



โครงการ

การเรียนการสอนเพื่อเสริมประสบการณ์

ชื่อโครงการ ระบบจัดหมวดหมู่เจตนาของประโยคและการสกัดชื่อเฉพาะจากประวัติ
สนทนาเพื่อฝึกสอนแชทบ็อต

An intent clustering and named entity extraction from
chatlogs for training chatbot

ชื่อนิสิต	นายณัฐกิตติ์ สมอร่าม	583 36213 23
	นางสาวพิชชาภา พะวงษ์	583 36477 23

ภาควิชา คณิตศาสตร์และวิทยาการคอมพิวเตอร์
สาขาวิชา วิทยาการคอมพิวเตอร์

ปีการศึกษา 2561

บทคัดย่อและฉบับเต็มของวิทยานิพนธ์ของงานทางวิชาการที่ขึ้นชื่อว่าเป็นคลังข้อมูลความรู้ (CUIR)
คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
เป็นแฟ้มข้อมูลของนิสิตเจ้าของโครงการงานทางวิชาการที่ส่งผ่านทางคณะที่สังกัด

The abstract and full text of senior projects in Chulalongkorn University Intellectual Repository(CUIR)
are the senior project authors' files submitted through the faculty.

ระบบจัดหมวดหมู่เจตนาของประโยคและการสกัดชื่อเฉพาะจากประวัติสนทนาเพื่อฝึกสอนแชทบ็อต

นายณัฐกิตต์ สมอรัมย์
นางสาวพิชชาภา พะวงษ์

โครงการนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรบัณฑิต
สาขาวิชาวิทยาการคอมพิวเตอร์ ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์
คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2561
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

An intent clustering and named entity extraction system from chatlogs
for training chatbot

Nattakit Somaram

Pitchapa Pawong

A Project Submitted in Partial Fulfillment of the Requirements
for the Degree of Bachelor of Science Program in Computer Science

Department of Mathematics and Computer Science

Faculty of Science


Chulalongkorn University

Academic Year 2018


Copyright of Chulalongkorn University

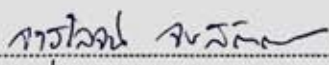
หัวข้อโครงการ ระบบจัดหมวดหมู่เจตนาของประโยคและการสกัดชื่อเฉพาะ
จากประวัติสนทนาเพื่อฝึกสอนแชตบ็อต
โดย นายณัฐกิตต์ สมอร่วม
นางสาวพิชชาภา พะวงษ์
สาขาวิชา วิทยาการคอมพิวเตอร์
อาจารย์ที่ปรึกษาโครงการหลัก ผู้ช่วยศาสตราจารย์ ดร.จิตยา หวานวารี

ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
อนุมัติให้นับโครงการฉบับนี้เป็นส่วนหนึ่ง ของการศึกษาตามหลักสูตรปริญญาบัณฑิต ในรายวิชา
2301499 โครงการวิทยาศาสตร์ (Senior Project)


..... หัวหน้าภาควิชาคณิตศาสตร์
(ศาสตราจารย์ ดร.กฤษณะ เนียมมณี) และวิทยาการคอมพิวเตอร์

คณะกรรมการสอบโครงการ


..... อาจารย์ที่ปรึกษาโครงการหลัก
(ผู้ช่วยศาสตราจารย์ ดร.จิตยา หวานวารี)


..... กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.จารุโลจน์ จงสถิตย์วัฒนา)

ภิกขุณ ปักซี่
..... กรรมการ
(ผู้ช่วยศาสตราจารย์ ดร.ภควรรณ ปักซี่)

ณัฐกิตต์ สมอรัมย์, พิชชาภา พะวงษ์: ระบบจัดหมวดหมู่เจตนาของประโยคและการสกัดชื่อเฉพาะจากประวัติสนทนาเพื่อฝึกสอนแชทบ็อต (An intent clustering and named entity extraction system from chatlogs for training chatbot) อ.ที่ปรึกษาโครงการหลัก : ผู้ช่วยศาสตราจารย์ ดร.จิตยา หวานวารี, 84 หน้า.

โครงการ “ระบบจัดหมวดหมู่เจตนาของประโยคและการสกัดชื่อเฉพาะจากประวัติสนทนาเพื่อฝึกสอนแชทบ็อต” มีจุดมุ่งหมายเพื่อช่วยลดความยุ่งยาก และอำนวยความสะดวกในการพัฒนาแชทบ็อต และสามารถนำไปเป็นแนวทางการออกแบบการตอบคำถามของแชทบ็อตได้สะดวกยิ่งขึ้น

ขั้นตอนการพัฒนา ระบบจัดหมวดหมู่เจตนาของประโยคและการสกัดชื่อเฉพาะจากประวัติสนทนาเพื่อฝึกสอนแชทบ็อต เริ่มต้นด้วยการเก็บข้อมูลประวัติสนทนา แล้วนำข้อมูลที่ได้อามาเตรียมข้อมูลให้เป็นคำย่อย เมื่อได้คำย่อยแล้วนำคำย่อยมาสร้างเป็นถ่วงคำ จากนั้นนำถ่วงคำไปสร้างเวกเตอร์แทนประโยคทั้งหมด 3 รูปแบบ ได้แก่ การเข้ารหัสข้อมูลวันฮ็อต การนับความถี่ และ Term Frequency – Invert Document Frequency เมื่อได้เวกเตอร์แทนประโยคจากทั้ง 3 รูปแบบแล้วจึงนำมาลดมิติของข้อมูลด้วยการวิเคราะห์องค์ประกอบหลัก และการจัดสรรตรีเคิลแฝง แล้วจัดหมวดหมู่เจตนาของประโยค ส่วนการสกัดชื่อเฉพาะนั้นจะเป็นการนำชุดของคำย่อยมาผ่านวิธีการเทียบคำที่ยาวที่สุด เมื่อได้หมวดหมู่ของเจตนาและชื่อเฉพาะแล้วก็นำมาสร้างแชทบ็อตผ่าน Dialogflow และผสมการทำงาน Dialogflow กับ Firebase และ SendGrid

ผลลัพธ์ที่ได้จากการพัฒนาพบว่า การสร้างเวกเตอร์แทนประโยคด้วยวิธีการ Term Frequency – Invert Document frequency ให้ผลลัพธ์ที่ดีที่สุด การสกัดชื่อเฉพาะที่พัฒนาให้ความแม่นยำที่ 77% และแชทบ็อตสามารถตอบคำถามสำหรับแต่ละเจตนาได้ รวมถึงสามารถดึงข้อมูลรายละเอียดวิชาจากฐานข้อมูลใน Firebase มาประกอบการตอบคำถาม และส่งจดหมายอิเล็กทรอนิกส์ไปยังผู้ดูแลได้

ภาควิชา...คณิตศาสตร์และวิทยาการคอมพิวเตอร์...ลายมือชื่อนิสิต...*ณัฐกิตต์ สมอรัมย์*.....
 ลายมือชื่อนิสิต...*พิชชาภา พะวงษ์*.....
 สาขาวิชา...วิทยาการคอมพิวเตอร์...ลายมือชื่อ อ.ที่ปรึกษาโครงการหลัก...*จิตยา หวานวารี*.....
 ปีการศึกษา...2561.....

5833621323, 5833647723: MAJOR COMPUTER SCIENCE

KEYWORDS : INTENT CLUSTERING / NAMED ENTITY EXTRACTION / CHATBOT

NATTAKIT SOMARAM, PITCHAPA PAWONG: AN INTENT CLUSTERING AND NAMED
ENTITY EXTRACTION SYSTEM FROM CHATLOGS FOR TRAINING CHATBOT. ADVISOR :
ASST. PROF. DITTAYA WANVARIE, Ph.D., 84 pp.

The purpose of “An intent clustering and named entity extraction system from chatlogs for training chatbot” is to reduce the complications and facilitating the development of chatbot. It can be used as a guideline to design a chatbot interaction.

The process of developing an intent clustering and named entity extraction system from chatlogs for training chatbot: Collecting chatlogs from the facebook page. Then, we tokenize and create bag-of-words. After that, we create vector that represent of sentence by 3 methods, 1-hot encoding, frequency count, term frequency-invert document frequency. After that, we use Principal Component Analysis and Latent Dirichlet frequency to reduce vector dimension. Next, we extract named entity by longest-matching method. Finally, we integrate dialogflow with Messenger, Firebase, and SendGrid.

We found that using term frequency invert document frequency gives the best result and named entity extraction give a result of 77%. Finally, chatbot can answer questions for each intent, bring courses information from database in Firebase to answer questions and send e-mail to admin.

Department : Mathematics and Computer Science Student's Signature nattakit
Student's Signature Pitchapa
Field of Study : Computer Science Advisor's Signature Dittha
Academic Year : 2018

กิตติกรรมประกาศ

การจัดทำโครงการเรื่อง “ระบบจัดหมวดหมู่เจตนาของประโยคและการสกัดชื่อเฉพาะจากประวัติสนทนาเพื่อฝึกสอนแชทบ็อต” สามารถสำเร็จลุล่วงไปได้ด้วยดี เนื่องจากได้รับการอนุเคราะห์และความช่วยเหลือจากหน่วยงาน คณาจารย์ และบุคลากรต่าง ๆ ตลอดระยะเวลาการดำเนินการทำโครงการ ผู้พัฒนาจึงขอขอบพระคุณในความอนุเคราะห์และสนับสนุนที่มอบให้ ดังต่อไปนี้

ขอขอบพระคุณ ผู้ช่วยศาสตราจารย์ ดร.ชิตยา หวานวารี อาจารย์ที่ปรึกษาโครงการที่คอยให้คำแนะนำ คำปรึกษา และแนวทางสำหรับการดำเนินงาน ตลอดจนปรับปรุงแก้ไขข้อผิดพลาดและข้อบกพร่องต่าง ๆ ด้วยความเอาใจใส่ตลอดการดำเนินการทำโครงการ

ขอขอบพระคุณ ผู้ช่วยศาสตราจารย์ ดร.จาร์โลจน์ จงสถิตย์วัฒนา และผู้ช่วยศาสตราจารย์ ดร.ภควรรณ ปักซี่ ซึ่งเป็นคณะกรรมการสอบครั้งนี้ ที่คอยชี้แนะให้เห็นถึงข้อผิดพลาดและข้อบกพร่องต่าง ๆ ที่ควรแก้ไข และแนะนำแนวทางที่ก่อให้เกิดประโยชน์กับโครงการขึ้นนี้

ขอขอบพระคุณ อาจารย์ ดร.นฤมล ประทานวณิช อาจารย์ประจำภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์ จุฬาลงกรณ์มหาวิทยาลัย ที่คอยแนะนำแนวทางสำหรับการดำเนินการทำโครงการขึ้นนี้

ขอขอบพระคุณศูนย์นวัตกรรมการเรียนรู้ จุฬาลงกรณ์มหาวิทยาลัย ที่ให้การสนับสนุนข้อมูลจากเพจ Chula MOOC รวมถึงความต้องการของหน่วยงาน และคำแนะนำต่าง ๆ ที่ก่อให้เกิดประโยชน์กับโครงการขึ้นนี้

ขอขอบพระคุณครอบครัวที่คอยช่วยสนับสนุน ดูแลเอาใจใส่ เป็นกำลังใจ และเป็นแรงผลักดันให้โครงการขึ้นนี้สามารถสำเร็จลุล่วงไปได้

ขอขอบคุณเพื่อน ๆ ทุกคนที่คอยช่วยเหลือ ให้คำปรึกษาเมื่อเกิดปัญหา และเป็นกำลังใจให้กัน จนโครงการขึ้นนี้สามารถสำเร็จลุล่วงไปได้

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	ง
บทคัดย่อภาษาอังกฤษ	จ
กิตติกรรมประกาศ	ฉ
สารบัญ	ช
สารบัญตาราง	ฅ
สารบัญภาพ	ญ
บทที่ 1 บทนำ.....	11
1.1 ความเป็นมาและเหตุผลของโครงการ.....	11
1.2 วัตถุประสงค์ของโครงการ.....	12
1.3 ขอบเขตของโครงการ.....	12
1.4 ขั้นตอนการดำเนินงาน.....	13
1.5 ประโยชน์ที่คาดว่าจะได้รับ.....	15
1.6 โครงสร้างของรายงาน.....	15
บทที่ 2 ทฤษฎี และซอฟต์แวร์ที่เกี่ยวข้อง.....	16
2.1 ซอฟต์แวร์ที่เกี่ยวข้อง	16
2.2 ทฤษฎีที่เกี่ยวข้อง	18
บทที่ 3 วิธีการพัฒนา.....	29
3.1 โครงสร้างระบบโดยรวม	29
3.2 เครื่องมือที่ใช้ในการพัฒนา.....	30
3.3 วิธีการเก็บข้อมูล	30
3.4 การเตรียมข้อมูล	31
3.5 การสร้างคุณค่า.....	32

3.6 การสร้างเวกเตอร์.....	33
3.7 การสกัดชื่อเฉพาะ.....	34
3.8 การลดมิติของข้อมูล.....	34
3.9 การจัดหมวดหมู่.....	34
3.10 การวัดความไม่บริสุทธิ์ของข้อมูล.....	35
3.11 การส่งออกระบบมาใช้งาน.....	35
3.12 การสร้างไฟล์เจสันเพื่อป้อนให้กับ Dialogflow.....	35
3.13 การกำหนดเจตนา (Intent) ลง Dialogflow.....	36
3.14 การกำหนดคำเฉพาะ (Entity) ลง Dialogflow.....	37
3.15 การผสมการทำงาน Dialogflow เข้ากับ Messenger บน Facebook.....	38
3.16 การผสมการทำงาน Dialogflow เข้ากับ Firebase.....	39
3.17 การส่งจดหมายอิเล็กทรอนิกส์ผ่าน SendGrid.....	40
บทที่ 4 ผลการพัฒนา.....	42
4.1 ผลการจัดหมวดหมู่เจตนาของประโยค.....	42
4.2 ผลการสกัดชื่อเฉพาะ.....	46
4.3 ผลการทำงานของแช็ตบ็อต.....	46
บทที่ 5 ข้อสรุปและปัญหาที่พบ.....	52
5.1 ข้อสรุป.....	52
5.2 ปัญหาที่พบและแนวทางการแก้ไข.....	52
รายการอ้างอิง.....	54
ภาคผนวก ก แบบเสนอหัวข้อโครงการ รายวิชา 2301399 Project Proposal ปีการศึกษา 2561 .	57
ภาคผนวก ข รายชื่อคำหยุด.....	64
ภาคผนวก ค ภาพหน้าจอการทำงานจริงของแช็ตบ็อต.....	76
ประวัติผู้เขียน.....	83

สารบัญตาราง

	หน้า
ตารางที่ 2.1 ตารางแสดงลักษณะของข้อมูล.....	21
ตารางที่ 2.2 ตารางแสดงข้อมูลที่ผ่านการจัดสรรของดีรีเคลแฝง	26
ตารางที่ 2.3 เมทริกซ์คอนฟิวชัน	27
ตารางที่ 4.1 ผลการจัดหมวดหมู่เจตนาของประโยค	42
ตารางที่ 5.1 ปัญหาในการพัฒนาและแนวทางการแก้ไข	53

สารบัญภาพ

	หน้า
ภาพที่ 2.1 การทำงานของ Dialogflow.....	16
ภาพที่ 2.2 ฐานข้อมูลบน Firebase.....	18
ภาพที่ 2.3 ขั้นตอนการทำเทคนิคการจัดหมวดหมู่ข้อมูลแบบเคมีน (K-mean Clustering).....	20
ภาพที่ 2.4 แบบจำลองแผนภาพของการจัดสรรของดีรีเคิลแฝง.....	25
ภาพที่ 3.1 ปุ่ม Upload Intent ที่หน้าจัดการเจตนาทั้งหมดของ Dialogflow.....	36
ภาพที่ 3.2 ปุ่ม Choose File ที่หน้า Upload Intent File.....	37
ภาพที่ 3.3 แถบ Entities ที่หน้าส่วนควบคุมของ Dialogflow.....	37
ภาพที่ 3.4 ปุ่ม Upload Entity ที่หน้าจัดการคำเฉพาะทั้งหมดใน Dialogflow.....	37
ภาพที่ 3.5 ปุ่ม Choose File ที่หน้า Upload Entity.....	38
ภาพที่ 3.6 หน้าสำหรับการรับโทเคนการเข้าถึง (Access Tokens) Messenger.....	38
ภาพที่ 3.7 หน้าแสดงค่าสำหรับการผสานงาน Dialogflow เข้ากับ Messenger บน Facebook....	39
ภาพที่ 3.8 หน้าสำหรับเติมค่า Callback URL และค่า Verify Token.....	39
ภาพที่ 3.9 ที่อยู่ของฐานข้อมูลแบบทันทีทันใดของ Firebase.....	39
ภาพที่ 3.10 หน้าสำหรับสร้าง API key บน SendGrid.....	40
ภาพที่ 4.1 ค่าความแปรปรวนของข้อมูลที่สร้างเวกเตอร์ด้วยวิธีการเข้ารหัสแบบวันฮ็อตและลดขนาด เวกเตอร์ด้วยการวิเคราะห์ห้วงค์ประกอบหลัก.....	43
ภาพที่ 4.2 ค่าความแปรปรวนของข้อมูลที่สร้างเวกเตอร์ด้วยวิธีการนับความถี่และลดขนาดเวกเตอร์ ด้วยการวิเคราะห์ห้วงค์ประกอบหลัก.....	43
ภาพที่ 4.3 ค่าความแปรปรวนของข้อมูลที่สร้างเวกเตอร์ด้วยวิธี Term Frequency – Inverse Document Frequency และลดขนาดเวกเตอร์ด้วยการวิเคราะห์ห้วงค์ประกอบหลัก.....	44
ภาพที่ 4.4 ค่าความแปรปรวนของข้อมูลที่สร้างเวกเตอร์ด้วยวิธีการเข้ารหัสแบบวันฮ็อตและลดขนาด เวกเตอร์ด้วยการจัดสรรดีรีเคิลแฝง.....	44
ภาพที่ 4.5 ค่าความแปรปรวนของข้อมูลที่สร้างเวกเตอร์ด้วยวิธีการนับความถี่และลดขนาดเวกเตอร์ ด้วยการจัดสรรดีรีเคิลแฝง.....	45
ภาพที่ 4.6 ค่าความแปรปรวนของข้อมูลที่สร้างเวกเตอร์ด้วยวิธี Term Frequency – Inverse Document Frequency และลดขนาดเวกเตอร์ด้วยการจัดสรรดีรีเคิลแฝง.....	45
ภาพที่ 4.7 ผลการสกัดชื่อเฉพาะ.....	46

บทที่ 1

บทนำ

1.1 ความเป็นมาและเหตุผลของโครงการ

ในปัจจุบันเครือข่ายสังคม (Social Network) ได้เข้ามามีบทบาทสำคัญทางด้านการติดต่อสื่อสารเป็นอย่างมาก เนื่องจากสามารถเพิ่มความสะดวกในการติดต่อกับผู้คนที่อยู่ห่างไกล และสามารถส่งสารเมื่ออยู่ในสถานการณ์ที่มีข้อจำกัดด้านเวลา

การติดต่อกับหน่วยงานหรือร้านค้าต่าง ๆ มักใช้ระบบส่งข้อความทันที (Instant Messaging) บนเครือข่ายสังคมเป็นตัวกลางในการติดต่อ เช่น Messenger ใน Facebook หรือ Direct Message (DM) ใน Twitter และ Instagram ซึ่ง Instant Messaging คือแอปพลิเคชันส่งข้อความทันทีไปยังบัญชีเป้าหมายซึ่งได้รับความนิยมทั้งบน Web Application และ Mobile Application เนื่องจากใช้งานได้ง่ายและสามารถติดต่อได้ตลอดเวลา

Chula MOOC [1] เป็นผู้ให้บริการการเรียนรู้ออนไลน์ ซึ่งมีช่องทางการติดต่อผ่านเพจ Chula MOOC ใน Facebook ที่เป็นเครือข่ายสังคมออนไลน์ที่ได้รับความนิยม ในปัจจุบัน Chula MOOC มีการเติบโตมากขึ้น ทำให้มีผู้สนใจและติดต่อผ่าน Messenger ใน Facebook มากขึ้นอย่างรวดเร็ว เนื่องจากจำนวนพนักงานที่น้อย ทำให้ไม่สามารถรองรับผู้ที่สนใจได้ ยิ่งไปกว่านั้น Chula MOOC สามารถเรียนรู้ได้ 24 ชั่วโมงไม่จำกัดสถานที่ ทำให้ผู้ดูแลไม่สามารถดูแลผู้ใช้งานหากอยู่นอกเวลาการทำงาน

แชทบอทจึงถูกนำมาปรับใช้เพื่อให้บริการ โดยการพัฒนาแชทบอทนั้นสามารถพัฒนาขึ้นเอง และสามารถบริหารทรัพยากรได้เอง หรือใช้งาน Platform สำเร็จรูปที่ให้บริการฟรี ซึ่งทั้งสองรูปแบบต้องฝึกแชทบอทให้เข้าใจถึงประโยคที่ผู้ติดต่อส่งมาด้วยตนเอง

การพัฒนาแชทบอทมีแนวทางการพัฒนาที่อาศัยกฎเกณฑ์หรือเงื่อนไขที่กำหนด (Rule-based approach) และอาศัยเจตนาของประโยคหรือข้อความ (Intent-based approach) [2] ซึ่งการพัฒนาโดยอาศัยเจตนาของประโยคหรือข้อความนั้นเป็นที่นิยมในปัจจุบันเนื่องจากมีการนำการประมวลผลภาษาธรรมชาติ (Natural Language Processing) มาใช้เพื่อให้แชทบอทนั้นสามารถเข้าใจและตอบสนองได้ถูกต้องและหลากหลายมากยิ่งขึ้น โดยอาศัยกฎเกณฑ์หรือเงื่อนไขที่กำหนดร่วมกับเจตนาของประโยคหรือข้อความทำให้สามารถตอบข้อความได้แม่นยำและหลากหลายมากขึ้น

การโต้ตอบของแชทบอทอาศัยการวิเคราะห์เจตนาของประโยคที่ผู้ใช้ป้อน และระบุการตอบสนองที่เหมาะสมกับเจตนา นั้น นอกจากนี้ ในประโยคที่ผู้ใช้ป้อนยังอาจมีชื่อเฉพาะ (Named

Entity) คือ นิพจน์ที่ใช้เรียกหรือระบุสิ่งใด ๆ เช่น ชื่อบุคคล ชื่อองค์กร ชื่อวิชา เป็นต้น และสามารถใช้เป็นคำสำคัญ (keyword) ในการค้นหาข้อมูลสำหรับการตอบสนองได้ [3] ซึ่งแชตบ็อตจะกำหนดให้การตอบสนองเปลี่ยนตามนิพจน์ที่ผู้ใช้ป้อน เช่น “คอร์สวิชาภาษาไทยจะเปิดเมื่อไหร่” มีเจตนาของข้อความ “การสอบถามรายวิชา” และมีชื่อเฉพาะเป็น “วิชาภาษาไทย” การตอบสนองข้อความของผู้ใช้จึงเป็นการหาข้อมูลเวลาเปิดของวิชาภาษาไทยมาตอบให้ผู้ใช้

Dialogflow เป็นระบบแชตบ็อตที่สามารถวิเคราะห์เจตนาของประโยคหรือข้อความที่ผู้ใช้ป้อนได้หากมีตัวอย่างของประโยคในเจตนาที่มากเพียงพอ แต่การสร้างตัวอย่างประโยคหรือการระบุเจตนาของประโยคตัวอย่างยังจำเป็นต้องทำโดยมนุษย์ ตัวอย่างประโยคนั้นสามารถรวบรวมได้จากประวัติสนทนาที่ผ่านมา แต่ยังคงจำเป็นต้องให้มนุษย์กำกับหรือระบุเจตนาของประโยคตัวอย่างอยู่ เช่น เจตนา “สอบถามกำหนดเปิดรายวิชา” จะต้องกำหนดตัวอย่างประโยค เช่น “คอร์สวิชาภาษาไทยเปิดเมื่อไหร่” “คอร์สBigDataจะเปิดอีกรอบเมื่อไหร่คะ” “วิชานำเสนอจะเปิดอีกตอนไหนคับ” เป็นต้น Dialogflow จะเรียนรู้ความคล้ายคลึงของประโยคที่กำหนดไว้ โดยอาศัยคุณลักษณะของประโยค เช่น ความยาวของประโยค ตำแหน่งคำศัพท์ในประโยค เป็นต้น ดังนั้นเมื่อผู้ใช้งานพิมพ์ข้อความมาว่า “คอร์สการตลาดจะเปิดเมื่อไหร่ครับ” จะถูกจัดให้อยู่ในเจตนา “สอบถามกำหนดเปิดรายวิชา” ซึ่งการสร้างแชตบ็อตขึ้นมาใหม่สำหรับแต่ละงานยังคงเป็นเรื่องที่ต้องใช้แรงงานมาก นอกจากนี้ การกำหนดชื่อเฉพาะก็ยังขึ้นกับลักษณะการสนทนานั้น ๆ ด้วย เช่น ในแชตบ็อตของ Chula MOOC ชื่อเฉพาะอาจเป็นรายวิชาที่เปิดสอน หรือรายวิชาที่ผู้เรียนสนใจ แต่ในแชตบ็อตของระบบซื้อขายสินค้า ชื่อเฉพาะอาจเป็นชื่อหรือรุ่นของสินค้า ด้วยเหตุนี้ ผู้พัฒนาจึงต้องการพัฒนาระบบจัดหมวดหมู่เจตนาของประโยคและสกัดชื่อเฉพาะจากประวัติสนทนาเพื่อนำไปใช้เป็นประโยคฝึกสอนในการฝึกสอนการวิเคราะห์เจตนาของประโยคในแชตบ็อต และนำไปป้อนให้กับ Dialogflow

1.2 วัตถุประสงค์ของโครงการ

1. พัฒนาระบบจัดหมวดหมู่เจตนาของประโยคจากประวัติสนทนา
2. พัฒนาระบบสกัดชื่อเฉพาะจากประวัติสนทนา
3. ประยุกต์ใช้ระบบแชตบ็อตสำเร็จรูป Dialogflow ให้ทำงานคู่กับ Messenger บน Facebook

1.3 ขอบเขตของโครงการ

1. ผู้ใช้งานต้องกำหนดจำนวนเจตนาของประโยคและชื่อหมวดหมู่เจตนาของประโยคเอง
2. แชตบ็อตยอมรับข้อความภาษาไทย และและคำภาษาอังกฤษบางคำ เช่น course certificate law เป็นต้น
3. แชตบ็อตพัฒนาขึ้นด้วย Dialogflow

4. แชตบ็อดสามารถประมวลผล 1 เจตนาต่อ 1 ข้อความ โดย 1 ข้อความคือการป้อนข้อมูลของผู้ใช้ 1 ครั้ง
5. แชตบ็อดสามารถจัดหมวดหมู่ของเจตนาของประโยคที่เกี่ยวข้องกับรายวิชา ใบประกาศนียบัตร การแจ้งปัญหา และคำถามที่พบบ่อย
6. แชตบ็อดสามารถตอบคำถามการสอบถามเกี่ยวกับรายวิชา ได้แก่
 1. เวลาเปิดรายวิชา เช่น “คอร์สวิชาภาษาฝรั่งเศสจะเปิดอีกเมื่อไหร่ครับ”
 2. รายวิชาที่เปิดให้ลงทะเบียน เช่น “ตอนนี้มีวิชาอะไรเปิดบ้าง”
 3. รายละเอียดของรายวิชา เช่น “คอร์สวิชาภาษาฝรั่งเศสมีกำหนดคะแนนยังไงครับ”
7. แชตบ็อดสามารถตอบคำถามเกี่ยวกับใบประกาศนียบัตร ได้แก่
 1. เงื่อนไขการได้รับใบประกาศนียบัตร เช่น “ทำยังไงถึงได้ใบประกาศนียบัตร”
 2. เวลาประกาศรายชื่อผู้ได้รับใบประกาศนียบัตร เช่น “วิชาภาษาฝรั่งเศสจะประกาศรายชื่อคนได้ใบประกาศนียบัตรเมื่อไหร่ครับ”
8. แชตบ็อดสามารถรับร้องเรียนการแจ้งปัญหา และสรุปเรื่องส่งไปยังผู้ดูแลเพจทางจดหมายอิเล็กทรอนิกส์เพื่อดำเนินการแก้ปัญหาต่อไป ได้แก่
 1. แจ้งการเปลี่ยนข้อมูลส่วนตัวใน myCourseVille และใบประกาศนียบัตร
 2. แจ้งปัญหาการตกหล่นของใบประกาศนียบัตร
 3. แจ้งปัญหาดาวนโหลดไฟล์ใบประกาศนียบัตรไม่ได้
9. แชตบ็อดสามารถตอบคำถามที่พบบ่อย (FAQ) ได้แก่
 1. Chula MOOC คืออะไร
 2. มีค่าใช้จ่ายหรือไม่
 3. Chula MOOC เหมาะกับใคร
 4. ลงทะเบียนเรียนอย่างไร

1.4 ขั้นตอนการดำเนินงาน

ระบบจัดหมวดหมู่เจตนาของประโยคและการสกัดชื่อเฉพาะจากประวัติสนทนาเพื่อฝึกสอนแชตบ็อด มีขั้นตอนการดำเนินงานดังต่อไปนี้

1. รวบรวมและวิเคราะห์ความต้องการจากเจ้าหน้าที่ศูนย์นวัตกรรมการศึกษา จุฬาลงกรณ์มหาวิทยาลัย
2. ศึกษาค้นคว้าข้อมูลเกี่ยวกับการหาเจตนาประโยคของผู้ใช้งาน
3. ศึกษาเครื่องมือ โปรแกรม และเทคนิคที่ใช้ในโครงการ
4. รวบรวมประวัติสนทนา เพื่อนำมาเป็นข้อมูลในการพัฒนาระบบ

5. กำหนดขอบเขต และวิธีการดำเนินงานของโครงการ
6. วิเคราะห์และออกแบบระบบ
7. พัฒนาระบบ
8. ทดสอบระบบ
9. ปรับปรุงและแก้ไขระบบ
10. สรุปผลการดำเนินงาน
11. จัดทำเอกสาร

ตารางที่ 1.1 ตารางเวลาการดำเนินงาน

วิธีการดำเนินงาน	ปี พ.ศ. 2561					ปี พ.ศ. 2562			
	ส.ค.	ก.ย.	ต.ค.	พ.ย.	ธ.ค.	ม.ค.	ก.พ.	มี.ค.	เม.ย.
1. รวบรวมและวิเคราะห์ความต้องการจากเจ้าหน้าที่ศูนย์นวัตกรรมการเรียนรู้ จุฬาลงกรณ์มหาวิทยาลัย									
2. ศึกษาค้นคว้าข้อมูลเกี่ยวกับการหาเจตนาประโยชน์ของผู้ใช้งาน									
3. ศึกษาเครื่องมือ โปรแกรม และเทคนิคที่ใช้ในโครงการ									
4. รวบรวมประวัติสนทนา เพื่อนำมาเป็นข้อมูลในการพัฒนาระบบ									
5. กำหนดขอบเขต และวิธีการดำเนินงานของโครงการ									
6. วิเคราะห์และออกแบบระบบ									
7. พัฒนาระบบ									
8. ทดสอบระบบ									

9. ปรับปรุงและแก้ไขระบบ									
10. สรุปผลการดำเนินงาน									
11. จัดทำเอกสาร									

1.5 ประโยชน์ที่คาดว่าจะได้รับ

ประโยชน์ต่อผู้ใช้

1. สามารถเพิ่มความมั่นใจให้ผู้ติดต่อว่าปัญหาที่ได้แจ้งมานั้นจะได้รับการช่วยเหลือหรือแก้ไข เนื่องจากตอบสนองได้รวดเร็ว
2. สามารถเพิ่มความสะดวกให้ผู้ดูแลเพจ โดยลดการตอบคำถามที่ถามบ่อย
3. สามารถลดระยะเวลาในการตอบคำถามและแก้ปัญหาให้ผู้ติดต่อ
4. สามารถนำระบบไปพัฒนาและใช้ร่วมกับแชทบ็อตอื่น ๆ ได้

ประโยชน์ต่อผู้พัฒนาระบบ

1. มีความรู้และความเข้าใจในกระบวนการพัฒนาซอฟต์แวร์
2. ฝึกกระบวนการคิดวิเคราะห์ ทำงานอย่างมีระเบียบแบบแผน มีความรับผิดชอบต่องาน และทำงานอย่างตรงต่อเวลา
3. มีความรู้ในการใช้ภาษาและเครื่องมือที่ใช้ในการพัฒนา ได้แก่ Python JavaScript Dialogflow Firebase และ SendGrid

1.6 โครงสร้างของรายงาน

บทที่ 2 จะกล่าวถึงทฤษฎี และซอฟต์แวร์ที่เกี่ยวข้องกับโครงการระบบจัดหมวดหมู่เจตนาของประโยคและการสกัดชื่อเฉพาะจากประวัติสนทนาเพื่อฝึกสอนแชทบ็อต

บทที่ 3 จะกล่าวถึงขั้นตอนการพัฒนา ซึ่งประกอบไปด้วยโครงสร้างระบบโดยรวม เครื่องมือที่ใช้ในการพัฒนา วิธีการเก็บข้อมูล การเตรียมข้อมูล การสร้างถ่วงคำ การสร้างเวกเตอร์ การลดมิติของข้อมูล การจัดหมวดหมู่ การสกัดชื่อเฉพาะ การวัดความไม่บริสุทธิ์ (Impurity) ของข้อมูล การส่งออกระบบ การสร้างไฟล์เจสันเพื่อนำเข้า Dialogflow การกำหนดเจตนา Dialogflow การกำหนดคำเฉพาะ Dialogflow การผสานการทำงาน Dialogflow เข้ากับ Messenger บน Facebook การผสานการทำงาน Dialogflow เข้ากับ Firebase และการส่งจดหมายอิเล็กทรอนิกส์ผ่าน SendGrid

บทที่ 4 จะกล่าวถึงผลการพัฒนา

บทที่ 5 จะกล่าวถึงข้อสรุปและปัญหาที่พบ

บทที่ 2

ทฤษฎี และซอฟต์แวร์ที่เกี่ยวข้อง

บทนี้จะกล่าวถึงทฤษฎี และซอฟต์แวร์ที่เกี่ยวข้องกับระบบจัดหมวดหมู่เจตนาของประโยค และการสกัดชื่อเฉพาะจากประวัติสนทนาเพื่อฝึกสอนแชทบ็อต

2.1 ซอฟต์แวร์ที่เกี่ยวข้อง

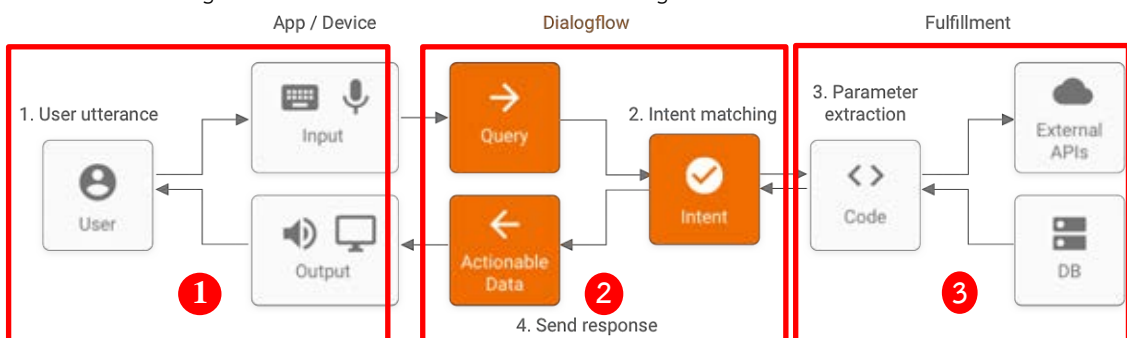
2.1.1 Dialogflow

Dialogflow [1] ให้บริการโดยบริษัทกูเกิล (Google) เป็นซอฟต์แวร์ที่ช่วยสร้างส่วนต่อประสาน (Interface) การสนทนาบนผลิตภัณฑ์หรือบริการต่าง ๆ เช่น Messenger Line Slack เป็นต้น โดยอาศัยการประมวลผลภาษาธรรมชาติมาใช้เพื่อให้สามารถเข้าใจข้อมูลนำเข้าที่เป็นภาษาธรรมชาติได้

เมื่อลองพิจารณาประโยคเช่น “อยากรู้เนื้อหาวิชาคะ” และ “วิชาคนรักสุนัขเรียนเกี่ยวกับอะไรบ้างคะ” จะเห็นว่าทั้งสองประโยคนั้นสื่อถึงความต้องการอยากรู้เนื้อหาวิชาคนรักสุนัขเหมือนกัน แต่มีรูปแบบของประโยคที่แตกต่างกัน ซึ่ง Dialogflow สามารถเข้าใจได้ว่าทั้งสองประโยคตัวอย่างนั้นมีเจตนาเดียวกันคือสอบถามข้อมูลรายวิชา จากการเรียนรู้ตัวอย่างประโยคที่ป้อนให้โดยอาศัยการประมวลผลภาษาธรรมชาติที่มีอยู่ใน Dialogflow

นอกจากนี้ Dialogflow ยังสามารถกำหนดค่าเฉพาะลงไปเพื่อใช้เป็นพารามิเตอร์ (parameter) สำหรับการใส่ประโยชน์จากฐานข้อมูลหรือส่วนต่อประสานโปรแกรมประยุกต์ (API) ได้โดยการเขียนโปรแกรมในส่วนเติมเต็ม (Fulfillment) ของ Dialogflow

โครงการนี้จึงใช้ Dialogflow เป็นตัวสร้างแชทบ็อตโดยประสานการทำงานกับ Messenger ของเพจ CHULA MOOC โดย Dialogflow มีการทำงานดังนี้



ภาพที่ 2.1 การทำงานของ Dialogflow

ที่มา : <https://dialogflow.com/docs/images/intro/fulfillment-diagram.png>

จากภาพที่ 2.1 สามารถอธิบายการทำงานของ Dialogflow ได้เป็น 3 ส่วนดังนี้

1. ส่วนเชื่อมต่อประสาน : Dialogflow จะสามารถเชื่อมต่อกับแพลตฟอร์มต่าง ๆ ภายนอกได้ เช่น Messenger Line เป็นต้น โดยอาศัยส่วน Integration ของ Dialogflow สำหรับรับข้อความจากผู้ใช้
2. ส่วนประมวลผล : ส่วนนี้จะจับคู่ข้อความที่ผู้ใช้ส่งมายังแช็ตบ๊อตนั้นเข้ากับเจตนาที่เหมาะสมกันจากที่ได้ป้อนข้อมูลไว้ในแต่ละเจตนา ซึ่งองค์ประกอบของแต่ละเจตนา นั้นประกอบด้วย
 - วลีการฝึกสอน (Training phrases) : ส่วนสำหรับป้อนข้อความ ตัวอย่างของเจตนา นั้น ๆ เพื่อให้ Dialogflow ได้เรียนรู้จากข้อความ ตัวอย่าง และทำให้สามารถจับคู่ข้อความที่ผู้ใช้ส่งมายังแช็ตบ๊อตได้
 - การดำเนินการและพารามิเตอร์ (Action and Parameters) : ส่วน สำหรับกำหนดค่าเฉพาะที่ปรากฏอยู่ในข้อความตัวอย่างในวลีการ ฝึกสอนให้เป็นพารามิเตอร์สำหรับใช้ประโยชน์เพิ่มเติมผ่านส่วนเพิ่มเติม

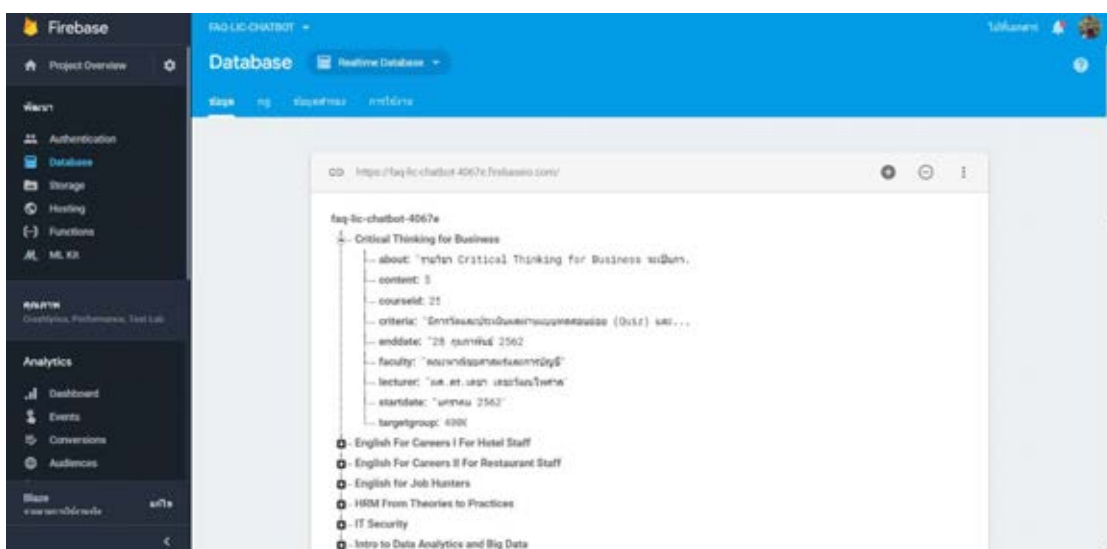
ในส่วนนี้สามารถกำหนดได้ว่าจะให้เจตนาที่ส่งข้อความไปยังส่วนเพิ่มเติมหรือไม่ ถ้า กำหนดไม่ส่งต่อไปยังส่วนเพิ่มเติมแล้ว Dialogflow จะส่งการตอบสนอง (Response) คืนไปยังส่วนเชื่อมต่อประสานเพื่อส่งให้กับผู้ใช้งาน
3. ส่วนเพิ่มเติม : ส่วนเพิ่มเติมเป็นระบบภายนอกจาก Dialogflow ซึ่งมีการเชื่อมต่อกันเพื่อแลกเปลี่ยนข้อมูล เช่น การเรียกข้อมูลจากฐานข้อมูลเพื่อนำมาประกอบกับ คำตอบให้ Dialogflow นั้นสามารถตอบโต้กับผู้ใช้งานอย่างมีประสิทธิภาพมากขึ้น การส่งจดหมายอิเล็กทรอนิกส์เพื่อรายงานข้อมูลให้กับผู้ดูแลระบบ เมื่อการทำงานของระบบภายนอกเสร็จสิ้นแล้วจะส่งคำตอบ คืนให้กับส่วนประมวลผลและส่งต่อ ไปยังผู้ใช้งาน

2.1.2 Firebase

Firebase [10] คือ แพลตฟอร์มสำเร็จรูปที่ให้บริการ Backend-as-a-Service (BaaS) เช่น ฐานข้อมูลแบบทันที (Realtime Database) ระบบยืนยันตัวตน

(Authentication Service) ซึ่งค่าใช้จ่ายขึ้นอยู่กับเงื่อนไขการใช้งานและปริมาณการใช้งาน อีกทั้ง Firebase สามารถติดตามใช้งานร่วมกับ Dialogflow ได้

โครงการนี้ใช้ Firebase เพื่อเป็นฐานข้อมูลแบบ NoSQL บนคลาวด์ (Cloud) ที่จัดเก็บข้อมูลรายละเอียดวิชาที่เปิดสอนใน Chula MOOC ในฐานข้อมูลแบบทันทีในรูปแบบของเจสัน (JSON) ดังภาพที่ 2.2 และเชื่อมต่อกับ Dialogflow โดยเรียกข้อมูลจาก Firebase ผ่าน Fulfillment ของ Dialogflow



ภาพที่ 2.2 ฐานข้อมูลบน Firebase

2.1.3 SendGrid

SendGrid ทำหน้าที่ให้บริการ SMTP (Simple Mail Transfer Protocol) ซึ่งเป็นโปรโตคอลที่ใช้ในการส่งจดหมายอิเล็กทรอนิกส์ในเครือข่ายอินเทอร์เน็ตบนคลาวด์

โครงการนี้ใช้ SendGrid เพื่อเป็นส่วนต่อประสานโปรแกรมประยุกต์ที่เชื่อมต่อกับ Dialogflow โดยเรียกใช้งานผ่านส่วนเติมเต็มของ Dialogflow

2.2 ทฤษฎีที่เกี่ยวข้อง

2.2.1 เทคนิคการจัดกลุ่มข้อมูลแบบเคมีน (K-mean Clustering)

เทคนิคการจัดกลุ่มข้อมูลแบบเคมีนต้องการที่จะแบ่งข้อมูลที่มีจำนวน n จุดให้ออกเป็น k กลุ่ม โดยการคำนวณระยะห่างระหว่างจุดข้อมูล มีขั้นตอนดังต่อไปนี้

1. สุ่มจุดสมมติขึ้นมาจำนวน k จุด ให้เป็นจุดกลางข้อมูลของกลุ่มที่ k

2. คำนวณระยะทางระหว่างจุดข้อมูลและจุดสมมติ ทั้ง k จุดซึ่งมีวิธีการคำนวณ 3 วิธี ได้แก่

- a. คำนวณระยะทางแบบชิตีบล็อก (City block) ดังสมการที่ 2.1

$$d(X, Y) = \sum_{j=1}^k |x_j - y_j|$$

โดย X และ Y คือ เวกเตอร์ที่มีขนาด k
 x_j คือ ข้อมูลตัวที่ j ของเวกเตอร์ x
 y_j คือ ข้อมูลตัวที่ j ของเวกเตอร์ y

สมการที่ 2.1 สมการคำนวณระยะทางแบบชิตีบล็อก

- b. คำนวณระยะทางแบบยูคลิด (Euclidean distance) ดังสมการที่ 2.2

$$d(X, Y) = \sqrt{\sum_{j=1}^k (x_j - y_j)^2}$$

โดย X และ Y คือ เวกเตอร์ที่มีขนาด k
 x_j คือ ข้อมูลตัวที่ j ของเวกเตอร์ x
 y_j คือ ข้อมูลตัวที่ j ของเวกเตอร์ y

สมการที่ 2.2 สมการคำนวณระยะทางแบบยูคลิด

- c. คำนวณระยะทางแบบเชบีเชฟ (Chebychev Distance) ดังสมการที่ 2.3

$$d(X, Y) = \left(\sum_{j=1}^k |x_j - y_j|^p \right)^{1/p}$$

โดย X และ Y คือ เวกเตอร์ที่มีขนาด k
 x_j คือ ข้อมูลตัวที่ j ของเวกเตอร์ x
 y_j คือ ข้อมูลตัวที่ j ของเวกเตอร์ y
 p คือ จำนวนจริงใด ๆ ที่มีค่าไม่เท่ากับ 0

สมการที่ 2.3 สมการคำนวณระยะทางแบบเชบีเชฟ

3. คำนวณจุดกลาง (Centroid) ของกลุ่มข้อมูล k กลุ่ม ดังสมการที่ 2.4

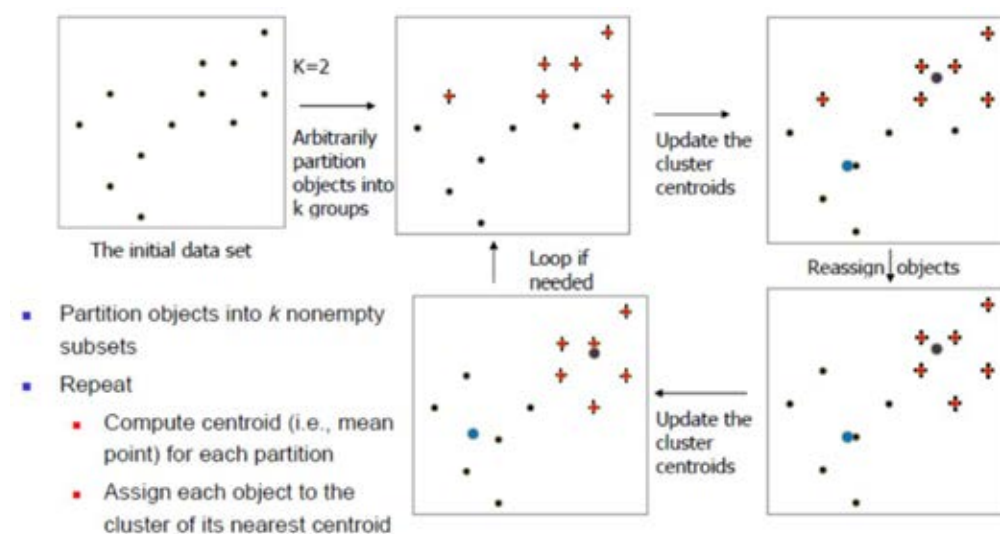
$$m_i = \frac{1}{|S_i|} \sum_{x_j \in S_i} x_j$$

โดย m_i คือ จุดกลางข้อมูลกลุ่มที่ i
 $|S_i|$ คือ จำนวนสมาชิกภายในกลุ่มที่ i
 x_j คือ ข้อมูลภายในกลุ่มที่ i

สมการที่ 2.4 สมการคำนวณจุดกลาง

4. กำหนดกลุ่มข้อมูลใหม่
 5. ทำซ้ำข้อ 2. - 4. จนกว่าจุดกลางจะไม่มีเปลี่ยนแปลง

ซึ่งมีภาพรวมขั้นตอนการทำเทคนิคการจัดกลุ่มข้อมูลแบบเคมีน ดังภาพที่ 2.3



ภาพที่ 2.3 ขั้นตอนการทำเทคนิคการจัดหมวดหมู่ข้อมูลแบบเคมีน (K-mean Clustering)

ที่มา: <https://wipawanblog.files.wordpress.com/2014/06/chapter-8-clustering-k-means.pdf>

สำหรับโครงการนี้จะใช้ชุดคำสั่งจาก Scikit-learn ในการจัดกลุ่มข้อมูลแบบเคมีนซึ่งมีการคำนวณระยะทางแบบยุคลิด

2.2.2 ถุงคำ (Bag-of-Words)

เทคนิค Bag-Of-Words [2,12] เป็นวิธีการที่แทนข้อมูลด้วยเวกเตอร์จากคำทั้งหมดที่ปรากฏในเอกสาร เช่น มีประโยค 2 ประโยคดังนี้ “ใบcerพิมพ์นามสกุลผิดแก้ไขได้ยังไง” และ “คอร์สนำเสนอผลงานราคาเท่าไหร่” เมื่อผ่านการตัดคำย่อยจะสามารถแทนข้อมูลด้วยเวกเตอร์ดังนี้ [ใบ, cer, พิมพ์, นามสกุล, ผิด, แก้ไข, ยังไง, คอร์ส, นำเสนอ, ผลงาน, ราคา, เท่าไหร่]

2.2.3 Term Frequency - Invert Document Frequency (TF-IDF)

Term Frequency (TF) [13] เป็นการวัดความสำคัญของคำโดยนับการปรากฏของคำที่สนใจ คำที่สนใจปรากฏบ่อยแสดงว่ามีความสำคัญมาก ซึ่งมีวิธีการคำนวณดังสมการที่

2.5

$$tf(t, d) = 0.5 + 0.5 \cdot \frac{f_{t,d}}{\max\{f_{t',d} : t' \in d\}}$$

โดย $f_{t,d}$ คือ ความถี่ของคำ t ที่ปรากฏในเอกสาร d
 $\max\{f_{t',d} : t' \in d\}$ คือ ความถี่ที่มากที่สุดที่ t' ปรากฏในเอกสาร d

สมการที่ 2.5 สมการ Term Frequency

Inverse Document Frequency (IDF) เป็นการวัดความสำคัญของคำโดยสนใจว่า คำปรากฏบ่อยในเอกสารอื่น ๆ หรือไม่ ยิ่งปรากฏบ่อยแสดงว่ามีความสำคัญน้อย ซึ่งมีวิธีการคำนวณดังสมการที่ 2.6

$$idf(t, D) = \log \frac{N}{1 + |\{d \in D : t \in d\}|}$$

โดย N คือ จำนวนเอกสารทั้งหมด
 $|\{d \in D : t \in d\}|$ คือ จำนวนเอกสารที่มีคำ t ปรากฏ เพื่อป้องกันค่าเป็น 0 จึงมีการเพิ่มค่า 1 ให้กลายเป็น $1 + |\{d \in D : t \in d\}|$

สมการที่ 2.6 สมการ Inverse Document Frequency

2.2.4 การวิเคราะห์องค์ประกอบหลัก (Principal Component Analysis: PCA)

การวิเคราะห์องค์ประกอบหลัก [9] คือ กระบวนการลดความซับซ้อนของข้อมูล ในกรณีที่มีข้อมูลมีขนาดใหญ่มาก การคำนวณด้วยวิธีปกติอาจทำให้เสียเวลามาก การวิเคราะห์องค์ประกอบหลักจะเข้ามาช่วยลดระยะเวลาการคำนวณตรงนี้ได้ สำหรับในบทนี้จะยกตัวอย่างข้อมูลและแสดงขั้นตอนของการวิเคราะห์องค์ประกอบหลัก ดังนี้

กำหนดให้ข้อมูลมีลักษณะดังตารางที่ 2.1

ตารางที่ 2.1 ตารางแสดงลักษณะของข้อมูล

ประโยค	คอร์ส	เปิด	ข้อมูล
คอร์สภาษาอังกฤษเปิดเมื่อไร	1	1	0
อยากรู้ข้อมูลวิชาภาษารัสเซีย	0	0	1
ใบ cer วิชาภาษารัสเซียจะได้เมื่อไร	0	0	0

สมมติว่าต้องการลดขนาดมิติข้อมูลให้เหลือ 2 มิติจาก 3 มิติ

1. นอร์มัลไลเซชัน (Normalization) ข้อมูลด้วยค่าเฉลี่ยเลขคณิตของแต่ละคอลัมน์ เพื่อให้ข้อมูลตั้งต้นมีความแตกต่างกันน้อยลงดังสมการที่ 2.7

$$x_{ij} = x_{ij} - \bar{x}_i$$

โดย	i	คือ คอลัมน์ที่ i
	x_{ij}	คือ ข้อมูลตัวที่ j ของคอลัมน์ที่ i
	\bar{x}_i	คือ ค่าเฉลี่ยเลขคณิตของแต่ละคอลัมน์

สมการที่ 2.7 สมการนอร์มัลไลเซชัน

ผลลัพธ์ที่ได้ คือ ข้อมูลชุดใหม่ที่มีผ่านการนอร์มัลไลเซชันแล้ว

จากข้อมูลในตารางที่ 2.1 เราสามารถเขียนในรูปแบบของเมทริกซ์ได้ดังนี้

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

กำหนดให้เมทริกซ์ดังกล่าวคือ เมทริกซ์ A

2. คำนวณเมทริกซ์ความแปรปรวนร่วมเกี่ยว (Covariance Matrix) ของเมทริกซ์ A จะได้เมทริกซ์ผลลัพธ์เป็นเมทริกซ์จัตุรัสซึ่งสามารถนำไปคำนวณเวกเตอร์ลักษณะเฉพาะและค่าลักษณะเฉพาะในขั้นต่อไปได้

เมทริกซ์ความแปรปรวนร่วมเกี่ยว คือ

$$C^{m \times n} = (c_{i,j}, c_{i,j} = cov(dim_i, dim_j))$$

โดย	$C^{m \times n}$	คือ เมทริกซ์ขนาด $m \times n$
	$c_{i,j}$	คือ สมาชิกตำแหน่งที่ i, j ภายในเมทริกซ์
	dim_i	คือ คอลัมน์ที่ i ของเมทริกซ์ A
	dim_j	คือ คอลัมน์ที่ j ของเมทริกซ์ A

สมการที่ 2.8 สมการเมทริกซ์แปรปรวนร่วมเกี่ยว

ค่าความแปรปรวนร่วมเกี่ยวสามารถคำนวณได้ดังสมการที่ 2.9

$$cov(x, y) = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{(n - 1)}$$

โดย	x_i	คือ สมาชิกตัวที่ i ของคอลัมน์ x
	\bar{x}	คือ ค่าเฉลี่ยเลขคณิตของคอลัมน์ x
	y_i	คือ สมาชิกตัวที่ i ของคอลัมน์ y
	\bar{y}	คือ ค่าเฉลี่ยเลขคณิตของคอลัมน์ y
	n	คือ จำนวนสมาชิกของคอลัมน์ x หรือ y

สมการที่ 2.9 สมการค่าความแปรปรวนร่วมเกี่ยว

จากเมทริกซ์ A สามารถคำนวณค่าได้ดังตารางที่ 2.2

ตารางที่ 2.2 แสดงค่าที่คำนวณจากเมทริกซ์ A

X	Y	Z	$X - \bar{X}$	$Y - \bar{Y}$	$Z - \bar{Z}$	$(X - \bar{X})(Y - \bar{Y})$	$(X - \bar{X})(Z - \bar{Z})$	$(Y - \bar{Y})(Z - \bar{Z})$
1	1	0	0.7776	0.7776	-0.3334	0.6046	-0.2592	-0.2592
0	0	1	-0.3334	-0.3334	0.7776	0.1111	-0.2592	-0.2592
0	0	0	-0.3334	-0.3334	-0.3334	0.1111	0.1111	0.1111

จากตารางที่ 2.2 แสดงค่าการคำนวณค่า $x - \bar{x}$ ของแต่ละคอลัมน์เพื่อใช้คำนวณค่าความแปรปรวนร่วม

ยกตัวอย่างการคำนวณในตำแหน่งที่ $c_{1,1}$ ของเมทริกซ์ C

จากสมการที่ 2.8 เราจะได้ว่า $c_{1,1}$ คือ $cov(dim_1, dim_1)$ ซึ่ง dim_1 คือ $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

จากสมการที่ 2.9 เรารู้ว่า x คือ dim_1 และรู้ว่า y คือ dim_1 ดังนั้นจะคำนวณค่า cov ได้ดังนี้

$$cov(x, y) = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{(n - 1)}$$

นำค่าจากในตารางที่ 2.2 ที่คำนวณไว้แล้ว

เมื่อ $i = 1$ จะได้ $x_i - \bar{x} = 0.7776$ และ $y_i - \bar{y} = 0.7776$

เมื่อ $i = 2$ จะได้ $x_i - \bar{x} = -0.3334$ และ $y_i - \bar{y} = -0.3334$

เมื่อ $i = 3$ จะได้ $x_i - \bar{x} = -0.3334$ และ $y_i - \bar{y} = -0.3334$

นำมาแทนค่าในสมการที่ 2.9 จะได้

$$= \frac{(0.7776 * 0.7776) + (-0.3334 * -0.3334) + (-0.3334 * -0.3334)}{(3 - 1)}$$

$$= \frac{0.8269}{2}$$

$$\therefore cov(x, y) = 0.4134$$

จะได้ว่า ตำแหน่งที่ $c_{1,1}$ ของเมทริกซ์ C มีค่า $cov = 0.4134$

เมื่อคำนวณค่า cov ครบทุกตำแหน่งแล้วจะได้เมทริกซ์ C ดังนี้

$$C = \begin{bmatrix} 0.4134 & 0.4134 & -0.2036 \\ 0.4134 & 0.4134 & -0.2036 \\ -0.2036 & -0.2036 & 0.4134 \end{bmatrix}$$

กำหนดให้เมทริกซ์ความแปรปรวนร่วมเกี่ยวกับได้คือ เมทริกซ์ cov

3. คำนวณเวกเตอร์ลักษณะเฉพาะ (Eigen Vector) และค่าลักษณะเฉพาะ (Eigen Value) ของเมทริกซ์ cov

เมื่อได้เมทริกซ์ผลลัพธ์หรือเมทริกซ์ cov แล้ว จะมีคุณสมบัติของเวกเตอร์ลักษณะเฉพาะและค่าลักษณะเฉพาะดังสมการที่ 2.10

$$CX = \lambda X$$

โดย C คือ เมทริกซ์ cov ของเมทริกซ์ A

X คือ เมทริกซ์เวกเตอร์ลักษณะเฉพาะ

l คือ เมทริกซ์ค่าลักษณะเฉพาะ

สมการที่ 2.10 คุณสมบัติของเวกเตอร์ลักษณะเฉพาะและค่าลักษณะเฉพาะ

จากเมทริกซ์ COV สามารถคำนวณเวกเตอร์ลักษณะเฉพาะและค่าลักษณะเฉพาะตามวิธีคำนวณใน [9] ซึ่งจะได้ผลลัพธ์ดังนี้

$$\text{เวกเตอร์ลักษณะเฉพาะ} = \begin{bmatrix} -0.6291 & -0.3228 & 0.7071 \\ -0.6291 & -0.3228 & -0.7071 \\ 0.4565 & -0.8897 & 0 \end{bmatrix}$$

$$\text{ค่าลักษณะเฉพาะ} = \begin{bmatrix} 0.9745 \\ 0.2656 \\ 0 \end{bmatrix}$$

ค่าลักษณะเฉพาะแปรผันตรงกับความสำเร็จของข้อมูล

4. เลือกเวกเตอร์ลักษณะเฉพาะที่มีค่าลักษณะเฉพาะสูงสุด 2 ค่า เนื่องจากต้องการลดขนาดให้เหลือเพียง 2 มิติ เพราะจำนวนเวกเตอร์ลักษณะเฉพาะที่เลือกคือขนาดของข้อมูลชุดใหม่ที่ได้ ซึ่งจะสามารถคำนวณเมทริกซ์แทนข้อมูลชุดใหม่ได้จากสมการที่ 2.11

$$D = AX$$

โดย D คือ เมทริกซ์แทนข้อมูลชุดใหม่

A คือ เมทริกซ์แทนข้อมูลชุดเก่า

X คือ เมทริกซ์เวกเตอร์ลักษณะเฉพาะที่มีขนาด $i \times k$
เมื่อ k คือจำนวนเวกเตอร์ลักษณะเฉพาะที่สนใจ

สมการที่ 2.11 สมการคำนวณ

$$D = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} -0.6291 & -0.3228 \\ -0.6291 & -0.3228 \\ 0.4565 & -0.8897 \end{bmatrix}$$

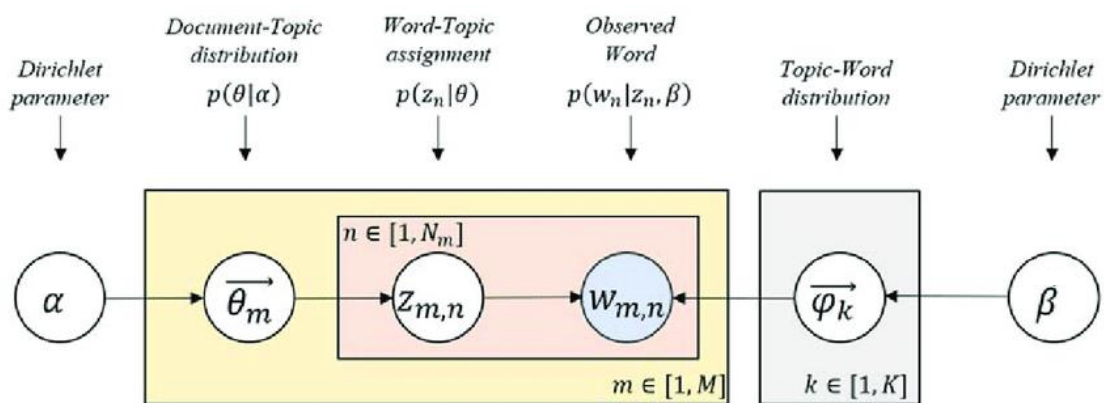
จากผลลัพธ์ที่ได้ เราจะได้เมทริกซ์ชุดใหม่ดังนี้

$$\begin{bmatrix} -1.2582 & 0 \\ 0.4565 & 0 \\ 0 & 0 \end{bmatrix}$$

การวิเคราะห์องค์ประกอบหลักนั้น นอกจากจะสามารถลดขนาดของมิติได้แล้วยังสามารถบ่งบอกถึงความสำคัญของมิติของข้อมูลได้อีกด้วย แต่ทั้งนี้การลดขนาดมิติข้อมูลอาจจะทำให้ผลลัพธ์การทำงานของระบบมีประสิทธิภาพลดลงได้

2.2.5 การจัดสรรของดีรีเคลแฝง (Latent Dirichlet allocation: LDA)

การจัดสรรของดีรีเคลแฝงนั้นใช้ในการจัดหมวดหมู่เอกสาร (Document) [10] โดยมีหลักการว่าเอกสารเกิดจากการปรากฏหัวข้อ (Topic) ในอัตราส่วนที่ต่างต่างกัน แต่ละหัวข้อจะประกอบไปด้วยคำ (Word) โดยแต่ละคำจะสามารถปรากฏได้ในหลายหัวข้อด้วยความน่าจะเป็นที่ต่างกันอย่างเช่นกัน ดังนั้น หากทราบถึงลักษณะการปรากฏหัวข้อในเอกสาร จะทำให้เราสามารถคาดเดาหัวข้อที่เอกสารควรถูกจัดไว้ได้ สามารถเขียนการจัดสรรของดีรีเคลแฝงให้อยู่ในแบบจำลองแผนภาพ (Graphical model) ได้ดังภาพที่ 2.4



ภาพที่ 2.4 แบบจำลองแผนภาพของการจัดสรรของดีรีเคลแฝง

ที่มา: Lee J, Kang J-H, Jun S, Lim H, Jang D, Park S. Ensemble Modeling for Sustainable Technology Transfer. Sustainability. 2018; 10(7):2278.

- โดย K คือ จำนวนหัวข้อทั้งหมด
- M คือ จำนวนเอกสารทั้งหมด
- N_m คือ จำนวนหัวข้อทั้งหมดที่ปรากฏในเอกสารที่ m
- α คือ Dirichlet parameter ซึ่งเป็นตัวแปรที่ต้องกำหนดค่าเอง ซึ่งกำหนด ความหลากหลายของหัวข้อที่จะปรากฏในเอกสาร ยิ่งค่า α สูงความน่าจะเป็นที่เอกสารจะมีหัวข้อปรากฏในเอกสารจะยิ่งมาก
- β คือ Dirichlet parameter ซึ่งเป็นตัวแปรที่ต้องกำหนดค่าเอง ซึ่งกำหนด ความหลากหลายของคำที่จะปรากฏในแต่ละหัวข้อ ยิ่งค่า β สูงความน่าจะเป็นที่คำจะปรากฏในหลายหัวข้อจะยิ่งมาก
- θ_m คือ การกระจายตัวของหัวข้อในเอกสาร m

- φ_k คือ การกระจายตัวของคำในหัวข้อที่ k
 $z_{m,n}$ คือ หัวข้อของคำในเอกสารที่ m ในหัวข้อที่ n
 $w_{m,n}$ คือ คำที่เราสนใจในหัวข้อที่ n ในเอกสารที่ m

โครงการนี้จะนำความสามารถในการจัดหมวดหมู่เอกสารมาใช้ในการลดขนาดมิติของข้อมูล โดยให้จำนวนหัวข้อที่เราสนใจคือขนาดของมิติลัพท์ที่เราอยากได้ ดังตัวอย่างต่อไปนี้

จากข้อมูลในตารางที่ 2.1 กำหนดให้มีหัวข้อแฝง 2 หัวข้อ นำข้อมูลดังกล่าวมาผ่านการจัดสรรของตรีเคิลแฝงจะได้ผลลัพธ์ดังตารางที่ 2.2

ตารางที่ 2.2 ตารางแสดงข้อมูลที่ผ่านการจัดสรรของตรีเคิลแฝง

ประโยค	หัวข้อแฝงที่ 1	หัวข้อแฝงที่ 2
คอร์สภาษาอังกฤษเปิดเมื่อไร	0.71	0.29
อยากรู้ข้อมูลวิชาภาษารัสเซีย	0.46	0.54
ใบ cer วิชาภาษารัสเซียจะได้เมื่อไร	0.91	0.09

จากตารางที่ 2.2 ตัวเลขที่ปรากฏคือความน่าจะเป็นที่ประโยคจะอยู่ในหัวข้อใด ๆ ซึ่งจะเห็นว่ามิติของข้อมูลถูกแทนที่ด้วยจำนวนหัวข้อที่เราสนใจ และผลลัพธ์คือข้อมูลที่ถูกลดมิติ

2.2.6 ความไม่บริสุทธิ์ของข้อมูล

ในโครงการนี้จะวัดความไม่บริสุทธิ์ของข้อมูลเพื่อนำไปวัดและเปรียบเทียบประสิทธิภาพของระบบด้วยค่าความแปรปรวนข้อมูล (Entropy) โดยค่าความแปรปรวนจะสูงขึ้นเมื่อภายในกลุ่มมีจำนวนสมาชิกหลายหมวดหมู่ปะปนอยู่ ซึ่งสามารถคำนวณได้จากสมการที่ 2.12

$$S = -\sum_{i=1}^n P_i \log P_i$$

- โดย n คือ จำนวนหมวดหมู่ทั้งหมด
 i คือ หมวดหมู่ที่ i
 P_i คือ ค่าความน่าจะเป็นของหมวดหมู่ที่ i

สมการที่ 2.12 สมการหาค่าความไม่บริสุทธิ์ของข้อมูล

2.2.7 การวัดผลประสิทธิภาพของระบบ

ในโครงการนี้จะใช้เมทริกซ์คอนฟิวชัน (Confusion Matrix) ในการแสดงผลการทำงานของระบบซึ่งเมทริกซ์คอนฟิวชันมีลักษณะดังตารางที่ 2.3

ตารางที่ 2.3 เมทริกซ์คอนฟิวชัน

Actual	Predict	
	True	False
True	True Positive (TP)	False Positive (FP)
False	False Negative (FN)	True Negative (TN)

จากตารางที่ 2.3 มีค่าต่าง ๆ ในตารางดังนี้

1. ค่า True Positive คือ จำนวนผลลัพธ์ที่ระบบให้คำตอบว่าเป็นจริงและคำตอบที่กำหนดไว้เป็นจริง
2. ค่า False Negative คือ จำนวนผลลัพธ์ที่ระบบให้คำตอบว่าเป็นเท็จแต่คำตอบที่กำหนดไว้เป็นจริง
3. ค่า False Positive คือ จำนวนผลลัพธ์ที่ระบบให้คำตอบว่าเป็นจริงแต่คำตอบที่กำหนดไว้เป็นเท็จ
4. ค่า True Negative คือ จำนวนผลลัพธ์ที่ระบบให้คำตอบว่าเป็นเท็จและคำตอบที่กำหนดไว้เป็นเท็จ

จากค่าในตาราง สามารถคำนวณเพื่อวัดประสิทธิภาพระบบได้ดังนี้

1. ความถูกต้องของระบบ (Accuracy) คำนวณจากจำนวนคำตอบที่ระบบตอบตรงกับที่ตั้งไว้เทียบกับจำนวนคำตอบทั้งหมด สามารถคำนวณได้จากสมการที่ 2.13

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN}$$

สมการที่ 2.13 สมการความถูกต้องของระบบ

2. ความแม่นยำของระบบ (Precision) คำนวณจากจำนวนคำตอบที่ระบบตอบว่าจริงและค่าที่กำหนดไว้เป็นจริงเทียบกับจำนวนคำตอบที่ระบบตอบว่าเป็นจริง สามารถคำนวณได้จากสมการที่ 2.14

$$Precision = \frac{TP}{TP + FP}$$

สมการที่ 2.14 สมการความแม่นยำของระบบ

3. ค่าเรียกคืน (Recall) คำนวณจากจำนวนคำตอบที่ระบบตอบว่าจริงและค่าที่กำหนดไว้เป็นจริง เทียบกับจำนวนคำตอบที่กำหนดว่าเป็นจริง สามารถคำนวณได้จากสมการที่ 2.15

$$Recall = \frac{TP}{TP + FN}$$

สมการที่ 2.15 สมการค่าเรียกคืน

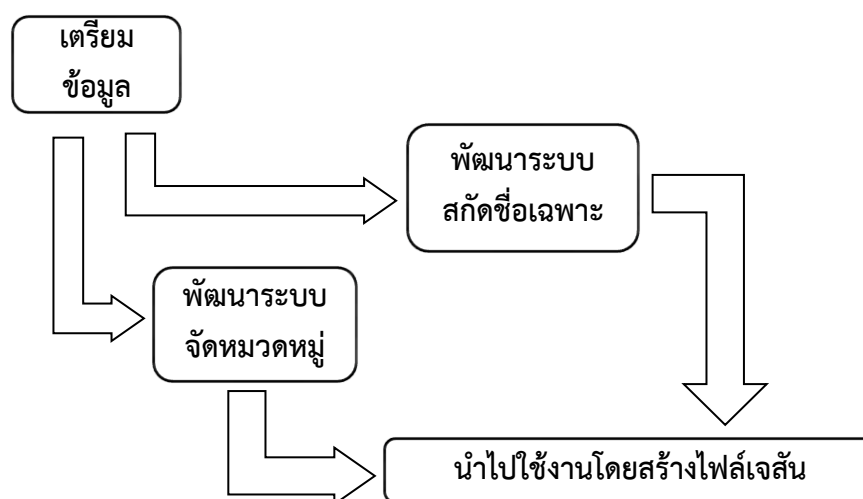
ค่าที่คำนวณได้ทั้ง 3 ค่าจะมีค่าสูงสุดที่ 1 และต่ำสุดที่ 0 ระบบจะมีประสิทธิภาพมากเมื่อคำนวณค่าได้ผลลัพธ์ที่เข้าใกล้ 1

บทที่ 3

วิธีการพัฒนา

ในบทนี้จะกล่าวถึงวิธีการพัฒนา ซึ่งประกอบไปด้วยโครงสร้างระบบโดยรวม เครื่องมือที่ใช้ในการพัฒนา วิธีการเก็บข้อมูล การเตรียมข้อมูล การสร้างถ่วงคำ การสร้างเวกเตอร์ การลดมิติของข้อมูล การจัดหมวดหมู่ การสกัดชื่อเฉพาะ การวัดความไม่บริสุทธิ์ของข้อมูล การส่งออกระบบ การสร้างไฟล์เจสันเพื่อนำเข้า Dialogflow การกำหนดเจตนา Dialogflow การกำหนดคำเฉพาะ Dialogflow การผสมผสานการทำงาน Dialogflow เข้ากับ Messenger บน Facebook การผสมผสานการทำงาน Dialogflow เข้ากับ Firebase และการส่งจดหมายอิเล็กทรอนิกส์ผ่าน SendGrid

3.1 โครงสร้างระบบโดยรวม



ภาพที่ 3.1 โครงสร้างระบบโดยรวม

จากภาพที่ 3.1 จะพัฒนาระบบโดยแบ่งออกเป็น 4 ส่วนดังนี้

1. เตรียมข้อมูล โดยมีขั้นตอนดังนี้
 1. เก็บข้อมูลซึ่งคือประวัติสนทนาใน Facebook เพจ Chula MOOC ซึ่งจะอธิบายในหัวข้อ 3.3
 2. เตรียมข้อมูลโดยนำประวัติสนทนามาผ่านกระบวนการต่าง ๆ เพื่อให้ข้อมูลสามารถนำไปใช้ได้ ซึ่งจะอธิบายในหัวข้อ 3.4
 3. สร้างถ่วงคำจากประโยคที่เตรียมไว้ ซึ่งจะอธิบายในหัวข้อ 3.5

4. สร้างเวกเตอร์แทนข้อมูลโดยมีวิธีการสร้างเวกเตอร์ 3 แบบเพื่อนำมาเปรียบเทียบและวิเคราะห์ ซึ่งจะอธิบายในหัวข้อ 3.6
2. พัฒนาระบบสกัดชื่อเฉพาะ ขั้นตอนนี้จะพัฒนาระบบสกัดชื่อเฉพาะโดยวิธีการเทียบคำที่ยาวที่สุด (Longest Matching) ซึ่งจะอธิบายในหัวข้อ 3.7
3. พัฒนาระบบจัดหมวดหมู่ โดยมีขั้นตอนดังนี้
 1. ลดมิติข้อมูล นำเวกเตอร์แทนข้อมูลที่เราสร้างไว้มาลดขนาดด้วย 2 แบบเพื่อเปรียบเทียบและวิเคราะห์ ซึ่งจะอธิบายในหัวข้อ 3.8
 2. จัดหมวดหมู่ข้อมูล ของเวกเตอร์ที่ไม่ได้ลดขนาด เวกเตอร์ที่ลดขนาดด้วยวิธีการวิเคราะห์องค์ประกอบหลัก และ การจัดสรรตรีเคิลแฝง ซึ่งจะอธิบายในหัวข้อ 3.9
 3. วัดความไม่บริสุทธิ์ของผลลัพธ์การจัดหมวดหมู่ของเวกเตอร์ที่ไม่ได้ลดขนาดทั้ง 3 รูปแบบเพื่อเปรียบเทียบประสิทธิภาพ และวัดความไม่บริสุทธิ์ของผลลัพธ์การจัดหมวดหมู่ของเวกเตอร์ที่ถูกลดขนาด เพื่อดูแนวโน้มว่าให้ผลลัพธ์อย่างไร ซึ่งจะอธิบายในหัวข้อ 3.10
4. นำไปใช้งานโดยสร้างเป็นไฟล์เจสัน โดยในส่วนนี้การสร้างไฟล์เจสันของระบบจัดหมวดหมู่เจตนาและไฟล์เจสันของระบบสกัดชื่อเฉพาะจะทำงานอย่างไม่เกี่ยวข้องกัน ซึ่งจะอธิบายในหัวข้อ 3.11

3.2 เครื่องมือที่ใช้ในการพัฒนา

การจัดหมวดหมู่เจตนาของประโยคและการสกัดชื่อเฉพาะในโครงการนี้ใช้ภาษาไพทอนเป็นหลักในการเขียนโปรแกรมของระบบขึ้นมาโดยใช้ชุดข้อมูลประวัติสนทนาที่มาจากเพจ Chula MOOC บน Facebook ในการจัดหมวดหมู่เจตนาของประโยคและการสกัดชื่อเฉพาะ ใช้เครื่องมือ จากคลังโปรแกรม (Library) ชื่อว่า pythainlp[19] เป็นตัวตัดคำของประโยค

สำหรับแช็ตบ็อตในโครงการนี้ใช้ Dialogflow ในการพัฒนาโดยเชื่อมต่อกับฐานข้อมูลบน Firebase และส่งจดหมายอิเล็กทรอนิกส์ด้วย SendGrid

3.3 วิธีการเก็บข้อมูล

การเก็บข้อมูลประวัติสนทนานำมาจากเพจ Chula MOOC บน Facebook ผ่าน Graph API ซึ่งให้บริการโดย Facebook ซึ่งมีขั้นตอนดังนี้

1. ค้นหา ID ของเพจ Chula MOOC จากเว็บไซต์ <https://findmyfbid.com>
2. ค้นหา Token ของเพจ Chula MOOC ที่ใช้สำหรับการร้องขอข้อมูล จากเว็บไซต์ <https://developers.facebook.com/>

3. ใช้ชุดคำสั่งที่ได้จาก Github[20] โดยใช้ชุดคำสั่ง

```
python fb-page-chat-download/run.py<PAGE_ID><OUTPUT_FILE><ACCESS_TOKEN>
```

3.1 รายละเอียดชุดข้อมูล

โครงการฉบับนี้ใช้ชุดข้อมูลประวัติสนทนาจากเพจ Chula MOOC โดยประวัติสนทนานี้เป็นบทสนทนายระหว่างผู้มาติดต่อเพจ Chula MOOC และเจ้าหน้าที่ดูแลเพจ ในบทสนทนาเป็นภาษาไทยที่มีการปนคำศัพท์ภาษาอังกฤษในบางบทสนทนา มีรายละเอียดดังนี้

จำนวนประโยคทั้งหมด	800	ประโยค
จำนวนประโยคที่สนใจทั้งหมด	450	ประโยค

ตัวอย่างประโยคที่สนใจ เช่น “ต้องการแก้ไขข้อมูลส่วนตัว” “อยากทราบว่าคอร์ส infographic จะเปิดให้ลงทะเบียนเมื่อไร” เป็นต้น

3.4 การเตรียมข้อมูล

ในขั้นตอนนี้จะนำประโยคคำถามที่ได้จากการเก็บข้อมูลมาผ่านกระบวนการดังต่อไปนี้

3.4.1 แบ่งประโยคความรวมให้เป็นประโยคย่อย

การแบ่งประโยคความรวมให้เป็นประโยคย่อยนั้นแบ่งโดยผู้จัดทำเอง ยกตัวอย่างเช่น แบ่งจาก “คอร์สนำเสนอผลงานเรียนกีโม่ และราคาเท่าไรครับ” เป็น “คอร์สนำเสนอผลงานเรียนกีโม่” และ “คอร์สนำเสนอผลงานราคาเท่าไรครับ” จากประโยคที่เราสนใจจำนวน 450 ประโยค เมื่อนำมาแบ่งแล้วจะได้ประโยคย่อยทั้งหมดจำนวน 470 ประโยค

3.4.2 แก้ไขคำผิด

การแก้ไขคำผิดผู้จัดทำจะตรวจและแก้ไขคำผิดเอง ยกตัวอย่างเช่น แก้จาก “Critixal Thinking for Business” เป็น “Critical Thinking for Business” เพื่อให้ผลลัพธ์การตัดคำมีประสิทธิภาพมากขึ้น เนื่องจากระบบตัดคำที่ใช้ยังไม่รองรับคำผิดบางคำ

3.4.3 การลบคำหยุด

การลบคำหยุดนั้นทำผ่านชุดคำสั่งที่ตรวจคำจากรายชื่อคำหยุดที่จัดทำขึ้น โดยมีรายละเอียดขั้นตอนดังนี้

1. นำเข้าประโยคที่ต้องการลบคำหยุด

```
raw_data = []
with open('s-in.txt') as inputfile:
    for line in inputfile:
        raw_data.append(line.strip().split('\n')[0])
```


- นำเข้ารายชื่อคำหยุดจาก pythainlp (รายละเอียดในภาคผนวก ข)

```
stopword = []
with open('stopword.txt') as inputfile:
    for line in inputfile:
        stopword.append(line.strip().split('\n')[0])
```

คำสั่งจะตรวจว่าประโยคนี้มีคำหยุดอยู่ในประโยคหรือไม่ และทำการลบออก
จากประโยคเมื่อตรวจพบ

- ชุดคำสั่งคืนค่าด้วยประโยคที่ผ่านการลบคำแล้ว

3.4.4 การตัดคำ

หลังจากชุดข้อมูลผ่านกระบวนการที่ 3.4.1 – 3.4.3 จนได้ประโยคที่ต้องการแล้วจะ
นำมาตัดออกเป็นคำย่อย โดยอาศัย pythainlp และได้ผลลัพธ์เป็นชุดคำย่อยจากประโยค

3.4.5 กำหนดเจตนาให้กับประโยค

ในขั้นตอนนี้จะกำหนดเจตนาทั้ง 14 เจตนาให้กับประโยคทั้งหมดเพื่อนำมาวัด
ประสิทธิภาพของระบบจัดหมวดหมู่ โดยแต่ละประโยคจะมีได้เพียง 1 เจตนาเท่านั้น ขั้นตอนนี้
นี้ทำด้วยผู้จัดทำเอง

3.4.6 กำหนดชื่อเฉพาะในประโยค

ในขั้นตอนนี้จะกำหนดชื่อเฉพาะในแต่ละประโยคเพื่อนำมาวัดประสิทธิภาพของ
ระบบสกัดชื่อเฉพาะ ได้ผลลัพธ์ว่ามี 61 ประโยคที่มีชื่อเฉพาะโดยแต่ละประโยคจะมีเพียง 1
คำเท่านั้น ซึ่งในขั้นตอนนี้จะทำด้วยผู้จัดทำเอง

3.5 การสร้างถุงคำ

เมื่อได้ชุดคำย่อยจากแต่ละประโยคแล้ว จะนำชุดคำย่อยทั้งหมดที่ไม่ซ้ำกันในทุกประโยคมา
รวมกันเป็นรายการ (list) เดียวกันเรียกว่า “ถุงคำ” โดยมีขั้นตอนดังต่อไปนี้

- สร้างรายการว่าง 1 รายการ

```
dataframe_text = pd.DataFrame(columns=['String', 'Intent'])
for text in input_text:
    text = text.split(",")
    df1 = pd.DataFrame([text], columns=['String', 'Intent'])
    dataframe_text = dataframe_text.append(df1)
```

- คำย่อยทั้งหมดลงไปรายการว่างนั้น

```

bag_of_word = []
dataframe_token = pd.DataFrame(columns=['Token', 'Intent'])
documents = []
intent_temp = dataframe_text['Intent'].tolist()
string_temp = dataframe_text['String'].tolist()
for count in range(len(string_temp)):
    result = tk.word_tokenize(string_temp[count], engine=
                              "newmm", whitespaces=False)
    documents.append(result)
    for j in result:
        bag_of_word.append(j)

```

คำสั่ง set() เพื่อกำจัดคำย่อยที่ซ้ำกัน ได้ผลลัพธ์เป็นลิสต์

```

bag_of_word = list(set(bag_of_word))
bag_of_word.sort()
bag_of_word = bag_of_word[13:]

```

โครงการนี้ได้ขนาดของลิสต์คือ 440 คำ ดังนั้นจึงได้เวกเตอร์ขนาด 440

3.6 การสร้างเวกเตอร์

หลังจากได้ลิสต์แล้ว จะนำลิสต์ที่ได้มาสร้างเวกเตอร์แทนประโยคโดยมีทั้งหมด 3 รูปแบบได้แก่

3.6.1 การเข้ารหัสข้อมูลวันฮ็อต (1-hot encoding)

รูปแบบนี้จะสร้างเวกเตอร์โดยตรวจสอบว่าคำในลิสต์ปรากฏในชุดคำย่อยหรือไม่ ถ้าปรากฏจะแทนด้วย 1 ถ้าไม่ปรากฏเลยจะแทนด้วย 0

3.6.2 การนับความถี่ (Frequency count)

รูปแบบนี้จะสร้างเวกเตอร์โดยตรวจสอบว่าคำในลิสต์ปรากฏในชุดคำย่อยกี่ครั้ง แล้วแทนด้วยจำนวนครั้งที่ปรากฏในประโยค ถ้าไม่ปรากฏเลยจะแทนด้วย 0

3.6.3 Term Frequency - Invert Document Frequency (TF-IDF)

รูปแบบนี้จะสร้างเวกเตอร์ด้วยวิธีการ Term Frequency - Invert Document Frequency โดยเชื่อว่าคำที่ปรากฏบ่อยจะมีความสำคัญมาก แต่ถ้าพบมากในหลายเอกสาร แสดงว่ามีความสำคัญน้อย

1. คำนวณ Term Frequency ตามสมการที่ 2.5
2. คำนวณ Invert Document Frequency ตามสมการที่ 2.6
3. คำนวณค่า Term Frequency – Invert Document Frequency โดยนำค่า Term Frequency และค่า Invert Document Frequency ของแต่ละคำย่อยในลิสต์มาทำการคูณกัน

3.7 การสกัดชื่อเฉพาะ

ขั้นตอนนี้ทำเพื่อสกัดคำที่เป็นชื่อเฉพาะที่อาจปรากฏในประวัติสนทนา ซึ่งในที่นี้คือชื่อวิชาที่อาจปรากฏได้ทั้งในรูปแบบย่อและรูปแบบเต็ม เช่น ภาษารัสเซียเบื้องต้นในชีวิตประจำวัน ซึ่งเป็นชื่อวิชาเต็ม สามารถปรากฏได้เป็น ภาษารัสเซีย ภาษารัสเซียเบื้องต้น หรือรัสเซียเบื้องต้น

กระบวนการนี้จะนำชุดของคำย่อยแต่ละประโยคมาเปรียบเทียบกับชื่อรายวิชาที่เปิดสอนใน Chula MOOC ที่เตรียมไว้ซึ่งเป็นชื่อเต็มโดยอาศัยการเทียบคำที่ยาวที่สุด ซึ่งมีขั้นตอนดังนี้

1. เลือกคำย่อยจากชุดคำย่อย n คำ โดยคำนึงถึงลำดับ
2. เปรียบเทียบคำย่อยที่เลือกกับชื่อรายวิชาที่เตรียมไว้ทั้งหมด
3. ลดความยาวคำย่อยที่เลือกเป็น $n-1$ คำ
4. เก็บคำย่อยที่ยาวที่สุดที่เป็นไปได้ที่ตรงกับส่วนหนึ่งหรือทั้งหมดของชื่อรายวิชา
5. วัดประสิทธิภาพของวิธีการด้วยค่าความแม่นยำของระบบ

3.8 การลดมิติของข้อมูล

เมื่อได้เวกเตอร์แทนประโยคจากทั้ง 3 รูปแบบแล้วจะมาทำการลดมิติของข้อมูลเพื่อลดความซับซ้อนลงโดยใช้ชุดโปรแกรม Scikit-learn โดยมีทั้งหมด 2 รูปแบบ ได้แก่

3.8.1 การวิเคราะห์องค์ประกอบหลัก

```
from sklearn.decomposition import PCA
PCA(n_components=<number>).fit_transform(<data>)
```

3.8.2 การจัดสรรของดีรีเคลแฝง

```
from sklearn.decomposition
import LatentDirichletAllocation
LDA(n_components=<number>).fit_transform(<data>)
```

ขั้นตอนนี้ทำขึ้นเพื่อหากต้องการหาว่าการลดมิติข้อมูลจะช่วยให้ระบบสามารถจัดหมวดหมู่ได้ดีขึ้นหรือไม่

3.9 การจัดหมวดหมู่

กระบวนการนี้จะใช้ชุดคำสั่งจาก Scikit-learn ดังนี้

```
from sklearn.cluster import KMeans
KMeans(init='k-means++', n_clusters=<number>).fit_predict(<data>)
```

3.10 การวัดความไม่บริสุทธิ์ของข้อมูล

ในกระบวนการนี้จะใช้ค่าความแปรปรวนในการวัดความไม่บริสุทธิ์ของข้อมูล โดยมีขั้นตอนดังนี้

1. นำเข้ารายการผลลัพธ์จากการจัดหมวดหมู่ซึ่งคือประโยคที่กำกับเจตนาของประโยคแล้ว

```
entropy = entropy(<data_result>)
```

2. คำนวณค่า entropy จากรายการที่นำเข้า

```
def(<data_result>):
    prob_list = []
    counter = collections.Counter(<data_result>)
    for i in counter.values():
        prob_list.append(i / len(<data_result>))
    entropy = 0
    for i in prob_list:
        entropy += i * math.log10(i)
    return entropy
```

3. คำนวณค่า entropy เฉลี่ยของระบบ

```
entropy_avg = 0
for list_result in result:
    entropy = entropy(list_result)
    entropy_avg += entropy / size_kmean
```

3.11 การส่งออกระบบมาใช้งาน

ในกระบวนการนี้จะส่งออกโมเดลที่ดีที่สุดเพื่อนำไปใช้จัดหมวดหมู่เจตนาของประโยคอื่น ๆ ที่อาจพบเจอในอนาคต ดังนี้

```
from sklearn.externals import joblib
joblib.dump(<model>, <filename>)
```

3.12 การสร้างไฟล์เจสันเพื่อป้อนให้กับ Dialogflow

เจสันไฟล์ที่สร้างมีโครงสร้างดังนี้

1. ไฟล์เจสันของเจตนา

```
{
  id: STRING
  name: STRING
  responses: [
    {
      messages: [
        {
          type: INTEGER
          platform: STRING
          lang: STRING
          speech: STRING
        }
      ]
    }
  ]
}
```

```

    ]
  }
]
priority: INTEGER
userSays: [
  {
    id: STRING
    data: [
      {
        text: STRING
        userDefined: BOOLEAN
      }
    ]
    isTemplate: BOOLEAN
    count: INTEGER
  }
]
}

```

2. ไฟล์เจสันของคำเฉพาะ

```

{
  id: STRING,
  name: STRING,
  entries: [
    {
      value: STRING,
      synonyms: [ STRING ]
    }
  ],
}

```

มีการทำงานในการสร้างไฟล์เจสันดังนี้

```

import json
with open(<filename>, 'w') as outfile:
    json.dump(<data>, outfile)

```

ไฟล์เจสันที่ได้สามารถนำไปใช้งานแยกกันได้ไม่เกี่ยวข้องกัน

3.13 การกำหนดเจตนา (Intent) ลง Dialogflow

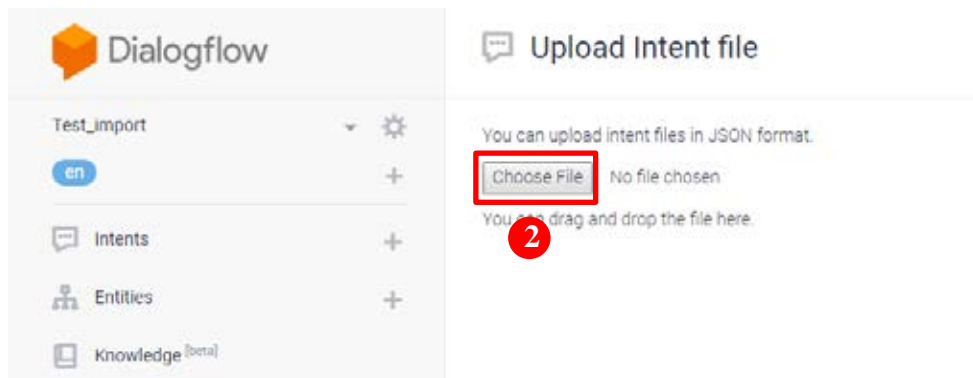
กระบวนการนี้เป็นการกำหนดเจตนาทั้งหมดที่ได้ลง Dialogflow และนำข้อความที่ตรงกับเจตนา นั้นมารวบรวมไว้ โดยมีขั้นตอนดังต่อไปนี้

1. กดปุ่ม เพิ่มเติม เลือก Upload Intent ดังภาพที่ 3.1



ภาพที่ 3.1 ปุ่ม Upload Intent ที่หน้าจัดการเจตนาทั้งหมดของ Dialogflow

- กดปุ่ม Choose File เพื่อเลือกไฟล์เจสันที่ต้องการอัปโหลด ดังภาพที่ 3.2

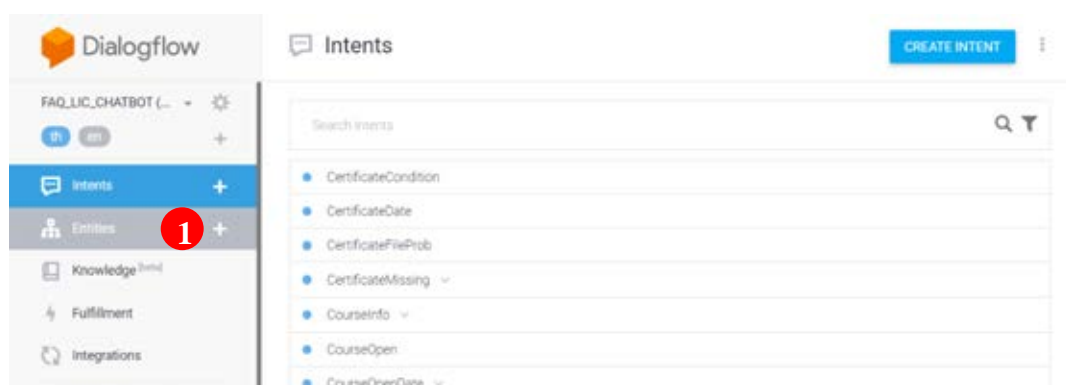


ภาพที่ 3.2 ปุ่ม Choose File ที่หน้า Upload Intent File

3.14 การกำหนดคำเฉพาะ (Entity) ลง Dialogflow

การกำหนดคำเฉพาะสำหรับ Dialogflow นั้น เป็นการนำชื่อเฉพาะที่จะได้จากการสกัดมา กำหนดจากข้อความที่ได้ใส่ไว้ในช่อง Training phrases โดยมีขั้นตอนดังต่อไปนี้

- เลือกแถบ Entities ที่หน้าส่วนควบคุมของ Dialogflow ดังภาพที่ 3.3



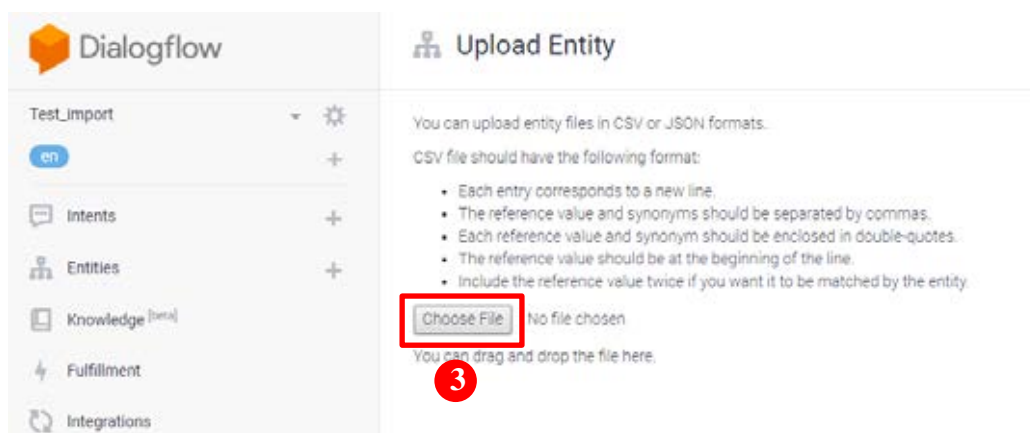
ภาพที่ 3.3 แถบ Entities ที่หน้าส่วนควบคุมของ Dialogflow

- กดปุ่ม เพิ่มเติม เลือก Upload Intent ดังภาพที่ 3.4



ภาพที่ 3.4 ปุ่ม Upload Entity ที่หน้าจัดการคำเฉพาะทั้งหมดใน Dialogflow

- กดปุ่ม Choose File เพื่อเลือกไฟล์เจสันที่ต้องการอัปโหลด ดังภาพที่ 3.5

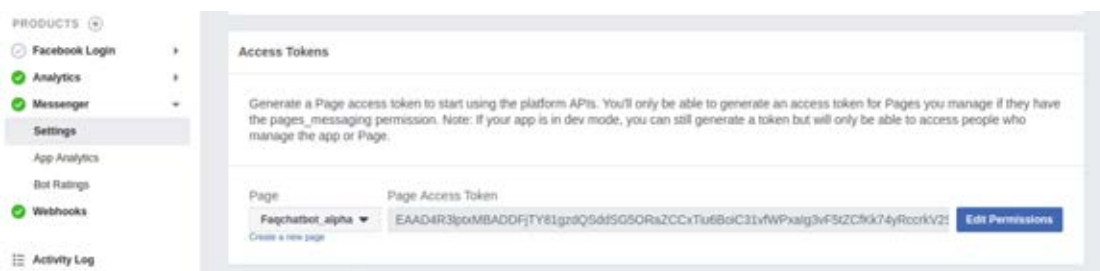


ภาพที่ 3.5 ปุ่ม Choose File ที่หน้า Upload Entity

3.15 การผสมการทำงาน Dialogflow เข้ากับ Messenger บน Facebook

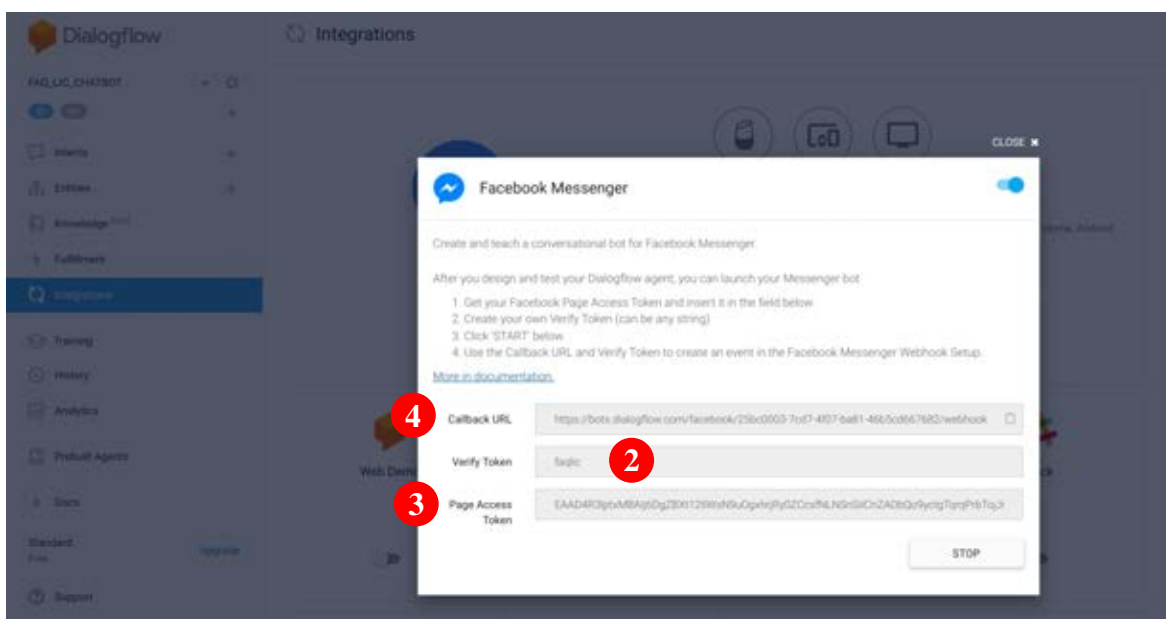
การผสมการทำงาน Dialogflow เข้ากับ Messenger บน Facebook มีขั้นตอนดังต่อไปนี้

1. เข้าไปยังเว็บไซต์ Facebook Developer แล้วเลือกแถบ Messenger เพื่อรับโทเคนการเข้าถึง (Access Tokens) ดังภาพที่ 3.6



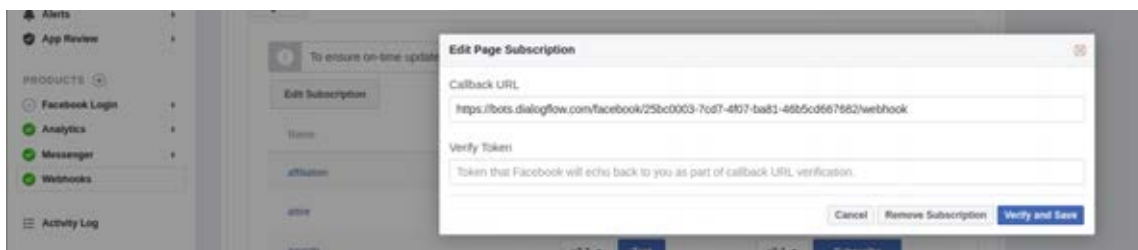
ภาพที่ 3.6 หน้าสำหรับการรับโทเคนการเข้าถึง (Access Tokens) Messenger

2. เข้า Dialogflow แล้วเลือกแถบ Integration จากนั้นเลือก Facebook Messenger เพื่อกำหนดค่า Verify Token ดังภาพที่ 3.7
3. นำค่าโทเคนการเข้าถึงที่ได้จาก Facebook Developer มาใส่ในช่อง Page Access Token ดังภาพที่ 3.7



ภาพที่ 3.7 หน้าแสดงค่าสำหรับการผสมงาน Dialogflow เข้ากับ Messenger บน Facebook

- นำค่า Callback URL และค่า Verify Token ที่ปรากฏดังภาพที่ 3.8 ไปใส่ในหน้าต่าง Edit Page Subscription ในแถบ Webhook บนเว็บไซต์ Facebook Developer จากนั้นกดปุ่ม Verify and Save ดังภาพที่ 3.8

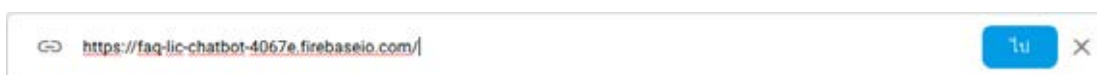


ภาพที่ 3.8 หน้าสำหรับเติมค่า Callback URL และค่า Verify Token

3.16 การผสมการทำงาน Dialogflow เข้ากับ Firebase

เนื่องจาก Dialogflow เป็นแพลตฟอร์มที่ทำงานร่วมกับ Firebase จึงไม่จำเป็นต้องมีการตั้งค่า แต่จะกล่าวถึงการตั้งค่าเพื่อดึงข้อมูลจากฐานข้อมูลแบบทันทีทันใด (Realtime Database) ซึ่งมีขั้นตอนดังต่อไปนี้

- คัดลอกที่อยู่ของฐานข้อมูลแบบทันทีทันใดจาก Firebase ดังภาพที่ 3.9



ภาพที่ 3.9 ที่อยู่ของฐานข้อมูลแบบทันทีทันใดของ Firebase

- นำเข้าสู่ชุดคำสั่ง firebase-admin ลงในส่วนเพิ่มเติมของ Dialogflow ดังนี้

```
const admin = require('firebase-admin')
```

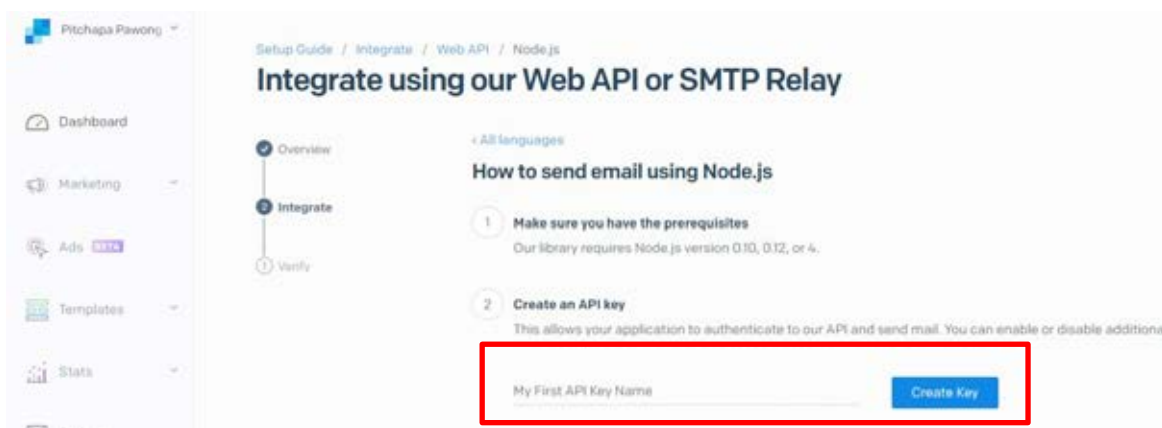
- กำหนดค่าเริ่มต้นการทำงานลงในส่วนเพิ่มเติมของ Dialogflow ดังนี้

```
admin.initializeApp({
  databaseURL: "<Realtime Database URL>"
});
```

3.17 การส่งจดหมายอิเล็กทรอนิกส์ผ่าน SendGrid

ในหัวข้อนี้จะกล่าวถึงการส่งจดหมายอิเล็กทรอนิกส์ผ่าน SendGrid ตั้งแต่ขั้นตอนการผสมการทำงานกับ Dialogflow ไปจนถึงการส่งจดหมายอิเล็กทรอนิกส์โดยมีขั้นตอนดังต่อไปนี้

- เข้าไปยังเว็บไซต์ <https://app.sendgrid.com/guide/integrate/langs/nodejs>
- กรอก API key ที่ช่อง My First API Key Name จากนั้นกดปุ่ม Create Key ดังภาพที่ 3.10



ภาพที่ 3.10 หน้าสำหรับสร้าง API key บน SendGrid

- คัดลอก API key ที่ได้มาเติมในชุดคำสั่งลงบนส่วนเพิ่มเติมของ Dialogflow ดังนี้

```
process.env.SENDGRID_API_KEY = "<API key>";
```

- ติดตั้งส่วนชุดคำสั่ง (Package) ของ SendGrid ที่ส่วนเพิ่มเติมของ Dialogflow สำหรับการส่งจดหมายอิเล็กทรอนิกส์ ดังนี้

```
npm install --save @sendgrid/mail
```

5. ส่งจดหมายอิเล็กทรอนิกส์ด้วยการเขียนรหัสลงในส่วนเติมเต็มของ Dialogflow ดังนี้

```
const sgMail = require("@sendgrid/mail");

sgMail.setApiKey(process.env.SENDGRID_API_KEY);
const msg = {
  to: "test@gmail.com",
  from: "test@hotmail.com",
  subject: "test",
  text: "test"
};
sgMail.send(msg);
```

ยกตัวอย่างซอร์สโค้ดสำหรับการส่งจดหมายอิเล็กทรอนิกส์ผ่าน SendGrid เมื่อมีข้อความการสนทนาจากผู้ใช้ที่ตรงกับเจตนาแจ้งปัญหาไม่สามารถดาวน์โหลดใบประกาศนียบัตรได้ ดังนี้

```
const sgMail = require("@sendgrid/mail");

function sendEmail(agent) {
  sgMail.setApiKey(process.env.SENDGRID_API_KEY);
  var usermsg = request.body.queryResult.queryText;
  var intentname = request.body.queryResult.intent.displayName;

  if (intentname==="CertificateFileProb"){
    const msg = {
      to: "ploypitch3338@gmail.com",
      from: "ploy_3338@hotmail.com",
      subject: "[Chatbot] แจ้งปัญหาการดาวน์โหลดไฟล์ Certificate",
      text: "ปัญหา:" + usermsg ,
    };
    console.log(msg);
    sgMail.send(msg);

    agent.add("รับทราบค่ะ จะทำการส่งเรื่องให้ผู้ดูแล");
    agent.add("*กรณีที่ไม่สามารถดาวน์โหลด Certificate of Completion ได้นั้นมีความเป็นไปได้ว่าผู้เรียนไม่ผ่านเกณฑ์การประเมินหรือเลยกำหนดเวลาเรียนได้ค่ะ");
  }
}
```

บทที่ 4

ผลการพัฒนา

ในบทนี้จะกล่าวถึง ผลของการดำเนินการพัฒนาระบบจัดหมวดหมู่เจตนาของประโยคและการสกัดชื่อเฉพาะจากประวัติสนทนาเพื่อฝึกสอนแชทบ็อต ประกอบด้วยผลการจัดหมวดหมู่เจตนาของประโยค ผลการสกัดชื่อเฉพาะ และผลการทำงานของแชทบ็อต

การวัดผลการจัดหมวดหมู่เจตนาของประโยคและการสกัดชื่อเฉพาะของระบบเป็นการเปรียบเทียบกับผลการจัดหมวดหมู่เจตนาของประโยคและการสกัดชื่อเฉพาะด้วยมือทั้งหมด โดยไม่มีการทำชุดข้อมูลสำหรับการทดสอบแยกไว้ และผลการทำงานของแชทบ็อตนั้นขึ้นกับความสามารถของ Dialogflow

4.1 ผลการจัดหมวดหมู่เจตนาของประโยค

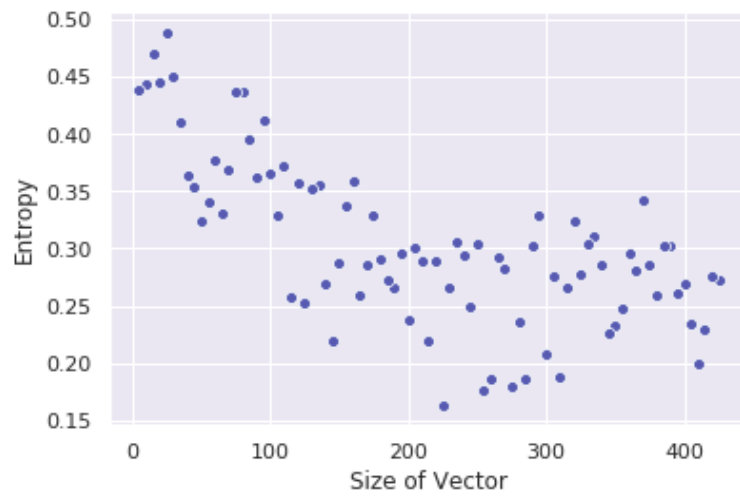
การจัดหมวดหมู่เจตนาของประโยคที่ผ่านการสร้างเวกเตอร์ด้วยวิธีการเคมีนด้วยการเข้ารหัสข้อมูลวันฮ็อต การนับความถี่ และ Term Frequency - Invert Document Frequency การจัดหมวดหมู่เจตนาของประโยคที่ผ่านการสร้างเวกเตอร์ด้วยกระบวนการเคมีนทั้งสามรูปแบบดังกล่าวมีการลดมิติของข้อมูลด้วยการวิเคราะห์องค์ประกอบหลัก และการจัดสรรของดีริเคลแฟงเพื่อหาแนวโน้มประสิทธิภาพของระบบเมื่อเวกเตอร์มีขนาดเล็กลง ซึ่งสามารถวัดค่าความไม่บริสุทธิ์โดยอาศัยค่าความแปรปรวนของข้อมูลที่เวกเตอร์ขนาด 440 ซึ่งเป็นขนาดของเวกเตอร์ตั้งต้นนั้น ได้ผลดังตารางที่ 4.1

ตารางที่ 4.1 ผลการจัดหมวดหมู่เจตนาของประโยค

	วิธีการสร้างเวกเตอร์		
	การเข้ารหัสแบบวันฮ็อต	การนับความถี่	TF-IDF
ความไม่บริสุทธิ์ของข้อมูล	0.2782	0.2929	0.0728

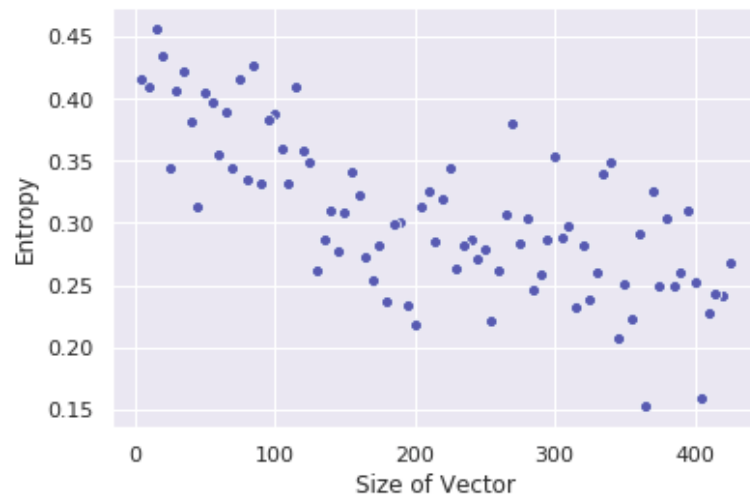
จากตารางที่ 4.1 พบว่าวิธีการสร้างเวกเตอร์ด้วย Term Frequency – Invert Document Frequency ให้ค่าความไม่บริสุทธิ์ของข้อมูลต่ำที่สุดคือ 0.0728 ถัดมาคือการเข้ารหัสแบบวันฮ็อต และการนับความถี่ที่ให้ค่าความไม่บริสุทธิ์ของข้อมูลที่ 0.2782 และ 0.2929 ตามลำดับ

สำหรับการทดลองลดขนาดมิติของเวกเตอร์ได้ให้ผลลัพธ์ดังนี้



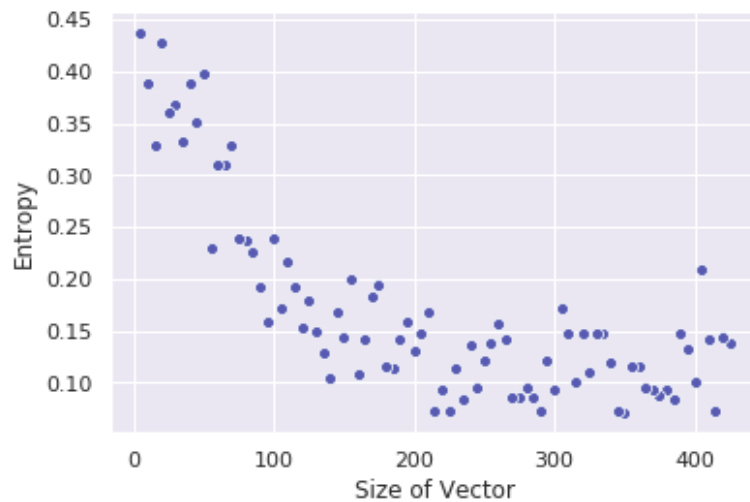
ภาพที่ 4.1 ค่าความแปรปรวนของข้อมูลที่สร้างเวกเตอร์ด้วยวิธีการเข้ารหัสแบบวันฮ็อตและลดขนาดเวกเตอร์ด้วยการวิเคราะห์องค์ประกอบหลัก

จากภาพที่ 4.1 พบว่าเมื่อมีการลดขนาดเวกเตอร์ด้วยการวิเคราะห์องค์ประกอบหลักที่สร้างด้วยวิธีการเข้ารหัสแบบวันฮ็อตจะมีแนวโน้มความไม่บริสุทธิ์ของข้อมูลสูงขึ้น



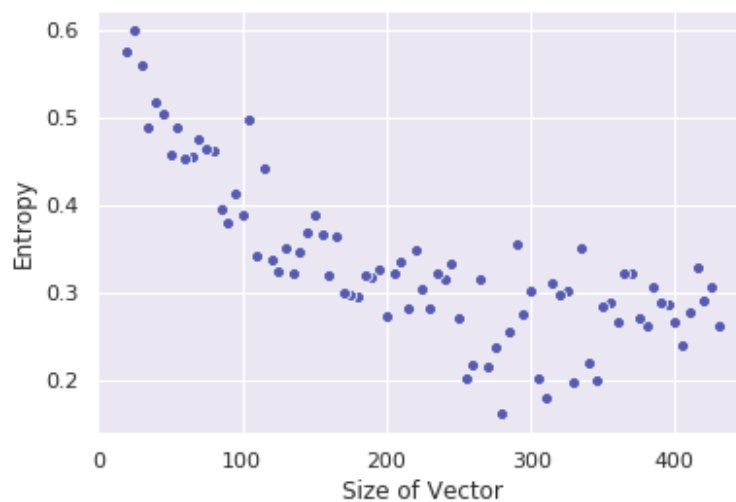
ภาพที่ 4.2 ค่าความแปรปรวนของข้อมูลที่สร้างเวกเตอร์ด้วยวิธีการนับความถี่และลดขนาดเวกเตอร์ด้วยการวิเคราะห์องค์ประกอบหลัก

จากภาพที่ 4.2 พบว่าเมื่อมีการลดขนาดเวกเตอร์ด้วยการวิเคราะห์องค์ประกอบหลักที่สร้างด้วยวิธีการนับความถี่จะมีแนวโน้มความไม่บริสุทธิ์ของข้อมูลสูงขึ้น



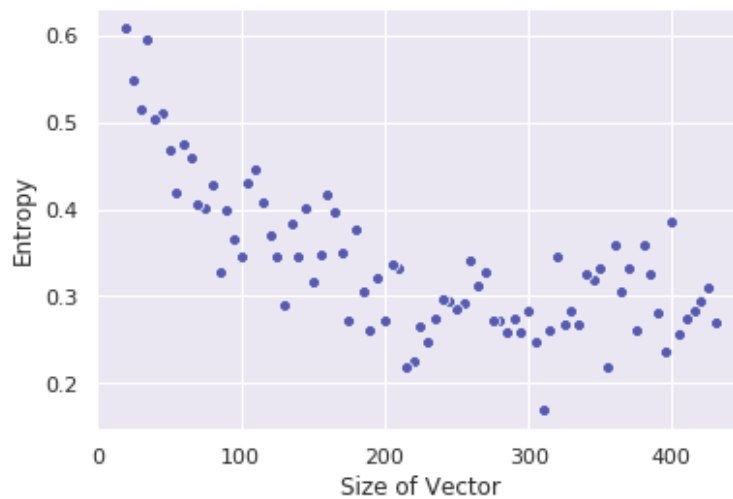
ภาพที่ 4.3 ค่าความแปรปรวนของข้อมูลที่สร้างเวกเตอร์ด้วยวิธี Term Frequency – Inverse Document Frequency และลดขนาดเวกเตอร์ด้วยการวิเคราะห์องค์ประกอบหลัก

จากภาพที่ 4.3 พบว่าเมื่อมีการลดขนาดเวกเตอร์วิธีการวิเคราะห์องค์ประกอบหลักที่สร้างด้วยวิธี Term Frequency – Inverse Document Frequency จะมีแนวโน้มความไม่บริสุทธิ์ของข้อมูลสูงขึ้น



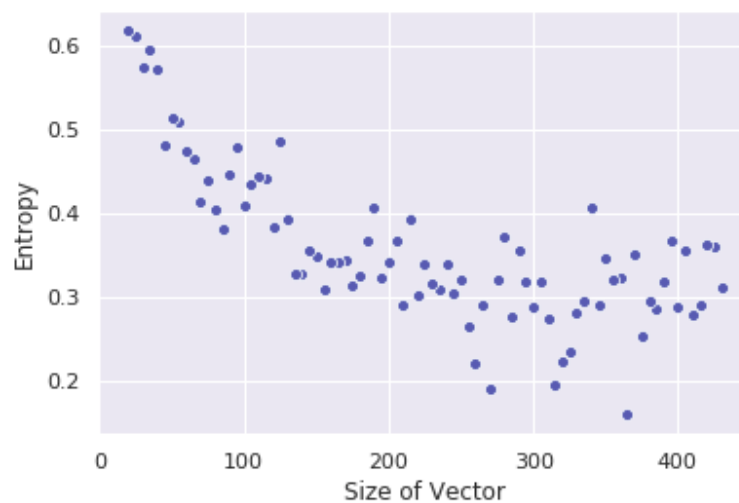
ภาพที่ 4.4 ค่าความแปรปรวนของข้อมูลที่สร้างเวกเตอร์ด้วยวิธีการเข้ารหัสแบบวันฮ็อตและลดขนาดเวกเตอร์ด้วยการจัดสรรดีรีเคิลแฝง

จากภาพที่ 4.4 พบว่า เมื่อมีการลดขนาดเวกเตอร์วิธีการจัดสรรดีรีเคิลแฝงที่สร้างด้วยวิธีการเข้ารหัสแบบวันฮ็อตจะมีแนวโน้มความไม่บริสุทธิ์ของข้อมูลสูงขึ้น



ภาพที่ 4.5 ค่าความแปรปรวนของข้อมูลที่สร้างเวกเตอร์ด้วยวิธีการนับความถี่และลดขนาดเวกเตอร์ด้วยการจัดสรรดีรีเคิลแฝง

จากภาพที่ 4.5 พบว่าเมื่อมีการลดขนาดเวกเตอร์การจัดสรรดีรีเคิลแฝงที่สร้างด้วยวิธีการนับความถี่จะมีแนวโน้มความไม่บริสุทธิ์ของข้อมูลสูงขึ้น



ภาพที่ 4.6 ค่าความแปรปรวนของข้อมูลที่สร้างเวกเตอร์ด้วยวิธี Term Frequency – Inverse Document Frequency และลดขนาดเวกเตอร์ด้วยการจัดสรรดีรีเคิลแฝง

จากภาพที่ 4.6 พบว่าเมื่อมีการลดขนาดเวกเตอร์การจัดสรรดีรีเคิลแฝงที่สร้างด้วยวิธีการ Term Frequency – Inverse Document Frequency จะมีแนวโน้มความไม่บริสุทธิ์ของข้อมูลสูงขึ้น

จากภาพที่ 4.1 – 4.6 ทำให้ได้ผลลัพธ์ว่าการลดมิติข้อมูลด้วยการวิเคราะห์องค์ประกอบหลัก และการจัดสรรดีรีเคิลแรงแฝงไม่สามารถเพิ่มประสิทธิภาพการจัดหมวดหมู่ข้อความได้ เมื่อเวกเตอร์มีขนาดลดลงส่งผลให้ประสิทธิภาพลดลงอย่างเห็นได้ชัด

4.2 ผลการสกัดชื่อเฉพาะ

การสกัดชื่อเฉพาะด้วยวิธีการเทียบคำที่ยาวที่สุดได้ผลลัพธ์ดังภาพที่ 4.7

Marked-word	47	14
	0	0
Nonmarked-word	Extract marked-word	Extract nonmarked-word

ภาพที่ 4.7 ผลการสกัดชื่อเฉพาะ

จากภาพที่ 4.7 การสกัดชื่อเฉพาะซึ่งเป็นทั้งชื่อย่อและชื่อเต็มเปรียบเทียบกับชื่อเต็มที่เตรียมไว้ด้วยวิธีการเทียบคำที่ยาวที่สุดพบว่าระบบสามารถสกัดชื่อเฉพาะได้ตรงกับที่กำหนดไว้ที่ 47 คำ และไม่ตรงกับที่กำหนดไว้ที่ 14 คำ ซึ่งสามารถคำนวณค่าความแม่นยำของระบบได้จากสมการที่ 2.14 เป็น $Precision = 47 / (47 + 14) = 0.7705$

4.3 ผลการทำงานของแช็ตบ็อต

ผลการทำงานของแช็ตบ็อตสำหรับแต่ละเจตนาได้ผลดังต่อไปนี้ และสามารถดูภาพหน้าจอการทำงานจริงของแช็ตบ็อต รวมทั้งการส่งจดหมายอิเล็กทรอนิกส์ได้ที่ภาคผนวก ค

1. ทักทาย

ผู้ใช้

สวัสดีค่ะ

แช็ตบ็อต

สวัสดีค่ะ 😊 ขอขอบคุณที่ให้ความสนใจกับ CHULA MOOC

2. สอบถามเวลาเปิดรายวิชา

ผู้ใช้ คอร์สคนรักสุนัข Dog Lover เปิดวันไหนคะ

แชทบ็อต คนรักสุนัข Dog Lover เปิดให้ลงทะเบียน 11 มีนาคม 2562 ค่ะ

และเปิดให้ลงทะเบียนเวลา 9.00 น. ของวันที่เปิดให้ลงทะเบียนค่ะ 😊

3. สอบถามรายวิชาที่เปิดสอน

ผู้ใช้ มีคอร์สอะไรเปิดตอนนี้บ้างคะ

แชทบ็อต วิชาที่เปิดให้ลงทะเบียนตอนนี้ ได้แก่ Practical Data Analytics Using Rapid Miner, คนรักสุนัข Dog Lover, สืบสานสารระมหาชาติ, มหากาพย์ อังกฤษ อัจฉกรต ศัพท์ทะเล อ่านทะเลง, การสร้าง Infographics What & How ?, Survival Thai, Thai on Campus ค่ะ

4. สอบถามรายละเอียดวิชา

ผู้ใช้ อยากรู้รายละเอียดวิชา English for careers I

แชทบ็อต กรุณาเลือกข้อมูลที่ต้องการทราบของรายวิชา English For Careers I For Hotel Staff

วันเริ่มลงทะเบียน

วันสิ้นสุดการเรียน

จำนวนบทเรียน

จำนวนผู้เรียนที่รับ

เกณฑ์การวัดผล

เกี่ยวกับเนื้อหา

อาจารย์ผู้สอน

ผู้ใช้ เกณฑ์การวัดผล

แชทบ็อต มีการวัดและประเมินผลผ่านแบบทดสอบย่อย (Quiz) และแบบทดสอบหลังเรียน (Posttest) โดยจะแบ่งเป็นคะแนนจาก Quiz เท่ากับ 50 คะแนน และคะแนนจาก Posttest เท่ากับ 50 คะแนน ทั้งนี้ผู้เรียนต้องทำคะแนนรวมทั้งหมดให้ได้ร้อยละ 60 ขึ้นไปจึงจะสามารถขอรับ Certificate of Completion ได้ หมายเหตุ ผู้เรียนจะมีสิทธิ์ทำแบบทดสอบหลังเรียน (Posttest) เมื่อเข้าร่วมกิจกรรมการเรียน (Course Progress) มากกว่า 80% ขึ้นไป

5. เงื่อนไขการได้รับใบประกาศนียบัตร

ผู้ใช้	ทำอย่างไรถึงได้ใบ cer
แชทบ็อต	<p>ผู้เรียนจะได้รับ Certificate of Completion เมื่อระบบทำการตรวจสอบว่าผู้เรียนมีคะแนนผ่านตามเกณฑ์การวัดและประเมินผล คือ ได้คะแนนมากกว่า 60% และผู้เรียนเข้าร่วมกิจกรรมการเรียนรู้ (Course Progress) มากกว่า 80% ขึ้นไป และจะปรากฏปุ่มให้ดาวน์โหลด Certificate of Completion ให้ที่หน้าบทเรียนและหน้า Score Detail ค่ะ</p> <p>*ในกรณีที่คะแนนไม่ถึงตามเกณฑ์หรือเรียนจบหลังเวลาที่กำหนด แต่ยังคงการรับ Certificate of Completion จะต้องรอลงทะเบียนในรายวิชาเดิมในรุ่นถัดไปเท่านั้น*</p>

6. เวลาประกาศรายชื่อผู้ได้รับใบประกาศนียบัตร

ผู้ใช้	รอรับ certificate ได้เมื่อไหร่ครับ
แชทบ็อต	<p>ผู้เรียนจะได้รับ Certificate of Completion หลังจากระบบได้ทำการตรวจสอบว่าผู้เรียนได้ผ่านเกณฑ์การประเมินเรียบร้อยแล้ว โดยจะปรากฏปุ่มให้ดาวน์โหลด Certificate of Completion ที่หน้าบทเรียนและหน้า Score Detail ค่ะ</p> <p>*ซึ่งแต่ละคนจะได้รับ Certificate of Completion ไม่พร้อมกัน ขึ้นกับเวลาที่ผู้เรียนผ่านเกณฑ์การประเมิน</p>

7. แจ้งการแก้ไขข้อมูล

ผู้ใช้	สามารถแก้ไข account info ได้มั้ยครับ
แชทบ็อต	<p>กรุณาเลือกสิ่งที่ต้องการแก้ไขข้อมูล</p> <div style="display: flex; justify-content: space-around; margin-top: 10px;"> <div style="border: 1px solid #007bff; border-radius: 10px; padding: 5px 15px; color: #007bff; text-decoration: none;">ข้อมูลส่วนตัวในระบบ</div> <div style="border: 1px solid #007bff; border-radius: 10px; padding: 5px 15px; color: #007bff; text-decoration: none;">ใบประกาศนียบัตร</div> </div>
ผู้ใช้	ข้อมูลส่วนตัวในระบบ
แชทบ็อต	รับทราบค่ะ รบกวนผู้เรียนแคปภาพหน้า Account Info ของผู้เรียน แล้วส่งมาในนี้ด้วยค่ะ
ผู้ใช้	(รูปภาพ)
แชทบ็อต	<p>ผู้เรียนต้องการแก้ไขข้อมูลส่วนใด</p> <div style="display: flex; justify-content: space-around; margin-top: 10px;"> <div style="border: 1px solid #007bff; border-radius: 10px; padding: 5px 15px; color: #007bff; text-decoration: none;">คำนำหน้าชื่อ</div> <div style="border: 1px solid #007bff; border-radius: 10px; padding: 5px 15px; color: #007bff; text-decoration: none;">ชื่อ-นามสกุล</div> <div style="border: 1px solid #007bff; border-radius: 10px; padding: 5px 15px; color: #007bff; text-decoration: none;">อาชีพ</div> <div style="border: 1px solid #007bff; border-radius: 10px; padding: 5px 15px; color: #007bff; text-decoration: none;">เพศ</div> <div style="border: 1px solid #007bff; border-radius: 10px; padding: 5px 15px; color: #007bff; text-decoration: none;">ปีเกิด</div> </div>

	<div style="border: 1px solid #0070C0; border-radius: 10px; padding: 5px; display: inline-block; margin-right: 10px;">ระดับการศึกษาสูงสุด</div> <div style="border: 1px solid #0070C0; border-radius: 10px; padding: 5px; display: inline-block;">e-mail address</div>
ผู้ใช้	คำนำหน้าชื่อ
แช็ตบ็อต	เลือกคำนำหน้าที่ถูกต้องของผู้เรียน <div style="display: flex; justify-content: space-around; margin-top: 10px;"> <div style="border: 1px solid #0070C0; border-radius: 10px; padding: 5px 15px;">Miss</div> <div style="border: 1px solid #0070C0; border-radius: 10px; padding: 5px 15px;">Mr.</div> <div style="border: 1px solid #0070C0; border-radius: 10px; padding: 5px 15px;">Mrs.</div> </div>
ผู้ใช้	Mr.
แช็ตบ็อต	ผู้เรียนต้องการแก้ไขข้อมูลส่วนตัวในระบบ โดยแก้ คำนำหน้าชื่อ เป็น Mr. ใช่หรือไม่ <div style="display: flex; justify-content: space-around; margin-top: 10px;"> <div style="border: 1px solid #0070C0; border-radius: 10px; padding: 5px 15px;">ใช่</div> <div style="border: 1px solid #0070C0; border-radius: 10px; padding: 5px 15px;">ไม่</div> </div>
ผู้ใช้	ใช่
แช็ตบ็อต	รับทราบค่ะ จะทำการส่งเรื่องให้ผู้ดูแลตรวจสอบและแก้ไขต่อไปค่ะ (ส่งจดหมายอิเล็กทรอนิกส์ไปยังผู้ดูแลระบบ)

8. แจ้งปัญหาตกหล่นใบประกาศนียบัตร

ผู้ใช้	ยังไม่ได้ cer
แช็ตบ็อต	คุณยังไม่ได้รับ Certificate ในวิชาใด ?
ผู้ใช้	ภาษาอาหรับเบื้องต้น
แช็ตบ็อต	รบกวนผู้เรียนให้ข้อมูลตามรูปแบบดังนี้ (พิมพ์ในข้อความเดียวกัน) Leaner ID : ชื่อ - นามสกุล : *กรณีที่ไม่ปรากฏปุ่มดาวน์โหลด Certificate of Completion นั้นมีความเป็นไปได้ว่าผู้เรียนไม่ผ่านเกณฑ์การประเมินหรือเลยกำหนดเวลาเรียนได้ค่ะ
ผู้ใช้	Leaner ID : CV999999 ชื่อ - นามสกุล : พิชชาภา พะวงษ์
แช็ตบ็อต	ผู้เรียนยังไม่ได้ Certificate of Completion ในวิชา ภาษาอาหรับเบื้องต้น Leaner ID : CV999999 ชื่อ - นามสกุล : พิชชาภา พะวงษ์ ข้อมูลถูกต้องตามนี้ใช่หรือไม่ <div style="display: flex; justify-content: space-around; margin-top: 10px;"> <div style="border: 1px solid #0070C0; border-radius: 10px; padding: 5px 15px;">ใช่</div> <div style="border: 1px solid #0070C0; border-radius: 10px; padding: 5px 15px;">ไม่</div> </div>

ผู้ใช้	ใช่
แชทบ็อต	รับทราบค่ะ จะทำการส่งเรื่องให้ผู้ดูแล (ส่งจดหมายอิเล็กทรอนิกส์ไปยังผู้ดูแลระบบ)

9. แจ้งปัญหาเปิดไฟล์ใบประกาศนียบัตรไม่ได้

ผู้ใช้	เปิดไฟล์ใบ cer ไม่ได้ค่ะ
แชทบ็อต	ผู้รับทราบค่ะ จะทำการส่งเรื่องให้ผู้ดูแล *กรณีที่ไม่สามารถดาวน์โหลด Certificate of Completion ได้นั้นมีความเป็นไปได้ว่าผู้เรียนไม่ผ่านเกณฑ์การประเมินหรือเลยกำหนดเวลาเรียนได้ค่ะ (ส่งจดหมายอิเล็กทรอนิกส์ไปยังผู้ดูแลระบบ)

10. คำถามที่พบบ่อยข้อที่ 1

ผู้ใช้	CHULA MOOC คืออะไร
แชทบ็อต	สำหรับการเรียนออนไลน์ CHULA MOOC จะเป็นการเรียนในรูปแบบการดูวิดีโอ และทำแบบทดสอบผ่านคอมพิวเตอร์จากสถานที่ใดก็ได้ตามความสะดวกของผู้เรียน และสามารถเรียนหลายวิชาได้ตามความสนใจ ภายในระยะเวลาที่กำหนดของวิชานั้นค่ะ

11. คำถามที่พบบ่อยข้อที่ 2

ผู้ใช้	มีค่าใช้จ่ายหรือไม่
แชทบ็อต	สำหรับคอร์สเรียนบน CHULA MOOC ผู้ที่สนใจสามารถลงทะเบียนเรียนได้ฟรี ไม่มีค่าใช้จ่ายค่ะ

12. คำถามที่พบบ่อยข้อที่ 3

ผู้ใช้	CHULA MOOC เหมาะกับใคร
แชทบ็อต	สำหรับคอร์สออนไลน์ CHULA MOOC นั้น เปิดให้กับทุกคนที่มีความสนใจกับคอร์สเรียนที่เราเปิดสอนค่ะ

13. คำถามที่พบบ่อยข้อที่ 4

ผู้ใช้ ลงทะเบียนเรียนอย่างไร

แชทบ็อด สำหรับวิธีการสมัครเรียนสามารถดูขั้นตอนได้จาก link นี้ได้เลยค่ะ >>

<https://mooc.chula.ac.th/how-to>

14. ขอบคุณ

ผู้ใช้ ขอบคุณค่ะ

แชทบ็อด



บทที่ 5

ข้อสรุปและปัญหาที่พบ

ในบทนี้จะกล่าวถึงการสรุปผลจากข้อมูลผลการพัฒนาในบทที่ 4 ของระบบจัดหมวดหมู่เจตนาของประโยคและการสกัดชื่อเฉพาะจากประวัติสนทนาเพื่อฝึกสอนแชตบ็อต รวมถึงปัญหาที่พบและแนวทางการแก้ไข

5.1 ข้อสรุป

จากข้อมูลผลการพัฒนาพบว่า การจัดหมวดหมู่เจตนาของประโยคโดยใช้การสร้างเวกเตอร์ด้วยวิธี Term Frequency – Inverse Document Frequency ให้ผลลัพธ์ที่ดีที่สุด และเมื่อทดสอบการลดขนาดเวกเตอร์ที่สร้างขึ้นจากทั้ง 3 วิธีด้วยการวิเคราะห์องค์ประกอบหลักและการจัดสรรตรีเคิลแฝงพบว่าเมื่อเวกเตอร์มีขนาดเล็กลงแล้วค่าความไม่บริสุทธิ์ของข้อมูลนั้นมีแนวโน้มที่สูงมากขึ้น ดังนั้นจึงสรุปได้ว่า การลดขนาดเวกเตอร์ไม่สามารถเพิ่มประสิทธิภาพของระบบในโครงการนี้ได้

การสกัดชื่อเฉพาะด้วยวิธีการเทียบคำที่ยาวที่สุดพบว่าระบบมีค่าความแม่นยำของระบบอยู่ที่ 0.7705 หรือ 77.05%

แชตบ็อตสามารถตอบคำถามได้ในทุกเจตนา ดึงข้อมูลรายละเอียดรายวิชาจากฐานข้อมูลมาใช้ในการตอบคำถาม และสรุปการแจ้งปัญหาจากผู้ใช้ส่งจดหมายอิเล็กทรอนิกส์ไปยังผู้ดูแลระบบได้ ทั้งนี้แชตบ็อตยังมีข้อจำกัดในการทำงานอยู่คือไม่สามารถตอบคำถามได้เมื่อเจอประโยคหรือคำที่ไม่เคยพบในประวัติสนทนา

5.2 ปัญหาที่พบและแนวทางการแก้ไข

สามารถสรุปปัญหาที่พบและแนวทางแก้ไขดังตารางที่ 5.1

ตารางที่ 5.1 ปัญหาที่พบและแนวทางการแก้ไข

ปัญหาที่พบ	แนวทางการแก้ไข
<p>1. จำนวนประวัติสนทนาที่สามารถดึงมาได้มีจำนวนค่อนข้างน้อย เนื่องจากการเปลี่ยนนโยบายการใช้งาน Graph API ของ Facebook ทำให้ไม่สามารถดึงประวัติสนทนาจากเพจ Chula MOOC เพิ่มเติมได้</p>	<p>1. หาข้อมูลเพิ่มเติมจากแหล่งอื่นนอกจากประวัติสนทนา เช่น ความคิดเห็นจากโพสต์บนเพจ Chula MOOC หรือการสอบถามจากจดหมายอิเล็กทรอนิกส์ เป็นต้น</p>
<p>2. จำนวนประโยคที่อยู่ในแต่ละหมวดหมู่เจตนามีจำนวนแตกต่างกันมากทำให้ยากต่อการสรุปการวัดความไม่บริสุทธิ์ของข้อมูล</p>	<p>2. หาประวัติสนทนาให้ได้มากขึ้น หาข้อมูลเพิ่มเติมจากแหล่งอื่น หรือมีการคัดเลือกข้อมูลให้แต่ละเจตนามีจำนวนใกล้เคียงกัน</p>
<p>3. จำเป็นต้องเปลี่ยนมาใช้ฐานข้อมูลบน Firebase แทนการใช้ฐานข้อมูลบนมอังกอดีบี (MongoDB)</p>	<p>3. เนื่องจาก Firebase นั้นสามารถผสานการทำงานเข้ากับ Dialogflow ได้สะดวกกว่า เพราะเป็นบริการจากกูเกิ้ลเช่นกัน ดังนั้นควรศึกษาการทำงานร่วมกับบริการอื่น ๆ ของ Dialogflow ให้ละเอียดถี่ถ้วนมากขึ้น</p>

รายการอ้างอิง

- [1] Dialogflow. (2018). “How does fulfillment work?”. Retrieved September 15, 2018, from <https://dialogflow.com/docs/intro/fulfillment>
- [2] Jocelyn D'Souza. (2562). **An Introduction to Bag-of-Words in NLP**, Retrieved October 17, 2018, from <https://medium.com/greyatom/an-introduction-to-bag-of-words-in-nlp-ac967d43b428>
- [3] Ashish Kathuria, Bernard J. Jansen, Carolyn Hafernik, Amanda Spink, (2010) **"Classifying the user intent of web queries using k-means clustering"**, Internet Research, Vol. 20 Issue: 5, pp.563-581, <https://doi.org/10.1108/10662241011084112>
- [4] Jason Brownlee. (2562). **A Gentle Introduction to the Bag-of-Words Model**. Retrieved September 15, 2018, from <https://machinelearningmastery.com/gentle-introduction-bag-words-model/>
- [5] Jiawei H, Mecheline K, Jian P. **Data Mining: Concepts and Techniques. 3rd ed.** San Francisco: Morgan Kaufmann Publishers Inc; 2011
- [6] Lindsay I Smith , **A tutorial on Principal Components Analysis**, University of Otago, 2002.
- [7] Lee J, Kang J-H, Jun S, Lim H, Jang D, Park S. **Ensemble Modeling for Sustainable Technology Transfer. Sustainability. 2018; 10(7):2278.**
- [8] Smith, L. I. (2002). **A tutorial on principal components analysis.**
- [9] Data School. (2562). **Simple guide to confusion matrix terminology.** Retrieved October 17, 2018, from <https://www.dataschool.io/simple-guide-to-confusion-matrix-terminology/>
- [10] Chula MOOC. (2561). **FAQ (คำถามที่พบบ่อย).** สืบค้นเมื่อ 10 สิงหาคม 2561, จาก <https://mooc.chula.ac.th/faq>.
- [11] Petch Kruapanich. (2561). **พัฒนาซอฟต์แวร์แบบ Rule-based approach VS AI based approach.** สืบค้นเมื่อ 25 พฤศจิกายน 2561, จาก <https://medium.com/@petchkruapanich/พัฒนาซอฟต์แวร์แบบ-rule-based-approach-vs-ai-based->

approach-3a32bee13ce3.

- [12] สุฤดี ฉัตรไตรมงคล, “การรู้จำและการจำแนกประเภทของชื่อเฉพาะภาษาไทย”, จุฬาลงกรณ์มหาวิทยาลัย, 2548
- [13] Jedsada Saengow. (2562). **[Firebase] คืออะไร มาดูวิธีสร้าง Project c และทำความรู้จักกับ Firebase**. สืบค้นเมื่อ 4 กุมภาพันธ์ 2562, จาก <https://medium.com/jed-ng/firebase-คือ-อะไร-มาดูวิธีสร้าง-Project-และทำความรู้จักกับ-Firebase>
- [14] Pool13433. (2562). **Firestore คืออะไร ขั้นตอนการเริ่มใช้งาน พร้อมโค้ดตัวอย่างแบบง่าย ๆ**, สืบค้นเมื่อ 10 มีนาคม 2562, จาก <https://www.poolsawat.com/firebase-tutorial/>
- [15] Lukkidd. (2562). **Word2Vec ทำอย่างไร**. สืบค้นเมื่อ 4 กุมภาพันธ์ 2562, จาก <https://lukkidd.com/word2vec-ทำอย่างไร>
- [16] lukkidd. (2562). **TF-IDF คำไหนสำคัญนะ?**. สืบค้นเมื่อ 4 กุมภาพันธ์ 2562, จาก <https://lukkidd.com/tf-idf-คำไหนสำคัญนะ-dd1e1568312e>
- [17] Pongsakorn Jirachanchaisiri. (2562). **รู้จักกับ Latent Dirichlet Allocation**. สืบค้นเมื่อ 20 ตุลาคม 2562, <https://medium.com/@pongsakorn.jrc/รู้จักกับ-latent-dirichlet-allocation-part-1-2495acfcda86>
- [18] KASIDIS SATANGMONGKOL. (2562). **อธิบาย Confusion Matrix ฉบับเข้าใจง่าย (มาก)**. สืบค้นเมื่อ 20 ตุลาคม 2560, <https://datarockie.blog/2018/04/30/confusion-matrix-explained/>
- [19] <https://github.com/PyThaiNLP/pythainlp>
- [20] <https://github.com/eisenjulian/fb-page-chat-download>

ภาคผนวก

ภาคผนวก ก

แบบเสนอหัวข้อโครงการ รายวิชา 2301399 Project Proposal

ปีการศึกษา 2561

ชื่อโครงการ (ภาษาไทย)	ระบบจัดหมวดหมู่เจตนาของประโยคและการสกัดชื่อเฉพาะจากประวัติสนทนาเพื่อฝึกสอนแชทบอต	
ชื่อโครงการ (ภาษาอังกฤษ)	An intent clustering and named entity extraction system from chatlogs for training chatbot	
อาจารย์ที่ปรึกษา	ผู้ช่วยศาสตราจารย์ ดร.จิตยา หวานวาริ	
ผู้ดำเนินการ	1. นายณัฐกิตต์ สมอร่าม	เลขประจำตัวนิสิต 5833621323
	2. นางสาวพิชชาภา พะวงษ์	เลขประจำตัวนิสิต 5833647723
	สาขาวิชาวิทยาการคอมพิวเตอร์ ภาควิชาคณิตศาสตร์และวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย	

หลักการและเหตุผล

ในปัจจุบันเครือข่ายสังคม (Social Network) ได้เข้ามามีบทบาทสำคัญทางด้านการติดต่อสื่อสารเป็นอย่างมาก เนื่องจากสามารถเพิ่มความสะดวกในการติดต่อกับผู้คน และยังมีความเร็วในการส่งสารเมื่ออยู่ในสถานการณ์ที่มีข้อจำกัดด้านเวลา

การติดต่อกับหน่วยงานหรือร้านค้าต่าง ๆ มักใช้ระบบส่งข้อความทันที (Instant Messaging) บนเครือข่ายสังคมเป็นตัวกลางในการติดต่อ เช่น Messenger ใน Facebook หรือ Direct Message (DM) ใน Twitter และ Instagram ซึ่ง Instant Messaging คือแอปพลิเคชันส่งข้อความทันทีไปยังบัญชีเป้าหมายซึ่งได้รับความนิยมทั้งบน Web Application และ Mobile Application เนื่องจากใช้งานได้ง่ายและสามารถติดต่อได้ตลอดเวลา

Chula MOOC [1] เป็นผู้ให้บริการการเรียนรู้ออนไลน์ ซึ่งมีช่องทางการติดต่อผ่านเพจ Chula MOOC ใน Facebook ที่เป็นเครือข่ายสังคมออนไลน์ที่ได้รับความนิยม ในปัจจุบัน Chula MOOC มีการเติบโตมากขึ้น ทำให้มีผู้สนใจและติดต่อผ่าน Messenger ใน Facebook มากขึ้นอย่างรวดเร็ว เนื่องด้วยจำนวนพนักงานที่น้อย ทำให้ไม่สามารถรองรับผู้ที่สนใจได้ ยิ่งไปกว่านั้น Chula

MOOC สามารถเรียนรู้ได้ 24 ชั่วโมงไม่จำกัดสถานที่ ทำให้ผู้ดูแลไม่สามารถดูแลผู้ใช้งานหากอยู่นอกเวลาการทำงาน

แชทบ็อตจึงถูกนำมาปรับใช้เพื่อให้บริการ โดยการพัฒนาแชทบ็อตนั้นสามารถพัฒนาขึ้นเอง และสามารถบริหารทรัพยากรได้เอง หรือใช้งาน Platform สำเร็จรูปที่ให้บริการฟรี ซึ่งทั้งสองรูปแบบต้องฝึกแชทบ็อตให้เข้าใจถึงประโยคที่ผู้ติดต่อส่งมาด้วยตนเอง

แชทบ็อตที่พัฒนาขึ้นนี้มีแนวทางการพัฒนาที่อาศัยกฎเกณฑ์หรือเงื่อนไขที่กำหนด (Rule-based approach) และอาศัยเจตนาของประโยคหรือข้อความ (Intent-based approach) [2] ซึ่งการพัฒนาโดยอาศัยเจตนาของประโยคหรือข้อความนั้นเป็นที่นิยมในปัจจุบันเนื่องจากการนำการประมวลผลภาษาธรรมชาติ (Natural Language Processing) มาใช้เพื่อให้แชทบ็อตนั้นสามารถเข้าใจและตอบสนองได้ถูกต้องและหลากหลายมากยิ่งขึ้นอาศัยกฎเกณฑ์หรือเงื่อนไขที่กำหนดร่วมกับอาศัยเจตนาของประโยคหรือข้อความ ทำให้สามารถตอบข้อความได้แม่นยำและหลากหลายมากขึ้น

การโต้ตอบของแชทบ็อตอาศัยการวิเคราะห์เจตนาของประโยคที่ผู้ใช้ป้อน และระบุการตอบสนองที่เหมาะสมกับเจตนา นั้น นอกจากนี้ ในประโยคที่ผู้ใช้ป้อนยังอาจมีชื่อเฉพาะ (Named Entity) คือ นิพจน์ที่ใช้เรียกหรือระบุสิ่งใด ๆ เช่น ชื่อบุคคล ชื่อองค์กร ชื่อวิชา เป็นต้น และสามารถใช้เป็นคำสำคัญ (keyword) ในการค้นหาข้อมูลสำหรับการตอบสนองได้ [3] ซึ่งแชทบ็อตจะกำหนดให้การตอบสนองเปลี่ยนตามนิพจน์ที่ผู้ใช้ป้อน เช่น “คอร์สวิชาภาษารัสเซียจะเปิดเมื่อไหร่” มีเจตนาของข้อความเป็น “การสอบถามรายวิชา” และมีชื่อเฉพาะเป็น “วิชาภาษารัสเซีย” การตอบสนองข้อความของผู้ใช้จึงเป็นการหาข้อมูลเวลาเปิดของวิชาภาษารัสเซียมาตอบให้ผู้ใช้

ระบบแชทบ็อต เช่น DialogFlow ChatFuel นั้นสามารถวิเคราะห์เจตนาของประโยคหรือข้อความที่ผู้ใช้ป้อนได้หากมีตัวอย่างของประโยคในเจตนา นั้นมากเพียงพอ แต่การสร้างตัวอย่างประโยคหรือการระบุเจตนาของประโยคตัวอย่างยังจำเป็นต้องทำโดยมนุษย์ ตัวอย่างประโยคนั้นสามารถรวบรวมได้จากประวัติสนทนาที่ผ่านมา แต่ยังคงจำเป็นต้องให้มนุษย์กำกับหรือระบุเจตนาของประโยคตัวอย่างอยู่ เช่น เจตนา “สอบถามกำหนดเปิดรายวิชา” จะต้องกำหนดตัวอย่างประโยค เช่น “คอร์สวิชาภาษารัสเซียเปิดเมื่อไหร่” “คอร์สBigDataจะเปิดอีกรอบเมื่อไหร่คะ” “วิชานาเสนอจะเปิดอีกตอนไหนคับ” เป็นต้น แชทบ็อตจะเรียนรู้ความคล้ายคลึงของประโยคที่กำหนดไว้ โดยอาศัยคุณลักษณะของประโยค เช่น ความยาวของประโยค ตำแหน่งคำศัพท์ในประโยค เป็นต้น เมื่อผู้ใช้งานพิมพ์ข้อความมาว่า “คอร์สการตลาดจะเปิดเมื่อไหร่ครับ” จะถูกจัดให้อยู่ในเจตนา “สอบถามกำหนดเปิดรายวิชา” และการสร้างแชทบ็อตขึ้นมาใหม่สำหรับแต่ละงานจึงยังคงเป็นเรื่องที่ต้องใช้แรงงานมาก นอกจากนี้ การกำหนดชื่อเฉพาะก็ยิ่งขึ้นกับลักษณะการสนทนานั้น ๆ ด้วย เช่น ในแชทบ็อตของ Chula MOOC ชื่อเฉพาะอาจเป็นรายวิชาที่เปิดสอน หรือรายวิชาที่ผู้เรียนสนใจ แต่ในแชทบ็อตของ

ระบบซื้อขายสินค้า ซึ่งเฉพาะอาจเป็นชื่อหรือรุ่นของสินค้า ด้วยเหตุนี้ ผู้พัฒนาจึงต้องการพัฒนาระบบจัดหมวดหมู่เจตนาของประโยคและสัณนิพนธ์ระบุนามจากประวัติการสนทนาเพื่อนำไปใช้เป็นประโยคฝึกสอนในการฝึกสอนการวิเคราะห์เจตนาของประโยคในแชตบ็อต

วัตถุประสงค์

โครงการมีวัตถุประสงค์เพื่อพัฒนาระบบเพื่อแก้ไขปัญหาที่ได้กล่าวมาข้างต้น ดังนี้

1. พัฒนาระบบจัดหมวดหมู่เจตนาของประโยคจากประวัติการสนทนา
2. พัฒนาระบบสกัดชื่อเฉพาะจากประวัติสนทนา
3. ประยุกต์ใช้ระบบแชตบ็อตสำเร็จรูป Dialogflow ให้ทำงานคู่กับ Messenger บน Facebook

ขอบเขตของโครงการ

1. ผู้ใช้งานต้องกำหนดจำนวนเจตนาของประโยคเอง
2. แชตบ็อตยอมรับข้อความภาษาไทย และคำภาษาอังกฤษบางคำ เช่น course certificate law เป็นต้น
3. แชตบ็อตพัฒนาขึ้นด้วย Dialogflow
4. แชตบ็อตสามารถประมวลผล 1 เจตนาต่อ 1 ข้อความ โดย 1 ข้อความ คือการป้อนข้อมูลของผู้ใช้ 1 ครั้ง
5. แชตบ็อตสามารถจัดหมวดหมู่ของเจตนาของประโยคที่เกี่ยวข้องกับรายวิชา ใบประกาศนียบัตร การแจ้งปัญหา และคำถามที่พบบ่อย
6. แชตบ็อตสามารถตอบคำถามการสอบถามรายวิชา ได้แก่
 1. เวลาเปิดรายวิชา เช่น “คอร์สวิชาภาษารัสเซียจะเปิดให้ลงทะเบียนก็โมงครับ”
 2. รายวิชาที่เปิดให้ลงทะเบียน เช่น “คอร์สวิชาภาษารัสเซียจะเปิดอีกเมื่อไหร่ครับ”
 3. รายละเอียดการเรียน เช่น “คอร์สวิชาภาษารัสเซียมีกำหนดคะแนนยังไงครับ”
7. แชตบ็อตสามารถตอบคำถามเกี่ยวกับใบประกาศนียบัตร ได้แก่
 1. เงื่อนไขการได้รับใบประกาศนียบัตร เช่น “เรียนจบแล้ว ต้องทำยังไงถึงได้ใบประกาศนียบัตร”
 2. เวลาประกาศรายชื่อผู้ได้รับใบประกาศนียบัตร เช่น “วิชาภาษารัสเซียจะประกาศรายชื่อคนได้ใบประกาศนียบัตรเมื่อไหร่ครับ”
8. แชตบ็อตสามารถรับร้องเรียนการแจ้งปัญหา และสรุปเรื่องส่งไปยังผู้ดูแลเพจทางจดหมายอิเล็กทรอนิกส์เพื่อดำเนินการแก้ปัญหาต่อไป โดยการแจ้งปัญหานั้นได้แก่

จุฬาลงกรณ์มหาวิทยาลัย									
2. ศึกษาค้นคว้าข้อมูลเกี่ยวกับการหาเจตนาประโยชน์ของผู้ใช้งาน									
3. ศึกษาเครื่องมือ โปรแกรม และเทคนิคที่ใช้ในโครงการงาน									
4. รวบรวมประวัติสนทนา เพื่อนำมาเป็นข้อมูลในการพัฒนาระบบ									
5. กำหนดขอบเขต และวิธีการดำเนินงานของโครงการงาน									
6. วิเคราะห์และออกแบบระบบหาเจตนาของประโยค									
7. พัฒนาระบบหาเจตนาของประโยค									
8. ทดสอบระบบ									
9. ปรับปรุงและแก้ไขระบบ									
10. สรุปผลการดำเนินงาน									
11. จัดทำเอกสาร									

ประโยชน์ที่คาดว่าจะได้รับ

ต่อผู้ใช้

1. สามารถเพิ่มความมั่นใจให้ผู้ติดต่อว่าปัญหาที่ได้แจ้งมานั้นจะได้รับการช่วยเหลือหรือแก้ปัญหาเนื่องจากตอบสนองได้รวดเร็ว
2. สามารถเพิ่มความสะดวกให้ผู้ดูแลเพจ โดยลดการตอบคำถามที่ถามบ่อย
3. สามารถลดระยะเวลาในการตอบคำถามและแก้ปัญหาให้ผู้ติดต่อ

4. สามารถนำระบบไปพัฒนาและใช้ร่วมกับแชทบอทอื่น ๆ ได้
ต่อผู้พัฒนาระบบ
1. มีความรู้และความเข้าใจในกระบวนการพัฒนาซอฟต์แวร์
2. ฝึกกระบวนการคิดวิเคราะห์ ทำงานอย่างมีระเบียบแบบแผน มีความรับผิดชอบต่องาน และทำงานอย่างตรงต่อเวลา
3. มีความรู้ในการใช้เครื่องมือที่ใช้ในการพัฒนา ใช้ Python JavaScript รวมไปถึงเครื่องมือสำเร็จรูป Dialogflow MongoDB และ Google Cloud Platform

อุปกรณ์และเครื่องมือที่ใช้

1. ฮาร์ดแวร์
 - 1.1 เครื่องคอมพิวเตอร์ที่ใช้ระบบปฏิบัติการ Windows 10 มีหน่วยประมวลผล 2.50 GHz Intel Core i7 หน่วยความจำหลัก 16 GB และหน่วยความจำสำรอง 1 TB
 - 1.2 เครื่องคอมพิวเตอร์ที่ใช้ระบบปฏิบัติการ Windows 10 มีหน่วยประมวลผล 2.20 GHz Intel Core i5 หน่วยความจำหลัก 4 GB และหน่วยความจำสำรอง 1 TB
2. ซอฟต์แวร์
 - 2.1 Dialogflow (Natural Language Understanding as a Service)
 - 2.2 Node.js Framework
 - 2.3 Firebase
 - 2.4 ThaipyNLP Package
3. เซอร์วิส
 - 3.1 Google Cloud Platform

งบประมาณ

- | | | |
|--|------|-------------------|
| 1. SSD ความจุ 256 GB จำนวน 2 ชิ้น | ราคา | 7,500 บาท |
| 2. Azure Cloud Platform | ราคา | 3,000 บาท |
| 3. ค่าถ่ายเอกสาร และทำรูปเล่ม | ราคา | 300 บาท |
| 4. กระดาษถ่ายเอกสาร A4 80 แกรม จำนวน 1 รีม | ราคา | 200 บาท |
| | รวม | <u>11,000 บาท</u> |

หมายเหตุ ทั้งนี้งบประมาณที่ตั้งไว้อาจเปลี่ยนแปลงตามความเหมาะสมและขออภัยทุกประการ

เอกสารอ้างอิง

- [1] Chula MOOC. (2561). **FAQ (คำถามที่พบบ่อย)**. สืบค้นเมื่อ 10 สิงหาคม 2561, จาก <https://mooc.chula.ac.th/faq>.
- [2] Petch Kruapanich. (2561). **พัฒนาเชทบอทแบบ Rule-based approach VS AI based approach**. สืบค้นเมื่อ 25 พฤศจิกายน 2561, จาก <https://medium.com/@petchkruapanich/พัฒนาเชทบอทแบบ-rule-based-approach-vs-ai-based-approach-3a32bee13ce3>.
- [3] สุฤดี นัทรไตรมงคล, “การรู้จำและการจำแนกประเภทของชื่อเฉพาะภาษาไทย”, จุฬาลงกรณ์มหาวิทยาลัย, 2548
- [4] Dialogflow. (2018). “**How does fulfillment work?**”. Retrieved September 15, 2018, from <https://dialogflow.com/docs/intro/fulfillment>.
- [5] Ashish Kathuria, Bernard J. Jansen, Carolyn Hafernik, Amanda Spink, (2010) "Classifying the user intent of web queries using k-means clustering", Internet Research, Vol. 20 Issue: 5, pp.563-581, <https://doi.org/10.1108/10662241011084112>
- [6] J. Wang, G. Cong, X. W. Zhao, X. Li, "Mining user intents in twitter: A semi-supervised approach to inferring intent categories for tweets", AAAI, 2015.
- [7] Z. Chen, B. Liu, M. Hsu, M. Castellanos, R. Ghosh, "Identifying intention posts in discussion forums", NAACL, pp. 1041-1050, 2013.
- [8] C. Cui and W. Mao, "Recognize user intents in online interactions from massive social media data," 2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA), Beijing, 2017, pp. 11-15.

ภาคผนวก ข

รายชื่อคำหยุด

รายชื่อคำหยุดที่จัดทำขึ้นเพื่อใช้ตรวจหาคำหยุดเหล่านี้ในประโยคและตัดออกจากประโยค มีดังต่อไปนี้

- | | | |
|----------------|---------------|-------------|
| - นับแต่นั้น | - ระยะๆ | - ครา |
| - มื้อ | - ยิ่งงั้น | - จรดกับ |
| - เหตุนั้น | - ดังกับว่า | - ยืนยัน |
| - เป็นเพราะ | - รับ | - ช่วงหลัง |
| - เสียยิ่งนัก | - ไหน | - ทุกๆ |
| - ช่วงก่อน | - อาจเป็นด้วย | - ไม่ค่อย |
| - บ่อยครั้ง | - เธอ | - ขาด |
| - ที่สุด | - หรือเปล่า | - รวม |
| - ได้ที่ | - ตลอดปี | - เชื่อว่า |
| - มา | - คราวนี้ | - อันจะ |
| - ครั้งที่ | - ทัน | - ขวางๆ |
| - คุณ | - ความ | - เหล่านี้ |
| - เพื่อให้ | - ยิ่งงั้น | - ข้างเคียง |
| - ได้ | - อนึ่ง | - เป็นต้น |
| - ยิ่งแล้ว | - ขณะที่ | - ประการใด |
| - ใน | - ภายนอก | - ยืมนาน |
| - ครั้งก่อน | - แรก | - พอเหมาะ |
| - พอแล้ว | - เริ่ม | - มื้อ |
| - อย่างไม่เสีย | - อย่างมาก | - แต่นั้น |
| - ครั้งหนึ่ง | - อันได้แก่ | - ผ่าน |
| - คราวนี้ | - สู่ | - ยาก |
| - ตลอดทั่ว | - เขา | - มัก |
| - จึงจะ | - เป็นดัง | - จนบัดนี้ |
| - ด้วยเพราะ | - คราวหลัง | - ออก |
| - รวมกัน | - แท้ | - ใคร่ |

- นี่	- เฉพาะ	- ก็ตามแต่
- ทุกอัน	- คราวใด	- ทั้งสิ้น
- นี่แหละ	- นับแต่ที่	- คราวๆ
- อื่น	- เหตุ	- ยอมรับ
- แห่งใด	- นอกนั้น	- ทุกครั้ง
- ใหม่ๆ	- กลับ	- เป็นอาทิ
- กลุ่มก้อน	- จ๊ะ	- ถึงบัดนั้น
- มิฉะนั้น	- เอา	- พวกนั้น
- ยิ่งแต่	- ฉะนี้	- เชื้อถือ
- นานๆ	- น้อยกว่า	- นั้น
- เหตุนี้	- จวบกับ	- ทั้งๆ
- ค่ะ	- ต้อง	- ครั้งละ
- ได้แต่	- พอเพียง	- ดังกับว่า
- ร่วมกัน	- อันละ	- ร่วมมือ
- บอกว่า	- หรือไง	- ดังเคย
- พอๆ	- เหล่านั้น	- แก่ใจ
- เนี่ย	- รวมด้วย	- นู่น
- ภายหน้า	- น้อยๆ	- คงจะ
- แต่ทว่า	- ช่วงหน้า	- จนกว่า
- อันๆ	- อย่างที่	- ได้รับ
- หรือ	- บางกว่า	- ภายภาคหน้า
- จริงๆจังๆ	- เมื่อนี้	- รวมๆ
- ต่างๆ	- แม้ว่าจะ	- รือว่า
- แค่เพียง	- ก่อนหน้า	- ฟัง
- ที่	- แต่ไหน	- ลง
- จากนี้	- เกือบๆ	- แห่ง
- ตลอดมา	- ข้าๆ	- โด
- ด้วยเหตุนี้	- ส่วนใหญ่	- เร็ว
- อย่างนี้	- เป็นเพียงว่า	- มุ่งหมาย
- เท่ากัน	- เสีย	- นี้นา
- ไว้	- ครั้งไหน	- ผู้

- | | | |
|--------------|----------------|-----------------|
| - ยิ่งขึ้น | - เป็นที่ | - รือ |
| - ส่วนที่ | - แต่ละ | - ซึ่ง |
| - เหตุไร | - ทุกที่ | - ตลอดกาลนาน |
| - จ๊ะ | - เพื่อที่ | - จัดงาน |
| - ครานั้น | - จัดหา | - ต่างก็ |
| - เต็มไปหมด | - นอกจาก | - บางคราว |
| - แห่งไหน | - และ | - ถึงแม้ |
| - ต่าง | - ยิ่งจะ | - สิ้น |
| - จนกระทั่ง | - เพียงเพราะ | - สูงๆ |
| - นั่นไง | - ไป | - ถูกๆ |
| - ก็แค่ | - ค่อยไปทาง | - กลุ่มๆ |
| - นับจากนี้ | - ตามๆ | - ครับ |
| - ค่อนข้างจะ | - ถึงเมื่อไร | - ซึ่งกันและกัน |
| - ก็ดี | - ซะก่อน | - ตรง |
| - ครั้งนั้น | - เท่านั้น | - เป็นด้วย |
| - คำ | - วันนี้ | - ล่าสุด |
| - ทั้งที่ | - เช่นเดียวกับ | - วันนั้น |
| - ภาคาฯ | - เสียก่อน | - อัน |
| - เล็ก | - จน | - เพียงใด |
| - ส่วนใด | - เมื่อก่อน | - ทั้งนั้นเพราะ |
| - แล้ว | - เพื่อว่า | - คราวไหน |
| - มัยนะ | - หากแม้ว่า | - จนตลอด |
| - มัยเนี่ย | - ที่แห่งนั้น | - จัง |
| - ผู้ใด | - พวกกู | - ช่วงนี้ |
| - ที่แล้ว | - ครั้งใด | - ช่วงระหว่าง |
| - หนอย | - จง | - จนเมื่อ |
| - ที่ใด | - แก | - นับตั้งแต่ |
| - คราวนั้น | - เกี่ยวกับ | - ทุกวันนี้ |
| - มัยนั้น | - ด้วยว่า | - เพิ่งจะ |
| - พยายาม | - ที่ไร | - ไม่ค่อยเป็น |
| - ก็จะ | - ทุกวัน | - เสรีจ |

- เคย
- ช่วงที่
- เกิด
- เช่นไร
- ฟัง
- ขณะใดๆ
- นี่ไง
- ครั้งหลัง
- สมัย
- จับ
- บางแห่ง
- เผื่อว่า
- เมื่อครั้ง
- ค่อนมาทาง
- ต่อกัน
- เชื่อมั่น
- เป็น
- จัดการ
- ส่วนน้อย
- ตาม
- พุด
- หารือ
- ในช่วง
- พอที่
- ทว่า
- นำ
- กันเถอะ
- ครั้งนี้
- เสมือนว่า
- พวกมึง
- อันที่จริง
- มักจะ
- บ่อยกว่า
- แม่
- หากว่า
- พวกโน้น
- ทุกชั้น
- ทำไร
- เห็นควร
- ใครจะ
- จริงจัง
- มองว่า
- วัน
- เป็นแต่
- เฉยๆ
- ไม่เป็นไร
- ตลอดไป
- แต่จะ
- ก่อน
- นี่แหละ
- แต่ก็
- เป็นอันมาก
- ปรับ
- นำมา
- สำหรับ
- จนแม่
- ในระหว่าง
- มั้ยละ
- เพิ่ง
- ยังไง
- อันเนื่องมาจาก
- ประการฉะนี้
- ผล
- ทาง
- ถึงเมื่อใด
- ครั้งคราว
- ตลอดทั้ง
- กันนะ
- คราไหน
- าล
- แค่ว่า
- ขณะนี้
- ให้แต่
- มีแต่
- ทั้งหลาย
- ยาว
- ใจ
- เป็นเพียง
- พร้อมกัน
- ทำให้
- เมื่อคราวก่อน
- เป็นๆ
- พวกที่
- ตามที่
- ครั้งกระนั้น
- คล้ายกันกับ
- ทุกคราว
- ข้างล่าง
- เปลี่ยนแปลง
- แหะละ
- เพียงเพื่อ
- คล้ายกับว่า
- ก็ตาม

- คล้ายกัน	- ทั้ง	- เช่นเดียวกัน
- ประมาณ	- เป็นที่	- ไฉน
- เป็นเพื่อ	- หาก	- ผิด
- ก็ตามที่	- จนแน่น	- ด้วยประการฉะนี้
- บอกแล้ว	- เต็มไปด้วย	- เช่นที่ว่า
- ที่ได้	- แห่งนี้	- เสียนี้
- เล็กน้อย	- จู่ๆ	- น้อย
- ถึง	- ตั้งเก่า	- แน่นว่า
- พวกมัน	- ตามด้วย	- ตั้งแต่
- พวกกัน	- สมัยนี้	- ถึงเมื่อ
- หาใช่	- ตลอดระยะเวลา	- ด้วย
- แยะ	- ไกล	- ภายหลัง
- หมดสิ้น	- คราว	- พร้อมด้วย
- ยังงี้	- เสียด้วย	- เพื่อ
- ขอ	- หน	- กระนั้น
- นะ	- คุณๆ	- ล้วนจน
- ถ้า	- ที่ว่า	- ถือว่า
- เป็นอันว่า	- ผ่านๆ	- ทุกเมื่อ
- จำเป็น	- คงอยู่	- ยิ่งขึ้นไป
- ด้วยเหตุที่	- ในเมื่อ	- แต่เมื่อ
- พบว่า	- บน	- เท่าไร
- พอสมควร	- ทั้งนั้นด้วย	- จรด
- ครบถ้วน	- เกี่ยวๆ	- ทั้งเป็น
- อันที่จะ	- ซึ่งกัน	- ที่เดียว
- เท่าใด	- นอกเหนือจาก	- ที่นี้
- ยิ่งเมื่อ	- เสรีจสิ้น	- บอก
- เมื่อใด	- คราวละ	- เหล่า
- บางที่	- ที่	- พบ
- รู้ว่า	- ยิ่งนัก	- นำ
- ขวาง	- จริงๆ	- แก่
- เป็นต้นไป	- ล้วน	- ด้วยเหตุว่า

- ทุกสิ่ง	- ส่วนนั้น	- ตนเอง
- อย่างไม่รู้ก็ได้	- เดียวกัน	- บัดดล
- ทุกอย่าง	- ไซ้ใหม่	- ช่วย
- นิดหน่อย	- เสียจนถึง	- ย่อม
- ก็คือ	- ตั้ง	- เฉกเช่น
- ข้าง	- ควร	- ทุกครา
- เพียงแค่	- ไม่	- ไม่ค่อยจะ
- คราที่	- รวมถึง	- อยาก
- พอที่	- ทัว	- นี้เอง
- บ้าง	- ตามแต่	- เน้น
- จนถึง	- ไหนๆ	- เปิด
- นี้แหละ	- ครั้งหลังสุด	- ข้างบน
- เช่น	- จวบ	- แต่ต้อง
- นอกจากนี้	- เคยๆ	- ทำไม
- เมื่อไร	- มิได้	- นัก
- ด้วยที่	- จาก	- ค่อยๆ
- เพราะฉะนั้น	- เกิน	- ยิ่งคง
- ที่ๆ	- ถึงอย่างไร	- ประสบ
- ครั้งครา	- ยืนยง	- ใหญ่
- ใหญ่โต	- เท่าไหร่	- กันเอง
- อาจเป็น	- พร้อมเพียง	- ตลอดกาล
- อย่างไม่	- ดังเคย	- ซ้ำๆ
- ให้ดี	- สิ้น	- เหลือ
- ช่วงแรก	- เดียว	- ยอม
- เยอะ	- ก็	- เยอะๆ
- ทีละ	- ใครๆ	- อยู่
- พร้อมที่	- สิ่งนั้น	- ใหม่
- เช่นก่อน	- ภายใน	- สูงกว่า
- นะ	- นาน	- ล้วนแต่
- เสรีจกกัน	- เต็มๆ	- เลย
- ข้างๆ	- เรียก	- ขึ้น

- เสียจนกระทั่ง	- ทรง	- ถึงแก่
- ที่นั่น	- ทั้งนั้น	- ปัจจุบัน
- นี้เอง	- คราวหนึ่ง	- สำคัญ
- หนึ่ง	- ทุกทาง	- ดังกล่าว
- เช่นเคย	- คุ้น	- เป็นการ
- ปฏิบัติ	- โดดๆ	- ง่ายๆ
- นั่นเป็น	- เกี่ยวกัน	- พวก
- เหตุผล	- แต่ไร	- เมื่อคราว
- ถูกต้อง	- สืบเนื่อง	- เช่นดั่งที่
- แค่นั้น	- เป็นเพราะว่า	- อันไหน
- เพื่อที่	- ข้าพเจ้า	- คือ
- ทำ	- หนักๆ	- ที่แท้
- ช่วง	- ครั้งๆ	- บางที
- แค่นี้	- ส่วนเกิน	- ใกล้เคียง
- นอกเหนือ	- กระผม	- ด้วยเหตุนี้
- ที่ไหน	- เมื่อเย็น	- คง
- คราวหนึ่ง	- กว้างขวาง	- นำพา
- ประการหนึ่ง	- มิ	- จะได้
- มัน	- สิ้นกาลนาน	- ทีเดียว
- ภาค	- กำหนด	- เสียนี้กระไร
- สูงส่ง	- ฝ่าย	- แค่ว่า
- เรื่อยๆ	- พอดี	- เพิ่มเติม
- ยิ่ง	- กำลัง	- สบาย
- ตั้ง	- ด้วยเหมือนกัน	- จัดแจง
- จด	- เช่นกัน	- มันๆ
- เราๆ	- สิ่งใด	- เสียแล้ว
- บัดนี้	- บัดนั้น	- ก็ต่อเมื่อ
- จัดทำ	- ให้	- เก็บ
- วันใด	- เกิด	- นับจากนั้น
- เพื่อที่จะ	- ทันที	- พวกเขา
- ช่วงๆ	- อย่างละ	- ด้าน

- มุ่ง	- ซึ่งได้แก่	- ญ
- ยกให้	- ดังเก่า	- อย่างน้อย
- ฟั่นๆ	- ของ	- พอ
- จำ	- บัดเดี๋ยวนี้	- ใหญ่ๆ
- บ่อย	- จัด	- แต่เดิม
- เข้า	- ส่วน	- เพียงแต่
- ผิดๆ	- จวบจน	- สิ่งนี้
- แค	- ชะจนกระทั่ง	- พอจะ
- ได้มา	- ถึงแม้ว่า	- ทุกคน
- นอก	- ค่อย	- มีง
- นี้	- เช่นดั่งเก่า	- สามารถ
- แสดงว่า	- มากมาย	- พอกัน
- เห็นว่า	- กว่า	- เรื่อย
- เพราะว่า	- เคย	- นั้นไว
- อย่างดี	- สมัยก่อน	- ทุกหน
- เท่าที่	- บางครา	- คราวโน้น
- ให้ไป	- จวนจะ	- คิดว่า
- ครบ	- ทุกตัว	- พอสม
- พอควร	- อาจจะ	- ๗
- ยืนยาว	- คิด	- ค่อนข้าง
- จำพวก	- สมัยโน้น	- หมด
- โดย	- หลังจาก	- ทิ้งปวง
- เห็น	- เร็วๆ	- ตลอดศก
- ช่วงนั้น	- นับ	- อย่างหนึ่ง
- กันใหม่	- รี	- ถ้าหาก
- จากนี้ไป	- เพราะ	- ทีเถอะ
- เปิดเผย	- คราวหน้า	- ในที่
- ปราบฏ	- ทุก	- ขณะนั้น
- เยอะแยะ	- เมื่อ	- จริง
- แต่เพียง	- เป็นแต่เพียง	- แล้วเสร็จ
- เมื่อคราวที่	- สิ่งไหน	- การ

- แล้วแต่	- เช่นนี้	- ไม่ใช่
- อันใด	- ทั้งนี้	- เสียยิ่ง
- ทีๆ	- เช่นที่เคย	- เนื่องจาก
- กระทั่ง	- ฉะนั้น	- ง่ายๆ
- เช่นดัง	- เพียง	- จากนั้น
- ราย	- ซึ้งๆ	- นับแต่
- นั้นแหละ	- อย่าง	- นั้นๆ
- เถอะ	- รับรอง	- กว้าง
- ไกลๆ	- ทำให้	- ครั้น
- ให้แก่	- อย่างยิ่ง	- ช้านาน
- นั่นเอง	- เรา	- พวกนี้
- ถ้าจะ	- แล้วกัน	- เห็นจะ
- ยิ่งกว่า	- มุ่งเน้น	- ฝ่ายใด
- เห็นแก่	- อย่างไร	- พวกท่าน
- ยิ่งใหญ่	- เช่นที่	- ขณะหนึ่ง
- กว้างๆ	- ก็แล้วแต่	- สมัยนั้น
- นอกจากนั้น	- กันดีกว่า	- เช่นใด
- อย่างเดียว	- เช่นดังว่า	- แต่ถ้า
- ต่างหาก	- พร้อมทั้ง	- พร้อมกับ
- แห่งนั้น	- ไม่ว่า	- ซึ่งก็
- จะ	- หมดกัน	- เพียงไร
- ช้า	- ซึ่งก็คือ	- ร่วมด้วย
- ใคร	- ณ	- อย่างนั้น
- ทุกที	- หลัง	- เท่านี้
- ครั้ง	- เกือบจะ	- เรียบ
- ง่าย	- ทั้งตัว	- ด้วยเหตุเพราะ
- กล่าวคือ	- สูงสุด	- ก็ได้
- ส่วนดี	- อย่างโน้น	- ว่า
- ช่วงท้าย	- ตลอดถึง	- หาความ
- พอตัว	- บางๆ	- ย่อย
- กำลังจะ	- น่าจะ	- อาจ

- ที่ละ	- หรือไม่	- วันไหน
- เข้าใจ	- ดั้งกับ	- ประกอบ
- ที่แท้จริง	- เช่นดังก่อน	- บ่อยๆ
- เสร็จแล้ว	- เช่นเมื่อ	- บางขณะ
- เล่าว่า	- ข้างต้น	- หรือไร
- อื่นๆ	- เอ็ง	- ไซ้
- ถึงจะ	- เสียมันเอง	- นิด
- รวด	- สิ่ง	- จนขณะนี้
- เชื้อ	- นับแต่นี้	- ทั้งคน
- มี	- ค่อน	- กัน
- เอง	- พา	- บางครั้ง
- นั้น	- ฉั้น	- ได้แก่
- ภายภาค	- ยัง	- บาง
- มาก	- นิดๆ	- เกี่ยวเนื่อง
- รวดเร็ว	- ทั้งหมด	- ไตๆ
- เสมือนกับ	- ให้มา	- ถึงบัดนี้
- สั้นๆ	- เผื่อ	- จวน
- หนอ	- ส่ง	- เผื่อจะ
- หรือยัง	- ทันใดนั้น	- ตลอดเวลา
- ที่จริง	- คราใด	- ข้า
- เมื่อครั้งก่อน	- พวกแก	- จนทั่ว
- แห่งโน้น	- คล้ายกับ	- เสียมัน
- พวกนั้น	- ระหว่าง	- รวมทั้ง
- ชะจน	- กันดีไหม	- ถือ
- มิใช่	- เมื่อเช้า	- ต่อ
- ไป	- ยิ่งจน	- เนี่ยเอง
- นอกจากที่	- กันและกัน	- เกินๆ
- เสีย	- มากกว่า	- จำ
- นำ	- เล็กๆ	- ครบครัน
- ไซ้	- ขณะใด	- จวนเจียน
- เสียจน	- สุดๆ	- สูง

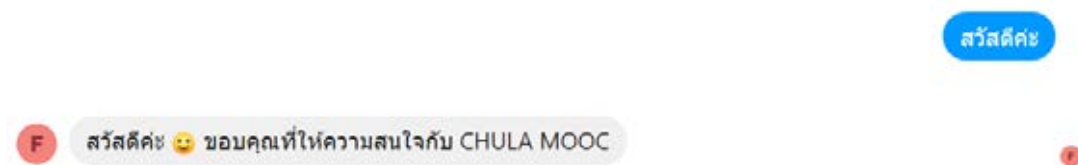
- อย่างเช่น	- พวกเขา	- พวกคุณ
- เกือบ	- ยาวนาน	- นาง
- แต่	- พร้อม	- ปรากฏว่า
- เมื่อวันวาน	- ถูก	- จึง
- เมื่อคืน	- นางสาว	- ได้
- ยก	- ทั้งที่	- แม้แต่
- เป็นอัน	- อดีต	- พวกฉัน
- อย่างไหน	- มอง	- ขึ้น
- แบบ	- ขณะเดียวกัน	- เท่า
- จึงเป็น	- ละ	- ส่วนมาก
- ส่วนด้อย	- แต่ว่า	- เมื่อนั้น
- แต่อย่างไร	- เกี่ยวข้อง	- ยังจะ
- ชะ	- เป็นที่สุด	- แม้กระทั่ง
- ฯลฯ	- เสร็จสมบูรณ์	- ระยะ
- ไกลๆ	- แคไหน	- ค่ะ
- กลุ่ม	- ภาย	- ก่อนหน้านี้
- ช่วงถัดไป	- คล้าย	- ตันฯ
- แสดง	- เป็นต้นมา	- ประการ
- เมื่อวาน	- อีก	- จัดตั้ง
- หลาย	- เช่นนั้น	- ๆ
- ชะจนถึง	- แต่ก่อน	- หาก मैं
- ทำๆ	- หาก मैं	- ด้วยกัน
- ตลอดทั่วถึง	- คล้ายว่า	- กระทำ
- ทันทีทันใด	- จึงๆ	- ชาว
- ดั่งกับ	- ที่ซึ่ง	- นาย
- กับ	- อย่งไรก็	- ด้วยเช่นกัน
- คราวก่อน	- ภายได้	- ตลอด
- เหลือเกิน	- ช่วงต่อไป	- สุด
- ถึงแม้จะ	- อันที่	- กล่าว
- แยะๆ	- ตั้ง	- ไร
- เปลี่ยน	- คราวที่	- จัดให้

- จำ
- ทุกแห่ง
- เพียงไหน
- เพียงพอ
- เมื่อไหร่
- อะไร
- เพิ่ม
- เท่ากับ
- ร่วม
- ปิด
- ตลอดวัน
- ขณะ
- ตรงๆ
- นอกจากว่า
- ตลอดจน
- หน่อย
- แต่ที่
- ตน
- ก่อนๆ
- แท้จริง
- เป็นอันๆ
- ตลอดทั่วทั้ง
- เช่นนั่นเอง
- ทั้งหมด

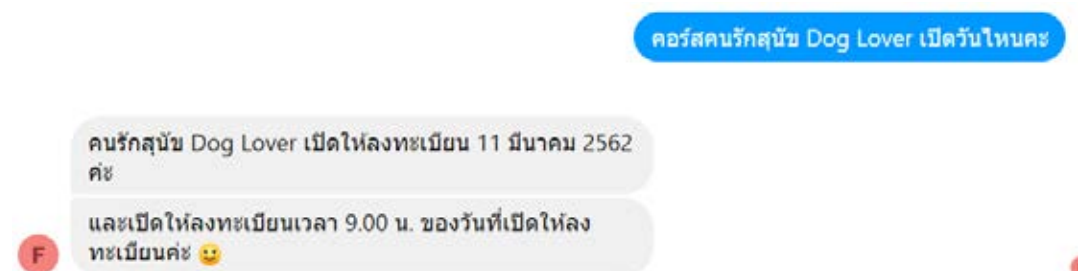
ภาคผนวก ค

ภาพหน้าจอการทำงานจริงของแช็ตบ็อต

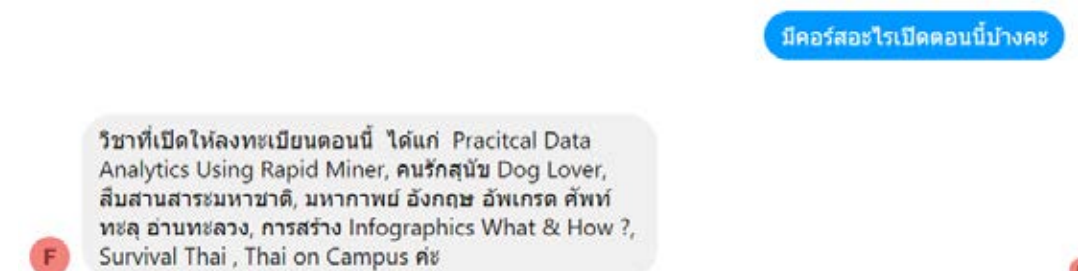
1. ภาพหน้าจอการทำงานจริงของแช็ตบ็อตสำหรับเจตนาทักทาย



2. ภาพหน้าจอการทำงานจริงของแช็ตบ็อตสำหรับเจตนาสอบถามเวลาเปิดรายวิชา



3. ภาพหน้าจอการทำงานจริงของแช็ตบ็อตสำหรับเจตนาสอบถามรายวิชาที่เปิดสอน



4. ภาพหน้าจอการทำงานจริงของแช็ตบ็อตสำหรับเจตนาสอบถามรายละเอียดวิชา

อยากดูรายละเอียดวิชา English for careers I

F กรุณาเลือกข้อมูลที่ต้องการทราบของรายวิชา English For Careers I For Hotel Staff

วันเริ่มลงทะเบียน วันสิ้นสุดการเรียน จำนวนบทเรียน จำนวนผู้เรียนที่รับ เกณฑ์การ >

อยากดูรายละเอียดวิชา English for careers I

F กรุณาเลือกข้อมูลที่ต้องการทราบของรายวิชา English For Careers I For Hotel Staff

< จำนวนบทเรียน จำนวนผู้เรียนที่รับ เกณฑ์การวัดผล เกี่ยวกับ/เนื้อหา อาจารย์ผู้สอน

เกณฑ์การวัดผล

F มีการวัดและประเมินผลผ่านแบบทดสอบย่อย (Quiz) และแบบทดสอบหลังเรียน (Posttest) โดยจะแบ่งเป็นคะแนนจาก Quiz เท่ากับ 50 คะแนน และคะแนนจาก Posttest เท่ากับ 50 คะแนน ทั้งนี้ผู้เรียนต้องทำคะแนนรวมทั้งหมดให้ได้ร้อยละ 60 ขึ้นไปจึงจะสามารถขอรับ Certificate of Completion ได้ หมายเหตุ ผู้เรียนจะมีสิทธิ์ทำแบบทดสอบหลังเรียน (Posttest) เมื่อเข้าร่วมกิจกรรมการเรียน (Course Progress) มากกว่า 80% ขึ้นไป

5. ภาพหน้าจอการทำงานจริงของแช็ตบ็อตสำหรับเจตนาเงื่อนไขการได้รับใบประกาศนียบัตร

ทำยังไงถึงได้ใบ cer

F ผู้เรียนจะได้รับ Certificate of Completion เมื่อระบบทำการตรวจสอบว่าผู้เรียนมีคะแนนผ่านตามเกณฑ์การวัดและประเมินผล คือ ได้คะแนนมากกว่า 60% และผู้เรียนเข้าร่วมกิจกรรมการเรียน (Course Progress) มากกว่า 80% ขึ้นไป และจะปรากฏปุ่มให้ดาวน์โหลด Certificate of Completion ให้ที่หน้าบทเรียนและหน้า Score Detail ค่ะ

F ในกรณีที่คะแนนไม่ถึงตามเกณฑ์หรือเรียนจบหลังเวลาที่กำหนด แต่ยังคงต้องการรับ Certificate of Completion จะต้องลงทะเบียนในรายวิชาเดิมในรุ่นถัดไปเท่านั้น

6. ภาพหน้าจอการทำงานจริงของแช็ตบ็อตสำหรับเจตนาเวลาประกาศรายชื่อผู้ได้รับใบประกาศนียบัตร

รอรับ certificate ได้เมื่อไหร่ครับ

ผู้เรียนจะได้รับ Certificate of Completion หลังจากระบบได้ทำการตรวจสอบว่าผู้เรียนได้ผ่านเกณฑ์การประเมินเรียบร้อยแล้ว โดยจะปรากฏปุ่มให้ดาวน์โหลด Certificate of Completion ที่หน้าบทเรียนและหน้า Score Detail ค่ะ

*ซึ่งแต่ละคนจะได้รับ Certificate of Completion ไม่พร้อมกัน ขึ้นกับเวลาที่ผู้เรียนผ่านเกณฑ์การประเมิน

7. ภาพหน้าจอการทำงานจริงของแช็ตบ็อตสำหรับเจตนาแจ้งการแก้ไขข้อมูล

สามารถแก้ไข account info ได้มั๊ยครับ

กรุณาเลือกสิ่งที่ต้องการแก้ไขข้อมูล

ข้อมูลส่วนตัวในระบบ ใบประกาศนียบัตร

ข้อมูลส่วนตัวในระบบ

รับทราบค่ะ รบกวนผู้เรียนแคปหน้าจอ Account Info ของผู้เรียน แล้วส่งมาใบนี้ด้วยค่ะ

ACCOUNT INFO

ผู้เรียนต้องการแก้ไขข้อมูลส่วนตัว

คำนำหน้าชื่อ ชื่อ-นามสกุล อาชีพ เพศ ปีเกิด ระดับการศึกษาสูงสุด e-mail >

7. ภาพหน้าจอการทำงานจริงของแชทบ็อตสำหรับเจตนาแจ้งการแก้ไขข้อมูล (ต่อ)

คำนำหน้าชื่อ

ผู้เรียนต้องการเปลี่ยนเป็นคำนำหน้าใด มีตัวเลือกดังนี้ Miss , Mr. , Mrs.

F เลือกคำนำหน้าที่ถูกต้องของผู้เรียน

Miss Mr. Mrs.

Mr.


ผู้เรียนต้องการแก้ไขข้อมูลส่วนตัวในระบบ โดยแก้ คำนำหน้าชื่อ เป็น Mr. ใช่หรือไม่

F ใช่ ไม่

ใช่

F รับทราบค่ะ จะทำการส่งเรื่องให้ผู้ดูแลตรวจสอบและแก้ไขต่อไปค่ะ

[Chatbot] แจ้งขอแก้ไขข้อมูล ปิดหน้าต่าง X

 ploy_3338@hotmail.com ผ่าน sendgrid.net
5:00 น.

🗨️ อังกฤษ > ไทย แปลข้อความ

ปัญหา : ผู้เรียนต้องการแก้ไขข้อมูลส่วนตัวในระบบ โดยแก้คำนำหน้าชื่อ เป็น Mr.
link รูปหน้า account info https://scontent.xx.fbcdn.net/v/t1.15752-9/60088258_501954761323961_631056585568813056_n.jpg?_nc_cat=100&_nc_ohc=MkIBVVDjva6IhS7xkhGNyCZBall&NMhHnSWoFz-Nw8TrSRKkqWNUim&_nc_ad=z-m&_nc_cid=0&_nc_zor=0&_nc_ht=scontent.xx&oh=a25537730640

🔍 ค้นพบ ➡ ส่งต่อ

8. ภาพหน้าจอการทำงานจริงของแช็ตบ็อตสำหรับเจตนาแจ้งปัญหาตกหล่นของใบประกาศนียบัตร

ยังไม่ได้ cer

F คุณแจ้งไม่ได้รับ Certificate ในวิชาใด ?

ภาษาอาหรับเบื้องต้น

รบกวนผู้เรียนให้ข้อมูลตามรูปแบบดังนี้ (พิมพ์ในข้อความเดียวกัน)
Learner ID :
ชื่อ - นามสกุล :

F *กรณีที่ไม่ปรากฏปุ่มดาวน์โหลด Certificate of Completion นั้นมีความเป็นไปได้ว่าผู้เรียนไม่ผ่านเกณฑ์การประเมินหรือเลขกำหนดเวลาเรียนได้ค่ะ

Learner ID : CV999999
ชื่อ - นามสกุล : พิชชาภา พวงษ์

F ผู้เรียนยังไม่ได้ Certificate of Completion ในวิชา ภาษาอาหรับเบื้องต้น
Learner ID : CV999999
ชื่อ-นามสกุล : พิชชาภา พวงษ์
ข้อมูลถูกต้องตามนี้ใช่หรือไม่

ใช่

F รับทราบค่ะ จะทำการส่งเรื่องให้ผู้ดูแล

[Chatbot] แจ้งปัญหา Certificate ตกหล่น กล่องจดหมาย X



ploy_3338@hotmail.com ผ่าน sendgrid.net

ถึงฉัน ▾

ปัญหา : ยังไม่ได้รับ Certificate วิชา ภาษาอาหรับเบื้องต้น Learner ID : CV999999 ชื่อ-นามสกุล : พิชชาภา พวงษ์

← ตอบ

➡ ส่งต่อ

9. ภาพหน้าจอการทำงานจริงของแชทบ็อตสำหรับเจตนาแจ้งปัญหาเปิดไฟล์ใบประกาศนียบัตรไม่ได้


เปิดไฟล์ใบ cer ไม่ได้ค่ะ

รับทราบค่ะ จะทำการส่งเรื่องให้ผู้ดูแล

*กรณีที่ไม่สามารถดาวน์โหลด Certificate of Completion ได้นั้นมีความเป็นไปได้ว่าผู้เรียนไม่ผ่านเกณฑ์การประเมินหรือเลยกำหนดเวลาเรียนได้ค่ะ

[Chatbot] แจ้งปัญหาการดาวน์โหลดไฟล์ Certificate

กส่งจดหมาย X

 ploy_3338@hotmail.com ผ่าน sendgrid.net
ถึง ฉัน

ปัญหา : เปิดไฟล์ใบ cer ไม่ได้ค่ะ

ตอบ ส่งต่อ

10. ภาพหน้าจอการทำงานจริงของแชทบ็อตสำหรับเจตนาคำถามที่พบบ่อยข้อที่ 1

CHULA MOOC คืออะไร

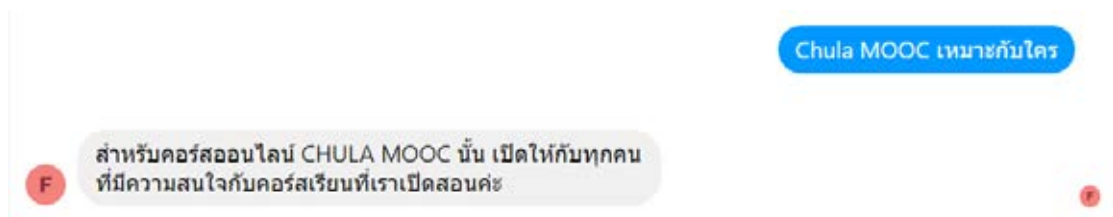
สำหรับการเรียนออนไลน์ CHULA MOOC จะเป็นการเรียนในรูปแบบการดูวิดีโอและทำแบบทดสอบผ่านคอมพิวเตอร์ จากสถานที่ใดก็ได้ตามความสะดวกของผู้เรียน และสามารถเรียนหลายวิชาได้ตามความสนใจ ภายในระยะเวลาที่กำหนดของวิชานั้นค่ะ

11. ภาพหน้าจอการทำงานจริงของแชทบ็อตสำหรับเจตนาคำถามที่พบบ่อยข้อที่ 2

มีค่าใช้จ่ายหรือไม่

สำหรับคอร์สเรียนบน CHULA MOOC ผู้ที่สนใจสามารถลงทะเบียนเรียนได้ฟรี ไม่มีค่าใช้จ่ายค่ะ

12. ภาพหน้าจอการทำงานจริงของแช็ตบ็อตสำหรับเจตนาคำถามที่พบบ่อยข้อที่ 3



13. ภาพหน้าจอการทำงานจริงของแช็ตบ็อตสำหรับเจตนาคำถามที่พบบ่อยข้อที่ 4



14. ภาพหน้าจอการทำงานจริงของแช็ตบ็อตสำหรับเจตนาขอบคุณ



ประวัติผู้เขียน



นายณัฐกิตต์ สมอรัมย์

ประวัติ

วัน เดือน ปีเกิด : 2 ตุลาคม 2539

สถานที่เกิด : กรุงเทพมหานคร

การศึกษา : สาขาวิทยาการคอมพิวเตอร์ ภาควิชาคณิตศาสตร์
และวิทยาการคอมพิวเตอร์ จุฬาลงกรณ์มหาวิทยาลัย

อีเมล : natt.sar1996@gmail.com



นางสาวพิชชาภา พะวงษ์

ประวัติ

วัน เดือน ปีเกิด : 16 สิงหาคม 2539

สถานที่เกิด : กรุงเทพมหานคร

การศึกษา : สาขาวิทยาการคอมพิวเตอร์ ภาควิชาคณิตศาสตร์
และวิทยาการคอมพิวเตอร์ จุฬาลงกรณ์มหาวิทยาลัย

อีเมล : ploypitch3338@gmail.com