

Data Mining คือ การวิเคราะห์ข้อมูลจาก big data เพื่อหาความสัมพันธ์ของข้อมูลที่ซ่อนอยู่ โดยการจำแนกประเภท รูปแบบ เชื่อมโยงข้อมูลที่สัมพันธ์กัน และหาความน่าจะเป็นที่จะเกิดขึ้น เพื่อให้ได้ความรู้ใหม่ๆ ที่สามารถนำไปใช้ในการประกอบการตัดสินใจต่างๆ

ขั้นตอนการทำเหมืองข้อมูล

1. Data Cleaning เป็นการคัดข้อมูลที่ไม่เกี่ยวข้องออกไป
2. Data Integration เป็นการรวมข้อมูลจากหลายแหล่งให้เป็นข้อมูลชุดเดียวกัน
3. Data Selection เป็นการดึงข้อมูลสำหรับวิเคราะห์จากแหล่งที่บันทึกไว้
4. Data Transformation เป็นการแปลงข้อมูลให้เหมาะกับการใช้งาน
5. Data Mining เป็นการค้นหารูปแบบที่เป็นประโยชน์จากข้อมูลที่มี
6. Pattern Evaluation เป็นการประเมินรูปแบบที่ได้จากการทำเหมืองข้อมูล
7. Knowledge Representation เป็นการนำเสนอความรู้ที่พบ

