การสกัดกิจกรรมสำคัญจากคำพิพากษาศาลฎีกา เพื่อสร้างรายงานประจำวันสำหรับตำรวจ

ว่าที่ ร.ต.ท.ณภัทร งามสดใส

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาวิทยาศาสตร์คอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2565
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

Extraction of important activities from Supreme Court judgments to generate police daily reports

Acting Pol.Lt. Napat Ngamsodsai

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Computer Science
Department of Computer Engineering
FACULTY OF ENGINEERING
Chulalongkorn University
Academic Year 2022

Copyright of Chulalongkorn University

หัวข้อวิทยานิพนธ์	การสกัดกิจกรรมสำคัญจากคำพิพากษาศาลฎีกา เพื่อสร้างรายงานประจำวันสำหรับตำรวจ
โดย	ว่าที่ ร.ต.ท.ณภัทร งามสดใส
สาขาวิชา	วิทยาศาสตร์คอมพิวเตอร์
อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก	ผู้ช่วยศาสตราจารย์ ดร.สุกรี สินธุภิญโญ

คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย อนุมัติให้นับวิทยานิพนธ์ฉบับนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

..... คณบดีคณะวิศวกรรมศาสตร์
(ศาสตราจารย์ ดร.สุพจน์ เตชวรสินสกุล)

คณะกรรมการสอบวิทยานิพนธ์

..... ประธานกรรมการ
(รองศาสตราจารย์ ดร.ณัฐพงศ์ ชินธเนศ)

..... อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(ผู้ช่วยศาสตราจารย์ ดร.สุกรี สินธุภิญโญ)

..... กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.นัทที นิภานันท์)

..... กรรมการภายนอกมหาวิทยาลัย
(ผู้ช่วยศาสตราจารย์ ดร.เด่นดวง ประดับสุวรรณ)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ณภัทร งามสดใส : การสกัดกิจกรรมสำคัญจากคำพิพากษาศาลฎีกา เพื่อสร้างรายงานประจำวันสำหรับตำรวจ. (Extraction of important activities from Supreme Court judgments to generate police daily reports) อ.ที่ปรึกษาหลัก : ผศ. ดร.สุกรี สินธุภิญโญ

งานวิจัยนี้มีเป้าหมายที่จะหาแนวทางแก้ปัญหาการบันทึกรายงานประจำวันของตำรวจ ที่จะช่วยให้พนักงานสอบสวนสามารถบันทึกรายงานประจำวันได้สะดวกมากยิ่งขึ้น โดยการสกัดสาระสำคัญทางคดีจากคำพิพากษาศาลฎีกาและคดีที่เกี่ยวข้อง เพื่อหาความสัมพันธ์ระหว่างคำทั้งหมดที่ปรากฏบนคำพิพากษา และระบุค่าความสำคัญเหล่านั้นโดยใช้เทคนิคทางกราฟ ได้แก่ ค่ากลางคั่นกลาง (Betweenness Centrality) ค่ากลางเพจแรงก์ (PageRank) ค่ากลางความเป็นศูนย์กลาง (Degree Centrality) ค่ากลางความใกล้ (Closeness Centrality) และ ค่ากลางไอเกนเวกเตอร์ (Eigenvector Centrality) เพื่อนำมาสร้างกราฟสรุปสำหรับการสร้างรายงานประจำวัน ทั้งนี้ วิธีการสร้างรายงานประจำวัน ผู้วิจัยออกแบบขั้นตอนวิธีสำหรับการท่องไปในกราฟสรุปเพื่อให้ได้ใจความสำคัญที่สามารถนำมาเป็นรายละเอียดหรือพฤติการณ์ที่ปรากฏอยู่บนรายงานประจำวัน

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

สาขาวิชา วิทยาศาสตร์คอมพิวเตอร์ ลายมือชื่อนิสิต .....
ปีการศึกษา 2565 ลายมือชื่อ อ.ที่ปรึกษาหลัก .....

6370083021 : MAJOR COMPUTER SCIENCE

KEYWORD: Text Summarization, Natural language processing, SOCIAL NETWORK ANALYSIS

Napat Ngamsodsai : Extraction of important activities from Supreme Court judgments to generate police daily reports. Advisor: SUKREE SINTHUPINYO

This research aims to solve the problem of the police daily record and help inquiry officers record the report conveniently. We extracted significant keywords from the Supreme Court's judgments and related cases. We use the relationship between individual words to identify those keywords. We apply social network graph techniques; Betweenness Centrality, PageRank Centrality, Degree Centrality, Closeness Centrality, and Eigenvector Centrality. To create the daily report, we designed a procedure for traversing a summarized graph. As a result, we got a gist which can be taken as a detail or circumstance to show in the daily report.

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

Field of Study: Computer Science Student's Signature .....

Academic Year: 2022 Advisor's Signature .....

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้ด้วยความอนุเคราะห์ของอาจารย์ ผู้ช่วยศาสตราจารย์ ดร. สุกรี สินธุภิญโญ อาจารย์ที่ปรึกษาวิทยานิพนธ์ ซึ่งท่านได้สละเวลาให้ความรู้ คำปรึกษา ตรวจสอบ แก้ไข ให้คำแนะนำทางการวิจัย ตลอดจนสนับสนุนให้การทำวิจัยครั้งนี้สำเร็จออกมาด้วยดี ข้าพเจ้าจึงขอกราบขอบพระคุณ อาจารย์ ผู้ช่วยศาสตราจารย์ ดร.สุกรี สินธุภิญโญ ไว้ ณ ที่นี้

ขอขอบพระคุณกรรมการสอบวิทยานิพนธ์ รองศาสตราจารย์ ดร.ณัฐพงศ์ ชินธเนศ ผู้ช่วยศาสตราจารย์ ดร.นัททที นิภานันท์ และ ผู้ช่วยศาสตราจารย์ ดร.เด่นดวง ประดับสุวรรณ ที่กรุณาสละเวลาให้คำแนะนำ ตรวจสอบและแก้ไขวิทยานิพนธ์ ซึ่งเป็นประโยชน์ต่อการพัฒนางานวิจัยนี้ และทำให้วิทยานิพนธ์ฉบับนี้สมบูรณ์ยิ่งขึ้น

ขอขอบพระคุณคณาจารย์ ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ที่ได้ประสิทธิ์ประสาทวิชา ให้มีความรู้ ความเข้าใจ จนสามารถนำมาประยุกต์ใช้ให้เกิดประโยชน์ในทุกๆด้าน

ขอขอบคุณบิดา มารดา และน้องสาวของผู้วิจัย ที่คอยสนับสนุน ให้กำลังใจเสมอมา

ขอขอบคุณผู้บังคับบัญชา พี่ๆ เพื่อนๆ ข้าราชการตำรวจ ที่คอยช่วยเหลือ ให้คำแนะนำ คำปรึกษา สนับสนุนข้อมูลที่เกี่ยวข้อง และเป็นประโยชน์ต่อการทำงานวิจัยชิ้นนี้

อนึ่ง ผู้วิจัยหวังว่า วิทยานิพนธ์ฉบับนี้ จะมีประโยชน์อยู่ไม่น้อย สำหรับข้อบกพร่องที่อาจจะเกิดขึ้น ผู้วิจัยยินดีที่รับฟังคำแนะนำจากทุกท่าน เพื่อเป็นประโยชน์ในการพัฒนางานวิจัยต่อไป

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ณภัทร งามสดใส

สารบัญ

หน้า

บทคัดย่อภาษาไทย ..... ค

บทคัดย่อภาษาอังกฤษ ..... ง

กิตติกรรมประกาศ ..... จ

สารบัญ ..... ฉ

สารบัญตาราง ..... ฌ

สารบัญรูปภาพ ..... ฎ

บทที่ 1 15

1.1 ความเป็นมาและความสำคัญของปัญหา 15

1.2 วัตถุประสงค์ของการวิจัย 18

1.3 ขอบเขตการวิจัย 18

1.4 ขั้นตอนและวิธีการดำเนินการวิจัย 19

1.5 ประโยชน์ที่คาดว่าจะได้รับ 19

บทที่ 2 20

2.1 การทบทวนวรรณกรรม 20

2.2 ทฤษฎีที่เกี่ยวข้อง 27

บทที่ 3 34

3.1 การเตรียมข้อมูลสำหรับงานวิจัย (Data Preparation) 34

3.2 การทำความสะอาดข้อมูล (Data Cleansing) 36

3.3 การประมวลผลข้อมูลเบื้องต้น (Data Preprocessing) 37

3.4 การสร้างกราฟเชิงความหมาย (Construction of the Graph) 41

3.5 การหาค่ากลาง (Centrality) 43

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

3.6 ขั้นตอนการท่องกราฟเพื่อสร้างแบบรายงานประจำวันสำหรับตำรวจ (Graph traverse procedure) 49
3.7 การวัดผล (Evaluation) 57
บทที่ 4 60
4.1 สภาพแวดล้อมและเครื่องมือที่ใช้ในการทดสอบ 60
4.2 การทดสอบ (Experiment) 60
4.2.1 การทดสอบผลลัพธ์ที่ถูกสกัดคำสำคัญด้วยวิธี TF-IDF และ LSA 61
4.2.2 การทดสอบผลลัพธ์เปรียบเทียบวิธีการทางกราฟเพื่อสกัดคำสำคัญ 62
4.2.3 การทดสอบแบบไขว้ (Cross Validation) 63
4.2.4 การทดสอบค่าเฉลี่ย 2 กลุ่ม (Paired – Samples T Test) 64
4.3 การสร้างผลลัพธ์จากการสกัดกิจกรรมของคำพิพากษาศาลฎีกา (Generate a police daily report) 66
4.3.1 สรุปผลลัพธ์ 85
4.3.2 การสร้างผลลัพธ์จากการสกัดกิจกรรมของคำพิพากษาศาลฎีกาเพิ่มเติม เฉพาะคดีอาญาความผิดฐานชิงทรัพย์ 95
4.3.3 การสร้างผลลัพธ์จากการสกัดกิจกรรมของคำพิพากษาศาลฎีกาเพิ่มเติม เฉพาะคดีอาญาความผิดฐานลักทรัพย์ 108
บทที่ 5 119
5.1 สรุปผลการวิจัย 119
5.2 ข้อจำกัดงานวิจัย 121
5.3 แนวทางวิจัยในอนาคต 121
บรรณานุกรม 123
ภาคผนวก 128
ภาคผนวก ก. ตัวอย่างข้อมูลคำพิพากษาศาลฎีกา และชุดคำสั่ง (Source Code) 129
ภาคผนวก ข. ภาพผลลัพธ์กราฟเมื่อผ่านกระบวนการสกัดกิจกรรมสำคัญ 131

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ประวัติผู้เขียน 136

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

สารบัญตาราง

หน้า

ตารางที่ 1 แสดงจำนวนประชากรไทย พ.ศ.2559-2563 16

ตารางที่ 2 ข้อมูลการรับแจ้งคดีอุบัติเหตุจราจร ระบบสารสนเทศสถานีตำรวจ 17

ตารางที่ 3 แสดงรายการชนิดของคำ (Part of Speech)..... 29

ตารางที่ 4 ตัวอย่างการตัดคำภาษาไทยด้วย PythaiNLP..... 38

ตารางที่ 5 ตัวอย่างการตัดประโยคภาษาไทยด้วย PythaiNLP 38

ตารางที่ 6 ตัวอย่างชื่อกลุ่มของ Named-Entity Tag ใน PythaiNLP 39

ตารางที่ 7 แสดงคุณลักษณะของกราฟ..... 52

ตารางที่ 8 ตารางแสดงคำทั้งหมดที่ความสำคัญต่อเครือข่าย 55

ตารางที่ 9 ตารางแสดงตัวอย่างคำสำคัญที่ได้จากคำพิพากษาศาลฎีกา 58

ตารางที่ 10 แสดงตัวอย่างผลลัพธ์การวัดผลเทียบกับคำพิพากษาศาลฎีกา 59

ตารางที่ 11 ผลลัพธ์การทดสอบการสกัดกิจกรรมสำคัญด้วยวิธี TF-IDF และ LSA ที่ N=3..... 61

ตารางที่ 12 ผลลัพธ์การทดสอบการสกัดกิจกรรมสำคัญด้วยวิธีการทางกราฟ ที่ N=3 62

ตารางที่ 13 ผลลัพธ์การทดสอบการสกัดกิจกรรมสำคัญ ด้วย 10-fold Cross Validation 64

ตารางที่ 14 ตารางผลลัพธ์ค่าเฉลี่ย 2 กลุ่มเทียบกับ TF-IDF 65

ตารางที่ 15 ตารางผลลัพธ์ค่าเฉลี่ย 2 กลุ่มเทียบกับ LSA..... 65

ตารางที่ 16 ตัวอย่างการท่องกราฟ และโหนดที่ถูกคัดเลือกด้วยวิธีค่ากลางคั่นกลาง..... 68

ตารางที่ 17 ตัวอย่างการท่องกราฟ และโหนดที่ถูกคัดเลือกด้วยวิธีค่ากลางความเป็นศูนย์กลาง 69

ตารางที่ 18 ตัวอย่างการท่องกราฟ และโหนดที่ถูกคัดเลือกด้วยวิธีค่ากลางความใกล้ 70

ตารางที่ 19 ตัวอย่างการท่องกราฟ และโหนดที่ถูกคัดเลือกด้วยวิธีค่ากลางไอเกนเวกเตอร์ 71

ตารางที่ 20 ตัวอย่างการท่องกราฟ และโหนดที่ถูกคัดเลือกด้วยวิธีค่ากลางแพจแรงก์ 72

ตารางที่ 21 แสดงรายละเอียดการท่องกราฟเพื่อสร้างรายงานประจำวัน..... 80

จุฬาลงกรณ์มหาวิทยาลัย Chulalongkorn University

ตารางที่ 22 แสดงรายละเอียดการท่องกราฟเพื่อสร้างรายงานประจำวัน..... 82
ตารางที่ 23 แสดงรายละเอียดการท่องกราฟเพื่อสร้างรายงานประจำวัน..... 84
ตารางที่ 24 แสดงรายการสรุปรายละเอียดโหนดที่ผู้ใช้เลือกทั้งหมด..... 86
ตารางที่ 25 ตารางแสดงผลลัพธ์การท่องกราฟ..... 87
ตารางที่ 26 ตารางแสดงผลลัพธ์การท่องกราฟ..... 89
ตารางที่ 27 ตารางแสดงผลลัพธ์การท่องกราฟ..... 90
ตารางที่ 28 ตารางแสดงผลลัพธ์การท่องกราฟ..... 93
ตารางที่ 29 แสดงรายละเอียดการท่องกราฟเพื่อสร้างรายงานประจำวัน..... 96
ตารางที่ 30 แสดงรายการสรุปรายละเอียดโหนดที่ผู้ใช้เลือกทั้งหมด..... 98
ตารางที่ 31 ตารางแสดงผลลัพธ์การท่องกราฟ..... 100
ตารางที่ 32 ตารางแสดงผลลัพธ์การท่องกราฟ..... 102
ตารางที่ 33 ตารางแสดงผลลัพธ์การท่องกราฟ..... 104
ตารางที่ 34 ตารางแสดงผลลัพธ์การท่องกราฟ..... 106
ตารางที่ 35 แสดงรายการสรุปรายละเอียดโหนดที่ผู้ใช้เลือกทั้งหมด..... 110
ตารางที่ 36 แสดงผลลัพธ์การท่องกราฟ..... 111
ตารางที่ 37 แสดงคำที่ได้สำหรับแม่แบบรายงานประจำวัน พร้อมตัวอย่างประโยค..... 113
ตารางที่ 38 แสดงคำที่ได้สำหรับแม่แบบรายงานประจำวัน พร้อมตัวอย่างประโยค..... 115
ตารางที่ 39 แสดงคำที่ได้สำหรับแม่แบบรายงานประจำวัน พร้อมตัวอย่างประโยค..... 116

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

สารบัญรูปภาพ

หน้า

ภาพที่ 1 จำนวนรถจดทะเบียนสะสมทั่วประเทศ ณ วันที่ 30 กันยายน 2560 – 2564 17
ภาพที่ 2 กระบวนการลดความซับซ้อนของคำศัพท์ 20
ภาพที่ 3 สถาปัตยกรรมของระบบการลดความซับซ้อน 21
ภาพที่ 4 ขั้นตอนการสร้างเอกสารสรุปด้วยการวิเคราะห์โครงสร้างความหมาย 22
ภาพที่ 5 ภาพตัวอย่างกราฟแสดงความสัมพันธ์เชิงความหมาย 23
ภาพที่ 6 ภาพตัวอย่างกราฟการเชื่อมต่อกันด้วยความสัมพันธ์ทางความหมาย 24
ภาพที่ 7 แผนภาพแบบจำลองวิธีการแยกสำหรับการสรุปข้อความโดยใช้กราฟ 25
ภาพที่ 8 ขั้นตอนการทำงานของการวิเคราะห์เครือข่ายของข้อความ 26
ภาพที่ 9 เครือข่ายข้อความตามคู่คำและกลุ่มคำ 26
ภาพที่ 10 คำหลัก และคำพึ่งพา 30
ภาพที่ 11 ความสัมพันธ์รูปแบบต่างๆ ระหว่างคำ 30
ภาพที่ 12 กราฟตัวแทนของการเชื่อมโยงเครือข่าย 31
ภาพที่ 13 แสดงแผนภาพการทำงานโดยรวม 34
ภาพที่ 14 แสดงชุดคำสั่ง HTML ของเว็บสืบค้นคำพิพากษาศาลฎีกา 35
ภาพที่ 15 ตัวอย่างแบบรายงานประจำวันเกี่ยวกับคดี 36
ภาพที่ 16 ตัวอย่างคำที่มักปรากฏในชั้นศาล 37
ภาพที่ 17 ตัวอย่างคำฟุ่มเฟือยภาษาไทย (Thai Stopword) 39
ภาพที่ 18 แสดงการรู้จำเอนทิตี้ด้วย PythaiNLP 40
ภาพที่ 19 แสดงต้นไม้แจงส่วน 41
ภาพที่ 20 แสดงความสัมพันธ์ระหว่างคำ 41
ภาพที่ 21 แสดงเดต้าเฟรมสำหรับสร้างกราฟ 42

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ภาพที่ 22 ซอร์สโค้ดการสร้างกราฟด้วย NetworkX 43
ภาพที่ 23 แสดงกราฟจากคำพิพากษาศาลฎีกาทั้งหมด 43
ภาพที่ 24 แสดงกราฟการหาค่ากลางด้วยวิธีค่ากลางคั่นกลาง 44
ภาพที่ 25 เดต้าเฟรมแสดงค่ากลางคั่นกลาง 44
ภาพที่ 26 แสดงกราฟการหาค่ากลางด้วยวิธีค่ากลางความเป็นศูนย์กลาง 45
ภาพที่ 27 เดต้าเฟรมแสดงค่ากลางความเป็นศูนย์กลาง 45
ภาพที่ 28 แสดงกราฟการหาค่ากลางด้วยวิธีค่ากลางความใกล้ 46
ภาพที่ 29 เดต้าเฟรมแสดงค่ากลางความใกล้ 46
ภาพที่ 30 แสดงกราฟการหาค่ากลางด้วยวิธีค่ากลางไอเกนเวกเตอร์ 47
ภาพที่ 31 แสดงเดต้าเฟรมค่ากลางไอเกนเวกเตอร์ 47
ภาพที่ 32 แสดงกราฟการหาค่าความสำคัญของกราฟด้วยวิธีค่ากลางเพจแรงก์ 48
ภาพที่ 33 แสดงเดต้าเฟรมค่ากลางเพจแรงก์ 48
ภาพที่ 34 แผนผังการทำงานของการสกัดกิจกรรมหรือคำสำคัญ 49
ภาพที่ 35 แสดงโหนดที่เชื่อมกับคำว่า “โจทก์” และมีความคล้ายคลึงสูงสุด 10 อันดับ 50
ภาพที่ 36 แสดงโหนดที่เชื่อมกับโจทก์ พร้อมค่าความคล้ายคลึง 51
ภาพที่ 37 โหนดที่เชื่อมกับโจทก์ และที่มีค่ากลางสูงสุด 51
ภาพที่ 38 กราฟสรุปโหนดที่มีค่ากลางสูงสุด จำนวน 2 โหนด 52
ภาพที่ 39 กราฟสรุปโดยค่ากลางความเป็นศูนย์กลาง 52
ภาพที่ 40 แสดงกราฟสรุป 53
ภาพที่ 41 แสดงกราฟสรุป 53
ภาพที่ 42 แสดงกราฟสรุป 54
ภาพที่ 43 แสดงกราฟสรุป 54
ภาพที่ 44 ซอร์สโค้ดการแบ่งชุดข้อมูลสำหรับ K-fold Cross Validation 63
ภาพที่ 45 การกำหนดโหนดเริ่มต้น และแสดงโหนดที่เกี่ยวข้องสำหรับการท่องกราฟ 67

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ภาพที่ 46 การท่องกราฟไปยังโหนดถัดไป 67
ภาพที่ 47 กราฟแสดงโหนดที่ถูกเลือก (สีแดง) และโหนดที่เกี่ยวข้อง (สีฟ้า) 73
ภาพที่ 48 โหนดที่เกี่ยวข้อง กับ “ผู้เสียหาย” และมีค่าความสำคัญสูงสุด 10 อันดับ 73
ภาพที่ 49 กราฟแสดงโหนดที่ถูกเลือก (สีแดง) และโหนดที่เกี่ยวข้อง (สีฟ้า) 74
ภาพที่ 50 โหนดที่เกี่ยวข้อง กับ “ขับ” และมีค่าความสำคัญสูงสุด 10 อันดับ 74
ภาพที่ 51 กราฟแสดงโหนดที่ถูกเลือก (สีแดง) และโหนดที่เกี่ยวข้อง (สีฟ้า) 75
ภาพที่ 52 โหนดที่เกี่ยวข้อง กับ “ขับรถ” และมีค่าความสำคัญสูงสุด 10 อันดับ 75
ภาพที่ 53 กราฟแสดงโหนดที่ถูกเลือก (สีแดง) และโหนดที่เกี่ยวข้อง (สีฟ้า) 76
ภาพที่ 54 โหนดที่เกี่ยวข้อง กับ “รถ” และมีค่าความสำคัญสูงสุด 10 อันดับ 76
ภาพที่ 55 กราฟแสดงโหนดที่ถูกเลือก (สีแดง) และโหนดที่เกี่ยวข้อง (สีฟ้า) 77
ภาพที่ 56 โหนดที่เกี่ยวข้อง กับ “รถยนต์” และมีค่ากลางสูงสุด 10 อันดับ 77
ภาพที่ 57 กราฟแสดงโหนดที่ถูกเลือก (สีแดง) และโหนดที่เกี่ยวข้อง (สีฟ้า) 78
ภาพที่ 58 โหนดที่เกี่ยวข้อง กับ “รถจักรยานยนต์” และมีค่ากลางสูงสุด 10 อันดับ 78
ภาพที่ 59 กราฟแสดงโหนดที่ถูกเลือก (สีแดง) และโหนดที่เกี่ยวข้อง (สีฟ้า) 79
ภาพที่ 60 โหนดที่เกี่ยวข้อง กับ “เดินรถ” และมีค่าความสำคัญสูงสุด 10 อันดับ 79
ภาพที่ 61 แสดงผลลัพธ์การท่องกราฟเพื่อสร้างแม่แบบรายงานประจำวันคดีจราจร 80
ภาพที่ 62 รายงานประจำวันที่ได้จากกระบวนการสกัดกิจกรรมสำคัญ 86
ภาพที่ 63 รายงานประจำวันที่ได้จากกระบวนการสกัดกิจกรรมสำคัญ 100
ภาพที่ 64 แสดงผลลัพธ์การท่องกราฟเพื่อสร้างแม่แบบรายงานประจำวันคดีอาญา (เหตุลักทรัพย์) 108
ภาพที่ 65 แสดงผลลัพธ์การท่องกราฟเพื่อสร้างแม่แบบรายงานประจำวันคดีอาญา (เหตุลักทรัพย์) 109
ภาพที่ 66 แสดงผลลัพธ์การท่องกราฟเพื่อสร้างแม่แบบรายงานประจำวันคดีอาญา (เหตุลักทรัพย์) 109
ภาพที่ 67 ตัวอย่างไฟล์ HTML จากเว็บไซต์สืบค้นคำพิพากษาศาลฎีกา 130

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ภาพที่ 68 แสดงผลลัพธ์กราฟหลังจากผ่านการสกัดกิจกรรมสำคัญด้วยวิธีค่ากลางคั่นกลาง..... 131
ภาพที่ 69 แสดงผลลัพธ์กราฟหลังจากผ่านการสกัดกิจกรรมสำคัญด้วยวิธีค่ากลางความเป็นศูนย์กลาง
..... 132
ภาพที่ 70 แสดงผลลัพธ์กราฟหลังจากผ่านการสกัดกิจกรรมสำคัญด้วยวิธีค่ากลางความใกล้ 133
ภาพที่ 71 แสดงผลลัพธ์กราฟหลังจากผ่านการสกัดกิจกรรมสำคัญด้วยวิธีค่ากลางไอเกนเวกเตอร์. 134
ภาพที่ 72 แสดงผลลัพธ์กราฟหลังจากผ่านการสกัดกิจกรรมสำคัญด้วยวิธีค่ากลางเพจแรงก์ 135

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บทที่ 1
บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

สำนักงานตำรวจแห่งชาติเป็นส่วนราชการมีฐานะเป็นนิติบุคคลอยู่ในบังคับบัญชาของนายกรัฐมนตรี มีอำนาจหน้าที่ในการป้องกันและปราบปรามการกระทำความผิดทางอาญา [1] และความผิดอื่นตามพระราชบัญญัติต่างๆ ข้าราชการตำรวจซึ่งดำรงตำแหน่งพนักงานสอบสวน มีอำนาจหน้าที่ในการสอบสวนคดีความทางอาญา และรวบรวมพยานหลักฐาน ซึ่งกระบวนการดังกล่าวนั้นเริ่มต้นจากผู้เสียหายจะต้องมาพบพนักงานสอบสวนเพื่อร้องทุกข์หรือกล่าวโทษความผิดทางอาญา

การแจ้งความรับคำร้องทุกข์ของประชาชน ข้าราชการตำรวจที่ปฏิบัติหน้าที่พนักงานสอบสวนจะเป็นผู้รับคำร้องทุกข์หรือกล่าวโทษความผิดทางอาญา และเจ้าหน้าที่เสมียนประจำวันจะเป็นผู้บันทึกรายงานประจำวัน โดยการจดบันทึกไว้เป็นหลักฐานในการสอบสวนความผิดอาญา รวมทั้งบันทึกคำร้องทุกข์ และกล่าวโทษ ซึ่งมักจะทำลงในสมุดรายงานประจำวันของสำนักงานตำรวจแห่งชาติ รายงานประจำวัน ซึ่งเป็นเอกสารที่ตำรวจจัดทำขึ้นสำหรับใช้บันทึกเหตุการณ์ต่าง ๆ ที่เกี่ยวข้องกับงานในหน้าที่ของตำรวจ และงานที่ตำรวจได้ปฏิบัติในวันหนึ่งๆ ไว้เป็นหลักฐานที่สถานีตำรวจ ตู้ยาม ที่พักสายตรวจประจำ หรือที่ทำการของเจ้าพนักงานตำรวจ เพื่อทราบว่าในแต่ละวันทีเหตุการณ์อย่างใดเกิดขึ้น และตำรวจได้จัดการไปอย่างไร และ ได้กำหนดประเภทของรายงานประจำวันไว้ 4 ประเภท ดังนี้ [2]

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

1) สมุดรายงานประจำวันเกี่ยวกับคดี
   สมุดที่ใช้บันทึกเกี่ยวกับการดำเนินการของพนักงานสอบสวนในสอบสวนคดี เพื่อใช้ยืนยันกับบันทึกการสอบสวนและบันทึกต่างๆ ใน สำนวนของพนักงานสอบสวน
2) สมุดรายงานประจำวันรับแจ้งเป็นหลักฐาน
   สมุดที่ใช้บันทึกเกี่ยวกับเรื่องราวต่างๆ ที่ผู้แจ้งมีประสงค์ที่จะแจ้งไว้เป็นหลักฐานเพื่อประกอบการยืนยัน
3) สมุดรายงานประจำวันรับแจ้งเอกสารหาย
   สมุดที่ใช้บันทึกการรับแจ้งเอกสารหาย เพื่อเป็นหลักฐาน ซึ่งไม่เกี่ยวกับคดีอาญา
4) สมุดรายงานประจำวันธุรการ

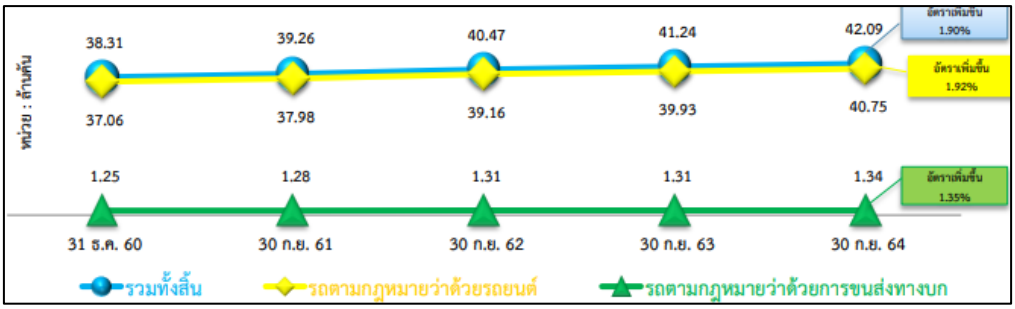
สมุดที่ใช้บันทึกการรับแจ้งไว้เป็นหลักฐานในการปฏิบัติหน้าที่ของตำรวจ เช่น การปฏิบัติหน้าที่เวรยาม การไปและกลับจากการตรวจท้องที่ การไปและกลับจากราชการ การลงทัณฑ์ทางวินัย การเลี้ยงอาหารผู้ต้องหา การเก็บรักษาของหลวง การคืนของส่วนตัวผู้ต้องหา การเก็บรักษาเงิน การส่งเงิน เป็นต้น

จากที่กล่าวมานั้น จะเห็นได้ว่า สมุดรายงานประจำวันเกี่ยวกับคดี เป็นสมุดที่ใช้บันทึกรายละเอียดที่เกี่ยวกับคดีทั้งหมดที่เกิดขึ้นภายในสถานีตำรวจ และใช้ยืนยันกับบันทึกการสอบสวนและบันทึกต่างๆ ในสำนวนของพนักงานสอบสวน สมุดรายงานประจำวันเกี่ยวกับคดีนั้นจึงมีความสำคัญอย่างยิ่งยวด การเป็นเอกสารเพื่อประกอบการดำเนินคดีแก่ผู้กระทำความผิดให้เป็นไปตามกระบวนการยุติธรรม ทั้งนี้ กระบวนการดำเนินการแจ้งความร้องทุกข์กล่าวโทษนั้น เริ่มด้วยการบันทึกรายงานประจำวันนั้น ซึ่งเป็นกระบวนการแรกของการแจ้งความร้องทุกข์หรือกล่าวโทษเพื่อให้ผู้ร้องทุกข์แจ้งข้อมูล และรายเอียดสิ่งที่เกิดขึ้นต่อพนักงานสอบสวน และพนักงานสอบสวนจะดำเนินการบันทึกรายละเอียดลงในสมุดประจำวันเกี่ยวกับคดี ในกรณีที่มีความผิดทางอาญาเกิดขึ้นและดำเนินการสอบสวน เพื่อรวบรวมพยานหลักฐาน และดำเนินคดีกับผู้ที่ถูกกล่าวโทษ

ตารางที่ 1 แสดงจำนวนประชากรไทย พ.ศ.2559-2563

ปี พ.ศ.	**2559**	**2560**	**2561**	**2562**	**2563**
จำนวนประชากร	65,931,550	66,188,503	66,413,979	66,558,935	66,186,727

ปัจจุบัน จำนวนประชากรของประเทศไทย [3] มีจำนวนเพิ่มมากขึ้นในทุกๆปี ดังตารางที่ 1 ประกอบกับรายงานสถิติขนส่ง ของกรมการขนส่งทางบก ประจำปีงบประมาณ พ.ศ.2560-2564 ปรากฏตามภาพที่ 1 มีรายการรถจดทะเบียนสะสม มีอัตราการเติบโตเพิ่มขึ้นร้อยละ 1.90 หรือประมาณ 0.95 ล้านคันต่อปี และข้อมูลการรับแจ้งคดีอุบัติเหตุจราจร ของปี พ.ศ.2563 และ 2564 มีจำนวนเพิ่มขึ้น ดังตารางที่ 2 แสดงให้เห็นว่าใน ปี พ.ศ. 2563 – 2564 มีจำนวนรถยนต์บนท้องถนนเพิ่มมากยิ่งขึ้น และการรับแจ้งคดีอุบัติเหตุจราจรของตำรวจ มีจำนวนเพิ่มมากยิ่งขึ้นตามไปด้วย



ภาพที่ 1 จำนวนรถจดทะเบียนสะสมทั่วประเทศ ณ วันที่ 30 กันยายน 2560 – 2564

ตารางที่ 2 ข้อมูลการรับแจ้งคดีอุบัติเหตุจราจร ระบบสารสนเทศสถานีตำรวจ

ปี พ.ศ.	จำนวนคดีอุบัติเหตุจราจร	เพิ่มขึ้น/ลดลง
2563	94,300	
2564	100,137	+ 6.189%

การดำเนินการตามกระบวนการยุติธรรม เมื่อมีข้อพิพาทเกิดขึ้น กล่าวคือ เมื่อมีการแจ้งความร้องทุกข์ต่อพนักงานสอบสวน พนักงานสอบสวนมีหน้าที่สอบสวน ซักถาม ถึงรายละเอียดสิ่งที่เกิดขึ้นและบันทึกรายละเอียดดังกล่าวลงในสมุดรายงานประจำวัน ในกระบวนการต่อไป เมื่อมีการสอบสวนทำสำนวนการสอบสวน รวบรวมพยานหลักฐานเสร็จเรียบร้อยแล้ว พนักงานสอบสวนมีหน้าที่ส่งสำนวนให้พนักงานอัยการ เพื่อนำคดีดังกล่าว เข้าสู่ชั้นพนักงานอัยการ และสิ้นสุดที่ศาลพิพากษาตัดสินคดีความ จากที่กล่าวมาจะเห็นว่า จุดเริ่มต้นของการดำเนินคดีตามกฎหมาย มีกระบวนการเริ่มจากพนักงานสอบสวน โดยเริ่มที่การบันทึกรายงานประจำวัน

ทั้งนี้ ผู้วิจัยได้มีแนวคิดที่จะนำคำพิพากษาศาลฎีกาของศาล มาวิเคราะห์และสกัดข้อความหรือกิจกรรมสำคัญที่สามารถระบุ จำแนก สิ่งที่จำเป็นต่อการดำเนินคดีความตามกฎหมายเพื่อย้อนไปถึงชั้นแรก คือ ชั้นพนักงานสอบสวนหรือตำรวจ ที่มีการลงบันทึกรายงานประจำวัน ผู้วิจัยเล็งเห็นปัญหาในการบันทึกรายงานประจำวันในสถานีตำรวจที่เกี่ยวกับคดีจราจร แนวโน้มที่เพิ่มขึ้นของจำนวนผู้จดทะเบียนรถสะสม และจำนวนคดีอุบัติเหตุจราจร มีอัตราการเปลี่ยนแปลงไปในทิศทางเดียวกัน จากที่กล่าวมาข้างต้นนั้น การบันทึกประจำวันเป็นสิ่งที่สำคัญอย่างมาก ผู้วิจัยจึงมีแนวความคิดที่จะนำคำพิพากษาศาลฎีกา ซึ่งเป็นศาลสูงสุด มารวบรวมเพื่อสังเคราะห์ ถึงรายละเอียดที่สำคัญต่อการดำเนินคดีตามกระบวนการยุติธรรม และนำผลลัพธ์จากการสังเคราะห์ข้อมูลดังกล่าวมาเป็นต้นแบบในการสร้างระบบบันทึกรายงานประจำวันสำหรับตำรวจเพื่อใช้รองรับการแจ้งความร้องทุกข์ที่สามารถประหยัดเวลาในการทำความเข้าใจ สามารถให้บริการประชาชนได้อย่างรวดเร็ว เมื่อประชาชนเดินทางมาแจ้งความร้องทุกข์ต่างๆ สามารถทำได้อย่างรวดเร็ว และแจ้งข้อมูลที่จำเป็น

ต่อพนักงานสอบสวนได้อย่างถูกต้อง สามารถรองรับคดีอุบัติเหตุจราจรที่เพิ่มมากขึ้น และภารกิจหน้าที่ของสำนักงานตำรวจแห่งชาติ

วิทยานิพนธ์นี้มีวัตถุประสงค์ในการสกัดกิจกรรมสำคัญจากคำพิพากษาศาลฎีกา เพื่อนำมาสร้างรายงานประจำวันสำหรับตำรวจ เพื่อแก้ไขปัญหาการบันทึกรายงานประจำวันเกี่ยวกับคดีจราจรในสถานีตำรวจ สามารถสร้างรายงานประจำวันเกี่ยวกับคดีจราจร จัดการกับรายละเอียดของข้อความหรือถ้อยคำที่ผู้แจ้งมาแจ้งความร้องทุกข์กับพนักงานสอบสวน และคัดเลือกข้อความรายละเอียดที่จำเป็นสำหรับการลงบันทึกรายงานประจำวันได้อย่างถูกต้อง เกิดความสะดวกสบาย และประโยชน์สูงสุดต่อประชาชนผู้เข้ารับบริการและพนักงานสอบสวนโดยตรงในการรับแจ้งความร้องทุกข์ ทั้งนี้วิทยานิพนธ์นี้เป็นการแก้ไขปัญหาเชิงเทคนิค โดยศึกษา วิเคราะห์คำ ข้อความ กิจกรรม ของคำพิพากษาศาลฎีกา และสกัดรายละเอียดสำคัญออกมา เพื่อเป็นเครื่องมือสำหรับช่วยสร้างรายงานประจำวัน เพื่อให้ระบบ สามารถให้บริการประชาชนได้อย่างรวดเร็วมากยิ่งขึ้น อีกทั้งช่วยเพิ่มประสิทธิภาพในด้านวิธีการดำเนินการของพนักงานสอบสวนในการรับคำร้องทุกข์กล่าวโทษ การทำสำนวนการสอบสวน และกระบวนการยุติธรรมในชั้นพนักงานอัยการ และศาล เพื่อให้เป็นไปตามแผนปฏิรูปประเทศด้านกระบวนการยุติธรรม อย่างมีนัยสำคัญ

1.2 วัตถุประสงค์ของการวิจัย

เพื่อศึกษาและวิเคราะห์คำพิพากษาศาลฎีกา และสกัดรายละเอียดสำคัญในการบันทึกรายงานประจำวันที่ช่วยอำนวยความสะดวกในการสร้างรายงานประจำวันเกี่ยวกับคดีจราจร ให้เกิดประสิทธิภาพ ประหยัดเวลาทั้งพนักงานสอบสวน และประชาชน พร้อมทั้ง เป็นต้นแบบในการนำไปต่อยอดพัฒนา เพื่อใช้ในภารกิจของสำนักงานตำรวจแห่งชาติ เพิ่มประสิทธิภาพในการสืบสวนสอบสวน และการให้บริการประชาชนในยุคไทยแลนด์ 4.0

1.3 ขอบเขตการวิจัย

การสกัดกิจกรรมสำคัญจากคำพิพากษาศาลฎีกา เพื่อสร้างรายงานประจำวันสำหรับตำรวจ เฉพาะคดีที่เป็นการรับคำร้องทุกข์ ตาม พ.ร.บ.จราจรทางบก พ.ศ.2522

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

1.4 ขั้นตอนและวิธีการดำเนินการวิจัย

1.4.1 ศึกษากระบวนการบันทึกรายงานประจำวันของพนักงานสอบสวน รายละเอียดที่จำเป็น แ

1.4.2 ศึกษาการเรียนรู้เกี่ยวกับการสร้างเครือข่ายสังคม และการสร้างกราฟเครือข่ายสังคม เพื่อวิเคราะห์ถึงข้อความหรือคำที่มีอิทธิพลต่อการแจ้งความร้องทุกข์คดีจราจร

1.4.3 รวบรวมคำพิพากษาศาลฎีกา

1.4.4 สร้างกราฟ เพื่อวิเคราะห์เครือข่ายสังคมของคำพิพากษาศาลฎีกา พร้อมสกัดรายละเอียดข้อความ หรือกิจกรรมที่สำคัญ เพื่อมาสร้างเป็นองค์ประกอบหลักของการบันทึกประจำวัน

1.4.5 จัดทำเอกสารที่เกี่ยวข้อง

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1.5.1 สามารถลดขั้นตอนการบันทึกรายงานประจำวัน และลดเวลาในการบันทึกรายงานประจำวันสำหรับพนักงานสอบสวน เพื่อให้สามารถให้บริการประชาชนได้อย่างรวดเร็ว

1.5.2 เป็นแนวทางและต้นแบบสำหรับการพัฒนาระบบงาน อื่นๆเพื่อใช้ในสถานีตำรวจ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บทที่ 2
แนวคิดและทฤษฎีที่เกี่ยวข้อง

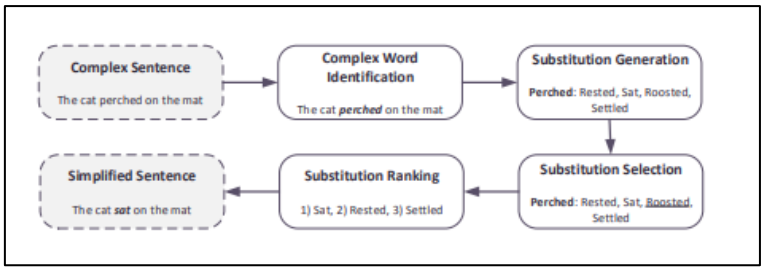
วิทยานิพนธ์นี้มุ่งเน้นในการเพิ่มประสิทธิภาพในการบันทึกรายงานประจำวันสำหรับสถานีตำรวจซึ่งจะส่งผลต่อประสิทธิภาพในการสืบสวนสอบสวนคดีของเจ้าหน้าที่ตำรวจ และเป็นการยกระดับการบริการและอำนวยความสะดวกแก่ประชาชน ดังนั้นงานวิจัยที่เกี่ยวข้องนั้น สามารถสรุปประเด็นที่เกี่ยวข้องได้ดังนี้

2.1 การทบทวนวรรณกรรม

การลดความซับซ้อนของข้อความ เป็นกระบวนการปรับเปลี่ยนภาษาธรรมชาติ เพื่อลดความซับซ้อน โดยนำข้อความมาประมวลผลด้วยกระบวนการตัดคำ และแทนที่ข้อความหรือคำดังกล่าว ด้วยคำที่มีความหมายเหมือนกันหรือ มีความหมายคล้ายคลึงกัน [4]

กระบวนการแรกของการประมวลผลภาษาธรรมชาติ คือ การรู้จำคำ และการสะสมคลังคำศัพท์ (lexicon) เป็นแหล่งทรัพยากรพื้นฐาน หรือ ฐานความรู้จำเป็นขั้นต้นทางภาษาที่ใช้สนับสนุนการประมวลผลภาษาด้วยคอมพิวเตอร์ และเมื่อลดความซับซ้อนของข้อความสำหรับแล้ว ยังสามารถรักษาความหมายของคำเดิมไว้ นอกจากนี้ การลดความซับซ้อนของข้อความ (Text Simplification) ควรกำหนดให้คำจำกัดความมีความครอบคลุมมากขึ้น [5]

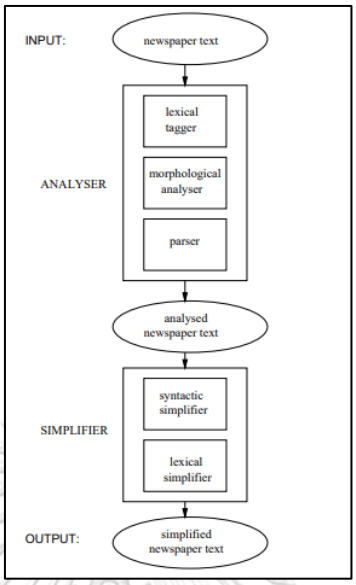
การลดความซับซ้อนของคำศัพท์ (Lexical Simplification) [6] เป็นกระบวนการย่อยกระบวนการหนึ่งของการลดความซับซ้อนของข้อความ (Text Simplification) โดยการทำงานในระดับคำ (Word-Level) เพื่อลดความซับซ้อน โดยการระบุและแทนคำซับซ้อนด้วยคำที่คำพ้องความหมาย โดยวิธีการดังกล่าวมี 4 ขั้นตอน ได้แก่ การระบุคำที่ซับซ้อน (Complex word Indentification) การสร้างคำเพื่อแทนที่ (Substitutions generation) การเลือกคำเพื่อแทนที่ (Substitution Selection) และ การจัดอันดับคำเพื่อแทนที่ (Substitution Ranking) รายละเอียดปรากฏดังภาพที่ 2



ภาพที่ 2 กระบวนการลดความซับซ้อนของคำศัพท์

การลดความซับซ้อนของคำโดยใช้กฎ (Rule – based lexical simplification) คือระบบการลดความซับซ้อนของคำหรือข้อความระบบแรกที่สามารถลดความซับซ้อนของข้อความ

ในหนังสือพิมพ์ภาษาอังกฤษเพื่อช่วยเหลือผู้อ่านที่มีความพิการทางสมอง โดยมีการทำงานดังภาพที่ 3 ประกอบไปด้วย 2 ส่วน ดังนี้ [7]



ภาพที่ 3 สถาปัตยกรรมของระบบการลดความซับซ้อน

ส่วนวิเคราะห์ (Analyser) ทำหน้าที่วิเคราะห์ด้วยการแยกวิเคราะห์ประโยคภาษาอังกฤษโดยใช้ Unication-based และ Part of Speech และเครื่องหมายวรรคตอนประกอบด้วย Lexical tagger Morphological Analyser และ Parser

ส่วนการทำข้อความให้ง่ายขึ้น (Simplifier) หลักจากวิเคราะห์ตามหลักภาษาศาสตร์แล้วจะเข้าสู่ขั้นตอน ประกอบด้วย Syntactic Simplifier คือการกำหนดกฎโดยใช้วากยสัมพันธ์เพื่อลดความซับซ้อนหรือทำให้เข้าใจง่าย และ Lexical Simplifier การวิเคราะห์ข้อความและสร้างคำพ้องความหมาย (Synonyms) และระบุความถี่ของแต่ละคำของคำพ้องความหมายนั้นๆ โดยคำที่มีความถี่สูงที่สุดจะเป็นคำที่ถูกเลือก

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

งานวิจัยในด้านการประมวลผลภาษาธรรมชาติได้เสนอวิธีที่ในการประมวลผลภาษาธรรมชาติ โดยนำเทคนิคต้นไม้ตัดสินใจ (Decision Tree) มาเป็นจำแนกประเภทของข้อความผลการวิจัยนั้น ระบุว่า กฎที่สร้างขึ้นด้วยเทคนิคด้านการประมวลผลภาษาธรรมชาติสำหรับการวิเคราะห์กลุ่มคำ การวิเคราะห์ชนิดกลุ่มคำ การวิเคราะห์ประโยคสามารถวิเคราะห์ได้ถูกต้อง [8]

วิธีการสรุปใจความสำคัญด้วย Term Frequency – Inverse Document Frequency (TF-IDF) [9] โดยใช้เป็นเครื่องมือสำหรับการกำหนดความสำคัญของคำหรือข้อความที่ปรากฏอยู่ในเอกสารสำหรับการสร้างระบบสรุปใจความสำคัญอัตโนมัติ โดยกำหนดจากความถี่ของคำที่ปรากฏ (Term Frequency) และค่าความผกผันในความถี่ของเอกสาร Inverse Document Frequency (IDF) เพื่อกำหนดน้ำหนักให้แต่ละคำที่ปรากฏในเอกสาร ผลลัพธ์ที่ได้จากงานวิจัยนี้พบว่า เมื่อนำ TF-IDF มาประยุกต์ใช้กับการสรุปใจความสำคัญสามารถให้ค่าความแม่นยำได้มากกว่าเครื่องมือสรุป

ใจความสำคัญอื่นๆ ที่มีอยู่ในออนไลน์ถึงร้อยละ 67 ทั้งนี้ ยังสามารถชี้ให้เห็นว่า TF-IDF สามารถกำหนดความสำคัญของคำต่างๆที่ปรากฏอยู่ในเอกสารได้ จนเกิดมาเป็นบทสรุป นอกจากนี้วิธีการสกัดคำสำคัญด้วย TF-IDF แล้วยังมีการใช้เทคนิค Latent Semantic Analysis (LSA) [10] เพื่อวิเคราะห์ความหมายที่แฝงอยู่ในข้อความ หรือจัดกลุ่มประเภทข้อความ ด้วยการสร้างเมทริกซ์ของคำศัพท์ที่ปรากฏในเอกสาร และแปลงเมทริกซ์ของคำศัพท์ทั้งหมดในเอกสารออกมาเป็นเมทริกซ์ย่อยๆ ด้วย Singular Value Decomposition (SVD) โดยผลลัพธ์ของงานวิจัยนี้ สรุปได้ว่าวิธีการสกัดคำสำคัญด้วย LSA นั้นมีประสิทธิภาพเทียบเท่ากับวิธีอื่นๆ

สำหรับการวิเคราะห์ข้อความในการรับแจ้งความหรือร้องทุกข์ของประชาชน และเกิดประโยชน์ต่อการดำเนินการของเจ้าหน้าที่ตำรวจในการสืบสวนสอบสวนอาชญากรรมที่เกิดขึ้นในสถานีตำรวจ ทั้งนี้ สามารถจัดการข้อมูลให้อยู่ในรูปแบบที่สามารถเข้าใจได้ง่าย โดยผ่านเทคนิคการประมวลผลดังกล่าว งานวิจัย [10] ได้เสนอวิธีการแยกประโยคจากเอกสาร เพื่อสร้างเอกสารสรุปหรือการสร้างบทคัดย่อเอกสารทั่วไป

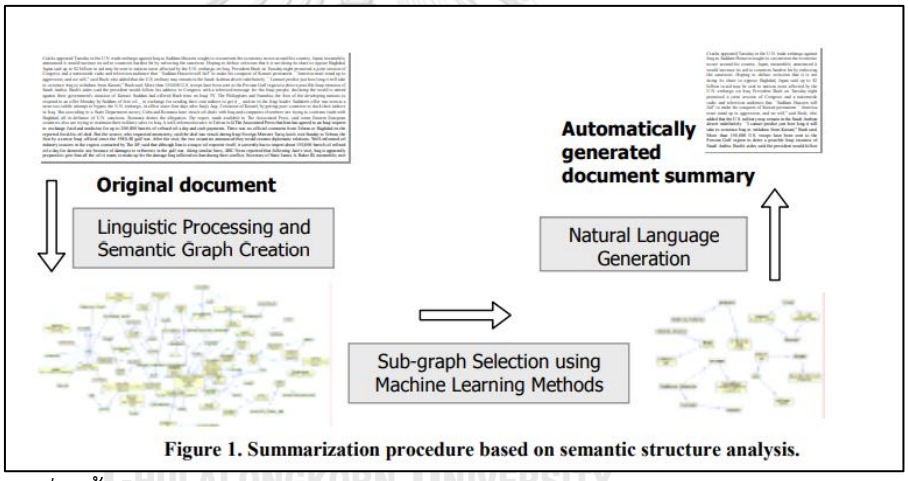


Figure 1. Summarization procedure based on semantic structure analysis.

ภาพที่ 4 ขั้นตอนการสร้างเอกสารสรุปด้วยการวิเคราะห์โครงสร้างความหมาย

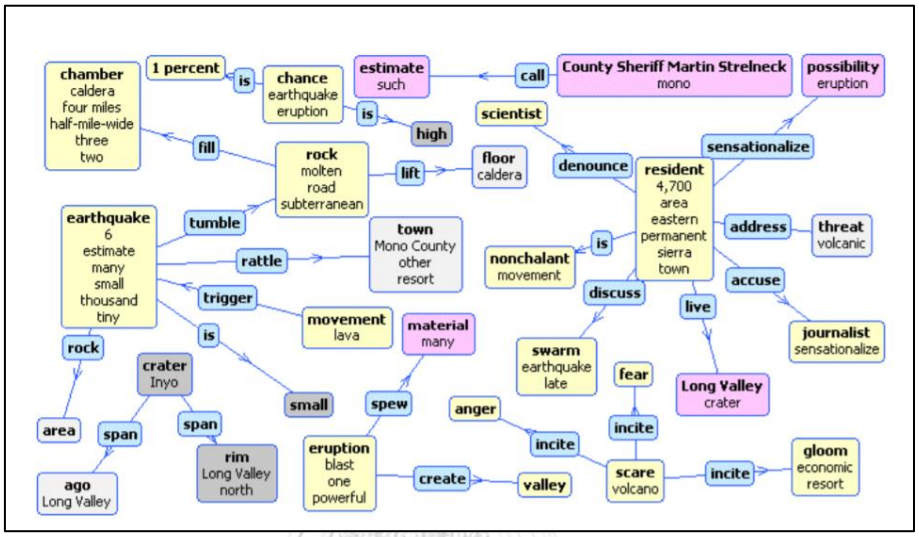
ด้วยวิธีการวิเคราะห์วากยสัมพันธ์ของข้อความที่สร้างการวิเคราะห์รูปแบบตรรกะสำหรับแต่ละประโยค โดยใช้ Subject-Objet-Predicate (SOP) จากแต่ละบุคคล เพื่อสร้างกราฟความหมายของเอกสารต้นฉบับ และข้อมูลสรุปที่ได้จากการนำเข้าข้อมูลที่เกี่ยวข้อง โดยกำหนดคุณลักษณะดังภาพที่ 4 ประกอบด้วย

1. คุณลักษณะทางภาษาศาสตร์ (Linguistic attributes) ได้แก่ การกำกับโครงสร้างประโยค (Part of speech tagging)

2. คุณลักษณะของกราฟเชิงความหมาย (Semantic graph attributes) คือการอธิบายคุณลักษณะของกราฟในแต่ละโหนด

3. โครงสร้างวาทกรรมของเอกสาร (Document Discourse Structure) คือการระบุตำแหน่งของประโยคในแต่ละเอกสาร รวมทั้ง ความถี่ และตำแหน่งของแต่ละคำในแต่ละประโยค

จากกลุ่มคุณลักษณะดังกล่าว จะนำมาประมวลผลด้วยโดยใช้วิธี Support Vector Machine เพื่อสร้างตัวจำแนกสำหรับการระบุ Subject-Objet-Predicate (SOP) โครงสร้าง และการย่อยคำหรือการลดความซับซ้อนของข้อความ (Text Simplification) จากเอกสาร

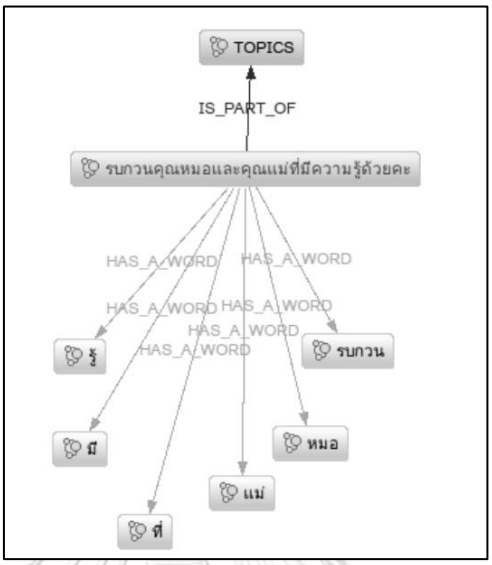


ภาพที่ 5 ภาพตัวอย่างกราฟแสดงความสัมพันธ์เชิงความหมาย

โดยผลการวิจัยของงานดังกล่าว ได้นำเสนอวิธีในการสรุปเอกสารซึ่งสามารถแสดงผลด้วยวิธีการสร้างกราฟแสดงความหมายของเอกสาร และนำไปประมวลผลด้วยกระบวนการเรียนรู้ด้วยเครื่อง โดยใช้เทคนิคซัพพอร์ตเวกเตอร์แมทชีน เพื่อแยกโครงสร้างดังภาพที่ 5 และ ย่อยความหมายหรือลดความซับซ้อน หรือปรับสารให้ง่าย ที่เหมาะสมสำหรับการสร้างรายงานสรุปผลเอกสารต่างๆ และในอนาคตจะมีการสำรวจความหมายทางเลือก หรือความอื่นๆ ของแต่ละคำ และข้อมูลสรุปปัญหาที่กว้างขึ้น รวมถึงการรวบรวมข้อมูลบทคัดย่อที่ถูกสร้างขึ้นด้วยมนุษย์ในแต่ละเอกสาร และการสรุปข้ามเอกสาร นอกจากนี้ งานวิจัย [11] เป็นการทดสอบการสกัดคำสำคัญจากบทย่อภาษาอังกฤษโดยใช้วิธีคำนวณค่า น้ำหนักของคำ (TF-IDF) เพื่อกำหนดคุณลักษณะที่ใช้ในการเลือกคำที่น่าจะเป็นคำสำคัญ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

งานวิจัย [12] ได้ใช้การแสดงผลแบบกราฟ โดยใช้ระบบฐานข้อมูลแบบกราฟ โดยแต่ละโหนด (Node) เชื่อมต่อกันด้วยความสัมพันธ์ทางความหมาย ของแต่ละคำที่ปรับสาร (Simplified) ออกมาแล้ว แสดงดังภาพที่ 6

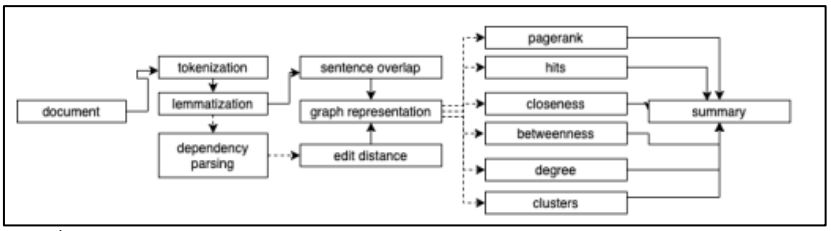


ภาพที่ 6 ภาพตัวอย่างกราฟการเชื่อมต่อกันด้วยความสัมพันธ์ทางความหมาย

งานวิจัย [13] การวิเคราะห์เครือข่ายทางสังคมของกระแสการเดินทางไปทำงานในภูมิภาคเมืองโคราช เสนอการนำกราฟหรือเครือข่ายหนึ่งๆ ที่ประกอบไปด้วยโหนด หรือ ศูนยรวมกิจกรรม และเส้นเชื่อมโยง (tiles) ระหว่างศูนย์รวมกิจกรรมหรือโหนดไม่ว่าจะเป็นตัวบุคคล ทีม องค์กร ฯลฯ วิธีการที่อยู่ในการวิเคราะห์เครือข่ายทางสังคมได้นำมาใช้เป็นสำหรับจัดระบบชุมชน โดยความหนาแน่นของเครือข่าย (Network Density) และค่ากลาง (Centrality) เพื่อหาว่าศูนย์กิจกรรมหรือโหนดใดมี “ความสำคัญ” ในเครือข่าย

วิธีการแยกสำหรับการสรุปข้อความโดยใช้กราฟ [14] อธิบายถึง การประมวลผลภาษาธรรมชาติซึ่งเครื่องมือที่สำคัญใน การนำเสนอข้อมูลที่มีลักษณะเป็นข้อความ ในรูปแบบแบบดิจิทัล (Digital Representation) เนื่องจากวิธีเขียนและพูดที่ปรากฏอยู่ในข้อความ หรือเอกสารทั่วๆไป นั้นมีความหลากหลายมาก ซึ่งทำให้ไม่สามารถระบุความถูกต้องของเอกสารได้อย่างแม่นยำ ดังนั้น งานวิจัยนี้ ได้ศึกษาอัลกอริทึมที่เกี่ยวข้องกับกราฟต่างๆ ที่สามารถนำใช้ได้ในการแก้ปัญหาการสรุปข้อความโดยใช้วิธีการสกัด (Extractive Approach) และใช้กราฟสำหรับการนำเสนอความสัมพันธ์ของแต่ละประโยค รวมทั้ง อัลกอริทึมที่เกี่ยวข้องสำหรับการระบุความสัมพันธ์ภายในกราฟ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University



ภาพที่ 7 แผนภาพแบบจำลองวิธีการแยกสำหรับการสรุปข้อความโดยใช้กราฟ

โดยการสรุปข้อความด้วยกราฟ (Graph-based Text Summarization) มีแบบจำลองดังภาพที่ 7 มีรายละเอียดหลัก ดังนี้

1. การประมวลผลข้อมูลเบื้องต้น (Data Pre-processing) ได้แก่ การโทเคนไนเซชัน (Tokenization) การหารากศัพท์ (Lemmatization) และการหาความสัมพันธ์แบบพึ่งพา (Dependency Parser)

2. เมทริกซ์ (Matrix) เมตริกหนึ่งที่ใช้สำหรับวัดความคล้ายคลึงของประโยคคือการตรวจสอบคำที่ทับซ้อนกัน

3. การสร้างกราฟ ที่ไม่มีทิศทาง (Undirected Graph) ประกอบด้วย โหนด สำหรับแต่ละประโยค และเส้นเชื่อม (Edge) ด้วยการคำนวณความคล้ายคลึงกันระหว่างประโยคแต่ละคู่

4. ขั้นตอนวิธี (Algorithm) ได้แก่ PageRank, Hits, Closeness, Betweenness เป็นเครื่องมือวัดความสำคัญของแต่ละประโยค

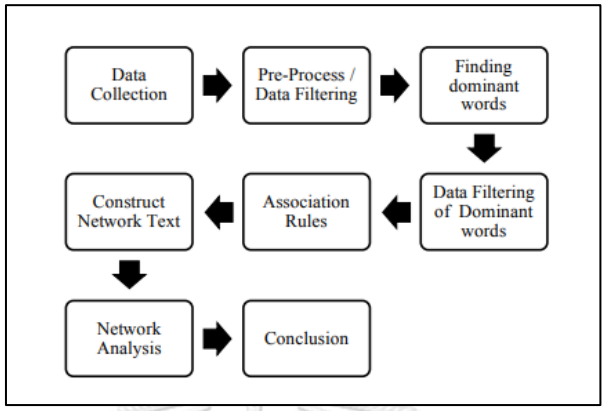
ผลงานวิจัยได้สรุปว่า ขั้นตอนวิธีแบบกราฟ (Graph-based Algorithm) สำหรับการคำนวณข้อมูลที่เกี่ยวข้องทั้งหมด สามารถสรุปเอกสารชุดทดสอบได้ดีที่สุด และช่วยแยกประโยคที่สำคัญที่สุดที่ จะประกอบเป็นบทสรุป

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

จากงานวิจัย [10] [13] [14] ที่เสนอกระบวนการสรุปข้อความด้วยวิธีการใช้กราฟ ประกอบกับการเรียนรู้ด้วยเครื่องนั้น พบว่า สามารถสรุปข้อความจากเอกสารได้โดยนำ SOP มาเป็นหลักในการแยกคำพร้อมกับ สรุปรายละเอียดของเอกสารตามแนวทาง SOP และ วิธีการหาความสำคัญของโหนด ที่อยู่ภายในเครือข่าย ด้วยกราฟ และมีการนำเทคนิคทางด้านการวิเคราะห์เครือข่ายทางสังคม (Social Network Analysis) เพื่อหาว่าโหนดใด มีอิทธิพลต่อเครือข่ายบ้าง เมื่อผู้วิจัยได้หลักการจากงานวิจัยทั้งสองแล้ว ผู้วิจัยเห็นว่า เทคนิคหรือขั้นตอนวิธีการดำเนินงานวิจัยของทั้งสอง ได้แก่ การสรุปข้อความ (Text Summarization) และ การวิเคราะห์เครือข่ายทางสังคม (Social Network Analysis) นั้น สามารถนำมาประยุกต์และใช้งานร่วมกันได้

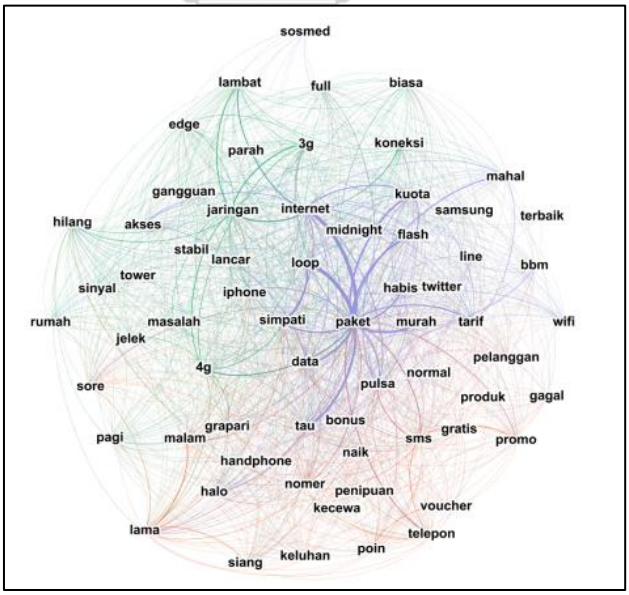
การวิเคราะห์ข้อความเครือข่ายเพื่อสรุปการสนทนาออนไลน์สำหรับความพยายามด้านการตลาดในอุตสาหกรรมโทรคมนาคม [15] ได้เสนอวิธีการสรุป ข้อมูลการสนทนาในรูปแบบข้อความเครือข่ายเพื่อรองรับการขยายตัวด้านการตลาด วิธีการนี้ให้มากขึ้น โดยวิธีการวิเคราะห์เครือข่าย

สังคม (SNA) เป็นวิธีการแก้ปัญหาในทางปฏิบัติสำหรับสร้างรูปแบบจากข้อมูลที่ไม่มีโครงสร้าง เนื่องจากความคล้ายคลึงของข้อมูลที่ไม่มีโครงสร้างในการวิเคราะห์ข้อความ และสามารถได้รับประโยชน์จากการสร้างการวิเคราะห์เครือข่ายสังคม



ภาพที่ 8 ขั้นตอนการทำงานของการวิเคราะห์เครือข่ายของข้อความ

จากภาพที่ 8 แสดงขั้นตอนการทำงานของการวิเคราะห์เครือข่ายของข้อความ โดยเริ่มจากการประมวลผลข้อมูลเบื้องต้น การกำจัดคำที่ไม่เกี่ยวข้อง อาทิ การลบคำสันธาน สรรพนาม ตัวเลข และการสร้างกราฟ โดยการใช้การวัดค่ากลาง (Centrality) เพื่อค้นหาคำที่ทรงอิทธิพลที่สุดในเครือข่ายข้อความเครือข่ายที่สร้างจากคู่คำ (Word Pairs) โดยคำนึงถึงระดับค่ากลาง



ภาพที่ 9 เครือข่ายข้อความตามคู่คำและกลุ่มคำ

ภาพที่ 9 สีของโหนดและเส้นเชื่อม (Edge) หมายถึงกลุ่ม (Cluster) คำที่ต่างกัน ขนาดของเส้นเชื่อมแปรผันตามน้ำหนักของความสัมพันธ์กันระหว่างโหนด ที่ถูกกำหนดด้วย กฎของความสัมพันธ์ (Association Rule) โดยบทสรุปของงานวิจัยนี้ สามารถสรุปได้ว่า วิธีการวิเคราะห์เครือข่าย

ทางสังคม (Social Network Analysis) สามารถช่วยวิเคราะห์ข้อความ และสร้างเครือข่าย ได้มีประโยชน์ในการสรุปข้อมูลจำนวนมาก และวิธีนี้ถูกพิสูจน์ว่า มีนัยยะในการเพิ่มประสิทธิภาพและประสิทธิผล สำหรับการสรุปการสนทนา เมื่อข้อมูลมีปริมาณเพิ่มขึ้น

2.2 ทฤษฎีที่เกี่ยวข้อง

สำหรับการสกัดกิจกรรมสำคัญจากคำพิพากษาศาลฎีกาเพื่อสร้างรายงานประจำวัน ด้วยคำพิพากษาศาลฎีกาค่อนข้างมีความซับซ้อน และมีข้อความที่หลากหลาย ผู้วิจัยจึงได้นำทฤษฎีที่เกี่ยวข้องกับการการลดความซับซ้อนของข้อความ (Text Simplification) มาเป็นเครื่องมือในการประมวลผลข้อมูลเบื้องต้น โดยทั้งนี้ ผู้วิจัยใช้แนวคิด ทฤษฎี เอกสาร เพื่อใช้เป็นแนวทางในการวิจัย ดังนี้

2.2.1 การลดความซับซ้อนของข้อความ (Text Simplification)

เป็นกระบวนการปรับเปลี่ยนภาษาธรรมชาติ เพื่อลดความซับซ้อน โดยนำข้อความมาประมวลผลด้วยกระบวนการตัดคำ และแทนที่ข้อความหรือคำดังกล่าว ด้วยคำที่มีความหมายเหมือนกันหรือมีความหมายคล้ายคลึงกัน [4]

2.2.2 การประมวลผลภาษาธรรมชาติ (Natural Language Processing)

การประมวลผลภาษาธรรมชาติ เป็นกระบวนการที่จะทำให้คอมพิวเตอร์เข้าใจภาษามนุษย์โดยรับอินพุตเป็นข้อควาในภาษาหนึ่งๆ แล้วทำความเข้าใจว่าผู้ป้อนข้อความกล่าวถึงอะไร [16] ซึ่งการใช้งาน NLP เกิดจากการนำทฤษฎีทางปัญญาประดิษฐ์มาใช้เพื่อวิเคราะห์และจำแนกคุณสมบัติหรือลักษณะทางภาษาศาสตร์ของประโยคหรือข้อความ

ขั้นตอนในการเข้าใจภาษาธรรมชาติ (Natural Language Understanding) มีขั้นตอนดังต่อไปนี้ [16]

1) การวิเคราะห์ทางองค์ประกอบ (Morphological Analysis) เป็นการวิเคราะห์ในหน่วยคำว่าคำๆ หนึ่งสามารถแยกออกเป็นหน่วยย่อยได้อะไรบ้าง เช่น ‘friendly’ มาจาก ‘friend’ กับ ‘ly’

2) การวิเคราะห์ทางวากยสัมพันธ์ (Syntactic Analysis) เป็นการวิเคราะห์ทางไวยากรณ์จุดมุ่งหมายก็เพื่อต้องการดูว่าประโยคที่รับเข้าซึ่งประกอบด้วยคำหลาย ๆ คำเรียงต่อกัน น้นมีโครงสร้างเชิงวากยสัมพันธ์เป็นอย่างไร

3) การวิเคราะห์ทางความหมาย (Semantic Analysis) เป็นการวิเคราะห์ความหมาย เมื่อได้โครงสร้างโดยการวิเคราะห์วากยสัมพันธ์มาแล้ว ก็จะกำหนดค่าของคำแต่ละคำว่าหมายถึงสิ่งใด

4) บูรณาการทางวจนิพน์ (Discourse integration) เป็นการพิจารณาความหมายของประโยคโดยดูจากประโยคข้างเคียง เนื่องจากคำบางคำในประโยคหนึ่งๆ จะเข้าใจความหมายได้ต้องดูประโยคก่อนหน้าหรือประโยคตามด้วย

5) การวิเคราะห์ทางปฏิบัติ (Pragmatic Analysis) เป็นการแปลความหมายของประโยคใหม่อีกครั้งว่าที่จริงแล้ว ผู้พูดตั้งใจจะหมายความว่าอย่างไรหรือต้องการสื่อความหมายอะไร

2.2.3 การตัดคำและการแปลงรูปคำ

เป็นหนึ่งในการวิเคราะห์ทางองค์ประกอบ (Morphological Analysis) การวิเคราะห์ในหน่วยคำ เนื่องจากภาษาไทยมีลักษณะที่เขียนตัวอักขระติดกันทั้งหมด ไม่มีเครื่องหมายวรรคตอนขั้นระหว่างคำ หากต้องการให้คอมพิวเตอร์ประมวลผลทางภาษา (Language Processing) จำเป็นต้องแบ่งคำออกเป็นส่วนๆ เพื่อแสดงถึงหน่วยของคำได้ โดยทั่วไปแล้วการตัดคำภาษาไทยมีวิธีการที่ใช้ในปัจจุบันเป็นหลักอยู่ 3 วิธี ได้แก่

1) การใช้กฎ
   การสร้างพยางค์ไทยที่ประกอบด้วยพยัญชนะ สระ วรรณยุกต์ ตัวสะกด ตัวการันต์ แนวทางนี้ทำได้ง่ายที่สุด ทำงานได้เร็วที่สุด แต่แบ่งคำพยางค์เดียวได้เท่านั้น ไม่สามารถจัดการกับคำหลายพยางค์ได้
2) การใช้พจนานุกรม
   ต้องทำรายการคำเอาไว้ล่วงหน้า เมื่อต้องการแบ่งคำก็เปรียบเทียบข้อความที่ต้องการแบ่งกับรายการคำที่เก็บไว้ในพจนานุกรม อาจพิจารณาตัดคำแบบยาวสุด คือการตัดคำที่เป็นไปได้ให้ได้คำยาวที่สุด หรือตัดคำให้ได้จำนวนคำน้อยสุด คือการตัดความโดยการเลือกรูปแบบการตัดที่จำนวนคำที่ตัดออกมาได้มีจำนวนน้อยที่สุด
3) การใช้เทคนิคการเรียนรู้ด้วยเครื่อง
   เป็นวิธีที่ได้รับความนิยมที่สุดในปัจจุบัน โดยการฝึกฝนระบบด้วยคลังข้อความขนาดใหญ่ที่มีการแบ่งคำไว้อย่างถูกต้องแล้ว เพื่อให้เครื่องได้เรียนรู้ด้วยตนเองจากการเก็บสถิติ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

การแปลงรูปแบบคำ เป็นการลดรูปแบบที่หลากหลาย (Word Variation) ของคำศัพท์ลงให้อยู่ในรูปแบบเดียวกัน ทำให้การตัดคำมีประสิทธิภาพ และสามารถระบุว่าคำใดเหมือนหรือแตกต่างกันได้ ในเอกสารภาษาไทยที่พบจริงคำที่มีความหมายเหมือนกันหรือคำเดยวกัน อาจเขียนได้หลายรูปแบบ กระบวนการที่นำมาใช้ได้แก่ การกำจัดช่องว่าง (Whitespace) การกำจัดอักขระพิเศษ การทำให้เป็นตัวอักษรพิมพ์เล็กสำหรับภาษาอังกฤษ

2.2.4 การกำกับโครงสร้างประโยค (Part of Speech Tagging)

ชนิดของคำพูด (Part of speech) คือการแบ่งชนิดของคำตามหน้าที่ โดยจะแบ่งได้เป็น 8 ชนิด ปรากฏตามตารางที่ 3 ได้แก่ คำนาม คำสรรพนาม คำกริยา คำกริยาวิเศษณ์ คำคุณศัพท์ คำบุพบท คำเชื่อม และคำอุทาน **ชนิดของคำพูด** จะทำให้เราสามารถเข้าใจ และใช้คำได้ถูกต้องมากขึ้น โดยการติดแท็กส่วนของคำพูด เป็นกระบวนการทำเครื่องหมายคำในข้อความไปยังส่วนของคำพูดโดยเฉพาะตามบริบทและคำจำกัดความ ในภาษาง่ายๆ เพื่อระบุส่วนของคำพูดและมักจะเป็นหมวดหมู่ทางไวยากรณ์อื่นๆ และใช้ในการค้นหาคลังข้อมูล และในเครื่องมือวิเคราะห์ข้อความและอัลกอริทึม

ตารางที่ 3 แสดงรายการชนิดของคำ (Part of Speech)

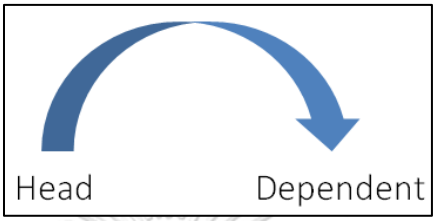
POS	DESCRIPTION
ADJ	adjective
ADP	adposition
ADV	adverb
AUX	auxiliary
CONJ	conjunction
CCONJ	coordinating conjunction
DET	determiner
INTJ	interjection
NOUN	noun
NUM	numeral
PART	particle
PRON	pronoun
PROPN	proper noun
PUNCT	punctuation
SCONJ	subordinating conjunction
SYM	symbol
VERB	verb

Chulalongkorn University

X	other
SPACE	space

2.2.5 การวิเคราะห์ประโยคเชิงพึ่งพิง (Dependency Parse)

เป็นการวิเคราะห์ความสัมพันธ์ของแต่ละคำในประโยคตามหลักภาษาไวยากรณ์ [17]



ภาพที่ 10 คำหลัก และคำพึ่งพา

ความสัมพันธ์แบบพึ่งพา (Dependency relation) จะกำกับระหว่างคำหลัก (Head) และคำพึ่งพา (Dependent) โดยคำหลัก จะเป็นตัวกำหนดว่า คำพึ่งพา ต้องเป็นคำชนิดใด มีความสัมพันธ์กันอย่างไร โดยลูกศรตั้งต้นมาจากคำหลัก และชี้ไปที่คำพึ่งพา ดังภาพที่ 10 และ 11 [18]

ตารางแสดงความสัมพันธ์รูปแบบต่าง ๆ ระหว่างคำ

	Nominals	Clauses	Modifier Words	Function Words
Core arguments	nsubj obj iobj	csubj ccomp xcomp		
Non-core dependents	obl vocative expl dislocated	advcl	advmod discourse	aux cop mark neg[2]
Nominal dependents	nmod appos nummod	acl	amod	det clf case
Coordination	MWE[1]	Loose	Special	Other
conj cc	fixed flat compound	list parataxis	orphan goeswith reparandum	punct root dep

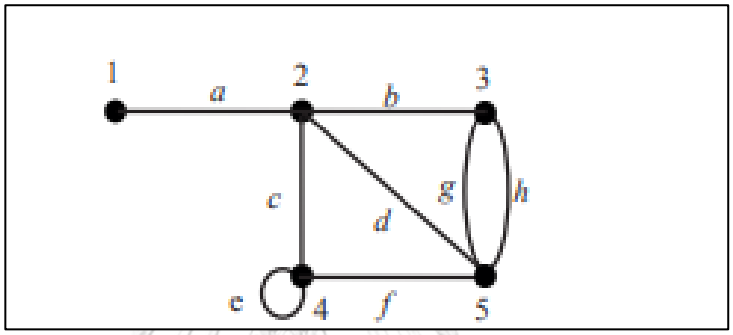
ภาพที่ 11 ความสัมพันธ์รูปแบบต่างๆ ระหว่างคำ

2.2.6 ทฤษฎีกราฟ (Graph Theory)

ในการแก้ไขโจทย์ปัญหา หรือเกมปริศนาบางข้อสามารถทำได้โดยง่ายด้วยการอธิบายปัญหานั้น ด้วยแผนภาพตัวแทน ปรากฏตามภาพที่ 12 ซึ่งเรียกว่า กราฟ (Graph) ประกอบด้วยจุดยอด (Vertex) และเส้นเชื่อม (Edge) ที่เชื่อมโยงบางจุดยอด นิยามของกราฟมีดังนี้ G = (V, E) ประกอบด้วยเซตจำกัด 2 เซต [19] ได้แก่

1) เซตของจุดยอด (V) ที่ไม่เป็นเซตว่างและเรียกสมาชิกใน V ว่า จุดยอด

2) เซตของเส้นเชื่อม (E) ที่อาจเป็นเซตว่างได้และ เรียกสมาชิกใน E ว่า เส้นเชื่อม



ภาพที่ 12 กราฟตัวแทนของการเชื่อมโยงเครือข่าย

2.2.7 ทฤษฎีเครือข่าย และ การสร้างเครือข่ายทางสังคม

เครือข่าย [20] หมายถึง โครงสร้างทางสังคมซึ่งเป็นความสัมพันธ์ของคน กลุ่ม หรือองค์การ ซึ่งส่งผลกระทบต่อความเชื่อและพฤติกรรม การศึกษาเครือข่ายจึงให้ความสนใจเรื่องค์ประกอบของความสัมพันธ์ที่เกิดขึ้นในเครือข่าย มากกว่าการศึกษาตัวองค์ประกอบของสมาชิก จึงถือว่าเป็นวิธีการมุ่งศึกษาความสัมพันธ์ และ การวิเคราะห์เครือข่ายทางสังคม (Social Network Analysis) เป็นกระบวนทัศน์ ที่ใช้วิธีการวิจัยแบบผสมผสานกันทั้งเชิงปริมาณและเชิงคุณภาพ ข้อมูลที่ใช้ในการวิเคราะห์เครือข่ายทางสังคมจะเป็นข้อมูลความสัมพันธ์ระหว่างสมาชิกใน เครือข่าย สามารถใช้ศึกษาได้ 3 ส่วน ได้แก่ เส้นเชื่อมโยง สมาชิกเครือข่าย และ เครือข่าย

2.2.8 การหาความสำคัญด้วยวิธีการทางกราฟ (Network Measure)

การวิเคราะห์การเชื่อมโยงของกราฟด้วยหลักการค่ากลาง (Centrality) [22] เป็นการวัดค่าความเป็นจุดศูนย์กลางของจุดยอดแต่ละจุด ซึ่งวิธีการที่นิยมนำมาใช้งานได้แก่

(1) ค่ากลางความเป็นศูนย์กลาง (Degree centrality) เป็นการวัดจำนวนเส้นความสัมพันธ์ของแต่ละบุคคลศูนย์กลาง โดยในเครือข่ายที่ศึกษาบุคคลศูนย์กลางที่มีค่าระดับความเป็นบุคคลศูนย์กลางสูงจะถือว่าเป็นบุคคล ศูนย์กลางของเครือข่าย ซึ่งเป็นบุคคลที่มีการเชื่อมโยงกับบุคคลอื่นในกลุ่มมากที่สุด

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

(2) ค่ากลางความใกล้ (Closeness centrality) เป็นการคำนวณผลรวมของเส้นความสัมพันธ์ (Ties) ที่ตรงที่สุดระหว่างบุคคลในเครือข่ายกับบุคคล ศูนย์กลางของเครือข่าย โดยการใช้การวัดระยะทาง ระหว่างบุคคลในเครือข่ายกับบุคคลศูนย์กลางของ เครือข่าย ซึ่งถ้าบุคคลศูนย์กลางใดมีค่าความใกล้ชิด กับบุคคลศูนย์กลางสูงแสดงว่า มีความสามารถใน การติดต่อสื่อสารและเข้าถึงข้อมูลของบุคคลในเครือ ข่ายได้อย่างรวดเร็ว และ

(3) ค่ากลางคั่นกลาง (Betweenness centrality) เป็นการคำนวณสัดส่วนของเส้นความสัมพันธ์ทั้งหมดที่มีการเชื่อมโยงผ่านไปจนถึงบุคคล ศูนย์กลางโดยบุคคลคั่นกลางกับบุคคลที่เป็น ศูนย์กลางของเครือข่ายจะเป็นบุคคลที่ใช้เป็นทาง ผ่าน ซึ่งบุคคลศูนย์กลางใดที่มีค่าสูงแสดงว่า มี ศักยภาพและความสามารถในการสื่อสารข้อมูลเชื่อม โยงผ่านไปจนถึงบุคคลศูนย์กลางอื่น

(4) ค่ากลางไอเกนเวกเตอร์ (Eigenvector Centrality) [23] หมายถึงความเป็นจุดศูนย์กลางโดยวัดจากไอเกนเวกเตอร์เฉพาะเป็นการคำ นวณค่าความเป็นจุดศูนย์กลางของเครือข่ายจากการวัดค่าอิทธิพลของจุดยอดในเครือข่าย โดยหากจุดยอดนั้นเชื่อมโยงกับจุดยอดอื่นที่มีค่าอิทธิพลสูงอยู่แล้ว ก็จะมีค่ากลางไอเกนเวกเตอร์ที่สูงกว่าจุดยอดที่เชื่อมโยงกับจุดยอดอื่นที่มีค่าอิทธิพลต่ำ

(5) เพจแรงก์ (PageRank Centrality) [21] เป็นวิธีการที่ใช้ในการจัดลำดับความสำคัญของหน้าเว็บจากกูเกิล (Google) โดยคำนวณจากหน้าเว็บหลาย ๆ หน้าที่เชื่อมโยงกันเป็นเครือข่าย การวัดค่าการจัดลำดับความสำคัญนี้กูเกิลได้กำหนดไว้ตั้งแต่ 0 ถึง 10 ยิ่งตัวเลขสูง ค่าเพจแรงก์จะยิ่งสูงเท่านั้น และเป็นเหตุให้ได้รับการจัดลำดับที่ดีกว่า วิธีการจัดลำดับ และการคำนวณค่าเพจแรงค์ นั้นจะทำการคำนวณจากเส้นเชื่อมของหน้าเว็บอื่นที่เชื่อมมายังหน้าเว็บนั้น (Inbound Link) โดยจะคำนึงถึงคุณภาพของเส้นเชื่อมที่มาเป็นสำคัญ ถ้าหากมีการเชื่อมต่อกบัหน้าเว็บที่มีเนื้อหาใกล้เคียงกันหรือคล้ายคลึงกันก็จะทำให้ค่าเพจแรงค์สูงขึ้น และหากหน้าเว็บที่ทำการเชื่อมมาหามีค่าเพจแรงค์สูง หน้าเว็บนั้นก็จะได้ค่าเพจแรงค์สูงไปด้วย

2.2.8 คำฝังตัว (Word Embedding)

1) คำฝังตัว [24] หรือการแปลงเชิงปริมาณที่เราแปลงคำ ให้อยู่ในรูปตัวเลข และนำไปสร้างเวกเตอร์คุณลักษณะ (Feature Vector) ขึ้นมาจากโทเคนที่เราได้ทำการสร้างไว้ โดยทำการสร้างเวกเตอร์คุณลักษณะขึ้นมาจากประโยคหรือเอกสารที่ มีอยู่ในข้อมูลของเรามาเพื่อทำการสร้างคุณลักษณะที่อยู่ในรูปของตัวเลขที่สามารถนำไปใช้ในการ คำนวณต่อได้ ซึ่งจุดเด่นของเวกเตอร์คุณลักษณะที่ได้จากคำฝังตัวนั้น จะสามารถนำไปใช้คำนวณความคล้ายคลึง (Similarity) กับคำอื่น ๆ ในบริบทของคำที่แตกต่างกันได้

2) Word2Vec [24] เป็นโมเดลแรกที่ถูกพัฒนาขึ้น โดย Google ที่มีการออกแบบโครงสร้างให้มีโครงข่ายประสาทเทียม 2 ชั้น โดยมีการนำข้อมูลจำนวนมากมาใช้ในการ ฝึกสอนให้ได้เวกเตอร์คุณลักษณะออกมา โดยเริ่มให้ได้ใช้งานครั้งแรกในปี2013 Word2Vec ส่วนมากถูกนำไปใช้กับงานทางด้าน Language Modeling ซึ่งเป็นแบบจำลองที่ใช้ใน การทำนายคำถัดไปของประโยคว่าควรที่จะเป็นคำใด โดยอาศัยหลักการของ คอนตินิวอัสแบกออฟเวิด (Continuous Bag of Words) สำหรับการใช้คำหลาย ๆ คำต่อกัน เพื่อทำนายคำที่อยู่ถัดไป และ คอนตินิวอัสสคริปแกรม (Continuous Skip-Gram) สำหรับการใช้คำหนึ่งคำในการทำนายคำอื่น ๆ ที่มีโอกาสเป็นคำถัดไปจากคำนี้

3) Thai2Vec [25] โมเดลที่ถูกฝึกฝนด้วย Word2Vec บนคลังคำศัพท์วิกิพีเดีย มีเวกเตอร์คำจำนวน 51,556 เวกเตอร์ พร้อมมิติข้อมูล 300 รายการ การทำ Embedding ของ Thai2Vec มีวิธีการคือ กำหนดให้มีจำนวนคำ OOV (Out-of-Vocabulary) คือ คำศัพท์ที่ไม่ได้อยู่ในพจนานุกรมทั่วไปที่พบในสภาพแวดล้อมการประมวลผลภาษาธรรมชาติ น้อยที่สุด

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บทที่ 3
แนวคิดและวิธีการดำเนินการวิจัย

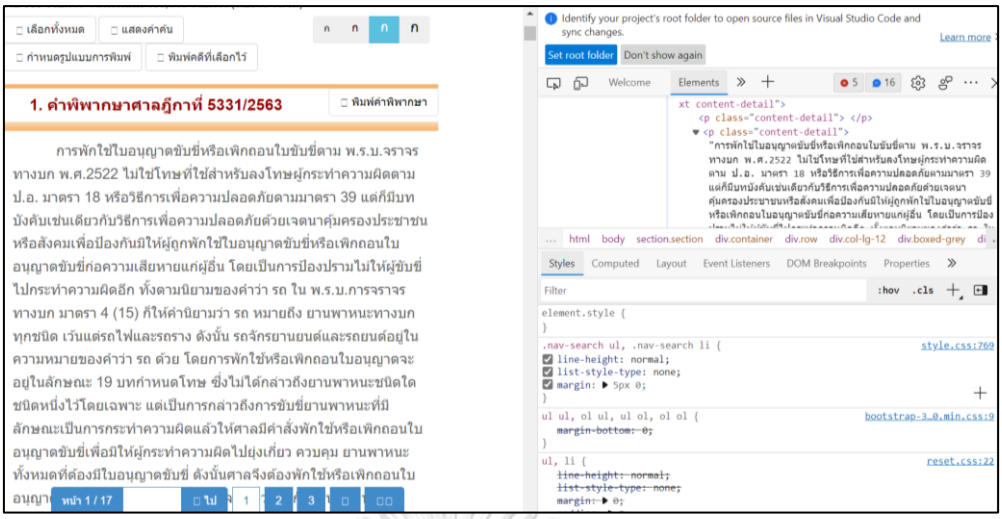
วิทยานิพนธ์นี้มีวัตถุประสงค์ในการสกัดกิจกรรมสำคัญจากคำพิพากษาศาลฎีกา เพื่อนำมาสร้างรายงานประจำวันสำหรับตำรวจ เพื่อแก้ไขปัญหาการบันทึกรายงานประจำวันเกี่ยวกับคดีจราจรในสถานีตำรวจ สามารถสร้างรายงานประจำวันเกี่ยวกับคดีจราจร เพื่อให้พนักงานสอบสวนหรือเจ้าหน้าที่ตำรวจสามารถนำกระบวนการดังกล่าวไปใช้กับการบันทึกรายงานประจำวันเกี่ยวกับคดีจราจรที่มีการบันทึกอยู่ในสถานีตำรวจทุกสถานีที่มีอำนาจสอบสวน โดยใช้ชุดข้อมูลสำหรับสกัดกิจกรรมสำคัญจากคำพิพากษาศาลฎีกา แล้วระบุถึงเอนทิตี้ที่สำคัญ รายละเอียดพฤติการณ์แห่งคดีที่พนักงานสอบสวนจำเป็นต้องบันทึกลงในรายงานประจำวัน และสร้างแบบรายงานประจำวันขึ้นมาเพื่อให้พนักงานสอบสวนหรือเจ้าหน้าที่ตำรวจผู้บันทึกสามารถลดเวลาการบันทึกแบบรายงานประจำวัน และบริการประชาชนได้รวดเร็วยิ่งขึ้น โดยแนวทางการดำเนินการวิจัยสามารถสรุปภาพรวมการทำงานได้ดังภาพที่ 13

ภาพที่ 13 แสดงแผนภาพการทำงานโดยรวม

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

3.1 การเตรียมข้อมูลสำหรับงานวิจัย (Data Preparation)

ผู้วิจัยได้สืบค้นคำพิพากษาศาลฎีกาจากเว็บไซต์เผยแพร่คำพิพากษาศาลฎีกา [26] โดยเลือกเฉพาะคำพิพากษาศาลฎีกาที่เกี่ยวข้อง ตาม พ.ร.บ.การจราจรทางบก พ.ศ.2522 ผู้วิจัยบันทึกไฟล์ประเภท Html ที่กรองผ่านหน้าเว็บจำนวนทั้งสิ้น 330 คำพิพากษาศาลฎีกา



ภาพที่ 14 แสดงชุดคำสั่ง HTML ของเว็บสืบค้นคำพิพากษาศาลฎีกา

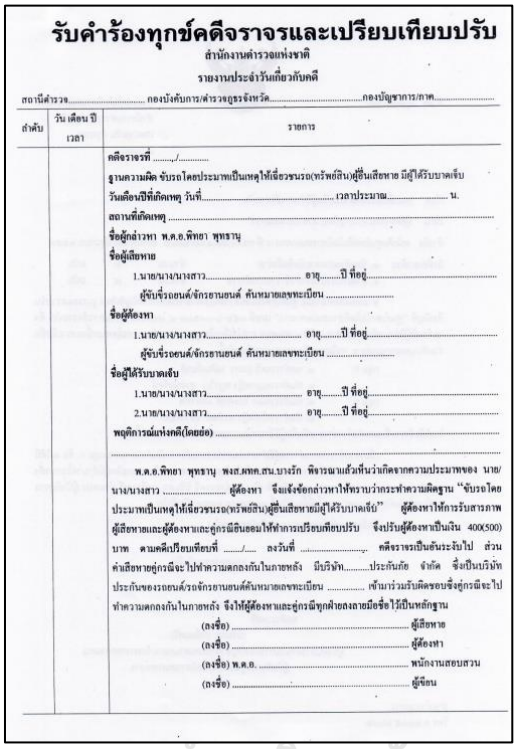
และทำการดึงข้อมูลจากเว็บไซต์ (Web Scraping) โดยใช้ BeautifulSoup ซึ่งเป็นโมดูลหนึ่งของภาษาไพทอน ในการดึงข้อมูลเฉพาะแท็ก (Tag) และคลาส (Class) รายละเอียดปรากฏตามภาพที่ 14 ที่มีเนื้อหาที่จำเป็นสำหรับการทำงานวิจัย เพื่อใช้ในการนำเข้าข้อมูลคำพิพากษาศาลฎีกา ประกอบด้วย เนื้อหาคำพิพากษาศาลฎีกา โดยเลือกเฉพาะคำพิพากษาศาลฎีกาแบบย่อ เนื้อหาคำพิพากษาศาลฎีกา เนื้อหาการตัดสิน และมาตราที่ใช้ปรับบทลงโทษ เมื่อผู้วิจัยสามารถดึงข้อมูลคำพิพากษาศาลฎีกามาได้แล้ว พบว่ามีจำนวนข้อมูลซ้ำอยู่จำนวนหนึ่ง ผู้วิจัยได้กำจัดโดยลบข้อมูลคำพิพากษาศาลฎีกาแต่ละรายการที่ซ้ำออก นอกจากนี้ยังพบว่ากรณีคำพิพากษาศาลฎีกาที่ถูกพิพากษาขึ้นก่อนที่จะมีการประกาศใช้ พ.ร.บ.จราจรทางบก พ.ศ.2522 จำเป็นต้องถูกตัดออกไปด้วย เนื่องจากรายละเอียดคำย่อคำพิพากษาศาลฎีกานั้น มิได้สอดคล้องกับบทบัญญัติกฎหมายที่ผู้วิจัยเลือกใช้ในการวิจัย

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

การศึกษาแบบรายงานประจำวันเกี่ยวกับคดี ที่มีการบันทึกลงในสมุดรายงานประจำวัน เพื่อศึกษารายละเอียดของข้อความที่มีการบันทึกลงในสมุดรายงานประจำวัน เพื่อกำหนดคุณลักษณะหรือคำสำคัญของข้อความที่ได้นำเข้ามา ดังนี้

1) ส่วนของข้อมูลคดี
   ประกอบด้วย เลขคดี และ ปีคดี ตัวอย่าง คดีอาญาที่ 1/2564
2) ส่วนของฐานความผิด
3) ส่วนของวันเวลา และสถานที่เกิดเหตุ
4) ข้อมูลผู้เสียหาย
5) พฤติการณ์แห่งคดี

จากที่กล่าวมาเป็นตัวอย่างประเภทข้อมูลที่มีการบันทึกลงในแบบรายงานประจำวัน ดังภาพที่ 15 จะเห็นได้ว่า ข้อมูลจะประกอบด้วย ข้อมูลฐานความผิด ข้อมูลสถานที่เกิดเหตุ และข้อมูลของกลาง เป็นหลัก และในส่วนที่สำคัญที่สุดคือ พฤติการณ์แห่งคดี จะเป็นการบันทึกข้อมูลรายละเอียดของคดีว่ามีลักษณะ ลำดับเหตุการณ์เป็นอย่างไร ซึ่งเป็นการบ่งบอกรายละเอียด ลำดับเหตุการณ์ ของสิ่งที่เกิดขึ้น



รับคำร้องทุกข์คดีจราจรและเปรียบเทียบปรับ

สำนักงานตำรวจแห่งชาติ

รายงานประจำวันเกี่ยวกับคดี

สถานีตำรวจ..... กองบังคับการ/ตำรวจภูธรจังหวัด..... กองบัญชาการ/กทม.....

ลำดับ	วัน เดือน ปี เวลา	รายการ
		คดีจราจรที่ ...../.....
		ฐานความผิด ขับรถโดยประมาทเป็นเหตุให้เฉี่ยวชนรถ(ทรัพย์สิน)ผู้อื่นเสียหาย มีผู้ได้รับบาดเจ็บ
		วันเดือนปีที่เกิดเหตุ วันที่..... เวลาประมาณ..... น.
		สถานที่เกิดเหตุ .....
		ชื่อผู้กล่าวหา พ.ต.อ.พิทยา พุทธานุ
		ชื่อผู้เสียหาย
		1.นาย/นาง/นางสาว..... อายุ.....ปี ที่อยู่.....
		ผู้ขับขี่รถยนต์/จักรยานยนต์ ทะเบียน .....
		ชื่อผู้ต้องหา
		1.นาย/นาง/นางสาว..... อายุ.....ปี ที่อยู่.....
		ผู้ขับขี่รถยนต์/จักรยานยนต์ ทะเบียน .....
		ชื่อผู้ได้รับบาดเจ็บ
		1.นาย/นาง/นางสาว..... อายุ.....ปี ที่อยู่.....
		2.นาย/นาง/นางสาว..... อายุ.....ปี ที่อยู่.....
		พฤติการณ์แห่งคดี(โดยย่อ) .....
		พ.ต.อ.พิทยา พุทธานุ พงส.สภค.สน.บางรัก พิจารณาแล้วเห็นว่าเกิดจากความประมาทของ นาย/นาง/นางสาว ..... ผู้ต้องหา จึงแจ้งข้อกล่าวหาให้ทราบว่ากระทำความผิดฐาน "ขับรถโดยประมาทเป็นเหตุให้เฉี่ยวชนรถ(ทรัพย์สิน)ผู้อื่นเสียหายมีผู้ได้รับบาดเจ็บ" ผู้ต้องหาให้การรับสารภาพ ผู้เสียหายและผู้ต้องหาและคู่กรณียินยอมให้ทำการเปรียบเทียบปรับ จึงปรับผู้ต้องหาเป็นเงิน 400(500) บาท ตามคดีเปรียบเทียบที่ ...../..... ลงวันที่ ..... คดีจราจรเป็นอันระงับไป ส่วนค่าเสียหายคู่กรณีจะไปทำความตกลงกันในภายหลัง มีบริษัท.....ประกันภัย จำกัด ซึ่งเป็นบริษัทประกันของรถยนต์/รถจักรยานยนต์คันหมายเลขทะเบียน ..... เข้ามาร่วมรับผิดชอบซึ่งคู่กรณีจะไปทำความตกลงกันในภายหลัง จึงให้ผู้ต้องหาและคู่กรณีทุกฝ่ายลงลายมือชื่อไว้เป็นหลักฐาน
		(ลงชื่อ) ..... ผู้เสียหาย
		(ลงชื่อ) ..... ผู้ต้องหา
		(ลงชื่อ) พ.ต.อ. ..... พนักงานสอบสวน
		(ลงชื่อ) ..... ผู้เขียน

ภาพที่ 15 ตัวอย่างแบบรายงานประจำวันเกี่ยวกับคดี

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

3.2 การทำความสะอาดข้อมูล (Data Cleansing)

ข้อมูลที่ได้รวบรวมมานั้น อาจมีบางข้อความที่มีการใช้อักขระพิเศษ หรือตัวเลขที่ไม่ได้เกี่ยวข้องกับข้อความแต่อย่างใด ผู้วิจัยจึงต้องการกำจัดอักขระพิเศษออกจากข้อความ ปัญหาของการวิเคราะห์ภาษาไทยที่มาจากการไม่มีตัวอักษรหรือสัญลักษณ์ที่แบ่งแยกคำแต่ละคำภายในข้อความหรือประโยค การที่จะสามารถจำแนกข้อความโดยที่ไม่ใช่การเปรียบเทียบความเหมือนของทุกลำดับตัวอักษรหรือการเปรียบเทียบข้อความนั้น จำเป็นที่จะต้องแยกหน่วยของข้อความให้เล็กลง เพื่อจะสามารถเปรียบเทียบในเชิงส่วนประกอบของโครงสร้างข้อความได้ การสกัดคำพิพากษาศาลฎีกาเพื่อนำมาสร้างรายงานประจำวันสำหรับตำรวจนั้น ซึ่งเมื่อมองในภาพรวมการตัดสินคดีความหรือพิพากษาคดีความจะเป็นขั้นตอนสุดท้ายของกระบวนการยุติธรรม โดยขั้นตอนแรกของกระบวนการยุติธรรมจะอยู่ที่ตำรวจหรือพนักงานสอบสวน เป็นผู้รับแจ้งความ/คำร้องทุกข์ของประชาชน โดยจุดประสงค์ของงานวิจัยนี้ คือการสกัดคำสำคัญเพื่อนำมาสร้างรายงานประจำวัน โดยจะมุ่งเป้าหมาย

ไปที่รายละเอียดของข้อมูลที่อยู่ในแบบรายงานประจำวัน เพราะฉะนั้น ข้อมูลคำพิพากษาศาลฎีกาที่มีอยู่นั้น จะเป็นส่วนของศาลเท่านั้น และจะมีคำใช้ในศาลปรากฏอยู่จำนวนมาก จึงจำเป็นต้องกำจัดคำที่มักปรากฏอยู่ในชั้นศาลออกเสียก่อน ก่อนที่จะดำดำเนินการในขั้นตอนต่อไป ผู้วิจัยได้รวบรวมคำศัพท์จากคู่มือล่ามในกระบวนการยุติธรรม [27] โดยเลือกเฉพาะคำศัพท์ในชั้นพิจารณาคดี และในชั้นคดีถึงที่สุด เพื่อกำจัดคำดังกล่าวๆ ในคำพิพากษาศาลฎีกาที่ผู้วิจัยได้รวบรวมมา ได้ตัวอย่างตามภาพที่ 16

'กรรโชกทรัพย์', 'กระทง', 'การกระทำโดยงดเว้น', 'การกระทำผิด ด้วยความจำเป็น', 'การก่อการร้าย', 'การขัดกันระหว่าง ประโยชน์ส่วนบุคคล และประโยชน์ส่วนรวม', 'การควบคุมตัว', 'การควบคุมผู้ต้องขัง', 'การคุ้มครองผู้แจ้งเบาะแส', 'การฆาตกรรม', 'การฉ้อโกง', 'การตรวจพยานหลักฐาน', 'การตรวจสอบทรัพย์สิน และหนี้สิน', 'การไต่สวนข้อเท็จจริง', 'การถอดถอนจากตำแหน่ง', 'การทำร้ายร่างกาย', 'การประชุมคดี', 'การปล่อยตัวชั่วคราว', 'การยักยอกทรัพย์', 'การรับทรัพย์สิน หรือประโยชน์อื่นใด โดยธรรมจรรยา'

ภาพที่ 16 ตัวอย่างคำที่มักปรากฏในชั้นศาล

3.3 การประมวลผลข้อมูลเบื้องต้น (Data Preprocessing)

การสร้างกราฟจากคำพิพากษาศาลฎีกาที่ผ่านกระบวนการประมวลผลเบื้องต้นแล้วเพื่อจับใจความ คำสำคัญ ความสัมพันธ์ทางความหมายที่สำคัญ และสกัดรายละเอียดสำคัญที่จำเป็น ทั้งนี้ขั้นตอนการสร้างกราฟเชิงความประกอบด้วย 3 ขั้นตอน ดังนี้

3.3.1 การวิเคราะห์เชิงภาษาศาสตร์ (Syntactic Analysis) เพื่อสกัดคุณลักษณะทางภาษาศาสตร์ ได้แก่ ของแต่ละประโยคในเอกสาร ด้วยการแบ่งข้อความออกเป็นคำ ด้วยการตัดคำภาษาไทย ในการฝึกฝนแบบจำลองสำหรับการเรียนรู้ด้วยเครื่องนั้น ข้อมูลที่ใช้ในงานวิจัยเป็นข้อความภาษาไทยที่เป็นรูปแบบการเขียนติดกัน จึงจำเป็นต้องแยกแต่ละคำที่ข้อความหรือเอกสารออกจากกันเพื่อนำไปใช้ประโยชน์ในการประมวลผลดังกล่าว เพื่อให้เข้าถึงข้อมูลระดับคำ (Word Level) โดยการตัดคำใช้เครื่องมือตัดคำ Newmm [28] ของ PythaiNLP ซึ่งเป็นการแบ่งกลุ่มคำภาษาไทยตามพจนานุกรม โดยใช้อัลกอริทึมการจับคู่สูงสุด (Maximum Matching) และคลัสเตอร์อักขระภาษาไทย (Thai Character Cluster) โดยมีตัวอย่างดังตารางที่ 4 และ 5

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตารางที่ 4 ตัวอย่างการตัดคำภาษาไทยด้วย PythaiNLP

ข้อความ	ผลลัพธ์การตัดคำด้วย PythaiNLP
โจทก์ฟ้องว่าจำเลยขับรถด้วยความประมาทน่าหวาดเสียว อันอาจเกิดอันตรายแก่บุคคลหรือทรัพย์สินของผู้อื่น และปราศจากความระมัดระวังซึ่งบุคคลในภาวะเช่นนั้นจักต้องมีตามวิสัยและพฤติการณ์ เป็นเหตุให้รถยนต์ที่จำเลยขับเฉี่ยวชนกับรถยนต์กระบะ	'โจทก์','ฟ้อง', 'ว่า', 'จำเลย', 'ขับรถ', 'ด้วย', 'ความประมาท', 'น่าหวาดเสียว', ' ', 'อัน', 'อาจ', 'เกิด', 'อันตราย', 'แก่', 'บุคคล', 'หรือ', 'ทรัพย์สิน', 'ของ', 'ผู้อื่น', ' ', 'และ', 'ปราศจาก', 'ความระมัดระวัง', 'ซึ่ง', 'บุคคล', 'ใน', 'ภาวะ', 'เช่นนั้น', 'จัก', 'ต้อง','มี', 'ตาม', 'วิสัย', 'และ', 'พฤติการณ์', ' ', 'เป็นเหตุให้', 'รถยนต์', 'ที่', 'จำเลย', 'ขับ', 'เฉี่ยว', 'ชน', 'กับ', 'รถยนต์', 'กระบะ'

ตารางที่ 5 ตัวอย่างการตัดประโยคภาษาไทยด้วย PythaiNLP

ข้อความ	ผลลัพธ์การตัดประโยคด้วย PythaiNLP
โจทก์ฟ้องว่าจำเลยขับรถด้วยความประมาทน่าหวาดเสียว อันอาจเกิดอันตรายแก่บุคคลหรือทรัพย์สินของผู้อื่น และปราศจากความระมัดระวังซึ่งบุคคลในภาวะเช่นนั้นจักต้องมีตามวิสัยและพฤติการณ์ เป็นเหตุให้รถยนต์ที่จำเลยขับเฉี่ยวชนกับรถยนต์กระบะ	ประโยคที่ 1 โจทก์ฟ้องว่าจำเลยขับรถด้วยความประมาทน่าหวาดเสียว ประโยคที่ 2 อันอาจเกิดอันตรายแก่บุคคลหรือทรัพย์สินของผู้อื่น และปราศจากความระมัดระวังซึ่งบุคคลในภาวะเช่นนั้นจักต้องมีตามวิสัยและพฤติการณ์ ประโยคที่ 3 เป็นเหตุให้รถยนต์ที่จำเลยขับเฉี่ยวชนกับรถยนต์กระบะ

1) การกำจัดคำฟุ่มเฟือย เริ่มจากการกำจัดคำที่ปรากฏบ่อยๆ หรือคำฟุ่มเฟือย (Stop Word) เพื่อให้เหลือแต่คำสำคัญ (Keyword) ของแต่ละข้อความ โดยใช้แพ็กเกจ corpus.thai_stopword ของ PythaiNLP [29] ซึ่งมีการรวบรวมคำภาษาไทยที่มักจะปรากฏบ่อยอยู่จำนวนกว่า 1000 รายการ ทั้งนี้ เพื่อต้องการให้ข้อความเหลือเพียงคำสำคัญ และเป็นการจำกัดขนาดของข้อมูลไม่ให้มีขนาดใหญ่จนเกินความจำเป็น รวมทั้งคำศัพท์ที่เกี่ยวข้องกับกระบวนการยุติธรรมหรือในชั้นพิจารณาคดี ที่ได้กล่าวไว้ในข้อ 3.2 โดยตัวอย่างคำฟุ่มเฟือย ปรากฏดังภาพที่ 17

เกี่ยวกับ', 'ก็ตามที', 'พวกนี้', 'ทั้งปวง', 'ทั้งสิ้น', 'แล้วกัน', 'ยิ่งจะ', 'สามารถ', 'ง่าย', 'ตลอดปี', 'ครั้งๆ', 'ซึ่ง', 'นับแต่นี้', 'ทันทีทันใด', 'ที่แล้ว'

ภาพที่ 17 ตัวอย่างคำฟุ่มเฟือยภาษาไทย (Thai Stopword)

2) การรู้จำชื่อเอนทิตี้ (Named-Entity Recognition) เป็นส่วนหนึ่งของการสกัดข้อมูลที่พยายามค้นหาและจัดประเภทของคำ ที่มีชื่อที่กล่าวถึงในข้อความที่ไม่มีโครงสร้างเป็นหมวดหมู่ที่กำหนดไว้ กระบวนการรู้จำในการหาตำแหน่ง และระบุประเภทหรือขอบเขต หรือชื่อเฉพาะของสิ่งต่างๆ ที่อยู่ในข้อความหรือเอกสาร แสดงดังตารางที่ 6

ตารางที่ 6 ตัวอย่างชื่อกลุ่มของ Named-Entity Tag ใน PythaiNLP

Named Entity tag	Examples
DATE	2/21/2004, 16 ก.พ., จันทร์
TIME	16.30 น., 5 วัน, 1-3 ปี
EMAIL	info@nrpsc.ac.th
LEN	30 กิโลเมตร, 5 กม.
LOCATION	ไทย, จ.ปราจีนบุรี, กำแพงเพชร
ORGANIZATION	กรมวิทยาศาสตร์การแพทย์, อย.
PERSON	น.พ.จรัล, นางประนอม ทองจันทร์
PHONE	1200, 0 2670 8888
URL	http://www.bangkokhealth.com/
ZIP	10400, 11130
Money	2.7 ล้านบาท, 2,000 บาท
LAW	พ.ร.บ.โรคระบาด พ.ศ.2499, รัฐธรรมนูญ

เมื่อผู้วิจัยสามารถระบุชื่อเอนทิตี้ ได้แล้ว จะได้ข้อความที่ติดแท็กของเอนทิตี้ประเภทต่างๆ ทั้งนี้ ผู้วิจัยต้องการเพียงรายละเอียดการกระทำความผิดที่มีอยู่ในคำ

พิพากษาศาลฎีกาเท่านั้น จากภาพที่ 18 จะเห็นได้ว่ามีการติดแท็ก ในส่วนของคำพิพากษาศาลฎีกา โดยระบุเป็นกลุ่มของเอนทิตี้ต่างๆ ยกตัวอย่าง <LAW></LAW> เป็นแท็ก ที่ระบุว่า คำหรือข้อความส่วนดังกล่าว เป็นชื่อของกฎหมาย ซึ่งจุดประสงค์คือต้องการเพียงแต่รายละเอียดการกระทำความผิดเท่านั้น ผู้วิจัยจึงกำจัดข้อความส่วนที่ไม่เกี่ยวข้องกับรายละเอียดดังกล่าวออก

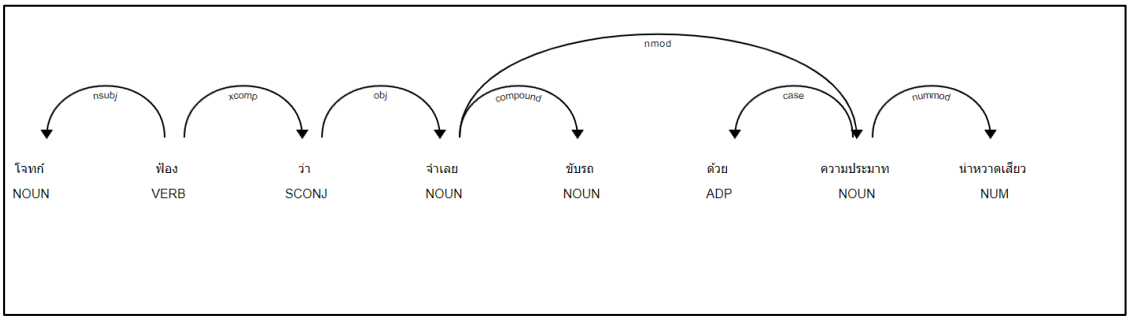
<PERSON>นาย ส.</PERSON> การพักใช้ใบอนุญาตขับขี่หรือเพิกถอนใบขับขี่ตาม <LAW>พ.ร.บ.จราจรทางบก พ.ศ.2522</LAW> ไม่ใช่โทษที่ใช้สำหรับลงโทษผู้กระทำความผิดตาม 18 หรือวิธีการเพื่อความปลอดภัยตามมาตรา 39 แต่ก็มีบทบังคับเช่นเดียวกับวิธีการเพื่อความปลอดภัยด้วยเจตนาคุ้มครองประชาชนหรือสังคม เพื่อป้องกันมิให้ผู้ถูกพักใช้ใบอนุญาตขับขี่หรือเพิกถอนใบอนุญาตขับขี่ก่อความเสียหายแก่ผู้อื่น โดยเป็นการป้องปรามไม่ให้ผู้ขับขี่ไปกระทำความผิดอีก ทั้งตามนิยามของคำว่า รถ ใน <LAW>พ.ร.บ.การจราจรทางบก มาตรา 4 (15)</LAW> ก็ให้คำนิยามว่า รถ หมายถึง <ORGANIZATION> ยานพาหนะทางบกทุกชนิด เว้นแต่รถไฟและรถราง</ORGANIZATION> ดังนั้น รถจักรยานยนต์และรถยนต์อยู่ในความหมายของคำว่า รถ ด้วย โดยการพักใช้หรือเพิกถอนใบอนุญาตจะอยู่ในลักษณะ <MONEY>19 บทกำหนดโทษ</MONEY> ซึ่งไม่ได้กล่าวถึงยานพาหนะชนิดใดชนิดหนึ่งไว้โดยเฉพาะ แต่เป็นการกล่าวถึงการขับขี่ยานพาหนะที่มีลักษณะเป็นการกระทำความผิดแล้วให้ศาลมีคำสั่งพักใช้หรือเพิกถอนใบอนุญาตขับขี่เพื่อมิให้ผู้กระทำความผิดไปยุ่งเกี่ยวควบคุม ยานพาหนะทั้งหมดที่ต้องมีใบอนุญาตขับขี่ ดังนั้นศาลจึงต้องพักใช้หรือเพิกถอนใบอนุญาตขับขี่ ไม่ว่าใบอนุญาตขับขี่นั้นจะเกี่ยวข้องกับยานพาหนะที่ผู้กระทำความผิดขับขี่ในขณะกระทำความผิดหรือไม่ก็ตาม แม้จำเลยมีใบอนุญาตขับขี่รถคนละประเภทกับรถที่จำเลยขับขณะกระทำความผิด ศาลก็มีอำนาจพักใช้ใบอนุญาตขับขี่หรือเพิกถอนใบอนุญาตขับขี่ของจำเลยได้

ภาพที่ 18 แสดงการรู้จำเอนทิตี้ด้วย PythaiNLP

3) การวิเคราะห์ทางภาษาศาสตร์ (Linguistic Analysis) โดยใช้เครื่องมือการประมวลผลภาษาธรรมชาติ ระบุชนิดของคำพูด (Part of speech) ของแต่ละคำที่ปรากฏในคำพิพากษาศาลฎีกาทั้ง ประกอบกับ โครงสร้างประโยคที่ประกอบด้วย ประธาน-กริยา-กรรม (Subject-Verb-Object) ซึ่งเป็นโครงสร้างประโยคที่ประธานมาก่อนเป็นลำดับแรก กริยาเป็นลำดับที่สอง และวัตถุเป็นลำดับที่สาม ผู้วิจัยได้เลือกเฉพาะคำที่เป็นคำนาม และกริยา เพื่อให้สอดคล้องกับโครงสร้างประโยคที่กล่าวถึง และกำจัดคำต่างๆ ที่มีคุณลักษณะเป็นคำเชื่อม (Conjunction) ตัวเลข (Numerical) ตัวกำหนด (Determiner) เพื่อให้เหลือแต่คำที่เป็นคำจำเป็นในโครงสร้างประโยค

4) การสร้างต้นไม้แจงส่วน (Parse Tree) การแปลงข้อความในประโยคเป็นต้นไม้แจงส่วน คือ แผนภาพที่แสดงโครงสร้างทางภาษาของแต่ละประโยค เพื่อระบุประเภทของคำประธาน กริยา กรรม การสร้างต้นไม้แจงส่วน ผู้วิจัยเลือกใช้ Spacy เป็น Library Open-Source

สำหรับการประมวลผลภาษาธรรมชาติขั้นสูง ที่สามารถตัดคำ (Tokenize) และระบุชนิดของคำพร้อมทั้งความสัมพันธ์แบบพึ่งพา (Dependency Relation) ของแต่ละโทเคนได้ ตามภาพที่ 19



ภาพที่ 19 แสดงต้นไม้แจงส่วน

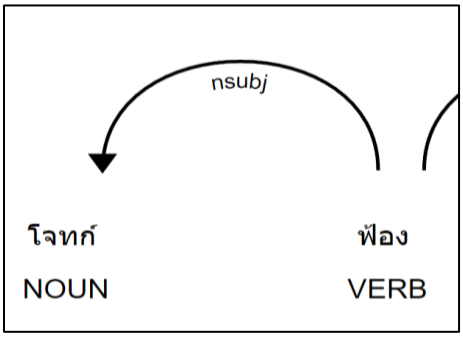
3.4 การสร้างกราฟเชิงความหมาย (Construction of the Graph)

การรวบรวมและผสานของแบบฟอร์มลอจิกที่เป็นคำประธาน และกรรม ของแต่ละโหนด ซึ่งอยู่ในคลาสหรือขอบเขตความหมายเดียวกัน และสร้างกราฟเชิงความหมาย ซึ่งสิ่งที่สำคัญที่สุดคือ โหนด และเส้นทางระหว่างแต่ละโหนด มีวิธีการ ดังนี้

1) การสกัดเอนทิตี้ (Entities Extraction) การแยกเอนทิตี้โดยกำหนดเป็นแต่ละคำของประโยคหรือข้อความโดยใช้คุณสมบัติของ Spacy สามารถนำข้อความภาษาไทยไปประมวลผลและระบุประเภทของชนิดคำ (Part Of Speech) หรือระบุภาคแสดงของแต่ละคำ เพื่อวิเคราะห์โครงสร้างการขึ้นแก่กัน (Dependency Relation) ของประโยค ซึ่งวิธีสกัดเอนทิตี้นี้ ทำในลักษณะเช่นเดียวกันการสร้างต้นไม้แจงส่วน โดย Spacy

2) การสกัดความสัมพันธ์ (Relation Extraction)

เมื่อผู้วิจัยสกัดเอนทิตี้ออกมาแล้วนั้น ลำดับต่อไปคือการสกัดความสัมพันธ์ของคำแต่ละคำเพื่อเป็นการสร้างเส้นทางระหว่างโหนดแต่โหนด



ภาพที่ 20 แสดงความสัมพันธ์ระหว่างคำ

จากภาพที่ 20 จะเห็นว่าทั้ง 2 โหนด ระหว่าง โจทก์ กับ ฟ้อง มีความสัมพันธ์กันแบบ nsubj (Nominal Subject) คือความสัมพันธ์ระหว่างกริยากับประธาน เมื่อผู้วิจัยสามารถสกัดเอนทิตี้ และความสัมพันธ์ของข้อความได้แล้ว ผู้วิจัยจำเป็นต้องสร้างตารางแสดงความสัมพันธ์ของแต่ละโหนด โดยกำหนดตัวแปร ดังนี้

1) โหนดต้นทาง (Source) เป็นชื่อสำหรับคอลัมน์ที่แสดงถึงโหนดที่มา
2) โหนดปลายทาง (Target) เป็นชื่อคอลัมน์ที่แสดงถึงปลายทางที่โหนดต้นทางจะเชื่อมไปถึง

เมื่อผู้วิจัยได้ Source กับ Target เป็นโหนด 2 โหนดแล้ว ตามทฤษฎีของกราฟนั้น เมื่อมีโหนดแล้วต้องมีเส้นเชื่อมระหว่างโหนด ซึ่งเมื่อผู้วิจัยใช้ Dependency Parser จะสามารถระบุถึงความสัมพันธ์ของโหนดทั้งสองได้ ผู้วิจัยเลือกใช้วิธีการวัดความคล้ายโคไซน์ (Cosine Similarity) เพื่อใช้ระบุเป็นค่าน้ำหนักของเส้นเชื่อมนั้น โดยใช้โมเดล Pretrained Word Embedding Thai2Vec ของกลุ่ม PyThaiNLP ที่เรียนรู้ด้วยข้อมูลจาก Thai Wikipedia จำนวน 60,005 embedding ขนาด 300 มิติ

	source	target	edge	dep	sents	no
618	การจับกุม	ตำรวจ	0.137878	compound	จำเลยขับรถจักรยานยนต์เมาสุราขับรถความเร็วปาดหน...	5035/2560
472	ตรี	โจทก์	0.039057	compound	จำเลยมีความผิดพรบจราจรทางบกพศมาตราตรีวรรคปอมาต...	5095/2560
4552	ขอให้	จำเลย	0.211214	obj	โจทก์ขอให้จำเลยเสพเมทแอมเฟตามีนผู้ขับขี่เสพเมท...	9932/2555
3563	บิดา	ผู้ตาย	0.163650	compound	ปวิอมาตราบัญญัติแพ่งจำต้องคำพิพากษาบทบัญญัติใด...	7426
1672	กรณี	เสพ	-0.009202	compound	พรบจราจรทางบกพศมาตราทวิวรรคบัญญัติผู้ขับขี่เสพ...	13947/2558
1338	พศ	วรรค	0.066719	compound	จำเลยผู้ขับขี่เสพเมทแอมเฟตามีนพรบยาเสพติดให้โท...	2908/2559
1370	ตำรวจ	มีอำนาจ	0.114781	compound	จำเลยผู้ขับขี่เสพเมทแอมเฟตามีนพรบยาเสพติดให้โท...	2908/2559
6300	ขับรถ	เมา	0.234438	compound	จำเลยขับรถเมาสุราจำเลยรับสารภาพจำเลยจำเลยขับรถ...	1411/2553
5037	ปี	เศษ	0.059877	acl	พรบเยาวชนครอบครัวพศมาตราบังคับจำเลยยื่นจำเลยปี...	12008/2554
1323	บท	ปอ	0.118337	compound	จำเลยผู้ขับขี่เสพเมทแอมเฟตามีนพรบยาเสพติดให้โท...	2908/2559

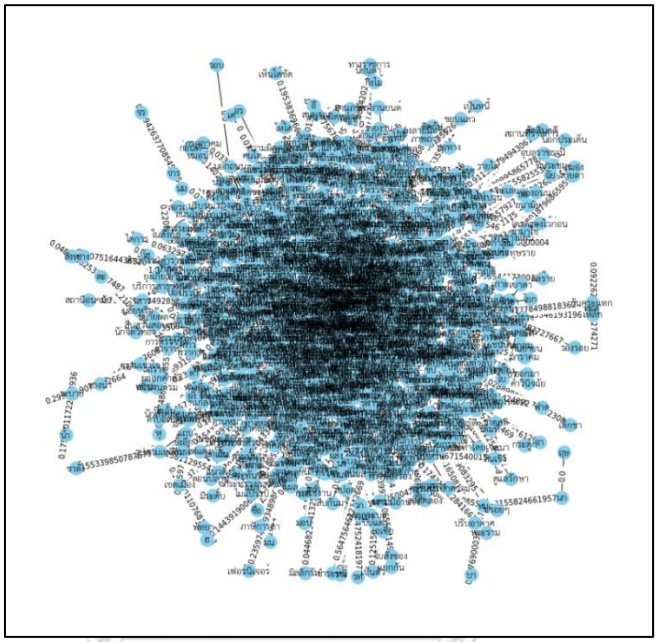
ภาพที่ 21 แสดงเดต้าเฟรมสำหรับสร้างกราฟ

เมื่อสร้างเดต้าเฟรมเสร็จแล้วดังภาพที่ 21 จึงสร้างกราฟด้วย NetworkX เป็นแพ็คเกจของ Python สำหรับการสร้าง จัดการ และศึกษาโครงสร้างแบบไดนามิก และหน้าที่ของเครือข่ายที่ซับซ้อน โดยมีวิธีการสร้างกราฟด้วย NetworkX ปรากฏรายละเอียดตามที่ 22 และผลลัพธ์ตามภาพที่ 23

```
import math

G=nx.from_pandas_edgelist(kg, "source", "target",edge_attr='edge')
pos = nx.spring_layout(G,k=5/math.sqrt(G.order()))
edge_labels = nx.get_edge_attributes(G,'edge')
plt.figure(1,figsize=(15,15))
nx.draw_networkx_edge_labels(G, pos, edge_labels, font_size=10)
nx.draw(G, with_labels=True, node_color='skyblue', edge_cmap=plt.cm.Blues, pos = pos,font_family='TH Sarabun New')
plt.show()
```

ภาพที่ 22 ซอร์สโค้ดการสร้างกราฟด้วย NetworkX

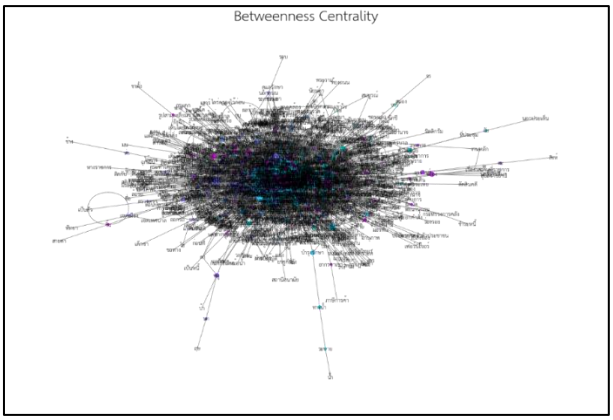


ภาพที่ 23 แสดงกราฟจากคำพิพากษาศาลฎีกาทั้งหมด

3.5 การหาค่ากลาง (Centrality)

เมื่อผู้วิจัยได้ส่วนประกอบของกราฟแล้ว จากคำพิพากษาศาลฎีกา เมื่อนำส่วนประกอบทั้งหมดมาสร้างกราฟ จะเห็นได้ว่ามีโหนดอยู่จำนวน 1340 โหนด และเส้นเชื่อมจำนวน 5271 เส้นเชื่อม เพราะฉะนั้นผู้วิจัยจำเป็นต้องหาค่ากลางของแต่ละโหนด เพื่อดูว่าโหนดใด มีค่ากลางสูง ซึ่งแสดงถึงโหนดดังกล่าว มีอิทธิพลต่อกราฟทั้งหมด โดยวิธีการวัดค่ากลาง มาใช้ได้แก่ ค่ากลางเพจแรงค์ ค่ากลางความเป็นศูนย์กลาง ค่ากลางความใกล้ ค่ากลางคั่นกลาง และค่ากลางไอเกนเวกเตอร์ ปรากฏตัวอย่างดังภาพที่ 24 - 33

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY



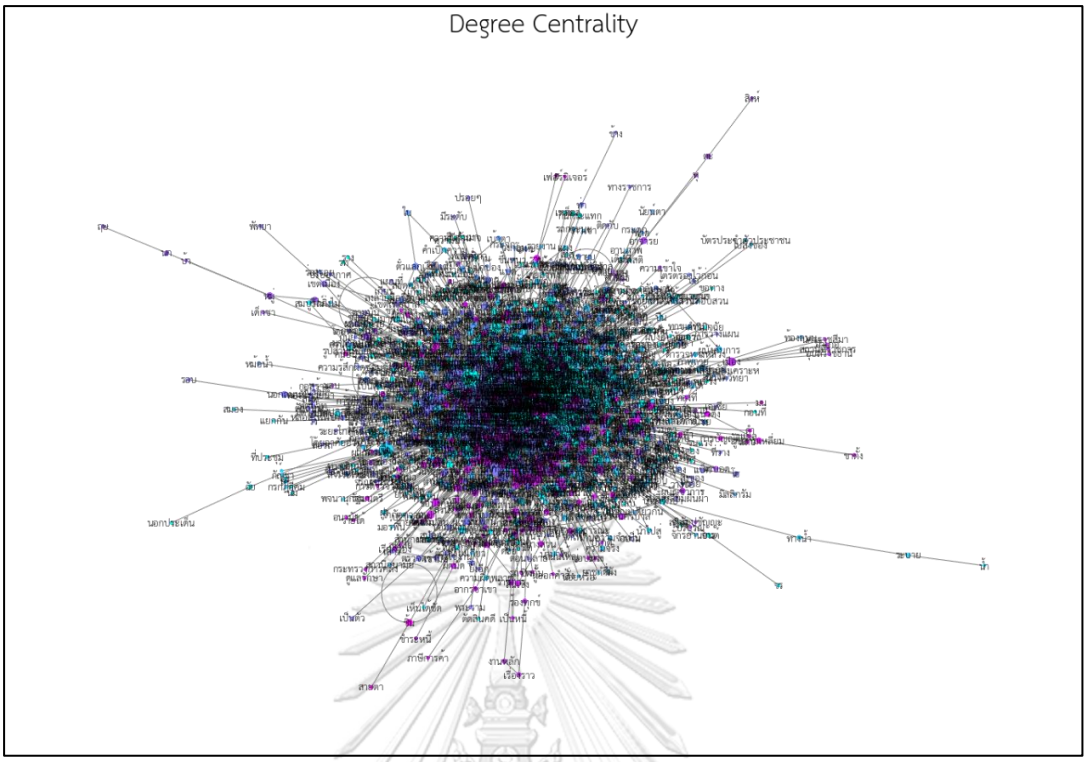
ภาพที่ 24 แสดงกราฟการหาค่ากลางด้วยวิธีค่ากลางคั่นกลาง

bet_df

	key	value
48	จำเลย	0.341549
22	รถ	0.093371
115	โจทก์	0.086892
46	ขับ	0.069384
33	รถยนต์	0.058170
10	มาตรา	0.049706
69	ถนน	0.039790
151	ผู้ตาย	0.039192
122	ขับรถ	0.038426
286	กรณี	0.034217

ภาพที่ 25 เดต้าเฟรมแสดงค่ากลางคั่นกลาง

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

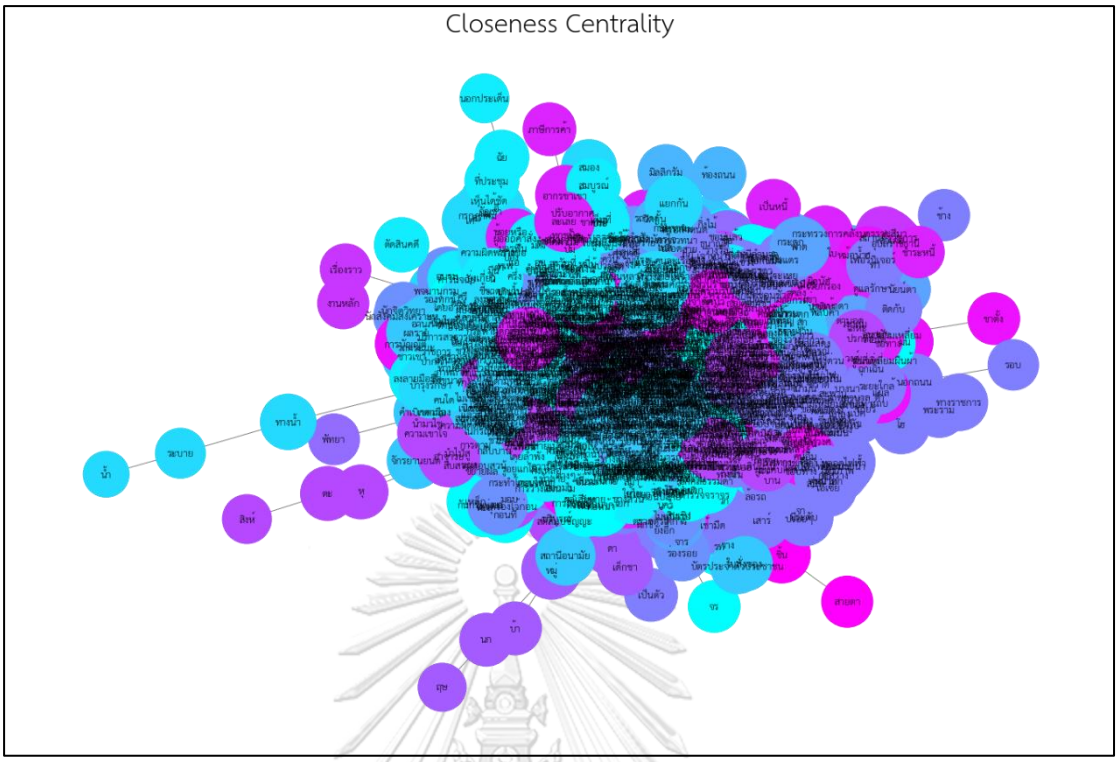


ภาพที่ 26 แสดงกราฟการหาค่ากลางด้วยวิธีค่ากลางความเป็นศูนย์กลาง

deg_df

	key	value
48	จำเลย	0.307435
115	โจทก์	0.144005
22	รถ	0.137307
46	ขับ	0.123242
10	มาตรา	0.115874
122	ขับรถ	0.095780
33	รถยนต์	0.094441
2	พร	0.076356
151	ผู้ตาย	0.076356
126	เป็นเหตุให้	0.072338

ภาพที่ 27 เดต้าเฟรมแสดงค่ากลางความเป็นศูนย์กลาง

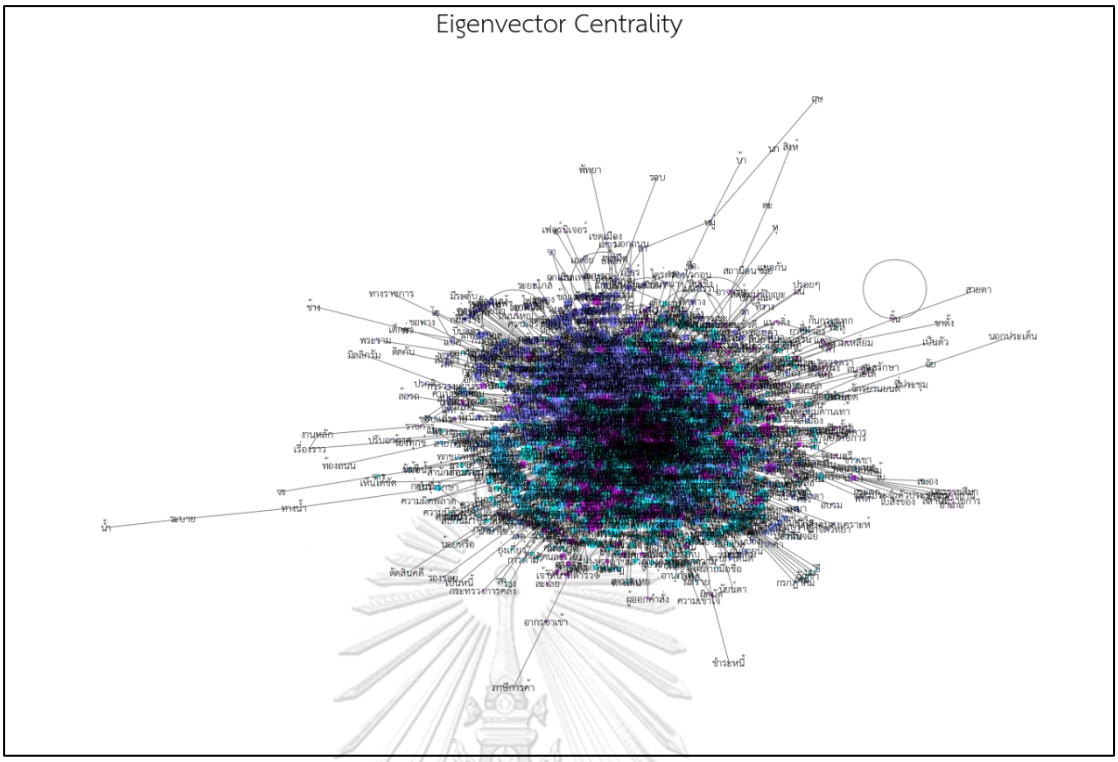


ภาพที่ 28 แสดงกราฟการหาค่ากลางด้วยวิธีค่ากลางความใกล้

close_df

	key	value
48	จำเลย	0.553783
115	โจทก์	0.475932
22	รถ	0.472917
46	ขับ	0.465689
10	มาตรา	0.465109
122	ขับรถ	0.461087
33	รถยนต์	0.448348
2	พร	0.445273
286	กรณี	0.441978
151	ผู้ตาย	0.441978

ภาพที่ 29 เดต้าเฟรมแสดงค่ากลางความใกล้

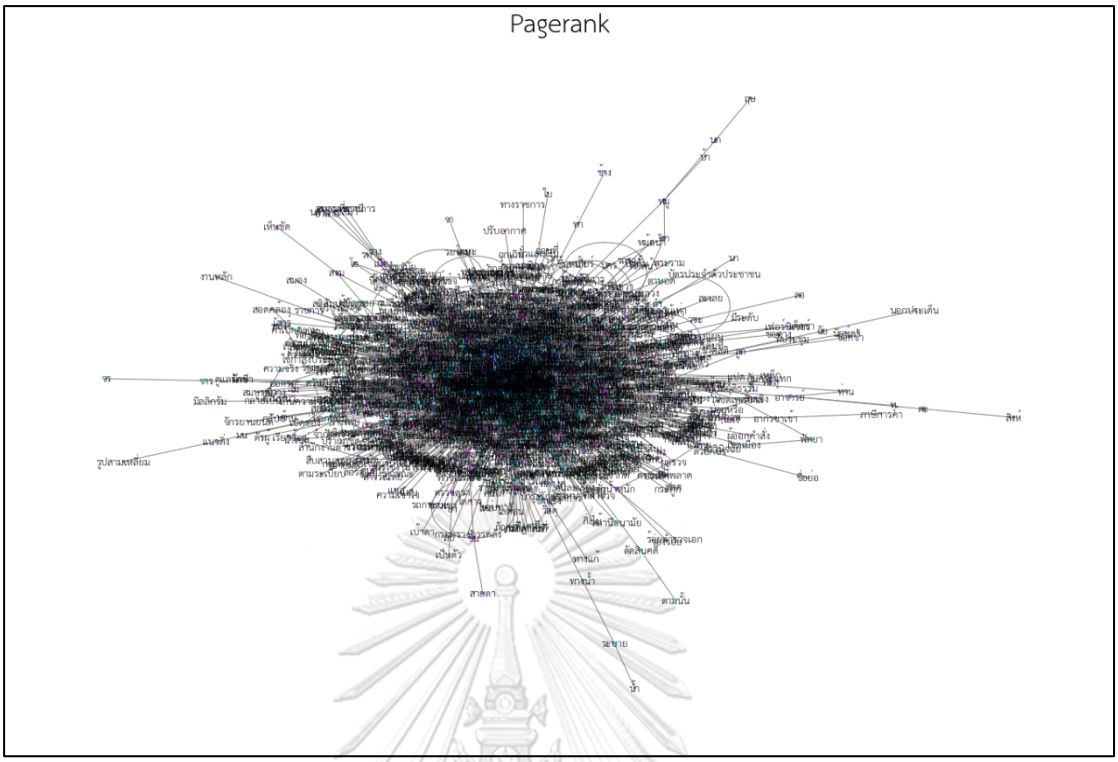


ภาพที่ 30 แสดงกราฟการหาค่ากลางด้วยวิธีค่ากลางไอเกนเวกเตอร์

eigen_df

	key	value
48	จำเลย	0.327446
115	โจทก์	0.206203
10	มาตรา	0.192163
122	ขับรถ	0.172885
22	รถ	0.172080
46	ขับ	0.159486
2	พร	0.153462
33	รถยนต์	0.132206
286	กรณี	0.130466
126	เป็นเหตุให้	0.126818

ภาพที่ 31 แสดงเดต้าเฟรมค่ากลางไอเกนเวกเตอร์



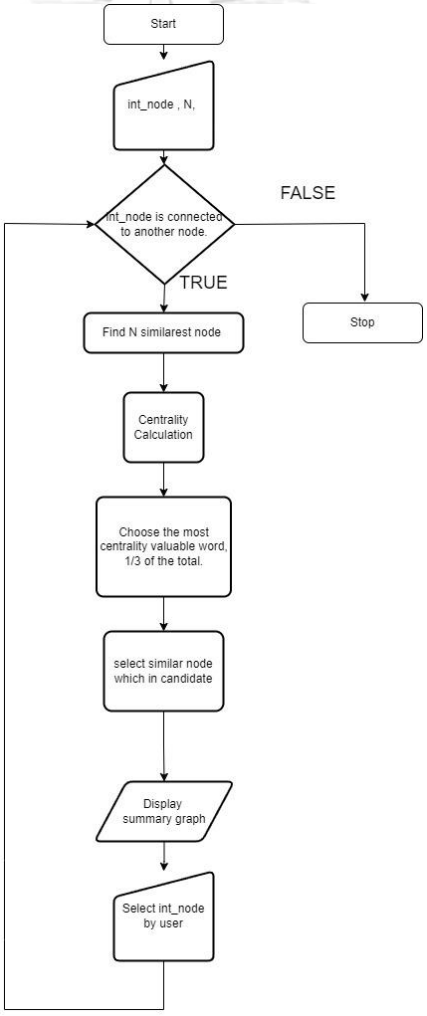
ภาพที่ 32 แสดงกราฟการหาค่าความสำคัญของกราฟด้วยวิธีค่ากลางเพจแรงก์

node	page_rank
จำเลย	0.034659
โจทก์	0.017655
รถ	0.015691
ขับ	0.012648
รถยนต์	0.010394
ขับรถ	0.009685
ผู้ตาย	0.008822
กรณี	0.008257
ถนน	0.007558
เป็นเหตุให้	0.007209

ภาพที่ 33 แสดงเดต้าเฟรมค่ากลางเพจแรงก์

3.6 ขั้นตอนการท่องกราฟเพื่อสร้างแบบรายงานประจำวันสำหรับตำรวจ (Graph traverse procedure)

การบันทึกประจำวันเกี่ยวกับคดีจราจร เริ่มต้นจากผู้เสียหายเดินทางเข้ามาที่สถานีตำรวจเพื่อพบกับพนักงานสอบสวน เพื่อแจ้งความร้องทุกข์ โดยปกติแล้ว พนักงานสอบสวนจะซักถามรายละเอียดข้อมูลของผู้แจ้ง/ผู้เสียหาย ที่มาแจ้งความร้องทุกข์ ดังนั้น ผู้วิจัยต้องศึกษาการไหลของข้อมูลหรือกราฟว่า ในการออกแบบรายงานประจำวันตำรวจนั้น พนักงานสอบสวนควรจะเริ่มซักถามในหัวข้อใดบ้าง เพื่อให้ได้รายละเอียดที่สำคัญ และครอบคลุม ดังนั้น การสร้างรายงานประจำวันจะต้องเป็นการสร้างคำถามในหัวข้อต่างๆ จากกราฟ ลักษณะเดียวกันกับการซักถามของพนักงานสอบสวนกับประชาชน โดยมีแผนผังการทำงานดังภาพที่ 34



ภาพที่ 34 แผนผังการทำงานของการสกัดกิจกรรมหรือคำสำคัญ

ขั้นตอนวิธีในการสร้างคำถามสำหรับรายงานประจำวัน ดังนี้

1) กำหนดโหนดเริ่มต้น และกำหนดชุดของคำที่มีความสำคัญต่อเครือข่ายทั้งหมด (คัดเลือกจากค่ากลาง สูงสุดจำนวนร้อยละ 33 ของจำนวนคำทั้งหมด)

2) ค้นหาโหนดที่เชื่อมกับโหนดตามข้อ 1) ทั้งหมด ที่มีค่าความคล้ายคลึงสูงที่สุด จำนวน N ตัว

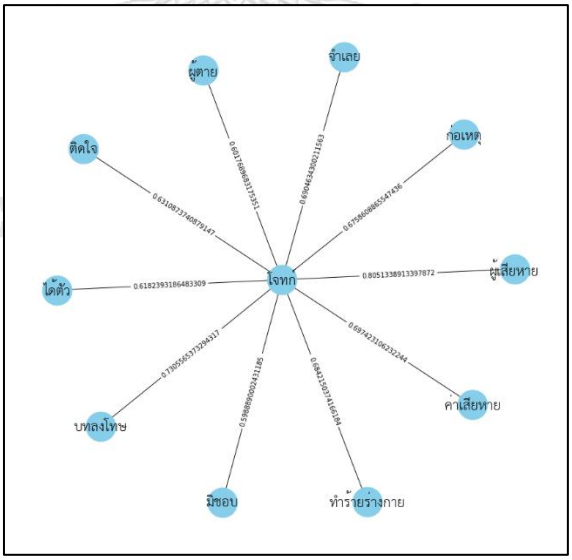
3) ระบุค่ากลางของโหนดทุกโหนดในข้อ 2)

4) เลือกโหนดตามข้อ 2 และอยู่ในชุดของคำที่มีความสำคัญต่อเครือข่าย

5) จาก 4) ถ้าโหนดใดที่มีค่ากลางสูงสุด และซ้ำกับโหนดที่ถูกเลือกไปแล้ว ให้เลือกโหนดที่มีค่ากลางรองลงมา

6) กำหนดให้โหนดที่มีค่ากลางสูงสุด เป็นโหนดเริ่มต้นสำหรับการทำงานรอบต่อไป

การวิเคราะห์จากกราฟ จะเริ่มวิเคราะห์ที่โหนด คำว่า โจทก์ หรือ จำเลย เนื่องจากเป็น โหนดที่มีเอนทิตี้ที่แสดงถึงบุคคล เมื่อผู้วิจัยนำโหนด “โจทก์” มาเปรียบเทียบกับโหนดอื่นๆ โดยใช้การ เพื่อดูว่าบุคคล (โจทก์) เชื่อมกับโหนดใดบ้าง โดยเลือกมา 10 ลำดับ โดยเรียงจากค่าน้ำหนักของเส้นเชื่อมที่มากที่สุด ผลลัพธ์ปรากฏตามภาพที่ 35



ภาพที่ 35 แสดงโหนดที่เชื่อมกับคำว่า “โจทก์” และมีความคล้ายคลึงสูงสุด 10 อันดับ

	source	target	edge
6458	โจทก์	ผู้เสียหาย	0.805134
2472	โจทก์	บทลงโทษ	0.730557
10177	โจทก์	ค่าเสียหาย	0.697423
4830	โจทก์	จำเลย	0.690463
11344	โจทก์	ทำร้ายร่างกาย	0.684215
4015	โจทก์	ก่อเหตุ	0.675861
8915	โจทก์	ติดใจ	0.631087
9037	โจทก์	ได้ตัว	0.618239
5917	โจทก์	ผู้ตาย	0.601769
2482	โจทก์	มิชอบ	0.598889

ภาพที่ 36 แสดงโหนดที่เชื่อมกับโจทก์ พร้อมค่าความคล้ายคลึง

เมื่อผู้วิจัยได้โหนดที่มีความคล้ายคลึงสูงสุด 10 อันดับ ตามภาพที่ 36 ขั้นตอนต่อไปคือการหาค่ากลางของโหนดทั้ง 10 โหนด เพื่อหาว่าโหนดใดมีค่ากลางงที่สุด โดยผู้วิจัย ยกตัวอย่างการหาค่ากลางโดยใช้ วิธีค่ากลางความเป็นศูนย์กลาง ผลลัพธ์ตามภาพที่ 37

	node	deg_centrality
0	ผู้เสียหาย	0.041611
1	บทลงโทษ	0.011409
2	ค่าเสียหาย	0.008054
3	จำเลย	0.308725
4	ทำร้ายร่างกาย	0.004027
5	ก่อเหตุ	0.000671
6	ติดใจ	0.001342
7	ได้ตัว	0.000671
8	ผู้ตาย	0.076510
9	มิชอบ	0.004027

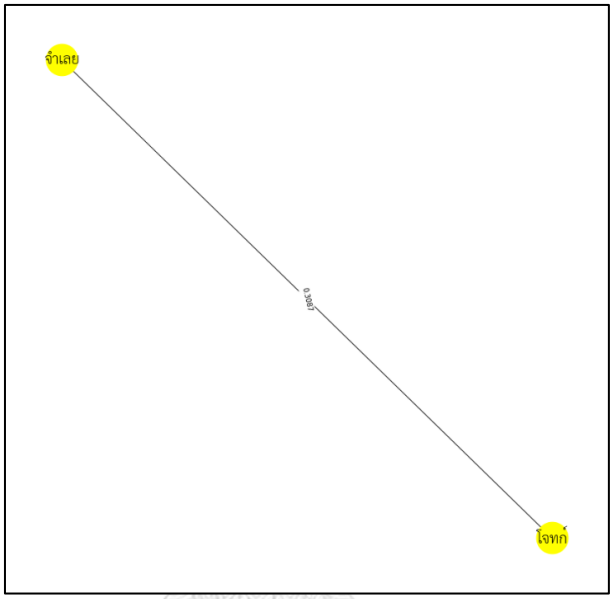
ภาพที่ 37 โหนดที่เชื่อมกับโจทก์ และที่มีค่ากลางสูงสุด

เมื่อผู้วิจัยได้โหนดที่มีค่ากลางสูงสุดแล้ว ขั้นตอนต่อไปเป็นการสร้างกราฟ โดยการกำหนดคุณลักษณะ (Attributes) ของกราฟ ซึ่งประกอบด้วย 3 ตัวแปร ได้แก่ โหนดต้นทาง โหนดปลายทาง และเส้นเชื่อม ดังตารางที่ 7

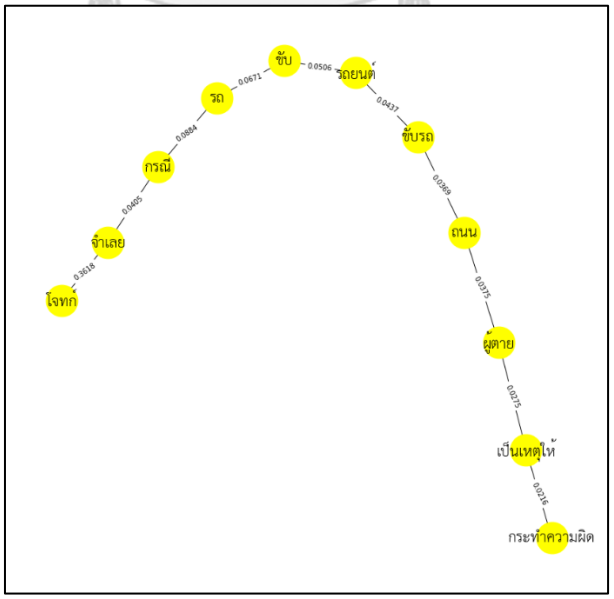
ตารางที่ 7 แสดงคุณลักษณะของกราฟ

Source	Target	Edge
โจทก์	จำเลย	0.3087248322147651

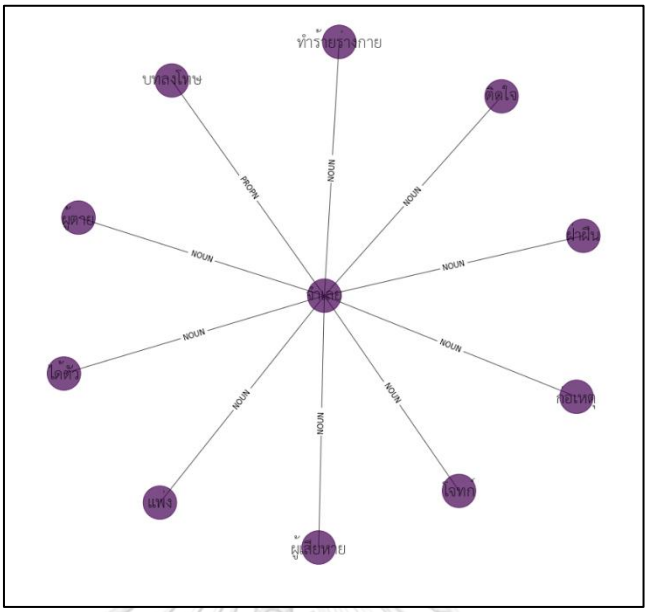
จากนั้น นำออปเจ็กต์ของคุณลักษณะดังกล่าวไปใช้สำหรับสร้างกราฟด้วย NetworkX จะได้ผลลัพธ์ดังภาพที่ 38



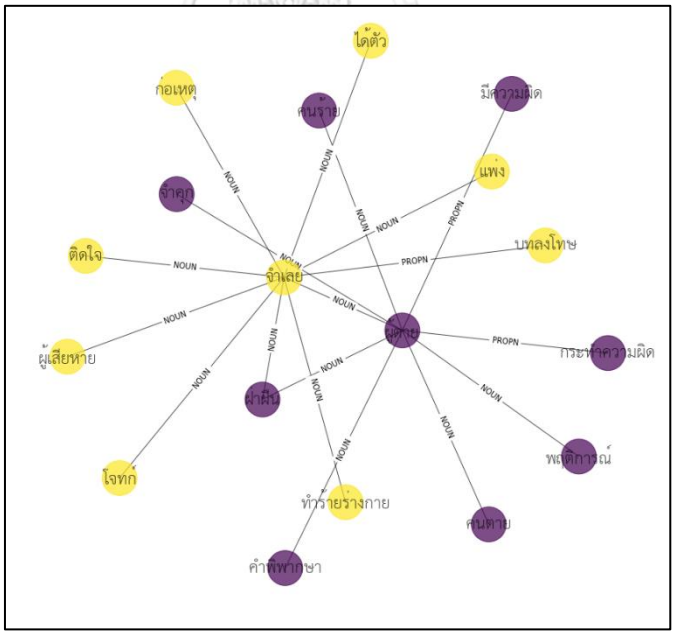
ภาพที่ 38 กราฟสรุปโหนดที่มีค่ากลางสูงสุด จำนวน 2 โหนด



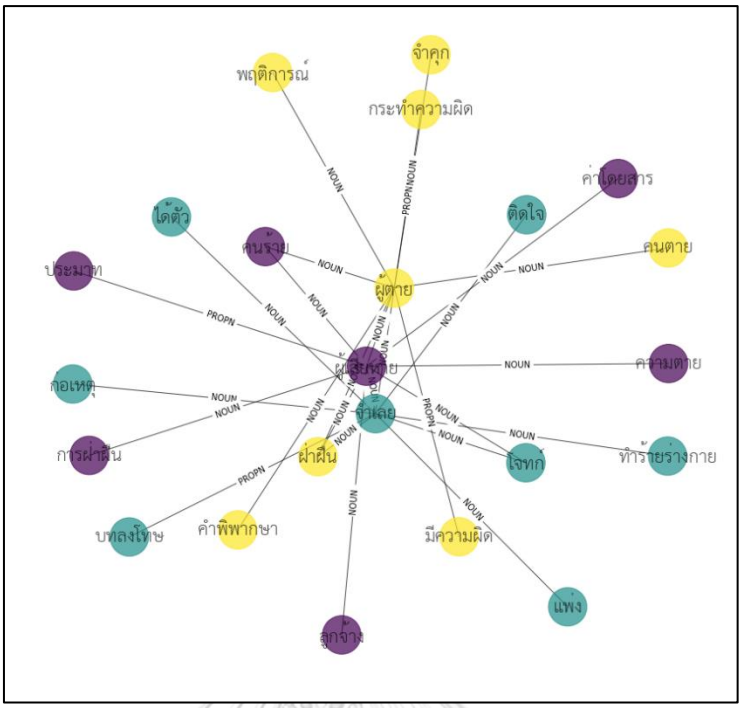
ภาพที่ 39 กราฟสรุปโดยค่ากลางความเป็นศูนย์กลาง



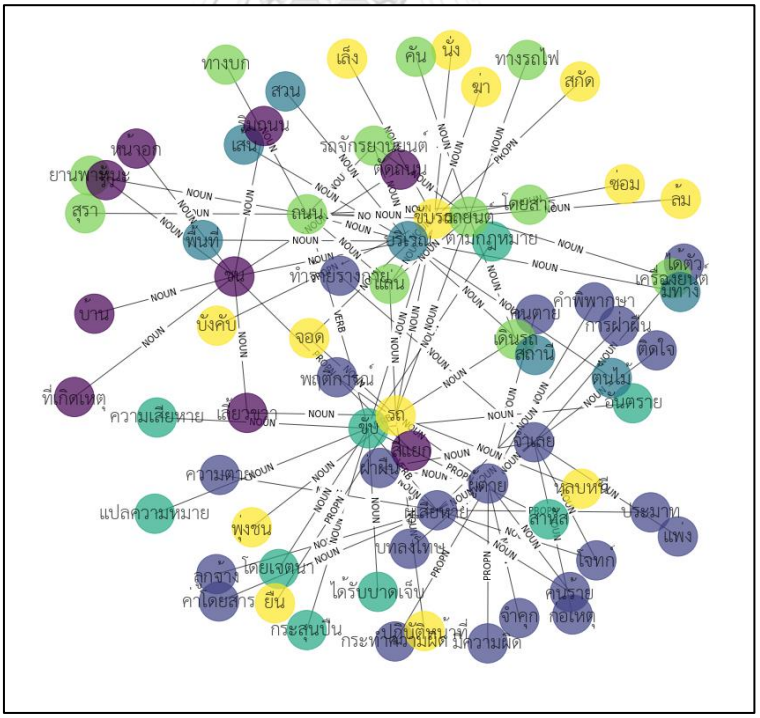
ภาพที่ 40 แสดงกราฟสรุป



ภาพที่ 41 แสดงกราฟสรุป



ภาพที่ 42 แสดงกราฟสรุป



ภาพที่ 43 แสดงกราฟสรุป

เมื่อเริ่มการสกัดกิจกรรมหรือคำสำคัญ นั้น ในแต่ละรอบ (Iteration) ของการทำงาน จะปรากฏโหนดที่มีความใกล้เคียง กับโหนดเริ่มต้น และเมื่อระบบทำงานต่อเนื่องไป จำนวนโหนดที่ได้ก็จะเพิ่มมากขึ้น ตามโหนดที่ผู้ใช้เลือกเป็นโหนดเริ่มต้น ในการทำงานแต่ละรอบผลลัพธ์การทำงานแต่ละรอบปรากฏดังภาพที่ 38 - 43 เมื่อผู้วิจัยสามารถสร้างกราฟสรุปเพื่อหาคำที่มีความสำคัญต่อเครือข่ายด้วยการใช้วิธีวัดความสำคัญของแต่ละคำ สามารถสรุปผลได้ดังตารางที่ 8

ตารางที่ 8 ตารางแสดงคำทั้งหมดที่ความสำคัญต่อเครือข่าย

วิธีวัดความสำคัญของเครือข่าย	**คำที่มีความสำคัญต่อเครือข่ายที่ K = 10**
ค่ากลางคั่นกลาง	กระทำความผิด กระสุนปืน การฝ่าฝืน ก่อเหตุ ขับ ขับรถ คนตาย คนร้าย ความตาย ความเสียหาย คัน คำพิพากษาศาลฎีกา ค่าโดยสาร ฆ่า จอด จำคุก จำเลย ชน ซ่อม ตัดถนน ตามกฎหมาย ติดใจ ต้นไม้ ถนน ทางบก ทางรถไฟ ทำร้ายร่างกาย ที่เกิดเหตุ นั่ง บทลงโทษ บริเวณ บังคับ บ้าน ปฏิบัติหน้าที่ ประมาท ผู้ตาย ผู้เสียหาย ฝ่าฝืน พฤติการณ์ พื้นที่ พุ่งชน มีความผิด มีทาง ยานพาหนะ ยืน รถ รถจักรยานยนต์ รถยนต์ รั้ว ริมถนน ลูกจ้าง ล้ม สกัด สถานี สวน สาหัส สี่แยก สุรา หน้าอก หลบหนี อันตราย เครื่องยนต์ เดินรถ เลี้ยวขวา เล็ง เส้น แปลความหมาย แพ่ง แล่น โจทก์ โดยสาร โดยเจตนา ได้ตัว ได้รับบาดเจ็บ
ค่ากลางความเป็นศูนย์กลาง	กระทำความผิด กระสุนปืน การฝ่าฝืน ก่อเหตุ ขับ ขับรถ คนตาย คนร้าย ความตาย ความเสียหาย คัน คำพิพากษาศาลฎีกา ค่าโดยสาร ฆ่า จอด จำคุก จำเลย ชน ซ่อม ตัดถนน ตามกฎหมาย ติดใจ ต้นไม้ ถนน ทางบก ทางรถไฟ ทำร้ายร่างกาย ที่เกิดเหตุ นั่ง บทลงโทษ บริเวณ บังคับ บ้าน ปฏิบัติหน้าที่ ประมาท ผู้ตาย ผู้เสียหาย ฝ่าฝืน พฤติการณ์ พื้นที่ พุ่งชน มีความผิด มีทาง ยานพาหนะ ยืน รถ รถจักรยานยนต์ รถยนต์ รั้ว ริมถนน ลูกจ้าง ล้ม สกัด สถานี สวน สาหัส สี่แยก สุรา หน้าอก หลบหนี อันตราย เครื่องยนต์ เดินรถ เลี้ยวขวา เล็ง เส้น แปลความหมาย แพ่ง แล่น โจทก์ โดยสาร โดยเจตนา ได้ตัว ได้รับบาดเจ็บ
ค่ากลางความใกล้	กฎกระทรวง กระทำความผิด กระสุนปืน การขนส่ง การฝ่าฝืน ก่อเหตุ ขับ ขับรถ คนตาย คนร้าย ความตาย ความเสียหาย คัน คำพิพากษาศาลฎีกา ค่าโดยสาร ฆ่า จราจร จอด จำคุก จำเลย ซ่อม ดุลพินิจ ตามกฎหมาย ติดใจ ถนน ทางบก ทางรถไฟ ทำร้ายร่างกาย นั่ง บทลงโทษ บังคับ ปฏิบัติหน้าที่ ประมาท

	ปว ผู้ตาย ผู้เสียหาย ฝ่าฝืน พฤติการณ์ พุ่งชน มอบอำนาจ มีความผิด ยานพาหนะ ยืน รถ รถจักรยานยนต์ รถชน รถยนต์ ลูกจ้าง ล้ม สกัด สาหัส สุรา หลบหนี อันตราย เครื่องยนต์ เดินรถ เป็นคุณ เป็นผล เลี้ยวขวา เล็ง เสพ แปล ความหมาย แพ่ง แล่น โจทก์ โดยสาร โดยเจตนา โดยเร็ว ใบอนุญาตขับขี่ ได้ตัว ได้รับบาดเจ็บ
ค่ากลางไอเกน เวกเตอร์	กฎกระทรวง กระทำความผิด กระสุนปืน การขนส่ง การฝ่าฝืน ก่อเหตุ ขับ ขับรถ คนตาย คนร้าย ความตาย ความเสียหาย คัน คำพิพากษาศาลฎีกา ค่าโดยสาร ฆ่า จราจร จอด จำคุก จำเลย ซ่อม ดุลพินิจ ตามกฎหมาย ติดใจ ถนน ทางบก ทางรถไฟ ทำร้ายร่างกาย นั่ง บทลงโทษ บังคับ ปฏิบัติหน้าที่ ประมาท ปว ผู้ตาย ผู้เสียหาย ฝ่าฝืน พฤติการณ์ พุ่งชน มอบอำนาจ มีความผิด ยานพาหนะ ยืน รถ รถจักรยานยนต์ รถชน รถยนต์ ลูกจ้าง ล้ม สกัด สาหัส สุรา หลบหนี อันตราย เครื่องยนต์ เดินรถ เป็นคุณ เป็นผล เลี้ยวขวา เล็ง เสพ แปล ความหมาย แพ่ง แล่น โจทก์ โดยสาร โดยเจตนา โดยเร็ว ใบอนุญาตขับขี่ ได้ตัว ได้รับบาดเจ็บ
ค่ากลางแพจ แรงก์	กระทำความผิด กระสุนปืน การฝ่าฝืน ก่อเหตุ ขับ ขับขี่ ขับรถ คนตาย คนร้าย ความตาย ความระมัดระวัง ความเสียหาย คัน คำพิพากษาศาลฎีกา ค่าโดยสาร ฆ่า จอด จำคุก จำเลย ซ่อม ซ้ายมือ ตามกฎหมาย ติดใจ ต้นไม้ ถนน ทางบก ทางรถไฟ ทำร้ายร่างกาย นั่ง บทลงโทษ บริเวณ บังคับ ปฏิบัติหน้าที่ ประมาท ผู้ตาย ผู้เสียหาย ฝ่าฝืน พฤติการณ์ พื้นที่ พุ่งชน มีความผิด มีทาง ยานพาหนะ ยืน รถ รถจักรยานยนต์ รถยนต์ ลูกจ้าง ล้ม สกัด สถานี สวน สาหัส สุรา หลบหนี หลบหลีก อันตราย เครื่องยนต์ เดินรถ เลี้ยว เลี้ยวขวา เล็ง เส้น แปล ความหมาย แพ่ง แล่น โจทก์ โดยสาร โดยเจตนา ได้ตัว ได้รับบาดเจ็บ ไฟหน้า

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

การสกัดคำสำคัญจากคำพิพากษาศาลฎีกานั้น เมื่อผู้วิจัยได้กราฟสรุปข้อความสำคัญที่สกัดออกมาแล้ว ขั้นตอนต่อไปคือการสร้างข้อความคำถามจากคำสำคัญที่ได้สกัดออกมา โดยคำที่สกัดออกมาจากคำพิพากษาศาลฎีกา

3.7 การวัดผล (Evaluation)

เมื่อผู้วิจัยสร้างแม่แบบสำหรับรายงานประจำวันจากกราฟสรุปแล้ว ต่อไปจะเป็นขั้นตอนสำหรับการวัดผล เพื่อดูว่าการสร้างรายงานประจำวันจากกระบวนการสกัดคำสำคัญจากคำพิพากษาศาลฎีกาที่ผู้วิจัยเสนอ โดยการหาความสำคัญโดยวิธีการทางกราฟ (Network Measure) ซึ่งประกอบด้วย เพจแรงค์ ค่ากลางความเป็นศูนย์กลาง ค่ากลางความใกล้ ค่ากลางคั่นกลาง ค่ากลางไอเกนเวกเตอร์ นั้น มีประสิทธิภาพเป็นอย่างไร เพื่อระบุว่า ผลลัพธ์ที่ได้นั้น มีความเหมาะสมและครอบคลุมการบันทึกประจำวันมากน้อยเพียงใด สำหรับขั้นตอนการวัดผลของการสร้างประจำวันจากกราฟสรุปคำพิพากษาศาลฎีกา ในวิธีนี้จะนำรายงานประจำวันที่ได้ มาวัดผลเทียบกับประจำวันของจริงที่ผู้วิจัยได้รวบรวมมาจากพนักงานสอบสวน โดยจะเปรียบเทียบข้อความที่สรุปมาแล้ว กับข้อมูลอ้างอิงหรือชุดข้อมูลอ้างอิง โดยการวัดผลมีผลลัพธ์ และสามารถเขียนเป็นสูตรคำนวณ ดังนี้

(1) ค่าความแม่นยำ (Precision) เป็นการคำนวณหาความแม่นยำ โดยคำนวณจากจำนวนคำในรายงานประจำวันที่สร้างขึ้นจากการสร้างกราฟสรุป ปรากฏคำที่ตรงตามคำที่อยู่ในคำพิพากษาศาลฎีกามากน้อยเพียงใด เทียบกับจำนวนคำสำคัญที่สกัดออกมาได้

(2) ค่าความระลึก (Recall) เป็นการคำนวณหาความครอบคลุมของรายงานประจำวันที่ถูกสร้างขึ้น กับคำที่อยู่ในคำพิพากษาศาลฎีกาว่ามีคำที่ปรากฏเหมือนกันมากน้อยเพียงใด

(3) ค่าวัดประสิทธิภาพ (F-measure) เป็นการหาค่าเฉลี่ยโดยคำนวณจากค่าความระลึกและค่าความแม่นยำ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

$$\textbf{ค่าความแม่นยำ (Precision)} = \frac{\text{จำนวนคำที่ทับซ้อนกันระหว่าง คำที่ระบบคัดเลือก และคำที่ปรากฏในคำพิพากษา}}{\text{จำนวนคำที่ระบบคัดเลือก}}$$

$$\textbf{ค่าความระลึก (Recall)} = \frac{\text{จำนวนคำที่ทับซ้อนกันระหว่าง คำที่ระบบคัดเลือก และคำที่ปรากฏในคำพิพากษา}}{\text{จำนวนคำในคำพิพากษาศาลฎีกา}}$$

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

การทดสอบเบื้องต้น (Preliminary experiment)

การทดสอบเบื้องต้น จะเป็นการวัดผลในลักษณะนี้นำคำสำคัญที่ถูกสกัด เทียบกับคำพิพากษาศาลฎีกา โดยกำหนดคำหรือข้อความสำคัญที่ผู้วิจัยสกัดได้จากกราฟสรุป โดยวิธีวัดความสำคัญ ตามกตารางที่ 8 จะเห็นได้ว่าคำที่มีความสำคัญต่อเครือข่ายโดยวิธีการวัดแต่ละแบบนั้น มีอยู่จำนวนมากและหลากหลาย และมีความหมายค่อนข้างครอบคลุมตาม พ.ร.บ.จราจรทางบก พ.ศ.2522 ซึ่งตามข้อเท็จจริงแล้ว ในการสร้างรายงานประจำวันก็ไม่จำเป็นต้องมีความหมายครอบคลุมในทุกๆ บทบัญญัติ มาตรา ตามกฎหมายดังกล่าว เพียงแต่ว่าผู้ใช้งานหรือพนักงานสอบสวนสามารถเลือกคำที่สำคัญและตรงตามวัตถุประสงค์และพฤติการณ์ของเรื่องที่จะบันทึก

ตารางที่ 9 ตารางแสดงตัวอย่างคำสำคัญที่ได้จากคำพิพากษาศาลฎีกา

วิธีวัดความสำคัญ	ตัวอย่างคำสำคัญ
ค่ากลางคั่นกลาง	ผู้ตาย ผู้เสียหาย ขับ ขับรถ รถ รถยนต์ ถนน บริเวณ ชน เป็นเหตุให้ ขอให้ บังคับ จราจร ทางบก รถชน ความประมาท เฉี่ยว รถจักรยานยนต์
ค่ากลางความเป็นศูนย์กลาง	ผู้ตาย ผู้เสียหาย ขับ ขับรถ รถ รถยนต์ ถนน รถจักรยานยนต์ เดินรถ ผู้ขับขี่ จราจร ทางบก รถชน ความประมาท เป็นเหตุให้ ขอให้ บังคับ กระทำความผิด
ค่ากลางความใกล้	ผู้ตาย ผู้เสียหาย ขับ ขับรถ รถ รถยนต์ ทางบก จราจร เสพ ผู้ขับขี่ รถจักรยานยนต์ เดินรถ จอด ชน เป็นเหตุให้ ขอให้ บังคับ กระทำความผิด
ค่ากลางไอเกนเวกเตอร์	ผู้ตาย ผู้เสียหาย ขับ ขับรถ รถ รถยนต์ ทางบก จราจร เสพ ผู้ขับขี่ รถจักรยานยนต์ เดินรถ จอด ชน เป็นเหตุให้ ขอให้ วินิจฉัย คำพิพากษาศาลฎีกา
ค่ากลางแพจแรงก์	ผู้ตาย ผู้เสียหาย ขับ ขับรถ รถ รถยนต์ ถนน รถจักรยานยนต์ เดินรถ ผู้ขับขี่ จราจร ทางบก รถชน ความประมาท เป็นเหตุให้ ขอให้ บังคับ กระทำความผิด

จะเห็นได้ว่า วิธีการสกัดคำสำคัญที่ยกตัวอย่างมา มีคำที่แตกต่างกัน ปรากฏดังตารางที่ 9 ซึ่งขึ้นอยู่กับขั้นตอนวิธีในการวัดความสำคัญของแต่ละวิธี ทั้งนี้ ผู้วิจัยจึงสร้างข้อความจากคำที่แตกต่างกันดังกล่าว เพื่อที่จะสร้างรายงานประจำวันที่มีความแตกต่างกันจากวิธีทางกราฟทั้ง 5 แบบ เพื่อทดสอบประสิทธิภาพของวิธีการทางกราฟแต่ละแบบ

ในส่วนการทดสอบเบื้องต้น ผู้วิจัยกำหนดให้เป็นการแจ้งความร้องทุกข์เพื่อลงบันทึกประจำวันคดีจราจร กรณีเกิดอุบัติเหตุรถเชี่ยวชนที่คู่กรณีทั้งสองฝ่ายขับรถยนต์ โดยผู้วิจัยทดสอบนำคำพิพากษาศาลฎีกามา 1 รายการ และทดสอบสกัดคำสำคัญ ได้ผลลัพธ์ดังตารางที่ 10

ตารางที่ 10 แสดงตัวอย่างผลลัพธ์การวัดผลเทียบกับคำพิพากษาศาลฎีกา

วิธีวัดความสำคัญ	ค่าความแม่นยำ	ค่าความระลึก	ค่าประสิทธิภาพ
ค่ากลางคั่นกลาง	1.0	0.666	0.799
ค่ากลางความเป็นศูนย์กลาง	1.0	0.694	0.819
ค่ากลางความใกล้	1.0	0.555	0.714
ค่ากลางไอเกนเวกเตอร์	1.0	0.555	0.714
ค่ากลางแพจแรงก์	1.0	0.722	0.838

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

บทที่ 4

การทดสอบ และผลการทดสอบ

4.1 สภาพแวดล้อมและเครื่องมือที่ใช้ในการทดสอบ

สภาพแวดล้อมและเครื่องมือที่ใช้ในการทดสอบ สามารถสรุปได้ ดังนี้

4.1.1 คอมพิวเตอร์ที่ใช้ในการทดสอบ

การทดสอบนี้ ทำบนเครื่องคอมพิวเตอร์ที่มีหน่วยประมวลผลกลาง Ryzen 5 4500U ความเร็ว 2.38GHz Ram 8 GB ทำงานบนระบบปฏิบัติการ Windows 10 64-bit

4.1.2 การเขียนโปรแกรม

1) การเขียนโปรแกรมเพื่อใช้ในการทดสอบนี้ถูกพัฒนาโดยใช้ภาษาไพทอน
2) การนำเข้าข้อมูลจากหน้าเว็บเพจ (Web Scraping) ด้วย BeautifulSoup ซึ่งเป็นไพทอนโมดูล (Python Module) ที่ใช้สำหรับการดึงข้อมูลออกมาจาก HTML หรือ XML
3) PythaiNLP สำหรับการประมวลผลภาษาธรรมชาติ ที่รองรับภาษาไทย ใช้สำหรับการประมวลผลข้อมูลเบื้องต้น ได้แก่ การโทเคนไนเซชัน, การกำกับโครงสร้างประโยค
4) การให้น้ำหนักของคำด้วย thai2vec โดยนำคำต่างๆ ในโมเดล และอยู่ในรูปแบบเวกเตอร์มาคำนวณหาค่าความคล้ายคลึงระหว่างโทเคน
5) NetworkX เป็นเครื่องมือสำหรับการสร้างกราฟ และคำนวณหาค่าความสำคัญทางกราฟ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

4.1.3 ข้อมูลที่ใช้ในการทดสอบ

คำพิพากษาศาลฎีกา เฉพาะความผิดตาม พ.ร.บ.จราจรทาง พ.ศ.2522 และที่แก้ไขเพิ่มเติม จำนวน 330 รายการ

4.2 การทดสอบ (Experiment)

ในส่วนนี้เป็นการทดสอบเพื่อวัดประสิทธิภาพผลลัพธ์ที่ได้ เทียบกับคำพิพากษาศาลฎีกา โดยการสร้างรายงานประจำวันจากวิธีการสกัดคำสำคัญต่างๆ เพื่อคัดเลือกคำสำคัญจากคำที่มีค่าคะแนนในแต่ละวิธีสูงที่สุด และจำลองคำที่เกี่ยวข้อง (Coverage Word) คือ เป็นการจำลองคำที่ผู้ใช้จะนำมากรอกในรายงานประจำวัน โดยกำหนดคำที่ปรากฏอยู่หน้า และหลังของคำที่ถูกคัดเลือกในคำพิพากษานั้น ๆ จำนวน N คำ เป็นคำที่ผู้ใช้จะนำมากรอกในรายงานประจำวัน เมื่อได้ผลลัพธ์แล้วนำมาเทียบกับคำพิพากษาศาลฎีกา เพื่อหาค่าความแม่นยำ ความค่าระลึก และค่าประสิทธิภาพ

ซึ่งจะคำนวณจากปริมาณคำที่ทับซ้อนกัน (Overlap) เมื่อเปรียบเทียบชุดของคำที่ระบบคัดเลือก กับชุดคำของคำพิพากษาศาลฎีกา โดยกำหนดวิธีการคำนวณความแม่นยำ ค่าความระลึก และค่าประสิทธิภาพ ดังนี้

$$\textbf{ค่าความแม่นยำ (Precision)} = \frac{\text{จำนวนคำที่ทับซ้อนกันระหว่าง คำที่ระบบคัดเลือก และคำที่ปรากฏในคำพิพากษา}}{\text{จำนวนคำที่ระบบคัดเลือก}}$$

$$\textbf{ค่าความระลึก (Recall)} = \frac{\text{จำนวนคำที่ทับซ้อนกันระหว่าง คำที่ระบบคัดเลือก และคำที่ปรากฏในคำพิพากษา}}{\text{จำนวนคำในคำพิพากษาศาลฎีกา}}$$

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

จากที่ผู้วิจัยได้ศึกษา ทบทวนงานวิจัยเกี่ยวกับการสกัดคำสำคัญของเอกสารวิธีต่างๆ ได้แก่ วิธี LSA และTF-IDF ผู้วิจัยได้นำวิธีการดังกล่าวมาประยุกต์ใช้กับการสกัดคำสัญจากคำพิพากษาศาลฎีกาเพื่อสร้างรายงานประจำวัน เพื่อเป็นเกณฑ์มาตรฐาน (Benchmark) ในการทดสอบนี้ได้สุ่มตัวอย่างของคำพิพากษาศาลฎีกามาจำนวน 1 รายการ และกำหนดให้มีคำที่เกี่ยวข้อง (Coverage Word) หน้าและหลังคำสำคัญ จำนวน 3 คำ (N=3)

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

4.2.1 การทดสอบผลลัพธ์ที่ถูกสกัดคำสำคัญด้วยวิธี TF-IDF และ LSA

ตารางที่ 11 ผลลัพธ์การทดสอบการสกัดกิจกรรมสำคัญด้วยวิธี TF-IDF และ LSA ที่ N=3

วิธีการ	ค่าความแม่นยำ	ค่าความระลึก	ค่าประสิทธิภาพ
TF-IDF	0.800	0.333	0.470
LSA	0.960	0.500	0.657

จากตารางที่ 11 แสดงให้เห็นว่ามีค่าความแม่นยำของผลลัพธ์จากที่ LSA สูงกว่าวิธี TF-IDF แสดงว่าคำสำคัญที่ถูกสกัดด้วยวิธี LSA ปรากฏอยู่ในคำพิพากษาศาลฎีกามากกว่าคำที่ถูกที่สกัดด้วยวิธี TF-IDF ส่วนค่าความระลึกวิธีการ LSA ทำได้ดีกว่า TF-IDF แสดงว่าคำสำคัญที่ถูกสกัดออกมาด้วยวิธี LSA มีความครอบคลุมกับคำพิพากษามากกว่าวิธี TF-IDF

4.2.2 การทดสอบผลลัพธ์เปรียบเทียบวิธีการทางกราฟเพื่อสกัดคำสำคัญ

ผู้วิจัยใช้วิธีค่ากลางเพจแรงค์ ค่ากลางความเป็นศูนย์กลาง ค่ากลางความใกล้ ค่ากลางคั่นกลาง และค่ากลางไอเกนเวกเตอร์ในการคัดเลือกและสกัดคำหรือกิจกรรมที่สำคัญออกมาจากคำพิพากษาศาลฎีกา โดยหาค่าความแม่นยำ ค่าความระลึก และค่าประสิทธิภาพ เทียบกับคำพิพากษาศาลฎีกา สำหรับการทดสอบนี้ผู้วิจัยได้เลือกคำพิพากษามา 1 รายการ เพื่อสร้างกราฟสรุปและสกัดคำสำคัญด้วยวิธีการท่องกราฟ

ตารางที่ 12 ผลลัพธ์การทดสอบการสกัดกิจกรรมสำคัญด้วยวิธีการทางกราฟ ที่ N=3

วิธีวัดความสำคัญของเครือข่ายด้วยค่ากลาง	ค่าความแม่นยำ	ค่าความระลึก	ค่าประสิทธิภาพ
ค่ากลางคั่นกลาง	0.781	0.675	0.724
ค่ากลางความเป็นศูนย์กลาง	0.741	0.621	0.676
ค่ากลางความใกล้	0.781	0.675	0.724
ค่ากลางไอเกนเวกเตอร์	0.757	0.675	0.714
ค่ากลางเพจแรงค์	0.781	0.675	0.724

จากตารางที่ 12 เป็นผลการทดสอบท่องกราฟสรุปที่ได้จากวิธีการทั้ง 5 วิธี และเลือกคำสำคัญที่มีค่ากลางสูงที่สุด จำนวน 10 ครั้ง ทั้งนี้ ได้เลือกคำที่ปรากฏอยู่หน้า และหลังของคำสำคัญในคำพิพากษานั้น ๆ จำนวน N คำ ซึ่งการทดสอบนี้กำหนดให้ N=3 ไปวัดผลเทียบกับคำพิพากษาศาลฎีกา จำนวน 1 รายการ พบว่า วิธีค่ากลางคั่นกลาง ค่ากลางความใกล้ และค่ากลางเพจแรงก์ มีค่าความแม่นยำเท่ากัน และสูงที่สุด แสดงให้เห็นว่าคำสำคัญที่ถูกสกัดด้วยวิธีดังกล่าวปรากฏอยู่ในคำพิพากษามากที่สุด ส่วนค่าความระลึก วิธีค่ากลางคั่นกลาง ค่ากลางความใกล้ ค่ากลางไอเกนเวกเตอร์ และ ค่ากลางเพจแรงค์ ให้ค่าความระลึกเท่ากัน และเป็นค่าที่สูงที่สุด ส่วนวิธีค่ากลางความเป็นศูนย์กลาง มีค่าความระลึกน้อยที่สุด แสดงให้เห็นว่า วิธีค่ากลางความเป็นศูนย์กลาง เมื่อสกัดคำสำคัญออกมา ผลลัพธ์ที่ได้มีความครอบคลุมกับคำพิพากษาศาลฎีกาน้อยกว่าวิธีอื่นที่กล่าวมา

จากการทดสอบดังตารางที่ 11 และ 12 จะเห็นได้ว่าเป็นการทดสอบแต่เพียงคำพิพากษาศาลฎีกาแค่รายการเดียว ซึ่งไม่อาจบอกผลลัพธ์ที่ได้โดยแน่ชัดว่า วิธีการท่องกราฟที่สร้างขึ้นมีประสิทธิภาพมากน้อยเพียงใด และในข้อเท็จจริง การรับแจ้งความร้องทุกข์ของพนักงานสอบสวนนั้น จะรับแจ้งความร้องทุกข์ในความผิดคดีจราจรทุกเรื่อง ตาม พ.ร.บ.จราจรทางบก พ.ศ.2522 เพราะฉะนั้น การทดสอบแค่คำพิพากษาศาลฎีกาเพียงแค่กรณีเดียว อาจจะยังไม่เพียงพอสำหรับการสร้างแม่แบบสำหรับการบันทึกรายงานประจำวัน จากเหตุผลนี้อาจส่งผลให้การสร้างรายงานประจำวันเกิดความไม่ครอบคลุมตามถึงรายละเอียด พฤติการณ์ ทั้งหมดของการรับแจ้งความร้องทุกข์ในเรื่องนั้น ๆ และไม่สามารถแสดงประสิทธิภาพของวิธีการต่างๆ ที่นำมาใช้ได้ ถ้าข้อมูลที่ถูกนำมาทดสอบมีการกระจายตัวไม่ดีพอ ดังนั้น จึงจำเป็นต้องมีการทดสอบแบบไขว้ (Cross Validation) เพื่อทดสอบว่า เมื่อผู้วิจัยแบ่งข้อมูลออกเป็นชุดทำให้ข้อมูลมีการกระจายตัว ผลลัพธ์ที่ได้จะเป็นอย่างไร และสามารถบ่งชี้ได้อย่างไรว่าข้อมูลหรือวิธีการสกัดคำสำคัญที่เลือกนำมาใช้ทั้งหมดมีประสิทธิภาพเป็นอย่างไร และครอบคลุมในกรณีที่เกิดเหตุ หรือคดีหลายๆ ประเภทหรือไม่

4.2.3 การทดสอบแบบไขว้ (Cross Validation)

เป็นเทคนิคในการวัดประสิทธิภาพการทำงานของแบบจำลอง โดยมีวิธีการที่ผู้วิจัยแบ่งข้อมูลเป็นจำนวน 10 ส่วนโดยการในแต่ละส่วนจะต้องมาจากสุ่มเพื่อที่จะให้ข้อมูลมีการกระจายตัวเท่าๆ กัน เพื่อสร้างชุดข้อมูลสำหรับการทดสอบกระบวนการสกัดกิจกรรมสำคัญ ซึ่งในแต่ละรอบการทำงานระบบจะแบ่งข้อมูลออกเป็น 2 ชุด ประกอบด้วย ชุดสำหรับการฝึกฝน เพื่อสร้างกราฟสรุปและทดสอบการสกัดกิจกรรมสำคัญ และชุดทดสอบ รายละเอียดซอร์สโค้ดการทำงานปรากฏตามภาพที่ 44 โดยในการทดสอบนี้ผู้วิจัยจะแบ่งการสุ่มชุดข้อมูล (K) ออกเป็น 10 ชุด (10-fold Cross Validation)

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

```
for train_index, test_index in kf.split(train_data):
    print('--------------start----------------')
    X_train = train_data.iloc[train_index].loc[:, features]
    X_test = train_data.iloc[test_index][features]

    y_train = train_data.iloc[train_index].loc[:,'target']
    y_test = train_data.loc[test_index]['target']

    df_train = X_train.assign(target=y_train)
    df_test = X_test.assign(target=y_test)
```

ภาพที่ 44 ซอร์สโค้ดการแบ่งชุดข้อมูลสำหรับ K-fold Cross Validation

ตารางที่ 13 ผลลัพธ์การทดสอบการสกัดกิจกรรมสำคัญ ด้วย 10-fold Cross Validation

วิธีการที่ใช้	ค่าความแม่นยำเฉลี่ย	ค่าความระลึกเฉลี่ย	ค่าประสิทธิภาพเฉลี่ย
TF-IDF	0.650	0.350	0.454
LSA	0.847	0.473	0.606
ค่ากลางคั่นกลาง	0.759	0.701	0.727
ค่ากลางความเป็นศูนย์กลาง	0.756	0.692	0.722
ค่ากลางความใกล้	0.751	0.672	0.709
ค่ากลางไอเกนเวกเตอร์	0.750	0.677	0.712
ค่ากลางเพจแรงค์	0.758	0.694	0.725

* ผลลัพธ์แตกต่างอย่างมีระดับนัยสำคัญที่ 0.01 ผลการวิเคราะห์แสดงดังตารางที่ 14 และ 15

จากผลลัพธ์การทดสอบตามตารางที่ 13 แสดงค่าเฉลี่ยของค่าความแม่นยำ ค่าความระลึก และค่าประสิทธิภาพ การทดสอบด้วย 10-fold Cross Validation วิธีการสกัดคำสำคัญด้วย LSA ให้ค่าความแม่นยำมากที่สุดคือ 0.847 แสดงว่า คำส่วนใหญ่ที่ถูกสกัดออกมา ตรงตามคำที่ปรากฏอยู่ในคำพิพากษาศาลฎีกา ในส่วนของค่าความระลึก วิธีค่ากลางคั่นกลางให้ค่าที่สูงสุด

4.2.4 การทดสอบค่าเฉลี่ย 2 กลุ่ม (Paired – Samples T Test)

เมื่อได้ผลลัพธ์จากการทดสอบด้วย 10-fold Cross Validation แล้ว ขั้นตอนต่อไปจะทำการทดสอบค่าเฉลี่ย 2 กลุ่มที่ไม่เป็นอิสระจากกัน เพื่อเปรียบเทียบระหว่างผลลัพธ์ของวิธีการสกัดคำสำคัญระหว่างวิธีที่นำมาเปรียบเทียบ ประกอบด้วย วิธี TF-IDF และ LSA เทียบกับวิธีการทางกราฟทั้ง 5 วิธี โดยนำผลลัพธ์ของแต่ละชุด มาเปรียบเทียบกันจำนวน 10 ชุด โดยมีรายละเอียดดังต่อไปนี้

(1) กำหนดสมมติฐาน

- สมมติฐานว่าง (Null Hypothesis: H_0) กำหนดขึ้นมาเพื่อแสดงให้เห็นว่าผลลัพธ์ระหว่างการสกัดความสำคัญทั้งสองวิธีไม่มีความแตกต่างกัน

- สมมติฐานทางเลือก (Alternative Hypothesis : H_1) กำหนดขึ้นมาเพื่อแสดงให้เห็นว่า ผลลัพธ์ระหว่างวิธีการสกัดความสำคัญทั้งสองวิธีมีความแตกต่างกัน

โดยมีสมการของสมมติฐาน ดังนี้

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

(2) กำหนดระดับนัยสำคัญที่ 0.01 ($\alpha = 0.01$)

เมื่อกำหนดสมมติฐาน และระดับนัยสำคัญ ก็จะมาคำนวณค่า p-value คือ ค่าความน่าจะเป็นที่จะปฏิเสธสมมติฐานว่าง (Null Hypothesis: H_0) ซึ่ง ได้ผลลัพธ์ดังตารางที่ 14 และ 15

ตารางที่ 14 ตารางผลลัพธ์ค่าเฉลี่ย 2 กลุ่มเทียบกับ TF-IDF

	TF-IDF	Betweenness Centrality	Degree Centrality	Closeness Centrality	Eigenvector Centrality	PageRank
Mean	0.4538	0.7274	0.7218	0.7218	0.7085	0.7110
Variance	0.0020	0.0007	0.0010	0.0010	0.0012	0.0009
p-value	-	0.00000002	0.00000002	0.00000001	0.00000001	0.00000002

ตารางที่ 15 ตารางผลลัพธ์ค่าเฉลี่ย 2 กลุ่มเทียบกับ LSA

	LSA	Betweenness Centrality	Degree Centrality	Closeness Centrality	Eigenvector Centrality	PageRank
Mean	0.6056	0.7274	0.7218	0.7218	0.7085	0.7110
Variance	0.0031	0.0007	0.0010	0.0010	0.0012	0.0009
p-value	-	0.00000002	0.00000993	0.00000993	0.00004919	0.00005558

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

ค่า p-value ของวิธีทางกราฟแต่ละวิธี เปรียบเทียบกับวิธี TF-IDF และ LSA มีค่าน้อยกว่าระดับนัยสำคัญที่ 0.01 ทุกวิธี ดังนั้นการทดสอบนี้ตัดสินใจปฏิเสธสมมติฐานว่าง สรุปได้ว่า ผลลัพธ์ระหว่างวิธีการสกัดความสำคัญด้วยวิธี TF-IDF และ LSA มีความแตกต่างกันกับผลลัพธ์ของการสกัดคำสำคัญด้วยวิธีการทางกราฟทั้ง 5 แบบ อย่างมีระดับนัยสำคัญที่ 0.01

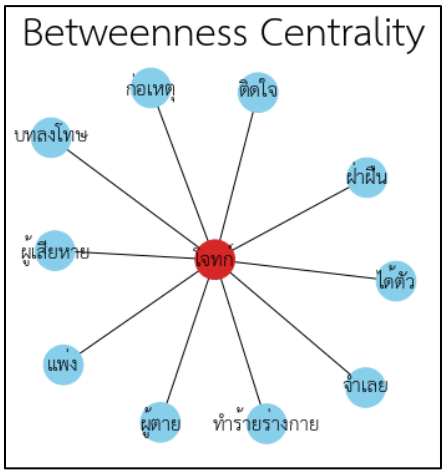
ทั้งนี้ จากการทดสอบดังกล่าว แสดงให้เห็นว่าทุกวิธีนั้น มีสามารถสกัดคำสำคัญออกจากคำพิพากษา ซึ่งแต่ละวิธีก็จะมีค่าความแม่นยำ และค่าความระลึกที่แตกต่างกันออกไป

4.3 การสร้างผลลัพธ์จากการสกัดกิจกรรมของคำพิพากษาศาลฎีกา (Generate a police daily report)

การสร้างผลลัพธ์จากการสกัดกิจกรรมของคำพิพากษาศาลฎีกา กล่าวคือ แม่แบบสำหรับรายงานประจำวัน นั้น ขั้นตอนวิธีในการสร้างกราฟของงานวิจัยนี้ ผู้วิจัยได้ออกแบบขั้นตอนการสร้างกราฟโดยให้ความสำคัญกับน้ำหนักของแต่ละโหนดเป็นสำคัญ โดยเลือกโหนดที่มีน้ำหนักเส้นเชื่อมสูงที่สุด ในที่นี้ผู้วิจัยให้ค่าน้ำหนักระหว่างโหนดด้วยค่าความคล้ายคลึงโดยการเปลี่ยนคำให้อยู่ในรูปแบบเวกเตอร์ และวัดค่าความเหมือนของเวกเตอร์ทั้งสองว่าไปในทิศทางเดียวกันมากน้อยเพียงใด และกำหนดให้โหนดตัวที่ถูกเลือกเป็นโหนดถัดไปสำหรับการท่องกราฟ หรือท่องกราฟออกไปจนกว่าระบบจะหยุดทำงานโดยผู้ใช้เป็นผู้กำหนด

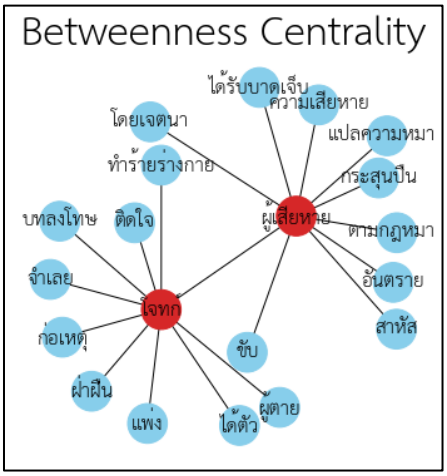
การสร้างข้อความจากคำพิพากษาศาลฎีกาทั้งหมด เพื่อนำคำได้มาเป็นตัวเลือกสำหรับสร้างแม่แบบของรายงานประจำวัน ประกอบด้วย ข้อมูลสถานที่เกิดเหตุ ชื่อผู้เสียหาย ชื่อผู้ต้องหา พฤติการณ์แห่งคดี เป็นหลัก และส่วนท้ายของเอกสาร ส่วนนี้จะเป็นส่วนของการสร้างผลลัพธ์รายงานประจำวันออกมา โดยการสร้างแม่แบบจะถูกสร้างขึ้นจากคำพิพากษาศาลฎีกา ในการสร้างแม่แบบของรายงานประจำวันนั้น จะเลือกคำที่ต้องการมาเพื่อเป็นคำสำคัญหรือหัวข้อในการบันทึกรายงานประจำวัน ที่จำเป็นและเกี่ยวข้องกับเหตุการณ์ของคดีที่เกิดขึ้น

เมื่อผู้วิจัยได้สร้างกราฟสรุปทั้งหมดได้แล้วนั้น การสร้างแม่แบบจะถูกเลือกจากโหนดที่ปรากฏ การกำหนดโหนดเริ่มต้น ขึ้นอยู่กับสถานะของผู้ที่เข้ามาแจ้งความ เช่น ผู้เสียหาย เป็นต้น ทั้งนี้ผู้วิจัยกำหนดโหนดเริ่มต้นไว้ที่ โหนด “ผู้เสียหาย” ซึ่งระบุถึง บุคคลที่เป็นฝ่ายร้องทุกข์หรือแจ้งความ ซึ่งก็ตกเป็นผู้เสียหาย เพื่อเลือกโหนดสำหรับการสร้างรายงานประจำวัน จากนั้นระบบจะคัดเลือกโหนดต่างๆ ที่มีความคล้ายคลึงมากที่สุด N ลำดับ และอยู่ในชุดของคำสำคัญ โดยชุดของคำสำคัญคัดเลือกจากคำที่มีค่ากลางมากที่สุดจำนวนร้อยละ 33 ของคำทั้งหมด รายละเอียดตามภาพที่ 45 ทั้งนี้การเลือกโหนดแต่ละครั้งนั้น ผู้ใช้จะเลือกคำที่เหมาะสมที่สุดสำหรับการบันทึกรายงานประจำวัน ซึ่งขึ้นอยู่กับตัวผู้ใช้ โดยเงื่อนไขของการทำงานของระบบนี้ จะหยุดก็ต่อเมื่อโหนดที่ผู้ใช้เลือก ไม่ได้เชื่อมกับโหนดใดอีกเลย หรือหากผู้ใช้พอใจกับรายละเอียดที่ได้จากที่กราฟได้สร้างแม่แบบออกมาแล้ว ก็สามารถหยุดได้ ทั้งนี้ ขึ้นอยู่กับสถานการณ์ที่เกิดขึ้นว่ามีความซับซ้อนมากน้อยเพียงใด โดยปรากฏขั้นตอนการทำงานของระบบดังตารางที่ 16 - 20



ภาพที่ 45 การกำหนดโหนดเริ่มต้น และแสดงโหนดที่เกี่ยวข้องสำหรับการท่องกราฟ

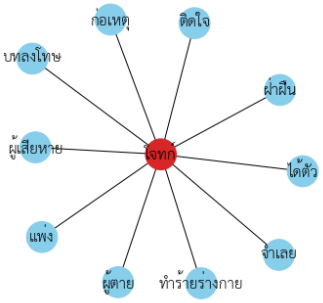
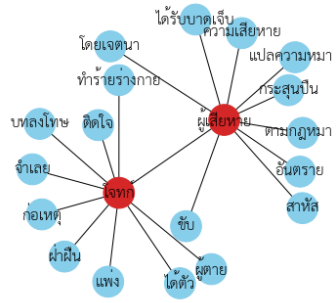
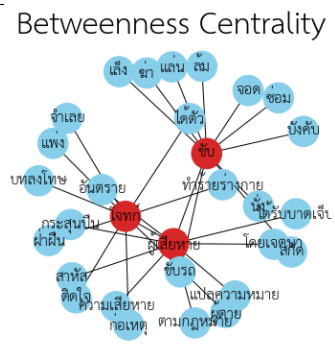
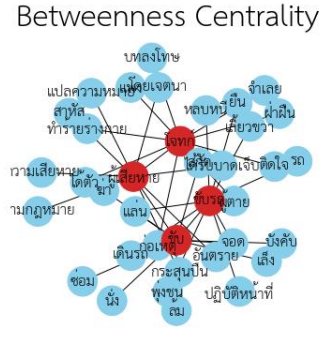
จากภาพที่ 45 เป็นการสร้างกราฟเพื่อสกัดกิจกรรมสำคัญ ผู้วิจัยกำหนดให้เป็นกรณีที่เกิดเหตุรถชนอันดับแรก ในการสร้างกราฟ โหนดเริ่มต้น จะถูกกำหนดให้เป็นคำว่า “โจทก์” เมื่อสร้างกราฟเสร็จแล้วในรอบแรก ผู้วิจัยเลือกโหนด “ผู้เสียหาย” จากโหนดดังกล่าว ก็จะผ่านกระบวนการเดิม และสร้างคัดเลือกโหนดที่มีความสำคัญหรือเกี่ยวข้องกับโหนดที่เลือกออกมา ดังภาพที่ 50 ซึ่งโหนดที่แสดงออกมานั้น จะถูกเลือกด้วยผู้ใช้ โดยระบบจะแสดงโหนดที่เกี่ยวข้องทั้งหมด และเรียงลำดับค่าความสำคัญของแต่ละโหนด เพื่อให้ผู้ใช้เลือกโหนดที่ต้องการเรียงจากโหนดที่มีค่าความสำคัญมากที่สุดลงมา 10 อันดับ และโหนดที่ถูกเลือกก็จะแตกหรือแสดงโหนดที่มีความสำคัญกับโหนดดังกล่าว ซึ่งสามารถแสดงโหนดทั้งหมดให้ผู้ใช้เลือกสำหรับการท่องกราฟไปอย่างต่อเนื่อง ซึ่งในแต่ละรอบการทำงาน ระบบจะแสดงกราฟที่แตกต่างกันออกไปตามโหนดที่ถูกเลือก ดังภาพที่ 46



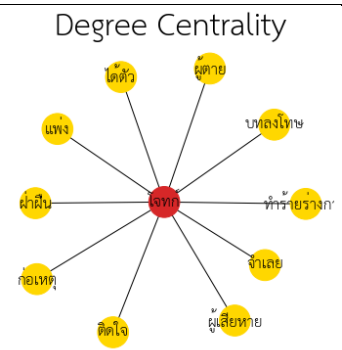
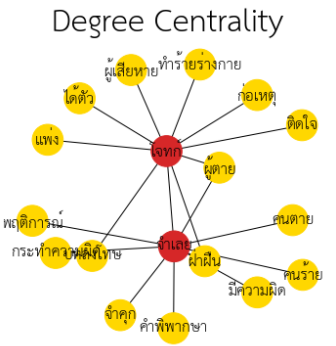
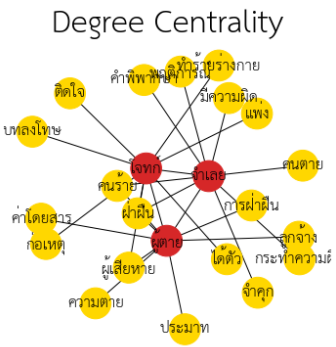
ภาพที่ 46 การท่องกราฟไปยังโหนดถัดไป

*** กำหนดให้ โหนดสีแดง คือคำที่เลือก**

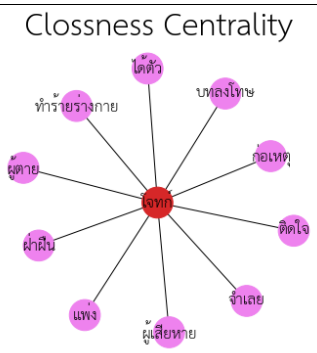
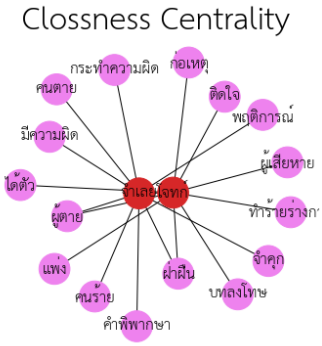
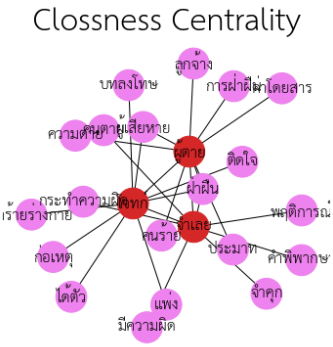
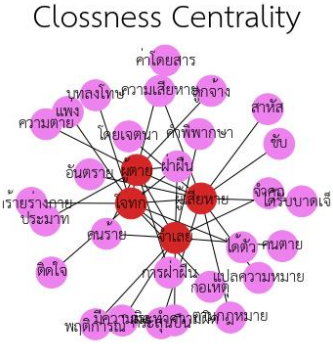
ตารางที่ 16 ตัวอย่างการท่องกราฟ และโหนดที่ถูกคัดเลือกด้วยวิธีค่ากลางคั่นกลาง

คำที่เลือก	กราฟที่ได้	คำสำคัญที่ได้
โจทก์	Betweenness Centrality 	ผู้เสียหาย บทลงโทษ จำเลย ทำร้ายร่างกาย ก่อเหตุ ฝ่าฝืน ติดใจ แพ่ง ได้ตัว ผู้ตาย
ผู้เสียหาย	Betweenness Centrality 	โดยเจตนา แปลความหมาย โจทก์ กระสุนปืน สาหัส ตามกฎหมาย ความเสียหาย ได้รับบาดเจ็บ อันตราย ขับ
ขับ	Betweenness Centrality 	ล้ม ขับรถ บังคับ สกัด ฆ่า แล่น ซ่อม จอด นั่ง เล็ง
ขับรถ	Betweenness Centrality 	แล่น เดินรถ หลบหนี พุ่งชน เลี้ยวขวา รถ ขับ ยืน ปฏิบัติหน้าที่ จอด

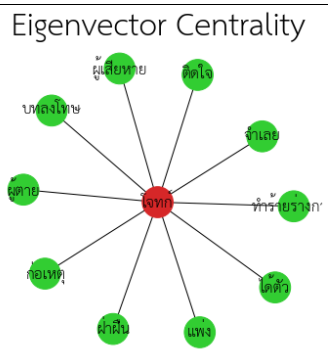
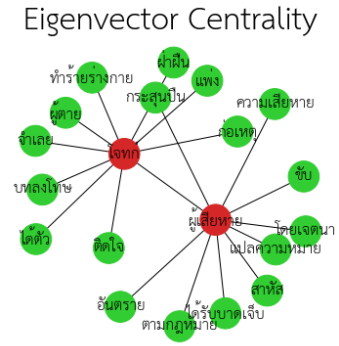
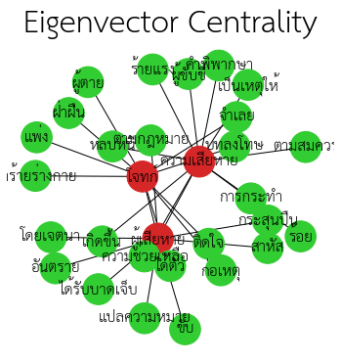
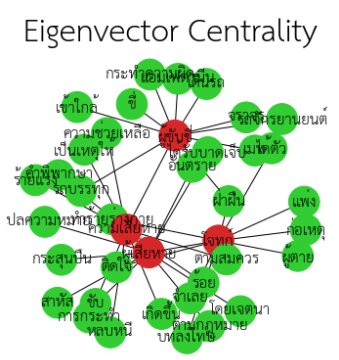
ตารางที่ 17 ตัวอย่างการท่องกราฟ และโหนดที่ถูกคัดเลือกด้วยวิธีค่ากลางความเป็นศูนย์กลาง

คำที่เลือก	กราฟที่ได้	คำสำคัญที่ได้
โจทก์		ผู้เสียหาย บทลงโทษ จำเลย ทำร้ายร่างกาย ก่อเหตุ ฝ่าฝืน ติดใจ แพ่ง ได้ตัว ผู้ตาย
จำเลย		โจทก์ คำพิพากษาศาลฎีกา ผู้ตาย มีความผิด กระทำความผิด พฤติการณ์ คนร้าย จำคุก ฝ่าฝืน คนตาย
ผู้ตาย		โจทก์ ผู้เสียหาย คนร้าย ฝ่าฝืน จำเลย ลูกจ้าง ประมาท ค่าโดยสาร ความตาย การฝ่าฝืน

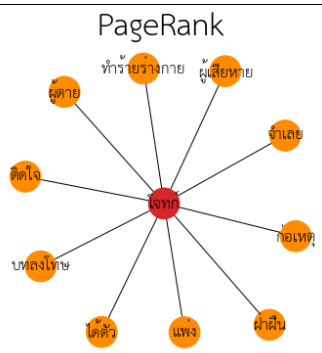
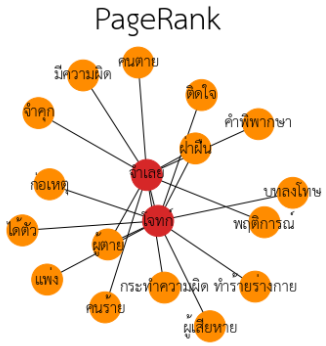
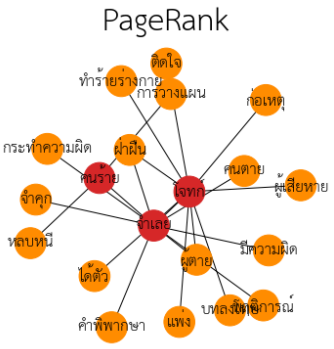
ตารางที่ 18 ตัวอย่างการท่องกราฟ และโหนดที่ถูกคัดเลือกด้วยวิธีค่ากลางความใกล้

คำที่เลือก	กราฟที่ได้	คำสำคัญที่ได้
โจทก์		ผู้เสียหาย บทลงโทษ จำเลย ทำร้ายร่างกาย ก่อเหตุ ฝ่าฝืน ติดใจ แพ่ง ได้ตัว ผู้ตาย
จำเลย		โจทก์ คำพิพากษาศาลฎีกา ผู้ตาย มีความผิด กระทำความผิด พฤติการณ์ คนร้าย จำคุก ฝ่าฝืน คนตาย
ผู้ตาย		โจทก์ ผู้เสียหาย คนร้าย ฝ่าฝืน จำเลย ลูกจ้าง ประมาท ค่าโดยสาร ความตาย การฝ่าฝืน
ผู้เสียหาย		โดยเจตนา แปลความหมาย โจทก์ กระสุนปืน สาหัส ตามกฎหมาย ความเสียหาย ได้รับบาดเจ็บ อันตราย ขับ

ตารางที่ 19 ตัวอย่างการท่องกราฟ และโหนดที่ถูกคัดเลือกด้วยวิธีค่ากลางไอเกนเวกเตอร์

คำที่เลือก	กราฟที่ได้	คำสำคัญที่ได้
โจทก์		ผู้เสียหาย บทลงโทษ จำเลย ทำร้ายร่างกาย ก่อเหตุ ฝ่าฝืน ติดใจ แพง ได้ตัว ผู้ตาย
ผู้เสียหาย		โจทก์ คำพิพากษาศาลฎีกา ผู้ตาย มีความผิด กระทำความผิด พฤติการณ์ คนร้าย จำคุก ฝ่าฝืน คนตาย
ความเสียหาย		การกระทำ คำพิพากษาศาลฎีกา ความช่วยเหลือ ร้ายแรง เป็นเหตุให้ หลบหนี ตามสมควร ร้อย เกิดขึ้น ผู้ขับขี่ ผู้ขับขี่
ผู้ขับขี่		แอมเฟตามีน จราจร ฝ่าฝืน รถจักรยานยนต์ เข้าใกล้ กระทำความผิด ขี่ เมท รถบรรทุก เดินรถ รถจักรยานยนต์

ตารางที่ 20 ตัวอย่างการท่องกราฟ และโหนดที่ถูกคัดเลือกด้วยวิธีค่ากลางแพจแรงก์

คำที่เลือก	กราฟที่ได้	คำสำคัญที่ได้
โจทก์		ผู้เสียหาย บทลงโทษ จำเลย ทำร้ายร่างกาย ก่อเหตุ ฝ่าฝืน ติดใจ แพ่ง ได้ตัว ผู้ตาย
จำเลย		โจทก์ คำพิพากษาศาลฎีกา ผู้ตาย มีความผิด กระทำความผิด พฤติการณ์ คนร้าย จำคุก ฝ่าฝืน คนตาย
คนร้าย		หลบหนี การวางแผน

ในส่วนต่อไปผู้วิจัยจะกำหนดสถานการณ์จำลอง (Scenario) สำหรับการรับแจ้งความร้องทุกข์เพื่อลงบันทึกประจำวัน

สถานการณ์จำลองที่ 1 - ผู้เสียหายกำลังขับรถไปตามถนนและเกิดเฉี่ยวชนกับรถจักรยานยนต์อีกคันหนึ่ง ในกรณีการทดสอบสถานการณ์จำลองที่ 1 คู่กรณีทั้งสองเดินทางมาที่สถานีตำรวจเพื่อเข้ามาแจ้งความ โดยบุคคลหนึ่งในเหตุการณ์จะมีสถานะเป็น ผู้เสียหาย ดังนั้น ระบบจึงเลือกโหนด **ผู้เสียหาย** เป็นโหนดแรกสำหรับการท่องกราฟ โดยผู้วิจัยจะทดสอบท่องกราฟที่ถูกสรุปด้วยวิธีทางกราฟทั้ง 5 แบบ รายละเอียด ดังนี้

1) วิธีค่ากลางคั่นกลาง

ภาพที่ 47 กราฟแสดงโหนดที่ถูกเลือก (สีแดง) และโหนดที่เกี่ยวข้อง (สีฟ้า)

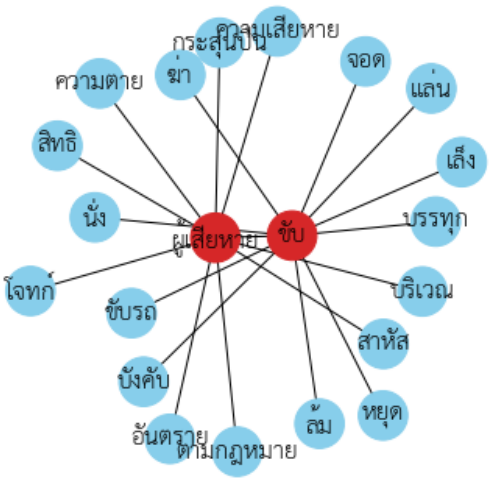
จากภาพที่ 47 เมื่อกำหนดให้โหนดเริ่มต้นเป็น **ผู้เสียหาย** ระบบจะแสดงคำที่เกี่ยวข้องขึ้นมาให้ประกอบด้วย **โดยเจตนา แปลความหมาย โจทก์ กระสุนปืน สาหัส ตามกฎหมาย ความเสียหาย ได้รับบาดเจ็บ อันตราย ขับ** ตามภาพที่ 48

	target	centrality_value
0	โจทก์	0.117508
1	ขับ	0.065891
2	บริเวณ	0.022868
3	ความเสียหาย	0.007796
4	อันตราย	0.005471
5	กระสุนปืน	0.004973
6	ตามกฎหมาย	0.004299
7	สิทธิ	0.004169
8	สาหัส	0.003943
9	ความตาย	0.003383

ภาพที่ 48 โหนดที่เกี่ยวข้อง กับ "ผู้เสียหาย" และมีค่าความสำคัญสูงสุด 10 อันดับ

ระบบแสดงโหนดที่เกี่ยวข้องขึ้นมา พร้อมค่าความสำคัญของแต่ละโหนด เพื่อให้ผู้ใช้สามารถเลือกโหนดที่ต้องการจะใช้ได้ โดยเลือกจากโหนดที่มีค่าความสำคัญสูงที่สุดก่อน หากว่าโหนดดังกล่าวไม่เกี่ยวข้อง ก็ให้ผู้ใช้เลือกโหนดถัดไปที่มีค่าความสำคัญรองลงมา ในกรณีนี้ คำที่มีค่าความสำคัญสูงที่สุด คือ โจทก์ ซึ่งคำว่าโจทก์หมายถึง ผู้ฟ้องคดี และเป็นคนเดียวกับผู้เสียหาย ที่ผู้วิจัยได้เลือกไปแล้ว

ดังนั้นผู้ใช้จึงไม่เลือกโหนดดังกล่าว เมื่อดูโหนดถัดไป คือ ขับ ซึ่งเกี่ยวข้องกับเหตุการณ์จำลอง ผู้ใช้จึงเลือกโหนด ขับ ระบบก็จะแสดงคำที่เกี่ยวข้อง ดังภาพที่ 49



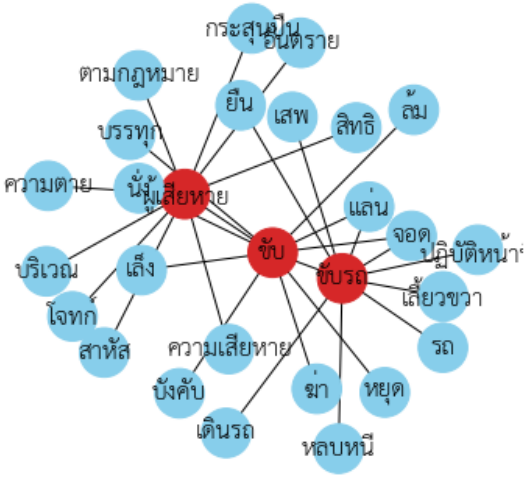
ภาพที่ 49 กราฟแสดงโหนดที่ถูกเลือก (สีแดง) และโหนดที่เกี่ยวข้อง (สีฟ้า)

คำที่มีความเกี่ยวข้องกับคำว่า **ขับ** ประกอบด้วย **ขับรถ บังคับ หยุด แล่น จอด บรรทุก นั่ง ล้ม ฆ่า เล็ง** เมื่อผู้วิจัยดูที่ค่าความสำคัญของโหนดที่ระบบแสดงขึ้นมา

	target	centrality_value
0	ขับรถ	0.035019
1	บังคับ	0.014745
2	หยุด	0.014201
3	แล่น	0.006828
4	จอด	0.006822
5	บรรทุก	0.006563
6	นั่ง	0.002862
7	ล้ม	0.002286
8	ฆ่า	0.001771
9	เล็ง	0.001190

ภาพที่ 50 โหนดที่เกี่ยวข้อง กับ “ขับ” และมีค่าความสำคัญสูงสุด 10 อันดับ

ในกรณีนี้ จากภาพที่ 50 แสดงคำที่มีค่าความสำคัญสูงที่สุด คือคำว่า **ขับรถ** ซึ่งคำที่เกี่ยวข้องกับสถานการณ์จำลองเนื่องจากเป็นเหตุรถชนที่คู่กรณีทั้งสองฝ่ายขับขี่ยานพาหนะ ผู้ใช้จึงเลือก **ขับรถ** เพื่อซักถามรายละเอียดเกี่ยวกับผู้ขับขี่ยานพาหนะ ชื่อ-สกุล ที่อยู่

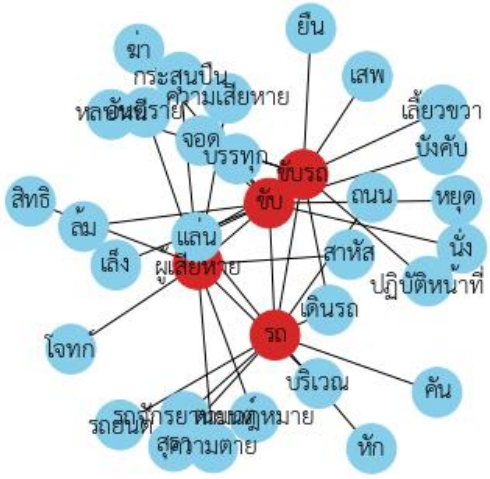


ภาพที่ 51 กราฟแสดงโหนดที่ถูกเลือก (สีแดง) และโหนดที่เกี่ยวข้อง (สีฟ้า)

	target	centrality_value
0	รถ	0.101515
1	ขับ	0.065891
2	เดินรถ	0.010612
3	เสพ	0.007079
4	แล่น	0.006828
5	จอด	0.006822
6	หลบหนี	0.003038
7	ยืน	0.002662
8	ปฏิบัติหน้าที่	0.002271
9	เลี้ยวขวา	0.000532

ภาพที่ 52 โหนดที่เกี่ยวข้อง กับ “ขับรถ” และมีค่าความสำคัญสูงสุด 10 อันดับ

เมื่อเลือกโหนดแล้วระบบก็จะแสดงกราฟ และคำที่เกี่ยวข้องขึ้นมา ดังภาพที่ 51 และ 52 ประกอบด้วย **รถ ขับ เดินรถ เสพ แล่น จอด หลบหนี ยืน ปฏิบัติหน้าที่ เลี้ยวขวา** คำที่มีค่าความสำคัญสูงสุด คือคำว่ารถ ผู้ใช้จึงเลือกโหนด รถ เพื่อสอบถามข้อมูลเกี่ยวกับรถที่ผู้เสียหายขับขี่

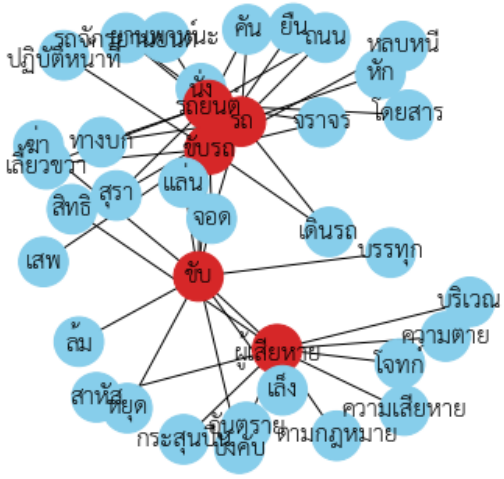


ภาพที่ 53 กราฟแสดงโหนดที่ถูกเลือก (สีแดง) และโหนดที่เกี่ยวข้อง (สีฟ้า)

	target	centrality_value
0	รถยนต์	0.057557
1	ถนน	0.044556
2	ทางบก	0.033707
3	รถจักรยานยนต์	0.022497
4	คัน	0.011033
5	เดินรถ	0.010612
6	แล่น	0.006828
7	จอด	0.006822
8	หัก	0.002377
9	สุรา	0.000679

ภาพที่ 54 โหนดที่เกี่ยวข้อง กับ “รถ” และมีค่าความสำคัญสูงสุด 10 อันดับ

เมื่อเลือกคำว่า **รถ** ระบบแสดงโหนดที่เกี่ยวข้อง ดังภาพที่ 53 และ 54 ในกรณีนี้คำที่ระบบแสดงขึ้นเป็นลำดับแรกคือ โหนด **รถยนต์** ผู้ใช้จึงเลือกโหนด **รถยนต์** เพื่อเป็นข้อคำถามรายละเอียดเกี่ยวกับรถยนต์ว่า ยี่ห้ออะไร รุ่นอะไร สีอะไร เป็นต้น

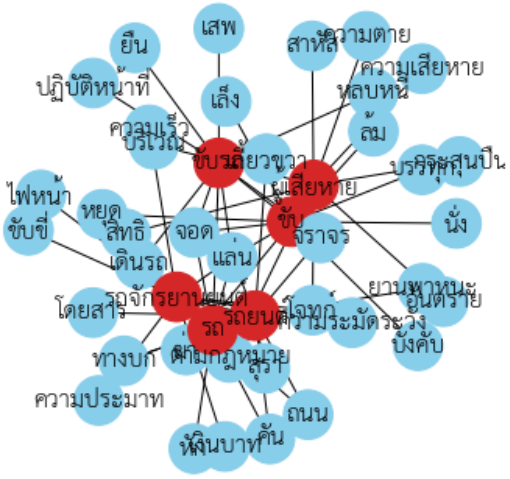


ภาพที่ 55 กราฟแสดงโหนดที่ถูกเลือก (สีแดง) และโหนดที่เกี่ยวข้อง (สีฟ้า)

	target	centrality_value
0	ถนน	0.044556
1	ทางบก	0.033707
2	จราจร	0.025102
3	รถจักรยานยนต์	0.022497
4	คัน	0.011033
5	แล่น	0.006828
6	ยานพาหนะ	0.004875
7	โดยสาร	0.001663
8	สุรา	0.000679
9	เลี้ยวขวา	0.000532

ภาพที่ 56 โหนดที่เกี่ยวข้อง กับ “รถยนต์” และมีค่ากลางสูงสุด 10 อันดับ

เมื่อเลือกโหนด **รถยนต์** แล้ว ระบบจะแสดงกราฟ และโหนดที่เกี่ยวข้องขึ้นมา ดังภาพที่ 55 และ 56 ผู้ใช้ **รถจักรยานยนต์** ซึ่งมีค่าความสำคัญสูงเป็นลำดับที่ 4 ผู้ใช้สามารถนำไปใช้เป็นข้อคำถามถึงลักษณะ หมายเลขทะเบียน สี ยี่ห้อ ของรถจักรยานยนต์หรือยานพาหนะอื่น ในอุบัติเหตุที่เกิดขึ้นได้ จากนั้นให้ผู้ใช้เลือกคำว่า **รถจักรยานยนต์**

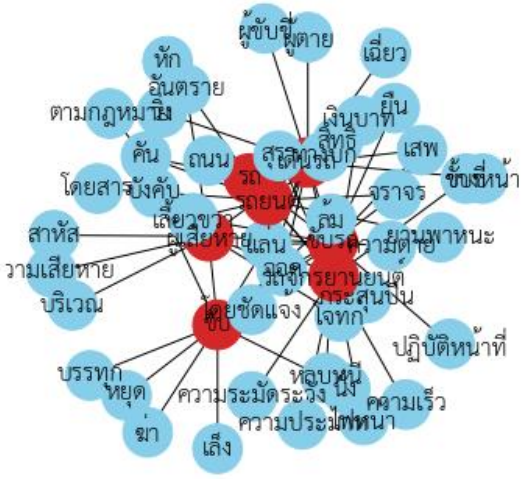


ภาพที่ 57 กราฟแสดงโหนดที่ถูกเลือก (สีแดง) และโหนดที่เกี่ยวข้อง (สีฟ้า)

	target	centrality_value
0	โจทก์	0.117508
1	เดินรถ	0.010612
2	ความเร็ว	0.009378
3	แล่น	0.006828
4	ความระมัดระวัง	0.004489
5	ความประมาท	0.004220
6	ขับขี่	0.001991
7	เงินบาท	0.001399
8	ไฟหน้า	0.001329
9	สุรา	0.000679

ภาพที่ 58 โหนดที่เกี่ยวข้อง กับ “รถจักรยานยนต์” และมีค่ากลางสูงสุด 10 อันดับ

เมื่อเลือกโหนด **รถจักรยานยนต์** แล้ว ระบบจะแสดงโหนดที่เกี่ยวข้องขึ้นมาดังภาพที่ 57 และ 58 ประกอบด้วย **โจทก์ เดินรถ ความเร็ว แล่น ความระมัดระวัง ความประมาท ขับขี่ เงินบาท ไฟหน้า สุรา** ซึ่งแต่ละโหนดแสดงสิ่งที่เกี่ยวข้องกับอุบัติเหตุ สภาพแวด ล้อม ลักษณะการขับรถ ในกรณีนี้เป็นเหตุรถเฉี่ยวชน ผู้ใช้จึงเลือกโหนด **เดินรถ** เพื่อเป็นข้อคำถามระหว่างผู้เสียหายว่า **รถยนต์** ที่ถูกเฉี่ยวชน ขับมาอย่างไร หรือขับไปในทิศทางใด

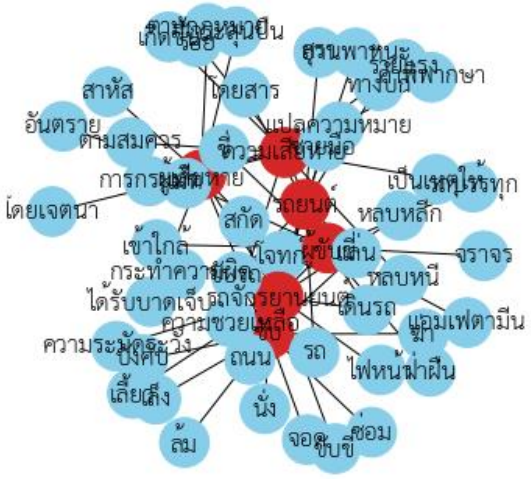


ภาพที่ 59 กราฟแสดงโหนดที่ถูกเลือก (สีแดง) และโหนดที่เกี่ยวข้อง (สีฟ้า)

	target	centrality_value
0	ถนน	0.044556
1	ผู้ตาย	0.033770
2	ผู้ขับขี่	0.031961
3	เสพ	0.007079
4	แล่น	0.006828
5	จอด	0.006822
6	เฉี่ยว	0.004748
7	ข้างหน้า	0.001727
8	โดยชัดแจ้ง	0.000488
9	วิ่ง	0.000420

ภาพที่ 60 โหนดที่เกี่ยวข้อง กับ “เดินรถ” และมีค่าความสำคัญสูงสุด 10 อันดับ

เมื่อเลือกโหนด **เดินรถ** ระบบจะแสดงโหนดที่เกี่ยวข้องดังภาพที่ 59 และ 60 ประกอบด้วย **ถนน ผู้ตาย ผู้ขับขี่ เสพ แล่น จอด เฉี่ยว ข้างหน้า โดยชัดแจ้ง วิ่ง** คำที่มีค่าความสำคัญเป็นลำดับแรก คือคำว่า **ถนน** ซึ่งสามารถบอกได้ว่า ผู้เสียหาย หรือคู่กรณี เดินรถอยู่บนถนนอะไร โดยทั้งหมดนี้คือพฤติการณ์แห่งคดี ที่สามารถบอกรายละเอียดสิ่งที่เกิดขึ้นในเหตุการณ์ดังกล่าว ทั้งนี้ระบบจะแสดงโหนดที่เพื่อให้ผู้ใช้สามารถนำโหนดมาเป็นข้อคำถามเพื่อให้รายงานประจำวันเกิดความครอบคลุมมากขึ้นได้ ขึ้นอยู่กับสถานการณ์ และพนักงานสอบสวน โดยตัวอย่างผลลัพธ์การท่องกราฟทั้งหมดปรากฏดังภาพที่ 61

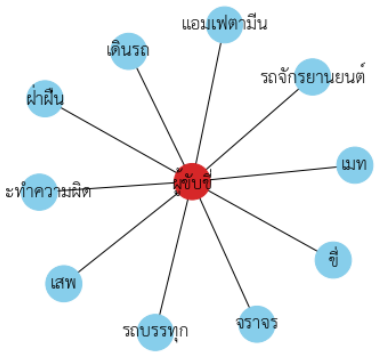


ภาพที่ 61 แสดงผลลัพธ์การท่องกราฟเพื่อสร้างแม่แบบรายงานประจำวันคดีจราจร

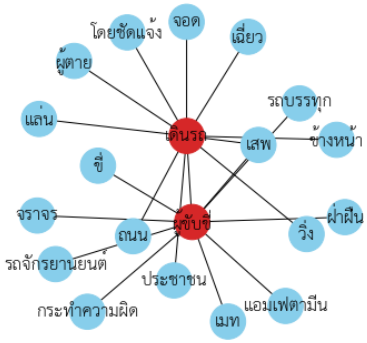
ผู้ใช้จะเลือกไปอย่างต่อเนื่อง โดยระบบกำหนดเงื่อนไขว่า เมื่อโหนดที่เลือกนั้น ไม่มีโหนดใดเชื่อมต่อไปอีกเลย จึงไม่สามารถท่องกราฟต่อไปอีกได้ หรือ หากคำสำคัญที่ได้มาเพียงพอกับการรับแจ้งความแล้ว ในกรณีต่างๆ ที่เกิดขึ้นแล้ว ขึ้นอยู่กับผู้ใช้ หรือดุลยพินิจของพนักงานสอบสวน

สถานการณ์จำลองที่ 2 - นายเอ จอดรถอยู่ริมถนนในเวลากลางคืน หลังจากนั้น นายบี ขับขี่รถยนต์มาชนท้ายรถของนายเอ นายเอได้รับบาดเจ็บสาหัส นำส่งโรงพยาบาล จากนั้น เจ้าหน้าที่นำตามนายบีมาที่สถานีตำรวจเพื่อทำการสอบสวนเบื้องต้น ในกรณีนี้ นายบี เป็นผู้ขับขี่รถยนต์มาชนรถของนายเอ ระบบกำหนดให้โหนดเริ่มแรก คือ ผู้ขับขี่ และเริ่มการท่องกราฟเพื่อสร้างรายงานประจำวันปรากฏรายละเอียดดังตารางที่ 21

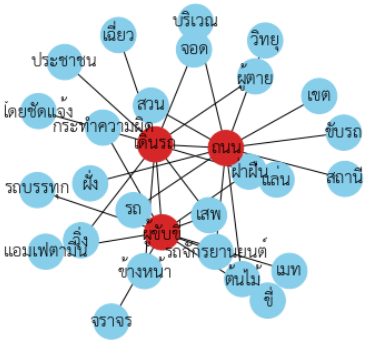
ตารางที่ 21 แสดงรายละเอียดการท่องกราฟเพื่อสร้างรายงานประจำวัน

กราฟ	โหนดที่ได้	คำอธิบาย
		ระบบแสดงโหนดที่เกี่ยวข้องกับ **ผู้ขับขี่** โดยเรียงตามค่าความสำคัญสูงสุด 10 อันดับ โดยอันดับแรกคือคำว่า **จราจร** ซึ่งไม่ได้เกี่ยวข้องกับเหตุการณ์ที่เกิดขึ้น ผู้ใช้

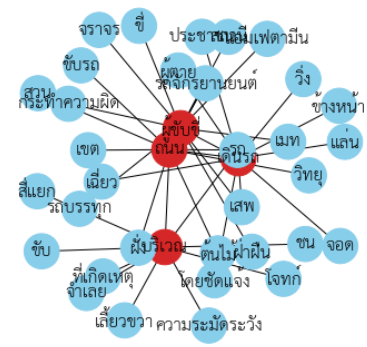
	target	centrality_value
0	จราจร	0.025102
1	รถจักรยานยนต์	0.022497
2	กระทำความผิด	0.019578
3	เดินรถ	0.010612
4	เสพ	0.007079
5	เมท	0.005537
6	แอมเฟตามีน	0.005438
7	รถบรรทุก	0.004274
8	ฝ่าฝืน	0.002823
9	ขี่	0.002123

<table>
<tr>
<td></td>
<td></td>
<td>จึงไม่ได้เลือดโหนดดังกล่าว จากกรณีนี้ผู้ใช้เลือกโหนดลำดับที่ 4 คือ เดินรถ เพื่อถามผู้ขับขี่ว่า เดินรถมาย่างไร</td>
</tr>
<tr>
<td> 
</td>
<td>

	target	centrality_value
0	ถนน	0.044556
1	ผู้ตาย	0.033770
2	เสพ	0.007079
3	แล่น	0.006828
4	จอด	0.006822
5	เฉี่ยว	0.004748
6	ข้างหน้า	0.001727
7	ประชาชน	0.000637
8	โดยชัดแจ้ง	0.000488
9	วิ่ง	0.000420

</td>
<td>ระบบแสดงโหนดที่เกี่ยวข้องกับ เดินรถ ในอันดับแรก คือคำว่า ถนน ซึ่งมีความเกี่ยวข้องกับ เดินรถ ว่า เดินรถมาบนถนนอะไร ผู้ใช้จึงเลือกโหนดดังกล่าว</td>
</tr>
<tr>
<td> 
</td>
<td>

	target	centrality_value
0	รถ	0.101515
1	ขับรถ	0.035019
2	บริเวณ	0.022868
3	รถจักรยานยนต์	0.022497
4	ฝั่ง	0.002636
5	เขต	0.002077
6	สถานี	0.001702
7	สวน	0.001183
8	ต้นไม้	0.001162
9	วิทยุ	0.000525

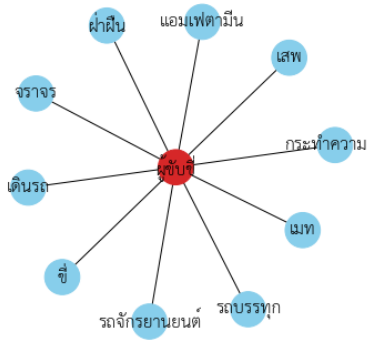
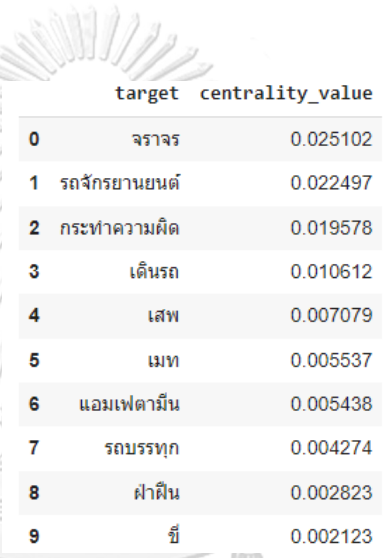
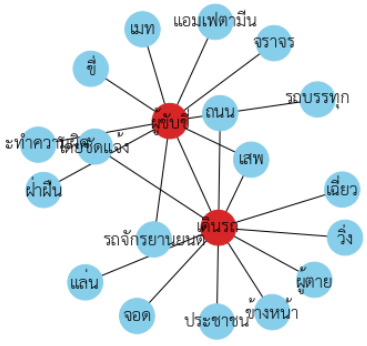
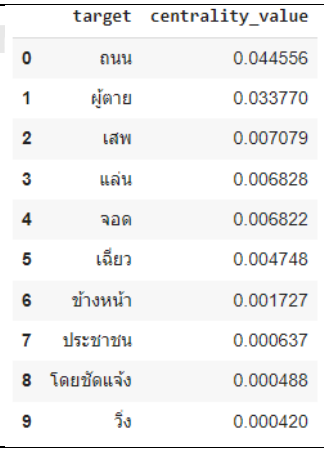
</td>
<td>ระบบแสดงโหนดที่เกี่ยวข้องกับ เดินรถ โดยโหนดที่มีค่าความสำคัญสูงที่สุด คือ รถ ซึ่งผู้ใช้ไม่ได้เลือก แต่ผู้ใช้เลือกโหนดลำดับที่ 3 เพื่อถามว่า ถนน ที่ผู้ขับขี่เดินรถมานั้น อยู่บริเวณใดของถนน</td>
</tr>
<tr>
<td> 
</td>
<td>

	target	centrality_value
0	จำเลย	0.347938
1	โจทก์	0.117508
2	รถ	0.101515
3	ขับ	0.065891
4	ชน	0.015300
5	สีแยก	0.006316
6	ที่เกิดเหตุ	0.005576
7	ความระมัดระวัง	0.004489
8	เลี้ยวขวา	0.000532

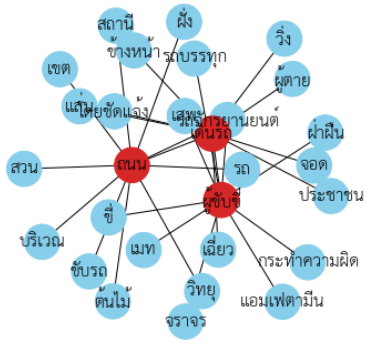
</td>
<td>ระบบแสดงโหนดที่เกี่ยวข้อง กับ บริเวณ ผู้ใช้เลือกโหนดลำดับที่ 5 คือ โหนด ชน และนำไปถามผู้ขับขี่ว่า ชนในลักษณะอย่างไรมีความเสียหายอะไรเกิดขึ้นบ้างนอกเหนือจากคนที่ถูกนำส่ง รพ.</td>
</tr>
</table>

สถานการณ์จำลองที่ 3 - นาย ก. ขับรถมาด้วยความเร็วสูง และชนกับ นาย ข. ที่กำลังข้ามถนนอยู่ ถึงแก่ความตายทันที เจ้าหน้าที่ตำรวจ นำตัวนาย ก. มาที่สถานีตำรวจ เพื่อสอบสวน และตรวจแอลกอฮอล์พบว่า มีปริมาณแอลกอฮอล์ในร่างกายเกินกว่าที่กฎหมายกำหนด ในกรณีนี้ นาย ก. เป็นผู้ขับขี่รถยนต์มาชนรถของนายเอ ระบบกำหนดให้โหนดเริ่มแรก คือ ผู้ขับขี่ และเริ่มการท่องกราฟเพื่อสร้างรายงานประจำวัน รายละเอียดการท่องกราฟ ปรากฏรายละเอียดตามตารางที่ 22

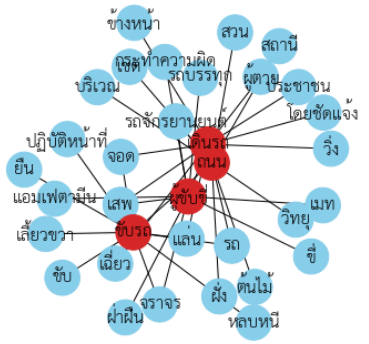
ตารางที่ 22 แสดงรายละเอียดการท่องกราฟเพื่อสร้างรายงานประจำวัน

กราฟ	โหนดที่ได้	คำอธิบาย
		ระบบแสดงโหนดที่เกี่ยวข้องกับ **ผู้ขับขี่** โดยเรียงตามค่าความสำคัญสูงสุด 10 อันดับ โดยอันดับแรกคือคำว่า **จราจร** ซึ่งไม่ได้เกี่ยวข้องกับเหตุการณ์ที่เกิดขึ้น ผู้ใช้จึงไม่ได้เลือดโหนดดังกล่าว จากกรณีนี้ผู้ใช้เลือกโหนดลำดับที่ 4 คือ **เดินรถ** เพื่อถามผู้ขับขี่ว่า **เดินรถ**มาอย่างไร
		ระบบแสดงโหนดที่เกี่ยวข้องกับ เดินรถ ในอันดับแรก คือคำว่า **ถนน** ซึ่งมีความเกี่ยวข้องกับ **เดินรถ** ว่าเดินรถมาบนถนนอะไร ผู้ใช้จึงเลือกโหนดดังกล่าว

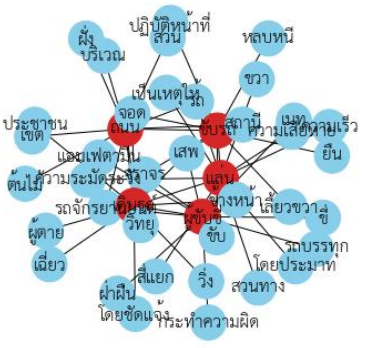
จุฬาลงกรณ์มหาวิทยาลัย

<table>
<tr>
<td> 
</td>
<td>

	target	centrality_value
0	รถ	0.101515
1	ขับรถ	0.035019
2	บริเวณ	0.022868
3	รถจักรยานยนต์	0.022497
4	ฝั่ง	0.002636
5	เขต	0.002077
6	สถานี	0.001702
7	สวน	0.001183
8	ต้นไม้	0.001162
9	วิทยุ	0.000525

</td>
<td>ระบบแสดงโหนดที่เกี่ยวข้องกับ เดินรถ โดยโหนดที่มีค่าความสำคัญสูงที่สุด คือ รถ แต่ในกรณีดังกล่าว ผู้ใช้เลือกโหนดลำดับที่ 2 ขับรถ เพื่อใช้ถามผู้ขับขี่ ว่า ขับรถยี่ห้ออะไร สีอะไร</td>
</tr>
<tr>
<td> 
</td>
<td>

	target	centrality_value
0	รถ	0.101515
1	ขับ	0.065891
2	จราจร	0.025102
3	เสพ	0.007079
4	แล่น	0.006828
5	จอด	0.006822
6	หลบหนี	0.003038
7	ยืน	0.002662
8	ปฏิบัติหน้าที่	0.002271
9	เลี้ยวขวา	0.000532

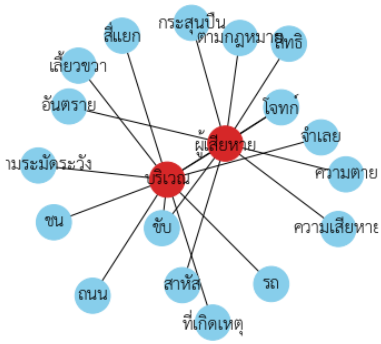
</td>
<td>ผู้ใช้เลือกโหนด แล่น เพื่อสอบถามผู้ขับขี่ว่า ขับรถมาอย่างไร</td>
</tr>
<tr>
<td> 
</td>
<td>

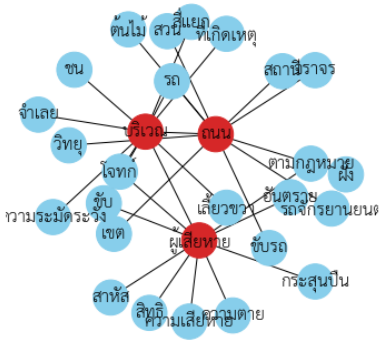
	target	centrality_value
0	รถ	0.101515
1	เป็นเหตุให้	0.024983
2	ความเร็ว	0.009378
3	ความเสียหาย	0.007796
4	สี่แยก	0.006316
5	ความระมัดระวัง	0.004489
6	สวนทาง	0.003549
7	ข้างหน้า	0.001727
8	ขวา	0.001583
9	โดยประมาท	0.000639

</td>
<td>เมื่อเลือกโหนด แล่น ระบบแสดงโหนดที่เกี่ยวข้องมาทั้งหมด 10 อันดับ ผู้ใช้เลือกโหนดลำดับที่ 3 ความเร็ว เพื่อถามผู้ขับขี่ว่า ขณะขับขี่แล่นรถมาด้วยความเร็วเท่าไหร่</td>
</tr>
</table>

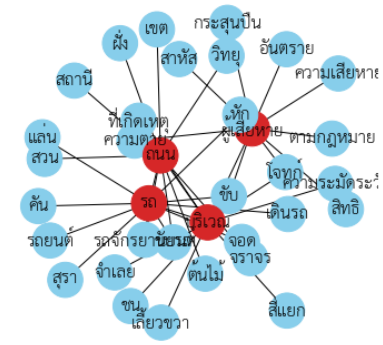
สถานการณ์จำลองที่ 4 - นายสมชาย กำลังขับรถกลับบ้าน ขณะที่กำลังเลี้ยวเข้าซอยบ้าน พบรถยนต์ยี่ห้อโตโยต้า สีขาว จอดขวางอยู่บริเวณหน้าปากซอย ทำให้ไม่สามารถขับรถเข้าซอยได้ จึงเดินทางไปที่สถานีตำรวจ เพื่อแจ้งความกรณี มีรถยนต์จอดรถขวางทางสาธารณะ ทำให้ตนเอง และประชาชนคนอื่น ได้รับความเดือดร้อน จากสถานการณ์จำลองนี้ กำหนดให้นายสมชาย เป็นผู้เสียหาย โดยการท่องกราฟจะเริ่มที่โหนดผู้เสียหาย มีปรากฏรายละเอียดดังตารางที่ 23

ตารางที่ 23 แสดงรายละเอียดการท่องกราฟเพื่อสร้างรายงานประจำวัน

<table>
<tr><th>กราฟ</th><th>โหนดที่ได้</th><th>คำอธิบาย</th></tr>
<tr>
<td>
</td>
<td>
<table>
<tr><th></th><th>target</th><th>centrality_value</th></tr>
<tr><td>0</td><td>โจทก์</td><td>0.117508</td></tr>
<tr><td>1</td><td>ขับ</td><td>0.065891</td></tr>
<tr><td>2</td><td>บริเวณ</td><td>0.022868</td></tr>
<tr><td>3</td><td>ความเสียหาย</td><td>0.007796</td></tr>
<tr><td>4</td><td>อันตราย</td><td>0.005471</td></tr>
<tr><td>5</td><td>กระสุนปืน</td><td>0.004973</td></tr>
<tr><td>6</td><td>ตามกฎหมาย</td><td>0.004299</td></tr>
<tr><td>7</td><td>สิทธิ</td><td>0.004169</td></tr>
<tr><td>8</td><td>สาหัส</td><td>0.003943</td></tr>
<tr><td>9</td><td>ความตาย</td><td>0.003383</td></tr>
</table>
</td>
<td>ระบบแสดงโหนด ผู้เสียหาย และโหนดที่มีความเกี่ยวข้องทั้งหมด ผู้ใช้เลือกโหนด บริเวณ เพื่อสอบถามผู้เสียหายว่าเหตุการณ์ที่เกิดขึ้น อยู่บริเวณใด หรือสถานที่ใด</td>
</tr>
<tr>
<td> 
</td>
<td>
<table>
<tr><th></th><th>target</th><th>centrality_value</th></tr>
<tr><td>0</td><td>จำเลย</td><td>0.347938</td></tr>
<tr><td>1</td><td>โจทก์</td><td>0.117508</td></tr>
<tr><td>2</td><td>รถ</td><td>0.101515</td></tr>
<tr><td>3</td><td>ขับ</td><td>0.065891</td></tr>
<tr><td>4</td><td>ถนน</td><td>0.044556</td></tr>
<tr><td>5</td><td>ชน</td><td>0.015300</td></tr>
<tr><td>6</td><td>สี่แยก</td><td>0.006316</td></tr>
<tr><td>7</td><td>ที่เกิดเหตุ</td><td>0.005576</td></tr>
<tr><td>8</td><td>ความระมัดระวัง</td><td>0.004489</td></tr>
<tr><td>9</td><td>เลี้ยวขวา</td><td>0.000532</td></tr>
</table>
</td>
<td>ผู้ใช้เลือกโหนดลำดับที่ 5 ถนน เพื่อซักรายละเอียดเพิ่มเติมจาก โหนด บริเวณ ว่าอยู่ที่ถนนใด</td>
</tr>
</table>

<table>
<tr>
<td> 
</td>
<td>

	target	centrality_value
0	รถ	0.101515
1	ขับรถ	0.035019
2	จราจร	0.025102
3	รถจักรยานยนต์	0.022497
4	ผัง	0.002636
5	เขต	0.002077
6	สถานี	0.001702
7	สวน	0.001183
8	ต้นไม้	0.001162
9	วิทยุ	0.000525

</td>
<td>โหนดที่มีค่าความสำคัญสูงที่สุดคือโหนด รถ ผู้ใช้เลือกโหนดดังกล่าว เพื่อสอบถามรายละเอียดเกี่ยวกับรถยนต์ในเหตุการณ์ ว่ามียี่ห้ออะไร สีอะไร รุ่นอะไร และหมายเลขทะเบียนอะไร</td>
</tr>
<tr>
<td> 
</td>
<td>

	target	centrality_value
0	ขับ	0.065891
1	รถยนต์	0.057557
2	ขับรถ	0.035019
3	รถจักรยานยนต์	0.022497
4	คัน	0.011033
5	เดินรถ	0.010612
6	แล่น	0.006828
7	จอด	0.006822
8	หัก	0.002377
9	สุรา	0.000679

</td>
<td>ผู้ใช้เลือกโหนด จอด เพื่อถามผู้เสียหาย ว่า รถที่จอดกีดขวางทางจราจร จอดอย่างไร บนขาวแดง ใช่หรือไม่ เป็นต้น</td>
</tr>
</table>

4.3.1 สรุปผลลัพธ์

การท่องกราฟจากแต่ละโหนด ผู้ใช้จะนำคำจากโหนดที่เลือกไปใช้เป็นข้อคำถามสำหรับซักถามหรือสอบสวนผู้เสียหาย เมื่อได้ข้อความหรือประโยคที่ได้มาจากการซักถามทั้งหมดจะถูกรวบรวมเป็นแบบรายงานประจำวัน ประกอบด้วย รายละเอียดผู้เสียหาย พฤติการณ์แห่งคดี เป็นต้น ทั้งนี้กำหนดให้สร้างรายงานประจำจากสถานการณ์จำลองที่ 4 รายละเอียดปรากฏดังตารางที่ 24

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ตารางที่ 24 แสดงรายการสรุปรายละเอียดโหนดที่ผู้ใช้เลือกทั้งหมด

ลำดับ	โหนดที่เลือก	ประโยค
(1)	**ผู้เสียหาย**	ชื่อ นายสมชาย ใจดี อายุ 50 ปี อยู่บ้านเลขที่ 999 ซอย พหลโยธิน 14 แขวงสามเสนใน เขตพญาไท กทม. หมายเลขโทรศัพท์ 09 099 0990
(2)	**บริเวณ**	จุดเกิดเหตุ บริเวณหน้าคอนโดพหลโยธินพาร์ค
(3)	**ถนน**	ถนน พหลโยธิน ซอย พหลโยธิน 14
(4)	**รถ**	รถยนต์ยี่ห้อโตโยต้า สีขาว รุ่น ฟอร์จูนเนอร์ หมายเลขทะเบียน 1 กก 1111 กทม.
(5)	**จอด**	ขณะนายสมชายกำลังขับรถกลับบ้าน และกำลังจะเลี้ยวเข้าคอนโดของตน พบรถคันดังกล่าว จอดอยู่บริเวณริมฟุตบาททางเข้าคอนโดของตน โดยไม่ได้ใส่เกียร์ว่างไว้ ทำให้ไม่สามารถเข็นรถยนต์ให้พ้นจากทางเข้าคอนโดไปได้

จากตารางที่ 24 แสดงข้อความ ประโยคที่ได้จากการซักถามระหว่างผู้เสียหายทั้งหมด ขั้นตอนต่อไป ผู้วิจัยก็จะนำข้อความที่ได้ นำมารวบรวมให้เป็นรายงานประจำวัน ทั้งนี้ผู้ใช้สามารถกำหนดรายละเอียดข้อความที่สร้างขึ้น และหรือสามารถจัดกลุ่มประเภทของรายละเอียดได้ ขึ้นอยู่กับดุลยพินิจของผู้ใช้ ทั้งนี้ ผู้วิจัยสรุปรายละเอียดที่ได้จากการซักถามระหว่างพนักงานสอบสวน และผู้เสียหาย ปรากฏตัวอย่างดังภาพที่ 62

(2) บริเวณหน้าคอนโดพหลโยธินพาร์ค **(3)** ซอย พหลโยธิน 14 แขวงสามเสนใน เขตพญาไท กทม.

(1) นายสมชาย ใจดี อายุ 50 ปี อยู่บ้านเลขที่ 999 ซอย พหลโยธิน 14 แขวงสามเสนใน เขตพญาไท กทม. หมายเลขโทรศัพท์ 09 099 0990

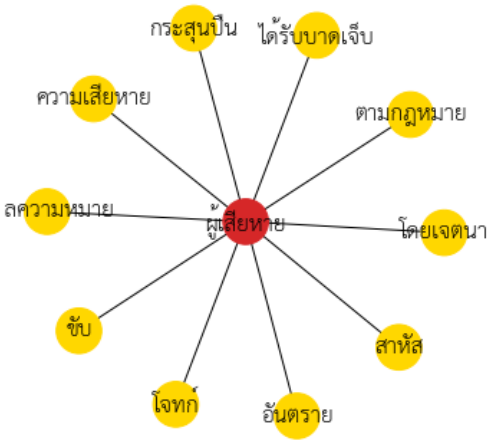
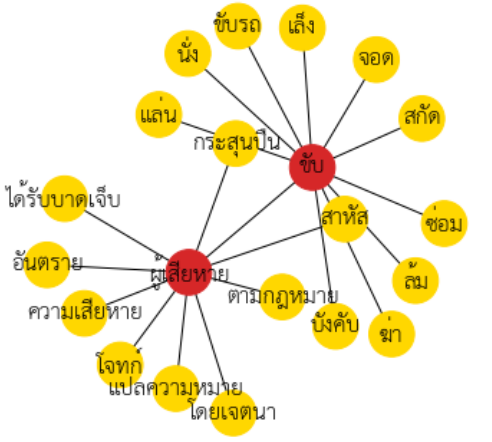
(5) ขณะนายสมชายกำลังขับรถกลับบ้าน และกำลังจะเลี้ยวเข้าคอนโดของตน พบรถ **(4)**รถยนต์ยี่ห้อโตโยต้า สีขาว รุ่น ฟอร์จูนเนอร์ หมายเลขทะเบียน 1 กก 1111 กทม. คันดังกล่าว จอดอยู่บริเวณริมฟุตบาททางเข้าคอนโดของตน โดยไม่ได้ใส่เกียร์ว่างไว้ ทำให้ไม่สามารถเข็นรถยนต์ให้พ้นจากทางเข้าคอนโดไปได้

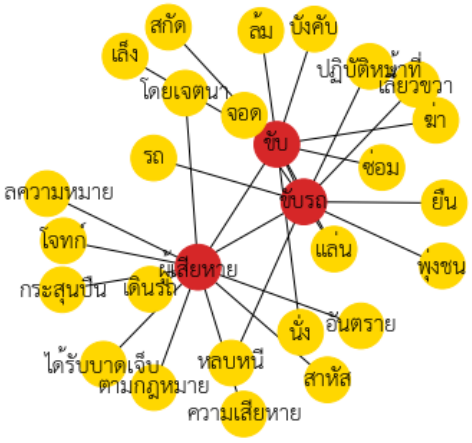
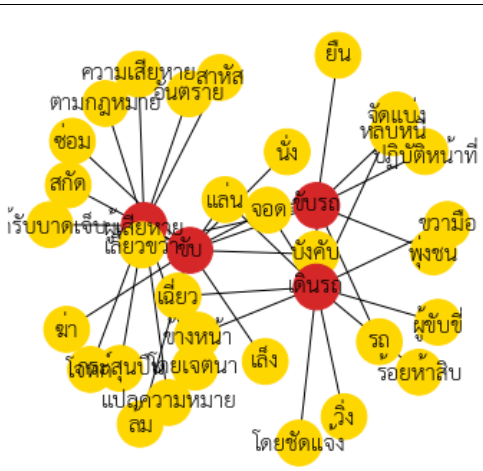
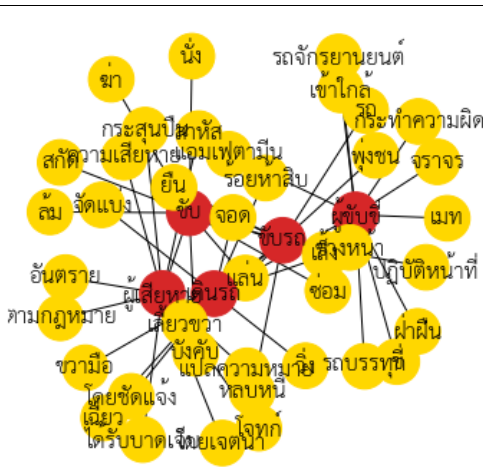
ภาพที่ 62 รายงานประจำวันที่ได้จากกระบวนการสกัดกิจกรรมสำคัญ

ส่วนต่อไป จะทดสอบท่องกราฟจากกราฟที่ถูกสรุปด้วย วิธีทางกราฟแบบที่เหลือ มีรายละเอียดดังตารางที่ 25 - 28

2) วิธีค่ากลางความเป็นศูนย์กลาง

ตารางที่ 25 ตารางแสดงผลลัพธ์การท่องกราฟ

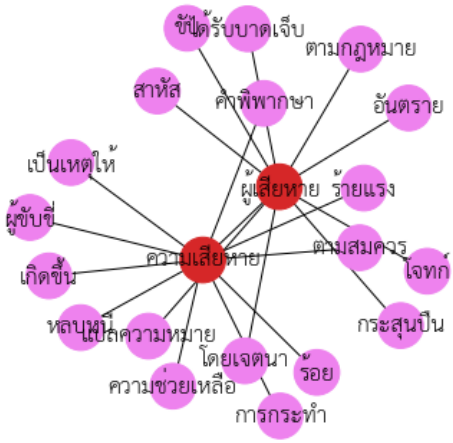
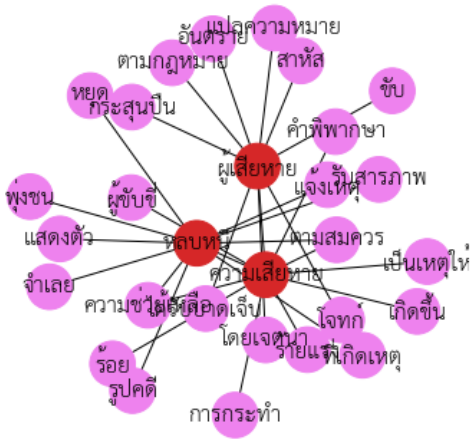
กราฟ	โหนดที่ผู้ใช้เลือก	โหนดที่เกี่ยวข้อง
	ผู้เสียหาย	โดยเจตนา แปลความหมาย โจทก์ กระสุนปืน สาหัส ตามกฎหมาย ความเสียหาย ได้รับบาดเจ็บ อันตราย ขับ
	ขับ	ล้ม ขับรถ บังคับ สกัด ฆ่า แล่น ซ่อม จอด นั่ง เล็ง

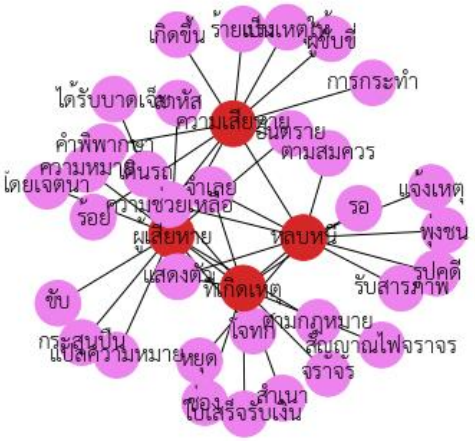
	ขับรถ	แล่น เดินรถ หลบหนี พุ่งชน เลี้ยวขวา รถ ขับ ยืน ปฏิบัติหน้าที่ จอด
	เดินรถ	จอด แล่น วิ่ง จัดแบ่ง ขวามือ เฉี่ยว ผู้ขับขี่ ข้างหน้า ร้อยห้าสิบ โดยชัดแจ้ง
	ผู้ขับขี่	แอมเฟตามีน จราจร ฝ่าฝืน รถจักรยานยนต์ เข้าใกล้ กระทำความผิด ขี่ เมท รถบรรทุก เดินรถ

มหาวิทยาลัย
RN UNIVERSITY

3) วิธีค่ากลางความใกล้

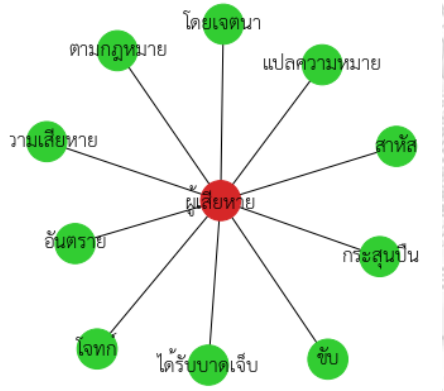
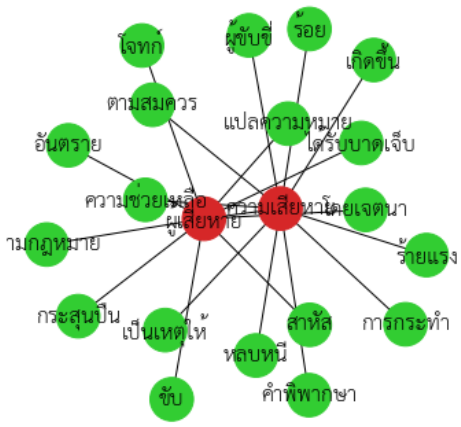
ตารางที่ 26 ตารางแสดงผลลัพธ์การท่องกราฟ

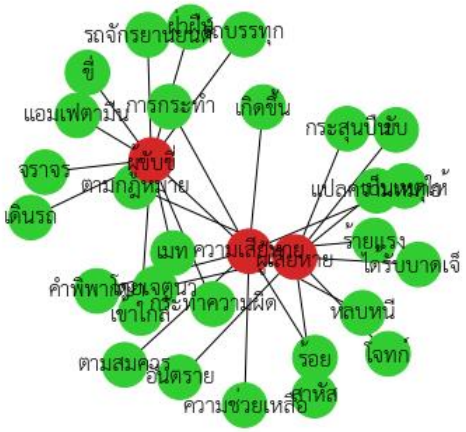
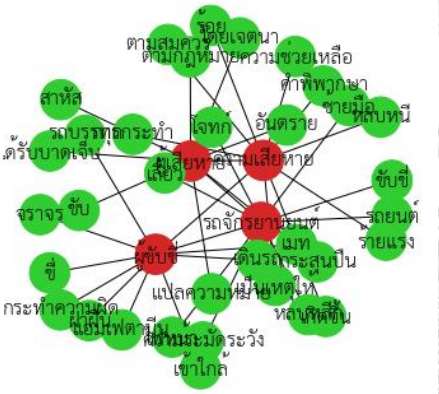
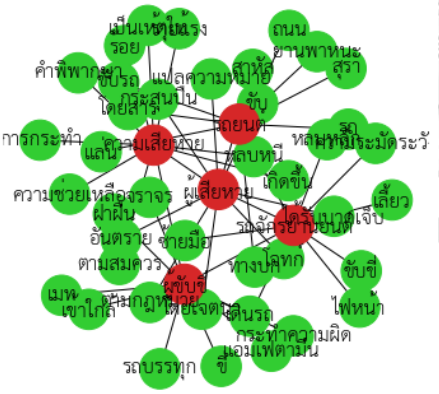
กราฟ	โหนดที่ผู้ใช้เลือก	โหนดที่เกี่ยวข้อง
	ผู้เสียหาย	โดยเจตนา แปลความหมาย โจทก์ กระสุนปืน สาหัส ตามกฎหมาย ความเสียหาย ได้รับบาดเจ็บ อันตราย ขับ
	ความเสียหาย	การกระทำ คำพิพากษาศาลฎีกา ความช่วยเหลือ ร้ายแรง เป็นเหตุให้ หลบหนี ตามสมควร ร้อย เกิดขึ้น ผู้ขับขี่
	หลบหนี	พุ่งชน ที่เกิดเหตุ แสดงตัว หยุด รับสารภาพ ความช่วยเหลือ จำเลย ตามสมควร แจ้งเหตุ รูปคดี

	ที่เกิดเหตุ	แสดงตัว จราจร จำเลย สำเนา สัญญาณไฟ จราจร เดินรถ ใบเสร็จรับเงิน รอ ช่อง ความหมาย

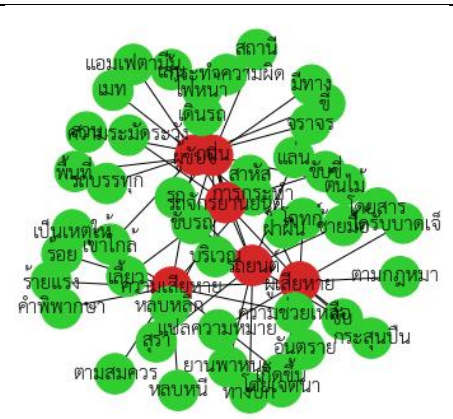
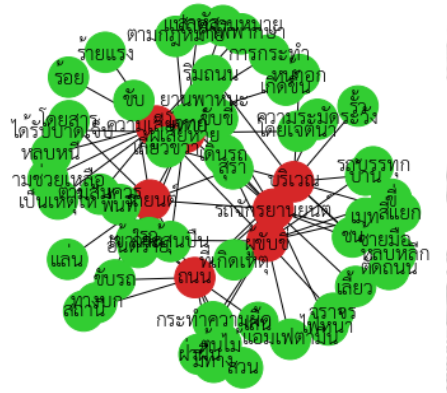
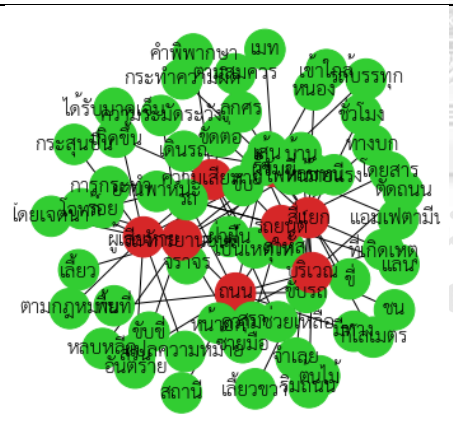
4) วิธีค่ากลางไอเกนเวกเตอร์

ตารางที่ 27 ตารางแสดงผลลัพธ์การท่องกราฟ

กราฟ	โหนดที่ผู้ใช้เลือก	โหนดที่เกี่ยวข้อง
	ผู้เสียหาย	โดยเจตนา แปลความหมาย โจทก์ กระสุนปืน สาหัส ตามกฎหมาย ความเสียหาย ได้รับบาดเจ็บ อันตราย ขับ
	ความเสียหาย	การกระทำ คำพิพากษา ศาลฎีกา ความช่วยเหลือ ร้ายแรง เป็นเหตุให้ หลบหนี ตามสมควร ร้อย เกิดขึ้น ผู้ขับขี่

	ผู้ขับขี่	แอมเฟตามีน จราจร ฝ่าฝืน รถจักรยานยนต์ เข้าใกล้ กระทำความผิด ขี่ เมท รถบรรทุก เดินรถ
	รถจักรยานยนต์	รถยนต์ ไฟหน้า ผู้เสียหาย ความระมัดระวัง หลบหลีก เดินรถ โจทก์ ขับขี่ ซ้ายมือ เลี้ยว
	รถยนต์	รถ รถจักรยานยนต์ ยานพาหนะ ขับรถ ทางบก โดยสาร ขับ ถนน สุรา แล่น

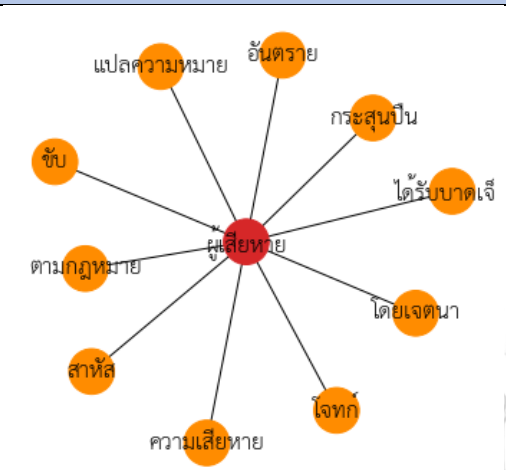
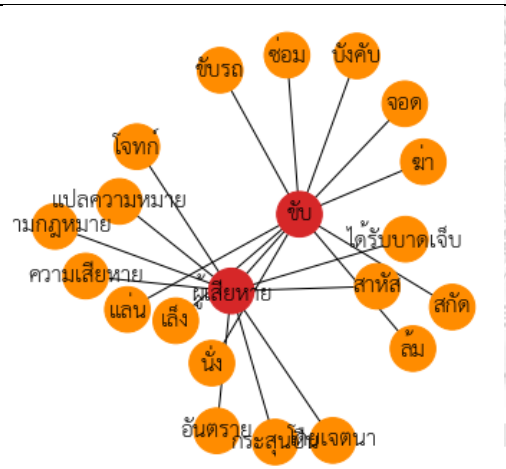
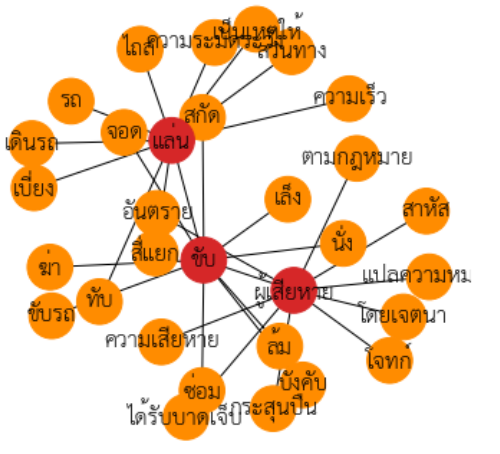
จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

	ถนน	เส้น รถ มีทาง ขับรถ สวน บริเวณ สถานี รถจักรยานยนต์ ต้นไม้ พื้นที่
	บริเวณ	ริมถนน สี่แยก ชน หน้าอก ที่เกิดเหตุ บ้าน ถนน รั้ว ตัดถนน เลี้ยว ขวา
	สี่แยก	กิโลเมตร ที่เกิดเหตุ หนอง ขับรถ ฝ่าฝืน ลูกศร จำเลย ขัดต่อ เป็นเหตุให้ ชั่วโมง

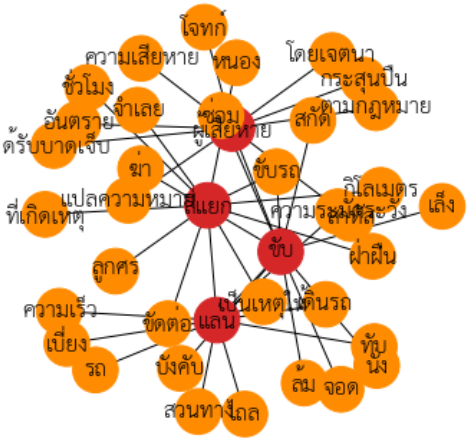
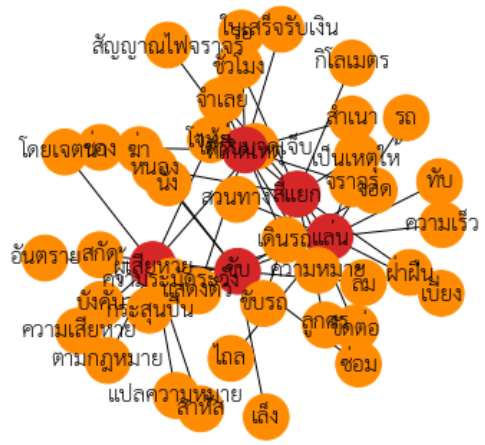
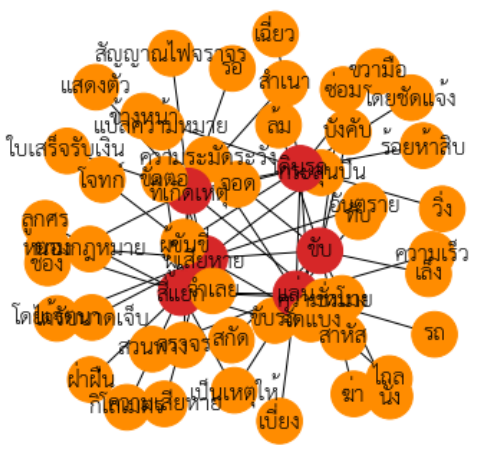
จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

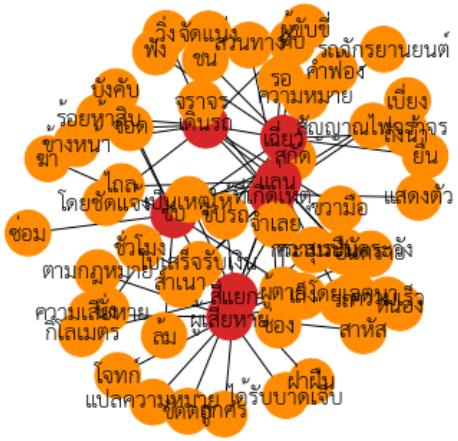
5) วิธีค่ากลางเพจแรงก์

ตารางที่ 28 ตารางแสดงผลลัพธ์การท่องกราฟ

กราฟ	โหนดที่ผู้ใช้เลือก	โหนดที่เกี่ยวข้อง
	ผู้เสียหาย	โดยเจตนา แปลความหมาย โจทก์ กระสุนปืน สาหัส ตามกฎหมาย ความเสียหาย ได้รับบาดเจ็บ อันตราย ขับ
	ขับ	ล้ม ขับรถ บังคับ สกัด ฆ่า แล่น ซ่อม จอด นั่ง เล็ง
	แล่น	เดินรถ เบี่ยง สวนทาง ความเร็ว รถ ความระมัดระวัง เป็นเหตุให้ สี่แยก ไถล ทับ

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

	สี่แยก	กิโลเมตร ที่เกิดเหตุ หนอง ขับรถ ฝ่าฝืน ลูกศร จำเลย ขัดต่อ เป็นเหตุให้ ชั่วโมง
	ที่เกิดเหตุ	แสดงตัว จราจร จำเลย สำเนา สัญญาณไฟ จราจร เดินรถ ใบเสร็จรับเงิน รอ ช่อง ความหมาย
	เดินรถ	จอด แล่น วิ่ง จัดแบ่ง ขวามือ เฉี่ยว ผู้ขับขี่ ข้างหน้า ร้อยห้าสิบ โดย ชัดแจ้ง

	เฉี่ยว	ถังน้ำ คำฟ้อง รถจักรยานยนต์ ผู้ตาย ขับรถ ยื่น ขับ จำเลย ฟัง ชน

เมื่อผู้วิจัยได้ขั้นตอนสำหรับการสร้างแม่แบบรายงานประจำวัน ผู้วิจัยได้นำวิธีการสร้างแม่แบบรายงานประจำวัน ดังกล่าว มาใช้กับการรับแจ้งความร้องทุกข์ในกรณีอื่น ๆ สำหรับการสร้างผลลัพธ์เพิ่มเติม เพื่อทดสอบว่า ขั้นตอนวิธีที่ผู้วิจัยได้สร้างขึ้นมา สามารถนำไปใช้กับคำพิพากษาศาลฎีกา ในกรณีอื่นๆ ได้หรือไม่ โดยผู้วิจัยเลือกขอบเขตการรับแจ้งความกรณีถูกชิงทรัพย์ เพื่อทดสอบว่าวิธีดังกล่าว สามารถหรือมีความเหมาะสมกับการใช้สร้างแม่แบบรายงานประจำวันในกรณีอื่นนอกจากคดีจราจรหรือไม่ โดยอันดับแรก ผู้วิจัยได้สร้างกราฟของคำพิพากษาศาลฎีกาโดยเลือกเฉพาะที่เป็นคำพิพากษาศาลฎีกาเกี่ยวกับการชิงทรัพย์ (ป.อาญา มาตรา 339) และการลักทรัพย์ (ป.อาญา มาตรา 335) เมื่อได้กราฟของคำพิพากษาศาลฎีกาทั้งหมดแล้ว ก็ดำเนินการสร้างแม่แบบรายงานประจำวันโดยการท่องกราฟในลักษณะเดียวกันกับคดีจราจร โดยกำหนดสถานการณ์จำลองมีรายละเอียด ดังนี้

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

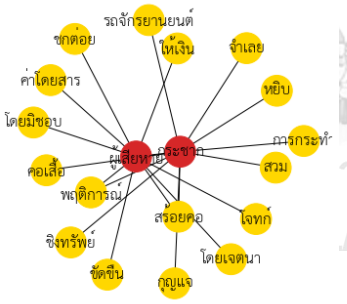
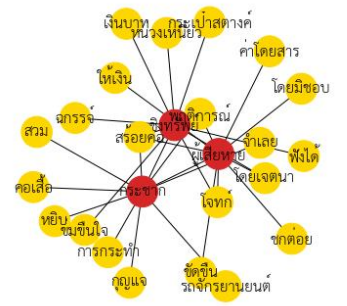
4.3.2 การสร้างผลลัพธ์จากการสกัดกิจกรรมของคำพิพากษาศาลฎีกาเพิ่มเติม เฉพาะคดีอาญาความผิดฐานชิงทรัพย์

ส่วนนี้จะกำหนดสถานการณ์สำหรับการรับแจ้งความร้องทุกข์เพื่อลงบันทึกประจำวัน

สถานการณ์จำลองที่ 5 - น.ส.เอ กำลังเดินอยู่ริมถนน ถูกชายวัยกลางคน กระชากกระเป๋าถือแล้ววิ่งหนีไป

ในกรณีสถานการณ์จำลองที่ 5 น.ส.เอ ถูกชายวันกลางคน กระชากกระเป๋าถือ จึงเดินทางมาที่สถานีตำรวจเพื่อเข้าแจ้งความ จากสถานการณ์นี้ น.ส.เอ เป็นผู้เสียหาย เนื่องจากถูกกระชากกระเป๋า ดังนั้น ระบบจึงเลือกโหนด **ผู้เสียหาย** เป็นโหนดแรกสำหรับการท่องกราฟ โดยผู้วิจัยทดสอบขั้นตอนวิธีดังกล่าวกับกราฟที่ถูกสรุปด้วย Degree Centrality มีรายละเอียดดังตารางที่ 29

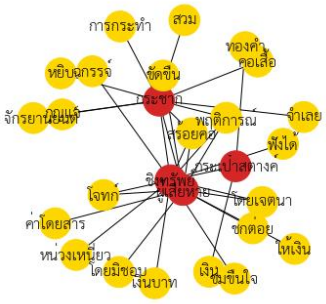
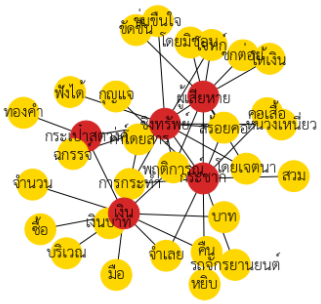
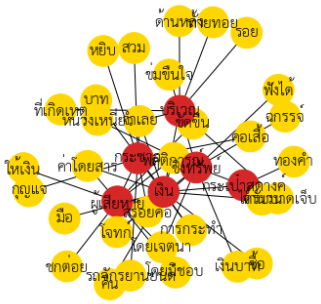
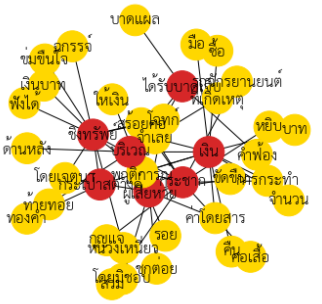
ตารางที่ 29 แสดงรายละเอียดการท่องกราฟเพื่อสร้างรายงานประจำวัน

กราฟ	โหนดที่ได้	คำอธิบาย
		เมื่อเลือกโหนด **ผู้เสียหาย** แล้ว ระบบจะแสดงโหนดที่เกี่ยวข้องมากที่สุด โหนดที่ระบบแสดงขึ้นมาเป็นลำดับแรกคือ โจทก์ ซึ่งมีค่าความสำคัญมากที่สุด แต่จากการสอบสวนเบื้องต้น ผู้เสียหายแจ้งว่าตนถูกกระชากกระเป๋า ผู้ใช้จึงเลือกโหนดลำดับที่ 4 คือ **กระชาก** เพื่อเป็นการถาม น.ส.เอ ว่า ถูกกระชากกระเป๋าในลักษณะอย่างไร
		เลือกโหนด **กระชาก** แล้วระบบก็จะแสดงโหนดที่เกี่ยวข้อง จากกรณีดังกล่าว ผู้ใช้ (ตำรวจ) พิจารณาเบื้องต้นแล้ว เห็นว่าเป็นเหตุชิงทรัพย์ ผู้ใช้จึงเลือกโหนดที่เกี่ยวข้องลำดับที่ 2 คือ **ชิงทรัพย์** ระบบก็จะแสดงโหนดที่เกี่ยวข้องขึ้นมา
		เมื่อระบบแสดงโหนดที่เกี่ยวข้องขึ้นมา ผู้ใช้เลือกโหนดลำดับที่ 6 **กระเป๋าสตางค์** เพื่อถามผู้เสียหายว่า **กระเป๋าสตางค์** ที่ถูกชิงทรัพย์ไป มีของมีค่าอะไรบ้าง

	target	centrality_value
0	โจทก์	0.104348
1	พฤติการณ์	0.032411
2	สร้อยคอ	0.024506
3	กระชาก	0.018972
4	ขัดขืน	0.010277
5	โดยเจตนา	0.007115
6	ค่าโดยสาร	0.007115
7	ให้เงิน	0.005534
8	ชกต่อย	0.004743
9	โดยมิชอบ	0.003162

	target	centrality_value
0	จำเลย	0.327273
1	ชิงทรัพย์	0.103557
2	รถจักรยานยนต์	0.086166
3	การกระทำ	0.064032
4	พฤติการณ์	0.032411
5	สร้อยคอ	0.024506
6	หยิบ	0.021344
7	สวม	0.008696
8	กุญแจ	0.005534
9	คอเสื้อ	0.003953

	target	centrality_value
0	โจทก์	0.104348
1	พฤติการณ์	0.032411
2	สร้อยคอ	0.024506
3	เงินบาท	0.015020
4	ข่มขืนใจ	0.013439
5	กระเป๋าสตางค์	0.007905
6	โดยเจตนา	0.007115
7	ฟังได้	0.007115
8	หน่วงเหนี่ยว	0.003162
9	ฉกรรจ์	0.003162

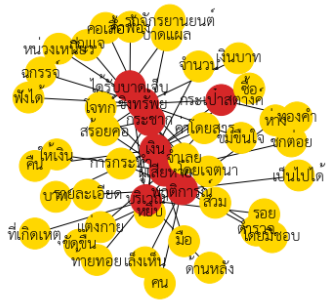
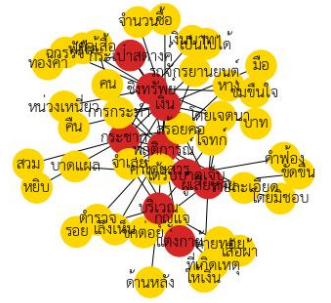
		ระบบแสดงโหนดที่เกี่ยวข้องกับ **กระเป๋าสตางค์** ประกอบด้วย **เงิน** และ **ทองคำ** ผู้ใช้เลือก **เงิน** ซึ่งมีค่าความสำคัญสูงที่สุด เพื่อถามว่ามี**เงิน**อยู่ใน**กระเป๋าสตางค์**เท่าไหร่
		ผู้ใช้เลือกโหนด **บริเวณ** ซึ่งมีค่าความสำคัญเป็นอยู่ในลำดับที่ 4 และถามผู้เสียหายว่า**บริเวณ**ถูก**ชิงทรัพย์**คือที่ใด
		ผู้ใช้เลือกโหนดลำดับที่ 5 **ได้รับบาดเจ็บ** เพื่อถามผู้เสียหายว่า **ได้รับบาดเจ็บ**ในส่วนใดบ้างหรือไม่ เพื่อจักได้ลงบันทึกรายละเอียดในรายงานประจำวันให้ชัดเจน
		เมื่อเลือกโหนด **ได้รับบาดเจ็บ** ระบบไม่ได้แสดงโหนดที่ผู้ใช้ต้องการ จึงย้อนกลับไปที่โหนด **บริเวณ** และเลือกโหนดลำดับที่ 3 **พฤติการณ์** เพื่ออยากทราบรายละเอียดเพิ่มเติมจากเหตุการณ์ที่เกิดขึ้น

	target	centrality_value
0	เงิน	0.040316
1	ทองคำ	0.028458

	target	centrality_value
0	จำเลย	0.327273
1	การกระทำ	0.064032
2	พฤติการณ์	0.032411
3	บริเวณ	0.020553
4	คืน	0.018182
5	ซื้อ	0.015020
6	บาท	0.013439
7	จำนวน	0.009486
8	ค่าโดยสาร	0.007115
9	มือ	0.005534

	target	centrality_value
0	จำเลย	0.327273
1	ที่เกิดเหตุ	0.026087
2	ด้านหลัง	0.012648
3	รอย	0.011067
4	ได้รับบาดเจ็บ	0.006324
5	ท้ายทอย	0.003953

	target	centrality_value
0	โจทก์	0.104348
1	บาดแผล	0.021344
2	คำฟ้อง	0.012648

		ระบบแสดงโหนดที่เกี่ยวข้องผู้ใช้เลือกลำดับที่ 7 **แต่งกาย** เพื่อถามหาลักษณะรูปพรรณของคนร้าย
		ระบบแสดงโหนดที่เกี่ยวข้องเพียงโหนดเดียวคือ **เสื้อผ้า** ซึ่งมีความเกี่ยวข้องกับเครื่องแต่งกาย ซึ่งสามารถให้ผู้เสียหายระบุได้ว่า คนร้ายแต่งกายด้วยเสื้อผ้าลักษณะแบบใด

	target	centrality_value
0	จำเลย	0.327273
1	การกระทำ	0.064032
2	คน	0.040316
3	ตำรวจ	0.026087
4	ห่าง	0.007115
5	เล็งเห็น	0.006324
6	แต่งกาย	0.005534
7	รายละเอียด	0.004743
8	เป็นไปได้	0.003953

	target	centrality_value
0	เสื้อผ้า	0.006324

โดยผู้ใช้จะเลือกคำสำคัญที่ระบบแสดงขึ้นมา เพื่อเป็นหัวข้อคำถาม สำหรับซักถามระหว่างผู้ใช้(ตำรวจ) กับ ผู้เสียหาย (น.ส.เอ) โดยระบบจะแสดงผลคำที่เกี่ยวข้องในลักษณะนี้ไปอย่างต่อเนื่อง เมื่อระบบทำงานเสร็จสิ้นหรือหยุดการทำงานแล้วนั้น ส่วนต่อไป ผู้วิจัยจะสรุปรายละเอียดของคำสำคัญที่ได้จากกราฟ และสร้างประโยคขึ้นมาเป็นตัวอย่าง ดังตารางที่ 30

ตารางที่ 30 แสดงรายการสรุปรายละเอียดโหนดที่ผู้ใช้เลือกทั้งหมด

ลำดับ	โหนดที่เลือก	ประโยค
1	**ผู้เสียหาย**	น.ส.เอ (นามสมมติ) อายุ 28 ปี อาศัยอยู่บ้านเลขที่ 11/1 ถนนอยู่บ้านเลขที่ 999 ซอยพหลโยธิน 14 แขวงสามเสนใน เขตพญาไท กทม. หมายเลขโทรศัพท์ 09 099 0990
2	**ชิงทรัพย์**	1. ลักทรัพย์โดยใช้กาลังประทุษร้าย หรือขู่เข็ญวาในทันใดนั้น จะใช้กาลังประทุษร้าย เพื อให้ความสะดวกแก่ การลักทรัพย์หรือการพาทรัพย์นั้นไป 2. ชิงทรัพย์เป็นเหตุให้ได้รับอันตรายแก่กายหรือจิตใจ
3	**กระชาก**	ถูกชายวัยกลางคน กระชากกระเป๋าถือ

4	**กระเป๋าสตางค์**	ในกระเป๋าถือ มีกระเป๋าสตางค์ และโทรศัพท์มือถือ ไอโฟน 14 มูลค่า 30,000 บาท พร้อมทั้ง เอกสารสำคัญต่าง ๆ
5	**เงิน**	ในกระเป๋าสตางค์มีเงินสดจำนวน 20,000 บาท
6	**บริเวณ**	สี่แยกราชประสงค์ ฝั่งถนนราชดำริ มุ่งหน้าแยกประตูน้ำ เขตราชเทวี กทม.
7	**ได้รับบาดเจ็บ**	ได้รับบาดเจ็บบริเวณแขนข้างขวา มีบาดแผลถูกสายกระเป๋าเสียดสี เลือดออก
8	**พฤติการณ์**	น.ส.เอ กำลังเดินทางกลับบ้าน จากบริเวณหน้าห้างเซ็นทรัลเวิร์ล ฝั่งถนนราชดำริ ระหว่างนั้น มีชายวัยกลางคน วิ่งเข้ามาจากทางด้านหลัง และกระชากกระเป๋าถือของตนแล้ววิ่งหนีไปทางแยกประตูน้ำ
9	**แต่งกาย**	ชายวัยกลางคน สวมเสื้อยืดสีขาว กางเกงยีนส์ สวมหน้ากากอนามัยสีขาว

เมื่อผู้ใช้ได้รายละเอียดจากผู้เสียหายแล้ว ผู้วิจัยได้รวบรวมรายละเอียดตามตารางที่ 30 เพื่อจำลองเป็นรายงานประจำวันได้ตัวอย่าง ปรากฏดังภาพที่ 63

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

(2) 1. ลักทรัพย์โดยใช้กาลังประทุษร้าย หรือขู่เข็ญวาในทันใดนั้น จะใช้กำลังประทุษร้าย เพื่อให้ความสะดวกแก่ การลักทรัพย์หรือการพาทรัพย์นั้นไป

2. ชิงทรัพย์เป็นเหตุให้ได้รับอันตรายแก่กายหรือจิตใจ

(6) สี่แยกราชประสงค์ ฝั่งถนนราชดำริ มุ่งหน้าแยกประตูน้ำ เขตราชเทวี กทม.

(1) น.ส.เอ (นามสมมติ) อายุ 28 ปี อาศัยอยู่บ้านเลขที่ 11/1 ถนนอยู่บ้านเลขที่ 999 ซอย พหลโยธิน 14 แขวงสามเสนใน เขตพญาไท กทม. หมายเลขโทรศัพท์ 09 099 0990

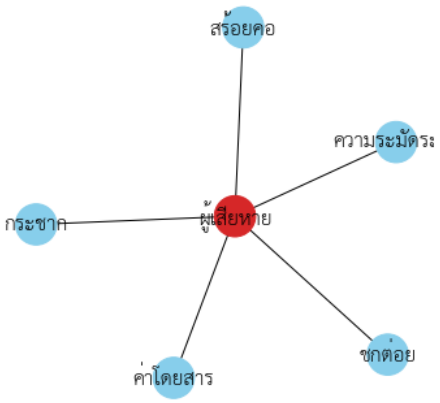
(8) น.ส.เอ กำลังเดินทางกลับบ้าน จากบริเวณหน้าห้างเซ็นทรัลเวิร์ล ฝั่งถนนราชดำริ ระหว่างนั้น มี **(9)** ชายวัยกลางคน สวมเสื้อยืดสีขาว กางเกงยีนส์ สวมหน้ากากอนามัยสีขาว วิ่งเข้ามาจากทางด้านหลัง และ **(3)** กระชากกระเป๋าถือ **(8)** ในกระเป๋าถือ มีกระเป๋าสตางค์ **(5)** ในกระเป๋าสตางค์มีเงินสดจำนวน 20,000 บาทและ โทรศัพท์มือถือ ไอโฟน 14 มูลค่า 30,000 บาท พร้อมทั้ง เอกสารสำคัญต่าง ๆ ของตนแล้ววิ่งหนีไปทางแยกประตูน้ำ **(7)** ได้รับบาดเจ็บบริเวณแขนข้างขวา มีบาดแผลถูกสายกระเป๋าเสียดสี เลือดออก

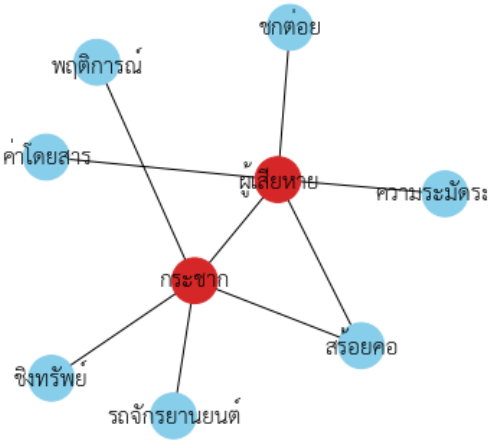
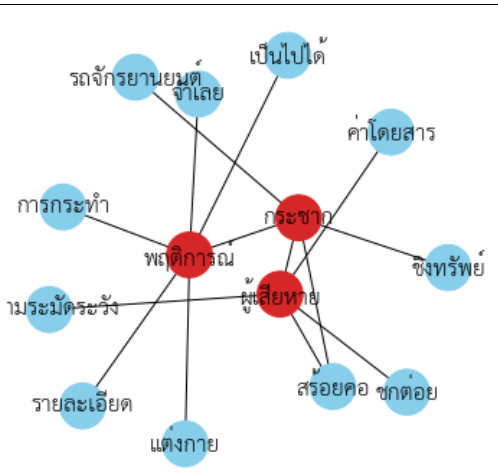
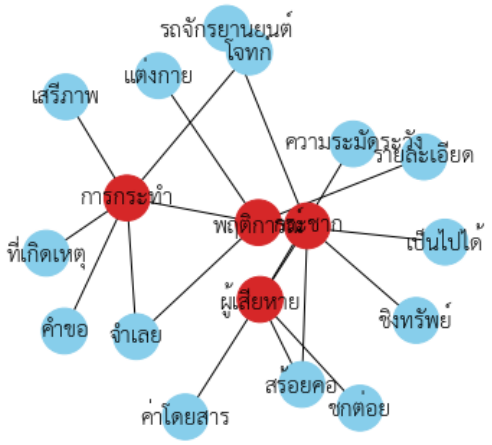
ภาพที่ 63 รายงานประจำวันที่ได้จากกระบวนการสกัดกิจกรรมสำคัญ

ส่วนวิธีทางกราฟวิธีที่เหลือ เมื่อท่องกราฟทั้งหมดได้ผลลัพธ์ ดังตารางที่ 31 - 34

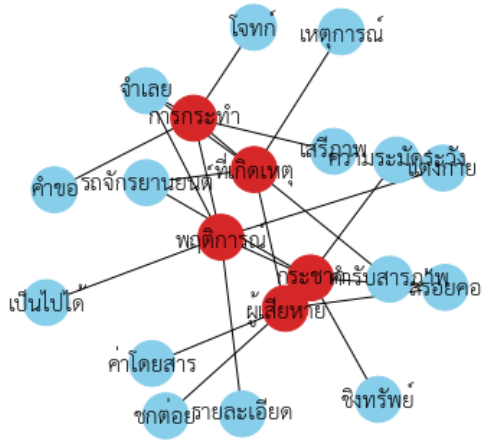
1) วิธีค่ากลางคั่นกลาง

ตารางที่ 31 ตารางแสดงผลลัพธ์การท่องกราฟ

กราฟ	โหนดที่ผู้ใช้เลือก	โหนดที่เกี่ยวข้อง
	ผู้เสียหาย	ความระมัดระวัง สร้อยคอ ค่าโดยสาร กระชาก ชกต่อย

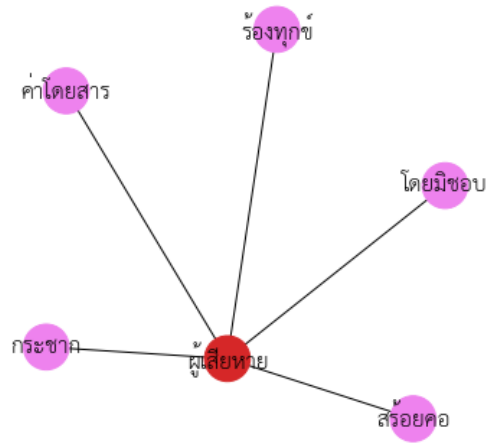
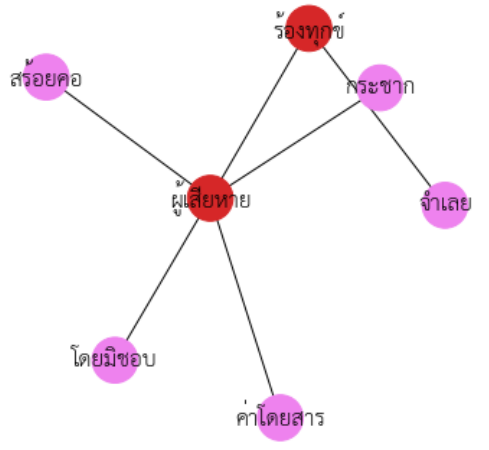
	กระชาก	ผู้เสียหาย สร้อยคอ พฤติการณ์ ชิงทรัพย์ รถจักรยานยนต์
	พฤติการณ์	จำเลย การกระทำ แต่งกาย รายละเอียด เป็นไปได้
	การกระทำ	คำขอ โจทก์ เสรีภาพ จำเลย ที่เกิดเหตุ

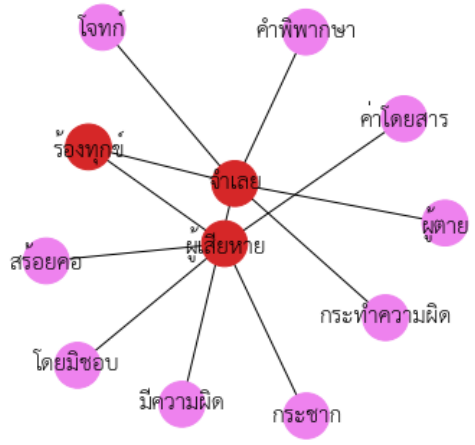
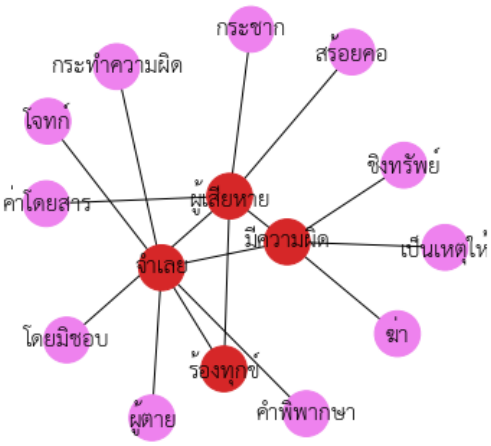
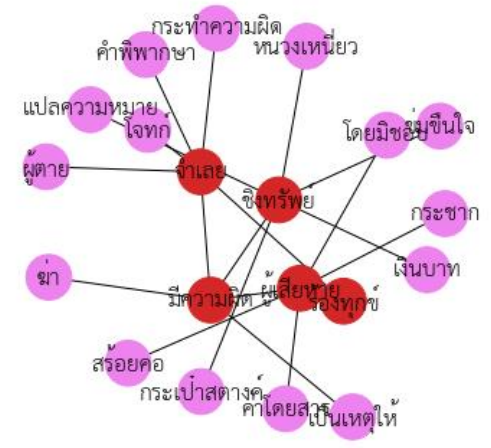
มหาวิทยาลัย
RN UNIVERSITY

	ที่เกิดเหตุ	ผู้เสียหาย รถจักรยานยนต์ จำเลย คำรับสารภาพ เหตุการณ์

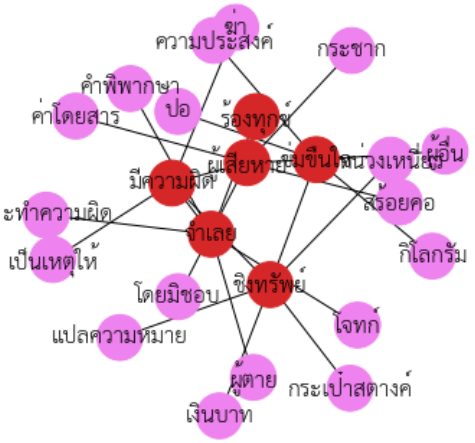
2) วิธีค่ากลางความใกล้

ตารางที่ 32 ตารางแสดงผลลัพธ์การท่องกราฟ

กราฟ	คำสำคัญที่ผู้ใช้เลือก	คำที่เกี่ยวข้อง
	ผู้เสียหาย	ร้องทุกข์ สร้อยคอ โดยมิชอบ ค่าโดยสาร กระชาก
	ร้องทุกข์	จำเลย

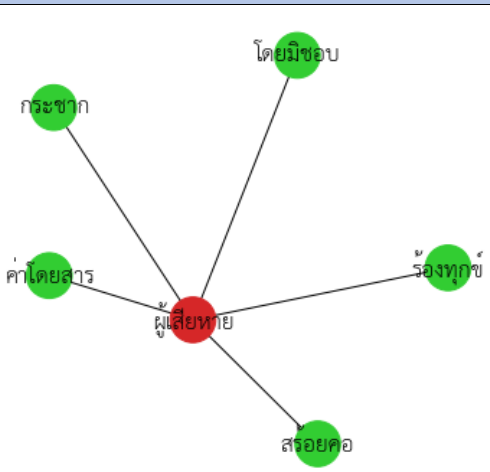
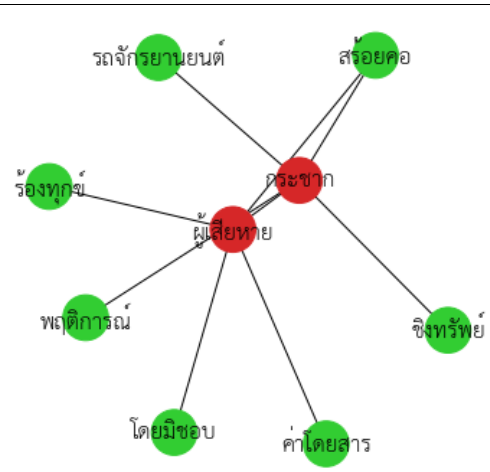
	จำเลย	โจทก์ คำพิพากษาศาลฎีกา ผู้ตาย มีความผิด กระทำความผิด
	มีความผิด	จำเลย ผู้เสียหาย ฆ่า เป็นเหตุให้ ชิงทรัพย์
	ชิงทรัพย์	เงินบาท หน่วงเหนี่ยว กระเป๋าสตางค์ ข่มขืนใจ แปลความหมาย

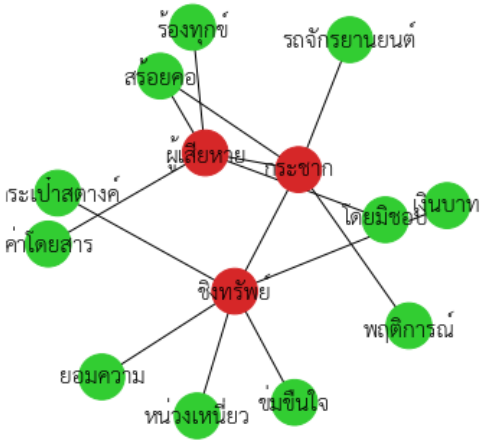
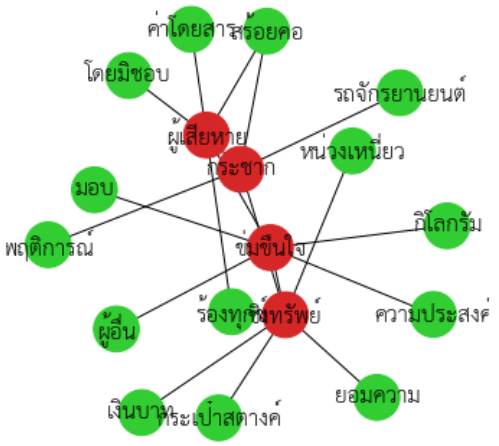
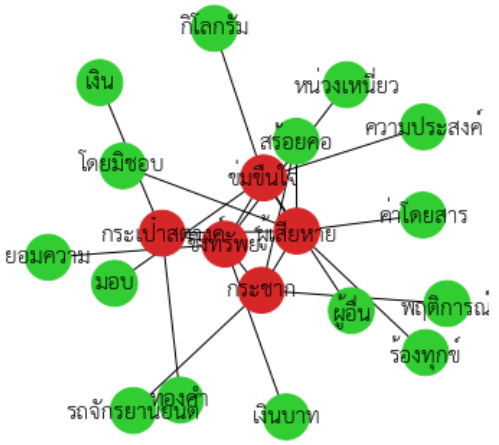
มหาวิทยาลัย
RN UNIVERSITY

	ข่มขืนใจ	ผู้เสียหาย ความประสงค์ ผู้อื่น กิโลกรัม

3) ค่ากลางไอเกนเวกเตอร์

ตารางที่ 33 ตารางแสดงผลลัพธ์การท่องกราฟ

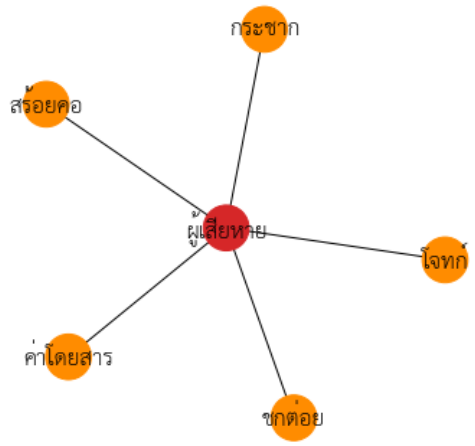
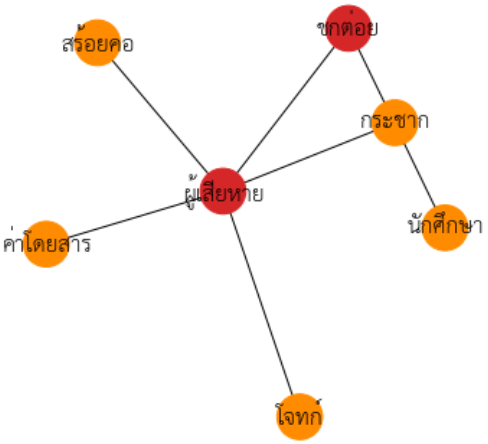
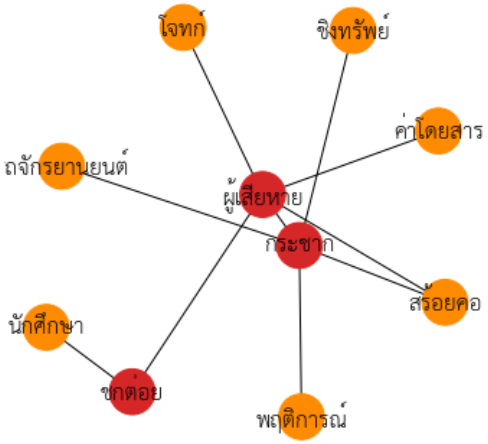
กราฟ	โหนดที่ผู้ใช้เลือก	โหนดที่เกี่ยวข้อง
	ผู้เสียหาย	ร้องทุกข์ สร้อยคอ โดยมิชอบ ค่าโดยสาร กระชาก
	กระชาก	ผู้เสียหาย สร้อยคอ พฤติการณ์ ชิงทรัพย์ รถจักรยานยนต์

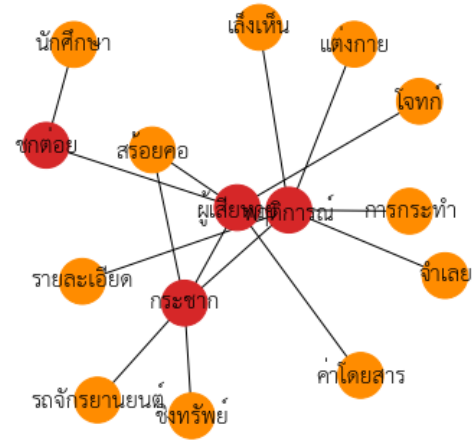
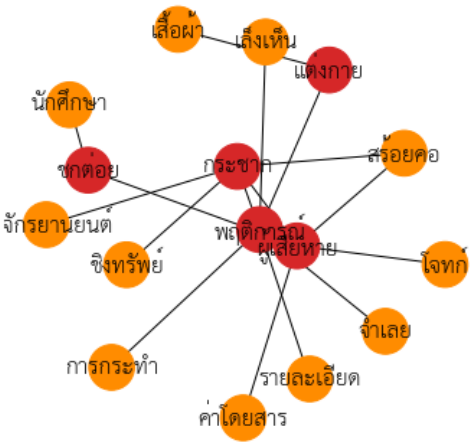
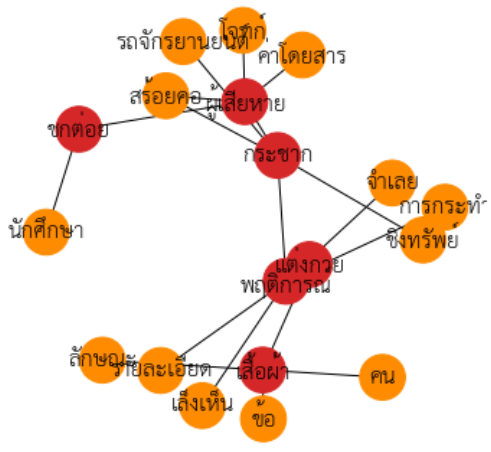
	ชิงทรัพย์	เงินบาท หน่วงเหนี่ยว กระเป๋าสตางค์ ยอม ความ ข่มขืนใจ
	ข่มขืนใจ	ผู้เสียหาย ความประสงค์ มอบ ผู้อื่น กิโลกรัม
	กระเป๋าสตางค์	ผู้เสียหาย เงิน ทองคำ

มหาวิทยาลัย
RN UNIVERSITY

4) ค่ากลางเพจแรงก์

ตารางที่ 34 ตารางแสดงผลลัพธ์การท่องกราฟ

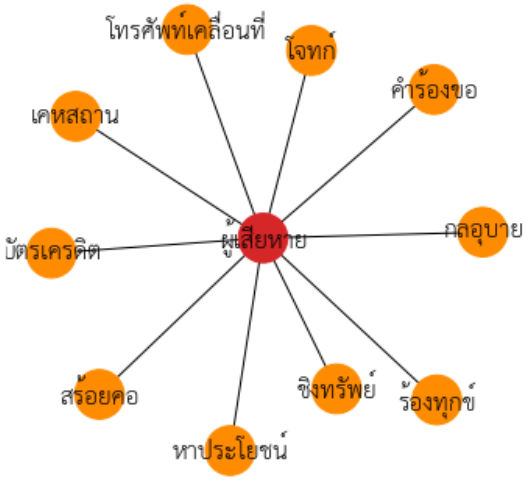
กราฟ	โหนดที่ผู้ใช้เลือก	โหนดที่เกี่ยวข้อง
	ผู้เสียหาย	สร้อยคอ ค่าโดยสาร กระชาก ชกต่อย โจทก์
	ชกต่อย	นักศึกษา
	กระชาก	ผู้เสียหาย สร้อยคอ พฤติการณ์ ชิงทรัพย์ รถจักรยานยนต์

	พฤติการณ์	จำเลย การกระทำ เล็งเห็น แต่งกาย รายละเอียด
	แต่งกาย	เสื้อผ้า
	เสื้อผ้า	ลักษณะ คน ข้อ

4.3.3 การสร้างผลลัพธ์จากการสกัดกิจกรรมของคำพิพากษาศาลฎีกาเพิ่มเติม เฉพาะคดีอาญาความผิดฐานลักทรัพย์

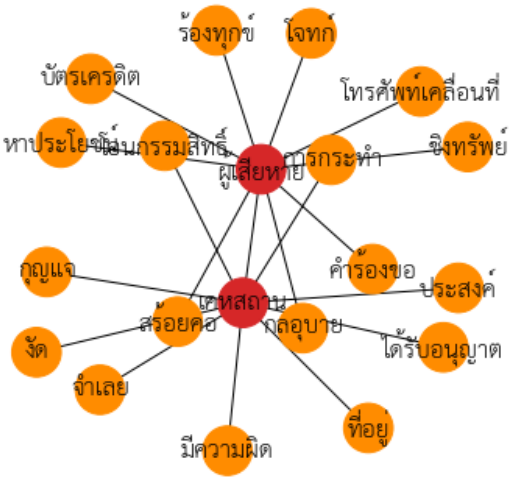
สถานการณ์จำลองที่ 6 - นายเอไปเที่ยวต่างตังหวัดในช่วงหยุดเทศกาล และเมื่อกลับมาที่บ้านพบว่า บ้านถูกงัด จึงเดินทางมาแจ้งความ ณ สถานีตำรวจ

ในกรณีนี้ นายเอพบว่า บ้านของตนถูกงัด เกิดความเสียหาย สถานะของนายเอ จึงเป็นผู้เสียหาย ดังนี้ ขั้นตอนการท่องกราฟเพื่อสร้างรายงานประจำวัน จึงเริ่มต้นที่โหนด **ผู้เสียหาย** ผู้วิจัยทดสอบท่องกราฟจากกราฟที่สรุปด้วยวิธีค่ากลางเพจแรงก์ มีรายละเอียด ดังนี้



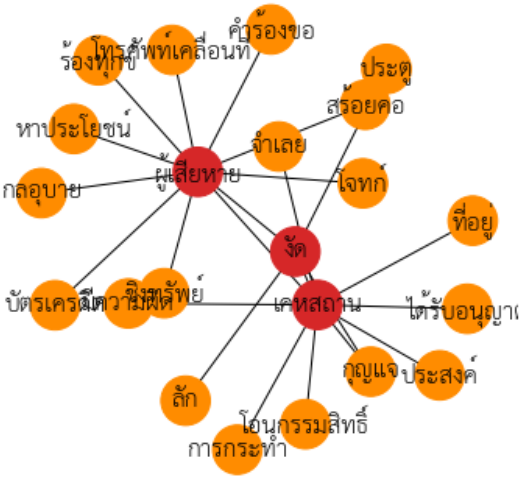
ภาพที่ 64 แสดงผลลัพธ์การท่องกราฟเพื่อสร้างแม่แบบรายงานประจำวันคดีอาญา (เหตุลักทรัพย์)

จากภาพที่ 64 เมื่อเลือกโหนด **ผู้เสียหาย** ระบบจะแสดงโหนดที่เชื่อมกับ ผู้เสียหาย ประกอบด้วย **ร้องทุกข์ บัตรเครดิต คำร้องขอ กลอุบาย สร้อยคอ หาประโยชน์ โจทก์ โทรศัพท์เคลื่อนที่ ชิงทรัพย์ เคหสถาน** จากการสอบสวนเบื้องต้น ผู้เสียหายแจ้งว่า บ้านถูกงัด ดังนั้น ผู้ใช้จึงเลือกโหนด **เคหสถาน** เพื่อสอบถามผู้เสียหายถึง รายละเอียดของ เคหสถาน หรือบ้านของผู้เสียหาย



ภาพที่ 65 แสดงผลลัพธ์การท่องกราฟเพื่อสร้างแม่แบบรายงานประจำวันคดีอาญา (เหตุลักทรัพย์)

จากภาพที่ 65 เมื่อผู้ใช้เลือกโหนด เคหสถาน ระบบจะแสดงโหนดที่เกี่ยวข้อง **ประกอบด้วย โอนกรรมสิทธิ์ ผู้เสียหาย งัด มีความผิด ได้รับอนุญาต กุญแจ จำเลย ที่อยู่ การกระทำ ประสงค์** กรณีนี้ กำหนดให้ผู้ใช้เลือกโหนดจำนวน 2 โหนด โดยโหนดแรกผู้ใช้เลือกโหนด **ที่อยู่** เพื่อถามว่าบ้านของผู้เสียหายตั้งอยู่ที่ใด และโหนดที่สองเลือกโหนด งัด เพื่อเป็นข้อคำถามว่า บ้านถูกงัดอย่างไร มีรอยงัดตรงไหน



ภาพที่ 66 แสดงผลลัพธ์การท่องกราฟเพื่อสร้างแม่แบบรายงานประจำวันคดีอาญา (เหตุลักทรัพย์)

จากภาพที่ 66 เมื่อเลือกคำว่า **งัด** ระบบแสดงโหนดที่เกี่ยวข้อง ประกอบด้วย **ผู้เสียหาย กุญแจ ประตู ลัก** ในกรณีนี้ผู้ใช้เลือกโหนด **ลัก** เพื่อที่จะนำคำดังกล่าวไปถามผู้เสียหายว่า มีทรัพย์สินอะไรถูกขโมยไปบ้างหรือไม่

จากการทำงานของระบบฯ ผู้ใช้เห็นว่าข้อมูลที่ได้เพียงพอต่อการลงบันทึกประจำวันแล้ว จึงหยุดการทำงานของระบบ และรวบรวมรายละเอียดของคำสำคัญที่ได้จากกราฟ พร้อมสร้างประโยคขึ้นมาเป็นตัวอย่าง ดังตารางที่ 35

ตารางที่ 35 แสดงรายการสรุปรายละเอียดโหนดที่ผู้ใช้เลือกทั้งหมด

ลำดับ	โหนดที่เลือก	ประโยค
1	**ผู้เสียหาย**	นายเอ อายุ 30 ปี
2	**เคหสถาน**	เป็นบ้านทาวน์โฮม 3 ชั้น
3	**ที่อยู่**	200/1 ถนนกาญจนาภิเษก อำเภอบางใหญ่ จังหวัดนนทบุรี
4	**งัด**	เมื่อนายเอ กลับถึงบ้าน พบว่ามีรอยงัดด้วยของแข็ง ที่บริเวณประตูรั้ว และประตูภายในบ้าน และถูกเปิดทิ้งไว้
5	**ลัก**	นายเอ เดินเข้าไปในบ้านและตรวจสอบทรัพย์สินของตน พบว่า ทองคำแท่งน้ำหนัก 5 บาท และพระเครื่องสมเด็จ หายไป

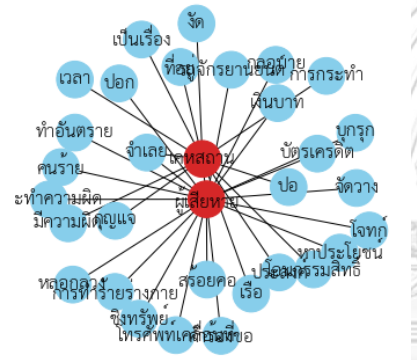
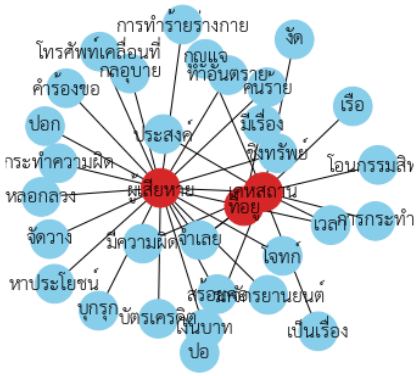
จากตัวอย่างรายงานประจำวันที่ได้นั้น พนักงานสอบสวนสามารถนำข้อมูลรายงานประจำวันข้อนี้ ไปให้เจ้าหน้าที่ฝ่ายสืบสวน เพื่อสืบสวนหาตัวผู้กระทำความผิด หรือเจ้าหน้าที่พิสูจน์หลักฐานเพื่อประสานงานในการดำเนินการเก็บลายนิ้วมือแฝงต่อไป

ส่วนวิธีทางกราฟวิธีที่เหลือ เมื่อท่องกราฟแล้วได้ผลลัพธ์ดังตารางที่ 36 - 39

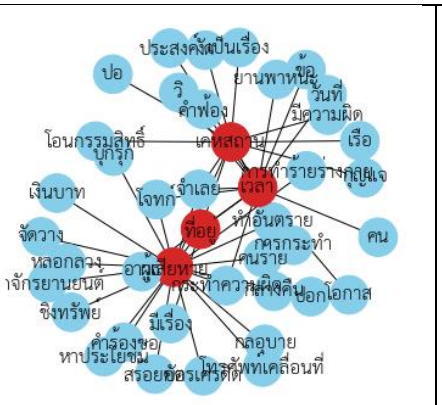
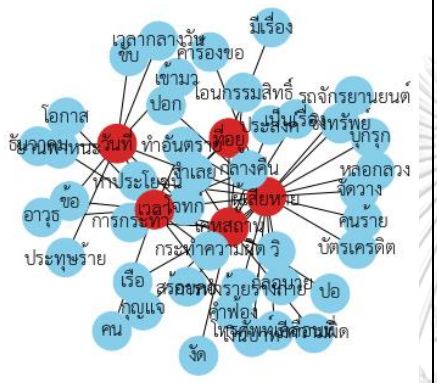
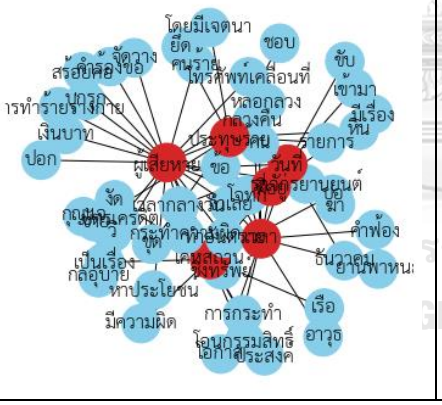
CHULALONGKORN UNIVERSITY

1) วิธีค่ากลางคั่นกลาง

ตารางที่ 36 แสดงผลลัพธ์การท่องกราฟ

กราฟ	โหนดที่ผู้ใช้เลือก	โหนดที่เกี่ยวข้อง
	ผู้เสียหาย	บัตรเครดิต คำร้องขอ กลอุบาย สร้อยคอ หาประโยชน์ โจทก์ โทรศัพท์เคลื่อนที่ ชิงทรัพย์ เคหสถาน เงินบาท ทำอันตราย การทำร้ายร่างกาย กระทำความผิด รถจักรยานยนต์ หลอกลวง คนร้าย บุกรุก จำเลย จัดวาง
	เคหสถาน	โอนกรรมสิทธิ์ ผู้เสียหาย งัด มีความผิด กุญแจ จำเลย ที่อยู่ การกระทำ ประสงค์ เป็นเรื่อง เวลา เรือ
	ที่อยู่	ผู้เสียหาย เวลา มีเรื่อง เคหสถาน จำเลย

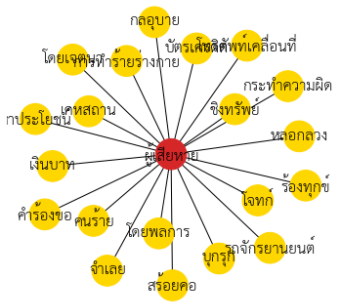
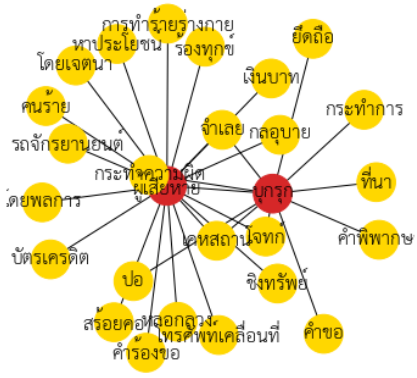
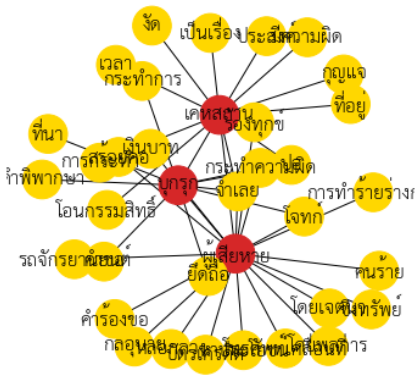
จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

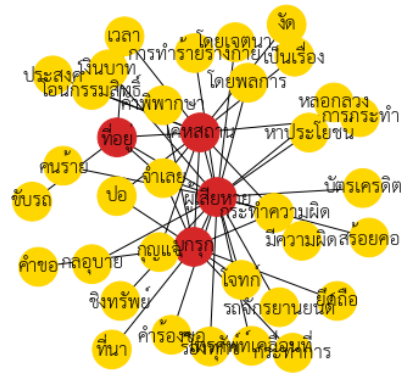
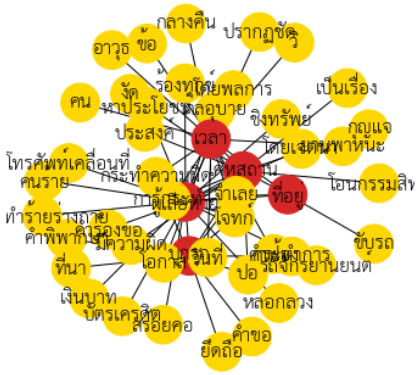
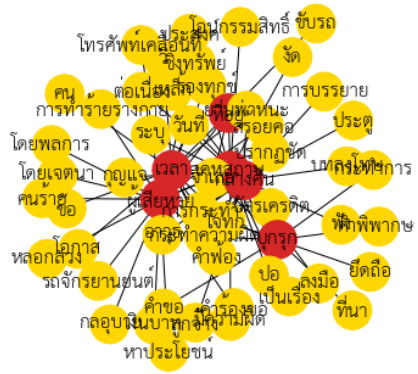
	เวลา	วันที่ วิ โอกาส การกระทำ อาวุธ กลางคืน จำเลย คน คำฟ้อง ยานพาหนะ โจทก์ ข้อ ทำอันตราย กระทำความผิด
	วันที่	เวลา ธันวาคม เวลากลางวัน โจทก์ ข้อ จำเลย ประทุษร้าย ขับ เข้ามา
	ประทุษร้าย	โดยมีเจตนา ผู้เสียหาย ยึด ฆ่า ขุด ขาย ชอบ รายการ หิน

ณ์มหาวิทยาลัย
KORN UNIVERSITY

2) วิธีค่ากลางความเป็นศูนย์กลาง

ตารางที่ 37 แสดงคำที่ได้สำหรับแม่แบบรายงานประจำวัน พร้อมตัวอย่างประโยค

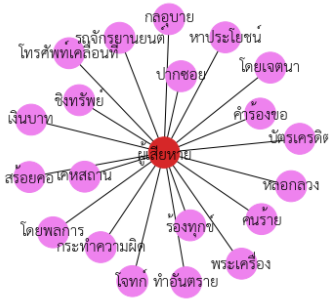
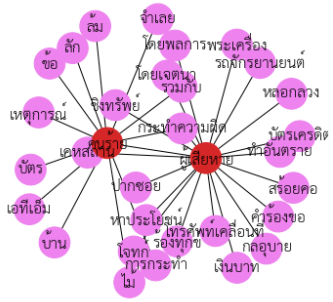
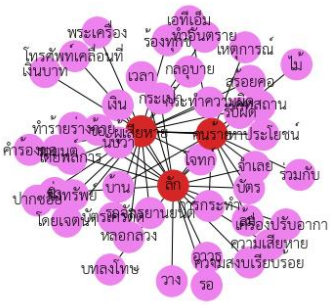
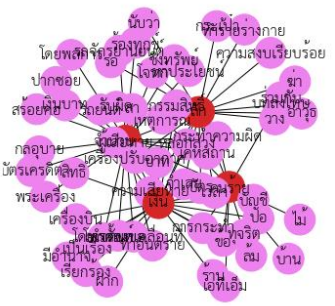
กราฟ	โหนดที่ผู้ใช้เลือก	โหนดที่เกี่ยวข้อง
	ผู้เสียหาย	โดยเจตนา ร้องทุกข์ บัตรเครดิต คำร้องขอ กลอุบาย สร้อยคอ หาประโยชน์ โจทก์ โทรศัพท์เคลื่อนที่ ชิงทรัพย์ เคหสถาน เงินบาท การทำร้ายร่างกาย โดยพลการ กระทำความผิด รถจักรยานยนต์ หลอกลวง คนร้าย บุกรุก จำเลย
	บุกรุก	กระทำความผิด กระทำการโจทก์ ยึดถือ คำพิพากษาศาล ฎีกา คำขอ ที่นา เคหสถาน จำเลย
	เคหสถาน	โอนกรรมสิทธิ์ ผู้เสียหาย งัด มีความผิด กุญแจ จำเลย ที่อยู่ การกระทำ ประสงค์ เป็นเรื่อง เวลา

	ที่อยู่	ขับรถ ผู้เสียหาย เวลา เคหสถาน จำเลย
	เวลา	วันที่ วิ โอกาส การกระทำ อาวุธ กลางคืน จำเลย ปรากฏ ชัด คน คำฟ้อง ยานพาหนะ โจทก์ ข้อ กระทำความผิด
	กลางคืน	ยานพาหนะ ต่อเนื่อง ลงมือ วันที่ ปรากฏชัด บทลงโทษ โจทก์ จำเลย การบรรยาย เคหสถาน คำฟ้อง กระทำ ความผิด ฟัง เหล็ก ลูกจ้าง ระบุ อาวุธ ประตู การกระทำ
	ประตู	ห้องน้ำ บ้าน ที่เกิดเหตุ ขึ้นไป จำเลย ผู้เสียหาย หลุด เคหสถาน เหล็ก เป็นเหตุให้

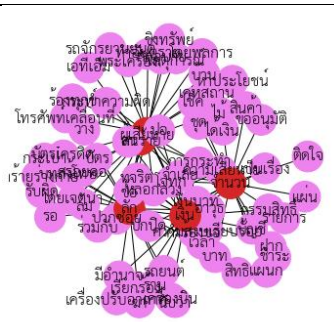
ณ์มหาวิทยาลัย
IGKORN UNIVERSITY

3) วิธีค่ากลางความใกล้

ตารางที่ 38 แสดงคำที่ได้สำหรับแม่แบบรายงานประจำวัน พร้อมตัวอย่างประโยค

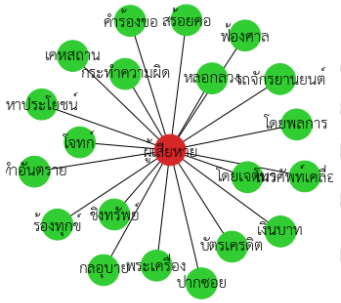
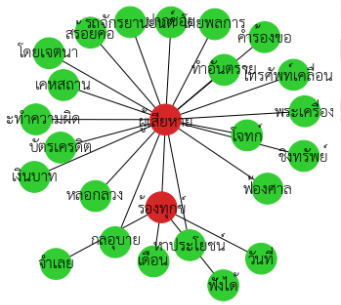
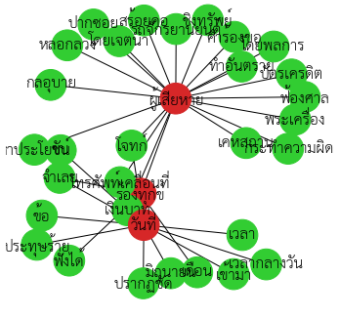
กราฟ	โหนดที่ผู้ใช้เลือก	โหนดที่เกี่ยวข้อง
	ผู้เสียหาย	โดยเจตนา ร้องทุกข์ บัตรเครดิต คำร้องขอ กลอุบาย ปากซอย สร้อยคอ หาประโยชน์ พระเครื่อง โจทก์ โทรศัพท์เคลื่อนที่ ชิงทรัพย์ เคหสถาน เงินบาท ทำอันตราย โดยพลการ กระทำความผิด รถจักรยานยนต์ หลอกลวง คนร้าย
	คนร้าย	ผู้เสียหาย จำเลย กระทำความผิด การกระทำ เหตุการณ์ ล้ม เอทีเอ็ม บัตร บ้าน ไม้ ลัก ร่วมกับ ข้อ
	ลัก	วาง ผู้เสียหาย รอ บัตร รับผิด ความสงบเรียบร้อย เครื่องปรับอากาศ จำเลย บทลงโทษ ความเสียหาย อาวุธ โจทก์ กระเป๋า ฆ่า รถยนต์ คนร้าย เวลา เงิน ทำร้ายร่างกาย นับว่า
	เงิน	จำนวน กรรมสิทธิ์ สิทธิ ทุจริต เรียกร้อง เครื่องบิน บัญชี บาท เวลา จำเลย การกระทำ ข้อ หลอกลวง ฝาก ผู้เสียหาย มีอำนาจ ร้าน เป็นเรื่อง

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

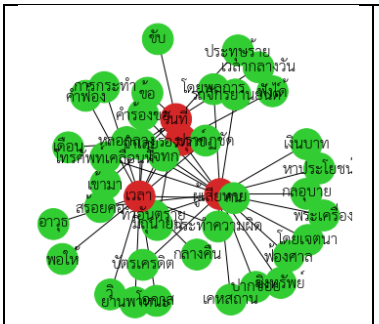
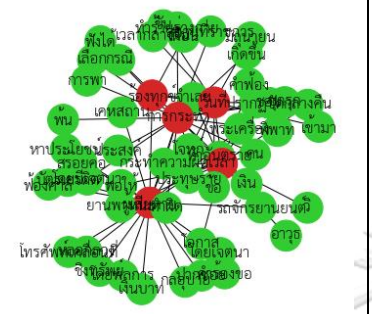
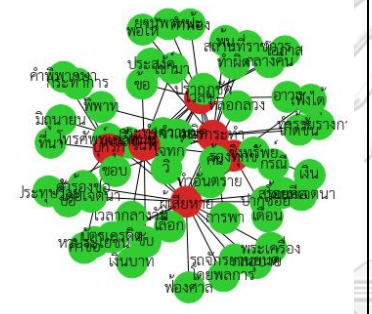
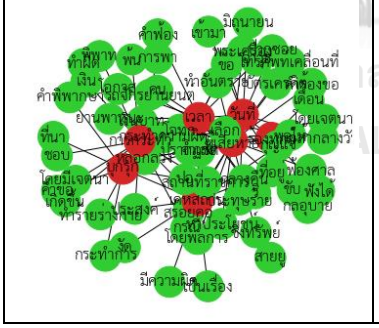
	จำนวน	ชุด เงิน ได้เงิน แผ่น รายการ แผนก บัญชี การกระทำ ขออนุมัติ ผู้เสียหาย โจทก์ สินค้า เช็ค จำเลย ติดใจ บาท เงินบาท ชำระ ปกปิด

4) วิธีค่ากลางไอเกนเวกเตอร์

ตารางที่ 39 แสดงคำที่ได้สำหรับแม่แบบรายงานประจำวัน พร้อมตัวอย่างประโยค

กราฟ	โหนดที่ผู้ใช้เลือก	โหนดที่เกี่ยวข้อง
	ผู้เสียหาย	โดยเจตนา ร้องทุกข์ บัตรเครดิต คำร้องขอ กลอุบาย ปากซอย สร้อยคอ หาประโยชน์ พระเครื่อง โจทก์ ฟ้องศาล โทรศัพท์เคลื่อนที่ ชิงทรัพย์ เคหสถาน เงินบาท ทำอันตราย โดยพลการ กระทำความผิด รถจักรยานยนต์ หลอกลวง
	ร้องทุกข์	ฟังได้ จำเลย เดือน วันที่
	วันที่	เวลา มิถุนายน เวลากลางวัน โจทก์ ข้อ จำเลย ประทุษร้าย ปรากฏชัด ขับ เข้ามา

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

	เวลา	วันที่ วิ โอกาส การกระทำ พอ ให้ อาวุธ กลางคืน จำเลย ปรากฏชัด คน คำฟ้อง ยานพาหนะ โจทก์ ข้อ ทำ อันตราย กระทำความผิด
	การกระทำ	กรณี กระทำความผิด โจทก์ จำเลย ทำร้ายร่างกาย ผู้เสียหาย บุกรุก การพา พิพาท สถานที่ราชการ ทำผิด เกิดขึ้น ประสงค์ เวลา พ้น คน เงิน เลือก ชอบ โดยมีเจตนา
	บุกรุก	กระทำความผิด กระทำการ โจทก์ คำพิพากษาศาลฎีกา คำ ขอ ที่นา เคหสถาน จำเลย
	เคหสถาน	สายยู ผู้เสียหาย งัด มีความผิด กุญแจ จำเลย ที่อยู่ การกระทำ ประสงค์ เป็นเรื่อง เวลา

การท่องกราฟจากสถานการณ์จำลอง ทั้งหมด ระบบเริ่มท่องกราฟตั้งแต่โหนดแรกไปอย่างต่อเนื่อง โดยการคัดเลือกโหนดเริ่มต้นนั้น ผู้ใช้จะเป็นคนเลือก จากสถานการณ์จำลองที่ผู้วิจัยได้กำหนดขึ้น เป็นเหตุที่ผู้เสียหาย เดินทางเข้ามาแจ้งความที่สถานีตำรวจ เมื่อกำหนดโหนดเริ่มต้นแล้ว ระบบก็จะเริ่มการท่องเข้าไปในกราฟจากโหนดเริ่มต้น โดยเมื่อท่องกราฟผ่านแต่ละโหนด ระบบจะแสดงโหนดที่เกี่ยวข้องออกมา ผู้ใช้สามารถนำคำที่แสดงออกมา มาเป็นข้อคำถามสำหรับถามผู้เสียหาย ได้อย่างต่อเนื่อง โดยคำที่แสดงออกมาจะถูกเลือกจากคำที่เกี่ยวข้องทั้งหมด โดยเรียงค่าความสำคัญของแต่ละโหนดจากมากไปน้อย สูงสุด 10 อันดับ เพื่อให้ผู้ใช้สามารถเลือกโหนดที่จะใช้

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ต่อได้อย่างต่อเนื่อง และหยุดก็ต่อเมื่อไม่มีคำที่ผู้ใช้เลือกหรือต้องการแล้ว ทั้งนี้ ขึ้นอยู่กับผู้ใช้หากคำสำคัญที่ระบบสร้างขึ้น มีความครอบคลุม หรือเพียงพอต่อสถานการณ์ที่เกิดขึ้นแล้ว เมื่อดูที่ผลสรุปจากกระบวนการนี้ จะเห็นได้ว่า ระบบสามารถแสดงคำสำคัญที่เกี่ยวข้องในลักษณะเครือข่ายของกราฟทั้งหมด และเมื่อทดสอบสร้างประโยคขึ้นมาจากคำสำคัญที่ผู้ใช้เลือกนั้น สามารถเป็นข้อคำถามสำหรับผู้เข้ามาแจ้งความได้ พร้อมทั้ง ช่วยให้ผู้ใช้เลือกข้อคำถามที่จะซักถามผู้เสียหายต่อไปได้ จนเกิดเป็นแม่แบบสำหรับการบันทึกรายงานประจำวัน ที่สามารถระบุรายละเอียดของเหตุการณ์ที่เกิดขึ้นได้หรือพฤติการณ์แห่งคดี และเป็นเครื่องมือที่ช่วยชี้นำ (Guideline) ในการรับแจ้งความของพนักงานสอบสวน

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

บทที่ 5
การสรุปผล

วิทยานิพนธ์ฉบับนี้ นำเสนอการสกัดกิจกรรมสำคัญจากคำพิพากษาศาลฎีกา เพื่อนำมาสร้างรายงานประจำวันสำหรับตำรวจ โดยใช้ข้อมูลคำพิพากษาศาลฎีกา ในคดีที่เกี่ยวข้องกับการกระทำความผิด ตามพระราชบัญญัติการจราจรทางบก พ.ศ.2522 ประกอบด้วยวิธีการทางกราฟในรูปแบบต่างๆ เพื่อหาว่าคำ หรือกิจกรรมใด ที่ปรากฏในคำพิพากษาศาลฎีกา มีความสำคัญและส่งอิทธิพลต่อกราฟทั้งหมด เพื่อนำมาสร้างแม่แบบสำหรับการสร้างรายงานประจำวันสำหรับตำรวจ พร้อมทั้งเสนอแนวทางการเพิ่มประสิทธิภาพของการสกัดคำสำคัญและกิจกรรมสำคัญเพื่อให้สามารถสร้างรายงานประจำวันได้เกิดประสิทธิภาพ และจากการทดสอบสามารถสรุปผลการวิจัย และข้อจำกัดงานวิจัย และแนวทางในการพัฒนาต่อไป ดังนี้

5.1 สรุปผลการวิจัย

จากที่ผู้วิจัยได้ศึกษา และทบทวนวรรณกรรม งานวิจัยที่เกี่ยวข้องทำให้ผู้วิจัยได้ทราบถึงกระบวนการขั้นตอนการสรุปใจความสำคัญของข้อความด้วยวิธีการต่างๆ ทำให้ผู้วิจัยมีความคิดที่จะได้นำแนวทาง ความรู้ และทฤษฎีจากที่ได้ศึกษามาประยุกต์ร่วมกับงานวิจัยนี้ เพื่อสกัดคำสำคัญหรือกิจกรรมสำคัญจากคำพิพากษาศาลฎีกา และผู้วิจัยได้ออกแบบขั้นตอนวิธีสำหรับการท่องกราฟให้ได้ใจความสำคัญ ที่สามารถนำมาเป็นรายละเอียดหรือพฤติการณ์ที่ปรากฏอยู่บนรายงานประจำวันจนสุดท้ายเกิดเป็นแม่แบบสำหรับรายงานประจำวัน สำหรับการสรุปผลการวิจัย ผู้วิจัยได้แยกเป็นประเด็นการสรุปผล ดังนี้

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

1) ส่วนของวิธีการที่นำมาใช้สำหรับงานวิจัยนี้ ผู้วิจัยได้ทำการศึกษา ทบทวนวรรณกรรมที่เกี่ยวข้องกับการสกัดคำสำคัญ และการสรุปใจความสำคัญของเอกสาร โดยได้นำวิธีการที่ได้ศึกษามาทดสอบการงานวิจัยนี้ ประกอบด้วย

- วิธี TF-IDF เป็นการคำนวณความถี่ของคำที่ปรากฏในคำพิพากษาศาลฎีกา ที่ใช้คัดเลือกคำที่มีความสำคัญต่อคำพิพากษาศาลฎีกาด้วยความถี่ของคำนั้น ๆ วิธี TF-IDF เป็นวิธีหาคำสำคัญในเอกสารวิธีการหนึ่ง โดยคำนวณจากความถี่ของคำที่ปรากฏ เทียบกับจำนวนเอกสารทั้งหมด และวิธี Latent Semantic Analysis (LSA) ที่ใช้เวกเตอร์ของจำนวนคำที่ปรากฏในเอกสารเพื่อหาคำสำคัญที่ซ่อนอยู่ในคำพิพากษาทั้งหมด สำหรับการทดสอบ ผู้วิจัยคัดเลือกจำนวนคำสำคัญจากคำที่มีค่าคะแนนสูงสุด 10 ลำดับ ของคำสำคัญทั้งหมด ซึ่งเมื่อนำคำสำคัญจากทั้งสองวิธีมาวัดผล เทียบกับคำพิพากษาศาลฎีกาทั้งหมดด้วยวิธี 10-fold Cross Validation พบว่าทั้งวิธี TF-IDF และ LSA

มีค่าความแม่นยำ (Precision) เฉลี่ยอยู่ที่ 0.650 และ 0.847 ตามลำดับ และผลลัพธ์ค่าความระลึก (Recall) วิธี TF-IDF และ LSA ให้ค่าความระลึกเฉลี่ยเท่ากับ 0.350 และ 0.473 ตามลำดับ

- วิธีการทางกราฟ (Graph-based Approach) เป็นการนำเสนอคำพิพากษาศาลฎีกาในรูปแบบกราฟ และวิเคราะห์โดยใช้เทคนิคการวิเคราะห์เครือข่ายทางสังคม เพื่อหาความสำคัญทาง ประกอบด้วย วิธีค่ากลางเพจแรงค์ ค่ากลางความเป็นศูนย์กลาง ค่ากลางความใกล้ ค่ากลางคั่นกลาง และค่ากลางไอเกนเวกเตอร์ ผลการทดสอบเมื่อนำคำสำคัญที่ได้จากการสกัดคำสำคัญมาเทียบกับคำพิพากษาศาลฎีกาทั้งหมด โดยใช้วิธีการทดสอบแบบไขว้ เมื่อทดสอบด้วย 10-fold Cross Validation วิธีการ LSA ให้ค่าความแม่นยำสูงที่สุด ที่ 0.847 เมื่อเทียบกับวิธีการแบบอื่น ๆ ส่วนผลลัพธ์ค่าความระลึก วิธีวัดค่ากลางคั่นกล่าง ให้ผลลัพธ์สูงที่สุดที่ 0.701 เมื่อเทียบกับวิธีการอื่น แสดงให้เห็นว่า คำสำคัญที่ได้จากวิธีการดังกล่าว นั้นมีความครอบคลุมกับกราฟของคำพิพากษาศาลฎีกาทั้งหมดมากที่สุด

จากผลการทดสอบวิธีการสกัดคำสำคัญทั้งหมด ได้มีการเปรียบเทียบผลเฉลี่ยของทั้ง 2 วิธีหลัก โดยการทดสอบค่าเฉลี่ย 2 กลุ่ม ระหว่างผลลัพธ์ที่ได้จากวิธีการทางสถิติ และวิธีการทางกราฟ โดยกำหนดระดับนัยสำคัญที่ 0.01 สามารถสรุปผลได้ว่า ผลลัพธ์ระหว่างวิธีการสกัดความสำคัญด้วยวิธี TF-IDF และ LSA มีความแตกต่างกันกับผลลัพธ์ของการสกัดคำสำคัญด้วยวิธีการทางกราฟทั้ง 5 แบบ

2) ในส่วนของขั้นตอนวิธีการสร้างแม่แบบรายงานประจำวันนั้น การสร้างรายงานประจำวันจากคำสำคัญที่ได้จากกราฟ โดยวิธีการคัดเลือกโหนดที่สำคัญจากกราฟทั้ง 5 วิธี สามารถคัดเลือกคำจากคำพิพากษาศาลฎีกาทั้งหมด ผู้วิจัยพัฒนาขั้นตอนวิธีสำหรับการท่องกราฟสรุปของคำพิพากษาศาลฎีกาโดยแสดงโหนดที่มีความสำคัญกับเครือข่าย เพื่อให้ผู้ใช้เลือกโหนดหรือคำสำคัญดังกล่าวไปใช้งานได้อย่างต่อเนื่อง และทุกๆ โหนดที่ผู้ใช้เลือก ระบบจะแสดงโหนดอื่นๆ ที่เกี่ยวข้องและมีความสำคัญกับคำพิพากษาศาลฎีกาทั้งหมด เพื่อทำให้ผู้ใช้รู้ได้ว่า เมื่อเลือกโหนดที่แสดงแล้ว ผู้ใช้ควรจะต้องเลือกโหนดใดต่อไปในการซักถาม หรือสอบสวนกับผู้ร้องทุกข์หรือผู้เสียหาย เพื่อเป็นเครื่องมือในการบันทึกรายงานประจำวัน จนเกิดเป็นแม่แบบสำหรับรายงานประจำวัน เมื่อได้ผลลัพธ์จากงานวิจัย ผู้วิจัยได้ทำการสัมภาษณ์เจ้าหน้าที่ตำรวจผู้มีประสบการณ์เกี่ยวกับการบันทึกรายงานประจำวัน และงานคดี เพื่อสอบถามถึงปัญหาที่พบในการทำงาน และข้อเสนอแนะเกี่ยวกับงานวิจัย หากมีการนำวิธีการนี้ มาใช้ในการรับแจ้งความในปัจจุบันจะมีผลดีต่อการปฏิบัติงานของเจ้าหน้าที่หรือไม่ จากการสัมภาษณ์ได้ข้อสรุปว่า งานวิจัยนี้ มีส่วนช่วยในการบันทึกรายงานประจำวัน ให้มีเกิดความครอบคลุมมากยิ่งขึ้น และเป็นเครื่องมือที่ช่วยแนะนำผู้ที่ยังไม่มีประสบการณ์ในการทำงานในด้านนี้มากนัก แต่มีข้อสังเกตในส่วนการสร้างคำสำคัญ ขึ้นมา หากคำสำคัญที่ถูกสร้าง

ขึ้นมาในบางครั้ง เมื่อนำไปใช้อาจทำให้เกิดช่องโหว่ของรูปคดี เพราะฉะนั้น ผู้ใช้ควรมีประสบการณ์ในด้านกฎหมายและการสอบสวนอยู่บ้าง

ผู้วิจัยได้สร้างกราฟเพิ่มเติม ในส่วนของคำพิพากษาศาลฎีกาที่เป็นความผิดฐานชิงทรัพย์และลักทรัพย์ เพื่อทดสอบว่า ขั้นตอนวิธีที่ผู้วิจัยสร้างขึ้นมามีประสิทธิภาพครอบคลุมกับคดีประเภทอื่นๆ หรือไม่ ซึ่งจากการทดสอบ พบว่า การสร้างแม่แบบสำหรับรายงานประจำวัน เมื่อกำหนดคำเริ่มต้น ระบบสามารถสร้างคำที่เกี่ยวข้องกับคำเริ่มต้น ผู้ใช้สามารถเลือกคำที่ระบบแสดงขึ้น เพื่อใช้เป็นข้อคำถาม สำหรับการซักถามระหว่างตำรวจ และ ผู้ที่เข้ามาแจ้งความร้องทุกข์ได้ โดยเมื่อผู้ใช้เลือกคำที่ต้องการจากที่ระบบสร้างขึ้นมาให้แล้วนั้น ระบบก็จะนำคำที่ผู้ใช้เลือกมาประมวลผลและแสดงคำที่เกี่ยวข้องขึ้นมาอีก เพื่อช่วยให้การบันทึกรายงานประจำวันมีความสะดวกรวดเร็วมากยิ่งขึ้น โดยเมื่อผู้ใช้เลือกใช้งานจากคำนี้แล้ว คำถามต่อไปหรือหัวข้อที่สำคัญต่อไปควรจะเป็นอะไรบ้าง เพื่อให้ผู้ใช้สามารถนำคำเหล่านี้มาใช้ถามตอบไปได้อย่างต่อเนื่อง จนเกิดเป็นแม่แบบสำหรับการบันทึกรายงานประจำวัน ทั้งนี้ คำพิพากษาศาลฎีกานั้น เป็นกระบวนการในชั้นศาล ซึ่งกว่าคดีความจะถึงเข้าสู่ชั้นศาลได้ จะเริ่มต้นที่พนักงานสอบสวน (ตำรวจ) เป็นผู้บันทึกรายงานประจำวัน สอบสวนคดี และทำสำนวนการสอบสวนส่งไปยังพนักงานอัยการ ดังนั้น หากมีการพัฒนาระบบที่ช่วยสร้างรายงานประจำวันที่ช่วยอำนวยความสะดวกแก่พนักงานสอบสวนนั้น จะช่วยให้ผู้ปฏิบัติบันทึกรายงานที่ได้เกิดประสิทธิภาพ ลดภาระงานของผู้ปฏิบัติ และสามารถให้บริการประชาชนได้รวดเร็วมากยิ่งขึ้น

5.2 ข้อจำกัดงานวิจัย

การสร้างกราฟใช้การวิเคราะห์จากคำพิพากษาศาลฎีกา ซึ่งในคำพิพากษาศาลฎีกามักจะประกอบไปด้วยคำศัพท์ในชั้นศาล จึงไม่อาจครอบคลุมในรายละเอียดของรายงานประจำวันทั้งหมด ในกรณีที่มีการบันทึกรายงานประจำวันคดีอาญา และการบันทึกรายงานประจำวันรูปแบบอื่น ทั้งนี้ ผู้วิจัยได้กำจัดคำออกไปในบางส่วนแล้ว จากรายการคำศัพท์ต่างๆ เพื่อให้กราฟเหลือแต่คำที่สำคัญและลดคำที่ไม่เกี่ยวข้องออกไป

5.3 แนวทางวิจัยในอนาคต

ด้วยงานวิจัยนี้เป็นแนวทางและต้นแบบสำหรับการพัฒนาระบบงาน อื่นๆเพื่อใช้ในสถานีตำรวจ ดังนี้

5.3.1 การพัฒนาโปรแกรมประยุกต์สำหรับการสร้างรายงานประจำวัน โดยนำเทคนิคของงานวิจัยนี้ไปเป็นต้นแบบสำหรับการสร้างรายงานประจำวัน จากคำสำคัญโดยประชาชนสามารถกรอกข้อมูลเบื้องต้น ได้ก่อนที่จะไปแจ้งความที่สถานีตำรวจ เพื่อให้เกิดความรวดเร็วเมื่อไปถึงสถานีตำรวจ ลดระยะเวลา และขั้นตอนการปฏิบัติบนสถานีตำรวจ

5.3.2 การนำเทคนิคของงานวิจัยนี้ไปใช้ในการวิเคราะห์เครือข่ายสังคมของข้อมูลคดี เพื่อสกัดคำสำคัญหรือกิจกรรมสำคัญ และนำผลลัพธ์ที่ได้ไปใช้ในงานสืบสวนสอบสวน อาทิ การสกัดคำสำคัญเพื่อสร้างแผนประทุษกรรม (ข้อมูลที่อธิบาย วิธีการในการกระทำความผิดของอาชญากร)

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

บรรณานุกรม

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

[1] พ.ร.บ.ตำรวจแห่งชาติ, 2547.

[2] *ประมวลระเบียบการตำรวจเกี่ยวกับคดี ลักษณะ 12,* กรมตำรวจ, 2537.

[3] “จำนวนประชากรจากการทะเบียน จำแนกตามกลุ่มอายุ รายจังหวัด และภาค พ.ศ. 2554 - 2563,” สำนักงานสถิติแห่งชาติ, 2563.

[4] ป. ปิติอลงกรณ์, ช. จิรรัตนชาญ และ ณ. บูรณประภานนท์, *การประมวลผลคำไทยด้วยคอมพิวเตอร์: ก้าวแรกของระบบประมวลผลสารสนเทศอัจฉริยะ,* 2542.

[5] A. Siddharthan, A survey of research in text simplification, Department of Computing Science, University of Aberdeen, 2015.

[6] L. Specia, “Translating from complex to simplified sentences,” ใน *International Conference on Computational Processing of the Portuguese Language*, 2010.

[7] J. Carroll, G. Minnen, Y. Canning, S. Devlin และ J. Tait, Practical Simplification of English Newspaper Text to Assist Aphasic Readers, University of Sussex, 1998.

[8] ม. นาคใหญ่, การวิเคราะห์ประโยคข้อมูลโดยอิงฐานความรู้ กรณีศึกษาการครองราชย์สมัยกรุงศรีอยุธยา, มหาวิทยาลัยศิลปากร, 2549.

[9] Christian, Hans; Agus, Mikhael Pramodana; Suhartono, Derwin;, “SINGLE DOCUMENT AUTOMATIC TEXT SUMMARIZATION USING TERM FREQUENCY-INVERSE DOCUMENT FREQUENCY (TF-IDF),” *ComTech Computer Mathematics and Engineering Applications,* 2016.

[10] J. Leskovec, . M.-F. Natasa และ G. Marko , “Extracting Summary Sentences Based on the Document,” 2005.

[11] ท. อุทัยสุริ, การสกัดคำสำคัญจากบทคัดย่อภาษาอังกฤษ, มหาวิทยาลัยศิลปากร, 2556.

[12] S. Poltree และ K. Saikaew, “Thai Web Forum Topic Suggestion Using Thai

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

WordNet Graph," ใน *Proceedings of the Third International Conference on Knowledge and Smart Technologies 2011*, 2011.

[13] ว. เรืองฤทธิ์, "การวิเคราะห์เครือข่ายทางสังคมของกระแสการเดินทางไปทำงานในภูมิภาคเมืองโคราช," *Veridian E-Journal, SU Vol.5 No. 3,* pp. 254-266, 2555.

[14] K. Kastriot และ . O. Milenko, "Extractive approach for text summarization using graphs," Ljubljana, Slovenia, 2021.

[15] M. P. J. P. R. H. Andry Alamsyah, "Network Text Analysis to Summarize Online Conversations for Marketing Intelligence Efforts in Telecommunication Industry," ใน *2016 4th International Conference on Information and Communication Technology (ICoICT)*, Bandung, Indonesia, 2016.

[16] บ. กิจศิริกุล, ปัญญาประดิษฐ์ เอกสารคำสอนวิชา 2110654, ภาควิชาวิศวกรรมคอมพิวเตอร์ จุฬาลงกรณ์มหาวิทยาลัย, 2546.

[17] "ทำความรู้จัก Dependency Parser กับภาษาไทย," [ออนไลน์]. Available: https://medium.com/super-ai-engineer/%E0%B8%97%E0%B8%B3%E0%B8%84%E0%B8%A7%E0%B8%B2%E0%B8%A1%E0%B8%A3%E0%B8%B9%E0%B9%89%E0%B8%88%E0%B8%B1%E0%B8%81-dependency-parser-%E0%B8%81%E0%B8%B1%E0%B8%9A%E0%B8%A0%E0%B8%B2%E0%B8%A9%E0%B8%B2%E0%B9%84%E0%B8%97%E0%B8%A2-.

[18] ว. อรุณมานะกุล และ ธ. แจ่มไพบูลย์, คู่มือการกำกับข้อมูล Thai Dependency Tree Bank ตามแนว Universal Dependencies v.2, ศูนย์วิจัยการประมวลผลภาษาและวัจนะ ภาควิชาภาษาศาตร์ คณะอักษรศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย, 2563.

[19] อ. อินทรสิทธิ์ และ น. มะกาเจ, "ทฤษฎีกราฟเบื้องต้นและการประยุกต์," *วารสารวิชาการพระจอมเกล้าพระนครเหนือ,* %13, 2558.

Chulalongkorn University

[20] ม. ปัจฉิมนันท, “การวิเคราะห์เครือข่ายทางสังคมกับการวิจัยด้านการสื่อสารในองค์การ,” *วารสารสถาบันพระปกเกล้า ,* %1 ฉบับที่ 3 , pp. หน้า 5-18, 2560.

[21] S. B. Lawrence Page, The PageRank Citation Ranking: Bringing Order to the Web, 1999.

[22] ศ. ประกิ่ง, ส. พิมพ์ชารี, ป. ทอดทอง, พ. โคตรสุวรรณ์ และ ป. อุตรนคร, “ศูนย์กลางเครือข่ายทางสังคมในการแลกเปลี่ยนข้อมูลของสมาชิกกลุ่ม กรณีศึกษากลุ่มปลูกผักปลอดภัยจากสารพิษบ้านหม้อ อำเภอซำสูง จังหวัดขอนแก่น,” *แก่นเกษตร,* %1พิเศษ 1, pp. 1065-1070, 2562.

[23] ธ. พุกเส็ง, “การวิเคราะห์โครงข่ายของสายน้ำ ด้วยทฤษฎีกราฟบนหลักการความเป็นจุดศูนย์กลาง: กรณีศึกษาพื้นที่สุ่มเสี่ยงต่อการเกิดปัญหาน้ำ ท่วมรอการระบายในจังหวัดจันทบุร,” 2020.

[24] ภ. ภูมิค้า, “เทคนิคการเรียนรู้เชิงลึกเพื่อวิเคราะห์ความรู้สึกจากผู้ใช้ผลิตภัณฑ์,” มหาวิทยาลัยเทคโนโลยีสุรนารี, 2562.

[25] P. NETISOPAKUL, “Word Similarity Datasets for Thai:Construction and Evaluation,” 2016.

[26] [ออนไลน์]. Available: http://deka.supremecourt.or.th/.

[27] กรมคุ้มครองสิทธิและเสรีภาพ, คู่มือล่ามในกระบวนการยุติธรรม.

[28] [ออนไลน์]. Available: https://pythainlp.github.io/docs/2.0/api/tokenize.html.

[29] [ออนไลน์]. Available: https://pythainlp.github.io/dev-docs/api/tag.html.

[30] ISKANDER AKHMETOV, ALEXANDER GELBUKH และ RUSTAM MUSSABAYEV, “Greedy Optimization Method for Extractive Summarization of Scientific Articles,” IEEE, December 2021.

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ภาคผนวก

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ภาคผนวก ก.
ตัวอย่างข้อมูลคำพิพากษาศาลฎีกา และชุดคำสั่ง (Source Code)

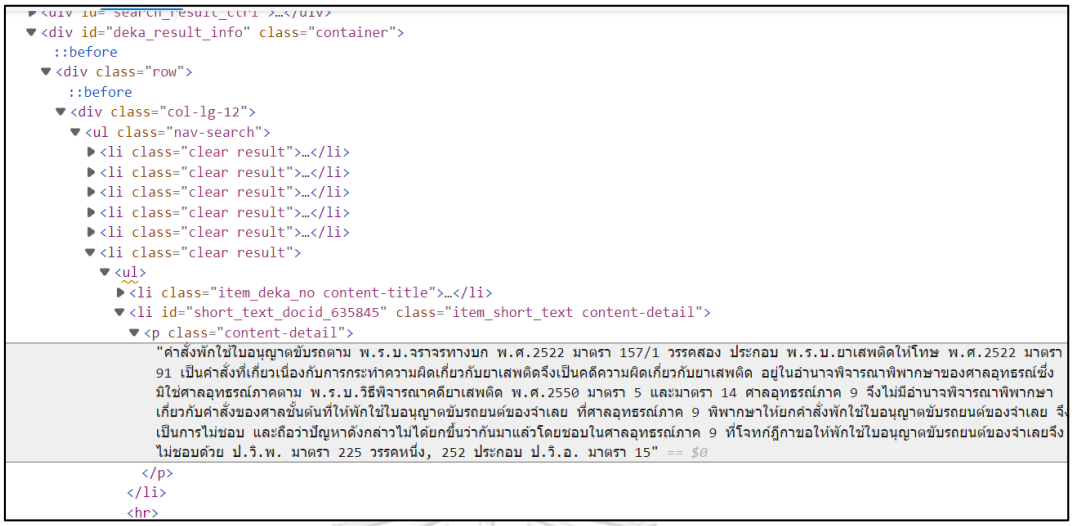
1. แหล่งสืบค้นข้อมูลคำพิพากษาศาลฎีกาจาก http://deka.supremecourt.or.th/

2. ตัวอย่างคำพิพากษาศาลฎีกาที่ 169/2562 –

การที่โจทก์ฟ้องว่าจำเลยขับรถในทางก่อให้เกิดความเสียหายแก่บุคคลแล้วไม่หยุดรถให้การช่วยเหลือและแสดงตัวต่อเจ้าพนักงานเป็นเหตุให้บุคคลนั้นถึงแก่ความตาย ตาม พ.ร.บ.จราจรทางบก พ.ศ.2522 โจทก์ร่วมไม่ใช่ผู้ได้รับความเสียหายเนื่องจากการกระทำความผิดฐานนี้โดยตรง จึงไม่ใช่ผู้เสียหายตามกฎหมายและไม่มีสิทธิอุทธรณ์ขอให้ลงโทษจำเลยในความผิดตาม พ.ร.บ. ดังกล่าวได้ และเมื่อโจทก์มิได้บรรยายฟ้องอ้างเหตุว่าการที่จำเลยไม่หยุดรถให้ความช่วยเหลือตามสมควร ไม่แสดงตัวและไม่แจ้งเหตุต่อพนักงานเจ้าหน้าที่ที่ใกล้เคียงทันทีอันเป็นการไม่ปฏิบัติตามมาตรา 78 แห่ง พ.ร.บ.จราจรทางบกฯ นั้น เป็นเหตุให้บุคคลอื่นถึงแก่ความตาย คดีจึงไม่อาจลงโทษจำเลยตาม พ.ร.บ. จราจรทางบกฯ มาตรา 160 วรรคสอง ได้ และปัญหานี้แม้คู่ความจะมิได้อุทธรณ์ฎีกา แต่เป็นปัญหาที่เกี่ยวกับความสงบเรียบร้อย ศาลฎีกามีอำนาจยกขึ้นวินิจฉัยเองได้ ตาม ป.วิ.อ. มาตรา 195 วรรคสอง ประกอบมาตรา 225 ส่วนที่จำเลยฎีกาว่า ขณะเกิดเหตุ รถจำเลยแล่นอยู่ในช่องเดินรถที่ 1 มิได้วิ่งคร่อมช่องเดินรถที่ 1 และที่ 2 การที่รถจำเลยเฉี่ยวชนกับรถจักรยานยนต์ของผู้ตาย ผู้ตายจึงมีส่วนประมาทอยู่ด้วย ทำนองว่าจำเลยมิได้กระทำความผิดตามฟ้องนั้น เป็นข้อเท็จจริงที่ยกขึ้นใหม่และขัดแย้งกับคำให้การรับสารภาพของจำเลย จึงเป็นข้อที่มิได้ยกขึ้นว่ากันมาแล้วโดยชอบในศาลชั้นต้นและศาลอุทธรณ์ภาค 7 ต้องห้ามมิให้ฎีกาตาม ป.วิ.พ. มาตรา 225 วรรคหนึ่ง และมาตรา 252 ประกอบ ป.วิ.อ. มาตรา 15 ศาลฎีกาไม่รับวินิจฉัย

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

3. ตัวอย่างรูปแบบของไฟล์ HTML สำหรับการดึงข้อมูลคำพิพากษาศาลฎีกา



ภาพที่ 67 ตัวอย่างไฟล์ HTML จากเว็บไซต์สืบค้นคำพิพากษาศาลฎีกา

4. ข้อมูลคำพิพากษาศาลฎีกาที่ใช้ทั้งหมด

https://drive.google.com/drive/folders/1zJPPR-y9EyyTBVr-9RIaHt5xdhsPRTyZ?usp=sharing

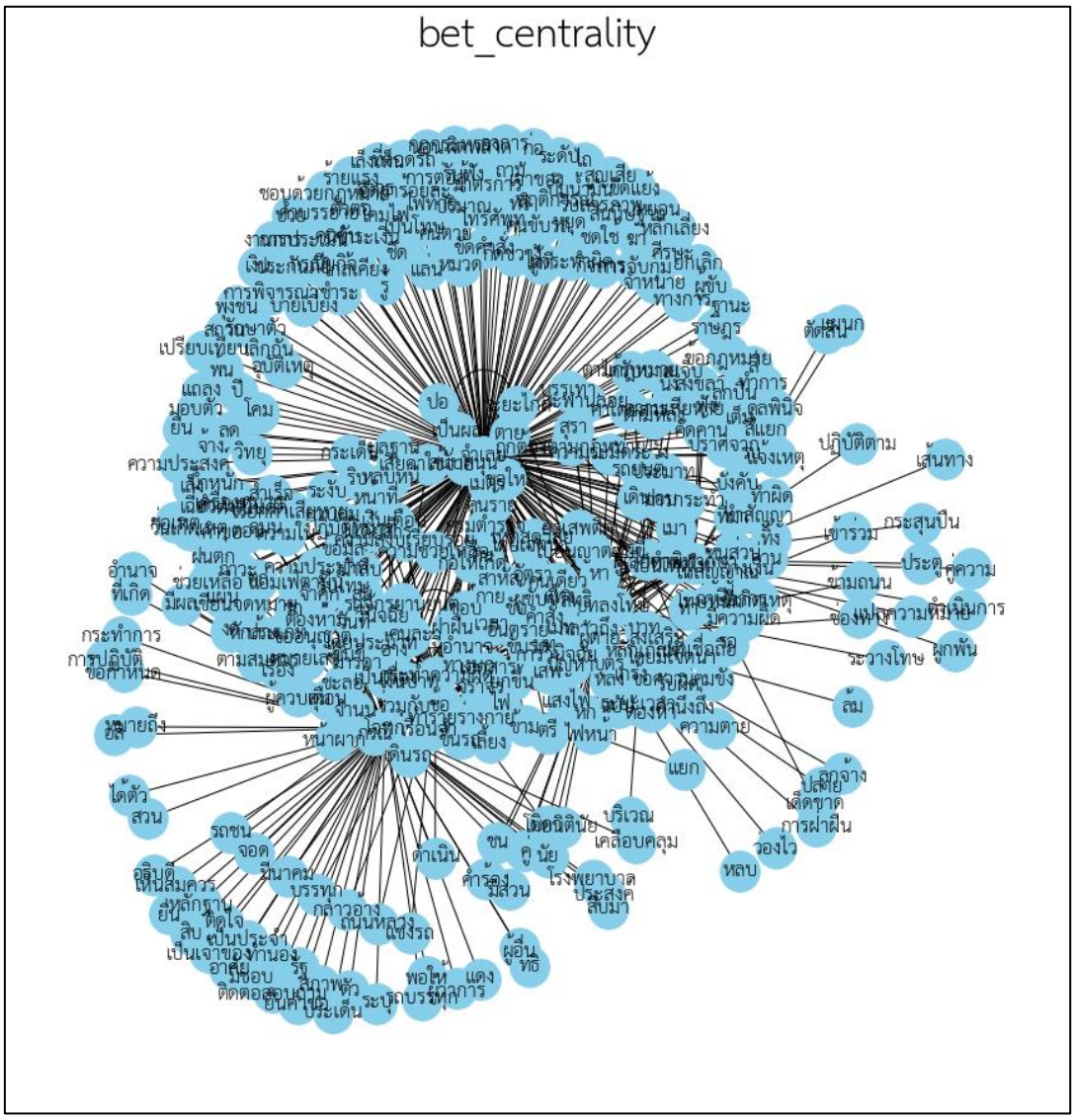
5. ชุดคำสั่ง (Source Code) ภาษา Python บน Google Colab

https://colab.research.google.com/drive/1hVWoOfGSNwvKVD0rjFrVS9_ixblujz2u?usp=sharing

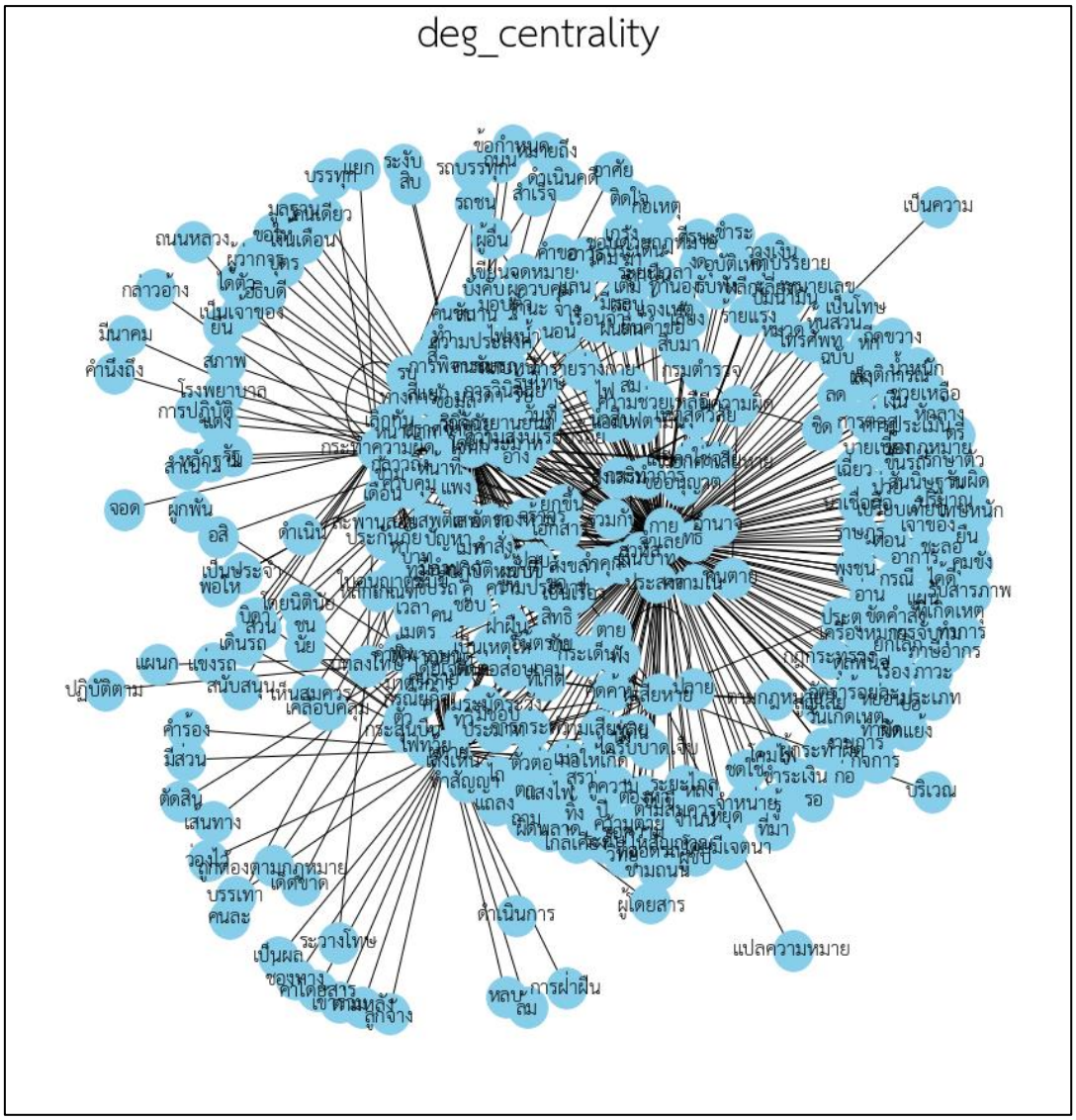
จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University

ภาคผนวก ข.

ภาพผลลัพธ์กราฟเมื่อผ่านกระบวนการสกัดกิจกรรมสำคัญ

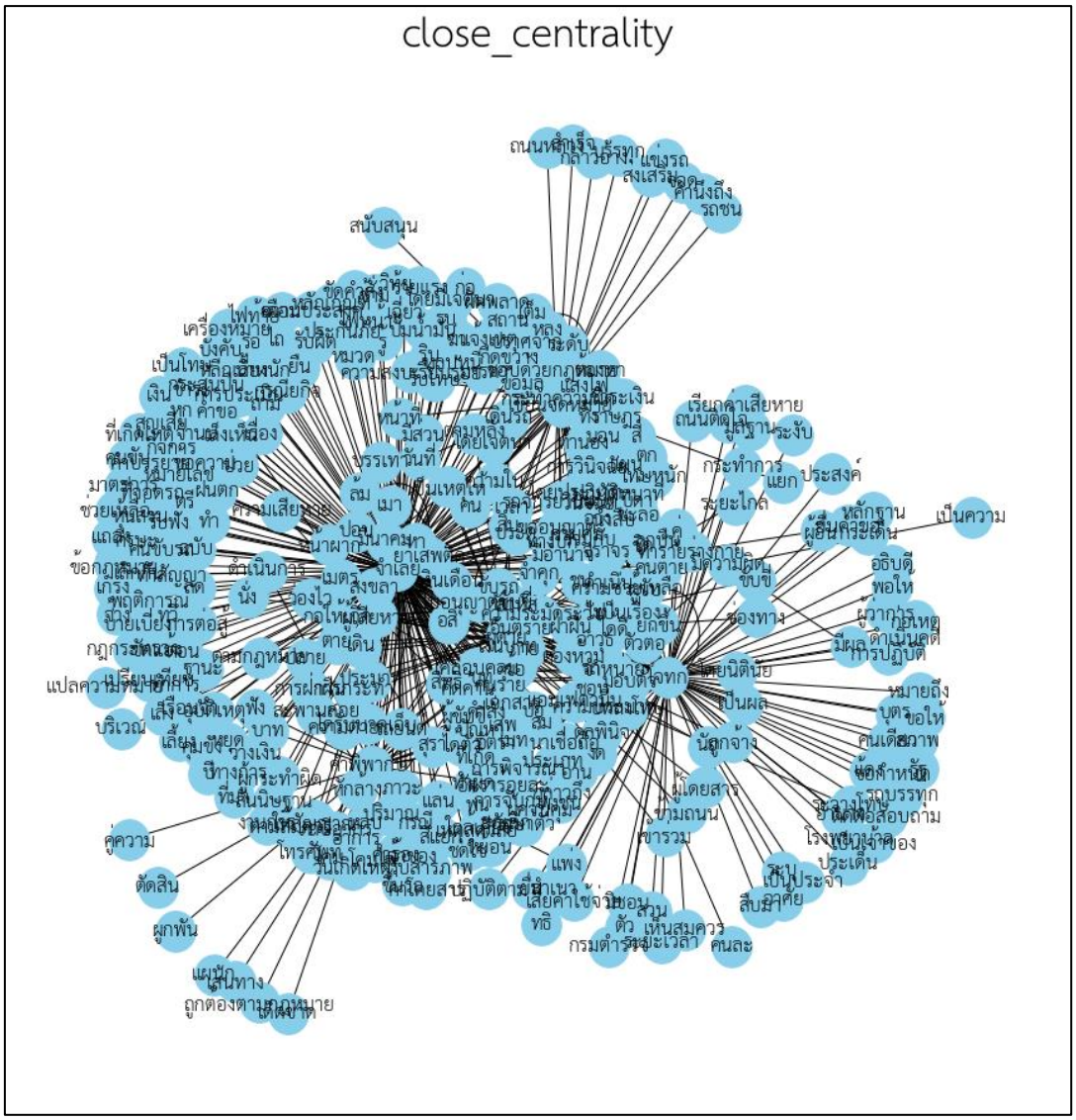


ภาพที่ 68 แสดงผลลัพธ์กราฟหลังจากผ่านการสกัดกิจกรรมสำคัญด้วยวิธีค่ากลางคั่นกลาง



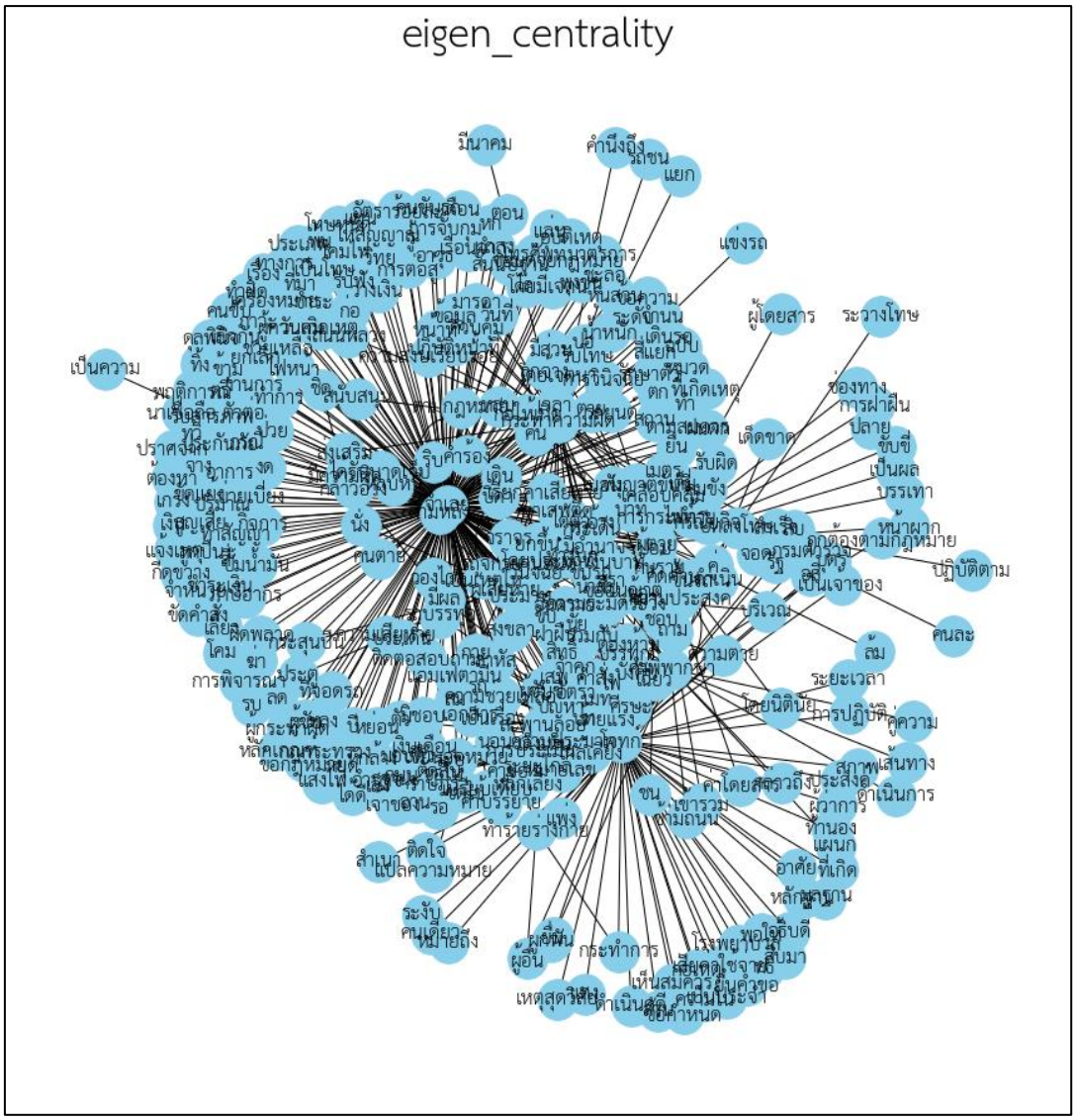
CHULALONGKORN UNIVERSITY

ภาพที่ 69 แสดงผลลัพธ์กราฟหลังจากผ่านการสกัดกิจกรรมสำคัญด้วยวิธีค่ากลางความเป็นศูนย์กลาง



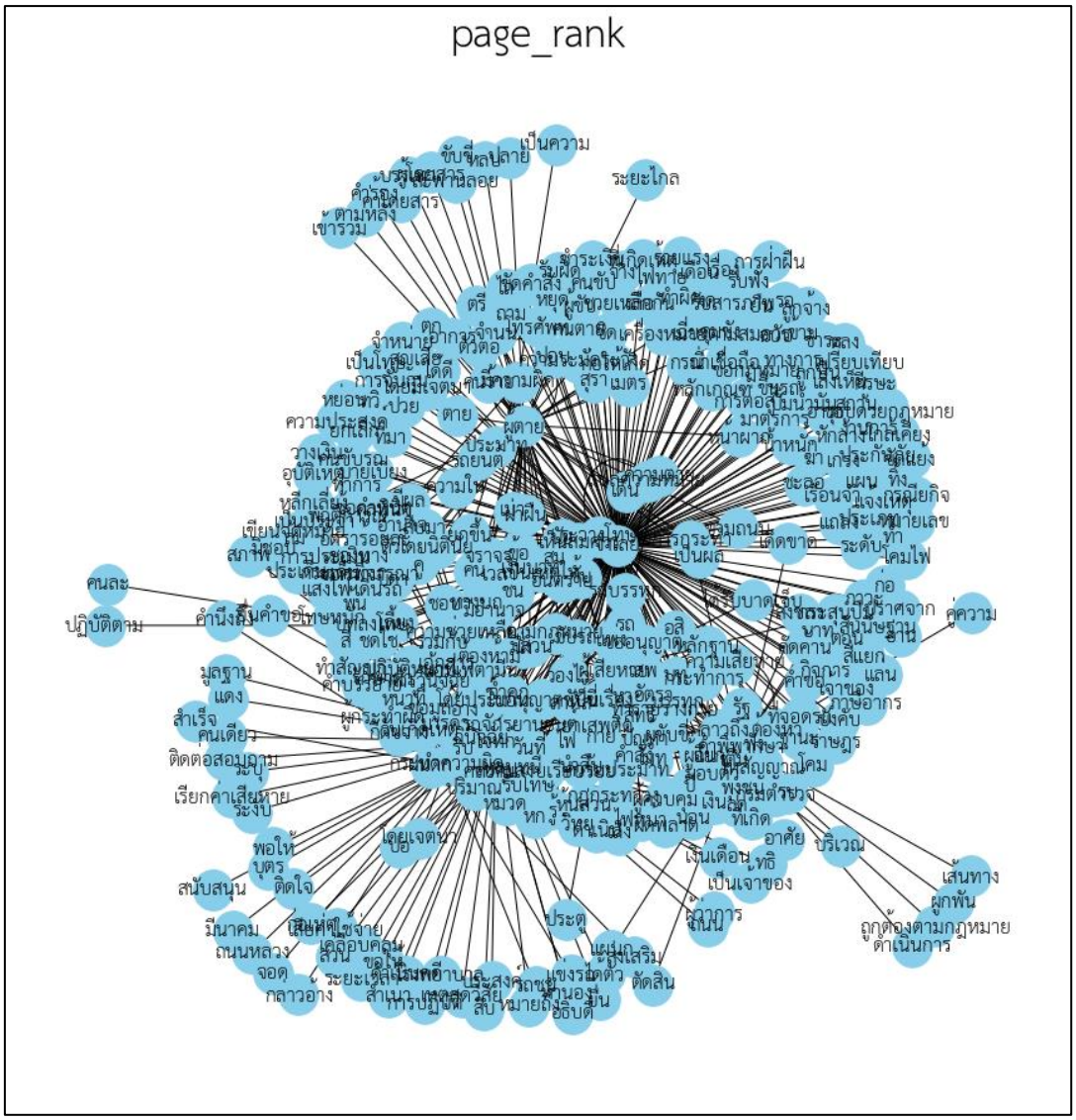
CHULALONGKORN UNIVERSITY

ภาพที่ 70 แสดงผลลัพธ์กราฟหลังจากผ่านการสกัดกิจกรรมสำคัญด้วยวิธีค่ากลางความใกล้



Chulalongkorn University

ภาพที่ 71 แสดงผลลัพธ์กราฟหลังจากผ่านการสกัดกิจกรรมสำคัญด้วยวิธีค่ากลางไอเกนเวกเตอร์



Chulalongkorn University

ภาพที่ 72 แสดงผลลัพธ์กราฟหลังจากผ่านการสกัดกิจกรรมสำคัญด้วยวิธีค่ากลางเพจแรงก์

ประวัติผู้เขียน

ชื่อ-สกุล	ว่าที่ ร้อยตำรวจโท ณภัทร งามสดใส
วัน เดือน ปี เกิด	19 กรกฎาคม 2538
สถานที่เกิด	กรุงเทพมหานคร
วุฒิการศึกษา	ปริญญาตรี วิทยาศาสตรบัณฑิต (วิทยาการคอมพิวเตอร์) มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
ผลงานตีพิมพ์	ผศ.ดร.สุกรี สินธุภิญโญ และ ร.ต.ท.ณภัทร งามสดใส “การสกัดกิจกรรมสำคัญจากคำพิพากษาศาลฎีกา เพื่อสร้างรายงานประจำวันสำหรับตำรวจ” วารสารเทคโนโลยีสารสนเทศ, ปีที่ 19, ฉบับที่ 1 เดือน มกราคม – มิถุนายน , 2566

จุฬาลงกรณ์มหาวิทยาลัย
Chulalongkorn University