

บทที่ 5

สรุปผลและข้อเสนอแนะ

การหาปัจจัยที่มีอิทธิพลต่อความสำเร็จของนักศึกษาโดยใช้เทคนิค Decision Tree (Classification) ผ่าน RapidMiner และเผยแพร่ผลลัพธ์ผ่านเว็บไซต์ พร้อมการแสดงผลด้วยกราฟและตาราง เพื่อให้ผู้ใช้งานทั่วไปสามารถเข้าถึงข้อมูลและเข้าใจแนวโน้มตลาดแรงงาน ข้อจำกัดของเว็บไซต์ ปัญหาและอุปสรรคต่างๆ ของการจัดทำโครงการ และข้อเสนอแนะเพื่อโครงการต่อไปดังนี้

5.1 บทสรุปผลโครงการ

โครงการฉบับนี้ถูกจัดทำขึ้นโดยใช้กระบวนการวิเคราะห์ข้อมูล CRISP-DM ซึ่งครอบคลุมขั้นตอนตั้งแต่การทำความเข้าใจธุรกิจ การรวบรวมและสำรวจข้อมูล การเตรียมข้อมูล การสร้างโมเดล การประเมินประสิทธิภาพ และการนำผลไปใช้งานจริง ข้อมูลที่ใช้ในการวิเคราะห์ประกอบด้วย ตำแหน่งงาน (Job Title), เงินเดือน (Salary in USD), สถานที่ทำงาน (Company Location) และปัจจัยอื่น ๆ ที่เกี่ยวข้องกับ Data Scientist ผู้วิจัยได้ทำการ ทำความสะอาดข้อมูล (Data Cleaning) และ แปลงข้อมูลให้อยู่ในรูปแบบที่เหมาะสม (Data Transformation) เช่น การจัดกลุ่มเงินเดือนเป็น High, Medium, Low การจัดกลุ่มตำแหน่งงาน และการจัดกลุ่มภูมิภาค จากนั้นได้สร้าง โมเดลการเรียนรู้ของเครื่อง (Supervised Learning) 2 เทคนิค คือ Decision Tree และ Random Forest ผ่านโปรแกรม RapidMiner เพื่อจำแนกระดับเงินเดือน ผลการประเมินประสิทธิภาพด้วยตัวชี้วัด Accuracy, Precision, Recall พบว่า Decision Tree เป็นโมเดลที่เหมาะสมที่สุด เนื่องจากมี Accuracy สูงสุดและเข้าใจง่ายสำหรับการอธิบายกฎการตัดสินใจ โมเดลที่ได้ถูกนำไปประยุกต์ใช้งานใน การวิเคราะห์ตลาดแรงงาน (Labor Market Analysis) โดยสามารถจำแนกและพยากรณ์ช่วงเงินเดือนของ Data Scientist ตามตำแหน่งงาน ระดับประสบการณ์ และภูมิภาคการทำงาน พร้อมนำเสนอผลลัพธ์ผ่าน เว็บไซต์ ในรูปแบบกราฟและตาราง เพื่อให้ผู้ใช้งานทั่วไปสามารถเข้าถึงและเข้าใจข้อมูลได้ง่าย

5.2 ข้อจำกัดเว็บไซต์

5.2.1 ผู้ใช้ไม่สามารถอัปโหลดข้อมูลการทำนายเงินเดือนตัวเองได้

5.3 ปัญหาและอุปสรรคของโครงการ

5.3.1 การรวบรวมข้อมูลบางส่วนใช้เวลานาน และข้อมูลที่ได้จากแหล่งต่าง ๆ มีความไม่สม่ำเสมอ ทำให้ต้องใช้เวลามากในการทำความสะดวกและปรับให้อยู่ในรูปแบบที่เหมาะสม

5.4.2 การเลือกวิธีการประมวลผลและการปรับพารามิเตอร์ของโมเดลเป็นเรื่องที่ซับซ้อน หากปรับไม่เหมาะสมอาจทำให้ผลลัพธ์ไม่แม่นยำ

5.4.3 ข้อมูลไม่เหมาะกับการจัดกลุ่ม

5.4.4 ไม่มีความรู้เรื่องโมเดล ทำให้ต้องใช้เวลาในการศึกษาเพิ่มเติม

5.4 ข้อเสนอแนะ

5.4.2 ควรมีการเก็บรวบรวมข้อมูลเพิ่มเติมและอัปเดตข้อมูลใหม่ ๆ อยู่เสมอ เพื่อให้โมเดลสามารถเรียนรู้จากข้อมูลล่าสุดและเพิ่มความแม่นยำ

5.4.2 ควรพัฒนาเว็บไซต์ให้มีระบบที่ผู้ใช้งานสามารถอัปโหลดข้อมูลใหม่เข้ามาและให้โมเดลทำการทำนายได้อัตโนมัติ เพื่อเพิ่มความสะดวกในการใช้งาน

5.4.3 ควรศึกษาข้อมูลและกระบวนการสร้างโมเดลอย่างละเอียด เพื่อป้องกันความผิดพลาดจากข้อมูลที่ไม่สมบูรณ์หรือโมเดลที่ไม่เหมาะสม