

การสรุปใจความสำคัญของข้อความแบบสกดสำหรับชาวท่องเที่ยวภาษาไทย



น.ส.ศรัญญา นาทองห่อ

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาวิศวกรรมซอฟต์แวร์ ภาควิชาวิศวกรรมคอมพิวเตอร์

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2563

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

Extractive Text Summarization for Thai Travel News



A Thesis Submitted in Partial Fulfillment of the Requirements  
for the Degree of Master of Science in Software Engineering

Department of Computer Engineering

FACULTY OF ENGINEERING

Chulalongkorn University

Academic Year 2020

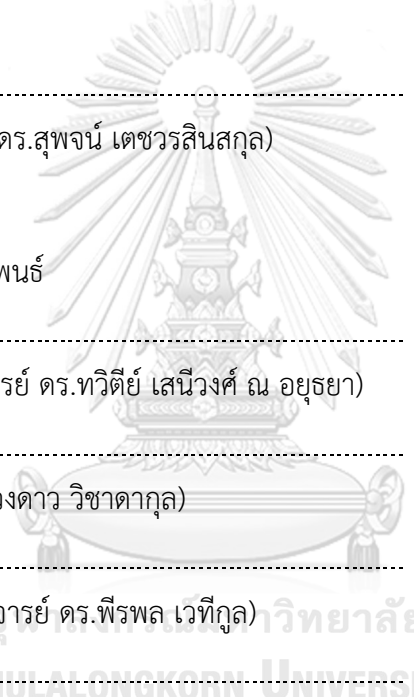
Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	การสรุปใจความสำคัญของข้อความแบบสก็ดสำหรับข่าว ท้องถิ่นภาษาไทย
โดย	น.ส.ศรัญญา นาทองห่อ
สาขาวิชา	วิศวกรรมซอฟต์แวร์
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	อาจารย์ ดร.ดวงดาว วิชาตากุล

---

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้รับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่ง  
ของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

.....	คณบดีคณะวิศวกรรมศาสตร์ (ศาสตราจารย์ ดร.สุพจน์ เตชวรสินสกุล)
คณะกรรมการสอบวิทยานิพนธ์	ประธานกรรมการ (รองศาสตราจารย์ ดร.ทวีชัย เสนิงค์ ณ อยุธยา)
.....	อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก (อาจารย์ ดร.ดวงดาว วิชาตากุล)
.....	กรรมการ (ผู้ช่วยศาสตราจารย์ ดร.พีรพล เวทีกุล)
.....	กรรมการภายนอกมหาวิทยาลัย (รองศาสตราจารย์ ดร.เบญจพร ลิ้มธรรมาภรณ์)



ศรัญญา นาทองห่อ : การสรุปใจความสำคัญของข้อความแบบสกัดสำหรับข่าวท่องเที่ยว  
ภาษาไทย. ( Extractive Text Summarization for Thai Travel News) อ.ที่ปรึกษา  
หลัก : อ. ดร.ดวงดาว วิชาตากุล

ปัจจุบันเทคโนโลยีทางด้านคอมพิวเตอร์มีความสำคัญต่อการดำเนินชีวิตประจำวันของ  
มนุษย์เป็นอย่างมากและยังถือว่าเป็นเครื่องมือที่ใช้ในการอำนวยความสะดวกให้แก่มนุษย์มากมาย  
โดยเฉพาะทางด้านการสื่อสารผ่านสังคมออนไลน์ เพื่อลดเวลาในการอ่านข่าวหรืออ่านบทความ  
และข่าวออนไลน์ต่างๆ จากการวิจัยที่ผ่านมาได้มีการศึกษาและพัฒนาการสรุปใจความสำคัญของ  
ภาษาไทยเป็นจำนวนมาก ในงานวิจัยนี้ได้นำเสนอวิธีการสรุปใจความสำคัญจากข่าวการท่องเที่ยว  
ภาษาไทย 2 วิธีคือการเลือกประโยคจากการจัดกลุ่มประโยคด้วยเคมีนและการเลือกประโยคด้วยวิธี  
หาค่าสำคัญประโยคจากหัวข้อข่าว โดยมีการพัฒนาและสร้างคลังข้อมูลรายการคำประสมเพื่อช่วย  
เพิ่มประสิทธิภาพในการตัดคำ โดยการทดลองนี้ใช้ข้อมูลข่าวการท่องเที่ยวไทย ทั้งหมด 400 ข่าว  
สำหรับใช้ทดลองในการสรุปใจความสำคัญ และ 5,000 ข่าวสำหรับการสร้างคลังข้อมูลรายการคำ  
ประสม การวัดประสิทธิภาพของวิธีการที่นำเสนอ มีการวัดประสิทธิภาพการสรุปใจความสำคัญโดย  
การเปรียบเทียบผลจากการสรุปที่ได้จากผู้เชี่ยวชาญด้านภาษาไทยเทียบกับผลสรุปที่ได้จากวิธีการที่  
นำเสนอ จากงานวิจัยนี้ในขั้นตอนการสร้างคำประสมได้คำประสมทั้งหมด จำนวน 2,340 คำ ผล  
การทดลองพบว่าวิธีตัดคำด้วยคัตคำร่วมกับตัดคำประสมได้ผลดีกว่าการตัดคำจากคัตคำเพียงอย่าง  
เดียว และการสรุปใจความสำคัญโดยใช้การคำนวณค่าน้ำหนักของคำสำคัญโดยหาค่าความถี่ของคำ  
จากหัวข้อข่าวเพียงอย่างเดียวและเลือกประโยคเรียงลำดับจากผลรวมความถี่ของคำสำคัญจาก  
หัวข้อข่าวมีประสิทธิภาพและความแม่นยำสูงสุดโดยมีค่าความแม่นยำ ค่าความระลึกและค่าวัด  
ประสิทธิภาพอยู่ที่ 0.8097 0.8367 และ 0.8216 ตามลำดับและเมื่อใช้คัตคำร่วมกับการตัดคำแบบ  
เอ็นแกรมโดยวิธีการสรุปใจความสำคัญแบบเดียวกันได้ค่าความแม่นยำ ค่าความระลึกและค่าวัด  
ประสิทธิภาพอยู่ที่ 0.8119 0.8398 และ 0.8242 ตามลำดับที่อัตราการบีบอัดร้อยละ 20

สาขาวิชา วิศวกรรมซอฟต์แวร์

ลายมือชื่อนิสิต .....

ปีการศึกษา 2563

ลายมือชื่อ อ.ที่ปรึกษาหลัก .....

# # 5970962421 : MAJOR SOFTWARE ENGINEERING

KEYWORD: Extractive Text Summarization

Sarunya Nathonghor : Extractive Text Summarization for Thai Travel News.

Advisor: Duangdao Wichadakul, Ph.D.

The daily newspaper has abundant data that users do not have enough time for reading them. It is difficult to identify relevant information to satisfy the information needed by users. Automatic summarization can reduce the problem of information overloading. Previously, many methodologies have been proposed for English and other languages. However, there were only a few research results in Thai text summarization due to the lack of a corpus in the Thai language and complicated grammar. In this research, we compared two methodologies of document summarization in Thai language. The first method is sentence selection by k-means and the second method is the sentence extraction function based on keyword score. In this research, we also created the 2,340 compound noun corpus from 5,000 news to improve the document summarization and collected 400 Thai travel news for evaluating the summarization. The sentence extraction function based on keyword score summation got the highest accuracy, recall and F1-score at 0.8097 0.8367 and 0.8216 respectively. Applying n-gram with the same summarization method, the accuracy, precision, and recall became 0.8119 0.8398 and 0.8242 with the 20% compression rate.

Field of Study: Software Engineering

Student's Signature .....

Academic Year: 2020

Advisor's Signature .....

## กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้ด้วยความอนุเคราะห์ของอาจารย์ ดร.ดวงดาว วิชาตากุล อาจารย์ที่ปรึกษาวิทยานิพนธ์ ซึ่งท่านได้สละเวลาให้ความรู้ ให้คำปรึกษา ตรวจสอบ แก้ไข ให้คำแนะนำทางการวิจัย และสนับสนุนจนทำให้การวิจัยในครั้งนี้สำเร็จออกมาด้วยดี ข้าพเจ้าจึงขอกราบขอบพระคุณ อาจารย์ ดร.ดวงดาว วิชาตากุล ไว้ ณ ที่นี้

ขอขอบพระคุณกรรมการสอบวิทยานิพนธ์ รองศาสตราจารย์ ดร.ทวีติย์ เสนีวงศ์ ณ อยุธยา ผู้ช่วยศาสตราจารย์ ดร.พีรพล เวทีกุล และรองศาสตราจารย์ ดร.เบญจพร ลิ้มธรรมมาภรณ์ ที่กรุณาสละเวลาให้คำแนะนำ ตรวจสอบและแก้ไขวิทยานิพนธ์ ซึ่งเป็นประโยชน์ต่อการพัฒนางานวิจัยนี้

ขอขอบพระคุณท่านอาจารย์ทุกท่าน ที่ได้ประสิทธิ์ประสาทวิชาให้มีความรู้ จนสามารถนำมาประยุกต์ใช้ให้เกิดประโยชน์ในทุกๆด้าน อีกทั้งรวมถึงบุคลากรทุกท่านในภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ที่คอยให้ข้อมูล คำแนะนำและความช่วยเหลือในระหว่างที่ผู้วิจัยกำลังศึกษาตลอดจนสอบวิทยานิพนธ์สำเร็จลุล่วง

ขอบคุณเพื่อนๆ พี่ๆ น้องๆ ในภาควิชาฯ ที่ได้ให้คำแนะนำช่วยเหลือ และเป็นแรงสนับสนุนแก่ผู้วิจัยเสมอมา

ขอบคุณบิดา มารดา คุณตา คุณยาย และพี่สาวของผู้วิจัย ที่คอยสนับสนุน คอยให้กำลังใจเสมอมาซึ่งเป็นสิ่งที่มีค่าที่สุดแก่ผู้วิจัย

ขอบคุณหัวหน้างาน พี่ๆ เพื่อนๆ น้องๆ บริษัทซีเกทเทคโนโลยี (ประเทศไทย) จำกัด ที่คอยช่วยเหลือและเป็นกำลังใจให้ผู้วิจัยเสมอมา

อนึ่ง ผู้วิจัยหวังว่า วิทยานิพนธ์ฉบับนี้จะมีประโยชน์อยู่ไม่น้อย สำหรับข้อบกพร่องต่างๆที่อาจจะเกิดขึ้น ผู้วิจัยยินดีที่จะรับฟังคำแนะนำจากทุกท่านเพื่อเป็นประโยชน์ในการพัฒนางานวิจัยต่อไป

ศรัณญา นาทองท้อ

## สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	ค
บทคัดย่อภาษาอังกฤษ.....	ง
กิตติกรรมประกาศ.....	จ
สารบัญ.....	ฉ
สารบัญตาราง.....	ฌ
สารบัญรูปภาพ.....	ญ
บทที่ 1 บทนำ.....	1
1.1. ที่มาและความสำคัญของปัญหา.....	1
1.2. วัตถุประสงค์.....	2
1.3. ขอบเขตงานวิจัย.....	2
1.4. ขั้นตอนและวิธีดำเนินงานวิจัย.....	3
1.5. ประโยชน์ที่คาดว่าจะได้รับ.....	3
1.6. บทความวิชาการที่ได้รับการตีพิมพ์.....	3
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง.....	4
2.1 ทฤษฎีที่เกี่ยวข้อง.....	4
2.1.1 การสรุปใจความสำคัญ (Text Summarization).....	4
2.1.2 แนวความคิดด้านการสรุปความ.....	4
2.1.3 การตัดคำ (Word Segmentation).....	5
2.1.4 การกรองคำหยุด (Stop Word Elimination).....	6
2.1.5 การแบ่งประโยค (Sentence Segmentation).....	6
2.1.6 ความถี่ของคำ (Term Frequency).....	6

2.1.7 คำประสม (Compound Noun) .....	6
2.1.8 เอ็นแกรม (N-Gram).....	7
2.1.9 การวิเคราะห์ความหมายแฝงของคำ (Latent Semantic Analysis).....	7
2.1.10 ความคล้ายคลึงแบบโคไซน์ (Cosine Similarity).....	9
2.1.11 การจัดกลุ่มแบบเคมีน (K-means Clustering).....	10
2.1.12 คัดคำ (Cutkum).....	11
2.1.13 การวัดประสิทธิภาพของการสรุปใจความสำคัญ.....	13
2.2 งานวิจัยที่เกี่ยวข้อง .....	14
2.2.1 งานวิจัยที่เกี่ยวกับการหาคำสำคัญและการตัดคำ .....	14
2.2.2 งานวิจัยที่เกี่ยวกับการย่อเอกสารและสรุปใจความสำคัญ.....	15
2.2.3 งานวิจัยการประมวลผลภาษาธรรมชาติอื่นๆที่นำมาประยุกต์ใช้ .....	18
บทที่ 3 แนวคิดและวิธีการ .....	21
3.1 ภาพรวมของวิธีการ.....	21
3.2 การสร้างรายการคำประสม.....	23
3.3 ขั้นตอนการเตรียมข้อมูลเพื่อสรุปใจความสำคัญ.....	29
3.4 การหาน้ำหนักของคำและการเลือกประโยค .....	32
บทที่ 4 การทดลองและผลการทดลอง .....	41
4.1 สภาพแวดล้อมและเครื่องมือที่ใช้ในการทดลอง .....	41
4.1.1 คอมพิวเตอร์ที่ใช้ในการทดลอง .....	41
4.1.2 การเขียนโปรแกรม .....	41
4.1.3 ข้อมูลที่ใช้ในการทดลอง .....	43
4.2 การทดลองตัดคำประสม.....	43
4.3 การทดลองเปรียบเทียบการตัดคำด้วยคัตคำร่วมกับคำประสมและการตัดคำด้วยคัตคำร่วมกับ เอ็นแกรม.....	52



4.4 การทดลองเพื่อหาอัตราการย่อความ .....	53
4.5 การทดลองเปรียบเทียบผลของแนวคิดที่นำเสนอ.....	54
4.6 การทดลองเปรียบเทียบผลกับงานวิจัยที่เกี่ยวข้อง.....	55
บทที่ 5 สรุปผลการวิจัย.....	59
5.1 สรุปผลการวิจัย.....	59
5.2 แนวทางวิจัยในอนาคต.....	60
ภาคผนวก ก รายการคำประสมที่ได้จากข่าวท่องเที่ยวไทย.....	62
ภาคผนวก ข ชุดข้อมูลข่าวที่ใช้ในการทดลอง.....	83
ภาคผนวก ค ตัวอย่างข่าวที่ได้จากการสรุปใจความสำคัญของข่าวท่องเที่ยวไทย .....	84
บรรณานุกรม.....	98
ประวัติผู้เขียน.....	100



## สารบัญตาราง

หน้า

ตารางที่ 2.1 เมทริกซ์ความสัมพันธ์ของคำกับประโยคที่พบในเอกสาร .....	8
ตารางที่ 2.2 เปรียบเทียบประสิทธิภาพของเครื่องมือตัดคำ .....	11
ตารางที่ 2.3 เปรียบเทียบประสิทธิภาพของเครื่องมือตัดคำ (ต่อ).....	12
ตารางที่ 2.4 เปรียบเทียบประสิทธิภาพของเครื่องมือตัดคำ (ต่อ) .....	12
ตารางที่ 2.5 ตัวอย่างการเลือกประโยคจาก Matrix $V^T$ ของ Gong และคณะ .....	15
ตารางที่ 2.6 ตัวอย่าง Matrix $\Sigma$ จากงานวิจัยของ Murray และคณะ .....	16
ตารางที่ 2.7 ตัวอย่างการเลือกประโยคจากงานวิจัยของ Murray และคณะ .....	17
ตารางที่ 2.8 เปรียบเทียบงานวิจัยที่เกี่ยวข้องด้านการสรุปใจความสำคัญของเอกสารภาษาไทย .....	20
ตารางที่ 3.1 ตัวอย่างจำนวนครั้งของคำที่เกิดขึ้นในแต่ละประโยค .....	35
ตารางที่ 3.2 ตัวอย่าง Term - Sentence Matrix จากการคำนวณความถี่ของคำ .....	36
ตารางที่ 3.3 ตัวอย่าง Matrix $V$ .....	37
ตารางที่ 3.4 ตัวอย่าง Term-Sentence Matrix จากความถี่ของคำจากหัวข้อข่าวกับประโยค.....	40
ตารางที่ 4.1 รายละเอียดชุดข้อมูลข่าวที่ใช้ในการทดลอง .....	43
ตารางที่ 4.2 ตารางแสดงผลการสรุปใจความสำคัญโดยแบ่งชุดข้อมูลตามจำนวนคำประสม .....	51
ตารางที่ 4.3 ผลการเปรียบเทียบการสรุปใจความสำคัญโดยการตัดคำประสมและการตัดคำด้วยเอ็น แกรม .....	52
ตารางที่ 4.4 ผลการทดลองใช้อัตราส่วนที่แตกต่างกันในการสรุปใจความสำคัญ.....	54
ตารางที่ 4.5 ผลการเปรียบเทียบแนวคิดที่นำเสนอโดยใช้ 400 ข่าวการท่องเที่ยวไทย .....	54
ตารางที่ 4.6 เปรียบเทียบผลการทดลองกับงานวิจัยที่เกี่ยวข้อง (อัตราการบีบอัดร้อยละ 20).....	55
ตารางที่ 4.7 เปรียบเทียบผลการทดลองกับงานวิจัยที่เกี่ยวข้อง (อัตราการบีบอัดร้อยละ 30).....	56
ตารางที่ 4.8 เปรียบเทียบผลการทดลองกับงานวิจัยที่เกี่ยวข้อง (อัตราการบีบอัดร้อยละ 40).....	57

## สารบัญรูปภาพ

หน้า

รูปที่ 2.1 ลักษณะการแยกตัวประกอบของเมตริกซ์ด้วย SVD .....	8
รูปที่ 2.2 ผลลัพธ์ของการแยกตัวประกอบของเมตริกซ์ด้วย SVD .....	9
รูปที่ 3.1 ภาพรวมของแนวคิดที่นำเสนอ .....	22
รูปที่ 3.2 ขั้นตอนการสร้างรายการคำประสม .....	24
รูปที่ 3.3 ตัวอย่างข้อความเนื้อหาข่าวท่องเที่ยวจากไฟล์ .html.....	25
รูปที่ 3.4 ตัวอย่างข้อความเนื้อหาข่าวท่องเที่ยวจากไฟล์ .txt.....	25
รูปที่ 3.5 ตัวอย่างชุดของคำหลังผ่านการตัดคำ ลบคำหยุด และคำซ้ำ .....	26
รูปที่ 3.6 ตัวอย่างการค้นหาชนิดของคำในพจนานุกรมออนไลน์จากเว็บไซต์ Longdo .....	27
รูปที่ 3.7 ตัวอย่างคำประสมที่ได้จากการค้นหาในพจนานุกรมออนไลน์จากเว็บไซต์ Longdo.....	28
รูปที่ 3.8 จำนวนรายการคำประสมแยกตามพยัญชนะไทย .....	28
รูปที่ 3.9 จำนวนประโยคจากข่าว 5 ชุด .....	29
รูปที่ 3.10 ตัวอย่างรายการคำประสมหมวดหมู่ ๗ .....	30
รูปที่ 3.11 ตัวอย่างชุดคำที่ผ่านโปรแกรมตัดคำและการตัดคำโดยใช้คำประสม.....	30
รูปที่ 3.12 จำนวนคำประสมในข้อมูลชุดทดลอง.....	31
รูปที่ 3.13 ตัวอย่างการแบ่งประโยค.....	32
รูปที่ 4.1 การใช้ไอโฟนเซอร์สในขั้นตอนต่างๆ .....	42
รูปที่ 4.2 ข่าวการท่องเที่ยวไทยฉบับเต็ม .....	48
รูปที่ 4.3 การตัดคำระหว่างการตัดคำด้วยคัตคำโดยไม่ตัดคำประสม และการตัดคำด้วยคัตคำร่วมกับ การตัดคำประสมจากตัวอย่างข่าว .....	50

# บทที่ 1

## บทนำ

### 1.1. ที่มาและความสำคัญของปัญหา

ปัจจุบันเทคโนโลยีทางด้านคอมพิวเตอร์มีความสำคัญต่อการดำเนินชีวิตประจำวันของมนุษย์เป็นอย่างมากและยังถือว่าเป็นเครื่องมือที่ใช้ในการอำนวยความสะดวกให้แก่มนุษย์มากมาย โดยเฉพาะทางด้านการสื่อสารผ่านสังคมออนไลน์ เนื่องจากสะดวกสบาย และตอบสนองความต้องการของมนุษย์ได้ดีกว่าการเสียค่าใช้จ่ายในการติดต่อผ่านทางโทรศัพท์ในอดีตที่ผ่านมา และนอกจากนั้นเทคโนโลยีการนำเสนอข้อมูลผ่านโลกอินเทอร์เน็ตก็ได้เข้ามามีส่วนร่วมในชีวิตประจำวันของมนุษย์มากขึ้นเช่นกัน เพราะจากสถานการณ์ปัจจุบัน มนุษย์ใช้ชีวิตอย่างเร่งรีบ และส่วนใหญ่สูญเสียเวลาไปกับการอยู่บนยานพาหนะมากกว่าการใช้ชีวิตทำอย่างอื่น ดังนั้นจึงเสาะหาช่องทางออนไลน์เป็นหลัก เพื่อลดเวลาในการอ่านข่าวหรืออ่านบทความต่างๆ จากหัวข้อวิจัยที่ผ่านมาจะเห็นได้ว่าการสรุปใจความสำคัญนั้นมีผู้ให้ความสนใจเป็นอย่างมาก แต่เนื่องจากภาษาไทยเป็นภาษาที่ยากและมีไวยากรณ์ที่ไม่ตายตัวเหมือนภาษาอื่นๆ จึงทำให้ยังไม่มีหัวข้องานวิจัยที่ชัดเจนเกี่ยวกับเทคนิคการสรุปใจความสำคัญของภาษาไทย ซึ่งส่วนใหญ่หัวข้องานวิจัยที่ผ่านมาจะเน้นไปที่เทคนิคการสรุปใจความสำคัญด้วยวิธีการดึงคำจากต้นฉบับ (Extraction) มากกว่าการวิธีการสรุปใจความแบบภาพรวม (Abstraction) โดยการที่สรุปใจความสำคัญด้วยวิธีการดึงคำจากต้นฉบับต้องอาศัยเครื่องมือตัดคำที่มีคุณภาพเพื่อส่งผลให้การดึงสาระสำคัญของบทความออกมาทำสรุปนั้นมีความหมายและเข้าใจได้ง่ายขึ้น

จากการวิจัยที่ผ่านมาได้มีการศึกษาและพัฒนาการสรุปใจความสำคัญของภาษาไทยมากมาย โดยที่การสรุปใจความสำคัญนั้นจะถูกต้องหรือไม่ขึ้นอยู่กับวิธีการตัดคำเป็นหลัก เพราะหน่วยคำเป็นหน่วยที่เล็กที่สุดของประโยค หากตัดคำผิดไปอาจจะทำให้การสรุปใจความสำคัญมีความหมายที่ผิดไปด้วย นอกจากนี้งานวิจัยส่วนใหญ่ไม่ได้เน้นการตัดคำประสมมากนักเนื่องจากถูกจำกัดด้วยข้อมูลทางคลังภาษา ทำให้การพัฒนาด้านการตัดคำและสรุปใจความสำคัญของภาษาไทยยังไม่ครอบคลุมกฎไวยากรณ์ทั้งหมด และไม่สามารถบอกได้ว่า ณ ปัจจุบันนี้มีงานวิจัยที่สามารถสรุปใจความสำคัญของภาษาไทยได้ถูกต้องและมีประสิทธิภาพเพียงพอ

จากปัญหาข้างต้น ผู้วิจัยจึงเล็งเห็นที่จะพัฒนาประสิทธิภาพการสรุปใจความสำคัญของภาษาไทยด้วยวิธีการดึงใจความสำคัญของประโยคจากต้นฉบับ (Extraction) โดยในงานวิจัยนี้ นำเสนอ 2 วิธี คือ 1. การสรุปใจความสำคัญโดยคำนวณค่าน้ำหนักของคำสำคัญโดยใช้ความถี่ของคำ

และเลือกประโยคจากลำดับประโยคจากผลรวมน้ำหนักของคำ 2. การสรุปใจความสำคัญของภาษาไทยโดยอาศัยหลักการคำนวณหาค่าน้ำหนักของคำสำคัญจากประโยคด้วยวิธีการแยกตัวประกอบเมตริกซ์และการจัดกลุ่มประโยคด้วยการจัดกลุ่มแบบเคมีน เพื่อเพิ่มประสิทธิภาพของโมเดล ผู้วิจัยได้เพิ่มการตัดคำประสมร่วมด้วยโดยสร้างคลังข้อมูลรายการคำประสมที่สกัดจากข่าวท่องเที่ยวภาษาไทยขึ้นมา ในงานวิจัยนี้ผู้วิจัยเลือกกำหนดจำนวนประโยคผลลัพธ์ที่ได้จากการสรุปใจความสำคัญอยู่ที่ระดับร้อยละ 80 70 และ 60 ของจำนวนประโยคทั้งหมดใน 1 ข่าว และเปรียบเทียบการสรุปใจความสำคัญของข่าวจากอัลกอริทึมกับกับการสรุปใจความสำคัญกับมนุษย์ว่าแบบไหนได้ผลที่ดีกว่ากัน

## 1.2. วัตถุประสงค์

- 1) เพื่อสร้างคลังข้อมูลรายการคำประสมที่สกัดจากข่าวท่องเที่ยวภาษาไทย
- 2) เพื่อเพิ่มประสิทธิภาพวิธีการสรุปใจความสำคัญแบบสกัดสำหรับข่าวท่องเที่ยวภาษาไทย

## 1.3. ขอบเขตงานวิจัย

วิทยานิพนธ์ฉบับนี้แบ่งขอบเขตงานวิจัยออกเป็น 2 ส่วนคือ 1. การสร้างคลังข้อมูลรายการคำประสมที่สกัดจากข่าวท่องเที่ยวภาษาไทย 2. การสรุปใจความสำคัญของข่าวการท่องเที่ยวภาษาไทย

- 1) ขอบเขตการสร้างคลังข้อมูลรายการคำประสมที่สกัดจากข่าวท่องเที่ยวภาษาไทย โดยเลือกใช้ข่าวการท่องเที่ยวและบทความการท่องเที่ยวจาก <https://mgronline.com/travel> ในช่วงระหว่างปีพ.ศ. 2558 ถึงปี พ.ศ. 2563 จำนวน 5,000 ข่าว
- 2) ขอบเขตการสรุปใจความสำคัญ
  - เป็นการสรุปใจความสำคัญของข่าวท่องเที่ยวไทยจากการรวบรวมข่าวการท่องเที่ยวและบทความการท่องเที่ยวจาก <https://www.thairath.co.th/tags/การท่องเที่ยว> จำนวน 400 ข่าว แบ่งออกเป็น 5 ชุด ชุดละ 80 ข่าว
  - ใช้เครื่องมือตัดคำในการตัดคำจากข่าว
  - การสรุปใจความสำคัญเฉพาะภาษาไทยเท่านั้น
  - การสรุปใจความสำคัญของมนุษย์ใช้การเลือกหมายเลขประโยคที่กำกับไว้ของแต่ละข่าวเท่านั้น ไม่ใช่การสร้างประโยคใหม่
  - ใช้มาตรฐานการวัดประสิทธิภาพ F1 Score โดยเป็นการเปรียบเทียบการเลือกประโยคที่ใช้ในการสรุปจากแนวคิดที่นำเสนอและผู้เชี่ยวชาญเท่านั้น

#### 1.4. ขั้นตอนและวิธีดำเนินงานวิจัย

- 1) ศึกษาความรู้และทฤษฎีที่เกี่ยวข้องกับการสรุปใจความสำคัญภาษาไทย
  - ศึกษาหลักการสรุปใจความสำคัญภาษาไทย
  - ศึกษางานวิจัยที่เกี่ยวข้อง
  - ศึกษาการใช้โปรแกรมตัดคำในการตัดคำภาษาไทย
  - เก็บรวบรวมบทความข่าวด้านท่องเที่ยวจากเว็บไซต์
  - ศึกษาขั้นตอนการเลือกคำสำคัญเพื่อหาความสัมพันธ์ของชื่อบทความและเนื้อหาบทความหรือข่าว
  - ศึกษาหลักการการจัดกลุ่มประโยคเพื่อหาวิธีที่มีประสิทธิภาพในการจัดกลุ่มประโยคที่ใกล้เคียงกันมากที่สุด
- 2) ออกแบบแนวคิดและการวัดประสิทธิภาพของโมเดล
- 3) พัฒนาระบบตามแนวคิด
- 4) ทดสอบและประเมินผล
- 5) จัดทำบทความวิชาการ
- 6) ออกแบบแนวคิดเสริมและปรับปรุงระบบให้ดีขึ้น
- 7) สรุปผลการทดลองและจัดทำรูปเล่มวิทยานิพนธ์
- 8) สอบและเผยแพร่วิทยานิพนธ์

#### 1.5. ประโยชน์ที่คาดว่าจะได้รับ

- 1) ได้คลังข้อมูลรายการคำประสมที่สกัดจากข่าวท่องเที่ยวภาษาไทย
- 2) มีวิธีการที่สามารถสรุปใจความสำคัญของข่าวการท่องเที่ยวไทยที่แม่นยำมากขึ้น ทำให้สามารถลดเวลาในการอ่านข่าว และเข้าใจเนื้อหาได้อย่างถูกต้อง

#### 1.6. บทความวิชาการที่ได้รับการตีพิมพ์

ส่วนหนึ่งของวิทยานิพนธ์ได้รับการตีพิมพ์และนำเสนอในงานประชุมวิชาการระดับนานาชาติ Sarunya Nathonghor and Duangdao Wichadukul. “Extractive Text Summarization for Thai Travel News Based on Keyword Scored in Thai Language” ในการประชุมวิชาการ “2020 2nd International Conference on Information Technology and Computer Communications” ระหว่างวันที่ 12-14 สิงหาคม 2563 ณ กรุงกัวลาลัมเปอร์ ประเทศมาเลเซีย (ออนไลน์)

## บทที่ 2

### ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

#### 2.1 ทฤษฎีที่เกี่ยวข้อง

##### 2.1.1 การสรุปใจความสำคัญ (Text Summarization)

การสรุปใจความสำคัญ (Hahn, 2000) เป็นเทคนิคการสรุปเนื้อหาของเอกสารโดยประกอบด้วย 3 ขั้นตอน คือ ขั้นตอนการวิเคราะห์คำ ขั้นตอนการหาประเด็นสำคัญ และขั้นตอนการสังเคราะห์ประโยค โดยสามารถแบ่งวิธีการสรุปใจความสำคัญตามวิธีการสร้างได้ 2 วิธีคือ วิธีการดึงคำจากต้นฉบับ (Extraction) ซึ่งเป็นการสร้างใจความสำคัญโดยใช้เทคนิคทางด้านสถิติ และการค้นหาคำสำคัญ (Keyword-Based) มาใช้ในการตัดคำจากเอกสารต้นฉบับ และอีกวิธีคือวิธีการสรุปแบบภาพรวม (Abstraction) ซึ่งเป็นการวิเคราะห์คำจากต้นฉบับด้วยหลักการทางภาษาศาสตร์ (Semantic) โดยเป็นการถอดความหรือแปลความหมายจากต้นฉบับ ผลที่ได้มีประสิทธิภาพและถูกต้องมากกว่าวิธีการดึงคำจากต้นฉบับ แต่เนื่องจากวิธีการสรุปแบบภาพรวมยังต้องอาศัยคลังข้อมูลขนาดใหญ่และในปัจจุบันคลังข้อมูลของภาษาไทยยังมีไม่เพียงพอ ดังนั้นผู้วิจัยจึงเลือกพัฒนาต่อยอดการสรุปใจความสำคัญโดยวิธีดึงคำจากต้นฉบับและเพิ่มประสิทธิภาพโดยตัดคำประสมรวมด้วยซึ่งพิจารณาจากบริบทรอบข้างคำนั้นๆ

##### 2.1.2 แนวความคิดด้านการสรุปความ

แนวความคิดด้านการสรุปความ ประกอบไปด้วย วัตถุประสงค์ (Intent) จุดความสนใจ (Focus) และขอบเขต (Coverage)

- 1) วัตถุประสงค์ (Intent) คือการนำการสรุปความไปประยุกต์ใช้ในการชี้แนะ การให้ข้อมูล และการประเมินค่า
  - การสรุปความเพื่อชี้แนะจะเป็นการสรุปคร่าวๆเพื่อให้ผู้อ่านเข้าใจประเด็นหลักของเอกสารฉบับเต็มหรือการให้คำชี้แนะสั้นๆเกี่ยวกับหัวเรื่องของเอกสาร
  - การสรุปใจความสำคัญแบบให้ข้อมูลเป็นเสมือนตัวแทนของเอกสารฉบับเต็มแต่จะเก็บเฉพาะรายละเอียดที่สำคัญไว้เท่านั้น
  - การสรุปแบบประเมินค่า เป็นการเก็บเนื้อหาที่ตรงกับหัวข้อที่ผู้แต่งกำหนด
- 2) จุดความสนใจ (Focus) คือ ขอบเขตของข้อมูลที่สนใจ
- 3) ขอบเขต (Coverage) คือ การสรุปความที่บ่งบอกว่ามีเอกสารฉบับเดียว หรือมีเอกสารหลายฉบับ

แนวคิดด้านการสรุปความในงานวิจัยนี้ คือ การสรุปใจความสำคัญโดยให้ข้อมูลเป็นเสมือนตัวแทนของเอกสารฉบับเต็มแต่จะเก็บเฉพาะรายละเอียดที่สำคัญ และเก็บเนื้อหาที่ตรงกับหัวข้อที่ผู้แต่งกำหนดโดยมีขอบเขตที่สนใจเกี่ยวกับข่าวการฟ้องเกี่ยวจากข่าวฉบับเดียว

### 2.1.3 การตัดคำ (Word Segmentation)

การตัดคำ (วีรัช ศรีเลิศล้ำวาณิช, 2536) คือการตัดแบ่งข้อความที่มีความยาวต่อเนื่องกันออกเป็นหน่วยที่เล็กที่สุดคือหน่วยของคำ (Morpheme) โดยมีจุดประสงค์คือต้องการแบ่งสายอักขระของตัวอักษรเพื่อหาขอบเขตของแต่ละหน่วยคำ ซึ่งวิธีการในการตัดคำขึ้นอยู่กับลักษณะของภาษานั้นๆ

การเขียนภาษาไทยจะเป็นลักษณะการเขียนที่ต่อเนื่องกันไปไม่สามารถหาขอบเขตของคำได้ ซึ่งแตกต่างจากภาษาอังกฤษที่มีช่องว่างในการแบ่งคำแต่ละคำอย่างชัดเจน ดังนั้นการตัดคำภาษาไทยจึงมีจุดมุ่งหมายในการหาขอบเขตของคำให้ถูกต้องที่สุด โดยวิธีการตัดคำภาษาไทยสามารถแบ่งได้เป็น 3 ยุคคือ ยุคการใช้กฎ (Rule-base) ยุคการใช้พจนานุกรม (Dictionary) และยุคการใช้คลังข้อมูล (Corpus) ดังนี้

#### 1) ยุคการใช้กฎ (Rule-base)

เป็นยุคที่มีการนำกฎเข้ามาใช้ในการตัดพยางค์ เนื่องจากพยางค์มีกฎเกณฑ์ที่แน่นอนมากกว่าคำ โดยกฎที่นำมาใช้นั้นได้นำมาจากหลักไวยากรณ์ภาษาไทย เช่น กฎการหาขอบเขต และการหาขอบเขตหลัง (Surin Charnyapornpong, 1983) ซึ่งจากการนำกฎเข้ามาใช้ทำให้สามารถแบ่งพยางค์ได้ถูกต้องเป็นจำนวนมาก ตัวอย่างการใช้กฎ เช่น การขึ้นย่อหน้าเป็นตัวบ่งชี้ถึงการสิ้นสุดข้อความ การเว้นวรรคเป็นตัวบ่งชี้ถึงการสิ้นสุดคำหรือประโยค เป็นต้น

#### 2) ยุคการใช้พจนานุกรม (Dictionary)

เป็นยุคที่สองที่พัฒนามาจากยุคของการใช้กฎในการตัดคำ โดยยุคนี้เป็นยุคที่คิดค้นวิธีการแบ่งคำ ซึ่งผลลัพธ์จากยุคแรกนั้นแสดงให้เห็นว่าการแบ่งคำจากการใช้กฎเพียงอย่างเดียวไม่สามารถหาขอบเขตของคำได้ ดังนั้นจึงต้องอาศัยการนำพจนานุกรมมาใช้ร่วมกับกฎในการตัดคำด้วย เช่น เก็บคำภาษาไทยไว้ในพจนานุกรม แล้วป้อนข้อความเข้าไปค้นหาและเทียบสายอักขระกับคำในพจนานุกรม เพื่อหาว่าควรตัดคำในข้อความดังกล่าวว่าประกอบด้วยคำใดบ้าง ตัวอย่างการเลือกใช้พจนานุกรม เช่น การเลือกตัดคำโดยเลือกตัดคำแบบที่ยาวที่สุด และการเลือกตัดคำโดยเลือกคำที่เหมือนที่สุดในพจนานุกรม เป็นต้น

#### 3) ยุคการใช้คลังข้อมูล (Corpus)

ยุคที่สามเป็นยุคของการใช้คลังข้อมูลซึ่งยุคนี้ได้มีการพัฒนาคำขึ้นใหม่ นอกจากการใช้กฎและการใช้พจนานุกรมแล้ว ในยุคนี้ยังมีการนำเอาความรู้จากคลังข้อมูลคำศัพท์เข้ามาประยุกต์ใช้อีกด้วย



ตัวอย่างเช่น การนำค่าสถิติของการใช้คำภายในคลังข้อมูลคำศัพท์และลักษณะไวยากรณ์ที่ใช้ในคลังข้อมูลคำศัพท์มาคำนวณหาการตัดคำเพื่อให้มีความสมบูรณ์มากขึ้น เป็นต้น

#### 2.1.4 การกรองคำหยุด (Stop Word Elimination)

คำหยุดคือ กลุ่มคำที่ไม่มีความหมายในภาษาไทย แต่พบในเอกสาร เช่นคำว่า “โดย” “ซึ่ง” “นั้น” “แล้ว” “เป็นอย่างไร” “กระนั้น” เป็นต้น ดังนั้นในงานวิจัยนี้จึงต้องเลือกลบคำหยุดออกด้วย เพื่อช่วยลดคำที่ไม่มีความเกี่ยวข้องกับประโยคออกจากการหาคำสำคัญของประโยคเพื่อนำมาใช้สรุปใจความสำคัญภาษาไทย

#### 2.1.5 การแบ่งประโยค (Sentence Segmentation)

ประโยค (วีระชาติ ศิริไกรวัฒนาวงศ์, 2547) คือ กลุ่มคำที่นำมาเรียงต่อกันตามระเบียบของภาษามีใจความสมบูรณ์ โดยหนึ่งประโยคจะประกอบไปด้วย ภาคประธานและภาคแสดงเป็นหลัก โดยภาคประธานของประโยคคือ ผู้กระทำ ส่วนภาคแสดงคือ วิธีการกระทำที่แสดงให้เห็นว่าภาคประธานทำอะไร อยู่ที่ไหน หรือมีวิธีการทำอย่างไร

#### 2.1.6 ความถี่ของคำ (Term Frequency)

ความถี่ของคำ (Term Frequency) เป็นวิธีที่ใช้ในการคำนวณหาค่าความสำคัญของคำที่ปรากฏขึ้นในประโยค โดยคำนวณจากจำนวนครั้งของคำนั้นที่ปรากฏขึ้นในประโยคหารด้วยจำนวนคำทั้งหมดในประโยคนั้นๆ

งานวิจัยนี้ได้นำวิธีการหาความถี่ของคำมาใช้ในการคำนวณหาค่าความสำคัญของคำที่ปรากฏขึ้นในประโยค โดยในงานวิจัยนี้ได้แบ่งน้ำหนักของคำ ดังสมการที่ (2.1)

$$tf(\text{term, sentence}) = \frac{\text{จำนวนคำนั้นที่เกิดขึ้นในประโยค}}{\text{จำนวนของคำทั้งหมดในประโยค}} \quad (2.1)$$

โดยที่

tf	แทน ความถี่ของคำคำนั้นที่ปรากฏในประโยค
term	แทน คำ
sentence	แทน ประโยค

#### 2.1.7 คำประสม (Compound Noun)

คำประสม คือคำที่เกิดจากคำ 2 คำขึ้นไปมารวมกันทำให้เกิดคำที่มีความหมายใหม่ โดยคำประสมในภาษาไทยเกิดจากชนิดของคำนำมาประสมกัน เช่น คำนามประสมคำนาม ได้แก่ พ่อตา แม่ยาย ลูกน้อง หน้าม้า เป็นต้น คำนามประสมคำกริยา ได้แก่ กลัวยับ หมอดู บ้านพัก ไข่ทอด เป็นต้น คำนามประสมคำลักษณะนาม ได้แก่ เพื่อนฝูง วงกบ ดวงตา เป็นต้น เพื่อเพิ่มประสิทธิภาพของการ

สรุปใจความสำคัญของภาษาไทย ผู้วิจัยจึงสนใจที่จะตัดคำประสมร่วมด้วยเพื่อเพิ่มประสิทธิภาพในการตัดคำให้แม่นยำมากยิ่งขึ้น

### 2.1.8 เอ็นแกรม (N-Gram)

เอ็นแกรม คือ ลำดับที่ต่อเนื่องกันของสายอักขระของคำ หรือเสียงที่เกิดร่วมกันเป็นประโยค ซึ่งคำว่า “แกรม” หมายถึง หน่วยที่ใช้ในการสร้างแบบจำลอง และแกรมมีได้หลายขนาด ตั้งแต่ 1 แกรม จนถึง n แกรม ตัวอย่างเช่น

ประโยค	จังหวัดกาญจนบุรีเป็นจังหวัดหนึ่งในภาคกลาง
ตัดคำ	จังหวัด กาญจนบุรี เป็น จังหวัด หนึ่งใน ภาคกลาง
2 แกรม	จังหวัดกาญจนบุรี กาญจนบุรีเป็น เป็นจังหวัด จังหวัดหนึ่ง หนึ่งในภาคกลาง
3 แกรม	จังหวัดกาญจนบุรีเป็น กาญจนบุรีเป็นจังหวัด เป็นจังหวัดหนึ่ง จังหวัดหนึ่งใน หนึ่งในภาคกลาง

งานวิจัยนี้ส่วนหนึ่งนั้นนำเสนอการสร้างคลังข้อมูลของรายการคำประสมที่สกัดจากข่าวท่องเที่ยวภาษาไทยโดยใช้หลักการนำคำที่ติดกัน 2 ตำแหน่งไปค้นหาในพจนานุกรม เพื่อเพิ่มประสิทธิภาพของการสรุปใจความสำคัญแบบสกัดของข่าวท่องเที่ยวไทย ผู้วิจัยจึงสนใจที่จะเปรียบเทียบการตัดคำโดยใช้เอ็นแกรมโดยไม่สนใจความหมายในเชิงพจนานุกรมเพื่อเปรียบเทียบกับผลจากการตัดคำประสมว่าวิธีการใดมีประสิทธิภาพมากกว่า โดยงานวิจัยนี้เลือกใช้การตัดคำขนาด 2 แกรมและ 3 แกรม

### 2.1.9 การวิเคราะห์ความหมายแฝงของคำ (Latent Semantic Analysis)

การวิเคราะห์ความหมายแฝงของคำ เป็นวิธีทางสถิติที่ใช้วิเคราะห์หาค่าความสัมพันธ์ของคำที่ปรากฏร่วมกัน เพื่อดึงความหมายของคำจากเอกสารหรือข้อความที่มีขนาดใหญ่ Deerwester และคณะ (Deerwester, Dumais, Furnas, Landauer, & Harshman, 1990) เป็นผู้พัฒนา หลักการวิเคราะห์วิธีนี้จะทำให้เห็นถึงความสัมพันธ์ที่ซ่อนอยู่ระหว่างคำกับคำ คำกับเอกสาร คำกับประโยคในรูปแบบของเมตริกซ์ที่มีขนาดใหญ่ได้ คอนเซปต์ที่ซ่อนอยู่ในเนื้อหาสามารถจำแนกโดยลักษณะการใช้คำ ที่มีความใกล้เคียงกันไว้ด้วยกัน โดยการวิเคราะห์ความหมายแฝงของคำนี้มีการนำอัลกอริทึมการแยกค่าแบบเดี่ยวมาประยุกต์ใช้และมีขั้นตอนหลักดังนี้

1) อัลกอริทึมการแยกค่าแบบเดี่ยว (Singular Value Decomposition, SVD)

อัลกอริทึมการแยกค่าแบบเดี่ยว (De Lathauwer, De Moor, Vandewalle, & Applications, 2000) ขั้นตอนแรกเริ่มจากการสร้างเมตริกซ์  $A$  ที่มีขนาด  $m \times n$  ดังตารางที่ 2.1 โดยที่  $m$  คือแถว (Row) แทนจำนวนคำ (Terms) และ  $n$  คือสดมภ์ (Column) แทนประโยคในเอกสารที่นำมาวิเคราะห์ ซึ่งคำในแต่ละแถวนั้นต้องไม่ซ้ำกัน แต่ละช่องแสดงความถี่ของคำที่ปรากฏในแต่ละประโยค

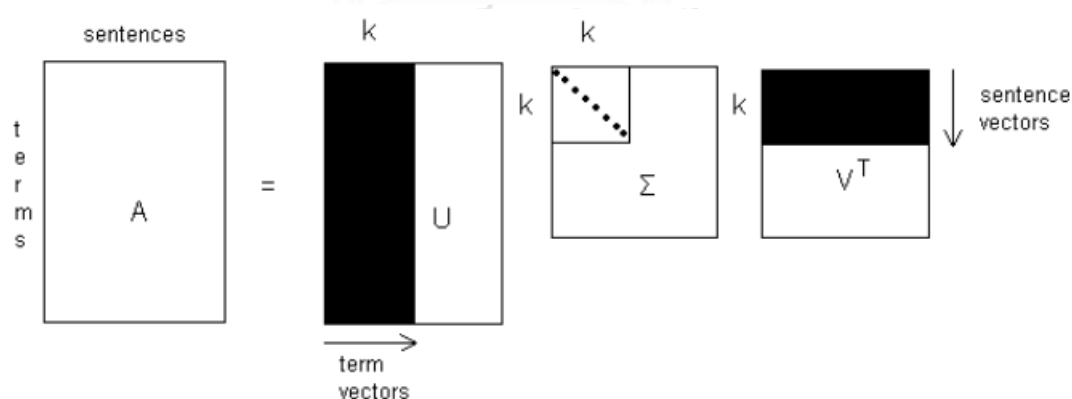
ตารางที่ 2.1 เมตริกซ์ความสัมพันธ์ของคำกับประโยคที่พบในเอกสาร

	ประโยค 1	ประโยค 2	ประโยค 3	...	ประโยค n
คำ 1	d(1,1)	d(2,1)	d(3,1)		d(1,n)
คำ 2					...
คำ 3					...
...					...
คำ m	d(m, 1)	...	...	...	d(m,n)

จากนั้นนำเมตริกซ์  $A$  มาผ่านอัลกอริทึมการแยกค่าแบบเดี่ยวในสมการ (2.2)

$$A \approx U\Sigma V^T \quad (2.2)$$

โดยที่เมตริกซ์  $V$  คือเมตริกซ์มุมฉาก (Orthogonal) เมตริกซ์  $\Sigma$  คือผลคูณในแนวทแยงและเมตริกซ์  $U$  คือค่าประมาณการความหมายแฝงของคำศัพท์แสดงดังรูปที่ 2.1



รูปที่ 2.1 ลักษณะการแยกตัวประกอบของเมตริกซ์ด้วย SVD (Deerwester et al., 1990)

เมื่อนำอัลกอริทึมการแยกค่าแบบเดี่ยวมาประยุกต์กับการวิเคราะห์ความหมายแฝงของคำจะได้อธิบายดังรูปที่ 2.2

$$\begin{aligned}
 & \text{Matrix } A \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 10 & 11 & 12 \end{bmatrix} \\
 & \approx \begin{bmatrix} -0.1409 & -0.8247 & 0.4872 \\ -0.3439 & -0.4263 & -0.4630 \\ -0.5470 & -0.0278 & -0.5355 \\ -0.7501 & 0.3706 & 0.5113 \end{bmatrix} \times \begin{bmatrix} 25.4624 & 0 & 0 \\ 0 & 1.2907 & 0 \\ 0 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} -0.5045 & -0.5746 & -0.6445 \\ 0.7608 & 0.0571 & -0.6465 \\ 0.4082 & -0.8165 & 0.4082 \end{bmatrix} \\
 & \qquad \qquad \qquad U \qquad \qquad \qquad \Sigma \qquad \qquad \qquad V^T
 \end{aligned}$$

รูปที่ 2.2 ผลลัพธ์ของการแยกตัวประกอบของเมตริกซ์ด้วย SVD (Lee, Park, Ahn, Kim, & Management, 2009)

ในขั้นตอนนี้เมื่อแยกตัวประกอบเมตริกซ์ด้วยอัลกอริทึมการแยกค่าแบบเดี่ยว สามารถแยกได้ทั้งหมด 3 เมตริกซ์ ดังนี้

1. Matrix U (Term - Concept) เมตริกซ์เชิงตั้งฉากแทนความสัมพันธ์ของคำและคอนเซปต์ ใช้เพื่อแสดงให้เห็นว่าแต่ละคำอยู่ในคอนเซปต์ใด โดยดูจากคอมลันน์ใดมีค่าสูงที่สุด โดยค่ามากหมายถึงความเป็นสมาชิกในกลุ่มนี้มาก
2. Matrix  $\Sigma$  (Concept-Strength) เมตริกซ์ทแยงมุมใช้เพื่อบอกขนาดความสำคัญของแต่ละคอนเซปต์
3. Matrix  $V^T$  (Concept-Sentence) เมตริกซ์เชิงตั้งฉากแทนความสัมพันธ์ของคอนเซปต์และประโยค ใช้เพื่อบอกประโยคสำคัญของแต่ละกลุ่มคอนเซปต์โดยดูจากคอลัมน์ใดมีค่าสูงหมายถึงประโยคนั้นมีความสำคัญมากในคอนเซปต์นั้นๆ

จากงานวิจัยของ Deerwester และคณะ (Deerwester et al., 1990) มีแนวคิดที่ว่าคำที่ปรากฏอยู่ในบริบทต่าง ๆ นั้น มีคุณสมบัติซ่อนอยู่ซึ่งคุณสมบัติเหล่านั้นจะสามารถทำให้เข้าใจความหมายและความสัมพันธ์ของคำนั้นๆ ได้

จากแนวคิดนี้ ผู้วิจัยจึงนำมาประยุกต์ใช้ในการหาความสัมพันธ์ของคำและประโยค โดยมีแนวคิดที่ว่าคำที่ปรากฏในประโยคที่แตกต่างกันมีความสัมพันธ์กันอยู่ กล่าวคือประโยคใดๆ มีคำๆ นั้นปรากฏอยู่ ประโยคอื่นๆ จะมีความหมายหรือมีความสัมพันธ์กัน

### 2.1.10 ความคล้ายคลึงแบบโคไซน์ (Cosine Similarity)

ความคล้ายคลึงแบบโคไซน์ (Cosine Similarity) เป็นทฤษฎีที่ใช้วัดความคล้ายคลึงกันของประโยคในรูปแบบของเวกเตอร์ ซึ่งวิธีนี้จะใช้วัดมุมที่กระทำต่อกันระหว่างประโยค โดยประโยคที่ทำมุมต่อกันน้อยแสดงว่ามีความคล้ายคลึงกันมาก และในทางกลับกันประโยคที่ทำมุมต่อกันมากแสดงว่า จะมีความคล้ายคลึงกันของประโยคน้อยหรืออาจจะไม่มีความเกี่ยวข้องกันเลย ความคล้ายคลึงของ

ประโยคในเอกสารหาได้จากสมการ (2.3) โดย  $\cos \theta$  คือเวกเตอร์ของประโยค A คูณกับเวกเตอร์ ของประโยค B ซึ่งการคูณกันของเวกเตอร์นี้เรียกว่า Dot Product และหารด้วยผลคูณของขนาด (Magnitude) ของเวกเตอร์ A และ B เพื่อต้องการให้ค่าของความคล้ายคลึงกันอยู่ระหว่าง 0 ถึง 1 ดังสมการที่ (2.4)

$$\text{Similarity}(A, B) = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} \quad (2.3)$$

$$\frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}} \quad (2.4)$$

โดย

A แทน เวกเตอร์ของประโยคต้นทาง

B แทน เวกเตอร์ของประโยคปลายทาง

ผลลัพธ์ที่ได้จากการคำนวณค่าคล้ายคลึงกันจะมีค่าอยู่ระหว่าง 0 ถึง 1 โดยที่ค่า 1 หมายถึง ความเหมือนกันและค่า 0 หมายถึง ไม่มีความเกี่ยวข้องกันซึ่งในการจัดกลุ่มประโยคจะเลือกค่าที่ใกล้ 1 มากที่สุดมาใช้ในการพิจารณาในขั้นตอนต่อไป

### 2.1.11 การจัดกลุ่มแบบเคมีน (K-means Clustering)

การจัดกลุ่มแบบเคมีน (K-means Clustering) เป็นการจัดกลุ่มแบบแบ่งส่วนโดยแบ่งกลุ่ม ข้อมูลที่มีลักษณะคล้ายกันอยู่กลุ่มเดียวกันด้วยการหาค่าที่เหมาะสมที่สุดโดยมีขั้นตอนดังนี้

- 1) สุ่มข้อมูลขึ้นมา K ตัวเท่ากับจำนวนกลุ่มที่ต้องการแบ่งเพื่อใช้เป็นตัวแทนในการแบ่งกลุ่ม
- 2) แบ่งกลุ่มโดยวัดค่าความคล้ายคลึงกันซึ่งในงานวิจัยนี้ใช้การวัดความคล้ายคลึงแบบ โคไซน์ระหว่างข้อมูลที่ต้องการจัดกลุ่มกับข้อมูลที่โดนสุ่มมาเป็นข้อมูลตัวแทนและจัดกลุ่มให้แต่ละประโยคอยู่กลุ่มที่มีค่าความคล้ายคลึงกันกับตัวแทนกลุ่มนั้นมากที่สุด
- 3) เกลี่ยข้อมูลแต่ละกลุ่มเพื่อหาตัวแทนกลุ่มใหม่

จากนั้นวนซ้ำเรื่อยๆในขั้นตอนที่ 2) และขั้นตอนที่ 3) จนกว่าข้อมูลสมาชิกของแต่ละกลุ่ม ไม่มีการเปลี่ยนแปลง

ซึ่งจากแนวคิดนี้ ผู้วิจัยจึงนำมาประยุกต์ใช้กับการแบ่งกลุ่มของประโยค โดยอ้างอิงจากแนวคิดที่ว่าประโยคที่คล้ายกันจัดอยู่กลุ่มเดียวกัน โดยกำหนดให้ K คือจำนวนกลุ่มประโยคที่ถูกเลือก ให้นำไปใช้ในการสรุปใจความสำคัญ ยกตัวอย่างเช่น มีประโยคทั้งหมด 100 ประโยค กำหนดการสรุปใจความสำคัญอยู่ที่ระดับร้อยละ 20 กล่าวคือ ผลลัพธ์ที่ได้จะมีทั้งหมด 80 ประโยคจาก 100 ประโยค

เพราะฉะนั้น K จะเท่ากับ 80 โดยแบ่งเป็น 80 กลุ่ม และเลือกตัวแทนประโยคของแต่ละกลุ่มจากค่าน้ำหนักของประโยคที่มากที่สุดมาใช้ในการสรุปใจความสำคัญในงานวิจัยนี้

### 2.1.12 คัดคำ (Cutkum)

คัดคำ (Pucktada Treeratpituk, 2017) คือโปรแกรมตัดคำภาษาไทยที่ผู้วิจัยนำมาใช้ในงานวิจัยนี้เนื่องจากมีค่าความแม่นยำสูงที่สุดเมื่อเทียบกับการเทคนิคการตัดคำแบบอื่นๆ ดังตารางที่ 2.2 คัดคำถูกพัฒนาขึ้นด้วยภาษาไพธอน ใช้หลักการเรียนรู้โดยใช้โครงข่ายประสาทเทียมแบบวนซ้ำ (Recurrent Neural Network) ใช้ข้อมูลจากคลังข้อมูล BEST2010 ของศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ โดยที่คลังข้อมูลมีการตัดคำและตีความวิภาคของคำไว้แล้วในการฝึกฝนโมเดล

ตารางที่ 2.2 เปรียบเทียบประสิทธิภาพของเครื่องมือตัดคำ  
(ที่มา: นายศักยภาพ ผิวเหลือง)

Tools	Number of Word				Correct Word			
	Article	Encyclopedia	News	Novel	Article	Encyclopedia	News	Novel
Test Set	1185090	1113342	1629548	1610350	1185090	1113342	1629548	1610350
PyICU	873422	930211	1274981	1326186	633478	708271	869479	998351
Lexto	1057455	1059362	1582817	1382589	814843	837504	1123291	1152377
Cutkum	1204978	1167399	1733088	1543675	1113956	1060378	1511066	1125931
veer66	1117702	1201293	1714747	1767681	715191	727541	1098774	1055206
SWATH	835419	871698	1234453	1265280	641235	680646	887806	987289

จากตารางที่ 2.2 เปรียบเทียบประสิทธิภาพของเครื่องมือตัดคำนั้น ใช้ข้อมูลทดสอบจากคลังข้อมูล BEST2010 โดยชุดข้อมูลสำหรับทดสอบ (Test set) ประกอบไปด้วยข้อมูล 4 ประเภทคือ ชุดข้อมูลบทความ ชุดข้อมูลสารานุกรม ชุดข้อมูลข่าว และชุดข้อมูลนวนิยาย ซึ่งจำนวนคำจากตารางที่ 2.2 ข้อมูลชุดทดสอบสามารถแยกตามประเภทของชุดข้อมูลดังนี้ ชุดข้อมูลบทความจำนวน 1,185,090 คำ ชุดข้อมูลสารานุกรมจำนวน 1,113,342 คำ ชุดข้อมูลข่าวจำนวน 1,629,548 คำและชุดข้อมูลนวนิยายจำนวน 1,610,350 คำ

จากตารางที่ 2.2 จะเห็นได้ว่าคัดคำสามารถตัดคำและได้จำนวนคำที่ตรงกับข้อมูลที่นำมาทดสอบมากที่สุด และเมื่อนำมาเทียบกับคำที่ถูกตัดแล้วแสดงให้เห็นว่ามีความถูกต้องของผลลัพธ์ใกล้เคียงกับข้อมูลชุดตรวจสอบมากที่สุดอีกด้วย

ตารางที่ 2.3 เปรียบเทียบประสิทธิภาพของเครื่องมือตัดคำ (ต่อ)

(ที่มา: นายศักยภาพ ผิวเหลือง)

Tools	Recall %				Precision %			
	Article	Encyclopedia	News	Novel	Article	Encyclopedia	News	Novel
PyICU	53.45399927	63.61666047	53.35706589	61.99590151	72.5282853	76.14089707	68.19544762	75.2798627
Lexto	68.75790024	75.2243246	68.93267335	71.56065452	77.0569906	79.05739492	70.96783772	83.34920935
Cutkum	93.99758668	95.24279152	92.72914943	69.91840283	92.44616914	90.83252598	87.18922524	72.93834518
veer66	60.34908741	65.34748532	67.42814572	65.52650045	63.98762819	60.56315986	64.0779077	59.69436793
SWATH	54.10854872	61.13539236	54.48173359	61.3089701	76.75609485	78.08277637	71.9189795	78.02928996

จากตารางที่ 2.3 เป็นการเปรียบเทียบประสิทธิภาพของเครื่องมือตัดคำโดยใช้ค่าความครบถ้วน (Recall) สมการที่ (2.5) ค่าความแม่นยำ (Precision) สมการที่ (2.6)

$$\text{Recall \%} = \frac{\text{Correct Word}}{\text{Referenced Word}} \times 100 \quad (2.5)$$

$$\text{Precision \%} = \frac{\text{Correct Word}}{\text{Output Word}} \times 100 \quad (2.6)$$

โดยที่

**Correct Word** หมายถึง จำนวนคำที่โปรแกรมตัดคำได้ถูกต้อง

**Output Word** หมายถึง จำนวนคำทั้งหมดที่โปรแกรมตัดคำได้

**Referenced Word** หมายถึง จำนวนคำที่ถูกแบ่งในไฟล์เอกสารชุดข้อมูลทดสอบ

ตารางที่ 2.4 เปรียบเทียบประสิทธิภาพของเครื่องมือตัดคำ (ต่อ)

(ที่มา: นายศักยภาพ ผิวเหลือง)

Tools	F1 score			
	Article	Encyclopedia	News	Novel
PyICU	61.54717582	69.31760517	59.87056765	67.99514802
Lexto	72.67127304	77.09324418	69.93545254	77.00638068
Cutkum	93.21542316	92.98539378	89.87389655	71.39645374
veer66	62.11511939	62.86442571	65.71035151	62.47461909
SWATH	63.47262002	68.57756015	61.99760405	68.66592712

ตารางที่ 2.4 เป็นการเปรียบเทียบประสิทธิภาพของเครื่องมือตัดคำโดยนำค่าความครบถ้วน (Recall) และค่าความแม่นยำ (Precision) ที่คำนวณได้จากสมการที่ (2.5) และสมการที่ (2.6) มาคำนวณเพื่อหาประสิทธิภาพความถูกต้องในการตัดคำโดยวัดค่า F-measure ดังสมการที่ (2.7)

$$F\text{-measure} = \frac{(2 \times \text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})} \quad (2.7)$$

จากตารางที่ 2.4 แสดงให้เห็นว่าคัตคำมีประสิทธิภาพในการตัดคำสูงสุดเมื่อเทียบกับโปรแกรมตัดคำอื่นๆ ดังนั้นในงานวิจัยนี้ผู้วิจัยจึงเลือกใช้โปรแกรมคัตคำเข้ามาช่วยในการตัดคำเพื่อใช้ในการสรุปใจความสำคัญ

### 2.1.13 การวัดประสิทธิภาพของการสรุปใจความสำคัญ

การวัดประสิทธิภาพการสรุปใจความสำคัญ มีการใช้มาตรฐานการวัดหลายแบบ ซึ่งในงานวิจัยนี้ใช้การวัดตามหลักมาตรฐานทั่วไป ซึ่งประกอบไปด้วยการวัดค่าความถูกต้อง วัดค่าความแม่นยำ ค่าความระลึกรวมถึง F-measure ซึ่งแสดงดังนี้

- 1) ค่าความถูกต้อง (Accuracy) เป็นการคำนวณข้อมูลที่ได้จากทุกคลาสรวมกัน จากสมการที่ (2.8)

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (2.8)$$

โดย TP = True Positive คือ ประโยคที่ระบบบอกใช่ มนุษย์บอกใช่

TN = True Negative คือ ประโยคที่ระบบบอกไม่ใช่ มนุษย์บอกไม่ใช่

FP = False Positive คือ ประโยคที่ระบบบอกใช่ มนุษย์บอกไม่ใช่

FN = False Negative คือ ประโยคที่ระบบบอกไม่ใช่ มนุษย์บอกใช่

- 2) ค่าความแม่นยำ (Precision) เป็นการคำนวณหาความแม่นยำจากตัวแปรใน 1) ซึ่งจากการคำนวณนี้เป็นการหาค่าความแม่นยำของระบบกล่าวได้ว่าค่ายิ่งมากระบบยิ่งมีความแม่นยำมาก สามารถคำนวณจากสมการที่ (2.9)

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2.9)$$

- 3) ค่าความระลึกรวม (Recall) เป็นการคำนวณหาค่าความครอบคลุมของระบบ หมายถึงระบบไม่มีความตกหล่น หรือตกหล่นน้อย จากตัวแปรใน 1) สามารถคำนวณได้จากสมการที่ (2.10)

$$\text{Recall} = \frac{TP}{TP + FN} \quad (2.10)$$



- 4) ค่าวัดประสิทธิภาพ (F-measure หรือ F1 score หรือ F-score) เป็นการหาค่าเฉลี่ย โดยคำนวณจากค่าความระลึกและค่าความแม่นยำ ดังสมการที่ (2.11)

$$F1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (2.11)$$

## 2.2 งานวิจัยที่เกี่ยวข้อง

### 2.2.1 งานวิจัยที่เกี่ยวกับการหาคำสำคัญและการตัดคำ

เนื่องจากขั้นตอนในการสรุปใจความสำคัญมีทั้งหมด 3 ขั้นตอนหลักคือ 1.ขั้นตอนการเตรียมข้อมูลซึ่งต้องสร้างพจนานุกรม 2. ขั้นตอนการวิเคราะห์ประโยคเลือกคำสำคัญของเอกสาร 3. ขั้นตอนการสรุปใจความสำคัญ ดังนั้นในหัวข้องานวิจัยที่เกี่ยวข้องจะจัดเรียงตามหัวข้อดังนี้

- 1) งานวิจัยหัวข้อ “การเพิ่มประสิทธิภาพของการตัดคำภาษาไทยโดยการพิจารณาบริบทรอบข้าง” (กรรณิกา บุญเกษม, 2554)

ผู้วิจัยได้นำเสนอการเพิ่มประสิทธิภาพของการตัดคำภาษาไทยโดยพิจารณาบริบทรอบข้าง โดยใช้ชุดข้อมูลจากคลังข้อมูล BEST2010 ของศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ (ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ) มาสร้างเป็นพจนานุกรมและกำหนดรูปแบบของการค้นหาในพจนานุกรมทั้งหมด 3 รูปแบบคือ 1.คำที่ค้นพบในพจนานุกรม 1 รูปแบบ 2.คำที่ค้นพบในพจนานุกรมมากกว่า 1 รูปแบบ และ 3. คำที่ค้นไม่พบในพจนานุกรม โดยสนใจเฉพาะรูปแบบที่ 2. คือคำที่ค้นพบในพจนานุกรมมากกว่า 1 รูปแบบเท่านั้น โดยมีการระบุตำแหน่งของคำเริ่มต้นและคำสิ้นสุดไว้ด้วย ซึ่งการตัดคำก่ากวมในขั้นตอนนี้จะตัดคำที่สามารถเป็นได้ทุกรูปแบบโดยอ้างอิงจากพจนานุกรมที่สร้างขึ้นจากคลังข้อมูล BEST2010 และคำนวณค่าความน่าจะเป็นโดยอ้างอิงค่าจากชุดความถี่ของการเกิดคำร่วมกันแล้วนำมาคำนวณค่าความน่าจะเป็นด้วยทฤษฎีของเบย์

- 2) งานวิจัยหัวข้อ “การค้นหาคำสำคัญในเอกสารภาษาไทยโดยใช้เทคนิคการค้นหาในรูปแบบความสัมพันธ์ในเหมืองข้อมูล” (วัฒน์พัชร แก้วลิ้ม, 2557)

ผู้วิจัยได้นำเสนอวิธีการค้นหาคำสำคัญของภาษาไทยโดยใช้วิธีการหารูปแบบคำที่ปรากฏในเอกสารตามลำดับการเกิดของคำ โดยใช้เทคนิคเหมืองข้อมูลการสร้างตารางความสัมพันธ์ของข้อมูลในรูปแบบ Sequential Pattern Mining โดยพิจารณาจากตัวแปรหลายตัว คือ จำนวนคำในประโยค ลำดับการเกิดคำนั้นๆในแต่ละประโยค ความถี่ของคำ เพื่อนำมาเพิ่มประสิทธิภาพในการคัดเลือกคำเพื่อมาใช้เป็นคำสำคัญของประโยคเพื่อสร้างแบบจำลองข้อมูล และใช้เทคนิคการตัดสินใจแบบ

Decision Tree โดยใช้ข้อมูลข่าวจากเว็บไซต์ จำนวน 200 ข่าวในการฝึกสอนและอีก 50 ข่าวในการทดสอบและเปรียบเทียบการวัดผลกับวิธีการอื่นๆอีก 3 วิธีได้แก่ การใช้เทคนิคโครงข่ายประสาทเทียมด้วยการเรียนรู้แบบ Back-Propagation การใช้เทคนิคทางสถิติประเมินค่าความสำคัญหรือน้ำหนักของวลีในเอกสาร (TF-IDF) และการสกัดวลีสำคัญออกจากเอกสารโดยหาวลีที่เกิดร่วมกันในเอกสาร ผลลัพธ์ที่ได้จากการวัดผลด้วยวิธี Precision and Coverage ร่วมกับการนำค่าที่มีการกำกับค่าสำคัญไว้แล้วในหน้าเว็บเพจของข้อมูลข่าวที่นำมาทดสอบ มาวัดผลรวมกันสรุปว่า การค้นหาคำสำคัญด้วยวิธี Sequential Pattern Mining ร่วมกับต้นไม้ตัดสินใจได้ผลลัพธ์ที่แม่นยำมากกว่าวิธีอื่นๆ

## 2.2.2 งานวิจัยที่เกี่ยวกับการย่อเอกสารและสรุปใจความสำคัญ

- 1) งานวิจัยหัวข้อ “Generic text summarization using relevance measure and latent semantic analysis” (Gong & Liu, 2001)

งานวิจัยนี้ได้เสนอวิธีการสรุปใจความเอกสารภาษาอังกฤษโดยอาศัยหลักการคำนวณหาค่าน้ำหนักของคำสำคัญจากประโยคด้วยวิธีการแยกตัวประกอบเมตริกซ์ในการวิเคราะห์ความหมายแฝงของคำหลังจากแยกตัวประกอบของเมตริกซ์ด้วย SVD แล้ว งานวิจัยนี้ผู้วิจัยเลือกใช้ Matrix  $V^T$  ที่เกิดจากความสัมพันธ์ของคอนเซปต์และประโยคมาใช้ในการเลือกประโยคในการสรุปใจความสำคัญ โดยในงานวิจัยนี้ผู้วิจัยเลือก 1 ประโยคจาก 1 คอนเซปต์มาใช้ในการสรุปใจความสำคัญ โดยให้เหตุผลว่าแถวแรกของคอนเซปต์จะมีความสำคัญมากที่สุด และเลือกประโยคที่สองจากคอนเซปต์รองลงมาจนกว่าจะครบตามจำนวนประโยคที่ต้องการนำมาใช้สรุปใจความสำคัญ ข้อเสียของวิธีนี้คือ ถ้ามีจำนวนคอนเซปต์น้อย แต่ต้องการเลือกประโยคมาสรุปอย่างน้อย 10 คอนเซปต์เพื่อเอาประโยคที่น้ำหนักมากที่สุดของแต่ละคอนเซปต์ไปเป็นตัวแทนประโยค ประโยคที่ได้จากคอนเซปต์หลังๆจะไม่ใช่อประโยคที่สำคัญจริงอาจส่งผลให้การสรุปใจความสำคัญนั้นยังไม่แม่นยำมากพอ

ตารางที่ 2.5 ตัวอย่างการเลือกประโยคจาก Matrix  $V^T$  ของ Gong และคณะ

	S1	S2	S3	S4	S5	S6	S7	S8
Con1	-0.166277307	0.329278815	-0.048739176	0.33778914	-0.462897161	0.0440453	-0.034253403	0.435819947
Con2	-0.015556927	0.063506319	-0.164748164	-0.155911683	0.058741849	0.514714681	-0.352328975	-0.017878014
Con3	-0.230676508	0.525413948	0.305102954	-0.098260335	0.086227894	-0.182684281	-0.371939014	-0.552924509
Con4	-0.155738675	0.335634671	0.275945112	-0.322929008	0.540902628	0.039373298	0.39363517	0.398600494
Con5	-0.061187891	0.113425283	0.093486746	-0.030109106	0.037498008	-0.012156418	-0.046670737	-0.023402442

จากตารางที่ 2.5 ตัวอย่างการเลือกประโยคของ Gong และคณะ เลือกจากน้ำหนักของประโยคในแต่ละคอนเซปต์ที่มีน้ำหนักมากที่สุดมาใช้ในการสรุปใจความสำคัญ



ตารางที่ 2.7 ตัวอย่างการเลือกประโยคจากงานวิจัยของ Murray และคณะ

	S1	S2	S3	S4	S5	S6	S7	S8
Con1	-0.166277307	0.329278815	-0.048739176	0.33778914	-0.462897161	0.0440453	-0.034253403	0.435819947
Con2	-0.015556927	0.063506319	-0.164748164	-0.155911683	0.058741849	0.514714681	-0.352328975	-0.017878014
Con3	-0.230676508	0.525413948	0.305102954	-0.098260335	0.086227894	-0.182684281	-0.371939014	-0.552924509
Con4	-0.155738675	0.335634671	0.275945112	-0.322929008	0.540902628	0.039373298	0.39363517	0.398600494
Con5	-0.061187891	0.113425283	0.093486746	-0.030109106	0.037498008	-0.012156418	-0.046670737	-0.023402442

จากตารางที่ 2.6 ของ Matrix  $\Sigma$  จะสังเกตได้ว่ามีค่าความสำคัญของแต่ละคอนเซปต์ที่แตกต่างกัน ดังนั้นการคำนวณของ Murray และคณะ จะคำนวณจาก ค่าที่ได้จาก Matrix  $\Sigma$  คิดเป็นร้อยละของจำนวนคอนเซปต์ทั้งหมด แล้วนำมาเลือกประโยคจาก Matrix  $V^T$  โดยที่ 1 คอนเซปต์สามารถมีได้มากกว่า 1 ประโยคดังตารางที่ 2.7 จำนวนประโยคขึ้นอยู่กับค่าของ Matrix  $\Sigma$  ในแต่ละคอนเซปต์นั้นๆ โดยแต่ละคอนเซปต์จะมีค่าต่างกัน

- 4) งานวิจัยหัวข้อ “การเลือกประโยคสำคัญในการสรุปความภาษาไทยโดยใช้แบบจำลองแบบลำดับชั้น” (เพ็ญศิริ ลีตระกูล, 2551)

ผู้วิจัยนำเสนอเทคนิคในการย่อความโดยใช้ค่าความถี่ของคำทั้งหมดที่เกิดขึ้นในเอกสารและคำนวณหาน้ำหนักของคำ เพื่อสร้างเป็นกลุ่มคำสำคัญขึ้นมา จากนั้นคำนวณหาความคล้ายกันของประโยคแบบโคไซน์และระยะห่างของแต่ละประโยค แล้วใช้การค้นหาแบบ  $A^*$  เพื่อหาลำดับของประโยคที่มีระยะสั้นที่สุดจากกลุ่มคำสำคัญ โดยเริ่มต้นคำนวณที่ประโยคชื่อเรื่อง แล้วทำการแบ่งกลุ่มประโยค โดยประโยคใดที่มีค่าซ้ำกันจะเป็นกลุ่มเดียวกัน ถ้าไม่มีค่าซ้ำกันจะจัดแยกเป็นอีกกลุ่มหนึ่ง จากนั้นเลือกตัวแทนของแต่ละกลุ่มประโยค โดยเลือกประโยคลำดับที่อยู่แกนกลางของกลุ่มเป็นประโยคสำคัญ สำหรับผลการทดลองข่าวเกษตรกรจำนวน 10 ข่าว โดยประเมินจากการเลือกประโยคของคนสามคน ได้ผลออกมาดีสุด โดยผลค่าความแม่นยำเท่ากับ 68.57% ค่าระลึกเท่ากับ 51.95% และค่าประสิทธิภาพโดยรวมของระบบเท่ากับ 56.72%

- 5) งานวิจัยหัวข้อ “การย่อความเอกสารภาษาไทยเชิงความหมายโดยใช้กราฟโนภาพ” (มงคล ยูสเปรมานนท์, 2551)

จากงานวิจัยนี้ ผู้วิจัยนำเสนอวิธีการย่อความสำหรับภาษาไทยโดยใช้ความสัมพันธ์เชิงความหมายของประโยคที่ได้จากประโยคหัวเรื่องและประโยคเนื้อเรื่อง ซึ่งได้จากเทคนิคการแปลงประโยคให้อยู่ในรูปกราฟโนภาพ (Conceptual Graph) แล้วใช้ตัวดำเนินการของกราฟในการรวมและลดรูปเพื่อใช้สร้างส่วนของการย่อความโดยมีกระบวนการและขั้นตอนการสร้างเริ่มต้นจาก การกำหนดขอบเขตของประโยคของเอกสาร โดยผู้ที่มีความรู้เกี่ยวกับการใช้หลักภาษาไทย เนื่องจากลักษณะของภาษาไทยที่ไม่มีมาตรฐานในการกำหนดขอบเขตของประโยค จากนั้นจึงนำประโยคแต่ละประโยคเข้าสู่กระบวนการตัดคำโดยใช้โปรแกรมตัดคำ SWATH โดยใช้อัลกอริทึมการตัดคำแบบเลือกคำที่ยาว

ที่สุด หลังจากนั้นทำการกำกับหน้าที่ของคำ และทำการหาคำที่สอดคล้องกันทางความหมายโดยใช้พจนานุกรม สร้างเป็นดัชนีคำ จากนั้นจึงนำคำที่ได้จากการตัดคำจากประโยคแต่ละประโยคที่ได้ทำการกำกับคำอย่างถูกต้องแล้ว มาสร้างเป็นกราฟมโนภาพจากกฎที่สร้างขึ้น แล้วจึงใช้ตัวดำเนินการของกราฟเพื่อรวมกราฟเข้าด้วยกัน จากดัชนีของคำที่สร้างไว้ แล้วลดรูปกราฟ เพื่อให้ได้เนื้อความที่กระชับ ขั้นตอนสุดท้ายคือการแปลงกราฟที่มีความสัมพันธ์กันออกมาในรูปของข้อความ สรุปเป็นบทย่อความ

บทความที่ใช้ทดสอบและประเมินผลนั้น ผู้วิจัยเลือกใช้บทความข่าวเกษตรกรรมจำนวน 30 บทความและทำการย่อบทความ แบ่งประโยคและเรียบเรียงให้เป็นประโยคอย่างง่ายโดยคณะอาจารย์ผู้สอนภาษาไทย จำนวน 3 ท่าน เพื่อใช้เป็นบทย่อความมาตรฐานที่ใช้ประเมินผลการย่อความที่ได้จากงานวิจัย งานวิจัยนี้วัดประสิทธิภาพโดยเปรียบเทียบกับผลการทดลองของวิธี A\*Search กับวิธีที่นำเสนอได้ค่าความแม่นยำเฉลี่ยอยู่ที่ร้อยละ 42.49, 55.94 และค่าความระลึกเฉลี่ยร้อยละ 33.51, 70.52 และประสิทธิภาพโดยรวมของระบบอยู่ที่ร้อยละ 35.79 และ 60.06 ตามลำดับ ซึ่งการใช้กราฟมโนภาพเข้ามาช่วยเลือกประโยคได้ดีกว่าวิธี A\*Search โดยได้ชุดของประโยคตรงตามอาจารย์ผู้สอนภาษาไทยเลือกมากกว่า

### 2.2.3 งานวิจัยการประมวลผลภาษาธรรมชาติอื่น ๆ ที่นำมาประยุกต์ใช้

- 1) งานวิจัยหัวข้อ “การวิเคราะห์ความหมายแฝงและการเรียนรู้เครื่องจักรสำหรับระบบถามตอบอัตโนมัติภาษาไทย” (กมลวรรณ โพธิ์สาย, 2552)

งานวิจัยนี้เป็นงานวิจัยเกี่ยวกับระบบถามตอบอัตโนมัติภาษาไทยโดยใช้การหาความหมายแฝง (Latent Similarity Analysis: LSA) และแบบจำลองคอนดิชันนอลแรนดอมฟิลด์ (Conditional Random Fields: CRFs) เพื่อเลือกคำตอบในฐานข้อมูล และผู้วิจัยได้เลือกการวิเคราะห์การหาความหมายแฝงมาเพื่อใช้ในการวิเคราะห์ประโยคที่มีความหมายคล้ายคลึงกันระหว่างประโยคคำถามและเอกสารคำตอบ โดยที่งานวิจัยนี้ผู้วิจัยได้เลือกพิจารณาความสัมพันธ์ 3 รูปแบบคือ ความสัมพันธ์ที่มีความหมายเหมือนกัน (Synonym) ความสัมพันธ์ที่มีความหมายทั่วไป (Hypernym) และความสัมพันธ์ที่มีความหมายเฉพาะเจาะจง (Hyponym) ซึ่งผู้วิจัยมองว่าความสัมพันธ์ทั้งสามแบบนี้ตรงประเด็นมากกว่าความสัมพันธ์แบบอื่นๆ และได้ใช้วิธีการแบ่งประเภทของประโยคคำถามโดยใช้แบบจำลองคอนดิชันนอลแรนดอมฟิลด์ ส่วนการเลือกประโยคคำตอบนั้นผู้วิจัยใช้โมเดลในการเรียนรู้เข้ามาช่วยในการตัดสินใจ โดยเลือกจากลักษณะเฉพาะของประเภทคำถามและคำตอบที่สังเกตได้มาเลือกคำตอบให้กับมนุษย์

สำหรับข้อมูลที่นำมาใช้ในการประเมินประสิทธิภาพของระบบนั้น ผู้วิจัยได้รวบรวมข้อมูลจากเอกสารเกี่ยวกับการตอบปัญหาคอมพิวเตอร์เบื้องต้นจากเว็บไซต์ต่างๆจำนวน 500 ชุด และถูกแบ่ง

เนื้อหาให้ได้ใจความเพื่อตอบปัญหาของมนุษย์ ซึ่งคำถามจะประกอบไปด้วย คำถามประเภทอะไร (What) ประเภทที่ไหน (Where) ประเภทเท่าไร (How) และประเภททำไม (Why) และให้มนุษย์เพียง 1 คนเป็นผู้เลือกคำตอบให้กับคำถามทั้งหมดเพียง 1 คำถามต่อ 1 คำตอบที่คิดว่าถูกต้องที่สุดเท่านั้น จากการประเมินผล โดยใช้การประเมินผลวัดค่าความแม่นยำ (Precision) และค่าความครบถ้วน (Recall) และนำค่าทั้งหมดที่ได้มาคำนวณหาค่า F-Measure โดยประเมินจากส่วนคำตอบที่เหมือนกันจากมนุษย์และระบบเลือกให้ ผลที่ได้คือ ระบบสามารถตอบคำถามประเภทอะไร (What) ได้ประสิทธิภาพสูงสุดถึง 84% เนื่องจากคำถามประเภทอะไรเป็นคำถามเชิงความหมายและผู้วิจัยได้นำวิธีการวิเคราะห์เชิงความหมายมาวิเคราะห์จึงทำให้ได้ประสิทธิภาพสูงสุด จากงานวิจัยนี้สามารถสรุปได้ว่าประสิทธิภาพของขั้นตอนการเลือกเอกสารคำตอบโดยใช้วิธีการหาความหมายแฝง (LSA) ได้ประสิทธิภาพสูงสุดถึง 88% และในส่วนของเลือกระบบคำถามโดยใช้แบบจำลองคอนดิชันนอลแรนดอมฟิลด์นั้นมีประสิทธิภาพสูงถึง 99% เนื่องจากลักษณะการเลือกระบบคำถามที่ได้เลือกมาพิจารณานั้นตรงจุดในการแบ่งประเภทคำถามคือส่วนใหญ่คำถามที่เลือกมาเป็นคำถามเชิงความหมายจึงทำให้การแยกประสิทธิภาพแทบไม่มีข้อบกพร่องเลย ส่วนการเลือกระบบคำตอบโดยใช้แบบจำลองคอนดิชันนอลแรนดอมฟิลด์ได้ประสิทธิภาพ 76% เนื่องจากลักษณะของคำตอบที่เลือกมาอยู่ในเกณฑ์ดีแต่เนื่องจากใช้มนุษย์เพียง 1 คนในการพิจารณาเลือกคำตอบอาจจะทำให้เลือกคำตอบไม่ตรงจุด หรืออาจกล่าวได้ว่าการเลือกคำตอบจากมนุษย์เพียง 1 คนนั้นส่งผลต่อประสิทธิภาพของการเลือกคำตอบเป็นอย่างมาก

ตารางที่ 2.8 เปรียบเทียบงานวิจัยที่เกี่ยวข้องด้านการสรุปใจความสำคัญของเอกสารภาษาไทย

งานวิจัย	เทคนิคที่ใช้	ผลสรุป
“A Practical Text Summarizer by Paragraph Extraction for Thai” (Chuleerat, 2003)	รวมคุณสมบัติเฉพาะที่ (Local Property) และคุณสมบัติโดยรวม (Global Property) เข้าด้วยกันโดยใช้คุณสมบัติเฉพาะที่สำหรับแบ่งคำ กลุ่มคำส่วนคุณสมบัติโดยรวมใช้ในการหาความสัมพันธ์ระหว่างประโยค	งานวิจัยนี้ทดลองกับข่าว เกษตรกรรม ข่าวทั่วไปและ นิตยสารแพชั่น มีผลค่าความแม่นยำ 60%, 51.8% และ 53.0% ความครบถ้วน 44%, 38.5% และ 33.0% ประสิทธิภาพความถูกต้อง 50.9%, 43.1% และ 40.4% ตามลำดับ
“การเลือกประโยคสำคัญ ในการสรุปความภาษาไทย โดยใช้แบบจำลองแบบ ลำดับชั้น” (เพ็ญศิริ ลีตระกูล, 2551)	ใช้ค่าความถี่ของคำในการคำนวณหา ค่าน้ำหนักของคำ และใช้โคไซน์ในการหาความคล้ายคลึงกันของ ประโยคและใช้การค้นหาแบบ A* ในการหาลำดับประโยค ประโยคไหนมี คำซ้ำกันจะจับเป็นกลุ่มเดียวกัน	งานวิจัยนี้ทดลองกับข่าว เกษตรกรรม 10 ข่าว มีผลค่าความแม่นยำเท่ากับ 68.57% ค่าระลึกเท่ากับ 51.95% และ ค่าประสิทธิภาพโดยรวมของ ระบบเท่ากับ 56.72%
“การย่อความเอกสาร ภาษาไทยเชิงความหมาย โดยใช้กราฟมโนภาพ” (มงคล ยูสเปรมานนท์, 2551)	ใช้แนวคิดกราฟมโนภาพในการหา ความสัมพันธ์เชิงความหมายระหว่าง ชื่อเรื่องกับเนื้อเรื่อง	งานวิจัยนี้ทดลองกับข่าว เกษตรกรรม ผลลัพธ์ได้จากการสรุปเลือกจากเส้นทางที่ กราฟผ่านโหนดที่มีความสำคัญ นั้นอยู่ ค่าความแม่นยำเฉลี่ยอยู่ที่ 42.49%, 55.94% และ ประสิทธิภาพโดยรวมของ ระบบอยู่ที่ 35.79% และ 60.06% ตามลำดับ

## บทที่ 3

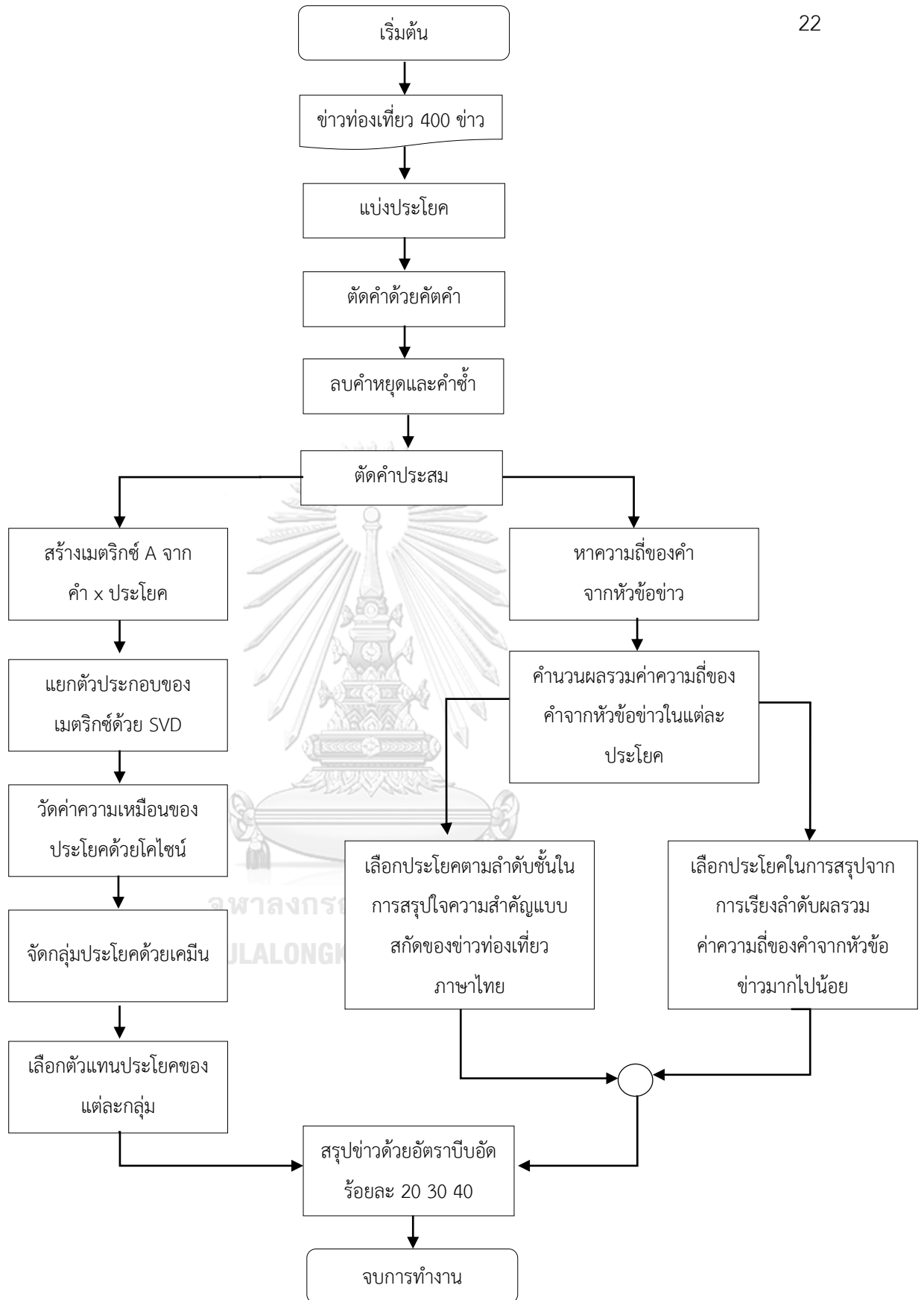
### แนวคิดและวิธีการ

วิทยานิพนธ์นี้ได้ทำการสรุปใจความสำคัญของข่าวท่องเที่ยวไทย โดยแบ่งการดำเนินงานออกเป็น 2 ส่วนคือ 1. การสร้างรายการคำประสมจากข่าวท่องเที่ยว 2. การสรุปใจความสำคัญแบบสกัดของข่าวท่องเที่ยวไทยโดยใช้วิธีการหาคำสำคัญจากการแยกตัวประกอบเมตริกซ์และจัดกลุ่มประโยคด้วยเคมีน และการสรุปใจความสำคัญโดยใช้การคำนวณค่าน้ำหนักของคำสำคัญโดยหาความถี่ของคำและเลือกประโยคเป็นลำดับขั้น จากการศึกษางานวิจัยที่ผ่านมาพบว่าการสรุปใจความสำคัญนั้นจะมีประสิทธิภาพหรือไม่ขึ้นอยู่กับการตัดคำด้วยเช่นกัน เนื่องจากคำเป็นหน่วยที่เล็กที่สุดของประโยคและคำยังสามารถแบ่งออกได้เป็นหลายความหมาย สำหรับงานวิจัยนี้ผู้วิจัยใช้ข้อมูลนำเข้าคือข่าวการท่องเที่ยว ดังนั้นเพื่อให้สรุปใจความสำคัญมีประสิทธิภาพมากยิ่งขึ้น ผู้วิจัยจึงได้สร้างคลังข้อมูลรายการคำประสมที่สกัดจากข่าวท่องเที่ยวภาษาไทยเพื่อนำมาเพิ่มความหมายของคำที่ได้จากการตัดคำจากโปรแกรมตัดคำ และนำมาใช้ในการหาค่าน้ำหนักของประโยค เพื่อเลือกประโยคมาใช้ในการสรุปใจความสำคัญ ในบทที่ 3 นี้จะแสดงภาพรวมของวิธีการ และขั้นตอนการทำงานหลักของวิธีการสรุปใจความสำคัญที่นำเสนอ

#### 3.1 ภาพรวมของวิธีการ

ภาพรวมของวิธีการที่งานวิจัยต้องการนำเสนอ แบ่งการดำเนินงานออกเป็น 2 ส่วนคือ การสร้างคลังข้อมูลรายการคำประสมที่สกัดจากข่าวท่องเที่ยวภาษาไทยและการสรุปใจความสำคัญแบบสกัดของข่าวท่องเที่ยวไทย ดังรูปที่ 3.1

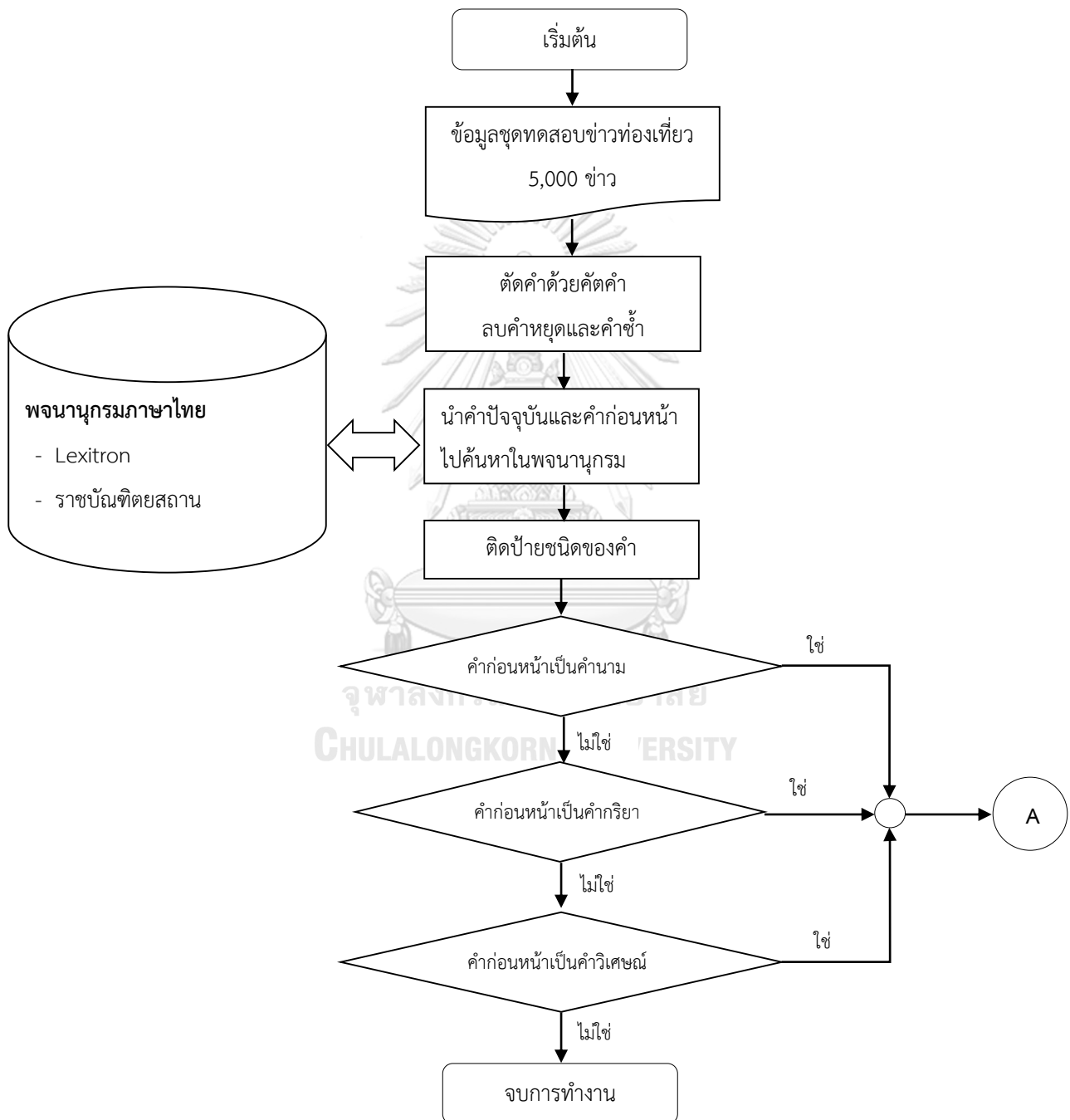




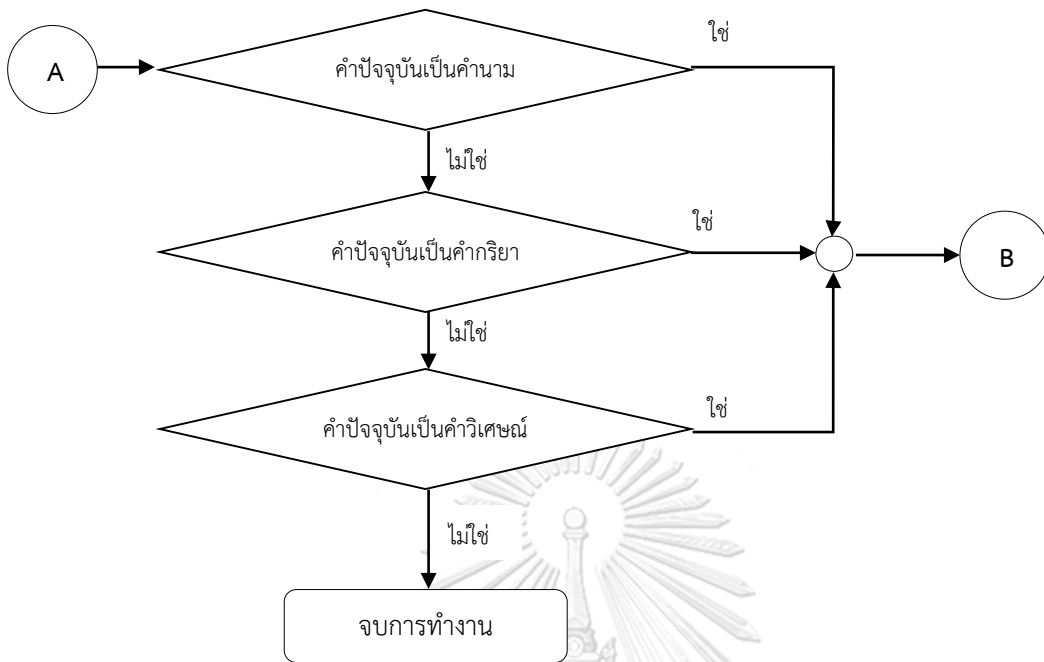
รูปที่ 3.1 ภาพรวมของแนวคิดที่นำเสนอ

### 3.2 การสร้างรายการคำประสม

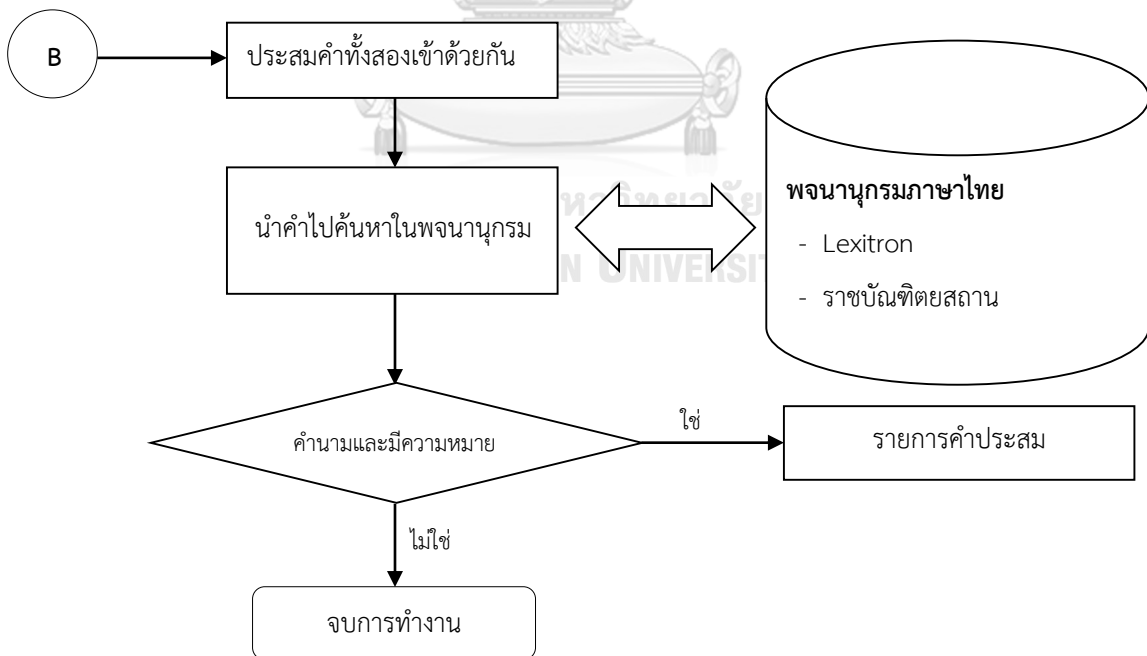
ในขั้นตอนการเตรียมข้อมูลรายการคำประสมใช้ข้อมูลข่าวการท่องเที่ยวไทยจาก <https://www.manager.co.th/travel> และ <https://news.kapook.com/travel> จำนวน 5,000 ข่าว โดยมีขั้นตอน (ก) (ข) และ (ค) ในรูปที่ 3.2 ตามลำดับ



(ก) ขั้นตอนการสร้างรายการคำประสม



(ข) ขั้นตอนการสร้างรายการคำประสม (ต่อ)



(ค) ขั้นตอนการสร้างรายการคำประสม (ต่อ)

รูปที่ 3.2 ขั้นตอนการสร้างรายการคำประสม

จากรูปที่ 3.2 แสดงขั้นตอนการเตรียมข้อมูลข่าวจากเว็บไซต์มาสร้างรายการคำประสม โดยแปลงข่าวจากไฟล์ .html ให้เป็นข้อความตัวอักษรธรรมดาและเก็บเป็นไฟล์ .txt

```

306 <div class="article-content">
307 <div class="m-detail-container">
308 <div class="
<strong>Facebook :Travel @ Manager</strong></span></U></A><br>เป็นอุทยานแห่งชาติประเภทชายฝั่งผสมหมู่เกาะในทะเล ซึ่งก็หมายถึงว่ามีพื้นที่ครอบคลุมทั้ง
รวมไปถึงเกาะต่างๆ ในท้องทะเล โดยเกาะที่นักท่องเที่ยวรู้จักกันดีก็คือเกาะเสม็ด ที่ขึ้นชื่อว่าเป็นเกาะที่สวยงามที่สุดในระยอง และยังมีเกาะจันทร์ เกาะทะลุ เกาะกัญ
width="700" data-height="467" alt="กางเต็นท์นอนฟังเสียงคลื่น"><br>ส่วนใหญ่แล้วนักท่องเที่ยวมักจะเลือกไปเที่ยวชมน้ำใสทรายสวยกันที่เกาะเสม็ดเสี
หญ้า" </strong>ส่วนที่ยูบมนแผ่นดินกัน <br> <br>จากตัวเมืองระยองเลยเลียบผ่านหาดแม่รำพึงใช้เวลาขับรถไม่เกิน 30 นาที ก็มาถึงเขาแหลมหญ้าแล้ว นั
src="https://pics.mgonline.com/pics/Images/561000003495503.JPEG" data-width="700" data-height="467" alt="เอาเต็นท์มาเองหรื
width="700" data-height="467" alt="สะพานไม้จุดถ่ายภาพยอดเยี่ยม"><br>ที่นี้เราสามารถนำเต็นท์มากางนอนเล่นฟังเสียงทะเล หรือเข้าเดินกับทางอุทย
ต้นมะขาม) บริเวณท่าเรือ และบริเวณอาคาร 80 พรรษา นอกจากนี้ก็ยังสามารถใช้บริการบ้านพักของทางอุทยานฯ เพื่อการพักผ่อนกับครอบครัวและเพื่อนฝูงได้ด้วย
สีขาวสะอาดตาติดกับสีฟ้าของน้ำทะเล ถือเป็นแลนด์มาร์กของเขาแหลมหญ้าที่ว่าได้ ใครๆ ก็มักจะมาถ่ายรูปกันตรงนี้ ยิ่งช่วงเย็นๆ ก็ยิ่งได้บรรยากาศแดดรมลมดกที่โร
height="467" alt="บรรยากาศยามเย็น "><br><br>บริเวณเสียบทะเลจะเป็นการเดินทางไปตามโขดหินแกรนิต ตรงไหนที่เดินค่อนข้างยากทางอุทยานฯ ได้ทำสะพานไม้ไว้ให้เดินได้อย่างสะดวก
กับหิน ไม้ดีต้นไม้อายุขึ้นเนื่องจากบริเวณเป็นปลายแหลมที่มีลมพัดแรงตลอดเวลา ต้นไม้ใหญ่ทานกระแสนลมไม่ได้จึงมีเพียงทงหยูขึ้นเป็นบริเวณกว้างงั้นที่มาจาก
data-height="467" alt="วิวดูงามไม่น้อย "><br><br>ด้วยวิวที่สวยงามและอากาศริมทะเลที่สะอาดสดชื่น ทำให้อากาศว
น้อย<br><br><span style="color: #0000ff;"><strong>"อุทยานแห่งชาติเขา
รายละเอียดได้ที่โทร.<strong>0 3865 3034</strong></span><br><br><span style="color: #ff0000;"><strong>"สามารถส่งข้อมูลข่าวสารด้านการท่องเที่ยวใ
HREF="https://www.facebook.com/travel.astvmanager/?ref=bookmarks" CLASS="innerlink" TARGET="_blank"><U> <span style="col
</div>

```

รูปที่ 3.3 ตัวอย่างข้อความเนื้อหาข่าวท่องเที่ยวจากไฟล์ .html

Topic: "เขาแหลมหญ้า" เทียบทะเลระยองสุดชิล กางเต็นท์นอนฟังเสียงคลื่น  
อีกหนึ่งแหล่งท่องเที่ยวทางทะเลที่อยู่ไม่ไกลกรุงเทพฯ และยังมีบรรยากาศสุดชิล แคมป์ยังสามารถกางเต็นท์  
ริมทะเลชมบรรยากาศและฟังเสียงคลื่นได้อย่างเต็มที่ก็คือ "อุทยานแห่งชาติเขาแหลมหญ้า-หมู่เกาะเสม็ด"  
อ.เมือง จ.ระยอง "อุทยานแห่งชาติเขาแหลมหญ้า-หมู่เกาะเสม็ด" เป็นอุทยานแห่งชาติประเภทชายฝั่งผสม  
หมู่เกาะในทะเล ซึ่งก็หมายถึงว่ามีพื้นที่ครอบคลุมทั้งบนแผ่นดินและในท้องทะเล บนแผ่นดินก็คือบริเวณ  
เขาแหลมหญ้าและชายทะเลด้านทิศตะวันตกของเขาแหลมหญ้า (อยู่ระหว่างหาดแม่รำพึงและอ่าวเพ) รวม  
ไปถึงเกาะต่างๆ ในท้องทะเล ...

รูปที่ 3.4 ตัวอย่างข้อความเนื้อหาข่าวท่องเที่ยวจากไฟล์ .txt

จากรูปที่ 3.3 และ รูปที่ 3.4 แสดงตัวอย่างเนื้อหาข่าวท่องเที่ยวจาก จากไฟล์ .html และ  
ไฟล์ .txt ซึ่งเป็นข้อความตัวอักษรธรรมดาและคัดลอกมาเฉพาะหัวข้อข่าวและเนื้อหาของข่าวซึ่งจะ  
เป็นข้อมูลเข้าเพื่อใช้ในการสร้างคำประสมหรือเพื่อสรุปใจความ

จากไฟล์ข้อความธรรมดาในรูปที่ 3.4 นำมาตัดคำโดยใช้โปรแกรมตัดคำจะได้ชุดคำดังรูปที่  
3.5 (ก) และ (ข) ตามลำดับ

[อีก, 'หนึ่ง', 'แหล่ง', 'ท่องเที่ยว', 'ทาง', 'ทะเล', 'ที่', 'อยู่', 'ไม่', 'ไกล', 'กรุงเทพฯ', ' ', 'และ', 'ยัง', 'มี', 'บรรยากาศ', 'สุดชิล', ' ', 'แถม', 'ยัง', 'สามารถ', 'กาง', 'เต็นท์', 'ริม', 'ทะเล', 'ชมบรรยากาศ', 'และ', 'ฟัง', 'เสียง', 'คลื่น', 'ได้', 'อย่าง', 'เต็มที่', 'ก็', 'คือ', ' ', ' ', ' ', 'อุทยานแห่งชาติ', 'เขาแหลมหญ้า', ' ', 'หญ้า', 'เกาะ', 'เสม็ด', ' ', ' ', 'อ.เมือง', ' ', 'จ.ระยอง', ' ', ' ', ' ', 'อุทยานแห่งชาติเขา ', 'แหลม ', 'หญ้า', ' ', ' ', 'หญ้า', 'เกาะ', 'เสม็ด', ' ', ' ', ' ', 'เป็น', 'อุทยาน', 'แห่งชาติ', 'ประเภท', 'ชาย', 'ฝั่ง', 'ผสม', 'หญ้า', 'เกาะ', 'ใน', 'ทะเล', ' ', ' ', 'ซึ่ง', 'ก็', 'หมาย', 'ถึง', 'ว่า', 'มี', 'พื้นที่', 'ครอบคลุม', 'ทั้ง', 'บน', 'แผ่นดิน', 'และ', 'ใน', 'ท้อง', 'ทะเล', ' ', 'บน', 'แผ่นดิน', 'ก็', 'คือ', 'บริเวณ', 'เขา', 'แหลม', 'หญ้า', 'และ', 'ชาย', 'ทะเล', 'ด้าน', 'ทิศ', 'ตะวันออก', 'ตก', 'ของ', 'เขา', 'แหลม', 'หญ้า', ' ', ' ', ' ', 'อยู่', 'ระหว่าง', 'หาด', 'แม่', 'รำพึง', 'และ', 'อ่าว', 'เพ', ' ', ' ', 'รวม', 'ไป', 'ถึง', 'เกาะ', 'ต่าง', 'ๆ', ' ', ' ', 'ใน', 'ท้อง', 'ทะเล', ' ', ' ', 'โดย', 'เกาะ', 'ที่', 'นัก', 'ท่องเที่ยว', ' ', 'รู้จัก', 'กัน', 'ดี', 'ก็', 'คือ', 'เกาะ', 'เสม็ด', ' ', ' ', 'ที่', 'ขึ้น', 'ชื่อ', 'ว่า', 'เป็น', 'เกาะ', 'ที่', 'สวยงาม', 'ที่สุด', 'ใน', 'ระยอง', ' ', ' ', 'และ', 'ยัง', 'มี', 'เกาะ', 'จันทร์', ' ', ' ', 'เกาะ', 'ทะเล', ' ', ' ', 'เกาะ', 'กุฎี', ' ', ' ', 'เกาะ', 'มะขาม', ' ', ' ', 'และ', 'เกาะปลาย', 'ดิน', ' ', ' ', 'อีก', 'ด้วย', 'ส่วน', 'ใหญ่', 'แล้ว', 'นัก', 'ท่องเที่ยว', 'นัก', 'จะ', 'เลือก', 'ไป', 'เที่ยว', 'ชม', 'น้ำ', 'ใส', 'ทราย', 'สวย', 'กัน', 'ที่', 'เกาะ', 'เสม็ด', 'เสีย', 'มาก', 'กว่า', ' ', ' ', 'จน', 'อาจ', 'หลงลืม',

(ก) ผลลัพธ์การตัดคำที่ได้จากโปรแกรมตัดคำ

[แหล่ง, 'ท่องเที่ยว', 'ทะเล', 'กรุงเทพฯ', 'บรรยากาศ', 'สุดชิล', 'แถม', 'กาง', 'เต็นท์', 'ริม', 'ชมบรรยากาศ', 'ฟัง', 'เสียง', 'คลื่น', 'เต็มที่', 'อุทยานแห่งชาติ', 'เขาแหลมหญ้า', 'หญ้า', 'เกาะ', 'เสม็ด', 'อ.เมือง', 'จ.ระยอง', 'อุทยานแห่งชาติเขา', 'แหลม', 'หญ้า', 'อุทยาน', 'แห่งชาติ', 'ประเภท', 'ชาย', 'ฝั่ง', 'ผสม', 'หมาย', 'พื้นที่', 'ครอบคลุม', 'แผ่นดิน', 'ท้อง', 'บริเวณ', 'แหลม', 'ทิศ', 'ตะวันออก', 'ตก', 'หาด', 'แม่', 'รำพึง', 'อ่าว', 'เพ', 'รู้จัก', 'ดี', 'ชื่อ', 'สวยงาม', 'ระยอง', 'จันทร์', 'ทะเล', 'เกาะ', 'กุฎี', 'เกาะมะขาม', 'เกาะปลาย', 'ดิน', 'เลือก', 'เที่ยว', 'ชม', 'น้ำ', 'ใส', 'ทราย', 'สวย', 'หลงลืม', 'แพ', 'ทัศนียภาพ', 'ตัว', 'เมือง', 'เกาะ', 'เสีย', 'เวลา', 'จับ', 'รถ', 'นาที', 'ค่า', 'ธรรมเนียม', 'คน', 'บาท', 'เด็ก', 'นอน', 'เล่น', 'เช่า', 'อุทยานฯ', 'สัมผัส', 'ลาน', 'บริการ', 'จุด', 'ครอง', 'ราชย์', 'ปี', 'ต้น', 'มะขาม', 'ท่า', 'เรือ', 'อาคาร', 'พรรษา', 'บ้าน', 'พัก', 'พักผ่อน', 'ครอบครัว', 'เพื่อนฝูง', 'นิยม', 'ถ่าย', 'รูป', 'ถนน', 'ยื่น', 'ปลาย', 'ศาลา', 'ไลฟ์', 'การ์ด', 'สี', 'ขาว', 'สะอาด', 'ตา', 'ตัด', 'ฟ้า', 'แลนด์', 'มาร์ก', 'เย็น', 'แดด', 'ร่ม', 'ลม', 'โรแมนติก', 'กิจกรรม', 'พลาด', 'หาย', 'เยือน', 'เดิน', 'เท้า', 'ศึกษา', 'เส้นทาง', 'ธรรมชาติ', 'ราว', 'กิโลเมตร', 'วง', 'กลม', 'วน', 'ขวา', 'ลัดเลาะ', 'ใจ', 'หิน', 'ซ้าย', 'ไม้', 'ใบ', 'หัว', 'ท่า', 'ป้าย', 'ข้อมูล', 'อ่าน', 'แกรนิต', 'สะพาน', 'สะดวก', 'กึ่ง', 'กลาง', 'โล่ง', 'เนิน', 'ทุ่ง', 'สลับ', 'เนิน', 'พัด', 'แรง', 'ทาน', 'กระแส', 'ภาพ', 'ประทับใจ', 'ทิวทัศน์', 'อากาศ', 'สดชื่น', 'ชวน', 'หยุด', 'หน้า', 'ร้อน', 'สนใจ', 'เขาแหลม', 'กัน', 'ถ.หาดแม่รำพึง', 'ตำบลเพ', 'อำเภอเมืองระยอง', 'จังหวัดระยอง', 'สอภาม', 'รายละเอียด', 'โทร]

(ข) ผลลัพธ์หลังจากกำจัดคำหยุดและคำซ้ำ

รูปที่ 3.5 ตัวอย่างชุดของคำหลังผ่านการตัดคำ ลบคำหยุด และคำซ้ำ

หลังจากได้ชุดคำภาษาไทยจากรูปที่ 3.5 แล้ว นำคำทั้งหมดมากำกับประเภทของคำเฉพาะที่เกี่ยวข้องกับกฎการประสมคำ ที่มีประกอบด้วยคำนาม คำกริยา และคำวิเศษณ์ประสมกัน ดังนั้นในงานวิจัยนี้จึงเลือกติดป้ายกำกับชนิดของคำเฉพาะคำ 3 ประเภทนี้เท่านั้นโดยนำคำที่ได้จากชุดคำทั้งหมดไปค้นหาในเว็บไซต์ออนไลน์ Longdo (<https://dict.longdo.com/search>) โดยใช้โปรแกรมที่เขียนขึ้นเพิ่มเติมเพื่อวนลูปค้นหาประเภทของคำจากเว็บไซต์ลองดู และเลือกชนิดของคำจากพจนานุกรมออนไลน์ของ Lexitron และพจนานุกรมฉบับราชบัณฑิตยสถาน พ.ศ. 2554 ตามที่กำกับไว้ในเว็บไซต์ ดังรูปที่ 3.6

แหล่ง

แปลศัพท์ | ดัดโดย NEW

บล็อก ป้อนคำใหม่ ทาปใหม่

Search result for แหล่ง (104 entries) (0.0163 seconds)

ลองค้นหาคำในรูปแบบอื่นๆ เพื่อให้ได้ผลลัพธ์มากขึ้นหรือน้อยลง: -แหล่ง-, \*แหล่ง\*

Thai-English: NECTEC's Lexitron-2 Dictionary [with local updates]

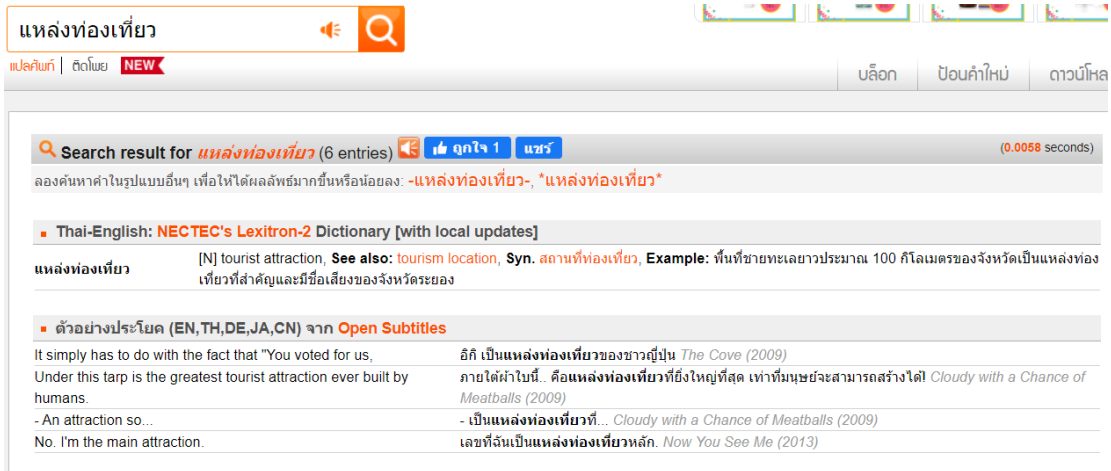
แหล่ง	[N] source, <b>Syn.</b> ถิ่น, ที่อยู่, ที่
แหล่ง	[N] centre, <b>See also:</b> location, spot, area, venue, position, place, <b>Syn.</b> ถิ่น, ที่อยู่, บริเวณ, ศูนย์รวม, แห่ง, ที่, <b>Example:</b> กลุ่มขบวนการค้ายาเสพติดได้หันมาผลิตกัญชาขายในรูปแบบใหม่โดยมีแหล่งอยู่ตามชายแดนป่ากันดารของประเทศกัมพูชา
แหล่งน้ำ	[N] water resource
แหล่งน้ำ	[N] water source, <b>See also:</b> source of water, <b>Example:</b> อุตสาหกรรมเยื่อและกระดาษจัดเป็นอุตสาหกรรมชนิดหนึ่งซึ่งสร้างมลภาวะต่อแหล่งน้ำได้รุนแรงที่สุด
แหล่งพัก	[N] shelter, <b>Syn.</b> ที่พัก
แหล่งข่าว	[N] informed source, <b>Syn.</b> ศูนย์ข่าว, <b>Example:</b> ข่าวนี้ได้จากแหล่งข่าวที่ไวใจได้, <b>Count unit:</b> แห่ง
แหล่งผลิต	[N] source of production, <b>Example:</b> ขบวนการค้ายาเสพติดได้ผลิตกัญชาขายในรูปแบบใหม่ โดยมีแหล่งผลิตอยู่ตามชายแดนป่ากันดารของประเทศกัมพูชา, <b>Thai definition:</b> ต้นต่อที่ผลิต
แหล่งผลิต	[N] source of production, <b>Example:</b> ขบวนการค้ายาเสพติดได้ผลิตกัญชาขายในรูปแบบใหม่ โดยมีแหล่งผลิตอยู่ตามชายแดนป่ากันดารของประเทศกัมพูชา, <b>Thai definition:</b> ต้นต่อที่ผลิต
แหล่งหล้า	[N] earth, <b>See also:</b> land, <b>Syn.</b> พื้นแผ่นดิน, พื้นโลก, โลก
แหล่งที่มา	[N] source, <b>See also:</b> origin, root, derivation, base, <b>Syn.</b> ที่มา, มูลเหตุ, <b>Example:</b> ตำรวจกำลังสาวถึงแหล่งที่มาของสินบน 5 ล้านบาท

ไทย-ไทย: พจนานุกรมฉบับราชบัณฑิตยสถาน พ.ศ. ๒๕๕๔

แหล่ง	(แหล่ง) น. ถิ่น, ที่อยู่, บริเวณ, ศูนย์รวม, บ่อเกิด, แห่ง, ที่
แหล่งหล้า	น. พื้นแผ่นดิน
แหล่งเสื่อมโทรม	น. บริเวณที่คนอาศัยอยู่อย่างแออัด ประกอบด้วยบ้านเรือนที่ทรุดโทรมไม่ถูกสุขลักษณะ
ชม	แหล่งที่เกิดที่เก็บ เช่น ชมทรัพย์
สมาคม	แหล่งหรือที่ประชุมของบุคคลหลายคนมารวมกันด้วยมีจุดประสงค์เพื่อประโยชน์บางประการ เช่น เข้าสมาคม สมาคมศิษย์เก่า
สำนัก	แหล่งศึกษาอบรม เช่น สำนักวิปัสสนาวัดป่าก่า สำนักที่สาปายโมกข์
อุ	แหล่งที่เกิด เช่น อุข้าว อุน้ำ

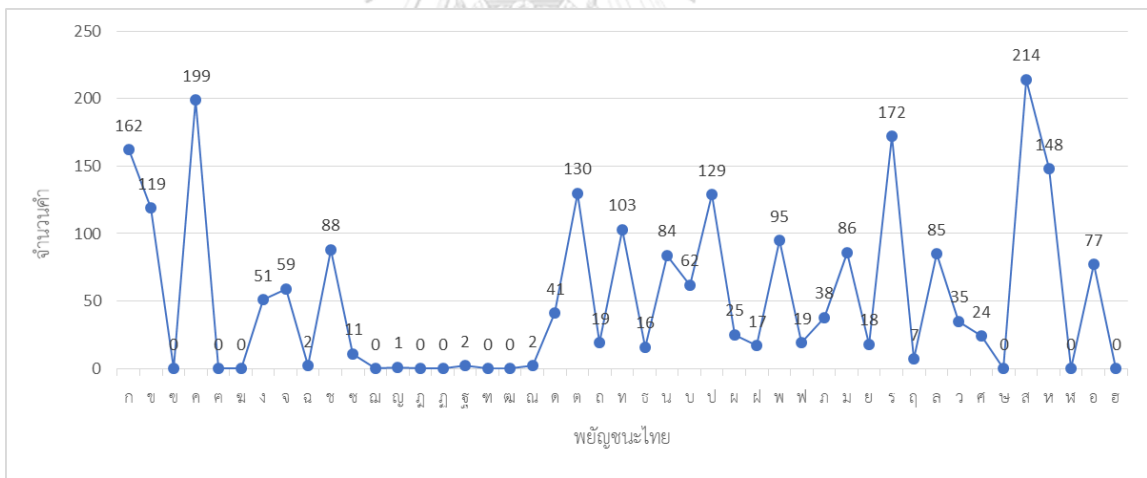
รูปที่ 3.6 ตัวอย่างการค้นหาชนิดของคำในพจนานุกรมออนไลน์จากเว็บไซต์ Longdo

หลังจากได้ชนิดของคำแล้ว ขั้นตอนต่อไปคือขั้นตอนการประสมคำ คำประสมเกิดจากคำมูลประสมกันทำให้เกิดคำที่มีความหมายใหม่ ดังนั้นในงานวิจัยนี้ผู้วิจัยเลือกประสมคำจากคำทั้ง 3 ชนิด ได้แก่ คำนาม คำกริยา และคำวิเศษณ์ และนำคำใหม่ที่เกิดจากการประสมไปค้นหาในพจนานุกรมออนไลน์อีกครั้ง ดังรูปที่ 3.7 เลือกคำใหม่ที่เป็นคำนามและมีความหมายจัดเก็บเป็นรายการคำประสม



รูปที่ 3.7 ตัวอย่างคำประสมที่ได้จากการค้นหาในพจนานุกรมออนไลน์จากเว็บไซต์ Longdo

เมื่อเสร็จสิ้นขั้นตอนี้แล้วจะได้ชุดรายการคำประสมทั้งหมด 2,340 คำจากทั้งหมด 5,000 คำ เช่น กองทัพเรือ กระดาษแก้ว ข้าวดี เป็นต้น โดยสามารถดูรายการคำประสมได้จาก ภาคผนวก ก. ในงานวิจัยนี้ผู้วิจัยได้เปรียบเทียบการสรุปใจความสำคัญระหว่างกรณีที่มีคำประสมและไม่มีคำประสม ซึ่งจะแสดงรายละเอียดในขั้นตอนต่อไป

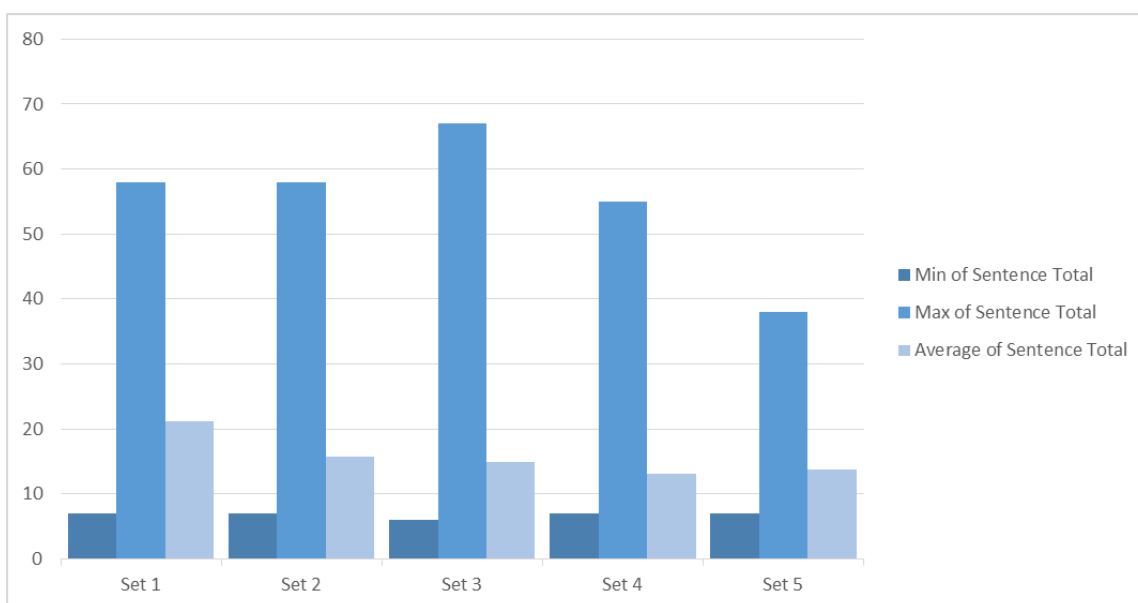


รูปที่ 3.8 จำนวนรายการคำประสมแยกตามพยัญชนะไทย

จากรูปที่ 3.8 ผลลัพธ์ที่ได้จากการประสมคำแยกตามพยัญชนะไทย หมวดตัวอักษร ส มีจำนวนคำมากที่สุดคือ 214 คำ และมีจำนวน 10 หมวดตัวอักษรที่ไม่พบคำประสมจากข่าวการท่องเที่ยวไทยในงานวิจัยนี้เลย ได้แก่ ข ฅ ฉ ฎ ฏ ฐ ฒ ษ ฝ และ ฮ

### 3.3 ขั้นตอนการเตรียมข้อมูลเพื่อสรุปใจความสำคัญ

ขั้นตอนการเตรียมข้อมูลเพื่อสรุปใจความ เป็นขั้นตอนการเตรียมคำในประโยคเพื่อใช้ในการหาหน้าหนึ่งของประโยค โดยมีวิธีเหมือนกับขั้นตอนการสร้างรายการคำประสม โดยเริ่มจาก รวบรวมข่าวท่องเที่ยวทั้งหมด 400 ข่าว โดยแบ่งเป็น 5 ชุด ชุดละ 80 ข่าว โดยมีความยาวของแต่ละข่าวไม่เท่ากันดังรูปที่ 3.9



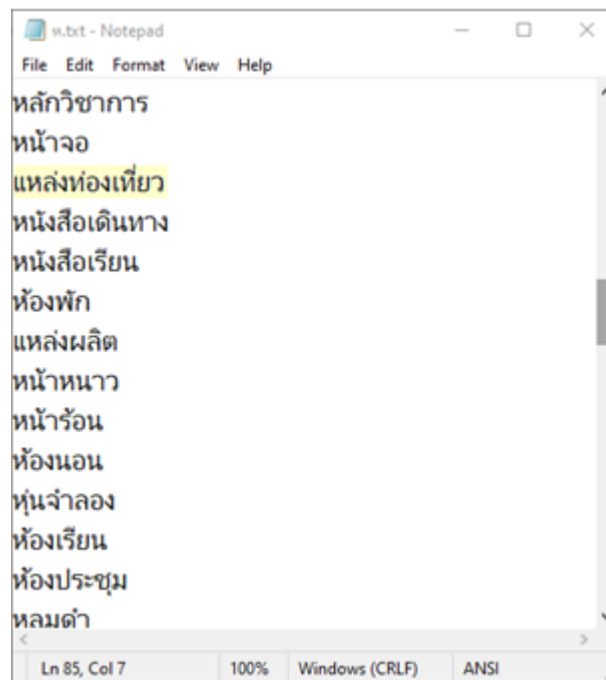
รูปที่ 3.9 จำนวนประโยคจากข่าว 5 ชุด

หลังจากนั้นแต่ละข่าวจะถูกตัดคำด้วยโปรแกรมตัดคำ ตามด้วยขั้นตอนการลบคำหยุดและคำซ้ำออก ซึ่งผลลัพธ์ดังตัวอย่างรูปที่ 3.5 ข้างต้น ผลลัพธ์ที่ได้เป็นผลลัพธ์จากโปรแกรมตัดคำเพียงอย่างเดียว ดังนั้นในขั้นตอนนี้จะมีการตัดคำประสมร่วมด้วย โดยนำคำจากข่าว คำปัจจุบัน + คำก่อนหน้า 1 ตำแหน่ง ไปค้นหาในรายการคำประสมที่เตรียมไว้ ถ้ามีในรายการคำประสมให้ตัดคำโดยรวมสองคำเหลือเพียง 1 คำ เช่น

ชุดคำจากโปรแกรมตัดคำ ['แหล่ง', 'ท่องเที่ยว', 'กรุงเทพฯ', 'บรรยากาศ', 'แถม', 'ริม', 'ชมบรรยากาศ', 'เต็มที', 'อุทยานแห่งชาติ', 'เขาแหลมหญ้า', 'หมู', 'เกาะ']

คำปัจจุบัน คือ 'ท่องเที่ยว' คำก่อนหน้า 1 ตำแหน่ง คือ 'แหล่ง' นำคำทั้งสองคือ 'แหล่ง' + 'ท่องเที่ยว' ไปค้นหาในรายการคำประสมที่เตรียมไว้ดังรูปที่ 3.10 ถ้าพบคำทั้งสองในรายการคำประสมให้รวมเป็นคำเดียว ดังรูปที่ 3.11



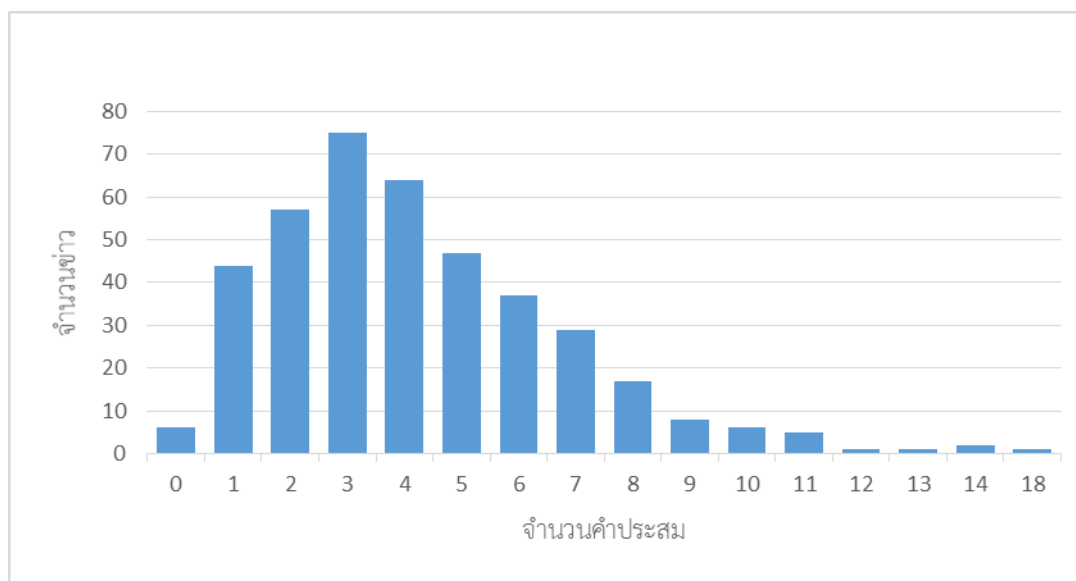


รูปที่ 3.10 ตัวอย่างรายการคำประสมหมวดหมู่ ห

[แหล่งท่องเที่ยว, 'กรุงเทพฯ', 'บรรยากาศ', 'แถม', 'ริม', 'ชมบรรยากาศ', 'เต็มที่', 'อุทยานแห่งชาติ', 'เขาแหลม  
 หญ้า', 'หมู่เกาะ', 'เสียด', 'อ.เมือง', 'จ.ระยอง', 'ประเภท', 'ชาย', 'ฝั่ง', 'ผสม', 'ทมาย', 'พื้นที่', 'ครอบคลุม',  
 'แผ่นดิน', 'ห้อง', 'บริเวณ', 'ทิศ', 'ตะวันตก', 'หาด', 'แม่', 'รำพึง', 'อ่าว', 'เพ', 'รู้จัก', 'ดี', 'ชื่อ', 'สวยงาม', 'จันทร์',  
 'ทะเล', 'เกาะภูเก็ต', 'เกาะมะขาม', 'เกาะปลาย', 'ดิน', 'เลือก', 'ชม', 'น้ำ', 'ใส', 'ทราย', 'สวย', 'หลงลืม', 'แพ',  
 'ทัศนียภาพ', 'ตัวเมือง', 'เลาะ', 'เลียบ', 'เวลา', 'จับ', 'รถ', 'นาทึ', 'ค่าธรรมเนียม', 'คน', 'บาท', 'เด็ก', 'เล่น', 'เช่า',  
 'อุทยานฯ', 'สัมผัส', 'ลาน', 'บริการ', 'จุด', 'ครอง', 'ราชย์', 'ปี', 'ต้นมะขาม', 'ท่าเรือ', 'อาคาร', 'พรรษา', 'บ้านพัก',  
 'พักผ่อน', 'ครอบครัว', 'เพื่อนฝูง', 'นิยม', 'ถ่าย', 'รูป', 'ถนน', 'ยื่น', 'ปลาย', 'ศาลา', 'ไลฟ์', 'การ์ด', 'สี', 'ขาว',  
 'สะอาด', 'ตา', 'ตัด', 'ฟ้า', 'แลนด์', 'มาร์ก', 'เย็น', 'แดด', 'ร่ม', 'สม', 'โรแมนติก', 'กิจกรรม', 'พลาดหาก', 'เยือน',  
 'เดิน', 'เท้า', 'ศึกษา', 'เส้นทาง', 'ธรรมชาติ', 'ราว', 'กิโลเมตร', 'วง', 'กลม', 'วน', 'ขวา', 'ลัดเลาะ', 'เข็ดหิน', 'ซ้าย',  
 'ไม้ใบ', 'หัว', 'ท่า', 'ป้าย', 'ข้อมูล', 'อ่าน', 'แกรนิต', 'สะพาน', 'สะดวก', 'กึ่งกลาง', 'โล่ง', 'เนิน', 'ทุ่ง', 'สลั', 'เนื่อง',  
 'พัด', 'แรง', 'ทาน', 'กระแสน', 'ภาพ', 'ประทับใจ', 'ทิวทัศน์', 'อากาศริม', 'สดชื่น', 'ชวน', 'หยุด', 'หน้าร้อน', 'สนใจ',  
 'เขาแหลม', 'กัน', 'ถ.หาดแม่รำพึง', 'ตำบลเพ', 'อำเภอเมืองระยอง', 'จังหวัดระยอง', 'สอภาม', 'รายละเอียด',  
 'โทร']

รูปที่ 3.11 ตัวอย่างชุดคำที่ผ่านโปรแกรมคัดคำและการตัดคำโดยใช้คำประสม

จากรูปที่ 3.11 ตัวอย่างชุดคำที่ผ่านโปรแกรมคัดคำและตัดคำประสมร่วมด้วย หลังจากขั้น  
 ตอนนี้จะได้คำของประโยคที่พร้อมใช้ในการคำนวณหาคำน่าหนักของประโยคในขั้นต่อไป



รูปที่ 3.12 จำนวนค่าประสมในข้อมูลชุดทดลอง

จากรูปที่ 3.12 แสดงจำนวนข่าวที่พบค่าประสมตั้งแต่ 0-18 ค่า จากทั้งหมด 400 ข่าวที่ใช้เป็นข้อมูลชุดทดลอง ซึ่งสามารถแจกแจงรายละเอียดได้ดังนี้

-	จำนวนค่าประสม	0	ค่า	มีทั้งหมด	6	ข่าว
-	จำนวนค่าประสม	1	ค่า	มีทั้งหมด	44	ข่าว
-	จำนวนค่าประสม	2	ค่า	มีทั้งหมด	57	ข่าว
-	จำนวนค่าประสม	3	ค่า	มีทั้งหมด	75	ข่าว
-	จำนวนค่าประสม	4	ค่า	มีทั้งหมด	64	ข่าว
-	จำนวนค่าประสม	5	ค่า	มีทั้งหมด	47	ข่าว
-	จำนวนค่าประสม	6	ค่า	มีทั้งหมด	37	ข่าว
-	จำนวนค่าประสม	7	ค่า	มีทั้งหมด	29	ข่าว
-	จำนวนค่าประสม	8	ค่า	มีทั้งหมด	17	ข่าว
-	จำนวนค่าประสม	9	ค่า	มีทั้งหมด	8	ข่าว
-	จำนวนค่าประสม	10	ค่า	มีทั้งหมด	6	ข่าว
-	จำนวนค่าประสม	11	ค่า	มีทั้งหมด	5	ข่าว
-	จำนวนค่าประสม	12	ค่า	มีทั้งหมด	1	ข่าว
-	จำนวนค่าประสม	13	ค่า	มีทั้งหมด	1	ข่าว
-	จำนวนค่าประสม	14	ค่า	มีทั้งหมด	2	ข่าว
-	จำนวนค่าประสม	18	ค่า	มีทั้งหมด	1	ข่าว

### 3.4 การหาหน้าหนักของคำและการเลือกประโยค

การหาหน้าหนักของประโยคเป็นการประยุกต์ใช้อัลกอริทึมการแยกค่าแบบเดี่ยวและวิธีการหาความถี่ของคำเพื่อหาความสัมพันธ์ของคำกับประโยคว่ามีความสัมพันธ์มากน้อยเพียงใด ยิ่งค่ามากหมายถึงคำนั้นมีความสัมพันธ์กับประโยคมาก และเลือกประโยคเหล่านั้นมาเป็นประโยคหลักในการสรุปใจความสำคัญ

หลังจากที่ได้ชุดของคำในแต่ละข่าวเรียบร้อยแล้ว ขั้นตอนต่อไปเป็นการแบ่งประโยคหลักเกณฑ์ของการแบ่งประโยคภาษาไทยในปัจจุบันไม่มีหลักเกณฑ์ที่ตายตัวทำให้ยากต่อการแบ่งประโยค ดังนั้นในงานวิจัยนี้จึงอาศัยเกณฑ์การแบ่งประโยคจากผู้เชี่ยวชาญด้านภาษาไทย 2 ท่านจากคณะศิลปศาสตร์ หลักสูตรภาษาไทยและการสื่อสาร มหาวิทยาลัยอุบลราชธานี โดยมีการแบ่งประโยคออกเป็นสองส่วน ส่วนแรกคือ ส่วนของหัวข้อข่าว ส่วนที่สองเป็นส่วนเนื้อหาของเนื้อหาข่าว ดังรูปที่ 3.13 โดยกำหนดให้  $S_x$  เป็นรหัสกำกับประโยค โดย  $x$  คือลำดับของประโยค ตัวอย่างเช่น  $S_1$  หมายถึงประโยคที่ 1 และ  $S_2$  หมายถึงประโยคที่ 2 เป็นต้น

Topic: "เขาแหลมหญ้า" เทียวทะเลระยองสุดชิล ทางเดินต้นองฟังเสียงคลื่น

$S_1$  = "อีกหนึ่งแหล่งท่องเที่ยวทางทะเลที่อยู่ไม่ไกลกรุงเทพฯ และยังมีบรรยากาศสุดชิล แคมป์สามารถทางเดินที่ริมทะเลชมบรรยากาศและฟังเสียงคลื่นได้อย่างเต็มที่ก็คือ "อุทยานแห่งชาติเขาแหลมหญ้า-หมู่เกาะเสม็ด" อ.เมือง จ.ระยอง"

$S_2$  = "อุทยานแห่งชาติเขาแหลมหญ้า-หมู่เกาะเสม็ด" เป็นอุทยานแห่งชาติประเภทชายฝั่งผสมหมู่เกาะในทะเล"

$S_3$  = "ซึ่งก็หมายถึงว่ามีพื้นที่ครอบคลุมทั้งบนแผ่นดินและในท้องทะเล บนแผ่นดินก็คือบริเวณเขาแหลมหญ้าและชายทะเลด้านทิศตะวันตกของเขาแหลมหญ้า (อยู่ระหว่างหาดแม่รำพึงและอ่าวเพ) รวมไปถึงเกาะต่างๆ ในท้องทะเล"

รูปที่ 3.13 ตัวอย่างการแบ่งประโยค

จากขั้นตอนนี้ ข้อมูลที่เตรียมไว้ ทั้งชุดคำ และชุดประโยค ได้ถูกจัดเตรียมไว้เรียบร้อยแล้ว ขั้นตอนต่อไปเป็นการคำนวณหาหน้าหนักของประโยคและการเลือกประโยคมาใช้ในการสรุปใจความสำคัญ งานวิจัยนี้ได้นำวิธีการหาค่าความสัมพันธ์ของคำและประโยคมาประยุกต์ใช้ 2 วิธีดังนี้

### 3.4.1 การหาความสำคัญของคำและประโยคด้วยการหาความถี่ของคำ และการหาค่าความสัมพันธ์ของคำกับประโยคโดยใช้อัลกอริทึมการแยกค่าแบบเดี่ยว (SVD) และจัดกลุ่มประโยคด้วยเคมีน

#### 1) การหาความสำคัญของคำและประโยคด้วยการหาความถี่ของคำและการหาค่าความสัมพันธ์ของคำกับประโยคโดยใช้อัลกอริทึมการแยกค่าแบบเดี่ยว (SVD)

ขั้นตอนแรกเริ่มจากการสร้างเมตริกซ์  $A$  ที่มีขนาด  $m \times n$  ดังตารางที่ 3.1 โดยที่  $m$  คือแถวแทนจำนวนคำ (Terms) และ  $n$  คือสตมภ์แทนประโยคในข่าว 1 ข่าว ซึ่งแต่ละคำในแถวนั้นจะไม่ซ้ำกัน แต่ละช่องแสดงจำนวนครั้งที่คำนั้นๆปรากฏในแต่ละประโยค

จากตารางที่ 3.1 ได้จำนวนครั้งของคำที่เกิดขึ้นในแต่ละประโยค ซึ่งความสั้นยาวของประโยคมีผลต่อจำนวนครั้งของคำที่เกิดขึ้นเนื่องจากประโยคที่ยาวกว่ามีโอกาสในการพบความซ้ำของคำมากกว่า ดังนั้นจึงต้องนำเอาวิธีคำนวณหาค่าความถี่ของคำเข้ามาช่วยเพื่อหาความสำคัญของคำ จากขั้นตอนนี้สามารถประยุกต์ใช้สูตรของการคำนวณหาค่าความถี่ของคำจากสมการ (2.1) จะได้ Term - Sentence Matrix ดังตัวอย่างตารางที่ 3.2

หลังจากได้ Term - Sentence Matrix จากตารางที่ 3.2 แล้วสามารถนำเมตริกซ์ดังกล่าวมาทำการหาค่าความสัมพันธ์โดยใช้อัลกอริทึมการแยกค่าแบบเดี่ยวจากสมการ (2.2) เพื่อนำเวกเตอร์ประโยคที่ได้จาก Matrix  $V$  ดังตัวอย่างตารางที่ 3.3 ไปจัดกลุ่มประโยคแบบเคมีน

ตัวอย่างการคำนวณตารางที่ 3.1 และ ตารางที่ 3.2

$S1 =$  "อีกหนึ่งแหล่งท่องเที่ยวทางทะเลที่อยู่นอกเมืองกรุงเทพฯ และยังมีบรรยากาศสุดชิล แลมนยังสามารถกางเต็นท์ริมทะเลชมบรรยากาศและฟังเสียงคลื่นได้อย่างเต็มที่ก็คือ 'อุทยานแห่งชาติเขาแหลมหญ้า-หมู่เกาะเสม็ด' อ.เมือง จ.ระยอง"

ตัดคำด้วยคัตคำและลบคำหยุดออกจากประโยคได้ดังนี้

$S1 =$  ['หนึ่ง', 'แหล่ง', 'ท่องเที่ยว', 'ทะเล', 'อยู่', 'นอก', 'กรุงเทพฯ', 'บรรยากาศ', 'สุดชิล', 'แลมน', 'ยัง', 'สามารถ', 'กาง', 'เต็นท์', 'ริม', 'ทะเล', 'ชมบรรยากาศ', 'ฟัง', 'เสียง', 'คลื่น', 'อย่าง', 'เต็มที่', 'คือ', 'อุทยาน', 'แห่ง', 'ชาติ', 'เขาแหลม', 'หญ้า', 'หมู่', 'เกาะ', 'เสม็ด', 'อ.เมือง', 'จ.ระยอง']

จากผลของการตัดคำด้วยคัตคำของประโยคที่ 1 ได้จำนวนคำในประโยคหลังจากตัดคำหยุดออกทั้งสิ้น 33 คำ เมื่อนำมาคำนวณหาค่าความถี่ของคำจากสมการ (2.1) จะได้

$$tf(\text{แหล่ง}, S1) = \frac{1}{33} = 0.0303$$

$$tf(\text{ทะเล}, S1) = \frac{2}{33} = 0.0606$$

## 2) การจัดกลุ่มประโยคด้วยเคมิน

หลังจากนั้นใช้ Matrix  $V$  มาใช้ในขั้นตอนการจัดกลุ่มประโยคด้วยเคมิน เพื่อหาประโยคที่สำคัญของแต่ละกลุ่มโดยมีขั้นตอนดังนี้

1. สุ่มข้อมูลขึ้นมา  $K$  ตัวเท่ากับจำนวนกลุ่มที่ต้องการแบ่งเพื่อใช้เป็นตัวแทนในการแบ่งกลุ่ม ในงานวิจัยนี้ค่า  $K$  คือ  $K$  จำนวนกลุ่มประโยคที่ถูกเลือกมาเป็นประโยคที่ใช้สรุปใจความสำคัญ กล่าวคือจำนวนประโยคจากอัตราการย่อยร้อยละ 20 30 และ 40 ของประโยคตั้งต้น เช่น 1 ข่าวมีประโยคทั้งหมดจำนวน 100 ประโยค ผลลัพธ์จากการสรุปใจความสำคัญคือ อัตราการย่อยร้อยละ 20 จะได้ประโยคสรุปทั้งหมด 80 ประโยค อัตราการย่อยร้อยละ 30 จะได้ประโยคสรุปทั้งหมด 70 และ อัตราการย่อยร้อยละ 40 จะได้ประโยคสรุปทั้งหมด 60 ประโยค เป็นต้น
2. สุ่มเลือกประโยคมาเป็นประโยคตั้งต้นของแต่ละกลุ่ม (Centroid) โดยใช้ค่าจากเวกเตอร์ประโยคของ Matrix  $V$
3. แบ่งกลุ่มโดยวัดค่าความคล้ายคลึงกันแบบโคไซน์ระหว่างประโยคที่ต้องการจัดกลุ่มกับประโยคที่โดนสุ่มมาเป็นประโยคตัวแทนและจัดกลุ่มให้แต่ละประโยคอยู่กลุ่มที่มีค่าความคล้ายคลึงกันมากที่สุด
4. เฉลี่ยประโยคแต่ละกลุ่มเพื่อหาตัวแทนกลุ่มใหม่  
จากนั้นวนทำซ้ำเรื่อยๆในขั้นตอนที่ 2, 3 และ 4 จนกว่าประโยคสมาชิกของแต่ละกลุ่มไม่มีการเปลี่ยนแปลง
5. เลือกตัวแทนประโยคของแต่ละกลุ่มโดยเลือกประโยคที่มีค่าความสำคัญมากที่สุดเป็นตัวแทนกลุ่มละ 1 ประโยค
6. เรียงลำดับประโยคผลลัพธ์จากลำดับของประโยคดั้งเดิม

ตารางที่ 3.1 ตัวอย่างจำนวนครั้งของค่าที่เกิดขึ้นในแต่ละประโยค

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16	S17	S18	S19	S20	S21	S22
เซาแหลม	1	0	2	0	0	1	0	1	0	1	0	0	2	2	0	0	0	4	1	0	1	1
หญ้า	1	1	2	0	0	1	1	1	0	1	0	0	2	3	0	0	0	4	1	0	1	1
เที่ยว	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ทะเล	2	1	3	0	0	0	0	0	0	1	0	0	2	2	0	1	0	0	1	1	1	0
ระยอง	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
สุดขีด	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
กาง	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0
เด่นที่	1	0	0	0	0	0	0	0	0	2	1	0	0	0	0	0	0	0	0	0	0	0
นอน	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
ฟัง	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
เสียง	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
คลื่น	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
แหล่ง	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
ท่องเที่ยว	1	0	0	1	0	1	0	0	1	0	0	0	1	0	0	0	0	0	1	1	1	0
กรุงเทพฯ	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
บรรยากาศ	1	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0
แดม	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ริม	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0

ตารางที่ 3.2 ตัวอย่าง Term - Sentence Matrix จากการคำนวณความถี่ของคำ

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16	S17	S18	S19	S20	S21	S22
เขาแหลม	0.0303	0	0.0606	0	0	0.0385	0	0.0588	0	0.0476	0	0	0.0392	0.037	0	0	0	0.08	0.0476	0	0.0588	0.0556
พญา	0.0303	0.0625	0.0606	0	0	0.0385	0.0833	0.0588	0	0.0476	0	0	0.0392	0.0556	0	0	0	0.08	0.0476	0	0.0588	0.0556
เที่ยว	0	0	0	0	0	0.0385	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ทะเล	0.0606	0.0625	0.0909	0	0	0	0	0	0	0.0476	0	0	0.0392	0.037	0	0.1111	0	0	0.0476	0.0833	0.0588	0
ระยอง	0	0	0	0.0667	0	0	0	0.0588	0	0	0	0	0	0	0	0	0	0	0	0	0	0
สุดขีด	0.0303	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
กก	0.0303	0	0	0	0	0	0	0	0	0.0476	0.0455	0	0	0	0	0	0	0	0	0	0	0
เด่นท์	0.0303	0	0	0	0	0	0	0	0	0.0952	0.0455	0	0	0	0	0	0	0	0	0	0	0
นอน	0	0	0	0	0	0	0	0	0	0.0476	0	0	0	0	0	0	0	0	0	0	0	0
ฟัง	0.0303	0	0	0	0	0	0	0	0	0.0476	0	0	0	0	0	0	0	0	0	0	0	0
เสียง	0.0303	0	0	0	0	0	0	0	0	0.0476	0	0	0	0	0	0	0	0	0	0	0	0
คลื่น	0.0303	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
แหล่ง	0.0303	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0588	0
ท่องเที่ยว	0.0303	0	0	0.0667	0	0.0385	0	0	0.0714	0	0	0	0.0196	0	0	0	0	0	0.0476	0.0833	0.0588	0
กรุงเทพฯ	0.0303	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
บรรยากาศ	0.0303	0	0	0	0	0	0	0	0	0.0476	0	0	0.0196	0	0	0	0	0	0	0	0	0
แถบ	0.0303	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ริม	0.0303	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0588	0

ตารางที่ 3.3 ตัวอย่าง Matrix V

	Con1	Con2	Con3	Con4	Con5	Con6	Con7	Con8	Con9	Con10	Con11	Con12	Con13	Con14	Con15	Con16	Con17	Con18	Con19	Con20	Con21	Con22
S1	-0.152	-0.186	-0.309	-0.052	-0.004	-0.114	-0.122	-0.218	-0.008	-0.226	-0.201	-0.026	-0.161	-0.273	-0.137	-0.435	-0.261	-0.257	-0.332	-0.221	-0.205	-0.18
S2	0.0782	0.1202	0.2062	0.0276	0.0003	0.1052	0.1369	0.2222	0.0171	0.12	5E-05	-0.049	0.086	0.0101	-0.55	-0.134	-0.556	0.1754	0.049	-0.319	0.1459	0.2155
S3	-0.067	-0.127	-0.133	-0.124	0.006	-0.141	-0.113	-0.129	-0.016	-0.067	0.3743	0.0567	-0.097	0.1031	-0.406	0.46	0.2393	0.0048	0.0752	-0.487	-0.127	-0.19
S4	0.032	-0.052	-0.028	-0.014	-0.012	-0.044	-0.066	-0.157	-0.029	0.0449	0.7438	-0.02	0.0065	-0.119	0.295	-0.034	-0.495	0.1248	0.0004	0.1672	-0.05	-0.119
S5	-0.1	-0.081	-0.012	-0.06	-5E-04	0.0387	0.1122	0.2437	0.054	0.0119	0.1208	0.0558	0.0029	-0.004	0.5026	-0.308	0.184	0.1853	0.0412	-0.647	-0.017	0.2213
S6	0.2985	0.0814	-0.012	-0.094	-0.053	0.023	0.0799	-0.378	-0.233	0.6592	0.0721	0.2337	0.0456	-0.038	-0.105	-0.299	0.2309	-0.057	-0.113	-0.074	0.0972	-0.075
S7	-0.01	-9E-04	0.0055	0.0116	-0.997	-0.003	-0.013	0.0361	0.0267	-0.025	-0.006	-0.005	-0.003	-0.048	-0.004	0.0318	0.0046	-0.003	0.0098	0.0005	-0.005	0.0126
S8	0.0147	-0.137	-0.093	0.1323	0.0237	0.0052	-0.101	0.3214	0.7081	0.2129	0.2241	0.2175	-0.092	-0.052	-0.209	-0.17	0.182	-0.051	-0.166	0.1969	-0.092	0.0513
S9	-0.185	-0.044	-0.097	0.4326	-0.004	0.2467	0.0373	0.0736	-0.481	-0.236	0.2736	0.1721	-0.09	0.0134	-0.246	-0.235	0.2399	0.0882	-0.027	0.1746	-0.119	0.2711
S10	0.0307	0.3232	-0.126	0.378	-0.007	0.3356	0.2293	-0.509	0.4134	-0.2	0.0134	-0.081	0.0646	0.0448	0.0132	-0.068	0.0355	0.0478	0.2064	-0.146	0.0417	-0.119
S11	-0.242	-0.145	0.2254	-0.503	-0.022	-0.037	0.527	-0.161	0.1067	-0.257	0.1065	0.0776	0.0406	0.1579	-0.146	-0.203	0.1278	0.2005	-0.136	0.1968	0.0868	-0.075
S12	0.0634	0.4107	0.3431	-0.112	0.016	-0.207	-0.045	-0.216	0.0861	-0.178	0.1207	0.167	-0.217	-0.153	0.0373	0.0977	0.0237	-0.329	-0.169	-0.047	-0.11	0.5396
S13	-0.105	-0.107	-0.096	-0.228	-0.002	-0.086	-0.574	-0.286	0.1011	-0.085	-0.045	0.0354	0.3294	0.0178	-0.106	-0.212	0.0575	0.2489	0.3326	0.0822	0.1285	0.3366
S14	0.0803	-0.121	-0.005	-0.018	-0.006	0.0018	0.0262	-0.084	0.0278	0.1466	0.1346	-0.879	-0.125	0.0609	-0.08	-0.083	0.1885	-0.012	-0.134	0.0395	-0.074	0.2553
S15	-0.013	-0.639	0.0255	0.1097	0.001	0.2631	0.1407	-0.307	0.0739	0.1402	-0.162	0.1733	-0.123	-0.013	0.0983	0.331	-0.186	0.0083	-0.093	-0.024	0.025	0.368
S16	-0.021	-0.03	-0.118	-0.008	0.0014	-0.245	0.3347	-8E-04	0.0002	0.1623	-0.072	0.0217	0.244	-0.02	-0.034	-0.025	-0.076	-0.089	0.4182	0.0753	-0.712	0.1354



### 3.4.2 การหาความสำคัญของคำด้วยการหาความถี่ของคำและเลือกประโยคจากหัวข้อข่าว

#### 1) การหาความสำคัญของคำและประโยคด้วยการหาความถี่ของคำ

ขั้นตอนแรกเริ่มจากการสร้างเมตริกซ์ A ที่มีขนาด  $m \times n$  กำหนดให้แถวคือ คำที่ได้จากหัวข้อข่าวกับสดมภ์คือประโยคในข่าว จะได้ Term - Sentence Matrix คือตารางแสดงจำนวนครั้งของคำจากหัวข้อข่าวที่เกิดขึ้นในแต่ละประโยค หลังจากนั้นคำนวณหาความถี่ของคำที่เกิดขึ้นจากหัวข้อข่าวในแต่ละประโยค จากสมการที่ (2.1) ดังตารางที่ 3.4

ตัวอย่างการคำนวณตารางที่ 3.4

S1 = "อีกหนึ่งแหล่งท่องเที่ยวทางทะเลที่อยู่ไม่ไกลกรุงเทพฯ และยังมีบรรยากาศสุดชิล แลมนยังสามารถกางเต็นท์ริมทะเลชมบรรยากาศและฟังเสียงคลื่นได้อย่างเต็มที่ก็คือ 'อุทยานแห่งชาติเขาแหลมหญ้า-หมู่เกาะเสม็ด' อ.เมือง จ.ระยอง"

ตัดคำด้วยคัตคำและลบคำหยุดออกจากประโยคได้ดังนี้

S1 = ['หนึ่ง', 'แหล่ง', 'ท่องเที่ยว', 'ทะเล', 'อยู่', 'ไกล', 'กรุงเทพฯ', 'บรรยากาศ', 'สุดชิล', 'แลม', 'ยัง', 'สามารถ', 'กาง', 'เต็นท์', 'ริม', 'ทะเล', 'ชมบรรยากาศ', 'ฟัง', 'เสียง', 'คลื่น', 'อย่าง', 'เต็มที่', 'คือ', 'อุทยาน', 'แห่ง', 'ชาติ', 'เขาแหลม', 'หญ้า', 'หมู่', 'เกาะ', 'เสม็ด', 'อ.เมือง', 'จ.ระยอง']

หัวข้อข่าว คือ "เขาแหลมหญ้า" เที่ยวทะเลระยองสุดชิล กางเต็นท์นอนฟังเสียงคลื่น

ตัดคำและลบคำหยุดได้ดังนี้

['เขาแหลม', 'หญ้า', 'เที่ยว', 'ทะเล', 'ระยอง', 'สุดชิล', 'กาง', 'เต็นท์', 'นอน', 'ฟัง', 'เสียง', 'คลื่น']

จากผลของการตัดคำด้วยคัตคำของประโยคที่ 1 ได้จำนวนคำในประโยคหลังจากตัดคำหยุดออกทั้งสิ้น 33 คำ เมื่อนำมาคำนวณหาความถี่ของคำจากสมการ (2.1) จะได้

$$tf(\text{เขาแหลม}, S1) = \frac{1}{33} = 0.0303$$

#### 2) การเลือกประโยคจากหัวข้อข่าว

หลังจากได้ตารางที่ 3.4 แล้ว ขั้นตอนต่อไปคือการเลือกประโยคที่จะใช้ในการสรุปใจความสำคัญ ในงานวิจัยนี้ผู้วิจัยทำการเปรียบเทียบการเลือกประโยคจาก 2 วิธีดังต่อไปนี้

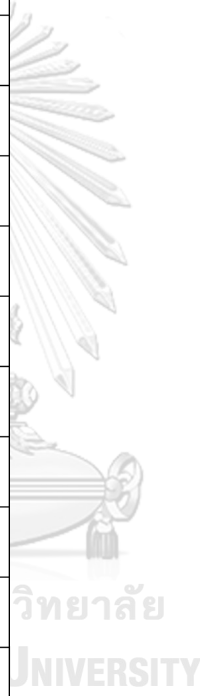
##### 2.1 เลือกประโยคจากความคล้ายคลึงกันของประโยคหลัก

1. เลือกประโยคหลักจากผลรวมของความถี่ของคำในหัวข้อข่าวที่เกิดขึ้นในแต่ละประโยค กล่าวได้ว่าผลรวมค่าความถี่ของคำสูงหมายถึงมีคำสำคัญที่เกิดขึ้นในประโยคนั้นเป็นจำนวนมาก ยิ่งค่ามากยิ่งหมายถึงประโยคนั้นอธิบายเนื้อหาของข่าวได้ดี

2. นำประโยคตั้งต้นไปหาค่าความคล้ายคลึงแบบโคไซน์กับประโยคอื่นๆในข่าว โดยใช้เวกเตอร์ประโยคตั้งต้นกับเวกเตอร์ประโยคลำดับถัดไปจนครบจำนวนประโยคทั้งหมดในข่าว
  3. เลือกประโยคที่มีค่าความคล้ายคลึงกันกับประโยคหลักมากที่สุดมาใช้ในการเป็นประโยคหลักในครั้งถัดไป  
จากนั้นวนทำซ้ำเรื่อยๆในขั้นตอนที่ 2-3 จนกว่าครบทุกประโยคจะได้ชุดของประโยคมา 1 ชุดโดยเรียงลำดับตามความคล้ายคลึงกันของประโยคก่อนหน้า
  4. เลือกจำนวนประโยคจากอัตราการบีบอัดที่ระบุไว้ คือร้อยละ 20 30 และ 40 ตามลำดับ
  5. เรียงลำดับประโยคที่ได้ตามลำดับของประโยคดั้งเดิม
- 2.2 เลือกประโยคจากการเรียงลำดับผลรวมของความถี่ของคำจากหัวข้อข่าว
1. เรียงลำดับประโยคจากผลรวมในแต่ละคอลัมน์ กล่าวคือประโยคที่มีความสัมพันธ์กับคำในหัวข้อข่าวมากที่สุดคือประโยคที่อธิบายเนื้อหาของข่าวได้ดีที่สุด โดยวิธีการนี้จะเลือกจำนวนประโยคจากอัตราการบีบอัดที่กำหนดไว้ในงานวิจัยนี้คือ ร้อยละ 20 30 และ 40 ตามลำดับ
  2. หลังจากได้ชุดของประโยคแล้ว นำมาจัดเรียงลำดับของประโยคตามลำดับประโยคตั้งต้น สุดท้ายก็จะได้ชุดประโยคที่เป็นบทสรุปใจความสำคัญของข่าวแล้วนำมาเปรียบเทียบกับบทสรุปใจความสำคัญของมนุษย์

ตารางที่ 3.4 ตัวอย่าง Term-Sentence Matrix จากความถี่ของคำจากหัวข้อเกี่ยวกับประโยค

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16	S17	S18	S19	S20	S21	S22
เขานกลม	0.0303	0	0.0606	0	0	0.0385	0	0.0588	0	0.0476	0	0	0.0392	0.037	0	0	0	0.08	0.0476	0	0.0588	0.0556
หญ้า	0.0303	0.0625	0.0606	0	0	0.0385	0.0833	0.0588	0	0.0476	0	0	0.0392	0.0556	0	0	0	0.08	0.0476	0	0.0588	0.0556
เขียว	0	0	0	0	0	0.0385	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ทะเล	0.0606	0.0625	0.0909	0	0	0	0	0	0	0.0476	0	0	0.0392	0.037	0	0.1111	0	0	0.0476	0.0833	0.0588	0
ระยอง	0	0	0	0.0667	0	0	0	0.0588	0	0	0	0	0	0	0	0	0	0	0	0	0	0
สุดขีด	0.0303	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
กาง	0.0303	0	0	0	0	0	0	0	0	0.0476	0.0455	0	0	0	0	0	0	0	0	0	0	0
เดินที่	0.0303	0	0	0	0	0	0	0	0	0.0952	0.0455	0	0	0	0	0	0	0	0	0	0	0
นอน	0	0	0	0	0	0	0	0	0	0.0476	0	0	0	0	0	0	0	0	0	0	0	0
ฟัง	0.0303	0	0	0	0	0	0	0	0	0.0476	0	0	0	0	0	0	0	0	0	0	0	0
เสียง	0.0303	0	0	0	0	0	0	0	0	0.0476	0	0	0	0	0	0	0	0	0	0	0	0
คลื่น	0.0303	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<b>RRรวม</b>	<b>0.303</b>	<b>0.125</b>	<b>0.2121</b>	<b>0.0667</b>	<b>0</b>	<b>0.1154</b>	<b>0.0833</b>	<b>0.1765</b>	<b>0</b>	<b>0.4286</b>	<b>0.0909</b>	<b>0</b>	<b>0.1176</b>	<b>0.1296</b>	<b>0</b>	<b>0.1111</b>	<b>0</b>	<b>0.16</b>	<b>0.1429</b>	<b>0.0833</b>	<b>0.1765</b>	<b>0.1111</b>



## บทที่ 4

### การทดลองและผลการทดลอง

วิทยานิพนธ์ฉบับนี้ได้แบ่งการทดลองออกเป็น 2 ช่วง คือ 1.ขั้นตอนการสร้างรายการคำประสมจากข่าวท่องเที่ยวภาษาไทย 2. การสรุปใจความสำคัญของข่าวท่องเที่ยวภาษาไทย และเปรียบเทียบผลการสรุปใจความสำคัญของข่าวท่องเที่ยวไทยกับวิธีที่มีอยู่ก่อนหน้า

#### 4.1 สภาพแวดล้อมและเครื่องมือที่ใช้ในการทดลอง

สภาพแวดล้อมและเครื่องมือที่ใช้ในการทดลอง สามารถสรุปได้ดังนี้

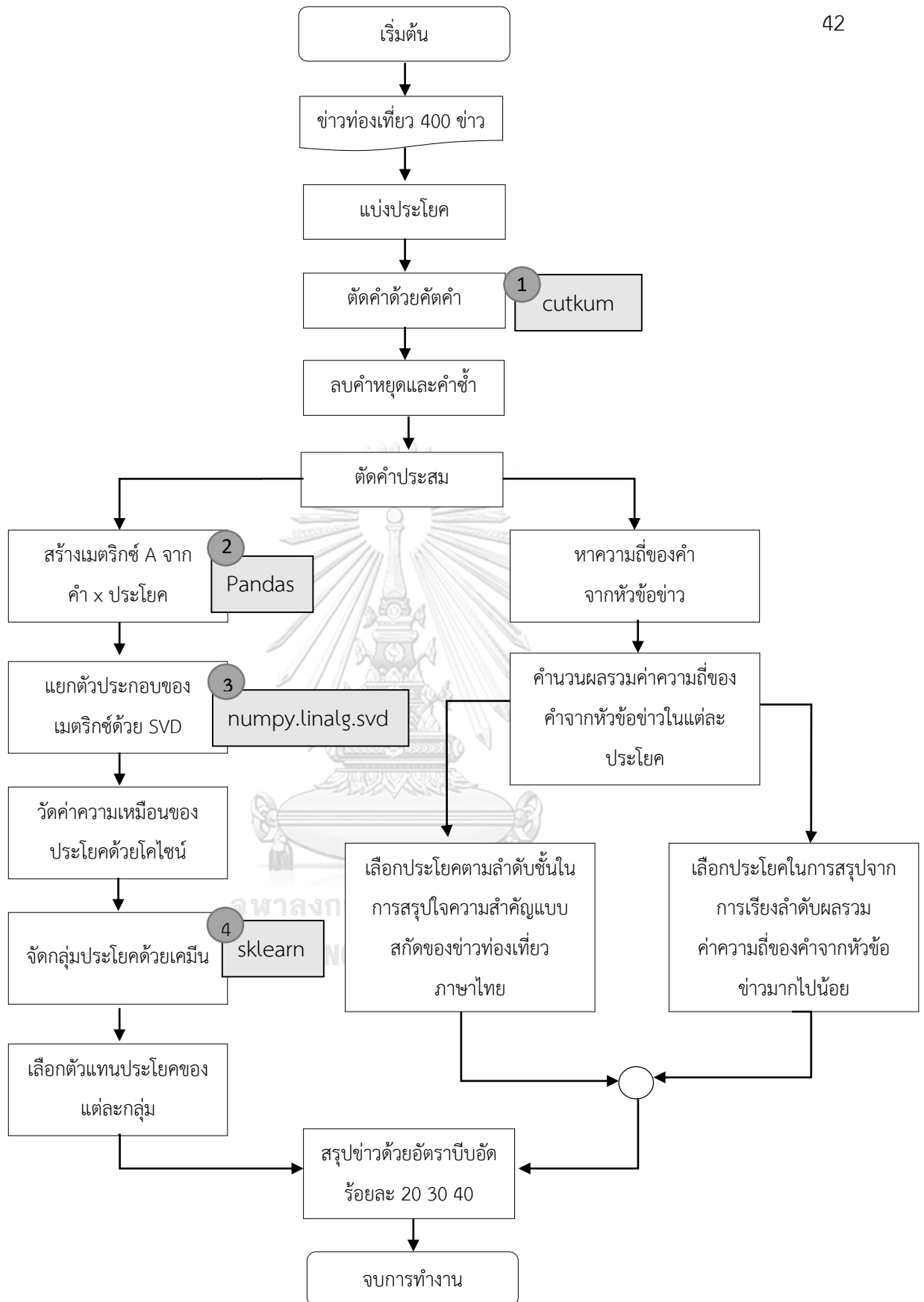
##### 4.1.1 คอมพิวเตอร์ที่ใช้ในการทดลอง

การทดลองต่อไปนี้นำบนเครื่องคอมพิวเตอร์ที่มีหน่วยประมวลผลกลาง Intel Core i7 ความเร็ว 1.9 GHz หน่วยความจำขนาด 16 GB โดยใช้ระบบปฏิบัติการ Windows 10

##### 4.1.2 การเขียนโปรแกรม

การเขียนโปรแกรมเพื่อใช้ในการทดลองนี้ถูกพัฒนาโดยใช้ภาษา python 3 และได้ใช้ไอออนเพนเซอร์สในขั้นตอนต่างๆ บางส่วนดังต่อไปนี้ และดังแสดงในรูปที่ 4.1

- 1) ตัดคำภาษาไทยโดย Cutkum
- 2) สร้าง Matrix โดย Pandas (McKinney & Computing, 2011)
- 3) แยกตัวประกอบเมตริกซ์การแยกค่าแบบเดี่ยว (SVD) โดย numpy.linalg.svd (Oliphant, 2006)
- 4) จัดกลุ่มประโยคเคมินโดย sklearn (Bahmani, Moseley, Vattani, Kumar, & Vas-silvitskii, 2012)



รูปที่ 4.1 การใช้ไอออนเพนซอร์สในขั้นตอนต่างๆ

#### 4.1.3 ข้อมูลที่ใช้ในการทดลอง

ข้อมูลข่าวภาษาไทยที่ใช้ในการทดลองแบ่งออกเป็นข่าวที่ใช้สำหรับทำรายการคำประสมทั้งหมด 5,000 ข่าว และข่าวที่ใช้ในการสรุปใจความสำคัญทั้งหมด 400 ข่าว แบ่งเป็น 5 ชุด ชุดละ 80 ข่าว มาจาก <https://www.manager.co.th/travel> และ <https://news.kapook.com/travel> (รายการข่าวอยู่ในภาคผนวก ข) ข่าวที่นำมาใช้ทดลองทั้งหมดสามารถแจกแจงได้ดังตารางที่ 4.1

ตารางที่ 4.1 รายละเอียดชุดข้อมูลข่าวที่ใช้ในการทดลอง

ชุดข้อมูลข่าว	จำนวนประโยคเนื้อหา	จำนวนประโยคเนื้อหา	จำนวนประโยคเนื้อหา
	ต่ำสุด (ประโยค)	เฉลี่ย (ประโยค)	สูงสุด (ประโยค)
ชุดที่ 1	7	21	58
ชุดที่ 2	7	16	58
ชุดที่ 3	6	15	67
ชุดที่ 4	7	13	55
ชุดที่ 5	7	13	38

#### 4.2 การทดลองตัดคำประสม

ในงานวิจัยนี้เพื่อเพิ่มประสิทธิภาพของการสรุปใจความสำคัญของข่าวมากขึ้น ผู้วิจัยได้ทดลองตัดคำประสมร่วมด้วย โดยมีสมมติฐานหลักคือการสรุปใจความสำคัญภาษาไทยนั้นจะแม่นยำมากขึ้น ส่วนหนึ่งมาจากการพิจารณาตัดคำประสมร่วมด้วย เพื่อเป็นพิสูจน์สมมติฐานว่าถูกต้องหรือไม่ ผู้วิจัยจึงทำการทดลองการสรุปใจความสำคัญโดยใช้การตัดคำอย่างเดี่ยวเปรียบเทียบกับการสรุปใจความสำคัญโดยใช้การตัดคำโดยตัดคำประสมร่วมด้วย รูปที่ 4.2 แสดงตัวอย่างข่าวต้นฉบับที่มีจำนวนประโยคทั้งสิ้น 51 ประโยค แต่ละประโยคกำกับด้วยหมายเลขประโยค  $S_x$  โดย  $S$  แทนประโยค และ  $x$  คือลำดับของประโยคในข่าวนั้นๆ

<p>Topic: แม้วชายท่องเที่ยวควบรำลึกสึนามิ</p> <p>S1 = รำลึกหนึ่งปีสึนามิ น้ำตาท่วมอันดามัน ชาวต่างชาติขอร้องอย่ารบกวน เข้อยีสึนามิยังร้องระงมไม่ได้รับการเหลียวแลจากภาครัฐ</p> <p>S2 = กลุ่มเรือประมงขนาดกลางยื่นหนังสือร้องเรียนถูกลอยแพ ชาวเกาะพิพิธวงคามันสัญญารัฐบาลหนึ่งปีฟื้นฟูทุกอย่างเสร็จ แต่จนบัดนี้ฝั่งเมืองก็ยังไม่สรุบไม่ได้</p> <p>S3 = นายกฯ ยอมรับความช่วยเหลืออาจมีปัญหาตกหล่น แต่ไม่วยาคยไม่ทำดีที่สุดแล้ว</p> <p>S4 = เผยหวังให้บรรยากาศการท่องเที่ยวฟื้นคืน มท.1 เตือนเคลียร์ปัญหาที่ดิงพังงา ชิดเส้น ม.ค.ปีหน้าเรียบร้อย</p> <p>S5 = งานรำลึกครบรอบ 1 ปีสึนามิ ตลอดวานนี้ (26 ธ.ค.) 4 จังหวัดชายฝั่งอันดามัน ยกเว้นระนอง และสตูล รัฐบาลได้จัดพิธีรำลึกและไว้อาลัยผู้เสียชีวิตจากเหตุการณ์สึนามิ 7 จุด มีผู้สูญเสียญาติมิตรทั้งชาวไทยและต่างชาติ เข้าร่วมหลายพันคน</p> <p>S6 = บรรยากาศโดยรวมเต็มไปด้วยความเศร้าสลด ก่อนพิธีการรำลึกไว้อาลัยผู้เสียชีวิตจากภัยพิบัติสึนามิ จะเริ่มขึ้นอย่างเป็นทางการในเวลา 08.00 น.</p> <p>S7 = นาย Bill Paterson เอกอัครราชทูตออสเตรเลีย ประจำประเทศไทย พร้อมด้วยครอบครัวผู้ประสบภัยสึนามิที่เสียชีวิตไปจำนวน 23 คน ได้ร่วมประกอบพิธีรำลึก 1 ปี สึนามิ ที่บริเวณชายหาดป่าตอง หน้าโรงแรมภูเก็ตแกรนด์รีสอร์ทแอนด์สปา</p> <p>S8 = ต่อมา เวลาประมาณ 09.00น. พ.ต.ท.ทักษิณ ชินวัตร นายกรัฐมนตรี พร้อมคณะ ได้เดินทางมาพบปะสนทนาตัวแทนญาติของผู้ประสบภัยที่บริเวณสวนสาธารณะโลมา และวางช่อดอกไม้สด ที่บริเวณศูนย์รวมใจ ใช้เวลาประมาณ 15 นาทีก่อนที่จะเดินทางไปที่จังหวัดพังงา</p> <p>S9 = นายกรัฐมนตรีกล่าวกับตัวแทนสื่อมวลชนต่างประเทศ ว่า อยากจะให้ผู้เสียชีวิตทั้งคนไทยและต่างประเทศได้มีชีวิตใหม่ในภพหน้า ในส่วนของการช่วยเหลือผู้ประสบภัยในพื้นที่เสี่ยงภัย รัฐบาลโดยหน่วยงานต่างๆ และท้องถิ่นพยายามดำเนินการช่วยเหลืออย่างทัดเทียมกันหรือต่อเนื่อง และให้ครอบคลุมเท่าที่ควรจะเป็น</p> <p>S10 = ในขณะที่องค์กรภาคเอกชน มูลนิธิต่างๆ เลือกที่จะให้การช่วยเหลือในชุมชน หรือหมู่บ้านตามความเหมาะสมเช่นเดียวกัน และสิ่งที่หวังหลังจากวันครบรอบการเกิดเหตุสึนามิ คือ อยากจะให้บรรยากาศการท่องเที่ยวได้กลับฟื้นคืนมาดังเดิม</p> <p>S11 = สำหรับบรรยากาศที่ บริเวณหาดป่าตอง จ.ภูเก็ต ซึ่งมีผู้เข้าร่วมงานกว่า 2,500 คน</p> <p>S12 = ในขณะที่ผู้ร่วมรำลึกเหตุการณ์กำลังวางพวงดอกไม้ได้มีหญิงชาวต่างประเทศรายหนึ่งขึ้นไปบนเวทีและใช้ไมโครโฟนพูดแทนญาติผู้สูญหายและเสียชีวิต ว่า วันนี้เป็นวันครบรอบ 1 ปี สึนามิ ทุกคนล้วนมาเพื่อรำลึกนึกถึงผู้จากไปด้วยความเศร้าสละเทือนใจ แต่กลับถูกสื่อสารมวลชนแขนงต่างๆ มารอทำข่าวและถ่ายภาพเป็นจำนวนมาก อยากบอกว่าทุกคนไม่อยากมีภาพปรากฏบนหน้าหนังสือพิมพ์หรือสื่ออื่นๆ พร้อมกับขอร้องอย่าติดตามทำข่าวและถ่ายภาพอย่างใกล้ชิด</p> <p>S13 = นายกรัฐมนตรี ได้เดินทางไปทำพิธีวางศิลาฤกษ์ อนุสรณ์สถานสึนามิ ที่บริเวณหาดเล็ก อุทยานแห่งชาติเขาหลัก-ลำรู่</p> <p>S14 = โดยมีรัฐมนตรี และญาติของผู้เสียชีวิตจากเหตุการณ์สึนามิ รวมทั้งแขกที่รัฐบาลเชิญมาร่วมพิธีกว่า 500 คน</p>
---

(ก.) ตัวอย่างข่าวการท่องเที่ยวฉบับเต็มแสดงประโยคที่ 1-14

S15 = โนงานดังก่อสร้าง มีการบรรเลงวงดุริยางค์ โดยวง Dr.Sax Chamber Orchestra จากมหาวิทยาลัยมหิดล ซึ่งขณะการประกวดระดับนานาชาติและการเดี่ยวแซ็กโซโฟนโดยศิลปินระดับโลก ชาวสวีเดน Mr.Johan Stengard และการอ่านบทกวีเป็นภาษาอังกฤษโดย มาริน่า อิมบราโนวิช ศิลปินระดับโลกชาวอเมริกัน ผู้ที่อยู่ในเหตุการณ์และประสบอุบัติเหตุในวันที่ 26 ธ.ค.2547 และสุดท้าย มีการฉายแบบการประกวดรูปแบบการก่อสร้างอนุสรณ์สถานสันามิทั้ง 5 แบบ ที่ได้ผ่านการคัดเลือกจากประเทศสเปน ฟินแลนด์ จีน สหรัฐอเมริกา และออสเตรเลีย

S16 = ต่อจากนั้น มีการประกอบพิธีไว้อาลัยผู้เสียชีวิตทั้งหมดที่หาดช่องฟ้า บางเนียง อ.ตะกั่วป่า จ.พังงา ซึ่งทุลกระหม่อมหญิงอุบลรัตนราชกัญญา สิริวัฒนาพรรณวดี เสด็จเป็นประธาน

S17 = ส่วนช่วงค่ำ ที่จังหวัดภูเก็ต ได้จัดพิธีจุดเทียน 100,000 เล่ม ที่บริเวณริมหาดป่าตองระยะทาง 3 กม. โดยองค์การบริหารส่วนจังหวัดภูเก็ต ( อบจ.ภูเก็ต ) ร่วมกับการท่องเที่ยวแห่งประเทศไทย ( ททท. ) และกลุ่มศิลปินภูเก็ต จัดงาน “Light Up Phuket” ที่บริเวณสวนสาธารณะโลมา

S18 = สำหรับบรรยากาศรำลึก 1 ปีสันามิภาคประชาชน ที่ จ.พังงา ชาวบ้านบ้านน้ำเค็ม ประมาณ 200 คน ร่วมทำพิธีทำบุญตักบาตร อุทิศส่วนกุศล ให้กับผู้ที่เสียชีวิตจากพิบัติภัยสันามิในตอนเช้า โดยมีพิธีกรรมทางศาสนาพุทธ และศาสนาคริสต์ บรรยากาศเป็นไปด้วยความโศกเศร้า

S19 = นางบัวผัน ภุมสนธิ อายุ 55 ปี ชาวบ้านบ้านน้ำเค็ม กล่าวระหว่างนำภาพของ “แสงเดือน ชูเฉลิม” ลูกสาวที่เสียชีวิตมาวางบริเวณเต็นท์พิธี ว่า ตนสูญเสียลูกสาวในเช้าวันเกิดเหตุ หลังจากที่ลูกสาวผู้เป็นคนหารายได้หลักในครอบครัวเพิ่งกลับมาจากกรุงเทพฯ

S20 = ได้แต่เอ่ยหักแม่ไม่กี่คำก็โดนคลื่นยักษ์ซัดไปพร้อมกับคนอื่น ๆ

S21 = “คิดถึงลูกสาวและหลานชายที่เสียชีวิตอยู่ทุกวัน ยังทำใจยังไม่ได้ พอลูกคนนี้เสียไปก็ไม่มีใครส่งเงินมาเลี้ยงชีพ การช่วยเหลือของภาครัฐถือว่าไม่เพียงพอ ถึงได้ก็ไม่พอ เราก็วิ่งเต้นไม่เป็นอย่างคนอื่นอย่างเขา” นางบัวผัน กล่าว

S22 = ส่วน นายวิเชษฐ หอมชื่น ชาวบ้านบ้านน้ำเค็มอีกคน ระบุว่า รู้สึกเสียใจกับสูญเสียหลาน และเพื่อนบ้าน วันนี้อยู่มาทำบุญ แต่ตนมองว่าต้องแปรวิกฤตให้เป็นโอกาส คนที่อยู่ก็ต้องสู้ต่อไป

S23 = จากนั้น เมื่อเวลาประมาณ 16.00 น. หลังจากนายกรัฐมนตรี เดินทางออกจากบ้านน้ำเค็มและงานรำลึกของภาครัฐเสร็จสิ้นลง ชาวบ้านบ้านน้ำเค็มจำนวน 300 คน ได้จัดงานรำลึกเหตุการณ์ 1 ปี สันามิ ฉบับชาวบ้านขึ้นที่บริเวณหน้าอนุสาวรีย์สันามิบ้านน้ำเค็ม โดยใช้พื้นที่ลาดปูนบ้านที่ได้รับความเสียหายจากคลื่นยักษ์สันามิ เป็นเวทีการแสดง โดยมีนักแสดงเยาวชนหลากหลายรุ่นมาร่วมเต้นประกอบเพลงหลายสิบ ทั้งมีพิธีทำบุญโดยการโยนไม้จันทน์เข้ากองไฟ และปล่อยโคมไฟขึ้นสู่ท้องฟ้าในที่สุด

S24 = “เป็นข่าวลือมาจากพวกข้าราชการ ว่า ไม่อยากให้ชาวบ้านทำเวทีขึ้นมา ทางชาวบ้านก็เลยต้องอาศัยพื้นบ้านเก่า นอกจากนี้เขายังไม่อยากให้เราจัด จริงๆ อยากจะให้ชาวบ้านเอากันไปที่บางเนียง (สถานที่จัดงานหลักของภาครัฐ จ.พังงา) แต่เราก็ไม่ไป เลยจัดตอนเช้าก่อนงานกับตอนเย็นหลังงานของทางการ” ชาวบ้านบ้านน้ำเค็มคนหนึ่ง ระบุ

(ข.) ตัวอย่างข่าวการท่องเที่ยวฉบับเต็มแสดงประโยคที่ 15-24



- S25 = ชาวบ้านคนเดิม ยังบอกด้วยว่า งานที่ทางการต้องการจัดชาวบ้านรับรู้ก่อนหน้าประมาณ 1 สัปดาห์ ในขณะที่ ก่อนหน้าที่จะมีการจัดงานไม่กี่วัน ทางอำเภอได้สั่งให้ผู้ใหญ่บ้านรวบรวมรายชื่อแนบบัตรประจำตัวประชาชนผู้ที่เป็ญาติผู้เสียชีวิต และผู้ได้รับบาดเจ็บจากสึนามิ โดยว่ากันว่าจะออกบัตรร่วมงานสำหรับแขกที่รัฐบาลเชิญ และสามารถนำไปรับการช่วยเหลือได้ 2 พันบาท
- S26 = “แต่จริงๆ ก็ไม่มีการตรวจอะไรเลย เข้าได้หมด เป็นเหมือนเป็นแรงจูงใจให้ชาวบ้านไปร่วมงาน บางรายมาจากต่างจังหวัดกันเป็นคันรด ส่วนคนบ้านน้ำเค็มก็ทะเลาะกันเกือบตาย เพื่อแย่งบัตรนี้ เพราะไม่ได้ทุกคน” ชาวบ้านบ้านน้ำเค็ม บอกเล่า
- S27 = ส่วน นางนิลนันท์ พกกระย้า วัย 35 ปี ชาวบ้านบ้านน้ำเค็ม กล่าวเสริมว่า งานดังกล่าวเป็นงานที่ชาวบ้านจัดขึ้นมาเพื่อระลึกถึงชาวบ้านบ้านน้ำเค็มที่เสียชีวิต ด้วยเห็นว่าเมื่อพวกเขาตายที่นี่ ก็ต้องจัดงานไว้อาลัยที่นี้ ส่วนทางการมาจัดก็มาร่วมจำนวนไม่น้อยหลายรายได้เห็นหน้านายกาฯ บางคนก็ไม่เห็น
- S28 = “เราเป็นคนบ้านน้ำเค็ม คนที่ตายบ้างก็เกิดที่นี้ แต่ตายที่นี้ เราไม่อยากไปร่วมงานที่อื่นหรอก แต่ถ้านายกาฯ มา เราก็มารับ”
- S29 = ด้าน นายวิรัตน์ นวลทอง ประธานกลุ่มตัวอย่างประมงขนาดกลางบ้านน้ำเค็ม กล่าวว่า ถ้าชาวบ้านไม่จัดในปีนี้ ก็คงยากที่จะจัดในปีต่อไป
- S30 = นายวิรัตน์ กล่าวด้วยว่า วันนี้ตัวแทนของกลุ่มได้มีโอกาสยื่นหนังสือให้กับทีมงานของนายกรัฐมนตรี ที่มาเปิดอนุสาวรีย์ ในประเด็นการขออุทธรณ์เงินช่วยเหลือสำหรับเรือประมงขนาดกลางที่มีอยู่ 88 ลำ ซึ่งก่อนหน้านี ทางกระทรวงเกษตรฯ ได้รับปากว่าจะเพิ่มการช่วยเหลือจาก 10% เป็น 60% แต่ก็ไม่มีการดำเนินการใดๆ ทั้งสิ้น รวมทั้งความช่วยเหลือแพปลาที่รับเฉพาะเรือขนาดกลาง ก็ยังไม่ได้รับการช่วยเหลือ
- S31 = “เขาก็รับปาก แต่ก็รับปากแบบขอไปที ซึ่งเราก็เข้าใจ หลังจากนั้นต้องขยันติดตาม” ตัวแทนกลุ่มเรือประมงขนาดกลาง กล่าว
- S32 = สำหรับบรรยากาศที่บ้านบางเนียง บริเวณที่เรือตรวจการณ์ชายฝั่ง ต.813 ขึ้นมาเกยตื้นเมื่อครั้งเหตุการณ์สึนามิ บรรยากาศเป็นไปด้วยความโศกเศร้า มีญาติผู้เสียชีวิตนับพันคนมาเข้าร่วมพิธีรำลึก
- S33 = นางสุดสงวน ลิมคูลย์ วัย 73 ปี ผู้เสียญาติไปในเหตุการณ์ กล่าวว่ ตั้งใจมาเพื่อร่วมรำลึกเพราะสูญเสียญาติพี่น้องไป 4 คน และขณะนี้อยู่กับพี่สาวซึ่งพิการขา และบ้านที่พังเพ็ียงบางส่วนยังคงไม่ได้รับการซ่อมแซม
- S34 = “เงินช่วยเหลือไม่ได้ชักแดงหนึ่ง เพราะว่าไปอยู่ต่างอำเภอ ไม่ได้กลับไปอยู่บ้านเดิม กลัวคลื่น เขาบอกให้ติดตาม แต่ทั้ง อบต. อำเภอ และจังหวัดก็โยนให้ไปหาที่นั่นที่นี้ สุดท้ายไม่ได้ ตอนนีก็ไม่สนใจแล้ว” นางสุดสงวน กล่าว
- S35 = สำหรับกระดานแสดงความรู้สึกได้รับความสนใจจากผู้ร่วมงานจำนวนมาก โดยข้อความส่วนใหญ่จะเป็นการแสดงความรู้สึกต่อการสูญเสียหลากหลายภาษา เช่น “อาลัยยัง น่องแป้ครีบ แมร์รักลูกตลอด ลูกกลับมาเกิด มาเป็นลูกของแม่ใหม่ะครับ”, “ขอให้ดวงวิญญาณทุกคนไปสู่สุคติ” ฯลฯ
- S36 = ที่ เกาะพีพี จ.กระบี่ นายสุรเกียรติ์ เสถียรไทย รองนายกรัฐมนตรี เปิดเผยหลังเป็นประธานพิธีรำลึกที่เกาะพีพี ว่า เสียใจกับเหตุการณ์ที่เกิดขึ้น ขณะนี้ผ่านมา 1 ปีแล้ว รัฐบาลและเอกชน รวมถึงชาวบ้านในพื้นที่ประสบภัยได้ร่วมมือกันฟื้นฟูจนกลับมาสู่สภาพดั้งเดิม

(ค.) ตัวอย่างข่าวการท่องเที่ยวฉบับเต็มแสดงประโยคที่ 25-36

- S37 = อย่างไรก็ตาม ชาวบ้านบนเกาะพีพี ได้ส่งบทกลอนต่อ นายสุรเกียรติ์ เรียกเรื่องผ่านถึงนายกรัฐมนตรี ให้เร่งฟื้นฟูตามที่เคยรับปากไว้ว่า ภายใน 1 ปี เกาะพีพีจะฟื้นทุกอย่าง แต่ขณะนี้ฝั่งเมืองยังไม่สามารถหาข้อสรุปได้ชัดเจน
- S38 = นายสุเทพ เทือกสุบรรณ เลขาธิการพรรคประชาธิปัตย์ กล่าวถึงกรณีที่รัฐบาลระบุว่าได้ช่วยเหลือประชาชนที่ได้รับผลกระทบจากเหตุการณ์คลื่นยักษ์สึนามิครบถ้วนแล้ว เหลือเพียงที่พัก โดยแสดงความกังวลว่า นายกฯ อาจได้ข้อมูลไม่ครบถ้วน เพราะจากการติดตามตรวจสอบยังพบว่าประชาชนยังเดือดร้อน มีปัญหาที่ทำกิน พื้นฟูอาชีพ โดยเฉพาะผู้ประกอบการรายเล็กที่ความช่วยเหลือยังไม่ถึง
- S39 = รวมถึงเด็กกำพร้าและผู้สูงอายุที่ลูกหลานเสียชีวิต รัฐบาลอย่ามัวสร้างภาพโฆษณาชวนเชื่อ ขอให้ลงมือของจริง
- S40 = พ.ต.ท.ทักษิณ ชินวัตร นายกรัฐมนตรี ให้สัมภาษณ์อีกครั้งถึงการช่วยผู้ประสบภัยที่มีการร้องเรียนว่าช่วยเหลือไม่ทั่วถึงว่า คนที่เข้าไปแจ้งว่าได้รับผลกระทบส่วนใหญ่ ทางจังหวัดโดยผู้ว่าราชการจังหวัดจะเป็นคนตรวจสอบ และจะมีชื่อส่งมาที่สำนักปลัดสำนักนายกรัฐมนตรี ซึ่งตนเชื่อว่าจะให้ความช่วยเหลือครบทุกคน ยกเว้นเรื่องบ้านที่ยังสร้างไม่เสร็จ แต่เขาก็รู้แล้วว่า เขาจะได้บ้านที่ไหน
- S41 = แต่มีภาคเอกชนบางรายที่ไปสร้างบ้านเกินความต้องการ และบังเอิญบริเวณนั้นยังสร้างไม่เต็ม ยังมีเหลืออยู่ อย่างบ้านไอทีวี ก็ยังมีเหลือนิดหน่อย จึงได้บอกกับผู้ว่าฯ ให้ช่วยอนุโลม เพราะนี่ถือเป็นส่วนหนึ่งของการแก้ปัญหาความยากจนที่เป็นนโยบายของรัฐบาล ซึ่งตนคิดว่า น่าจะครบถ้วน
- S42 = ส่วนที่ไม่ครบถ้วนจริงๆ อาจจะไม่ใช่การเสียหายโดยตรง เช่น เรือประมง เรามีหลักเกณฑ์ช่วยเหลืออยู่ลำละ 5 หมื่นบาท แต่ปรากฏว่า เรือสำล้นราคาแพง ประชาชนก็คาดหวังว่าจะได้รับความช่วยเหลือเป็นล้าน เรื่องนี้จึงอยู่ระหว่างที่หน่วยงานที่เกี่ยวข้องจะเข้าไปช่วยดูแล ตนได้กำชับไปว่าอย่าไปใช้กฎเกณฑ์ที่ตายตัว เพราะความเสียหายไม่เท่ากัน จึงได้มอบให้กระทรวงเกษตรฯ ไปดูแลเฉพาะคนที่เสียหายมา ก็อาจจะต้องช่วยมากหน่อย
- S43 = สำหรับ การฟื้นฟูสภาพจิตใจ ทางกรมสุขภาพจิตก็เข้าไปดูแลแล้ว และทำร่วมกับกรมการศาสนา แต่สิ่งที่เป็นปัญหา ก็คือเมื่อนักท่องเที่ยวยังไม่กลับมา คนเหล่านี้ก็ไม่รู้จะทำมาหากินอะไร ทำให้สภาพจิตใจแย่ พอนักท่องเที่ยวไม่มา รายได้ก็ขาดหายไปตรงนี้เป็น 1 ปีที่ผ่านมา แต่ตอนนี้นักท่องเที่ยวเพิ่งจะกลับเข้ามา และนี่คือ สิ่งที่ตนเชิญเขามาเพื่อมากระตุ้นเพื่อให้เห็นสภาพปกติ และคิดว่าปีหน้าจะดีกว่านี้
- S44 = ที่ผ่านมานักท่องเที่ยวมีน้อย ซึ่งอาจเกิดจากสภาพจิตใจเขาส่วนหนึ่ง ก็ต้องเข้าไปดูแล และต้องดูแลในเรื่องรายได้ของเขาด้วย ปีนี้ น่าจะดี และจะพยายามเร่งทำต่อไป
- S45 = ผู้สื่อข่าวถามว่า การจัดงานครั้งนี้ประชาชนไม่ค่อยเห็นด้วย เพราะเขาไม่มีส่วนร่วม นายกฯ กล่าวว่า มีหลายส่วนบางที่ชาวบ้านก็มา บางที่เราก็จัดเป็นทางการ ส่วนชาวบ้านที่จัดเองก็มี
- S46 = เมื่อถามว่ามีข่าวว่าเจ้าหน้าที่ห้ามไม่ให้ชาวบ้านมายื่นหนังสือร้องเรียนให้นายกฯ พ.ต.ท.ทักษิณ กล่าวว่า ความจริงตนได้รับเรื่องร้องเรียนมาหลายครั้งแล้ว อย่างบางรายบอกว่าได้บ้านแล้วแต่ยังไม่ได้ทีวี เพราะมีเอกชนให้ทีวีด้วย จึงอยากได้บ้าง ซึ่งก็ต้องมีความพอดีด้วย

(ง.) ตัวอย่างข่าวการท่องเที่ยวฉบับเต็มแสดงประโยคที่ 37-46

<p>S47 = ส่วนที่มีข่าวว่าเจ้าหน้าที่ไม่ค่อยไปหาชาวบ้าน ได้แต่รอให้ชาวบ้านมาหาเอง นายกฯ กล่าวว่าไป ตนได้สั่งให้เจ้าหน้าที่ลงไปหาประชาชน แต่อาจมีตักค้างอยู่บ้าง แต่จะฟังอะไรก็ต้องฟังหูไว้หู เพราะบางคนขอเพื่อให้เท่าเทียมกับคนอื่นทั้งที่ความเสียหายไม่เท่ากัน เราก็ต้องดูแต่คิดว่าความช่วยเหลือรัฐบาลช่วยเหลืออย่างดีที่สุดแล้ว ซึ่งตอนนี้เงินก็ยังเหลืออยู่ แต่เราอยากให้เป็นระบบ เพราะจะได้ไม่วุ่น</p> <p>S48 = ผู้สื่อข่าวถามว่าเงินช่วยเหลือสึนามิ สามารถตรวจสอบได้เลยหรือไม่ พ.ต.ท.ทักษิณ กล่าวว่า ตรวจสอบได้ และตนได้ให้นายวิษณุ เครืองาม รองนายกรัฐมนตรีทำบัญชีทั้งหมดแล้ว ซึ่งตอนนี้มีเงินเหลืออยู่ 100 กว่าล้านบาท</p> <p>S49 = พล.อ.อ.คงศักดิ์ วันทนา รมว.มหาดไทย กล่าวถึงปัญหาการจัดสรรที่ดินให้กับประชาชนที่ประสบภัยสึนามิ ซึ่งขณะนี้มีหลายพื้นที่ตกไปอยู่ในมือของนายทุนเพื่อทำเป็นแหล่งท่องเที่ยวว่า ได้ตรวจสอบแล้ว ขณะนี้มีเพียง 3 หมู่บ้าน ใน จ.พังงา ประมาณ 140 หลังคาเรือน ที่ยังเป็นข้อพิพาทอยู่ ซึ่งคาดว่าในเดือนม.ค.49 กรมที่ดินน่าจะดูแลเรียบร้อย</p> <p>S50 = “ขอยืนยันว่าการจัดเอกสารสิทธิที่ดินจะต้องเป็นไปตามกฎหมายไม่ใช่ให้เฉพาะนายทุน และปัญหาเรื่องที่ดินที่เกิดขึ้นอาจเป็นเพราะเจ้าของที่ดินที่ยังอยากเป็นเจ้าของที่ดินอยู่ แต่ยังไม่สำเร็จ จึงเป็นปัญหาอยู่” พล.อ.อ.คงศักดิ์ กล่าว</p> <p>S51 = ส่วนเรื่องที่รัฐบาลมุ่งแต่พัฒนาด้านการท่องเที่ยวแต่ลืมแก้ปัญหาพื้นฐาน เช่น การสร้างอาชีพให้กับประชาชนที่ตกงานนั้น พล.อ.อ.คงศักดิ์ กล่าวว่า ต้องทำไปพร้อม ๆ กัน และเห็นว่ากลุ่มที่ยังมีปัญหาเรื่องอาชีพจะเป็นกลุ่มที่เคยทำงานโรงแรม ซึ่งปัจจุบันโรงแรมที่เสียหายไปกับเหตุการณ์สึนามิ ขณะนี้ยังสร้างไม่เสร็จเหลือประมาณร้อยละ 30 จึงทำให้เกิดปัญหาการขาดงาน ซึ่งกระทรวงจะสำรวจทุกพื้นที่จุดไหนใครเดือดร้อน จะลงทะเบียนและแก้ไขปัญหาให้ตรงจุด</p>
---

(จ.) ตัวอย่างข่าวการท่องเที่ยวฉบับเต็มแสดงประโยคที่ 47-51

รูปที่ 4.2 ข่าวการท่องเที่ยวไทยฉบับเต็ม

จากตัวอย่างข่าวรูปที่ 4.2 สามารถนำมาตัดคำประสมได้ทั้งหมด 18 คำ ดังรูปที่ 4.3 (ข) ได้แก่ เรือประมง หนังสือเรียน ผังเมือง ภัยพิบัติ ดอกไม้ มวลชน องค์กรเอกชน หมู่บ้าน บทกวี ตอนเช้า ศาสนาพุทธ แสงเดือน ลูกสาว กองไฟ แรงจูงใจ น้ำเค็ม ดวงวิญญาณ เด็กกำพร้า

จุฬาลงกรณ์มหาวิทยาลัย  
CHULALONGKORN UNIVERSITY

[แม้ว่าชาย, 'ห้องเที่ยว', 'ควบร่าลีสกินามิ', 'ร่าลิก', 'ปี', 'สีนามิ', 'น้ำตา', 'ท่วม', 'อันตามัน', 'ชาติ', 'ขอรื่อง', 'อย่า', 'รบกวน', 'เหยื่อ', 'สี', 'นามิ', 'รื่อง', 'ระงม', 'เหลียว', 'แล', 'รัฐ', 'เรือ', 'ประมง', 'ขนาด', 'กลาง', 'ยื่น', 'หนังสือ', 'เรียน', 'ลอย', 'แพ', 'เกาะ', 'พิพิทวง', 'มันสัญญา', 'รัฐบาล', 'พินฟู', 'บัต', 'ฝั่ง', 'เมือง', 'สรุป', 'นายก', 'ช่วยเหลือ', 'ปัญหา', 'ตก', 'หล่น', 'วาย', 'คุย', 'ไม้', 'ทำ', 'ดี', 'เคย', 'หวัง', 'บรรยากาศ', 'พิน', 'คิน', 'มท.', 'เดิน', 'เคลียร์', 'ดิน', 'พังกา', 'ซิด', 'เส้น', 'ม.ค.', 'หน้า', 'เรียบริอย', 'งาน', 'รอบ', 'วาน', 'ธ.ค.', 'จังหวัด', 'ชาย', 'ฝั่ง', 'ยกเว้น', 'ระนอง', 'สตุล', 'พีซี', 'อาลัย', 'ชีวิต', 'เหตุการณ์', 'จุด', 'สูญเสีย', 'ญาติมิตร', 'ไทย', 'พิน', 'คน', 'เต็ม', 'เศรษฐาสล', 'พิธีการ', 'ภัย', 'พิบัติ', 'ทางการ', 'เวลา', 'น.', 'นาย', 'Bill', 'Paterson', 'เอกอัครราชทูต', 'ออสเตรเลีย', 'ประจำ', 'ประเทศไทย', 'ครอบครัว', 'จำนวน', 'บริเวณ', 'ขาดป่าตอง', 'โรง', 'แรมภูเก็ตแกรนด์รีสอร์ทแอนด์', 'สปา', 'พ.ต.ท.ทักษิณ', 'ชิน วัตร', 'รัฐมนตรี', 'คณะ', 'เดินทาง', 'ปะ', 'สนทนา', 'ตัว', 'แทน', 'ญาติ', 'สวนสาธารณะ', 'โลมา', 'วาง', 'ข้อ', 'ตอก', 'ไม้', 'สด', 'ศูนย์รวมใจ', 'นาท', 'จังหวัดพังกา', 'เรือ', 'มวล', 'ชน', 'ประเทศ', 'ภพหน้า', 'พื้นที่', 'เสียง', 'หน่วย', 'ท้องถิ่น', 'ดำเนิน', 'หักเทียม', 'ต่อเนื่อง', 'ครอบคลุม', 'องค์กร', 'เอกชน', 'มูลนิธิต่าง', 'เลือก', 'ชุมชน', 'หมู่', 'บ้าน', 'เหมาะสม', 'เดิม', 'สำหรับ', 'จ.ภูเก็ต', 'ร่า', 'ลิก', 'พวง', 'หญิง', 'เวที', 'ไมโครโฟน', 'สูญหาย', 'นิก', 'เคราะห์เสียใจ', 'สื่อสาร', 'แขนง', 'รอ', 'ข่าว', 'ถ่าย', 'ภาพ', 'หนังสือพิมพ์', 'ติดตาม', 'ใกล้ชิด', '\*', 'วางศิลาฤกษ์', 'อนุสรณ์', 'สถานสีนามิ', 'จ.พังกา', 'ศิลาฤกษ์', 'หาด', 'อุทยานแห่งชาติเขา', 'หลัก', 'สำรู่', 'แขก', 'เชิญ', 'บรรเลง', 'วงดุริยางค์', 'วง', 'Dr.Sax', 'Chamber', 'Orchestra', 'มหาวิทยาลัยมหิดล', 'ชนะ', 'ประกวด', 'ระดับ', 'นาชาติ', 'เดี่ยวแซ็กโซโฟน', 'ศิลปิน', 'โลก', 'สวีเดน', 'Mr.Johan', 'Stengard', 'อ่าน', 'บท', 'กวี', 'ภาษา', 'อังกฤษ', 'มารี', 'อิมบราโน วิต', 'ศิลปิน', 'อเมริกัน', 'สุดท้าย', 'ฉาย', 'รูปแบบ', 'ก่อสร้าง', 'คัดเลือก', 'ประเทศสเปน', 'พินแลนด์', 'จีน', 'สหรัฐอเมริกา', 'ช่อง', 'ฟ้า', 'เนียง', 'อตะกัว', 'ป่า', 'ทุล', 'กระท่อมหญิงอุบลรัตนราชกัญญา', 'สิริวัฒนาพรรณวดี', 'เสด็จ', 'ประธาน', 'คำ', 'จังหวัดภูเก็ต', 'เทียน', 'เล่ม', 'ริม', 'กม.', 'องค์การบริหารส่วนจังหวัดภูเก็ต', 'อบจ.ภูเก็ต', 'แห่งประเทศไทย', 'ททท.', 'กลุ่มศิลปินภูเก็ต', 'Light', 'Up', 'Phuket', 'ไร่', 'สีนามิภาค', 'ประชาชน', 'บ้านน้ำเค็ม', 'บุญ', 'ตักบาตร', 'อุทิศ', 'กุศล', 'พิบัติภัย', 'ตอน', 'เข้า', 'พิธีกรรม', 'ศาสนา', 'พุทธ', 'คริสต์', 'บรรยากาศ', 'โคกเศรษฐี', 'นางบัวผัน', 'ภุมสนิพ', 'อายุ', 'แสง', 'เดือน', 'ซูเลลา', 'ลูก', 'สาว', 'เดิน', 'หา ราย', 'กรุงเทพฯ', 'เอ๋ย', 'หัก', 'แม่', 'กั', 'โดน', 'คลิ่น', 'ยักซ์', 'ซัด', 'หลาน', 'ทำใจ', 'เงิน', 'เสียง', 'ชีพ', 'วัง', 'นางบัวผัน', 'นายวิชเร ศวร์', 'หอมชื่น', 'ระบุ', 'รู้สึก', 'เสียใจ', 'เพื่อนบ้าน', 'แปร', 'วิกฤต', 'โอกาส', 'สู้', 'น้ำเค็ม', 'รัฐเสร็จสิ้น', 'ฉบับ', 'อนุสาวรีย์สีนามิบ้าน น้ำเค็ม', 'พิน', 'ลาด', 'ปูน', 'เสียหาย', 'ยาวชน', 'หลากรุ่น', 'เพลง', 'สิบ', 'โยน', 'จันทร์', 'กอง', 'ไฟ', 'ปล่อง', 'โคม', 'ห้อง', 'สื่อ', 'ราชการ', 'อาศัย', 'พินบ้าน', 'เก่า', 'เฮกัน', 'สถานที่', 'จัดตอน', 'เย็น', 'ต้องการ', 'รู้', 'สัปดาห์', 'อำเภอ', 'สั่ง', 'รวบรวม', 'ชื่อ', 'แนบ', 'บัตร', 'ด้วย', 'บาดเจ็บ', 'บาท', 'ตรวจ', 'เหมือน', 'แรง', 'จงใจ', 'คั้น', 'รถ', 'น้ำ', 'เค็ม', 'ทะเลาะ', 'ตาย', 'แย่ง', 'บ้านน้ำ', 'เล่า', 'นาง นิลนันท์', 'พวก', 'ระย้า', 'วัย', 'เสริม', 'ระลึก', 'บาง', 'หอรก', 'นายวิรัตน์', 'นวลทอง', 'นายวิรัตน์', 'โอกาส', 'ทีม', 'อนุสาวรีย์', 'ประเต็น', 'อุทธรณ์', 'สำ', 'กระทรวงเกษตรฯ', 'ปาก', 'ดำเนินการ', 'โต', 'แพปลา', 'รับปาก', 'ขยัน', 'บ้านบางเนียง', 'การณ', 'ต.', 'เกย', 'ตื่นเมื่อ', 'นางสุดสงวน', 'ลิมตุลย์', 'ตั้งใจ', 'ญาติพี่น้อง', 'ที่', 'พิการ', 'ขา', 'พังก', 'ซ่อมแซม', 'ซัก', 'แดง', 'กลัว', 'อบต.', 'หา', 'สนใจ', 'นางสุดสงวน', 'กระดาน', 'ข้อ', 'หลากหลาย', 'น้อง', 'แป้', 'รัก', 'ดวง', 'วิญญาณ', 'สุข', 'คดี', 'พิพิ', 'เร่งรัฐ', 'เกาะพิพิ', 'จ. กระบี่', 'นายสุรเกียรติ', 'เสถียรไทย', 'รอง', 'สภาพ', 'กลอน', 'นายสุรเกียรติ', 'เรียกรื่อง', 'เร่ง', 'ชัดเจน', 'สุเทพ', 'แนว', 'สร้าง', 'นาย สุเทพ', 'เทือกสุบรรณ', 'เลขาธิการ', 'พรรคประชาธิปัตย์', 'กรณ', 'กระทบ', 'หัก', 'กังวล', 'ข้อมูล', 'ตรวจสอบ', 'เด็กร้อน', 'ทำกิน', 'อาชีพ', 'เด็ก', 'กำพร้า', 'ลูกหลาน', 'มั่วสร้าง', 'โฆษณา', 'ชวน', 'ดู', 'ทักษิณ', 'แก้', 'สัมภาษณ์', 'รื่องเรียน', 'ทั่วถึง', 'แจ้ง', 'สำนัก

(ก) การตัดคำด้วยคัตคำโดยไม่ตัดคำประสม

[แม่ัวชาย, 'ห้องเที่ยว', 'ควบร่าลิกสินามิ', 'ปี', 'สินามิ', 'น้ำตา', 'ท่วม', 'อันตามัน', 'ชาติ', 'ขอร้อง', 'อย่า', 'รบกวน', 'เหยื่อ', 'สี', 'นามิ', 'ร้อง', 'ระงม', 'เหลียว', 'แล', 'รัฐ', 'เรือประมง', 'ขนาด', 'กลาง', 'ยื่น', 'หนังสือเรียน', 'ลอย', 'แพ', 'เกาะ', 'พิพิทวง', 'มันสัญญา', 'รัฐบาล', 'พินฟู', 'บัด', 'ฝั่งเมือง', 'สรุป', 'นายก', 'ช่วยเหลือ', 'ปัญหา', 'ตก', 'หล่น', 'วาย', 'คุย', 'ไม้', 'ท่า', 'ดี', 'เผยแพร่', 'หวัง', 'บรรยากาศ', 'พิน', 'คิน', 'เดิน', 'เคลียร์', 'คิน', 'พังกา', 'ซิด', 'เส้น', 'หน้า', 'เรียบร้อย', 'งาน', 'รอบ', 'วาน', 'จังหวัด', 'ชาย', 'ฝั่ง', 'ยกเว้น', 'ระนอง', 'สตูล', 'พิธิ', 'อาลัย', 'ชีวิต', 'เหตุการณ์', 'จุด', 'สูญเสีย', 'ญาติมิตร', 'ไทย', 'พัน', 'คน', 'เต็ม', 'เศร้าสลด', 'พิธีการ', 'ภัยพิบัติ', 'ทางการ', 'เวลา', 'นาย', 'Bill', 'Paterson', 'เอกอัครราชทูต', 'ออสเตรเลีย', 'ประจำ', 'ประเทศไทย', 'ครอบครัว', 'จำนวน', 'บริเวณ', 'หาดป่าตอง', 'โรง', 'แรมภูเก็ตแกรนด์ริสอร์ทแอนด์', 'สปา', 'พ.ต.ท.ทักษิณ', 'ชินวัตร', 'รัฐมนตรี', 'คณะ', 'เดินทาง', 'ปะ', 'สนทนา', 'ตัว', 'แทน', 'ญาติ', 'สวนสาธารณะ', 'โลมา', 'วาง', 'ข้อ', 'ดอกไม้', 'สด', 'ศูนย์รวมใจ', 'นาทิจ', 'จังหวัดพังกา', 'สื่อ', 'มวลชน', 'ประเทศ', 'ภพหน้า', 'พื้นที่', 'เสียง', 'หน่วย', 'ท้องถิ่น', 'ดำเนิน', 'หัดเทียม', 'ต่อเนื่อง', 'ครอบคลุม', 'องค์กรเอกชน', 'มูลนิธิต่าง', 'เลือก', 'ชุมชน', 'หมู่บ้าน', 'เหมาะสม', 'เดิม', 'สำหรับ', 'จ.ภูเก็ต', 'ร่า', 'ลึก', 'พวง', 'หญิง', 'เวที', 'ไมโครโฟน', 'สูญหาย', 'นึก', 'เศร้าสะเทือนใจ', 'สื่อสาร', 'แขนง', 'รอ', 'ข่าว', 'ถ่าย', 'ภาพ', 'หนังสือพิมพ์', 'ติดตาม', 'ใกล้ชิด', 'วางศิลาฤกษ์', 'อนุสรณ์', 'สถานสินามิ', 'จ.พังกา', 'ศิลาฤกษ์', 'หาด', 'อุทยานแห่งชาติเขา', 'หลัก', 'สำรู่', 'แขก', 'เชิญ', 'บรรเลง', 'วงดุริยางค์', 'Dr.Sax', 'Chamber', 'Orchestra', 'มหาวิทยาลัยมหิดล', 'ชนะ', 'ประกวด', 'ระดับ', 'นานาชาติ', 'เดี่ยวแซ็กโซโฟน', 'ศิลปิน', 'โลก', 'สวีเดน', 'Mr.Johan', 'Stengard', 'อ่าน', 'บทกวี', 'ภาษา', 'อังกฤษ', 'มารี', 'อิมบรานโนวิต', 'ศิลปิน', 'อเมริกัน', 'สุดท้าย', 'ฉาย', 'รูปแบบ', 'ก่อสร้าง', 'คัดเลือก', 'ประเทศสเปน', 'ฟินแลนด์', 'จีน', 'สหรัฐอเมริกา', 'ช่อง', 'ฟ้า', 'เนียง', 'อ.ตะกั่ว', 'ป่า', 'ทูล', 'กระหม่อมหญิงอุบลรัตนราชกัญญา', 'สิริวัฒนาพรรณวดี', 'เสด็จ', 'ประธาน', 'คำ', 'จังหวัดภูเก็ต', 'เทียน', 'เล่ม', 'ริม', 'องค์การบริหารส่วนจังหวัดภูเก็ต', 'อบจ.ภูเก็ต', 'แห่งประเทศไทย', 'กลุ่มศิลปินภูเก็ต', 'Light', 'Up', 'Phuket', 'ไร่', 'สินามิภาค', 'ประชาชน', 'บ้านน้ำเค็ม', 'บุญ', 'ตักบาตร', 'อุทิศ', 'กุศล', 'พิธีกรรม', 'ตอนเช้า', 'พิธีกรรม', 'ศาสนาพุทธ', 'คริสต์', 'บรรยากาศ', 'โคกเศร้า', 'นางบัวผัน', 'ภุมสนิท', 'อายุ', 'แสงเดือน', 'ชุลลลา', 'ลูกสาว', 'เดินท์', 'หาราย', 'กรุงเทพฯ', 'เอ่ย', 'พัก', 'แม่', 'กั', 'โดน', 'คลื่น', 'ยักษ์', 'ซัด', 'หลาน', 'ทำใจ', 'เงิน', 'เสียง', 'ชีพ', 'วัง', 'นางบัวผัน', 'นายวัชเรศวร์', 'หอมชื่น', 'ระบุ', 'รู้สึก', 'เสียใจ', 'เพื่อนบ้าน', 'แปร', 'วิกฤต', 'โอกาส', 'สู้', 'น้ำเค็ม', 'รัฐเสร็จสิ้น', 'ฉบับ', 'อนุสาวรีย์สินามิบ้านน้ำเค็ม', 'พิน', 'ลาด', 'ปูน', 'เสียหาย', 'เยาวชน', 'หลากรุ่น', 'เพลง', 'สิบ', 'โยน', 'จันทร์', 'กองไฟ', 'ปล่อย', 'คอม', 'ห้อง', 'สื่อ', 'ราชการ', 'อาศัย', 'พินบ้าน', 'เก่า', 'เฮกัน', 'สถานที่', 'จัดตอน', 'เย็น', 'ต้องการ', 'รู้', 'สัปดาห์', 'อำเภอ', 'สั่ง', 'รวบรวม', 'ชื่อ', 'แนบ', 'บัตร', 'ด้วย', 'บาดเจ็บ', 'บาท', 'ตรวจ', 'เหมือน', 'แรงจูงใจ', 'คั้น', 'รถ', 'น้ำเค็ม', 'ทะเลาะ', 'ตาย', 'แย่ง', 'บ้านน้ำ', 'เล่า', 'นางนิลนันท์', 'พวก', 'ระย้า', 'วัย', 'เสริม', 'ระลึก', 'บาง', 'ห rokok', 'นายวิรัตน์', 'นวลทอง', 'นายวิรัตน์', 'โอกาส', 'หิม', 'อนุสาวรีย์', 'ประเด็น', 'อุทธรณ์', 'สำ', 'กระทรวงเกษตรฯ', 'ปาก', 'ดำเนินการ', 'โต', 'แพปลา', 'รับปาก', 'ขยัน', 'บ้านบางเนียง', 'การณ', 'เกย', 'ตื่นเมื่อ', 'นางสุดสงวน', 'ลิมคสุย์', 'ตั้งใจ', 'ญาติพี่น้อง', 'พิการ', 'ขา', 'พังก', 'ซ่อมแซม', 'ซัก', 'แดง', 'กลัว', 'หา', 'สนใจ', 'นางสุดสงวน', 'กระดาน', 'ข้อ', 'หลากหลาย', 'น้อง', 'แป้', 'รัก', 'ดวงวิญญาณ', 'สุข', 'คดี', 'พิพิ', 'เร่งรัฐ', 'เกาะพิพิ', 'จ.กระบี่', 'นายสุรเกียรติ์', 'เสถียรไทย', 'รอง', 'สภาพ', 'กลอน', 'นายสุรเกียรติ์', 'เรียกร้อง', 'เร่ง', 'ชัดเจน', 'สุเทพ', 'นะ', 'สร้าง', 'นายสุเทพ', 'เทือกสุบรรณ', 'เลขาธิการ', 'พรรคประชาธิปัตย์', 'กรณ', 'กระหบ', 'พัก', 'กังวล', 'ข้อมูล', 'ตรวจสอบ', 'เดือดร้อน', 'ทำกิน', 'อาชีพ', 'เด็กกำพร้า', 'ลูกหลาน', 'มีสร้าง', 'โฆษณา', 'ชวน', 'ดู', 'ทักษิณ', 'แก่', 'สัมภาษณ์', 'ร้องเรียน', 'ทั่วถึง', 'แจ้ง', 'สำนักปลัด', 'สำนักนายกรัฐมนตรี', 'เรื่อง', 'บ้านเกิน', 'บังเอิญ', 'ไอ', 'ทีวี', 'อนุโลม', 'นโยบาย', 'หลักเกณฑ์', 'สำละ', 'หมื่น', 'ราคา', 'แพง', 'คาดหวัง', 'ล้าน', 'ดูแล', 'ทำซั้บ', 'กฎเกณฑ์', 'มอบ', 'จิตใจ', 'กรมสุขภาพจิต', 'กรมการศาสนา', 'มหาากิน', 'ย่าแย้', 'ขาดหาย,

(ข) การตัดคำด้วยคัตคำโดยตัดคำประสมร่วมด้วย

รูปที่ 4.3 การตัดคำระหว่างการตัดคำด้วยคัตคำโดยไม่ตัดคำประสม และการตัดคำด้วยคัตคำร่วมกับ การตัดคำประสมจากตัวอย่างข่าว

จากสมมติฐานหลักข้างต้นคือการสรุปใจความสำคัญภาษาไทยนั้นจะแม่นยำมากขึ้น ส่วนหนึ่งมาจากการพิจารณาตัดคำประสมร่วมด้วย เพื่อเป็นการพิสูจน์สมมติฐานว่าเป็นจริงหรือไม่ ในงานวิจัยนี้ได้ทดลองสรุปใจความสำคัญของข่าวทั้งหมด 400 ข่าว โดยใช้ข้อมูลเข้าเป็นข่าวที่ใช้การตัดคำอย่างเดียวเปรียบเทียบกับตัดคำร่วมกับคำประสม โดย 400 ข่าวนี้เมื่อมีการตัดคำร่วมกับคำประสมสามารถแบ่งข่าวออกได้เป็น 3 ชุด ชุดที่ 1 คือ ข่าวที่มีคำประสมตั้งแต่ 0-5 คำ (จำนวน 293 ข่าว) ชุดที่ 2 คือ ข่าวที่มีคำประสม 6-10 คำ (จำนวน 97 ข่าว) และชุดที่ 3 คือข่าวที่มีคำประสมตั้งแต่ 11 คำขึ้นไป (จำนวน 10 ข่าว) แสดงดังตารางที่ 4.2

ตารางที่ 4.2 ตารางแสดงผลการสรุปใจความสำคัญโดยแบ่งชุดข้อมูลตามจำนวนคำประสม

ชุดข้อมูล	ตัดคำประสม	ค่าความถูกต้อง	ค่าความแม่นยำ	ค่าความระลึก	ค่าวัดประสิทธิภาพ
ชุดที่ 1	มีคำประสม	0.6808	0.6274	0.7479	0.6824
ชุดที่ 1	ไม่มีคำประสม	0.6200	0.6673	0.7871	<b>0.7223</b>
ชุดที่ 2	มีคำประสม	0.6971	0.7574	0.8514	<b>0.8017</b>
ชุดที่ 2	ไม่มีคำประสม	0.6907	0.7517	0.8437	0.7950
ชุดที่ 3	มีคำประสม	0.6000	0.5863	0.8166	<b>0.6825</b>
ชุดที่ 3	ไม่มีคำประสม	0.5827	0.5690	0.8141	0.6698

จากตารางที่ 4.2 แสดงผลการสรุปใจความสำคัญของข่าวการท่องเที่ยวไทยทั้งหมด 400 ข่าว ระหว่างการตัดคำแบบมีและไม่มีคำประสม โดยชุดข่าวที่มีจำนวนคำประสมในข่าวระหว่าง 6-10 คำ หรือมากกว่านั้น มีแนวโน้มที่จะได้ผลการสรุปใจความสำคัญที่ดีกว่าการสรุปใจความสำคัญแบบไม่มีคำประสม สำหรับข่าวชุดที่ 1 การตัดคำแบบไม่มีคำประสมมีประสิทธิภาพสูงสุดที่ 0.7223 ส่วนข่าวชุดที่ 2 และข่าวชุดที่ 3 การตัดคำแบบมีคำประสมร่วมด้วยให้ประสิทธิภาพสูงที่สุดคือ 0.8017 และ 0.6825 ตามลำดับ

#### 4.3 การทดลองเปรียบเทียบการตัดคำด้วยคัตคำร่วมกับคำประสมและการตัดคำด้วยคัตคำร่วมกับเอ็นแกรม

เพื่อเปรียบเทียบประสิทธิภาพของการตัดคำด้วยคัตคำร่วมกับตัดคำประสมและการตัดคำด้วยคัตคำร่วมกับตัดคำด้วยเอ็นแกรมว่าวิธีการตัดคำด้วยคัตคำร่วมกับการตัดคำวิธีการใดที่ให้ประสิทธิภาพดีกว่า ในงานวิจัยนี้ ผู้วิจัยจึงเลือกทดสอบเปรียบเทียบประสิทธิภาพของการสรุปใจความสำคัญแบบสกัดของภาษาไทยระหว่างการตัดคำประสมและการตัดคำด้วยเอ็นแกรม โดยเลือกขนาด 2 แกรมและ 3 แกรมเนื่องจากเป็นขนาดที่มีการทดลองกันอย่างแพร่หลายในงานด้านภาษาธรรมชาติของภาษาไทย โดยได้ทำการทดลองกับวิธีที่นำเสนอ 2 วิธี คือ 1. การจัดกลุ่มประโยคด้วยเคมีน และ 2. การเลือกประโยคด้วยวิธีหาค่าสำคัญของประโยคจากหัวข้อข่าว ด้วยอัตราบัพัตทั้ง 3 ระดับคือ ร้อยละ 20 30 และ 40 ได้ผลดังตารางที่ 4.3 ซึ่งเป็นค่าเฉลี่ยของจากทั้งสองวิธีและค่าเฉลี่ยจากอัตราบัพัตทั้ง 3 ระดับรวมกัน

ตารางที่ 4.3 ผลการเปรียบเทียบการสรุปใจความสำคัญโดยการตัดคำประสมและการตัดคำด้วยเอ็นแกรม

ชุดข้อมูล	ตัดคำ	ค่าความถูกต้อง	ค่าความแม่นยำ	ค่าความระลึก	ค่าวัดประสิทธิภาพ
ชุดที่ 1	คำประสม	0.6163	0.7302	0.7134	0.7217
ชุดที่ 2	คำประสม	0.6205	0.7298	0.7256	0.7277
ชุดที่ 3	คำประสม	0.6340	0.7395	0.7411	0.7403
ชุดที่ 4	คำประสม	0.5950	0.6332	0.7668	0.6935
ชุดที่ 5	คำประสม	0.6286	0.7323	0.7386	0.7354
<b>ค่าเฉลี่ย 5 ชุด</b>	<b>คำประสม</b>	0.6189	0.7130	0.7371	0.7237
ชุดที่ 1	2-แกรม	0.6188	0.7309	0.7169	0.7238
ชุดที่ 2	2-แกรม	0.6178	0.7261	0.7280	0.7270
ชุดที่ 3	2-แกรม	0.6450	0.7460	0.7524	0.7492
ชุดที่ 4	2-แกรม	0.6379	0.6646	0.7966	0.7246
ชุดที่ 5	2-แกรม	0.6317	0.7373	0.7383	0.7378
<b>ค่าเฉลี่ย 5 ชุด</b>	<b>2-แกรม</b>	0.6302	0.7210	0.7464	0.7325
ชุดที่ 1	3-แกรม	0.6218	0.7294	0.7299	0.7292
ชุดที่ 2	3-แกรม	0.6210	0.7266	0.7359	0.7310

ชุดข้อมูล	ตัดคำ	ค่าความถูกต้อง	ค่าความแม่นยำ	ค่าความระลึก	ค่าวัดประสิทธิภาพ
ชุดที่ 3	3-แกรม	0.6445	0.7447	0.7546	0.7495
ชุดที่ 4	3-แกรม	0.6406	0.6668	0.7992	0.7270
ชุดที่ 5	3-แกรม	0.6290	0.7381	0.7318	0.7348
<b>ค่าเฉลี่ย 5 ชุด</b>	<b>3-แกรม</b>	<b>0.6314</b>	<b>0.7211</b>	<b>0.7503</b>	<b>0.7343</b>

จากตารางที่ 4.3 สามารถสรุปได้ว่าจากข่าวการท่องเที่ยวไทยทั้งหมด 400 ข่าว การตัดคำด้วย 3 แกรมมีค่าเฉลี่ยค่าความถูกต้อง ค่าเฉลี่ยความแม่นยำ ค่าเฉลี่ยความระลึกและค่าเฉลี่ยประสิทธิภาพดีที่สุด ส่วนค่าเฉลี่ยของการตัดคำประสมแบบมีความหมายจากพจนานุกรมมีประสิทธิภาพน้อยที่สุด เนื่องจากคลังข้อมูลที่มีจำกัด อย่างไรก็ตามงานวิจัยนี้ทำให้มีคลังข้อมูลรายการคำประสมเพื่อนำไปใช้ในพัฒนาในงานวิชาการด้านภาษาไทยเพิ่มขึ้น

#### 4.4 การทดลองเพื่อหาอัตราการย่อความ

เพื่อให้ได้จำนวนประโยคในบทสรุปใจความของข่าวการท่องเที่ยวไทยที่สอดคล้องกับผู้เชี่ยวชาญมากที่สุด จึงได้ทำการทดลองการสรุปใจความสำคัญทั้ง 2 วิธีคือ 1. การหาความสำคัญของคำและประโยคด้วยการหาความถี่ของคำ การหาค่าความสัมพันธ์โดยใช้อัลกอริทึมการแยกค่าแบบเดี่ยว (SVD) และจัดกลุ่มประโยคด้วยเคมีน และ 2. การหาความสำคัญของคำด้วยการหาความถี่ของคำและเลือกประโยคจากหัวข้อข่าว โดยทั้ง 2 วิธีมีการตัดคำประสมร่วมด้วย โดยทดลองในอัตราการสรุปที่แตกต่างกัน ตั้งแต่อัตราการย่อร้อยละ 20, 30 และ 40 กล่าวคือเหลือเนื้อหาข่าวหลังจากย่อแล้วร้อยละ 80, 70 และ 60 ตามลำดับ และเปรียบเทียบผลสรุปเฉลี่ยที่ได้จากวิธีการสรุปใจความสำคัญทั้ง 2 วิธีข้างต้นกับการสรุปใจความสำคัญจากข่าวของผู้เชี่ยวชาญด้านภาษาไทย ซึ่งค่าเฉลี่ยของค่าความแม่นยำ ค่าเฉลี่ยของค่าความระลึก และค่าเฉลี่ยของค่าประสิทธิภาพของผลการสรุปใจความสำคัญของข่าวการท่องเที่ยวไทยทั้งหมด 400 ข่าว จากการทดลอง 5 ครั้ง ครั้งละ 80 ข่าว ดังแสดงในตารางที่ 4.4



ตารางที่ 4.4 ผลการทดลองใช้อัตราส่วนที่ต่างกันในการสรุปใจความสำคัญ

ค่าบีบอัดร้อยละ	ค่าความถูกต้อง	ค่าความแม่นยำ	ค่าความระลึก	ค่าวัดประสิทธิภาพ
20	0.7060	0.8049	0.8347	0.8180
30	0.6053	0.7119	0.7368	0.7230
40	0.5453	0.6222	0.6397	0.6301

จากตารางที่ 4.4 ผลการทดลองใช้อัตราส่วนที่ต่างกันในการสรุปใจความสำคัญ สรุปได้ว่ายิ่งย่อมากค่าเฉลี่ยของค่าความถูกต้อง ค่าความแม่นยำ ค่าเฉลี่ยของค่าความระลึกและค่าเฉลี่ยของค่าประสิทธิภาพยิ่งลดลงไปด้วย ดังนั้นเพื่อความแม่นยำจึงเลือกใช้อัตราการบีบอัดร้อยละ 20 คือย่อให้เหลือเนื้อหาข่าวร้อยละ 80 ของข่าวตั้งต้น เพื่อความแม่นยำและประสิทธิภาพของเนื้อหาข่าวที่ดี

#### 4.5 การทดลองเปรียบเทียบผลของแนวคิดที่นำเสนอ

เพื่อทดสอบแนวคิดที่นำเสนอนั้นมีส่วนช่วยให้การสรุปใจความสำคัญของข่าวการท่องเที่ยวมีประสิทธิภาพที่ดีขึ้น ผู้วิจัยจึงทดลองกับวิธีที่นำเสนอ 2 วิธี คือ 1. การจัดกลุ่มประโยคด้วยเคมีน และ 2. การเลือกประโยคด้วยวิธีหาคำสำคัญของประโยคจากหัวข้อข่าว โดยการทดลองนี้มีอัตราการบีบอัดร้อยละ 20 ตามที่ได้ทดลอง ในหัวข้อที่ 4.4 ใช้การวัดผลโดยเปรียบเทียบกับข่าวที่ได้จากการสรุปใจความสำคัญจากผู้เชี่ยวชาญด้านภาษาไทย ได้ผลเฉลี่ยของค่าความแม่นยำ ค่าความระลึกและประสิทธิภาพจาก F1-score จาก 400 ข่าวการท่องเที่ยวไทย โดยแบ่งเป็นการทดสอบทั้งหมด 5 ครั้ง ครั้งละ 80 ข่าว ได้ผลดังตารางที่ 4.5

ตารางที่ 4.5 ผลการเปรียบเทียบแนวคิดที่นำเสนอโดยใช้ 400 ข่าวการท่องเที่ยวไทย

วิธีหาน้ำหนักของคำสำคัญ	การเลือกประโยค	อัลกอริทึมตัดคำ	ค่าความแม่นยำ (Precision)	ค่าความระลึก (Recall)	ค่าประสิทธิภาพ (F-Score)
SVD	K-means	ตัดคำ	0.7993	0.8245	0.8103
SVD	K-means	ตัดคำ + คำประสม	0.7950	0.82	0.8060
หัวข้อข่าว	Sentence Similarity	ตัดคำ	0.8071	0.8343	0.8192

วิธีหาน้ำหนักของคำสำคัญ	การเลือกประโยค	อัลกอริทึมตัดคำ	ค่าความแม่นยำ (Precision)	ค่าความระลึก (Recall)	ค่าประสิทธิภาพ (F-Score)
หัวข้อข่าว	Sentence Similarity	ตัดคำ + คำประสม	0.8074	0.8347	0.8194
หัวข้อข่าว	Sentence Ranked	ตัดคำ	0.8066	0.8329	0.8181
หัวข้อข่าว	Sentence Ranked	ตัดคำ + คำประสม	<b>0.8097</b>	<b>0.8367</b>	<b>0.8216</b>

จากผลการทดลองตารางที่ 4.5 เป็นผลการทดลองจากทั้งหมด 5 ครั้ง ครั้งละ 80 ข่าวรวมทั้งหมด 400 ข่าว พบว่าค่าเฉลี่ยความแม่นยำ ค่าเฉลี่ยความระลึกและค่าเฉลี่ยประสิทธิภาพของการคำนวณน้ำหนักของคำสำคัญด้วยการหาความถี่ของคำและเลือกประโยคจากการเรียงลำดับผลรวมของความถี่ของคำจากหัวข้อข่าวโดยมีการตัดคำประสมร่วมด้วยมีประสิทธิภาพและความแม่นยำสูงสุด

#### 4.6 การทดลองเปรียบเทียบผลกับงานวิจัยที่เกี่ยวข้อง

การทดลองนี้เลือกใช้อัตราบีบอัดร้อยละ 20 ตามที่ได้ทดลองในหัวข้อที่ 4.3 รวมทั้งใช้ทดสอบอัตราบีบอัดร้อยละ 30 และ 40 ด้วย เพื่อนำมาทดลองเปรียบเทียบกับงานวิจัยที่เกี่ยวข้องโดยใช้ข้อมูลข่าวการท่องเที่ยวไทยตั้งต้นชุดเดียวกัน ทั้งหมด 400 ข่าว สามารถสรุปผลได้ดังตารางที่ 4.6 ตารางที่ 4.7 และ ตารางที่ 4.8

ตารางที่ 4.6 เปรียบเทียบผลการทดลองกับงานวิจัยที่เกี่ยวข้อง (อัตราการบีบอัดร้อยละ 20)

วิธีหาน้ำหนักของคำสำคัญ	การเลือกประโยค	อัลกอริทึมตัดคำ	ค่าความแม่นยำ	ค่าความระลึก	ค่าประสิทธิภาพ
SVD	Gong และ คณะ	ตัดคำ	0.8031	0.8286	0.8143
SVD	Gong และ คณะ	ตัดคำ + คำประสม	0.8041	0.8294	0.8152
SVD	Gong และ คณะ	3-แกรม	<b>0.8172</b>	0.8449	0.8294

วิธีหาน้ำหนัก ของคำสำคัญ	การเลือก ประโยค	อัลกอริทึมตัด คำ	ค่าความ แม่นยำ	ค่าความ ระลึก	ค่า ประสิทธิภาพ
SVD	Murray และ คณะ	ตัดคำ	0.8081	0.8537	0.8284
SVD	Murray และ คณะ	ตัดคำ + คำ ประสม	0.8102	0.8544	0.8296
SVD	Murray และ คณะ	3-แกรม	0.8074	<b>0.8571</b>	<b>0.8298</b>
หัวข้อข่าว	Sentence Ranked	ตัดคำ	0.8066	0.8329	0.8181
หัวข้อข่าว	Sentence Ranked	ตัดคำ + คำ ประสม	0.8097	0.8367	0.8216
หัวข้อข่าว	Sentence Ranked	3-แกรม	0.8119	0.8398	0.8242

จากตารางที่ 4.6 พบว่าค่าเฉลี่ยความแม่นยำของงานวิจัยของ Gong และคณะ กับ ประสิทธิภาพการตัดคำแบบ 3-แกรม มีค่าความแม่นยำสูงสุด ส่วนค่าเฉลี่ยความระลึก และค่าเฉลี่ย ประสิทธิภาพของงานวิจัย Murray และคณะ กับประสิทธิภาพการตัดคำด้วยตัดคำแบบ 3-แกรม ประสิทธิภาพสูงสุดสำหรับการสรุปใจความสำคัญของข่าวการท่องเที่ยวไทยสำหรับการปิดที่ร้อยละ 20

ตารางที่ 4.7 เปรียบเทียบผลการทดลองกับงานวิจัยที่เกี่ยวข้อง (อัตราการปิดที่ร้อยละ 30)

วิธีหาน้ำหนัก ของคำสำคัญ	การเลือก ประโยค	อัลกอริทึม ตัดคำ	ค่าความ แม่นยำ	ค่าความ ระลึก	ค่า ประสิทธิภาพ
SVD	Gong และ คณะ	ตัดคำ	0.7080	0.7289	0.7173
SVD	Gong และ คณะ	ตัดคำ + คำ ประสม	0.7016	0.7225	0.7110
SVD	Gong และ คณะ	3-แกรม	<b>0.7265</b>	0.7502	<b>0.7371</b>

วิธีหาน้ำหนัก ของคำสำคัญ	การเลือก ประโยค	อัลกอริทึม ตัดคำ	ค่าความ แม่นยำ	ค่าความ ระลึก	ค่า ประสิทธิภาพ
SVD	Murray และ คณะ	ตัดคำ	0.7135	0.7515	0.7307
SVD	Murray และ คณะ	ตัดคำ + คำ ประสม	0.7158	0.7555	0.7337
SVD	Murray และ คณะ	3-แกรม	0.7160	<b>0.7746</b>	0.7346
หัวข้อความ	Sentence Ranked	ตัดคำ	0.7180	0.7391	0.7274
หัวข้อความ	Sentence Ranked	ตัดคำ + คำ ประสม	0.7211	0.7626	0.7356
หัวข้อความ	Sentence Ranked	3-แกรม	0.7240	0.7476	0.7346

จากตารางที่ 4.7 พบว่าค่าเฉลี่ยความแม่นยำและค่าเฉลี่ยประสิทธิภาพของงานวิจัยของ Gong และคณะ กับประสิทธิภาพการตัดคำแบบ 3-แกรม มีความแม่นยำและประสิทธิภาพสูงสุด สำหรับการสรุปใจความสำคัญของข่าวการท่องเที่ยวไทยสำหรับการปิดที่ร้อยละ 30

ตารางที่ 4.8 เปรียบเทียบผลการทดลองกับงานวิจัยที่เกี่ยวข้อง (อัตราการปิดที่ร้อยละ 40)

วิธีหาน้ำหนัก ของคำสำคัญ	การเลือก ประโยค	อัลกอริทึม ตัดคำ	ค่าความ แม่นยำ	ค่าความ ระลึก	ค่า ประสิทธิภาพ
SVD	Gong และ คณะ	ตัดคำ	0.6094	0.6248	0.6164
SVD	Gong และ คณะ	ตัดคำ + คำ ประสม	0.6062	0.6219	0.6133
SVD	Gong และ คณะ	3-แกรม	0.6527	0.6518	0.6414
SVD	Murray และ คณะ	ตัดคำ	0.6169	0.6381	0.6262

วิธีหาหน้าหนัก ของคำสำคัญ	การเลือก ประโยค	อัลกอริทึม ตัดคำ	ค่าความ แม่นยำ	ค่าความ ระลึก	ค่า ประสิทธิภาพ
SVD	Murray และ คณะ	ตัดคำ + คำ ประสม	0.6227	0.6459	0.6331
SVD	Murray และ คณะ	3-แกรม	0.6309	0.6883	0.6577
หัวข้อข่าว	Sentence Ranked	ตัดคำ	0.6401	0.6556	0.6471
หัวข้อข่าว	Sentence Ranked	ตัดคำ + คำ ประสม	0.6411	0.6570	0.6483
หัวข้อข่าว	Sentence Ranked	3-แกรม	<b>0.6500</b>	<b>0.6887</b>	<b>0.6585</b>

จากตารางที่ 4.8 ทดสอบด้วยอัตราการบีบอัดที่มากขึ้นเป็นร้อยละ 40 พบว่า พบว่าค่าเฉลี่ยความแม่นยำ ค่าเฉลี่ยประสิทธิภาพและค่าเฉลี่ยความระลึกของแนวคิดที่นำเสนอกับการตัดคำแบบ 3-แกรม มีประสิทธิภาพดีที่สุด สำหรับการสรุปใจความสำคัญของข่าวการท่องเที่ยวไทยสำหรับการบีบอัดที่ร้อยละ 40

## บทที่ 5

### สรุปผลการวิจัย

#### 5.1 สรุปผลการวิจัย

ในงานวิจัยนี้ได้นำเสนอวิธีการสรุปใจความสำคัญจากข่าวการท่องเที่ยวภาษาไทย 2 วิธีคือการเลือกประโยคจากการจัดกลุ่มประโยคด้วยเคมีนและการเลือกประโยคด้วยวิธีหาคำสำคัญประโยคจากหัวข้อข่าว โดยมีการพัฒนาและสร้างคลังข้อมูลรายการคำประสมเพื่อช่วยเพิ่มประสิทธิภาพในการตัดคำ โดยการทดลองนี้ใช้ข้อมูลข่าวการท่องเที่ยวไทย ทั้งหมด 400 ข่าวสำหรับใช้ทดลองในการสรุปใจความสำคัญ และ 5,000 ข่าวสำหรับการสร้างคลังข้อมูลรายการคำประสม และในการทดสอบวิธีการที่นำเสนอ โดยการเปรียบเทียบผลจากการสรุปที่ได้จากผู้เชี่ยวชาญด้านภาษาไทย เทียบกับผลสรุปที่ได้ จากงานวิจัยนี้ในขั้นตอนการสร้างคำประสมได้คำประสมจากการประสมระหว่าง คำนาม คำกริยา และคำวิเศษณ์ ทั้งหมด จำนวน 2,340 คำ เมื่อทดลองเปรียบเทียบการตัดคำร่วมกับคำประสมและการตัดคำร่วมกับเอ็นแกรม การตัดคำด้วย 3 แกรมมีค่าเฉลี่ยค่าความถูกต้อง ค่าเฉลี่ยความแม่นยำ ค่าเฉลี่ยความระลึกลับและค่าเฉลี่ยประสิทธิภาพดีที่สุดใน ส่วนค่าเฉลี่ยของการตัดคำประสมมีประสิทธิภาพน้อยที่สุด เนื่องจากคลังข้อมูลที่มีจำกัด อย่างไรก็ตามงานวิจัยนี้ทำให้มีคลังข้อมูลรายการคำประสมที่มีความหมายในพจนานุกรมเพื่อนำไปใช้ในพัฒนาในงานวิชาการด้านภาษาไทยเพิ่มขึ้นและเมื่อทดลองเปรียบเทียบจากวิธีการที่นำเสนอร่วมกับการตัดคำประสม ผลการทดลองพบว่าวิธีตัดคำด้วยคัตคำร่วมกับตัดคำประสมจากชุดข่าวที่มีจำนวนคำประสมในช่วง 6-10 คำ หรือมากกว่า นั้น มีแนวโน้มที่จะได้ผลการสรุปใจความสำคัญที่ดีกว่าการสรุปใจความสำคัญแบบไม่มีคำประสมและการสรุปใจความสำคัญโดยใช้การคำนวณค่าน้ำหนักของคำสำคัญโดยหาค่าความถี่ของคำจากหัวข้อข่าวและเลือกประโยคเรียงลำดับจากผลรวมความถี่ของคำสำคัญจากหัวข้อข่าวมีประสิทธิภาพและความแม่นยำสูงสุดโดยมีค่าความแม่นยำ ค่าความระลึกลับและค่าวัดประสิทธิภาพอยู่ที่ 0.8097 0.8367 และ 0.8216 ตามลำดับ และเมื่อทดสอบกับการตัดคำโดยใช้ 3-แกรมในการตัดคำร่วมกับคัตคำจะมีประสิทธิภาพเพิ่มขึ้นเป็น 0.8119 0.8398 และ 0.8242 ตามลำดับที่อัตราการบีบอัดร้อยละ 20

เมื่อเปรียบเทียบวิธีการที่นำเสนอกับวิธีจากงานวิจัยที่เกี่ยวข้องพบว่าการสรุปใจความสำคัญจากงานของ Gong และ Murray มีประสิทธิภาพความแม่นยำสูงที่สุดและมีประสิทธิภาพดีที่สุดในสำหรับการสรุปใจความสำคัญของข่าวการท่องเที่ยวไทย เนื่องจากงานของ Murray และคณะเป็นการคำนวณหาค่าความสัมพันธ์ของคอนเซ็ปต์กับประโยคมาใช้ในการสรุปใจความสำคัญ โดยแต่ละคอนเซ็ปต์สามารถเลือกได้มากกว่า 1 ประโยค โดยการเลือกจำนวนประโยคในแต่ละคอนเซ็ปต์จะขึ้นอยู่กับ

กับการคำนวณของ Matrix  $\Sigma$  ดังนั้นการเลือกประโยคของ Murray จึงได้ประโยคที่มาจากคอนเซปต์อันดับต้นๆ และจากคอนเซปต์ต้นๆ นั้นเองทำให้ประโยคนั้นเป็นประโยคที่สำคัญที่นำมาใช้ในการสรุปใจความสำคัญที่แม่นยำและมีประสิทธิภาพ อย่างไรก็ตามเมื่อทดสอบด้วยอัตราการบิบัติที่มากขึ้นเป็นร้อยละ 40 พบว่าจากอัตราการบิบัติที่ร้อยละ 40 พบว่า ค่าเฉลี่ยความแม่นยำ ค่าเฉลี่ยความแม่นยำและค่าเฉลี่ยประสิทธิภาพของงานวิจัยที่นำเสนอที่มีการตัดคำแบบ 3-แกรมร่วมกับตัดคำให้ความแม่นยำมากที่สุด ดังนั้นอาจจะสรุปได้ว่าวิธีการที่ใช้ความถี่ของคำจากหัวข้อข่าวเป็นชุดคำเริ่มต้นเพื่อหาประโยคที่มีความสำคัญในการสรุปใจความมีแนวโน้มที่จะมีประสิทธิภาพที่ดีกว่าวิธีการแยกคำแบบเดี่ยวเมื่อต้องการอัตราการบิบัติที่สูงขึ้น ทั้งนี้ประสิทธิภาพของวิธีการที่ใช้ความถี่ของคำก็ขึ้นกับชุดคำที่ใช้ในการแสดงแต่ละหัวข้อข่าวด้วย

จากผลการทดลองทั้งหมดนี้สามารถสรุปได้ว่า การตัดคำด้วยตัดคำและตัดคำประสมร่วมด้วยในงานวิจัยนี้ให้ดีกว่าการตัดคำจากตัดคำเพียงอย่างเดียวและการสรุปใจความสำคัญโดยใช้การคำนวณหาความถี่ของคำและเลือกประโยคจากการเรียงลำดับผลรวมของความถี่ของคำที่เกิดขึ้นนั้นได้ผลการทดลองที่ดีโดยเฉพาะเมื่อมีอัตราการบิบัติมาก การสร้างคลังข้อมูลคำประสมจากงานวิจัยนี้สามารถนำไปประยุกต์ใช้ในการสร้างคำประสมจากแหล่งอื่นๆ ได้ จากการแก้ไขโปรแกรมในส่วนของ การเลือกตำแหน่งของคำที่ติดกันหรือเพิ่มประเภทของคำที่ต้องการนำมาประสมได้ โดยอาจจะเพิ่มจำนวนตำแหน่งของคำเพื่อใช้ในการสร้างคลังรายการคำประสมเพื่อใช้ในการพัฒนาการสรุปใจความสำคัญหรืองานวิจัยอื่นๆ เกี่ยวกับภาษาไทยต่อไป

## 5.2 แนวทางวิจัยในอนาคต

- 1) พัฒนาค้นคลังข้อมูลรายการคำประสม โดยใช้วิธีการเลือกคำจากข่าวหรือบทความใดๆ โดยเลือกคำที่ติดกันมากกว่า 1 ตำแหน่งและเพิ่มลักษณะการประสมคำ เพื่อนำมาใช้ในการพัฒนางานวิจัยด้านภาษาไทยต่อไป
- 2) ทดลองเปรียบเทียบกับอัลกอริทึมการตัดคำที่ใช้วิธีการตัดคำแบบสอดคล้องมากที่สุด (Maximal Matching) โดยทดลองใช้กับพจนานุกรมภาษาไทยจากเว็บไซต์ Longdo เพื่อเปรียบเทียบและพัฒนาค้นคลังข้อมูลรายการคำประสมที่สกัดจากข่าวท่องเที่ยวภาษาไทยต่อไป



ภาคผนวก

จุฬาลงกรณ์มหาวิทยาลัย  
CHULALONGKORN UNIVERSITY



ภาคผนวก ก

รายการค่าประสมที่ได้จากข้าวทองเที่ยวไทย

ตารางที่ ก.1 รายการค่าประสม 2,340 ค่าจากข้าว 5,000 ข้าว

กองทัพเรือ	ข้าวดี	คนไทย	งาช้าง
กระดาดแก้ว	ข้าวตอก	คณะรัฐมนตรี	งวงช้าง
แก้วน้ำ	ขอบฟ้า	ค่าธรรมเนียม	เงินทอง
กำแพงแก้ว	ขบวนรถ	คุณภาพชีวิต	งานเทศกาล
ไก่นก	แขนขา	โครงการวิจัย	งานศิลปะ
เกรียบปากหม้อ	ชุมทรัพย์	เครื่องจักร	งานช่าง
เกลียวคลื่น	ขนมปัง	ค่ายทหาร	งานบุญ
ก้อนหิน	โชดหิน	คณะสถาปัตยกรรมศาสตร์	เงินปี
กลิ่นหอม	ไข่เจียว	คนเมือง	งานปี
กระบอกปิ่น	ข้าวแกง	คูเมือง	งานวิจัย
กองทัพ	ข้าวโลก	คณะกรรมการ	เงินบาท
กองทหาร	ขบวนแห่	คะแนนเสียง	งานหลัก
กระเป่าเสื้อ	ข้อเท็จจริง	คุ่มทุน	งานประจำ
กระแสน้ำ	ข้าวกล้อง	โครงสร้างพื้นฐาน	เงินสด
กองทุน	ข้อมูลข่าวสาร	โครงข่าย	งานเฉลิมฉลอง
กระบี่กระบอง	ขาแข้ง	คนงาน	เงินหมุนเวียน
กรงนก	ข้อดี	ค่าคืน	งานหนัก
กองข้าว	เขี้ยวเล็บ	ค่าเงิน	งานพิเศษ
กองทัพอากาศ	ข้อกฎหมาย	คณะกรรมการพัฒนาการเศรษฐกิจและสังคมแห่งชาติ	เงินทุน
ก้าชธรรมชาติ	เขตร้อน	ค่าตัว	งานวัด
กระแสนเงิน	ข้อสงสัย	เครื่องหมาย	งานพิธี
กองไฟ	ข้อห้าม	ค่านิยม	เงินสมทบ
กล้วยไม้	ขนมหวาน	คู่แข่ง	งานบ้าน
กำแพงเมือง	ข้าวเหนียว	เครื่องดนตรี	เงินอุดหนุน

กาบกล้วย	ข้อสรุป	เครื่องทอง	เงินสนับสนุน
กากบาท	ข้อจำกัด	คณะเศรษฐศาสตร์	เงินสำรอง
กระจกใส	ข่าวลือ	คู่มือ	งานฉลอง
กระดานโต้คลื่น	ข่าวสุก	คนใกล้ชิด	เงินผ่อน
แกนกลาง	ข้อหา	คนชั่ว	งานก่อสร้าง
กลิ่นอาย	ข้อยุติ	เคล็ดลับ	งานแต่งงาน
กลิ่นคาว	ขจร	เครื่องประดับ	เงินเดือน
กิ่งกลาง	ข้อบังคับ	คนธรรมดา	งานติดต่อ
กระเบื้องเคลือบ	ข้อกล่าวหา	เครื่องดื่ม	งานสัมพันธ์
กองเชียร์	ขีดจำกัด	ค่าเช่า	เงินประกัน
กรณีพิเศษ	ขี้เมา	เครื่องบิน	งานอุตสาหกรรม
กลางคืน	ขนมไทย	คู่แข่ง	งานฝีมือ
กินรวบ	ข้อซักถาม	คนดู	แง่ดี
ก๊าซ คาร์บอนไดออกไซด์	ข่าวผิด	คำรักษาพยาบาล	งานอดิเรก
กระเป๋าเดินทาง	ขี้เหล็ก	คาถา	เงินปลีก
กลิ่นไอ	ข้อสันนิษฐาน	เครื่องอุปโภค	เงินช่วยเหลือ
กองทัพบก	เข้มงวด	คนพิการ	งานหลวง
กระเบื้องเพชร	ขนมเค้ก	ค่าโดยสาร	งานศิลปกรรม
กรณีศึกษา	เขตเทศบาล	เครื่องสำอาง	งูพิษ
ไกทอด	ข้อเรียกร้อง	คูหา	งานประติมากรรม
กระแสไฟฟ้า	เขตสงวน	เครื่องจักสาน	เงินสะสม
กระดาดแข็ง	ข้ามคืน	ค่าตอบแทน	งานจักสาน
แกงมัสมั่น	ข้อแตกต่าง	คำรักษา	งานรื่นเริง
กิ่งก้าน	ไขไก่	เครื่องมือสื่อสาร	เงินก้อน
กระแสน้ำ	ขวานฟ้า	คนกลาง	งานค้นคว้า
กรอบประตู	เขตเมือง	เครื่องเงิน	เงินโอน
กรณีพิพาท	ขนมจีบ	คณะกรรมการ	งานสอน
แกงส้ม	ข่าวต้ม	เครื่องปรุง	วงการศิลปะ

กระป๋องน้ำ	ข้อสังเกต	คู่หู	วิทยาศาสตร์ ธรรมชาติ
กำแพงเพชร	เข็มกลัด	เครือข่าย	วัดถุมงคล
กั๊กหันลม	ข้อมือ	คนเก่ง	เวลาอาหาร
ไก่อ่าง	ข้าวบิณฑ์	โคมไฟ	วัดอุประสงค์
กระบอกฉีต	ข้าวเกรียบ	คนพื้นเมือง	วัดอุติบ
กระถางรูป	ชากรรไกร	คณะรัฐศาสตร์	เวลาว่าง
กระเช้าไฟฟ้า	ขอบตา	เครื่องไฟฟ้า	เวลาพักผ่อน
กิ่งไม้	แขกเมือง	ค่าแรง	วัวควาย
กระแสบรรพชาดำรัส	ข้าวหลาม	คนจีน	วิชาชีพ
กระแสดม	ไข่ต้ม	คนชรา	วิทยาเขต
กายกรรม	ขุมสมบัติ	คณะสงฆ์	เวลาเรียน
กัณฑ์เทศน์	ข้าวเย็น	เครื่องยศ	วัฒนธรรมแห่งชาติ
กล้าไม้	ขบวนรถไฟ	คณะท้าว	วัดอุระเบต
กระบอกเสียง	เขตป่า	คอกลม	วิทยุโทรทัศน์
กิจกรรม	ไข่ป่า	คนภูเขา	วัสดุก่อสร้าง
กิริยามารยาท	ไข่ทอด	คบไฟ	วันเสาร์
กลีบดอก	ข้อบกพร่อง	คอสพาน	วงจรชีวิต
กฎจรรยา	ข้าวเม่า	คู่พิพาท	แว่นสายตา
กล่องโฟม	ขนมขบเคี้ยว	คู่แข่งชั้น	เวรกรรม
เกียรติประวัติ	ไข่ปลา	คนรถ	แว่นตา
กระดาษหนังสือพิมพ์	ซีโต้	คลังสมอง	วงกลม
กระสุนปืน	ข้าวโพดคั่ว	ครูฝึก	วงการ
กุญแจมือ	ขบวนการ	ครัวเรือน	วงใน
กั้นปล่อง	ของสูง	ค่าบำรุง	วชิราวุธ
การท่องเที่ยวแห่ง ประเทศไทย	ของหลวง	เครื่องต้น	วิชาชีพ
กระเพาะปลา	ขอฉาย	ค่าเดินทาง	วัฒนธรรม
กล่องจุลทรรศน์	ขอบข่าย	ค่าเฉลี่ย	วิเศษนานุประโยค
แกงร้อน	ขอบเขต	ค่าตัว	แว่นแคว้น

กล้องโทรทัศน์	ช่างกรอง	คนญี่ปุ่น	เวลาเย็น
กินรี	ข้าวของ	เครื่องร่อน	เวทิมวย
กำไรสุทธิ	ข้าวเปลือก	คู่สมรส	ไวน์ขาว
ก้วยเตี่ยวผัด	ข้าวสาร	ค่าเรือ	วิทยาลัยครู
ก้านใบ	ข้าวสาร	เครื่องเซ่น	แวดล้อมแห่งชาติ
เกาเหลา	ข้าศึก	คนฟัง	วิทยุสื่อสาร
กระบวนแห่	ซีกบ	คะแนนเต็ม	อาหารสัตว์
ไก่อวง	ซิดคั่น	คนเหนือ	อนุสรณ์สถาน
กรอบรูป	ไซข้อ	คนโบราณ	แอ่งน้ำ
กระดาศสี่	ขนมจีน	เครื่องเคลือบ	อุ้งเท้า
กล้วยไข่	ไข่แดง	ค่าจ้าง	อาคารบ้านเรือน
ก้านขด	ข้อยกเว้น	คู่สัญญา	อรุณสวัสดิ์
กระดาศทอง	เขตชนบท	เครื่องมือแพทย์	เอกภาพ
กรรมการตัดสิน	ชุดแต่ง	คนหนุ่ม	โอกาสดี
การร้อง	เข้มทิศ	เครื่องสังเว	องค์กรเอกชน
กระดาศปอนด์	ข้อพิสูจน์	คำสอน	อาหารทะเล
กระจกส่อง	ข้อแนะนำ	โค้งสุดท้าย	อุทยานแห่งชาติ
กาน้ำ	ข้อเท้า	คู่บ่าวสาว	อาหารคาว
ก้ามปู	เขตแดน	เครื่องเคียง	อุตสาหกรรม ท่องเที่ยว
กงเวียน	ขนกี้ว	คณะดนตรี	องค์กรอิสระ
กงจักร	ข้าวปลุก	คดีอาญา	อายุยืน
กฎบัตร	ข้อต่อรอง	คันทนา	อาหารประจำ
กรรมฐาน	ขอนไม้	คณะกรรมการอาหาร และยา	อักษรย่อ
กรรมพันธุ์	ข้าวสวย	คณะกรรมการ วัฒนธรรมแห่งชาติ	อำนาจซื้อ
กรรมวิธี	ขนมครก	ค่าภาษี	อาสาสมัคร
กระโจมไฟ	ซี้เถ้า	คู่แต่งงาน	อาคารสำนักงาน
กระดานดำ	ขุนศึก	เครื่องสูบ	อาหารเช้า

กระต๊อ	ขนมชั้น	คอหนัง	องค์การขนส่งมวลชน กรุงเทพฯ
กระบวนกร	ข้าวเจ้า	เครื่องบูชา	อ้อมกอด
กริยานุเคราะห์	เขี้ยวตะขาบ	เครื่องราง	อาหารว่าง
กริยาวิเศษณานุ ประโยค	ขาม้า	เค้าเดิม	อาการป่วย
กวีนิพนธ์	ไขดาว	โครงเรื่อง	อาหารแห้ง
กสิกร	ชั้นหวาน	เครื่องมือ	อู่เรือ
กะหล่ำปลี	ข้อแลกเปลี่ยน	โครงร่าง	อาหารเย็น
กากหมู	ขั้นตอนวิธี	เครื่องบงบอก	อาหารหวาน
กาพย์กลอน	ข้าวท้องถิ่น	เครื่องจักรกล	องค์ประชุม
กายินทรีย์	ข้าวหมก	คชสาร	ไอคิว
กาเยนทรีย์	เข็มนาฬิกา	คณิตศาสตร์	อารมณ์ขัน
การเมือง	ข้อครหา	คฤหาสน์	อาหารค่ำ
กาลเทศะ	คลองส่งน้ำ	เครื่องเงิน	ไอน้ำ
กาลีกันน้ำ	คลื่นกระทบฝั่ง	เครื่องครัว	อาหารเสริม
กิจจานุกิจ	คลื่นใต้น้ำ	เครื่องคู่	อัตราดอกเบี้ย
เกลือแร่	คลื่นปานกลาง	คณวิทย์ศาสตร์	ไออุ่น
แก้แค้น	คลื่นยาว	คอพับ	เอกสารสิทธิ์
เกียรติคุณ	คลื่นลูกใหม่	เครื่องแก้ว	อุตสาหกรรมหนัก
กลืนตัว	คลื่นวิทยุ	คนแนะนำ	อัตราแลกเปลี่ยน
กึ่งแห่ง	คลื่นสั้น	ค่าขนส่ง	อาหารจีน
เกลียดปลา	ค่าน้ำนม	คนโดยสาร	อุตสาหกรรมเกษตร
เกือกม้า	คำปิดปาก	คู่ครอง	โอกาสหน้า
เก้าอี้หนัง	คำเสียหาย	คูปอง	อาจารย์พิเศษ
กรดไขมัน	คำขวัญ	คนกรุง	องค์กร
กอหญ้า	คำตาย	คันธนู	องค์การ
กระเป่าหิ้ว	คำแปล	เครื่องกิน	องค์กรกีฬา
กินดิบ	คำเสนอ	คำเรียน	อดีตจัน
กฏระเบียบ	คุณค่า	เครื่องสาน	อนุประโยค

กระดุกอ่อน	คุณธรรม	เครื่องปรุงรส	อบายมุข
กุศลจิต	คุณประโยชน์	คณรองคมนตรี	อรรถคดี
กรอบหน้า	คุณภาพ	เครื่องญาติ	อรรถประโยชน์
ไก่อ้วน	คุณลักษณะ	คู่ซี้	ออมสิน
กะหรี	คุณวุฒิ	คุณสมบัติพิเศษ	อักษรวิธี
แกงจืด	คุณศัพท์	คนป่วย	อัคคีภัย
แก้คัน	คุณสมบัติ	ครัวกลาง	อัญประกาศ
กฎกระทรวง	คุณานุประโยค	ครัวไฟ	อัญมณี
แกมเขียว	คู่กัด	คนสวย	อัญรูป
กระบวนการยุติธรรม	คู่ขา	คลื่นเสียง	อุณหภูมิต่ำ
กระดานหมากรุก	คู่ชีวิต	คติชน	อุตมคติ
กระเทียมหอม	เครื่องเครา	เครื่องบันทึก	อุตรดิตถ์
กระเป๋าสตางค์	เครื่องใช้	เครื่องแกง	อุทกภัย
แกงเลียง	เครื่องแบบ	คริสต์ศาสนา	อุทกวิทยา
ก้านดอก	เครื่องสาย	เครื่องว่าง	อุทกศาสตร์
กรามข้าง	เค้าเงื่อน	คุณแม่	อุบัติเหตุ
กรงขัง	โคบาล	คันไถ	เอกพจน์
กองเรือ	โครงสร้าง	คณวณศาสตร์	อุตสาหกรรมแห่ง ประเทศไทย
กองกลาง	คนตาย	เครื่องพิสูจน์	อาคารชุด
ก้านรูป	คณะ วิศวกรรมศาสตร์	คนโสด	ไอซี
กาแฟสด	คูน้า	ค่ามาตรฐาน	อาหารสด
กล้วยหอม	คนป่า	บ้านเรือน	อึ้งมือ
เกลือสินเธาว์	คอซอง	บ้านเมือง	อักษรศาสตร์
กิตติคุณ	เค้าโครง	โบราณวัตถุ	เอกสารอ้างอิง
กล่องดนตรี	เคราะห์กรรม	ใบแดง	อาการไอ
กระดานลื่น	คั้งน้ำ	ใบพัด	อกไก่
ก้างปลา	เค้าแมว	บ่าวสาว	อารามหลวง
ผังเมือง	ทิศเหนือ	โบราณคดี	อักษรเลข

แผงลอย	ท้องฟ้า	บ้านไร่	ประสบการณ์นิยม
ผักสด	ทุ่งนา	ใบไม้	เปลือกหอย
ผักกระเฉด	ท้องทะเล	บทกวี	ประตุน้ำ
ผักตบชวา	ทรัพยากรธรรมชาติ	ใบยา	ปุ๋ยคอก
ผลิตภัณฑ์อุตสาหกรรม	ทะเลสาบ	บุคลิกภาพ	ประชาชนลาว
ผ้าไตร	เทศบาลตำบล	ใบสมัคร	ปากคลอง
ผ้าจีน	ทิวแถว	บ้านพัก	ป่าเบญจพรรณ
ผักชี	ท่าเรือ	ใบอนุญาต	ป่าละเมาะ
เผ่าพันธุ์	ท่อนซุง	บทเรียน	เปลือกไม้
ผมปลอม	ทะเลทราย	เปื้อน	ปลายเดือน
ผ้าแดง	ทัพเรือ	บ้านเช่า	ประธานกรรมการ
ผลิตภัณฑ์	ท่อนไม้	โบราณกาล	ปลาหมึก
ผลิตภัณฑ์	ทรายขาว	โบราณสถาน	แปลงผัก
ผ้าขนหนู	ทุ่งหญ้า	บทวิเคราะห์	ข้อปราคาการ
ผ้าเช็ดตัว	ทหารเรือ	ใบปลิว	ประวัติการณ์
ผู้ผลิต	ทหารช่าง	บุรีรัมย์	ปัจจัยพื้นฐาน
ผู้สืบสันดาน	ท่าอากาศยาน	บุตรหลาน	ปูนซีเมนต์
ผู้อำนวยเพลง	ทานตะวัน	บัญชีชื่อ	ปัญหาจราจร
ผงซुरส	ท่าอากาศยานแห่ง ประเทศไทย	ใบताल	ป่าดง
ผ้าถุง	ทรัพยากรธรณี	บทสรุป	ปลาไหล
ผักบุ้ง	ไทยธรรม	ใบระกา	ปรากฏการณ์ ธรรมชาติ
ผักสลัด	ทองสัมฤทธิ์	บัญชีออมทรัพย์	ประเทศไทย
ผิวเมีย	ทัพอากาศ	บาทหลวง	ประเทศญี่ปุ่น
ผ้าลาย	ทีมงาน	ใบประกาศ	ประเทศสมาชิก
ฤดูฝน	ท้องถนน	บทละคร	ปากเสียง
ฤดูแล้ง	ทะเบียนสมรส	ใบขับขี	ประสาทสัมผัส
ฤดูหนาว	เที่ยวบิน	บัญญัติไตรยางค์	ปัจจัยหลัก
ฤดูร้อน	ทุนหมุนเวียน	บัวลอย	ป่าดิบ

ฤกษ์ดี	ทอหทัยใจ	บาทวิถี	ปีกลาย
ฤดูใบไม้ร่วง	โทษหนัก	บ้านเกิดเมืองนอน	ปูนปั้น
ฤกษ์เปลี่ยนแปลง	ทหารติดตาม	บุญคุณ	แปดสิบ
ลำน้ำ	ทาสี	บุพเพสันนิวาส	ปีมะเส็ง
ลำธาร	ไทยเดิม	เบาหวาน	ประเทศ สหรัฐอเมริกา
ลมแรง	ทุนทรัพย์	เบี้ยกัณฑ์	ปีวอก
ลำคลอง	ท้องสนามหลวง	เบี้ยทำขวัญ	ปีเถาะ
ลำต้น	เที่ยงคืน	เบี้ยบน	ประสาทตา
ลูกตา	เทียนชัย	เบี้ยล่าง	ปริญญาเอก
ลำตัว	เทียนไข	เบี้ยเลี้ยง	ปัญหาลพิษ
เลขท้าย	ห้องทุ่ง	แบบอย่าง	ป่าฝน
เลือดเนื้อเชื้อไข	เทียนแพ	บุญบารมี	ปู่ย่า
ลูกหลาน	ห้องน้ำ	ใบจอบ	ปีนัง
ลายเส้น	ห้องแบน	บานชื่น	ปลอกหมอน
ลูกท้าว	ทะเลปานกลาง	บ้านสวน	เปิดพะโล้
โลหะปราสาท	ทานบารมี	ใบชา	ประเทศจีน
ลมพายุ	ทะเลหลวง	บุหรี	ปีชาล
ลำยอง	เทียนพรรษา	บายโมง	แป้งสด
ลายมือ	ทรายแก้ว	ใบเมี่ยง	ปลาส้ม
ลูกสาว	ทุนลอย	บัญชีดำ	ปูม่า
ลูกไฟ	ท้ายเขื่อน	บายคล้อย	โปรแกรมควบคุม
ลมแล้ง	ทศกัณฐ์	บ้านเดิม	ป่าสงวน
ลมมรสุม	ทหารอากาศ	บุตรสาว	ประเทศชาติ
ลมหนาว	ทุ่งโล่ง	บ้านจัดสรร	ปู่ยี่หวภาพ
ลักษณะเด่น	ทองเปลว	บ้านช่อง	ปูชนียบุคคล
โลหะหนัก	ทองธรรมชาติ	ใบบุญ	ปากแม่น้ำ
ลมหายใจ	ทะเลบ้าน	ใบเบิก	ปริญญาตรี
ละครร้อง	ทองใบ	นายกรัฐมนตรี	ปีกุน



ลัทธิคอมมิวนิสต์	เที่ยวแห่งประเทศไทย	น้ำตาล	ปีแสง
เลขจำนวน	ห้องอืด	นมหนู	เปียโน
ลูกกรง	ท้ายเรือ	น้ำนม	ปัญหาสังคม
ลูกอ้นทะ	ท่าเทียบเรือ	น้ำชา	เป้าประสงค์
ลูกบ้าน	ทับเกษตร	น้ำฝน	ประเทศอินเดีย
ลำตัด	ท่าน้ำ	น้ำทะเล	ปากน้ำ
ลวดหนาม	ทวีปเอเชีย	นาสวน	ป้ายประกาศ
ลูกครึ่ง	ทุนเดิม	นามสกุล	ประตูด้าย
ลานวัด	ทหารพราน	น่านฟ้า	ปากท้อง
ลับแล	ทองหยอด	นายกเทศมนตรี	ปืนลม
ลูกกวาด	ไทยทาน	น้ำตาลทราย	ปากอ่าว
ละครลิง	ทุนรัฐบาล	น้ำปลา	ปลากัด
ลายเซ็น	ทรงกลม	น้ำดื่ม	ป้อมตำรวจ
เลขหมาย	ทรพีช	น้ำจืด	ปลาช่อน
ลิงอุรังอุตัง	ทรวดทรง	แนวตั้ง	ปลาตะเพียน
ลำไส้	ทรัพย์สินสมบัติ	น้ำเค็ม	ปืนพก
ละครแห่งชาติ	ทรัพย์สิน	น้ำพุ	ประเพณีนิยม
เล็บมือ	ทรายแป้ง	น้ำตก	ปลัดกระทรวง
ล้อถีบ	ทศพร	น้ำค้าง	ประจำหลัก
ลมปราณ	ทหารเกณฑ์	น้ำหลาก	ประตูเมือง
ลูกศิษย์	ทอดมัน	น้ำท่วม	เปลือกโลก
ลูกเรือ	ทัศนสถาน	น้ำตาย	ป้อมยาม
ลูกเต๋า	ทางการ	น้ำสะอาด	ประเภทลาน
ลูกน้ำ	ทางเท้า	น้ำหน้า	ประวัติศาสตร์
ลิงแสม	ทิ้งกระจาด	น้ำเปล่า	ปากหม้อ
ลานบ้าน	ทัศนุทิศ	นมข้น	ประเทศอุตสาหกรรม
ลำแสง	พืชชาติ	น้ำแข็ง	ปฏิกริยา
ละครรำ	พืชสระ	น้ำสังข์	ปรปักษ์
ลูกชิ้น	ทุนนิยม	นาเมือง	ปริมาณ

ลหุโทษ	ทูตานุทูต	น้ำป่า	ประชากร
ลายลักษณ์อักษร	เทศบาลนคร	น้ำเกลือ	ประชาทัณฑ์
ลูกขนไก่	เท้าคู่	น้ำร้อน	ประเทศราช
ลูกเขย	เทียนอบ	นครพนม	ประภาคาร
ลูกค้า	โทษานุโทษ	นกกระจอก	ประสบการณ์
ลูกจ้าง	ห้องหมู	น้ำท่า	ปรากฏการณ์
ลูกบาศก์เซนติเมตร	ทองคำเปลว	น้ำลาย	ปริมณฑล
ลูกพี่ลูกน้อง	โทรพิมพ์	น้ำตา	ปริมาตร
เล่ห์เหลี่ยม	ทับทิม	น้ำยา	ปลาเห็ด
เลือดเนื้อ	ท่อพลาสติก	น้ำกะทิ	ปากกา
ลูกดอก	ทองเอก	น่านน้ำ	ปาฐกถา
ลายไทย	เท้าแขน	นาเกลือ	ปูนดิบ
ลมพิษ	ทวีปอเมริกาเหนือ	น้ำเย็น	ปูนแดง
ลูกระเบิด	ทวารหนัก	น้ำมันก๊าด	ปูนสุก
เลขคี	เท้าหน้า	น้องชาย	แป้งเหล้า
ล้อรถ	ทะเลเพลิง	น้ำจิ้ม	แปดเหลี่ยม
ลูกค้าประจำ	ไถยยาม	น้องสาว	ปูนขาว
ละครสัตว์	ยามเช้า	นกแก้ว	ปลอกคอ
ลำปาง	ยอดไม้	แนวเขต	ปลาร้า
โลกหน้า	เยื่อกระดาษ	น้ำประปา	ปูนเปียก
โลหะผสม	ยามค่ำคืน	น้ำมันรถ	ปลายจมูก
ลักษณะดิน	ยุคเก่า	น้ำพระพุทธรมณต์	ป่าไม้
ลำดับ	ยอดจำหน่าย	น้ำพริก	ปัญญาประดิษฐ์
ลายแทง	ยอดขาย	น้ำดิบ	เปลวไฟ
ลูกทิม	ใยสังเคราะห์	น้ำดี	ปิงหา
ลำพูน	ยาทา	น้ำเชื้อ	ปลาแนม
ลายฉลุ	ยวดยาน	น้ำทิพย์	ปล่องไฟ
ลูกไม้	ยัญพิธี	นกหวีด	ประเพณีพื้นบ้าน
ละครทีวี	ยากันยุง	นครหลวง	ปลาเค็ม
ลมทะเล	ยาพิษ	นพรัตน์	เปรี้ยวหวาน

ลิ้นชัก	ยาสูบ	นมผง	ปัจจัยเสี่ยง
จานรอง	ยาฉีต	จุดเด่น	ป้ายชื่อ
จุดกำเนิด	ยาสมุนไพรร	จุดเริ่มต้น	ป้ายราคา
จอมทัพ	ยุงลาย	จุดสิ้นสุด	ประกันสังคม
จุดศูนย์กลาง	ย่าค่า	งานบิน	ปลายลื่น
จอภาพ	เจ้าพ่อหลักเมือง	จุดขาย	ปลาหว
จิตสำนึก	เจ้าบ่าว	จุดประสงค์	ปีชวด
เจ้าหน้าที่ตำรวจ	จอเงิน	จุดผ่อนปรน	ป่าชุมชน
เจ้าของงาน	จำปา	จุดแข็ง	เปลวเพลิง
เจ้าของบ้าน	จักรกล	จุดหลัก	ปีกซ้าย
เจ้าของที่ดิน	จตุสดมภ์	เจ้าของร้าน	ป่าทึบ
จุดหมาย	จรรยาบรรณ	เจ้าวัด	ปีระกา
เจ้าของกิจการ	จักรยาน	จุดวิกฤติ	ปีจอ
เจ็ดสิบ	จานผี	เจ้าพ่อ	ปุ๋ยเคมี
จิตวิทยา	จำนวนนับ	จำนวนคู่	ศาลาท่าน้ำ
จุฬามาศ	จิตใจ	จุดยุทธศาสตร์	หน่วยงาน
จุดยืน	ศิลปะพื้นบ้าน	จำนวนเต็ม	หาดทราย
เจ็ดเหลี่ยม	ศาสนาพราหมณ์	จุดอ่อน	หมู่เกาะ
เจ้ากี้เจ้าการ	ศูนย์กลาง	เจ้าคณะ	หน้าตัก
เจ้าบ้าน	ศิษย์เก่า	จุดดี	หน้าฝน
ใจดำ	ศีลธรรม	หมอลำ	หินปูน
เงินชวบน้ำ	ศาลาประชาคม	ห้องสมุด	ห้างสรรพสินค้า
เจ้าของภาษา	เศษสตางค์	ไหมพรม	หมู่บ้าน
เจ้าถิ่น	ศาสนาคริสต์	ห้องโถง	หลักธรรม
งานชาม	ศตวรรษ	แหล่งข่าว	หน้าแข็ง
เจ้าหน้า	ศัลยกรรม	หน้าบ้าน	ห้องแถว
จุดสัมผัส	ศาลเตี้ย	แหล่งน้ำ	หีบศพ
เจ้าแม่	ศาลทหาร	หมู่คณะ	หนังตะลุง
จอแก้ว	ศาลาวัด	แหล่งอารยธรรม	หนุ่มสาว
จอแบน	ศิลปกรรม	หน้ากาก	หลักเมือง

ศาสนาพุทธ	เศษส่วน	หลักแหล่ง	เหรียญทอง
ศาลเจ้า	ศัพท์บัญญัติ	หลักวิชาการ	หมาป่า
ศึกสงคราม	ศาลพระภูมิ	หน้าจ่อ	หมวดหมู่
เศรษฐกิจอุตสาหกรรม	หอนาฬิกา	แหล่งท่องเที่ยว	หนองน้ำ
ศาลากลาง	หลักประกัน	หนังสือเดินทาง	หน่วยราชการ
ศิลาจารึก	หินย้อย	หนังสือเรียน	หม้อดิน
ทางเครื่อง	หาบเร่	ห้องพัก	ห้องน้ำ
ห้างร้าน	หน้าตา	แหล่งผลิต	หญิงสาว
หมู่บูชา	หนังสือคู่มือ	หน้าหนาว	หน่วยงานราชการ
หมวย่าง	หินอ่อน	หน้าร้อน	เหล็กไหล
หมายเลข	หน่อไม้	ห้องนอน	เหล็กตัด
หนังสือแนะนำ	ห้วยหนอง	หุ่นจำลอง	เหล็กหล่อ
หมอนขวาน	ห้วงเวลา	ห้องเรียน	หน้าเตา
ห้องส้วม	แหล่งกำเนิด	ห้องประชุม	ห้องครัว
ห้องเครื่อง	หุ่นกระบอก	หลุมดำ	เหล้าไวน์
หลักยึด	หोजดหมายเหตุ	ห่อหมก	หินทราย
หลอดไฟ	หน่วยงานรัฐ	ห้องอาหาร	หมูแดง
แหล่งข้อมูล	หน่วยเลือกตั้ง	เหมืองแร่	หนังสือพิมพ์
หนามเตย	หินงอก	หวานเย็น	ห้องชุด
หลักหน่วย	แหล่งเสื่อมโทรม	หมูป่า	เหล็กเส้น
แหล่งพัก	หุ่นเชิด	หลักเขต	หมายประกาศ
หลักกฎหมาย	หกเหลี่ยม	หุบเหว	หูกู้
หน้าไม้	หนทาง	ห้องทดลอง	ไหมทอง
ทางเลข	หนังสือ	หลักสูตร	หลักพระพุทธศาสนา
หมอนอิง	หน้าม้า	หน้าอก	หมูแฮม
ทางกระรอก	หนังสือ	หนังสือราชการ	หน้าแว่น
หอรขัง	หลวงพ่อ	หน้าท้อง	หมูสับ
หูกวาง	หลอดน็อน	หม้อน้ำ	หูฟัง
หลุมศพ	หลักทรัพย์	หน้าจั่ว	หน้าแล้ง
ห้องปฏิบัติการ	ห้องทับ	หลานชาย	หมอดู

ไหลทวิป	หอบหืด	หัวปลี	ทางเสื่อ
หน้าปก	หัวเทียน	หัวแร้ง	หุ่นกู่
หม่อมหลวง	หนังยาง	แหล่งพลังงาน	เหตุการณ์
ชายหาด	โคมหน้า	ตอนเหนือ	ถ้วยชาม
ชายแดน	ชากดึกดำบรรพ์	ต้นไม้	ถุงพลาสติก
ชั้นดิน	ซ้ายมือ	ตัวเมีย	ถ้วยรางวัล
ช่างฝีมือ	ชากศพ	ตัวเมีย	ถิ่นทุรกันดาร
ช่างกลึง	ชากสัตว์	ตาราง	ถ้วยแกลง
ชนชาติ	ช่องประมุล	ตู้โทรศัพท์	ถุงทอง
ชายทะเล	ซีดีรอม	ต้นไม้	ถิ่นกำเนิด
ชานอ้อย	ซีกโลก	ตู้กระจก	ถุงมือ
ชีวิตจิตใจ	ซอด้วง	ตอนต้น	ถ้วยระ
ชั้นเอก	เณรหน้าไฟ	ต้นตำรับ	ถ้วยตวง
เชิงเทิน	เณรทางนาค	ตอนท้าย	ถ้วยคำ
เชื้อเพลิง	ดอกไม้	ตอนปลาย	ถ้วยอก
ชุดราตรี	ดวงอาทิตย์	ต้นมะขาม	ถ่านไฟเก่า
ชื่อเสียง	ดอกเห็ด	ตลาดน้ำ	ถ่านหิน
ชั้นปี	เด็กหนุ่ม	ตลาดโลก	เถรานุเถระ
เข้าตรู่	ดอกเบ็ญ	ตู้เย็น	ถ้วยเขียว
ชนพื้นเมือง	ดวงวิญญาณ	ตัวเลข	โถส้วม
ช่องแคบ	เด็กกำพร้า	ต้นทุ่น	ถนนหนทาง
ชั้นล่าง	ดินเค็ม	ตอนบ่าย	ถมเงิน
ชั้นสัมพัทธ์	ดาวเด่น	ตลาดนัด	ญาติสนิท
ข่อม่วง	ดินเผา	ตัวดี	ฐานทัพ
ชวนชม	เดือนเต็ม	ต้นเดือน	ฐานข้อมูล
ชื่อเรื่อง	ดาวเทียม	ตอนเที่ยง	ร้านอาหาร
ชิงช้าสวรรค์	ดอกกุหลาบ	ตอนเช้า	ริมแม่น้ำ
ชัยชนะ	ดวงแก้ว	ต้นไทร	แร่ธาตุ
ชื่อสกุล	ดวงกลม	ตอนเย็น	ริมทะเล
ชนไก่	ดวงตา	ตัวหลัก	โรงแรม

ช่องทางสื่อสาร	ดาวดิงส์	ตีนกา	รถม้า
ชั้นเรียน	ดินสอพอง	ต้มยำ	เรือนแพ
ชิ้นงาน	ดวงประทีป	ตลาดหลักทรัพย์	ระดับน้ำ
ชุมชนเมือง	ดีเปรสชัน	ตอนกลาง	รถราง
ชั้นลอย	ดาวเรือง	ตะวันตก	ร้านกาแฟ
ช่องลม	ดอกฟ้า	ดีกระฟ้า	ริมน้ำ
ชั้นดาดฟ้า	ดวงชะตา	ตัวโดยสาร	เรือพาย
ชั้นครู	เด็กชาย	ตลาดสด	เรือนกระจก
ชายหนุ่ม	ดาวเหนือ	ตัวกระตุ้น	รถกระบะ
ชาตินิยม	เด็กหญิง	ตัวเลือก	โรงงาน
ชั้นเอก	ด้ายดิบ	ตัวกลาง	โรคตับ
ชนชั้น	ดินแดน	ตัวย่อ	ริมถนน
ชกกะเย่อ	ดินเหนียว	ตุ้ย	โรงละคร
ช่องว่าง	ดินปืน	ตัวอ่อน	เรือไฟ
ชั้นมัธยมศึกษา	ดวงเดือน	ตรุษจีน	โรงพยาบาล
ชื่อตัว	ดวงจันทร์	ตัวหนังสือ	โรคมะเร็ง
ชื่อย่อ	ดีแดง	ตู้ปลา	โรงหนัง
ช่างภาพ	ดีบุก	ตราไปรษณียากร	รองประธาน
ช่างน้ำ	เด็กแดง	โต้วาที	ไร่นา
ชั้นต้น	ดอกดิน	ต้นเงิน	เรือประมง
เขलयศิก	เด็กอ่อน	ตัวอักษร	รูปภาพ
ช่องคลอด	ดินขาว	ตัวไม้	รถจักรยาน
ชายป่า	ดาวอังคาร	ตัวปัญหา	รายการวิทยุ
ชานเมือง	ดาวเคราะห์	ต้นกล้วย	รองอธิบดี
ชานชาลา	ดินดำ	ตัวแปร	ระดับชาติ
ชานบ้าน	สีน้ำตาล	เตาแก๊ส	รายการโทรทัศน์
ช่องทางเดิน	สีน้ำ	ตอนค่ำ	รูปคน
ชั้นเชิง	สัตว์ป่า	ตัวเร่ง	เรือพระที่นั่ง
ชนแห่งประเทศไทย	แสงไฟ	ไต่ปลา	รถทัวร์
เชิงกราน	แสงแดด	ตำรวจหลวง	ระบบดิจิทัล

ซีตัน	ไต้เดือน	ตันเรือ	โรงเรียน
ชนบท	แสงอาทิตย์	ต้มส้ม	ร้านค้า
ชนมพรรษา	สีแดง	ตะเกียงเจ้าพายุ	รูปปั้น
ชลประทาน	สระน้ำ	ตัวละคร	แรงจูงใจ
ช่องทาง	แสงสี	ตาเปล่า	รอยยิ้ม
ข้อหา	สีตัวน้ำ	ตัวสื่อ	โรคระบาด
ข้อโต้แย้ง	สภาพอากาศ	ต้นกล้า	รถโดยสาร
ชะวากทะเล	เสาหลัก	ตัวเอก	เรือจ้าง
ชักกระปี่สีท่า	สวนสัตว์	ตารางเวลา	แรงกระตุ้น
ชายเฟื่อง	สวนพฤกษศาสตร์	เตาเผา	เรื่องเล่า
ชานา	สวนยาง	ต้นตาล	แรงบันดาลใจ
ชีววิทยา	สงครามโลก	เตาถ่าน	รูปหล่อ
ชุมนุมชน	สินค้าเกษตร	ตกค้า	รถเช่า
ชุมพล	สำนักงาน	ตาแดง	เรื่องอุกเหิน
เชิงชั้น	สื่อสิ่งพิมพ์	ตู้ไปรษณีย์	โรงกลั่น
เชียงใหม่	สถาบันวิจัย	ต้มข่า	แรงผลักดัน
ชาเย็น	สินค้าหลัก	ตัวสำรอง	ราคาขาย
ชั้นพิเศษ	สวนสาธารณะ	ไต้ฝุ่น	แรงงาน
ข้อฟ้า	สวนน้ำ	ตู้เอทีเอ็ม	ระเบียบวินัย
ชั้นตรี	สนามกีฬา	ตำราเรียน	ระดับทะเล
เชิงชาย	สวัสดิการสังคม	ตุ๊กตุ๊ก	รูปกรวย
ชาติตระกูล	สิทธิประโยชน์	ตำรวจนครบาล	รสชาติ
ชายผ้า	แสงเดือน	ตราสินค้า	รถแข่ง
เชื้อชาติ	สภาพเศรษฐกิจ	ตำรวจน้ำ	รูปโฉม
ห้องจรรยา	สระบุรี	ต้นร้ายปลายดี	รังนก
ซีลาว	เสาไฟฟ้า	ตราสาร	รูปร่างหน้าตา
ห้องสัญญาณ	สังคมเศรษฐกิจ	ตลาดมืด	เรือใบ
ช่างปั้น	สถาบันเทคโนโลยี ราชวมงคล	ตะวันตกเฉียงใต้	ร้อยแก้ว
ชั้นอนุบาล	สนามหลวง	ตะวันตกเฉียงเหนือ	ร้อยเอ็ด

ชุมชนแออัด	สุขภาพอนามัย	ตะวันออกกลาง	รถตู้
ช่างหล่อ	สำนักข่าว	ตะวันออกเฉียงใต้	ร้านรวง
ฝาผนัง	สถานที่ท่องเที่ยว	ตะวันออกเฉียงเหนือ	โรงไฟฟ้า
ฝาเท้า	เสื่อยืด	ดับเพลิง	รอยขีด
ฝ่ามือ	สะพานแขวน	ตัวคุณร่วมน้อย	ร้อยตรี
ฝาแฝด	สภาพแวดล้อม	ตัวหารร่วมมาก	รถเข็น
ฝนตก	สาวงาม	ตาข่าย	ระดับชั้น
ฝุ่นละออง	เส้นโค้ง	ตาขาว	โรงงาน
ฝีมือแรงงาน	สูตรสำเร็จ	ตารางสอน	รอยตีน
ฝนฟ้า	สถานที่ศักดิ์สิทธิ์	ตำรวจภูธร	รถไฟ
ฝันดี	เสือโคร่ง	โต๊ะจีน	รถแท็กซี่
ฝูงบิน	สิทธิพิเศษ	เตาไฟ	ระบอบ ประชาธิปไตย
ฝนหลวง	สังคมศึกษา	ตัวอาคาร	เรือสำเภา
ฝากระดาน	สัมตำ	ตำรวจเทศกิจ	รูปหน้า
ฝีมือ	สินค้าพื้นเมือง	ตึกแถว	รถพ่วง
ฝูงชน	สินไหมทดแทน	ต้นแขน	รูปถ่าย
ฝาเรือน	สังคมแห่งชาติ	ตัวควบคุม	รูปแกะสลัก
ฝนแล้ง	สภาวัฒนธรรม แห่งชาติ	เตียงนอน	ระเบิดเวลา
ฝุ่นผง	สินค้าฟุ่มเฟือย	ตู้หนังสือ	รอยเท้า
ธงชาติ	สารุชน	เต้านม	รถจักร
ธนาคารแห่งประเทศไทย	สถานีอนามัย	ตัวพระ	โรงเรียนเทศบาล
ธรรมชาติมนุษย์	สีย้อม	ตัวเต็ง	ระนาดเอก
ธนาคารพาณิชย์	สวนป่า	ตะบองเพชร	รถพยาบาล
ธุรกิจท่องเที่ยว	สื่อผสม	ตราประทับ	รวงข้าว
ธนาคารออมสิน	ไส้กรอก	ตัวรถไฟ	รองคนบตี
ธาตุอาหาร	เส้นชัย	ต้นอ่อน	แรงต้านทาน
ธรรมเนียม	สัตว์โลก	ต้นน้ำ	ร่องน้ำ



อุดมคสมาทาน	สนามม้า	ตลาดแรงงาน	รูปร่าง
ธุรกิจ	สภาวะเศรษฐกิจ	ต้นเรื่อง	โรคเรื้อน
ธรรมกาย	สถาบัน พระมหากษัตริย์	ตาตำ	รางน้ำ
ธรรมจักร	แสงจันทร์	ต้มเครื่อง	รถกระเช้า
ธรรมชาตินิยม	สถานศึกษา	ต้นเสียง	รถลาก
ธนาคารโลก	เสียงหัวเราะ	ตำรวจจราจร	แรงดัน
ธรรมสังเวช	สถานีวิทยุ	ต้นขา	รูปพรรณ
ธรรมนิยมประเพณี	แสงสว่าง	ต้นตอ	เรือนร่าง
พันธุ์ไม้	สำนักงานสถิติ แห่งชาติ	เตาอบ	ริมตลิ่ง
เพื่อนมนุษย์	สัญญาจ้าง	ต้นเพลิง	ระลอกคลื่น
พื้นดิน	เสียงดนตรี	ตัวยืน	รสเปรี้ยว
พื้นถิ่น	สังคมนิยม	ตุ้มหู	โรงอาหาร
พืชไร่	สถาบันเทิง	ตัวขีด	ราชการประจำ
พลบค่ำ	แสงตะวัน	เตาไฟฟ้า	ระเบิดทำลาย
พืชพรรณ	สถาบันครอบครัว	ตาชั่ง	ระบบขับถ่าย
พรรณไม้	สนุกสนานร่าเริง	ตลาดหุ้น	รางวัลชมเชย
พงไพร	สถานีขนส่ง	ตอนจบ	รอยสัก
พดลม	สื่อโฆษณา	ฝั่งเมือง	ราชสำนัก
พระศาสนา	สถานีรถไฟ	แผงลอย	รูปแบบมาตรฐาน
พาณิชย์นาวี	สร้อยคอ	ผักสด	โรงรถ
เพศชาย	สาวไทย	ผักกระเฉด	ร้อยเอก
พืชสวน	สถานีโทรทัศน์	ผักตบชวา	รองเง็ง
พระที่นั่ง	สัตว์ปีก	ผลิตภัณฑ์อุตสาหกรรม	แรงกดดัน
พันธุ์พืช	สะพานไม้	ผ้าไตร	ระบบประสาท
พระภิกษุสงฆ์	สีไฟ	ผ้าชิ้น	ราคาทุน
พ่อเมือง	สปู่เหลว	ผักชี	โรงสี
พ่อครัว	สภาพสังคม	เผ่าพันธุ์	รถถัง
พระอารามหลวง	เส้นเลือด	ผมปลอม	รถยนต์

พันเอก	เสียงฮา	ผ้าแดง	รหัสผ่าน
พระอาทิตย์ตกดิน	สวนหย่อม	ผลลัพท์	รักสามเส้า
พิธีแต่งงาน	เสนาธิการทหาร	ผลิตภัณฑ์	รัชกาล
เพื่อนสนิท	สะพานลอย	ผ้าขนหนู	รัชทายาท
พื้นที่ราบ	สภาพคล่อง	ผ้าเช็ดตัว	รัฐมนตรี
เพลงใหม่	สำนักงบประมาณ	ผู้ผลิต	รัฐวิสาหกิจ
พัดโบก	สนามกอล์ฟ	ผู้สืบสันดาน	รัสเซีย
พันจ่า	สัมพัทธ์	ผู้อำนวยการเพลง	รากดิน
พ่อแม่	สินค้าอุตสาหกรรม	ผงชูรส	ร่างกาย
พลังงานไฟฟ้า	สนามหญ้า	ผ้าถุง	ราชการ
พื้นคอนกรีต	สื่อโทรทัศน์	ผักบุง	ราชประเพณี
เพื่อนเดินทาง	สีหน้า	ผักสลัด	ราชพิธี
พืชเศรษฐกิจ	สถานีตำรวจ	ผิวเมีย	ราชโยกการ
พื้นราบ	สหภาพแรงงาน	ผ้าลาย	ราชวงศ์
เพลงศพ	สวนสนุก	ปากฟ้า	ราชโองการ
พิษงู	เส้นขนาน	ฟักทอง	ราชาภิเษก
เพศหญิง	สาขาวิชา	ไฟเขียว	ราชินูปถัมภ์
พายุฝน	เสียงอ่าน	ฟ้าร้อง	ราชูปถัมภ์
พืชผัก	สื่อกลาง	ฟ้าหลัว	ราชโฆปถัมภ์
พวงกุญแจ	เสียงร้อง	ฟ้าคะนอง	ราชโฆวาท
พุทธธรรม	สวนผัก	ฟ้าลั่น	เรือนตะเกียง
พวงมาลัย	สถานที่ราชการ	ไฟฉาย	โรคจิต
พี่สาว	แสงสะท้อน	ไฟถนน	โรงรับจำนำ
พ่อแม่	สถานที่ก่อสร้าง	ฟ้าสาง	เรียนสาธิต
พดยศ	สิงโตทะเล	แฟ้มข้อมูล	ระดับปานกลาง
พื้นน้ำ	สภาพทั่วไป	ไฟแดง	รสเค็ม
เพื่อนเจ้าบ่าว	เสบียงอาหาร	ฟ้าผ่า	โรงหล่อ
พิธีสมรส	เสียงแห่งประเทศไทย	ไฟเบอร์	แรงลม
พระรูป	สำมะโนประชากร	ฟองสบู่	รังไข่

พี่น้อง	สัญญาเช่า	พินแท้	รูปวาด
พฤติกรรมนิยม	สวนครัว	พินฟาง	โรงพัก
พ่อพันธุ์	สุพรรณบุรี	ไฟฟ้า	รองคุณภาพ
พังกา	สุขอนามัย	ไฟฟ้า	ร่างกาย
พินทราย	เสียงสวด	ภาวะเศรษฐกิจ	โรคติดต่อ
พวงมาลา	สมาชิกวุฒิสภา	ภัยธรรมชาติ	รอยแตก
พระเจ้า	สิ่งแวดล้อม	ภาวะวิกฤติ	รถมอเตอร์ไซค์
พื้นโลก	เสียงสะท้อน	ภาวะสงคราม	เรือบิน
พายุกมุน	ลิบเอ็ด	ภูเขาไฟ	รถไฟ
พจนานุกรม	เสากระโดง	ภัยพิบัติ	รถไฟฟ้า
พยาธิวิทยา	สำนักงาน คณะกรรมการ วัฒนธรรมแห่งชาติ	ภาพพจน์	ราชการลับ
พระพุทธรูป	สมาชิกสภาผู้แทน ราษฎร	ภัยแล้ง	แรงขับ
พฤติการณ์	สวนพฤกษชาติ	ภาคเอกชน	ระบบทุนนิยม
พลเมือง	สถานที่บันเทิง	ภาพวาด	ราชการทหาร
พสกนิกร	สถานการณ์	ภูมิปัญญาพื้นบ้าน	รถไฟฟ้ามหานคร
พ่อค้า	สถานที่	ภาพเหมือน	รถเก๋ง
พ่อค้าคนกลาง	สมณศักดิ์	ภาพถ่าย	ระวางโทษ
พาหุรัด	สมาชิกภาพ	ภาษาไทย	โรงนา
พุทธศาสนิกชน	สรรพนาม	ภาษาท้องถิ่น	รายการอาหาร
เพลงชาติ	ส่วนประกอบ	ภัยอันตราย	เรือแจว
เพศสัมพันธ์	สะพานไฟ	ภาษีอากร	โรงจำนำ
เพื่อนฝูง	สังฆกรรม	ภาวะฉุกเฉิน	รุ่นพี่
แพทยศาสตร์	สังฆทาน	ภาพพิมพ์	รถชน
ไพร่พล	สังฆนายก	ภาพนิ่ง	วางรถไฟ
พ่อตา	สังฆเภท	ภาวะการณ์	แม่ไม้
พินกระเบื้อง	สังฆราช	ภาพล้อ	มัชฌิมาปฏิปทา
พินหญ้า	สังฆวาส	ภาษาถิ่น	มวยไทย

พี่ชาย	สัญลักษณ์	ภาษาพื้นเมือง	มัธยมต้น
เพื่อนกิน	สัตว์เลี้ยง	ภาวะถดถอย	มือสี
พลังน้ำ	สันดอน	ภาษาโบราณ	ไม้ทะเล
พิมพ์ดีด	สาธารณชน	ภัตตาคาร	ไม้ซุง
พริกแดง	สาธารณรัฐ	ภัตตาคาร	เมืองอุตสาหกรรม
พุทธเจ้า	สาธารณรัฐประชาชน จีน	ภาคพื้น	เมล็ดข้าว
พื้นที่เพาะปลูก	สาธารณูปโภค	ภูเขา	มือซ้าย
พลังงาน	สามง่าม	ภูเขาน้ำแข็ง	แม่ครัว
พระนอน	สามเหลี่ยม	ภูมิศาสตร์	แมลงทับ
พระพุทธรูป	สารบรรณ	ภาษาจีน	ไม้หมอน
พ่อบ้าน	สาวไส้	ภาษาปาก	ไม้ตัด
พันธุ์ผสม	สำนักสงฆ์	ภาษาเครื่อง	เมืองเหนือ
พระอาทิตย์	สามะโนครัว	ภาษาที่ดิน	เมืองร้อน
พระเครื่อง	สิ่งก่อสร้าง	ภาพเคลื่อนไหว	มวยปล้ำ
พระพุทธรูป	สิ่งของ	ภาพสะท้อน	ไม้ต้น
เพื่อนซี้	สิ่งศักดิ์สิทธิ์	ไม้กระป๋อง	มนุษย์ต่างดาว
พลังจิต	สิทธิบัตร	เม็ดทราย	ม้าหมุน
พรมแดน	สินบน	เมฆหมอก	ไม้ขีด
พุ่มไม้	ลิบเหลี่ยม	มิงมิตร	ไม้ประดับ
นักปราชญ์	ลิวเสียน	มุกดาหาร	มวลอากาศ
นามานูประโยค	สี่เหลี่ยม	เมืองท่องเที่ยว	มกุฎราชกุมาร
นายท้าย	สี่เหลี่ยมขนมเปียก ปูน	มหาวิทยาลัยมหิดล	มนุษย์อวกาศ
น้ำกลั่น	สี่เหลี่ยมข้าวหลาม ตัด	ไม้ดอก	มโนธรรม
น้ำขาว	สี่เหลี่ยมจัตุรัส	มาตรการป้องกัน	มโนศิลา
น้ำซิม	สี่เหลี่ยมผืนผ้า	ไม้เต็ด	มวลสาร
น้ำประสานทอง	สุภาพบุรุษ	มิตรสหาย	มหाराช
น้ำผึ้ง	สุภาพสตรี	แม่พิมพ์	มะเขือเทศ

น้ำมัน	เสรีภาพ	เมารี	มาตรฐาน
เนื้อเยื่อ	เสียนศึก	ใส่ไก่	มือถือ
เนื้อเรื่อง	เสียนหนาม	สงครามศาสนา	มุขยประโยชน์
ในหลวง	เสือนอนกิน	สถาบันอุดมศึกษา	มุ่มฉาก
น้ำเงิน	เสื่อแสง	สัตยาติญาณ	เมืองนอก
น้ำมูก	เสาเข็ม	สำนักพิมพ์	แม่ค้า
นกฮูก	สัตว์ประหลาด	สุขภาพกาย	แม่น้ำ
นาดำ	เส้นก๋วยเตี๋ยว	สภาพภูมิอากาศ	แม่เปี้ย
น้ำแกง	สะพานปลา	สมุดบัญชี	ไม้บรรทัด
น้ำมันงา	สีชา	สาวรุ่น	ไม้หน้า
น้ำดื่ม	สถานที่สาธารณะ	เส้นด้าย	แม่ลาย
นมสด	สถานพยาบาล	สมอเรือ	มะพร้าวอ่อน
น้ำกิน	สถานที่พักผ่อน	สิทธิเสรีภาพ	มัดจำ
น้ำมือ	สีน้ำเงิน	สมองไหล	แม่ผิว
นาฬิกาแดด	สีน้ำมัน	ไม้กระดาน	มุขเด็ด
มวลดชน	สีฝุ่น	แมงกะพรุน	แมวป่า
เมืองหลวง	เสื่อดำ	ม้านั่ง	เมืองร้าง
มูลโค	สำนักพระราชวัง	มวยผม	มือดี
ม้าน้ำ	สัญลักษณ์	มวลมนุษย์	มุ่มถนน
แม่เหล็ก	สภาพร่างกาย	มือขวา	เสียงประกาศ
เมืองกรุง	สัตว์เลื้อยคลาน	ไม้ล้มลุก	เสื่อคลุม
ไม้ใบ	เสาธง	แม่ยาย	สวนเกษตร
ไม้เท้า	สตูดิโอ	ไม้ขอน	เสื่อฝน
มหาวิทยาลัยขอนแก่น	เสียงกึกก้อง	แม่ซื้อ	เส้นประสาท
มดแดง	มือปืน	เม็ดเงิน	เมล็ดพันธุ์

ภาคผนวก ข  
ชุดข้อมูลข่าวที่ใช้ในการทดลอง

ข้อมูลทดลองทั้ง 5 ชุด รวม 400 ข่าว สามารถเข้าถึงได้จาก

[https://drive.google.com/drive/u/0/folders/1wK1-gnk014nu3ODUyVOEN8J\\_COk2f5LO](https://drive.google.com/drive/u/0/folders/1wK1-gnk014nu3ODUyVOEN8J_COk2f5LO)

ข้อมูลข่าวที่ใช้สร้างคำประสมทั้งหมด 5,000 ข่าว สามารถเข้าถึงได้จาก

<https://drive.google.com/drive/u/0/folders/1aXGrXF7Y9BX5NLbiYsvrTc7FEAm1Gouj>

รายการคำประสมที่ได้จาก หัวข้อ 3.2 สามารถเข้าถึงได้จาก

<https://drive.google.com/drive/u/0/folders/16Zb5RGVn3wCnUBTgB63s50qgVO7Dhrk2>

ข้อมูลการสรุปใจความสำคัญของผู้เชี่ยวชาญภาษาไทย จำนวน 400 ชุดสามารถเข้าถึงได้จาก

[https://drive.google.com/drive/u/0/folders/1n1svddXskMulEmSRc6kd\\_O2T8k7YPo-k](https://drive.google.com/drive/u/0/folders/1n1svddXskMulEmSRc6kd_O2T8k7YPo-k)

## ภาคผนวก ค

ตัวอย่างข่าวที่ได้จากการสรุปใจความสำคัญของข่าวท่องเที่ยวไทย

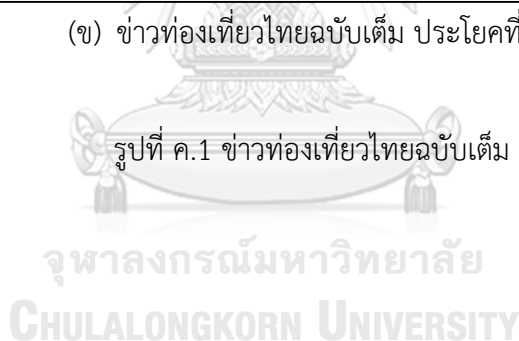
ในภาคผนวก ค. แสดงการสรุปใจความสำคัญโดยใช้การคำนวณค่าน้ำหนักของคำสำคัญโดยหาค่าความถี่ของคำจากหัวข้อข่าวและเลือกประโยคเรียงลำดับจากผลรวมความถี่ของคำสำคัญจากหัวข้อข่าว

<p>Topic: ถึงเวลาส่องแก่งน้ำใส ที่ “แก่งหินเพิง” จ.ปราจีนบุรี</p> <p>S1 = ' ททท.ชวนนักท่องเที่ยวสนุกไปกับกิจกรรมส่องแก่งหินเพิง แก่งน้ำใสสะอาดที่มีระดับความง่าย-ยากของสายน้ำอยู่ที่ 3-5 ซึ่งจะมีสายน้ำเหมาะสำหรับการส่องแก่งเฉพาะในหน้าฝนเท่านั้น'</p> <p>S2 = ' นางสาวเบญจวรรณ สุนทรวรกุล ผู้อำนวยการสำนักงาน ททท.ภาคกลางเขต 8 เปิดเผยว่า กิจกรรมส่องแก่งหินเพิง เป็นกิจกรรมท่องเที่ยวที่มีมนต์เสน่ห์ไม่เหมือนแก่งที่อื่น ๆ ด้วยสายธารที่มีความใสสะอาด มีความรุนแรงที่พอเหมาะกับนักท่องเที่ยวทุกเพศทุกวัย'</p> <p>S3 = ' ซึ่งหากนักท่องเที่ยวที่มาเป็นครอบครัวและสนใจที่จะเดินทางมาท่องเที่ยวในช่วงฤดูฝนนี้ สำหรับแก่งหินเพิง ถือกำเนิดเกิดจากลำน้ำใสใหญ่ในเขตอุทยานฯ เขาใหญ่ สิ้นสุดที่ปลายน้ำ ณ ต.สะพานหิน อ.นาดี จ.ปราจีนบุรี และไหลลงสู่ม่าน้ำปราจีนบุรีไหลรวมกับแม่น้ำนครนายกกรมเป็นแม่น้ำบางปะกง'</p> <p>S4 = ' โดยในฤดูฝนระหว่างเดือนมิถุนายนถึงตุลาคมของทุกปี แก่งหินเพิงจะมีปริมาณน้ำที่เหมาะสมแก่การส่องแก่ง ในประเภทเรือยาง นั่งได้ประมาณ 8-10 คน ผ่านแก่งต่างๆ ที่มีระดับความง่ายของสายน้ำอยู่ที่ 3-5 มีแก่งหินที่สามารถทำให้หัวใจของทุกคนเต้นแรง กระแสน้ำไหลพาเรืออย่างแสนปะทะโชดหิน ทำให้เกิดกระแสน้ำวนไหลแรง'</p> <p>S5 = ' ผู้นำส่องแก่งจึงต้องใช้ฝีมือและทักษะในการพาย รวมทั้งต้องใช้ความระมัดระวังเป็นพิเศษ ร่วมแรงร่วมใจช่วยกันพายเพื่อให้พ้นกับการปะทะกับแก่งหินและป้องกันเรือพลิกคว่ำ'</p> <p>S6 = 'และบางจุดกระแสน้ำจะเบา ทุกคนก็จะสนุกสนานกับลำธารใสที่สามารถแวะพักเหนื่อยเล่นน้ำใสเย็นกันได้ทั้งนี้เพื่อความสนุกสนานควบคู่กับความปลอดภัย'</p> <p>S7 = ' ก่อนจะส่องแก่งหินเพิงทุกครั้ง นักท่องเที่ยวจึงจำเป็นต้องรับฟังและเรียนรู้ทักษะการบังคับเรือเบื้องต้นก่อน'</p> <p>S8 = 'รวมทั้งการช่วยเหลือตนเองในกรณีพลัดตกเรือจากเจ้าหน้าที่ผู้ที่ได้รับการอบรมและมีความชำนาญด้านการส่องแก่งโดยเฉพาะ'</p> <p>S9 = ' หลังจากฟังการบรรยายและฝึกทักษะการพายเรือแล้ว นักท่องเที่ยวจะต้องเดินเท้าจากบริเวณที่ทำการหน่วยพิทักษ์อุทยานแห่งชาติเขาใหญ่ที่ 9 (ขย.9) ไปยังต้นแก่ง ระยะทางประมาณ 2.50 กิโลเมตร'</p> <p>S10 = 'ระหว่างทางนักท่องเที่ยวจะได้ศึกษาระบบนิเวศที่เกื้อหนุนการดำรงอยู่ของสิ่งต่าง ๆ ที่ช่วยเพิ่มจุดสนใจจนเป็นเส้นทางเดินชมธรรมชาติที่คงความอุดมสมบูรณ์ของพรรณไม้'</p> <p>S11 = 'เมื่อถึงจุดเริ่มต้นของการส่องแก่ง จะผ่านแก่ง 6 แก่ง ได้แก่ แก่งหินเพิง แก่งวังหนามล้อม แก่งวังบอน แก่งลูกเสือ แก่งวังไทร แก่งงูเห่า และมีจุดหมายปลายทางที่ท่าเรือบริเวณ ขย.9 มีอัตราค่าบริการส่องแก่ง รายบุคคล คนละประมาณ 450 บาท เหมาลำ ลำละประมาณ 3,500 บาท / 8 คน'</p> <p>S12 = 'ด้านความปลอดภัยนั้น ด้วยลักษณะของแก่งหินเพิงถือว่ามีความยากระดับ 3 - 5'</p> <p>S13 = 'การส่องต้องใช้ทักษะและความระมัดระวัง ความปลอดภัยจึงเป็นสิ่งสำคัญมาก'</p>
---

(ก) ข่าวท่องเที่ยวไทยฉบับเต็ม ประโยคที่ 1-13

- S14 = 'ตั้งนั้น ตามแก่งต่าง ๆ ทั้ง 6 แก่ง จึงมีหน่วยรักษาความปลอดภัยประจำทุกจุด ทุกคนผ่านการฝึกอบรมและทดสอบการช่วยเหลือผู้ประสบเหตุทางน้ำ มีชูชีพ เข็ม อุปกรณ์ พร้อมเจ้าหน้าที่ดูแลความปลอดภัยอย่างใกล้ชิด'
- S15 = 'สำหรับการเดินทางสู่แก่งหินเพ็ง จากตัวจังหวัดปราจีนบุรี ใช้ถนนสุวรรณศร ทางหลวงหมายเลข 33'
- S16 = 'หากมาจากกรุงเทพ ฯ ตรงไปทางอำเภอบ้านทรบุรี บริเวณหลัก กม.201 ให้สังเกตสถานีบริการน้ำมันและโรงเรียนวัดสระคู ทางซ้ายมือ มีถนน รพช. สระคู – สะพานหิน ให้เลี้ยวซ้ายวิ่งเข้าไปประมาณ 17 กิโลเมตร ถึงโรงเรียนบ้านหนองแหน เลี้ยวซ้ายไป อีก 7 กิโลเมตร ถึง ขญ.9 '
- S17 = 'นอกเหนือจากกิจกรรมส่องแก่งหินเพ็งแล้ว ในพื้นที่ใกล้เคียงยังมีแหล่งท่องเที่ยวอื่น ๆ ที่สามารถจัดสรรให้อยู่ในโปรแกรมท่องเที่ยวได้'
- S18 = 'อาทิตร์ เริ่มต้นที่กรุงเทพฯ ฯ เดินทางออกจากบ้านแต่เช้า แวะเที่ยวและชมแหล่งท่องเที่ยวของจังหวัดนครนายก อาทิตร์ น้ำตกสวย ๆ (สาริกา นางรอง อุทยานวังตะไคร้) หรือชมเขื่อนขุนด่านปราการชล เขื่อนคอนกรีตบดอัดยาวที่สุดในโลก, เรียนรู้เรื่องเศรษฐกิจพอเพียงตามแนวพระราชดำริที่ "ศูนย์ภูมิรักษ์ธรรมชาติ" '
- S19 = 'จากนั้นมุ่งหน้าเข้าตัวเมืองปราจีนบุรี ชมพิพิธภัณฑ์การแพทย์แผนไทยที่ "ตึกเจ้าพระยาอภัยภูเบศร" '
- S20 = 'ออกจากตัวเมืองผ่าน "พิพิธภัณฑ์อยู่สุขสุวรรณ" ชมแหล่งสะสมตะเกียงโบราณที่มีจำนวนมากที่สุดในประเทศไทย'
- S21 = ' เดินทางถึงแก่งหินเพ็ง'
- S22 = 'จากนั้นเตรียมตัวเพื่อส่องแก่งทดสอบกำลังใจ เข้าที่พักหรือจะเดินทางกลับโดยสวัสดิภาพ'
- S23 = 'ทั้งนี้นักท่องเที่ยวที่สนใจส่องแก่งหินเพ็งสามารถสอบถามและขอรับตัวอย่างเส้นทางท่องเที่ยวได้ที่ สำนักงาน ททท.ภาคกลางเขต 8 โทร.0-3731-2282, 0-3731-2284, 0-3731-5664'

(ข) ข่าวท่องเที่ยวไทยฉบับเต็ม ประโยคที่ 15-19





**Topic: ถึงเวลาส่งแก่งน้ำใส ที่ “แก่งหินเพ็ง” จ.ปราจีนบุรี**

S1 = ' ททท.ชวนนักท่องเที่ยวสนุกไปกับกิจกรรมส่งแก่งหินเพ็ง แก่งน้ำใสสะอาดที่มีระดับความง่าย-ยากของสายน้ำอยู่ที่ 3-5 ซึ่งจะมีสายน้ำเหมาะสำหรับการส่งแก่งเฉพาะในหน้าฝนเท่านั้น'

S2 = ' นางสาวเบญจวรรณ สุนทรวรกุล ผู้อำนวยการสำนักงาน ททท.ภาคกลางเขต 8 เปิดเผยว่า กิจกรรมส่งแก่งหินเพ็ง เป็นกิจกรรมท่องเที่ยวที่มีมนต์เสน่ห์ไม่เหมือนแก่งที่อื่น ๆ ด้วยสายธารที่มีความใสสะอาด มีความรุนแรงที่พอเหมาะกับนักท่องเที่ยวทุกเพศทุกวัย'

S3 = ' ซึ่งหากนักท่องเที่ยวที่มาเป็นครอบครัวและสนใจที่จะเดินทางมาท่องเที่ยวในช่วงฤดูฝนนี้ สำหรับแก่งหินเพ็ง ถือกำเนิดเกิดจากลำน้ำใสใหญ่ในเขตอุทยานฯ เขาใหญ่ สิ้นสุดที่ปลายน้ำ ณ ต.สะพานหิน อ.นาดี จ.ปราจีนบุรี และไหลลงสู่แม่น้ำปราจีนบุรีไหลรวมกับแม่น้ำนครนายกกรมเป็นแม่น้ำบางปะกง'

S4 = ' โดยในฤดูฝนระหว่างเดือนมิถุนายนถึงตุลาคมของทุกปี แก่งหินเพ็งจะมีปริมาณน้ำที่เหมาะสมแก่การส่งแก่ง ในประเภทเรือยาง นั่งได้ประมาณ 8-10 คน ผ่านแก่งต่างๆ ที่มีระดับความง่าย-ยากของสายน้ำอยู่ที่ 3-5 มีแก่งหินที่สามารถทำให้หัวใจของทุกคนเต้นแรง กระแสน้ำไหลพาเรืออย่างเล่นปะทะโขดหิน ทำให้เกิดกระแสน้ำวนไหลแรง'

S5 = ' ผู้นำส่งแก่งจึงต้องใช้ฝีมือและทักษะในการพาย รวมทั้งต้องใช้ความระมัดระวังเป็นพิเศษ ร่วมแรงร่วมใจช่วยกันพายเพื่อให้พ้นกับการปะทะกับแก่งหินและป้องกันเรือพลิกคว่ำ'

S6 = ' และบางจุดกระแสน้ำจะเบา ทุกคนก็จะสนุกสนานกับลำธารใสที่สามารถแวะพักเหนื่อยเล่นน้ำใสเย็นกันได้ทั้งนี้เพื่อความสนุกสนานควบคู่กับความปลอดภัย'

S7 = ' ก่อนจะส่งแก่งหินเพ็งทุกครั้ง นักท่องเที่ยวจึงจำเป็นต้องรับฟังและเรียนรู้ทักษะการบังคับเรือเบื้องต้นก่อน'

S8 = ' รวมทั้งการช่วยเหลือตนเองในกรณีพลัดตกเรือจากเจ้าหน้าที่ผู้ที่ได้รับการอบรมและมีความชำนาญด้านการส่งแก่งโดยเฉพาะ'

S9 = ' หลังจากฟังการบรรยายและฝึกทักษะการพายเรือแล้ว นักท่องเที่ยวจะต้องเดินเท้าจากบริเวณที่ทำการหน่วยพิทักษ์อุทยานแห่งชาติเขาใหญ่ที่ 9 (ขญ.9) ไปยังต้นแก่ง ระยะทางประมาณ 2.50 กิโลเมตร'

S10 = ' ระหว่างทางนักท่องเที่ยวจะได้ศึกษาระบบนิเวศที่เกื้อหนุนการดำรงอยู่ของสิ่งต่าง ๆ ที่ช่วยเพิ่มจุดสนใจจนเป็นเส้นทางเดินชมธรรมชาติที่คงความอุดมสมบูรณ์ของพรรณไม้'

S11 = ' เมื่อถึงจุดเริ่มต้นของการส่งแก่ง จะผ่านแก่ง 6 แก่ง ได้แก่ แก่งหินเพ็ง แก่งวังนวมล้อม แก่งวังบอน แก่งลูกเสือ แก่งวังไทร แก่งงูเห่า และมีจุดหมายปลายทางที่ท่าเรือบริเวณ ขญ.9 มีอัตราค่าบริการส่งแก่ง รายบุคคล คนละประมาณ 450 บาท เหมาลำ ลำละประมาณ 3,500 บาท / 8 คน'

S12 = ' ด้านความปลอดภัยนั้น ด้วยลักษณะของแก่งหินเพ็งถือว่ามีความยากระดับ 3 - 5'

S13 = ' การส่งแก่งต้องใช้ทักษะและความระมัดระวัง ความปลอดภัยจึงเป็นสิ่งสำคัญมาก'

(ก) ข่าวที่ได้จากการสรุปของผู้เชี่ยวชาญในอัตราการปี้อัตร้อยละ 20

<p>S14 = 'ตั้งนั้น ตามแก่งต่าง ๆ ทั้ง 6 แก่ง จึงมีหน่วยรักษาความปลอดภัยประจำทุกจุด ทุกคนผ่านการฝึกอบรมและทดสอบการช่วยเหลือผู้ประสบเหตุทางน้ำ มีชูชีพ เข็มอก อุปกรณ์ พร้อมเจ้าหน้าที่ดูแลความปลอดภัยอย่างใกล้ชิด'</p> <p>S15 = 'สำหรับการเดินทางสู่แก่งหินเพ็ง จากตัวจังหวัดปราจีนบุรี ใช้ถนนสุวรรณศร ทางหลวงหมายเลข 33'</p> <p>S16 = 'หากมาจากกรุงเทพ ฯ ตรงไปทางอำเภอบึงนารางบุรี บริเวณหลัก กม.201 ให้สังเกตสถานีบริการน้ำมันและโรงเรียนวัดสระคู ทางซ้ายมือ มีถนน รพช. สระคู – สะพานหิน ให้เลี้ยวซ้ายวิ่งเข้าไปประมาณ 17 กิโลเมตร ถึงโรงเรียนบ้านหนองแหน เลี้ยวซ้ายไป อีก 7 กิโลเมตร ถึง ขย.9 '</p> <p>S17 = 'นอกเหนือจากกิจกรรมล่องแก่งหินเพ็งแล้ว ในพื้นที่ใกล้เคียงยังมีแหล่งท่องเที่ยวอื่น ๆ ที่สามารถจัดสรรให้อยู่ในโปรแกรมท่องเที่ยวได้'</p> <p>S18 = 'อาทิต์ เริ่มต้นที่กรุงเทพฯ ฯ เดินทางออกจากบ้านแต่เช้า แวะเที่ยวและชมแหล่งท่องเที่ยวของจังหวัดนครนายก อาทิต์ น้ำตกสวย ๆ (สาริกา นางรอง อุทยานวังตะไคร้) หรือชมเขื่อนขุนด่านปราการชล เขื่อนคอนกรีตบดอัดยาวที่สุดในโลก, เรียนรู้เรื่องเศรษฐกิจพอเพียงตามแนวพระราชดำริที่ "ศูนย์ภูมิรักษ์ธรรมชาติ" '</p> <p>S19 = 'จากนั้นมุ่งหน้าเข้าตัวเมืองปราจีนบุรี ชมพิพิธภัณฑ์การแพทย์แผนไทยที่ "ตึกเจ้าพระยาอภัยภูเบศร" '</p> <p>S20 = 'ออกจากตัวเมืองผ่าน "พิพิธภัณฑ์อยู่สุขสุวรรณ" ชมแหล่งสะสมตะเกียงโบราณที่มีจำนวนมากที่สุดในประเทศไทย'</p> <p>S21 = ' เดินทางถึงแก่งหินเพ็ง'</p> <p>S22 = 'จากนั้นเตรียมตัวเพื่อล่องแก่งทดสอบกำลังใจ เข้าที่พักหรือจะเดินทางกลับโดยสวัสดิภาพ'</p> <p>S23 = 'ทั้งนี้ นักท่องเที่ยวที่สนใจล่องแก่งหินเพ็งสามารถสอบถามและขอรับตัวอย่างเส้นทางท่องเที่ยวได้ที่ สำนักงาน ททท.ภาคกลางเขต 8 โทร.0-3731-2282, 0-3731-2284, 0-3731-5664'</p>
--

(ข) ข่าวดังที่ได้จากการสรุปของผู้เชี่ยวชาญในอัตราการปีบอัดร้อยละ 20 (ต่อ)

รูปที่ ค.2 ข่าวดังที่ได้จากการสรุปของผู้เชี่ยวชาญในอัตราการปีบอัดร้อยละ 20

จุฬาลงกรณ์มหาวิทยาลัย  
CHULALONGKORN UNIVERSITY

**Topic: ถึงเวลาส่องแก่งน้ำใส ที่ “แก่งหินเพ็ง” จ.ปราจีนบุรี**

S1 = ' ททท.ชวนนักท่องเที่ยวสนุกไปกับกิจกรรมส่องแก่งหินเพ็ง แก่งน้ำใสสะอาดที่มีระดับความง่าย-ยากของสายน้ำอยู่ที่ 3-5 ซึ่งจะมีสายน้ำเหมาะสำหรับการส่องแก่งเฉพาะในหน้าฝนเท่านั้น'

S2 = ' นางสาวเบญจวรรณ สุนทรวรกุล ผู้อำนวยการสำนักงาน ททท.ภาคกลางเขต 8 เปิดเผยว่า กิจกรรมส่องแก่งหินเพ็ง เป็นกิจกรรมท่องเที่ยวที่มีมนต์เสน่ห์ไม่เหมือนแก่งที่อื่น ๆ ด้วยสายธารที่มีความใสสะอาด มีความรุนแรงที่พอเหมาะกับนักท่องเที่ยวทุกเพศทุกวัย'

S3 = ' ซึ่งหากนักท่องเที่ยวที่มาเป็นครอบครัวและสนใจที่จะเดินทางมาท่องเที่ยวในช่วงฤดูฝนนี้ สำหรับแก่งหินเพ็ง ถือกำเนิดเกิดจากลำน้ำใสใหญ่ในเขตอุทยานฯ เขาใหญ่ สิ้นสุดที่ปลายน้ำ ณ ต.สะพานหิน อ.นาดี จ.ปราจีนบุรี และไหลลงสู่แม่น้ำปราจีนบุรีไหลรวมกับแม่น้ำนครนายกกรมเป็นแม่น้ำบางปะกง'

S4 = ' โดยในฤดูฝนระหว่างเดือนมิถุนายนถึงตุลาคมของทุกปี แก่งหินเพ็งจะมีปริมาณน้ำที่เหมาะสมแก่การส่องแก่ง ในประเภทเรือยาง นั่งได้ประมาณ 8-10 คน ผ่านแก่งต่างๆ ที่มีระดับความง่าย-ยากของสายน้ำอยู่ที่ 3-5 มีแก่งหินที่สามารถทำให้หัวใจของทุกคนเต้นแรง กระแสน้ำไหลพาเรืออย่างแสนปะทะโขดหิน ทำให้เกิดกระแสน้ำวนไหลแรง'

S5 = ' ผู้นำส่องแก่งจึงต้องใช้ฝีมือและทักษะในการพาย รวมทั้งต้องใช้ความระมัดระวังเป็นพิเศษ ร่วมแรงร่วมใจช่วยกันพายเพื่อให้พ้นกับการปะทะกับแก่งหินและป้องกันเรือพลิกคว่ำ'

S6 = ' และบางจุดกระแสน้ำจะเบา ทุกคนก็จะสนุกสนานกับลำธารใสที่สามารถแวะพักเหนื่อยเล่นน้ำใสเย็นกันได้ทั้งนี้เพื่อความสนุกสนานควบคู่กับความปลอดภัย'

S7 = ' ก่อนจะส่องแก่งหินเพ็งทุกครั้ง นักท่องเที่ยวจึงจำเป็นต้องรับฟังและเรียนรู้ทักษะการบังคับเรือเบื้องต้นก่อน'

S8 = ' รวมทั้งการช่วยเหลือตนเองในกรณีพลัดตกเรือจากเจ้าหน้าที่ผู้ที่ได้รับการอบรมและมีความชำนาญด้านการส่องแก่งโดยเฉพาะ'

S9 = ' หลังจากฟังการบรรยายและฝึกทักษะการพายเรือแล้ว นักท่องเที่ยวจะต้องเดินเท้าจากบริเวณที่ทำการหน่วยพิทักษ์อุทยานแห่งชาติเขาใหญ่ที่ 9 (ขญ.9) ไปยังต้นแก่ง ระยะทางประมาณ 2.50 กิโลเมตร'

S10 = ' ระหว่างทางนักท่องเที่ยวจะได้ศึกษาระบบนิเวศที่เกื้อหนุนการดำรงอยู่ของสิ่งต่าง ๆ ที่ช่วยเพิ่มจุดสนใจจนเป็นเส้นทางเดินชมธรรมชาติที่คงความอุดมสมบูรณ์ของพรรณไม้'

S11 = ' เมื่อถึงจุดเริ่มต้นของการส่องแก่ง จะผ่านแก่ง 6 แก่ง ได้แก่ แก่งหินเพ็ง แก่งวังหนามล้อม แก่งวังบอน แก่งลูกเสือ แก่งวังไทร แก่งงูเห่า และมีจุดหมายปลายทางที่ท่าเรือบริเวณ ขญ.9 มีอัตราค่าบริการส่องแก่ง รายบุคคล คนละประมาณ 450 บาท เหมาลำ ลำละประมาณ 3,500 บาท / 8 คน'

S12 = ' ด้านความปลอดภัยนั้น ด้วยลักษณะของแก่งหินเพ็งถือว่ามีความยากระดับ 3 - 5'

S13 = ' การส่องต้องใช้ทักษะและความระมัดระวัง ความปลอดภัยจึงเป็นสิ่งสำคัญมาก'

(ก) ข่าวที่ได้จากการสรุปของผู้เชี่ยวชาญในอัตราการปี้อัตราร้อยละ 30

<p>S14 = 'ดังนั้น ตามแก่งต่าง ๆ ทั้ง 6 แก่ง จึงมีหน่วยรักษาความปลอดภัยประจำทุกจุด ทุกคนผ่านการฝึกอบรมและทดสอบการช่วยเหลือผู้ประสบเหตุทางน้ำ มีชูชีพ เข็มอก อุปกรณ์ พร้อมเจ้าหน้าที่ดูแลความปลอดภัยอย่างใกล้ชิด'</p> <p>S15 = 'สำหรับการเดินทางสู่แก่งหินเพ็ง จากตัวจังหวัดปราจีนบุรี ใช้ถนนสุวรรณศร ทางหลวงหมายเลข 33'</p> <p>S16 = 'หากมาจากกรุงเทพ ฯ ตรงไปทางอำเภอบึงนารางบุรี บริเวณหลัก กม.201 ให้สังเกตสถานีบริการน้ำมันและโรงเรียนวัดสระคู ทางซ้ายมือ มีถนน รพช. สระคู – สะพานหิน ให้เลี้ยวซ้ายวิ่งเข้าไปประมาณ 17 กิโลเมตร ถึงโรงเรียนบ้านหนองแหน เลี้ยวซ้ายไปอีก 7 กิโลเมตร ถึง ขญ.9 '</p> <p>S17 = 'นอกเหนือจากกิจกรรมล่องแก่งหินเพ็งแล้ว ในพื้นที่ใกล้เคียงยังมีแหล่งท่องเที่ยวอื่น ๆ ที่สามารถจัดสรรให้อยู่ในโปรแกรมท่องเที่ยวได้'</p> <p>S18 = 'อาทิตร์ เริ่มต้นที่กรุงเทพฯ ฯ เดินทางออกจากบ้านแต่เช้า แวะเที่ยวและชมแหล่งท่องเที่ยวของจังหวัดนครนายก อาทิตร์ น้ำตกสวย ๆ (สาริกา นางรอง อุทยานวังตะไคร้) หรือชมเขื่อนขุนด่านปราการชล เขื่อนคอนกรีตบดอัดยาวที่สุดในโลก, เรียนรู้เรื่องเศรษฐกิจพอเพียงตามแนวพระราชดำริที่ "ศูนย์ภูมิรักษ์ธรรมชาติ" '</p> <p>S19 = 'จากนั้นมุ่งหน้าเข้าตัวเมืองปราจีนบุรี ชมพิพิธภัณฑ์การแพทย์แผนไทยที่ "ตึกเจ้าพระยาอภัยภูเบศร" '</p> <p>S20 = 'ออกจากตัวเมืองผ่าน "พิพิธภัณฑ์อยู่สุขสุวรรณ" ชมแหล่งสะสมตะเกียงโบราณที่มีจำนวนมากที่สุดในประเทศไทย'</p> <p>S21 = ' เดินทางถึงแก่งหินเพ็ง'</p> <p>S22 = 'จากนั้นเตรียมตัวเพื่อล่องแก่งทดสอบกำลังใจ เข้าที่พักหรือจะเดินทางกลับโดยสวัสดิภาพ'</p> <p>S23 = 'ทั้งนี้ นักท่องเที่ยวที่สนใจล่องแก่งหินเพ็งสามารถสอบถามและขอรับตัวอย่างเส้นทางท่องเที่ยวได้ที่ สำนักงาน ททท.ภาคกลางเขต 8 โทร.0-3731-2282, 0-3731-2284, 0-3731-5664'</p>
--

(ข) ชาวที่ได้จากการสรุปของผู้เชี่ยวชาญในอัตราการปีบอัดร้อยละ 30 (ต่อ)

รูปที่ ค.3 ชาวที่ได้จากการสรุปของผู้เชี่ยวชาญในอัตราการปีบอัดร้อยละ 30

จุฬาลงกรณ์มหาวิทยาลัย  
CHULALONGKORN UNIVERSITY

**Topic: ถึงเวลาส่องแก่น้ำใส ที่ “แก่งหินเพ็ง” จ.ปราจีนบุรี**

S1 = 'ททท.ชวนนักท่องเที่ยวสนุกไปกับกิจกรรมส่องแก่งหินเพ็ง แก่งน้ำใสสะอาดที่มีระดับความง่าย-ยากของสายน้ำอยู่ที่ 3-5 ซึ่งจะมีสายน้ำเหมาะสำหรับการส่องแก่งเฉพาะในหน้าฝนเท่านั้น'

S2 = ' นางสาวเบญจวรรณ สุนทรวรกุล ผู้อำนวยการสำนักงาน ททท.ภาคกลางเขต 8 เปิดเผยว่า กิจกรรมส่องแก่งหินเพ็ง เป็นกิจกรรมท่องเที่ยวที่มีมนต์เสน่ห์ไม่เหมือนแก่งที่อื่น ๆ ด้วยสายธารที่มีความใสสะอาด มีความรุนแรงที่พอเหมาะกับนักท่องเที่ยวทุกเพศทุกวัย'

S3 = 'ซึ่งหากนักท่องเที่ยวที่มาเป็นครอบครัวและสนใจที่จะเดินทางมาท่องเที่ยวในช่วงฤดูฝนนี้ สำหรับแก่งหินเพ็ง ถือกำเนิดเกิดจากลำน้ำใสใหญ่ในเขตอุทยานฯ เขาใหญ่ สิ้นสุดที่ปลายน้ำ ณ ต.สะพานหิน อ.นาดี จ.ปราจีนบุรี และไหลลงสู่แม่น้ำปราจีนบุรีไหลรวมกับแม่น้ำนครนายกกรมเป็นแม่น้ำบางปะกง'

S4 = ' โดยในฤดูฝนระหว่างเดือนมิถุนายนถึงตุลาคมของทุกปี แก่งหินเพ็งจะมีปริมาณน้ำที่เหมาะสมแก่การส่องแก่ง ในประเภทเรือยาง นั่งได้ประมาณ 8-10 คน ผ่านแก่งต่างๆ ที่มีระดับความง่ายของสายน้ำอยู่ที่ 3-5 มีแก่งหินที่สามารถทำให้หัวใจของคุณเต้นแรง กระแสน้ำไหลพาเรืออย่างแสนปะทะโขดหิน ทำให้เกิดกระแสน้ำวนไหลแรง'

S5 = 'ผู้มาส่องแก่งจึงต้องใช้ฝีมือและทักษะในการพาย รวมทั้งต้องใช้ความระมัดระวังเป็นพิเศษ ร่วมแรงร่วมใจช่วยกันพายเพื่อให้พ้นกับการปะทะกับแก่งหินและป้องกันเรือพลิกคว่ำ'

S6 = 'และบางจุดกระแสน้ำจะเบา ทุกคนก็จะสนุกสนานกับลำธารใสที่สามารถแวะพักเหนื่อยเล่นน้ำใสเย็นกันได้ทั้งนี้เพื่อความสนุกสนานควบคู่กับความปลอดภัย'

S7 = ' ก่อนจะส่องแก่งหินเพ็งทุกครั้ง นักท่องเที่ยวจึงจำเป็นต้องรับฟังและเรียนรู้ทักษะการบังคับเรือเบื้องต้นก่อน'

S8 = 'รวมทั้งการช่วยเหลือตนเองในกรณีพลัดตกเรือจากเจ้าหน้าที่ผู้ที่ได้รับการอบรมและมีความชำนาญด้านการส่องแก่งโดยเฉพาะ'

S9 = ' หลังจากฟังการบรรยายและฝึกทักษะการพายเรือแล้ว นักท่องเที่ยวจะต้องเดินเท้าจากบริเวณที่ทำการหน่วยพิทักษ์อุทยานแห่งชาติเขาใหญ่ที่ 9 (ขญ.9) ไปยังต้นแก่ง ระยะทางประมาณ 2.50 กิโลเมตร'

S10 = 'ระหว่างทางนักท่องเที่ยวจะได้ศึกษาระบบนิเวศที่เกื้อหนุนการดำรงอยู่ของสิ่งต่าง ๆ ที่ช่วยเพิ่มจุดสนใจจนเป็นเส้นทางเดินชมธรรมชาติที่คงความอุดมสมบูรณ์ของพรรณไม้'

S11 = 'เมื่อถึงจุดเริ่มต้นของการส่องแก่ง จะผ่านแก่ง 6 แก่ง ได้แก่ แก่งหินเพ็ง แก่งวังนวมล้อม แก่งวังบอน แก่งลูกเสือ แก่งวังไทร แก่งงูเห่า และมีจุดหมายปลายทางที่ท่าเรือบริเวณ ขญ.9 มีอัตราค่าบริการส่องแก่ง รายบุคคล คนละประมาณ 450 บาท เหมาลำ ลำละประมาณ 3,500 บาท / 8 คน'

S12 = 'ด้านความปลอดภัยนั้น ด้วยลักษณะของแก่งหินเพ็งถือว่ามีความยากระดับ 3 - 5'

S13 = 'การส่องต้องใช้ทักษะและความระมัดระวัง ความปลอดภัยจึงเป็นสิ่งสำคัญมาก'

(ก) ชาวที่ได้จากการสรุปของผู้เชี่ยวชาญในอัตราการบีบอัดร้อยละ 40

S14 = 'ตั้งนั้น ตามแก่งต่าง ๆ ทั้ง 6 แก่ง จึงมีหน่วยรักษาความปลอดภัยประจำทุกจุด ทุกคนผ่านการฝึกอบรมและทดสอบการช่วยเหลือผู้ประสบเหตุทางน้ำ มีชูชีพ เข็มอก อุปกรณ์ พร้อมเจ้าหน้าที่ดูแลความปลอดภัยอย่างใกล้ชิด'
S15 = 'สำหรับการเดินทางสู่แก่งหินเพ็ง จากตัวจังหวัดปราจีนบุรี ใช้ถนนสุวรรณศร ทางหลวงหมายเลข 33'
S16 = 'หากมาจากกรุงเทพ ฯ ตรงไปทางอำเภอบึงนารางบุรี บริเวณหลัก กม.201 ให้สังเกตสถานีบริการน้ำมันและโรงเรียนวัดสระคู ทางซ้ายมือ มีถนน รพช. สระคู – สะพานหิน ให้เลี้ยวซ้ายวิ่งเข้าไปประมาณ 17 กิโลเมตร ถึงโรงเรียนบ้านหนองแหน เลี้ยวซ้ายไป อีก 7 กิโลเมตร ถึง ขญ.9 '
S17 = 'นอกเหนือจากกิจกรรมส่องแก่งหินเพ็งแล้ว ในพื้นที่ใกล้เคียงยังมีแหล่งท่องเที่ยวอื่น ๆ ที่สามารถจัดสรรให้อยู่ในโปรแกรมท่องเที่ยวได้'
S18 = 'อาทิตร์ เริ่มต้นที่กรุงเทพฯ ฯ เดินทางออกจากบ้านแต่เช้า แวะเที่ยวและชมแหล่งท่องเที่ยวของจังหวัดนครนายก อาทิตร์ น้ำตกสวย ๆ (สาริกา นางรอง อุทยานวังตะไคร้) หรือชมเขื่อนขุนด่านปราการชล เขื่อนคอนกรีตบดอัดยาวที่สุดในโลก, เรียนรู้เรื่องเศรษฐกิจพอเพียงตามแนวพระราชดำริที่ “ศูนย์ภูมิรักษ์ธรรมชาติ” '
S19 = 'จากนั้นมุ่งหน้าเข้าตัวเมืองปราจีนบุรี ชมพิพิธภัณฑ์การแพทย์แผนไทยที่ “ตึกเจ้าพระยาอภัยภูเบศร” '
S20 = 'ออกจากตัวเมืองผ่าน “พิพิธภัณฑ์อยู่สุขสุวรรณ” ชมแหล่งสะสมตะเกียงโบราณที่มีจำนวนมากที่สุดในประเทศไทย'
S21 = ' เดินทางถึงแก่งหินเพ็ง'
S22 = 'จากนั้นเตรียมตัวเพื่อส่องแก่งทดสอบกำลังใจ เข้าที่พักหรือจะเดินทางกลับโดยสวัสดิภาพ'
S23 = 'ทั้งนี้ นักท่องเที่ยวที่สนใจส่องแก่งหินเพ็งสามารถสอบถามและขอรับตัวอย่างเส้นทางท่องเที่ยวได้ที่ สำนักงาน ททท.ภาคกลางเขต 8 โทร.0-3731-2282, 0-3731-2284, 0-3731-5664'

(ข) ข่าวดังที่ได้จากการสรุปของผู้เชี่ยวชาญในอัตราการปีบอัดร้อยละ 40 (ต่อ)

รูปที่ ค.4 ข่าวดังที่ได้จากการสรุปของผู้เชี่ยวชาญในอัตราการปีบอัดร้อยละ 40

จุฬาลงกรณ์มหาวิทยาลัย  
CHULALONGKORN UNIVERSITY

- S1 = 'ททท.ชวนนักท่องเที่ยวสนุกไปกับกิจกรรมส่องแก่งหินเพ็ง แก่งน้ำใสสะอาดที่มีระดับความง่าย-ยากของสายน้ำอยู่ที่ 3-5 ซึ่งจะมีสายน้ำเหมาะสำหรับการส่องแก่งเฉพาะในหน้าฝนเท่านั้น'
- S2 = 'นางสาวเบญจวรรณ สุนทรวรกุล ผู้อำนวยการสำนักงาน ททท.ภาคกลางเขต 8 เปิดเผยว่า กิจกรรมส่องแก่งหินเพ็ง เป็นกิจกรรมท่องเที่ยวที่มีมนต์เสน่ห์ไม่เหมือนแก่งที่อื่น ๆ ด้วยสายธารที่มีความใสสะอาด มีความรุนแรงที่พอเหมาะกับนักท่องเที่ยวทุกเพศทุกวัย'
- S3 = 'ซึ่งหากนักท่องเที่ยวที่มาเป็นครอบครัวและสนใจที่จะเดินทางมาท่องเที่ยวในช่วงฤดูฝนนี้ สำหรับแก่งหินเพ็ง ถือกำเนิดเกิดจากลำน้ำใสใหญ่ในเขตอุทยานฯ เขาใหญ่ สิ้นสุดที่ปลายน้ำ ณ ต.สะพานหิน อ.นาดี จ.ปราจีนบุรี และไหลลงสู่ม่น้ำปราจีนบุรีไหลรวมกับแม่น้ำนครนายกกรมเป็นแม่น้ำบางปะกง'
- S4 = 'โดยในฤดูฝนระหว่างเดือนมิถุนายนถึงตุลาคมของทุกปี แก่งหินเพ็งจะมีปริมาณน้ำที่เหมาะสมแก่การส่องแก่ง ในประเภทเรือยาง นั่งได้ประมาณ 8-10 คน ผ่านแก่งต่างๆ ที่มีระดับความง่ายของสายน้ำอยู่ที่ 3-5 มีแก่งหินที่สามารถทำให้หัวใจของทุกคนเต้นแรง กระแสน้ำไหลพาเรืออย่างเล่นปะทะโขดหิน ทำให้เกิดกระแสน้ำวนไหลแรง'
- S5 = 'ผู้นำส่องแก่งจึงต้องใช้ฝีมือและทักษะในการพาย รวมทั้งต้องใช้ความระมัดระวังเป็นพิเศษ ร่วมแรงร่วมใจช่วยกันพายเพื่อให้พ้นกับการปะทะกับแก่งหินและป้องกันเรือพลิกคว่ำ'
- S6 = 'และบางจุดกระแสน้ำจะเบา ทุกคนก็จะสนุกสนานกับลำธารใสที่สามารถแวะพักผ่อนเล่นน้ำใสเย็นกันได้ทั้งนี้เพื่อความสนุกสนานควบคู่กับความปลอดภัย'
- S7 = 'ก่อนจะส่องแก่งหินเพ็งทุกครั้ง นักท่องเที่ยวจึงจำเป็นต้องรับฟังและเรียนรู้ทักษะการบังคับเรือเบื้องต้นก่อน'
- S8 = 'รวมทั้งการช่วยเหลือตนเองในกรณีลัดตกเรือจากเจ้าหน้าที่ผู้ที่ได้รับการอบรมและมีความชำนาญด้านการส่องแก่งโดยเฉพาะ'
- S9 = 'หลังจากฟังการบรรยายและมีทักษะการพายเรือแล้ว นักท่องเที่ยวจะต้องเดินเท้าจากบริเวณที่ทำการหน่วยพิทักษ์อุทยานแห่งชาติเขาใหญ่ที่ 9 (ขญ.9) ไปยังต้นแก่ง ระยะทางประมาณ 2.50 กิโลเมตร'
- S10 = 'ระหว่างทางนักท่องเที่ยวจะได้ศึกษาระบบนิเวศที่เกื้อหนุนการดำรงอยู่ของสิ่งต่าง ๆ ที่ช่วยเพิ่มจุดสนใจจนเป็นเส้นทางเดินชมธรรมชาติที่คงความอุดมสมบูรณ์ของพรรณไม้'
- S11 = 'เมื่อถึงจุดเริ่มต้นของการส่องแก่ง จะผ่านแก่ง 6 แก่ง ได้แก่ แก่งหินเพ็ง แก่งวังหนามล้อม แก่งวังบอน แก่งลูกเสือ แก่งวังไทร แก่งงูเห่า และมีจุดหมายปลายทางที่ท่าเรือบริเวณ ขญ.9 มีอัตราค่าบริการส่องแก่ง รายบุคคล คนละประมาณ 450 บาท เหมาลำ ลำละประมาณ 3,500 บาท / 8 คน'
- S12 = 'ด้านความปลอดภัยนั้น ด้วยลักษณะของแก่งหินเพ็งถือว่ามีความยากระดับ 3 - 5'
- S13 = 'การส่องต้องใช้ทักษะและความระมัดระวัง ความปลอดภัยจึงเป็นสิ่งสำคัญมาก'

(ก) ชาวที่ได้จากการสรุปด้วยวิธีหาคำสำคัญจากหัวข้อข่าวในอัตรการบีบอัดร้อยละ 20

- S14 = 'ตั้งนั้น ตามแก่งต่าง ๆ ทั้ง 6 แก่ง จึงมีหน่วยรักษาความปลอดภัยประจำทุกจุด ทุกคนผ่านการฝึกอบรมและทดสอบการช่วยเหลือผู้ประสบเหตุทางน้ำ มีชูชีพ เชือก อุปกรณ์ พร้อมเจ้าหน้าที่ดูแลความปลอดภัยอย่างใกล้ชิด'
- S15 = 'สำหรับการเดินทางสู่แก่งหินเพิง จากตัวจังหวัดปราจีนบุรี ใช้ถนนสุวรรณศร ทางหลวงหมายเลข 33'
- S16 = 'หากมาจากกรุงเทพ ฯ ตรงไปทางอำเภอบึงนครบุรี บริเวณหลัก กม.201 ให้สังเกตสถานีบริการน้ำมันและโรงเรียนวัดสระคู ทางซ้ายมือ มีถนน รพช. สระคู – สะพานหิน ให้เลี้ยวซ้ายวิ่งเข้าไปประมาณ 17 กิโลเมตร ถึงโรงเรียนบ้านหนองแห่น เลี้ยวซ้ายไปอีก 7 กิโลเมตร ถึง ชญ.9 '
- S17 = 'นอกเหนือจากกิจกรรมล่องแก่งหินเพิงแล้ว ในพื้นที่ใกล้เคียงยังมีแหล่งท่องเที่ยวอื่น ๆ ที่สามารถจัดสรรให้อยู่ในโปรแกรมท่องเที่ยวได้'
- S18 = 'อาทิ เริ่มต้นที่กรุงเทพฯ ฯ เดินทางออกจากบ้านแต่เช้า แวะเที่ยวและชมแหล่งท่องเที่ยวของจังหวัดนครนายก อาทิ น้ำตกสวย ๆ (สาริกา นางรอง อุทยานวังตะไคร้) หรือชมเขื่อนขุนด่านปราการชล เขื่อนคอนกรีตบดอัดยาวที่สุดในโลก, เรียนรู้เรื่องเศรษฐกิจพอเพียงตามแนวพระราชดำริที่ “ศูนย์ภูมิรักษ์ธรรมชาติ” '
- S19 = 'จากนั้นมุ่งหน้าเข้าตัวเมืองปราจีนบุรี ชมพิพิธภัณฑ์การแพทย์แผนไทยที่ “ตึกเจ้าพระยาอภัยภูเบศร” '
- S20 = 'ออกจากตัวเมืองผ่าน “พิพิธภัณฑ์อยู่สุขสุวรรณ” ชมแหล่งสะสมตะเกียงโบราณที่มีจำนวนมากที่สุดในประเทศไทย'
- S21 = 'เดินทางถึงแก่งหินเพิง'
- S22 = 'จากนั้นเตรียมตัวเพื่อล่องแก่งทดสอบกำลังใจ เข้าที่พักหรือจะเดินทางกลับโดยสวัสดิภาพ'
- S23 = 'ทั้งนี้นักท่องเที่ยวที่สนใจล่องแก่งหินเพิงสามารถสอบถามและขอรับตัวอย่างเส้นทางท่องเที่ยวได้ที่ สำนักงาน ททท.ภาคกลางเขต 8 โทร.0-3731-2282, 0-3731-2284, 0-3731-5664'

(ข) ข่าวดังที่ได้จากการสรุปด้วยวิธีหาคำสำคัญจากหัวข้อข่าวในอัตราการปี้อัดร้อยละ 20 (ต่อ)

รูปที่ ค.5 ข่าวดังที่ได้จากการสรุปด้วยวิธีหาคำสำคัญจากหัวข้อข่าวในอัตราการปี้อัดร้อยละ 20

จุฬาลงกรณ์มหาวิทยาลัย  
CHULALONGKORN UNIVERSITY



- S1 = 'ททท.ชวนนักท่องเที่ยวสนุกไปกับกิจกรรมส่องแก่งหินเพ็ง แก่งน้ำใสสะอาดที่มีระดับความง่าย-ยากของสายน้ำอยู่ที่ 3-5 ซึ่งจะมีสายน้ำเหมาะสำหรับการส่องแก่งเฉพาะในหน้าฝนเท่านั้น'
- S2 = 'นางสาวเบญจวรรณ สุเนตรวรกุล ผู้อำนวยการสำนักงาน ททท.ภาคกลางเขต 8 เปิดเผยว่า กิจกรรมส่องแก่งหินเพ็ง เป็นกิจกรรมท่องเที่ยวที่มีมนต์เสน่ห์ไม่เหมือนแก่งที่อื่น ๆ ด้วยสายธารที่มีความใสสะอาด มีความรุนแรงที่พอเหมาะกับนักท่องเที่ยวทุกเพศทุกวัย'
- S3 = 'ซึ่งหากนักท่องเที่ยวที่มาเป็นครอบครัวและสนใจที่จะเดินทางมาท่องเที่ยวในช่วงฤดูฝนนี้ สำหรับแก่งหินเพ็ง ถือกำเนิดเกิดจากลำน้ำใสใหญ่ในเขตอุทยานฯ เขาใหญ่ สิ้นสุดที่ปลายน้ำ ณ ต.สะพานหิน อ.นาดี จ.ปราจีนบุรี และไหลลงสู่น้ำปราจีนบุรีไหลรวมกับแม่น้ำนครนายกกรมเป็นแม่น้ำบางปะกง'
- S4 = 'โดยในฤดูฝนระหว่างเดือนมิถุนายนถึงตุลาคมของทุกปี แก่งหินเพ็งจะมีปริมาณน้ำที่เหมาะสมแก่การส่องแก่ง ในประเภทเรือยาง นั่งได้ประมาณ 8-10 คน ผ่านแก่งต่างๆ ที่มีระดับความง่ายของสายน้ำอยู่ที่ 3-5 มีแก่งหินที่สามารถทำให้หัวใจของทุกคนเต้นแรง กระแสน้ำไหลพาเรือยางเล่นปะทะโขดหิน ทำให้เกิดกระแสสีขาวไหลแรง'
- S5 = 'ผู้นำส่องแก่งจึงต้องใช้ฝีมือและทักษะในการพาย รวมทั้งต้องใช้ความระมัดระวังเป็นพิเศษ ร่วมแรงร่วมใจช่วยกันพายเพื่อให้พ้นกับการปะทะกับแก่งหินและป้องกันเรือพลิกคว่ำ'
- S6 = 'และบางจุดกระแสน้ำจะเบา ทุกคนก็จะสนุกสนานกับลำธารใสที่สามารถแวะพักเหนื่อยเล่นน้ำใสเย็นกันได้ทั้งนี้เพื่อความสนุกสนานควบคู่กับความปลอดภัย'
- S7 = 'ก่อนจะส่องแก่งหินเพ็งทุกครั้ง นักท่องเที่ยวจึงจำเป็นต้องรับฟังและเรียนรู้ทักษะการบังคับเรือเบื้องต้นก่อน'
- S8 = 'รวมทั้งการช่วยเหลือตนเองในกรณีพลัดตกเรือจากเจ้าหน้าที่ผู้ที่ได้รับการอบรมและมีความชำนาญด้านการส่องแก่งโดยเฉพาะ'
- S9 = 'หลังจากฟังการบรรยายและฝึกทักษะการพายเรือแล้ว นักท่องเที่ยวจะต้องเดินเท้าจากบริเวณที่ทำการหน่วยพิทักษ์อุทยานแห่งชาติเขาใหญ่ที่ 9 (ขญ.9) ไปยังต้นแก่ง ระยะทางประมาณ 2.50 กิโลเมตร'
- S10 = 'ระหว่างทางนักท่องเที่ยวจะได้ศึกษาระบบนิเวศที่เกื้อหนุนการดำรงอยู่ของสิ่งต่าง ๆ ที่ช่วยเพิ่มจุดสนใจจนเป็นเส้นทางเดินชมธรรมชาติที่คงความอุดมสมบูรณ์ของพรรณไม้'
- S11 = 'เมื่อถึงจุดเริ่มต้นของการส่องแก่ง จะผ่านแก่ง 6 แก่ง ได้แก่ แก่งหินเพ็ง แก่งวังนวมล้อม แก่งวังบอน แก่งลูกเสือ แก่งวังไทร แก่งงูเห่า และมีจุดหมายปลายทางที่ท่าเรือบริเวณ ขญ.9 มีอัตราค่าบริการส่องแก่ง รายบุคคล คนละประมาณ 450 บาท เหมาลำ ลำละประมาณ 3,500 บาท / 8 คน'
- S12 = 'ด้านความปลอดภัยนั้น ด้วยลักษณะของแก่งหินเพ็งถือว่ามีความยากระดับ 3 - 5'
- S13 = 'การส่องต้องใช้ทักษะและความระมัดระวัง ความปลอดภัยจึงเป็นสิ่งสำคัญมาก'

(ก) ข่าวที่ได้จากการสรุปด้วยวิธีหาคำสำคัญจากหัวข้อข่าวในอัตราการใช้บัตรร้อยละ 30

- S14 = 'ตั้งนั้น ตามแก่งต่าง ๆ ทั้ง 6 แก่ง จึงมีหน่วยรักษาความปลอดภัยประจำทุกจุด ทุกคนผ่านการฝึกอบรมและทดสอบการช่วยเหลือผู้ประสบเหตุทางน้ำ มีชูชีพ เชือก อุปกรณ์ พร้อมเจ้าหน้าที่ดูแลความปลอดภัยอย่างใกล้ชิด'
- S15 = 'สำหรับการเดินทางสู่แก่งหินเพ็ง จากตัวจังหวัดปราจีนบุรี ใช้ถนนสุวรรณศร ทางหลวงหมายเลข 33'
- S16 = 'หากมาจากกรุงเทพ ฯ ตรงไปทางอำเภอบึงนครบุรี บริเวณหลัก กม.201 ให้สังเกตสถานีบริการน้ำมันและโรงเรียนวัดสระคู ทางซ้ายมือ มีถนน รพช. สระคู – สะพานหิน ให้เลี้ยวซ้ายวิ่งเข้าไปประมาณ 17 กิโลเมตร ถึงโรงเรียนบ้านหนองแหน เลี้ยวซ้ายไปอีก 7 กิโลเมตร ถึง ชญ.9 '
- S17 = 'นอกเหนือจากกิจกรรมล่องแก่งหินเพ็งแล้ว ในพื้นที่ใกล้เคียงยังมีแหล่งท่องเที่ยวอื่น ๆ ที่สามารถจัดสรรให้อยู่ในโปรแกรมท่องเที่ยวได้'
- S18 = 'อาทิ เริ่มต้นที่กรุงเทพฯ ฯ เดินทางออกจากบ้านแต่เช้า แวะเที่ยวและชมแหล่งท่องเที่ยวของจังหวัดนครนายก อาทิ น้ำตกสวย ๆ (สาริกา นางรอง อุทยานวังตะไคร้) หรือชมเขื่อนขุนด่านปราการชล เขื่อนคอนกรีตบดอัดยาวที่สุดในโลก, เรียนรู้เรื่องเศรษฐกิจพอเพียงตามแนวพระราชดำริที่ “ศูนย์ภูมิรักษ์ธรรมชาติ” '
- S19 = 'จากนั้นมุ่งหน้าเข้าตัวเมืองปราจีนบุรี ชมพิพิธภัณฑ์การแพทย์แผนไทยที่ “ตึกเจ้าพระยาอภัยภูเบศร” '
- S20 = 'ออกจากตัวเมืองผ่าน “พิพิธภัณฑ์อยู่สุขสุวรรณ” ชมแหล่งสะสมตะเกียงโบราณที่มีจำนวนมากที่สุดในประเทศไทย'
- S21 = 'เดินทางถึงแก่งหินเพ็ง'
- S22 = 'จากนั้นเตรียมตัวเพื่อล่องแก่งทดสอบกำลังใจ เข้าที่พักหรือจะเดินทางกลับโดยสวัสดิภาพ'
- S23 = 'ทั้งนี้นักท่องเที่ยวที่สนใจล่องแก่งหินเพ็งสามารถสอบถามและขอรับตัวอย่างเส้นทางท่องเที่ยวได้ที่ สำนักงาน ททท.ภาคกลางเขต 8 โทร.0-3731-2282, 0-3731-2284, 0-3731-5664'

(ข) ข่าวดังที่ได้จากการสรุปด้วยวิธีหาคำสำคัญจากหัวข้อข่าวในอัตราการปี้อัดร้อยละ 30 (ต่อ)

รูปที่ ค.6 ข่าวดังที่ได้จากการสรุปด้วยวิธีหาคำสำคัญจากหัวข้อข่าวในอัตราการปี้อัดร้อยละ 30

- S1 = 'ททท.ชวนนักท่องเที่ยวสนุกไปกับกิจกรรมส่องแก่งหินเพ็ง แก่งน้ำใสสะอาดที่มีระดับความง่าย-ยากของสายน้ำอยู่ที่ 3-5 ซึ่งจะมีสายน้ำเหมาะสำหรับการส่องแก่งเฉพาะในหน้าฝนเท่านั้น'
- S2 = 'นางสาวเบญจวรรณ สุเนตรวรกุล ผู้อำนวยการสำนักงาน ททท.ภาคกลางเขต 8 เปิดเผยว่า กิจกรรมส่องแก่งหินเพ็ง เป็นกิจกรรมท่องเที่ยวที่มีมนต์เสน่ห์ไม่เหมือนแก่งที่อื่น ๆ ด้วยสายธารที่มีความใสสะอาด มีความรุนแรงที่พอเหมาะกับนักท่องเที่ยวทุกเพศทุกวัย'
- S3 = 'ซึ่งหากนักท่องเที่ยวที่มาเป็นครอบครัวและสนใจที่จะเดินทางมาท่องเที่ยวในช่วงฤดูฝนนี้ สำหรับแก่งหินเพ็ง ถือกำเนิดเกิดจากลำน้ำใสใหญ่ในเขตอุทยานฯ เขาใหญ่ สิ้นสุดที่ปลายน้ำ ณ ต.สะพานหิน อ.นาดี จ.ปราจีนบุรี และไหลลงสู่ม่น้ำปราจีนบุรีไหลรวมกับแม่น้ำนครนายกกรมเป็นแม่น้ำบางปะกง'
- S4 = 'โดยในฤดูฝนระหว่างเดือนมิถุนายนถึงตุลาคมของทุกปี แก่งหินเพ็งจะมีปริมาณน้ำที่เหมาะสมแก่การส่องแก่ง ในประเภทเรือยาง นั่งได้ประมาณ 8-10 คน ผ่านแก่งต่างๆ ที่มีระดับความง่ายของสายน้ำอยู่ที่ 3-5 มีแก่งหินที่สามารถทำให้หัวใจของทุกคนเต้นแรง กระแสน้ำไหลพาเรืออย่างเล่นปะทะโขดหิน ทำให้เกิดกระแสน้ำวนไหลแรง'
- S5 = 'ผู้นำส่องแก่งจึงต้องใช้ฝีมือและทักษะในการพาย รวมทั้งต้องใช้ความระมัดระวังเป็นพิเศษ ร่วมแรงร่วมใจช่วยกันพายเพื่อให้พ้นกับการปะทะกับแก่งหินและป้องกันเรือพลิกคว่ำ'
- S6 = 'และบางจุดกระแสน้ำจะเบา ทุกคนก็จะสนุกสนานกับลำธารใสที่สามารถแวะพักผ่อนเล่นน้ำใสเย็นกันได้ทั้งนี้เพื่อความสนุกสนานควบคู่กับความปลอดภัย'
- S7 = 'ก่อนจะส่องแก่งหินเพ็งทุกครั้ง นักท่องเที่ยวจึงจำเป็นต้องรับฟังและเรียนรู้ทักษะการบังคับเรือเบื้องต้นก่อน'
- S8 = 'รวมทั้งการช่วยเหลือตนเองในกรณีลัดตกเรือจากเจ้าหน้าที่ผู้ที่ได้รับการอบรมและมีความชำนาญด้านการส่องแก่งโดยเฉพาะ'
- S9 = 'หลังจากฟังการบรรยายและมีทักษะการพายเรือแล้ว นักท่องเที่ยวจะต้องเดินเท้าจากบริเวณที่ทำการหน่วยพิทักษ์อุทยานแห่งชาติเขาใหญ่ที่ 9 (ขย.9) ไปยังต้นแก่ง ระยะทางประมาณ 2.50 กิโลเมตร'
- S10 = 'ระหว่างทางนักท่องเที่ยวจะได้ศึกษาระบบนิเวศที่เกื้อหนุนการดำรงอยู่ของสิ่งต่าง ๆ ที่ช่วยเพิ่มจุดสนใจจนเป็นเส้นทางเดินชมธรรมชาติที่คงความอุดมสมบูรณ์ของพรรณไม้'
- S11 = 'เมื่อถึงจุดเริ่มต้นของการส่องแก่ง จะผ่านแก่ง 6 แก่ง ได้แก่ แก่งหินเพ็ง แก่งวังหนามล้อม แก่งวังบอน แก่งลูกเสือ แก่งวังไทร แก่งงูเห่า และมีจุดหมายปลายทางที่ท่าเรือบริเวณ ขย.9 มีอัตราค่าบริการส่องแก่ง รายบุคคล คนละประมาณ 450 บาท เหมาลำ ลำละประมาณ 3,500 บาท / 8 คน'
- S12 = 'ด้านความปลอดภัยนั้น ด้วยลักษณะของแก่งหินเพ็งถือว่ามีความยากระดับ 3 - 5'
- S13 = 'การส่องต้องใช้ทักษะและความระมัดระวัง ความปลอดภัยจึงเป็นสิ่งสำคัญมาก'

(ก) ชาวที่ได้จากการสรุปด้วยวิธีหาคำสำคัญจากหัวข้อข่าวในอัตรการบีบอัดร้อยละ 40

- S14 = 'ตั้งนั้น ตามแก่งต่าง ๆ ทั้ง 6 แก่ง จึงมีหน่วยรักษาความปลอดภัยประจำทุกจุด ทุกคนผ่านการฝึกอบรมและทดสอบการช่วยเหลือผู้ประสบเหตุทางน้ำ มีชูชีพ เชือก อุปกรณ์ พร้อมเจ้าหน้าที่ดูแลความปลอดภัยอย่างใกล้ชิด'
- S15 = 'สำหรับการเดินทางสู่แก่งหินเพิง จากตัวจังหวัดปราจีนบุรี ใช้ถนนสุวรรณศร ทางหลวงหมายเลข 33'
- S16 = 'หากมาจากกรุงเทพฯ ฯ ตรงไปทางอำเภอบึงนครบุรี บริเวณหลัก กม.201 ให้สังเกตสถานีบริการน้ำมันและโรงเรียนวัดสระคู ทางซ้ายมือ มีถนน รพช. สระคู – สะพานหิน ให้เลี้ยวซ้ายวิ่งเข้าไปประมาณ 17 กิโลเมตร ถึงโรงเรียนบ้านหนองแหน เลี้ยวซ้ายไปอีก 7 กิโลเมตร ถึง ขณ.9 '
- S17 = 'นอกเหนือจากกิจกรรมล่องแก่งหินเพิงแล้ว ในพื้นที่ใกล้เคียงยังมีแหล่งท่องเที่ยวอื่น ๆ ที่สามารถจัดสรรให้อยู่ในโปรแกรมท่องเที่ยวได้'
- S18 = 'อาทิต์ เริ่มต้นที่กรุงเทพฯ ฯ เดินทางออกจากบ้านแต่เช้า แวะเที่ยวและชมแหล่งท่องเที่ยวของจังหวัดนครนายก อาทิต์ น้ำตกสวย ๆ (สาริกา นางรอง อุทยานวังตะไคร้) หรือชมเขื่อนขุนด่านปราการชล เขื่อนคอนกรีตบดอัดยาวที่สุดในโลก, เรียนรู้เรื่องเศรษฐกิจพอเพียงตามแนวพระราชดำริที่ “ศูนย์ภูมิรักษ์ธรรมชาติ” '
- S19 = 'จากนั้นมุ่งหน้าเข้าตัวเมืองปราจีนบุรี ชมพิพิธภัณฑ์การแพทย์แผนไทยที่ “ตึกเจ้าพระยาอภัยภูเบศร” '
- S20 = 'ออกจากตัวเมืองผ่าน “พิพิธภัณฑ์อยู่สุขสุวรรณ” ชมแหล่งสะสมตะเกียงโบราณที่มีจำนวนมากที่สุดในประเทศไทย'
- S21 = 'เดินทางถึงแก่งหินเพิง'
- S22 = 'จากนั้นเตรียมตัวเพื่อล่องแก่งทดสอบกำลังใจ เข้าที่พักหรือจะเดินทางกลับโดยสวัสดิภาพ'
- S23 = 'ทั้งนี้นักท่องเที่ยวที่สนใจล่องแก่งหินเพิงสามารถสอบถามและขอรับตัวอย่างเส้นทางท่องเที่ยวได้ที่ สำนักงาน ททท.ภาคกลางเขต 8 โทร.0-3731-2282, 0-3731-2284, 0-3731-5664'

(ข) ข่าวดังที่ได้จากการสรุปด้วยวิธีหาคำสำคัญจากหัวข้อข่าวในอัตราการปี้อัดร้อยละ 40 (ต่อ)

รูปที่ ค.7 ข่าวดังที่ได้จากการสรุปด้วยวิธีหาคำสำคัญจากหัวข้อข่าวในอัตราการปี้อัดร้อยละ 40

## บรรณานุกรม

- Bahmani, B., Moseley, B., Vattani, A., Kumar, R., & Vas-silvitskii, S. J. V. E. (2012). Sklearn k-means. 5, 622-633.
- Chuleerat, J. a. C., K. (2003). *A Practical Text Summarizer by Paragraph Extraction for Thai*. Paper presented at the The Sixth International Workshop on Information Retrieval with Asian Languages (IRAL 2003), Sapporo, Japan.
- De Lathauwer, L., De Moor, B., Vandewalle, J. J. S. j. o. M. A., & Applications. (2000). A multilinear singular value decomposition. 21(4), 1253-1278.
- Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. J. J. o. t. A. s. f. i. s. (1990). Indexing by latent semantic analysis. 41(6), 391-407.
- Gong, Y., & Liu, X. (2001). *Generic text summarization using relevance measure and latent semantic analysis*. Paper presented at the Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval.
- Hahn, U. a. M., I. (2000). The Challenges of Automatic Summarization. *Computer*, 33(11), 29-36. doi:10.1109/2.881692
- Lee, J.-H., Park, S., Ahn, C.-M., Kim, D. J. I. P., & Management. (2009). Automatic generic document summarization based on non-negative matrix factorization. 45(1), 20-34.
- McKinney, W. J. P. f. H. P., & Computing, S. (2011). pandas: a foundational Python library for data analysis and statistics. 14(9).
- Murray, G., Renals, S., & Carletta, J. (2005). Extractive summarization of meeting recordings.
- Oliphant, T. E. (2006). *A guide to NumPy* (Vol. 1): Trelgol Publishing USA.
- Pucktada Treeratpituk. (2017). Cutkum: Thai Word-Segmentation with LSTM in Tensorflow. Retrieved from <https://github.com/pucktada/cutkum>
- Surin Charnyapornpong. (1983). *A Thai Syllable Separation Algorithm*. (Master's Thesis), Asian Institute of Technology,

- เพ็ญศิริ ลีตระกูล. (2551). การเลือกประโยคสำคัญในการสรุปความภาษาไทยโดยใช้แบบจำลองแบบลำดับชั้น. (วิทยานิพนธ์ปริญญามหาบัณฑิต), มหาวิทยาลัยธรรมศาสตร์,
- กมลวรรณ โพธิ์สาย. (2552). การวิเคราะห์ความหมายแฝงและการเรียนรู้เครื่องจักรสำหรับระบบถามตอบอัตโนมัติภาษาไทย. (วิทยานิพนธ์ปริญญามหาบัณฑิต), มหาวิทยาลัยธรรมศาสตร์,
- กรรณิกา บุญเกษม. (2554). การเพิ่มประสิทธิภาพการตัดคำภาษาไทยโดยพิจารณาบริบทรอบข้าง. (วิทยานิพนธ์ปริญญามหาบัณฑิต), มหาวิทยาลัยศิลปากร,
- มงคล ยูสเปรมานนท์. (2551). การย่อความเอกสารภาษาไทยเชิงความหมายโดยใช้กราฟมโนภาพ. (วิทยานิพนธ์ปริญญามหาบัณฑิต), มหาวิทยาลัยธรรมศาสตร์,
- วัฒนพัชร แก้วลิ้ม. (2557). การค้นหาคำสำคัญในเอกสารภาษาไทยโดยใช้เทคนิคการค้นหารูปแบบความสัมพันธ์ในเหมืองข้อมูล. (วิทยานิพนธ์ปริญญามหาบัณฑิต), มหาวิทยาลัยศิลปากร,
- วิรัช ศรีเลิศล้ำวาณิช. (2536). การตัดคำภาษาไทยในระบบแปลภาษา. กรุงเทพฯ: ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ.
- วีระชาติ ศิริไกรวัฒนาวงศ์. (2547). แบบท่องจำหลักการใช้ภาษาไทย. เชียงใหม่: โรงเรียนมงฟอร์ตวิทยาลัย.

## ประวัติผู้เขียน

ชื่อ-สกุล	ศรัญญา นาทองห่อ
วัน เดือน ปี เกิด	24 มกราคม 2534
สถานที่เกิด	ขอนแก่น
วุฒิการศึกษา	วิทยาศาสตรบัณฑิต สาขาเทคโนโลยีสารสนเทศและการสื่อสาร คณะ วิทยาศาสตร์ มหาวิทยาลัยขอนแก่น
ที่อยู่ปัจจุบัน	จังหวัดขอนแก่น



จุฬาลงกรณ์มหาวิทยาลัย  
CHULALONGKORN UNIVERSITY