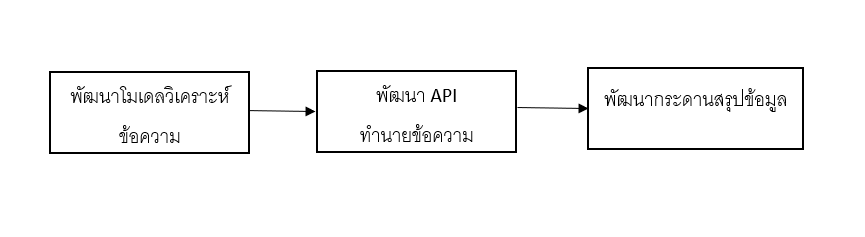
**บทที่ 3**

**วิธีการดำเนินโครงการ**

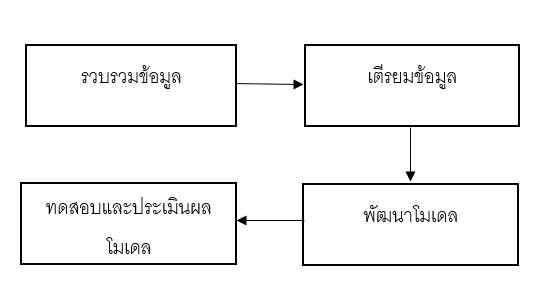
ในการจัดทำโครงการเรื่อง การวิเคราะห์ความคิดเห็นของนักเรียน นักศึกษาวิทยาลัยอาชีวศึกษานครปฐมบนสื่อสังคมออนไลน์โดยใช้การประมวลผลภาษาธรรมชาติ ประกอบด้วยการทำงานหลัก 3 ส่วน ได้แก่ การพัฒนาโมเดลวิเคราะห์ข้อความการพัฒนา APIและการพัฒนากระดานสรุปข้อมูล(Dashboard) บนเว็บแอปพลิเคชัน ในบทนี้จะเสนอรายละเอียดการดำเนินงานตามผังการดำเนินงานภาพที่ 3-1



ภาพที่ 3-1\*\*การดำงานหลักของโครงการ

3.1 การพัฒนาโมเดลวิเคราะห์ข้อความ

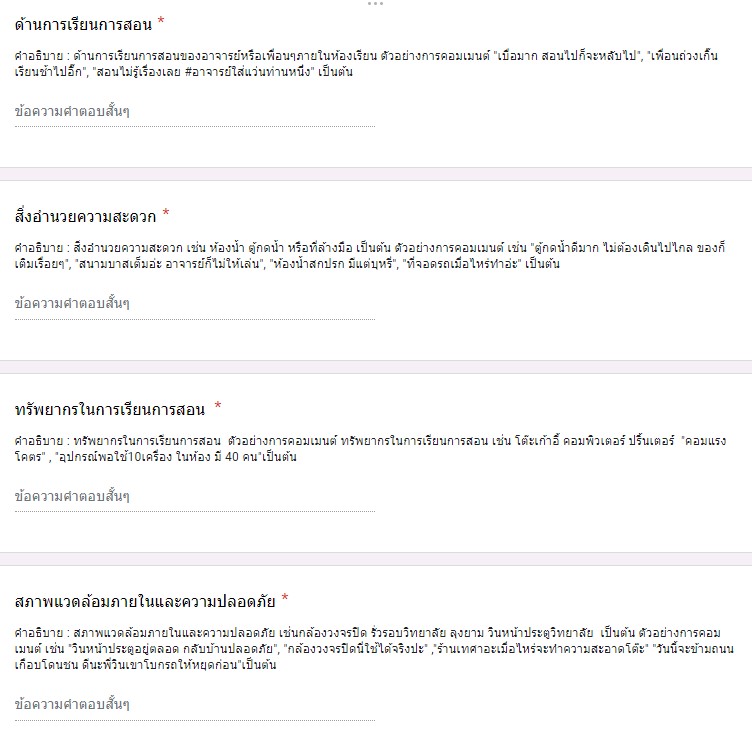
การพัฒนาโมเดลวิเคราะห์ข้อความ ประกอบด้วยขั้นตอนการรวบรวมข้อมูล การเตรียมข้อมูล การพัฒนาโมเดลจำแนกข้อความ (Text Classification) ตลอดจนการทดสอบและประเมินโมเดล ซึ่งมีผังการดำเนินงานดังภาพที่ 3-2



ภาพที่ 3-2\*\*การพัฒนาโมเดลวิเคราะห์ข้อความ

3.1.1 การรวบรวมข้อมูล

คณะผู้จัดทำได้ทำการรวบรวมข้อมูลจากนักเรียนนักศึกษาวิทยาลัยอาชีวศึกษานครปฐม จำนวน 200 คน มีจำนวนข้อมูล 800 ข้อความโดยใช้แอปพลิเคชันกูเกิ้ลฟอร์ม (Google Form) ซึ่งจำแนกประแภทข้องข้อความ 4 ด้าน ได้แก่ ด้านการเรียนการสอน ด้านสิ่งอำนวยความสะดวก ด้านทรัพยากรในการเรียนการสอน ตลอดจนด้านสภาพแวดล้อมและความปลอดภัย จำนวนหัวข้อละ 200 ข้อความ ข้อความส่วนใหญ่ที่รวบรวมจะใช้ภาษาที่ไม่เป็นทางการ อาจมีโครงสร้างประโยคที่ไม่ถูกต้องตามไวยกรณ์ภาษาไทยผ่านแบบฟอร์ม ดังภาพที่ 3-3

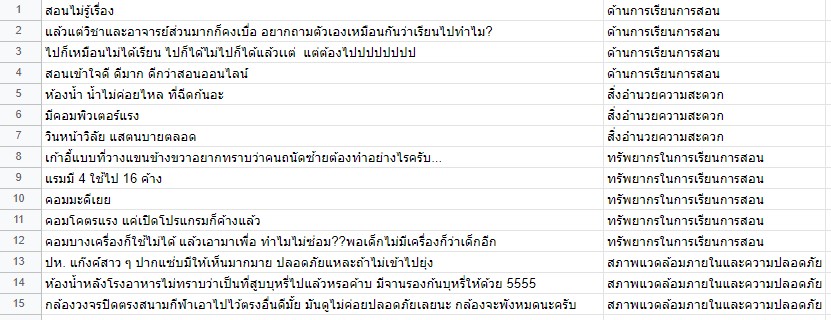


ภาพที่ 3-3\*\*การรวบรวมข้อมูล

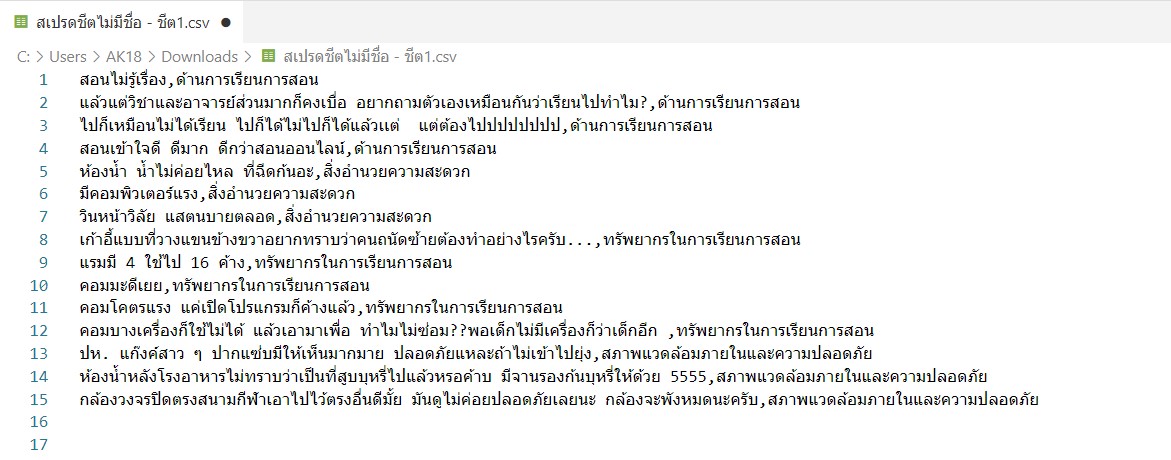
3.1.2 การเตรียมข้อมูล

ขั้นตอนการเตรียมข้อมูลจะประกอบด้วยการเตรียมข้อมูล 2 ส่วน ได้แก่ การเตีรยมข้อมูลสำหรับโมเดล และการเตรียมข้อมูลที่ดึงจากทวิตเตอร์เพื่อนำมาจำแนกประเภทข้อความ และวิเคราะห์ความรู้สึกของข้อความ

3.1.2.1 การเตีรยมข้อมูลสำหรับโมเดล เบื้องต้นคณะผู้จัดทำได้คัดกรองข้อความที่มีเนื้อหาไม่เหมาะสม เป็นค่าว่าง ไม่สื่อความหมาย และไม่ตรงตามหัวข้อที่กำหนด นำข้อมูลนามสกุลไฟล์ .xls ภาพที่ 3-4 ให้อยู่ในรูปแบบนามสกุลไฟล์ .csv ภาพที่ 3-5



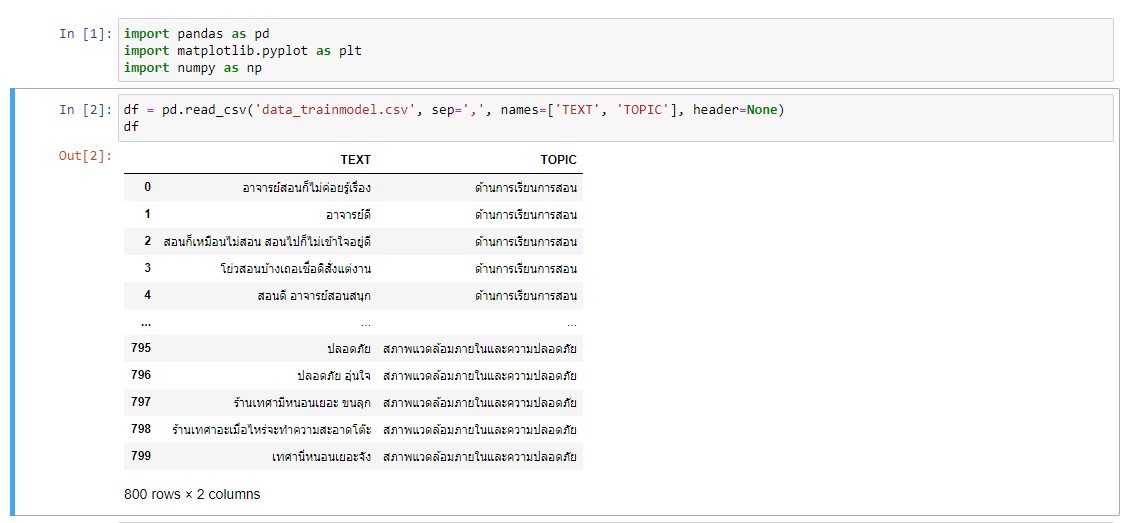
ภาพที่ 3-4\*\*การเตรียมข้อมูล – นามสกุลไฟล์ .xls



ภาพที่ 3-5\*\*การเตรียมข้อมูล – นามสกุลไฟล์ .csv

จากนั้นจึงนำข้อความทั้งหมดผ่านกระบวนการการประมวลภาษาธรรมชาติ (Natural Language Processing) ได้แก่ การตัดคำที่ไม่สื่อความหมาย (Stop word) ตลอดจนตัดสัญลักษณ์ต่างๆ และนำข้อความไปแปลงให้อยู่ในรูปของเวกเตอร์ (Word Vectorization) และจัดทำ Bag-of-Words (BOW) ซึ่งคณะผู้จัดทำใช้ภาษา Python ในการพัฒนา โดยมีรายละเอียดดังต่อไปนี้

1.\*\*ทำการอ่านข้อมูลจากไฟล์ CSV ที่เก็บข้อความดังกล่าวเข้าตัวแปร เมื่อแสดงออกมาดู จะประกอบด้วย 2 ฟิลด์ คือ TEXT และ TOPIC ดังภาพที่ 3-6



ภาพที่ 3-6\*\*การเตรียมข้อมูลโดยใช้ภาษา Python

2.\*\*นำฟิลด์ TOPIC ลงรหัส (Encode) ให้อยู่ในรูปแบบของตัวเลข ภาพที่ 3-7



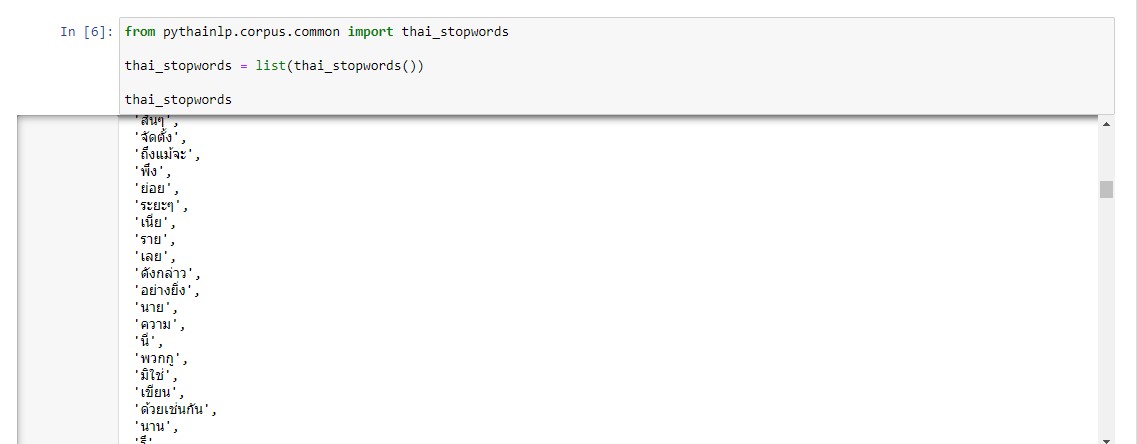
ภาพที่ 3-7\*\*การเตรียมข้อมูลโดยใช้ภาษา Python

โดยฟิลด์ TOPIC\_ID หมายเลขต่างๆ จะแทนหัวข้อความคิดเห็นดังตารางที่ 3-

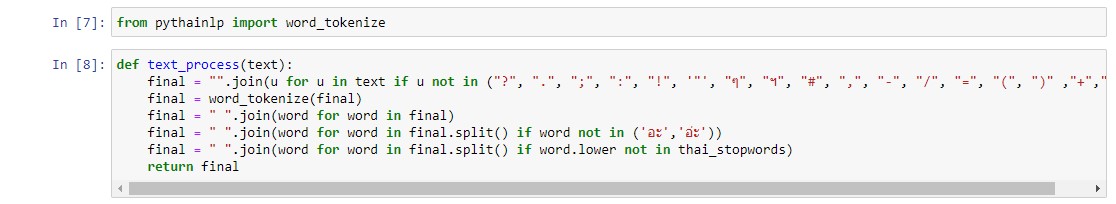
**ตารางที่ 3-...** การลงรหัสหัวข้อความคิดเห็นให้เป็นตัวเลข

|  |  |
| --- | --- |
| **หัวข้อความคิดเห็น** | **TOPIC\_ID** |
| ด้านการเรียนการสอน | 0 |
| ด้านสิ่งอำนวยความสะดวก | 1 |
| ด้านทรัพยากรในการเรียนการสอน | 2 |
| ด้านสภาพแวดล้อมและความปลอดภัย | 3 |

3.\*\*นำไลบรารี่ PyThaiNLP มาใช้ในการตัดคำที่ไม่สื่อความหมาย เช่น ฉะนั้น เมื่อนั้นและ จะ เป็นต้น จากนั้นสร้างฟังก์ชั่นในการตัดคำ และเครื่องหมายวรรคตอนดังภาพที่ 3-8 และ ภาพที่ 3-9

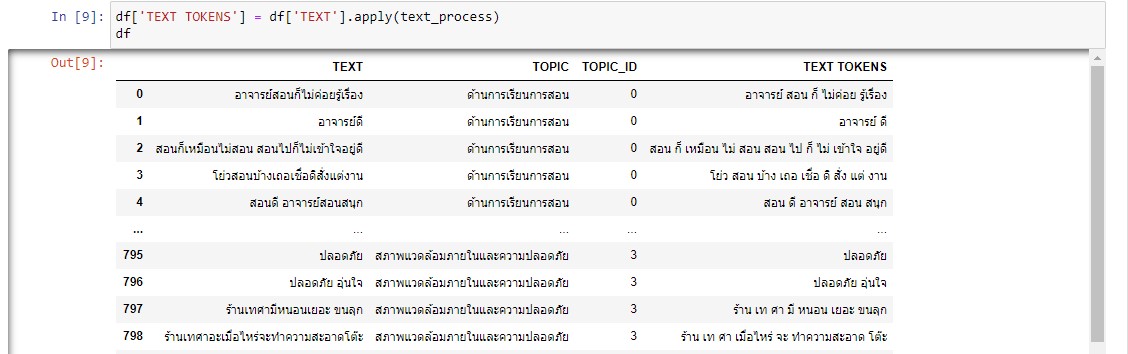


ภาพที่ 3-8\*\*รายการคำ Stop Word ที่จะนำมาใช้ในการตัดคำด้วยฟังก์ชัน thai\_stopwords



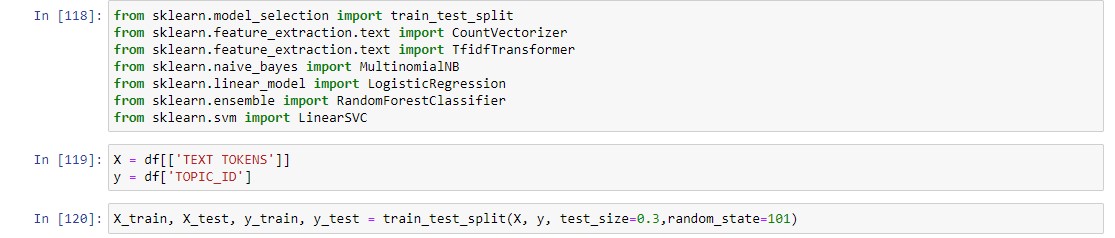
ภาพที่ 3-9\*\*การตัดคำด้วยฟังก์ชัน thai\_stopwords ด้วยไลบรารี่ PyThaiNLP

4.\*\*นำฟิลด์ TEXT มาตัดคำและเครื่องหมายวรรคตอน จากนั้นจึงเก็บไว้ที่ฟิลด์ TEXT TOKENS ดังภาพที่ 3-10



ภาพที่ 3-10\*ผลการใช้ไลบรารี่ thai\_stopwords ในการตัดคำและเครื่องหมายวรรคตอนในข้อความ

5.\*\*แบ่งชุดข้อมูลออกเป็น 2 ส่วน ได้แก่ ชุดข้อมูลสำหรับฝึกโมเดล (Traing Set) และชุดข้อมูลสำหรับทดสอบโมเดล (Testing Set) โดยสลับตำแหน่งเพื่อสุ่มข้อความ และแบ่งชุดข้อมูลในอัตราส่วน 70:30 ดังภาพที่ 3-11 โดยผู้จัดทำใช้ไลบรารี่ Scikit Learn ในการแบ่งชุดข้อมูล (train\_stest\_split)



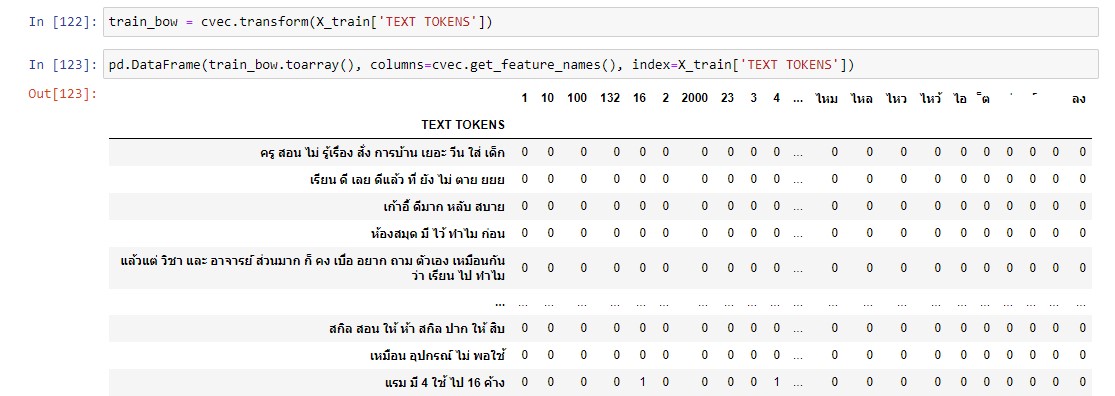
ภาพที่ 3-11\*\*การแบ่งชุดข้อมูลสำหรับฝึก และทดสอลโมเดล

6.\*\*Word Vectorizer และ Bag-of-Words (BOW) โดยเริ่มจากใช้ Count Vectorizer ของ Scikit Learn มาช่วยในการจัดการดึงคำทั้งหมดออกมาจากข้อความ และจัดเก็บในรูปแบบ Vector ดังภาพที่ 3-12



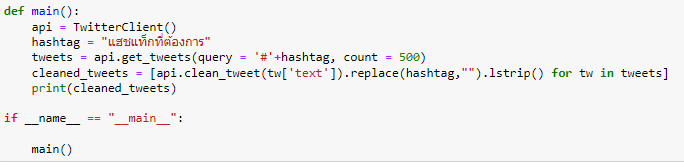
ภาพที่ 3-12\*\*จัดการดึงคำทั้งหมดออกมาจากข้อความ

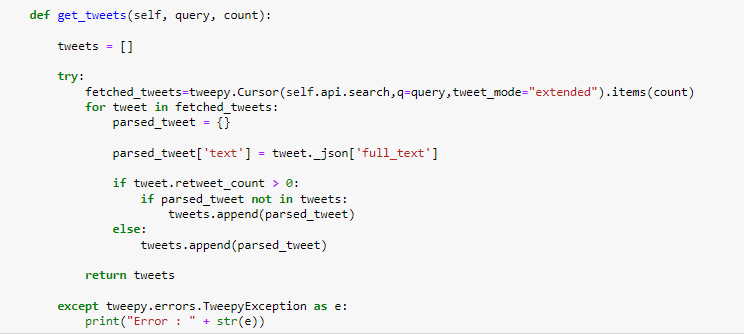
จากนั้นทำการสร้าง Bag-of-Words (BOW) ที่เปรียบเสมือนกับตารางที่มีแถวเป็นข้อความ คอลัมน์เป็นคำทั้งหมด และค่าคือจำนวนคำที่ปรากฏในข้อความดังกล่าว ภาพที่ 3-13



ภาพที่ 3-13\*\*ตารางคอลัมน์ที่เป็นคำทั้งหมด

3.1.2.2 การเตรียมข้อมูลที่ดึงจากทวิตเตอร์ประกอบด้วย ในส่วนของการรวบรวมข้อมูลจากทวิตเตอร์นั้น คณะผู้จัดทำได้ใช้ไลบรารี่ Tweepy ในการค้นหาทวีตตามแฮชแท็กที่ต้องการ ดังรูปที่ 3-11(มั่วไม่ให้หลง) มีฟังก์ชันการทำงานดังรูปที่ 3-12 และการเตรียมข้อมูลทวีตเพื่อให้อยู่ในรูปแบบข้อมูลที่เหมาะสมสำหรับการจำแนกประเภทของข้อความ และวิเคราะห์ความรู้สึกจากข้อความต่อไป ได้แก่ การตัดกำจัดเครื่องหมายวรรคตอน การตัดส่วนที่อ้างถึงผู้ใช้อื่น (Mention) การตัดข้อความที่เป็นจุดเชื่อมโยง (URL) และสัญลักษณ์การแสดงอารมณ์ (Emoji) ดังรูปที่ 3-13



ภาพที่ 3-11 การค้นหาทวีตตามแฮชแท็กที่ต้องการ เพื่อนำเข้ากระบวนการตัดคำ

ภาพที่ 3-11 ฟังก์ชันการทำงานของการดึงข้อมูลทวีต



ภาพที่ 3-12 ฟังก์ชันการเตรียมข้อมูลทวีตที่นำเข้า

3.1.3 การพัฒนาโมเดลการจำแนกประเภทข้อความ (Topic Classification)

คณะผู้จัดทำได้เลือก 4 อัลกอริทึมมาใช้ในการเทรนโมเดล ได้แก่ Random Forest, Linear SVM, Naive Bayes และ Logistic Regression มีรายละเอียดอัลกอริทึมดังนี้

3.1.3.1 Random Forest เป็นอัลกอรึธึมที่จะสร้างกฏขึ้นมาเองโดยการสุ่ม เช่นถ้ามี y = {0,1} และ x เป็นตัวเลข 0-100 เขียนเป็นฟังก์ชั่นได้ดังนี้ y = f(x) กฎการทำนาย(Prediction) สามารถแสดงได้ ดังภาพที่ 3-......

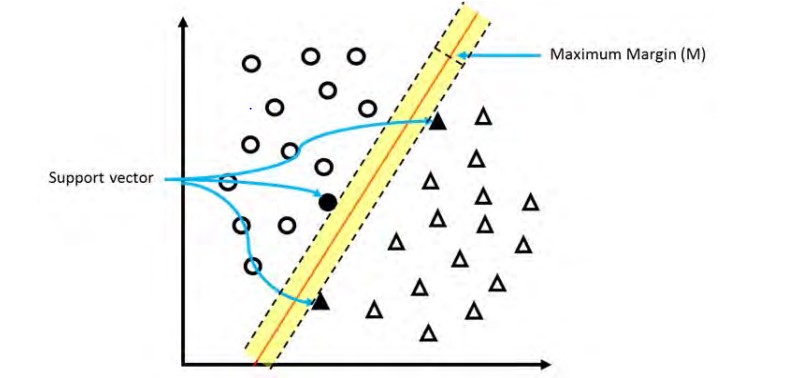
Text

Description automatically generated with medium confidence

ภาพที่ 3-…. กฎที่สุ่มโดยอัลกอริทึมของ Random Forest

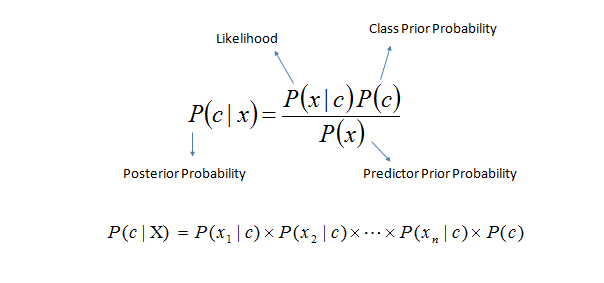
กฎดังภาพที่ 3-.... เกิดมาจากกระบวนการเรียนรู้ ของ Decision Tree ส่วน Random Forest คือ การสร้างโมเดลจาก Decision Tree หลายโมเดลย่อยๆ ตั้งแต่ 500 จนถึง หรือมากกว่า 1,000 โมเดลเพื่อช่วยในการทำนายค่า y ตอนทำนายก็ให้แต่ละ Decision Tree ทำนายของใครของมัน และผลลัพธ์ที่ทำนายได้จะถูกเลือกโดยการโหวตของ Decision Tree ที่อยู่ภายใน Random Forest

3.1.3.2 Linear SVM เป็นเทคนิครูปแบบหนึ่งในการจำแนกข้อมูลบนปริภูมิคุณลักษณะ (Features Space) เป็นสองกลุ่มด้วยระนาบ โดยระนาบนี้จะมีขนาดคู่ขนาดอยู่ทั้งสองข้าง โดยที่ระนาบคู่ชนาดนี้จะสัมผัสกับข้อมูลที่อยู่ใกล้ระนาบหลักที่สุด(Support Vector) และมีระยะห่าง(Margin) จากระนาบหลักมากที่สุด (Maximum Margin) ตัวอย่างเช่นถ้าเรามีข้อมูลที่เราทำการเทรน แล้วแบ่งเป็น 2 กลุ่มโดยเราจะเลือกใช้เส้น Linear Model ที่ไม่จำรูปแบบของข้อมูลที่ฝึกมากเกินไป(Over Fitting) โดยที่เส้นโมเดลที่เราเลือกจะมีระยะห่างระหว่าง 2 คลาสห่างกันมากที่สุด ดังภาพที่ 3-14



ภาพที่ 3-14\*\* Linear SVM

3.1.3.3 Naive Bayes เป็นอัลกอริทึมในการจำแนกหหมวดหมู่โดยใช้หลักความ น่าจะเป็น ที่สามารถคาดการณ์ผลลัพธ์และสามารถอธิบายได้โดยทำการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรเพื่อใช้ในการสร้างเงื่อนไขความน่าจะเป็น ดังสมการที่ 3-1



(3-1)

คำอธิบายสมการ แทนตัวแปร 3 ตัว ดังนี้

ก) c หมายถึง Class

ข) X หมายถึง Attribute

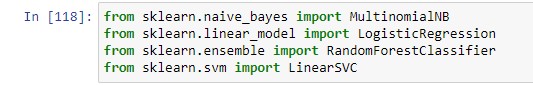
ค) P หมายถึง Probability (ความน่าจะเป็น)

ง) P(c|x) Posterior probability หมายถึง ความน่าจะเป็นที่ข้อมูลที่มี แอตทริบิวต์เป็น x จะมีคลาส C P(x|c) Likelihood คือ ความน่าจะเป็นที่ข้อมูลที่มีคลาส C และมีแอตทริบิวต์ x

จ) P(c) Prior probability คือ จำนวน Classที่อาจจะเกิดขึ้น/จำนวนClass ทั้งหมด หรือความน่าจะเป็นของ Class C

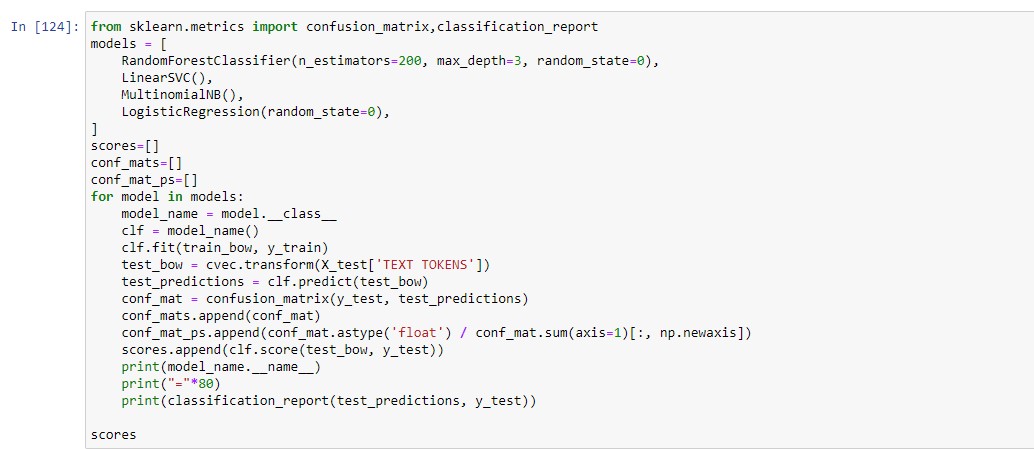
ฉ) P(x) Predictor Prior probability คือ จำนวน Attribute ทั้งหมด

3.1.3.4 Logistic Regression การวิเคราะห์การถดถอยเป็นเทคนิคการสร้างแบบจําลองเชิงคาดการณ์ชนิดหนึ่งซึ่งใช้ในการหาความสัมพันธ์ระหว่างตัวแปรที่ขึ้นต่อกัน (โดยปกติเรียกว่าตัวแปร "Y" ) และตัวแปรอิสระหนึ่งตัว (ตัวแปร "X" ) หรือชุดของตัวแปรอิสระ เมื่อใช้ตัวแปรอิสระสองตัวขึ้นไปเพื่อทํานายหรืออธิบายผลลัพธ์ของตัวแปรที่ขึ้นต่อกัน เป็นอัลกอริทึมการจําแนกประเภท มันถูกใช้เพื่อทํานายผลลัพธ์แบบไบนารีตามชุดของตัวแปรอิสระ เช่น "ใช่" หรือ "ไม่", "ผ่าน" หรือ "ล้มเหลว" เป็นต้น แต่ก็มี Multinomial logistic regression ที่สามารถใช้ได้เมื่อมีตัวแปรที่ขึ้นกับประเภทหนึ่งตัวแปรที่มีระดับที่ไม่ได้เรียงลําดับตั้งแต่สองระดับขึ้นไป เช่น รถไฟ รถบัส รถราง และจักรยาน เป็นต้น คณะผู้จัดทำได้เลือกใช้ไลบรารี่ Scikit Learn ซึ่งเป็นไลบรารีในภาษา Python สำหรับการพัฒนาโปรแกรมโดยใช้การเรียนรู้ของเครื่อง จุดเด่นคือฟังก์ชันในการแบ่งประเภทข้อมูล การแบ่งกลุ่มข้อมูล การวิเคราะห์การถดถอย ฯลฯ และเครื่องมือที่ใช้ในการเทรนโมเดลด้วยอัลกอริทึมดังกล่าว คือ RandomForestClassifier LinearSVC MultinomialNB และ LogisticRegression ดังภาพที่ 3-15



ภาพที่ 3-15\* เครื่องมือที่ใช้ในการเทรนโมเดลด้วยอัลกอริทึมต่าง ๆ

จากนั้นทำการทดสอบโมเดลโดยมีชุดข้อมูลในการฝึกเดียวกัน เพื่อหาอัลกอริทึมที่มีประสิทธิภาพ และมีความแม่นยำมากที่สุดโดยอัลกอริทึม ดังภาพที่ 3-16



ภาพที่ 3-16\* ทดสอบโมเดล

3.1.4 การจำแนกความรู้สึกของข้อความ (Sentiment Analysis)

คณะผู้จัดทำได้ใช้โมเดลที่ได้รับการฝึกมาเรียบร้อยแล้ว (Pre-Trained Model) ชื่อ WangchanBERTa ในการจำแนกความรู้สึกของข้อความ ที่พัฒนาโดยสถาบันวิจัยปัญญาประดิษฐ์ประเทศไทย (Thailand Artificial Intelligence Research Institute) [[WangchanBERTa: Pre-trained Thai Language Model | airesearch.in.th](https://airesearch.in.th/releases/wangchanberta-pre-trained-thai-language-model/?fbclid=IwAR08tLV_gr7TKU37xTs7dPoe8k8AcUe0GA-fOZN9LOhq_xCZS7SfBA7IQU0)] มาวิเคราะห์ประเภทความรู้สึกของข้อความ ได้แก่ ด้านบวก (Positive) กลาง (Neutral) และด้านลบ (Negative) เพื่อนำไปสรุปแนวโน้มของความคิดเห็น และนำไปแสดงบนกระดานสรุปข้อมูล (Dashboard) ต่อไป ผลลัพธ์ที่ได้จากโมเดล WangchanBERTa สามารถแสดงได้ดังภาพที่ 3-.....

Graphical user interface, text, application, email

Description automatically generated

ภาพที่ 3-.... ผลลัพธ์การทำนายความรู้สึกของข้อความจากโมเดล WangchanBERTa

3.1.5 การทดสอบและประเมินผลโมเดล

ในการทดสอลและประเมินผลโมเดลจำแนกประเภทข้อความที่พัฒนาโดยอัลกอริทึมทั้ง 4 ประเภท คณะผู้จัดทำได้ใช้ค่าความแม่นยำ (Accuracy) เป็นเกณฑ์ในการประเมินประสิทธิภาพของโมเดล โดยค่าความแม่นยำเป็นการเปรียบเทียบร้อยละของประเภทข้อความที่โมเดลทำนายได้ ซึ่งสามารถแจกแจงเป็นตัวเลขในรูปแบบเมทริกซ์ (Confusion Matrix) ดังตารางที่ 3-........

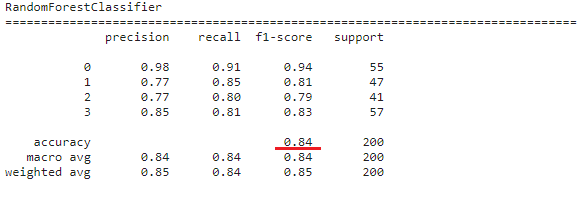
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | | **คำที่ทำนาย (Predicted)** | | | |
| **การเรียนการสอน (0)** | **สิ่งอำนวยความสะดวก (1)** | **ทรัพยากรในการเรียนการสอน (2)** | **สภาพแวดล้อมและความปลอดภัย (3)** |
| **ค่าที่แท้จริง (Actual)** | **การเรียนการสอน (0)** | P00 | P01 | P02 | P03 |
| **สิ่งอำนวยความสะดวก (1)** | P10 | P11 | P12 | P13 |
| **ทรัพยากรในการเรียนการสอน (2)** | P20 | P21 | P22 | P23 |
| **สภาพแวดล้อมและความปลอดภัย (3)** | P30 | P31 | P32 | P33 |

ตารางที่ 3-..........

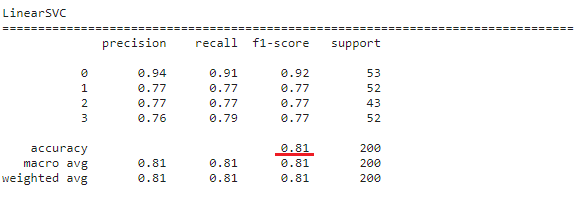
โดยแต่ละช่องใน Confusion Matrix ได้แก่ค่า Pij ซึ่งหมายถึง จำนวนข้อความที่อยู่ในประเภท i แต่โมเดลทำนายป็นประเภท j กรณีที่โมเดลทำนายได้ถูกต้องค่า i จะเท่ากับค่า j ดังนั้นการหาค่าความแม่นยำ (Accuracy) สามารถหาได้จากสูตรงคำนวณดังสมการที่ 3-.......

(3-….)

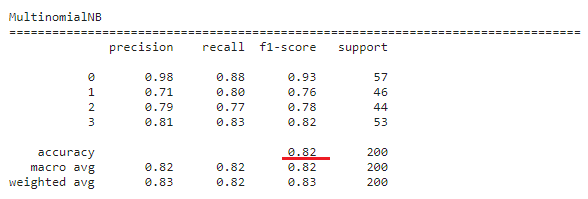
คณะผู้จัดทำโครงการได้ใช้ฟังก์ชั่น Classification Report ในไลบรารี่ของ Scikit Learn เพื่อแสดงผลการทดสอบความแม่นยำของแต่ละอัลกอริทึม ซึ่งแสดงได้ดังภาพที่ 3-.....



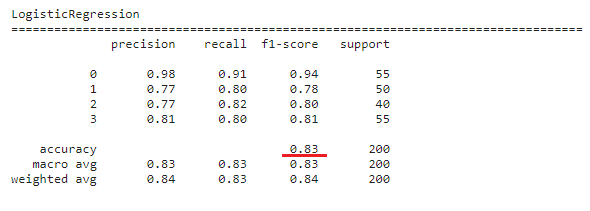
ภาพที่ 3-17\* รายงานผลความแม่นยำในการทำนายข้อความของอัลกอริทึม RandomForestClassifier



ภาพที่ 3-18\* รายงานผลความแม่นยำในการทำนายข้อความของอัลกอริทึม LinearSVC



ภาพที่ 3-19\*รายงานผลความแม่นยำในการทำนายข้อความของอัลกอริทึม MultinomialNB



ภาพที่ 3-20\*รายงานผลความแม่นยำในการทำนายข้อความของอัลกอริทึม LogisticRegression

จากภาพที่ 3-... ถึง 3-.... พบว่าโมเดลที่พัฒนาด้วยอัลกอริทึม RandomForestClassifier มีความแม่นยำร้อยละ 84 รองลงมาได้แก่โมเดล LogisticRegression มีความแม่นยำร้อยละ 83 ตามด้วย MultinomialNB มีความแม่นยำอร้อยละ 82 และ LinearSVC มีความแม่นยำอยู่ที่ร้อยละ 81 คณะผู้จัดทำจึงเลือกโมเดลที่พัฒนาโดยอัลกอริทึม RandomForestClassifier มาใช้ในการสร้างAPI จำแนกข้อความต่อไป สำหรับรายละเอียดความแม่นยำของโมเดลที่ได้เลือกใช้คณะผู้จัดทำจะอภิปรายในบทที่ 4

3.2 การพัฒนาเว็บแอปพลิเคชัน และกระดานสรุปข้อมูล (Dashboard)

การพัฒนาเว็บแอปพลิเคชัน ประกอบด้วยขั้นตอนการการวิเคราะห์ระบบ และการออกแบบส่วนติดต่อผู้ใช้ของเว็บแอปพลิเคชัน NVC Tweets Analytics เพื่อนำไปพัฒนาแอปพลิเคชันต่อไป ซึ่งมีรายละเอียดดังต่อไปนี้

3.2.1 การวิเคราะห์ระบบ

การวิเคราะห์ระบบ NVC Tweets Analytics ทางคณะผู้จัดทำได้ปรึกษาผู้เชี่ยวชาญด้านวิทยาการข้อมูล และร่วมกันวิเคราะห์ความต้องการของผู้ใช้งานสรุปได้ว่าเว็บแอปพลิเคชัน NVC Tweets Analytics ควรประกอบด้วย 3 ฟังก์ชันหลัก ได้แก่ การแสดงกระดานสรุปข้อมูล (View Dashboard) การทดลองการทำนายข้อมูลด้วยโมเดล (Predict Text) และการดึงข้อมูลการทวีตด้วยแฮชแท็กมาเก็บยังฐานข้อมูลเพื่อนำมาวิเคราะห์ข้อมูลและสร้างเป็นกระดานสรุปข้อมูล (Get Tweets) ภาพที่ 3-..... แสดง Use Case Diagram ของระบบ NVC Tweets Analystics

Graphical user interface, application

Description automatically generated

3.2.2 การออกแบบส่วนติดต่อผู้ใช้ของเว็บแอปพลิเคชัน

คณะผู้จัดทำได้ทำการออกแบบส่วนติดต่อผู้ใช้งานของระบบ NVC Tweets Analytics ประกอบด้วยหน้ากระดานสรุปข้อมูล (Dashboard) หน้าทดสอบการใช้งานโมเดล และหน้าการดึงข้อมูลเพื่อนำเข้าฐานข้อมูลตามแฮชแท็ก ซึ่งมีรายละเอียดดังต่อไปนี้

3.2.1.1 หน้ากระดานสรุปข้อมูล

หน้ากระดานสรุปข้อมูลจะแสดงการวิเคราะห์ และการจำแนกประเภทของหัวเรื่องทั้งหมด แสดงเป็นแผนภูมิ และร้อยละของจำนวนข้อความในแต่ละหัวข้อ ที่หมายเลข 1 ของภาพที่ 3-15 รวมถึงการแสดงแผนภูมิการวิเคราะห์ความรู้สึกจากข้อความของแต่ละหัวเรื่อง แสดงร้อยความรู้สึกของจำนวนข้อความในแต่ละหัวเรื่องนั้น ที่หมายเลข 2 ของภาพที่ 3-16

Graphical user interface, application, website

Description automatically generated

3.2.1.2 หน้าทดสอบการใช้งานโมเดล

หน้านี้ผู้ใช้จะสามรถทดลองโมเดล โดยการใส่ข้อความลงในกล่องข้อความเพื่อนำข้อความนั้นไปวิเคราะห์การจำแนกประเภทตามหัวข้อ และวิเคราะห์ความรู้สึกของข้อความ ดังภาพที่....... เมื่อกดปุ่มวิเคราะห์ข้อความหากมีอยู่ในระหว่างกระบวนการวิเคราะห์ข้อความ หน้าเว็บจะแสดงหน้าการโหลด ดังภาพที่........ และเมื่อเสร็จสิ้นจะแสดงผลการวิเคราะห์ ดังภาพที่.........

Graphical user interface, application

Description automatically generated

ภาพที่ .....

Graphical user interface, application

Description automatically generated

ภาพที่ ......

Graphical user interface, text

Description automatically generated with medium confidence

ภาพที่ ....

3.2.1.2 การดึงข้อมูลเพื่อนำเข้าฐานข้อมูลตามแฮชแท็ก

ในหน้านี้ผู้ใช้งานจะใส่ข้อความแฮชแท็กที่ต้องการลงในกล่องข้อความ ดังภาพที่....... เมื่อกดปุ่มดึงข้อมูล โปรแกรมจะแสดงหน้าโหลด ดังภาพที่ ....(ข้างบน) และดึงข้อมูลพร้อมกับนำข้อมูลไปวิเคราะห้วยโมเดลที่พัฒนาขึ้น และบันทึกผลการวิเคราะห์ลงฐานข้อมูล เมื่อเสร็จสิ้นจะแสดงผลการแจ้งเตือน ดังภาพที่......

Graphical user interface, text, application

Description automatically generated

ภาพที่ .....

Text

Description automatically generated with medium confidence

ภาพที่....

3.2.3 การพัฒนาเว็บแอปพลิเคชัน

ในการพัฒนาเว็บแอปพลิเคชัน NVC Tweets Analytics ทางคณะผู้จัดทำได้ใช้เฟรมเวิร์กฟลาสก์ (Flask) ซึ่งเป็นเฟรมเวิร์กในการพัฒนาเว็บแอปพลิเคชันด้วยภาษา Python เนื่องจากสะดวกในการเรียกใช้โมเดลวิเคราะห์ข้อความ การดึงข้อความจากทวีตลงฐานข้อมูล MongoDB โดยใช้ไลบรารี PyMongo นอกจากนี้ยังใช้ Chart.js ในการสร้างแผนภูมิเพื่อแสดงในหน้ากระดานสรุปข้อมูล (Dashboard) ตลอดจนการสร้างหน้าเว็บไซต์ด้วย Boostrap 5 เพื่อรองรับการแสดงผลแบบ Responsive

3.3 สรุป

บทนี้ได้นำเสนอเกี่ยวกับขั้นตอนการดำเนินงานของโครงการ ได้แก่ การรวบรวมและเตรียมข้อมูลเพื่อใช้สำหรับการพัฒนาโมเดลวิเคราะห์ข้อความ และการเตรียมข้อมูลที่ดึงจากทวิตเตอร์เพื่อนำไปวิเคราะห์และสร้างกระดานสรุปข้อมูล (Dashboard) การวิเคราะห์ความต้องการข ออกแบบ และพัฒนาเว็บแอปพลิเคชัน NVC Tweets Analytics โดยผลการประเมินโมเดลที่พัฒนา และผลการสร้างเว็บแอปพลิเคชันดังกล่าว จะอภิปรายในบทถัดไป