้าค่ายิ่งสูงแสดงถึงประสิทธิภาพที่ดีมากตามหมายความว่าคำหลายคำที่ปรากฏภายในหัวข้อนั้นมีความสัมพันธ์กันอยู่จริงในเอกสารหรือชุดเอกสาร โดยค่าแต่ละค่าที่แสดงในภาพที่ 4.175 คือค่าเฉลี่ยของ Cv ของทุกผลลัพธ์หัวข้อที่คำนวนได้แยกตามประเภทข้อมูลและบริษัทในแต่ละเทคนิค จากการประเมินประสิทธิภาพเชิงปริมาณของแบบจำลองหัวข้อทั้ง 3 เทคนิคในแต่ละบริษัท แสดงให้เห็นว่า เทคนิค LDA ทำค่า Cv โดยเฉลี่ยสูงเป็นอันดับ 1 รองลงมาเป็น BERTopic และ Top2vec ตามลำดับ

## บทที่ 5

### สรุป

#### 5.1 สรุปผลการทดลอง

โครงงานนี้เป็นการศึกษาวิธีการวิเคราะห์หาคำที่มีนัยสำคัญและหัวข้อในข้อความและการจัดกลุ่มข้อความที่มีลักษณะคล้ายคลึงกันอยู่ในกลุ่มเดียวกัน ด้วยเทคนิค 3 รูปแบบ คือ การจัดสรรเดรีเคลฟอง (Latent Dirichlet Allocation) , Top2Vec และ BERTopic ซึ่งให้เห็นถึงพนักงานบริษัทต่างๆที่เข้าร่วมให้กับ Glassdoor

โดยผู้วิจัยมีแนวทางการสรุปผลลัพธ์หัวข้อทั้ง 3 บริษัท โดยจะสรุปเป็นกลุ่มหัวข้อที่ได้ทำการตีความและจัดกลุ่มแล้ว ซึ่งผู้วิจัยมีข้อสันนิษฐานว่า หากกลุ่มหัวข้อที่ผู้วิจัยได้จัดกลุ่มแล้ว จากผลลัพธ์หัวข้อของทั้ง 3 เทคนิค LDA, Top2vec, BERTopic ในแต่ละเทคนิคก็กลุ่มหัวข้อที่คล้ายคลึงกันก็จะนำมาสรุป เป็นกลุ่มสมมติฐานเดียวกัน หมายความว่ากลุ่มหัวข้อนั้นมีความหมายคล้ายคลึงกันจึงสรุปเป็นกลุ่มสมมติฐานเดียวกัน

##### 5.1.1 สรุปกลุ่มหัวข้อที่คล้ายคลึงในแต่ละเทคนิคแบ่งตามบริษัทและประเภทข้อมูล

###### รีวิวหัวเรื่องของบริษัท Apple

สรุปกลุ่มสมมติฐาน	ตัวอย่างของคำที่พบได้มากในกลุ่มหัวข้อ
กลุ่มหัวข้อแนะนำว่าเป็น ที่ทำงานที่ดี	Great, amazing, best, company, work, place
กลุ่มหัวข้อประสบการณ์ของทำงาน ที่บริษัท	Experience
ตารางที่ 5.1 กลุ่มหัวเรื่องที่คล้ายคลึงในแต่ละเทคนิคของ Apple	

###### รีวิวประโยชน์ที่ได้ของบริษัท Apple

สรุปกลุ่มสมมติฐาน	ตัวอย่างของคำที่พบได้มากในกลุ่มหัวข้อ
กลุ่มหัวข้องค์กร	Employees, company, career
กลุ่มหัวข้อworklifebalance	job
กลุ่มหัวข้อสวัสดิการ	Benefits, opportunities
ตารางที่ 5.2 กลุ่มประโยชน์ที่คล้ายคลึงในแต่ละเทคนิคของ Apple	

###### รีวิวผลเสียที่ได้ของบริษัท Apple

สรุปกลุ่มสมมติฐาน	ตัวอย่างของคำที่พบได้มากในกลุ่มหัวข้อ
กลุ่มหัวข้อมีผลเสีย	Don't, really, cons, bad, always
กลุ่มหัวข้องค์กร	Employees, company, management, managers
กลุ่มหัวข้อworklifebalance	Job, work
ตารางที่ 5.3 กลุ่มผลเสียที่คล้ายคลึงในแต่ละเทคนิคของ Apple	

รีวิวหัวเรื่องของบริษัท Google

สรุปกลุ่มสมมติฐาน	ตัวอย่างของคำที่พบได้มากในกลุ่มหัวข้อ
กลุ่มหัวข้อแนะนำว่าเป็นที่ทำงานที่ดี	Great, amazing, best, company, work, place
กลุ่มหัวข้อประสบการณ์ของทำงานที่บริษัท	Experience, culture, people, nice, work
ตารางที่ 5.4 กลุ่มหัวข้อหัวเรื่องที่คล้ายคลึงในแต่ละเทคนิคของ Google	

รีวิวประโยชน์ที่ได้ของบริษัท Google

สรุปกลุ่มสมมติฐาน	ตัวอย่างของคำที่พบได้มากในกลุ่มหัวข้อ
กลุ่มหัวข้อวัฒนธรรม	Culture, employees, company, colleagues, coworkers
กลุ่มหัวข้อเงินเดือน	salary, pay
กลุ่มหัวข้อสวัสดิการ	Opportunities, compensation, perks, benefits
กลุ่มหัวข้อWorklifebalance	Career, work, working, balance
กลุ่มหัวข้อการเรียนรู้	Learn, smart
ตารางที่ 5.5 กลุ่มประโยชน์ที่คล้ายคลึงในแต่ละเทคนิคของ Google	

รีวิวผลเสียที่ได้ของบริษัท Google

สรุปกลุ่มสมมติฐาน	ตัวอย่างของคำที่พบได้มากในกลุ่มหัวข้อ
กลุ่มหัวข้อมิ่มีผลเสีย	nothing, none, think, bad, really
กลุ่มหัวข้อองค์กร	Management, politics, Career, growth, company
กลุ่มหัวข้อworklifebalance	job, working, life, process, hard, slow
กลุ่มหัวข้อช้าในการทำงาน	Long, time, hours
ตารางที่ 5.6 กลุ่มผลเสียที่คล้ายคลึงในแต่ละเทคนิคของ Google	

รีวิวหัวเรื่องของบริษัท Microsoft

สรุปกลุ่มสมมติฐาน	ตัวอย่างของคำที่พบได้มากในกลุ่มหัวข้อ
กลุ่มหัวข้อแนะนำว่าเป็นที่ทำงานที่ดี	Excellent, great, amazing, best, company, work, place
กลุ่มหัวข้อประสบการณ์ของทำงานที่บริษัท	experience, culture
ตารางที่ 5.7 กลุ่มหัวข้อหัวเรื่องที่คล้ายคลึงในแต่ละเทคนิคของ Microsoft	

รีวิวประโยชน์ที่ได้ของบริษัท Microsoft

สรุปกลุ่มสมมติฐาน	ตัวอย่างของคำที่พบได้มากในกลุ่มหัวข้อ
กลุ่มหัวข้อวัฒนธรรม	Culture, diversity, company, environment, people, management
กลุ่มหัวข้อเงินเดือน	Salary, pay
กลุ่มหัวข้อสวัสดิการ	Compensation, lots, benefits
กลุ่มหัวข้อWorklifebalance	Projects, work, balance, flexible
กลุ่มหัวข้อการเรียนรู้	Learn, technology

ตารางที่ 5.8 กลุ่มประโยชน์ที่คล้ายคลึงในแต่ละเทคนิคของ Microsoft

รีวิผลเสียที่ได้ของบริษัท Microsoft

สรุปกลุ่มสมมติฐาน	ตัวอย่างของคำที่พบได้มากในกลุ่มหัวข้อ
กลุ่มหัวข้อองค์กร	Company, employees, managers, growth, politics, internal
กลุ่มหัวข้อworklifebalance	work, life, balance
กลุ่มหัวข้อแรงกดดัน	pressure
กลุ่มหัวข้อช่วงเวลาการทำงาน	Long, times, hours
กลุ่มหัวข้อเงินเดือน	Salary, pay, slow

ตารางที่ 5.9 กลุ่มผลเสียที่คล้ายคลึงในแต่ละเทคนิคของ Microsoft

เมื่อพิจารณาพบว่าการตั้งสมมติฐานเบื้องต้นและผลการทำนายทั้ง 3 บริษัท พบร่วมกับงานบริษัท ต่างๆ ที่เข้าร่วม 3 ส่วน คือ หัวเรื่อง ผลประโยชน์ที่ได้รับ ผลเสียที่ได้รับ โดยมีเกี่ยวกับเรื่องที่พูดในทิศทางเดียวกันทั้ง 3 บริษัท

หัวเรื่อง

- กลุ่มหัวข้อแนะนำว่าเป็นที่ทำงานที่ดี
- กลุ่มหัวข้อประสบการณ์ทำงานที่บริษัท

ผลประโยชน์ที่ได้รับ

- กลุ่มหัวข้อวัฒนธรรม
- กลุ่มหัวข้อ worklifebalance
- กลุ่มหัวข้อ สวัสดิการ

ผลเสียที่ได้รับ

- กลุ่มหัวข้อองค์กร
- กลุ่มหัวข้อworklifebalance

### 5.1.2 สรุปผลการประเมินประสิทธิภาพเชิงคุณภาพของแบบจำลองหัวข้อในแต่ละเทคนิค

โดยเบื้องต้นได้พิจารณาจากผลลัพธ์หัวข้อทั้ง 3 เทคนิคผู้วิจัยมีความเห็นว่าเทคนิค LDA กับ BERTopic เมน้ำสมกับ data รีวิวการทำงานภายบริษัทต่างๆ ในชุดข้อมูล Glassdoor นี้ เพราะสามารถตีความและตั้งสมมติฐานได้ไม่ยากลำบากนักและให้กลุ่มผลลัพธ์หัวข้อที่คล้ายคลึงกันในสองเทคนิคดังกล่าว

เนื่องจากข้อความใน data รีวินี้ซึ่งผู้รีวิวแต่ละคนพูดไปในทางทำนองเดียวกัน ทำให้ผลลัพธ์ของ Top2Vec คำที่นำเสนอในแต่ละผลลัพธ์หัวข้อนั้นซ้ำกันอย่างมากทั้งในความหมายและรูปคำ ส่งผลให้การวิเคราะห์หรือตั้งสมมติฐานได้ยากลำบากผู้วิจัยจึงมีข้อสรุปเกี่ยวกับเทคนิค top2vec ว่าอาจจะไม่เหมาะสมสมกับ data นี้(ข้อความรีวิวของชุดข้อมูล Glassdoor)

หากต้องเลือกมา 1 แบบจำลองใน 3 เทคนิคดังกล่าว วิจัยมีความเห็นว่าจะเลือกใช้เทคนิคการจัดตีรีเคลฟ่ง (Latent Dirichlet Allocation) เพราะ data รีวิวที่ผู้วิจัยเลือกใช้ไม่ได้ซับซ้อนและเพียงต้องการจะสืบอุอกมาเป็นรูปของแต่ละบริษัท ซึ่งผลลัพธ์ของ BERTopic จะแสดงหัวข้อไว้เป็นจำนวนมากโดยความหมายแต่ละหัวข้อจะมีความคล้ายคลึงกัน ผู้วิจัยจึงมีความเห็นว่าควรเลือกใช้เทคนิคการจัดตีรีเคลฟ่ง (Latent Dirichlet Allocation) กับ data รีวิวของชุดข้อมูล Glassdoor นี้ ซึ่งผู้ที่จะศึกษาในเรื่องนี้ควรจะพิจารณากำหนดจำนวนหัวข้อให้ลดลง ตามความเหมาะสมและความสนใจของผู้ที่จะศึกษา

## รายการอ้างอิง

- [1] วรอัปสร บุบพา และเอกสาร รัฐกานจน. (2565). การสร้างแบบจำลองหัวข้อ (Topic Modeling) และวิเคราะห์ความรู้สึก (Sentiment Analysis) ของผู้ใช้บริการศูนย์ซื้อขายสินทรัพย์ดิจิทัล (Digital Asset Exchange) กรณีศึกษาเพจเฟซบุ๊คศูนย์ซื้อขายสินทรัพย์ดิจิทัลในประเทศไทย.  
สถาบันบัณฑิตพัฒนบริหารศาสตร์. เข้าถึงจาก,  
[https://as.nida.ac.th/wp-content/uploads/2021/11/6210422030\\_Woraupsorn\\_Paper-IS.pdf](https://as.nida.ac.th/wp-content/uploads/2021/11/6210422030_Woraupsorn_Paper-IS.pdf)
- [2] ณิชาภา รัตนปรีชาเวช (2562). ความสัมพันธ์ระหว่างประเด็นต่างๆ  
ที่ถูกพูดถึงในโซเชียลมีเดียกรณีศึกษา เว็บ บอร์ดกลุ่มแม่และเด็ก. สถาบันบัณฑิตพัฒนบริหารศาสตร์
- [3] ปิยารณ พองพโลย (2563). การวิเคราะห์หัวข้อความภาษาไทยเกี่ยวกับการตั้งครรภ์ด้วยวิธีการ  
สร้างแบบจำลองหัวข้อ [วิทยานิพันธ์ระดับบัณฑิตศึกษา] คณะวิทยาศาสตร์ สาขาวิชาการข้อมูล  
มหาวิทยาลัยศรีนครินทรวิโรฒ. เข้าถึงจาก,  
<http://ir-thesis.swu.ac.th/dspace/bitstream/123456789/1243/1/gs621130237.pdf>

### แหล่งค้นคว้าจากอินเทอร์เน็ต

- [4] Nini G. (2018). วิถีประเมินประสิทธิภาพของโมเดล. สืบค้นเมื่อ 9 ตุลาคม 2566. จาก  
[https://medium.com/@gunanini784/%E0%B8%A7%E0%B8%B4%E0%B8%96%E0%B8%95%E0%B8%9B%E0%B8%A3%E0%B8%80%E0%B9%80%E0%B8%A1%E0%B8%B4%E0%B8%99%E0%B8%9B%E0%B8%A3%E0%B8%80%E0%B8%80%E0%B8%AA%E0%B8%B4%E0%B8%97%E0%B8%98%E0%B8%B4%E0%B8%A0%E0%B8%B2%E0%B8%9E%E0%B8%82%E0%B8%AD%E0%B8%87%E0%B9%82%E0%B8%A1%E0%B9%80%E0%B8%94%E0%B8%A5-cf98823a1a02](https://medium.com/@gunanini784/%E0%B8%A7%E0%B8%B4%E0%B8%96%E0%B8%95%E0%B8%9B%E0%B8%A3%E0%B8%80%E0%B9%80%E0%B8%A1%E0%B8%B4%E0%B8%99%E0%B8%9B%E0%B8%A3%E0%B8%80%E0%B8%AA%E0%B8%B4%E0%B8%97%E0%B8%98%E0%B8%B4%E0%B8%A0%E0%B8%B2%E0%B8%9E%E0%B8%82%E0%B8%AD%E0%B8%87%E0%B9%82%E0%B8%A1%E0%B9%80%E0%B8%94%E0%B8%A5-cf98823a1a02)
- [5] Sagar Pundir. TOP2VEC: New way of topic modelling. (2020). สืบค้นเมื่อ 12 ตุลาคม 2566. จาก  
<https://towardsdatascience.com/top2vec-new-way-of-topic-modelling-bea165eeac4a>
- [6] BERTopic. BERTopic. สืบค้นเมื่อ 17 ตุลาคม 2566. จาก  
<https://maartengr.github.io/BERTopic/>

- [7] Surapong Kanoktipsatharporn. (2020). Part of Speech Tagging คืออะไร และ Named-Entity Recognition / Tagging คืออะไร สอน POS Tagging, NER ภาษาไทย – PyThaiNLP ep.4. สืบคันเมื่อ 20 ตุลาคม 2566. จาก <https://www.bualabs.com/archives/4112/what-is-part-of-speech-tagging>
- [8] CHAKRIT. (2019). tf-idf-ทำงานยังไง. สืบคันเมื่อ 2 พฤษภาคม 2566. จาก <https://www.softnix.co.th/2019/05/28/tf-idf-ทำงานยังไง/>
- [9] Emil Rijcken. (2023). Cv Topic Coherence Explained. สืบคันเมื่อ 6 พฤษภาคม 2566. จาก <https://towardsdatascience.com/c%E1%B5%A5-topic-coherence-explained-fc70e2a85227>
- [10] Eugenia Anello. (2023). Topic Modeling Approaches: Top2Vec vs BERTopic. สืบคันเมื่อ 8 พฤษภาคม 2566. จาก <https://www.kdnuggets.com/2023/01/topic-modeling-approaches-top2vec-bertopic.html>
- [11] Roman Egger, Joanne Yu. (2022). A Topic Modeling Comparison Between LDA, NMF, Top2Vec, and BERTopic to Demystify Twitter Posts. สืบคันเมื่อ 10 พฤษภาคม 2566 จาก <https://www.frontiersin.org/articles/10.3389/fsoc.2022.886498/full#T4>
- [12] Sasiwit Chaiyadecha. (2021). News classification #2: รู้จักกับ Word embedding และ Word2Vec. สืบคันเมื่อ 13 พฤษภาคม 2566. จาก <https://lengyi.medium.com/word-embedding-word2vec-nlp-model-dbc4c892dfb9>
- [13] lukkidd. (2018). Word Embedding และ Word2Vec คืออะไร. สืบคันเมื่อ 14 พฤษภาคม 2566. จาก <https://lukkidd.com/word-embedding-%E0%B9%81%E0%B8%A5%E0%B8%B0-word2vec-%E0%B8%84%E0%B8%B7%E0%B8%AD%E0%B8%AD%E0%B8%B0%E0%B9%84%E0%B8%A3-e60bdf6d78d3>