FM-BIS-01

แบบข้อเสนอโครงงาน

หลักสูตรระบบสารสนเทศทางธุรกิจ

สาขาบริหารธุรกิจ ปีการศึกษา 2567

1. ชื่อโครงงาน

การศึกษาปัจจัยที่ส่งผลต่อการเกิดอุบัติเหตุบนท้องถนน และการเปรียบเทียบประสิทธิภาพตัวแบบการจำแนกข้อมูล บนชุดข้อมูลอุบัติเหตุบนโครงข่ายถนนของกระทรวงคมนาคม ปี 2562 – 2567

Study of Factors Affecting Road Accidents and Comparison of Data Classification Efficiency on the Dataset of Accidents on the Ministry of Transport's Road Network, years 2019 – 2024

2. ผู้จัดทำ

น.ส. ธิดา กิจสวัสดิ์ไพบูลย์ รหัสนักศึกษา 66541207065-8

น.ส. เนตรนภา เมืองมาหล้า รหัสนักศึกษา 66541207069-0

3. อาจารย์ที่ปรึกษาโครงการ

อาจารย์วรการ ใจดี

4. ความสำคัญและที่มาของปัญหา

ในยุคปัจจุบันเทคโนโลยีดิจิทัลเข้ามาเป็นส่วนหนึ่งในการดำรงชีวิตประจำวันของคนเราเป็นอย่างมาก หน่วยงานต่าง ๆ ไม่ว่าจะเป็นภาคธุรกิจ องค์กรต่าง ๆ รวมถึงหน่วยงานภาครัฐเอง ต่างก็มีการนำเทคโนโลยีต่าง ๆ เข้ามาช่วยเพื่อปรับปรุงประสิทธิภาพการดำเนินงานและยกระดับการให้บริการไปสู่ประชาชนให้สามารถเขาถึงบริการต่าง ๆ ได้สะดวก และรวดเร็วยิ่งขึ้น จึงทำให้เกิดข้อมูลสารสนเทศขึ้นอย่างมากมายมหาศาล โดยเฉพาะอย่างยิ่งข้อมูลของภาครัฐที่จัดเก็บในรูปแบบของฐานข้อมูลนับเป็นเวลาหลายสิบปี และกำลังเพิ่มขึ้นอีกในปัจจุบันและอนาคต ข้อมูลเหล่านี้สามารถนำมาใช้ประโยชน์ได้ทั้งสำหรับภาครัฐ ภาคเอกชน และประชาชนซึ่งทำให้เกิดเป็นแนวคิด วิธีการรูปแบบทางธุรกิจ รวมถึงวิถีชีวิตใหม่ๆ การแบ่งปันข้อมูลเพื่อนำมาใช้ประโยชน์ และสร้างนวัตกรรมนั้นซึ่งถือเป็นสัญญาณเริ่มต้นของการขับเคลื่อนเศรษฐกิจด้วยข้อมูล อันจะช่วยให้เศรษฐกิจของประเทศเติบโตขึ้น และเพิ่มขีดความสามารถในการแข่งขันในเวทีโลกได้ ที่ผ่านมาหลาย ๆ ประเทศมีการผลักดันและดำเนินการเกี่ยวกับการเปิดเผยข้อมูลสู่สาธารณะกันอย่างจริงจัง เช่นเดียวกันกับในประเทศไทยหลายๆ หน่วยงานก็เริ่มเห็นถึงความสำคัญ และประโยชน์ของการเปิดเผยข้อมูลกันมากขึ้น ดังนั้นสำนักงานรัฐบาลอิเล็กทรอนิกส์ (องค์การมหาชน) (สรอ.) ซึ่งมีภารกิจหลักในการพัฒนารัฐบาล

อิเล็กทรอนิกส์ของประเทศให้มีคุณภาพ จึงได้พัฒนาระบบศูนย์กลางข้อมูลภาครัฐ หรือ data.go.th เพื่อเป็นการส่งเสริมและผลักดันให้เกิดการเปิดเผยข้อมูลภาครัฐสู่สาธารณะอย่าง เป็นรูปธรรม เป็นช่องทางให้ผู้ใช้บริการทั้งภาคประชาชน ภาคธุรกิจเอกชน รวมถึงหน่วยงานของรัฐ สามารถค้นหาและเข้าถึงข้อมูลที่มีคุณภาพของภาครัฐได้ง่าย โดยมีเป้าหมายสูงสุดเพื่อส่งเสริมให้ เกิดธรรมาภิบาล ส่งเสริมให้เกิดความร่วมมือของภาครัฐ และสร้างการมีส่วนร่วมของภาคประชาชน รวมถึงแก้ปัญหาการทุจริตคอรัปชั่น การเมือง และเศรษฐกิจอย่างบูรณาการรวมกันได้ศึกษาปัจจัยที่ส่งผลต่อการเกิดอุบัติเหตุ บนท้องถนนและเปรียบเทียบประสิทธิภาพ การจำแนกข้อมูล คือ การนำข้อมูลที่ได้ถูกรวบรวมไว้และถูกเผยแพร่ผ่านเว็บไซต์ Open Government Data of Thailand ซึ่งเป็นเว็บไซต์ที่เก็บรวบรวมชุดข้อมูลต่าง ๆ ของประเทศไทยซึ่งเป็นข้อมูลที่เปิดเผยได้ เพื่อให้ผู้ใช้บริการสามารถนำชุดข้อมูลไปศึกษาหรือวิเคราะห์ให้เกิดประโยชน์ต่อไปได้ การวิเคราะห์ข้อมูลปริมาณมหาศาลแสดงให้เห็นประโยชน์ต่าง ๆ ก็คือ ความรวดเร็วและประสิทธิภาพ ขณะที่หลายปีก่อนหน้านั้นหลายธุรกิจต้องรวบรวมข้อมูล ดำเนินการวิเคราะห์ และค้นหาข้อมูลที่สามารถนำมาใช้ประกอบการตัดสินใจในอนาคตได้ ส่วนในทุกวันนี้ธุรกิจสามารถระบุข้อมูลเชิงลึกสำหรับการกระทำที่ต้องการการตัดสินใจอย่างทันท่วงที ความสามารถในการทำงานได้เร็วขึ้น และยังคงไว้ซึ่งความคล่องตัว ช่วยให้หลายองค์กรสร้างความได้เปรียบด้านการแข่งขันอย่างที่ไม่เคยมีมาก่อน

อุบัติเหตุทางถนนเป็นปัญหาที่มีผลกระทบอย่างมากต่อชีวิตและทรัพย์สินของประชาชนทั่วโลก เป็นสาเหตุหลักของการบาดเจ็บและการเสียชีวิต ซึ่งส่งผลให้เกิดความสูญเสียทั้งในด้านมนุษย์และเศรษฐกิจ อุบัติเหตุเหล่านี้ไม่เพียงแต่ส่งผลกระทบต่อผู้ที่เกี่ยวข้องโดยตรง แต่ยังสร้างความเครียดและภาระให้กับครอบครัวและสังคมโดยรวม ด้วยเหตุนี้ จึงมีการนำข้อมูลปัจจัยการเกิดอุบัติเหตุ ประจำปี 2562-2567 มาทำการวิเคราะห์ เนื่องจากต้องการทราบปัจจัยที่ทำให้เกิดอุบัติเหตุ ที่นำไปสู่การบาดเจ็บและเสียชีวิต ซึ่งการวิเคราะห์ ปัจจัยต่าง ๆ จากข้อมูลที่ได้มานั้นถือเป็น เรื่องสำคัญต่อการป้องกันภัยจากอุบัติเหตุต่าง ๆ เพื่อหาทางป้องกันและลดอัตราการบาดเจ็บและเสียชีวิตแต่ด้วยจำนวนข้อมูลที่มีปริมาณมหาศาลทำให้ยากต่อการนำเสนอให้เข้าใจได้ง่าย ซึ่งถือเป็นข้อจำกัดหากต้องการนำข้อมูลชุดนี้ไปใช้เพื่อเผยแพร่ให้แก่บุคคลภายนอกหรือกลุ่มผู้ใช้ข้อมูลได้รับรู้

จากปัญหาข้างต้นดังนั้นคณะผู้จัดทำจึงได้นำชุดข้อมูลที่ได้มานั้นเข้าสู่กระบวนการจำแนกตามรายการชุดข้อมูล โดยการใช้ขีดความสามารถในการประมวลผลของโปรแกรมคอมพิวเตอร์เพื่อเพิ่มความน่าเชื่อถือในการวิเคราะห์ และประมวลผลข้อมูลให้เหมาะสมตรงกับความต้องการเป็นการแยกแยะสิ่งที่จะพิจารณาออกเป็นส่วนย่อยที่มีความสัมพันธ์กัน เพื่อทำความเข้าใจแต่ละส่วนให้ชัดเจน รวมทั้งการสืบค้นความสัมพันธ์ของส่วนต่าง ๆ เพื่อดูว่าส่วนประกอบปลีกย่อยนั้นสามารถเข้ากันได้หรือไม่ มีความสัมพันธ์เกี่ยวเนื่องกันอย่างไร โดยคณะผู้จัดทำได้ใช้เทคนิคการทำเหมืองข้อมูลในการศึกษาปัจจัยที่ส่งผลต่อการเกิดอุบัติเหตุบนท้องถนนและเปรียบเทียบประสิทธิภาพตัวแบบการจำแนกข้อมูล ดังนี้ (1) เทคนิคที่ใช้วิเคราะห์ปัจจัยการเกิดอุบัติเหตุ ได้แก่ เทคนิค Chi-Square ใช้

ประเมินค่าของคุณลักษณะ 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 เป็นการสร้างแบบจำลองด้วยวิธีการต้นไม้ตัดสินใจขึ้นมาหลายๆแบบจำลองโดยวิธีการสุ่มตัวแปร แล้วนำผลที่ได้แต่ละแบบจำลองมารวมกันพร้อมนับจำนวนผลที่มีจำนวนซ้ำกันมากที่สุด , เทคนิค Voting เป็นการนำข้อมูลการเรียนรู้ชุดเดียวกันให้โมเดลแต่ละโมเดลทำนาย แล้วโหวตว่าคลาสคำตอบใดที่แต่ละโมเดลทำนายผลมากที่สุด, เทคนิค Naïve Bayes เป็นการจำแนกข้อมูลที่อาศัยหลักการความน่าจะเป็นตามทฤษฎีของเบย์(Bayes's theorem) โดยเรียนรู้ปัญหาที่เกิดขึ้นแล้วนำมาสร้างเงื่อนไขการจำแนกข้อมูลใหม่ , เทคนิค Gradient Boosting เป็นการปรับปรุงประสิทธิภาพของแบบจำลองต้นไม้ตัดสินใจให้มีประสิทธิภาพมากขึ้น โดยการสุ่มสร้างต้นไม้ตัดสินใจหลายสิบต้นจนถึงหลายร้อยต้น และทำการประเมินผลต้นไม้แต่ละต้นจนกว่าจะได้ต้นไม้ตัดสินใจที่มีประสิทธิภาพมากที่สุด และเทคนิค Deep Learning เป็นการเรียนรู้ด้วยการเลียนแบบการทำงานของโครงข่ายประสาทมนุษย์ และเรียนรู้ข้อมูลตัวอย่าง ข้อมูลดังกล่าวถูกนำไปใช้จัดหมวดหมู่ข้อมูลและทางคณะผู้จัดทำยังได้ทำการนำเสนอข้อมูลแบบ Visualization ด้วยการแสดงผลข้อมูลในรูปแบบของภาพด้วยโปรแกรม Power BI และเผยแพร่ข้อมูลสารสนเทศนี้บน Web Browser ที่เป็นที่นิยมในยุคอินเตอร์เน็ต คือ การเผยแพร่ทางสื่อออนไลน์ โดยใช้ภาษา HTML ในการพัฒนาหน้าเว็บไซต์สำหรับการเปิดเผยข้อมูลภาครัฐที่ได้มาจาก data.go.th หรือใน Open Government Data of Thailand เพื่อให้ผู้ใช้งานสามารถเข้าถึงข้อมูลได้อย่างเหมาะสม ซึ่งประโยชน์ที่ผู้ใช้งานทั่วไปจะได้รับจะเป็นในเรื่องของความเข้าใจ การรับรู้ และการตระหนักถึงความปลอดภัยในเหตุการณ์ต่าง ๆ สำหรับผู้ที่ทำการค้นหาและศึกษาข้อมูลก็สามารถดำเนินการผ่าน Web Browser ที่ได้เผยแพร่ข้อมูลสารสนเทศของผู้บาดเจ็บและเสียชีวิตประจำปี 2562–2567 ที่ผ่านการสรุปจำแนกข้อมูลตามประเภทต่าง ๆ ได้สะดวกรวดเร็วยิ่งขึ้น รวมถึงมีความถูกต้อง สะดวก ลดความซ้ำซ้อน เพิ่มประสิทธิภาพให้กับการศึกษาค้นคว้า และทันต่อเวลา

5. วัตถุประสงค์

1. เพื่อศึกษาปัจจัยที่ส่งผลต่อการเกิดอุบัติเหตุ บนท้องถนน
2. เพื่อเปรียบเทียบประสิทธิภาพการจำแนกข้อมูลบนชุดข้อมูลการเกิดอุบัติเหตุบนท้องถนน

6. ประโยชน์ที่จะได้รับ

1. ได้ผลการศึกษาปัจจัยที่ส่งผลต่อการเกิดอุบัติเหตุ บนท้องถนน
2. ได้ผลการเปรียบเทียบประสิทธิภาพการจำแนกข้อมูลบนชุดข้อมูลการเกิดอุบัติเหตุบนท้องถนน

7. วรรณกรรมที่เกี่ยวข้อง

7.1 วรการ ใจดี และ นรินทร์ จิวิตัน (2566) ปัจจัยที่ส่งผลต่อการตัดสินใจเข้าศึกษาต่อในระดับปริญญาตรี คณะบริหารธุรกิจและศิลปศาสตร์ โดยใช้เทคนิคการคัดเลือกคุณลักษณะและเทคนิคการทำเหมืองข้อมูล การวิจัยนี้ศึกษาปัจจัยที่มีผลต่อการตัดสินใจเข้าศึกษาต่อในระดับปริญญาตรี คณะบริหารธุรกิจและศิลปศาสตร์ โดยวิเคราะห์ข้อมูลจากการรับสมัครนักศึกษามหาวิทยาลัยเทคโนโลยีราชมงคลล้านนาเชียงใหม่ ช่วงปีการศึกษา 2563–2565 ซึ่งมีทั้งหมด 2,509 รายการ โดยใช้เทคนิคการคัดเลือกคุณลักษณะ 3 เทคนิค และเทคนิคการทำเหมืองข้อมูล 6 เทคนิค ผลการศึกษาแสดงให้เห็นว่าปัจจัยสำคัญที่ส่งผลต่อการตัดสินใจ ได้แก่ สาขาวิชาหรือหลักสูตรที่เลือก สายการเรียนเดิม และวุฒิการศึกษาเดิม ส่วนเทคนิคการทำเหมืองข้อมูลที่มีความแม่นยำสูงสุดคือ เทคนิคการโหวต ด้วยความถูกต้อง 73.44% มากกว่าเทคนิคอื่น ๆ เช่น เทคนิคป่าสุ่มและเทคนิคการถดถอยโลจีสติก.

การใช้เทคนิคเหมืองข้อมูลเป็นการค้นหาองค์ความรู้ (Knowledge) จากข้อมูลที่มีอยู่มาทำการวิเคราะห์ปัจจัยที่ส่งผลต่อการตัดสินใจเข้าศึกษาต่อในระดับปริญญาตรี คณะบริหารธุรกิจและศิลปศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนาเชียงใหม่ ซึ่งในการวิจัยครั้งต่อไป คณะผู้วิจัยมีความเห็นว่า 1. ควรเพิ่มปัจจัยอื่น ๆ ที่อาจจะมีผลต่อการเข้าศึกษาต่อในระดับปริญญาตรี เช่น ปัจจัยด้านผู้ปกครอง อาชีพ ผู้ปกครอง รายได้ผู้ปกครอง ปัจจัยด้านข้อมูลส่วนตัวนักศึกษา เช่น งานอดิเรก ความสามารถ อาชีพในอนาคต เป็นต้น 2. ควรมีการใช้เทคนิคการพยากรณ์อื่น ๆ เข้ามาเปรียบเทียบ เช่น เทคนิคการเรียนรู้แบบรวมกลุ่ม (Ensemble Model) ประเภทต่าง ๆ เพื่อเพิ่มประสิทธิภาพให้กับตัวแบบพยากรณ์ รวมถึงมีการปรับปรุงพารามิเตอร์ของแต่ละเทคนิคการพยากรณ์ให้ประสิทธิภาพมากขึ้น 3. ควรมีการใช้เทคนิคการปรับสมดุลข้อมูลที่หลากหลาย เพื่อเพิ่มประสิทธิภาพของตัวแบบพยากรณ์ เช่น วิธีการ สุ่มเพิ่ม (Oversampling) วิธีการผสมผสาน (Hybrid Method) เป็นต้น 4. ควรมีการทดสอบเทคนิคการคัดเลือกคุณลักษณะและเทคนิคการพยากรณ์หลายๆ ครั้ง เพื่อให้ได้ประสิทธิภาพมากขึ้นและเพื่อไม่ให้เกิดปัญหา overfiting

7.2 วรการ ใจดี และ นพณัฐ วรรณภีร์ (2563) การศึกษาปัจจัยที่ส่งผลต่อการสำเร็จการศึกษาตามแผนของนักศึกษาระดับปริญญาตรี โดยใช้เทคนิคการคัดเลือกคุณลักษณะบนชุดข้อมูลที่ไม่สมดุล งานวิจัยนี้ได้จัดทำขึ้นโดยมีวัตถุประสงค์ 2 ประการคือ (1) เพื่อที่ศึกษาปัจจัยที่ส่งผลต่อการสำเร็จการศึกษา ตามแผนของนักศึกษา และ (2) เพื่อศึกษาแนวทางและข้อปรับปรุงในการแก้ไขปัญหาการเรียนการสอนในแต่ละรายวิชาที่ ส่งผลต่อการสำเร็จการศึกษาตามแผนของนักศึกษา โดยปัจจัยหรือ

ข้อมูลหลักในการวิจัยนี้ ผู้วิจัยจะใช้ข้อมูลระดับคะแนน ในแต่ละวิชา ข้อมูลเกรดเฉลี่ย ข้อมูลเพศของนักศึกษาชั้นปี ที่ 1–2 โดยผู้วิจัยได้เก็บรวบรวมข้อมูลนักศึกษาจากสาขาวิชา ระบบสารสนเทศทางคอมพิวเตอร์ คณะบริหารธุรกิจและศิลป์ศาสตร์ มหาวิทยาลัยเทคโนโลยีราชมงคลล้านนาเชียงใหม่ จำนวน 358 คน จากการรวบรวมข้อมูลเบื้องต้นพบว่า ชุดข้อมูลมีความไม่สมดุลกัน (Imbalanced Datasets) โดยมีจำนวนกลุ่มนักศึกษาที่สำเร็จการศึกษาตามแผนมากกว่ากลุ่มนักศึกษาที่ไม่สำเร็จการศึกษาตามแผน ดังนั้นผู้วิจัยจึงได้ทำการแก้ปัญหาโดยใช้วิธีการปรับสมดุลให้กับชุดข้อมูลโดยใช้เทคนิคการสุ่มเพิ่มตัวอย่างข้อมูลกลุ่มน้อย (Synthetic Minority Over–sampling Technique : SMOTE) ก่อนที่จะนำเข้าสู่กระบวนการวิเคราะห์ ปัจจัยโดยจะใช้ เทคนิคการคัดเลือกคุณลักษณะ (Feature Selection) 3 เทคนิค คือ (1) Chi–Square Feature Selection (2) Information Gain Feature Selection และ (3) Correlation Based Feature Selection จากการทดลองพบว่าปัจจัยที่มีความสำคัญมากที่สุดและเป็นอันดับหนึ่งในทุกเทคนิคการคัดเลือกคุณลักษณะคือ วิชาการเขียนโปรแกรมคอมพิวเตอร์ 1 ซึ่งวิชานี้เป็นวิชาพื้นฐานและ เป็นวิชาบังคับของสาขาระบบสารสนเทศทางคอมพิวเตอร์และเป็นวิชาที่นักศึกษาส่วนใหญ่ได้ระดับคะแนนในวิชานี้ค่อนข้างน้อย ซึ่งอาจส่งผลให้รายวิชานี้ได้ลำดับความสำคัญมาเป็นอันดับหนึ่งในทุกเทคนิคการคัดเลือกคุณลักษณะและอาจเป็นปัจจัยหลักที่ส่งผลต่อการสำเร็จการศึกษาตามแผนของนักศึกษาอย่างมีนัยสำคัญ.

เทคนิคการคัดเลือกคุณลักษณะทั้งสามเทคนิคที่ผู้วิจัยได้นำเสนอ นั้นเป็นเพียงบางเทคนิคของการคัดเลือกคุณลักษณะจากหลากหลายเทคนิคที่มีอยู่ในปัจจุบัน และยังมีข้อจำกัดในเรื่องของปัจจัยข้อมูลที่มีเพียงข้อมูลระดับคะแนนรายวิชา เกรดเฉลี่ยเพศของนักศึกษาเท่านั้น ซึ่งอาจจะมีปัจจัยในด้านอื่น ๆ ที่อาจ ส่งผลต่อการสำเร็จการศึกษาของนักศึกษาได้ เช่น ข้อมูลด้านสถานภาพทางครอบครัว สถานการณ์กู้ยืมเงิน ข้อมูลการศึกษาในอดีต เป็นต้น ในอนาคตหากมีการพัฒนางานวิจัยให้มีประสิทธิภาพมากขึ้น อาจจะต้องมีการเก็บรวมรวมข้อมูลเพื่อให้ได้ปัจจัยที่เพิ่มมากขึ้นและใช้เทคนิคในการคัดเลือกคุณลักษณะในหลายๆ เทคนิคมากขึ้นเพื่อสามารถนำผลของการทดลองไปใช้ได้อย่างมีประสิทธิภาพสูงสุด และในงานวิจัยนี้ผู้วิจัยยังไม่ได้มีกำดำเนินการในการนำคุณลักษณะที่คัดเลือกได้ไปหาค่าความแม่นยำในความถูกต้องของแต่ละเทคนิคการคัดเลือกคุณลักษณะ ที่ใช้ ซึ่งผู้วิจัยจะมีการดำเนินการในหัวข้อดังกล่าวในงานวิจัยถัดไป

7.3 จิรกิตติ์ จินะกัณฑ์ และ สุรางคณา หมื่นคำแปง (2564) การศึกษานี้วิเคราะห์ข้อมูลผู้เสียชีวิตจากอุบัติเหตุทางถนนระหว่างปี 2554–2563 โดยใช้กระบวนการ CRISP–DM และเทคนิคการทำเหมืองข้อมูลแบบ Classification กับโมเดล Decision Tree (J48) พบว่าความถูกต้องในการจำแนกประเภท "ดื่ม" และ "ไม่ดื่ม" คือ 67.37% พร้อมทั้งสร้างเว็บไซต์เผยแพร่ข้อมูลในรูปแบบแผนภาพผ่าน Google Data Studio โดยใช้ HTML, CSS, PHP และ JavaScript เว็บไซต์นี้มีประโยชน์ในการศึกษาและสามารถนำข้อมูลไปปรับปรุงเพื่อเพิ่มประสิทธิภาพความปลอดภัยในอนาคต.

การวิเคราะห์ข้อมูลผู้เสียชีวิตจากอุบัติเหตุทางถนน ปี 2554 – 2563 ที่จัดทำเสร็จสิ้นแล้วแม้ว่าจะสามารถจัดทำและแสดงผลได้ตรงตามวัตถุประสงค์และขอบเขตที่กำหนดไว้ แต่ยังมีอีกหลายข้อจำกัดที่ยังสามารถพัฒนาต่อได้ ซึ่งหากจะมีผู้พัฒนาให้เว็บไซต์เผยแพร่ข้อมูลแสดงผลและทำงานได้อย่างมีประสิทธิภาพมากยิ่งขึ้นผู้พัฒนาจะต้องปรับปรุงในส่วนต่าง ๆ 1. ควรมีการพัฒนาหน้าการแสดงผลต่าง ๆ ของเว็บไซต์ให้เหมาะสม สะดวก และเข้าใจง่ายมากยิ่งขึ้น เพื่อลดความผิดพลาดและเพิ่มประสิทธิภาพในการนำเสนอข้อมูล 90 2. ควรมีการพัฒนาในส่วนของเว็บไซต์ให้มีการรองรับบนโทรศัพท์ได้ในอนาคต 3. ควรมีการพัฒนาการปรับเปลี่ยนรูปแบบมุมมองการแสดงกราฟข้อมูลในรูปแบบต่าง ๆ ได้ 4. ควรทำให้ระบบสามารถแยกดูรายงานยอดสรุปของแต่ละจังหวัดได้

7.4 มนิศรา เทพฝั้น และศิริขวัญ ขันนาแล (2562) การจัดทำโครงงานนี้มีวัตถุประสงค์เพื่อวิเคราะห์ ข้อมูลผู้บาดเจ็บและเสียชีวิตในเทศกาลปีใหม่ระหว่างปี 2551–2558 และเผยแพร่ข้อมูลผ่านเว็บไซต์ โดยใช้กระบวนการ CRISP–DM และเทคนิคการทำเหมืองข้อมูลแบบ Classification ด้วยโมเดล Decision Tree (J48) ซึ่งมีความถูกต้อง 70.12% และสร้างกฎการจำแนกประเภทได้ 9 กฎ เว็บไซต์ที่สร้างขึ้นมีลักษณะการเผยแพร่ข้อมูลในรูปแบบแผนภาพผ่าน Tableau Public และใช้ HTML, CSS, PHP, และ JavaScript โดยประกอบด้วยข้อมูลเกี่ยวกับอุบัติเหตุ การวิเคราะห์ข้อมูล การสรุปผล รายปี และแนวทางป้องกันอุบัติเหตุ เว็บไซต์นี้สามารถนำไปปรับปรุงเพิ่มเติมเพื่อเพิ่มประสิทธิภาพในการส่งเสริมความปลอดภัยของบุคคลทั่วไปได้ในอนาคต.

การวิเคราะห์ข้อมูลผู้ประสบอุบัติเหตุในช่วงเทศกาลตั้งแต่ปี 2551 – 2558 เพื่อเผยแพร่ บนเว็บไซต์ที่จัดทำเสร็จสิ้นแล้วนี้แม้จะสามารถทำงานและแสดงผลข้อมูลได้ถูกตองตรงตาม วัตถุประสงค์และขอบเขตที่ตั้งไว้ แต่ก็ยังมีข้อจำกัดหลายประการ ซึ่งหากจะพัฒนาใหเว็บไซต์ เผยแพร่ขอมูลนี้แสดงผลขอมูลและทำงานได้อย่างมีประสิทธิภาพมากขึ้น ผู้พัฒนาควรจะตอง ปรับปรุงในส่วนต่าง ๆ ดังต่อไปนี้ 1. หากมีชุดข้อมูลเพิ่มเติมควรมีการปรับปรุงข้อมูลใหม่ ๆ อยู่เสมอ 2. ควรจัดเก็บข้อมูลด้วยตัวเองเพื่อให้ได้ข้อมูลที่ตรงต่อความต้องการในการวิเคราะห์ ข้อมูลในครั้งต่อไป 3. ควรมีการพัฒนาหน้าการแสดงผลต่าง ๆ ของเว็บไซต์ให้เหมาะสม สะดวก และ เข้าใจง่ายมากยิ่งขึ้น เพื่อลดความผิดพลาดและเพิ่มประสิทธิภาพในการนำเสนอข้อมูล

7.5 นรินทร์ จิวิตัน (2565) การเปรียบเทียบประสิทธิภาพในการทำนายผลความไม่สมดุลของข้อมูลการเสียชีวิตจากอุบัติเหตุบนโครงข่ายถนน โดยใช้เทคนิคการทำเหมืองข้อมูล งานวิจัยนี้มุ่งศึกษาและวิเคราะห์ปัจจัยที่ส่งผลต่อการเสี่ยงชีวิตจากอุบัติเหตุบนโครงข่ายถนน โดยใช้เทคนิคต่างๆ ในการจำแนกข้อมูลเพื่อทำนายความเสี่ยง มีการใช้เทคนิคหลายประการในการสร้างแบบจำลอง เช่น Gradient Boosting, Decision Trees, Naive Bayes, Random Forest, และ Neural Networks ซึ่งการศึกษาใช้กระบวนการ CRISP–DM และได้รวบรวมข้อมูลจากศูนย์เทคโนโลยีสารสนเทศและการสื่อสาร สำนักงานปลัดกระทรวงคมนาคม ระหว่างวันที่ 1 มกราคม 2562 ถึง 30 มิถุนายน 2564 จำนวน 51,384 แถวการวิเคราะห์พบว่าเทคนิค Naive Bayes มีประสิทธิภาพต่ำที่สุดที่ 72.23% ขณะที่เทคนิค

อื่นๆ เช่น Random Forest มีความแม่นยำรวม 73.35% และ Neural Networks มีค่าเฉลี่ยความแม่นยำ 72.10% โดยการใช้ K-Fold Cross Validation แบ่งข้อมูลออกเป็น 30 ส่วนเพื่อประเมินประสิทธิภาพของแต่ละเทคนิคผลการวิจัยสามารถใช้เป็นข้อมูลพื้นฐานในการป้องกันอุบัติเหตุและเตรียมความพร้อมในการจัดการความเสี่ยงบนโครงข่ายถนนได้อย่างมีประสิทธิภาพ.

เพื่อให้ได้ข้อสรุปของผลการวิเคราะห์ข้อมูลที่ครอบคลุมมากขึ้นอาจเปรียบเทียบประสิทธิภาพโดยใช้อัลกอริทึมประเภทอื่นเพิ่มเติม เช่น วิธีเคเนียเรสเนเบอร์ วิธีโครงข่ายประสาทเทียม วิธีซัพพอร์ตเวกเตอร์แมชชีน เป็นต้น มีการเปรียบเทียบกับการใช้เครื่องมือสำเร็จรูปในการช่วยวิเคราะห์อื่น เช่น Microsoft Power BI, Tableau, Pentaho Big Data Integration

8. ขอบเขต

8.1 ขอบเขตการวิเคราะห์ข้อมูล

8.1.1 เก็บรวบรวมข้อมูล (Data Integration) จากข้อมูลอุบัติเหตุบนท้องถนนตั้งแต่ปี 2562 ถึง 2567

8.1.2 ทำการคัดเลือก (Data Selection) โดยเลือกคอลัมน์ที่มีความเชื่อมโยงกันและสามารถนำมาสร้างโมเดลต่างๆได้

8.1.3 ทำการทำความสะอาดข้อมูล (Data Cleaning) โดยการตรวจสอบข้อมูลเปลี่ยนแปลงข้อมูลให้ได้ข้อมูลที่ถูกต้องและสมบูรณ์

8.1.4 แปลงข้อมูลให้อยู่ในรูปแบบเพื่อใช้ในการวิเคราะห์ ให้นำข้อมูลมีความสมบูรณ์เพื่อนำมา แปลงข้อมูลให้อยู่ในรูปแบบข้อมูลที่สามารถนำมาวิเคราะห์โดยเทคนิคการทำเหมืองข้อมูล

8.1.5 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

8.1.6 ประเมินรูปแบบ (Pattern Evaluation) ประเมินผลประสิทธิภาพของโมเดล ด้วยตาราง Confusion Matrix กำหนดให้ค่าที่ยอมรับและดีที่สุด คือ 80

8.1.7 นำเสนอข้อมูลภาพ (Data Visualization) นำเสนอและเผยแพร่ข้อมูลผ่านเว็บไซต์โดยเว็บไซต์จะพัฒนาโดยใช้ภาษา HTML และชุดคำสั่ง CSS เพื่อนำเข้า วิเคราะห์ข้อมูลด้วยโมเดลที่เลือกใช้และผ่านการทดสอบประสิทธิภาพของการวิเคราะห์ข้อมูลแล้ว และนำเสนอข้อมูลแบบ Visualization สำหรับแสดงผลข้อมูลในรูปแบบภาพจากข้อมูลที่ผ่านการวิเคราะห์แล้วภายในเว็บไซต์

8.2 ขอบเขตผู้ใช้งานทั่วไปบนเว็บบราวเซอร์

8.2.1 สามารถดูข้อมูลสารสนเทศผู้บาดเจ็บเผยแพร่ผ่านทางเว็บไซต์

- สารสนเทศสถิติผู้บาดเจ็บแต่ละจังหวัด
- สารสนเทศสถิติจังหวัดที่มีผู้บาดเจ็บสูงที่สุด

8.2.2 สามารถดูข้อมูลสารสนเทศผู้เสียชีวิตเผยแพร่ผ่านทางเว็บไซต์

- สารสนเทศสถิติผู้เสียชีวิตแต่ละจังหวัด
- สารสนเทศสถิติจังหวัดที่มีผู้เสียชีวิตสูงที่สุด

8.2.3 สามารถดูข้อมูลสารสนเทศปัจจัยที่มีผลต่อการเกิดอุบัติเหตุ

8.2.4 สามารถดูผลการเปรียบเทียบประสิทธิภาพการจำแนกข้อมูล

8.2.5 สามารถกรองข้อมูลสารสนเทศผ่านระบบของ Power BI บน Web browser ได้

9. เครื่องมือที่ใช้ในการพัฒนาโปรแกรม

9.1 Hardware

9.1.1 เครื่องคอมพิวเตอร์โน้ตบุ๊ครุ่น Lenovo

9.2 Software

9.2.1 โปรแกรม RapidMiner Studio ใช้ในการวิเคราะห์ข้อมูล

9.2.2 โปรแกรม Xampp ใช้ในการจำลอง web server

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

9.2.4 โปรแกรม Microsoft word ใช้จัดทำเอกสารประกอบโครงงาน

9.2.5 โปรแกรม Microsoft excel ใช้เตรียมข้อมูล

9.2.6 โปรแกรม Power BI ใช้นำเสนอข้อมูลรูปแบบของภาพ

9.2.7 โปรแกรม Sublime Text 3 ใช้เขียนเขียนโค้ด HTML และ CSS

10. สถานที่ใช้ในการดำเนินการศึกษาและรวบรวมข้อมูล

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

11. 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.จัดทำเอกสารประกอบโครงงาน			←	—	—	—	—	—	→

12. นิยามศัพท์เฉพาะ (ถ้ามี)

.....

.....

.....

13. ความเห็นของอาจารย์ที่ปรึกษา

o เห็นชอบและอนุญาตให้นักศึกษาทำโครงการได้

o ไม่อนุญาตให้นักศึกษาทำโครงการนี้

ความเห็นเพิ่มเติม.....

.....

ลงชื่อ .....

(.....)

14. ความเห็นเจ้าของหน่วยงาน/ระบบงาน (ตามดุลยพินิจของอาจารย์ที่ปรึกษา)

o เห็นชอบและอนุญาตให้นักศึกษาทำโครงการได้

o ไม่อนุญาตให้นักศึกษาทำโครงการนี้

ความเห็นเพิ่มเติม.....

.....

ลงชื่อ .....

(.....)

ตำแหน่ง .....เบอร์โทรติดต่อ .....

หน่วยงาน/บริษัท/ห้างร้าน .....