

## สรุปความรู้ที่ได้จากการโครงการแลกเปลี่ยนเรียนรู้ ครั้งที่ 2 Data Mining

วันที่ 16 กุมภาพันธ์ 2558

จัดโดยศูนย์เทคโนโลยีสารสนเทศ มหาวิทยาลัยแม่โจ้

### 1. ความหมายและลักษณะการเหมืองข้อมูล

เหมืองข้อมูล (Data Mining) คือ กระบวนการที่นำมากระทำกับข้อมูลจำนวนมากเพื่อค้นหารูปแบบและความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูลนั้นๆ

### 2. ประเภทข้อมูลที่ใช้ทำเหมืองข้อมูล

การทำเหมืองข้อมูลสามารถนำข้อมูลที่มาจากหลากหลายประเภท ได้แก่

- Relational Database เป็นฐานข้อมูลที่จัดเก็บอยู่ในรูปแบบของตาราง โดยในแต่ละตารางจะประกอบไปด้วยแถวและคอลัมน์ ความสัมพันธ์ของข้อมูลทั้งหมดสามารถแสดงได้โดย Entity Relationship Model
- Data Warehouses เป็นการเก็บรวบรวมข้อมูลจากหลายแหล่งมาเก็บไว้ในรูปแบบเดียวกันและรวบรวมไว้ในที่ๆ เดียวกัน
- Transactional Database ประกอบด้วยข้อมูลที่แต่ละ Transaction แทนด้วยเหตุการณ์ในขณะใดขณะหนึ่ง เช่น ใบเสร็จรับเงิน จะเก็บข้อมูลในรูปแบบชื้อลูกค้าและรายการสินค้าที่ลูกค้ารายชื้อ
- Advanced Database เป็นฐานข้อมูลที่จัดเก็บในรูปแบบอื่นๆ เช่น ข้อมูลแบบ Object-Oriented ข้อมูลที่เป็น Text File ข้อมูลมัลติมีเดีย ข้อมูลในรูปของ Web

### 3. ขั้นตอนการทำเหมืองข้อมูล

การทำเหมืองข้อมูล ประกอบด้วยขั้นตอนการทำงานย่อยที่จะเปลี่ยนข้อมูลดิบให้กลายเป็นความรู้ ประกอบด้วยขั้นตอนดังนี้

- Data Cleaning ขั้นตอนสำหรับการคัดข้อมูลที่เกี่ยวข้องออกไป
- Data Integration ขั้นตอนการรวมข้อมูลที่มีหลายแหล่งให้เป็นข้อมูลชุดเดียวกัน
- Data Selection ขั้นตอนการดึงข้อมูลสำหรับการวิเคราะห์จากแหล่งที่บันทึกไว้
- Data Transformation ขั้นตอนการแปลงข้อมูลให้เหมาะสมสำหรับการใช้งาน
- Data Mining ขั้นตอนการค้นหารูปแบบที่เป็นประโยชน์จากข้อมูลที่มีอยู่
- Pattern Evaluation ขั้นตอนการประเมินรูปแบบที่ได้จากการทำเหมืองข้อมูล
- Knowledge Representation ขั้นตอนการนำเสนอความรู้ที่ค้นพบและนำเสนอด้วยวิธีที่เข้าใจได้

#### 4. งานของเหมืองข้อมูล

เทคนิคการทำเหมืองข้อมูลจะเป็นปัญหาที่ต้องใช้เหตุผล ไม่มีเทคนิคหรือเครื่องมือเพียงชนิดเดียวของเหมืองข้อมูลที่เหมาะสมกับงานทุกชนิด ซึ่งเทคนิคที่ใช้จะแตกต่างกันไปขึ้นอยู่กับชนิดของงาน รูปแบบของงานที่ได้ภายหลังการทำเหมืองข้อมูลมี 6 งาน ดังนี้

1. การจัดหมวดหมู่ (Classification)
2. การประเมินค่า (Estimation)
3. การทำนายล่วงหน้า (Prediction)
4. การจัดกลุ่มโดยอาศัยความใกล้ชิด (Affinity Group)
5. การรวมตัว (Clustering)
6. การบรรยาย (Description)

#### 5. บทสรุป

การทำเหมืองข้อมูลได้ถูกนำไปประยุกต์ใช้ในงานหลายประเภท เช่น ด้านธุรกิจช่วยในการตัดสินใจของผู้บริหาร ด้านวิทยาศาสตร์และการแพทย์ หรือด้านเศรษฐกิจและสังคม การทำเหมืองข้อมูลมีการพัฒนาจากเดิมที่มีการจัดเก็บข้อมูลอย่างง่าย มาสู่การจัดเก็บในรูปแบบข้อมูลที่สามารถดึงข้อมูลสารสนเทศมาใช้จนถึงการสามารถค้นพบความรู้ที่ซ่อนอยู่ในข้อมูลนั้นๆ

#### 6. แหล่งอ้างอิง

<http://www.no-poor.com/dssandos/Chapter5-dss.htm>

<http://alaska.reru.ac.th/text/datamining.pdf>

ผู้สรุปข้อมูล นางสาวพัชรี ยางยี่น