

ระบบวิเคราะห์ข้อมูลความคิดเห็นที่มีต่อมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
จากเว็บไซต์สำหรับวัยรุ่นโดยเฉพาะ
SENTIMENT DATA ANALYTIC OF KING MONGKUT'S UNIVERSITY OF TECHNOLOGY
NORTH BANGKOK FROM TEEN ONLY SOCIETY WEBSITE

นายรัชิต กองวงศ์
นายพีรณัฐ เทศทรัพย์

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรเทคโนโลยีบัณฑิต
สาขาเทคโนโลยีสารสนเทศ ภาควิชาเทคโนโลยีสารสนเทศ
คณะเทคโนโลยีและการจัดการอุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
ปีการศึกษา 2562
ลิขสิทธิ์ของมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

ระบบวิเคราะห์ข้อมูลความคิดเห็นที่มีต่อมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
จากเว็บไซต์สำหรับวัยรุ่นโดยเฉพาะ

นายรัชิต กองวงศ์
นายพีรณัฐ เทศทรัพย์

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรเทคโนโลยีบัณฑิต
สาขาเทคโนโลยีสารสนเทศ ภาควิชาเทคโนโลยีสารสนเทศ
คณะเทคโนโลยีและการจัดการอุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
ปีการศึกษา 2562
ลิขสิทธิ์ของมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

SENTIMENT DATA ANALYTIC OF KING MONGKUT'S UNIVERSITY OF TECHNOLOGY
NORTH BANGKOK FROM TEEN ONLY SOCIETY WEBSITE

MR. RAKSIT KONGWONG
MR. PEERANUT THADSUP

PROJECT REPORT SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE BACHELOR'S DEGREE OF TECHNOLOGY IN INFORMATION TECHNOLOGY
DEPARTMENT OF INFORMATION TECHNOLOGY
FACULTY OF INDUSTRIAL TECHNOLOGY AND MANAGEMENT
KING MONGKUT'S UNIVERSITY OF TECHNOLOGY NORTH BANGKOK
2019

COPYRIGHT OF MONGKUT'S UNIVERSITY OF TECHNOLOGY NORTH BANGKOK



ใบรับรองปริญญาานิพนธ์
คณะเทคโนโลยีและการจัดการอุตสาหกรรม
มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

เรื่อง ระบบวิเคราะห์ข้อมูลความคิดเห็นที่มีต่อมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
จากเว็บไซต์สำหรับวัยรุ่นโดยเฉพาะ

โดย นายรักชิต กองวงศ์
นายพีรณัฐ เทศทรัพย์

ได้รับอนุมัติให้นับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรเทคโนโลยีบัณฑิต
สาขาวิชาเทคโนโลยีสารสนเทศ

_____ คณบดี

(รองศาสตราจารย์ ดร.อนิราช มิ่งขวัญ)

คณะกรรมการสอบปริญญาานิพนธ์

ประธานกรรมการ

(ผู้ช่วยศาสตราจารย์สุปิตี กุลจันทร์)

กรรมการ

(อาจารย์ ดร.ประติษฐ์ ทิทัศน์เสถียรกุล)

กรรมการ

(ผู้ช่วยศาสตราจารย์ ดร.วันที ประจวบศุภกิจ)

ชื่อผู้จัดทำ : นายรัชิต กองวงศ์
: นายพีรณัฐ เทศทรัพย์
ชื่อปริญญาบัตร : ระบบวิเคราะห์ข้อมูลความคิดเห็นที่มีต่อมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ จากเว็บไซต์สำหรับวัยรุ่นโดยเฉพาะ
สาขาวิชา : เทคโนโลยีสารสนเทศ
: มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
ที่ปรึกษาปริญญาบัตร : ผู้ช่วยศาสตราจารย์ ดร.วันนี ประจวบศุภกิจ
ปีการศึกษา : 2562

บทคัดย่อ

โครงการพิเศษระบบวิเคราะห์ข้อมูลความคิดเห็นที่มีต่อมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือจากเว็บไซต์สำหรับวัยรุ่นโดยเฉพาะนี้ มีวัตถุประสงค์เพื่อใช้เก็บข้อมูลความคิดเห็นในเว็บไซต์สำหรับวัยรุ่นโดยเฉพาะและนำมาวิเคราะห์ข้อมูลความคิดเห็นที่มีต่อมหาวิทยาลัยฯจะทำให้มหาวิทยาลัยฯ หรือผู้ดูแลระบบสามารถวิเคราะห์ข้อมูลหรือเหตุการณ์ที่เกี่ยวข้อง รวมถึงสามารถสรุปความคิดเห็นว่ามีความหมายเชิงบวก หรือเชิงลบ ตามช่วงเวลาต่าง ๆ ในรูปแบบของ Dashboard นอกจากนี้สามารถนำระบบมาช่วยในการตัดสินใจของผู้บริหารต่อไปได้

โครงการนี้พิเศษนี้จะเก็บข้อมูลความคิดเห็นในรูปแบบอัตโนมัติโดยจะพัฒนาในฝั่ง Server จาก NodeJs จากนั้นนำข้อมูลที่ได้เก็บลงในและนำข้อมูลที่ได้มาวิเคราะห์ความเห็นเชิงบวกเชิงลบและจัดประเภทของความคิดเห็นด้วย Python จากนั้นนำข้อมูลที่วิเคราะห์ได้มาแสดงผลให้ผู้ใช้งานดูด้วย PHP, JavaScript และอาศัย HTML, CSS เป็นตัวช่วยในการพัฒนา

(ปริญญาบัตรมีจำนวนทั้งสิ้น 64 หน้า)

อาจารย์ที่ปรึกษาปริญญาบัตร

Name : Mr. Raksit KongWong
: Mr. Peeranut Thadsup
Project Title : Sentiment Data Analytic of King Mongkut's University of
Technology North Bangkok from Teen Only Society Website
Major Field : Information Technology
: King Mongkut's University of Technology North Bangkok
Project Advisor : Asst. Prof. Wanthanee Prachuabsupakij, Ph.D.
Academic Year : 2562

Abstract

This project titled as "Sentiment Data Analytic of King Mongkut's University of Technology North Bangkok from Teen Only Society Website". It is intended to store comments of a teen only society and to analyze the comment data of the university. In order to help administration get and analyze data with related data. Including the ability to summarize opinions that are meaningful, positive or negative at different times in the form of Dashboard moreover, the system can be used to help management decisions in the future.

This project is processed automatically to collect any comments from teen websites, which were developed by nodejs on the server side and then sent to store the database and analyzed the positive and negative comments and categorized the comments using Python language. After that, the analyzed data will be displayed for users by using PHP, JavaScript and HTML, CSS as a development aid.

(Total 64 page)

Project Advisor

กิตติกรรมประกาศ

ระบบวิเคราะห์ข้อมูลความคิดเห็นที่มีต่อมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ จากเว็บไซต์สำหรับวัยรุ่นโดยเฉพาะ สำเร็จลงได้ด้วยดีเนื่องจากได้รับการสนับสนุน คำปรึกษาจาก คณะอาจารย์ คำแนะนำจากเพื่อน ๆ และความอดทนกับความพยายามในการทำงาน

คณะผู้จัดทำต้องขอขอบพระคุณ อาจารย์ประจำสาขาวิชาเทคโนโลยีสารสนเทศ และ โดยเฉพาะอย่างยิ่งผู้ช่วยศาสตราจารย์ ดร.วันทนี ประจวบศุภกิจ ซึ่งเป็นอาจารย์ที่ปรึกษาที่ได้กรุณา คอยให้คำปรึกษาแนะนำ ตลอดการดำเนินปริญญานิพนธ์จนสำเร็จลุล่วงไปได้ด้วยดี และขอบคุณ เพื่อน ๆ พี่ ๆ ที่คอยช่วยเหลือให้คำแนะนำต่างๆ จึงขอกราบพระคุณคณะอาจารย์ทุกท่านเป็นอย่างสูง

สุดท้ายนี้คณะผู้จัดทำขอขอบคุณ บิดา มารดา ผู้ซึ่งมีพระคุณอย่างสูงสุดที่ให้ความอุปการะ ผู้จัดทำมาโดยตลอดรวมทั้งผู้มีพระคุณทุกท่าน คณะผู้จัดทำรู้สึกซาบซึ้งเป็นอย่างยิ่ง จึงใคร่ ขอขอบพระคุณเป็นอย่างสูงมา ณ โอกาสนี้ด้วย

รักชิต กองวงศ์

พีรณัฐ เทศทรัพย์

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	ข
บทคัดย่อภาษาอังกฤษ	ค
กิตติกรรมประกาศ	ง
สารบัญตาราง	ช
สารบัญภาพ	ฅ
บทที่ 1 บทนำ	
1.1 ความเป็นมาและความสำคัญของปัญหา	1
1.2 วัตถุประสงค์ของการจัดทำโครงการพิเศษ	1
1.3 ขอบเขตของการทำโครงการพิเศษ	2
1.4 วิธีการดำเนินจัดทำโครงการพิเศษ	2
1.5 อุปกรณ์ที่ต้องใช้ในการใช้งานโครงการพิเศษ	4
1.6 ผลที่คาดว่าจะได้รับ	4
บทที่ 2 ทฤษฎีที่เกี่ยวข้อง	
2.1 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง	6
2.2 งานวิจัยที่เกี่ยวข้อง	7
2.3 การเปรียบเทียบงานวิจัยที่เกี่ยวข้อง	15
บทที่ 3 การออกแบบและการพัฒนา	
3.1 ระเบียบวิเคราะห์ระบบงาน	16
3.2 การออกแบบการทำงานระบบ	17
3.3 การออกแบบของระบบ	18
3.4 ขั้นตอนการออกแบบระบบ	21
3.5 Activity diagram	25
3.6 ER-Diagram	26
3.7 การออกแบบฐานข้อมูล	28
3.8 การออกแบบเว็บแอปพลิเคชัน	31

สารบัญ (ต่อ)

	หน้า
บทที่ 4 ผลการดำเนินงาน	
4.1 ขั้นตอนการศึกษาค้นคว้าและเก็บรวบรวมข้อมูล	38
4.2 ขั้นตอนวิเคราะห์เว็บไซต์เด็กดี	38
4.3 ขั้นตอนการออกแบบฐานข้อมูล	39
4.4 ขั้นตอนการดึงข้อมูลจากเว็บไซต์เด็กดี	39
4.5 ขั้นตอนการสรุปผลเพื่อแสดงบน เว็บแอปพลิเคชัน	39
4.6 การจัดกลุ่ม	39
4.7 ผลการดำเนินงาน	39
4.8 ลักษณะของโปรแกรม	40
บทที่ 5 สรุปผลการดำเนินงาน และ ข้อเสนอแนะ	
5.1 สรุปผลการดำเนินการ	61
5.2 ปัญหา อุปสรรค และการแก้ไข	62
5.3 การแก้ปัญหา	62
5.4 ข้อเสนอแนะ	62
บรรณานุกรม	63

สารบัญตาราง

ตารางที่	หน้า
2-1 ปริมาณข้อมูลการวิเคราะห์ความคิดเห็นภาษาไทย เกี่ยวกับการรีวิวสินค้าออนไลน์	7
2-2 การเปรียบเทียบงานวิจัยที่เกี่ยวข้อง	15
3-1 แสดง Confusion Matrix ของ API aiforthai	19
3-2 แสดง Confusion Matrix ของ API repustate	19
3-3 แสดง Confusion Matrix ในการจัดประเภทความคิดเห็น	20
3-4 Use Case Id 1	22
3-5 Use Case Id 2	22
3-6 Use Case Id 3	23
3-7 Use Case Id 4	23
3-8 Use Case Id 5	24
3-9 pre_web	24
3-10 pre_comment	28
3-11 analytic_web	28
3-12 analytic_comment	29
3-13 sentiment	29
3-14 Comment	30
3-15 user	30

สารบัญภาพ

ภาพที่	หน้า
2-1 สร้างแบบจำลองโดยใช้อัลกอริทึม K-Nearest Neighbor	9
2-2 สร้างแบบจำลองโดยใช้อัลกอริทึม Naïve Bayes	9
2-3 สร้างแบบจำลองโดยใช้อัลกอริทึม Decision Tree	9
2-4 สร้างแบบจำลองโดยใช้อัลกอริทึม SVM	9
2-5 สถาปัตยกรรมระบบวิเคราะห์ความคิดเห็นของผู้บริโภคจากเว็บไซต์รีวิวร้านอาหาร	10
2-6 ขั้นตอนการตัดคำจากประโยคความคิดเห็น	11
2-7 การเปรียบเทียบคำที่ตัดได้กับคำศัพท์ในโดเมนร้านอาหาร	12
2-8 ตัวอย่างผลลัพธ์ที่ได้จากการวิเคราะห์เหมืองความคิดเห็น	14
3-1 Diagram การทำงานของระบบ	17
3-2 User Case Diagram	21
3-3 Activity diagram	25
3-4 ER-Diagram	26
3-5 การออกแบบหน้าจอหลัก	31
3-6 การออกแบบที่มาของระบบ	32
3-7 การออกแบบเพื่อวัดประสิทธิภาพของระบบ	33
3-8 การออกแบบหน้าที่ทีมงาน	34
3-9 การออกแบบหน้าแสดงผลกระทู้	35
3-10 การออกแบบหน้าแสดงผลความคิดเห็น	36
3-11 การออกแบบหน้า dashbord	37
4-1 เข้าสู่ระบบ	40
4-2 หน้าจอสมัครสมาชิก	41
4-3 ภาพแสดงหน้าจอหลัก	42
4-4 แสดงที่มาของระบบ	43
4-5 แสดงประสิทธิภาพของระบบ	44
4-6 การแสดงผลกระทู้เด็กดี	45
4-7 แสดงข้อมูลแต่ละความคิดเห็น	46
4-8 หน้าเว็บเด็กดี	47
4-9 จัดการสมาชิก	48

สารบัญภาพ (ต่อ)

ภาพที่	หน้า
4-10 แก๊วไฮโดรเจน	49
4-11 การเพิ่มข้อมูล	50
4-12 แสดงความคิดเห็น 10 อันดับแรกความคิดเห็นในเชิงบวก	51
4-13 แสดงความคิดเห็น 10 อันดับแรกความคิดเห็นในเชิงลบ	52
4-14 วิเคราะห์หัวข้อกระทู้	53
4-15 แสดงค่าความถูกต้องวิเคราะห์หัวข้อกระทู้	54
4-16 วิเคราะห์เนื้อหากระทู้	55
4-17 แสดงค่าความถูกต้องวิเคราะห์เนื้อหากระทู้	56
4-18 วิเคราะห์ความคิดเห็น	57
4-19 แสดงค่าความถูกต้องวิเคราะห์ความคิดเห็น	58
4-20 จัดกลุ่ม	59
4-21 ค่าความถูกต้องการจัดกลุ่ม	60

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ในปัจจุบันอินเทอร์เน็ตได้กลายเป็นช่องทางหลักในการติดต่อสื่อสารของมนุษย์ ประชากร ส่วนใหญ่มักคุ้นเคยกับการใช้ชีวิตในสังคมออนไลน์ (Social Media) รูปแบบต่าง ๆ เช่น สื่อสังคมออนไลน์ กระทั่งหรือเว็บบอร์ดต่าง ๆ โดยเว็บไซต์สำหรับวัยรุ่นโดยเฉพาะ เป็นเว็บไซต์ที่ก่อตั้งเมื่อวันที่ 31 ธันวาคม พ.ศ. 2542 โดยกลุ่มนักเรียนชั้นมัธยมศึกษาตอนปลาย มีจุดมุ่งหมายในการสร้างเว็บไซต์ สำหรับการพบปะพูดคุยของเพื่อนวัยเดียวกัน และปัจจุบันเป็นเว็บไซต์หนึ่งที่ได้รับคามนิยมมากในหมู่เด็กวัยรุ่นโดยเฉพาะในระดับชั้นมัธยมศึกษาทั้งตอนต้นและตอนปลาย จากสถิติการใช้เว็บไซต์นี้พบว่าสถิติการเข้าถึงลำดับที่ 29 จากสถิติการใช้งานของเว็บไซต์ในประเทศไทย ซึ่งเว็บดังกล่าวเป็นบอร์ดข้อความที่ เป็นสถานที่ในการแลกเปลี่ยนความคิดเห็นของ ผู้คน โดยเฉพาะวัยรุ่น เป็นแหล่งในการสืบค้นข้อมูล ผู้ใช้อินเทอร์เน็ตจำนวนมากได้ร่วมสร้างข้อมูลหลากหลายรูปแบบ เช่น ข้อความ รูปภาพ ฯลฯ ทำให้มี ข้อมูลจำนวนมากมหาศาลบนโลกออนไลน์ซึ่งสามารถเก็บรวบรวม คัดลอก หรือประมวลผลด้วย คอมพิวเตอร์ได้

มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ เป็นมหาวิทยาลัยหนึ่งที่ได้เปิดการเรียนการสอนในระดับอุดมศึกษา ซึ่งมีกลุ่มเป้าหมายหลักคือนักเรียนในระดับมัธยม และระดับ ปวช. ซึ่งกลุ่มเป้าหมายกลุ่มนี้ได้มีการเข้าใช้งานในเว็บไซต์สำหรับวัยรุ่นโดยเฉพาะ เป็นจำนวนมากผ่านการสอบถามข้อมูล การแสดงความคิดเห็น เพื่อเป็นข้อมูลในการศึกษาต่อในอนาคต ดังนั้น หากได้มีการจัดทำระบบวิเคราะห์ข้อมูลความคิดเห็นที่มีต่อมหาวิทยาลัยฯ จากเว็บไซต์สำหรับวัยรุ่นโดยเฉพาะ ผู้ดูแลระบบสามารถวิเคราะห์ข้อมูลหรือเหตุการณ์ที่เกี่ยวข้อง รวมถึงสามารถสรุปความคิดเห็นว่ามีความหมายเชิงบวก หรือเชิงลบ ตามช่วงเวลาต่าง ๆ ในรูปแบบของ Dashboard

1.2 วัตถุประสงค์ของการจัดทำโครงการพิเศษ

1.2.1 เพื่อศึกษากระบวนการการจัดเก็บข้อมูลความคิดเห็นจากเว็บไซต์เด็กดีที่เกี่ยวข้องกับมหาวิทยาลัยพระจอมเกล้าพระนครเหนือ

1.2.2 เพื่อนำข้อมูลจากเว็บบอร์ดมาวิเคราะห์ความคิดเห็นเชิงบวกและเชิงลบ

1.2.3 เพื่อสร้าง Dashboard เพื่อสรุปข้อมูลความคิดเห็นและสถิติการแสดงความคิดเห็นตามช่วงเวลาที่แตกต่างกัน

1.3 ขอบเขตของการทำโครงการพิเศษ

1.3.1 ระบบจัดทำในรูปแบบของเว็บแอปพลิเคชันที่สามารถดึงข้อมูลจากเว็บไซต์เด็กดีตามข้อความที่กำหนดให้ได้ เช่น กำหนดข้อความคำว่า มจพ. (ปัจจุบันมี 1,242 รายการ) หรือ KMUTNB (ปัจจุบันมี 41 รายการ) โดยผู้จัดทำได้จัดเก็บ ณ วันที่ 1 กรกฎาคม 2562 ถึง 10 ตุลาคม 2562

1.3.2 ระบบสามารถกำหนดวันและเวลาที่จะดึงข้อมูลจากเว็บไซต์สำหรับวัยรุ่นโดยเฉพาะ

1.3.3 ระบบสามารถคัดกรองเฉพาะความคิดเห็นเท่านั้นที่จะถูกนำมาวิเคราะห์ข้อมูล

1.3.4 ระบบสามารถสรุปความคิดเห็นว่าเป็นความคิดเห็นเชิงบวกและเชิงลบได้

1.3.5 ระบบสามารถแสดงข้อมูลเชิงสถิติในรูปแบบของ Dashboard เพื่อสรุปความคิดเห็นตามช่วงวันที่ที่แสดงความคิดเห็นที่ต้องการได้

1.3.6 กรณีที่พบความคิดเห็นที่เป็นลบ ผู้ดูแลระบบสามารถติดตามไปยังกระทู้ที่เป็นที่มาของความคิดเห็นนั้นในเว็บไซต์สำหรับวัยรุ่นโดยเฉพาะ เพื่อให้ผู้ที่เกี่ยวข้องสามารถเข้าไปดำเนินการชี้แจงได้ต่อไป

1.4 วิธีการดำเนินจัดทำโครงการพิเศษ

1.4.1 ขอบเขตของการทำโครงการพิเศษ ภาคเรียนที่ 1/2562

1.4.1.1 สร้างระบบ เพื่อดึงข้อมูลบอร์ดข้อความ ความคิดเห็น วันที่ และเวลา ออกมาในรูปแบบ Window Service ประกอบไปด้วย

4.1.1.1 จำนวนบอร์ดข้อความ

4.1.1.2 จำนวนความคิดเห็นของบอร์ดข้อความ

4.1.1.3 ข้อมูลการแสดงความคิดเห็นในบอร์ดข้อความ

4.1.1.4 ข้อมูลบอร์ดข้อความ เวลา - วันที่ - เดือน - ปี

4.1.1.5 นำข้อมูลที่ดึงมาจากบอร์ด จัดเก็บเข้าสู่ฐานข้อมูล

1.4.1.2 ออกแบบฐานข้อมูลเพื่อเก็บข้อมูลความคิดเห็นในเว็บไซต์เด็กดีที่เกี่ยวข้องกับมหาวิทยาลัยพระจอมเกล้าพระนครเหนือ

1.4.1.3 พัฒนาระบบเข้าใช้งานบนเว็บแอปพลิเคชัน

1.4.1.3.1 ผู้ดูแลระบบ

1.4.1.3.1.1 สามารถเพิ่มลบแก้ไขข้อมูลของผู้ใช้งานได้

1.4.1.3.1.2 สามารถแสดงข้อมูลที่ดึงออกมาจากเว็บไซต์สำหรับวัยรุ่นโดยเฉพาะ และแสดงบนเว็บแอปพลิเคชัน

1.4.1.3.2 ผู้ใช้ทั่วไป

1.4.1.3.2.1 แสดงข้อมูลของผู้ใช้

1.4.2 ขอบเขตของการทำโครงการพิเศษ ภาคเรียนที่ 2/2562

1.4.2.1 สามารถวิเคราะห์ข้อมูลเชิงบวกหรือเชิงลบ

1.4.2.1.1 สามารถรับรู้ได้ว่าวันที่เวลาของบอร์ดข้อความ

1.4.2.1.2 สามารถรับรู้ความคิดเห็นในเชิงลบที่เยอะที่สุด 10 อันดับแรก

1.4.2.1.3 สามารถรับรู้ความคิดเห็นในเชิงบวกที่เยอะที่สุด 10 อันดับแรก

1.4.2.1.4 สามารถติดตามเข้าไปยังบอร์ดข้อความเชิงลบเพื่อให้ผู้เกี่ยวข้องชี้แจงข้อมูลเชิงลบได้

1.4.2.1.5 นำข้อความจากเว็บบอร์ดมาผ่านกระบวนการทำเหมืองข้อความ (text mining) เพื่อสรุปข้อความความคิดเห็นในเชิงสถิติ

1.4.2.2 พัฒนา Web Application สำหรับแสดงผลการวิเคราะห์และ Dashboard

1.4.2.2.1 ผู้ดูแลระบบสามารถ

1.4.2.2.1.1 แสดงข้อมูลเชิงสถิติเกี่ยวกับความคิดเห็นเชิงบวกและเชิงลบ

1.4.2.2.1.2 แสดงข้อมูลเชิงสถิติตามช่วงเวลาที่มีความคิดเห็นบนบอร์ดข้อความ

1.4.2.2.1.3 แสดงข้อมูลความคิดเห็นในเชิงลบที่เยอะที่สุด 10 อันดับแรก

1.4.2.2.1.4 แสดงข้อมูลความคิดเห็นในเชิงบวกที่เยอะที่สุด 10 อันดับแรก

1.4.2.2.1.5 สามารถติดตามเข้าไปยังบอร์ดข้อความเพื่อชี้แจงข้อมูลเชิงลบได้

1.4.2.2.1.6 สามารถแสดงผลการวิเคราะห์เหมืองข้อความเพื่อสรุปข้อความความคิดเห็นในเชิงสถิติ

1.4.2.2.2 ผู้ใช้ทั่วไปสามารถ

1.4.2.2.2.1 แสดงข้อมูลเชิงสถิติตามช่วงเวลาของความคิดเห็นบนบอร์ดข้อความ

1.4.2.2.2.2 แสดงเชิงสถิติตามความคิดเห็นเยอะที่สุด Top10 เชิงลบ

1.4.2.2.2.3 แสดงเชิงสถิติตามความคิดเห็นเยอะที่สุด Top10 เชิงบวก

1.5 อุปกรณ์ที่ต้องใช้ในการใช้งานโครงการพิเศษ

1.5.1 Hardware

1.5.1.1 เครื่องคอมพิวเตอร์(Computer)

1.5.1.2 เครื่องพิมพ์ (Printer)

1.5.1.3 เครื่องเซิร์ฟเวอร์หรือเว็บโฮสติ้ง

1.5.2 Software และ ภาษาที่ใช้เขียนโปรแกรม

1.5.2.1 Nodejs

1.5.2.2 visual studio code

1.5.2.3 python

1.5.2.4 PHP

1.5.2.5 Css

1.5.2.6 Javascript

1.5.2.7 bootstrap

1.6 ผลที่คาดว่าจะได้รับ

1.6.1 สามารถดึงข้อมูลจากบอร์ดข้อความที่เกี่ยวข้องกับ มหาวิทยาลัยพระจอมเกล้าพระนครเหนือ มาแสดงบนเว็บแอปพลิเคชัน

1.6.2 แสดงข้อมูลเชิงวิเคราะห์จากข้อความบนบอร์ดข้อความว่าเป็นความคิดเห็นเชิงบวกหรือเชิงลบ และข้อมูลเชิงสถิติต่าง ๆ ได้

1.6.3 สามารถให้ผู้บริหารได้นำข้อมูลไปวิเคราะห์และสามารถตัดสินใจได้

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ทฤษฎีที่เกี่ยวข้อง

2.1.1 Big data analytics คือ กระบวนการวิเคราะห์ชุดข้อมูลขนาดใหญ่เพื่อค้นหารูปแบบความสัมพันธ์ ของข้อมูลเหล่านั้นที่ซ่อนอยู่ข้างใน หาสิ่งเชื่อมโยงที่เชื่อมข้อมูลเหล่านั้นเข้าไว้ด้วยกัน หาเทรนด์ทางการตลาด หาความต้องการของลูกค้า และข้อมูลอื่น ๆ ที่เป็นประโยชน์ต่อธุรกิจ ผลการวิเคราะห์ข้อมูลสามารถนำไปสู่การทำแผนการตลาดที่มีประสิทธิภาพ โอกาสในการสร้างผลกำไร การให้บริการที่ดีมากขึ้นแก่ลูกค้า การปรับปรุงการทำงานให้เกิดประสิทธิภาพ ความได้เปรียบเหนือคู่แข่ง ในการแข่งขันทางการตลาด และผลประโยชน์ทางธุรกิจด้านอื่น ๆ (Wikipedia , 2557)

2.1.2 Dashboard คือ หน้าแรกที่รวมสิ่งต่าง ๆ ที่สำคัญ บน Application ะบบ Mobile เอามาไว้ในที่เดียว นักลงทุนสามารถเห็นความเคลื่อนไหวของตลาดผ่านหน้าจอนี้ได้ทันที (helponline , 2559)

2.1.3 Social Network (โซเชี่ยล เน็ตเวิร์ค) คือ เครือข่ายสังคมออนไลน์ หรือเป็นการบริการที่เชื่อมโยงหลายคนเข้าไว้ด้วยกันผ่านอินเทอร์เน็ต ตัวอย่างของ Social Network ได้แก่ Facebook Twitter Hi5 Blogger เป็นต้น ซึ่งเปรียบเหมือน สังคมจำลองเสมือนจริงนั่นเอง และในปัจจุบัน นอกจาก Social Network จะเป็นสังคมออนไลน์แล้ว ยังเป็นเครื่องมือทางการตลาดซึ่งเป็นที่นิยมอีกด้วย (softmelt , 2558)

2.1.4 nodejs คือ JavaScript มีการ Compiled เป็น Byte Code ด้วย V8 Engine ของ Google. Debug ได้ ต่างจาก JavaScript ในยุคแรก ๆ ทำให้แก้ปัญหได้ง่ายขึ้น รันได้บนทุกระบบปฏิบัติการยอดนิยม มาพร้อมกับเทคโนโลยีที่เรียกว่า Non – Blocking I/O (beyourcyber , 2561)

2.1.5 ภาษา พีเอชพี(PHP) คือ ภาษาคอมพิวเตอร์ในลักษณะเซิร์ฟเวอร์-ไซด์ สคริปต์ โดยลิขสิทธิ์อยู่ในลักษณะโอเพนซอร์ส ภาษาพีเอชพีใช้สำหรับจัดทำเว็บไซต์ และแสดงผลออกมาในรูปแบบ HTML มีฐานโครงสร้างคำสั่งมาจากภาษา ภาษาซี ภาษาจาวา และ ภาษาเพิร์ล ซึ่ง ภาษา

พีเอชพี นั้นง่ายต่อการเรียนรู้ ซึ่งเป้าหมายหลักของภาษานี้ คือให้นักพัฒนาเว็บไซต์สามารถเขียนเว็บเพจ ที่มีความตอบโต้ได้อย่างรวดเร็ว (pasaphp , ภาษา php)

2.1.6 NoSQL คือ เทคโนโลยีฐานข้อมูลที่ถูกออกแบบมาสำหรับงานเฉพาะทางบางอย่างที่ SQL ยังไม่สามารถตอบโจทย์ได้ดีเพียงพอ เมื่อพูดถึง NoSQL จะได้ยินชื่อเว็บไซต์ที่ใหญ่ ๆ ติดพ่วงมาด้วย เช่น Facebook, Twitter, FourSquare, Digg และอื่น ๆ ทำให้เรารู้ว่า NoSQL เป็นระบบฐานข้อมูล สำหรับงานที่ต้องรองรับข้อมูลขนาดใหญ่ ๆ รองรับการทำงานแบบได้ง่าย (softmelt , 2558)

2.1.7 API ย่อมาจาก Application Programming Interface คือ ช่องทางการเชื่อมต่อ , ช่องทางหนึ่งที่จะเชื่อมต่อกับเว็บไซต์ผู้ให้บริการAPIจากที่อื่นเป็นตัวกลางที่ทำให้โปรแกรมประยุกต์เชื่อมต่อกับโปรแกรมประยุกต์อื่นหรือเชื่อมการทำงานเข้ากับระบบปฏิบัติการตัวอย่าง เช่น Twitter มีหลายเว็บ ที่มีการ เชื่อมโยงข้อมูลกับ twitter ทั้งเป็นการอ่านข้อมูลจาก twitter หรือ ส่งข้อมูลเข้า twitter เองก็ตาม ซึ่งล้วนอาศัยการเชื่อมต่อแลกเปลี่ยนข้อมูลกัน ด้วย API นั้นเองและอีกหนึ่งตัวอย่างเช่น Google Maps API คือบริการของGoogle อีกรูปแบบหนึ่งที่เราสามารถนำข้อมูลของ GoogleMaps ที่ทาง Googleให้บริการโดยส่วนมากจะนำมาใช้กับเว็บไซต์ ของบริษัทหรือเว็บไซต์ห้างร้านต่าง ๆ เพื่อเป็นอีกช่องทางที่ลูกค้ารู้ว่าบริษัทฯ หรือห้างร้านนั้น (mindphp , 2560)

2.1.8 Python Python คือชื่อภาษาที่ใช้ในการเขียนโปรแกรมภาษาหนึ่ง ซึ่งถูกพัฒนาขึ้นมาโดยไม่ยึดติดกับแพลตฟอร์มกล่าวคือสามารถรันภาษาPythonได้ทั้งบนระบบ Unix, Linux, Windows, NT, Windows2000,WindowsXPหรือแม้แต่ระบบFreeBSDอีกอย่างหนึ่งภาษาตัวนี้เป็น OpenSource เหมือนอย่าง PHP ทำให้ทุกคนสามารถที่จะนำ Python มาพัฒนาโปรแกรมของเราได้ฟรีๆโดยไม่ต้องเสียค่าใช้จ่ายและความเป็น Open Source ทำให้มีคนเข้ามาช่วยกันพัฒนาให้ Python มีความสามารถสูงขึ้น และใช้งานได้ครอบคลุมกับทุกลักษณะงาน (Aosof , 2561)

2.1.9 text mining เป็นเทคนิคเพื่อค้นหารูปแบบ (pattern) ของจากข้อความจำนวนมากศาสตร์โดยอัตโนมัติโดยใช้ขั้นตอนวิธีจากวิชาสถิติ การเรียนรู้ของเครื่อง และ การรู้จำแบบ หรือในอีกนิยามหนึ่งการทำเหมืองข้อความ คือกระบวนการที่กระทำกับข้อความ (โดยส่วนใหญ่จะมีจำนวนมาก) เพื่อค้นหารูปแบบ แนวทาง และความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อความนั้น โดยอาศัยหลักสถิติ การรู้จำการเรียนรู้ของเครื่อง หลักคณิตศาสตร์ หลักการ ประมวลผลเอกสาร (Document Processing) หลักการประมวลผลข้อความ (Text Processing) และหลักการประมวลผลภาษาธรรมชาติ (Natural Language Processing) (Wikipedia , 2557)

2.1.10 puppeteer เป็น Node library ที่มีชุดของ API สำหรับการควบคุม Google Chrome หรือ Chromium ผ่าน DevTool protocol ทำงานในรูปแบบ headless โดยdefault แต่สามารถ

กำหนดให้ทำงานแบบ full (non-headless) Chrome ได้ ที่สำคัญไม่ต้องทำงานผ่าน Web Driver อีกต่อไป (psu , 2561)

2.2 งานวิจัยที่เกี่ยวข้อง

2.2.1 การวิเคราะห์ความคิดเห็นภาษาไทยเกี่ยวกับการรีวิวสินค้าออนไลน์

งานวิจัยนี้ได้นำเสนอวิธีการจำแนกความคิดเห็นโดยการวิเคราะห์ความคิดเห็นภาษาไทย เกี่ยวกับการรีวิว สินค้าออนไลน์ ด้านการบริการห้องพัก โรงแรม รีสอร์ท จาก Agoda Thailand และ Twitter Thailand ที่จัดทะเบียนหลักทรัพย์ โดยเก็บข้อมูล จำนวน 2,890 ข้อมูล โดยใช้เทคนิคเหมืองข้อความ วิเคราะห์ความคิดเห็นภาษาไทยเกี่ยวกับการรีวิว สินค้าออนไลน์ และสร้างแบบจำลองด้วยอัลกอริทึม 4 วิธี ได้แก่ซัพพอร์ตเวกเตอร์แมชชีน ต้นไม้ตัดสินใจ นาอ์ฟเบย์ และเคเนียร์เซนเบอร์ เพื่อเปรียบเทียบ ประสิทธิภาพการวิเคราะห์ความคิดเห็น ภาษาไทย เกี่ยวกับการรีวิว สินค้าออนไลน์ จากการทดลองพบว่า คุณลักษณะที่ดีที่สุดคือ ซัพพอร์ตเวกเตอร์แมชชีนระดับรองลงมาเป็น นาอ์ฟเบย์ , ต้นไม้ตัดสินใจ และเคเนียร์เซนเบอร์ ตามลำดับ (รวิศดา, 2561)

2.2.1.1 การเตรียมข้อมูล

เป็นขั้นตอนการรวบรวมการวิเคราะห์ความคิดเห็นภาษาไทยเกี่ยวกับการรีวิว สินค้าออนไลน์ที่นำมาสร้างแบบจำลองเพื่อการเรียนรู้ จำนวน 2,890 ข้อความ จากแหล่งข่าวสาธารณะที่ได้ทำการศึกษามา ดังตารางที่ 2-1

ตารางที่ 2-1 ปริมาณข้อมูลการวิเคราะห์ความคิดเห็นภาษาไทย เกี่ยวกับการรีวิวสินค้าออนไลน์

แหล่งข่าวสาธารณะ ภาษาไทย	จำนวน ข้อความ
Agoda Thailand	1,573
Twitter Thailand	1,317

2.2.1.2 การแยกวิเคราะห์เอกสาร (Document Parser)

ขั้นตอนการดึงข้อมูลจากแหล่งข้อมูลโดยใช้โปรแกรม Import.io1 จากนั้นทำการคัดเลือกข้อมูลโดยการจัดข้อมูลที่ซ้ำกันออก การแปลงข้อมูลให้อยู่ในรูปแบบที่เหมาะสมสำหรับการสร้างแบบจำลองโดยการนำข้อความมาตัดคำผ่านโปรแกรมการตัดคำภาษาไทย (ThaiLexemeTokenizer:LexTo)

2.2.1.3 การสกัดคุณลักษณะ (Feature extraction)

การดึงคุณลักษณะ (Feature) ของ ความคิดเห็นออกมา ซึ่งการดึงคุณลักษณะออกมานั้น ต้องกำหนดก่อนว่าจะใช้อะไร เป็นตัวแทนคุณลักษณะ ของความคิดเห็น และใช้ค่าใดแทนคุณลักษณะความคิดเห็นนั้น โดยการสร้างคำจากข้อความที่เตรียมไว้ (Tokenize) กรองจำนวนคำตัวอักษรของคำ(Filter Tokenize) โดยกำหนดคำที่มีความยาวตั้งแต่ 2 ถึง 25 ตัวอักษรและกรอกคำหยุด (Filter Stopword) เป็นการนำคำที่ไม่มีนัยสำคัญออกโดยไม่ทำให้ความหมาย ของข่าวสารเปลี่ยนแปลง ในขั้นตอนนี้นำโปรแกรม Rapidminer3 มาช่วยในการสกัดคุณลักษณะของ ความคิดเห็นภาษาไทยเกี่ยวกับการรีวิวสินค้าออนไลน์

2.2.1.4 การกำจัดคำหยุด (Stop-Word List Removal)

การนำคำที่ไม่มีนัยสำคัญออกโดยไม่ทำให้ ความหมายของเอกสารเปลี่ยนแปลงคำที่ไม่มีนัยสำคัญ หมายถึงคำที่ใช้กันโดยทั่วไปไม่มี ความหมายสำคัญต่อเอกสาร เมื่อตัดออกจากเอกสาร แล้วไม่ทำให้ ใจความของเอกสารเปลี่ยนแปลง ตัวอย่างเช่น คำบุพบทเป็นคำที่ใช้เชื่อมคำหรือกลุ่มคำ ให้สัมพันธ์กัน คำสันธานเป็นคำที่ทำหน้าที่เชื่อมคำกับคำ คำสรรพนามเป็นคำที่ใช้แทนคำนามที่กล่าวถึง มาแล้วในประโยค เป็นต้น จึงถือได้ว่าคำหยุดเป็น คุณลักษณะที่ไม่มีประโยชน์ในการจำแนก

2.2.1.5 การหารากศัพท์ (Stemming)

การหารูปเดิมของคำหรือหาคำที่มีความหมายคล้ายกัน เพื่อปรับรวมให้เป็นคำเดียวกัน การหารากศัพท์ เป็นกระบวนการที่ควรทำก่อนการ จัดทำดัชนี ทำให้สามารถลดขนาดของดัชนีลงและเพิ่ม ประสิทธิภาพในการค้นคืน หรือการจำแนกหมวดหมู่

2.2.1.6 การสร้างแบบจำลองสำหรับการเรียนรู้อัลกอริทึมการจำแนก 4 วิธี

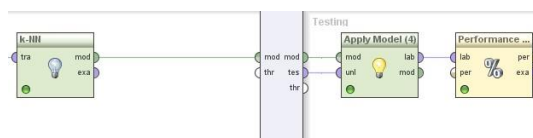
2.2.1.6.1 Naïve Bayes ดังภาพที่ 2-1

2.2.1.6.2 SVM ดังภาพที่ 2-2

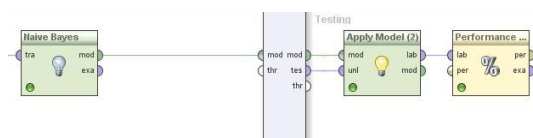
2.2.1.6.3 K-Nearest Neighbor ดังภาพที่ 2-3

2.2.1.6.4 Decision Tree ดังภาพที่ 2-4

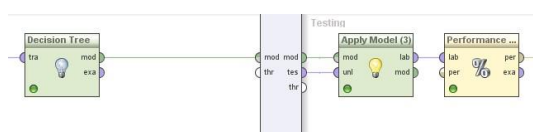
เพื่อเปรียบเทียบ ประสิทธิภาพการจำแนกหมวดหมู่



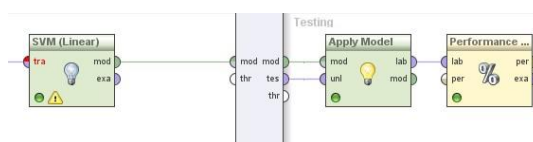
ภาพที่ 2-1 สร้างแบบจำลองโดยใช้อัลกอริทึม K-Nearest Neighbor



ภาพที่ 2-2 สร้างแบบจำลองโดยใช้อัลกอริทึม Naïve Bayes



ภาพที่ 2-3 สร้างแบบจำลองโดยใช้อัลกอริทึม Decision Tree



ภาพที่ 2-4 สร้างแบบจำลองโดยใช้อัลกอริทึม SVM

2.2.1.7 การจำแนกหมวดหมู่ (Classifier)

การจำแนกหมวดหมู่ความคิดเห็นภาษาไทย เกี่ยวกับการรีวิวสินค้าออนไลน์ ในงานวิจัยนี้จำแนกออกเป็น 5 หมวด

2.2.1.7.1 ความสะอาด

2.2.1.7.2 สิ่งอำนวยความสะดวก

2.2.1.7.3 ความสะดวกสบายและคุณภาพห้องพัก

2.2.1.7.4 การให้บริการ

2.2.1.7.5 คำนึงถึงราคาที่จ่าย

2.2.1.8 การวัดประสิทธิภาพ (Testing and Evaluation)

การทดสอบแบบจำลองการจำแนกหมวดหมู่ข้อความความคิดเห็นภาษาไทยเกี่ยวกับการรีวิวสินค้า ออนไลน์แหล่งข้อมูลสาธารณะภาษาไทย พิจารณาจากค่าความถูกต้องโดยใช้ วิธี การประเมิน ความสามารถของแบบจำลอง วัดที่ประสิทธิภาพของ การจำแนก ข้อมูลตามแนวคิดทางด้านการค้นคืน สารสนเทศ ซึ่งก็คือ การวัดค่าความถูกต้อง (Accuracy) คำนวณได้ดัง สมการ

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TF} + \text{FP} + \text{FN} + \text{TN}}$$

2.2.2 การวิเคราะห์ความคิดเห็นเกี่ยวกับร้านอาหารบนเว็บไซต์รีวิว

ระบบจะวิเคราะห์ข้อความด้วยเทคนิคการตัดคำ การวิเคราะห์ประเภทคำ และรูปประโยค เพื่อหาความหมายเชิงบวกหรือเชิงลบของประโยค แล้วจึงคำนวณและแสดงผลสรุปออกมาเป็นค่าระดับความพึงพอใจสำหรับแต่ละด้านในรูปแบบกราฟิกเพื่อให้ผู้บริโภคใช้ประกอบการตัดสินใจได้ง่ายยิ่งขึ้นระบบวิเคราะห์ความคิดเห็นของผู้บริโภคจากเว็บไซต์รีวิวร้านอาหารนี้จะสอดคล้องกับสถาปัตยกรรมดังภาพ (ปราชญ์ภาคย์ , 2561)



ภาพที่ 2-5 สถาปัตยกรรมระบบวิเคราะห์ความคิดเห็นของผู้บริโภคจากเว็บไซต์รีวิวร้านอาหาร

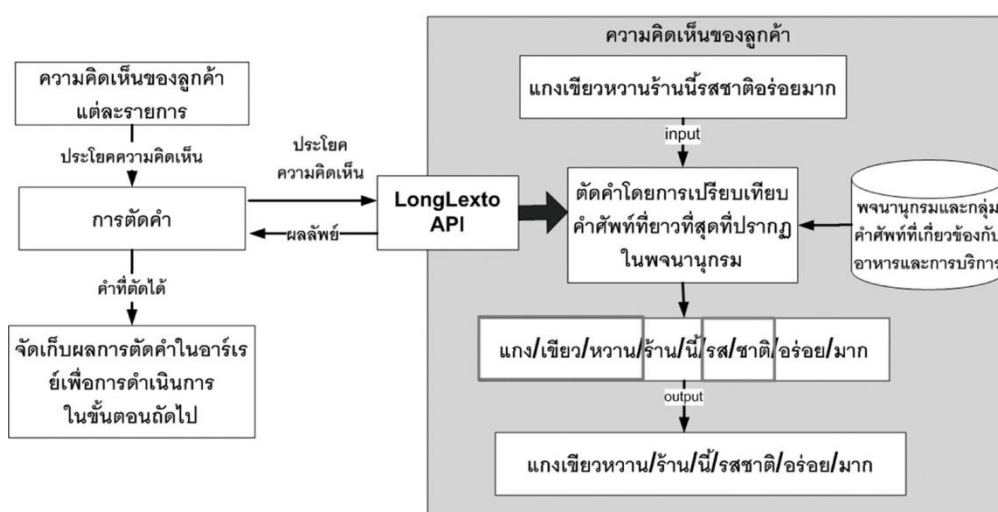
2.2.2.1 การรวบรวมข้อมูลความคิดเห็น (Collect comments)

ความคิดเห็นเกี่ยวกับร้านอาหารจากเว็บไซต์รีวิวจะถูกรวบรวมและจัดเก็บลงในตารางความคิดเห็นเกี่ยวกับร้านอาหาร ที่จะประกอบด้วยแอททริบิวต์ รหัสความคิดเห็น ชื่อร้าน และ

รายละเอียดของความคิดเห็น ตัวอย่างความคิดเห็นได้แก่ “ร้านโภข้าวอาหารกึ่งนั้น ๆ ไม่ค่อยประทับใจเท่าไร เด็กเสิร์ฟไม่ค่อยสนใจบริการ” เป็นต้น

2.2.2.2 ตัดคำ (Tokenization)

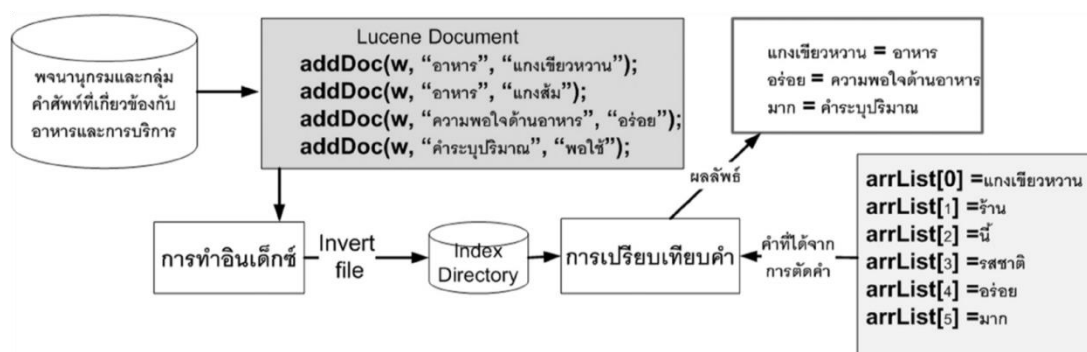
ระบบจะนำประโยคความคิดเห็นแต่ละรายการมาทำการตัดคำโดยใช้เครื่องมือ LongLexto API การตัดคำ จะใช้วิธีการตัดคำแบบยาวที่สุดโดยเทียบกับพจนานุกรม LEXiTRON โดยการตัดคำจะเริ่มพิจารณาจากซ้ายไปขวา ตามหลักการเขียนภาษาไทย จากนั้นทำการนำ “คำหยุด” (Stop Word) ซึ่งเป็นคำที่ไม่จำเป็นสำหรับการวิเคราะห์ ออกเพื่อลดขนาดของข้อมูลที่ใช้ในการประมวลผล จากภาพที่ 2-6 เมื่อประโยคความคิดเห็น “แกงเขียวหวานร้านนี้รสชาติอร่อยมาก” มาตัดคำจะพบว่าสำหรับ “แกงเขียวหวาน” สามารถตัดคำตามพจนานุกรมได้ 2 แบบ คือคำว่า “แกง” และ คำว่า “แกงเขียวหวาน” แต่เมื่อใช้หลักการตัดคำแบบยาวที่สุด จึงเลือกตัดคำเป็น “แกงเขียวหวาน” เป็นต้น ซึ่งผลลัพธ์ของคำที่ตัดได้เหล่านี้มาเก็บไว้ในตัวแปรอาร์เรย์ชื่อ ArrList เพื่อให้ง่ายต่อขั้นตอนการสะกดคำต่อไป



ภาพที่ 2-6 ขั้นตอนการตัดคำจากประโยคความคิดเห็น

2.2.2.3 การสกัดคำ (Word extraction)

ในขั้นตอนนี้จะเกี่ยวข้องกับการเปรียบเทียบคำที่ตัดได้ในประโยค กับคำศัพท์ในโดเมนร้านอาหาร เพื่อทำการระบุวาทนิพจน์หรือคำที่ตัดได้นั้นเป็นคำประเภทใด จากนั้นจะทำการวิเคราะห์รูปแบบประโยคที่สมบูรณ์ (ตามหลักไวยากรณ์) เพื่อการนำไปใช้ในการคำนวณความคิดเห็นในขั้นตอนต่อไป



ภาพที่ 2-7 การเปรียบเทียบคำที่ตัดได้กับคำศัพท์ในโดเมนร้านอาหาร

2.2.2.4 การคำนวณคะแนนความคิดเห็น

การคำนวณคะแนนความพึงพอใจจากความคิดเห็นของผู้ใช้จะอาศัยค่าน้ำหนักของค่าระบุนปริมาณ โดยค่าน้ำหนักนี้ได้มาจากการสำรวจเว็บไซต์รีวิวร้านอาหารต่าง ๆ ที่ให้ผู้ใช้แสดงความคิดเห็นและมีค่าเหล่านั้น ปรากฏอยู่จากนั้นคำนวณออกมาเป็นค่าระดับน้ำหนัก เช่น คำว่า “ที่สุด” มีค่าน้ำหนักเท่ากับ 2 คำว่า “มาก” มีค่าน้ำหนักเท่ากับ 1.6 หากไม่ปรากฏค่าระบุนปริมาณในประโยค ให้ค่าน้ำหนักเท่ากับ 1. เป็นต้น โดยค่าคะแนน ของประโยคความคิดเห็นจะเป็นบวกหรือลบ ขึ้นอยู่กับความพึงพอใจและน้ำหนักของค่าระบุนปริมาณ เช่นถ้าผู้ใช้ เขียนว่า “ซอสสปรกที่สุด” ค่าคะแนนด้านความสะอาดก็จะเป็น - 2 (สปรก ความพึงพอใจติดลบ, ที่สุด ค่าน้ำหนักของค่าระบุนปริมาณ = 2)

2.2.3 การวิเคราะห์เหมือนความคิดเห็นบนเครือข่ายสังคมออนไลน์

การประมวลผลภาษาธรรมชาติสำหรับข้อความภาษาไทยจำเป็นต้องตัดประโยคออกเป็นคำก่อน เนื่องจากภาษาไทยมีรูปแบบการเขียนคำยาวต่อกัน ไม่มีการเว้นวรรคระหว่างคำ ดังเช่นภาษาอังกฤษ โปรแกรมตัดคำสำหรับข้อความภาษาไทยที่เปิดให้ดาวน์โหลดได้ฟรี ซึ่งพัฒนาโดยหน่วยปฏิบัติการ วิทยาการมนุษย์ภาษาศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ (NECTEC) คือ โปรแกรมตัดคำ LexTo (ThaiLexeme Tokenizer) และ TLexs (Thai Lexeme Analyser) เมื่อข้อความถูกตัดคำออกมาด้วย โปรแกรมแล้ว จึงนำคำมาวิเคราะห์และแทนข้อความให้อยู่ในรูปแบบที่เป็นโครงสร้าง ด้วยแนวทางใดแนวทางหนึ่ง หรือทั้งสองแนวทางร่วมกัน (กานดา , 2561)

2.2.3.1 การวิเคราะห์เหมือนข้อความ

ข้อมูลแบ่งออกเป็น 2 รูปแบบคือ (1) ข้อมูลที่เป็น โครงสร้าง (structured data) ซึ่งการประมวลผลข้อมูลที่เป็นโครงสร้างเรียกว่าการวิเคราะห์เหมือนข้อมูล และ (2) ข้อมูลที่ไม่เป็นโครงสร้าง (unstructured data) หรือไม่มี โครงสร้างที่ชัดเจน (implicit structured data)

ซึ่งส่วนใหญ่ มักอยู่ในรูปแบบของข้อความหรือภาษาธรรมชาติ และ เรียกกระบวนการวิเคราะห์ข้อความว่าการวิเคราะห์เหมือง ข้อความ (text mining) โดยเป็นการนำความรู้ด้านการ วิเคราะห์เหมืองข้อมูลมาประยุกต์ใช้ หน่วยงาน Cross Industry Standard Process for Data Mining (CRISP DM) ได้นำเสนอกระบวนการ วิเคราะห์เหมืองข้อมูล

2.2.3.2 การสร้างแบบจำลองการวิเคราะห์ข้อมูล

แบบจำลองการวิเคราะห์ข้อมูลถูกสร้างจากการ เรียนรู้รูปแบบข้อมูลที่มีอยู่และรู้ผลเฉลยแล้ว เรียกว่าชุดข้อมูลเรียนรู้ (training set) เพื่อนำไปวิเคราะห์อธิบาย หรือ ทำนายรูปแบบของข้อมูลที่ยังไม่เกิดขึ้นหรือยังไม่รู้ผลเฉลย เรียกว่าชุดข้อมูลทดสอบ (test set) ซึ่งการสร้างแบบจำลอง การวิเคราะห์ข้อมูลแบ่งเป็น 2 เทคนิคหลักคือ (1) การเรียนรู้แบบมีผู้สอน (supervised learning) มีวิธีที่ใช้ เช่น การจำแนกข้อมูล (classification) สำหรับกระบวนการ วิเคราะห์เหมืองข้อความเรียกกระบวนการดังกล่าวว่า การจำแนกข้อความ (text classification หรือ text categorization) ยังมีปริมาณชุดข้อมูลเรียนรู้น้อย จะยิ่งทำให้แบบจำลองการจำแนกข้อความที่มีความถูกต้องสูง

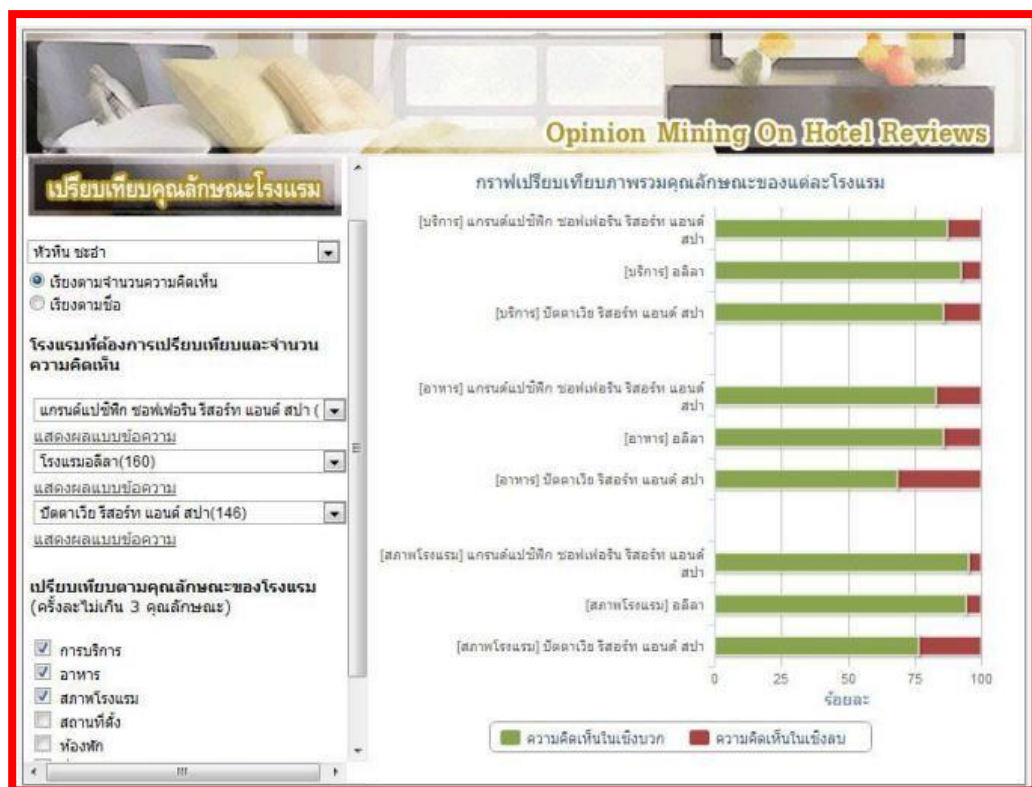
2.2.3.3 การประยุกต์ใช้การวิเคราะห์เหมืองความคิดเห็น

การประยุกต์ใช้การวิเคราะห์เหมืองความคิดเห็นข้อความบนเครือข่ายสังคมออนไลน์แบ่งเป็น 3 ประเภท ได้แก่ ข้อเท็จจริง ความคิดเห็น และข้อเสนอแนะ โดย (1) ข้อเท็จจริง หมายถึงข้อความหรือเหตุการณ์ที่เป็นมาหรือที่เป็นอยู่ตามจริง (2) ความคิดเห็นหมายถึง ความเห็นหรือข้อวินิจฉัยหรือความเชื่อที่แสดงออกตามที่ เห็น รู้ คิดและ (3) ข้อเสนอแนะหมายถึงข้อคิดเห็นเชิงแนะนำที่เสนอเพื่อพิจารณา ชี้แจงให้ทำหรือปฏิบัติตาม

ตัวอย่าง“เมื่อคืนได้ดู รายการพื้นที่ชีวิต ไม่ชอบพิธีกรเลย อยากให้ปรับปรุงเรื่อง การใช้ภาษาหน่อย” จากประโยคตัวอย่างสามารถจำแนก ประเภทของบทวิจารณ์ได้ ดังนี้

ข้อเท็จจริง : “เมื่อคืนได้ดูรายการพื้นที่ชีวิต” ความคิดเห็น : “ไม่ชอบพิธีกรเลย”

ข้อเสนอแนะ : “อยากให้ปรับปรุงเรื่องการใช้ ภาษาหน่อย” ธุรกิจในปัจจุบันให้ความสำคัญกับความคิดเห็น ของผู้บริโภคบนเครือข่ายสังคมออนไลน์มากขึ้นแต่ปัญหา คือบางความคิดเห็นมีข้อความแสดงความคิดเห็นค่อนข้างยาว และมีเพียงบางประโยคเท่านั้นที่ผู้อ่านให้ความสนใจ และข้อความความคิดเห็นของแต่ละบุคคลมีความหลากหลายในแต่ละมุมมองที่แตกต่างกันไป จึงเสียเวลามากหากต้องอ่าน ทุกความคิดเห็นกว่าจะสรุปความคิดเห็นในมุมมองเฉพาะ เรื่องที่สนใจได้ หลายธุรกิจจึงได้นำวิธีการวิเคราะห์เหมือง ความคิดเห็นไปประยุกต์ใช้ เพื่อสรุปความพึงพอใจของ ผู้บริโภคหรือแม้กระทั่งผู้บริโภคเองที่ค้นหาความพึงพอใจ ของบุคคลอื่นที่แสดงความคิดเห็นไว้บนเครือข่ายสังคม ออนไลน์



ภาพที่ 2-8 ตัวอย่างผลลัพธ์ที่ได้จากการวิเคราะห์เหมืองความคิดเห็น

ตารางที่ 2-2 การเปรียบเทียบงานวิจัยที่เกี่ยวข้อง

งานวิจัยที่เกี่ยวข้อง	รีวิวสินค้าออนไลน์	ร้านอาหารเว็บไซต์รีวิว	เหมือนความคิดเห็นบนเครือข่ายสังคมออนไลน์
วิธีการดึงข้อมูล	แบบสอบถาม	แบบสอบถาม	แบบสอบถาม
การตัดคำ	Lexto	Lexto	Lexto
การประเมินประสิทธิภาพ	SVM 83.38%	หาค่าเฉลี่ยจากแบบสอบถาม 3.27	SVM 79.21%
แสดงผล	เว็บแอปพลิเคชัน	เว็บแอปพลิเคชัน	เว็บแอปพลิเคชัน

2.3 การเปรียบเทียบงานวิจัยที่เกี่ยวข้อง

จากตารางที่ 2-2 ได้เปรียบเทียบวิธีการดึงข้อมูล , การวิเคราะห์คำ , การประเมินประสิทธิภาพ และการแสดงผล ปรากฏว่า การดึงข้อมูลนั้นได้จากแบบสอบถามเหมือนกันทั้ง 3 งานวิจัย

การวิเคราะห์คำมีการใช้โปรแกรม Lexto สำหรับการตัดคำเพื่อสำหรับไปวิเคราะห์โดยทั้ง 3 งานวิจัยจะใช้เหมือนกัน เป็นซอฟต์แวร์ฟรีที่พัฒนาโดย ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ

การประเมินประสิทธิภาพของความคิดเห็นนั้นจะใช้ Support-vector machine ทั้ง 2 งานวิจัยได้แก่ รีวิวสินค้าออนไลน์และเหมือนความคิดเห็นบนเครือข่ายสังคมออนไลน์โดยค่าความแม่นยำไม่ต่างกันมากนักซึ่งรีวิวสินค้าจะมากกว่าอยู่ที่ 83.38 % และร้านอาหารเว็บไซต์รีวิวนั้นจะประเมินจากแบบสอบถามที่สุ่มจากนิสิตจำนวน 20 คน

การแสดงผลการวิเคราะห์ความคิดเห็นนั้นจะออกมาเป็น เว็บแอปพลิเคชัน เหมือนกันทั้ง 3 งานวิจัย

บทที่ 3

การออกแบบและการพัฒนา

3.1 ระบบวิเคราะห์ระบบงาน

ระบบงานนี้มีการออกแบบการทำงานเพื่อให้สอดคล้องกับวัตถุประสงค์และขอบเขตของปริญญาานิพนธ์โดยได้นำความรู้ทางด้านการเขียนเว็บแอปพลิเคชัน การประมวลภาษาธรรมชาติ และการทำเว็บเซิร์ฟเวอร์มาประยุกต์ใช้เพื่อพัฒนา ระบบวิเคราะห์ข้อมูลความคิดเห็นที่มีต่อมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือจากเว็บไซต์สำหรับวัยรุ่นโดยเฉพาะ ซึ่งมีรูปแบบการทำงานด้วยกันหลายส่วน โดยแต่ละส่วนได้มีการออกแบบที่แตกต่างกันออกไป ทุกส่วนของการออกแบบจึงมีความสำคัญอย่างมาก เพื่อให้ปริญญาานิพนธ์นี้เสร็จสมบูรณ์ ให้ประโยชน์แก่ผู้อ่าน และระบบวิเคราะห์ข้อมูลความคิดเห็นที่มีต่อมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือจากเว็บไซต์สำหรับวัยรุ่นโดยเฉพาะ สามารถช่วยให้ผู้ใช้งานได้รับความสะดวกสบายในชีวิตประจำวันมากขึ้น โดยแบ่งส่วนของการออกแบบทั้งหมดดังนี้

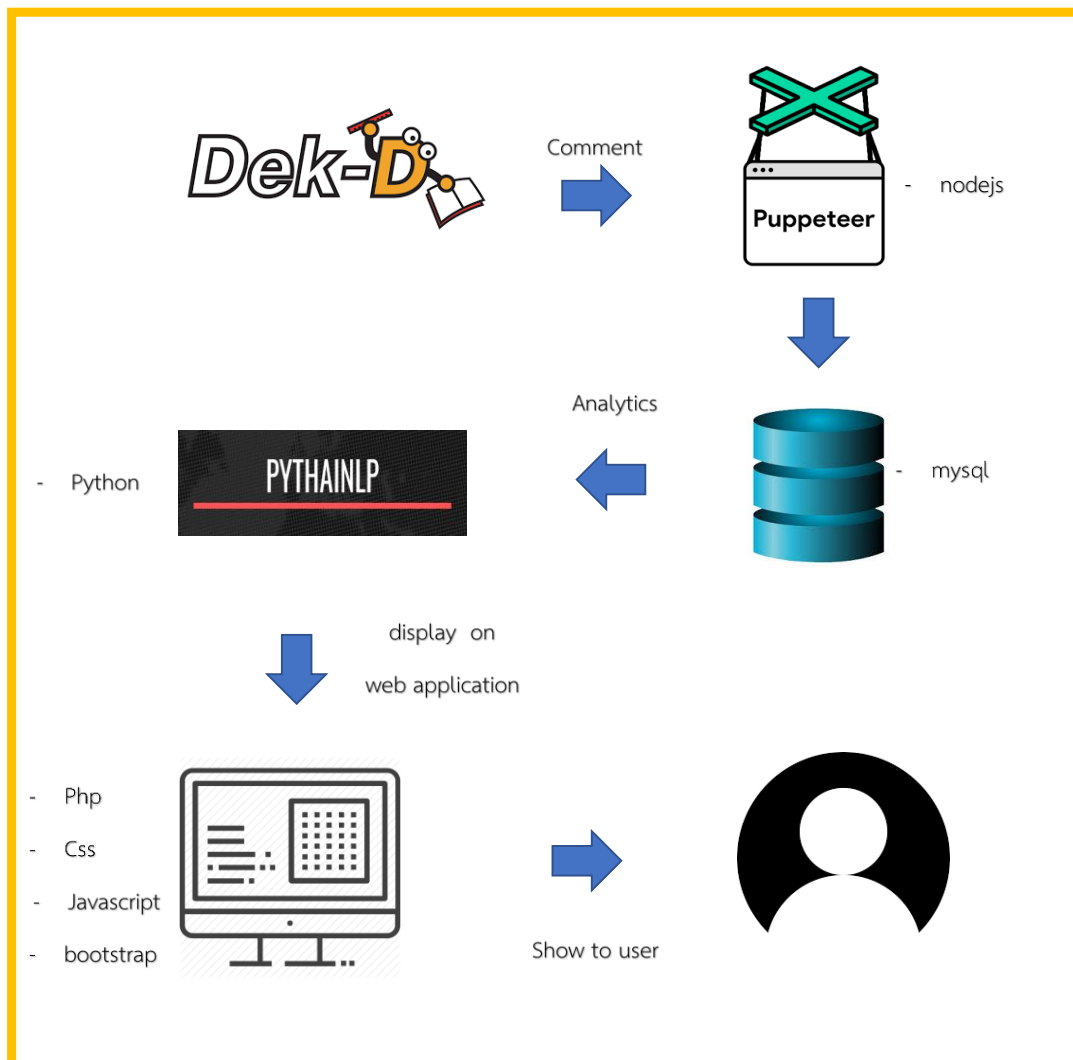
3.1.1 การวิเคราะห์ระบบงานเดิม

ระบบงานเดิมของมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือเน้นการที่ต้องวิเคราะห์ความคิดเห็นที่มีต่อมหาวิทยาลัยนั้น ต้องเก็บข้อมูลความคิดเห็นในแต่ละกระทู้ของเว็บที่เป็นเป้าหมายซึ่งการเก็บข้อมูลความคิดเห็นนั้นจำเป็นต้องใช้คนเปิดทีละกระทู้จนได้ข้อมูลความคิดเห็นที่ต้องการสำหรับวิเคราะห์ความคิดเห็น หลังจากนั้นก็นำข้อมูลที่ได้ไปวิเคราะห์ด้วยการใช้คนคัดการณเอาจว่าความคิดเห็นที่ได้เป็นในเชิงบวกหรือเชิงลบ ซึ่งวิธีการดังกล่าวนี้ใช้เวลาสิ้นเปลืองมากในการวิเคราะห์ข้อมูลได้อาจจะกินเวลาไปถึง 1 สัปดาห์ หรือ อาจมากกว่านั้น

3.1.2 การวิเคราะห์ระบบงานใหม่

ระบบงานใหม่ สำหรับการเก็บข้อมูลความคิดเห็นในแต่ละกระทู้ ผู้พัฒนาจะทำการสร้างอินเทอร์เน็ตบอทสำหรับเก็บข้อมูลโดยเฉพาะ มันจะทำหน้าที่อ่านข้อมูลในแต่ละกระทู้และจัดเก็บลงฐานข้อมูลเองอัตโนมัติจากนั้น ผู้พัฒนาจะสร้างระบบสำหรับการวิเคราะห์ความคิดเห็นของแต่ละกระทู้โดยมันจะทำโดยมันจะทำการแยกความคิดเห็นว่าเป็นเชิงบวกหรือเชิงลบ จากข้อมูลที่ได้มาในฐานข้อมูลและสรุปเป็นข้อมูลเชิงสถิติบนหน้าจอเว็บแอปพลิเคชัน เพื่อความสะดวกแก่ผู้ใช้งานในการที่จะตรวจสอบข้อมูลที่วิเคราะห์ได้จากระบบ ซึ่งวิธีดังกล่าวจะลดเวลาลงไปมากเทียบกับระบบงานเดิม

3.2 การออกแบบการทำงานของระบบ



ภาพที่ 3-1 Diagram การทำงานของระบบ

3.3 การออกแบบของระบบ จะแบ่งได้ส่วนต่าง ๆ ได้ดังนี้

3.3.1 ในส่วนของการจัดเก็บข้อมูลจะเป็นการเก็บข้อมูลจากกระทู้ของเว็บไซต์เด็กดีโดย Key word คือ มจพ. และ KMUTNB โดยตัวที่ใช้ในการกวาดข้อมูลคือ puppeteer เป็น Library ของภาษา Nodejs ที่ทำหน้าที่เปรียบเสมือนอินเทอร์เน็ตบอทในการดูแลและจัดเก็บข้อมูลของเว็บไซต์ต่าง ๆ จากนั้นนำข้อมูลที่ได้อาจจัดเก็บลงฐานข้อมูล

3.3.2 ในส่วนของการวิเคราะห์ความรู้สึกเราจะใช้ PYTHAINLP เป็น Library ของ Python

3.3.2.1 Word Segmentation สำหรับการตัดคำ ในภาษาไทย

3.3.2.2 Stop – Word List Removal คือการนำคำที่ไม่มีนัยสำคัญออก โดยที่ไม่ทำให้ความหมายของเอกสารเปลี่ยนแปลงไป

3.3.2.3 Feature Selection โดยใช้ค่าการเพิ่มของข้อมูล (Information Gain : IG) เป็นตัวเลือกในการเลือกคุณลักษณะซึ่งการคำนวณค่า Information Gain นี้ จะใช้ค่า Entropy ซึ่งเป็นการวัดความแตกต่างหรือการกระจายกระจายของข้อมูลถ้า

3.3.3 ในส่วนของการแสดงผลจะเป็นในรูปแบบ เว็บไซต์แอปพลิเคชันโดยใช้ PHP ในการออกแบบจะมีการ Login เพื่อเช็คว่าเป็น ผู้ดูแลระบบสามารถหรือเป็นผู้ใช้ทั่วไปสามารถ โดยการแสดงผลจะขึ้นอยู่กับสถานะ ผู้ใช้งานดังนี้

3.3.3.1 ผู้ดูแลระบบสามารถ

3.3.3.1.1 แสดงข้อมูลเชิงสถิติเกี่ยวกับความคิดเห็นเชิงบวกและเชิงลบ

3.3.3.1.2 แสดงข้อมูลเชิงสถิติตามช่วงเวลาที่มีความคิดเห็นบนบอร์ดข้อความ

3.3.3.1.3 แสดงเชิงสถิติตามความคิดเห็นเชิงลบเยอะที่สุด 10 อันดับแรก

3.3.3.1.4 แสดงเชิงสถิติตามความคิดเห็นเชิงบวกเยอะที่สุด 10 อันดับแรก

3.3.3.1.5 สามารถเข้าติดตามเข้าไปยังบอร์ดข้อความเพื่อชี้แจงข้อมูลเชิงลบได้

3.3.3.1.6 สามารถแสดงผลการวิเคราะห์เหมืองข้อความใช้สรุปข้อความความ

คิดเห็นในเชิงสถิติ

3.3.3.2 ผู้ใช้ทั่วไปสามารถ

3.3.3.2.1 แสดงข้อมูลเชิงสถิติตามช่วงเวลาที่มีความคิดเห็นบนบอร์ดข้อความ

3.3.3.2.2 แสดงเชิงสถิติตามความคิดเห็นเชิงลบเยอะที่สุด 10 อันดับแรก

3.3.3.2.3 แสดงเชิงสถิติตามความคิดเห็นเชิงบวกเยอะที่สุด 10 อันดับแรก

3.4 การใช้ API ในการวิเคราะห์ความคิดเห็น

ตารางที่ 3-1 แสดง Confusion Matrix ของ API aiforthai ซึ่งมีค่าความถูกต้องเท่ากับร้อยละ 70

จากการตรวจสอบของผู้พัฒนาระบบ

aiforthai		ความคิดเห็นเชิงบวก	ความคิดเห็นเชิงลบ	รวม
	ความคิดเห็นเชิงบวก	60	1	189
	ความคิดเห็นเชิงลบ	4	25	45
	รวม	57	4	615

ตารางที่ 3-2 แสดง Confusion Matrix ของ API repustate ซึ่งมีค่าความถูกต้องเท่ากับร้อยละ 50

จากการตรวจสอบของผู้พัฒนาระบบ

repustate		ความคิดเห็นเชิงบวก	ความคิดเห็นเชิงลบ	รวม
	ความคิดเห็นเชิงบวก	40	1	382
	ความคิดเห็นเชิงลบ	3	15	22
	รวม	78	14	445

ซึ่ง API ที่เราได้เลือกนั้นมีอยู่ 2 ที่ คือ aiforthai ซึ่งเป็นของคนไทยสร้างกับ repustate ซึ่งเป็นของต่างประเทศสร้างจากตารางที่ 3-1 และตารางที่ 3-2 คือการเปรียบเทียบความถูกต้องความคิดเห็นของ API กับ สิ่งที่ผู้จัดทำพิจารณา (โดยมนุษย์) หลังการเปรียบเทียบพบว่า aiforthai มีค่าประสิทธิภาพที่สูงกว่า repustate ดังนั้น ผู้พัฒนาจึงเลือกใช้ API ของ aiforthai ในการจำแนกความคิดเห็นเชิงบวกและเชิงลบของระบบ

3.5 การใช้ API ในการจัดกลุ่ม

ตารางที่ 3-3 แสดง Confusion Matrix ในการจัดประเภทความคิดเห็นของ API aiforthai ซึ่งมีค่าความถูกต้องเท่ากับร้อยละ 79

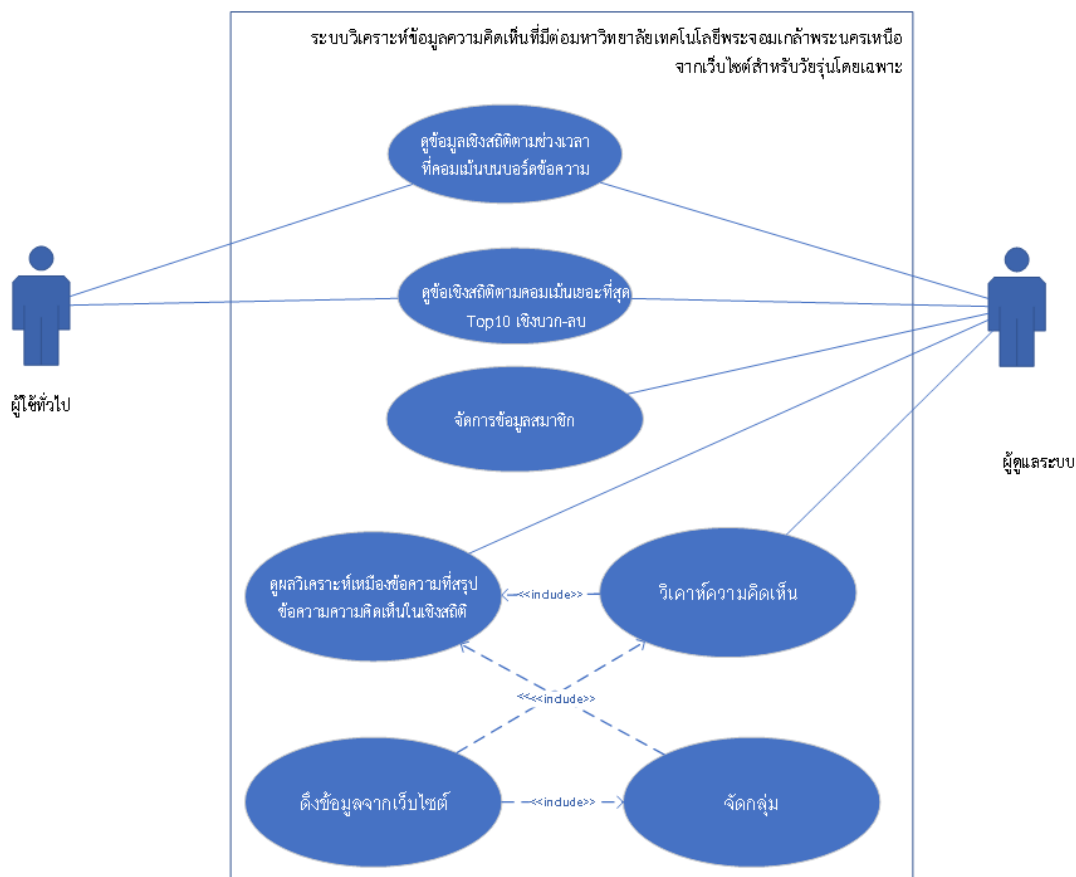
จากการตรวจสอบของผู้พัฒนาระบบ

aiforthai		การศึกษา	หอพัก	หาเพื่อน	กิจกรรม	อื่นๆ
	การศึกษา	676	2	13	0	12
	หอพัก	0	40	10	0	1
	หาเพื่อน	39	2	11	0	0
	กิจกรรม	3	1	0	3	0
	อื่นๆ	117	32	27	1	242

ในการจัดกลุ่มนั้นเราได้ใช้ API ของ aiforthai ในการจัดกลุ่มซึ่งผลลัพธ์ที่ได้ทำให้ได้กลุ่มมามากมายเราจึงยุบรวมอยู่ 5 กลุ่มดังนี้คือ การศึกษา, หอพัก, กิจกรรม, อื่นๆ

3.4 ขั้นตอนการออกแบบระบบ

3.4.1 Use Case Diagram



ภาพที่ 3-2 Use Case Diagram

จากภาพที่ 3-2 ของ Use case ในฝั่งของผู้ใช้ทั่วไปจะสามารถเข้าไปดูในส่วนข้อมูลเชิงสถิติตามช่วงเวลาที่แสดงความคิดเห็นและดูในส่วนการจัดอันดับความเห็นในแง่บวกและในแง่ลบได้ ส่วนผู้ดูแลระบบสามารถทำได้ทุกอย่างในส่วนของผู้ใช้ทั่วไปและมีฟังก์ชันเสริมคือสามารถเข้าไปเว็บเด็กดีในกรณีที่มีความคิดเห็นเชิงลบมากเกินไปและดูผลการวิเคราะห์จากเหมือนข้อความได้

3.4.2 คำอธิบาย Use Case Diagram

ตารางที่ 3-4 Use Case Id 1

Use Case Title : ดูข้อมูลเชิงสถิติตามช่วงเวลา ที่ความคิดเห็นบนบอร์ดข้อความ	Use Case Id : 1
Primary Actor : ผู้ดูแลระบบ	
Stakeholder Actor : ผู้ใช้ทั่วไป	
Main Flow : ผู้ใช้ทั่วไปและผู้ดูแลระบบสามารถดูข้อมูลเชิงสถิติตามเวลาที่บุคคลทั่วไปได้ สามารถความคิดเห็น หรือตั้งกระทู้ในเว็บไซต์ได้โดยจะแสดงในรูปแบบ dashboard ออกมาในกราฟบนหน้าเว็บใน ระบบ	

ตารางที่ 3-5 Use Case Id 2

Use Case Title : เข้าลิงค์ไปยังบอร์ดข้อความ เพื่อชี้แจงข้อมูลเชิงลบได้	Use Case Id : 2
Primary Actor : ผู้ดูแลระบบ	
Stakeholder Actor :	
Main Flow : ในกรณีที่เกิดข้อมูลเชิงลบจำนวนมากในกระทู้ของ เว็บเด็กดีผู้ดูแลระบบสามารถเข้ายังลิงค์ของ กระทู้นั้นเพื่อชี้แจงข้อผิดพลาดในเว็บกระทู้นั้น	

ตารางที่ 3-6 ตาราง Use Case Id 3

Use Case Title : ดูผลวิเคราะห์เหมือนข้อความที่สรุปข้อความความคิดเห็นในเชิงสถิติ	Use Case Id : 3
Primary Actor : ผู้ดูแลระบบ	
Stakeholder Actor :	
Main Flow : หลังจากที่ได้มีการเข้ากระบวนการวิเคราะห์ความคิดเห็น สามารถนำข้อมูลการวิเคราะห์ความคิดเห็นมาแสดงผลในเชิงสถิติได้ โดยแสดงบนเว็บแอปพลิเคชันในรูปแบบของ Dashboard	

ตารางที่ 3-7 ตาราง Use Case Id 4

Use Case Title : ดูข้อมูลเชิงสถิติตามความคิดเห็นเยอะที่สุด Top10 เชิงบวก-ลบ	Use Case Id : 4
Primary Actor : ผู้ดูแลระบบ	
Stakeholder Actor : ผู้ใช้ทั่วไป	
Main Flow : ผู้ใช้ทั่วไปและผู้ดูแลระบบสามารถดูข้อเชิงสถิติตามความคิดเห็นที่ระบบได้จัดอันดับความรู้สึกว่าความคิดเห็นที่บุคคลทั่วไปได้แสดงความคิดเห็นในเว็บเด็กดีนั้นเป็นในทางแง่บวกหรือแง่ลบโดย จัดอันดับเป็น 10 อันดับแรก	

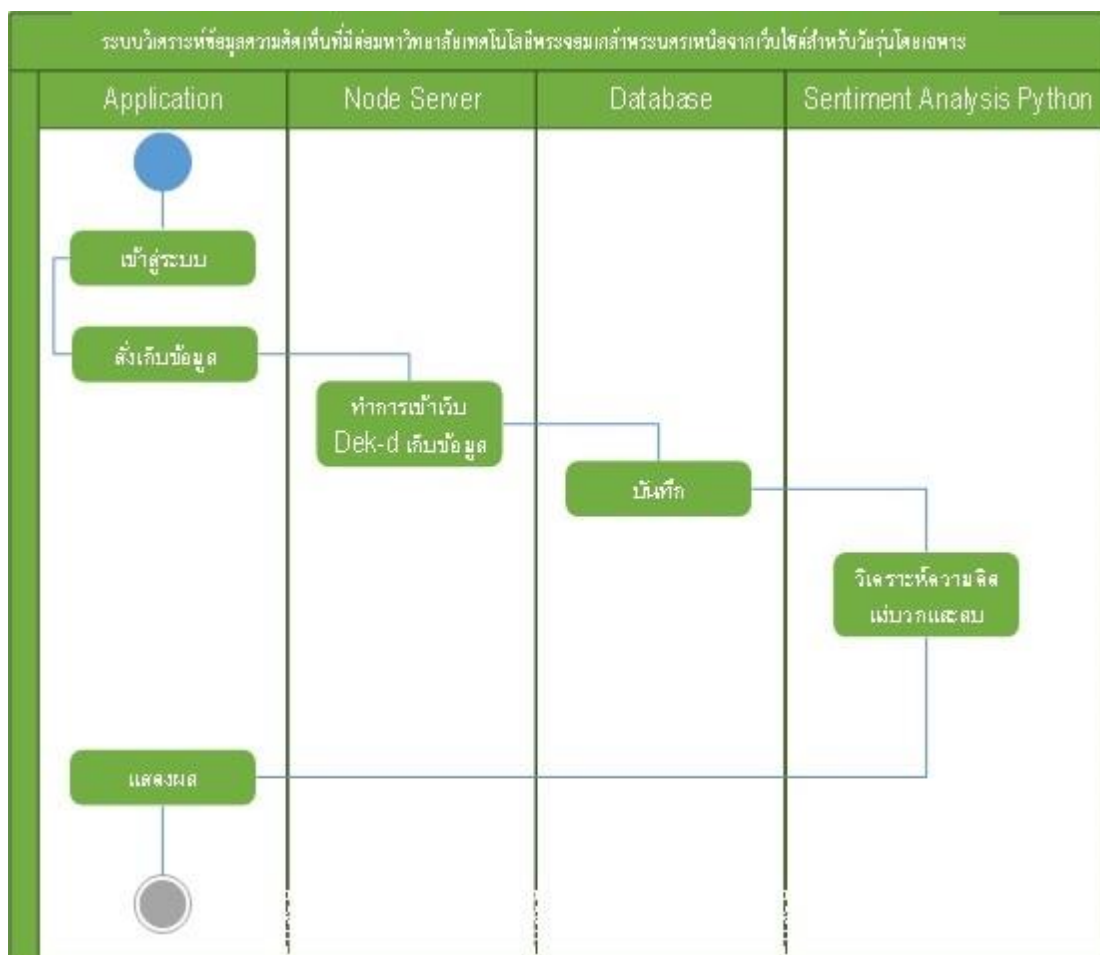
ตารางที่ 3-8 ตาราง Use Case Id 5

Use Case Title : การดึงข้อมูลข้อมูลจากเว็บไซต์	Use Case Id : 5
Primary Actor : ผู้ดูแลระบบ	
Stakeholder Actor :	
Main Flow : ผู้ดูแลระบบสามารถดึงข้อมูลจากเว็บไซต์ได้โดยผู้ดูแลระบบสามารถกรอก Keyword ที่ต้องการดึงมาได้และหลังจากที่ดึงมา ระบบจะสามารถกรองข้อมูลที่เกี่ยวข้องกับมหาวิทยาลัยพระจอมเกล้าพระนครเหนือให้อัตโนมัติ	

ตารางที่ 3-9 ตาราง Use Case Id 6

Use Case Title : การวิเคราะห์ความคิดเห็น	Use Case Id : 6
Primary Actor : ผู้ดูแลระบบ	
Stakeholder Actor :	
Main Flow : หลังจากผ่านการดึงข้อมูลจากเว็บไซต์ได้เรียบร้อยแล้วนั้น ผู้ดูแลระบบจะสามารถนำข้อมูลมาวิเคราะห์ความคิดเห็นจากข้อมูลที่ได้ดึงมาว่าความคิดเห็นนั้นเป็นไปทางเชิงบวก เชิงลบหรือ ปกติ	

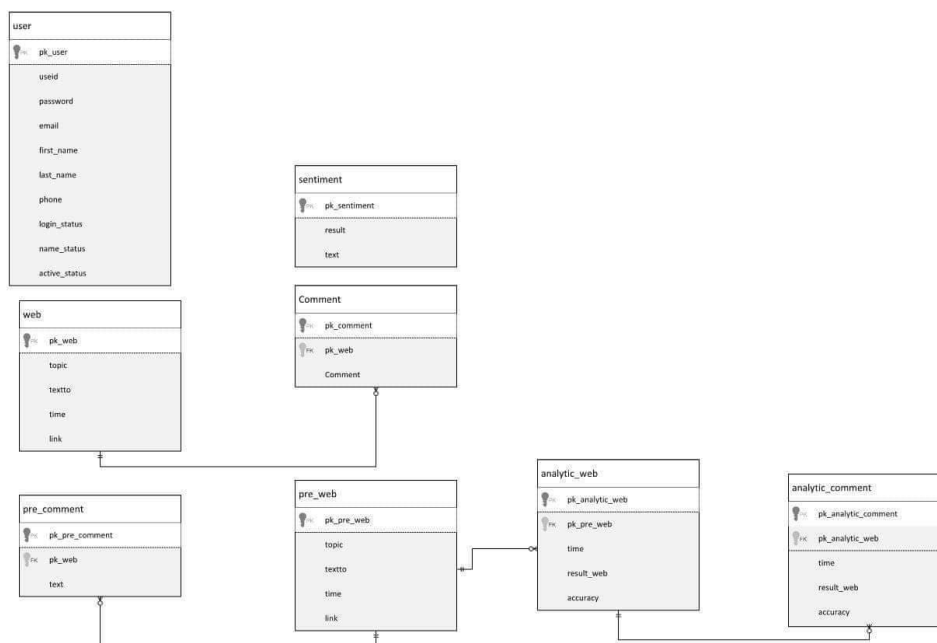
3.5 Activity diagram



ภาพที่ 3-3 Activity diagram

จากภาพที่ 3-3 Activity diagram จะเป็นการทำงานของตัว Application เมื่อเริ่มการทำงานของเว็บ Server จากฝั่ง Nodejs จะทำการเก็บข้อมูลในเว็บ จากนั้นก็บันทึกลงในฐานข้อมูลแล้วนำข้อมูลในฐานข้อมูลมาวิเคราะห์ความคิดเห็นแบ่งกลุ่มเมื่อวิเคราะห์เสร็จก็แสดงผลออกในเว็บ Application

3.6 ER-Diagram



ภาพที่ 3-4 ER-Diagram

จากภาพที่ 3-4 ER-Diagram ประกอบด้วย 6 ตารางคือ pre_web เป็นตารางที่เอาไว้สำหรับเก็บกระทู้ ตาราง pre_comment ไว้สำหรับเก็บ ความคิดเห็นของกระทู้นั้น ตาราง analytic_web เป็นตารางสำหรับ ข้อมูลผลวิเคราะห์ความรู้สึกของกระทู้ ตาราง analytic_comment เป็นตารางสำหรับ ข้อมูลผลวิเคราะห์ความรู้สึกของความคิดเห็น ตารางสำหรับ sentiment เป็นตารางสำหรับเก็บข้อมูลความรู้สึก ตาราง user เป็น ตารางสำหรับเก็บข้อมูลผู้ใช้

3.7 การออกแบบฐานข้อมูล

3.7.1 ตาราง pre_web

ตารางที่ 3-10 ตาราง pre_web

ชื่อฟิลด์	ความหมาย	ชนิดของข้อมูล	ขนาด	ประเภทยุติ
pk_pre_web	รหัสของกระทู้	Integer	10	Primary Key
Topic	หัวข้อกระทู้	Varchar	150	Null
Text	เนื้อหากระทู้	Varchar	250	Null
Time	เวลาแต่ละกระทู้	Varchar	10	Null
Link	ลิงค์เข้าเว็บใน กระทู้	Varchar	100	Null

3.7.2 ตาราง pre_comment

ตารางที่ 3-11 ตาราง pre_comment

ชื่อฟิลด์	ความหมาย	ชนิดของข้อมูล	ขนาด	ประเภทยุติ
pk_pre_comment	รหัสของความคิดเห็น	Integer	10	Primary Key
pk_pre_web	รหัสของกระทู้	Integer	10	Foreign Key
Comment	ความคิดเห็นกระทู้	Varchar	250	Null

3.7.3 ตาราง analytic_web

ตารางที่ 3-12 ตาราง analytic_web

ชื่อฟิลด์	ความหมาย	ชนิดของข้อมูล	ขนาด	ประเภทคีย์
pk_analytic_web	รหัสของการวิเคราะห์กระทู้	Integer	10	Primary Key
pk_pre_web	รหัสกระทู้	Integer	10	Foreign Key
Time	เวลาในการวิเคราะห์	Varchar	10	Null
result_web	ผลลัพธ์การวิเคราะห์ในแต่ละกระทู้	Varchar	20	Null
Accuracy	ค่าความถูกต้อง	Float	10	Null

3.7.4 ตาราง analytic_comment

ตารางที่ 3-13 ตาราง analytic_comment

ชื่อฟิลด์	ความหมาย	ชนิดของข้อมูล	ขนาด	ประเภทคีย์
pk_analytic_comment	รหัสของการวิเคราะห์ความคิดเห็น	Integer	10	Primary Key
pk_analytic_web	รหัสของการวิเคราะห์กระทู้	Integer	10	Foreign Key
time	เวลาในการวิเคราะห์	Varchar	10	Null
result_comment	ผลลัพธ์การวิเคราะห์ในแต่ละความคิดเห็น	Varchar	20	Null
accuracy	ค่าความถูกต้อง	Float	10	N

3.7.2 ตาราง sentiment

ตารางที่ 3-14 ตาราง Comment

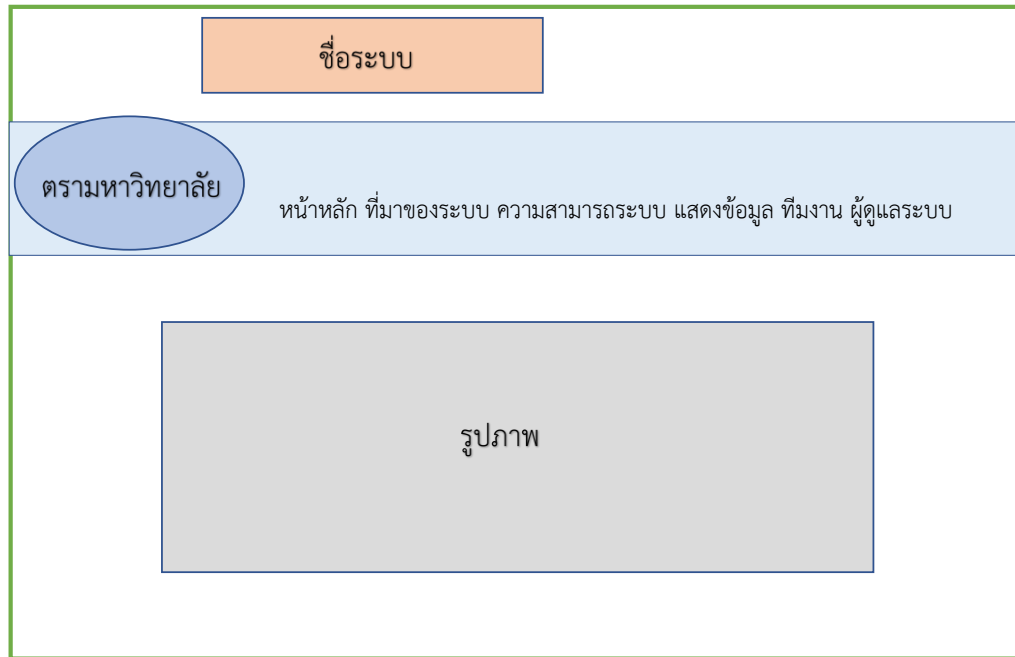
ชื่อฟิลด์	ความหมาย	ชนิดของข้อมูล	ขนาด	ประเภทคีย์
pk_sentiment	รหัสของความรู้สึก	Integer	10	Primary Key
pk_analytic_web	รหัสของการวิเคราะห์ความคิดเห็น	Integer	10	Foreign Key
result	ผลลัพธ์ความรู้สึก	Varchar	20	Null
text	ข้อมูลคำไว้วางใจความรู้สึก	Varchar	250	Null

3.7.2 ตาราง user

ตารางที่ 3-15 ตาราง user

ชื่อฟิลด์	ความหมาย	ชนิดของข้อมูล	ขนาด	ประเภทคีย์
pk_user	รหัสของผู้ใช้	Integer	10	Primary Key
userid	ชื่อสำหรับใช้งานของผู้ใช้	Varchar	150	Null
password	รหัสผ่านสำหรับผู้ใช้งาน	Varchar	250	Null
email	อีเมลล์ของผู้ใช้งาน	Varchar	100	Null
first_name	ชื่อจริงผู้ใช้งาน	Varchar	100	Null
last name	นามสกุลผู้ใช้งาน	Varchar	100	Null
phone	เบอร์โทรผู้ใช้งาน	Char	10	Null
login status	สถานะการเข้าใช้งาน	Varchar	20	Null
name status	สถานะของผู้ใช้งาน	Varchar	20	Null
active_status	สถานะการทำงาน	Varchar	20	Null

3.8 การออกแบบเว็บแอปพลิเคชัน



ภาพที่ 3-5 การออกแบบหน้าหลัก

จากภาพที่ 3-5 หน้าหลักนั้นประกอบไปด้วยชื่อระบบซึ่งจัดตำแหน่งไว้อยู่บนสุดและเอียงทางซ้ายเล็กน้อยพอ ตรามหาวิทยาลัยจะอยู่ตำแหน่งทางซ้ายลงจากชื่อระบบเล็กน้อย เมนูของระบบจะอยู่ตำแหน่งถัดจากตรามหาวิทยาลัยจากนั้นตำแหน่งรูปภาพคือตำแหน่งลงมาจากเมนูซึ่งรูปภาพจะเป็นรูปภาพที่เกี่ยวข้องกับมหาวิทยาลัย



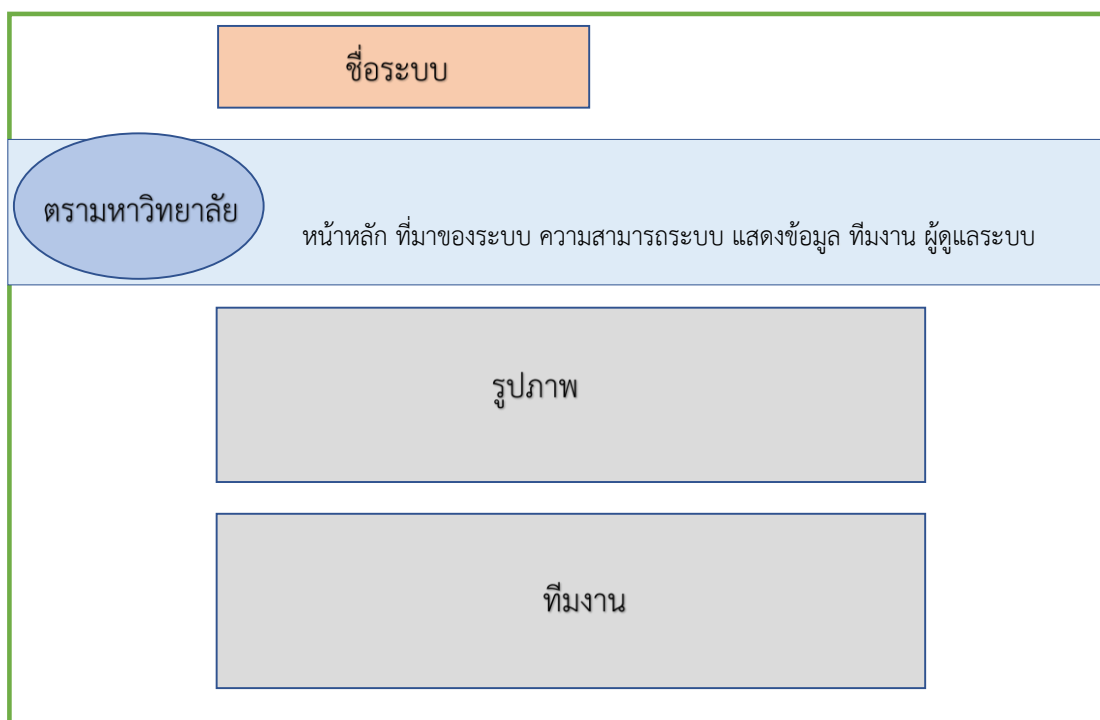
ภาพที่ 3-6 การออกแบบที่มาของระบบ

จากภาพที่ 3-6 จะเหมือนกับหน้าหลักโดยจะมีที่มาของระบบไว้ได้รูปภาพเนื้อหาจะกล่าวถึงที่มาและความสำคัญของระบบนี้



ภาพที่ 3-7 การออกแบบเพื่อวัดประสิทธิภาพของระบบ

จากภาพที่ 3-7 จะเหมือนกับหน้าหลักโดยจะเพิ่มประสิทธิภาพของระบบไว้ล่างรูปภาพโดยเนื้อหาจะกล่าวถึงอัลกอริทึม ภาษา และแพ็คเกจ ต่าง ๆ ที่นำมาใช้กับระบบนี้



ภาพที่ 3-8 การออกแบบหน้าทีมงาน

จากภาพที่ 3-8 จะเหมือนกับหน้าหลักโดยจะเพิ่มส่วนของทีมงานเข้าไว้ล่างรูปภาพจะบอกเกี่ยวกับทีมงานที่จัดทำระบบพร้อมรูปภาพของทีมงาน

ชื่อระบบ

ตรามหาวิทยาลัย

หน้าหลัก ที่มาของระบบ ความสามารถระบบ แสดงข้อมูล ที่มงาน ผู้ดูแลระบบ

ลำดับ	เวลา	เนื้อหา	ความคิดเห็น
x	dd/mm/yyyy	xxxxxxxxxxxxxxxxxx	คลิก
x	dd/mm/yyyy	xxxxxxxxxxxxxxxxxx	คลิก

previou	1	2	next
---------	---	---	------

ภาพที่ 3-9 การออกแบบหน้าแสดงข้อมูลกระทู้

จากภาพที่ 3-9 จะเหมือนกับหน้าหลักโดยจะเพิ่มการแสดงผลที่ดึงมาในแต่ละกระทู้โดยจะเพิ่ม การแบ่งเลขหน้าไว้ข้างล่างและปุ่มคลิกตรงช่องความคิดเห็นเพื่อดูความคิดเห็นของแต่ละกระทู้ได้

หน้าเนื้อหาเว็บขอบบน 1.5 นิ้วค่ะ

35

ชื่อระบบ

ตรามหาวิทยาลัย

หน้าหลัก ที่มาของระบบ ความสามารถระบบ แสดงข้อมูล ที่มงาน ผู้ดูแลระบบ

ลำดับ	ความคิดเห็น	ลิง
x	xxxxxxxxxxxxxxxxxx	xxxxxxxxxxxxxxxxxx
x	xxxxxxxxxxxxxxxxxx	xxxxxxxxxxxxxxxxxx

previou

1

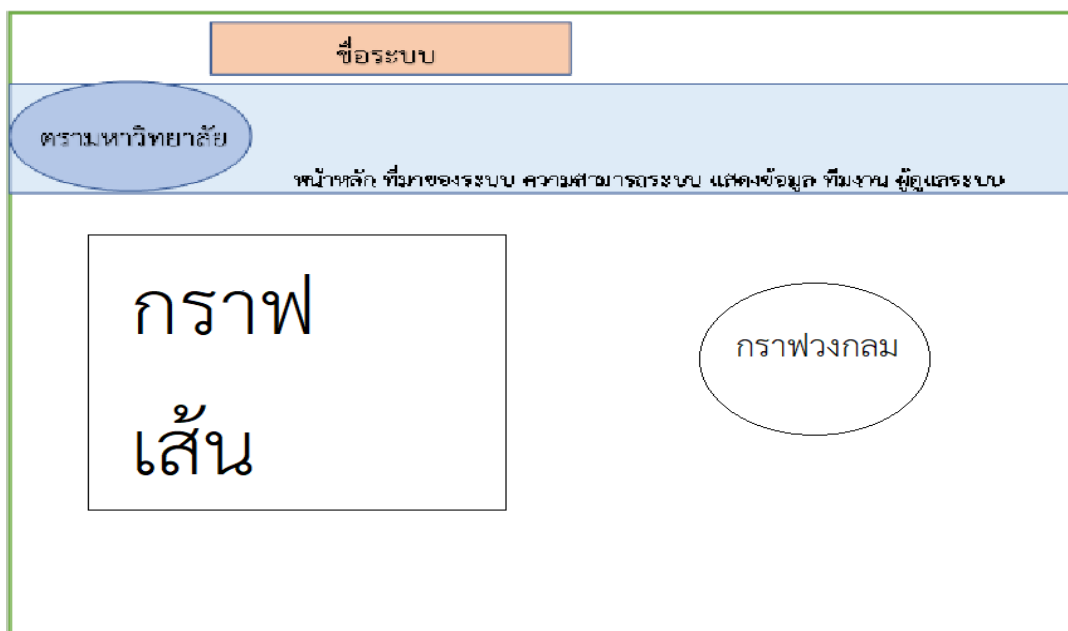
2

next

back

ภาพที่ 3-10 การออกแบบหน้าแสดงข้อมูลความคิดเห็น

จากภาพที่ 3-10 จะเหมือนกับหน้าหลักโดยจะเพิ่มการแสดงผลข้อมูลที่ดึงมาในแต่ละความคิดเห็นที่มาจากกระทู้ได้



ภาพที่ 3-11 การออกแบบหน้า dashbord

จากภาพที่ 3-11 เป็นการออกแบบหน้า dashboard ความคิดเห็นในแต่ละเดือนในรูปแบบกราฟเส้นและสรุปยอดรวมในรูปแบบกราฟวงกลม

บทที่ 4

ผลการดำเนินงาน

ในการปฏิบัติงานของโครงการพิเศษนี้ได้มีการกำหนด และวางแผนการดำเนินงานเพื่อที่จะสามารถดำเนินงานไปได้อย่างสำเร็จตรงตามระยะเวลาที่กำหนดไว้ และตรงไปตามวัตถุประสงค์ของโครงการที่ได้กำหนดไว้โดยมีดังต่อไปนี้

แผนการดำเนินงานโครงการพิเศษ

แผนการดำเนินงานของโครงการพิเศษ ระบบวิเคราะห์ข้อมูลความคิดเห็นที่มีต่อมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือจากเว็บไซต์สำหรับวัยรุ่นโดยเฉพาะ

1. การศึกษาค้นคว้าและเก็บรวบรวมข้อมูล
2. วิเคราะห์เว็บไซต์เด็กดี
3. ออกแบบฐานข้อมูล
4. ดึงข้อมูลจากเว็บไซต์เด็กดี
5. วิเคราะห์ความคิดเห็น
6. จัดกลุ่มของกระทู้
7. แสดงข้อมูลบน เว็บแอปพลิเคชัน

4.1 ขั้นตอนการศึกษาค้นคว้าและเก็บรวบรวมข้อมูล

ในการศึกษาค้นคว้าและการเก็บรวบรวมข้อมูลสำหรับโครงการพิเศษนี้ได้มีการศึกษาและเก็บรวบรวมข้อมูล เพื่อนำมาออกแบบและสร้างเว็บแอปพลิเคชันในการแสดงข้อมูล โดยการศึกษาค้นคว้าในส่วนนี้จะเกี่ยวข้องกับทฤษฎีการเขียนโปรแกรมโดยใช้ภาษา php , css , nodejs, python ซึ่งเป็นภาษาหลักในการพัฒนาเว็บแอปพลิเคชันซึ่งจะต้องมีการฝึกเขียนให้เกิดความคุ้นเคยกับรูปแบบของภาษาแต่ละภาษาที่นำมาใช้ โดยขั้นตอนการศึกษา การเขียนเว็บแอปพลิเคชันนี้จะใช้เวลาค่อนข้างมาก เพราะข้อมูลและภาษาที่ใช้มีความซับซ้อนในการใช้งาน

4.2 ขั้นตอนวิเคราะห์เว็บไซต์เด็กดี

ในส่วนของขั้นตอนการวิเคราะห์ข้อมูลนี้ในทางด้านการเขียนโปรแกรมเพื่อให้สามารถทำงานควบคู่ไปกับเว็บได้นั้น ต้องศึกษาโครงสร้างของเว็บเพื่อที่จะเขียนโปรแกรมได้ง่ายต่อการดึงข้อมูล และนำข้อมูลไปเก็บไว้ยังฐานข้อมูลโดยจะเริ่มศึกษา Document Object Model โครงสร้างต้นไม้ของ

HTML ของเว็บไซต์เด็กดี คือการมองส่วนหน้าเว็บไซต์เด็กดีให้เป็น Object และเมื่อเราวิเคราะห์โครงสร้าง Object เหล่านั้นได้เราก็สามารถเข้าถึงข้อมูลแล้วเก็บลงในฐานข้อมูลได้

4.3 ขั้นตอนการออกแบบฐานข้อมูล

การออกแบบ ฐานข้อมูล เป็นส่วนที่สำคัญมากผู้พัฒนาจึงจำเป็นต้องออกแบบฐานข้อมูลให้สอดคล้องกับข้อมูล que ผู้พัฒนาได้ไปดึงจากเว็บไซต์เด็กดี เพื่อที่จะนำข้อมูลความคิดเห็นจากฐานข้อมูลไปใช้ในโอกาสต่อไปได้ในอนาคต

4.4 ขั้นตอนการดึงข้อมูลจากเว็บไซต์เด็กดี

ในขั้นตอนการดึงข้อมูลนั้น ผู้พัฒนาได้ใช้ nodejs และ แพคเกจ Puppeteer ในเก็บข้อมูลความคิดเห็นของแต่ละกระทู้ โดยจะออกมาในรูปแบบอินเทอร์เน็ทบอท สำหรับเก็บข้อมูลอัตโนมัติและส่งเข้าไปยังฐานข้อมูล และนำข้อมูลที่เก็บมานั้นแสดงบนเว็บแอปพลิเคชัน php, css, html, javascript

4.5 วิเคราะห์ความคิดเห็น

ในการวิเคราะห์ความคิดเห็นนั้น ผู้พัฒนาได้นำ API ของ aiforthai.in.th มาช่วยในการวิเคราะห์ความคิดเห็นของกระทู้, เนื้อหากระทู้, และหัวข้อกระทู้ จากนั้นนำผลลัพธ์และค่าความถูกต้องที่ได้เก็บลงบนฐานข้อมูล

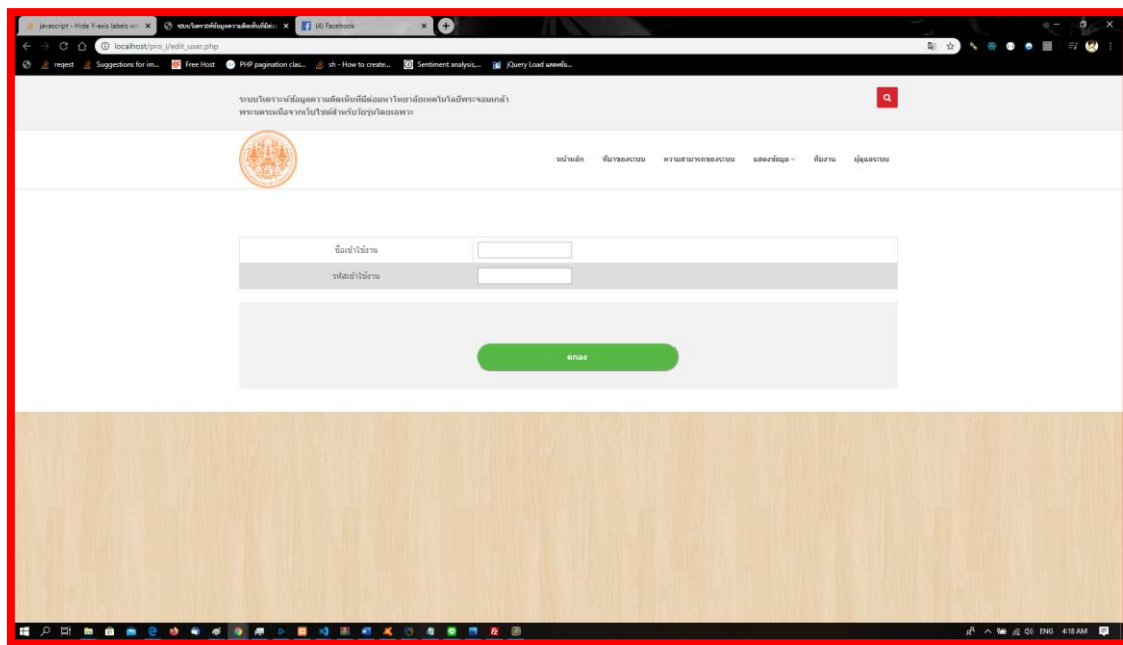
4.6 การจัดกลุ่ม

ในการจัดกลุ่มนั้นผู้พัฒนาได้นำ API ของ aiforthai.in.th มาช่วยในการจัดกลุ่มของหัวข้อกระทู้เนื่องจากผลลัพธ์ของการจัดกลุ่มมีมากมายผู้พัฒนาจึงยุบรวมเหลืออยู่ 5 กลุ่ม การศึกษา, หอพัก, หาเพื่อน และอื่น จากนั้นนำผลลัพธ์และค่าความถูกต้องที่ได้เก็บลงบนฐานข้อมูล

4.7 ผลของการดำเนินการ

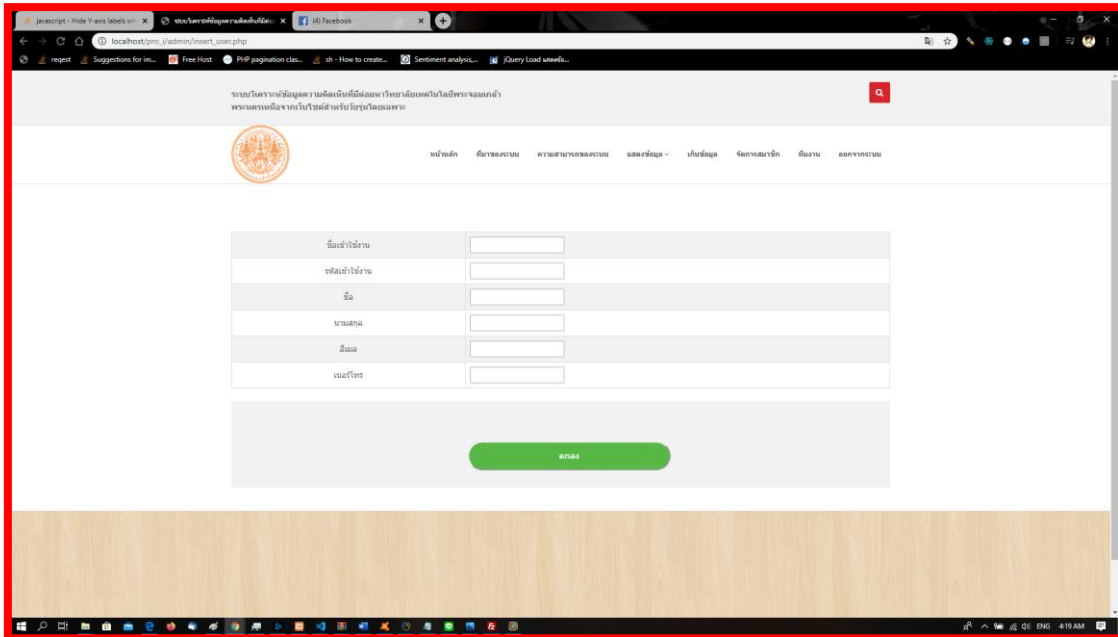
จากการดำเนินการตามโครงการพิเศษนี้ สามารถแสดงผลตามวัตถุประสงค์ที่จัดทำในภาคเรียนที่ 2/2562 คือ สามารถแสดงข้อมูลบนเว็บได้

4.8 ลักษณะของโปรแกรม



ภาพที่ 4-1 เข้าสู่ระบบ

จากภาพที่ 4-1 เป็นหน้าจอเข้าสู่ระบบของระบบวิเคราะห์ข้อมูลความคิดเห็นที่มีต่อมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือจากเว็บไซต์สำหรับวัยรุ่นโดยเฉพาะ



ระบบบริหารงานบุคคล มหาวิทยาลัยราชภัฏวไลยอลงกรณ์

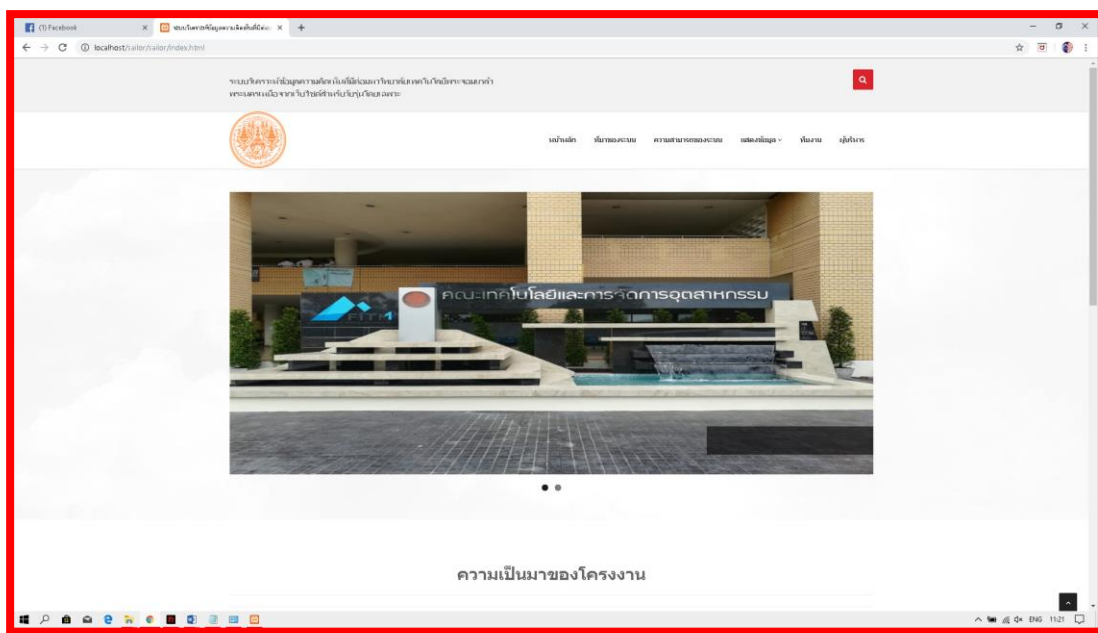
หน้าหลัก | ข่าวสาร | ความสามารถของระบบ | ติดต่อเรา | เกี่ยวกับเรา | ติดต่อเรา | ติดต่อเรา

ชื่อเข้าใช้งาน	<input type="text"/>
รหัสเข้าใช้งาน	<input type="password"/>
ชื่อ	<input type="text"/>
นามสกุล	<input type="text"/>
อีเมล	<input type="text"/>
เบอร์โทร	<input type="text"/>

ภาพที่ 4-2 หน้าจอสมัครสมาชิก

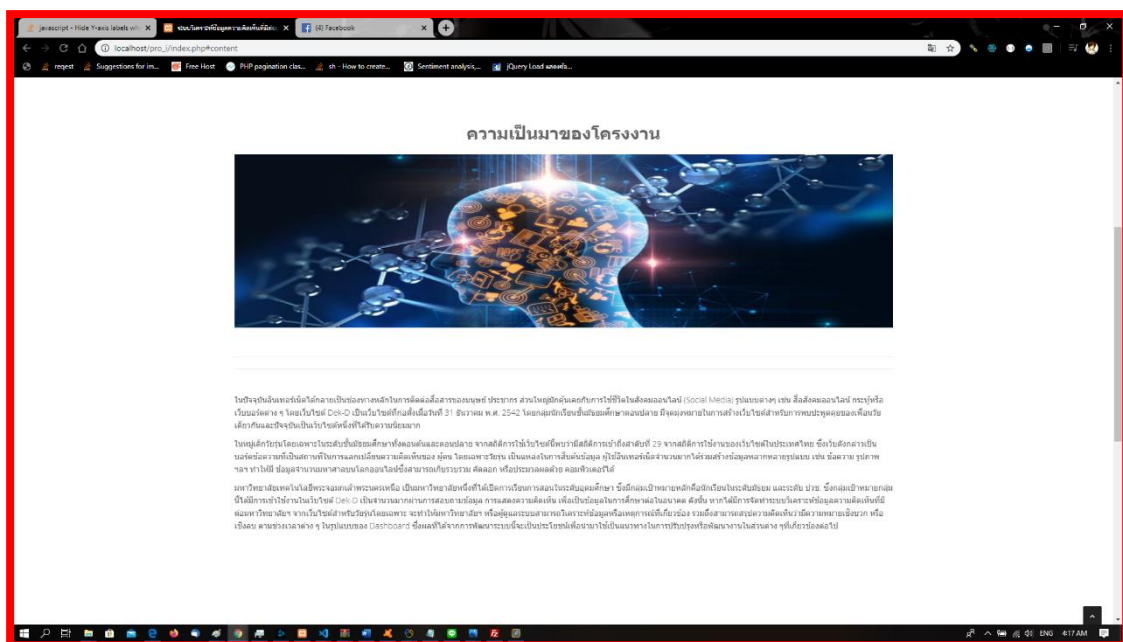
จากภาพที่ 4-2 เป็นหน้าจอสมัครสมาชิกของระบบซึ่งประกอบด้วย

- ชื่อที่ใช้เข้าสู่ระบบ
- รหัสผ่านที่เอาไว้เข้าสู่ระบบ
- ชื่อจริง และ นามสกุล
- อีเมล
- เบอร์โทร



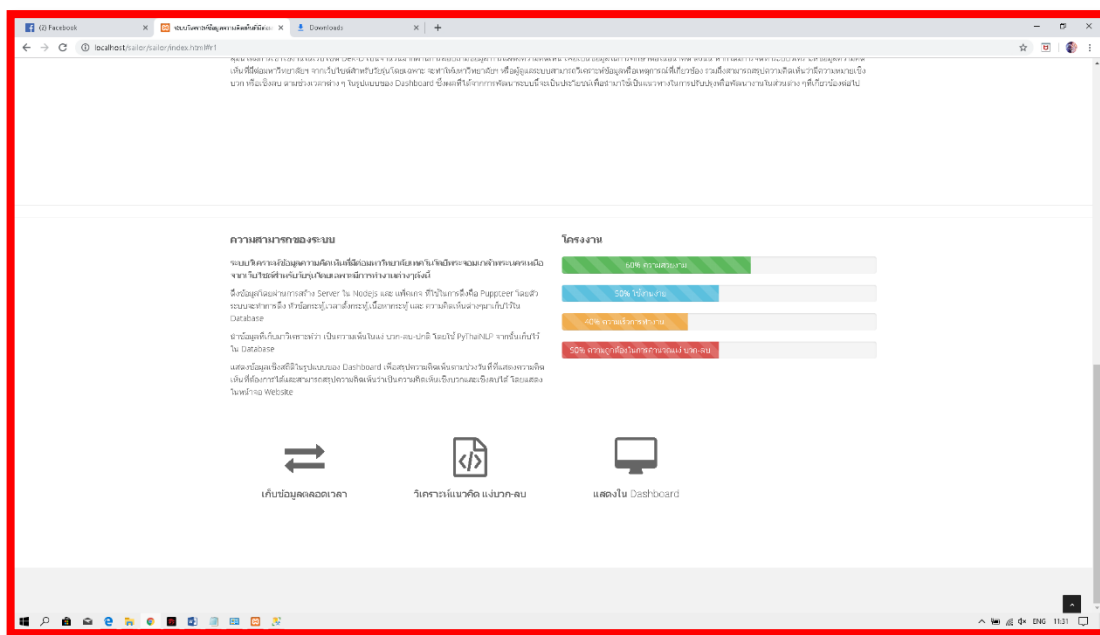
ภาพที่ 4-3 ภาพแสดงหน้าหลัก

จากภาพที่ 4-3 เป็นหน้าจอหลักของระบบงานมีเมนูประกอบด้วย หน้าหลัก ที่มาของระบบ ความสามารถของระบบ การแสดงข้อมูล ทีมงาน และส่วนของผู้บริหาร



ภาพที่ 4-4 แสดงที่มาระบบ

จากภาพที่ 4-4 จะเป็นการอธิบายที่มาและความสำคัญต่างที่อยู่ในระบบ



ภาพที่ 4-5 แสดงประสิทธิภาพของระบบ

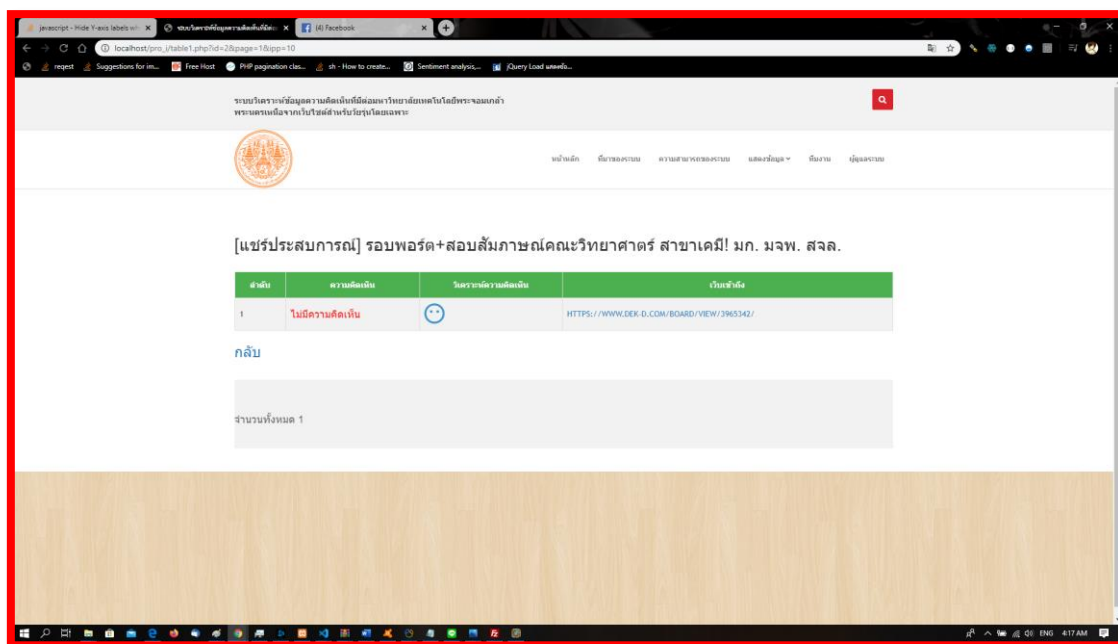
จากภาพที่ 4-5 จะเป็นการบอกรายละเอียดประสิทธิภาพของระบบ ภาษาที่ใช้งาน แพ็คเกจต่างๆ อัลกอริทึม ต่าง ๆ ที่ใช้งานภายในระบบนี้

ลำดับ	หัวข้อ	เวลา	สถานะ	การตีความ	วิเคราะห์ความรู้สึก	วิเคราะห์ความรู้สึก	รวมทั้งหมด
1	[โปรดระบุการ] รอยพริ้ว-รอยยิ้มภายใต้แว่นตา สายตาดี มก. มช. สด.	16/03/63 15:10	▶ ดูเนื้อหา	การตีความ	😊	😊	คลิก
2	สมาชิกในวงสังคมออนไลน์ของ มช. 42	17/03/63 14:07	▶ ดูเนื้อหา	การตีความ	😊	😊	คลิก
3	ขอไปขอเอกสารการศึกษารายละเอียดของงานและพิจารณาจาก มช.	08/03/63 23:04	▶ ดูเนื้อหา	การตีความ	😊	😊	คลิก
4	มช. ลงนามบันทึกข้อตกลงความร่วมมือความร่วมมือทางวิชาการกับ วิทยาลัย กุฎี จ.บุรีรัมย์ (พหุภาค)	02/03/63 14:11	▶ ดูเนื้อหา	การตีความ	😊	😊	คลิก
5	เอกสารงานด้านสังคมศึกษา 2.000คนลงนามในเอกสาร มช. นำมาใช้จริง	01/03/63 18:13	▶ ดูเนื้อหา	การตีความ	😊	😊	คลิก
6	มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือจัดงานวันวิชาการ มช. "61 ปี มช. นวัตกรรมก้าวหน้า เพื่ออนาคตสังคม"	26/02/63 16:08	▶ ดูเนื้อหา	การตีความ	😊	😊	คลิก
7	วิทยาลัย (2 ปี) วิทยาลัย มช. กรุงเทพฯ	14/02/63 22:46	▶ ดูเนื้อหา	การตีความ	😊	😊	คลิก
8	หนังสือคณะกรรมการ/คณะกรรมการในชุดที่ (ชุดที่ 2)	05/02/63 20:30	▶ ดูเนื้อหา	การตีความ	😊	😊	คลิก
9	เอกสารงานด้านสังคมศึกษา 2.000คนลงนามในเอกสาร มช. นำมาใช้จริง	05/02/63 19:48	▶ ดูเนื้อหา	การตีความ	😊	😊	คลิก
10	เอกสารงานด้านสังคมศึกษา 2.000คนลงนามในเอกสาร มช. นำมาใช้จริง	04/02/63 09:57	▶ ดูเนื้อหา	การตีความ	😊	😊	คลิก

Page: 1
จำนวนทั้งหมด 1164

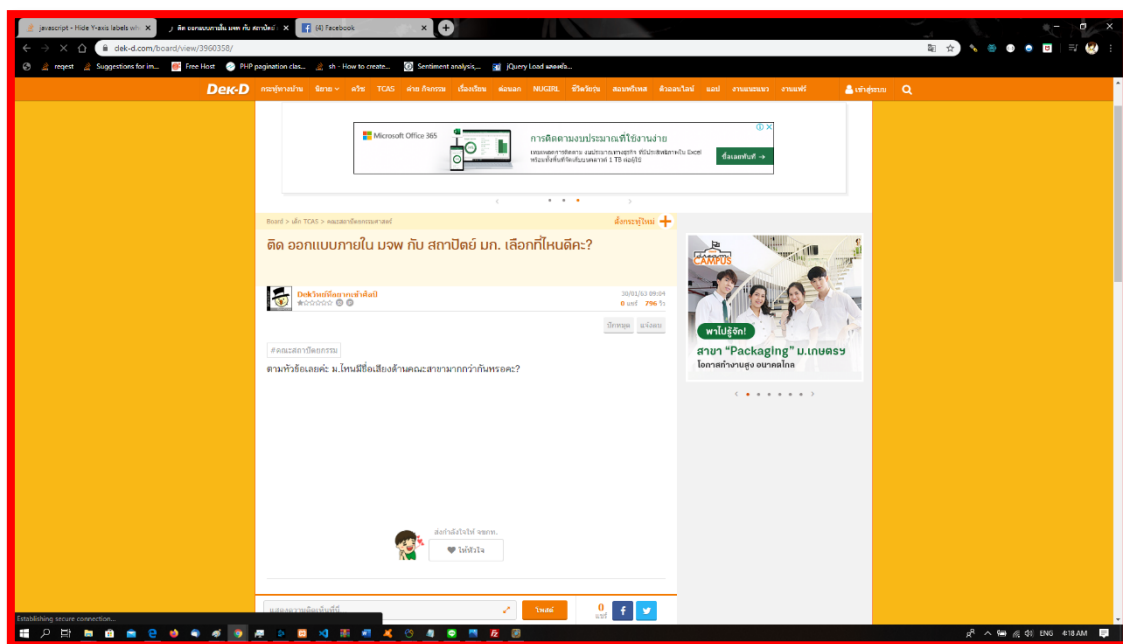
ภาพที่ 4-6 การแสดงข้อมูลกระทู้ได้เกิด

จากภาพที่ 4-6 จะเป็นการแสดงกระทู้ได้เกิดที่ระบบได้เก็บเข้ามาในฐานข้อมูลได้มีการสร้างเลขหน้าในการแบ่งจำนวนข้อมูลเพื่อไม่ให้ข้อมูลนั้นมีการแสดงเยอะเกินไป และเนื้อหาได้ซ่อนปุ่มเพื่อแสดงเนื้อหาเอาไว้เพราะตัวเนื้อหานั้นมีตัวอักษรที่เยอะเกินไป มีการแสดงค่าที่วิเคราะห์ความคิดเห็นของหัวข้อกระทู้และเนื้อหาในกระทู้ในรูปแบบของ ภาพแสดงอารมณ์ ถ้าเอาเมาส์ไปชี้ที่ภาพจะแสดงค่าความถูกต้องออกมา นอกจากนี้ยังมีปุ่ม “คลิก” ตรงความคิดเห็นเพื่อแสดงข้อมูลความคิดเห็นของแต่ละกระทู้ได้



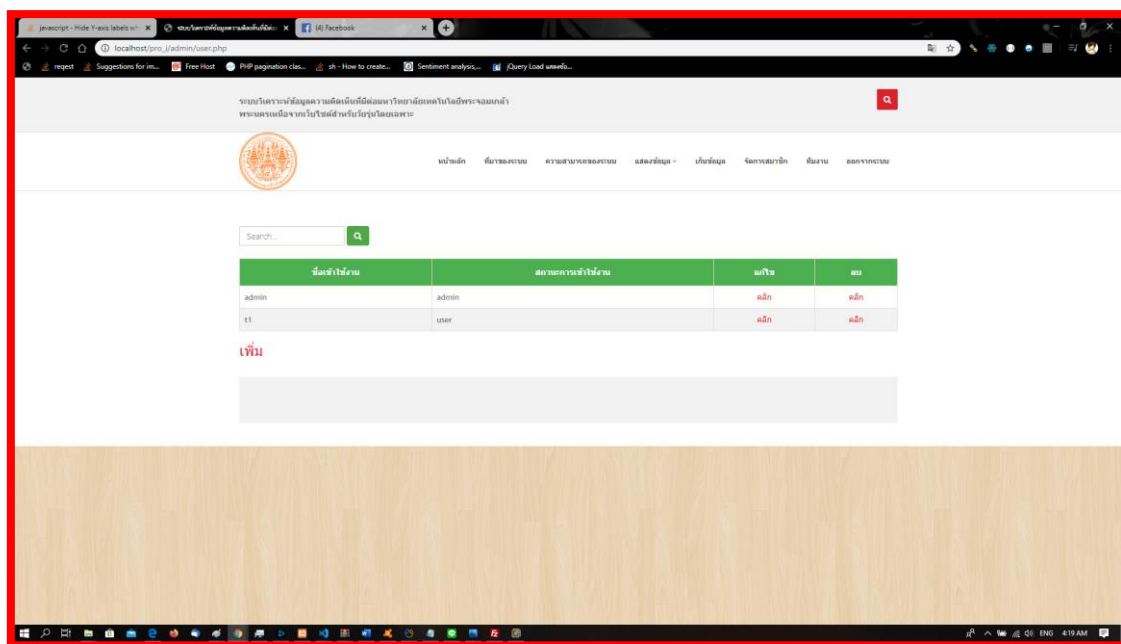
ภาพที่ 4-7 แสดงข้อมูลแต่ละความคิดเห็น

จากภาพที่ 4-7 เมื่อกดปุ่มคลิกของกระทู้เพื่อดูความคิดเห็นจะถูกนำมามานี้ซึ่งจะมีความคิดเห็นต่างๆที่แสดงในแต่ละกระทู้และมีลิงก์เว็บไซต์ไว้สำหรับเข้าไปดูกระทู้ที่ทางระบบได้เก็บมาได้ มีการแสดงค่าที่วิเคราะห์ความคิดเห็นในรูปแบบของ ภาพแสดงอารมณ์ ถ้าเอาเมาส์ไปชี้ที่ภาพจะแสดงค่าความถูกต้องออกมา นอกจากนี้ยังมีปุ่มกลับเพื่อกลับดูข้อมูลตารางกระทู้ที่เคยคลิกมาได้



ภาพที่ 4-8 หน้าเว็บเด็กดี

จากภาพที่ 4-8 จากภาพเกิดจากที่กดเพื่อเข้าถึงเว็บไซต์ จากตารางความคิดเห็นของระบบ



ภาพที่ 4-9 จัดการสมาชิก

จากภาพที่ 4-9 เป็นการจัดการสมาชิกในส่วนของ Admin ซึ่งจะมีในส่วนเพิ่ม แก้ไข ลบ

จากรายการ ขาดข้อมูล หรือเป็นข้อมูลที่ผิดพลาด กรุณาตรวจสอบข้อมูลก่อนจะทำการแก้ไข
 พิจารณาตามเนื้อหาจากใบรายชื่อสำหรับปรับปรุงข้อมูล

หน้าหลัก | ข่าวสารระบบ | ความสามารถของระบบ | แหล่งข้อมูล | เกี่ยวกับเรา | ติดต่อเรา | ข้อมูลระบบ | ติดต่อเรา

ชื่อใช้งาน	<input type="text" value="admin"/>
รหัสใช้งาน	<input type="password" value="*****"/>
อีเมล	<input type="text" value="admin"/>
เบอร์โทร	<input type="text" value="admin"/>
ชื่อ	<input type="text" value="admin@admin.com"/>
นามสกุล	<input type="text" value="00000000"/>

ภาพที่ 4-10 แก้ไขสมาชิก

จากภาพที่ 4-10 จะเป็นการแก้ไขสมาชิก

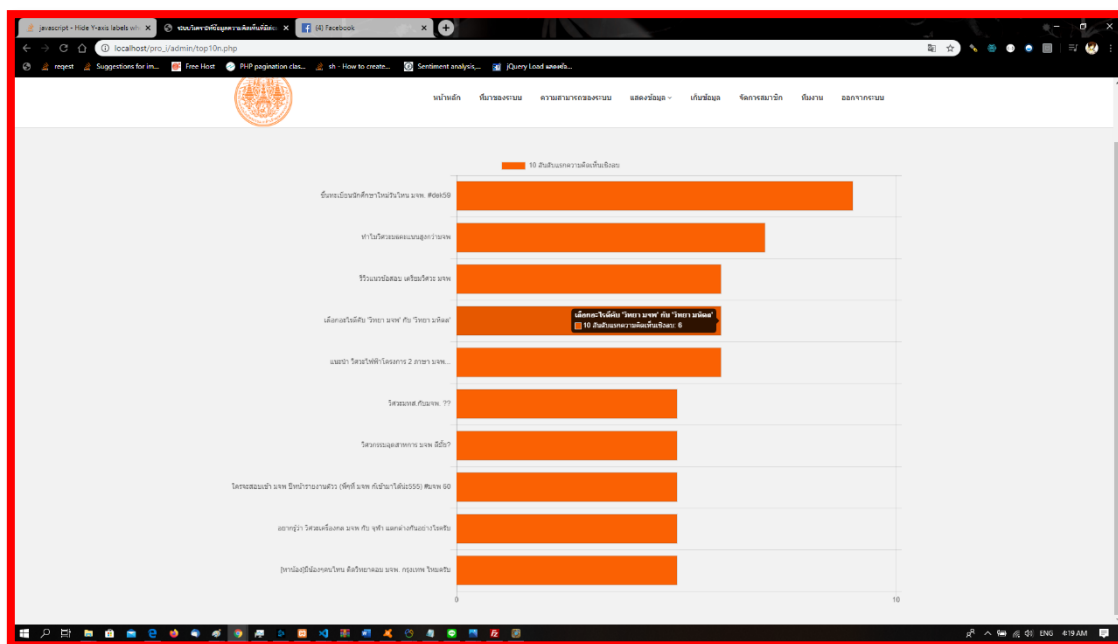
ภาพที่ 4-11 การเพิ่มข้อมูล

จากภาพที่ 4-11 เป็นการเพิ่มข้อมูลโดยเราสามารถที่จะพิมพ์ Keyword ที่ต้องการที่จะเพิ่มข้อมูลลง Database ได้ที่ ช่องกรอก Keyword และจะมีสถานการณ์ทำงานที่กำลังเพิ่มข้อมูลอยู่ตรงช่อง สถานะตั้งข้อมูล



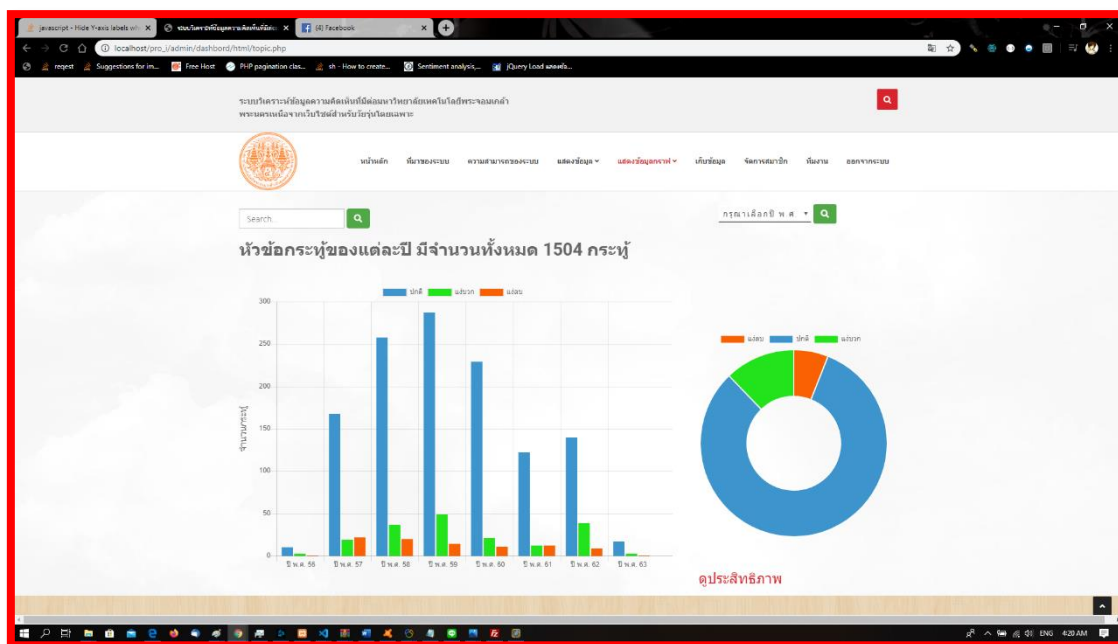
ภาพที่ 4-12 แสดงความคิดเห็น 10 อันดับแรกความคิดเห็นในเชิงบวก

จากภาพที่ 4-12 จะเป็นการแสดงความคิดเห็น 10 อันดับแรกความคิดเห็นในเชิงบวกในรูปแบบของกราฟแท่งแบบแนวนอนถ้าเราคลิกที่กราฟของกระทุ้เราจะสามารถไปยังหน้าความคิดเห็นของกระทุ้นั้นได้



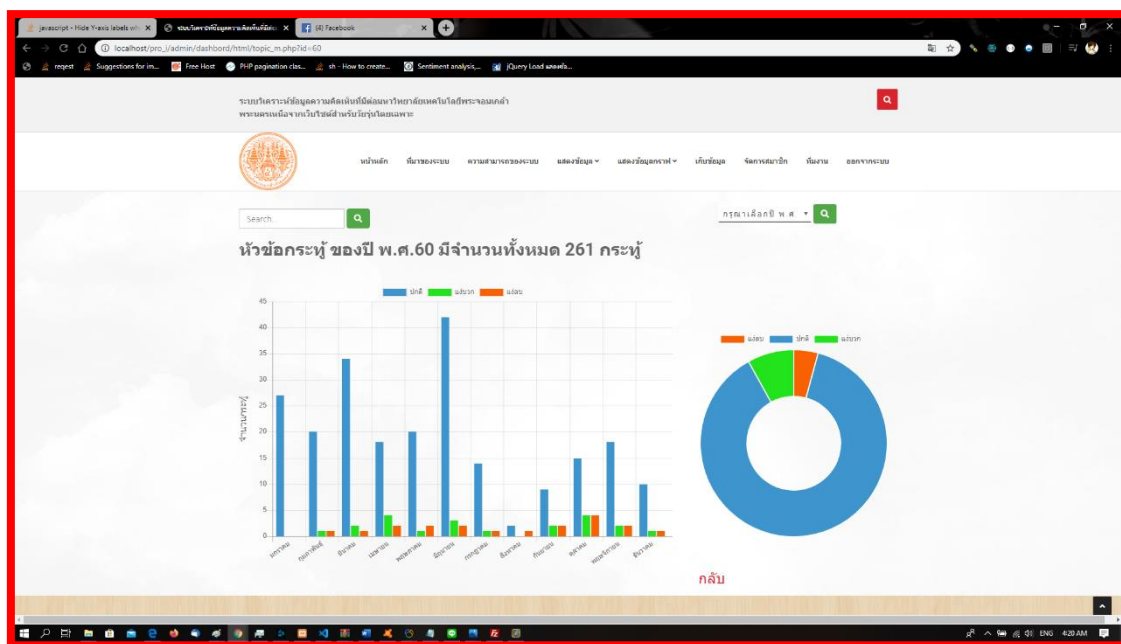
ภาพที่ 4-13 แสดงความคิดเห็น 10 อันดับแรกความคิดเห็นในเชิงลบ

จากภาพ 4-13 จะเป็นการแสดงความคิดเห็น 10 อันดับแรกความคิดเห็นในเชิงลบในรูปแบบของกราฟแท่งแบบแนวนอนถ้าเราคลิกที่กราฟของกระทู้เราจะสามารถไปยังหน้าความคิดเห็นของกระทู้นั้นได้



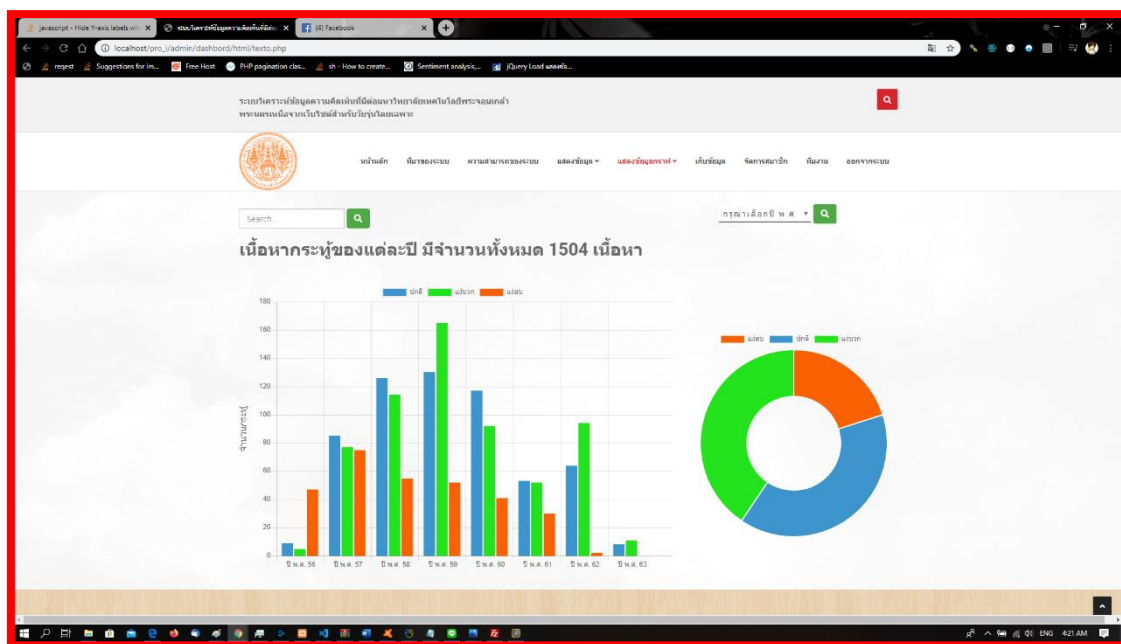
ภาพที่ 4-14 วิเคราะห์หัวข้อกระทู้

จากภาพที่ 4-14 จะเป็นการแสดงกราฟของการวิเคราะห์หัวข้อกระทู้โดยกราฟเส้นนั้นจะเป็นการบอกจำนวนความถี่ของหัวข้อกระทู้ในแต่ละปีตามความรู้สึกที่โปรแกรมได้วิเคราะห์มาได้โดยเส้นสีฟ้าคือความรู้สึกปกติเส้น สีเขียวความรู้สึกในบวกเส้นสีแดงความรู้สึกในแง่ลบ ส่วนกราฟวงกลมจะบอกจำนวนรวมในแต่ละปี นอกจากนี้ยังมีปุ่ม “ดูประสิทธิภาพ” เพื่อดูประสิทธิภาพที่ API ได้วิเคราะห์มาได้



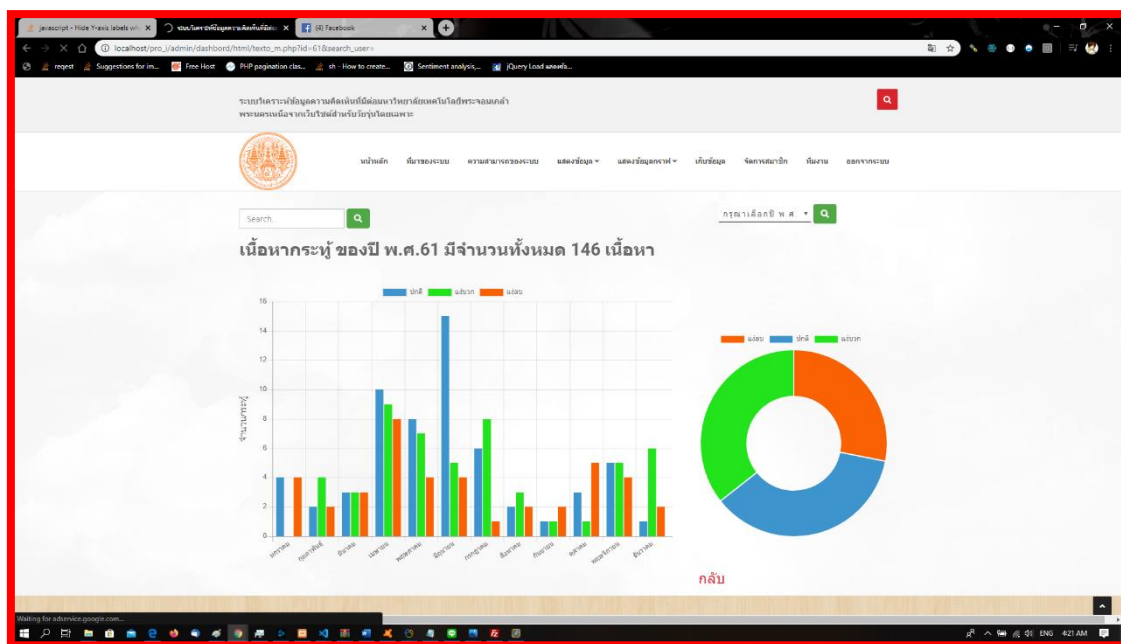
ภาพที่ 4-15 แสดงค่าความถูกต้องวิเคราะห์หัวข้อกระทู้

จากภาพ 4-15 แสดงค่าความถูกต้องวิเคราะห์หัวข้อกระทู้ จะเป็นการแสดงค่าความถูกต้องที่ API ได้วิเคราะห์หัวข้อกระทู้ในรูปแบบกราฟแท่งและกราฟวงกลม โดยถ้าเป็น เชนกปกติ ค่าที่คืนกลับมาจะเป็น 0 นอกจากนี้ยังมีปุ่มกลับเพื่อกลับดูข้อมูลตารางกระทู้ที่เคยคลิกมาได้



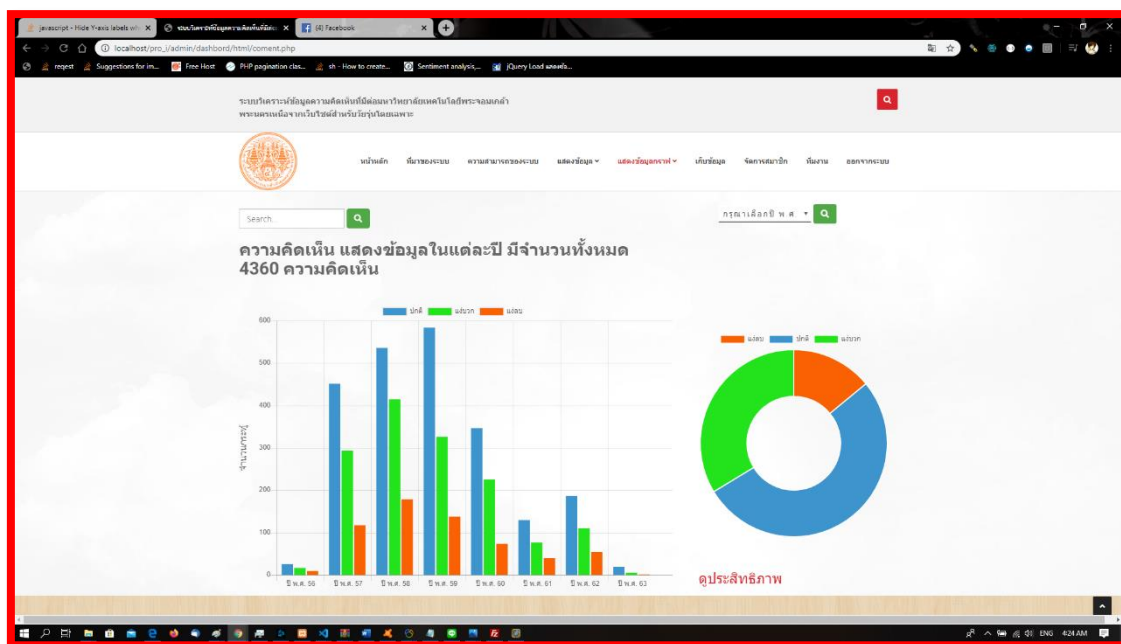
ภาพที่ 4-16 วิเคราะห์ของเสียจากอาหาร

จากภาพที่ 4-16 จะเป็นการแสดงกราฟของการวิเคราะห์ของเสียจากอาหารโดยกราฟเส้นนั้นจะ เป็นการบอกจำนวนความถี่ของของเสียในแต่ละปีตามความรู้สึกที่โปรแกรมได้วิเคราะห์มาได้โดย เส้นสีฟ้าคือความรู้สึกปกติ เส้น สีเขียวความรู้สึกในบวก เส้นสีแดงความรู้สึกในแง่ลบ ส่วนกราฟ วงกลมจะบอกจำนวนรวมในแต่ละปี นอกจากนี้ยังมีปุ่ม “ดูประสิทธิภาพ” เพื่อดูประสิทธิภาพที่ API ได้วิเคราะห์มาได้



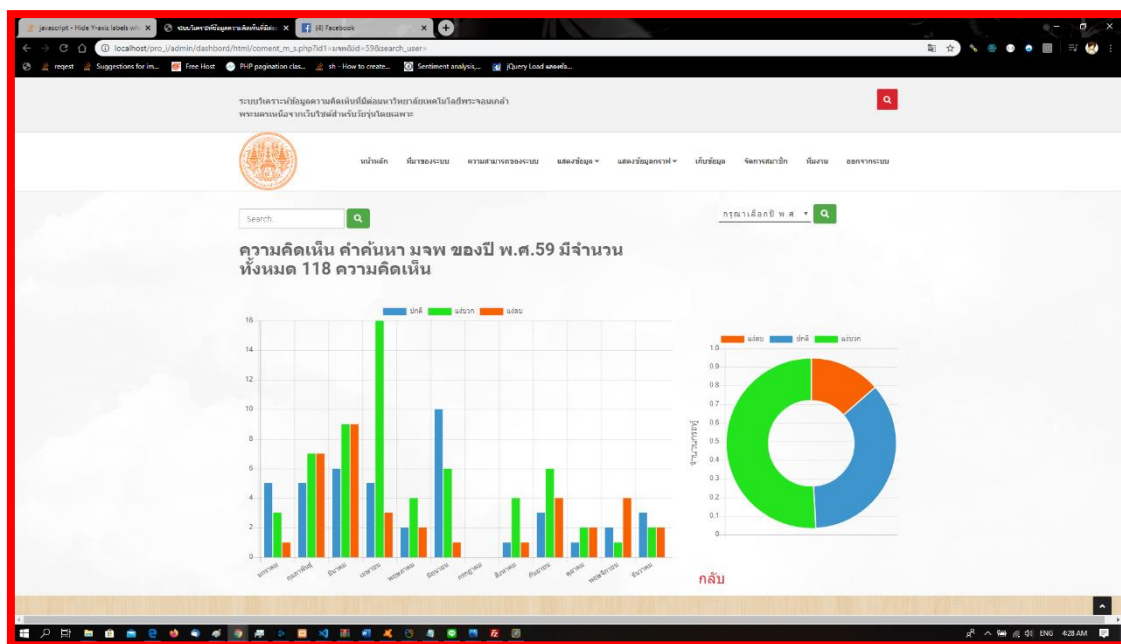
ภาพที่ 4-17 แสดงค่าความถูกต้องวิเคราะห์เนื้อหาการทูล

จากภาพที่ 4-17 แสดงค่าความถูกต้องวิเคราะห์เนื้อหาการทูลจะเป็นการแสดงค่าความถูกต้องที่ API ได้วิเคราะห์หัวข้อการทูลในรูปแบบกราฟแท่งและกราฟวงกลม โดยถ้าเป็นเชิงปกติ ค่าที่คืนกลับมาจะเป็น 0 นอกจากนี้ยังมีปุ่มกลับเพื่อกลับดูข้อมูลตารางการทูลที่เคยคลิกมาได้



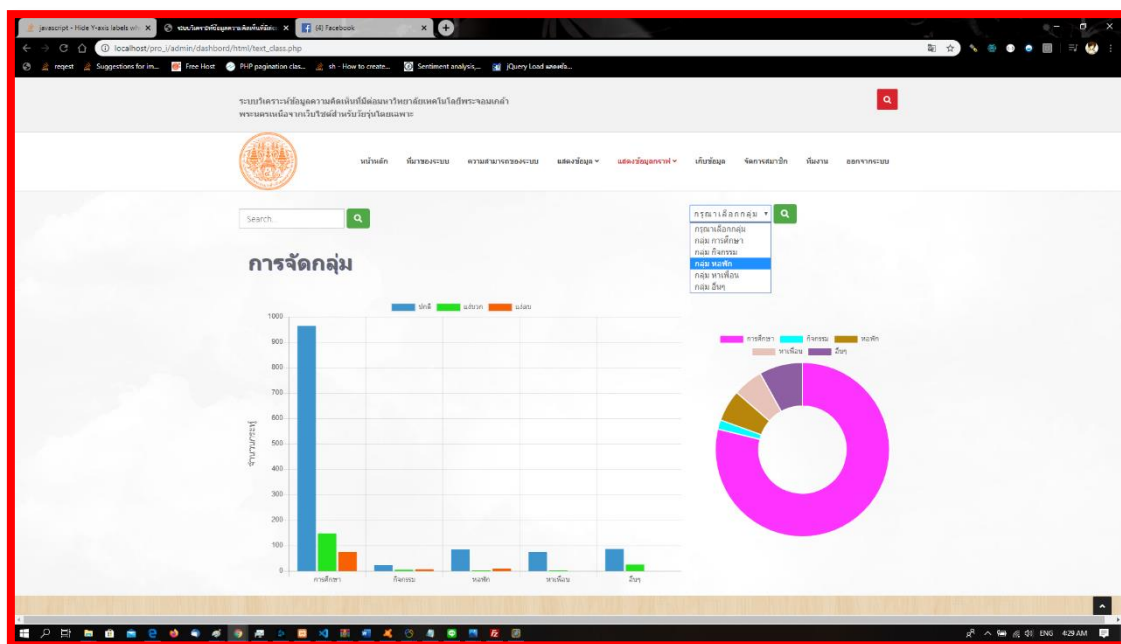
ภาพที่ 4-18 วิเคราะห์ความคิดเห็น

จากภาพที่ 4-18 จะเป็นการแสดงกราฟของวิเคราะห์ความคิดเห็นโดยกราฟเส้นนั้นจะเป็นการบอกจำนวนความถี่ของเนื้อหากระทู้ในแต่ละปีตามความรู้สึกที่โปรแกรมได้วิเคราะห์มาได้โดยเส้นสีฟ้าคือความรู้สึกปกติเส้น สีเขียวความรู้สึกในบวกเส้นสีแดงความรู้สึกในแง่ลบ ส่วนกราฟวงกลมจะบอกจำนวนรวมในแต่ละปี นอกจากนี้ยังมีปุ่ม “ดูประสิทธิภาพ” เพื่อดูประสิทธิภาพที่ API ได้วิเคราะห์มาได้



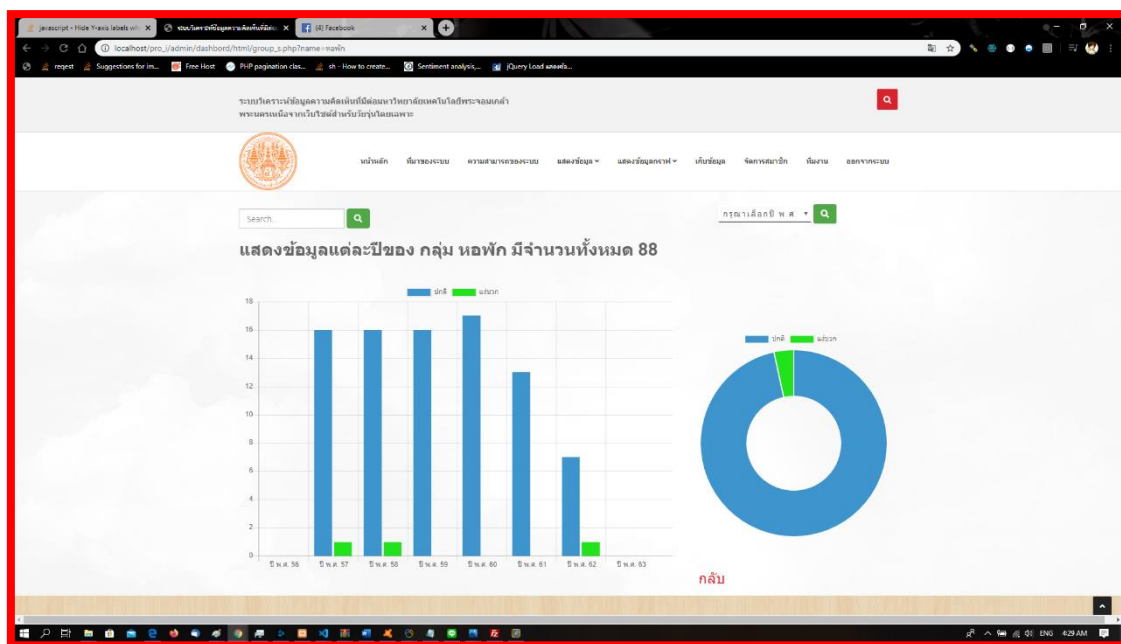
ภาพที่ 4-19 แสดงค่าความถูกต้องวิเคราะห์ความคิดเห็น

จากภาพที่ 4-19 แสดงค่าความถูกต้องวิเคราะห์ความคิดเห็นจะเป็นการแสดงความถูกต้องที่ API ได้วิเคราะห์หัวข้อกระทู้ในรูปแบบกราฟแท่งและกราฟวงกลม โดยถ้าเป็นเชิงปกติ ค่าที่คืนกลับมาจะเป็น 0 นอกจากนี้ยังมีปุ่มกลับเพื่อกลับดูข้อมูลตารางกระทู้ที่เคยคลิกมาได้



ภาพที่ 4-20 จัดกลุ่ม

จากภาพที่ 4-20 จะเป็นการแสดงกราฟของการจัดกลุ่มโดยกราฟแท่งนั้นจะเป็นการบอกจำนวนความถี่ของเนื้อหากระตุ้ในแต่ละกลุ่มตามความรู้สึกที่โปรแกรมได้วิเคราะห์มาได้โดยแท่งสีฟ้าคือความรู้สึกปกติแท่งสีเขียวความรู้สึกในบวกแท่งสีแดงความรู้สึกในแง่ลบ ส่วนกราฟวงกลมจะบอกจำนวนรวมในแต่ละปี นอกจากนี้ยังมีปุ่ม “ดูประสิทธิภาพ” เพื่อดูประสิทธิภาพที่ API ได้วิเคราะห์มา



ภาพที่ 4-21 ค่าความถูกต้องการจัดกลุ่ม

จากภาพที่ 4-21 แสดงค่าความถูกต้องวิเคราะห์ความคิดเห็นจะเป็นการแสดงค่าความถูกต้องที่ API ได้วิเคราะห์หัวข้อกระทู้ในรูปแบบกราฟแท่งและกราฟวงกลม โดยถ้าเป็นเชิงปกติ นอกจากนี้ยังมีปุ่มกลับเพื่อกลับดูข้อมูลตารางกระทู้ที่เคยคลิกมาได้

บทที่ 5

สรุปผลการดำเนินการและข้อเสนอแนะ

5.1 สรุปผลการดำเนินการ

ผลจากการดำเนินงานระบบวิเคราะห์ข้อมูลความคิดเห็นที่มีต่อมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือจากเว็บไซต์สำหรับวัยรุ่นโดยเฉพาะพบว่าสามารถเก็บรวบรวมข้อมูลจากเว็บสำหรับวัยรุ่นโดยเฉพาะได้เป็นจำนวนมากโดยจำนวนกระทู้ได้เก็บรวบรวมทั้งสิ้นเป็น จำนวน 1,232 กระทู้ และ ความคิดเห็นมีทั้งหมด 3,705 ความคิดเห็น ในการวิเคราะห์หัวข้อกระทู้พบว่าหัวข้อกระทู้ ความรู้สึกในเชิงบวกจำนวน 143 กระทู้ ในเชิงลบ 78 กระทู้ และ ปกติ 1,011 กระทู้ การวิเคราะห์เนื้อหากระทู้พบว่าเนื้อหากระทู้ความรู้สึกในเชิงบวกจำนวน 473 ในเชิงลบ 225 และ ปกติ 543 การวิเคราะห์ความคิดเห็นพบว่ามีความคิดเห็นความรู้สึกในเชิงบวกจำนวน 501 ความคิดเห็น ในเชิงลบ 1234 ความคิดเห็น และ ปกติ 1942 ความคิดเห็น และการจัดกลุ่มพบว่ามีการศึกษา 693 กระทู้ หอพัก 50 กระทู้ หาเพื่อน 53 กระทู้ กิจกรรม 13 กระทู้ อื่นๆ 423 กระทู้ โดยข้อมูลที่เก็บรวบรวมนี้สามารถแสดงบนหน้าจอเว็บแอปพลิเคชันได้อย่างถูกต้อง โดยผู้จัดทำได้จัดเก็บ ณ วันที่ 1 กรกฎาคม 2562 ถึง 10 กุมภาพันธ์ 2563

โปรแกรมแสดงข้อมูลความคิดเห็นที่มีต่อมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือจากเว็บไซต์สำหรับวัยรุ่นโดยเฉพาะ มีความสามารถใช้งานได้ดังต่อไปนี้

- 5.1.1 สามารถแสดงข้อมูลบอร์ดข้อความได้
- 5.1.2 สามารถติดตามเข้าไปยังบอร์ดข้อความได้
- 5.1.3 สามารถแสดงเวลา วัน เดือน ปี ของบอร์ดข้อความได้
- 5.1.4 สามารถแสดง จำนวนความคิดเห็นได้
- 5.1.5 สามารถแสดงจำนวนบอร์ดข้อความได้
- 5.1.6 สามารถค้นหาบอร์ดข้อความที่ต้องการในระบบจากชื่อของบอร์ดข้อความได้
- 5.1.7 แสดง 10 อันดับแรกความคิดเห็นเชิงบวกและเชิงลบในรูปแบบกราฟได้
- 5.1.8 แสดงค่าความรู้สึกในรูปแบบของภาพได้
- 5.1.9 แสดงกราฟสถิติของการวิเคราะห์ความคิดเห็นในแต่ละเดือนได้

5.2 ปัญหา อุปสรรค และการแก้ไข

จากการศึกษาความเป็นไปได้ในการทำโครงการพิเศษ จนกระทั่งได้พัฒนาโปรแกรมจนเสร็จถึงเป้าหมายแล้ว ได้มีปัญหและอุปสรรคเกิดขึ้นดังต่อไปนี้

5.2.1 การออกแบบหน้าจอโปรแกรมให้สะดวกและความสวยงามเหมาะแก่การใช้งานของผู้ใช้

5.2.2 การออกแบบการทำงานให้มีความรวดเร็วและเหมาะสมกับการใช้งาน โดยต้องคำนึงถึงข้อมูลที่นำมาแสดง เพราะมีจำนวนมาก

5.2.3 การวางแผนและออกแบบและแบ่งงาน เป็นสิ่งสำคัญมาก เนื่องจากเวลาในการทำโครงการพิเศษมีเวลาที่จำกัด

5.3 การแก้ปัญหา

5.3.1 ได้มีศึกษาการออกแบบเว็บไซต์การใช้งาน CSS และ Javascript และได้นำเอา และได้มีการนำเอา bootstrap มาเพื่อช่วยเพิ่มความสวยงามให้กับเว็บไซต์

5.3.2 หลีกเลี่ยงการใช้ `Select * from table` และ ใช้ `Select field_name from table` เพื่อลดปริมาณการโหลด Data ผ่าน Network และ IO ช่วยให้ Query ได้รวดเร็วยิ่งขึ้น อีกทั้งยังนำเอา index มาใช้งานเพื่อเพิ่มความรวดเร็วในการค้นหาข้อมูลยิ่งขึ้นไปอีก

5.3.2 ผู้พัฒนาจึงมีการวางแผนการทำงานในโครงการพิเศษโดยการสร้าง บันทึกตารางงานโครงการและจัดสรรเวลาให้เพียงพอต่อการทำงานในครั้งนี้

5.4 ข้อเสนอแนะ

5.4.1 ควรศึกษาค้นคว้าข้อมูลในอินเทอร์เน็ตและคำแนะนำที่เกี่ยวข้องกับการพัฒนาระบบอย่างละเอียด

5.4.2 โปรแกรมมีความซับซ้อนมากและมีการจัดเก็บข้อมูลจำนวนมาก ในอนาคตอาจจะต้องมีการสร้างฐานข้อมูลขึ้นมาใหม่เพื่อเพิ่มประสิทธิภาพในการเก็บข้อมูลและรองรับจำนวนข้อมูลที่มากขึ้น

บรรณานุกรม

Bigdataexperience .(2559). Big data analytics สำคัญยังไงและช่วยอะไรเราได้บ้าง?.
[ออนไลน์]. [สืบค้นวันที่ 5 กรกฎาคม 2562]. จาก <http://bigdataexperience.org/what-is-big-data-analytics/>

helponline.efinancethai .(2559) .Dashboard คืออะไร.[ออนไลน์].
[สืบค้นวันที่ 5 กรกฎาคม 2562].จากhttp://helponline.efinancethai.com/index.php?option=com_content&view=article&id=1261

Wikipedia .(2553).เด็กดีดอตคอม[ออนไลน์].[สืบค้นวันที่ 5 กรกฎาคม 2562].
จาก <https://th.wikipedia.org/wiki/เด็กดี.คอม>

xn—12 .(2561). Social Network คืออะไร .[ออนไลน์].[สืบค้นวันที่ 5 กรกฎาคม 2562].
จาก <https://www.xn--12cg1cxchd0a2gzc1c5d5a.net/social-network/>

Aosof .(2561). Python คืออะไร .[ออนไลน์].[สืบค้นวันที่ 5 กรกฎาคม 2562].
จาก <https://www.aosoft.co.th/article/322/Python/>

beyourcyber.(2561). node.js คืออะไร [ออนไลน์].[สืบค้นวันที่ 5 กรกฎาคม 2562].
จาก <https://beyourcyber.com/2016/node-js-is-programming-language-by-javascript/>

pasaphp.blogspot.(2561). ภาษา php [ออนไลน์].[สืบค้นวันที่ 6 กรกฎาคม 2562].
จาก <http://pasaphp.blogspot.com/>

softmelt.(2558). รู้จักกับ SQL, NoSQL และ NewSQL.[ออนไลน์]
[สืบค้นวันที่ 6 กรกฎาคม 2562]. จาก <https://www.softmelt.com/>

article.php?id=57mindphp.(2560). API คืออะไร.[ออนไลน์].[สืบค้นวันที่ 6 กรกฎาคม 2562].
จาก <https://www.mindphp.com/คู่มือ/73-คืออะไร/2038-api-คืออะไร.html>

บรรณานุกรม

ปราชญ์ภาคย์ เหล่าสังข์สุข.(2561).การวิเคราะห์ความคิดเห็นเกี่ยวกับร้านอาหารบนเว็บไซต์รีวิว [ออนไลน์].[สืบค้นวันที่ 6 กรกฎาคม 2562]. จาก https://www.kmutt.ac.th/jif/public_html/article_detail.php?ArticleID=210442

รวิศดา เทศเมือง และ นิเวศ จิระวิจิตชัยน(2561).การวิเคราะห์ความคิดเห็นภาษาไทยเกี่ยวกับการรีวิวสินค้าออนไลน์[ออนไลน์].[สืบค้นวันที่ 6 กรกฎาคม 2562]. จาก <https://www.tcithaijo.org/index.php/tsujournal/article/download/90081/70809/>

กานดา แผ้ววัฒนากุล.(2561).การวิเคราะห์เหมือนความคิดเห็นบนเครือข่ายสังคมออนไลน์ [ออนไลน์].[สืบค้นวันที่ 6 กรกฎาคม 2562]. จาก <https://www.tci-thaijo.org/index.php/stou-sms-pr/article/view/16934/15262>

wikipedia (2557).การทำเหมืองข้อความ[ออนไลน์].[สืบค้นวันที่ 6 กรกฎาคม 2562]. จาก <https://th.wikipedia.org/wiki/การทำเหมืองข้อความ>

sysadmin.psu.(2561). Puppeteer ควบคุมและบันทึกการใช้งาน chrome สำหรับ UI testing [ออนไลน์].[สืบค้นวันที่ 6 กรกฎาคม 2562]. จาก <https://sysadmin.psu.ac.th/2018/07/08/puppeteer-ควบคุมและบันทึกการ>