

## บทที่ 1

### บทนำ

#### 1.1 ความสำคัญและที่มาของปัญหา

ในยุคปัจจุบันเทคโนโลยีดิจิทัลเข้ามาเป็นส่วนหนึ่งในการดำรงชีวิตประจำวันของคนเราเป็นอย่างมาก หน่วยงานต่าง ๆ ไม่ว่าจะเป็นภาครัฐกิจ องค์กรต่าง ๆ รวมถึงหน่วยงานภาครัฐเองต่างก็มีการนำเทคโนโลยีต่าง ๆ เข้ามาช่วยเพื่อปรับปรุงประสิทธิภาพการดำเนินงานและยกระดับการให้บริการไปสู่ประชาชนให้สามารถเข้าถึงบริการต่าง ๆ ได้สะดวก และรวดเร็วยิ่งขึ้น จึงทำให้เกิดข้อมูลสารสนเทศขึ้นอย่างมากมายมหาศาล โดยเฉพาะอย่างยิ่งข้อมูลของภาครัฐที่จัดเก็บในรูปแบบของฐานข้อมูลนับเป็นเวลาหลายสิบปี และกำลังเพิ่มขึ้นอีกในปัจจุบันและอนาคต ข้อมูลเหล่านี้สามารถนำมาใช้ประโยชน์ได้ทั้งสำหรับภาครัฐ ภาคเอกชน และประชาชนซึ่งทำให้เกิดเป็นแนวคิด วิธีการ รูปแบบทางธุรกิจ รวมถึงวิถีชีวิตใหม่ๆ การแบ่งปันข้อมูลเพื่อนำมาใช้ประโยชน์และสร้างนวัตกรรมนั้น ถือเป็นสัญญาณเริ่มต้นของการขับเคลื่อนเศรษฐกิจด้วยข้อมูล อันจะช่วยให้เศรษฐกิจของประเทศเติบโตขึ้น และเพิ่มขีดความสามารถในการแข่งขันในเวทีโลกได้ ที่ผ่านมามีหลาย ๆ ประเทศมีการผลักดัน และดำเนินการเกี่ยวกับการเปิดเผยข้อมูลสู่สาธารณะกันอย่างจริงจัง เช่นเดียวกันกับในประเทศไทย หลายๆ หน่วยงานก็เริ่มเห็นถึงความสำคัญ และประโยชน์ของการเปิดเผยข้อมูลกันมากขึ้น ดังนั้นสำนักงานรัฐบาลอิเล็กทรอนิกส์ (องค์การมหาชน) (สรอ.) ซึ่งมีภารกิจหลักในการพัฒนารัฐบาลอิเล็กทรอนิกส์ของประเทศให้มีคุณภาพ จึงได้พัฒนาระบบศูนย์กลางข้อมูลภาครัฐ หรือ data.go.th เพื่อเป็นการส่งเสริมและผลักดันให้เกิดการเปิดเผยข้อมูลภาครัฐสู่สาธารณะอย่างเป็นรูปธรรม เป็นช่องทางให้ผู้ใช้บริการทั้งภาคประชาชน ภาคธุรกิจ เอกชน รวมถึงหน่วยงานของรัฐ สามารถค้นหาและเข้าถึงข้อมูลที่มีคุณภาพของภาครัฐได้ง่าย โดยมีเป้าหมายสูงสุดเพื่อส่งเสริมให้เกิดธรรมาภิบาล ส่งเสริมให้เกิดความร่วมมือของภาครัฐและสร้างการมีส่วนร่วมของภาคประชาชน รวมถึงแก้ปัญหาการทุจริตคอร์รัปชัน การเมือง และเศรษฐกิจอย่างบูรณาการร่วมกัน ได้ศึกษาปัจจัยที่ส่งผลต่อการเกิดอุบัติเหตุบนท้องถนน และ

เปรียบเทียบประสิทธิภาพการจำแนกข้อมูล คือ การนำข้อมูลที่ได้ถูกรวบรวมไว้และถูกเผยแพร่ผ่านเว็บไซต์ Open Government Data of Thailand ซึ่งเป็นเว็บไซต์ที่เก็บรวบรวมชุดข้อมูลต่าง ๆ ของประเทศไทย เป็นข้อมูลที่เปิดเผยได้ เพื่อให้ผู้ใช้บริการสามารถนำชุดข้อมูลไปศึกษาหรือวิเคราะห์ให้เกิดประโยชน์ต่อไปได้ การวิเคราะห์ข้อมูลปริมาณมหาศาลแสดงให้เห็นประโยชน์ต่าง ๆ ก็คือ ความรวดเร็วและประสิทธิภาพ ขณะที่หลายปีก่อนหน้านั้นหลายธุรกิจต้องรวบรวมข้อมูลดำเนินการวิเคราะห์ และค้นหาข้อมูลที่สามารถนำมาใช้ประกอบการตัดสินใจในอนาคตได้ ส่วนในทุกวันนี้ธุรกิจสามารถระบุข้อมูลเชิงลึกสำหรับการกระทำที่ต้องการการตัดสินใจอย่างทันท่วงทีความสามารถในการทำงานได้เร็วขึ้น และยังคงไว้ซึ่งความคล่องตัว ช่วยให้หลายองค์กรสร้างความได้เปรียบด้านการแข่งขันอย่างไม่เคยมีมาก่อน

อุบัติเหตุทางถนนเป็นปัญหาที่มีผลกระทบอย่างมากต่อชีวิตและทรัพย์สินของประชาชนทั่วโลก เป็นสาเหตุหลักของการบาดเจ็บและการเสียชีวิต ซึ่งส่งผลให้เกิดความสูญเสียทั้งในด้านมนุษย์และเศรษฐกิจ อุบัติเหตุเหล่านี้ไม่เพียงแต่ส่งผลกระทบต่อผู้ที่เกี่ยวข้องโดยตรง แต่ยังสร้างความเครียดและภาระให้กับครอบครัวและสังคมโดยรวม ด้วยเหตุนี้ จึงมีการนำข้อมูลปัจจัยการเกิดอุบัติเหตุ ประจำปี 2562-2567 มาทำการวิเคราะห์ เนื่องจากต้องการทราบปัจจัยที่ทำให้เกิดอุบัติเหตุที่นำไปสู่การบาดเจ็บและเสียชีวิต ซึ่งการวิเคราะห์ ปัจจัยต่าง ๆ จากข้อมูลที่ได้มานั้นถือเป็นเรื่องสำคัญต่อการป้องกันภัยจากอุบัติเหตุต่าง ๆ เพื่อหาทางป้องกันและลดอัตราการบาดเจ็บและเสียชีวิต แต่ด้วยจำนวนข้อมูลที่มีปริมาณมหาศาลทำให้ยากต่อการนำเสนอให้เข้าใจได้ง่าย ซึ่งถือเป็นข้อจำกัดหากต้องการนำข้อมูลชุดนี้ไปใช้เพื่อเผยแพร่ให้แก่บุคคลภายนอกหรือกลุ่มผู้ใช้ข้อมูลได้รับรู้

จากปัญหาข้างต้น คณะผู้จัดทำจึงได้นำชุดข้อมูลที่ได้มานั้นเข้าสู่กระบวนการจำแนกตามรายการชุดข้อมูล โดยการใช้ขีดความสามารถในการประมวลผลของโปรแกรมคอมพิวเตอร์ เพื่อเพิ่มความน่าเชื่อถือในการวิเคราะห์ และประมวลผลข้อมูลให้เหมาะสมตรงกับความต้องการเป็นการแยกแยะสิ่งที่จะพิจารณาออกเป็นส่วนย่อยที่มีความสัมพันธ์กัน เพื่อทำความเข้าใจแต่ละส่วนให้ชัดเจน รวมทั้งการสืบค้นความสัมพันธ์ของส่วนต่าง ๆ เพื่อดูว่าส่วนประกอบปลีกย่อยนั้นสามารถเข้ากันได้หรือไม่ มีความสัมพันธ์เกี่ยวเนื่องกันอย่างไร โดยคณะผู้จัดทำได้ใช้เทคนิคการทำเหมืองข้อมูลในการศึกษาปัจจัยที่ส่งผลต่อการเกิดอุบัติเหตุบนท้องถนน และเปรียบเทียบ

ประสิทธิภาพตัวแบบการจำแนกข้อมูล ดังนี้ (1) เทคนิคที่ใช้วิเคราะห์ปัจจัยการเกิดอุบัติเหตุ ได้แก่ เทคนิค Chi-Square ใช้ประเมินค่าของคุณลักษณะ เพื่อศึกษาว่าการแจกแจงความถี่ของตัวแปรคุณลักษณะเป็นไปตามรูปแบบที่กำหนดไว้หรือไม่ , เทคนิค Correlation Based Feature Selection (CFS) ใช้เลือกคุณสมบัติของคุณลักษณะโดยใช้การพิจารณาบนพื้นฐานความสัมพันธ์ของกลุ่มคุณลักษณะที่ได้จากการประเมินค่าความสามารถในการคาดการณ์ และยังสามารถจัดการกับคุณลักษณะที่ไม่เกี่ยวข้องได้ และ เทคนิค Information Gain เป็นเทคนิคการเลือกคุณลักษณะเพื่อประเมินค่าในการแบ่งข้อมูล และ (2) เทคนิคการจำแนกข้อมูล ได้แก่ เทคนิค Decision Tree แบบจำลองที่ใช้จำแนกประเภทข้อมูลหรือจำแนกกฎ , เทคนิค Random Forest เป็นการสร้างแบบจำลองด้วยวิธีการต้นไม้ตัดสินใจขึ้นมาหลายๆแบบจำลอง โดยวิธีการสุ่มตัวแปร แล้วนำผลที่ได้แต่ละแบบจำลองมารวมกันพร้อมนับจำนวนผลที่มีจำนวนซ้ำกันมากที่สุด , เทคนิค Voting เป็นการนำข้อมูลการเรียนรู้ชุดเดียวกันให้โมเดลแต่ละโมเดลทำนาย แล้วโหวตว่าคลาสคำตอบใดที่แต่ละโมเดลทำนายผลมากที่สุด, เทคนิค Naïve Bayes เป็นการจำแนกข้อมูลที่อาศัยหลักการความน่าจะเป็นตามทฤษฎีของเบย์(Bayes's theorem) โดยเรียนรู้ปัญหาที่เกิดขึ้นแล้วนำมาสร้างเงื่อนไขการจำแนกข้อมูลใหม่ , เทคนิค Gradient Boosting เป็นการปรับปรุงประสิทธิภาพของแบบจำลองต้นไม้ตัดสินใจให้มีประสิทธิภาพมากขึ้น โดยการสุ่มสร้างต้นไม้ตัดสินใจหลายสิบต้นจนถึงหลายร้อยต้น และทำการประเมินผลต้นไม้แต่ละต้นจนกว่าจะได้ต้นไม้ตัดสินใจที่มีประสิทธิภาพมากที่สุด และเทคนิค Deep Learning เป็นการเรียนรู้ด้วยการเลียนแบบการทำงานของโครงข่ายประสาทมนุษย์ และเรียนรู้ข้อมูลตัวอย่างข้อมูลดังกล่าวถูกนำไปใช้จัดหมวดหมู่ข้อมูลและทางคณะผู้จัดทำยังได้ทำการนำเสนอข้อมูลแบบ Visualization ด้วยการแสดงผลข้อมูลในรูปแบบของภาพด้วยโปรแกรม Power BI และเผยแพร่ข้อมูลสารสนเทศนี้บน Web Browser ที่เป็นที่ยอมรับในยุคอินเทอร์เน็ตคือ การเผยแพร่ทางสื่อออนไลน์ โดยใช้ภาษา HTML ในการพัฒนาเว็บไซต์สำหรับการเปิดเผยข้อมูลภาครัฐที่ได้มาจาก data.go.th หรือใน Open Government Data of Thailand เพื่อให้ผู้ใช้งานสามารถเข้าถึงข้อมูลได้อย่างเหมาะสม ซึ่งประโยชน์ที่ผู้ใช้งานทั่วไปจะได้รับจะเป็นในเรื่องของความเข้าใจ การรับรู้ และการตระหนักถึงความปลอดภัยในเหตุการณ์ต่าง ๆ สำหรับผู้ที่ทำการค้นหาและศึกษาข้อมูลก็สามารถดำเนินการผ่าน Web Browser ที่ได้เผยแพร่ข้อมูลสารสนเทศของผู้บาดเจ็บและเสียชีวิตประจำปี 2562-2567 ที่ผ่านการสรุปจำแนกข้อมูลตามประเภทต่าง ๆ ได้สะดวกรวดเร็วยิ่งขึ้นรวมถึงมีความถูกต้อง สะดวก ลดความซ้ำซ้อน เพิ่มประสิทธิภาพให้กับการศึกษาค้นคว้า และทันต่อเวลา

## 1.2 วัตถุประสงค์

1. เพื่อศึกษาปัจจัยที่ส่งผลต่อการเกิดอุบัติเหตุ บนท้องถนน
2. เพื่อเปรียบเทียบประสิทธิภาพการจำแนกข้อมูลบนชุดข้อมูลการเกิดอุบัติเหตุบนท้องถนน

## 1.3 ประโยชน์ที่จะได้รับ

1. ได้ผลการศึกษาปัจจัยที่ส่งผลต่อการเกิดอุบัติเหตุบนท้องถนน
2. ได้ผลการเปรียบเทียบประสิทธิภาพการจำแนกข้อมูลบนชุดข้อมูลการเกิดอุบัติเหตุบนท้องถนน

## 1.4 ขอบเขต

### 1.4.1 ขอบเขตการวิเคราะห์ข้อมูล

1.4.1.1 เก็บรวบรวมข้อมูล (Data Integration) จากข้อมูลอุบัติเหตุบนท้องถนนตั้งแต่ปี 2562 ถึง 2567

1.4.1.2 ทำการคัดเลือก (Data Selection) โดยเลือกคอลัมน์ที่มีความเชื่อมโยงกันและสามารถนำมาสร้างโมเดลต่างๆได้

1.4.1.3 ทำการทำความสะอาดข้อมูล (Data Cleaning) โดยการตรวจสอบข้อมูล เปลี่ยนแปลงข้อมูลให้ได้ข้อมูลที่ถูกต้องและสมบูรณ์

1.4.1.4 แปลงข้อมูลให้อยู่ในรูปแบบเพื่อใช้ในการวิเคราะห์ ให้นำข้อมูลมีความสมบูรณ์เพื่อนำมา แปลงข้อมูลให้อยู่ในรูปแบบข้อมูลที่สามารถนำมาวิเคราะห์โดยเทคนิคการทำเหมืองข้อมูล

1.4.1.5 ทำเหมืองข้อมูล (Data Mining Engine) ใช้เทคนิคการทำเหมืองข้อมูลในการศึกษาปัจจัยที่ส่งผลต่อการเกิดอุบัติเหตุบนท้องถนนและเปรียบเทียบประสิทธิภาพตัวแบบการจำแนกข้อมูล ได้แก่ เทคนิค Chi-Square , เทคนิค Correlation Based Feature Selection (CFS) และ เทคนิค Information gain, เทคนิค Decision Tree ,เทคนิค Random Forest , เทคนิค Voting , เทคนิค Naïve Bayes , เทคนิค Gradient Boosting และเทคนิค Deep Learning

1.4.1.6 ประเมินรูปแบบ (Pattern Evaluation) ประเมินผลประสิทธิภาพของโมเดล ด้วยตาราง Confusion Matrix กำหนดให้ค่าที่ยอมรับและดีที่สุดคือ 80

1.4.1.7 นำเสนอข้อมูลภาพ (Data Visualization) นำเสนอและเผยแพร่ข้อมูลผ่านเว็บไซต์ โดยเว็บไซต์จะพัฒนาโดยใช้ภาษา HTML และชุดคำสั่ง CSS เพื่อนำเข้า วิเคราะห์ข้อมูลด้วย โมเดลที่เลือกใช้และผ่านการทดสอบประสิทธิภาพของการวิเคราะห์ข้อมูลแล้ว และนำเสนอ ข้อมูลแบบ Visualization สำหรับแสดงผลข้อมูลในรูปแบบภาพจากข้อมูลที่ได้ผ่านการวิเคราะห์ แล้วภายในเว็บไซต์

#### 1.4.2 ขอบเขตผู้ใช้งานทั่วไปบนเว็บเบราว์เซอร์

1.4.2.1 สามารถดูข้อมูลสารสนเทศผู้บาดเจ็บเผยแพร่ผ่านทางเว็บไซต์

- สารสนเทศสถิติผู้บาดเจ็บแต่ละจังหวัด
- สารสนเทศสถิติจังหวัดที่มีผู้บาดเจ็บสูงที่สุด

1.4.2.2 สามารถดูข้อมูลสารสนเทศผู้เสียชีวิตเผยแพร่ผ่านทางเว็บไซต์

- สารสนเทศสถิติผู้เสียชีวิตแต่ละจังหวัด
- สารสนเทศสถิติจังหวัดที่มีผู้เสียชีวิตสูงที่สุด

1.4.2.3 สามารถดูข้อมูลสารสนเทศปัจจัยที่มีผลต่อการเกิดอุบัติเหตุ

1.4.2.4 สามารถดูผลการเปรียบเทียบประสิทธิภาพการจำแนกข้อมูล

1.4.2.5 สามารถกรองข้อมูลสารสนเทศผ่านระบบของ Power BI บน Web browser ได้

### 1.5 เครื่องมือที่ใช้ในการพัฒนาโปรแกรม

1.5.1 Hardware

1.5.1.1 เครื่องคอมพิวเตอร์โน้ตบุ๊กรุ่น Lenovo IdeaPad S540-14API

1.5.2 Software

1.5.2.1 โปรแกรม RapidMiner Studio ใช้ในการวิเคราะห์ข้อมูล

1.5.2.2 โปรแกรม Xampp ใช้ในการจำลอง web server

1.5.2.3 โปรแกรม FileZilla โปรแกรมที่ใช้ติดต่อกับ FTP Server เพื่ออัปโหลด (Upload) ไฟล์ (File)

1.5.2.4 โปรแกรม Microsoft word ใช้จัดทำเอกสารประกอบโครงการงาน

1.5.2.5 โปรแกรม Microsoft excel ใช้เตรียมข้อมูล

1.5.2.6 โปรแกรม Power BI ใช้นำเสนอข้อมูลรูปแบบของภาพ

1.5.2.7 โปรแกรม Sublime Text 3 ใช้เขียนเขียนโค้ด HTML และ CSS

## 1.6 สถานที่ใช้ในการดำเนินการศึกษาและรวบรวมข้อมูล

- มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนา
- แหล่งรวบรวมข้อมูลเว็บไซต์ Open Government Data of Thailand หรือ data.go.th

## 1.7 ระยะเวลาในการดำเนินการ

แผนการดำเนินงาน	2567						2568		
	ก.ค	ส.ค	ก.ย	ต.ค	พ.ย	ธ.ค	ม.ค	ก.พ	มี.ค
1.ศึกษาและกำหนดความต้องการของข้อมูล	↔								
2.ตรวจสอบความถูกต้องของข้อมูล		↔							
3.วิเคราะห์ข้อมูลจำแนกตามกลุ่ม		↔							
4.ออกแบบระบบ			↔						
5.เขียนและทดสอบระบบ				↔	↔				
6.ตรวจสอบระบบโดยรวม					↔				
7.ประเมินการใช้งานระบบ						↔	↔		
8.จัดทำเอกสารประกอบโครงการ			↔	↔	↔	↔	↔	↔	↔

## 1.8 บทสรุป

จากบทหน้าที่ได้กล่าวมาในข้างต้นทั้งหมดนั้น คณะผู้จัดทำได้เล็งเห็นความสำคัญของการนำเทคโนโลยีมาประยุกต์ใช้ในการวิเคราะห์ข้อมูลผู้บาดเจ็บและเสียชีวิตบนท้องถนนประจำปี 2562-2567 ด้วยกระบวนการวิเคราะห์ข้อมูลด้วย CRISP-DM จากเทคนิคการทำเหมืองข้อมูลแบบ Classification ในรูปแบบของแผนภูมิต้นไม้ หรือที่เรียกว่า Decision Tree และเผยแพร่ข้อมูลสารสนเทศบน Web Browser โดยใช้ภาษา HTML และ CSS ในการพัฒนาหน้า

เว็บไซต์สำหรับการเปิดเผยข้อมูลภาครัฐจากเว็บไซต์ data.go.th หรือใน Open Government Data of Thailand เนื่องจากเทคโนโลยีปัจจุบันเข้ามาช่วยในการจัดการข้อมูลหรือดำเนินงานต่างๆ ให้มีความสะดวกสบายเป็นอย่างมากอีกทั้งผู้ใช้งานยังเข้าถึงข้อมูลได้ง่ายและมีประสิทธิภาพ