

ระบบผู้ช่วยสั่งงานด้วยเสียงสำหรับถอดปริมาณงานก่อสร้างอาคาร

จิตาภรณ์ พอบุตรดี*

คณะวิทยาศาสตร์และวิศวกรรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์

วิทยาเขตเฉลิมพระเกียรติ จังหวัดสกลนคร

Received: 9 July 2021

Revised: 12 November 2022

Accepted: 14 November 2022

บทคัดย่อ

การถอดปริมาณงานก่อสร้างเป็นขั้นตอนที่สำคัญในการหาต้นทุนงานก่อสร้างซึ่งเป็นงานที่สำคัญในงานด้านวิศวกรรมโยธา ถึงแม้ว่าการถอดปริมาณงานก่อสร้างเป็นกระบวนการที่มีลักษณะการกระทำซ้ำและในแต่ละการกระทำซ้ำนั้นมีขั้นตอนที่ไม่ซับซ้อน แต่การถอดปริมาณงานก่อสร้างก็เป็นขั้นตอนที่ใช้เวลาค่อนข้างมาก การนำเทคโนโลยีปัญญาประดิษฐ์มาประยุกต์ใช้จะช่วยพัฒนาระบบงานให้มีประสิทธิภาพขึ้นได้ งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาผู้ช่วยสั่งงานด้วยเสียงสำหรับถอดปริมาณงานก่อสร้างอาคารด้วยภาษาไพทอน โดยใช้ไลบรารีหลัก 3 ไลบรารี คือ (1) Speech Recognition 3.8.1 (2) gTTS 2.2.2 และ (3) Playsound 1.2.2 ทำการพัฒนาการป้อนข้อมูลโดยใช้เสียงใน 2 รูปแบบ ได้แก่ (1) การป้อนข้อมูลลักษณะถามข้อมูลทีละข้อ และ (2) การป้อนข้อมูลโดยการให้ข้อมูลทั้งหมดขององค์อาคารครั้งเดียว ผลการวิจัยพบว่า โปรแกรมสามารถรับข้อมูลองค์อาคารจากเสียงพูดได้ดี โดยการป้อนข้อมูลลักษณะถามข้อมูลทีละข้อให้ความถูกต้องสูงกว่าแต่ใช้เวลาในการทำงานนานกว่า โดยมีค่าอัตราค่าผิดพลาดของค่าอยู่ระหว่าง 12.36 และ 22.45 ในขณะที่รูปแบบโดยการให้ข้อมูลทั้งหมดขององค์อาคารครั้งเดียวมีค่าอัตราค่าผิดพลาดของค่าอยู่ระหว่าง 20.54 และ 28.76 โปรแกรมที่พัฒนาขึ้นจะทำให้การทำงานสะดวกสบายและมีประสิทธิภาพมากยิ่งขึ้น

คำสำคัญ: การถอดปริมาณงาน การก่อสร้าง ปัญญาประดิษฐ์

* ผู้ประสานงานหลัก; อีเมล: titaporn.pob@ku.th

Voice Assistant System for Construction Quantity Take-off

Titaporn Pobutdee*

Faculty of Science and Engineering, Kasetsart University
Chalermphrakiat Sakon Nakhon Province Campus

Received: 9 July 2021

Revised: 12 November 2022

Accepted: 14 November 2022

Abstract

Quantity take-off is a critical step in determining construction costs, which is an important task in civil engineering. Although the quantity take-off is a repetitive procedure and each iteration has a simple task, quantity take-off is a time-consuming process. The application of artificial intelligence technology will help develop more efficient work systems. The objective of this research is to develop a voice assistant for construction quantity take-off with Python. The three main libraries were used: (1) Speech Recognition 3.8.1, (2) gTTS 2.2.2, and (3) playsound. 1.2.2. The voice input was developed in two types: (1) Information of the building member input one-by-one (2) Data entry by giving all information of the building member one time. The results showed that the program was able to receive the building member information from the sound of speech. Information of the building member input one-by-one gives higher accuracy but takes longer time. The input one-by-one method provides word error rate between 12.36 and 22.45, while all information of the building member one time method provides word error rate between 20.54 and 28.76. The developed program will make the work more comfortable and efficient.

Keywords: Construction Estimation, Construction, Artificial Intelligent

* Corresponding Author; E-mail: titaporn.pob@ku.th

ความเป็นมาและความสำคัญของปัญหา

การถอดปริมาณงานก่อสร้างเป็นขั้นตอนที่สำคัญในการหาต้นทุนงานก่อสร้าง ซึ่งเป็นงานที่เกี่ยวข้องทั้งในการดำเนินการกำหนดราคากลางงานก่อสร้างของผู้ว่าจ้างและการกำหนดราคาเพื่อยื่นประมูลงานก่อสร้างของผู้รับเหมา ทั้งนี้การถอดปริมาณงานก่อสร้างอาคารดำเนินการด้วยวิธีการหลัก 2 วิธี ได้แก่ (1) การถอดปริมาณงานด้วยมือซึ่งดำเนินการโดยการอ่านแบบ 2 มิติ แล้วทำการคำนวณโดยใช้เครื่องคิดเลข หรืออาจจะใช้โปรแกรมช่วยในการคำนวณ รวมทั้งในบางกรณีอาจใช้อุปกรณ์ดิจิทัลเทอร์หรือการสแกนแบบก่อสร้างเข้าสู่โปรแกรมเขียนแบบ (CAD) เพื่อช่วยในการหาขนาดองค์อาคาร และ (2) การถอดปริมาณงานจากแบบจำลองสารสนเทศอาคาร (BIM) ซึ่งเป็นแบบจำลอง 3 มิติ เมื่อเปรียบเทียบวิธีการทั้ง 2 วิธีการดังกล่าวเฉพาะในขั้นตอนการถอดปริมาณงานก่อสร้างจะพบว่าการถอดปริมาณงานจากแบบจำลองสารสนเทศอาคาร (BIM) เป็นวิธีการที่มีความสะดวกรวดเร็วกว่า แต่หากพิจารณาขั้นตอนโดยรวมจะพบว่า หากเริ่มต้นจากแบบก่อสร้าง 2 มิติ การถอดปริมาณงานจากแบบจำลองสารสนเทศอาคาร (BIM) มีขั้นตอนที่ยุ่งยากและซับซ้อนมากกว่า รวมทั้งต้องใช้เครื่องมือที่มีราคาสูงกว่าและบุคลากรที่มีประสบการณ์มากกว่า

อย่างไรก็ตาม การถอดปริมาณงานด้วยมือต้องใช้เวลาในการทำงาน โดยเวลาที่ใช้ในการถอดแบบจะขึ้นอยู่กับขนาดและรายละเอียดของอาคาร ดังนั้น เพื่อช่วยให้การทำงานสะดวกสบายมากขึ้นจึงได้มีการนำโปรแกรมคอมพิวเตอร์มาช่วยในการถอดปริมาณงานแทนการใช้เครื่องคิดเลขคำนวณ อย่างไรก็ตามหากพิจารณาจะพบว่าการใช้งานโปรแกรมยังต้องมีขั้นตอนการป้อนข้อมูลรายละเอียดขององค์อาคารซึ่งยังคงต้องพึ่งพาการนำเข้าข้อมูลผ่านทางคีย์บอร์ด โดยผู้ถอดปริมาณงานต้องอ่านแบบและต้องคีย์ข้อมูลที่แตกต่างกันเพียงเล็กน้อยขององค์อาคารแต่ละชั้นในเวลาเดียวกัน ซึ่งเป็นขั้นตอนที่ผู้ใช้งานยังไม่ได้รับความสะดวกนัก

การพัฒนาระบบผู้ช่วยสั่งงานด้วยเสียง (Voice Assistant) ซึ่งเป็นเทคโนโลยีที่ใช้ปัญญาประดิษฐ์ (Artificial Intelligence) สำหรับถอดปริมาณงานก่อสร้างอาคารเป็นช่องทางหนึ่งในการพัฒนาวิธีการถอดปริมาณงานก่อสร้างด้วยมือ โดยการใช้ประโยชน์จากปัญญาประดิษฐ์ที่ทำให้คอมพิวเตอร์สามารถรับเข้าข้อมูลจากเสียงของมนุษย์ สำหรับตัวอย่างที่เห็นได้ชัดเจนถึงความสะดวกสบายของการสั่งงานด้วยเสียงคือการพิมพ์งานด้วยเสียงผ่านโปรแกรม Google Docs ซึ่งทำให้การพิมพ์เอกสารทำได้ง่ายกว่าแต่เดิมมาก ผู้ใช้งานสามารถแต่งผลงานได้ในทุกอริยบทโดยไม่จำเป็นต้องนั่งพิมพ์ผ่านคีย์บอร์ดที่หน้าเครื่องคอมพิวเตอร์ที่ตั้งอยู่บนโต๊ะเหมือนเช่นในอดีต

สำหรับการพัฒนาระบบผู้ช่วยสั่งงานด้วยเสียงสำหรับงานวิจัยนี้จะทำให้ผู้ที่ทำการถอดปริมาณงานสามารถอ่านแบบและป้อนข้อมูลองค์อาคารด้วยเสียงแทนการป้อนข้อมูลผ่านทางคีย์บอร์ด ทำให้การทำงานซึ่งแต่เดิมต้องจดจ้องแบบก่อสร้าง จอคอมพิวเตอร์และคีย์บอร์ดกลายเป็นการทำงานที่ใช้เสียงสั่งการเป็นหลัก โดยยังคงรักษาข้อดีของวิธีการทำงานถอดปริมาณงานก่อสร้างด้วยมือที่สามารถดำเนินการได้ง่ายและสามารถใช้บุคลากรที่มีประสบการณ์ไม่มาก รวมทั้งคอมพิวเตอร์สำหรับรองรับการทำงานของโปรแกรมเป็นเพียงคอมพิวเตอร์รุ่นปกติที่ใช้งานตามสำนักงานทั่วไป ดังนั้น การพัฒนาระบบผู้ช่วยสั่งงานด้วยเสียงสำหรับถอดปริมาณงานก่อสร้างอาคารจะเป็นการปรับปรุงการทำงานการถอดปริมาณงานก่อสร้างอาคารด้วยมือให้มีประสิทธิภาพมากยิ่งขึ้น

วัตถุประสงค์ของการวิจัย

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาระบบผู้ช่วยสั่งงานด้วยเสียงสำหรับถอดปริมาณงานก่อสร้างอาคารด้วยภาษาไพทอน โดยมีขอบเขตการถอดปริมาณงานขององค์อาคารจำนวน 7 ประเภท ได้แก่ (1) ฐานรากสี่เหลี่ยมจัตุรัส (2) ฐานรากสี่เหลี่ยมผืนผ้า (3) เสาหน้าตัดสี่เหลี่ยมจัตุรัส (4) เสาหน้าตัดสี่เหลี่ยมผืนผ้า (5) คาน (6) พื้นทางเดียว และ (7) พื้นสองทาง และจำกัดเฉพาะการนำเข้าข้อมูลองค์อาคารและการคำนวณปริมาณงานคอนกรีต ไม้แบบ และ เหล็กเสริม

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

งานวิจัยในอดีตพบว่า การถอดปริมาณงานด้วยการใช้งานแบบจำลองสารสนเทศอาคาร (BIM) ให้ผลของปริมาณงานในหมวดงานโครงสร้างคอนกรีตเสริมเหล็กและหมวดงานสถาปัตยกรรมมีความแตกต่างกันไม่เกินร้อยละ 3 (Sukhamwang et al., 2019) และเมื่อเปรียบเทียบระหว่างการใช้ Microsoft Excel ในการคำนวณถอดปริมาณวัสดุแบบทั่วไป และใช้โปรแกรม Autodesk Revit (Free Software for Students & Educators Revit Autodesk) ทำการวางแผนการตัดขนาดวัสดุ (Bar-Cut List) พบว่าระบบสารสนเทศของอาคารสามารถทำได้ดีกว่าระบบเดิมที่สามารถถอดปริมาณเหล็กเสริมได้สูงสุดร้อยละ 17 (Tibodee et al., 2021) อย่างไรก็ตาม การใช้ระบบ BIM ยังต้องใช้ทรัพยากรในการดำเนินการค่อนข้างมากทั้งซอฟต์แวร์ที่มีราคาแพง รวมทั้งจากการวิจัยของ Chunhamaneewat and Tuntisewee (2020) ได้ให้ความเห็นว่าในประเทศไทยมีฝ่ายและหน่วยงานที่สามารถใช้ BIM ได้อย่างเต็มประสิทธิภาพได้น้อยอันมีผลมาจากความพร้อมของบริษัทด้านบุคลากรที่จะสามารถใช้งานแบบจำลองสารสนเทศได้จากงานวิจัยของ Suriyamongkol and Tuntisewee (2020) พบว่า การขาดแคลนอุปกรณ์ software hardware และระบบอินเทอร์เน็ตที่เหมาะสมเป็นอุปสรรคในการนำระบบ BIM มาใช้ ดังนั้น การถอดปริมาณงานก่อสร้างจากแบบ 2 มิติที่ดำเนินการด้วยคนยังคงเป็นวิธีที่มีการใช้งานอย่างแพร่หลายในปัจจุบันและในอนาคตอีกระยะหนึ่ง

การพัฒนาการถอดปริมาณงานด้วยมือโดยการพัฒนาระบบการนำเข้าข้อมูลแบบก่อสร้างเข้าสู่คอมพิวเตอร์ด้วยเสียงจะช่วยให้การทำงานมีความสะดวกสบายมากขึ้น ทั้งนี้ระบบสั่งงานด้วยเสียงมีการพัฒนาขึ้นมาหลายระบบโดยผู้พัฒนาหลายรายและมีการนำไปประยุกต์ใช้ในด้านต่างๆ รวมทั้งมีงานวิจัยที่ทำการประเมินประสิทธิภาพของระบบ อาทิ (1) Julius เป็นระบบที่พัฒนาขึ้นในประเทศญี่ปุ่น (Lee and Kawahara, 2009) (2) Microsoft Speech API (SAPI) พัฒนาขึ้นโดยบริษัทไมโครซอฟท์ (Sharma, 2013) (3) CMU Sphinx เป็นระบบที่พัฒนาโดยมหาวิทยาลัยคาร์เนกีเมลลอน (Plonkowski and Urbanovich, 2014) (4) Kaldi เป็นซอฟต์แวร์โอเพนซอร์สเขียนขึ้นด้วยภาษา C++ (Ali et. al, 2014) (5) Google Cloud Speech API ซึ่งพัฒนาขึ้นโดยกูเกิล (Anggraini, 2018) (6) IBM® (Speech to Text) ซึ่งให้บริการบนระบบคลาวด์ของไอบีเอ็ม และ (7) WIT ซึ่งพัฒนาโดย Facebook ทั้งนี้แต่ละระบบมีประสิทธิภาพการทำงานที่แตกต่างกันและรองรับภาษาที่แตกต่างกัน

ประสิทธิภาพของระบบสั่งงานด้วยเสียงถูกวัดในรูปของความถูกต้องของการแปลงข้อความเสียงเป็นข้อความซึ่งมีวิธีการวัดได้หลายวิธี Xu et al. (2021) นำเสนอวิธีการวัดความถูกต้องของระบบประกอบด้วย (1) อัตราค่าผิดพลาดของคำ (Word error rate, WER) (2) อัตราค่าผิดพลาดของการจับคู่ (Match error rate, MER) (3) การสูญเสียข้อมูลของคำ (Word information lost, WIL) และ (4) การรักษาข้อมูลของคำ (Word information preserve,

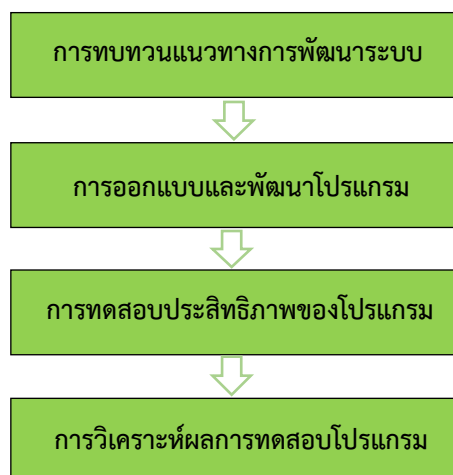
WIP) ทั้งนี้ Lancu (2019) กล่าวว่า วิธีการที่มีการใช้อย่างแพร่หลาย ได้แก่ วิธีอัตราค่าผิดพลาดของคำ (word error rate) ซึ่งเป็นการเปรียบเทียบจำนวนคำที่ผิดกับจำนวนคำทั้งหมด

สำหรับงานวิจัยด้านการประยุกต์ใช้ระบบสั่งงานด้วยเสียงในประเทศไทยมีไม่มากนัก Muangpool et al. (2015) ใช้เครื่องมือกูเกิลเอพีไอพัฒนาแอปพลิเคชันเพื่อแปลงเสียงคำพูดออกมาเป็นข้อความตัวอักษรผ่านทางโทรศัพท์หรือแท็บเล็ตบนระบบปฏิบัติการแอนดรอยด์ ในรูปแบบของการแปลงข้อความที่เป็นการอ่านบทความ และการแปลงบทสนทนาในชีวิตประจำวัน Namahoot (2014) พัฒนาระบบสารสนเทศสำหรับการสืบค้นข้อมูลการท่องเที่ยวในจังหวัดพิษณุโลกโดยใช้เสียงในการค้นหาข้อมูล ส่วนงานวิจัยที่เกี่ยวข้องกับการถอดปริมาณงานก่อสร้าง ได้แก่ Olanrewaju et al. (2020) พัฒนาระบบสั่งงานด้วยเสียงสำหรับถอดปริมาณงานก่อสร้างโดยใช้ภาษาวิชาวเบสิกในการพัฒนา อย่างไรก็ตาม งานวิจัยดังกล่าวนี้พัฒนาเฉพาะการสั่งงานด้วยเสียงเพื่อเรียกใช้คำสั่งต่างๆ เช่น การเปิดเครื่องคิดเลข การเปิดเว็บไซต์ ยังไม่มีส่วนที่เกี่ยวข้องกับการกรอกข้อมูลรายละเอียดขององค์อาคาร

วิธีดำเนินการวิจัย

ขั้นตอนการวิจัย

การดำเนินการวิจัยแบ่งขั้นตอนการดำเนินงานออกเป็น 4 ขั้นตอน ได้แก่ (1) การทบทวนแนวทางการพัฒนาระบบ (2) การออกแบบและพัฒนาโปรแกรม (3) การทดสอบประสิทธิภาพของโปรแกรม และ (4) การวิเคราะห์ผลการทดสอบ ดังแสดงในภาพที่ 1



ภาพที่ 1 ขั้นตอนการดำเนินการวิจัย

การทบทวนแนวทางในการพัฒนา

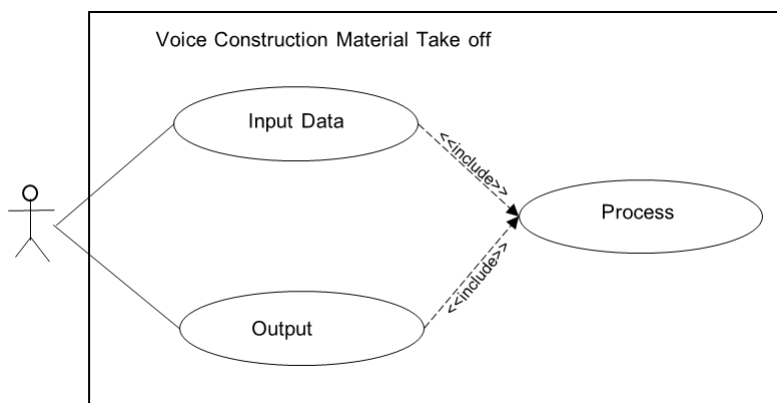
เทคโนโลยีการสั่งงานด้วยเสียงทำหน้าที่เป็นตัวเชื่อมประสานระหว่างอุปกรณ์กับผู้ใช้งาน (User Interface) ซึ่งเป็นการอำนวยความสะดวกในการติดต่อระหว่างมนุษย์กับอุปกรณ์ในขณะที่มือหรือสายตาไม่ว่าง หรือต้องการความคล่องตัวในการทำกิจกรรมอื่น รวมทั้งไม่ต้องการใช้คีย์บอร์ดในการติดต่อกับคอมพิวเตอร์ (Department of Special Investigation, 2021) การพัฒนาระบบสั่งงานด้วยเสียงสามารถดำเนินการโดยใช้ภาษาในการพัฒนาได้หลากหลาย เช่น ภาษาซี ภาษาเบสิก รวมทั้งภาษาไพทอน อย่างไรก็ตาม เนื่องจากภาษาไพทอนเป็นซอฟต์แวร์เสรี

(Open source software) ดังนั้น จึงได้รับความนิยมในปัจจุบันและมีการพัฒนาความสามารถขึ้นอย่างต่อเนื่อง รวมทั้งมีไลบรารีมาตรฐานจำนวนมาก ทำให้สามารถใช้โปรแกรมไพทอนสำหรับพัฒนาทางซอฟต์แวร์ที่หลากหลาย โดยเฉพาะการพัฒนาโปรแกรมการคำนวณทางด้านวิศวกรรมรวมทั้งทางการเรียนรู้ของเครื่อง (machine learning)

การพัฒนาระบบผู้ช่วยสั่งงานด้วยเสียงด้วยภาษาไพทอนเป็นการเรียกใช้งาน API ทั้งนี้ กูเกิลเอพีไอ (Google API) เป็น API หนึ่งที่มีความนิยม โดยกูเกิลเอพีไอที่เกี่ยวข้องกับการสั่งงานด้วยเสียง ประกอบด้วย Google Speech To Text API เป็น API ที่นำเสียงพูดที่ได้จากการบันทึกผ่านไมโครโฟนหรือช่องทางอื่นๆ แปลงเป็นข้อความ โดยถูกพัฒนาขึ้นใน ปี ค.ศ.2008 เพื่อให้งานการค้นหาคำด้วยเสียงสำหรับไอโฟนและพัฒนาขึ้นสู่ระบบคลาวด์ใน ปี ค.ศ. 2017 โดยพัฒนาจากฐานข้อมูลขนาดใหญ่และใช้วิธีการการเรียนรู้ของเครื่อง รองรับหลากหลายภาษารวมทั้งภาษาไทย โดยเมื่อเปรียบเทียบประสิทธิภาพระหว่าง (1) Google Cloud Speech API (2) IBM® (Speech to Text) และ (3) WIT พบว่า Google Cloud Speech API ให้ค่าความถูกต้องสูงที่สุด (Filippidou and Moussiades, 2020) นอกจากนี้ยังมี Google Text To Speech API เป็น API ที่ใช้สำหรับสังเคราะห์เสียงพูดจากข้อความที่กำหนดไว้ ทั้งนี้สำหรับภาษาไพทอน API ที่ใช้สำหรับสังเคราะห์เสียงพูด ได้แก่ gTTS ซึ่งเป็นโมดูลที่เรียก Google TTS (Text-to-Speech) API เข้ามาใช้งาน โดยจะสร้างไฟล์เสียงสังเคราะห์ขึ้นมาในรูปแบบไฟล์ mp3 ไม่จำกัดความยาว รองรับภาษาต่างๆ รวมทั้งภาษาไทย

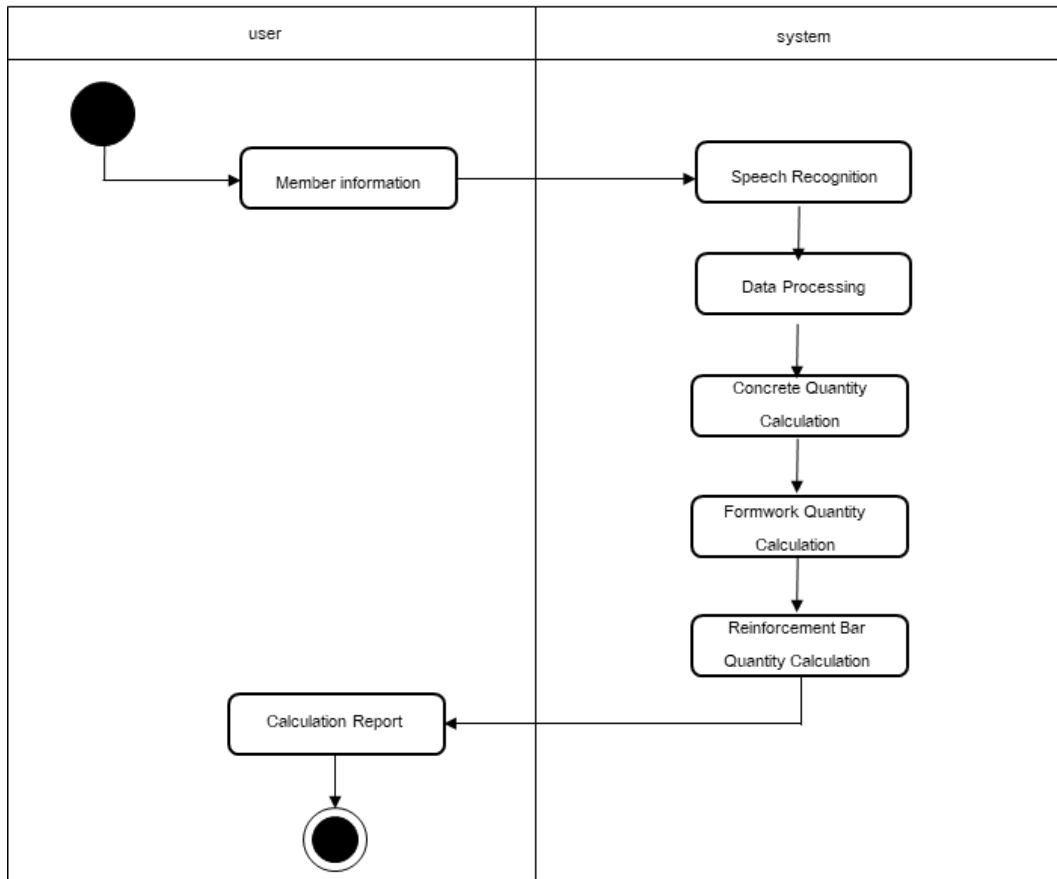
การออกแบบและพัฒนาโปรแกรม

1. การวิเคราะห์ระบบงาน Use case diagram โดย Actor ได้แก่ ผู้ใช้ และ use case ประกอบด้วย (1) การนำเข้าข้อมูล ได้แก่ ข้อมูลเกี่ยวกับองค์อาคาร (2) การประมวลผล และ (3) การรายงานผลการทำงาน โดยผู้ใช้นำเข้าข้อมูลองค์อาคารในระบบ จากนั้นระบบจะทำการประมวลผลข้อมูลเสียงและทำการคำนวณปริมาณงาน ได้แก่ ปริมาณคอนกรีต เหล็กเสริม และไม้แบบ จากนั้นระบบจะแสดงผลการประมวลผลให้ผู้ใช้ทราบ ดังแสดงในภาพที่ 2



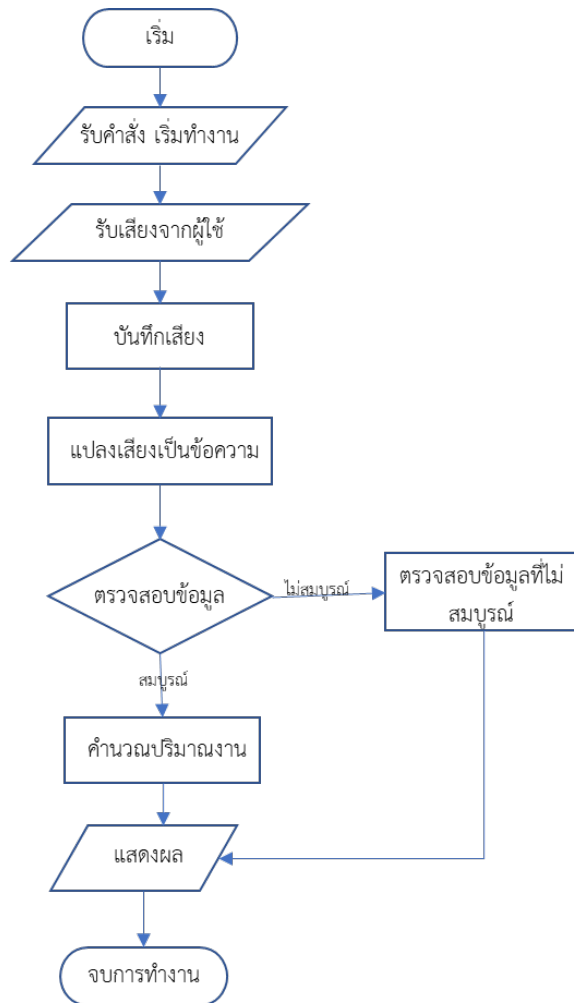
ภาพที่ 2 Use case diagram ของโปรแกรม

2. Activity Diagram แสดงรายละเอียดฟังก์ชัน การทำงานของโปรแกรม ดังแสดงในภาพที่ 3 โดยเมื่อผู้ใช้งานเรียกใช้โปรแกรม ผู้ใช้ทำการป้อนข้อมูลองค์อาคาร ระบบจะทำการแปลงเสียงพูดเป็นตัวอักษร จากนั้นทำการประมวลผลข้อความ และจึงทำการคำนวณปริมาณงานคอนกรีต ไม้แบบ และเหล็กเสริม สำหรับข้อมูลองค์อาคาร และแสดงผลการคำนวณปริมาณงาน



ภาพที่ 3 Activity diagram ของโปรแกรม

3. การพัฒนาระบบสั่งงานด้วยเสียงโดยใช้ไลบรารีหลัก 3 ไลบรารี คือ (1) Speech Recognition 3.8.1 สำหรับแปลงเสียงพูดเป็นตัวอักษร (2) gTTS 2.2.2 สำหรับแปลงข้อความเป็นไฟล์เสียง และ (3) playsound 1.2.2 สำหรับเล่นไฟล์เสียงที่สังเคราะห์ขึ้น โดยมีขั้นตอนหลักในการทำงานดังแสดงในภาพที่ 4 โดยการทำงานของโปรแกรมเริ่มจากรับคำสั่งให้เริ่มทำงานโดยผู้ใช้คลิกที่ปุ่มสั่งงาน โปรแกรมจะเริ่มบันทึกเสียง โดยตั้งค่าให้หยุดบันทึกเมื่อไม่มีเสียงเป็นเวลา 2 วินาที จากนั้นโปรแกรมจะแปลงเสียงเป็นข้อความเก็บไว้ในตัวแปร ขั้นตอนต่อมาคือการตรวจสอบข้อมูล ทั้งนี้โปรแกรมได้ถูกออกแบบให้นำเข้าข้อมูลใน 2 ลักษณะคือ (1) การป้อนข้อมูลลักษณะถามข้อมูลที่ละข้อ (2) การป้อนข้อมูลโดยการให้ข้อมูลทั้งหมดขององค์อาคารครั้งเดียว ดังนั้น ขั้นตอนการตรวจสอบข้อมูลของทั้ง 2 รูปแบบจึงมีความแตกต่างกัน ผลจากการตรวจสอบหากพบว่าข้อมูลไม่สมบูรณ์ระบบจะแสดงผลถึงข้อมูลที่ขาดหายไป ส่วนกรณีที่ข้อมูลสมบูรณ์โปรแกรมจะทำการคำนวณปริมาณงาน ได้แก่ ปริมาณคอนกรีต ปริมาณไม้แบบ และปริมาณเหล็กเสริม



ภาพที่ 4 รูปแบบการทำงานของระบบสั่งงานด้วยเสียง

การตรวจสอบข้อมูลข้อความเสียง แบ่งออกเป็น 2 รูปแบบ ตามลักษณะของการนำเข้าสู่ข้อมูลสำหรับกรณีป้อนข้อมูลลักษณะถามข้อมูลทีละข้อ ระบบจะทำการตรวจสอบข้อความเสียงว่าข้อมูลที่ผู้ใช้งานเข้ามามีความครบถ้วนและมีความถูกต้องหรือไม่ เช่น ข้อมูลขนาดขององค์อาคารต้องเป็นข้อมูลตัวเลข ชนิดเหล็กเสริมเป็นข้อความประเภทเหล็กเสริม เป็นต้น ส่วนกรณีการป้อนข้อมูลโดยการให้ข้อมูลทั้งหมดขององค์อาคารครั้งเดียว ระบบจะทำการประมวลผลชุดข้อมูลที่ได้จากการแปลงข้อมูลเสียงเป็นข้อความ จากนั้นจึงทำการประมวลผลข้อมูลข้อความเพื่อหาคำสำคัญที่แสดงรายละเอียดขององค์อาคาร เช่น คาน กว้าง ยาว ลึก เหล็กบน เหล็กล่าง เป็นต้น จากนั้นจึงทำการตัดข้อมูลข้อความดังกล่าวนี้ไว้ในตัวแปรที่บันทึกรายละเอียดขององค์อาคาร แล้วจึงทำการตรวจสอบความถูกต้องของข้อมูลเช่นเดียวกับกรณีป้อนข้อมูลลักษณะถามข้อมูลทีละข้อ เมื่อตรวจสอบข้อมูลแล้วเสร็จ หากไม่พบข้อผิดพลาดระบบจะทำการประมวลผลต่อไป สำหรับการเตรียมรูปแบบการสนทนาพิจารณาจากลักษณะการทำงานของกรอบแบบก่อสร้างขององค์อาคารแต่ละประเภท ทั้งนี้ Namahoot (2014) พบว่าในสภาวะแวดล้อมแบบไม่มีเสียงรบกวนคำสั่งที่สั้นมีความถูกต้องมากกว่าคำสั่งที่ยาว ส่วนในสภาวะแวดล้อมแบบมีเสียงรบกวน คำสั่งที่สั้นจะมีความถูกต้องน้อยกว่าคำสั่งที่ยาว ดังนั้น การเตรียมรูปแบบของบทสนทนาจึงต้องพิจารณาให้มีความกระชับ และมีการโต้ตอบที่เหมาะสมรวมทั้งคำนึงถึงระยะเวลาที่ใช้ในการกรอกข้อมูลแต่ละประเภท

4. การพัฒนาโปรแกรมส่วนติดต่อกับผู้ใช้ใช้การอินเทอร์เฟซแบบกราฟิกโดยใช้ Tkinter ซึ่งเป็นโมดูลที่ได้รวบรวมฟังก์ชันต่างๆ ให้นำมาใช้สร้างคอมโพเนนต์ได้อย่างหลากหลาย (Thearmontree, 2020) โดยกำหนดให้มีทั้งส่วนเมนู (Menu) และส่วนปุ่มกด (Button) และให้แสดงผลตามประเภทขององค์อาคาร

การทดสอบประสิทธิภาพของโปรแกรม โดยใช้ผู้ทดสอบจำนวน 10 คน ได้แก่ ชาย 5 คน และหญิง 5 คน โดยให้ทดลองใช้งานโปรแกรมสำหรับแต่ละองค์ประกอบขององค์อาคารคนละ 3 ครั้ง ทั้ง 2 รูปแบบ ซึ่งจะทำให้ได้ผลการทดสอบสำหรับองค์อาคารแต่ละประเภทและรูปแบบการป้อนข้อมูลจำนวน 30 ตัวอย่าง โดยผู้ทำการทดสอบแต่ละคนจะได้ทดลองใช้โปรแกรมก่อนในแต่ละองค์อาคารเป็นจำนวนคนละ 3 ครั้ง

การวิเคราะห์ผลการทดสอบโปรแกรม ทำการวิเคราะห์โดยใช้ค่าอัตราค่าผิดพลาดของคำ (word error rate) ซึ่งสามารถหาได้จากสมการที่ (1)

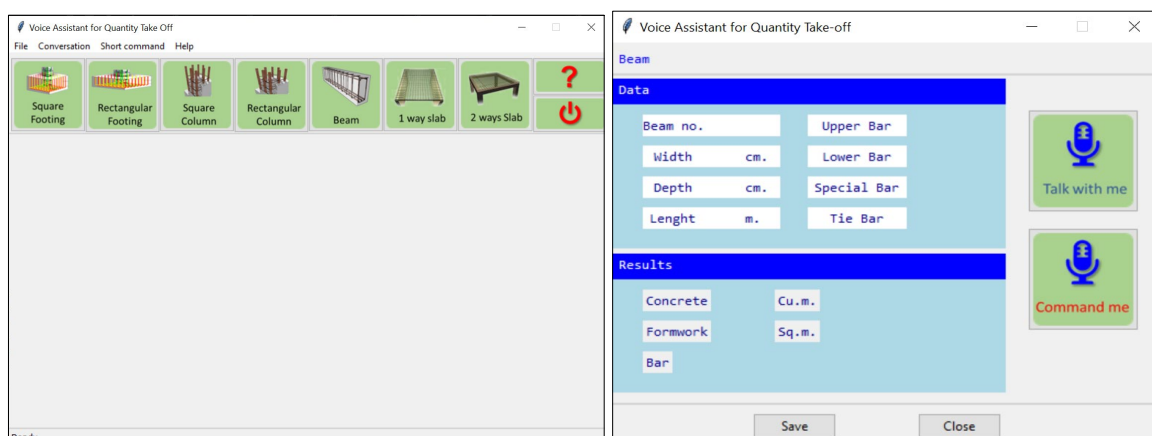
$$WER = (I + D + S) / N \quad (1)$$

โดยที่ I หมายถึง คำที่ถูกแทรกขึ้นมาจากข้อความเดิม (Inserted words), D หมายถึง คำที่หายไปจากข้อความเดิม (Deleted words), S หมายถึง คำที่ถูกแทนที่ไปจากคำเดิม (Substituted words) และ N หมายถึง จำนวนคำทั้งหมด ทั้งนี้อัตราค่าผิดพลาดของคำจะแสดงในหน่วยร้อยละ โดยค่าอัตราค่าผิดพลาดของคำมีค่าสูงจะแสดงว่าระบบมีประสิทธิภาพต่ำ

ผลการวิจัย

1. การพัฒนาส่วนติดต่อกับผู้ใช้

การพัฒนาโปรแกรมในส่วนติดต่อกับผู้ใช้ (user interface) มีวัตถุประสงค์เพื่อให้ใช้งานได้ง่าย จึงได้ออกแบบให้ผู้ใช้สามารถเลือกใช้งานได้ในรูปแบบเมนูและเรียกใช้ได้จากปุ่มกด โดยในเนื้อหาเมนูนอกจากจะเป็นการเรียกใช้งานเพื่อกรอกข้อมูลแล้ว ยังมีส่วนของรายละเอียดวิธีการป้อนข้อมูลของแต่ละองค์อาคารให้ผู้ใช้ทราบถึงรายละเอียดของข้อมูล ดังแสดงในภาพที่ 5



ภาพที่ 5 หน้าจอหลักของโปรแกรม

2. การเตรียมรูปแบบของบทสนทนา

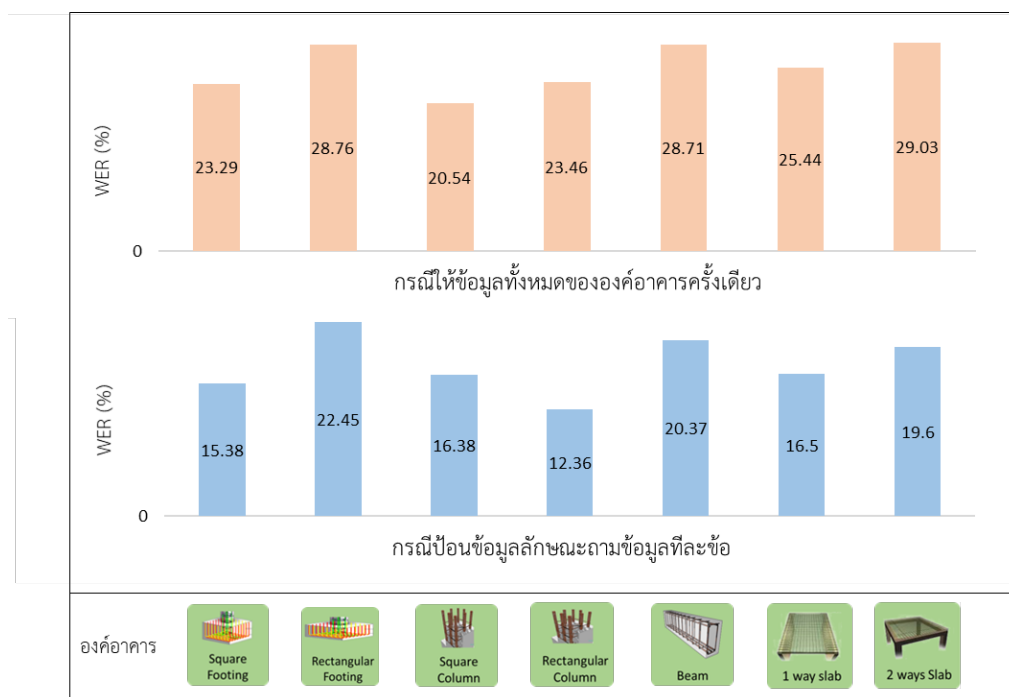
งานวิจัยนี้ได้แบ่งรูปแบบการสนทนาออกเป็น 2 รูปแบบ คือ (1) การป้อนข้อมูลลักษณะถามข้อมูลทีละข้อ (2) การป้อนข้อมูลโดยการให้ข้อมูลทั้งหมดขององค์อาคารครั้งเดียว เพื่อทดสอบประสิทธิภาพของโปรแกรมในการตอบสนองที่แตกต่างกัน นอกจากนี้การตอบโต้ของระบบเป็นประโยคที่แตกต่างกัน 4 รูปแบบ ตัวอย่าง เช่น (1) หมายเลขคานเป็นเท่าไรคะ (2) คานชื่ออะไรคะ (3) คานหมายเลขอะไรคะ และ (4) ขอหมายเลขคานด้วยคะ โดยระบบจะเลือกรูปแบบการตอบสนองที่แตกต่างกันในแต่ละบทสนทนาเพื่อให้การสนทนามีลักษณะเป็นธรรมชาติมากกว่าการโต้ตอบด้วยประโยคเดิมซ้ำๆ แต่เนื่องจากบทสนทนาแต่ละรูปแบบจะใช้ระยะเวลาในการแสดงผลที่แตกต่างกัน ดังนั้น งานวิจัยจึงได้ทำการวิเคราะห์ระยะเวลาของบทสนทนาแต่ละบทเพื่อพิจารณาว่าการรวมบทสนทนาย่อยเข้าด้วยกันจะมีระยะเวลารวมเท่าใด โดยพบว่าเวลารวมของฐานรากจัตุรัสมีรูปแบบการสนทนาทั้งหมด 256 รูปแบบ โดยระยะเวลาของบทสนทนามีค่าน้อยกว่า 13000 มิลลิวินาที ฐานรากสี่เหลี่ยมผืนผ้ามีรูปแบบการสนทนาทั้งหมด 4096 รูปแบบและระยะเวลามีค่าน้อยกว่า 20000 มิลลิวินาที เสาจัตุรัสมีรูปแบบการสนทนาทั้งหมด 1024 รูปแบบ โดยมีระยะเวลาสูงสุดประมาณ 16000 มิลลิวินาที เสาสี่เหลี่ยมผืนผ้ามีรูปแบบการสนทนาทั้งหมด 4096 รูปแบบ มีระยะเวลาน้อยกว่า 18500 มิลลิวินาที คานมีรูปแบบการสนทนาทั้งหมด 65536 รูปแบบ มีค่าน้อยกว่า 24500 มิลลิวินาที และพื้นมีรูปแบบการสนทนาทั้งหมด 4096 รูปแบบ มีค่าน้อยกว่า 19000 มิลลิวินาที การตอบสนองของระบบจะทำการสุ่มรูปแบบของบทสนทนาย่อยเพื่อโต้ตอบกับผู้ใช้จน โดยผลรวมของระยะเวลาในการโต้ตอบของระบบสูงสุดคือองค์อาคารประเภทคานมีระยะเวลาโต้ตอบของระบบน้อยกว่า 24500 มิลลิวินาที สำหรับรูปแบบการป้อนข้อมูลโดยการให้ข้อมูลทั้งหมดขององค์อาคารครั้งเดียว ดังแสดงในตารางที่ 1

ตารางที่ 1 การป้อนข้อมูลโดยการให้ข้อมูลทั้งหมดขององค์อาคารครั้งเดียว

องค์อาคาร	บทสนทนา	
	ระบบ	ผู้ใช้
1) ฐานรากสี่เหลี่ยมจัตุรัส	ขอข้อมูลฐานราก	ฐานราก <เช่น F1> ขนาด <เช่น 120 > ลึก <เช่น 30 > เหล็กเสริม <เช่น 6 DB 16>
2) ฐานรากสี่เหลี่ยมผืนผ้า	ขอข้อมูลฐานราก	ฐานราก <เช่น F1> กว้าง <เช่น 80 > ยาว <เช่น 120 > ลึก <เช่น 30 > เหล็กเสริมยาว <เช่น 6 DB 16> เหล็กเสริมสั้น <เช่น 8 DB 16>
3) เสาหน้าตัดสี่เหลี่ยมจัตุรัส	ข้อมูลเสาด้วยค้	เสา <เช่น C1> ขนาด <เช่น 30> สูง <เช่น 3> เหล็กแกน <เช่น 4 DB 16> ปลอก <เช่น 1 RB 6>
4) เสาหน้าตัดสี่เหลี่ยมผืนผ้า	ข้อมูลเสาด้วยค้	เสา <เช่น C1> กว้าง <เช่น 40> หนา <เช่น 30> สูง <เช่น 3> เหล็กแกน <เช่น 4 DB 16> ปลอก <เช่น 1 RB 6>
5) คาน	ข้อมูลคานด้วยค้	คาน <เช่น B1> กว้าง <เช่น 20> ลึก <เช่น 40> ยาว <เช่น 3> เหล็กบน <เช่น 2 DB 16> เหล็กล่าง <เช่น 2 DB 16> ปลอก <เช่น 1 RB 6>
6) พื้นทางเดียว	ขอข้อมูลพื้นทางเดียว	พื้น <เช่น S1> กว้าง <เช่น 200> ยาว <เช่น 400> หนา <เช่น 10> เหล็กยาว <เช่น RB 9 @ 20> เหล็กสั้น <เช่น RB 9 @ 20>
7) พื้นสองทาง	ขอข้อมูลพื้นสองทาง	พื้น <เช่น S1> กว้าง <เช่น 200> ยาว <เช่น 400> หนา <เช่น 10> เหล็กยาว <เช่น RB 9 @ 20> เหล็กสั้น <เช่น RB 9 @ 20>

3. ผลการทดสอบประสิทธิภาพของโปรแกรม

การทดสอบประสิทธิภาพของโปรแกรม พบว่ารูปแบบการป้อนข้อมูลลักษณะถามข้อมูลทีละข้อให้ค่าอัตราค่าผิดพลาดของคำอยู่ระหว่าง 12.36 และ 22.45 ทั้งนี้ความผิดพลาดส่วนใหญ่เป็นกรณีคำที่ถูกแทนที่ไปจากคำเดิม เช่น การใช้คำว่า “RB” และ “DB” แทนชนิดของเหล็กทำให้การแปลผลผิดพลาด เนื่องจากการแปลข้อมูลเสียงเป็นข้อความโปรแกรมแปลเป็นภาษาไทย เช่น “อาบี” หรือ “ดีบี” หรือผิดไปจากคำเดิม เช่น “บีบี” สำหรับรูปแบบโดยการให้ข้อมูลทั้งหมดขององค์อาคารครั้งเดียว พบว่า ให้ค่าอัตราค่าผิดพลาดของคำอยู่ระหว่าง 20.54 และ 28.76 โดยค่าอัตราค่าผิดพลาดของคำของรูปแบบโดยการให้ข้อมูลทั้งหมดขององค์อาคารครั้งเดียวมีค่ามากกว่ารูปแบบการป้อนข้อมูลลักษณะถามข้อมูลทีละข้อ เนื่องจากการป้อนรูปแบบโดยการให้ข้อมูลทั้งหมดขององค์อาคารครั้งเดียวข้อมูลมีความยาวมากกว่า ดังนั้น ผลของการออกเสียงและการเว้นจังหวะในการพูดมีผลต่อการแปลผลซึ่งสอดคล้องกับ Namahoot (2014) ซึ่งพบว่าในสภาวะแวดล้อมแบบไม่มีเสียงรบกวนคำสั่งที่สั้นมีความถูกต้องมากกว่าคำสั่งที่ยาว นอกจากนี้ปัญหานี้ในการแปลผลข้อมูลเหล็กเสริมซึ่งมีความถูกต้องน้อยที่สุดเนื่องจากการใช้คำว่า “RB” และ “DB” เกิดขึ้นเช่นเดียวกับการป้อนรูปแบบการป้อนข้อมูลลักษณะถามข้อมูลทีละข้อ รวมทั้งผลการทดสอบยังสอดคล้องกับ Lancu (2019) ซึ่งทำการทดสอบประสิทธิภาพของ Google Speech-to-Text API โดยพบว่าค่าอัตราค่าผิดพลาดของคำร้อยละ 9.93 ถึง 57.19 ดังแสดงในภาพที่ 6



ภาพที่ 6 ประสิทธิภาพของโปรแกรม

ค่าอัตราค่าผิดพลาดของคำเป็นค่าที่สะท้อนถึงความถูกต้องของการแปลผลข้อมูลจากรูปแบบเสียงเป็นข้อความสำหรับการประมวลผลซึ่งอาจส่งผลกระทบต่อปริมาณงานหากมีการนำข้อมูลผิดพลาดไปประมวลผลอย่างไรก็ตาม ในขั้นตอนการดำเนินการถอดปริมาณงานก่อสร้าง ผู้ใช้งานสามารถตรวจสอบผลการแปลงเสียงเป็นข้อความก่อนที่จะนำไปประมวลผล หากพบว่าการแปลผลผิดผู้ใช้งานต้องทำการป้อนข้อมูลใหม่ เมื่อถูกต้องแล้วจึงนำไป

ประมวลผลต่อไป ดังนั้น ความถูกต้องของการแปลผลข้อมูลจากรูปแบบเสียงเป็นข้อความจึงไม่ส่งผลต่อความถูกต้องของการถอดปริมาณงานก่อสร้าง แต่จะส่งผลต่อเวลาการทำงานที่เพิ่มขึ้นเนื่องจากต้องป้อนข้อมูลใหม่

เมื่อเปรียบเทียบการใช้งานระบบผู้ช่วยสั่งงานด้วยเสียงสำหรับถอดปริมาณงานก่อสร้างอาคารที่พัฒนาขึ้น และการถอดปริมาณงานก่อสร้างด้วยแบบจำลองสารสนเทศอาคาร (BIM) ในประเด็นต่างๆ หากพิจารณาในภาพรวมจะพบว่าระบบผู้ช่วยสั่งงานด้วยเสียงสำหรับถอดปริมาณงานก่อสร้างอาคารที่พัฒนาขึ้นใช้ทรัพยากรในการทำงานน้อยกว่า อย่างไรก็ตาม การเปรียบเทียบประสิทธิภาพในการทำงานเชิงปริมาณ เช่น เวลาที่ใช้ในการทำงาน เวลาในการประมวลผลของคอมพิวเตอร์ ไม่อยู่ในขอบเขตของงานวิจัยนี้ ดังแสดงในตารางที่ 2

ตารางที่ 2 การเปรียบเทียบการใช้งานระบบและการใช้งานแบบจำลองสารสนเทศอาคาร (BIM)

ประเด็นเปรียบเทียบ	แบบจำลองสารสนเทศอาคาร (BIM)	ผู้ช่วยสั่งงานด้วยเสียงสำหรับถอดปริมาณงานก่อสร้างอาคาร
แบบก่อสร้างที่ใช้สำหรับถอดปริมาณงาน	3 มิติ ต้องเปิดโดยใช้โปรแกรมคอมพิวเตอร์	2 มิติ ทั้งรูปแบบกระดาษและไฟล์คอมพิวเตอร์
ประสิทธิภาพของเครื่องคอมพิวเตอร์ที่ใช้	ประสิทธิภาพสูงและควรมีการดจอ	เครื่องคอมพิวเตอร์ทั่วไป
ความยากง่ายในการเรียนรู้	ต้องสามารถใช้งานโปรแกรมในการทำงาน เช่น autodesk revit	เรียนรู้ได้ง่าย
ประสบการณ์ในการถอดแบบก่อสร้าง	สามารถใส่บุคลากรที่มีประสบการณ์ในการถอดแบบก่อสร้างไม่มาก แต่ต้องเข้าใจการกำหนดพารามิเตอร์ในโปรแกรมอย่างดี และผู้ตรวจสอบควรเข้าใจการทำงานของโปรแกรมเพื่อตรวจสอบว่าการกำหนดพารามิเตอร์ต่างๆ ถูกต้องหรือไม่	สามารถใส่บุคลากรที่มีประสบการณ์ไม่มากทำการอ่านแบบและสั่งการด้วยเสียงเพื่อนำเข้าข้อมูลและมีผู้ที่มีประสบการณ์ในการถอดแบบก่อสร้างตรวจสอบ

อภิปรายผล

การถอดปริมาณงานก่อสร้างเป็นขั้นตอนที่ใช้เวลาในการดำเนินการค่อนข้างมาก ถึงแม้จะมีการใช้โปรแกรมคอมพิวเตอร์ช่วยในการถอดแบบแล้วก็ตาม ผู้ที่ทำการถอดปริมาณงานยังคงต้องอ่านแบบก่อสร้างและป้อนข้อมูลเข้าเครื่องคอมพิวเตอร์ผ่านทางคีย์บอร์ด ซึ่งหากทำงานต่อเนื่องเป็นเวลานานทำให้เกิดความเมื่อยล้าและเกิดข้อผิดพลาดขึ้นได้ง่าย งานวิจัยนี้ได้ทำการพัฒนาระบบสั่งการด้วยเสียง (Voice Assistant) สำหรับประยุกต์ใช้ในการถอดปริมาณงานก่อสร้างด้วยภาษาไพทอนโดยใช้ไลบรารีหลัก 3 ไลบรารี คือ (1) Speech Recognition 3.8.1 สำหรับแปลงเสียงพูดเป็นตัวอักษร (2) gTTS 2.2.2 สำหรับแปลงข้อความเป็นไฟล์เสียง และ (3) playsound 1.2.2 สำหรับเล่นไฟล์เสียงที่สังเคราะห์ขึ้น โดยทำการทดสอบการป้อนข้อมูลด้วยรูปแบบการสนทนา 2 รูปแบบ คือ (1) การป้อนข้อมูลลักษณะถามข้อมูลทีละข้อ (2) การป้อนข้อมูลโดยการให้ข้อมูลทั้งหมดขององค์อาคารครั้งเดียว ผลการทดสอบพบว่า การทดสอบประสิทธิภาพของโปรแกรมพบว่ารูปแบบการป้อนข้อมูลลักษณะถามข้อมูลทีละข้อให้ค่าอัตราค่าผิดพลาด

ของคำอยู่ระหว่าง 12.36 และ 22.45 ในขณะที่รูปแบบโดยการให้ข้อมูลทั้งหมดขององค์อาคารครั้งเดียวให้ค่าอัตราค่าผิดพลาดของคำอยู่ระหว่าง 20.54 และ 28.76 ทั้งนี้ความผิดพลาดส่วนใหญ่เป็นกรณีคำที่ถูกแทนที่ไปจากคำเดิม ทั้งนี้ระบบที่พัฒนาขึ้นช่วยให้การทำงานสะดวกมากขึ้น โดยผู้ใช้งานสามารถอ่านแบบก่อสร้างและป้อนข้อมูลโดยการพูดแทนการคีย์ข้อมูลผ่านทางคีย์บอร์ด รวมทั้งระบบที่พัฒนาขึ้นรองรับการทำงานด้วยภาษาไทยทำให้การใช้งานง่ายขึ้น อย่างไรก็ตาม ระบบที่พัฒนาขึ้นยังเป็นเพียงต้นแบบ ยังต้องมีการปรับปรุงระบบการสนทนาระหว่างเครื่องและผู้ใช้งานให้มีความกระชับมากยิ่งขึ้น เพื่อลดเวลาในการทำงานและลดข้อผิดพลาดของการแปลผลคำพูดลง

ข้อเสนอแนะ

1. ข้อเสนอแนะในการนำผลวิจัยไปใช้

1.1 งานวิจัยนี้พัฒนาขึ้นสำหรับการถอดแบบอาคารคอนกรีตเสริมเหล็กและรูปแบบของอาคารเป็นรูปแบบทั่วไป ยังไม่ครอบคลุมองค์อาคารที่มีรูปร่างพิเศษและองค์อาคารประเภทคอนกรีตอัดแรง

1.2 เสี่ยงรบกวนเป็นประเด็นที่ต้องพิจารณาในขณะที่ใช้โปรแกรมซึ่งส่งผลกระทบต่อการทำงานของโปรแกรม ดังนั้นการทำงานในพื้นที่ที่มีเสียงรบกวนต่ำหรือใช้อุปกรณ์เช่นไมโครโฟนจะทำให้การแปลผลมีประสิทธิภาพมากยิ่งขึ้น

2. ข้อเสนอแนะในการวิจัยครั้งต่อไป

2.1 ควรพัฒนาประสิทธิภาพของโปรแกรมในการประมาณราคาในส่วนของการสถาปัตยกรรมซึ่งเป็นองค์ประกอบสำคัญขององค์ประกอบหนึ่งในการประมาณราคางานอาคาร

2.2 ควรพัฒนาโปรแกรมให้ทำงานร่วมกับโปรแกรม Microsoft Excel เพื่อสรุปปริมาณงานและคำนวณราคาก่อสร้างจะทำให้การทำงานมีประสิทธิภาพมากยิ่งขึ้น

เอกสารอ้างอิง

- Ali, A., Zhang, Y., Cardinal, P., Dahak, N., Vogel, S. and Glass, J. (2014). A complete KALDI recipe for building Arabic speech recognition systems. *The Proceedings of 2014 IEEE Spoken Language Technology Workshop (SLT)*, 7-10 December 2014 at South Lake Tahoe (USA), 525-529.
- Anggraini, N., Kurniawan, A., Wardhani, L.K. and Hakiem, N. (2018). Speech Recognition Application for the Speech Impaired using the Android-based Google Cloud Speech API. *TELKOMNIKA*, 16(6), 2733-2739.
- Chunhamaneewat, S. and Tuntisewee, K. (2020) . A study of process to use 3D model (BIM) for design and take off quantities material of complementary wall. *The Proceedings of the 25th National Convention on Civil Engineering*, 15-17 July 2020 at Chonburi, 1-9. (in Thai)
- Department of Special Investigation, Ministry of Justice. (2021) The Study of The Speech Recognition [Online]. Retrieved April 10, 2021, from: <http://www.dsi.go.th/Files/25620919/การศึกษาระบบรู้จำเสียงพูดอัตโนมัติ.pdf>. (in Thai)
- Filippidou, F. and Moussiades, L. (2020). A Benchmarking of IBM, Google and Wit Automatic Speech Recognition Systems. *Artificial Intelligence Applications and Innovations*, 583, 73–82.

- Iancu, B. (2019). Evaluating Google Speech-to-Text API's Performance for Romanian e-Learning. *Informatica Economică*, 23(1), 17-25.
- Lee, A. and Kawahara, T. (2009). Recent development of open-source speech recognition engine julius. *The Proceedings of Asia-Pacific Signal and Information Processing Association, 2009 Annual Summit and Conference (APSIPA ASC 2009)*, 5 October 2009 at Hokkaido University (Japan), 131-137.
- Muangpool, T., Intharasomba, O. and Maneechot, A. (2015). Voice to Text Translation Application for Deafness using Google API. *The Proceedings of the National and International Conference*, 3-4 December 2015 at Uttaradit Rajabhat University, 1-8. (in Thai)
- Namahoot, S.J. (2014). Tourist Information Searching System by Speech. *MUT Journal of Business Administration*, 11(2), 1-20. (in Thai)
- Olanrewaju, O.L., Sandanayake, M., and Babarinde, S.A. (2020). Voice Assisted Key-In Building Quantities Estimation System. *Journal of Engineering Project and Production Management*, 10(2), 114-122.
- Plonkowski, M. and Urbanovich, P. (2014). Tuning a CMU Sphinx-III Speech Recognition System for Polish Language. *Przegląd Elektrotechniczny*, 90(4), 181-184.
- Sharma, K., Prasad, T.V. and Prasad, S.V.A.V. (2013). Hindi speech enabled windows application using Microsoft SAPI. *International journal of computer engineering & technology (IJCET)*, 4(2), 425-436.
- Sukhamwang, A., Nusen, S., Nusen, P., and Kaewmorachoen, M. (2019). Guideline of Construction Quantity Takeoff on BIM-based Model According to Master Format Categories. *KKU Research Journal (Graduate Studies)*, 19(2), 92-105. (in Thai)
- Suriyamongkol, S. and Tuntisewee, K. (2020). Factors Influencing Adoption of Building Information Modeling in Government Sector : A Case Study of Construction Department, Ministry of Defense. *The Proceedings of the 25th National Convention on Civil Engineering*, 15-17 July 2020 at Chonburi, 1-9. (in Thai)
- Thearmontree, S. (2020). *PYTHON Programming Language guide*. Nonthaburi: IDC Premier. (in Thai)
- Tibodee, P., Sudsin, T. and Thodsakai, T. (2021). The Apply of Using Building Information Modeling (BIM) to Present Material Takeoff: Case Study Educational Building. *Journal of Architecture Design and Construction*, 3(1), 57-68. (in Thai)
- Xu, B., Tao, C., Feng, Z., Raqui, Y. and Ranwez, S. (2021). A Benchmarking on Cloud based Speech-To-Text Services for French Speech and Background Noise Effect. *The Proceedings of 6th National Conference on Practical Applications of Artificial Intelligence*, 2021 at France, 1-6.