Science and Technology
RMUTT Journal
Vol.6 No.1 (2016) : 102-109
www.sci.rmutt.ac.th/stj
Online ISSN 2229-1547

การจำแนกจดหมายอิเล็กทรอนิกส์ที่เป็นสแปมโดยใช้เทคนิคเหมืองข้อมูล

Email Spam Classification Using Data Mining Techniques

จุฑาทิพย์ ทิพย์พูล[1] และ นิเวศ จิระวิชิตชัย[1*]

[1]หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาระบบสารสนเทศคอมพิวเตอร์
คณะเทคโนโลยีสารสนเทศ มหาวิทยาลัยศรีปทุม 61 ถนนพหลโยธิน เขตจตุจักร กรุงเทพมหานคร 10900
*E-mail: nivet.ch@spu.ac.th

บทคัดย่อ

ปัจจุบันการสื่อสารผ่านทางจดหมายอิเล็กทรอนิกส์ (e-mail) เป็นอีกหนึ่งช่องทางที่ได้รับความนิยม เนื่องจากมีการใช้งานที่ง่าย สะดวก รวดเร็ว และไม่เสียค่าใช้จ่าย ด้วยเหตุนี้จึงเป็นช่องทางให้ผู้ไม่ประสงค์ดี ใช้ในการส่งจดหมายอิเล็กทรอนิกส์ที่เป็น สแปม เพื่อจุดประสงค์อื่น เช่น การโฆษณา หรือการหลอกลวง บทความนี้นำเสนอการศึกษาการเปรียบเทียบประสิทธิภาพการจำแนกอีเมล์ที่เป็นสแปมโดยใช้เทคนิคการทำเหมืองข้อมูล ซึ่งประกอบด้วยต้นไม้ตัดสินใจ (Decision tree) นาอีฟนาอีฟเบย์ (Naive Bayes) และเคเนียเรสเนเบอร์ (K-Nearest Neighbor) ผลการทดลองเมื่อเปรียบเทียบจากความค่าถูกต้อง (Accuracy) พบว่าวิธีนาอีฟนาอีฟเบย์ให้ค่าความถูกต้องสูงสุดเท่ากับ 92.48 %

คำสำคัญ: จดหมายอิเล็กทรอนิกส์สแปม เหมืองข้อมูล

Abstract

Today the communication by electronic mail (e-mail) is another channel that has been popular because it is simple, easy to use, fast and free of charge. This is why it is the way that others use to send e-mail spam, for the purpose of advertising or deception. This article presents a study comparison the performance of electronic mail classification with data mining technique such as Decision tree, Naive Bayes,

Received: December 11, 2015
Revised: April 22, 2016
Accepted: April 29, 2016

and K-Nearest Neighbor. The results of this research found naive bayes algorithm provides the highest accuracy of 92.48%.

Keywords: Spam Classification, Data Mining

1. บทนำ

การสื่อสารผ่านทางจดหมายอิเล็กทรอนิกส์เป็นช่องทางติดต่อสื่อสารที่สำคัญ เนื่องจากปัจจุบันจำนวนประชากรที่สามารถเข้าถึงการให้บริการอินเทอร์เน็ต (Internet) นั้นมีมากขึ้นอย่างต่อเนื่อง สาเหตุจากปัจจัยหลายด้าน เช่น อุปกรณ์คอมพิวเตอร์และค่าบริการอินเทอร์เน็ตนั้นมีราคาถูก ด้วยเหตุนี้จึงเป็นช่องทางให้ผู้ไม่ประสงค์ดีใช้ในการส่งการจำแนกจดหมายอิเล็กทรอนิกส์ที่เป็นสแปม เพื่อจุดประสงค์บางอย่าง เช่น การโฆษณา จึงสร้างความรำคาญให้ผู้ใชเป็นอย่างมาก จากปัญหาข้างต้นจึงได้มีแนวคิดที่จะจำแนกการจำแนกจดหมายอิเล็กทรอนิกส์ที่เป็นสแปมหรือไม่ โดยการศึกษาการเปรียบเทียบประสิทธิภาพ โดยเลือกเฉพาะขั้นตอนวิธีที่เป็นที่นิยมใช้กับข้อมูลประเภทข้อความ คือ วิธีการต้นไม้ตัดสินใจ วิธีการของนาอีฟเบย์ และวิธีการค้นหาเพื่อนบ้านใกล้ที่สุด k ตัว ผ่านโปรแกรมสำเร็จรูป RapidMiner Studio 6.4

2. วิธีดำเนินการวิจัย

การศึกษาเรื่อง การเปรียบเทียบประสิทธิภาพการจำแนกจดหมายอิเล็กทรอนิกส์ที่เป็นสแปมโดยใช้วิธีการต้นไม้ตัดสินใจ วิธีการของนาอีฟเบย์ และวิธีการค้นหาเพื่อนบ้านใกล้ที่สุด k ตัว โดยใช้เทคโนโลยีการทำเหมืองข้อมูล เริ่มจากการสมัครสมาชิกในเว็บไซต์ เป้าหมาย เช่น เว็บไซต์เกี่ยวกับการขายสินค้า เว็บไซต์เกี่ยวกับการสมัครงาน และเว็บไซต์เกี่ยวกับการใช้ฟรีโปรแกรม เป็นต้น หลังจากได้รับจดหมายอิเล็กทรอนิกส์ที่เป็นสแปมแล้วจึงทำให้อยู่ในรูปแบบของ .txt ไฟล์ แล้วจึงนำเข้าประมวลผลในโปรแกรมสำเร็จรูป RapidMiner Studio 6.4 ต่อไป

2.1 การศึกษาทฤษฎีวิธีการของการทำเหมืองข้อมูล (Data mining Algorithms)

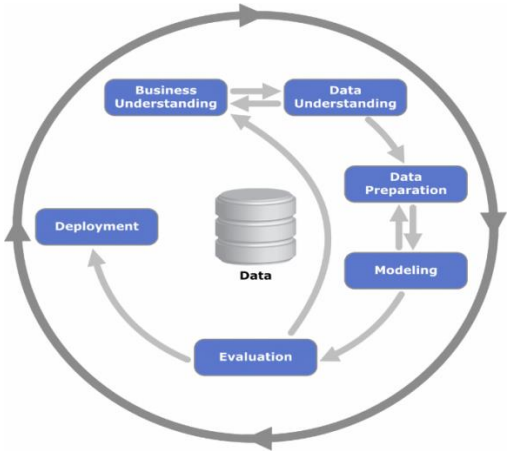
การศึกษาค้นคว้า เรื่องการทำเหมืองข้อมูล (data mining) [2-3] เป็นเทคนิค เพื่อค้นหารูปแบบ (pattern) ความสัมพันธ์ของข้อมูลจำนวนมหาศาลที่ซ่อนอยู่ในชุดข้อมูลนั้นแบบอัตโนมัติ โดยอาศัยหลักการทางสถิติ และหลักคณิตศาสตร์ โดยสนใจเฉพาะเทคนิคของการจำแนกประเภทข้อมูล (Data Classification)

การทำเหมืองข้อมูลในครั้งนี้มุ่งเน้นศึกษาวิธีการจำแนกประเภทข้อมูล (Data Classification) คือ กระบวนการสร้างแบบจำลอง (model) เพื่อจัดการข้อมูลให้อยู่ในกลุ่มที่กำหนด เช่น จัดกลุ่มประเภทนักเรียนดีมากดี ปานกลางและไม่ดี โดยพิจารณาจากประวัติและผลการรียน หรือจัดกลุ่มประเภทของลูกค้าว่าน่าเชื่อถือหรือไม่ โดยพิจารณาจากข้อมูลที่มีอยู่

การสร้างแบบจำลองการจำแนกประเภทข้อมูล (Data Classification model) จะเกิดขึ้นมาจากการหาความสัมพันธ์ของข้อมูลในฐานข้อมูลขนาดใหญ่ โดยข้อมูลทั้งหมดจะมีการแบ่งออกเป็น 2 กลุ่ม คือ กลุ่มข้อมูลเรียนรู้ (Training set) เป็นชุดข้อมูลที่มีบทบาทในการสร้างแบบจำลองการจำแนกประเภทข้อมูลขึ้นมา และกลุ่มข้อมูลทดสอบ (Test set) เป็นชุดข้อมูลประเมินความถูกต้องของแบบจำลองการจำแนกประเภทข้อมูล

แบบจำลองการจำแนกประเภทข้อมูลได้ถูกนำมาประยุกต์ใช้งานอย่างหลากหลาย เช่น การวิเคราะห์หุ้นในแต่ละบริษัทมีแนวโน้มเป็นอย่างไร นำมาวิเคราะห์โดยมีปัจจัยที่เกี่ยวข้อง เช่น การเติบโตของรายได้ ความสามารถในการควบคุมต้นทุน ความผันแปรของรายได้และกำไร และผู้บริหาร หรือการพยากรณ์อากาศ การจัดการความสัมพันธ์ของลูกค้า (Customer relationship management) และอื่นๆ

ขั้นตอนการทำเหมืองข้อมูลแบบ CRISP-DM ประกอบด้วย 6 ขั้นตอนดังแสดงในรูปที่ 2 โดยแต่ละขั้นตอนจะเป็นขั้นตอนที่ต่อเนื่องกัน คือ ขั้นตอนต่อไปต้องมีผลลัพธ์จากขั้นตอนก่อนหน้าจากด้วยลูกศรที่เชื่อมระหว่างแต่ละขั้นตอน เช่น จากผลลัพธ์จากขั้นตอนการเตรียมข้อมูล (Data Preparation) แล้วสารมารถสร้างโมเดลจำแนกประเภทข้อมูลในขั้นตอน Modeling และอาจทำซ้ำเพื่อเข้มูลที่ถูกต้องมากที่สุด



รูปที่ 1 แสดงขั้นตอนในกระบวนการ CRISP-DM

- Business Understanding การเข้าใจปัญหาและทำการวิเคราะห์เหมืองข้อมูล กล่าวคือ การเข้าใจจุดประสงค์ของ SMS ที่เป็นสแปม เช่น การโฆษณา การสมัครงาน และการหลอกเพื่อหวังข้อมูลส่วนบุคคล เป็นต้น

- Data Understanding การเก็บรวบรวมข้อมูล และนำข้อมูลที่ได้มาตรวจสอบดูความถูกต้อง

- Data Preparation ทำการแปลงข้อมูลดิบ (raw data) เป็นข้อมูลที่สามารถวิเคราะห์ได้ โดยการแปลงข้อมูลเป็นรูปแบบที่ถูกต้อง (data cleaning) เช่น การแปลงข้อมูลเป็นมาตราฐานเดียวกัน หรือการกำจัดข้อมูลว่าง (null value) เป็นต้น

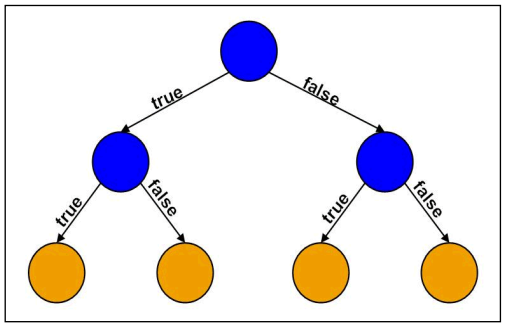
- Modeling ทำการวิเคราะห์ข้อมูลด้วยวิธีการเหมืองข้อมูล ในที่นี้ได้เลือกมา 3 วิธี คือ วิธีการต้นไม้ตัดสินใจ วิธีการของนาอีฟเบย์ และวิธีการค้นหาเพื่อนบ้านใกล้ที่สุดเคตัว เพราะเหมาะสมกับการจำแนกประเภทข้อมูลประเภทตัวอักษร (Text Classification)

- Evaluation คือ การประเมินประสิทธิภาพของผลลัพธ์จากแบบจำลองข้างต้นว่าครอบคลุมและสามารถให้ผลลัพธ์ตามเป้าหมายที่ตั้งไว้ในขั้นตอนแรกหรือไม่

- Deployment นำผลลัพธ์ หรือองค์ความรู้ที่ได้มาใช้งานต่อไป

2.1.1 วิธีการต้นไม้ตัดสินใจ

วิธีการต้นไม้ตัดสินใจ เป็นโครงสร้างข้อมูลชนิดเป็นลำดับชั้น (hierarchy) ใช้สนับสนุนการตัดสินใจ โดยจะมีลักษณะคล้ายต้นไม้กลับหัวที่มีโหนดราก (root node) ในตำแหน่งด้านบน และโหนดใบ (leaf node) ในตำแหน่งด้านล่างของต้นไม้



รูปที่ 2 วิธีการต้นไม้ตัดสินใจ

ภายในต้นไม้จะประกอบไปด้วยโหนดต่างๆ ซึ่งแต่ละโหนดจะมีคุณลักษณะ (attribute) เป็นตัวทดสอบ กิ่งของต้นไม้ (branch) แสดงถึงค่าที่เป็นไปได้ของคุณลักษณะที่ถูกเลือกทดสอบ และใบ ซึ่งเป็นสิ่งที่อยู่ล่างสุดของต้นไม้ตัดสินใจแสดงถึงกลุ่มของข้อมูล (class) ก็คือผลลัพธ์ที่ได้จากการทำนาย โหนดที่อยู่บนสุดของต้นไม้เรียกว่าโหนดราก [4-5]

2.1.2 วิธีการของนาอีฟเบย์

วิธีการของนาอีฟเบย์เป็นวิธีการรียนรู้ที่ใช้หลักของความน่าจะเป็น ตามกฎทฤษฎีของนาอีฟเบย์เข้ามาช่วยในการเรียนรู้ เพื่อหาสมมติฐานหนึ่งๆ ร่วมกับข้อมูล การเรียนรู้แบบนาอีฟเบย์ อาศัยหลักการของการคำนวณความน่าจะเป็นของแต่ละสมมติฐาน โดยการเรียนรู้แบบนาอีฟเบย์เป็นการเรียนรู้เพิ่มได้ เนื่องจากตัวอย่างใหม่ที่ได้มาถูกนำมาปรับเปลี่ยนการแจกแจง ซึ่งมีผลต่อการเพิ่ม หรือลดความน่าจะเป็น ทำให้มีการเรียนรู้ที่เปลี่ยนไป วิธีการนี้ตัวแบบจะถูกปรับเปลี่ยนไปตามตัวอย่างใหม่ที่ได้โดยผนวกกับความรู้เดิมที่มี ซึ่งการทำนายค่าคลาสเป้าหมายของตัวอย่างใช้ความน่าจะเป็นมากที่สุดของทุกสมมติฐาน [4-5]

2.1.3 วิธีการค้นหาเพื่อนบ้านใกล้ที่สุด k ตัว

วิธีการค้นหาเพื่อนบ้านใกล้ที่สุด k ตัว เป็นวิธีที่ใช้ในการจัดแบ่งคลาส โดยวิธีการนี้จะตัดสินใจว่า คลาสใดที่จะแทนเงื่อนไข หรือกรณีใหม่ๆ ได้บ้าง โดยการตรวจสอบจำนวนบางจำนวนของกรณี หรือเงื่อนไขที่เหมือนกัน หรือใกล้เคียงกันมากที่สุด โดยจะหาผลรวม (Count Up) ของจำนวนเงื่อนไข หรือกรณีต่างๆ สำหรับแต่ละคลาส และกำหนดเงื่อนไขใหม่ๆ ให้คลาสที่เหมือนกันกับคลาสที่ใกล้เคียงกันมากที่สุด [4-5]

2.2 การเลือกใช้โปรแกรมสำเร็จรูป

RapidMiner [1] เป็นแพลตฟอร์มซอฟต์แวร์ที่พัฒนา โดยบริษัทที่มีชื่อเดียวกัน เป็นเครื่องมือสำหรับวิเคราะห์ การเรียนรู้ของเครื่อง (machine learning) การทำเหมืองข้อมูล (data mining) การทำเหมืองข้อความ (text mining) การวิเคราะห์เชิงพยากรณ์ (predictive analytics) และการวิเคราะห์เชิงธุรกิจ (business analytics) สร้างเพื่อใช้สำหรับธุรกิจ การประยุกต์ใช้ในอุตสาหกรรม การวิจัยการศึกษาการฝึกอบรม และการพัฒนาแอพลิเคชัน สนับสนุนทุกขั้นตอนของกระบวนการทำเหมืองข้อมูลรวมทั้งการแสดงผล การตรวจสอบและการเพิ่มประสิทธิภาพ เวอร์ชั่น RapidMiner ที่ใช้ คือ RapidMiner Studio 6.4 ซึ่งเป็นฟรีเวอร์ชั่นและใช้งานง่ายสำหรับผู้ที่เริ่มใช้เบื้องต้นจนถึงขั้นระดับสูง เนื่องจากมีโอเปอเรชั่นให้เลือกใช้หลากหลายและอยู่ในรูปแบบลากและวาง (drag and drop)

2.3 การออกแบบระบบ

การออกแบบระบบการจำแนกจดหมายอิเล็กทรอนิกส์ที่เป็นสแปม โดยใช้โปรแกรม RapidMiner Studio 6.4 เพื่อใช้ในการวิเคราะห์ข้อมูลต่างๆ ใช้ โอเปอเรเตอร์ [4] ดังต่อไปนี้

2.3.1 Process Document From Files (Text Processing)

- Tokenize
- Filter Tokens (by Length)

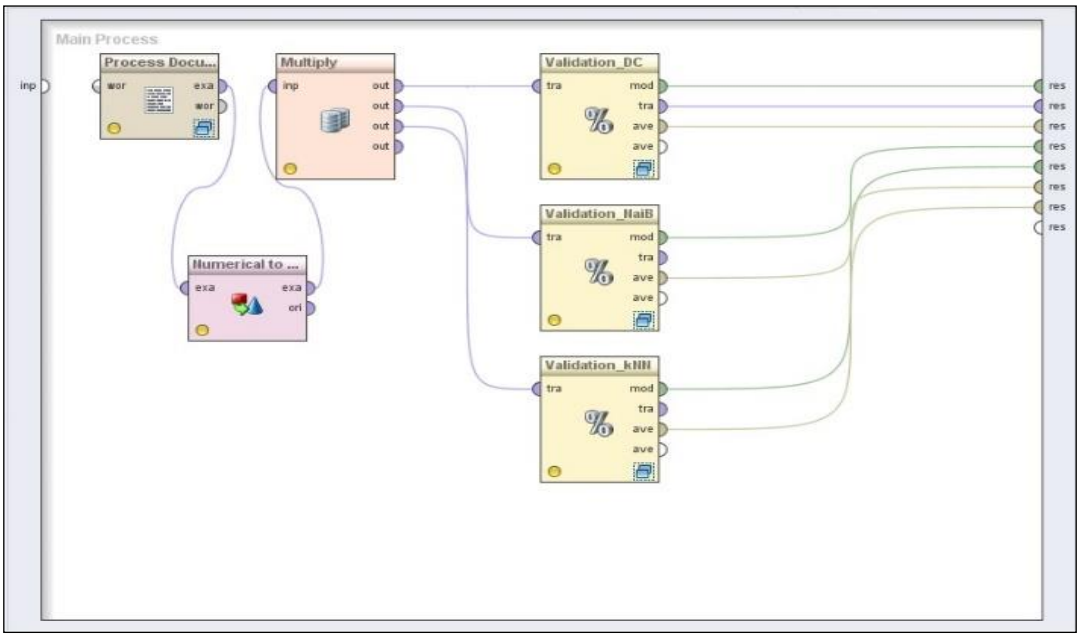
- Stem (poter)
- Filter stopwords (English)

2.3.2 Numerical to Binominal

2.3.3 Multiply (RapidMiner Studio Core)

2.3.4 X-Validation (RapidMiner Studio Core)

- Validation Naive Bayes
- Validation k-NN
- Validation Decition Tree



รูปที่ 3 แสดงการออกแบบระบบการเปรียบเทียบประสิทธิภาพการจำแนกจดหมายอิเล็กทรอนิกส์ที่เป็นสแปม

3. ผลการวิจัยและวิจารณ์ผลการวิจัย

การรายงานผลประสิทธิภาพการจำแนกจดหมายอิเล็กทรอนิกส์ที่เป็นสแปม รายละเอียด ดังนี้

3.1 วิธีการต้นไม้ตัดสินใจ (Decision tree)

Performance

accuracy: 81.84% +/- 4.03% (mikro: 81.84%)

precision: 84.02% +/- 9.60% (mikro: 82.39%) (positive class: nonspam)

recall: 94.66% +/- 7.10% (mikro: 94.66%) (positive class: nonspam)

f_measure: 88.25% +/- 1.73% (mikro: 88.10%) (positive class: nonspam)

	true spam	true nonspam	class precision
pred. spam	757	196	79.43%
pred. nonspam	743	3476	82.39%
class recall	50.47%	94.66%	

รูปที่ 4 การรายงานผลประสิทธิภาพการจำแนกจดหมายอิเล็กทรอนิกส์ที่เป็นสแปมโดยใช้วิธีการต้นไม้ตัดสินใจ

- ค่า true spam คือ ค่าที่ทำนายว่าเป็นสแปมจริง คือ 757 ข้อความ และทำนายผิด 743 ข้อความจาก 1,500 ข้อความ
- ค่า true nonspam คือ ค่าที่ทำนายว่าไม่เป็นสแปมจริง คือ 3,476 ข้อความ และทำนายผิด 196 ข้อความ จาก 3,672 ข้อความ

3.2 วิธีการของนาอีฟเบย์ (Naive Bayes)

Performance

accuracy: 92.48% +/- 1.22% (mikro: 92.48%)

precision: 94.99% +/- 1.43% (mikro: 94.96%) (positive class: nonspam)

recall: 94.42% +/- 1.40% (mikro: 94.42%) (positive class: nonspam)

f_measure: 94.69% +/- 0.86% (mikro: 94.69%) (positive class: nonspam)

	true spam	true nonspam	class precision
pred. spam	1316	205	86.52%
pred. nonspam	184	3467	94.96%
class recall	87.73%	94.42%	

รูปที่ 5 การรายงานผลประสิทธิภาพการจำแนกจดหมายอิเล็กทรอนิกส์ที่เป็นสแปมโดยใช้วิธีการของนาอีฟเบย์

- ค่า true spam คือ ค่าที่ทำนายว่าเป็นสแปมจริง คือ 1,316 ข้อความ และทำนายผิด 184 ข้อความจาก 1,500 ข้อความ
- ค่า true nonspam คือ ค่าที่ทำนายว่าไม่เป็นสแปมจริง คือ 3,467 ข้อความ และทำนายผิด 205ข้อความ จาก 3,672 ข้อความ

3.3 วิธีการค้นหาเพื่อนบ้านใกล้ที่สุด k ตัว (K-Nearest Neighbour)

Performance

accuracy: 86.68% +/- 1.09% (mikro: 86.68%)

precision: 98.66% +/- 0.51% (mikro: 98.66%) (positive class: nonspam)

recall: 82.35% +/- 1.36% (mikro: 82.35%) (positive class: nonspam)

f_measure: 89.77% +/- 0.90% (mikro: 89.77%) (positive class: nonspam)

	true spam	true nonspam	class precision
pred. spam	1459	648	69.25%
pred. nonspam	41	3024	98.66%
class recall	97.27%	82.35%	

รูปที่ 6 การรายงานผลประสิทธิภาพการจำแนกจดหมายอิเล็กทรอนิกส์ที่เป็นสแปมโดยใช้วิธีการค้นหาเพื่อบ้านใกล้ที่สุด k ตัว

- ค่า true spam คือ ค่าที่ทำนายว่าเป็นสแปมจริง คือ 1,459 ข้อความ และทำนายผิด 41 ข้อความจาก 1,500 ข้อความ
- ค่า true nonspam คือ ค่าที่ทำนายว่าไม่เป็นสแปมจริง คือ 3,024 ข้อความ และทำนายผิด 648ข้อความ จาก 3,672 ข้อความ

4. สรุปผลการวิจัย

จากการเก็บรวบรวมข้อมูลแบ่งเป็นจดหมายอิเล็กทรอนิกส์ที่เป็นสแปมจำนวน 1,500 ข้อความและจดหมายอิเล็กทรอนิกส์ที่ไม่เป็นสแปม จำนวน 3,672 ข้อความ รวมทั้งหมด 5,172 ข้อความ โดยมีคุณลักษณะ (attribute) จำนวน 920 คุณลักษณะ ซึ่งเกิดจากคำที่ไม่ซ้ำกันจำนวน 916 คำ และมีคุณลักษณะพิเศษ (special attribute) จำนวน 4 คุณลักษณะ ได้แก่ ฉลาก (label) ซึ่งบ่งบอกว่าเป็นจดหมายอิเล็กทรอนิกส์ที่เป็นสแปมหรือไม่ ชื่อไฟล์ (metadata file) ที่อยู่ของไฟล์ (metadata path) วันที่ได้รับจดหมายอิเล็กทรอนิกส์ (metadata date) ซึ่งใน 4 คุณลักษณะพิเศษนี้ มีเพียงคุณลักษณะฉลากเท่านั้นที่นำมาร่วมคำนวน ผลลัพธ์แสดงดังตาราง

ตารางที่ 1 แสดงการเปรียบเทียบผลสรุปค่าความถูกต้อง (accuracy) ค่าความแม่นยำ (Precision) ค่าความอ่อนไหว (recall) และค่าเฉลี่ยของความแม่นยำและค่าความอ่อนไหว (f measure) ของทั้ง 3 อัลกอริทึม

Performance	Decision tree (%)	Naive Bayes (%)	K-Nearest Neighbour (%)
accuracy	81.84	92.48	86.68
Precision	84.02	94.99	98.66
recall	94.66	94.42	82.35
f measure	88.25	94.69	89.77

จากตารางที่ 1 ได้สรุปการเปรียบเทียบประสิทธิภาพการจำแนกจดหมายอิเล็กทรอนิกส์ที่เป็นสแปมโดยใช้เทคนิคการทำเหมืองข้อมูล ซึ่งประกอบด้วยต้นไม้ตัดสินใจ (Decision tree) นาอีฟนาอีฟเบย์ (Naive Bayes) และเคเนียเรสเนเบอร์ (K-Nearest Neighbor) ผลการทดลองเมื่อเปรียบเทียบจากความค่าถูกต้อง (Accuracy) พบว่าวิธีนาอีฟนาอีฟเบย์ ให้ค่าความถูกต้องสูงสุดเท่ากับ 92.48 % ดังนั้นจึงสมารถสรุปได้ว่า แบบจำลองที่ได้จากการทดลองนี้สามารถวิเคราะห์และทำนายหาที่มาของจดหมายอิเล็กทรอนิกส์ที่เป็นสแปมได้ถูกต้องในระดับที่ยอมรับได้ และสามารถนำไปประยุกต์ใช้ ซึ่งจะเป็นการลดภาระของเครื่องแม่ข่าย (Server) ในองค์กรธุรกิจในการกรองจดหมายอิเล็กทรอนิกส์ที่เป็นสแปม ได้อย่างมีประสิทธิภาพ และสามารถนำทรัพยากรขององค์กรไปใช้ให้เกิดประโยชน์สูงสุดกับองค์กรต่อไป

5. เอกสารอ้างอิง

[1] Markus Hofmann, Ralf Klinkenberg. Rapidminer Data Mining Use Cases and Business Analytics Applications: CRC Press, 2013

[2] Charu C. Aggarwal. Data Classification Algorithms and Applications: CRC Press, 2014

[3] Rajan Chattamvelli . Data Mining Algorithms: Alpha Science International Ltd., 2011

[4] นิเวศ จิระวิชิตชัย, ปริญญา สงวนสัตย์, พยุง มีสัจ. การพัฒนาประสิทธิภาพการจัดหมวดหมู่เอกสารภาษาไทยอัตโนมัติ: วารสารพัฒนบริหารศาสตร์ สถาบันบัณฑิตพัฒนบริหารศาสตร์ ปี (2554)

[5] นิเวศ จิระวิชิตชัย. แบบจำลองการจำแนกเอกสารภาษาไทยอัตโนมัติ: วารสารวิชาการเทคโนโลยีอุตสาหกรรม มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ ปี (2555)

[6] นายเอกสิทธิ์ พัชรวงศ์ศักดา, Introduction to Business Analytics with RapidMiner Studio 6 (ฉบับภาษาไทย), กรุงเทพฯ: บริษัท เอเชีย ดิจิตอลการพิมพ์ จำกัด (2558)