**ระบบแชทบอทแนะนำการศึกษาต่อในคณะเทคโนโลยีสารสนเทศ**

**สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง**

**IT KMITL Admission Chatbot**

**โดย**

**พีรดา มั่นศิลป์**

**PEERADA MUNSILP**

**วิลดา สุขกล่ำ**

**WINLADA SOOKGLUM**

**ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรบัณฑิต**

**สาขาวิชาวิทยาการข้อมูลและการวิเคราะห์เชิงธุรกิจ**

**คณะเทคโนโลยีสารสนเทศ**

**สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง**

**ภาคเรียนที่ 1 ปีการศึกษา 2563**

**ระบบแชทบอทแนะนำการศึกษาต่อในคณะเทคโนโลยีสารสนเทศ**

**สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง**

**(IT KMITL Admission Chatbot)**

**โดย**

**พีรดา มั่นศิลป์  
PEERADA MUNSILP**

**วิลดา สุขกล่ำ  
WINLADA SOOKGLUM**

**อาจารย์ที่ปรึกษา**

**ผู้ช่วยศาสตราจารย์ ดร.บัณฑิต ฐานะโสภณ  
รองศาสตราจารย์ ดร.พรฤดี เนติโสภากุล**

**ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรบัณฑิต**

**สาขาวิชาวิทยาการข้อมูลและการวิเคราะห์เชิงธุรกิจ**

**คณะเทคโนโลยีสารสนเทศ**

**สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง**

**ภาคเรียนที่ 1 ปีการศึกษา 2563**

**IT KMITL ADMISSION CHATBOT**

**PEERADA MUNSILP**

**WINLADA SOOKGLUM**

**A PROJECT SUBMITTED IN PARTIAL FULFILLMENT**

**OF THE REQUIREMENT FOR THE DEGREE OF**

**BACHELOR OF SCIENCE PROGRAM IN INFORMATION TECHNOLOGY**

**FACULTY OF INFORMATION TECNOLOGY**

**KING MONGKUT’S INSTITUTE OF TECHNOLOGY LADKRABANG**

**1/2020**

**COPYRIGHT 2020**

**FACULTY OF INFORMATION TECHNOLOGY**

**KING MONGKUT’S INSTITUTE OF TECHNOLOGY LADKRABANG**

**ใบรับรองปริญญานิพนธ์ ประจำปีการศึกษา 2563**

**คณะเทคโนโลยีสารสนเทศ**

**สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง**

**เรื่อง** **ระบบแชทบอทแนะนำการศึกษาต่อในคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง**

**IT KMITL ADMISSION CHATBOT**

**ผู้จัดทำ**

**1.  นางสาวพีรดา มั่นศิลป์ รหัสนักศึกษา 60070153**

**2.  นางสาววิลดา สุขกล่ำ รหัสนักศึกษา 60070160**

**........................................อาจารย์ที่ปรึกษา**

(.......................................)

**.........................................อาจารย์ที่ปรึกษาร่วม**

(.......................................)

**ใบรับรองโครงงาน (PROJECT)**

**เรื่อง**

**ระบบแชทบอทแนะนำการศึกษาต่อในคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง IT KMITL ADMISSION CHATBOT**

**นางสาวพีรดา มั่นศิลป์ รหัสประจำตัว 60070153**

**นางสาววิลดา สุขกล่ำ รหัสประจำตัว 60070160**

ขอรับรองว่ารายงานฉบับนี้ ข้าพเจ้าไม่ได้คัดลอกมาจากที่ใด

รายงานฉบับนี้ได้รับการตรวจสอบและอนุมัติให้เป็นส่วนหนึ่งของ

การศึกษาวิชาโครงงาน หลักสูตรวิทยาศาสตรบัณฑิต (เทคโนโลยีสารสนเทศ)

ภาคเรียนที่ 1 ปีการศึกษา 2563

......…...……….……...……….

        (นางสาวพีรดา มั่นศิลป์)

......…...……….……...……….

        (นางสาววิลดา สุขกล่ำ)

**หัวข้อโครงงาน** ระบบแชทบอทแนะนำการศึกษาต่อในคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

**นักศึกษา**  นางสาวพีรดา มั่นศิลป์ รหัสนักศึกษา 60070153

นางสาววิลดา สุขกล่ำ รหัสนักศึกษา 60070160

**ปริญญา**  วิทยาศาสตรบัณฑิต

**สาขาวิชา**  วิทยาการข้อมูลและการวิเคราะห์เชิงธุรกิจ

**ปีการศึกษา** 2563

**อาจารย์ที่ปรึกษา** ผู้ช่วยศาสตราจารย์ ดร.บัณฑิต ฐานะโสภณ

รองศาสตราจารย์ ดร.พรฤดี เนติโสภากุล

**บทคัดย่อ**

การค้นหาตัวเองของนักเรียนชั้นมัธยมศึกษาตอนปลาย ว่าจะเรียนต่อด้านไหนในระดับอุดมศึกษา ยังคงเป็นปัญหาสำหรับนักเรียนบางส่วน ที่กำลังลังเลและตัดสินใจไม่ได้ จึงทำให้เกิดความกังวลแก่นักเรียน ในงานวิจัยนี้ทางคณะผู้จัดทำได้ศึกษาวิธีการต่างๆ ที่จะช่วยให้นักเรียนที่สนใจจะศึกษาต่อในคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ได้ตัดสินใจง่ายขึ้นว่านักเรียนคนนั้นเหมาะกับการเรียนคณะดังกล่าวมากน้อยเพียงใด โดยศึกษาและเก็บข้อมูลจากกลุ่มตัวอย่างที่เป็นรุ่นพี่ในคณะ มีทั้งกลุ่มคนที่ชอบเรียนและไม่ชอบเรียนทางด้านคอมพิวเตอร์ เช่น เหตุผลในการอยากเข้าคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง วิชาที่ชอบเรียนตอนมัธยมศึกษาตอนปลาย พื้นฐานการเขียนโปรแกรม เป็นต้น ปัจจัยเหล่านี้สามารส่งผลต่อความสุขในการเรียนได้ อันเนื่องมาจากพื้นฐานความชอบและความสนใจของแต่ละบุคคล นอกจากนี้ยังมีบทความประสบการณ์และการแนะแนวต่างๆของรุ่นพี่ในคณะ ไม่ว่าจะเป็นเรื่องเรียน เรื่องกิจกรรม หรือแม้แต่สังคมและสภาพแวดล้อม เพื่อให้นักเรียนสามารถประเมินตัวเองได้ง่ายขึ้นว่าเหมาะกับการเรียนคณะนี้มากน้อยแค่ไหน งานวิจัยนี้จะช่วยให้ผู้ใช้งานสามารถคลายความลังเลและสามารถตัดสินใจได้ง่ายขึ้นในการศึกษาต่อคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง อีกทั้งยังช่วยลดโอกาสที่นักศึกษาจะลาออกเพื่อไปเข้าคณะใหม่ ซึ่งจะทำให้เสียเวลาได้อีกด้วย

**Project Title** IT KMITL Admission Chatbot

**Student** Ms. Peerada Munsilp Student ID 60070153

Ms. Winlada Sookglum Student ID 60070160

**Degree** Bachelor of Science

**Program** Information Technology

**Academic Year** 2020

**Advisor**  Asst. Prof. Dr. Bundit Thanasopon  
 Assoc. Prof. Dr. Ponrudee Netisopakul

**ABSTRACT**

The self-search of a high school student about what area to study in higher education is still

be a problem for some students who are hesitating and cannot decide that causing concern for

students.

In this research, the organizers studied about methods that will help students who interested

to study in the Faculty of Information Technology (KMITL), make easier decide how well the student

would be suitable for studying in the faculty. By studying and collecting data from a sample group of

seniors in the faculty. There are both groups of people who like and do not like to study in the

computer field. The example of questions are, the reason why they decided to study in the Faculty of

Information Technology (KMITL), subjects they like to study in high school, programming skill or

experience, etc. These factors can affect the happiness of studying. Based on individual preferences

and interests. In addition, there are articles about experiences and guidance of the seniors in the

faculty such as studying in the faculty, activities in the faculty or society and environment. To make it

easier for students to assess themselves how well they are suitable for studying in this faculty.

This research will help users alleviate their hesitations and make decisions more easily in

their studies at the the Faculty of Information Technology (KMITL). It also reduces the chance of

students leaving to enter a new faculty or university, which will also waste of time.

**กิตติกรรมประกาศ**

โครงงานฉบับนี้สำเร็จสมบูรณ์ได้ด้วยดีเพราะได้รับความกรุณาชี้แนะและช่วยเหลืออย่างดียิ่งจาก ผู้ช่วยศาสตราจารย์ ดร.บัณฑิต ฐานะโสภณ อาจารย์ที่ปรึกษา และรองศาสตราจารย์ ดร.พรฤดี เนติโสภากุล อาจารย์ที่ปรึกษาร่วม ที่ให้คำแนะนำ รวมถึงแก้ไขตรวจสอบข้อบกพร่องมาโดยตลอดจนสำเร็จ ผู้วิจัยขอกราบขอบพระคุณด้วยความเคารพอย่างสูงไว้ ณ โอกาสนี้

กราบขอบพระคุณคณาจารย์ คณะเทคโนโลยีสารสนเทส สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบังทุกท่าน ที่สั่งสอนและมอบความรู้ต่างๆ ที่เป็นประโยชน์ในการทำโครงงานฉบับนี้ จนสำเร็จลุล่วงด้วยดี

ขอขอบคุณเพื่อนๆ รุ่นพี่ และรุ่นน้อง ที่ได้มีส่วนเกี่ยวข้องในการชี้แนะ ให้คำปรึกษา และให้การช่วยเหลือผู้วิจัยในการทำโครงงานฉบับนี้

สุดท้ายนี้ กราบขอบพระคุณบิดามารดาและครอบครัวเป็นอย่างยิ่ง ที่คอยสนับสนุนและให้กำลังใจผู้วิจัยอย่างดีเสมอมา

พีรดา มั่นศิลป์

วิลดา สุขกล่ำ

**บทที่ 1**

**บทนำ**

**1.1 ความเป็นมาและความสำคัญของปัญหา**

สืบเนื่องจากผู้วิจัยได้มองเห็นถึงปัญหาของนักเรียนชั้นมัธยมศึกษาตอนปลายและนักศึกษา

บางส่วนในการศึกษาต่อในระดับอุดมศึกษา ที่เกิดความลังเลใจในการศึกษาต่อ เนื่องจากไม่รู้ว่าชอบ

หรือไม่ชอบอะไร ทำให้ตัดสินใจยาก เมื่อตัดสินใจผิดพลาดก็จะทำให้เสียเวลาและหมดกำลังใจ ส่งผล

ให้เกิดความเครียด ซึ่งความเครียดของนักเรียนหรือนักศึกษาที่เกี่ยวกับการศึกษาต่อนั้น สามารถส่งผล

ต่อด้านร่างกาย จิตใจ พฤติกรรมและอารมณ์ของนักศึกษา ซึ่งการศึกษาในระดับอุดมศึกษาเป็นช่วงวัยที่

นักศึกษาต้องมีการปรับตัวให้เหมาะสมกับสภาพแวดล้อมด้านต่างๆ ไม่ว่าจะเป็นด้านสังคม หรือด้าน

การเรียน

ในงานวิจัยชิ้นนี้ทางคณะผู้จัดทำได้จัดทำระบบแชทบอทแนะนำการศึกษาต่อในคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง เพื่อช่วยให้นักเรียนที่สนใจจะศึกษาต่อในคณะดังกล่าวได้ตัดสินใจง่ายขึ้นว่านักเรียนคนนั้นเหมาะกับการเรียนคณะดังกล่าวมากน้อยเพียงใด โดยศึกษาและเก็บขอ้มูลจากกลุ่มตัวอย่างที่เป็นรุ่นพี่ในคณะ มีทั้งกลุ่มคนที่ชอบเรียนและไม่ชอบเรียนทางด้านคอมพิวเตอร์ ซึ่งหลายๆปัจจัยสามารถส่งผลต่อความสุขในการเรียนได้ อัน

เนื่องมาจากพื้นฐานความชอบและความสนใจของแต่ละบุคคลเพื่อให้นักเรียนสามารถประเมินตัวเอง

ได้ง่ายขึ้นว่าเหมาะกับการเรียนคณะนี้มากน้อยแค่ไหน

**1.2 จุดมุ่งหมายและจุดประสงค์ของการพัฒนา**

การจัดทำระบบแชทบอทแนะนำการศึกษาต่อในคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง มีจุดมุ่งหมายและจุดประสงค์ดังต่อไปนี้

1.2.1 เพื่อนำความรู้ที่ได้จากการศึกษามาประยุกต์ใช้ เช่น Natural Language Processing (NLP) เพื่อใช้ในการประมวลผลและวิเคราะห์ประโยคที่ได้รับมา Machine Learning เพื่อสร้างโมเดลสำหรับทดสอบค่าความแม่นยำของการทำนายข้อมูลที่ได้รับมา

1.2.2 เพื่อเป็นแนวทางการตัดสินใจศึกษาต่อในคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ให้แก่น้องๆชั้นมัธยมศึกษา และบุคคลทั่วไปที่สนใจ

1.2.3 เพื่อเป็นแนวทางการใช้ชีวิตการเรียนแก่นักศึกษาคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

**1.3 ขอบเขตโครงงาน**

การจัดทำระบบแชทบอทแนะนำการศึกษาต่อในคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง จะต้องมีการกำหนดขอบเขตของงานดังนี้

1.3.1 กลุ่มเป้าหมายของผู้ใช้บริการระบบคือ ผู้ที่สนใจจะศึกษาต่อในคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ผู้ปกครองที่มีความต้องการทราบถึงรายละเอียดต่างๆเกี่ยวกับคณะ และนักศึกษาคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

1.3.2 ระบบที่ใช้ในการติดต่อสื่อสารกับนักศึกษาคือ Line Official

**1.4 ขั้นตอนการออกแบบและพัฒนาระบบ**

เพื่อให้การจัดทำระบบแชทบอทแนะนำการศึกษาต่อในคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง ดำเนินไปอย่างมีระเบียบแบบแผน จึงจำเป็นที่จะต้องมีการวางแผนขั้นตอนการดำเนินงานและพัฒนาระบบดังต่อไปนี้

* + 1. กำหนดปัญหา ที่มาและวัตถุประสงค์ของโครงการ
    2. กำหนดขอบเขตของโครงงาน
    3. ศึกษาเครื่องมือที่เกี่ยวข้อง
    4. ศึกษาทฤษฎี ข้อมูลงานวิจัยที่เกี่ยวข้องกับความเครียด ปัจจัยที่ทำให้เกิดความเครียดของนักศึกษา และการบรรเทาความเครียด
    5. เก็บข้อมูลจากกลุ่มตัวอย่างซึ่งเป็นนักศึกษาคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
    6. การพัฒนาระบบ

1. พัฒนา algorithm ต่างๆที่ใช้ในการวิเคราะห์ เช่น การตรวจจับคำ
2. พัฒนา model ในการเรียนรู้ของ bot
3. นำ Algorithm และ Model ที่พัฒนาไว้เชื่อมต่อกับ ระบบ Line Official ด้วย Heroku
4. นำ Intent และ รูปแบบการตอบกลับใส่ใน Dialogflow
   * 1. ทดสอบระบบ
5. ทดสอบว่า Chatbot สามารถตอบคำถามที่สอดคล้องกับข้อความที่รับเข้ามา
6. ทดสอบว่าข้อมูลที่สื่อสารกันระหว่างนักศึกษาและ Chatbot จะสามารถเก็บเข้าฐานข้อมูลได้
   * 1. ประเมินผล
     2. เรียบเรียงเอกสารประกอบโครงงาน

**1.5 ประโยชน์ที่คาดว่าจะได้รับ**

ผู้วิจัยได้จัดทำระบบแชทบอทแนะนำการศึกษาต่อในคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบังขึ้น โดยมีเป้าหมายให้ระบบสามารถนำไปใช้ประโยชน์ได้ดังต่อไปนี้

* + 1. ช่วยเป็นแนวทางการตัดสินใจศึกษาต่อในคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
    2. เพื่อให้นักเรียนผู้ที่สนใจเข้าศึกษาต่อคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง หรือ บุคคลทั่วไป ได้รู้จักคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบังมากขึ้น
    3. ช่วยในการค้นหาตัวเองว่าชอบและอยากเข้าคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง จริงๆหรือไม่

**บทที่ 2**

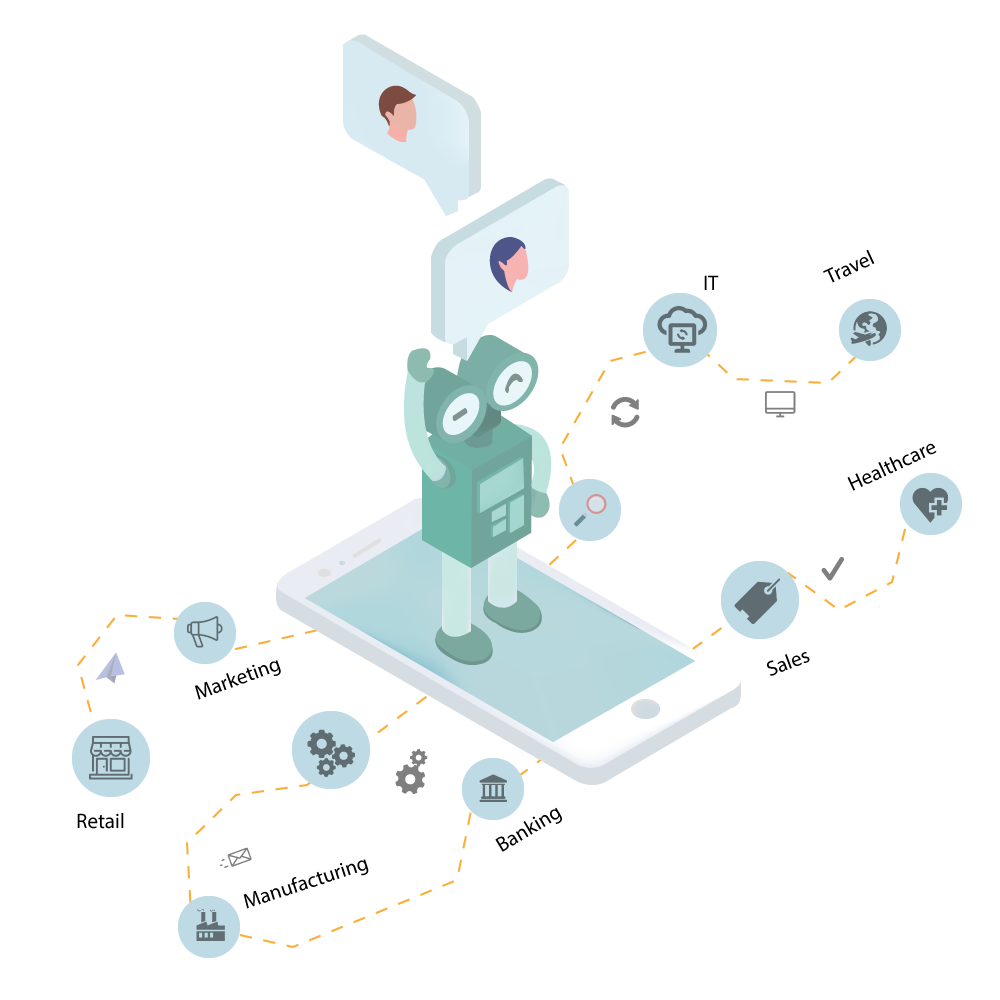
**แนวคิด ทฤษฎี และงานวิจัยที่เกี่ยวข้อง**

**2.1 แนวคิดเกี่ยวกับ Chatbot**

**2.1.1 ความหมายของ Chatbot**

Chatbot คือโปรแกรมคอมพิวเตอร์ชนิดหนึ่ง ถูกพัฒนาขึ้นมาให้มีบทบาทในการตอบกลับการสนทนาด้วยตัวอักษรแบบอัตโนมัติ เป็นการจำลองบทสนทนาของมนุษย์แบบ real-time ปัจจุบันแชทบอทเป็นหนึ่งในตัวช่วยขององค์กรต่างๆ ซึ่งจะช่วยตอบข้อความ ตอบคำถาม หรือให้ข้อมูล ความช่วยเหลือ ความบันเทิง และอื่นๆ ขึ้นอยู่กับการตั้งค่าของผู้พัฒนา [1][2] โดย Chat Bot นั้น ได้มีการพัฒนาขึ้นมา 3 แบบ

1. Rule-Based approach คือ แนวทางการพัฒนา ChatBot เป็นการโต้ตอบกับผู้ใช้แบบมีเงื่อนไข ที่ผู้พัฒนาได้กำหนดเอาไว้ [3]
2. Intellectually Independent Chatbots คือ แนวทางการพัฒนา ChatBot โดยใช้ Machine Learning ซึ่งช่วยให้ Chatbot เรียนรู้จากอินพุตของผู้ใช้ ทำให้เกิดการตอบกลับและเรียนรู้ด้วยตัวเองให้เข้าใจคำถามของผู้ใช้งานมากขึ้น [4]
3. AI-Based Approach คือ แนวทางการพัฒนา ChatBot โดยใช้ปัญญาประดิษฐ์ (AI) และ Machine Learning เข้ามาช่วยให้ ChatBot นั้นฉลาดและเข้าใจภาษาของมนุษย์มากขึ้น ให้ความสำคัญกับความตั้งใจ (Intent) และบริบท (Context) ซึ่งจำเป็นต้องใช้เทคโนโลยีอย่าง National Language Processing เข้ามาช่วย โดยการพัฒนาแบบ AI-Based Approach จะได้รับความนิยมมาก หลายบริษัทได้นำไปพัฒนา เช่น IBM, Microsoft, Google, Facebook, Amazon เป็นต้น [3]



รูปที่ 2.1 แสดงการประยุกต์ใช้แชทบอทในอุตสาหกรรมต่างๆ  
(**ที่มา**: <https://smartsensesolutions.com/expert-data-scientists>)

**2.1.2 ความเป็นมาและพัฒนาการ**

เริ่มต้นจาก THE TURING TEST โดย Alan Turing ได้สร้างทฤษฎีที่มีชื่อว่าเครื่องอัจฉริยะอย่างแท้จริง (Truly Intelligent Machine) โดยสิ่งนี้จะกลายมาเป็นเรื่องที่ผู้คนจะแยกออกได้ยากว่า บทสนทนาข้อความที่เกิดขึ้นนั้นเป็นของมนุษย์หรือว่าหุ่นยนต์ ซึ่งความคิดเหล่านี้ได้ถูกพัฒนาต่อเนื่องจนกลายมาเป็นรากฐานสำคัญของการพัฒนาแชทบอท [5]  
 ปีคริสตศักราช 1966 บอทตัวแรกของโลกคือ ELIZA ซึ่งถูกสร้างขึ้นเพื่อจำลองการสนทนาของมนุษย์ โดยมีการตั้งค่าคำตอบที่เหมาะสมกับคู่สนทนาในรูปแบบถามมาตอบไป [4][5]

ต่อมาในปีคริสตศักราช 1980 JABBERWACKY ได้ถูกพัฒนาขึ้น ซึ่งแชทบอทตัวนี้ถูกออกแบบมาเพื่อจำลองการพูดคุยของมนุษย์ตามธรรมชาติ [4][5]

ต่อมาในปีคริสตศักราช 1995 A.L.I.C.E. แชทบอทออนไลน์ยอดนิยม เป็นบอทประมวลผลทางภาษาเป็นระบบคอมพิวเตอร์ ที่เชื่อมต่อกับอินเทอร์เน็ตภาษาศาสตร์ประดิษฐ์ (Artificial Linguistic) เป็นกระบวนการประมวลผลภาษาธรรมชาติของบอท สามารถนำรูปแบบการใช้กฎของการเรียนรู้ด้วยตัวเองไปใช้กับมนุษย์ได้เพื่อที่จะได้มีการเข้าใจบทสนทนามากขึ้น [4][5]

**2.2 แนวคิดเกี่ยวกับความเครียด**

**2.2.1 ความหมายของความเครียด**

ความเครียด (stress) คือสภาวะทางอารมณ์ความรู้สึก ที่ถูกกดดัน บีบคั้น ซึ่งในแต่ละบุคคลจะมีวิธีในการปรับตัวให้ผ่านไปได้ด้วยตนเอง แต่สำหรับในบางสถานการณ์ที่ไม่สามารถหาทางออกจากสถานการณ์ที่ตึงเครียดนั้นได้ จึงทำให้ส่งผลกระทบต่อการใช้ชีวิตประจำวัน ทั้งของตนเองและผู้คนรอบข้าง และเมื่อมนุษย์ต้องเจอสถานการณ์ตึงเครียดโดยที่หาทางออกไม่ได้บ่อยครั้ง อาจทำให้ตกอยู่ใน ภาวะซึมเศร้า หรือ เป็นโรควิตกกังวล ได้ เป็นเหตุให้ต้องเข้าพบจิตแพทย์เพื่อรับการรักษาอย่างถูกวิธี แต่ในปัจจุบันนี้ด้วยสภาพของเศรษฐกิจ และสังคม ที่ต้องดิ้นรนหาเงินเลี้ยงชีพ โดยอาจได้รับหรือไม่ได้รับผลกระทบจากสถานการณ์โควิด-19 ทำให้ยิ่งเกิดความเครียดที่สะสม หาทางออกไม่ได้ จนนำไปสู่การเกิดเหตุการณ์ที่ไม่คาดฝันขึ้นในที่สุด [7]

**2.2.2 การสังเกตความเครียดเบื้องต้น**

1. นอนไม่หลับ  ความเครียดอาจส่งผลให้นอนไม่หลับ ถ้าหากนอนไม่หลับเป็นเวลานาน จะส่งผลต่อสุขภาพกายและใจโดยรวมจนเริ่มเข้าสูภาาวะซึมเศร้า หรือความเครียดรุนแรง
2. พฤติกรรมเปลี่ยนไป ไม่ร่าเริง ไม่พูดคุย  นิ่งเงียบ เบื่อหน่าย และปิดกั้นตัวเอง
3. เศร้าหมอง ไม่มีความสุข วิตกกังวลกับเรื่องต่างๆ จนแสดงออกทางสีหน้า ท่าทาง หรือคำพูด
4. ความเครียดอาจแสดงออกเป็นอาการทางกาย เช่น หายใจถี่ขึ้น หรือ กลั้นหายใจโดยไม่รู้สึกตัว  ปวดท้อง อาเจียน ปวดศีรษะ เป็นต้น
5. ในกรณีรุนแรง ผู้ป่วยอาจจะพูดว่าอยากตาย บางครั้งอาจฟังเหมือนเป็นการพูดเล่น ดังนั้นจึงควรใส่ใจผู้พูดให้มากขึ้น ไม่ว่าจะพบการตัดพ้อเช่นนี้ ในโลกโซเชียลหรือพูดขึ้นลอยๆ [7]



รูปที่ 2.2 สาเหตุของการเกิดความเครียดต่างๆ

(**ที่มา**: <http://memory-alertide.lnwshop.com/>)

**2.2.3 การรักษาอาการเครียด**

หากความเครียดรบกวนการใช้ชีวิตประจำวัน รวมถึงกระทบต่อการทำงาน หรือมีผลต่อผู้อื่น การพบจิตแพทย์เพื่อปรึกษาและรักษาอย่างถูกวิธีเป็นสิ่งที่ควรทำอย่างยิ่ง โดยจิตแพทย์จะทำการรักษาโดย

1. แพทย์พูดคุยซักประวัติและตรวจร่างกายเพื่อวินิจฉัยทางการแพทย์และหาสาเหตุของความเครียด
2. ในกรณีความเครียดส่งผลทางกาย เช่น นอนไม่หลับ ปวดท้อง หรือปวดศีรษะ จิตแพทย์อาจให้รับประทานยาเพื่อบรรเทาอาการ
3. การให้คำปรึกษาโดยจิตแพทย์ ผู้เชี่ยวชาญเฉพาะทาง ด้วยการสื่อสารให้เกิดความเข้าใจและหาสาเหตุของปัญหา ชี้แนะอย่างถูกวิธีเพื่อคลายความเครียด
4. จิตบำบัด ซึ่งต้องได้รับการบำบัดโดยจิตแพทย์ [7]

**2.3 แนวคิดเกี่ยวกับการบรรเทาความเครียด**

**2.3.1 การนำความคิดเชิงบวกมาใช้**

ความคิดเชิงบวกเป็นเครื่องมือที่มีอิทธิพลในการเปลี่ยนความคิดเชิงลบ ให้กลายเป็นเชิงบวกที่สามารถสร้างแรงบันดาลใจ ตัวอย่างของความคิดเชิงบวกคือการยืนยัน การยืนยันที่สามารถใช้ได้ที่ซึ่งได้รับการออกแบบมาเพื่อลดความเครียดและเพิ่มพลังให้กับตนเอง อีกทั้งยังช่วยเติมเต็มศักยภาพให้กับตนเองอีกด้วย ยกตัวอย่างประโยคของการยืนยันเช่น ฉันใจเย็น ฉันผ่อนคลาย ฉันให้อภัยผู้อื่น ฉันชอบออกกำลังกาย ฉันสามารถทำสิ่งใดๆให้สำเร็จได้ ฉันมีความสุข เป็นต้น [8]

**2.3.2 การเปลี่ยนมุมมองของร่างกาย ความคิด และความรู้สึก**

นักวิจัยทางด้านความเครียดพบว่า ร่างกาย ความคิด และความรู้สึกของมนุษย์ สามารถเปลี่ยนแปลงได้ โดยที่สภาวะทางจิตใจมีผลกระทบอย่างมีนัยสำคัญต่อสภาวะทางกายภาพ ยกตัวอย่างเช่น ถ้าสมองเชื่อว่าสิ่งต่างๆกำลังท่วมท้นและไม่สามารถรับมือกับความเครียดได้ มนุษย์จะโน้มน้าวตนเองไปสู่มุมมองนั้น และยิ่งทำให้สถานการณ์ยากขึ้นด้วยต่อการจัดการ ตัวอย่างวิธีเปลี่ยนมุมมองของร่างกาย ความคิด และความรู้สึก [9]

1. การขจัดความเครียดด้วยรอยยิ้ม

พลังของรอยยิ้มอันเป็นผลต่อความรู้สึก จะสังเกตเห็นว่าในขณะที่ยิ้ม จะทำให้รู้สึกมีความสุขมากขึ้นโดยอัตโนมัติ เห็นถึงการเปลี่ยนแปลงของความรู้สึก และเมื่อใดที่พบว่าตนเองรู้สึกเศร้าหรือเสียใจ จงบังคับให้ตนเองยิ้มแล้วจะเห็นการเปลี่ยนแปลงทางด้านความคิดและการเคลื่อนไหว เพราะวิธีที่เร็วที่สุดในการเปลี่ยนทางจิตวิทยาคือการเปลี่ยนสรีรวิทยา [9]

**2.3.3 หายใจเข้าลึกๆ**

สิ่งที่ต้องทำเพื่อการมีชีวิตอยู่ของมนุษย์ 3 สิ่งที่สำคัญที่สุดคือ การหายใจ การดื่มน้ำ และการกินอาหาร โดยปกติมนุษย์สามารถอยู่ได้เป็นเวลาหลายสัปดาห์โดยที่ไม่กินอาหาร และไม่ดื่มน้ำ แต่สามารถอยู่ได้เพียงไม่กี่นาทีโดยไม่หายใจ การทำให้ร่างกายมั่นใจว่ามีปริมาณของออกซิเจนมากพอเป็นสิ่งที่จำเป็น เพื่อให้สามารถรับมือกับความเครียดที่ต้องพบเจอในชีวิตประจำวันได้ดีมากขึ้น เนื่องจากทุกเซลล์ในร่างกายต้องการออกซิเจนในการทำงานอย่างถูกต้อง เมื่อใดที่เผชิญกับสถานการณ์ความตึงเครียดไม่ว่าทางใดทางนึง ให้ลองถอยกลับมาหายใจเข้าลึกๆเพื่อประเมินสถานการณ์และคิดวิธีการแก้ปัญหาได้อย่างมีประสิทธิภาพ วิธีการทดสอบว่าความเครียดส่งผลต่อการหายใจหรือไม่ด้วยการตอบคำถามดังต่อไปนี้  
 1) เคยเป็นโรคลมหายใจ หรือ หายใจถี่หรือไม่?  
 2) พบว่าตนเองมักหายใจสั้นๆ อย่างรวดเร็วหรือไม่?  
 3) รู้สึกหายใจไม่ออกหรือไม่?  
 4) เคยทรมานจากความคิดที่คลุมเครือ หรือ มีหมอกในสมองหรือไม่?  
 5) รู้สึกว่ายาก หรือ อึดอัดที่จะสูดลมหายใจเข้าลึกๆหรือไม่?

หากพบว่าตนเองตอบ ใช่ ต่อคำถามทั้งหมด หรือเกือบทั้งหมด นั่นนับว่าเป็นสัญญาณเตือนที่ต้องได้รับความช่วยเหลือ เนื่องจากความเครียดมีผลกระทบต่อการหายใจ [10][11]

**2.3.4 การเรียนรู้ที่จะปฏิเสธ**

ในชีวิตประจำวันมนุษย์จะพบเจอสถานการณ์ที่ทำให้เกิดการตัดสินใจขึ้น และส่วนมากมนุษย์จะมีปัจจัยหลายด้านที่ทำให้ไม่สามารถปฏิเสธสิ่งที่ไม่อยากทำ หรือไม่สบายใจได้ ซึ่งปัจจัยเหล่านั้นส่งผลทำให้เกิดการต่อต้านในการปฏิเสธ ตัวอย่างปัจจัยบางประการที่ทำให้พบว่ายากที่มนุษย์จะพูดปฏิเสธ [12]  
 1) ไม่อยากเสียมารยาท กับผู้สูงอายุ หรือผู้อาวุโส หรือในพื้นที่ที่มีวัฒนธรรมที่เฉพาะเจาะจง ซึ่งเป็นปัญหาที่ละเอียดอ่อน ทั้งนี้ก็คงไม่มีใครอยากเสียมารยาทโดยการบอกปฏิเสธ และบ่อยครั้งที่มันทำให้มนุษย์ไม่สามารถพูดว่า ไม่ได้  
 2) ไม่ต้องการให้ผู้คนผิดหวังและต้องการที่จะช่วยเหลืออย่างแท้จริง หากมนุษย์คนใดที่มีจิตใจที่ดี มนุษย์ผู้นั้นจะพยายามดิ้นรนและหาหนทางช่วยเหลืออย่างเต็มที่ในกรณีที่เป็นไปได้ แม้ว่าการช่วยเหลือนี้จะต้องเสียเวลาในการดำเนินชีวิตของตนเองก็ตาม และเหตุนี้ทำให้เรียกว่าได้ว่า เป็นเหตุให้เกิดความเครียด  
 3) ไม่ต้องการที่จะทำให้เกิดความขัดแย้งใดๆ บางทีบางสถานการณ์ที่ต้องตอบว่า ใช่ เพียงเพราะต้องการหลีกเลี่ยงการเผชิญหน้าจากความขัดแย้ง และมีชีวิตที่ง่ายขึ้น  
 4) ไม่ต้องการการแปลกแยก มนุษย์ทุกคนพบว่าตนเองในสถานการณ์นึง ในบางช่วงเวลาของชีวิต ไม่ต้องการแยกตนเองออกจากกลุ่ม เนื่องจากการแสดงความคิดเห็นทางด้านตรงข้ามกับกลุ่ม ดังนั้นมนุษย์จึงต้องเห็นด้วยกับคำพูด / ความคิดเห็นของผู้อื่น  
 5) ความกังวลที่ว่า เมื่อพูดออกไปว่า ไม่ อาจเป็นเหตุให้หันหลังให้กับโอกาสที่เข้ามาได้ [12]

ปัจจัยเหล่านี้ล้วนเป็นความเชื่อที่ผิด ที่จริงแล้วนั้นล้วนเกี่ยวกับวิธีที่มนุษย์ปฏิเสธไม่ได้ มากกว่าความเป็นจริงที่มนุษย์พูด ซึ่งส่งผลต่อผลลัพธ์ที่ตามมา โดยธรรมชาติของมนุษย์มักมีเวลาที่ยากในตอนที่จะกล่าวคำว่า ปฏิเสธ กับผู้อื่นไม่ว่าจะด้วยกรณีหรือปัจจัยใด และการบอกปฏิเสธยังเป็นความรู้สึกที่เหมือนปล่อยให้ผู้อื่นผิดหวัง แท้จริงแล้วจะมีสิ่งเดียวเท่านั้นที่สามารถทำร้ายมนุษย์ได้นั่นคือ ตัวมนุษย์เอง ซึ่งการปฏิเสธและถูกปฏิเสธ มักจะนำไปสู่ความเครียดและความรู้สึกที่ไม่ดีหรือกระวนกระวายตลอดเวลาหากไม่ได้รับการตอบสนองที่เหมาะสม สิ่งที่สำคัญอีก 1 ประการคือ การเรียนรู้ที่จะพูดโดยไม่รู้สึกผิด การตอบว่า ‘ใช่’ นั้น ไม่ใช่คำตอบที่ดีต่อสุขภาพเสมอไป เพราะเมื่อมนุษย์มีความมุ่งมั่นที่จะทำสิ่งนั้นให้ได้มากเกินไป หรือยึดติดมากเกินไป จะทำให้ตกอยู่ในภายใต้ความเครียดที่มากเกินไป ซึ่งจะมีผลกระทบต่อสุขภาพทำให้มีแนวโน้มที่จะรู้สึกแย่ลง และถึงขั้นป่วยหนัก ดังนั้น มนุษย์จึงต้องเรียนรู้ที่จะปฏิเสธ สิ่งที่ควรคำนึงถึงเมื่อจำเป็นต้องพูดปฏิเสธ  
 1) ปฏิเสธด้วยความมั่นใจ พูดออกไปอย่างมั่นใจและไม่กลัว และพยายามหลีกเลี่ยงวลีเช่น ‘ฉันไม่แน่ใจ’ เนื่องจากอาจถูกทำให้สับสนและไขว้เขวได้   
 2) ปฏิเสธด้วยความกะทัดรัดชัดเจนว่าทำไมคุณถึงพูดปฏิเสธ  
 3) อย่าพูดตรงๆ นอกจากนี้อย่าพยายามหาข้อแก้ตัวเพื่อให้หลุดพ้นจากคำมั่นสัญญา เพราะความซื่อสัตย์สุจริตเป็นสิ่งที่ดีที่สุด  
 4) อย่าพูดอย่างเคารพ บางสิ่งอาจไม่ถูกใจหรือไม่ได้อย่างที่หวัง ให้ชมเชยความพยายามและทำให้พวกเขาทราบว่า เราเคารพความพยายามของเขา ที่สามารถกระทำได้ในเวลา  
 5) อย่าพูดซ้ำๆ [13]

ในส่วนนี้ผู้จัดทำได้นำไปใช้ในการประยุกต์รูปแบบของคำและประโยคในการโต้ตอบกับ ผู้ใช้งาน เพื่อลดความเสี่ยงในการตอบคำถามของบอทที่ผิดพลาดและอาจส่งผลกระทบให้เกิดความรู้สึกที่ไม่ดีหรือแย่ลง หรือไปกระตุ้นความคิดด้านลบของผู้ใช้งานอันเป็นเหตุให้เกิดเหตุกาณ์ไม่คาดฝันได้ ห(เช่น การคิดสั้นฆ่าตัวตาย หรือการไปทำร้ายผู้อื่น เป็นต้น)

**2.4 เครื่องมือที่เกี่ยวข้อง**

**2.4.1 Machine Learning**

Machine Learning คือ เป็นแขนงความรู้ที่เติบโตมาจาก Artificial Intelligence ที่ให้คอมพิวเตอร์เรียนรู้ได้โดยไม่ต้องโปรแกรมโดยตรง มาจากประสบการณ์ (Experience) ที่เรียนรู้ด้วยงาน (Task) และมีประสิทธิภาพ (Performance) เช่น การแยกประเภทของอีเมล์ เป็นต้น ซึ่งจะมีการเรียนรู้ทั้งหมด 4 แบบ ซึ่งการเรียนรู้แต่ละแบบจะมีการเรียนรู้ในระดับย่อยที่แตกต่างกันออกไป ดังนี้

1) Supervised Learning

Supervised Learning คือ การเรียนรู้ที่ต้องมีผู้สอน ก่อนที่คอมพิวเตอร์จะสามารถหาคำตอบหรือผลลัพธ์ได้ด้วยตัวมันเอง หลังจากที่ได้เรียนรู้จากข้อมูลที่ใส่และสอนไปก่อนหน้านี้ เมื่อใส่ข้อมูลเพิ่มเติมลงไปมันก็จะบอกผลลัพธ์ได้ทันทีตามที่ถูกสอน ซึ่งหากข้อมูลที่สอนมีจำนวนมาก ในครั้งต่อไปก็จะยิ่งให้ผลลัพธ์ที่แม่นยำมากยิ่งขึ้นเช่นกัน

1.1) การเรียนรู้แบบ Classification

การที่สอนให้แบ่งกลุ่มข้อมูลสำหรับข้อมูลที่เป็นข้อมูลแบบไม่ต่อเนื่อง เช่น มีคนไข้ที่เป็นเนื้องอกในสมอง อยากทราบว่าเป็นเนื้องอกอันตรายหรือเนื้องอกที่ไม่อันตราย ซึ่งจำเป็นต้องใส่ชุดข้อมูลและสอนคอมพิวเตอร์ว่า ขนาดเนื้องอกที่อยู่ระหว่าง 1 – 2 มิลลิเมตร คือเนื้องอกไม่อันตรายและขนาดเนื้องอก 3-5 มิลลิเมตรเป็นเนื้องอกอันตราย คอมพิวเตอร์จะเรียนรู้การแบ่งกลุ่มตามที่สอน และเมื่อมีคนไข้ใหม่ที่เป็นเนื้องอกประมาณ 2 มิลลิเมตรคอมพิวเตอร์จะบอกได้เลยว่าเป็นเนื้องอกไม่อันตราย  
1.2) การเรียนรู้แบบ regression

เหมาะสำหรับข้อมูลที่เป็นข้อมูลที่ต่อเนื่อง เช่น ต้องการขายบ้านแต่ไม่ทราบว่าควรขายในราคาเท่าไหร่จึงจะเหมาะสมและไม่ขาดทุน จึงใส่ข้อมูลขนาดบ้าน ราคาบ้านแบบต่างๆลงในคอมพิวเตอร์ คอมพิวเตอร์จะเรียนรู้จากชุดข้อมูลที่มี และนำไปวิเคราะห์หาความสัมพันธ์ของสองตัวแปร (ขนาดบ้านและราคาบ้าน) ออกมาเป็นสมการการประมาณราคาบ้านที่สมควร

2) Unsupervised Learning

Unsupervised Learning คือ การเรียนรู้ที่ไม่ต้องอาศัยผู้สอน ใส่ข้อมูลลงไปในคอมพิวเตอร์โดยไม่ได้กำหนดผลลัพธ์ว่าจะต้องเป็นแบบไหน ให้คอมพิวเตอร์เรียนรู้ที่จะแยกกลุ่มได้ด้วยตัวเอง ซึ่งการเรียนรู้แบบ Unsupervised Learning มีการแบ่งเป็น 2 ประเภทหลักๆ คือ

2.1) Clustering

Clustering คือ การที่คอมพิวเตอร์สามารถเรียนรู้การแยกกลุ่มข้อมูลได้ด้วยตัวเอง เช่น การแบ่งส่วนตลาด (Market Segmentation) ซึ่งคือกระบวนการในการแบ่งหรือแยกลูกค้าออกเป็นกลุ่มย่อยๆ เพื่อให้ลูกค้าที่มีลักษณะความต้องการคล้ายคลึงกันอยู่ในกลุ่มเดียวกัน สามารถนำข้อมูลของลูกค้าที่มีใส่ลงในคอมพิวเตอร์เพื่อให้คอมพิวเตอร์แยกกลุ่มของลูกค้าให้โดยอัตโนมัติในวิธีการของ clustering ซึ่งจะช่วยให้เห็นรูปแบบการแบ่งกลุ่มลูกค้า และนำไปใช้ในการแบ่งเมื่อมีลูกค้าใหม่เข้ามา  
2.2) Non-clustering

ตัวอย่างของการเรียนรู้แบบนี้ คือ Cocktail Party Algorithm เป็นการเรียนรู้ของคอมพิวเตอร์ที่จะสามารถแยกเสียงที่ต้องการได้ ยกตัวอย่างคือเมื่อเข้าไปในงานเลี้ยง มีคนมากมายพูดคุยกันเสียงดังและมีเสียงเพลงอีกด้วย แล้วได้ยินเสียงของคู่สนทนาได้อย่างไร นั่นเพราะสมองสามารถที่จะแยกเสียงที่ต้องการจะฟังได้ประกอบกับใช้วิธีการอ่านปากไปด้วยจึงทำให้สามารถทำให้เข้าใจได้ ซึ่งตัว Machine Learning สามารถที่จะเรียนรู้การแบ่งแยกเสียงที่ต้องการคล้ายสมองได้เช่นกัน ซึ่งตอนนี้ทาง google ได้นำหลักการนี้มาพัฒนาเป็น Google Recognize สำหรับ Google Assistant [14]

**2.4.2 Natural Language Processing (NLP)**

Natural Language Processing เป็นส่วนหนึ่งของเทคโนโลยีปัญญาประดิษฐ์ (Artificial Intelligence) ที่ทำให้คอมพิวเตอร์เข้าใจภาษามนุษย์เพื่อวัตถุประสงค์ด้านการสื่อสาร

NLP เกี่ยวข้องกับ 2 กระบวนการ คือ

1. การเข้าใจภาษาธรรมชาติ (NLU) เป็นความสามารถของ chatbot ที่จะเข้าใจมนุษย์ เป็นกระบวนการแปลงข้อความเป็นข้อมูลที่มีโครงสร้าง
2. การสร้างภาษาธรรมชาติ (NLG) แปลงข้อมูลที่มีโครงสร้างเป็นข้อความ

NLP มีกระบวนการเรียนรู้ภาษาทั้งหมด 6 ขั้นตอน ดังนี้

1. Morphological Level ขั้นเข้าใจตัวอักษร ซึ่งถอดคำออกเป็นตัวอักษร หาพยัญชนะ สระ ตัวสะกดเพื่อทำความแม่นยำในขั้นต่อไป
2. Lexical Level ขั้นเข้าใจคำ เป็นขั้นตอนที่เริ่มหาความหมายของคำนั้นๆ เพื่อเตรียมสำหรับการทำความเข้าใจทั้งประโยค
3. Syntactic Level ขั้นเข้าใจประโยค โดยอ้างอิงจากการเข้าใจคำและลำดับโครงสร้างแบบแผนที่ได้เรียนรู้
4. Semantic Level ขั้นเข้าใจบริบทของคำในประโยค เข้าใจถึงความหมายของคำที่ใช้ในประโยค
5. Discourse Level ขั้นเข้าใจความเชื่อมโยงของประโยค เข้าใจความเชื่อมโยงระหว่างประโยคก่อนหน้าและประโยคปัจจุบัน
6. Pragmatic Level ขั้นเข้าใจความหมายของคำและประโยคอ้างอิงจากสถานการณ์หรือฐานความรู้เดิม สามารถเชื่อมโยงความรู้ใหม่กับความรู้เดิมให้ใกล้เคียงกับมนุษย์ [15]

**2.4.3 Text Classification**

Text Classification คือ กระบวนการของการกำหนดหมวดหมู่ให้กับข้อความตามเนื้อหา เป็นหนึ่งในงานพื้นฐานในการประมวลผลภาษาธรรมชาติ (NLP) ที่มีแอปพลิเคชันที่หลากหลายเช่น การวิเคราะห์ความเชื่อมั่น การติดฉลากหัวข้อ และการตรวจจับสแปม การทำ Text Classification สามารถใช้หลากหลายวิธีการ วิธีที่นิยมใช้ในปัจจุบันจะใช้ Machine Learning [16]

มีหลายวิธีในการจัดหมวดหมู่ข้อความอัตโนมัติซึ่งสามารถจัดกลุ่มเป็นระบบที่แตกต่างกัน 3 ประเภท คือ

1.) Rule-based systems

Rule-Based approach คือแนวทางการพัฒนาแชทบอท ที่จะโต้ตอบกับผู้ใช้ผ่านเงื่อนไข หรือกฎที่กำหนดเอาไว้ การพัฒนาบอทประเภทนี้จำเป็นต้องวางเงื่อนไขให้ครอบคลุม เพราะถ้ากำหนดเงื่อนไขไม่ครอบคลุมพอ หรือผู้ใช้งานถาม หรือตอบอะไรนอกเหนือจากเงื่อนไขที่เรากำหนดเอาไว้ แชทบอทจะไม่เข้าใจสิ่งที่ผู้ใช้งานจะสื่อสาร แต่อย่างไรก็ตาม Rule-Based Chatbot อาจจะไม่เหมาะ เพราะถ้าต้องการพัฒนาแชทบอทที่เปิดกว้าง ให้ผู้ใช้งานคุยกับบอทด้วย Natural Language Processing ซึ่งกระบวนการพัฒนาบางครั้งนักพัฒนาอาจใช้วิธี ตรวจสอบว่าถ้ามี keywords ดังนี้จะให้ตอบว่าอะไร [3]

2.) Machine Learning based systems

แทนที่จะอาศัยกฎที่สร้างขึ้นด้วยตนเอง การจัดหมวดหมู่ข้อความด้วย Machine Learning ที่จะจำแนกประเภทตามการสังเกตที่ผ่านมา ด้วยการใช้ตัวอย่างที่มีป้ายกำกับล่วงหน้าเป็นข้อมูลการฝึกอบรมอัลกอริทึม ของ Machine Leaning ให้สามารถเรียนรู้การเชื่อมโยงที่แตกต่างกันระหว่างข้อความ และคาดเดาเอาต์พุตเฉพาะ (เช่นแท็ก) สำหรับการป้อนข้อมูลเฉพาะ (เช่นข้อความ) อัลกอริทึมของ Machine Leaning ที่ได้รับความนิยมมากที่สุดสำหรับการสร้างแบบจำลองการจำแนกข้อความ คือ

1.) Naïve Bayes

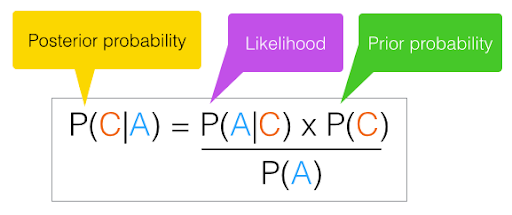
คือ การทำเหมืองข้อมูลในแบบ classifier ที่ถูกสร้างขึ้นโดยหลักความน่าจะเป็น Naïve Bayesian Classification จะใช้วิเคราะห์หาความน่าจะเป็นของสิ่งที่ยังไม่เคยเกิดขึ้น โดยการคาดเดาจากสิ่งที่เคยเกิด ขึ้นมาก่อน ใช้ทฤษฎีของ Bayes ในการแก้ปัญหา และสามารถเปรียบเทียบประสิทธิภาพได้กับเทคนิคแบบ Decision Tree และ Neural Network

สมการของ Bayes จะมี 3 ส่วนที่สำคัญ คือ

1.1.) Posterior probability หรือ P(C|A) คือ ค่าความน่าจะเป็นที่ข้อมูลที่มีแอตทริบิวต์เป็น A จะมีคลาส C

1.2.) Likelihood หรือ P(A|C) คือ ค่าความน่าจะเป็นที่ข้อมูล training data ที่มีคลาส C และมีแอตทริบิวต์ A โดยที่ A = a1 ∩ a2 … ∩ aMโดยที่ M คือจำนวนแอตทริบิวต์ใน training data

1.3.) Prior probability หรือ P(C) คือ ค่าความน่าจะเป็นของคลาส C [17]



**รูปที่** 2.3 สมการ Naive Bayes

(ที่มา: <http://dataminingtrend.com/2014/naive-bayes/>)

Multinomial Naïve Bayesian ใช้การแจกแจงแบบพหุนาม ซึ่งจะเหมาะสำหรับการนับคำในประโยคยาวๆ มากกว่า Naïve Bayesian ปกติที่ทำนายโมเดลจาก Features ที่เป็นอิสระจากกัน โดยมีสูตรคือ

นับคำ w ที่ปรากฎในคลาส c

นับคำทั้งหมดที่ปรากฎในคลาส c

โดยที่ w คือ คำ

c คือ คลาส

แต่มีโอกาสที่จะไม่เจอคำที่ต้องการ กล่าวคือไม่มีคำๆนั้นในประโยคเลยในข้อมูลtraining ทำให้จำนวนนับเป็น 0 ดังนั้นจึงต้องใช้วิธี Laplace Smoothing ในการจัดการ โดยมีค่า α (Alpha) เป็นตัวแทนค่าของคำที่ไม่ปรากฏ [3] ดังสูตรต่อไปนี้

โดยที่ V คือ array ของคำทุกคำในคลังคำศัพท์

2.) Support Vector Machines

เป็นอัลกอริทึมที่สามารถนำมาช่วยแก้ปัญหาการจำแนกข้อมูล ใช้ในการวิเคราะห์ข้อมูลและจำแนกข้อมูล โดยอาศัยหลักการของการหาสัมประสิทธิ์ของสมการเพื่อสร้างเส้นแบ่งแยกกลุ่มข้อมูลที่ถูกป้อนเข้าสู่กระบวนการสอนให้ระบบเรียนรู้ โดยเน้นไปยังเส้นแบ่งแยกแยะกลุ่มข้อมูลได้ดีที่สุด ซึ่งเกิดจากการที่นำค่าของกลุ่มข้อมูลมาวางลงในฟีเจอร์สเปซ (Feature Space) จากนั้นจึงหาเส้นที่ใช้แบ่งข้อมูลทั้งสองออกจากกันโดยจะสร้างเส้นแบ่ง (Hyperplane) ที่เป็นเส้นตรงขึ้นมา และเพื่อให้ทราบว่าเส้นตรงที่แบ่งสองกลุ่มออกจากกันนั้น เส้นตรงใดเป็นเส้นที่ดีที่สุด [18]



**รูปที่** 2.4 Support Vector Machine

(ที่มา: <https://www.javatpoint.com/machine-learning-support-vector-machine-algorithm>)

SVM เป็นแบบจำลองที่สร้าง Hyperplane ที่เหมาะสมถูกต้องในการแบ่งข้อมูลออกเป็น 2 คลาส หรือ Binary Classification โดยการสร้าง Hyperplane นั้นจะสอดคล้องกันกับ Hard-Margin และ Soft-Margin ซึ่งเป็นระยะขอบของเส้นแบ่ง Hyperplane มีสูตรคือ โดยที่ xi คือข้อมูลของคลาส w1 และ w2

* Hard-Margin

Hard-Margin หรือ Linear SVM นั้นเป็นการสร้างเส้น Hyperplane ที่ไม่ให้มีข้อมูลอยู่ในระหว่างเส้นประเลย หรือไม่ให้มีการทำนายผิดเลยนั่นเอง ตามรูปที่ ซึ่งขอบ Hyperplane 2 ฝั่งของแต่ละคลาสจะเท่ากับ 1 และ -1 ซึ่งจะทำให้สมการเป็น

ทำให้ได้สมการเป้าหมายที่ minimize ตัว Objective Function ของการOptimization เป็น

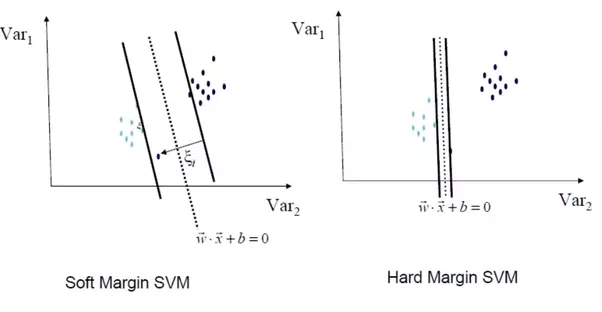
ภายใต้เงื่อนไข

* Soft-Margin

Soft-Margin เป็นการสร้างเส้น Hyperplane ที่มีข้อมูลอยู่ในระหว่างเส้นประ กล่าวคือยอมให้มี error หรือการทำนายผิดบ้างเล็กน้อย ทำให้ต้องเพิ่มตัวแปร Slack variable ζ(i) ≥ 0 ซึ่งกำหนดโดยค่า C (Hyperparameter) [5] ตามรูปที่ โดยมีสมการเป้าหมายที่ minimize ตัว Objective Function ของการOptimization เป็น

ภายใต้เงื่อนไข

ค่า C Hyperparameter นั้นแตกต่างกันออกไปขึ้นอยู่กับเป้าหมายของการ Optimize หากค่า C ยิ่งมาก ขอบ (Margin) ก็จะยิ่งแคบ หรือเป็นการสร้างขอบ (Margin) Hyperplane แบบ Hard Margin ทำให้ลดจำนวนของการทำนายผิด หรือ error แต่หากมีข้อผิดพลาดเกิดขึ้นกับข้อมูล เช่น การมี Outliers หรือ Noise ของข้อมูล ก็จะทำให้เส้นแบ่งนั้นผิดเพี้ยน

แต่หากค่า C ยิ่งน้อย ขอบ (Margin) จะกว้าง หรือเป็นการสร้างขอบ (Margin) Hyperplane แบบ Soft Margin ก็จะเป็นการเพิ่มจำนวนของการทำนายผิด ซึ่งเป็นการป้องกันเมื่อมีข้อมูลผิดพลาด ทำให้ปัจจุบันนิยมใช้วิธีแบบ Soft margin

รูปที่

https://medium.com/mmp-li/svm-อดีตเคยหวานปัจจุบันแอบเซง-machine-learning-101-6008753c780c

3.) Hybrid systems

ระบบผสมรวมส่วนขยายพื้นฐานที่ได้รับการฝึกฝนกับ Machine Learning และ Rule-based systems ซึ่งใช้เพื่อปรับปรุงผลลัพธ์ต่อไป ระบบผสมเหล่านี้สามารถปรับแต่งได้อย่างง่ายดายโดยการเพิ่มกฎเฉพาะสำหรับแท็กที่ตรงข้ามกัน ซึ่งยังไม่ได้รับการจัดรูปแบบอย่างถูกต้องโดยตัวแยกประเภทพื้นฐาน [16]

**2.4.4 N-Gram**

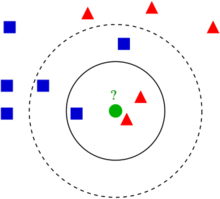
N-Gram คือ แบบจำลองที่ใช้คำนวณค่าความน่าจะเป็นของชุดอักขระ (Character Sequence) ที่เกิดขึ้นร่วมกันเป็นคำ หรือค่าความน่าจะเป็นของคำที่เขียนเรียงกัน (Word Sequence) ที่เกิดขึ้นร่วมกันเป็นประโยค โดยค่าความน่าจะเป็นของชุดอักขระหรือคำ ประมาณได้จากคลังข้อมูลที่สร้างไว้ ซึ่ง N-Gram ได้ใช้หลักการของสถิติในหลาย ๆ ด้านมาประยุกต์ใช้ [19]

**2.4.5 Machine Learning model อื่นๆ**

แบบจำลอง Machine Learning ที่นำมาประยุกต์ใช้ในงานวิจัยชิ้นนี้ ได้แก่

1. K-Nearest Neighbors (KNN)

เป็นวิธีการหาเพื่อนบ้านที่ใกล้ที่สุด ซึ่งใช้ในการแบ่งหมวดหมู่ข้อมูล จัดอยู่ใน Machine Learning ประเภท Supervised learning ซึ่ง KNN จะตัดสินใจว่าหมวดหมู่ใดที่จะแทนกรณีข้อมูลใหม่ๆที่ถูกป้อนเข้ามา โดยจะตรวจสอบจำนวน (K) ของเพื่อนบ้านสำหรับแต่ละกลุ่มที่อยู่ใกล้ที่สุด ทำการกำหนดกลุ่มให้กับข้อมูลใหม่ที่เหมือนกันกับกลุ่มที่ใกล้เคียงกันมากที่สุด เหมาะกับข้อมูลแบบตัวเลข แต่ตัวแปรที่เป็นค่าแบบไม่ต่อเนื่องก็สามารถทำได้เช่นกัน

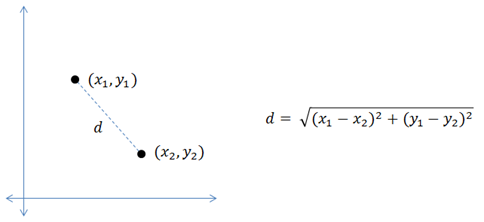


ตัวอย่างเช่น กำหนดให้จุดสีเขียวคือข้อมูลใหม่ที่ได้รับมา และต้องการจะแบ่งว่าจุดสีเขียวนั้นควรถูกจัดอยู่ในกลุ่มอะไร ระหว่างสามเหลี่ยมสีแดงหรือสี่เหลี่ยมสีน้ำเงิน

ถ้ากำหนดค่า k = 3 จุดสีเขียวจะถูกแบ่งอยู่ในกลุ่มสามเหลี่ยมสีแดง เนื่องจากสิ่งที่อยู่ใกล้เคียงกับจุดสีเขียวที่สุด 3 จุดนั้นมีจำนวนสามเหลี่ยมสีแดง 2 จุดซึ่งมากกว่าสี่เหลี่ยมสีน้ำเงินที่มีจำนวน 1 จุด

ถ้ากำหนดค่า k = 5 จุดสีเขียวจะถูกแบ่งอยู่ในกลุ่มสี่เหลี่ยมสีน้ำเงินเนื่องจากสิ่งที่อยู่ใกล้เคียงกับจุดสีเขียวที่สุด 5 จุดนั้นมีจำนวนสามเหลี่ยมสีแดง 2 จุดซึ่งน้อยกว่าสี่เหลี่ยมสีน้ำเงินที่มีจำนวน 3 จุด https://th.wikipedia.org/wiki/ขั้นตอนวิธีการค้นหาเพื่อนบ้านใกล้สุด\_k\_ตัว

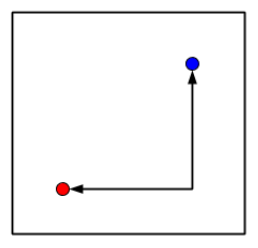
* 1. ขั้นตอนในการทำ KNN
     1. กำหนดค่า k (ควรเป็นเลขคี่)
     2. คำนวณระยะห่าง (Distance) ของข้อมูลที่ต้องการพิจารณากับกลุ่มข้อมูลตัวอย่างโดยมีประเภทของ Distance ดังต่อไปนี้

1. การวัดระยะทางแบบยุคลิด (Euclidean distance) คือวัดระยะทางปกติระหว่างจุดสองจุดในแนวเส้นตรง https://th.wikipedia.org/wiki/ระยะทางแบบยุคลิด

https://predictivehacks.com/tip-how-to-define-your-distance-function-for-hierarchical-clustering/

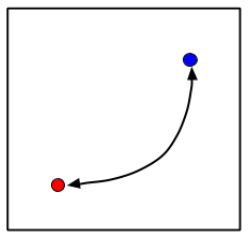
มีสูตรทั่วไปคือ

1. การวัดระยะทางแบบแมนฮัตตัน (Manhattan distance)

 <https://subscription.packtpub.com/book/big_data_and_business_intelligence/9781785882104/6/ch06lvl1sec40/measuring-distance-or-similarity>

มีสูตรทั่วไปคือ

1. การวัดระยะทางแบบ Minkowski (Minkowski distance)



<https://subscription.packtpub.com/book/big_data_and_business_intelligence/9781785882104/6/ch06lvl1sec40/measuring-distance-or-similarity>

มีสูตรทั่วไปคือ

* + 1. จัดเรียงลำดับของระยะห่าง และเลือกพิจารณาชุดข้อมูลที่อยู่ใกล้จุดที่ต้องการพิจารณาตามจำนวน K ที่ได้กำหนดไว้
    2. พิจารณาข้อมูลจำนวน k ชุด และสังเกตว่ากลุ่ม (class) ไหนที่ใกล้จุดที่พิจารณาเป็นจำนวนมากที่สุด
    3. กำหนดกลุ่ม (class) ให้กับจุดที่พิจารณา โดยเลือกกลุ่ม (class) ที่ใกล้จุดพิจารณามากที่สุด

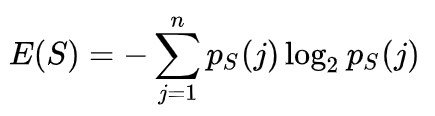
1. ต้นไม้ตัดสินใจ (Decision Tree)

เป็นอัลกอริธึมที่ใช้ทำนายประเภทของวตถุ โดยพิจารณาจากลักษณะของวัตถุ ซึ่ง inner node ของต้นไม้จะแสดงตัวแปร leaf node จะแสดงประเภทของวัตถุ และกิ่งจะแสดงค่าที่สามารถเป็นไปได้ของตัวแปรนั้นๆ

Decision tree จะทำการจัดกลุ่มของชุดข้อมูลที่ถูกนำเข้าในแต่ละกรรี โดยแต่ละnodeของต้นไม้ก็คือตัวแปรหรือ attribute ต่างๆของข้อมูลชุดนั้น

ในปัจจุบันใช้วิธีการสอนต้นไม่ตัดสินใจโดย ID3 ซึ่งสร้างต้นไม้ตัดสินใจจากบนลงล่าง

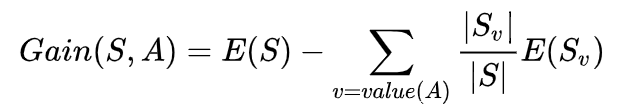
2.1) ขั้นตอนวิธีการสร้างต้นไม้การตัดสินใจ

2.1.1) เอนโทรปี (Entropy) เป็นสิ่งที่ใช้บอกถึงความไม่บริสุทธิ์ของข้อมูล

เมื่อ

S คือตัวอย่างที่ประกอบด้วยชุดของตัวแปรต้นและตัว  
แปรตามหลายๆกรณี

Ps(j) คืออัตราส่วนของกรณีใน S ที่ตัวแปรตามหรือผลลัพธ์มีค่า j

2.1.2) เกนความรู้ (Information Gain) เป็นค่าที่บอกความดีของตัวแปรต้นที่พิจารณา

เมื่อ

S คือตัวอย่างที่ประกอบด้วยชุดของตัวแปรต้นและตัว  
แปรตามหลายๆกรณี

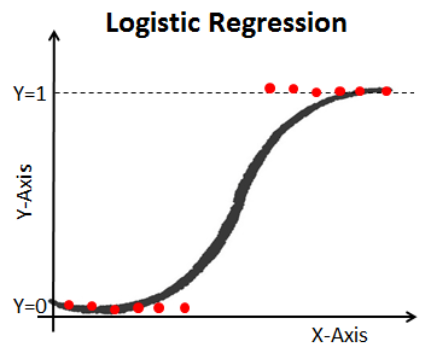
E คือ Entropy ของตัวอย่าง

A คือตัวแปรต้นที่พิจารณา

Sv คือตัวอย่างที่ A มีค่า v ทั้งหมด

2.1.3) นำตัวแปรต้นทั้งหมดมาหา Information Gain จากนั้นให้เลือกตัวที่มีค่า Information Gain สูงสุด และสร้างต้นไม้ที่มีรากเป็นตัวแปรต้นตัวนั้น

<https://th.wikipedia.org/wiki/ต้นไม้ตัดสินใจ>

1. การวิเคราะห์การถดถอยโลจิสติค (Logistic Regression)

คืออัลกอริธึมที่ใช้เพื่อศึกษาว่าตัวแปรอิสระใดบ้างที่สามารถอธิบายตัวแปรตามได้ หรือเพื่อทำนายว่าจะเกิดเหตุการณ์หนึ่งขึ้นหรือไม่ http://rdi.rmutsv.ac.th/rmutsvrj/download/year4-issue1-2555/p1.pdf

https://medium.com/@nonthakon/ machine-learning-ด้วย-python-การจำแนกประเภทด้วย-logistic-regression-ใน-scikit-learn-e78b8aa04517

โดยมีสูตรดังนี้

http://faculty.cas.usf.edu/mbrannick/regression/Logistic.html

โดย

P คือความน่าจะเป็นของการเกิดเหตุการณ์นั้นๆ

e คือ exponential function (e = 2.718)

a คือค่าของผลตอบแทน P เมื่อ X เป็นศูนย์

b คือค่าที่ปรับความเร็วของความน่าจะเป็นที่เปลี่ยนไปด้วยการเปลี่ยน X ตัวเดียว

1. การเคลื่อนลงตามความชัน (Gradient Descent Algorithm)

เป็นอัลกอริธึมที่ใช้หาค่าที่เหมาะสมที่สุดให้กับฟังก์ชันที่กำหนดขึ้น โดยใช้การเคลื่อนที่ซ้ำๆในทิศทางที่มีความชันมากที่สุด ใน Machine Learning นั้นใช้เพื่ออัปเดตค่าพารามิเตอร์ของแบบจำลอง ที่แสดงถึงค่าสัมประสิทธิ์ (Coefficients) ในการวิเคราะห์การถดถอยเชิงเส้น (Linear Regression)

<https://th.wikipedia.org/wiki/การ>เคลื่อนลงตามความชัน

https://ml-cheatsheet.readthedocs.io/en/latest/gradient\_descent.html

1. Random Forest

เป็นอัลกอริธึมของ Machine Learning ที่จัดอยู่ในประเภทของ Supervised Learning เป็นหนึ่งในวิธี Ensemble ซึ่งใช้สำหรับแบบจำลองประเภทการจัดหมวดหมู่ (Classification) และการถดถอย (Regression)

Random Forest เป็นอัลกอริธึมที่สร้างต้นไม้ตัดสินใจ (Decision Tree) หลายๆต้น จากข้อมูลตัวอย่าง จากนั้นก็ทำนายจากแต่ละต้น และเลือกวิธีการที่ดีที่สุดโดยใช้การโหวต

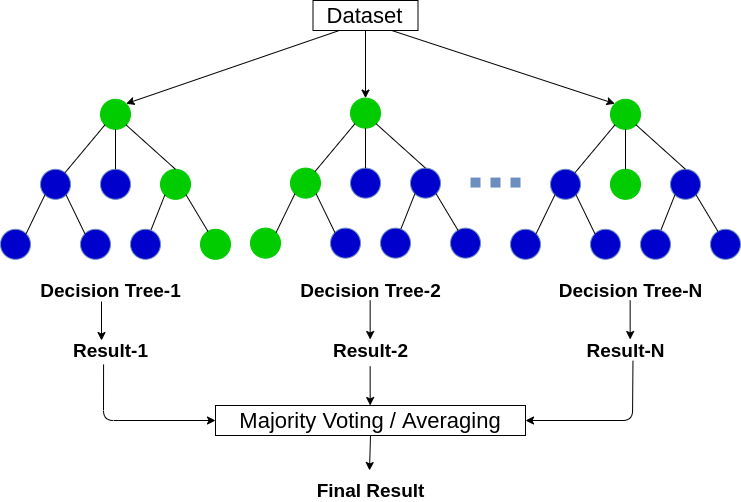
5.1.1) การทำงานของ Random Forest

1.) สุ่มตัวอย่างจากชุดข้อมูลที่มี

2.) อัลกอริธึมนี้จะสร้างต้นไม้ตัดสินใจ (Decision Tree) สำหรับทุกๆตัวอย่าง จากนั้นก็จะทำนายผลจากทุกๆต้น

3.) ทำการโหวตเลือกต้นไม้ที่ให้ผลลัพธ์จากการทำนายดีที่สุด

4.) เลือกผลลัพธ์การทำนายที่ได้รับการโหวตมากที่สุด

https://www.tutorialspoint.com/machine\_learning\_with\_python/machine\_learning\_with\_python\_classification\_algorithms\_random\_forest.htm

https://www.analyticsvidhya.com/blog/2020/05/decision-tree-vs-random-forest-algorithm/

1. Ensemble Learning

เป็นหนึ่งในเทคนิคของ Machine Learning ที่นำเทคนิคหรืออัลกอริธึมต่างๆที่ใช้มารวมกันอยู่ในแบบจำลองตัวเดียว เพื่อเพิ่มประสิทธิภาพในการทำนาย

https://blog.statsbot.co/ensemble-learning-d1dcd548e936

1. Feature Selection

เป็นกระบวนการในการลดจำนวนตัวแปร เพื่อปรับปรุงการทำนายของแบบจำลองให้มีประสิทธิภาพมากขึ้น ซึ่งจะแบ่งเป็น

7.1) Unsupervised

7.1.1) Correlation

7.2) Supervised

7.2.1) Wrapper

7.2.2) Filter

7.2.3) Intrinsic

https://machinelearningmastery.com/feature-selection-with-real-and-categorical-data/

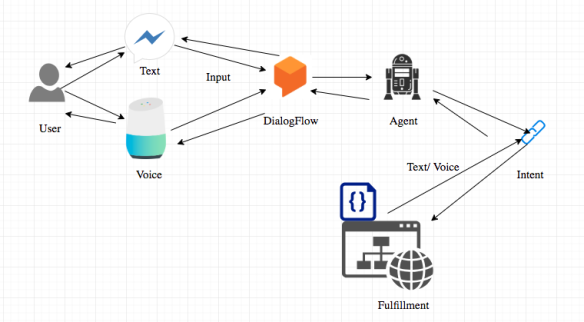
**2.5 เทคโนโลยีที่เกี่ยวข้อง**

**2.5.1 Dialogflow**

Dialogflow เป็นเทคโนโลยีการโต้ตอบสนทนาระหว่างมนุษย์และคอมพิวเตอร์ของ Google สำหรับสร้าง chatbot ที่มีการใช้ Machine Learning ด้านภาษาธรรมชาติ (Natural language) เช่น Assistant (โดย Speaktoit) ที่ได้สร้างเครื่องมือประมวลผลภาษาธรรมชาติ (Natural Language Processing) ปัจจุบัน Dialogflow รองรับการประมวลผลภาษาธรรมชาติ(Natural Language processing) มากกว่า 20 ภาษารวมถึงภาษาไทย มาช่วยในการทำความเข้าใจถึงความต้องการ (Intent) ของประโยคที่ได้รับจากแชทของผู้ใช้งาน และตอบคำถามความต้องการ (Intent) ของผู้ใช้งานได้อย่างถูกต้องและตรงประเด็นตามกฎของผู้พัฒนาที่ได้ถูกพัฒนาไว้ ซึ่ง Dialogflow จะช่วยเพิ่มความยืดหยุ่นของประโยคที่บอทรับมา

ใน Dialogflow สามารถจัดกลุ่มของประโยคที่มีความต้องการ (Intent) เหมือนกันเข้าด้วยกันเพื่อให้บอทได้เรียนรู้ และกำหนดคำตอบสำหรับแต่ละความต้องการ (Intent) ได้ ถ้าคำตอบที่ได้ถูกกำหนดไว้มีมากกว่า 1 ประโยค ทาง Dialogflow จะทำการสุ่มคำตอบในการตอบกลับให้กับผู้ใช้งานยกตัวอย่างเช่น Intent : แนะนำเมนูอาหาร ถ้าผู้ใช้งานพิมพ์มาบอกประโยคประมาณว่า กินไรดี หิวอะ ไม่อยากกินอันนี้เลย เป็นต้น และผู้พัฒนาได้กำหนดคำตอบไว้ประมาณว่า ข้าวมันไก่ ข้าวผัด สเต็ก กะเพราะ เป็นต้น เมื่อผู้ใช้งานพิมพ์โต้ตอบกับบอทด้วยคำว่า กินไรดี บอทจะทำความเข้าใจก่อนว่า คำ/ประโยคของผู้ใช้งานที่พิมพ์เข้ามาว่าอยู่ในความต้องการ (Intent) ไหนและทำการสุ่มคำตอบมาหนึ่งคำ/ประโยคที่ทางผู้พัฒนาได้กำหนดไว้ภายใต้ Intent นั้นๆ [20]

นอกจากนี้ Dialogflow สามารถเชื่อมต่อกับ Chat Platform ได้มากมาย เช่น Facebook Messanger, Twitter, Skype, Telegram และอีกมากมาย รวมถึง Application Line ที่ทางผู้จัดทำได้ใช้ในโครงงานนี้ด้วย



**รูปที่** 2.5 แสดงการทำงานคร่าวๆของ Dialogflowเมื่อมีข้อความเข้ามาจากผู้ใช้งาน

(ที่มา : <https://programmerbonn.home.blog/2019/01/20/dialogflow-chatbot-%E0%B8%97%E0%B8%B3%E0%B8%87%E0%B8%B2%E0%B8%A2/>)

**2.5.2 Heroku**

Heroku เกิดมาจากแนวคิดเพื่อสนับสนุนนักพัฒนาซอฟแวร์ (Software Developer) เพราะ heroku มีทุกอย่างที่จะช่วยให้นักพัฒนาซอฟแวร์ ไม่ต้องกังวลเกี่ยวกับเรื่อง environment ต่างๆในการรันโปรแกรม หรือการนำไปปรับใช้ขึ้นสู่ระบบ เพื่อให้มุ่งมั่นกับการพัฒนาโปรแกรม, การทำ software product

Heroku เป็นแพลตฟอร์มคลาวด์ที่เป็น Platform as a Service (Paas) ซึ่งช่วยให้บริษัทต่างๆสร้าง(build) ส่งมอบ(deliver) ตรวจสอบ(monitor) และปรับขนาด(scale) ของแอพพลิเคชั่น heroku เป็นวิธีที่เร็วที่สุดในการเปลี่ยนจากแนวคิดไปสู่ ​​URL ผ่านโครงสร้างพื้นฐานทั้งหมด โดย heroku นั้นรองรับภาษาโปรแกรมที่หลากหลาย เช่น Ruby, PHP, Node.js, Python, Java, Clojure, Scala และยังสามารถสร้าง buildpack สำหรับภาษาอื่นๆได้ เช่น Lua ที่รันอยู่บน OpenResty [21]

ผู้จัดทำใช้Heroku ในการเชื่อมต่อ code python กับ Line Chatbot ที่ได้ทำการสร้างขึ้นมา เพื่อแสดงผลลัพธ์ของchatbot และสามารถนำไปใช้งานเพื่อแนะนำการศึกษาต่อคระเทคดนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบังได้จริง

**2.5.3 Line**

LINE คือแอพพลิเคชั่นที่มีบริการ Messaging และ Voice Over IP ที่ถูกรวมเข้าไว้ด้วยกัน จึงทำให้เกิดเป็นแอพพลิชั่นที่สามารถ สร้างกลุ่ม ส่งข้อความ โพสต์รูปต่างๆ หรือจะโทรคุยกันแบบเสียงก็ได้ มีทั้งแบบวิดีโอและแค่เสียง โดยไม่ต้องเสียเงิน หากเราใช้งานโทรศัพท์ที่มีแพคเกจอินเทอร์เน็ตอยู่แล้ว แถมยังสามารถใช้งานร่วมกันระหว่าง iOS และ Android รวมทั้งระบบปฏิบัติการอื่น ๆ ได้อีกด้วย [22]

LINE OFFICIAL ACCOUNT หรือ LINE OA คือ บริการสร้างบัญชีทางการของ LINE ผู้ใช้บริการ หรือบริษัทสามารถสร้างบัญชีของตนเอง และส่งข้อความถึงผู้ใช้บริการจำนวนมากที่เป็นเพื่อนกับบัญชีได้ทั่วทั้งประเทศไทย รวมถึงโพสต์ข่าวสารบนไทม์ไลน์ และอื่นๆ ซึ่งเหมาะกับทั้งบริษัทหรือร้านค้าในทุกขนาดตั้งแต่ระดับ SMEs ไปจนถึงแบรนด์ขนาดใหญ่ โดยเริ่มต้นง่ายๆ จากการสร้าง LINE OFFICIAL ACCOUNT ซึ่งเป็นบัญชี LINE ที่เหมือนกับ Line ปกติทั่วไปที่ไว้สื่อสารกับเพื่อนและครอบครัว โดยสามารถส่งได้ทั้งข้อมูลทั่วไป กิจกรรมทางการขายและการตลาด หรือโปรโมชั่นพิเศษไปยังลูกค้า และแบรนด์สามารถจัดการข้อความหรือรูปภาพต่างๆ ได้ด้วยตนเองผ่านระบบจัดการคอนเทนต์ของ LINE ทั้งนี้ยังเพียบพร้อมด้วยฟีเจอร์ที่จะช่วยให้ลูกค้ารู้ความพิเศษของบริษัท แบรนด์และสินค้าได้อีกด้วย

บริการอื่นๆที่เกี่ยวข้องกับพัฒนาของไลน์ (LINE Developers)

1.) รับพัฒนา Audio message สามารถส่งไฟล์ข้อความเสียงได้ด้วย URL HTTPS

2.) รับพัฒนา Location message สามารถส่งพิกัดแผนที่ได้

3.) รับพัฒนา Messaging API

3.1) Text message

3.2) Sticker message

3.3) Image message

3.4) Video message

3.5) Audio message

3.6) Location message

3.7) Imagemap message

3.8) Template message

3.9) Flex Message  
4.) รับพัฒนา LINE BOT สำหรับองค์กร (PHP,Node.js) [23]

**2.5.4 Anaconda (Python distribution)**

เป็นเครื่องมือในการจัดการ package python ที่ผู้คนนิยมใช้งานกันทั่วโลก ซึ่งรวมpackage มากมายสำหรับทำ data preparation, data analysis รวมถึง machine learning

Anaconda Navigator ประกอบไปด้วย

1. Jupyter Lab

Jupyter Lab เป็น user interface ที่ยืดหยุ่นมากสำหรับ Project Jupyter [26]

1. Jupyter Notebook

Jupyter Notebook เป็น web-based interactive ใช้ในการสร้าง Jupyter notebook document ซึ่งเป็นรูปแบบJSON อีกทั้งยังสามารถแปลงรูปแบบเอกสารให้เป็นรูปแบบอื่นๆได้ เช่น HTML, presentation slides, PDF, Python [26]

1. QtConsole

QtConsole ใช้ในการปรับปรุง graphical user interface (GUI) [27]

1. Spyder

Spyder เป็น open source ของ integrated development environment (IDE) หรือเครื่องมือที่ช่วยในการพัฒนาโปรแกรม สำหรับการเขียนโปรแกรมในภาษา Python [26]

1. glue

glue เป็นpackage ในการเชื่อมโยง data visualization ที่เขียนในภาษา python [26]

1. Orange

Orange เป็นopen-source ในการทำ data visualization, machine learning และdata mining [26]

1. RStudio

RStudio เป็นintegrated development environment (IDE) หรือเครื่องมือที่ช่วยในการพัฒนาโปรแกรม สำหรับภาษา R ซึ่งเป็นภาษาในการเขียนโปรแกรมในเชิงสถิติ [26]

1. Visual Studio Code

Visual Studio Code เป็น Code Editor ที่ใช้ในการแก้ไขและปรับแต่งโค้ด [27][28]

**2.5.5 Git และ Github**

Git เป็นซอฟต์แวร์ที่ใช้ระบบVersion control สามารถจัดเก็บการเปลี่ยนแปลงของไฟล์ได้ซึ่งสามารถเรียกดูเวอร์ชั่นต่างๆของโปรเจคได้ [29][30]

Github คือเว็บไซต์ที่ให้บริการ Git ทำให้ผู้ใช้งานสามารถใช้ Git ร่วมกับผู้อื่นได้ โดยจะมี Repository (repo) หรือพื้นที่เก็บข้อมูล เป็นตำแหน่งที่เก็บไฟล์ทั้งหมดของโครงงานนั้นๆ ซึ่งแต่ละโครงงานจะมี repo เป็นของตัวเอง และสามารถเข้าถึงได้โดยURLที่ไม่ซ้ำกัน [31]

**บทที่ 3**

**วิธีการดำเนินงานวิจัย**

**3.1 ภาพรวมของระบบ**

ระบบแชทบอทแนะนำการศึกษาต่อในคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบังจะมีผู้ใช้งานคือ นักเรียนชั้นมัธยมศึกษาที่มีความสนใจจะศึกษาต่อในคณะเทคโนโลยีสารสนเทศ บุคคลทั่วไปหรือผู้ปกครองที่มีความสนใจจะสอบถามข้อมูลเกี่ยวกับคณะ และนักศึกษาคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ผู้ใช้งานจะสามารถใช้งานส่วนต่างๆของระบบแชทบอทได้ ดังนี้

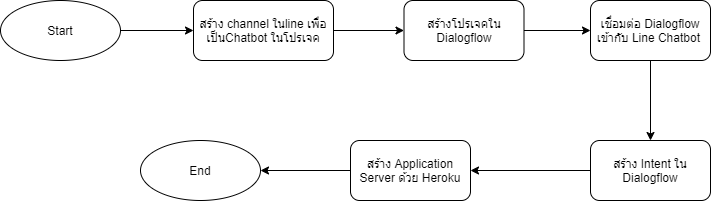
1. สามารถตรวจสอบข้อมูลหลักสูตรของสาขาและแขนงต่างๆได้
2. สามารถอ่านประสบการณ์ของรุ่นพี่ในคณะได้ เช่น การเรียน การทำกิจกรรม การใช้ชีวิตในสถาบัน เป็นต้น
3. สามารถตรวจสอบคะแนนที่ใช้ในการสอบเข้าเพื่อศึกษาต่อคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบังได้
4. สามารถเทียบคะแนน เพื่อดูโอกาสสอบติดคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบังได้
5. สามารถใช้งานการทำนายความเหมาะสมในการศึกษาต่อคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง



**รูปที่** 3.1 ตัวอย่างระบบแชทบอทแนะนำการศึกษาต่อในคณะเทคโนโลยีสารสนเทศ

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

**3.2 Flowchart ในการทำ Chatbot**



**รูปที่** 3.2Flowchart ในการทำ Chatbot

**3.3 ขั้นตอนการเก็บข้อมูล**

ผู้วิจัยได้ทำการเก็บข้อมูลจากกลุ่มตัวอย่างจำนวน 60 คน ซึ่งเป็นนักศึกษาคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง โดยใช้วิธีการสัมภาษณ์ซึ่งมีหัวข้อสัมภาษณ์ดังต่อไปนี้

* + 1. รอบการสมัคร
    2. เหตุผลที่ตัดสินใจศึกษาต่อในคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
    3. ตั้งใจจะเข้าศึกษาต่อในคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบังตั้งแต่แรกหรือไม่
    4. มีความสนใจจะศึกษาต่อในคณะอื่นๆ หรือสถาบันอื่นๆอีกหรือไม่ อะไรบ้าง
    5. พื้นฐานการเขียนโปรแกรม
    6. สายการเรียนในชั้นมัธยมศึกษาตอนปลาย
    7. ระดับความชอบของกลุ่มวิชาต่อไปนี้
  1. คณิตศาสตร์
  2. วิทยาศาสตร์
  3. สังคมศึกษา
  4. ภาษาต่างประเทศ
  5. พลศึกษา
  6. ศิลปะ
  7. เทคโนโลยี
     1. ระดับความชอบของการศึกษาเล่าเรียนในคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

โดยจะต้องนำตัวแปรต่างๆมาจำแนกความเหมาะและไม่เหมาะที่จะศึกษาต่อในคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง โดยพิจารณาจากตัวแปรดังนี้

* 1. GPA ของนักศึกษา
  2. ระดับความชอบของการศึกษาเล่าเรียนในคณะเทคโนโลยีสารสนเทศ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

**3.3.1 การจำแนกประเภท tag ข้อมูลตามประเภทของความเครียด**

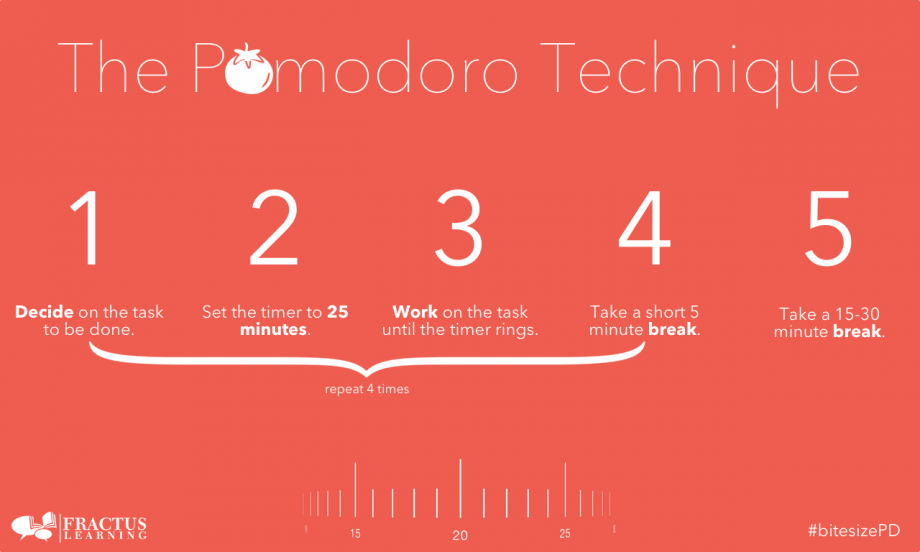
การทำ Text Classification นั้น ผู้วิจัยได้ทำการศึกษาค้นคว้าหาหัวข้อประเภทของความเครียดที่ครอบคลุมและเหมาะสมกับข้อมูล เพื่อนำข้อมูลที่ได้มาจำแนกประเภท โดยอ้างอิงถึงความเครียด 4 ประเภท เป็นความเครียดที่พบเจอได้บ่อย รวมถึงคำแนะนำวิธีจัดการความเครียดในแต่ละประเภทของ Karl Albrecht ซึ่งสามารถใช้เพื่อแนะนำวิธีการจัดการกับความเครียดเบื้องต้นตามแต่ละประเภทได้ ประกอบไปด้วย

1.) ความเครียดเกี่ยวกับเวลา (Time Stress)

เป็นประเภทความเครียดที่เกิดจากการกังวลเรื่องเวลาไม่พอ กังวลว่าจะไม่สามารถทำงานให้เสร็จได้ตามกำหนด หรือเร่งรีบเพื่อหลีกเลี่ยงการมาสาย

การจัดการกับความเครียดประเภทนี้ คือ

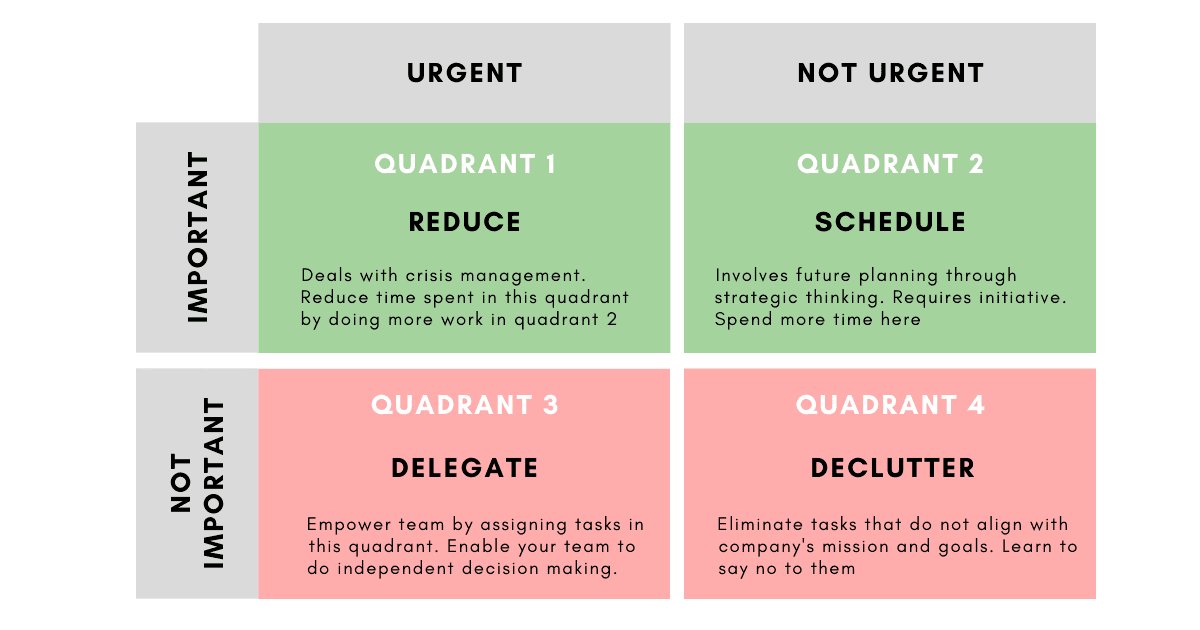
* 1. จดรายการสิ่งที่ต้องทำ หรือ To-Do Lists โดยให้ระบุสิ่งที่ต้องทำและเวลาที่จะต้องทำให้เสร็จ รวมถึงเรียงลำดับความสำคัญของงาน หรืออาจใช้ Pomodoro Technique ในการช่วยแก้ปัญหา

**รูปที่** 3.3 Pomodoro Technique

(ที่มา: <https://www.fractuslearning.com/tomato-students-stay-on-task/>)

จากรูปที่ 8 ขั้นตอนการบริหารจัดการเวลาด้วย Pomodoro Technique มีดังนี้

* + 1. วางแผนสิ่งที่ต้องการจะทำให้สำเร็จ
    2. จับเวลาถอยหลังไว้ที่ 25 นาที
    3. ทำตามแผนนั้นๆ จนครบ 25 นาที
    4. พัก 5 นาที
    5. ทำขั้นตอนที่ 1-4 วนไปเรื่อยๆ จนครบ 4 ครั้ง
    6. พักยาว 15-30 นาที
    7. ทำวนไปเรื่อยๆจนแผนงานนั้นเสร็จ
  1. ให้ความสำคัญกับการจัดลำดับความสำคัญของงาน ซึ่งอาจใช้เทคนิค Eisenhower Matrix เข้ามาช่วย

**รูปที่** 3.4 Eisenhower Matrix

(ที่มา: <https://www.techtello.com/eisenhower-productivity-matrix/>)

จากรูปที่ 3.4 ได้อธิบายเทคนิค Eisenhower Matrix ที่ช่วยในการจัดลำดับความสำคัญของงาน โดยมีรูปแบบเป็นตาราง 4 ช่อง ได้แก่

1.2.1 ช่องที่ 1 คืองานด่วนและสำคัญ งานที่อยู่ในช่องนี้ควรลงมือทำทันที นอกจากนี้ยังสามารถลดการใช้เวลาในงานส่วนนี้ได้โดยการทำงานในส่วนของช่องที่ 2 เพิ่ม

1.2.2) ช่องที่ 2 คืองานไม่ด่วนแต่สำคัญ งานที่อยู่ในช่องนี้ส่วนใหญ่มีความสัมพันธ์เกี่ยวเนื่องกันกับงานในอนาคต เพราะฉะนั้นควรให้ความสำคัญเช่นกัน

1.2.3) ช่องที่ 3 คืองานด่วนแต่ไม่สำคัญ งานในช่องนี้ สามารถมอบหมายให้ผู้อื่นรับผิดชอบแทนก่อนได้

1.2.4) ช่องที่ 4 คืองานไม่ด่วนและไม่สำคัญ เช่น งานที่ไม่เกี่ยวข้องกับงานของบริษัท, งานเลี้ยงสังสรรค์ เป็นต้น

* 1. เลือกเวลาทำงานในตอนที่มีพลังมากที่สุด เช่น เวลาเช้า บางคนจะมีสมาธิมากขึ้น เป็นต้น
  2. รู้จักปฏิเสธคนอื่น เมื่อมีงานที่ต้องทำ

1. ความเครียดจากการคาดการณ์ (Anticipatory stress)

เป็นประเภทความเครียดที่กังวลต่อเหตุการณ์ในอนาคต ความเครียดชนิดนี้ส่งผลให้มีอารมณ์หงุดหงิด เนื่องจากไม่สามารถทำอะไรได้และควบคุมไม่ได้

การจัดการกับความเครียดประเภทนี้ คือ

* 1. คิดหาผลลัพธ์ที่เป็นไปได้ทั้งหมด เพื่อทำแผนรับมือฉุกเฉิน
  2. ใช้เทคนิคการสร้างภาพเชิงบวก โดยจินตนาการถึงสถานการณ์ที่ดี และเหตุการณ์ต่อจากนั้น
  3. นั่งสมาธิ ซึ่งช่วยในการจดจ่อกับปัจจุบันมากขึ้น

1. ความเครียดจากสถานการณ์ (Situational stress)

เป็นประเภทความเครียดที่ไม่สามารถควบคุมได้ เช่น การทำอะไรผิดพลาด, เหตุการณ์ฉุกเฉิน, สถานการณ์ความขัดแย้ง เป็นต้น ส่งผลให้รู้สึกหมดหนทางในการแก้ปัญหา

การจัดการกับความเครียดประเภทนี้ คือ

* 1. ระแวดระวังตัวให้มากขึ้น
  2. เรียนรู้ที่จะจัดการกับความขัดแย้ง
  3. เรียนรู้ที่จะจำอาการทางร่างกายและและอารมณ์เมื่อเกิดความเครียด ทำให้สามารถควบคุมตัวเองดีขึ้น

1. ความเครียดจากการเผชิญหน้า (Encounter stress)

เป็นประเภทความเครียดที่เกิดจากการมีปฏิสัมพันธ์กับผู้อื่น เช่น บุคคลรอบตัว, ลูกค้า, เจ้านาย, เพื่อนร่วมงาน เป็นต้น

การจัดการกับความเครียดประเภทนี้ คือ

* 1. พัฒนาทักษะการสร้างความสัมพันธ์
     1. พัฒนาความฉลาดทางอารมณ์
     2. ยอมรับขีดจำกัดของตนเอง โดยการเรียนรู้ขีดจำกัดของตนเองเวลาสื่อสารกับแต่ละบุคคล [32][33]

ผู้วิจัยได้ค้นหาแท็กที่ผู้คนส่วนใหญ่ใช้ใน Twitter เพื่อระบายความเครียด และได้จัดการจำแนกแท็กต่างๆ ตามความเครียด 4 ประเภทของ Karl Albrecht ดังตารางต่อไปนี้

**ตารางที่** 3.1 ตารางจำแนกแท็กตามประเภทความเครียดของ Karl Albrecht

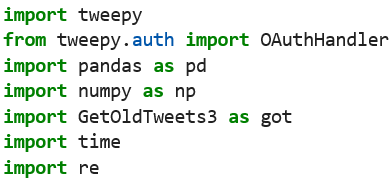
|  |  |  |  |
| --- | --- | --- | --- |
| **สาเหตุของความเครียด** | | | |
| ความเครียดจากเวลา | ความเครียด  จากสถานการณ์ | ความเครียดจาก  การคาดการณ์ล่วงหน้า | ความเครียด  จากการเผชิญหน้า |
| #เราจะผ่านมันไปด้วยกัน | #โควิด19 | #โควิด19 | #เครียด |
| #เคอร์ฟิว | #ไวรัสโคโรนา | #ไวรัสโคโรนา | #ปรึกษา |
| #เคอร์ฟิว24ชั่วโมง | #WFH | #การบินไทย | #ผิดหวัง |
| #การศึกษาฆ่าฉัน | #WorkFromHome | #ประกันสังคม | #เบื่อ |
| #dek63 | #อยู่บ้านหยุดเชื้อเพื่อชาติ | #พรกฉุกเฉิน | #ระบาย |
| #การศึกษา | #SocialDistancing | #ผนงรจตกม |  |
| #เครียด | #เราจะไม่ทิ้งกัน | #ฝุ่นPM25 |  |
| #เหนื่อยกับชีวิต | #เราจะผ่านมันไปด้วยกัน | #การศึกษาฆ่าฉัน |  |
| #เหนื่อย | #การบินไทย | #dek63 |  |
| #เบื่อ | #เคอร์ฟิว | #การศึกษา |  |
|  | #ประกันสังคม | #เครียด |  |
|  | #พรกฉุกเฉิน | #ปรึกษา |  |
|  | #รัฐบาลเฮงซวย | #ผิดหวัง |  |
|  | #รัฐบาลส้นตีน | #เบื่อ |  |
|  | #ผนงรจตกม | #ระบาย |  |
|  | #ฝุ่นPM25 | #โควิด19เราต้องรอด |  |
|  | #เคอร์ฟิว24ชั่วโมง |  |  |
|  | #การศึกษาฆ่าฉัน |  |  |
|  | #dek63 |  |  |
|  | #การศึกษา |  |  |
|  | #หุ้นตก |  |  |
|  | #เครียด |  |  |
|  | #เหนื่อยกับชีวิต |  |  |
|  | #เหนื่อย |  |  |
|  | #ปรึกษา |  |  |
|  | #ท้อ |  |  |
|  | #ผิดหวัง |  |  |
|  | #เบื่อ |  |  |
|  | #อยากหลุดพ้น |  |  |
|  | #ระบาย |  |  |
|  | #ซึมเศร้า |  |  |
|  | #โควิด19เราต้องรอด |  |  |
|  | #WFH |  |  |

จากตารางที่ 3.1 มีแท็กทั้งหมด 33 แท็ก และแท็ก 1 อัน สามารถอยู่ได้หลายประเภท เนื่องจากข้อมูลมีความหลากหลาย

**3.3.2) เริ่มต้นเก็บข้อมูลจาก Twitter โดยใช้ Python**

1.) Import Library ต่างๆที่ต้องใช้

* 1. tweepy: ใช้ในการดึงข้อมูลจาก Twitter
  2. Library GetOldTweets3: ช่วยในการดึงแท็กที่ต้องการระบุช่วงวันที่ ซึ่ง tweepy มีข้อจำกัด ทำให้ไม่สามารถดึงข้อมูลจากวันที่เก่าๆได้
  3. from tweepy.auth import OAuthHandler : ใช้ในการจัดการกับ API และ Token ต่างๆที่
  4. pandas as pd ใช้ในการเก็บผลลัพธ์เป็น csv file และ แสดงผลลัพธ์ในรูปแบบของตาราง (Dataframe) เพื่อให้ง่ายต่อการดูข้อมูล
  5. time ใช้เพื่อดูเวลาที่ใช้ในการดึงข้อมูลได้รับมาจากTwitter



**รูปที่** 3.5 การ Import Library ที่ใช้

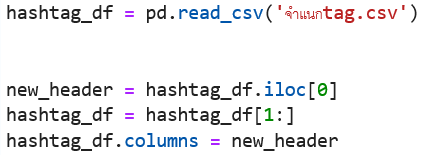
2.) ใช้ API key และ Access token ที่ได้รับจากTwitter มาใช้เพื่อดึงข้อมูลจาก twitter



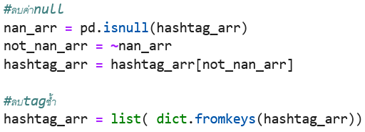
**รูปที่** 3.6 ใช้ API key และ Access token

3.) ตรวจสอบว่ามี tag มีอะไรบ้าง

* 1. เปิดไฟล์ที่ได้จำแนก tag ทวิตเตอร์ไว้ โดยใช้ Libary pandas ในการเปิดไฟล์ csv ออกมาเป็นรูปแบบของ DataFrame



**รูปที่** 3.7 Import Library เพื่อเปิดไฟล์ csv

* 1. ลบค่า null values และลบ tag ที่ซ้ำกัน

**รูปที่** 3.8 ลบค่า null values และ tag ที่ซ้ำ

จากการประมวลผลในรูปที่ 3.8 ได้ผลลัพธ์เป็นจำนวนแท็กทั้งหมด 33 แท็ก ดังนี้

**ตารางที่** 3.2 ผลลัพธ์จากการประมวลในรูปที่ 3.8

|  |  |  |  |
| --- | --- | --- | --- |
| #เราจะผ่านมันไปด้วยกัน | #โควิด19 | #เครียด | #โควิด19เราต้องรอด |
| #เคอร์ฟิว | #ไวรัสโคโรนา | #ปรึกษา | #เคอร์ฟิว24ชั่วโมง |
| #WFH | #การบินไทย | #ผิดหวัง | #การศึกษาฆ่าฉัน |
| #WorkFromHome | #ประกันสังคม | #เบื่อ | #dek63 |
| #อยู่บ้านหยุดเชื้อเพื่อชาติ | #พรกฉุกเฉิน | #ระบาย | #การศึกษา |
| #SocialDistancing | #ผนงรจตกม | #เราจะไม่ทิ้งกัน | #ฝุ่นPM25 |
| #เหนื่อยกับชีวิต | #เหนื่อย | #รัฐบาลเฮงซวย | #รัฐบาลส้นตีน |
| #หุ้นตก | #ปรึกษา | #ท้อ | #ซึมเศร้า |
| #อยากหลุดพ้น |  |  |  |

4.) ดึงข้อมูลจากtwitter

1.) ดึงข้อมูลแยกตามประเภท tag โดยใช้ โดยดึงข้อมูล tag ละ 50 ตัว จำนวน 32 Tag รวมเป็น 1,600 tweets

2.) โดยกำหนดข้อมูลที่ถูก tweet ตั้งแต่วันที่ 1 พฤษภาคม 2555 จนถึงวันที่ 5 พฤษภาคม 2563 และเป็นภาษาไทยเท่านั้น

2.1) Import เข้าสู่ Data Frame



**รูปที่** 3.9 Import ข้อมูลเข้าในรูปแบบของ Data Frame

2.2) Import ข้อมูลที่ได้เป็น csv file



**รูปที่** 3.10 บันทึกเป็นไฟล์ csv

5.) จำแนกข้อมูล

5.1) เมื่อได้เป็น csv File ออกมาแล้ว ต่อไปคือการจำแนกข้อมูลแยกตามประเภทของความเครียด 4 ประเภทของ Karl Albrecht โดยกำหนดให้ หมายเลข 1 คือ ความเครียดจากเวลา, หมายเลข 2 คือความเครียดจากสถานการณ์, หมายเลข 3 คือความเครียดจากการคากการณ์ล่วงหน้า และหมายเลข 4 คือความเครียดจากการเผชิญหน้า

5.2) บางข้อมูลสามารถแยกได้หลายประเภท และบางข้อมูลก็ไม่อยู่ในประเภทใดเลย ซึ่งเมื่อคณะผู้จัดทำได้ทำการจำแนกประเภทของแท็กแล้ว มีข้อมูลที่ใช้ได้เหลืออยู่เพียง 298 ค่า เหตุนี้เพราะข้อมูลที่ดึงมานั้นมีความซ้ำซ้อนกัน เนื่องจากภายในข้อความที่ถูกดึงออกมานั้น อาจมีจำนวนของแท็กมากกว่า 1 แท็ก และเป็นแท็กที่ทางคณะผู้จัดทำกำหนดที่จะดึงไว้ จึงทำให้ต้องลบตัวที่ซ้ำซ้อนออกไป และอีกหนึ่งสาเหตุคือ ข้อมูลที่ดึงออกมาไม่สามารถจำแนกออกเป็นประเภทได้ ทำให้ต้องลบออกไปเช่นกัน จำนวนข้อมูลเมื่อแยกตามประเภท ได้ผลลัพธ์คือ

5.2.1) ประเภทความเครียดแบบที่ 1 มีจำนวน 13 ข้อมูล

5.2.2) ประเภทความเครียดแบบที่ 2 มีจำนวน 159 ข้อมูล

5.2.3) ประเภทความเครียดแบบที่ 3 มีจำนวน 45 ข้อมูล

5.2.4) ประเภทความเครียดแบบที่ 4 มีจำนวน 81 ข้อมูล

จากข้อมูลข้างต้น จะเห็นได้ว่าข้อมูลเป็น imbalanced หมายถึงข้อมูลมีจำนวนคลาสที่แตกต่างกันมาก ทำให้การจำแนกประเภทข้อมูลไม่มีประสิทธิภาพ ดังเช่นในกรณีนี้มีข้อมูลที่ถูกจัดอยู่ในประเภทความเครียดแบบที่ 2 จำนวน 159 ข้อมูล แต่มีข้อมูลที่ถูกจัดอยู่ในประเภทความเครียดแบบที่ 1 เพียงจำนวน 13 ข้อมูลเท่านั้น วิธีจัดการกับข้อมูลที่เป็น imbalanced เช่น Sampling method เป็นวิธีการสุ่มตัวอย่าง เพื่อทำให้ข้อมูลมีความสมดุลกัน โดยอาจเลือกสุ่มลดข้อมูลให้มีจำนวนเท่าๆกันกับข้อมูลน้อย หรือสุ่มเพิ่มข้อมูลให้มีจำนวนเท่าๆกันกับข้อมูลมาก [37] แต่ทว่าข้อมูลที่ถูกจัดอยู่ในประเภทความเครียดแบบที่ 1 นั้นมีน้อยเกินไป ผู้วิจัยจึงดึงข้อมูลมาเพิ่มอีกจำนวน 467 ข้อมูล โดยเลือกดึงเฉพาะแท็กที่คาดว่าจะเกี่ยวข้องกับ ประเภทความเครียดแบบที่ 1, ประเภทความเครียดแบบที่ 3 และ ประเภทความเครียดแบบที่ 4 ดังตารางต่อไปนี้

**ตารางที่** 3.3 แท็กที่ใช้เพื่อดึงข้อมูลเพิ่มเติม

|  |  |  |  |
| --- | --- | --- | --- |
| #ปัญหาชีวิต | #ทะเลาะกับเพื่อน | #งานเยอะ | #ไปสาย |
| #ทำงานไม่ทัน | #กลัวทำไม่ได้ |  |  |

จำนวนข้อมูลเมื่อแยกตามประเภทหลังจากเพิ่มข้อมูลเข้าไป ได้ผลลัพธ์คือ

5.2.5) ประเภทความเครียดแบบที่ 1 มีจำนวน 121 ข้อมูล

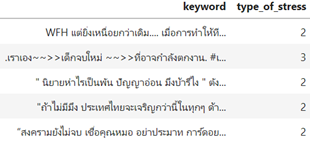
5.2.6) ประเภทความเครียดแบบที่ 2 มีจำนวน 160 ข้อมูล

5.2.7) ประเภทความเครียดแบบที่ 3 มีจำนวน 113 ข้อมูล

5.2.8) ประเภทความเครียดแบบที่ 4 มีจำนวน 125 ข้อมูล

จากข้อมูลข้างต้นจะเห็นได้ว่า จำนวนข้อมูลทั้ง 4 คลาส ไม่แตกต่างกันมาก

5.3) Save ข้อมูลที่ทำการจำแนกเป็น csv file

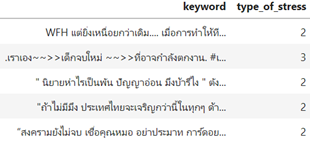


**รูปที่** 3.11 ตัวอย่างข้อมูล

**3.4 การวิเคราะห์ข้อมูล**

**3.4.1 Pre-processing ข้อมูล**

1.) Import ข้อมูลที่ได้ทำการจำแนกไว้



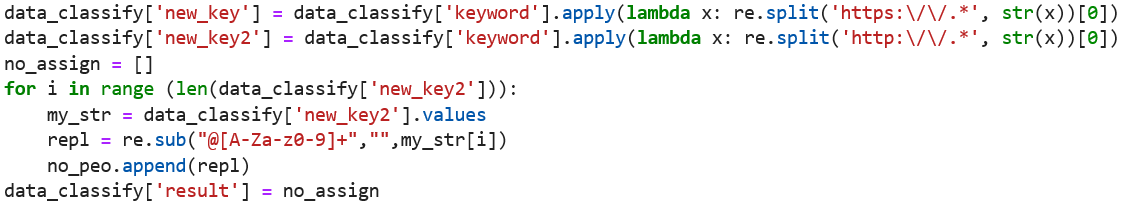
**รูปที่** 3.12 ข้อมูลที่ถูกจำแนกไว้

2.) Clean ข้อมูล เพื่อลบบางข้อความที่ไม่เกี่ยวข้อง

2.1) ลบคอลัมน์ข้อมูลที่มีค่า “type\_of\_stress” เป็น “-” ซึ่งหมายถึงไม่สามารถแยกประเภทข้อมูลนั้นๆได้



**รูปที่** 3.13 ลบข้อมูลที่ไม่เกี่ยวข้อง

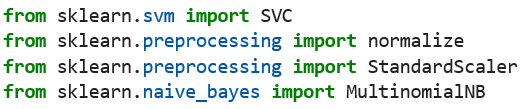
 2.2) ลบ URL ออกจากข้อมูล รวมถึง “@” ที่ผู้คนใช้ตอบกลับถึงอีกคนหนึ่ง

**รูปที่** 3.14 ลบคำต่างๆข้อมูลที่ไม่เกี่ยวข้อง

* 1. ลบคอลัมน์ที่ไม่ได้ใช้แล้ว และเปลี่ยนชื่อคอลัมน์

**รูปที่** 3.15 ลบคอลัมน์ที่ไม่ได้ใช้แล้ว และเปลี่ยนชื่อคอลัมน์

3.) Import Library ที่จะใช้ทำ Feature Scaling



**รูปที่** 3.16 Import Library ต่างๆที่ต้องใช้ในการ Pre-processing

3.1) Feauture Scaling

การจะปรับปรุงโมเดลให้ดีขึ้นได้นั้น จะใช้การปรับ scale ตัวแปรให้ใกล้เคียงกัน (Feauture Scaling) ซึ่งการปรับscaleจะช่วยลดความหลายหลายของตัวแปรต่างๆ ทั้งนี้ยังช่วยลด bias อีกด้วย [34] วิธีการทำ Feature Scaling ได้แก่

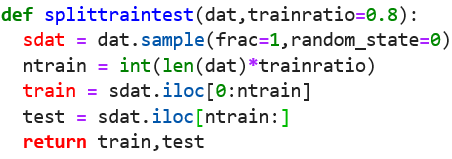
3.1.1) การ Normalization เป็นการ rescale ตัวแปร ให้อยู่ในช่วง [0,1] โดยในpython สามารถimport library “normalize” ได้ มีสมการคือ

norm\_X = (X - min\_X) / (max\_X - min\_X)

3.1.2) การ Standardize หรือเรียกอีกอย่างหนึ่งว่าการหาค่า z-score โดยในpython สามารถimport library “StandardScaler” ได้ มีสมการคือ

z\_score = (X - mean\_X) / sd\_X

4.) แบ่งข้อมูลออกสำหรับ Train 80%และTest 20% เหตุผลที่ต้องแบ่งข้อมูลออกเป็น Train Set และ Test Set เพื่อทดสอบความแม่นยำของโมเดล

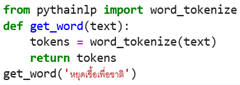


**รูปที่** 3.17 แบ่งข้อมูล

5.) ใช้ library pythainlp ซึ่งเป็น library ที่ช่วยในการตัดคำของภาษาไทย

5.1) สร้างฟังก์ชันที่ต้องการใช้งานดังนี้

5.1.1) Function “get\_word” ใช้สำหรับแยกภาษาไทยออกเป็นคำๆ

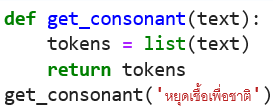


**รูปที่** 3.18 สร้างฟังก์ชั่น get world



**รูปที่** 3.19 ตัวอย่างการแยกคำ

5.1.2) Function “get\_consonant” ใช้สำหรับแยกภาษาไทยออกเป็นพยัญชนะ สระ และวรรณยุกต์

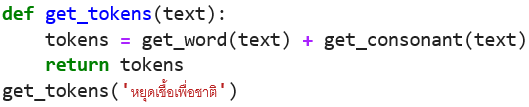


รูปที่ 3.20 สร้างฟังก์ชั่น get consonant



**รูปที่** 3.21 ตัวอย่างการแยกวรรณยุกต์ และสระ

5.1.3) Function “get\_tokens” เป็นfunctionที่รวมfunction “get\_word” และ “get\_consonant” เข้าด้วยกัน



**รูปที่** 3.22 สร้างฟังก์ชั่น get tokens

5.2) จากนั้นทำการแปลงข้อมูลเป็นเวกเตอร์โดยใช้ TfidfVectorizer และ CountVectorizer



**รูปที่** 3.23 แปลงข้อมูล

5.2.1) CountVectorizer และ TfidfVectorizer    
             Countvectorizer ใช้เพื่อนับจำนวนครั้งที่คำคำหนึ่งปรากฏขึ้น เรียกว่า term frequency เพื่อลดมิติข้อมูล ช่วยให้การประมวลผลข้อมูลมีประสิทธิภาพมากขึ้น

TfidfVectorizer หรือ TF-IDF ใช้เพื่อแยกคำที่ไม่ต้องการ หรือไม่จำเป็นต้องมีอยู่ในประโยคออกไปได้ เช่น stop words ต่างๆ และสามารถนำไปเข้าสู่กระบวนการทำText classification ได้

หลักการทำงานของ TfidfVectorizer หรือ TF-IDF

1. หา Term Frequency ซึ่งวิธีการคำนวณterm frequency มีหลายวิธีคือ
   1. Raw counts เป็นการนับจำนวนคำแต่ละคำที่มีอยู่ในเอกสารนั้นๆ มีสมการคือ f(t,d) โดย t คือterm หรือคำ และ d คือ document
   2. Term frequency เป็นการนำRaw count ที่ได้ไปหารกับจำนวนคำทั้งหมดในdocument
   3. Log normalization ช่วยnormalized ให้ค่าของข้อมูลไม่แตกต่างกันเกินไป มีสมการคือ log(1+f(t,d))
   4. Double Normalization ช่วยป้องกันไม่ให้เกิด bias เนื่องจากความยาวของdocument นั้นมีผลต่อการคำนวณ
2. หาค่า IDF (Inverse Document Frequency) ซึ่งเป็นค่าที่ช่วยหาความสำคัญของคำในdocument ซึ่งหากค่าของจำนวนคำที่พบน้อย ค่าIDFก็จะยิ่งสูง ซึ่งจะเป็นinverse กันกับค่าของ term frequency
3. จากนั้นให้นำค่า TF และ DF มาคูณกัน เพื่อให้น้ำหนักของTF และ IDF สามารถแยกคำสำคัญได้ [35][36]

5.3) กำหนดค่า train และ test ของข้อมูล



รูปที่ 3.24 กำหนดค่า train และ test ของ keyword



**รูปที่** 3.25 กำหนดค่า train และ test ของ intent

5.4) กำหนดช่วง N gram เป็น 3 และทำการ fit เพื่อให้ข้อมูลtrain สามารถทำนายผลลัพธ์ได้



**รูปที่** 3.26 กำหนดช่วง N-gram

5.5) จากนั้น transform ข้อมูลtrain และ test ของตัว keyword



**รูปที่** 3.27 transform ข้อมูลtrain และ test ของตัว keyword

**3.4.2 Text classification**

1.) Import library ของalgorithm ต่างๆ ที่ต้องใช้ ได้แก่

1.1) Naive Bayes

1.2) Support Vector Machine



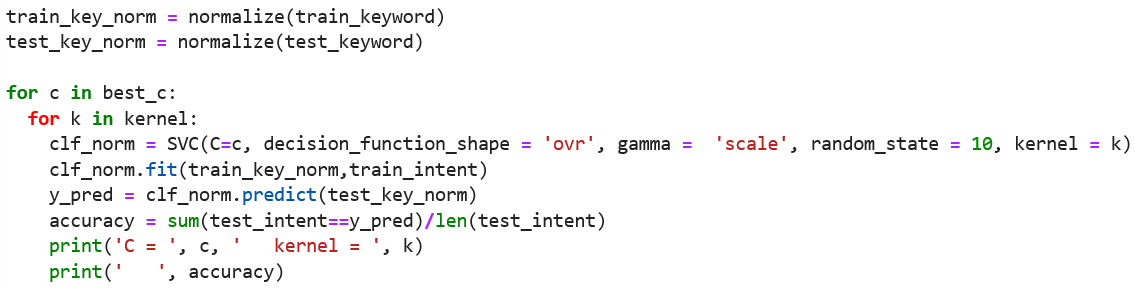
**รูปที่** 3.28 Import library ของalgorithm ต่างๆ

2.) ทดสอบอัลกอริทึม SVM

2.1) กำหนดค่าhyperparameter และkernel ที่ใช้

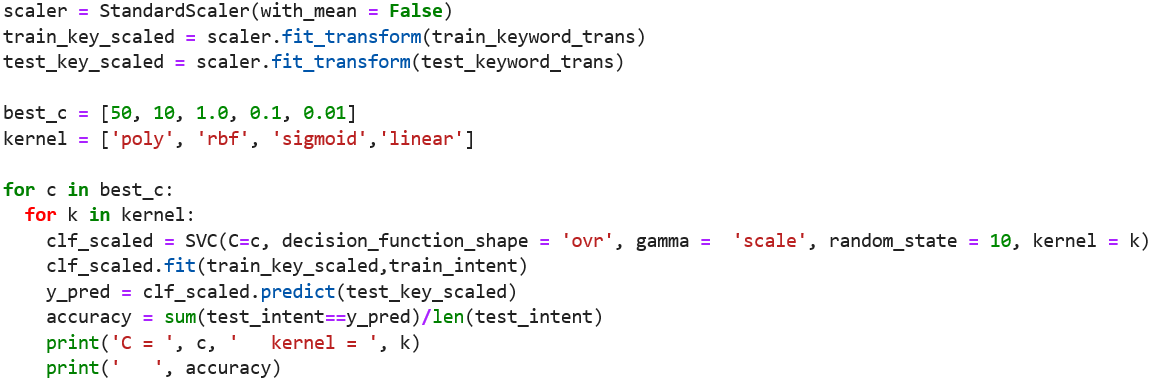


**รูปที่** 3.29 กำหนดค่าhyperparameter และkernel

2.2) Tune โมเดล เพื่อหาค่าhyperparameter ที่ดีที่สุด เมื่อทำNormalization

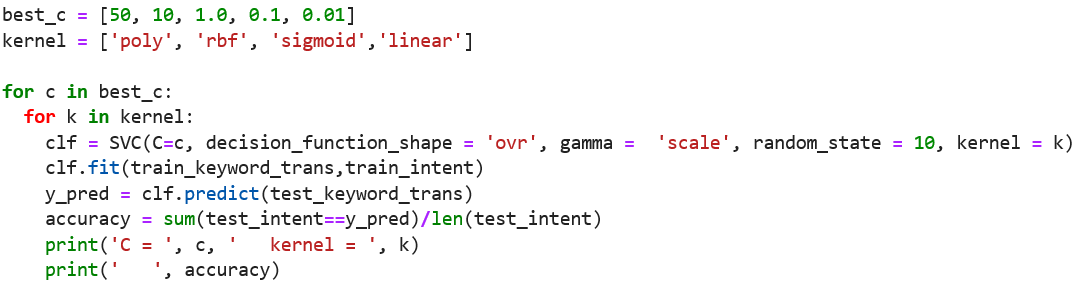
**รูปที่** 3.30 Tune โมเดล เพื่อหาค่าhyperparameter ที่ดีที่สุด

2.3) Tune โมเดล เพื่อหาค่าhyperparameter ที่ดีที่สุด เมื่อทำStandardization



**รูปที่** 3.31 Tune โมเดล เพื่อหาค่าhyperparameter ที่ดีที่สุด เมื่อทำ Standardization

2.3) Tune โมเดล เพื่อหาค่าhyperparameter ที่ดีที่สุด เมื่อทำไม่ได้ทำNormalizationหรือStandardization



**รูปที่** 3.32 Tune โมเดล เพื่อหาค่าhyperparameter ที่ดีที่สุด เมื่อทำไม่ได้ทำ feature scaling

3.) ทดสอบอัลกอริทึม Naïve Bayes

3.1) Import library เพิ่มเติม



**รูปที่** 3.33 Import library เพิ่มเติม

3.2) ใช้ CountVectorizer() นับจำนวนคำที่อยู่ในข้อมูล



**รูปที่** 3.34 ใช้CountVectorizer() นับจำนวนคำที่อยู่ในข้อมูล

3.3) ใช้ TfidTransformer() เพื่อเพิ่มความสำคัญของคำ



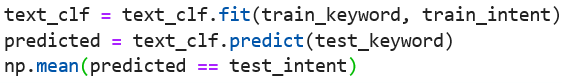
**รูปที่** 3.35 ใช้CountVectorizer() นับจำนวนคำที่อยู่ในข้อมูลเพื่อเพิ่มความสำคัญ

3.4) เริ่มใช้algorithm มาเทรนโมเดล และสร้างPipeline เพื่อจัดเตรียมข้อมูลให้อยู่ในรูปแบบที่เหมาะสม



**รูปที่** 3.36 เริ่มใช้algotithm มาเทรนโมเดล และสร้างPipeline

3.5) Fit ข้อมูลเทรน ทำการpredict แล้วหาความแม่นยำที่ได้



**รูปที่** 3.37 Fit ข้อมูลเทรน ทำการpredict แล้วหาความแม่นยำที่ได้

**บทที่ 4**

**ผลการทดลอง**

#### **4.1 ผลการทดลอง Model**

**4.1.1) Support Vector Machine (SVM)**

#### เปรียบเทียบผลลัพธ์จากการnormalization และ standardize สำหรับ Support Vector Machine algorithm

**ตารางที่** 4.1 เปรียบเทียบผลลัพธ์ของ SVM

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | C=50 | C=10 | C=1.0 | C=0.1 | C=0.01 |
| Non-standardize | Poly | 44.23 % | 44.23 % | 40.38 % | 24.04 % | 24.04 % |
|  | Rbf | 71.15 % | 71.15 % | 66.34 % | 24.04 % | 24.04 % |
|  | sigmoid | 72.11 % | 72.11 % | 75.00 % | 25.00 % | 24.04 % |
|  | linear | 73.07 % | 73.07 % | 75.00 % | 25.00 % | 24.04 % |
| Standardize | Poly | 27.88 % | 27.88 % | 24.03 % | 24.04 % | 24.04 % |
|  | Rbf | 54.80 % | 54.80 % | 28.84 % | 24.04 % | 24.04 % |
|  | Sigmoid | 71.15 % | 71.15 % | 37.5 % | 24.04 % | 24.04 % |
|  | linear | 71.15 % | 71.15 % | 71.15 % | 71.15 % | 71.15 % |
| Normalization | Poly | 44.23 % | 44.23 % | 40.38 % | 24.04 % | 24.04 % |
|  | Rbf | 71.15 % | 71.15 % | 66.34 % | 24.04 % | 24.04 % |
|  | Sigmoid | 72.11 % | 72.11 % | 75.00 % | 25.00 % | 24.04 % |
|  | linear | 73.07 % | 73.07 % | 75.00 % | 25.00 % | 24.04 % |

**4.1.2)** **Naïve Bayes**

สำหรับ Naïve Bayes ได้ผลลัพธ์ คือ 64.42 %

จากผลการทดลอง สามารถเห็นได้ว่า Algorithm ที่ทำการทำนายข้อมูลชุดนี้ได้ดีที่สุดคือ Support Vector Machine โดยกำหนดค่าต่างๆ ดังนี้

1.) Non-Standardize  
 1.1) C = 1.0, kernel = sigmoid

1.2) C = 1.0, kernel = linear

2.) Normalization

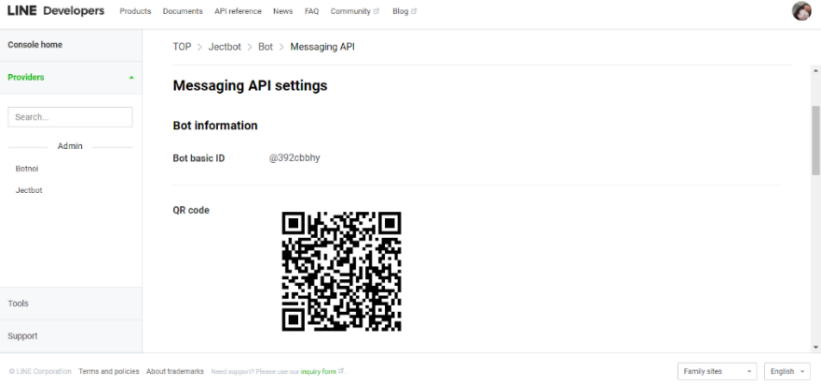
2.1) C = 1.0, kernel = sigmoid

2.1) C = 1.0, kernel = linear

ทั้ง 4 ตัว มีค่าความแม่นยำเท่ากัน คือ 75.00 % โ

**4.2 ผลการทดลองสร้างระบบแชทบอท**

ตัวอย่างระบบทดลองของแชทบอทบำบัดความเครียด



**รูปที่** 4.1 ระบบทดลองของแชทบอทบำบัดความเครียด

ตัวอย่างการพูดคุยระบบทดลองของระบบแชทบอทบำบัดความเครียด



**รูปที่** 4.2 การทดลองพูดคุยกับระบบทดลองของระบบบำบัดความเครียด

**บทที่ 5**

**สรุปผลการดำเนินงาน**

**5.1 สรุปผลการทดลอง**

ค่าความแม่นยำที่มากที่สุดคือ 75.00 % โดยผลลัพธ์นี้ ไม่เป็นที่น่าพอใจนัก โดยมีค่า confusion matrix คือ

**ตารางที่** 5.1 ตาราง confusion matrix

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | ความเครียดจากเวลา | ความเครียดจากสถานการณ์ | ความเครียดจากการคาดการณ์ล่วงหน้า | ความเครียดจากการเผชิญหน้า |
| ความเครียดจากเวลา | 22 | 2 | 3 | 2 |
| ความเครียดจากสถานการณ์ | 0 | 18 | 1 | 6 |
| ความเครียดจากการคาดการณ์ล่วงหน้า | 0 | 9 | 18 | 1 |
| ความเครียดจากการเผชิญหน้า | 0 | 5 | 0 | 17 |

จากตารางที่ 5.1 confusion matrix เป็นตารางที่ใช้วัดความสามารถในการจำแนกประเภทของ classification algorithm ระหว่างข้อมูลประเภทความเครียดที่เราใช้ทดสอบ (test) และผลลัพธ์ที่โมเดลทำนายออกมาได้ (predict) ซึ่งแต่ละเลขคือจำนวนที่โมเดลทำนายได้ เช่น ค่าความเครียดจากเวลา ได้ใช้ข้อมูลในการทดสอบจำนวน 29 ข้อมูล ซึ่งโมเดลสามารถทำนายได้ถูกต้องเพียง 22 ข้อมูลเท่านั้น ที่เหลืออีก 7 ข้อมูลคือทำนายผิดแยกไปตามประเภทความเครียดแบบต่างๆ โดยจะเห็นได้ว่าโมเดลยังทำนายไม่แม่นยำเท่าที่ควร โดยเฉพาะประเภทความเครียดที่ 1, 2 และ 3 การทำนายกระจายไปอยู่ส่วนอื่นๆค่อนข้างมาก

**5.2 วิเคราะห์ผลลัพธ์**

จากผลการทดลองที่ได้ค่าความแม่นยำออกมาค่อนข้างน้อย อาจเป็นเพราะว่า ข้อมูลที่มีอยู่นั้นไม่มากพอ และคลุมเครือ วิธีแก้ไขคือ เก็บข้อมูลให้เยอะขึ้นและหาประเภทความเครียดที่สามารถครอบคลุมได้มากกว่านี้

**5.3 การนำไปใช้**

จากการทำวิชาโครงงานในภาคการศึกษานี้ สามารถนำการทำ Text Classification ไปจำแนกตามความเครียดของแต่ละบุคคล เพื่อสร้าง intent ของ chatbot ย่อยลงไปให้ครอบคลุมมากขึ้น

**5.4 อุปสรรคที่เกิดขึ้น**

การดึงข้อมูลออกมาจาก Twitter ตามแท็กต่างๆที่กำหนดไว้ ทำให้มีข้อมูลที่ถูกดึงมาซ้ำกัน เพราะบางโพสต์ที่ดึงออกมาประกอบด้วยแท็กจำนวนหลายแท็ก ทำให้เกิดการซ้ำซ้อนของข้อมูล ซึ่งต้องทำการ Clean ข้อมูลก่อน และข้อมูลที่ดึงออกมาต้องนำมาจำแนกด้วยวิธีไล่ดูด้วยตัวเอง ซึ่งทำให้มีบางข้อมูลที่ดึงมาไม่สามารถจำแนกได้จึงต้องตัดออก ส่งผลให้จำนวนข้อมูลลดน้อยลงกว่าที่ตั้งไว้

**5.5 การวางแผนการทำงานในปีการศึกษา 2563 ภาคเรียนที่ 1 วิชาโครงงาน 2**

**5.5.1 ทำ Text Classification**

โดยนำข้อมูลจำแนกตาม Intent ย่อยๆ เพื่อปรับปรุงประสิทธิภาพในการจำแนกโดยการปรับประเภทในการแยกใหม่เป็น 6 ประเภท คือ

1.) Acute stress ความเครียดแบบเฉียบพลัน เป็นรูปแบบของความเครียดที่**ร**ุนแรงที่สุดซึ่งต้องการการตอบสนองทันทีต่อความท้าทายเหตุการณ์หรือความต้องการใหม่ของร่างกาย

2.) Episodic Acute stress การเกิดความเครียดเฉียบพลันบ่อยครั้ง ที่มักจะรู้สึกวิตกกังวลหงุดหงิด หรือมีอารมณ์แปรปรวนในระยะสั้น โดยผู้ที่เกิดความเครียดนี้จะมีทัศนะเชิงลบต่อทุกสิ่งและกังวลเรื่องเล็กๆ ทุกเรื่อง ซึ่งส่งผลกระทบด้านลบต่อสุขภาพ ผู้ที่เกิดความเครียดเหล่านี้ยอมรับว่าความเครียดเป็นส่วนหนึ่งของชีวิตและเป็นเรื่องยากสำหรับพวกเขาที่จะเปลี่ยนแปลงวิถีชีวิตของพวกเขา

3.) Chronic stress ความเครียดเรื้อรัง หากความเครียดเฉียบพลันยืดออกเป็นเวลานานจะกลายเป็นความเครียดเรื้อรัง ความเครียดประเภทนี้คงที่และไม่หายไปโดยง่าย ยกตัวอย่างเช่น ความเครียดที่เป็นผลมาจากปัญหาทางการเงิน การแต่งงานที่ไม่มีความสุข งานที่ไม่ดีหรือครอบครัวที่ไม่สมบูรณ์ เป็นต้น ซึ่งความเครียดเรื้อรังนี้เป็นอันตรายต่อสุขภาพและสามารถนำไปสู่โรคร้ายแรงต่างๆได้ ได้แก่ โรคหัวใจ และภาวะซึมเศร้า

4.) Phycical stress ความเครียดทางกายภาพ ความเครียดทางกายภาพอาจเกิดขึ้นจากการบาดเจ็บ อันมีสาเหตุมาจากเกิดการบาดเจ็บ การติดเชื้อหรือการผ่าตัดใดๆ การใช้แรงงานอย่างหนักมลภาวะต่อสิ่งแวดล้อม ออกซิเจนไม่เพียงพอต่อร่างกาย ร่างกายมีความอ่อนล้าอ่อนเพลีย ความไม่สมดุลของฮอร์โมน หรือ การขาดน้ำ เป็นต้น

5.) Phychological stress ความเครียดทางจิตวิทยา รวมถึงความเครียดทางอารมณ์ ที่เกิดจากความกลัว ความขุ่นมัว ความเศร้า ความโกรธ และความเครียดทางปัญญาจากความหึงหวงสิ่งที่แนบมาความวิตกกังวล การโจมตีทางจิตวทยาเพื่อให้เสียขวัญ หรือการวิจารณ์ตนเอง โดยทั่วไปหมายถึง ปฏิกิริยาทางอารมณ์และทางสรีรวิทยาที่เกิดขึ้นเมื่อบุคคลเจอสถานการณ์นั้นๆ

6.) Phychosocial stress ความเครียดทางจิตสังคม ความเครียดนี้มาจากความสัมพันธ์หรือปัญหาการแต่งงาน อาจรวมถึงความสัมพันธ์ของพนักงานกับนายจ้าง พี่กับน้องหรือครอบครัว การขาดการสนับสนุนทางสังคม การสูญเสียงาน การสูญเสียคนที่รัก และความโดดเดี่ยวอาจนำไปสู่ความเครียดประเภทนี้ได้

ซึ่งต่างจากของเดิมที่แบ่งเพียง 4 ประเภทเท่านั้น เพราะการแบ่งประเภทแบบเดิมอาจมีความหมายที่คลุมเครือและส่งผลทำให้ต้องลบข้อความที่ไม่ตรงกับประเภทใดๆออก ซึ่งความเป็นจริงอาจแยกประเภทได้

**5.5.2 กำหนดรูปแบบของคำและประโยค**

กำหนดรูปแบบของคำและประโยคที่ไว้ใช้สำหรับการตอบกลับ กับผู้ใช้งาน ซึ่งอ้างอิงได้จากการศึกษาแนวคิดและบทความที่เกี่ยวข้อง เพื่อเพิ่มประสิทธิภาพและความน่าเชื่อถือในคำตอบของการตอบกลับของบอท

**5.5.3 นำ Intent ใส่เข้าใน Dialogflow**

เพื่อใช้เป็นที่รับข้อความที่ได้รับมาจาก Chat Room ของ Line Chatbot ซึ่งรับข้อความผ่าน Messaging API ผ่าน webhook และจากนั้น Dialogflow จะทำหน้าที่เข้าใจความต้องการ (Intent) ที่อยู่ในข้อความที่รับมา และตัดสินใจว่า ความต้องการ (Intent) นั้นคืออะไร โดยใช้ Machine Learning (Classification model)

**5.5.4 ทดสอบระบบด้วยกลุ่มตัวอย่าง**

เพื่อวัดประสิทธิภาพของประโยคในการตอบกลับของบอท ต่อ ความรู้สึกของผู้ใช้งาน และเพื่อนำข้อผิดพลาด หรือจุดที่คาดไม่ถึงมาพัฒนาบอท ให้ดีมากยิ่งขึ้น

**บรรณานุกรม**

[1] Krungsri Plearn Plearn. **ยกระดับธุรกิจให้ง่ายขึ้นด้วย “Chatbot” สุดยอดผู้ช่วยอัจฉริยะแห่งยุค 4.0**

. ค้นเมื่อ 17 กุมภาพันธ์ 2563, จาก [https://www.krungsri.com/bank/th/plearn-plearn/chatbot](https://www.krungsri.com/bank/th/plearn-plearn/chatbot-%E0%B8%AA%E0%B8%B8%E0%B8%94%E0%B8%A2%E0%B8%AD%E0%B8%94-%E0%B8%9C%E0%B8%B9%E0%B9%89%E0%B8%8A%E0%B9%88%E0%B8%A7%E0%B8%A2-%E0%B8%AD%E0%B8%B1%E0%B8%88%E0%B8%89%E0%B8%A3%E0%B8%B4%E0%B8%A2%E0%B8%B0.html)

[2] ไม่ปรากฏชื่อผู้แต่ง. (2561). [**Chat Bot (แชทบอท) คืออะไร โปรแกรมสนทนาอัตโนมัติ**](https://www.mindphp.com/%E0%B8%9A%E0%B8%97%E0%B8%84%E0%B8%A7%E0%B8%B2%E0%B8%A1/240-ai-machine-learning/5766-chat-bot.html)

. ค้นเมื่อ 17 กุมภาพันธ์ 2563, จาก <https://www.mindphp.com/บทความ/240-ai-machine-learning/5766-chat-bot.html>

[3] [Petch Kruapanich](https://medium.com/@petchkruapanich?source=post_page-----3a32bee13ce3----------------------). (2651). **พัฒนาแชทบอทแบบ Rule-based approach VS AI based approach**

. ค้นเมื่อ 17 กุมภาพันธ์ 2563, จาก [https://medium.com/readmoreth/พัฒนาแชทบอทแบบ-rule-based-approach-vs-ai-based-approach-3a32bee13ce3](https://medium.com/readmoreth/%E0%B8%9E%E0%B8%B1%E0%B8%92%E0%B8%99%E0%B8%B2%E0%B9%81%E0%B8%8A%E0%B8%97%E0%B8%9A%E0%B8%AD%E0%B8%97%E0%B9%81%E0%B8%9A%E0%B8%9A-rule-based-approach-vs-ai-based-approach-3a32bee13ce3)

[4] ไม่ปรากฏชื่อผู้แต่ง. **What Are Bots? How Do Chatbots Work**. ค้นเมื่อ 21 กุมภาพันธ์ 2563, จาก <https://botscrew.com/blog/what-are-bots/>

[5] M[itusha Arya](https://chatbotslife.com/@mitushaarya?source=post_page-----d5a8689cf52f----------------------). (2562). **A brief history of Chatbots**. ค้นเมื่อ 17 กุมภาพันธ์ 2563, จาก <https://chatbotslife.com/a-brief-history-of-chatbots-d5a8689cf52f>  
[6] [Kittithuch Anantakan](https://medium.com/@kittithuch?source=post_page-----6774767c5aaf----------------------). (2562). **Chatbot ที่เรารู้จักมันเกิดขึ้นมาได้ยังไงกันนะ**. ค้นเมื่อ 17 กุมภาพันธ์ 2563, จาก [https://medium.com/convolab/chatbot ที่เรารู้จักมันเกิดขึ้นมาได้ยังไงกันนะ-6774767c5aaf](https://medium.com/convolab/chatbot%20ที่เรารู้จักมันเกิดขึ้นมาได้ยังไงกันนะ-6774767c5aaf)  
[7] [นพ. โกวิทย์ นพพร](https://www.samitivejhospitals.com/th/doctors/kowit-nopporn/). (2561). [**Mental Health ความเครียดสะสม เสี่ยงฆ่าตัวตาย**](https://www.samitivejhospitals.com/th/%e0%b8%84%e0%b8%a7%e0%b8%b2%e0%b8%a1%e0%b9%80%e0%b8%84%e0%b8%a3%e0%b8%b5%e0%b8%a2%e0%b8%94%e0%b8%aa%e0%b8%b0%e0%b8%aa%e0%b8%a1/). ค้นเมื่อ 28 กุมภาพันธ์ 2563, จาก [https://www.samitivejhospitals.com/th/ความเครียดสะสม/](https://www.samitivejhospitals.com/th/%E0%B8%84%E0%B8%A7%E0%B8%B2%E0%B8%A1%E0%B9%80%E0%B8%84%E0%B8%A3%E0%B8%B5%E0%B8%A2%E0%B8%94%E0%B8%AA%E0%B8%B0%E0%B8%AA%E0%B8%A1/)  
[8] Louisa Valvano. (2562). **Using affirmations to change your mindset.** ค้นเมื่อ 27 มีนาคม 2563, จาก https://www.stress.org.uk/using-affirmations-to-change-your-mindset/  
[9] Louisa Valvano. (2562). **Simple ways to change your mind.** ค้นเมื่อ 27 มีนาคม 2563, จาก https://www.stress.org.uk/simple-ways-to-change-your-mind/  
[10] Louisa Valvano. (2562). **Assess your breathing.** ค้นเมื่อ 27 มีนาคม 2563, จาก https://www.stress.org.uk/assess-your-breathing/  
[11] Louisa Valvano. (2562). **Deep breathinh 4-7-8 approach.** ค้นเมื่อ 27 มีนาคม 2563, จาก https://www.stress.org.uk/deep-breathing-4-7-8-approach/  
[12] Louisa Valvano. (2562). **Why do we find it hard to say no.** ค้นเมื่อ 27 มีนาคม 2563, จาก https://www.stress.org.uk/why-do-we-find-it-hard-to-say-no/  
[13] Louisa Valvano. (2562). **Learn to say no.** ค้นเมื่อ 27 มีนาคม 2563, จาก https://www.stress.org.uk/learn-to-say-no/https://www.stress.org.uk/learn-to-say-no/

[14] ปลั๊กไฟ. **Machine Learning**. ค้นเมื่อ 20 กุมภาพันธ์ 2563, จาก <https://ilog.ai/machine-learning-ml/?gclid=CjwKCAiAy9jyBRA6EiwAeclQhBDGM9a75zDXdH4Ag-Ml5BXuuLfZ1HFqr-QkJTWf_hEbwyn2hjMH_BoCA_oQAvD_BwE>

[15] DIGITAL VENTURES. (2561). **Natural Language Processing เทคโนโลยีเชื่อมโยงปัญญาประดิษฐ์กับมนุษย์ด้วย “ภาษา”**. ค้นเมื่อ 19 กุมภาพันธ์ 2563, จาก <http://www.dv.co.th/blog-th/get-to-know-natural-language-processing-nlp/>   
[16] MonkeyLearn. (2562). **Text Classification**. ค้นเมื่อ 8 พฤษภาคม 2563, จาก https://monkeylearn.com/text-classification/

[17] ผศ.วิภาวรรณ บัวทอง. (2557). **Naïve Bayes**. ค้นเมื่อ 8 พฤษภาคม 2563, จาก https://wipawanblog.files.wordpress.com/2014/06/chapter-7-naive-bayes.pdf   
[18] [Jaruwit Pratancheewin](https://www.glurgeek.com/author/jaruwit-ptcw/). (2562). **เรียนรู้และทำความเข้าใจเรื่อง Support Vector Machine (SVM) คืออะไร**. ค้นเมื่อ 8 พฤษภาคม 2563, จาก https://www.glurgeek.com/education/support-vector-machine/

[19] ejeepss. (2550). **N-Gram**. ค้นเมื่อ 18 กุมภาพันธ์ 2563, จาก https://ejeepss.wordpress.com/2007/06/20/n-gram/

[20] [Petch Kruapanich](https://medium.com/@petchkruapanich?source=post_page-----4bd3a8c550de----------------------) . (2561). **ลองทำแชทบอทง่ายๆด้วย Dialogflow กันเถอะ**. ค้นเมื่อ 19 กุมภาพันธ์ 2563, จาก [https://medium.com/readmoreth/ลองทำแชทบอทลงทะเบียนง่ายๆด้วย-dialogflow-กันเถอะ-4bd3a8c550de](https://medium.com/readmoreth/ลองทำแชทบอทลงทะเบียนง่ายๆด้วย-dialogflow-กันเถอะ-4bd3a8c550de%20)

[21] ไม่ปรากฏชื่อผู้แต่ง. **Heroku**. ค้นเมื่อ 21 กุมภาพันธ์ 2563, จาก <https://www.heroku.com/what>  
[22] ไม่ปรากฏชื่อผู้แต่ง. (2562). **ไลน์ออฟฟิเชี่ยลแอคเคานต์ คืออะไร?.** ค้นเมื่อ 8 พฤษภาคม 2563, จากhttps://digitalcontent.top/knowledge/what-is-a-line-official-account/   
[23] กระปุกดอทคอม. (2556). **Line คืออะไร มารู้จัก ไลน์ กัน - ไลน์ คืออะไร วิธีสมัคร line ข่าวไลน์**. ค้นเมื่อ 8 พฤษภาคม 2563, จาก https://line.kapook.com/view64457.html   
[24] Job\_The\_Gamer. (2561). **Anaconda Distribution เครื่องมือจัดการแพคเกจ Python ยอดนิยม, ประกาศรวม VS Code เข้ามาในตัวติดตั้ง**. ค้นเมื่อ 8 พฤษภาคม 2563, จาก <https://www.blognone.com/node/100302>

[25] [ผศ.ดร.สายชล ใจเย็น.](https://medium.com/@saichonjaiyen?source=post_page-----4dbf02068792----------------------) (2561). **การติดตั้ง Anaconda บน Windows**. ค้นเมื่อ 8 พฤษภาคม 2563, จาก <https://medium.com/@saichonjaiyen/%E0%B8%81%E0%B8%B2%E0%B8%A3%E0%B8%95%E0%B8%B4%E0%B8%94%E0%B8%95%E0%B8%B1%E0%B9%89%E0%B8%87-anaconda-%E0%B8%9A%E0%B8%99-windows-4dbf02068792>

[26] [Wikipedia Contributors.](https://medium.com/@saichonjaiyen?source=post_page-----4dbf02068792----------------------) (2563). **Anaconda (Python distribution)**. ค้นเมื่อ 8 พฤษภาคม 2563, จาก https://en.wikipedia.org/wiki/Anaconda\_(Python\_distribution)

[27] [Github Contributors.](https://medium.com/@saichonjaiyen?source=post_page-----4dbf02068792----------------------) (2563). **The Qt Console for Jupyter**. ค้นเมื่อ 8 พฤษภาคม 2563, จาก <https://qtconsole.readthedocs.io/en/latest/>

[28] [ไม่ปรากฏชื่อผู้แต่ง.](https://medium.com/@saichonjaiyen?source=post_page-----4dbf02068792----------------------) (2560). **รู้จักกับ Visual Studio Code (วิชวล สตูดิโอ โค้ด) โปรแกรมฟรีจากค่ายไมโครซอฟท์** . ค้นเมื่อ 8 พฤษภาคม 2563, จาก <https://www.mindphp.com/%E0%B8%9A%E0%B8%97%E0%B8%84%E0%B8%A7%E0%B8%B2%E0%B8%A1/microsoft/4829-visual-studio-code.html>

[29] Chai Phonbopit. (2558). **Git คืออะไร? + พร้อมสอนใช้งาน Git และ Github**. ค้นเมื่อ 8 พฤษภาคม 2563, จาก <https://devahoy.com/blog/2015/08/introduction-to-git-and-github/>

[30] Thanatcha Kromsang. (2560). **เรียนรู้ Git และ Github ฉบับเด็กมหาลัย**. ค้นเมื่อ 8 พฤษภาคม 2563, จาก <https://medium.com/@thanatchakromsang/%E0%B9%80%E0%B8%A3%E0%B8%B5%E0%B8%A2%E0%B8%99%E0%B8%A3%E0%B8%B9%E0%B9%89-git-%E0%B9%81%E0%B8%A5%E0%B8%B0-github-%E0%B8%89%E0%B8%9A%E0%B8%B1%E0%B8%9A%E0%B9%80%E0%B8%94%E0%B9%87%E0%B8%81%E0%B8%A1%E0%B8%AB%E0%B8%B2%E0%B8%A5%E0%B8%B1%E0%B8%A2-7311034c6527>

[31] [KORBIN BROWN.](https://medium.com/@saichonjaiyen?source=post_page-----4dbf02068792----------------------) (2562). **What Is GitHub, and What Is It Used For?**. ค้นเมื่อ 8 พฤษภาคม 2563, จาก <https://www.howtogeek.com/180167/htg-explains-what-is-github-and-what-do-geeks-use-it-for/>  
[32] Patty Mulder[.](https://medium.com/@saichonjaiyen?source=post_page-----4dbf02068792----------------------) (2561). **4 Types of Stress by Karl Albrecht**. ค้นเมื่อ 8 พฤษภาคม 2563, จาก <https://www.toolshero.com/human-resources-hr/4-types-of-stress-by-karl-albrecht/>

[33] ดร.ปิยนันท์ สวัสดิ์ศฤงฆาร[.](https://medium.com/@saichonjaiyen?source=post_page-----4dbf02068792----------------------) (2562). **4 Types of Stress [Karl Albrecht]**. ค้นเมื่อ 8 พฤษภาคม2563, จาก https://www.toolshero.com/human-resources-hr/4-types-of-stress-by-karl-albrecht/  
[34] Kasidis Satangmongkol. (2562). **รีวิวเทคนิค Normalization และ Standardization**. ค้นเมื่อ 8 พฤษภาคม 2563, จาก <https://datarockie.com/2019/11/07/comparison-normalization-standardization/>

[35] [CHAKRIT.](https://medium.com/@saichonjaiyen?source=post_page-----4dbf02068792----------------------) (2562). **TF-IDF ทำงานอย่างไร**. ค้นเมื่อ 8 พฤษภาคม 2563, จาก <https://www.softnix.co.th/2019/05/28/tf-idf-%E0%B8%97%E0%B8%B3%E0%B8%87%E0%B8%B2%E0%B8%99%E0%B8%A2%E0%B8%B1%E0%B8%87%E0%B9%84%E0%B8%87/>

[36] [lukkiddd.](https://medium.com/@saichonjaiyen?source=post_page-----4dbf02068792----------------------) (2560). **TF-IDF คำไหนสำคัญนะ?**. ค้นเมื่อ 8 พฤษภาคม 2563, จาก <https://www.softnix.co.th/2019/05/28/tf-idf-%E0%B8%97%E0%B8%B3%E0%B8%87%E0%B8%B2%E0%B8%99%E0%B8%A2%E0%B8%B1%E0%B8%87%E0%B9%84%E0%B8%87/>

[37] Weerasak Thachai[.](https://medium.com/@saichonjaiyen?source=post_page-----4dbf02068792----------------------) (2559). **จัดการข้อมูล Imbalanced ใน Scikit-learn**. ค้นเมื่อ 8 พฤษภาคม 2563, จาก https://medium.com/espressofx-notebook/%E0%B8%88%E0%B8%B1%E0%B8%94%E0%B8%81%E0%B8%B2%E0%B8%A3%E0%B8%82%E0%B9%89%E0%B8%AD%E0%B8%A1%E0%B8%B9%E0%B8%A5-imbalanced-%E0%B9%83%E0%B8%99-scikit-learn-c22f4c18ebb5