

การจัดหมวดหมู่ข้อความจากแบบประเมินผลแบบอัตโนมัติ
ด้วยวิธีการเรียนรู้ของเครื่อง

สงวน แก้วขาว

สารนิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชาวิศวกรรมข้อมูลขนาดใหญ่
วิทยาลัยนวัตกรรมด้านเทคโนโลยีและวิศวกรรมศาสตร์
มหาวิทยาลัยธุรกิจบัณฑิตย์

พ.ศ.2564

**MACHINE LEARNING FOR AUTOMATED TEXT CATEGORIZATION
FROM EVALUATION**

SA-NGUAN KEAWKAO

**A Thematic Paper Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Engineering
Department of Big Data Engineering,
College of Innovative Technology and Engineering
Dhurakij Pundit University**

2021

หัวข้อสารนิพนธ์	การจัดหมวดหมู่ข้อความจากแบบประเมินผลแบบอัตโนมัติ ด้วยวิธีการเรียนรู้ของเครื่อง
ชื่อผู้เขียน	สงวน แก้วขาว
อาจารย์ที่ปรึกษา	ดร.เอกสิทธิ์ พัทธวงษ์ศักดิ์ดา
สาขาวิชา	วิศวกรรมข้อมูลขนาดใหญ่
ปีการศึกษา	2563

บทคัดย่อ

การรับรู้ผลของระดับความพึงพอใจ หรือการนำผลที่ได้จากการประเมินไปสู่การเรียนรู้ เพื่อปรับปรุงประสิทธิภาพหรือแก้ไขปัญหาในด้านต่างๆ ของกระบวนการทำงานในแต่ละงานเป็นเรื่องที่หน่วยงาน หรือองค์กรต้องการ การสรุป รวบรวมยอด จากชุดข้อมูลจากการประเมิน ในปัจจุบันรูปแบบของการประเมินผลความพึงพอใจมีหลากหลายรูปแบบ รวมถึงวิธีการตอบคำถาม แต่มีชุดคำถามอีกประเภทที่น่าสนใจนั่นคือ การเปิดให้กรอกข้อความที่เป็น Text Area ซึ่งข้อมูลที่ได้จะเป็นเชิงรูปแบบของประโยค ที่สามารถบอกถึงความต้องการของผู้กรอกโดยอิสระ

ปัญหาของการลำดับความสำคัญ เป็น Pain Point ที่อีกอย่างหนึ่งที่ส่งผลให้เกิดการแก้ปัญหาไม่ตรงตามปัญหาที่เกิดขึ้น ส่งผลต่อความเชื่อมั่น และคุณภาพของงาน และเรื่องเวลา อีกประเด็นคือ ปัญหาจากพนักงาน ผู้ประสานงาน ผู้รับผิดชอบโครงการ ที่สถาบันรับรองคุณภาพสถานพยาบาล มีกิจกรรมหลักคือ การจัดกิจกรรม การประชุมหลักสูตร สัมมนา รวมถึงการจัดประชุมวิชาการที่มีผู้เข้าร่วมงานจำนวนมาก พบว่าขั้นตอนของการประเมินผลกิจกรรม ใช้เวลาจัดกลุ่มของชุดคำถามและข้อเสนอแนะยาก การใช้คนในการจัดการ และพบว่าเทคโนโลยีมีโอกาสช่วย และมีโอกาสพัฒนา ให้เกิดความเชื่อมั่นต่อความถูกต้องของข้อมูล การสรุปผล โดยเฉพาะข้อมูลที่มาก และใช้เวลาสั้นลง ทำให้เพิ่มคุณค่าของงานได้

เครื่องมือที่จัดทำขึ้น เป็นการจัดหมวดหมู่ข้อความจากแบบประเมินผลแบบอัตโนมัติด้วยวิธีการเรียนรู้ของเครื่อง จึงนำเอาข้อความจากการประเมินผลมาอ่าน และสร้างสร้างคำตอบ Labeling โดยนักวิชาการจำนวน 15 ท่าน โดยใช้การอ่านและรวมแสดงความคิดเห็น เพื่อตกผลึกเพื่อสร้างเป็นข้อมูลต้นแบบจำนวน 17,000 ชุดข้อความ หลังจากนั้นใช้วิธีการสร้างชุดคำสั่งในการจัดการข้อมูล Data Exploration และ Classification Data เพื่อแปลงค่า Class ของ Labeling ที่เป็นตัวเลข และจัดการกับ

ข้อความที่ไม่สามารถใช้งานได้ ("m\d|[|\^|\\$|\.\|?|*|+|\(|\)|\{|}\|'|") จากนั้นใช้ Deep cut Tokenize เพื่อตัดคำเนื่องจากเครื่องมือที่เป็น Model ที่จะใช้คือ LSTM ต้องการข้อมูลที่จะนำเข้า Model ได้ต้องเป็นตัวเลขแบบ Array จึงต้องใช้ wiki.th ในการแปลงข้อมูล เพื่อให้ได้ค่า X, Y จากนั้นจึงสร้าง Model เพื่อรับค่าเข้าเพื่อทำการ (Train Test) โดยให้ Lookup=50 และ epochs=25 พบว่ามีความแม่นยำร้อยละ

68.91



Thematic Paper Title	MACHINE LEARNING FOR AUTOMATED TEXT CATEGORIZATION FROM EVALUATION
Author	Sa-nguan Keawkao
Thematic Paper Advisor	Dr. Eakasit Pacharawongsahda
Department	Big Data Engineering
Academic Year	2020

ABSTRACT

Applying the perception of satisfaction level result or the assessment results to improve work efficiency, solve problems in different aspects of work system are need for organization development.

At present, there are many forms of satisfaction evaluation including; answer the set of question. However, “text comment” is a interested evaluation form. The result of data will be show in term of sentence which specify personal need by the set of question.

Prioritization is the pain point of this topic. Since the importance and urgent issue are not be in line with the organization problem. This will be affected the trust, quality of work and time spent.

The main activities of HAI are organizing activities, academic courses, seminars including the academic event with many participants. The problem that found from project coordinator are that the process of activity evaluation . It takes difficult by using human management for categorizing the text and suggestions from participant. Thus, technology is the opportunity for developing the trust and data accurate. To summarize the results of big set of data. Timeless can be increasing the value of work.

This tool will be automatic categorize the text of evaluation form. The process are read the text from evaluation form, labeling answer from 15 person of academic and create Labeling responses by 15 person of academic and combining comments to crystallize and create a prototype of 17,000 texts.

Then use the data exploration and classification data management to convert class numerical Labeling and text that cannot be used ("n|d| | [| | ^ | \$ | \. | \? | * | \+ | \ (| \) | \ { | \} | ':") Then use Deep cut Tokenize to wrap words because the Model tool that will be used is LSTM requires data to be imported. Model must be Array numbers, so wiki.th is required to convert the data. To get X, Y, then create a model to receive the entrance fee to do (Train Test), where Lookup = 50 and epochs = 25 found to be 68.91% accurate.



กิตติกรรมประกาศ

สารนิพนธ์นี้เป็นการเรียนรู้จากการบูรณาการข้อมูลจากแหล่งข้อมูลจริงกับสถานที่ทำงานที่ปฏิบัติหน้าที่ของผู้ทำงานวิจัย รวมถึงเจ้าหน้าที่นักวิชาการจนสำเร็จได้ ด้วยความอนุเคราะห์จากบุคลากรหน่วยงานสถาบันรับรองคุณภาพสถานพยาบาล (องค์การมหาชน) ที่ได้ให้คำแนะนำและอุทิศตนเพื่อวิเคราะห์ และแลกเปลี่ยนเรียนรู้ในการจัดทำ Labeling เพื่อเป็นข้อมูลในการ training Data และขอขอบพระคุณคณะอาจารย์ประจำหลักสูตรวิศวกรรมศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมข้อมูลขนาดใหญ่วิทยาลัยนวัตกรรมด้านเทคโนโลยีและวิศวกรรมศาสตร์ มหาวิทยาลัยธุรกิจบัณฑิตย์ ดร.ชนภัทร ฆังคะจิตร (ผู้อำนวยการหลักสูตร) ผู้ช่วยศาสตราจารย์ ดร.ดวงใจ จิตคงชื่น (อาจารย์ประจำหลักสูตร) ดร.รัฐศิลป์ รานอกภานุวัชร (อาจารย์ผู้สอน) ที่สำคัญท่านอาจารย์ที่ปรึกษา ดร.เอกสิทธิ์ พัทธวงศ์ศักดิ์ ที่ให้ความอนุเคราะห์ในการให้ความรู้ ประสบการณ์ และองค์ความรู้ที่สำคัญสู่การปฏิบัติ รวมถึงเพื่อนนักศึกษาที่ให้คำปรึกษา ผู้วิจัยใคร่ขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ โอกาสนี้

สงวน แก้วขาว

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	๗
บทคัดย่อภาษาอังกฤษ.....	๖
กิตติกรรมประกาศ.....	๗
สารบัญตาราง.....	๗
สารบัญภาพ.....	๘
บทที่	
1. บทนำ.....	1
1.1 ที่มาและความสำคัญของปัญหา.....	1
1.2 วัตถุประสงค์.....	3
1.3 ขอบเขตของการวิจัย.....	3
2. ทฤษฎี และงานวิจัยที่เกี่ยวข้อง.....	4
2.1 องค์ประกอบ รูปแบบ ลักษณะของแบบประเมินผล.....	4
2.2 การวิเคราะห์ข้อมูลเชิงคุณภาพ Qualitative Data Analysis Techniques.....	5
2.3 การวิเคราะห์ข้อมูล Data Mining Techniques.....	9
2.4 Machine Learning and LSTM Model.....	13
2.5 องค์ประกอบของ API.....	14
2.6 Web Application.....	15
2.7 Data Text Mining.....	16
3. ระเบียบวิธีวิจัย.....	19
3.1 แนวทางการวิจัยและพัฒนา.....	19
4. ผลการศึกษา.....	41
4.1 Business Intelligence (BI).....	25
4.2 การแสดงผลของการวิเคราะห์ข้อมูล.....	26
4.3 การแสดงข้อมูลแบบจำแนกปัญหา.....	28
4.4 จากการนำเสนอข้อมูลแบบรายปี.....	29
4.5 ผลประเมินจากผู้ร่วมทดสอบการใช้งาน.....	29

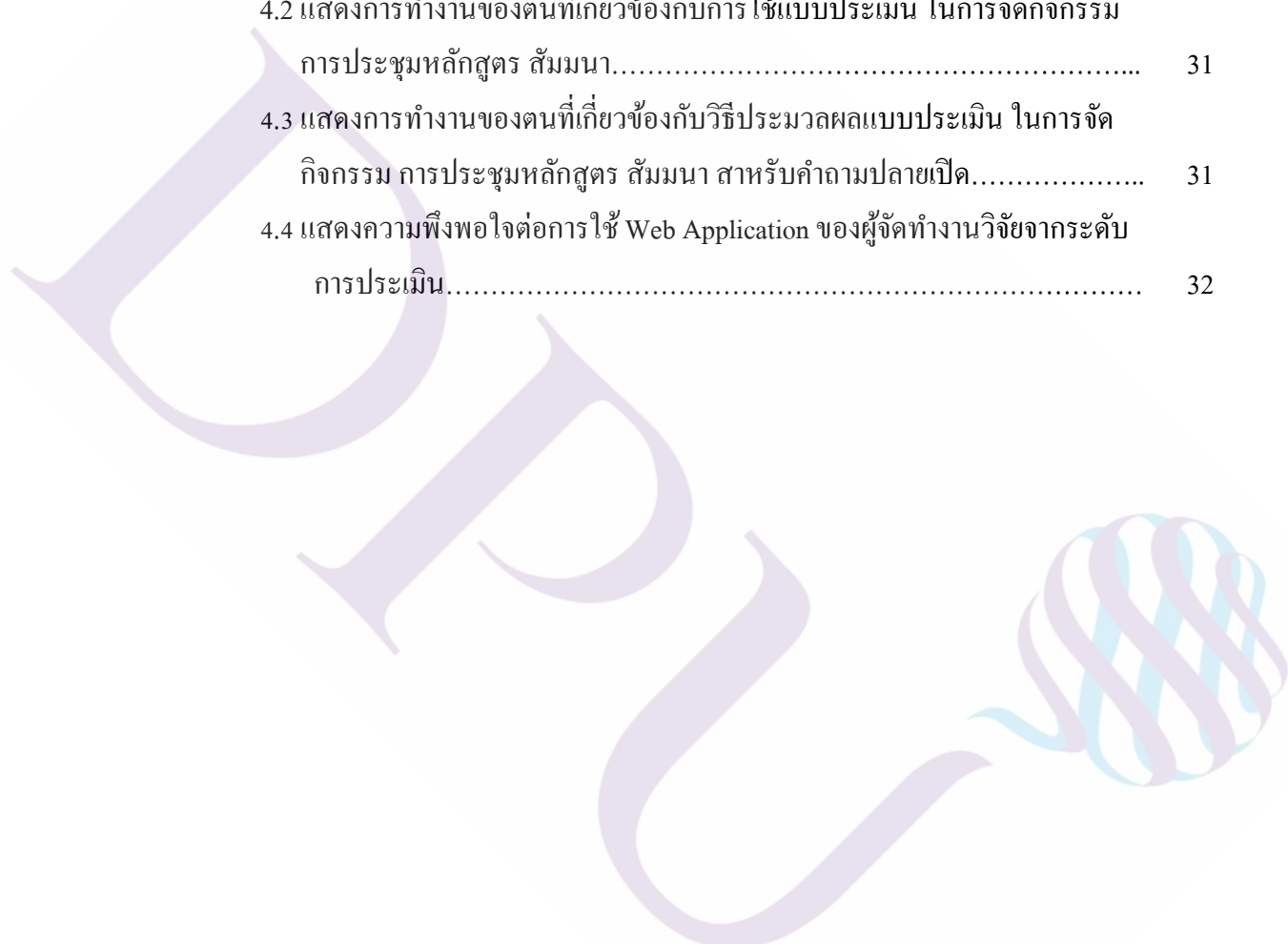
สารบัญ (ต่อ)

บทที่	หน้า
5. บทสรุปและข้อเสนอแนะ.....	34
5.1 สรุปผลการศึกษา.....	34
5.2 ข้อเสนอแนะ.....	34
บรรณานุกรม.....	36
ภาคผนวก.....	38
ประวัติผู้เขียน.....	45



สารบัญตาราง

ตารางที่	หน้า
2.1 การวิเคราะห์คำหลักเพื่อจัดกลุ่มคำ.....	6
2.2 แสดงการจำแนกสารระบบ.....	7
4.1 แสดงข้อมูลการรับรู้เทคโนโลยีเกี่ยวกับ Bigdata.....	30
4.2 แสดงการทำงานของคนที่เกี่ยวข้องกับการใช้แบบประเมิน ในการจัดกิจกรรม การประชุมหลักสูตร สัมมนา.....	31
4.3 แสดงการทำงานของคนที่เกี่ยวข้องกับวิธีประมวลผลแบบประเมิน ในการจัด กิจกรรม การประชุมหลักสูตร สัมมนา สำหรับคำถามปลายเปิด.....	31
4.4 แสดงความพึงพอใจต่อการใช้ Web Application ของผู้จัดทำงานวิจัยจากระดับ การประเมิน.....	32



สารบัญภาพ

ภาพที่	หน้า
2.1 แสดง Association rule.....	10
2.2 แสดง algorithm Hierarchical Clustering ตัวอย่างการ clustering จาก dataset iris..	11
2.3 แสดง algorithm K-Means ตัวอย่างการ clustering จาก dataset iris.....	12
2.4 แสดง Recurrent Neural Networks (RNNs).....	13
2.5 แสดง Application Program Interface (API).....	15
2.6 แสดงการจัดการสารสนเทศด้วยกระบวนการ Text Mining.....	16
2.7 แสดงการเครื่องมือจัดการ Text Mining.....	17
3.1 แสดงกระบวนการ CRISP-DM.....	18
3.2 แสดงชุดคำสั่งของการจัดทำ Data Preparation.....	19
3.3 แสดงกระบวนการทำงานของโปรแกรม.....	20
3.4 แสดงชุดคำสั่งของการจัดทำ Model.....	21
3.5 แสดง Accuracy Model = 68.91%.....	21
3.6 แสดงการนำข้อมูลที่เป็น Text เข้าสู่ Model และทำการ Predict.....	22
3.7 แสดงการออกแบบ web Application.....	23
3.8 แสดงการเลือกใช้ Server.....	24
4.1 แสดงการนำข้อความเข้าสู่การ Predict และ Success Predicted.....	25
4.2 แสดงการนำข้อมูลการจัดเรียง Priority จากมากไปน้อย.....	26
4.3 แสดงประโยชน์ข้อความที่น่าเข้า Prediction.....	27
4.4 แสดงการจำแนกผลของการ Predict.....	28
4.5 แสดงกราฟเปรียบเทียบปัญหาเป็นกราฟเส้น เพื่อเปรียบเทียบรายปี.....	29

บทที่ 1

บทนำ

1.1 ที่มาและความสำคัญของปัญหา

ในสถานการณ์ของการสังคมที่มีการแข่งขันเรื่องของการตอบสนอง ต่อผู้รับบริการหรือกลุ่มลูกค้ามีความจำเป็น และสำคัญอย่างยิ่ง ในสร้างการรับรู้ผลของระดับความพึงพอใจ หรือเสียงสะท้อน เพื่อนำผลที่ได้จากการประเมินไปสู่การเรียนรู้ เพื่อปรับปรุงประสิทธิภาพ ตามวงล้อของการพัฒนาคุณภาพ 3c, PDCA ในการวางแผนแก้ไขปัญหาในด้านต่างๆ ของกระบวนการทำงานในแต่ละส่วนงาน ซึ่งเป็นเรื่องที่ หน่วยงาน หรือองค์กรต้องการ ข้อมูลการสรุป รวบรวม และวิเคราะห์ จากชุดข้อมูล แหล่งต่างๆ หรือจากการประเมิน ในปัจจุบันรูปแบบของการประเมินผลความพึงพอใจมีหลากหลายรูปแบบ รวมถึงวิธีการตอบคำถาม แต่มีชุดคำถามอีกประเภทที่น่าสนใจนั่นคือ การเปิดให้กรอกข้อความที่เป็น Text Area ซึ่งข้อมูลที่ได้จะเป็นเชิงรูปแบบของประโยค ที่สามารถบอกถึงความต้องการของผู้กรอกโดยอิสระ

ปัญหาของการลำดับความสำคัญ เป็น Pain Point ที่อีกอย่างหนึ่งส่งผลให้เกิดการแก้ปัญหาไม่ตรงตามปัญหาที่เกิดขึ้น ส่งผลต่อความเชื่อมั่น และคุณภาพของงาน และเรื่องเวลา อีกประเด็นคือปัญหาจากพนักงาน ผู้ประสานงาน ผู้รับผิดชอบ โครงการ ที่สถาบันรับรองคุณภาพสถานพยาบาล มีกิจกรรมหลักคือ การจัดกิจกรรม การประชุมหลักสูตร สัมมนา รวมถึงการจัดประชุมวิชาการที่มีผู้เข้าร่วมงานจำนวนมาก พบว่าขั้นตอนของการประเมินผลกิจกรรม ใช้เวลาจัดกลุ่มของชุดคำถามและข้อเสนอแนะยาก การใช้คนในการจัดการ และพบว่าเทคโนโลยีมีโอกาช่วย และมีโอกาสพัฒนา ให้เกิดความเชื่อมั่นต่อความถูกต้องของข้อมูล การสรุปผล โดยเฉพาะข้อมูลที่มาก และการใช้เวลาสั้นลงทำให้เพิ่มคุณค่าของงานได้

การวิเคราะห์และจัดกลุ่มข้อมูลจากแบบประเมินที่มีลักษณะคำถามปลายเปิดเพื่อสรุปโอกาสพัฒนาที่สำคัญ และจัดลำดับความสำคัญ นำมาสู่การปรับปรุงกิจกรรมให้มีคุณภาพมากขึ้น ตอบสนองต่อความต้องการของลูกค้าที่มีความหลากหลาย

จากข้อความที่เป็นรูปแบบของชุดประโยค ในอดีตหน่วยงานจะสรุปโดยนักวิชาการหรือเจ้าของโครงการ สิ่งที่ได้จะเป็นการสรุปเป็นข้อๆ หรือจัดรวมให้เป็นเป็นกลุ่ม ซึ่งใช้เวลาในการอ่านข้อความทั้งหมดแล้วจึงสรุปโดยประสบการณ์ของผู้อ่านเอง ในปัจจุบันมีเทคโนโลยีการต่างๆที่สามารถยกระดับการอ่าน การตัด การวิเคราะห์ข้อความ รวมถึงสามารถแปลงค่าต่างๆได้ Artificial Intelligence (AI) เป็นการประยุกต์ใช้เทคโนโลยีที่จัดการกับ Pour Point ของมนุษย์ที่สามารถก้าวข้ามเรื่องที่มีมนุษย์ทำซ้ำๆ ในเรื่องเดิมๆได้

ประเภทของคำถามปลายเปิด (Open-ended question) ผู้วิจัยจะถามคำถามในลักษณะที่ให้โอกาสผู้ตอบได้ แสดงความคิดเห็นของตนเองอย่างอิสระ ผู้ตอบจะใช้คำพูดของตนเองซึ่งจะช่วยให้ผู้วิจัยได้ทราบถึงแนวคิด และความรู้สึกของผู้ตอบคำถามอย่างกว้างๆ ส่วนข้อเสีย ได้แก่ ผู้วิจัยมีความลำบากในการตีความหมายของข้อมูล การประมวลผลข้อมูล และการสรุปผลข้อมูล เพราะผู้ตอบจะตอบคำตอบหลากหลาย ซึ่งมีความยุ่งยากในการจัดหมวดหมู่ วิธีการวิจัยเบื้องต้น

เครื่องมือที่จัดทำขึ้น เป็นการเรียนรู้ในการจัดหมวดหมู่ข้อความอัตโนมัติจากแบบประเมินผล จึงนำเอาข้อความจากการประเมินผลมาอ่าน และสร้างสร้างคำตอบ Labeling โดยนักวิชาการจำนวน 15 ท่าน โดยใช้การอ่านและรวมแสดงความคิดเห็น เพื่อตกผลึกเพื่อสร้างเป็นข้อมูลต้นแบบจำนวน 17,000 ชุดข้อความ หลังจากนั้นใช้วิธีการสร้างชุดคำสั่งในการจัดการข้อมูล Data Exploration และ Classification Data เพื่อแปลงค่า Class ของ Labeling ที่เป็นตัวเลข และจัดการกับข้อความที่ไม่สามารถใช้งานได้ ("`\n|d|[\|^\|$\|.\|?|*|+|(|)\|{|}': "`) จากนั้นใช้ Deep cut Tokenize เพื่อตัดคำเนื่องจากเครื่องมือที่เลือกศึกษาเป็น Model คือ LSTM ต้องการข้อมูลที่จะนำเข้า Model ได้ ต้องเป็นตัวเลขแบบ Array Vector จึงต้องใช้ wiki.th ในการแปลงข้อมูล เพื่อให้ได้ค่า X,Y จากนั้นจึงสร้าง Model เพื่อรับค่าเข้าเพื่อทำการ (Train Test) โดยให้ Lookup=50 และ epochs=25 พบว่ามีความแม่นยำร้อยละ 68.91

คำสำคัญ : รูปแบบของประโยค, รูปแบบของชุดประโยค, Pour Point, จัดหมวดหมู่ตามชุดคำตอบ

1.2 วัตถุประสงค์

เพื่อสร้างระบบที่สามารถช่วยบริหาร จัดการเกี่ยวกับการจัดหมวดหมู่ของปัญหา จัดทำ Machine learning ในการเรียนรู้การจัดกลุ่มคำโดยใช้เทคโนโลยีที่เกี่ยวกับ AI เข้าช่วยงาน สร้าง Visualization ในการติดตาม และวิเคราะห์ปัญหาที่สำคัญและเร่งด่วน นำไปสู่การเรียนรู้กระบวนการทำ Machine learning จากงานวิจัยที่ศึกษา บูรณาการเข้าสู่งานประจำ

1.3 ขอบเขตของการวิจัย

1. จัดทำเตรียมข้อมูล และเรียนรู้การทำงานของ Model เพื่อสร้างการเรียนรู้เพื่อให้สามารถ Prediction Data
2. พัฒนา Model เพื่อใช้การเรียนรู้ในการจัดหมวดหมู่ข้อความอัตโนมัติจากแบบประเมินผล
3. จัดทำ Application on Web
4. จัดทำระบบ Visualization เพื่อประกอบการตัดสินใจ

บทที่ 2

ทฤษฎี และผลงานที่เกี่ยวข้อง

ในการศึกษาเอกสารงานวิจัยที่เกี่ยวข้องกับการจัดทำ การจัดหมวดหมู่ข้อความจากแบบประเมินผลแบบอัตโนมัติด้วยวิธีการเรียนรู้ของเครื่อง Machine learning for automated text categorization from evaluation โดยมีหัวข้อดังนี้

- 2.1 องค์ประกอบ รูปแบบ ลักษณะของแบบประเมินผล
- 2.2 การวิเคราะห์ข้อมูลเชิงคุณภาพ Qualitative Data Analysis Techniques
- 2.3 การวิเคราะห์ข้อมูล Data Mining Techniques
- 2.4 Machine Learning and LSTM Model
- 2.5 องค์ประกอบของ API
- 2.6 Web Application
- 2.7 Data Text Mining

2.1 องค์ประกอบ รูปแบบ ลักษณะของแบบประเมินผล

ศิริชัย กาญจนวาสี (2538) กล่าวว่า การประเมินหมายถึง การตัดสินคุณค่าของสิ่งที่มีมุ่งประเมิน นิยามนี้เป็นที่ยอมรับกันอย่างกว้างขวางในปัจจุบัน เป็นผลมาจากแนวคิดการบุกเบิกของ Scriven(1967, 1974, 1976 : อ้างอิงใน ศิริชัย กาญจนวาสี,2538) ซึ่งว่าเป้าหมายของการประเมินอยู่ที่การตัดสินคุณค่า โดยผู้ประเมินจะต้องมีความเชี่ยวชาญในหลักการสังเกตและเหตุผล ถ้าผู้ประเมินยังไม่ได้ตัดสินคุณค่าของสิ่งที่ประเมิน ถือว่าผู้ประเมินยังทำหน้าที่ไม่สมบูรณ์ แนวคิดดังกล่าวเป็นที่ยอมรับโดยนักทฤษฎีการประเมินในเวลาต่อมา

การประเมินตามนัยดังกล่าวได้ถูกนำไปใช้ปฏิบัติโดยมีรูปแบบแตกต่างกัน เช่น การตัดสินคุณค่าโดยผู้เชี่ยวชาญ (Professional Judgment) ซึ่งเสนอโดย Scriven(1973: อ้างอิงใน ศิริชัย กาญจน

วาลี, 2538) การตัดสินคุณค่าตามมาตรฐานวิชาชีพ (Professional Review) เช่นการรับรองวิทยฐานะ การพิจารณาผลงานทางวิชาการ การสัมภาษณ์คัดเลือกบุคลากร

การเก็บรวบรวมข้อมูลและการเตรียมข้อมูล จะมีลักษณะของการเก็บที่มุ่งเน้นเพื่อให้ได้ข้อมูลที่มีความสำคัญและจำเป็น รูปแบบของคำถาม คือลักษณะของคำถามที่มีในแบบสอบถามที่ผู้วิจัยได้สร้างขึ้น จำแนกออกเป็นประเภทดังนี้ (1) คำถามปลายเปิด (Open-ended Question) ผู้วิจัยจะถามคำถามในลักษณะที่ให้โอกาสผู้ตอบได้ แสดงความคิดเห็นของตนเองอย่างอิสระ ผู้ตอบจะใช้คำพูดของตนเองซึ่งจะช่วยให้ผู้วิจัยได้ทราบถึงแนวคิด และความรู้สึกของผู้ตอบคำถามอย่างกว้างๆ ส่วนข้อเสีย ได้แก่ ผู้วิจัยมีความลำบากในการตีความหมายของข้อมูล การประมวลผลข้อมูล และการสรุปผลข้อมูล เพราะผู้ตอบจะตอบคำตอบหลากหลาย ซึ่งมีความยุ่งยากในการจัดหมวดหมู่ วิธีการวิจัยเบื้องต้น (2) คำถามแบบปลายปิด (Closed-ended Question) แบบสอบถามรูปแบบนี้ จะมีลักษณะคล้ายกับข้อสอบแบบให้เลือกตอบ คือ ในแบบสอบถาม จะมีคำถามและคำตอบที่กำหนดไว้ก่อนล่วงหน้าแล้วให้ผู้ตอบแบบสอบถามเลือกที่จะตอบข้อใดข้อหนึ่งที่ตรงกับความคิดของผู้ตอบมากที่สุด ในคำตอบนั้น จำนวนตัวเลือกอาจมีหลายลักษณะ Multiple Response คณะศิลปศาสตร์และวิทยาศาสตร์ มหาวิทยาลัยธุรกิจบัณฑิตย์ (2549,127-128)

2.2 การวิเคราะห์ข้อมูลเชิงคุณภาพ Qualitative Data Analysis Techniques

จากการศึกษาเรื่องการวิเคราะห์ข้อมูลเชิงคุณภาพ Qualitative Data Analysis Techniques เพื่อนำไปสู่ของการนิยามหรือ Labeling ผู้วิจัยพบว่าข้อดีของการนิยามศัพท์ หรือประโยคจะทำให้สามารถบรรลุเป้าหมายของการประเมินในแง่ของการตัดสินประเด็นสำคัญ ที่อาศัยการวิเคราะห์ การมองให้เห็นคุณค่าของสิ่งที่อ่านหรือการวิเคราะห์ จึงจะมีความสมบูรณ์ในแง่ของการวัด และการตัดสิน รวมถึงวิธีการที่จะรับรู้ถึงเป้าหมายของเนื้อหา ร่วมกับประสบการณ์ของผู้เชี่ยวชาญ จะทำให้ผลของการประเมินและนิยาม มีน้ำหนักน่าเชื่อถือ และเป็นประโยชน์ แต่ถ้าผู้ประเมินหรือผู้ที่จัดทำ Labeling ขาดคุณสมบัติ อาจก่อให้เกิดความไม่เที่ยงตรงของข้อมูลได้ ดังนั้นการตัดสินจนไม่ใช่ทางออกเสมอไป ผู้วิจัยจึงใช้วิธีการ Brainstorming หรือการระดมความคิดเพื่อให้ได้ผลลัพธ์เกี่ยวกับเรื่องใด เรื่องหนึ่ง ในที่นี้หมายถึงการนิยามประโยค Labeling ของงานวิจัยเรื่องนี้ โดยผ่านกระบวนการ KM (Knowledge Management) ซึ่งได้ศึกษาจากกระบวนการและวิธีทางวิชาการดังนี้

การวิเคราะห์ข้อมูลเชิงคุณภาพ Qualitative Data Analysis Techniques เป็นวิธีการสร้างข้อสรุปจากข้อมูลจำนวนหนึ่งซึ่งมักไม่ใช่สถิติในการวิเคราะห์ ทั้งนี้การวิเคราะห์ข้อมูลเชิงคุณภาพ อาจ

ใช้กับการวิจัยเชิงปริมาณที่ผู้วิจัยมีการเก็บรวบรวมข้อมูลเชิงคุณภาพ เช่น แบบสอบถามปลายเปิด การสัมภาษณ์ การสังเกต จดบันทึก (เอี่ยมพร หลินเจริญ เทคนิคการวิเคราะห์ข้อมูลเชิงคุณภาพคณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารคาม, 18-29) เทคนิควิเคราะห์ข้อมูลเชิงคุณภาพสามารถแบ่งออกได้ดังนี้

1. การวิเคราะห์โดยการจำแนกชนิดข้อมูล (Typological Analysis)

การวิเคราะห์โดยการจำแนกชนิดข้อมูล คือ การจำแนกข้อมูลเป็นชนิด (Typologies) คำว่า “Typologies” หมายถึง ขั้นตอนของเหตุการณ์ที่เกิดขึ้นอย่างต่อเนื่อง ซึ่งในกรณีที่นักวิจัยทำการเก็บรวบรวม ข้อมูลแล้ว นักวิจัยจำเป็นต้องจัดระบบข้อมูลโดยอาศัยหลักเกณฑ์ ที่ผู้วิจัยกำหนดขึ้น ซึ่งการจำแนกหรือการ จัดกลุ่มข้อมูลนี้ แบ่งได้เป็น 2 ประเภท ได้แก่

การจำแนกข้อมูลในระดับจุลภาคการจำแนก ข้อมูลระดับนี้แบ่งเป็น 2 ประเภทย่อย ได้แก่ การวิเคราะห์คำ หลัก (Domain Analysis) กับการวิเคราะห์สารระบบ (Taxonomy Analysis) การวิเคราะห์คำ หลัก หมายถึงการจำแนกจัดกลุ่มคำชุดหนึ่งให้อยู่ภายใต้คำอีกชุดหนึ่ง ซึ่งคำดังกล่าวนี้มีความหมายครอบคลุมคำชุดนั้น ๆ หรือ อาจกล่าวได้ว่า การวิเคราะห์คำ หลัก เป็นการจัดกลุ่มคำชุดหนึ่งให้อยู่ร่วมกันโดยอาศัยลักษณะความสัมพันธ์บางอย่างของคำแต่ละคำ ที่นำ มาใช้จัดกลุ่ม ทั้งนี้ลักษณะความสัมพันธ์บางอย่างที่ว่าเป็นความสัมพันธ์เชิงวัฒนธรรมที่บุคคลแต่ละสังคมเป็นผู้จัดจำแนกเช่น คำว่า “คุณลักษณะครูที่ดี” นั้นมิได้หมายถึงครูที่สอนเท่านั้น แต่คำ นี้หมายรวมถึง ด้านจิตวิทยาการเรียนการสอน สามารถทำงานร่วมกับผู้อื่นได้ เป็นต้น ซึ่งถ้าเปรียบเทียบกับการวิเคราะห์ข้อมูลเชิงปริมาณก็คือการวิเคราะห์องค์ประกอบ (Factor Analysis) เพื่อจัดกลุ่มตัวแปรย่อยให้อยู่ภายใต้องค์ประกอบ

ตารางที่ 2.1 การวิเคราะห์คำ หลักเพื่อจัดกลุ่มคำ

กลุ่มคำ	ความสัมพันธ์	คำหลัก
<ul style="list-style-type: none"> - มีความสามารถด้านการสอน - มีคุณธรรมจริยธรรมและบุคลิกที่เป็นแบบอย่างแก่นักเรียน - มีความเข้าใจด้านจิตวิทยาการเรียนการสอน - ร่วมมือกับผู้อื่นอย่างสร้างสรรค์ - ปฏิบัติหน้าที่อย่างมีเกียรติมีศักดิ์ศรี 	<p>เป็นส่วนหนึ่งของ</p>	<p>คุณลักษณะของครูที่ดี</p>

สำหรับการวิเคราะห์สารระบบ นั้นมีความหมายคล้ายคลึงกับการวิเคราะห์คำหลัก เพียงแต่มีความแตกต่างกันที่ว่าการวิเคราะห์จำแนกสารระบบจะมุ่งเน้นแสดงให้เห็นถึงความสัมพันธ์ ระหว่างกลุ่มคำย่อย ๆ ด้วยกันเองและคำหลักในภาพรวมทั้งหมด ลักษณะการจำแนกจัดกลุ่มคำหรือกลุ่มข้อมูล จะมีความซับซ้อนและมีระดับความสัมพันธ์ระหว่างคำต่าง ๆ สูงกว่าการวิเคราะห์คำหลัก ดังตัวอย่างในตารางที่ 2.2

ตารางที่ 2.2 แสดงการจำแนกสารระบบ

กลุ่มคำ	คำหลัก	คำหลัก
คุณลักษณะของครูที่ดี	<ul style="list-style-type: none"> - มีความสามารถด้านการสอน - มีคุณธรรมจริยธรรมและบุคลิกที่เป็นแบบอย่างแก่นักเรียน - มีความเข้าใจด้านจิตวิทยาการเรียนการสอน - ร่วมมือกับผู้อื่นอย่างสร้างสรรค์ - ปฏิบัติหน้าที่อย่างมีเกียรติมีศักดิ์ศรี 	<ul style="list-style-type: none"> - วางตัวเหมาะสมกับกาลเทศะ - มีสติและอารมณ์ที่มั่นคง - ไม่เห็นแก่ตัวและเอาเปรียบผู้อื่น - มีส่วนร่วมในกิจกรรมของชุมชนอยู่เสมอ - ช่วยงานธุรการหรืองานอื่นที่นอกเหนือจากภาระงานสอน - ปฏิบัติตามคำสั่งของผู้บังคับบัญชา

2. การจำแนกข้อมูลในระดับมหภาค การจำแนกข้อมูลในระดับมหภาค เป็นการจำแนกข้อมูลตามเหตุการณ์ (Event) หรือ การวิเคราะห์เหตุการณ์ ตามเรื่องราว (Event Analysis) ที่ปรากฏ ซึ่งการวิเคราะห์ข้อมูลในระดับมหภาค แบ่งได้เป็น 2 วิธีคือ การวิเคราะห์เหตุการณ์แบบอิงทฤษฎี และแบบไม่อิงทฤษฎี

2.1 การวิเคราะห์เหตุการณ์แบบอิงทฤษฎี คือ การแยกชนิดในเหตุการณ์นั้น ๆ โดย การยึดแนวคิดหรือกรอบการจำแนกเหตุการณ์โดยอาศัยทฤษฎีใดทฤษฎีหนึ่งเป็นกรอบการจำแนก ซึ่งกรอบการจำแนกเหตุการณ์ที่นิยมใช้หรือมีลักษณะกลาง ๆ ที่มักนำ มาใช้ร่วมกันคือ การวิเคราะห์เหตุการณ์

ออกเป็น 6 ประเภท คือ ใคร ทำอะไร ที่ไหน เมื่อไร อย่างไรและทำไม นอกจากนี้ยังสามารถแยกชนิดออกเป็น 6 ชนิด ตามกรอบปรากฏการณ์ ได้ดังนี้ ดังนี้

2.1.1 การกระทำ (Acts) คือ พฤติกรรมของบุคคลในเหตุการณ์หรือสถานการณ์ที่เกิดขึ้น ในช่วงระยะเวลาใดเวลาหนึ่ง ไม่ยาวนานหรือต่อเนื่อง

2.1.2 กิจกรรม (Activities) คือ เหตุการณ์หรือสถานการณ์ขอบประเพณีที่เกิดขึ้นใน ลักษณะต่อเนื่อง กิจกรรมเป็นการกระทำ ที่เกี่ยวข้องกับคนหลายคนมากกว่าจะเป็นส่วนบุคคลหรือส่วนตัวและมักใช้เวลา นานเป็นวัน สัปดาห์ เดือนหรือช่วงฤดู

2.1.3 ความหมาย (Meaning) คือ การให้ความหมายของการกระทำ หรือกิจกรรม โดยคำพูดที่แสดงออกของผู้ให้ข้อมูลอธิบาย ให้คำจำกัดความและทิศทางของการกระทำต่อสิ่งของ เหตุการณ์ และคุณลักษณะต่าง ๆ ของมนุษย์ เช่น ทำตามวัฒนธรรม ความเชื่อ บรรทัดฐาน อุดมการณ์หรืออคติต่าง ๆ ทั้งทางบวกและทางลบ

2.1.4 การมีส่วนร่วมในกิจกรรม (Participation) คือ การที่บุคคลมีความผูกพันและเข้า ร่วมกิจกรรม สถานการณ์หรือเหตุการณ์ที่เกิดขึ้น

2.1.5 ความสัมพันธ์ (Relationship) คือ ความเกี่ยวข้องระหว่างบุคคลหลาย ๆ คน ในสังคมหรือเหตุการณ์ที่ศึกษาที่เกิดขึ้นในช่วงเวลา อาจเป็นความสัมพันธ์ของกิจกรรมที่เชื่อมโยงรูปแบบต่าง ๆ ไม่ว่าจะเป็นทางแนวตั้งหรือแนวนอน ทางมิตรหรือคู่อริ ทางปกป้องหรือละเลย รวมทั้งความสัมพันธ์เกี่ยวกับ สถานภาพของมนุษย์ เช่น ความเป็นบิดามารดา เครือญาติ ความเป็นครูกับศิษย์หรือเจ้านายกับลูกน้อง เป็นต้น

2.1.6 สถานการณ์หรือสภาพการณ์ (Setting) คือ สถานการณ์หรือสภาพการณ์ที่การ กระทำ หรือกิจกรรมที่กำลังศึกษาอยู่ เกิดขึ้น ซึ่งอาจเป็นสถานที่ เช่น ในห้องประชุม ห้องเรียนหรือสถาบัน เช่น โรงเรียน โรงพยาบาล มหาวิทยาลัย ตำ นักงานเขตพื้นที่การศึกษา ตำ นักงานปลัดกระทรวงหรืออาจเป็นระบบ เช่น ระบบราชการ ระบบเอกชน เป็นต้น

2.2 การวิเคราะห์เหตุการณ์แบบไม่อิงทฤษฎี คือ การจำแนกข้อมูลในเหตุการณ์หนึ่ง ๆ ที่จะวิเคราะห์ตามความเหมาะสมกับข้อมูล ซึ่งอาจใช้สามัญสำนึกหรือประสบการณ์ของผู้วิจัย ซึ่งผู้วิจัยจะจำแนก ข้อมูลเป็นชนิดง่าย ๆ ตามประเภทที่ผู้วิจัยสังเกต เมื่อจำแนกข้อมูลเป็นชนิดแล้ว ผู้วิจัยจะพิจารณา ความ สม่่าเสมอของการเกิดของข้อมูลต่าง ๆ ซึ่งจะเป็พื้นฐานในการอธิบายสาเหตุของปรากฏการณ์ นอกจากนี้ แม้ว่าผู้วิจัยจะจำแนกเองแต่ก็ต้องอาศัยมุมมองของบุคคลในสนามวิจัยเป็นพื้นฐานในการจำแนก

3. การวิเคราะห์โดยการเปรียบเทียบเหตุการณ์ (Constant Comparison) การวิเคราะห์โดยการเปรียบเทียบเหตุการณ์ คือ การใช้วิธีการเปรียบเทียบ โดยการนำข้อมูลมาเทียบ เป็นปรากฏการณ์ วิธีการนี้สามารถทำได้โดยการทำผู้วิจัยสังเกต หรือรวบรวมข้อมูลได้หลาย ๆ อย่างแล้วนำมาแยกตามชนิดนำมาเปรียบเทียบกันโดยทำตารางหาความสัมพันธ์จากสิ่งต่าง ๆ เหล่านั้นและสรุปผลออกมาผลที่ได้จากการวิเคราะห์ด้วยวิธีการนี้จะทำให้ได้ข้อสรุปที่มีความเป็นนามธรรมมากขึ้นและครอบคลุมหรือ สามารถใช้อ้างอิงเหตุการณ์ที่เหมาะสม

4. การวิเคราะห์ส่วนประกอบ (Component Analysis) การวิเคราะห์ส่วนประกอบของข้อมูล เป็นการวิเคราะห์คุณสมบัติของส่วนประกอบของข้อมูลแต่ละชุดแล้วนำคุณสมบัติของส่วนประกอบของข้อมูล มาเปรียบเทียบเพื่อหาลักษณะร่วมที่เหมือนกันและแตกต่างกันหลังจากนั้นจึงทำการสรุปบรรยายให้เห็นถึงความหมายของข้อมูลเหล่านั้น โดยการวิเคราะห์ส่วนประกอบจะกระทำ ได้ก็ต่อเมื่อมีข้อมูลตั้งแต่สองชุดขึ้นไป

5. การวิเคราะห์สรุปอุปนัย (Analytic Induction) การวิเคราะห์แบบอุปนัย คือ การตีความสร้างข้อสรุปข้อมูลจากสิ่งที่เป็นรูปธรรมหรือปรากฏการณ์ที่มองเห็นที่เก็บรวบรวมมาได้จากข้อมูลตั้งแต่ 2 ชุดขึ้นไป เช่น การปฏิบัติงาน พฤติกรรมการสอน ตลอดจนการดำเนินชีวิต ความเป็นอยู่ ฯลฯ

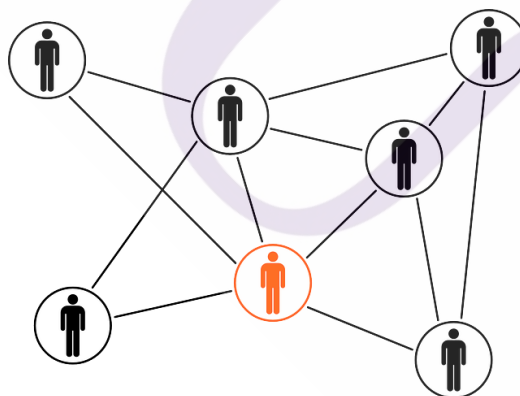
2.3 การวิเคราะห์ข้อมูล Data Mining Techniques

Data Mining เป็นเทคนิคในการวิเคราะห์ข้อมูลอย่างหนึ่ง ซึ่งมาจากคำว่า “เหมืองข้อมูล” Eakasit Pacharawongsakda, PH.D., An Introduction to Data Mining Techniques (2557, น8-9) ซึ่งเป็นศัพท์ที่ใช้เปรียบกับการขุดเหมืองแร่ทั่วไป โดยในการขุดเหมืองแร่นั้นสิ่งที่ต้องการก็คือแร่ที่มีค่า เช่น เพชร พลอย ต่างๆ ในการขั้นตอนการทำเหมืองแร่นั้นจะต้องระเบิดภูเขาหลายๆลูก เพื่อค้นหาแร่ที่ต้องการ ซึ่งแร่ที่พบนั้น ก็ออกมาน้อยมาเมื่อเทียบกับหินที่โดนระเบิดจากภูเขา เช่นเดียวกันเมื่อในองค์กร หรือบริษัทมีภูเขาของข้อมูลที่มีขนาดมหาศาล บริษัทจึงต้องมีการขุดค้นหาลงไปข้อมูลเหล่านั้น เพื่อให้ได้สิ่งที่มีค่าในข้อมูลเหล่านั้น ตามหลักวิชาการมีการให้ความหมายไว้หลายอย่างของ Data Mining เช่น “The exploration and analysis of large quantities of data in order discover meaningful patterns and rules เป็นการวิเคราะห์ข้อมูลเพื่อค้นหารูปแบบ (Patterns) หรือกฎ (rules) ที่เกิดขึ้นในฐานข้อมูลขนาดใหญ่”, Extraction of interesting (non-trivial, previously, unknown and

potential useful) information from data in large databases เป็นกระบวนการดึงข่าวสารที่น่าสนใจ และมีประโยชน์แต่ไม่เคยรู้มาก่อนจากฐานข้อมูลขนาดใหญ่ สรุปได้ว่า **Data Mining** คือการค้นหาสิ่งที่มีประโยชน์จากฐานข้อมูลที่มีขนาดใหญ่ จากข้อมูลแบบที่มีโครงสร้าง Structured data และข้อมูลแบบที่ไม่มีโครงสร้าง Unstructured data ซึ่งการวิเคราะห์ข้อมูลด้วยเทคนิค Data Mining แบ่งออกได้ 2 ประเภทหลักๆคือ

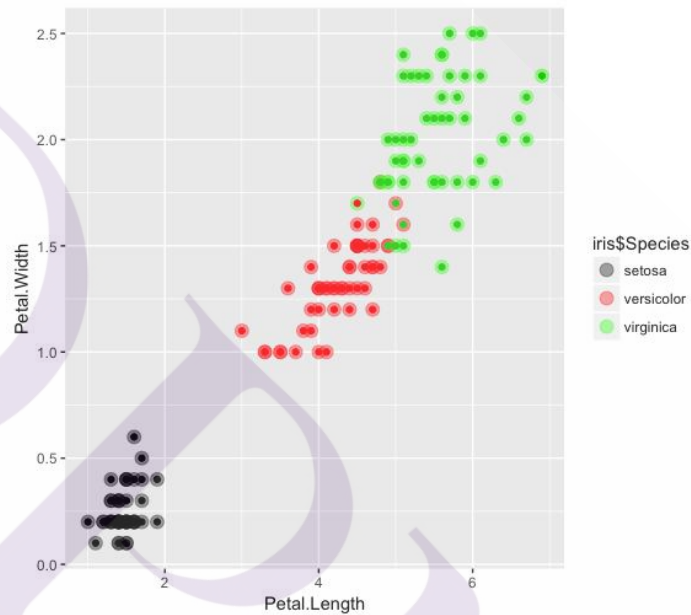
- เทคนิคการเรียนรู้แบบไม่มีผู้สอน (Unsupervised learning)
- เทคนิคการเรียนรู้แบบมีผู้สอน (Supervised learning)

Association Rules ศศ.วิภาวรรณ บัวทอง(2557) กล่าวว่า เป็นเทคนิคหนึ่งของ Data Mining คือการค้นหาความสัมพันธ์ของข้อมูล จากข้อมูลขนาดใหญ่ที่มีอยู่เพื่อนำไปหารูปแบบที่เกิดขึ้นบ่อยๆ (frequent pattern) และใช้ในการวิเคราะห์ความสัมพันธ์หรือทำนายปรากฏการณ์ต่างๆ ฐานข้อมูลที่ใช้ในการทำเหมืองความสัมพันธ์ (Association Mining) มักเป็นฐานข้อมูลประเภท Transaction Database ผลลัพธ์ที่ได้เป็นกฎความสัมพันธ์ (Association Rule) สามารถเขียนได้ ในรูปเซตของรายการที่เป็นเหตุ ไปสู่เซตของรายการที่เป็นผล ซึ่งมี รากฐานมาจากการวิเคราะห์ตะกร้าตลาด (Market Basket Analysis) เช่น ลูกค้าที่ซื้อผ้าอ้อมส่วนใหญ่มักจะซื้อเบียร์ด้วย ข้อมูลที่นำมาใช้จะอยู่ในรูปแบบ Nominal หรือ Ordinal เท่านั้น



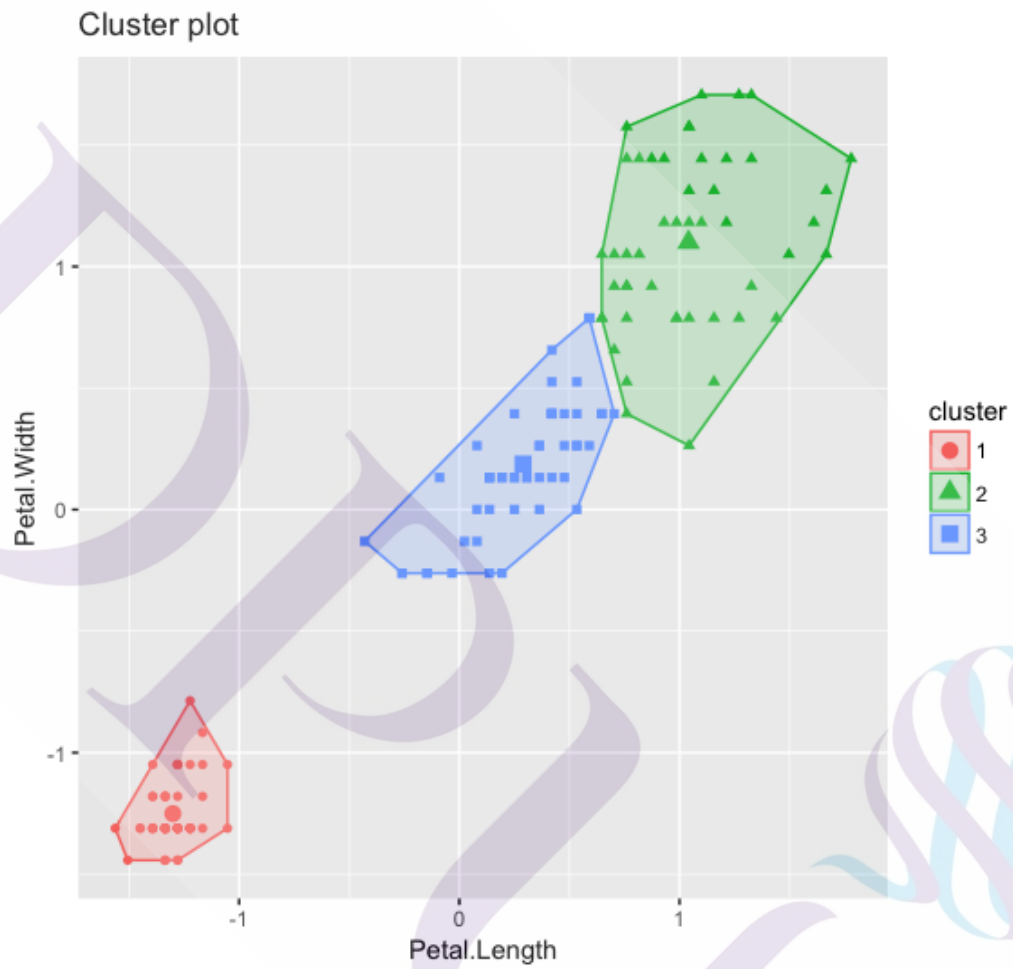
ภาพที่ 2.1 แสดง Association rule

Clustering เป็นเทคนิคการแบ่งกลุ่มข้อมูลหรือการจัดกลุ่มของข้อมูล เป็นหนึ่งใน method ของ Unsupervised Learning โดยทั่วไปแล้วก็จะมี algorithm เช่น K-Means หรือ Hierarchical Clustering ซึ่งทั้งสอง algorithm นี้จะเหมาะสำหรับ cluster ที่รวมกันเป็นกลุ่มก้อน (compact) และแยกออกจากกันอย่างชัดเจน



ภาพที่ 2.2 แสดง algorithm Hierarchical Clustering ตัวอย่างการ clustering จาก dataset iris

(Cluster ของดอกไม้ 3 species ใน genus Iris โดยการใช้ Hierarchical Clustering)



ภาพที่ 2.3 แสดง algorithm K-Means ตัวอย่างการ clustering จาก dataset iris

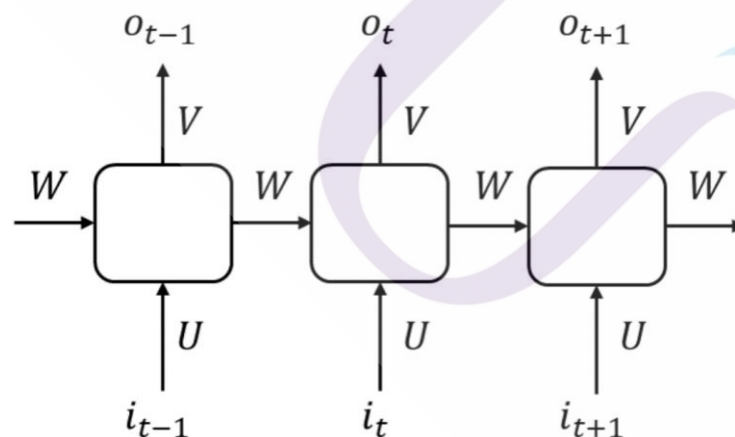
(Cluster ของดอกไม้ 3 species ใน genus Iris โดยการใช้ K-means โดยที่ $k = 3$)

Classification เป็นกระบวนการสร้างโมเดล จัดการข้อมูลให้อยู่กลุ่มที่กำหนดมาให้ เช่น จัดกลุ่มนักเรียนว่า ดีมาก ดี ปานกลาง ไม่ดี โดยพิจารณาจากประวัติและผลการเรียน หรือแบ่งประเภทของลูกค้าว่าเชื่อถือได้ หรือเชื่อถือไม่ได้โดยพิจารณาจากข้อมูลที่มีอยู่ ซึ่งอัลกอริทึมที่ใช้ในการจำแนกที่มีอยู่คือ

- Decision Tree induction
- Naïve Bayes method
- K-nearest neighbor (K-NN)
- Neural Network

2.4 Machine Learning and LSTM Model

Recurrent Neural Networks (RNNs) หมายถึง การหวนย้อนคืนกลับมา ซึ่งการทำงานของ RNNs คือการเอาผลลัพธ์ที่ได้จากการคำนวณย้อนกลับมาใช้เป็นข้อมูลขาเข้าอีกครั้ง ดังรูป ซึ่งมีประโยชน์อย่างมากในข้อมูลที่มีความต่อเนื่อง เช่น time series ข้อมูลเสียง ข้อความ หรือรูปภาพ



ภาพที่ 2.4 แสดง Recurrent Neural Networks (RNNs)

ภาพที่ 2.4 แสดง Recurrent Neural Networks (RNNs)

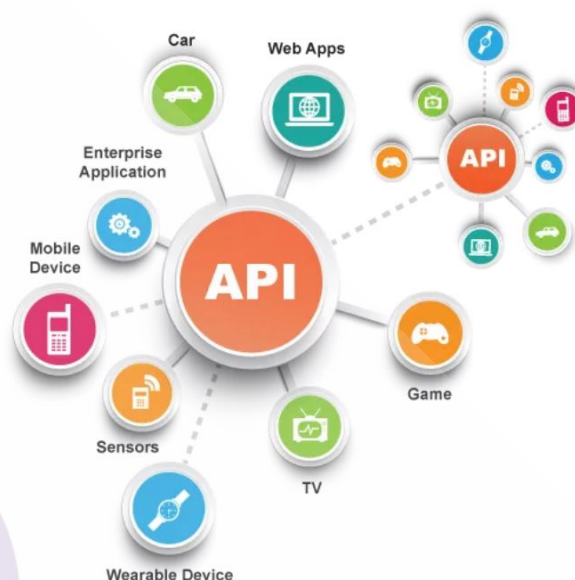
ในแต่ละ node ของ RNNs จะมีข้อมูลขาเข้า 2 อย่าง ได้แก่ input ณ node นั้น ๆ และผลลัพธ์ที่ได้จากการคำนวณใน node ก่อนหน้า ซึ่งทั้ง 2 ข้อมูลจะถูกนำมารวมเข้าด้วยกันและออกผลลัพธ์มาเป็น 2 ทางคือ ผลลัพธ์ที่ออก ณ node นั้น ๆ และออกเพื่อไปเข้าเป็นข้อมูลขาเข้าใน node ถัดไป ข้อดีของ RNNs คือ มีการใช้ข้อมูลก่อนหน้าในการทำนายสิ่งที่จะเกิดขึ้นในอนาคต นั้นหมายถึง อะไรที่ เคยเกิดขึ้นในอดีตย่อมส่งผลกระทบต่อเหตุการณ์ที่จะเกิดขึ้นในอนาคตด้วย แต่ RNNs จะมีข้อดีในการทำงานของข้อมูลที่มีความต่อเนื่อง แต่ยังมีข้อเสียของ RNNs ที่ต้องพบเจอคือ สามารถย้อนกลับได้แค่เพียงในช่วงระยะเวลาสั้น ๆ เท่านั้น ซึ่งปัญหาหลัก ๆ ของ RNNs นั้นเกิดมาจากค่า gradient ที่เริ่มน้อยลงในข้อมูลที่มีความยาวมากขึ้น จนแทบจะไม่สามารถเห็นความเปลี่ยนแปลงของ gradient ได้เลย ซึ่งปัญหานี้ถูกเรียกว่า Vanishing Gradient Problem

สืบเนื่องจากปัญหาที่เกิดขึ้นใน RNNs เกี่ยวกับค่า gradient ที่มีค่าน้อยลงจากการทำงานของ back-propagation จึงได้มีการคิดค้น machine learning ตัวใหม่ที่ใช้หลักการคล้าย ๆ เดิม แต่เปลี่ยนตัวฟังก์ชันด้านในให้มีความเสถียรและมีประสิทธิภาพมากขึ้น ซึ่งนั่นก็คือ *Long Short-Term Memory* หรือเรียกย่อ ๆ ว่า LSTMs

LSTMs เกิดขึ้นครั้งแรกในปี 1997 โดย Hochreiter และ Schmidhuber ซึ่งหลักการของมันก็ง่าย ๆ คือการเก็บ ‘สถานะ’ ของแต่ละ node เอาไว้เพื่อย้อนกลับมาจะได้รู้ว่าค่านี้แท้จริงแล้วเป็นอะไรมาก่อน สิ่งที่ทำให้ LSTMs โดดเด่นขึ้นมา นั่นก็คือการที่มันสามารถเลือกได้ว่า ข้อมูลไหนที่ควรที่จะจดจำ ข้อมูลไหนที่ควรที่จะกำจัดทิ้งออกไป ผ่านการ ‘ลืม’ ของสถานะใน node นั้น ๆ

2.5 องค์ประกอบของ API

Application Program Interface (API) ซึ่งคือ คำสั่ง (Code) ที่อนุญาตให้ software program สามารถสื่อสารระหว่างกันได้ API เป็นช่องทางสำหรับขอใช้บริการคำสั่ง จาก operation system (OS) หรือ application อื่นๆ ซึ่งใช้งาน โดยติดตั้ง function และเรียกใช้งานตาม document ที่เขียนไว้



ภาพที่ 2.5 แสดง Application Program Interface (API)

ส่วนประกอบของ APIs สร้างขึ้นจากส่วนสำคัญ 2 อย่าง คือ ข้อกำหนดที่จะอธิบายการแลกเปลี่ยนข้อมูลระหว่าง program ซึ่งทำออกมาในลักษณะ document เพื่อบอกว่า request/response ต้องเป็นอย่างไร Software ที่เขียนขึ้นตามข้อกำหนด และทำการเผยแพร่ออกไปให้ใช้งาน โดยปกติแล้ว applications ที่มี APIs จะต้องถูกเขียนเป็นภาษา programming และ พัฒนาเพิ่มได้ง่าย จึงจำเป็นต้องมีการตรวจสอบโครงสร้าง API เพราะฉะนั้น API ที่ดี ผู้ที่ออกแบบต้องให้ความสำคัญในการ test เพื่อตรวจสอบ logic ที่สามารถเกิดขึ้นได้จากการใช้งาน ปัจจุบันนี้ API ถูกใช้งานใน application เพื่อสื่อสารกับ user

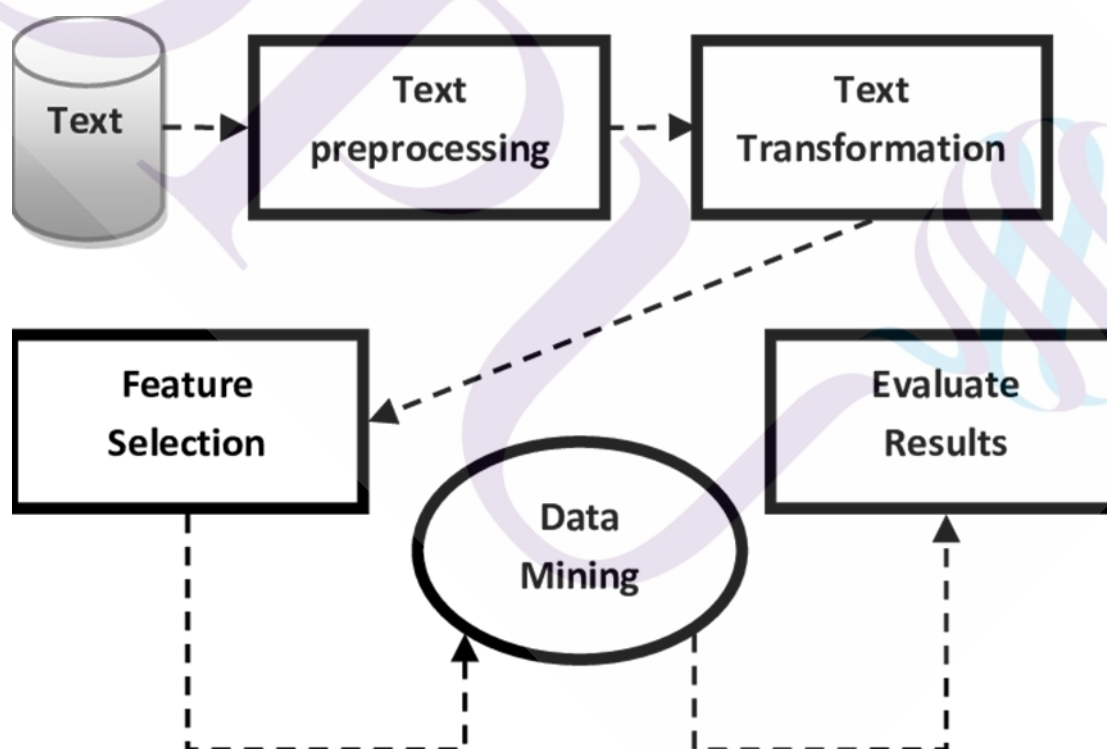
2.6 Web Application

Web Application คือ การพัฒนาระบบงานในเว็บไซต์ ซึ่งมีระบบการไหลเวียนในแบบออนไลน์ (Online) ทั้งแบบโลคอล (Local) ภายในวงแลน (Lan) และ โกลบอล (Global) ออกไปยังเครือข่ายอินเทอร์เน็ต ทำให้เหมาะสำหรับงานที่ต้องการข้อมูลแบบเรียลไทม์ (Real Time) เป็นการสร้างแอปพลิเคชันหรือโปรแกรมที่ถูกออกแบบมาเพื่อการทำงานเฉพาะด้าน และถูกเขียนขึ้นมาเพื่อเป็น

เบราว์เซอร์สำหรับการใช้งานเว็บเพจต่าง ๆ สามารถทำงานได้ในทุกอุปกรณ์ Web application เป็นโปรแกรมซอฟต์แวร์คอมพิวเตอร์ที่ทำงานบนเว็บเซิร์ฟเวอร์ซึ่งแตกต่างจาก โปรแกรมซอฟต์แวร์ที่ใช้บนคอมพิวเตอร์ สมาร์ทโฟน และแท็บเล็ต ที่ได้รับการจัดเก็บไว้ในระบบปฏิบัติการ (OS) ของอุปกรณ์ เว็บแอปพลิเคชันสามารถเข้าถึงได้ โดยใช้ผ่านเว็บเบราว์เซอร์ที่มีการเชื่อมต่ออินเทอร์เน็ต เว็บแอปพลิเคชันเป็นโปรแกรมที่ใช้บนอุปกรณ์ได้ในทันที ในการใช้งานไม่จำเป็นต้องติดตั้งบนคอมพิวเตอร์

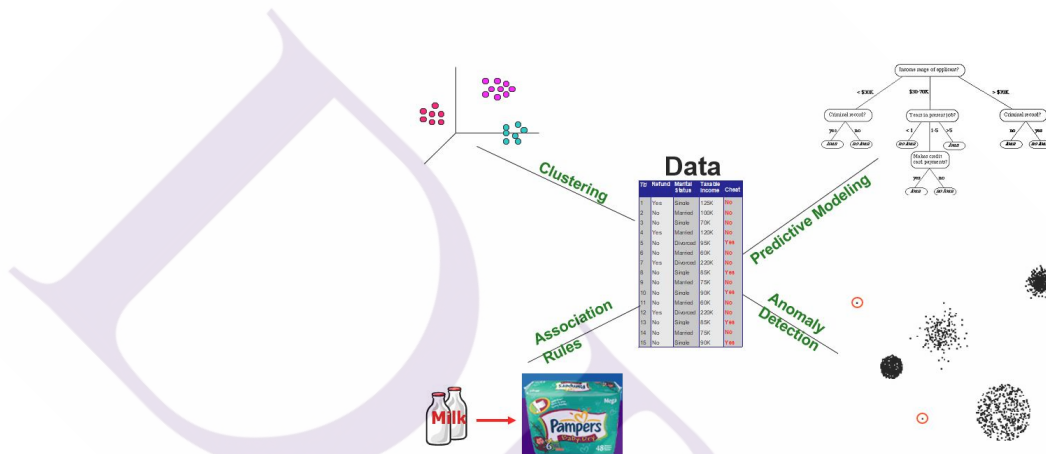
2.7 Data Text Mining

Text Mining เป็นกระบวนการที่มีการเข้าไปค้นหาพฤติกรรม หรือรูปแบบของคำ จากฐานข้อมูลขนาดใหญ่ (Large Textual Information) ที่อยู่ในรูปแบบ unstructured data to structured ให้เป็นสารสนเทศที่สามารถจัดการได้ตามรูปแบบต่างๆ โดยมีกระบวนการดังรูป



ภาพที่ 2.6 แสดงการจัดการสารสนเทศด้วยกระบวนการ Text Mining

จากภาพจะเห็นว่า การนำข้อมูลขนาดใหญ่ และเป็นทั้งข้อมูล unstructured data to structured จำเป็นจะต้องมี Process ในการทำข้อความให้อยู่ในรูปแบบที่สามารถส่งเข้ากระบวนการต่างๆ เช่น Clustering, Association Rules, Anomaly Detection, Predictive Modeling ต่างได้ เพื่อให้ได้ Results ดังภาพข้างล่างนี้



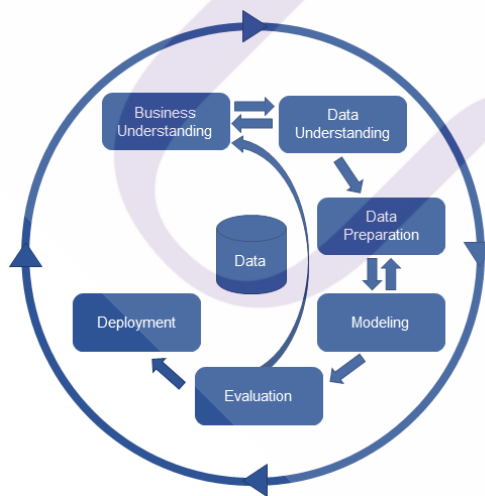
ภาพที่ 2.7 แสดงการเครื่องมือจัดการ Text Mining

บทที่ 3

ระเบียบวิธีวิจัย

3.1. แนวทางการวิจัยและพัฒนา

รวบรวมการรับรู้ผลของระดับความพึงพอใจ หรือการนำผลที่ได้จากการประเมินไปสู่การเรียนรู้ เพื่อปรับปรุงประสิทธิภาพหรือแก้ไขปัญหาในด้านต่างๆ ของกระบวนการทำงานในแต่ละงานเป็นเรื่องที่ หน่วยงาน หรือองค์กรต้องการ การสรุป รวบรวมยอด จากชุดข้อมูลจากการประเมิน ในปัจจุบันรูปแบบของการประเมินผลความพึงพอใจมีหลากหลายรูปแบบ รวมถึงวิธีการตอบคำถาม แต่มีชุดคำถามอีกประเภทที่น่าสนใจนั่นคือ การเปิดให้กรอกข้อความที่เป็น text ซึ่งข้อมูลที่ได้จะเป็นเชิงรูปแบบของประโยค ที่สามารถบอกถึงความต้องการของผู้กรอกโดยอิสระ ผู้ทำงานวิจัยจึงได้ศึกษา และออกแบบ เครื่องมือที่จะช่วยเรียนรู้ในการจัดหมวดหมู่ข้อความอัตโนมัติจากแบบประเมินผล โดยมีขั้นตอนการศึกษาและเรียนรู้ด้วยกระบวนการวิเคราะห์ข้อมูลด้วย CRISP-DM ดังนี้



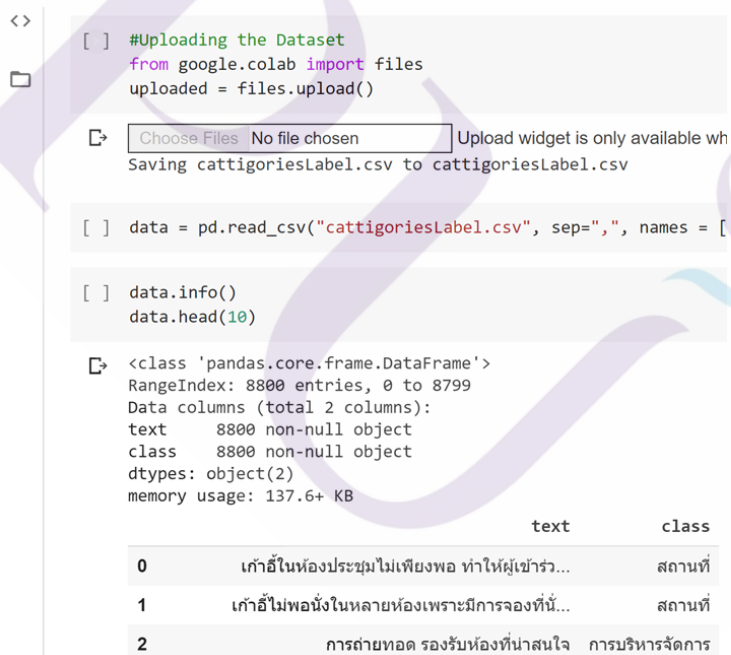
ภาพที่ 3.1 แสดงกระบวนการ CRISP-DM

3.1.1 Data Understanding

จากปัญหาของการ Priority เรื่องที่สำคัญ เป็นจุด Pain Point สำคัญและเร่งด่วนไม่ตรงตามปัญหาที่เกิดขึ้น ส่งผลต่อความเชื่อมั่น และคุณภาพของงาน และเรื่องเวลา อีกประเด็นคือปัญหาจากหน้างาน ผู้ประสานงาน ผู้รับผิดชอบโครงการ ที่สถาบันรับรองคุณภาพสถานพยาบาล มีกิจกรรมหลักคือการจัดกิจกรรม การประชุมหลักสูตร สัมมนา รวมถึงการจัดประชุมวิชาการที่มีผู้เข้าร่วมงานจำนวนมาก พบว่าขั้นตอนของการประเมินผลกิจกรรม ใช้เวลาจัดกลุ่มของชุดคำถามและข้อเสนอแนะยาก การใช้คนในการจัดการ และพบว่าเทคโนโลยีมีโอกาช่วย และมีโอกาสพัฒนา ให้เกิดความเชื่อมั่นต่อความถูกต้องของข้อมูล การสรุปผล โดยเฉพาะข้อมูลที่มาก และใช้เวลาสั้นลงทำให้เพิ่มคุณค่าของงานได้ ผู้ทำงานวิจัยจึงได้ศึกษา และออกแบบ เครื่องมือที่จะช่วยเรียนรู้ในการจัดหมวดหมู่ข้อความอัตโนมัติจากแบบประเมินผล

3.1.2 Data Preparation

ผู้วิจัยได้ใช้วิธีการเขียนชุดคำสั่งจากระบบปฏิบัติการ Google Colab โคนนำเอาข้อความจากการประเมินผลมาทำการ Data Exploration



```
<>
[ ] #Uploading the Dataset
from google.colab import files
uploaded = files.upload()

[ ] data = pd.read_csv("cattigoriesLabel.csv", sep=",", names = [

[ ] data.info()
data.head(10)

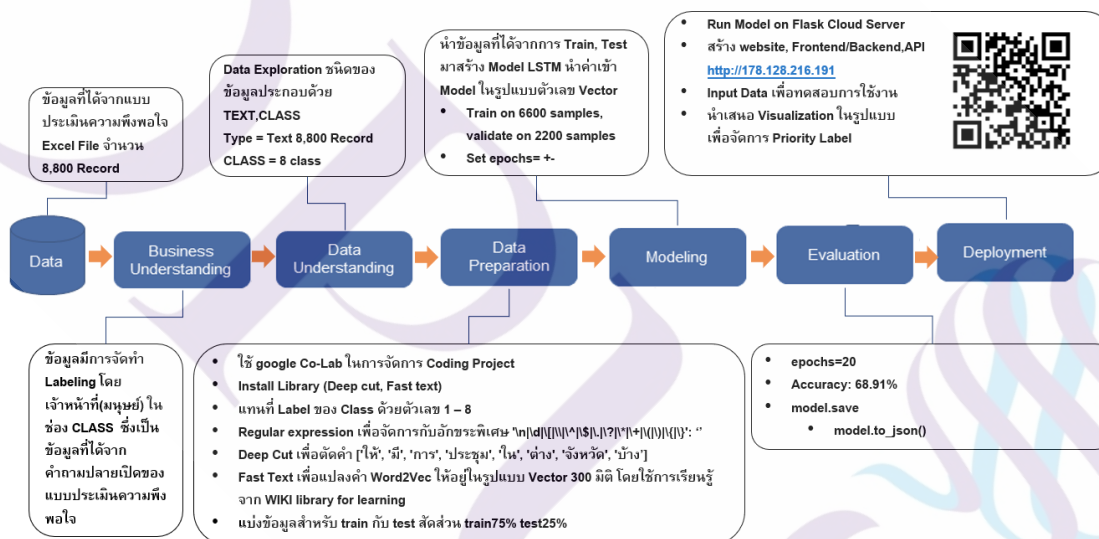
[ ] <class 'pandas.core.frame.DataFrame'>
RangeIndex: 8800 entries, 0 to 8799
Data columns (total 2 columns):
text      8800 non-null object
class     8800 non-null object
dtypes: object(2)
memory usage: 137.6+ KB

      text      class
0  เก้าอี้ในห้องประชุมไม่เพียงพอ ทำให้ผู้เข้าร่วม...  สถานที่
1  เก้าอี้ไม่พอนั่งในหลายห้องเพราะมีการจองที่นั่ง...  สถานที่
2  การถ่ายทอด รองรับห้องที่น่าสนใจ  การบริหารจัดการ
```

ภาพที่ 3.2 แสดงชุดคำสั่งของการจัดทำ Data Preparation

3.1.1.1. หลังจากนั้นผู้วิจัยได้ดำเนินการสรุปจำนวนของข้อความและ Labeling จากสร้างคำตอบ Labeling โดยนักวิชาการจำนวน 15 ท่าน โดยใช้การอ่านและรวมแสดงความคิดเห็น เพื่อตกผลึกเพื่อสร้างเป็นข้อมูลต้นแบบจำนวน 17,000 ชุดข้อความ หลังจากนั้นใช้วิธีการสร้างชุดคำสั่งในการจัดการข้อมูล Data Exploration และ Classification Data เพื่อแปลงค่า Class ของ Labeling ที่เป็นตัวเลขและจัดการกับข้อความที่ไม่สามารถใช้งานได้ (`"\n|d|[|\\|^\|$\|.\|?|*|+|\(|\)|\{|}": "`)

3.1.1.2. จากนั้นใช้ Deep cut Tokenize เพื่อตัดคำเนื่องจากเครื่องมือที่เป็น Model ที่จะใช้คือ LSTM ต้องการข้อมูลที่จะนำเข้า Model ได้ต้องเป็นตัวเลขแบบ Array จึงต้องใช้ wiki.th ในการแปลงข้อมูล เพื่อให้ได้ค่า X,Y จากนั้นจึงสร้าง Model เพื่อรับค่าเข้าเพื่อทำการ (Train Test)



ภาพที่ 3.3 แสดงกระบวนการทำงานของโปรแกรม

3.1.3 Modeling

3.1.3.1. ผู้วิจัยได้ใช้ Model LSTM ในการรับค่า x_{train} , y_{train} โดยกำหนดค่า Dense=8 ให้ Lookup=50 และ epochs=20 พบว่ามีความแม่นยำร้อยละ 68.91%

```
[ ] model_path = 'comment_model.h5'
if os.path.exists(model_path):
    # load json and create model
    json_file = open('comment_model.json', 'r')
    loaded_model_json = json_file.read()
    json_file.close()
    model = model_from_json(loaded_model_json)
    # load weights into new model
    model.load_weights("comment_model.h5")
    print("Loaded model from disk")
else:
    # create and fit the LSTM network
    look_back = 300
    model = Sequential()
    model.add(LSTM(5, input_shape=(1, look_back)))
    model.add(Dense(8))
    model.compile(loss='mean_squared_error', optimizer='adam', metrics=['accuracy'])
    history = model.fit(X_train, y_train, validation_data=(X_test, y_test), epochs=20, batch_size=1, verbose=2)
    # Final evaluation of the model
    scores = model.evaluate(X_test, y_test, verbose=0)
    print("Accuracy: %.2f%%" % (scores[1]*100))

Train on 6600 samples, validate on 2200 samples
Epoch 1/20
- 35s - loss: 0.0806 - acc: 0.5355 - val_loss: 0.0722 - val_acc: 0.6105
Epoch 2/20
- 34s - loss: 0.0664 - acc: 0.6465 - val_loss: 0.0663 - val_acc: 0.6555
```

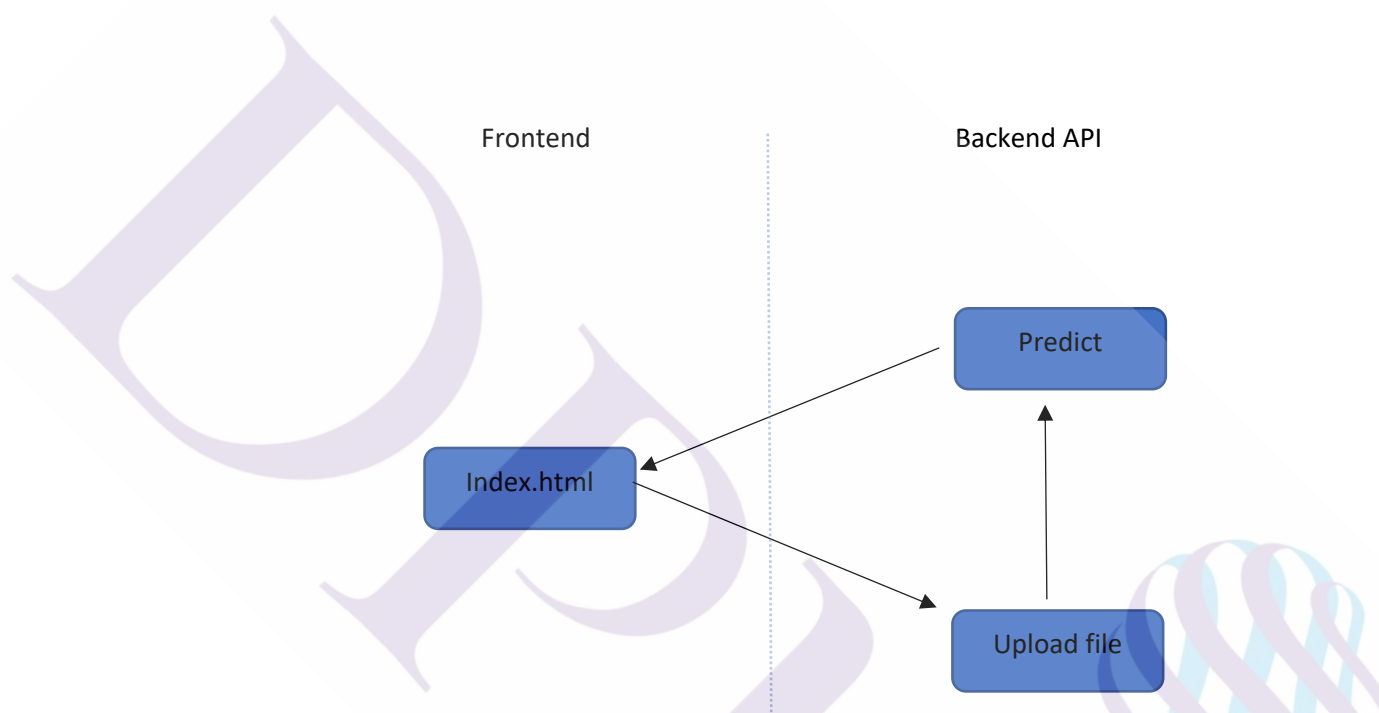
ภาพที่ 3.4 แสดงชุดคำสั่งของการจัดทำ Model

```
Epoch 20/20
- 34s - loss: 0.0502 - acc: 0.7356 - val_loss: 0.0597 - val_acc: 0.6891
Accuracy: 68.91%
```

ภาพที่ 3.5 แสดง Accuracy Model = 68.91%

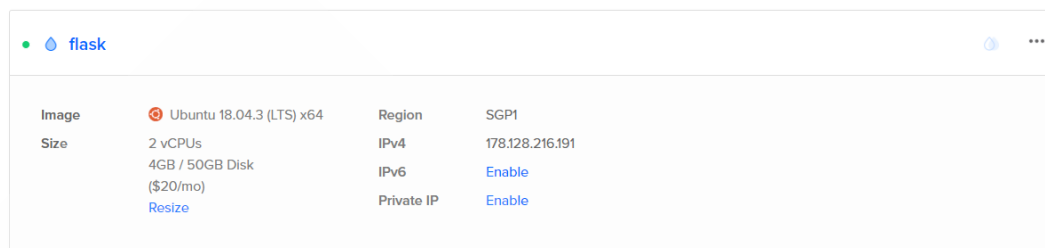
3.1.5 Deployment

จากกระบวนการข้างต้น ผู้วิจัยได้จัดเตรียมการนำข้อมูลเข้าสู่ Model เป็นที่เรียบร้อยแล้ว แต่การนำไปใช้ประโยชน์จะไม่เหมาะกับการใช้งานแบบทำให้เห็นภาพรวมได้ จึงได้ออกแบบการทำ Web Application โดยการนำ Model ที่ได้จากการ Save ออกมาจาก Google CoLab โดยออกแบบให้มี website, Frontend/Backend เพื่อให้มีการใช้งานผ่าน API และ GUI ที่ง่ายต่อการใช้งาน



ภาพที่ 3.7 แสดงการออกแบบ web Application

ซึ่ง Server ที่ใช้งานเป็นสถาปัตยกรรมแบบ Flask ระบบปฏิบัติการแบบ Ubuntu 18.04.3(LTS)x64 Ram 4GB HDD 50GB



ภาพที่ 3.8 แสดงการเลือกใช้ Server

3.1.5.1 การใช้งานผ่าน web application ผู้วิจัยได้ออกแบบใช้ Graphic User Interface (GUI) ให้มีความทันสมัยน่าใช้งาน และที่สำคัญจะต้องเข้าใจงาน โดยประกอบด้วย Menu ที่สำคัญ 4 อย่างคือ

3.1.5.1.1 Project การอธิบายถึงแนวคิดปัญหาภาพรวมของงานที่ทำ

3.1.5.1.2 Model ช่องทางการนำเข้าสู่ข้อความ หรือไฟล์ข้อมูลที่จะทำการ Prediction

Data

3.1.5.1.3 Performance การแสดงข้อมูลผ่าน Business Intelligence (BI)

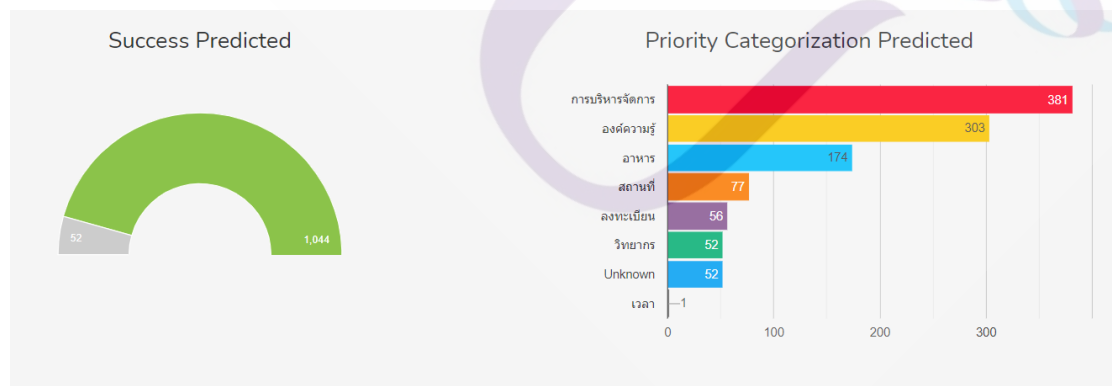
3.1.5.1.4 Education องค์กรความรู้ต่างที่ผู้วิจัยได้ทำการศึกษา

บทที่ 4 ผลการศึกษา

ผลการศึกษาการจัดหมวดหมู่ข้อความจากแบบประเมินผลแบบอัตโนมัติด้วยวิธีการเรียนรู้ของเครื่อง หลังจากที่มีการนำเข้าสู่การทำสอบผ่าน Web Application สามารถสรุปและลำดับความสำคัญของ ปัญหา Priority เรื่องที่เป็น Pain Point ที่ผู้ใช้งานจะสามารถนำไปประกอบการตัดสินใจได้อย่างมีประสิทธิภาพ โดยได้นำเสนอต่อนักวิชาการ สถาบันรับรองคุณภาพสถานพยาบาล โดยจัดทำเป็น Business Intelligence (BI) ประกอบด้วยรายละเอียดดังนี้

4.1 Business Intelligence (BI)

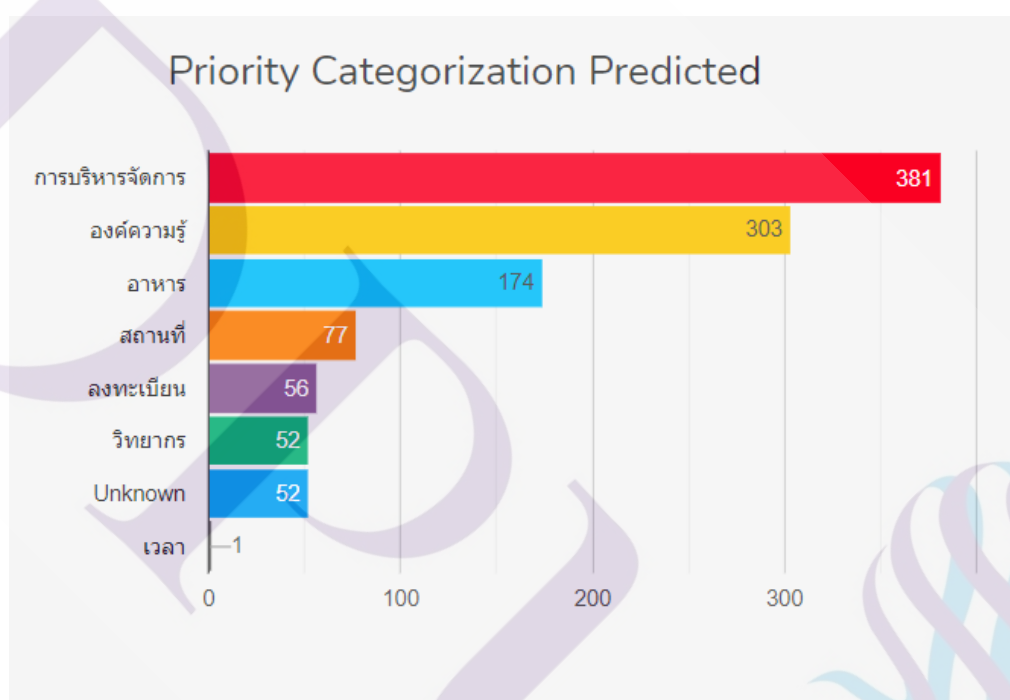
ผู้วิจัยได้ออกแบบการนำเสนอข้อมูลเพื่อให้เห็น การนำข้อความเข้าสู่การ Predict จากจำนวน 1,096 ประโยค และผลของการประมวลผล Success Predicted ที่สำเร็จจำนวน คิดเป็นร้อยละ ซึ่งใช้เวลาโดยเฉลี่ยในการประมวลผลต่อคำใช้เวลา 0.81 วินาที/คำ



ภาพที่ 4.1 แสดงการนำข้อความเข้าสู่การ Predict และ Success Predicted

4.2 การแสดงผลของการวิเคราะห์ข้อมูล

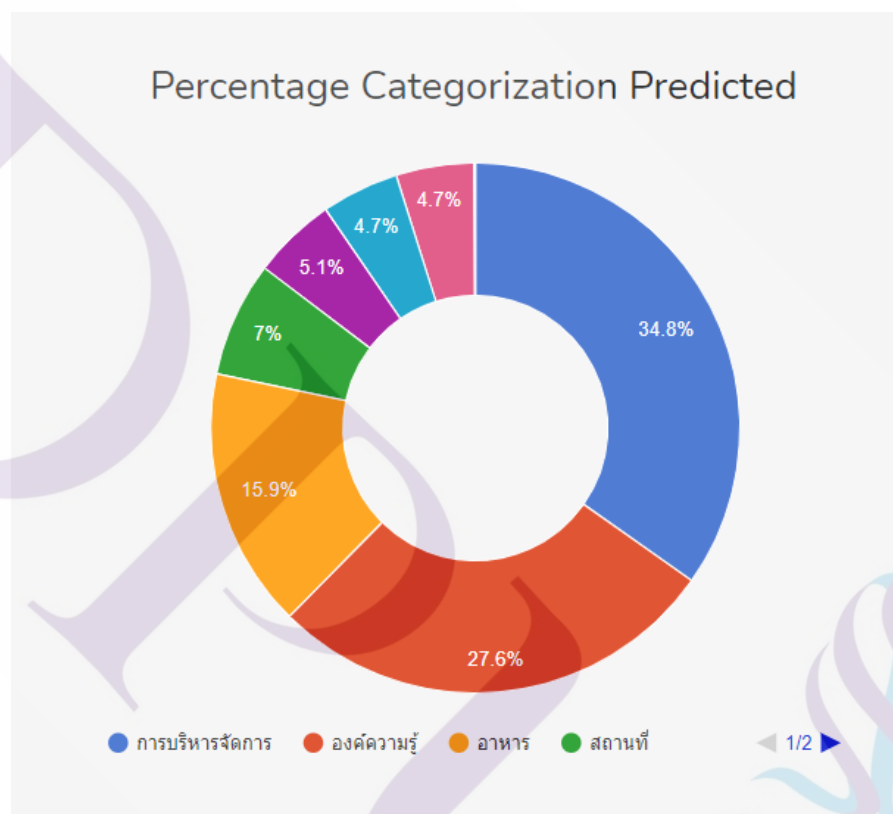
การแสดงผลของการวิเคราะห์ข้อมูลให้สามารถระบุลำดับความสำคัญ เป็นลำดับของการนำไปจัดการกับปัญหาที่เกิดขึ้น และนอกจากนั้นยังสามารถเข้าดูประโยชน์ข้อความของหมวดหมู่การ Priority แล้วจากการเลือกไปยังกลุ่ม Label



ภาพที่ 4.2 แสดงการนำข้อมูลการจัดเรียง Priority จากมากไปน้อย

4.3 การแสดงข้อมูลแบบจำแนกปัญหา

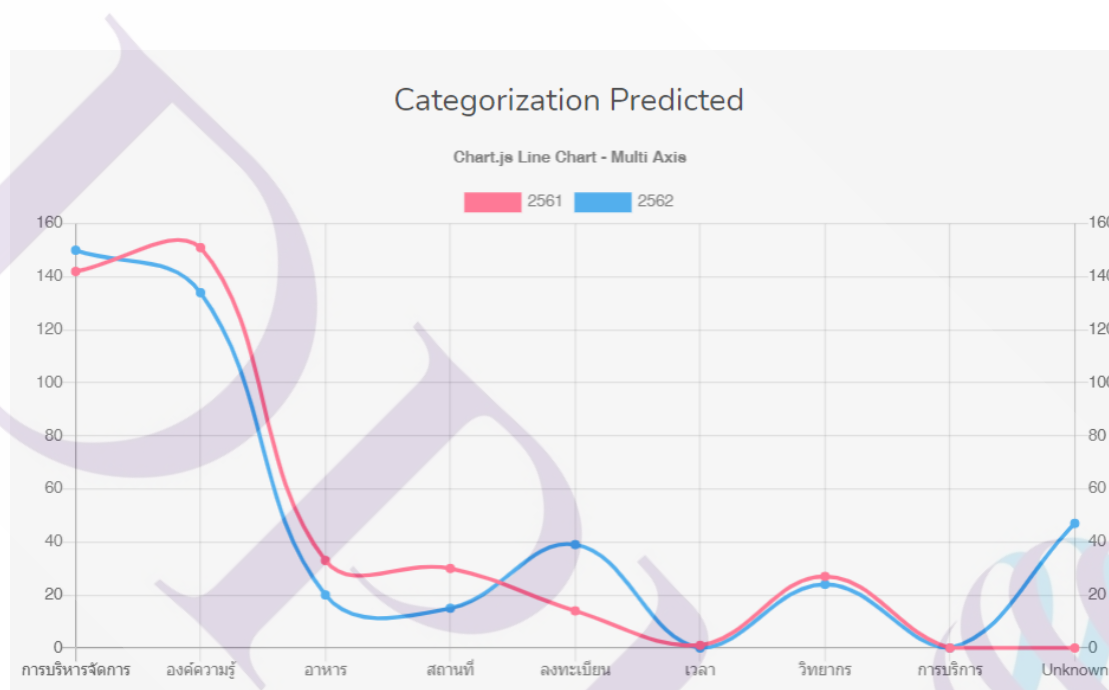
การแสดงข้อมูลแบบจำแนกปัญหาเพื่อให้เห็นสัดส่วนของข้อมูล Percentage Categorization Predicted โดยการนำผลที่ได้มานำเสนอตาม Labeling ที่จัดเตรียมไว้ใน Model



ภาพที่ 4.4 แสดงการจำแนกผลของการ Predict

4.4 จากการนำเสนอข้อมูลแบบรายปี

จากการนำเสนอข้อมูลที่ผ่านมาข้างต้น ผู้วิจัยได้เรียนรู้ว่าการมองภาพรวมของแต่ละปีมีความสำคัญอย่างยิ่งจึงได้ออกแบบกราฟที่สามารถแสดงภาพรวมทั้งปี เมื่อมีข้อมูลที่มากกว่า 2 ปีขึ้นไปมาเปรียบเทียบก็จะสามารถมองเห็นภาพของแนวโน้ม เพื่อนำไปสู่การตัดสินใจได้ง่ายมากขึ้น ผู้วิจัยจึงได้ออกแบบกราฟเปรียบเทียบปัญหาเป็นกราฟเส้น เพื่อเปรียบเทียบรายปี



ภาพที่ 4.5 แสดงกราฟเปรียบเทียบปัญหาเป็นกราฟเส้น เพื่อเปรียบเทียบรายปี

4.5 ผลประเมินจากผู้ร่วมทดสอบการใช้งาน

ผลประเมินจากผู้ร่วมทดสอบการใช้งานและประเมินการนำเสนอต่อนักวิชาการ จำนวน 2 ครั้ง ดังนี้

4.5.1 เป็นการเรียนรู้จากการใช้เครื่องมือ Machine Learning โดยมีนักวิชาการและผู้ปฏิบัติงานของสถาบันรับรองคุณภาพสถานพยาบาล (องค์การมหาชน) จำนวน 30 ท่าน เป็นการเรียนรู้การใช้งาน

และทดสอบ จึงเลือกใช้การประเมินแบบ Discussion เพื่อหาโอกาสในการพัฒนา ซึ่งสรุปได้ว่ามี 2 ประเด็นสำคัญ คือ

4.5.1.1 การนำไปใช้ประโยชน์ต่อระบบสาธารณสุขในแง่ของการรับรู้ความต้องการของการให้บริการ และรับบริการต่อประชาชน ผู้ป่วยและญาติได้

4.5.1.2 การประยุกต์ใช้ข้อมูลที่แสดงใน Business Intelligence (BI) นี้สามารถประกอบการตัดสินใจได้ในงานที่สำคัญของระบบสาธารณสุขได้ในอนาคต

4.5.2 นำเสนอต่อเวทีการประชุมวิชาการ Knowledge Management โดยมีเจ้าหน้าที่ประจำสถาบันรับรองคุณภาพจำนวน 72 คน และได้มีการให้ประเมินแบบสอบถามความพึงพอใจต่อการจัดการกับปัญหาของงานวิจัยดังกล่าวผลของการประเมินมีดังนี้

ตารางที่ 4.1 แสดงข้อมูลการรับรู้เทคโนโลยีเกี่ยวกับ Bigdata

ท่านรู้จักเทคโนโลยีเกี่ยวกับ Bigdata หรือไม่	จำนวน	ร้อยละ (%)
เคย	53	73.6
ไม่เคย	19	26.4
รวม	72	100

จากตารางพบว่า กลุ่มเจ้าหน้าที่ มีการติดตามข้อมูล ข่าวสาร การก้าวกระโดดของเทคโนโลยีในสมัยปัจจุบันซึ่งจากการการประเมินผลด้วยคำถาม ท่านรู้จักเทคโนโลยีเกี่ยวกับ Bigdata หรือไม่ พบว่ามีถึงร้อยละ 73.6 ที่ทราบเกี่ยวกับเทคโนโลยีด้านนี้

ตารางที่ 4.2 แสดงการทำงานของตนเองที่เกี่ยวข้องกับการใช้แบบประเมิน ในการจัดกิจกรรมการประชุม
หลักสูตร สัมมนา

ท่านใช้แบบประเมินความพึงพอใจต่อการจัด กิจกรรม การประชุมหลักสูตร สัมมนา ในงาน ของท่านหรือไม่	จำนวน	ร้อยละ (%)
ใช้	72	100
ไม่ใช้	0	0
รวม	72	100

จากตารางพบว่าหน่วยงานหรือเจ้าหน้าที่ประจำหน่วยงานมีการใช้แบบประเมินความพึงพอใจอยู่แล้วหลังจากการมีการจัดประชุม หรือสัมมนาต่างๆ ซึ่งพบว่ามีถึงร้อยละ 100

ตารางที่ 4.3 แสดงการทำงานของตนเองที่เกี่ยวข้องกับวิธีประมวลผลแบบประเมิน ในการจัดกิจกรรม การ
ประชุมหลักสูตร สัมมนา สำหรับคำถามปลายเปิด

ท่านเลือกใช้วิธีการการประมวลผลแบบประเมิน ความพึงพอใจกับคำถามปลายเปิดในข้อใดมากที่สุด	จำนวน	ร้อยละ (%)
นับความถี่ของข้อความ	35	48.6
Word Cloud	5	6.9
อ่านและจัดกลุ่มประโยค	32	44.4
อื่นๆ	0	0
รวม	72	100

จากตารางพบว่าพฤติกรรมการประมวลผลแบบสอบถามของเจ้าหน้าที่เลือกใช้วิธีนับความถี่ของข้อความมากที่สุดถึงร้อยละ 48.6 และมีการอ่านและจัดจัดกลุ่มประโยคที่มีการตอบแบบสอบถามเข้ามาถึงร้อยละ 44.4

ตารางที่ 4.4 แสดงความพึงพอใจต่อการใช้ Web Application ของผู้จัดทำงานวิจัยจากระดับการประเมิน (1 = น้อยที่สุด, 2 = น้อย, 3 = ปานกลาง, 4 = มาก, 5 = มากที่สุด)

การประเมินความพึงพอใจต่อการใช้ Web Application	ระดับผลการประเมิน	ร้อยละ (%)
การใช้งาน web application <u>ง่าย</u> ต่องานใช้งานอยู่ในระดับใด	มาก	73.3
ท่านคิดว่า <u>ความแม่นยำ</u> ต่อการประมวลผลของประโยค หรือข้อความที่ท่านได้ทดสอบอยู่ในระดับใด	ปานกลาง	52.4
การใช้งาน web application <u>สะดวก</u> ต่องานใช้งานอยู่ในระดับใด	มากที่สุด	82.5
การใช้งาน web application <u>รวดเร็ว</u> ต่องานใช้งานอยู่ในระดับใด	มาก	79.2
การใช้งาน web application <u>การแสดงผลแบบ Business Intelligence</u> ต่องานใช้งานอยู่ในระดับใด	มากที่สุด	85.5
การใช้งาน web application <u>การจัดลำดับความสำคัญของปัญหา</u> ต่องานใช้งานอยู่ในระดับใด	มาก	83
ความพึงพอใจต่อภาพรวมของการใช้ web application ของท่านอยู่ในระดับใด	มาก	76.5

จากตารางผู้วิจัยมีเจตนาที่จะสื่อสารการใช้ Machine learning for automated text categorization from evaluation การเรียนรู้ในการจัดหมวดหมู่ข้อความอัตโนมัติจากแบบประเมินผลในภาพรวมของการใช้งานเพื่อแก้ไขปัญหาของงานวิจัยนี้ ซึ่งร้อยละ 73.3 พบว่ามีการ web application งานต่อการใช้งานในระดับ มาก ความแม่นยำจากการประเมินของผู้ใช้อยู่ที่ระดับ ปานกลาง หรือร้อยละ 52.4 ในมิติของความสะดวกต่อการใช้งานอยู่ในระดับมากที่สุด และความรวดเร็วอยู่ในระดับ มาก ในแง่ของการนำไปใช้งานเกี่ยวกับ BI (Business Intelligence) เพื่อนำไปประกอบการจัด Priority ลำดับความสำคัญของปัญหาพบว่าความพึงพอใจอยู่ในระดับ มากที่สุด ร้อยละ 85.5 โดยสรุปภาพรวมของงานวิจัยนี้และการจัดทำ web application อยู่ที่ระดับ มาก ร้อยละ 76.5



บทที่ 5

บทสรุป และข้อเสนอแนะ

5.1 สรุปผลการวิจัย

จากการทำวิจัยของเพื่อสร้างระบบที่สามารถช่วยบริหาร จัดการเกี่ยวกับการจัดหมวดหมู่ของปัญหา จัดทำ Machine learning ในการเรียนรู้การจัดกลุ่มคำโดยใช้เทคโนโลยีที่เกี่ยวกับ AI เข้าช่วยงาน สร้าง Visualization ในการติดตาม และวิเคราะห์ปัญหาที่สำคัญและเร่งด่วน นำไปสู่การเรียนรู้กระบวนการทำ Machine learning จากงานวิจัยที่ศึกษา บูรณาการเข้าสู่งานประจำ พบว่าการ ประยุกต์นำไปใช้ประโยชน์เป็นเรื่องน่าใจมาก สำหรับการพัฒนาระบบสารสนเทศของระบบสาธารณสุข ของการทำงานในสถาบันรับรองคุณภาพสถานพยาบาล โดยเฉพาะการปรับใช้การรับรู้ความรู้สึก หรือความต้องการของประชาชน ผู้ป่วย หรือนุคลากรทางแพทย์ เพื่อให้ทันต่อการแก้ปัญหา ของการจัดการเรื่องสำคัญ จำเป็น และเร่งด่วน

จากผลการสำรวจความพึงพอใจที่ผ่านมาจะเห็นได้ว่าร้อยละของการความพึงพอใจในภาพรวม ต่อการใช้งานของ web application คิดเป็นร้อยละ 76.5 เป็นจุดที่สามารถคาดการณ์ได้ว่า หากมีการนำไปใช้ประยุกต์ต่อจะเกิดประโยชน์อย่างยิ่ง

5.2 ข้อเสนอแนะ

ปัญหาของการ Priority เรื่องที่สำคัญ เป็นจุด Pain Point สำคัญและเร่งด่วน ที่ส่งผลให้ไม่ตรงตามปัญหาที่เกิดขึ้น ส่งผลต่อความเชื่อมั่น และคุณภาพของงาน เรื่องเวลา และประเด็นการสร้าง Labeling ด้วยมนุษย์ในการสร้างอาจจะมีผลจุดเสี่ยงเรื่องการความเชื่อมั่น ของผู้อ่านประโยชน์นั้น แต่อย่างไรก็ดี การสร้าง Labeling จากมนุษย์ เป็นสิ่งที่ต้องใช้ประสบการณ์ ดังเช่น Model ที่ผู้วิจัยได้เลือก LSTM ใช้งานเรื่องการการเรียนรู้ หรือการตัดสินใจจากประสบการณ์เป็นสิ่งที่มีความโดยตลอด ตัวอย่างเช่น การฟังเสียงพูดของคนพูดออกออกมา และทำให้เราได้ยิน เราจะใช้ประสบการณ์เดิมที่เราเรามี

ประมวลผลได้แล้วว่า ผู้พูดเป็นคนอาศัยอยู่ในภูมิภาคใด โดยสรุป Model LSTM เป็น Model ที่เหมาะสมกับงานวิจัย ที่ใช้เรียนรู้ และแก้ปัญหาเกี่ยวกับ การจัดหมวดหมู่ข้อความจากแบบประเมินผลแบบอัตโนมัติด้วยวิธีการเรียนรู้ของเครื่อง Machine learning for automated text categorization from evaluation





บรรณานุกรม

บรรณานุกรม

คณะศิลปศาสตร์และวิทยาศาสตร์ มหาวิทยาลัยธุรกิจบัณฑิตย์ (2549,127-128)

Scriven(1967, 1974, 1976 : อ้างอิงใน ศิริชัย กาญจนวาสี,2538)

เอี่ยมพร หลินเจริญ เทคนิคการวิเคราะห์ข้อมูลเชิงคุณภาพคณะศึกษาศาสตร์ มหาวิทยาลัยมหาสารกา
,18-29)

Association Rules ผศ.วิภาวรรณ บัวทอง(2557)

Mirko Sokovic , Dusko Pavletic , Karmen KERN Pipan (2010) Quality improvement methodologies
PDCA cycle, RADAR matrix, DMAIC and DFSS, 476-483

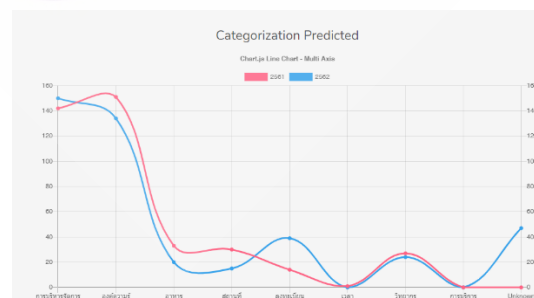
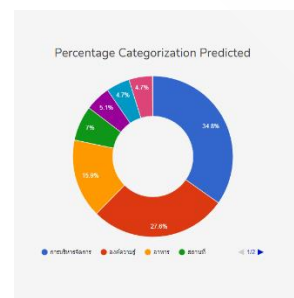
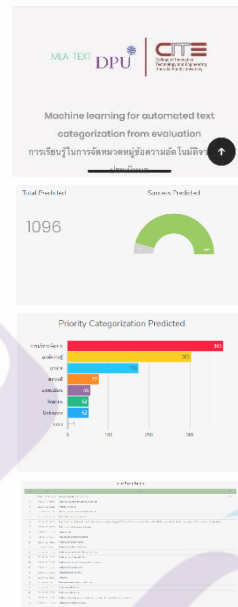
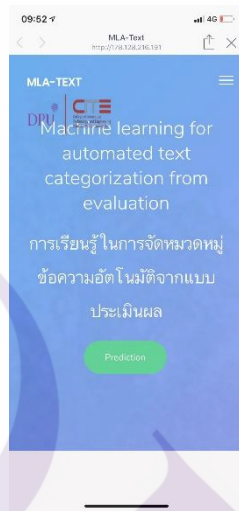
Eakasit Pacharawongsakda, PH.D., An Introduction to Data Mining Techniques (2557, น8-9)

Mirko Sokovic , Dusko Pavletic , Karmen KERN Pipan (2010) Quality improvement methodologies -
PDCA cycle, RADAR matrix, DMAIC and DFSS, 476-483

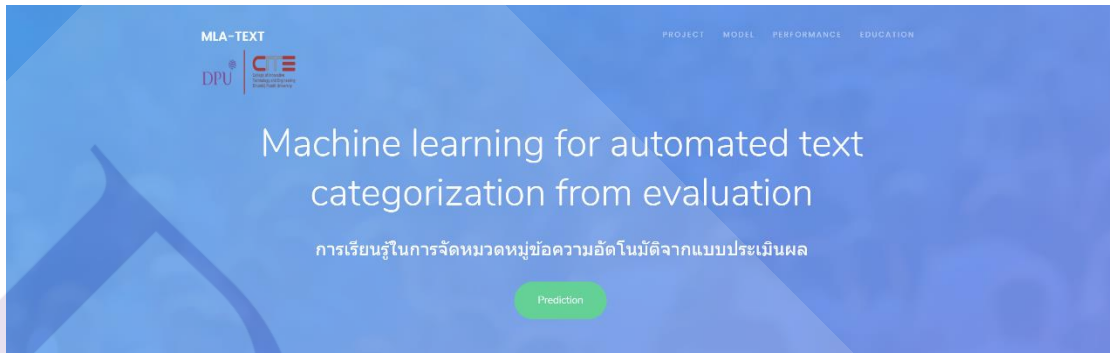


ภาคผนวก

การใช้งาน Web Application ผ่าน Mobile Device



การใช้งาน Web Application ผ่าน Web Browser



My Project

การเรียนรู้ของระดับความพึงพอใจ หรือการนำผลที่ได้จากการประเมินไปสู่ การเรียนรู้ เพื่อปรับปรุงประสิทธิภาพหรือแก้ไขปัญหาในด้านต่างๆ ของกระบวนการทำงาน ในแต่ละงานเป็นเรื่องที่ หน่วยงานหรือองค์กรต้องการ การสรุป รวบรวมยอด จากชุดข้อมูล จากการประเมิน ในปัจจุบันรูปแบบของการประเมินผลความพึงพอใจมีหลากหลายรูปแบบ รวมถึงวิธีการตอบคำถาม แต่มีชุดคำถามอีกประเภทที่น่าสนใจขึ้นคือ การเปิดให้กรอกข้อความที่เป็น TEXT ซึ่งข้อมูลที่ได้อาจเป็นเชิงรูปแบบของประโยค ที่สามารถบอกถึงความต้องการของผู้กรอกโดยอิสระ จากข้อความที่เป็นรูปแบบของชุดประโยค ในอดีตหน่วยงานจะสรุปโดยนักวิชาการหรือเจ้าของโครงการ สิ่งที่ได้จะเป็นการสรุปเป็นข้อๆ หรือจัดรวมให้เป็นเป็นกลุ่ม ซึ่งใช้เวลาในการอ่านข้อความทั้งหมดแล้วจึงสรุปโดยประสบการณ์ของผู้อ่านเอง ปัจจุบันมีเทคโนโลยีการต่างๆที่สามารถยกระดับ การอ่าน การตัด การวิเคราะห์ข้อความ รวมถึงสามารถแปลงค่าต่างๆได้ Artificial Intelligence (AI) เป็นการประยุกต์ใช้เทคโนโลยีที่จัดการกับ Pour

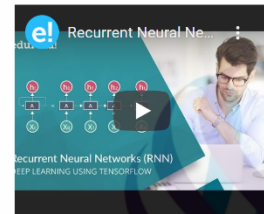


Machine learning automated

ขั้นตอนในการจัดทำ
 ออกแบบการประเมินผลของการจัดประชุมโดยเก็บข้อมูลที่สำคัญของผู้เข้าร่วมประชุมเพื่อนำมาวิเคราะห์ในส่วนต่าง ของความคิดเห็น ส่วนสำคัญของเก็บข้อมูล เพื่อให้ได้ Data Set ที่จำนวนหนึ่งที่เป็นค่าและค่า ข้อเสนอแนะ คำติชมต่างๆ ซึ่งประโยคข้อความเหล่านี้เป็น TEXT เมื่อนำมาวิเคราะห์จะพบว่าข้อความจำนวนมาก ผู้มีปฏิบัติงาน (นักวิชาการ) จะต้องทำการ อ่านด้วยตัวเองแล้วสรุป จัดกลุ่ม เพื่อหาความสัมพันธ์โดยกระบวนการใช้คนการทำงานในขั้นตอนนี้จึงเกิดการระบุ Class ของข้อมูลเก็บขึ้นจากการประมวลผลของคน โดยจัดให้อยู่ใน 8 กลุ่ม ดังนี้ [การบริหารจัดการ, องค์กรความรู้, อาหาร, 'สถานที่', ลงทะเบียน, เวลา, 'วิทยากร', การบริการ] จากนั้นนำข้อมูลเข้าสู่กระบวนการ Data Exploration เพื่อดูข้อมูล และใช้ wiki ในการตัดคำ รวมถึงแทน Class ที่ได้ทั้ง 8 คำ เป็นตัวเลข เพื่อนำเข้าสู่กระบวนการ Train/Test โดยได้ปรับให้ค่า Accuracy ที่สูงพอที่จำทำการ Prediction ได้แม่นยำที่สุด โดย library ที่ใช้คือ fasttext.



Model



LSTM หรือชื่อเต็มคือ Long Short-Term Memory เป็นโครงข่ายประสาทเทียมแบบหนึ่งที่ถูกออกแบบมาสำหรับการประมวลผลลำดับ (sequence)

ทดสอบการนำข้อความเข้าสู่ Model

Result

นำข้อมูลเข้าสู่ Model ด้วย file .tex หรือ CSV

 No file chosen



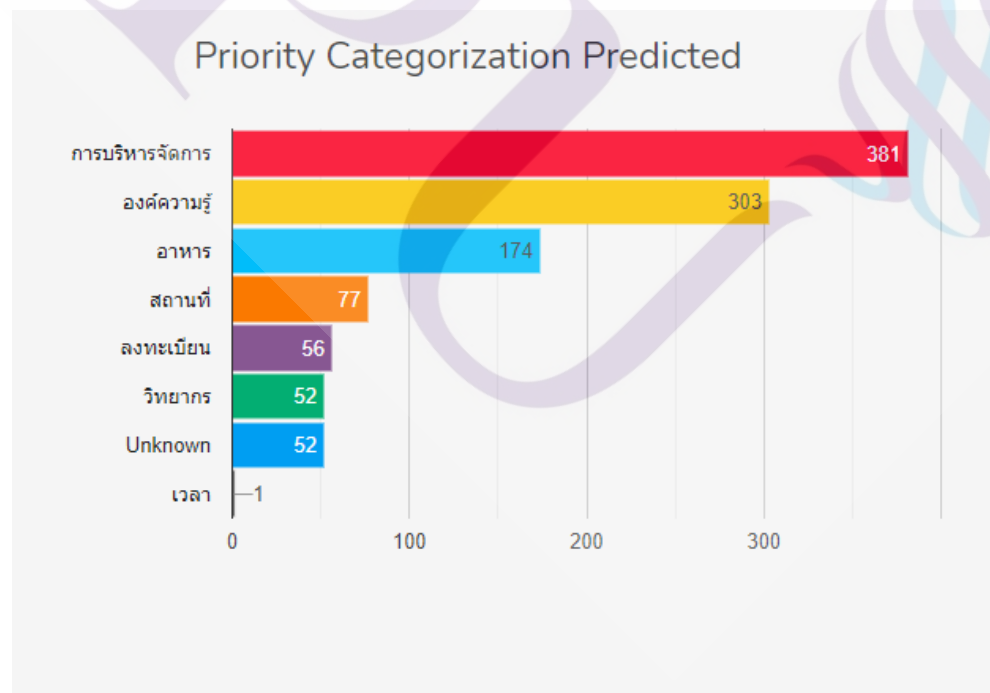
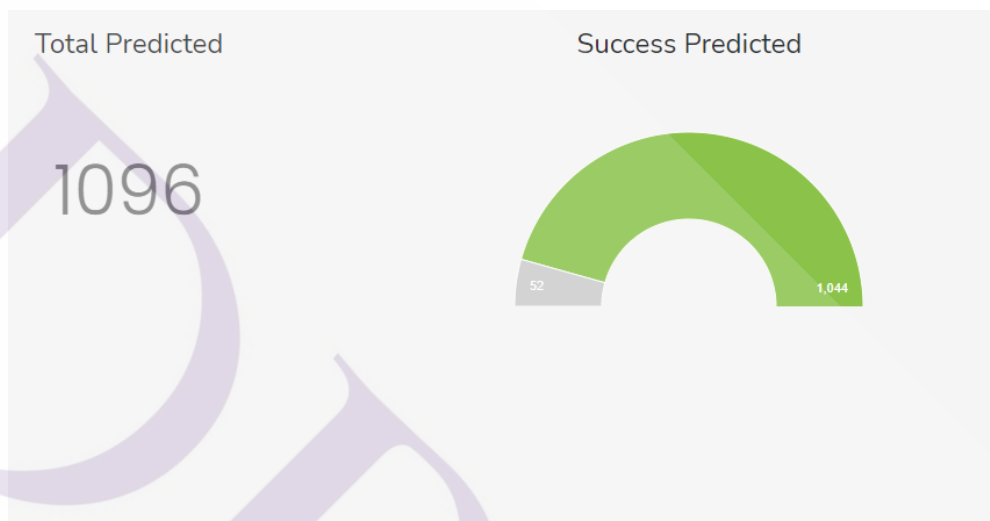
Machine learning for automated text categorization from evaluation
 การเรียนรู้ในการจัดหมวดหมู่ข้อความอัตโนมัติจากแบบประเมินผล

Sa-nguan Keawkao

Faculty of engineering, Dhurakij Pundit University, 110/1-4 Prachachuen Rd., Laksi, Bangkok 10210

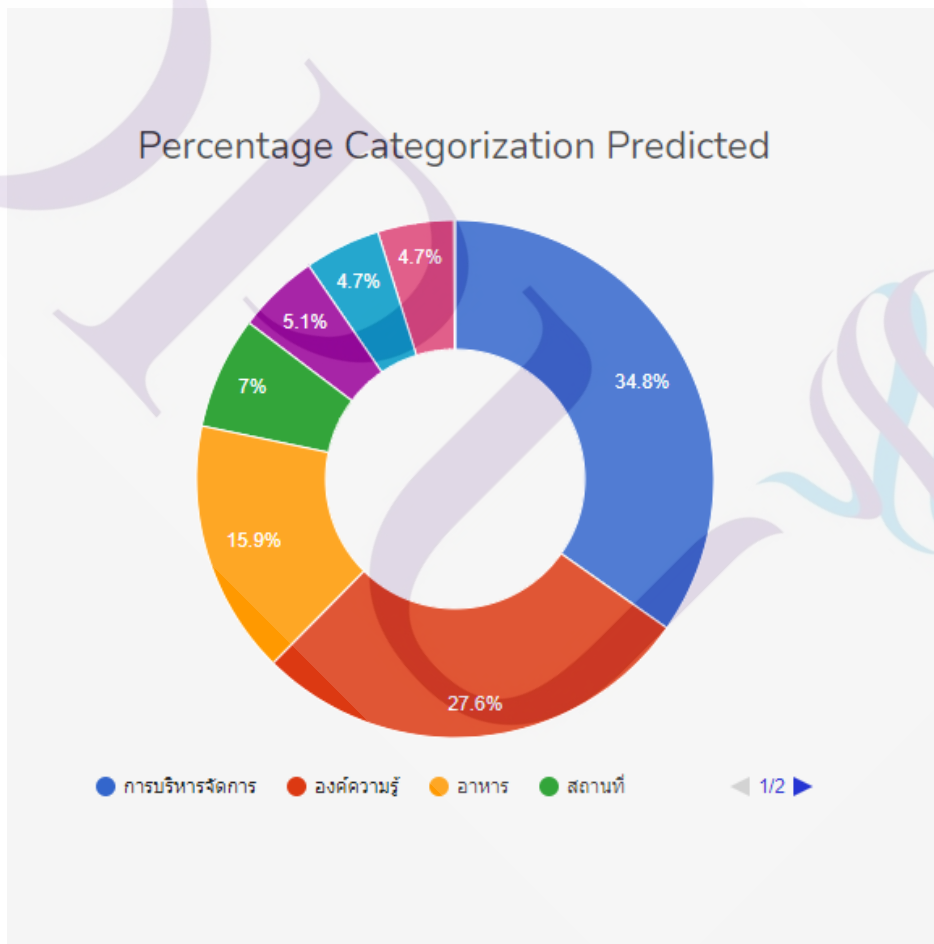
Copyright© 2020 All rights reserved | This template is made with by Colorlib
 Demo Images: Unsplash, Pexels

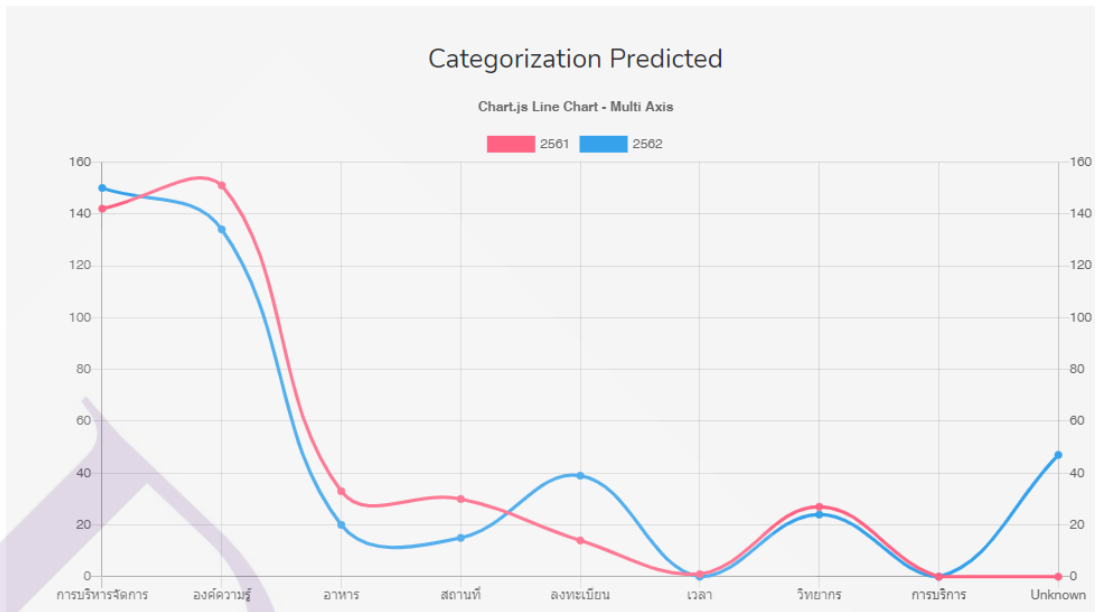
การแสดงผล Business Intelligence



การบริหารจัดการ

ร	วันที่	หัวข้อ	ผู้
1	2020-11-27 08:22:51	ได้ผู้เข้าประชุมมีสารพันการบริหารจัดการ	2561
2	2020-11-23 12:14:46	เสริมพลังใจให้เพื่อพล และเสริมพลังชีวิตวิชาการ	
3	2020-11-23 12:13:25	ตั้งชื่อประชุมใหม่คือพล	
4	2020-11-23 12:11:19	เข้าศึกษาต่อ 15-18 จะขอภาคใต้ระดับต้นและกลาง	
5	2020-11-23 12:10:24	ตั้งชื่อภาคใหม่คือพลและพลวิชาการ	
6	2020-11-23 12:08:15	ตั้งชื่อภาคใหม่คือ พล ซึ่งเป็นทุกปี เช่น พล 2P safety สภาฯให้จัดใหม่ยกกว่าให้ไปไกล (ตั้งชื่อจากเรื่องจกนพว่มีตัว) พล ใจไว้ให้มีลักษณะพลที่พลังเพิ่มเต็ม ส่วนคนงานจากที่ประชุมสามารถมีให้พลวงที่	
7	2020-11-23 12:05:18	ตั้งชื่อภาคใหม่คือพลและพลวิชาการ	
8	2020-11-23 11:59:36	ตั้งชื่อภาคใหม่	
9	2020-11-23 11:58:27	ตั้งชื่อคือพล ในห้องประชุมบางห้อง	
10	2020-11-23 11:56:33	ตั้งชื่อพลสำหรับประชุมวิชาการประชุม	
11	2020-11-23 11:56:31	ตั้งชื่อพลคือชื่ออื่น ไม่สะดวกเลย	
12	2020-11-23 11:56:28	ตั้งชื่อใหม่ สงครามชีวิตที่ไม่ได้จบเป็นอนันตภาค	
13	2020-11-23 11:56:27	ตั้งชื่อใหม่ เจ้าภาพที่อุดมไปด้วยสิ่ง	
14	2020-11-23 11:56:25	ตั้งชื่อใหม่ประชุมภาคใหม่คือพล และใช้มีการขยายพล	
15	2020-11-23 11:56:24	ตั้งชื่อใหม่ประชุมพลคือพล	
16	2020-11-23 11:56:23	ตั้งชื่อใหม่ประชุมใหม่คือพล	
17	2020-11-23 11:56:23	ตั้งชื่อใหม่คือ	
18	2020-11-23 11:56:22	ตั้งชื่อใหม่ภาคใหม่คือชื่อประชุมวิชาการ	
19	2020-11-23 11:56:22	ตั้งชื่อใหม่ภาคใหม่คือพล	
20	2020-11-23 11:56:22	ตั้งชื่อใหม่ประชุมใหม่คือพล	
21	2020-11-23 11:56:22	ตั้งชื่อในการเข้าศึกษาเพื่อพล สภาฯให้เลิกสารให้กลับเข้าชื่อ (หากมีได้ lecture จะดีมาก)	
22	2020-11-23 11:56:22	ตั้งชื่อใหม่และชื่อประชุมใหม่คือพล	
๗๖	๒๐๒๐-๑๑-๒๓ ๑๕:๕๖:๗๗	ตั้งชื่อใหม่ประชุมใหม่คือพล	





ประวัติผู้เขียน

ชื่อ-นามสกุล
ประวัติการศึกษา

นายสงวน แก้วขาว
ปี 2551
สำเร็จการศึกษาระดับปริญญาตรี
สาขาวิชาคอมพิวเตอร์ธุรกิจ
คณะเทคโนโลยีสารสนเทศ
มหาวิทยาลัยธุรกิจบัณฑิตย์

