| | |
|---|---|
| **ชื่อเรื่องภาษาไทย** | การวิเคราะห์และพัฒนาเว็บแอปพลิเคชันเพื่อทำนายความเสี่ยงของอุบัติเหตุบนท้องถนนด้วยการเรียนรู้ของเครื่อง |
| **ชื่อภาษาอังกฤษ** | Analysis and Development of a Web Application for Predicting Road Accident Risk Using Machine Learning |
| **ผู้วิจัย** | นางสาวสุวิมล ชื่นทิม |
| **ที่ปรึกษาวิทยานิพนธ์** | ผศ.ดร.กัมปนาท ปิยะธำรงชัย |
| **ประเภทสารนิพนธ์ วิทยานิพนธ์** | วิทยาศาสตรบัณฑิต สาขาวิชาภูมิศาสตร์, มหาวิทยาลัยนเรศวร,2567 |

| | | |
|---|---|---|
| (ผู้ช่วยศาสตราจารย์ ดร.กัมปนาท ปิยะธำรงชัย) | (อาจารย์ธัญญาลัคน์ จันทร์สมบัติ) | (ผู้ช่วยศาสตราจารย์ ดร. รังสรรค์ เกตุอิอด) |
| อาจารย์ที่ปรึกษา | ประธานบริหารหลักสูตร | หัวหน้าภาควิชา |
| วิทยานิพนธ์ระดับปริญญาตรี | วิทยาศาสตรบัณฑิต สาขาภูมิศาสตร์ | ทรัพยากรธรรมชาติและสิ่งแวดล้อม |

## ABSTRACT

This research aims to (1) study and assess the risk levels of road accidents in Thailand using accident data from 2019 to 2024, (2) analyze factors influencing the severity of accidents to gain insight into factors associated with accident occurrences, and (3) develop a web application for predicting road accident risk using machine learning techniques. This study employs a quantitative research approach, utilizing real accident data from Thailand, which is applied to a Random Forest algorithm to build a predictive model for accident severity. The results are presented through a web application. Descriptive statistics are used to describe the data features, and model performance is evaluated by dividing data samples to test accuracy. The findings show that machine learning techniques can effectively predict accident risk, and the information obtained can be used to inform policy and enhance preventive measures to reduce accident occurrences more effectively.



## บทคัดย่อ

บทความวิจัยนี้มีวัตถุประสงค์เพื่อ (1) ศึกษาและประเมินระดับความเสี่ยงในการเกิดอุบัติเหตุบนถนนในประเทศไทย โดยใช้ข้อมูลอุบัติเหตุในช่วงปี 2562 ถึง 2567 (2) วิเคราะห์ปัจจัยที่ส่งผลต่อความรุนแรงของอุบัติเหตุ เพื่อให้เข้าใจถึงปัจจัยที่เกี่ยวข้องกับการเกิดอุบัติเหตุ และ (3) พัฒนาเว็บแอปพลิเคชันสำหรับทำนายความเสี่ยงของอุบัติเหตุบนถนนโดยใช้เทคนิคการเรียนรู้ของเครื่อง การวิจัยนี้เป็นการวิจัยเชิงปริมาณ โดยใช้ข้อมูลอุบัติเหตุจริงในประเทศไทย ซึ่งนำมาประยุกต์ใช้กับอัลกอริทึม ป่าสุ่ม ในการสร้างแบบจำลองทำนายความรุนแรงของอุบัติเหตุ และแสดงผลในรูปแบบเว็บแอปพลิเคชัน สำหรับการวิเคราะห์ข้อมูลใช้สถิติพรรณนาในการอธิบายคุณลักษณะข้อมูล รวมถึงการประเมินประสิทธิภาพของแบบจำลองด้วยการแบ่งกลุ่มตัวอย่างข้อมูลเพื่อตรวจสอบความแม่นยำ ผลการวิจัยพบว่าการใช้เทคนิคการเรียนรู้ของเครื่องช่วยให้สามารถทำนายความเสี่ยงในการเกิดอุบัติเหตุได้อย่างมีประสิทธิภาพ โดยข้อมูลที่ได้สามารถนำไปใช้ประโยชน์ในการกำหนดนโยบายและวางมาตรการป้องกันการเกิดอุบัติเหตุให้มีประสิทธิภาพมากยิ่งขึ้น



ลิขสิทธิ์มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved

# กิตติกรรมประกาศ

วิทยานิพนธ์ระดับปริญญาตรีฉบับนี้เสร็จสมบูรณ์ด้วยการสนับสนุนและคำแนะนำจากหลายฝ่ายอย่างดียิ่ง โดยเฉพาะอย่างยิ่งอาจารย์ที่ปรึกษาดร.กัมปนาท ปิยะธำรงชัย ผู้ให้คำปรึกษา แนะนำและตรวจแก้ไขผลงานตลอดการดำเนินงานวิจัย ผู้วิจัยรู้สึกซาบซึ้งและขอขอบพระคุณอาจารย์เป็นอย่างสูง

ขอขอบพระคุณ ดร.กัมปนาท ปิยะธำรงชัย ที่ได้ให้คำแนะนำอันมีค่าและสละเวลาอันมีค่าตรวจทานผลงานในทุกขั้นตอน นอกจากนี้ยังต้องขอขอบคุณทุกท่านที่เกี่ยวข้องในการให้คำแนะนำเชิงเทคนิคและการสนับสนุนในด้านการพัฒนาและทดสอบเว็บแอปพลิเคชัน ซึ่งมีความสำคัญอย่างยิ่งต่อความสำเร็จของโครงการวิจัย “การวิเคราะห์และพัฒนาเว็บแอปพลิเคชันเพื่อทำนายความเสี่ยงของอุบัติเหตุบนท้องถนนด้วยการเรียนรู้ของเครื่อง”

สุดท้ายนี้ ขอขอบคุณครอบครัวและเพื่อน ๆ ที่ให้กำลังใจและการสนับสนุนทางจิตใจตลอดเส้นทางการทำวิจัย ผู้วิจัยรู้สึกซาบซึ้งในความกรุณาของทุกท่าน จึงขอกราบขอบพระคุณทุกท่านไว้อีกครั้ง ณ ที่นี้

สุวิมล ชื่นทิม

ลิขสิทธิ์ มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved

## ที่มาและความสำคัญ

อุบัติเหตุบนถนนเป็นหนึ่งในปัญหาที่มีความสำคัญระดับโลก โดยเฉพาะในยุคที่การคมนาคมเป็นปัจจัยหลักที่ขับเคลื่อนเศรษฐกิจและสังคม จำนวนผู้เสียชีวิตและบาดเจ็บจากอุบัติเหตุบนถนนเพิ่มขึ้นอย่างต่อเนื่องในหลายประเทศ ส่งผลกระทบทั้งในด้านชีวิตและทรัพย์สิน รวมถึงสร้างภาระทางเศรษฐกิจและสังคมอย่างมหาศาลการลดจำนวนอุบัติเหตุและความรุนแรงของอุบัติเหตุเป็นเป้าหมายสำคัญของหลายๆ หน่วยงานที่เกี่ยวข้อง ทั้งในภาครัฐบาลและภาคเอกชน แม้ว่าจะมีการบังคับใช้กฎหมายและมาตรการต่างๆ เพื่อเพิ่มความปลอดภัย แต่จำนวนอุบัติเหตุก็ยังคงเกิดขึ้นอย่างต่อเนื่อง สาเหตุของอุบัติเหตุบนถนนสามารถแบ่งออกเป็นหลายปัจจัย เช่น ความประมาทของผู้ขับขี่ การใช้ความเร็วเกินกำหนด การขับรถภายใต้อิทธิพลของแอลกอฮอล์หรือสารเสพติด สภาพถนนที่ไม่ปลอดภัย และการขาดการบังคับใช้กฎหมายอย่างเคร่งครัด นอกจากนี้ความซับซ้อนของการจราจรในเมืองใหญ่ สภาพอากาศที่ไม่เอื้ออำนวย และพฤติกรรมการขับขี่ที่ไม่เหมาะสมของผู้ใช้ถนน ต่างก็เป็นปัจจัยที่ส่งผลให้เกิดอุบัติเหตุได้เช่นกัน

ประเทศไทยเป็นหนึ่งในประเทศที่มีอัตราการเกิดอุบัติเหตุทางรถสูงที่สุดในโลก การเสียชีวิตและบาดเจ็บจากอุบัติเหตุทางรถกลายเป็นปัญหาสำคัญที่มีผลกระทบทั้งต่อครอบครัว ผู้คนในชุมชน และเศรษฐกิจของประเทศ โดยเฉพาะในช่วงเทศกาลหรือวันหยุดยาว ที่มีการเดินทางของประชาชนจำนวนมาก ทำให้จำนวนอุบัติเหตุเพิ่มขึ้นอย่างรวดเร็ว ผลกระทบของอุบัติเหตุบนถนนไม่เพียงแต่จำกัดอยู่ที่ความเสียหายทางร่างกายและจิตใจของผู้ประสบเหตุ แต่ยังรวมถึงผลกระทบทางเศรษฐกิจที่เกิดจากการสูญเสียแรงงาน และค่าใช้จ่ายในการรักษาพยาบาล รวมถึงผลกระทบทางสังคม เช่น การสูญเสียบุคคลอันเป็นที่รัก ในปี 2566 ประเทศไทยมีสถิติอุบัติเหตุทางถนนที่ยังคงสูงอยู่ โดยเฉพาะในช่วงเทศกาลและวันหยุดยาว ซึ่งเป็นช่วงเวลาที่มักเกิดอุบัติเหตุจำนวนมาก จากข้อมูลจนถึงเดือนตุลาคม 2566 มีรายงานการเกิดอุบัติเหตุบนท้องถนนรวมประมาณ 159,000 ครั้ง ซึ่งก่อให้เกิดการบาดเจ็บเล็กน้อยถึง 61.86% ของทั้งหมด ส่วนอุบัติเหตุที่ทำให้เกิดบาดเจ็บสาหัสมีสัดส่วนประมาณ 20% และมีผู้เสียชีวิตอย่างต่อเนื่อง ซึ่งจากจำนวนสถิติการเกิดอุบัติเหตุจากการรวบรวมพบว่า ในช่วงเทศกาลสงกรานต์ที่ผ่านมา (11-17 เมษายน 2566) มีรายงานว่าเกิดอุบัติเหตุรวมกว่า 2,203 ครั้งทั่วประเทศ ส่งผลให้มีผู้เสียชีวิต 264 ราย และบาดเจ็บ 2,208 คน ซึ่งถือว่าเป็นการเกิดอุบัติเหตุสูงสุดในปี 2566 และสูงขึ้นกว่าปี 2565 จำนวน 286 ราย โดยสาเหตุมาจากการขับรถเร็วเกินกำหนดและการขับขี่ขณะมึนเมาเป็นสาเหตุสำคัญของการเกิดอุบัติเหตุ ซึ่งแสดงให้เห็นถึงความจำเป็นในการเพิ่มความเข้มงวดในการบังคับใช้กฎหมายและการรณรงค์เพื่อความปลอดภัยในการขับขี่

ลิขสิทธิ์มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved

ในปัจจุบัน เทคโนโลยีที่ก้าวหน้ามีบทบาทสำคัญในการพัฒนาระบบความปลอดภัยของยานพาหนะ รวมถึงการวิจัยและการพัฒนารถยนต์ที่มีระบบช่วยขับอัตโนมัติและเทคโนโลยีความปลอดภัยขั้นสูง เช่น ระบบเตือนการชน ระบบเบรกอัตโนมัติ และระบบควบคุมความเร็ว ซึ่งถูกออกแบบมาเพื่อช่วยลดความเสี่ยงและความรุนแรงของอุบัติเหตุโดยการนำ Machine Learning (ML) มาใช้ในการวิเคราะห์อุบัติเหตุรถยนต์ได้รับความนิยมมากขึ้น เนื่องจากสามารถช่วยทำนายและป้องกันอุบัติเหตุได้อย่างมีประสิทธิภาพ โดยการประยุกต์ใช้ ML ในการวิเคราะห์ข้อมูลอุบัติเหตุรถยนต์สามารถทำได้ในหลายด้าน อาทิการทำนายอุบัติเหตุ ML สามารถทำนายความน่าจะเป็นของการเกิดอุบัติเหตุในสภาพแวดล้อมต่างๆ ได้ การตรวจจับพฤติกรรมการขับขี่ที่ไม่ปลอดภัย โดยการใช้เซ็นเซอร์และอุปกรณ์วัดต่างๆ เช่น กล้องวิดีโอและเซ็นเซอร์ในรถยนต์ร่วมกับอัลกอริทึม Machine Learning ช่วยในการตรวจจับพฤติกรรมการขับขี่ที่เสี่ยง เช่น การขับเร็วเกินกำหนด การเบรกกระทันหัน การเปลี่ยนเลนที่ไม่ปลอดภัย หรือการขับรถโดยใช้โทรศัพท์มือถือ การออกแบบระบบช่วยขับขี่ นำมาใช้ทำระบบควบคุมความเร็วอัตโนมัติ (Adaptive Cruise Control), และระบบช่วยจอดรถ โดยการเรียนรู้จากพฤติกรรมและข้อมูลการขับขี่ในสถานการณ์ต่างๆ การวิเคราะห์สาเหตุของอุบัติเหตุ สามารถวิเคราะห์ข้อมูลขนาดใหญ่จากเหตุการณ์อุบัติเหตุหลายล้านกรณี การพัฒนานโยบายและมาตรการความปลอดภัย การวิเคราะห์ข้อมูลโดยใช้ Machine Learning ช่วยให้ผู้กำหนดนโยบายสามารถพัฒนามาตรการความปลอดภัยที่ตรงจุด เช่น การกำหนดเขตจำกัดความเร็ว การติดตั้งไฟจราจร และการวางแผนเส้นทางการขนส่งที่ปลอดภัยมากขึ้น การบำรุงรักษาเชิงพยากรณ์ ML ถูกนำมาใช้ในด้านการบำรุงรักษารถยนต์ เช่น การพยากรณ์ว่าชิ้นส่วนใดอาจเสียหายหรือเสื่อมสภาพ ซึ่งจะช่วยลดความเสี่ยงของอุบัติเหตุที่เกิดจากปัญหาทางกลไก

จากความสำคัญและที่มาของปัญหาที่กล่าวมาข้างต้น ผู้วิจัยจึงมีความสนใจที่จะศึกษา การวิเคราะห์และพัฒนาเว็บแอปพลิเคชันเพื่อทำนายความเสี่ยงของอุบัติเหตุบนท้องถนนด้วยการเรียนรู้ของเครื่อง โดยใช้เครื่องมือMachine Learning (ML) มาใช้ในการวิเคราะห์อุบัติเหตุรถยนต์เพื่อเป็นแนวทางในการทำความเข้าใจถึงเหตุปัจจัย การพัฒนานโยบายและมาตรการความปลอดภัย ที่มีผลต่อการเกิดอุบัติเหตุ ซึ่งจะเป็นประโยชน์ในการวางแผนป้องกัน เพื่อลดจำนวนอุบัติเหตุ ลดความเสี่ยงและความรุนแรงของอุบัติเหตุในประเทศไทยได้

ลิขสิทธิ์มหาวิทยาลัยนเรศวร
Copyright© by Naresuan University
All rights reserved

**วัตถุประสงค์**

1) เพื่อวิเคราะห์ลักษณะการเกิดอุบัติเหตุทางถนนในประเทศไทย

2) เพื่อสร้างแบบจำลองทำนายความเสี่ยงในการเกิดอุบัติเหตุโดยใช้การเรียนรู้ของเครื่อง

3) เพื่อพัฒนาเว็บแอปพลิเคชันแสดงข้อมูลอุบัติเหตุทางถนนน ปี2019-2024 และเพื่อทำนายความเสี่ยงในการเกิดอุบัติเหตุ

**คำถามงานวิจัย**

1) การนำ Machine Learning มาวิเคราะห์ข้อมูลอุบัติเหตุบนถนนสามารถช่วยทำนายความเสี่ยงของอุบัติเหตุได้อย่างไร?

**ประโยชน์ที่คาดว่าจะได้รับ**

1) เว็บแอปพลิเคชันสำหรับการประเมินและทำนายความเสี่ยงของอุบัติเหตุบนท้องถนน ที่สามารถใช้งานได้จริง ซึ่งจะช่วยให้ผู้ใช้ถนนและหน่วยงานที่เกี่ยวข้องสามารถนำข้อมูลไปใช้ในการตัดสินใจและวางแผนเพื่อลดความเสี่ยงของอุบัติเหตุได้อย่างมีประสิทธิภาพ

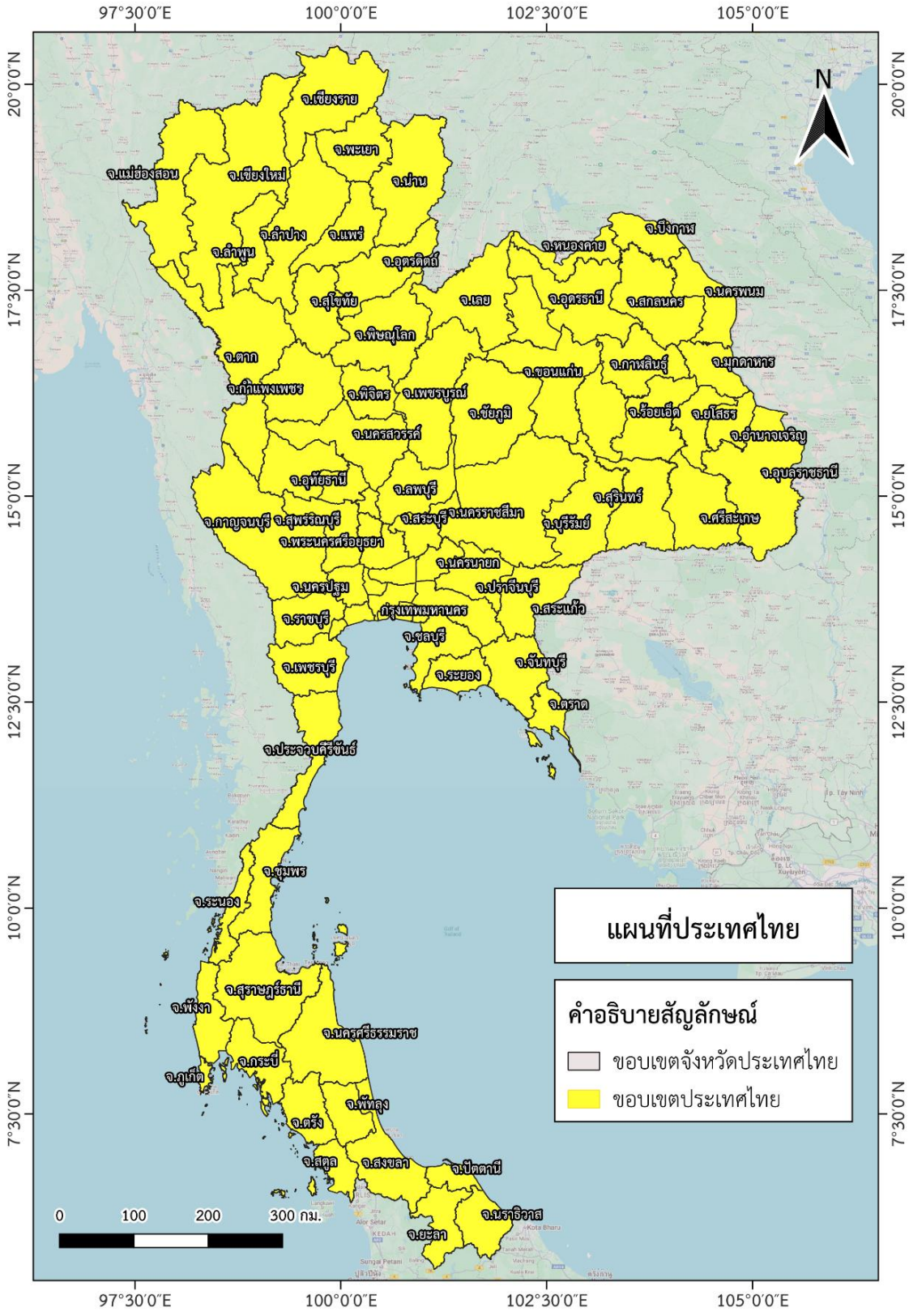
2) โมเดลการทำนายความเสี่ยงที่มีประสิทธิภาพที่มีความแม่นยำ ซึ่งสามารถนำไปประยุกต์ใช้ในสถานการณ์จริงเพื่อลดอุบัติเหตุและเพิ่มความปลอดภัยบนท้องถนนในประเทศไทย

ลิขสิทธิ์ มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved

**ขอบเขตของการศึกษา**

**ขอบเขตพื้นที่**

พื้นที่ศึกษาครอบคลุมประเทศไทย มีพื้นที่ 513,115 ตารางกิโลเมตร ตั้งอยู่ใจกลางคาบสมุทรอินโดจีนในเอเชียตะวันออกเฉียงใต้ ตำแหน่งที่ตั้งของประเทศได้มีอิทธิพลต่อสังคมและวัฒนธรรมไทย ประเทศไทยยังเป็นเส้นทางทางบกเพียงทางเดียวจากทวีปเอเชียไปยังประเทศมาเลเซียและประเทศสิงคโปร์ ประเทศไทยตั้งอยู่ในช่วงพิกัด 5°37′N - 20°27′N และ 97°22′E - 105°37′E




**ภาพ 1**. แผนที่แสดงพื้นที่ศึกษา

**ขอบเขตด้านเนื้อหาและข้อมูล**

1) การนำข้อมูลสถิติอุบัติเหตุที่เกิดขึ้นในประเทศไทย ระหว่างปี พ.ศ. 2562 ถึง พ.ศ. 2567 รวมถึงข้อมูลที่เกี่ยวข้องกับปัจจัยที่ส่งผลต่อความรุนแรงของอุบัติเหตุ เช่น ประเภทถนน สภาพอากาศ เวลาในการเกิดเหตุ ประเภทของยานพาหนะ และปัจจัยอื่นๆ ที่อาจส่งผลต่ออุบัติเหตุ วิเคราะห์ข้อมูลปัจจัยที่อาจส่งผลต่อความรุนแรงของอุบัติเหตุ เพื่อหาความสัมพันธ์และระดับผลกระทบของปัจจัยต่างๆ ต่อความรุนแรงของอุบัติเหตุ โดยใช้เทคนิคการวิเคราะห์ข้อมูลและการสร้างโมเดลเพื่อประเมินความเสี่ยงในการเกิดอุบัติเหตุ
2) การพัฒนาโมเดลทำนายความเสี่ยงของอุบัติเหตุใช้เทคนิคการเรียนรู้ของเครื่อง (Machine Learning) โดยเฉพาะอัลกอริทึม Random Forest ในการสร้างแบบจำลองที่สามารถทำนายความรุนแรงของอุบัติเหตุได้จากข้อมูลปัจจัยที่รวบรวมมาเพื่อประเมินความแม่นยำและประสิทธิภาพในการทำนาย
3) การพัฒนาและทดสอบเว็บแอปพลิเคชัน พัฒนาเว็บแอปพลิเคชันที่สามารถนำเสนอผลลัพธ์การทำนายความเสี่ยงของอุบัติเหตุบนท้องถนน โดยมีการแสดงผลข้อมูลในรูปแบบที่เข้าใจง่าย เพื่อให้ผู้ใช้งานสามารถนำข้อมูลไปใช้ในการวางแผนและป้องกันอุบัติเหตุบนถนน

**เอกสารงานวิจัยที่เกี่ยวข้อง**

Shakil Ahmed., Md Akbar Hossain., Sayan Kumar Ray., Md Mafijul Islam Bhuiyan., Saifur Rahman Sabuj. (2023) **“A study on road accident prediction and contributing factors using explainable machine learning models: analysis and performance”** งานวิจัยนี้มีเป้าหมายหลักในการพัฒนาโมเดลการทำนายอุบัติเหตุบนท้องถนน โดยใช้เทคนิคการเรียนรู้ของเครื่องที่สามารถอธิบายได้ (Explainable Machine Learning Models) การวิจัยนี้มุ่งเน้นไปที่การเข้าใจปัจจัยต่างๆ ที่มีผลต่อการเกิดอุบัติเหตุ เพื่อช่วยในการวางแผนและดำเนินการป้องกันอุบัติเหตุในอนาคตอย่างมีประสิทธิภาพ ข้อมูลที่ใช้ในการศึกษาได้มาจากฐานข้อมูลอุบัติเหตุที่มีการบันทึกไว้ ซึ่งครอบคลุมข้อมูลที่หลากหลาย เช่น ประเภทของถนน สภาพอากาศ ช่วงเวลาของวัน ลักษณะการเกิดอุบัติเหตุ และจำนวนยานพาหนะที่เกี่ยวข้อง โมเดลการทำนายที่พัฒนาในงานวิจัยนี้ได้ใช้เทคนิคการเรียนรู้ของเครื่องหลายแบบ รวมถึงการเรียนรู้เชิงลึก (Deep Learning) และการ

ลิขสิทธิ์ มหาวิทยาลัยนเรศวร
Copyright© by Naresuan University
All rights reserved

ตัดสินใจต้นไม้ (Decision Trees) เพื่อตรวจสอบว่าปัจจัยใดบ้างที่มีผลสำคัญต่อการเกิดอุบัติเหตุ ผลการวิเคราะห์พบว่า โมเดลที่ใช้สามารถทำนายการเกิดอุบัติเหตุได้ด้วยความแม่นยำที่สูง โดยเฉพาะในสภาวะที่มีปัจจัยเสี่ยงสูง เช่น ถนนลื่นในสภาพอากาศฝนตกหรือการขับขี่ในเวลากลางคืน นอกจากนี้ เทคนิคการเรียนรู้ของเครื่องที่สามารถอธิบายได้ยังช่วยให้เห็นภาพชัดเจนเกี่ยวกับปัจจัยที่มีผลต่อการเกิดอุบัติเหตุ เช่น ประเภทของถนนที่มีความเสี่ยงสูง หรือการขับขี่ที่ไม่ปลอดภัยในช่วงเวลาที่มีการจราจรหนาแน่น

Jonathan J. Rolison., Shirley Regev., Salissou Moutari., Aidan Feeney. (2018) **“What are the factors that contribute to road accidents? An assessment of law enforcement views, ordinary drivers’ opinions, and road accident records”** งานวิจัยนี้วิเคราะห์ปัจจัยที่ก่อให้เกิดอุบัติเหตุบนท้องถนน โดยพิจารณามุมมองจากสามกลุ่มหลัก ได้แก่ ผู้บังคับใช้กฎหมาย, ผู้ขับขี่ทั่วไป, และข้อมูลจากบันทึกอุบัติเหตุ การวิจัยนี้มุ่งเน้นไปที่การเปรียบเทียบความคิดเห็นและข้อมูลเชิงสถิติ เพื่อหาความสัมพันธ์ระหว่างปัจจัยที่มีผลต่อการเกิดอุบัติเหตุในมุมมองต่างๆ ซึ่งครอบคลุมปัจจัยเช่น ความเร็วในการขับขี่, การดื่มแอลกอฮอล์, สภาพถนน, และสภาพอากาศข้อมูลจากการสัมภาษณ์ผู้บังคับใช้กฎหมายและผู้ขับขี่ทั่วไปถูกนำมาเปรียบเทียบกับข้อมูลบันทึกอุบัติเหตุที่เกิดขึ้นจริง โดยใช้การวิเคราะห์ทางสถิติ เพื่อประเมินความแตกต่างระหว่างมุมมองของผู้คนและข้อมูลที่บันทึกไว้ การศึกษานี้ยังได้ประเมินการรับรู้ของผู้ขับขี่เกี่ยวกับปัจจัยเสี่ยงและความสำคัญของมาตรการความปลอดภัยที่ผู้บังคับใช้กฎหมายมองว่าเป็นสิ่งจำเป็น ผลการวิเคราะห์ชี้ให้เห็นว่า แม้ว่าผู้บังคับใช้กฎหมายและบันทึกอุบัติเหตุจะชี้ให้เห็นถึงความสำคัญของความเร็วและการดื่มแอลกอฮอล์เป็นปัจจัยหลักในการเกิดอุบัติเหตุ แต่ผู้ขับขี่ทั่วไปมักจะมองว่าเหตุผลที่เกิดอุบัติเหตุเป็นเพราะสภาพถนนและสภาพอากาศมากกว่า งานวิจัยยังพบว่า มีความแตกต่างในวิธีการประเมินความเสี่ยงของการขับขี่ในแต่ละกลุ่ม ซึ่งอาจนำไปสู่ความเข้าใจที่ไม่ตรงกันระหว่างผู้ขับขี่และผู้บังคับใช้กฎหมายการปรับปรุงความเข้าใจร่วมกันระหว่างผู้ขับขี่และผู้บังคับใช้กฎหมายเกี่ยวกับปัจจัยที่ทำให้เกิดอุบัติเหตุ จะช่วยลดอุบัติเหตุได้อย่างมีประสิทธิภาพ นอกจากนี้ ควรมีการเสริมสร้างมาตรการความปลอดภัยที่ตรงกับความเป็นจริงที่ผู้ขับขี่และผู้บังคับใช้กฎหมายเห็นพ้องต้องกัน เพื่อลดความเสี่ยงในการเกิดอุบัติเหตุบนท้องถนน

Tian-Jing Sun., Si-Jia Liu ., Fang-Ke Xie ., Xiao-Fei Huang ., Jian-Xiu Tao ., Yuan-Lan Lu ., Tian-Xi Zhang ., An-Yong Yu. (2024) **“Influence of road types on road traffic accidents in northern Guizhou Province, China”** งานวิจัยนี้วิเคราะห์อิทธิพลของประเภท

ลิขสิทธิ์ มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved

ถนนต่อการเกิดอุบัติเหตุในมณฑลกุ้ยโจวตอนเหนือ ประเทศจีน โดยมุ่งเน้นไปที่การเปรียบเทียบความเสี่ยงของการเกิดอุบัติเหตุในถนนประเภทต่างๆ เช่น ทางหลวง ถนนในเมือง และถนนชนบท งานวิจัยนี้ตั้งเป้าที่จะทำความเข้าใจว่าประเภทถนนแต่ละประเภทมีความเสี่ยงในการเกิดอุบัติเหตุแตกต่างกันอย่างไร และปัจจัยใดบ้างที่ส่งผลกระทบต่อความปลอดภัยในการเดินทาง ข้อมูลที่ใช้ในการวิจัยมาจากบันทึกอุบัติเหตุที่เกิดขึ้นในมณฑลกุ้ยโจวตอนเหนือในช่วงเวลาหนึ่ง โดยมีการจำแนกประเภทของถนนและลักษณะการเกิดอุบัติเหตุ จากนั้นทำการวิเคราะห์เชิงสถิติเพื่อระบุปัจจัยที่ส่งผลต่อการเกิดอุบัติเหตุในแต่ละประเภทถนน นอกจากนี้ยังได้มีการศึกษาเพิ่มเติมเกี่ยวกับลักษณะของถนน เช่น ความกว้าง สภาพพื้นผิวถนน และการควบคุมความเร็ว

Yunzhi Shi's., Raj Biswas., Mehdi Noori., Michael Kilberry. (2021) **"Predicting Road Accident Risk Using Geospatial Data and Machine Learning (Demo Paper)"** งานวิจัยนี้มุ่งเน้นไปที่การพัฒนาโมเดลการทำนายความเสี่ยงของการเกิดอุบัติเหตุทางถนน โดยใช้ข้อมูลเชิงพื้นที่ (Geospatial Data) ร่วมกับเทคนิคการเรียนรู้ของเครื่อง (Machine Learning) เป้าหมายหลักของงานวิจัยนี้คือการวิเคราะห์และทำความเข้าใจความสัมพันธ์ระหว่างข้อมูลเชิงพื้นที่ เช่น ตำแหน่งที่ตั้ง ถนน และปัจจัยสิ่งแวดล้อม กับความเสี่ยงของการเกิดอุบัติเหตุ เพื่อสร้างโมเดลที่สามารถทำนายความเสี่ยงได้อย่างแม่นยำ ข้อมูลที่ใช้ในการศึกษามาจากแหล่งข้อมูลเชิงพื้นที่ที่ครอบคลุมพื้นที่การศึกษา เช่น ข้อมูลแผนที่ถนน สภาพภูมิประเทศ และข้อมูลประชากร รวมถึงบันทึกอุบัติเหตุที่เกิดขึ้นในพื้นที่ดังกล่าว งานวิจัยนี้ใช้เทคนิคการเรียนรู้ของเครื่องหลายรูปแบบในการวิเคราะห์ข้อมูล เช่น Random Forest, Gradient Boosting, และ Neural Networks เพื่อทำนายความเสี่ยงของการเกิดอุบัติเหตุในพื้นที่ต่างๆ การใช้ข้อมูลเชิงพื้นที่ร่วมกับเทคนิคการเรียนรู้ของเครื่องช่วยเพิ่มความแม่นยำในการทำนายความเสี่ยงของอุบัติเหตุ โมเดลที่พัฒนาขึ้นสามารถระบุพื้นที่ที่มีความเสี่ยงสูงได้อย่างแม่นยำ โดยเฉพาะในพื้นที่ที่มีความซับซ้อนทางภูมิประเทศ

Behram Wali., Asad J. Khattak., Thomas Karnowski. (2020) **"The relationship between driving volatility in time to collision and crash-injury severity in a naturalistic driving environment"** งานวิจัยนี้มุ่งเน้นศึกษาความสัมพันธ์ระหว่างความไม่แน่นอนในการขับขี่ในช่วงเวลาที่เกิดการชนและความรุนแรงของการบาดเจ็บที่เกิดจากการชนในสภาพแวดล้อมการขับขี่ตามธรรมชาติ โดยมุ่งวิเคราะห์ปัจจัยที่เพิ่มความไม่แน่นอนในการขับขี่ เช่น การเปลี่ยนแปลงของสภาพอากาศ การจราจรที่หนาแน่น หรือการขับขี่ในพื้นที่ที่ไม่คุ้นเคย และวิธีที่ปัจจัยเหล่านี้ส่งผลต่อระดับความรุนแรงของการบาดเจ็บเมื่อเกิดอุบัติเหตุ จากการรวบรวมข้อมูลการ

ลิขสิทธิ์มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved

ชนในสภาพแวดล้อมการขับขี่ตามธรรมชาติ โดยครอบคลุมปัจจัยต่างๆ เช่น เวลาที่เกิดเหตุ ลักษณะของถนน สภาพอากาศ และลักษณะการขับขี่ของผู้ขับขี่ ข้อมูลนี้ถูกนำมาวิเคราะห์เพื่อหาความสัมพันธ์ระหว่างระดับความไม่แน่นอนในการขับขี่กับความรุนแรงของการบาดเจ็บ โดยใช้เทคนิคการวิเคราะห์ทางสถิติ เช่น การวิเคราะห์การถดถอย (Regression Analysis) และการวิเคราะห์ความเสี่ยง (Risk Analysis) พบว่า ความไม่แน่นอนในการขับขี่มีความสัมพันธ์อย่างมีนัยสำคัญกับความรุนแรงของการบาดเจ็บที่เกิดจากการชน โดยเฉพาะอย่างยิ่งในกรณีที่มีการเปลี่ยนแปลงของสภาพแวดล้อมในการขับขี่ เช่น การขับขี่ในช่วงเวลาที่มีฝนตกหนัก หรือในพื้นที่ที่มีสภาพถนนไม่ดี การวิจัยยังชี้ให้เห็นว่า ผู้ขับขี่ที่ต้องเผชิญกับความไม่แน่นอนในช่วงเวลาที่เกิดการชนมีแนวโน้มที่จะได้รับบาดเจ็บที่รุนแรงมากขึ้นเมื่อเทียบกับผู้ขับขี่ในสภาวะที่มีความแน่นอนสูงกว่า

**นิยามศัพท์เฉพาะ**

**อุบัติเหตุ (Accident)** หมายถึง เหตุการณ์ที่เกิดขึ้นโดยไม่ตั้งใจ ซึ่งนำไปสู่ผลกระทบที่ไม่พึงประสงค์ เช่น การบาดเจ็บ เสียชีวิต หรือความเสียหายต่อทรัพย์สินหรือสิ่งแวดล้อม อุบัติเหตุอาจเกิดขึ้นในหลายบริบท ทั้งในชีวิตประจำวันและในสภาพแวดล้อมที่แตกต่างกัน

**การทำนาย (Prediction)** หมายถึง กระบวนการในการใช้ข้อมูลที่มีอยู่เพื่อตัดสินใจหรือคาดการณ์ผลลัพธ์ในอนาคต โดยทั่วไปจะใช้เทคนิคการวิเคราะห์ข้อมูลและการเรียนรู้ของเครื่องในการสร้างแบบจำลองที่สามารถทำนายได้

**การเรียนรู้ของเครื่อง (Machine Learning)** หมายถึง สาขาหนึ่งของปัญญาประดิษฐ์ (AI) ที่ใช้เทคนิคทางคณิตศาสตร์และสถิติในการพัฒนาอัลกอริทึมที่สามารถเรียนรู้จากข้อมูลและปรับปรุงประสิทธิภาพในการทำนายโดยอัตโนมัติ

ลิขสิทธิ์มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved

**ประเภทของการเรียนรู้ของเครื่อง:**

- การเรียนรู้แบบมีผู้สอน (Supervised Learning): ใช้ข้อมูลที่มีการกำกับ (labelled data) เพื่อฝึกอบรมโมเดล เช่น การใช้ข้อมูลอุบัติเหตุที่มีการระบุความรุนแรงในการฝึกอบรมโมเดล เพื่อทำนายความรุนแรงในข้อมูลใหม่

- การเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning): ใช้ข้อมูลที่ไม่มีการกำกับ (unlabeled data) เพื่อค้นหารูปแบบหรือโครงสร้าง เช่น การแบ่งกลุ่มอุบัติเหตุตามลักษณะการเกิดเหตุ
- การเรียนรู้แบบกึ่งมีผู้สอน (Semi-supervised Learning): ใช้ข้อมูลทั้งที่มีการกำกับและไม่มีการกำกับ โดยมักจะใช้ข้อมูลที่มีการกำกับในปริมาณน้อยเพื่อเสริมข้อมูลที่ไม่มีการกำกับ
- การเรียนรู้แบบเสริมกำลัง (Reinforcement Learning): การเรียนรู้ที่ใช้การเรียนรู้จากการปฏิสัมพันธ์ (Interaction) ระหว่างผู้เรียน (Agent) กับสิ่งแวดล้อม (Environment) โดยการกำหนดเงื่อนไขบางอย่างให้กับคอมพิวเตอร์ และทำให้คอมพิวเตอร์ทำตามเงื่อนไขนั้นให้ได้

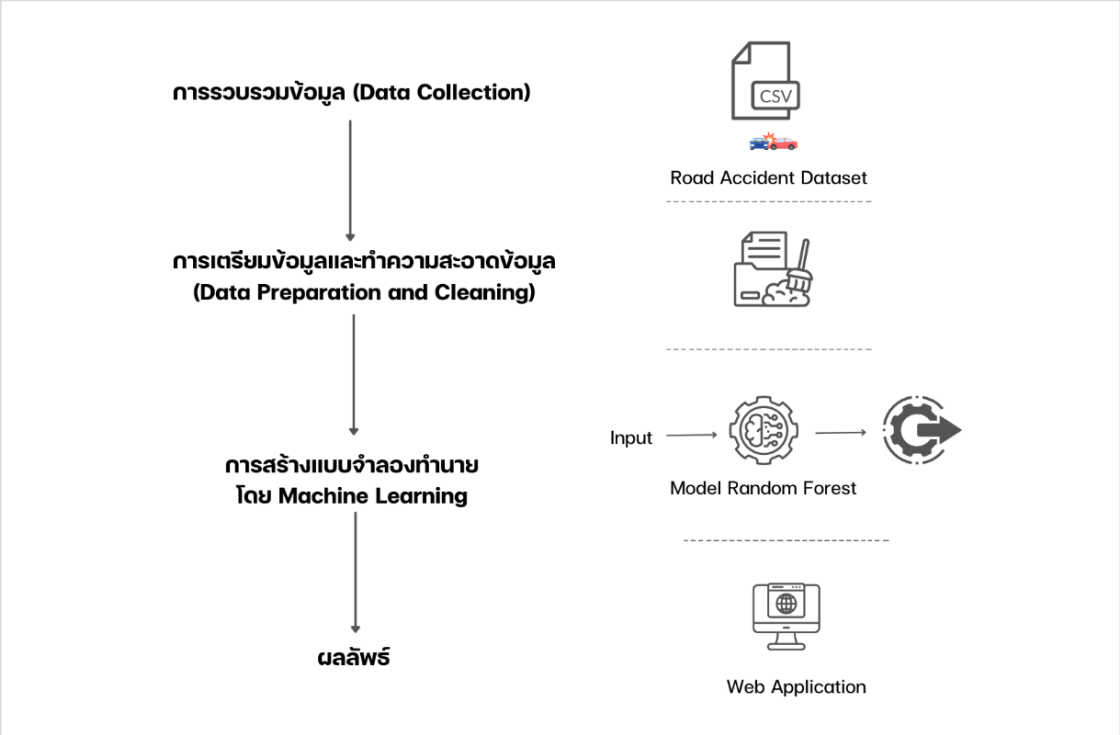
**โมเดล (Model)** หมายถึง แบบจำลองทางคณิตศาสตร์ที่ใช้ในการอธิบายและทำนายพฤติกรรมหรือผลลัพธ์จากข้อมูล โมเดลสามารถมีหลายประเภท เช่น โมเดลเชิงเส้น โมเดลต้นไม้การตัดสินใจ (Decision Tree) หรือโมเดล Random Forest

**การสร้างและใช้งานโมเดล**:

- **การสร้างโมเดล**: ใช้ข้อมูลที่เก็บรวบรวมเพื่อฝึกอบรมโมเดล โดยการเลือกอัลกอริทึมที่เหมาะสม เช่น Random Forest, Support Vector Machine (SVM) หรือ Neural Networks
- **การทดสอบและประเมินผลโมเดล**: หลังจากสร้างโมเดลแล้ว จะต้องทดสอบเพื่อประเมินความแม่นยำและประสิทธิภาพในการทำนาย โดยใช้ข้อมูลทดสอบ (test data) ที่แยกออกจากชุดข้อมูลที่ใช้ในการฝึกอบรม
- **การปรับปรุงโมเดล**: หากโมเดลไม่สามารถทำนายได้ดี อาจต้องปรับแต่งพารามิเตอร์หรือเปลี่ยนแปลงอัลกอริทึมที่ใช้

ลิขสิทธิ์ มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved

**กรอบแนวความคิด**




**ภาพ 2.** กรอบแนวความคิด

การศึกษาครั้งนี้ใช้ข้อมูลอุบัติเหตุของประเทศไทยในช่วง 6 ปี (2019–2024) ที่รวบรวมจากระบบวิเคราะห์อุบัติเหตุ ของสำนักงานปลัดกระทรวงคมนาคม ข้อมูลนี้ยังอยู่ในพอร์ทัลข้อมูลเปิดอีกด้วย ข้อมูลที่ได้เกี่ยวกับชุดข้อมูลอุบัติเหตุเป็นชุดข้อมูลทั้งหมด 6 ชุดซึ่งประกอบด้วยปัจจัยที่ก่อให้เกิดอุบัติเหตุ ปี (2019–2024) มีข้อมูลรวมกันทั้งหมดประมาณ 84,627 records ดังนั้นเราได้เลือกข้อมูลอุบัติเหตุ เช่น วันที่เกิดอุบัติเหตุ เดือนที่เกิดอุบัติเหตุ ชั่วโมงที่เกิดอุบัติเหตุ ประเภทของอุบัติเหตุ ตำแหน่งที่เกิดอุบัติเหตุ ปัจจัยด้านสิ่งแวดล้อม ประเภทของยานพาหนะ ปัจจัยของยานพาหนะ และปัจจัยส่วนบุคคล (หรือผู้ใช้) ที่มีอิทธิพลต่อความรุนแรงของอุบัติเหตุ ในงานนี้ เราพิจารณาประเภทของอุบัติเหตุตามความรุนแรง ตามข้อมูลของเรา มีความรุนแรงของอุบัติเหตุสี่ประเภท และคำจำกัดความของอุบัติเหตุมีดังนี้:

ลิขสิทธิ์มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved

- อุบัติเหตุร้ายแรง: อุบัติเหตุทางถนนที่ส่งผลให้มีผู้เสียชีวิต

- อุบัติเหตุบาดเจ็บสาหัส: อุบัติเหตุทางถนนที่ฝ่ายใดฝ่ายหนึ่งต้องได้รับการรักษาพยาบาลและนำส่งโรงพยาบาล

- อุบัติเหตุที่ได้รับบาดเจ็บปานกลาง: อุบัติเหตุทางถนนซึ่งไม่มีใครต้องการความช่วยเหลือทางการแพทย์ แต่ได้รับบาดเจ็บเล็กน้อยและมีบาดแผลเล็กน้อย หรือความเสียหายทางทรัพย์สินไม่มาก

- อุบัติเหตุที่ได้รับบาดเจ็บเล็กน้อย: อุบัติเหตุที่มีความรุนแรงต่ำ หรือไม่มีการบาดเจ็บที่รุนแรง

**สถิติเชิงพรรณนาข้อมูลอุบัติเหตุทางถนน**

ตารางที่ 1 แสดงสถิติเชิงพรรณนาของตัวแปรในแง่ของจำนวนและเปอร์เซ็นต์ของตัวแปรตาม ตัวแปรตามจะถูกแบ่งออกเป็น 4 ประเภทตามความรุนแรงของการบาดเจ็บ ได้แก่ เสียชีวิต ร้ายแรง ปานกลาง และเล็กน้อย จากอุบัติเหตุบนท้องถนน 84,627 records สถิติสำหรับตัวแปรอิสระแสดงอยู่ในตารางที่ 2 ตัวแปรอิสระเหล่านี้ถูกเลือกโดยอิงตามหมวดหมู่ต่างๆ ที่ เช่น ปัจจัยด้านมนุษย์ สภาพแวดล้อม วันในสัปดาห์ และอื่นๆ จำนวนอุบัติเหตุสูงสุดมากกว่า 20,000 ครั้ง เกิดขึ้นในปี 2020 และปี 2021 แต่เมื่อพิจารณาจากจำนวนทั้งหมดแล้ว พบว่าในปี 2022 มีอุบัติเหตุลงลงอย่างเห็นได้ชัด เนื่องจากมีข้อมูลอุบัติเหตุที่น้อยกว่าปีอื่นๆ และโอกาสเกิดอุบัติเหตุ ปี 2023-2024 จำนวนอุบัติเหตุกลับมาเพิ่มขึ้น เนื่องจากการเดินทางและกิจกรรมต่าง ๆ ฟื้นตัว นอกจากนี้ ข้อมูลยังระบุว่าเดือนเมษายนมีอัตราการเกิดอุบัติเหตุสูงสุด 13.4% จำนวนอุบัติเหตุเพิ่มสูงในช่วงสงกรานต์ เนื่องจากมีการเดินทางกลับภูมิลำเนาและการเฉลิมฉลอง ทำให้เกิดการขับขี่ในสภาพมึนเมารองลงมาจะเป็นเดือนมกราคม 12.2% และเดือนธันวาคม 10.2% ช่วงเทศกาลปีใหม่มีการเดินทางท่องเที่ยวและกลับบ้าน ทำให้การขับขี่ระยะไกลเพิ่มขึ้น ส่งผลให้จำนวนอุบัติเหตุเพิ่มขึ้นเช่นกัน และข้อมูลระบุว่าวันศุกร์และวันเสาร์เป็นวันที่เกิดอุบัติเหตุร้ายแรงที่สุดในสัปดาห์ มีจำนวนอุบัติเหตุสูงอย่างต่อเนื่องทุกปี เนื่องจากเป็นวันสิ้นสัปดาห์ที่มีการเดินทางและเฉลิมฉลองมาก ในแต่ละปีตั้งแต่ 2019 ถึง 2024 อุบัติเหตุตั้งแต่เวลา 07.00 – 09.59 น. อุบัติเหตุเริ่มเพิ่มขึ้นเนื่องจากช่วงเวลาที่คนเดินทางไปทำงานหรือโรงเรียน และเวลา 15.00 – 18.59 น. เป็นช่วงที่มีจำนวนอุบัติเหตุสูงสุด ซึ่งสอดคล้องกับเวลาเดินทางกลับบ้านและการจราจรหนาแน่น ช่วงเวลาที่เกิดอุบัติเหตุน้อยที่สุดจะเป็นเวลาตั้งแต่ 00.00 – 05.59 น.

ลิขสิทธิ์มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved

ช่วงกลางคืนถึงเช้ามืดมีจำนวนอุบัติเหตุต่ำเนื่องจากการจราจรเบาบาง ประเภทการชนที่พบบ่อยที่สุด คือ พลิกคว่ำ/ตกถนนในทางตรงเนื่องจากการขับขี่ด้วยความเร็วสูงและการควบคุมรถที่ผิดพลาด ชนท้ายเกิดจากการเว้นระยะห่างไม่เพียงพอและการเบรกกะทันหัน และพลิกคว่ำในทางโค้งมักเกิดเมื่อผู้ขับขี่ใช้ความเร็วสูงในบริเวณที่ต้องเลี้ยว สุดท้ายจากข้อมูลที่ได้ศึกษา สถิติอุบัติเหตุจำแนกตามประเภทยานพาหนะ รถปิคอัพและรถยนต์ส่วนบุคคล เกิดอุบัติเหตุบ่อยที่สุด (รถปิคอัพ > 40,000 ครั้ง, รถยนต์ส่วนบุคคล 30,000 ครั้ง) เนื่องจากการใช้งานแพร่หลายและเป็นยานพาหนะหลักของคนไทย รองลงมาจะเป็นรถจักรยานยนต์แม้จำนวนอุบัติเหตุจะน้อยกว่ารถยนต์ แต่ความรุนแรงของบาดเจ็บสูง

ตาราง 1 สถิติเชิงพรรณนาของตัวแปรตาม

| **ตัวแปรตาม** | | |
|---|---|---|
| **ความรุนแรง** | **ตัวแปร** | **เปอร์เซ็นต์ %** |
| | ถึงแก่ชีวิต | 8.3 |
| | บาดเจ็บสาหัส | 5.9 |
| | บาดเจ็บปานกลาง | 23.8 |
| | บาดเจ็บเล็กน้อย | 62.0 |

ลิขสิทธิ์ มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved

ตาราง 2 สถิติเชิงพรรณนาของตัวแปรต้น

| **ตัวแปรต้น** |
|---|
| Model 1 |
| **วัน** อาทิตย์ จันทร์ อังคาร พุธ พฤหัสบดี ศุกร์ เสาร์ |
| **เดือน** มกราคม กุมภาพันธ์ มีนาคม เมษายน พฤษภาคม มิถุนายน<br>กรกฎาคม สิงหาคม กันยายน ตุลาคม พฤศจิกายน ธันวาคม |
| **ชั่วโมง** |
| • 00.00 - 00.59 น. • 01.00 - 01.59 น. • 02.00 - 02.59 น.• 03.00 - 03.59 น. |
| • 04.00 - 04.59 น. • 05.00 - 05.59 น. • 06.00 - 06.59 น. • 07.00 - 07.59 น. |
| • 08.00 - 08.59 น. • 09.00 - 09.59 น. • 10.00 - 10.59 น. • 11.00 - 11.59 น. |
| • 12.00 - 12.59 น. • 13.00 - 13.59 น. • 14.00 - 14.59 น.• 15.00 - 15.59 น. |
| • 16.00 - 16.59 น. • 17.00 - 17.59 น. • 18.00 - 18.59 น. • 19.00 - 19.59 น. |
| • 20.00 - 20.59 น. • 21.00 - 21.59 น. • 22.00 - 22.59 น. • 23.00 - 23.59 น. |
| **พิกัด** |
| LATITUDE LONGITUDE |

ลิขสิทธิ์ มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved

**ตัวแปรต้น**

Model 2

**วัน** อาทิตย์ จันทร์ อังคาร พุธ พฤหัสบดี ศุกร์ เสาร์

**เดือน** มกราคม กุมภาพันธ์ มีนาคม เมษายน พฤษภาคม มิถุนายน

กรกฎาคม สิงหาคม กันยายน ตุลาคม พฤศจิกายน ธันวาคม

**ชั่วโมง**

• 00.00 - 00.59 น. • 01.00 - 01.59 น. • 02.00 - 02.59 น.• 03.00 - 03.59 น.

• 04.00 - 04.59 น. • 05.00 - 05.59 น. • 06.00 - 06.59 น. • 07.00 - 07.59 น.

• 08.00 - 08.59 น. • 09.00 - 09.59 น. • 10.00 - 10.59 น. • 11.00 - 11.59 น.

• 12.00 - 12.59 น. • 13.00 - 13.59 น. • 14.00 - 14.59 น.• 15.00 - 15.59 น.

• 16.00 - 16.59 น. • 17.00 - 17.59 น. • 18.00 - 18.59 น. • 19.00 - 19.59 น.

• 20.00 - 20.59 น. • 21.00 - 21.59 น. • 22.00 - 22.59 น. • 23.00 - 23.59 น.

**จังหวัด**

เชียงราย น่าน พะเยา เชียงใหม่ แม่ฮ่องสอน แพร่ ลำปาง ลำพูน อุตรดิตถ์ กรุงเทพมหานคร พิษณุโลก สุโขทัย เพชรบูรณ์ พิจิตร กำแพงเพชร นครสวรรค์ ลพบุรี ชัยนาท อุทัยธานี สิงห์บุรี อ่างทอง สระบุรี พระนครศรีอยุธยา สุพรรณบุรี นครนายก ปทุมธานี นนทบุรี นครปฐม สมุทรปราการ สมุทรสาคร สมุทรสงคราม หนองคาย นครพนม สกลนคร อุดรธานี หนองบัวลำภู เลย มุกดาหาร กาฬสินธุ์ ขอนแก่น อำนาจเจริญ ยโสธร ร้อยเอ็ด มหาสารคาม ชัยภูมิ นครราชสีมา บุรีรัมย์ สุรินทร์ ศรีสะเกษ อุบลราชธานี สระแก้ว ปราจีนบุรี ฉะเชิงเทรา ชลบุรี ระยอง จันทบุรี ตราด ตาก กาญจนบุรี ราชบุรี เพชรบุรี ประจวบคีรีขันธ์ ชุมพร ระนอง สุราษฎร์ธานี นครศรีธรรมราช กระบี่ พังงา ภูเก็ต พัทลุง ตรัง ปัตตานี สงขลา สตูล นราธิวาส และยะลา

ลิขสิทธิ์ มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved

**ตัวแปรต้น**

Model 3

**วัน** อาทิตย์ จันทร์ อังคาร พุธ พฤหัสบดี ศุกร์ เสาร์

**เดือน** มกราคม กุมภาพันธ์ มีนาคม เมษายน พฤษภาคม มิถุนายน
กรกฎาคม สิงหาคม กันยายน ตุลาคม พฤศจิกายน ธันวาคม

**ชั่วโมง**

• 00.00 - 00.59 น. • 01.00 - 01.59 น. • 02.00 - 02.59 น.• 03.00 - 03.59 น.

• 04.00 - 04.59 น. • 05.00 - 05.59 น. • 06.00 - 06.59 น. • 07.00 - 07.59 น.

• 08.00 - 08.59 น. • 09.00 - 09.59 น. • 10.00 - 10.59 น. • 11.00 - 11.59 น.

• 12.00 - 12.59 น. • 13.00 - 13.59 น. • 14.00 - 14.59 น.• 15.00 - 15.59 น.

• 16.00 - 16.59 น. • 17.00 - 17.59 น. • 18.00 - 18.59 น. • 19.00 - 19.59 น.

• 20.00 - 20.59 น. • 21.00 - 21.59 น. • 22.00 - 22.59 น. • 23.00 - 23.59 น.

**ตัวแปรต้น**

Model 4

**วัน** อาทิตย์ จันทร์ อังคาร พุธ พฤหัสบดี ศุกร์ เสาร์

**เดือน** มกราคม กุมภาพันธ์ มีนาคม เมษายน พฤษภาคม มิถุนายน
กรกฎาคม สิงหาคม กันยายน ตุลาคม พฤศจิกายน ธันวาคม

**ชั่วโมง**

• 00.00 - 00.59 น. • 01.00 - 01.59 น. • 02.00 - 02.59 น.• 03.00 - 03.59 น.

• 04.00 - 04.59 น. • 05.00 - 05.59 น. • 06.00 - 06.59 น. • 07.00 - 07.59 น.

ลิขสิทธิ์มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved

• 08.00 - 08.59 น. • 09.00 - 09.59 น. • 10.00 - 10.59 น. • 11.00 - 11.59 น.

• 12.00 - 12.59 น. • 13.00 - 13.59 น. • 14.00 - 14.59 น.• 15.00 - 15.59 น.

• 16.00 - 16.59 น. • 17.00 - 17.59 น. • 18.00 - 18.59 น. • 19.00 - 19.59 น.

• 20.00 - 20.59 น. • 21.00 - 21.59 น. • 22.00 - 22.59 น. • 23.00 - 23.59 น.

**ประเภทรถ**

รถจักรยานยนต์,รถสามล้อเครื่อง,รถยนต์นั่งส่วนบุคคล,รถตู้,รถปิคอัพโดยสาร

รถโดยสารมากกว่า4ล้อ,รถปิคอัพบรรทุก4ล้อ,รถบรรทุก6ล้อ,รถบรรทุกไม่เกิน10ล้อ

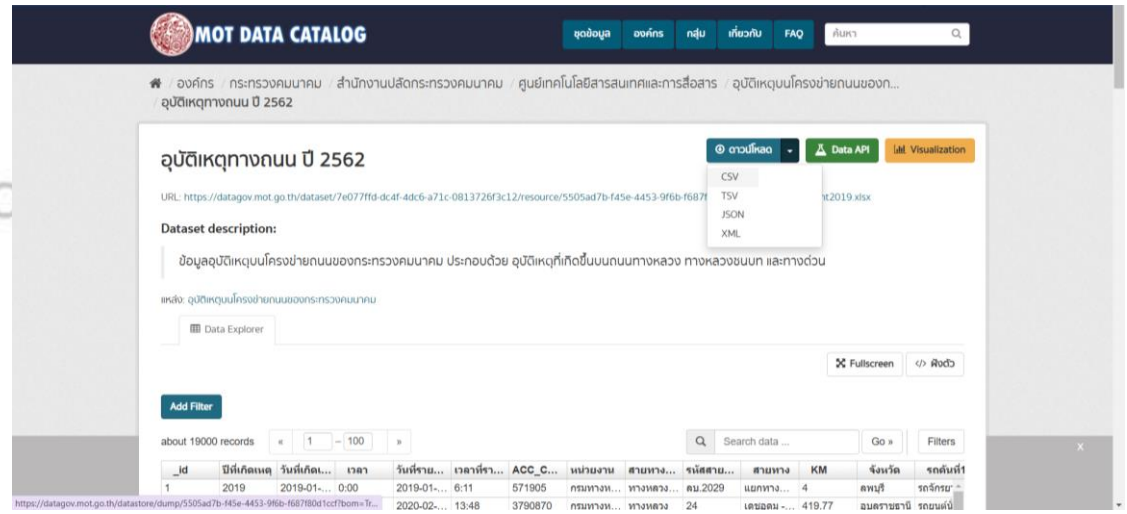
รถบรรทุกมากกว่า10ล้อ,รถอีแต๋น อื่นๆ

---

**เครื่องมือที่ใช้ในการวิจัย**

เครื่องมือที่ใช้ในการวิจัย มีดังต่อไปนี้

1. ข้อมูลสำนักงานปลัดกระทรวงคมนาคม ข้อมูลอุบัติเหตุ 2019-2024 ได้แก่ วันที่เกิดอุบัติเหตุ ปีที่เกิดอุบัติเหตุ เวลาที่เกิดอุบัติเหตุ สายทาง จังหวัด พิกัดทางภูมิศาสตร์ ประเภทการชน ประเภทรถอื่นๆ เว็บไซต์ : https://datagov.mot.go.th/dataset/roadaccident




**ภาพ 3.** ข้อมูลอุบัติเหตุจากสำนักงานปลัดกระทรวงคมนาคม

2. โปรแกรม Jupyter

เป็นแพลตฟอร์มที่ใช้ในการเขียนและรันโค้ด Pythonเป็นภาษาการเขียนโปรแกรมที่ใช้อย่างแพร่หลายในเว็บแอปพลิเคชัน การพัฒนาซอฟต์แวร์ วิทยาศาสตร์ข้อมูล และแมชชีนเลิร์นนิง (ML) นักพัฒนาใช้ Python เนื่องจากมีประสิทธิภาพ เรียนรู้ง่าย และสามารถทำงานบนแพลตฟอร์มต่างๆ ได้มากมาย การพัฒนาเว็บฝั่งเซิร์ฟเวอร์ประกอบด้วยฟังก์ชันแบ็คเอนด์ที่ซับซ้อนซึ่งเว็บไซต์ดำเนินการเพื่อแสดงข้อมูลต่อผู้ใช้ ตัวอย่างเช่น เว็บไซต์ต้องโต้ตอบกับฐานข้อมูล สื่อสารกับเว็บไซต์อื่น และปกป้องข้อมูลเมื่อส่งข้อมูลผ่านเครือข่าย

Python มีคุณสมบัติต่อไปนี้ทำให้ภาษาการเขียนโปรแกรม Python มีเอกลักษณ์ที่ไม่เหมือนใคร ภาษาที่แปลผลแล้วPython เป็นภาษาที่แปลผลแล้ว ซึ่งหมายความว่าสามารถเรียกใช้โค้ดทีละบรรทัดได้โดยตรง หากมีข้อผิดพลาดในโค้ดโปรแกรม ก็จะหยุดทำงานทันที ดังนั้นโปรแกรมเมอร์จึงสามารถค้นหาข้อผิดพลาดในโค้ดได้อย่างรวดเร็ว ภาษาที่ใช้งานง่าย Python ใช้คำที่เหมือนในภาษาอังกฤษ ซึ่งแตกต่างจากภาษาการเขียนโปรแกรมอื่นๆ เนื่องจาก Python ไม่ใช้วงเล็บปีกกา แต่จะใช้การเยื้องแทน ภาษาที่ระบุประเภทแบบไดนามิก โปรแกรมเมอร์ไม่ต้องระบุประเภทตัวแปรเมื่อเขียนโค้ดเนื่องจาก Python จะกำหนดไว้ที่รันไทม์ ด้วยเหตุนี้ คุณจึงสามารถเขียนโปรแกรม Python ได้รวดเร็วขึ้น ภาษาระดับสูงPython มีความใกล้เคียงกับภาษามนุษย์มากกว่าภาษาการเขียนโปรแกรมอื่นๆ ดังนั้นโปรแกรมเมอร์จึงไม่ต้องกังวลกับฟังก์ชันการทำงานพื้นฐานต่างๆ เช่น สถาปัตยกรรมและการจัดการหน่วยความจำ ภาษาเชิงอ็อบเจกต์ Python ถือว่าทุกสิ่งเป็นอ็อบเจกต์ แต่ก็ยังรองรับการเขียนโปรแกรมประเภทอื่นๆ ด้วย เช่น การเขียนโปรแกรมเชิงโครงสร้างและเชิงฟังก์ชัน

**ไลบรารี Python**

ไลบรารีคือชุดของโค้ดที่ใช้บ่อยซึ่งนักพัฒนาสามารถใช้ในโปรแกรม Python เพื่อหลีกเลี่ยงการเขียนโค้ดขึ้นใหม่ทั้งหมด ตามค่าเริ่มต้นแล้ว Python จะมาพร้อมกับไลบรารีมาตรฐาน ซึ่งมีฟังก์ชันที่นำกลับมาใช้ใหม่ได้มากมาย นอกจากนี้ยังมีไลบรารี Python มากกว่า 137,000 รายการสำหรับการประยุกต์ใช้ต่างๆ รวมถึงการพัฒนาเว็บ วิทยาศาสตร์ข้อมูล และแมชชีนเลิร์นนิง (ML)

**ไลบรารี Python ที่ใช้ในงานMachine Learning**

**Matplotlib** นักพัฒนาซอฟต์แวร์ใช้ Matplotlib เพื่อลงจุดข้อมูลในกราฟิกสองมิติและสามมิติ (2D และ 3D) คุณภาพสูง ซึ่งมักจะใช้ในงานทางวิทยาศาสตร์ Matplotlib ช่วยให้คุณสามารถแสดงข้อมูล

ลิขสิทธิ์ มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved

เป็นภาพโดยแสดงผลในแผนภูมิต่างๆ เช่น แผนภูมิแท่งและแผนภูมิเส้น คุณยังสามารถลงจุดแผนภูมิได้หลายรายการพร้อมกัน และกราฟิกสามารถใช้งานได้ในทุกแพลตฟอร์มอีกด้วย

**Pandas**

Pandas มีโครงสร้างข้อมูลที่ปรับให้เหมาะสมและยืดหยุ่น ซึ่งคุณสามารถใช้เพื่อจัดการชุดข้อมูลเวลาและข้อมูลที่มีโครงสร้าง เช่น ตารางและอาร์เรย์ ตัวอย่างเช่น คุณสามารถใช้ Pandas เพื่ออ่าน เขียน ผสานรวม กรอง และจัดกลุ่มข้อมูลได้ หลายคนจึงใช้สำหรับงานด้านวิทยาศาสตร์ข้อมูล การวิเคราะห์ข้อมูล และ ML

**NumPy**

NumPy เป็นไลบรารียอดนิยมที่นักพัฒนาใช้เพื่อสร้างและจัดการอาร์เรย์ จัดการรูปร่างเชิงตรรกะ และดำเนินการคำนวณพีชคณิตเชิงเส้นได้อย่างง่ายดาย โดย NumPy รองรับการทำงานร่วมกับภาษาต่างๆ มากมาย เช่น C และ C++

**Request**

ไลบรารีคำขอมีฟังก์ชันที่เป็นประโยชน์ที่จำเป็นสำหรับการพัฒนาเว็บ โดยคุณสามารถใช้ไลบรารีนี้เพื่อส่งคำขอ HTTP, เพิ่มส่วนหัว, เพิ่มพารามิเตอร์ URL, เพิ่มข้อมูล และทำงานอื่นๆ อีกมากมายเมื่อสื่อสารกับเว็บแอปพลิเคชัน

**OpenCV-Python**

OpenCV-Python เป็นไลบรารีที่นักพัฒนาใช้ในการประมวลผลรูปภาพสำหรับการประยุกต์ใช้เชิงคอมพิวเตอร์วิทัศน์ ซึ่งมีฟังก์ชันมากมายสำหรับงานประมวลผลภาพ เช่น การอ่านและเขียนรูปภาพพร้อมกัน, การสร้างสภาพแวดล้อม 3 มิติจาก 2 มิติ และการบันทึกภาพและวิเคราะห์ภาพจากวิดีโอ

ลิขสิทธิ์ มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved

**Keras**

Keras เป็นไลบรารีนิวรัลเน็ตเวิร์กเชิงลึกของ Python ที่รองรับการประมวลผลข้อมูล การแสดงภาพข้อมูล และอื่นๆ อีกมากมายได้อย่างยอดเยี่ยม Keras รองรับนิวรัลเน็ตเวิร์กมากมาย ซึ่งมีโครงสร้างแบบโมดูลที่ให้ความยืดหยุ่นในการเขียนแอปพลิเคชันด้านนวัตกรรม

**Scikit-learn**

Scikit-learn เป็น Python module ที่รวบรวม Machine learning algorithms ต่างๆ ทั้งการเรียนรู้แบบมีผู้สอนและไม่มีผู้สอน

cikit-learn นำเสนอชุดเครื่องมือที่ครอบคลุมสำหรับการประมวลผลข้อมูล การเลือกแบบจำลอง การประเมินผล และการปรับแต่งแบบจำลอง ไลบรารีนี้มีคุณสมบัติที่สำคัญหลายอย่าง

- **ความหลากหลายของ Model**: Scikit-learn รองรับแบบจำลองการเรียนรู้ของเครื่องจักรที่หลากหลาย รวมถึงการเรียนรู้ภายใต้การดูแล (supervised learning) เช่น การจำแนกประเภท (classification) และการถดถอย (regression) รวมถึงการเรียนรู้โดยไม่มีการดูแล (unsupervised learning) เช่น การจัดกลุ่ม (clustering) และการลดมิติ (dimensionality reduction)
- **การประมวลผลข้อมูลที่ง่าย**: Scikit-learn มาพร้อมกับเครื่องมือที่หลากหลายสำหรับการทำความสะอาดข้อมูล การแปลง และการปรับปรุงข้อมูลเพื่อให้เหมาะสมกับแบบจำลอง
- **การประเมินและการปรับแต่ง Model**: มีเครื่องมือในการประเมินประสิทธิภาพของแบบจำลองและการปรับแต่งพารามิเตอร์เพื่อให้ได้ผลลัพธ์ที่ดีที่สุด

การเริ่มต้นใช้งาน Scikit-learn นั้นง่ายมาก เพียงติดตั้งผ่าน pip install:

pip install scikit-learn หลังจากติดตั้ง, คุณสามารถเริ่มใช้งานได้ทันที

**วิทยาศาสตร์ข้อมูลและแมชชีนเลิร์นนิง**

วิทยาศาสตร์ข้อมูลคือการศึกษาข้อมูลเพื่อดึงข้อมูลเชิงลึกที่มีความหมายสำหรับธุรกิจ เป็นแนวทางสหสาขาวิชาที่ผสมผสานหลักการและแนวทางปฏิบัติต่างๆ จากสาขาวิชาคณิตศาสตร์ สถิติ ปัญญาประดิษฐ์ และวิศวกรรมคอมพิวเตอร์ เพื่อวิเคราะห์ข้อมูลจำนวนมาก การวิเคราะห์นี้ช่วยให้นักวิทยาศาสตร์ข้อมูลสามารถถามและตอบคำถามต่างๆ ได้ เช่น เกิดอะไรขึ้น ทำไมถึงเกิดขึ้น จะเกิดอะไรขึ้น และทำอะไรกับผลลัพธ์ได้บ้างเราใช้วิทยาศาสตร์ข้อมูลเพื่ออะไรบ้าง

ลิขสิทธิ์มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved

วิทยาศาสตร์ข้อมูลใช้เพื่อศึกษาข้อมูลใน 4 วิธีหลัก ดังนี้

**1. การวิเคราะห์แบบพรรณนา** จะตรวจสอบข้อมูลเพื่อค้นหาข้อมูลเชิงลึกว่าเกิดอะไรขึ้นหรือเกิดอะไรในสภาพแวดล้อมข้อมูล โดยมีลักษณะเฉพาะตัวในการใช้การแสดงข้อมูลเป็นภาพ เช่น

แผนภูมิวงกลม แผนภูมิแท่ง กราฟเส้น ตาราง หรือการบรรยายที่สร้างขึ้น ตัวอย่างเช่น บริการจองเที่ยวบินอาจบันทึกข้อมูล อาทิ จำนวนตั๋วที่จองในแต่ละวัน การวิเคราะห์แบบพรรณนาจะเผยให้เห็นยอดการจองที่เพิ่มขึ้น ยอดการจองที่ต่ำลง และเดือนที่บริการนี้มีผลการดำเนินงานสูง

**2. การวิเคราะห์แบบวินิจฉัย** คือการตรวจสอบข้อมูลแบบเจาะลึกหรือแบบละเอียดเพื่อทำความเข้าใจถึงสาเหตุ โดยมีลักษณะเฉพาะตัวในการใช้เทคนิคต่างๆ เช่น การเจาะลึก การค้นพบข้อมูล การทำเหมืองข้อมูล และการหาความสัมพันธ์ อาจมีการดำเนินการและการเปลี่ยนแปลงข้อมูลหลายรายการในชุดข้อมูลที่กำหนดเพื่อค้นหารูปแบบที่ไม่ซ้ำกันในแต่ละเทคนิค

**3. การวิเคราะห์แบบคาดการณ์** ใช้ข้อมูลในอดีตเพื่อสร้างการคาดการณ์ที่แม่นยำเกี่ยวกับรูปแบบข้อมูลที่อาจเกิดขึ้นในอนาคต โดยมีลักษณะเฉพาะตัวในการใช้เทคนิคต่างๆ เช่น แมชชีนเลิร์นนิง การพยากรณ์ การจับคู่รูปแบบ และการสร้างแบบจำลองเชิงคาดการณ์ ในแต่ละเทคนิคเหล่านี้ มีการฝึกฝนคอมพิวเตอร์ให้ทำวิศวกรรมย้อนกลับการสัมพันธ์ที่เป็นเหตุเป็นผลในข้อมูล ตัวอย่างเช่น ทีม

**4. การวิเคราะห์แบบให้คำแนะนำ** จะต่อยอดจากข้อมูลที่คาดการณ์ ไม่เพียงแต่คาดการณ์สิ่งที่น่าจะเกิดขึ้น แต่ยังแนะนำการตอบสนองที่เหมาะสมที่สุดต่อผลลัพธ์นั้นด้วย โดยสามารถวิเคราะห์ผลกระทบที่อาจเกิดขึ้นจากตัวเลือกต่างๆ และแนะนำแนวทางปฏิบัติที่ดีที่สุด และใช้การวิเคราะห์กราฟ การจำลอง การประมวลผลเหตุการณ์ที่ซับซ้อน นิวรัลเน็ตเวิร์ค และกลไกการแนะนำจากแมชชีนเลิร์นนิง

**แมชชีนเลิร์นนิง (ML)**

จะสอนคอมพิวเตอร์ให้เรียนรู้จากข้อมูลโดยอัตโนมัติและทำนายได้อย่างแม่นยำใช้ Python สำหรับงานด้านวิทยาศาสตร์ข้อมูลต่างๆ ดังต่อไปนี้:

ลิขสิทธิ์ มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved

- การแก้ไขและลบข้อมูลที่ไม่ถูกต้อง ซึ่งเรียกว่าการทำความสะอาดข้อมูล
- การแยกและเลือกคุณสมบัติต่างๆ ของข้อมูล
- การระบุประเภทข้อมูล ซึ่งเป็นการเพิ่มชื่อที่มีความหมายสำหรับข้อมูล
- การค้นหาสถิติต่างๆ จากข้อมูล

- การแสดงข้อมูลด้วยภาพโดยใช้แผนภูมิและกราฟ เช่น แผนภูมิเส้น กราฟแท่ง ฮิสโทแกรม และแผนภูมิวงกลม

**การทำงานของแมชชีนเลิร์นนิง**

เราจะกำหนดชุดข้อมูลสำหรับ **การฝึกโมเดล (Training) และชุดข้อมูลสำหรับ การทดสอบโมเดล (Testing)** โดยใช้ฟังก์ชัน train_test_split()

X_train , X_test, y_train, y_test ค่าที่ได้รับจาก train_test_split() จะถูกแบ่งเป็น 4 ชุดข้อมูล:

- X_train: ข้อมูลสำหรับฝึกโมเดล
- X_test: ข้อมูลสำหรับทดสอบโมเดล
- y_train: ผลลัพธ์ที่ถูกต้องของข้อมูลฝึกโมเดล
- y_test: ผลลัพธ์ที่ถูกต้องของข้อมูลทดสอบ

**อัลกอริทึมของแมชชีนเลิร์นนิงมีกี่ประเภท**

อัลกอริทึมสามารถแบ่งออกได้ตามรูปแบบการเรียนรู้ที่แตกต่างกันสี่ประเภท โดยขึ้นอยู่กับผลลัพธ์ที่คาดหวังและประเภทของข้อมูลที่ป้อนเข้า

**1. แมชชีนเลิร์นนิงแบบมีผู้สอน**

นักวิทยาศาสตร์ข้อมูลจัดหาอัลกอริทึมที่มีข้อมูลการฝึกฝนที่มีการระบุประเภทและกำหนดไว้เพื่อประเมินความสัมพันธ์กัน โดยข้อมูลตัวอย่างจะระบุทั้งข้อมูลที่ป้อนเข้าและผลลัพธ์ของอัลกอริทึม ตัวอย่างเช่น รูปภาพของตัวเลขที่เขียนด้วยลายมือจะมีคำอธิบายประกอบเพื่อระบุว่าตรงกับหมายเลขใด ระบบแมชชีนเลิร์นนิงแบบมีผู้สอนจะสามารถจดจำกลุ่มพิกเซลและรูปร่างที่เกี่ยวข้องกับแต่ละตัวเลขได้ หากมีตัวอย่างที่เพียงพอ ซึ่งในที่สุดแล้วระบบจะจำตัวเลขที่เขียนด้วยลายมือได้ โดยสามารถแยกความแตกต่างระหว่างตัวเลข 9 และ 4 หรือ 6 และ 8 ได้อย่างน่าเชื่อถือ โดยจุดแข็งของแมชชีนเลิร์นนิงแบบมีผู้สอนคือความเรียบง่ายและความสะดวกในการออกแบบ ซึ่งมีประโยชน์เมื่อต้องคาดการณ์ชุดผลลัพธ์ที่เป็นไปได้แบบจำกัด โดยแบ่งข้อมูลออกเป็นหมวดหมู่ หรือรวมผลลัพธ์จากอัลกอริทึมแมชชีนเลิร์นนิงอีกสองแบบเข้าด้วยกัน อย่างไรก็ตาม การระบุประเภทชุดข้อมูลที่ไม่มีการระบุประเภทหลายล้านรายการก็อาจถือเป็นเรื่องที่ท้าทาย เรามาลองเจาะลึกลงไปอีกหน่อยกัน

ลิขสิทธิ์ มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved

**การระบุประเภทข้อมูลคืออะไร**

การระบุประเภทข้อมูลเป็นกระบวนการในการจัดหมวดหมู่ข้อมูลที่ป้อนเข้าด้วยค่าผลลัพธ์ที่กำหนดไว้ที่สอดคล้องกัน โดยจำเป็นต้องมีข้อมูลการฝึกฝนที่มีการระบุประเภทสำหรับแมชชีนเลิร์นนิงแบบมีผู้สอน ตัวอย่างเช่น จะต้องมีการติดแท็กรูปภาพของแอปเปิ้ลและกล้วยหลายล้านรายการด้วยคำว่า "แอปเปิ้ล" หรือ "กล้วย" จากนั้นแอปพลิเคชันแมชชีนเลิร์นนิงจะสามารถใช้ข้อมูลการฝึกฝนนี้ในการทายชื่อของผลไม้เมื่อให้ภาพผลไม้ได้ อย่างไรก็ตาม การระบุประเภทข้อมูลใหม่นับล้านรายการอาจเป็นงานที่ใช้เวลานานและท้าทาย บริการที่ทำงานร่วมกับคนหมู่มาก เช่น Amazon Mechanical Turk สามารถแก้ไขข้อจำกัดของอัลกอริทึมแมชชีนเลิร์นนิงแบบมีผู้สอนได้ในระดับหนึ่ง โดยบริการเหล่านี้ช่วยให้เข้าถึงแรงงานจำนวนมากซึ่งกระจายอยู่ทั่วโลกได้ในราคาไม่แพง โดยช่วยลดความท้าทายในการได้มาซึ่งข้อมูลให้น้อยลง

**2. แมชชีนเลิร์นนิงแบบไม่มีผู้สอน**

อัลกอริทึมแมชชีนเลิร์นนิงแบบไม่มีผู้สอนจะฝึกฝนกับข้อมูลที่ไม่มีการระบุประเภท ซึ่งจะสแกนผ่านข้อมูลใหม่ โดยพยายามสร้างการเชื่อมต่อที่มีความหมายระหว่างข้อมูลที่ป้อนเข้าและผลลัพธ์ที่กำหนดไว้ล่วงหน้า โดยอัลกอริทึมดังกล่าวสามารถระบุรูปแบบและจัดหมวดหมู่ข้อมูลได้ ตัวอย่างเช่น อัลกอริทึมแบบไม่มีผู้สอนสามารถจัดกลุ่มบทความข่าวจากเว็บไซต์ข่าวต่างๆ ออกเป็นหมวดหมู่ทั่วไปต่างๆ เช่น กีฬา อาชญากรรม ฯลฯ ซึ่งสามารถใช้การประมวลผลภาษาธรรมชาติเพื่อทำความเข้าใจความหมายและอารมณ์ในบทความดังกล่าวได้ ในแวดวงการค้าปลีก แมชชีนเลิร์นนิงแบบไม่มีผู้สอนสามารถค้นหารูปแบบในการซื้อสินค้าของลูกค้า และให้ผลการวิเคราะห์ข้อมูล เช่น ลูกค้ามักจะซื้อขนมปังหากซื้อเนยด้วย แมชชีนเลิร์นนิงแบบไม่มีผู้สอนมีประโยชน์อย่างยิ่งสำหรับการจดจำรูปแบบ การตรวจจับความผิดปกติ และการจัดกลุ่มข้อมูลเป็นหมวดหมู่โดยอัตโนมัติ เนื่องจากข้อมูลการฝึกฝนไม่ต้องอาศัยการระบุประเภท จึงตั้งค่าได้ง่าย นอกจากนี้ อัลกอริทึมเหล่านี้ยังสามารถใช้เพื่อล้างและประมวลผลข้อมูลสำหรับการสร้างโมเดลเพิ่มเติมโดยอัตโนมัติได้อีกด้วย โดยข้อจำกัดของวิธีการนี้คือไม่สามารถคาดการณ์ได้อย่างแม่นยำ นอกจากนี้ยังไม่สามารถแยกแยะผลลัพธ์ของข้อมูลบางประเภทแยกจากกันได้อีกด้วย

ลิขสิทธิ์มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved

**3. แมชชีนเลิร์นนิงแบบกึ่งมีผู้สอน**

ตามที่ชื่อระบุไว้ วิธีการนี้จะผสมผสานแมชชีนเลิร์นนิงแบบมีผู้สอนและแมชชีนเลิร์นนิงแบบไม่มีผู้สอนเข้าด้วยกัน โดยเทคนิคนี้อาศัยการใช้ข้อมูลที่มีการระบุประเภทจำนวนเล็กน้อยและข้อมูลที่

ไม่มีการระบุประเภทจำนวนมากในการฝึกฝนระบบ ขั้นแรก เราจะใช้ข้อมูลที่มีการระบุประเภทเพื่อฝึกฝนอัลกอริทึมแมชชีนเลิร์นนิงเพียงบางส่วน หลังจากนั้น อัลกอริทึมที่ได้รับการฝึกฝนเพียงบางส่วนจะระบุประเภทข้อมูลที่ไม่มีการระบุประเภท โดยกระบวนการนี้เรียกว่าการระบุประเภทเทียม จากนั้นโมเดลดังกล่าวจะได้รับการฝึกฝนใหม่เกี่ยวกับการผสมผสานข้อมูลที่เป็นผลลัพธ์โดยไม่ได้เขียนโปรแกรมไว้อย่างชัดเจน ข้อดีของวิธีการนี้คือ คุณไม่ต้องใช้ข้อมูลที่มีการระบุประเภทจำนวนมาก ซึ่งเหมาะเป็นอย่างยิ่งเมื่อต้องทำงานร่วมกับข้อมูลบางประเภท เช่น เอกสารขนาดยาว ซึ่งอาจใช้เวลานานเกินกว่าที่มนุษย์จะอ่านและระบุประเภทได้

**4. การเรียนรู้แบบเสริมกำลัง**

การเรียนรู้แบบเสริมกำลังเป็นวิธีการที่มีค่ารางวัลแนบมากับขั้นตอนต่างๆ ที่อัลกอริทึมต้องดำเนินการตลอดกระบวนการ ดังนั้นเป้าหมายของโมเดลดังกล่าวคือการสะสมคะแนนให้ได้มากที่สุดและบรรลุเป้าหมายสุดท้ายได้ในที่สุด โดยการประยุกต์ใช้การเรียนรู้แบบเสริมกำลังในทางปฏิบัติส่วนใหญ่ในทศวรรษที่ผ่านมาอยู่ในแวดวงของวิดีโอเกม อัลกอริทึมการเรียนรู้แบบเสริมกำลังที่ล้ำสมัยได้บรรลุผลลัพธ์ที่น่าประทับใจทั้งในเกมสุดคลาสสิกและเกมสมัยใหม่ ซึ่งมักจะเอาชนะคู่ต่อสู้ที่เป็นมนุษย์ได้เป็นส่วนใหญ่ แม้ว่าวิธีการนี้จะได้ผลดีที่สุดในสภาพแวดล้อมข้อมูลที่ไม่แน่นอนและซับซ้อน แต่ก็ไม่ค่อยมีการนำมาใช้มากนักในบริบททางธุรกิจ เนื่องจากไม่มีประสิทธิภาพสำหรับงานที่กำหนดไว้อย่างชัดเจน และอคติของนักพัฒนาอาจส่งผลต่อผลลัพธ์ด้วย เพราะในขณะที่นักวิทยาศาสตร์ข้อมูลออกแบบรางวัลต่างๆ สิ่งเหล่านี้ก็อาจมีอิทธิพลต่อผลลัพธ์ได้เช่นกัน

3. โปรแกรม visual studio code

ลิขสิทธิ์ มหาวิทยาลัยนเรศวร
Copyright© by Naresuan University
All rights reserved

**การวิเคราะห์ข้อมูล**

1. การวิเคราะห์สถิติเชิงพรรณนาและการวิเคราะห์ภาพรวมการเกิดอุบัติเหตุในประเทศไทย

2. การวิเคราะห์ข้อมูลจะทำโดยใช้ Machine Learning โดยใช้ Random Forest เพื่อทำนายความรุนแรงของอุบัติเหตุจากฟีเจอร์ต่างๆ ที่ใช้สำหรับฝึกโมเดล โดยประกอบไปด้วย คอลลัมจากdata เช่น วัน เดือน ปี พิกัด ข้อมูลจะถูกแบ่งออกเป็นชุดฝึกและชุดทดสอบเพื่อประเมินความแม่นยำของโมเดล

โดยมีโมเดลทั้ง 4 โมเดล

ตัวแปรต้น X

Model 1 วัน, เดือน, ชั่วโมง, พิกัด

Model 2 วัน, เดือน, ชั่วโมง, จังหวัด

Model 3 วัน, เดือน, ชั่วโมง,

Model 4 วัน, เดือน, ชั่วโมง, ประเภทรถ

ตัวแปรตาม Y

-อุบัติเหตุร้ายแรง: อุบัติเหตุทางถนนที่ส่งผลให้มีผู้เสียชีวิต

- อุบัติเหตุบาดเจ็บสาหัส: อุบัติเหตุทางถนนที่ฝ่ายใดฝ่ายหนึ่งต้องได้รับการรักษาพยาบาลและนำส่งโรงพยาบาล

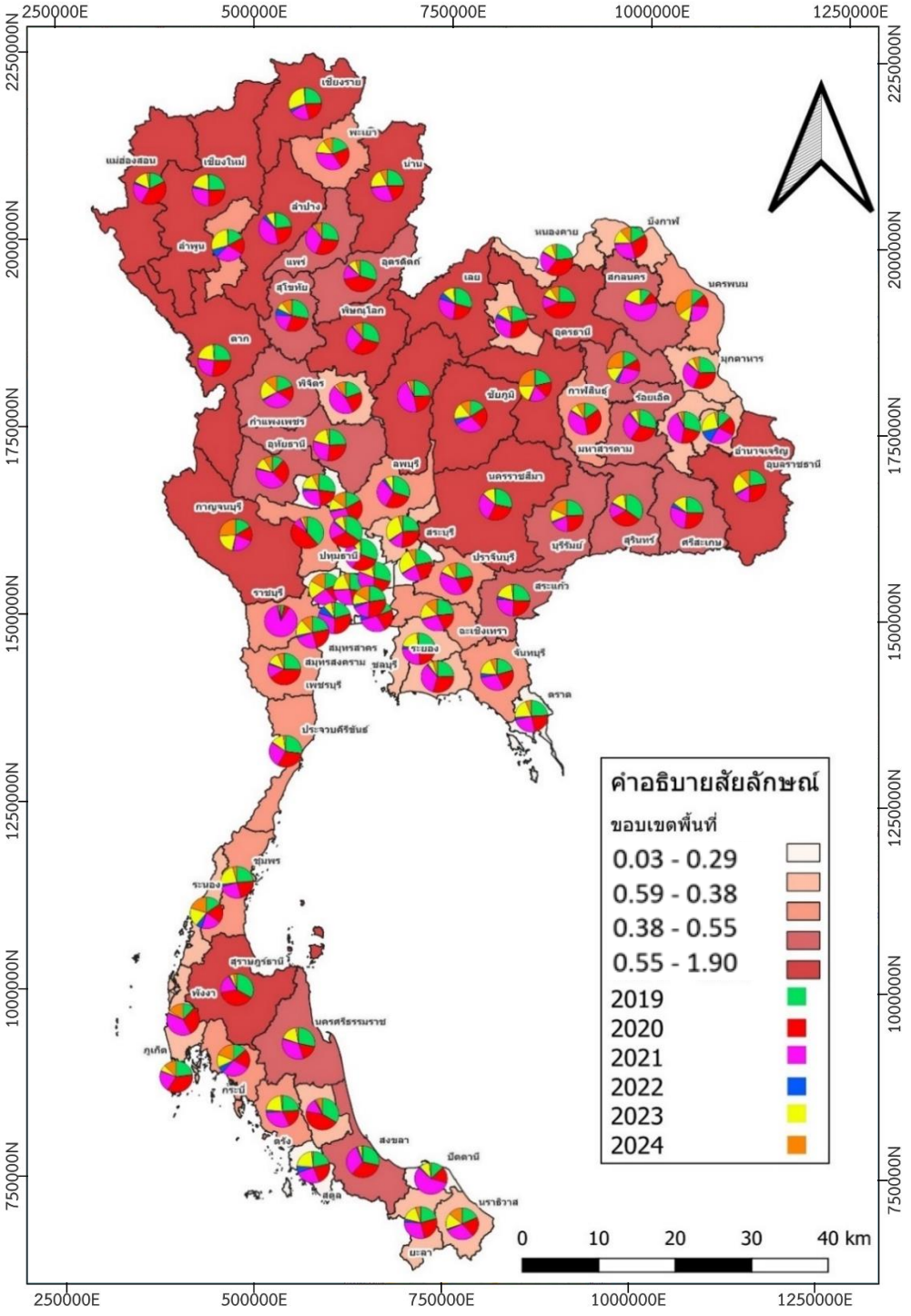
- อุบัติเหตุที่ได้รับบาดเจ็บปานกลาง: อุบัติเหตุทางถนนซึ่งไม่มีใครต้องการความช่วยเหลือทางการแพทย์ แต่ได้รับบาดเจ็บเล็กน้อยและมีบาดแผลเล็กน้อย หรือความเสียหายทางทรัพย์สินไม่มาก

- อุบัติเหตุที่ได้รับบาดเจ็บเล็กน้อย: อุบัติเหตุที่มีความรุนแรงต่ำ หรือไม่มีการบาดเจ็บที่รุนแรง

3.นอกจากนี้ยังจะมีการใช้เครื่องมือ visual studio code เพื่อทำการทำนายความเสี่ยงของอุบัติเหตุบนท้องถนน

ลิขสิทธิ์ มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved

## ผลการวิเคราะห์ข้อมูล



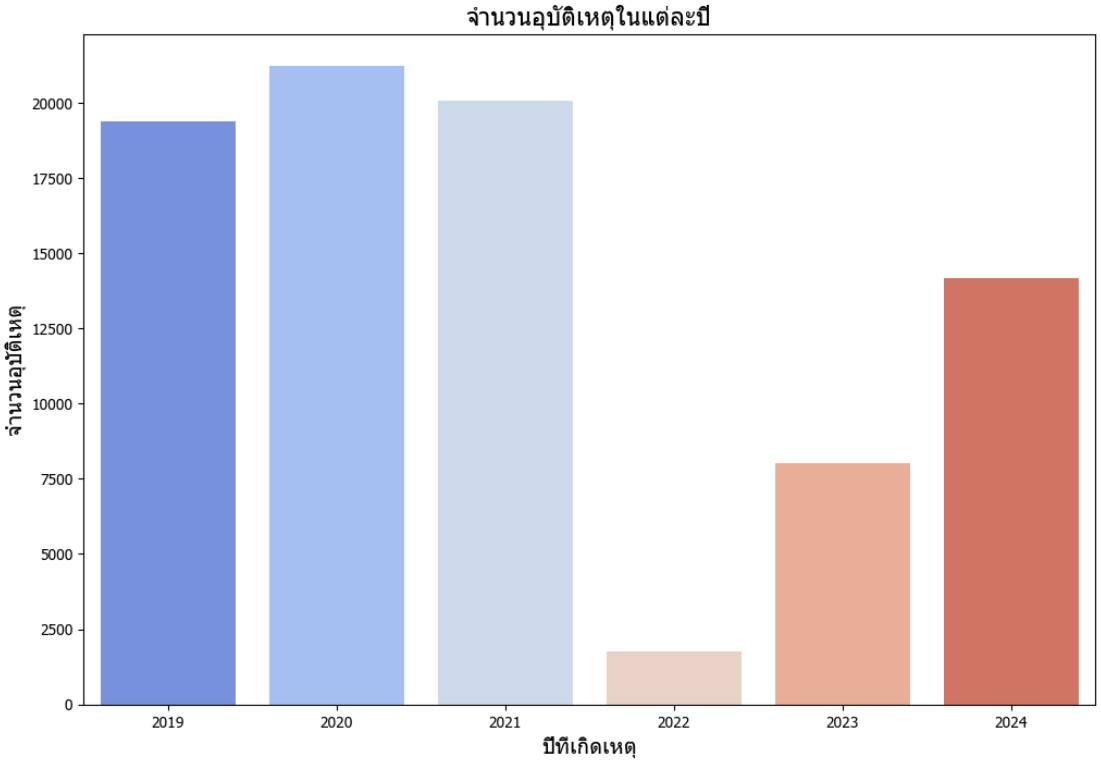

**ภาพ 4.** แผนที่แสดงการเกิดอุบัติเหตุในแต่ละปีของแต่ละจังหวัด

จังหวัดที่เกิดอุบัติเหตุมากที่สุดในแต่ละปี ดังนี้

ปี 2019 กรุงเทพมหานคร 1528 ครั้ง ปี 2020 กรุงเทพมหานคร 1689 ครั้ง

ปี 2021 กรุงเทพมหานคร 1199 ครั้ง ปี 2022 นครราชสีมา 103 ครั้ง

ปี 2023 กรุงเทพมหานคร 776 ครั้ง ปี 2024 กรุงเทพมหานคร 1613 ครั้ง




**ภาพ 5.** กราฟแสดงจำนวนอุบัติเหตุที่เกิดขึ้นในปี 2019-2024

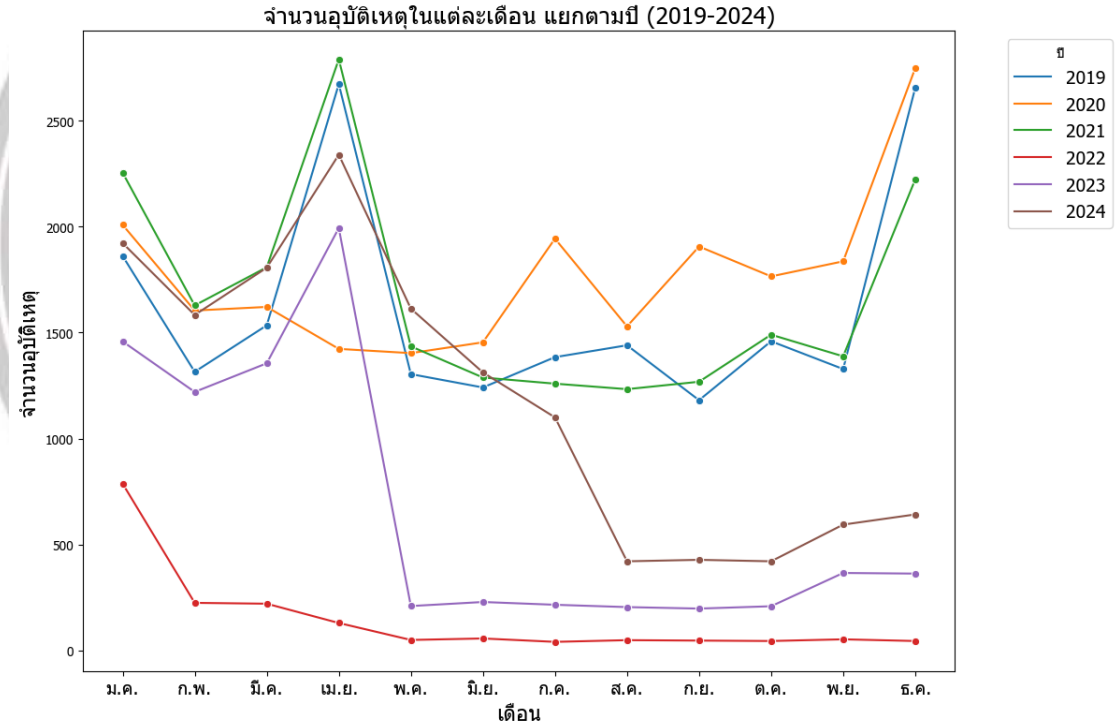
ลิขสิทธิ์ มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved

กราฟนี้แสดงข้อมูลจำนวนอุบัติเหตุที่เกิดขึ้นในแต่ละปี ระหว่างปี 2019 ถึง 2024 โดยใช้แถบสีเพื่อเปรียบเทียบจำนวนอุบัติเหตุในแต่ละปี ดังนี้:

ปี 2019 - 2021จำนวนอุบัติเหตุในแต่ละปีอยู่ในระดับสูงใกล้เคียงกัน โดยปี 2020 มีจำนวนอุบัติเหตุสูงที่สุดในกราฟ ซึ่งอาจเป็นเพราะการเดินทางและกิจกรรมทางเศรษฐกิจยังคงดำเนินไปอย่างปกติในช่วงต้นปี ก่อนที่จะมีการระบาดของโรค COVID-19 อย่างกว้างขวาง

ปี 2022 จำนวนอุบัติเหตุลดลงอย่างมากเมื่อเทียบกับปีก่อนหน้า สาเหตุหลักอาจมาจากการมีข้อมูลอุบัติเหตุที่น้อยกว่าปีอื่นๆ

ปี 2023 – 2024 จำนวนอุบัติเหตุเริ่มกลับมาสูงขึ้นอีกครั้งในปี 2023 และเพิ่มมากขึ้นในปี 2024 ซึ่งอาจเป็นเพราะการผ่อนคลายมาตรการ COVID-19 ประชาชนกลับมาเดินทางและทำกิจกรรมตามปกติมากขึ้น ส่งผลให้มีการเกิดอุบัติเหตุเพิ่มขึ้น และส่งผลให้การจราจรเบาบางลงทั่วประเทศหลังจากการระบาดของ COVID-19 คลี่คลาย มาตรการควบคุมโรคต่าง ๆ ก็ผ่อนคลายลง ประชาชนกลับมาเดินทางมากขึ้น ทำให้จำนวนอุบัติเหตุกลับมาเพิ่มขึ้นในปี 2023 และปี 2024 กราฟนี้สะท้อนให้เห็นถึงการเปลี่ยนแปลงของจำนวนอุบัติเหตุในช่วงเวลาที่มีสถานการณ์พิเศษ ซึ่งเป็นปัจจัยที่ไม่คาดคิดแต่ส่งผลต่อจำนวนอุบัติเหตุอย่างมีนัยสำคัญ



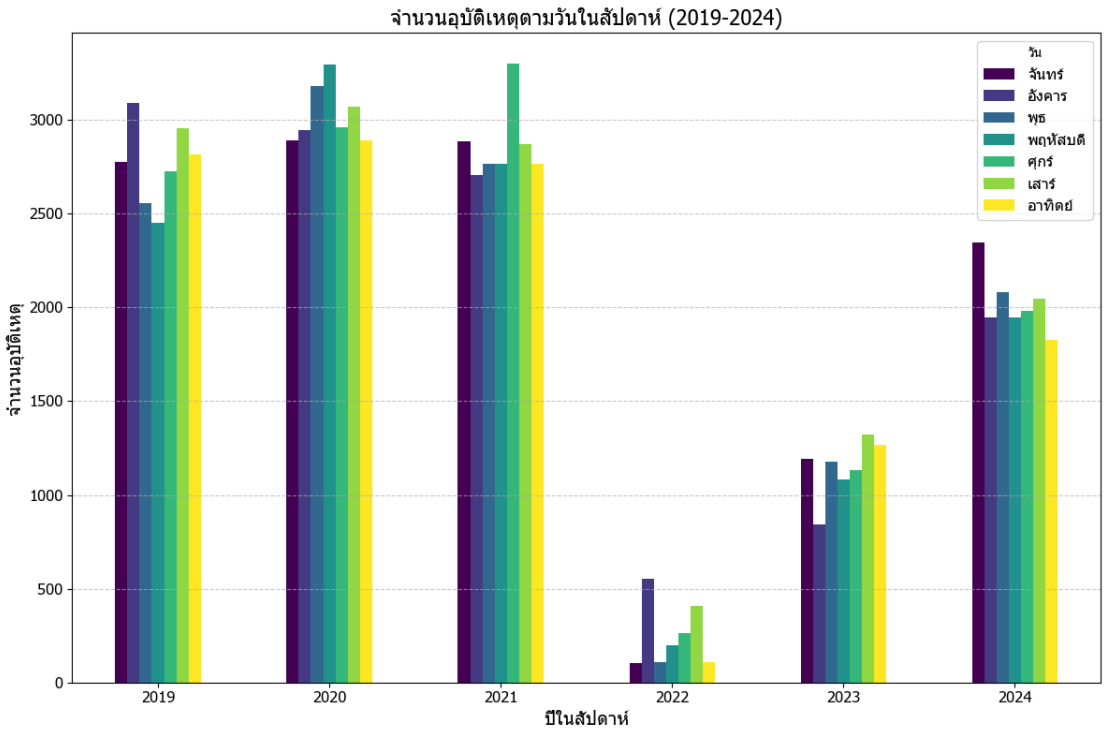

**ภาพ 6.** กราฟเส้นแสดงจำนวนอุบัติเหตุในแต่ละเดือนแยกตามรายปี 2019-2024

กราฟนี้แสดงจำนวนอุบัติเหตุที่เกิดขึ้นในแต่ละเดือน แยกตามปี ตั้งแต่ปี 2019 ถึงปี 2024 โดยมีแนวโน้มและลักษณะการเปลี่ยนแปลงที่สามารถสรุปได้ดังนี้:

ลิขสิทธิ์ มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved

ในทุกๆปีเดือนเมษายนจะมีจำนวนอุบัติเหตุสูงมากที่สุด ซึ่งสอดคล้องกับเทศกาลสงกรานต์ในประเทศไทย ที่มีการเดินทางมากขึ้นและเป็นช่วงที่เกิดอุบัติเหตุบ่อยครั้ง สาเหตุหลักมาจากการเดินทางกลับภูมิลำเนาและการเฉลิมฉลองที่ทำให้เกิดการขับขี่ในสภาพมึนเมา รองลงมาจะเป็นเดือนธันวาคมจำนวนอุบัติเหตุกลับมามีแนวโน้มสูงขึ้นอีกครั้ง ช่วงเทศกาลมักเป็นช่วงเวลาที่มีอุบัติเหตุทางถนนเกิดขึ้นสูง เนื่องจากเป็นเดือนที่ผู้คนจำนวนมากเดินทางเพื่อเฉลิมฉลองช่วงเทศกาลสิ้นปี เช่น คริสต์มาสและปีใหม่การเดินทางทั้งในระยะใกล้และระยะไกลการกลับบ้านเพื่อไปพบครอบครัว หรือ

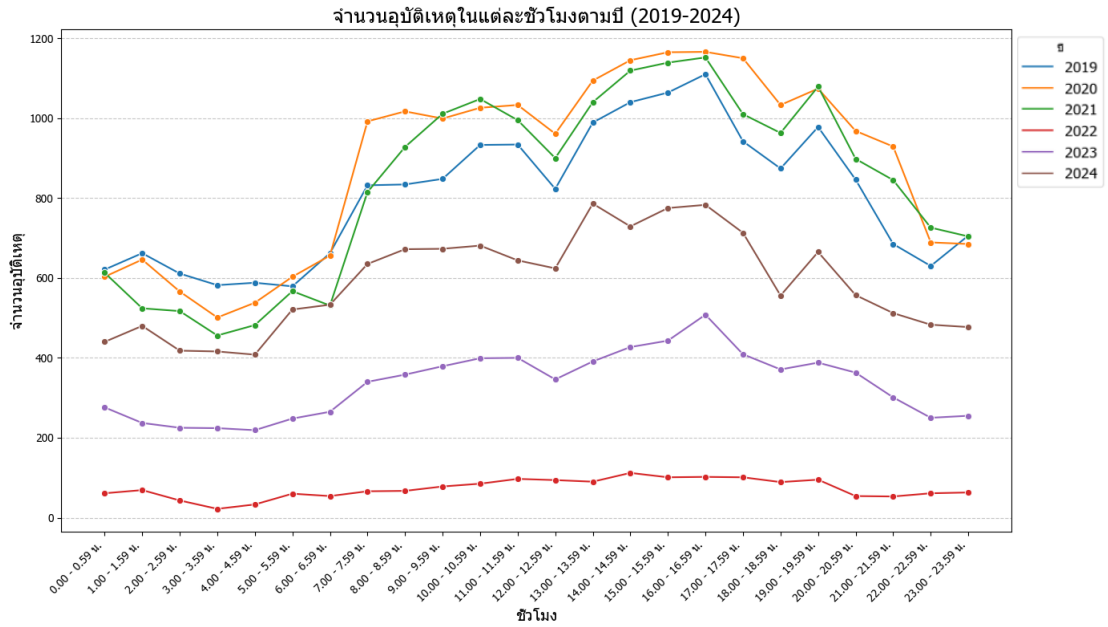
การท่องเที่ยวในช่วงวันหยุดยาว ล้วนส่งผลให้จำนวนอุบัติเหตุบนถนนเพิ่มสูงขึ้น ปัจจัยหลักที่ทำให้เดือนธันวาคมเป็นช่วงที่มีอุบัติเหตุสูงนั้นประกอบด้วยหลายประการ ในบางพื้นที่ในประเทศไทย เดือนธันวาคมเป็นช่วงที่มีสภาพอากาศเย็นหรือมีหมอกบาง ซึ่งอาจทำให้ทัศนวิสัยในการขับขี่ลดลง ถนนบางการเดินทางเพื่อกลับบ้านและท่องเที่ยวส่งผลให้เกิดการขับขี่ระยะไกลและเกิดอุบัติเหตุบ่อยขึ้น




**ภาพที่ 7.** กราฟแสดงจำนวนอุบัติเหตุตามในวันสัปดาห์แยกตามรายปี 2019-2024

แนวโน้มอุบัติเหตุรายวัน วันศุกร์และวันเสาร์ มีจำนวนอุบัติเหตุสูงอย่างต่อเนื่องทุกปี เนื่องจากเป็นวันสิ้นสัปดาห์ที่มีการเดินทางและเฉลิมฉลองมาก รองลงมาคือวันอาทิตย์โดยจำนวนอุบัติเหตุเพิ่มสูงขึ้นในปี 2024 ซึ่งอาจสะท้อนถึงการเดินทางท่องเที่ยวหรือกลับบ้าน

ลิขสิทธิ์ มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved



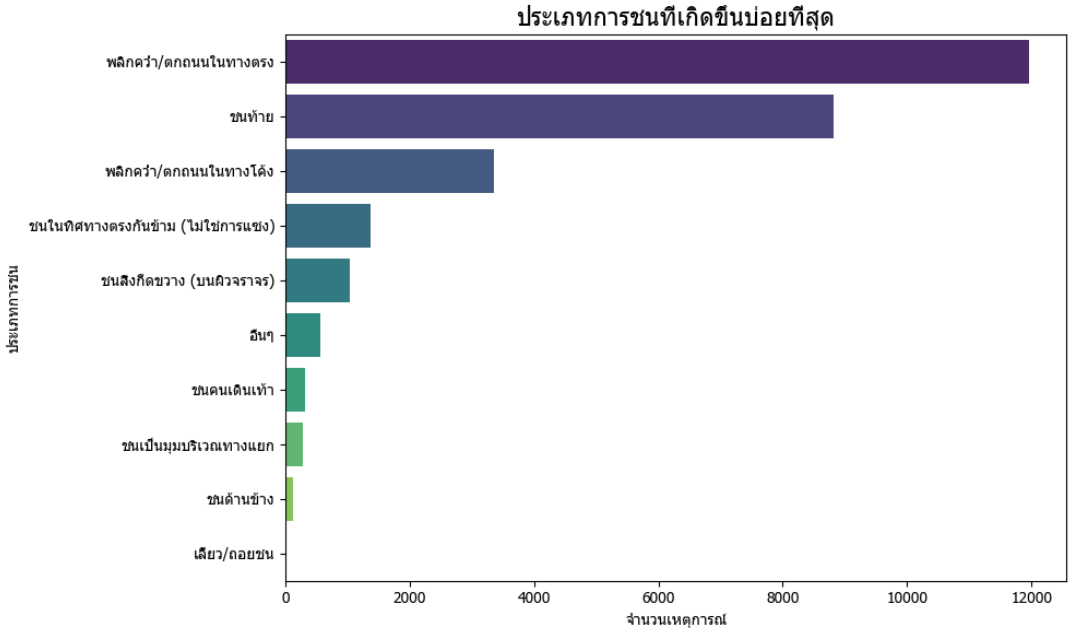

**ภาพ 8.** กราฟแสดงจำนวนอุบัติเหตุในแต่ละชั่วโมงแยกตามรายปี 2019-2024

ในแต่ละปีตั้งแต่ 2019 ถึง 2024 จำนวนอุบัติเหตุในช่วงเช้าประมาณ 6.00-8.00 น. และช่วงเย็นตั้งแต่ 16.00-19.00 น. มีแนวโน้มที่จะสูงที่สุด โดยสาเหตุที่เป็นไปได้อาจจะมาจากช่วงเวลานี้เป็นช่วงที่คนเดินทางไปทำงานหรือกลับบ้าน ส่งผลให้เกิดอุบัติเหตุบ่อยครั้ง

**กำหนดชั่วโมงดังนี้**

- 00.00 - 00.59 น. • 01.00 - 01.59 น. • 02.00 - 02.59 น.• 03.00 - 03.59 น.
- 04.00 - 04.59 น. • 05.00 - 05.59 น. • 06.00 - 06.59 น. • 07.00 - 07.59 น.
- 08.00 - 08.59 น. • 09.00 - 09.59 น. • 10.00 - 10.59 น. • 11.00 - 11.59 น.
- 12.00 - 12.59 น. • 13.00 - 13.59 น. • 14.00 - 14.59 น.• 15.00 - 15.59 น.
- 16.00 - 16.59 น. • 17.00 - 17.59 น. • 18.00 - 18.59 น. • 19.00 - 19.59 น.
- 20.00 - 20.59 น. • 21.00 - 21.59 น. • 22.00 - 22.59 น. • 23.00 - 23.59 น.

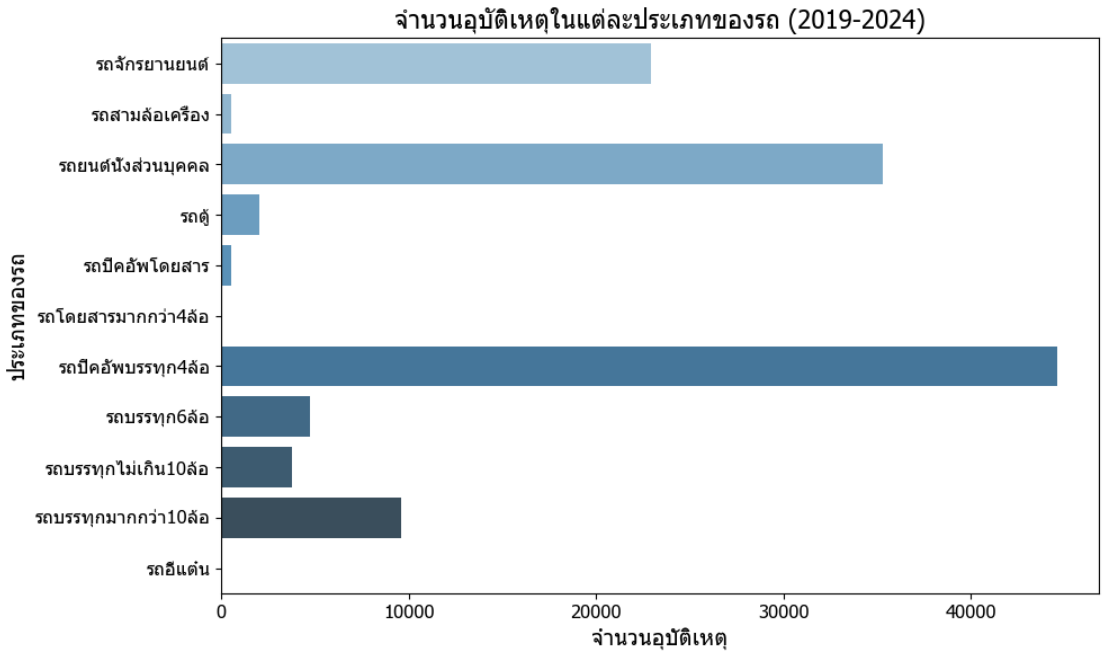
ลิขสิทธิ์ มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved




**ภาพ 9.** กราฟแสดงประเภทการชนที่เกิดขึ้นบ่อยที่สุด ตั้งแต่ปี 2019-2024

ประเภทของการชนที่เกิดขึ้นบ่อยที่สุด โดยมีการจัดลำดับจากจำนวนเหตุการณ์สูงสุดไปต่ำสุด อันดับหนึ่ง พลิกคว่ำ/ตกถนนในทางตรง เป็นประเภทที่เกิดขึ้นมากที่สุด โดยมีจำนวนเหตุการณ์เกิน 12,000 ครั้ง เนื่องจากการขับขี่ด้วยความเร็วสูงและการควบคุมรถที่ผิดพลาด อันดับสอง ชนท้าย เป็นประเภทที่เกิดบ่อยรองลงมา ประมาณ 8,000 ครั้ง เกิดจากการเว้นระยะห่างไม่เพียงพอและการเบรกกะทันหัน และอันดับที่สาม คือ พลิกคว่ำ/ตกถนนในทางโค้ง ประมาณ 4,000 ครั้ง มักเกิดเมื่อผู้ขับขี่ใช้ความเร็วสูงในบริเวณที่ต้องเลี้ยว

ลิขสิทธิ์ มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved




**ภาพ 10.** กราฟแสดงจำนวนอุบัติเหตุในแต่ละประเภทของรถ ตั้งแต่ปี 2019-2024

รถปิคอัพและรถยนต์ส่วนบุคคล*เกิดอุบัติเหตุบ่อยที่สุด (รถปิคอัพ > 40,000 ครั้ง, รถยนต์ส่วนบุคคล 30,000 ครั้ง) เนื่องจากการใช้งานแพร่หลายและเป็นยานพาหนะหลักของคนไทย รถจักรยานยนต์ แม้จำนวนอุบัติเหตุจะน้อยกว่ารถยนต์ แต่ความรุนแรงของบาดเจ็บสูง รถบรรทุกขนาดใหญ่ เกิดอุบัติเหตุบ่อย และความเสียหายรุนแรงเมื่อเกิดการชน รถโดยสารและรถตุ๊กตุ๊กเกิดอุบัติเหตุน้อยเมื่อเทียบกับยานพาหนะประเภทอื่น

ประเภทของยานพาหนะ รถจักรยานยนต์มีความเสี่ยงสูงต่อการเสียหลักและบาดเจ็บเนื่องจากไม่มีโครงสร้างป้องกัน พบว่ามีจำนวนอุบัติเหตุที่รุนแรงมากกว่าเมื่อเทียบกับรถยนต์ รถยนต์ส่วนบุคคลเป็นยานพาหนะที่ใช้มาก ทำให้เกิดอุบัติเหตุบ่อย รถปิคอัพ รถประเภทนี้นิยมใช้ทั้งในชีวิตประจำวันและเชิงพาณิชย์ ทำให้เสี่ยงต่อการเกิดอุบัติเหตุสูง รถบรรทุก ขนาดและน้ำหนักมากทำให้หยุดรถลำบากเมื่อเกิดเหตุฉุกเฉิน หากเกิดอุบัติเหตุ มักมีความรุนแรงและสร้างความเสียหายมาก

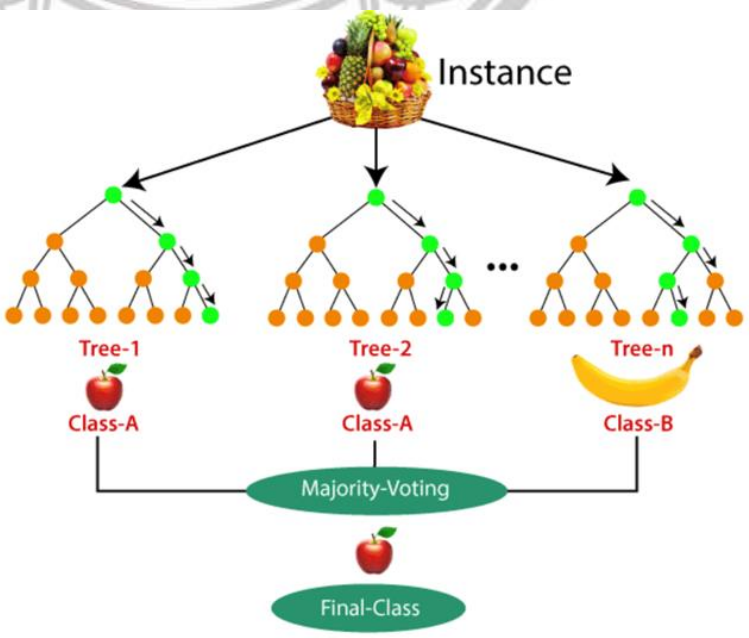
ลิขสิทธิ์มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved

ในการวิจัยนี้ เราตรวจสอบการทำนายอุบัติเหตุทางถนนในฐานะปัญหาการจำแนกประเภทที่สามารถจำแนกประเภทความรุนแรงของอุบัติเหตุออกเป็นสี่ประเภท ได้แก่ เสียชีวิต ร้ายแรง ปากลาง และเล็กน้อย เนื่องจากมีมากกว่าสองคลาส (หรือหลายคลาส) ที่เกี่ยวข้องกับปัญหาการวิจัยกาทำนายความรุนแรงของอุบัติเหตุทางถนนนี้ จึงกลายเป็น ปัญหา การจำแนกประเภทหลายคลาสการวิจัยนี้ได้พิจารณาอัลกอริทึม ML แบบรวมเพื่อวิเคราะห์ชุดข้อมูลอุบัติเหตุทางถนน อัลกอริทึม ML แบบรวมจะ

รวมโมเดลพื้นฐาน หลายตัวเข้าด้วยกันเพื่อให้ได้ ประสิทธิภาพที่เหมาะสมที่สุดและลดการกระจายของการทำนาย เพื่อปรับปรุงการยอมรับการทำนายอุบัติเหตุทางถนน จำเป็นต้องทำความเข้าใจปัจจัยเหล่านี้และอิทธิพลของปัจจัยเหล่านี้ที่มีต่อโมเดล ซึ่งเป็นประเด็นหลักของการวิจัยนี้ ดังนั้น จึงต้องวิเคราะห์ค่า Shapley เพื่อทำความเข้าใจถึงการมีส่วนร่วมของคุณลักษณะต่อตัวแปรเป้าหมาย ผลการวิจัยถูกนำมาใช้เพื่อกำหนดความสัมพันธ์พื้นฐานของปัจจัยที่ส่งผลต่ออุบัติเหตุทางถนน แสดงขั้นตอนของงานนี้ งานวิจัยนี้วิเคราะห์ประสิทธิภาพของโมเดล ML ต่างๆ ในกลุ่ม โดยพิจารณาจากการทำนายความรุนแรงของอุบัติเหตุทางถนนอย่างแม่นยำ การทำความเข้าใจความแม่นยำของการทำนาย และการคำนวณคะแนน F1 และการเรียกคืน หัวข้อย่อยต่อไปนี้จะอธิบายวิธี ML ทั้งหมดที่สำรวจในเอกสารนี้โดยย่อ

อัลกอริทึม (Random Forest : RF)

RF เป็นอัลกอริทึมการเรียนรู้แบบมีผู้ดูแล ML ที่ใช้ต้นไม้ตัดสินใจเป็นชุดซึ่งใช้กันอย่างแพร่หลาย โดยอาศัยต้นไม้ตัดสินใจเป็นชุดเพื่อลดความเสี่ยงในการติดตั้งเกิน ประสิทธิภาพของ RF ขึ้นอยู่กับไฮเปอร์พารามิเตอร์หลักสามประการ ได้แก่ จำนวนต้นไม้ ขนาดโหนด และจำนวนฟีเจอร์ที่สุ่มตัวอย่าง การวิเคราะห์ความสำคัญของฟีเจอร์สามารถประเมินได้โดยการประมาณปัจจัยที่มีส่วนทำให้เกิดอุบัติเหตุบนท้องถนนและอิทธิพลของปัจจัยเหล่านี้ต่อความรุนแรงของการบาดเจ็บ




**ภาพ 11.** หลักการสร้างโมเดล Random Forest

หลักการของ Random Forest คือ สร้าง model จาก Decision Tree หลายๆ model ย่อย ๆ (ตั้งแต่ 10 model ถึง มากกว่า 1000 model) โดยแต่ละ model จะได้รับ data set ไม่เหมือนกัน ซึ่งเป็น subset ของ data set ทั้งหมด ตอนทำ prediction ก็ให้แต่ละ Decision Tree ทำ prediction ของใครของมัน และคำนวณผล prediction ด้วยการ vote output ที่ ถูกเลือกโดย Decision Tree มากที่สุด (กรณี classification) หรือ หาค่า mean จาก output ของแต่ละ Decision Tree (กรณี regression)

Decision Tree แต่ละ model ใน Random Forest ถือว่าเป็น weak learner — ประมาณว่าเป็น model ที่ไม่เก่งเท่าไหร่ แต่พอนำเอาแต่ละ Decision Tree มาทำ prediction ร่วมกัน ก็จะได้ model รวมที่มีความเก่ง และแม่นยำมากกว่า Decision Tree ที่ทำ prediction แบบเดี่ยว ๆ

**หลักการทำ Random Forest คือ**

1. sample ข้อมูล (bootstrapping) จาก data set ทั้งหมด ให้ได้ข้อมูลออกมา n ชุด ที่ไม่เหมือนกัน ตามจำนวน Decision Tree ใน Random Forest เช่น data set ตั้งต้นมีอยู่ 10 feature (X1,X2,...,X10) แต่ละ Decision Tree จะได้ feature ไปไม่เหมือนกัน และ จะได้ข้อมูลไม่ครบทุก row ด้วยจาก data set ทั้งหมดด้วย (X1 -> X1',X2->X2',...)
2. สร้าง model Decision Tree สำหรับแต่ละชุดข้อมูล
3. ทำ aggregation ผลลัพธ์ จากแต่ละ model (bagging) เช่น voting ในกรณี classification หรือ หาค่า mean ในกรณี regression

ลิขสิทธิ์ มหาวิทยาลัยนเรศวร

**ขั้นตอนการวิเคราะห์ด้วยเทคนิควิธีการเรียนรู้ของเครื่อง (Machine Learning)**

1. **การโหลดข้อมูล ทุกโมเดลเริ่มจากการโหลดข้อมูลอุบัติเหตุจากไฟล์ CSV**

Copyright by Naresuan University

```python
import pandas as pd

# โหลดข้อมูลจากไฟล์
data = pd.read_csv('acc_data.csv')
```

**ภาพ 12.** การโหลดข้อมูล

- import pandas as pd: นำเข้าไลบรารี Pandas ซึ่งใช้สำหรับจัดการข้อมูลในรูปแบบ DataFrame
- pd.read_csv('acc_data.csv'): อ่านไฟล์ CSV ซึ่งมีข้อมูลอุบัติเหตุ
- ไฟล์นี้ควรประกอบด้วย: คอลัมน์ฟีเจอร์ (Features) เช่น วัน, เดือน, ชั่วโมง, จังหวัด และตัวแปรเป้าหมาย (Target) เช่น ระดับความรุนแรงของอุบัติเหตุ

**2. การเตรียมข้อมูล**

ข้อมูลจะถูกแปลงเป็นตัวเลข (numeric) เพื่อให้โมเดลทำงานได้:

```
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import LabelEncoder

# แปลงข้อมูล 'ความรุนแรง' เป็นตัวเลข
encoder = LabelEncoder()
data['ความรุนแรง'] = encoder.fit_transform(data['ความรุนแรง'])

# แยก X (features) และ Y (target)
X = data[['วัน', 'เดือน', 'ชั่วโมง']]
y = data['ความรุนแรง']

# แบ่งข้อมูลเป็นชุด train/test
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

**ภาพ 13.** การเตรียมข้อมูล

2.1 LabelEncoder():

- ใช้แปลงตัวแปรเป้าหมาย ('ความรุนแรง') จากข้อความ เช่น "อุบัติเหตุเล็กน้อย" เป็นตัวเลข เช่น 0, 1, 2, 3
- ทำให้โมเดลเข้าใจข้อมูลเป้าหมายในรูปแบบตัวเลขได้ง่ายขึ้น

ลิขสิทธิ์มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved

2.2 แยกฟีเจอร์และเป้าหมาย:

- X: ประกอบด้วยฟีเจอร์ที่ใช้สร้างโมเดล เช่น วัน, เดือน, ชั่วโมง
- y: เป็นตัวแปรเป้าหมาย 'ความรุนแรง'

2.3 การแบ่งข้อมูลด้วย train_test_split:

- test_size=0.2: แบ่งข้อมูล 20% สำหรับทดสอบ และ 80% สำหรับเทรนโมเดล

- random_state=42: ใช้ควบคุมการสุ่มเพื่อให้ได้ผลลัพธ์เดียวกันทุกครั้งที่รัน

ผลลัพธ์: ชุดข้อมูลสำหรับเทรน (X_train, y_train) และชุดข้อมูลสำหรับทดสอบ (X_test, y_test)

**3. การสร้างโมเดล Random Forest**

```
from sklearn.ensemble import RandomForestClassifier

# สร้างโมเดล Random Forest
model = RandomForestClassifier(n_estimators=100, random_state=42)

# เทรนโมเดล
model.fit(X_train, y_train)
```

**ภาพ 14.** การสร้างโมเดล

3.1 RandomForestClassifier:

- Random Forest เป็นเทคนิคการเรียนรู้ของเครื่องที่รวมผลจากต้นไม้การตัดสินใจ (Decision Trees) หลาย ๆ ต้น
- n_estimators=100: ใช้ต้นไม้ 100 ต้นในโมเดล
- random_state=42: เพื่อควบคุมการสุ่ม

3.2 model.fit(X_train, y_train):

3.3เทรนโมเดลด้วยข้อมูลฝึก (Training Data) ที่แยกไว้ก่อนหน้านี้

ผลลัพธ์: โมเดลที่เทรนแล้ว พร้อมใช้ทำนายผลลัพธ์

ลิขสิทธิ์ มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved

**4. การทำนายผลและประเมินประสิทธิภาพ**

```python
from sklearn.metrics import classification_report, accuracy_score

# ทำนายผลบนชุดทดสอบ
y_pred = model.predict(X_test)

# ประเมินผล
print(classification_report(y_test, y_pred, target_names=encoder.classes_))
print(f'Accuracy: {accuracy_score(y_test, y_pred)}')
```

**ภาพ 15.** การทำนายผลและประเมินประสิทธิภาพ

4.1 model.predict(X_test):

- ใช้โมเดลที่เทรนแล้วในการทำนายผลลัพธ์ระดับความรุนแรงของข้อมูลในชุดทดสอบ

4.2 classification_report:

- Precision: ความแม่นยำของโมเดลในการทำนายแต่ละหมวดหมู่
- Recall: ความสามารถของโมเดลในการตรวจจับตัวอย่างจริงในแต่ละหมวดหมู่
- F1-Score: ค่าเฉลี่ยถ่วงน้ำหนักระหว่าง Precision และ Recall
- Support: จำนวนตัวอย่างในแต่ละหมวดหมู่

4.3 accuracy_score:

- แสดงค่า Accuracy ของโมเดล ซึ่งเป็นสัดส่วนของการทำนายที่ถูกต้องทั้งหมด

ผลลัพธ์: รายงานประสิทธิภาพของโมเดลและค่า Accuracy

**5. การใช้งานโมเดล**

หลังจากประเมินผล หากโมเดลทำงานได้ดี คุณสามารถบันทึกโมเดลไว้ใช้งานในอนาคต:

ลิขสิทธิ์ มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved

```python
import joblib

# บันทึกโมเดล
joblib.dump(model, 'road_accident_model.pkl')

# โหลดโมเดลกลับมาใช้
loaded_model = joblib.load('road_accident_model.pkl')
```

**ภาพ 16.** การใช้งานโมเดล

| คุณลักษณะ | Model 1 | Model 2 | Model 3 | Model 4 |
|---|---|---|---|---|
| ค่าความถูกต้อง | 61% | 44% | 61% | 58% |

ผลลัพธ์สำหรับการทำนายจากทั้ง 4 โมเดล

| คุณลักษณะ | Model 1 | Model 2 | Model 3 | Model 4 |
|---|---|---|---|---|
| Accuracy | 0.61 | 0.44 | 0.61 | 0.58 |
| Precision | อุบัติเหตุถึงแก่ชีวิต: 0.27 | อุบัติเหตุถึงแก่ชีวิต: 0.14 | อุบัติเหตุถึงแก่ชีวิต: 0.20 | อุบัติเหตุถึงแก่ชีวิต: 0.25 |
| | อุบัติเหตุปานกลาง: 0.38 | อุบัติเหตุปานกลาง: 0.46 | อุบัติเหตุปานกลาง: 0.28 | อุบัติเหตุปานกลาง: 0.30 |
| | อุบัติเหตุรุนแรง: 0.20 | อุบัติเหตุรุนแรง: 0.15 | อุบัติเหตุรุนแรง: 0.00 | อุบัติเหตุรุนแรง: 0.09 |
| | อุบัติเหตุเล็กน้อย: 0.64 | อุบัติเหตุเล็กน้อย: 0.50 | อุบัติเหตุเล็กน้อย: 0.62 | อุบัติเหตุเล็กน้อย: 0.66 |
| Recall | อุบัติเหตุถึงแก่ชีวิต: 0.04 | อุบัติเหตุถึงแก่ชีวิต: 0.07 | อุบัติเหตุถึงแก่ชีวิต: 0.00 | อุบัติเหตุถึงแก่ชีวิต: 0.12 |
| | อุบัติเหตุปานกลาง: 0.18 | อุบัติเหตุปานกลาง: 0.51 | อุบัติเหตุปานกลาง: 0.03 | อุบัติเหตุปานกลาง: 0.18 |
| | อุบัติเหตุรุนแรง: 0.02 | อุบัติเหตุรุนแรง: 0.09 | อุบัติเหตุรุนแรง: 0.00 | อุบัติเหตุรุนแรง: 0.02 |
| | อุบัติเหตุเล็กน้อย: 0.91 | อุบัติเหตุเล็กน้อย: 0.54 | อุบัติเหตุเล็กน้อย: 0.98 | อุบัติเหตุเล็กน้อย: 0.86 |
| F1-Score | อุบัติเหตุถึงแก่ชีวิต: 0.07 | อุบัติเหตุถึงแก่ชีวิต: 0.10 | อุบัติเหตุถึงแก่ชีวิต: 0.00 | อุบัติเหตุถึงแก่ชีวิต: 0.16 |
| | อุบัติเหตุปานกลาง: 0.24 | อุบัติเหตุปานกลาง: 0.48 | อุบัติเหตุปานกลาง: 0.06 | อุบัติเหตุปานกลาง: 0.22 |
| | อุบัติเหตุรุนแรง: 0.04 | อุบัติเหตุรุนแรง: 0.11 | อุบัติเหตุรุนแรง: 0.00 | อุบัติเหตุรุนแรง: 0.03 |
| | อุบัติเหตุเล็กน้อย: 0.75 | อุบัติเหตุเล็กน้อย: 0.52 | อุบัติเหตุเล็กน้อย: 0.76 | อุบัติเหตุเล็กน้อย: 0.74 |
| Macro Average | Precision: 0.37 | Precision: 0.31 | Precision: 0.27 | Precision: 0.32 |
| | Recall: 0.29 | Recall: 0.31 | Recall: 0.25 | Recall: 0.29 |
| | F1-Score: 0.27 | F1-Score: 0.30 | F1-Score: 0.20 | F1-Score: 0.29 |
| Weighted Average | Precision: 0.52 | Precision: 0.42 | Precision: 0.47 | Precision: 0.50 |
| | Recall: 0.61 | Recall: 0.44 | Recall: 0.61 | Recall: 0.58 |
| | F1-Score: 0.53 | F1-Score: 0.43 | F1-Score: 0.48 | F1-Score: 0.53 |

**ภาพ 17.** ผลการวิเคราะห์

ลิ
Copyri
y

**Accuracy** Model 1 และ Model 3 มี Accuracy สูงที่สุดที่ 0.61 ในขณะที่ Model 2 มี Accuracy ต่ำที่สุดที่ 0.44

**Precision** Model 1 มี Precision สูงในหมวดหมู่ "อุบัติเหตุเล็กน้อย" ที่ 0.64 และต่ำใน "อุบัติเหตุถึงแก่ชีวิต" ที่ 0.27 Model 4 มี Precision ที่ดีขึ้นในหมวดหมู่ "อุบัติเหตุเล็กน้อย" ที่ 0.66

Model 3 มี Precision ต่ำในหลายหมวดหมู่ โดยเฉพาะใน "อุบัติเหตุรุนแรง" ซึ่งเป็น 0.00

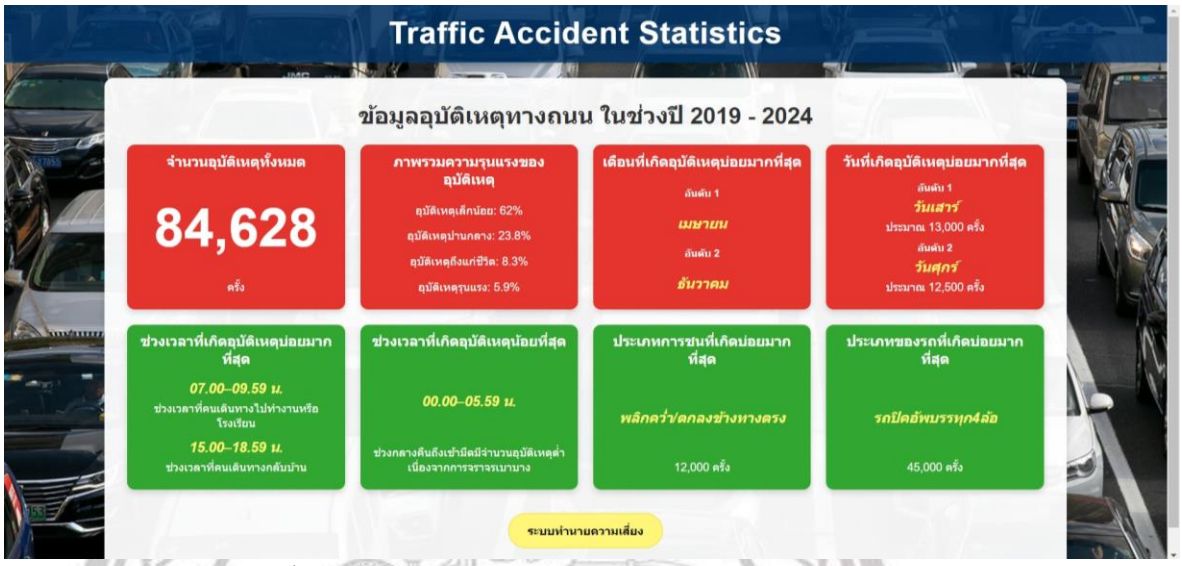
**Recall** Model 1 มี Recall สูงในหมวดหมู่ "อุบัติเหตุเล็กน้อย" ที่ 0.91 แสดงว่าโมเดลสามารถจำแนกประเภทนี้ได้ดี Model 3 มี Recall ต่ำมากในหมวดหมู่ "อุบัติเหตุถึงแก่ชีวิต" และ "อุบัติเหตุรุนแรง" (0.00)

**F1-Score** Model 1 มี F1-Score สูงสุดในหมวดหมู่ "อุบัติเหตุเล็กน้อย" ที่ 0.75 โมเดลที่มี F1-Score ต่ำสุดในหมวดหมู่ "อุบัติเหตุถึงแก่ชีวิต" คือ Model 3 ที่ 0.00

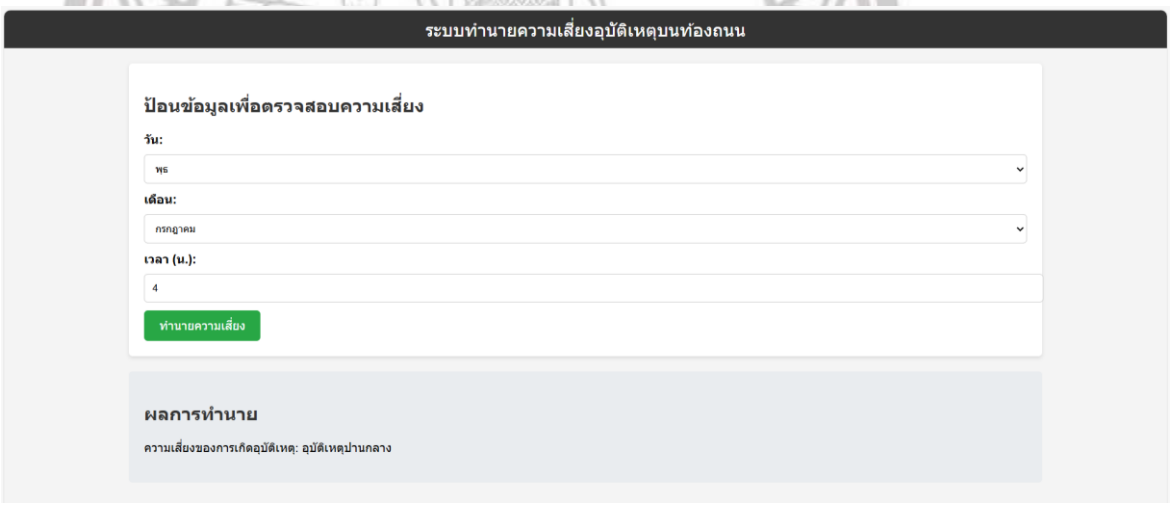
**Macro Average และ Weighted Average** Model 1 มี Macro Average ที่ดีที่สุดใน Precision และ F1-Score แต่ยังมี Recall ที่ต่ำ Model 2 มี Macro Average ต่ำกว่าโมเดลอื่นในเกือบทุกค่าที่วัดได้

Model 1 และ Model 3 มีประสิทธิภาพที่ดีในด้าน Accuracy แต่ Model 1 มี Precision และ Recall ที่ดีกว่า Model 3 Model 4 แสดงถึง Precision ที่สูงใน "อุบัติเหตุเล็กน้อย" เช่นเดียวกับ Model 1 แต่มี Performance ที่ลดลงในหมวดหมู่อื่น Model 2 มีประสิทธิภาพต่ำที่สุดในกลุ่มนี้ โดยเฉพาะในด้าน Accuracy และ Recall ของหมวดหมู่ที่สำคัญ

ลิขสิทธิ์ มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved




**ภาพ 18.** เว็บแอปพลิเคชันแสดงข้อมูลอุบัติเหตุทางถนนน ปี2019-2024




**ภาพ 19.** เว็บแอปพลิเคชันเพื่อทำนายความเสี่ยงของอุบัติเหตุบนท้องถนน

จากโมเดลที่เราได้ทำการทำนายจากทั้ง 4 โมเดล ได้เลือก 1 โมเดลในการมาทำ เว็บแอปพลิเคชันเพื่อทำนายความเสี่ยงของอุบัติเหตุบนท้องถนน นั้นก็คือโมเดลที่ 3 มีค่าความถูกต้องอยู่ที่ 61% โดยกำหนดตัวแปรต้น ได้แก่ วันในสัปดาห์ เดือน ชั่วโมงในวัน เป็นปัจจัยที่มีความสำคัญในการวิเคราะห์แนวโน้มการเกิดอุบัติเหตุ เนื่องจากการเกิดอุบัติเหตุมักมีความสัมพันธ์กับเวลา (เช่น เวลาที่มีการจราจรหนาแน่น) โดยไม่จำเป็นต้องพิจารณาตำแหน่งที่แน่นอน โดยเว็บแอปพลิเคชันจะทำนายด้วยการป้อนข้อมูล เช่น วันพุธ เดือน กรกฎาคม เวลาที่เราต้องการ แล้วกดทำนายผล เว็บแอปพลิเคชันจะทำนายผลออกมาว่า ความเสี่ยงของการเกิดอุบัติเหตุอยู่ในระดับไหน เช่น อุบัติเหตุที่มีความรุนแรงระดับปานกลาง

ลิขสิทธิ์มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved

## สรุปและอภิปรายผล

การวิเคราะห์และพัฒนาเว็บแอปพลิเคชันเพื่อทำนายความเสี่ยงของอุบัติเหตุบนท้องถนนด้วยการเรียนรู้ของเครื่อง งานวิจัยเกี่ยวกับการคาดการณ์อุบัติเหตุทางถนนด้วยการใช้เทคโนโลยีการเรียนรู้ของเครื่องจักร (Machine Learning) เป็นประเด็นสำคัญในยุคปัจจุบันที่การจราจรมีความซับซ้อนมากขึ้น สังคมต้องการเครื่องมือที่สามารถช่วยลดจำนวนอุบัติเหตุและผลกระทบที่เกิดจากอุบัติเหตุเหล่านั้น งานวิจัยของคุณเกี่ยวกับการพัฒนาและประเมินผลโมเดล Machine Learning เพื่อคาดการณ์ความรุนแรงของอุบัติเหตุมีความสำคัญอย่างยิ่ง เนื่องจากให้ข้อมูลเชิงลึกที่สามารถนำไปใช้ในเชิงปฏิบัติได้ โดยเฉพาะในระดับหน่วยงานที่รับผิดชอบด้านความปลอดภัยบนท้องถนน ในบทความนี้จะมีการอภิปรายผลลัพธ์ของงานวิจัย เปรียบเทียบกับงานวิจัยอื่น และเสนอแนะแนวทางเพื่อการพัฒนาเพิ่มเติม งานวิจัยของคุณได้ประเมินประสิทธิภาพของโมเดล Machine Learning ทั้งหมด 4 แบบ โดยใช้ชุดตัวชี้วัด ได้แก่ Accuracy, Precision, Recall, และ F1-Score ซึ่งตัวชี้วัดเหล่านี้ถูกแบ่งตามระดับความรุนแรงของอุบัติเหตุ (อุบัติเหตุถึงแก่ชีวิต, อุบัติเหตุปานกลาง, อุบัติเหตุรุนแรง, และอุบัติเหตุเล็กน้อย) ผลการทดลองที่สำคัญสรุปได้ดังนี้:

1. Accuracy: โมเดลที่ให้ค่าความแม่นยำ (Accuracy) สูงสุดคือ Model 1 และ Model 3 ซึ่งทั้งคู่มีค่าเท่ากับ 0.61 ในขณะที่ Model 2 ให้ค่า Accuracy ต่ำที่สุดคือ 0.44 แสดงให้เห็นว่า Model 1 และ Model 3 สามารถทำนายผลได้ในภาพรวมที่น่าพอใจ อย่างไรก็ตาม ค่า Accuracy เพียงอย่างเดียวไม่สามารถสะท้อนถึงความสามารถของโมเดลในการจัดการกับข้อมูลที่มีความไม่สมดุล (imbalanced data) ได้อย่างแท้จริง

2. Precision และ Recall:

- Precision และ Recall สำหรับกรณี "อุบัติเหตุเล็กน้อย" มีค่าสูงที่สุดในทุกโมเดล โดยเฉพาะใน Model 3 ซึ่งมีค่า Precision เท่ากับ 0.98 และ Recall เท่ากับ 0.62 แสดงให้เห็นว่าโมเดลสามารถทำนายกรณีที่มีความรุนแรงต่ำได้อย่างมีประสิทธิภาพ

- ในทางกลับกัน กลุ่ม "อุบัติเหตุถึงแก่ชีวิต" และ "อุบัติเหตุรุนแรง" มีค่า Precision และ Recall ต่ำมากในทุกโมเดล เช่น Model 1 มีค่า Recall เพียง 0.04 สำหรับอุบัติเหตุถึงแก่ชีวิต แสดงให้เห็นว่าโมเดลยังมีข้อจำกัดในการจัดการกับข้อมูลที่มีความไม่สมดุล

ลิขสิทธิ์มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved

3. F1-Score: ค่า F1-Score ของอุบัติเหตุที่มีความรุนแรงสูง เช่น อุบัติเหตุถึงแก่ชีวิต และอุบัติเหตุรุนแรง อยู่ในระดับต่ำมาก (0.07 และ 0.04 ใน Model 1) แสดงให้เห็นว่าโมเดลยังไม่สามารถทำนายกรณีที่มีความรุนแรงสูงได้ดีเท่าที่ควร

จากการเปรียบเทียบงานวิจัยนี้มุ่งเน้นการวิเคราะห์และพัฒนาเว็บแอปพลิเคชันเพื่อทำนายความเสี่ยงของอุบัติเหตุบนท้องถนนด้วยการเรียนรู้ของเครื่อง ข้อมูลอุบัติเหตุทางถนนในประเทศไทย โดยใช้ Random Forest เป็นโมเดลหลัก มีจุดที่สามารถเปรียบเทียบกับงานวิจัยเรื่อง ปัจจัยที่ส่งผลต่อการเกิดอุบัติเหตุทางถนน: การประเมินมุมมองของเจ้าหน้าที่บังคับใช้กฎหมาย ความคิดเห็นของผู้ขับขี่ทั่วไป และข้อมูลบันทึกอุบัติเหตุ ของ Jonathan J. Rolison et al. (2018) และ งานวิจัยเรื่อง การพยากรณ์ความเสี่ยงของอุบัติเหตุทางถนนโดยใช้ข้อมูลภูมิสารสนเทศและการเรียนรู้ของเครื่อง ของ Yunzhi Shi et al. (2021) ได้อย่างน่าสนใจในมิติของตัวแปรที่ใช้ศึกษา วิธีการพัฒนาระบบ และผลลัพธ์ที่ได้ ตัวแปรที่ศึกษา: ความคล้ายคลึงและความแตกต่าง ตัวแปรในงานวิจัยนี้มุ่งเน้นปัจจัยเชิงกายภาพและเชิงเวลา เช่น วัน เดือน เวลา และประเภทของยานพาหนะ ซึ่งสะท้อนถึงลักษณะเฉพาะของอุบัติเหตุในประเทศไทย ตัวแปรเหล่านี้ได้รับการคัดเลือกเพื่อวิเคราะห์และคาดการณ์ระดับความรุนแรงของอุบัติเหตุ ในทางกลับกันงานของ Jonathan J. Rolison et al. (2018) ใช้ตัวแปรที่มาจากมุมมองพฤติกรรมมนุษย์ เช่น ทัศนคติของผู้ขับขี่ต่อความเสี่ยง การละเมิดกฎจราจร และมุมมองของเจ้าหน้าที่บังคับใช้กฎหมาย รวมถึงข้อมูลอุบัติเหตุจริง ตัวแปรเหล่านี้เน้นการสำรวจปัจจัยเชิงพฤติกรรมและการรับรู้ ซึ่งต่างจากงานวิจัยนี้ที่เน้นปัจจัยเชิงสถิติและกายภาพ ในส่วนของงานวิจัยโดย Yunzhi Shi et al. (2021) ตัวแปรที่ใช้มุ่งเน้นข้อมูลเชิงพื้นที่ เช่น ตำแหน่งทางภูมิศาสตร์ ความหนาแน่นของการจราจร และโครงสร้างพื้นฐาน ตัวแปรเหล่านี้ช่วยวิเคราะห์พื้นที่เสี่ยงต่อการเกิดอุบัติเหตุ ข้อมูลที่ได้มีความละเอียดสูงและเชื่อมโยงกับมิติทางภูมิศาสตร์ ซึ่งมีความต่างกันที่ไม่ได้รวมข้อมูลเชิงพื้นที่เข้ามา การเปรียบเทียบ 1. ตัวแปรเชิงกายภาพ เช่น สภาพถนน มีความใกล้เคียงกับตัวแปรในงานของ Yunzhi Shi et al. (2021) แม้ว่าอีกฝ่ายจะเพิ่มมิติข้อมูลเชิงพื้นที่เพื่อวิเคราะห์พื้นที่เสี่ยง 2. ตัวแปรเชิงพฤติกรรมในงานของ Jonathan J. Rolison et al. (2018) ต่างจากวิจัยนี้อย่างสิ้นเชิง เพราะมุ่งเน้นการสำรวจพฤติกรรมมนุษย์และความคิดเห็น ซึ่งไม่ปรากฏในงานวิจัยนี้ แบบจำลองที่ใช้และประสิทธิภาพ งานวิจัยนี้ใช้ Random Forest ซึ่งมีความสามารถในการจัดการข้อมูลที่ซับซ้อนและหลากหลาย โมเดลนี้ให้ความสำคัญกับการวิเคราะห์ตัวแปรที่ส่งผลต่อความรุนแรงของอุบัติเหตุ โดยให้ความแม่นยำประมาณ 80% ซึ่งถือว่าเพียงพอสำหรับการคาดการณ์ในบริบทของข้อมูลขนาดใหญ่ ในขณะที่งานของ Jonathan J. Rolison et al. (2018) ไม่ได้ใช้แบบจำลอง Machine Learning แต่ใช้การวิเคราะห์เชิงสถิติเพื่อเปรียบเทียบข้อมูลอุบัติเหตุจริงกับความคิดเห็น

ลิขสิทธิ์ มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved

ของผู้ขับขี่และเจ้าหน้าที่ ผลลัพธ์ช่วยให้เข้าใจปัจจัยเชิงพฤติกรรมและการรับรู้ แต่ไม่ได้มุ่งเน้นไปที่การพยากรณ์ สำหรับงานของ Yunzhi Shi et al. (2021) ใช้โมเดลที่ซับซ้อนกว่า เช่น Gradient Boosting และ Neural Networks ซึ่งเหมาะกับข้อมูลเชิงพื้นที่และการวิเคราะห์ความสัมพันธ์เชิงซ้อน โมเดลเหล่านี้ให้ความแม่นยำสูงกว่า 90% ในบางชุดข้อมูล อย่างไรก็ตามความซับซ้อนในการพัฒนาและการตีความผลลัพธ์อาจเป็นข้อจำกัดในการใช้งานจริง การเปรียบเทียบ 1. โมเดล Random Forest ในงานนี้มีข้อได้เปรียบในด้านความเรียบง่ายและความง่ายต่อการปรับใช้ ทำให้เหมาะสำหรับการคาดการณ์ในบริบทของข้อมูลขนาดใหญ่ 2. โมเดล Neural Networks ในงานของ Yunzhi Shi et al. (2021) มีความแม่นยำสูงกว่า แต่มีความซับซ้อนในการพัฒนา 3. งานของ Jonathan J. Rolison et al. (2018) เน้นการวิเคราะห์เชิงสถิติแบบดั้งเดิม และไม่ได้เปรียบเทียบในแง่ของประสิทธิภาพโมเดล Machine Learning บทสรุปของการเปรียบเทียบ การเปรียบเทียบงานวิจัยนี้กับงานวิจัยอื่นสะท้อนให้เห็นถึงแนวทางที่แตกต่างกันในการศึกษาอุบัติเหตุทางถนน ตัวแปรในงานของวิจัยนี้มุ่งเน้นปัจจัยเชิงสถิติและกายภาพ ซึ่งเหมาะกับบริบทของข้อมูลในประเทศไทย ในขณะที่งานของ Jonathan J. Rolison et al. (2018) ใช้ตัวแปรเชิงพฤติกรรม และงานของ Yunzhi Shi et al. (2021) ใช้ตัวแปรเชิงพื้นที่ ในแง่ของแบบจำลอง Random Forest ที่ใช้ในงานวิจัยมีข้อได้เปรียบในด้านความง่ายต่อการใช้งานและการตีความผลลัพธ์ แม้ว่าความแม่นยำอาจต่ำกว่าการใช้ Neural Networks หรือ Gradient Boosting ก็ตาม อย่างไรก็ตาม การบูรณาการตัวแปรหลายมิติ เช่น การเพิ่มข้อมูลเชิงพฤติกรรมหรือเชิงพื้นที่ อาจช่วยเพิ่มความแม่นยำและประสิทธิภาพของการพยากรณ์ในอนาคต ข้อเสนอแนะสำหรับการพัฒนาในอนาคต

**ข้อเสนอแนะเพื่อการพัฒนาในอนาคต**

1. การใช้แบบจำลอง Machine Learning อื่นๆ การพัฒนางานวิจัยในอนาคตควรพิจารณาเปรียบเทียบผลลัพธ์ของแบบจำลอง Random Forest กับแบบจำลอง Machine Learning อื่นๆ ที่มีศักยภาพสูง เช่น Neural Network ซึ่งเป็นโมเดลที่สามารถจับความสัมพันธ์ที่ซับซ้อนระหว่างปัจจัยต่างๆ ได้อย่างมีประสิทธิภาพ โดยเฉพาะอย่างยิ่งในกรณีที่มีข้อมูลจำนวนมากและข้อมูลมีความหลากหลาย Neural Network มีความสามารถในการประมวลผลเชิงลึก เช่น การวิเคราะห์ปัจจัยที่ไม่เป็นเชิงเส้น (Non-linear Relationships) และการประยุกต์ใช้โครงข่ายประสาทเทียมแบบหลายชั้น (Deep Learning) อาจช่วยให้ผลลัพธ์ของการทำนายมีความแม่นยำและครอบคลุมมากยิ่งขึ้น

ลิขสิทธิ์ มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved

2. เพิ่มปัจจัยสนับสนุน รวมข้อมูลที่เกี่ยวข้อง เช่น สภาพถนน (เช่น ทางลาดชัน หรือโค้งอันตราย) สภาพอากาศ (เช่น ฝนตก หรือหมอกหนา) และพฤติกรรมของคนขับ (เช่น การใช้ความเร็วเกินกำหนด หรือการขับขี่ขณะมึนเมา) เพื่อเพิ่มความแม่นยำในการคาดการณ์

3. การทดสอบโมเดลเพิ่มเติม ทดลองใช้โมเดลที่จัดการกับข้อมูลไม่สมดุลได้ดี เช่น XGBoost, LightGBM หรือการผสมผสานเทคนิค Ensemble Learning ซึ่งอาจช่วยเพิ่มค่าตัวชี้วัดหลัก เช่น Recall และ F1-Score

งานวิจัยนี้เป็นก้าวสำคัญในการนำเทคโนโลยี Machine Learning มาประยุกต์ใช้กับปัญหาที่มีความสำคัญในชีวิตประจำวัน ผลลัพธ์ที่ได้แม้จะมีข้อจำกัด แต่สามารถเป็นพื้นฐานสำหรับการพัฒนาเครื่องมือช่วยตัดสินใจในเชิงปฏิบัติ ตัวอย่างเช่น การใช้ข้อมูลที่ได้เพื่อวางแผนการจัดการทรัพยากรด้านความปลอดภัย เช่น การจัดเจ้าหน้าที่หรือป้ายเตือนในพื้นที่ที่มีความเสี่ยงสูง การพัฒนาแอปพลิเคชันที่ให้คำแนะนำแก่ผู้ขับขี่เพื่อหลีกเลี่ยงความเสี่ยง ในอนาคต การวิจัยนี้สามารถนำไปสู่การพัฒนาระบบที่สามารถใช้งานได้ในระดับพื้นที่หรือระดับประเทศ โดยการใช้เทคโนโลยีที่ทันสมัย และการปรับปรุงข้อมูลที่ใช้ในการวิจัยให้ครอบคลุมและสมบูรณ์ยิ่งขึ้น

ลิขสิทธิ์ มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved

**บรรณานุกรม**


Behram Wali., Asad J. Khattak., Thomas Karnowski. (2020). The relationship between driving volatility in time to collision and crash-injury severity in a naturalistic driving environment. สืบค้น 10 สิงหาคม 2567 จาก https://www.sciencedirect.com/science/article/pii/S2213665720300269

Jonathan J. Rolison., Shirley Regev., Salissou Moutari., Aidan Feeney. (2018). What are the factors that contribute to road accidents? An assessment of law enforcement views, ordinary drivers' opinions, androad accident records. สืบค้น 10 สิงหาคม 2567 จาก https://www.sciencedirect.com/science/article/pii/S0001457518300873

Shakil Ahmed., Md Akbar Hossain., Sayan Kumar Ray., Md Mafijul Islam Bhuiyan., Saifur Rahman Sabuj. (2023). A study on road accident prediction and contributing factors using explainable machine learning models: analysis and performance. สืบค้น 10 สิงหาคม 2567 จาก https://www.sciencedirect.com/science/article/pii/S2590198223000611

Tian-Jing Sun., Si-Jia Liu ., Fang-Ke Xie ., Xiao-Fei Huang ., Jian-Xiu Tao ., Yuan-Lan Lu ., Tian-Xi Zhang ., An-Yong Yu. (2024) Influence of road types on road traffic accidents in northern Guizhou Province, China. สืบค้น 10 สิงหาคม 2567 จาก https://www.sciencedirect.com/science/article/pii/S1008127520302029

Yunzhi Shi's., Raj Biswas., Mehdi Noori., Michael Kilberry. (2021). Predicting Road Accident Risk Using Geospatial Data and Machine Learning (Demo Paper). สืบค้น 10 สิงหาคม 2567 จาก https://www.researchgate.net/publication/355918225_Predicting_Road_Accident_Risk_Using_Geospatial_Data_and_Machine_Learning_Demo_Paper



ลิขสิทธิ์ มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved

**ประวัติผู้วิจัย**

**ชื่อ – สกุล** สุวิมล ชื่นทิม

**วัน เดือน ปี เกิด** 27 กรกฎาคม 2545

**ที่อยู่ปัจจุบัน** 66/1 ม.3 ตำบลท่านางงาม อำเภอบางระกำ
จังหวัดพิษณุโลก 65140

**ประวัติการศึกษา**

พ.ศ. 2564 – ปัจจุบัน วท.บ.(ภูมิศาสตร์) มหาวิทยาลัยนเรศวร เกรดเฉลี่ย 2.96

พ.ศ. 2561 – 2563 ระดับมัธยมศึกษาตอนปลาย (ศิลป์คำนวณ)
โรงเรียนเตรียมอุดมศึกษา ภาคเหนือ ตำบลหัวรอ
อำเภอเมืองพิษณุโลก จังหวัดพิษณุโลก 65000 เกรดเฉลี่ย 3.45

พ.ศ.2558 – 2560 ระดับมัธยมศึกษาตอนต้น โรงเรียนเซนต์นิโกลาส
228 วิสุทธิกษัตริย์ ตำบลในเมือง อำเภอเมืองพิษณุโลก
จังหวัดพิษณุโลก 65000

พ.ศ.2552 – 2557 ระดับประถมศึกษา โรงเรียนบ้านบางแก้ว ตำบลท่านางงาม
อำเภอบางระกำ จังหวัดพิษณุโลก 65140

ลิขสิทธิ์ มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved

**การเข้าร่วมกิจกรรม**

1) โครงการส่งเสริมคุณธรรมจริยธรรมและเชิดชูความสำเร็จ คณะเกษตรศาสตร์ทรัพยากรธรรมชาติและสิ่งแวดล้อม มหาวิทยาลัยนเรศวร วันที่ 26 มกราคม พ.ศ.2566

2) โครงการพัฒนาทักษะการเรียนรู้ตลอดชีวิต ที่อุทยานแห่งชาติภูหินร่องกล้า วันที่ 10-11 กุมภาพันธ์ พ.ศ.2566

3) กิจกรรมบูรณาการวัฒนธรรมและภูมิปัญญาท้องถิ่นด้านการเกษตรกับองค์ความรู้เทคโนโลยีสมัยใหม่ คณะเกษตรศาสตร์ฯ มหาวิทยาลัยนเรศวร วันที่ 17 สิงหาคม พ.ศ.2566

**รางวัลที่ได้รับ**

1) ได้รับรางวัล”ดีเด่น” ประเภทการจัด visual hierarchy and layout โครงการประกวดแผนที่แห่งประเทศไทย ครั้งที่ 1

**ผลงาน**

ผลงาน ช่องยูทูปเพื่อความรู้เกี่ยวกับรายวิชาภูมิศาสตร์ สารสนเทศภูมิศาสตร์และการประยุกต์ใช้ช่อง “Anne” โดยจัดทำสื่อการสอน YouTube ดังนี้

1) การประมวลผลภาพถ่ายอากาศยานไร้คนขับ โดยโปรแกรม Agisoft Metashape พื้นที่ศึกษา คณะเกษตรศาสตร์ มหาวิทยาลัยนเรศวร

2) กระบวนการโฟโตแกรมเมตรีเชิงเลขจากข้อมูลภาพถ่ายทางอากาศ Frame camera โดยโปรแกรม ERDAS IMAGINE

3) การประมวลผลภาพถ่ายอากาศยานไร้คนขับ โดยโปรแกรม Agisoft Metashape พื้นที่ศึกษา เหมือง แม่เมาะ จ.ลำปาง

4) การจำแนกสิ่งปกคลุมดินจากภาพดาวเทียม Sentinel-2 โดยโปรแกรม SNAP พื้นที่ศึกษา ตำบลวัด พริกและตำบลท่าโพธิ์

5) การจำแนกสิ่งปกคลุมดิน Unspervised Classification โดยโปรแกรม SNAP

6) การตรวจสอบความถูกต้อง Supervised Classification โดยโปรแกรม SNAP

7) การประเมินความรุนแรงของพื้นที่เผาไหม้จากไฟป่า โดยโปรแกรม SNAP พื้นที่ศึกษา อุทยานแห่งชาติภูกระดึง จ.เลย

8) สาธิตการสอนใช้งาน Google Earth Engine เบื้องต้น

9) Mini Project การประเมินความเสียหายจากอุทกภัย กรณีศึกษาอำเภอบางระกำ จังหัดพิษณุโลก ด้วย Google Earth Engine

ลิขสิทธิ์ มหาวิทยาลัยนเรศวร
Copyright by Naresuan University
All rights reserved