

การพัฒนาแชทบอทปัญญาประดิษฐ์ภาษาไทยเพื่อสนับสนุน
การให้คำปรึกษาวิชาการสำหรับนักศึกษาระดับอุดมศึกษา
THE DEVELOPMENT OF THAI ARTIFICIAL INTELLIGENCE CHATBOT
FOR SUPPORTING ACADEMIC CONSULTANCY
FOR TERTIARY STUDENTS

ธินาภณ นิตยวิทย์ และทศนันท์ ตรีนันทรรัตน์*

Thinaphan Nithiyuwit and Tassanan Treenuntharath*

บทคัดย่อ

การคงไว้ของจำนวนนักศึกษาในมหาวิทยาลัยรัฐและเอกชนในยุควิกฤติที่จำนวนนักศึกษาลดลงอย่างต่อเนื่อง ประกอบด้วยสิ่งสำคัญคือการมีระบบการให้คำปรึกษาที่ดีและมีความต่อเนื่อง ซึ่งเป็นโจทย์สำคัญในการวิจัยครั้งนี้ โดยมีวัตถุประสงค์ 1) เพื่อศึกษาชุดข้อมูลและฐานความรู้ของแชทบอทปัญญาประดิษฐ์ภาษาไทยเพื่อสนับสนุนการให้คำปรึกษาวิชาการสำหรับนักศึกษาระดับอุดมศึกษา 2) เพื่อพัฒนาแชทบอทปัญญาประดิษฐ์ภาษาไทยเพื่อสนับสนุนการให้คำปรึกษาวิชาการสำหรับนักศึกษาระดับอุดมศึกษา และ 3) เพื่อประเมินประสิทธิภาพของแชทบอทปัญญาประดิษฐ์ภาษาไทยเพื่อสนับสนุนการให้คำปรึกษาวิชาการสำหรับนักศึกษาระดับอุดมศึกษา ขั้นตอนการดำเนินงานวิจัยแบ่งได้ 3 ขั้นตอน คือ 1) เก็บรวบรวมข้อมูลจากอาจารย์ที่ปรึกษา นักศึกษา และผู้อำนวยการฝ่ายวิชาการและงานทะเบียน มหาวิทยาลัยราชภัฏเพชรบูรณ์ และมหาวิทยาลัยพิษณุโลก จำนวน 54 คน เพื่อนำมาพัฒนาชุดข้อมูลและฐานความรู้ของแชทบอทปัญญาประดิษฐ์ 2) พัฒนาแชทบอทปัญญาประดิษฐ์ภาษาไทย ประกอบด้วย ฐานข้อมูลเชิงสัมพันธ์ ฐานความรู้เชิงความหมายออนโทโลยี การประมวลผลภาษาธรรมชาติ และเว็บอินเทอร์เฟซในการติดต่อกับแชทบอทปัญญาประดิษฐ์ และ 3) ประเมินประสิทธิภาพแชทบอท พบว่า แชทบอทปัญญาประดิษฐ์มีประสิทธิภาพในการทำนายคำตอบการสนทนาด้วย K-NN ที่ K=3 และ K=5 ได้ถูกต้อง 100% สืบค้นความรู้จากออนโทโลยีได้ถูกต้อง 93.6% และผู้มีความพึงพอใจต่อแชทบอทในภาพรวมอยู่ในระดับมาก

คำสำคัญ: แชทบอทปัญญาประดิษฐ์ภาษาไทย การให้คำปรึกษาวิชาการ นักศึกษาระดับอุดมศึกษา

คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏเพชรบูรณ์ อำเภอเมือง จังหวัดเพชรบูรณ์ 67000

Faculty of Science and Technology, Phetchabun Rajabhat University, Muang District, Phetchabun Province 67000

*corresponding author e-mail: tassanan.t@pcru.ac.th

Received: 29 August 2020; Revised: 9 October 2020; Accepted: 24 October 2020

Abstract

In crisis era, that makes students continuously decrease. The important way that helps solve the problem is to have an excellent consulting system. This research aims 1) to study the dataset and knowledge base of the Thai artificial intelligence chatbot (TAICB) for supporting academic consultancy for tertiary students, 2) to develop the TAICB for supporting academic consultancy for tertiary students, and 3) to assess the performance of TAICB for supporting academic consultancy for tertiary students. This research was classified into three steps 1) gathering data from 54 advisors, students, and education evaluation directors of Phetchabun rajabhat university and Phitsanulok university for developing the dataset and knowledge base of the TAICB. 2) developing the TAICB consists of the relationship database, ontology knowledge base, natural language processing concept and web interface to connect with the TAICB, and 3) evaluating the performance of TAICB. This research found that the artificial intelligence chatbot has accuracy in predicting the conversation answers by K-NN (K=3 and K=5) 100%, searching knowledge from ontology has accuracy 93.6% and users satisfied to the TAICB at a high level.

Keywords: Thai artificial intelligence chatbot, Academic consultancy, Tertiary students

บทนำ

กลไกของระบบให้คำปรึกษาในมหาวิทยาลัยของทั้งภาครัฐและเอกชนที่มีการบริหารจัดการอย่างมีประสิทธิภาพ และนักศึกษาสามารถเข้าถึงได้ เป็นบทบาทสำคัญต่อความสำเร็จของการพัฒนานักศึกษาให้สามารถปรับตัวเมื่อเข้ามาสู่สถานศึกษาใหม่ และเป็นระบบตัวกลางที่ช่วยให้อาจารย์ที่ปรึกษาสามารถให้การช่วยเหลือเมื่อนักศึกษาประสบปัญหา ให้เรียนอย่างมีความสุข และสำเร็จ การศึกษาอย่างมีคุณภาพ คือ เป้าหมายของการทำหน้าที่ของอาจารย์ที่ปรึกษา อย่างไรก็ตามระบบให้คำปรึกษาบางส่วนที่ยังไม่มีประสิทธิภาพ ส่วนหนึ่งเกิดจากการใช้ระบบอาจารย์ที่ปรึกษาที่ทำงานด้วยมือ (Manual) ทำให้เกิดปัญหาตามมา ได้แก่ 1) การจัดเก็บข้อมูลไม่เป็นระเบียบ 2) การให้ข้อมูลที่ไม่ทันต่อเวลาใช้งาน และ 3) เกิดความล่าช้าและใช้เวลาทำงานนาน (Setthachotsombut, 2018) จึงเป็นการเพิ่มภาระงานให้แก่อาจารย์ที่ปรึกษา จากปัญหาข้างต้นทำให้สถาบันการศึกษาส่วนใหญ่หันมาใช้โซเชียลในการให้คำปรึกษาแก่นักศึกษามากขึ้น ไม่ว่าจะเป็น เฟซบุ๊ก และไลน์ แต่ทั้งนี้ข้อมูลที่ได้บันทึกในโซเชียลต่าง ๆ ก็ไม่ได้มีการเชื่อมต่อกับฐานข้อมูล และบางครั้งการให้คำปรึกษาอาจจะไม่ทันต่อเวลาเมื่ออาจารย์ไม่ได้ใช้โซเชียลตลอดเวลา รวมทั้งรัฐบาลได้มุ่งเน้นให้มีการนำเทคโนโลยีสารสนเทศ

เพื่อการสื่อสารมาใช้ในหลักสูตรการเรียนการสอนมากยิ่งขึ้น (Ministry of Information and Communication, 2011) ดังนั้นการใช้แชทบอทแบบปัญญาประดิษฐ์ภาษาไทยจะสามารถช่วยลดปัญหาข้างต้นได้

แชทบอท (Chatbot) คือโปรแกรมคอมพิวเตอร์ชนิดหนึ่ง ที่ถูกพัฒนาขึ้นมาให้มีบทบาทในการตอบกลับการสนทนาด้วยตัวอักษรแบบอัตโนมัติผ่าน Messaging Application เสมือนการโต้ตอบของคนจริง ๆ หรืออาจเรียกง่าย ๆ ว่า โปรแกรมตอบกลับอัตโนมัติ ซึ่งเวลานี้มีการนำไปประยุกต์ใช้มากมาย เช่น Nookhong & Kaewrattanapat (2017) ได้พัฒนาระบบอัตโนมัติเพื่อช่วยเหลือและให้คำปรึกษาเกี่ยวกับข้อมูลมหาวิทยาลัย แบ่งออกเป็น 4 กลุ่ม ได้แก่ การศึกษา ค่าใช้จ่าย ปัญหาส่วนตัว และปัญหาอื่น ๆ และเพื่อประเมินประสิทธิภาพของระบบให้คำปรึกษาอัตโนมัติ ระบบได้รับการพัฒนารูปแบบของเว็บแอปพลิเคชันผ่านมือถือ และใช้ THSplitlib สำหรับการตัดคำภาษาไทย ร่วมกับเทคนิค Dictionary-Based Approach และผลการประเมินประสิทธิภาพระบบอยู่ในระดับดี และ Villegas et al. (2020) นำเสนอวิธีการใช้ปัญญาประดิษฐ์เพื่อระบุตัวแปรและวิเคราะห์ข้อมูลของนักเรียน เพื่อจำแนกตามความต้องการ และนำไปสู่การตัดสินใจในการพัฒนาและส่งเสริมการเรียนรู้ของของนักเรียนแต่ละคน ซึ่งสถาปัตยกรรมนี้จะมีการรวมหลายระบบที่มีความจำเป็นทางการศึกษาเข้าด้วยกันเพื่อการประมวลผล โดยเฉพาะระบบจัดการเรียนรู้ (LMS: Learning Management System) โดยใช้ความสามารถของ AI และ NLP ในการรู้จำบทสนทนาของนักเรียน และปฏิสัมพันธ์ตอบคำถามอย่างถูกต้องและเหมาะสม

จากปัญหาข้างต้นผู้วิจัยสนใจที่จะพัฒนาแชทบอทปัญญาประดิษฐ์ภาษาไทยเพื่อสนับสนุนการให้คำปรึกษาวิชาการสำหรับนักศึกษาระดับอุดมศึกษา เพื่อช่วยในการจัดเก็บ วิเคราะห์ และตอบคำถามเชิงแนะนำแก่นักศึกษาระดับอุดมศึกษา รวมทั้งการวิเคราะห์ความรู้สึก เพื่อสร้างคำตอบและคำสนทนาที่เหมาะสมกับสภาพจิตใจและอารมณ์ของนักศึกษา ซึ่งนวัตกรรมนี้จะเป็นเครื่องมือที่ช่วยรับทราบปัญหา และตอบสนองการให้คำปรึกษาแก่นักศึกษาได้ทุกสถานที่ ทุกเวลา และสามารถให้คำแนะนำที่อยู่ในขอบเขตความสนใจของนักศึกษาได้อย่างแท้จริง

วิธีดำเนินการวิจัย

การวิจัยครั้งนี้ เป็นการวิจัยเชิงปฏิบัติการ (Action research) ที่มีเป้าหมายที่จะแก้ปัญาที่เกิดขึ้นและพัฒนากิจกรรมโดยการปรับปรุงเปลี่ยนแปลง ซึ่งแบ่งขั้นตอนการดำเนินงานออกเป็น 3 ขั้นตอน ดังนี้

ขั้นตอนที่ 1 การศึกษาชุดข้อมูลและฐานความรู้ของแชทบอทปัญญาประดิษฐ์ภาษาไทย เพื่อสนับสนุนการให้คำปรึกษาวิชาการสำหรับนักศึกษาระดับอุดมศึกษา โดยทำการเก็บรวบรวมข้อมูลกับอาจารย์ที่ปรึกษาสาขาวิชาวิทยาการคอมพิวเตอร์ และสาขาวิชาเทคโนโลยีสารสนเทศและการสื่อสาร คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏเพชรบูรณ์ และอาจารย์ที่ปรึกษาสาขาวิชาเทคโนโลยีสารสนเทศ คณะบริหารธุรกิจและการบัญชี มหาวิทยาลัยพิษณุโลก โดยสุ่มเลือกแบบเจาะจงจากผู้ที่มีมติชอบเป็นที่ปรึกษาของนักศึกษาชั้นปีที่ 2 และ 3 ในสาขาวิชาดังกล่าว จำนวน 6 คน และนักศึกษาชั้นปีที่ 2 และ 3 จำนวน 46 คน เพื่อรวบรวมชุดข้อมูล ในการนำมาฝึกสอนโมเดล

การสนทนาของแชทบอท โดยครอบคลุมไดอะล็อกการสนทนาในหัวข้อคำทักทาย คำให้กำลังใจ การสอบถาม และแนะนำการแก้ไขปัญหาทางการเรียน/วิชาการ จำนวน 150 รายการ จากนั้นเก็บรวบรวมข้อมูลจากเอกสารข้อกำหนดในฝ่ายวิชาการและงานทะเบียน จากผู้อำนวยการฝ่ายวิชาการและงานทะเบียนของทั้งสองมหาวิทยาลัย จำนวน 2 คน รวมกลุ่มตัวอย่างที่ใช้รวบรวมข้อมูลในขั้นตอนนี้ จำนวนทั้งหมด 54 คน

ขั้นตอนที่ 2 การพัฒนาแชทบอทปัญญาประดิษฐ์ภาษาไทยเพื่อสนับสนุนการให้คำปรึกษาวิชาการสำหรับนักศึกษาระดับอุดมศึกษา สามารถแสดงขั้นตอนการดำเนินงาน ดังภาพที่ 1 (Figure 1)

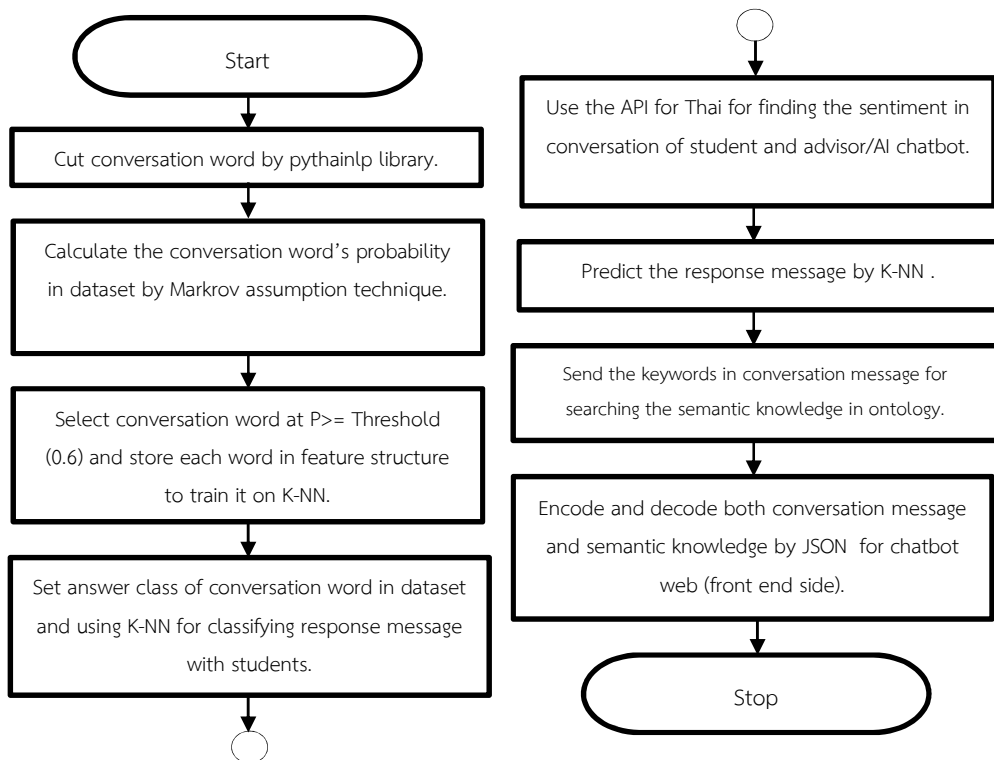


Figure 1 Flowchart of the artificial intelligence chatbot development

จากภาพที่ 1 (Figure 1) ขั้นตอนการทำงานของแชทบอท จะเริ่มตั้งแต่การตัดข้อความสนทนาเพื่อนำมาคำนวณค่าความน่าจะเป็นเพื่อรู้จำรูปแบบข้อความสนทนาด้วยเทคนิค Markov จากนั้นทำการเลือกรูปแบบที่มีค่าความน่าจะเป็นมากกว่า/เท่ากับค่า Threshold (0.6) มาสร้างชุดข้อมูลในการฝึกสอนแชทบอท เพื่อทำนายข้อความตอบกลับด้วยอัลกอริทึม K-NN และส่งคำสำคัญไปคิวรีข้อมูลที่เก็บในออนโทโลยีเพื่อตอบกลับการสนทนา โดยการติดต่อระหว่างส่วนประมวลผลและหน้าจอแชทบอทจะใช้ JSON เป็นมาตรฐานในการแลกเปลี่ยนข้อมูล

ขั้นตอนที่ 3 การประเมินประสิทธิภาพของแชทบอทปัญญาประดิษฐ์ภาษาไทยเพื่อสนับสนุนการให้คำปรึกษาวิชาการสำหรับนักศึกษาระดับอุดมศึกษา โดยการประเมินประสิทธิภาพแบ่งออกเป็น

2 ส่วน คือ 1) การประเมินประสิทธิภาพในการตอบคำถามของแชทบอท โดยการหาความถูกต้องในการทำนายข้อความตอบกลับ ซึ่งงานวิจัยนี้ใช้เทคนิค 5-fold cross validation ในการแบ่งชุดข้อมูลบทสนทนาที่ผ่านขั้นตอนของเทคนิค Markov (Threshold ≥ 0.6) และให้คลาสคำตอบเรียบร้อยแล้วจำนวน 100 รายการ รวมทั้งมีการแบ่งรายการข้อมูลฝึกสอนและข้อมูลทดสอบในอัตรา 70:30 ทั้งนี้จะได้ชุดข้อมูลในการสร้างโมเดลผ่านอัลกอริทึม K-NN รอบละ $100/5 = 20$ รายการ แบ่งเป็นข้อมูลฝึกสอนจำนวน 14 รายการ และข้อมูลทดสอบ จำนวน 6 รายการ จากนั้นทำการหาค่าเฉลี่ยของร้อยละความถูกต้องของการทดสอบทั้ง 5 รอบ รวมทั้งทำการหาประสิทธิภาพการสืบค้นความรู้เชิงความหมายในออนโทโลยีในการได้ตอบการสนทนา โดยการหาความเที่ยงตรง (Precision) ค่าความแม่นยำ (Recall) และค่าการวัดประสิทธิภาพ (F-Measure) (Manning et al., 2009; Goutte & Gaussier, 2005) และ 2) การประเมินความพึงพอใจผู้ใช้ต่อระบบ จำนวน 54 คน โดยใช้ค่าเฉลี่ย และค่าส่วนเบี่ยงเบนมาตรฐาน

ผลการวิจัย

การดำเนินการวิจัยและพัฒนาแชทบอทปัญญาประดิษฐ์ภาษาไทยเพื่อสนับสนุนการให้คำปรึกษาวิชาการสำหรับนักศึกษาระดับอุดมศึกษา สามารถสรุปผลการวิจัยที่สอดคล้องกับวัตถุประสงค์ ดังนี้

1. ผลการศึกษาชุดข้อมูลและฐานความรู้ของแชทบอทปัญญาประดิษฐ์ภาษาไทย เพื่อสนับสนุนการให้คำปรึกษาวิชาการสำหรับนักศึกษาระดับอุดมศึกษา

โดยสามารถสรุปผลการรวบรวมชุดข้อมูลตัวอย่างของการสนทนาระหว่างอาจารย์ที่ปรึกษาและนักศึกษา เพื่อนำเข้าสู่การฝึกสอนโมเดล ดังตารางที่ 1 (Table 1)

Table 1 Example data of conversation between advisors and students (Conversation owner T is advisor and S is student)

Conversation word	Cut word	Conversation owner	Sentiment
Greeting (From all 11 messages)			
สวัสดีคะนักศึกษา (Hello.)	สวัสดี (Hello), ค่ะ (yes), นักศึกษา (student)	T	Neutral
สวัสดีครับ/ค่ะ (Hello.)	สวัสดี (Hello), ครับ (yes), ค่ะ (yes)	S	Neutral
Encouragement (From all 2 messages)			
เอาใจช่วยนะคะ (Wish success.)	เอาใจช่วย (wish success), นะคะ (yes),	T	Positive (75%)
สู้ๆ ค่ะ (Keep fighting.)	สู้ (fight), ะ (yes)	T	Positive (75%)

Table 1 Example data of conversation between advisors and students (Conversation owner T is advisor and S is student) (Cont.)

Conversation word	Cut word	Conversation owner	Sentiment
Asking the problem (From all 17 messages)			
การเรียนรู้ที่ผ่านมาเป็นอย่างไรบ้างคะ (How was your study last semester?)	การเรียนรู้ (study), ที่ผ่านมา (last semester), เป็น (was), อย่างไรบ้าง (How), ค่ะ (yes)	T	Neutral
ฉิวเฉียดครับ (Barely)	ฉิวเฉียด (Barely), ครับ (yes)	S	Neutral
ก็ยากค่ะครู แต่ว่า ก็พอไหวค่ะ (It's difficult but I quite can do it.)	ก็ (it), ยาก (difficult), ค่ะ (yes), ครู (teacher), แต่ว่า (but), ก็ (it), พอ, ไหว (can), ค่ะ (yes)	S	Negative (60%)
Solving problem (From all 13 messages)			
อยากให้ครูช่วยแนะนำอะไรบ้างคะ (What do you want for suggestion?)	อยาก (want), ให้ (offer), ครู (teacher), , ช่วย (help), แนะนำ (suggestion), อะไร (what), บ้าง, ค่ะ (yes)	T	Positive (66.7%)
อยากให้ช่วยแนะนำการดำเนินชีวิตต่อเมื่อ จบการศึกษาไป ในโลกที่มีวิกฤต ค่อนข้าง แย่ในตอนนี้ค่ะ (Way of life when I graduated.)	อยาก (want), ให้ (offer), ช่วย (help), แนะนำ (suggestion), การดำเนินชีวิต (Way of life), ต่อเมื่อ (when), จบ (graduated), การศึกษา (study), ไป (go), ใน (in), โลก (word), ที่ (that), มี (have), วิกฤต (critical), ค่อนข้าง (rather), แย่ (badly), ใน (in), ตอนนี้ (now), ค่ะ (yes)	S	Negative (66.7%)

ผลการสร้างฐานความรู้เชิงความหมายด้วยออนโทโลยี ในการแนะนำการเรียนรู้/วิชาการแก่นักศึกษาที่มีความสอดคล้องกับบทสนทนา ดังภาพที่ 2 (Figure 2)

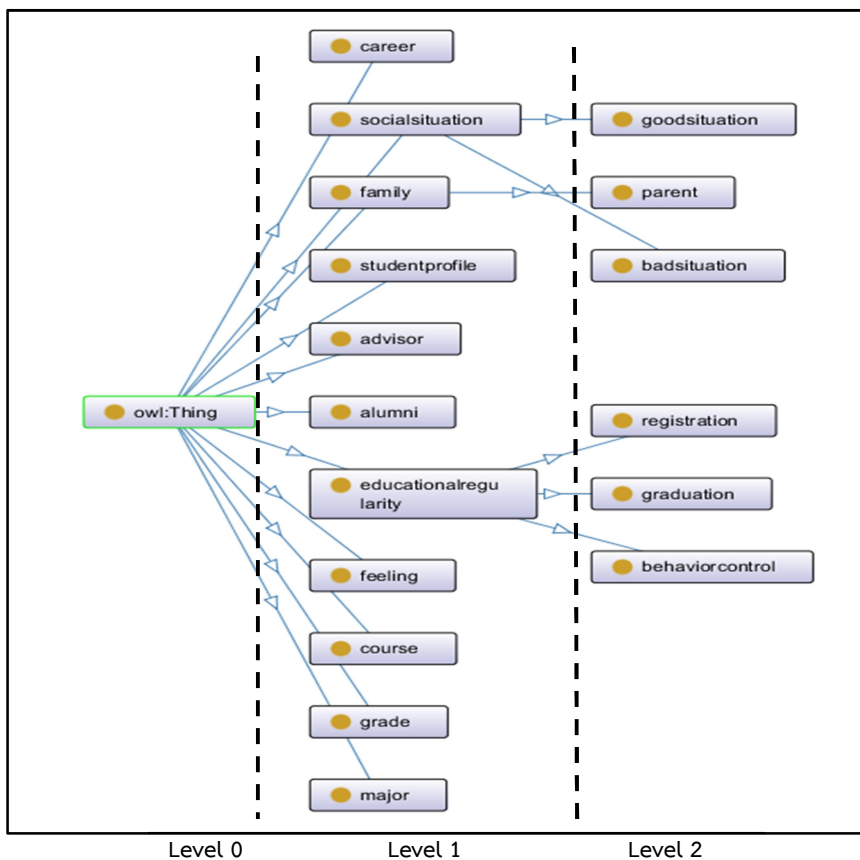


Figure 2 Ontology semantic knowledge base in study and academic recommendation

จากภาพที่ 2 (Figure 2) ฐานความรู้เชิงความหมาย ประกอบด้วยโหนดความรู้จำนวนทั้งหมด 18 โหนด แบ่งเป็น 3 ระดับชั้น ครอบคลุมการให้คำแนะนำที่สอดคล้องกับคำสำคัญที่สกัดได้จากบทสนทนาระหว่างอาจารย์ที่ปรึกษากับนักศึกษาแบ่งออกเป็น 5 กลุ่มเรื่อง ได้แก่ 1) ประวัตินักเรียนและผลการเรียน 2) เนื้อหารายวิชาเรียน 3) ศิษย์เก่าต้นแบบ 4) ระเบียบการศึกษาของมหาวิทยาลัย และ 5) สถานการณ์ที่เกิดขึ้นในสังคม

2. ผลการพัฒนาแชทบอทปัญญาประดิษฐ์ภาษาไทยเพื่อสนับสนุนการให้คำปรึกษาวิชาการสำหรับนักศึกษาระดับอุดมศึกษา สามารถสรุปผลการพัฒนาแบ่งออกเป็น 4 ส่วน ดังนี้

2.1 ผลการตัดข้อความสนทนา เพื่อนำมาหาค่าความน่าจะเป็น (P) ของแต่ละข้อความสนทนาจาก Dataset บทสนทนาของอาจารย์ที่ปรึกษาและนักศึกษา โดยใช้อัลกอริทึม Markov assumption (Fosler-Lussier, 1998) และเลือกข้อความสนทนาที่มีค่า $p \geq$ ค่า Threshold (0.6) เพื่อนำ Pattern ไปใช้ประโยชน์ในการทำนายข้อความตอบกลับ ดังนี้

Table 2 Finding the example of conversation's pattern

Conversation category	Advisor's conversation	Student's conversation for training	Student's conversation (use for finding P)	Probability (P) Threshold = 0.6	Conversation pattern
Asking the problem	การเรียนรู้ที่ผ่านมาเป็นอย่างไรบ้างคะ (How was your study last semester?) (การเรียนรู้ที่ผ่าน (study), ที่ผ่าน มา (in last semester), เป็น (was), อย่างไรบ้าง (How), ค่ะ (yes))	ก็ยากค่ะครู (It's difficult.), ก็พอไหวค่ะ (I quite can do it), ดีมากเลยครั้บครู (Very good), การเรียนที่ผ่าน มาดีค่ะอาจารย์ (The past studies were good, teacher.), ค่อนข้างยากนิดๆ ครั้บแต่ก็ไม่วายจนเกินไปครั้บ (It is a little difficult, but not too poor.),ฉิวเฉียดครั้บ (Barely) ก็ยากค่ะครู แต่ ว่า ก็พอไหวค่ะ (It's difficult but I quite can do it.)	ฉิวเฉียดครั้บ (Barely) ก็ยากค่ะครู (It's difficult.) ยาก (difficult), ก็ (it) ค่ะ (yes), ยาก (difficult) ครู (teacher), ค่ะ (yes) <stop>, ครู (teacher) AVG	ฉิวเฉียด (Barely), <start> ครั้บ (yes) 1* 0.5 0.83 0.67* 0.33 0.33 0.67* 0.47	ฉิวเฉียด..... (Barely) ก็.....ครู (It....teacher)

จากตารางที่ 2 (Table 2) สามารถนำ Pattern ของบทสนทนาเพื่อไปทำนายข้อความตอบกลับ 2 บทสนทนา คือ ข้อความ “ฉิวเฉียด....” มีค่าเฉลี่ยความน่าจะเป็นเท่ากับ 0.83 และ “ก็.....ครู” มีค่าเฉลี่ยความน่าจะเป็นเท่ากับ 0.47 หากพิจารณาจะพบว่า สิ่งที่เป็นช่องว่าง (.....) จะสามารถเป็นข้อความใด ๆ มาเติม จะขึ้นอยู่กับบทสนทนาที่นำเข้ามาฝึกสอนเพิ่มในรอบต่อไป

2.2 ผลการกำหนด Class คำตอบของข้อความได้ตอบการสนทนาใน Dataset เพื่อฝึกสอน และใช้อัลกอริทึม K-NN ในการฝึกการจำแนกข้อความได้ตอบการสนทนา

ทั้งนี้ผู้วิจัยได้นำ API ของ API for Thai หาค่า Sentiment ของข้อความโต้ตอบของนักศึกษาและแชทบอท เพื่อรับรู้ความรู้สึกของผู้สนทนาในส่วนข้อความว่า “ฉิวเฉียด” “ก็ยากครู” “ก็ไม่ยากครู” เพื่อนำ Sentiment มาสร้างเป็นคุณลักษณะ (Feature) ร่วมกับ Feature ที่สร้างขึ้นจากการตัดข้อความ จากนั้นให้กลุ่มตัวอย่างอาจารย์ที่ปรึกษาให้คลาสคำตอบ ผูกสอนและทำนายด้วยอัลกอริทึม K-NN ดัง Table 3 ซึ่ง K-NN เป็นอัลกอริทึมการเรียนรู้ที่ไม่ซับซ้อน และมีประสิทธิภาพสำหรับการจำแนกข้อมูลและการรับรู้วัตถุ โดย K-NN จะคำนวณระยะทางโดยปริภูมิแบบยุคลิด (Sun et al., 2009) ดังสมการที่ 1

$$D(p, q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2} \quad (1)$$

Table 3 Dataset for training conversation's pattern by K-NN

Feature						Sentiment	Label
ฉิว (Barely)	เฉียด (Barely)	ก็ (it)	ไม่ (not)	ยาก (difficult)	ครู (teacher)		
1	1					Neutral	สู้ๆ ค่ะ (fighting)
		1	1	1		Positive	เยี่ยมมากค่ะ (very good)
		1		1	1	Negative	สู้ๆ ค่ะ (fighting)

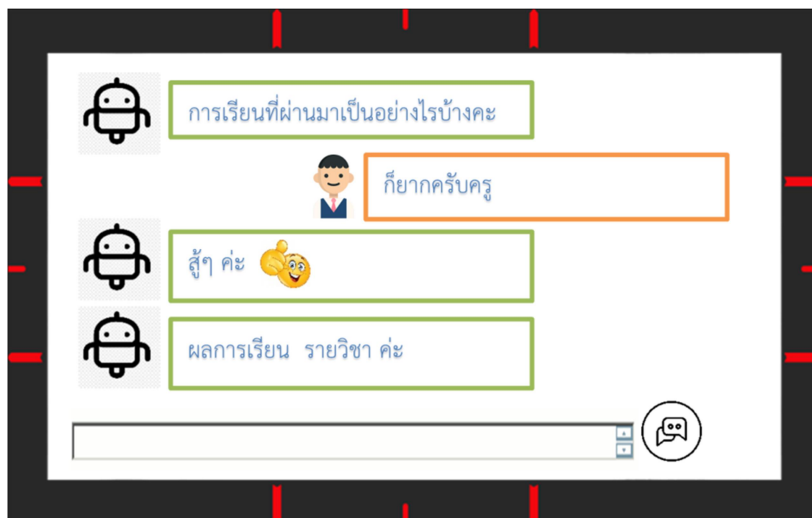
จากตารางที่ 3 (Table 3) ชุดข้อมูล (Dataset) จะส่งเข้าสู่ขั้นตอนการฝึกสอนด้วย K-NN เพื่อทำนายข้อความตอบกลับ ยกตัวอย่างเช่น เมื่อครูถามว่า “การเรียนที่ผ่านมาเป็นอย่างไรบ้างคะ” นักศึกษาตอบว่า “ก็ยากครู.....” แชทบอทจะทำการทำนายด้วย K-NN และตอบกลับมามีคำว่า “สู้ๆ ค่ะ”

2.3 ผลการส่งคำสำคัญใน Pattern บทสนทนา ไปคิวรีเพื่อค้นหาความรู้เชิงความหมายในออนโทโลยี โดยในขั้นตอนนี้ผู้วิจัยใช้ภาษา Sparql ในการคิวรีข้อมูลเชิงความหมาย เช่นเดียวกับ Tupmongkol & Pilabutr (2016) ใช้ภาษา Sparql ในการเข้าถึงโครงสร้าง RDF ของออนโทโลยี ในการสืบค้นข้อมูลเชิงความหมายของการท่องเที่ยวไทย ซึ่งได้แสดงคำสั่งและผลการสืบค้นออนโทโลยี ดังนี้

Table 4 Sparql command and result of ontology searching for response message from the AI Chat bot

Advisor/AI chat bot's conversation	Student's conversation	Sparql command	Result of ontology searching
การเรียนรู้, ที่ผ่านมา, เป็น, ก็, ยาก, ครู อย่างไรบ้าง, ค่ะ (How was your study last semester?)	(It's difficult, teacher.)	SELECT ?keyword ?knowledge1 ?knowledge2 {?knowledge1 rdfs: sub class of ?keyword. ?knowledge2 rdfs: domain ?keyword. FILTER(?keyword=node:การเรียนรู้ ?keyword=node:เรียน)}	-Node=Grade, Course - The AI chatbot will set question about grade and course in which the student encounter problems.
สู้ๆ, ค่ะ (Keep fighting)		SELECT {?keyword ?emoji ?keyword node:send_emoji ?emoji FILTER(?keyword=node:สู้ๆ)}	Node = emoji - The AI chat bot will send encouraging emoji.

2.4 จากนั้นระบบจะส่งข้อความตอบกลับผ่านการเข้ารหัสแบบ JSON มาจากส่วนการประมวลผล (Backend) ด้วย Flask framework ซึ่งเป็นไลบรารีใน Python ที่ใช้ในการเชื่อมต่อระหว่างส่วนติดต่อผู้ใช้กับส่วนการจัดการข้อมูล (Mitpitak & Nanthajirapong, 2019) ซึ่งในขั้นตอนนี้ Web UI ของแชทบอท จะทำการถอดรหัสข้อมูลบทสนทนา และความรู้เชิงความหมายที่ได้รับมาแสดงผล/โต้ตอบผู้ใช้ดังภาพที่ 3 (Figure 3)

**Figure 3** Web UI of AI chatbot

3. ผลการประเมินประสิทธิภาพในการตอบคำถาม และประเมินความพึงพอใจผู้ใช้ต่อแชทบอทปัญญาประดิษฐ์

ผู้วิจัยทำการทดสอบประสิทธิภาพความถูกต้องในการทำนายคำตอบการสนทนาของแชทบอทที่พัฒนาด้วย K-NN โดยใช้เทคนิค 5-folds cross validation กับชุดข้อมูล จำนวน 100 รายการ และความถูกต้องและครบถ้วนในการสืบค้นความรู้จากออนโทโลยี โดยกำหนดบทสนทนาของนักศึกษาจำนวน 5 ข้อความ แล้วพิมพ์ให้แชทบอทตอบบทสนทนา จากนั้นทำการประเมินประสิทธิภาพด้วยการพิจารณาคำตอบของแชทบอทที่แสดงออกมาว่ามีจำนวนค่าที่ตรง/ไม่ตรงกับความต้องการหรืออยู่นอกขอบเขตการสนทนา ผลที่ได้ดังตารางที่ 5-6 (Table 5-6)

Table 5 AI chatbot performance testing on K-NN

Folded process	K-NN			
	K=3	K=5	K=7	K=9
1	100.0%	100.0%	100.0%	98.0%
2	100.0%	100.0%	98.0%	98.0%
3	100.0%	100.0%	98.0%	97.5%
4	100.0%	100.0%	95.0%	97.5%
5	100.0%	100.0%	95.0%	95.0%
Accuracy total	100.0%	100.0%	97.2%	97.2%

จากตารางที่ 5 (Table 5) แชทบอทสามารถทำนายคำตอบการสนทนาด้วยอัลกอริทึม K-NN ได้แม่นยำมากที่สุด เท่ากับ 100% ที่ K=3 และ 5

Table 6 Performance Assessment result in response message of AI chatbot

List of test	Result of semantic searching		
	number of	Number of	Number of
	words in	words in	words which is
	interested	interested	responded but it
	domain which	domain which is	is out of
	is responded	not responded	domain.
	(A)	(B)	(C)
1. สวัสดีค่ะครู (Hello teacher.) (สวัสดี, ค่ะ)	2	0	0
2. ปรึกษาเรื่องการเรียนครับ (Consult the students 's studying) (ผลการเรียน, รายวิชา, ค่ะ)	3	0	0
3. ข้อสอบยากมากครับครู (The examination is very difficult.) (สู้ๆ, ค่ะ, อืมม)	3	1 ** รายวิชา (Course)	0

Table 6 Performance Assessment result in response message of AI chatbot (Cont.)

List of test	Result of semantic searching		
	number of words in interested domain which is responded (A)	Number of words in interested domain which is not responded (B)	Number of words which is responded but it is out of domain. (C)
4. ช่วงนี้หนูเครียดจังเลย (Now I'm very nervous.) (สู้ๆ, ค่ะ, อีโมจิ, การเรียน)	4	1 ** ทางบ้าน (Problem at home)	0
5. ขอบคุณครูค่ะ/ครับ (Thank you teacher.) (ค่ะ, บายๆ, อีโมจิ)	3	0	0
Sum	15	2	0
Mean	3	0.4	0
Precision = $A/(A+C) \times 100$		88.2%	
Recall = $A/(A+B) \times 100$		100%	
F-measure = $2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$		93.6%	

จากตารางที่ 6 (Table 6) สรุปได้ว่า แชทบอทตอบคำถามตรงกับขอบเขตความสนใจ โดยมีค่า F-measure เท่ากับ 93.6%

ผลการประเมินความพึงพอใจแชทบอทปัญญาประดิษฐ์ภาษาไทยเพื่อสนับสนุนการให้คำปรึกษาวิชาการสำหรับนักศึกษาระดับอุดมศึกษา ในภาพรวม อยู่ในระดับมาก ($\bar{x}=4.06$, S.D.=0.68) และสามารถเรียงประเด็นที่ผู้ให้พึงพอใจสูงสุด 3 อันดับแรก ได้แก่ 1) การออกแบบแชทบอทใช้งานง่าย อยู่ในระดับมากที่สุด ($\bar{x}=4.52$, S.D.=0.76) 2) ฐานความรู้ของแชทบอทให้ข้อมูลที่เป็นประโยชน์ อยู่ในระดับมาก ($\bar{x}=4.21$, S.D.=0.55) และ 3) ช่วยผ่อนคลายความเครียดได้ อยู่ในระดับมาก ($\bar{x}=4.18$, S.D.=0.67) และประเด็นที่ผู้ให้พึงพอใจต่ำสุด คือ ความเร็วในการตอบคำถามของแชทบอท อยู่ในระดับมาก ($\bar{x}=3.67$, S.D.=0.72)

อภิปรายผล

ชุดข้อมูลที่ใช้ฝึกสอนแชทบอทแบ่งออกเป็น 3 ส่วน ได้แก่ คำทักทาย คำให้กำลังใจ และคำแนะนำในการแก้ปัญหาด้านการเรียน เป็นต้น และการสร้างฐานความรู้เชิงความหมายของแชทบอทครอบคลุมการให้คำแนะนำที่สอดคล้องกับคำสำคัญที่สกัดได้จากบทสนทนาระหว่างอาจารย์ที่ปรึกษากับนักศึกษา แบ่งออกเป็น 5 กลุ่มเรื่อง ได้แก่ 1) ประวัตินักเรียนและผลการเรียน 2) เนื้อหารายวิชาเรียน 3) ศิษย์เก่าต้นแบบ 4) ระบบการศึกษาของมหาวิทยาลัย และ 5) สถานการณ์ที่เกิดขึ้นในสังคม ซึ่งเป็นตามหลักของการให้คำปรึกษานักศึกษาระดับปริญญาตรี ที่สรุปสาระสำคัญไว้ 4 ด้าน ได้แก่ 1) ด้านวิชาการ 2)

ด้านการปรับตัว 3) ด้านการพัฒนาบุคลิกภาพ และ 4) ด้านการทำกิจกรรมของนักศึกษา ซึ่งการใช้ระบบจัดการฐานความรู้เชิงความหมายแทนระบบจัดการฐานข้อมูลเชิงสัมพันธ์เพียงอย่างเดียวจะช่วยให้เช็ตบอทมีประสิทธิภาพในการค้นคืนความรู้ที่มีความเกี่ยวข้องกันในเชิงความหมายในการตอบการสนทนาเชิงการให้ความรู้ไปพร้อมกัน โดยออนโทโลยีจะช่วยในการจัดการโครงสร้างของข้อมูลที่ไม่มรูปแบบโครงสร้างได้ดี และช่วยเชื่อมโยงฐานข้อมูลที่กระจายอยู่ในหลายแหล่งรวมศูนย์เพื่อบริหารจัดการในโดเมนออนโทโลยีไว้แหล่งเดียว ช่วยให้เช็ตบอทมีความรอบรู้มากยิ่งขึ้น ทั้งนี้งานวิจัยของ Altinok (2018) ได้ประยุกต์ใช้ออนโทโลยีในระบบการจัดการบทสนทนาสำหรับการธนาคารและการเงิน ซึ่งจะช่วยให้การสนทนาผ่านโดเมนออนโทโลยี มีศักยภาพที่ดีในการแยกแยะบริบทของชุดผลิตภัณฑ์และความจำเพาะของคำถาม ชื่อเอนทิตี คำกริยา และงานวิจัยของ Bhalotia & Bisen (2018) พบว่า แนวคิดเชิงความหมาย จะช่วยจัดการความคล้ายคลึงกันในประโยคสนทนาได้ ในกรณีที่คำถาม Q1 และ Q2 หมายถึงสิ่งเดียวกัน โดยจะเลือกข้อมูลที่มีคะแนนความสอดคล้องในเชิงความหมายกับข้อความนำเข้าสู่สูงสุด หรือ $\geq$ ค่า Threshold เป็นข้อความตอบกลับ และการใช้เทคนิคดังกล่าวเพื่อหารูปแบบของบทสนทนา ทั้งนี้ในงานวิจัยนี้ได้มีการวัดประสิทธิภาพของเช็ตบอทที่สร้างขึ้น 3 ค่า คือ ค่าความเที่ยงตรง (Precision = 88.2%), ค่าความแม่นยำ (Recall = 100%) และค่าการวัดประสิทธิภาพ (F-Measure = 93.6%) ซึ่งทั้ง 3 ค่า สูง ดังนั้นจึงบ่งชี้ได้ว่าเช็ตบอทที่สร้างขึ้นนี้มีประสิทธิภาพสูงสอดคล้องกับงานวิจัยของ Netsuwan (2015) ศึกษาเรื่อง การจำแนกข่าวภาษาอังกฤษด้านอาชญากรรมออนไลน์ด้วยเทคนิคการทำเหมืองข้อความ โดยได้วัดประสิทธิภาพของแบบจำลองด้วย 3 ค่า คือ Precision, Recall และ F-Measure ที่มีค่าสูงกว่าร้อยละ 80 เช่นกัน และในการวิจัยนี้มีการทำนายข้อความตอบกลับด้วยโมเดล K-NN เกิดประสิทธิภาพที่ดี ซึ่ง K-NN เป็นอัลกอริทึมที่นิยมใช้เพราะเข้าใจง่าย การเรียนรู้ที่ไม่ซับซ้อน และมีประสิทธิภาพสำหรับการจำแนกข้อมูลและการรับรู้วัตถุ (Sun et al., 2009) K-NN ถูกนำมาใช้ในการทำนายร่วมกับกลไกของภาษาธรรมชาติ Hussain et al. (2020) ได้พัฒนาเช็ตบอททำนายโรคโดยใช้ภาษาธรรมชาติ ร่วมกับอัลกอริทึมการเรียนรู้ของเครื่อง สำหรับการจำแนกสองประเภท ได้แก่ ต้นไม้การตัดสินใจและ K-NN ซึ่งประสิทธิภาพความถูกต้องในการทำนายด้วย K-NN บนเช็ตบอท เท่ากับ 93.6%

สรุปผลการวิจัย

เช็ตบอทปัญญาประดิษฐ์ภาษาไทยเพื่อสนับสนุนการให้คำปรึกษาวิชาการสำหรับนักศึกษาระดับอุดมศึกษา ได้ทำการการรวบรวมชุดข้อมูลตัวอย่างของการสนทนาระหว่างอาจารย์ที่ปรึกษา และนักศึกษา เพื่อนำเข้าสู่การฝึกสอนโมเดล เพื่อสร้างฐานความรู้เชิงความหมายด้วยออนโทโลยี ในการแนะนำการเรียน/วิชาการแก่นักศึกษาที่มีความสอดคล้องกับบทสนทนา หลังจากนั้นพัฒนาเช็ตบอทปัญญาประดิษฐ์ภาษาไทยเพื่อสนับสนุนการให้คำปรึกษาวิชาการสำหรับนักศึกษาระดับอุดมศึกษาสามารถสรุปผลการพัฒนาแบ่งออกเป็น 4 ส่วน ได้แก่ 1) ส่วนการตัดข้อความสนทนา เพื่อนำมาหา

ค่าความน่าจะเป็น (P) ของแต่ละข้อความสนทนา จาก Dataset บทสนทนาของอาจารย์ที่ปรึกษาและนักศึกษา โดยใช้อัลกอริทึม Markov assumption และเลือกข้อความสนทนาที่มีค่า $p \geq$ ค่า Threshold (0.6) เพื่อนำ Pattern ไปใช้ประโยชน์ในการทำนายข้อความตอบกลับ 2) ส่วนการกำหนด Class คำตอบของข้อความโต้ตอบการสนทนาใน Dataset เพื่อฝึกสอน และใช้อัลกอริทึม K-NN ในการฝึกการจำแนกข้อความโต้ตอบการสนทนา 3) ส่วนการส่งคำสำคัญใน Pattern บทสนทนา ไปคิวรีเพื่อค้นหาความรู้เชิงความหมายในออนโทโลยี โดยในขั้นตอนนี้อยู่ใช้ภาษา Sparql ในการคิวรีข้อมูลเชิงความหมาย และ 4) ส่วนการประมวลผลด้วย Flask framework ซึ่งเป็นไลบรารีใน Python ที่ใช้ในการเชื่อมต่อระหว่างส่วนติดต่อผู้ใช้กับส่วนการจัดการข้อมูล โดยในขั้นตอนนี้อยู่ Web UI ของแชทบอทจะทำการถอดรหัสข้อมูลบทสนทนา และความรู้เชิงความหมายที่ได้รับมาแสดงผล/โต้ตอบผู้ใช้ ซึ่งจากการทดสอบประสิทธิภาพโดยกำหนดบทสนทนาของนักศึกษา จำนวน 5 ข้อความ แล้วพิมพ์ให้แชทบอทตอบบทสนทนา จากนั้นทำการประเมินประสิทธิภาพด้วยพิจารณาคำตอบของแชทบอทที่แสดงออกมาว่ามีจำนวนคำที่ตรง/ไม่ตรงกับความต้องการหรืออยู่นอกขอบเขตการสนทนา ซึ่งให้ผลลัพธ์ว่ามีความเที่ยงตรง (Precision = 88.2%) ความแม่นยำ (Recall = 100%) และการวัดประสิทธิภาพ (F-Measure = 93.6%)

กิตติกรรมประกาศ

ขอขอบพระคุณมหาวิทยาลัยราชภัฏเพชรบูรณ์ที่ได้ให้โอกาสในการทำงานวิจัยนี้ รวมทั้งกลุ่มตัวอย่างอาจารย์ที่ปรึกษา นักศึกษา สาขาวิชาวิทยาการคอมพิวเตอร์และสาขาวิชาเทคโนโลยีสารสนเทศ และการสื่อสาร คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยราชภัฏเพชรบูรณ์ อาจารย์ที่ปรึกษา นักศึกษา สาขาวิชาเทคโนโลยีสารสนเทศ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยพิษณุโลก และผู้อำนวยการฝ่ายวิชาการและงานทะเบียนของทั้งสองมหาวิทยาลัย ที่ได้ให้ข้อมูลที่เป็นประโยชน์ในงานวิจัยนี้

เอกสารอ้างอิง

- Altinok D. An *Ontology-Based Dialogue Management System for Banking and Finance Dialogue Systems*. Proceedings of 1st Financial Narrative Processing Workshop at The LREC; 2018 May 7-12; Miyazaki, Japan. 2018: 1-9.
- Bhalotia S, Bisen S. Implementation of a Chat Bot System using AI and NLP. *International Journal of Innovative Research in Computer Science & Technology*, 2018; 6(3): 26-30.
- Fosler-Lussier E. *Markov Models and Hidden Markov Models: A Brief Tutorial*. International Computer Science Institute, Berkeley, California, 1998.
- Goutte C, Gaussier E. A *Probabilistic Interpretation of Precision, Recall and F-Score, With Implication For Evaluation*. Proceeding of 27th European Conference on IR Research (ECIR 2005); 2005 March 21-23; Santiago de Compostela, Spain; 2005: 345-359.

- Hussain H, Aswani K, Gupta M. et al. Implementation of Disease Prediction Chatbot and Report Analyzer using the Concepts of NLP, Machine Learning and OCR. *International Research Journal of Engineering and Technology*, 2020; 7(4): 1814-1819.
- Kerschberg L, Chowdhury M, Damiano A. et al. *Knowledge Sifter: Ontology-Driven Search over Heterogeneous Databases*. Proceedings of 16th International Conference on Scientific and Statistical Database Management; 2004 May 7-12; Virginia, USA. 2004: 431-432.
- Manning C, Raghavan P, Schütze H. *Measuring Information Retrieval*. 2009. Available at: <https://nlp.stanford.edu/IR-book/pdf/08eval.pdf>. Accessed December 20, 2019.
- Ministry of Information and Communication. *Technology policy framework Information and communication period 2011-2020*. Bangkok: Ministry of Information and Communication; 2011.
- Mitpitak A, Nanthajirapong N. *Voice Command System for Software Testing and Quality Improvement*. Proceedings of the 7th Academic Science and Technology Conference 2019; 2019 June 7; Bangkok, Thailand. 2019: 1675-1683.
- Netsuwan T. *Online Classified English News on Crime with Text Mining Techniques*. Proceedings of National academic conference 11th Computer and Information Technology; 2015 July; Thailand. 2015: 61-66.
- Nookhong J, Kaewrattanapat N. The automatic consulting system in higher education institution. *International Journal of Information and Education Technology*. 2017; 7(10): 712-715.
- Setthachotsombut N. Using the advisor system and study plan management system for services efficiency. *Graduated Studies Journal*. 2018; 15(69): 9-17.
- Sun B, Du J, Gao T. *Study on the Improvement of K-Nearest-Neighbor Algorithm*. Proceedings of International Conference on Artificial Intelligence and Computational Intelligence; Shanghai, China; 2009: 390-393.
- The Office of Academic Promotion. *A Guide for Bachelor's Degree Advisor, Revised Version 2017*. The Office of Academic Promotion, Phetchabun Rajabhat University. 2017.
- Tupmongkol S, Pilabutr S. Thailand's tourism semantic web with ontology (RDF and SPARQL). *APHEIT Journal Science & Technology*. 2016; 5(2): 5-11.
- Villegas-CH W, Arias-Navarrete A, Palacios-Pacheco X. Proposal of an architecture for the integration of a chatbot with artificial intelligence in a smart campus for the improvement of Learning. *Sustainability*. 2020; 12(4): 1-20.