การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษา

สำนักหอสมุดกลาง

โดย

นางสาวชนิสรา ลือวิพันธ์

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาเภสัชศาสตรมหาบัณฑิต

สาขาวิชาสารสนเทศศาสตร์ทางสุขภาพ

ภาควิชาสารสนเทศศาสตร์ทางสุขภาพ

บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร

ปีการศึกษา 2555

ลิขสิทธิ์ของบัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร

การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษา

สำนักหอสมุดกลาง

โดย

นางสาวชนิสรา ลือวิพันธ์

มหาวิทยาลัย ศิลปากร

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาเภสัชศาสตรมหาบัณฑิต

สาขาวิชาสารสนเทศศาสตร์ทางสุขภาพ

ภาควิชาสารสนเทศศาสตร์ทางสุขภาพ

บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร

ปีการศึกษา 2555

ลิขสิทธิ์ของบัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร

AUTOMATIC TEXT CATEGORIZATION FOR DRUG MONOGRAPH BASED ON PHARMACOLOGIC-THERAPEUTIC CLASSES

สำนักหอสมุดกลาง

By

Miss Chanisara Lueviphan

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree

Master of Pharmacy Program in Health Informatics

Department of Health-Related Informatics

Graduate School, Silpakorn University

Academic Year 2012

Copyright of Graduate School, Silpakorn University

บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร อนุมัติให้วิทยานิพนธ์เรื่อง “การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษา” เสนอโดย นางสาวชนิสรา ลือวิพันธ์ เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาเภสัชศาสตรมหาบัณฑิต สาขาวิชาสารสนเทศศาสตร์ทางสุขภาพ

.....

(ผู้ช่วยศาสตราจารย์ ดร. ปานใจ ธารทัศนวงศ์)

คณบดีบัณฑิตวิทยาลัย

วันที่.....เดือน.....พ.ศ.....

สำนักหอสมุดกลาง

อาจารย์ที่ปรึกษาวิทยานิพนธ์

เภสัชกร รองศาสตราจารย์ ดร. วีรยุทธ์ เลิศนที

คณะกรรมการตรวจสอบวิทยานิพนธ์

.....ประธานกรรมการ

(เภสัชกร ผู้ช่วยศาสตราจารย์ ดร. สรวง รุ่งประกายพรรณ)

...../...../.....

.....กรรมการ

(เภสัชกร รองศาสตราจารย์ ดร. วีรยุทธ์ เลิศนที)

...../...../.....

.....กรรมการ

(ดร. อิทธิพันธ์ เมธเศรษฐ)

...../...../.....

52363201 : สาขาวิชาสารสนเทศศาสตร์ทางสุขภาพ
คำสำคัญ : การจัดแบ่งประเภทเอกสารโดยอัตโนมัติ / การจัดแบ่งประเภทเอกสารชนิดหนึ่งหมวดหมู่ / การจัดแบ่งประเภทเอกสารชนิดหลายหมวดหมู่ / เอกสารข้อมูลยา

ชนิสรา ลือวิพันธ์ : การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษา. อาจารย์ที่ปรึกษาวิทยานิพนธ์ : ภก.รศ.ดร. วีรยุทธ์ เลิศนที. 91 หน้า.

ปัจจุบันเอกสารข้อมูลยามีจำนวนเพิ่มขึ้นมาก การจัดแบ่งประเภทเอกสารโดยอัตโนมัติจึงมีบทบาทสำคัญในการจัดแบ่งประเภทเอกสารให้อยู่ในกลุ่มที่เหมาะสม เอกสารข้อมูลยาสามารถจัดแบ่งประเภทตามกลุ่มเภสัชวิทยาและการรักษาให้อยู่ตั้งแต่หนึ่งกลุ่มจนถึงหลายกลุ่ม ในการศึกษานี้นำเอกสารข้อมูลยามาจากเว็บไซต์ RxList และ DailyMed โดยใช้ตัวจัดแบ่งประเภท 3 วิธีคือ วิธีเบย์อย่างง่าย วิธีเพื่อนบ้านใกล้ที่สุด และวิธีแบบเซนทรอยด์ในการจัดแบ่งประเภทเอกสารชนิดหนึ่งกลุ่มและชนิดหลายกลุ่ม การจัดแบ่งประเภทแบบไบนารี่และการจัดแบ่งประเภทแบบครั้งละหลายกลุ่มถูกนำมาใช้ด้วย การจัดแบ่งประเภทเอกสารแบบ 2 ขั้นตอนถูกพัฒนาขึ้นสำหรับการจัดแบ่งประเภทเอกสารชนิดหลายกลุ่ม ในขั้นที่ 1 เอกสารข้อมูลยาจะถูกจัดแบ่งประเภทตามจำนวนกลุ่มที่เป็นไปได้ ผลจากขั้นที่ 1 จะถูกใช้ในการจัดแบ่งประเภทเอกสารข้อมูลยาตามกลุ่มเภสัชวิทยาและการรักษาของขั้นที่ 2 การจัดแบ่งประเภทแบบสมบูรณ์ของทั้ง 2 ขั้นตอนจะถูกใช้ในการประเมิน จากผลการทดลองวิธีเพื่อนบ้านใกล้ที่สุดเป็นตัวจัดแบ่งประเภทที่ดีที่สุดในการจัดแบ่งประเภทเอกสารข้อมูลยาตามจำนวนกลุ่ม ในขณะที่วิธีแบบเซนทรอยด์เป็นตัวจัดแบ่งประเภทที่ดีที่สุดในการจัดแบ่งประเภทเอกสารข้อมูลยาตามกลุ่มเภสัชวิทยาและการรักษา ซึ่งการจัดแบ่งประเภทแบบ 2 ขั้นตอนสามารถประยุกต์ใช้ในการจัดแบ่งประเภทเอกสารชนิดหลายกลุ่มได้ โดยจะมีประสิทธิภาพดีกว่าการจัดแบ่งประเภทเอกสารแบบดั้งเดิมสำหรับเอกสารชนิดหนึ่งกลุ่มและชนิดหลายกลุ่ม ดังเช่นการจัดแบ่งประเภทแบบไบนารี่

ภาควิชาสารสนเทศศาสตร์ทางสุขภาพ บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร

ลายมือชื่อนักศึกษา..... ปีการศึกษา 2555

ลายมือชื่ออาจารย์ที่ปรึกษาวิทยานิพนธ์.....

52363201: MAJOR: HEALTH INFORMATICS

KEY WORD: TEXT CATEGORIZATION, SINGLE-LABEL CLASSIFICATION, MULTI-LABEL CLASSIFICATION, DRUG MONOGRAPH CLASSIFICATION

CHANISARA LUEVIPHAN: AUTOMATIC TEXT CATEGORIZATION FOR DRUG MONOGRAPH BASED ON PHARMACOLOGIC-THERAPEUTIC CLASSES. THESIS ADVISOR: ASSOC. PROF. VERAYUTH LERTNATTEE, Ph.D. 91 pp.

With an increasing number of documents for drug information, automatic categorization is an important task for organizing these documents into suitable classes. Drug monographs can be classified into one or multiple classes based on Pharmacologic-Therapeutic classes. In this research, a set of drug monographs was drawn from the RxList and the DailyMed websites. Three classifiers, naïve Bayes, k-nearest neighbor, and centroid-based were used in single-label and multi-label classification. Binary classification and multiclass classification were used. Two-step text classification was developed for multi-label classification. The first step, drug monographs were classified into the number of possible classes. The results from the first step were applied on the second step based on their pharmacologic-therapeutic classes. The complete classification of the two-step was used for evaluation. From the experimental results, k-nearest neighbor was the best classifier to categorize drug monographs into possible classes while centroid-based classifier was the best classifier to categorize drug monographs based on pharmacologic-therapeutic classes. In conclusion, two-step text classification could be applied for multi-label classification. It was better performance than the classical method for single-label and multi-label documents, i.e., binary classification.

Department of Health-Related Informatics Graduate School, Silpakorn University

Student's signature..... Academic Year 2012

Thesis Advisor's signature.....

กิตติกรรมประกาศ

การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษา ได้รับการสนับสนุนจาก 2 ทุน คือทุนวิจัยในโครงการวิจัยเรื่อง “การพัฒนาตัวสืบค้นข้อมูลยาผ่านเครือข่ายอินเทอร์เน็ตสำหรับบุคลากรสาธารณสุขและบุคคลทั่วไป” SURDI 52/01/23 และ SURDI 53/01/12 จากสถาบันวิจัยและพัฒนา มหาวิทยาลัยศิลปากร และทุนอุดหนุนการทำวิจัย จากงบประมาณแผ่นดิน (หมวดเงินอุดหนุนทั่วไป) และเงินงบประมาณโครงการพิเศษของบัณฑิตวิทยาลัย (สมทบ) ประจำปีงบประมาณ 2554 จึงทำให้เกิดโครงการวิทยานิพนธ์นี้ ข้าพเจ้าขอขอบพระคุณผู้เกี่ยวข้องในหน่วยงานดังกล่าวทุกท่านที่ให้การสนับสนุน ประสานงานในทุกๆ ด้านจนกระทั่งการศึกษาในครั้งนี้สำเร็จลงด้วยดี

ข้าพเจ้าขอกราบขอบพระคุณ ภก.รศ.ดร. วีรยุทธ์ เลิศนที อาจารย์ที่ปรึกษาวิทยานิพนธ์ที่ได้กรุณาให้โอกาส ให้คำแนะนำ ให้กำลังใจ ช่วยเหลือตลอดจนแก้ไขข้อบกพร่องต่างๆ ในการทำวิทยานิพนธ์จนสำเร็จลงได้ ขอขอบพระคุณ ภก.รศ.ดร. สรวง รุ่งประกายพรรณ ประธานกรรมการตรวจสอบวิทยานิพนธ์ และดร. อิทธิพันธ์ เมธเศรษฐ กรรมการตรวจสอบวิทยานิพนธ์ที่กรุณาให้คำแนะนำตรวจทานแก้ไขวิทยานิพนธ์อย่างสมบูรณ์

สุดท้ายนี้ต้องขอบพระคุณคุณพ่อ น้องชาย คุณป้า พี่ๆ น้องๆ และเพื่อนๆ ทุกคนที่เป็นแรงผลักดันและเป็นกำลังใจให้สามารถทำวิทยานิพนธ์นี้สำเร็จลุล่วงไปด้วยดี

สารบัญ

หน้า

บทคัดย่อภาษาไทย.....ง
บทคัดย่อภาษาอังกฤษ.....จ
กิตติกรรมประกาศ.....ฉ
สารบัญตาราง.....ฌ
สารบัญรูป.....ฏ
บทที่
1 บทนำ.....1
ความเป็นมาและความสำคัญของปัญหา.....1
วัตถุประสงค์ของการวิจัย.....3
ขอบเขตของการวิจัย.....3
นิยามศัพท์เฉพาะ.....4
ประโยชน์ที่ได้รับ.....5
2 วรรณกรรมที่เกี่ยวข้อง.....6
การจัดแบ่งประเภทเอกสารโดยอัตโนมัติ.....6
เอกสารข้อมูลยา.....18
งานวิจัยที่เกี่ยวข้อง.....19
3 วิธีดำเนินการวิจัย.....27
การคัดเลือกเอกสารข้อมูลยา.....27
การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติ.....31
การประเมินผลการจัดแบ่งประเภทและประสิทธิภาพตัวจัดแบ่งประเภท.....35
4 ผลการวิเคราะห์ข้อมูล.....38
การประเมินตัวจัดแบ่งประเภทเพื่อการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติ
ตามกลุ่มเภสัชวิทยาและการรักษา.....38
การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษา

สำนักหอสมุดกลาง
จุฬาลงกรณ์มหาวิทยาลัย

บทที่ หน้า
โดยวิธีแบบไบนารี่.....48
การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษา
โดยวิธีแบบครั้งละหลายกลุ่ม.....52
การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษา
ชนิดหลายกลุ่ม.....54
5 สรุป อภิปรายผล และข้อเสนอแนะ.....70
สรุปผลการศึกษา.....70
อภิปรายผลการศึกษา.....71
ข้อเสนอแนะ.....73
รายการอ้างอิง.....75
ภาคผนวก.....78
ประวัติผู้วิจัย.....91

สำนักหอสมุดกลาง
มหาวิทยาลัย ศิลปากร

สารบัญตาราง

ตารางที่ หน้า

2.1 Confusion matrix ของผลการจัดแบ่งประเภทเอกสารแบบครั้งละสองกลุ่ม.....14

2.2 Confusion matrix ของผลการจัดแบ่งประเภทเอกสารแบบครั้งละหลายกลุ่ม.....16

3.1 กลุ่มเภสัชวิทยาและการรักษาจากหนังสือ AHFS Drug Information 2012 จำนวน 11 กลุ่มจากเอกสารข้อมูลยาคัดกรองได้.....29

3.2 จำนวนเอกสารข้อมูลยาตามกลุ่มเภสัชวิทยาและการรักษาชนิด 1 กลุ่ม.....29

3.3 จำนวนเอกสารข้อมูลยาตามกลุ่มเภสัชวิทยาและการรักษาชนิด 1 และ 2 กลุ่ม.....30

3.4 จำนวนเอกสารข้อมูลยาตามกลุ่มเภสัชวิทยาและการรักษาชนิด 1, 2 และ 3 กลุ่ม.....30

3.5 จำนวนเอกสารข้อมูลยาในเอกสารข้อมูลยาชนิด 1, 2 และ 3 กลุ่ม.....31

4.1 ผลการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบครั้งละหลายกลุ่มด้วยวิธีเบย์อย่างง่าย.....40

4.2 ผลการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบครั้งละหลายกลุ่มด้วยวิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k=3....41

4.3 ผลการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบครั้งละหลายกลุ่มด้วยวิธีแบบเซนทรอยด์.....42

4.4 การเปรียบเทียบค่าความถูกต้องของการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีครั้งละหลายกลุ่มด้วยวิธีเบย์อย่างง่าย วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k = 3 และวิธีแบบเซนทรอยด์...43

4.5 ผลการประเมินการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบครั้งละหลายกลุ่มด้วยวิธีเบย์อย่างง่าย.....44

4.6 ผลการประเมินการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบครั้งละหลายกลุ่มด้วยวิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k=3.....45

4.7 ผลการประเมินการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบครั้งละหลายกลุ่มด้วยวิธีแบบเซนทรอยด์.....46

ตารางที่ หน้า

4.8 การเปรียบเทียบค่า F1 ของการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบครั้งละหลายกลุ่มด้วยด้วยวิธีเบย์อย่างง่าย วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k = 3 และวิธีแบบเซนทรอยด์.....47

4.9 ผลการจัดจำนวนกลุ่มของการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาแบบไบนารี่โดยวิธีแบบเซนทรอยด์.....49

4.10 ผลการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาชนิดหนึ่งกลุ่มแบบไบนารี่โดยวิธีแบบเซนทรอยด์.....49

4.11 ผลการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาชนิด 2 กลุ่มแบบไบนารี่โดยวิธีแบบเซนทรอยด์.....50

4.12 ผลการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติชนิดหลายกลุ่มตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบไบนารี่ด้วยวิธีแบบเซนทรอยด์.....51

4.13 ผลการประเมินการจัดแบ่งประเภทเอกสารโดยอัตโนมัติชนิดหลายกลุ่มตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีไบนารี่ด้วยวิธีแบบเซนทรอยด์.....51

4.14 ผลการเปรียบเทียบการจัดแบ่งประเภทเอกสารข้อมูลยาตามกลุ่มเภสัชวิทยาและการรักษาชนิดหลายกลุ่มโดยวิธีแบบครั้งละหลายกลุ่มและวิธีแบบไบนารี่.....53

4.15 ข้อดีข้อเสียของการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาชนิดหนึ่งกลุ่มโดยวิธีแบบไบนารี่และวิธีแบบครั้งละหลายกลุ่ม.....53

4.16 ผลของขั้นที่ 1 การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามจำนวนกลุ่ม [1, 2].....56

4.17 ผลการประเมินผลของขั้นที่ 1 การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามจำนวนกลุ่ม [1, 2].....56

4.18 ผลของขั้นที่ 2 การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาจำนวน 11 กลุ่ม.....58

4.19 การเปรียบเทียบค่า F1 ของผลขั้นที่ 2 การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาจำนวน 11 กลุ่มด้วยวิธีเบย์

ตารางที่ หน้า

อย่างง่าย วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k = 3 และวิธีแบบเซนทรอยด์.....59
4.20 ผลการจัดแบ่งประเภทเอกสารข้อมูลยาชนิดหลายกลุ่ม.....61
4.21 ผลของขั้นที่ 1 การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามจำนวนกลุ่ม [1, 3].....63
4.22 ผลของขั้นที่ 2 การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติเป็น 11 กลุ่มตามกลุ่มเภสัชวิทยาและการรักษา.....64
4.23 การเปรียบเทียบค่า F1 ของผลขั้นที่ 2 การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติเป็น 11 กลุ่มตามกลุ่มเภสัชวิทยาและการรักษาในการจัดแบ่งประเภทเอกสารข้อมูลยาวิธีเบย์อย่างง่าย วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k = 3 และวิธีแบบเซนทรอยด์.....65
4.24 ผลการจัดแบ่งประเภทเอกสารข้อมูลยาชนิดหลายกลุ่ม.....67

.

สำนักหอสมุดกลาง
มหาวิทยาลัยศิลปากร

สารบัญรูป

รูปที่ หน้า

2.1 สมการความน่าจะเป็นของการจัดแบ่งประเภทเอกสารแบบเบย์อย่างง่าย.....11

2.2 สมการความน่าจะเป็นตามกฎของเบย์.....11

2.3 สูตรความคล้ายคลึงของโคไซน์.....13

2.4 การคำนวณค่าความถูกต้อง.....15

2.5 การคำนวณค่าความแม่นยำ.....15

2.6 การคำนวณค่าการเรียกคืน.....15

2.7 การคำนวณค่า F1.....15

2.8 การคำนวณค่าความถูกต้อง.....15

2.9 การคำนวณค่าความแม่นยำของกลุ่ม A.....17

2.10 การคำนวณค่าการเรียกคืนของกลุ่ม A.....17

2.11 การคำนวณค่า F1ของกลุ่ม A.....17

2.12 กระบวนการจัดแบ่งประเภทเอกสาร.....20

2.13 กลุ่มคำและบทบาทในเหมืองข้อมูล Biocaster.....23

3.1 การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาชนิด 1 กลุ่มโดยวิธีแบบครั้งละหลายกลุ่ม.....32

3.2 การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบไบนารี่.....33

3.3 การจัดแบ่งประเภทเอกสารแบบ 2 ขั้นตอน.....35

4.1 การสรุปผลการจัดแบ่งประเภทเอกสารข้อมูลยาชนิดหลายกลุ่มแบบ 2 ขั้นตอนในเอกสารข้อมูลยาชนิด 1 กลุ่มและ 2 กลุ่ม.....61

4.2 การสรุปผลการจัดแบ่งประเภทเอกสารข้อมูลยาชนิดหลายกลุ่มแบบ 2 ขั้นตอน.....68

ภาพผนวก

ก.1 กลุ่มเภสัชวิทยาและการรักษาของหนังสือ AHFS Drug Information 2013.....79

ก.2 ตัวอย่างเอกสารข้อมูลยา (drug monograph).....80

บทที่ 1
บทนำ

1. ความเป็นมาและความสำคัญของปัญหา

เทคโนโลยีและสารสนเทศต่างๆ มีการพัฒนาขึ้นอย่างรวดเร็วและครอบคลุมไปทั่วโลก ทำให้ความรู้ต่างๆ จากในอดีตและความรู้ใหม่ๆ ในปัจจุบันมีจำนวนเพิ่มขึ้นและมีการพัฒนาขึ้นเป็นอย่างมาก จึงเกิดข้อมูลข่าวสารต่างๆ จำนวนมากและสามารถแพร่กระจายได้ทั่วโลกผ่านเครือข่ายอินเทอร์เน็ต ซึ่งจะทำให้ผู้ใช้สามารถสืบค้นและค้นคว้าหาข้อมูลข่าวสารต่างๆ ได้อย่างง่ายดายและรวดเร็ว ข้อมูลข่าวสารทางสุขภาพก็เป็นอีกข้อมูลหนึ่งซึ่งมีจำนวนเพิ่มขึ้นอย่างรวดเร็ว ผู้ที่ใช้เครือข่ายอินเทอร์เน็ตสามารถสืบค้นหาข้อมูลที่เป็นประโยชน์ต่อสุขภาพได้ ทั้งเรื่องโรค การรักษา การป้องกันและการดูแลสุขภาพของตนเองให้ปราศจากโรค ซึ่งข้อมูลเหล่านี้จะเป็นประโยชน์ต่อมนุษย์เป็นอย่างมาก จะทำให้มนุษย์มีสุขภาพที่ดี ปราศจากโรคภัย ทำให้มีคุณภาพชีวิตที่ดีได้ ผู้ที่ใช้ข้อมูลทางสุขภาพมีทั้งบุคลากรทางสายการแพทย์ ผู้ป่วย รวมทั้งบุคคลทั่วไปที่สนใจ แต่ผู้ที่มีความเกี่ยวข้องโดยตรงคือบุคลากรทางสายการแพทย์โดยเฉพาะอย่างยิ่งเภสัชกร ซึ่งบุคคลเหล่านี้จะเป็นผู้ที่ให้ข้อมูลทางด้านการดูแลรักษาสุขภาพ โรคและการรักษาแก่ผู้ป่วย เพื่อให้ผู้ป่วยมีคุณภาพชีวิตที่ดีขึ้น แต่เนื่องจากข้อมูลทางด้านสุขภาพที่มีมากขึ้น ทำให้การค้นคว้าหาข้อมูลนั้นต้องใช้เวลานาน ดังนั้นจึงเกิดการจัดแบ่งประเภทของเอกสารขึ้น เพื่อจัดแบ่งเอกสารข้อมูลที่มีอยู่ให้เป็นกลุ่มหรือหมวดหมู่ ง่ายต่อการค้นคว้า และสามารถค้นคว้าได้อย่างรวดเร็ว

ในอดีตการจัดแบ่งประเภทเอกสารมีการกระทำโดยมนุษย์ แต่เนื่องจากเอกสารที่มีจำนวนมากขึ้นในปัจจุบัน การจัดแบ่งประเภทเอกสารจึงต้องใช้แรงงานมนุษย์เป็นจำนวนมาก และเอกสารบางประเภทยังต้องใช้ทักษะและความสามารถเฉพาะทางของมนุษย์ในการจัดแบ่งประเภทเอกสารอีกด้วย ดังเช่นเอกสารทางด้านสุขภาพ รวมทั้งเอกสารข้อมูลยา ซึ่งเป็นเอกสารที่ให้ข้อมูลที่เกี่ยวกับยารักษาโรค การออกฤทธิ์ของยา ข้อควรระวังในการใช้ยา และผลข้างเคียงของยา เป็นต้น ซึ่งข้อมูลเหล่านี้มีความจำเป็นมากในการรักษาโรคของผู้ป่วย ซึ่งจะต้องมีการใช้ยาที่ถูกโรค ถูกวิธี และ

ถูกเวลา ดังนั้นเภสัชกรจึงควรที่จะสามารถค้นคว้าข้อมูลยาได้อย่างรวดเร็วและถูกต้อง ซึ่งการจัดแบ่งประเภทเอกสารให้เป็นหมวดหมู่ จะช่วยให้เภสัชกรสามารถค้นคว้าหาข้อมูลได้รวดเร็วขึ้น จากที่เอกสารข้อมูลยานี้เป็นเอกสารเฉพาะทาง ผู้ที่จะสามารถจัดแบ่งเอกสารข้อมูลยาได้นั้น จะต้องเป็นผู้ที่มีทักษะความรู้ในด้านยาจึงจะสามารถจัดแบ่งประเภทเอกสารข้อมูลยาได้อย่างถูกต้อง จึงต้องใช้ผู้ที่มีความเชี่ยวชาญ ทักษะ และความรู้เรื่องยา มาจัดแบ่งประเภทเอกสาร และจากเอกสารที่มีจำนวนเพิ่มมากขึ้นในปัจจุบัน จึงต้องใช้แรงงานมนุษย์ในการจัดแบ่งประเภทเอกสารเป็นจำนวนมาก ดังนั้น การจัดแบ่งประเภทเอกสารโดยอัตโนมัติจึงเป็นอีกทางเลือกหนึ่งที่จะทำให้สามารถจัดแบ่งประเภทเอกสารได้อย่างรวดเร็ว

การจัดแบ่งประเภทเอกสารโดยอัตโนมัติ (Text categorization, text classification, topic spotting, automatic text categorization, automatic text classification) คือการจัดแบ่งประเภทของเอกสารโดยคอมพิวเตอร์ตามหมวดหมู่หรือกลุ่มที่ได้จัดแบ่งไว้ก่อนแล้ว โดยใช้หลักการเรียนรู้ด้วยเครื่อง (machine learning) โดยมนุษย์จะต้องเป็นผู้สอนคอมพิวเตอร์ให้เรียนรู้ด้วยตนเองจากหมวดหมู่ที่กำหนดและจัดแบ่งประเภทไว้ก่อนแล้ว ซึ่งเป็นการใช้ตัวจัดแบ่งประเภท (classifier) ในการสร้างแบบจำลองหรือตัวแบบ (model) ขึ้น หลังจากนั้นคอมพิวเตอร์จึงจะสามารถจัดแบ่งประเภทเอกสารได้ด้วยตนเอง โดยวิธีการนี้จะทำให้สามารถจัดแบ่งประเภทเอกสารที่มีจำนวนมากได้อย่างรวดเร็วและลดแรงงานมนุษย์ที่ใช้ในการจัดแบ่งประเภทเอกสารได้อีกด้วย [1, 2, 3, 4, 5]

ปัจจุบันมีการพัฒนาการจัดแบ่งประเภทเอกสารโดยอัตโนมัติในหลายส่วนที่สำคัญ เช่น ตัวแทนของเอกสาร การสร้างแบบจำลอง และการประเมินแบบจำลอง เป็นต้น [1, 6, 7, 8] และเริ่มนำการจัดแบ่งประเภทมาจัดแบ่งประเภทเอกสารต่างๆ เช่น Son Doan และคณะได้ศึกษาบทบาทของชื่อ (name entities) ในการจัดแบ่งประเภทของรายงานข่าวโรคระบาด เป็นต้น [9] แต่การนำมาใช้ในเรื่องที่เกี่ยวข้องกับทางด้านสุขภาพหรือทางการแพทย์ยังมีไม่มากนัก ตัวอย่างดังงานวิจัยของ Anastasia Sotelsek-margalef , Julio Villena-román และ Iii Madrid ได้ศึกษาถึงการจัดแบ่งประเภทเอกสารบันทึกการรักษาทางคลินิกและทางรังสีของผู้ป่วยในการจัดแบ่งประเภทอาการของผู้ป่วยตาม ICD-9-CM เพื่อสร้าง MIDAS (Medical Diagnosis Assistant) ระบบผู้เชี่ยวชาญที่ช่วยในการวินิจฉัยอาการของผู้ป่วย โดยใช้การสกัดข้อมูลสารสนเทศ (Information extraction) และการเรียนรู้ด้วยเครื่องจากประวัติการรักษาของผู้ป่วย แล้วนำมาจัดแบ่งประเภทโดยใช้ Weka เป็นเครื่องมือเพื่อช่วยในการเรียนรู้ แล้วทดสอบการจัดแบ่งประเภทโดยใช้วิธีต้นไม้ตัดสินใจ J48 และ

เพื่อนบ้านใกล้ที่สุดเป็นตัวจัดแบ่งประเภท ผลปรากฏว่าสามารถจัดแบ่งประเภทอาการจากประวัติการรักษาได้ [10]

เนื่องจากการนำการจัดแบ่งประเภทเอกสารโดยอัตโนมัติไปใช้ในเรื่องที่เกี่ยวกับยาสุขภาพ และเรื่องอื่นๆ ทางการแพทย์นั้นยังมีไม่มากนัก ผู้จัดทำวิทยานิพนธ์จึงสนใจที่จะทดลองนำการจัดแบ่งประเภทเอกสารโดยอัตโนมัติมาใช้ในเอกสารข้อมูลยา

2. วัตถุประสงค์ของการวิจัย

2.1. เพื่อศึกษาการจัดแบ่งประเภทเอกสารโดยอัตโนมัติสามารถใช้จัดแบ่งประเภทของเอกสารข้อมูลยาตามกลุ่มเภสัชวิทยาและการรักษาได้

2.2. เพื่อศึกษาหาตัวจัดแบ่งประเภทที่ดีที่สุดในการจัดแบ่งประเภทเอกสารข้อมูลยาตามกลุ่มเภสัชวิทยาและการรักษา

2.3. เพื่อศึกษาวิธีการจัดแบ่งประเภทเอกสารชนิดหนึ่งกลุ่ม (single-label categorization) ในเอกสารข้อมูลยาตามกลุ่มเภสัชวิทยาและการรักษา

2.4. เพื่อศึกษาวิธีการจัดแบ่งประเภทเอกสารชนิดหลายกลุ่ม (multi-label categorization)ในเอกสารข้อมูลยาตามกลุ่มเภสัชวิทยาและการรักษา

3. ขอบเขตของการวิจัย

3.1. คัดเลือกเอกสารแหล่งข้อมูลยาจาก www.google.com ได้แหล่งข้อมูลยาจำนวน 2 เว็บไซต์ คือ RxList (http://www.rxlist.com) และ DailyMed (http://dailymed.nlm.nih.gov) นำมาทำชุดข้อมูล (data sets) ที่จะใช้ในการจัดแบ่งประเภท โดยจัดแบ่งประเภทตามกลุ่มเภสัชวิทยาและการรักษาของหนังสือ AHFS Drug Information 2012 โดยแหล่งข้อมูลยาที่นำมาทำชุดข้อมูลนี้ ต้องมีคุณลักษณะดังนี้ คือ

3.1.1. เป็นไฟล์ .htm หรือ .html เนื่องจากสามารถนำข้อมูลภายในไฟล์มาใช้ได้ทันที

3.1.2. ข้อมูลในแต่ละไฟล์ มีลักษณะที่เป็นโครงสร้างข้อมูลตามลักษณะของข้อมูลยา (drug monograph) ตามหนังสือ AHFS Drug Information

3.1.3. จำนวนไฟล์ในแหล่งข้อมูลมีจำนวนมากพอ โดยมีประมาณ 2,000-5,000 ไฟล์ เพื่อที่จะนำมาคัดกรองให้ได้จำนวนมากพอที่จะนำมาใช้ได้

3.2. การจัดแบ่งประเภทนี้อ้างอิงตามกลุ่มเภสัชวิทยาและการรักษาของหนังสือ AHFS Drug Information 2012 เนื่องจากเป็นหนังสือที่นิยมใช้ในการค้นคว้าหาข้อมูลยาของเภสัชกร โดยจะใช้จำนวนทั้งหมด 11 กลุ่ม

3.3. ตัวจัดแบ่งประเภทที่ใช้จำนวน 3 วิธี คือ วิธีเบย์อย่างง่าย (naïve Bayes classifier), วิธีเพื่อนบ้านใกล้ที่สุด (k-nearest neighbor) และวิธีแบบเซนทรอยด์ (centroid-based classifier)

สำนักหอสมุดกลาง

4. นิยามศัพท์เฉพาะ

4.1 การจัดแบ่งประเภทเอกสารโดยอัตโนมัติ (text categorization, text classification) คือการจัดแบ่งประเภทของเอกสารโดยคอมพิวเตอร์ตามกลุ่มหรือหมวดหมู่ที่ได้จัดแบ่งไว้ก่อนแล้ว โดยใช้หลักการเรียนรู้ด้วยเครื่อง โดยมนุษย์จะต้องเป็นผู้สอนคอมพิวเตอร์ให้เรียนรู้ด้วยตนเองจากกลุ่มหรือหมวดหมู่ที่ได้กำหนดและจัดแบ่งประเภทไว้แล้ว ซึ่งจะสร้างแบบจำลองขึ้น หลังจากนั้นคอมพิวเตอร์จึงจะสามารถจัดแบ่งประเภทได้จากแบบจำลองที่สร้างขึ้น โดยใช้ตัวจัดแบ่งประเภท (classifier)

4.2 ตัวจัดแบ่งประเภทหรือวิธีการจัดแบ่งประเภท (classifier) เป็นอัลกอริทึมของการเรียนรู้ด้วยเครื่อง สามารถนำมาใช้จัดแบ่งประเภทได้ เช่น วิธีเบย์อย่างง่าย (naïve Bayes), วิธีแบบเซนทรอยด์ (centroid-based) และวิธีเพื่อนบ้านใกล้ที่สุด (k-nearest neighbor)

4.3 แบบจำลองหรือตัวแบบ (model) เป็นแบบจำลองที่สร้างขึ้นเพื่อนำไปใช้ในการจัดแบ่งประเภท

4.4 การจัดแบ่งประเภทเอกสารชนิดหนึ่งกลุ่ม (single-label classification) คือ เอกสาร 1 เอกสารจะสามารถกำหนดให้อยู่ได้เพียง 1 กลุ่มเท่านั้น

4.5 การจัดแบ่งประเภทเอกสารชนิดหลายกลุ่ม (multi-label classification) คือ เอกสาร 1 เอกสารจะสามารถกำหนดให้อยู่ได้ตั้งแต่ 1 กลุ่มขึ้นไป

4.6 การจัดแบ่งประเภทแบบไบนารี่ (binary classification) หรือการจัดแบ่งประเภทแบบครั้งละสองกลุ่ม คือ ในการจัดแบ่งประเภท 1 ครั้ง สามารถจัดแบ่งประเภทได้ 2 กลุ่ม

4.7 การจัดแบ่งประเภทแบบครั้งละหลายกลุ่ม (multiclass classification) คือ ในการจัดแบ่งประเภท 1 ครั้ง สามารถจัดแบ่งประเภทได้มากกว่า 2 กลุ่ม

4.8 การจัดแบ่งประเภทเอกสารแบบ 2 ขั้นตอน (two-step classification) เป็นการจัดแบ่งประเภทเอกสารข้อมูลยาชนิดหลายกลุ่มที่ได้พัฒนาขึ้น มีขั้นตอนดังนี้

ขั้นที่ 1 การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามจำนวนกลุ่ม [a, b]

เมื่อกำหนดให้ [a, b] คือช่วงของจำนวนกลุ่มที่จัดแบ่งประเภทได้

a คือ จำนวนกลุ่มที่น้อยที่สุดที่เอกสารสามารถจัดแบ่งประเภทอยู่ได้

b คือ จำนวนกลุ่มที่มากที่สุดที่เอกสารสามารถจัดแบ่งประเภทอยู่ได้

ขั้นที่ 2 การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติเป็น 11 กลุ่มตามกลุ่มเภสัชวิทยาและการรักษา

5. ประโยชน์ที่ได้รับ

5.1. เข้าใจการจัดแบ่งประเภทของเอกสารข้อมูลยาตามหนังสือ AHFS Drug Information 2012

5.2. เข้าใจการเก็บรวบรวมข้อมูลเพื่อจัดแบ่งประเภทเอกสารโดยอัตโนมัติ

5.3. รู้จักใช้เครื่องมือในการจัดแบ่งประเภทเอกสารโดยอัตโนมัติ

5.4. สามารถพัฒนาอัลกอริทึมในการจัดแบ่งประเภทเอกสารให้เหมาะสมกับการจัดแบ่งประเภทเอกสารข้อมูลยา

สำนักวิทยบริการฯ
มหาวิทยาลัย ศิลปากร

บทที่ 2
วรรณกรรมที่เกี่ยวข้อง

ในงานวิจัยเรื่องการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัตินี้ ผู้วิจัยได้ศึกษาและกำหนดเนื้อหาของวรรณกรรมที่เกี่ยวข้องดังหัวข้อต่อไปนี้

1. การจัดแบ่งประเภทเอกสารโดยอัตโนมัติ
2. เอกสารข้อมูลยา
3. งานวิจัยที่เกี่ยวข้อง

1. การจัดแบ่งประเภทเอกสารโดยอัตโนมัติ

ในยุค 80 มีการใช้เทคนิคของวิศวกรรมการเรียนรู้ (knowledge engineering) ในการจัดแบ่งประเภทเอกสาร (text categorization, text classification) ซึ่งเป็นการใช้ระบบผู้เชี่ยวชาญ (expert system) มาช่วยในการตัดสินใจในการจัดแบ่งประเภทเอกสาร แต่มีอุปสรรคการได้มาขององค์ความรู้ที่เกิดคอขวดขึ้น (knowledge acquisition bottleneck) คือจากความรู้ที่มีเพิ่มขึ้นและมีการเปลี่ยนแปลงอยู่ตลอดเวลาทำให้การจัดการความรู้โดยวิศวกรความรู้ (knowledge engineer) ที่มีระบบผู้เชี่ยวชาญเป็นผู้ช่วยนั้นไม่สามารถทำได้ทันท่วงทีไปพร้อมกับการเปลี่ยนแปลงของความรู้ ดังนั้นในยุค 90 จึงมีเทคนิคขั้นสูงของการเรียนรู้ด้วยเครื่อง (machine learning) เกิดขึ้น เพื่อเพิ่มประสิทธิภาพในการจัดการข้อมูลจำนวนมากที่ไม่สามารถกระทำได้โดยมนุษย์ โดยจะมีการให้ข้อมูลสอน (training data) สำหรับสอนให้คอมพิวเตอร์เรียนรู้ เพื่อให้คอมพิวเตอร์สามารถเรียนรู้และรู้ถึงสิ่งสำคัญภายในเอกสารต่างๆ ได้ด้วยตนเองโดยอัตโนมัติ [1, 2]

จากเอกสารที่มีเพิ่มมากขึ้น ทำให้มีการเก็บรวมรวมเอกสารต่างๆ เหล่านั้นไว้ แต่การนำข้อมูล (data) สารสนเทศ (information) หรือความรู้ (knowledge) ที่สำคัญภายในเอกสารที่มีมากมายมหาศาลเหล่านั้นมาใช้ ทำให้ต้องใช้เวลานานและสิ้นเปลืองแรงงานมนุษย์ ดังนั้นจึงเกิดการ

สกัดข้อมูลหรือสังเคราะห์ข้อมูลขึ้น เพื่อหาสารสนเทศหรือความรู้ที่สำคัญ เพื่อให้สามารถนำสารสนเทศและความรู้เหล่านั้นมาใช้ประโยชน์ได้ต่อไป ซึ่งสามารถเรียกวิธีนี้ว่า การทำเหมืองข้อมูล (data mining) [3, 4] การทำเหมืองข้อมูลนี้จะเป็นการหาความสัมพันธ์หรือรูปแบบของข้อมูลที่มีอยู่จริง แต่ถูกซ่อนไว้ภายในข้อมูลจำนวนมาก ซึ่งการทำเหมืองข้อมูลจะมี 2 รูปแบบ คือ การเรียนรู้แบบมีผู้สอน (supervised learning) และการเรียนรู้แบบไม่มีผู้สอน (unsupervised clustering) [3]

การเรียนรู้แบบมีผู้สอน จะมีการสอนคอมพิวเตอร์จากชุดข้อมูลสอนซึ่งมีการกำหนดกลุ่มหรือหมวดหมู่ของข้อมูลไว้แล้ว และสร้างแบบจำลองหรือตัวแบบ (model) จากชุดข้อมูลสอน (training set) และชุดข้อมูลทดสอบ (test set) ซึ่งการจัดแบ่งประเภท (classification) จะเป็นการเรียนรู้แบบมีผู้สอนแบบหนึ่ง โดยจะมีการสร้างแบบจำลองของการจัดแบ่งประเภทขึ้น ซึ่งแบบจำลองที่เกิดขึ้นนี้จะสามารถประยุกต์ใช้จัดแบ่งประเภทชุดข้อมูลที่ยังไม่ทราบกลุ่มหรือหมวดหมู่ได้ ส่วนการเรียนรู้แบบไม่มีผู้สอน จะไม่มีการสอนคอมพิวเตอร์ แต่ให้คอมพิวเตอร์เรียนรู้จากชุดข้อมูลที่ไม่มีการกำหนดกลุ่มหรือหมวดหมู่ไว้ แล้วจึงจัดกลุ่มหรือแบ่งกลุ่มจากชุดข้อมูลนั้น ซึ่งข้อมูลที่อยู่ในกลุ่มเดียวกันจะมีความคล้ายคลึงกันมากกว่าอีกกลุ่มหนึ่ง [2, 3] ดังนั้นการจัดแบ่งประเภทเอกสารโดยอัตโนมัติ (automatic text categorization, automatic text classification) จึงเป็นการเรียนรู้แบบมีผู้สอน

การจัดแบ่งประเภทเอกสารโดยอัตโนมัติ คือการจัดแบ่งประเภทของเอกสารโดยคอมพิวเตอร์ตามกลุ่มหรือหมวดหมู่ที่ได้จัดแบ่งไว้ก่อนแล้ว โดยใช้หลักการเรียนรู้ด้วยเครื่อง โดยมนุษย์จะต้องเป็นผู้สอนคอมพิวเตอร์ให้เรียนรู้ด้วยตนเองจากกลุ่มหรือหมวดหมู่ที่ได้กำหนดและจัดแบ่งประเภทไว้แล้ว ซึ่งจะสร้างแบบจำลองขึ้น หลังจากนั้นคอมพิวเตอร์จึงจะสามารถจัดแบ่งประเภทได้จากแบบจำลองที่สร้างขึ้น โดยใช้ตัวจัดแบ่งประเภท (classifier) โดยวิธีการนี้จะทำให้สามารถจัดแบ่งประเภทเอกสารที่มีจำนวนมากได้อย่างรวดเร็วและลดแรงงานมนุษย์ที่ใช้ในการจัดแบ่งประเภทเอกสารได้อีกด้วย [1, 3, 10]

การจัดแบ่งประเภทเอกสารโดยอัตโนมัติจะมีการกำหนดค่าเป็นค่าตรรกกะ (Boolean value) ของ $\langle d_p, c_k \rangle \in DxC$ เมื่อกำหนดให้

D เป็นชุดเอกสารที่นำมาจัดแบ่งประเภท ซึ่ง $D = \{d_1,...,d_p\}$ เมื่อ p คือจำนวนเอกสารที่มีอยู่

C เป็นกลุ่มหรือหมวดหมู่ที่กำหนดไว้ก่อนแล้ว ซึ่ง $C = \{c_1,...,c_k\}$ เมื่อ k คือจำนวนกลุ่มของเอกสาร

$<d_p, c_k>$ จะมีค่าเป็นจริง (T) เมื่อเอกสาร d_p อยู่ในกลุ่ม c_k

และ$<d_p, c_k>$ จะมีค่าเป็นเท็จ (F) เมื่อเอกสาร d_p ไม่อยู่ในกลุ่ม c_k

การจัดแบ่งประเภทเอกสารโดยอัตโนมัติมี 2 ชนิด คือ การจัดแบ่งประเภทเอกสารชนิดหนึ่งกลุ่ม (single-label categorization) และการจัดแบ่งประเภทเอกสารชนิดหลายกลุ่ม (multi-label categorization)

การจัดแบ่งประเภทเอกสารชนิดหนึ่งกลุ่ม คือเอกสาร 1 เอกสารจะสามารถกำหนดให้อยู่ได้เพียง 1 กลุ่มเท่านั้น ดังนั้นเอกสาร $d_p \in D$ จะถูกกำหนดให้อยู่ในกลุ่ม c_k ได้เพียง 1 กลุ่ม เมื่อ c_k เป็นสมาชิกของ C แต่เนื่องจากเอกสารบางชนิดไม่สามารถจัดให้อยู่ในกลุ่มใดกลุ่มหนึ่งได้ จึงเกิดการจัดแบ่งประเภทเอกสารชนิดหลายกลุ่มขึ้น ซึ่งเอกสาร $d_p \in D$ จะถูกกำหนดให้อยู่ในกลุ่ม C ได้หลายกลุ่มตั้งแต่ 0 ถึง k

ในอดีตการจัดแบ่งประเภทเอกสารโดยอัตโนมัติมักทำในวิธีแบบไบนารี่ หรือเรียกว่า การจัดแบ่งประเภทแบบไบนารี่ (binary classification) เมื่อกำหนดให้

D เป็นชุดเอกสารที่นำมาจัดแบ่งประเภท ซึ่ง $D = \{d_1,...,d_p\}$ เมื่อ p คือจำนวนเอกสารที่มีอยู่

C เป็นกลุ่มหรือหมวดหมู่ที่กำหนดไว้ก่อนแล้ว ซึ่ง $C = \{c_k, \bar{c}_k\}$ เมื่อ k คือจำนวนกลุ่มของเอกสาร และ $\bar{c}_k$ คือไม่อยู่ในกลุ่ม k

การจัดแบ่งประเภทเอกสารแบบไบนารี่ มีการจัดแบ่งประเภทเอกสารโดยในการจัดแบ่งประเภท 1 ครั้ง เอกสาร 1 เอกสารจะจัดแบ่งประเภทให้อยู่ในกลุ่มนั้นหรือไม่อยู่ในกลุ่มนั้น คือ d_p จะถูกกำหนดให้อยู่ในกลุ่ม c_k, หรือ $\bar{c}_k$ แต่มีการจัดแบ่งประเภทเอกสารอีกแบบหนึ่งซึ่งในการจัดแบ่งประเภทเอกสาร 1 ครั้งจะสามารถจัดแบ่งประเภทเอกสารให้อยู่ได้ในหลายกลุ่ม ซึ่งจะเรียกการจัดแบ่งประเภทแบบนี้ว่าการจัดแบ่งประเภทแบบครั้งละหลายกลุ่ม (multiclass classification) โดยการจัดประเภทแบบครั้งละหลายกลุ่มนี้ ในการจัดแบ่งประเภทเอกสาร 1 ครั้ง เอกสาร d_p จะถูกจัดแบ่งประเภทให้อยู่ในกลุ่ม c_k มากกว่า 1 กลุ่มขึ้นไป [1]

การจัดแบ่งประเภทเอกสารโดยอัตโนมัติจะมีขั้นตอนการทำ 3 ขั้นตอนใหญ่ๆ คือ การเตรียมชุดข้อมูล การจัดแบ่งประเภท และการประเมิน ซึ่งมีรายละเอียดดังนี้

1.1. การเตรียมชุดข้อมูล

ชุดข้อมูลที่นำมาใช้ในการจัดแบ่งประเภท จะต้องมีกระบวนการในการเตรียมข้อมูล (preprocessing) ให้ได้ข้อมูลที่มีประโยชน์สามารถนำไปประมวลผลการจัดแบ่งประเภทได้ เพื่อเพิ่มประสิทธิภาพของการจัดแบ่งประเภทเอกสาร เช่น การทำความสะอาดข้อมูล (data cleaning) และ การปรับเปลี่ยนรูปแบบข้อมูล (data transformation) [3, 4] ซึ่งในงานวิจัยชิ้นนี้ได้ใช้การตัดคำหยุด (stopwords), การตัดโครงสร้างคำสั่งของภาษา HTML และการให้น้ำหนักคำ (term weighting) ในการเตรียมชุดข้อมูล

คำหยุด คือคำขยายหรือคำที่ไม่มีความหมายในตัวเอง เช่น a, an, the, and, but, can, about, of, on, in และ at เป็นต้น [2] โครงสร้างคำสั่งของภาษา HTML เช่น <html>, <b>, , <br>, และ <p> เป็นต้น ซึ่งการตัดคำหยุดและโครงสร้างคำสั่งของภาษา HTML นี้จะทำให้คำที่เหลืออยู่ในเอกสารเป็นคำที่มีความสำคัญของเอกสารนั้น ส่วนการให้น้ำหนักคำจะเป็นการเพิ่มประสิทธิภาพของคำที่มีความสำคัญของเอกสาร การให้น้ำหนักคำมีหลายรูปแบบ แต่ในการศึกษาครั้งนี้จะใช้การให้น้ำหนักคำแบบ *tfidf* หรือ $tf \times idf$ โดยจะให้น้ำหนักแก่คำที่มีจำนวนมากในเอกสารนั้น แต่มีจำนวนน้อยในเอกสารอื่น ๆ ซึ่ง *tf* หรือ term frequency คือความถี่ของคำที่ปรากฏอยู่ในเอกสาร และ *idf* หรือ inverse document frequency คือส่วนกลับของจำนวนเอกสารที่มีคำนั้นปรากฏอยู่เทียบกับจำนวนเอกสารทั้งหมด $\log\frac{N}{n}$ [1, 6] นอกจากนี้เรายังเพิ่มประสิทธิภาพโดยใช้การกระจายน้ำหนักของคำ (term distribution) ในการศึกษานี้เราใช้ส่วนเบี่ยงเบนมาตรฐาน (standard deviation, sd) เพื่อลดคำที่ไม่สัมพันธ์กับกลุ่ม โดยคำที่มีการกระจายน้ำหนักมากจะมีส่วนเบี่ยงเบนมาตรฐานสูง ซึ่งจะปรากฏในหลายเอกสาร แต่เมื่อใช้หารน้ำหนักของคำด้วยส่วนเบี่ยงเบนมาตรฐาน $\frac{tf \times idf}{sd}$ หรือ *tfidf/sd* จะทำให้ลดการกระจายน้ำหนักของคำลง ทำให้คำที่ไม่สัมพันธ์กับกลุ่มมีจำนวนลดลง การจัดแบ่งประเภทจึงมีประสิทธิภาพมากขึ้น [11]

ในการแบ่งชุดข้อมูลออกเป็นชุดข้อมูลสอนและชุดข้อมูลทดสอบนั้น ในการศึกษานี้ได้สนใจที่จะใช้การตรวจสอบไขว้ (cross-validation) มาช่วยในการจัดแบ่งชุดข้อมูลเพื่อลดขีดจำกัดในเรื่องจำนวนข้อมูลหรือเอกสารที่นำมาใช้จัดแบ่งประเภทไม่เพียงพอ และสามารถตรวจสอบประสิทธิภาพของวิธีการจัดแบ่งประเภทได้อีกด้วย โดยการตรวจสอบไขว้จะเป็นวิธีที่แบ่งเอกสารออกเป็นส่วน ส่วนละเท่าๆ กัน ในรอบที่ 1 ให้เอกสารส่วนที่ 1 เป็นชุดข้อมูลทดสอบ และ

เอกสารส่วนที่เหลือเป็นชุดข้อมูลสอน แล้วจัดแบ่งประเภท ในรอบที่ 2 ให้นำเอกสารชุดที่ 2 เป็นชุดข้อมูลทดสอบ และเอกสารส่วนที่เหลือเป็นชุดข้อมูลสอน โดยเอกสารส่วนที่ 1 จากครั้งแรกนั้นจะถูกรวมเข้าไปเป็นชุดข้อมูลสอนด้วย แล้วจึงจัดแบ่งประเภท โดยเอกสารที่เป็นชุดข้อมูลทดสอบในแต่ละครั้งจะไม่ซ้ำกัน และทำไปเรื่อยๆ จนครบรอบที่กำหนด เช่น ถ้ากำหนดการตรวจสอบแบบไขว้ 5 รอบ (5-fold cross-validation) จะต้องทำทั้งหมด 5 รอบ เป็นต้น [2]

1.2. การจัดแบ่งประเภท

หลังจากเตรียมข้อมูลที่จะนำมาใช้จัดแบ่งเอกสารเรียบร้อยแล้ว จึงสามารถนำมาจัดแบ่งประเภทเอกสารได้ ซึ่งการจัดแบ่งประเภทเอกสารจะต้องแบ่งชุดข้อมูลออกเป็นชุดข้อมูลสอนและชุดข้อมูลทดสอบ โดยตัวจัดแบ่งประเภทหรือวิธีการจัดแบ่งประเภทนี้จะสร้างแบบจำลองจากชุดข้อมูลสอนที่มีการกำหนดกลุ่มหรือหมวดหมู่ไว้แล้ว ซึ่งสามารถกล่าวได้ว่าตัวจัดแบ่งประเภทนี้ใช้วิธีการเรียนรู้ (learning algorithm) ในการสร้างแบบจำลองเพื่อให้ได้แบบจำลองที่ดีที่สุดที่มีความสัมพันธ์ระหว่างข้อมูลและกลุ่มที่กำหนดไว้แล้ว โดยแบบจำลองที่ได้จะต้องจัดแบ่งประเภทชุดข้อมูลทดสอบซึ่งตัวจัดแบ่งประเภทไม่เคยเห็นมาก่อนได้อย่างถูกต้อง ตัวอย่างของตัวจัดแบ่งประเภทที่ใช้สร้างแบบจำลอง เช่น วิธีต้นไม้ตัดสินใจ (decision trees), กฎความสัมพันธ์ (association rules), โครงข่ายประสาทเทียม (artificial neural network, ANN), และเบย์อย่างง่าย (naïve Bayes, NB) [5] นอกจากวิธีการจัดแบ่งประเภทแล้วยังต้องมีเครื่องมือที่มาช่วยให้สามารถจัดแบ่งประเภทได้โดยง่าย ดังนั้นจึงสนใจใช้ Rainbow เป็นเครื่องมือจัดแบ่งประเภทในการศึกษาครั้งนี้

1.2.1. วิธีการจัดแบ่งประเภทเอกสาร

วิธีที่ใช้ในการจัดแบ่งประเภทเอกสารโดยอัตโนมัตินั้น จะใช้เทคนิคของการเรียนรู้ด้วยเครื่องมาเป็นตัวจัดแบ่งประเภทของเอกสาร ซึ่งมีเทคนิคหรือวิธีที่สามารถนำมาใช้ได้มีหลายวิธีด้วยกัน เช่น วิธีต้นไม้ตัดสินใจ (decision trees), กฎความสัมพันธ์ (association rules), โครงข่ายประสาทเทียม (artificial neural network, ANN), ซัพพอร์ตเวกเตอร์แมชชีน (support vector machines, SVMs), เบย์อย่างง่าย (naïve Bayes, NB) , แบบเซนทรอยด์ (centroid-based, CB) และ เพื่อนบ้านใกล้ที่สุด (k-nearest neighbor, k-NN) เป็นต้น ในการศึกษาครั้งนี้จะใช้ 3 วิธี คือ เบย์อย่างง่าย, แบบเซนทรอยด์ และเพื่อนบ้านใกล้ที่สุด

1.2.1.1. วิธีการจัดแบ่งประเภทเอกสารโดยอัตโนมัติแบบเบย์อย่างง่าย

การจัดแบ่งประเภทเอกสารโดยอัตโนมัติแบบเบย์อย่างง่าย เป็นการจัดแบ่งประเภทโดยประมาณความน่าจะเป็นที่มีเงื่อนไขเกี่ยวข้องกับกลุ่มนั้น โดยใช้หลักความน่าจะเป็นบนพื้นฐานของกฎของเบย์ (Bayes' theorem) ซึ่งทุกคำในเอกสารจะถูกสมมุติให้เป็นอิสระต่อกันและจะกำหนดให้อยู่ในกลุ่มหรือหมวดหมู่ (class) *y* เมื่อเอกสาร $X = \{X_1, X_2, ..., X_d\}$ มีคำ *d* ซึ่งสามารถเขียนเป็นสมการได้ ดังรูปที่ 2.1 และสมการความน่าจะเป็นตามกฎของเบย์ ดังรูปที่ 2.2

$$P(X|Y = y) = \prod_{i=1}^{d} P(X_i|Y = y)$$

รูปที่ 2.1 แสดงสมการความน่าจะเป็นของการจัดแบ่งประเภทเอกสารแบบเบย์อย่างง่าย

$$P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)}$$

รูปที่ 2.2 แสดงสมการความน่าจะเป็นตามกฎของเบย์

จากกฎของเบย์ในรูปที่ 2.2 กำหนดให้ *X* เป็นเหตุการณ์หรือในที่นี้คือเอกสาร และ *Y* เป็นสมมุติฐานหรือในที่นี้คือกลุ่ม ดังนั้น

P(Y|X) หรือ posterior probability ของ *Y* เป็นสิ่งที่ต้องการทราบ ซึ่งหมายถึงค่าความน่าจะเป็นของ *Y* หลังจากที่มี *X* เกิดขึ้น

P(Y) หรือ prior probability ของ *Y* หรือ priori probability ของ *Y* คือค่าความน่าจะเป็นของ *Y* ก่อนที่จะมี *X* เกิดขึ้น

P(X|Y) หรือ posterior probability ของ *X* คือค่าความน่าจะเป็นของ X หลังจากมี *Y* เกิดขึ้น

P(X) หรือ prior probability ของ *X* หรือ priori probability ของ *X* คือค่าความน่าจะเป็นของ *X* ก่อนที่จะมี *Y* เกิดขึ้น

ในขั้นตอนของการเรียนรู้ (training phase) จะมีการเรียนรู้ความน่าจะเป็นของ *Y* หลังจากที่มี *X* เกิดขึ้น คือ *P(Y|X)* จากชุดข้อมูลสอนหรือชุดข้อมูลสำหรับเรียนรู้ แล้วจึงหาความน่าจะเป็นของชุดข้อมูลทดสอบด้วยแบบจำลองที่ได้ [1, 5]

1.2.1.2. วิธีการจัดแบ่งประเภทเอกสารโดยอัตโนมัติแบบเซนทรอยด์

การจัดแบ่งประเภทเอกสารโดยอัตโนมัติด้วยวิธีแบบเซนทรอยด์ เป็นการใช้วิธีของแบบจำลองเวกเตอร์สเปซ (vector space model, VSM) ในการจัดแบ่งประเภท โดยจะค้นหาเอกสารต้นแบบ (prototype document) จากเอกสารทั้งหมดที่ได้กำหนดกลุ่มหรือหมวดหมู่ไว้แล้ว ซึ่งจะเรียกว่า เซนทรอยด์ (centroid) หรือ ร็อคคิโอ (Rocchio) ซึ่งกำหนดให้ ชุดของเวกเตอร์เอกสารจำนวน *p* คือ $D = \{\vec{d}_1, \ldots, \vec{d}_p\}$ จัดให้อยู่ในกลุ่ม *C* จำนวน *k* กลุ่ม $C = \{C_1, \ldots, C_k\}$ ซึ่งจะใช้ D_{Cj} แทน ชุดของเวกเตอร์เอกสารที่จัดอยู่ในกลุ่ม C_j เมื่อ $1 \le j \le k$ โดยเซนทรอยด์ของกลุ่ม C_j จะแทนด้วยเวกเตอร์ $\vec{c}_j$ เมื่อต้องการจัดแบ่งเอกสารที่ไม่ทราบกลุ่ม จะใช้เวกเตอร์เอกสารนั้นเทียบกับเวกเตอร์เอกสารต้นแบบทั้งหมด ถ้าใกล้เคียงกับเวกเตอร์เอกสารต้นแบบใด ก็จะจัดได้ว่าเป็นกลุ่มที่เวกเตอร์เอกสารต้นแบบนั้นอยู่ [7, 12] เมื่อต้องการเพิ่มประสิทธิภาพของการการจัดแบ่งประเภท จึงใช้การให้น้ำหนักคำ *tfidf/sd* มาช่วย ซึ่งเอกสารหนึ่งจะประกอบด้วยคำที่มีการให้น้ำหนักแล้ว $\vec{d}_j = \{w_{1j}, w_{2j} \ldots, w_{kj}\}$ เมื่อ w_{ij} คือน้ำหนักที่ให้แก่คำ t_i ในเอกสารนั้น นอกจากนี้ยังทำนอร์มัลไลเซชั่น (normalization) ซึ่งจะมีข้อดีต่อเอกสารที่มีขนาดใหญ่ ทำให้ความถี่ของคำ (*tf*) สูงขึ้นและเพิ่มจำนวนคำที่สามารถเป็นตัวแทนในเอกสารได้มากขึ้น ซึ่งการทำนอร์มัลไลเซชั่นที่นิยมกันมากคือการทำนอร์มัลไลเซชั่นโดยใช้โคไซน์ (cosine normalization) คือการนำขนาดของเวกเตอร์ $\sqrt{\sum_i w_i^2}$ ไปหารจำนวนคำของทุกคำ เมื่อ w_i คือน้ำหนักของคำ t_i ก่อนทำนอร์มัลไลเซชั่น [11]

1.2.1.3. วิธีการจัดแบ่งประเภทเอกสารโดยอัตโนมัติแบบเพื่อนบ้านใกล้ที่สุด

การจัดแบ่งประเภทเอกสารโดยอัตโนมัติด้วยวิธีแบบเพื่อนบ้านใกล้ที่สุด เรียกว่าเป็นการเรียนรู้แบบขี้เกียจ (lazy learners) เนื่องจากตัวจัดแบ่งประเภทนี้จะไม่สร้างแบบจำลองขึ้นมาก่อนล่วงหน้า แต่จะรอจนกว่าจะทราบชุดข้อมูลทดสอบแล้วจึง

เริ่มกระบวนการทำงาน โดยวิธีเพื่อนบ้านใกล้ที่สุดจะใช้การเปรียบเทียบชุดข้อมูลทดสอบที่คล้ายคลึงกับชุดข้อมูลสอนมากที่สุด ดังนั้นวิธีเพื่อนบ้านใกล้ที่สุดจึงต้องคำนวณหามุมระหว่างเวกเตอร์ของข้อมูลทดสอบและเวกเตอร์ของข้อมูลตัวอย่าง เพื่อหามุมที่ใกล้ที่สุด โดยใช้ความคล้ายคลึงของโคไซน์ (cosine similar) เพื่อหาขนาดของมุมของเอกสาร ดังสูตรในรูปที่ 2.3 ในการจัดแบ่งประเภทเอกสารโดยอัตโนมัติด้วยวิธีเพื่อนบ้านใกล้ที่สุด ชุดข้อมูลสอนจะประกอบด้วยคำจำนวน *n* คำ แต่ละเอกสารจะมีสเปซ (space) *n* มิติ และเมื่อให้เอกสารที่ไม่ทราบกลุ่ม ตัวจัดแบ่งประเภทเพื่อนบ้านใกล้ที่สุด *k* ตัว (k-nearest neighbor classifier) จะค้นหาสเปซที่มีเอกสารในชุดข้อมูลสอนจำนวน *k* เอกสาร ซึ่งใกล้เคียงกับเอกสารจากชุดทดสอบ ดังนั้นจึงเรียกเอกสารในชุดข้อมูลสอนจำนวน *k* เอกสารนี้ว่าเพื่อนบ้านใกล้ที่สุด k ตัวของเอกสารที่ไม่ทราบกลุ่ม [4, 5]

$$similarity = \cos(\theta) = \frac{A \cdot B}{\|A\|\|B\|} = \frac{\sum_{i=1}^{n} A_i \times B_i}{\sqrt{\sum_{i=1}^{n} (A_i)^2} \times \sqrt{\sum_{i=1}^{n} (B_i)^2}}$$

รูปที่ 2.3 แสดงสูตรความคล้ายคลึงของโคไซน์ (cosine similarity)

1.2.2. เครื่องมือที่ใช้ในการจัดแบ่งประเภท

เครื่องมือที่สามารถนำมาใช้ในการจัดแบ่งประเภทเอกสารมีหลายชนิด เช่น WEKA, KNIME และ Rainbow ซึ่งในการศึกษาครั้งนี้ ได้สนใจที่จะใช้ Rainbow ในการจัดแบ่งประเภทเอกสารโดยอัตโนมัติ Rainbow เป็นโปรแกรมที่สร้างขึ้นเพื่อจัดแบ่งประเภทเอกสาร โดยภายใน Rainbow มีวิธีการจัดแบ่งประเภทหลายวิธี เช่น วิธีเบย์อย่างง่าย, เพื่อนบ้านใกล้ที่สุด และ TFIDF/Rocchio เป็นต้น [13]

1.3. การประเมิน

หลังจากจัดแบ่งประเภทเอกสารแล้ว จึงต้องประเมินผลจากการจัดแบ่งประเภท และประเมินประสิทธิภาพของการจัดแบ่งประเภท ซึ่งจากในการศึกษาครั้งนี้จะใช้การประเมิน 2 แบบ คือ การประเมินของการจัดแบ่งประเภทเอกสารแบบไบนารี่และการจัดแบ่งประเภทเอกสารแบบครั้งละหลายกลุ่ม

โดยในการศึกษาครั้งนี้จะจัดแบ่งประเภท 4 แบบ คือการประเมินตัวจัดแบ่งประเภทเพื่อการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษา การจัดแบ่งประเภทเอกสารข้อมูลยาตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบไบนารี่ การจัดแบ่งประเภทเอกสารข้อมูลยาตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบครั้งละหลายกลุ่ม และการจัดแบ่งประเภทเอกสารข้อมูลยาตามกลุ่มเภสัชวิทยาและการรักษาชนิดหลายกลุ่ม ซึ่งการประเมินการจัดแบ่งประเภทเอกสารแบบครั้งละสองกลุ่มจะใช้กับการจัดแบ่งประเภทเอกสารข้อมูลยาแบบไบนารี่ และการจัดแบ่งประเภทเอกสารแบบครั้งละหลายกลุ่มจะใช้กับการจัดแบ่งประเภทเอกสารข้อมูลยาชนิดหนึ่งกลุ่มและชนิดหลายกลุ่ม

1.3.1. การจัดแบ่งประเภทเอกสารแบบครั้งละสองกลุ่ม

จะสามารถเขียนผลการจัดแบ่งประเภทเป็นตาราง confusion matrix ได้ โดยตารางนี้จะเป็นตารางสรุปรวมจำนวนของผลการจัดแบ่งประเภทโดยอัตโนมัติที่สามารถจัดแบ่งประเภทได้ทั้งถูกและผิดของแต่ละกลุ่มเอกสาร (*C*) ดังตารางที่ 2.1

ตารางที่ 2.1 แสดง Confusion matrix ของผลการจัดแบ่งประเภทเอกสารแบบครั้งละสองกลุ่ม

ผลการจัดแบ่งประเภทเอกสารโดยมนุษย์	**ผลการจัดแบ่งประเภทเอกสารโดยคอมพิวเตอร์**	
	ถูก	**ผิด**
ถูก	TP	FN
ผิด	FP	TN

จำนวนของผลการจัดแบ่งประเภทของเอกสารทั้งถูกและผิดจะถูกบันทึกลงในตารางที่ 2.1 ซึ่งคำในแต่ละช่องจะมีความหมายต่างๆ ดังนี้

True positive (TP) คือ จำนวนเอกสารที่อยู่ในกลุ่มนั้น และคอมพิวเตอร์จัดแบ่งประเภทให้อยู่ในกลุ่มนั้นได้

False positive (FP) คือ จำนวนเอกสารที่ไม่อยู่ในกลุ่มนั้น แต่คอมพิวเตอร์จัดแบ่งประเภทให้อยู่ในกลุ่มนั้น

False negative (FN) คือ จำนวนเอกสารที่อยู่ในกลุ่มนั้น แต่คอมพิวเตอร์ไม่สามารถ

จัดแบ่งประเภทให้อยู่ในกลุ่มนั้นได้

True negative (TN) คือ จำนวนเอกสารที่ไม่อยู่ในกลุ่มนั้น และคอมพิวเตอร์จัดแบ่งประเภทให้ไม่อยู่ในกลุ่มนั้นได้

จากตารางที่ 2.1 นี้ สามารถนำค่าจำนวนเอกสารในตารางมาคำนวณเป็นค่าความถูกต้อง (accuracy) ความแม่นยำ (precision) ค่าการเรียกคืน (recall) และค่า F1 ได้ ดังรูปที่ 2.4, 2.5, 2.6 และ 2.7

$$Acc = \frac{TP + TN}{TP + FP + TN + FN}$$

รูปที่ 2.4 แสดงการคำนวณค่าความถูกต้อง

$$Pr = \frac{TP}{TP + FP}$$

รูปที่ 2.5 แสดงการคำนวณค่าความแม่นยำ

$$Re = \frac{TP}{TP + FN}$$

รูปที่ 2.6 แสดงการคำนวณค่าการเรียกคืน

$$F_1 = \frac{2 \times Pr \times Re}{Pr + Re}$$

รูปที่ 2.7 แสดงการคำนวณค่า F1

ค่าความถูกต้อง (accuracy, Acc) คือ ค่าที่ได้จากผลการจัดแบ่งประเภทเอกสารโดยอัตโนมัติที่คอมพิวเตอร์สามารถจัดแบ่งได้อย่างถูกต้องเทียบกับจำนวนเอกสารทั้งหมด

ความแม่นยำ (precision, Pr) คือ ค่าที่ได้จากผลการจัดแบ่งประเภทเอกสารโดยอัตโนมัติที่คอมพิวเตอร์สามารถจัดแบ่งประเภทเอกสารที่อยู่ในกลุ่มนั้นได้อย่างถูกต้องเทียบกับจำนวนเอกสารทั้งหมดที่คอมพิวเตอร์จัดแบ่งประเภทให้อยู่ในกลุ่มนั้น

ค่าการเรียกคืน (recall, Re) คือ ค่าที่ได้จากผลการจัดแบ่งประเภทเอกสารโดยอัตโนมัติที่คอมพิวเตอร์สามารถจัดแบ่งประเภทเอกสารที่อยู่ในกลุ่มนั้นได้อย่างถูกต้องเทียบกับจำนวนเอกสารทั้งหมดที่มนุษย์จัดแบ่งประเภทให้อยู่ในกลุ่มนั้น

ค่า F1 เป็นค่าเฉลี่ยฮาร์โมนิคของความแม่นยำและค่าเรียกคืน ถ้าค่าความแม่นยำและค่าการเรียกคืนสูง ค่า F1 ก็จะสูงด้วย ซึ่งค่า F1 ที่ดีจะมีค่าใกล้เคียงหรือเท่ากับ 1 และค่า F1 ที่ไม่ดีจะมีค่าใกล้เคียงหรือเท่ากับ 0 [5]

1.3.2. การจัดแบ่งประเภทเอกสารแบบครั้งละหลายกลุ่ม

จะสามารถเขียนผลการจัดแบ่งประเภทเป็นตาราง confusion matrix ได้ โดยตารางนี้จะเป็นตารางสรุปรวมจำนวนของผลการจัดแบ่งประเภทโดยอัตโนมัติที่สามารถจัดแบ่งประเภทได้ทั้งถูกและผิดของแต่ละกลุ่มเอกสาร (*C*) ดังตารางที่ 2.2

ตารางที่ 2.2 แสดง Confusion matrix ของผลการจัดแบ่งประเภทเอกสารแบบครั้งละหลายกลุ่ม

ผลการจัดแบ่งประเภทเอกสารโดยมนุษย์	**ผลการจัดแบ่งประเภทเอกสารโดยคอมพิวเตอร์**		
	กลุ่ม A	**กลุ่ม B**	**กลุ่ม C**
กลุ่ม A	TP_A	e_{AB}	e_{AC}
กลุ่ม B	e_{BA}	TP_B	e_{BC}
กลุ่ม C	e_{CA}	e_{BC}	TP_C

จำนวนของผลการจัดแบ่งประเภทของเอกสารทั้งถูกและผิดจะถูกบันทึกลงในตารางที่ 1 ซึ่งคำศัพท์ในแต่ละช่องจะมีความหมายต่างๆ ดังนี้

True positive (TP) คือ จำนวนเอกสารที่อยู่ในกลุ่มนั้น และคอมพิวเตอร์จัดแบ่งประเภทให้อยู่ในกลุ่มนั้นได้

Error (e) คือ จำนวนเอกสารที่คอมพิวเตอร์จัดแบ่งประเภทให้ผิดกลุ่ม เช่น e_{AB} เป็นจำนวนเอกสารที่อยู่ในกลุ่ม A แต่คอมพิวเตอร์จัดแบ่งให้เอกสารเหล่านี้อยู่ในกลุ่ม B เป็นต้น

จากตารางที่ 1 นี้ สามารถนำค่าจำนวนเอกสารในตารางมาคำนวณเป็นค่าความถูกต้อง (accuracy) ความแม่นยำ (precision) ค่าการเรียกคืน (recall) และค่า F1 ได้ ดังรูปที่ 2.8, 2.9, 2.10 และ 2.11

$$Acc = \frac{TP_A + TP_B + TP_C}{TP_A + TP_B + TP_C + e_{AB} + e_{AC} + e_{BA} + e_{BC} + e_{CA} + e_{CB}}$$

รูปที่ 2.8 แสดงการคำนวณค่าความถูกต้อง

$$Pr_A = \frac{TP_A}{TP_A + e_{BA} + e_{CA}}$$

รูปที่ 2.9 แสดงการคำนวณค่าความแม่นยำของกลุ่ม A

$$Re_A = \frac{TP_A}{TP_A + e_{AB} + e_{AC}}$$

รูปที่ 2.10 แสดงการคำนวณค่าการเรียกคืนของกลุ่ม A

$$F_{1A} = \frac{2 \times Pr_A \times Re_A}{Pr_A + Re_A}$$

รูปที่ 2.11 แสดงภาพการคำนวณค่า F1 ของกลุ่ม A

ค่าความถูกต้อง (accuracy, Acc) คือ ค่าที่ได้จากผลการจัดแบ่งประเภทเอกสารโดยอัตโนมัติที่คอมพิวเตอร์สามารถจัดแบ่งได้อย่างถูกต้องเทียบกับจำนวนเอกสารทั้งหมด

ความแม่นยำ (precision, Pr) คือ ค่าที่ได้จากผลการจัดแบ่งประเภทเอกสารโดยอัตโนมัติที่คอมพิวเตอร์สามารถจัดแบ่งประเภทเอกสารที่อยู่ในกลุ่มนั้นได้อย่างถูกต้องเทียบกับจำนวนเอกสารทั้งหมดที่คอมพิวเตอร์จัดแบ่งประเภทให้อยู่ในกลุ่มนั้น เช่น ในรูปที่ 2.9 เป็นภาพแสดงการคำนวณค่าความแม่นยำของกลุ่ม A หากต้องการหาค่าความแม่นยำของการจัดแบ่งประเภทจะต้องใช้ค่าความแม่นยำของทุกกลุ่มเฉลี่ยกัน

ค่าการเรียกคืน (recall, Re) คือ ค่าที่ได้จากผลการจัดแบ่งประเภทเอกสารโดยอัตโนมัติที่คอมพิวเตอร์สามารถจัดแบ่งประเภทเอกสารที่อยู่ในกลุ่มนั้นได้อย่างถูกต้องเทียบกับจำนวนเอกสารทั้งหมดที่มนุษย์จัดแบ่งประเภทให้อยู่ในกลุ่มนั้น เช่น ในรูปที่ 2.10 เป็นภาพแสดงการคำนวณค่าการเรียกคืนของกลุ่ม A หากต้องการหาค่าการเรียกคืนของการจัดแบ่งประเภทจะต้องใช้ค่าการเรียกคืนของทุกกลุ่มเฉลี่ยกัน

ค่า F1 เป็นค่าเฉลี่ยฮาร์โมนิคของความแม่นยำและค่าเรียกคืน ถ้าค่าความแม่นยำและค่าการเรียกคืนสูง ค่า F1 ก็จะสูงด้วย ซึ่งค่า F1 ที่ดีจะมีค่าใกล้เคียงหรือเท่ากับ 1 และค่า F1 ที่ไม่ดีจะมีค่าใกล้เคียงหรือเท่ากับ 0 ในรูปที่ 2.11 เป็นภาพแสดงการคำนวณค่า F1 ของกลุ่ม A หากต้องการหาค่า F1 ของการจัดแบ่งประเภทจะต้องใช้ค่า F1 ของทุกกลุ่มเฉลี่ยกัน [5]

2. เอกสารข้อมูลยา (Drug monograph)

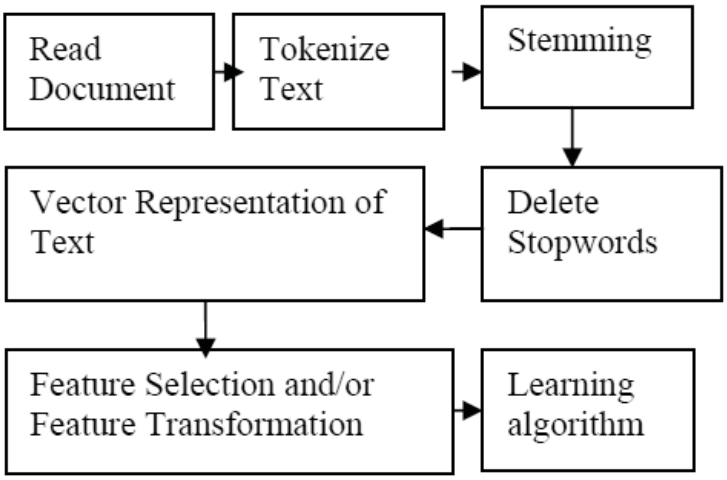
ปัจจุบันมีข้อมูลยาเพิ่มขึ้นมากมาย และมีการบันทึกข้อมูลยาเป็นเอกสาร เพื่อสะดวกในการเก็บข้อมูลและค้นคว้าหาข้อมูล การบันทึกข้อมูลยาแบบเอกสารข้อมูลยาหรือ Drug monograph เป็นการบันทึกในลักษณะของหัวข้อของเนื้อหาที่สำคัญ เช่น ชื่อสามัญ ชื่อการค้า วิธีใช้ กลไกการออกฤทธิ์ อาการข้างเคียง และการเกิดพิษ โดยหนังสือที่มีการรวบรวมข้อมูลยาเป็นจำนวนมากและนิยมใช้มากในการค้นคว้าหาข้อมูลที่สำคัญ คือ หนังสือ AHFS Drug information ซึ่งมีการจัดแบ่งประเภทของยาออกเป็นกลุ่มตามเภสัชวิทยาและการรักษา (Pharmacologic-Therapeutic classification) ต่างๆ เช่น Anti-infective agents, Central nervous system agents, Autonomic drugs, Hormones and synthetic substitutes, Cardiovascular drugs และ Skin and mucous membrane agents เป็นต้น ซึ่งการจัดแบ่งประเภทของยาตามกลุ่มเภสัชวิทยาและการรักษาของ

หนังสือเล่มนี้ได้มีการจัดแบ่งประเภทออกเป็น 4 ระดับ โดยในระดับที่ 1 มีจำนวนกลุ่มเภสัชวิทยาและการรักษา 31 กลุ่ม รายชื่อกลุ่มเภสัชวิทยาและการรักษาในระดับที่ 1 แสดงในภาพผนวก ก.1 และตัวอย่างเอกสารข้อมูลยาแสดงในภาพผนวก ก.2 นอกจากนี้การจัดแบ่งประเภทในหนังสือนี้ยังเป็นการจัดแบ่งประเภทแบบยาหนึ่งตัวอยู่ได้หลายกลุ่มอีกด้วย เนื่องจากมียาบางชนิดที่สามารถอยู่ได้หลายกลุ่มการรักษา เพราะมีฤทธิ์การรักษาที่หลากหลาย ดังนั้นเราจึงสนใจที่จะทดสอบการจัดแบ่งประเภทเอกสารตามกลุ่มเภสัชวิทยาและการรักษาของหนังสือ AHFS Drug information 2012 [14]

3. งานวิจัยที่เกี่ยวข้อง

ด้วยข้อมูลยาที่มีจำนวนเพิ่มมากขึ้นในปัจจุบัน ทำให้มีการเก็บบันทึกไว้เป็นเอกสารเพื่อการค้นหาความรู้หรือข้อมูลที่ต้องการ แต่จากเอกสารข้อมูลยาที่มีจำนวนมากทำให้ต้องใช้การจัดแบ่งประเภทเอกสาร แต่การจัดแบ่งประเภทเอกสารโดยมนุษย์จะทำให้ต้องใช้เวลานาน และเนื่องจากเป็นข้อมูลยาซึ่งเป็นข้อมูลเฉพาะ จึงต้องใช้ผู้เชี่ยวชาญทางด้านยาในการจัดแบ่งประเภท ดังนั้นการจัดแบ่งประเภทเอกสารโดยอัตโนมัติจึงมีประโยชน์มากสำหรับเอกสารข้อมูลยา แต่เนื่องจากการจัดแบ่งประเภทเอกสารในด้านยายังมีไม่มากนัก ผู้จัดทำจึงได้สนใจการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติ ซึ่งจากงานวิจัยต่างๆ พบว่า มีงานวิจัยที่เกี่ยวข้องกับการจัดแบ่งประเภทดังนี้

ในปี ค.ศ. 2005 M Ikonomakis, Sotiris Kotsiantis และ Vasilis Tampakas ได้ศึกษาการจัดแบ่งประเภทเอกสารโดยใช้เทคนิคการเรียนรู้ด้วยเครื่อง การจัดแบ่งประเภทเอกสารได้รวมถึงการจัดแบ่งประเภทแบบหัวข้อ (topic based text classification) และการจัดแบ่งประเภทตามชนิด (text genre-based classification) ตัวอย่างของชนิดของเอกสาร เช่น บทความวิทยาศาสตร์ ข่าวกีฬา และบทวิจารณ์ภาพยนตร์ เป็นต้น กระบวนการการจัดแบ่งประเภทเอกสารแสดงในรูปที่ 2.12 ในงานวิจัยฉบับนี้ได้กล่าวถึง การแทนเอกสารด้วยเวกเตอร์สเปซ (vector space document representations) การคัดเลือกคุณลักษณะ (feature selection) การปรับเปลี่ยนคุณลักษณะ (feature transformation) เทคนิคการเรียนรู้ด้วยเครื่อง ซึ่งรวมถึงวิธีเบย์อย่างง่าย ซัพพอร์ตเวกเตอร์แมชชีน เพื่อนบ้านใกล้ที่สุด และการประเมินผลด้วยค่าความถูกต้อง ค่าความแม่นยำ และค่าการเรียกคืน [15]



รูปที่ 2.12 แสดงภาพกระบวนการจัดแบ่งประเภทเอกสาร

ในปีค.ศ. 2012 Vandana Korde และ C Namrata Mahender ได้สำรวจถึงเรื่องการจัดแบ่งประเภทเอกสารและตัวจัดแบ่งประเภทอีกครั้ง ในงานวิจัยได้กล่าวถึงความหมายของการจัดแบ่งประเภทเอกสารซึ่งเป็นส่วนสำคัญในการทำเหมืองข้อความ โดยได้กล่าวถึงวิธีการจัดแบ่งประเภทเอกสารเมื่อมีการกำหนดกลุ่มให้กับเอกสารไว้ก่อน Vandana และ C Namrata ได้กล่าวว่ามีกระบวนการในการจัดแบ่งประเภทเอกสารทั้งหมด 6 ขั้นตอน คือ

1. การรวบรวมข้อมูลเอกสาร (Document Collection) ในเอกสารที่มีนามสกุลไฟล์ .html, .pdf และ .doc เป็นต้น
2. การเตรียมข้อมูลก่อนประมวลผล (Pre-Processing) เช่น การตัดคำจากข้อความ (Tokenization) การตัดคำหยุด การเปลี่ยนเป็นรากศัพท์ (Stemming word)
3. การทำดัชนี (Indexing)
4. การคัดเลือกคุณลักษณะ (Feature selection)
5. การจัดแบ่งประเภท โดยมีวิธีการจัดแบ่งประเภทหลายวิธี เช่น ต้นไม้ตัดสินใจ (Decision Tree), เพื่อนบ้านใกล้ที่สุด k ตัว (K-nearest neighbor) และซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machines) เป็นต้น
6. การประเมินผล เช่น การใช้ค่าความแม่นยำ ค่าการเรียกคืน และค่าความถูกต้อง

ในงานวิจัยชิ้นนี้ ได้กล่าวถึงตัวจัดแบ่งประเภทที่มีใช้ในการจัดแบ่งประเภทเอกสารทั้งหมด 12 วิธี คือ อัลกอริทึมของร็อคคิโอ (Rocchio's algorithm), เพื่อนบ้านใกล้ที่สุด k ตัว, เบย์อย่างง่าย, ต้นไม้ตัดสินใจ, กฎการตัดสินใจ (decision rule), ซัพพอร์ตเวกเตอร์แมชชีน, โครงข่ายประสาทเทียม (neural network), การถดถอยเชิงเส้นกำลังสองน้อยที่สุด (linear least squares fit), การลงคะแนนเสียง (voting), กฎความสัมพันธ์ (associative classifier), ตัวจัดแบ่งประเภทแบบเซนทรอยด์และตัวจัดแบ่งประเภทที่มีการเพิ่มเติม (additional classifier) นอกจากนี้ผู้วิจัยยังได้เปรียบเทียบข้อดีและข้อเสียของตัวจัดแบ่งประเภทไว้อีกด้วย [16]

Susan Dumais และคณะได้ศึกษาถึงอัลกอริทึมของการเรียนรู้โดยอุปนัยและตัวแทนสำหรับการจัดแบ่งประเภทเอกสาร โดยได้เปรียบเทียบอัลกอริทึมของการเรียนรู้โดยอุปนัยจำนวน 5 อัลกอริทึม คือ วิธีค้นหาจากความคล้าย (find similar), เบย์อย่างง่าย (naïve Bayes), เครือข่าย (bayesian networks) และซัพพอร์ตเวกเตอร์แมชชีน โดยได้ทดลองในเรื่องของความเร็วของการเรียนรู้ ความเร็วในขณะที่จัดแบ่งประเภท และความถูกต้องของการจัดแบ่งประเภท การจัดแบ่งประเภทเอกสารจะสร้างแบบจำลองโดยการเรียนรู้จากชุดข้อมูลสอนตามคำกล่าวที่ว่า "ฉันรู้เมื่อฉันเห็น (I know it when I see it)" ซึ่งจะง่ายต่อการสร้างและการปรับให้เป็นปัจจุบันเนื่องจากไม่ต้องมีทักษะด้านการเขียนโปรแกรมและการเขียนกฎ ดังนั้นจึงง่ายต่อผู้ใช้ในการกำหนดกลุ่มหรือหมวดหมู่ ในการทดลองนี้ใช้ชุดข้อมูล Reuters-21578 (ModApte split) จาก Reuters ซึ่งใช้เอกสารทั้งหมด 12,902 เรื่อง จัดประเภทให้อยู่ในกลุ่มทั้งหมด 118 กลุ่ม โดยแต่ละเอกสารจะมีคำศัพท์เฉลี่ย 200 คำต่อเอกสาร เมื่อทดลองจัดแบ่งประเภทแล้วพบว่า วิธีค้นหาจากความคล้ายสามารถเรียนรู้จากชุดข้อมูลสอนได้เร็วที่สุด รองลงมาคือซัพพอร์ตเวกเตอร์แมชชีนและเบย์อย่างง่ายตามลำดับ แต่เมื่อมีการใช้ SMO ร่วมกับซัพพอร์ตเวกเตอร์แมชชีน จะพบว่าซัพพอร์ตเวกเตอร์แมชชีนสามารถเรียนรู้ได้เร็วที่สุด ในเรื่องความเร็วของการจัดแบ่งประเภทเอกสารอัตโนมัตินั้น วัดจากค่า breakeven point (คือค่าเฉลี่ยของค่าความแม่นยำและค่าการเรียกคืน) พบว่าทุกวิธีสามารถจัดแบ่งประเภทได้รวดเร็วมาก และความเร็วในขั้นตอนการเรียนรู้จะใช้เวลานานกว่าขั้นตอนการจัดแบ่งประเภท ส่วนเรื่องความถูกต้องนั้น วิธีซัพพอร์ตเวกเตอร์แมชชีนมีความถูกต้องมากที่สุด รองลงมาคือต้นไม้ตัดสินใจและเครือข่ายเบย์ ตามลำดับ นอกจากนี้ยังได้ศึกษาในเรื่องของขนาดตัวอย่าง เมื่อทดลองใช้ขนาดตัวอย่างทั้งหมดคิดเป็น 100% และลดขนาดลงให้เหลือ10% 5% และ 1% จะพบว่าค่า breakeven point

จะลดลงไปด้วยตามลำดับ โดยขนาดตัวอย่างที่เหมาะสมจะต้องมีจำนวน 20 ตัวอย่าง ซึ่งจะทำให้การแสดงผลของการจัดแบ่งประเภทเอกสารโดยอัตโนมัติมีค่าคงที่ [8]

Alberto Holts และคณะได้ทำการศึกษาการจัดแบ่งประเภทแบบไบนารี่โดยใช้หลักการเรียนรู้ด้วยเครื่อง ในงานวิจัยนี้สนใจวิธีการจัดแบ่งประเภทแบบเบย์อย่างง่ายและซัพพอร์ตเวกเตอร์แมชชีน โดยใช้การให้น้ำหนักคำของตัวแทนของเอกสาร 5 ชนิด คือ ไบนารี่ (binary) ความถี่ (frequency) *tf* *tf·idf* และ *tf·rf* เครื่องมือที่ใช้ในการจัดแบ่งประเภทด้วยวิธีซัพพอร์ตเวกเตอร์แมชชีน คือ LibSVM และใช้ *m-estimator* สำหรับวิธีเบย์อย่างง่าย ในการศึกษานี้จะใช้เอกสารภาษาสเปนจาก Chilean newspaper La Tercera จำนวน 1,000 เอกสารจากกลุ่มกีฬา โลก ธุรกิจ และชิลี โดยใช้เอกสารจำนวน 2/3 เป็นชุดข้อมูลสอนและ 1/3 เป็นชุดข้อมูลทดสอบ และเอกสารภาษาอังกฤษจาก ModApte split ของชุดข้อมูล Reuters-21578 โดยมีจำนวนเอกสารชุดข้อมูลสอนและชุดข้อมูลทดสอบเท่ากันกับเอกสารภาษาสเปน จากผลการทดลองพบว่าทั้งสองภาษาสามารถจัดแบ่งประเภทของเอกสารได้ดีใกล้เคียงกันและการให้น้ำหนักคำแต่ละแบบไม่มีความแตกต่างอย่างมีนัยสำคัญในการจัดแบ่งประเภทแบบไบนารี่ [17]

Son Doan และคณะได้ศึกษาบทบาทของชื่อ (name entities) ในการจัดแบ่งประเภทของรายงานข่าวโรคระบาด โดยมีสมมุติฐานว่ากลุ่มที่บทบาทมีความสำคัญต่อชื่อจะทำให้มีความหมายของความรู้มากขึ้น ซึ่งชุดข้อมูลรายงานข่าวจะถูกเก็บอยู่ในเหมืองข้อมูล BioCaster โดยจะมีชื่อทั้งหมด 18 ชื่อ และบทบาท 3 บทบาท คือ case, transmission และ therapeutic พร้อมทั้งกำหนดบทบาทให้กับชื่อดังรูปที่ 2.13 และจัดแบ่งประเภทเอกสารเป็น 4 กลุ่มคือ alert, publish, check และ reject ซึ่งเอกสารที่ถูกจัดประเภทให้อยู่ในกลุ่ม alert, publish และ check จะเป็นเอกสารที่บทบาทมีความสำคัญต่อชื่อ (relevant) และเอกสารที่ถูกจัดประเภทให้อยู่ในกลุ่ม reject จะเป็นเอกสารที่บทบาทไม่มีความสำคัญต่อชื่อ (reject)

NE	Role	NE	Role(s)
PERSON	Case	ANATOMY	Transmission
ORGANIZATION	None	SYMPTOM	None
LOCATION	None	CONTROL	None
TIME	None	CHEMICAL	Therapeutic
DISEASE	None	BACTERIA	None
CONDITION	None	PRODUCT	Transmission, therapeutic
NONHUMAN	Transmission	DNA	None
VIRUS	None	RNA	None
OUTBREAK	None	PROTEIN	None

รูปที่ 2.13 แสดงกลุ่มคำและบทบาทในเหมืองข้อมูล BioCaster

จากเอกสารรายงานข่าวทั้งหมด 1,000 เอกสาร สามารถจัดแบ่งเอกสารโดยมนุษย์ได้โดยจัดเป็นเอกสารที่บทบาทมีความสำคัญต่อชื่อจำนวน 350 เอกสาร และเอกสารที่บทบาทไม่มีความสำคัญต่อชื่อ 650 เอกสาร การแบ่งชุดข้อมูลสอนและชุดข้อมูลทดสอบใช้วิธีการตรวจสอบแบบไขว้ 10 รอบ (10-fold cross-validation) และใช้วิธีเบย์อย่างง่ายและซัพพอร์ตเวกเตอร์แมชชีนในการจัดแบ่งประเภทเอกสาร โดยใช้ความถี่ของคำเป็นการให้น้ำหนักแก่คำและลิเนียร์เคอร์เนิล (linear kernel) เพื่อสร้างแบบจำลองซัพพอร์ตเวกเตอร์แมชชีน ในการทดลองนี้ใช้ระบบปฏิบัติการลีนุกซ์ (Linux OS) และชุดเครื่องมือโบ (Bow toolkit) สำหรับการจัดแบ่งประเภทด้วยเบย์อย่างง่ายและซัพพอร์ตเวกเตอร์แมชชีน นอกจากนี้งานวิจัยชิ้นนี้ยังได้ศึกษาคุณลักษณะของตัวแทน โดยศึกษาคุณลักษณะของบทบาท คุณลักษณะของชื่อและบทบาท และคุณลักษณะของกลุ่มคำสำหรับการจัดแบ่งประเภทตามความหมาย โดยกลุ่มคำที่ศึกษามีตัวอย่างคือ doctor, nurse, physician และ surgeon ซึ่งเป็นกลุ่มคำนามที่ถูกจัดให้อยู่ในประเภทของความหมายด้านอาชีพทางการแพทย์ medical_occupation และ detect, find, discover และ confirm เป็นกลุ่มคำกริยาที่ถูกจัดให้อยู่ในประเภท detect_verbs ซึ่งมีกลุ่มคำนามจำนวน 25 คำและกลุ่มคำกริยาจำนวน 49 คำ หลังจากทดลองจัดแบ่งประเภทแล้วจึงประเมินผลด้วยค่าความถูกต้อง ค่าความแม่นยำ ค่าการเรียกคืน และค่า F นอกจากนี้ยังประเมินการตรวจสอบแบบไขว้ 10 รอบด้วย t-test ($p<0.05$) จากผลการศึกษาพบว่า การใช้ชื่อทั้งหมดร่วมกับบทบาทและกลุ่มคำนามพร้อมทั้งกลุ่มคำกริยาจะทำให้จัดแบ่งประเภทได้ดีที่สุด สำหรับบทบาททั้งสามกรณี เป็นบทบาทที่เพิ่มการแสดงออกในการจัดแบ่งประเภทได้ดี ส่วน transmission จะเพิ่มในขณะที่ therapeutic จะทำให้การจัดแบ่งประเภทลดลงเล็กน้อย ในการทดลองนี้ได้กล่าวว่าวิธีซัพพอร์ตเวกเตอร์แมชชีนสามารถจัดแบ่งประเภทเอกสารได้ดีกว่าเบย์อย่างง่าย

แต่เบย์อย่างง่ายทำให้ค่าการเรียกคืนสูงกว่าซัพพอร์ตเวกเตอร์แมชชีน ดังนั้นการจัดแบ่งประเภทเอกสารโดยอัตโนมัติด้วยวิธีเบย์อย่างง่ายจึงนิยมใช้อย่างกว้างขวาง [9]

Grigorios Tsomakas และ Ioannis Katakis ได้รวบรวมความรู้เกี่ยวกับการจัดแบ่งประเภทเอกสารชนิดหลายกลุ่ม (multi-label classification) ในอดีตมีการจัดแบ่งประเภทเอกสารชนิดหนึ่งกลุ่ม (single-label classification) แต่มีบางเอกสารจะต้องกำหนดให้อยู่ได้มากกว่าหนึ่งกลุ่ม เช่น บทความในหนังสือพิมพ์ได้กล่าวถึงการฉายภาพยนตร์เรื่อง *Da Vinci Code* ของโบสถ์คริสเตียน ซึ่งสามารถจัดแบ่งประเภทได้ทั้งกลุ่มสังคม/ศาสนา และศิลปะ/ภาพยนตร์ นอกจากนี้ยังมีการวินิจฉัยโรคซึ่งผู้ป่วยในตัวอย่างเป็นทั้งโรคเบาหวานและโรคต่อมลูกหมากในขณะเดียวกัน ในงานวิจัยชิ้นนี้ได้กล่าวถึงวิธีของการจัดแบ่งประเภทชนิดหลายกลุ่ม 2 แบบคือ วิธีการเปลี่ยนแปลงปัญหา (problem transformation methods) และวิธีการดัดแปลงอัลกอริทึม (algorithm adaptation methods) [18]

Araken M Santos, Anne M P Canuto และ Antonino Feitosa Neto ได้ศึกษาวิเคราะห์เปรียบเทียบวิธีการจัดแบ่งประเภทของการจัดแบ่งประเภทชนิดหลายกลุ่มในการประยุกต์ใช้ในด้านต่างๆ การจัดแบ่งประเภทชนิดหลายกลุ่ม มี 2 วิธีคือวิธีการเปลี่ยนแปลงปัญหา (problem transformation methods) และวิธีการดัดแปลงอัลกอริทึม (algorithm adaptation methods) ในวิธีแรกคือการเปลี่ยนแปลงปัญหานี้ เดิมการจัดแบ่งประเภทมีเฉพาะการจัดแบ่งประเภทชนิดหนึ่งกลุ่ม ดังนั้นจึงเปลี่ยนแปลงปัญหาการจัดแบ่งประเภทโดยเปลี่ยนชุดข้อมูลจากชนิดหลายกลุ่มให้เป็นชุดข้อมูลชนิดหนึ่งกลุ่มแทน ซึ่งใช้วิธีการเปลี่ยนแปลงแบบต่างๆ ดังนี้ Binary Relevance (BR), Label Powerset (LP), Random k-labelsets (RAkEL), Classifier Chains (CC), Pruned Sets (PS), Ensembles of Classifier Chains (ECC) และ Ensembles of Pruned Sets (EPS) และทดสอบโดยใช้วิธีการจัดแบ่งประเภท 5 วิธี คือ วิธีเพื่อนบ้านใกล้ที่สุด, ต้นไม้ตัดสินใจ, ซัพพอร์ตเวกเตอร์แมชชีน, เบย์อย่างง่าย และ multilayer perceptron (MLP) พร้อมทั้งใช้การวัดประเมินผลทั้ง 11 แบบ คือ HammingLoss, ค่าความถูกต้อง, ค่าความแม่นยำ, ค่าการเรียกคืน, ค่า F1, ค่าความถูกต้องในซับเซต, Micro-F1, Macro-F1, One-error, ค่าการครอบคลุม (coverage) และค่าความแม่นยำเฉลี่ย และในส่วนของวิธีการดัดแปลงอัลกอริทึมในการศึกษานี้ได้เปรียบเทียบ Multi-Label k Nearest Neighbours (ML-kNN) และ Back-Propagation Multi-Label Learning (BPMLL) พบว่าการเปลี่ยนแปลงปัญหาเมื่อใช้ Binary Relevance (BR), Label Powerset (LP) จะ

ทำให้มีการเรียนรู้ที่ดี แต่ไม่แตกต่างกับวิธีอื่นๆ อย่างมีนัยสำคัญ ส่วนตัวจัดแบ่งประเภทที่ดีที่สุดคือซัพพอร์ตเวกเตอร์แมชชีน ซึ่งให้ผลการจัดแบ่งประเภทที่ดีที่สุดในทุกชุดข้อมูล สำหรับวิธีการดัดแปลงอัลกอริทึม ML-kNN เป็นวิธีที่ดีที่สุดในทุกชุดข้อมูลซึ่งให้ผลที่แตกต่างกับ BPMLL อย่างมีนัยสำคัญ และเมื่อเปรียบเทียบระหว่างวิธีการเปลี่ยนแปลงปัญหาและวิธีการดัดแปลงอัลกอริทึมพบว่า วิธีการดัดแปลงอัลกอริทึมจะให้ผลที่ดีกว่าวิธีการเปลี่ยนแปลงปัญหา โดยแตกต่างอย่างมีนัยสำคัญ ดังนั้นวิธีการดัดแปลงอัลกอริทึมจะเป็นวิธีที่ดีที่สุดในการจัดแบ่งประเภทเอกสารชนิดหลายกลุ่มเมื่อเปรียบเทียบกับวิธีการเปลี่ยนแปลงปัญหา [19]

Thanaruk Theeramunkong และ Verayuth Lertnattee ได้ศึกษาการจัดแบ่งประเภทเอกสารแบบหลายมิติ (multi-dimensional text classification) เดิมการจัดแบ่งประเภทเอกสารจะมีการจัดกลุ่มของเอกสารเป็นแบบเรียบ (flat category) และแบบลำดับขั้น (hierarchical category) แต่การจัดแบ่งประเภทเอกสารจำนวนมากจะมีความซับซ้อนมาก ซึ่งอาจไม่สามารถจัดกลุ่มให้เป็นแบบเรียบหรือแบบลำดับขั้นได้ ดังนั้นผู้วิจัยจึงทดสอบการจัดแบ่งประเภทเอกสารแบบหลายมิติ ซึ่งแต่ละเอกสารจะถูกกำหนดให้อยู่ในชุดของกลุ่ม โดยใช้ตัวจัดแบ่งประเภทคือวิธีเพื่อนบ้านใกล้ที่สุด, เบย์อย่างง่าย และวิธีแบบเซนทรอยด์ และชุดข้อมูลที่ใช้คือ WebKB และ Drug information collection ซึ่งเป็นชุดข้อมูลที่มีการจัดกลุ่มเป็นแบบหลายมิติ ในการทดลองได้จัดแบ่งประเภทเอกสารแบบเรียบ จัดแบ่งประเภทเอกสารแบบลำดับขั้น และจัดแบ่งประเภทเอกสารแบบหลายมิติ พบว่าใน WebKB วิธีการจัดแบ่งประเภทเอกสารแบบหลายมิติจะให้ผลดีกว่าวิธีการจัดแบ่งประเภทเอกสารแบบเรียบและวิธีการจัดแบ่งประเภทเอกสารแบบลำดับขั้น แต่ใน Drug information collection บางครั้งวิธีการจัดแบ่งประเภทเอกสารแบบลำดับขั้นจะสามารถจัดแบ่งประเภทได้ดีกว่าวิธีอื่น [20]

นอกจากนี้ Thanaruk Theeramunkong และ Verayuth Lertnattee ได้ศึกษาเกี่ยวกับการจัดแบ่งประเภทเอกสารข้อมูลยาในหลายมิติ โดยศึกษาทั้งหัวข้อของยา (drug topics) และกลุ่มการรักษาด้วยวิธีเพื่อนบ้านใกล้ที่สุด, เบย์อย่างง่าย และวิธีแบบเซนทรอยด์ โดยการจัดแบ่งประเภททั้งแบบ เรียบและแบบลำดับขั้น ซึ่งพบว่าการจัดแบ่งประเภทแบบหลายมิติ มีความเหมาะสมมากที่จะนำมาใช้ในงานเอกสารทางข้อมูลยา [21] อย่างไรก็ตาม งานวิจัยชิ้นนี้ได้ศึกษาถึงการจัดแบ่งประเภทเอกสารโดยอัตโนมัติในหลายมิติ และการจัดแบ่งประเภทยังเป็นการจัดกลุ่มตามสรีรวิทยาอยู่มาก ซึ่งจากการใช้ข้อมูลยาของเภสัชกรและบุคลากรทางการแพทย์นั้นจะมีการใช้ประโยชน์จากการ

จัดกลุ่มตามเภสัชวิทยาและการรักษามากกว่า ดังนั้นผู้วิจัยจึงได้สนใจที่จะศึกษาการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษา

บทที่ 3
วิธีดำเนินการวิจัย

ในการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามเภสัชวิทยาและการรักษานี้ มีวิธีทดลอง 3 ขั้นตอนใหญ่ คือ การคัดเลือกเอกสารข้อมูลยา การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติ และการประเมินผลการจัดแบ่งประเภทและประสิทธิภาพตัวจัดแบ่งประเภท ซึ่งมีรายละเอียดดังต่อไปนี้

1. การคัดเลือกเอกสารข้อมูลยา

1.1. การค้นหาและคัดเลือกแหล่งข้อมูลยา

ค้นหาแหล่งข้อมูลยาที่มีเอกสารข้อมูลยาตามคุณลักษณะที่ต้องการ โดยวิธีดังนี้

1.1.1. ใช้ตัวสืบค้นข้อมูล www.google.com และใช้คำสำคัญคือ “drug monograph” จะได้แหล่งข้อมูลมาจำนวนหนึ่ง

1.1.2. คัดเลือกแหล่งข้อมูลโดยดูจากข้อมูลภายในแหล่งข้อมูลที่ต้องการ โดยข้อมูลภายในแหล่งข้อมูลจะต้องมีลักษณะเป็นเอกสารข้อมูลยา (drug monograph) และมีนามสกุล .htm หรือ .html

1.1.3. ดึงข้อมูลจากแหล่งข้อมูลโดยใช้ HTTrack ซึ่งเป็นโปรแกรมที่สามารถดาวน์โหลดข้อมูลจากเว็บไซต์มาเก็บไว้ที่คอมพิวเตอร์ได้

1.1.4. เนื่องจากแหล่งข้อมูลยามีจำนวนข้อมูลยาที่กระจายหลากหลายขนาด จึงค้นหาเฉพาะแหล่งข้อมูลที่มีจำนวนข้อมูลยาประมาณ 2,000-5,000 ไฟล์ ซึ่งเป็นจำนวนข้อมูลยาที่ทำให้ได้แหล่งข้อมูลที่มีขนาดใกล้เคียงกัน

1.1.5. คัดเลือกหาชื่อยาที่มีเหมือนกันในแต่ละเว็บไซต์ที่ได้

จากการค้นหาและคัดเลือกแหล่งข้อมูลยา จึงได้แหล่งข้อมูลยาจำนวน 2 เว็บไซต์ คือ RxList (http://www.rxlist.com) และ DailyMed (http://dailymed.nlm.nih.gov)

1.2. การคัดกรองเอกสารข้อมูลยา

จากแหล่งข้อมูลยาที่ได้ทั้ง 2 เว็บไซต์ คือ RxList และ DailyMed ให้คัดกรองเอกสารข้อมูลยาที่มีคุณลักษณะที่ต้องการจากแหล่งข้อมูลยา โดยเลือกเอกสารข้อมูลยาที่มียาตัวเดียวกันในทั้ง 2 แหล่งข้อมูล เช่น diclofenac จะต้องมีทั้งใน RxList และ DailyMed โดยเอกสารที่คัดกรองมาได้นี้ จะมีเอกสารข้อมูลยาที่สามารถจัดแบ่งประเภทตามกลุ่มเภสัชวิทยาและการรักษาได้ทั้งแบบ 1 กลุ่ม 2 กลุ่ม และ 3 กลุ่ม

เอกสารข้อมูลยาที่คัดกรองมาได้จะสามารถจัดแบ่งประเภทตามกลุ่มเภสัชวิทยาและการรักษาได้จำนวน 11 กลุ่ม ซึ่งอ้างอิงจากหนังสือ AHFS Drug Information 2012 ดังตารางที่ 3.1 โดยยาทุกตัวจะสามารถจัดแบ่งประเภทให้อยู่ได้ในกลุ่มใดกลุ่มหนึ่ง และยาบางตัวจะสามารถจัดแบ่งประเภทให้อยู่ได้มากที่สุด 3 กลุ่ม มีตารางแสดงจำนวนเอกสารข้อมูลยาในแต่ละกลุ่มตามกลุ่มเภสัชวิทยาและการรักษา ตารางที่ 3.2 แสดงจำนวนเอกสารข้อมูลยาที่สามารถจัดแบ่งประเภทให้อยู่ในกลุ่มเภสัชวิทยาและการรักษาได้เอกสารละ 1 กลุ่ม หรือเรียกเอกสารชนิด 1 กลุ่ม ตารางที่ 3.3 แสดงจำนวนเอกสารข้อมูลยาที่สามารถจัดแบ่งประเภทให้อยู่ในกลุ่มเภสัชวิทยาและการรักษาได้เอกสารละ 2 กลุ่ม หรือเรียกเอกสารชนิด 2 กลุ่ม ตารางที่ 3.4 แสดงจำนวนเอกสารข้อมูลยาที่สามารถจัดแบ่งประเภทให้อยู่ในกลุ่มเภสัชวิทยาและการรักษาได้เอกสารละ 3 กลุ่ม หรือเรียกเอกสารชนิด 3 กลุ่ม ตารางที่ 3.5 แสดงจำนวนเอกสารข้อมูลยาในเอกสารชนิด 1 กลุ่ม, 2 กลุ่ม และ 3 กลุ่ม

ตารางที่ 3.1 แสดงกลุ่มเภสัชวิทยาและการรักษาจากหนังสือ AHFS Drug Information 2012 จำนวน 11 กลุ่มจากเอกสารข้อมูลยาคัดกรองได้

ลำดับ	กลุ่มเภสัชวิทยาและการรักษา
1	Anti-infective agents
2	Antineoplastic drugs
3	Autonomics drugs
4	Blood formulation, coagulation, and thrombosis agents
5	Cardiovascular
6	CNS
7	EENT preparations
8	Gastrointestinal drugs
9	Hormones and synthetic substitutes
10	Skin and mucous membrane
11	Vitamins

ตารางที่ 3.2 แสดงจำนวนเอกสารข้อมูลยาตามกลุ่มเภสัชวิทยาและการรักษาชนิด 1 กลุ่ม

กลุ่มเภสัชวิทยาและการรักษา	จำนวนเอกสารข้อมูลยา
Anti-infective agents	38
Antineoplastic drugs	48
Autonomics drugs	40
Blood formulation, coagulation, and thrombosis agents	38
Cardiovascular	62
CNS	42
EENT preparations	12
Gastrointestinal drugs	46
Hormones and synthetic substitutes	36
Skin and mucous membrane	20
Vitamins	20
รวม	402

ตารางที่ 3.3 แสดงจำนวนเอกสารข้อมูลยาตามกลุ่มเภสัชวิทยาและการรักษาชนิด 1 และ 2 กลุ่ม

กลุ่มเภสัชวิทยาและการรักษา	**จำนวนเอกสารข้อมูลยา**
Anti-infective agents	50
Antineoplastic drugs	52
Autonomics drugs	48
Blood formulation, coagulation, and thrombosis agents	38
Cardiovascular	62
CNS	50
EENT preparations	38
Gastrointestinal drugs	50
Hormones and synthetic substitutes	50
Skin and mucous membrane	40
Vitamins	20
รวม	498

ตารางที่ 3.4 แสดงจำนวนเอกสารข้อมูลยาตามกลุ่มเภสัชวิทยาและการรักษาชนิด 1, 2 และ 3 กลุ่ม

กลุ่มเภสัชวิทยาและการรักษา	**จำนวนเอกสารข้อมูลยา**
Anti-infective agents	54
Antineoplastic drugs	52
Autonomics drugs	48
Blood formulation, coagulation, and thrombosis agents	38
Cardiovascular	62
CNS	50
EENT preparations	48
Gastrointestinal drugs	50
Hormones and synthetic substitutes	56
Skin and mucous membrane	50
Vitamins	20
รวม	528

ตารางที่ 3.5 แสดงจำนวนเอกสารข้อมูลยาในเอกสารข้อมูลยาชนิด 1, 2 และ 3 กลุ่ม

ประเภทของเอกสารข้อมูลยา	จำนวนเอกสารข้อมูลยา
ชนิด 1 กลุ่ม	402
ชนิด 2 กลุ่ม	48
ชนิด 3 กลุ่ม	10
รวม	460

2. การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติ

ในการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติ ได้ศึกษาใน 4 วิธี คือ การประเมินตัวจัดแบ่งประเภทเพื่อการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษา การจัดแบ่งประเภทเอกสารข้อมูลยาตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบไบนารี่ การจัดแบ่งประเภทเอกสารข้อมูลยาตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบครั้งละหลายกลุ่ม และการจัดแบ่งประเภทเอกสารข้อมูลยาตามกลุ่มเภสัชวิทยาและการรักษาชนิดหลายกลุ่ม ซึ่งมีขั้นตอนการจัดแบ่งประเภทเอกสาร 4 ขั้นตอน ดังนี้

1. เตรียมข้อมูลโดย

การตัดคำหยุด เช่น a, an ,the เป็นต้น

การตัดโครงสร้าง HTML เช่น <html>, <b> เป็นต้น

หลังจากตัดคำหยุดและตัดโครงสร้าง HTML แล้วจะเหลือคำศัพท์ของชุดข้อมูลที่ใช้ในการจัดแบ่งประเภทจำนวน 29,571 คำ

การให้น้ำหนักของคำ ซึ่งในแต่ละตัวจัดแบ่งประเภทจะมีการให้น้ำหนักที่แตกต่างกัน ดังนี้

วิธีเบย์อย่างง่าย ใช้ทฤษฎีความน่าจะเป็นแบบเบย์

วิธีแบบเซนทรอยด์ ใช้ tfidf/sd และใช้ความคล้ายคลึงของโคไซน์

วิธีเพื่อนบ้านใกล้ที่สุด โดย k เป็น 3 และใช้ความคล้ายคลึงของโคไซน์

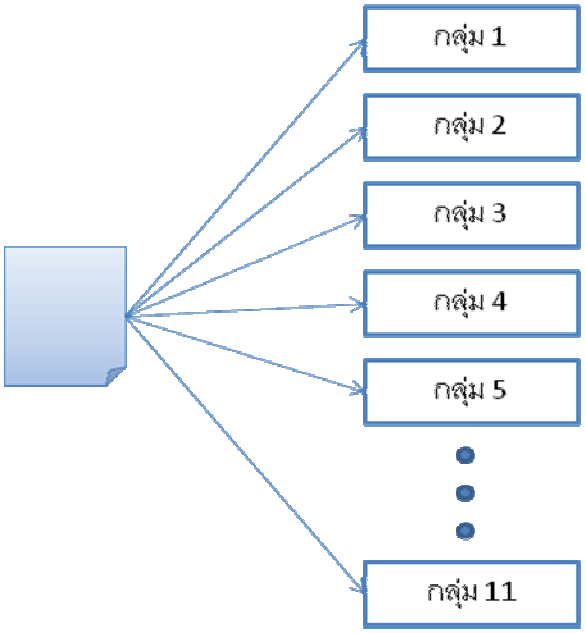
2. จัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติทั้ง 3 วิธี โดยใช้ Rainbow
3. ทดสอบด้วย 5-fold cross validation

4. นำผลที่ได้ไปประเมินและสรุปผล

จากขั้นตอนการจัดแบ่งประเภทเอกสาร 4 ขั้นตอน ซึ่งนำมาใช้ในการทดสอบการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษา 4 วิธี ดังต่อไปนี้

2.1. การประเมินตัวจัดแบ่งประเภทเพื่อการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษา

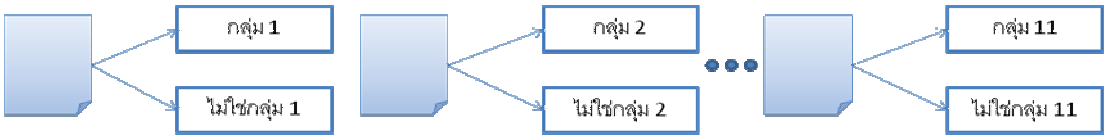
การประเมินตัวจัดแบ่งประเภทนี้ จัดทำขึ้นเพื่อทดสอบหาตัวจัดแบ่งประเภทที่ดีในการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษา จึงใช้ตัวจัดแบ่งประเภท 3 วิธี คือ วิธีเบย์อย่างง่าย วิธีเพื่อนบ้านใกล้ที่สุด และวิธีแบบเซนทรอยด์มาทดสอบจัดแบ่งประเภทเอกสารข้อมูลยาตามกลุ่มเภสัชวิทยาและการรักษา โดยศึกษาในเอกสารข้อมูลยาที่สามารถจัดแบ่งประเภทให้อยู่ในกลุ่มเภสัชวิทยาและการรักษาได้เอกสารละ 1 กลุ่มหรือเรียกเอกสารข้อมูลยาชนิด 1 กลุ่ม จำนวน 402 เอกสาร จำนวนเอกสารข้อมูลยาในแต่ละกลุ่มเภสัชวิทยาและการรักษาชนิด 1 กลุ่ม แสดงในตารางที่ 3.2 และใช้วิธีแบบครั้งละหลายกลุ่มในการทดสอบ ภาพวิธีการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบครั้งละหลายกลุ่มแสดงดังรูปที่ 3.1



รูปที่ 3.1 แสดงภาพการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาชนิด 1 กลุ่มโดยวิธีแบบครั้งละหลายกลุ่ม

2.2. การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบไบนารี่

การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัตินี้ ใช้วิธีแบบไบนารี่ซึ่งเป็นวิธีดั้งเดิมในการจัดแบ่งประเภทเอกสาร โดยจะทดสอบจัดแบ่งประเภทเอกสารข้อมูลยาตามกลุ่มเภสัชวิทยาและการรักษาจำนวน 11 กลุ่ม ดังตารางที่ 3.1 เพื่อจัดแบ่งประเภทเอกสารข้อมูลยาตามกลุ่มเภสัชวิทยาและการรักษาชนิดหลายกลุ่ม และประเมินข้อดีข้อเสียที่ได้ในการจัดแบ่งประเภทเอกสารข้อมูลยาแบบไบนารี่ ภาพวิธีการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบไบนารี่ แสดงดังรูปที่ 3.2 และเอกสารข้อมูลยาที่นำมาจัดแบ่งประเภทนี้จะใช้จำนวน 450 เอกสาร ซึ่งเป็นเอกสารข้อมูลยาชนิด 1 กลุ่ม จำนวน 402 เอกสาร และจำนวนเอกสารข้อมูลยาชนิด 2 กลุ่ม จำนวน 48 เอกสาร จำนวนเอกสารข้อมูลยาในแต่ละกลุ่มเภสัชวิทยาและการรักษาชนิด 1 กลุ่มและ 2 กลุ่มแสดงในตารางที่ 3.3 โดยในการจัดแบ่งประเภทเอกสารนี้ จะมีเอกสารข้อมูลยาชนิด 2 กลุ่มเป็นตัวแทนในเอกสารข้อมูลยาชนิดหลายกลุ่ม และการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาแบบไบนารี่นี้ จะใช้วิธีแบบเซนทรอยด์ในการจัดแบ่งประเภท



รูปที่ 3.2 แสดงภาพการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบไบนารี่

หลังจากจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาชนิดหลายกลุ่มโดยใช้เอกสารข้อมูลยาชนิด 2 กลุ่มเป็นตัวแทนด้วยวิธีแบบครั้งละหลายกลุ่มแล้ว จึงสามารถนำผลมาเปรียบเทียบกับผลของการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาชนิดหลายกลุ่มโดยใช้เอกสารข้อมูลยาชนิด 2 กลุ่มเป็นตัวแทนด้วยวิธีแบบไบนารี่ได้ เพื่อเปรียบเทียบการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัช

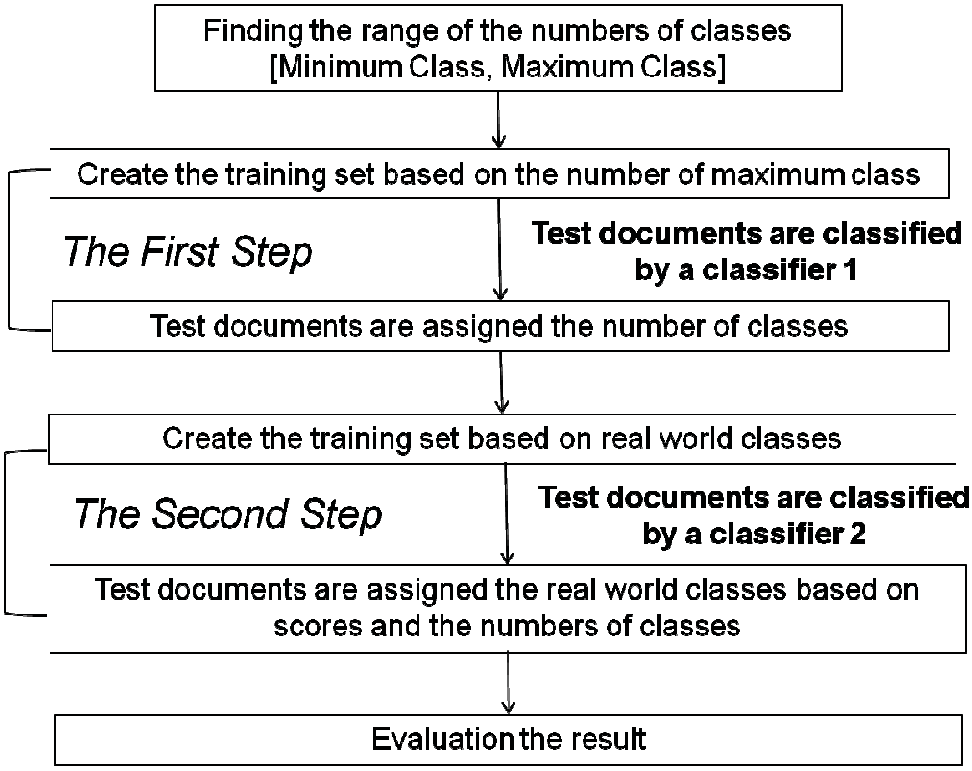
วิทยาและการรักษาแบบครั้งละหลายกลุ่มและแบบไบนารี่ และประเมินข้อดีข้อเสียที่ได้ในการจัดแบ่งประเภทเอกสารข้อมูลยาแบบครั้งละหลายกลุ่มและไบนารี่

2.3. การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบครั้งละหลายกลุ่ม

จากวิธีที่ 2.1 ได้ทำการประเมินตัวจัดแบ่งประเภทเพื่อการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาไปแล้ว โดยใช้เอกสารข้อมูลยาชนิด 1 กลุ่มและวิธีแบบครั้งละหลายกลุ่มในการทดสอบ เพื่อจัดแบ่งประเภทเอกสารข้อมูลยาตามเภสัชวิทยาและการรักษาชนิด 1 กลุ่ม ดังนั้นในขั้นตอนนี้จึงทดสอบโดยใช้เอกสารชนิดหลายกลุ่ม ซึ่งจะใช้เอกสารชนิด 2 กลุ่มเป็นตัวแทนของเอกสารชนิดหลายกลุ่ม เพื่อจัดแบ่งประเภทเอกสารข้อมูลยาตามกลุ่มเภสัชวิทยาและการรักษาชนิดหลายกลุ่ม และประเมินข้อดีข้อเสียที่ได้ในการจัดแบ่งประเภทเอกสารข้อมูลยาแบบครั้งละหลายกลุ่ม จำนวนเอกสารข้อมูลยาที่ใช้ศึกษามีจำนวน 450 เอกสาร ซึ่งเป็นเอกสารข้อมูลยาประเภท 1 กลุ่ม จำนวน 402 เอกสาร และเอกสารข้อมูลยาประเภท 2 กลุ่ม จำนวน 48 เอกสาร จำนวนเอกสารข้อมูลยาในแต่ละกลุ่มเภสัชวิทยาและการรักษาชนิด 1 กลุ่ม, 2 กลุ่ม และ 3 กลุ่มแสดงในตารางที่ 3.3 และจัดแบ่งประเภทโดยวิธีแบบครั้งละหลายกลุ่ม ซึ่งตัวจัดแบ่งประเภททั้ง 3 วิธี คือวิธีเบย์อย่างง่าย วิธีเพื่อนบ้านใกล้ที่สุดและวิธีแบบเซนทรอยด์ในการจัดแบ่งประเภทเพื่อหาตัวจัดแบ่งประเภทที่ดีที่สุดในการจัดแบ่งประเภทเอกสารข้อมูลยาชนิดหลายกลุ่มที่มีเอกสารชนิด 2 กลุ่มเป็นตัวแทน

2.4. การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาชนิดหลายกลุ่ม

การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัตินี้ จะศึกษาในเอกสารข้อมูลยาชนิดหลายกลุ่ม ซึ่งจะมีการจัดแบ่งประเภทเอกสารแบบ 2 ขั้นตอน (two-step classification) ดังรูปที่ 3.3 ซึ่งแต่ละขั้นจะมีวิธีการศึกษาที่แตกต่างกันคือ



รูปที่ 3.3 แสดงภาพการจัดแบ่งประเภทเอกสารแบบ 2 ขั้นตอน

ขั้นที่ 1 การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามจำนวนกลุ่ม [a, b]

เมื่อกำหนดให้ [a, b] คือช่วงของจำนวนกลุ่มที่จัดแบ่งประเภทได้

a คือ จำนวนกลุ่มที่น้อยที่สุดที่เอกสารสามารถจัดแบ่งประเภทอยู่ได้

b คือ จำนวนกลุ่มที่มากที่สุดที่เอกสารสามารถจัดแบ่งประเภทอยู่ได้

ในขั้นที่ 1 นี้จะจัดทำเพื่อหาจำนวนกลุ่มของเอกสารข้อมูลยาที่สามารถจัดแบ่งประเภทตามกลุ่มเภสัชวิทยาและการรักษาได้ และทดสอบหาตัวจัดแบ่งประเภทที่ดีที่สุดของขั้นที่ 1 นี้ โดยทดสอบจากตัวจัดแบ่งประเภท 3 วิธีคือ วิธีเบย์อย่างง่าย วิธีเพื่อนบ้านใกล้ที่สุด และวิธีแบบเซนทรอยด์ในการจัดแบ่งประเภท เมื่อได้ตัวจัดแบ่งประเภทที่ดีที่สุดแล้วจึงใช้ตัวจัดแบ่งประเภทนั้นจัดแบ่งประเภทเอกสารข้อมูลยาเพื่อหาจำนวนกลุ่มตามกลุ่มเภสัชวิทยาและการรักษา

ขั้นที่ 2 การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาจำนวน 11 กลุ่ม

หลังจากที่จัดแบ่งประเภทเอกสารข้อมูลยาในขั้นที่ 1 แล้ว จะได้จำนวนกลุ่มที่เอกสารข้อมูลยาสามารถจัดแบ่งประเภทได้ และนำเอกสารข้อมูลยาที่จัดจำนวนกลุ่มได้ถูกต้องมาใช้ทดสอบต่อในขั้นที่ 2 ซึ่งจะเป็นการจัดแบ่งประเภทของเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษา เพื่อทดสอบหากลุ่มเภสัชวิทยาและการรักษาที่เอกสารข้อมูลยาสามารถจัดแบ่งประเภทได้โดยจำนวนกลุ่มตามกลุ่มเภสัชวิทยาและการรักษาที่นำมาศึกษามีจำนวน 11 กลุ่มจากหนังสือ AHFS Drug Information 2012 และทดสอบหาตัวจัดแบ่งประเภทที่ดีที่สุดของขั้นที่ 2 นี้จากตัวจัดแบ่งประเภท 3 วิธีคือ วิธีเบย์อย่างง่าย วิธีเพื่อนบ้านใกล้ที่สุด และวิธีแบบเซนทรอยด์ เมื่อได้ตัวจัดแบ่งประเภทที่ดีที่สุดแล้วจึงใช้ตัวจัดแบ่งประเภทนั้นเพื่อจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษา

การจัดแบ่งประเภทเอกสารข้อมูลยาแบบ 2 ขั้นตอนจะศึกษาในการจัดแบ่งประเภทเอกสารข้อมูลยาชนิด 1 กลุ่ม, 2 กลุ่ม และ 3 กลุ่ม โดยวิธีแรกจะศึกษาในเอกสารข้อมูลยาชนิด 1 กลุ่มและ 2 กลุ่ม ส่วนวิธีที่ 2 จะศึกษาในเอกสารข้อมูลยาชนิด 1 กลุ่ม, 2 กลุ่ม และ 3 กลุ่ม

2.4.1. การจัดแบ่งประเภทเอกสารแบบ 2 ขั้นตอนในเอกสารชนิด 1 กลุ่มและชนิด 2 กลุ่ม

ในการศึกษานี้ จะทดสอบในเอกสารข้อมูลยาชนิด 1 กลุ่มและชนิด 2 กลุ่ม เอกสารข้อมูลยาที่นำมาศึกษานี้จะสามารถจัดแบ่งประเภทตามกลุ่มเภสัชวิทยาและการรักษาให้อยู่ได้น้อยที่สุดจำนวน 1 กลุ่ม และมากที่สุดจำนวน 2 กลุ่มตามหนังสือ AHFS Drug Information 2012 โดยในขั้นที่ 1 จะจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามจำนวนกลุ่ม [1, 2] และขั้นที่ 2 จะจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาจำนวน 11 กลุ่ม เอกสารที่ใช้ทดสอบจำนวน 450 เอกสาร แบ่งเป็นเอกสารชนิด 1 กลุ่ม จำนวน 402 เอกสาร และเอกสารชนิด 2 กลุ่ม จำนวน 48 เอกสาร

2.4.2. การจัดแบ่งประเภทเอกสารแบบ 2 ขั้นตอนในเอกสารชนิด 1 กลุ่ม, ชนิด 2 กลุ่ม และชนิด 3 กลุ่ม

ในการศึกษานี้ จะทดสอบในเอกสารข้อมูลยาชนิด 1 กลุ่ม, 2 กลุ่ม และ 3 กลุ่ม เอกสารข้อมูลยาที่นำมาศึกษานี้จะสามารถจัดแบ่งประเภทตามกลุ่มเภสัชวิทยาและการรักษาให้อยู่ได้น้อยที่สุด 1 กลุ่ม และมากที่สุด 3 กลุ่มของหนังสือ AHFS Drug Information

2012 โดยในขั้นที่ 1 จะจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามจำนวนกลุ่ม [1, 3] และขั้นที่ 2 จะจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาจำนวน 11 กลุ่ม เอกสารที่ใช้ทดสอบจำนวน 460 เอกสาร แบ่งเป็นเอกสารชนิด 1 กลุ่ม จำนวน 402 เอกสาร เอกสารชนิด 2 กลุ่ม จำนวน 48 เอกสาร และเอกสารชนิด 3 กลุ่ม จำนวน 10 เอกสาร

3. การประเมินผลการจัดแบ่งประเภทและประสิทธิภาพตัวจัดแบ่งประเภท

การจัดแบ่งประเภทเอกสารโดยอัตโนมัติทั้ง 3 วิธีคือ การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบไบนารี่ การจัดแบ่งประเภทเอกสารข้อมูลยาตามกลุ่มเภสัชวิทยาและการรักษาชนิดหนึ่งกลุ่มโดยวิธีแบบครั้งละหลายกลุ่ม และการจัดแบ่งประเภทเอกสารข้อมูลยาตามกลุ่มเภสัชวิทยาและการรักษาชนิดหลายกลุ่ม ซึ่งการประเมินประสิทธิภาพของทั้ง 3 แบบ จะใช้การประเมินจากค่าความถูกต้อง (accuracy) ความแม่นยำ (precision) ค่าการเรียกคืน (recall) และค่า F1

ในการจัดแบ่งประเภทเอกสารข้อมูลยาตามกลุ่มเภสัชวิทยาและการรักษาชนิดหลายกลุ่ม จะมีการประเมินแตกต่างจากจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติชนิดหนึ่งกลุ่ม เนื่องจากมีขั้นตอนการจัดแบ่งประเภท 2 ขั้นตอน ดังนั้นในขั้นที่ 1 จึงประเมินจากผลการจัดแบ่งประเภทที่สามารถจัดแบ่งจำนวนกลุ่มของเอกสารข้อมูลยาได้ถูกต้อง และในขั้นที่ 2 ซึ่งเป็นการจัดแบ่งประเภทตามกลุ่มเภสัชวิทยาและการรักษานั้น เนื่องจากบางเอกสารจะสามารถจัดแบ่งได้จำนวน 2 กลุ่มและ 3 กลุ่ม ดังนั้นจึงประเมินเฉพาะจำนวนเอกสารที่สามารถจัดแบ่งประเภทได้ถูกต้องทุกกลุ่มเท่านั้น เรียกการประเมินแบบนี้ว่า การประเมินการจัดแบ่งประเภทเอกสารแบบสมบูรณ์ (complete classification) โดยใช้การประเมินจากค่าความถูกต้อง ความแม่นยำ ค่าการเรียกคืน และค่า F1

ในการประเมินผลการจัดแบ่งประเภทเอกสารนี้ จะมีทั้งการประเมินในด้านจำนวนเอกสารที่จัดแบ่งประเภทตามกลุ่มเภสัชวิทยาและการรักษา และการประเมินตัวจัดแบ่งประเภทเอกสารโดยเปรียบเทียบจากค่า F1 เฉลี่ย

บทที่ 4
ผลการวิเคราะห์ข้อมูล

การจัดแบ่งประเภทเอกสารโดยอัตโนมัติมี 4 วิธีคือ การประเมินตัวจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษา การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบไบนารี่ การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบครั้งละหลายกลุ่ม และการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาชนิดหลายกลุ่ม มีผลการวิเคราะห์ดังต่อไปนี้

1. การประเมินตัวจัดแบ่งประเภทเพื่อการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษา

ในการประเมินตัวจัดแบ่งประเภทนี้ จะใช้ตัวจัดแบ่งประเภท 3 วิธีคือ วิธีเบย์อย่างง่าย วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k = 3 และวิธีแบบเซนทรอยด์ ทดสอบในเอกสารชนิด 1 กลุ่ม จำนวน 402 เอกสาร โดยใช้วิธีแบบครั้งละหลายกลุ่มในการจัดแบ่งประเภท จะได้ผลการทดลองของการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบครั้งละหลายกลุ่มด้วยวิธีเบย์อย่างง่าย วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k = 3 และวิธีแบบเซนทรอยด์เป็นตัวจัดแบ่งประเภทจะได้ ผลการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาแสดงในตารางที่ 4.1, 4.2 และ 4.3 การเปรียบเทียบค่าความถูกต้องของการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีครั้งละหลายกลุ่มในการจัดแบ่งประเภทเอกสารข้อมูลยาวิธีเบย์อย่างง่าย วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k = 3 และวิธีแบบเซนทรอยด์ แสดงในตารางที่ 4.4 พบว่าจากเอกสารข้อมูลยาชนิด 1 กลุ่ม จำนวน 402 เอกสาร จะสามารถจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาด้วยวิธีเบย์อย่างง่ายได้ถูกต้องจำนวน 369

เอกสาร จัดแบ่งประเภทผิดจำนวน 33 เอกสาร คิดเป็นค่าความถูกต้อง 91.79% วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k=3 จะจัดแบ่งได้ถูกต้องจำนวน 367 เอกสาร จัดแบ่งประเภทผิดจำนวน 35 เอกสาร คิดเป็นค่าความถูกต้อง 91.29% วิธีแบบเซนทรอยด์ได้ถูกต้องจำนวน 378 เอกสาร คิดเป็นค่าความถูกต้อง 94.03% ดังนั้นจะเห็นได้ว่าเมื่อเปรียบเทียบจากค่าความถูกต้อง วิธีแบบเซนทรอยด์จะเป็นวิธีที่มีค่าความถูกต้องมากที่สุดจาก 3 วิธีในการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาของเอกสารชนิด 1 กลุ่มโดยวิธีครั้งละหลายกลุ่ม

จากผลการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบครั้งละหลายกลุ่มสามารถคำนวณเป็นค่าความแม่นยำ ค่าการเรียกคืน และค่า F1 ได้ และมีผลการประเมินด้วยค่าความแม่นยำ ค่าการเรียกคืนของการจัดแบ่งประเภทด้วยวิธีเบย์อย่างง่าย วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k = 3 และวิธีแบบเซนทรอยด์ แสดงในตารางที่ 4.5, 4.6 และ 4.7 การเปรียบเทียบค่า F1 ของการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบครั้งละหลายกลุ่มด้วยวิธีเบย์อย่างง่าย วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k = 3 และวิธีแบบเซนทรอยด์เป็นตัวจัดแบ่งประเภท แสดงในตารางที่ 4.8 พบว่าวิธีการจัดแบ่งประเภททั้ง 3 วิธีมีค่า F1 เฉลี่ย 0.92 จะเห็นได้ว่าวิธีการจัดแบ่งประเภททั้ง 3 วิธี สามารถใช้จัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาชนิด 1 กลุ่มโดยวิธีครั้งละหลายกลุ่มได้ ซึ่งในวิธีเบย์อย่างง่ายมีค่า F1 เฉลี่ย 0.90 วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k=3 มีค่า F1 เฉลี่ย 0.91 และวิธีแบบเซนทรอยด์มีค่า F1 เฉลี่ย 0.94 จากค่า F1 เฉลี่ยของแต่ละตัวจัดแบ่งประเภทจะเห็นได้ว่าวิธีแบบเซนทรอยด์มีค่า F1 เฉลี่ยมากที่สุด จึงเป็นวิธีที่ดีที่สุดจาก 3 วิธีในการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาชนิด 1 กลุ่มโดยวิธีครั้งละหลายกลุ่ม

เมื่อพิจารณาในด้านแต่ละกลุ่มของกลุ่มเภสัชวิทยาและการรักษาจะพบว่าแต่ละกลุ่มในตารางที่ 4.8 จากกลุ่มเภสัชวิทยาและการรักษาจำนวน 11 กลุ่ม จะสามารถจัดแบ่งประเภทเอกสารได้ดีโดยมีค่า F1 เฉลี่ย 0.92 ซึ่งกลุ่มที่จัดแบ่งประเภทได้ดีที่สุดซึ่งมีค่า F1 เฉลี่ย 0.97 จำนวน 3 กลุ่มคือ กลุ่ม Anti-infective agents, กลุ่ม Blood formulation, coagulation, and thrombosis agents, และกลุ่ม Cardiovascular และกลุ่มที่จัดแบ่งประเภทได้ดีน้อยที่สุดซึ่งมีค่า F1 เฉลี่ย 0.85 จำนวน 1 กลุ่มคือ กลุ่ม EENT preparations

ตารางที่ 4.1 แสดงผลการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบครั้งละหลายกลุ่มด้วยวิธีเบย์อย่างง่าย

จำนวนเอกสารที่ถูกจัดแบ่งโดยมนุษย์	**จำนวนเอกสารที่ถูกจัดแบ่งโดยคอมพิวเตอร์**											
กลุ่มเภสัชวิทยาและการรักษา	**กลุ่มเภสัชวิทยาและการรักษา**											
	Anti-infective agents	Antineoplastic drugs	Autonomics drugs	Blood formulation, coagulation, and thrombosis agents	Cardiovascular	CNS	EENT preparations	Gastrointestinal drugs	Hormones and synthetic substitutes	Skin and mucous membrane	Vitamins	รวม
Anti-infective agents	38	0	0	0	0	0	0	0	0	0	0	38
Antineoplastic drugs	1	45	0	0	0	0	0	0	2	0	0	48
Autonomics drugs	0	0	35	0	1	1	1	1	0	1	0	40
Blood formulation, coagulation, and thrombosis agents	0	0	0	34	2	0	2	0	0	0	0	38
Cardiovascular	0	0	0	0	62	0	0	0	0	0	0	62
CNS	0	0	1	2	0	37	1	1	0	0	0	42
EENT preparations	0	0	0	0	1	0	11	0	0	0	0	12
Gastrointestinal drugs	0	0	2	0	1	2	2	37	0	1	1	46
Hormones and synthetic substitutes	0	0	0	0	0	0	0	0	36	0	0	36
Skin and mucous membrane	0	0	1	0	0	0	0	0	0	19	0	20
Vitamins	0	2	1	0	0	0	1	0	1	0	15	20
รวม	39	47	40	36	67	40	18	39	39	21	16	402

ตารางที่ 4.2 แสดงผลการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบครั้งละหลายกลุ่มด้วยวิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k=3

จำนวนเอกสารที่ถูกจัดแบ่งโดยมนุษย์	**จำนวนเอกสารที่ถูกจัดแบ่งโดยคอมพิวเตอร์**											
กลุ่มเภสัชวิทยาและการรักษา	**กลุ่มเภสัชวิทยาและการรักษา**											
	Anti-infective agents	Antineoplastic drugs	Autonomics drugs	Blood formulation, coagulation, and thrombosis agents	Cardiovascular	CNS	EENT preparations	Gastrointestinal drugs	Hormones and synthetic substitutes	Skin and mucous membrane	Vitamins	รวม
Anti-infective agents	38	0	0	0	0	0	0	0	0	0	0	38
Antineoplastic drugs	0	43	0	0	0	0	0	0	4	0	1	48
Autonomics drugs	2	0	36	0	0	1	0	1	0	0	0	40
Blood formulation, coagulation, and thrombosis agents	0	0	0	38	0	0	0	0	0	0	0	38
Cardiovascular	0	0	0	0	59	0	1	1	0	0	1	62
CNS	1	0	2	0	0	35	1	1	0	0	2	42
EENT preparations	0	0	0	0	0	0	12	0	0	0	0	12
Gastrointestinal drugs	0	0	2	0	0	7	0	35	0	2	0	46
Hormones and synthetic substitutes	0	1	0	0	0	0	0	0	35	0	0	36
Skin and mucous membrane	0	0	0	0	0	1	0	0	0	19	0	20
Vitamins	0	2	0	0	0	0	1	0	0	0	17	20
รวม	41	46	40	38	59	44	15	38	39	21	21	402

ตารางที่ 4.3 แสดงผลการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบครั้งละหลายกลุ่มด้วยวิธีแบบเซนทรอยด์

จำนวนเอกสารที่ถูกจัดแบ่งโดยมนุษย์	**จำนวนเอกสารที่ถูกจัดแบ่งโดยคอมพิวเตอร์**											
กลุ่มเภสัชวิทยาและการรักษา	**กลุ่มเภสัชวิทยาและการรักษา**											
	Anti-infective agents	Antineoplastic drugs	Autonomics drugs	Blood formulation, coagulation, and thrombosis agents	Cardiovascular	CNS	EENT preparations	Gastrointestinal drugs	Hormones and synthetic substitutes	Skin and mucous membrane	Vitamins	รวม
Anti-infective agents	38	0	0	0	0	0	0	0	0	0	0	38
Antineoplastic drugs	0	44	0	0	0	0	0	0	4	0	0	48
Autonomics drugs	2	0	37	0	0	0	0	1	0	0	0	40
Blood formulation, coagulation, and thrombosis agents	0	0	0	37	1	0	0	0	0	0	0	38
Cardiovascular	0	0	0	0	62	0	0	0	0	0	0	62
CNS	0	0	3	0	0	38	1	0	0	0	0	42
EENT preparations	0	0	0	0	0	0	12	0	0	0	0	12
Gastrointestinal drugs	1	0	2	0	1	3	1	37	0	0	1	46
Hormones and synthetic substitutes	0	0	0	0	1	0	0	0	35	0	0	36
Skin and mucous membrane	0	0	0	0	0	0	0	0	0	20	0	20
Vitamins	0	2	0	0	0	0	0	0	0	0	18	20
รวม	41	46	42	37	65	41	14	38	39	20	19	402

ตารางที่ 4. 4 แสดงการเปรียบเทียบค่าความถูกต้องของการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีครั้งละหลายกลุ่มด้วยวิธีเบย์อย่างง่าย วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k = 3 และวิธีแบบเซนทรอยด์

การประเมิน	**วิธีเบย์อย่างง่าย**	**วิธีเพื่อนบ้านใกล้ที่สุด (k=3)**	**วิธีแบบเซนทรอยด์**
จำนวนเอกสารทั้งหมดที่สามารถจัดแบ่งประเภทถูก	369	367	378
จำนวนเอกสารทั้งหมดที่สามารถจัดแบ่งประเภทผิด	33	35	24
ค่าความถูกต้อง	91.79%	91.29%	94.03%

ตารางที่ 4.5 แสดงผลการประเมินการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบครั้งละหลายกลุ่มด้วยวิธีเบย์อย่างง่าย

การประเมิน	**กลุ่มเภสัชวิทยาและการรักษา**										
	Anti-infective agents	Antineoplastic drugs	Autonomics drugs	Blood formulation, coagulation, and thrombosis agents	Cardiovascular	CNS	EENT preparations	Gastrointestinal drugs	Hormones and synthetic substitutes	Skin and mucous membrane	Vitamins
ค่าความแม่นยำ	0.97	0.96	0.88	0.94	0.93	0.93	0.61	0.95	0.92	0.90	0.75
ค่าการเรียกคืน	1.00	0.94	0.88	0.89	1.00	0.88	0.92	0.80	1.00	0.95	0.75
ค่า F1	0.99	0.95	0.88	0.92	0.96	0.90	0.73	0.87	0.96	0.93	0.83
จำนวนเอกสารทั้งหมดที่สามารถจัดแบ่งประเภทถูก	369										
จำนวนเอกสารทั้งหมดที่สามารถจัดแบ่งประเภทผิด	33										
ค่าความถูกต้อง	91.79%										

ตารางที่ 4.6 แสดงผลการประเมินการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบครั้งละหลายกลุ่มด้วยวิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k=3

การประเมิน	**กลุ่มเภสัชวิทยาและการรักษา**										
	Anti-infective agents	Antineoplastic drugs	Autonomics drugs	Blood formulation, coagulation, and thrombosis agents	Cardiovascular	CNS	EENT preparations	Gastrointestinal drugs	Hormones and synthetic substitutes	Skin and mucous membrane	Vitamins
ค่าความแม่นยำ	0.93	0.93	0.90	1.00	1.00	0.80	0.80	0.92	0.90	0.90	0.81
ค่าการเรียกคืน	1.00	0.90	0.90	1.00	0.95	0.83	1.00	0.76	0.97	0.95	0.85
ค่า F1	0.96	0.91	0.90	1.00	0.98	0.81	0.89	0.83	0.93	0.93	0.83
จำนวนเอกสารทั้งหมดที่สามารถจัดแบ่งประเภทถูก	367										
จำนวนเอกสารทั้งหมดที่สามารถจัดแบ่งประเภทผิด	35										
ค่าความถูกต้อง	91.29%										

ตารางที่ 4.7 แสดงผลการประเมินการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบครั้งละหลายกลุ่มด้วยวิธีแบบเซนทรอยด์

การประเมิน	**กลุ่มเภสัชวิทยาและการรักษา**										
	Anti-infective agents	Antineoplastic drugs	Autonomics drugs	Blood formulation, coagulation, and thrombosis agents	Cardiovascular	CNS	EENT preparations	Gastrointestinal drugs	Hormones and synthetic substitutes	Skin and mucous membrane	Vitamins
ค่าความแม่นยำ	0.93	0.96	0.88	1.00	0.95	0.93	0.86	0.97	0.90	1.00	0.95
ค่าการเรียกคืน	1.00	0.92	0.93	0.97	1.00	0.90	1.00	0.80	0.97	1.00	0.90
ค่า F1	0.96	0.94	0.90	0.99	0.98	0.92	0.92	0.88	0.93	1.00	0.92
จำนวนเอกสารทั้งหมดที่สามารถจัดแบ่งประเภทถูก	378										
จำนวนเอกสารทั้งหมดที่สามารถจัดแบ่งประเภทผิด	24										
ค่าความถูกต้อง	94.03%										

ตารางที่ 4.8 แสดงการเปรียบเทียบค่า F1 ของการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบครั้งละหลายกลุ่มด้วยด้วยวิธีเบย์อย่างง่าย วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k = 3 และวิธีแบบเซนทรอยด์

กลุ่มเภสัชวิทยาและการรักษา	**F1**			
	วิธีการจัดแบ่งประเภทเอกสาร			**เฉลี่ย**
	วิธีเบย์อย่างง่าย	**วิธีเพื่อนบ้านใกล้ที่สุด (k=3)**	**วิธีแบบเซนทรอยด์**	
Anti-infective agents	0.99	0.96	0.96	0.97
Antineoplastic drugs	0.95	0.91	0.94	0.93
Autonomics drugs	0.88	0.90	0.90	0.89
Blood formulation, coagulation, and thrombosis agents	0.92	1.00	0.99	0.97
Cardiovascular	0.96	0.98	0.98	0.97
CNS	0.90	0.81	0.92	0.88
EENT preparations	0.73	0.89	0.92	0.85
Gastrointestinal drugs	0.87	0.83	0.88	0.86
Hormones and synthetic substitutes	0.96	0.93	0.93	0.94
Skin and mucous membrane	0.93	0.93	1.00	0.95
Vitamins	0.83	0.83	0.92	0.86
เฉลี่ย	0.90	0.91	0.94	0.92

2. การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบไบนารี่

การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาจำนวน 11 กลุ่มโดยวิธีแบบไบนารี่ เมื่อใช้วิธีแบบเซนทรอยด์ในการจัดแบ่งประเภทเอกสาร ได้ผลการจัดจำนวนกลุ่มของการจัดแบ่งประเภทดังตารางที่ 4.9 เอกสารข้อมูลยาที่นำมาจัดแบ่งประเภทนี้มีจำนวนทั้งหมด 450 เอกสาร แบ่งเป็นเอกสารข้อมูลยาชนิด 1 กลุ่มจำนวน 402 เอกสาร และเอกสารข้อมูลยาชนิด 2 กลุ่มจำนวน 48 เอกสารซึ่งเป็นตัวแทนเอกสารข้อมูลยาชนิดหลายกลุ่ม พบว่ามีเอกสารข้อมูลยาชนิด 1 กลุ่มที่จัดแบ่งประเภทโดยอัตโนมัติเป็น 1 กลุ่มได้อย่างถูกต้องจำนวน 367 เอกสาร จัดแบ่งประเภทผิดจำนวน 35 เอกสาร คิดเป็นค่าความถูกต้อง 91.29% และเอกสารข้อมูลยาชนิด 2 กลุ่มที่จัดแบ่งประเภทโดยอัตโนมัติเป็น 2 กลุ่มได้อย่างถูกต้องจำนวน 29 เอกสาร จัดแบ่งประเภทผิดจำนวน 19 เอกสาร คิดเป็นค่าความถูกต้อง 60.42%

จากตารางที่ 4.9 มีเอกสารข้อมูลยาชนิด 1 กลุ่มที่จัดแบ่งประเภทโดยอัตโนมัติให้อยู่ใน 1 กลุ่มได้อย่างถูกต้องจำนวน 367 เอกสาร ซึ่งผลการจัดแบ่งประเภทตามกลุ่มเภสัชวิทยาและการรักษาแบบไบนารี่ เมื่อใช้วิธีแบบเซนทรอยด์เป็นตัวจัดแบ่งประเภท แสดงในตารางที่ 4.10 ซึ่งจัดแบ่งประเภทได้อย่างถูกต้องจำนวน 358 เอกสาร และจัดแบ่งประเภทผิดกลุ่มตามกลุ่มเภสัชวิทยาและการรักษาจำนวน 9 เอกสาร คิดเป็นค่าความถูกต้องของเอกสารข้อมูลยาชนิด 1 กลุ่มที่จัดจำนวนกลุ่มและจัดแบ่งประเภทถูก 97.55% ของเอกสารข้อมูลยาชนิด 1 กลุ่มที่จัดเป็น 1 กลุ่มได้อย่างถูกต้อง ดังนั้นจึงมีเอกสารข้อมูลยาชนิด 1 กลุ่มที่จัดแบ่งประเภทเอกสารโดยอัตโนมัติได้ถูกต้องจำนวน 358 เอกสารจาก 402 เอกสาร คิดเป็นค่าความถูกต้องทั้งหมด 89.05% ของการจัดแบ่งประเภทเอกสารข้อมูลยาชนิด 1 กลุ่ม

จากตารางที่ 4.9 มีเอกสารข้อมูลยาชนิด 2 กลุ่มที่จัดแบ่งประเภทโดยอัตโนมัติเป็น 2 กลุ่มได้อย่างถูกต้องจำนวน 29 เอกสาร ซึ่งผลการจัดแบ่งประเภทตามกลุ่มเภสัชวิทยาและการรักษาแบบไบนารี่ เมื่อใช้วิธีแบบเซนทรอยด์เป็นตัวจัดแบ่งประเภท แสดงในตารางที่ 4.11 ซึ่งจัดแบ่งประเภทตามกลุ่มเภสัชวิทยาและการรักษาได้อย่างถูกต้องจำนวน 29 เอกสาร คิดเป็นค่าความถูกต้องของเอกสารข้อมูลยาที่จัดจำนวนกลุ่มและจัดแบ่งประเภทถูก 100.00% ของเอกสารข้อมูลยาชนิด 2 กลุ่มที่จัดแบ่งประเภทโดยอัตโนมัติเป็น2 กลุ่มได้อย่างถูกต้อง ดังนั้นจึงมีเอกสารข้อมูลยาชนิด 2 กลุ่ม

ที่จัดแบ่งประเภทเอกสารโดยอัตโนมัติได้ถูกต้องจำนวน 29 เอกสารจาก 48 เอกสาร คิดเป็นค่าความถูกต้องทั้งหมด 60.42% ของการจัดแบ่งประเภทเอกสารข้อมูลยาชนิด 2 กลุ่ม

ตารางที่ 4.9 แสดงผลการจัดจำนวนกลุ่มของการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาแบบไบนารี่โดยวิธีแบบเซนทรอยด์

ประเภทเอกสาร	**เอกสารที่ถูกจัดแบ่งประเภทโดยคอมพิวเตอร์**	
	จำนวนกลุ่ม	**จำนวนเอกสาร**
เอกสารข้อมูลยา ชนิด 1 กลุ่ม จำนวน 402 เอกสาร	ไม่จัดอยู่ในกลุ่มใดกลุ่มหนึ่ง	22
	จัดอยู่ใน 1 กลุ่ม	367
	จัดอยู่ใน 2 กลุ่ม	13
	จัดอยู่ใน 3 กลุ่ม	0
เอกสารข้อมูลยา ชนิด 2 กลุ่ม จำนวน 48 เอกสาร	ไม่จัดอยู่ในกลุ่มใดกลุ่มหนึ่ง	3
	จัดอยู่ใน 1 กลุ่ม	11
	จัดอยู่ใน 2 กลุ่ม	29
	จัดอยู่ใน 3 กลุ่ม	5

ตารางที่ 4.10 แสดงผลการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาชนิดหนึ่งกลุ่มแบบไบนารี่โดยวิธีแบบเซนทรอยด์

การประเมิน	**จำนวนเอกสารข้อมูลยาชนิดหนึ่งกลุ่ม**
จำนวนเอกสารข้อมูลยาที่จัดจำนวนกลุ่มและจัดแบ่งประเภทถูก	358
จำนวนเอกสารข้อมูลยาที่จัดจำนวนกลุ่มและจัดแบ่งประเภทผิด	9
จำนวนเอกสารข้อมูลยาของการจัดจำนวนกลุ่มทั้งหมด	367
ค่าความถูกต้องของการจัดจำนวนกลุ่มทั้งหมด	91.29%
ค่าความถูกต้องของการจัดแบ่งประเภทตามกลุ่มเภสัชวิทยาและการรักษา	97.55%
ค่าความถูกต้องทั้งหมด	89.05%

ตารางที่ 4.11 แสดงผลการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาชนิด 2 กลุ่มแบบไบนารี่โดยวิธีแบบเซนทรอยด์

การประเมิน	**จำนวนและร้อยละของเอกสารข้อมูลยาชนิด 2 กลุ่ม**
จำนวนเอกสารทางยาที่จัดจำนวนกลุ่มและจัดแบ่งประเภทถูก	29
จำนวนเอกสารข้อมูลยาที่จัดจำนวนกลุ่มและจัดแบ่งประเภทผิด	0
จำนวนเอกสารข้อมูลยาของการจัดจำนวนกลุ่มทั้งหมด	29
ค่าความถูกต้องของการจัดจำนวนกลุ่มทั้งหมด	60.42%
ค่าความถูกต้องของการจัดแบ่งประเภทตามกลุ่มเภสัชวิทยาและการรักษา	100.00%
ค่าความถูกต้องทั้งหมด	60.42%

จากการศึกษาการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบไบนารี่นี้ เมื่อใช้วิธีแบบเซนทรอยด์ในการจัดแบ่งประเภท จากตารางที่ 4.12 พบว่า ในเอกสารข้อมูลยาชนิด 1 กลุ่ม มีค่าความถูกต้องในการจัดแบ่งประเภท 89.05% และเอกสารข้อมูลยาชนิด 2 กลุ่ม มีค่าความถูกต้องในการจัดแบ่งประเภท 60.42% ซึ่งจากเอกสารข้อมูลยาทั้งหมด 450 เอกสาร สามารถจัดแบ่งประเภทได้ถูกต้องจำนวน 387 เอกสาร จึงมีค่าความถูกต้องทั้งหมดของการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาด้วยวิธีแบบไบนารี่คิดเป็น 86.00% ทำให้เห็นว่าการจัดแบ่งประเภทของเอกสารข้อมูลยาชนิดหลายกลุ่มที่มีเอกสารข้อมูลยาชนิด 2 กลุ่มเป็นตัวแทนด้วยวิธีแบบไบนารี่นั้น สามารถจัดแบ่งประเภทเอกสารได้ถูกต้อง

จากตารางที่ 4.13 แสดงผลการประเมินการจัดแบ่งประเภทเอกสารโดยอัตโนมัติชนิดหลายกลุ่ม เมื่อมีเอกสารชนิด 2 กลุ่มเป็นตัวแทนโดยวิธีแบบไบนารี่ และใช้วิธีแบบเซนทรอยด์เป็นตัวจัดแบ่งประเภท โดยดูจากค่าความแม่นยำ ค่าการเรียกคืน และค่า F1 พบว่าการจัดแบ่งประเภทเอกสารนี้สามารถจัดแบ่งประเภทเอกสารได้ดีโดยมีค่าความแม่นยำเฉลี่ย 0.90 ค่าการเรียกคืนเฉลี่ย 0.85 และค่า F1 เฉลี่ย 0.88

ตารางที่ 4.12 แสดงผลการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติชนิดหลายกลุ่มตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบไบนารี่ด้วยวิธีแบบเซนทรอยด์

การประเมิน	**ร้อยละของเอกสารข้อมูลยา**
ค่าความถูกต้องทั้งหมดของเอกสารชนิด 1 กลุ่ม	89.05%
ค่าความถูกต้องทั้งหมดของเอกสารชนิด 2 กลุ่ม	60.42%
ค่าความถูกต้องทั้งหมด	86.00%

ตารางที่ 4.13 แสดงผลการประเมินการจัดแบ่งประเภทเอกสารโดยอัตโนมัติชนิดหลายกลุ่มตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีไบนารี่ด้วยวิธีแบบเซนทรอยด์

กลุ่มเภสัชวิทยาและการรักษา	**จำนวนที่จัดแบ่งประเภทถูก**	**ความแม่นยำ**	**การเรียกคืน**	**F1**
Anti-infective agents	40	0.89	0.98	0.93
Antineoplastic drugs	44	0.90	0.90	0.90
Autonomics drugs	38	0.86	0.84	0.85
Blood formulation, coagulation, and thrombosis agents	36	1.00	0.95	0.97
Cardiovascular	60	0.98	0.97	0.98
CNS	32	0.89	0.68	0.77
EENT preparations	23	0.77	0.77	0.77
Gastrointestinal drugs	36	0.94	0.78	0.86
Hormones and synthetic substitutes	34	0.89	0.81	0.85
Skin and mucous membrane	26	0.87	0.87	0.87
Vitamins	17	0.94	0.85	0.89
เฉลี่ย	386 (รวม)	0.90	0.85	0.88

3. การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบครั้งละหลายกลุ่ม

จากวิธีที่ 1 ได้ประเมินตัวจัดแบ่งประเภทเพื่อการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาซึ่งได้ใช้วิธีแบบครั้งละหลายกลุ่มและเอกสารชนิด 1 กลุ่มในการทดลอง โดยต้องการหาตัวจัดแบ่งประเภทที่ดีที่สุดจากการทดสอบตัวจัดแบ่งประเภท 3 วิธีคือ วิธีเบย์อย่างง่าย วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k=3 และวิธีแบบเซนทรอยด์ พบว่าวิธีแบบเซนทรอยด์เป็นวิธีที่ดีที่สุด ดังนั้นในวิธีที่ 2 นี้จึงจะใช้วิธีแบบเซนทรอยด์ทดสอบการจัดแบ่งประเภทโดยวิธีแบบครั้งละหลายกลุ่มในเอกสารข้อมูลยาชนิดหลายกลุ่มซึ่งใช้เอกสารชนิด 2 กลุ่มเป็นตัวแทน

เมื่อทดลองศึกษาการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาชนิดหลายกลุ่มโดยวิธีแบบครั้งละหลายกลุ่มและวิธีแบบไบนารี่แล้วสามารถเปรียบเทียบได้ดังตารางที่ 4.14 พบว่าจากเอกสารข้อมูลยาชนิด 1 กลุ่ม จำนวน 402 เอกสาร เมื่อจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบไบนารี่จะจัดแบ่งประเภทได้ถูกจำนวน 358 เอกสาร คิดเป็นค่าความถูกต้อง 89.05% ส่วนวิธีแบบครั้งละหลายกลุ่มจัดแบ่งประเภทได้ถูกจำนวน 379 เอกสาร คิดเป็นค่าความถูกต้อง 94.28% และเอกสารข้อมูลยาชนิด 2 กลุ่มจำนวน 48 เอกสาร เมื่อจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบไบนารี่จะจัดแบ่งประเภทได้ถูกจำนวน 29 เอกสาร คิดเป็นค่าความถูกต้อง 60.42% ส่วนวิธีแบบครั้งละหลายกลุ่มจัดแบ่งประเภทได้ถูกจำนวน 46 เอกสาร คิดเป็นค่าความถูกต้อง 95.83% ดังนั้นเอกสารข้อมูลยาทั้งหมด 450 เอกสาร เมื่อจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาชนิดหลายกลุ่มโดยวิธีแบบไบนารี่ จะจัดแบ่งประเภทเอกสารข้อมูลยาได้ถูกต้องจำนวน 387 เอกสาร คิดเป็นค่าความถูกต้อง 86.00% และเมื่อจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาชนิดหลายกลุ่มโดยวิธีครั้งละหลายกลุ่ม จะจัดแบ่งประเภทเอกสารข้อมูลยาได้ถูกต้องจำนวน 425 เอกสาร คิดเป็นค่าความถูกต้อง 94.44% จากการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาชนิดหลายกลุ่มในวิธีแบบไบนารี่และวิธีแบบครั้งละหลายกลุ่ม พบว่าวิธีแบบครั้งละหลายกลุ่มจะสามารถจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติได้ถูกต้องมากกว่าวิธีแบบไบนารี่ ดังนั้นหากทราบจำนวนกลุ่มที่เอกสารข้อมูลยาสามารถจัดแบ่งประเภทตามกลุ่มเภสัชวิทยาและการรักษาได้ จะทำให้วิธีแบบ

ครั้งละหลายกลุ่มเป็นวิธีที่จัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาชนิดหลายกลุ่มได้ดีกว่าวิธีแบบไบนารี่

ตารางที่ 4.14 แสดงผลการเปรียบเทียบการจัดแบ่งประเภทเอกสารข้อมูลยาตามกลุ่มเภสัชวิทยาและการรักษาชนิดหลายกลุ่มโดยวิธีแบบครั้งละหลายกลุ่มและวิธีแบบไบนารี่

การประเมิน	**วิธีแบบไบนารี่**	**วิธีแบบครั้งละหลายกลุ่ม**
เอกสารข้อมูลยาชนิดหนึ่งกลุ่มที่จัดแบ่งประเภทถูก (402)	358 (89.05%)	379 (94.28%)
เอกสารข้อมูลยาชนิดสองกลุ่มที่จัดแบ่งประเภทถูก (48)	29 (60.42%)	46 (95.83%)
เอกสารข้อมูลยาที่จัดแบ่งประเภทถูก	387 (86.00%)	425 (94.44%)

ตารางที่ 4.15 แสดงข้อดีข้อเสียของการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาชนิดหนึ่งกลุ่มโดยวิธีแบบไบนารี่และวิธีแบบครั้งละหลายกลุ่ม

วิธีแบบไบนารี่	**วิธีแบบครั้งละหลายกลุ่ม**
<u>ข้อดี</u>-จัดแบ่งประเภทเอกสารข้อมูลยาชนิดหลายกลุ่มได้	<u>ข้อดี</u>-การจัดแบ่งประเภทง่ายกว่าและใช้เวลาน้อยกว่า
<u>ข้อเสีย</u> 1. จำนวนกลุ่มที่จัดแบ่งประเภทได้จะน้อยกว่าหรือมากกว่าจำนวนกลุ่มจริง หรือไม่จัดอยู่ในกลุ่มใดเลย 2. การจัดแบ่งประเภทยุ่งยากกว่าและใช้เวลามากกว่า	<u>ข้อเสีย</u>-จัดแบ่งประเภทเอกสารข้อมูลยาชนิดหลายกลุ่มได้ยาก ต้องใช้วิธีอื่นมาช่วยในการจัดแบ่งประเภท เช่น การทำ threshold และการทำสัดส่วนคะแนน

เมื่อนำวิธีการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบไบนารี่และวิธีแบบครั้งละหลายกลุ่มมาเปรียบเทียบข้อดีข้อเสียดังตารางที่ 4.6 พบว่า

วิธีการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบไบนารี่มีข้อดีคือสามารถจัดแบ่งประเภทเอกสารข้อมูลยาชนิดหลายกลุ่มได้ ในขณะที่วิธีการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบครั้งละหลายกลุ่มจะจัดแบ่งประเภทเอกสารข้อมูลยาชนิดหลายกลุ่มได้ยาก ต้องใช้วิธีอื่นมาช่วยในการจัดแบ่งประเภท เช่น การทำ Threshold หรือการทำสัดส่วนคะแนน และข้อดีของวิธีการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบครั้งละหลายกลุ่ม คือการจัดแบ่งประเภทจะสามารถทำได้ง่ายกว่าและใช้เวลาน้อยกว่าวิธีการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบไบนารี่ ในขณะที่วิธีการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบไบนารี่จะจัดแบ่งประเภทได้ยุ่งยากกว่าและใช้เวลามากกว่าวิธีการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบครั้งละหลายกลุ่ม นอกจากนี้ข้อเสียอีกข้อหนึ่งของวิธีการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบไบนารี่คือมีจำนวนกลุ่มที่จัดแบ่งประเภทได้จะน้อยกว่าหรือมากกว่าจำนวนกลุ่มจริง หรือไม่จัดอยู่ในกลุ่มใดเลย

4. การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาชนิดหลายกลุ่ม

จากการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาโดยวิธีแบบไบนารี่และวิธีแบบครั้งละหลายกลุ่ม พบว่าหากทราบจำนวนกลุ่ม วิธีแบบครั้งละหลายกลุ่มจะเป็นวิธีที่ดีในการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาชนิดหลายกลุ่ม ดังนั้นจึงพัฒนาวิธีการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาชนิดหลายกลุ่มเป็นวิธีการจัดแบ่งประเภทเอกสารแบบ 2 ขั้นตอน (Two-step text classification) คือ

ขั้นที่ 1 การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามจำนวนกลุ่ม [a, b]

เมื่อกำหนดให้ [a, b] คือช่วงของจำนวนกลุ่มที่จัดแบ่งประเภทได้

a คือ จำนวนกลุ่มที่น้อยที่สุดที่เอกสารสามารถจัดแบ่งระเภทอยู่ได้

b คือ จำนวนกลุ่มที่มากที่สุดที่เอกสารสามารถจัดแบ่งประเภทอยู่ได้

ขั้นที่ 2 การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาจำนวน 11 กลุ่ม

การจัดแบ่งประเภทเอกสารแบบ 2 ขั้นตอนนี้ จะทดสอบในการจัดแบ่งประเภทเอกสารข้อมูลยาชนิดหลายกลุ่ม โดยได้ทำการจัดแบ่งประเภทเอกสารแบบ 2 ขั้นตอนในเอกสารข้อมูลยาชนิด 1 กลุ่มและชนิด 2 กลุ่ม และการจัดแบ่งประเภทเอกสารแบบ 2 ขั้นตอนในเอกสารข้อมูลยาชนิด 1 กลุ่ม, ชนิด 2 กลุ่ม และชนิด 3 กลุ่ม

4.1 การจัดแบ่งประเภทเอกสารแบบ 2 ขั้นตอนในเอกสารข้อมูลยาชนิด 1 กลุ่ม และชนิด 2 กลุ่ม

ในวิธีการจัดแบ่งประเภทนี้ จะใช้เอกสารข้อมูลยาทั้งหมด 450 เอกสาร แบ่งเป็นข้อมูลยาชนิด 1 กลุ่ม จำนวน 402 เอกสาร และเอกสารข้อมูลยาชนิด 2 กลุ่ม จำนวน 48 เอกสาร สามารถจัดแบ่งประเภทเอกสารแบบ 2 ขั้นตอนได้ ดังนี้

4.1.1. ขั้นที่ 1 การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามจำนวนกลุ่ม [1, 2]

เมื่อจัดแบ่งประเภทเอกสารข้อมูลยาตามจำนวนกลุ่ม [1, 2] ด้วยวิธีเบย์อย่างง่าย วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k=3 และวิธีแบบเซนทรอยด์ เพื่อหาตัวจัดแบ่งประเภทที่ดีที่สุดในการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามจำนวนกลุ่ม จะสามารถจัดแบ่งประเภทเอกสารได้ผลดังตาราง 4.16 จากเอกสารข้อมูลยา 450 เอกสาร วิธีเบย์อย่างง่าย สามารถจัดแบ่งประเภทได้ถูก 421 เอกสาร จัดแบ่งประเภทผิด 29 เอกสาร คิดเป็นค่าความถูกต้อง 93.56% วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k=3 สามารถจัดแบ่งประเภทได้ถูก 442 เอกสาร จัดแบ่งประเภทผิด 8 เอกสาร คิดเป็นค่าความถูกต้อง 98.22% และวิธีแบบเซนทรอยด์สามารถจัดแบ่งประเภทได้ถูก 441 เอกสาร จัดแบ่งประเภทผิด 9 เอกสาร คิดเป็นค่าความถูกต้อง 98.00% จะเห็นได้ว่าในขั้นที่ 1 นี้ วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k=3 เป็นวิธีที่มีค่าความถูกต้องมากที่สุด

ตารางที่ 4.16 ผลของขั้นที่ 1 การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามจำนวนกลุ่ม [1, 2]

การประเมิน	**วิธีเบย์อย่างง่าย**	**วิธีเพื่อนบ้านใกล้ที่สุด (k=3)**	**วิธีแบบเซนทรอยด์**
จำนวนเอกสารทั้งหมดที่สามารถจัดแบ่งประเภทถูก	421	442	441
จำนวนเอกสารทั้งหมดที่สามารถจัดแบ่งประเภทผิด	29	8	9
ค่าความถูกต้อง	93.56%	98.22%	98.00%

จากผลการประเมินการจัดแบ่งประเภทเอกสารข้อมูลยาตามจำนวนกลุ่ม [1, 2] จะแสดงในค่าความแม่นยำ ค่าการเรียกคืน และค่า F1 ดังตารางที่ 4.17 พบว่า วิธีเบย์อย่างง่ายมีค่า F1 เฉลี่ย 0.86 วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k=3 มีค่า F1 เฉลี่ย 0.96 และวิธีแบบเซนทรอยด์มีค่า F1 เฉลี่ย 0.95 ดังนั้นจึงพบว่า วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k=3 เป็นตัวจัดแบ่งประเภทที่ดีที่สุดจาก 3 วิธีของการจัดแบ่งประเภทเอกสารแบบ 2 ขั้นตอน ในขั้นที่ 1 จัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามจำนวนกลุ่ม [1, 2]

ตารางที่ 4.17 แสดงผลการประเมินผลของขั้นที่ 1 การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามจำนวนกลุ่ม [1, 2]

การประเมิน	**วิธีการจัดแบ่งประเภทเอกสาร**								
	วิธีเบย์อย่างง่าย			**วิธีเพื่อนบ้านใกล้ที่สุด (k=3)**			**วิธีแบบเซนทรอยด์**		
	ประเภทเอกสาร			**ประเภทเอกสาร**			**ประเภทเอกสาร**		
	1 กลุ่ม	**2 กลุ่ม**	**เฉลี่ย**	**1 กลุ่ม**	**2 กลุ่ม**	**เฉลี่ย**	**1 กลุ่ม**	**2 กลุ่ม**	**เฉลี่ย**
ค่าความแม่นยำ	0.99	0.64	0.82	0.99	0.90	0.95	0.99	0.90	0.95
ค่าการเรียกคืน	0.94	0.94	0.94	0.99	0.94	0.96	0.99	0.91	0.95
ค่า F1	0.96	0.76	0.86	0.99	0.92	0.96	0.99	0.90	0.95

4.1.2. ขั้นที่ 2 การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาจำนวน 11 กลุ่ม

เมื่อจัดแบ่งประเภทเอกสารข้อมูลยาตามจำนวนกลุ่ม [1, 2] แล้ว จึงได้ตัวจัดแบ่งประเภทที่ดีที่สุดของขั้นที่ 1 คือ วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k=3 ดังนั้นเมื่อจัดแบ่งประเภทเอกสารด้วยวิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k=3 แล้วจะมีเอกสารข้อมูลยาที่จัดแบ่งประเภทได้ถูกต้องจากขั้นที่ 1 จำนวน 442 เอกสาร แล้วจึงนำมาจัดแบ่งประเภทเอกสารในขั้นที่ 2 ด้วยตัวจัดแบ่งประเภท 3 วิธี คือ วิธีเบย์อย่างง่าย วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k=3 และวิธีแบบเซนทรอยด์ เพื่อหาตัวจัดแบ่งประเภทที่ดีที่สุดในขั้นที่ 2 จะสามารถจัดแบ่งประเภทเอกสารได้ผลดังตาราง 4.18 จากเอกสารข้อมูลยา 442 เอกสาร วิธีเบย์อย่างง่ายสามารถจัดแบ่งประเภทได้ถูก 392 เอกสาร จัดแบ่งประเภทผิด 50 เอกสาร คิดเป็นค่าความถูกต้อง 88.69% วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k=3 สามารถจัดแบ่งประเภทได้ถูก 408 เอกสาร จัดแบ่งประเภทผิด 34 เอกสาร คิดเป็นค่าความถูกต้อง 91.63% และวิธีแบบเซนทรอยด์สามารถจัดแบ่งประเภทได้ถูก 412 เอกสาร จัดแบ่งประเภทผิด 30 เอกสาร คิดเป็นค่าความถูกต้อง 93.21% จะเห็นได้ว่าในขั้นที่ 2 นี้ วิธีแบบเซนทรอยด์เป็นวิธีที่มีค่าความถูกต้องมากที่สุด

จากผลการประเมินการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาจำนวน 11 กลุ่ม ซึ่งจะแสดงในค่าความแม่นยำ ค่าการเรียกคืน และค่า F1 ดังตารางที่ 4.19 พบว่า ทั้ง 3 วิธีมีค่า F1 เฉลี่ยคือ 0.90 จึงสามารถจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาจำนวน 11 กลุ่มได้ โดยวิธีเบย์อย่างง่ายมีค่า F1 เฉลี่ย 0.86 วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k=3 มีค่า F1 เฉลี่ย 0.92 และวิธีแบบเซนทรอยด์มีค่า F1 เฉลี่ย 0.93 ดังนั้นจึงพบว่า วิธีแบบเซนทรอยด์ เป็นตัวจัดแบ่งประเภทที่ดีที่สุดในการจัดแบ่งประเภทเอกสารแบบ 2 ขั้นตอน ในขั้นที่ 2 การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาจำนวน 11 กลุ่ม

ตารางที่ 4.18 แสดงผลของขั้นที่ 2 การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาจำนวน 11 กลุ่ม

การประเมิน	**วิธีเบย์อย่างง่าย**	**วิธีเพื่อนบ้านใกล้ที่สุด (k=3)**	**วิธีแบบเซนทรอยด์**
จำนวนเอกสารทั้งหมดที่สามารถจัดแบ่งประเภทถูก	392	408	412
จำนวนเอกสารทั้งหมดที่สามารถจัดแบ่งประเภทผิด	50	34	30
ค่าความถูกต้อง	88.69%	91.63%	93.21%

ตารางที่ 4.19 แสดงการเปรียบเทียบค่า F1 ของผลขั้นที่ 2 การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาจำนวน 11 กลุ่มด้วยวิธีเบย์อย่างง่าย วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k = 3 และวิธีแบบเซนทรอยด์

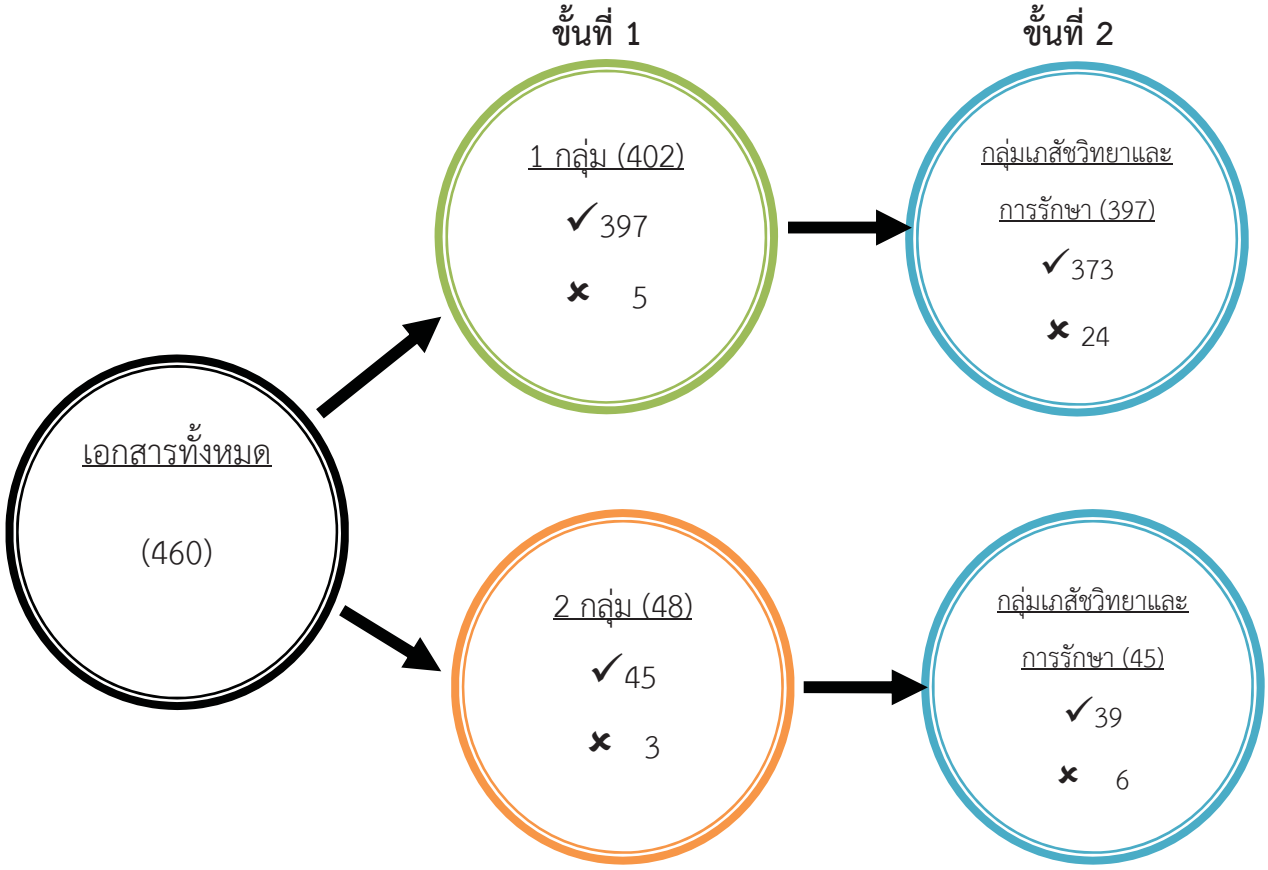
กลุ่มเภสัชวิทยาและการรักษา	**F1**			
	วิธีการจัดแบ่งประเภทเอกสาร			**เฉลี่ย**
	วิธีเบย์อย่างง่าย	**วิธีเพื่อนบ้านใกล้ที่สุด (k=3)**	**วิธีแบบเซนทรอยด์**	
Anti-infective agents	0.96	0.97	0.99	0.97
Antineoplastic drugs	0.92	0.92	0.92	0.92
Autonomics drugs	0.85	0.91	0.86	0.87
Blood formulation, coagulation, and thrombosis agents	0.92	1.00	1.00	0.97
Cardiovascular	0.96	0.98	0.98	0.97
CNS	0.85	0.85	0.89	0.86
EENT preparations	0.68	0.93	0.92	0.84
Gastrointestinal drugs	0.87	0.85	0.92	0.88
Hormones and synthetic substitutes	0.89	0.93	0.92	0.91
Skin and mucous membrane	0.83	0.93	0.93	0.90
Vitamins	0.83	0.86	0.90	0.86
เฉลี่ย	0.86	0.92	0.93	0.90

จากการจัดแบ่งประเภทเอกสารแบบ 2 ขั้นตอนของเอกสารข้อมูลยาชนิด 1 กลุ่มและชนิด 2 กลุ่ม พบว่า การจัดแบ่งประเภทเอกสารนี้สามารถทำได้ดี โดยในขั้นที่ 1 การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามจำนวนกลุ่ม [1, 2] มีตัวจัดแบ่งประเภทที่ดีที่สุดคือ วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k=3 และในขั้นที่ 2 การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาจำนวน 11 กลุ่มมีตัวจัดแบ่งประเภทที่ดีที่สุดคือ วิธีแบบเซนทรอยด์

จากการจัดแบ่งประเภทเอกสารแบบ 2 ขั้นตอนที่ได้ศึกษาไปในเอกสารข้อมูลยาชนิด 1 กลุ่มและชนิด 2 กลุ่ม สามารถนำมาเขียนเป็นตารางและภาพแสดงการสรุปผลการจัดแบ่งประเภทเอกสารแบบ 2 ขั้นตอนในเอกสารข้อมูลยาชนิด 1 กลุ่มและชนิด 2 กลุ่มได้ดังตารางที่ 4.20 และรูปที่ 4.1 โดยในขั้นที่ 1 การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามจำนวนกลุ่ม [1, 2] จะใช้วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k=3 เป็นตัวจัดแบ่งประเภท ซึ่งเป็นตัวจัดแบ่งประเภทที่ดีที่สุดในขั้นที่ 1 และขั้นที่ 2 การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาจำนวน 11 กลุ่ม จะใช้วิธีวิธีแบบเซนทรอยด์ซึ่งเป็นตัวจัดแบ่งประเภทที่ดีที่สุดในขั้นที่ 2 จากตารางที่ 4.20 และรูปที่ 4.1 พบว่าจากเอกสารทั้งหมดจำนวน 450 เอกสาร เมื่อนำมาจัดแบ่งประเภทในขั้นที่ 1 สามารถจัดแบ่งประเภทเอกสารข้อมูลยาชนิด 1 กลุ่มจำนวน 402 เอกสารได้ถูกจำนวน 397 เอกสาร คิดเป็นค่าความถูกต้อง 98.76% และเอกสารข้อมูลยาชนิด 2 กลุ่มจำนวน 48 เอกสารได้ถูกจำนวน 45 เอกสารคิดเป็นค่าความถูกต้อง 93.75% หลังจากนั้นจึงนำเอกสารข้อมูลยาที่จัดแบ่งประเภทได้ถูกต้องจากขั้นที่ 1 มาจัดแบ่งประเภทในขั้นที่ 2 คือการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาจำนวน 11 กลุ่ม ซึ่งจากเอกสารข้อมูลยาชนิด 1 กลุ่มที่จัดแบ่งประเภทในขั้นที่ 1 ได้ถูกต้องจำนวน 397 เอกสาร จะนำมาจัดแบ่งประเภทเอกสารข้อมูลยาในขั้นที่ 2 โดยจัดแบ่งประเภทถูกจำนวน 373 เอกสาร คิดเป็นค่าความถูกต้องของขั้นที่ 2 จำนวน 93.95% และเอกสารข้อมูลยาชนิด 2 กลุ่มที่จัดแบ่งประเภทได้ถูกต้องในขั้นที่ 1จำนวน 45 เอกสารนำมาจัดแบ่งประเภทในขั้นที่ 2 ซึ่งสามารถจัดแบ่งได้ถูกต้องจำนวน 39 เอกสาร คิดเป็นค่าความถูกต้องของขั้นที่ 2 จำนวน 86.67% ดังนั้นค่าความถูกต้องทั้งหมดของทั้ง 2 ขั้นตอนในเอกสารชนิด 1 กลุ่มเป็นจำนวน 92.79% และเอกสารชนิด 2 กลุ่มเป็นจำนวน 81.25% และจากเอกสารข้อมูลยาทั้งหมดจำนวน 450 เอกสาร เอกสารชนิด 1 กลุ่มจัดแบ่งประเภทได้ถูกต้องจำนวน 373 เอกสาร และเอกสารชนิด 2 กลุ่มจัดแบ่งประเภทได้ถูกต้องจำนวน 39 เอกสาร ดังนั้นจัดแบ่งประเภทได้ถูกต้องทั้งหมดจำนวน 412 เอกสาร คิดเป็นค่าความถูกต้องของทั้ง 2 ขั้นตอนทุกกลุ่มจำนวน 91.56% ซึ่งการจัดแบ่งประเภทเอกสารแบบ 2 ขั้นตอนนี้สามารถจัดแบ่งประเภทเอกสารชนิด 1 กลุ่มและ 2 กลุ่มได้เป็นอย่างดี

ตารางที่ 4.20 แสดงผลการจัดแบ่งประเภทเอกสารข้อมูลยาชนิดหลายกลุ่ม

การประเมิน	จำนวนกลุ่มเภสัชวิทยาและการรักษา	
	1 กลุ่ม	2 กลุ่ม
จำนวนเอกสารทั้งหมด	402	48
จำนวนเอกสารที่จัดแบ่งประเภทถูกในขั้นที่ 1 ด้วยวิธีเพื่อนบ้านใกล้ที่สุด	397	45
จำนวนเอกสารที่จัดแบ่งประเภทถูกในขั้นที่ 2 ด้วยวิธีแบบเซนทรอยด์	373	39
ค่าความถูกต้องของขั้นที่ 1	98.76%	93.75%
ค่าความถูกต้องของขั้นที่ 2	93.95%	86.67%
ค่าความถูกต้องของทั้ง 2 ขั้นตอน	92.79%	81.25%
ค่าความถูกต้องของทั้ง 2 ขั้นตอนทุกกลุ่ม	91.56%	



รูปที่ 4.1 แสดงภาพการสรุปผลการจัดแบ่งประเภทเอกสารข้อมูลยาชนิดหลายกลุ่มแบบ 2 ขั้นตอนใน เอกสารข้อมูลยาชนิด 1 กลุ่มและ 2 กลุ่ม

4.2 การจัดแบ่งประเภทเอกสารแบบ 2 ขั้นตอนในเอกสารข้อมูลยาชนิด 1 กลุ่ม, ชนิด 2 กลุ่ม และชนิด 3 กลุ่ม

ในวิธีการจัดแบ่งประเภทนี้ จะใช้เอกสารข้อมูลยาทั้งหมด 460 เอกสาร แบ่งเป็นข้อมูลยาชนิด 1 กลุ่ม จำนวน 402 เอกสาร เอกสารข้อมูลยาชนิด 2 กลุ่ม จำนวน 48 เอกสาร และเอกสารข้อมูลยาชนิด 3 กลุ่ม จำนวน 10 เอกสาร สามารถจัดแบ่งประเภทเอกสารแบบ 2 ขั้นตอนได้ ดังนี้

4.2.1. ขั้นที่ 1 การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามจำนวนกลุ่ม [1, 3]

จากเอกสารข้อมูลยาทั้งหมดจำนวน 460 เอกสาร ซึ่งมีเอกสารข้อมูลยาชนิด 1 กลุ่ม จำนวน 402 เอกสาร, เอกสารข้อมูลยาชนิด 2 กลุ่ม จำนวน 48 เอกสาร และเอกสารข้อมูลยาชนิด 3 กลุ่ม จำนวน 10 เอกสาร เมื่อทำการทดลองโดยทั้ง 3 วิธีคือ วิธีเบย์อย่างง่าย วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k=3 และวิธีแบบเซนทรอยด์ แล้วจะได้ผลดังตารางที่ 4.21 พบว่าวิธีเบย์อย่างง่ายสามารถจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามจำนวนกลุ่มได้ถูกจำนวน 426 เอกสาร คิดเป็นค่าความถูกต้อง92.61% วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k=3 สามารถจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามจำนวนกลุ่มได้ถูกจำนวน 447 เอกสารคิดเป็นค่าความถูกต้อง 97.17% และวิธีแบบเซนทรอยด์สามารถจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติได้ถูกจำนวน 445 เอกสารคิดเป็นค่าความถูกต้อง 96.74% จะเห็นได้ว่าวิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k=3 จะสามารถจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามจำนวนกลุ่มได้ถูกต้องมากที่สุด ดังนั้นในขั้นที่ 1 นี้จึงจะใช้วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k=3 เป็นตัวจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามจำนวนกลุ่ม

ตารางที่ 4.21 ผลของขั้นที่ 1 การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามจำนวนกลุ่ม [1, 3]

การประเมิน	วิธีเบย์อย่างง่าย	วิธีเพื่อนบ้านใกล้ที่สุด (k=3)	วิธีแบบเซนทรอยด์
จำนวนเอกสารทั้งหมดที่สามารถจัดแบ่งประเภทถูก	426	447	445
จำนวนเอกสารทั้งหมดที่สามารถจัดแบ่งประเภทผิด	34	13	15
ค่าความถูกต้อง	92.61%	97.17%	96.74%

4.2.2. ขั้นที่ 2 การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาจำนวน 11 กลุ่ม

จากการทดลองในขั้นที่ 1 พบว่าวิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k=3 เป็นตัวจัดแบ่งประเภทที่ดีที่สุดในการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามจำนวนกลุ่ม ซึ่งสามารถจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามจำนวนกลุ่มได้ถูกจำนวน 447 เอกสาร ดังนั้นในขั้นที่ 2 การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติเป็น 11 กลุ่มตามกลุ่มเภสัชวิทยาและการรักษานี้ ซึ่งต้องใช้จำนวนเอกสารที่จัดแบ่งประเภทถูกจากขั้นที่ 1 มาจัดแบ่งประเภท จึงใช้เอกสารที่จัดแบ่งประเภทถูกจากขั้นที่ 1 เมื่อใช้วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k=3 จำนวน 447 เอกสาร ซึ่งผลของการจัดแบ่งประเภทขั้นที่ 2 การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติเป็น 11 กลุ่มตามกลุ่มเภสัชวิทยาและการรักษาแสดงในตารางที่ 4.22 เมื่อจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติเป็น 11 กลุ่มตามกลุ่มเภสัชวิทยาและการรักษาโดยทั้ง 3 วิธีคือ วิธีเบย์อย่างง่าย วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k=3 และวิธีแบบเซนทรอยด์ พบว่าวิธีเบย์อย่างง่ายจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติได้ถูกจำนวน 396 เอกสารคิดเป็นค่าความถูกต้อง 88.59% วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k=3 จัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติได้ถูกจำนวน 411 เอกสารคิดเป็นค่าความถูกต้อง 91.55% และวิธีแบบเซนทรอยด์จะจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติได้ถูกจำนวน 420 เอกสารคิดเป็นค่าความถูกต้อง 93.96% จะเห็นได้ว่าวิธีแบบเซนทรอยด์เป็นตัวจัดแบ่งประเภทที่ดีที่สุดจากทั้ง 3 วิธีในการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติเป็น 11 กลุ่มตามกลุ่มเภสัชวิทยาและการรักษา

ตารางที่ 4.22 ผลของขั้นที่ 2 การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติเป็น 11 กลุ่มตามกลุ่มเภสัชวิทยาและการรักษา

การประเมิน	**วิธีเบย์อย่างง่าย**	**วิธีเพื่อนบ้านใกล้ที่สุด (k=3)**	**วิธีแบบเซนทรอยด์**
จำนวนเอกสารทั้งหมดที่สามารถจัดแบ่งประเภทถูก	396	411	420
จำนวนเอกสารทั้งหมดที่สามารถจัดแบ่งประเภทผิด	51	36	27
ค่าความถูกต้อง	88.59%	91.95%	93.96%

เมื่อทดลองประเมินผลด้วยค่า F1 การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติเป็น 11 กลุ่มตามกลุ่มเภสัชวิทยาและการรักษาในการจัดแบ่งประเภทเอกสารข้อมูลยาวิธีแบบเซนทรอยด์ วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k = 3 และวิธีเบย์อย่างง่ายดังตารางที่ 4.12 พบว่าทั้ง 3 วิธีมีค่า F1 เฉลี่ย 0.91 เห็นได้ว่าทั้ง 3 วิธีสามารถใช้จัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติเป็น 11 กลุ่มตามกลุ่มเภสัชวิทยาและการรักษาได้ ซึ่งวิธีเบย์อย่างง่ายมีค่า F1 เฉลี่ย 0.87 วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k=3 มีค่า F1 เฉลี่ย 0.91 และวิธีแบบเซนทรอยด์มีค่า F1 เฉลี่ย 0.94 จะเห็นได้ว่าจากทั้ง 3 วิธี มีวิธีแบบเซนทรอยด์เป็นวิธีที่ดีที่สุดในการการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติชนิดหลายกลุ่มในขั้นที่ 2 ซึ่งจะจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติเป็น 11 กลุ่มตามกลุ่มเภสัชวิทยาและการรักษา

ตารางที่ 4.23 แสดงการเปรียบเทียบค่า F1 ของผลขั้นที่ 2 การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติเป็น 11 กลุ่มตามกลุ่มเภสัชวิทยาและการรักษาในการจัดแบ่งประเภทเอกสารข้อมูลยาวิธีเบย์อย่างง่าย วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k = 3 และวิธีแบบเซนทรอยด์

กลุ่มเภสัชวิทยาและการรักษา	F1			
	วิธีการจัดแบ่งประเภทเอกสาร			เฉลี่ย
	วิธีเบย์อย่างง่าย	วิธีเพื่อนบ้านใกล้ที่สุด (k=3)	วิธีแบบเซนทรอยด์	
Anti-infective agents	0.93	0.97	0.97	0.96
Antineoplastic drugs	0.90	0.92	0.94	0.92
Autonomics drugs	0.87	0.91	0.92	0.90
Blood formulation, coagulation, and thrombosis agents	0.92	1.00	0.99	0.97
Cardiovascular	0.96	0.98	0.98	0.97
CNS	0.88	0.84	0.93	0.88
EENT preparations	0.64	0.92	0.90	0.82
Gastrointestinal drugs	0.87	0.83	0.88	0.86
Hormones and synthetic substitutes	0.93	0.93	0.93	0.93
Skin and mucous membrane	0.82	0.92	0.95	0.90
Vitamins	0.83	0.83	0.92	0.86
เฉลี่ย	0.87	0.91	0.94	0.91

เมื่อพิจารณาในด้านแต่ละกลุ่มของกลุ่มเภสัชวิทยาและการรักษาจะพบว่าแต่ละกลุ่มในตารางที่ 4.23 จากกลุ่มเภสัชวิทยาและการรักษาจำนวน 11 กลุ่ม จะสามารถจัดแบ่งประเภทเอกสารได้ดี โดยมีกลุ่มที่จัดแบ่งประเภทได้ดีที่สุดซึ่งมีค่า F1 เฉลี่ย 0.97 จำนวน 2 กลุ่มคือ กลุ่ม Blood formulation, coagulation, and thrombosis agents, และกลุ่ม Cardiovascular

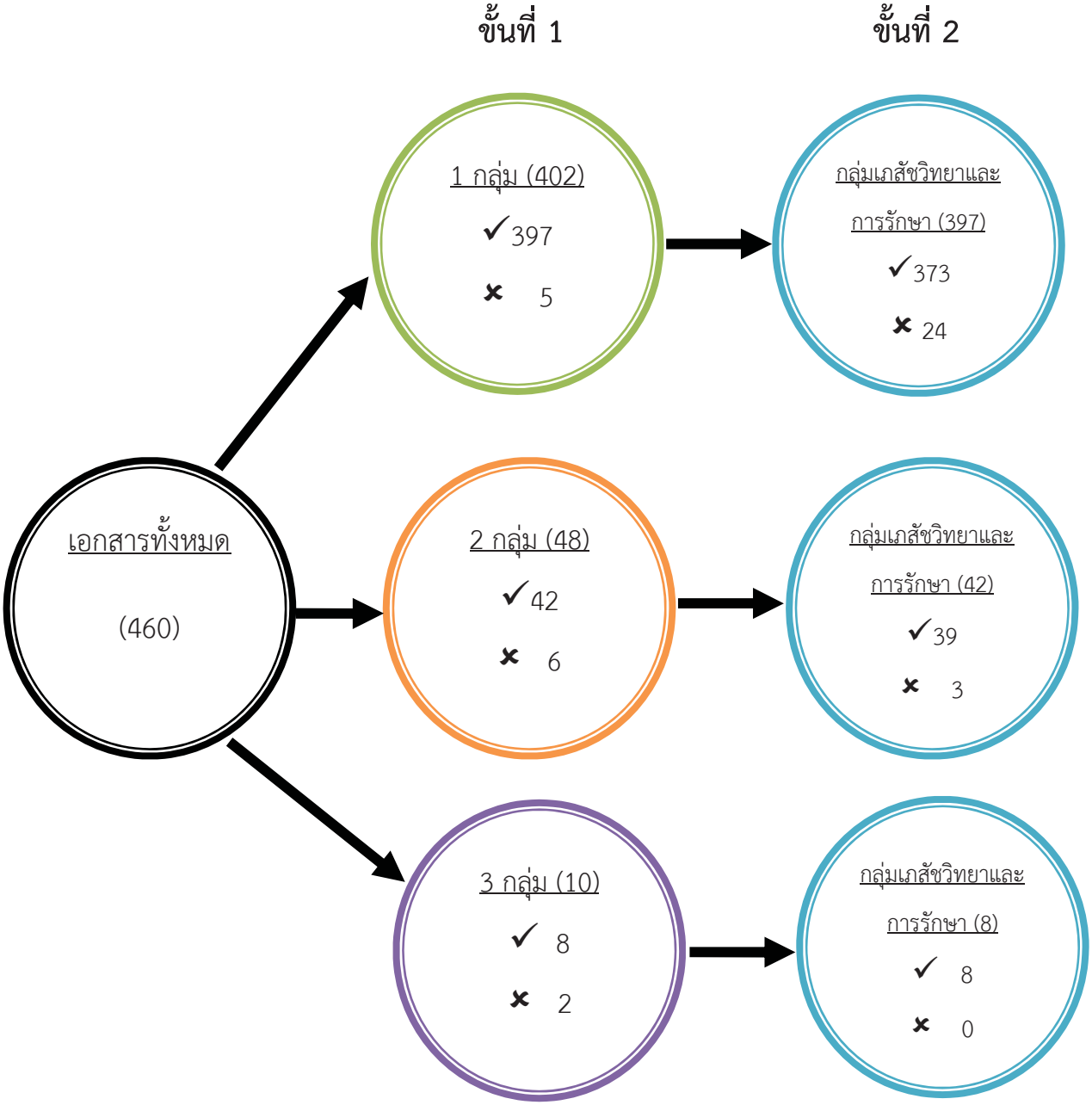
และมีกลุ่มที่จัดแบ่งประเภทได้ดีน้อยที่สุดซึ่งมีค่า F1 เฉลี่ย 0.82 จำนวน 1 กลุ่มคือ กลุ่ม EENT preparations

จากการจัดแบ่งประเภทเอกสารแบบ 2 ขั้นตอนในเอกสารข้อมูลยาชนิด 1 กลุ่ม, 2 กลุ่มและ 3 กลุ่มที่ได้ศึกษาไปแล้ว สามารถนำมาเขียนเป็นตารางและภาพแสดงการสรุปผลได้ดังตารางที่ 4.24 และรูปที่ 4.2 ซึ่งแสดงผลการจัดแบ่งประเภทเอกสารแบบ 2 ขั้นตอนในเอกสารข้อมูลยาชนิด 1 กลุ่ม, 2 กลุ่มและ 3 กลุ่มสามารถทำได้โดยในขั้นที่ 1 การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามจำนวนกลุ่ม [1, 3] จะใช้วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k=3 เป็นตัวจัดแบ่งประเภท ซึ่งวิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k=3 นี้เป็นตัวจัดแบ่งประเภทในขั้นที่ 1 ที่ดีที่สุดใน 3 วิธีจากที่ได้ศึกษามาแล้ว และขั้นที่ 2 การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาจำนวน 11 กลุ่ม จะใช้วิธีวิธีแบบเซนทรอยด์เป็นตัวจัดแบ่งประเภทซึ่งวิธีแบบเซนทรอยด์นี้เป็นตัวจัดแบ่งประเภทในขั้นที่ 2 ที่ดีที่สุดใน 3 วิธี จากตารางที่ 4.24 และรูปที่ 4.2 พบว่าจากเอกสารทั้งหมดจำนวน 460 เอกสาร เมื่อนำมาจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามจำนวนกลุ่มเป็น 3 ประเภท คือ 1 กลุ่ม, 2 กลุ่ม และ 3 กลุ่ม จะเห็นได้ว่าสามารถจัดแบ่งประเภทเอกสารข้อมูลยาชนิด 1 กลุ่มจำนวน 402 เอกสารได้ถูกจำนวน 397 เอกสารคิดเป็นค่าความถูกต้อง 98.76% เอกสารข้อมูลยาชนิด 2 กลุ่มจำนวน 48 เอกสารได้ถูกจำนวน 42 เอกสารคิดเป็นค่าความถูกต้อง 87.50% และเอกสารข้อมูลยาชนิด 3 กลุ่มจำนวน 10 เอกสารได้ถูกจำนวน 8 เอกสารคิดเป็นค่าความถูกต้อง 80.00% และนำเอกสารข้อมูลยาที่จัดแบ่งประเภทตามจำนวนกลุ่มที่สามารถจัดแบ่งประเภทได้ถูกต้องนี้มาจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษา โดยเอกสารข้อมูลยาชนิด 1 กลุ่มจะสามารถจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาได้ถูกจำนวน 373 เอกสารคิดเป็นค่าความถูกต้อง 93.95% และเอกสารข้อมูลยาชนิด 2 กลุ่มจะสามารถจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาได้ถูกจำนวน 39 เอกสารคิดเป็นค่าความถูกต้อง 92.86% และเอกสารข้อมูลยาชนิด 3 กลุ่มจะสามารถจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาได้ถูกจำนวน 8 เอกสารคิดเป็นค่าความถูกต้อง 100.00% ดังนั้นเอกสารข้อมูลยาชนิด 1 กลุ่มจะสามารถจัดแบ่งประเภทเอกสารโดยอัตโนมัติได้อย่างถูกต้องทั้ง 2 ขั้นตอนคิดเป็น 92.79% เอกสารข้อมูลยาชนิด 2 กลุ่มจะสามารถจัดแบ่งประเภทเอกสารโดยอัตโนมัติได้อย่างถูกต้องทั้ง 2 ขั้นตอนคิดเป็น

81.25% เอกสารข้อมูลยาชนิด 3 กลุ่มจะสามารถจัดแบ่งประเภทเอกสารโดยอัตโนมัติได้อย่างถูกต้องทั้ง 2 ขั้นตอนคิดเป็น 80.00% และจากเอกสารข้อมูลยาทั้งหมด 460 เอกสาร เอกสารข้อมูลยาชนิด 1 กลุ่มสามารถจัดแบ่งประเภทได้ถูกจำนวน 373 เอกสาร เอกสารข้อมูลยาชนิด 2 กลุ่มสามารถจัดแบ่งประเภทได้ถูกจำนวน 39 เอกสาร และเอกสารชนิด 3 กลุ่มสามารถจัดแบ่งประเภทได้ถูกจำนวน 8 เอกสาร จึงมีค่าความถูกต้องของทั้ง 2 ขั้นตอนในทุกกลุ่มคิดเป็น 91.30% ซึ่งจะเห็นได้ว่าการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาชนิดหลายกลุ่มแบบ 2 ขั้นตอนสามารถจัดแบ่งประเภทเอกสารข้อมูลยาชนิดหลายกลุ่มได้เป็นอย่างดี

ตารางที่ 4.24 แสดงผลการจัดแบ่งประเภทเอกสารข้อมูลยาชนิดหลายกลุ่ม

การประเมิน	**จำนวนกลุ่มเภสัชวิทยาและการรักษา**		
	1 กลุ่ม	**2 กลุ่ม**	**3 กลุ่ม**
จำนวนเอกสารทั้งหมด	402	48	10
จำนวนเอกสารที่จัดแบ่งประเภทถูกในขั้นที่ 1 ด้วยวิธีเพื่อนบ้านใกล้ที่สุด	397	42	8
จำนวนเอกสารที่จัดแบ่งประเภทถูกในขั้นที่ 2 ด้วยวิธีแบบเซนทรอยด์	373	39	8
ค่าความถูกต้องของขั้นที่ 1	98.76%	87.50%	80.00%
ค่าความถูกต้องของขั้นที่ 2	93.95%	92.86%	100.00%
ค่าความถูกต้องของทั้ง 2 ขั้นตอน	92.79%	81.25%	80.00%
ค่าความถูกต้องของทั้ง 2 ขั้นตอนทุกกลุ่ม	91.30%		



รูปที่ 4.2 แสดงภาพการสรุปผลการจัดแบ่งประเภทเอกสารข้อมูลยาชนิดหลายกลุ่มแบบ 2 ขั้นตอน

เนื่องจากเอกสารข้อมูลยาสามารถจัดแบ่งประเภทตามกลุ่มเภสัชวิทยาและการรักษาให้อยู่ได้ในหลายกลุ่มตามหนังสือ AHFS Drug Information 2012 ซึ่งเอกสารข้อมูลยาจะต้องจัดอยู่ในจำนวนกลุ่มที่น้อยที่สุดคือ 1 กลุ่ม และจากเอกสารข้อมูลยาที่นำมาศึกษาจะจัดให้อยู่ในจำนวนกลุ่มที่มากที่สุดได้จำนวน 3 กลุ่ม ดังนั้นในการจัดแบ่งประเภทเอกสารข้อมูลยาชนิดหลายกลุ่มจึงต้องรู้จำนวนกลุ่มที่เอกสารข้อมูลยาสามารถอยู่ได้ก่อน แล้วจึงจัดแบ่งประเภทตามเภสัชวิทยาและการรักษาได้ ดังนั้นในการศึกษาครั้งนี้จึงได้พัฒนาการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาชนิดหลายกลุ่มแบบ 2 ขั้นตอนขึ้น คือในขั้นที่ 1 เป็นการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามจำนวนกลุ่ม [a, b] และขั้นที่ 2 การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาจำนวน 11 กลุ่ม ซึ่งในตัวจัดแบ่งประเภทที่นำมาศึกษาทั้ง 3 วิธีคือ วิธีเบย์อย่างง่าย วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k=3 และวิธีแบบเซนทรอยด์ พบว่าตัวจัดแบ่งประเภทที่ดีที่สุดในขั้นที่ 1 คือ วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k=3 และตัวจัดแบ่งประเภทที่ดีที่สุดในขั้นที่ 2 คือวิธีแบบเซนทรอยด์

บทที่ 5
สรุป อภิปรายผล และข้อเสนอแนะ

ในงานวิทยานิพนธ์ชิ้นนี้ได้ศึกษาการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติทั้ง 2 ชนิด คือ การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติชนิดหนึ่งกลุ่มและการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติชนิดหลายกลุ่ม ซึ่งศึกษาด้วยวิธีการจัดแบ่งประเภททั้ง 3 วิธี คือ วิธีเบย์อย่างง่าย วิธีเพื่อนบ้านใกล้ที่สุด และวิธีแบบเซนทรอยด์ จากผลการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติจะสามารถสรุป อภิปรายผล และมีข้อเสนอแนะ ดังต่อไปนี้

1. สรุปผลการศึกษา

จากผลการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษา จะสามารถสรุปผลได้ดังนี้

1.1. การจัดแบ่งประเภทเอกสารโดยอัตโนมัติสามารถใช้จัดแบ่งประเภทเอกสารที่เป็นเอกสารข้อมูลยาได้ และสามารถจัดแบ่งประเภทได้ดี วิธีการที่ใช้จัดแบ่งประเภทเอกสารที่ใช้ทั้ง 3 วิธี สามารถจัดแบ่งประเภทได้อย่างมีประสิทธิภาพ

1.2. การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติชนิดหนึ่งกลุ่มนั้น สามารถจัดแบ่งประเภทเอกสารข้อมูลยาตามกลุ่มเภสัชวิทยาและการรักษาได้เป็นอย่างดี โดยจากกลุ่มเภสัชวิทยาและการรักษาทั้ง 11 กลุ่ม จะมีกลุ่มเภสัชวิทยาและการรักษา 3 กลุ่มที่สามารถจัดแบ่งประเภทได้ดีที่สุด คือ Anti-infective agents, Blood formulation, coagulation, and thrombosis agents และ Cardiovascular ในส่วนของวิธีการจัดแบ่งประเภทมีวิธีที่ดีที่สุด คือ วิธีแบบเซนทรอยด์

1.3. การจัดแบ่งประเภทเอกสารแบบไบนารี่สามารถใช้จัดแบ่งประเภทเอกสารชนิดหนึ่งกลุ่มและชนิดหลายกลุ่มได้ แต่การจัดแบ่งประเภทในเอกสารชนิดหลายกลุ่มจะจัดแบ่งประเภทได้ไม่ดีนักและเมื่อเปรียบเทียบกับการจัดแบ่งประเภทเอกสารแบบ 2 ขั้นตอนจะเห็นได้ว่า การจัดแบ่งประเภทเอกสารแบบ 2 ขั้นตอนจะมีการจัดแบ่งประเภทเอกสารที่มีประสิทธิภาพดีกว่า มีความถูกต้องมากกว่าการจัดแบ่งประเภทเอกสารแบบไบนารี่ทั้งชนิดหนึ่งกลุ่มและชนิดหลายกลุ่ม

1.4. การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติชนิดหลายกลุ่ม เมื่อนำมาจัดประเภทเอกสารข้อมูลยาชนิดที่อยู่ได้หลายกลุ่มเภสัชวิทยาและการรักษา จะพบปัญหาว่าไม่สามารถระบุจำนวนกลุ่มทางเภสัชวิทยาและการรักษาที่เหมาะสมกับเอกสารข้อมูลยาได้ ดังนั้นในงานวิทยานิพนธ์นี้จึงได้พัฒนาวิธีการจัดแบ่งประเภทเอกสารแบบ 2 ขั้นตอนขึ้น โดยสามารถจัดแบ่งประเภทเอกสารข้อมูลยาชนิดหลายกลุ่มได้ ซึ่งจากการศึกษาใช้เอกสารข้อมูลยาชนิด 1 กลุ่ม, ชนิด 2 กลุ่ม และชนิด 3 กลุ่ม โดยขั้นที่ 1 การจัดแบ่งประเภทเอกสารตามจำนวนกลุ่ม [a, b] ในการทดลองจะเป็นการค้นหาว่าเอกสารข้อมูลยานั้นๆ สามารถจัดแบ่งเอกสารให้อยู่ในกลุ่มเภสัชวิทยาและการรักษาได้กี่กลุ่ม ซึ่งจากผลการศึกษาพบว่าขั้นตอนนี้สามารถจัดแบ่งประเภทเอกสารได้ดีและมีวิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k=3 เป็นตัวจัดแบ่งประเภทที่ดีที่สุดในขั้นนี้ หลังจากนั้นจึงจัดแบ่งประเภทในขั้นที่ 2 ซึ่งเป็นการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาจำนวน 11 กลุ่ม เป็นการค้นหาว่าจะสามารถจัดแบ่งประเภทเอกสารข้อมูลยานั้นให้อยู่ในกลุ่มเภสัชวิทยาและการรักษาใดได้บ้าง จากผลการศึกษาพบว่าขั้นตอนนี้สามารถจัดแบ่งประเภทเอกสารข้อมูลยาตามกลุ่มเภสัชวิทยาและการรักษาได้ดีและมีวิธีแบบเซนทรอยด์เป็นตัวจัดแบ่งประเภทที่ดีที่สุดในขั้นที่ 2 นี้

2. อภิปรายผลการศึกษา

จากการใช้เอกสารข้อมูลยาในการทดลองพบว่า เอกสารข้อมูลยา 1 เอกสารจะสามารถจัดให้อยู่ในกลุ่มเภสัชวิทยาและการรักษาได้หลายกลุ่ม โดยเอกสารที่คัดเลือกมาจะสามารถอยู่ในกลุ่มเภสัชวิทยาและการรักษาได้ตั้งแต่ 1 กลุ่มถึง 3 กลุ่ม ตามหนังสือ AHFS Drug Information 2012

ในการศึกษาครั้งนี้ได้ทดลองจัดแบ่งประเภทเอกสารโดยอัตโนมัติในเอกสารข้อมูลยา โดยใช้วิธีการจัดแบ่งประเภทเอกสารแบบไบนารี่และการจัดแบ่งประเภทเอกสารแบบครั้งละหลายกลุ่ม จากการศึกษาพบว่า วิธีการจัดแบ่งประเภทเอกสารแบบไบนารี่สามารถจัดแบ่งประเภทเอกสารข้อมูลยาชนิดหลายกลุ่มได้ แต่การจัดแบ่งประเภทแบบไบนารี่จะมีวิธีการจัดแบ่งประเภทที่ยุ่งยากและใช้เวลานานในการจัดแบ่งประเภท นอกจากนี้จำนวนกลุ่มที่ได้จากการจัดแบ่งประเภทจะสามารถจัดแบ่งประเภทได้น้อยกว่าหรือมากกว่าจำนวนกลุ่มจริง ซึ่งในบางครั้งอาจไม่สามารถจัดให้อยู่ในกลุ่มใดได้เลย ส่วนวิธีการจัดแบ่งประเภทเอกสารแบบครั้งละหลายกลุ่มจะมีวิธีการในการจัดแบ่งประเภทง่ายกว่าและใช้เวลาน้อยกว่า แต่ในการจัดแบ่งประเภทเอกสารแบบครั้งละหลายกลุ่มในเอกสารข้อมูลยาชนิดหลายกลุ่มจะทำได้ยาก เนื่องจากจะต้องใช้วิธีอื่นมาช่วยในการจัดแบ่งประเภท เช่น การทำ

Threshold และการทำสัดส่วนคะแนน ดังนั้นหากทราบจำนวนกลุ่มของเอกสารที่จะนำมาจัดแบ่งประเภทตามเอกสารข้อมูลยาและการรักษาจึงจะสามารถใช้วิธีแบบครั้งละหลายกลุ่มในการจัดแบ่งประเภทเอกสารข้อมูลยาตามกลุ่มเภสัชวิทยาและการรักษาได้ แต่จากการศึกษาเบื้องต้น ด้วยการทำ Threshold และการทำสัดส่วนคะแนน ยังไม่สามารถหาวิธีการที่เหมาะสมในการตัดสินจำนวนกลุ่มทางเภสัชวิทยาและการรักษาที่ถูกต้องได้อย่างมีประสิทธิภาพ หลังจากนั้นจึงพัฒนาการจัดแบ่งประเภทเอกสารแบบ 2 ขั้นตอนขึ้น โดยขั้นที่ 1 จะเป็นการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามจำนวนกลุ่ม [a, b]

เมื่อกำหนดให้ [a, b] คือช่วงของจำนวนกลุ่มที่จัดแบ่งประเภทได้

a คือ จำนวนกลุ่มที่น้อยที่สุดที่เอกสารสามารถจัดแบ่งประเภทอยู่ได้

b คือ จำนวนกลุ่มที่มากที่สุดที่เอกสารสามารถจัดแบ่งประเภทอยู่ได้

เมื่อทราบจำนวนกลุ่มของเอกสารข้อมูลยาแล้ว จึงสามารถนำเอกสารข้อมูลยาเหล่านั้นมาจัดแบ่งประเภทในขั้นที่ 2 ได้ซึ่งเป็นการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาจำนวน 11 กลุ่ม

วิธีการจัดแบ่งประเภทเอกสารข้อมูลยาที่นำมาศึกษามี 3 วิธีคือ วิธีเบย์อย่างง่าย วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k=3 และวิธีแบบเซนทรอยด์เมื่อนำวิธีการจัดแบ่งประเภททั้ง 3 วิธีมาทดสอบจัดแบ่งประเภทเอกสารข้อมูลยาพบว่า ทั้ง 3 วิธีสามารถจัดแบ่งประเภทเอกสารได้ดี ซึ่งในการจัดแบ่งประเภทเอกสารแบบ 2 ขั้นตอนนั้น ในขั้นที่ 1 การจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามจำนวนกลุ่ม [a, b] มีตัวจัดแบ่งประเภทที่จัดแบ่งประเภทได้ดีที่สุดในขั้นนี้คือ วิธีเพื่อนบ้านใกล้ที่สุดเมื่อ k=3 และในขั้นที่ 2 เป็นการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษาจำนวน 11 กลุ่ม มีตัวจัดแบ่งประเภทที่ดีที่สุดคือ วิธีแบบเซนทรอยด์

ในการจัดแบ่งประเภทเอกสารข้อมูลยาชนิดหลายกลุ่มแบบ 2 ขั้นตอนที่ได้ศึกษานี้ หากต้องการให้มีประสิทธิภาพดี จะต้องมีการจัดแบ่งประเภทเอกสารที่ดีในทั้ง 2 ขั้นตอน ซึ่งในการศึกษานี้มีข้อจำกัดคือ จำนวนเอกสารข้อมูลยาไม่เท่ากันในแต่ละชนิด โดยจำนวนเอกสารข้อมูลยาชนิด 2 กลุ่มและ 3 กลุ่มมีจำนวนน้อยกว่าเอกสารข้อมูลยาชนิด 1 กลุ่ม ซึ่งจะทำให้การจัดแบ่งประเภทเอกสารข้อมูลยาชนิดหลายกลุ่มแบบ 2 ขั้นตอนในขั้นที่ 1 ซึ่งเป็นการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามจำนวนกลุ่มเกิดความคลาดเคลื่อนขึ้นได้

ในส่วนของการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติตามกลุ่มเภสัชวิทยาและการรักษานี้ สามารถจัดแบ่งประเภทเอกสารข้อมูลยาได้ แต่ยังมีเอกสารข้อมูลยาที่ถูกจัดให้อยู่ผิดกลุ่ม ซึ่งอาจเนื่องมาจากเอกสารข้อมูลยาเหล่านี้มีคำที่สามารถบ่งชี้ว่าอยู่ในกลุ่มอื่นได้ เช่น เอกสารข้อมูลยาบางชนิดที่อยู่ในกลุ่ม EENT preparations และ Anti-infective agents บางเอกสารอาจถูกจัดอยู่ในกลุ่ม Anti-infective agents และ Antineoplastic drugs เนื่องจากยาบางชนิดที่เป็นยาเตรียมเพื่อใช้ในส่วนของตาหูคอจมูกจะมียาชนิดที่เป็นส่วนออกฤทธิ์ในการฆ่าเชื้อ ดังนั้นภายในเอกสารข้อมูลยาจึงได้มีการกล่าวถึงยาที่เป็นส่วนออกฤทธิ์ในการฆ่าเชื้อ และอาจกล่าวในด้านฤทธิ์ทางเภสัชวิทยา โดยในบางเอกสารอาจมีการกล่าวถึงการทดสอบการฆ่าเชื้อของยาตัวนั้น ซึ่งจะมีคำที่เกี่ยวกับการยับยั้งการเจริญเติบโต โดยคำเหล่านี้จะมีความสำคัญในกลุ่ม Antineoplastic drugs ซึ่งเป็นกลุ่มของยาที่ยับยั้งการเจิญเติบโตของเซลล์ เช่นยารักษามะเร็ง ในขณะที่คำที่เกี่ยวกับกลุ่ม EENT preparations นั้นจะปรากฏอยู่น้อย เช่นในส่วนของข้อบ่งใช้และวิธีใช้ ดังนั้นในเอกสารดังกล่าวจึงอาจถูกจัดแบ่งประเภทผิดเป็นกลุ่ม Antineoplastic drugs ได้

เอกสารข้อมูลยาบางชนิดจัดอยู่ในกลุ่ม CNS แต่ถูกจัดประเภทผิดกลุ่ม โดยจัดให้อยู่ในกลุ่ม Cardiovascular เนื่องจากเอกสารข้อมูลยาชนิดนั้น เป็นยาที่มีฤทธิ์ต่อระบบประสาทส่วนกลางแต่มีผลต่อระบบหัวใจและหลอดเลือดด้วยซึ่งจะเกี่ยวกับกลุ่ม Cardiovascular เนื้อหาภายในเอกสารจะมีคำที่มีความสำคัญกับกลุ่ม Cardiovascular ด้วย เช่น ยาที่มีผลทำให้เกิด hypotension หรือภาวะความดันโลหิตต่ำ หรือยาที่อาจทำให้เกิด bradycardia ซึ่งเป็นภาวะหัวใจเต้นช้าได้ ยาบางชนิดที่อยู่ในกลุ่ม CNS จะมีการกล่าวถึงขนาดของยาซึ่งจะมีผลต่อระบบหัวใจและหลอดเลือดได้ ดังนั้นอาจทำให้เกิดการจัดแบ่งประเภทผิดกลุ่มได้

3. ข้อเสนอแนะ

ในการศึกษาจัดแบ่งประเภทเอกสารข้อมูลยาในครั้งต่อไป ควรจะต้องมีการศึกษาในเอกสารข้อมูลยาที่มีจำนวนมากขึ้น ซึ่งอาจจะมาจากหลายแหล่งข้อมูลมากขึ้น เพื่อเพิ่มจำนวนข้อมูลสอนและข้อมูลทดสอบ เพื่อให้สามารถจัดแบ่งประเภทเอกสารข้อมูลยาได้แม่นยำและถูกต้องมากขึ้น โดยเฉพาะเอกสารข้อมูลยาชนิดหลายกลุ่ม

นอกจากนี้ในการจัดแบ่งประเภทเอกสารข้อมูลยาโดยอัตโนมัติชนิดหลายกลุ่มของการศึกษาครั้งนี้ ได้มีการศึกษาเฉพาะเอกสารข้อมูลยาที่สามารถจัดแบ่งประเภทได้ถูกต้องทั้งหมด ทั้ง

จำนวนกลุ่มและกลุ่มเภสัชวิทยาและการรักษาที่สามารถจัดแบ่งได้ ซึ่งยังคงมีเอกสารข้อมูลยาบางเอกสารที่สามารถจัดแบ่งประเภทจำนวนกลุ่มได้ถูกต้องแต่กลุ่มเภสัชวิทยาและการรักษานั้นจัดแบ่งได้ถูกเพียงบางส่วนเท่านั้น ซึ่งสิ่งนี้จึงมีควรความจำเป็นในการศึกษาเพิ่มเติม

รายการอ้างอิง

[1] Fabrizio, Sebastiani. (2002). “Machine learning in automated text categorization.” **ACM Computing Serveys** 34, 1 (March): 1-47.

[2] Witten, Ian H. and Eibe Frank. (2005). **Data mining practical machine learning tools and techniques** 2nd ed. United States of America: Morgan Kaufmann.

[3] Roiger, Richard J., and Michael W. Geatz. (2003). **Data mining a tutorial-based primer** Boston: Addison Wesley.

[4] Jiawei Han, and Micheline Kamber. (2006). **Data mining concepts and techniques** 2nd ed. United States of America: Morgan Kaufmann.

[5] Pang-Ning Tan, Michael Steinbach, and Vipin Kumar. (2006). **Introduction to data mining**. United States of America: Addison-Wesley.

[6] Salton, Gerard, and Christopher Buckley. (1988). “Term-weighting approaches in automatic text retrieval.” **Information Procession & Management** 24, 5 (January): 513-523.

[7] Hu Guan, Jingyu Zhou, and Minyi Guo. (2009). “A class-feature-centroid classifier for text categorization.” **Proceedings of the 18th international conference on World Wide Web** (April): 201-210.

[8] Dumais, Susan, John Platt, and Davis Heckerman. (1998). “Inductive learning algorithms and representations for text categorization.” **Proceedings of the seventh international conference on Information and knowledge management**: 148-155.

[9] Son Doan et al. (2009). “Towards role-based filtering of disease outbreak reports.” **Journal of Biomedical Informatics** 42, 5 (October): 773-780.

[10] Anastasia Sotelsek-margalef , Julio Villena-román , and Iii Madrid. (2008). “MIDAS: An Information-Extraction Approach to Medical Text Classification.” **Procesamiento del lenguaje Natural** 41: 97-104.

[11] Verayuth Lertnattee, and Thanaruk Theeramunkong. (2004). “Effect of term distributions on centroid-based text categorization.” **Information Sciences** 158, 1: 89-115.

[12] Ana Cardoso-Cachopo, and Arlindo L. Oliveira. (2007). “Semi-supervised single-label text categorization using centroid-based classifiers.” **Proceedings of the 2007 ACM symposium on applied computing** (March): 844-851.

[13] McCallum, Andrew K. (1996). **Bow: A toolkit for statistical language modeling, text retrieval, classification and clustering**. Accessed April 19. Available from http://www.cs.cmu.edu/~mccallum/bow

[14] American Society of Health-System Pharmacists. (2012). **AHFS Drug Information 2012**. Maryland: The American Society of Health-System Pharmacists, Inc.

[15] M. Ikonomakis, Sotiris Kotsiantis, and Vasilis Tampakas. (2005). “Text classification using machine learning techniques.” **WSEA Transactions on Computers** 4, 8 (August): 966-974.

[16] Vandana Korde and C Namrata Mahender. (2012). “Text classification and classifiers: A survey.” **International Journal of Artificial Intelligence & Applications (IJAIA)** 3, 2 (March): 85-99.

[17] Holts, Alberto, Claidio Riquelme, and Rodrigo Alfaro. (2010). “Automated text binary classification using machine learning approach.” **XXIX International Conference of the Chilean Computer Science Society**: 212-217.

[18] Grigorios Tsoumakas, and Ioannis Katakis. (2007). “Multi-label classification: An overview.” **International Journal of Data Warehousing and mining**: 1-13.

[19] Araken M Santos, Anne M P Canuto, and Antonino Feitosa Neto. (2011). “A comparative analysis of classification methods to multi-label tasks in different application domains.” **International Journal of Computer Information Systems and Industrial Management Applications** 3: 218-227.

[20] Thanaruk Theeramunkong, and Verayuth Lertnattee, (2002). “Multi-dimensional text classification.” **Proceedings of the 19th international conference on Computational linguistics** 1: 1-7.

[21] Verayuth Lertnattee, and Thanaruk Theeramunkong. (2004) “Multidimensional Text Classification for Drug Information.” **IEEE Transactions on Information Technology in Biomedicine** 8, 3 (September): 306 – 312.

ภาคผนวก ก

First Tier Classifications

The following table lists the first tier in the Classification. All marketed drug products fall under one or more of these classifications:

AHFS Class Number	AHFS Class Description
4:00	Antihistamine Drugs
8:00	Anti-infective Agents
10:00	Antineoplastic Agents
12:00	Autonomic Drugs
16:00	Blood Derivatives
20:00	Blood Formation, Coagulation, and Thrombosis Agents
24:00	Cardiovascular Drugs
26:00	Cellular Therapy
28:00	Central Nervous System Agents
32:00	Contraceptives (foams, devices)
34:00	Dental Agents
36:00	Diagnostic Agents
38:00	Disinfectants (for agents used on objects other than skin)
40:00	Electrolytic, Caloric, and Water Balance
44:00	Enzymes
48:00	Respiratory Tract Agents
52:00	Eye, Ear, Nose, and Throat (EENT) Preparations
56:00	Gastrointestinal Drugs
60:00	Gold Compounds
64:00	Heavy Metal Antagonists
68:00	Hormones and Synthetic Substitutes
72:00	Local Anesthetics
76:00	Oxytocics
78:00	Radioactive Agents
80:00	Serums, Toxoids, and Vaccines
84:00	Skin and Mucous Membrane Agents
86:00	Smooth Muscle Relaxants
88:00	Vitamins
92:00	Miscellaneous Therapeutic Agents
94:00	Devices
96:00	Pharmaceutical Aids

ภาพผนวก ก.1 แสดงกลุ่มเภสัชวิทยาและการรักษาของหนังสือ AHFS Drug Information 2013 ที่มา: http://www.ahfsdruginformation.com/class/index.aspx เข้าถึงเมื่อวันที่ 19 เมษายน 2556

Omeprazole

56:28.36 Proton-pump Inhibitors (AHFS primary)

Omeprazole Magnesium

56:28.36 Proton-pump Inhibitors (AHFS primary)

■ Omeprazole, commonly referred to as an acid- or proton-pump inhibitor, is a gastric antisecretory agent.

Uses

Omeprazole is used in adults for the short-term treatment of active duodenal and benign gastric ulcer. Omeprazole also is used in combination with clarithromycin (dual therapy) or with amoxicillin and clarithromycin (triple therapy) for the treatment of *Helicobacter pylori* infection and duodenal ulcer disease in adults. Omeprazole also has been used in other multiple-drug regimens (with or without clarithromycin)† for the treatment of *H. pylori* infection associated with peptic ulcer disease. Omeprazole is used in adults and children 1 year of age and older for short-term treatment and symptomatic relief of gastroesophageal reflux disease (e.g., erosive esophagitis, heartburn) and as maintenance therapy following healing of erosive esophagitis to reduce its recurrence. The drug also is used as *self-medication* for short-term treatment and symptomatic relief of frequent heartburn in adults. Omeprazole is used for the long-term treatment of pathologic GI hypersecretory conditions in adults. Omeprazole also is used to decrease the risk of upper GI bleeding in critically ill adults.

■ Duodenal Ulcer

Acute Therapy

Omeprazole immediate- and delayed-release capsules, omeprazole immediate-release oral suspension, and omeprazole magnesium delayed-release oral suspension are used in adults for the short-term treatment of endoscopically or radiographically confirmed active duodenal ulcer. Antacids may be used concomitantly as needed for pain relief. In controlled studies in patients with endoscopically confirmed duodenal ulcers, reported rates of ulcer healing for omeprazole were substantially higher than those for placebo. In a multicenter, double-blind study in patients with endoscopically confirmed duodenal ulcer, reported rates of ulcer healing for an oral omeprazole dosage of 20 mg each morning or placebo were 41 or 13%, respectively, at 2 weeks and 75 or 27%, respectively, at 4 weeks. Omeprazole also produced greater reductions in daytime and nocturnal pain and antacid consumption than did placebo, with complete relief of pain in most patients usually occurring within 4 weeks after initiation of omeprazole therapy.

Omeprazole appears to be at least as effective as histamine H_2-receptor antagonists for short-term treatment of active duodenal ulcer. In a multicenter, controlled study in patients with endoscopically confirmed duodenal ulcers, 42 or 34% of ulcers were healed following oral administration of omeprazole 20 mg each morning or ranitidine 150 mg twice daily, respectively, for 2 weeks and 82 or 63%, respectively, were healed after 4 weeks of therapy. In another multicenter, controlled study in patients with endoscopically confirmed duodenal ulcers, ulcer healing occurred faster in patients given omeprazole 20 or 40 mg daily compared with patients given ranitidine 150 mg twice daily. Ulcer healing rates averaged 83 or 53% at 2 weeks, 97–100 or 82% at 4 weeks, and 100 or 94% at 8 weeks with the omeprazole regimens or ranitidine 150 mg twice daily, respectively. In several studies, ulcer healing was less likely in patients who were smokers and in those with large ulcers than in other patients.

Most patients with duodenal ulcer respond to omeprazole therapy during the initial 4-week course of therapy; an additional 4 weeks of therapy may contribute to healing in some patients.

Omeprazole delayed-release capsules and omeprazole magnesium delayed-release oral suspension are used in combination with clarithromycin and amoxicillin (triple therapy) for the treatment of *H. pylori* infection and duodenal ulcer disease in adults. Omeprazole delayed-release capsules and omeprazole magnesium delayed-release oral suspension also are used in combination with clarithromycin (dual therapy) in adults for the treatment of *H. pylori* infection and duodenal ulcer disease. Omeprazole also has been used in other multiple-drug regimens† for the treatment of *H. pylori* infection associated with peptic ulcer disease. Current epidemiologic and clinical evidence supports a strong association between gastric infection with *H. pylori* and the pathogenesis of duodenal and gastric ulcers; long-term *H. pylori* infection also has been implicated as a risk factor for gastric cancer. For additional information on the association of this infection with these and other GI conditions, see Helicobacter pylori Infection, under Uses, in Clarithromycin 8:12.12.92. Conventional antiulcer therapy with histamine H_2-receptor antagonists, proton-pump inhibitors, sucralfate, and/or antacids heals ulcers but generally is ineffective in eradicating *H. pylori*, and such therapy is associated with a high rate of ulcer recurrence (e.g., 60–100% per year). The American College of Gastroenterology (ACG), the National Institutes of Health (NIH), and most clinicians currently recommend that *all* patients with initial or recurrent duodenal or gastric ulcer and documented *H. pylori* infection receive anti-infective therapy for treatment of the infection. Although 3-drug regimens consisting of a bismuth salt (e.g., bismuth subsalicylate) and 2 anti-infective agents (e.g., tetracycline or amoxicillin plus metronidazole) administered for 10–14 days have been effective in eradicating the infection, resolving associated gastritis, healing peptic ulcer, and preventing ulcer recurrence in many patients with *H. pylori*-associated peptic ulcer disease, current evidence principally from studies in Europe suggests that 1 week of such therapy provides comparable *H. pylori* eradication rates. Other regimens that combine one or more anti-infective agents (e.g., clarithromycin, amoxicillin) with a bismuth salt and/or an antisecretory agent (e.g., omeprazole, lansoprazole, histamine H_2-receptor antagonist) also have been used successfully for *H. pylori* eradication, and the choice of a particular regimen should be based on the rapidly evolving data on optimal therapy, including consideration of the patient's prior exposure to anti-infective agents, the local prevalence of resistance, patient compliance, and costs of therapy.

Current evidence suggests that inclusion of a proton-pump inhibitor (e.g., omeprazole, lansoprazole) in anti-*H. pylori* regimens containing 2 anti-infectives enhances effectiveness, and limited data suggest that such regimens retain good efficacy despite imidazole (e.g., metronidazole) resistance. Therefore, the ACG and many clinicians currently recommend 1 week of therapy with a proton-pump inhibitor and 2 anti-infective agents (usually clarithromycin and amoxicillin or metronidazole), or a 3-drug, bismuth-based regimen (e.g., bismuth-metronidazole-tetracycline) concomitantly with a proton-pump inhibitor, for treatment of *H. pylori* infection.

Therapy with an antisecretory drug and a single anti-infective agent (i.e., "dual therapy") also has been used successfully for treatment of *H. pylori* infection. However, while some studies demonstrate that certain 2-drug anti-*H. pylori* regimens (e.g., clarithromycin-omeprazole, ranitidine bismuth citrate-omeprazole, amoxicillin-omeprazole) can successfully eradicate *H. pylori* infection and prevent recurrence of duodenal ulcer at least in the short term (e.g., at 6 months following completion of anti-*H. pylori* therapy), the ACG and some clinicians currently state that anti-*H. pylori* regimens consisting of at least 3 drugs (e.g., 2 anti-infective agents plus a proton-pump inhibitor) are recommended because of enhanced *H. pylori* eradication rates, decreased failures due to resistance, and shorter treatment periods compared with those apparently required with 2-drug regimens. Additional randomized, controlled studies comparing various anti-*H. pylori* regimens are needed to clarify optimum drug combinations, dosages, and durations of treatment for *H. pylori* infection. For a more complete discussion of *H. pylori* infection, including details about the efficacy of various regimens and rationale for drug selection, see Uses: Helicobacter pylori Infection, in Clarithromycin 8:12.12.92.

■ Gastric Ulcer

Acute Therapy

Omeprazole immediate- and delayed-release capsules, omeprazole immediate-release oral suspension, and omeprazole magnesium delayed-release oral suspension are used in adults for the short-term treatment and symptomatic relief of active benign gastric ulcer. In controlled studies in patients with endoscopically confirmed gastric ulcers, reported rates of ulcer healing with omeprazole therapy were substantially higher than those with placebo. In a multicenter, double-blind study in patients with endoscopically confirmed gastric ulcer, reported rates of ulcer healing with omeprazole 20 or 40 mg daily or placebo were 48, 56, or 31%, respectively, at 4 weeks and 75, 83, or 48%, respectively, at 8 weeks. In patients with an ulcer larger than 1 cm in size, the percentage of patients with healed ulcers at 8 weeks was greater with the 40-mg dosage than with the 20-mg dosage of omeprazole. Otherwise, for patients with smaller ulcers, no difference in ulcer healing rates between the 40- and 20-mg dosages was observed.

In a multicenter, comparative study in patients with endoscopically confirmed gastric ulcer, ulcer healing occurred at 4 weeks in 64 or 78% of patients receiving omeprazole 20 or 40 mg daily, respectively, compared with 56% of those receiving ranitidine 150 mg twice daily; at 8 weeks, 82, 91, or 78% of patients receiving omeprazole 20 mg daily, omeprazole 40 mg daily, or ranitidine 150 mg twice daily, respectively, had healed ulcers.

■ Crohn's Disease-associated Ulcers

Although evidence currently is limited, proton-pump inhibitors have been used for gastric acid-suppressive therapy as an adjunct in the symptomatic treatment of upper GI Crohn's disease†, including esophageal, gastroduodenal, and jejunoileal disease. The drugs have been used for symptomatic relief of upper GI symptoms and to promote healing of Crohn's disease-associated peptic ulcer disease. Most evidence of efficacy to date has been from case studies in patients with Crohn's-associated peptic ulcer disease unresponsive to other therapies (e.g., histamine H_2-receptor antagonists, cytoprotective agents, antacids, and/or sucralfate). Omeprazole (20 or 40 mg daily) was associated with resolution of symptoms and ulcer healing within about 2 and 4 weeks, respectively, in some patients, while others required several months of acid-suppressive therapy. Subsequent symptomatic relief may be maintained with prolonged acid-suppressive therapy with a proton-pump inhibitor or H_2-receptor antagonist, with or without an immunosuppressive agent (e.g., azathioprine). Adjunctive inhibition

ภาพผนวก ก.2 แสดงตัวอย่างเอกสารข้อมูลยา (drug monograph)

ที่มา: http://www.ahfsdruginformation.com/class/index.aspx เข้าถึงเมื่อวันที่ 19 เมษายน 2556

of gastric acid secretion is likely to be more effective in promoting ulcer healing in Crohn's disease than corticosteroid therapy. Pending accumulation of more definitive evidence, some experts and clinicians state that therapy with a proton-pump inhibitor may be a useful adjunct to provide symptomatic relief and promote ulcer healing in patients with upper GI Crohn's disease.

For further information on the management of Crohn's Disease, see Uses: Crohn's Disease, in Mesalamine 56:36.

■ Gastroesophageal Reflux

Omeprazole delayed-release capsules and omeprazole magnesium delayed-release oral suspension are used in adults and children 1 year of age and older, and omeprazole immediate-release capsules and immediate-release oral suspension are used in adults for the short-term treatment and symptomatic relief of gastroesophageal reflux disease (GERD) (e.g., erosive esophagitis, heartburn) and as maintenance therapy following healing of erosive esophagitis to prevent its recurrence. Safety and efficacy of omeprazole immediate-release capsules and immediate-release oral suspension have not been established in pediatric patients. Omeprazole magnesium delayed-release capsules are used in adults as *self-medication* for the short-term treatment and symptomatic relief of frequent heartburn.

GERD is considered to be a chronic disease, and many patients with GERD require long-term, even lifelong, treatment. Typical GERD symptoms include heartburn and/or regurgitation, often occurring after meals, especially large and/or fatty meals. The symptoms often are aggravated by recumbency or bending, and are relieved by antacids. GERD symptoms generally are controlled by appropriate medical therapy. Suppression of gastric acid secretion is considered by the ACG to be the mainstay of treatment for GERD, and a proton-pump inhibitor or histamine H_2-receptor antagonist is used to achieve acid suppression, control symptoms, and prevent complications of the disease. The ACG states that proton-pump inhibitors are more effective than histamine H_2-receptor antagonists for acute therapy of GERD and also are appropriate as maintenance therapy in many patients with the disease. Lifestyle modifications (e.g., elevation of the head of the bed, decreased dietary fat intake, smoking cessation, avoidance of recumbency for 3 hours after a meal, avoidance of foods that increase reflux, weight loss) should be initiated and continued throughout the course of treatment.

Acute Therapy

Omeprazole delayed-release capsules and omeprazole magnesium delayed-release oral suspension are used in adults and children 1 year of age and older, and omeprazole immediate-release capsules and immediate-release oral suspension are used in adults for the short-term (4–8 weeks) treatment of endoscopically diagnosed erosive esophagitis in patients with GERD. Omeprazole delayed-release capsules and omeprazole magnesium delayed-release oral suspension are used in adults and children 1 year of age and older, and omeprazole immediate-release capsules and immediate-release oral suspension are used in adults for the short-term (4–8 weeks) treatment of symptomatic GERD (e.g., heartburn). Potential benefits of omeprazole in gastroesophageal reflux and esophagitis are thought to result principally from reduced acidity of gastric contents induced by the drug and resultant reduced irritation of esophageal mucosa; the drug can effectively relieve symptoms of esophagitis (e.g., heartburn) and promote healing of ulcerative and erosive lesions.

Drug Selection Considerations

The ACG states that proton-pump inhibitors are more effective (i.e., provide more frequent and more rapid symptomatic relief and healing of esophagitis) than histamine H_2-receptor antagonists in the treatment of GERD. Although higher doses and more frequent administration of histamine H_2-receptor antagonists appear to increase their efficacy, such dosages are less effective than proton-pump inhibitor therapy. In addition, the ACG states that proton-pump inhibitors provide greater control of acid reflux than do prokinetic agents (e.g., cisapride [no longer commercially available in the US], metoclopramide) without the risk of severe adverse effects associated with these agents. Correction of esophageal and gastric motility defects that cause GERD might theoretically control the disease and make suppression of normal gastric acid secretion unnecessary, and prokinetic agents have been used in the treatment of GERD. However, cisapride was withdrawn from the US market because of its association with serious cardiac arrhythmias and death (see Cisapride 56:32), and metoclopramide frequently is associated with adverse CNS effects (e.g., restlessness, drowsiness, fatigue, lassitude). The ACG states that the frequent occurrence of adverse CNS effects has appropriately decreased regular use of metoclopramide for treatment of GERD. Cisapride or metoclopramide therapy appears to provide symptomatic relief and esophageal healing as effectively as a standard dosage of a histamine H_2-receptor antagonist, and improved efficacy has been reported when a prokinetic agent has been used in combination with a histamine H_2-receptor antagonist. Bethanechol, a cholinergic drug that increases GI motility, may increase lower esophageal sphincter pressure to a small degree, but the ACG states that the drug has limited efficacy in the treatment of GERD.

The ACG states that a histamine H_2-receptor antagonist administered daily in divided doses is effective in many patients with less severe GERD, and over-the-counter (OTC) antacids and histamine H_2-receptor antagonists are appropriate for *self-medication* as initial therapy in such individuals. A histamine H_2-receptor antagonist is particularly useful when taken before certain activities (e.g., heavy meal, exercise) that may result in acid reflux symptoms in some patients.

Other Considerations

The ACG states that initial empiric therapy including suppression of gastric acid secretion and lifestyle modification is appropriate for patients with typical symptoms of uncomplicated GERD, and a diagnosis of GERD is reasonably assumed in those who respond to such therapy. Diagnostic testing (e.g., endoscopy, endoscopic biopsy, ambulatory pH testing, esophageal manometry) may be indicated when empiric drug therapy is unsuccessful, continuous medical therapy is required for symptomatic relief, chronic symptoms occur in patients at risk for esophageal metaplasia (e.g., Barrett's epithelium), or manifestations suggestive of complicated disease (e.g., dysphagia, bleeding, weight loss, choking [acid causing cough, shortness of breath, or hoarseness], chest pain) occur. In patients with symptoms refractory to empiric drug therapy, the diagnosis of GERD should be carefully confirmed with diagnostic testing before chronic, high-dose acid-suppression therapy or antireflux surgery is undertaken. Higher dosage and a longer therapeutic trial of a gastric antisecretory agent may be required in patients with atypical or extraesophageal symptoms (e.g., chronic chest pain, cough, hoarseness, asthma, dental erosions).

Clinical Trials

In a controlled study in patients with manifestations of GERD (e.g., heartburn) and the absence of erosive esophageal lesions, symptomatic improvement with omeprazole was better than that with placebo. Complete resolution of heartburn was reported in 56, 36, or 14% of patients with endoscopically confirmed GERD and in 46, 31, or 13% of all enrolled patients after up to 4 weeks of therapy with omeprazole 20 mg daily, omeprazole 10 mg daily, or placebo, respectively.

In an uncontrolled, open-label study of 113 pediatric patients 2–16 years of age with a history of symptoms suggestive of nonerosive GERD, patients received an omeprazole dosage of 10 or 20 mg once daily (based on body weight) either as an intact capsule or as an open capsule in applesauce. The number and intensity of either pain-related symptoms or vomiting/regurgitation episodes was successfully reduced in 60 or 59% of those receiving omeprazole 10 or 20 mg, respectively. In another uncontrolled study in 12 children 1–2 years of age with a history of clinically diagnosed GERD, administration of omeprazole (0.5–1.5 mg/kg as an opened capsule in 8.4% sodium bicarbonate solution) for 8 weeks reduced episodes of vomiting/regurgitation from baseline by at least 50% in 9 patients (75%).

In controlled studies in patients with endoscopically diagnosed erosive esophagitis and symptoms of GERD, reported rates of healing with omeprazole were higher than those with placebo or an H_2-receptor antagonist. Healing rates from a controlled study were 39, 45, or 7% at 4 weeks and 74, 75, or 14% at 8 weeks for omeprazole 20 mg daily, 40 mg daily, or placebo, respectively. In controlled studies in patients with esophagitis, reported rates of healing were 57–74 or 27–43% at 4 weeks and 78–87 or 28–56% at 8 weeks in patients given omeprazole or ranitidine, respectively. Patients receiving omeprazole reported faster relief of daytime and nocturnal heartburn than those receiving placebo or an H_2-receptor antagonist. Omeprazole also has been shown to be effective in promoting healing and providing symptomatic relief in a substantial proportion of patients who failed to respond to an adequate course of relatively high dosages of an H_2-receptor antagonist.

In an uncontrolled, open-label dose-titration study in 57 pediatric patients aged 1–16 years of age with erosive esophagitis, omeprazole dosages of 0.7–3.5 mg/kg daily were required to promote healing. Dosages were initiated at 0.7 mg/kg daily and if therapeutic goals (intraesophageal pH below 4 for less than 6% of a 24-hour period) were not achieved after 5–14 days of treatment, the dosage was increased to 1.4 mg/kg daily. Based on additional measurements of intraesophageal pH and/or presence of pathologic acid reflux, the dosages were increased up to a maximum dosage of 3.5 mg/kg or 80 mg daily. After titration of omeprazole dosage, patients remained on treatment for 3 months (healing phase); patients with persistent erosive esophagitis after 3 months received a discretionary dosage increase and treatment for an additional 3 months. Erosive esophagitis was healed in 90% of children completing the first course of treatment in the healing phase of the study; 5% received a second treatment course. Healing occurred in 44% of the patients receiving omeprazole 0.7 mg/kg daily, and an additional 28% were healed with 1.4 mg/kg daily. After 3 months of treatment, 33% of the children had no overall symptoms, 57% had mild reflux symptoms, and 40% had less frequent regurgitation or vomiting.

Most patients with GERD respond to omeprazole therapy during an initial 8-week course of therapy; however, an additional 4 weeks of therapy may contribute to healing and symptomatic improvement in some patients. Short-term omeprazole therapy for the treatment of GERD will not prevent recurrence following discontinuance of the drug. If symptomatic GERD or erosive esophagitis recur, the manufacturers state that additional 4- to 8-week courses of omeprazole may be given. However, the ACG states that chronic therapy with a proton-pump inhibitor is appropriate in many patients with GERD.

Maintenance Therapy

Omeprazole delayed-release capsules and omeprazole magnesium delayed-release oral suspension are used in adults and children 1 year of age and older, and

omeprazole immediate-release capsules and immediate-release oral suspension are used in adults as maintenance therapy following healing of erosive esophagitis to reduce recurrence of the disease. In a multicenter, double-blind study, endoscopically documented remission of esophagitis was maintained at 6 months in 70, 34, or 11% of patients receiving omeprazole 20 mg daily, 20 mg on 3 consecutive days each week, or placebo, respectively. In another multicenter, double-blind study in patients with endoscopically confirmed healed esophagitis, endoscopic remission of esophagitis was maintained at 12 months in 77, 58, or 46% of patients receiving omeprazole 20 mg daily, 10 mg daily, or ranitidine 150 mg twice daily, respectively. However, patients with initial grade 3 or 4 erosive esophagitis required 20 mg of omeprazole daily for maintenance of healing.

In an uncontrolled, open-label study in 46 pediatric patients, maintenance dosages were half the dosages that were required for promotion of healing in 54% of the children studied. The remaining patients required a dosage increase (0.7 to a maximum of 2.8 mg/kg daily) for all or part of the maintenance period. There was no relapse of erosive esophagitis in 41% of the patients, and no symptoms occurred in 63% of the pediatric patients receiving omeprazole maintenance therapy.

Because GERD is a chronic condition, the ACG states that continuous therapy to control symptoms and prevent complications of the disease is appropriate, and chronic, even lifelong, use of a proton-pump inhibitor is effective and appropriate as maintenance therapy in many patients with GERD. Although neither medical nor surgical therapy of GERD appears to result in regression of Barrett's epithelium in the esophagus, chronic use of a proton-pump inhibitor at full dosage decreases the recurrence of esophageal strictures, increases the interval between symptomatic relapses, and may improve esophageal motility. In a double-blind, controlled study, antisecretory therapy had no clinically important effect on Barrett's mucosa in 106 patients receiving omeprazole (40 mg twice daily for 12 months, followed by 20 mg twice daily for 12 months) or ranitidine (300 mg twice daily for 24 months). Although neosquamous epithelium developed during antisecretory therapy, complete elimination of Barrett's mucosa was not achieved.

The frequent marked improvement in symptoms associated with full dosage of a proton-pump inhibitor generally is followed by rapid recurrence of symptoms once the drug is discontinued, and reduced-dosage regimens (e.g., every other day, "weekend" dosage) have not been shown to be consistently effective for maintenance therapy. In addition, many patients initially responding to proton-pump inhibitors experience symptomatic relapse and failure of esophageal healing when switched subsequently to a histamine H_2-receptor antagonist or prokinetic agent (e.g., cisapride, metoclopramide). Furthermore, prokinetic agents have been associated with severe adverse effects. Cisapride has been withdrawn from the US market because of its association with serious cardiac arrhythmias and death (see Cisapride 56:32), and metoclopramide frequently is associated with CNS adverse effects (e.g., restlessness, drowsiness, fatigue, lassitude) and may cause irreversible tardive dyskinesia with prolonged use. Once-daily administration of a histamine H_2-receptor antagonist at full dosage is *not* considered to be appropriate therapy for GERD. Although antacids and lifestyle modifications may provide long-term symptomatic control in up to 20% of patients with GERD, frequent symptomatic relapses may occur despite appropriate therapy in up to 50% of patients with chronic gastroesophageal reflux.

Self-Medication

Omeprazole magnesium delayed-release capsules are used in adults 18 years of age or older as *self-medication* for short-term (14 days) treatment and symptomatic relief of frequent (e.g., 2 or more days a week) heartburn. Because 1–4 days may be required for complete relief of symptoms, omeprazole for *self-medication* is not intended for the immediate relief of heartburn, and other agents (e.g., antacids, histamine H_2-receptor antagonists) may be needed for initial relief. However, some individuals may experience complete relief of symptoms within 24 hours of taking the first dose of omeprazole. In 2 controlled studies, 50% of patients receiving omeprazole 20 mg daily experienced no heartburn during the first day of therapy, and the percentage of patients experiencing complete relief continued to increase in subsequent days; 30% of those receiving placebo experienced no heartburn during the first day of therapy. Omeprazole should not be used for *self-medication* of occasional heartburn (i.e., heartburn that occurs once weekly or less frequently) or for prevention of occasional meal- or beverage-induced heartburn.

■ Pathologic GI Hypersecretory Conditions

Omeprazole delayed-release capsules and omeprazole magnesium delayed-release oral suspension are used in adults for the long-term treatment of pathologic GI hypersecretory conditions (e.g., Zollinger-Ellison syndrome, multiple endocrine adenomas, systemic mastocytosis). The drug reduces gastric acid secretion and associated symptoms (including diarrhea, anorexia, and pain) in patients with these conditions. In dosages ranging from 20 mg every other day to 360 mg daily, omeprazole can maintain basal acid secretion below 5 or 10 mEq/hour in patients who have or have not undergone gastric surgery, respectively. In addition, dosages ranging from 20–360 mg daily have been effective in resolving acid-related pathology in most patients with Zollinger-Ellison syndrome, including those whose symptoms were unresponsive to H_2-receptor antagonist therapy.

■ Upper GI Bleeding

Omeprazole immediate-release oral suspension is used to decrease the risk of upper GI bleeding in critically ill adults. Efficacy of omeprazole was evaluated in a controlled, double-blind randomized clinical trial in critically ill patients who were randomized to receive either omeprazole immediate-release oral suspension (2 doses of 40 mg 6–8 hours apart on the first day, then 40 mg daily) via a gastric tube or IV cimetidine (300 mg loading dose, then 50–100 mg/hour continuously) for up to 14 days. The primary efficacy end point of the study was clinically important upper GI bleeding (defined as bright red blood that did not clear after tube adjustment and 5–10 minutes of lavage or positive test for occult blood in gastric aspirate ["coffee ground material"] for 8 consecutive hours on days 1 and 2, or for 2–4 hours on days 3–14 that did not clear with 100 mL of lavage). Omeprazole was at least as effective as IV cimetidine in preventing clinically important upper GI bleeding. In the intent-to-treat population, clinically important gastric bleeding occurred in 3.9% of patients receiving omeprazole and in 5.5% of those receiving IV cimetidine.

Dosage and Administration

■ Administration

Omeprazole immediate-release and delayed-release capsules and omeprazole magnesium delayed-release tablets for *self-administration* are administered orally; oral suspensions of the drug are administered orally or through a gastric tube. To avoid decomposition of omeprazole in the acidic pH of the stomach, the commercially available delayed-release capsules and delayed-release oral suspension contain enteric-coated granules of the drug, and the immediate-release capsules and immediate-release oral suspension contain sodium bicarbonate.

Omeprazole immediate-release capsules (Zegerid®) are administered orally and must be swallowed intact with water; other liquids should not be used. The capsules should not be opened and mixed with food. Both the 20- and 40-mg capsules contain the same amount of sodium bicarbonate (1100 mg). Therefore, two 20-mg capsules are not equivalent to and should not be substituted for one 40-mg capsule.

Patients should be advised that the delayed-release capsules must be swallowed intact and *not* chewed or crushed. However, for adult and pediatric patients with difficulty swallowing, the delayed-release capsule may be opened, the contents carefully emptied on and mixed with a tablespoon of applesauce in a bowl, and the mixture swallowed immediately with a glass of cool water to ensure complete swallowing of the pellets. The applesauce should not be hot and should be soft enough to be swallowed without chewing. The applesauce and omeprazole enteric-coated pellet mixture should *not* be stored for future use. The manufacturer states that the 40-mg capsule, but not the 20-mg capsule, is bioequivalent when administered with or without applesauce. When the contents of a 20-mg capsule were administered with applesauce, the peak plasma omeprazole concentration decreased by 25%, but the area under the concentration-time curve (AUC) was not substantially changed. However, the clinical importance of this is unknown.

Tablets used for *self-medication* must be swallowed intact with a glass of water; the tablets should *not* be chewed or crushed and should *not* be crushed in food.

Omeprazole magnesium powder for delayed-release oral suspension (Prilosec®) should be reconstituted prior to administration by pouring the contents of a single-dose packet containing 2.5 or 10 mg of the drug into a small cup containing 5 or 15 mL, respectively, of water. The suspension should be stirred well and allowed to thicken for 2–3 minutes. Within 30 minutes of preparation, the mixture should be stirred and consumed. If any material remains in the cup after the mixture is ingested, additional water should be added, mixed, and ingested immediately. If the delayed-release oral suspension is to be administered through a nasogastric or gastric tube, the contents of a 2.5- or 10-mg packet should be mixed with 5 or 15 mL of water, respectively, in a catheter-tipped syringe and then shaken immediately. The mixture should be allowed to thicken for 2–3 minutes. The mixture should be administered within 30 minutes of reconstitution; prior to administration, the syringe should be shaken again and the mixture injected into the stomach through the nasogastric or gastric tube (French size 6 or larger). The syringe should be refilled with additional water (5 or 15 mL, respectively), shaken, and used to flush any remaining drug mixture from the nasogastric or gastric tube into the stomach.

Omeprazole powder for immediate-release oral suspension (Zegerid®) should be reconstituted prior to administration by pouring the contents of a single-dose packet containing 20 or 40 mg of the drug into a small cup containing 15–30 mL of water. The 20- and 40-mg powder for oral suspension packets contain the same amount of sodium bicarbonate (1680 mg). Therefore, two 20-mg packets are not equivalent to and should not be substituted for one 40-mg packet. The suspension should be stirred well and ingested immediately. The cup should be refilled with water and the contents ingested to ensure complete consumption of the dose. The manufacturer states that omeprazole powder for immediate-release oral suspension should not be mixed with any liquids (other than water) or foods. If the oral suspension is to be administered through a nasogastric or orogastric tube, the contents of each packet should be reconstituted with approximately 20 mL of water, stirred well and administered immediately. An

appropriate-sized syringe should be used to instill the suspension into the tube. The suspension should then be flushed through the tube with 20 mL of water.

Following administration of delayed-release capsules of omeprazole with meals, the rate of GI absorption is reduced. Therefore, omeprazole should be taken before meals; administration up to 2 minutes prior to a meal reportedly has no adverse effect on oral bioavailability. However, since an acidic environment in the parietal cell canaliculi is required for conversion of proton-pump inhibitors (e.g., omeprazole) to their active sulfenamide metabolites, the American College of Gastroenterology suggests that proton-pump inhibitors are most effective when given about 30 minutes prior to meals; effectiveness may be compromised if these drugs are administered during the basal state (e.g., to fasting patients at bedtime) or concomitantly with other antisecretory agents (e.g., anticholinergics, histamine H_2-receptor antagonists, somatostatin analogs, misoprostol). The manufacturer states that delayed-release preparations of omeprazole should be administered at least 1 hour before a meal. Antacids may be administered concomitantly with the delayed-release preparations of omeprazole.

The manufacturer states that immediate-release preparations of omeprazole (Zegerid®) should be administered on an empty stomach at least 1 hour prior to a meal. For patients receiving continuous feedings via a nasogastric or orogastric tube, enteral feeding should be stopped temporarily for 3 hours before, and for 1 hour after administration of omeprazole immediate-release oral suspension. Also, the manufacturer states that antacids, antacid/alginic acid combinations, histamine H_2-receptor antagonists, or histamine H_2-receptor antagonist and antacid combinations may be used for "breakthrough" symptoms; however, efficacy of these agents for this use has not been established.

Omeprazole usually is administered once daily in the morning; however, administering the drug in divided doses (e.g., every 12 hours) has been reported to improve efficacy in patients receiving more than 80 mg daily.

■ Dosage

The manufacturer states that omeprazole dosage adjustments based on age are not necessary in geriatric patients. However, since the bioavailability of omeprazole appears to be increased substantially in Asians, the manufacturer states that dosage adjustment should be considered in Asian patients, especially when such patients are receiving long-term omeprazole therapy for maintenance of healing of erosive esophagitis. There is no evidence from the omeprazole prescription safety database that Asians experience excess risk from omeprazole, or that accumulation of omeprazole in the blood is harmful when used over a short period of time (e.g., 14 days of *self-medication*) in Asian patients.

Dosage of omeprazole magnesium is expressed in terms of omeprazole.

Duodenal Ulcer

For the short-term treatment of active duodenal ulcer, the usual adult dosage of omeprazole is 20 mg once daily. Therapy should be continued until healing occurs, usually within 2–4 weeks; some patients may benefit from an additional 4 weeks of therapy. Occasionally, dosages up to 40 mg daily may be necessary in patients who have been poorly responsive to therapy with H_2-receptor antagonists.

When omeprazole is used in combination with clarithromycin (dual therapy) for the treatment of *Helicobacter pylori* infection in patients with active duodenal ulcer, the usual adult dosage of omeprazole is 40 mg once daily (in the morning) for 14 days. In patients who have an active ulcer present at the time anti-*H. pylori* therapy is initiated, an additional 14 days of therapy with omeprazole 20 mg once daily is recommended for ulcer healing and symptom relief. When omeprazole is used in combination with clarithromycin and amoxicillin (triple therapy) for the treatment of *H. pylori* infection in patients with active duodenal ulcer, the usual adult dosage of omeprazole is 20 mg *twice* daily (morning and evening) for 10 days. In patients who have an active ulcer present at the time anti-*H. pylori* therapy is initiated, an additional 18 days of therapy with omeprazole 20 mg *once* daily is recommended for ulcer healing and symptom relief. Multiple-drug regimens currently recommended by the American College of Gastroenterology (ACG) and many clinicians for the treatment of *H. pylori* infection consist of a proton-pump inhibitor (e.g., omeprazole) and 2 anti-infective agents (e.g., clarithromycin and amoxicillin or metronidazole) or a 3-drug, bismuth-based regimen (e.g., bismuth-metronidazole-tetracycline) concomitantly with a proton-pump inhibitor; when omeprazole has been used in these regimens, dosages of 20 mg once daily to 80 mg twice daily (generally 20 mg twice daily) for 7–28 days have been used.While the minimum duration of therapy required to eradicate *H. pylori* infection with these 3- or 4-drug regimens has not been fully elucidated, the ACG and many clinicians state that treatment for longer than 1 week probably is not necessary. However, more prolonged therapy is recommended for patients with complicated, large, or refractory ulcers; therapy in such patients should be continued at least until successful eradication of *H. pylori* has been confirmed. (See Uses: Helicobacter pylori Infection, in Clarithromycin 8:12.12.92.)

Gastric Ulcer

For the short-term treatment of active benign gastric ulcer, the usual adult dosage of omeprazole is 40 mg once daily for 4–8 weeks.

Gastroesophageal Reflux

For the short-term, symptomatic treatment of gastroesophageal reflux disease (GERD) without erosive esophageal lesions, the usual adult dosage of omeprazole is 20 mg once daily for 4 weeks. For the short-term treatment of erosive esophagitis, the usual adult dosage of omeprazole is 20 mg once daily for 4–8 weeks. Occasionally, dosages up to 40 mg daily may be necessary in some patients. Therapy is continued until healing occurs, usually within 4–8 weeks; an additional 4 weeks of therapy (up to 12 weeks for a single course) may contribute to healing and symptomatic improvement in some patients. If erosive esophagitis or symptomatic GERD (heartburn) recurs, the manufacturer states that additional 4- to 8-week courses of omeprazole may be considered. However, the American College of Gastroenterology (ACG) states that chronic, even lifelong, therapy with a proton-pump inhibitor is appropriate in many patients with GERD.

For maintenance therapy following healing of erosive esophagitis to reduce recurrence, the usual adult dosage of omeprazole is 20 mg daily. Safety and efficacy of omeprazole maintenance therapy for longer than 1 year have not been established.

For the treatment of symptomatic GERD or erosive esophagitis and for maintenance of healing of erosive esophagitis in pediatric patients 1–16 years of age, a dosage of 5 mg of omeprazole daily is recommended for pediatric patients weighing at least 5 kg but less than 10 kg, 10 mg daily is recommended for those weighing at least 10 kg but less than 20 kg, and 20 mg daily is recommended for those weighing 20 kg or more. Omeprazole was administered as a single daily dose for 4 weeks in one study of children with symptomatic nonerosive GERD. On a mg/kg basis, the dosage of omeprazole required to heal erosive esophagitis in pediatric patients is greater than that required in adults. In an uncontrolled open-label study, dosages of 0.7–3.5 mg/kg daily (up to a maximum dosage of 80 mg daily) for 3–6 months were required for healing in children 1–16 years of age; a dosage of 0.7 mg/kg daily resulted in healing of erosive esophagitis in 44% of children, but a dosage of 1.4 mg/kg daily was required for healing to occur in an additional 28% of the children. In an uncontrolled open-label study of 46 pediatric patients, dosages of omeprazole for maintenance therapy following healing of erosive esophagitis were half those required for initial healing in 54% of children, but the remainder required an increased dosage (0.7 to a maximum of 2.8 mg/kg daily) for all or part of the maintenance period; maintenance therapy was continued for about 2 years.

Self-Medication

For *self-medication* to relieve symptoms of frequent heartburn in adults 18 years of age or older, an omeprazole dosage of 20 mg once daily in the morning for 14 days is recommended. For *self-medication*, the manufacturer recommends that the dosage of omeprazole not exceed 20 mg in 24 hours. In addition, the drug should not be used for *self-medication* for longer than 14 days of *continuous* use and individuals should not exceed one course of therapy every 4 months unless otherwise directed by a clinician.

Pathologic GI Hypersecretory Conditions

For the treatment of pathologic GI hypersecretory conditions (e.g., Zollinger-Ellison syndrome, multiple endocrine adenomas, systemic mastocytosis), dosages of omeprazole should be individualized according to patient response and tolerance. The usual initial adult dosage is 60 mg (as delayed-release capsules or delayed-release oral suspension) once daily. Subsequent omeprazole dosage should be adjusted as tolerated and necessary to adequately suppress gastric acid secretion, and therapy continued as long as clinically necessary. Daily dosages exceeding 80 mg should be administered in divided doses.

Oral dosages ranging from 20 mg every other day to 360 mg daily (given in 3 divided doses) have been necessary to maintain basal gastric acid secretion at less than 10 mEq/hour in patients without a history of gastric surgery and less than 5 mEq/hour in those who have undergone gastric surgery; determination of gastric acid secretion during the hour prior to a dose may be useful in establishing optimum dosage. Omeprazole has been given continuously for more than 5 years in some patients with Zollinger-Ellison syndrome.

Upper GI Bleeding

For reduction of risk of upper GI bleeding in critically ill adults, the initial loading dose of omeprazole is 40 mg (as immediate-release oral suspension) followed by another 40-mg dose after 6–8 hours on the first day; thereafter, 40 mg (as immediate-release oral suspension) is administered once daily for up to 14 days. Safety and efficacy of omeprazole immediate-release oral suspension in critically ill patients for longer than 14 days have not been established.

■ Dosage in Renal and Hepatic Impairment

Although pharmacokinetics may be altered in patients with renal impairment, dosage adjustment does not appear necessary in patients with such impairment. However, the manufacturers state that dosage adjustment should be considered in patients with hepatic impairment, particularly in such patients receiving long-term omeprazole therapy for maintenance of healing of erosive esophagitis. Some clinicians recommend that such patients with hepatic dysfunction receiving dosages exceeding 20 mg daily should be monitored closely for possible adverse effects.

Cautions

Omeprazole generally is well tolerated. The most frequent adverse effects associated with omeprazole therapy involve the GI tract (e.g., diarrhea, nausea, constipation, abdominal pain, vomiting, flatulence) and the CNS (e.g., headache, dizziness). In short-term studies, the incidence of reported adverse effects was similar in patients receiving omeprazole or placebo. In addition, while the most common effects have been reported in 1–7% of patients receiving omeprazole, they were considered by investigators as being possibly, probably, or definitely related to the drug in only 0.2–2.4% of patients. Overall, the frequency and type of adverse effects produced by omeprazole appear to be similar to those produced by ranitidine, and the frequency of omeprazole-induced effects does not appear to be affected by age in adults. In dose-ranging studies, a relationship between doses ranging from 10–60 mg and the frequency of adverse effects was not observed. Adverse effects were severe enough to result in discontinuance of omeprazole therapy in less than 2% of patients in clinical studies. The manufacturer states that the adverse event profile of omeprazole in pediatric patients is similar to that in adults. However, the most frequently reported adverse effects in pediatric patients were respiratory effects, which were reported in about 75 or 18% of those 1–2 or 2–16 years of age, respectively; fever was frequently reported (33%) in children 1–2 years of age, and accidental injuries were frequently reported (about 4%) in those 2–16 years of age.

In controlled clinical trials with combined omeprazole-clarithromycin or omeprazole-clarithromycin-amoxicillin therapy, no adverse drug experiences peculiar to these combinations were noted.

In a controlled clinical trial, the adverse event profile was similar for critically ill patients receiving either omeprazole immediate-release suspension or IV cimetidine for up to 14 days. The most frequent adverse effects reported in patients receiving omeprazole were pyrexia (20.2%), hypokalemia (12.4%), nosocomial pneumonia (11.2%), hyperglycemia (10.7%), thrombocytopenia (10.1%), hypomagnesemia (10.1%), and hypotension (9.6%).

■ GI Effects

Diarrhea, abdominal pain, nausea, vomiting, constipation, flatulence, and acid regurgitation are the most frequent adverse GI effects reported with omeprazole therapy, occurring in about 1–5% of patients. Constipation, diarrhea, and gastric hypomotility occurred in 4.5, 3.9, and 1.7%, respectively, of critically ill patients receiving omeprazole immediate-release oral suspension or in 4.4, 8.3, and 3.3%, respectively, of those receiving IV cimetidine in a controlled clinical trial. Dysphagia, abdominal swelling, anorexia, irritable colon, fecal discoloration, pancreatitis (sometimes fatal), esophageal candidiasis, mucosal atrophy of the tongue, taste perversion, dry mouth, stomatitis, and microscopic colitis have been reported during postmarketing surveillance in patients receiving omeprazole; a causal relationship to the drug was not established in many cases. Benign gastric fundic polyps have been reported rarely and appear to resolve upon discontinuation of omeprazole therapy. Long-term administration of omeprazole has produced dose-related increases in gastric carcinoid tumors and enterochromaffin-like (ECL) cell hyperplasia in rats. Carcinoid tumors also have been observed in rats subjected to fundectomy or long-term treatment with other proton-pump inhibitors or high dosages of H_2-receptor antagonists. Gastric biopsy specimens obtained from patients in long-term studies with omeprazole have demonstrated an increased frequency of ECL cell hyperplasia. However, no cases of ECL cell carcinoid tumor, dysplasia, or neoplasia were found. (See Cautions: Mutagenicity and Carcinogenicity.)

Adverse GI effects observed in controlled trials with combined omeprazole and clarithromycin therapy that were not reported with omeprazole monotherapy include taste perversion in 15% of such patients and tongue discoloration in 2%.

As with other agents that elevate intragastric pH, administration of omeprazole for 14 days in healthy individuals increased the intragastric concentration of viable bacteria. The pattern of bacteria isolated was similar to that of saliva. Alterations in the intragastric bacterial flora were reversible following discontinuance of omeprazole. Treatment with proton-pump inhibitors may result in a slight increase in the risk of GI infections caused by such organisms as *Salmonella* and *Campylobacter* species.

Available data suggest a possible association between use of proton-pump inhibitors and risk of *Clostridium difficile* infection, including *C. difficile*-associated diarrhea and colitis (CDAD; also known as antibiotic-associated diarrhea and colitis or pseudomembranous colitis). In most observational studies to date, the risk of *C. difficile* infection in patients exposed to proton-pump inhibitors has ranged from 1.4–2.75 times that in patients not exposed to proton-pump inhibitors; however, some observational studies have found no increase in risk. Although many of the cases occurred in patients who had other risk factors for CDAD, including advanced age, comorbid conditions, and/or use of broad-spectrum anti-infectives, the US Food and Drug Administration (FDA) concluded that a contributory role for proton-pump inhibitors could not be definitively ruled out. The mechanism by which proton-pump inhibitors might increase the risk of CDAD has not been elucidated. Although it has been suggested that reduction of gastric acidity by gastric antisecretory agents might facilitate colonization with *C. difficile*, some studies have raised questions about this proposed mechanism or have suggested that the observed association is the result of confounding with other risk factors for CDAD. FDA also is reviewing the risk of CDAD in patients exposed to histamine H_2-receptor antagonists.

CDAD can be serious in patients who have one or more risk factors for *C. difficile* infection and are receiving concomitant therapy with a proton-pump inhibitor; colectomy and, rarely, death have been reported. FDA recommends that patients receive proton-pump inhibitors at the lowest effective dosage and for the shortest possible time appropriate for their clinical condition. Patients experiencing persistent diarrhea should be evaluated for CDAD and should be managed with appropriate supportive therapy (e.g., fluid and electrolyte management), anti-infective therapy directed against *C. difficile* (e.g., metronidazole, vancomycin), and surgical evaluation as clinically indicated.

■ Nervous System Effects

Headache and dizziness are the most common adverse nervous system effects of omeprazole, occurring in 6.9 and 1.5%, respectively, of patients in US clinical studies. In a controlled clinical trial in critically ill patients, agitation occurred in 3.4 or 8.8% of patients receiving omeprazole immediate-release oral suspension or IV cimetidine, respectively. Asthenia has been reported in 1.1–1.3% of patients receiving omeprazole; in controlled studies, the incidence of this effect was similar in patients receiving omeprazole, ranitidine, or placebo. Psychic and sleep disturbances, including depression, agitation, aggression, hallucinations, confusion, insomnia, nervousness, tremors, apathy, somnolence, anxiety, and dream abnormalities, have been reported during postmarketing surveillance in patients receiving omeprazole; a causal relationship to the drug was not established in many cases. Other infrequent nervous system effects for which a causal relationship may not have been established include pain, fatigue, malaise, vertigo, paresthesia, and hemifacial dysesthesia.

■ Respiratory Effects

Upper respiratory tract infections and cough have occurred in 1.9 and 1.1%, respectively, of patients receiving omeprazole; in controlled studies, the incidence of these effects was similar in patients receiving omeprazole, ranitidine, or placebo. Acute respiratory distress syndrome, respiratory failure, and pneumothorax occurred in 3.4, 1.7, and 0.6%, respectively, of critically ill patients receiving omeprazole immediate-release oral suspension or in 3.9, 3.3, and 4.4%, respectively, of those receiving IV cimetidine in a controlled clinical trial. Epistaxis and pharyngeal pain have been reported during postmarketing surveillance in patients receiving omeprazole; a causal relationship to the drug was not established. Adverse respiratory effects have been reported in about 75% of children 1–2 years of age and in about 18% of those 2–16 years of age.

Other adverse respiratory effects observed in controlled trials with combined omeprazole and clarithromycin therapy that were not reported with omeprazole monotherapy were rhinitis in 2% of patients, pharyngitis in 1%, and flu syndrome in 1%.

Community-acquired Pneumonia

Administration of gastric antisecretory agents (e.g., proton-pump inhibitors, H_2-receptor antagonists) has been associated with an increased risk for developing certain infections (e.g., community-acquired pneumonia). A possible association between chronic administration of gastric acid-suppressive drugs and occurrence of community-acquired pneumonia has been evaluated using a large Dutch database (Integrated Primary Care Information [IPCI]) containing information on approximately 500,000 patients, 364,683 of whom (average follow-up: 2.7 years) were selected for evaluating any such association. During the 8-year population-based, case-control study, gastric acid suppressants were first prescribed in 19,459 individuals (12,337 received proton-pump inhibitors [mean duration of use: 5 months] and 10,177 received H_2-receptor antagonists [mean duration of use: 2.8 months]; some individuals received both drugs). Most patients did not undergo endoscopy and were treated empirically for upper GI symptoms. In this study, first occurrence of pneumonia (confirmed by radiography or microbiologic testing in 18% of patients) was reported in 5551 individuals; development of pneumonia occurred in 185 individuals while receiving gastric acid suppressants and in 292 individuals who had discontinued such use.

The adjusted relative risk for development of pneumonia (or the incidence rate) was 0.6, 2.3 and 2.5 per 100 person-years for individuals not receiving acid-suppressive drugs, for those receiving H_2-receptor antagonists, and for those receiving proton-pump inhibitors, respectively. Patients using gastric acid suppressants developed community-acquired pneumonia 4.5 (95% confidence interval of 3.8–5.1) times more often than those who never used such drugs. When evaluating use of all gastric acid suppressants, current use of the drugs was associated with a small (27%) overall increase in the risk of pneumonia (adjusted odds ratio 1.27 and 95% confidence interval of 1.06–1.54). Higher risks were observed for current users of proton-pump inhibitors and H_2-receptor antagonists; the adjusted relative risk for developing community-acquired pneumonia was 1.89 (95% confidence interval of 1.36–2.62) or 1.63 (95% confidence interval of 1.07–2.48), respectively, for these classes of drugs compared with those who discontinued using these agents. Estimates for developing pneumonia were higher (2.2 [95% confidence interval of 1.4–3.5] for proton-pump inhibitors and 1.7 [95% confidence interval of 0.8–2.9] for H_2-receptor antagonists) when only laboratory-confirmed cases of pneumonia were considered for analysis.

Although there was variation among individual proton-pump inhibitors and individual H_2-receptor antagonists, the numbers were small and the heterogeneity was not considered significant. For patients currently receiving proton-pump inhibitors, a dose-response relationship for developing pneumonia was observed; individuals using more than one defined daily dose of these drugs had a 2.3-fold increased risk for developing pneumonia compared with those who discontinued gastric acid suppressants. Such a dose-response relationship for developing pneumonia was not observed in patients receiving H_2-receptor antagonists; however, dose variation of these drugs was limited. Among current users of proton-pump inhibitors or H_2-receptor antagonists, the risk for developing pneumonia was most pronounced among those who initiated such therapies within the past 30 days.

Although the exact mechanism for development of community-acquired pneumonia in patients receiving gastric acid suppressants has not been fully elucidated, it has been suggested that reduction of gastric acid secretion by acid suppressive therapy and consequent increases of gastric pH may result in a favorable environment for the development of infection. Intragastric acidity constitutes a major nonspecific defense mechanism of the stomach to ingested pathogens; when gastric pH is less than 4, most pathogens are killed, while at higher gastric pH, pathogens may survive. Since intragastric pH should be maintained above 4 for several hours for the effective management of upper GI symptoms, acid suppressive therapy may lead to insufficient elimination or even increased colonization of ingested pathogens. Some evidence indicates that acid-suppressive therapy may result in nosocomial infections.

It should be considered that certain patients (e.g., those with pleuritic chest pain, hypothermia, systolic hypotension, tachypnea, diabetes mellitus, neoplastic disease, neurologic disease, bacteremia, leukopenia, multilobar pulmonary infiltrate) are at increased risk for developing infections and in these individuals community-acquired pneumonia may be associated with increased mortality. Some clinicians state that gastric acid-suppressive drugs should be used in patients in whom community-acquired pneumonia may be severe (e.g., those with asthma or chronic obstructive lung disease, immunocompromised patients, pediatric or geriatric individuals) only when clearly needed and the lowest effective dose should be employed.

■ Musculoskeletal Effects

Back pain has been reported in about 1% of patients receiving omeprazole. Other musculoskeletal effects have been reported during postmarketing surveillance; a causal relationship to the drug was not established in many cases. Such effects include muscle cramps, myalgia, muscle weakness, joint pain, and leg pain. Bone fracture also has been reported during postmarketing surveillance in patients receiving omeprazole.

Findings from several observational studies suggest that therapy with proton-pump inhibitors, particularly in high dosages (i.e., multiple daily doses) and/or for prolonged periods of time (i.e., one year or longer), may be associated with an increased risk of osteoporosis-related fractures of the hip, wrist, or spine. The magnitude of the risk is not clear. To date, most of the observational studies assessing fracture risk in patients receiving proton-pump inhibitor therapy have limited the study population to individuals at least 50 years of age or older. Some of the studies found that the risk of hip fracture was increased with use of higher dosages of the drugs (e.g., average dosage of at least 1.5 "pills" daily) or with long-term use of the drugs, particularly long-term use of high dosages (e.g., dosages exceeding 1.75 times the average daily dosage for more than one year). Study results relating duration of proton-pump inhibitor use to emergence of increased fracture risk have been variable; one study found an increased risk of hip fracture after more than one year of use, whereas another study found an increased risk of hip fracture or osteoporosis-related fracture after 5 or 7 years of use, respectively. In yet another study, an increased risk of hip or spinal fracture was observed when the drugs were last used within the previous year but not when last use was more distant. One study that excluded patients with major risk factors for hip fracture found no relationship between proton-pump inhibitor use and hip fracture occurrence. Because these observational studies relied extensively on claims data from computerized administrative databases, the clinical relevance of reported findings is difficult to determine. FDA states that a causal relationship between proton-pump inhibitor use and fracture occurrence has not been established. To further evaluate this safety issue, FDA intends to analyze data obtained from several large, long-term, placebo-controlled trials of bisphosphonates in women at risk for osteoporosis-related fractures to assess risk of fractures based on use or nonuse of proton-pump inhibitors.

The mechanism by which proton-pump inhibitors may increase risk of fractures has not been elucidated, but it has been suggested that the mechanism may involve decreased insoluble calcium absorption secondary to increased gastric pH. Results of 3 observational studies showed no consistent relationship between proton-pump inhibitor use and bone mineral density. Additional studies evaluating effects of these drugs on bone homeostasis, including effects on biomarkers of bone formation and resorption, are ongoing.

■ Hypomagnesemia

Hypomagnesemia, symptomatic and asymptomatic, has been reported rarely in patients receiving long-term therapy (for at least 3 months or, in most cases, for longer than one year) with proton-pump inhibitors, including omeprazole. (See Cautions: Precautions and Contraindications.) On March 2, 2011, after reviewing reports of hypomagnesemia in patients receiving proton-pump inhibitors (i.e., 38 cases from the Adverse Event Reporting System [AERS] database, 23 cases from the medical literature [at least 8 of which have been identified in AERS]), FDA confirmed that long-term use (in most cases, longer than one year) of proton-pump inhibitors may be associated with an increased risk of hypomagnesemia; the incidence of this adverse effect could not be quantified because hypomagnesemia is likely underrecognized and underreported. The mechanism responsible for hypomagnesemia associated with long-term use of proton-pump inhibitors is unknown; however, long-term use of these agents may be associated with changes in intestinal absorption of magnesium.

Clinically serious adverse effects associated with hypomagnesemia, which are similar to manifestations of hypocalcemia, include tetany, seizures, tremors, carpopedal spasm, arrhythmias (e.g., atrial fibrillation, supraventricular tachycardia), and abnormal QT interval. Other reported adverse effects include paresthesia, muscle weakness, muscle cramps, lethargy, fatigue, and unsteadiness. Manifestations of hypomagnesemia secondary to proton-pump inhibitor therapy may not be present in all patients. Hypomagnesemia also produces impaired parathyroid hormone secretion, which may lead to hypocalcemia. In most patients, treatment of hypomagnesemia required magnesium replacement and discontinuance of the proton-pump inhibitor. Following discontinuance of the proton-pump inhibitor, hypomagnesemia resolved within a median of one week; upon rechallenge, hypomagnesemia recurred within a median of 2 weeks. In a few patients in whom reinitiation of proton-pump inhibitor therapy was necessary, use of pantoprazole, the least potent proton-pump inhibitor, in combination with oral magnesium supplements resulted in acceptable control of GI discomfort (e.g., dyspepsia, reflux symptoms) without causing recurrent hypomagnesemia; further study is needed to establish the role, if any, of pantoprazole in patients with proton-pump inhibitor-induced hypomagnesemia.

Because nonprescription (over-the-counter, OTC) proton-pump inhibitors are marketed at low dosages and are only intended for a 14-day course of treatment up to 3 times per year, FDA states that there is very little risk of hypomagnesemia when these preparations are used in accordance with the directions on the labeling.

■ Hepatic Effects

Mild and, rarely, marked increases in serum ALT (SGPT), AST (SGOT), γ-glutamyltransferase (GGT, γ-glutamyltranspeptidase, GGTP), alkaline phosphatase, and bilirubin concentrations have been reported during postmarketing surveillance in patients receiving omeprazole, but in many cases a causal relationship has not been established. In a controlled clinical trial in critically ill patients, abnormal liver function test results (not otherwise specified) occurred in 1.7 or 3.3% of patients receiving omeprazole immediate-release oral suspension or IV cimetidine, respectively. Rare occurrences of symptomatic liver disease have been reported, including hepatocellular, cholestatic, or mixed hepatitis, jaundice, liver necrosis, hepatic failure, and hepatic encephalopathy. Fatalities have been reported in some patients with hepatic failure or liver necrosis.

■ Dermatologic and Sensitivity Reactions

In a controlled clinical trial in critically ill patients, rash and decubitus ulcer occurred in 5.6 and 3.4%, respectively, of patients receiving omeprazole immediate-release oral suspension or in 6.1 and 2.8%, respectively, of those receiving IV cimetidine.

Rash has been reported during postmarketing surveillance in patients receiving omeprazole; severe generalized reactions such as toxic epidermal necrolysis (some fatal), Stevens-Johnson syndrome, erythema multiforme, exfoliative dermatitis, and lichenoid eruptions have been reported. Other adverse dermatologic effects reported during postmarketing surveillance include skin inflammation, urticaria, purpura and/or petechiae (some cases with rechallenge) angioedema, pruritus, photosensitivity, alopecia, dry skin, and hyperhidrosis. Hypersensitivity reactions, including anaphylaxis, anaphylactic shock, angioedema, bronchospasm, interstitial nephritis, and urticaria, have been reported with omeprazole therapy. In many cases, a causal relationship to omeprazole has not been established.

■ Hematologic Effects

Short-term use of omeprazole does not appear to be associated with substantial changes in hematologic parameters. However, in a controlled clinical trial of critically ill patients, thrombocytopenia, anemia, and aggravated anemia occurred in 10.1, 7.9 and 2.2%, respectively, of those receiving omeprazole immediate-release oral suspension, or in 6.1, 7.7, and 3.9%, respectively, of those receiving IV cimetidine. Agranulocytosis (occasionally fatal) has been reported rarely with omeprazole therapy, but a causal relationship to the drug is uncertain. Other adverse hematologic effects reported during postmarketing surveillance include pancytopenia, thrombocytopenia, neutropenia, leukopenia, anemia, and leukocytosis. Hemolytic anemia has been reported rarely in patients receiving omeprazole. In many cases, a causal relationship with omeprazole has not been established.

■ Genitourinary Effects

Acute interstitial nephritis (some cases with positive rechallenge), urinary tract infection, microscopic pyuria, urinary frequency, elevated serum creatinine concentration, proteinuria, hematuria, glycosuria, testicular pain, and gynecomastia have been reported during postmarketing surveillance in patients receiving omeprazole; in many cases a causal relationship to the drug has not been established.

Sexual disturbances (e.g., priapism) have been reported occasionally in patients receiving omeprazole.

■ Cardiovascular Effects

In a controlled clinical trial of critically ill patients, hypotension and hypertension occurred in 9.6 and 7.9%, respectively, of patients receiving omeprazole immediate-release oral suspension, or in 6.6 and 3.3%, respectively, of those receiving IV cimetidine. Atrial fibrillation, ventricular tachycardia, bradycardia, supraventricular tachycardia and tachycardia (not otherwise specified) occurred in 6.2, 4.5, 3.9, 3.4, and 3.4%, respectively, of patients receiving omeprazole or in 3.9, 3.3, 2.8, 1.1, and 3.3%, respectively, of patients receiving IV cimetidine.

Chest pain, angina pectoris, tachycardia, bradycardia, palpitation, elevated blood pressure, and peripheral edema have been reported during postmarketing surveillance in patients receiving omeprazole; a causal relationship to the drug has not been established in many cases.

Although preliminary safety data from 2 long-term clinical trials comparing omeprazole or esomeprazole with antireflux surgery in patients with severe gastroesophageal reflux disease (GERD) raised concerns about a potential increased risk of cardiac events (myocardial infarction, heart failure, and sudden death) in patients receiving these drugs, FDA has reviewed safety data from these and other studies of the drugs and has concluded that long-term use of omeprazole or esomeprazole is not likely to be associated with an increased risk of such cardiac events. FDA has concluded that the apparent increase in cardiac events observed in the early analyses is not a true effect of the drugs.

In one study (a 14-year study comparing omeprazole with antireflux surgery in 298 patients), death from cardiac causes (heart failure, sudden death) or nonfatal myocardial infarction occurred in 8 or 9 patients, respectively, randomized to receive omeprazole and in 2 or 2 patients, respectively, randomized to undergo surgery. However, the findings may have been biased by baseline differences between the 2 groups, since patients in the surgery group tended to be younger and healthier and were less likely to have a history of myocardial infarction than those receiving omeprazole. In addition, some patients withdrew from the study prior to undergoing surgery, and several underwent surgery and also received drug therapy. Fewer than half of the patients remained in the study until its completion. Preliminary data from the second study (an ongoing study comparing esomeprazole with antireflux surgery in 554 patients) also suggested a difference in occurrence of cardiac events between treatment groups; however, after 5 years of follow-up, a similar number of patients in each treatment group had experienced cardiac-related events. FDA reviewed safety data from 14 additional comparative studies of omeprazole (including 4 placebo-controlled studies) and indicated that these studies do not suggest that omeprazole is associated with an increased risk of cardiac events. None of the studies were designed to assess cardiac risk, and patient follow-up in the studies was incomplete.

■ Ocular Effects

Blurred vision, ocular irritation, dry eye syndrome, optic atrophy, anterior ischemic optic neuropathy, optic neuritis, and double vision have been reported during postmarketing surveillance in patients receiving omeprazole; in many cases a causal relationship to the drug has not been established.

■ Other Adverse Effects

In a controlled clinical trial of critically ill patients, hypophosphatemia, hypocalcemia, fluid overload, and hyponatremia occurred in 6.2, 6.2, 5.1, and 3.9%, respectively, of patients receiving omeprazole immediate-release oral suspension or in 3.9, 5.5, 7.7, and 2.8%, respectively, of patients receiving IV cimetidine. Hypoglycemia, hyperkalemia, and hypernatremia occurred in 3.4, 2.2, and 1.7%, respectively, of those receiving omeprazole or in 4.4, 3.3, and 5%, respectively, of patients receiving IV cimetidine. Hyperpyrexia and edema occurred in 4.5 and 2.8%, respectively, of patients receiving omeprazole or in 1.7 and 6.1%, respectively, of patients receiving IV cimetidine. Sepsis (not otherwise specified), oral candidiasis, urinary tract infection, and candidal infection (not otherwise specified) occurred in 5.1, 3.9, 2.2, and 1.7%, respectively, of patients receiving omeprazole or in 5, 0.6, 3.3, and 3.9%, respectively, of patients receiving IV cimetidine.

Hyponatremia, hypoglycemia, weight gain, fever, and tinnitus have been reported during postmarketing surveillance in patients receiving omeprazole, but in many cases were not attributed to the drug. Acute gout also has been reported during omeprazole therapy.

Accidental injury occurred in about 4% of pediatric patients 2–16 years of age receiving omeprazole in clinical studies.

Limited evidence suggests that omeprazole therapy may cause a dose-dependent reduction in cyanocobalamin absorption, although conflicting data also exist. In one study, absorption of protein-bound cyanocobalamin decreased from a median value of 2.2 or 2.3% at baseline to 0.8 or 0.5% in healthy men receiving 20 or 40 mg, respectively, of omeprazole daily for 2 weeks. (See Cautions: Precautions and Contraindications.)

■ Precautions and Contraindications

Symptomatic response to omeprazole should not be interpreted as precluding the presence of gastric malignancy.

While available endoscopic and histologic examinations of gastric biopsy specimens from humans exposed short-term to omeprazole have failed to reveal any associated risk, a dose-related increase in gastric carcinoid tumors has been observed during long-term exposure in animals, and further data from humans are needed to rule out the possibility of an increased risk of tumors during long-term exposure to the drug. (See Cautions: Mutagenicity and Carcinogenicity.)

Atrophic gastritis occasionally has been noted in gastric corpus biopsies from patients receiving long-term treatment with omeprazole.

The possibility that gastric acid-suppressive therapy may increase the risk of community-acquired pneumonia should be considered. (See Respiratory Effects: Community-acquired Pneumonia, in Cautions.)

Findings from several observational studies suggest that therapy with proton-pump inhibitors, particularly in high dosages and/or for prolonged periods of time, may be associated with an increased risk of osteoporosis-related fractures of the hip, wrist, or spine. (See Cautions: Musculoskeletal Effects.) Although controlled studies are required to confirm these findings, patients should receive proton-pump inhibitors at the lowest effective dosage and for the shortest possible time appropriate for their clinical condition. Individuals using omeprazole for *self-medication* should be advised that they should use the drug only as directed for no longer than 14 days of *continuous* use and that they should not exceed one course of therapy every 4 months. Patients who are at risk for osteoporosis-related fractures should receive an adequate intake of calcium and vitamin D and should have their bone health assessed and managed according to current standards of care.

Long-term use (in most cases, longer than one year) of proton-pump inhibitors may be associated with an increased risk of hypomagnesemia. In patients expected to receive long-term therapy with a proton-pump inhibitor or in those receiving a proton-pump inhibitor concomitantly with digoxin or drugs that may cause hypomagnesemia (e.g., diuretics), clinicians should consider measurement of serum magnesium concentrations prior to initiation of prescription proton-pump inhibitor therapy and periodically thereafter. Patients receiving proton-pump inhibitors should be advised to seek immediate care if manifestations of hypomagnesemia (e.g., arrhythmias, tetany, tremor, seizures) occur; in children, abnormal heart rates may cause fatigue, upset stomach, dizziness, and lightheadedness. Patients receiving nonprescription proton-pump inhibitors should be advised to follow the manufacturer's directions on the package carefully; if therapy with a nonprescription proton-pump inhibitor is continued for longer than the maximum recommended duration (an unlabeled [off-label] use), patients should be informed of the potential increased risk of hypomagnesemia.

Although preliminary safety data from 2 long-term clinical trials comparing omeprazole or esomeprazole with antireflux surgery in patients with severe gastroesophageal reflux disease (GERD) raised concerns about a potential increased risk of cardiac events (myocardial infarction, heart failure, and sudden death) in patients receiving these drugs, FDA has reviewed safety data from these and other studies of the drugs and has concluded that long-term use of omeprazole or esomeprazole is not likely to be associated with an increased risk of such cardiac events. FDA has concluded that the apparent increase in cardiac events observed in the early analyses is not a true effect of the drugs and recommends that clinicians continue to prescribe and patients continue to use these drugs in the manner described in the manufacturers' labelings. (See Cautions: Cardiovascular Effects.)

Because available data suggest a possible association between use of proton-pump inhibitors and risk of *Clostridium difficile* infection, patients should be advised to contact a clinician if persistent watery stools, abdominal pain, and fever occur.

Limited data suggest that omeprazole therapy may cause a dose-dependent reduction in cyanocobalamin absorption. (See Cautions: Other Adverse Effects.) Whether such a reduction in cyanocobalamin absorption can result in a deficiency of the vitamin has not been determined, although it has been suggested that pending further study, serum cyanocobalamin concentrations should be monitored in patients receiving long-term therapy with omeprazole.

Increases in intragastric pH may result in hypergastrinemia, enterochromaffin-like cell hyperplasia, and increased serum chromogranin A (CgA) concentrations. Increased CgA concentrations may produce false-positive results for diagnostic tests for neuroendocrine tumors. Clinicians should temporarily discontinue omeprazole therapy before assessing CgA concentrations and consider repeating the test if initial CgA concentrations are high.

Each 20- or 40-mg packet of omeprazole powder for immediate-release oral suspension (Zegerid®) contains 1680 mg of sodium bicarbonate (460 mg [20 mEq] of sodium). Each 20- or 40-mg immediate-release capsule of omeprazole (Zegerid®) contains 1100 mg of sodium bicarbonate (304 mg [13 mEq] of sodium). The sodium content of these preparations should be taken into consideration in patients whose sodium intake must be restricted; increased sodium intake may produce edema and weight increase. Sodium bicarbonate may cause metabolic alkalosis, seizures, and tetany, and chronic use with calcium or milk may cause milk-alkali syndrome. Acute toxicity associated with sodium bicarbonate overdose may include hypocalcemia, hypokalemia, hypernatremia, and seizures. Sodium bicarbonate should be used with caution in patients with Bartter's syndrome, hypokalemia, respiratory alkalosis, or acid-base abnormalities. Sodium bicarbonate is contraindicated in patients with metabolic alkalosis or hypocalcemia.

In clinical trials in patients who received combined clarithromycin-omeprazole therapy for *H. pylori* infection, some *H. pylori* isolates demonstrated an increase in clarithromycin MICs over time, indicating decreased susceptibility and increasing resistance to the drug. Susceptibility testing should be performed if possible in patients with *H. pylori* infection in whom therapy with combined clarithromycin-omeprazole fails (i.e., as determined in clinical trials by a positive result for *H. pylori* on culture or histologic testing 4 weeks following completion of therapy); if resistance to clarithromycin is demonstrated or susceptibility testing is not possible, alternative anti-infective therapy should be instituted. The American College of Gastroenterology (ACG) states that clarithromycin or metronidazole should not be used subsequently in patients with *H. pylori* infection who fail therapy that includes these drugs since resistance consistently emerges during such unsuccessful therapy. (See Uses: *Helicobacter pylori* Infection, in Clarithromycin 8:12.12.92.)

Patients should be advised to consult their clinician before using omeprazole for *self-medication* if they are taking warfarin, an antifungal agent (e.g., ketoconazole), diazepam, or digoxin. Patients with heartburn that has persisted for more than 3 months or heartburn in conjunction with lightheadedness, sweating, or dizziness should consult their clinician before using omeprazole for *self-medication*. Patients should be advised to consult their clinician before using omeprazole for *self-medication* if they are experiencing chest or shoulder pain with lightheadedness, shortness of breath, sweating, or pain spreading to arms, neck, or shoulders. Those with frequent chest pain, unexplained weight loss, nausea and vomiting, stomach pain, or frequent wheezing (especially with heartburn) also should consult their clinician before using omeprazole for *self-medication*. Patients should discontinue taking omeprazole for *self-medication* and consult their clinician if heartburn persists, or worsens after 14 days of therapy, or a course of treatment is needed more frequently than every 4 months. Patients with difficulty or pain with swallowing, vomiting with blood, or bloody or blackened stools should not use omeprazole for *self-medication*; such manifestations should be reported promptly to a clinician, since they may be indicative of a serious condition requiring alternative treatment. Women who are pregnant or breast feeding should consult their clinician before using omeprazole for *self-medication*.

Omeprazole is contraindicated in patients with known hypersensitivity to the drug, esomeprazole, or other substituted benzimidazoles (e.g., lansoprazole, pantoprazole, rabeprazole), or any ingredient in the formulation.

■ Pediatric Precautions

Safety and efficacy of omeprazole (omeprazole delayed-release capsules and omeprazole magnesium delayed-release oral suspension) have been established for the treatment of symptomatic gastroesophageal reflux disease (GERD), erosive esophagitis, and maintenance of healing of erosive esophagitis in pediatric patients 1–16 years of age. Use in pediatric patients is supported by adequate and well-controlled studies in adults and additional pharmacokinetic data and clinical and safety studies in children. (See Gastroesophageal Reflux: Clinical Trials and see Gastroesophageal Reflux: Maintenance Therapy, in Uses.)

Safety and efficacy of omeprazole in pediatric patients younger than 1 year of age have not been established. Safety and efficacy of omeprazole for *self-medication* in those younger than 18 years of age have not been established. Safety and efficacy of omeprazole immediate-release capsules and oral suspension (Zegerid®) have not been established in pediatric patients younger than 18 years of age.

■ Geriatric Precautions

In US and European clinical trials, more than 2000 patients treated with omeprazole were 65 years of age or older. Although no overall differences in efficacy or safety were observed between geriatric and younger patients, and other clinical experience revealed no evidence of age-related differences, the possibility that some older patients may exhibit increased sensitivity to the drug cannot be ruled out.

Although elimination of omeprazole may be somewhat delayed and oral bioavailability increased in the elderly, clinically important differences in the pharmacokinetic profile of omeprazole between geriatric individuals and younger adults generally do not appear to exist. Therefore, dosage adjustment solely on the basis of age generally is not required for geriatric patients.

The adverse effect profile of omeprazole is similar in geriatric patients and those 65 years of age and younger.

■ Mutagenicity and Carcinogenicity

No evidence of mutagenicity was observed in vivo in the rat liver DNA damage assay or in some in vitro test systems, including the microbial (Ames test) and mammalian (mouse lymphoma) assays. Omeprazole was positive for clastogenic effects in an in vitro human lymphocyte chromosome aberration assay, in 1 of 2 in vivo mouse micronucleus tests, and in an in vivo bone marrow cell chromosomal aberration assay.

In animals, long-term administration of relatively high dosages of omeprazole results in morphologic changes in the gastric mucosa. Such changes observed in rats during long-term (24-month) administration of the drug include dose-related increases in gastric carcinoid tumors and enterochromaffin-like (ECL) cell hyperplasia, which are thought to represent exaggerated physiologic responses occurring secondary to profound inhibition of gastric acid secretion and subsequent hypergastrinemia and reversible hypertrophy of oxyntic mucosa. While such changes have not been observed following short-term administration of the drug in humans, additional long-term data are needed to rule out the possibility of an increased risk of gastric tumors in patients receiving long-term omeprazole therapy. In two 24-month studies in rats given omeprazole dosages of 1.7, 3.4, 13.8, 44, and 140.8 mg/kg daily (about 0.7–57 times the human dosage of 20 mg daily based on body surface area), the drug caused a dose-related increase in gastric ECL cell carcinoids in both male and female rats; the increase in carcinoids occurred more frequently in female rats. In addition, ECL cell hyperplasia was observed in both male and female rats receiving omeprazole. In female rats given omeprazole dosages of 13.8 mg/kg daily (about 6 times the human dosage of 20 mg daily based on body surface area) for 1 year and then observed for another year without the drug, carcinoids were not detected but ECL hyperplasia occurred in 94% of rats given omeprazole versus 10% of controls at the end of 1 year; at the end of the second year, hyperplasia was observed in 46% of rats given omeprazole versus 26% of controls. Gastric adenocarcinoma was reported in one rat; similar tumors were not seen in male or female rats treated for 2 years. For this strain of rat no similar tumor had been noted historically, but the finding of this tumor in only one rat is difficult to interpret. In a 1-year toxicity study in Sprague-Dawley rats, brain astrocytomas were found in a small number of males (but not in females) given omeprazole at dosage levels of 0.4, 2, and 16 mg/kg daily (about 0.2–6.5 times the human dosage of 20 mg daily based on body surface area). In a 2-year carcinogenicity study in Sprague-Dawley rats, no astrocytomas were found in males or females at 140.8 mg/kg daily (about 57 times the human dosage of 20 mg daily based on body surface area). Long-term carcinogenicity studies (78 weeks) in mice did not demonstrate increased tumor occurrence; however, the manufacturer states that the study was inconclusive. The drug was not carcinogenic in a 26-week p53± transgenic mouse study.

A number of patients with Zollinger-Ellison syndrome receiving long-term therapy with omeprazole have developed gastric carcinoids; however, Zollinger-Ellison syndrome is known to be associated with such tumors, and these findings are believed to be related to the underlying disease rather than to omeprazole therapy. Gastric corpus biopsy specimens obtained from more than 3000 patients in long-term studies with omeprazole have demonstrated an increased frequency of ECL cell hyperplasia (including micronodular hyperplasia of argyrophil cells) in association with increased plasma gastrin concentrations and progression to subatrophic or atrophic gastritis. However, no evidence of ECL cell carcinoids, dysplasia, or neoplasia has been observed in these patients, and it has been suggested that the development of mucosal cell hyperplasia may be related to the severity and natural progression of gastritis rather than to hypergastrinemia.

■ Pregnancy, Fertility, and Lactation

Pregnancy

Omeprazole crosses the placenta in animals and in humans. Reproductive studies in rats or rabbits using omeprazole dosages up to 138 or 69 mg/kg daily (about 56 times the human dosage of 20 mg daily based on body surface area), respectively, have not revealed evidence of teratogenicity. However, in rabbits given omeprazole dosages of 6.9–69.1 mg/kg daily (about 5.6–56 times the human dosage of 20 mg daily based on body surface area), dose-related increases in embryolethality, fetal resorptions, and pregnancy loss occurred. In rats, dose-related embryo/fetal toxicity and postnatal developmental toxicity were observed in offspring resulting from administration of omeprazole in dosages of 13.8–138 mg/kg daily (about 5.6–56 times the human dosage of 20 mg daily based on body surface area) to parents.

There are no adequate and controlled studies using omeprazole in pregnant women. Most reported experience with omeprazole during human pregnancy has been first trimester exposure; duration of use (i.e., intermittent, long-term) rarely has been specified. A review of published data (considered fair in quality and quantity) by the Teratogen Information System (TERIS) concluded that therapeutic dosages of omeprazole during pregnancy are unlikely to pose a substantial teratogenic risk. Data from cohort studies have not demonstrated that omeprazole exposure is associated with a statistically significant increase in the rate of major birth defects; however, the studies lacked the power to detect small increases in birth defects or in rare malformations. Therefore, additional study is needed.

A population-based retrospective cohort epidemiologic study using data from the Swedish Medical Birth Registry reported on outcomes in infants whose mothers used omeprazole during pregnancy; most (about 86%) were exposed to omeprazole during the first trimester, 4% during and beyond the first trimester, and about 10% were exposed only after the first trimester of pregnancy. Exposure to omeprazole was not associated with increased risk of any malformation (odds ratio 0.82 and 95% confidence interval of 0.50–1.34), low birth weight, or low Apgar score. The number of infants born with ventricular septal defects and the number of stillborn infants was slightly higher in the omeprazole-exposed infants than the expected number in the normal population, but both effects may be random. In an earlier study using data from the Swedish Medical Birth Registry, exposure to proton-pump inhibitors was not associated with increased risk of congenital malformation.

In a retrospective cohort study, the incidence of congenital malformations in women who received omeprazole or histamine H_2-antagonists (cimetidine or ranitidine) in the first trimester of pregnancy was compared with a control

group of women who were not exposed to acid-suppressant drugs. The overall malformation rate was 4.4% (95% confidence interval of 3.6–5.3), the malformation rate for nonexposed women was 3.8% (95% confidence interval of 3–4.9), and the malformation rate associated with omeprazole exposure was 3.6% (95% confidence interval of 1.5–8.1). The relative risk of malformations associated with first-trimester exposure to omeprazole (compared with nonexposed women) was 0.9 (95% confidence interval of 0.3–2.2). The study could effectively rule out a relative risk greater than 2.5 for all malformations. Rates of preterm delivery or growth retardation did not differ between the groups.

A controlled prospective observational study followed women exposed to omeprazole, disease-paired controls exposed to histamine H_2-receptor antagonists, and controls exposed to nonteratogenic agents (e.g., acetaminophen, dental radiation) during pregnancy; major congenital malformations occurred in 4, 2.8, and 2%, respectively, of live births, or in 5.1%, 3.1%, and 3%, respectively, of live births when exposure occurred during the first trimester of pregnancy. Rates of spontaneous and elective abortions, preterm deliveries, gestational age at delivery, and mean birth weight did not differ between the groups. The study lacked statistical power to detect a small increase in major malformations; the sample size had 80% power to detect a fivefold increase in the major malformation rate.

The manufacturers state that several studies reported that adverse short-term effects were not observed in infants when a single oral or IV dose of omeprazole was administered to pregnant women as premedication for cesarean section under general anesthesia.

Omeprazole immediate-release capsules and oral suspension contain sodium bicarbonate; chronic use of sodium bicarbonate may lead to systemic alkalosis, and increased sodium intake may produce edema and weight increase.

Because there are no adequate and controlled studies using omeprazole in pregnant women, and because studies to date in animals and pregnant women cannot rule out the possibility of harm, the drug should be used during pregnancy only when clearly needed and only when the potential benefits justify the possible risk to the fetus.

Fertility

Reproductive studies in rats using omeprazole dosages of up to 138 mg/kg daily (about 56 times the human dose of 20 mg daily based on body surface area) have not revealed evidence of impaired fertility.

Lactation

Omeprazole is distributed into human milk; following oral administration of omeprazole 20 mg in one lactating woman, the peak concentration of the drug in breast milk was less than 7% of the peak serum concentration. Because of the potential for serious adverse reactions to omeprazole in nursing infants, and because of the potential for tumorigenicity shown in animal studies, a decision should be made whether to discontinue nursing or the drug, taking account the importance of the drug to the woman. In addition, omeprazole immediate-release capsules and oral suspension contain sodium bicarbonate, which should be used with caution in nursing mothers.

Drug Interactions

■ Drugs Affecting or Metabolized by Hepatic Microsomal Enzymes

Omeprazole can prolong the elimination of diazepam, warfarin (the *R*-isomer), phenytoin, cyclosporine, disulfiram, and benzodiazepines, and the possibility that dosages of these and other drugs that are metabolized by cytochrome P-450 (CYP)-mediated oxidation in the liver may require adjustment should be considered when concomitant omeprazole therapy is initiated or discontinued.

Omeprazole is extensively metabolized by the CYP enzyme system. Drugs that induce CYP2C19 or CYP3A4 (e.g., rifampin, St. John's wort [*Hypericum perforatum*]) can substantially decrease omeprazole concentrations. In a crossover study, concomitant use of omeprazole and St. John's wort (300 mg 3 times daily for 14 days) in healthy men resulted in decreased exposure to omeprazole; peak plasma concentrations and area under the concentration-time curve (AUC) of omeprazole were decreased by 37.5 and 37.9%, respectively, in poor CYP2C19 metabolizers and by 49.6 and 43.9%, respectively, in extensive metabolizers. Concomitant use of omeprazole with St. John's wort or rifampin should be avoided.

■ Drugs that Cause Hypomagnesemia

In patients receiving diuretics (i.e., loop or thiazide diuretics) or other drugs that may cause hypomagnesemia, monitoring of magnesium concentrations should be considered prior to initiation of prescription proton-pump inhibitor therapy and periodically thereafter. (See Cautions: Hypomagnesemia.)

■ Gastric pH-dependent Drugs

The possibility that omeprazole-induced increases in gastric pH may affect the bioavailability of drugs such as ketoconazole, ampicillin esters, iron salts, erlotinib, or digoxin (where gastric acidity is an important determinant in oral absorption) should be considered. Concomitant use of omeprazole 20 mg once daily and digoxin in healthy individuals increased digoxin bioavailability by 10% (up to 30% in some individuals); monitoring for manifestations of digoxin toxicity may be required when omeprazole is used concomitantly.

■ Antiretroviral Agents

Atazanavir

Concomitant use of omeprazole 40 mg once daily and atazanavir (with or without low-dose ritonavir) results in a substantial decrease in plasma concentrations of atazanavir and possible loss of the therapeutic effect of the antiretroviral agent. The manufacturers of omeprazole state that concomitant use of omeprazole with atazanavir is not recommended. If atazanavir is administered in an antiretroviral treatment-naive patient receiving a proton-pump inhibitor, a *ritonavir-boosted* regimen of 300 mg of atazanavir once daily with ritonavir 100 mg once daily with food is recommended. The dose of the proton-pump inhibitor should be administered approximately 12 hours before *ritonavir-boosted* atazanavir; the dose of the proton-pump inhibitor should not exceed omeprazole 20 mg daily (or equivalent). Concomitant use of proton-pump inhibitors with atazanavir is not recommended in antiretroviral treatment-experienced patients.

Darunavir

Concomitant use of omeprazole with *ritonavir-boosted* darunavir resulted in decreased plasma concentrations of omeprazole; plasma concentrations of darunavir were not affected. However, no dosage adjustment of either drug is required when omeprazole is used concomitantly with *ritonavir-boosted* darunavir.

Fosamprenavir

Concomitant use of esomeprazole and fosamprenavir increased the AUC of esomeprazole by about 55% but did not substantially alter plasma concentrations of amprenavir (active metabolite of fosamprenavir). Concomitant use of esomeprazole and *ritonavir-boosted* fosamprenavir did not substantially affect concentrations of either amprenavir or esomeprazole. No dosage adjustment is required when proton-pump inhibitors are used concomitantly with fosamprenavir (with or without ritonavir).

Lopinavir

Concomitant use of omeprazole with the fixed combination of lopinavir and ritonavir (lopinavir/ritonavir) did not have a clinically important effect on plasma concentrations or AUC of lopinavir. No dosage adjustment is required when proton-pump inhibitors are used concomitantly with lopinavir/ritonavir.

Nelfinavir

Concomitant use of omeprazole 40 mg once daily (given 30 minutes before a nelfinavir dose) and nelfinavir 1.25 g twice daily decreased peak plasma concentrations and AUCs of nelfinavir by 37 and 36%, respectively, and of its major active metabolite M8 by 89 and 92%, respectively. The manufacturers of omeprazole state that concomitant use with nelfinavir is not recommended.

Raltegravir

Concomitant use of omeprazole with raltegravir resulted in substantially increased peak plasma concentration and AUC of raltegravir; however, no dosage adjustment is recommended when proton-pump inhibitors are used concomitantly with raltegravir.

Rilpivirine

Concomitant use of omeprazole with rilpivirine resulted in decreased plasma concentrations and AUC of rilpivirine. Concomitant use of rilpivirine and proton-pump inhibitors is contraindicated.

Saquinavir

Concomitant use of omeprazole 40 mg once daily and saquinavir 1 g twice daily (with ritonavir 100 mg twice daily) increased peak serum concentrations and AUC of saquinavir by 75 and 82%, respectively. If omeprazole is used concomitantly with *ritonavir-boosted* saquinavir, clinical and laboratory monitoring for saquinavir toxicity is recommended, and reduction of saquinavir dosage should be considered.

Tipranavir

Concomitant use of omeprazole with *ritonavir-boosted* tipranavir resulted in decreased plasma concentrations of omeprazole; plasma concentrations of tipranavir were not affected. An increase in omeprazole dosage may be required when omeprazole is used concomitantly with *ritonavir-boosted* tipranavir.

■ Cilostazol

Concomitant use of omeprazole (40 mg daily for one week) and cilostazol in healthy individuals resulted in increased peak plasma concentrations and AUC of cilostazol (by 18 and 16%, respectively) and one of its active metabolites (by 29 and 69%, respectively). Therefore, reduction of cilostazol dosage (from 100 mg twice daily to 50 mg twice daily) should be considered during such concomitant use.

■ Clarithromycin

Concomitant administration of clarithromycin (500 mg 3 times daily) and omeprazole (40 mg daily) in healthy men resulted in increases of 30, 89, and 34% in the peak plasma concentration, AUC, and elimination half-life, respectively, of omeprazole. Increases in omeprazole AUC and half-life had a modest effect on gastric pH; mean 24-hour gastric pH was 5.2 when omeprazole was administered alone versus 5.7 with concomitant administration of clarithromycin. Acid suppression resulting from omeprazole appears to enhance the activity of anti-infective therapy against *H. pylori*. Serum concentrations and AUCs of clarithromycin and 14-hydroxyclarithromycin also are increased by concomitant administration of omeprazole. (See Pharmacokinetics: Absorption, in Clarithromycin 8:12.12.92.)

■ Clopidogrel

Because omeprazole inhibits CYP2C19, concurrent use of omeprazole with clopidogrel, which is metabolized to its active metabolite by CYP2C19, reduces plasma concentrations of clopidogrel's active metabolite and potentially may reduce clopidogrel's clinical efficacy.

In a crossover clinical trial in healthy individuals who received clopidogrel (a 300-mg loading dose, followed by 75 mg daily) alone or with omeprazole (80 mg administered at the same time as the clopidogrel dose) for 5 days, exposure to the active metabolite of clopidogrel was decreased by 46% and 42% on days 1 and 5, respectively, when the drugs were administered simultaneously. In addition, mean inhibition of platelet aggregation was reduced by 47% at 24 hours and by 30% on day 5. When administration of the 2 drugs (at the same dosages) was separated by 12 hours in another study, results were similar. Concomitant use of clopidogrel and omeprazole also has been associated with decreased antiplatelet effects as determined by vasodilator-stimulated phosphoprotein (VASP) phosphorylation.

The clinical importance of these effects has not yet been established but reduction in clopidogrel's effectiveness in preventing cardiovascular events is possible. Several observational studies involving large numbers of patients suggest that proton-pump inhibitors reduce the effectiveness of clopidogrel in preventing cardiovascular events (e.g., recurrent myocardial infarction, rehospitalization for acute coronary syndromes, urgent target vessel revascularization, death). However, data discounting the clinical importance of an interaction between clopidogrel and proton-pump inhibitors also have been reported. Some experts, including the American College of Cardiology (ACC) and the American Heart Association (AHA), state that additional data from large, prospective trials are needed to fully elucidate the clinical consequences, if any, of the observed interaction between clopidogrel and certain proton-pump inhibitors, including omeprazole.

Proton-pump inhibitors vary in their potency for inhibiting CYP2C19. The change in inhibition of adenosine diphosphate (ADP)-induced platelet aggregation associated with concomitant use of proton-pump inhibitors with clopidogrel is related to the change in exposure to the active metabolite of clopidogrel. In pharmacokinetic and pharmacodynamic studies in healthy individuals, concomitant use of dexlansoprazole, lansoprazole, or pantoprazole had less effect on the antiplatelet activity of clopidogrel than did concomitant use of omeprazole or esomeprazole. For further information on interactions between proton-pump inhibitors and clopidogrel, see Drug Interactions: Proton-Pump Inhibitors, in Clopidogrel Bisulfate 20:12.18. Concomitant use of clopidogrel and either omeprazole or esomeprazole should be avoided. Administration of the drugs at separate times will not prevent the interaction.

The decision to use a proton-pump inhibitor concomitantly with clopidogrel should be based on the assessed risks and benefits in individual patients. The American College of Cardiology Foundation/American College of Gastroenterology/American Heart Association (ACCF/ACG/AHA) states that the reduction in GI bleeding risk with proton-pump inhibitors is substantial in patients with risk factors for GI bleeding (e.g., advanced age; concomitant use of warfarin, corticosteroids, or nonsteroidal anti-inflammatory agents [NSAIAs]; *H. pylori* infection) and may outweigh any potential reduction in the cardiovascular efficacy of antiplatelet treatment associated with a drug-drug interaction. In contrast, ACCF/ACG/AHA states that patients without such risk factors receive little if any absolute risk reduction from proton-pump inhibitor therapy, and the risk/benefit balance may favor use of antiplatelet therapy without a proton-pump inhibitor in these patients. (See Drug Interactions: Proton Pump Inhibitors, in Clopidogrel Bisulfate 20:12.18.)

If concomitant proton-pump inhibitor use is considered necessary, use of an agent with little or no CYP2C19-inhibitory activity should be considered. Alternatively, treatment with a histamine H_2-receptor antagonist (ranitidine, famotidine, nizatidine) may be considered, although such agents may not be as effective as a proton-pump inhibitor in providing gastric protection; cimetidine should *not* be used since it also is a potent CYP2C19 inhibitor. There currently is no evidence that histamine H_2-receptor antagonists (other than cimetidine) or other drugs that reduce gastric acid (e.g., antacids) interfere with the antiplatelet effects of clopidogrel.

■ Digoxin

Hypomagnesemia (e.g., resulting from long-term use of proton-pump inhibitors) sensitizes the myocardium to digoxin and, thus, may increase the risk of digoxin-induced cardiotoxic effects. In patients receiving digoxin, monitoring of magnesium concentrations should be considered prior to initiation of prescription proton-pump inhibitor therapy and periodically thereafter.

■ Methotrexate

Data from case reports, population pharmacokinetic studies, and retrospective analyses suggest that concomitant use of proton-pump inhibitors, including omeprazole, with methotrexate may result in increased serum methotrexate concentrations and possibly result in toxicity. Increased serum concentrations and delayed clearance of methotrexate and/or its metabolite hydroxymethotrexate, with or without symptoms of methotrexate toxicity, have been reported in patients receiving methotrexate (usually at doses of 300 mg/m^2 to 12 g/m^2) concomitantly with a proton-pump inhibitor. Although most of the reported cases occurred in patients receiving high doses of methotrexate, toxicity also has been reported in patients receiving low dosages of methotrexate (e.g., 15 mg per week) concomitantly with a proton-pump inhibitor. No formal studies of interactions between high-dose methotrexate and proton-pump inhibitors have been conducted to date.

The manufacturer of omeprazole states that temporary discontinuance of proton-pump inhibitor therapy may be considered in some patients receiving high-dose methotrexate therapy. Some clinicians recommend either withholding proton-pump inhibitor therapy for several days before and after methotrexate administration or substituting a histamine H_2-receptor antagonist for the proton-pump inhibitor when acid suppressive therapy is indicated during methotrexate therapy. Pending further evaluation, some clinicians state that these recommendations should extend to patients receiving low-dose methotrexate.

■ Tacrolimus

Concomitant use of omeprazole with tacrolimus may result in increased serum concentrations of tacrolimus.

■ Voriconazole

Concomitant use of omeprazole (40 mg daily for 7 days) with voriconazole (a combined inhibitor of CYP2C19 and CYP3A4; 400 mg every 12 hours for one day, then 200 mg for 6 days) in healthy individuals increased peak plasma concentrations and AUC of omeprazole by an average of twofold and fourfold, respectively. Dosage adjustment of omeprazole usually is not required but may be considered in patients receiving high dosages (up to 240 mg daily), such as those with Zollinger-Ellison syndrome.

■ Warfarin

Increases in international normalized ratio (INR) and prothrombin time have been reported in patients receiving warfarin concomitantly with a proton-pump inhibitor, including omeprazole. Because such increases may lead to abnormal bleeding and death, monitoring of INR and prothrombin time may be necessary in patients receiving warfarin and a proton-pump inhibitor concomitantly.

Description

Omeprazole is a substituted benzimidazole gastric antisecretory agent. Omeprazole is structurally and pharmacologically related to esomeprazole, lansoprazole, pantoprazole, and rabeprazole. Omeprazole is a racemic mixture of *R*- and *S*-isomers; esomeprazole is the *S*-isomer of omeprazole. The drugs are chemically and pharmacologically unrelated to H_2-receptor antagonists, antimuscarinics, or prostaglandin analogs.

Because the omeprazole molecule is acid labile, the drug is administered orally as a delayed-release capsule, tablet, or oral suspension or as a buffered immediate-release capsule or oral suspension. The commercially available omeprazole delayed-release capsules increase oral bioavailability by delaying absorption until after the capsule leaves the stomach; peak plasma concentrations of omeprazole occur 30 minutes to 3.5 hours after administration. Following administration of omeprazole magnesium delayed-release oral suspension, the peak plasma concentration and area under the concentration-time curve (AUC) of the drug are 88 and 87%, respectively, of the values achieved following oral administration of omeprazole delayed-release capsules. Omeprazole immediate-release capsules and oral suspension are rapidly absorbed immediate-release formulations that contain sodium bicarbonate to neutralize gastric acid; mean peak plasma concentrations of omeprazole occur at about 30 minutes (range 10–90 minutes) after oral administration of a single dose or repeated doses of the immediate-release capsule or oral suspension on an empty stomach (1 hour prior to a meal).

Omeprazole binds to hydrogen/potassium adenosine triphosphatase (H^+K^+-exchanging ATPase) in gastric parietal cells; inactivation of this enzyme system (also known as the proton, hydrogen, or acid pump) blocks the final step in the secretion of hydrochloric acid by these cells. Therefore, gastric antisecretory agents such as omeprazole and lansoprazole are commonly referred to as acid- or proton-pump inhibitors. Omeprazole, a weak base, does not directly inhibit this enzyme system, but instead, it concentrates under the acid conditions of the parietal cell secretory canaliculi, where the drug undergoes rearrangement to its active sulfenamide metabolite; this metabolite then reacts with sulfhydryl groups of H^+K^+-exchanging ATPase, inactivating the proton pump. Because the sulfenamide metabolite forms an irreversible covalent bond to H^+K^+-exchanging ATPase, acid secretion is inhibited until additional enzyme is synthesized, resulting in a prolonged duration of action. In an animal model, the pharmacologic effect of the drug at this enzyme was shown to correlate directly with sulfenamide formation.

Omeprazole inhibits basal and stimulated gastric acid secretion; in addition, because the drug inhibits the final step in the secretory pathway, it inhibits such secretion regardless of the stimulus. The degree of inhibition of gastric acid secretion is related to the dose and duration of therapy, but omeprazole is a more potent inhibitor of such secretion than are H_2-receptor antagonists. Following oral administration of omeprazole, inhibition of gastric acid secretion is apparent within 1 hour, peaks within 2 hours, and persists for up to 72 hours. Inhibition of gastric acid secretion increases with continuous drug administration and reaches a plateau after about 4 days of omeprazole therapy. Although omeprazole has a short terminal plasma half-life,

the drug has a long duration of action (presumably secondary to prolonged binding of the drug to H^+K^+-exchanging ATPase). Following continuous oral administration of omeprazole, 78 and 58–80% of basal gastric acid secretion is inhibited 2–6 and 24 hours, respectively, after a 20-mg dose, and 94 and 80–93% of basal gastric acid secretion is inhibited, respectively, after a 40-mg dose. Following continuous oral administration, 79 and 50–59% of peak gastric acid output is inhibited 2–6 and 24 hours, respectively, after a 20-mg dose, and 88 and 62–68% of peak gastric acid output is inhibited, respectively, after a 40-mg dose. Oral administration of omeprazole 10–40 mg daily has reduced 24-hour intragastric acidity by 100% in some patients. Following discontinuance of omeprazole therapy, gastric acid secretion returns to baseline over a 3–5 day period.

Following oral administration of omeprazole 20 or 40 mg daily (as the immediate-release oral suspension) in healthy individuals, the median decrease in 24-hour integrated gastric acidity from baseline was 82 or 84%, respectively, the percent of time during a 24-hour period that the gastric pH exceeded 4 was 51 or 77%, respectively, and the median 24-hour gastric pH was 4.2 or 5.2, respectively. In critically ill patients receiving omeprazole 40 mg daily as the immediate-release oral suspension via nasogastric or orogastric tube, the median daily gastric pH was above 4 in at least 95% of patients over the course of a 14-day trial. Gastric pH was above 4 in 99% of patients 1–2.5 hours after the first dose, and in 92% of patients 6 hours after the first dose.

Omeprazole increases plasma gastrin concentrations; this increase occurs via a negative feedback mechanism resulting from decreased gastric acid secretion. Because of omeprazole's greater potency as an inhibitor of gastric acid secretion, the drug also causes secondary increases in plasma gastrin concentrations that exceed those produced by H_2-receptor antagonists. For example, administration of omeprazole 20 mg once daily for 1–2 weeks results in a 1.3- to 3.6-fold increase in plasma gastrin concentration, whereas administration of an H_2-receptor antagonist usually results in only a 1.1- to 1.8-fold increase. Plasma gastrin concentrations return to pretreatment values with 1–4 weeks after discontinuing omeprazole therapy. Despite omeprazole-induced reductions in gastric acid secretion, the drug does not contribute appreciably to increased plasma gastrin concentrations in most patients with Zollinger-Ellison syndrome, since gastrin is produced principally by the tumor rather than in response to achlorhydria in such patients. Omeprazole also indirectly causes a dose-dependent reduction in pepsin secretion by decreasing the volume of gastric acid secretion. A systematic dose-dependent effect on basal or stimulated pepsin secretion has not been observed in humans; basal pepsin output is low and pepsin activity is decreased when intragastric pH is maintained above 4. The drug does not appear to affect intrinsic factor secretion.

Omeprazole can suppress gastric *Helicobacter pylori* (formerly *Campylobacter pylori* or *C. pyloridis*) in patients with duodenal ulcer and/or reflux esophagitis infected with the organism. Combined therapy with omeprazole and one or more appropriate anti-infectives (e.g., clarithromycin, amoxicillin), can effectively eradicate *H. pylori* gastric infection. (See Duodenal Ulcer: Acute Therapy, in Uses.) Omeprazole does not appear to affect gastric emptying or lower esophageal sphincter pressure.

Short-term administration (2–4 weeks) of omeprazole dosages of 30–40 mg daily does not appear to affect thyroid function, carbohydrate metabolism, or plasma/serum concentrations of parathyroid hormone (parathormone), cortisol, estradiol, testosterone, prolactin, cholecystokinin, or secretin. However, the drug may decrease antral somatostatin concentrations.

Therapy with proton-pump inhibitors, particularly in high dosages and/or for prolonged periods of time, may be associated with an increased risk of osteoporosis-related fractures of the hip, wrist, or spine. (See Cautions: Musculoskeletal Effects.) The mechanism by which these drugs may increase risk of such fractures has not been elucidated but may involve decreased insoluble calcium absorption secondary to increased gastric pH.

SumMon® (see Users Guide). For additional information on this drug until a more detailed monograph is developed and published, the manufacturer's labeling should be consulted. It is *essential* that the labeling be consulted for detailed information on the usual cautions, precautions, and contraindications concerning potential drug interactions and/or laboratory test interferences and for information on acute toxicity.

Preparations

Excipients in commercially available drug preparations may have clinically important effects in some individuals; consult specific product labeling for details.

Omeprazole

Oral

Capsules

20 mg	**Zegerid®**, Santarus

Omeprazole

40 mg	
	Zegerid®, Santarus
Capsules, delayed-release (containing enteric- coated granules)	
10 mg*	
	Omeprazole Delayed-release Capsules
	Prilosec®, AstraZeneca
20 mg*	
	Omeprazole Delayed-release Capsules
	Prilosec®, AstraZeneca
40 mg*	
	Omeprazole Delayed-release Capsules
	Prilosec®, AstraZeneca
For suspension, powder	
20 mg/packet	
	Zegerid®, Santarus
40 mg/packet	
	Zegerid®, Santarus

Omeprazole Magnesium

Oral

For suspension, delayed-release (containing enteric-coated granules)	
2.5 mg (of omeprazole) per packet	
	Prilosec®, AstraZeneca
10 mg (of omeprazole) per packet	
	Prilosec®, AstraZeneca
Tablets, delayed-release	
20 mg (of omeprazole)	
	Prilosec® OTC, Procter & Gamble

**available from one or more manufacturer, distributor, and/or repackager by generic (nonproprietary) name*

† Use is not currently included in the labeling approved by the US Food and Drug Administration.

Selected Revisions December 8, 2012, © Copyright, January 1, 1994, American Society of Health-System Pharmacists, Inc.

ประวัติผู้วิจัย

ชื่อ-สกุล	ชนิสรา ลือวิพันธ์
ที่อยู่	49 ถ.ประชาอุทิศ เขตราษฎร์บูรณะ กรุงเทพมหานคร 10140
ที่ทำงาน	ภาควิชาสารสนเทศศาสตร์ทางสุขภาพ คณะเภสัชศาสตร์ มหาวิทยาลัยศิลปากร
ประวัติการศึกษา	พ.ศ. 2550 สำเร็จการศึกษาปริญญาเภสัชศาสตรบัณฑิต จาก มหาวิทยาลัยศิลปากร
	พ.ศ. 2552 ศึกษาต่อระดับปริญญามหาบัณฑิต สาขาสารสนเทศศาสตร์ทางสุขภาพ บัณฑิตวิทยาลัย มหาวิทยาลัยศิลปากร
ประวัติการทำงาน	พ.ศ. 2550 – ปัจจุบัน อาจารย์ ภาควิชาสารสนเทศศาสตร์ทางสุขภาพ คณะเภสัชศาสตร์ มหาวิทยาลัยศิลปากร