



การพยากรณ์การเสียชีวิตจากอุบัติเหตุทางถนนในช่วงเทศกาลปีใหม่

โดย

นายพิเศษ คุปตินทร

การค้นคว้าอิสระนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

วิทยาศาสตรมหาบัณฑิต

สาขาวิชานโยบายและการบริหารดิจิทัล

วิทยาลัยนวัตกรรม มหาวิทยาลัยธรรมศาสตร์

ปีการศึกษา 2561

ลิขสิทธิ์ของมหาวิทยาลัยธรรมศาสตร์

การพยากรณ์การเสียชีวิตจากอุบัติเหตุทางถนนในช่วงเทศกาลปีใหม่

โดย

นายพิเศษ คุปตินทร



การค้นคว้าอิสระนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

วิทยาศาสตรมหาบัณฑิต

สาขาวิชานโยบายและการบริหารดิจิทัล

วิทยาลัยนวัตกรรม มหาวิทยาลัยธรรมศาสตร์

ปีการศึกษา 2561

ลิขสิทธิ์ของมหาวิทยาลัยธรรมศาสตร์

PREDICTION OF ROAD DEATHS DURING NEW YEAR FESTIVAL

By

MR. PISES CUPTINTORN



AN INDEPENDENT STUDY SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE

DIGITAL POLICY AND MANAGEMENT

COLLEGE OF INNOVATION

THAMMASAT UNIVERSITY

ACADEMIC YEAR 2018

COPYRIGHT OF THAMMASAT UNIVERSITY

มหาวิทยาลัยธรรมศาสตร์

วิทยาลัยนวัตกรรมการ

การค้นคว้าอิสระ

ของ

นายพิเศษ คุปตินทร

เรื่อง

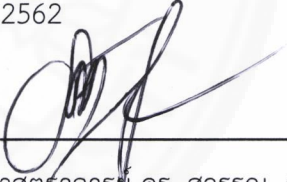
การพยากรณ์การเสียชีวิตจากอุบัติเหตุทางถนนในช่วงเทศกาลปีใหม่

ได้รับการตรวจสอบและอนุมัติ ให้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

วิทยาศาสตรมหาบัณฑิต

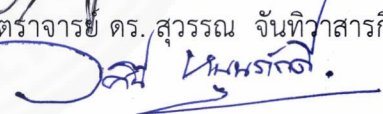
เมื่อวันที่ 1 มิถุนายน พ.ศ. 2562

ประธานกรรมการสอบการค้นคว้าอิสระ



(ผู้ช่วยศาสตราจารย์ ดร. สุวรรณ จันทนาสารกิจ)

กรรมการและอาจารย์ที่ปรึกษาการค้นคว้าอิสระ



(ดร. วศินี หนูนภักดี)

กรรมการสอบการค้นคว้าอิสระ



(ดร. มานิต สาธิตสมิตพงษ์)

คณบดี



(ผู้ช่วยศาสตราจารย์ ดร. ประวิทย์ เขมะสุนันท์)

หัวข้อการค้นคว้าอิสระ	การพยากรณ์การเสียชีวิตจากอุบัติเหตุทางถนนในช่วงเทศกาลปีใหม่
ชื่อผู้เขียน	นายพิเศษ คุปตินทร
ชื่อปริญญา	วิทยาศาสตรมหาบัณฑิต
สาขาวิชา/คณะ/มหาวิทยาลัย	สาขาวิชานโยบายและการบริหารจัดการดิจิทัล วิทยาลัยนวัตกรรมการ มหาวิทยาลัยธรรมศาสตร์
อาจารย์ที่ปรึกษาการค้นคว้าอิสระ	ดร. วศินี หนูนภักดี
ปีการศึกษา	2561

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อที่ทำการศึกษาปัจจัย และคุณลักษณะของข้อมูลต่างๆ ในการเสียชีวิตทางถนนในช่วงเทศกาลปีใหม่ โดยผู้ทำวิจัยใช้เครื่องมือเครื่องการเรียนรู้ (Machine learning) เพื่อทำการวิเคราะห์ข้อมูล ซึ่งผู้ทำวิจัยได้ใช้เทคนิคการเรียนรู้ของเครื่องทั้งแบบมีการสอนเครื่องเรียนรู้ (Supervised learning) และไม่มีการสอนเครื่องเรียนรู้ (Unsupervised learning) โดยที่ผู้ทำวิจัยได้ทำการสร้างแบบจำลองการพยากรณ์โดยใช้ Decision tree และ Naïve Bayes พร้อมทั้งมีการแบ่งลักษณะประชากรที่โอกาสในการเสียชีวิตทางถนนในช่วงเทศกาลปีใหม่โดยใช้ K-mean

ผลการวิจัยโดยใช้เครื่องเรียนรู้ ได้มีการสร้างแบบจำลองการพยากรณ์ โดยที่เทคนิค Decision tree นั้นค่าความแม่นยำอยู่ที่ 98.6766 % และค่าประสิทธิภาพของแบบจำลองอยู่ที่ 0.984 ส่วนเทคนิค Naïve Bayes นั้นค่าความแม่นยำอยู่ที่ 98.6500 % และค่าประสิทธิภาพของแบบจำลองอยู่ที่ 0.982 สำหรับการแบ่งลักษณะประชากรที่มีโอกาสเสียชีวิตทางถนนในช่วงเทศกาลปีใหม่ มีค่าความแม่นยำอยู่ที่ 56% โดยทำการแบ่งประชากรออกเป็น 2 กลุ่ม คือเพศชาย และเพศหญิง ซึ่งลักษณะข้อมูลที่มีความแตกต่างกันอย่างชัดเจน คือ ช่วงอายุ การดื่มแอลกอฮอล์ และรถคู่กรณี ผลการวิจัยครั้งนี้สามารถนำไปเป็นแนวทางในการออกมาตรการต่าง ๆ เพื่อลดการสูญเสียชีวิตและทรัพย์สินอันเกิดจากอุบัติเหตุในช่วงเทศกาลปีใหม่ เช่น การรณรงค์เมาไม่ขับ การเพิ่มพูนความรู้ การปฐมพยาบาล เป็นต้น

คำสำคัญ: การเรียนรู้ของเครื่อง, การเสียชีวิต, การเกิดอุบัติเหตุ,แบบจำลองการพยากรณ์

Independent Study Title	PREDICTION OF ROAD DEATHS DURING NEW YEAR FESTIVAL
Author	Mr. Pises Cuptintorn
Degree	Master of Science
Department/Faculty/University	Digital Policy and Management College of Innovation Thammasat University
Independent Study Advisor	Wasinee Noonpakdee, D.Sc.
Academic Year	2018

ABSTRACT

This independent study aims to examine some factors and features of road deaths during the New Year festival in Thailand. Machine learning technique is applied in order to analyse data using Decision tree and Naïve Bayes for supervised learning and K – mean for unsupervised learning.

The results of decision tree indicate that the accuracy of the model is 98.6766% and the efficiency is 0.984. For Naïve Bayes, the accuracy of model is 98.6500% and the efficiency is 0.982. Regarding K – mean, the accuracy of people who died during the New Year festival is 56%. K – mean divided people into two groups which are male and female with the different characteristics of age, drinking and party vehicle. The results of this research can be used as a guideline for issuing measure to reduce the loss of life and property caused by accidents during the New Year festival such as anti-drunk driving campaign, and Basic First Aid knowledge.

Keywords: Machine learning , Accident , Death , Predictive Model

กิตติกรรมประกาศ

การค้นคว้าอิสระนี้สำเร็จลุล่วงไปได้ด้วยดี ผู้วิจัยต้องกราบขอบพระคุณอาจารย์ที่ปรึกษา อาจารย์ ดร.วศินี หนูนภักดี เป็นอย่างสูงไว้ ณ ที่นี้ ที่ท่านได้กรุณาเสียสละเวลาในการให้คำปรึกษา และเปิดโอกาสให้ได้ทำการค้นคว้าวิจัยในเรื่องที่ผู้วิจัยมีความต้องการศึกษาอย่างอิสระ โดยคอยให้คำแนะนำความรู้ต่าง ๆ พร้อมทั้งแนวทางในการดำเนินการศึกษาวิจัย ตลอดจนการตรวจสอบและแก้ไขจนการค้นคว้าอิสระครั้งนี้สมบูรณ์ และขอขอบพระคุณ ผศ.ดร.จิรพล สังข์โพธิ์ และผู้ช่วยศาสตราจารย์ ดร. สุวรรณ จันทิวาสารกิจ ประธานกรรมการสอบการค้นคว้าอิสระ ดร.มานิต สาธิตสมิตพงษ์ กรรมการสอบการค้นคว้าอิสระ ที่ให้คำชี้แนะแนวทางการปรับปรุงและคำแนะนำเพิ่มเติม เพื่อให้งานวิจัยฉบับนี้มีความถูกต้องสมบูรณ์ยิ่งขึ้น

ผู้วิจัยขอขอบพระคุณคณาจารย์ทุกท่านในหลักสูตรที่ให้ความรู้และถ่ายทอดประสบการณ์และคำแนะนำต่าง ๆ ทั้งในเรื่องของความรู้ทางวิชาการ และการดำเนินชีวิตในสังคม ตลอดจนช่วยสร้างโอกาสการเรียนรู้ใหม่ ๆ และได้พบกับโอกาสที่ดีต่าง ๆ ในชีวิตการทำงาน และขอขอบคุณเพื่อน ๆ พี่ ๆ น้อง ๆ CIO ทุกคนที่คอยช่วยเหลือและสนับสนุนซึ่งกันและกันมาโดยตลอด

หากผลการศึกษานี้มีข้อบกพร่องประการใด ผู้ศึกษาขอน้อมรับไว้เพื่อปรับปรุง แก้ไขในการศึกษาครั้งต่อไป

นายพิเศษ คุปตินทร

(4)

สารบัญ

	หน้า
บทคัดย่อภาษาไทย	(1)
บทคัดย่อภาษาอังกฤษ	(2)
กิตติกรรมประกาศ	(3)
สารบัญ	(4)
สารบัญตาราง	(8)
สารบัญภาพ	(10)
บทที่ 1 บทนำ	1
1.1 ที่มาและความสำคัญของปัญหา	1
1.2 วัตถุประสงค์ของการวิจัย	3
1.3 ขอบเขตของการวิจัย	3
1.4 ขอบเขตด้านประชากรและกลุ่มตัวอย่าง	3
1.5 ระเบียบวิธีวิจัย	4
1.6 ประโยชน์ที่คาดว่าจะได้รับ	4
1.7 นิยามศัพท์เฉพาะ	4
บทที่ 2 แนวคิด ทฤษฎีและงานวิจัยที่เกี่ยวข้อง	5
2.1 แนวคิดและทฤษฎีเกี่ยวกับการเกิดอุบัติเหตุ	5
2.1.1 ความหมาย	5

	(5)
2.1.2 ประเภทของอุบัติเหตุ	6
2.1.3 ความสูญเสียจากอุบัติเหตุ	6
2.1.4 สาเหตุที่ก่อให้เกิดอุบัติเหตุจราจร	6
2.1.5 สภาพแวดล้อมบนถนนที่ไม่ปลอดภัย	7
2.1.6 เหตุสุดิวสัย	8
2.1.7 ทฤษฎีอุบัติเหตุเกี่ยวข้องกับงานวิจัย	8
2.1.8 ชุดปฏิบัติการฉุกเฉิน	10
2.1.9 แนวคิดทฤษฎีเกี่ยวกับความหมายและกระบวนการนโยบาย สาธารณะ (Public Policy)	11
2.2 หลักการและทฤษฎีเกี่ยวกับการเรียนรู้ของเครื่อง (Machine Learning)	12
2.2.1 ความหมายของการเรียนรู้ของเครื่อง	12
2.2.2 กระบวนการวิเคราะห์ข้อมูลด้วย CRISP-DM	13
2.2.3 ประเภทของการเรียนรู้ของเครื่อง	14
2.2.4 การแบ่งข้อมูลเพื่อนำมาทดสอบประสิทธิภาพของแบบจำลอง	17
2.2.5 การคัดเลือกคุณลักษณะ	19
2.2.6 การวัดประสิทธิภาพของโมเดลการพยากรณ์	20
2.3 เอกสารและงานวิจัยที่เกี่ยวข้อง	22
2.4 กระบวนการในการวิจัย	23
บทที่ 3 วิธีการวิจัย	24
3.1 ขอบเขตของการวิจัยและกลุ่มตัวอย่างที่ใช้ในการวิจัย	24
3.2 เครื่องมือที่ใช้ในการวิจัย	24
3.2.1 เครื่องมือที่ใช้ในการทำการเรียนรู้ของเครื่อง (Machine Learning)	24
3.2.2 การทดสอบเนื้อหาความในการวิจัย (Content Validity Testing)	25
3.3 ขั้นตอนการดำเนินงาน	25
3.3.1 การศึกษาและการวางแผนการเก็บข้อมูล	25
3.3.2 การเก็บรวบรวมข้อมูล	27
3.3.3 การสำรวจข้อมูลและจัดการข้อมูล	27
3.3.4 การวิเคราะห์และออกแบบตัวแบบการทำนาย	27

	(6)
3.3.5 การอธิบายผลและการนำไปใช้	28
3.4 ระยะเวลาในการวิจัย	28
บทที่ 4 ผลการวิจัยและอภิปรายผล	31
4.1 ผลการวิจัย	31
4.1.1 ผลของการดำเนินงานในแต่ละขั้นตอน	31
4.1.2 ผลของการจัดกลุ่มลักษณะของประชากรที่โอกาสเสียชีวิตทางถนน ในช่วงเทศกาลปีใหม่	64
4.1.3 ผลของการจัดแบบจำลองพยากรณ์	66
4.2 การอภิปรายผล	67
4.2.1 ผลของการวิจัยเปรียบเทียบกับงานวิจัยอื่น	67
4.2.2 ผลของการวิจัยเทียบกับแนวคิดทฤษฎีโดมิโน	68
บทที่ 5 สรุปผลการวิจัยและข้อเสนอแนะ	70
5.1 สรุปผลการวิจัย	70
5.1.1 แบบจำลองการพยากรณ์	70
5.1.2 การจัดกลุ่มข้อมูล	71
5.2 แนวทางการนำไปใช้	71
5.3 ข้อจำกัดงานวิจัย	71
5.4 ประโยชน์ที่ได้รับ	72
5.4.1 ด้านวิชาการ	72
5.4.2 ด้านผู้วิจัย	72
5.4.3 ด้านภาครัฐหรือหน่วยงานที่เกี่ยวข้อง	72
5.5 ข้อเสนอแนะ	73
5.6 แนวทางการทำวิจัยในขั้นถัดไป	73

(7)

รายการอ้างอิง

74

ประวัติผู้เขียน

76



สารบัญตาราง

ตารางที่	หน้า
1.1 แสดงการเกิดอุบัติเหตุทางถนนสาเหตุจากบุคคลในปี 2549 – 2558	1
2.1 ประเภทของบุคคลตามลักษณะพื้นฐาน 2 ประเภท	10
2.2 ตารางสรุปการเรียนรู้ของเครื่อง (Machine learning)	15
2.3 ตารางแสดงตัวอย่าง Confusion Matrix	21
2.4 ตารางแสดงผลสรุปปัจจัยที่มีผลต่อการเกิดอุบัติเหตุ	22
3.1 แสดงข้อมูลด้านประชากร	25
3.2 แสดงข้อมูลด้านสิ่งแวดล้อม	26
3.3 แสดงข้อมูลด้านนโยบายสาธารณะ	26
3.4 แสดงข้อมูลข้อมูลผลการรักษา	27
3.5 ระยะเวลาในการทำงานวิจัย	29
4.1 แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Decision tree (16 ตัวแปร)	33
4.2 แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Decision tree (15 ตัวแปร)	34
4.3 แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Decision tree (14 ตัวแปร)	36
4.4 แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Decision tree (13 ตัวแปร)	38
4.5 แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Decision tree (12 ตัวแปร)	40
4.6 แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Decision tree (11 ตัวแปร)	42
4.7 แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Decision tree (10 ตัวแปร)	44
4.8 แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Decision tree (9 ตัวแปร)	46

4.9	แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Decision tree (8 ตัวแปร)	48
4.10	สรุปค่า Accuracy และ F – measure ของแบบจำลองพยากรณ์	49
4.11	แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Naïve Bayes (16 ตัวแปร)	50
4.12	แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Naïve Bayes (15 ตัวแปร)	51
4.13	แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Naïve Bayes (14 ตัวแปร)	53
4.14	แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Naïve Bayes (13 ตัวแปร)	55
4.15	แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Naïve Bayes (12 ตัวแปร)	57
4.16	แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Naïve Bayes (11 ตัวแปร)	59
4.17	แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Naïve Bayes (10 ตัวแปร)	61
4.18	แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Naïve Bayes (9 ตัวแปร)	62
4.19	แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Naïve Bayes (8 ตัวแปร)	63
4.20	สรุปค่า Accuracy และ F – measure ของการทำ Naïve Bayes	64
4.21	แสดงการจัดกลุ่มประชากรตามคุณลักษณะข้อมูลตาม Decision tree โดยวิธี K - mean	65
4.22	แสดงการจัดกลุ่มประชากรตามคุณลักษณะข้อมูลตาม Naïve Bayes โดยวิธี K - mean	66
4.23	แสดงการเปรียบเทียบค่า Accuracy และ F – measure โดยวิธี Decision tree กับ Naïve Bayes	67

สารบัญภาพ

ภาพที่	หน้า
1.1 แสดงข้อมูลการเกิดอุบัติเหตุและเสียชีวิตในปี 2558 – 2560	2
2.1 การเกิดอุบัติเหตุตามแนวคิดทฤษฎีโดมิโน	9
2.2 แสดงการเปรียบเทียบระหว่าง Traditional programming กับ การเรียนรู้ของเครื่อง	12
2.3 แสดงกระบวนการ CRISP-DM	14
2.4 แสดงตัวอย่างการแบ่งข้อมูลแบบ Self-consistency Test	17
2.5 แสดงตัวอย่างการแบ่งข้อมูลแบบ Split Test	18
2.6 แสดงตัวอย่างการแบ่งข้อมูลแบบ Cross-validation Test	18
2.7 แสดงรูปแบบวิธีการคัดเลือกคุณลักษณะ	20
2.8 แผนภาพแสดงกรอบงานวิจัย	23
3.1 แสดงขั้นตอนการออกแบบจำลองพยากรณ์	28

บทที่ 1

บทนำ

1.1 ที่มาและความสำคัญของปัญหา

ปัจจุบันปัญหาเรื่องอุบัติเหตุทางจราจรทางบกถือว่าเป็นปัญหาใหญ่ที่มีความสำคัญเป็นอย่างมาก โดยการเสียชีวิตจากอุบัติเหตุทางจราจรทางบกถือว่าเป็นอันดับ 9 ของโลก ซึ่งผู้ที่เสียชีวิตบนท้องถนนเกิดกับคนเดินเท้า ผู้ขี่จักรยาน และผู้ขี่จักรยานยนต์ และจากการเปิดเผยของเว็บไซต์เวิลด์แอตลาสในปี 2560 ประเทศที่มีผู้เสียชีวิตจากอุบัติเหตุมากที่สุดคือ ประเทศไทย มีอัตราการเสียชีวิตอยู่ที่ 36.2 รายต่อแสนประชากร โดยสาเหตุของการเกิดอุบัติเหตุนั้นมีสาเหตุมาจากหลายอย่าง ตัวอย่างเช่น สาเหตุจากบุคคลที่เมาแล้วขับขี่ยานพาหนะ สาเหตุจากสิ่งแวดล้อมที่ไม่เอื้ออำนวยหรือประเภทของยานพาหนะก็มีส่วนทำให้เกิดอุบัติเหตุ และเกิดการสูญเสียของชีวิต จากข้อมูลของสำนักงานสถิติแห่งชาติ หรือ สสช. ได้ทำการเก็บการเกิดอุบัติเหตุทั่วราชอาณาจักรในช่วงปี 2549 – 2558 มีปัจจัยที่ทำให้เกิดอุบัติเหตุอยู่ 3 อย่าง คือ 1.ปัจจัยจากบุคคล 2.ปัจจัยจากสิ่งแวดล้อม 3.ปัจจัยจากอุปกรณ์ที่ใช้ในการขับขี่ และมีแนวโน้มที่จะเพิ่มมากขึ้นทุกปี ซึ่งสอดคล้องกับข้อมูลที่เว็บไซต์เวิลด์แอตลาสเปิดเผยออกมา.

ตารางที่ 1.1

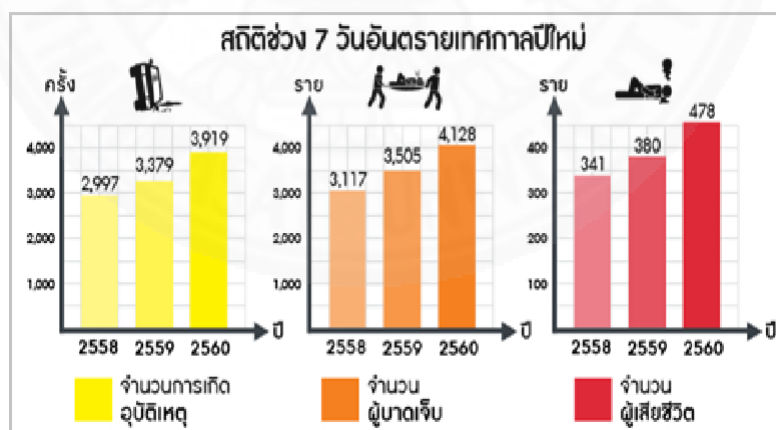
แสดงการเกิดอุบัติเหตุทางถนนสาเหตุจากบุคคลในปี 2549 – 2558

สาเหตุการเกิดอุบัติเหตุ	2549	2550	2551	2552	2553	2554	2555	2556	2557	2558
สาเหตุจากบุคคล	34,670	39,652	33,477	36,753	39,733	46,244	45,036	43,193	41,580	43,423
สาเหตุจากสิ่งแวดล้อม	9,828	9,960	8,679	11,129	12,626	18,440	17,530	18,028	14,946	14,554
สาเหตุจากอุปกรณ์ที่ใช้ขับขี่	6,369	6,580	6,711	8,557	9,037	13,610	12,411	10,557	9,885	10,000

ที่มา: <http://service.nso.go.th>

จากตารางที่ 1.1 ข้อมูลการเกิดอุบัติเหตุจะแสดงให้เห็นว่าการเกิดอุบัติเหตุในประเทศไทยตลอดทั้งปีมีจำนวนที่ค่อนข้างสูง แต่ถ้าหากลงรายละเอียดจำนวนการเกิดอุบัติเหตุจะพบว่าช่วงเวลาส่วนใหญ่ที่เกิดอุบัติเหตุเป็นจำนวนมากจะพบว่าเป็นช่วงเทศกาลที่ประชาชนนิยมกลับภูมิลำเนาเป็นส่วนใหญ่ซึ่งก็คือ ช่วงเทศกาลปีใหม่และสงกรานต์ จากข้อมูลของศูนย์อำนวยการความปลอดภัยทางถนนในปี 2561 ในช่วง 28 ธันวาคม 2560 ถึง 3 มกราคม 2561 มีผู้บาดเจ็บจำนวน 4,068 คน

แม้ว่าทางคณะกรรมการจัดทำแผนและมาตรการลดการสูญเสียจากภัยบนท้องถนน สภานิติบัญญัติแห่งชาติได้เปิดตัวโครงการ “หยุดตายบนท้องถนน ด้วยอาสาตางจรจร” โดยเป็นมาตรการที่ใช้เทคโนโลยีอย่างเช่นกล้องหน้ารถยนต์ผสานกับการสื่อสารโซเชียลมีเดียและประชาชนคอยช่วยกันสอดส่องดูแลผู้กระทำความผิดกฎหมายจราจร หรือแม้กระทั่งศูนย์อำนวยการความปลอดภัยทางถนน (ศปถ.) ได้มีแผนป้องกันและลดอุบัติเหตุทางถนนในช่วงเทศกาลปีใหม่ “ขับรถมีน้ำใจ รักษาวินัยจราจร” โดยมุ่งขับเคลื่อน 6 มาตรการลดปัจจัยเสี่ยงอุบัติเหตุทางถนน ได้แก่ 1.มาตรการลดปัจจัยเสี่ยงด้านคน 2. มาตรการลดปัจจัยเสี่ยงด้านถนนและสิ่งแวดลอม 3.มาตรการลดปัจจัยเสี่ยงด้านยานพาหนะ 4.มาตรการดูแลความปลอดภัยแก่นักท่องเที่ยว 5.มาตรการด้านความปลอดภัยทางน้ำ 6. มาตรการด้านการช่วยเหลือหลังเกิดอุบัติเหตุ แต่การเกิดอุบัติเหตุในช่วงเทศกาลปีใหม่นี้ก็ยังมีแนวโน้มเพิ่มมากขึ้นทุกปี



ภาพที่ 1.1 แสดงข้อมูลการเกิดอุบัติเหตุและเสียชีวิตในปี 2558 – 2560

ที่มา: <http://www.bltbangkok.com>

ดังนั้นผู้ทําวิจัยจึงเล็งเห็นว่าปัญหาอุบัติเหตุจราจรทางบกในช่วงเทศกาลปีใหม่ของไทยเป็นปัญหาที่มีความสำคัญยิ่งนัก ซึ่งเป็นสาเหตุของการสูญเสียทางด้านชีวิตและทรัพย์สิน ซึ่งบุคคลากร

ในประเทศชาติถือว่าเป็นส่วนสำคัญในการพัฒนาประเทศ นอกจากนี้รัฐบาลยังต้องสูญเสียงบประมาณในการบริหารประเทศเพื่อมารณรงค์ ป้องกัน แก้ไขปัญหา และรักษาผู้บาดเจ็บจากอุบัติเหตุที่นับวันจะยังมีจำนวนเพิ่มมากขึ้นเรื่อย ๆ อีกทั้งปัญหาการเกิดอุบัติเหตุจากรางทางบกในช่วงเทศกาลปีใหม่ 7 วันระว่างอันตราย มีแนวโน้มที่จะทวีความรุนแรงมากขึ้นเรื่อยๆ ซึ่งปัญหานี้หน่วยงานเพียงหน่วยงานเดียวไม่สามารถแก้ไขปัญหาได้ อันต้องเกิดจากการร่วมมือของทุกภาคฝ่ายเพื่อช่วยลดหรือบรรเทาปัญหานี้ ผู้ทำวิจัยเล็งเห็นความสำคัญของปัญหานี้ จึงเกิดแนวความคิดการทำวิจัยเรื่อง “การพยากรณ์การเกิดอุบัติเหตุจากรางทางบกในช่วงเทศกาลปีใหม่ 7 วันระว่างอันตราย” เพื่อที่จะหาวิธีป้องกัน บรรเทา ของการสูญเสียชีวิตและทรัพย์สิน

1.2 วัตถุประสงค์ของการวิจัย

1. เพื่อศึกษาปัจจัยที่ทำให้เสียชีวิตจากอุบัติเหตุทางท้องถนนในช่วงเทศกาลปีใหม่โดยใช้ Machine learning
2. เพื่อศึกษาถึงลักษณะของประชากรที่มีโอกาสที่จะเสียชีวิตบนท้องถนนในเทศกาลปีใหม่
3. เพื่อเป็นแนวทางในการช่วยลดการเกิดอุบัติเหตุและการสูญเสียที่เกิดจากอุบัติเหตุ

1.3 ขอบเขตของการวิจัย

1. ขอบเขตของพื้นที่ที่ทำการศึกษา คือ อุบัติเหตุที่เกิดขึ้นทั่วประเทศไทย โดยอาศัยข้อมูลของสถาบันการแพทย์ฉุกเฉินแห่งชาติ ช่วงปี พ.ศ. 2551 - 2558
2. ศึกษาแนวทางพร้อมสร้างตัวแบบการพยากรณ์การเกิดอุบัติเหตุในช่วง 7 วันระว่างอันตรายของเทศกาลปีใหม่ โดยใช้ข้อมูลปี 2551 ถึง 2558

1.4 ขอบเขตด้านประชากรและกลุ่มตัวอย่าง

ข้อมูลการเกิดอุบัติเหตุทำการเก็บข้อมูลแบบทุติยภูมิ จากข้อมูลที่ยินยอมให้การเปิดเผยจากสถาบันการแพทย์ฉุกเฉินแห่งชาติในช่วงระยะเวลา พ.ศ. 2551 – 2558 หรือ 8 ปี โดยข้อมูลที่นำมาทำการวิจัยนั้นเป็นข้อมูลที่เกิดขึ้นทั่วประเทศไทยในช่วงเทศกาลปีใหม่

1.5 ระเบียบวิธีวิจัย

การวิจัยครั้งนี้เป็นการวิจัยเชิงปริมาณ (Quantitative Research) โดยใช้ข้อมูลแบบทุติยภูมิ (Secondary Data) คือนำมาจากสถาบันการแพทย์ฉุกเฉินแห่งชาติ ผนวกกับข้อมูลด้านสภาพอากาศจาก <https://www.wunderground.com> เพื่อนำมาทำการวิเคราะห์เพื่อหาวิธีป้องกันการเกิดอุบัติเหตุ และลดการสูญเสียจากอุบัติเหตุในช่วงเทศกาลปีใหม่ โดยมีการนำเรื่องของการเรียนรู้ของเครื่อง (Machine Learning) มาช่วยในการสร้างตัวแบบการพยากรณ์ เพื่อที่จะนำผลลัพธ์ที่ได้ไปใช้ในการออกมาตรการป้องกัน และการรณรงค์ของทางภาครัฐให้ตรงจุดกับปัญหาการเกิดอุบัติเหตุ

1.6 ประโยชน์ที่คาดว่าจะได้รับ

1. สามารถช่วยลดการสูญเสียชีวิตและทรัพย์สินอันเกิดจากอุบัติเหตุในช่วงเทศกาลปีใหม่
2. ทางภาครัฐและหน่วยงานที่เกี่ยวข้องเกี่ยวกับการลดอุบัติเหตุในช่วงเทศกาลปีใหม่นำผลการวิจัยไปใช้ เพื่อออกมาตรการ การรณรงค์ หรือเป็นแนวทางในการลดการเกิดอุบัติเหตุและลดการสูญเสีย
3. ภาครัฐสามารถใช้งบประมาณที่จัดสรรไว้สำหรับการดูแลในช่วงเทศกาลปีใหม่ได้อย่างคุ้มค่ามากที่สุด
4. สามารถใช้เป็นแนวทางให้สำหรับงานวิจัยอื่นที่จะทำการวิจัยเกี่ยวกับการเกิดอุบัติเหตุโดยใช้เครื่องมือการเรียนรู้ของเครื่อง (Machine Learning) ในงานวิจัยอื่นๆต่อไป

1.7 นิยามศัพท์เฉพาะ

1. อุบัติเหตุจราจร หมายถึง อุบัติเหตุที่เกิดจากการเดินทางบนท้องถนนที่เกิดขึ้นและมีส่วนเกี่ยวข้องกับผู้เดินถนน ผู้ขับขี่ และผู้โดยสารยานพาหนะทุกประเภท
2. สาเหตุการเกิดอุบัติเหตุ หมายถึง เหตุการณ์ สถานการณ์ การกระทำ หรือด้วยปัจจัยด้านต่างๆ ที่ได้เกิดขึ้นหรือกระทำที่ทำให้เกิดขึ้นก่อนหรือขณะเกิดอุบัติเหตุในการขับขี่ยานพาหนะ

บทที่ 2

แนวคิด ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

การศึกษาเรื่อง “การพยากรณ์การเสียชีวิตจากอุบัติเหตุทางถนนในช่วงเทศกาลปีใหม่” ในครั้งนี้ทางผู้ทำการวิจัยได้ทำการศึกษา แนวคิด ทฤษฎีและงานวิจัยที่เกี่ยวข้อง เพื่อที่จะนำมาเป็นกรอบแนวคิดในการทำงานวิจัยดังต่อไปนี้

- 2.1 แนวคิดและทฤษฎีเกี่ยวกับการเกิดอุบัติเหตุ
- 2.2 หลักการและทฤษฎีเกี่ยวกับการเรียนรู้ของเครื่อง (Machine Learning)
- 2.3 เอกสารและงานวิจัยที่เกี่ยวข้อง
- 2.4 กระบวนการในการวิจัย

2.1 แนวคิดและทฤษฎีเกี่ยวกับการเกิดอุบัติเหตุ

2.1.1 ความหมาย

ในส่วนของความหมายของ “อุบัติเหตุ” มีนักวิชาการหลายท่านที่มีการให้นิยามความหมายของอุบัติเหตุดังต่อไปนี้ วิจิตร บุญยะโทตระ (2527) ให้ความหมายไว้ว่า อุบัติเหตุเป็นเหตุการณ์ที่เกิดขึ้นโดยไม่ได้คาดคิดล่วงหน้า โดยการเกิดอุบัติเหตุ นั้นก่อให้เกิดผลเสียหายหลายอย่างตามมา ได้แก่ ร่างกาย จิตใจ ทรัพย์สิน และชีวิต

คณะกรรมการป้องกันอุบัติเหตุแห่งชาติ (2529) ให้นิยามความหมายไว้ว่า อุบัติเหตุเป็นเหตุการณ์ที่ทุกคนไม่อยากจะเกิดขึ้นทั้งต่อตัวเอง ญาติพี่น้อง บุคคลในครอบครัว และมิตรสหาย ซึ่งการเกิดอุบัติเหตุ นั้นเกิดจากการที่ไม่เจตนาให้เกิดขึ้น แต่เกิดจากผลพวงของการประมาท เลินเล่อ ไม่มีสติ และการขาดจิตสำนึกด้านความปลอดภัย

Thygerson (1977) ให้นิยามความหมายไว้ว่า อุบัติเหตุเป็นเหตุการณ์ที่ไม่มีใครคาดคิดที่จะให้เกิดขึ้นหรือปรารถนาที่จะให้เกิดขึ้น แต่เกิดจากการกระทำของบุคคลอื่นเกิดจากความประมาท การกระทำสิ่งที่ก่อให้เกิดความไม่ปลอดภัย ซึ่งก่อให้เกิดผลกระทบที่ไม่ต้องการหรือพึงประสงค์

Stresser (1976) ให้นิยามความหมายไว้ว่า อุบัติเหตุเป็นเหตุการณ์ที่เกิดจากความไม่ตั้งใจให้เกิดหรือคาดหวังว่าจะให้เกิด แต่อุบัติเหตุเกิดมาจากการกระทำที่ไม่คำนึงถึงความปลอดภัยจนส่งผลต่อการสูญเสียทรัพย์สินและชีวิต

2.1.2 ประเภทของอุบัติเหตุ

ประเภทของอุบัติเหตุสามารถแบ่งออกมาได้ 4 ประเภท ได้แก่

1. อุบัติเหตุในบ้าน ได้แก่ การลื่นล้ม การถูกมีดบาด ไฟไหม้ ไฟฟ้าดูด น้ำร้อน ลวก เป็นต้น
2. อุบัติเหตุจากการทำงาน ได้แก่ การเกิดอุบัติเหตุในที่ทำงาน และนอกที่ทำงาน
3. อุบัติเหตุในที่สาธารณะ ได้แก่ การพลัดตกรถสาธารณะ การพลัดตกน้ำ การเล่นกีฬา เป็นต้น
4. อุบัติเหตุจากการจราจร ได้แก่ ทางบก ทางน้ำ และทางอากาศ โดยอุบัติเหตุที่มีจำนวนมากที่สุดในทางปริมาณคือ อุบัติเหตุจากการจราจรทางบก

2.1.3 ความสูญเสียจากอุบัติเหตุ

ความสูญเสียจากอุบัติเหตุสามารถแบ่งออกมาได้ 2 ประเภท ได้แก่ ความสูญเสียทางตรง และความสูญเสียทางอ้อม

2.1.3.1 ความสูญเสียทางตรง (Direct Loss) คือ ค่าใช้จ่ายต่างๆที่เกิดจากอุบัติเหตุ เช่น ค่ารักษาพยาบาล ค่าดูแลผู้บาดเจ็บ ค่าใช้จ่ายชดเชย ค่าทำศพ เป็นต้น

2.1.3.2 ความสูญเสียทางอ้อม (Indirect Loss) คือ การสูญเสียที่ไม่อาจมองเห็นได้ชัดเจน ได้แก่

1. ทางด้านร่างกาย เช่น การตาย การเจ็บป่วย การพิการ เป็นต้น
2. ทางด้านจิตใจ เช่น สภาพจิตใจ อารมณ์ความรู้สึก กำลังใจ การเสียชีวิต เป็นต้น
3. ทางด้านเศรษฐกิจ เช่น การที่ผู้ประสบอุบัติเหตุไม่สามารถไปทำงานได้ตามปกติ การขาดรายได้ของผู้ประสบอุบัติเหตุ ค่าใช้จ่ายต่างๆที่ตามมาหลังการเกิดอุบัติเหตุ เป็นต้น
4. ด้านสังคม เช่น เรื่องความสัมพันธ์รอบตัวของผู้ประสบอุบัติเหตุ หรือแม้กระทั่งการสูญเสียอวัยวะ เป็นต้น

2.1.4 สาเหตุที่ก่อให้เกิดอุบัติเหตุจราจร

สาเหตุที่ก่อให้เกิดอุบัติเหตุจราจรสามารถแบ่งได้ 4 ปัจจัยหลัก คือ การกระทำที่ไม่ปลอดภัยของผู้ขับขี่ สภาพยานพาหนะที่ไม่ปลอดภัย สภาพสิ่งแวดล้อมที่ไม่ปลอดภัย และเหตุสุดวิสัย

2.1.4.1 การกระทำที่ไม่ปลอดภัยของผู้ขับขี่ ประกอบด้วยลักษณะส่วนบุคคลของผู้ขับขี่ และความพร้อมของผู้ขับขี่

2.1.4.2 ลักษณะส่วนบุคคลของผู้ขับขี่ มักเป็นปัจจัยหลักของการเกิดอุบัติเหตุ เนื่องจากผู้ขับขี่เป็นผู้ควบคุมยานพาหนะ โดยปัจจัยที่เกี่ยวข้องกับลักษณะส่วนตัวของผู้ขับขี่ ได้แก่ เพศ อายุ สถานภาพสมรส ระดับการศึกษา อาชีพ รายได้ ประสบการณ์ในการขับขี่ และประวัติการได้รับอุบัติเหตุจากการขับขี่

1. ความพร้อมของผู้ขับขี่รถจักรยานยนต์ ประกอบด้วย สภาพร่างกาย และสภาวะทางจิต

2. สภาพร่างกาย คือ ความพร้อมของร่างกายในการขับขี่ยานพาหนะ มีปัจจัยที่ส่งผลต่อความพร้อมของสภาวะร่างกายดังนี้ แอลกอฮอล์ ยา ความเหนื่อยล้า โรคร ความผิดปกติทางหูมีส่วนเกี่ยวข้องกับการเกิดอุบัติเหตุ และความผิดปกติทางสายตา

3. สภาวะทางจิต ได้แก่ ความเครียด ความกดดัน อารมณ์ที่มีความวิตกกังวล สภาวะทางจิตนั้นอาจก่อให้เกิดอุบัติเหตุได้ เช่น การกะเวลาในการเดินทางไม่เหมาะสม กับระยะทาง ทำให้เกิดการชะล่าใจ หรือขาดการระมัดระวัง จนก่อให้เกิดการฝ่าฝืนทางกฎหมายและขาดประสิทธิภาพในการควบคุมยานพาหนะ

2.1.4.3 สภาพยานพาหนะที่ไม่ปลอดภัย คือ สภาพของยานพาหนะที่ก่อให้เกิดอุบัติเหตุสามารถแบ่งได้ดังนี้ คือ ยานพาหนะที่มีสภาพชำรุดบกพร่อง ยานพาหนะไม่เป็นไปตามมาตรฐานความปลอดภัย

2.1.4.4 ยานพาหนะที่มีสภาพชำรุดหรือบกพร่องอันเนื่องมาจากอุปกรณ์ต่างๆ เช่น สภาพของยาง ระบบเบรค ระบบไฟ คันบังคับ

2.1.4.5 ยานพาหนะไม่เป็นไปตามมาตรฐานความปลอดภัย เช่น การใช้เครื่องยนต์ที่ผิดประเภท น้ำหนักของยานพาหนะ ตัวถัง และไม่มีอุปกรณ์รักษาความปลอดภัย

2.1.5 สภาพแวดล้อมบนถนนที่ไม่ปลอดภัย

ประกอบด้วยปัจจัยหลัก 4 ปัจจัย คือ ลักษณะทางกายภาพของถนน ทางชีวสังคม ทางเคมี และทางธรรมชาติ

2.1.5.1 ลักษณะทางกายภาพของถนน ประกอบด้วยปัจจัยดังต่อไปนี้ พื้นผิวถนน จำนวนช่องเดินรถ ความกว้างของช่องเดินรถ แนวกั้นกลางถนน ไหล่ทาง เครื่องกั้นทาง แสงสว่าง และทางแยกที่ยังไม่มีมาตรฐานการออกแบบที่แน่นอน

2.1.5.2 ทางชีวสังคม เป็นสิ่งที่ต้องอาศัยความร่วมมือจากผู้ใช้ถนนทั้งถนน เช่น การมีน้ำใจในการขับขี่ยานพาหนะ และคนเดินเท้าบนท้องถนนมักได้รับอันตรายจากอุบัติเหตุบ่อย และมากที่สุดเนื่องจากไม่มีเครื่องป้องกันอันตรายจากอุบัติเหตุ ในทางกลับกันบางทีคนเดินเท้าบนท้องถนนก็อาจมีส่วนร่วมทำให้เกิดอุบัติเหตุด้วยเช่นกัน เช่น วิ่งข้ามถนนตัดหน้ายานพาหนะที่กำลังใช้งานบนท้องถนน

2.1.5.3 ทางเคมี มักเป็นสิ่งที่เกิดจากการกระทำของคน เช่น คว้นจากท่อไอเสีย คว้นจากการเผาสิ่งของข้างทาง หรือเสียงจากเครื่องยนต์ที่ทำลายสมาธิการขับขี่ เป็นต้น

2.1.5.4 ทางธรรมชาติ ในบางครั้งธรรมชาติก็ก่อให้เกิดอุบัติเหตุบนท้องถนนได้เช่นกัน เช่น ลูกเห็บ หมอก หรือน้ำท่วม

2.1.6 เหตุสุวิสัย

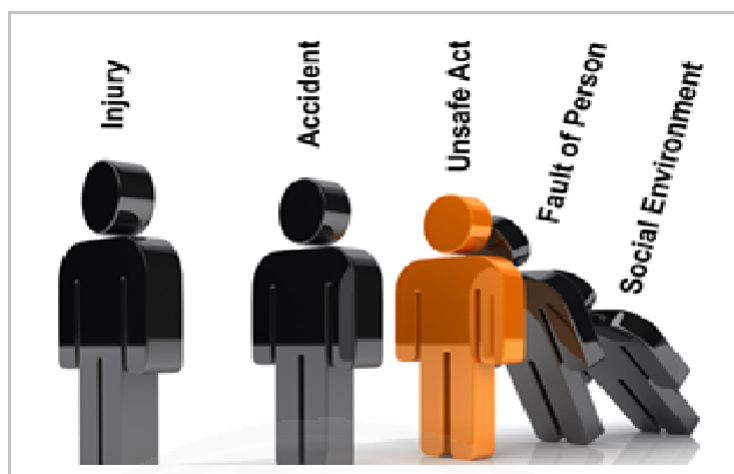
ในบางครั้งอาจมีเหตุการณ์ที่ไม่คาดคิดว่าจะสามารถเป็นต้นเหตุของการเกิดอุบัติเหตุได้ เช่น สุนัขวิ่งตัดหน้ารถยนต์

2.1.7 ทฤษฎีอุบัติเหตุเกี่ยวข้องกับงานวิจัย

2.1.7.1 แนวคิดทฤษฎีโดมิโน (Domino theory)

ผู้คิดทฤษฎีนี้คือ เฮ็นริช ในปี 1959 โดยนายเฮ็นริชได้กล่าวว่า “การกระทำที่ไม่ปลอดภัยเป็นสาเหตุที่สำคัญในการเกิดอุบัติเหตุ สาเหตุของการเกิดอุบัติเหตุเกิดจากการกระทำของคนมากกว่าสภาวะของแวดล้อมหรือสภาพการทำงาน” ดังนั้นจึงมีการเปรียบเทียบการกระทำที่ไม่ปลอดภัยให้เหมือนกับตัวโดมิโนทั้ง 5 ตัว ซึ่งหากมีตัวใดตัวหนึ่งล้มย่อมส่งต่อตัวที่อยู่ถัดไป โดยที่โดมิโนทั้ง 5 ตัว ประกอบด้วยดังต่อไปนี้

1. สภาพแวดล้อมหรือภูมิหลังของบุคคล (Social Environment of Background) ได้แก่ สภาพครอบครัว ฐานะความเป็นอยู่การศึกษา เป็นต้น
2. ความบกพร่องผิดปกติของบุคคล (Defect of Person) ได้แก่ ได้แก่ การปฏิบัติงานโดยขาดความยั้งคิด อารมณ์รุนแรง การขาดความรอบคอบ หรือการเพิกเฉยต่อความปลอดภัย เป็นต้น
3. การกระทำหรือสภาพการณ์ที่ไม่ปลอดภัย (Unsafe Acts/Unsafe Conditions) ได้แก่ ขาดเครื่องป้องกันจุดอันตราย แสงสว่างไม่เพียงพอ หรือมีสารเคมีฟุ้งกระจาย เป็นต้น
4. อุบัติภัย (Accident) คือเหตุการณ์ที่ไม่คาดคิดว่าจะเกิดขึ้น ได้แก่ ลื่น หกล้ม เดินสะดุด สิ่งของพลัดหล่นจากที่สูง ไฟไหม้ เป็นต้น
5. การบาดเจ็บหรือเสียหาย (Injury or Damage) ได้แก่ การบาดเจ็บ หรือสูญเสียอวัยวะ เช่น กระดูกหักหรือแตก ได้รับบาดเจ็บ หรือแม้กระทั่งถึงแก่ความตาย เป็นต้น



ภาพที่ 2.1 การเกิดอุบัติเหตุตามแนวคิดทฤษฎีโดมิโน

ที่มา: <https://www.scglogistics.co.th/blog/detail/209>

2.1.7.2 ทฤษฎีความเอนเอียงในการเกิดอุบัติเหตุ (Accident-proneness)

ปี 1919 กรีนวูด และ วูดส์ (Wood and Greenwood) ได้ทำการศึกษาการเกิดอุบัติเหตุของพนักงานในประเทศอังกฤษ โดยอธิบายถึงเหตุการณ์การเกิดอุบัติเหตุที่บ่อยๆ หรือซ้ำๆ คือ การที่บางคนเกิดอุบัติเหตุ มากกว่าคนอื่นๆ แต่ความโน้มเอียงที่จะทำให้เกิดอุบัติเหตุอธิบายถึงว่า “ทำไมบุคคลนั้นจึงเกิดอุบัติเหตุมากกว่าคนอื่นๆ” โดยความโน้มเอียงที่จะทำให้บุคคลเกิดอุบัติเหตุ สามารถแบ่งออกเป็น 2 ประเภทได้ดังนี้

1. ความโน้มเอียงที่จะเกิดอุบัติเหตุมีกำหนดเวลาในช่วงสั้นๆ โดยเกิดผลในระยะวิกฤต เช่น ความเครียด ความเหนื่อยล้า ซึ่งเหตุการณ์เหล่านี้จะเกิดในระยะเวลาเพียงสั้นๆ หรือชั่วขณะ และเมื่อสภาวะเหล่านั้นผ่านไปบุคคลก็จะปรับตัวเข้าสู่สภาวะปกติ

2. ความโน้มเอียงที่จะทำให้เกิดอุบัติเหตุของแต่ละบุคคล มีสาเหตุใหญ่มาจากแหล่งภายในที่ประกอบด้วย บุคลิก สภาพจิตใจและสภาพร่างกาย

ในปี 1926 Karl Marbe นักจิตวิทยาชาวเยอรมันได้เสนอความโน้มเอียงในการเกิดอุบัติเหตุ อันเป็นลักษณะบุคลิกภาพที่มีแนวโน้มจะทำให้บุคคลได้รับอุบัติเหตุ โดยแนวคิดนี้ได้ทำการศึกษาลักษณะธรรมชาติของบุคคลที่มีส่วนหรือเป็นสาเหตุทำให้เกิดอุบัติเหตุได้ ซึ่งแยกประเภทบุคคลนี้อาจเรียกว่า เป็นปัจจัยซึ่งเป็นสาเหตุของการเกิดอุบัติเหตุไว้เป็น 2 กลุ่ม ดังนี้ ผู้ที่มีลักษณะ X ได้แก่ ผู้มีความโน้มเอียงที่จะไม่เกิดอุบัติเหตุ และผู้ที่มีลักษณะ Y ได้แก่ ผู้มีความโน้มเอียงที่จะเกิดอุบัติเหตุ

ตารางที่ 2.1

ประเภทของบุคคลตามลักษณะพื้นฐาน 2 ประเภท

บุคคลประเภทเอ็กซ์ (Type X) มีความเอนเอียงที่จะไม่เกิดอุบัติเหตุ(Non-Accident-Prone)	บุคคลประเภทวาย (Type Y) มีความเอนเอียงที่จะเกิดอุบัติเหตุ(Accident-Prone)
1. มีระเบียบแบบแผน	1. ไม่มีระเบียบแบบแผน
2. ผู้ที่มีเป้าหมายในการดำรงชีวิต	2. ผู้ที่ไม่มีเป้าหมายในการดำรงชีวิต
3. ผู้ที่พอใจในชีวิตประจำวัน	3. ผู้ที่ไม่พอใจในชีวิตประจำวัน
4. ผู้ที่เคารพสิทธิและความคิดเห็นของผู้อื่น	4. ผู้ที่ไม่สนใจในสิทธิและความคิดเห็นของผู้อื่น
5. ผู้ที่ไม่เผด็จการ	5. ผู้ที่ไม่มีมนุษยสัมพันธ์
6. ผู้ที่ไม่ชอบโต้เถียงหรือทะเลาะวิวาท	6. ผู้ที่ระงับอารมณ์ ความรู้สึกเกลียดชังยาก
7. ผู้ที่นึกถึงผู้อื่น	7. ผู้ที่นึกถึงแต่ตัวเอง

2.1.8 ชุดปฏิบัติการฉุกเฉิน

ชุดปฏิบัติการด้านการแพทย์ฉุกเฉิน การรับรู้ถึงภาวะการณเจ็บป่วยฉุกเฉินจนถึงการดำ เนินการให้ผู้ป่วยฉุกเฉินได้รับการบำบัดรักษาให้พ้นภาวะฉุกเฉิน ซึ่งรวมถึงการประเมิน การจัดการ การประสานงาน การควบคุม ดูแล การติดต่อสื่อสาร การลำเลียงหรือขนส่งผู้ป่วย การตรวจวินิจฉัย และการบำบัดรักษาพยาบาล ผู้ป่วยฉุกเฉินทั้งนอกสถานพยาบาลและในสถานพยาบาลสามารถแบ่งออกเป็น 4 ชุดปฏิบัติการได้ดังนี้

1. ชุดปฏิบัติการฉุกเฉินเบื้องต้น (First response unit : FR)

เป็นชุดปฏิบัติการฉุกเฉินที่มีความสามารถในการประเมินและปฐมพยาบาล ได้แก่ การตามกระตุก การห้ามเลือด การช่วยฟื้นคืนชีพขั้นต้น รวมถึงการบริหารยาสามัญประจำบ้าน และการเคลื่อนย้ายผู้ป่วยอย่างถูกวิธี สื่อสารประสานงานกับศูนย์รับแจ้งเหตุ สั่งการจังหวัดและร่วมปฏิบัติงานกับหน่วยปฏิบัติการในระดับที่สูงกว่า โดยไม่มีการปฏิบัติที่จัดอยู่ในขอบเขตของการประกอบวิชาชีพเวชกรรม โดยอยู่ภายใต้การควบคุมของแพทย์หรือพยาบาลประจำศูนย์รับแจ้งเหตุ และสั่งการจังหวัดตามที่กฎหมายกำหนด ตลอดจนมีทักษะการสื่อสารและประสานกับผู้เกี่ยวข้อง เพื่อการจัดการภาวะฉุกเฉินให้มีประสิทธิผล หรือสังกัดอยู่ในองค์กรปกครองส่วนท้องถิ่น หรือหน่วยงานที่องค์กรปกครองส่วนท้องถิ่นมอบหมายให้ดำเนินการอย่างเป็นทางการ หรือหน่วยงานของรัฐอื่นๆ (เช่น ตำรวจดับเพลิง อุทยานแห่งชาติ เป็นต้น) หรือองค์กรอื่นที่สถาบันการแพทย์ฉุกเฉินแห่งชาติรับรอง โดยขึ้นทะเบียนกับสำนักงานระบบการแพทย์ฉุกเฉินจังหวัดสามารถปฏิบัติงานได้ตลอด 24 ชั่วโมง

2. ชุดปฏิบัติการฉุกเฉินระดับต้น (Basic life support unit : BLS)

เป็นชุดปฏิบัติการฉุกเฉินที่มีความสามารถในการประเมินสถานการณ์และสภาพผู้เจ็บป่วยฉุกเฉินให้การปฐมพยาบาล การช่วยฟื้นคืนชีพขั้นพื้นฐาน เคลื่อนย้ายผู้ป่วยอย่างถูกวิธี พร้อมทั้งยังสามารถสื่อสารประสานงานกับศูนย์รับแจ้งเหตุและสั่งการจังหวัดร่วมปฏิบัติงานกับหน่วยปฏิบัติการในระดับที่สูงกว่า และหน่วยงานที่เกี่ยวข้อง โดยขึ้นทะเบียนกับสำนักงานระบบการแพทย์ฉุกเฉินจังหวัด และสามารถปฏิบัติงานได้ตลอด 24 ชั่วโมง

3. ชุดปฏิบัติการฉุกเฉินระดับกลาง (Intermediate Life Support unit : ILS)

เป็นชุดปฏิบัติการฉุกเฉินที่มีความสามารถในการประเมินสถานการณ์และสภาพผู้เจ็บป่วยฉุกเฉินเคลื่อนย้ายผู้ป่วย อย่างถูกวิธี ช่วยเหลือภาวะคุกคามต่อชีวิตขั้นพื้นฐาน สื่อสารประสานงานกับศูนย์รับแจ้งเหตุและสั่งการจังหวัด และร่วมปฏิบัติงานกับหน่วยปฏิบัติการในระดับที่สูงกว่าและหน่วยงานที่เกี่ยวข้อง โดยขึ้นทะเบียนกับสำนักงานระบบการแพทย์ฉุกเฉินจังหวัด สามารถปฏิบัติงานได้ตลอด 24 ชั่วโมง

4. ชุดปฏิบัติการฉุกเฉินระดับสูง (Advanced Life Support Unit : ALS)

เป็นชุดปฏิบัติการฉุกเฉินที่มีความสามารถในการประเมินสถานการณ์และสภาพผู้เจ็บป่วยฉุกเฉินเคลื่อนย้ายผู้ป่วยอย่างถูกวิธี ช่วยเหลือภาวะคุกคามต่อชีวิตขั้นสูง รวมทั้งสามารถสื่อสารประสานงานกับศูนย์รับแจ้งเหตุและสั่งการจังหวัด และร่วมปฏิบัติกับหน่วยงานที่เกี่ยวข้อง โดยขึ้นทะเบียนกับสำนักงานระบบการแพทย์ฉุกเฉินจังหวัด สามารถปฏิบัติงานได้ตลอด 24 ชั่วโมง

2.1.9 แนวคิดทฤษฎีเกี่ยวกับความหมายและกระบวนการนโยบายสาธารณะ (Public Policy)

นักวิชาการมีการให้ความหมายของคำว่า “นโยบายสาธารณะ” ไว้อย่างมากมาย โดยมีทั้งนักวิชาการชาวไทยและชาวต่างประเทศให้ความหมายไว้ดังต่อไปนี้

มยุรี อนุมานราชชน กล่าวว่ นโยบายสาธารณะคือแนวทางกว้าง ๆ ที่รัฐบาลได้เลือกหรือตัดสินใจเลือกโดยกำหนดไว้ล่วงหน้า เพื่อให้มีกิจกรรมหรือการกระทำต่าง ๆ เกิดขึ้น เพื่อให้สิ่งนั้นสามารถบรรลุเป้าหมายที่กำหนดไว้

ถวัลย์ วรเทพพิพิงค์ กล่าวว่ นโยบายสาธารณะเป็นแนวทางรัฐบาลเลือกที่จะปฏิบัติ โดยที่รัฐบาลมีวัตถุประสงค์เพื่อแก้ไขปัญหาซึ่งกำลังเกิดขึ้น หรือป้องกันปัญหาที่อาจจะเกิดขึ้นในอนาคต โดยหวังว่าจะก่อให้เกิดผลลัพธ์ตามที่รัฐบาลคาดหวัง

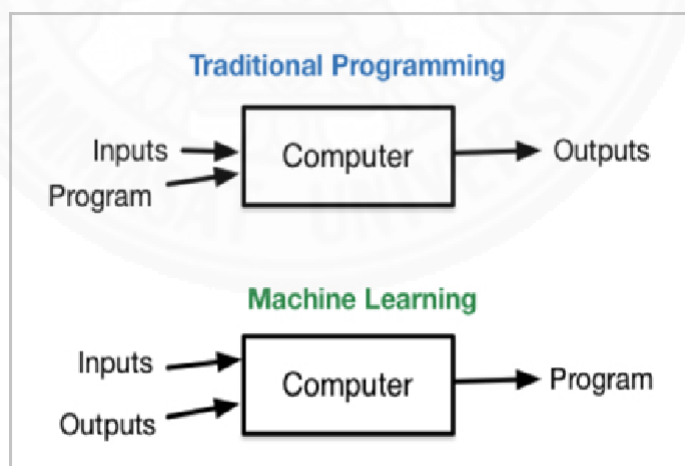
James Anderson ได้ทำการนิยามความหมายของนโยบายสาธารณะ ไว้ว่า คือ แนวทางการกระทำของรัฐโดยหวังที่จะแก้ไขปัญหาสิ่งที่กำลังเกิดขึ้น เช่น ความยากจน การผูกขาดทางอุตสาหกรรม หรือการเกิดอุบัติเหตุ

ดังนั้นจะเห็นได้ว่า “นโยบายสาธารณะ” แม้ว่าจะมีนักวิชาการหลายท่านได้ให้นิยามความไว้หลายท่าน แต่ค่อนข้างจะมีความหมายไปในทิศทางเดียวกัน คือเป็นแนวทางที่ทางภาครัฐออกมากำหนดแนวทางไว้เพื่อที่จะแก้ไขปัญหา หรืออุปสรรคที่ทางรัฐบาลเล็งเห็นว่ากำลังเกิดขึ้น โดยที่หวังว่าจะสามารถที่จะป้องกัน หรือบรรเทาความเดือดร้อนที่จะเกิดขึ้น โดยที่หน่วยงานทางภาครัฐจะต้องเข้ามาดำเนินการหรือควบคุมดูแล เพื่อที่จะให้เกิดผลประโยชน์กับประชาชนให้มากที่สุด

2.2 หลักการและทฤษฎีเกี่ยวกับการเรียนรู้ของเครื่อง (Machine Learning)

2.2.1 ความหมายของการเรียนรู้ของเครื่อง

การเรียนรู้ของเครื่อง (Machine Learning) เป็นสาขาหนึ่งของปัญญาประดิษฐ์ที่พัฒนามาจากการศึกษาการเรียนรู้จำแบบเกี่ยวข้องกับการศึกษาและการสร้างอัลกอริทึมที่สามารถเรียนรู้ข้อมูล และสามารถทำนายข้อมูลได้เมื่อมีข้อมูลเข้าสามารถทำนายหรือตัดสินใจได้โดยปราศจากการทำงานตามลำดับคำสั่งโปรแกรม ซึ่งการเรียนรู้ของเครื่องนั้นเป็นการรวบรวมศาสตร์ไว้อย่างหลากหลายแขนง อาทิเช่น วิทยาการคอมพิวเตอร์ วิศวกรรม และที่สำคัญที่สุดนั่นคือ สถิติ พร้อมทั้งยังสามารถนำไปประยุกต์กับศาสตร์วิชาแขนงอื่นได้เช่นกัน ไม่ว่าจะเป็นชีววิทยา การเงินการธนาคาร เคมี และหลักการตลาด เป็นต้น



ภาพที่ 2.2 แสดงการเปรียบเทียบระหว่าง Traditional programming กับ การเรียนรู้ของเครื่อง
ที่มา: <https://medium.com>

2.2.2 กระบวนการวิเคราะห์ข้อมูลด้วย CRISP-DM

กระบวนการมาตรฐานในการวิเคราะห์ข้อมูลด้วย CRISP-DM พัฒนาขึ้นในปี ค.ศ. 1996 โดยความร่วมมือกันของ 3 บริษัท ได้แก่ DaimlerChrysler SPSS และ NCR เรียกว่า “CRISP-DM” หรือ “Cross-Industry Standard Process for Data Mining” โดยกระบวนการ CRISP-DM จะประกอบด้วย 6 กระบวนการ โดยแต่ละขั้นตอนจะเป็นขั้นตอนที่ต่อเนื่องกัน หมายความว่าผลลัพธ์จากขั้นตอน ณ ขณะนั้นจะเป็นข้อมูลตั้งต้นให้กับขั้นตอนถัดไป

จากภาพ 2.3 สามารถอธิบายกระบวนการได้ดังต่อไปนี้

1. Business Understanding หรือกระบวนการเข้าใจในปัญหาหรือโจทย์ที่ต้องการหาคำตอบหรือคำอธิบายเพื่อที่จะวางแผนในการทำงาน

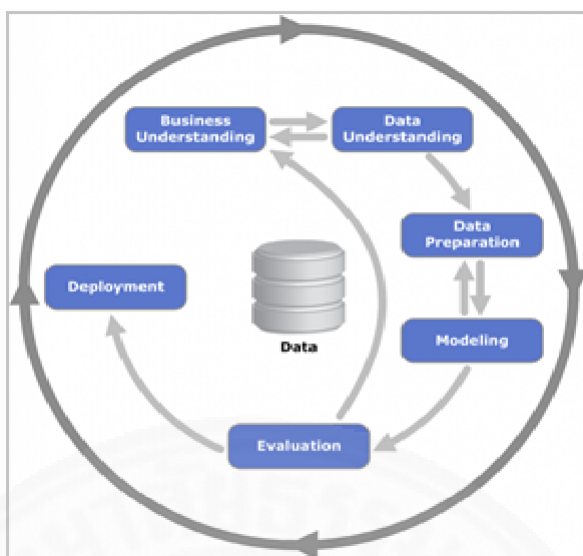
2. Data Understanding หรือกระบวนการเข้าใจในข้อมูลที่มี โดยขั้นตอนนี้จะเป็นขั้นตอนที่ได้จากการเก็บข้อมูล หลังจากนั้นจะต้องพิจารณาว่าข้อมูลที่มีนั้นเพียงพอต่อการนำไปวิเคราะห์ต่อหรือไม่ หรือข้อมูลที่มีนั้นส่วนไหนที่ไม่จำเป็นต้องการนำไปวิเคราะห์

3. Data Preparation หรือกระบวนการเตรียมข้อมูล เป็นขั้นตอนของการแปลงข้อมูลให้พร้อมเพื่อนำไปวิเคราะห์โดยอาจจะต้องมีการทำข้อมูลให้ถูกต้อง (data cleaning) ยกตัวอย่างเช่น การแปลงข้อมูลให้อยู่ในช่วงเดียวกัน การเติมข้อมูลที่ขาดหายไป หรือการทำข้อมูลให้อยู่ในรูปแบบเดียวกัน เป็นต้น ซึ่งขั้นตอนนี้จะเป็นขั้นตอนที่ใช้เวลาในการทำมากที่สุดใน 6 กระบวนการของ CRISP-DM

4. Modeling หรือกระบวนการขั้นตอนวิเคราะห์ข้อมูล เป็นขั้นตอนที่พยายามหาคำตอบโดยการสร้างแบบจำลองขึ้นมาโดยอาศัยเทคนิคต่างๆ ในบางครั้งอาจจำเป็นต้องกลับไปในขั้นตอนการเตรียมข้อมูลอีกครั้ง เพื่อที่ว่าจะได้ตัวแบบจำลองที่เหมาะสมที่สุดกับโจทย์ที่เราต้องการหาคำตอบ

5. Evaluation หรือกระบวนการวัดประสิทธิภาพของแบบจำลองว่ามีความถูกต้องแม่นยำมากน้อยเพียงใดด้วยวิธีการต่างๆ เช่น Self-consistency test , Split test , Cross-validation test เป็นต้น

6. Deployment หรือกระบวนการนำแบบจำลองไปใช้จริงในการแก้ปัญหาจากขั้นตอนที่ 1



ภาพที่ 2.3 แสดงกระบวนการ CRISP-DM

ที่มา: <http://dataminingtrend.com>

2.2.3 ประเภทของการเรียนรู้ของเครื่อง

2.2.3.1 การเรียนรู้แบบมีผู้สอน (Supervised learning)

เป็นเทคนิคการเรียนรู้ของเครื่องโดยการแบ่งชุดข้อมูลที่มีออกเป็น 2 ชุด คือ ชุดข้อมูลที่ใช้สอนให้เครื่องเรียนรู้ (Training Data) และชุดข้อมูลที่ใช้ทดสอบ (Test Data) โดยที่ชุดข้อมูลที่ใช้สอนให้เครื่องเรียนรู้มีเป้าหมายเพื่อที่หาแบบจำลองในการพยากรณ์และหาผลลัพธ์ที่ได้ออกมา ส่วนชุดข้อมูลที่ใช้ทดสอบนั้นมีเป้าหมายเพื่อใช้ในการทดสอบประสิทธิภาพและความถูกต้องของแบบจำลองพยากรณ์ว่ามีความน่าเชื่อถือมากน้อยเพียงใด โดยหลักการของการเรียนรู้แบบมีผู้สอนมีหลายวิธีการ เช่น Decision Tree , Naïve Bayes , Neural Network เป็นต้น

2.2.3.2 การเรียนรู้แบบไม่มีผู้สอน (Unsupervised learning)

เป็นเทคนิคการเรียนรู้อีกรูปแบบหนึ่งของการเรียนรู้ของเครื่องโดยที่ไม่จำเป็นต้องมีการสอน แต่จะเป็นการใช้ชุดข้อมูลที่มีทั้งหมดในการให้เครื่องทำการเรียนรู้ โดยไม่จำเป็นต้องระบุผลที่ต้องการ โดยการเรียนรู้แบบไม่มีผู้สอน นั้นตรงกันข้ามกับการเรียนรู้แบบมีผู้สอน คือการเรียนรู้แบบไม่มีผู้สอนนี้จะไม่มีการระบุผลที่ต้องการไว้ก่อน ให้เครื่องพยายามหาความสัมพันธ์จากข้อมูลเอาเอง โดยการเรียนรู้แบบไม่มีผู้สอนมีหลายวิธีการ เช่น คือ การแบ่งกลุ่มข้อมูล (Clustering) และการหากฎความสัมพันธ์ (Association Rules)

ตารางที่ 2.2

ตารางสรุปการเรียนรู้ของเครื่อง (Machine learning)

Supervised	Classification	<p>Decision Tree คือ เป็นเทคนิคการนำข้อมูลมาสร้างแบบจำลองการพยากรณ์ในรูปแบบโครงสร้างต้นไม้ โดยเทคนิคนี้สามารถสร้างแบบจำลองการจัดหมวดหมู่ได้จากกลุ่มตัวอย่างข้อมูลที่กำหนดไว้ล่วงหน้า และพยากรณ์กลุ่มของรายการที่ยังไม่เคยนำมาจัดหมวดหมู่</p>
		<p>Naïve Bayes คือ เป็นหลักการที่ใช้ความน่าจะเป็นเข้ามาเกี่ยวข้อง โดยหลักความน่าจะเป็นของ Bayes's Theorem และสมมติฐาน จะเป็นการกำหนดให้มีความเป็นอิสระต่อกัน</p> <p>Posterior probability หรือ $P(C A)$ คือ ค่าความน่าจะเป็นที่ข้อมูลที่มิแอตทริบิวต์เป็น A จะมีคลาส C</p> <p>Likelihood หรือ $P(A C)$ คือ ค่าความน่าจะเป็นที่ข้อมูล training data ที่มีคลาส C และมีแอตทริบิวต์ A โดยที่ $A = a_1 \cap a_2 \dots \cap a_M$ โดยที่ M คือจำนวนแอตทริบิวต์ใน training data</p> <p>Prior probability หรือ $P(C)$ คือ ค่าความน่าจะเป็นของคลาส C</p>
		<p>Neural Networks คือ หลักการที่มีแนวคิดมาจากโครงข่ายประสาทเทียมนั้นมาจากเลียนแบบการทำงานของระบบประสาทของมนุษย์ จะจำลองการทำงานเหมือนกลุ่มเซลล์ประสาทที่เชื่อมโยงกัน สามารถรับรู้หลายๆ ได้ในเวลาเดียวกัน โดยจะมีการประมวลผลแบบขนาน (Parallel Network) ทำให้ระบบสามารถตัดสินใจได้ใกล้เคียงกับมนุษย์ โดยที่ Input network คือข้อมูลที่รับจากภายนอก ในขั้นต่อมาคือ Hidden network คือ ทำหน้าที่ในการประเมินชุดข้อมูลจาก Input network แล้วนำส่งต่อไปที่ชั้น Output network และในส่วนสุดท้ายคือ Output network คือ เป็นส่วนที่นำข้อมูลออกมาแสดง</p>

ตารางที่ 2.2

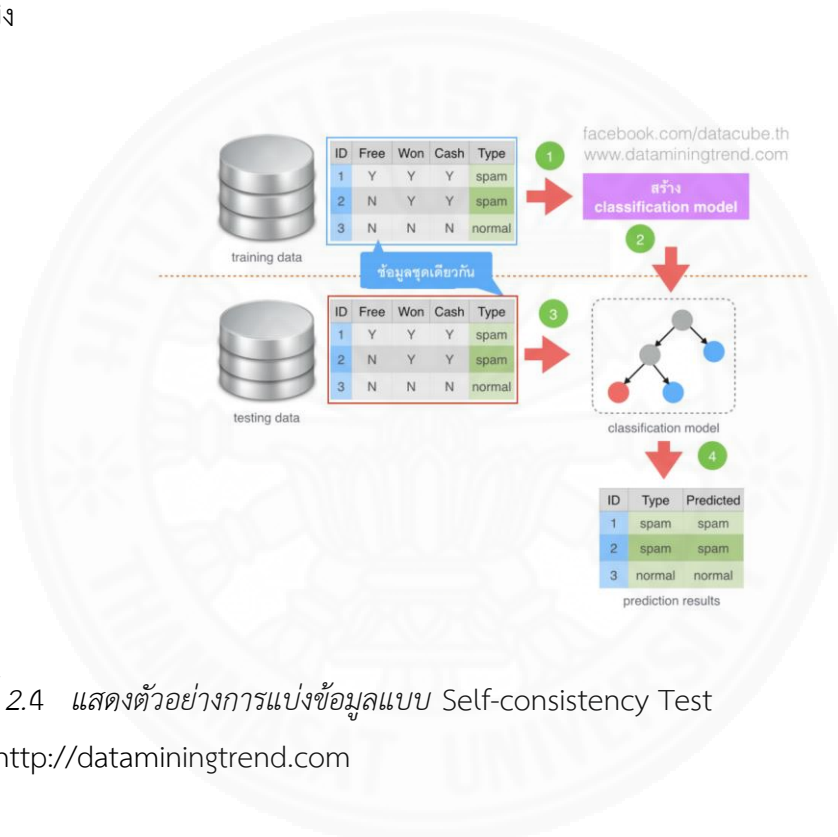
ตารางสรุปการเรียนรู้ของเครื่อง (Machine learning) (ต่อ)

		<p>Random forest คือ การสร้างโมเดลจากโมเดลย่อยๆ หลาย decision tree โดยอาจตั้งแต่ 10 – 1000 โมเดล โดยที่แต่ละโมเดลย่อยนั้นจะได้รับข้อมูลที่ใช้ในการสอน (training data) ไม่เหมือนกัน หลังจากนั้นแต่ละโมเดลจะทำการพยากรณ์ความแม่นยำของตัวเองออกมาด้วยการ vote output เพื่อหาคำตอบ</p>
	Regression	<p>การวิเคราะห์การถดถอย(regression) และการวิเคราะห์สหสัมพันธ์ (Correlation Analysis) เป็นการศึกษาเกี่ยวกับความสัมพันธ์ของตัวแปร โดยมีวัตถุประสงค์หลักของการวิเคราะห์การถดถอยคือ ความต้องการประมาณค่าของตัวแปร โดยเรียกว่า ตัวแปรตาม (Dependent Variable) ซึ่งนิยมเขียนแทนด้วย Y ซึ่งตัวแปรตามจะต้องอาศัยค่าจากตัวแปรอื่นโดยเรียกว่า ตัวแปรอิสระ (Independent Variable) นิยมเขียนแทนด้วย X หรืออาจสรุปได้ว่าตัวแปรตามขึ้นอยู่กับค่าของตัวแปรอิสระ หากความสัมพันธ์ของ Y และ X เป็นเชิงเส้นตรง เราเรียกว่า การถดถอยเชิงเส้นอย่างง่าย (Simple Linear Regression)</p>
Unsupervised	Clustering	<p>เป็นหนึ่งในหลักการการเรียนรู้แบบไม่มีผู้สอน โดยทั่วไปแล้วก็จะมีขั้นตอนวิธีการ เช่น K-Means หรือ Hierarchical Clustering ซึ่งทั้งสองหลักการนี้จะเหมาะสำหรับ cluster ที่รวมกันเป็นกลุ่มก้อน และแยกออกจากกันอย่างชัดเจน</p>
	Association Rule	<p>เป็นหลักการที่ค้นหาความสัมพันธ์ของข้อมูลจากข้อมูลที่มีอยู่เพื่อนำไปหารูปแบบที่เกิดขึ้นบ่อยๆ(frequent pattern) และใช้ในการวิเคราะห์ความสัมพันธ์หรือทำนายปรากฏการณ์ต่างๆ โดยในหลักการนี้จะมีค่า support และค่า confidence โดยที่ค่า support คือ ตัววัดประสิทธิภาพของแต่ตัวแปร และค่า confidence คือค่าวัดประสิทธิภาพของกฎความสัมพันธ์</p>

2.2.4 การแบ่งข้อมูลเพื่อนำมาทดสอบประสิทธิภาพของแบบจำลอง

2.2.4.1 Self Consistency Test

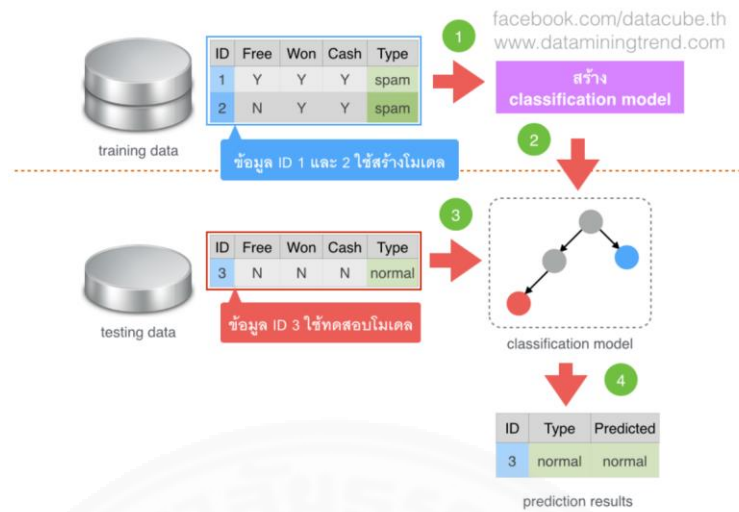
เป็นวิธีของการใช้ข้อมูลในการสร้างแบบจำลองกับข้อมูลในการทดสอบเป็นชุดเดียวกัน โดยที่ประสิทธิภาพจากการที่วัดออกมาได้จะมีค่าค่อนข้างสูง อันเนื่องมาจากมีการใช้ข้อมูลชุดเดียวกันในการสร้างแบบจำลองและใช้ในการวัดประสิทธิภาพ ดังนั้นวิธีนี้จึงเหมาะจะใช้ในการทดสอบประสิทธิภาพเพื่อดูแนวโน้มของโมเดลที่สร้างขึ้น ถ้าได้ผลการวัดที่น้อย แสดงว่าแบบจำลองที่สร้างขึ้นมาไม่มีความเหมาะสมกับข้อมูล จึงยังไม่ควรนำไปทดสอบกับชุดข้อมูลอื่นเป็นอย่างยิ่ง



ภาพที่ 2.4 แสดงตัวอย่างการแบ่งข้อมูลแบบ Self-consistency Test
ที่มา: <http://dataminingtrend.com>

2.2.4.2 Split Test

เป็นการแบ่งข้อมูลด้วยการสุ่มออกเป็น 2 ส่วน คือ 70% หรือ 80% ไว้สำหรับในการสร้างแบบจำลองการพยากรณ์ และ 20% หรือ 30% ไว้สำหรับในทดสอบความแม่นยำของแบบจำลองที่สร้างขึ้นมา โดยการทดสอบแบบ Split Test นี้ทำการสุ่มข้อมูลเพียงครั้งเดียว ซึ่งถ้าหากครั้งที่ทำการสุ่มข้อมูลที่ใช้ในการทดสอบที่มีลักษณะคล้ายกับข้อมูลที่ใช้สร้างแบบจำลองก็จะทำให้ผลการวัดประสิทธิภาพได้ออกมาดี ในทางกลับกันถ้าหากการสุ่มข้อมูลที่ใช้ในการทดสอบที่มีลักษณะแตกต่างกับข้อมูลที่ใช้สร้างแบบจำลองมากก็จะทำให้ผลการวัดประสิทธิภาพได้ออกมาไม่ดี ดังนั้นหากต้องการใช้วิธี Split Test นี้จึงควรทำการสุ่มหลายๆ ครั้ง

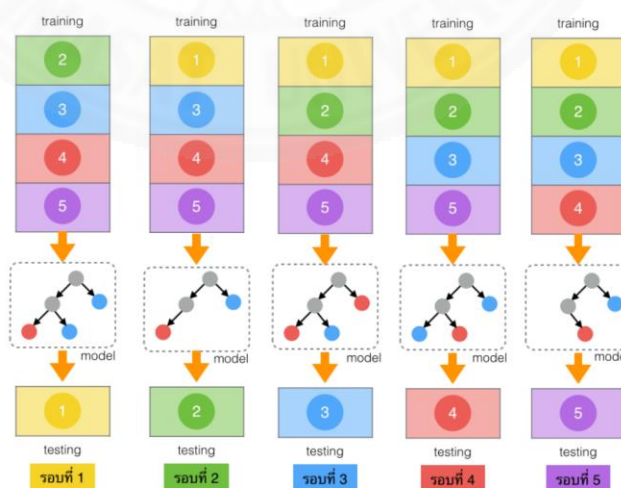


ภาพที่ 2.5 แสดงตัวอย่างการแบ่งข้อมูลแบบ Split Test

ที่มา: <http://dataminingtrend.com>

2.2.4.3 Cross-validation Test

วิธี Cross-validation นี้จะทำการแบ่งข้อมูลออกเป็นหลายส่วน โดยสามารถกำหนดได้ว่าต้องการแบ่งข้อมูลออกเป็นกี่ส่วนเท่าๆกัน หลังจากนั้นข้อมูลหนึ่งส่วนจะถูกใช้เป็นตัวทดสอบประสิทธิภาพของแบบจำลอง ซึ่งจะมีการทำวนไปเช่นนี้จนครบจำนวนที่แบ่งไว้ โดยวิธีนี้เป็นวิธีที่เหมาะสมและนิยมใช้ในการทำวิจัยเนื่องจากผลที่ได้มีความน่าเชื่อถือมากที่สุด ใน 3 วิธีของการแบ่งข้อมูลเพื่อนำมาทดสอบประสิทธิภาพของแบบจำลอง



ภาพที่ 2.6 แสดงตัวอย่างการแบ่งข้อมูลแบบ Cross-validation Test

ที่มา: <http://dataminingtrend.com>

2.2.5 การคัดเลือกคุณลักษณะ

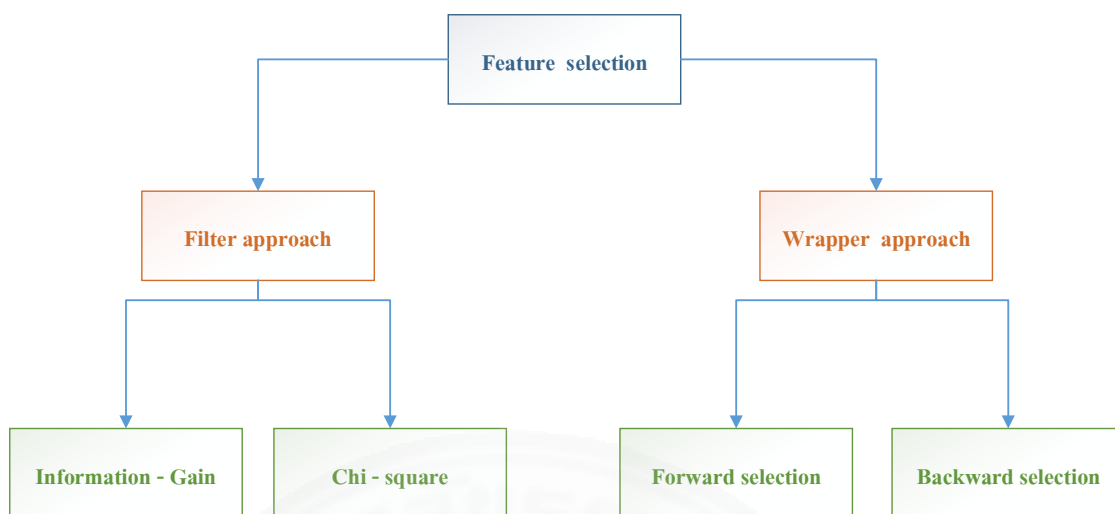
การคัดเลือกคุณลักษณะ (Feature selection) คือ การคัดเลือกคุณสมบัติที่ช่วยในการลดจำนวนตัวแปรที่จะใช้ในการพยากรณ์ข้อมูลในแบบประเภทข้อมูลการจำแนกข้อมูล (classification) เพื่อให้จำนวนตัวแปรนั้นมีจำนวนให้น้อยที่สุดและไม่ส่งผลต่อประสิทธิภาพของความแม่นยำของแบบจำลองในการพยากรณ์ โดยขั้นตอนนี้เรียกว่าการคัดเลือกคุณลักษณะ หรือ feature selection สามารถแบ่งได้เป็น 2 กลุ่มใหญ่ดังนี้

2.2.5.1 Filter approach คือ การคัดเลือกคุณลักษณะด้วยการใช้วิธีคำนวณหาค่าน้ำหนักซึ่งอาจจะเป็นค่าความสัมพันธ์ระหว่างแต่ละฟีเจอร์กับผลของการจำแนกประเภทของข้อมูล โดยที่ค่าน้ำหนักที่คำนวณได้แล้วเลือกคุณลักษณะที่มีค่าน้ำหนักมากกว่าที่ต้องการมาใช้งานต่อไป ซึ่งเทคนิคในการคำนวณค่าน้ำหนักของคุณลักษณะต่างๆ มีหลายวิธี ได้แก่ Information Gain , Chi-Square หรือ Correlation

2.2.5.2 Wrapper approach คือ เป็นวิธีการคัดเลือกคุณลักษณะด้วยการสร้างโมเดล (classification model) ขึ้นมาจากกลุ่มของคุณลักษณะที่ได้มีการกำหนดไว้และทำการวัดประสิทธิภาพการทำงานของโมเดล พร้อมทั้งทำการเลือกกลุ่มของคุณลักษณะที่ทำให้โมเดลมีประสิทธิภาพมากที่สุดมาใช้งาน ได้แก่ โมเดลที่ให้ค่าความถูกต้อง (accuracy) มากที่สุด การคัดเลือกคุณลักษณะด้วยวิธีนี้สามารถแบ่งได้เป็น 2 แบบใหญ่ๆ คือ

1. Forward Selection คือ การสร้างโมเดลโดยการเพิ่มคุณลักษณะของข้อมูลเข้าไปทีละ 1 คุณลักษณะ โดยที่คุณลักษณะที่มีการใส่เข้าไบนั้นจะต้องทำให้ประสิทธิภาพของโมเดลที่ดีขึ้น และจะทำการเพิ่มคุณลักษณะอื่นๆ เข้ามาใส่เพิ่มต่อไปเรื่อยๆจนประสิทธิภาพของโมเดลที่ได้นั้นไม่สูงกว่าหรือเท่าเดิมจากประสิทธิภาพสูงสุดที่ทำได้ จึงจะหยุดการเพิ่มคุณลักษณะลงในโมเดล

2. Backward Elimination คือ การสร้างโมเดลโดยเริ่มจากคุณลักษณะของข้อมูลทั้งหมดที่มีนำมาหาประสิทธิภาพของโมเดล และหลังจากนั้นจะทำการตัดคุณลักษณะของข้อมูลออกไปทีละคุณลักษณะ โดยที่ถ้าประสิทธิภาพของโมเดลดีขึ้น จะทำการตัดคุณลักษณะอื่นๆต่อไปเรื่อยๆจนกว่าประสิทธิภาพของโมเดลที่ได้ ไม่สูงกว่าเท่าเดิมจากประสิทธิภาพสูงสุดที่ทำได้ จึงหยุดการตัดคุณลักษณะออกจากโมเดล



ภาพที่ 2.7 แสดงรูปแบบวิธีการคัดเลือกคุณลักษณะ

ที่มา: <http://dataminingtrend.com>

2.2.6 การวัดประสิทธิภาพของโมเดลการพยากรณ์

ตัววัดประสิทธิภาพที่นิยมใช้ในการวัดประสิทธิภาพของตัวแบบพยากรณ์แบบแบ่งประเภท (classification) มีอยู่ 5 ค่า ดังนี้

1. Recall คือ การวัดความถูกต้องของตัวแบบพยากรณ์ แล้วคำนวณจากการที่ตัวแบบพยากรณ์สามารถตอบคำถามถูกในส่วนที่สนใจได้มากน้อยเพียงใดจากค่าที่ถูกต้องจริงทั้งหมด สามารถคำนวณได้จาก $Recall = True\ Positive / (True\ Positive + False\ Negative)$

2. Precision คือ การวัดความแม่นยำของตัวแบบพยากรณ์ แล้วคำนวณจากการที่ตัวแบบพยากรณ์สามารถตอบคำถามได้ถูกในส่วนที่ไม่ได้สนใจมากน้อยเพียงใดจากค่าที่ถูกต้องจริงทั้งหมด สามารถคำนวณได้จาก $Precision = True\ Positive / (True\ Positive + False\ Positive)$

3. F-measure คือ การวัดค่า Precision และ Recall พร้อมกันของตัวแบบพยากรณ์ โดยที่ทำการพิจารณาแยกทีละผลลัพธ์ที่สนใจและไม่สนใจ ซึ่งใช้เป็นตัวบอกระดับความเหมาะสมของการนำข้อมูลไปใช้ว่ามีความเหมาะสมเพียงใด พร้อมทั้งบอกระดับประสิทธิภาพของตัวแบบพยากรณ์ สามารถคำนวณได้จาก $F\text{-measure} = (2 \times Precision \times Recall) / (Precision + Recall)$

4. Accuracy คือ การวัดความถูกต้องของตัวแบบพยากรณ์ โดยพิจารณารวมทุกผลลัพธ์ที่เกิดขึ้นทั้งหมด เพื่อดูว่ามีความแม่นยำในการทำนายผลถูกต้องมากหรือน้อยเพียงใด สามารถคำนวณได้จาก $Accuracy = (True\ Positive + Negative\ Positive) / \text{จำนวนชุดข้อมูลทั้งหมด}$

5. ROC Graph & Area คือ กราฟที่แสดงความสัมพันธ์ระหว่างข้อมูลที่ทำนายถูก (แกน Y) และทำนายผิด (แกน X) เพื่อแสดงให้เห็นพื้นที่ความถูกต้องที่ตัวแบบพยากรณ์ได้จากชุดข้อมูลที่ใช้ในการทดสอบ

ซึ่งการคำนวณค่าต่างๆจากข้างต้นนั้นมีการนำค่าต่างๆมาจากตาราง confusion matrix

ตารางที่ 2.3

ตารางแสดงตัวอย่าง Confusion Matrix

Predict / Actual	Yes	No
Yes	TP	FP
No	FN	TN

จากตารางที่ 2.1 ค่าที่แสดงในตารางสามารถอธิบายได้ดังนี้

True Positive (TP) คือ สิ่งที่โปรแกรมทำนายว่าจริง และคนบอกว่ามันจริง

True Negative (TN) คือ สิ่งที่โปรแกรมทำนายว่าไม่จริง และคนบอกว่ามันไม่

จริง

False Positive (FP) คือ สิ่งที่โปรแกรมทำนายว่าจริง แต่คนบอกว่าไม่จริง

False Negative (FN) คือ สิ่งที่โปรแกรมทำนายว่าไม่จริง แต่คนบอกว่าจริง

2.3 เอกสารและงานวิจัยที่เกี่ยวข้อง

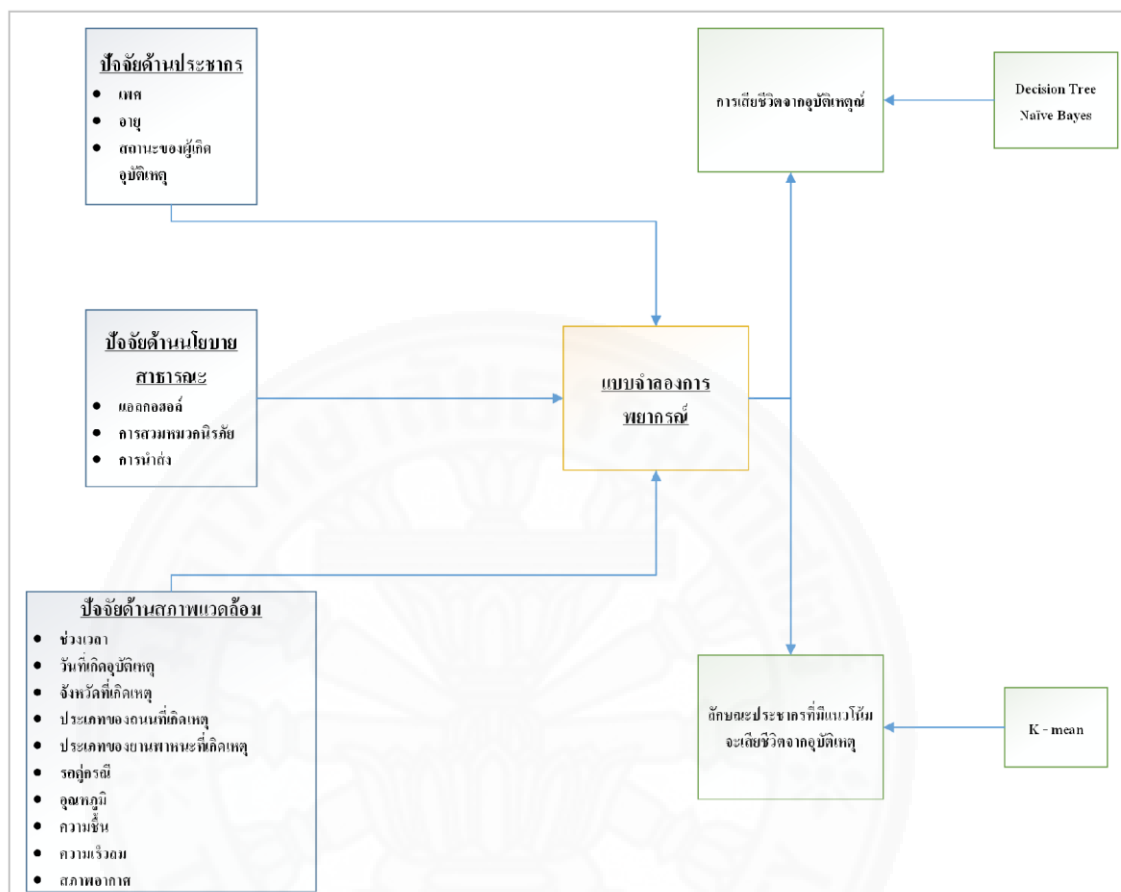
สรุปเอกสารและงานวิจัยที่เกี่ยวข้องกับการเกิดอุบัติเหตุจราจรทางบกที่เกี่ยวข้องได้ดังต่อไปนี้

ตารางที่ 2.4

ตารางแสดงผลสรุปปัจจัยที่มีผลต่อการเกิดอุบัติเหตุ

ผู้วิจัย	อายุ	เพศ	สถานะของผู้ประสบอุบัติเหตุ	ประเภทของยานพาหนะที่เกิดอุบัติเหตุ	ประสบการณ์ของผู้ขับขี่	การดื่มแอลกอฮอล์	บริเวณที่เกิดอุบัติเหตุ	ความเร็ว	การสวมหมวกหรือคาดเข็มขัดนิรภัย	สภาพแวดล้อมต่างๆ เช่น แสง บริเวณที่เกิดเหตุ
Chaozhong (2009)	X	X		X	X			X		X
Saraee and other (2005)	X	X		X						
Lluis Sanmiquel (2015)	X				X					
Tibebe Beshah	X	X					X			
Sami and other (2009)						X	X	X		
S.Krishnaveni (2011)	X	X	X	X						X
WHO						X		X	X	
ดร.ณิ ศรีมณีรัตน์ (2558)	X	X		X			X	X		

2.4 กระบวนการในการวิจัย



ภาพที่ 2.8 แผนภาพแสดงกรอบงานวิจัย

จากภาพที่ 2.13 เป็นการแสดงกรอบงานวิจัยโดยที่งานวิจัยนี้ศึกษาปัจจัย 3 ปัจจัย ได้แก่ ปัจจัยด้านประชากร ปัจจัยด้านนโยบายสาธารณะ และปัจจัยด้านสภาพแวดล้อม โดยที่ผู้วิจัยนั้นจะใช้วิธีการ Decision Tree และ Naïve Bayes เพื่อสร้างแบบจำลองการพยากรณ์การเกิดอุบัติเหตุในช่วงเทศกาลปีใหม่เพื่อศึกษาถึงตัวแปรในแต่ละปัจจัย ว่าในการเกิดอุบัติเหตุทางถนนในช่วงเทศกาลปีใหม่เกี่ยวข้องกับตัวแปรใดบ้างพร้อมทั้งเพื่อเป็นการเปรียบเทียบประสิทธิภาพของทั้ง 2 วิธี และใช้วิธี Cluster เพื่อแบ่งกลุ่มประชากรออกเป็น 2 กลุ่ม คือ ชาย และ หญิงเพื่อศึกษาว่าในแต่ละกลุ่มนั้นประชากรแต่ละกลุ่มจะมีลักษณะเป็นอย่างไร

บทที่ 3

วิธีการวิจัย

งานวิจัยนี้เป็นการวิจัยเพื่อหาแบบจำลองการพยากรณ์การเกิดอุบัติเหตุทางถนนในช่วงเทศกาลปีใหม่ ซึ่งผู้วิจัยได้เลือกใช้ข้อมูลทุติยภูมิมาใช้ในการทำวิจัย เพื่อหาปัจจัยที่สัมพันธ์กับการเกิดอุบัติเหตุในช่วงเทศกาลปีใหม่ โดยเลือกใช้ขั้นตอนการทำวิทยาศาสตร์ข้อมูล (Data Science Process) โดยมีขั้นตอนในการวิจัยดังนี้

- 3.1 ขอบเขตของการวิจัยและกลุ่มตัวอย่างที่ใช้ในการวิจัย
- 3.2 เครื่องมือที่ใช้ในการวิจัย
- 3.3 ขั้นตอนการดำเนินงาน
- 3.4 ระยะเวลาในการวิจัย

3.1 ขอบเขตของการวิจัยและกลุ่มตัวอย่างที่ใช้ในการวิจัย

ข้อมูลการเกิดอุบัติเหตุทางถนนในช่วงเทศกาลปีใหม่ได้ทำการจัดเก็บโดยสถาบันการแพทย์ฉุกเฉินแห่งชาติ และเผยแพร่โดยสำนักงานพัฒนารัฐบาลดิจิทัล (EGA) ซึ่งเป็นข้อมูลแบบทุติยภูมิ โดยเป็นข้อมูลที่เกิดอุบัติเหตุทางถนนในช่วงเทศกาลปีใหม่ตั้งแต่ปี 2551 - 2558 ซึ่งเป็นข้อมูลการเกิดอุบัติเหตุทั่วประเทศไทย ข้อมูลที่มีการเก็บรวบรวมนั้นประกอบไปด้วยข้อมูลด้านต่างๆ เช่น เพศ อายุ สถานที่เกิดเหตุ วันเวลาที่เกิดเหตุ ประเภทของยานพาหนะ และพฤติกรรมของผู้ขับขี่ เป็นต้น ในที่นี้การนำข้อมูลในการสร้างตัวแบบจำลองพยากรณ์การเกิดอุบัติเหตุทางถนนในช่วงเทศกาลปีใหม่เพื่อที่จะช่วยให้มีจำนวนการเกิดอุบัติเหตุและการสูญเสียจากอุบัติเหตุมีจำนวนที่ลดลง

3.2 เครื่องมือที่ใช้ในการวิจัย

3.2.1 เครื่องมือที่ใช้ในการทำการเรียนรู้ของเครื่อง (Machine Learning)

ผู้วิจัยได้ใช้โปรแกรมที่ใช้ในการประมวลผล (Software Tools) คือ Weka โดยโปรแกรม Weka ย่อมาจาก Waikato Environment for Knowledge Analysis พัฒนาขึ้นโดยใช้ภาษา java เป็นโปรแกรมประเภทโอเพ่นซอส คือเปิดให้ทุกคนสามารถใช้ได้อย่างไม่มีค่าใช้จ่าย โดยที่โปรแกรม Weka เป็นเครื่องมือที่ใช้ในการสร้างแบบจำลองการพยากรณ์เพื่อหาคำตอบตามที่งานวิจัยนี้ต้องการ ซึ่งสามารถติดตั้งได้ทั้งบนระบบปฏิบัติการ Windows, Linux และ Mac OS

3.2.2 การทดสอบเนื้อหาความในการวิจัย (Content Validity Testing)

ผู้วิจัยได้ขอคำแนะนำจากผู้ทรงคุณวุฒิได้แก่ อาจารย์ที่ปรึกษาการค้นคว้าอิสระ ผู้เชี่ยวชาญด้านการวิจัยและคณะกรรมการพิจารณา เพื่อที่จะเป็นการยืนยันและตรวจสอบว่างานวิจัยในครั้งนี้มีผลการวิจัยที่ตรงกับวัตถุประสงค์ของงานวิจัย อีกทั้งยังเป็นการให้ความมั่นใจว่าผลของการวิจัยนี้มีความถูกต้องและแม่นยำ

3.3 ขั้นตอนการดำเนินงาน

การวิจัยครั้งนี้ผู้ทำการวิจัยได้ทำการแบ่งขั้นตอนการดำเนินงานไว้ทั้งหมด 5 ขั้นตอน ดังนี้

1. การศึกษาและการวางแผนการเก็บข้อมูล
2. การเก็บรวบรวมข้อมูล
3. การสำรวจข้อมูลและจัดการข้อมูล
4. การวิเคราะห์และออกแบบตัวแบบการทำนาย
5. การอธิบายผลและการนำไปใช้

3.3.1 การศึกษาและการวางแผนการเก็บข้อมูล

การศึกษางานวิจัยครั้งนี้นั้น จะทำการเก็บรวบรวมข้อมูลเพื่อต้องการที่จะทราบ ว่าปัจจัยของการเกิดอุบัติเหตุและตัวแปรใดบ้างที่มีความสัมพันธ์เกี่ยวกับการเกิดอุบัติเหตุในช่วงเทศกาลปีใหม่ และการเสียชีวิตในช่วงเทศกาลปีใหม่นั้นกลุ่มประชากรมีลักษณะเป็นอย่างไรบ้าง และในส่วนของการวางแผนการเก็บข้อมูลนั้น ข้อมูลของกลุ่มเป้าหมายจะประกอบไปด้วยรายละเอียดดังต่อไปนี้

ข้อมูลด้านประชากร

ตารางที่ 3.1

แสดงข้อมูลด้านประชากร

ที่	ชื่อข้อมูล (EN)	ชื่อข้อมูล (TH)	ประเภท
1	Gender	เพศ	Text
2	Age	อายุ	Number
3	Status of Accident	สถานะของผู้ประสบอุบัติเหตุ	Text

ข้อมูลด้านสิ่งแวดล้อม

ตารางที่ 3.2

แสดงข้อมูลด้านสิ่งแวดล้อม

ที่	ชื่อข้อมูล (EN)	ชื่อข้อมูล (TH)	ประเภท
1	Time	ช่วงเวลา	Time
2	Date	วันที่เกิดเหตุ	Date
3	Province	จังหวัดที่เกิดเหตุ	Text
4	Type of road	ประเภทของถนนที่เกิดเหตุ	Text
5	Vehicle of vehicle incident	ประเภทยานพาหนะที่เกิด ประสพเหตุ	Text
6	Party vehicle	รถคู่กรณี	Text
7	Temperature	อุณหภูมิ	Number
8	Humidity	ความชื้น	Percentage
9	Wind speed	ความเร็วลม	Number
10	Condition	สภาพอากาศ	Text

ข้อมูลด้านกฎหมาย

ตารางที่ 3.3

แสดงข้อมูลด้านนโยบายสาธารณะ

ที่	ชื่อข้อมูล (EN)	ชื่อข้อมูล (TH)	ประเภท
1	Drinking	ดื่มแอลกอฮอล์	Text
2	Defensive measure	มาตรการป้องกัน	Text
3	Send by	การนำเสนอ	Text

ข้อมูลด้านสถานะของเป้าหมาย

ตารางที่ 3.4

แสดงข้อมูลข้อมูลผลการรักษา

ที่	ชื่อข้อมูล (EN)	ชื่อข้อมูล (TH)	ประเภท
1	Status	เสียชีวิตหรือไม่	Text

3.3.2 การเก็บรวบรวมข้อมูล

ข้อมูลการเกิดอุบัติเหตุทางถนนในช่วงเทศกาลปีใหม่ได้ทำการจัดเก็บโดยสถาบันการแพทย์ฉุกเฉินแห่งชาติ และเผยแพร่โดยสำนักงานพัฒนารัฐบาลดิจิทัล (EGA) ซึ่งเป็นข้อมูลแบบทุดิยภูมิ โดยเป็นข้อมูลที่เกิดอุบัติเหตุทางถนนในช่วงเทศกาลปีใหม่ตั้งแต่ปี 2551 - 2558 ซึ่งเป็นข้อมูลการเกิดอุบัติเหตุทั่วประเทศไทย และข้อมูลด้านสภาพอากาศจาก <https://www.wunderground.com>

3.3.3 การสำรวจข้อมูลและจัดการข้อมูล

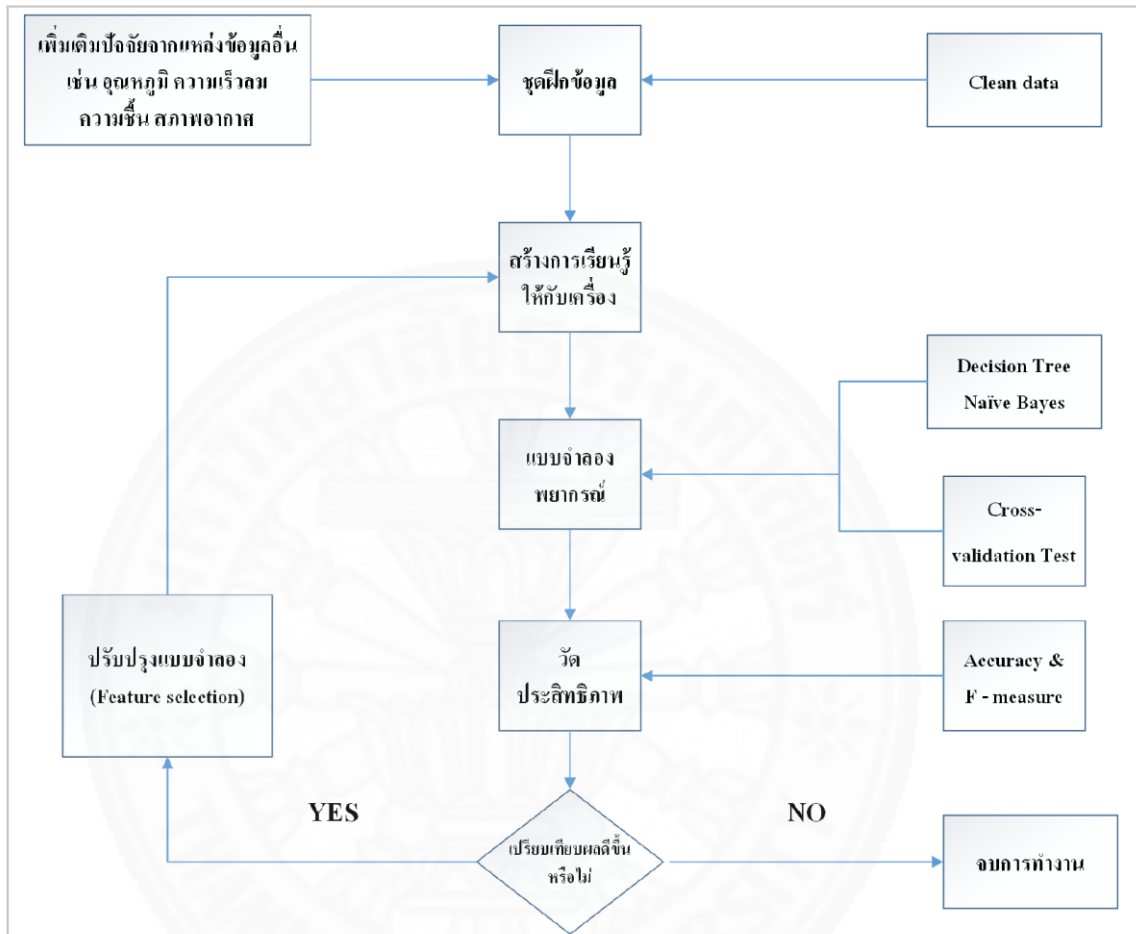
งานวิจัยนี้ใช้วิธีการตัดข้อมูลที่มีลักษณะข้อมูลที่ระบุข้อมูลไม่ถูกต้องหรือไม่ครบถ้วน ออกจากชุดข้อมูลที่จะนำมาวิเคราะห์

3.3.4 การวิเคราะห์และออกแบบตัวแบบการทำนาย

เนื่องจากข้อมูลที่มีนั้นเป็นข้อมูลที่บันทึกเป็นภาษาไทยทั้งหมด ผู้ทำวิจัยจึงต้องแปลงข้อมูลทั้งหมดให้อยู่ในรูปภาษาอังกฤษ เนื่องจากโปรแกรม Weka ไม่รองรับภาษาไทย จึงต้องแปลงข้อมูลให้เหมาะสมเพื่อนำไปใช้ในการวิเคราะห์ โดย ขั้นตอนการออกแบบตัวแบบการทำนายและการวิเคราะห์เป็นไปตามภาพที่ 3.1

จากภาพ 3.1 เป็นการแสดงขั้นตอนการออกแบบตัวแบบการจำลองพยากรณ์ โดยเริ่มจากการนำข้อมูลที่มีอยู่มาทำเป็นชุดฝึกหัดเพื่อให้เครื่องการเรียนรู้ทำการเรียนรู้โดยใช้ขั้นตอนวิธี (Algorithm) พร้อมทั้งจำแนกประเภทข้อมูล (Classification) หลังจากนั้นให้เครื่องทำการเรียนรู้สร้างตัวแบบจำลองพยากรณ์ แล้วนำตัวแบบที่ได้ไปทำการทดสอบด้วยชุดข้อมูลทดสอบ เพื่อวัดประสิทธิภาพของแบบจำลองพยากรณ์ แล้วนำผลไปวัดผลประสิทธิภาพเพื่อใช้เปรียบเทียบกับตัวแบบถัดไป กรณีที่ทำการเปรียบเทียบผลลัพธ์ของการวัดประสิทธิภาพแล้วให้ผลลัพธ์ที่ดีกว่าตัวแบบก่อนหน้า จะทำการนำคุณลักษณะของชุดข้อมูลออกไปทีละตัวแล้วทำการทดสอบซ้ำไปจนกว่าจะไม่ได้ผลลัพธ์ที่ดีกว่าผลลัพธ์เดิมแล้วจึงทำการหยุดทำซ้ำ ในส่วนของการทำการแบ่งกลุ่มข้อมูลจะใช้วิธี K -

mean ในการจัดกลุ่มของข้อมูลตามเพศ เพื่อศึกษาว่าแต่ละเพศมีคุณลักษณะการเกิดอุบัติเหตุต่างกัน หรือเหมือนกันตรงไหนบ้าง



ภาพที่ 3.1 แสดงขั้นตอนการออกแบบจำลองพยากรณ์

3.3.5 การอธิบายผลและการนำไปใช้

การอธิบายคุณลักษณะของการจัดกลุ่มชุดข้อมูลที่ทำให้เกิดอุบัติเหตุ พร้อมทั้งมีการสร้างแบบจำลองการพยากรณ์จากตัวแบบที่ได้ทำการสร้างไว้ในข้อ 3.3.4

3.4 ระยะเวลาในการวิจัย

งานวิจัยครั้งนี้ผู้วิจัยได้ทำการเก็บข้อมูลทุติยภูมิและค้นคว้าข้อมูลที่เกี่ยวข้องจากหนังสือ บทความ วิทยานิพนธ์ สารนิพนธ์ การค้นคว้าอิสระ บทความวารสาร และสื่ออิเล็กทรอนิกส์ที่เกี่ยวข้อง

รวมถึงทำการสำรวจข้อมูลที่นำมาใช้ในงานวิจัยจนแล้วเสร็จมาเป็นการวิเคราะห์ผลการวิจัยครั้งนี้ตั้งแต่เดือนตุลาคม 2561 – พฤษภาคม 2562 รวมระยะเวลา 8 เดือน

ตารางที่ 3.5

ระยะเวลาในการทำงานวิจัย

ขั้นตอนการดำเนินงาน	ระยะเวลา							
	ต.ค. 61	พ.ย. . 61	ธ.ค. 61	ม.ค. 62	ก.พ. . 62	มี.ค. 62	เม.ษ. 62	พ.ค. . 62
1. เตรียมหัวข้อการค้นคว้าอิสระ/งานวิจัย								
1.1 คิดหัวข้อ ศึกษาแนวความคิดและข้อมูลที่เกี่ยวข้อง								
1.2 เขียนเค้าโครงงานวิจัย								
1.3 เสนอหัวข้อและเค้าโครงให้อาจารย์ที่ปรึกษาพิจารณาอนุมัติ								
1.4 เสนอหัวข้องานวิจัย								
1.5 วิทยาลัยประกาศผลหัวข้อพร้อมรายชื่ออาจารย์ที่ปรึกษา								
2. แนวคิดและรูปแบบงานวิจัย								
2.1 พัฒนาแนวคิดและรูปแบบงานวิจัย								
2.2 จัดทำรายละเอียดขั้นตอนและระเบียบวิธีวิจัย								
3. ดำเนินงานวิจัย								
3.1 สอบ Defense งานวิจัย (ครั้งที่ 1)								
4. ปรับปรุง/แก้ไขผลการวิจัย								
4.1 ปรับปรุงแก้ไขงานวิจัยตามคำแนะนำของคณะกรรมการสอบงานวิจัย (ครั้งที่ 1)								

ตารางที่ 3.5

ระยะเวลาในการทำงานวิจัย (ต่อ)

ขั้นตอนการดำเนินงาน	ระยะเวลา							
	ต.ค. 61	พ.ย. . 61	ธ.ค. 61	ม.ค. 62	ก.พ. . 62	มี.ค. 62	เม.ษ. 62	พ.ค. . 62
4.2 รวบรวมข้อมูลทุติยภูมิ พร้อมทำงานวิจัย								
4.3 สรุปผลการวิจัย								
4.4 วิเคราะห์ข้อมูล-ทดสอบผลการวิจัย								
5. เขียนรายงานวิจัย								
5.1 เขียนรายงานวิจัย								
6. เสนอรายงานวิจัย								
6.1 สอบ Defend งานวิจัยครั้งที่ 2								
7. ปรับปรุงแก้ไข								
7.1 ปรับปรุงแก้ไขงานวิจัยตามคำแนะนำของคณะกรรมการ สอบงานวิจัย (ถ้ามี)								
8. ส่งผลงานวิจัย								
8.1 ส่งผลงานวิจัยฉบับสมบูรณ์								

บทที่ 4

ผลการวิจัยและอภิปรายผล

ในการวิจัยการพยากรณ์การเสียชีวิตจากอุบัติเหตุทางถนนในช่วงเทศกาลปีใหม่ ผู้วิจัยได้ทำการแบ่งขั้นตอนการดำเนินงานออกเป็น 5 ขั้นตอน ซึ่งได้มีการจัดทำพร้อมทั้งสรุปผลลัพธ์จากการดำเนินงานออกเป็น 2 หัวข้อดังนี้ คือ 1.ผลของการจัดทำตัวแบบพยากรณ์ และ 2.ผลของการจัดกลุ่มลักษณะประชากรที่มีโอกาสเสียชีวิตทางถนนในช่วงเทศกาลปีใหม่

4.1 ผลการวิจัย

4.1.1 ผลของการดำเนินงานในแต่ละขั้นตอน

4.1.1.1 การศึกษาและการวางแผนการเก็บข้อมูล

ขั้นตอนนี้เป็นขั้นตอนการกำหนดวัตถุประสงค์ของงานวิจัยที่เราต้องการศึกษาหรือคำถามต้องการหาคำตอบ เพื่อคว้าข้อมูลที่ได้จาก <https://data.go.th/default.aspx> และ <https://www.wunderground.com> นำมาใช้ในการวิจัยนั้นมีความเพียงพอต่อการทำงานวิจัยหรือไม่ โดยงานวิจัยนี้มีวัตถุประสงค์หรือมุ่งเน้นเพื่อลดการเสียชีวิตของประชากรที่จะเสียชีวิตบนท้องถนนในเทศกาลปีใหม่ไทย ดังนั้นการตั้งคำถามของงานวิจัยนี้จึงจะมุ่งเน้นไปที่การเสียชีวิตบนท้องถนนในเทศกาลปีใหม่ระหว่างวันที่ 28 ธันวาคม ถึง 3 มกราคม ซึ่งผู้ทำการวิจัยได้ตั้งคำถามเพื่อให้สอดคล้องกับวัตถุประสงค์ของงานวิจัยไว้ดังนี้

1. การหาตัวแบบพยากรณ์การเสียชีวิตทางถนนในเทศกาลปีใหม่
2. การหาลักษณะของประชากรที่มีโอกาสที่จะเสียชีวิตทางถนนในเทศกาลปีใหม่ จากคำถามข้างต้นทางผู้วิจัยจำเป็นต้องมีข้อมูลเพื่อทำการหาคำตอบของคำถาม โดยแบ่งคุณลักษณะของข้อมูลออกเป็น 4 ส่วนคือ

1. ข้อมูลปัจจัยด้านประชากร ได้แก่ เพศ อายุ สถานะของผู้เกิดอุบัติเหตุ
2. ข้อมูลปัจจัยด้านนโยบายสาธารณะ ได้แก่ การดื่มแอลกอฮอล์ การสวมหมวกนิรภัยหรือขาดเข็มขัดนิรภัย การนำส่ง

3. ข้อมูลปัจจัยด้านสภาพแวดล้อม ได้แก่ ช่วงเวลา วันที่เกิดอุบัติเหตุ จังหวัดที่เกิดเหตุ ประเภทของถนนที่เกิดเหตุ ประเภทของยานพาหนะที่เกิดเหตุ รัศมีถนน อุณหภูมิ ความชื้น ความเร็วลม และสภาพอากาศ

4. ข้อมูลคลาสของผลลัพธ์ที่สนใจ ได้แก่ สถานะของประชากรที่เกิดอุบัติเหตุ

4.1.1.2 การเก็บรวบรวมข้อมูล

ขั้นตอนนี้เป็นการรวบรวมข้อมูลเพื่อนำมาใช้ในการตอบคำถามโดยใช้คุณลักษณะข้อมูลที่จำเป็นจากข้อ 4.1 ซึ่งเป็นข้อมูลทุติยภูมิจาก <https://data.go.th/default.aspx> และเป็นระยะเวลาย้อนหลัง 8 ปี (พ.ศ.2551-2558) โดยมีจำนวนข้อมูลทั้งสิ้น 241,950 รายการ

4.1.1.3 การสำรวจข้อมูลและจัดการข้อมูล

ขั้นตอนนี้จะเป็นรวบรวมข้อมูลและเตรียมข้อมูลให้เหมาะสมกับการนำไปใช้ในการสร้างตัวแบบการพยากรณ์ และการจัดกลุ่มข้อมูลประชากรโดยมีขั้นตอนในการจัดเตรียมข้อมูลก่อนนำไปทำการวิเคราะห์และสร้างตัวแบบดังต่อไปนี้

1. แปลงข้อมูลจากภาษาไทยเป็นภาษาอังกฤษทั้งหมด
2. คัดรายการของข้อมูลที่มีข้อมูลไม่สมบูรณ์ออกจากชุดข้อมูล
3. คัดคุณลักษณะที่ไม่เกี่ยวข้องออก เช่น รหัสของจังหวัด จำนวนวัน

ของการรักษา

4. การปรับเปลี่ยนคุณลักษณะของข้อมูลให้เหมาะสมกับการนำไปใช้เพื่อสร้างตัวแบบการพยากรณ์ หรือ การจัดกลุ่ม

5. การสร้างคุณลักษณะใหม่เพิ่มขึ้นมาจากคุณลักษณะเดิมที่มี เพื่อสร้างตัวแบบการพยากรณ์ หรือ การจัดกลุ่ม ได้แก่ ความเร็วลม ความชื้น อุณหภูมิ สภาพอากาศ

6. หลังจากที่มีการจัดการข้อมูลแล้วทำให้เหลือข้อมูลทั้งหมด 175,230 รายการ

4.1.1.4 การวิเคราะห์และออกแบบตัวแบบการทำนาย

ขั้นตอนนี้เป็นขั้นตอนการนำข้อมูลที่มีการเตรียมการไว้เรียบร้อยแล้วนำมาใช้ โดยใช้โปรแกรม Weka ในการทำการวิเคราะห์ข้อมูลเพื่อหาตัวแบบการทำนายและหาลักษณะของประชากรที่มีโอกาสเสียชีวิตทางถนนในเทศกาลปีใหม่

1. การสร้างตัวแบบการพยากรณ์การเสียชีวิตทางถนนในเทศกาลปีใหม่

ใช้ชุดข้อมูลจำนวน 175,230 รายการ มาทำการสร้างตัวแบบการพยากรณ์ โดยใช้วิธีของ Decision Tree และ Naives Bayes พร้อมทั้งทำการปรับปรุงตัวแบบการพยากรณ์เพื่อหาชุดของลักษณะข้อมูลที่เหมาะสมเพื่อให้ได้ตัวแบบการพยากรณ์ที่มีประสิทธิภาพในการทำนายสูงสุด โดยใช้ค่าของ F-measure และ Accuracy ซึ่งเป็นตัววัดค่าของประสิทธิภาพของตัวแบบพยากรณ์และค่าความแม่นยำของตัวแบบพยากรณ์ตามลำดับ พร้อมทั้งมีการปรับปรุงตัวแบบด้วยเทคนิคการคัดเลือกคุณลักษณะของข้อมูลแบบการถอดลักษณะข้อมูลไปที่ละลักษณะแล้วทำการเปรียบเทียบกันด้วยค่าประสิทธิภาพของตัวแบบและค่าความแม่นยำของตัวแบบพยากรณ์ของแต่ละชุดข้อมูลที่แตกต่างกัน ถ้าตัวแบบที่ได้หลังจากการถอดลักษณะข้อมูลไป 1 ลักษณะ มีค่าของประสิทธิภาพมากกว่าเดิม จะมีการทำการทดสอบต่อโดยถอดลักษณะข้อมูลถัดไปเรื่อย ๆ จนค่าของประสิทธิภาพและค่าความแม่นยำของตัวแบบพยากรณ์ที่ได้เท่าเดิมหรือน้อยกว่าจึงจะหยุด และใช้ตัวแบบที่ให้ค่าประสิทธิภาพสูงที่สุดในการสร้างตัวแบบการพยากรณ์

การทดสอบโดยใช้ขั้นตอนวิธี Decision Tree
 รอบที่ 1 ใช้ลักษณะข้อมูลทั้งหมด

ตารางที่ 4.1

แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Decision tree (16 ตัวแปร)

ครั้งที่	ตัวแปร															Accuracy	Precision	Recall	F - measure	
	เพศ	อายุ	สถานะของผู้เกิดอุบัติเหตุ	แอลกอฮอล์	มาตรการ	การนำส่ง	ช่วงเวลา	วันที่เกิดเหตุ	จังหวัดที่เกิดเหตุ	ประเภทของถนนที่เกิดเหตุ	ประเภทของยานพาหนะที่เกิดเหตุ	รถคู่กรณี	อุณหภูมิ	ความชื้น	ความเร็วลม					สภาพอากาศ
1																	98.6749	0.986	0.987	0.982

รอบที่ 2 ถอดลักษณะข้อมูลออกไป 1 ลักษณะ

ตารางที่ 4.2

แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Decision tree (15 ตัวแปร)

ตัวแปร	ตัวแปร															Accuracy	Precision	Recall	F - measure	
	เพศ	อายุ	สถานะของผู้เกิดอุบัติเหตุ	แอลกอฮอล์	มาตรการ	การนำส่ง	ช่วงเวลา	วันที่เกิดเหตุ	จังหวัดที่เกิดเหตุ	ประเภทของถนนที่เกิดเหตุ	ประเภทของยานพาหนะที่เกิดเหตุ	รถคู่กรณี	อุณหภูมิ	ความชื้น	ความเร็วลม					สภาพอากาศ
1																	98.6749	0.986	0.987	0.982
2																	98.6749	0.986	0.987	0.982
3																	98.6750	0.986	0.987	0.982
4																	98.6739	0.986	0.987	0.982
5																	98.6749	0.986	0.987	0.982
6																	98.4111	0.984	0.984	0.992
7																	98.6733	0.986	0.987	0.982
8																	98.6749	0.986	0.987	0.982

ตารางที่ 4.2

แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Decision tree (15 ตัวแปร) (ต่อ)

ปี ปีฐาน	ตัวแปร															Accuracy	Precision	Recall	F - measure	
	เพศ	อายุ	สถานะของผู้เกิดอุบัติเหตุ	แอลกอฮอล์	มาตการ	การนำส่ง	ช่วงเวลา	วันที่เกิดเหตุ	จังหวัดที่เกิดเหตุ	ประเภทของถนนที่เกิดเหตุ	ประเภทของยานพาหนะที่เกิดเหตุ	รถคู่กรณี	อุณหภูมิ	ความชื้น	ความเร็วลม					สภาพอากาศ
9																	98.6733	0.986	0.987	0.982
10																	98.6684	0.986	0.987	0.982
11																	98.6749	0.986	0.987	0.982
12																	98.6506	0.987	0.987	0.982
13																	98.6753	0.986	0.987	0.982
14																	98.6749	0.986	0.987	0.982
15																	98.67	0.986	0.987	0.982
16																	98.6749	0.986	0.987	0.982

รอบที่ 3 ถอดลักษณะข้อมูลออกไป 2 ลักษณะ

ตารางที่ 4.3

แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Decision tree (14 ตัวแปร)

วิธี เชิง คณิต	ตัวแปร														Accuracy	Precision	Recall	F - measure		
	เพศ	อายุ	สถานะของผู้เกิดอุบัติเหตุ	แอลกอฮอล์	มาตรการ	การนำส่ง	ช่วงเวลา	วันที่เกิดเหตุ	จังหวัดที่เกิดเหตุ	ประเภทของถนนที่เกิดเหตุ	ประเภทของยานพาหนะที่เกิดเหตุ	รถจักรยานยนต์	อุณหภูมิ	ความชื้น					ความเร็วลม	สภาพอากาศ
1																	98.6749	0.986	0.987	0.982
2																	98.6749	0.986	0.987	0.982
3																	98.6756	0.986	0.987	0.982
4																	98.6742	0.986	0.987	0.982
5																	98.6749	0.986	0.987	0.982
6																	98.4117	0.984	0.984	0.972
7																	98.6733	0.986	0.987	0.982
8																	98.6749	0.986	0.987	0.982
9																	98.6730	0.986	0.987	0.982

ตารางที่ 4.3

แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Decision tree (14 ตัวแปร) (ต่อ)

ครั้งที่	ตัวแปร														Accuracy	Precision	Recall	F - measure		
	เพศ	อายุ	สถานะของผู้เกิดอุบัติเหตุ	แอลกอฮอล์	มาตรการ	การนำส่ง	ช่วงเวลา	วันที่เกิดเหตุ	จังหวัดที่เกิดเหตุ	ประเภทของถนนที่เกิดเหตุ	ประเภทของยานพาหนะที่เกิดเหตุ	รถคู่กรณี	อุณหภูมิ	ความชื้น					ความเร็วลม	สภาพอากาศ
10																	98.6684	0.986	0.987	0.982
11																	98.6749	0.986	0.987	0.982
12																	98.6506	0.987	0.987	0.982
13																	98.6749	0.986	0.987	0.982
14																	98.6722	0.986	0.987	0.982
15																	98.6752	0.986	0.987	0.982

รอบที่ 4 ถอดลักษณะข้อมูลออกไป 3 ลักษณะ

ตารางที่ 4.4

แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Decision tree (13 ตัวแปร)

ครั้งที่	ตัวแปร															Accuracy	Precision	Recall	F - measure	
	เพศ	อายุ	สถานะของผู้เกิดอุบัติเหตุ	แอลกอฮอล์	มาตรการ	การนำส่ง	ช่วงเวลา	วันที่เกิดเหตุ	จังหวัดที่เกิดเหตุ	ประเภทของถนนที่เกิดเหตุ	ประเภทของยานพาหนะที่เกิดเหตุ	รถคู่กรณี	อุณหภูมิ	ความชื้น	ความเร็วลม					สภาพอากาศ
1																	98.6755	0.986	0.987	0.982
2																	98.6751	0.986	0.987	0.982
3																	98.6742	0.986	0.987	0.982
4																	98.6749	0.986	0.987	0.982
5																	98.4225	0.984	0.984	0.972
6																	98.6733	0.986	0.987	0.982
7																	98.6753	0.986	0.987	0.982
8																	98.6750	0.986	0.987	0.982
9																	98.6684	0.986	0.987	0.982

ตารางที่ 4.4

แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Decision tree (13 ตัวแปร) (ต่อ)

รุ่น	ตัวแปร															Accuracy	Precision	Recall	F - measure	
	เพศ	อายุ	สถานะของผู้เกิดอุบัติเหตุ	แอลกอฮอล์	มาตรการ	การนำส่ง	ช่วงเวลา	วันที่เกิดเหตุ	จังหวัดที่เกิดเหตุ	ประเภทของถนนที่เกิดเหตุ	ประเภทของยานพาหนะที่เกิดเหตุ	รถคู่กรณี	อุณหภูมิ	ความชื้น	ความเร็วลม					สภาพอากาศ
10																	98.6751	0.96	0.987	0.982
11																	98.6516	0.987	0.987	0.982
12																	98.6749	0.986	0.987	0.982
13																	98.6758	0.986	0.987	0.984
14																	98.6755	0.986	0.987	0.983

รอบที่ 5 ถอดลักษณะข้อมูลออกไป 4 ลักษณะ

ตารางที่ 4.5

แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Decision tree (12 ตัวแปร)

ครั้งที่	ตัวแปร														Accuracy	Precision	Recall	F - measure	
	เพศ	อายุ	สถานะของผู้เกิดอุบัติเหตุ	แอลกอฮอล์	มาตรการ	การนำส่ง	ช่วงเวลา	วันที่เกิดเหตุ	จังหวัดที่เกิดเหตุ	ประเภทของถนนที่เกิดเหตุ	ประเภทของยานพาหนะที่เกิดเหตุ	รถจักรยานยนต์	อุณหภูมิ	ความชื้น					ความเร็วลม
1																98.6758	0.986	0.987	0.982
2																98.6758	0.986	0.987	0.982
3																98.6742	0.986	0.987	0.982
4																98.6749	0.986	0.987	0.982
5																98.4225	0.984	0.984	0.972
6																98.6733	0.986	0.987	0.982
7																98.6753	0.986	0.987	0.982
8																98.6751	0.986	0.987	0.982

ตารางที่ 4.5

แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Decision tree (12 ตัวแปร) (ต่อ)

ลำดับ	ตัวแปร														Accuracy	Precision	Recall	F - measure	
	เพศ	อายุ	สถานะของผู้เกิดอุบัติเหตุ	แอลกอฮอล์	มาตรการ	การนำส่ง	ช่วงเวลา	วันที่เกิดเหตุ	จังหวัดที่เกิดเหตุ	ประเภทของถนนที่เกิดเหตุ	ประเภทของยานพาหนะที่เกิดเหตุ	รถคู่กรณี	อุณหภูมิ	ความชื้น					ความเร็วลม
9																98.6695	0.986	0.987	0.982
10																98.6757	0.986	0.987	0.982
11																98.6716	0.987	0.987	0.982
12																98.6749	0.986	0.987	0.982
13																98.6760	0.986	0.987	0.984

รอบที่ 6 ถอดลักษณะข้อมูลออกไป 5 ลักษณะ

ตารางที่ 4.6

แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Decision tree (11 ตัวแปร)

ลำดับ	ตัวแปร															Accuracy	Precision	Recall	F - measure	
	เพศ	อายุ	สถานะของผู้เกิดอุบัติเหตุ	แอลกอฮอล์	มาตรการ	การนำส่ง	ช่วงเวลา	วันที่เกิดเหตุ	จังหวัดที่เกิดเหตุ	ประเภทของถนนที่เกิดเหตุ	ประเภทของยานพาหนะที่เกิดเหตุ	รถคู่กรณี	อุณหภูมิ	ความชื้น	ความเร็วลม					สภาพอากาศ
1																	98.6758	0.986	0.987	0.982
2																	98.6758	0.986	0.987	0.982
3																	98.6750	0.986	0.987	0.982
4																	98.6757	0.986	0.987	0.982
5																	98.4267	0.984	0.984	0.972
6																	98.6740	0.986	0.987	0.982
7																	98.6753	0.986	0.987	0.982
8																	98.6751	0.986	0.987	0.982

ตารางที่ 4.6

แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Decision tree (11 ตัวแปร) (ต่อ)

ชื่อ ตัวแปร	ตัวแปร														Accuracy	Precision	Recall	F - measure		
	เพศ	อายุ	สถานะของผู้เกิดอุบัติเหตุ	แอลกอฮอล์	มาตรการ	การนำส่ง	ช่วงเวลา	วันที่เกิดเหตุ	จังหวัดที่เกิดเหตุ	ประเภทของถนนที่เกิดเหตุ	ประเภทของยานพาหนะที่เกิดเหตุ	รถคู่กรณี	อุณหภูมิ	ความชื้น					ความเร็วลม	สภาพอากาศ
9																	98.6702	0.986	0.987	0.982
10																	98.6757	0.986	0.987	0.982
11																	98.6516	0.987	0.987	0.982
12																	98.6763	0.986	0.987	0.984

รอบที่ 7 ถอดลักษณะข้อมูลออกไป 6 ลักษณะ

ตารางที่ 4.7

แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Decision tree (10 ตัวแปร)

ชุดข้อมูล	ตัวแปร															Accuracy	Precision	Recall	F - measure	
	เพศ	อายุ	สถานะของผู้เกิดอุบัติเหตุ	แอลกอฮอล์	มาตรการ	การนำส่ง	ช่วงเวลา	วันที่เกิดเหตุ	จังหวัดที่เกิดเหตุ	ประเภทของถนนที่เกิดเหตุ	ประเภทของยานพาหนะที่เกิดเหตุ	รถคู่กรณี	อุณหภูมิ	ความชื้น	ความเร็วลม					สภาพอากาศ
1																	98.6749	0.986	0.987	0.982
2																	98.676	0.986	0.987	0.982
3																	98.6749	0.986	0.987	0.982
4																	98.6760	0.986	0.987	0.982
5																	98.4111	0.984	0.984	0.972
6																	98.6760	0.986	0.987	0.982
7																	98.6760	0.986	0.987	0.982
8																	98.6765	0.986	0.987	0.984

ตารางที่ 4.7

แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Decision tree (10 ตัวแปร) (ต่อ)

ปี ครั้งที่	ตัวแปร														Accuracy	Precision	Recall	F - measure		
	เพศ	อายุ	สถานะของผู้เกิดอุบัติเหตุ	แอลกอฮอล์	มาตรการ	การนำส่ง	ช่วงเวลา	วันที่เกิดเหตุ	จังหวัดที่เกิดเหตุ	ประเภทของถนนที่เกิดเหตุ	ประเภทของยานพาหนะที่เกิดเหตุ	รถคู่กรณี	อุณหภูมิ	ความชื้น					ความเร็วลม	สภาพอากาศ
9																	98.6695	0.986	0.987	0.982
10																	98.6760	0.986	0.987	0.982
11																	98.6505	0.987	0.987	0.982

รอบที่ 8 ถอดลักษณะข้อมูลออกไป 7 ลักษณะ

ตารางที่ 4.8

แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Decision tree (9 ตัวแปร)

ครั้งที่	ตัวแปร															Accuracy	Precision	Recall	F - measure	
	เพศ	อายุ	สถานะของผู้เกิดอุบัติเหตุ	แอลกอฮอล์	มาตรการ	การนำส่ง	ช่วงเวลา	วันที่เกิดเหตุ	จังหวัดที่เกิดเหตุ	ประเภทของถนนที่เกิดเหตุ	ประเภทของยานพาหนะที่เกิดเหตุ	รถคู่กรณี	อุณหภูมิ	ความชื้น	ความเร็วลม					สภาพอากาศ
1																	98.676	0.986	0.987	0.982
2																	98.6766	0.986	0.987	0.984
3																	98.6749	0.986	0.987	0.982
4																	98.6755	0.986	0.987	0.982
5																	98.4111	0.984	0.984	0.972
6																	98.6728	0.986	0.987	0.982
7																	98.6749	0.986	0.987	0.982

ตารางที่ 4.8

แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Decision tree (9 ตัวแปร) (ต่อ)

วิธี ซึ่งใช้	ตัวแปร															Accuracy	Precision	Recall	F - measure
	เพศ	อายุ	สถานะของผู้เกิดอุบัติเหตุ	แอลกอฮอล์	มาตรการ	การนำส่ง	ช่วงเวลา	วันที่เกิดเหตุ	จังหวัดที่เกิดเหตุ	ประเภทของถนนที่เกิดเหตุ	ประเภทของยานพาหนะที่เกิดเหตุ	รถคู่กรณี	อุณหภูมิ	ความชื้น	ความเร็วลม				
8																98.6695	0.986	0.987	0.982
9																98.6760	0.986	0.987	0.982
10																98.6506	0.987	0.987	0.982

รอบที่ 9 ถอดลักษณะข้อมูลออกไป 8 ลักษณะ

ตารางที่ 4.9

แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Decision tree (8 ตัวแปร)

ร.น. ๒๖๕	ตัวแปร															Accuracy	Precision	Recall	F - measure	
	เพศ	อายุ	สถานะของผู้เกิดอุบัติเหตุ	แอลกอฮอล์	มาตรการ	การนำส่ง	ช่วงเวลา	วันที่เกิดเหตุ	จังหวัดที่เกิดเหตุ	ประเภทของถนนที่เกิดเหตุ	ประเภทของยานพาหนะที่เกิดเหตุ	รถคู่กรณี	อุณหภูมิ	ความชื้น	ความเร็วลม					สภาพอากาศ
1																	98.6766	0.986	0.987	0.982
3																	98.6755	0.986	0.987	0.982
4																	98.6765	0.986	0.987	0.982
5																	98.4111	0.984	0.984	0.972
6																	98.6728	0.986	0.987	0.982
7																	98.6749	0.986	0.987	0.982
8																	98.6695	0.986	0.987	0.982
9																	98.6763	0.986	0.987	0.982
10																	98.6506	0.987	0.987	0.982

จากการใช้โปรแกรม Weka ในการวิเคราะห์ข้อมูลนั้นสามารถสรุปผลได้
ดังต่อไปนี้

ตารางที่ 4.10

สรุปค่า Accuracy และ F - measure ของแบบจำลองพยากรณ์

รอบที่	ค่า Accuracy	ค่า F - measure	คุณลักษณะที่ถูกถอดออก
1	98.6749	0.982	-
2	98.6753	0.982	อุณหภูมิ
3	98.6756	0.982	สถานะของผู้เกิดเหตุ
4	98.6758	0.984	ความเร็วลม
5	98.6760	0.984	สภาพอากาศ
6	98.6763	0.984	ความชื้น
7	98.6765	0.984	จังหวัดที่เกิดเหตุ
8	98.6766	0.984	อายุ
9	98.6766	0.982	-

เพราะฉะนั้นสามารถสรุปผลของวิเคราะห์ข้อมูลโดยที่คุรลักษณะที่มีผลต่อการเสียชีวิตทางถนนในช่วงเทศกาลปีใหม่นั้น ได้แก่ เพศ การดื่มแอลกอฮอล์ การคาดเข็มขัดนิรภัยหรือหมอกนิรภัย การนำส่ง ช่วงเวลา วันที่เกิดเหตุประเภทของถนนที่เกิดเหตุ ประเภทของยานพาหนะที่เกิดและรถคู่กรณี โดยที่ประสิทธิภาพของแบบจำลองพยากรณ์ เท่ากับ 0.984 และค่าความแม่นยำของแบบจำลองพยากรณ์ เท่ากับ 98.6766

การทดสอบโดยใช้ขั้นตอนวิธี Naives Bayes
 รอบที่ 1 ใช้ลักษณะข้อมูลทั้งหมด

ตารางที่ 4.11

แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Naïve Bayes (16 ตัวแปร)

ครั้งที่	ตัวแปร															Accuracy	Precision	Recall	F - measure	
	เพศ	อายุ	สถานะของผู้เกิดอุบัติเหตุ	แอลกอฮอล์	มาตรการ	การนำส่ง	ช่วงเวลา	วันที่เกิดเหตุ	จังหวัดที่เกิดเหตุ	ประเภทของถนนที่เกิดเหตุ	ประเภทของยานพาหนะที่เกิดเหตุ	รถคู่กรณี	อุณหภูมิ	ความชื้น	ความเร็วลม					สภาพอากาศ
1																	98.4138	0.98	0.984	0.981

รอบที่ 2 ทำการถอดลักษณะข้อมูลออกไป 1 ลักษณะ

ตารางที่ 4.12

แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Naïve Bayes (15 ตัวแปร)

ครั้งที่	ตัวแปร															Accuracy	Precision	Recall	F - measure	
	เพศ	อายุ	สถานะของผู้เกิดอุบัติเหตุ	แอลกอฮอล์	มาตรการ	การนำส่ง	ช่วงเวลา	วันที่เกิดเหตุ	จังหวัดที่เกิดเหตุ	ประเภทของถนนที่เกิดเหตุ	ประเภทของยานพาหนะที่เกิดเหตุ	รถคู่กรณี	อุณหภูมิ	ความชื้น	ความเร็วลม					สภาพอากาศ
1																	98.3889	0.98	0.984	0.981
2																	98.4669	0.98	0.985	0.981
3																	98.3981	0.98	0.984	0.981
4																	98.4192	0.98	0.984	0.981
5																	98.5525	0.982	0.986	0.982
6																	98.2838	0.972	0.983	0.976
7																	98.4257	0.98	0.984	0.981
8																	98.4187	0.98	0.984	0.981

ตารางที่ 4.12

แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Naïve Bayes (15 ตัวแปร) (ต่อ)

ครั้งที่	ตัวแปร														Accuracy	Precision	Recall	F - measure		
	เพศ	อายุ	สถานะของผู้เกิดอุบัติเหตุ	แอลกอฮอล์	มาตรการ	การนำส่ง	ช่วงเวลา	วันที่เกิดเหตุ	จังหวัดที่เกิดเหตุ	ประเภทของถนนที่เกิดเหตุ	ประเภทของยานพาหนะที่เกิดเหตุ	รถจักรยานยนต์	อุณหภูมิ	ความชื้น					ความเร็วลม	สภาพอากาศ
9																	98.403	0.98	0.984	0.981
10																	98.5238	0.981	0.985	0.982
11																	98.5568	0.982	0.986	0.982
12																	98.5422	0.982	0.985	0.982
13																	98.5503	0.982	0.986	0.982
14																	98.5422	0.982	0.985	0.982
15																	98.5533	0.983	0.986	0.982
16																	98.5503	0.982	0.986	0.982

รอบที่ 3 ทำการถอดลักษณะข้อมูลออกไป 2 ลักษณะ

ตารางที่ 4.13

แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Naïve Bayes (14 ตัวแปร)

ครั้งที่	ตัวแปร														Accuracy	Precision	Recall	F – measure	
	เพศ	อายุ	สถานะของผู้เกิดอุบัติเหตุ	แอลกอฮอล์	มาตรการ	การนำส่ง	ช่วงเวลา	วันที่เกิดเหตุ	จังหวัดที่เกิดเหตุ	ประเภทของถนนที่เกิดเหตุ	ประเภทของยานพาหนะที่เกิดเหตุ	รถจักรยาน	อุณหภูมิ	ความชื้น					ความเร็วลม
1																98.5442	0.98	0.985	0.982
2																98.5833	0.98	0.985	0.982
3																98.5503	0.98	0.984	0.982
4																98.5590	0.98	0.984	0.982
5																98.6327	0.982	0.986	0.982
6																98.3727	0.972	0.983	0.976
7																98.5612	0.98	0.984	0.982
8																98.5617	0.98	0.984	0.982

ตารางที่ 4.13

แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Naive Bayes (14 ตัวแปร) (ต่อ)

เครื่องที่	ตัวแปร														Accuracy	Precision	Recall	F – measure		
	เพศ	อายุ	สถานะของผู้เกิดอุบัติเหตุ	แอลกอฮอล์	มาตรการ	การนำส่ง	ช่วงเวลา	วันที่เกิดเหตุ	จังหวัดที่เกิดเหตุ	ประเภทของถนนที่เกิดเหตุ	ประเภทของยานพาหนะที่เกิดเหตุ	รถคู่กรณี	อุณหภูมิ	ความชื้น					ความเร็วลม	สภาพอากาศ
9																	98.5530	0.98	0.984	0.982
10																	98.6067	0.981	0.985	0.982
11																	98.6256	0.982	0.985	0.982
12																	98.5533	0.982	0.986	0.982
13																	98.5427	0.982	0.985	0.982
14																	98.5553	0.983	0.986	0.982
15																	98.5909	0.982	0.986	0.982

รอบที่ 4 ทำการถอดลักษณะข้อมูลออกไป 3 ลักษณะ

ตารางที่ 4.14

แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Naïve Bayes (13 ตัวแปร)

ครั้งที่	ตัวแปร													Accuracy	Precision	Recall	F - measure		
	เพศ	อายุ	สถานะของผู้เกิดอุบัติเหตุ	แอลกอฮอล์	มาตรการ	การนำส่ง	ช่วงเวลา	วันที่เกิดเหตุ	จังหวัดที่เกิดเหตุ	ประเภทของถนนที่เกิดเหตุ	ประเภทของยานพาหนะที่เกิดเหตุ	รถคู่กรณี	อุณหภูมิ					ความชื้น	ความเร็วลม
1																98.5492	0.98	0.985	0.982
2																98.5783	0.98	0.985	0.982
3																98.5533	0.98	0.984	0.982
4																98.6290	0.98	0.984	0.982
5																98.3897	0.972	0.983	0.976
6																98.5672	0.98	0.984	0.982
7																98.5657	0.98	0.984	0.982
8																98.5230	0.98	0.984	0.982

ตารางที่ 4.14

แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Naive Bayes (13 ตัวแปร) (ต่อ)

ครั้งที่	ตัวแปร															Accuracy	Precision	Recall	F - measure	
	เพศ	อายุ	สถานะของผู้เกิดอุบัติเหตุ	แอลกอฮอล์	มาตรการ	การนำส่ง	ช่วงเวลา	วันที่เกิดเหตุ	จังหวัดที่เกิดเหตุ	ประเภทของถนนที่เกิดเหตุ	ประเภทของยานพาหนะที่เกิดเหตุ	รถคู่กรณี	อุณหภูมิ	ความชื้น	ความเร็วลม					สภาพอากาศ
9																	98.6207	0.981	0.985	0.982
10																	98.6266	0.982	0.985	0.982
11																	98.5378	0.982	0.986	0.982
12																	98.5569	0.982	0.985	0.982
13																	98.5649	0.983	0.986	0.982
14																	98.6355	0.982	0.986	0.982

รอบที่ 5 ทำการถอดลักษณะข้อมูลออกไป 4 ลักษณะ

ตารางที่ 4.15

แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Naïve Bayes (12 ตัวแปร)

ครั้งที่	ตัวแปร														Accuracy	Precision	Recall	F - measure	
	เพศ	อายุ	สถานะของผู้เกิดอุบัติเหตุ	แอลกอฮอล์	มาตรการ	การนำส่ง	ช่วงเวลา	วันที่เกิดเหตุ	จังหวัดที่เกิดเหตุ	ประเภทของถนนที่เกิดเหตุ	ประเภทของยานพาหนะที่เกิดเหตุ	รถจักรยาน	อุณหภูมิ	ความชื้น					ความเร็วลม
1																98.5992	0.98	0.985	0.982
2																98.5783	0.98	0.985	0.982
3																98.5934	0.98	0.984	0.982
4																98.6304	0.98	0.984	0.982
5																98.3843	0.972	0.983	0.976
6																98.5431	0.98	0.984	0.982
7																98.5666	0.98	0.984	0.982

ตารางที่ 4.15

แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Naive Bayes (12 ตัวแปร) (ต่อ)

ครั้งที่	ตัวแปร															Accuracy	Precision	Recall	F - measure	
	เพศ	อายุ	สถานะของผู้เกิดอุบัติเหตุ	แอลกอฮอล์	มาตรการ	การนำส่ง	ช่วงเวลา	วันที่เกิดเหตุ	จังหวัดที่เกิดเหตุ	ประเภทของถนนที่เกิดเหตุ	ประเภทของยานพาหนะที่เกิดเหตุ	รถจักรยานยนต์	อุณหภูมิต่ำ	ความชื้น	ความเร็วลม					สภาพอากาศ
8																	98.5548	0.98	0.984	0.982
9																	98.6306	0.981	0.985	0.982
10																	98.6245	0.982	0.985	0.982
11																	98.6166	0.982	0.986	0.982
12																	98.6067	0.982	0.985	0.982
13																	98.6381	0.983	0.986	0.982

รอบที่ 6 ทำการถอดลักษณะข้อมูลออกไป 5 ลักษณะ

ตารางที่ 4.16

แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Naïve Bayes (11 ตัวแปร)

ครั้งที่	ตัวแปร															Accuracy	Precision	Recall	F – measure
	เพศ	อายุ	สถานะของผู้เกิดอุบัติเหตุ	แอลกอฮอล์	มาตรการ	การนำส่ง	ช่วงเวลา	วันที่เกิดเหตุ	จังหวัดที่เกิดเหตุ	ประเภทของถนนที่เกิดเหตุ	ประเภทของยานพาหนะที่เกิดเหตุ	รถคู่กรณี	อุณหภูมิ	ความชื้น	ความเร็วลม				
1																98.6142	0.98	0.985	0.982
2																98.5963	0.98	0.985	0.982
3																98.5994	0.98	0.984	0.982
4																98.6304	0.98	0.984	0.982
5																98.4330	0.973	0.983	0.976
6																98.5671	0.98	0.984	0.982
7																98.5697	0.98	0.984	0.982
8																98.5548	0.98	0.984	0.982

ตารางที่ 4.16

แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Naïve Bayes (11 ตัวแปร) (ต่อ)

วิธี ร.	ตัวแปร														Accuracy	Precision	Recall	F - measure	
	เพศ	อายุ	สถานะของผู้เกิดอุบัติเหตุ	แอลกอฮอล์	มาตรการ	การนำส่ง	ช่วงเวลา	วันที่เกิดเหตุ	จังหวัดที่เกิดเหตุ	ประเภทของถนนที่เกิดเหตุ	ประเภทของยานพาหนะที่เกิดเหตุ	รถคู่กรณี	อุณหภูมิ	ความชื้น					ความเร็วลม
9																98.6236	0.981	0.985	0.982
10																98.6335	0.982	0.985	0.982
11																98.6386	0.982	0.986	0.982
12																98.6167	0.982	0.985	0.982

รอบที่ 7 ทำการถอดลักษณะข้อมูลออกไป 6 ลักษณะ

ตารางที่ 4.17

แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Naïve Bayes (10 ตัวแปร)

ครั้งที่	ตัวแปร															Accuracy	Precision	Recall	F - measure	
	เพศ	อายุ	สถานะของผู้เกิดอุบัติเหตุ	แอลกอฮอล์	มาตรการ	การนำส่ง	ช่วงเวลา	วันที่เกิดเหตุ	จังหวัดที่เกิดเหตุ	ประเภทของถนนที่เกิดเหตุ	ประเภทของยานพาหนะที่เกิดเหตุ	รถคู่กรณี	อุณหภูมิ	ความชื้น	ความเร็วลม					สภาพอากาศ
1																	98.6222	0.98	0.985	0.982
2																	98.6003	0.98	0.985	0.982
3																	98.6034	0.98	0.984	0.982
4																	98.6254	0.98	0.984	0.982
5																	98.4510	0.973	0.983	0.976
6																	98.5679	0.98	0.984	0.982
7																	98.5847	0.98	0.984	0.982
8																	98.5588	0.98	0.984	0.982
9																	98.6406	0.981	0.985	0.982
10																	98.6445	0.982	0.985	0.982
11																	98.61873	0.982	0.985	0.982

รอบที่ 8 ทำการถอดลักษณะข้อมูลออกไป 7 ลักษณะ

ตารางที่ 4.18

แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Naïve Bayes (9 ตัวแปร)

ครั้งที่	ตัวแปร														Accuracy	Precision	Recall	F - measure	
	เพศ	อายุ	สถานะของผู้เกิดอุบัติเหตุ	แอลกอฮอล์	มาตรการ	การนำส่ง	ช่วงเวลา	วันที่เกิดเหตุ	จังหวัดที่เกิดเหตุ	ประเภทของถนนที่เกิดเหตุ	ประเภทของยานพาหนะที่เกิดเหตุ	รถจักรยาน	อุณหภูมิต่ำ	ความชื้น					ความเร็วลม
1																98.6289	0.98	0.985	0.982
2																98.6462	0.98	0.985	0.982
3																98.6240	0.98	0.984	0.982
4																98.6338	0.98	0.984	0.982
5																98.4024	0.973	0.983	0.976
6																98.6278	0.98	0.984	0.982
7																98.6316	0.98	0.984	0.982
8																98.6305	0.98	0.984	0.982
9																98.6489	0.981	0.985	0.982
10																98.6500	0.982	0.987	0.982

รอบที่ 9 ทำการถอดลักษณะข้อมูลออกไป 8 ลักษณะ

ตารางที่ 4.19

แสดงค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองพยากรณ์โดยใช้วิธี Naïve Bayes (8 ตัวแปร

ครั้งที่	ตัวแปร															Accuracy	Precision	Recall	F - measure	
	เพศ	อายุ	สถานะของผู้เกิดอุบัติเหตุ	แอลกอฮอล์	มาตรการ	การนำส่ง	ช่วงเวลา	วันที่เกิดเหตุ	จังหวัดที่เกิดเหตุ	ประเภทของถนนที่เกิดเหตุ	ประเภทของยานพาหนะที่เกิดเหตุ	รถจักรยาน	อุณหภูมิ	ความชื้น	ความเร็วลม					สภาพอากาศ
1																	98.6484	0.98	0.985	0.982
2																	98.6462	0.98	0.985	0.982
3																	98.6240	0.98	0.984	0.982
4																	98.6404	0.98	0.984	0.982
5																	98.4278	0.973	0.983	0.976
6																	98.6484	0.98	0.984	0.982
7																	98.6500	0.98	0.987	0.982
8																	98.6457	0.98	0.984	0.982
9																	98.6473	0.981	0.985	0.982

จากการใช้โปรแกรม Weka ในการวิเคราะห์ข้อมูลนั้นสามารถสรุปผลได้ดังต่อไปนี้

ตารางที่ 4.20

สรุปค่า Accuracy และ F – measure ของการทำ Naïve Bayes

รอบที่	ค่า Accuracy	ค่า F – measure	คุณลักษณะที่ถูกถอดออก
1	98.4138	0.981	-
2	98.5568	0.982	ประเภทของยานพาหนะที่เกิดเหตุ
3	98.6327	0.982	มาตรการ
4	98.6355	0.982	สภาพอากาศ
5	98.6381	0.982	ความเร็วลม
6	98.6386	0.982	อุณหภูมิ
7	98.6445	0.982	ความชื้น
8	98.6500	0.982	รถคู่กรณี
9	98.6500	0.982	-

เพราะฉะนั้นสามารถสรุปผลของวิเคราะห์ข้อมูลโดยที่คุณลักษณะที่มีผลต่อการเสียชีวิตทางถนนในช่วงเทศกาลปีใหม่นั้น ได้แก่ เพศ อายุ สถานะของผู้เกิดอุบัติเหตุ แอลกอฮอล์ การนำส่งช่วงเวลา วันที่เกิด จังหวัดที่เกิดเหตุ และประเภทของถนนที่เกิดเหตุ โดยที่ประสิทธิภาพของแบบจำลองพยากรณ์ เท่ากับ 0.982 และค่าความแม่นยำของแบบจำลองพยากรณ์ เท่ากับ 98.6500

4.1.2 ผลของการจัดกลุ่มลักษณะของประชากรที่โอกาสเสียชีวิตทางถนนในช่วงเทศกาลปีใหม่

1. ผลของการจัดกลุ่มประชากรโดยอิงจากลักษณะข้อมูลของวิธี Decision tree

ตารางที่ 4.21

แสดงการจัดกลุ่มประชากรตามคุณลักษณะข้อมูลตาม Decision tree โดยวิธี K - mean

ลักษณะข้อมูล	ชุดที่ 1 (เพศชาย)	ชุดที่ 2 (เพศหญิง)
วันที่	31 ธันวาคม	31 ธันวาคม
ช่วงเวลา	18.01 – 19.00	19.01 – 20.00
ประเภทของถนน	ทางหลวง	นอกเมือง
ประเภทของยานพาหนะที่เกิดเหตุ	มอเตอร์ไซด์	มอเตอร์ไซด์
รถคู่กรณี	รถปิคอัพ	ไม่มี
การคาดเข็มขัดนิรภัยหรือหมวกนิรภัย	ไม่สวมหมวก	ไม่สวมหมวก
การดื่มแอลกอฮอล์	ดื่ม	ไม่ดื่ม
การนำส่ง	อาสาสมัคร	ญาติผู้ป่วย
สถานะ	เสียชีวิต	เสียชีวิต

จากตาราง 4.21 แสดงเกี่ยวกับการแบ่งกลุ่มประชากรที่มีโอกาสเสียชีวิต โดยแบ่งออกเป็น 2 กลุ่ม คือเพศชาย และเพศหญิง โดยที่ข้อมูลทั้งสองกลุ่มมีความคล้ายคลึงกันคือ วันที่เกิดเหตุคือเป็นวันที่ 31 ธันวาคม และช่วงเวลาที่เกิดอุบัติเหตุจะมีความใกล้เคียงกันมาก คือ 18.01 - 19.00 และ 19.01 - 20.00 โดยสถานที่เกิดเหตุมักเป็นถนนที่ใช้ความเร็วได้มากคือ ถนนทางหลวง และถนนนอกเมือง ซึ่งผู้เกิดเหตุมักขี่รถจักรยานยนต์โดยไม่สวมหมวกนิรภัย และการนำส่งถึงแพทย์มักไม่ได้ส่งด้วยการแพทย์ คือส่งด้วยอาสาสมัคร และนำส่งกันเอง โดยที่ลักษณะข้อมูลที่มีความแตกต่างกันชัดเจนนั้นคือ เพศชายมักดื่มแอลกอฮอล์และรถคู่กรณีมักเป็นรถปิคอัพ ส่วนเพศหญิงจะไม่ดื่มและไม่มียานพาหนะ โดยที่การจัดกลุ่มประชากรนี้มีค่าความแม่นยำอยู่ที่ประมาณ 56.53%

2. ผลของการจัดกลุ่มประชากรโดยอิงจากลักษณะข้อมูลของวิธี Naive Bayes

ตารางที่ 4.22

แสดงการจัดกลุ่มประชากรตามคุณลักษณะข้อมูลตาม Naive Bayes โดยวิธี K - mean

ลักษณะข้อมูล	ชุดที่ 1 (เพศชาย)	ชุดที่ 2 (เพศหญิง)
วันที่	31 ธันวาคม	31 ธันวาคม
ช่วงเวลา	21.01 – 22.00	18.01 – 19.00
ประเภทของถนน	ทางหลวง	นอกเมือง
จังหวัดที่เกิดเหตุ	นครราชสีมา	เชียงใหม่
สถานะของผู้เกิดอุบัติเหตุ	คนขับ	คนขับ
อายุ	30.8	40.23
การดื่มแอลกอฮอล์	ดื่ม	ไม่ดื่ม
การนำส่ง	ญาติผู้ป่วย	ญาติผู้ป่วย
สถานะ	เสียชีวิต	เสียชีวิต

จากตาราง 4.22 แสดงเกี่ยวกับการแบ่งกลุ่มประชากรที่มีโอกาสเสียชีวิต โดยแบ่งออกเป็น 2 กลุ่ม คือเพศชาย และเพศหญิง โดยที่ข้อมูลทั้งสองกลุ่มมีความคล้ายคลึงกันคือ วันที่เกิดเหตุคือเป็นวันที่ 31 ธันวาคม และช่วงเวลาที่เกิดอุบัติเหตุจะมีความใกล้เคียงกันคือ 21.01 – 22.00 และ 18.01 – 19.00 โดยสถานที่เกิดเหตุมักเป็นถนนที่ใช้ความเร็วได้มากคือ ถนนทางหลวงและถนนนอกเมือง โดยที่เกิดเหตุในจังหวัดที่มักเป็นเมืองใหญ่หรือแหล่งท่องเที่ยว นอกจากนี้ผู้เสียชีวิตมักเป็นคนขับซึ่งและการนำส่งนั้นมักเป็นคนที่มีความรู้เรื่องการปฐมพยาบาลเบื้องต้นน้อย นั่นคือ ญาติผู้ป่วย โดยที่ลักษณะข้อมูลที่มีความแตกต่างกันชัดเจนนั้นคือ อายุของผู้เสียชีวิต และการดื่มแอลกอฮอล์

4.1.3 ผลของการจัดแบบจำลองพยากรณ์

ผลของการทำแบบจำลองการพยากรณ์เพื่อทำนายการเสียชีวิตทางถนนในช่วงเทศกาลปีใหม่ โดยใช้วิธีของ Decision Tree และ Naive Bayes เปรียบเทียบกัน พร้อมทั้งทำการปรับปรุงประสิทธิภาพของแบบจำลองและค่าความแม่นยำของแบบจำลองด้วยวิธีการถอดลักษณะของข้อมูลที่ละตัวเพื่อหาตัวแบบที่ประสิทธิภาพและความแม่นยำที่ดีที่สุดของทั้งสองขั้นตอนวิธี หลังจากนั้นจึงทำการเปรียบเทียบตัวแบบที่ดีที่สุดของแต่ละขั้นตอนวิธี นำมาเปรียบเทียบประสิทธิภาพและความแม่นยำที่ได้จากทั้งสองขั้นตอนวิธี ได้ดังตาราง

ตารางที่ 4.23

แสดงการเปรียบเทียบค่า Accuracy และ F - measure โดยวิธี Decision tree กับ Naïve Bayes

ลำดับที่	วิธีการ	Accuracy	Precision	Recall	F - measure
1	Decision Tree	98.6766	0.987	0.987	0.984
2	Naïve Bayes	98.6500	0.982	0.987	0.982

จากตาราง 4.23 แสดงให้เห็นว่าแบบจำลองการพยากรณ์การเสียชีวิตทางถนนในช่วงเทศกาลปีใหม่ด้วยวิธีแบบ Decision Tree ด้วยลักษณะข้อมูล เพศ การดื่มแอลกอฮอล์ การคาดเข็มขัดนิรภัยหรือหมอกนิรภัย การนำส่ง ช่วงเวลา วันที่เกิดเหตุ ประเภทของถนนที่เกิดเหตุ ประเภทของยานพาหนะที่เกิด และรถคู่กรณี ในการสร้างแบบจำลองการพยากรณ์ให้ค่าประสิทธิภาพที่ดีกว่าแบบ Decision Forest โดยมีค่าความถูกต้องของการทำนายจากชุดข้อมูลทดสอบอยู่ที่ 98.6766% และตัวแบบมีประสิทธิภาพที่จะนำไปใช้กับข้อมูลใหม่ที่ไม่เคยเห็นมาก่อนอยู่ที่ 98.4%

4.2 การอภิปรายผล

4.2.1 ผลของการวิจัยเปรียบเทียบกับงานวิจัยอื่น

งานวิจัยฉบับนี้เป็นการศึกษาถึงความสัมพันธ์ของปัจจัยด้านต่างๆ ได้แก่ ด้านประชากร ด้านนโยบายสาธารณะ และด้านสภาพแวดล้อม ที่ส่งผลทำให้ประชากรมีโอกาสที่จะเสียชีวิตจากอุบัติเหตุทางถนนในช่วงเทศกาลปีใหม่ โดยจากผลการวิจัยนั้นพบว่าตัวแปรที่มีผลต่อการเสียชีวิตจากอุบัติเหตุทางถนนในช่วงเทศกาลปีใหม่ ได้แก่ เพศ การดื่มแอลกอฮอล์ การคาดเข็มขัดนิรภัยหรือหมอกนิรภัย การนำส่ง ช่วงเวลา วันที่เกิดเหตุ ประเภทของถนนที่เกิดเหตุ ประเภทของยานพาหนะที่เกิด และรถคู่กรณี โดยที่ตัวแปรในส่วนของ การนำส่งนั้นมีผลต่อการเสียชีวิตมากที่สุด ซึ่งสามารถสังเกตได้จากเมื่อมีการตัวแปรการนำส่งออกไป จะทำให้ทั้งค่าความแม่นยำและค่าประสิทธิภาพของแบบจำลองมีค่าน้อยลงอย่างเห็นได้ชัด

โดยงานวิจัยฉบับนี้นั้นผลที่ได้มีความสอดคล้องกับงานวิจัยอื่นๆ ซึ่งจากงานวิจัยของ Saraee, M, Kerry, J, Llyod, M and Markey, C (2005) , S.Krishnaveni (2011) , Tibebe Beshah, Shawndra Hill ซึ่งงานวิจัยของทั้ง 2 ฉบับนี้เป็นการศึกษาเกี่ยวกับการเกิดอุบัติเหตุโดยใช้วิธีการของ Data mining โดยใช้เครื่องมือ Weka ในการหาความสัมพันธ์นั้นพบว่าสอดคล้องกับ

งานวิจัยฉบับนี้ คือ เพศ ประเภทของถนนที่เกิดเหตุมีความสอดคล้องกัน นอกจากนี้ในงานวิจัยของ Sarae, M, Kerry, J, Llyod, M and Markey, C (2005) ยังมีสิ่งที่สอดคล้องกับงานวิจัยฉบับนี้ใน ส่วนของ เพศชายมีแนวโน้มที่จะเกิดอุบัติเหตุและมักเกิดบนถนนในส่วนของถนนทางหลวง (High way) แต่มีตัวแปรที่ไม่สอดคล้องกับงานวิจัยฉบับนี้ คือ ตัวแปรด้านอายุ ซึ่งงานวิจัยของ Sarae and other (2005) พบว่าอายุมีผลต่อการเกิดอุบัติเหตุ

ในส่วนของตัวแปรของประเภทของยานพาหนะที่เกิด และรถคู่กรณี มีผลต่อการเกิดอุบัติเหตุและเสียชีวิตในงานวิจัยของ Sarae and other (2005) และ S.Krishnaveni (2011) พบว่ามีผลต่อการเกิดอุบัติเหตุและอาจนำไปสู่การเสียชีวิต ซึ่งสอดคล้องกับงานวิจัยฉบับนี้ที่แสดงให้เห็นว่า ประเภทของยานพาหนะที่เกิด และรถคู่กรณี มีผลต่อการเกิดอุบัติเหตุและเสียชีวิต

ในส่วนของตัวแปรการดื่มแอลกอฮอล์ และการคาดเข็มขัดนิรภัยหรือสวมหมวกนิรภัยนั้น งานวิจัยของ Sami and other (2009) และบทความของ WHO มีความสอดคล้องกับงานวิจัยฉบับนี้โดยที่พบว่า การดื่มแอลกอฮอล์มีผลต่อการเกิดอุบัติเหตุเป็นอย่างมาก โดยที่ WHO พบว่าสามารถลดการชนและเสียชีวิตที่มีสาเหตุจากการดื่มแอลกอฮอล์ได้ประมาณ 20% และหากมีการสวมหมวกนิรภัยหรือคาดเข็มขัดนิรภัยจะสามารถลดการเสียชีวิตลงได้ประมาณ 40% และลดการบาดเจ็บอย่างรุนแรงได้ถึง 70%

ในส่วนของตัวแปรช่วงเวลา วันที่เกิดเหตุ และการนำส่งนั้นยังไม่พบว่ามีงานวิจัยที่นำมาอ้างอิงในครั้งนี้นั้นมีความเกี่ยวข้องกับการเกิดอุบัติเหตุและการเสียชีวิต

นอกจากในงานวิจัยนี้ที่มีค่า accuracy ที่สูงมากจนเกินไป ผู้วิจัยมีความคิดเห็นว่าเกิดจากปัญหาที่มีชื่อว่า “Accuracy Paradox” คือเกิดจากความไม่สมดุลของปริมาณข้อมูลในส่วนของคลาสที่เป็นบวกและลบ หรือสามารถอธิบายได้ว่า ปริมาณข้อมูลของคลาสผู้เสียชีวิตกับคลาสของผู้ไม่เสียชีวิตมีปริมาณที่ห่างกันมากจนเกินไป ทำให้ Machine learning สามารถคำนวณค่าความแม่นยำได้สูงเกินความเป็นจริง

4.2.2 ผลของการวิจัยเทียบกับแนวคิดทฤษฎีโดมิโน

ทฤษฎีโดมิโน (Domino Theory) ของการเกิดอุบัติเหตุ สามารถเชื่อมโยงได้กับปรัชญาความปลอดภัยของ H.W. Heinrich เกี่ยวกับสาเหตุของอุบัติเหตุได้ทฤษฎีโดมิโน กล่าวว่า การบาดเจ็บและความเสียหายต่างๆ เป็นผลที่สืบเนื่องโดยตรงมาจากอุบัติเหตุและอุบัติเหตุเป็นผลมาจากการกระทำที่ไม่ปลอดภัยหรือสภาพการณ์ที่ไม่ปลอดภัย ซึ่งเปรียบได้เหมือนตัวโดมิโนที่เรียงกันอยู่ 5 ตัวใกล้กัน เมื่อตัวที่หนึ่งล้มย่อมมีผลทำให้ตัวโดมิโนถัดไปล้มตามกันไปด้วย ตัวโดมิโนทั้งห้าตัว ได้แก่

1. สภาพแวดล้อมหรือภูมิหลังของบุคคล (Social Environment of Background) ได้แก่ สภาพครอบครัว ฐานะความเป็นอยู่การศึกษา อายุ เป็นต้น

2. ความบกพร่องผิดปกติของบุคคล (Defect of Person) ได้แก่ การปฏิบัติงานโดยขาดความยั้งคิด อารมณ์รุนแรง การขาดความรอบคอบ หรือการเพิกเฉยต่อความปลอดภัย เป็นต้น

3. การกระทำหรือสภาพการณ์ที่ไม่ปลอดภัย (Unsafe Acts/Unsafe Conditions) ได้แก่ ขาดเครื่องป้องกันจุดอันตราย แสงสว่างไม่เพียงพอ หรือมีสารเคมีฟุ้งกระจาย เป็นต้น

4. อุบัติภัย (Accident) คือเหตุการณ์ที่ไม่คาดคิดว่าจะเกิดขึ้น ได้แก่ ลื่นหกล้ม เดินสะดุด สิ่งของผลัดหล่นจากที่สูง ไฟไหม้ เป็นต้น

5. การบาดเจ็บหรือเสียหาย (Injury or Damage) ได้แก่ การบาดเจ็บหรือสูญเสียอวัยวะ เช่น กระดูกหักหรือแตก ได้รับความแผล หรือแม้กระทั่งถึงแก่ความตาย เป็นต้น

งานวิจัยนี้เป็นการสร้างแบบจำลองการพยากรณ์การเสียชีวิตจากอุบัติเหตุบนท้องถนนในช่วงเทศกาลปีใหม่นั้นมีความสอดคล้องกับทฤษฎีโดมิน่าทั้ง 3 ขั้นตอนดังต่อไปนี้

ขั้นตอนที่ 1 สภาพแวดล้อมหรือภูมิหลังของบุคคล (Social Environment of Background) คือ อายุ เพศ สถานะของผู้เกิดอุบัติเหตุ มีผลต่อการเกิดอุบัติเหตุ

ขั้นที่ 2 ความบกพร่องผิดปกติของบุคคล (Defect of Person) ได้แก่ การดื่มแอลกอฮอล์ มีผลต่อการเกิดอุบัติเหตุ

ขั้นที่ 3 การกระทำหรือสภาพการณ์ที่ไม่ปลอดภัย (Unsafe Acts/Unsafe Conditions) ได้แก่ ช่วงเวลา วันที่เกิดเหตุ ประเภทของถนนที่เกิดเหตุ ประเภทของยานพาหนะที่เกิด และการสวมหมวกนิรภัยหรือการคาดเข็มขัดนิรภัย มีผลต่อการเกิดอุบัติเหตุ

ในส่วนของตัวแปรเรื่องการนำส่งนั้น ผู้วิจัยคิดว่าเป็นตัวป้องกันไม่ให้เกิดการล้มต่อนื่องของโดมิโนตัวที่ 4 กับโดมิโนตัวที่ 5 ซึ่งก็คือหลังจากที่เกิดอุบัติเหตุแล้วพยายามไม่ให้เกิดการสูญเสียอวัยวะหรือชีวิต

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

การศึกษาค้นคว้าวิจัยนี้ เป็นการศึกษาเพื่อเป็นแนวทางในการแก้ปัญหาเรื่องการเสียชีวิตทางถนนในช่วงเทศกาลปีใหม่ โดยที่ปัญหานี้เป็นปัญหาที่เกิดขึ้นมาอย่างยาวนานแม้ว่าทางรัฐบาลพยายามที่จะแก้ปัญหานี้แต่ก็ยังไม่ประสบผลสำเร็จ ในงานวิจัยนี้จึงได้ทำการศึกษาแยกออกเป็น 2 รูปแบบ คือ ลักษณะประชากรที่มีโอกาสที่จะเสียชีวิตทางถนนในช่วงเทศกาลปีใหม่ และการสร้างแบบจำลองพยากรณ์สำหรับการเสียชีวิตทางถนนในช่วงเทศกาลปีใหม่เพื่อศึกษาว่ามีคุณลักษณะใดที่มีผลต่อการเสียชีวิต โดยการใช้ข้อมูลของสถาบันการแพทย์ฉุกเฉินแห่งชาติช่วงปี พ.ศ. 2551 - 2558 เป็นเวลาทั้งหมด 8 ปี

ในขั้นตอนของการดำเนินงานนั้น ทางผู้วิจัยได้แบ่งขั้นตอนการดำเนินงานออกเป็น 5 ขั้นตอน โดยเริ่มจาก การศึกษาและการวางแผนการเก็บข้อมูล, การเก็บรวบรวมข้อมูล, การสำรวจข้อมูลและจัดการข้อมูล, การวิเคราะห์และออกแบบจำลองการพยากรณ์, การอธิบายผลและการนำไปใช้ โดยที่ทางผู้วิจัยใช้การเรียนรู้ของเครื่องทำการเรียนรู้จากชุดข้อมูลที่มี โดยแบ่งเป็นการเรียนรู้ของเครื่องแบบมีผู้สอน คือ การแบ่งชุดข้อมูลที่มีออกเป็น 2 ส่วน คือข้อมูลส่วนที่ใช้ในการสอนเครื่อง และข้อมูลส่วนที่ใช้ในการทดสอบการให้เครื่องเรียนรู้จากข้อมูลทั้งหมดที่มีและอีกส่วนใช้วิธีการเรียนรู้ของเครื่องแบบไม่มีผู้สอน เพื่อทำการจัดกลุ่มที่มีลักษณะข้อมูลคล้ายกันให้อยู่ในกลุ่มเดียวกัน

5.1 สรุปผลการวิจัย

จากการนำข้อมูลการเกิดอุบัติเหตุของสถาบันการแพทย์ฉุกเฉินแห่งชาติ ผนวกกับข้อมูลเรื่อง ความเร็วลม สภาพอากาศ ความชื้น และ อุณหภูมิ จาก <https://www.wunderground.com> ในช่วงระยะเวลา 8 ปี (พ.ศ.2551-2558) โดยการเก็บข้อมูลแบบทุติยภูมิ เพื่อมาทำการศึกษาโดยใช้การเรียนรู้ของเครื่อง เพื่อหาแนวทางในการลดปัญหาการเสียชีวิตจากอุบัติเหตุทางถนนในช่วงเทศกาลปีใหม่ พร้อมทั้งแบบจำลองการพยากรณ์การเสียชีวิตทางถนนในช่วงเทศกาลปีใหม่ ซึ่งสามารถสรุปได้ดังนี้

5.1.1 แบบจำลองการพยากรณ์

ผลที่ได้จากการสร้างแบบจำลองการพยากรณ์ โดยวิธีที่เหมาะสมในการวิจัยนี้คือ Decision Tree ซึ่งให้ค่าประสิทธิภาพของตัวแบบพยากรณ์อยู่ที่ 98.4% และค่าความแม่นยำอยู่ที่

98.6766% โดยในการพยากรณ์มีการใช้ข้อมูลทั้งจากข้อมูลด้านประชากร ด้านนโยบายสาธารณะ และด้านสภาพแวดล้อม

5.1.2 การจัดกลุ่มข้อมูล

ผลที่ได้จากการจัดกลุ่มข้อมูลเกี่ยวกับประชากรที่มีโอกาสที่จะเสียชีวิตในช่วงเทศกาลปีใหม่โดยสามารถแบ่งออกมาได้เป็น 2 กลุ่มคือ เพศชาย และเพศหญิง ส่วนใหญ่จะมีลักษณะที่เหมือนหรือคล้ายคลึงกัน จะมีบางลักษณะข้อมูลที่แตกต่างกันอย่างชัดเจน เช่น ช่วงอายุ การดื่ม แอลกอฮอล์ และรถคู่กรณี ซึ่งค่าความยำอยู่ที่ 56.33% ซึ่งอาจยังมีลักษณะข้อมูลอื่นที่อาจมีผลต่อความแม่นยำ

5.2 แนวทางการนำไปใช้

ผู้วิจัยได้มีการทำแนวทางเกี่ยวกับการวิจัยการจำลองแบบพยากรณ์การเสียชีวิตทางถนนในช่วงเทศกาลปีใหม่ไว้ และการจัดกลุ่มประชากรที่มีโอกาสที่จะเสียชีวิตทางถนนในช่วงเทศกาลปีใหม่ไว้ โดยผู้วิจัยหวังว่างานวิจัยครั้งนี้จะสามารถช่วยให้หน่วยงานต่างๆของทางภาครัฐและเอกชนที่ส่วนเกี่ยวข้องกับการลดการเกิดอุบัติเหตุและการสูญเสียชีวิตทางถนนในช่วงเทศกาลปีใหม่ มีแนวทางเพื่อนำไปประยุกต์ใช้เพื่อที่จะลดการสูญเสียให้มากที่สุดเท่าที่จะเป็นไปได้ และจากผลการวิจัยในครั้งนี้ทำให้ทราบว่าปัจจัยที่มีผลต่อการเสียชีวิตที่มีผลมากที่สุด คือ การนำส่ง โดยวิธีการนำส่งนั้นจะขึ้นอยู่กับบุคคลที่ทำการโทรแจ้งกับหน่วยทางการแพทย์จากนั้นหน่วยงานทางการแพทย์จะประเมินอาการผู้ป่วยจากผู้โทรแจ้งจึงจะส่งหน่วยแพทย์ที่เหมาะสมกับอาการของผู้ป่วย ซึ่งจะเห็นได้ว่าถ้าหากผู้โทรแจ้งมีความรู้เกี่ยวกับการสังเกตอาการผู้ได้รับบาดเจ็บน้อยอาจมีผลต่อการส่งหน่วยแพทย์ได้ นอกจากนี้หากมีการปฐมพยาบาลเบื้องต้นที่ถูกวิธีอาจมีผลทำให้อัตราการเสียชีวิตของผู้ได้รับบาดเจ็บลดลง ดังนั้นผู้วิจัยเห็นว่าทางภาครัฐควรที่จะมีการจัดสอนการเรียนการสอนหรือการอบรมเรื่องการปฐมพยาบาลเบื้องต้นให้กับนักเรียน นักศึกษา และประชาชนทั่วไป เพื่อที่จะไม่ให้มีการปฐมพยาบาลเบื้องต้นผิดวิธี หรือไม่ได้มีการปฐมพยาบาลเบื้องต้นกับผู้ได้รับบาดเจ็บเลย

5.3 ข้อจำกัดงานวิจัย

การวิจัยครั้งนี้ผู้ทำการวิจัยพบเจอข้อจำกัดในการทำงานวิจัยดังต่อไปนี้

1. คุณลักษณะข้อมูลยังมีไม่ครบ ในข้อมูลที่ได้มานั้นการบันทึกข้อมูลยังขาดคุณลักษณะข้อมูลที่อาจส่งผลต่อการเกิดอุบัติเหตุและทำให้เสียชีวิตอีก เช่น ความเร็วของการขับขี่ของผู้เกิดอุบัติเหตุ สภาพยานพาหนะของผู้เกิดอุบัติเหตุ ลักษณะของถนนที่เกิดเหตุ เป็นต้น

2. ข้อมูลที่ได้จากสถาบันการแพทย์ฉุกเฉินแห่งชาติยังไม่เป็นปัจจุบัน เนื่องจากข้อมูลที่ใช้ในการวิเคราะห์นั้นผู้วิจัยใช้ข้อมูลในปี 2551 – 2558 ซึ่งถือว่ายังไม่เป็นปัจจุบัน

3. คอมพิวเตอร์ของผู้ใช้ทำวิจัยไม่สามารถดำเนินการได้บางวิธีการ เนื่องจากข้อมูลที่มีจำนวนค่อนข้างมาก และเครื่องคอมพิวเตอร์ของผู้ทำวิจัยมี Ram 4 GB และ CPU I5 เมื่อทำการดำเนินวิธีการ Random Forest ไม่สามารถดำเนินการออกมาได้ มักเกิดข้อผิดพลาดที่โปรแกรมแสดงออกมา

5.4 ประโยชน์ที่ได้รับ

การทำวิจัยครั้งนี้ผู้ทำวิจัยได้ใช้การเรียนรู้ของเครื่องเข้ามาช่วยในการวิเคราะห์ พยากรณ์ และจัดลักษณะกลุ่มของประชากร เกี่ยวกับเรื่องการเสียชีวิตทางถนนในช่วงเทศกาลปีใหม่ และยังสามารถก่อให้เกิดประโยชน์ในด้านต่าง ๆ ดังนี้

5.4.1 ด้านวิชาการ

งานวิจัยครั้งนี้ผู้ทำวิจัยได้ใช้โปรแกรม Weka ซึ่งเป็นโปรแกรมเกี่ยวกับการใช้เครื่องของการเรียนรู้เพื่อมาช่วยวิเคราะห์ พยากรณ์ และจัดกลุ่มของประชากร ซึ่งอาจเป็นประโยชน์ต่อการแก้ปัญหาเรื่องอุบัติเหตุและการลดการสูญเสียในภายภาคหน้า หรืออาจเป็นส่วนหนึ่งของงานวิจัยที่ผู้อื่นนำไปปรับเปลี่ยนดัดแปลงเพื่อเป็นการต่อยอด

5.4.2 ด้านผู้วิจัย

ผู้ทำวิจัยได้ทำการศึกษาข้อมูลเกี่ยวกับการเกิดอุบัติเหตุตามทฤษฎีต่างๆ พร้อมทั้งมีการนำหลักการของกระบวนการทางวิทยาศาสตร์ด้านข้อมูล (Data science) มาผนวกกันเพื่อหาคุณลักษณะที่อาจก่อให้เกิดการเสียชีวิตในช่วงเทศกาลปีใหม่ รวมถึงเทคนิคต่างๆของการทำกระบวนการทางวิทยาศาสตร์ด้านข้อมูล (Data science) ซึ่งผู้ทำวิจัยคาดว่าหลักการและความรู้เกี่ยวกับกระบวนการทางวิทยาศาสตร์ด้านข้อมูล (Data science) จะเป็นประโยชน์ในภายภาคหน้าต่อสายงานอาชีพของผู้ทำวิจัย

5.4.3 ด้านภาครัฐหรือหน่วยงานที่เกี่ยวข้อง

ผู้ทำวิจัยคาดหวังว่างานวิจัยครั้งนี้จะเป็นประโยชน์ต่อทางภาครัฐและหน่วยงานที่เกี่ยวข้องที่มีหน้าที่ในการดูแลเกี่ยวกับอุบัติเหตุในช่วงเทศกาลปีใหม่ เพื่อเป็นแนวทางหรืออาจทำให้เกิดความคิดใหม่เพื่อที่จะนำไปประยุกต์กับแผนงานของภาครัฐหรือหน่วยงานที่เกี่ยวข้องกับเรื่องนี้เพื่อที่จะแก้ไขปัญหาที่ออกให้ดียิ่งขึ้น

5.5 ข้อเสนอแนะ

ผู้วิจัยจึงได้จัดทำข้อเสนอแนะสำหรับงานวิจัยในครั้งนี้ เพื่อเป็นแนวทางในการปรับปรุงพัฒนางานวิจัย หรือใช้เป็นแนวทางเพื่อประยุกต์ในงานวิจัยนี้ต่อไปในอนาคต

1. ผลลัพธ์ของงานวิจัยนี้ได้มีการกำหนดปัจจัยและคุณลักษณะของข้อมูลไว้ระดับหนึ่ง แต่อาจยังไม่ครอบคลุมถึงปัจจัยละคุณลักษณะของการเสียชีวิตจากการเสียชีวิตทางถนนในช่วงเทศกาลปีใหม่ก็เป็นไปได้
2. ในส่วนของขั้นตอนในการดำเนินการสร้างแบบจำลองการพยากรณ์นั้น ยังมีมีเทคนิคอื่นๆที่ผู้ทำวิจัยยังไม่ได้นำมาใช้ เช่น Random Forest , Neural network ซึ่งอาจทำให้ได้ผลลัพธ์ที่ดียิ่งขึ้น รวมทั้งการปรับแต่งค่าข้อมูลต่างๆ หรือพารามิเตอร์ต่างๆ ที่อาจทำให้ผลการวิจัยมีประสิทธิภาพเพิ่มมากขึ้น
3. ควรมีการทำการคัดเลือกข้อมูลปริมาณคลาสของผู้ที่เสียชีวิตกับคลาสของผู้ไม่เสียชีวิตให้มีปริมาณข้อมูลใกล้เคียงกัน เพื่อป้องกันการเกินปัญหา “Accuracy Paradox” คือการที่มีค่าความแม่นยำของโมเดลสูงจนเกินไป

5.6 แนวทางการทำวิจัยในขั้นถัดไป

จากการทำวิจัยในครั้งนี้ ผู้ทำวิจัยได้มีการศึกษาและทดลองเพื่อทำการสร้างแบบจำลองพยากรณ์การเสียชีวิตทางถนนในช่วงเทศกาลปีใหม่ เพื่อที่จะเป็นการพัฒนาวิทยาศาสตร์ข้อมูลในด้านนี้ ผู้ทำวิจัยจึงขอเสนอแนะผู้ที่ต้องการศึกษาหรือต่อยอดงานวิจัยนี้ หรือในลักษณะคล้าย ๆ กับผู้วิจัย จึงได้วางแนวทางในการวิจัยขั้นต่อไปไว้ดังนี้

1. การวิจัยในครั้งนี้มีข้อจำกัดในเรื่องของการเก็บรวบรวมข้อมูล เนื่องจากข้อมูลที่ได้มานั้นยังขาดคุณลักษณะข้อมูลอื่นอีก ดังนั้นหากต้องการทำงานวิจัยต่อยอดนั้นควรมีคุณลักษณะข้อมูลอื่นเข้ามาเพิ่ม เช่น เรื่องความเร็ว เรื่องการใช้อุปกรณ์ขณะขับขี่หรือไม่ เรื่องสภาพยานพาหนะของผู้ขับขี่ หรือสภาพจิตใจของผู้ขับขี่ เป็นต้น ดังนั้นหากมีการเพิ่มคุณลักษณะข้อมูลเหล่านี้เข้าไป ผู้ทำวิจัยเชื่อว่าน่าจะทำให้งานวิจัยนี้มีความสมบูรณ์เพิ่มมากขึ้น
2. ในทำวิจัยครั้งนี้ปัญหาเรื่องของความเหมาะสมในการจัดการข้อมูล ตั้งแต่การกรองการคัดออก รวมถึงการแยกค่าผิดปกติออกจากชุดข้อมูล ซึ่งสิ่งเหล่านี้จำเป็นต้องใช้ประสบการณ์และความชำนาญในการทำ และเรียนรู้ความเหมาะสมในการจัดการเพื่อให้ได้ชุดข้อมูลที่ดีพอจะนำไปใช้ในการสร้างแบบจำลองพยากรณ์

รายการอ้างอิง

สื่ออิเล็กทรอนิกส์

ดร. เอกสิทธิ์ พัทธวงศ์ศักดิ์, การคัดเลือก feature (feature selection) ด้วยวิธี Information Gain, <http://dataminingtrend.com/2014/data-mining-techniques/feature-selection-information-gain/>, สืบค้นเมื่อวันที่ 15 ธันวาคม 2561

บทความวารสาร

Bouckaert Remco, Eibe Frank, Mark Hall, Richard Kirkby, Peter Reutemann, and AlexSeewald, (2008) .WEKA Manual for Version 3-6-0. University of Waikato, New Zealand.

Brijesh Kumar Baradwaj, Saurabh Pal,” Mining Educational Data to Analyze Students Performance”.

Maninder Singh, Amrit Kaur, (2013). A Review on Road Accident in Traffic System using Data Mining Techniques, India.

Lluís Sanmiquel, Josep M. Rossell , Carla Vintró, (2015) Study of Spanish mining accidents using data mining techniques, Spain.

S.Krishnaveni, (2011) A Perspective Analysis of Traffic Accident using DataMining Techniques, India.

Articles

Domingos, Pedro & Michael Pazzani (1997) "On the optimality of the simple Bayesian classifier under zero-one loss". Machine Learning, 29:103–137.

E. Frank and I. H. Witten. (1998) Generating accurate rule sets without global optimization. In Proc. of the Int^l Conf. on Machine Learning, pages 144–151. Morgan Kaufmann Publishers Inc.

Han, Jiawei and Kamber, Micheline. (2006),” Data Mining: concepts and Techniques. San Fransisco”,

Mehmed Kantardzic (2003), “Data Mining: Concepts,Models, Methods, and Algorithms” ISBN13:9780471228523, John Wiley & Sons Publisher.

Miaou, S.P. and Harry, L. (1993) , “Modeling vehicle accidents and highway geometric design relationships”. Accidents Analysis and Prevention, (6), pp. 689–709.27.

Desktop Reference for Crash Reduction Factors Report No. FHWA-SA-07-015, Federal Highway Administration September, 2007 <http://www.ite.org/safety/issuebriefs/Desktop%20Reference%20Complete.pdf>

National Highway Traffic Safety Administration, Traffic Safety Facts 2005, 2007, P. 54. <http://www.nrd.nhtsa.dot.gov/Pubs/TSF2006.PDF>

S. Shanthi, R. Geetha Ramani " Feature Relevance Analysis and Classification of Road Traffic Accident Data through Data Mining Techniques" Proceedings of the World Congress on Engineering and Computer Science 2012 Vol. I WCECS 2012, October 24-26, 2012, San Francisco, USA.



ประวัติผู้เขียน

ชื่อ	นาย พิเศษ คุปตินทร
วันเดือนปีเกิด	18 กันยายน 2532
ตำแหน่ง	ผู้ตรวจสอบเทคโนโลยีสารสนเทศ
ประวัติการทำงาน	2562 – ปัจจุบัน : ผู้ตรวจสอบเทคโนโลยีสารสนเทศ บริษัท ธนาकारกรุงเทพ จำกัด (มหาชน) 2560 – 2561 : ผู้ดูแลความปลอดภัยเทคโนโลยี สารสนเทศ เชี่ยวชาญพิเศษ บริษัท คอมพิวเตอร์ยูเนียน จำกัด 2557 – 2560 : วิศวกรดูแลระบบ บริษัท ซีดีจี ซิสเต็ม จำกัด 2556 – 2557 : วิศวกรดูแลระบบ บริษัท เมโทร คอนเนค จำกัด