

# การวิเคราะห์วิวัฒนาการของ Data Mining จากอดีตสู่ปัจจุบัน

ผศ. ทิพย์วรรณ พู่เพ็ญ

สาขาเทคโนโลยีสารสนเทศ คณะวิทยาการคอมพิวเตอร์และเทคโนโลยีสารสนเทศ

มหาวิทยาลัยราชภัฏรำไพพรรณี

Data Mining หรือการทำเหมืองข้อมูล ได้ถือกำเนิดขึ้นอย่างเป็นทางการเป็นระบบในทศวรรษ 1990 โดยมีเป้าหมายเพื่อสกัดองค์ความรู้ที่ซ่อนอยู่ในฐานข้อมูลขนาดใหญ่ผ่านกระบวนการที่เป็นที่รู้จักในนาม Knowledge Discovery in Databases (KDD) ตลอดระยะเวลาที่สามทศวรรษที่ผ่านมา Data Mining ได้พัฒนาและเปลี่ยนแปลงไปตามบริบททางเทคโนโลยีและความต้องการเชิงธุรกิจ บทความนี้มุ่งสำรวจวิวัฒนาการของ Data Mining ตั้งแต่ยุคเริ่มต้นที่อาศัยเทคนิคทางสถิติและ Machine Learning แบบดั้งเดิม ไปจนถึงการเปลี่ยนผ่านไปสู่ Data Science และ AI ที่ขับเคลื่อนด้วย Big Data และ Deep Learning ในยุคดิจิทัลปัจจุบัน ผลการศึกษาชี้ให้เห็นว่า แม้คำว่า “Data Mining” จะถูกใช้น้อยลงในเชิงการตลาดและการศึกษา แต่แก่นแท้ของศาสตร์ดังกล่าวยังคงมีความสำคัญและเป็นรากฐานของระบบอัจฉริยะยุคใหม่ เช่น ระบบแนะนำ (Recommendation Systems) การวิเคราะห์ความรู้สึก (Sentiment Analysis) ความปลอดภัยไซเบอร์ และการพยากรณ์ทางการแพทย์ นอกจากนี้ แนวโน้มในอนาคตยังชี้ไปสู่การทำเหมืองข้อมูลแบบเรียลไทม์ (Real-time Data Mining) การสร้างแบบจำลองที่สามารถอธิบายได้ (Explainable AI) และการบูรณาการด้านจริยธรรมและการตรวจสอบอคติ (Bias Detection) เพื่อตอบสนองสังคมที่ขับเคลื่อนด้วยข้อมูลอย่างมีความรับผิดชอบ

## บทนำ

ในศตวรรษที่ 21 ข้อมูลถูกยกย่องว่าเป็น “ทรัพยากรเชิงกลยุทธ์” ที่มีมูลค่าไม่ต่างจากทรัพยากรธรรมชาติ เช่น น้ำมันหรือแร่ธาตุ องค์กรทั้งภาครัฐและเอกชนต่างตระหนักว่าข้อมูลที่สะสมและเกิดขึ้นอย่างต่อเนื่องในปริมาณมหาศาล สามารถนำมาใช้เพื่อสร้างข้อได้เปรียบเชิงการแข่งขัน พัฒนานวัตกรรม และยกระดับคุณภาพชีวิตของสังคม กระบวนการที่ทำให้ข้อมูลดิบแปรเปลี่ยนเป็นความรู้เชิงลึกที่ใช้ประโยชน์ได้จริง จึงเป็นหัวใจสำคัญของการบริหารจัดการยุคดิจิทัล

Data Mining หรือการทำเหมืองข้อมูล เกิดขึ้นเพื่อตอบโจทย์ดังกล่าว โดยผสมองค์ความรู้จากหลายศาสตร์ ไม่ว่าจะเป็น สถิติศาสตร์ (เพื่อการวิเคราะห์เชิงปริมาณ) วิทยาการคอมพิวเตอร์ (เพื่อการจัดเก็บและประมวลผล) และ ปัญญาประดิษฐ์ (เพื่อการเรียนรู้และพยากรณ์) เข้าด้วยกัน อย่างไรก็ตาม การเปลี่ยนแปลงอย่างรวดเร็วของเทคโนโลยี โดยเฉพาะ Big Data, Internet of Things (IoT) และ Deep Learning ทำให้คำว่า Data Mining ถูกตีความใหม่ กลายเป็นส่วนหนึ่งของ Data Science และ Artificial Intelligence ที่มีขอบเขตกว้างและทันสมัยมากกว่า

## ประวัติศาสตร์และวิวัฒนาการของ Data Mining

ยุคบุกเบิก (1960–1980) จุดกำเนิดจากสถิติและฐานข้อมูล

ก่อนที่คำว่า Data Mining จะปรากฏในเชิงวิชาการ นักวิจัยได้พัฒนาวิธีการวิเคราะห์ข้อมูลด้วยเทคนิคสถิติและคณิตศาสตร์ เช่น การถดถอยเชิงเส้น (Linear Regression) การวิเคราะห์องค์ประกอบหลัก (Principal Component Analysis: PCA) และทฤษฎีความน่าจะเป็นของเบย์ (Bayes' Theorem) ขณะเดียวกัน การพัฒนาระบบฐานข้อมูลเชิงสัมพันธ์ (Relational Databases) และภาษา SQL ในทศวรรษ 1970 ทำให้การจัดเก็บและดึงข้อมูลจำนวนมากเป็นไปอย่างเป็นระบบและมีประสิทธิภาพ

ยุคเฟื่องฟู (1990) การนิยามและความรุ่งเรืองของ Data Mining

ยุค 1990 ถือเป็นช่วงที่ Data Mining ถูกยอมรับอย่างกว้างขวาง ทั้งในเชิงทฤษฎีและการประยุกต์ โดยถูกกำหนดให้เป็นขั้นตอนหลักของกระบวนการ KDD ซึ่งประกอบด้วย 5 ขั้นตอน ได้แก่ (1) การเลือกข้อมูล (2) การเตรียมข้อมูล (3) การแปลงข้อมูล (4) การทำเหมืองข้อมูล และ (5) การตีความผลลัพธ์ ตัวอย่างความก้าวหน้าในยุคนี้ได้แก่

- Association Rule Mining ด้วยอัลกอริธึม Apriori ซึ่งกลายเป็นมาตรฐานในการทำ Market Basket Analysis
- Decision Trees (เช่น ID3, C4.5) ที่ใช้ในการจำแนกและพยากรณ์ข้อมูล
- Clustering Algorithms เช่น K-Means ที่ช่วยค้นหาโครงสร้างแฝงของข้อมูล
- Neural Networks ที่แม้ยังไม่ซับซ้อน แต่ได้วางรากฐานของ Deep Learning

ยุค Big Data และการประยุกต์ใช้เชิงอุตสาหกรรม (2000–2010)

การมาของอินเทอร์เน็ต เครือข่ายสังคมออนไลน์ และระบบธุรกรรมดิจิทัล ทำให้เกิดข้อมูลที่มีคุณสมบัติ 3V (Volume, Velocity, Variety) หรือที่ต่อมายายเป็น 5V (รวม Veracity และ Value) Data Mining จึงถูกพัฒนาให้รองรับข้อมูลขนาดใหญ่และซับซ้อนผ่านเทคโนโลยีใหม่ เช่น Hadoop และ Spark นอกจากนี้ยังถูกประยุกต์ใช้อย่างแพร่หลายในธุรกิจ

- การเงิน การตรวจจับการฉ้อโกงบัตรเครดิต, Credit Scoring
- การค้าออนไลน์ Recommendation Systems
- การแพทย์ การคาดการณ์โรคและการวิเคราะห์จีโนม
- ยุคปัจจุบัน (2010–ปัจจุบัน): การบูรณาการสู่ Data Science และ AI

ในช่วงทศวรรษที่ผ่านมา Data Mining ถูกหลอมรวมเข้ากับ Data Science ซึ่งเป็นสหวิทยาการที่ครอบคลุมตั้งแต่การเก็บข้อมูล การประมวลผล การสร้างโมเดล ไปจนถึงการสื่อสารผลลัพธ์ การเติบโตของ Machine Learning และ Deep Learning ทำให้โมเดลสามารถประมวลผลข้อมูลที่ไม่มีโครงสร้าง เช่น ข้อความ รูปภาพ และเสียง ได้อย่างมีประสิทธิภาพ ตัวอย่างที่ชัดเจนคือการใช้ Convolutional Neural Networks (CNNs) ในการจำแนกรูปภาพ และ Recurrent Neural Networks (RNNs) ในการประมวลผลภาษาธรรมชาติ

### การเปลี่ยนผ่านสู่ Data Science และ AI

แม้คำว่า Data Mining จะปรากฏน้อยลงในเชิงการตลาดและการศึกษา แต่แก่นแท้ของศาสตร์ยังคงดำรงอยู่ และได้กลายเป็นส่วนประกอบหลักของระบบ AI และการวิเคราะห์ข้อมูลยุคใหม่ เช่น

- Deep Learning พัฒนามาจาก Neural Networks
- AutoML ยกกระดับอัลกอริธึม Data Mining ให้ทำงานโดยอัตโนมัติ
- Anomaly Detection สำหรับความปลอดภัยไซเบอร์
- Sentiment Analysis สำหรับการวิเคราะห์โซเชียลมีเดีย

การเปลี่ยนแปลงของคำศัพท์เกิดขึ้นทั้งจากเหตุผลด้าน ภาพลักษณ์ทางการตลาด และ การขยายขอบเขตการทำงาน ของนักวิเคราะห์ข้อมูลยุคใหม่

## บทสรุปและแนวโน้มอนาคต

แม้คำว่า Data Mining จะค่อย ๆ ถูกแทนที่โดยคำศัพท์ที่กว้างกว่าอย่าง Data Science และ AI แต่รากฐานและเทคนิคของศาสตร์นี้ยังคงมีบทบาทสำคัญต่อการประยุกต์ใช้งานในโลกดิจิทัล แนวโน้มอนาคตของ Data Mining และ AI ได้แก่

- Real-time Data Mining การวิเคราะห์ข้อมูลแบบสตรีมจาก IoT และธุรกรรมดิจิทัล
- Explainable AI (XAI) โมเดลที่ไม่เพียงแม่นยำแต่ยังตีความได้
- Ethical AI: การตรวจสอบและลดอคติที่อาจเกิดในข้อมูลหรือโมเดล
- Cross-disciplinary Applications การประยุกต์ใช้ในด้านสาธารณสุข สิ่งแวดล้อม และการศึกษา
- AI Governance การจัดการข้อมูลและอัลกอริธึมอย่างมีจริยธรรมและความโปร่งใส

ดังนั้น Data Mining ไม่ได้หายไป แต่ได้วิวัฒนาการจนกลายเป็นหัวใจที่ขับเคลื่อน Data Science และ AI โดยยังคงเป็นศาสตร์ที่นักวิจัยและผู้ปฏิบัติงานด้านเทคโนโลยีสารสนเทศต้องเข้าใจและประยุกต์ใช้อย่างมีวิจารณญาณในอนาคต