

การปกปิดข้อมูลเสียงพูดเพื่อปกป้องข้อมูลที่ระบุตัวบุคคล

ณัฐธินา ชัยศิริพานิช¹ และ ประวิตรนันท์ บุตรโพธิ์²

¹คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง กรุงเทพฯ

²คณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง กรุงเทพฯ

Emails: 60070135@it.kmitl.ac.th, 60070148@it.kmitl.ac.th

บทคัดย่อ

ข้อมูลถือเป็นสิ่งสำคัญในการนำไปใช้ประโยชน์ให้กับองค์กร แต่การนำข้อมูลในองค์กรมาใช้มักมีข้อจำกัดในเรื่องข้อมูลส่วนบุคคล เช่น ข้อมูลลูกค้า เมื่อข้อมูลนั้นถูกนำไปวิเคราะห์ อาจส่งผลให้มีการลักลอบนำข้อมูลไปใช้ในทางที่ไม่ถูกต้อง ซึ่งข้อมูลจากการบันทึกเสียงบทสนทนาการทำธุรกรรมธนาคารทางโทรศัพท์ก็ถือว่ามีข้อมูลส่วนบุคคลเป็นจำนวนมาก จึงยังไม่สามารถนำข้อมูลเหล่านั้นมาใช้ประโยชน์ได้

ผู้จัดทำจึงดำเนินการปกปิดข้อมูลส่วนบุคคลจากไฟล์เสียงบทสนทนา โดยแปลงไฟล์เสียงพูดให้อยู่ในรูปแบบข้อความ จากนั้นวิเคราะห์คำที่เป็นข้อมูลส่วนบุคคลและระยะเวลาในไฟล์เสียง และแทนที่คำนั้นด้วยเสียงรบกวนในไฟล์เสียง เพื่อให้สามารถนำข้อมูลที่ผ่านมาการปกปิดข้อมูลส่วนบุคคลแล้วไปใช้ประโยชน์ในการวิเคราะห์ทางด้านอื่น ๆ ได้

คำสำคัญ – ข้อมูลส่วนบุคคล; ศูนย์บริการข้อมูลลูกค้าทางโทรศัพท์ (Call Center); การประมวลผลภาษาธรรมชาติ (Natural Language Processing); นิพจน์ระบุนาม (Named Entities); การแปลงเสียงพูดให้อยู่ในรูปแบบข้อความ (Speech-to-Text)

1. บทนำ

ปัจจุบันการละเมิดข้อมูลส่วนบุคคลนั้นเกิดขึ้นได้หลายรูปแบบ ซึ่งการละเมิดข้อมูลส่วนบุคคลจากการบันทึกบทสนทนาการทำธุรกรรมกับทางธนาคารก็ถือเป็นหนึ่งในปัญหาการละเมิดสิทธิส่วนบุคคลเช่นกัน ทางผู้จัดทำได้เล็งเห็นถึงความสำคัญของการรักษาข้อมูลส่วนบุคคลของลูกค้าในการทำธุรกรรมกับทางธนาคารผ่านศูนย์บริการข้อมูลลูกค้าทางโทรศัพท์โดยจะมีการทำการตรวจสอบการสนทนาบางส่วนในไฟล์บันทึกเสียง โดยเฉพาะส่วนที่เป็นข้อมูลส่วนบุคคลของลูกค้า เช่น ชื่อ - นามสกุล วันเกิด เบอร์โทรศัพท์ เลขที่บัญชี และเลขหน้าบัตรเครดิต หรือเดบิต ก่อนจะนำข้อมูลผลลัพธ์เหล่านั้นไปใช้ประโยชน์ในด้านอื่น ๆ โดยมีการทดลองสร้างระบบในการปกปิดข้อมูลที่ระบุตัวบุคคล ซึ่งภายในระบบจะดำเนินการแปลงไฟล์เสียงพูดให้อยู่ในรูปแบบข้อความ จากนั้นวิเคราะห์คำที่เป็นข้อมูลส่วนบุคคลพร้อมกับเก็บค่าของระยะเวลาที่พูดในไฟล์เสียงนั้น ระบบจะทำการแทรกเสียงรบกวนแทนที่คำพูดที่เป็นข้อมูลส่วนบุคคล และผลลัพธ์ที่ได้คือไฟล์เสียงที่ผ่านการปกปิดข้อมูลส่วนบุคคลแล้ว ซึ่งสามารถนำไปใช้ประโยชน์ในการวิเคราะห์ทางด้านอื่น ๆ ต่อไป

2. แนวคิด ทฤษฎี เทคโนโลยีและเครื่องมือที่เกี่ยวข้อง

ผู้จัดทำได้แบ่งรายละเอียดของแนวคิด ทฤษฎี เทคโนโลยีและเครื่องมือที่เกี่ยวข้องที่จะนำมาใช้กับการพัฒนาระบบเป็น 3 ส่วนหลัก ๆ ดังนี้

2.1 แนวคิดที่เกี่ยวข้อง

2.2.1 สิทธิความเป็นส่วนตัวส่วนบุคคล

มีการบัญญัติรับรองสิทธิดังกล่าวในรัฐธรรมนูญแห่งราชอาณาจักรไทย พ.ศ. 2560 มาตรา 32 บัญญัติไว้ว่า “สิทธิของบุคคลในครอบครัว เกียรติยศ ชื่อเสียง ตลอดจนความเป็นอยู่ส่วนบุคคล ย่อมได้รับความคุ้มครอง การกล่าวหรือไขข่าวแพร่หลายซึ่งข้อความหรือภาพไม่ว่าด้วยวิธีใดไปยังสาธารณชนอันเป็นการละเมิดหรือกระทบถึงสิทธิของบุคคลในครอบครัว เกียรติยศ ชื่อเสียง หรือความเป็นอยู่ส่วนบุคคล จะกระทำได้ เว้นแต่กรณีที่เป็น ประโยชน์ต่อสาธารณะ บุคคลย่อมมีสิทธิได้รับความคุ้มครองจากการแสวงประโยชน์โดยมิชอบจากข้อมูลส่วนบุคคลที่เกี่ยวกับตน ทั้งนี้ ตามที่กฎหมายบัญญัติ” [1]

2.2 ทฤษฎีที่เกี่ยวข้อง

2.2.1 การรู้จำเสียงพูด (Speech Recognition)

เป็นสิ่งที่ช่วยให้โปรแกรมสามารถประมวลผลคำพูดของมนุษย์ให้อยู่ในรูปแบบลายลักษณ์อักษร โดยเน้นที่การแปลงเสียงพูดจากรูปแบบคำพูดเป็นข้อความ [2]

2.2.2 การประมวลผลภาษาธรรมชาติ (Natural Language Processing: NLP)

เป็นศาสตร์ที่ศึกษาเกี่ยวกับการทำให้คอมพิวเตอร์สามารถสื่อสารโต้ตอบด้วยภาษาของมนุษย์ และทำให้คอมพิวเตอร์เข้าใจภาษามนุษย์มากขึ้น เช่น Siri, Google Assistant และ Alexa [3]

2.2.3 การรู้จำนิพจน์ระบุนาม (Named Entity Recognition)

การรู้จำนิพจน์ระบุนาม (Named Entity Recognition) เป็นขั้นตอนที่สำคัญสำหรับการพัฒนาระบบประมวลผลเอกสาร โดยเฉพาะอย่างยิ่งสำหรับระบบที่เกี่ยวข้องกับการเข้าถึงข้อมูล เช่น ระบบสกัดสารสนเทศ (Information Extraction) หรือในระบบค้นคืนเอกสาร (Information Retrieval) [4] ผู้จัดทำได้นำทฤษฎีนี้มาใช้ในการพัฒนาระบบส่วนของการตรวจจับคำที่เป็นข้อมูลส่วนบุคคล

2.2.4 Jaccard's Coefficient Similarity

เป็นสถิติประยุกต์แนวคิดในทฤษฎีเซตเพื่อใช้เปรียบเทียบความคล้ายคลึงและความหลากหลายของกลุ่มตัวอย่าง แนวคิดของค่าสัมประสิทธิ์ Jaccard's Coefficient Similarity คือ การวัดค่าความคล้ายคลึงระหว่างกลุ่มประชากร 2 กลุ่ม โดยคำนวณจากขนาดของประชากรที่ทั้งสองกลุ่มมีตัวอย่างร่วมกัน [5] ดังสมการที่ 1

$$Jaccard(X, Y) = \frac{|X \cap Y|}{|X \cup Y|} \quad (1)$$

2.2 เทคโนโลยีและเครื่องมือที่เกี่ยวข้อง

2.2.1 Cloud Speech to Text by Google Cloud

กูเกิลคลาวด์แพลตฟอร์มเป็นเว็บเซอร์ฟเวอร์ที่ให้บริการคลาวด์ ซึ่งภายในกูเกิลคลาวด์แพลตฟอร์มนั้นมีบริการแยกย่อยอีก เช่น Cloud Speech to Text, Cloud Storage, Compute Engine, และ Machine Learning เป็นต้น ทั้งนี้การใช้งานกูเกิลคลาวด์แพลตฟอร์มจะคิดค่าใช้จ่ายตามจำนวนการใช้งาน

ทางผู้จัดทำเลือกบริการ Cloud Storage ในการเก็บไฟล์เสียง และใช้ Cloud Speech to Text ไลบรารี Speech ในการแปลงเสียงพูดให้อยู่ในรูปแบบข้อความ [6]

2.2.2 Stanford Named Entity Recognizer (Stanford NER)

เป็นการประยุกต์ใช้จากภาษาจาวาสำหรับการรู้จำนิพจน์ระบุนาม (Named Entity Recognizer) ซึ่งเป็นการจัดประเภทของคำในข้อความ เช่น ชื่อสิ่งของ ชื่อบุคคล และบริษัท เป็นการกำหนดโครงสร้างการสกัดคุณสมบัติที่เหมาะสมสำหรับการรู้จำนิพจน์ระบุนาม [7]

2.2.3 Natural Language Toolkit (NLTK)

เป็นแพลตฟอร์มที่นิยมในโปรแกรมภาษาไพทอน เพื่อทำงานกับข้อมูลภาษาของมนุษย์ พร้อมกับชุดของไลบรารีที่ช่วยในการประมวลข้อความ แบ่งประเภทของคำ (Classification) การแบ่งโทเคนของคำ (Tokenization) การตัดคำ (Stemming) การติดแท็กคำ (Tagging) และการแยกวิเคราะห์คำ (Parsing) [8]

2.2.4 spaCy

เป็นไลบรารีสำหรับการทำการประมวลผลภาษาธรรมชาติขั้นสูงในภาษาไพทอน โดยที่ spaCy จะทำความเข้าใจข้อความจำนวนมากสามารถใช้ในการดำเนินการสกัดข้อมูล (Information Extraction) หรือระบบการทำความเข้าใจภาษาธรรมชาติเพื่อดำเนินการประมวลผลข้อความล่วงหน้าสำหรับการเรียนรู้เชิงลึก (Deep Learning) [9]

2.2.5 Regular Expressions

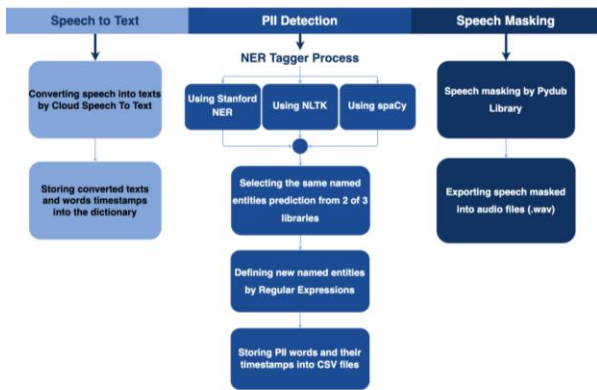
เป็นสัญลักษณ์ที่ใช้ระบุชุดของอักขระตัวอักษร เมื่อชุดของอักขระตัวอักษรที่เฉพาะเจาะจงนั้นอยู่ในชุดอักขระตัวอักษรที่มีการกำหนดให้เป็น Regular Expressions โดยทั่วไปแล้วจะใช้สัญลักษณ์ “*”, “+”, “?”, “()” และ “|” ในการกำหนดเงื่อนไขของชุดตัวอักษร [10]

2.2.6 Pydub

เป็นไลบรารีหนึ่งของภาษาไพทอนที่ใช้ในการจัดการกับข้อมูลที่เป็นไฟล์เสียง

3. ขั้นตอนการพัฒนาระบบ

ผู้จัดทำได้นำเทคโนโลยีและเครื่องมือมาประยุกต์ใช้โดยมีการดำเนินงานตามแผนงานการปกป้องข้อมูลทีละบุคคล ดังรูปที่ 1



รูปที่ 1. แผนงานการปกป้องข้อมูลที่ระบุตัวบุคคล

การดำเนินงานประกอบด้วย 3 ส่วนหลัก ดังรูปที่ 3.1 มีรายละเอียดการดำเนินงาน ดังนี้

3.1 การแปลงเสียงพูดให้อยู่ในรูปแบบข้อความ

3.1.1 ประยุกต์ใช้ Cloud Speech to Text

สร้างโปรเจกต์บนกูเกิลคลาวด์ และเปิดใช้งาน API 2 ประเภท คือ Cloud Storage และ Cloud Speech to Text จากนั้นอัปโหลดไฟล์เสียงขึ้นบน Cloud Storage ดังรูปที่ 2

NAME	SIZE	TYPE	CREATED TIME	STORAGE CLASS	LAST MODIFIED	PUBLIC ACCESS	ENCRYPTION	RETENTION EXPIRATION DATE
Amanda	3 MB	audio/mpeg	New 10, 2020...	Standard	New 10, 20...	Not public	Google-managed key	—
Calvin	2.8 MB	audio/mpeg	New 10, 2020...	Standard	New 10, 20...	Not public	Google-managed key	—
Calvin	1.3 MB	audio/mpeg	New 10, 2020...	Standard	New 10, 20...	Not public	Google-managed key	—
Calvin	4 MB	audio/mpeg	New 10, 2020...	Standard	New 10, 20...	Not public	Google-managed key	—
Michael	1.3 MB	audio/mpeg	New 10, 2020...	Standard	New 10, 20...	Not public	Google-managed key	—
Nancy	3.2 MB	audio/mpeg	New 27, 2020...	Standard	New 27, 20...	Not public	Google-managed key	—
Nelson	2.7 MB	audio/mpeg	New 10, 2020...	Standard	New 10, 20...	Not public	Google-managed key	—
Robert	3.4 MB	audio/mpeg	New 10, 2020...	Standard	New 10, 20...	Not public	Google-managed key	—
Robert	2.4 MB	audio/mpeg	New 10, 2020...	Standard	New 10, 20...	Not public	Google-managed key	—
Robert	2.8 MB	audio/mpeg	New 30, 2020...	Standard	New 30, 20...	Not public	Google-managed key	—
Robert	2 MB	audio/mpeg	New 30, 2020...	Standard	New 30, 20...	Not public	Google-managed key	—
Robert	2.9 MB	audio/mpeg	New 30, 2020...	Standard	New 30, 20...	Not public	Google-managed key	—
Robert	1.7 MB	audio/mpeg	New 30, 2020...	Standard	New 30, 20...	Not public	Google-managed key	—
Robert	2.7 MB	audio/mpeg	New 30, 2020...	Standard	New 30, 20...	Not public	Google-managed key	—
Robert	1.7 MB	audio/mpeg	New 30, 2020...	Standard	New 30, 20...	Not public	Google-managed key	—

รูปที่ 2. อัปโหลดไฟล์เสียงขึ้นบน Cloud Storage

ต่อมาทำการแปลงเสียงพูดให้อยู่ในรูปแบบข้อความโดยการสร้างฟังก์ชันต่าง ๆ ได้แก่ ฟังก์ชัน print_word_offsets ระบุเวลาที่คำในบทสนทนา (Timestamp) โดยระบุเวลาที่เริ่มต้นในแต่ละคำ และเวลาที่สิ้นสุดของคำ ๆ นั้น โดยหน่วยเวลาเป็นวินาทีเพื่อให้ง่ายต่อการเข้าใจ ฟังก์ชัน print_sentences แสดงผลประโยคที่ผ่านการแปลงเสียงพูดให้อยู่ในรูปแบบข้อความและแสดงผลค่าความเชื่อมั่น (Confidence) ส่วนสุดท้ายในขั้นตอนการแปลงเสียงพูดให้อยู่ในรูปแบบของข้อความในฟังก์ชัน speech_to_text ใช้โมดูลของไลบรารี Speech ในการแปลงข้อมูลเสียงพูดให้อยู่ในรูปแบบข้อความ และระบุไฟล์

```

data = []
start_all = []
end_all = []
word_all = []

def speech_to_text(config, audio):
    client = speech.SpeechClient()
    operation = client.long_running_recognize(config=config, audio=audio)
    response = operation.result(timeout=90)
    return print_sentences(response)

def print_sentences(response):
    for result in response.results:
        best_alternative = result.alternatives[0]
        transcript = best_alternative.transcript
        confidence = best_alternative.confidence
        data.append(transcript)
        print("-" * 80)
        print(f"Transcript: {transcript}")
        print(f"Confidence: {confidence:0%}")
        print_word_offsets(best_alternative)

def print_word_offsets(alternative):
    start_end_words = [], [], []
    for word in alternative.words:
        start_s = word.start_time.total_seconds()
        start.append(start_s)
        end_s = word.end_time.total_seconds()
        end.append(end_s)
        word = word.word
        words.append(word)
        print(f"{{start_s:>7.3f}} | {{end_s:>7.3f}} | {{word}}")

    start_all.append(start_s)
    end_all.append(end_s)
    word_all.append(word)
    return resultdict
speech_to_text(config, audio)
  
```

รูปที่ 3 ฟังก์ชันการแปลงเสียงพูดให้อยู่ในรูปแบบข้อความ

3.1.2 บันทึกไฟล์ให้อยู่ในรูปแบบ Dictionary

บันทึกเป็นประเภไฟล์ JSON เพื่อทำการตรวจจับคำที่เป็นข้อมูลส่วนบุคคลในขั้นตอนต่อไป

3.2 การตรวจจับคำที่เป็นข้อมูลส่วนบุคคลจากข้อมูลรูปแบบข้อความ

3.2.1 กระบวนการตรวจจับนิพจน์ระบุนาม

ขั้นตอนนี้มีการใช้ไลบรารีทั้งหมด 3 ไลบรารี เพื่อเพิ่มความแม่นยำในการตรวจจับนิพจน์ระบุนาม ได้แก่ Stanford NER, NLTK และ spaCy มีกระบวนการดำเนินงาน ดังนี้

- พัฒนาไลบรารี Stanford NER โดยเลือกประเภทของนิพจน์ระบุนามในการติดแท็กบทสนทนาทั้งหมดเป็นจำนวน 5 ประเภท ได้แก่ PERSON, ORGANIZATION, LOCATION, DATE และ MONEY ซึ่งในฟังก์ชันมีการทำ Word Tokenization เพื่อแยกโทเค็นของคำในข้อความ ต่อมาทำการติดแท็กนิพจน์ระบุนามจากอัลกอริทึมของ Stanford NER จากนั้นสร้างเงื่อนไขเก็บเฉพาะโทเค็นที่เป็นนิพจน์ระบุนามเท่านั้น จากนั้นจึงแก้ไขประเภทของนิพจน์ระบุนามที่ถูกติดแท็ก เพื่อให้ประเภทของนิพจน์ระบุนามตรงกับไลบรารีอื่น ๆ เช่น คำว่า “ORG” ที่ติดแท็กไว้ จะดำเนินการเปลี่ยนเป็นคำว่า “ORGANIZATION” จากนั้นทำการจับคู่โทเค็นที่ไลบรารีแบ่งออกมาเทียบกับโทเค็นที่ Cloud Speech to Text แบ่งไว้ เพื่อให้แน่ใจว่าโทเค็นที่ถูกติดแท็กนั้นตรงกับระยะเวลาที่ Cloud Speech to Text ทำนายออกมา และเก็บค่าของคำที่ติดแท็กได้ พร้อมกับประเภทของนิพจน์ระบุนาม

- พัฒนาไลบรารี NLTK ผู้จัดทำได้เลือกการติดแท็กในบทสนทนาเป็นจำนวนทั้งหมด 6 ประเภท ได้แก่ ORGANIZATION, PERSON,

LOCATION, GPE, DATE และ MONEY โดยเริ่มจากการทำ Word Tokenization จากนั้นทำการติดแท็กนิพจน์ระบุนามจาก อัลกอริทึม NLTK ซึ่งต้องมีการติดแท็กส่วนของประโยค (Part-of-Speech) ก่อนจึงจะติดแท็กได้ และกระบวนการหลังจากนั้นมีวิธีการทำเช่นเดียวกันกับ Stanford NER คือ เลือกโทเค็นที่เป็น นิพจน์ระบุนาม และทำการเปลี่ยนประเภทนิพจน์ระบุนามให้ เหมือนกันทุกไลบรารี จากนั้นจับคู่โทเค็นที่ไลบรารีแบ่งเทียบกับ โทเค็นของ Cloud Speech to Text และเก็บค่าของโทเค็น

- พัฒนาไลบรารี spaCy โดยผู้จัดทำได้เลือกการติดแท็กในบท สทนาเป็นจำนวนทั้งหมด 6 ประเภท ได้แก่ ORGANIZATION, PERSON, LOCATION, GPE, DATE และ MONEY ในฟังก์ชันมี การใช้อัลกอริทึมของ spaCy ซึ่งในอัลกอริทึมจะดำเนินการ วิเคราะห์ข้อความต่าง ๆ อัตโนมัติ ซึ่งสามารถเรียกดูค่าได้จาก อัลกอริทึมได้ทันที และกระบวนการหลังจากนั้นมีวิธีการทำ เช่นเดียวกันกับ Stanford NER และ NLTK คือ เลือกโทเค็นที่เป็น นิพจน์ระบุนาม และทำการเปลี่ยนประเภทนิพจน์ระบุนามให้ เหมือนกันทุกไลบรารี จากนั้นจับคู่โทเค็นที่แบบจำลองแบ่งเทียบกับ โทเค็นของ Cloud Speech to Text และเก็บค่าของโทเค็น

3.2.2 เลือกการทำนายประเภทของนิพจน์ระบุนามที่เหมือนกัน ตั้งแต่ 2 ใน 3 ของไลบรารี

เลือกการทำนายประเภทของนิพจน์ระบุนามที่เหมือนกันตั้งแต่ 2 จาก 3 ไลบรารีขึ้นไปจากการสร้างฟังก์ชันจับคู่โทเค็นที่มีการ ทำนายนิพจน์ระบุนามค่าเดียวกัน และเก็บค่าของโทเค็นนั้นใหม่ เพื่อนำไปใช้วิเคราะห์กระบวนการถัดไป ในที่นี้ ผู้จัดทำขอแทน ผลลัพธ์ของกระบวนการนี้ว่าค่าทำนายจริง

3.2.3 สร้างประเภทของนิพจน์ระบุนามเพิ่มเพื่อติดแท็กเลขที่เป็น ข้อมูลส่วนบุคคลโดยใช้ Regular Expressions

ขั้นตอนนี้มีการดึงโทเค็นค่าของ Cloud Speech to Text เฉพาะที่เป็นเลขมาตรวจสอบเงื่อนไขเพื่อติดแท็กเลขที่เป็นข้อมูลส่วนบุคคล เท่านั้น โดยแบ่งประเภทของเลขที่เป็นข้อมูลส่วนบุคคลไว้ 5 ประเภท คือ IDCARD (เลขบัตรประชาชน 13 หลัก) PHONENUM (เบอร์โทรศัพท์ 10 หลัก) ACCNUM (เลขบัญชี 9 หลัก) CARDNUM (เลขบัตรเครดิต หรือบัตรเครดิต 16 หลัก) และ PIINUM (เลขอื่น ๆ ที่ไม่เข้าเงื่อนไขประเภทก่อนหน้านี้ แต่มีตั้งแต่ 9 หลักขึ้นไป มีไว้ใน กรณีที่ Cloud Speech to Text แปลงเป็นข้อความออกมาได้ไม่ แม่นยำ) จากนั้นนำค่าที่ได้ไปรวมกับค่าทำนายจริง และเก็บค่า นั้นไว้ในรูปแบบไฟล์ CSV เพื่อนำไปดำเนินการต่อในขั้นถัดไป

3.2.4 เก็บค่าต่าง ๆ ให้อยู่ในรูปของไฟล์ CSV

หลังจากทำนายนิพจน์ระบุนามทั้งหมดแล้ว จึงจัดเก็บค่าเหล่านั้น ให้อยู่ในรูปแบบตารางและบันทึกเป็นไฟล์ CSV โดยมีจำนวน ทั้งหมด 5 คอลัมน์ ได้แก่ ลำดับโทเค็น โทเค็นค่า เวลาที่เริ่มพูด โทเค็นนั้นในไฟล์เสียง เวลาที่พูดโทเค็นนั้นจบ และประเภทของ นิพจน์ระบุนาม

3.3 การแทนที่เสียงที่เป็นข้อมูลส่วนบุคคลด้วยเสียงรบกวน

นำข้อมูลไฟล์ CSV มาใช้ในการดำเนินงาน คือ นำคอลัมน์ของเวลา เริ่มต้นของคำพูดและเวลาสิ้นสุดมาใช้ โดยแปลงให้ค่าของเวลาอยู่ใน หน่วยของมิลลิวินาที จากนั้นแทนที่เสียงรบกวนในช่วงเวลาที่ ค่ารวมไว้ จากนั้นดำเนินการบันทึกไฟล์เสียงที่มีการแทนที่ข้อมูล ส่วนบุคคลแล้วเป็นไฟล์ประเภท .wav เพื่อที่จะสามารถนำข้อมูลไป วิเคราะห์ประโยชน์ในด้านอื่น ๆ ต่อได้

3.5 การประเมินผล (Evaluation)

มีการประเมินผล 2 กระบวนการหลัก ๆ คือ ประเมินผลความ แม่นยำของการแปลงเสียงพูดให้อยู่ในรูปแบบข้อความโดยนำ แนวคิดของ Jaccard's Coefficient Similarity มาประยุกต์ใช้ในการ ประเมินผล และกระบวนการประเมินผลความแม่นยำของการ ตรวจจับคำที่เป็นข้อมูลส่วนบุคคลจากข้อมูลรูปแบบข้อความ จาก การสร้างผลเฉลยของการทำนายข้อความและโทเค็นต่าง ๆ เพื่อใช้ ตรวจสอบความแม่นยำในการทำนายของไลบรารีทั้งหมดโดยวัด จากค่า Accuracy

4. การทดลองและผลลัพธ์ที่ได้จากระบบ

ผู้จัดทำได้แบ่งการแสดงผลการทดลองและผลลัพธ์ที่ได้เป็น 4 หัวข้อหลัก มีรายละเอียด ดังนี้

4.1 การวิเคราะห์ข้อมูลที่สร้างขึ้น

ในการเตรียมข้อมูลเกิดปัญหาที่ไม่สามารถหาชุดข้อมูล จากแหล่งข้อมูลสาธารณะมาพัฒนาระบบได้เนื่องจากข้อมูลเหล่านี้ มีข้อจำกัดในเรื่องของข้อมูลส่วนบุคคล จึงต้องสร้างชุดข้อมูลขึ้นมา เอง โดยมีรายละเอียดการสร้างข้อมูล ดังนี้

4.1.1 สร้างบทสนทนาระหว่างลูกค้าและศูนย์บริการข้อมูลลูกค้า ทางโทรศัพท์

สร้างชุดข้อมูลในรูปแบบข้อความเป็นจำนวนทั้งหมด 23 บท สทนาเพื่อใช้ในการพัฒนาและประเมินผลระบบ จากการ วิเคราะห์ประโยคในบทสนทนาคิดเป็น 566 ประโยค มีจำนวน เฉลี่ยทั้งหมด 24.61 ประโยค แบ่งย่อยลงไปเป็นการวิเคราะห์คำที่ ยังไม่ผ่านการทำความสะอาดข้อมูลมีทั้งหมด 4,095 คำ โดยที่ใน 1 บทสนทนาจะมีค่าเป็นจำนวนเฉลี่ย 178.04 คำ และหากวิเคราะห์

คำผ่านการทำความสะอาดข้อมูลแล้ว กล่าวคือ ดำเนินการตัดเครื่องหมายวรรคตอนและ Stop words บางส่วนออก มีทั้งหมด 1732 คำ ค่าเฉลี่ยใน 1 บทสนทนาจะมีค่าโดยเฉลี่ยจำนวน 75.3 คำ

4.1.2 นำข้อมูลในรูปแบบข้อความมาดำเนินการบันทึกเสียง

เนื่องจากบทสนทนาเป็นบทสนทนาภาษาอังกฤษ จึงต้องนำบทสนทนาไปบันทึกเสียงโดยใช้ระบบสั่งการด้วยเสียงของระบบปฏิบัติการ iOS หรือ “สิริ” (Siri) ในการช่วยอ่านบทสนทนาเหล่านั้น ใน 1 บทสนทนาจะประกอบไปด้วยเสียงของพนักงานและลูกค้า โดยที่เสียงของพนักงานจะมีเพียงเพศเดียว คือ เพศหญิง โดยใช้เสียงของ “Siri Female” และในส่วนเสียงลูกค้าจะแบ่งออกเป็น 2 เพศ ได้แก่ เพศหญิง ใช้เสียงของ “Samantha” และเพศชาย ใช้เสียงของ “Siri Male”

4.1.3 แปลงประเภทของไฟล์เสียงบทสนทนา

แปลงประเภทไฟล์เสียงจาก “.m4a” ให้อยู่ในประเภทไฟล์ “.wav” เพื่อสะดวกต่อการนำไปประมวลผลโดยได้ดำเนินการแปลงบนเว็บไซต์ที่ชื่อว่า “Convert MP4 to WAV” [11]

4.1 การแปลงเสียงพูดให้อยู่ในรูปแบบข้อความ

```
{ 'transcript': "Hello, you have called virtual bank. This is Linda speaking. How may I help you? Hi Linda. I was just at your bill branch and I think I left my debit card in the ATM machine. Okay. Do you have your debit card number? I don't know. Okay. Well, do you have the checking account number associated with the debit card, but I do have are you ready? I will give you what I have got 760-545-6789. Okay. That's +765-450-600-7089. Correct? What is your identification number? 774-589-6589 665 okay. I have +774-580-960-5896 65 and what is your name sir? It is Robert. Appel board. Okay. I have Robert Applebaum yet. And what is your date of birth Mr. Appelbaum, July 7th, 1974. Okay. July 7th, 1974. Yes, and your phone number. It is 610-265-1715. Okay, I have 610-265-1715. Yes. Okay, Mr. Appelbaum. I have just this pending your card. If it is in the machine, we will contact you as lift the suspension 00. Thank you, sir. Thank you u.", 'values': { 'start': [0.0, 0.4, 1.2, 1.3, 1.8, 2.2, 2.4, 3.2, 3.4, 3.8, 4.3, 5.3, 5.3, 5.5, 5.7, 6.2, 6.8, 7.2, 8.0, 8.2, 8.3, 8.7, 8.8, 9.0, 9.5, 9.8, 9.8, 10.0, 10.2, 10.4, 10.7, 11.1, 11.2, 11.6, 11.7, 11.8, 12.3, 13.1, 14.2, 14.2, 14.4, 14.6, 15.0, 15.1, 15.4, 16.4, 16.5, 16.7, 18.2, 18.9, 19.2, 19.3, 19.4, 19.6, 19.9, 20.5, 20.8, 21.1, 21.8, 21.9, 22.3, 22.4, 23.1, 23.3, 23.4, 23.6, 24.6, 24.8, 25.1, 25.9, 26.1, 26.2, 26.5, 26.6, 26.7, 26.8, 27.2, 30.6, 31.8, 32.7, 36.0, 37.1, 37.2, 37.3, 37.5, 38.1, 38.9, 39.0, 39.1, 39.2, 39.3, 39.4, 39.5, 39.6, 39.7, 39.8, 39.9, 40.0, 40.1, 40.2, 40.3, 40.4, 40.5, 40.6, 40.7, 40.8, 40.9, 41.0, 41.1, 41.2, 41.3, 41.4, 41.5, 41.6, 41.7, 41.8, 41.9, 42.0, 42.1, 42.2, 42.3, 42.4, 42.5, 42.6, 42.7, 42.8, 42.9, 43.0, 43.1, 43.2, 43.3, 43.4, 43.5, 43.6, 43.7, 43.8, 43.9, 44.0, 44.1, 44.2, 44.3, 44.4, 44.5, 44.6, 44.7, 44.8, 44.9, 45.0, 45.1, 45.2, 45.3, 45.4, 45.5, 45.6, 45.7, 45.8, 45.9, 46.0, 46.1, 46.2, 46.3, 46.4, 46.5, 46.6, 46.7, 46.8, 46.9, 47.0, 47.1, 47.2, 47.3, 47.4, 47.5, 47.6, 47.7, 47.8, 47.9, 48.0, 48.1, 48.2, 48.3, 48.4, 48.5, 48.6, 48.7, 48.8, 48.9, 49.0, 49.1, 49.2, 49.3, 49.4, 49.5, 49.6, 49.7, 49.8, 49.9, 50.0, 50.1, 50.2, 50.3, 50.4, 50.5, 50.6, 50.7, 50.8, 50.9, 51.0, 51.1, 51.2, 51.3, 51.4, 51.5, 51.6, 51.7, 51.8, 51.9, 52.0, 52.1, 52.2, 52.3, 52.4, 52.5, 52.6, 52.7, 52.8, 52.9, 53.0, 53.1, 53.2, 53.3, 53.4, 53.5, 53.6, 53.7, 53.8, 53.9, 54.0, 54.1, 54.2, 54.3, 54.4, 54.5, 54.6, 54.7, 54.8, 54.9, 55.0, 55.1, 55.2, 55.3, 55.4, 55.5, 55.6, 55.7, 55.8, 55.9, 56.0, 56.1, 56.2, 56.3, 56.4, 56.5, 56.6, 56.7, 56.8, 56.9, 57.0, 57.1, 57.2, 57.3, 57.4, 57.5, 57.6, 57.7, 57.8, 57.9, 58.0, 58.1, 58.2, 58.3, 58.4, 58.5, 58.6, 58.7, 58.8, 58.9, 59.0, 59.1, 59.2, 59.3, 59.4, 59.5, 59.6, 59.7, 59.8, 59.9, 60.0, 60.1, 60.2, 60.3, 60.4, 60.5, 60.6, 60.7, 60.8, 60.9, 61.0, 61.1, 61.2, 61.3, 61.4, 61.5, 61.6, 61.7, 61.8, 61.9, 62.0, 62.1, 62.2, 62.3, 62.4, 62.5, 62.6, 62.7, 62.8, 62.9, 63.0, 63.1, 63.2, 63.3, 63.4, 63.5, 63.6, 63.7, 63.8, 63.9, 64.0, 64.1, 64.2, 64.3, 64.4, 64.5, 64.6, 64.7, 64.8, 64.9, 65.0, 65.1, 65.2, 65.3, 65.4, 65.5, 65.6, 65.7, 65.8, 65.9, 66.0, 66.1, 66.2, 66.3, 66.4, 66.5, 66.6, 66.7, 66.8, 66.9, 67.0, 67.1, 67.2, 67.3, 67.4, 67.5, 67.6, 67.7, 67.8, 67.9, 68.0, 68.1, 68.2, 68.3, 68.4, 68.5, 68.6, 68.7, 68.8, 68.9, 69.0, 69.1, 69.2, 69.3, 69.4, 69.5, 69.6, 69.7, 69.8, 69.9, 70.0, 70.1, 70.2, 70.3, 70.4, 70.5, 70.6, 70.7, 70.8, 70.9, 71.0, 71.1, 71.2, 71.3, 71.4, 71.5, 71.6, 71.7, 71.8, 71.9, 72.0, 72.1, 72.2, 72.3, 72.4, 72.5, 72.6, 72.7, 72.8, 72.9, 73.0, 73.1, 73.2, 73.3, 73.4, 73.5, 73.6, 73.7, 73.8, 73.9, 74.0, 74.1, 74.2, 74.3, 74.4, 74.5, 74.6, 74.7, 74.8, 74.9, 75.0, 75.1, 75.2, 75.3, 75.4, 75.5, 75.6, 75.7, 75.8, 75.9, 76.0, 76.1, 76.2, 76.3, 76.4, 76.5, 76.6, 76.7, 76.8, 76.9, 77.0, 77.1, 77.2, 77.3, 77.4, 77.5, 77.6, 77.7, 77.8, 77.9, 78.0, 78.1, 78.2, 78.3, 78.4, 78.5, 78.6, 78.7, 78.8, 78.9, 79.0, 79.1, 79.2, 79.3, 79.4, 79.5, 79.6, 79.7, 79.8, 79.9, 80.0, 80.1, 80.2, 80.3, 80.4, 80.5, 80.6, 80.7, 80.8, 80.9, 81.0, 81.1, 81.2, 81.3, 81.4, 81.5, 81.6, 81.7, 81.8, 81.9, 82.0, 82.1, 82.2, 82.3, 82.4, 82.5, 82.6, 82.7, 82.8, 82.9, 83.0, 83.1, 83.2, 83.3, 83.4, 83.5, 83.6, 83.7, 83.8, 83.9, 84.0, 84.1, 84.2, 84.3, 84.4, 84.5, 84.6, 84.7, 84.8, 84.9, 85.0, 85.1, 85.2, 85.3, 85.4, 85.5, 85.6, 85.7, 85.8, 85.9, 86.0, 86.1, 86.2, 86.3, 86.4, 86.5, 86.6, 86.7, 86.8, 86.9, 87.0, 87.1, 87.2, 87.3, 87.4, 87.5, 87.6, 87.7, 87.8, 87.9, 88.0, 88.1, 88.2, 88.3, 88.4, 88.5, 88.6, 88.7, 88.8, 88.9, 89.0, 89.1, 89.2, 89.3, 89.4, 89.5, 89.6, 89.7, 89.8, 89.9, 90.0, 90.1, 90.2, 90.3, 90.4, 90.5, 90.6, 90.7, 90.8, 90.9, 91.0, 91.1, 91.2, 91.3, 91.4, 91.5, 91.6, 91.7, 91.8, 91.9, 92.0, 92.1, 92.2, 92.3, 92.4, 92.5, 92.6, 92.7, 92.8, 92.9, 93.0, 93.1, 93.2, 93.3, 93.4, 93.5, 93.6, 93.7, 93.8, 93.9, 94.0, 94.1, 94.2, 94.3, 94.4, 94.5, 94.6, 94.7, 94.8, 94.9, 95.0, 95.1, 95.2, 95.3, 95.4, 95.5, 95.6, 95.7, 95.8, 95.9, 96.0, 96.1, 96.2, 96.3, 96.4, 96.5, 96.6, 96.7, 96.8, 96.9, 97.0, 97.1, 97.2, 97.3, 97.4, 97.5, 97.6, 97.7, 97.8, 97.9, 98.0, 98.1, 98.2, 98.3, 98.4, 98.5, 98.6, 98.7, 98.8, 98.9, 99.0, 99.1, 99.2, 99.3, 99.4, 99.5, 99.6, 99.7, 99.8, 99.9, 100.0] } }
```

รูปที่ 4. ตัวอย่างการแปลงข้อมูลเสียงให้อยู่ในรูปแบบข้อความโดยใช้ Cloud Speech to Text

จากรูปที่ 4 แปลงให้อยู่ในรูปแบบของ Dictionary และสร้างคีย์ที่ชื่อว่า transcript เพื่อเก็บข้อความในบทสนทนาทั้งหมด ในส่วนของโทเค็นคำ ได้มีการสร้างคีย์ที่ชื่อว่า values เพื่อเก็บค่าของเวลาที่เริ่มพูดโทเค็นนั้น ๆ (start) เวลาที่พูดจบ (end) และโทเค็นนั้น ๆ (word) และประเมินผลความแม่นยำในการทำนายของแบบจำลองโดยการนำข้อมูลบทสนทนาจริงเทียบกับข้อมูลที่แบบจำลองทำนายโดยใช้ Jaccard's Coefficient Similarity ดังนี้

```
"Hello, you have called virtual bank, this is Linda speaking. How may I help you? Hi Linda. I was just at your Vill e branch and I think I left my Debit c ard in the ATM machine. Okay. Do you h ave your Debit card number? I don't ha ve. Okay, well do you have the checkin g account number associated with the D ebit card? That I do have. Are you rea dy? I will give you what I have got. 7 65-456-789. Okay. That's 765-456-789. Correct. What is your identification n umber? 774-589-658-9665. Okay, I have 774-589-658-9665 and what is your name sir? It is Robert Applebaum. Okay. I h ave Robert Applebaum. Yes. And what is your date of birth Mr. Applebaum? July 7th, 1974. Okay. July 7th, 1974. Yes. And your phone number? It is 610-265-1 715. Okay. I have 610-2651715. Yes. Ok ay Mr. Applebaum. I have just suspende d your card. If it is in the machine, we will contact you and lift the suspe nsion. Oh, thank you, Sure. Thank yo u."
```

รูปที่ 5. ข้อมูลบทสนทนาจริง

```
"Hello, you have called virtual bank. This is Linda speaking. How may I help you? Hi Linda. I was just at your bill branch and I think I left my debit card in the ATM machine. Okay. Do you have your debit card number? I don't k now. Okay. Well, do you have the checking acc ount number associated with the debit card, b ut I do have are you ready? I will give you w hat I have got 760-545-6789. Okay. That's +76 5-450-600-7089. Correct? What is your identif ication number? 774-589-6589 665 okay. I have +774-580-960-5896 65 and what is your name si r? It is Robert. Appel board. Okay. I have Rob ert Applebaum yet. And what is your date of b irth Mr. Appelbaum, July 7th, 1974. Okay, Jul y 7th, 1974. Yes, and your phone number. It i s 610-265-1715. Okay, I have 610-265-1715. Ye s. Okay, Mr. Appelbaum. I have just this pend ed your card. If it is in the machine, we wil l contact you as lift the suspension 00. Than k you, sir. Thank you."
```

รูปที่ 6. บทสนทนาที่แบบจำลองทำนาย

ตารางที่ 1. ตารางแสดงค่าความแม่นยำของการแปลงเสียงพูดให้อยู่ในรูปของข้อความ

Algorithm	Accuracy (%)
Cloud Speech to Text	57.02

จากตารางที่ 1 ความแม่นยำในการทำนายคิดเป็นร้อยละ 57.02 ซึ่งเมื่อเทียบบทสนทนาที่ 5 และ 6 พบว่าสิ่งที่ส่งผลให้ค่าความแม่นยำของแบบจำลองไม่สูงนั้นส่วนใหญ่ขึ้นอยู่กับเครื่องหมายวรรคตอนของข้อมูลบทสนทนาทั้งสอง ดังนั้น จึงสร้างฟังก์ชันตัดเครื่องหมายวรรคตอนของบทสนทนาทั้งสองออก เพื่อประเมินผลค่าความแม่นยำใหม่ ดังรูปที่ 7 และ 8 และ ตารางที่ 2

```
'Hello you have called virtual bank thi s is Linda speaking How may I help yo u? Hi Linda I was just at your Ville b ranch and I think I left my Debit card in the ATM machine Okay Do you have y our Debit card number? I dont have Oka y well do you have the checking account number associated with the Debit card? That I do have Are you ready? I will g ive you what I have got 765456789 Oka y Thats 765456789 Correct What is yo ur identification number? 7745896589665 Okay I have 7745896589665 and what is y our name sir? It is Robert Applebaum O kay I have Robert Applebaum Yes And what is your date of birth Mr Applebau m? July 7th 1974 Okay July 7th 1974 Yes And your phone number? It is 61026 51715 Okay I have 6102651715 Yes Ok ay Mr Applebaum I have just suspended your card If it is in the machine we w ill contact you and lift the suspension 00 thank you Sure Thank you '
```

รูปที่ 7. ข้อมูลบทสนทนาจริงที่ผ่านการทำความสะอาด

'Hello you have called virtual bank Th is is Linda speaking How may I help yo u? Hi Linda I was just at your bill br anch and I think I left my debit card i n the ATM machine Okay Do you have yo ur debit card number? I dont know Okay Well do you have the checking account n umber associated with the debit card bu t I do have are you ready? I will give you what I have got 7605456789 Okay T hats 765450607089 Correct? What is yo ur identification number? 7745896589 66 5 okay I have 774589605896 65 and wha t is your name sir? It is Robert Appel board Okay I have Robert Applebaum yet And what is your date of birth Mr Appe lbaum July 7th 1974 Okay July 7th 1974 Yes and your phone number It is 610265 1715 Okay I have 6102651715 Yes Okay Mr Appelbaum I have just this pended your card If it is in the machine we w ill contact you as lift the suspension 00 Thank you sir Thank you '

รูปที่ 8. บทสนทนาที่แบบจำลองทำนาย ที่ผ่านการทำความสะอาด

ตารางที่ 2. ตารางแสดงค่าความแม่นยำของการแปลงเสียงพูดให้อยู่ในรูปของข้อความ (ใหม่)

Algorithm	Accuracy (%)
Cloud Speech to Text	71.43

จากตารางที่ 2 ความแม่นยำในการทำนายคำพูดของแบบจำลอง คิดเป็นร้อยละ 71.43 สามารถเห็นได้ชัดว่าค่าความแม่นยำสูงขึ้น อย่างชัดเจน เมื่อตัดเครื่องหมายวรรคตอนออกเบื้องต้น

4.3 การตรวจจับคำที่เป็นข้อมูลส่วนบุคคลจากข้อมูลรูปแบบข้อความ

เมื่อนำข้อมูลในรูปแบบข้อความที่ได้จาก Cloud Speech to Text เข้าสู่ฟังก์ชันต่าง ๆ ของไลบรารี Stanford NER, NLTK และ spaCy จากนั้นนำเข้าสู่ฟังก์ชันของการเลือกคำทำนายจริง และสร้างนิพจน์ระบุนามเพิ่มเติมสำหรับเลขที่เป็นข้อมูลส่วนบุคคลโดยใช้ Regular Expressions ดังที่ได้กล่าวไว้ในบทขั้นตอนการพัฒนา ระบบ จึงเก็บค่าของการทำนายของทุก ๆ ไลบรารีไว้ในรูปแบบ ตาราง ดังรูปที่ 9

	word	start_time	end_time	stanford_pred	nltk_pred	spacy_pred	real_ents
indx							
0	Hello,	0.0	0.4	DATE	LOCATION	O	O
1	you	0.4	1.2	O	O	O	O
2	have	1.2	1.3	O	O	O	O
3	called	1.3	1.8	O	O	O	O
4	virtual	1.8	2.2	O	O	O	O
5	bank.	2.2	2.4	O	O	O	O
6	This	2.4	3.2	O	O	O	O
7	is	3.2	3.4	O	O	O	O
8	Linda	3.4	3.8	PERSON	PERSON	PERSON	PERSON
9	speaking.	3.8	4.3	O	O	O	O

รูปที่ 9. ตารางการทำนายประเภทของนิพจน์ระบุนาม

จากรูปที่ 9 เก็บค่าการทำนายของโทเค็นไว้ในตารางเดียวกันตามประเภทของนิพจน์ระบุนาม แลวดังได้มีการทำนายเป็นคำว่า “O” หมายความว่าโทเค็นนั้นไม่ได้เป็นนิพจน์ระบุนาม และมีการเก็บค่าการทำนายทั้งหมด 4 คอลัมน์ ได้แก่ stanford_pred (ค่าที่แบบจำลอง Stanford NER ทำนาย) nltk_pred (ค่าที่ NLTK

ทำนาย) spacy_pred (ค่าที่ spaCy ทำนาย) และ real_ents (ค่าทำนายจริง) นอกจากนี้ ยังได้เก็บบันทึกค่าการทำนายจริง เฉพาะโทเค็นที่มีการติดแท็กนิพจน์ระบุนามขึ้นอีก 1 ตาราง เพื่อบันทึกให้อยู่ในรูปแบบไฟล์ CSV และนำไปปกปิดเสียงในขั้นตอนถัดไป ดังรูปที่ 10

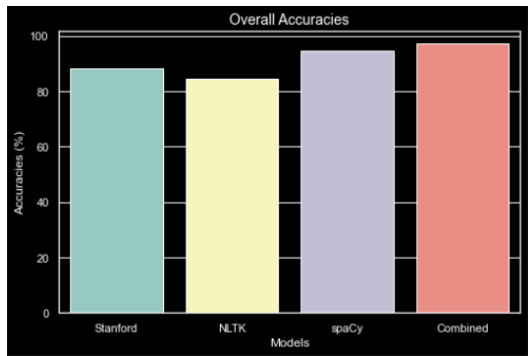
	word	start_time	end_time	real_ents
indx				
8	Linda	3.4	3.8	PERSON
16	Linda.	6.8	7.2	PERSON
34	ATM	11.7	11.8	ORGANIZATION
76	760-545-6789.	27.2	30.6	PHONENUM
79	+765-450-600-7089.	32.7	35.7	IDCARD
86	774-589-6589	38.9	42.7	PHONENUM
91	+774-580-960-5896	45.4	49.0	IDCARD
101	Robert.	51.9	52.3	PERSON
107	Robert	55.0	55.4	PERSON
108	Applebaum,	55.4	56.0	PERSON
118	Appelbaum,	59.8	60.3	PERSON
119	July	60.3	61.6	DATE
120	7th,	61.6	62.1	DATE
121	1974.	62.1	63.5	DATE
123	July	64.9	66.0	DATE
124	7th,	66.0	66.6	DATE
125	1974.	66.6	68.3	DATE
133	610-265-1715.	71.9	75.4	PHONENUM
137	610-265-1715.	77.4	80.7	PHONENUM
141	Appelbaum.	83.1	83.6	PERSON

รูปที่ 10. ตารางค่าทำนายจริงเฉพาะที่มีการติดแท็กนิพจน์ระบุนาม

มีการประเมินผลความแม่นยำในการทำนายนิพจน์ระบุนามของแต่ละไลบรารี ซึ่งวัดจากค่า Accuracies โดยนำข้อมูลโทเค็นที่ Cloud Speech to Text แบ่งมาทำการเจดนิพจน์ระบุนามจริง เพื่อที่จะนำไปประเมินผลความแม่นยำจากค่า Accuracies ของการทำนายนิพจน์ระบุนามในทุก ๆ ไลบรารี

ตารางที่ 3. ตารางการแสดงผลค่าความแม่นยำของการทำนายนิพจน์ระบุนามทุกประเภท

Libraries	Accuracies
Stanford NER	88.17
NLTK	84.62
spaCy	94.67
Combined and Regex	97.04



รูปที่ 11. กราฟการประเมินผลความแม่นยำของแต่ละไลบรารี

จากตารางที่ 3 สามารถสรุปได้ ดังนี้

- ความแม่นยำของการทำนายนิพจน์ระบุนามของ Stanford NER คิดเป็นร้อยละ 88.17
- ความแม่นยำของการทำนายนิพจน์ระบุนามของ NLTK คิดเป็นร้อยละ 84.62
- ความแม่นยำของการทำนายนิพจน์ระบุนามของ spaCy คิดเป็นร้อยละ 94.67
- ความแม่นยำของการทำนายนิพจน์ระบุนามของการรวมไลบรารีและการทำ Regular Expressions คิดเป็นร้อยละ 97.04

สังเกตได้ว่า เมื่อดำเนินการรวมการทำนายของแต่ละแบบจำลองเข้าด้วยกัน และสร้างเงื่อนไขจาก Regular Expressions นั้น ส่งผลให้ค่าความแม่นยำในการทำนายนิพจน์ระบุนามสูงที่สุด

4.3 การแทนที่เสียงที่เป็นข้อมูลส่วนบุคคลด้วยเสียงรบกวน

เมื่อทำการแทนที่เสียงแล้วจึงดำเนินการบันทึกไฟล์เสียงที่ผ่านการแทนที่เสียงที่เป็นข้อมูลส่วนบุคคลด้วยเสียงรบกวนเป็นไฟล์เสียงประเภท .wav

5. บทสรุป

5.1 สรุปผลโครงการ

การสร้างข้อมูลบทสนทนาในรูปแบบข้อความ เนื้อหาข้อมูลส่วนบุคคลของบทสนทนาประกอบด้วย ชื่อ - นามสกุล เลขที่บัญชี เลขบัตรเดบิต หรือ เครดิต เลขบัตรประชาชน วันเกิด ที่อยู่ และเบอร์โทรศัพท์ จากนั้นนำข้อมูลบทสนทนาในรูปแบบข้อความมาดำเนินการบันทึกเสียงและบันทึกเป็นไฟล์เสียง เพื่อนำไปใช้ในการพัฒนาระบบ

ส่วนของการแปลงเสียงพูดให้อยู่ในรูปแบบข้อความนั้น หากเป็นการประเมินผลโดยไม่คำนึงถึงความถูกต้องของเครื่องหมายวรรคตอน ถือว่าค่าของความแม่นยำอยู่ในระดับที่ดี อาจจะมีการแปลงชื่อบุคคลที่ไม่ตรงกับข้อมูลบทสนทนาจริงเล็กน้อย อาจเป็นสาเหตุมาจากเสียงที่ใช้ในการดำเนินการ

บันทึกเสียงที่แต่ละบุคคลมีสำเนียงการพูดที่ไม่เหมือนกัน เช่น นามสกุล Applebaum เมื่อเป็นเสียงของ Siri Male ทางแบบจำลองแปลงได้เป็น 2 โทเค็น คือ “Appel” และ “board.” แต่เมื่อเป็นเสียงของ “Siri Female” ทางแบบจำลองกลับแปลงคำได้ถูกต้อง จึงสรุปได้ว่าบางครั้งสำเนียงการพูดของแต่ละบุคคลอาจส่งผลต่อความแม่นยำของการแปลงข้อมูลเสียงให้อยู่ในรูปแบบข้อความ นอกจากนี้ ยังมีการแปลงเลขที่ผิดพลาดไปบ้าง เช่น เมื่อสิริพูดว่า “oh” ในบางครั้งแบบจำลองจะแปลงเป็นเลข “0” ซึ่งส่งผลให้ความแม่นยำของแบบจำลองลดลง

ขั้นตอนต่อมาเป็นการตรวจจับคำที่เป็นข้อมูลส่วนบุคคลจากข้อมูลรูปแบบข้อความ ได้ทำการทดลอง ทั้งหมด 3 ไลบรารี

- Stanford NER สามารถติดแท็กบุคคล และค่าเงินได้ค่อนข้างแม่นยำ แต่ในการติดแท็กวันที่ ด้วยข้อจำกัดที่ไม่มีการติดแท็กตัวเลขที่เป็นประเภท Cardinal จึงส่งผลให้มีการติดแท็กตัวเลขธรรมดา เป็นประเภทของวันที่ (Date) ทำให้ความแม่นยำลดลง
- NLTK สามารถติดแท็กองค์กรได้แม่นยำมากที่สุด แต่แบบจำลองนี้มักมีการติดแท็กที่ผิดพลาดตรงส่วนของสถานที่ กล่าวคือ หากโทเค็นนั้นขึ้นต้นด้วยตัวอักษรพิมพ์ใหญ่ เช่น คำว่า “Hello” แบบจำลองจะติดแท็กเป็นสถานที่ที่ทันที
- spaCy จากผลลัพธ์การประเมินผลความแม่นยำ สังเกตได้ว่าส่วนใหญ่แล้ว spaCy มีค่าความแม่นยำสูงในการติดแท็กโทเค็น แต่หากให้สรุปเป็นรายประเภท สรุปได้ว่าสามารถติดแท็กบุคคล สถานที่ วันที่ และค่าเงินได้ดีที่สุด แต่เนื่องจากการติดแท็กของไลบรารีนี้ยังมีความไม่แม่นยำอยู่บ้าง ทางผู้จัดทำจึงมีความเห็นว่าควรรวมไลบรารีเข้าด้วยกันเพื่อเพิ่มประสิทธิภาพในการติดแท็ก

ในส่วนของการรวมไลบรารีเข้าด้วยกันมีความแม่นยำค่อนข้างสูง ซึ่งเฉลี่ยแล้วคิดเป็นร้อยละ 90 ถือเป็นค่าความแม่นยำที่น่าพึงพอใจ

การตรวจจับเลขที่เป็นข้อมูลส่วนบุคคลโดยใช้ Regular Expressions ก็มีความแม่นยำค่อนข้างสูง แต่ในบางครั้งอาจไม่แม่นยำอย่างสมบูรณ์เนื่องจากรูปแบบการแปลงตัวเลขของ Cloud Speech to Text อาจแบ่งโทเค็นได้ไม่ตรงกับตัวเลขที่ควรจะเป็น เช่น เลขบัตรเดบิต หรือบัตรเครดิต 16 หลัก ทางแบบจำลองอาจมีรูปแบบการแปลงตัวเลขได้เพียงแค่ 13 หลัก แล้วจึงแบ่งเลขอีก 3 หลักหลังเป็นอีกโทเค็น ซึ่งในเงื่อนไขมักจะติดแท็กเลขที่มากกว่า 9 หลักขึ้นไปโดยไม่สนใจเครื่องหมายต่าง ๆ เช่น +111-111-111-1111 หรือ 111-111-1111 เป็นต้น แต่หากพิจารณาถึงภาพรวมของค่าความแม่นยำแล้ว ถือเป็นที่น่าพึงพอใจ

ขั้นตอนสุดท้ายเป็นการแทนที่เสียงที่เป็นข้อมูลส่วนบุคคลด้วยเสียงรบกวน ในบางช่วงที่มีการแทนที่เสียงรบกวนนั้น

อาจเกินหรือขาดจากระยะเวลาของคำพูดที่เป็นข้อมูลส่วนบุคคล แต่โดยภาพรวมแล้วถือว่าปิดบังคำพูดที่เป็นข้อมูลส่วนบุคคลได้ดี

5.2 ปัญหาในการพัฒนาโครงการ

- 1) ปัญหาในการเตรียมข้อมูล โครงการนี้ไม่สามารถหาชุดข้อมูลสาธารณะในการนำไปพัฒนาระบบได้เนื่องจากข้อมูลที่ใช้ดำเนินงานนั้นเป็นข้อมูลส่วนบุคคล ผู้จัดทำจึงจำเป็นต้องสร้างชุดข้อมูลขึ้นเองเพื่อพัฒนาระบบนี้
- 2) ปัญหาในการพัฒนาระบบหลัก ๆ คือ ความแม่นยำของการแปลงข้อมูลเสียงให้อยู่ในรูปแบบข้อความ หากบางบทสนทนามีความแม่นยำที่ไม่สูงมากพอ เมื่อนำข้อมูลเหล่านั้นเข้าสู่กระบวนการตรวจจับคำที่เป็นข้อมูลส่วนบุคคลจากข้อมูลรูปแบบข้อความ ส่งผลให้ไลบรารีไม่สามารถติดแท็กประเภทของโทเค็นที่ควรจะมีพจน์ระบุนามได้ เช่น ชื่อบุคคล หรือ ส่วนเล็ก ๆ ของเลขที่เป็นข้อมูลสำคัญ ส่งผลให้เป็นปัญหาต่อการปิดบังคำที่เป็นข้อมูลส่วนบุคคลในขั้นตอนสุดท้ายได้
- 3) การแทนที่เสียงที่เป็นข้อมูลส่วนบุคคลด้วยเสียงรบกวน ในบางช่วงของการแทนที่เสียงรบกวนอาจเกินหรือขาดจากระยะเวลาของคำพูดที่เป็นข้อมูลส่วนบุคคล

5.3 แนวทางในการพัฒนาต่อ

- 1) โครงการนี้สามารถต่อยอดโดยการนำระบบที่พัฒนาไปใช้ในด้านของการปกปิดข้อมูลที่เป็นส่วนบุคคลในหน่วยงานและองค์กรที่ต้องการประยุกต์ใช้ระบบได้ทั้งในรูปแบบไฟล์เสียง และข้อมูลที่เป็นข้อความ
- 2) สามารถนำผลลัพธ์ที่ได้จากการพัฒนาระบบไปวิเคราะห์และพัฒนาต่อในด้านอื่น ๆ ได้ โดยไม่ต้องคำนึงถึงสิทธิส่วนบุคคล เนื่องจากมีการปกปิดข้อมูลส่วนบุคคลแล้ว

เอกสารอ้างอิง

- [1] ศ. สวัสดิ์พงศ์ธาดา. “ความเป็นส่วนตัว (Privacy)”. [Online]. Available: <https://angsila.cs.buu.ac.th/~58160640/887420/hw/hw8.pdf>. 2015.
- [2] IBM. “What is Speech Recognition?”. [Online]. Available: <https://www.ibm.com/cloud/learn/speech-recognition>. 2020.
- [3] Flame Sillawat. “การเปิดใช้งาน Cloud Speech API”. [Online]. Available: <https://bit.ly/3orAjle>. 2018.

- [4] รัฐภูมิ ดันสุตะพานิช. “การสกัดความสัมพันธ์ระหว่างนิพจน์ระบุนามในภาษาไทย”. (วิทยานิพนธ์ปริญญาวิทยาศาสตรมหาบัณฑิต, 2552).
- [5] ศุภวัจน์ แด่รุ่งเรือง. “การตรวจเทียบภายนอกหาลักษณะการกลไกในงานวิชาการโดยใช้แบบจำลองซัพพอร์ทเวกเตอร์แมชชีนและการวัดค่าความคล้ายของข้อความ”. (วิทยานิพนธ์ปริญญาอักษรศาสตรดุษฎีบัณฑิต จุฬาลงกรณ์มหาวิทยาลัย, 2560).
- [6] A. Geitgey Natural Language Processing is Fun!”. [Online]. Available: <https://bit.ly/36Vpsk8>. 2018.
- [7] C. Dishmon. “Named Entity Recognition with Stanford NER Tagger”. [Online]. Available: <https://bit.ly/3LXHvKB>. 2020.
- [8] NLTK. “Natural Language Toolkit”. [Online]. Available: <https://www.nltk.org/>. 2020.
- [9] spaCy. “spaCy 101: Everything you need to know”. [Online]. Available: <https://spacy.io/usage/spacy-101>. 2020.
- [10] R. Cox. “Regular Expression Matching Can Be Simple And Fast (but is slow in Java, Perl, PHP, Python, Ruby, ...)”. [Online]. Available: <https://swtch.com/~rsc/regexp/regexp1.html>. 2007.
- [11] “MP4 to WAV online file converter”. [Online]. Available: <https://audio.online-convert.com/convert/mp4-to-wav>.