****

**เว็บแอปพลิเคชันปัญญาประดิษฐ์ตรวจวัดคำตอบ เพื่อประเมินการสอบสัมภาษณ์กองทุนเงินให้กู้ยืมเพื่อการศึกษา (กยศ.) มหาวิทยาลัยราชภัฏจันทรเกษม**

**Artificial Intelligence Web Application Measures Answers To measure**

**the criteria for interviews with the Student Loan Fund (KBA)**

**Chandrakasem Rajabhat University**

**โดย**

**นายศุภณัฐ ชินราช**

**งานวิจัยนี้เป็นส่วนหนึ่งของรายวิชา การวิจัยทางวิทยาการคอมพิวเตอร์**

**ระดับปริญญาตรี สาขาวิทยาการคอมพิวเตอร์**

**คณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏจันเกษม**

**ปีการศึกษา 2564**

หัวข้อวิจัย : เว็บแอปพลิเคชันปัญญาประดิษฐ์ตรวจวัดคำตอบ เพื่อประเมินการสอบสัมภาษณ์

กองทุนเงินให้กู้ยืมเพื่อการศึกษา (กยศ.) มหาวิทยาลัยราชภัฏจันทรเกษม

Artificial Intelligence Web Application Measures Answers To measure

the criteria for interviews with the Student Loan Fund (KBA)

Chandrakasem Rajabhat University

ผู้วิจัย : ศุภณัฐ ชินราช

รหัสประจำตัว : 6111500143

สาขาวิชา : วิทยาการคอมพิวเตอร์

อาจารย์ที่ปรึกษา : อาจารย์ชัยศิริ สนิทพลกลาง

สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยราชภัฏจันทรเกษม อนุมัติให้นับ

โครงงานวิจัยนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาบัณทิต

ลงชื่อ..........................................................

( ผู้ช่วยศาสตราจารย์ รวินทร์ ไชยสิทธิพร)

ประธานหลักสูตร

คณะกรรมการสอบงานวิจัย ลงชื่อ..........................................................

( ผู้ช่วยศาสตราจารย์ รวินทร์ ไชยสิทธิพร)

ประธาน

ลงชื่อ..........................................................

( อาจารย์ ชัยศิริ สนิทพลกลาง)

อาจารย์ที่ปรึกษา

ลงชื่อ..........................................................

( ผู้ช่วยศาสตราจารย์ วรรณา วิโรจน์แดนไทย )

กรรมการ

**บทคัดย่อ**

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อพัฒนาเว็บแอปพลิเคชันปัญญาประดิษฐ์ตรวจวัดคำตอบ เพื่อประเมินการสอบสัมภาษณ์กองทุนเงินให้กู้ยืมเพื่อการศึกษา (กยศ.) มหาวิทยาลัยราชภัฏจันทรเกษม

การพัฒนาเว็ปแอพลิเคชันสำหรับประเมินการสัมภาษณ์โดยระบบที่ผู้วิจัยสร้าง

ขึ้นมาเพื่อนำเทคโนโลยีปัญญาประดิษฐ์มาใช้สามารถตรวจจับคำตอบ คัดแยก จำแนกประเภท รวมถึงตัวเลขจำนวน และนำไปสู่การประเมินผลสอบของตัวผู้สัมภาษณ์ ยังช่วยลดการแพร่ระบาดของเชื้อโควิด จากการเว้นระยะห่างทางสังคมในการเข้าคิวสอบสัมภาษณ์ เฟรมเวิร์ค(framework) ที่ใช้ได้แก่ รีแอค(react) บูทส์แตป

(boost) ภาษาที่ใช้ได้แก่ ไทป์สคิป(Typescript) โหนดเจเอส(Node JS) อัลกอริทึมประญาประดิษฐ์ที่ใช้พัฒนาได้แก่ เทคทูสปีด(Take Two Speed) สปีดทูเทค(Speed ​​2 Tech) แทคซักเกตชัน(tactics) คอนดิชันนอลแรมดอมฟิล(Conventional Random fil) เครื่องมือที่ใช้พัฒนา ได้แก่ วิชัวสตูดิโอโค้ด(Visual Studio Code) ยานซีแอลไอ(CLI) ผลที่ได้จากการพัฒนาเว็บแอปพลิเคชันปัญญาประดิษฐ์ตรวจวัดคำตอบ ได้รับระบบประเมินผลการสอบสัมภาษณ์ และนำเทคโนโลยีปัญญาประดิษฐ์มาใช้ประโยชน์ร่วม

**กิตติกรรมประกาศ**

งานวิจัยการพัฒนาเว็บแอปพลิเคชันปัญญาประดิษฐ์ตรวจวัดคำตอบ เพื่อประเมินการสอบสัมภาษณ์

กองทุนเงินให้กู้ยืมเพื่อการศึกษา (กยศ.) มหาวิทยาลัยราชภัฏจันทรเกษม บนระบบเว็บบราวเซอร์ สำเร็จลุล่วงได้ด้วยความกรุณาจาก อาจารย์ชัยศิริ สนิทพลกลาง อาจาร์ยที่ปรึกษางานวิจัยที่ได้ให้

คำปรึกษา แนวคิด ตลอดจนให้โอกาสแก้ไขข้อบกพร่องต่างๆ ตลอดระยะเวลาพัฒนาวิจัย จนงานวิจัยเสร็จสมบูรณ์ ผู้วิจัยขอกราบ ขอบพระคุณเป็นอย่างสูง

ขอขอบพระคุณอาจารย์ทุกท่านภายในสาขาวิชาวิทยาการคอมพิวเตอร์ที่ให้คำปรึกษาต่างๆ

รวมทั้งข้อเสนอแนะ และกำลังใจที่ดีเสมอมา ทำให้งานวิจัยครั้งนี้สำเร็จลุล่วง

ขอขอบพระคุณบิดามารดา ที่สนับสนุนและให้กำลังใจจนงานวิจัยสำเร็จด้วยดี คุณค่า และ

ประโยชน์อันพึงมีจากการวิจัยนี้ ผู้วิจัยขอน้อมบูชาพระคุณบิดามารดา และบูรพาจารย์ทุกท่าน ที่ทำให้การ

ศึกษาวิจัยนี้สำเร็จลุล่วงได้ดี

นายศุภณัฐ ชินราช

**สารบัญ**

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  | | **หน้า** |
| **บทคัดย่อ** | | **ก** |
| **กิตติกรรมประกาศ** | | **ข** |
| **สารบัญ** | | **ค** |
| **สารบัญตาราง** | | **จ** |
| **สารบัญภาพ** | | **ฉ** |
| **บทที่ 1 บทนำ** | | **1** |
|  | ความเป็นมาและความสำคัญของปัญหา | 1 |
|  | วัตถุประสงค์ | 2 |
|  | หลักการ ทฤษฎี เหตุผล | 2 |
|  | ระยะเวลาดำเนินการ | 3 |
|  | แผนการดำเนินงาน ขอบเขตการศึกษา | 3 |
|  | ประโยชน์ที่ได้รับ | 5 |
|  | คำนิยามศัพท์เฉพาะ | 5 |
| **บทที่ 2 เอกสารและงานวิจัยที่เกี่ยวข้อง** | | **7** |
|  | การเรียนรู้ของเครื่อง | 7 |
|  | การจัดหมวดหมู่ | 10 |
|  | อัลกอริทึมการจำแนกประเภทหลายฉลาก | 12 |
|  | งานวิจัยที่เกี่ยวข้อง | 15 |
| **บทที่ 3 วิธีดำเนินการวิจัย** | | **17** |
|  | ศึกษาความเป็นไปได้ของปัญญาประดิษฐ์ |  |
|  | ออกแบบระบบ |  |
|  | พัฒนาเว็บแอปพลิเคชัน |  |
|  | ทดสอบและนำผลงานขึ้นระบบพร้อมประเมิน |  |
|  | สรุปและคำนวณ |  |

**สารบัญ (ต่อ)**

|  |  |  |
| --- | --- | --- |
|  |  | **หน้า** |
| **บทที่ 4 ผลการดำเนินงาน** | |  |
| ขั้นตอนการใช้งานแอพพลิเคชัน | |  |
| ผลการประเมินแอพพลิเคชัน | |  |
| **บทที่ 5 บทสรุป** | |  |
| สรุปผลการดำเนินงาน | |  |
| ข้อจำกัดของการวิจัย | |  |
| ข้อเสนอแนะ | |  |
| **บรรณานุกรม** | |  |
| **ประวัติผู้วิจัย** | |  |

**สารบัญตาราง**

|  |  |  |
| --- | --- | --- |
| **ตารางที่** |  | **หน้า** |
| 2.1 |  |  |
| 2.2 |  |  |
| 2.3 |  |  |
| 4.1 |  |  |
| 4.2 |  |  |
| 4.3 |  |  |
| 4.4 |  |  |

**สารบัญรูปภาพ**

|  |  |  |
| --- | --- | --- |
| ภาพที่ |  | **หน้า** |
| 1.1 |  |  |
| 1.2 |  |  |
|  |  |  |
|  |  |  |

**บทที่ 1**

**บทนำ**

**1.1 ความเป็นมาและความสำคัญของปัญหา**

มหาวิทยาลัยราชภัฏจันทรเกษมเป็นหน่วยงานราชการประเภทสถานศึกษา ในทุกๆปีจะมีการรับสมัครผู้ขอทุนการศึกษา แบ่งออกเป็น 7 ประเภท ได้แก่ 1.ทุนรางวัลผลการเรียนดี 2.ทุนช่วยเหลือผู้ขาดแคลน โดยเรามุ่งเน้นไปที่ทุนช่วยเหลือผู้ขาดแคลนประเภทกองทุนเงินให้กู้ยืมเพื่อการศึกษา (กยศ.) ทุนที่มีวัตถุประสงค์ช่วยการสนับสนุนและส่งเสริมการศึกษา โดยการให้กู้ยืมแก่นักเรียนนักศึกษา ที่ขาดแคลนทุนทรัพย์ และนักเรียนนักศึกษาต้องชำระหนี้คืนพร้อมดอกเบี้ยตามอัตราที่กำหนดเมื่อจบการศึกษาแล้ว (อ้างอิง 2563 : กองทุนเงินให้กู้ยืมเพื่อการศึกษา กยศ.) การรับสมัครมีหลายรูปแบบ อาทิเช่น การยื่นผลคะแนนสอบวัดระดับการศึกษา และการสอบสัมภาษณ์ 3.ทุนส่งเสริมการศึกษาเฉพาะทาง 4.ทุนสนับสนุนกิจกรรมเสริมหลักสูตร 5.ทุนสำหรับผู้ด้อยโอกาสทางสังคมและวัฒนธรรม 6.ทุนเงินยืมเพื่อการลงทุนพัฒนาบุคคล 7.ทุนการศึกษาเพื่อส่งเสริมธุรกิจ

การสัมภาษณ์ขอทุนเงินให้กู้ยืมเพื่อการศึกษาในปัจจุบัน จากสถิติข้อมูล กยศ. เดือนมิถุนายน 2563 มีผู้เข้าสอบสัมภาษณ์ปริมาณ 5,771,655 ราย เข้าสัมภาษณ์ 10 รายต่อวัน (อ้างอิง 2563 : กองทุนเงินให้กู้ยืมเพื่อการศึกษา กยศ www.studentloan.or.th ) ผู้ถูกสัมภาษณ์ประสบปัญหาการรอสอบสัมภาษณ์ทุนเป็นเวลานาน เสียค่าเดินทาง และเสี่ยงต่อการแพร่เชื้อระบาดโควิดเนื่องจากเป็นที่รวมตัวกันของคนหมู่มาก ในส่วนของปัญหาผู้รับสัมภาษณ์ คือ ผู้ถูกสัมภาษณ์มีจำนวนมากก่อให้เกิดการนั่งเป็นเวลานาน ส่งผลกระทบต่อกล้ามเนื้อของผู้รับสัมภาษณ์ อาการปวด อาทิเช่น ศีรษะ คอ ไหล่ หลัง ขา อาจส่งผลไปถึงกระดูกสันหลังทับเส้นประสาทได้ (อ้างอิง 2561 : Coach Bank Chira: เคล็ดลับง่ายๆรักษาโรคด้วยอาหาร และการเป็นเศรษฐีความสุข ) ปัญหาของการรับฟังผู้ถูกสัมภาษณ์นำไปสู่ผลกระทบต่อสภาวะทางจิตใจของผู้รับสัมภาษณ์ จากทฤษฎีโดยธรรมชาติ ภาวะทางจิตใจมีการปรับเปลี่ยนไปตามสิ่งที่ได้รับรู้จากการฟังปัญหาเชิงลบของบุคคลหนึ่ง ( อ้างอิง 2563 : จากเว็บไซต์ RamaMental.mahidol.ac.th ภาควิชาจิตเวชศาสตร์คณะแพทยศาสตร์โรงพยาบาลรามาธิบดี มหาวิทยาลัยมหิดล ) จากสถานการณ์ปัญหาข้างต้น อารมณ์ความรู้สึกของผู้รับสัมภาษณ์ ส่งผลถึงการประเมินคะแนนผู้ถูกสัมภาษณ์ทำให้เกิดความคลาดเคลื่อนได้ และยังเสี่ยงต่อการแพร่ระบาดโควิด

ผู้วิจัยเล็งเห็นถึงความสำคัญดังกล่าวจึงได้นำเอาวิทยาการทางคอมพิวเตอร์สมัยใหม่ที่เรียกว่าแมชชีนเลิร์นนิ่ง (Machine Learning) ซึ่งสามารถตรวจจับคำตอบจากผู้สัมภาษณ์ทางการออกเสียง และตรวจจับตัวเลข ประเภทของคำตอบ พร้อมประเมินคะแนนโดยเกณฑ์สัมภาษณ์กองทุนเงินให้กู้ยืมเพื่อการศึกษา (กยศ.) เข้ามาช่วยแก้ปัญหาโดยการนำมาสร้างเว็บแอปพลิเคชันหุ่นยนต์ประเมินคะแนนสำหรับการสอบสัมภาษณ์ เพื่อจัดการคะแนนของผู้ถูกสัมภาษณ์อย่างถูกต้องตามมาตรฐาน อีกทั้งรองรับปัญหาข้างต้น และท้ายที่สุดนำผลที่ได้รับการประเมินเข้ายื่นคะแนนต่อกับทางมหาวิทยาลัยราชภัฏจันทรเกษม

**1.2 วัตถุประสงค์**

เพื่อพัฒนาหุ่นยนต์ประเมินคะแนนให้มีความสามารถทางด้านด้านการคัดแยกคำตอบ ประเภท ตัวเลข จากการเรียนรู้ด้วยเสียง ใช้ประโยชน์ในการเข้าเกณฑ์มาตรฐานของ กยศ. ทั้งในด้านความแม่นยำของการประเมินคะแนนและเวลาที่ใช้สัมภาษณ์

**1.3 หลักการ ทฤษฎี เหตุผล**

1.3.1 แปลงเสียงพูดให้เป็นข้อความ (Speech to Text) เป็นการแปลงเสียงพูดให้เป็นข้อความ โดยในบทความนี้เลือกใช้โมดูล speech\_recognition ที่เรียกใช้ API มาจากคลาวด์ของ Google โดยการใช้งานต้องติดตั้งไลบรารี react-speech-recognition สามารถเรียนรู้จากเสียงคำพูดภาษามนุษย์ รองรับทั้งภาษาไทยและภาษาอังกฤษ แล้วแปลงเป็นข้อความ

1.3.2 Text to Speech เป็นกระบวนการแปลงจากข้อความเป็นเสียงพูด โดยในตัวอย่างบทความนี้เลือกใช้บริการของ "วาจา" ผู้ผลิตโดยองกรณ์ Nectac เพื่ออ่านข้อความออกมาเป็นเสียง ซึ่งผลลัพธ์ที่ได้จาก วาจา เป็นไฟล์เสียง MP3 โดยหลักการเลียนแบบเสียงมนุษย์ ใช้เสียงสังเคราะห์

1.3.3 การรูจําชื่อเฉพาะภาษาไทย (Conditional Random Fields) วิธีสำหรับการทำ Labeling และการทำ Segmentation สำหรับข้อมูลแบบลำดับ (Sequence Data) คือการใช้ Hidden Markov Models (HMMs) หรือ เครื่องที่ทำงานตามสถานะ จำกัดตามความจะเป็น (Probabilistic Finite-state Automata) เพื่อที่จะระบุลำดับที่เป็นไปได้จากประเภทข้อมูลนำเข้า ซึ่งอาจแทนด้วยรูปแบบ, สัญลักษณ์หรือ คุณสมบัติ(Features) ที่กําหนดโดย ที่รูปแบบที่ถูกสร้างจาก Hidden Markov Models (HMMs) นั้นสามารถถูกกำหนดด้วยการกระจาย ความน่าจะเป็นร่วมกันของตัวแปรเชิงสุ่มX และY โดยที่ X คือข้อมูลนำเข้า (ObservationSequence) ส่วน Y คือลักษณะเฉพาะหรือผลลัพธ์ (label Sequence) ซึ่งตรงกับคุณสมบัติของข้อมูลนำเข้า X แบบจำลองฮิดเดนมาร์คอฟนี้เป็น Generative Model และใช้คำความน่าจะเป็นร่วม(Joint Probability) ระหว่างข้อมูลนําเข้า(X) และ ลักษณะเฉพาะหรือผลลัพธ์(Y) ซึ่งอาจพบปัญหาว่าไม่ สามารถจับคู่ความสัมพันธ์ระหว่างคุณสมบัติของลักษณะเฉพาะต่างๆกับข้อมูลนำเข้า

1.3.4 การจำแนกประเภทหลายฉลาก (TAG SUGGESTION) การจำแนกรูปแบบนี้ คลาสทั้งหมดที่เป็นคำตอบได้มีมากกว่าสองคลาส และข้อมูลอาจถูกจำแนกให้อยู่ได้มากกว่าหนึ่งคลาส อาทิเช่น การจำแนกคำพูดที่ได้จากการสัมภาษณ์ คำตอบผู้ถูกสัมภาษณ์จัดได้ในหลายประเภท หรือการจำแนกประเภทรูปภาพรูปหนึ่งอาจจะเป็นไปได้ทั้งรูปขาวดำ รูปทิวทัศน์ รูปวาด เป็นต้น

**1.4 ระยะเวลาดำเนินการ**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ลำดับ | กิจกรรม | ระยะเวลา | | | |
| มกราคม 2565 | กุมภาพัน 2565 | มีนาคม 2565 | เมษายน 2565 |
| 1 | รวบรวมข้อมูล |  |  |  |  |
| 2 | เตรียมข้อมูล |  |  |  |  |
| 3 | สร้างโมเดล |  |  |  |  |
| 4 | สร้างเว็บแอปพลิเคชัน |  |  |  |  |
| 6 | ประเมินผล |  |  |  |  |
| 7 | นำไปใช้งาน |  |  |  |  |

**1.5 แผนการดำเนินงาน ขอบเขตการศึกษา**

1.5.1 แผนการดำเนินงาน

1.5.1.1 รวบรวมข้อมูล

- ศึกษารูปแบบของการสมัครขอทุนการศึกษาของผู้สมัครที่ทำให้เกิดคะแนนของผู้ได้รับทุนการศึกษา

- ศึกษาข้อมูลและใช้เทคนิคการจัดหมวดหมู่ (Classification) ในการสร้างเว็บแอปพลิเคชันเพื่อวิเคราะห์และหาสิ่งสำคัญจากข้อมูล และนำมาสู่การให้คะแนนอย่างเที่ยงตรงแม่นยำ

- ตั้งคำถามที่อยู่ในเกณฑ์การประเมินที่ใช้ในการกู้ กยศ.

1.5.1.2 เตรียมข้อมูล

- คำตอบให้เป็นตัวเลขทั้งหมด เพื่อให้เข้าถึงได้ง่ายและเรียกใช้ได้อย่างรวดเร็ว เป็นประโยชน์ต่อการทำงานขั้นถัดไป ในการบวกคะแนน

1.5.1.3 สร้างโมเดล

- จัดทำโมเดลเกณฑ์ให้คะแนนการรับสมัครตามเกณฑ์การประเมิน กยศ.

1.5.1.4 สร้างเว็บแอปพลิเคชัน

- พัฒนาเทคโนโลยีเว็บแอปพลิเคชันคัดกรองผู้สอบเข้าสัมภาษณ์ขอทุนการศึกษาด้วยระบบการให้คะแนนจำแนกตัวเลขจากคำตอบ (Conditional Random Fields) มหาวิทยาลัยราชภัฏจันทรเกษม

1.5.1.5 ประเมินผลเว็บแอปพลิเคชัน

- การประเมินผลโดยผู้เชี่ยวชาญที่นำหุ่นยนต์ไปใช้จริง - พัฒนาเทคโนโลยีเว็บแอปพลิเคชันคัดกรองผู้สอบเข้าสัมภาษณ์ขอทุนการศึกษาด้วยระบบการให้คะแนนจำแนกประเภทหลายฉลาก (Multi-Label Classification) มหาวิทยาลัยราชภัฏจันทรเกษม

1.5.1.6 ประเมินผลเว็บแอปพลิเคชัน

- การประเมินผลโดยผู้เชี่ยวชาญที่นำหุ่นยนต์ไปใช้จริง

1.5.2 ขอบเขตการศึกษา

ข้อมูลที่นำมาใช้ทดสอบการจำแนกในการวิจัยนี้ คือ ข้อมูลเชิงทั่วไปขนาดใหญ่ จากแบบสอบถามออนไลน์ Google form ประกอบด้วย ชุดข้อมูลเรียนรู้ และชุดข้อมูลที่ไม่ได้นำมาเรียนรู้ อาจจัดอยู่ได้ในหลายคลาส คลาสเหล่านี้มีความสัมพันธ์ตามโครงสร้างที่กำหนดมาให้ โดยโมเดลการจัดโครงสร้างคือการจำแนกประเภทหลายฉลาก (Multi-Label Classification)

1.5.3 ฮาร์ดแวร์ (Hardware) ที่ใช้

1.5.3.1 คอมพิวเตอร์ Spec ขั้นต่ำ

OS : Windows 7/8.1/10 (64-bit versions)

Processor : Intel Dual core or AMD at 2.8 GHz

Memory : 4 GB RAM

Graphics : nVidia GeForce 8600/9600GT

Storage : 20 GB available space

1.5.3.2 โทรศัพท์ Spec ขั้นต่ำ  
OS : Android 7.0

Internet : 4G

Battery : 30%

ROM : 16 GB

RAM : 4 GB

1.5.4 ซอฟท์แวร์ (Software) ที่ใช้

1.5.4.1 Editor : Visual Studio Code, Repl

1.5.4.2 Database : 000webhost, phpMyAdmin, Mysqli

**1.6 ประโยชน์ที่คาดว่าจะได้รับ**

1.6.1 ได้อัลกอริทึมวิเคราะห์ข้อมูลผู้สอบสัมภาษณ์ทุนการศึกษาที่มีประสิทธิภาพมากที่สุด

1.6.2 ได้ระบบการวิเคราะห์ข้อมูลการสมัครเข้าร่วมรับทุนเพิ่มนำไปช่วยในการตัดสินใจของเจ้าหน้าที่ทางมหาวิทยาลัยราชภัฏจันทรเกษม

1.6.4 ได้ความสะดวกสบายในการคาการณ์ให้คะแนน ลดระยะดวลาในการสอบสัมภาษณ์ของนักศึกษาในแต่ละปีการศึกษา

1.6.5 สามารถนำการวิเคราะห์มาช่วยในการสอบสัมภาษณ์ทุนการศึกษาตามมารตรฐาน ในแต่ละปีการศึกษา

1.6.7 สามารถประหยัดจำนวนผู้รับสัมภาษณ์

**1.7 คำนิยามศัพท์เฉพาะ**

1.7.1 การเรียนรู้ของเครื่อง (Machine Learning) เป็นสาขาหนึ่งของปัญญาประดิษฐ์ที่พัฒนามาจากการศึกษาการรู้จำรูปแบบ เกี่ยวข้องกับการศึกษาและการสร้างอัลกอริทึมที่สามารถเรียนรู้ข้อมูลและทำนายข้อมูลได้ อัลกอริทึมนั้นจะทำงานโดยอาศัยโมเดลที่สร้างมาจากชุดข้อมูลตัวอย่างขาเข้าเพื่อการทำนายหรือตัดสินใจในภายหลัง แทนที่จะทำงานตามลำดับของคำสั่งโปรแกรมคอมพิวเตอร์

1.7.2 การจัดหมวดหมู่ (Classification) เป็น Model ประเภท Supervised Model หมายถึง Model ที่ต้องมี Target หรือ ตัวแปรที่ใช้วัดเป้าหมาย เป็นตัวตั้งต้นให้เรียนรู้ โดย Target ของ Classification จะเป็นแบ่งออกเป็นกลุ่ม หรือมีลักษณะเป็น Discrete เช่น yes/no, A/B/C เป็นต้น ดังนั้น ในการประเมินผลลัพธ์ที่ได้จาก Classification Model จะสามารถวัดค่าความแม่นยำ หรือ Accuracy ได้

1.7.3 การจำแนกประเภทหลายฉลาก (TAG SUGGESTION) การจำแนกรูปแบบนี้ คลาสทั้งหมดที่เป็นคำตอบได้มีมากกว่าสองคลาส และข้อมูลอาจถูกจำแนกให้อยู่ได้มากกว่าหนึ่งคลาส อาทิเช่น การจำแนกคำพูดที่ได้จากการสัมภาษณ์ คำตอบผู้ถูกสัมภาษณ์จัดได้ในหลายประเภท หรือการจำแนกประเภทรูปภาพรูปหนึ่งอาจจะเป็นไปได้ทั้งรูปขาวดำ รูปทิวทัศน์ รูปวาด เป็นต้น ฟังก์ชัน Tag Suggestion เป็น 1 ฟังก์ชันที่ระบบงานต่างๆ ควรติดตั้งเพื่อบริการผู้ใช้งาน โดย Tag Suggestion จะช่วยให้ผู้ใช้ระบุ Tag กำกับเนื้อหาบทความได้สะดวก โดยระบบจะวิเคราะห์ Tag จากเนื้อหาที่นำเข้า แล้วแสดงผลการวิเคราะห์ให้ผู้ใช้ได้เลือก ชุดข้อมูลในการเรียนรู้ ถูกยกนำมาจากพันทิป (Pantip)

1.7.4 การรูจําชื่อเฉพาะภาษาไทย (Conditional Random Fields) วิธีสำหรับการทำ Labeling และการทำ Segmentation สำหรับข้อมูลแบบลำดับ (Sequence Data) คือการใช้ Hidden Markov Models (HMMs) หรือ เครื่องที่ทำงานตามสถานะ จำกัดตามความจะเป็น (Probabilistic Finite-state Automata) เพื่อที่จะระบุลำดับที่เป็นไปได้จากประเภทข้อมูลนำเข้า ซึ่งอาจแทนด้วยรูปแบบ, สัญลักษณ์หรือ คุณสมบัติ(Features) ที่กําหนดโดย ที่รูปแบบที่ถูกสร้างจาก Hidden Markov Models (HMMs) นั้นสามารถถูกกำหนดด้วยการกระจาย ความน่าจะเป็นร่วมกันของตัวแปรเชิงสุ่มX และY โดยที่ X คือข้อมูลนำเข้า (ObservationSequence) ส่วน Y คือลักษณะเฉพาะหรือผลลัพธ์ (label Sequence) ซึ่งตรงกับคุณสมบัติของข้อมูลนำเข้า X แบบจำลองฮิดเดนมาร์คอฟนี้เป็น Generative Model และใช้คำความน่าจะเป็นร่วม(Joint Probability) ระหว่างข้อมูลนําเข้า(X) และ ลักษณะเฉพาะหรือผลลัพธ์(Y) ซึ่งอาจพบปัญหาว่าไม่ สามารถจับคู่ความสัมพันธ์ระหว่างคุณสมบัติของลักษณะเฉพาะต่างๆกับข้อมูลนำเข้า

**บทที่ 2**

**เอกสารและงานวิจัยที่เกี่ยวข้อง**

การประเมินคะแนนผู้สอบสัมภาษณ์ โดยการสร้างโมเดลในการจำแนกกลุ่มคะแนนของการตอบคำถามและวัดผลที่อาศัยหลักการ การจัดหมวดหมู่ (Classification) ในการวิเคราะห์ข้อมูลที่มีขนาดใหญ่ ด้วยอัลกอริทึมการจำแนกประเภทหลายฉลาก (Multi-Label Classification) ผู้วิจัยได้รวบรวมข้อมูลเอกสาร และงานวิจัยที่เกี่ยวข้อง เพื่อเป็นข้อมูลในการวิจัย ซึ่งมีรายละเอียดดังนี้

2.1 การเรียนรู้ของเครื่อง (Machine Learning)

2.2 การจัดหมวดหมู่ (Classification)

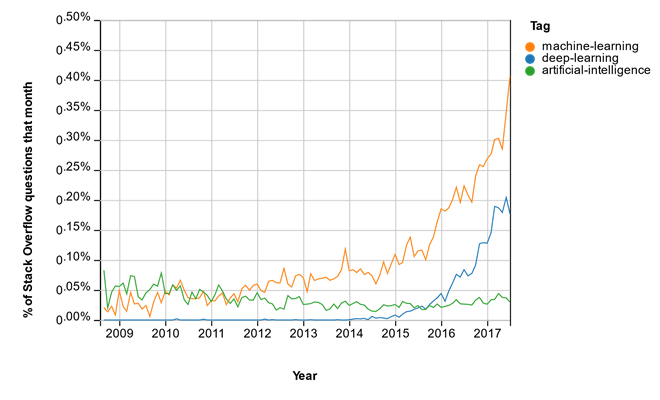
2.3 อัลกอริทึมการจำแนกประเภทหลายฉลาก (Multi-Label Classification)

2.4 งานวิจัยที่เกี่ยวข้อง

**2.1 การเรียนรู้ของเครื่อง (Machine Learning)**

การเรียนรู้ของเครื่อง นำมาใช้เพื่อวิเคราะห์, เข้าใจ และหารูปแบบของข้อมูล หนึ่งในแนวคิดหลักภายใต้ การเรียนรู้ของเครื่อง คือการที่คอมพิวเตอร์สามารถถูกสอนอย่างอัตโนมัติซึ่งสามารถทำได้อย่างหมดจดหรือเป็นไปไม่ได้สำหรับที่มนุษย์จะทำ และยังมีช่องโหว่ที่ชัดเจนจากการวิเคราะห์ยุคก่อนคือการที่การเรียนรู้ของเครื่อง สามารถตัดสินใจได้ด้วยการแทรกแซงจากมนุษย์เพียงเล็กน้อย

การเรียนรู้ของเครื่อง ใช้ข้อมูลเพื่อส่งต่อเข้าไปในอัลกอริทึมซึ่งสามารถเข้าใจความสัมพันธ์ระหว่างข้อมูลขาเข้าและขาออกได้ เมื่อ เครื่องจักรสิ้นสุดการเรียนรู้แล้ว มันสามารถทำนายมูลค่าหรือประเภทของข้อมูลใหม่ได้



ภาพที่ 2-1 กระกระบวนการเรียนรู้ของเครื่อง

ที่มา : <https://www.guru99.com/images/tensorflow/083018_0454_MachineLear2.png>

2.1.1 การเรียนรู้ของเครื่อง (machine learning - ML) เป็นการศึกษาอัลกอริทึมของคอมพิวเตอร์ที่มีการพัฒนาการเรียนรู้ของเครื่องถูกมองว่าเป็นส่วนหนึ่งของปัญญาประดิษฐ์ โดยอัลกอริทึมสร้างแบบจำลองทางคณิตศาสตร์จากข้อมูลตัวอย่าง (เรียกว่า ข้อมูลสอน) เพื่อที่จะคาดการณ์หรือตัดสินใจได้อย่างชัดเจน

การเรียนรู้ของเครื่องพัฒนามาจากการศึกษาการรู้จำแบบ เกี่ยวข้องกับการศึกษาและการสร้างอัลกอริทึมที่สามารถเรียนรู้ข้อมูลและทำนายข้อมูลได้ อัลกอริทึมนั้นจะทำงานโดยอาศัยโมเดลที่สร้างมาจากชุดข้อมูลตัวอย่างขาเข้าเพื่อการทำนายหรือตัดสินใจในภายหลัง แทนที่จะทำงานตามลำดับของคำสั่งโปรแกรมคอมพิวเตอร์ การเรียนรู้ของเครื่องมีเกี่ยวข้องอย่างมากกับสถิติศาสตร์ เนื่องจากทั้งสองสาขาศึกษาการวิเคราะห์ข้อมูลเพื่อการทำนายเช่นกัน นอกจากนี้ยังมีความสัมพันธ์กับสาขาการหาค่าเหมาะที่สุดในทางคณิตศาสตร์ที่แงของวิธีการ ทฤษฎี และการประยุกต์ใช้ การเรียนรู้ของเครื่องสามารถนำไปประยุกต์ใช้งานได้หลากหมาย ไม่ว่าจะเป็นการกรองอีเมล์ขยะ การรู้จำตัวอักษร เครื่องมือค้นหา และคอมพิวเตอร์วิทัศน์

2.2.2 ขั้นตอนการสัมมนา (machine learning - ML)

ค้นคว้าเทคโนโลยีปัญญาประดิษฐ์ ที่จะใช้นำมาคัดแยกคำตอบ ประกอบกับการหาตัวเลขในคำตอบผู้สัมภาษณ์ เพื่อเป็นเข้าสู่การคิดคำนวณคะแนนของชุดคำถามนั้น

เลือกใช้เทคโนโลยีปัญญาประดิษฐ์การคัดแยกคำ จากอัลกอริทึมการรูจําชื่อเฉพาะภาษาไทย โดยชุดข้อมูลที่นำมาเรียนรู้มาจากพันทิป (Pantip) เพื่อเข้ากระบวนการคัดแยกคำที่มาจากภาษาไทย คิดความน่าจะเป็น ที่ใกล้เคียงใช้เฉพาะสกอที่มีความถูกต้อง 98% บวกกับเว้นวรรคตัวอักษร เปรียบเทียบอักษรกับรหัสยูนิโค้ด คัดแยกอักษรพิเศษ อักขระ ตัวเลข ท้ายที่สุดนำผลลัพธ์มาใส่ชนิด

นำเทคโนโลยีปัญญาประดิษฐ์ที่ได้ค้นคว้ามา พัฒนาหลังบ้านที่ถูกเขียนด้วยภาษาโปรแกรมมิ่งโหนดเจเอส (NodeJs) เพื่อส่งข้อมูลเข้าเว็บไซต์ เก็บข้อมูลผู้สัมภาษณ์

นำหลังบ้านที่ถูกพัฒนาต่อจากเทคโนโลยีปัญญาประดิษฐ์มาสาร้างเว็บแอปพลิเคชันหุนยนต์สัมภาษณ์กองทุนกู้ยืมการศึกษา (กยศ.)

เก็บข้อมูลผู้สัมภาษณ์ทั้งผ่านเกณฑ์ และไม่ผ่านเกณฑ์ลงฐานข้อมูลแบบออนไลน์

2.2.3 ประโยชน์ของ (machine learning - ML)

สามารถคัดแยกประโยคเป็นคำได้ ระบุประเภทของคำ คือ ตัวเลข อักษณพิเศษ อักขระ สามารถจัดหมวดหมู่ของคำที่ใกล้เคียงประโยคการสัมภาษณ์ได้

เรียนรู้เสียงพร้อมแปลเป็นคำพูดแทนการพิมพ์ ได้ทั้งภาษาไทยและภาษาอังกฤษ และสามารถแปลงข้อความเป็น เสียงแทนการอ่านได้

กระบวนการของ Machine Learning

สมมติว่าคุณต้องการจะสร้างโปรแกรมที่มีการจดจำวัตถุขึ้นมา เพื่อที่จะฝึก(train) model นั้น,คุณจะต้องใช้ตัวแบ่งประเภท(classifier) classifier ใช้คุณลักษณะ(feature)ของวัตถุเพื่อพยายามหาประเภทของวัตถุนั้น

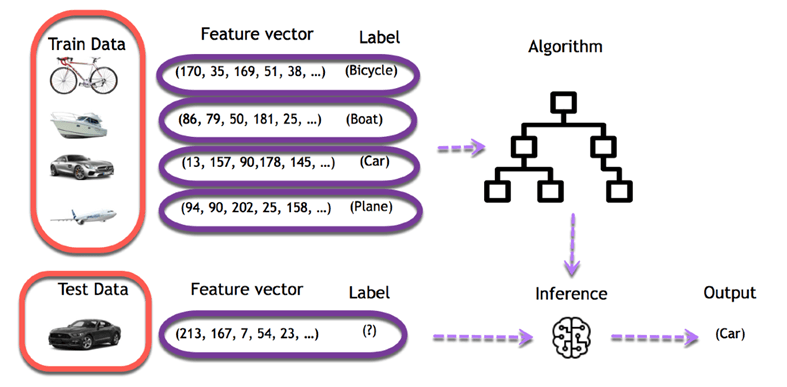
ในตัวอย่าง, classifier จะถูก train เพื่อตรวจจับ(detect)ถ้ารูปเป็น : จักรยาน, เรือ, รถยนต์, เครื่องบิน

วัตถุนี้คือประเภทของวัตถุที่แตกต่างกัน classifierจำเป็นต้องจดจำ เพื่อที่จะสร้าง classifier ขึ้น,คุณจำเป็นต้องมีข้อมูลจำนวนหนึ่งเป็นข้อมูลขาเข้าและกำหนดให้เลยว่าข้อมูลแต่ละอันเป็นประเภทอะไร อัลกอริทึมนี้จะนำข้อมูลไปหารูปแบบ(pattern)แล้วแบ่งประเภทของข้อมูลตามประเภทต่าง ๆ

การกระทำรูปแบบนี้เรียกว่า การเรียนรู้แบบมีผู้สอน(Supervised Learning) ใน Supervised Learning,ข้อมูลสำหรับการฝึก (training data) คุณจะส่งต่อไปในอัลกอริทึมพร้อมกับทำสัญลักษณ์เพื่อบอกผลลัพธ์ไว้แล้ว

การ train อัลกอริทึมจำเป็นต้องทำตามวิธีการดังนี้ : เก็บข้อมูล, Train classifier, ทำการทำนาย

ขั้นตอนแรกเป็นขั้นตอนที่จำเป็นมาก การเลือกข้อมูลที่ถูกต้องจะนำมาซึ่งอัลกอริทึมที่ประสบความสำเร็จหรือล้มเหลว ข้อมูลซึ่งคุณเลือกมาเพื่อ train จะถูกเรียกว่า คุณลักษณะ (feature) ในตัวอย่างของวัตถุ,feature คือพิกเซล(pixel)ของรูป

แต่ละรูปในแถวของข้อมูลขณะที่แต่ละ pixel เป็นหลักแทน ถ้ารูปของคุณมีขนาด 28 x 28 ชุดข้อมูลจะมีขนาด 784 หลัก ในรูปภาพด้านล่าง,แต่ละรูปจะถูกแปลงเป็นเวกเตอร์ของคุณลักษณะ(feature vector) การทำสัญลักษณ์ไว้(label)เป็นการบอกคอมพิวเตอร์ว่าอะไรอยู่ในรูปภาพนั้น

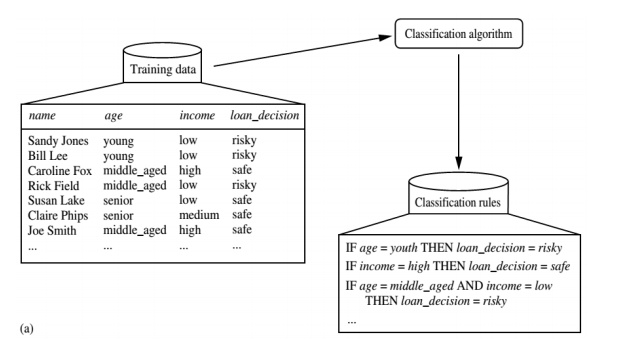
ภาพที่ 2-2 ประเภทของตัวแบบอัลกอริทึมในการเรียนรูปของเครื่อง

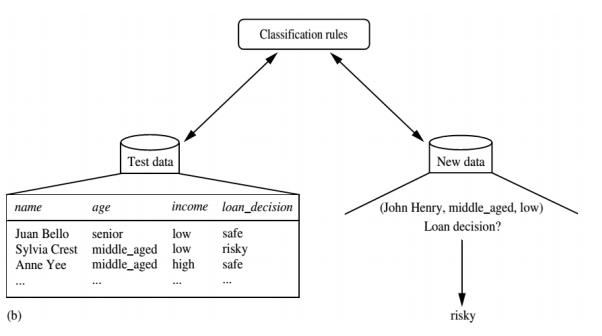
ที่มา : <https://www.guru99.com/images/tensorflow/083018_0454_MachineLear3.png>

วัตถุประสงค์หลักจะเปนการใช้ training data เพื่อแบ่งประเภทของชนิดของวัตถุ ในขั้นตอนแรกประกอบไปด้วยการสร้าง feature เป็นหลัก แล้วต่อมา,ขั้นตอนที่ 2 เกี่ยวข้องกับการเลือกอัลกอริทึมเพื่อ train model นั้น เมื่อ train เสร็จแล้ว,model ดังกล่าว จะทำนายว่ามีสิ่งใดอยู่ในรูปภาพบ้าง

หลังจากนั้น , มันง่ายมากที่จะใช้ model นั้นไปทำนายรูปภาพอื่น ๆ ต่อไป สำหรับการนำรูปภาพใหม่ ๆ เข้าไปสู่ model นั้น ,machine นั้นจะทำนายประเภทของวัตถุนั้นว่าอยู่ประเภทไหน ยกตัวอย่างเช่น คุณมีรูปภาพใหม่ทั้งหมดอยู่โดยปราศจากการ label ไว้ จึงนำไปใช้กับ model ดังกล่าว สำหรับมนุษย์มันเป็นสิ่งง่ายดายมากตอบว่ารูปภาพนั้นมีรถอยู่แต่ machine ใช้ความรู้ที่เพิ่งกล่าวมาทำได้มากสุดแค่ทำนายว่ามีรถอยู่ในรูปภาพนั้นเท่านั้นเอง

**2.2 การจัดหมวดหมู่ (Classification)**

การจำแนกข้อมูลจะประกอบไปด้วยสองกระบวนการหลัก ดังแสดงตัวอย่างในรูปที่ 6.1 ที่เป็นการกู้-ยืมเงิน โดยจากรูปที่ 6.1(a) จะเป็นกระบวนการสร้างตัวจำแนกข้อมูลจากชุดของข้อมูลที่เป็นอินพุต ที่ซึ่งแต่ละเรคคอร์ดของข้อมูลที่ทำการพิจารณาจะประกอบไปด้วยเซตของแอทริบิวที่บ่งบอกถึงคุณลักษณะของบุคคลที่ทำการกู้-ยืมเงิน และหมวดหมู่ของบุคคลนั้นๆว่ามีความปลอดภัยหรือมีความเสี่ยงในการให้กู้-ยืมเงินหรือไม่โดยกระบวนการสร้างตัวจำแนกข้อมูลมักถูกเรียกว่า ‘learning’ หรือ ‘training’ ที่เกิดจากการนำเอาขั้นตอนวิธีสำหรับการจำแนกข้อมูลมาดำเนินการกับข้อมูล ข้อมูลเรคคอร์ด X หนึ่งๆ ในชุดข้อมูลที่ทำการพิจารณา

ภาพที่ 2-3 การจัดหมวดหมู่

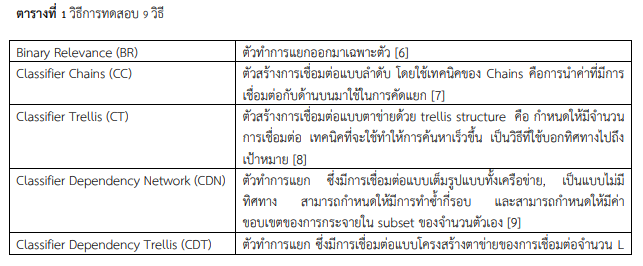
ที่มา : Data mining-การทำเหมืองข้อมูล. 2562 : 2

รูปที่ 6.1 ตัวอย่างกระบวนการในการจำแนกข้อมูล (a) การเรียนรู้จากข้อมูลเพื่อสร้างตัวจำแนกข้อมูล (b) การทดสอบตัวจำแนกข้อมูลเพื่อวัดความถูกต้อง

ประกอบไปด้วยเซตของแอทริบิว 𝑋 = (x1, x2, … , x𝑛), 𝑛 แอทริบิวที่บ่งบอกถึงคุณลักษณะต่างๆของข้อมูลเรค-คอร์ด 𝑋 นอกจากนั้นเรคคอร์ด 𝑋 ยังมีข้อมูลอีกหนึ่งแอทริบิวที่บ่งบอกถึงหมวดหมู่ของข้อมูล (class label attribute) โดยแอทริบิวหมวดหมู่ข้อมูลจะเป็นข้อมูลแบบไม่ต่อเนื่อง (discrete-valued) โดยชุดข้อมูลที่เป็นอินพุตสำหรับทำการสร้างตัวจำแนกข้อมูลจะถูกเรียกว่า “ชุดข้อมูลสำหรับสอน (training data)ตัวอย่าง (samples/instances) ชุดข้อมูล (data points) or สิ่งของ (objects)” เป็นต้น หมายเหตุ—เนื่องจากแต่ละเรคคอร์ดในชุดข้อมูลที่เป็นอินพุตจะมีแอทริบิวหมวดหมู่ข้อมูลแนบอยู่ด้วย ดังนั้น การจำแนกข้อมูลด้วยข้อมูลลักษณะนี้จะเรียกว่า การเรียนรู้แบบมีผู้สอน (Supervised learning—คือ การสร้างตัวจำแนกข้อมูลจะถูกสอนโดยแอทริบิวหมวดหมู่ข้อมูลต่างๆที่ถูกแนบอยู่ในแต่ละเรคคอร์ดของชุดข้อมูล โดยการเรียนรู้แบบมีผู้สอนจะแตกต่างกับการเรียนรู้แบบไม่มีผู้สอน (Unsupervised learning หรือ clustering) ที่จะไม่ทราบถึงหมวดหมู่ของข้อมูล ตัวอย่างเช่น ในการวิเคราะห์ข้อมูลการกู้-ยืมเงินที่ไม่มีหมวดหมู่ข้อมูลที่บ่งบอกว่าการกู้ยืมครั้งหนึ่งๆมีความเสี่ยงหรือไม่ เราจะสามารถวิเคราะห์ข้อมูลได้จากการเชื่อมโยงเรคคอร์ดของการกู้-ยืมเงินที่มีลักษณะใกล้เคียงกันหรือเหมือนกันให้อยู่ในกลุ่มเดียวกัน เป็นต้นขั้นตอนที่สองของการจำแนกข้อมูล (ดังแสดงในรูปที่ 6.1(b)) จะเป็นการเรียกใช้ตัวจำแนกข้อมูลที่สร้างขึ้นจากขั้นตอนที่หนึ่งเพื่อทำการจำแนกข้อมูล โดยในตอนเริ่มต้น ตัวจำแนกข้อมูลจะถูกทดสอบและประเมินค่าความถูกต้อง (ถ้าเราใช้ชุดข้อมูลสำหรับสอนในการทดสอบตัวจำแนกข้อมูลจะทำให้ความถูกต้องจำมีค่าค่อนข้างสูง เนื่องจากตัวจำแนกข้อมูลที่สร้างขึ้นจะเหมาะกับข้อมูลชุดนั้นเป็นอย่างมาก (overfit) แต่ถ้าเราใช้ชุดข้อมูลที่แตกต่างออกไปในการทดสอบ (test set) โดยชุดข้อมูลที่ใช้จะต้องมีแอทริบิวหมวดหมู่ข้อมูลแนบอยู่ด้วย จะทำให้เราทราบค่าความถูกต้องของตัวจำแนกข้อมูลได้) โดยค่าความถูกต้องของตัวจำแนกข้อมูลที่ถูกสร้างขึ้นจะเป็นเปอร์เซ็นต์ของตัวจำแนกข้อมูลที่สามารถจำแนกข้อมูลได้อย่างถูกต้อง (ตัวจำแนกข้อมูลบ่งบอก

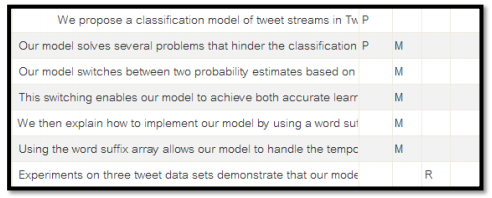
ถึงหมวดหมู่ข้อมูลได้เหมือนกับหมวดหมู่ข้อมูลที่ถูกแนบมากับข้อมูลเรคคอร์ดหนึ่งๆ)เมื่อค่าความถูกต้องของตัวจำแนกข้อมูลมีค่าที่น่าพึงพอใจหรือยอมรับได้ เราจะใช้ตัวจำแนกข้อมูลในการจำแนกหรือบ่งบอกถึงหมวดหมู่ข้อมูลที่เข้ามาใหม่ที่ซึ่งเราไม่ทราบหมวดหมู่ข้อมูลมาก่อน (ข้อมูลที่เข้ามาใหม่จะถูกเรียกว่า ‘unknown’ หรือ ‘previously unseen’ data) ตัวอย่างเช่น ตัวจำแนกข้อมูลที่ถูกสร้างขึ้นในรูปที่ 6.1(a) จะถูกใช้เพื่อตัดสินใจการให้กู้-ยืมเงินของเอกสารที่ยื่นเข้ามาใหม่ว่าจะให้กู้-ยืมหรือไม่

**2.3 อัลกอริทึมการจำแนกประเภทหลายฉลาก (Multi-Label Classification)**

การเรียนรู้จากชุดของตัวอย่าง ปัญหาการจำแนกแบบมัลติเลเบลจัดหมวดหมู่แบบไบนารี(Binary Relevance Methods) ปัญหาการจำแนกประเภทหลายชั้น (multi-label classification) โดยผู้วิจัยเลือกเทคนิคที่ใช้ในการจำแนกข้อมูลแบบไบนารี โดยใช้วิธีการจำแนกแบบมัลติเลเบล จำนวน 9 รูปแบบ ดังนี้**ตารางที่ 1** วิธีการทดสอบ 9 วิธี

ภาพที่ 2-4 อัลกอริทึมการจำแนกประเภทหลายฉลาก

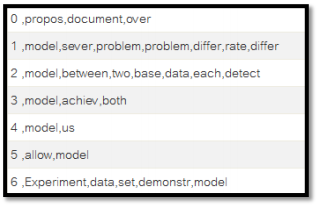
ที่มา : Data mining-การทำเหมืองข้อมูล. 2562 : 56



ภาพที่ 2-5 ตัวอย่างการจำแนกบริบทหน้าที่แต่ละประโยค

ที่มา : การประชุมวิชาการระดับชาติ วิทยาศาสตร์และเทคโนโลยีระหว่างสถาบัน. 2560 : 3

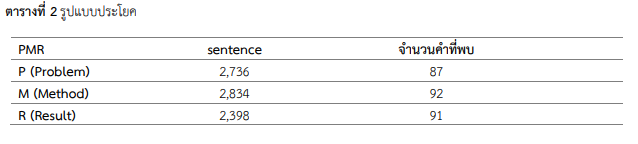
การกำหนดคุณลักษณะ ทำการแปลงข้อมูลโดยใช้กระบวนการ NLP (Natural language processing)ให้อยู่ในลักษณะที่เหมาะสมในการวิเคราะห์ข้อมูลสำหรับด าเนินการและการจำแนกบริบทหน้าที่ของประโยคโดยการหารากศัพท์ (Stemming) และการกำจัดคำหยุด (Stopping) เมื่อได้ประโยคจากขั้นตอนการสกัดคำและการหาบริบทหน้าที่แล้ว ต่อไปจึงน าคำเหล่านี้มาหารากศัพท์โดยใช้อัลกอริทึม Porter[16] ซึ่งเป็นอัลกอริทึมที่ได้รับความนิยม โดยปรับเปลี่ยนคำท้าย (Suffix) ในภาษาอังกฤษ แล้วจึงกำหนดให้ตัวอักษรตัวเล็กทั้งหมด เช่น car, cars, car's, cars = car การกำจัดคำหยุด เป็นกระบวนการการกำจัดคำที่ไม่จำเป็นสำหรับการวิเคราะห์ออกให้เหลือเฉพาะคำที่ความสำคัญเท่านั้น คำเหล่านี้เรียกว่า“Stoplist” ตัวอย่างของคำที่เป็นคำหยุด เช่น a,about, above, across, after, again, against เป็นต้น จากนั้นผู้วิจัยทำการลดขนาดคุณลักษณะด้วยค่าความถี่ (frequency)ของคำโดยการเลือกคำที่เกิดขึ้นมากที่มากที่สุด 100 คำในบทคัดย่อที่ได้รวบรวมเพื่อเป็นตัวคัดเลือกคุณลักษณะ จากนั้นกำหนดขนาดคุณลักษณะโดย D={w1,w2,…wn} และน าข้อมูลที่ทำการแปลงค่าเรียบร้อยแล้วน าไปสู่ขั้นตอนที่ 3.2 ต่อไป



ภาพที่ 2-6 ตัวอย่างการกำหนดคุณลักษณะ

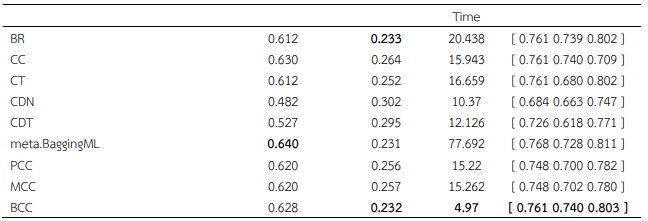
ที่มา : การประชุมวิชาการระดับชาติ วิทยาศาสตร์และเทคโนโลยีระหว่างสถาบัน. 2560 : 4

โดยแบ่งข้อมูลออกเป็น 3 กลุ่มคือ P (Problem)= 2,736 M (Method)= 2,834 R(Result)= 2,398 ทดสอบด้วย 9 วิธีดังกล่าวข้างต้น โดยผู้ศึกษาได้ทดลองกับค่าพารามิเตอร์ให้เหมาะสมกับการจำแนกข้อมูลในแต่ละวิธี การทดสอบได้ค่า 4 ค่าคือ ค่าความถูกต้อง (Accuracy), ค่าระยะที่หายไป (Hamming loss), เวลาในการสร้าง (BuildTime), ค่าความถูกต้องในแต่ละ label (Accuracy (per label)) มาประเมินประสิทธิ์ภาพอัตราการจำแนกโดยคิดเป็นร้อยละความถูกต้องเมื่อเทียบกับจำนวนที่ถูกต้องโดยใช้เครื่องมือในการทดสอบคือโปรแกรม Meka [15] และใช้วิธี10 Cross-validationในการหาค่าต่างๆเพื่อประเมินประสิทธิภาพในครั้งนี้

**ตารางที่ 2** รูปแบบประโยค 

ภาพที่ 2-6 ตัวอย่างการกำหนดคุณลักษณะ

ที่มา : การประชุมวิชาการระดับชาติ วิทยาศาสตร์และเทคโนโลยีระหว่างสถาบัน. 2560 : 4

**ตารางที่3** แสดงผลการทดลอง Multi-label 9 วิธี

สรุปผลการทดลอง

การใช้วิธีการ Multi-label ในแต่ล่ะวิธีจำเป็นที่จะต้องเลือกวิธี classification ที่เหมาะสมในแต่ล่ะวิธีด้วยเช่นกัน โดยผลการทดสอบนั้นได้แสดงให้เห็นถึงการเลือกใช้การ classification ในรูปแบบที่แตกต่างกันออกไปร่วมกับวิธีMulti-label โดยเลือกใช้ทั้ง 9 วิธีเป็นตัวทดสอบกระบวนการ การทดลองในครั้งนี้ในแต่ละวิธีมีข้อดีข้อเสียที่แตกต่างกันออกไปขึ้นอยู่กับผู้ใช้งานต้องการที่จะเน้นค่าอะไรเป็นพิเศษ อาจจะนำวิธีอื่นๆเพื่อมาทดสอบโดยการเปลี่ยนแปลงค่าของวิธี Multi-label เป็นวิธีอื่นๆหรือเปรียบเทียบกับข้อมูลชนิดอื่นๆ เพื่อเปรียบเทียบประสิทธิภาพกับวิธี classification เพื่อให้ได้วิธีที่เหมาะสมและได้ค่าความถูกต้องที่ดีที่สุดในการทดลองกับข้อมูลในแต่ละประเภทที่แตกต่างกันออกไปในงานวิจัยในครั้งต่อไป ซึ่งงานวิจัยชิ้นนี้ต้องการเน้นที่เรื่องของค่าAccuracy ซึ่งจากการทดสอบ ค่า Accuracy อยู่ในช่วง 0.482-0.640 ซึ่งยังให้ผลที่ไม่น่าพอใจสำหรับการทดลองกับข้อมูลชุดนี้ดังนั้นผู้วิจัยจึงต้องท าการทดสอบเพื่อหาวิธีอื่นๆที่สามารถให้ค่า Accuracy ที่สูงกว่าเดิมเพื่อให้ได้การท านายข้อมูลที่ถูกต้องแม่นยำมากยิ่งขึ้น

**2.4 งานวิจัยที่เกี่ยวข้อง**

วัชเรศ ขันธิโชต (2560) การจำแนกประโยคด้วยตัวจำแนกแบบมัลติเลเบล A Study of Sentence Classification Using Multi-label Classifiers วันที่ 25 พฤษภาคม (2560) วิทยานิพนธ์เล่มนี้มุ่งเน้นในการสร้างข้อมูลสนับสนุนการค้นหาเอกสารงานวิจัยเพื่อให้ตรงตามความต้องการของผู้อ่านมากที่สุด โดยมีมุมมองในประเด็นของคำสำคัญ (keyword) ที่ถูกน ามาใช้ในการสืบค้นว่า โดยส่วนใหญ่แล้วมนุษย์มักใช้คำที่คุ้นเคย หรือคำที่เป็นที่รู้จักโดยทั่วไปในการค้นหา ซึ่งคำเหล่านี้เมื่อปรากฏในเอกสารต่างๆ อาจมีบริบทหน้าที่ที่แตกต่างกัน เช่น คำ neural networkอาจเป็นปัญหาที่พบในเอกสารหรืองานวิจัยกลุ่มหนึ่ง แต่อาจเป็นวิธีการที่น ามาแก้ปัญหาในอีกกลุ่มของเอกสารหรืองานวิจัยอื่นได้เป็นต้น ด้วยเหตุนี้ผู้วิจัยมีความเห็นว่าหากระบบสืบค้นสามารถระบุบริบทหน้าที่ของคำสำคัญ อันได้แก่ ปัญหา วิธีการ ผลลัพธ์หรือข้อสรุป ในเอกสารแต่ละชิ้นได้ ย่อมเกิดประโยชน์ต่อผู้อ่านโดยตรง

อัครา ประโยชน (2553) การค้นหาเทคนิคเหมืองข้อมูลเพื่อสร้างโมเดลการวเคราะห์โรคอัตโนมัติ วัตถุประสงค์ของงานวิจัยครั้งนี้ มุ่งเน้นการค้นหาเทคนิคด้านเหมืองข้อมูล เพื่อสร้างโมเดลการวิเคราะห์โรคอัตโนมัติเพื่อค้นหาอลกอริทึมที่เหมาะสมที่สุดสสำหรับฐานข้อมลทางการแพทย์ โดยใช้การ Classification function Network, Support Vector Machine, K-Nearest Neighbor, Decision Tree, Ripperรวมถึงการศกษาเปรียบเทียบการลดคุณลักษณะที่เหมาะสมด้วย วิธี Correlation-based Feature Subset Selection (CFS) และวิธี Feature selection method based on correlation measure and relevance & redundancy analysis (FCBF) รวมถึงทดสอบกับอัลกอริทึมประเภท Single learning กับ Multiple learning โดยเพิ่มประสทธิภาพด้วยวิธี Bagging และ Boosting โดยทดสอบกับข้อมูลทางการแพทย์ทั้ง 13 ชุด

ดร.นิตยา เกิดประสพ (2560) เทคนิคการจำแนกข้อมูลที่พัฒนาสำหรับชุดข้อมูลที่ไม่สมดุลของภาวะข้อเข่าเสื่อมในผู้สูงอาย งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาตัวแบบการพยากรณ์ภาวะข้อเข่าเสื่อมในผู้สูงอายุ จากข้อมูลแบบบันทึกการประเมินข้อเข่าเสื่อม โรงพยาบาลส่งเสริมสุขภาพตำบลบ้านหาร อำเภอท่าศาลา จังหวัดนครศรีธรรมราช จำนวน370 เรคอร์ด และมีข้อมูล 4 คลาส ได้แก่คลาส 0 ยังไม่พบอาการผิดปกติ 200 เรคอร์ด คลาส 1 เริ่มมีอาการข้อเข่าเสื่อม 115 เรคอร์ดคลาส 2 มีอาการโรคข้อเข่าเสื่อมระดับปานกลาง 39 เรคอร์ด และคลาส 3 เป็นโรคข้อเข่าเสื่อมระดับรุนแรง 16 เรคอร์ด สำหรับการวินิจฉัยทางการแพทย์ ข้อมูลกลุ่มน้อย คือ ข้อมูลที่สนใจและการจำแนกผิดพลาดเกิดขึ้นได้สูงกว่าข้อมูลกลุ่มมาก ซึ่งข้อมูลชุดนี้มีจำนวนรวมของคลาส 0 และคลาส 1 สูงกว่าคลาส 2 และคลาส 3 เป็นจำนวนมาก จึงเกิดความไม่สมดุลของข้อมูล ส่งผลให้การจำแนกข้อมูลผิดพลาดได้การปรับความไม่สมดุลของข้อมูลคลาส 2 และคลาส 3 ทำได้ด้วยเทคนิคการปรับเพิ่มข้อมูลด้วยวิธีสุ่ม โดยใช้วิธี ADASYN และSMOTE และใช้งานวิธีการตรวจสอบไขว้แบบ 10 กลุ่ม ในการแบ่งเป็นชุดข้อมูลสอนและชุดข้อมูลทดสอบ จากนั้นจำแนกข้อมูลด้วย multi-class imbalanced data classification

พุทธิพร ธนธรรมเมธ (2560) ตัวแบบการพยากรณ์ที่นำเสนอในงานวิจัยนี้เป็นหนึ่งในเทคนิคของการทำ Classification ข้อมูลที่มีจุดเด่นคือ ค้นหาตัวแบบที่เหมาะสมต่อการนำไปประยุกต์ใช้สำหรับการพยากรณ์ภาวะโรคข้อเข่าเสื่อมในผู้สูงอายุเพื่อทำให้สามารถวินิจฉัยภาวะโรคข้อเข่าเสื่อมในระดับปานกลางและในระดับรุนแรงได้ถูกต้องเพิ่มขึ้นและรวดเร็ว โดยใช้ข้อคำถามเพียง 7 ข้อ จาก 14 ข้อ (ข้อคำถามในแบบประเมินข้อเข่าเสื่อม 12 ข้อ รวมเพศและ BMI) ซึ่งการลดจำนวนข้อคำถามยังคงทำให้พยากรณ์ภาวะโรคข้อเข่าเสื่อมได้ถูกต้อง และมีข้อดีคือ ทำให้ลดจำนวนข้อมูลที่ต้องเก็บรวบรวมจากผู้ป่วยลดเวลาการทำงานของหน่วยพยาบาลปฐมภูมิ และลดเวลาการกรอกข้อมูลของผู้ป่วย แต่มีข้อจำกัดของจำนวนชุดข้อมูล ดังนั้นการพัฒนางานวิจัยให้มีประสิทธิภาพเพิ่มขึ้นจะต้องเพิ่มปริมาณข้อมูลที่นำมาใช้สร้างตัวแบบเพื่อให้มีข้อมูลที่หลากหลาย ซึ่งจะส่งผลให้ตัวแบบมีประสิทธิภาพเพิ่มขึ้น

เยาวเรศ ศิริสถิตกุล (2560) การประเมินเทคนิคการปรับเพิ่มข้อมูลด้วยวิธีสุ่มโดยนำวิธี Classification และ Multi-Label มาใช้ปรับสมดุลข้อมูล ซึ่งเป็นการเพิ่มจำนวนข้อมูลกลุ่มน้อย พบว่าวิธี ADASYN สามารถเพิ่มจำนวนข้อมูลกลุ่มน้อยของคลาส 2 และคลาส 3 ได้สูงกว่าวิธี SMOTE เนื่องจากในการเพิ่มข้อมูลเทียมนั้น วิธี ADASYN จะพิจารณาจากข้อมูลจริงที่ยากต่อการแบ่งกลุ่มข้อมูล ซึ่งไม่จำเป็นต้องพิจารณาข้อมูลทุกตัวที่อยู่ในกลุ่มน้อย ถ้าข้อมูลตัวใดยากต่อการแบ่งกลุ่มก็ให้ค่าน้ำหนักข้อมูลนั้นมากและสร้างชุดข้อมูลเทียมขึ้นมาบริเวณนั้น ๆ ซึ่งทำให้มีการปรับขอบเขตของเส้นการตัดสินใจในการแบ่งกลุ่มให้ดี[27] และพบว่าการปรับเพิ่มข้อมูลด้วยวิธีสุ่มด้วยวิธี ADASYN ทำให้ประสิทธิภาพของตัวแบบสูงกว่าวิธี SMOTE และสอดคล้องกับ Wang และคณะซึ่งพบว่าวิธี SMOTE ส่งผลต่อการทำนายต่ำกว่าวิธี CSC และในงานของ Shoorangiz และคณะ [34]ได้นำวิธี SMOTE และ ADASYN มาใช้ในการทำนายEEG ของภาวะหลับในหรือการหลับระยะสั้น ๆ ผลการทำนายแสดงค่า ROC ของทั้ง 2 วิธีอยู่ในระดับ 90 %เท่ากัน วิธี ADASYN ให้ค่า sensitivity ที่ 76 % แต่วิธีSMOTE อยู่ที่ 70 % ดังนั้น วิธี SMOTE จึงไม่ได้เป็นวิธีที่มีประสิทธิภาพดีที่สุดในการเพิ่มจำนวนข้อมูลสำหรับการทำนาย

**บทที่ 3**

**วิธีดำเนินการวิจัย**

วิธีดำเนินงาน) เว็บแอปพลิเคชันปัญญาประดิษฐ์ตรวจวัดคำตอบ เพื่อประเมินการสอบสัมภาษณ์กองทุนเงินให้กู้ยืมเพื่อการศึกษา (กยศ.) มหาวิทยาลัยราชภัฏจันทรเกษม

ผู้วิจัยได้กำหนดวิธีดำเนินงานออกเป็น 5 ส่วน ดังนี้

3.1 ศึกษาความเป็นไปได้ของปัญญาประดิษฐ์

3.2 ออกแบบระบบ

3.3 พัฒนาเว็บแอปพลิเคชัน

3.4 ทดสอบและนำผลงานขึ้นระบบพร้อมประเมิน

3.5 สรุปและดำเนินงาน

**3.1 ศึกษาความเป็นไปได้ของปัญญาประดิษฐ์**

3.1.1 รวบรวมข้อมูล

เก็บรวบรวมข้อมูลจากงานวิจัย ของกลุ่มวิจัยอุปกรณ์สเปกโทรสโกปีและเซนเซอร์ (SSDRG) เว็บแอปพลิเคชันปัญญาประดิษฐ์ตรวจวัดคำตอบ เพื่อประเมินการสอบสัมภาษณ์กองทุนเงินให้กู้ยืมเพื่อการศึกษา (กยศ.) มหาวิทยาลัยราชภัฏจันทรเกษม

3.1.2 ศึกษาความเป็นไปได้ของปัญหาฃ

ศึกษาเทคโนโลยีปัญญาประดิษฐ์การเรียนรู้ด้วยเสียงเพื่อแปลงเป็นข้อความอักษร ศึกษาเทคโนโลยีปัญญาประดิษฐ์เปลี่ยนข้อความส่งออกเป็นเสียง และความเป็นไปได้อื่น ๆ ของปัญญาประดิษฐ์

**3.2 ออกแบบระบบ**

ผู้วิจัยออกแบบระบบเว็บแอปพลิเคชันปัญญาประดิษฐ์ ซึ่งสามารถตรวจจับคำตอบจากผู้สัมภาษณ์ทางการออกเสียง และตรวจจับตัวเลข ประเภทของคำตอบ พร้อมประเมินคะแนนโดยเกณฑ์สัมภาษณ์กองทุนเงินให้กู้ยืมเพื่อการศึกษา (กยศ.) โดยมีผู้ใช้เป็นนักศึกษาของมหาวิทยาลัยราชภัฏจันทรเกษม

**3.3 พัฒนาเว็บแอปพลิเคชัน**

ผู้วิจัยพัฒนาเว็บแอปพลิเคชันปัญญาประดิษฐ์ตรวจวัดคำตอบ เพื่อประเมินการสอบสัมภาษณ์

กองทุนเงินให้กู้ยืมเพื่อการศึกษา (กยศ.) มหาวิทยาลัยราชภัฏจันทรเกษม โดยการสร้างกระบวนการวิธีให้คะแนนพร้อมทั้งประเมินเกนฑ์วัดผล ดังนี้

(1) สถานะครอบครัวเป็นปัจจุบันเป็นอย่างไรครับ มาจากเงื่อนไข ผู้สัมภาษณ์อธิบายคำตอบผ่านไมโครโฟน ตัวปัญญาประดิษฐ์จะสามารถจับคำตอบที่กล่าวมา คัดแยกผ่านอัลกอริทึมแทกสักเกตชันเมื่อได้แทกที่ตรงกับหมวดหมู่ เลิก อย่าร้าง สถานะครอบครัวแยกย้ายกันอยู่ ตัวหลังบ้านจะทำการเพิ่มคะแนนในข้อนี้เป็นคะแนน 1 คะแนน อ้างอิงจากเกณฑ์การประเมินทางเว็บไซต์ทางการ กยศ.

(2) ผู้สัมภาษณ์มีสัญชาติประเทศอะไรครับ มาจากเงื่อนไข ผู้สัมภาษณ์อธิบายคำตอบผ่านไมโครโฟน ตัวปัญญาประดิษฐ์จะสามารถจับคำตอบที่กล่าวมา คัดแยกผ่านอัลกอริทึมแทกสักเกตชันเมื่อได้แทกที่ตรงกับหมวดหมู่ประเทศไทย Thailand ไทย สัญชาติไทย ตัวหลังบ้านจะทำการเพิ่มคะแนนในข้อนี้เป็นคะแนน 1 คะแนน อ้างอิงจากเกณฑ์การประเมินทางเว็บไซต์ทางการ กยศ.

(3) ผลการเรียนเกรดเฉลี่ยเท่าไหร่ครับ มาจากเงื่อนไข ผู้สัมภาษณ์อธิบายคำตอบผ่านไมโครโฟน ตัวปัญญาประดิษฐ์จะสามารถจับคำตอบที่กล่าวมา คัดแยกผ่านอัลกอริทึมการรูจําชื่อเฉพาะภาษาไทย เมื่อได้คัดแยกตัวเลข พร้อมชนิดข้อความ ใช้กับอัลกอริทึมแทกสักเกตชัน คัดแยกหมวดหมู่ตรงกับหมวด เกรด ผลการเรียน คะแนนการเรียน ตัวหลังบ้านจะทำการเพิ่มคะแนนในข้อนี้เป็น จำนวนตัวเลขเกรด โดยไม่เกิน 4.00 คะแนน อ้างอิงจากเกณฑ์การประเมินทางเว็บไซต์ทางการ กยศ.

(4) ครอบครัวมีรายได้ 1 ปี ทั้งหมดเท่าไหร่ครับ มาจากเงื่อนไข ผู้สัมภาษณ์อธิบายคำตอบผ่านไมโครโฟน ตัวปัญญาประดิษฐ์จะสามารถจับคำตอบที่กล่าวมา คัดแยกผ่านอัลกอริทึมการรูจําชื่อเฉพาะภาษาไทย เมื่อได้คัดแยกตัวเลข พร้อมชนิดข้อความ ใช้กับอัลกอริทึมแทกสักเกตชัน คัดแยกหมวดหมู่ตรงกับหมวด รายได้ต่อไป รายได้ครอบครัว รายได้ครอบครัวต่อปี ตัวเลขรายได้ ตัวหลังบ้านจะทำการเพิ่มคะแนนในข้อนี้เป็น จำนวนตัวเลขรายได้ต่อปี โดยนำรายได้ต่อปีมาหาร 360,000 นำมาใช้เป็นคะแนน อ้างอิงจากเกณฑ์การประเมินทางเว็บไซต์ทางการ กยศ.

(5) ผู้สัมภาษณ์ได้ทำงานพิเศษ หรือไม่ ถ้าใช่อธิบายเกี่ยวกับงานของผู้สัมภาษณ์ครับ มาจากเงื่อนไข ผู้สัมภาษณ์อธิบายคำตอบผ่านไมโครโฟน ตัวปัญญาประดิษฐ์จะสามารถจับคำตอบที่กล่าวมา คัดแยกผ่านอัลกอริทึมแทกสักเกตชันเมื่อได้แทกที่ตรงกับหมวดหมู่งาน พาร์ทไทม์ สายอาชีพ ตัวหลังบ้านจะทำการเพิ่มคะแนนในข้อนี้เป็นคะแนน 1 คะแนน อ้างอิงจากเกณฑ์การประเมินทางเว็บไซต์ทางการ กยศ.

(6) เคยสำเร็จการศึกษาจากที่มหาลัยใดมาก่อนหรือไม่ครับ มาจากเงื่อนไข ผู้สัมภาษณ์อธิบายคำตอบผ่านไมโครโฟน ตัวปัญญาประดิษฐ์จะสามารถจับคำตอบที่กล่าวมา คัดแยกผ่านอัลกอริทึมแทกสักเกตชันเมื่อได้แทกที่ตรงกับหมวดหมู่ปฏิเสธ ไม่เคย กำลังศึกษาอยู่ ไม่มีวุฒิปริญญาตรี เทียบเท่าปริญญาตรี ตัวหลังบ้านจะทำการเพิ่มคะแนนในข้อนี้เป็นคะแนน 1 คะแนน อ้างอิงจากเกณฑ์การประเมินทางเว็บไซต์ทางการ กยศ.

(7) ผู้สัมภาษณ์เรียนอยู่สาขาอะไรครับ และทำไมถึงเลือกมาเรียนสาขา (ชื่อสาขา) ครับ มาจากเงื่อนไข ผู้สัมภาษณ์อธิบายคำตอบผ่านไมโครโฟน ตัวปัญญาประดิษฐ์จะสามารถจับคำตอบที่กล่าวมา คัดแยกผ่านอัลกอริทึมแทกสักเกตชันเมื่อได้แทกที่ตรงกับหมวดหมู่ปฏิเสธ ไม่เคย กำลังศึกษาอยู่ ไม่มีวุฒิปริญญาตรี เทียบเท่าปริญญาตรี ตัวหลังบ้านจะทำการเพิ่มคะแนนในข้อนี้เป็นคะแนน 1 คะแนน อ้างอิงจากเกณฑ์การประเมินทางเว็บไซต์ทางการ กยศ.

(8) ท่านเคยทำประโยชน์ต่อสังคม / สาธารณะ อะไรบ้าง รวมทั้งสิ้นกี่ชั่วโมงครับ มาจากเงื่อนไข ผู้สัมภาษณ์อธิบายคำตอบผ่านไมโครโฟน ตัวปัญญาประดิษฐ์จะสามารถจับคำตอบที่กล่าวมา คัดแยกผ่านอัลกอริทึมการรูจําชื่อเฉพาะภาษาไทย เมื่อได้คัดแยกตัวเลข พร้อมชนิดข้อความ ใช้กับอัลกอริทึมแทกสักเกตชัน คัดแยกหมวดหมู่ตรงกับหมวด จิตอาสา ทำประโยชน์ต่อสังคม ชั่วโมงจิตอาสา ตัวหลังบ้านจะทำการเพิ่มคะแนนในข้อนี้เป็น จำนวนตัวเลขรายชั่วโมงจิตอาสา โดยนำรายชั่วโมงจิตอาสา มาใช้เป็นคะแนน อ้างอิงจากเกณฑ์การประเมินทางเว็บไซต์ทางการ กยศ.

(9) ผู้สัมภาษณ์มีจำนวนกี่คนในครอบครัว มาจากเงื่อนไข ผู้สัมภาษณ์อธิบายคำตอบผ่านไมโครโฟน ตัวปัญญาประดิษฐ์จะสามารถจับคำตอบที่กล่าวมา คัดแยกผ่านอัลกอริทึมการรูจําชื่อเฉพาะภาษาไทย เมื่อได้คัดแยกตัวเลข พร้อมชนิดข้อความ ใช้กับอัลกอริทึมแทกสักเกตชัน คัดแยกหมวดหมู่ตรงกับหมวด จิตอาสา ทำประโยชน์ต่อสังคม ชั่วโมงจิตอาสา ตัวหลังบ้านจะทำการเพิ่มคะแนนในข้อนี้เป็น จำนวนตัวเลขรายชั่วโมงจิตอาสา โดยนำรายชั่วโมงจิตอาสา มาใช้เป็นคะแนน อ้างอิงจากเกณฑ์การประเมินทางเว็บไซต์ทางการ กยศ.

ผู้พัฒนานำเกณฑ์การให้คะแนนข้างต้นมาพัฒนาระบบหลังบ้านตามขั้นตอนกระบวนการดังนี้

**3.4 ทดสอบและนำผลงานขึ้นระบบพร้อมประเมิน**

**3.5 สรุปและดำเนินงาน**

จากการทดสอบส่วนใหญ่ไม่ผ่านเกณฑ์การประเมิน และส่วนน้อยที่ผ่านการประเมิน ทางผู้พัฒนาได้

รวบรวมคำตอบของผู้ใช้นำมาเก็บลงในฐานข้อมูล และระบบสามารถจำแนกการปลอมแปลงอีเมล 14 ครั้ง จาก 67 ครั้ง

**บทที่ 4**

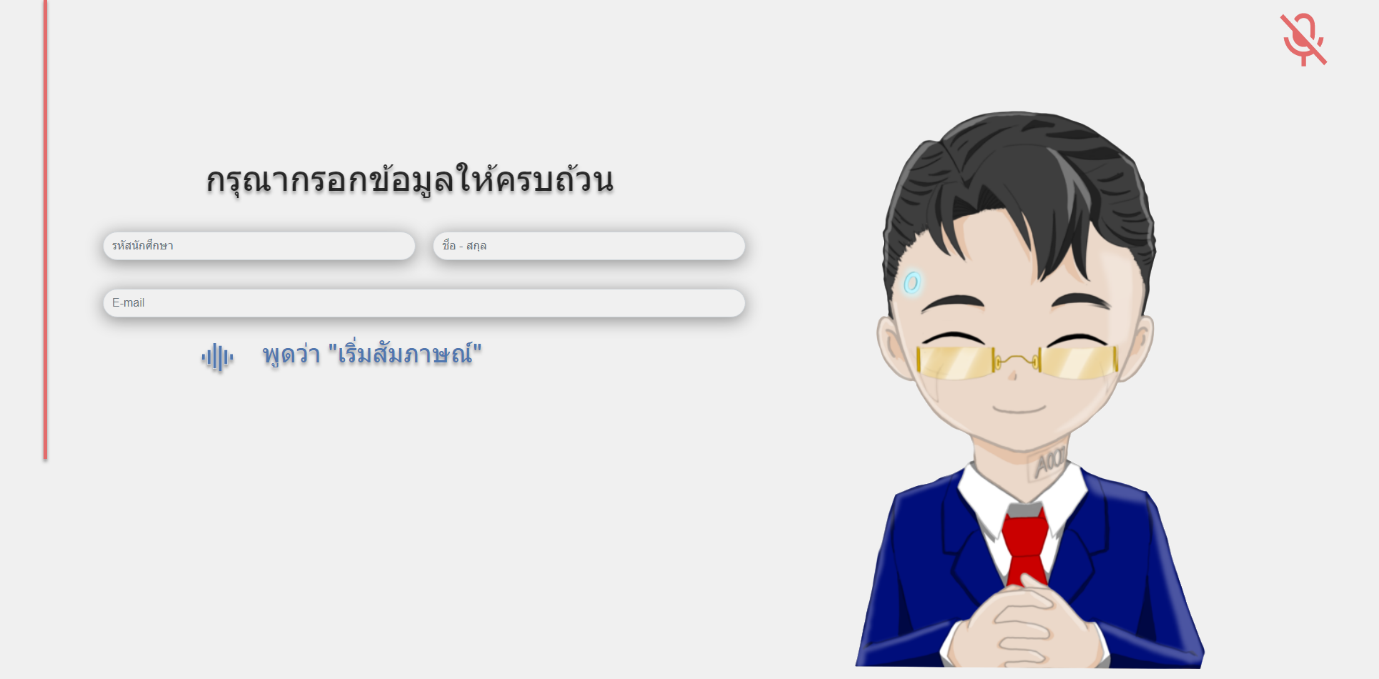
**ผลการดำเนินงาน**

การพัฒนาเว็บแอปพลิเคชันปัญญาประดิษฐ์ตรวจวัดคำตอบ เพื่อประเมินการสอบสัมภาษณ์กองทุนเงินให้กู้ยืมเพื่อการศึกษา (กยศ.) มหาวิทยาลัยราชภัฏจันทรเกษม ผู้วิจัยดืทำการรวบรวมข้อมูล โดยมีการดำเนินงานดังนี้

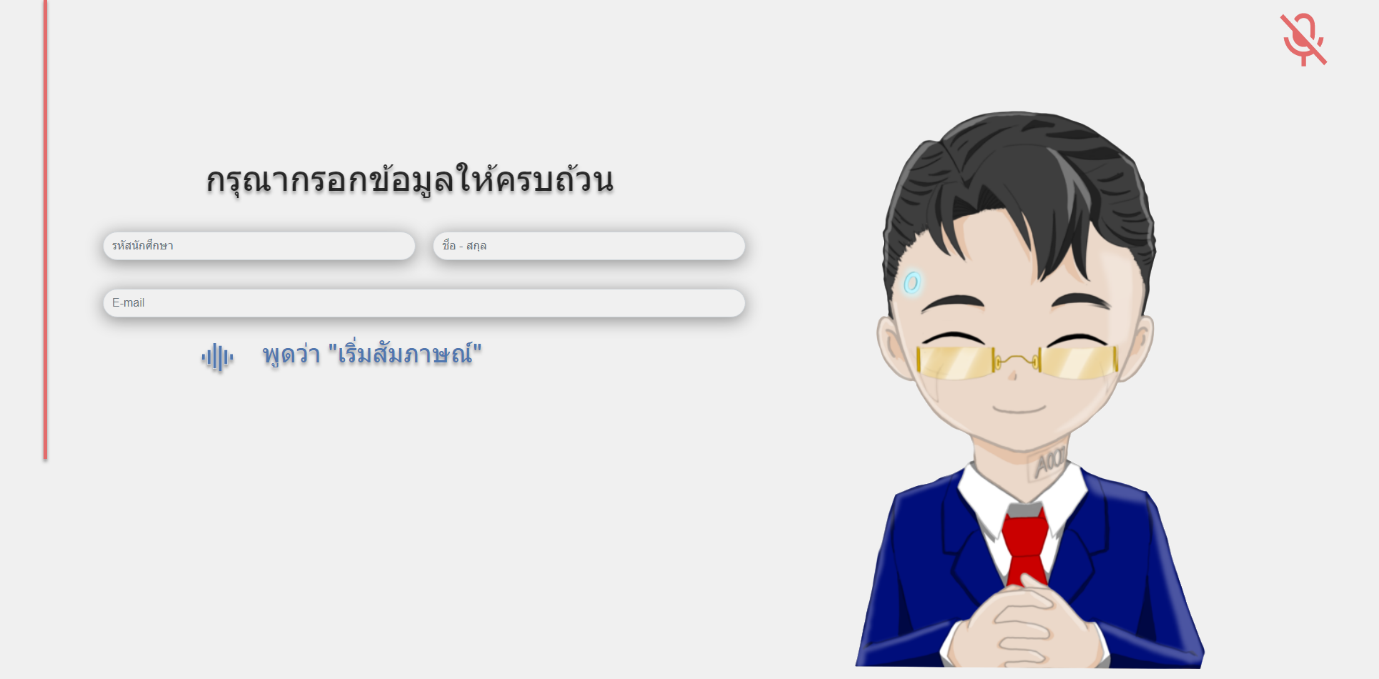
4.1 ขั้นตอนการใช้งานแอปพลิเคชัน

4.2 ผลการประเมินแอปพลิเคชัน

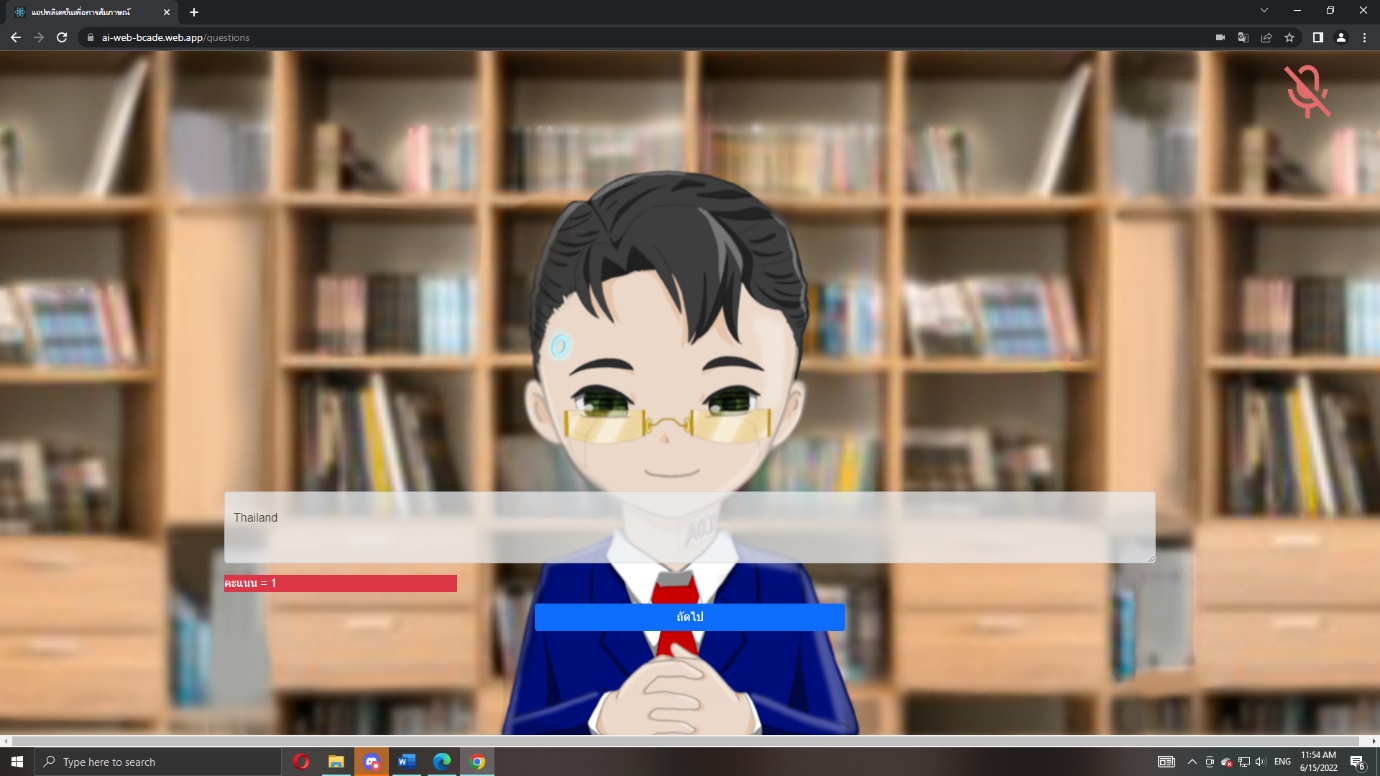
**4.1 ขั้นตอนการใช้งานแอปพลิเคชัน**

1. เข้าใช้งานผ่านเว็บไซต์บราวเซอร์ พิมพ์ค้นหาว่า <https://ai-web-bcade.web.app/>

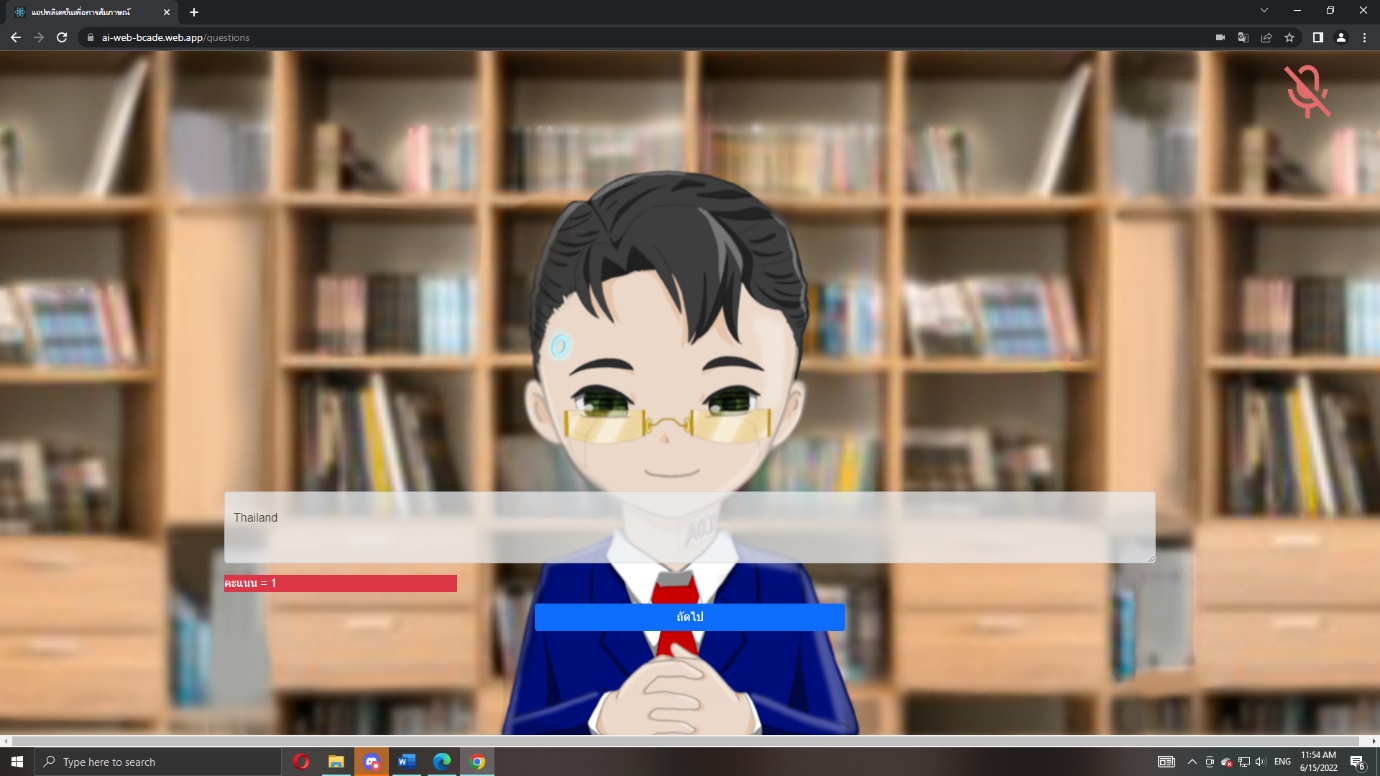
รูปที่ 4.1 เว็บแอปพลิเคชันเพื่อการสัมภาษณ์

2. เมื่อเข้าใช้งานเว็บแอปพลิเคชันจะพบหน้าแรกที่แสดงข้อความให้กรอกข้อมูล รหัสนักศึกษา ชื่อ - สกุล และ E-mail ให้ครบถ้วน แล้วพูดคำว่า “เริ่มสัมภาษณ์”

รูปที่ 4.2 หน้าแรกของแอพพลิเคชัน

 3. เมื่อกรอกข้อมูลครบถ้วนในหน้าแรกเสร็จแล้วจะเข้ามาสู่หน้าที่สอง โดยมีคำถามที่หนึ่งว่าสถานะครอบครัวเป็นยังไงครับ แก่ผู้สัมภาษณ์แต่ละคำถามและคำตอบขึ้นอยู่กับผู้สัมภาษณ์ตอบตรงกับคำถามก็ได้คะแนน 1 คะแนน และกดถัดไป

รูปที่ 4.3 เข้าสู่หน้าคำถามที่ 1 สถานะครอบครัวเป็นปัจจุบันเป็นอย่างไรครับ

 4. จากกดถัดไปจะมีคำถามที่สองว่าผู้สัมภาษณ์มีสัญชาติประเทศอะไรครับ และตอบตรงกับคำถามก็ได้คะแนน 1 คะแนน และกดถัดไป

รูปที่ 4.4 เข้าสู่หน้าคำถามที่ 2 ผู้สัมภาษณ์มีสัญชาติประเทศอะไรครับ

5. กดถัดไปจะมีคำถามที่สามว่าผู้สัมภาษณ์มีผลการเรียนเกรดเฉลี่ยเท่าไหร่ครับ และตอบตรงกับคำถามก็ได้คะแนน 1 คะแนน และกดถัดไป

รูปที่ 4.5 เข้าสู่หน้าคำถามที่ 3 ผลการเรียนเกรดเฉลี่ยเท่าไหร่ครับ

6. กดถัดไปจะมีคำถามที่สี่ว่าผู้สัมภาษณ์เคยสำเร็จการศึกษาจากที่มหาลัยใดมาก่อนหรือไม่ครับ และตอบตรงกับคำถามก็ได้คะแนน 1 คะแนน และกดถัดไป

รูปที่ 4.6 เข้าสู่หน้าคำถามที่ 4 เคยสำเร็จการศึกษาจากที่มหาลัยใดมาก่อนหรือไม่ครับ

7. กดถัดไปจะมีคำถามที่ห้าว่าผู้สัมภาษณ์มีครอบครัวมีรายได้ 1 ปี ทั้งหมดเท่าไหร่ครับ และตอบตรงกับคำถามก็ได้คะแนน 1 คะแนน และกดถัดไป

รูปที่ 4.7 เข้าสู่หน้าคำถามที่ 5 ครอบครัวมีรายได้ 1 ปี ทั้งหมดเท่าไหร่ครับ

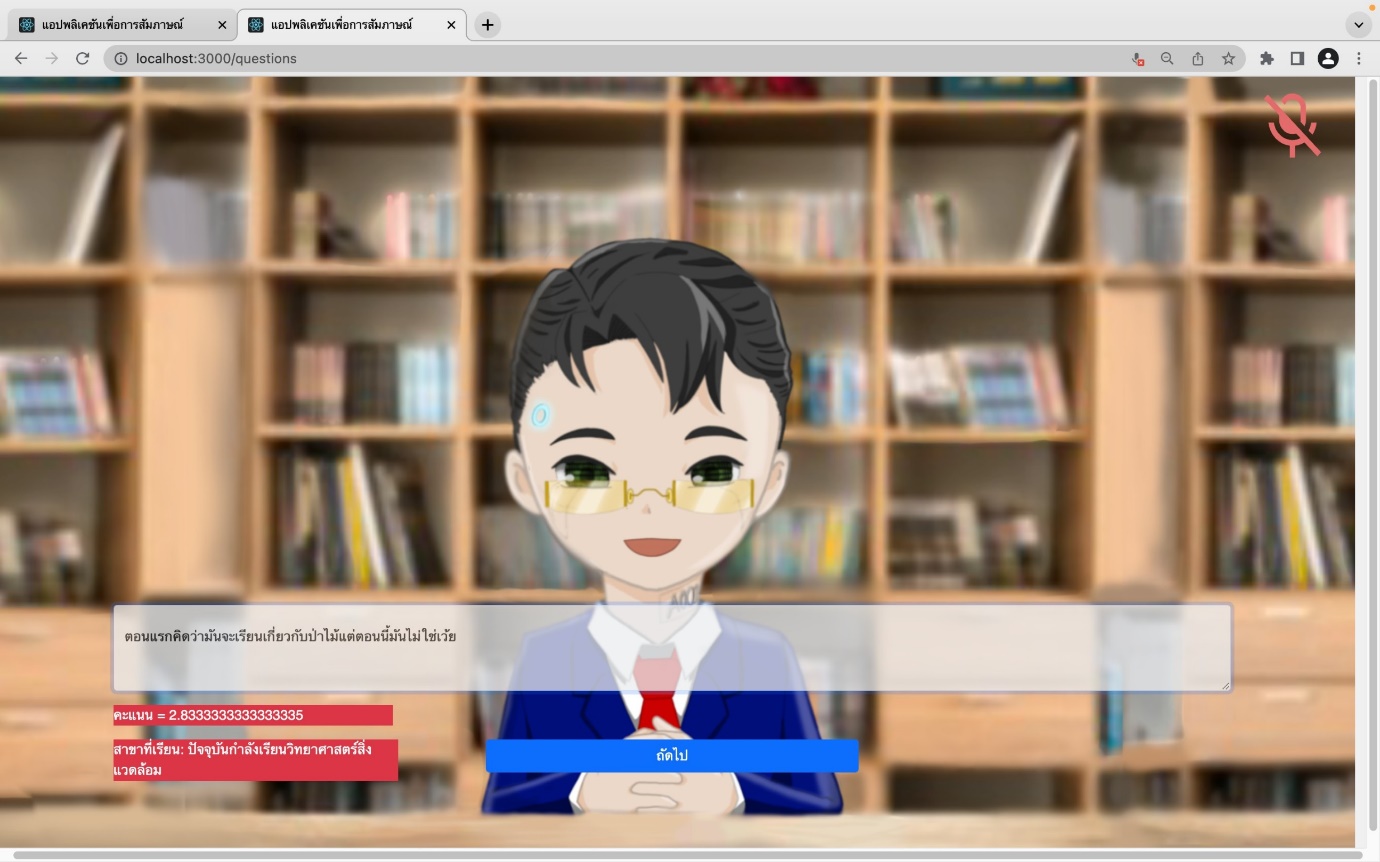
8. กดถัดไปจะมีคำถามที่หกว่าผู้สัมภาษณ์ได้ทำงานพิเศษ หรือไม่ ถ้าใช่อธิบายเกี่ยวกับงานของผู้สัมภาษณ์ครับ และตอบตรงกับคำถามก็ได้คะแนน 1 คะแนน และกดถัดไป

รูปที่ 4.8 เข้าสู่หน้าคำถามที่ 6 ผู้สัมภาษณ์ได้ทำงานพิเศษ หรือไม่

ถ้าใช่อธิบายเกี่ยวกับงานของผู้สัมภาษณ์ครับ

9. กดถัดไปจะมีคำถามที่เจ็ดว่าผู้สัมภาษณ์เรียนอยู่สาขาอะไรครับ และตอบตรงกับคำถามก็ได้คะแนน 1 คะแนน และกดถัดไป

รูปที่ 4.9 เข้าสู่หน้าคำถามที่ 7 ผู้สัมภาษณ์เรียนอยู่สาขาอะไรครับ

10.จากดถัดไปจะมีคำถามที่แปดว่าผู้สัมภาษณ์ทำไมถึงเลือกมาเรียนสาขา (ชื่อสาขา) ครับ ครับ และตอบตรงกับคำถามก็ได้คะแนน 1 คะแนน และกดถัดไป

รูปที่ 4.10 เข้าสู๋หน้าคำถามที่ 8 ทำไมถึงเลือกมาเรียนสาขา (ชื่อสาขา) ครับ

11.จากดถัดไปจะมีคำถามที่เก้าว่าท่านเคยทำประโยชน์ต่อสังคม / สาธารณะ อะไรบ้าง รวมทั้งสิ้นกี่ชั่วโมงครับ และตอบตรงกับคำถามก็ได้คะแนน 1 คะแนน และกดถัดไป

รูปที่ 4.11 เข้าสู่หน้าคำถามที่ 9 ท่านเคยทำประโยชน์ต่อสังคม / สาธารณะ อะไรบ้าง รวมทั้งสิ้นกี่ชั่วโมงครับ

11.จากดถัดไปจะมีคำถามที่สิบว่าผู้สัมภาษณ์มีจำนวนกี่คนในครอบครัว และตอบตรงกับคำถามก็ได้คะแนน 1 คะแนน และกดถัดไป

รูปที่ 4.11 เข้าสู๋หน้าคำถามที่ 10 ผู้สัมภาษณ์มีจำนวนกี่คนในครอบครัว

 12.ขั้นตอนสุดท้ายของการตอบแบบสัมภาษณ์แอพพลิเคชันคือการส่งอีเมลผลการสัมภาษณ์พร้อมทั้งเก็บผลสัมภาษณ์บันทึกลงระบบฐานข้อมูล

รูปที่ 4.12 หน้าสุดท้ายขั้นตอนส่งอีเมลผลการสัมภาษณ์

**4.2 ผลการประเมินแอพพลิเคชัน**

ผู้วิจัยได้ใช้แบบประเมินคะแนนสำหรับการสอบสัมภาษณ์ เพื่อจัดการคะแนนของผู้ถูกสัมภาษณ์อย่างถูกต้องตามมาตรฐาน อีกทั้งรองรับปัญหาข้างต้น และท้ายที่สุดนำผลที่ได้รับการประเมินเข้ายื่นคะแนนต่อกับทางมหาวิทยาลัยราชภัฏจันทรเกษมซึ่งมีผู้สัมภาษณ์ตอบแบบสอบถามรวมทั้งสิ้น 81 คน แบ่งเป็น ข้อมูลเพศ ข้อมูลอายุ ดังตารางที่ 4.1 ถึงตารางที่ 4.2 และสรุปผลคะแนนจาการตอบผู้สัมภาษณ์ ได้ดังตารางที่ 4.3

**ตารางที่ 4.1** ตารางสรุปเพศของผู้ตอบแบบสอบถาม

|  |  |  |
| --- | --- | --- |
| **เพศ** | **จำนวน (คน)** | **ร้อยละ** |
| หญิง | 34 | 41.98 |
| ชาย | 47 | 58.02 |
| **รวม** | 81 | 100 |

**ตารางที่ 4.2** ตารางสรุปช่วงอายุของผู้ตอบแบบสัมภาษณ์

|  |  |  |
| --- | --- | --- |
| **อายุ** | **จำนวน (คน)** | **ร้อยละ** |
| อายุ 19 - 21 ปี | 29 | 35.80 |
| อายุ 22 – 24 ปี | 34 | 41.98 |
| อายุ 24 – 27 ปี | 16 | 19.75 |
| อายุ 28 ปีขึ้นไป | 2 | 2.47 |
| **รวม** | 81 | 100 |

**ตารางที่ 4.3** สรุปผลคะแนนจาการสอบสัมภาษณ์

|  |  |  |  |
| --- | --- | --- | --- |
| **คะแนน** | **จำนวน (คน)** | **ร้อยละ** | **ผลของผู้สัมภาษณ์** |
| 0 - 48 คะแนน | 30 | 37.04 | ไม่ผ่าน |
| 48 คะแนนขึ้นไป | 51 | 62.96 | ผ่าน |
| **รวม** | 81 | 100 |  |

จากตารางที่ 4.1 พบว่ามีผู้สัมภาษณ์ตอบคำถามมากที่สุดเป็นเพศชาย จำนวน 47 คน คิดเป็นร้อยละ 58.02 และเพศหญิง จำนวน 34 คน คิดเป็นร้อยละ 41.98

จากตารางที่ 4.2 พบว่ามีผู้สัมภาษณ์ตอบคำถามที่มีช่วงอายุมากที่สุด คือ อายุ 22 – 24 ปี จำนวน 34 คน คิดเป็นร้อยละ 41.98 รองลงมาอยู่ในช่วงอายุ 19 – 21 ปี จำนวน 29 คน คิดเป็นร้อยละ 35.80 ช่วงอายุ 24 – 27 ปี จำนวน 16 คน คิดเป็นร้อยละ 19.75 และน้อยที่สุดในช่วงอายุ 28 ปีขึ้นไป จำนวน 2 คน คิดเป็นร้อยละ 2.47 ตามลำดับ

จากตารางที่ 4.3 สรุปผลคะแนนการสอบสัมภาษณ์ พบว่า ผลสอบของผู้สัมภาษณ์ที่ผ่านเกณฑ์จาการตอบคำถามจำนวน 51 คน ที่ได้คะแนนมากว่า 48 ขึ้นไป คิดเป็นร้อยละ 62.96 และผลสอบของผู้สัมภาษณ์ที่ไม่ผ่านเกณฑ์จากการตอบคำถามจำนวน 30 คน ที่ได้คะแนน 0 – 48 คะแนน คิดเป็นร้อยละ 37.04

**บทที่ 5**

**บทสรุป**

การพัฒนาเว็บแอปพลิเคชันปัญญาประดิษฐ์ตรวจวัดคำตอบ เพื่อประเมินการสอบสัมภาษณ์กองทุนเงินให้กู้ยืมเพื่อการศึกษา (กยศ.) มหาวิทยาลัยราชภัฏจันทรเกษม ผู้วิจัยได้ทำการเก็บรวบรวมข้อมูล และสรุปผลคะแนนจาการตอบสัมภาษณ์

5.1 สรุปผลการดำเนินงาน

5.2 ข้อจำกัดของการวิจัย

5.3 ข้อเสนอแนะ

**5.1 สรุปผลการดำเนินงาน**

5.1.1 เพื่อประเมินการสอบสัมภาษณ์กองทุนเงินให้กู้ยืมเพื่อการศึกษา (กยศ.)

การวิจัยเพื่อพัฒนาเว็บแอปพลิเคชันปัญญาประดิษฐ์ตรวจวัดคำตอบ เพื่อประเมินการสอบสัมภาษณ์กองทุนเงินให้กู้ยืมเพื่อการศึกษา (กยศ.) มหาวิทยาลัยราชภัฏจันทรเกษม โดยระบบที่ผู้วิจัยสร้างขึ้นมาเพื่อนำเทคโนโลยีปัญญาประดิษฐ์มาใช้สามารถตรวจจับคำตอบ คัดแยก จำแนกประเภท รวมถึงตัวเลขจำนวน และนำไปสู่การประเมินผลสอบของตัวผู้สัมภาษณ์ ยังช่วยลดการแพร่ระบาดของเชื้อโควิดวิด จากการเว้นระยะห่างทางสังคมในการเข้าคิวสอบสัมภาษณ์ เฟรมเวิร์ค(framework) ที่ใช้ได้แก่ รีแอค(react) บูทส์แตป (boost) ภาษาที่ใช้ได้แก่ ไทป์สคิป(Typescript) โหนดเจเอส(Node JS) อัลกอริทึมประญาประดิษฐ์ที่ใช้พัฒนาได้แก่ เทคทูสปีด(Take Two Speed) สปีดทูเทค(Speed ​​2 Tech) แทคซักเกตชัน(tactics) คอนดิชันนอลแรมดอมฟิล(Conventional Random fil) เครื่องมือที่ใช้พัฒนา ได้แก่ วิชัวสตูดิโอโค้ด(Visual Studio Code) ยานซีแอลไอ(CLI)

ในส่วนของการพัฒนาแอพพลิเคชันผู้วิจัยได้พัฒนาเว็บแอปพลิเคชันปัญญาประดิษฐ์ตรวจวัดคำตอบ ได้รับระบบประเมินผลการสอบสัมภาษณ์ และนำเทคโนโลยีปัญญาประดิษฐ์มาใช้ประโยชน์ร่วม โดยมีรายละเอียดังนี้ การประเมินผลการสอบสัมภาษณ์จากตรวจวัดคำตอบของผู้สัมภาษณ์ที่ตอบตรงกับคำถามในแอพพลิเคชันซึ่งคะแนนที่ได้รับแต่ละข้อจะขึ้นอยู่คำตอบของผู้สัมภาษณ์เท่านั้นและใช้เกณฑ์ของการประเมินการสอบสัมภาษณ์ของกองทุนเงินให้กู้ยืมเพื่อการศึกษา (กยศ.) ของทางมหาวิทยาลัยราชภัฏจันทรเกษมที่จะกำหนดเกณฑ์จากผู้สัมภาษณ์ได้ตอบทางอีเมลว่าข้อมูลของผู้สัมภาษร์ผ่านหรือไม่ผ่านในการสอบสัมภาษณ์โดยมีเกณฑ์อยู๋ที่ระดับ 48 คะแนนขึ้นไป

5.1.2 คะแนนการเมินของผู้สัมภาษณ์ต่อเว็บแอพพลิเคชันปัญญาประดิษฐ์ตรวจวัดคำตอบ เพื่อประเมินการสอบสัมภาษณ์กองทุนเงินให้กู้ยืมเพื่อการศึกษา (กยศ.)

จากการพัฒนาแอพพลิเคชันเว็บแอปพลิเคชันปัญญาประดิษฐ์ตรวจวัดคำตอบ ได้รับระบบประเมินผลการสอบสัมภาษณ์ มีผู้สอบสัมภาษณ์ทั้งสิ้น 81 คน พบว่า ผู้สอบสัมภาษณ์ส่วนใหญ่ผ่านการสอบสัมภาษณ์ของเว็บแอพพลิเคชันอยู่ในระดับคะแนน 48 คะแนนขึ้นไป จำนวน 51 คน ทำให้เห็นว่าผู้สอบสัมภาษณ์ใช้งานรู้สึกว่าเป็นแอพพลิเคชันที่น่าสนใจเหมือนกับการได้พูดคุยกับคนเป็นๆและใช้งานง่าย วิธีการสอบไม่ซับซ้อนยุ่งยากและมีคะแนนจากการตอบคำถามบอกทุกครั้งเพื่อให้ผู้สอบสามารถเปลี่ยนคำตอบได้เพื่อได้คะแนนจาการสอบสัมภาษณ์มากขึ้น

**5.2 ข้อจำกัดของการวิจัย**

การพัฒนาเว็บแอปพลิเคชันปัญญาประดิษฐ์ มีข้อจำกัดดังนี้

5.2.1 เว็บแอปพลิเคชันสามารถใช้ได้ด้วยการพูดด้วยเสียงเท่านั้น

5.2.2 ต้องตอบให้ตรงคำถามที่ผู้สัมภาษณ์ได้ถามมากที่สุดเพื่อที่จะได้รับคะแนนถึงจะผ่านเกณฑ์ทางเว็บแอปพลิเคชันกำหนด

**5.3 ข้อเสนอแนะ**

5.3.1 ควรมีการเพิ่มเติมในแอปพลิเคชันของรูปแบบระบบปฏิบัติการด้วยการที่พิมพ์เพิ่มเข้าไปในระบบเผื่อบางคนไม่สามารถสื่อสารหรือพูดได้

**บรรณานุกรม**

**ประวัติผู้วิจัย**

ชื่อ-นามสกุล นายศุภณัฐ ชินราช

รหัสประจำตัวนักศึกษา 6111500143

วัน เดือน ปี 10 มิถุนายน 2542

ประวัติการศึกษา ระดับประถมศึกษา ศึกษาที่ โรงเรียน

ระดับมัธยมศึกษา ศึกษาที่ โรงเรียน

ปีที่เข้าศึกษา 2561

อีเมล suphanat.ch61@chandra.ac.th