

สรุปบทเรียนการพัฒนาความรู้ E-learning รอบที่ ๑ ประจำปี ๒๕๖๘

เรื่อง การวิเคราะห์ข้อมูล (Data Analytics)

การวิเคราะห์ข้อมูล (Data Analytics) คือ กระบวนการทางวิทยาศาสตร์ที่ใช้เทคนิคและหลักการเชิงสถิติ คณิตศาสตร์ และวิทยาการคอมพิวเตอร์ในการประมวลผลและวิเคราะห์ชุดข้อมูลขนาดใหญ่และซับซ้อน เพื่อสังเคราะห์ข้อมูลให้สามารถนำมาใช้ประโยชน์ได้ กระบวนการนี้ช่วยค้นหารูปแบบ แนวโน้ม และข้อมูลเชิงลึกที่อยู่ในข้อมูลจำนวนมาก ซึ่งสามารถนำไปใช้ในการคาดการณ์แนวโน้มที่อาจเกิดขึ้น วิเคราะห์ผลกระทบของการตัดสินใจ และสนับสนุนการตัดสินใจในด้านต่าง ๆ เช่น การพัฒนากลยุทธ์ทางการตลาด การปรับปรุงประสิทธิภาพการดำเนินงาน และการสร้างนวัตกรรม

ประเภทของการวิเคราะห์ข้อมูล

1. การวิเคราะห์แบบพื้นฐาน (Descriptive Analytics): การวิเคราะห์ข้อมูลพื้นฐานที่เกิดขึ้นแล้วให้เข้าใจได้ง่าย ไม่ซับซ้อน มุ่งเน้นการสรุปและอธิบายลักษณะของข้อมูลที่เก็บรวบรวม เพื่อให้เห็นภาพรวมและความสัมพันธ์ของข้อมูล โดยใช้สถิติเบื้องต้น เช่น ค่าเฉลี่ย มัธยฐาน ฐานนิยม เป็นต้น การวัดการกระจายของข้อมูล เช่น พิสัย ส่วนเบี่ยงเบนมาตรฐาน เป็นต้น และการนำเสนอข้อมูลในรูปแบบกราฟ แผนภูมิ และตาราง เพื่อให้เห็นภาพรวมและแนวโน้มของข้อมูลอย่างรวดเร็วและชัดเจน ตัวอย่างการวิเคราะห์เชิงพรรณนา ได้แก่ การสรุปผลการสำรวจความคิดเห็นของลูกค้า โดยแสดงจำนวนและร้อยละของผู้ที่พึงพอใจในผลิตภัณฑ์ หรือการวิเคราะห์คะแนนสอบของนักเรียน เพื่อหาค่าเฉลี่ยและการกระจายของคะแนน

2. การวิเคราะห์แบบพยากรณ์ (Predictive Analytics): การวิเคราะห์เพื่อคาดการณ์หรือพยากรณ์สิ่งที่จะเกิดขึ้นหรือมีแนวโน้มจะเกิดขึ้นโดยใช้ข้อมูลที่เกิดขึ้นแล้วในอดีตมาวิเคราะห์เพื่อคาดการณ์ในรูปแบบจำลองทางสถิติหรือเทคโนโลยีปัญญาประดิษฐ์ เช่น การพยากรณ์ยอดขาย การทำนายผู้ใช้บริการในอนาคต เป็นต้น

2.1 การวิเคราะห์การถดถอย (Regression Analysis): เป็นเทคนิคทางสถิติที่ใช้สร้างความสัมพันธ์ระหว่างตัวแปรอิสระ (Independent Variable) และตัวแปรตาม (Dependent Variable) เพื่อใช้ในการพยากรณ์หรือทำนายค่าของตัวแปรตาม ซึ่งเหมาะสำหรับการวิเคราะห์ข้อมูลเชิงตัวเลขที่มีค่าไม่จำกัด เช่น ข้อมูลหุ้น

(1) Linear regression: การทำนายโดยใช้สมการเชิงเส้นเพื่อแสดงความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตาม โดยสมมติว่าความสัมพันธ์ระหว่างตัวแปรเป็นเชิงเส้นตรง

(2) Polynomial regression: การทำนายที่มีความซับซ้อนมากขึ้นโดยใช้สมการพหุนาม ใช้เมื่อความสัมพันธ์ระหว่างตัวแปรอิสระและตัวแปรตามไม่เป็นเชิงเส้น ใช้ในกรณีที่ข้อมูลมีแนวโน้มเป็นโค้ง เช่น การเติบโตของประชากร เป็นต้น

2.2 การจำแนกประเภทข้อมูล (Classification): เป็นเทคนิคในการวิเคราะห์ข้อมูลที่ใช้สำหรับการทำนายผลลัพธ์ที่มีลักษณะเป็นกลุ่มหรือค่าที่ไม่ต่อเนื่อง โดยใช้ข้อมูลในอดีตเพื่อสร้างแบบจำลองที่สามารถจำแนกข้อมูลใหม่เข้าสู่กลุ่มที่เหมาะสม เช่น การวินิจฉัยโรค

(1) Decision trees: ใช้โครงสร้างต้นไม้ในการตัดสินใจ ใช้ในการตัดสินใจโดยแบ่งข้อมูลออกเป็นลำดับขั้นตอน

(2) Logistic Regression: ใช้สำหรับการพยากรณ์ผลลัพธ์ที่เป็น สองกลุ่ม (Binary Classification) เช่น เป็นโรคหรือไม่เป็นโรค การใช้ฟังก์ชัน Sigmoid เพื่อจำแนกผลลัพธ์ระหว่าง 0 และ 1

3. ระบบข้อเสนอแนะ (Recommendation system): การวิเคราะห์เพื่อแนะนำการดำเนินการที่เหมาะสม โดยพิจารณาจากพฤติกรรมหรือข้อมูลในอดีต ตัวอย่างเช่น ระบบแนะนำสินค้าของแพลตฟอร์มอีคอมเมิร์ซ หรือระบบแนะนำเนื้อหาบนแพลตฟอร์มสตรีมมิ่ง

การวิเคราะห์ข้อมูลเป็นกระบวนการสำคัญในการเปลี่ยนข้อมูลดิบให้เป็นข้อมูลเชิงลึกที่สามารถนำไปใช้ประโยชน์ได้ ช่วยให้เราสามารถคาดการณ์แนวโน้ม กำหนดกลยุทธ์ และตัดสินใจได้อย่างแม่นยำ