

โลกแห่งดิจิทัล ยุคแห่งข้อมูลข่าวสาร

ความมหัศจรรย์ของข้อมูลทางการแพทย์ สู่การค้นพบความรู้ที่มีค่าด้วยการทำเหมืองข้อมูล

ศราวุธ แดงมาก ค.อ.ม. (เทคโนโลยีคอมพิวเตอร์)

สถานวิทยามะเร็งศิริราช, คณะแพทยศาสตร์ศิริราชพยาบาล, มหาวิทยาลัยมหิดล, กรุงเทพมหานคร 10700.

“โลกปัจจุบัน เป็นโลกดิจิทัล” “ยุคปัจจุบัน เป็นยุคแห่งข้อมูลข่าวสาร” หลายคนเคยได้ยินคำกล่าวเหล่านี้ และคงไม่สามารถปฏิเสธว่า คำกล่าวเหล่านี้ไม่เป็นความจริง เพราะข้อมูลข่าวสารต่างๆ เกิดขึ้นมากมายในชีวิตประจำวัน และอยู่ในรูปแบบของดิจิทัลแทบทั้งสิ้น ซึ่งข้อดีของดิจิทัลคือ มีความถูกต้องแม่นยำ สามารถนำไปประมวลผลในลักษณะต่างๆ ด้วยเครื่องมือต่างๆ ได้ง่ายและรวดเร็ว¹ จึงไม่แปลกที่หน่วยงานต่างๆ ธุรกิจต่างๆ โรงพยาบาล ฯลฯ นำระบบดิจิทัลไปใช้ในการดำเนินงานกันอย่างกว้างขวางแพร่หลาย แม้กระทั่งในระดับชาติ รัฐบาลยังให้ความสำคัญกับข้อมูลข่าวสารในยุคดิจิทัล ด้วยการพัฒนาขยายโครงข่ายอินเทอร์เน็ตความเร็วสูงทั้งภายในประเทศและระหว่างประเทศ² ยิ่งแสดงให้เห็นชัดเจนว่า ประเทศไทยให้ความสำคัญกับการพัฒนาระบบเครือข่ายและอินเทอร์เน็ต เพื่อเพิ่มศักยภาพและรองรับการสื่อสารข้อมูลข่าวสารในยุคดิจิทัลที่เพิ่มขึ้น

ในทางกลับกันนั้นพบว่า ปริมาณข้อมูลมีมากมายมหาศาล และเพิ่มมากขึ้นทุก ๆ ปี แต่ปริมาณข้อมูลที่เราวิเคราะห์ออกมา หรือสกัดได้ออกมาเป็นโน้ล

(Knowledge) หรือว่าเป็นความรู้ที่ออกมานั้นมีปริมาณน้อยมาก ๆ ถ้าเทียบกับปริมาณข้อมูลที่มีอยู่³ สะท้อนให้เห็นว่าธุรกิจ องค์กร หน่วยงาน โรงพยาบาล หรือแม้แต่หน่วยงานระดับชาติก็ตาม มีระบบการเก็บข้อมูลข่าวสารต่างๆ มีฐานข้อมูล และการสื่อสารข้อมูลมากมาย แต่การวิเคราะห์ข้อมูลหรือการหาความรู้ใหม่ๆ จากข้อมูลที่ตนเองมีอยู่นั้นน้อยมาก โดยเฉพาะอย่างยิ่งข้อมูลหรือระบบฐานข้อมูล ที่เกี่ยวข้องกับการดำเนินงานทางการแพทย์สมัยใหม่นั้น มีการเก็บรวบรวมข้อมูลอยู่มากมาย เชื่อมโยงกันหลายระบบ ซึ่งหากมีการนำข้อมูลต่างๆ มาสำรวจและวิเคราะห์ตามกระบวนการในการทำเหมืองข้อมูล ก็อาจจะพบความรู้ใหม่ๆ ที่มีค่าหรืออาจจะสามารถตอบคำถามบางอย่างที่เราไม่เคยหาคำตอบได้มาก่อน ก่อให้เกิดประโยชน์สูงสุดต่อการรักษาผู้ป่วยต่อไป

บทความนี้จะกล่าวถึง ความหมาย และประโยชน์ของเหมืองข้อมูล รูปแบบของการทำเหมืองข้อมูล กระบวนการในการทำเหมืองข้อมูล ข้อสรุปและข้อเสนอแนะในการค้นพบความรู้ที่มีค่าด้วยการทำเหมืองข้อมูล

ความหมาย และ ประโยชน์ของเหมืองข้อมูล

หากกล่าวถึง การสกัดเอาแร่ที่มีค่า หรือ วัสดุทางธรณีวิทยาอื่น ๆ จากใต้ผืนแผ่นดิน ก็คงหนีไม่พ้นการทำเหมืองแร่⁴ ในทำนองเดียวกันหากเราต้องการสกัดหาความรู้ใหม่ ๆ จากฐานข้อมูล หรือ ข้อมูลข่าวสารที่สะสมไว้มากมาย เราจะต้องใช้วิธีการทำเหมืองข้อมูล (Data mining) ดังนั้นเหมืองข้อมูล จึงเป็นการสกัดข้อมูลขนาดใหญ่ โดยอาศัยเทคโนโลยีที่มีประสิทธิภาพ เพื่อให้ได้มาซึ่งข้อมูลความรู้ใหม่ ๆ ที่เป็นประโยชน์ต่อองค์กร โดยความรู้ที่สำคัญของเหมืองข้อมูลได้แก่ การคาดการณ์แนวโน้ม การพยากรณ์เหตุการณ์ล่วงหน้าในอนาคต เพื่อเป็นประโยชน์ในการตัดสินใจ⁵ สำหรับในทางการแพทย์แล้วการทำเหมืองข้อมูลถือว่าเป็นเครื่องมือที่ดีและมีประโยชน์มาก เพราะผลลัพธ์ที่ได้จากการทำเหมืองข้อมูลอาจช่วยในการค้นหาข้อบกพร่องต่าง ๆ ตลอดจนค้นหาหนทางที่ดีกว่าในการรักษาพยาบาล นอกจากนี้ยังช่วยลดค่าใช้จ่ายต่าง ๆ ที่เกิดขึ้นในระหว่างการรักษาพยาบาล และการดำเนินกิจการในโรงพยาบาลอีกด้วย โดยที่ผ่านมามีงานวิจัยที่ประยุกต์ใช้เหมืองข้อมูลกับการแพทย์มากยิ่งขึ้น เช่น การประเมินความเสี่ยงการเกิดมะเร็งปอด เพื่อต้องการทราบว่าใช้ชีวิตอย่างไร จะมีความเสี่ยงเกิดมะเร็งปอด โดยใช้วิธีการตอบคำถาม และนำคำตอบที่ได้ไปเทียบกับข้อมูลที่มีอยู่ โดยอาศัยวิธีการทางเหมืองข้อมูลในการวิเคราะห์ข้อมูลเพื่อให้ได้ผลลัพธ์ออกมาว่ามีความเสี่ยงมากน้อยระดับใด⁷ การทำนายสาเหตุจากรูปแบบข้อมูลต่าง ๆ ของการกลับมารักษาซ้ำในโรงพยาบาลภายในระยะเวลาอันสั้น โดยใช้ฐานข้อมูลของโรงพยาบาลที่มีอยู่ เพื่อค้นหาต้นตอสาเหตุของปัญหา ที่ทำให้ผู้ป่วยต้องกลับมารักษาในโรงพยาบาลอีกในระยะเวลาอันสั้น ซึ่งเป็นสาเหตุหนึ่งของค่าใช้จ่ายที่เพิ่มขึ้น⁸ นอกจากนี้ยังมีการประยุกต์ใช้เหมืองข้อมูลในการคาดการณ์การคลอดก่อนกำหนดของมารดาที่มีคุณลักษณะต่าง ๆ โดยอาศัยข้อมูลที่มีอยู่ในอดีตเป็นต้น⁹

รูปแบบของการทำเหมืองข้อมูล

การทำเหมืองข้อมูลสามารถประยุกต์ใช้ให้เหมาะสมกับงานต่าง ๆ ได้หลายรูปแบบ โดยมองที่

ผลลัพธ์ของงานที่ต้องการจากนั้นมาพิจารณาว่า ต้องใช้การทำเหมืองข้อมูลรูปแบบใด มาช่วยจัดการงานเพื่อให้บรรลุผลลัพธ์ที่ต้องการ ซึ่งในการทำเหมืองข้อมูล มีรูปแบบต่าง ๆ ดังต่อไปนี้

1. การจำแนกประเภทข้อมูล (Classification) เป็นการสร้างตัวแบบจากข้อมูลที่มีการจำแนกประเภทแล้ว เพื่อใช้ตัวแบบนั้นในการจำแนกข้อมูลที่ไม่ทราบประเภท โดยตัวแบบอาจสร้างขึ้นโดยวิธีการเรียนรู้โดยเครื่อง (Machine Learning) หรือใช้วิธีการเชิงสถิติ เช่น Regression เป็นต้น ซึ่งการจำแนกประเภทข้อมูลมีเทคนิคที่สำคัญได้แก่ ตัวแบบเพื่อนบ้านที่ใกล้ที่สุดเคตว (K-nearest neighbors), ต้นไม้การตัดสินใจ (Decision Tree), ตัวแบบการจำแนกประเภทข้อมูลแบบเบย์ส (Bayes Classifier), ตัวแบบจำแนกข้อมูลแบบเครือข่ายประสาท (Neural Network), ตัวแบบ Support Vector Machine (SVM) ส่วนตัวอย่างของงานที่ใช้รูปแบบเหมืองข้อมูลการจำแนกประเภทข้อมูลได้แก่ การจำแนกโอกาสที่ผู้ป่วยจะเป็นโรค เป็นต้น

2. การจัดกลุ่มข้อมูล (Clustering) เป็นการแบ่งกลุ่มข้อมูลออกเป็นกลุ่ม โดยข้อมูลที่อยู่ในกลุ่มเดียวกันจะมีลักษณะที่คล้ายคลึงกัน และข้อมูลที่อยู่ต่างกลุ่ม จะมีลักษณะที่แตกต่างกัน การจัดกลุ่มข้อมูลต้องอาศัยการวัดความคล้ายคลึง หรือความแตกต่างระหว่างกลุ่ม 2 ตัว ซึ่งทำได้หลายแบบ ขึ้นกับชนิดของข้อมูล และการกำหนดความคล้ายคลึงของข้อมูล ตามเป้าประสงค์ในการจัดกลุ่มข้อมูล ซึ่งการจัดกลุ่มข้อมูล มีเทคนิคที่สำคัญได้แก่ วิธีการจับกลุ่มแบบ K-Means, Hierarchical clustering, การจับกลุ่มข้อมูลด้วยวิธี DBSCAN, การประเมินการจับกลุ่ม (Cluster Validity) ส่วนตัวอย่างของงานที่ใช้รูปแบบเหมืองข้อมูลการจัดกลุ่มข้อมูลได้แก่ การจัดกลุ่มของยีนส์ในโครโมโซมที่ทำหน้าที่คล้ายคลึงกัน เป็นต้น

3. การสร้างกฎความสัมพันธ์ (Association Rules) เป็นการหากฎที่บอกถึงความสัมพันธ์ระหว่างข้อมูลซึ่งมักเกิดขึ้นพร้อม ๆ กันอยู่เสมอ โดยกฎความสัมพันธ์ที่สร้างได้จะระบุถึงความสัมพันธ์ว่า เมื่อพบเหตุการณ์หนึ่งหรือหลายเหตุการณ์เกิดขึ้นจะมีโอกาสสูงที่เหตุการณ์อีกอย่างหนึ่งหรืออีกหลายเหตุการณ์จะเกิดขึ้นด้วย ซึ่ง

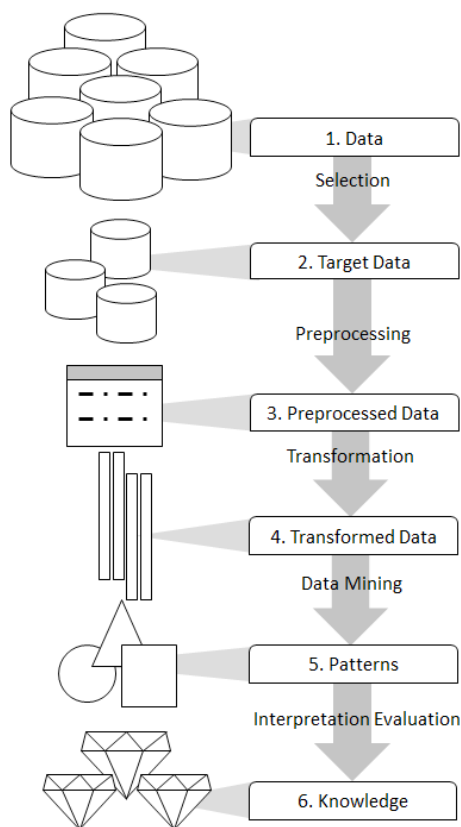
การสร้างกฎความสัมพันธ์ มีเทคนิคที่สำคัญได้แก่ การหากฎความสัมพันธ์โดยวิธี Apriori, FP-Tree, การสร้างกฎความสัมพันธ์จาก Frequent Itemsets, multiple Minimum Support ส่วนตัวอย่างของงานที่ใช้รูปแบบเหมือนข้อมูลการสร้างกฎความสัมพันธ์ได้แก่ การหาความสัมพันธ์ในการเกิดโรคจากเหตุการณ์หรือสภาพแวดล้อม เป็นต้น¹⁰

4. การคาดคะเน (Estimation) มีลักษณะคล้ายกับการจำแนกประเภทข้อมูล คือมีการสร้างตัวแบบเพื่อคาดคะเนค่าของตัวแปรที่ไม่ทราบค่า จากค่าตัวแปรอื่นที่ทราบค่า ข้อแตกต่าง คือ ตัวแปรที่จะถูกคาดคะเนจะมีชนิดเป็นค่าต่อเนื่อง หรือเป็นจำนวนเต็ม ส่วนตัวอย่างของงานที่ใช้รูปแบบเหมือนข้อมูลการคาดคะเนได้แก่ การประมาณระดับคะแนนที่ผู้สอบจะได้จากคุณสมบัติของผู้สอบ เป็นต้น

5. การค้นหาข้อมูลที่มีความผิดปกติ (Data Anomaly Detection) เป็นการค้นหาข้อมูลที่มีลักษณะผิดปกติ หรือมีความแตกต่างกับข้อมูลส่วนใหญ่ (Outlier Data) ข้อมูลเหล่านี้มีจำนวนน้อยปะปนอยู่กับข้อมูลส่วนใหญ่ ซึ่งอาจส่งผลให้ผลการวิเคราะห์ข้อมูลมีความแม่นยำน้อยลง การค้นหาข้อมูลผิดปกติ อาจช่วยในงานบางอย่าง ที่ต้องการตรวจสอบข้อมูลที่ผิดปกติ เช่น การจําแนกรูขุมที่มึความผิดปกติ (Fraud Detection) การตรวจจับการส่งผ่านข้อมูลผิดปกติในเครือข่ายคอมพิวเตอร์ (Intruder Detection) เป็นต้น¹¹

กระบวนการในการทำเหมืองข้อมูล

เมื่อรับทราบแล้วว่า ต้องประยุกต์ใช้วิธีการทำเหมืองข้อมูลรูปแบบใด เพื่อให้บรรลุผลลัพธ์ที่ต้องการ สิ่งต่อมาคือ การลงมือทำเหมืองข้อมูล ซึ่งไม่ว่าจะเป็นงานด้านใด เหมืองข้อมูลก็จะมีกระบวนการหลักอยู่ 6 ขั้นตอน เริ่มตั้งแต่การเตรียมข้อมูล เรื่อยไปจนถึงการสกัดออกมาเป็นโน้ต หรือความรู้ที่มีค่า ดังนั้นเพื่อให้ได้ผลลัพธ์ที่แม่นยำจึงจำเป็นต้องพิจารณาข้อมูลที่มีอยู่ตามกระบวนการในการทำเหมืองข้อมูลดังนี้



รูปที่ 1. ภาพรวมของกระบวนการการทำเหมืองข้อมูลตามหลักของ Knowledge Discovery in Databases หรือ KDD

ที่มา : ผู้เขียนจัดทำขึ้นเอง

1. Data คือข้อมูลต่างๆ ที่เก็บไว้ในฐานข้อมูลในแต่ละองค์กรที่สามารถเลือก (Selection) ไปเป็นข้อมูลเป้าหมาย (Target Data) ที่จะใช้ในการทำเหมืองข้อมูลต่อไป

2. Target Data คือข้อมูลเป้าหมาย หมายถึงข้อมูลที่เลือกมาใช้ในการทำเหมืองข้อมูล โดยส่วนใหญ่จะเป็นข้อมูลในรูปแบบที่มีโครงสร้าง กล่าวคือเป็นข้อมูลที่เก็บไว้เป็นรูปแบบที่แน่นอน เช่น ข้อมูลอายุ เพศ วันเดือนปีเกิด ข้อมูลรหัสโรค เป็นต้น ทั้งนี้หากจะทำเหมืองข้อมูลในงานด้านใด ก็จะต้องเลือกข้อมูลให้สัมพันธ์กับงานด้านนั้น เช่นหากจะพยากรณ์การเกิดมะเร็งบางชนิด ก็อาจจะต้องมีข้อมูลการสูบบุหรี่ หรือ

การตีความแอลกอริธึมเป็นต้น¹² โดยอาจจะต้องมีงานวิจัยรองรับ หรือมีผู้เชี่ยวชาญเฉพาะด้านในการเลือกข้อมูลเป้าหมาย เพื่อนำไปใช้ในการทำเหมืองข้อมูลให้ได้ผลลัพธ์ที่แม่นยำมากยิ่งขึ้นทั้งนี้เมื่อเลือกข้อมูลเป้าหมายได้สำเร็จแล้ว ก็จะไปสู่การเตรียมการ (Preprocessing) เพื่อเข้าสู่กระบวนการเตรียมข้อมูล (Preprocessed data) ต่อไป

3. Preprocessed Data คือกระบวนการเตรียมข้อมูล เนื่องจากข้อมูลที่จะนำมาทำเหมืองข้อมูลนี้ได้มาจากฐานข้อมูลขนาดใหญ่ บางฐานข้อมูลมีโครงสร้างข้อมูลที่ไม่ค่อยสมบูรณ์ เช่น ข้อมูลรหัสโรคอาจมีช่องว่างซึ่งอาจเกิดจากความผิดพลาดระหว่างบันทึกข้อมูลหรือการคีย์ข้อมูล ดังนั้นกระบวนการเตรียมข้อมูลคือการทำให้ข้อมูลให้มีความสมบูรณ์มากยิ่งขึ้น ทั้งนี้อาจจะต้องมีการซ่อมแซมข้อมูล หรือ ตัดข้อมูลบางรายการที่ไม่มีความหมาย เพื่อจะทำให้กระบวนการในการวิเคราะห์ข้อมูลมีความแม่นยำมากยิ่งขึ้น ซึ่งในขั้นตอนนี้เมื่อสำเร็จแล้วจะนำไปสู่การแปลงข้อมูลให้เหมาะสม (Transformation) ตามกระบวนการแปลงข้อมูล (Transformed Data) ต่อไป

4. Transformed Data คือกระบวนการแปลงข้อมูล เมื่อทำการเตรียมข้อมูลเสร็จเรียบร้อยแล้ว ขั้นตอนนี้จะเป็นกระบวนการแปลงข้อมูล เพื่อให้เหมาะสมกับการวิเคราะห์ข้อมูล เช่น ข้อมูลวันเดือนปีเกิดของผู้ป่วย อาจจะต้องแปลงข้อมูลให้เป็นอายุในปัจจุบัน และอาจรวมไปถึงการจัดกลุ่มอายุของผู้ป่วย ข้อมูลโรคบางชนิดอาจจะต้องทำการจัดกลุ่มโรคเพื่อลดการกระจายของข้อมูล เป็นต้น โดยกระบวนการแปลงข้อมูลเมื่อทำสำเร็จแล้ว ก็จะนำไปสู่กระบวนการทำเหมืองข้อมูล (Data mining) เพื่อให้ได้มาซึ่งรูปแบบต่างๆ ของข้อมูล (Patterns) ต่อไป

5. Patterns คือรูปแบบของข้อมูล ในขั้นตอนนี้จะจะเป็นกระบวนการทำเหมืองข้อมูล โดยอาศัยรูปแบบของการทำเหมืองข้อมูลในลักษณะต่างๆ ตามผลลัพธ์ของงานที่ต้องการ เช่น การจำแนกประเภทข้อมูล (Classification) การจัดกลุ่มข้อมูล (Clustering) การสร้างกฎความสัมพันธ์ (Association Rules) การคาดคะเน

(Estimation) การค้นหาข้อมูลที่มีความผิดปกติ (Data Anomaly Detection) ซึ่งในการทำเหมืองข้อมูลดังกล่าวมีโปรแกรมที่เป็นเครื่องมือช่วยในการประมวลผลที่เป็นที่นิยมแพร่หลายคือ โปรแกรมเวก้า (Weka) ซึ่งเป็นโปรแกรมประเภทฟรีแวร์ สามารถใช้งานได้ทุกระบบปฏิบัติการ และมีความง่ายในการใช้งาน รวมถึงมีคู่มือการใช้งานที่แพร่หลายทั่วไป

เมื่อประมวลผลโดยโปรแกรมดังกล่าวแล้วผลลัพธ์ที่ได้ก็จะออกมาเป็นรูปแบบข้อมูลในลักษณะต่างๆ ทั้งนี้จะต้องมีการประเมินค่าความถูกต้องและความแม่นยำ ของรูปแบบข้อมูลในลักษณะต่างๆ ที่ได้ออกมา (Interpretation Evaluation) เพื่อนำผลลัพธ์ที่ดีที่สุดออกไปใช้งานเป็นความรู้ใหม่ (Knowledge) ต่อไป

6. Knowledge คือการนำเสนอผลลัพธ์ที่ดีที่สุดหรือ ความรู้ใหม่ที่ค้นพบได้ จากการทำเหมืองข้อมูล ทั้งนี้อาจจะแสดงเป็นกราฟ หรือตาราง เพื่อให้เกิดความเข้าใจตรงไปตรงมา

ข้อสรุปและข้อเสนอแนะในการค้นพบความรู้ที่มีค่าด้วยการทำเหมืองข้อมูล

จากข้อมูลข่าวสาร ที่เคยเก็บอยู่ในรูปแบบของเอกสาร ผู้การเปลี่ยนผ่านเข้าสู่โลกแห่งดิจิทัล ยุคแห่งข้อมูลข่าวสารทุกอย่างถูกจัดเก็บให้อยู่ในรูปแบบของดิจิทัลเป็นส่วนใหญ่ มีการสื่อสารข้อมูลถึงกัน ก่อให้เกิดความมหาศาลของข้อมูลข่าวสารต่างๆ ซึ่งนับรวมไปถึงข้อมูลทางการแพทย์ด้วย และหากเทียบกับในแง่หรือความรู้ที่สกัดออกมาได้ จากปริมาณข้อมูลที่มีอยู่ในปัจจุบันนั้นถือว่าน้อยมาก ความรู้ใหม่ๆ ที่อาจซ่อนอยู่ในฐานข้อมูลขนาดใหญ่เปรียบเสมือนแร่ที่มีค่า หากต้องการนำออกมาต้องอาศัยกระบวนการในการทำเหมืองข้อมูลหรือ Data Mining ผ่านรูปแบบของการทำเหมืองข้อมูล และเทคนิคต่างๆ เพื่อสกัดความรู้หรือผลลัพธ์ออกมาใช้ให้เกิดประโยชน์ในงานด้านต่างๆ ต่อไป

คณะแพทยศาสตร์ศิริราชพยาบาล ถือว่าเป็นหน่วยงานหนึ่งที่มีปริมาณข้อมูลมหาศาลทั้งข้อมูลทางการแพทย์ และข้อมูลด้านการบริหารงานต่างๆ

ดังนั้นจึงเหมาะสมเป็นอย่างยิ่ง หากจะมีการประยุกต์ใช้เหมืองข้อมูล เพื่อสกัดความรู้ใหม่ๆ หรือคำตอบบางอย่าง ออกมาใช้ประโยชน์ได้อย่างมีคุณค่า

เอกสารอ้างอิง

1. information.go.ke [Internet]. Kenya: Advantage of Digital Signal Over Analog Signal Broadcasting; c2014 [cited 2016 June 30]. Available from: <http://www.information.go.ke/?p=690>.
2. กระทรวงเทคโนโลยีสารสนเทศและการสื่อสาร. ข่าวกระทรวง. [อินเทอร์เน็ต]. กรุงเทพฯ; 2559 [เข้าถึงเมื่อ 30 มิย. 2559]. เข้าถึงจาก: <http://goo.gl/5ZBSEo>.
3. พีรพล เวทีกุล. เหมืองข้อมูล (Data Mining). [อินเทอร์เน็ต]. กรุงเทพฯ; 2556 [เข้าถึงเมื่อ 30 มิย. 2559]. เข้าถึงจาก: <https://goo.gl/iQr2MP>.
4. Wikipedia. Mining [Internet]. 2016 [cited 2016 July 6]. Available from: <https://en.wikipedia.org/wiki/Mining>.
5. Thearling K. An Introduction to Data Mining Discovering hidden value in your data warehouse. [Internet]. Boston; 2012 [cited 2016 July 6]. Available from: <http://goo.gl/DNNby>.
6. Crockett D, Eliason B. What Is Data Mining in Health care? [Internet]. UT; 2016 [cited 2016 December 1]. Available from: <https://goo.gl/ilzwBo>.
7. Ahmed K, Emran AA, Jesmin T, Mukti RF, Rahman MZ, Ahmed F. Early detection of lung cancer risk using data mining. Asian Pac J Cancer Prev 2013;14:595-8.
8. Lee EW. Selecting the best prediction model for hospital readmission within 28 days: data mining approach using hospital database. International Journal of Integrated Care [Internet]. 2014 [cited 2016 December 1]; 14: [1] Available from: <http://www.ijic.org/articles/abstract/10.5334/ijic.1783/>
9. Desikan P, Hsu KW, Srivastava J. Data mining for healthcare management. 2011 Siam International Conference on data mining. [Internet]. 2011 [cited 2016 December 1]; [32]. Available from: <https://www.siam.org/meetings/sdm11/dmhm.pdf>.
10. กิตติศักดิ์ สุมาลย์. การคัดกรองสุขภาพเบื้องต้นโดยใช้เทคนิคเหมืองข้อมูล [สารนิพนธ์วิทยาศาสตรมหาบัณฑิต]. กรุงเทพฯ: มหาวิทยาลัยธุรกิจบัณฑิต; 2555. หน้า 1
11. สุรพงศ์ เอื้อวัฒนามงคล. การทำเหมืองข้อมูล (Data Mining). พิมพ์ครั้งที่ 1. กรุงเทพฯ: ห้างหุ้นส่วนจำกัด บางกอกบลิ๊อค; 2557. หน้า 3
12. Kharya S. Using data mining techniques for diagnosis and prognosis of cancer disease. International Journal of Computer Science [Internet]. 2012 [cited 2016 December 1];2: [57] Available from: <http://airccse.org/journal/ijcseit/papers/2212ijcseit06.pdf>.