



การวิเคราะห์ระดับความรุนแรงของอุบัติเหตุทางถนนโดยใช้การเรียนรู้ของเครื่อง

นายสุรัชย์ เดชมี

การค้นคว้าอิสระนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

วิทยาศาสตรมหาบัณฑิต

สาขาวิชาวิทยาศาสตร์ข้อมูลเพื่อนวัตกรรม

บัณฑิตวิทยาลัย มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

ปีการศึกษา 2567

ลิขสิทธิ์ของมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

การวิเคราะห์ระดับความรุนแรงของอุบัติเหตุทางถนนโดยใช้การเรียนรู้ของเครื่อง



การค้นคว้าอิสระนี้เป็นส่วนหนึ่งของการศึกษาหลักสูตร

วิทยาศาสตรมหาบัณฑิต

สาขาวิชาวิทยาศาสตรข้อมูลเพื่อนวัตกรรม

บัณฑิตวิทยาลัย มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

ปีการศึกษา 2567

ลิขสิทธิ์ของมหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ



ใบรับรองการค้นคว้าอิสระ

บัณฑิตวิทยาลัย มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

เรื่อง การวิเคราะห์ระดับความรุนแรงของอุบัติเหตุทางถนนโดยใช้การเรียนรู้ของเครื่อง

โดย นายสุรัชย์ เดชมี

ได้รับอนุมัติให้นับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชา
วิทยาศาสตร์ข้อมูลเพื่อนวัตกรรม

..... คณบดีบัณฑิตวิทยาลัย / หัวหน้าภาควิชา
(ผู้ช่วยศาสตราจารย์ ดร.สุนันทา สดสี)

คณะกรรมการสอบการค้นคว้าอิสระ

.....
(ดร.ภัทรารุณี แสงศิริ)

ประธานกรรมการ

.....
(ผู้ช่วยศาสตราจารย์ ดร.ผุสดี บุญรอด)

อาจารย์ที่ปรึกษา

.....
(ผู้ช่วยศาสตราจารย์ ดร.ศิพาณิชย์ นุชิตประสิทธิ์ชัย)

กรรมการ

ชื่อ : นายสุรัชย์ เดชมี
ชื่อการค้นคว้าอิสระ : การวิเคราะห์ระดับความรุนแรงของอุบัติเหตุทางถนนโดยใช้
การเรียนรู้ของเครื่อง
สาขาวิชา : วิทยาศาสตร์ข้อมูลเพื่อนวัตกรรม
มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
อาจารย์ที่ปรึกษาการค้นคว้าอิสระหลัก : ผู้ช่วยศาสตราจารย์ ดร.ผุสดี บุญรอด
ปีการศึกษา : 2567

บทคัดย่อ

อุบัติเหตุทางถนนเป็นปัญหาที่สำคัญในประเทศไทย ซึ่งมีอัตราการเกิดเหตุและเสียชีวิตที่สูงในทุกปี ทำให้เกิดการสูญเสียต่อชีวิตและทรัพย์สินเป็นจำนวนมาก การค้นคว้าอิสระนี้จึงมีวัตถุประสงค์เพื่อพัฒนาแบบจำลองวิเคราะห์ระดับความรุนแรงของอุบัติเหตุทางถนน และพัฒนาระบบสารสนเทศสำหรับวิเคราะห์อุบัติเหตุทางถนน ซึ่งใช้ชุดข้อมูลอุบัติเหตุของกระทรวงคมนาคม ตั้งแต่ปี พ.ศ. 2562 - 2567 ทำการพัฒนาแบบจำลองโดยประยุกต์ใช้การเรียนรู้ของเครื่อง ได้แก่ ต้นไม้สินใจ (Decision Tree) ป่าสุ่ม (Random Forest) เอ็กซ์ ตริ้ม เกรเดียนต์ บูสต์ติง (Extreme Gradient Boosting: XGBoost) และไลต์ เกรเดียนต์ บูสต์ติง แมชชีน (Light Gradient Boosting Machine: LightGBM) และทำการเพิ่มประสิทธิภาพแบบจำลองโดยเพิ่มข้อมูลที่ไม่สมดุล (Resampling) จากผลการวิจัยพบว่าแบบจำลองเอ็กซ์ ตริ้ม เกรเดียนต์ บูสต์ติง ที่ผ่านการเพิ่มประสิทธิภาพแบบการปรับน้ำหนักของคลาสที่ 1 ให้ค่าความถูกต้อง (Accuracy) 85% ค่าความแม่นยำ (Precision) 68% ค่าความระลึก (Recall) 67% และค่าคะแนนเอฟหนึ่ง (F1-score) 68% ในส่วนการคัดเลือกปัจจัยที่ส่งผลต่อความรุนแรงโดยใช้ แชปลีย์ แอ็ดดิทิฟเอ็กซ์พลานเนชันส์ (SHapley Additive exPlanations: SHAP) พบว่า ยานพาหนะที่เกิดเหตุประเภทรถมอเตอร์ไซค์ ส่งผลต่อระดับความรุนแรงของอุบัติเหตุทางถนนมากที่สุด นอกจากนี้ได้นำแบบจำลองที่พัฒนาขึ้นนำไปต่อยอดพัฒนาระบบสารสนเทศสำหรับการแสดงผลข้อมูลอุบัติเหตุทางถนน โดยจากการประเมินคุณภาพระบบสารสนเทศของผู้เชี่ยวชาญ และการประเมินความพึงพอใจที่มีต่อระบบสารสนเทศของผู้ใช้งานทั่วไป พบว่าภาพรวมอยู่ในระดับมาก ซึ่งสามารถสรุปได้ว่าระบบสารสนเทศที่พัฒนาขึ้นรองรับความต้องการของผู้ใช้งาน และเกิดประโยชน์สามารถนำไปใช้งานในด้านต่าง ๆ ตามความต้องการของผู้ใช้งานได้

(มีจำนวนทั้งสิ้น 94 หน้า)

คำสำคัญ : อุบัติเหตุทางถนน ระดับความรุนแรง การเรียนรู้ของเครื่อง

อาจารย์ที่ปรึกษาการค้นคว้าอิสระหลัก

Name : Mr. Surachai Detmee
Independent Study Title : An Analysis of Road Accident Severity Using Machine Learning
Major Field : Data Science for Innovation
King Mongkut's University of Technology North
Bangkok
Independent Study Advisor : Assistant Professor Dr. Pudsadee Boonrawd
Academic Year : 2024

ABSTRACT

Road accidents are a significant problem in Thailand, with consistently high rates of incidents and fatalities each year, leading to substantial loss of life and property. This independent study aims to develop a model for analyzing the severity levels of road accidents and to create an information system for road accident analysis. The study utilizes accident data from the Ministry of Transport between 2019 and 2024. The model is developed using machine learning techniques, including Decision Tree, Random Forest, Extreme Gradient Boosting (XGBoost), and Light Gradient Boosting Machine (LightGBM), with enhanced model performance through resampling to address data imbalance. The research findings indicate that the optimized XGBoost model, which includes class weight adjustment for class 1, achieved an accuracy of 85%, precision of 68%, recall of 67%, and an F1-score of 68%. Feature importance analysis using SHapley Additive exPlanations (SHAP) revealed that motorcycle-related accidents had the highest impact on the severity of road accidents. Furthermore, the developed model was integrated into an information system for visualizing road accident data. Evaluation by experts on system quality and satisfaction assessment from general users indicated that the system was rated highly overall. It can be concluded that the developed information system meets user requirements and is beneficial for various applications as needed.

(Total 94 Pages)

Keywords: Road Accidents, Severity Levels, Machine Learning

Advisor

กิตติกรรมประกาศ

กิตติกรรมประกาศฉบับนี้ข้าพเจ้าจัดทำขึ้นเพื่อขอกราบขอบพระคุณแต่ผู้มีส่วนที่เกี่ยวข้องที่ให้คำแนะนำ ช่วยเหลือ และให้คำปรึกษาในการจัดทำการค้นคว้าอิสระในครั้งนี้ โดยขอขอบคุณอาจารย์ที่ปรึกษา ผู้ช่วยศาสตราจารย์ ดร.ผุสดี บุญรอด ที่ให้คำแนะนำ องค์ความรู้ และแนวทางในการค้นคว้าอิสระครั้งนี้ ด้วยความเอาใจใส่และทุ่มเทเป็นอย่างยิ่ง ซึ่งช่วยให้ข้าพเจ้าสามารถดำเนินการค้นคว้าอิสระจนสำเร็จลุล่วงไปได้ด้วยดี

ขอขอบคุณ คณะกรรมการสอบและอาจารย์ทุกท่าน ที่ให้คำแนะนำอันเป็นประโยชน์ต่อการปรับปรุงงานการค้นคว้าอิสระให้มีคุณภาพยิ่งขึ้นอนาคต โดยมี ดร.ภัทราวดี แสงศิริ ประธานกรรมการ ผู้ช่วยศาสตราจารย์ ดร.ผุสดี บุญรอด กรรมการและอาจารย์ที่ปรึกษา และผู้ช่วยศาสตราจารย์ ดร.ศิพาณิชย์ นุชิตประสิทธิ์ชัย กรรมการ รวมถึง ผู้เชี่ยวชาญที่ให้ความอนุเคราะห์ในการประเมินระบบสารสนเทศ และผู้เข้าร่วมการประเมินความพึงพอใจทุกท่าน ซึ่งมีส่วนสำคัญต่อการพัฒนาระบบให้ตรงกับวัตถุประสงค์ของการค้นคว้าอิสระ และตรงตามการใช้งาน

ขอขอบคุณ ผู้ให้ข้อมูลและหน่วยงานที่เกี่ยวข้อง คือศูนย์เทคโนโลยีสารสนเทศและการสื่อสาร สำนักงานปลัดกระทรวงคมนาคม กระทรวงคมนาคม ประเทศไทย ที่เอื้อเฟื้อข้อมูลด้านอุบัติเหตุทางถนน ซึ่งเป็นข้อมูลสำคัญต่อการวิเคราะห์และพัฒนาแบบจำลอง อีกทั้งบุคลากรของศูนย์เทคโนโลยีสารสนเทศและการสื่อสาร ยังให้คำแนะนำต่าง ๆ ที่มีประโยชน์เป็นอย่างยิ่งต่อการค้นคว้าอิสระในครั้งนี้

สุดท้ายนี้ ข้าพเจ้าขอขอบคุณ ครอบครัวและเพื่อน ที่คอยให้กำลังใจ สนับสนุน และเป็นแรงผลักดันสำคัญที่ช่วยให้ข้าพเจ้ามีความมุ่งมั่นจนสามารถดำเนินการศึกษาวิจัยนี้จนสำเร็จลุล่วงไปได้ด้วยดีข้าพเจ้าขอมอบผลของความสำเร็จในครั้งนี้ให้แก่ทุกท่านด้วยความซาบซึ้งเป็นอย่างยิ่ง

สุรัชย์ เดชมี

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	ง
บทคัดย่อภาษาอังกฤษ	จ
กิตติกรรมประกาศ	ฉ
สารบัญ	ช
สารบัญตาราง	ฌ
สารบัญภาพ	ฎ
บทที่ 1 บทนำ	1
1.1 ความเป็นมาและความสำคัญของปัญหา	1
1.2 วัตถุประสงค์การวิจัย	3
1.3 ขอบเขตการวิจัย	3
1.4 นิยามศัพท์เฉพาะ	4
1.5 ประโยชน์ที่ได้รับ	5
บทที่ 2 เอกสารและงานวิจัยที่เกี่ยวข้อง	6
2.1 อุบัติเหตุทางถนน	6
2.2 การเรียนรู้ของเครื่อง	7
2.3 การวัดประสิทธิภาพแบบจำลอง	11
2.4 งานวิจัยที่เกี่ยวข้อง	13
บทที่ 3 วิธีการดำเนินวิจัย	20
3.1 การศึกษาและวิเคราะห์ปัญหาที่เกี่ยวข้อง	22
3.2 การเก็บและรวบรวมข้อมูล	22
3.3 การพัฒนาแบบจำลอง	26
3.4 การวัดประสิทธิภาพแบบจำลอง	30
3.5 การพัฒนาระบบสารสนเทศ	31
3.6 การประเมินคุณภาพและความพึงพอใจระบบสารสนเทศ	35
บทที่ 4 ผลการดำเนินงานวิจัย	38
4.1 ผลการพัฒนาแบบจำลอง	38
4.2 ผลการพัฒนาระบบสารสนเทศ	46

สารบัญ (ต่อ)

	หน้า
4.3 ผลการประเมินคุณภาพและผลการประเมินความพึงพอใจ ต่อระบบสารสนเทศ	48
บทที่ 5 สรุปผล อภิปรายผล และข้อเสนอแนะการวิจัย	55
5.1 สรุปผลการวิจัย	55
5.2 อภิปรายผลการวิจัย	58
5.3 ข้อเสนอแนะการวิจัย	59
บรรณานุกรม	61
ภาคผนวก ก รายชื่อผู้เชี่ยวชาญ และหนังสือขอความอนุเคราะห์เป็นผู้เชี่ยวชาญ	65
ภาคผนวก ข แบบประเมินคุณภาพระบบสารสนเทศสำหรับผู้เชี่ยวชาญระบบ สารสนเทศการแสดงผลข้อมูลอุบัติเหตุทางถนนในประเทศไทย	71
ภาคผนวก ค แบบประเมินความพึงพอใจระบบสารสนเทศสำหรับผู้ใช้งานทั่วไประบบ สารสนเทศการแสดงผลข้อมูลอุบัติเหตุทางถนนในประเทศไทย	79
ภาคผนวก ง คู่มือการใช้งานระบบสารสนเทศ	87
ประวัติผู้เขียน	91

สารบัญตาราง

ตารางที่	หน้า
2-1 งานวิจัยที่เกี่ยวข้องกับการพัฒนาแบบจำลอง	16
3-1 คุณลักษณะข้อมูลอุบัติเหตุทางถนนปี พ.ศ. 2562 – 2567	22
3-2 การทำงานของผู้ใช้งานระบบสารสนเทศส่วนการค้นหาข้อมูล	32
3-3 รายละเอียดการทำงานของระบบสารสนเทศในการแสดงผลข้อมูล	34
3-4 เกณฑ์การให้คะแนนแบบประเมินคุณภาพของผู้เชี่ยวชาญ	36
3-5 เกณฑ์การให้คะแนนแบบประเมินความพึงพอใจผู้ใช้งานทั่วไป	36
4-1 ผลการพัฒนาแบบจำลองแบบไม่เพิ่มประสิทธิภาพ	38
4-2 ผลการพัฒนาแบบจำลองปรับปรุงประสิทธิภาพโดย SMOTE	42
4-3 ผลการพัฒนาแบบจำลองปรับปรุงประสิทธิภาพโดย SMOTETomek	43
4-4 ผลการพัฒนาแบบจำลองปรับปรุงประสิทธิภาพโดยการปรับน้ำหนัก	44
4-5 ผลการประเมินคุณภาพด้านความสามารถของระบบสารสนเทศตามความต้องการของผู้ใช้งาน	48
4-6 ผลการประเมินคุณภาพของผู้เชี่ยวชาญด้านการทำงานตามฟังก์ชัน	49
4-7 ผลการประเมินคุณภาพของผู้เชี่ยวชาญด้านความสะดวกและง่ายต่อการใช้งานระบบสารสนเทศ	50
4-8 สรุปผลการประเมินคุณภาพระบบสารสนเทศโดยผู้เชี่ยวชาญทั้ง 3 ด้าน	51
4-9 ผลการประเมินความพึงพอใจด้านความสามารถของระบบสารสนเทศตามความต้องการของผู้ใช้งาน	52
4-10 ผลการประเมินความพึงพอใจด้านการทดสอบการทำงานตามฟังก์ชัน	53
4-11 ผลการประเมินความพึงพอใจด้านความสะดวกและง่ายต่อการใช้งานระบบ	54
4-12 สรุปผลการประเมินความพึงพอใจของระบบสารสนเทศโดยผู้ใช้งานทั่วไปทั้ง 3 ด้าน	55
ข-1 แบบประเมินคุณภาพด้านความสามารถของระบบสารสนเทศตามความต้องการของผู้ใช้งาน	78
ข-2 การประเมินความพึงพอใจด้านความสะดวกและง่ายต่อการใช้งานระบบ	79
ข-3 แบบประเมินคุณภาพด้านความสะดวกและง่ายต่อการใช้งานระบบสารสนเทศ	80

สารบัญตาราง (ต่อ)

ตารางที่	หน้า
ค-1 การประเมินความพึงพอใจด้านความสามารถของระบบสารสนเทศ ตามความต้องการของผู้ใช้งาน	86
ค-2 การประเมินความพึงพอใจด้านการทดสอบการทำงานตามฟังก์ชันงาน	86
ค-3 การประเมินความพึงพอใจด้านความสะดวกและง่ายต่อการใช้งานระบบ	87



สารบัญรูปภาพ

ภาพที่	หน้า
1-1 จำนวนผู้เสียชีวิตทั่วโลกตามรายงานขององค์การอนามัยโลก	1
2-1 กระบวนการทำงานการเรียนรู้ของเครื่องแบบมีผู้สอน	8
2-2 กระบวนการทำงานการเรียนรู้ของเครื่องแบบไม่มีผู้สอน	9
2-3 กระบวนการทำงานการเรียนรู้ของเครื่องแบบเสริมกำลัง	10
2-4 รูปแบบตารางสับสน	11
3-1 กรอบแนวคิดการวิจัย	21
3-2 การแปลงข้อมูลแบบ One-Hot Encoding	25
3-3 จำนวนชุดข้อมูลฝึกและชุดข้อมูลทดสอบ	26
3-4 จำนวนชุดข้อมูลฝึกในการแบ่งออกแต่ละคลาสในระดับความรุนแรง	27
3-5 สร้างแบบจำลองด้วยต้นไม้ตัดสินใจด้วย Google Colab	28
3-6 การสร้างแบบจำลองด้วยป่าการสุ่มใน Google Colab	29
3-7 การสร้างแบบจำลองด้วยเอนกซ์ตรึม เกรเดียนต์ บูสตั้ง ใน Google Colab	29
3-8 การสร้างแบบจำลองด้วยไลท์ เกรเดียนต์ บูสตั้ง แมชชีน ใน Google Colab	30
3-9 แผนภาพการทำงานผู้ไ้ระบบสารสนเทศการแสดงผลข้อมูลอุบัติเหตุ ในประเทศไทย	32
3-10 ผังแสดงเหตุการณ์การค้นหาข้อมูลอุบัติเหตุในระบบสารสนเทศ	33
3-11 ผังแสดงเหตุการณ์การการแสดงผลข้อมูลอุบัติเหตุ	34
3-12 ออกแบบหน้าจแสดงผลข้อมูลระบบข้อมูลอุบัติเหตุในประเทศไทย	35
4-1 ตารางสับสนของต้นไม้ตัดสินใจ	39
4-2 ตารางสับสนของป่าการสุ่ม	40
4-3 ตารางสับสนของเอนกซ์ตรึม เกรเดียนต์ บูสตั้ง	40
4-4 ตารางสับสนของไลท์ เกรเดียนต์ บูสตั้ง แมชชีน	41
4-5 ปัจจัยที่ส่งผลต่อความรุนแรงด้วย SHAP	46
4-6 หน้าจแสดงผลระบบสารสนเทศส่วนการค้นหาข้อมูล	47
4-7 แผนที่การแสดงจุดเกิดเหตุและแผนภูมิเปรียบเทียบข้อมูลอุบัติเหตุรายปี	47
4-8 ข้อมูลการแสดงผลอุบัติเหตุแยกตามประเภทข้อมูล	48

สารบัญรูปภาพ (ต่อ)

ภาพที่	หน้า
4-9 ผลการประเมินคุณภาพระบบสารสนเทศโดยผู้เชี่ยวชาญในภาพรวม ทั้ง 3 ด้าน	52
4-10 ผลการประเมินความพึงพอใจระบบสารสนเทศโดยผู้ใช้งานทั่วไปในภาพรวม ทั้ง 3 ด้าน	56
ง-1 ระบบสารสนเทศแสดงข้อมูลทั้งหมด	90
ง-2 ส่วนการค้นหาข้อมูล	91
ง-3 ข้อมูลอุบัติเหตุหลังจากระบุเงื่อนไขการค้นหา	92
ง-4 ข้อมูลการเกิดเหตุบนแผนที่	93



บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

อุบัติเหตุทางถนนเป็นปัญหาที่รุนแรงและส่งผลกระทบทั่วโลก จากรายงาน Global Status Report on Road Safety 2023 ขององค์การอนามัยโลก (World Health Organization: WHO) พบว่าในปี พ.ศ. 2564 มีผู้เสียชีวิตจากอุบัติเหตุทางถนนมากถึง 1.19 ล้านคน [1] เมื่อเทียบกับปีก่อนหน้าตามภาพที่ 1-1



ภาพที่ 1- 1 จำนวนผู้เสียชีวิตทั่วโลกตามรายงานขององค์การอนามัยโลก

จากภาพที่ 1-1 พบว่าจำนวนผู้เสียชีวิตในปี พ.ศ. 2564 มีจำนวนผู้เสียชีวิตน้อยที่สุดอยู่ที่ 1.19 ล้านคนเมื่อเทียบกับจำนวนผู้เสียชีวิตในปีก่อนหน้า แต่จำนวนดังกล่าวยังคงสูงและเป็นหนึ่งในปัญหาด้านความปลอดภัยของโลก โดยเฉพาะอย่างยิ่งในกลุ่มเยาวชนอายุตั้งแต่ 5 - 29 ปี ซึ่งเป็นวัยที่มีผู้เสียชีวิตสูงสุด โดยการเสียชีวิตจากอุบัติเหตุทางถนนไม่เพียงแต่ส่งผลกระทบโดยตรงต่อครอบครัวของผู้ประสบเหตุ แต่ยังลดศักยภาพในการพัฒนาทรัพยากรมนุษย์ของประเทศนั้น ๆ ซึ่งปัจจัยที่ส่งผลต่ออัตราการเสียชีวิตในหลายประเทศเกิดจากโครงสร้างพื้นฐานที่ไม่ปลอดภัย ขาดมาตรการความปลอดภัยที่มีประสิทธิภาพ และการบังคับใช้กฎหมายที่ยังไม่มีความเข้มงวดเพียงพอ นอกจากนี้การเพิ่มขึ้นของจำนวนยานพาหนะบนถนนไม่ว่าจะเป็น รถยนต์ รถมอเตอร์ไซด์ หรือรถบรรทุกก็ตามโดยมี

มาตรการความปลอดภัยความปลอดภัยที่เป็นการรองรับการเพิ่มขึ้นของรถบนถนนยังไม่เพียงพอ ในการลดการเกิดอุบัติเหตุ ซึ่งเป็นอีกหนึ่งปัจจัยสำคัญที่ทำให้อัตราการเสียชีวิตจากอุบัติเหตุทางถนน ยังคงอยู่ในระดับสูง นอกจากนี้เมื่อพิจารณาถึงสาเหตุที่ทำให้เกิดผู้เสียชีวิต พบว่าผู้ใช้รถจักรยานยนต์ ยานยนต์ที่มีสองล้อและสามล้อเป็นกลุ่มที่มีผู้เสียชีวิตจากอุบัติเหตุทางถนนมากที่สุด โดยมีสัดส่วนอยู่ที่ร้อยละ 30 ของจำนวนผู้เสียชีวิตทั้งหมด รองลงมาคือผู้โดยสารในรถยนต์ ร้อยละ 25 และคนเดินเท้า ร้อยละ 21 [1] จากการสูญเสียดังกล่าวทำให้หลายประเทศทั่วโลกตั้งเป้าหมายที่จะลดจำนวนการเกิดอุบัติเหตุและผู้เสียชีวิตให้มีจำนวนที่ลดลง ซึ่งรวมไปถึงการใช้มาตรการทางด้านกฎหมาย และการนำเทคโนโลยีเข้ามาช่วยก็ตาม

ประเทศไทยยังคงเป็นหนึ่งในประเทศที่มีอัตราการเกิดอุบัติเหตุอยู่ในระดับที่สูงต่อเนื่อง จากรายงานขององค์การอนามัยโลกพบว่า ในปี พ.ศ. 2556 ประเทศไทยมีผู้เสียชีวิตจากอุบัติเหตุทางถนนจำนวน 26,312 คน ซึ่งมากเป็นอันดับที่ 3 ของโลก แม้ในปี พ.ศ. 2558 จะมีจำนวนผู้เสียชีวิตลดลงเป็น 24,237 คน แต่ประเทศไทยยังคงอยู่ในอันดับที่ 2 ของโลก และในปี พ.ศ. 2561 แม้จะลดลงเหลือ 22,491 คน แต่อันดับยังสูงถึงอันดับที่ 9 ของโลก และในปี พ.ศ. 2566 ประเทศไทยมีผู้เสียชีวิตทางอุบัติเหตุจำนวน 18,218 คน จากจำนวนผู้เสียชีวิตดังกล่าวประเทศไทยยังเป็นประเทศที่มีผู้เสียชีวิตจากอุบัติเหตุทางถนนที่สูง นอกจากนี้รายงานสถานการณ์อุบัติเหตุทางถนนประจำปี 2567 กรมควบคุมโรค กระทรวงสาธารณสุข ได้รายงานผู้เสียชีวิตจากอุบัติเหตุตั้งแต่ปี พ.ศ. 2561 – 2565 ว่าตลอด 5 ปี มีประชากรประเทศไทยเสียชีวิตจากอุบัติเหตุเฉลี่ยปีละ 18,400 รายต่อปี ซึ่งช่วงอายุที่มีอัตราการเสียชีวิตมากที่สุดคือ 20 – 24 ปี [2] โดยอุบัติเหตุส่วนใหญ่ที่มีผู้เสียชีวิตเกิดจากการใช้รถจักรยานยนต์ และมีสาเหตุหลักอื่น ๆ ที่เกี่ยวข้องได้แก่ การขับเร็วเกินกำหนด การดื่มเครื่องดื่มแอลกอฮอล์ขณะขับขี่ และการไม่ปฏิบัติตามกฎจราจร

จากข้อมูลอุบัติเหตุทางถนนข้างต้นหน่วยงานต่าง ๆ ที่เกี่ยวข้องทั้งประเทศไทยและต่างประเทศได้นำเทคโนโลยีต่าง ๆ มาปรับใช้เพื่อเพิ่มความปลอดภัยทางถนน เช่น ระบบช่วยเหลือผู้ขับขี่ขั้นสูง (Advanced Driver Assistance Systems: ADAS) มีความสามารถในการตรวจจับสิ่งกีดขวาง เตือนผู้ขับขี่เมื่อมีความเสี่ยงในการเกิดอุบัติเหตุการเบรกอัตโนมัติเพื่อลดการชน และการพัฒนาระบบเฝ้าระวังการขับขี่และการตรวจจับการกระทำผิดกฎจราจรเพื่อควบคุมและลดการเกิดอุบัติเหตุ อีกทั้งยังใช้การวิเคราะห์ข้อมูลจากอุบัติเหตุเพื่อศึกษาถึงปัญหาที่เกิดขึ้นโดยการนำเทคโนโลยีมาช่วยในการวิเคราะห์เช่น การนำระบบตามกฎเกณฑ์ (Rule-based Systems) มาช่วยในการตรวจจับการกระทำผิดกฎจราจรหรือการประเมินความเสี่ยงเบื้องต้น ทำให้สามารถควบคุมสถานการณ์ได้อย่างรวดเร็วและมีประสิทธิภาพ การนำการวิเคราะห์อนุกรมเวลา (Time Series Analysis) มาช่วยในการวิเคราะห์ข้อมูลอุบัติเหตุ โดยใช้ในการตรวจจับรูปแบบจากข้อมูลการเกิดอุบัติเหตุที่สะสมมาในช่วงเวลาต่าง ๆ เพื่อคาดการณ์ความเป็นไปได้ของการเกิดอุบัติเหตุในอนาคต [3] และการเรียนรู้ของ

เครื่อง (Machine Learning) ซึ่งเป็นเทคโนโลยีที่สามารถวิเคราะห์ข้อมูลจำนวนมากศาสตร์ที่เกี่ยวข้องกับอุบัติเหตุทางถนน ด้วยการนำข้อมูลที่ซับซ้อนมาวิเคราะห์เพื่อระบุปัจจัยเสี่ยงและคาดการณ์โอกาสเกิดอุบัติเหตุในอนาคตอย่างแม่นยำ สามารถช่วยในการประมวลผลข้อมูลจากอุบัติเหตุในอดีต และสามารถระบุความเสี่ยง ความรุนแรงของอุบัติเหตุได้อย่างมีประสิทธิภาพ

จากประเด็นดังกล่าวข้างต้น การค้นคว้าอิสระนี้จึงมุ่งเน้นพัฒนาแบบจำลองการวิเคราะห์ระดับความรุนแรงของอุบัติเหตุทางถนน โดยใช้การเรียนรู้ของเครื่องสร้างแบบจำลองที่เหมาะสมในการวิเคราะห์ระดับความรุนแรงจากข้อมูลอุบัติเหตุทางถนนในประเทศไทย และพัฒนาเว็บแอปพลิเคชันที่ช่วยในการประมวลผลและแสดงผลการวิเคราะห์ได้อย่างมีประสิทธิภาพ เข้าถึงได้ง่าย เพื่อให้ผู้ใช้ทั่วไป ผู้ที่สนใจศึกษาหรือหน่วยงานที่เกี่ยวข้องสามารถประเมินความเสี่ยง และมีมาตรการป้องกันการเกิดอุบัติเหตุได้อย่างเหมาะสม

1.2 วัตถุประสงค์การวิจัย

1.2.1 เพื่อพัฒนาแบบจำลองการวิเคราะห์ระดับความรุนแรงของอุบัติเหตุทางถนน

1.2.2 เพื่อพัฒนาระบบสารสนเทศสำหรับวิเคราะห์อุบัติเหตุทางถนน

1.3 ขอบเขตการวิจัย

1.3.1 การค้นคว้าอิสระนี้ใช้ชุดข้อมูลอุบัติเหตุบนโครงข่ายถนนของกระทรวงคมนาคม ปี พ.ศ. 2562 – 2567 โดยข้อมูลชุดนี้เป็นข้อมูลเปิดเผยสาธารณะจากระบบรายงานข้อมูลอุบัติเหตุของกระทรวงคมนาคม (Transport Accident Management Systems: TRAMS) โดยศูนย์เทคโนโลยีสารสนเทศและการสื่อสาร สำนักงานปลัดกระทรวงคมนาคม มีจำนวน 133,090 ข้อมูล [4]

1.3.2 การค้นคว้าอิสระนี้ใช้วิธีการเรียนรู้ของเครื่อง ในรูปแบบการเรียนรู้แบบมีผู้สอน สำหรับการวิเคราะห์ระดับความรุนแรงในการเกิดอุบัติเหตุทางถนนโดยใช้แบบจำลองการเรียนรู้ของเครื่อง ได้แก่ ต้นไม้ตัดสินใจ ป่าการสุ่ม เอ็กซ์ตริม เกรเดียนต์ บูสต์ และโลท เกรเดียนต์ บูสต์ แมชชีน โดยใช้การวัดประสิทธิภาพของแบบจำลอง เพื่อระบุขั้นตอนวิธีที่ดีที่สุดในการวิเคราะห์ระดับความรุนแรงในการเกิดอุบัติเหตุทางถนน เพื่อนำแบบจำลองมาประยุกต์ใช้กับการพัฒนาระบบสารสนเทศในการแสดงผลข้อมูลอุบัติเหตุในประเทศไทย

1.3.3 การค้นคว้าอิสระนี้มีการทดสอบและประเมินประสิทธิภาพแบบจำลองที่ถูกนำมาใช้ในการวิเคราะห์ระดับความรุนแรงของอุบัติเหตุทางถนน โดยใช้มาตรวัดที่สำคัญคือ ค่าความถูกต้อง (Accuracy) ค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) และค่าคะแนนเอฟหนึ่ง (F1-Score) เพื่อเปรียบเทียบแบบจำลองที่ให้ผลลัพธ์ที่ดีและเหมาะสมที่สุด

1.3.4 ขอบเขตด้านฮาร์ดแวร์ (Hardware) มีรายละเอียดขั้นต่ำ ดังนี้

1.3.4.1 หน่วยประมวลผลกลาง (Central Processing Unit : CPU) i5-9300H 2.40 GHz

1.3.4.2 หน่วยความจำหลัก(Random Access Memory : RAM) 8 กิกะไบต์ (Gigabyte: GB)

1.3.4.3 พื้นที่จัดเก็บ (Storage) 256 กิกะไบต์ (Gigabyte: GB)

1.3.5 ขอบเขตด้านซอฟต์แวร์ (Software) มีรายละเอียด ดังนี้

1.3.5.1 บริการออนไลน์ Google Colab สำหรับการสร้าง และฝึกฝนขั้นตอนวิธี

1.3.5.2 โปรแกรม Microsoft Excel สำหรับการจัดการกับข้อมูล

1.3.5.3 โปรแกรม Visual Studio Code (VS Code) สำหรับการเขียนโปรแกรมพัฒนาระบบสารสนเทศ

1.3.5.4 ภาษา HyperText Markup Language (HTML) สำหรับการพัฒนาโครงสร้างระบบสารสนเทศ

1.3.5.5 Cascading Style Sheets (CSS) สำหรับการกำหนดและจัดการรูปแบบการแสดงผลของระบบสารสนเทศ

1.3.5.6 ภาษา JavaScript สำหรับการพัฒนาระบบสารสนเทศด้านการทำงานของผู้ใช้งาน

1.3.5.7 ภาษา Hypertext Preprocessor (PHP) สำหรับการพัฒนาระบบสารสนเทศสำหรับการเรียกใช้งาน Application Programming Interface (API)

1.3.5.8 ภาษา Python สำหรับการพัฒนาขั้นตอนวิธีการเรียนรู้ของเครื่อง

1.4 นิยามศัพท์เฉพาะ

1.4.1 อุบัติเหตุทางถนน (Road Accident) หมายถึง อุบัติเหตุที่เกิดขึ้นบนท้องถนนที่ทำให้เกิดความเสียหายต่อชีวิต ร่างกาย หรือทรัพย์สินของผู้ใช้ถนน ซึ่งอาจเกิดจากหลายสาเหตุ เช่น ความผิดพลาดของผู้ขับขี่ ความบกพร่องของยานพาหนะ สภาพถนน หรือสภาพอากาศที่ไม่เหมาะสม ส่งผลให้เกิดการบาดเจ็บ เสียชีวิต หรือทรัพย์สินเสียหาย [5]

1.4.2 ระดับความรุนแรงของอุบัติเหตุทางถนน (Severity of Road Accidents) หมายถึง การจัดกลุ่มอุบัติเหตุทางถนนตามความเสียหายที่เกิดขึ้น ซึ่งสามารถแบ่งออกได้เป็นหลายระดับตามความรุนแรงของผลกระทบต่อชีวิตและทรัพย์สินของผู้เกี่ยวข้อง

1.4.3 การเรียนรู้ของเครื่อง (Machine Learning) หมายถึง สาขาหนึ่งของวิทยาศาสตร์คอมพิวเตอร์ที่เกี่ยวข้องกับการสร้างและพัฒนาอัลกอริทึมและขั้นตอนวิธีที่ทำให้คอมพิวเตอร์สามารถเรียนรู้จากข้อมูลได้โดย การเรียนรู้ของเครื่องเป็นการนำข้อมูลที่มีอยู่มาใช้ในการฝึกสอนให้

คอมพิวเตอร์สามารถสร้างรูปแบบ (Pattern) หรือกฎเกณฑ์ (Rules) ที่สามารถนำไปใช้ในการคาดการณ์หรือการตัดสินใจในอนาคต

1.5 ประโยชน์ที่ได้รับ

1.5.1 ได้แบบจำลองสำหรับการวิเคราะห์ระดับความรุนแรงในการเกิดอุบัติเหตุทางถนน ซึ่งแบบจำลองดังกล่าวจะช่วยให้หน่วยงานที่เกี่ยวข้องเข้าใจสาเหตุที่ส่งผลให้เกิดความรุนแรงของอุบัติเหตุได้อย่างชัดเจน

1.5.2 ได้ระบบสารสนเทศสำหรับวิเคราะห์ระดับความรุนแรงของอุบัติเหตุทางถนน โดยนำเสนอข้อมูลจากการวิเคราะห์ในรูปแบบแดชบอร์ดรายงาน ที่มีความเป็นมิตรกับผู้ใช้ ทำให้หน่วยงานภาครัฐ เอกชน และผู้ที่สนใจสามารถเข้าถึงและศึกษาข้อมูลได้อย่างง่ายดาย สะดวก และรวดเร็ว

1.5.3 ได้แนวทางในการออกนโยบายและมาตรการป้องกันเพื่อลดความรุนแรงในการเกิดอุบัติเหตุทางถนน ซึ่งจะสนับสนุนหน่วยงานภาครัฐและผู้ที่เกี่ยวข้องในการสร้างสภาพแวดล้อมที่ปลอดภัยยิ่ง

บทที่ 2

เอกสารและงานวิจัยที่เกี่ยวข้อง

การวิเคราะห์ระดับความรุนแรงของอุบัติเหตุทางถนนโดยใช้การเรียนรู้ของเครื่อง ผู้จัดทำ การค้นคว้าอิสระได้ศึกษาแนวคิดทฤษฎีและงานวิจัยที่เกี่ยวข้องมาเป็นแนวทางในงานค้นคว้าอิสระ ดังต่อไปนี้

- 2.1 อุบัติเหตุทางถนน
- 2.2 การเรียนรู้ของเครื่อง
- 2.3 การวัดประสิทธิภาพการเรียนรู้ของเครื่อง
- 2.4 งานวิจัยที่เกี่ยวข้อง

2.1 อุบัติเหตุทางถนน

อุบัติเหตุทางถนนเป็นเหตุการณ์ที่เกิดขึ้นบนถนนที่เกี่ยวข้องกับการจราจรบนท้องถนน ซึ่งรวมถึงการชนกันระหว่างยานพาหนะกับยานพาหนะ ยานพาหนะกับบุคคล หรือยานพาหนะกับวัตถุอื่น ๆ โดยเหตุการณ์เหล่านี้มักส่งผลให้เกิดความเสียหายทั้งทางกายภาพและทรัพย์สิน รวมถึงการบาดเจ็บหรือการเสียชีวิตของผู้ที่เกี่ยวข้อง ผลกระทบจากอุบัติเหตุทางถนนยังสามารถส่งผลกระทบต่อชีวิตประจำวันของบุคคลและสังคมโดยรวม โดยอุบัติเหตุในแต่ละครั้งมีระดับความรุนแรงที่ต่างกันออกไป ดังนี้

2.1.1 ระดับความรุนแรงของอุบัติเหตุทางถนนเป็นการประเมินความรุนแรงที่เกิดขึ้นจากเหตุการณ์การชนหรือการประสบอุบัติเหตุที่เกี่ยวข้องกับยานพาหนะ โดยความรุนแรงของอุบัติเหตุสามารถจัดแบ่งได้เป็นหลายระดับ ซึ่งมีผลกระทบที่แตกต่างกันต่อผู้ที่เกี่ยวข้อง เช่น ผู้ขับขี่ ผู้โดยสาร และผู้เดินทางที่ได้รับผลกระทบความรุนแรงของอุบัติเหตุ สามารถแบ่งออกเป็น 2 ระดับ

2.1.1.1 ระดับที่ 1 เกิดอุบัติเหตุที่มีผู้บาดเจ็บแต่ไม่มีผู้เสียชีวิต

2.1.1.2 ระดับที่ 2 เกิดอุบัติเหตุมีผู้เสียชีวิต

2.1.2 การป้องกันการเกิดอุบัติเหตุทางถนน เป็นหนึ่งในภารกิจที่สำคัญที่สุดในการสร้างความปลอดภัยในการเดินทาง การพัฒนาและการปรับปรุงมาตรการป้องกันอุบัติเหตุทางถนนมีความสำคัญทั้งในด้านการลดความสูญเสียชีวิตและการบาดเจ็บ รวมถึงการลดค่าใช้จ่ายที่เกิดจากอุบัติเหตุ การศึกษาและการนำแนวทางที่มีประสิทธิภาพมาใช้เป็นวิธีที่สำคัญในการลดอุบัติเหตุทางถนน

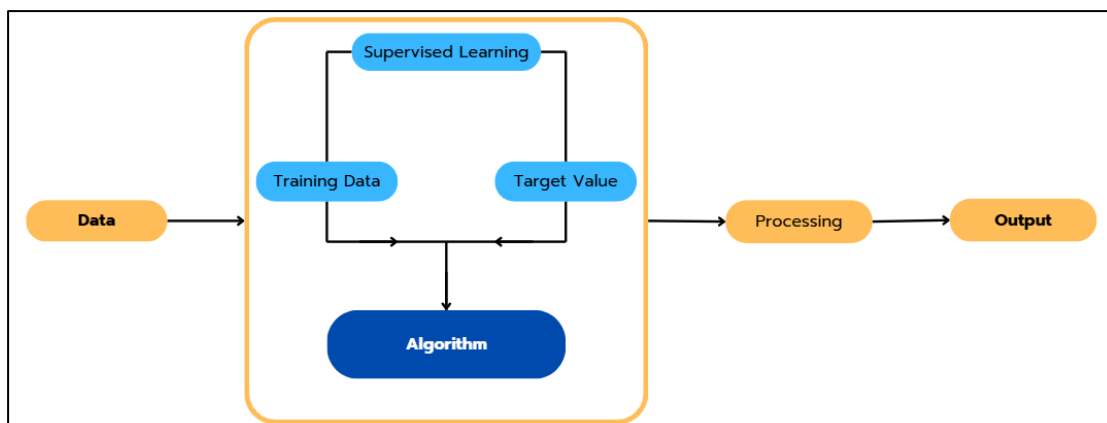
ซึ่งสามารถทำได้ผ่านหลายกลยุทธ์และทฤษฎีที่ได้รับการพัฒนาอย่างต่อเนื่อง ซึ่งการป้องกันการเกิดอุบัติเหตุทางถนนนั้น สามารถทำได้ทั้งเชิงนโยบายของภาครัฐและการนำเทคโนโลยีมาช่วยในการป้องกันการเกิดอุบัติเหตุ เช่น การเพิ่มมาตรการในการบังคับใช้กฎหมายทำให้จำนวนอุบัติเหตุและจำนวนผู้เสียชีวิตลดลง และ การใช้ระบบที่ช่วยตรวจจับและแจ้งเตือนเมื่อรถยนต์ขับเร็วเกินกว่าความเร็วที่กำหนด (SpeedAlert) เพื่อป้องกันไม่ให้ผู้ขับขี่ขับรถเร็วเกินไป ซึ่งช่วยลดความเสี่ยงจากการเกิดอุบัติเหตุที่เกิดจากความเร็ว [6]

2.1.3 การนำเทคโนโลยีมาช่วยในการลดอุบัติเหตุทางถนน มีบทบาทสำคัญในการเพิ่มความปลอดภัยในการขับขี่และลดอุบัติเหตุทางถนน ซึ่งไม่เพียงแต่ช่วยลดการบาดเจ็บและความเสียหายทางทรัพย์สิน แต่ยังช่วยเพิ่มความปลอดภัยให้กับผู้ใช้รถอีกด้วย โดยปัจจุบันมีการนำเทคโนโลยีมาใช้อย่างมากมายและกว้างขวาง เช่น ระบบป้องกันอุบัติเหตุทางถนนอัจฉริยะที่ช่วยผู้ขับขี่ในการหลีกเลี่ยงการชนโดยการให้ข้อความเตือนล่วงหน้า โดยใช้การเรียนรู้ของเครื่อง

2.2 การเรียนรู้ของเครื่อง

การเรียนรู้ของเครื่อง (Machine Learning) เป็นกระบวนการที่คอมพิวเตอร์สามารถเรียนรู้จากข้อมูลและประสบการณ์จากข้อมูลต่าง แล้วสามารถตัดสินใจหรือพยากรณ์ได้อย่างแม่นยำ โดยการนำการเรียนรู้ของเครื่องมาใช้ในการวิเคราะห์อุบัติเหตุทางถนนนั้นเป็นการเสริมสร้างความเข้าใจในปัจจัยต่าง ๆ ที่มีผลต่อการเกิดอุบัติเหตุ โดยเฉพาะข้อมูลที่มีความซับซ้อนและหลากหลายในด้านปัจจัยที่ส่งผลต่อระดับความรุนแรง การวิเคราะห์ข้อมูลด้วยการเรียนรู้ของเครื่อง สามารถช่วยระบุแนวโน้มและปัจจัยเสี่ยงได้อย่างดี [7] นอกจากนี้ยังช่วยในการพัฒนาแบบจำลองที่สามารถคาดการณ์เหตุการณ์ในอนาคตได้อย่างแม่นยำยิ่งขึ้น ทำให้การวางแผนและการตอบสนองต่ออุบัติเหตุมีประสิทธิภาพมากขึ้น ซึ่งการเรียนรู้ของเครื่องแบ่งออกเป็น 3 ประเภท ดังนี้

2.2.1 การเรียนรู้แบบมีผู้สอน (Supervised Learning) เป็นการใช้ชุดข้อมูลที่มีการระบุผลลัพธ์ที่ถูกต้องหรือ ป้ายกำกับ (Label) หรือการจัดกลุ่มเอาไว้ล่วงหน้าในชุดข้อมูลนั้น ๆ แล้วจะทำการเปรียบเทียบผลจากป้ายกำกับซึ่งแบ่งออกเป็น 2 ประเภทคือ การจำแนกประเภท (Classification) เหมาะสำหรับข้อมูลที่เป็นกลุ่มหรือหมวดหมู่ เช่น การทำนายการเป็นโรคหรือไม่เป็น การแยกประเภทของสินค้า การจำแนกภาพของสัตว์ เป็นต้น และการถดถอย (Regression) เหมาะสำหรับข้อมูลที่เป็นค่าตัวเลขต่อเนื่อง เช่น การทำนายราคารถยนต์ การทำนายราคาก่อนบ้าน การทำนายอุณหภูมิ เป็นต้น [8] โดยกระบวนการทำงานของการเรียนรู้แบบมีผู้สอนดังภาพที่ 2-1

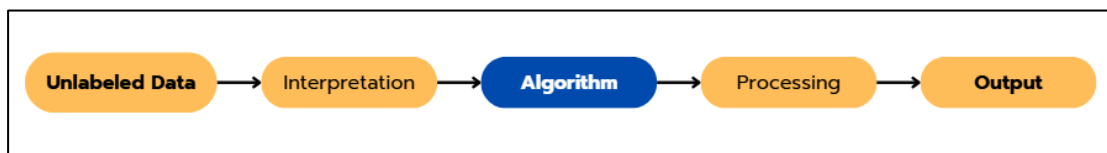


ภาพที่ 2-1 กระบวนการทำงานการเรียนรู้ของเครื่องแบบมีผู้สอน

จากภาพที่ 2-1 แสดงให้เห็นกระบวนการของการเรียนรู้แบบมีผู้สอน ซึ่งเริ่มต้นจากการนำข้อมูลเข้าสู่กระบวนการเรียนรู้ โดยข้อมูลที่ป้อนเข้ามาจะประกอบไปด้วยชุดข้อมูลฝึกสอนที่มีทั้งตัวแปรอินพุตและค่าผลลัพธ์ที่ถูกต้องกำกับอยู่ โดยข้อมูลจะถูกส่งเข้าสู่แบบจำลองที่ทำหน้าที่เรียนรู้รูปแบบความสัมพันธ์ระหว่างข้อมูลเข้าและผลลัพธ์เมื่อข้อมูลเข้าสู่กระบวนการเรียนรู้ แบบจำลองจะทำการวิเคราะห์และสร้างแบบจำลองโดยการจับคู่ข้อมูลเข้ากับค่าผลลัพธ์ที่กำหนดไว้ กระบวนการนี้จะดำเนินไปอย่างต่อเนื่องจนกระทั่งแบบจำลองสามารถทำนายผลลัพธ์ได้ใกล้เคียงกับค่าจริงมากที่สุด ซึ่งแบบจำลองที่ผ่านการฝึกฝนแล้วจะถูกนำมาใช้กับข้อมูลใหม่เพื่อทำนายหรือจำแนกผลลัพธ์ ระบบสามารถประมวลผลและให้ผลลัพธ์ที่แม่นยำตามข้อมูลที่ได้เรียนรู้ ซึ่งอาจอยู่ในรูปของการคาดการณ์ค่าหรือการจำแนกประเภทของข้อมูล เช่น การพยากรณ์ระดับความรุนแรงของอุบัติเหตุ หรือการจัดกลุ่มลักษณะของอุบัติเหตุในแต่ละพื้นที่

2.2.2 การเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) เป็นกระบวนการเรียนรู้ของแบบจำลองที่ไม่ต้องอาศัยผลลัพธ์ล่วงหน้าหรือฉลากกำกับกับข้อมูล แต่ให้แบบจำลองค้นหารูปแบบความสัมพันธ์ และโครงสร้างที่ซ่อนอยู่ในข้อมูลได้ด้วยตัวเอง กระบวนการนี้มักใช้เมื่อต้องการวิเคราะห์ข้อมูลที่ยังไม่มีการแบ่งกลุ่มหรือไม่มีการกำหนดผลลัพธ์ที่แน่ชัด การเรียนรู้ประเภทนี้มีบทบาทสำคัญในการค้นหาความเชื่อมโยงที่ซับซ้อนภายในชุดข้อมูลขนาดใหญ่ โดยแบบจำลองจะระบุรูปแบบที่เกิดขึ้นเองโดยอัตโนมัติ ซึ่งสามารถนำไปใช้ในหลายบริบท เช่น การจัดกลุ่มข้อมูล (Clustering) เพื่อจำแนกประเภทของอุบัติเหตุโดยพิจารณาจากลักษณะของเหตุการณ์ หรือการค้นหาแบบแผนพฤติกรรมการขับขี่ที่ชี้ถึงความเสี่ยงในการเกิดอุบัติเหตุ การนำการเรียนรู้แบบไม่มีผู้สอนมาใช้ในการวิเคราะห์ข้อมูลอุบัติเหตุ ช่วยให้สามารถค้นหาแนวโน้มของอุบัติเหตุที่เกิดขึ้นบ่อย ประเมินความเสี่ยงของพื้นที่ต่าง ๆ และเข้าใจปัจจัยที่อาจส่งผลต่อการเกิดอุบัติเหตุได้ดีขึ้น ซึ่งเป็นแนวทาง

สำคัญในการวางแผนและพัฒนาแนวทางป้องกันอุบัติเหตุในอนาคต [9] โดยกระบวนการทำงานของการเรียนรู้แบบไม่มีผู้สอนดังภาพที่ 2-2

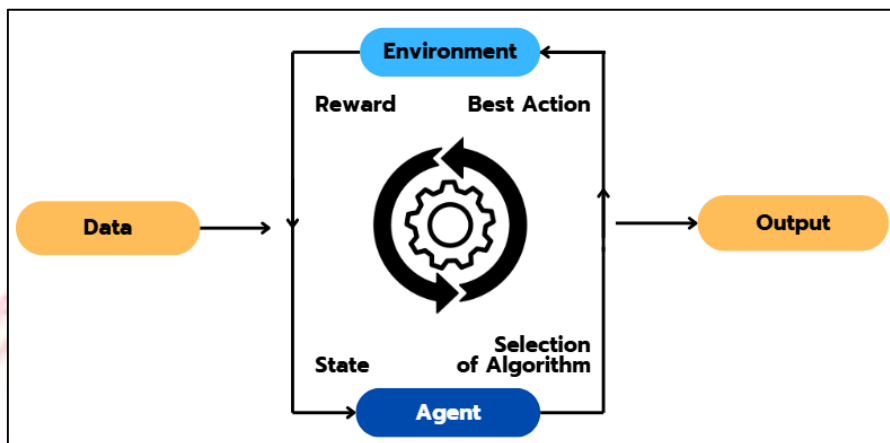


ภาพที่ 2-2 กระบวนการทำงานการเรียนรู้ของเครื่องแบบไม่มีผู้สอน

จากภาพที่ 2-2 แสดงให้เห็นกระบวนการทำงานการเรียนรู้ของเครื่องแบบไม่มีผู้สอน โดยข้อมูลที่ไม่มีป้ายกำกับ (Unlabeled Data) ซึ่งเป็นข้อมูลที่ยังไม่มีการจัดหมวดหมู่หรือกำหนดค่าผลลัพธ์ที่ชัดเจน ข้อมูลจะถูกส่งเข้าสู่กระบวนการตีความ (Interpretation) ซึ่งแบบจำลองพยายามวิเคราะห์และทำความเข้าใจโครงสร้างหรือรูปแบบที่ซ่อนอยู่ภายในข้อมูล โดยไม่มีการชี้นำจากตัวแปรเป้าหมาย หลังจากนั้นข้อมูลที่ผ่านการตีความจะเข้าสู่กระบวนการเรียนรู้ผ่านแบบจำลอง ซึ่งทำหน้าที่ค้นหารูปแบบและความสัมพันธ์ระหว่างข้อมูล กระบวนการนี้ช่วยให้แบบจำลองสามารถจัดกลุ่มข้อมูลที่มีความคล้ายคลึงกันหรือค้นหากลุ่มของข้อมูลที่มีลักษณะร่วมกัน หลังจากที่ได้แบบจำลองประมวลผลข้อมูลเสร็จสิ้น ข้อมูลจะถูกส่งต่อไปยังขั้นตอนการประมวลผล (Processing) ซึ่งเป็นการปรับปรุงและสรุปผลลัพธ์ที่ได้จากแบบจำลองเพื่อให้สามารถนำไปใช้งานต่อได้อย่างมีประสิทธิภาพ

2.2.3 การเรียนรู้แบบเสริมกำลัง (Reinforcement Learning) เป็นกระบวนการที่แบบจำลองเรียนรู้ผ่านการทดลองและการปรับปรุงพฤติกรรมโดยอาศัยรางวัล (Reward) และบทลงโทษ (Penalty) ที่ได้รับจากสภาพแวดล้อม แบบจำลองจะค่อยๆ พัฒนาแนวทางตัดสินใจที่เหมาะสมที่สุดผ่านการโต้ตอบกับสิ่งแวดล้อม เพื่อให้เกิดผลลัพธ์ที่ดีที่สุดในระยะยาว แนวทางนี้สามารถนำมาใช้ในระบบที่ต้องการการตัดสินใจแบบพลวัต (Dynamic) และการปรับตัวอย่างต่อเนื่อง เช่น การสร้างระบบอัจฉริยะสำหรับการจัดการจราจร เพื่อลดความเสี่ยงจากอุบัติเหตุในพื้นที่ที่มีความเสี่ยงสูง ระบบสามารถเรียนรู้และคาดการณ์สภาพการจราจรแบบตามเวลาจริง (Real Time) พร้อมทั้งปรับเปลี่ยนสัญญาณไฟจราจรหรือแนะนำเส้นทางที่ปลอดภัยที่สุดให้กับผู้ใช้ถนน การนำการเรียนรู้แบบเสริมกำลังมาใช้กับระบบจราจรช่วยให้สามารถระบุปัจจัยเสี่ยงได้อย่างแม่นยำ และสร้างแนวทางตอบสนองที่เหมาะสมต่อสถานการณ์ต่าง ๆ เช่น การปรับเวลาสัญญาณไฟเพื่อบรรเทาความแออัด ลดจุดเสี่ยงที่อาจเกิดอุบัติเหตุ หรือส่งสัญญาณเตือนให้ผู้ขับขี่ระมัดระวังในบริเวณที่มีความเสี่ยงสูง การประยุกต์ใช้แนวทางการเรียนรู้ของเครื่องในลักษณะนี้ทำให้สามารถบริหารจัดการการจราจรได้อย่างมีประสิทธิภาพ

ประสิทธิภาพมากขึ้น ส่งผลให้ลดอุบัติเหตุ เพิ่มความปลอดภัยบนท้องถนน และสร้างสภาพแวดล้อมการขับขี่ที่ดีขึ้น [10] โดยกระบวนการทำงานของการเรียนรู้แบบเสริมกำลังภาพที่ 2-3



ภาพที่ 2-3 กระบวนการทำงานการเรียนรู้ของเครื่องแบบเสริมกำลัง

จากภาพที่ 2-3 แสดงให้เห็นกระบวนการทำงานการเรียนรู้ของเครื่องแบบเสริมกำลัง โดยข้อมูล ถูกส่งเข้าสู่ระบบไปที่ตัวแทน (Agent) และตัวแทนจะทำหน้าที่เลือกแบบจำลองที่เหมาะสมเพื่อใช้ในการประมวลผลและตัดสินใจในแต่ละสถานการณ์ ตัวแทนจะรับรู้สถานะ (State) ของสภาพแวดล้อมในขณะนั้นและทำการตัดสินใจเลือกการกระทำที่ดีที่สุดเพื่อให้ได้ผลลัพธ์ที่เหมาะสมที่สุด เมื่อดำเนินการแล้วตัวแทนจะได้รับผลลัพธ์ ซึ่งเป็นผลที่เกิดขึ้นจากการกระทำที่เลือก สภาพแวดล้อมจะทำหน้าที่ให้ผลตอบแทน (Reward) หรือค่าการวัดผลที่สะท้อนว่าการกระทำนั้นดีหรือไม่ดี โดยหากการกระทำที่เลือกส่งผลดีต่อเป้าหมายของระบบ ตัวแทนจะได้รับค่าตอบแทนเชิงบวก และหากเป็นการกระทำที่ไม่เหมาะสม อาจได้รับค่าตอบแทนเชิงลบ หรือไม่มีรางวัลเลย กระบวนการนี้เกิดขึ้นอย่างต่อเนื่อง โดยตัวแทนจะใช้ค่าตอบแทนที่ได้รับมาปรับปรุงกลยุทธ์และเลือกการกระทำที่ดียิ่งขึ้นในรอบถัดไป ทำให้แบบจำลองสามารถเรียนรู้และพัฒนาแนวทางการตัดสินใจที่มีประสิทธิภาพมากขึ้นเรื่อย ๆ

จากการค้นคว้าเกี่ยวกับการเรียนรู้ของเครื่อง ทั้งการเรียนรู้แบบมีผู้สอน การเรียนรู้แบบไม่มีผู้สอน และการเรียนรู้แบบเสริมกำลัง พบว่าแต่ละวิธีมีคุณลักษณะเฉพาะและเหมาะสมกับลักษณะของข้อมูลและปัญหาที่ต้องการแก้ไขแตกต่างกันไปสำหรับการเรียนรู้แบบมีผู้สอนนั้น อาศัยข้อมูลที่มีผลลัพธ์หรือป้ายกำกับที่ชัดเจนอยู่แล้ว ทำให้สามารถใช้ข้อมูลตัวอย่างในการสร้างแบบจำลองที่สามารถจำแนกและทำนายผลลัพธ์ได้อย่างแม่นยำ ดังนั้นการค้นคว้าอิสระในครั้งนี้ ใช้ข้อมูลจากกระทรวงคมนาคม ประเทศไทย โดยสามารถระบุระดับความรุนแรงของอุบัติเหตุได้อย่างชัดเจน จึง

เหมาะสมกับการใช้การเรียนรู้แบบมีผู้สอน โดยเฉพาะในรูปแบบของการจำแนกประเภท (Classification) มากกว่าการใช้แนวทางอื่น ๆ [11]

2.3 การวัดประสิทธิภาพแบบจำลอง

การวัดประสิทธิภาพของแบบจำลองการเรียนรู้ (Learning Performance) เป็นขั้นตอนที่ช่วยให้ผู้ค้นคว้าอิสระทราบถึงแบบจำลองที่พัฒนาขึ้นสามารถทำงานได้ดีตามที่ต้องการหรือไม่ โดยการวัดประสิทธิภาพของแบบจำลองมีตัวชี้วัด ดังนี้

2.3.1 ตารางสับสน (Confusion Matrix)

ตารางสับสน คือเครื่องมือสำหรับการวัดประสิทธิภาพของแบบจำลองในการจำแนกประเภท โดยตารางสับสนจะทำงานร่วมกับชุดข้อมูล โดยการเปรียบเทียบค่าจริงและค่าที่คาดการณ์จากแบบจำลอง โดยรูปแบบตารางดังภาพที่ 2-4

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

ภาพที่ 2-4 รูปแบบตารางสับสน [12]

จากภาพที่ 2-4 สามารถอธิบายได้ดังนี้

2.3.1.1 True Positive (TP) = สิ่ง que แบบจำลองทำนายหรือวิเคราะห์ว่าจริง และเป็นสิ่งที่เกิดขึ้นจริงนั่นคือสิ่งที่ทำนายหรือวิเคราะห์ตรงกับความเป็นจริง

2.3.1.2 True Negative (TN) = สิ่ง que แบบจำลองทำนายหรือวิเคราะห์ว่าไม่จริง และเป็นสิ่งที่เกิดขึ้นนั้นไม่ นั่นคือสิ่งที่ทำนายหรือวิเคราะห์ตรงกับความเป็นจริงจริง

2.3.1.3 False Positive (FP) = สิ่ง que แบบจำลองทำนายหรือวิเคราะห์ว่าจริง และเป็นสิ่งที่ไม่เกิดขึ้นจริงนั่นคือสิ่งที่ทำนายหรือวิเคราะห์ไม่ตรงกับความเป็นจริง

2.3.1.4 False Negative (FN) = สิ่งที่เกิดแบบจำลองทำนายหรือวิเคราะห์ว่าไม่จริง และ เป็นสิ่งที่เกิดขึ้นจริง นั่นคือสิ่งที่ทำนายหรือวิเคราะห์ไม่ตรงกับความเป็นจริง

2.3.2 ค่าความถูกต้อง

เป็นการวัดค่าความถูกต้องของแบบจำลองว่าสามารถทำนายหรือวิเคราะห์ได้อย่างถูกต้อง โดย พิจารณาจากจำนวนที่แบบจำลองทำนายถูก เมื่อเทียบกับจำนวนกรณีทั้งหมดของชุดข้อมูล ดังสมการ ที่ 2-1

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (2-1)$$

2.3.3 ค่าความแม่นยำ

เป็นการวัดความแม่นยำที่ใช้ประเมินความถูกต้องของการทำนายหรือการวิเคราะห์ ในกรณีที่ แบบจำลองระบุข้อมูลเป็นบวกและถูกต้อง (TP) กับจำนวนทั้งหมดของกรณีที่แบบจำลองทำนายหรือ วิเคราะห์ว่าข้อมูลนั้นเป็นบวก (TP + FP) ดังสมการที่ 2-2

$$Precision = \frac{TP}{TP+FP} \quad (2-2)$$

2.3.4 ค่าความระลึก

เป็นการวัดค่าแบบจำลองในการระบุข้อมูลที่เป็นบวกทั้งหมดในชุดข้อมูลให้ครบถ้วน โดยเมื่อมี ข้อมูลจริงอยู่ในการจำแนกข้อมูลที่เป็นบวก ค่าความระลึกจะบอกถึงความสามารถของแบบจำลองใน การตรวจจับข้อมูลได้ถูกต้อง ดังสมการที่ 2-3

$$Recall = \frac{TP}{TP+FN} \quad (2-3)$$

2.3.5 ค่าคะแนนเอฟหนึ่ง

เป็นการวัดที่ใช้การประเมินประสิทธิภาพของแบบจำลองในการจำแนกประเภทโดยเอาค่า ความแม่นยำ และค่าความระลึกมารวมกันโดยใช้ค่าเฉลี่ยน้ำหนักแบบฮาร์มอนิก (Harmonic Mean) ดังสมการที่ 2-4

$$F1 - Score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (2-4)$$

การวัดประสิทธิภาพของแบบจำลองไม่ได้จำกัดเพียงแค่การพิจารณาค่า ค่าความถูกต้องเพียงอย่างเดียว แต่ยังต้องพิจารณา ค่าความแม่นยำ ค่าความระลึก และค่าคะแนนเอฟหนึ่ง และการวิเคราะห์ ตารางสับสน เพื่อให้ได้ข้อมูลที่ครอบคลุมและช่วยในการปรับปรุงแบบจำลองให้สามารถจำแนกระดับความรุนแรงของอุบัติเหตุบนท้องถนนได้อย่างถูกต้องและน่าเชื่อถือมากยิ่งขึ้น

2.3 งานวิจัยที่เกี่ยวข้อง

อุบัติเหตุทางถนนเป็นปัญหาที่มีผลกระทบอย่างรุนแรงทั้งในระดับบุคคลและระดับสังคม โดยส่งผลกระทบต่อชีวิตผู้คนและสร้างความเสียหายทางเศรษฐกิจในระดับมหภาค จากการศึกษาพบว่างานวิจัยได้ใช้เทคโนโลยีการเรียนรู้ของเครื่องเพื่อวิเคราะห์แนวโน้มของอุบัติเหตุและพยากรณ์ระดับความรุนแรงของเหตุการณ์เหล่านี้ เพื่อนำไปสู่แนวทางป้องกันและลดอุบัติเหตุได้อย่างเป็นระบบ งานวิจัยที่ได้ศึกษาในครั้งนี้นำเสนอแนวทางที่หลากหลายทั้ง แบบจำลอง ข้อมูล และข้อดีข้อเสียของแต่ละงานวิจัย

การศึกษากการวิเคราะห์ความรุนแรงของอุบัติเหตุทางถนนโดยใช้การเรียนรู้ของเครื่องพบว่าผู้วิจัยได้ทำการศึกษาปัจจัยที่ส่งผลต่อการเกิดอุบัติเหตุ หรือการพยากรณ์อุบัติเหตุทางถนนโดย Meet Gorasia และคณะ ได้ศึกษาและวิเคราะห์ข้อมูลอุบัติเหตุ โดยเลือกใช้แบบจำลองต้นไม้ตัดสินใจ เค-เพื่อนบ้านใกล้สุด (k-Nearest Neighbour: kNN) และ ป่าสุ่ม [13] ผลการศึกษาพบว่าป่าสุ่มมีค่าความแม่นยำสูงสุดที่ 87 ซึ่งปัจจัยสำคัญที่ส่งผลต่อระดับความรุนแรงของอุบัติเหตุ ได้แก่ พฤติกรรมของผู้ขับขี่ สภาพแวดล้อมของถนน สภาพอากาศ และช่วงเวลาที่เกิดอุบัติเหตุ โดยพบว่าอุบัติเหตุร้ายแรงมักเกิดขึ้นในเวลากลางวัน จากงานวิจัยดังกล่าวผู้วิจัยได้เสนอให้มีการนำข้อมูลแบบเรียลไทม์จากเซ็นเซอร์จราจรเข้ามาใช้ เพื่อเพิ่มประสิทธิภาพของแบบจำลอง โดยผลการวิจัยสอดคล้อง Shakil Ahmed ได้ศึกษาการพยากรณ์ระดับความรุนแรงของอุบัติเหตุ และวิเคราะห์ปัจจัยที่เกี่ยวข้อง โดยใช้การเรียนรู้ของเครื่องเปรียบเทียบกับแบบจำลองอื่น ๆ เช่น ป่าสุ่ม ป่าการตัดสินใจ การเสริมแรงแบบปรับตัว (Adaptive Boosting) เอ็กซ์ตรีม เกรเดียนต์ บูสต์ติ้ง โลท เกรเดียนต์ บูสต์ติ้ง แมชชีน และ แคตทิเกอริคอล บูสต์ติ้ง (Categorical Boosting) [14] พบว่าป่าสุ่ม ให้ค่าความแม่นยำสูงสุดที่ 81.45% และเมื่อนำการค้นหปัจจัยแบบ แซปเลียย์ แอ็ดดิทิฟเอ็กซ์พลานชันส์ (Shapley

Additive Explanations: SHAP) มาใช้ในการวิเคราะห์ปัจจัย พบว่า ประเภทของถนน เป็นปัจจัยที่มีผลกระทบต่อความรุนแรงของอุบัติเหตุมากที่สุด รองลงมาคือ จำนวนยานพาหนะที่เกี่ยวข้อง การบริโภคสารเสพติดและแอลกอฮอล์ และ สภาพแวดล้อมของอุบัติเหตุ

ซึ่งสอดคล้องกับ Kavana V H และคณะได้ศึกษาการพยากรณ์ระดับความรุนแรงของอุบัติเหตุ โดยเปรียบเทียบโมเดล ซิงเกิลโหมด แมชชีนเลิร์นนิง (Single-mode Machine Learning) และ เอนเซมเบิลโหมด แมชชีนเลิร์นนิง (Ensemble-mode Machine Learning) [15] ผลลัพธ์พบว่า ป่าสุ่มให้ค่าความถูกต้องสูงที่สุดสำหรับการจำแนกแบบทวิภาค (Binary) ที่ 86.64% และสำหรับการจำแนกข้อมูลออกเป็นหลายคลาส (Multi Class) ให้ค่าความแม่นยำที่ 67.67% นอกจากนี้ ยังพบว่าปัจจัยที่มีผลกระทบต่อความรุนแรงของอุบัติเหตุ ได้แก่ จำนวนยานพาหนะที่เกี่ยวข้อง สภาพอากาศ ถนน และความเร็วของยานพาหนะ ซึ่งเหตุทางปัจจัยสอดคล้องกับ Dheya Mustafa ที่ได้ศึกษาและเปรียบเทียบแบบจำลองหลายตัว เช่น เครื่องจักรเวกเตอร์สนับสนุน (Support Vector Machine) เค-เพื่อนบ้านใกล้สุด เนฟ เบย์ส (Naïve Bayes) การถดถอยแบบโลจิสติก (Logistic Regression) และต้นไม้ตัดสินใจ [16] โดยพบว่าปัจจัยที่มีผลกระทบต่อความรุนแรงของอุบัติเหตุ ได้แก่ จำนวนยานพาหนะที่เกี่ยวข้อง สภาพอากาศ ลักษณะของถนน และพฤติกรรมของผู้ขับขี่ โดยเฉพาะความเร็วของยานพาหนะขณะเกิดเหตุ ซึ่งแบบจำลองที่ให้ผลที่ดีที่สุดคือ การถดถอยแบบโลจิสติกมีความแม่นยำที่ 98.1% จากนั้นผู้วิจัยยังเสนอให้นำ การเรียนรู้เชิงลึก (Deep Learning) และ โครงข่ายประสาทเทียม (Neural Networks) มาทดลองเพิ่มเติมเพื่อปรับปรุงความแม่นยำของแบบจำลองโดย Umar Ilyasu ได้ศึกษาการเกิดอุบัติเหตุในไนจีเรียโดยใช้ ป่าสุ่ม ต้นไม้ตัดสินใจ และเพื่อนบ้านที่ใกล้เคียงที่สุดเคตัว [17] ผลการทดลองพบว่าป่าสุ่มให้ค่าความแม่นยำสูงสุดที่ 85% ในขณะที่ ต้นไม้ตัดสินใจ และเค-เพื่อนบ้านใกล้สุด มีค่าความแม่นยำต่ำว่ามาก การวิเคราะห์ รีซีฟเวอร์ โอเปอเรตติ้ง แคร็กเทอริสติก เคิร์ฟ (Receiver Operating Characteristic Curve) แสดงให้เห็นว่าป่าสุ่มมีความสามารถในการพยากรณ์สูงสุด เนื่องจากมีอัตราการทำนายถูกในกลุ่มบวก (True Positive Rate) สูง และ อัตราการแจ้งผิดว่าเป็นบวก (False Positive Rate) ต่ำ

เมื่อเปรียบเทียบทุกงานวิจัยที่กล่าวมาจะพบว่าหลายงานมีข้อสรุปที่คล้ายคลึงกัน ในเรื่องของแบบจำลองที่ให้ผลลัพธ์ดีที่สุดคือป่าสุ่มและ ปัจจัยที่มีผลกระทบต่อความรุนแรงของอุบัติเหตุ เช่น สภาพถนน สภาพอากาศ พฤติกรรมของผู้ขับขี่ และจำนวนยานพาหนะที่เกี่ยวข้อง ซึ่งจากข้อเสนอแนะจะพบว่างานวิจัยที่ศึกษาในบริบทของแต่ละประเทศมีข้อจำกัดในเรื่องของข้อมูลที่ใช้ ซึ่งมักมาจากแหล่งเดียวเท่านั้น ทำให้ผลลัพธ์อาจไม่สามารถใช้ได้กับทุกประเทศ ข้อเสนอแนะที่สำคัญจากงานวิจัยหลาย

ขึ้นคือ การนำข้อมูลจากหลายประเทศหรือการใช้ ข้อมูลจราจรแบบตามเวลาปัจจุบัน (Real-Time Traffic Data) และ การเรียนรู้ของเครื่องบนพื้นฐานอินเทอร์เน็ตของสรรพสิ่ง (IoT-based Machine Learning) เพื่อเพิ่มความแม่นยำของแบบจำลอง และสามารถนำไปใช้จริงในการลดอุบัติเหตุได้อย่างมีประสิทธิภาพมากขึ้น

ในขณะเดียวกันมีผู้วิจัยได้ทำการศึกษา ระบบตามกฎเกณฑ์ ที่ใช้ในการวิเคราะห์ปัจจัยต่าง ๆ ที่เกี่ยวข้องกับอุบัติเหตุทางถนน Ching-Hsue Cheng และคณะ [18] ได้พัฒนาแบบจำลองในการจำแนกประเภทที่ใช้หลักการตามกฎเพื่อสำรวจปัจจัยที่เกี่ยวข้องกับอุบัติเหตุ โดยอิงจากข้อมูลสาธารณะในเมืองเกาหยวน ประเทศไต้หวัน พบว่าพฤติกรรมการขับขี่ของผู้ขับขี่ที่มีส่วนสำคัญ คิดเป็น 66.19% ของสาเหตุอุบัติเหตุ นอกจากนี้ยังมีปัจจัย เช่น ประเภทยานพาหนะ และประเภทใบขับขี่ที่ส่งผลต่อความรุนแรงของอุบัติเหตุ โดย Berrones-Sanz และคณะ [19] ได้ทำการวิเคราะห์ข้อมูลอุบัติเหตุจราจรในเม็กซิโก โดยเน้นการใช้เทคโนโลยีในการบันทึกข้อมูล พบว่ามีการเก็บข้อมูลจำนวนมากแต่ยังไม่ถูกวิเคราะห์อย่างเต็มที่ ซึ่งส่งผลต่อการวิเคราะห์อุบัติเหตุ การใช้เทคนิคการเรียนรู้ของเครื่องเพื่อค้นหารูปแบบและอนุมานจากข้อมูลกลายเป็นสิ่งสำคัญในการปรับปรุงความปลอดภัยทางถนน และ Datu, N. H. [20] ได้ใช้เทคนิคการชดเชยความสัมพันธ์ในเมืองอังกะเลส ประเทศฟิลิปปินส์ โดยพบว่าช่วงเวลาและสถานที่ที่เกิดอุบัติเหตุสูงสุดอยู่ในช่วงเวลา 20:01 น. ถึง 00:00 น. ของวันเสาร์ การสร้างกฎความสัมพันธ์ช่วยให้สามารถระบุแนวโน้มและลักษณะของอุบัติเหตุที่สำคัญการศึกษาเหล่านี้ชี้ให้เห็นถึงความสำคัญของการวิเคราะห์ข้อมูลและเทคโนโลยีในการปรับปรุงความปลอดภัยทางถนน โดยการพิจารณาปัจจัยที่เกี่ยวข้องสามารถนำไปสู่การพัฒนามาตรการเพื่อป้องกันและลดอุบัติเหตุได้อย่างมีประสิทธิภาพมากขึ้น

การศึกษาปัจจัยที่ส่งผลต่อการเกิดอุบัติเหตุด้วยการวิเคราะห์อนุกรมเวลา โดยงานวิจัยของ Shafay et al. [21] ได้ดำเนินการศึกษาอุบัติเหตุทางถนนที่แผนกผู้บาดเจ็บของโรงพยาบาล Liaquat University ในเมืองไฮเดอราบาด ประเทศปากีสถาน โดยใช้การออกแบบ การศึกษาเชิงพยากรณ์ (Prospective Study) และการสุ่มตัวอย่างต่อเนื่องแบบความน่าจะเป็นในการเก็บข้อมูล 966 กรณีระหว่างเดือนมกราคม พ.ศ. 2563 ถึงธันวาคม พ.ศ. 2565 นักวิจัยใช้กราฟดัชนีตามฤดูกาลและกราฟอนุกรมเวลา (Time Series) เพื่อแสดงแนวโน้มการเกิดอุบัติเหตุตามฤดูกาล ซึ่งผลการศึกษาชี้ให้เห็นว่ามีจำนวนกรณีสูงสุดในเดือนธันวาคม ในขณะเดียวกัน การศึกษาโดย Frank B. K. Twenefour และคณะ [22] มุ่งเน้นการวิเคราะห์อุบัติเหตุจราจรในประเทศกานา โดยใช้ข้อมูลระหว่างเดือนมกราคม พ.ศ. 2533 ถึงธันวาคม พ.ศ. 2562 ผ่านการสร้างแบบจำลองอนุกรมเวลา (Time Series

Model) ในแบบจำลองถดถอยอัตโนมัติและค่าเฉลี่ยเคลื่อนที่ (Autoregressive Moving Average : ARMA) เพื่อวิเคราะห์แนวโน้มและการคาดการณ์อุบัติเหตุ โดยผลการศึกษาชี้ให้เห็นถึงการคาดการณ์อุบัติเหตุในช่วงเวลาต่าง ๆ รวมถึงแนวโน้มที่เพิ่มขึ้นของอุบัติเหตุทางถนน การใช้วิธีการอนุกรมเวลาในทั้งสองงานวิจัยนี้แสดงให้เห็นถึงความสำคัญในการเข้าใจลักษณะการเกิดอุบัติเหตุและปัจจัยที่มีผลกระทบในแต่ละช่วงเวลา เพื่อพัฒนามาตรการด้านความปลอดภัยบนถนนอย่างมีประสิทธิภาพ โดยทั้งสองการศึกษาเสริมสร้างความรู้เกี่ยวกับอุบัติเหตุทางถนนในแต่ละประเทศและแนวทางในการแก้ไขปัญหา ซึ่งการใช้การเรียนรู้ของเครื่อง ระบบตามกฎเกณฑ์ และการวิเคราะห์อนุกรมเวลา ในการศึกษาเกี่ยวกับอุบัติเหตุทางถนนมีบทบาทสำคัญในการวิเคราะห์และพยากรณ์ความรุนแรงของอุบัติเหตุ การเรียนรู้ของเครื่องสามารถวิเคราะห์ข้อมูลขนาดใหญ่และค้นหารูปแบบที่ซับซ้อนได้อย่างแม่นยำ โดยใช้ขั้นตอนวิธีต่าง ๆ เช่น ป่าสุ่ม และ เนฟ เบย์ส เป็นต้น ซึ่งเหมาะสำหรับการพยากรณ์และวิเคราะห์ความรุนแรงของอุบัติเหตุตามปัจจัยที่เกี่ยวข้อง เช่น พฤติกรรมการขับขี่และสภาพอากาศในขณะนั้น ในอีกวิธีหนึ่งระบบตามกฎเกณฑ์ช่วยให้สามารถกำหนดกฎที่ชัดเจนในการจำแนกประเภทอุบัติเหตุ แต่มีข้อจำกัดในการจัดการกับข้อมูลที่ซับซ้อนและการสร้างโมเดลที่ยืดหยุ่นน้อยกว่า ส่วนการวิเคราะห์อนุกรมเวลาจะมุ่งเน้นที่การวิเคราะห์แนวโน้มในช่วงเวลาต่าง ๆ ซึ่งอาจไม่เหมาะสมในการพยากรณ์หรือวิเคราะห์ความรุนแรงของอุบัติเหตุในบริบทที่ต้องการข้อมูลเชิงลึกจากหลายปัจจัย ดังนั้นการเลือกใช้การเรียนรู้ของเครื่องจึงเหมาะสมในการวิจัยที่มีข้อมูลมากและมีปัจจัยที่หลากหลาย เนื่องจากมีความสามารถในการจัดการข้อมูลที่ซับซ้อนและการค้นหารูปแบบที่มีความสัมพันธ์กันได้อย่างมีประสิทธิภาพ อีกทั้งยังสามารถปรับปรุงขั้นตอนวิธีได้อย่างต่อเนื่อง ทำให้การวิเคราะห์มีความแม่นยำและเชื่อถือได้มากยิ่งขึ้น จากการศึกษางานวิจัยที่เกี่ยวข้องผู้จัดทำการค้นคว้าอิสระได้สรุปงานวิจัยที่เกี่ยวข้อง ดังตารางที่ 2-1

ตารางที่ 2-1 งานวิจัยที่เกี่ยวข้องกับการพัฒนาแบบจำลอง

แบบจำลอง	ข้อดี	ข้อจำกัด
ต้นไม้ตัดสินใจ [13], [16], [17], [22], [24], [29], [28]	แบบจำลองสามารถเข้าใจง่าย รองรับทั้งข้อมูลเชิงตัวเลขและ หมวดหมู่ในการพัฒนา แบบจำลองไม่ต้องปรับแต่ง คุณลักษณะมากสามารถทำงาน ได้แม้มีข้อมูลมีค่าว่าง	แบบจำลองนี้เกิดการเรียนรู้มากเกินไปของแบบจำลอง (Overfitting) ได้ง่าย ไม่เหมาะกับข้อมูลที่มีความสัมพันธ์เชิงเส้นสูง

ตารางที่ 2-1 (ต่อ)

แบบจำลอง	ข้อดี	ข้อจำกัด
ป่าสุ่ม [13], [14], [16], [17], [19], [28], [29]	สามารถลดการเรียนรู้มากเกินไปของแบบจำลองได้ดี มีระดับความสำคัญของตัวแปรให้วิเคราะห์ได้	ใช้เวลาและหน่วยความจำสูงและยากต่อการตีความผลรวมของหลายต้นไม้ โดยที่ถ้ามีข้อมูลที่ไม่เกี่ยวข้อง (Noisy Features) อาจไม่มีแนวโน้มเกิดขึ้น
เค-เพื่อนบ้านใกล้สุด [13], [17], [24], [28]	ไม่ต้องมีการฝึกแบบจำลอง (lazy learner) และง่ายต่อการดำเนินการ (Implement) โดยที่แบบจำลองรองรับการจัดกลุ่มเชิงพื้นที่ได้ดี	แบบจำลองมีประสิทธิภาพต่ำเมื่อข้อมูลมีจำนวนมาก มีความอ่อนไหวต่อค่าผิดปกติ (Outlier) และต้องเลือกค่า เค ที่เหมาะสมเสมอ
เนฟ เบย์ส [13], [16], [17], [28]	แบบจำลองมีความเร็วและประสิทธิภาพสูงกับข้อความ (Text classification) สามารถใช้ได้แม้ข้อมูลมีค่าว่าง	สมมติฐานความเป็นอิสระของคุณลักษณะ (Independence) มักไม่เป็นจริง และไม่สามารถเรียนรู้จากความสัมพันธ์ของคุณลักษณะโดยที่ประสิทธิภาพอาจต่ำเมื่อมีคุณลักษณะที่ต้องพึ่งพากัน
การถดถอยแบบโลจิสติก [13], [17], [28]	แบบจำลองเข้าใจง่าย เป็นแบบจำลองเชิงเส้นช่วยวิเคราะห์ความสัมพันธ์ของผลลัพธ์ได้ดี	แบบจำลองไม่เหมาะกับข้อมูลที่ไม่เป็นข้อมูลแบบเชิงเส้นไม่เหมาะเมื่อเจอภาวะที่ตัวแปรอิสระมีความสัมพันธ์กันสูง (Multicollinearity)
เครื่องจักรเวกเตอร์สนับสนุน [17], [19], [28], [35]	- แบบจำลองมีประสิทธิภาพสูงในข้อมูลที่ระยะห่าง (Margin) ที่ชัดเจน ซึ่งรองรับเคอร์เนล (Kernel) เพื่อจัดการข้อมูลที่ไม่เป็นเชิงเส้น สามารถใช้ได้แม้จำนวนข้อมูลไม่มาก	แบบจำลองเรียนรู้ยากโดยเฉพาะข้อมูลที่มีขนาดใหญ่ต้องเลือกเคอร์เนลและพารามิเตอร์อย่างเหมาะสม โดยที่แบบจำลองนี้ไม่เหมาะกับข้อมูลมีส่วนที่ไม่เกี่ยวข้องมาก

ตารางที่ 2-1 (ต่อ)

แบบจำลอง	ข้อดี	ข้อจำกัด
แบบจำลองแบบเสริมแรง [14], [19], [23], [28], [29], [30], [31]	แบบจำลองมีความแม่นยำสูงโดยใช้การเสริมพลัง (Boosting) สามารถจัดการกับข้อมูลสูญหายได้ดีสามารถใช้ได้ทั้ง การจำแนกประเภท และการจัดกลุ่ม เพิ่มประสิทธิภาพจากแบบจำลองที่มีความสามารถต่ำ (Weak Learner) หลายตัว เหมาะสำหรับข้อมูลไม่สมดุล	แบบจำลองใช้ทรัพยากรสูงในการเรียนรู้ต้องการการปรับพารามิเตอร์จำนวนมากเพื่อเพิ่มประสิทธิภาพแบบจำลอง
โครงข่ายประสาทเทียม (Artificial Neural Network) [16], [29], [32]	มีความยืดหยุ่นสูงและสามารถเรียนรู้ความสัมพันธ์ซับซ้อนได้เหมาะกับข้อมูลขนาดใหญ่และรูปแบบที่ไม่ชัดเจน	ต้องการข้อมูลมากและใช้ทรัพยากรในการประมวลผลสูงปรับพารามิเตอร์ในแต่ละตัวยาก
กฎความสัมพันธ์ของข้อมูล (Association Rule) [20], [30]	ค้นหาความสัมพันธ์หรือรูปแบบที่ซ่อนอยู่ในข้อมูลได้ และสามารถเข้าใจง่าย เมื่อนำเสนอเป็นกฎ (if-then)	แบบจำลองใช้เวลามากเมื่อข้อมูลใหญ่แบบจำลองต้องเลือกค่าขั้นต่ำ support/confidence อย่างเหมาะสม และมีค่าผิดพลาดมากหากไม่มีการกรองข้อมูล
ระบบตามกฎเกณฑ์ [18], [25], [26]	แบบจำลองเข้าใจง่าย ใช้กฎที่ชัดเจน ตรงไปตรงมาเหมาะกับระบบผู้เชี่ยวชาญ (Expert system) และสามารถควบคุมได้ดี (Deterministic) เหมาะกับงานวิเคราะห์เชิงนโยบาย	แบบจำลองต้องการผู้เชี่ยวชาญสร้างกฎเองโดยที่ใช้การสร้างกฎจากมนุษย์ทำให้ไม่สามารถปรับตัวได้อัตโนมัติ หากมีข้อมูลที่เพิ่มขึ้นอาจทำให้แบบจำลองผิดพลาด

ตารางที่ 2-1 (ต่อ)

แบบจำลอง	ข้อดี	ข้อจำกัด
การวิเคราะห์อนุกรมเวลา [21], [22], [34]	การวิเคราะห์อนุกรมเวลา เหมาะ กับการพยากรณ์ข้อมูลเชิงเวลา สามารถอธิบายได้ดี เข้าใจง่าย	การวิเคราะห์อนุกรมเวลา ต้องการข้อมูลต่อเนื่องตาม ช่วงเวลาไม่สามารถจัดการข้อมูล ที่มีหลายตัวแปรได้ดีมากและไม่ เหมาะกับข้อมูลที่มีความสัมพันธ์ ไม่ที่ไม่เป็นเชิงเส้น

จากงานวิจัยที่เกี่ยวข้องทั้งหมดได้กล่าวการใช้การเรียนรู้ของเครื่อง ระบบตามกฎเกณฑ์ และการวิเคราะห์อนุกรมเวลา ในการศึกษาเกี่ยวกับอุบัติเหตุทางถนนมีบทบาทสำคัญในการวิเคราะห์และพยากรณ์ความรุนแรงของอุบัติเหตุ การเรียนรู้ของเครื่องสามารถวิเคราะห์ข้อมูลขนาดใหญ่และค้นหาแบบที่ซับซ้อนได้อย่างแม่นยำ โดยใช้ขั้นตอนวิธีต่าง ๆ เช่น ป่าการสุ่ม และ Naïve Bayes เป็นต้น ซึ่งเหมาะสำหรับการพยากรณ์และวิเคราะห์ความรุนแรงของอุบัติเหตุตามปัจจัยที่เกี่ยวข้อง เช่น พฤติกรรมการขับขี่และสภาพอากาศในขณะที่ ในอีกวิธีหนึ่งระบบตามกฎเกณฑ์ช่วยให้สามารถกำหนดกฎที่ชัดเจนในการจำแนกประเภทอุบัติเหตุ แต่มีข้อจำกัดในการจัดการกับข้อมูลที่ซับซ้อนและการสร้างโมเดลที่ยืดหยุ่นน้อยกว่า ส่วนการวิเคราะห์อนุกรมเวลาจะมุ่งเน้นที่การวิเคราะห์แนวโน้มในช่วงเวลาต่าง ๆ ซึ่งอาจไม่เหมาะสมในการพยากรณ์หรือวิเคราะห์ความรุนแรงของอุบัติเหตุในบริบทที่ต้องการข้อมูลเชิงลึกจากหลายปัจจัย ดังนั้นการเลือกใช้การเรียนรู้ของเครื่องจึงเหมาะสมในการวิจัยที่มีข้อมูลมากและมีปัจจัยที่หลากหลาย เนื่องจากมีความสามารถในการจัดการข้อมูลที่ซับซ้อนและการค้นหารูปแบบที่มีความสัมพันธ์กันได้อย่างมีประสิทธิภาพ อีกทั้งยังสามารถปรับปรุงขั้นตอนวิธีได้อย่างต่อเนื่อง ทำให้การวิเคราะห์มีความแม่นยำและเชื่อถือได้มากยิ่งขึ้น

บทที่ 3

วิธีการดำเนินการวิจัย

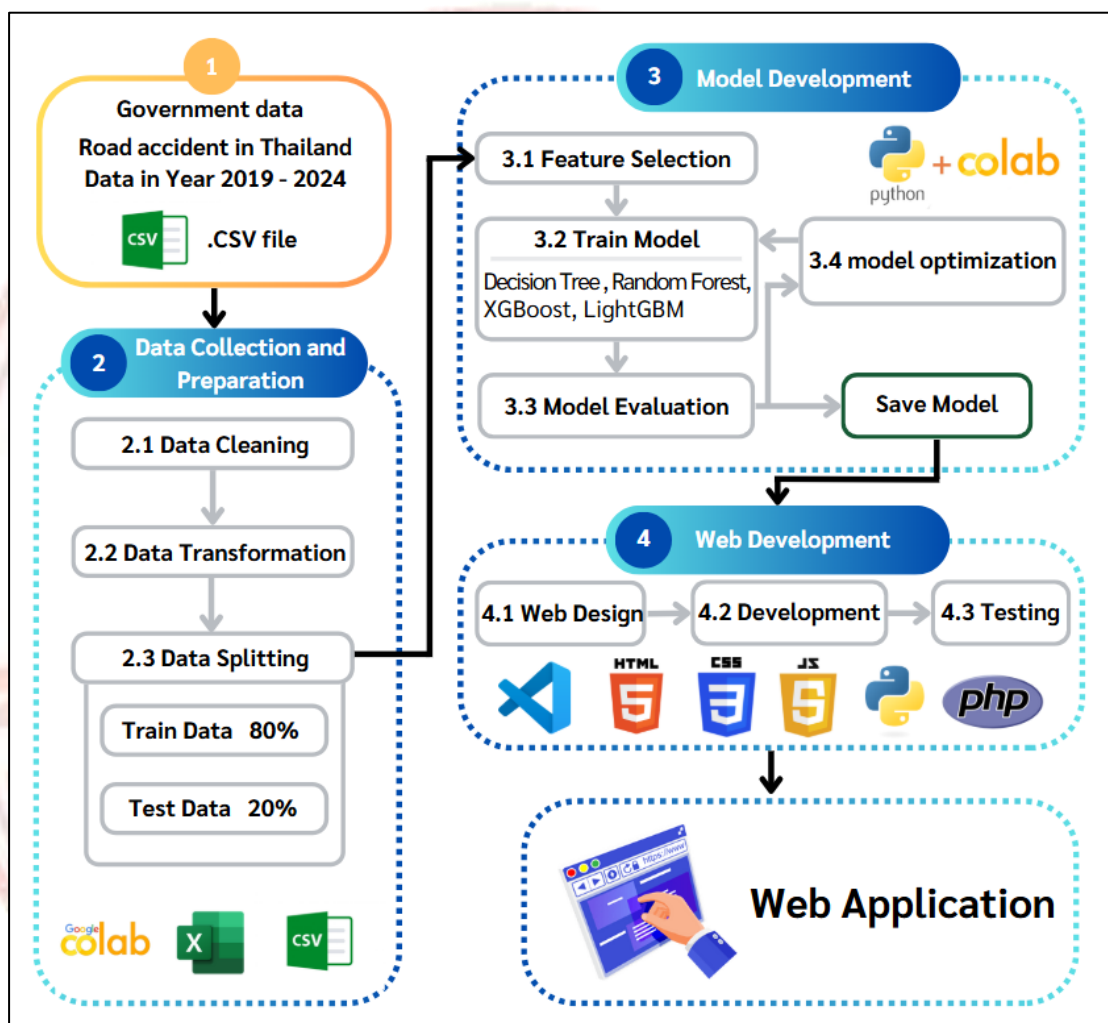
การศึกษาทฤษฎีและงานวิจัยที่เกี่ยวข้องพบว่าการเรียนรู้ของเครื่องแบบมีผู้สอนในรูปแบบการจำแนกประเภทมีความเหมาะสมต่อการจัดกลุ่มข้อมูล ผู้ค้นคว้าอิสระจึงนำมาประยุกต์ใช้กับการวิเคราะห์ระดับความรุนแรงของอุบัติเหตุทางถนน เนื่องจากข้อมูลอุบัติเหตุมีการจัดกลุ่มข้อมูลของระดับความรุนแรง โดยมีวิธีการดำเนินการค้นคว้าอิสระประกอบไปด้วยขั้นตอนดังต่อไปนี้

- 3.1 การศึกษาและวิเคราะห์ปัญหาที่เกี่ยวข้อง
- 3.2 การเก็บและรวบรวมข้อมูล
- 3.3 การพัฒนาแบบจำลอง
- 3.4 การวัดประสิทธิภาพแบบจำลอง
- 3.5 การพัฒนาระบบสารสนเทศ
- 3.6 การประเมินคุณภาพและความพึงพอใจระบบสารสนเทศ

3.1 การศึกษาและวิเคราะห์ปัญหาที่เกี่ยวข้อง

ผู้จัดทำการค้นคว้าอิสระได้ศึกษาและทบทวนวรรณกรรม (Literature Review) ที่เกี่ยวข้องกับปัญหาอุบัติเหตุทางถนนทั้งในประเทศไทยและต่างประเทศ พบว่าในประเทศทั่วโลกรวมทั้งประเทศไทยประสบปัญหาการเกิดอุบัติเหตุทางถนนส่งผลให้มีผู้เสียชีวิตอย่างมาก จากปัญหาดังกล่าวทำให้มีการนำเทคโนโลยีเข้ามาช่วยในการลดการเกิดอุบัติเหตุโดยเน้นการประยุกต์ใช้เทคโนโลยีที่ทันสมัยในการวิเคราะห์ข้อมูลและค้นหาข้อมูลเชิงลึกจากเหตุการณ์อุบัติเหตุทางถนนซึ่งประกอบด้วยเทคโนโลยีที่สำคัญ ได้แก่ การเรียนรู้ของเครื่อง ที่นำมาใช้ในการจำแนกประเภทอุบัติเหตุ ทำนายความเสี่ยง และระบุปัจจัยเสี่ยงที่อาจส่งผลต่อความรุนแรงของอุบัติเหตุ การนำระบบตามกฎเกณฑ์ มาวิเคราะห์ข้อมูลเชิงเหตุผลที่สามารถประมวลผลข้อมูลอุบัติเหตุ เพื่อสร้างกฎเกณฑ์ที่ช่วยในการป้องกันและลดอุบัติเหตุ ตลอดจนการใช้การวิเคราะห์อนุกรมเวลา เพื่อตรวจสอบแนวโน้มและรูปแบบของการเกิดอุบัติเหตุในช่วงเวลาต่าง ๆ ทั้งในระดับรายวัน รายเดือน หรือรายปี โดยวิธีการดังกล่าวมีผู้วิจัยที่หลากหลายได้ใช้หลายวิธีตามความเหมาะสมของข้อมูลและวัตถุประสงค์ของการวิจัย การประยุกต์ใช้วิธีการที่หลากหลายนี้สามารถนำมาปรับใช้ในบริบทของข้อมูลในประเทศไทยได้

ผู้จัดทำการค้นคว้าอิสระจึงได้นำการเรียนรู้ของเครื่องแบบการเรียนรู้แบบมีผู้สอนมาใช้ในการคว่ำอิสระเรื่องการวิเคราะห์ระดับความรุนแรงของอุบัติเหตุทางถนนโดยใช้การเรียนรู้ของเครื่องซึ่งเป็นขั้นตอนวิธีที่เหมาะสมกับชุดข้อมูลและสามารถนำไปพัฒนาระบบสารสนเทศได้ดังกรอบแนวคิดของการค้นคว้าอิสระดังภาพที่ 3-1



ภาพที่ 3-1 กรอบแนวคิดการวิจัย

จากภาพที่ 3-1 กรอบแนวคิดการวิจัย ซึ่งเป็นขั้นตอนทั้งหมดในการศึกษาค้นคว้าอิสระในครั้งนี้ โดยเริ่มการจากการรวบรวมปัญหาและข้อมูลที่เกี่ยวข้อง หลังจากนั้นจะเข้าสู่ขั้นตอนที่ 2 คือการจัดการกับข้อมูลที่ใช้ในการศึกษาการค้นคว้าอิสระ เมื่อจัดการกับข้อมูลเพื่อให้พร้อมแล้วจะเข้าสู่กระบวนการที่ 3 ในการพัฒนาแบบจำลองและหลังจากนั้นขั้นตอนที่ 4 การพัฒนาระบบสารสนเทศ ซึ่งจะทำได้ระบบสารสนเทศตามวัตถุประสงค์ของการศึกษาค้นคว้าอิสระในครั้งนี้

3.2 การเก็บรวบรวมข้อมูล

ผู้จัดทำการค้นคว้าอิสระได้ทำการรวบรวมข้อมูล โดยข้อมูลที่ใช้คือชุดข้อมูลอุบัติเหตุบนโครงข่ายถนนของกระทรวงคมนาคมตั้งแต่ปี พ.ศ. 2562 – 2567 รายการ ซึ่งเป็นข้อมูลการเกิดอุบัติเหตุทางถนนที่ทางกระทรวงคมนาคมได้เก็บรวบรวมเอาไว้มีจำนวนทั้งสิ้น 131,393 รายการ ซึ่งประกอบไปด้วยคุณลักษณะจำนวน 37 คุณลักษณะดังตารางที่ 3-2

ตารางที่ 3-1 คุณลักษณะข้อมูลอุบัติเหตุทางถนนปี พ.ศ. 2562 – 2567

ชื่อคุณลักษณะ	รายละเอียดข้อมูล	ประเภทข้อมูล	ตัวอย่าง
ปีที่เกิดเหตุ	ปีค.ศ. ที่เกิดอุบัติเหตุ	number	2023
วันที่เกิดเหตุ	เดือน/วัน/ปีค.ศ. ที่เกิดอุบัติเหตุ	date	9/20/2023
เวลา	เวลาที่เกิดอุบัติเหตุ	time	0:10
วันที่รายงาน	เดือน/วัน/ปีค.ศ. ที่รายงานการเกิดอุบัติเหตุ	date	10/30/2023
เวลาที่รายงาน	เวลาที่รายงานการเกิดอุบัติเหตุ	time	5:42
ACC_CODE	รหัสอุบัติเหตุจากระบบ TRAMS	number	7536059
หน่วยงาน	หน่วยงานที่รายงานการเกิดอุบัติเหตุ	text	กรมทางหลวง
รหัสสายทาง	หมายเลขสายทาง	text	1112
สายทาง	ชื่อสายทาง	text	พิจิตร - บ่อแก้ว
KM	กิโลเมตรที่เกิดอุบัติเหตุ	number	3
จังหวัด	จังหวัดที่เกิดอุบัติเหตุ	text	กำแพงเพชร
รถคันที่ 1	ประเภทของยานพาหนะที่เกิดอุบัติเหตุ	text	รถจักรยานยนต์
บริเวณที่เกิดเหตุ	บริเวณที่เกิดเหตุ/ลักษณะทาง	text	ทางตรง+ไม่มี ความลาดชัน
มูลเหตุ สันนิษฐาน	มูลเหตุสันนิษฐานในการเกิดอุบัติเหตุ	text	ขับเร็วเกิน อัตรากำหนด
ลักษณะการ เกิดเหตุ	ลักษณะการเกิดอุบัติเหตุ	text	ชนคนเดินเท้า
สภาพอากาศ	สภาพอากาศวันที่เกิดอุบัติเหตุ	text	แจ่มใส
LATITUDE	พิกัดละติจูด	number	15.97965797

ตารางที่ 3-1 (ต่อ)

LONGITUDE	พิกัดลองจิจูด	number	99.78557912
รถที่เกิดเหตุ	จำนวนรถที่เกิดอุบัติเหตุทั้งหมด	number	1
รถและคนที่เกิดเหตุ	จำนวนรถและคนเดินเท้าที่เกิดอุบัติเหตุทั้งหมด	number	1
รถจักรยานยนต์	จำนวนรถจักรยานยนต์ที่เกิดอุบัติเหตุ	number	0
รถสามล้อเครื่อง	จำนวนรถสามล้อเครื่องที่เกิดอุบัติเหตุ	number	0
รถยนต์นั่งส่วนบุคคล	จำนวนรถยนต์นั่งส่วนบุคคล/รถยนต์นั่งสาธารณะที่เกิดอุบัติเหตุ	number	1
รถตู้	จำนวนรถตู้ที่เกิดอุบัติเหตุ	number	1
รถปิคอัพโดยสาร	จำนวนรถปิคอัพโดยสารที่เกิดอุบัติเหตุ	number	1
รถโดยสารมากกว่า 4 ล้อ	จำนวนรถโดยสารมากกว่า 4 ล้อที่เกิดอุบัติเหตุ	number	0
รถปิคอัพบรรทุก 4 ล้อ	จำนวนรถปิคอัพบรรทุก 4 ล้อที่เกิดอุบัติเหตุ	number	1
รถบรรทุก 6 ล้อ	จำนวนรถบรรทุก 6 ล้อที่เกิดอุบัติเหตุ	number	0
รถบรรทุกไม่เกิน 10 ล้อ	จำนวนรถบรรทุกมากกว่า 6 ล้อ ไม่เกิน 10 ล้อที่เกิดอุบัติเหตุ	number	1
รถบรรทุกมากกว่า 10 ล้อ	จำนวนรถบรรทุกมากกว่า 10 ล้อ (รถพ่วง) ที่เกิดอุบัติเหตุ	number	0
รถอีแต่น	จำนวนรถอีแต่นที่เกิดอุบัติเหตุ	number	1
รถอื่น ๆ	จำนวนรถอื่น ๆ ที่เกิดอุบัติเหตุ	number	1
คนเดินเท้า	จำนวนคนเดินเท้าที่เกิดอุบัติเหตุ	number	0
ผู้เสียชีวิต	จำนวนผู้เสียชีวิต	number	0
ผู้บาดเจ็บสาหัส	จำนวนผู้บาดเจ็บสาหัส	number	0
ผู้บาดเจ็บเล็กน้อย	จำนวนผู้บาดเจ็บเล็กน้อย	number	1
รวมจำนวนผู้บาดเจ็บ	จำนวนผู้บาดเจ็บทั้งหมด	number	0

จากตารางที่ 3-1 จะได้ข้อมูลในรูปแบบของไฟล์ .csv โดยข้อมูลดังกล่าวไม่สามารถที่จะนำไปวิเคราะห์ด้วยแบบจำลองได้ เนื่องจากรูปแบบของข้อมูลยังไม่ถูกต้อง โดยยังมีข้อมูลที่เป็นช่องว่าง (Missing Values) และรูปแบบของข้อมูลยังเป็นในรูปแบบข้อความซึ่งยังไม่สามารถนำไปใช้งานได้ เพื่อให้ข้อมูลสามารถนำไปวิเคราะห์ในแบบจำลองการเรียนรู้ของเครื่องได้ มีขั้นตอนการดำเนินการดังต่อไปนี้

3.2.1 คัดเลือกคุณลักษณะที่เกี่ยวข้อง

จากคุณลักษณะของข้อมูลที่มีจำนวน 37 คุณลักษณะดังตารางที่ 3-2 พบว่ามีคุณลักษณะที่ไม่เกี่ยวข้องกับการเกิดอุบัติเหตุคือ วันที่รายงาน เวลาที่รายงาน และ ACC_CODE เนื่องจากเป็นข้อมูลการรายงานโดยเจ้าหน้าที่ที่เกี่ยวข้องเป็นผู้รายงานซึ่งไม่เกี่ยวข้องกับอุบัติเหตุที่เกิดขึ้น จึงนำคุณลักษณะดังกล่าวออกไม่นำไปวิเคราะห์ ทำให้มีข้อมูลคุณลักษณะทั้งหมดที่นำไปใช้งาน 34 คุณลักษณะ

3.2.2 การทำความสะอาดข้อมูล (Data Cleaning)

การทำความสะอาดข้อมูล เป็นหนึ่งในขั้นตอนที่สำคัญที่สุดในการเตรียมข้อมูลสำหรับการสร้างแบบจำลอง เนื่องจากข้อมูลของกระทรวงคมนาคมที่ได้รับมีปัญหาหลายประการ เช่น ค่าที่หายไป ค่าผิดปกติ และข้อผิดพลาดในข้อมูล (Incorrect Data) โดยต้องดำเนินการทำความสะอาดข้อมูลเพื่อไม่ให้เกิดข้อผิดพลาดเมื่อต้องนำข้อมูลเข้าแบบจำลอง ซึ่งดำเนินการดังนี้

3.2.2.1 การลบแถวที่ไม่มีข้อมูลจะดำเนินการลบแถวข้อมูลทั้งแถว โดยพิจารณาจากคุณลักษณะ รถคันที่ 1 มุลเหตุสันนิษฐาน ละติจูด (Latitude) และ ลองจิจูด (Longitude) ซึ่งเป็นค่าพื้นฐานที่ไม่สามารถแทนค่าได้เพราะเป็นการระบุถึงคุณลักษณะในการเกิดเหตุโดยลบข้อมูลไปทั้งสิ้น 4,047 รายการ

3.2.2.2 การแทนค่าข้อมูลที่หายไปโดยการใช้เค-เพื่อนบ้านใกล้สุดโดยจะพิจารณาจากค่า K ที่ใกล้ที่สุดเพื่อประมาณค่าที่หายไปได้แก่ สภาพอากาศ ปีที่เกิดเหตุ วันที่เกิดเหตุ และเวลาที่เกิดเหตุ โดยนำข้อมูลเข้าไปที่ Google Colab และใช้ เค-เพื่อนบ้านใกล้สุดเพื่อเติมค่าที่หายไป

จากกระบวนการทั้งหมดดังกล่าว เมื่อเริ่มต้นชุดข้อมูลอุบัติเหตุทางถนนตั้งแต่ปี พ.ศ. 2562 – 2567 ของกระทรวงคมนาคม มีทั้งหมด 131,393 รายการ และหลังจากผ่านกระบวนการการทำความสะอาดข้อมูลแล้วเหลือ 127,346 รายการ คิดเป็น 96.92% ของข้อมูลทั้งหมดก่อนทำความสะอาด

3.2.1 การแปลงข้อมูล (Data Transformation)

การแปลงข้อมูลเป็นขั้นตอนที่สำคัญในการเตรียมข้อมูลก่อนที่จะนำเข้าแบบจำลองเพื่อนำไปวิเคราะห์ข้อมูล โดยข้อมูลอุบัติเหตุตั้งแต่ปี พ.ศ. 2562 – 2567 ของกระทรวงคมนาคมนั้นคุณลักษณะที่เป็นตัวอักษรทั้งหมด 12 คุณลักษณะจากทั้งหมด 34 คุณลักษณะ จึงจำเป็นต้องแปลงข้อมูลใน 12

คุณลักษณะนี้ให้เป็นตัวเลขเพื่อสามารถนำไปใช้ในการทำแบบจำลองได้อย่างเหมาะสม โดยคุณลักษณะ มูลเหตุสันนิษฐาน บริเวณที่เกิดเหตุ และลักษณะการเกิดเหตุ เป็นคุณลักษณะข้อมูลที่มีความหลากหลายมากที่สุด โดยนำมาแปลงข้อมูลในรูปแบบ One-Hot Encoding ซึ่งจะทำให้ข้อมูลในรูปแบบหมวดหมู่ให้อยู่ในรูปแบบของเวกเตอร์ที่ประกอบด้วยเลข 0 และ 1 โดยแต่ละเวกเตอร์จะแทนแต่ละหมวดหมู่ ตรงตามตำแหน่งของข้อมูล โดยจะใช้การแปลงข้อมูลผ่าน Google Colab ด้วย `pd.get_dummies` ดังภาพที่ 3-2

```
import pandas as pd
from google.colab import files
uploaded = files.upload()
filename = list(uploaded.keys())[0]
df = pd.read_csv(filename)
df_encoded = pd.get_dummies(df,
                             columns=['มูลเหตุสันนิษฐาน', 'บริเวณที่เกิดเหตุ', 'ลักษณะการเกิดเหตุ'],
                             prefix=['มูลเหตุ', 'บริเวณ', 'ลักษณะ'])
print("ข้อมูลหลัง One-Hot Encoding:")
print(df_encoded.head())
output_filename = "encoded_data.csv"
df_encoded.to_csv(output_filename, index=False, encoding='utf-8-sig')
files.download(output_filename)
```

ภาพที่ 3-2 การแปลงข้อมูลแบบ One-Hot Encoding

โดยคุณลักษณะที่เหลือทั้งหมดจะถูกแปลงข้อมูลให้เป็นตัวเลขโดยใช้การแปลงข้อมูลในรูปแบบ Label Encoding ตามชนิดของข้อมูลโดยไม่มีการจัดกลุ่มที่เหมือนกัน

3.2.3 การเพิ่มข้อมูล (Feature Engineering)

เมื่อพิจารณาจากชุดข้อมูลอุบัติเหตุทางถนนจากข้อมูลคุณลักษณะทั้งหมดแล้วพบว่า สามารถเพิ่มคุณลักษณะจากข้อมูลที่มีอยู่ได้ ดังนี้

3.2.3.1 คุณลักษณะ “วันหยุดราชการ” สามารถเพิ่มได้จากการเปรียบเทียบวันที่ของคุณลักษณะวันที่เกิดเหตุกับวันหยุดราชการของไทย ตั้งแต่ปี พ.ศ. 2562 – 2567 โดยถ้าวันที่เกิดเหตุตรงกับวันหยุดราชการให้ใส่ 0 และถ้าไม่ตรงให้ใส่ 1

3.2.3.2 คุณลักษณะ “วันหยุดสุดสัปดาห์” สามารถเพิ่มได้จากการเปรียบเทียบวันที่ของคุณลักษณะวันที่เกิดเหตุตรงกับวันเสาร์และวันอาทิตย์ โดยถ้าวันที่เกิดเหตุตรงกับวันเสาร์หรือวันอาทิตย์ให้ใส่ 0 และถ้าไม่ตรงให้ใส่ 1

3.2.3.3 คุณลักษณะ “ช่วงเวลา” สามารถเพิ่มได้จากการเปรียบเทียบคุณลักษณะเวลาที่เกิดเหตุโดย ถ้าเวลาที่เกิดเหตุอยู่ในช่วง 00:00 – 06:00 น. ช่วงเวลาจะเป็น 0 ช่วงเวลา 06:01 – 12:00 น. ช่วงเวลาจะเป็น 1 ช่วงเวลา 12:01 – 18:00 น. ช่วงเวลาจะเป็น 2 ช่วงเวลา 18:01 – 21:00 น. ช่วงเวลาจะเป็น 3 และช่วงเวลา 21:01 – 23:59 น. ช่วงเวลาจะเป็น 4

3.2.3.4 คุณลักษณะ “ระดับความรุนแรง” สามารถเพิ่มได้จากการเปรียบเทียบคุณลักษณะ ผู้บาดเจ็บเล็กน้อย ผู้บาดเจ็บสาหัส และ ผู้เสียชีวิต โดยถ้ามีผู้เสียชีวิต 1 คนขึ้นไป ระดับความรุนแรงจะเป็น 1 ถ้ามีผู้บาดเจ็บเล็กน้อยหรือผู้บาดเจ็บสาหัสและไม่มีผู้เสียชีวิต ระดับความรุนแรงจะเป็น 0

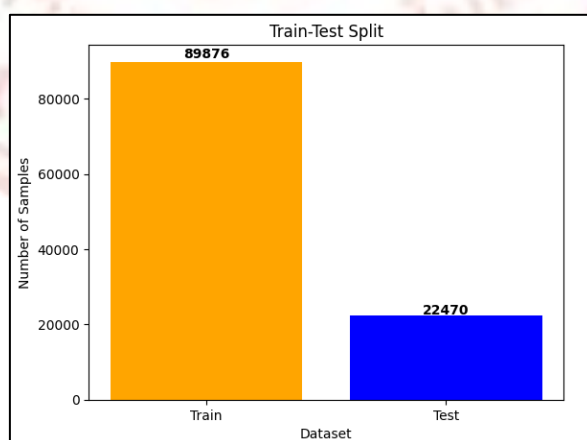
จากการเตรียมข้อมูลด้วย คัดเลือกคุณลักษณะที่เกี่ยวข้อง การทำความสะอาดข้อมูล การแปลงข้อมูล และการเพิ่มข้อมูล ทำให้ข้อมูลที่กำหนดสามารถนำไปวิเคราะห์ด้วยแบบจำลองของการเรียนรู้ของเครื่องได้ต่อไป

3.3 การพัฒนาแบบจำลอง

การพัฒนาแบบจำลองและการวัดประสิทธิภาพแบบจำลองเป็นการบวนการที่มีความสำคัญในการเรียนรู้ของเครื่อง ซึ่งในการค้นคว้าอิสระนี้ได้ใช้การเรียนรู้ของเครื่องแบบมีผู้สอน ซึ่งกระบวนการนี้จะเกิดขึ้นหลังจากการเก็บรวบรวมข้อมูลและการจัดการข้อมูลเสร็จเรียบร้อยแล้ว โดยมีขั้นตอนดังนี้

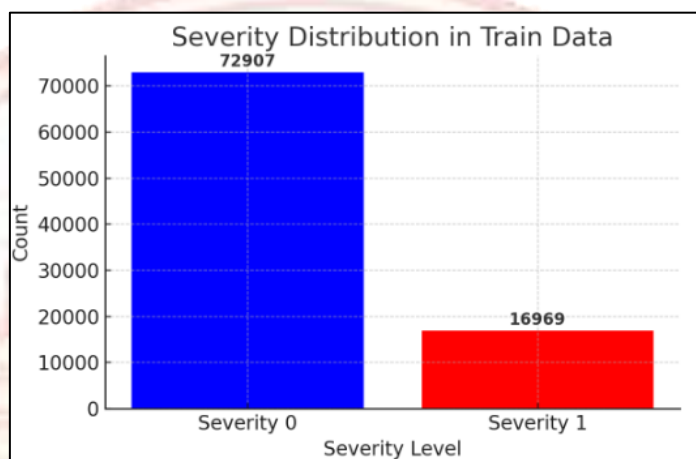
3.3.1 การแบ่งชุดข้อมูล (Train-Test Split)

การแบ่งชุดข้อมูล คือการเตรียมข้อมูลในการวิเคราะห์ด้วยแบบจำลองต่าง ๆ โดยแบ่งข้อมูลออกเป็น 2 ส่วนคือ ชุดข้อมูลสำหรับการฝึก (Training Set) และชุดข้อมูลสำหรับการทดสอบ (Testing Set) โดยการนำข้อมูลทั้งหมดมาแบ่งนั้นจะใช้ชุดข้อมูลสำหรับการสอนหรือฝึกแบบจำลองนั้นคือ ชุดข้อมูลสำหรับการฝึก โดยแบบจำลองจะหาความสัมพันธ์ของข้อมูล และอีกส่วนหนึ่งจะแบ่งเป็น ชุดข้อมูลสำหรับการทดสอบ โดยจะถูกเก็บเอาไว้เพื่อทดสอบประสิทธิภาพของแบบจำลองเมื่อได้ทำการฝึกแล้ว ในการค้นคว้าอิสระในครั้งนี้ได้แบ่งข้อมูลสำหรับการฝึก 80% และสำหรับชุดข้อมูลทดสอบ 20% ดังรูปภาพที่ 3-3



ภาพที่ 3-3 จำนวนชุดข้อมูลฝึกและชุดข้อมูลทดสอบ

จากการแบ่งข้อมูลเป็นชุดข้อมูลฝึกและชุดข้อมูลทดสอบ โดยมีข้อมูลคุณลักษณะ ระดับความรุนแรง (Severity) เป็นค่าเป้าหมาย (Target) ซึ่งสามารถตรวจสอบชุดข้อมูลการฝึกว่าการกระจายของคลาสในระดับความรุนแรงเพื่อให้รู้ถึงจำนวนที่ชุดข้อมูลฝึกในการเรียนรู้ของแต่ละคลาส ดังภาพที่ 3-5



ภาพที่ 3-4 จำนวนชุดข้อมูลฝึกในการแบ่งออกแต่ละคลาสในระดับความรุนแรง

จากภาพที่ 3-4 แสดงให้เห็นว่าชุดข้อมูลฝึกได้กระจายคลาสในระดับความรุนแรงเป็นระดับความรุนแรงที่ 0 ใช้ 72,907 ข้อมูล และระดับความรุนแรงที่ 1 ใช้ 16,969 ข้อมูล จากข้อมูลดังกล่าว พบว่าระดับความรุนแรงที่ 1 ใช้ข้อมูลในการฝึกน้อยที่สุด

3.3.2 การคัดเลือกคุณลักษณะ

การคัดเลือกคุณลักษณะจะทำการคัดเลือกคุณลักษณะที่เกี่ยวข้องหรือส่งผลกับแบบจำลองมากที่สุดโดยผู้จัดทำทำการค้นคว้าอิสระได้เลือกใช้ ลิสต์ แอ็บโซลูต ชรินเคจ แอนด์ ซีเล็คชัน โอเปอเรเตอร์ (Least Absolute Shrinkage and Selection Operator: Lasso) มีคุณสมบัติในการช่วยลดคุณลักษณะที่ไม่จำเป็นและไม่สำคัญในการสร้างแบบจำลองอีกทั้งยังช่วยลดปัญหาการเรียนรู้มากเกินไปซึ่งสามารถจำแนกคุณลักษณะที่เกี่ยวข้องได้จำนวน 22 คุณลักษณะ ได้แก่ ปีที่เกิดเหตุ (Year) วันหยุดสุดสัปดาห์ (Weekend) ช่วงเวลาที่เกิดเหตุ (Period) ละติจูด (Latitude) ลองจิจูด (Longitude) รถจักรยานยนต์ที่เกิดเหตุ (Motorcycle) รถสามล้อ (Tricycle) รถยนต์นั่งส่วนบุคคล (car) รถตู้ (Van) รถกระบะโดยสาร (Pickup Truck) รถบรรทุก 6 ล้อ (Truck 6 wheel) รถบรรทุกไม่เกิน 10 ล้อ (Trucks Less 10 wheels) รถบรรทุกมากกว่า 10 ล้อ (Trucks More 10 wheels) รถอื่น ๆ (Other Car) คนเดินเท้า (Pedestrians) หน่วยงานที่รับชอบ (Agency Encoded) จังหวัดที่

เกิดเหตุ (Province Encoded) รถที่เกิดเหตุ (Accident Car Encoded) มูลเหตุสันนิษฐาน (Assumption Encoded) สภาพอากาศ (Weather Encoded) และบริเวณที่เกิดเหตุ (Accident Area Encoded) ซึ่งจะนำเข้าสู่กระบวนการพัฒนาแบบจำลองต่อไป

3.3.3 การพัฒนาแบบจำลอง

เพื่อให้แบบจำลองสามารถเรียนรู้ชุดข้อมูลฝึกเพื่อหาความสัมพันธ์ระหว่างตัวแปร โดยการค้นคว้าอิสระนี้ได้ใช้ค่าเป้าหมายคือ ระดับความรุนแรง (Severity) ซึ่งมีคลาสระดับความรุนแรงที่ 0 และ 1 โดยนำคุณลักษณะที่เกี่ยวข้องโดยตรงกับผลลัพธ์ของ ระดับความรุนแรง คือ ผู้เสียชีวิต ผู้บาดเจ็บสาหัส และผู้บาดเจ็บเล็กน้อย ออกเพื่อป้องกันปัญหาการรั่วไหลของข้อมูล (Data Leakage) โดยในการค้นคว้าอิสระนี้ได้ใช้แบบจำลอง 4 แบบจำลองผ่าน Google Colab ดังนี้

3.3.3.1 ต้นไม้ตัดสินใจ (Decision Tree) ใช้ในการตัดสินใจจากข้อมูล โดยแสดงผลลัพธ์ของการตัดสินใจในแต่ละขั้นตอน กระบวนการสร้างต้นไม้จะเริ่มจากการเลือกคุณสมบัติที่สำคัญที่สุดในการแบ่งข้อมูล และทำการแบ่งข้อมูลออกเป็นกลุ่มย่อย ๆ ตามค่าของคุณสมบัติต้นไม้ตัดสินใจเป็นแบบที่เข้าใจง่ายและสามารถอธิบายได้ชัดเจน โดยการสร้างแบบจำลองด้วยต้นไม้ตัดสินใจได้ทำใน GoogleColab ดังภาพที่ 3-5

```
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.tree import DecisionTreeClassifier
from sklearn.metrics import confusion_matrix, classification_report
df = pd.read_csv(filename)
target_col = "Severity"
drop_cols = ["Deceased", "SeriouslyInjured", "MinorInjuries"]
feature_cols = [col for col in df.columns if col not in drop_cols + [target_col]]
X_train, X_test, y_train, y_test = train_test_split(df[feature_cols], df[target_col],
                                                    test_size=0.2, random_state=42, stratify=df[target_col])
model = DecisionTreeClassifier(random_state=42)
model.fit(X_train, y_train)
y_pred = model.predict(X_test)
```

ภาพที่ 3-5 สร้างแบบจำลองด้วยต้นไม้ตัดสินใจด้วย Google Colab

3.3.3.2 ป่าสุ่ม (Random Forest) เป็นการรวมกันของหลายต้นไม้ตัดสินใจซึ่งแต่ละต้นไม้จะได้รับการฝึกด้วยข้อมูลย่อยที่สุ่มเลือกมาจากชุดข้อมูลทั้งหมด และการตัดสินใจสุดท้ายจะพิจารณาจากผลการตัดสินใจของต้นไม้หลาย ๆ ต้นรวมกัน ซึ่งช่วยเพิ่มความแม่นยำและลดการเกิดการเรียนรู้มากเกินไปในการจำแนกประเภทและการคาดการณ์ โดยการสร้างแบบจำลองด้วยป่าสุ่มได้ทำใน Google Colab ดังภาพที่ 3-6

```

import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import confusion_matrix, classification_report
df = pd.read_csv(filename)
target_col = "Severity"
drop_cols = ["Deceased", "SeriouslyInjured", "MinorInjuries"]
feature_cols = [col for col in df.columns if col not in drop_cols + [target_col]]
X_train, X_test, y_train, y_test = train_test_split(df[feature_cols], df[target_col],
                                                    test_size=0.2, random_state=42, stratify=df[target_col])
model = RandomForestClassifier(n_estimators=100, random_state=42)
model.fit(X_train, y_train)
y_pred = model.predict(X_test)
conf_matrix = confusion_matrix(y_test, y_pred)
plt.figure(figsize=(5, 4))
sns.heatmap(conf_matrix, annot=True, fmt='d', cmap='Blues', xticklabels=['Pred 0', 'Pred 1'], yticklabels=['True 0', 'True 1'])

```

ภาพที่ 3-6 การสร้างแบบจำลองด้วยป่าการสุ่มใน Google Colab

3.3.2.3 เอ็กซ์ตรีม เกรเดียนต์ บูสต์ติง (Extreme Gradient Boosting: XGBoost) เป็นการปรับปรุงประสิทธิภาพของขั้นตอนวิธี ซึ่งจะสร้างการรวมผลลัพธ์ของต้นไม้ตัดสินใจเข้าด้วยกัน โดยในแต่ละขั้นตอนจะแก้ไขข้อผิดพลาดที่เกิดจากขั้นตอนวิธีก่อนหน้านี้ ด้วยการปรับน้ำหนักให้กับข้อมูลที่ถูกทำนายผิด ทำให้ในขั้นต่อไป ขั้นตอนวิธีจะมีความสามารถในการทำนายที่ดีขึ้น มีความสามารถในการจัดการกับความผันผวนในข้อมูลการฝึกมากเกินไปด้วยการลดความซับซ้อนของโมเดล (Regularization) ของข้อมูลและ ซึ่งช่วยให้มีค่าความถูกต้องและแม่นยำยิ่งขึ้น โดยการสร้างแบบจำลองด้วยเอ็กซ์ตรีม เกรเดียนต์ บูสต์ติง ได้ทำใน Google Colab ดังภาพที่ 3-7

```

import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from xgboost import XGBClassifier
from sklearn.metrics import confusion_matrix, classification_report
df = pd.read_csv(filename)
target_col = "Severity"
drop_cols = ["Deceased", "SeriouslyInjured", "MinorInjuries"]
feature_cols = [col for col in df.columns if col not in drop_cols + [target_col]]
X_train, X_test, y_train, y_test = train_test_split(df[feature_cols], df[target_col],
                                                    test_size=0.2, random_state=42, stratify=df[target_col])
model = XGBClassifier(n_estimators=100, learning_rate=0.1, random_state=42)
model.fit(X_train, y_train)
y_pred = model.predict(X_test)

```

ภาพที่ 3-7 การสร้างแบบจำลองด้วยเอ็กซ์ตรีม เกรเดียนต์ บูสต์ติง ใน Google Colab

3.3.2.4 โล้ท์ เกรเดียนต์ บูลตั่ง แมชชีน (Light Gradient Boosting Machine: LightGBM) เป็นขั้นตอนวิธีที่สร้างจากการรวมต้นไม้ตัดสินใจหลาย ๆ ต้น โดยจะค้นหาตัวแปรต้นที่มีผลต่อค่าตัวแปรตามที่น่าสนใจและแบ่งกลุ่มข้อมูลตามช่วงค่าของตัวแปรต้นนั้น ๆ หลังจากนั้นจะคำนวณค่าเฉลี่ยของตัวแปรตามในแต่ละกลุ่มเพื่อประเมินการลดความผิดพลาด โดยการสร้างแบบจำลองด้วยโล้ท์ เกรเดียนต์ บูลตั่ง แมชชีน ได้ทำใน Google Colab ดังภาพที่ 3-8

```
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from lightgbm import LGBMClassifier
from sklearn.metrics import confusion_matrix, classification_report
df = pd.read_csv(filename)
target_col = "Severity"
drop_cols = ["Deceased", "SeriouslyInjured", "MinorInjuries"]
feature_cols = [col for col in df.columns if col not in drop_cols + [target_col]]
X_train, X_test, y_train, y_test = train_test_split(df[feature_cols], df[target_col],
                                                    test_size=0.2, random_state=42, stratify=df[target_col])
model = LGBMClassifier(n_estimators=100, learning_rate=0.1, random_state=42)
model.fit(X_train, y_train)
y_pred = model.predict(X_test)
```

ภาพที่ 3-8 การสร้างแบบจำลองด้วยโล้ท์ เกรเดียนต์ บูลตั่ง แมชชีน ใน Google Colab

จากการพัฒนาแบบจำลองทั้ง 4 แบบจำลอง ทำให้ได้ผลลัพธ์ที่แตกต่างกันออกไป ตามความเหมาะสมของข้อมูลและแบบจำลองแต่ละประเภท โดยต้องนำผลการพัฒนาแบบจำลองทั้ง 4 แบบจำลองนำไปวัดประสิทธิภาพแบบจำลอง เพื่อวิเคราะห์หาแบบจำลองที่เหมาะสมกับการใช้งานที่สุด

3.4 การวัดประสิทธิภาพแบบจำลอง

การวัดประสิทธิภาพของแบบจำลองโดยใช้ ค่าความถูกต้อง ค่าความแม่นยำ ค่าความระลึกละ และ ค่าคะแนนเอฟหนึ่ง ควบคู่ไปกับตารางสับสน เพื่อเปรียบเทียบหาแบบจำลองที่ดีและเหมาะสมที่สุด สามารถนำไปพัฒนาหรือใช้งานได้อย่างมีประสิทธิภาพ โดยแบบจำลองจะต้องสัมพันธ์กับชุดข้อมูลซึ่งจำเป็นที่จะต้องเปรียบเทียบประสิทธิภาพระหว่างชุดข้อมูล โดยหากข้อมูลมีค่าความถูกต้อง หรือค่าอื่น ๆ ที่น้อยจนเกินไปแบบจำลองไม่สามารถเรียนรู้รูปแบบของข้อมูลได้ดีพอ ซึ่งอาจเกิดจากปัญหาข้อมูลที่ไม่สมดุล (Class Imbalance) หรือข้อมูลที่มีความกระจายตัวไม่ดี ส่งผลให้แบบจำลองมีแนวโน้มที่จะวิเคราะห์ค่าของกลุ่มที่มีจำนวนมากกว่าได้แม่นยำกว่ากลุ่มที่มีจำนวนน้อย ซึ่งหากการวัด

ประสิทธิภาพแบบจำลองพบว่า แบบจำลองเกิดชุดข้อมูลที่ไม่สมดุลกัน ทำให้ต้องแก้ไขปัญหาดังกล่าว ซึ่งการแก้ไขปัญหาคือทำให้ข้อมูลให้สมดุลกันสามารถใช้การทำการสุ่มตัวอย่างใหม่ (Resampling) และการปรับค่าโดยใช้คลาสสมดุล เพื่อปรับจำนวนชุดข้อมูลฝึกของแต่ละคลาสให้เท่ากัน ดังนี้

3.4.1 การเพิ่มจำนวนคลาสที่มีจำนวนน้อยให้มีมากยิ่งขึ้นด้วยการใช้ Synthetic Minority Over-sampling Technique (SMOTE) เป็นวิธีการเพิ่มตัวอย่างของคลาสส่วนน้อย (Minority Class) โดยใช้หลักการอินเทอร์โพล (Interpolation) ระหว่างจุดข้อมูลที่อยู่ใกล้กัน (Nearest Neighbors) ในคลาสส่วนน้อย ทำให้ข้อมูลในคลาสส่วนน้อยมีความหลากหลายมากขึ้น

3.4.2 การเพิ่มจำนวนคลาสที่มีจำนวนน้อยให้มีมากขึ้นด้วยการใช้ Synthetic Minority Over-sampling Technique ร่วมกับ Tomek Links (SMOTETomek) โดย SMOTE จะสร้างตัวอย่างเพิ่มในคลาสส่วนน้อยส่วน Tomek Links จะช่วยลบตัวอย่างที่อาจเป็นจุดที่ทำให้ขอบเขตการแบ่งคลาสไม่ชัดเจน (Noise) ทำให้ข้อมูลในคลาสส่วนน้อยมีความหลากหลายมากขึ้นโดยได้ผลลัพธ์ดังต่อไปนี้

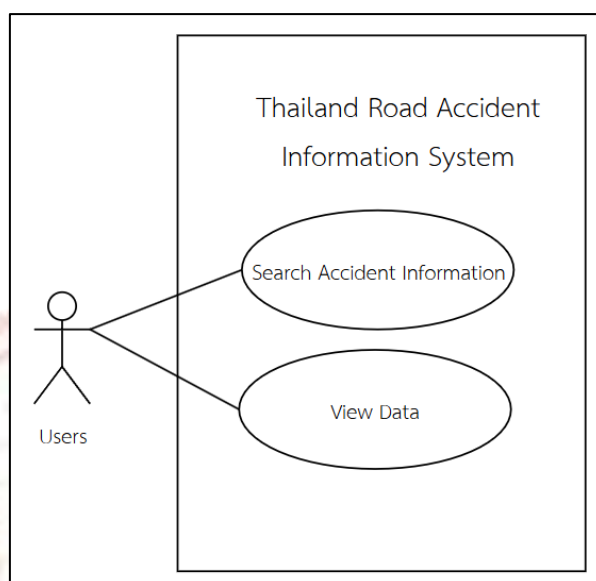
3.4.3 การปรับน้ำหนัก (weight) ของแต่ละคลาสให้อัตโนมัติ โดยน้ำหนักที่ให้แต่ละคลาสจะสัมพันธ์กับ สัดส่วนของข้อมูลรวมต่อข้อมูลในคลาสนั้นโดยเป็นการปรับโดยใช้คำสั่ง `class_weight='balanced'` ใน Google Colab โดยใช้ร่วมกับ Hyperparameter Tuning เพื่อให้แบบจำลองได้ใช้ข้อมูลฝึกอย่างเหมาะสมเพื่อทำให้ไม่เกิดการอคติ (Bias) ในคลาสที่มากกว่า

3.5 การพัฒนาระบบสารสนเทศ

3.5.1 การออกแบบระบบสารสนเทศ (Web Design)

การออกแบบระบบสารสนเทศเป็นการบวนการที่ต้องพัฒนาระบบให้ตอบสนองต่อความต้องการของผู้ใช้งานโดยแบ่งออกเป็น 2 กระบวนการดังนี้

3.5.1.1 การออกแบบฟังก์ชันการใช้งาน ผู้ออกแบบจะเป็นการกำหนดเริ่มต้นว่าระบบควรมีความสามารถอย่างไรบ้าง เพื่อให้สอดคล้องกับวัตถุประสงค์ของโครงการและความต้องการของผู้ใช้งาน โดยแผนภาพการทำงานผู้ใช้ระบบ (Use Case Diagram) จะแสดงภาพรวมของระบบโดยระบุผู้ใช้งาน (Actor) ที่จะใช้งานระบบ และใช้อะไร (Use Case) ภายในระบบสารสนเทศที่เกิดขึ้นดังภาพที่ 3-9



ภาพที่ 3-9 แผนภาพการทำงานของใช้ระบบสารสนเทศการแสดงผลข้อมูลอุบัติเหตุในประเทศไทย

จากภาพที่ 3-9 แผนภาพการทำงานของใช้ระบบสารสนเทศการแสดงผลข้อมูลอุบัติเหตุในประเทศไทยซึ่งมีผู้ใช้งานที่เกี่ยวข้องคือ ผู้ใช้งานทั่วไป โดยภายในระบบสารสนเทศประกอบไปด้วย การค้นหาข้อมูลอุบัติเหตุ (Search Accident Information) และการแสดงข้อมูลแผนภูมิ (View Data Via Graph Display) ซึ่งสามารถอธิบายแผนภาพการทำงานของใช้ระบบ คือ การค้นหาข้อมูลอุบัติเหตุ เป็นกระบวนการที่ผู้ใช้งานส่งเงื่อนไขการค้นหาคือ ปีที่เกิดเหตุเริ่มต้น ปีที่เกิดเหตุสิ้นสุด จังหวัดที่เกิดเหตุ และรถที่เกิดเหตุมาที่ระบบ ระบบจะทำการส่งเงื่อนไขการค้นหาไปที่ datagov.mot.go.th เพื่อเรียกข้อมูลที่ถูกต้องมาแล้วผล โดยระบบจะส่งพิกัดจากข้อมูลของ datagov.mot.go.th ไปที่ Openstreetmap เพื่อขอสิทธิ์ในการใช้งาน หลังจากนั้นจะแสดงผลผ่านหน้าจอในรูปแบบแผนภูมิต่าง ๆ ดังตารางที่ 3-2

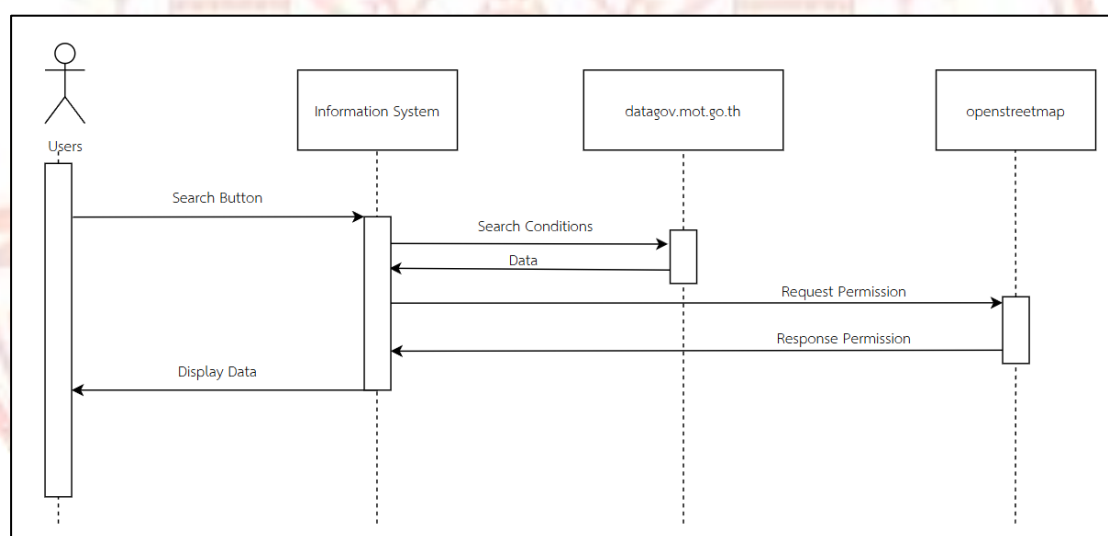
ตารางที่ 3-2 การทำงานของผู้ใช้งานระบบสารสนเทศส่วนการค้นหาข้อมูล

ชื่อユスเคス	หัวข้อ	รายละเอียด
Search Accident Information	ผู้ใช้งาน	ผู้ใช้งานทั่วไป
	คำอธิบาย	ตรวจสอบความถูกต้องของข้อมูลการเกิดเหตุ ผ่านกราฟที่แสดงผลต่าง ๆ
	การใช้งาน	1. ผู้ใช้งานสามารถเข้าใช้งานได้โดยไม่ต้อง Login 2. ผู้ใช้งานไปที่หัวข้อ “ค้นหาข้อมูล”

ตารางที่ 3-2 (ต่อ)

ชื่อยูสเคส	หัวข้อ	รายละเอียด
		3. ผู้ใช้งานระบุ ปีเริ่มต้น ปีสิ้นสุด หรือ จังหวัด หรือ ประเภทที่ เกิดเหตุ 4. ผู้ใช้กดปุ่ม “ค้นหา” 5. ระบบส่งเงื่อนไขไปที่ Datagov.mot.go.th 6. ระบบส่งละติจูด และ ลองจิจูด ไปที่ Openstreetmap เพื่อวาง พิกัดจุดลงบนแผนที่

จากตารางที่ 3-2 สามารถนำมาออกแบบแผนภาพซีควเอนซ์การค้นหาข้อมูลอุบัติเหตุได้
ดังภาพที่ 3-10



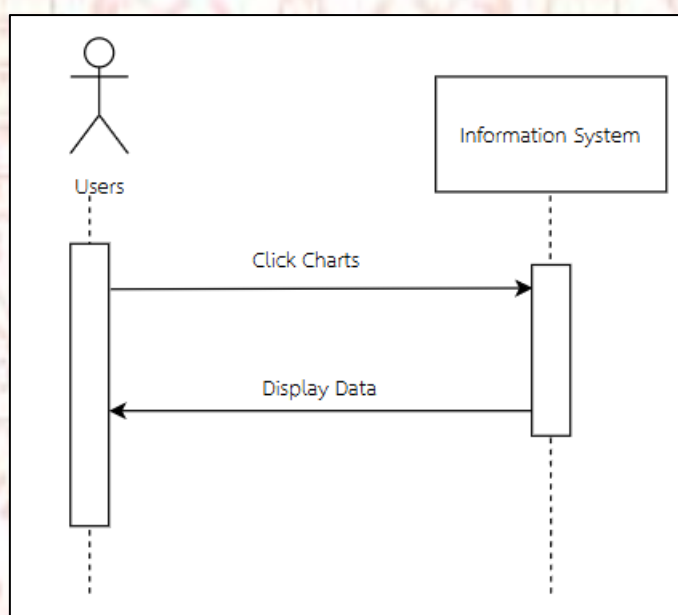
ภาพที่ 3-10 ผังแสดงเหตุการณ์การค้นหาข้อมูลอุบัติเหตุในระบบสารสนเทศ

จากการทำงานของระบบสารสนเทศในการค้นหาข้อมูลอุบัติเหตุจะสามารถตรวจสอบเส้นทางการเดินทางในการแสดงข้อมูลอุบัติเหตุผ่านเส้นทางที่ผู้ใช้งานเลือกได้ ดังตารางที่ 3-3

ตารางที่ 3-3 รายละเอียดการทำงานของระบบสารสนเทศในการแสดงผลข้อมูล

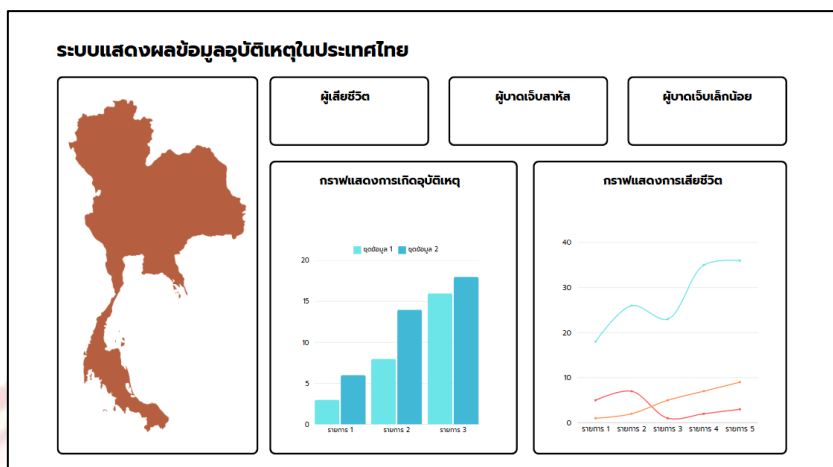
ชื่อยูสเคส	หัวข้อ	รายละเอียด
View Data	ผู้ใช้งาน	ผู้ใช้งานทั่วไป
	คำอธิบาย	ผู้ใช้งานพิจารณาข้อมูลผ่านระบบสารสนเทศ
	การใช้งาน	ผู้ใช้งานสามารถดูรายละเอียดข้อมูลอุบัติเหตุได้ผ่านแผนภูมิในรูปแบบต่าง ๆ โดยมีรายละเอียดคือ จุดเกิดเหตุ ข้อมูลการเกิดเหตุ ข้อมูลสภาพอากาศที่เกิดเหตุ และข้อมูลรถที่เกิดเหตุ

จากตารางที่ 3-3 สามารถนำมาออกแบบแผนภาพซีควเอนซ์การแสดงผลข้อมูลอุบัติเหตุได้
ดังภาพที่ 3-11



ภาพที่ 3-11 ผังแสดงเหตุการณ์การแสดงผลข้อมูลอุบัติเหตุ

3.5.1.2 การออกแบบหน้าจอแสดงผล เป็นหนึ่งในกระบวนการสำคัญของการพัฒนาระบบสารสนเทศ เนื่องจากหน้าจอเป็นจุดที่ผู้ใช้งานโต้ตอบกับระบบโดยตรง การออกแบบที่เหมาะสมและเป็นมิตร (User-friendly) จะช่วยให้ผู้ใช้งานสามารถทำงานได้อย่างมีประสิทธิภาพและเข้าใจการทำงานของระบบได้ง่าย ดังภาพที่ 3-12



ภาพที่ 3-12 ออกแบบหน้าจอแสดงผลข้อมูลระบบข้อมูลอุบัติเหตุในประเทศไทย

จากภาพที่ 3-14 เป็นการออกแบบหน้าจอแสดงผลโดยเป็นการแสดงผลในส่วนเริ่มต้นของหน้าระบบ โดยการแสดงผลในแผนที่จะสร้างจุดตามพิกัดละติจูด และลองจิจูด ที่เกิดเหตุลงบนแผนที่ รวมทั้งแสดงข้อมูลรายละเอียดจำนวนการอุบัติเหตุเปรียบเทียบรายปี

3.6 การประเมินคุณภาพและความพึงพอใจระบบสารสนเทศ

การพัฒนาระบบสารสนเทศ ผู้จัดทำการค้นคว้าอิสระได้เลือกใช้แนวทางการทดสอบระบบแบบแบล็กบ็อกซ์ (Black-box Testing) เพื่อประเมินความถูกต้องและความพร้อมในการใช้งานของระบบ โดยลักษณะของการทดสอบนี้เป็นการมองระบบสารสนเทศจากมุมมองของผู้ใช้งาน โดยไม่พิจารณาและไม่ตรวจสอบการเขียนชุดคำสั่งภายใน หรือขั้นตอนการประมวลผลของระบบสารสนเทศ โดยมุ่งเน้นไปที่การใช้งานของระบบสารสนเทศที่ผู้ใช้กรอกหรือเลือกข้อมูลต่าง ๆ (Input) แล้ว ระบบสามารถประมวลผลและแสดงผลลัพธ์ (Output) ได้ถูกต้องหรือไม่ เช่น เมื่อต้องการดูสถิติอุบัติเหตุในปีใดปีหนึ่ง ระบบจะแสดงข้อมูลรายปี รายจังหวัด หรือรายละเอียดการเกิดอุบัติเหตุอย่างถูกต้องหรือไม่ โดยการทำสอบจะแบ่งเป็น 2 ลักษณะ ดังนี้

3.6.1 การประเมินคุณภาพระบบสารสนเทศโดยผู้เชี่ยวชาญทางด้านคอมพิวเตอร์หรือระบบสารสนเทศจำนวน 3 ท่าน ซึ่งใช้แบบทดสอบประสิทธิภาพระบบสารสนเทศสำหรับผู้เชี่ยวชาญระบบสารสนเทศการแสดงผลข้อมูลอุบัติเหตุทางถนนในประเทศไทย โดยแบ่งออกเป็น 3 ตอนคือ ข้อมูลทั่วไปของผู้เชี่ยวชาญ ความคิดเห็นของผู้เชี่ยวชาญเกี่ยวกับคุณภาพของระบบสารสนเทศการแสดงผลข้อมูลอุบัติเหตุทางถนนในประเทศไทย และข้อเสนอแนะและข้อแนะนำอื่นๆ โดยแบบสอบถามความคิดเห็นนั้นจะแบ่งออกเป็น 3 ด้าน คือ ด้านความสามารถของระบบสารสนเทศตามความต้องการของ

ผู้ใช้งาน (Functional Requirement Test) ด้านการทำงานตามฟังก์ชัน (Functional Test) และ ด้านความสะดวกและง่ายต่อการใช้งานระบบสารสนเทศ (Usability Test)

3.6.2 การประเมินความพึงพอใจระบบสารสนเทศโดยผู้ใช้งานทั่วไปจำนวน 30 ท่าน ซึ่งแบบประเมินความพึงพอใจระบบสารสนเทศสำหรับผู้ใช้งานทั่วไประบบสารสนเทศการแสดงผลข้อมูลอุบัติเหตุทางถนนในประเทศไทยโดยแบ่งออกเป็น 3 ตอนคือ ข้อมูลทั่วไปของผู้ใช้งานทั่วไป ความคิดเห็นของผู้ใช้งานเกี่ยวกับความพึงพอใจของระบบสารสนเทศการแสดงผลข้อมูลอุบัติเหตุทางถนนในประเทศไทย และข้อเสนอแนะ โดยแบบสอบถามความคิดเห็นนั้นจะแบ่งออกเป็น 3 ด้าน คือ ด้านความสามารถของระบบสารสนเทศตามความต้องการของผู้ใช้งาน (Functional Requirement Test) ด้านการทดสอบการทำงานตามฟังก์ชันงาน (Functional Test) และ ด้านความสะดวกและง่ายต่อการใช้งานระบบสารสนเทศ (Usability Test)

นอกจากนี้เกณฑ์การให้คะแนนแบบทดสอบคุณภาพและแบบประเมินความพึงพอใจ ผู้จัดทำ การค้นคว้าอิสระได้ใช้แบบประเมินโดยอิงตามมาตรวัดไลเกิร์ต (Likert Scale) โดยแบ่งเกณฑ์การให้ คะแนนเชิงคุณภาพและเชิงปริมาณ 5 ระดับ โดยแสดงดังตารางที่ 3-4 และ ตารางที่ 3-5

ตารางที่ 3-4 เกณฑ์การให้คะแนนแบบประเมินคุณภาพของผู้เชี่ยวชาญ

เกณฑ์การให้คะแนน		การแปลผล
เชิงคุณภาพ	เชิงปริมาณ	
5	4.51 – 5.00	ระบบที่พัฒนาขึ้นมีคุณภาพมากที่สุด
4	3.51 – 4.50	ระบบที่พัฒนาขึ้นมีคุณภาพมาก
3	2.51 – 3.50	ระบบที่พัฒนาขึ้นมีคุณภาพปานกลาง
2	1.51 – 2.50	ระบบที่พัฒนาขึ้นมีคุณภาพน้อย
1	1.00 – 1.50	ระบบที่พัฒนาขึ้นมีคุณภาพน้อยมาก

ตารางที่ 3-5 เกณฑ์การให้คะแนนแบบประเมินความพึงพอใจผู้ใช้งานทั่วไป

เกณฑ์การให้คะแนน		การแปลผล
เชิงคุณภาพ	เชิงปริมาณ	
5	4.51 – 5.00	มีความพึงพอใจต่อระบบที่พัฒนาขึ้นมากที่สุด
4	3.51 – 4.50	มีความพึงพอใจต่อระบบที่พัฒนาขึ้นมาก
3	2.51 – 3.50	มีความพึงพอใจต่อระบบที่พัฒนาขึ้นปานกลาง
2	1.51 – 2.50	มีความพึงพอใจต่อระบบที่พัฒนาขึ้นน้อย
1	1.00 – 1.50	มีความพึงพอใจต่อระบบที่พัฒนาขึ้นน้อยที่สุด

เมื่อผู้เชี่ยวชาญและผู้ใช้งานทั่วไปได้ทำแบบประเมินแล้วเสร็จ ผู้จัดทำการค้นคว้าอิสระได้นำผลคะแนนที่ได้ มาวิเคราะห์ด้วยสถิติเชิงพรรณนา (Descriptive Statistics) ได้แก่ การหาค่าเฉลี่ย (Mean) ดังสมการที่ 3-1

$$\bar{x} = \frac{\sum x}{n} \quad (3-1)$$

โดยที่ $\bar{x}$ คือ ค่าเฉลี่ยผลการทดสอบประสิทธิภาพ
 $\sum x$ คือ ผลรวมทั้งหมดของการทดสอบประสิทธิภาพ
 n คือ จำนวนข้อคำถาม

นอกจากนี้ยังมีส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) เพื่อวัดการกระจายของการตอบแบบทดสอบประสิทธิภาพและความพึงพอใจซึ่งสามารถคำนวณได้ ดังสมการที่ 3-2

$$S.D. = \frac{\sqrt{\sum (x - \bar{x})^2}}{n-1} \quad (3-2)$$

โดยที่ $S.D.$ คือ ส่วนเบี่ยงเบนมาตรฐานผลการทดสอบประสิทธิภาพ
 x คือ คะแนนของข้อคำถาม
 $\bar{x}$ คือ ค่าเฉลี่ยผลการทดสอบประสิทธิภาพ
 n คือ จำนวนข้อคำถาม

โดยการหาค่าเฉลี่ยและส่วนเบี่ยงเบนมาตรฐานนั้น เพื่อให้ผู้ค้นคว้าอิสระสามารถทราบถึงแนวโน้มความคิดเห็นของผู้เชี่ยวชาญและผู้ใช้งาน เพื่อนำไปปรับปรุงระบบสารสนเทศให้สอดคล้องต่อความต้องการของผู้ใช้งาน

บทที่ 4

ผลการดำเนินงานวิจัย

การวิเคราะห์ระดับความรุนแรงของอุบัติเหตุทางถนนโดยใช้การเรียนรู้ของเครื่อง ได้ดำเนินการปรับปรุงประสิทธิภาพแบบจำลองและวัดประสิทธิภาพจากค่าความถูกต้อง ค่าความแม่นยำ ค่าความระลึก และค่าคะแนนเอฟหนึ่ง โดยการเปรียบเทียบแบบจำลองทั้ง 4 แบบจำลองได้แก่ ต้นไม้ตัดสินใจ ป่าสุ่ม เอ็กซ์ตรีม เกรเดียนต์ บูสตีง และโลท เกรเดียนต์ บูสตีง แมชชีน เพื่อหาผลลัพธ์ที่ดีที่สุด ดังนี้

4.1 ผลการพัฒนาแบบจำลอง

4.2 ผลการพัฒนาระบบสารสนเทศ

4.3 ผลการประเมินคุณภาพและผลการประเมินความพึงพอใจต่อระบบสารสนเทศ

4.1 ผลการพัฒนาแบบจำลอง

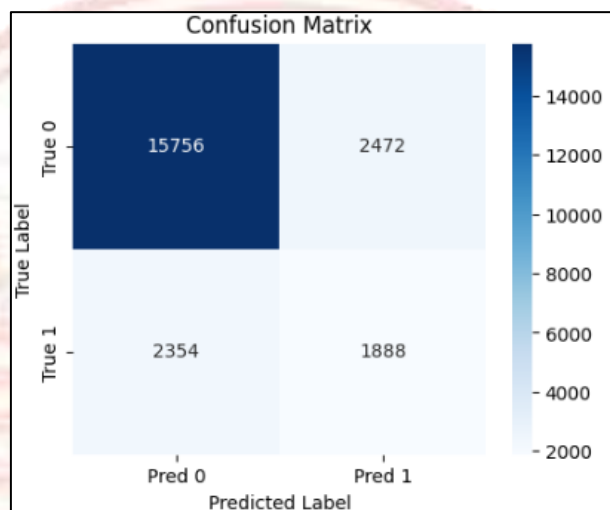
จากการพัฒนาแบบจำลองทั้ง 4 แบบจำลอง ได้แก่ ต้นไม้ตัดสินใจ ป่าสุ่ม เอ็กซ์ตรีม เกรเดียนต์ บูสตีง และโลท เกรเดียนต์ บูสตีง แมชชีน ในการวิเคราะห์ระดับความรุนแรงของอุบัติเหตุทางถนน ดังต่อไปนี้

4.1.1 ผลการพัฒนาแบบจำลองแบบไม่ปรับปรุงประสิทธิภาพ โดยเป็นการพัฒนาแบบจำลองผ่านการนำข้อมูลเข้ากระบวนการการเรียนรู้ของเครื่องทั้ง 4 แบบจำลอง โดยไม่มีการเพิ่มประสิทธิภาพแบบจำลอง โดยให้ผลลัพธ์การพัฒนาดังตารางที่ 4-1

ตารางที่ 4-1 ผลการพัฒนาแบบจำลองแบบไม่เพิ่มประสิทธิภาพ

แบบจำลอง	ระดับความรุนแรง	ค่าความถูกต้อง	ค่าความแม่นยำ	ค่าความระลึก	ค่าคะแนนเอฟหนึ่ง
ต้นไม้ตัดสินใจ	0	0.79	0.86	0.87	0.87
	1		0.44	0.44	0.44
ป่าสุ่ม	0	0.85	0.87	0.95	0.91
	1		0.64	0.40	0.49
เอ็กซ์ตรีม เกรเดียนต์ บูสตีง	0	0.85	0.87	0.96	0.91
	1		0.67	0.41	0.50
โลท เกรเดียนต์ บูสตีง แมชชีน	0	0.85	0.87	0.96	0.91
	1		0.69	0.40	0.50

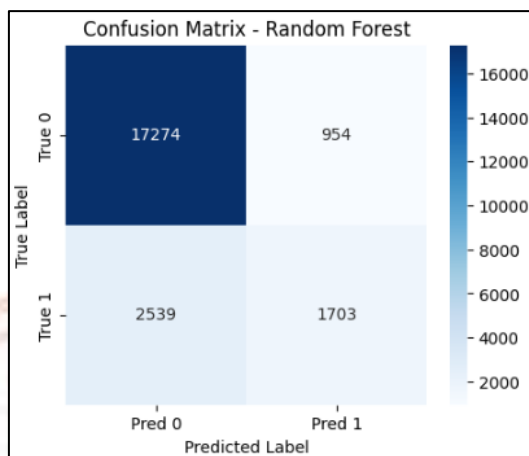
จากตารางที่ 4-1 พบว่าต้นไม้ตัดสินใจให้ค่าความถูกต้องอยู่ที่ 79% และแบบจำลองสามารถจำแนกในคลาส 0 ได้ดีโดยดูที่ค่า ความแม่นยำ และค่าความระลึกลับ ที่ค่อนข้างสูงแต่การจำแนกในคลาส 1 นั้นยังไม่ดีพอหรือยังมีปัญหาเนื่องจาก ค่าความแม่นยำ และค่าความระลึกลับนั้นมีค่าที่ต่ำมากซึ่งอาจเกิดชุดข้อมูลที่ไม่สมดุลได้



ภาพที่ 4-1 ตารางสับสนของต้นไม้ตัดสินใจ

จากภาพที่ 4-1 แสดงให้เห็นถึงประสิทธิภาพของแบบจำลองที่ใช้ในการจำแนกประเภทข้อมูลออกเป็นสองกลุ่ม ได้แก่ คลาส 0 และคลาส 1 ซึ่งจากผลลัพธ์แบบจำลองสามารถวิเคราะห์ค่าของคลาส 0 ได้อย่างแม่นยำ โดยมีจำนวน True Negative (TN) อยู่ที่ 15,756 รายการ ขณะที่จำนวน False Positive (FP) จำนวน 2,472 รายการ ซึ่งยังถือว่ามีความผิดพลาดอยู่บ้าง และ แบบจำลองจำแนกคลาส 1 ผิดพลาดสูงโดยมี False Negative (FN) 2,354 รายการ ซึ่งถึงแบบจำลองมีแนวโน้มที่จะมองข้ามตัวอย่างที่เป็นคลาส 1 ไปเป็นคลาส 0 มาก และ True Positive (TP) 1,888 รายการ ซึ่งเป็นจำนวนที่ต่ำเมื่อเทียบกับคลาส 0 ซึ่งแสดงว่าแบบจำลองนี้เกิดชุดข้อมูลที่ไม่สมดุลต้องได้รับการปรับปรุงแบบจำลอง

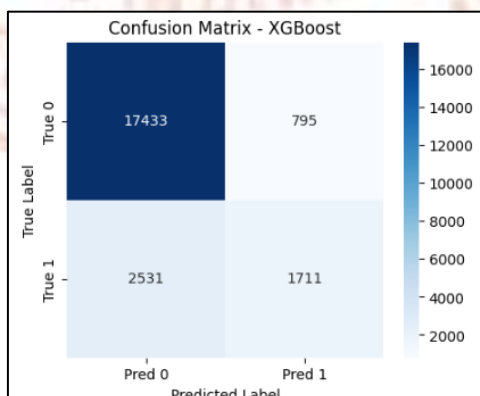
ซึ่งป่าการสุ่มให้ค่าความถูกต้องอยู่ที่ 84% โดยแบบจำลองป่าการสุ่มมีผลลัพธ์ที่ดีกว่าต้นไม้ตัดสินใจ โดยแบบจำลองสามารถจำแนกในคลาส 0 ได้ดีโดยดูที่ค่า ความแม่นยำ และค่าความระลึกลับ ที่ค่อนข้างสูงแต่การจำแนกในคลาส 1 นั้นยังไม่ดีพอหรือยังมีปัญหาเนื่องจาก ค่า ความแม่นยำ และค่าความระลึกลับนั้นมีค่าที่ต่ำถึงแม้จะมากกว่าต้นไม้ตัดสินใจก็ตาม ซึ่งอาจเกิดชุดข้อมูลที่ไม่สมดุลได้ โดยสามารถดูได้จากตารางสับสนดังภาพที่ 4-2



ภาพที่ 4-2 ตารางสับสนของป่าการสุ่ม

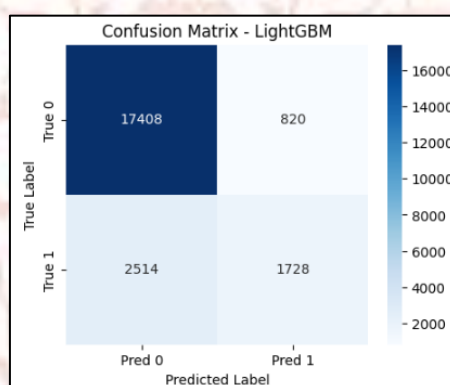
จากภาพที่ 4-2 แสดงให้เห็นถึงประสิทธิภาพของแบบจำลองที่ใช้ในการจำแนกประเภทข้อมูลออกเป็นสองกลุ่ม ได้แก่ คลาส 0 และคลาส 1 ซึ่งจากผลลัพธ์แบบจำลองสามารถวิเคราะห์ค่าของคลาส 0 ได้อย่างแม่นยำ โดยมีจำนวน True Negative (TN) อยู่ที่ 17,274 รายการ ขณะที่จำนวน False Positive (FP) จำนวน 954 รายการ ซึ่งยังถือว่ามีความผิดพลาดอยู่ และ แบบจำลองจำแนกคลาส 1 ผิดพลาดสูงโดยมี False Negative (FN) 2,539 รายการ ซึ่งถึงแบบจำลองมีแนวโน้มที่จะมองข้ามตัวอย่างที่เป็นคลาส 1 ไปเป็นคลาส 0 มาก และ True Positive (TP) 1,703 รายการ ซึ่งเป็นจำนวนที่ต่ำเมื่อเทียบกับคลาส 0 ซึ่งแสดงว่าแบบจำลองนี้เกิดชุดข้อมูลที่ไม่สมดุลต้องได้รับการปรับปรุงแบบจำลอง

นอกจากนี้เอ็กซ์ตรีม เกรเดียนต์ บูสตั๊ง ให้ค่าความถูกต้องอยู่ที่ 85% โดยแบบจำลองเอ็กซ์ตรีม เกรเดียนต์ บูสตั๊ง มีผลลัพธ์ที่ดีกว่าต้นไมัดัดสินใจเล็กน้อย โดยแบบจำลองสามารถจำแนกในคลาส 0 ได้ดีโดยดูที่ค่า ความแม่นยำ และค่าความระลึก ที่ค่อนข้างสูงแต่การจำแนกในคลาส 1 นั้นยังไม่ดีพอหรือยังมีปัญหาเนื่องจาก ค่า ความแม่นยำ และค่าความระลึกนั้นมีค่าที่ต่ำ ซึ่งอาจเกิดชุดข้อมูลที่ไม่สมดุลได้ โดยสามารถดูได้จากตารางสับสนดังภาพที่ 3-3



ภาพที่ 4-3 ตารางสับสนของเอ็กซ์ตรีม เกรเดียนต์ บูสตั๊ง

จากภาพที่ 4-3 แสดงให้เห็นถึงประสิทธิภาพของแบบจำลองที่ใช้ในการจำแนกประเภทข้อมูลออกเป็นสองกลุ่ม ได้แก่ คลาส 0 และคลาส 1 ซึ่งจากผลลัพธ์แบบจำลองสามารถวิเคราะห์ค่าของคลาส 0 ได้อย่างแม่นยำ โดยมีจำนวน True Negative (TN) อยู่ที่ 17,433 รายการ ขณะที่จำนวน False Positive (FP) จำนวน 795 รายการ ซึ่งยังถือว่ามีความผิดพลาดอยู่บ้าง และ แบบจำลองจำแนกคลาส 1 ผิดพลาดสูงโดยมี False Negative (FN) 2,531 รายการ ซึ่งถึงแบบจำลองมีแนวโน้มที่จะมองข้ามตัวอย่างที่เป็นคลาส 1 ไปเป็นคลาส 0 มาก และ True Positive (TP) 1,711 รายการ ซึ่งเป็นจำนวนที่ต่ำเมื่อเทียบกับคลาส 0 ซึ่งแสดงว่าแบบจำลองนี้เกิดชุดข้อมูลที่ไม่สมดุลต้องได้รับการปรับปรุงแบบจำลอง และพบว่าไลต์ เกรเดียนต์ บูสติง แมชชีน ให้ค่าความถูกต้องอยู่ที่ 85 % โดยแบบจำลอง ไลต์ เกรเดียนต์ บูสติง แมชชีน ให้ค่าผลลัพธ์ที่ใกล้เคียงกับเอ็กซ์ตรีม เกรเดียนต์ บูสติง โดยแบบจำลองสามารถจำแนกในคลาส 0 ได้ดีโดยดูที่ค่า ความแม่นยำ และค่าความระลึก ที่ค่อนข้างสูงแต่การจำแนกในคลาส 1 นั้นยังไม่ดีพอหรือยังมีปัญหาเนื่องจาก ค่า ความแม่นยำ และค่าความระลึกนั้นมีค่าที่ต่ำ ซึ่งอาจเกิดชุดข้อมูลที่ไม่สมดุลได้ โดยสามารถดูได้จากตารางสับสนดังภาพที่ 4-4



ภาพที่ 4-4 ตารางสับสนของไลต์ เกรเดียนต์ บูสติง แมชชีน

จากภาพที่ 4-4 แสดงให้เห็นถึงประสิทธิภาพของแบบจำลองที่ใช้ในการจำแนกประเภทข้อมูลออกเป็นสองกลุ่ม ได้แก่ คลาส 0 และคลาส 1 ซึ่งจากผลลัพธ์แบบจำลองสามารถวิเคราะห์ค่าของคลาส 0 ได้อย่างแม่นยำ โดยมีจำนวน True Negative (TN) อยู่ที่ 17,408 รายการ ขณะที่จำนวน False Positive (FP) จำนวน 820 รายการ ซึ่งยังถือว่ามีความผิดพลาดอยู่บ้าง และ แบบจำลองจำแนกคลาส 1 ผิดพลาดสูงโดยมี False Negative (FN) 2,514 รายการ ซึ่งถึงแบบจำลองมีแนวโน้มที่จะมองข้ามตัวอย่างที่เป็นคลาส 1 ไปเป็นคลาส 0 มาก และ True Positive (TP) 1,728 รายการ ซึ่งเป็นจำนวนที่ต่ำเมื่อเทียบกับคลาส 0 ซึ่งแสดงว่าแบบจำลองนี้เกิดชุดข้อมูลที่ไม่สมดุลต้องได้รับการปรับปรุงแบบจำลอง

จากการวัดประสิทธิภาพแบบจำลองพบว่าแบบจำลองต้นไม้ตัดสินใจให้ผลลัพธ์ที่แย่ที่สุด และแบบจำลองที่ให้ผลลัพธ์ที่ดีที่สุดคือ เอ็กซ์ตรีม เกรเดียนต์ บูสตีง และ โลท เกรเดียนต์ บูสตีง แมชชีน อย่างไรก็ตามแบบจำลองทั้ง 4 แบบนั้นวิเคราะห์จำแนกในคลาส 0 ได้เป็นอย่างดีในทุก ๆ แบบจำลอง และวิเคราะห์คลาส 1 ได้แย่กว่าในทุกแบบจำลอง ซึ่งเกิดจากคลาสของข้อมูลที่ไม่สมดุลกัน โดยชุดข้อมูลการฝึกมีคลาส 0 จำนวน 72,907 รายการ และคลาส 1 จำนวน 16,969 รายการ ซึ่งทั้ง 2 คลาสมีข้อมูลต่างกัน 55,938 ทั้งนี้แบบจำลองต้องได้รับการปรับปรุงประสิทธิภาพแบบจำลอง เพื่อให้แบบจำลองสามารถจำแนกคลาสทั้ง 2 ได้เป็นอย่างดี

4.1.2 ผลการพัฒนาแบบจำลองปรับปรุงประสิทธิภาพโดย SMOTE เป็นการพัฒนาแบบจำลองผ่านกระบวนการเรียนรู้ของเครื่องทั้ง 4 แบบจำลองเช่นเดียวกับข้อ 4.1.1 แต่ในขั้นตอนนี้ได้มีการปรับปรุงประสิทธิภาพของแบบจำลองด้วยการใช้เทคนิค SMOTE เพื่อแก้ไขปัญหาความไม่สมดุลของข้อมูลโดยการเพิ่มจำนวนข้อมูลในกลุ่มที่มีจำนวนน้อยให้ใกล้เคียงกับกลุ่มที่มีจำนวน โดยผลลัพธ์ของการพัฒนาแบบจำลองหลังการปรับปรุงด้วย SMOTE แสดงดังตารางที่ 4-2

ตารางที่ 4-2 ผลการพัฒนาแบบจำลองปรับปรุงประสิทธิภาพโดย SMOTE

แบบจำลอง	ระดับความรุนแรง	ค่าความถูกต้อง	ค่าความแม่นยำ	ค่าความระลึก	ค่าคะแนนเอฟหนึ่ง
ต้นไม้ตัดสินใจ	0	0.76	0.87	0.82	0.84
	1		0.38	0.50	0.43
ป่าสุ่ม	0	0.82	0.89	0.88	0.89
	1		0.51	0.54	0.52
เอ็กซ์ตรีม เกรเดียนต์ บูสตีง	0	0.81	0.90	0.86	0.88
	1		0.50	0.58	0.54
โลท เกรเดียนต์ บูสตีง แมชชีน	0	0.81	0.90	0.87	0.88
	1		0.51	0.60	0.57

จากตารางที่ 4-2 พบว่าต้นไม้ตัดสินใจสามารถจำแนกระดับความรุนแรงระดับ 0 ดีกว่าคลาส 1 โดยมีค่าความถูกต้องเท่ากับ 0.79 และค่าคะแนนเอฟหนึ่งเท่ากับ 0.87 โดยที่ความรุนแรงระดับ 1 ค่าประสิทธิภาพลดลงอย่างชัดเจน ซึ่งค่าความระลึก และ ค่าคะแนนเอฟหนึ่งอยู่ที่ 0.44 เมื่อปรับปรุงประสิทธิภาพด้วย SMOTE พบว่า ประสิทธิภาพโดยรวมไม่ได้ดีขึ้น โดยเฉพาะระดับความรุนแรง 1 ซึ่งค่าความระลึกเพิ่มขึ้น 0.50 แต่ค่าความแม่นยำและค่าคะแนนเอฟหนึ่งยังต่ำ นอกจากนี้แบบจำลองป่า

สุ่มเมื่อใช้ SMOTE พบว่า ค่าคะแนนเอฟหนึ่ง ในระดับความรุนแรง 1 เพิ่มขึ้นเล็กน้อยเป็น 0.52 แต่ประสิทธิภาพยังอยู่ในระดับต่ำเมื่อเทียบกับความรุนแรงระดับ 0 ซึ่งมีค่าคะแนนเอฟหนึ่ง อยู่ที่ 0.89 แสดงถึงความไม่สมดุลแม้ใช้ SMOTE เข้ามาช่วยเพิ่มประสิทธิภาพ อีกทั้งเอกซ์ตรึม เกรเดียนต์ บูสตั๋ง หลังจากใช้ SMOTE พบว่าประสิทธิภาพเพิ่มขึ้นเล็กน้อยในความรุนแรงระดับ 1 โดยค่าความระลึก เพิ่มจาก 0.44 เป็น 0.58 และ ค่าคะแนนเอฟหนึ่ง เพิ่มเป็น 0.54 และ โลท เกรเดียนต์ บูสตั๋ง แมชชีน ให้ผลการจำแนกที่คล้ายกับ เอกซ์ตรึม เกรเดียนต์ บูสตั๋ง เมื่อใช้ SMOTE แล้ว พบว่า ค่าคะแนนเอฟหนึ่ง สำหรับระดับ 1 เพิ่มขึ้นเป็น 0.57 ซึ่งสูงกว่าแบบจำลองอื่น ๆ

จากผลพบว่าแบบจำลองที่ให้ประสิทธิภาพดีที่สุดในการจำแนกความรุนแรงของอุบัติเหตุเมื่อใช้ SMOTE คือ โลท เกรเดียนต์ บูสตั๋ง แมชชีน ซึ่งสามารถเพิ่มค่าความระลึก และ ค่าคะแนนเอฟหนึ่งในระดับความรุนแรงที่ 1 ได้สูงที่สุดเมื่อเทียบกับแบบจำลองอื่น รองลงมาคือ เอกซ์ตรึม เกรเดียนต์ บูสตั๋ง ที่ตอบสนองต่อ SMOTE ได้ดีเช่นกัน ในขณะที่ป่าสุ่มแม้จะมีความแม่นยำสูงในระดับ 0 แต่ยังคงประสบปัญหาในการจำแนกระดับ 1 ส่วนแบบจำลอง Decision Tree มีประสิทธิภาพต่ำสุดโดยรวม และไม่ได้รับประโยชน์จากการใช้ SMOTE

4.1.3 ผลการพัฒนาแบบจำลองปรับปรุงประสิทธิภาพโดย SMOTETomek เป็นการผสมผสานระหว่าง SMOTE ซึ่งช่วยเพิ่มจำนวนตัวอย่างของกลุ่มข้อมูลที่มีจำนวนน้อยให้สมดุลมากขึ้น กับ Tomek Link ที่ช่วยลบตัวอย่างข้อมูลที่อาจเป็น Noise หรือค่าผิดปกติ (Outliers) ออกจากชุดข้อมูล โดยผลลัพธ์ของการพัฒนาแบบจำลองหลังการปรับปรุงด้วย SMOTETomek แสดงดังตารางที่ 4-3

ตารางที่ 4-3 ผลการพัฒนาแบบจำลองปรับปรุงประสิทธิภาพโดย SMOTETomek

แบบจำลอง	ระดับความรุนแรง	ค่าความถูกต้อง	ค่าความแม่นยำ	ค่าความระลึก	ค่าคะแนนเอฟหนึ่ง
ต้นไม้ตัดสินใจ	0	0.76	0.88	0.83	0.85
	1		0.40	0.50	0.44
ป่าสุ่ม	0	0.83	0.91	0.85	0.88
	1		0.50	0.62	0.55
เอกซ์ตรึม เกรเดียนต์ บูสตั๋ง	0	0.81	0.85	0.83	0.83
	1		0.34	0.76	0.47
โลท เกรเดียนต์ บูสตั๋ง แมชชีน	0	0.81	0.85	0.84	0.83
	1		0.45	0.80	0.47

จากตารางที่ 4-3 ต้นไม้ตัดสินใจเมื่อใช้ SMOTETomek พบว่าประสิทธิภาพในการจำแนก ระดับความรุนแรงที่ 0 อยู่ในระดับใกล้เคียงกับ SMOTE และแบบไม่เพิ่มประสิทธิภาพ โดยมีค่าคะแนนเอฟหนึ่ง เท่ากับ 0.85 ส่วนระดับ 1 มีค่าคะแนนเอฟหนึ่ง เท่ากับ 0.44 ซึ่ง เท่ากับการใช้ SMOTE และ สูงกว่าแบบไม่เพิ่มประสิทธิภาพเล็กน้อย นอกจากนี้แบบจำลองป่าสุ่มด้วยการเพิ่ม ประสิทธิภาพด้วย SMOTETomek ทำให้ค่าคะแนนเอฟหนึ่ง สำหรับระดับความรุนแรง 0 ลดลงอยู่ที่ 0.88 ส่วนระดับ 1 ได้ค่าคะแนนเอฟหนึ่ง เท่ากับ 0.55 ซึ่ง เพิ่มขึ้นจากทั้ง SMOTE และ แบบไม่ ปรับปรุง ซึ่งเอ็กซ์ตรีม เกรเดียนต์ บูสติงผลลัพธ์แสดงว่า SMOTETomek ไม่ได้เพิ่มประสิทธิภาพ ให้กับ เอ็กซ์ตรีม เกรเดียนต์ บูสติง มากนัก โดยระดับ 0 ได้ค่าคะแนนเอฟหนึ่ง เท่ากับ 0.83 ลดลง จาก SMOTE และแบบที่ไม่เพิ่มประสิทธิภาพ ขณะที่ระดับ 1 ได้ค่าคะแนนเอฟหนึ่ง เพียง 0.47 ซึ่ง ลดลงจาก SMOTE และ แบบไม่ปรับปรุง และโลท เกรเดียนต์ บูสติง แมชชี ระดับ 0 มีค่าคะแนนเอฟ หนึ่ง เท่ากับ 0.83 ลดลงกว่าการใช้ SMOTE และไม่เพิ่มประสิทธิภาพ ส่วนระดับ 1 มีค่าคะแนนเอฟ หนึ่ง เท่ากับ 0.47 ซึ่ง น้อยกว่าการใช้ SMOTE และ เท่ากับแบบไม่ปรับปรุง

จากผลลัพธ์พบว่า SMOTETomek ช่วยเพิ่มประสิทธิภาพได้เพียงบางส่วน โดยเฉพาะใน แบบจำลองป่าสุ่มที่ค่าคะแนนเอฟหนึ่ง ในระดับความรุนแรงที่ 1 เพิ่มขึ้นมากที่สุด แสดงถึงประโยชน์ ของเทคนิคนี้ในการจัดการกับข้อมูลไม่สมดุลในบางโมเดล อย่างไรก็ตาม แบบจำลองอื่นอย่าง โลท เกรเดียนต์ บูสติง แมชชีน และ เอ็กซ์ตรีม เกรเดียนต์ บูสติง กลับให้ผลลัพธ์ที่ ลดลงเมื่อใช้ SMOTETomek เมื่อเทียบกับการใช้ SMOTE เพียงอย่างเดียว

4.1.4 ผลการพัฒนาแบบจำลองปรับปรุงประสิทธิภาพโดยการปรับน้ำหนัก โดยใช้ Random Search เพื่อค้นหาค่า Hyperparameter ในแบบจำลองที่ช่วยทำให้แบบจำลองเรียนรู้ได้อย่างสมดุล โดยผลลัพธ์ของการพัฒนาแบบจำลองหลังการปรับปรุงประสิทธิภาพโดยการปรับน้ำหนัก แสดงดัง ตารางที่ 4-4

ตารางที่ 4-4 ผลการพัฒนาแบบจำลองปรับปรุงประสิทธิภาพโดยการปรับน้ำหนัก

แบบจำลอง	ระดับความรุนแรง	ค่าความถูกต้อง	ค่าความแม่นยำ	ค่าความระลึก	ค่าคะแนนเอฟหนึ่ง
ต้นไม้ตัดสินใจ	0	0.79	0.91	0.83	0.87
	1		0.47	0.65	0.54
ป่าสุ่ม	0	0.82	0.91	0.87	0.89
	1		0.53	0.61	0.57

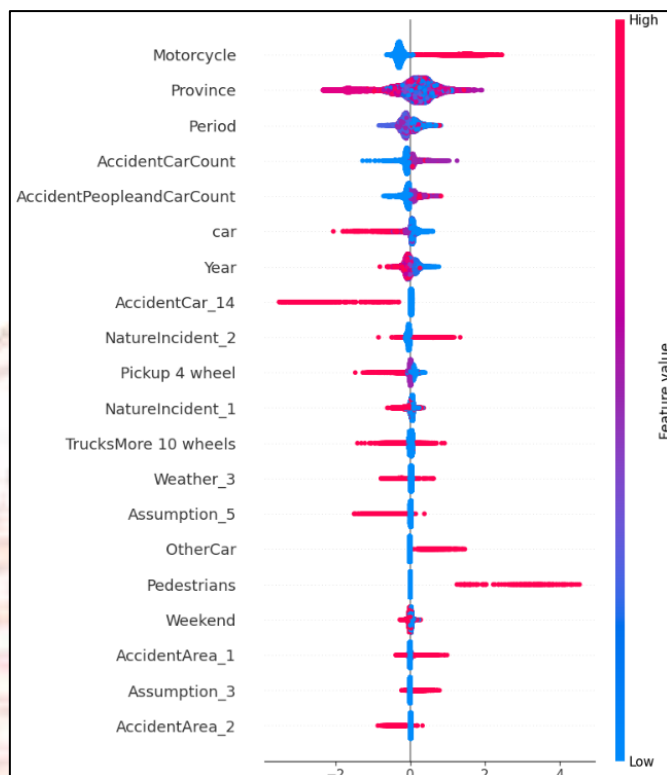
ตารางที่ 4-4 (ต่อ)

แบบจำลอง	ระดับความรุนแรง	ค่าความถูกต้อง	ค่าความแม่นยำ	ค่าความระลึก	ค่าคะแนนเอฟหนึ่ง
เอ็กซ์ตรีม เกรเดียนต์ บูสตีง	0	85	0.92	0.84	0.88
	1		0.68	0.67	0.68
โลท กรเดียนต์ บูสตีง แมชชีน	0	84	0.89	0.92	0.90
	1		0.59	0.53	0.54

จากตารางที่ 4-4 พบว่าต้นไม้ตัดสินใจจากการเพิ่มประสิทธิภาพด้วยการปรับน้ำหนักค่าคะแนนเอฟหนึ่ง ของความรุนแรงระดับ 1 เพิ่มขึ้นเป็น 0.54 ซึ่งสูงกว่าการใช้ SMOTE และ SMOTETomek แสดงให้เห็นว่าการปรับน้ำหนักช่วยเพิ่มประสิทธิภาพในการจำแนกระดับความรุนแรงที่น้อยได้ดีกว่าเทคนิคอื่น ๆ สำหรับแบบจำลองนี้ โดยยังคงค่าคะแนนเอฟหนึ่ง ในระดับ 0 ไว้ที่ 0.87 ซึ่งใกล้เคียงเดิม โดยที่ป่าสุ่มผลลัพธ์หลังจากการเพิ่มประสิทธิภาพด้วยการปรับน้ำหนัก พบว่าค่าคะแนนเอฟหนึ่ง ของความรุนแรงระดับ 1 เพิ่มขึ้นเป็น 0.57 ซึ่งมากกว่าการเพิ่มประสิทธิภาพในแบบอื่น ๆ นอกจากนี้เอ็กซ์ตรีม เกรเดียนต์ บูสตีง แบบจำลองนี้ค่าคะแนนเอฟหนึ่ง ของความรุนแรงระดับ 1 เท่ากับ 0.68 ซึ่งสูงกว่า SMOTETomek SMOTE และแบบไม่ปรับปรุง แสดงว่า เอ็กซ์ตรีม เกรเดียนต์ บูสตีง และ ค่าคะแนนเอฟหนึ่ง ของระดับ 0 ก็ยังอยู่ในระดับสูงคือ 0.88 และโลท เกรเดียนต์ บูสตีง แมชชี มีค่าคะแนนเอฟหนึ่ง ของระดับ 0 สูงที่สุดที่ 0.90 แต่สำหรับระดับ 1 ค่าคะแนนเอฟหนึ่ง อยู่ที่ 0.54 ซึ่ง ต่ำกว่าการใช้ SMOTE

จากผลการเปรียบเทียบพบว่า การปรับน้ำหนักเป็นเทคนิคที่ช่วยเพิ่มประสิทธิภาพในการจำแนกระดับความรุนแรงที่ 1 ได้ดีในหลายแบบจำลอง โดยเฉพาะต้นไม้ตัดสินใจและ เอ็กซ์ตรีม เกรเดียนต์ บูสตีง ที่ค่าคะแนนเอฟหนึ่ง สูงกว่าการใช้เทคนิค SMOTE และ SMOTETomek

เมื่อได้แบบจำลองที่เหมาะสม ผู้จัดทำการค้นคว้าอิสระได้ใช้ SHapley Additive exPlanations (SHAP) เพื่อหาปัจจัยที่ส่งผลต่อผลการวิเคราะห์ของแบบจำลอง โดย SHAP จะคำนวณค่าความสำคัญของคุณลักษณะแต่ละรายการ ผลการวิเคราะห์ด้วย SHAP ช่วยให้สามารถระบุคุณลักษณะที่มีอิทธิพลมากที่สุดต่อการวิเคราะห์ของแบบจำลองได้ ดังภาพที่ 4-5



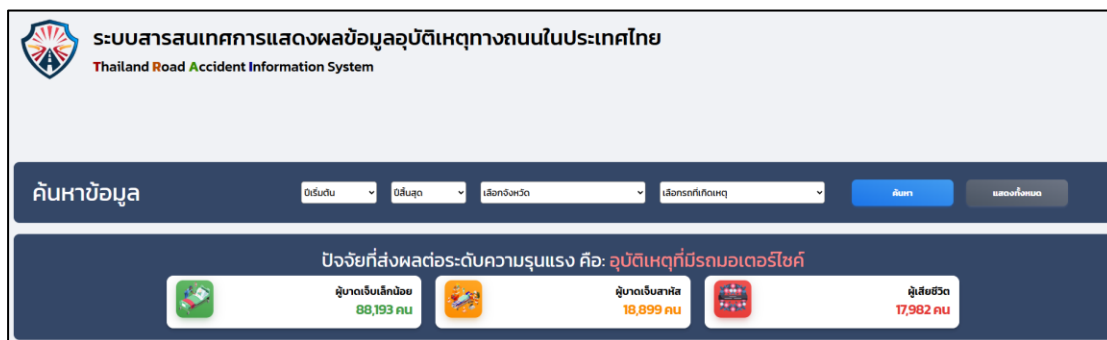
ภาพที่ 4-5 ปัจจัยที่ส่งผลต่อความรุนแรงด้วย SHAP

จากภาพที่ 4-5 พบว่าผลการวิเคราะห์ของ SHAP เป็นการอธิบายความสำคัญของคุณลักษณะที่มีผลต่อแบบจำลอง โดยคุณลักษณะที่ส่งผลต่อระดับความรุนแรงของอุบัติเหตุมากที่สุด ได้แก่ Motorcycle โดยมีค่า SHAP ที่กระจายทั้งในเชิงบวกและลบ ซึ่งสะท้อนให้เห็นว่าการเกิดอุบัติเหตุที่เกี่ยวข้องกับรถจักรยานยนต์เป็นปัจจัยสำคัญต่อระดับความรุนแรง รองลงมาคือ Province, Period และ AccidentCarCount ซึ่งแสดงถึงความสำคัญอย่างมีนัยต่อการวิเคราะห์ โดยเฉพาะในกลุ่มข้อมูลที่มีค่าของคุณลักษณะมีความหลากหลายสูงในขณะที่ยังตัวแปร เช่น Pedestrians, OtherCar และ Weekend มีค่า SHAP ที่กระจายตัวน้อยกว่า แสดงให้เห็นว่ามีผลต่อแบบจำลองน้อยกว่าเมื่อเทียบกับคุณลักษณะอื่น

4.2 ผลการพัฒนาระบบสารสนเทศ

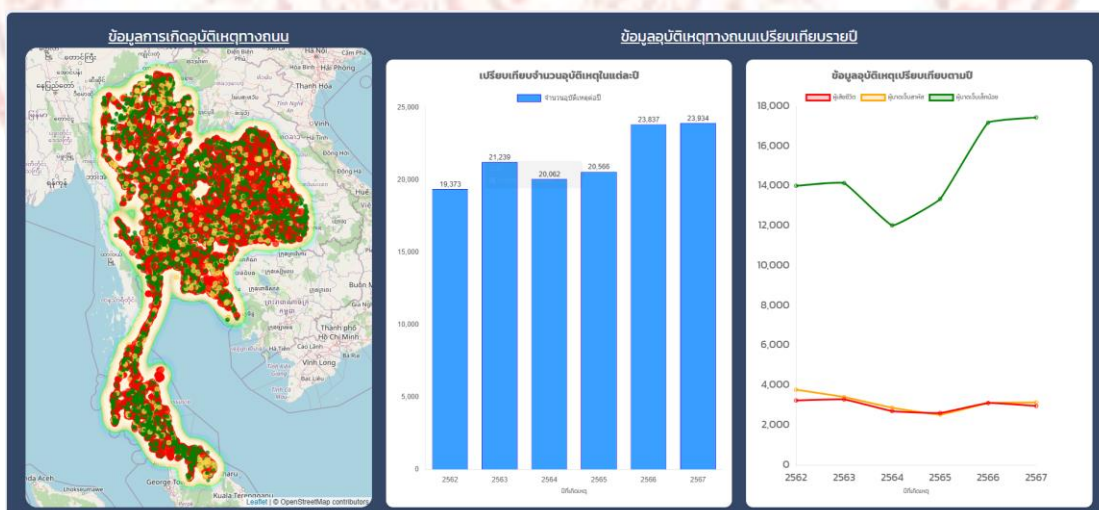
การพัฒนาระบบสารสนเทศนั้นได้พัฒนาระบบสารสนเทศการแสดงผลข้อมูลอุบัติเหตุทางถนนในประเทศไทย (Road Accident Dashboard) ดังนี้

4.2.1 ส่วนการค้นหาข้อมูล โดยระบบสารสนเทศจะแสดงส่วนในการค้นหาข้อมูลต่าง ๆ โดยแบ่งเป็น การตรวจสอบเส้นทางล่วงหน้า และ การค้นหาข้อมูล โดยผู้ใช้งานสามารถเลือกเงื่อนไขที่ต้องการค้นหา เพื่อแสดงข้อมูลตามที่ต้องการดังภาพที่ 4-6

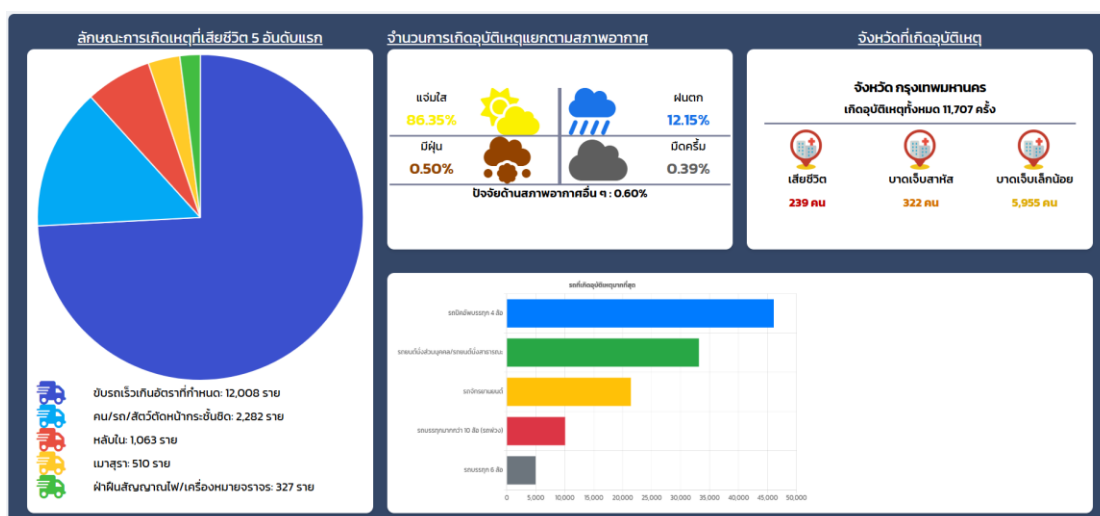


ภาพที่ 4-6 หน้าจอแสดงผลระบบสารสนเทศส่วนการค้นหาข้อมูล

4.2.2 ส่วนการแสดงผลข้อมูล โดยส่วนการแสดงผลข้อมูลจะแสดงข้อมูลตามเงื่อนไขที่ผู้ใช้งานเลือก โดยการแสดงผลจะแสดงในรูปแบบของแผนที่การเกิดอุบัติเหตุแบ่งเป็น 3 ระดับคือ การเกิดเหตุที่มีผู้บาดเจ็บเล็กน้อยในแผนที่แสดงสีเขียว การเกิดเหตุที่มีผู้บาดเจ็บสาหัสในแผนที่แสดงสีส้ม และการเกิดเหตุที่มีผู้เสียชีวิตในแผนที่แสดงสีแดง นอกจากนี้ยังมีการแสดงผลการเปรียบเทียบการเกิดอุบัติเหตุ ดังภาพที่ 4-7 และ ภาพที่ 4-8



ภาพที่ 4-7 แผนที่การแสดงจุดเกิดเหตุและแผนภูมิเปรียบเทียบข้อมูลอุบัติเหตุรายปี



ภาพที่ 4-8 ข้อมูลการแสดงผลอุบัติเหตุแยกตามประเภทข้อมูล

4.3 ผลการประเมินคุณภาพและผลการประเมินความพึงพอใจต่อระบบสารสนเทศ

4.3.1 ผลการประเมินคุณภาพระบบสารสนเทศโดยผู้เชี่ยวชาญ โดยได้รับความอนุเคราะห์จากผู้เชี่ยวชาญ 3 ท่านในการประเมินคุณภาพของระบบสารสนเทศ โดยการประเมินทั้ง 3 ด้าน โดยทั้ง 3 ด้านจะมีการถามคำถามที่แตกต่างกันออกไป ซึ่งมีผลการประเมินคุณภาพดังนี้

ตารางที่ 4-5 ผลการประเมินคุณภาพด้านความสามารถของระบบสารสนเทศตามความต้องการของผู้ใช้งาน

ด้านการประเมินคุณภาพ	ผลการวิเคราะห์		
	$\bar{x}$	S.D.	อภิปรายผล
1. ระบบสามารถระบุเส้นทางเริ่มต้นและเส้นทางปลายทางได้	4.30	0.20	มาก
2. ระบบสามารถแสดงข้อมูลอุบัติเหตุระหว่างเส้นทางที่ระบุได้	4.40	0.25	มาก
3. ระบบสามารถวิเคราะห์อุบัติเหตุที่เคยเกิดขึ้นระหว่างเส้นทางได้	4.60	0.40	มากที่สุด
4. ระบบสามารถแสดงข้อมูลอุบัติเหตุบนพิกัดของแผนที่ได้	4.30	0.30	มาก
5. ระบบสามารถสืบค้นข้อมูลต่าง ๆ	4.50	0.25	มาก

ตารางที่ 4-5 (ต่อ)

ด้านการประเมินคุณภาพ	ผลการวิเคราะห์		
	$\bar{x}$	S.D.	อภิปรายผล
6. ระบบสามารถแจ้งเตือนเมื่อไม่พบข้อมูลที่ต้องการ หรือพบข้อผิดพลาดระหว่างใช้งาน	4.30	0.25	มาก
7. ระบบสามารถตอบสนองของระบบต่อคำสั่งของผู้ใช้งาน	4.70	0.20	มากที่สุด
ผลการประเมินคุณภาพโดยภาพรวม	4.44	0.26	มาก

จากตารางที่ 4-5 ผลการประเมินคุณภาพด้านความสามารถของระบบสารสนเทศตามความต้องการของผู้ใช้งานโดยผู้เชี่ยวชาญนั้น พบว่าความสามารถของระบบในภาพรวมทั้งในส่วนของ การวิเคราะห์อุบัติเหตุระหว่างเส้นทาง การสืบค้นข้อมูลและการแจ้งเตือนของระบบ และการตอบสนองของระบบต่อคำสั่งผู้ใช้งานได้รับค่าเฉลี่ยสูงที่สุด ซึ่งผลการประเมินโดยรวมมีค่าเฉลี่ยเท่ากับ 4.44 คะแนน และค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.26 คะแนน จึงสามารถสรุปได้ว่าระบบสารสนเทศมีคุณภาพด้านความสามารถของระบบสารสนเทศตามความต้องการของผู้ใช้งานอยู่ในระดับมาก

ตารางที่ 4-6 ผลการประเมินคุณภาพของผู้เชี่ยวชาญด้านการทำงานตามฟังก์ชัน

ด้านการประเมินคุณภาพ	ผลการวิเคราะห์		
	$\bar{x}$	S.D.	อภิปรายผล
1. ระบบสามารถระบุเส้นทางเริ่มต้นและเส้นทางปลายทางได้อย่างถูกต้อง	4.30	0.20	มาก
2. ระบบสามารถแสดงข้อมูลอุบัติเหตุระหว่างเส้นทางที่ระบุได้อย่างถูกต้อง	4.40	0.25	มาก
3. ระบบสามารถวิเคราะห์อุบัติเหตุที่เคยเกิดขึ้นระหว่างเส้นทางได้อย่างถูกต้องเหมาะสม	4.60	0.30	มากที่สุด
4. ระบบสามารถแสดงข้อมูลอุบัติเหตุบนพิกัดของแผนที่ได้อย่างถูกต้อง	4.50	0.20	มาก
5. ระบบสามารถสืบค้นข้อมูลต่าง ๆ ได้อย่างถูกต้อง	4.40	0.35	มาก

ตารางที่ 4-6 (ต่อ)

ด้านการประเมินคุณภาพ	ผลการวิเคราะห์		
	$\bar{x}$	S.D.	อภิปรายผล
6. ระบบสามารถแจ้งเตือนเมื่อไม่พบข้อมูลที่ต้องการ หรือพบข้อผิดพลาดระหว่างใช้งานได้อย่างถูกต้อง	4.30	0.15	มาก
7. ระบบสามารถตอบสนองของระบบต่อคำสั่งของผู้ใช้งานได้อย่างถูกต้อง	4.70	0.25	มากที่สุด
ผลการประเมินคุณภาพโดยภาพรวม	4.41	0.24	มาก

จากตารางที่ 4-6 ผลการประเมินคุณภาพด้านการทำงานตามฟังก์ชันโดยผู้เชี่ยวชาญนั้น พบว่าความสามารถของระบบในภาพรวมทั้งในส่วนของกระบวนการระบุเส้นทาง การวิเคราะห์อุบัติเหตุที่เคยเกิดขึ้นระหว่างเส้นทาง การตอบสนองของระบบต่อคำสั่งของผู้ใช้งาน และการแสดงข้อมูลอุบัติเหตุบนพิกัดของแผนที่ สามารถทำงานได้อย่างถูกต้องได้รับค่าเฉลี่ยสูงสุด ซึ่งผลการประเมินโดยรวมมีค่าเฉลี่ยเท่ากับ 4.41 คะแนน และค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.24 คะแนน จึงสามารถสรุปได้ว่าระบบสารสนเทศมีคุณภาพด้านการทำงานตามฟังก์ชันอยู่ในระดับมาก

ตารางที่ 4-7 ผลการประเมินคุณภาพของผู้เชี่ยวชาญด้านความสะดวกและง่ายต่อการใช้งานระบบสารสนเทศ

ด้านการประเมินคุณภาพ	ผลการวิเคราะห์		
	$\bar{x}$	S.D.	อภิปรายผล
1. ความเหมาะสมของการใช้กราฟในการแสดงผลข้อมูล	4.40	0.15	มาก
2. ความเหมาะสมของการจัดวางตำแหน่งต่าง ๆ (Layout) ภายในหน้าเว็บไซต์	4.40	0.10	มาก
3. ความเหมาะสมของการเลือกใช้สี พื้นหลัง สีกราฟ สีตัวอักษร	4.40	0.20	มาก
4. ระบบสามารถใช้งานได้ง่าย ไม่ซับซ้อน	4.30	0.20	มาก
5. ระบบสามารถทำงานได้อย่างต่อเนื่องโดยไม่มีปัญหาการล่มหรือค้าง	4.30	0.30	มาก

ตารางที่ 4-7 (ต่อ)

ด้านการประเมินคุณภาพ	ผลการวิเคราะห์		
	$\bar{x}$	S.D.	อภิปรายผล
6. ระบบสามารถใช้งานได้ดีทั้งบนคอมพิวเตอร์และอุปกรณ์มือถือ หรืออุปกรณ์อื่น ๆ (Responsive)	4.10	0.10	มาก
ผลการประเมินคุณภาพโดยภาพรวม	4.31	0.18	มาก

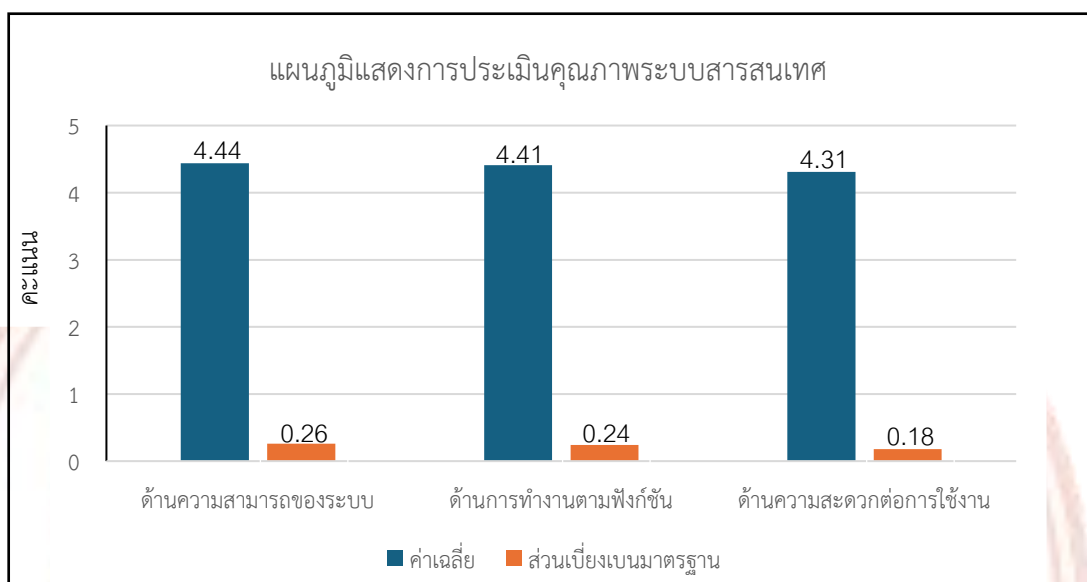
จากตารางที่ 4-7 ผลการประเมินคุณภาพด้านด้านความสะดวกและง่ายต่อการใช้งานระบบสารสนเทศ พบว่าความเหมาะสมของการใช้กราฟในการแสดงผลข้อมูล ความเหมาะสมของการจัดวางตำแหน่งต่าง ๆ ภายในหน้าเว็บไซต์ และความเหมาะสมของการเลือกใช้สี พื้นหลัง สีกราฟ สีตัวอักษร ได้ค่าเฉลี่ยสูงสุด ซึ่งผลการประเมินโดยรวมมีค่าเฉลี่ยเท่ากับ 4.31 คะแนน และค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.18 คะแนน จึงสามารถสรุปได้ว่าระบบสารสนเทศมีคุณภาพด้านด้านความสะดวกและง่ายต่อการใช้งานระบบสารสนเทศอยู่ในระดับมาก

ตารางที่ 4-8 สรุปผลการประเมินคุณภาพระบบสารสนเทศโดยผู้เชี่ยวชาญทั้ง 3 ด้าน

ด้านการประเมินคุณภาพ	ผลการวิเคราะห์		
	$\bar{x}$	S.D.	อภิปรายผล
1. ด้านความสามารถของระบบสารสนเทศตามความต้องการของผู้ใช้งาน	4.44	0.26	มาก
2. ด้านการทำงานตามฟังก์ชัน	4.41	0.24	มาก
3. ด้านความสะดวกและง่ายต่อการใช้งานระบบสารสนเทศ	4.31	0.18	มาก
ผลการประเมินคุณภาพโดยภาพรวม	4.39	0.23	มาก

จากตารางที่ 4-8 พบว่าด้านความสามารถของระบบสารสนเทศตามความต้องการของผู้ใช้งานมีค่าเฉลี่ยเท่ากับ 4.44 คะแนน และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.26 คะแนน ผลการประเมินอยู่ในระดับมาก และด้านการทำงานตามฟังก์ชันได้ค่าเฉลี่ยเท่ากับ 4.41 คะแนน ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.24 คะแนน ผลการประเมินอยู่ในระดับมาก นอกจากนี้ด้านความสะดวกและง่ายต่อการใช้งานได้รับค่าเฉลี่ยเท่ากับ 4.31 คะแนน ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.18 คะแนน ผลการประเมิน

อยู่ในระดับมาก เมื่อพิจารณาโดยรวมทั้งสามด้านแล้วพบว่าระบบสารสนเทศมีค่าเฉลี่ยเท่ากับ 4.39 คะแนนและส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.23 คะแนน ซึ่งสามารถสรุปได้ว่าระบบสารสนเทศที่พัฒนาขึ้นมีคุณภาพมาก ดังภาพที่ 4-9



ภาพที่ 4-9 ผลการประเมินคุณภาพระบบสารสนเทศโดยผู้เชี่ยวชาญในภาพรวมทั้ง 3 ด้าน

4.3.2 ผลการประเมินความพึงพอใจระบบสารสนเทศโดยผู้ใช้งานทั่วไปโดยได้รับความอนุเคราะห์จากผู้ใช้งานจำนวน 30 ท่าน ในการประเมินความพึงพอใจของระบบสารสนเทศ โดยมีการประเมิน 3 ด้าน โดยทั้ง 3 ด้านจะมีการถามคำถามที่แตกต่างกันออกไป ซึ่งมีผลการประเมินความพึงพอใจระบบสารสนเทศดังนี้

ตารางที่ 4-9 ผลการประเมินความพึงพอใจด้านความสามารถของระบบสารสนเทศตามความต้องการของผู้ใช้งาน

ด้านการประเมินความพึงพอใจ	ผลการวิเคราะห์		
	$\bar{x}$	S.D.	อภิปรายผล
1. ระบบสามารถระบุเส้นทางเริ่มต้นและเส้นทางปลายทางได้	4.10	0.30	มาก
2. ระบบสามารถแสดงข้อมูลอุบัติเหตุระหว่างเส้นทางที่ระบุได้	4.20	0.45	มาก

ตารางที่ 4-9 (ต่อ)

ด้านการประเมินความพึงพอใจ	ผลการวิเคราะห์		
	$\bar{x}$	S.D.	อภิปรายผล
3. ระบบสามารถแสดงข้อมูลอุบัติเหตุบนพิกัดของแผนที่ได้	3.80	0.40	ปานกลาง
4. ระบบสามารถสืบค้นข้อมูลต่าง ๆ	4.50	0.30	มาก
5. ระบบสามารถแจ้งเตือนเมื่อไม่พบข้อมูลที่ต้องการ หรือพบข้อผิดพลาดระหว่างใช้งาน	4.50	0.25	มาก
ผลการประเมินความพึงพอใจในภาพรวม	4.22	0.34	มาก

จากตารางที่ 4-9 ผลการประเมินความพึงพอใจด้านความสามารถของระบบสารสนเทศตามความต้องการของผู้ใช้งาน โดยผู้ใช้งานทั่วไปนั้น พบว่าสามารถแจ้งเตือนเมื่อไม่พบข้อมูลที่ต้องการ หรือพบข้อผิดพลาดระหว่างใช้งาน การสืบค้นข้อมูลและแสดงข้อมูลอุบัติเหตุระหว่างเส้นทางที่ระบุได้รับค่าเฉลี่ยสูงสุด ซึ่งผลการประเมินโดยรวมมีค่าเฉลี่ยเท่ากับ 4.22 คะแนน และค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.34 คะแนน จึงสามารถสรุปได้ผู้ใช้งานทั่วไปมีความพึงพอใจด้านความสามารถของระบบสารสนเทศตามความต้องการของผู้ใช้งานต่อระบบที่พัฒนาขึ้นมา

ตารางที่ 4-10 ผลการประเมินความพึงพอใจด้านการทดสอบการทำงานตามฟังก์ชัน

ด้านการประเมินความพึงพอใจ	ผลการวิเคราะห์		
	$\bar{x}$	S.D.	อภิปรายผล
1. ระบบสามารถระบุเส้นทางเริ่มต้นและเส้นทางปลายทางได้อย่างถูกต้อง	4.10	0.35	มาก
2. ระบบสามารถแสดงข้อมูลอุบัติเหตุระหว่างเส้นทางที่ระบุได้อย่างถูกต้อง	4.10	0.20	มาก
3. ระบบสามารถแสดงข้อมูลอุบัติเหตุบนพิกัดของแผนที่ได้อย่างถูกต้อง	4.30	0.15	มาก
4. ระบบสามารถสืบค้นข้อมูลต่าง ๆ ได้อย่างถูกต้อง	4.40	0.20	มาก

ตารางที่ 4-10 (ต่อ)

ด้านการประเมินความพึงพอใจ	ผลการวิเคราะห์		
	$\bar{x}$	S.D.	อภิปรายผล
5. ระบบสามารถแจ้งเตือนเมื่อไม่พบข้อมูลที่ต้องการ หรือพบข้อผิดพลาดระหว่างใช้งานได้อย่างถูกต้อง	4.20	0.10	มาก
ผลการประเมินความพึงพอใจในภาพรวม	4.41	0.22	มาก

จากตารางที่ 4-10 ผลการประเมินความพึงพอใจด้านด้านการทดสอบการทำงานตามฟังก์ชันงาน โดยผู้ใช้งานทั่วไปนั้น พบว่าความสามารถของระบบในภาพรวมทั้งในส่วนของการสืบค้นข้อมูล การแสดงข้อมูลอุบัติเหตุนกพิภพของแผนและแจ้งเตือนเมื่อไม่พบข้อมูลที่ต้องการ หรือพบข้อผิดพลาดระหว่างใช้งานได้อย่างถูกต้องได้รับค่าเฉลี่ยสูงสุด ซึ่งผลการประเมินโดยรวมมีค่าเฉลี่ยเท่ากับ 4.41 คะแนน และค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.22 คะแนน จึงสามารถสรุปได้ผู้ใช้งานทั่วไปมีความพึงพอใจด้านการทดสอบการทำงานตามฟังก์ชันงานต่อระบบที่พัฒนาขึ้นมาก

ตารางที่ 4-11 ผลการประเมินความพึงพอใจด้านความสะดวกและง่ายต่อการใช้งานระบบ

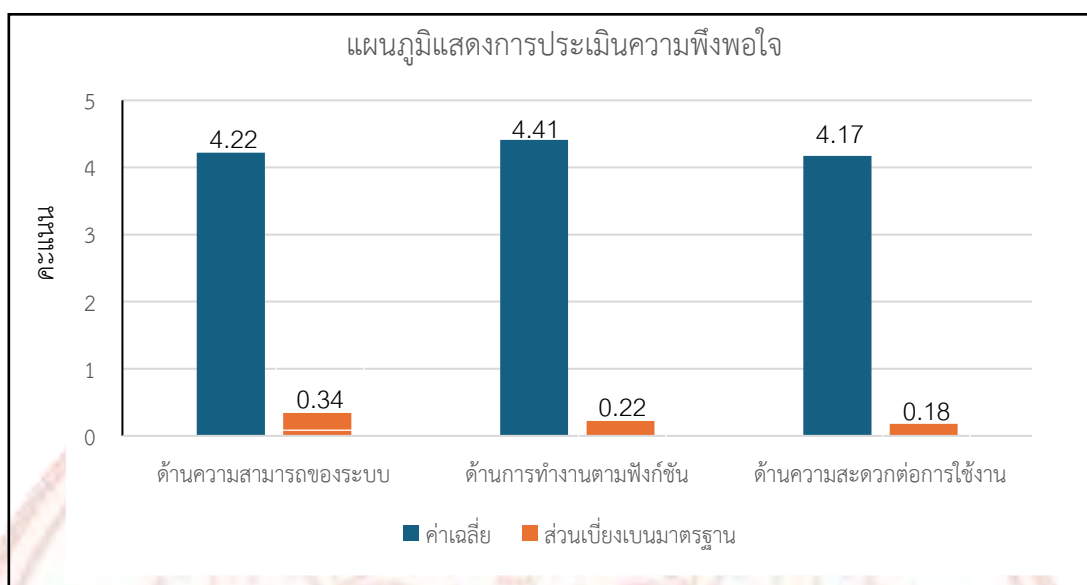
ด้านการประเมินความพึงพอใจ	ผลการวิเคราะห์		
	$\bar{x}$	S.D.	อภิปรายผล
1. ความเหมาะสมของการใช้กราฟในการแสดงผลข้อมูล	4.10	0.25	มาก
2. ความเหมาะสมของการจัดวางตำแหน่งต่าง ๆ (Layout) ภายในหน้าเว็บไซต์	4.30	0.15	มาก
3. ความเหมาะสมของการเลือกใช้สี พื้นหลัง สีกราฟ สีตัวอักษร	4.30	0.20	มาก
4. ระบบสามารถใช้งานได้ง่าย ไม่ซับซ้อน	4.10	0.20	มาก
5. ระบบสามารถใช้งานได้ทั้งบนคอมพิวเตอร์ และอุปกรณ์มือถือ หรืออุปกรณ์อื่น ๆ	4.10	0.10	มาก
ผลการประเมินความพึงพอใจในภาพรวม	4.17	0.18	มาก

จากตารางที่ 4-11 ผลการประเมินความพึงพอใจด้านความสะดวกและง่ายต่อการใช้งานระบบ โดยผู้ใช้งานทั่วไปนั้น พบว่าความเหมาะสมของการจัดวางตำแหน่งต่าง ๆ ภายในหน้าเว็บไซต์ และความเหมาะสมของการเลือกใช้สี พื้นหลัง สื่กราฟ สีตัวอักษร มีค่าเฉลี่ยสูงสุด ซึ่งผลการประเมินโดยรวมมีค่าเฉลี่ยเท่ากับ 4.17 คะแนน และค่าส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.18 คะแนน จึงสามารถสรุปได้ผู้ใช้งานทั่วไปมีความพึงพอใจด้านความสะดวกและง่ายต่อการใช้งานระบบต่อระบบที่พัฒนาขึ้นมา

ตารางที่ 4-12 สรุปผลการประเมินความพึงพอใจของระบบสารสนเทศโดยผู้ใช้งานทั่วไปทั้ง 3 ด้าน

ด้านการประเมินความพึงพอใจ	ผลการวิเคราะห์		
	$\bar{x}$	S.D.	อภิปรายผล
1. ด้านความสามารถของระบบสารสนเทศตามความต้องการของผู้ใช้งาน	4.22	0.34	มาก
2. ด้านการทดสอบการทำงานตามฟังก์ชัน	4.41	0.22	มาก
3. ด้านความสะดวกและง่ายต่อการใช้งานระบบสารสนเทศ	4.17	0.18	มาก
ผลการประเมินความพึงพอใจในภาพรวม	4.27	0.25	มาก

จากตารางที่ 4-12 พบว่าด้านความสามารถของระบบสารสนเทศตามความต้องการของผู้ใช้งานมีค่าเฉลี่ยเท่ากับ 4.22 คะแนน และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.34 คะแนน ผลการประเมินอยู่ในระดับมาก และด้านการทำงานตามฟังก์ชันได้ค่าเฉลี่ยเท่ากับ 4.41 คะแนน ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.22 คะแนน ผลการประเมินอยู่ในระดับมาก นอกจากนี้ด้านความสะดวกและง่ายต่อการใช้งานได้รับค่าเฉลี่ยเท่ากับ 4.17 คะแนน ส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.18 คะแนน ผลการประเมินอยู่ในระดับมาก เมื่อพิจารณาโดยรวมทั้งสามด้านแล้วพบว่าระบบสารสนเทศมีค่าเฉลี่ยเท่ากับ 4.27 คะแนนและส่วนเบี่ยงเบนมาตรฐานเท่ากับ 0.25 คะแนน ซึ่งสามารถสรุปได้ว่าผู้ใช้งานทั่วไปมีความพึงพอใจต่อการใช้งานระบบสารสนเทศที่พัฒนาขึ้นในระดับมาก ดังภาพที่ 4-10



ภาพที่ 4-10 ผลการประเมินความพึงพอใจระบบสารสนเทศโดยผู้ใช้งานทั่วไปในภาพรวมทั้ง 3 ด้าน

บทที่ 5

สรุปผล อภิปรายผล และข้อเสนอแนะการวิจัย

5.1 สรุปผลการวิจัย

จากการค้นคว้าอิสระเรื่องการวิเคราะห์ระดับความรุนแรงของอุบัติเหตุทางถนนโดยใช้การเรียนรู้ของเครื่อง โดยมีวัตถุประสงค์เพื่อพัฒนาแบบจำลองการวิเคราะห์ระดับความรุนแรงของอุบัติเหตุทางถนนและเพื่อพัฒนาระบบสารสนเทศสำหรับวิเคราะห์อุบัติเหตุทางถนน โดยสามารถสรุปผลการค้นคว้าอิสระได้ดังนี้

5.1.1 สรุปผลการพัฒนาแบบจำลอง

จากการพัฒนาแบบจำลองในการวิเคราะห์ระดับความรุนแรงของอุบัติเหตุด้วยการเรียนรู้ของเครื่อง โดยใช้แบบจำลองต้นไม้ตัดสินใจ (Decision Tree) ป่าสุ่ม (Random Forest) เอ็กซ์ตรีม เกรเดียนต์ บูสตีง (XGBoost) และโลท เกรเดียนต์ บูสตีง แมชชีน (LightGBM) โดยประเมินประสิทธิภาพแบบจำลองด้วยค่าความถูกต้อง (Accuracy) ค่าความแม่นยำ (Precision) ค่าความระลึก (Recall) และค่าคะแนนเอฟหนึ่ง (F1-score) ในการจำแนกระดับความรุนแรงออกเป็นระดับ 0 และระดับ 1 พบว่าแบบจำลองทั้ง 4 ที่ไม่ได้รับการปรับปรุงประสิทธิภาพให้ผลลัพธ์ที่ต่ำในระดับความรุนแรงที่ 1 ซึ่งเป็นกลุ่มที่มีข้อมูลน้อย โดยเฉพาะแบบจำลองต้นไม้ตัดสินใจที่มีค่าคะแนนเอฟหนึ่ง ค่าความแม่นยำ และค่าความระลึกเท่ากับ 0.44 ส่วนในระดับความรุนแรงที่ 0 ได้ค่าคะแนนเอฟหนึ่งเท่ากับ 0.86 สะท้อนถึงการเรียนรู้ของแบบจำลองไปที่ข้อมูลกลุ่มใหญ่มากกว่า โดยในแบบจำลองป่าสุ่ม ให้ผลลัพธ์ดีกว่าต้นไม้ตัดสินใจเพียงเล็กน้อย โดยในระดับความรุนแรงที่ 1 ได้ค่าคะแนนเอฟหนึ่งเท่ากับ 0.49 ค่าความแม่นยำเท่ากับ 0.64 ค่าความระลึกเท่ากับ 0.40 และระดับความรุนแรงที่ 0 ได้ค่าคะแนนเอฟหนึ่งเท่ากับ 0.91 ส่วนแบบจำลองเอ็กซ์ตรีม เกรเดียนต์ บูสตีง ในระดับความรุนแรงที่ 1 ได้ค่าคะแนนเอฟหนึ่ง เท่ากับ 0.50 ค่าความแม่นยำ 0.67 ค่าความระลึกเท่ากับ 0.41 และระดับความรุนแรงที่ 0 ได้ค่าคะแนนเอฟหนึ่งเท่ากับ 0.91 สำหรับโลท เกรเดียนต์ บูสตีง แมชชีน ให้ค่าใกล้เคียงกับ เอ็กซ์ตรีม เกรเดียนต์ บูสตีง โดยระดับความรุนแรงที่ 1 ได้ค่าคะแนนเอฟหนึ่งเท่ากับ 0.50 ค่าความแม่นยำเท่ากับ 0.69 ค่าความระลึกเท่ากับ 0.40 และระดับความรุนแรงที่ 0 ได้ค่าคะแนนเอฟหนึ่ง เท่ากับ 0.91 เช่นกัน ผลลัพธ์ของแบบจำลองทั้งหมดชี้ให้เห็นว่าแบบจำลองมีแนวโน้มจำแนกระดับความรุนแรงที่ 0 ได้ดีกว่าในทุกกรณีและจำแนกระดับความรุนแรงที่ 1 ได้ผลลัพธ์ที่น้อยกว่า โดยที่ระดับความรุนแรงที่ 1 มีจำนวนข้อมูลน้อยกว่าอย่างมีนัยสำคัญส่งผลให้เกิดปัญหาข้อมูลที่ไม่

สมดุลกัน (Class Imbalance) และจำเป็นต้องมีการปรับปรุงแบบจำลองเพิ่มเติมเพื่อให้สามารถเรียนรู้และจำแนกข้อมูลในระดับความรุนแรงที่ 1 ได้อย่างมีประสิทธิภาพ

จากปัญหาความไม่สมดุลของข้อมูลในระดับความรุนแรงที่ 1 ทำให้จำเป็นต้องดำเนินการปรับปรุงประสิทธิภาพของแบบจำลองด้วย การใช้ SMOTE การใช้ SMOTETomek และการปรับน้ำหนักของข้อมูล ซึ่งแต่ละวิธีส่งผลต่อค่าความแม่นยำ ค่าความระลึก และคะแนนเอฟหนึ่ง ที่แตกต่างกันไป โดยเฉพาะในระดับความรุนแรงที่ 1 ซึ่งเป็นคลาสที่เกิดผลกระทบจากความไม่สมดุลของข้อมูลจากการปรับปรุงประสิทธิภาพแบบจำลองโดยใช้ SMOTE พบว่าแบบจำลองต้นไม้ตัดสินใจให้ค่าคะแนนเอฟหนึ่งเท่ากับ 0.43 ค่าความแม่นยำเท่ากับ 0.38 และค่าความระลึกเท่ากับ 0.50 ในขณะที่ป่าสุ่มให้ค่าคะแนนเอฟหนึ่งเท่ากับ 0.52 ค่าความแม่นยำเท่ากับ 0.51 ค่าความระลึกเท่ากับ 0.54 และ เอ็กซ์ตรีม เกรเดียนต์ บูสติ้ง ให้ค่าคะแนนเอฟหนึ่งเท่ากับ 0.54 ค่าความแม่นยำเท่ากับ 0.50 ค่าความระลึกเท่ากับ 0.58 ส่วน โลท เกรเดียนต์ บูสติ้ง แมชชีน มีประสิทธิภาพดีที่สุดภายใต้เทคนิคนี้ โดยให้ ค่าคะแนนเอฟหนึ่งเท่ากับ 0.57 ค่าความแม่นยำเท่ากับ 0.51 และ ค่าความระลึกเท่ากับ 0.60 สำหรับการใช้ SMOTETomek ซึ่งเป็นการผสมการเพิ่มข้อมูลกับการลบข้อมูลที่ทับซ้อน พบว่าในแบบจำลองต้นไม้ตัดสินใจให้ผลใกล้เคียงเดิมคือค่าคะแนนเอฟหนึ่งเท่ากับ 0.44 ค่าความแม่นยำเท่ากับ 0.40 และค่าความระลึกเท่ากับ 0.50 ส่วนป่าสุ่มได้ค่าคะแนนเอฟหนึ่งเพิ่มขึ้นเล็กน้อยเป็น 0.55 ค่าความแม่นยำเท่ากับ 0.62 และค่าความระลึกเท่ากับ 0.50 ขณะที่แบบจำลอง เอ็กซ์ตรีม เกรเดียนต์ บูสติ้ง และ โลท เกรเดียนต์ บูสติ้ง แมชชีน แม้จะมีค่าความระลึกที่สูงคือ 0.76 และ 0.80 ตามลำดับ แต่ค่าความแม่นยำลดลงอย่างมากเหลือเพียง 0.34 และ 0.45 ตามลำดับ ทำให้ค่าคะแนนเอฟหนึ่งเท่ากับ 0.47 เท่ากันในทั้งสองแบบจำลอง แสดงให้เห็นว่า SMOTETomek แม้จะช่วยลดข้อมูลที่ไม่เกี่ยวข้อง (Noise) ได้ แต่อาจทำให้ข้อมูลจำแนกผิดพลาดมากขึ้นในกลุ่มข้อมูลที่มีความเฉพาะ นอกจากนี้การปรับปรุงประสิทธิภาพด้วยการปรับน้ำหนักของข้อมูลโดยใช้ค้นหาค่าพารามิเตอร์ที่ดีที่สุด (Hyperparameter Tuning) ด้วยการค้นหาค่าพารามิเตอร์แบบสุ่มร่วมกับการตรวจสอบไขว้ (Randomized Search Cross-Validation) ให้ผลลัพธ์ที่ดีที่สุดในภาพรวมของแบบจำลอง โดยเฉพาะในแบบจำลอง เอ็กซ์ตรีม เกรเดียนต์ บูสติ้ง ที่ให้ค่าคะแนนเอฟหนึ่งสูงถึง 0.68 ค่าความแม่นยำเท่ากับ 0.68 และค่าความระลึกเท่ากับ 0.67 ซึ่งดีกว่าทุกการปรับปรุงประสิทธิภาพที่ผ่านมา นอกจากนี้แบบจำลองป่าสุ่มให้ค่าคะแนนเอฟหนึ่งเท่ากับ 0.57 ค่าความแม่นยำเท่ากับ 0.53 และค่าความระลึกเท่ากับ 0.61 ขณะที่ต้นไม้ตัดสินใจเพิ่มขึ้นจากการใช้ SMOTEF โดยให้ค่าคะแนนเอฟหนึ่ง 0.54 ค่าความแม่นยำเท่ากับ 0.47 และค่าความระลึกเท่ากับ 0.65 และโลท เกรเดียนต์ บูสติ้ง แมชชีน ให้ค่าคะแนนเอฟหนึ่งเท่ากับ 0.54 ค่าความแม่นยำเท่ากับ 0.59 และค่าความระลึกเท่ากับ 0.53

จากการปรับปรุงประสิทธิภาพแบบจำลองพบว่าแบบจำลองที่ให้ผลลัพธ์ที่ดีที่สุดคือ เอ็กซ์ตรีม เกรเดียนต์ บูสติ้ง เมื่อใช้ร่วมกับการปรับปรุงประสิทธิภาพด้วยการปรับน้ำหนักของข้อมูลโดยค่าพารามิเตอร์ที่ให้ผลลัพธ์ที่ดีที่สุดประกอบด้วยความลึกสูงสุดของแบบจำลอง (Max Depth) เท่ากับ 9 ค่าอัตราการเรียนรู้ (Learning Rate) เท่ากับ 0.07 รอบการเรียนรู้ (Number of Estimators) เท่ากับ 113 และอัตราการสุ่มตัวอย่างข้อมูลการฝึกแบบจำลอง (Subsample) เท่ากับ 0.73 พบว่าค่าความแม่นยำในระดับความรุนแรงที่ 1 สูงถึง 0.68 ค่าความระลึกเท่ากับ 0.67 และค่าคะแนนเอฟหนึ่งเท่ากับ 0.68 ซึ่งสูงที่สุดเมื่อเทียบกับการปรับปรุงประสิทธิภาพและแบบจำลองอื่น ๆ นอกจากนี้ยังสามารถรักษาค่าประสิทธิภาพในระดับความรุนแรงที่ 0 ซึ่งให้ค่าคะแนนเอฟหนึ่งเท่ากับ 0.88 และค่าความถูกต้องโดยรวมสูงถึง 85% โดยเมื่อนำแบบจำลองดังกล่าวมาประเมินประสิทธิภาพด้วยวิธี 10-Fold Cross-Validation พบว่าค่าความแม่นยำในแต่ละรอบอยู่ระหว่าง 0.85 ถึง 0.86 และมีค่าเฉลี่ยอยู่ที่ 0.85 ซึ่งสะท้อนให้เห็นว่าแบบจำลองมีประสิทธิภาพในการจำแนกข้อมูลได้อย่างต่อเนื่องและความสมดุลในการจำแนกข้อมูลทั้งสองกลุ่มได้อย่างแม่นยำและครอบคลุม ส่งผลให้แบบจำลอง เอ็กซ์ตรีม เกรเดียนต์ บูสติ้ง ร่วมกับการปรับน้ำหนักข้อมูลเป็นทางเลือกที่เหมาะสมที่สุดสำหรับการจำแนกความรุนแรงของอุบัติเหตุในกรณีที่มีข้อมูลมีลักษณะไม่สมดุล โดยเมื่อนำแบบจำลองดังกล่าวไปเลือกคุณลักษณะที่ส่งผลต่อระดับความรุนแรงของอุบัติเหตุด้วย แซปลี่ แอ็ดดิทีฟ เอ็กซ์พลเนชันส์ (SHapley Additive exPlanations:SHAP) พบว่าคุณลักษณะที่มีผลมากที่สุดต่อการวิเคราะห์ระดับความรุนแรงของอุบัติเหตุคืออัตรามอเตอร์ไซค์ซึ่งแสดงให้เห็นว่าเมื่อค่าของคุณลักษณะนี้สูงขึ้น การวิเคราะห์ของแบบจำลองมีแนวโน้มเพิ่มขึ้นอย่างชัดเจน สะท้อนถึงความสัมพันธ์เชิงบวกระหว่างการเกี่ยวข้องกับรถจักรยานยนต์และระดับความรุนแรงของอุบัติเหตุ โดยที่จังหวัดที่เกิดเหตุ และช่วงเวลาที่เกิดเหตุ มีความสำคัญในลำดับถัดมา นอกจากนี้คุณลักษณะที่เกี่ยวข้องกับผลกระทบจากอุบัติเหตุ เช่น จำนวนรถที่เกิดเหตุและจำนวนคนกับรถที่เกิดเหตุ มีบทบาทสำคัญต่อระดับความรุนแรงเช่นกัน แสดงถึงความสัมพันธ์ระหว่างจำนวนรถหรือคนที่เกี่ยวข้องกับอุบัติเหตุ และความรุนแรงของเหตุการณ์ อย่างไรก็ตามคุณลักษณะบางตัวเช่น พื้นที่เกิดเหตุแบบทางตรง มูลเหตุสันนิฐานด้านการขาดความชำนาญ และอุบัติเหตุจากคนเดินถนนมีค่า SHAP ที่กระจายตัวน้อยมาก สะท้อนว่ามีอิทธิพลต่อแบบจำลองในระดับที่ต่ำ

5.1.2 สรุปผลการพัฒนาระบบสารสนเทศ

เมื่อการพัฒนาแบบจำลองทำได้แบบจำลองที่เหมาะสมที่สุดคือ เอ็กซ์ตรีม เกรเดียนต์ บูสติ้ง ร่วมกับการปรับปรุงประสิทธิภาพด้วยการปรับน้ำหนักแล้ว ได้นำแบบจำลองไปพัฒนาต่อยอดในระบบสารสนเทศ ซึ่งการพัฒนาระบบสารสนเทศนั้นได้พัฒนาโดยการเน้นการพัฒนาในรูปแบบการแสดงผลข้อมูลในรูปแบบแผนภูมิต่าง ๆ ตามความเหมาะสมของประเภทข้อมูล โดยผู้ใช้งานจะสามารถค้นหาข้อมูลอุบัติเหตุโดยเลือกเงื่อนไขต่าง ๆ ตามที่ต้องการดูข้อมูลในแต่ละประเภทได้

5.1.3 สรุปผลการประเมินคุณภาพและความพึงพอใจระบบสารสนเทศ

ในการประเมินจากผู้เชี่ยวชาญทั้ง 3 ท่านกับด้านการประเมินทั้ง 3 ด้านคือ ด้านความสามารถของระบบสารสนเทศตามความต้องการของผู้ใช้งาน ด้านการทำงานตามฟังก์ชัน และด้านความสะดวกและง่ายต่อการใช้งานระบบสารสนเทศ พบว่าในทุกท่านได้ค่าเฉลี่ยผลการประเมินคุณภาพโดยภาพรวมอยู่ที่ 4.39 คะแนน ซึ่งจัดอยู่ในระดับมาก และมีค่าเบี่ยงเบนมาตรฐานไปในทางเดียวกันซึ่งหมายถึงผู้เชี่ยวชาญทั้ง 3 ท่านโดยมีคะแนนอยู่ที่ 0.23 คะแนน โดยสรุปว่าผู้เชี่ยวชาญมีความเห็นไปในทิศทางเดียวกันซึ่งสอดคล้องกับการประเมินความพึงพอใจของผู้ใช้งาน ซึ่งมีผู้ประเมินทั้งสิ้น 30 ท่านประเมินความพึงพอใจใน 3 ด้านคือ ด้านความสามารถของระบบสารสนเทศตามความต้องการของผู้ใช้งาน ด้านการทดสอบการทำงานตามฟังก์ชันงาน และด้านความสะดวกและง่ายต่อการใช้งานระบบ พบว่าในทุกท่านได้ค่าเฉลี่ยผลการประเมินความพึงพอใจในภาพรวมอยู่ที่ 4.27 จัดอยู่ในระดับมาก และมีค่าเบี่ยงเบนมาตรฐานไปในทางเดียวกันซึ่งหมายถึงผู้เชี่ยวชาญทั้ง 30 ท่านมีความเห็นไปในทิศทางเดียวกันโดยมีคะแนนอยู่ที่ 0.25 คะแนน

5.2 อภิปรายผลการวิจัย

จากผลการค้นคว้าอิสระเรื่องการวิเคราะห์ระดับความรุนแรงของอุบัติเหตุทางถนนโดยใช้การเรียนรู้ของเครื่องสามารถนำมาอภิปรายผลการวิจัยได้ดังนี้

5.2.1 อภิปรายผลการพัฒนาแบบจำลอง

จากการพัฒนาแบบจำลองทั้ง 4 แบบจำลองคือต้นไม้ตัดสินใจ ป่าสุ่ม เอ็กซ์ตรีม เกรเดียนต์ บูสต์ และโลห์ เกรเดียนต์ บูสต์ แมชชีน พบว่าแบบจำลองทั้งหมดมีแนวโน้มให้ผลลัพธ์ที่ดีในระดับความรุนแรงที่ 0 แต่มีข้อจำกัดในการจำแนกระดับความรุนแรงที่ 1 ซึ่งเป็นกลุ่มข้อมูลที่มีปริมาณน้อยทำให้แบบจำลองไม่สามารถเรียนรู้ข้อมูลระดับความรุนแรงที่ 1 ได้ดีเพราะเกิดจากปัญหาข้อมูลไม่สมดุลกันของข้อมูล โดยเฉพาะแบบจำลองต้นไม้ตัดสินใจที่ให้ค่าคะแนนเอฟหนึ่งต่ำที่สุด สะท้อนว่าแบบจำลองนี้มีความซับซ้อนต่ำและไม่สามารถเรียนรู้ข้อมูลขนาดใหญ่ที่มีคลาสไม่เท่ากันได้ดี เพราะไม่สามารถจับลักษณะเฉพาะของข้อมูลกลุ่มน้อยได้อย่างมีประสิทธิภาพ โดยที่แบบจำลองป่าสุ่ม แม้จะมีโครงสร้างที่ซับซ้อนกว่าต้นไม้ตัดสินใจ แต่ไม่สามารถเพิ่มค่าความระลึกในระดับที่สูงขึ้นได้ ขณะที่เอ็กซ์ตรีม เกรเดียนต์ บูสต์ และโลห์ เกรเดียนต์ บูสต์ แมชชีน ซึ่งเป็นแบบจำลองเชิงกราดบูสต์ มีความสามารถในการเรียนรู้ข้อมูลเชิงลึกและเชิงลำดับ ส่งผลให้ให้ผลลัพธ์ที่ดีกว่าแบบจำลองอื่นเล็กน้อย แต่ยังมีข้อจำกัดในกรณีที่ไม่ได้รับการปรับปรุงข้อมูล

เมื่อพิจารณาการปรับปรุงประสิทธิภาพ พบว่า SMOTE สามารถช่วยเพิ่มค่าความระลึกในระดับความรุนแรงที่ 1 ได้อย่างชัดเจน โดยเฉพาะใน เอ็กซ์ตรีม เกรเดียนต์ บูสต์ และโลห์ เกรเดียนต์ บูสต์ แมชชีน ซึ่งมีโครงสร้างที่ยืดหยุ่นสูง อย่างไรก็ตามค่าความแม่นยำในหลายแบบจำลองกลับลดลงจากการสร้างข้อมูลสังเคราะห์ขึ้นมาในกลุ่มข้อมูลน้อย อาจส่งผลให้แบบจำลองเกิดความสับสน

ในเชิงลักษณะของคลาส และนำไปสู่การจำแนกผิดได้สูง ส่วนการปรับปรุงประสิทธิภาพด้วย SMOTETomek แม้มีวัตถุประสงค์เพื่อเพิ่มความสมดุลของข้อมูลและลดข้อมูลที่ไม่เกี่ยวข้อง แต่ในแบบจำลองที่ซับซ้อนสูง เช่น เอ็กซ์ตรีม เกรเดียนต์ บูสตีง และโลท กรเดียนต์ บูสตีง แมชชีน กลับให้ค่าความแม่นยำที่ต่ำลง แม้จะมีค่าความระลึกรู้สูงขึ้นก็ตาม แสดงให้เห็นถึงปัญหาการตัดข้อมูลที่อาจมีลักษณะสำคัญทิ้งไป โดยในขณะที่การปรับน้ำหนักของข้อมูลให้ผลลัพธ์ที่ดีและสม่ำเสมอมากที่สุด โดยเฉพาะแบบจำลอง เอ็กซ์ตรีม เกรเดียนต์ บูสตีง ซึ่งสามารถเรียนรู้จากข้อมูลกลุ่มน้อยได้ดี โดยไม่สูญเสียความแม่นยำในกลุ่มข้อมูลหลัก สะท้อนให้เห็นว่าการปรับปรุงประสิทธิภาพนี้เหมาะสมสำหรับชุดข้อมูลที่มีความไม่สมดุล และแบบจำลองที่สามารถปรับสมดุลระหว่างความแม่นยำและความครอบคลุมได้อย่างเหมาะสมจะมีความได้เปรียบในด้านประสิทธิภาพโดยรวมซึ่งสอดคล้องกับ Pichit Boonkrong และคณะในการศึกษาความสำคัญของการปรับสมดุลข้อมูลซึ่งพบว่า เอ็กซ์ตรีม เกรเดียนต์ บูสตีง ให้ผลลัพธ์ที่ดีที่สุด [28] นอกจากนี้ เมื่อทำการวิเคราะห์คุณลักษณะสำคัญด้วยวิธี SHAP พบว่าคุณลักษณะที่ส่งผลกระทบต่อระดับความรุนแรงของอุบัติเหตุมากที่สุดคือการเกี่ยวข้องกับรถจักรยานยนต์ ซึ่งสอดคล้องกับสถานการณ์ในประเทศไทยที่รถจักรยานยนต์เป็นยานพาหนะที่เกี่ยวข้องกับอุบัติเหตุจำนวนมากและมักมีความรุนแรงสูงสอดคล้องกับรายงานสถานการณ์อุบัติเหตุทางถนนของกรมควบคุมโรค กระทรวงสาธารณสุขประเทศไทยได้ระบุว่ารถจักรยานยนต์เป็นยานพาหนะที่เกิดเหตุสูงที่สุดถึงร้อยละ 83.86 จากกรณีที่เกิดเหตุทั้งหมด [2] รองลงมาคือจังหวัดที่เกิดเหตุและช่วงเวลาที่เกิดเหตุ สะท้อนถึงปัจจัยเชิงพื้นที่และเวลา ซึ่งเกี่ยวข้องกับโครงสร้างพื้นฐาน ความหนาแน่นของการจราจรโดยในจังหวัดที่เป็นถนนสายหลักมักจะเกิดอุบัติเหตุที่สูงเนื่องจากการจราจรที่หนาแน่น และในช่วงเวลากลางคืนส่งผลให้ทัศนวิสัยในการขับขี่ไม่ดีเท่ากลางวันสอดคล้องกับ Shakil Ahmed และคณะที่ทำนายระดับความรุนแรงของอุบัติเหตุบนถนนในประเทศนิวซีแลนด์โดยใช้ SHAP [14] พบว่าสภาพของแสงและการขับขี่เวลากลางคืนทำให้การตอบสนองต่อเหตุการณ์ฉุกเฉินลดลง ในขณะที่คุณลักษณะที่เกี่ยวข้องกับผลกระทบ เช่น จำนวนรถที่เกี่ยวข้องและจำนวนคนที่เกี่ยวข้อง ก็มีความสำคัญเช่นกัน โดยชี้ให้เห็นว่าเหตุการณ์ที่มีการเกี่ยวข้องของบุคคลและพาหนะมากมักสัมพันธ์กับความรุนแรงที่เพิ่มขึ้น ในทางกลับกัน คุณลักษณะบางตัว เช่น ลักษณะถนนตรงมูลเหตุจากความไม่ชำนาญ หรือการเกี่ยวข้องกับคนเดินถนน มีค่า SHAP กระจายตัวต่ำ แสดงว่ามีอิทธิพลต่อการตัดสินใจของแบบจำลองน้อย

5.2.2 อภิปรายผลการพัฒนาระบบสารสนเทศ

จากการพัฒนาแบบจำลองการทำนายระดับความรุนแรงของอุบัติเหตุ พบว่าแบบจำลองที่ให้ผลลัพธ์ที่ดีที่สุดคือ เอ็กซ์ตรีม เกรเดียนต์ บูสตีง โดยมีการปรับปรุงประสิทธิภาพเพิ่มเติมด้วยการปรับน้ำหนักข้อมูล ซึ่งช่วยลดปัญหาคลาสไม่สมดุลและเพิ่มความแม่นยำในการทำนายระดับความรุนแรงได้อย่างมีนัยสำคัญ สอดคล้องกับงานวิจัยที่ผ่านมา ที่ระบุว่า เอ็กซ์ตรีม เกรเดียนต์ บูสตีง เป็น

แบบจำลองที่มีความสามารถในการจัดการกับข้อมูลที่ไม่สมดุล [28] และสามารถเรียนรู้ความสัมพันธ์ที่ซับซ้อนได้ดี นอกจากนี้ การนำแบบจำลองที่ได้ไปพัฒนาต่อยอดในรูปแบบของระบบสารสนเทศ ยังแสดงให้เห็นถึงความเป็นไปได้ในการประยุกต์ใช้ผลการวิเคราะห์ในเชิงปฏิบัติ โดยระบบสารสนเทศถูกออกแบบให้สามารถแสดงผลข้อมูลในรูปแบบแผนภูมิต่าง ๆ ได้ตามเหมาะสมของข้อมูลซึ่งช่วยให้ผู้ใช้งานสามารถเข้าใจและวิเคราะห์ข้อมูลได้ง่ายยิ่งขึ้น โดยเฉพาะในด้านการ ค้นหาข้อมูลอุบัติเหตุ ที่ช่วยให้ผู้ใช้งานสามารถค้นหาข้อมูลตามที่ตนเองสนใจ เพื่อเป็นการช่วยให้ผู้ใช้สามารถเข้าถึงข้อมูลได้ในเชิงลึกมากยิ่งขึ้น

5.2.3 อภิปรายผลการประเมินคุณภาพและความพึงพอใจระบบสารสนเทศ

จากผลการประเมินจากผู้เชี่ยวชาญจำนวน 3 ท่านในด้านต่าง ๆ ได้แก่ ด้านความสามารถของระบบสารสนเทศตามความต้องการของผู้ใช้งาน ด้านการทำงานตามฟังก์ชัน และด้านความสะดวกและง่ายต่อการใช้งาน พบว่าค่าเฉลี่ยของผลการประเมินคุณภาพระบบอยู่ที่ 4.39 คะแนน ซึ่งจัดอยู่ในระดับ มาก แสดงให้เห็นว่าผู้เชี่ยวชาญมีมุมมองในเชิงบวกต่อคุณภาพของระบบสารสนเทศ โดยมีค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.23 ซึ่งอยู่ในระดับต่ำ แสดงถึงความเห็นที่สอดคล้องกันของผู้เชี่ยวชาญทั้ง 3 ท่านอย่างชัดเจน โดยเมื่อพิจารณาพร้อมกับผลการประเมินความพึงพอใจของผู้ใช้งานจำนวน 30 ราย ในประเด็นเดียวกัน ได้แก่ ความสามารถของระบบตามความต้องการ การทำงานตามฟังก์ชัน และความสะดวกในการใช้งาน พบว่ามีค่าเฉลี่ยความพึงพอใจรวมเท่ากับ 4.27 คะแนน ซึ่งอยู่ในระดับ มาก เช่นกัน และมีค่าเบี่ยงเบนมาตรฐานเท่ากับ 0.25 สะท้อนถึงความเห็นของผู้ใช้งานที่มีความสอดคล้องกันในระดับที่ดี นอกจากนี้ผลการประเมินทั้งจากผู้เชี่ยวชาญและผู้ใช้งานทั่วไปชี้ให้เห็นว่าระบบสารสนเทศที่พัฒนาขึ้นมีความเหมาะสมทั้งในแง่ของการออกแบบ การทำงาน และประสิทธิภาพใช้งาน ซึ่งถือเป็นปัจจัยสำคัญที่สะท้อนถึงความสำเร็จของกระบวนการพัฒนาระบบ โดยเฉพาะในบริบทของการนำเทคโนโลยีมาช่วยสนับสนุนการวิเคราะห์และจัดการข้อมูลอุบัติเหตุทางถนน นอกจากนี้ ความเห็นที่เป็นไปในทิศทางเดียวกันทั้งจากผู้เชี่ยวชาญและผู้ใช้งาน ยังเป็นตัวบ่งชี้ถึงความน่าเชื่อถือและความสามารถในการตอบสนองต่อความต้องการของผู้ใช้งานอย่างแท้จริง

5.3 ข้อเสนอแนะการวิจัย

5.3.1 จากผลการค้นคว้าอิสระพบว่า ข้อมูลที่ใช้ในการพัฒนาแบบจำลองมีลักษณะไม่สมดุลระหว่างคลาส โดยเฉพาะในกรณีของข้อมูลระดับความรุนแรง 1 ที่มีจำนวนน้อยกว่าข้อมูลระดับ 0 อย่างมีนัยสำคัญ ซึ่งส่งผลให้แบบจำลองไม่สามารถเรียนรู้และจำแนกข้อมูลในคลาสดังกล่าวได้อย่างมีประสิทธิภาพ ดังนั้นในการวิจัยในอนาคต ควรพิจารณาใช้ชุดข้อมูลที่มีความสมดุลมากขึ้น หรือมีการรวบรวมข้อมูลเพิ่มเติมในคลาสดังกล่าวที่มีจำนวนน้อย เพื่อเพิ่มความหลากหลายและความครอบคลุมของข้อมูล

5.3.2 จากการค้นคว้าอิสระแม้การปรับน้ำหนักของคลาส จะให้ผลลัพธ์ที่ดีที่สุดในงานวิจัยนี้ แต่ควรมีการทดลองเพิ่มเติมโดยใช้เทคนิคการจัดการข้อมูลที่ไม่สมดุลแบบอื่น เช่น ADASYN, Cluster Centroids หรือการใช้แบบจำลองแบบรวมกัน (Ensemble Models) แบบเฉพาะทาง เพื่อเปรียบเทียบประสิทธิภาพอย่างรอบด้าน นอกจากนี้ยังควรศึกษาการปรับค่าพารามิเตอร์ แบบ Grid Search หรือการปรับแต่งแบบเบย์เซียน (Bayesian Optimization) เพื่อให้ได้แบบจำลองที่เหมาะสมกับลักษณะของข้อมูลมากที่สุด และสามารถนำไปประยุกต์ใช้ในสถานการณ์จริงได้อย่างแม่นยำ

5.3.3 การพัฒนาระบบสารสนเทศในรูปแบบเว็บแอปพลิเคชันเพื่อแสดงผลข้อมูลและวิเคราะห์อุบัติเหตุทางถนนสามารถตอบสนองความต้องการของผู้ใช้งานได้ดีในระดับหนึ่ง อย่างไรก็ตาม ในการพัฒนาระบบต่อไปควรเพิ่มฟังก์ชันที่สามารถวิเคราะห์เชิงลึกมากขึ้น เช่น การเปรียบเทียบรายพื้นที่ การแสดงผลในรูปแบบแผนที่ GIS การวิเคราะห์แนวโน้มรายปี รวมถึงการเปิดให้ผู้ใช้งานสามารถปรับพารามิเตอร์การวิเคราะห์ได้ด้วยตนเอง นอกจากนี้ควรให้ความสำคัญกับการออกแบบระบบให้มีความยืดหยุ่น รองรับการขยายระบบในอนาคต และสามารถเชื่อมต่อกับฐานข้อมูลจากหน่วยงานภาครัฐ หรือแหล่งข้อมูลเปิด (Open Data) เพื่อให้ระบบมีความทันสมัยและใช้งานได้อย่างยั่งยืน

บรรณานุกรม

1. World Health Organization, United Nations specialized agency. Global Status Report on Road Safety 2023. 1st ed. Geneva: World Health Organization, 2023.
2. ควบคุมโรค, กรม. สถานการณ์อุบัติเหตุทางท้องถนนประเทศไทย. นนทบุรี : กรมควบคุมโรค กระทรวงสาธารณสุข, 2567.
3. Neumann, Tomasz. “Analysis of Advanced Driver-Assistance Systems for Safe and Comfortable Driving of Motor Vehicles.” Sensors (Basel, Switzerland), vol. 24, no. 19, 2024
4. สำนักงานปลัดกระทรวงคมนาคม. (2567). [ออนไลน์]. MOT DATA CATALOG. [สืบค้นวันที่ 20 สิงหาคม 2567]. จาก <https://datagov.mot.go.th/dataset/roadaccident>
5. Suchkov, A. I. “Provoked Road Traffic Accident: Investigative Tactics.” LEGAL ORDER: History, Theory, Practice, vol. 43, no. 4, 2024, pp. 154–157
6. Bhele, R., & Rajchal, P. (2023). “Accident analysis and identification of black spots on road section from Sanga to Dhulikhel in araniko highway.” Journal of Transportation Systems, 8(3), 18–32.
7. Asst. Professor, Dept. of Computer Science & Engineering, Rajeev Institute of Technology, Hassan, and Ruksar Parveen. “Road Accident Detection Using Maching Learning.” INTERANTIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT, vol. 09, no. 02, 2025, pp. 1–9,
8. Geetha, T. V., & Sendhilkumar, S. (2023). Machine learning: Concepts, techniques and applications. Philadelphia, PA: Chapman & Hall/CRC.
9. Santos, Daniel, et al. “Machine Learning Approaches to Traffic Accident Analysis and Hotspot Prediction.” Computers, vol. 10, no. 12, 2021, p. 157

10. PES Modern College of Engineering, Pune, and Pratham Shimpi. "Reinforcement Learning in Real Life Applications." INTERANTIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT, vol. 09, no. 01, 2025, pp. 1–9
11. Chowdhury, Shourav, et al. "Computer Vision-Based Road Accident Classification from Traffic Surveillance." Preprints, 2025. [Online]. Available: <https://www.preprints.org/manuscript/xxxx>
12. SCB TechX. "Confusion Matrix: The Key to Precision in Classification Analysis." SCB TechX, <https://scbtechx.io/th/blogs/tech-tips-for-life-confusion-matrix-the-key-to-precision-in-classification-analysis/>. Accessed 1 Feb. 2025.
13. Gorasia, Meet, et al. "Road Accident Analysis and Classification Using Machine Learning Algorithms." International Journal for Research in Applied Science and Engineering Technology, vol. 12, no. 5, 2024, pp. 472–476,
14. Ahmed, S., Hossain, M. A., Ray, S. K., Bhuiyan, M. M. I., & Sabuj, S. R. (2023). "A study on road accident prediction and contributing factors using explainable machine learning models: analysis and performance" Transportation Research Interdisciplinary Perspectives, 19(100814), 100814.
15. Kavana V H and Kavana V H "Machine Learning-Based Predictions of Traffic Accidents Severeness." JNNCE Journal of Engineering and Management, special issue SP02, Mar. 2025, pp. 118-123.
16. Mustafa, Dheya, et al. "Predictive Analysis of Jordan Road Accidents Severity Using Machine Learning." 2023 International Symposium on Networks, Computers and Communications (ISNCC), IEEE, 2023.
17. Iliyasu, U., Yakudu, M. M., & Abdulwasiiu, A. A. (2023). "Predictive analysis of road traffic accidents in Katsina State, Nigeria using machine learning algorithms: A study on factors and mitigation strategies." International Journal of Science for Global Sustainability, 9(2), 188–198.

18. Cheng, C.-H., Yang, J.-H., & Liu, P.-C. (2022). "Rule-based classifier based on accident frequency and three-stage dimensionality reduction for exploring the factors of road accident injuries". PloS One, 17(8), e0272956.
19. Berrones-Sanz, L., Perez-Diaz, E., Becerril, D., & Diaz, E. (2023). "Predicting crash injury severity in road freight flows with association rules algorithms." Acta Logistica, 10(3), 413–422.
20. Shafay, M., Anwar, H. N., Bibi, I., Samad, A., Akhtar, N., Rasheed, A., & Kazi, M. A. (2024). "Seasonal and time-series analysis of road Traffic Accidents: Seasonal analysis of road accidents. Pakistan Journal of Health Sciences", pp.121–125.
21. Datu, N. H. (2023). "Road traffic accidents analysis using association rule mining and descriptive analytics." AIP Conference Proceedings, 2701, 020003. AIP Publishing.
22. Twenefour, F. B. K., Ayitey, E., Kangah, J., & Brew, L. (2021). "Time Series analysis of road traffic accidents in Ghana. Asian Journal of Probability and Statistics." 12–20.
23. Wang, Xin. "Predicting and Analysing Road Accident Severity with Machine Learning Models and Resampling." Transactions on Computer Science and Intelligent Systems Research, vol. 5, 2024, pp. 928–934
24. Kumar, Anukali Pramod, and D. Teja Santosh. "Road Accident Severity Prediction Using Machine Learning Algorithms." International Journal of Computer Engineering in Research Trends, vol. 9, no. 9, 2022, pp. 175-183.
25. Pourroostaei Ardakani, Saeid, et al. "Road Car Accident Prediction Using a Machine-Learning-Enabled Data Analysis." Sustainability, vol. 15, no. 7, 2023, article 5939

26. Ballamudi, Venkata Koteswara Rao. "Road Accident Analysis and Prediction using Machine Learning Algorithmic Approaches." Asian Journal of Humanities and Applied Linguistics, i-Proclaim, vol. 5, no. 9, pp.195-213
27. Mallahi, Imad El, et al. "Severity of Road Accident Classification Using Machine Learning Models." Research Square, 2023,
28. Boonkrong, Pichit, and Teerawat Simmachan. "Does Data Balancing Play an Important Role in Classifying Road Accident Severity in Thailand?" The 10th International Conference on Digital Arts, Media and Technology (DAMT) and 8th ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (NCON), Jan. 2025
29. Mbarek, Abdelilah, et al. "Accident Black Spots Identification Based on Association Rule Mining." Bulletin of Electrical Engineering and Informatics, vol. 13, no. 3, 2024, pp. 2075–2085
30. Ballamudi, Koteswara Rao. "Road Accident Analysis and Prediction Using Machine Learning Algorithmic Approaches." Asian Journal of Humanity Art and Literature, vol. 6, no. 2, 2019, pp. 185–192
31. Wang, Xin. "Predicting and Analysing Road Accident Severity with Machine Learning Models and Resampling." Transactions on Computer Science and Intelligent Systems Research, vol. 5, 2024, pp. 928–934,
32. Alqatawna, Ali. "Highway Traffic Accidents Prediction Using Artificial Neural Network: A Case Study of Freeways in Spain." Journal of Engineering Research, vol. 4, no. 1, 2024, pp. 2–14
33. Davies, Owunari S., Ette H. Etuk, and Emeka Amos. "Time Series Analysis of Road Traffic Accident in Rivers State, Nigeria." Ianna Journal of Interdisciplinary Studies, vol. 1, no. 1, Oct. 2019, pp. 120–133.



ภาคผนวก ก

รายชื่อผู้เชี่ยวชาญ และหนังสือขอความอนุเคราะห์เป็นผู้เชี่ยวชาญ

รายชื่อผู้เชี่ยวชาญ

1. ชื่อ : นางสาวชลันดา มุลมี
ตำแหน่ง : หัวหน้ากลุ่มพัฒนาระบบคมนาคมอัจฉริยะ
สถานที่ทำงาน : สำนักงานปลัดกระทรวงคมนาคม
วุฒิการศึกษา : บริหารธุรกิจมหาบัณฑิต
สถานบันการศึกษา : มหาวิทยาลัยบูรพา สาขาวิชาการจัดการสาธารณะ
ความเชี่ยวชาญเฉพาะด้าน : การนำเทคโนโลยีอัจฉริยะมาใช้ในการพัฒนาระบบงาน และบริการด้านคมนาคมขนส่ง
2. ชื่อ : นางสาวจุฑารัตน์ คันธีระ
ตำแหน่ง : นักวิชาการคอมพิวเตอร์ชำนาญการ
สถานที่ทำงาน : สำนักงานปลัดกระทรวงคมนาคม
วุฒิการศึกษา : วิทยาศาสตร์มหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ
สถานบันการศึกษา : มหาวิทยาลัยศรีปทุม
ความเชี่ยวชาญเฉพาะด้าน : วิเคราะห์ ออกแบบ เว็บไซต์และโมบายแอปพลิเคชัน และการจัดการสารสนเทศองค์กร
3. ชื่อ : นายสิริวิทย์ เสมสวัสต์
ตำแหน่ง : นักวิชาการคอมพิวเตอร์ชำนาญการ
สถานที่ทำงาน : นักงานเลขาธิการวุฒิสภา
วุฒิการศึกษา : วิทยาศาสตร์มหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ
สถานบันการศึกษา : มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
ความเชี่ยวชาญเฉพาะด้าน : วิเคราะห์ ออกแบบ พัฒนาเว็บไซต์และโมบายแอปพลิเคชัน

ที่ อว ๗๑๐๗.๒๖๒



คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล
มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
๑๕๑๘ ถนนประชากราษฎร์ ๑ แขวงวงศ์สว่าง
เขตบางซื่อ กทม. ๑๐๘๐๐

๕ มีนาคม ๒๕๖๘

เรื่อง ขอความอนุเคราะห์เป็นผู้เชี่ยวชาญในการประเมินคุณภาพเพื่อประกอบการทำการค้นคว้าอิสระ

เรียน นางสาวชลันดา มุลมี หัวหน้ากลุ่มพัฒนาระบบคมนาคมอัจฉริยะ สำนักงานปลัดกระทรวงคมนาคม

เนื่องด้วยนายสุรชัย เดชมี รหัสประจำตัว ๖๖-๐๗๐๑๑๘-๕๗๑๗-๐ นักศึกษาระดับบัณฑิตศึกษา ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ กำลังดำเนินการจัดทำ การค้นคว้าอิสระ เรื่อง “การวิเคราะห์ระดับความรุนแรงของอุบัติเหตุทางถนนโดยใช้การเรียนรู้ของเครื่อง” โดยมี ผู้ช่วยศาสตราจารย์ ดร.มุสตี บุญรอด เป็นประธานคณะกรรมการที่ปรึกษาการค้นคว้าอิสระ ในการนี้นักศึกษามีความประสงค์ขอความอนุเคราะห์จากท่านเพื่อโปรดพิจารณาแบบประเมินเพื่อประกอบการทำการค้นคว้าอิสระดังกล่าว

จึงเรียนมาเพื่อโปรดพิจารณาให้ความอนุเคราะห์ จักขอบพระคุณยิ่ง

ขอแสดงความนับถือ

(ผู้ช่วยศาสตราจารย์ ดร.ศักดิ์ชาย ตั้งวรรณวิทย์)
รองคณบดีฝ่ายวิชาการและวิจัย

ภาควิชาเทคโนโลยีสารสนเทศ
โทร. ๐ ๒๕๕๕ ๒๐๐๐ ต่อ ๒๗๒๘
www.itd.kmutnb.ac.th

ที่ อว ๗๑๐๗/๒๖๓



คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล
มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
๑๕๑๘ ถนนประชากรราษฎร์ ๑ แขวงวงศ์สว่าง
เขตบางซื่อ กทม. ๑๐๘๐๐

๕ มีนาคม ๒๕๖๘

เรื่อง ขอความอนุเคราะห์เป็นผู้เชี่ยวชาญในการประเมินคุณภาพเพื่อประกอบการทำการค้นคว้าอิสระ

เรียน นางสาวจุฑารัตน์ คันธีระ นักวิชาการคอมพิวเตอร์ชำนาญการ สำนักงานปลัดกระทรวงคมนาคม

เนื่องด้วยนายสุรัชย์ เตชมี รหัสประจำตัว ๖๖-๐๗๐๑๑๘-๕๗๑๗-๐ นักศึกษาระดับบัณฑิตศึกษา ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ กำลังดำเนินการจัดทำการศึกษาค้นคว้าอิสระ เรื่อง “การวิเคราะห์ระดับความรุนแรงของอุบัติเหตุทางถนนโดยใช้การเรียนรู้ของเครื่อง” โดยมี ผู้ช่วยศาสตราจารย์ ดร.สุสดี บุญรอด เป็นประธานคณะกรรมการที่ปรึกษาการค้นคว้าอิสระ ในการนี้นักศึกษามีความประสงค์ขอความอนุเคราะห์จากท่านเพื่อโปรดพิจารณาแบบประเมินเพื่อประกอบการทำการค้นคว้าอิสระดังกล่าว

จึงเรียนมาเพื่อโปรดพิจารณาให้ความอนุเคราะห์ จักขอบพระคุณยิ่ง

ขอแสดงความนับถือ

(ผู้ช่วยศาสตราจารย์ ดร.ศักดิ์ชาย ตั้งวรรณวิทย์)
รองคณบดีฝ่ายวิชาการและวิจัย

ภาควิชาเทคโนโลยีสารสนเทศ
โทร. ๐ ๒๕๕๕ ๒๐๐๐ ต่อ ๒๗๒๘
www.itd.kmutnb.ac.th

ที่ อว ๗๑๐๗/๒๖๔



คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล
มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ
๑๕๑๘ ถนนประชากราษฎร์ ๑ แขวงวงศ์สว่าง
เขตบางซื่อ กทม. ๑๐๘๐๐

๕ มีนาคม ๒๕๖๘

เรื่อง ขอความอนุเคราะห์เป็นผู้เชี่ยวชาญในการประเมินคุณภาพเพื่อประกอบการทำการค้นคว้าอิสระ

เรียน นายสิริวิทย์ เสมสวัสดิ์ นักวิชาการคอมพิวเตอร์ชำนาญการ สำนักงานเลขาธิการวุฒิสภา

เนื่องด้วยนายสุรชัย เดชมี รหัสประจำตัว ๖๖-๐๗๐๑๑๘-๕๗๑๗-๐ นักศึกษาระดับบัณฑิตศึกษา ภาควิชาเทคโนโลยีสารสนเทศ คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ กำลังดำเนินการจัดทำการศึกษาอิสระ เรื่อง “การวิเคราะห์ระดับความรุนแรงของอุบัติเหตุทางถนนโดยใช้การเรียนรู้ของเครื่อง” โดยมี ผู้ช่วยศาสตราจารย์ ดร.สุสดี บุญรอด เป็นประธานคณะกรรมการที่ปรึกษาการค้นคว้าอิสระ ในการนี้นักศึกษามีความประสงค์ขอความอนุเคราะห์จากท่านเพื่อโปรดพิจารณาแบบประเมินเพื่อประกอบการทำการค้นคว้าอิสระดังกล่าว

จึงเรียนมาเพื่อโปรดพิจารณาให้ความอนุเคราะห์ จักขอบพระคุณยิ่ง

ขอแสดงความนับถือ

(ผู้ช่วยศาสตราจารย์ ดร.ศักดิ์ชาย ตั้งวรรณวิทย์)

รองคณบดีฝ่ายวิชาการและวิจัย

ภาควิชาเทคโนโลยีสารสนเทศ
โทร. ๐ ๒๕๕๕ ๒๐๐๐ ต่อ ๒๗๒๘
www.itd.kmutnb.ac.th



ภาคผนวก ข

แบบประเมินคุณภาพระบบสารสนเทศสำหรับผู้เชี่ยวชาญ
ระบบสารสนเทศการแสดงผลข้อมูลอุบัติเหตุทางถนนในประเทศไทย

**แบบประเมินคุณภาพระบบสารสนเทศสำหรับผู้เชี่ยวชาญ
ระบบสารสนเทศการแสดงผลข้อมูลอุบัติเหตุทางถนนในประเทศไทย**

นายสุรัชย์ เดชมี รหัสประจำตัวนักศึกษา 66-070118-5717-0

นักศึกษาระดับปริญญาโท สาขาวิชาวิทยาศาสตร์ข้อมูลเพื่อนวัตกรรม

คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

คำชี้แจง

แบบประเมินคุณภาพนี้มีวัตถุประสงค์เพื่อให้ผู้เชี่ยวชาญประเมินคุณภาพและแสดงความคิดเห็นของ ระบบสารสนเทศการแสดงผลข้อมูลอุบัติเหตุทางถนนในประเทศไทย จากการค้นคว้าอิสระ ระดับปริญญาโท สาขาวิชาวิทยาศาสตร์ข้อมูลเพื่อนวัตกรรม คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ ซึ่งข้อมูลจากการตอบแบบประเมินคุณภาพของท่านจะถูกนำไปใช้ในการค้นคว้าอิสระและพัฒนาระบบสารสนเทศการแสดงผลข้อมูลอุบัติเหตุทางถนนในประเทศไทยเท่านั้น โดยแบบประเมินคุณภาพแบ่งออกได้เป็น 3 ตอนดังนี้

ตอนที่ 1 ข้อมูลทั่วไปของผู้เชี่ยวชาญ

ตอนที่ 2 ความคิดเห็นของผู้เชี่ยวชาญเกี่ยวกับคุณภาพของระบบสารสนเทศการแสดงผลข้อมูลอุบัติเหตุทางถนนในประเทศไทย

ตอนที่ 3 ข้อเสนอแนะและข้อแนะนำอื่น ๆ

ตอนที่ 1 ข้อมูลทั่วไปของผู้เชี่ยวชาญ

คำชี้แจง โปรดทำเครื่องหมายถูก (✓) ลงในช่อง ☐ และกรอกข้อมูลที่ตรงกับสถานภาพที่เป็นจริงของท่าน

1. เพศ ☐ ชาย ☐ หญิง ☐ อื่น ๆ

2. เพศ ☐ น้อยกว่า 20 ปี ☐ 20 – 30 ปี
☐ 31 - 40 ปี ☐ 41 – 50 ปี
☐ 51 - 60 ปี ☐ มากกว่า 60 ปีขึ้นไป

3. ระดับการศึกษา ☐ ปริญญาตรี ☐ ปริญญาโท
☐ ปริญญาเอก ☐ อื่น ๆ (โปรดระบุ).....

4. สถาบันที่สำเร็จการศึกษา.....

.....

5. สถานที่ทำงาน.....

6. ตำแหน่งงาน.....

.....

7. ประสบการณ์ทำงาน

☐ น้อยกว่า 1 ปี ☐ ระหว่าง 1 – 5 ปี
☐ ระหว่าง 6 – 10 ปี ☐ ระหว่าง 11 – 15 ปี
☐ ระหว่าง 16 – 20 ปี ☐ มากกว่า 20 ปี

8. ความเชี่ยวชาญ.....

.....

ตอนที่ 2 ความคิดเห็นของผู้เชี่ยวชาญเกี่ยวกับคุณภาพของระบบสารสนเทศการแสดงผลข้อมูลอุบัติเหตุทางถนนในประเทศไทย

คำชี้แจง แบบประเมินคุณภาพของระบบสารสนเทศการแสดงผลข้อมูลอุบัติเหตุทางถนนในประเทศไทย ได้แบ่งการประเมินคุณภาพออกเป็น 3 ด้านหลัก ได้แก่ ด้านความสามารถของระบบสารสนเทศตามความต้องการของผู้ใช้งาน (Functional Requirement Test) ด้านการทำงานได้ตามฟังก์ชันงานของระบบสารสนเทศ (Functional Test) และด้านความสะดวกและง่ายต่อการใช้งานระบบสารสนเทศ (Usability Test) ซึ่งมีรายละเอียดดังนี้

- 1) ด้านความสามารถของระบบสารสนเทศตามความต้องการของผู้ใช้งาน เป็นการประเมินคุณภาพความสามารถของระบบสารสนเทศที่ระบบสมควรทำได้ตามความต้องการของผู้ใช้งาน
- 2) ด้านการทำงานได้ตามฟังก์ชันงานของระบบสารสนเทศ เป็นการประเมินคุณภาพความถูกต้องของการทำงานของระบบสารสนเทศว่าสามารถทำงานได้จนได้ผลลัพธ์ตรงตามฟังก์ชันงานนั้น ๆ
- 3) ด้านความสะดวกและง่ายต่อการใช้งานระบบสารสนเทศ เป็นการประเมินคุณภาพการออกแบบและการใช้งานระบบสารสนเทศ

โปรดทำเครื่องหมายถูก (✓) ในช่องระดับความพึงพอใจ ที่ตรงกับระดับความคิดเห็นของท่านมากที่สุด ซึ่งมีระดับความพึงพอใจจำนวน 5 ระดับ ดังนี้

- 5 หมายถึง ระบบที่พัฒนาขึ้นมีคุณภาพในระดับดีมาก
- 4 หมายถึง ระบบที่พัฒนาขึ้นมีคุณภาพในระดับดี
- 3 หมายถึง ระบบที่พัฒนาขึ้นมีคุณภาพในระดับปานกลาง
- 2 หมายถึง ระบบที่พัฒนาขึ้นมีคุณภาพในระดับน้อย
- 1 หมายถึง ระบบที่พัฒนาขึ้นมีคุณภาพในระดับน้อยมาก

ตารางที่ ข-1 แบบประเมินคุณภาพด้านความสามารถของระบบสารสนเทศตามความต้องการของผู้ใช้งาน

รายการประเมิน	ระดับคุณภาพ				
	5	4	3	2	1
1.1 ระบบสามารถระบุเส้นทางเริ่มต้นและเส้นทางปลายทางได้					
1.2 ระบบสามารถแสดงข้อมูลอุบัติเหตุระหว่างเส้นทางที่ระบุได้					
1.3 ระบบสามารถวิเคราะห์อุบัติเหตุที่เคยเกิดขึ้นระหว่างเส้นทางได้					
1.4 ระบบสามารถแสดงข้อมูลอุบัติเหตุบนพิกัดของแผนที่ได้					
1.5 ระบบสามารถสืบค้นข้อมูลต่าง ๆ					
1.6 ระบบสามารถเรียกใช้งาน API จาก datagov.mot.go.th ได้					
1.7 ระบบสามารถแสดงผลจากการระบุ latitude และ Longitude ได้จากการเรียกข้อมูลผ่าน API					
1.8 ระบบสามารถเรียกใช้งานไลบรารี (Library) หรือปลั๊กอิน (Plugin) ได้					
1.9 ระบบสามารถแจ้งเตือนเมื่อไม่พบข้อมูลที่ต้องการ หรือพบข้อผิดพลาดระหว่างใช้งาน					
1.10 ความสามารถของระบบสารสนเทศในภาพรวม					

ตารางที่ ข-2 การประเมินความพึงพอใจด้านความสะดวกและง่ายต่อการใช้งานระบบ

รายการประเมิน	ระดับความพึงพอใจ				
	5	4	3	2	1
2.1 ระบบสามารถระบุเส้นทางเริ่มต้นและเส้นทางปลายทางได้อย่างถูกต้อง					
2.2 ระบบสามารถแสดงข้อมูลอุบัติเหตุระหว่างเส้นทางที่ระบุได้อย่างถูกต้อง					
2.3 ระบบสามารถวิเคราะห์อุบัติเหตุที่เคยเกิดขึ้นระหว่างเส้นทางได้อย่างถูกต้อง					
2.4 ระบบสามารถแสดงข้อมูลอุบัติเหตุบนพิกัดของแผนที่ได้อย่างถูกต้อง					
2.5 ระบบสามารถสืบค้นข้อมูลต่าง ๆ ได้อย่างถูกต้อง					
2.6 ระบบสามารถเรียกใช้งาน API จาก datagov.mot.go.th ได้อย่างถูกต้อง					
2.7 ระบบสามารถแสดงผลจากการระบุ latitude และ Longitude ได้จากการเรียกข้อมูลผ่าน API ได้อย่างถูกต้อง					
2.8 ระบบสามารถเรียกใช้งานไลบรารี (Library) หรือปลั๊กอิน (Plugin) ได้อย่างถูกต้อง					
2.9 ระบบสามารถแจ้งเตือนเมื่อไม่พบข้อมูลที่ต้องการ หรือพบข้อผิดพลาดระหว่างใช้งานได้อย่างถูกต้อง					
2.10 ความสามารถของระบบสารสนเทศในภาพรวม					

ตารางที่ ข-3 แบบประเมินคุณภาพด้านความสะดวกและง่ายต่อการใช้งานระบบสารสนเทศ

รายการประเมิน	ระดับคุณภาพ				
	5	4	3	2	1
3.1 ความเหมาะสมของการใช้กราฟในการแสดงผลข้อมูล					
3.2 ความเหมาะสมของการจัดวางตำแหน่งต่าง ๆ (Layout) ภายในหน้าเว็บไซต์					
3.3 ความเหมาะสมของการเลือกใช้สีพื้นหลังและสีของกราฟ					
3.4 ความเหมาะสมของการเลือกใช้สีตัวอักษร ประเภทของตัวอักษร(font) และขนาดตัวอักษร					
3.5 ความเหมาะสมของข้อมูลที่ใช้ในการนำเสนอ					
3.6 ข้อความภายในกราฟให้ความหมายที่ชัดเจนและเข้าใจได้ง่าย					
3.7 ระบบสามารถใช้งานได้ง่าย ไม่ซับซ้อน					
3.8 ระบบมีความครอบคลุมและความหลากหลายของข้อมูลที่แสดงบน Dashboard					
3.9 ระบบสามารถทำงานได้อย่างต่อเนื่องโดยไม่มีปัญหาการล่มหรือค้าง					

ตารางที่ ข-3 (ต่อ)

รายการประเมิน	ระดับคุณภาพ				
	5	4	3	2	1
3.10 ระบบสามารถใช้งานได้ดีทั้งบนคอมพิวเตอร์และอุปกรณ์มือถือ หรืออุปกรณ์อื่น ๆ (Responsive)					
3.11 ระบบสามารถตอบสนองต่อคำสั่งของผู้ใช้งานได้อย่างรวดเร็ว					
3.12 ระบบสารสนเทศมีความสะดวกและง่ายต่อการใช้งานโดยภาพรวม					

ตอนที่ 3 ข้อเสนอแนะและข้อแนะนำอื่น ๆ

1. ด้านความสามารถของระบบสารสนเทศตามความต้องการของผู้ใช้งาน

.....

.....

2. ด้านการทดสอบการทำงานตามฟังก์ชันการใช้งาน

.....

.....

3. ด้านความสะดวกและง่ายต่อการใช้งานระบบ

.....

.....

4. ข้อเสนอแนะข้อแนะนำอื่น ๆ

.....

.....

.....

(.....)

ภาคผนวก ค

แบบประเมินความพึงพอใจระบบสารสนเทศสำหรับผู้ใช้งานทั่วไป
ระบบสารสนเทศการแสดงผลข้อมูลอุบัติเหตุทางถนนในประเทศไทย



แบบประเมินความพึงพอใจระบบสารสนเทศสำหรับผู้ใช้งานทั่วไป
ระบบสารสนเทศการแสดงผลข้อมูลอุบัติเหตุทางถนนในประเทศไทย

นายสุรัชย์ เดชมี รหัสประจำตัวนักศึกษา 66-070118-5717-0

นักศึกษาระดับปริญญาโท สาขาวิชาวิทยาศาสตร์ข้อมูลเพื่อนวัตกรรม

คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ

คำชี้แจง

แบบประเมินความพึงพอใจนี้มีวัตถุประสงค์เพื่อรวบรวมความคิดเห็น ความพึงพอใจและข้อเสนอแนะจากผู้ใช้งานทั่วไปต่อการใช้งานระบบสารสนเทศการแสดงผลข้อมูลอุบัติเหตุทางถนนในประเทศไทย ซึ่งเป็นการค้นคว้าอิสระ ระดับปริญญาโท สาขาวิชาวิทยาศาสตร์ข้อมูลเพื่อนวัตกรรม คณะเทคโนโลยีสารสนเทศและนวัตกรรมดิจิทัล มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ ข้อมูลจากการตอบแบบประเมินความพึงพอใจของท่านจะถูกนำไปใช้ในการค้นคว้าอิสระและพัฒนา ระบบสารสนเทศการแสดงผลข้อมูลอุบัติเหตุทางถนนในประเทศไทยเท่านั้น โดยแบบประเมินความพึงพอใจแบ่งออกได้เป็น 3 ตอนดังนี้

ตอนที่ 1 ข้อมูลทั่วไปของผู้ใช้งานทั่วไป

ตอนที่ 2 ความคิดเห็นของผู้ใช้งานทั่วไปเกี่ยวกับความพึงพอใจของระบบสารสนเทศการแสดงผลข้อมูลอุบัติเหตุทางถนนในประเทศไทย

ตอนที่ 3 ข้อเสนอแนะและข้อแนะนำอื่น ๆ

ตอนที่ 1 ข้อมูลทั่วไปของผู้ใช้งานทั่วไป

คำชี้แจง โปรดทำเครื่องหมายถูก (✓) ลงในช่อง ☐ และกรอกข้อมูลที่ตรงกับสถานภาพที่เป็นจริงของท่าน

1. เพศ ☐ ชาย ☐ หญิง ☐ อื่น ๆ

2. อายุ ☐ น้อยกว่า 20 ปี ☐ 20 – 30 ปี
☐ 31 - 40 ปี ☐ 41 – 50 ปี
☐ 51 - 60 ปี ☐ มากกว่า 60 ปีขึ้นไป

3. ระดับการศึกษา ☐ ปริญญาตรี ☐ ปริญญาโท
☐ ปริญญาเอก ☐ อื่น ๆ (โปรดระบุ).....

4. สถาบันที่สำเร็จการศึกษา.....

.....

5. สถานที่ทำงาน.....

6. ตำแหน่งงาน.....

.....

7. ประสบการณ์ทำงาน

☐ น้อยกว่า 1 ปี ☐ ระหว่าง 1 – 5 ปี
☐ ระหว่าง 6 – 10 ปี ☐ ระหว่าง 11 – 15 ปี
☐ ระหว่าง 16 – 20 ปี ☐ มากกว่า 20 ปี

ตอนที่ 2 ความคิดเห็นของผู้ใช้งานทั่วไปเกี่ยวกับความพึงพอใจของระบบสารสนเทศการแสดงผลข้อมูลอุบัติเหตุทางถนนในประเทศไทย

คำชี้แจง แบบประเมินความพึงพอใจระบบสารสนเทศการแสดงผลข้อมูลอุบัติเหตุทางถนนในประเทศไทย ที่พัฒนาขึ้น โดยแบ่งการประเมินความพึงพอใจออกเป็น 3 ด้านหลัก ได้แก่ ด้านความสามารถของระบบสารสนเทศตามความต้องการของผู้ใช้งาน (Functional Requirement Test) ด้านการทำงานได้ตามฟังก์ชันงานของระบบสารสนเทศ (Functional Test) และด้านความสะดวกและง่ายต่อการใช้งานระบบสารสนเทศ (Usability Test) ซึ่งมีรายละเอียดดังนี้

- 1) ด้านความสามารถของระบบสารสนเทศตามความต้องการของผู้ใช้งาน เป็นการทดสอบความสามารถของระบบสารสนเทศที่ระบบสมควรทำได้ตามความต้องการของผู้ใช้งาน
- 2) ด้านการทำงานได้ตามฟังก์ชันงานของระบบสารสนเทศ เป็นการทดสอบความถูกต้องของการทำงานของระบบสารสนเทศว่าสามารถทำงานได้จนได้ผลลัพธ์ตรงตามฟังก์ชันงานนั้น ๆ
- 3) ด้านความสะดวกและง่ายต่อการใช้งานระบบสารสนเทศ เป็นการทดสอบการออกแบบและการใช้งานระบบสารสนเทศ

โปรดทำเครื่องหมายถูก (✓) ในช่องระดับความพึงพอใจ ที่ตรงกับระดับความคิดเห็นของท่านมากที่สุด ซึ่งมีระดับความพึงพอใจจำนวน 5 ระดับ ดังนี้

- 5 หมายถึง มีความพึงพอใจต่อระบบที่พัฒนาขึ้นในระดับดีมาก
- 4 หมายถึง มีความพึงพอใจต่อระบบที่พัฒนาขึ้นในระดับดี
- 3 หมายถึง มีความพึงพอใจต่อระบบที่พัฒนาขึ้นในระดับปานกลาง
- 2 หมายถึง มีความพึงพอใจต่อระบบที่พัฒนาขึ้นในระดับน้อย
- 1 หมายถึง มีความพึงพอใจต่อระบบที่พัฒนาขึ้นในระดับน้อยมาก

ตารางที่ ค-1 การประเมินความพึงพอใจด้านความสามารถของระบบสารสนเทศตามความต้องการ
ของผู้ใช้งาน

รายการประเมิน	ระดับความพึงพอใจ				
	5	4	3	2	1
1.1 ระบบสามารถระบุเส้นทางเริ่มต้นและ เส้นทางปลายทางได้					
1.2 ระบบสามารถแสดงข้อมูลอุบัติเหตุ ระหว่างเส้นทางที่ระบุได้					
1.3 ระบบสามารถแสดงข้อมูลอุบัติเหตุบน พิกัดของแผนที่ได้					
1.4 ระบบสามารถสืบค้นข้อมูลต่าง ๆ ได้					
1.5 ระบบสามารถแจ้งเตือนเมื่อไม่พบข้อมูลที่ ต้องการ หรือพบข้อผิดพลาดระหว่างใช้งาน					
1.6 ความสามารถของระบบสารสนเทศใน ภาพรวม					

ตารางที่ ค-2 การประเมินความพึงพอใจด้านการทดสอบการทำงานตามฟังก์ชันงาน

รายการประเมิน	ระดับความพึงพอใจ				
	5	4	3	2	1
2.1 ระบบสามารถระบุเส้นทางเริ่มต้นและ เส้นทางปลายทางได้อย่างถูกต้อง					
2.2 ระบบสามารถแสดงข้อมูลอุบัติเหตุ ระหว่างเส้นทางที่ระบุได้อย่างถูกต้อง					
2.3 ระบบสามารถแสดงข้อมูลอุบัติเหตุบน พิกัดของแผนที่ได้อย่างถูกต้อง					
2.4 ระบบสามารถสืบค้นข้อมูลตามเงื่อนไขที่ ระบุได้อย่างถูกต้อง					

ตารางที่ ค-2 (ต่อ)

รายการประเมิน	ระดับความพึงพอใจ				
	5	4	3	2	1
2.5 ระบบสามารถแจ้งเตือนเมื่อไม่พบข้อมูลที่ต้องการ หรือพบข้อผิดพลาดระหว่างใช้งานได้อย่างถูกต้อง					
2.6 ความถูกต้องของระบบสารสนเทศในภาพรวม					

ตารางที่ ค-3 การประเมินความพึงพอใจด้านความสะดวกและง่ายต่อการใช้งานระบบ

รายการประเมิน	ระดับความพึงพอใจ				
	5	4	3	2	1
3.1 ความเหมาะสมของการใช้กราฟในการแสดงผลข้อมูล					
3.2 ความเหมาะสมของการจัดวางตำแหน่งต่าง ๆ (Layout) ภายในหน้าเว็บไซต์					
3.3 ความเหมาะสมของการเลือกใช้สีพื้นหลังและสีของกราฟ					
3.4 ความเหมาะสมของการเลือกใช้สีตัวอักษร ประเภทของตัวอักษร(font) และขนาดตัวอักษร					
3.5 ความเหมาะสมของข้อมูลที่ใช้ในการนำเสนอ					
3.6 ข้อความภายในกราฟให้ความหมายที่ชัดเจนและเข้าใจได้ง่าย					
3.7 ระบบสามารถใช้งานได้ง่าย ไม่ซับซ้อน					
3.8 ระบบสารสนเทศมีความสะดวกและง่ายต่อการใช้งานโดยภาพรวม					

ตอนที่ 3 ข้อเสนอแนะและข้อแนะนำอื่น ๆ

1. ด้านความสามารถของระบบสารสนเทศตามความต้องการของผู้ใช้งาน

.....

.....

.....

2. ด้านการทดสอบการทำงานตามฟังก์ชันการใช้งาน

.....

.....

.....

3. ด้านความสะดวกและง่ายต่อการใช้งานระบบ

.....

.....

.....

4. ข้อเสนอแนะข้อแนะนำอื่น ๆ

.....

.....

.....

(.....)



ภาคผนวก ง

คู่มือการใช้งานระบบสารสนเทศ

คู่มือการใช้งานระบบสารสนเทศ

ระบบสารสนเทศการแสดงผลข้อมูลอุบัติเหตุทางถนนในประเทศไทย มีคู่มือการใช้งานระบบสารสนเทศ โดยเมื่อเข้าสู่หน้าระบบสารสนเทศแล้ว ระบบจะเรียกใช้ข้อมูลทั้งหมดเพื่อมาแสดงผลในหน้าจอแสดงผล ดังภาพที่ ง-1



ภาพที่ ง-1 ระบบสารสนเทศแสดงข้อมูลทั้งหมด

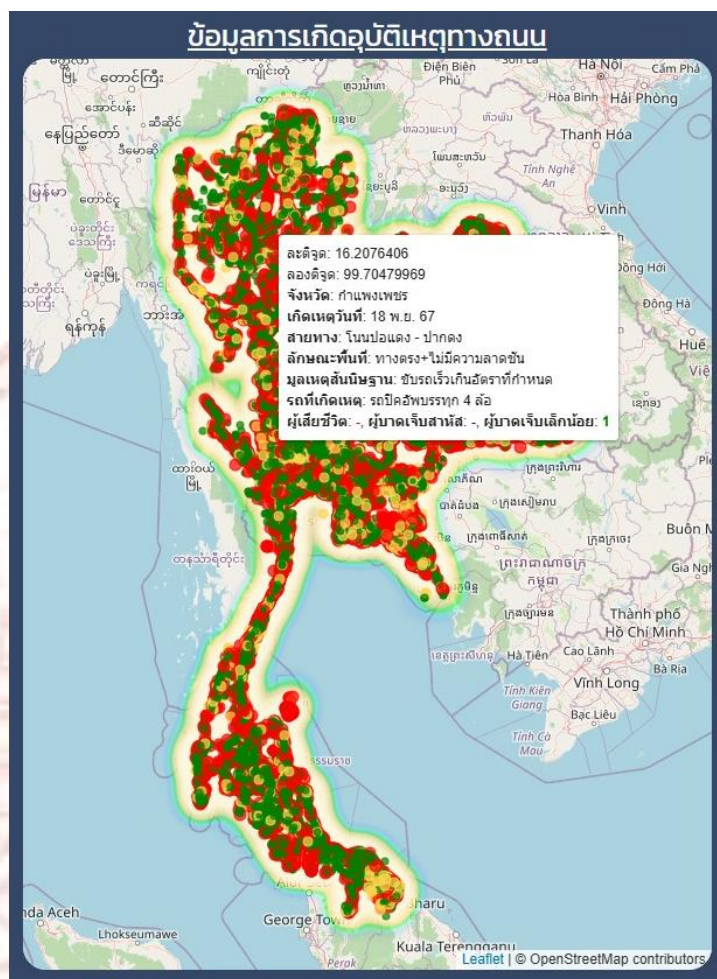
1. การค้นหาข้อมูลการเกิดอุบัติเหตุ

เมื่อผู้ใช้งานต้องการค้นหาข้อมูลอุบัติเหตุสามารถค้นหาได้ที่หัวข้อ “ค้นหาข้อมูล” โดยสามารถเลือกค้นหา ปีเริ่มต้น ปีสิ้นสุด จังหวัด และประเภทรถที่เกิดเหตุ ดังภาพที่ ง-2



ภาพที่ ง-2 ส่วนการค้นหาข้อมูล

จากภาพที่ ง-2 เมื่อผู้ใช้งานกรอกเงื่อนไขที่ต้องการค้นหาเรียบร้อยแล้วสามารถกดปุ่ม “ค้นหา” เพื่อแสดงข้อมูลตามเงื่อนไขที่ต้องการ โดยเมื่อกดปุ่มค้นหา จะแสดงดังภาพที่ ง-3



ภาพที่ ง-4 ข้อมูลการเกิดเหตุบนแผนที่

ประวัติผู้เขียน

ชื่อ	นายสุรชัย เดชมี
ชื่อการค้นคว้าอิสระ	การวิเคราะห์ระดับความรุนแรงของอุบัติเหตุทางถนนโดยใช้การเรียนรู้ของเครื่อง
สาขาวิชา	วิทยาศาสตร์ข้อมูลเพื่อนวัตกรรม ภาควิชาเทคโนโลยีสารสนเทศ
ประวัติ	ประวัติส่วนตัว เกิดวันที่ 18 มกราคม พ.ศ. 2540 สถานที่ติดต่อบ้านเลขที่ 29 หมู่ 4 ตำบลห้วยใหญ่ อำเภอหนองบัว จังหวัดนครสวรรค์ 60110 Email:s6607011857170@email.kmutnb.ac.th ประวัติการศึกษาระดับปริญญาตรี สำเร็จการศึกษาระดับปริญญาตรี วิทยาศาสตร์บัณฑิต คณะวิทยาศาสตร์ สาขาวิชาวิทยาการคอมพิวเตอร์ มหาวิทยาลัยนเรศวร ปี พ.ศ. 2562 ประวัติการทำงาน ปี พ.ศ. 2564 – 2567 ปฏิบัติหน้าที่นักวิชาการคอมพิวเตอร์ปฏิบัติการสำนักงานปลัดกระทรวงคมนาคม ปี พ.ศ. 2567 ถึงปัจจุบันปฏิบัติงานที่ สำนักงานเลขาธิการวุฒิสภา ตำแหน่ง นักวิชาการคอมพิวเตอร์ปฏิบัติการ