

บทที่ 5

สรุป และข้อเสนอแนะ

จากผลการวิเคราะห์ข้อมูลในโครงการเรื่อง การวิเคราะห์ปัจจัยที่ส่งผลต่อความล่าช้าของสายการบิน รวมถึงการพัฒนาเว็บไซต์ AirAnalytica เพื่อเผยแพร่ผลการวิเคราะห์ข้อมูล ผู้จัดทำได้ดำเนินการ สรุปผลการดำเนินโครงการ วิเคราะห์ข้อจำกัด ปัญหา และอุปสรรคที่พบตลอดกระบวนการทำงาน พร้อมทั้งเสนอแนะแนวทางในการปรับปรุงและพัฒนาโครงการให้มีประสิทธิภาพยิ่งขึ้นและสามารถนำไปประยุกต์ใช้ได้จริงในอนาคต ดังรายละเอียดต่อไปนี้

5.1 บทสรุปผลโครงการ

ผู้วิเคราะห์ข้อมูลได้ดำเนินโครงการเรื่อง การวิเคราะห์ปัจจัยที่ส่งผลต่อความล่าช้าของสายการบิน โดยใช้ชุดข้อมูล 2015 Flight Delays and Cancellations จากเว็บไซต์ Kaggle ซึ่งมีจำนวนกว่า 5,819,079 แถว และ 31 แอตทริบิวต์ เพื่อศึกษาปัจจัยที่ส่งผลต่อการล่าช้าในการมาถึงของเที่ยวบิน (Arrival Delay) ด้วยเทคนิคการทำเหมืองข้อมูล (Data Mining) และการเรียนรู้ของเครื่อง (Machine Learning)

การดำเนินการใช้กระบวนการ CRISP-DM ครอบคลุมตั้งแต่การทำความเข้าใจธุรกิจ การเตรียมข้อมูล การสร้างโมเดล และการประเมินผล โดยแบ่งการวิเคราะห์ออกเป็น 2 ส่วนหลัก ได้แก่

5.1.1 การจำแนกสถานะเที่ยวบิน (Classification) ใช้เทคนิค Decision Tree และ Random Forest เพื่อจำแนกสถานะเที่ยวบินออกเป็น 3 กลุ่ม ได้แก่ ON TIME, SHORT DELAY และ DELAY โดยโมเดล Decision Tree ให้ค่าความแม่นยำ (Accuracy) สูงสุดที่ $90.01\% \pm 0.27\%$ ซึ่งสามารถอธิบายความสัมพันธ์ระหว่างตัวแปรได้อย่างชัดเจน

5.1.2 การพยากรณ์ค่าความล่าช้า (Regression) ใช้เทคนิค Multiple Linear Regression เพื่อพยากรณ์ค่าความล่าช้าเชิงเวลา โดยเลือกใช้ตัวแปรที่มีค่าสหสัมพันธ์ (Correlation ≥ 0.4) ทั้งหมด 14 ตัว เช่น DEPARTURE_DELAY, AIRLINE_DELAY และ LATE_AIRCRAFT_DELAY ผลการประเมินแบบจำลองพบว่า $R^2 = 0.873$, MAE = 7.2 นาที และ RMSE = 9.875 นาที แสดงถึงความแม่นยำในระดับสูง

ผลลัพธ์ทั้งหมดได้ถูกนำเสนอผ่านเว็บไซต์ AirAnalytica ซึ่งพัฒนาโดยใช้ภาษา HTML, CSS และ JavaScript เพื่อให้ผู้ใช้งานสามารถเข้าถึงข้อมูล วิเคราะห์แนวโน้ม และพยากรณ์ความล่าช้าได้แบบโต้ตอบในลักษณะ Interactive Dashboard โดยมีหน้า Home, Dashboard, Survey, Result และ About ที่ออกแบบให้ใช้งานได้สะดวกและสอดคล้องกับวัตถุประสงค์ของโครงการ

โดยสรุป โครงการนี้สามารถสร้างแบบจำลองการทำนายความล่าช้าได้อย่างมีประสิทธิภาพ และนำเสนอผลในรูปแบบเว็บไซต์ที่เป็นมิตรต่อผู้ใช้ ช่วยให้หน่วยงานด้านการบินสามารถนำข้อมูลไปประยุกต์ใช้ในการปรับปรุงประสิทธิภาพการบริหารจัดการเที่ยวบินได้จริง

5.2 ข้อจำกัดของระบบ

5.2.1 เว็บไซต์ AirAnalytica เป็นเพียงระบบจำลองสำหรับนำเสนอผลการวิเคราะห์ ผู้ใช้ทั่วไปไม่สามารถปรับเปลี่ยนรูปแบบของแดชบอร์ดหรือโมเดลได้โดยตรง

5.2.2 ข้อมูลที่ใช้วิเคราะห์มาจากปี 2015 ซึ่งไม่ได้อัปเดตแบบเรียลไทม์ ทำให้ผลลัพธ์ไม่สะท้อนสถานการณ์การบินปัจจุบัน

5.2.3 ระบบไม่รองรับการอัปเดตข้อมูลใหม่โดยผู้ใช้งานทั่วไป การปรับปรุงข้อมูลต้องทำผ่านผู้ดูแลระบบ (Admin) เท่านั้น

5.2.4 การแสดงผลในบางอุปกรณ์ เช่น โทรศัพท์มือถือ อาจมีข้อจำกัดในการแสดงกราฟหรือแผนภูมิที่มีรายละเอียดมาก

5.3 ปัญหาและอุปสรรคของโครงการ

5.3.1 ชุดข้อมูลที่ใช้ในการวิเคราะห์เป็นข้อมูลทุติยภูมิจาก Kaggle ผู้วิเคราะห์ไม่สามารถควบคุมความครบถ้วนหรือความถูกต้องของข้อมูลได้ทั้งหมด

5.3.2 การจัดการข้อมูลขนาดใหญ่ (มากกว่า 5 ล้านแถว) ทำให้ต้องใช้กระบวนการสุ่มข้อมูล (Sampling 5%) เพื่อให้ประมวลผลได้ในเวลาที่เหมาะสม

5.3.3 การปรับสมดุลของข้อมูลบางคลาส เช่น SHORT DELAY ที่มีสัดส่วนเพียง 2% ของข้อมูลทั้งหมด ทำให้การจำแนกคลาสนี้มีความคลาดเคลื่อนสูง

5.3.4 การสร้างแบบจำลองต้องใช้เวลาในการปรับแต่งพารามิเตอร์และเลือกตัวแปรอย่างรอบคอบ เพื่อป้องกันปัญหา Overfitting หรือความเอนเอียงของโมเดล

5.3.5 ในขั้นตอนการพัฒนาเว็บไซต์ ต้องเชื่อมต่อผลลัพธ์การวิเคราะห์เข้ากับระบบหน้าเว็บซึ่งใช้หลายเทคโนโลยีร่วมกัน เช่น HTML, CSS, JavaScript ทำให้ต้องตรวจสอบความเข้ากันของโค้ดและการแสดงผลในเบราว์เซอร์หลายรูปแบบ

5.4 ข้อเสนอแนะ

5.4.1 ควรปรับปรุงชุดข้อมูลให้มีความทันสมัย โดยใช้ข้อมูลเที่ยวบินในช่วงปีปัจจุบัน และเพิ่มตัวแปรด้านสภาพอากาศหรือสภาพจราจรทางอากาศ เพื่อให้แบบจำลองสะท้อนสถานการณ์จริงมากขึ้น

5.4.2 ควรนำเทคนิค Machine Learning ขั้นสูง เช่น Gradient Boosting, XGBoost หรือ Neural Network มาทดลองเปรียบเทียบเพื่อเพิ่มความแม่นยำของโมเดล

5.4.3 ควรพัฒนาเว็บไซต์ AirAnalytica ให้สามารถเชื่อมต่อกับฐานข้อมูลแบบออนไลน์ เพื่ออัปเดตข้อมูลและผลการพยากรณ์แบบเรียลไทม์

5.4.4 ควรเพิ่มระบบแจ้งเตือนหรือบันทึกประวัติการอัปเดตข้อมูล เพื่อให้ผู้ใช้งานทราบถึงการปรับปรุงผลการวิเคราะห์แต่ละครั้ง

5.4.5 ควรเปิดให้หน่วยงานที่เกี่ยวข้อง เช่น สนามบินหรือสายการบิน เข้ามาใช้งานระบบร่วมกัน เพื่อสร้างฐานข้อมูลที่มีความหลากหลายและสนับสนุนการตัดสินใจเชิงนโยบายได้มากขึ้น