

การปกป้องข้อมูลที่ระบุตัวบุคคล

ณัฏฐณิชา 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 โดยทางผู้จัดทำจะดำเนินการแปลงไฟล์เสียงบทสนทนาให้อยู่ในรูปแบบข้อความ ตรวจสอบเนื้อหาของข้อความว่าคำใดมีรูปแบบที่เป็นข้อมูลส่วนบุคคล จากนั้นดำเนินการจับคู่คำกับเวลาในไฟล์บันทึกเสียง และดำเนินการปกปิดข้อความในส่วนนั้นออกไป

2. แนวคิด และเทคโนโลยีที่เกี่ยวข้อง

2.1 แนวคิดที่เกี่ยวข้อง

2.2.1 สิทธิความเป็นส่วนตัวส่วนบุคคล

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

2.2 เทคโนโลยีที่เกี่ยวข้อง

2.2.1 การรู้จำเสียงพูด (Speech Recognition)

เป็นสิ่งที่ช่วยให้โปรแกรมสามารถประมวลผลคำพูดของมนุษย์ให้อยู่ในรูปแบบลายลักษณ์อักษร โดยเน้นที่การแปลงเสียงพูดจากรูปแบบคำพูดเป็นข้อความ [2]

2.2.2 Cloud Speech to Text by Google Cloud

กูเกิลคลาวด์แพลตฟอร์มเป็นเว็บเซิร์ฟเวอร์ที่ให้บริการคลาวด์ ซึ่งภายในกูเกิลคลาวด์แพลตฟอร์มนั้นมีบริการที่แยกย่อยอีกมากมายให้ตรงตามลักษณะการใช้งาน เช่น Cloud Speech to Text, Cloud Storage, Compute Engine, Machine Learning เป็นต้น ทั้งนี้การใช้งานกูเกิลคลาวด์แพลตฟอร์มจะคิดค่าใช้จ่ายตามจำนวนการใช้งานทางผู้จัดทำเลือกบริการ Cloud Storage ในการเก็บไฟล์เสียง และใช้ Cloud Speech to Text ไสบริรี Speech ในการแปลงเสียงพูดให้อยู่ในรูปแบบข้อความ [3]

2.2.3 การประมวลผลภาษาธรรมชาติ (Natural Language Processing: NLP)

เป็นศาสตร์ที่ศึกษาเกี่ยวกับการทำให้คอมพิวเตอร์สามารถสื่อสารโต้ตอบด้วยภาษาของมนุษย์ และทำให้คอมพิวเตอร์เข้าใจภาษามนุษย์มากขึ้น เช่น Siri, Google Assistant และ Alexa [4]

2.2.4 Stanford Named Entity Recognizer (Stanford NER)

เป็นการประยุกต์ใช้จากภาษาจาวาสำหรับการรู้จำนิพจน์ระบุนาม (Named Entity Recognizer) ซึ่งเป็นการจัดประเภทของคำในข้อความ เช่น ชื่อสิ่งของ ชื่อบุคคล และบริษัท เป็นการกำหนดโครงสร้างการสกัดคุณสมบัติที่เหมาะสมสำหรับการรู้จำนิพจน์ระบุนาม [5]

2.2.5 Natural Language Toolkit (NLTK)

เป็นแพลตฟอร์มที่นิยมในโปรแกรมภาษาไพทอน เพื่อทำงานกับข้อมูลภาษาของมนุษย์ พร้อมกับชุดของไลบรารีที่ช่วยในการประมวลข้อความ แบ่งประเภทของคำ (Classification) การแบ่งโทเค็นของคำ (Tokenization) การตัดคำ (Stemming) การติดแท็กคำ (Tagging) และการแยกวิเคราะห์คำ (Parsing) [6]

2.2.6 spaCy

เป็นไลบรารีสำหรับการทำการประมวลผลภาษาธรรมชาติขั้นสูงในภาษาไพทอน โดยที่ spaCy จะทำความเข้าใจข้อความจำนวนมากสามารถใช้ในการดำเนินการสกัดข้อมูล (Information Extraction) หรือระบบการทำความเข้าใจภาษาธรรมชาติเพื่อดำเนินการประมวลผลข้อความล่วงหน้าสำหรับการเรียนรู้เชิงลึก (Deep Learning) [7]

2.2.7 Regular Expressions

เป็นสัญลักษณ์ที่ใช้ระบุชุดของอักขระตัวอักษร เมื่อชุดของอักขระตัวอักษรที่เฉพาะเจาะจงนั้นอยู่ในชุดอักขระตัวอักษรที่มีการ

กำหนดให้เป็น Regular Expressions โดยทั่วไปแล้วจะใช้สัญลักษณ์ “*”, “+”, “?”, “()” และ “|” ในการกำหนดเงื่อนไขของชุดตัวอักษร [8]

2.2.8 Pydub

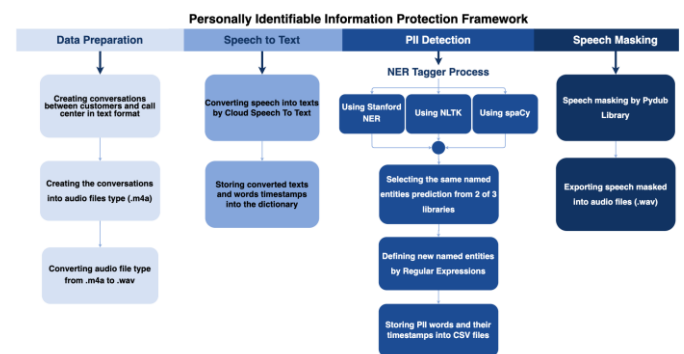
เป็นไลบรารีหนึ่งของภาษาไพทอนที่ใช้ในการจัดการกับข้อมูลที่เป็นไฟล์เสียง

2.2.9 Jaccard's Coefficient Similarity

เป็นสถิติประยุกต์แนวคิดในทฤษฎีเซตเพื่อใช้เปรียบเทียบความคล้ายคลึงและความหลากหลายของกลุ่มตัวอย่าง แนวคิดของค่าสัมประสิทธิ์ Jaccard's Coefficient Similarity คือ การวัดค่าความคล้ายคลึงระหว่างกลุ่มประชากร 2 กลุ่ม โดยคำนวณจากขนาดของประชากรที่ทั้งสองกลุ่มมีตัวอย่างร่วมกัน [9] ดังสมการที่ 1

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

3. ขั้นตอน และวิธีการดำเนินงานวิจัย



รูปที่ 1. แผนงานการปกป้องข้อมูลที่ระบุตัวบุคคล

จากรูปที่ 1 มีการแบ่งส่วนของการดำเนินงานเป็น 4 ส่วนหลัก ๆ คือ การเตรียมข้อมูล การแปลงเสียงพูดให้อยู่ในรูปแบบข้อความ การตรวจจับคำที่เป็นข้อมูลส่วนบุคคลจากข้อมูลรูปแบบข้อความ และการแทนที่เสียงที่เป็นข้อมูลส่วนบุคคลด้วยเสียงรบกวน

3.1 การเตรียมข้อมูล (Data Preparation)

3.1.1 สร้างบทสนทนาระหว่างลูกค้าและศูนย์บริการข้อมูลลูกค้าทางโทรศัพท์

ผู้จัดทำได้ดำเนินการสร้างชุดข้อมูลในรูปแบบข้อความจำนวนทั้งหมด 23 บทสนทนา จากการวิเคราะห์ประโยคในบทสนทนาคิดเป็น 566 ประโยค ค่าเฉลี่ยใน 1 บทสนทนาจะมีประโยคโดยเฉลี่ยจำนวน 24.61 ประโยค หากแบ่งย่อยลงไปเป็นการวิเคราะห์คำที่ยังไม่ผ่านการทำความสะอาดข้อมูลมีทั้งหมด 4,095 คำ ค่าเฉลี่ยใน 1 บทสนทนาจะมีคำโดยเฉลี่ยจำนวน 178.04 คำ และหาก

3.1.2 นำข้อมูลในรูปแบบข้อความมาดำเนินการบันทึกเสียง

3.2 การแปลงเสียงพูดให้อยู่ในรูปแบบข้อความ

ดำเนินการสร้างโปรเจกต์บนกูเกิลคลาวด์ และเปิดใช้งาน API 2 ตัว ได้แก่ Cloud Storage และ Cloud Speech to Text จากนั้นอัปโหลดไฟล์เสียงขึ้นบน Cloud Storage ดังรูปที่ 2

รูปที่ 2. อัปโหลดไฟล์เสียงขึ้นบน Cloud Storage

	word	start_times	end_times
0	Hello,	0.0	0.4
1	you	0.4	1.2
2	have	1.2	1.3
3	called	1.3	1.8
4	virtual	1.8	2.2
...
164	Thank	91.7	92.4
165	you,	92.4	92.5
166	sir.	92.5	93.4
167	Thank	93.4	94.5
168	you.	94.5	94.7

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 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 to give you are ready? I will give you what I have got 760-545-6789. O kay. That's +765-450-600-7889. Correct? What is your identification number? 774-589 585 665. I have +774-580-968-5896 65 and what is your name sir? It is Robert. Appel board. Okay.

Confidence: 72%

3.2.2 บันทึกไฟล์ให้อยู่ในรูปแบบ Dictionary

3.3 การตรวจจับคำที่เป็นข้อมูลส่วนบุคคลจากข้อมูลรูปแบบข้อความ

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

- พัฒนาไลบรารี NLTK ทางผู้จัดทำได้เลือกการติดแท็กในบทสนทนาเป็นจำนวนทั้งหมด 6 ประเภท ได้แก่ ORGANIZATION, PERSON, LOCATION, GPE, DATE และ MONEY โดยเริ่มจากการทำ Word Tokenization จากนั้นทำการติดแท็กนิพจน์ระบุนามจากอัลกอริทึม NLTK ซึ่งต้องมีการติดแท็กส่วนของประโยค (Part-of-Speech) ก่อนจึงจะติดแท็กได้ และกระบวนการหลังจากนั้นมิมีวิธีการทำเช่นเดียวกันกับ Stanford NER คือ เลือกโทเค็นที่เป็นนิพจน์ระบุนาม และทำการเปลี่ยนประเภทนิพจน์ระบุนามให้เหมือนกันทุกไลบรารี จากนั้นจับคู่โทเค็นที่ไลบรารีแบ่งเทียบกับโทเค็นของ Cloud Speech to Text และเก็บค่าของโทเค็น

- พัฒนาไลบรารี spaCy โดยทางผู้จัดทำได้ดำเนินการเลือกการติดแท็กในบทสนทนาเป็นจำนวนทั้งหมด 6 ประเภท ได้แก่ ORGANIZATION, PERSON, LOCATION, GPE, DATE และ MONEY ในฟังก์ชันมีการใช้อัลกอริทึมของ spaCy ซึ่งในอัลกอริทึมจะดำเนินการวิเคราะห์ข้อความต่าง ๆ อัตโนมัติ ซึ่งสามารถเรียกดูค่าได้จากอัลกอริทึมได้ทันที และกระบวนการหลังจากนั้นมีวิธีการทำเช่นเดียวกันกับ Stanford NER และ NLTK คือ เลือกโทเค็นที่เป็นนิพจน์ระบุนาม และทำการเปลี่ยนประเภทนิพจน์ระบุนามให้เหมือนกันทุกไลบรารี จากนั้นจับคู่โทเค็นที่แบบจำลองแบ่งเทียบกับโทเค็นของ Cloud Speech to Text และเก็บค่าของโทเค็น

3.3.2 เลือกการทำนายประเภทของนิพจน์ระบุนามที่เหมือนกันตั้งแต่ 2 ใน 3 ของไลบรารี

ดำเนินการเลือกการทำนายประเภทของนิพจน์ระบุนามที่เหมือนกันตั้งแต่ 2 จาก 3 ไลบรารีขึ้นจากการสร้างฟังก์ชันจับคู่โทเค็นที่มีการทำนายนิพจน์ระบุนามค่าเดียวกัน และเก็บค่าของโทเค็นนั้นใหม่ เพื่อนำไปใช้วิเคราะห์กระบวนการถัดไป ในที่นี้ทางผู้จัดทำขอแทนผลลัพธ์ของกระบวนการนี้ว่าค่าทำนายจริง

3.3.3 สร้างประเภทของนิพจน์ระบุนามเพิ่มเพื่อติดแท็กเลขที่เป็นข้อมูลส่วนบุคคลโดยใช้ Regular Expressions

ขั้นตอนนี้จะมีการดึงโทเค็นค่าของ Cloud Speech to Text เฉพาะที่เป็นเลขมาตรวจสอบเงื่อนไขเพื่อติดแท็กเลขที่เป็นข้อมูลส่วนบุคคลเท่านั้น โดยแบ่งประเภทของเลขที่เป็นข้อมูลส่วนบุคคลไว้ 5 ประเภท คือ IDCARD (เลขบัตรประชาชน 13 หลัก) PHONENUM (เบอร์โทรศัพท์ 10 หลัก) ACCNUM (เลขบัญชี 9 หลัก) CARDNUM (เลขบัตรเดบิต หรือบัตรเครดิต 16 หลัก) และ PIINUM (เลขอื่น ๆ ที่ไม่เข้าเงื่อนไขประเภทก่อนหน้านี้ แต่มีตั้งแต่ 9 หลักขึ้นไป มีไว้ในกรณีที่ Cloud Speech to Text แปลงเป็นข้อความออกมาได้ไม่แม่นยำ) จากนั้นนำค่าที่ได้ไปรวมกับค่าทำนายจริง และเก็บค่าที่ไว้ในรูปแบบไฟล์ CSV เพื่อนำไปดำเนินการต่อในขั้นถัดไป

3.3.4 เก็บค่าต่าง ๆ ให้อยู่ในรูปแบบของไฟล์ CSV

หลังจากดำเนินการทำนายนิพจน์ระบุนามทั้งหมดแล้ว จึงจัดเก็บค่าเหล่านั้นให้อยู่ในรูปแบบตารางและบันทึกเป็นไฟล์ CSV โดยมีจำนวนทั้งหมด 5 คอลัมน์ ได้แก่ ลำดับโทเค็น โทเค็นค่า เวลาที่เริ่มพูดโทเค็นนั้นในไฟล์เสียง เวลาที่พูดโทเค็นนั้นจบ และประเภทของนิพจน์ระบุนาม

3.4 การแทนที่เสียงที่เป็นข้อมูลส่วนบุคคลด้วยเสียงรบกวน

ขั้นตอนนี้มีการดึงข้อมูลไฟล์ CSV มาใช้ในการดำเนินงาน คือ นำคอลัมน์ของเวลาเริ่มต้นของคำพูดและเวลาสิ้นสุดมาใช้ โดยแปลงให้ค่าของเวลาอยู่ในหน่วยของมิลลิวินาที จากนั้นแทนที่เสียงรบกวนในช่วงเวลาที่คำนวณไว้ จากนั้นดำเนินการบันทึกไฟล์เสียงที่มีการแทนที่ข้อมูลส่วนบุคคลแล้วเป็นไฟล์ประเภท .wav

3.5 การประเมินผล (Evaluation)

มีการประเมินผล 2 กระบวนการหลัก ๆ คือ ประเมินผลความแม่นยำของการแปลงเสียงพูดให้อยู่ในรูปแบบข้อความโดยนำแนวคิดของ Jaccard's Coefficient Similarity มาประยุกต์ใช้ในการประเมินผล และกระบวนการประเมินผลความแม่นยำของการตรวจจับคำที่เป็นข้อมูลส่วนบุคคลจากข้อมูลรูปแบบข้อความ จากการสร้างผลเฉลยของการทำนายข้อความและโทเค็นต่าง ๆ เพื่อใช้ตรวจสอบความแม่นยำในการทำนายของไลบรารีทั้งหมด

4. ผลการดำเนินงานเบื้องต้น

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.", '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,
```

รูปที่ 5. ตัวอย่างการแปลงข้อมูลเสียงให้อยู่ในรูปแบบข้อความโดยใช้ Cloud Speech to Text

จากรูปที่ 5 ดำเนินการแปลงให้อยู่ในรูปแบบของ Dictionary และสร้างคีย์ที่ชื่อว่า transcript ไว้เก็บข้อความในบทสนทนาทั้งหมดในส่วนของโทเค็นค่า ได้มีการสร้างคีย์ที่ชื่อว่า values ไว้เก็บค่าของเวลาที่เริ่มพูดโทเค็นนั้น ๆ (start) เวลาที่พูดจบ (end) และโทเค็นนั้น ๆ (word) และประเมินผลความแม่นยำในการทำนายของแบบจำลอง โดยการนำข้อมูลบทสนทนาจริงเทียบกับข้อมูลที่แบบจำลองทำนายโดยใช้ 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 Ville branch and I think I left my Debit card in the ATM machine. Okay. Do you have your Debit card number? I don't have. Okay, well do you have the checking account number associated with the Debit card? That I do have. Are you ready? I will give you what I have got. 765-456-789. Okay. That's 765-456-789. Correct. What is your identification number? 774-589-658-9665. Okay, I have 774-589-658-9665 and what is your name sir? It is Robert Applebaum. Okay. I have 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 610-265-1715. Okay. I have 610-265-1715. Yes. Okay Mr. Applebaum. I have just suspended your card. If it is in the machine, we will contact you and lift the suspension. Oh, thank you, Sure. Thank you.

รูปที่ 6. 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 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."

รูปที่ 7. บทสนทนาที่แบบจำลองทำนาย

```
acc = Jaccard_Similarity(dict_, ori_text)
acc = acc*100

print('Accuracy of the conversation:', '%.2f' %acc, '%')
Accuracy of the conversation: 57.02 %
```

รูปที่ 8. ค่าของความแม่นยำในการทำนาย

จากรูปที่ 8 ความแม่นยำในการทำนายคิดเป็นร้อยละ 57.02 ซึ่งเมื่อเทียบบทสนทนารูปที่ 6 และ 7 จะพบว่าสิ่งที่ส่งผลให้ค่าความแม่นยำของแบบจำลองไม่สูงนั้นส่วนใหญ่แล้วขึ้นอยู่กับเครื่องหมายวรรคตอนของข้อมูลบทสนทนาทั้งสอง ดังนั้น จึงดำเนินการสร้างฟังก์ชันตัดเครื่องหมายวรรคตอนของบทสนทนาทั้งสองออก เพื่อประเมินผลค่าความแม่นยำใหม่ ดังรูปที่ 9, 10 และ 11

'Hello you have called virtual bank this is Linda speaking. How may I help you? Hi Linda. I was just at your Ville branch and I think I left my Debit card in the ATM machine. Okay. Do you have your Debit card number? I don't have. Okay. Well, do you have the checking account number associated with the Debit card? That I do have. Are you ready? I will give you what I have got 765456789. Okay. That's 765456789. Correct? What is your identification number? 7745896589665. Okay. I have 7745896589665 and what is your name sir? It is Robert Applebaum. Okay. I have 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 6102651715. Okay. I have 6102651715. Yes. Okay Mr. Applebaum. I have just suspended your card. If it is in the machine, we will contact you and lift the suspension. Oh, thank you Sure. Thank you.'

รูปที่ 9. ข้อมูลบทสนทนาจริงที่ผ่านการทำความสะอาด

'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 7605456789. Okay. That's 7654506007089. Correct? What is your identification number? 7745896589665 okay. I have 774580960589665 and what is your name sir? It is Robert Applebaum. 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 6102651715. Okay. I have 6102651715. 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.'

รูปที่ 10. บทสนทนาที่แบบจำลองทำนายที่ผ่านการทำความสะอาด

```
acc = Jaccard_Similarity(clean_text(dict_), clean_text(ori_text))
acc = acc*100

print('Accuracy of the conversation:', '%.2f' %acc, '%')
Accuracy of the conversation: 71.43 %
```

รูปที่ 11. ค่าของความแม่นยำในการทำนาย (ใหม่)

จากรูปที่ 11 ความแม่นยำในการทำนายคำพูดของแบบจำลองคิดเป็นร้อยละ 71.43 สามารถเห็นได้ชัดว่าค่าความแม่นยำสูงขึ้นอย่างชัดเจน เมื่อตัดเครื่องหมายวรรคตอนออกเบื้องต้น

4.2 การตรวจจับคำที่เป็นข้อมูลส่วนบุคคลจากข้อมูลรูปแบบข้อความ

เมื่อดำเนินการนำข้อมูลในรูปแบบข้อความที่ได้จาก Cloud Speech to Text มาเข้าฟังก์ชันต่าง ๆ ของไลบรารี Stanford NER, NLTK และ spaCy พร้อมกับนำเข้าฟังก์ชันของการเลือกคำทำนายจริง และสร้างนิพจน์ระบุนามเพิ่มสำหรับเลขที่เป็นข้อมูลส่วนบุคคลโดยใช้ Regular Expressions ทางผู้จัดทำก็ได้ดำเนินการเก็บค่าของการทำนายของทุก ๆ แบบจำลองไว้ในรูปแบบตาราง ดังรูปที่ 12

	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 12. ตารางการทำนายประเภทของนิพจน์ระบุนาม

จากรูปที่ 12 ได้เก็บค่าการทำนายของโทเค็นไว้ในตารางเดียวกันตามประเภทของนิพจน์ระบุนาม แถวใดมีการทำนายเป็นคำว่า “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 (ค่า

ทำนายจริง) นอกจากนี้ ยังได้ดำเนินการเก็บบันทึกค่าการทำนายจริง เฉพาะโทเค็นที่มีการติดแท็กนิพจน์ระบุนามขึ้นมาอีก 1 ตาราง เพื่อดำเนินการบันทึกให้อยู่ในรูปแบบไฟล์ CSV และนำไปกปิดเสียงในขั้นตอนถัดไป ดังรูปที่ 13

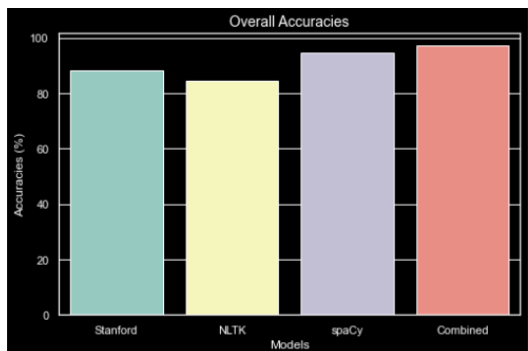
	word	start_time	end_time	real_ents
indx				
8	Linda	3.4	3.8	PERSON
16	Linda.	6.8	7.2	PERSON
34	ATM	11.7	11.8	ORGANIZATION
76	760-545-6789.	27.2	30.6	PHONENUM
79	+765-450-600-7089.	32.7	35.7	IDCARD
86	774-589-6589	38.9	42.7	PHONENUM
91	+774-580-960-5896	45.4	49.0	IDCARD
101	Robert.	51.9	52.3	PERSON
107	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 13. ตารางค่าทำนายจริงเฉพาะที่มีการติดแท็กนิพจน์ระบุนาม

ทางผู้จัดทำมีการประเมินผลความแม่นยำในการทำนายนิพจน์ระบุนามของแต่ละแบบจำลอง โดยการนำโทเค็นที่ Cloud Speech to Text แบ่งออกมา ไปทำการเจดลยนิพจน์ระบุนามจริง เพื่อที่จะนำไปประเมินผลความแม่นยำของการทำนายนิพจน์ระบุนามในทุก ๆ แบบจำลอง

Stanford Accuracy: 88.17%
 NLTK Accuracy: 84.62%
 spaCy Accuracy: 94.67%
 ** Combined Models and using Regular Expressions Accuracy: 97.04% **

รูปที่ 14. การประเมินผลความแม่นยำของแต่ละแบบจำลอง



รูปที่ 15. กราฟการประเมินผลความแม่นยำของแต่ละแบบจำลอง

จากรูปที่ 14 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

5.1.1 การแปลงเสียงพูดให้อยู่ในรูปแบบข้อความ

การแปลงเสียงพูดให้อยู่ในรูปแบบข้อความนั้น หากเป็นการประเมินผลโดยไม่คำนึงถึงความถูกต้องของเครื่องหมายวรรคตอน ถือว่าค่าของความแม่นยำอยู่ในระดับที่ดี อาจจะมีการแปลงชื่อบุคคลที่ไม่ตรงกับข้อมูลบทสนทนาจริงเล็กน้อย อาจเป็นสาเหตุมาจากเสียงที่ใช้ในการดำเนินการบันทึกเสียงที่แต่ละบุคคลมีสำเนียงการพูดที่ไม่เหมือนกัน เช่น นามสกุล Applebaum เมื่อเป็นเสียงของ Siri Male ทางแบบจำลองแปลงได้เป็น 2 โทเค็น คือ “Appel” และ “board.” แต่เมื่อเป็นเสียงของ “Siri Female” ทางแบบจำลองกลับแปลงคำได้ถูกต้อง จึงสรุปได้ว่าบางครั้งสำเนียงการพูดของแต่ละบุคคลอาจส่งผลต่อความแม่นยำของการแปลงข้อมูลเสียงให้อยู่ในรูปแบบข้อความ นอกจากนี้ ยังมีการแปลงเลขที่ผิดพลาดไปบ้าง เช่น เมื่อสรีพูดว่า “oh” ในบางครั้งแบบจำลองจะแปลงเป็นเลข “0” ซึ่งส่งผลให้ความแม่นยำของแบบจำลองลดลง

5.1.2 การตรวจจับคำที่เป็นข้อมูลส่วนบุคคลจากข้อมูลรูปแบบข้อความ

ในขั้นตอนนี้ ผู้จัดทำจะอธิบายรายละเอียดของแต่ละไลบรารี ดังนี้

- Stanford NER สามารถติดแท็กบุคคล และค่าเงินได้ค่อนข้างแม่นยำ แต่ในการติดแท็กวันที่ ด้วยข้อจำกัดที่ไม่มีการติดแท็กตัว

เลขที่เป็นประเภท Cardinal จึงส่งผลให้มีการติดแท็กตัวเลขธรรมดา เป็นประเภทของวันที่ (Date) ทำให้ความแม่นยำลดลง

- NLTK สามารถติดแท็กองค์กรได้แม่นยำมากที่สุด ส่วนนิพจน์ระบุนามประเภทอื่น ๆ มีความแม่นยำเฉลี่ยเท่า ๆ กันกับแบบจำลองอื่น ๆ แต่แบบจำลองนี้มักมีการติดแท็กที่ผิดพลาดตรงส่วนของสถานที่ กล่าวคือ หากโทเค็นนั้นขึ้นต้นด้วยตัวอักษรพิมพ์ใหญ่ เช่น คำว่า “Hello” แบบจำลองจะติดแท็กเป็นสถานที่ทันที

- spaCy จากผลลัพธ์การประเมินผลความแม่นยำ จะสังเกตได้ว่าส่วนใหญ่แล้ว spaCy มีค่าความแม่นยำสูงในการติดแท็กโทเค็น แต่หากให้สรุปเป็นรายประเภท จะสรุปได้ว่าสามารถติดแท็กบุคคล สถานที่ วันที่ และค่าเงินได้ดีที่สุด แต่เนื่องจากการติดแท็กของไลบรารีนี้ยังไม่แม่นยำอยู่บ้าง ทางผู้จัดทำจึงมีความเห็นว่าควรรวมไลบรารีเข้าด้วยกันเพื่อเพิ่มประสิทธิภาพในการติดแท็ก

ในส่วนของการรวมไลบรารีเข้าด้วยกันมีความแม่นยำค่อนข้างสูง ซึ่งเฉลี่ยแล้วคิดเป็นร้อยละ 90 ถือเป็นค่าความแม่นยำที่น่าพึงพอใจ

และการตรวจจับเลขที่เป็นข้อมูลส่วนบุคคลโดยใช้ 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

5.1.3 การแทนที่เสียงที่เป็นข้อมูลส่วนบุคคลด้วยเสียงรบกวน

การแทนที่เสียงที่เป็นข้อมูลส่วนบุคคลด้วยเสียงรบกวน ในบางช่วงที่มีการแทนที่เสียงรบกวนนั้นอาจเกินหรือขาดจากระยะเวลาของคำพูดที่เป็นข้อมูลส่วนบุคคล แต่โดยภาพรวมแล้วถือว่าปิดบังคำพูดที่เป็นข้อมูลส่วนบุคคลได้ดี

5.2 ปัญหาในการทำโครงงานและสรุปผล

โดยส่วนใหญ่แล้ว ปัญหาในการทำโครงงานนี้ คือ ความแม่นยำของการแปลงข้อมูลเสียงให้อยู่ในรูปแบบข้อความนั้น มีความแม่นยำในระดับปานกลางจนถึงค่อนข้างสูง แต่เมื่อเข้าสู่กระบวนการตรวจจับคำที่เป็นข้อมูลส่วนบุคคลจากข้อมูลรูปแบบข้อความ ส่งผลให้แบบจำลองไม่สามารถติดแท็กประเภทของโทเค็นที่ควรจะมีนิพจน์ระบุนามได้ เช่น ชื่อบุคคล หรือส่วนเล็ก ๆ ของเลขที่เป็นข้อมูลส่วนบุคคล จึงอาจส่งผลให้เป็นปัญหาต่อการปิดบังคำที่เป็นข้อมูลส่วนบุคคล

บุคคลในขั้นตอนสุดท้ายได้ และในการเป็นการแทนที่เสียงที่เป็นข้อมูลส่วนบุคคลด้วยเสียงรบกวน ในบางช่วงของการแทนที่เสียงรบกวนอาจเกินหรือขาดจากระยะเวลาของคำพูดที่เป็นข้อมูลส่วนบุคคล

5.3 แนวทางในการพัฒนาต่อ

ทางผู้จัดทำจะดำเนินการหาวิธีการเพิ่มค่าความแม่นยำของการแปลงข้อมูลเสียงให้อยู่ในรูปแบบข้อความให้มีความแม่นยำมากขึ้น เพื่อให้การติดแท็กโทเค็นตรงเงื่อนไขมากที่สุด และอาจมีการดำเนินการพัฒนาต่อเพิ่มในด้านของการตรวจจับข้อมูลส่วนบุคคล เช่น หลังจากติดแท็กโทเค็นนั้นแล้ว อาจมีการฝึกฝนแบบจำลองอื่น ๆ เพิ่มเติม เพื่อตรวจจับว่าโทเค็นนั้น ๆ เป็นข้อมูลส่วนบุคคลที่จำเป็นต้องปกปิดจริงหรือไม่ แต่ด้วยวิธีการนั้นอาจจะต้องดำเนินการสร้างชุดข้อมูลพร้อมกับการเฉลยผลการตรวจจับว่าเป็นข้อมูลส่วนบุคคลหรือไม่ เป็นจำนวนมาก เพื่อให้แบบจำลองสามารถทำนายได้อย่างแม่นยำ

เอกสารอ้างอิง

- [1] ศ. สวัสดิ์พงศ์ธาดา. “ความเป็นส่วนตัว (Privacy)”. [Online]. Available: <https://angasila.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] A. Geitgey Natural Language Processing is Fun!”. [Online]. Available: <https://bit.ly/36Vpsk8>. 2018.
- [5] C. Dishmon. “Named Entity Recognition with Stanford NER Tagger”. [Online]. Available: <https://bit.ly/3LXHvKB>. 2020.
- [6] NLTK. “Natural Language Toolkit”. [Online]. Available: <https://www.nltk.org/>. 2020.

- [7] spaCy. “spaCy 101: Everything you need to know”.
[Online]. Available:
<https://spacy.io/usage/spacy-101>. 2020.
- [8] 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.
- [9] ศุภวัจน์ แต่รุ่งเรือง. “การตรวจเทียบภายนอกหาการลักลอก
ในงานวิชาการโดยใช้แบบจำลองซัพพอร์ตเวกเตอร์
แมชชีนและการวัดค่าความคล้ายของข้อความ”.
(วิทยานิพนธ์ปริญญาอักษรศาสตรดุษฎีบัณฑิต
จุฬาลงกรณ์มหาวิทยาลัย, 2560).