การพยากรณ์การเกิดอุบัติเหตุบนท้องถนนโดยใช้เทคนิคอนุกรมเวลา

วิทยานิพนธ์

ของ

ปทิตญา บุญรักษา

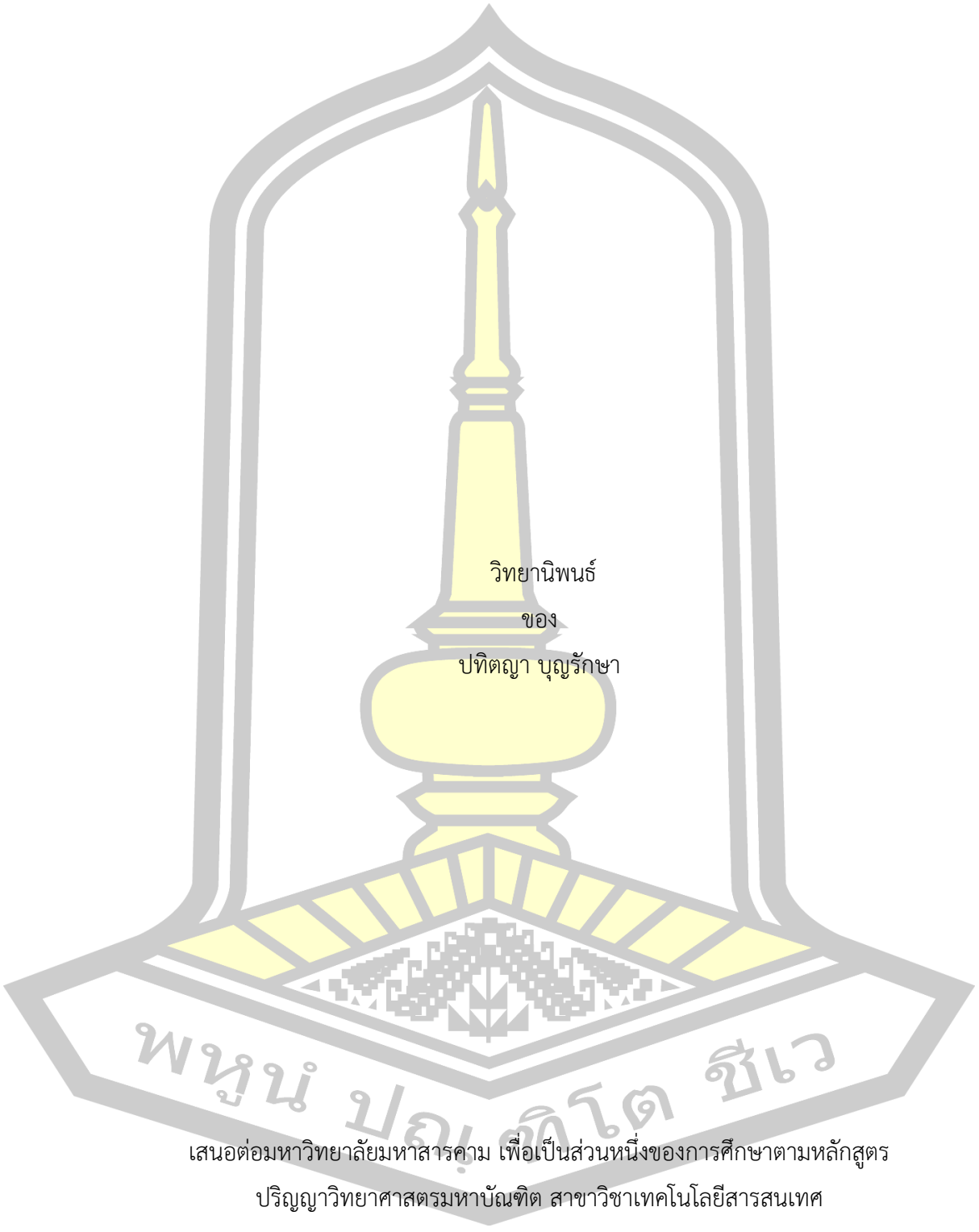
เสนอต่อมหาวิทยาลัยมหาสารคาม เพื่อเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ

ตุลาคม 2561

สงวนลิขสิทธิ์เป็นของมหาวิทยาลัยมหาสารคาม

การพยากรณ์การเกิดอุบัติเหตุบนท้องถนนโดยใช้เทคนิคอนุกรมเวลา



วิทยานิพนธ์

ของ

ปทิตญา บุญรักษา

เสนอต่อมหาวิทยาลัยมหาสารคาม เพื่อเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

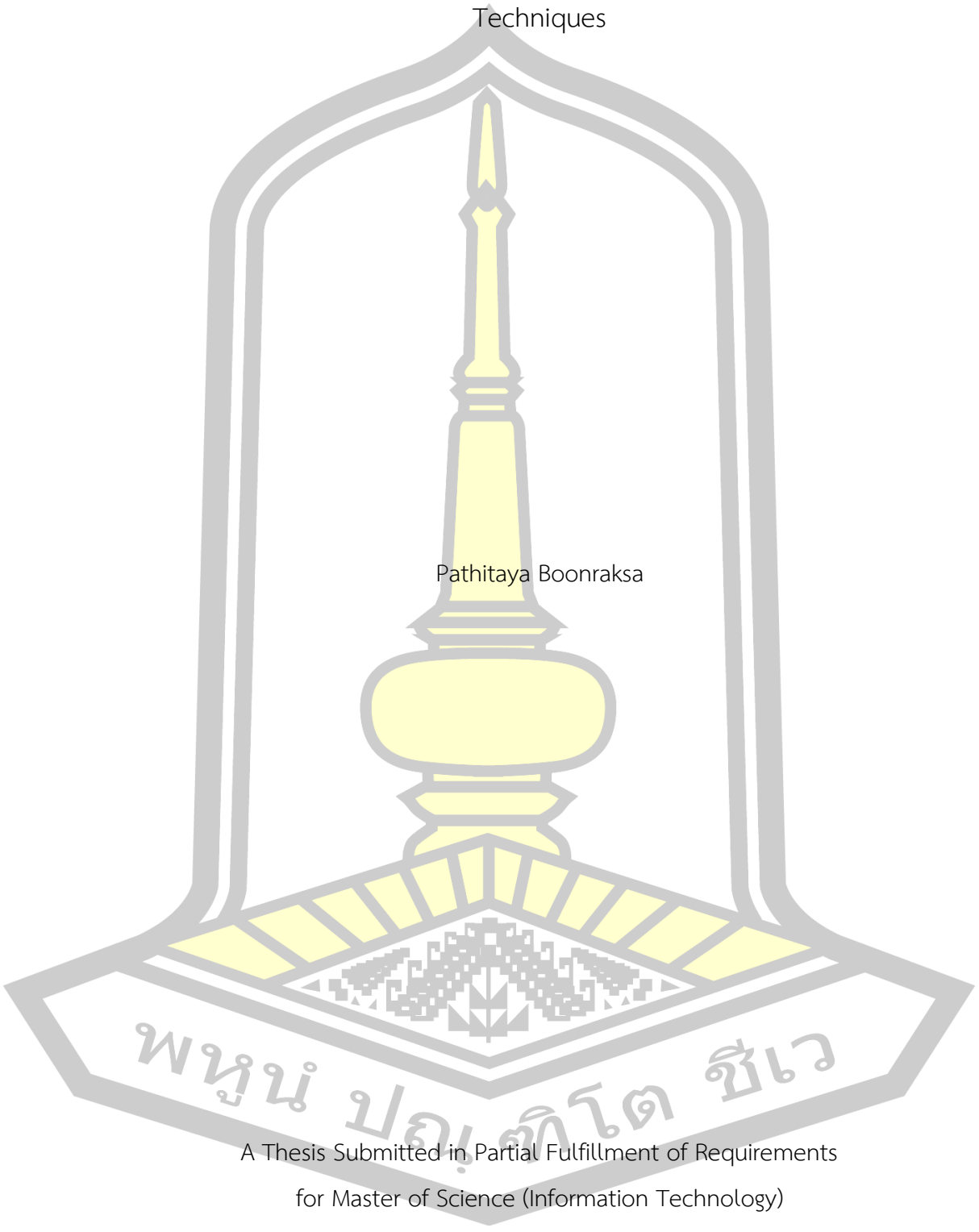
ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ

ตุลาคม 2561

สงวนลิขสิทธิ์เป็นของมหาวิทยาลัยมหาสารคาม

Predicting the Occurrence of Accidents on the Road using Time Series Techniques

Pathitaya Boonraksa



A Thesis Submitted in Partial Fulfillment of Requirements

for Master of Science (Information Technology)

October 2018

Copyright of Mahasarakham University

คณะกรรมการสอบวิทยานิพนธ์ ได้พิจารณาวิทยานิพนธ์ของนางสาวปทิตญา บุญรักษา แล้วเห็นสมควรรับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญา วิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ ของมหาวิทยาลัยมหาสารคาม

คณะกรรมการสอบวิทยานิพนธ์

.....ประธานกรรมการ
(ผศ. ดร. แกมกาญจน์ สมประเสริฐศรี)

.....อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(ผศ. ดร. จารี ทองคำ)

.....กรรมการ
(อ. ดร. สาธิต แสงประดิษฐ์)

.....กรรมการผู้ทรงคุณวุฒิภายนอก
(รศ. ดร. สิทธิชัย บุษหมั่น)

มหาวิทยาลัยอนุมัติให้รับวิทยานิพนธ์ฉบับนี้ เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญา วิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ ของมหาวิทยาลัยมหาสารคาม

.....
(ผศ. ดร. สุจิน บุตรดีสุวรรณ)
คณบดีคณะวิทยาการสารสนเทศ

.....
(ผศ. ดร. กริสน์ ชัยมูล)
คณบดีบัณฑิตวิทยาลัย

วัน.....เดือน.....ปี.....

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้ สำเร็จลงได้ด้วยดี เนื่องจากได้รับความกรุณาอย่างสูงจาก ผู้ช่วยศาสตราจารย์ ดร.จารี ทองคำ อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก ที่ได้กรุณาให้คำแนะนำปรึกษาตลอดจนปรับปรุงแก้ไขข้อบกพร่องต่าง ๆ ด้วยความเอาใจใส่อย่างดียิ่ง ผู้วิจัยตระหนักถึงความตั้งใจจริงและความทุ่มเทของอาจารย์ และขอกราบขอบพระคุณเป็นอย่างสูงไว้ ณ ที่นี้

ผู้วิจัยขอกราบขอบพระคุณ รศ.ดร.สิทธิชัย บุษหมั่น ผศ.ดร.แกมกาญจน์ สมประเสริฐศรี และ ดร.สาธิต แสงประดิษฐ์ ที่กรุณาเป็นกรรมการสอบวิทยานิพนธ์และให้คำปรึกษา ตลอดจนคำแนะนำต่างๆ รวมถึงทุกท่านที่มีส่วนช่วยให้งานวิจัยนี้สำเร็จลงได้มา ณ โอกาสนี้

ขอขอบพระคุณ ศูนย์ข้อมูลการจราจร จากสถานีตำรวจภูธรจังหวัดขอนแก่น ที่ให้ความอนุเคราะห์ ข้อมูลในการดำเนินงานวิจัยครั้งนี้จนทำให้งานวิจัยนี้สำเร็จลุล่วงไปด้วยดี

ผู้วิจัยมีความหวังว่า งานวิจัยฉบับนี้จะมีประโยชน์อยู่ไม่น้อย จึงขอมอบส่วนดีทั้งหมดนี้ ให้แก่เหล่าคณาจารย์ ที่ได้ให้คำแนะนำและแนวทางในการแก้ไขปัญหา

และขอมอบความกตัญญูกตเวทิตาคุณ แด่บิดา มารดา ครอบครัว ตลอดจนผู้ที่เป็นที่รักและผู้มีพระคุณทุกท่าน ที่เป็นแรงผลักดันและกำลังใจให้สามารถผ่านพ้นปัญหาไปได้ด้วยดี

สำหรับข้อบกพร่องต่าง ๆ ที่อาจจะเกิดขึ้นนั้น ผู้วิจัยก็ขออภัย และยินดีที่จะรับฟังคำแนะนำจากทุกท่านที่ได้เข้ามาศึกษา เพื่อเป็นประโยชน์ในการพัฒนางานวิจัยต่อไป

ปทิตญา บุญรักษา

พหูน ปณฺฑิโต ชีเว

ชื่อเรื่อง การพยากรณ์การเกิดอุบัติเหตุบนท้องถนนโดยใช้เทคนิคอนุกรมเวลา
ผู้วิจัย ปทิตญา บุญรักษา
อาจารย์ที่ปรึกษา ผู้ช่วยศาสตราจารย์ ดร. จารี ทองคำ
ปริญญา วิทยาศาสตรมหาบัณฑิต **สาขาวิชา** เทคโนโลยีสารสนเทศ
มหาวิทยาลัย มหาวิทยาลัยมหาสารคาม **ปีการศึกษา** 2561

บทคัดย่อ

การพยากรณ์การเกิดอุบัติเหตุโดยใช้เทคนิคอนุกรมเวลาให้มีความแม่นยำ เป็นงานวิจัยที่ท้าทาย งานวิจัยนี้จึงมีวัตถุประสงค์ เพื่อสร้างแบบจำลองที่มีประสิทธิภาพในการพยากรณ์โดยใช้เทคนิคอนุกรมเวลา คือ Linear regression (LR), Artificial Neural Networks: ANN (ANN), Sequential Minimal Optimization for Regression (SMOreg), Support Vector Machine for Regression (SVMR) และ Gussian Process (GP) และสามารถแสดงผลการพยากรณ์ผ่านทางเว็บไซต์ โดยใช้ข้อมูลจากงานอุบัติเหตุจราจรรายเดือนของ สถานีตำรวจอำเภอเมืองขอนแก่น ตั้งแต่เดือนมกราคม 2554 ถึงเดือนมิถุนายน 2560 จำนวน 78 เดือน การวัดประสิทธิภาพแบบจำลอง ได้ทดลองแบ่งชุดข้อมูลออกเป็นชุดข้อมูลเรียนรู้และ ชุดข้อมูลทดสอบ ทั้งหมดจำนวน 7 รอบ โดยการใช้หลักการ Sliding window และวัดประสิทธิภาพการพยากรณ์ของแบบจำลองด้วยค่า mean absolute error (MAE) และ root mean square error (RMSE) เพื่อวัดประสิทธิภาพของแบบจำลองนั้น

จากการศึกษาพบว่าเทคนิค SVMR สามารถพยากรณ์การเกิดอุบัติเหตุบนท้องถนนด้วยเทคนิคอนุกรมเวลา ได้ค่าความผิดพลาดต่ำที่สุด จากยานพาหนะที่ 1.42 จากเพศที่ 8.99 และจากการสูญเสีย 3.19 เมื่อเปรียบเทียบกับแบบจำลองที่สร้างด้วยเทคนิค LR,ANN, SMOreg และ Gussian Process เทคนิค SVMR มีความเหมาะสมมากที่สุดในการพยากรณ์การเกิดอุบัติเหตุบนท้องถนน

คำสำคัญ : การพยากรณ์, อุบัติเหตุ, อุบัติเหตุบนท้องถนน, อนุกรมเวลา

TITLE Predicting the Occurrence of Accidents on the Road using Time Series Techniques

AUTHOR Pathitaya Boonraksa

ADVISORS Assistant Professor Jaree Thongkam , Ph.D.

DEGREE Master of Science **MAJOR** Information Technology

UNIVERSITY Mahasarakham University **YEAR** 2018

ABSTRACT

Accident Forecasting using Time Series Techniques for Accuracy. It is a challenging research. This research aims to create efficient models for forecasting using time series technique. Linear regression (LR), Artificial Neural Networks: ANN (ANN), Sequential Minimal Optimization for Regression (SMOreg), Support Vector Machine for Regression (SVMR) และ Gussian Process (GP) .And can display the forecast results through the website. The website uses data from the monthly traffic accident's work. Police Station in Khon Kaen from January 2011 to June 2017 78 months. Modeling Performance Measurement Experiment the data into a learning set and A total of 7 test data sets using the sliding window principle.And measure the predictive efficiency of the model with the mean absolute error (MAE) and root mean square error (RMSE) to measure the efficiency of the model.The study found that SVMR techniques can predict road accidents by time series technique. The lowest error value. From vehicles at 1.42, from 8.99, and from a loss of 3.19, compared to LR, ANN, SMOreg and Gussian Process models, the SVMR technique is most suitable for predicting road accidents.

Keyword : Forecasting, Accident, Road Accidents, Time Series

สารบัญ

หน้า

บทคัดย่อภาษาไทย ..... ง

บทคัดย่อภาษาอังกฤษ ..... จ

กิตติกรรมประกาศ ..... ฉ

สารบัญ ..... ช

สารบัญตาราง ..... ฌ

สารบัญภาพ ..... ญ

บทที่ 1 บทนำ 1

1.1 หลักการและเหตุผล 1

1.2 วัตถุประสงค์ของการวิจัย 2

1.3 ความสำคัญของการวิจัย 3

1.4 ขอบเขตของงานวิจัย 3

1.5 นิยามศัพท์เฉพาะ 3

บทที่ 2 เอกสารและงานวิจัยที่เกี่ยวข้อง 4

2.1 ข้อมูลทั่วไปของจังหวัดขอนแก่น 4

2.2 อุบัติเหตุบนท้องถนน 5

2.3 เหมืองข้อมูล 10

2.4 อนุกรมเวลา 12

2.5 เทคนิคการพยากรณ์แบบอนุกรมเวลา 16

2.6 งานวิจัยที่เกี่ยวข้องกับการเกิดอุบัติเหตุบนท้องถนน 22

2.7 งานวิจัยที่เกี่ยวข้องกับเทคนิคอนุกรมเวลา 25

บทที่ 3 วิธีการดำเนินวิจัย 29

3.1 การศึกษาและวิเคราะห์ข้อมูล 29
3.2 การทำกระบวนการก่อนการสร้างแบบจำลอง 30
3.3 การสร้างแบบจำลอง 31
3.4 การวัดประสิทธิภาพ 38
บทที่ 4 ผลการศึกษา 40
4.1 การวิเคราะห์ข้อมูลข้อมูลการเกิดอุบัติเหตุ 40
4.2 การเปรียบเทียบการทำงานของแบบจำลอง 43
4.3 ผลการพยากรณ์การเกิดอุบัติเหตุบนท้องถนน ของจังหวัดขอนแก่น 49
4.4 การแสดงผลเว็บไซต์ของการพยากรณ์การเกิดอุบัติเหตุล่วงหน้า 58
บทที่ 5 สรุป วิจารณ์ และข้อเสนอแนะ 59
5.1 สรุป 59
5.2 วิจารณ์ 60
5.3 ข้อเสนอแนะ 61
บรรณานุกรม 62
ภาคผนวก 68
ประวัติผู้เขียน 78

พหุน ปณ ฑิโต ชีเว

สารบัญตาราง

หน้า

ตารางที่ 3.1 คำอธิบายตัวแปรและชนิดของตัวแปร 30

ตารางที่ 4.1 เปรียบเทียบค่า MAE ของ LR, ANN, SMOreg, SVMR และ GP โดยแบ่งตามยานพาหนะ 44

ตารางที่ 4.2 ตารางเปรียบเทียบค่า MAE ของ LR, ANN, SMOreg, SVMR และ GP โดยแบ่งตามเพศ 45

ตารางที่ 4.3 ตารางเปรียบเทียบค่า MAE ของ LR, ANN, SMOreg, SVMR และ GP โดยแบ่งตามการสูญเสียที่เกิดขึ้น 45

ตารางที่ 4.4 ตารางเปรียบเทียบค่า RMSE ของ ของ LR, ANN, SMOreg, SVMR และ GP แบ่งตามยานพาหนะ 46

ตารางที่ 4.5 ตารางเปรียบเทียบค่า RMSE ของ ของ LR, ANN, SMOreg, SVMR และ GP แบ่งตามเพศ 48

ตารางที่ 4.6 ตารางเปรียบเทียบค่า RMSE ของ ของ LR, ANN, SMOreg, SVMR และ GP แบ่งตามการสูญเสียที่เกิดขึ้น 48

พหุน ปณ ฑิโต ชีเว

สารบัญภาพ

หน้า

ภาพประกอบที่ 2.1 สัญลักษณ์ของจังหวัดขอนแก่น 4
ภาพประกอบที่ 2.2 แผนที่จังหวัดขอนแก่น 5
ภาพประกอบที่ 2.3 กราฟแสดงค่าแนวโน้มของผลผลิตเคมีภัณฑ์ชนิดหนึ่ง 14
ภาพประกอบที่ 2.4 กราฟแสดงยอดขายรายเดือนของห้างสรรพสินค้าแห่งหนึ่ง 14
ภาพประกอบที่ 2.5 กราฟการเกิดระยะต่าง ๆที่เกิดอย่างต่อเนื่องเป็นวัฏจักร 15
ภาพประกอบที่ 2.6 แบบแผนอนุกรมเวลา 16
ภาพประกอบที่ 2.7 การทำงานของ Linear Regression 17
ภาพประกอบที่ 2.8 แสดง Model ของ Neuron ในสมองมนุษย์ 18
ภาพประกอบที่ 2.9 รูปแบบการทำงานของ ANN แบบ Multi-Layer Prpceptron 19
ภาพประกอบที่ 2.10 หลักการทำงานของ Guassian Process 21
ภาพประกอบที่ 3.1 ตัวอย่างข้อมูลการเกิดอุบัติเหตุบนท้องของจังหวัดขอนแก่น 29
ภาพประกอบที่ 3.2 Data transformation 30
ภาพประกอบที่ 3.3 ตัวอย่างหลักการ Sliding Windows ที่ใช้ในงานวิจัย จำนวน 7 รอบ พ.ศ. 2560 39
ภาพประกอบที่ 4.1 จำนวนอุบัติเหตุที่เกิดจากยานพาหนะ จากมกราคม พ.ศ. 2554 40
ภาพประกอบที่ 4.2 มูลค่าความเสียหายที่เกิดขึ้น จากมกราคม พ.ศ. 2554 ถึง มิถุนายน พ.ศ.2560 41
ภาพประกอบที่ 4.3 จำนวนของเพศหญิง และชายที่เกิดอุบัติเหตุ มกราคม พ.ศ. 2554 ถึง มิถุนายน พ.ศ.2560 41
ภาพประกอบที่ 4.4 จำนวนความสูญเสียที่เกิดขึ้นของเพศชาย จากเกิดอุบัติเหตุ มกราคม พ.ศ. 2554 ถึง มิถุนายน พ.ศ.2560 42
ภาพประกอบที่ 4.5 จำนวนความสูญเสียที่เกิดขึ้นของเพศหญิง จากเกิดอุบัติเหตุ มกราคม พ.ศ. 2554 ถึง มิถุนายน พ.ศ.2560 42

ภาพประกอบที่ 4.6 ข้อมูลการเกิดอุบัติเหตุประเภทต่างๆ 49

ภาพประกอบที่ 4.7 ผลการพยากรณ์การเกิดอุบัติเหตุบนท้องถนน จากประเภทยานพาหนะ 50

ภาพประกอบที่ 4.8 ผลการพยากรณ์การเกิดอุบัติเหตุบนท้องถนน จากประเภทยานพาหนะ 51

ภาพประกอบที่ 4.9 ผลการพยากรณ์การเกิดอุบัติเหตุบนท้องถนน จากประเภทยานพาหนะ 52

ภาพประกอบที่ 4.10 ผลการพยากรณ์การเกิดอุบัติเหตุบนท้องถนน โดยแบ่งตามเพศ 53

ภาพประกอบที่ 4.11 ผลการพยากรณ์การเกิดอุบัติเหตุบนท้องถนน โดยแบ่งตามความสูญเสียที่เกิดขึ้น 54

ภาพประกอบที่ 4.12 ผลการพยากรณ์รายเดือน การเกิดอุบัติเหตุของ คนเดินเท้า ,รถจักรยาน ,รถสามล้อ และรถจักรยานยนต์ 55

ภาพประกอบที่ 4.13 ผลการพยากรณ์รายเดือน การเกิดอุบัติเหตุของ สามล้อเครื่อง,รถยนต์ ,รถตู้ , รถบัส, รถบรรทุกเล็ก, รถบรรทุก 6 ล้อ, รถบรรทุก10 ล้อ, และรถการเกษตร 56

ภาพประกอบที่ 4.14 ผลการพยากรณ์รายเดือน การเกิดอุบัติเหตุของรถแท็กซี่ และ รถอื่นๆ 57

ภาพประกอบที่ 4.15 ผลการพยากรณ์รายเดือน การสูญเสียเพศชายและหญิง 57

ภาพประกอบที่ 4.16 ผลการพยากรณ์รายเดือน การเสียชีวิตเพศชายและหญิง การบาดเจ็บสาหัสเพศชายและหญิง, บาดเจ็บเพศชายและหญิง 58

ภาพประกอบที่ 4.17 การแสดงผลการพยากรณ์การเกิดอุบัติเหตุล่วงหน้า 58

พหุน ปณฺ ฑิโต ชีเว

บทที่ 1
บทนำ

1.1 หลักการและเหตุผล

การเกิดอุบัติเหตุจราจรทางบกเป็นปัญหาที่สำคัญปัญหาหนึ่งของโลกและมีแนวโน้มความรุนแรงของการเกิดอุบัติเหตุจราจรทางบกเพิ่มมากขึ้น ซึ่งไม่เพียงจะก่อให้เกิดผลเสียต่อร่างกาย ชีวิต และทรัพย์สินแต่ยังสร้างความเสียหายทางเศรษฐกิจของประเทศ จากสถิติอุบัติเหตุจราจรจากสำนักงานตำรวจแห่งชาติ จะพบว่ามีผู้ประสบอุบัติเหตุ 328,400 ครั้ง ทั่วทั้งประเทศ ในปี พ.ศ. 2560 ซึ่งมีจำนวนผู้บาดเจ็บ 375,402 ราย ทุพพลภาพ 1,414 และเสียชีวิต 10,483 ราย โดยเฉพาะจังหวัดขอนแก่นเป็นเมืองที่ขนส่งสินค้าเส้นใหญ่นี้เพิ่มมากขึ้น มีความเสี่ยงการเกิดอุบัติเหตุบนท้องถนน ในปี พ.ศ. 2560 มีผู้ประสบอุบัติเหตุเกิดขึ้น 5,847 ครั้ง เป็นจำนวน 6,695 ราย มีจำนวนผู้บาดเจ็บ 6,460 ราย ทุพพลภาพ 30 ราย และเสียชีวิต 205 ราย [1] พบว่า สาเหตุหลักของการเกิดอุบัติเหตุปัจจัยที่ส่งผลให้เกิดอุบัติเหตุทางถนนประกอบไปด้วย ความผิดพลาดของผู้ใช้ถนน ความบกพร่องของรถ ถนนและสิ่งแวดล้อม สาเหตุหลักของการเกิดอุบัติเหตุ 3 อันดับแรก ได้แก่ ขับรถเร็วเกินอัตราที่กำหนด ตัดหน้ากระชั้นชิด และแซงรถผิดกฎหมาย โดยมีจำนวนมากกว่าหนึ่งในสามของสาเหตุอุบัติเหตุทั้งหมด 39.5% ซึ่ง สอดคล้องกับผลการศึกษาในประเทศสหรัฐอเมริกาที่พบว่า อุบัติเหตุที่เกี่ยวข้องกับความเร็วและส่งผลให้มีผู้เสียชีวิตมีสัดส่วน ประมาณ 1 ใน 3 ของอุบัติเหตุทั้งหมด [2] ในแต่ละปี มีจำนวนอุบัติเหตุทางถนนเฉลี่ย ประมาณ 15,000 ครั้ง [3] สถานการณ์ของปัญหาอุบัติเหตุการจราจรทางถนนยังเพิ่มขึ้นเรื่อยๆ ในทุกเดือน ไม่เพียงแต่เฉพาะช่วงเทศกาลเท่านั้น การเกิดอุบัติเหตุจราจรทางบกสามารถควบคุมป้องกันได้จากการวิจัยต่างประเทศ และการใช้บังคับกฎหมายสามารถลดการเกิดอุบัติเหตุจราจรทางบกได้เพียงน้อยมากเมื่อเทียบกับการเกิดอุบัติเหตุจราจรทางบกทั้งหมดโดยเป็นปัญหาสำคัญ สำหรับผู้ใช้ยานพาหนะในการคมนาคม โดยเฉพาะอย่างยิ่งในเขตที่มีการจราจรหนาแน่นคับคั่ง เช่น ในเขตตำบลในเมือง อำเภอเมือง จังหวัดขอนแก่น จากสถิติการเกิดอุบัติเหตุในแต่ละปีมักมีผู้เสียชีวิต และได้รับบาดเจ็บเป็นจำนวนมาก

ข้อมูลอุบัติเหตุจราจรทางบก เป็นข้อมูลอนุกรมเวลา ซึ่งเป็นชุดของข้อมูลที่เก็บรวบรวมตามระยะเวลารายวันต่อเนื่องกัน ข้อมูลดังกล่าวสามารถวิเคราะห์ได้ด้วยการวิเคราะห์อนุกรมเวลา (Time Series Analysis) [4] ซึ่งเป็นการพยากรณ์ที่อาศัยข้อมูลในอดีตมาพยากรณ์ การเปลี่ยนแปลงของข้อมูล เมื่อเวลาเปลี่ยนการเกิดอุบัติเหตุจะมีแนวโน้มสูงขึ้นหรือลดลง เพื่อให้ทางกรมทางบก โรงพยาบาล ได้นำไปช่วยในการจัดเตรียมบุคลากรทางการแพทย์ เครื่องมือแพทย์ และอื่นๆ ซึ่งการ

การวิเคราะห์อนุกรมเวลาโดยทั่วไปใช้การวิเคราะห์ถดถอย (Linear Regression) การคำนวณหาเส้นค่าเฉลี่ย (Moving Average) หรือ การวิเคราะห์การถดถอย (Regression) ในการพยากรณ์ข้อมูลอนุกรมเวลา

ปัจจุบันเทคโนโลยีของคอมพิวเตอร์เข้ามามีบทบาทในการใช้เป็นเครื่องมือที่ช่วยอำนวยความสะดวกในการคิดคำนวณ ซึ่งเทคนิคแบบอนุกรมเวลาเป็นเทคนิคการพยากรณ์โดยอาศัยข้อมูลในอดีตมาพยากรณ์อนาคต ซึ่งกล่าวได้ว่าการวิเคราะห์อนุกรมเวลาเป็นระเบียบทางสถิติที่สามารถแปลงประสบการณ์ในอดีตไปพยากรณ์เหตุการณ์ในอนาคต [6] โดยใช้เทคนิคในเหมืองข้อมูลที่นำมาใช้ในการวิเคราะห์ข้อมูลอนุกรมเวลา เช่น เทคนิค Linear regression (LR), Artificial Neural Network (ANN), Sequential Minimal Optimization for Regression (SMOreg), Support Vector Machine Regression (SVMR) และ Gussian Process (GP) ซึ่งนักวิจัยหลายท่านได้นำเอาเทคนิคมาใช้ในการรวิเคราะห์อนุกรมเวลา เช่น วราพร งามสุข[5] ได้ทำการศึกษาเพื่อเปรียบเทียบการพยากรณ์อนุกรมเวลาด้วยวิธีการถดถอย โดยใช้ข้อมูลสถิติจำนวนการเกิดอุบัติเหตุในประเทศไทยระหว่างเดือนมกราคม พ.ศ. 2545 ถึง เดือนธันวาคม พ.ศ. 2554 พบว่าค่าประสิทธิภาพที่ได้ RMSE เท่ากับ 51.59% ส่วนงานวิจัยของ Sven และ Rohit [6] ได้สร้างแบบจำลองเพื่อการพยากรณ์ข้อมูลอนุกรมเวลาในรูปแบบฤดูกาลจำนวน 8 ชุดข้อมูลด้วยโครงข่ายประสาทเทียม โดยทดลองกับโมเดลโครงข่ายประสาทเทียมแบบต่าง ๆ แสดงให้เห็นว่าโครงข่ายประสาทเทียมมีประสิทธิภาพในการแก้ปัญหาการพยากรณ์อนุกรมเวลาแบบมีฤดูกาลได้เป็นอย่างดี โดยจะต้องเลือกจำนวนชั้นของโครงข่ายประสาทเทียมให้เหมาะสมกับชุดข้อมูล ซึ่งจะเห็นว่ามีนักวิจัยจำนวนน้อยที่นำเหมืองข้อมูลมาสร้างแบบจำลองเพื่อการพยากรณ์ข้อมูลอนุกรมเวลาของการเกิดอุบัติเหตุ

ดังนั้นผู้วิจัยสนใจศึกษาการประยุกต์เทคนิคการพยากรณ์ที่เหมาะสมการวิเคราะห์อนุกรมเวลาในการพยากรณ์พื้นที่จังหวัดขอนแก่นเพื่อพยากรณ์การเกิดอุบัติเหตุบนท้องถนน โดยใช้ข้อมูลย้อนหลังตั้งแต่ พ.ศ. 2554 ถึง พ.ศ. 2560 เพื่อเป็นข้อมูลเบื้องต้นที่สามารถนำผลการพยากรณ์จากข้อมูลไปใช้ ในการวางแผนป้องกันและหลีกเลี่ยงการเกิดอุบัติเหตุในแต่ละเส้นทางของการสัญจร โดยได้ใช้หลักการ Sliding Window ในการแบ่งกลุ่มข้อมูลเป็นชุดข้อมูลเรียนรู้ และชุดข้อมูลทดสอบและวัดประสิทธิภาพการพยากรณ์ของแบบจำลอง เพื่อหาเทคนิคการพยากรณ์ที่มีความแม่นยำมากที่สุดในการพยากรณ์การเกิดอุบัติเหตุบนท้องถนน และยังพบว่าระบบที่สร้างขึ้นเพื่อแสดงผลการพยากรณ์นั้นสามารถใช้งานได้อย่างถูกต้อง

1.2 วัตถุประสงค์ของการวิจัย

1.2.1 เพื่อสร้างแบบจำลองในการพยากรณ์การเกิดอุบัติเหตุบนท้องถนนด้วยเทคนิคอนุกรมเวลาและเปรียบเทียบประสิทธิภาพแบบจำลอง

1.2.2 เพื่อแสดงผลของการพยากรณ์การเกิดอุบัติเหตุล่วงหน้าทางเว็บไซต์

1.3 ความสำคัญของการวิจัย

1.3.1 ได้เทคนิคที่สามารถพยากรณ์การเกิดอุบัติเหตุบนท้องถนนแบบอนุกรมเวลาอำเภอเมืองขอนแก่น

1.3.2 ได้ผลของการพยากรณ์ของการเกิดอุบัติเหตุล่วงหน้าในระยะเวลา 3 ปี ของอำเภอเมืองขอนแก่นทางเว็บไซต์

1.4 ขอบเขตของงานวิจัย

ข้อมูลการเกิดอุบัติเหตุ จากสถานีตำรวจจังหวัดขอนแก่น ระหว่างเดือนมกราคม พ.ศ. 2554 ถึง เดือนมิถุนายน พ.ศ. 2560 โดยมีตัวแปรที่ศึกษาประกอบด้วย เพศ ยานพาหนะ มูลค่าทรัพย์สินที่เสียหาย ความเสียหายที่เกิดขึ้นกับบุคคล เดือน และปี และสร้างแบบจำลองเพื่อพยากรณ์การเกิดอุบัติเหตุในอนาคต โดยพยากรณ์การเกิดอุบัติเหตุล่วงหน้า ในเขตอำเภอเมือง จังหวัดขอนแก่น

1.5 นิยามศัพท์เฉพาะ

1.5.1 เทคนิคอนุกรมเวลา (Time series forecasting) หมายถึง เทคนิคที่สามารถพยากรณ์การเกิดเหตุการณ์ในอนาคต

1.5.2 การพยากรณ์ หมายถึงการคาดการณ์การเกิดอุบัติเหตุในพื้นที่ อำเภอเมือง ขอนแก่น จากปี พ.ศ. 2561 ถึง พ.ศ. 2563

1.5.3 อุบัติเหตุบนท้องถนน หมายถึง อุบัติเหตุเกี่ยวกับยานยนต์ ซึ่งเป็นสาเหตุการตายและบาดเจ็บ ที่ผู้ประสบเหตุบนท้องถนน ด้วยจักรยานยนต์ หรือยานพาหนะต่างๆ รวมไปถึงคนเดินเท้าที่ถูกรถจักรยานยนต์หรือพาหนะอื่นๆ เฉี่ยวชน

พหูน ปณ ฑิโต ชีเว

บทที่ 2
เอกสารและงานวิจัยที่เกี่ยวข้อง

ในบทนี้ผู้วิจัยได้ศึกษาค้นคว้าและนำเสนอทฤษฎี และงานวิจัยที่เกี่ยวข้อง ในการพัฒนาแบบจำลองเพื่อพยากรณ์การเกิดอุบัติเหตุบนท้องถนนโดยใช้เทคนิคอนุกรมเวลา ยังมีการนำเสนอการวิเคราะห์ประสิทธิภาพในการพยากรณ์ในส่วนท้ายของบทนี้ และตามด้วยงานวิจัยที่เกี่ยวข้อง

2.1 ข้อมูลทั่วไปของจังหวัดขอนแก่น

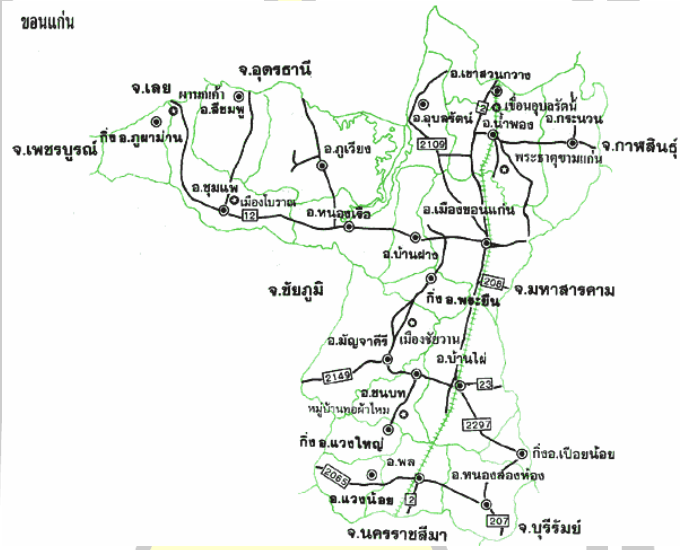
ภาพประกอบที่ 2.1 สัญลักษณ์ของจังหวัดขอนแก่น
ที่มา : ชมรมนักวิทยุสมัครเล่นแก่นทอง จังหวัดขอนแก่น [7]

จังหวัดขอนแก่นมีสภาพพื้นที่ลาดเอียงจากทิศตะวันตกไปทิศตะวันออกและทิศ ใต้ บริเวณที่สูงทางด้านตะวันตกมีสภาพพื้นที่เป็นเขาหินปูนตะปุ่มตะป่ำสลับกับ พื้นที่เป็นลูกคลื่นลอนลาดเล็กน้อย มีระดับความสูงประมาณ 200-250 เมตรจากระดับน้ำทะเล มีภูเขารูปแอ่งหรือภูเวียงวางตัวอยู่ติดอำเภอภูเวียง บริเวณที่สูงตอนกลางและด้านเหนือมีสภาพพื้นที่เป็นเทือกเขา ได้แก่ ภูเก้า ภูเม็ง ภูพานคำ เป็นแนวขวางมาจากด้านเหนือ แล้ววกลงมาทางตะวันตกเฉียงใต้ ไหล่เขาด้านนอกมีความสูงและลาดชันมาก สูงประมาณ 300-660 เมตร ไหล่เขาด้านในมีความลาดชันน้อย มีระดับความสูงประมาณ 220-250 เมตร

บริเวณแอ่งโคราช ครอบคลุมพื้นที่ทางด้านใต้จังหวัด สภาพพื้นที่เป็นลูกคลื่นลอนลาดเล็กน้อย มีความสูงประมาณ 150-200 เมตร มีบางส่วนเป็นเนิน สูงประมาณ 170-250 เมตร และลาดต่ำไปทางราบลุ่มที่ขนานกับลำน้ำชี มีความสูงประมาณ 130-150 เมตร จากนั้น พื้นที่จะลาดชันไปทางตะวันออก มีลักษณะเป็นลูกคลื่นลอนลาดมีความสูงประมาณ 200-250 เมตร และค่อนข้างราบ มีความสูงประมาณ 170 -180 เมตร [8]

อาณาเขตติดต่อ

ทิศเหนือ ติดต่อกับจังหวัดเลย จังหวัดหนองบัวลำภู และจังหวัดอุดรธานี
ทิศตะวันออก ติดต่อกับจังหวัดกาฬสินธุ์ และจังหวัดมหาสารคาม
ทิศใต้ ติดต่อกับจังหวัดนครราชสีมา และจังหวัดบุรีรัมย์
ทิศตะวันตก ติดต่อกับจังหวัดชัยภูมิ และจังหวัดเพชรบูรณ์ ดังภาพประกอบที่ 2.2



ภาพประกอบที่ 2.2 แผนที่จังหวัดขอนแก่น

ที่มา : มหาวิทยาลัยรามคำแหง สาขาวิทยบริการ จังหวัดขอนแก่น [9]

2.2 อุบัติเหตุบนท้องถนน

อุบัติเหตุบนท้องถนน หมายถึง ผู้ประสบเหตุบนท้องถนนด้วยจักรยานยนต์ หรือยานพาหนะอื่นๆ รวมไปถึงคนเดินเท้าที่ถูกรถจักรยานยนต์หรือพาหนะอื่นๆ เฉี่ยวชน ทั้งที่มีคู่กรณีและไม่มีคู่กรณี ซึ่งก่อให้เกิดการบาดเจ็บ หรือเสียชีวิต หรือทรัพย์สินเสียหาย [10]

อุบัติเหตุยานยนต์ในถนนหลวง [11] โดยเฉพาะอุบัติเหตุเกี่ยวกับยานยนต์ เป็นสาเหตุการตายและบาดเจ็บสูงสุด ของสถิติอุบัติเหตุ ที่รุนแรงทุกประเภท จากสถิติของกระทรวงสาธารณสุข รายงานไว้ว่า อัตราตายจากอุบัติเหตุยานยนต์ในถนนหลวงปี พ.ศ. 2515 ถึง พ.ศ. 2519 อยู่ในเกณฑ์ระหว่าง 10.0 ถึง 12.9 ต่อประชากรแสนคน เป็นสาเหตุการตายอันดับที่ 1 ของบรรดาอุบัติเหตุ

2.2.1 สาเหตุของการเกิดอุบัติเหตุบนท้องถนน

สาเหตุของการเกิดอุบัติเหตุบนท้องถนน ส่วนใหญ่เกิดขึ้นจากสาเหตุที่สำคัญ 3 ประการดังนี้ สาเหตุที่เกิดจากบุคคล ได้แก่ ผู้ใช้ยานพาหนะ ผู้โดยสาร และคนเดินเท้า

สาเหตุจากผู้ใช้ยานพาหนะ

1. มีความบกพร่องทางด้านร่างกาย เช่น ร่างกายอ่อนเพลีย ง่วงนอนหรือหลับใน มีโรคประจำตัว โรคลมชัก ตาบอดสี ตาพร่า น้ำตาลในเลือดต่ำ เป็นต้น

2. มีความบกพร่องทางด้านจิตใจและอารมณ์ เช่น มีความกลุ้มใจ วิตกกังวล อารมณ์หงุดหงิด มีความตึงเครียดทางอารมณ์ เป็นต้น

3. ขาดความรู้ ความชำนาญ และประสบการณ์ในการใช้รถบนท้องถนน เช่น ขาดความรู้เรื่องความเร็วในการขับ คาดคะเนความเร็วหรือกะระยะทางไม่ถูกต้อง ไม่มีความรู้ความชำนาญ ในเรื่องลักษณะของยานพาหนะที่ใช้ขับ ไม่รู้กฎจราจร เป็นต้น

4. ไม่ปฏิบัติตามกฎระเบียบหรือข้อบังคับ เช่น ขับรถเร็ว ขับรถตัดหน้าระยะกระชั้นชิด ขับรถล้ำช่องทางเดินรถ ขับรถแซงซ้ายหรือแซงขวาในที่คับขัน ขับรถตามระยะกระชั้นชิด ขับรถย้อนศรทางเดินรถ หยุดรถโดยกระชั้นชิด เป็นต้น

5. ไม่รู้จักป้องกันตนเอง เช่น ขับรถด้วยความประมาท ขาดความระมัดระวัง เสพยากระตุ้นประสาท ดื่มสุราขณะขับรถ เป็นต้น สำหรับเรื่องการดื่มสุรานั้นจากสถิติของสถาบันนิติเวชวิทยา กรมตำรวจ ปี พ.ศ. 2532 พบว่าผู้เสียชีวิตด้วยอุบัติเหตุจากการจราจรมีประวัติการดื่มสุราจำนวน 288 คน ซึ่งคิดเป็นร้อยละ 77.12

สาเหตุจากผู้โดยสารหรือคนเดินเท้า

1. การขาดความระมัดระวัง เช่น ผู้โดยสารขึ้นหรือลงรถโดยไม่ระมัดระวัง ในการปิดหรือเปิดประตูรถ วิ่งตัดหน้ารถ ลังเลใจในการข้ามถนน เป็นต้น

2. การไม่ปฏิบัติตามกฎจราจร เช่น ห้อยโหนรถโดยสารรถประจำทาง ไม่ขึ้นหรือลงขณะรถหยุด หรือที่ป้ายจอด ไม่ข้ามถนนตรงทางม้าลายหรือสะพานลอย ไม่เดินถนนตามบาทวิถีหรือทางเท้า เป็นต้น

3. ความรู้เท่าไม่ถึงการณ์ เช่น ข้ามถนนโดยออกจากหน้าหรือท้ายรถขณะที่รถยังจอดอยู่ วิ่งตัดหน้ารถ เป็นต้น

4. ความไม่สมบูรณ์ของร่างกายและจิตใจ เช่น สภาพร่างกายที่อ่อนเพลีย การดื่มสุราขณะขับรถ เป็นต้น

สาเหตุจากสิ่งแวดล้อม ได้แก่ สภาพของรถ สภาพถนน สภาพดินฟ้าอากาศ

สาเหตุจากสภาพของรถ

1. ยางระเบิดหรือยางแตก ทำให้รถเสียการทรงตัว พลิกคว่ำได้ง่าย โดยเฉพาะรถที่กำลังแล่นด้วยความเร็วสูง และถนนลื่น

2. เบรกแตก เบรกลื่น ทำให้รถไม่สามารถหยุดหรือชะลอความเร็วลงได้ตามความต้องการ

3. เพลาหลุดหรือเพลาขาด ทำให้รถหมดกำลังในการขับเคลื่อน รถจะไม่แล่น แม้ว่าจะเหยียบคันเร่งอย่างไรก็ตาม ทำให้ยากแก่การควบคุมความเร็ว และง่ายต่อการเกิดอุบัติเหตุ

4. คันส่งหลุดทำให้พวงมาลัยใช้การไม่ได้ ไม่สามารถควบคุมรถได้

5. อุปกรณ์ประจำรถชำรุดหรือขัดข้อง เช่น ไม่มีไฟหน้าหรือหลัง ไฟใหญ่มีข้างเดียวหรือไม่มีเลย ไฟเลี้ยวชำรุดไม่ได้ซ่อมแซมหรือแก้ไข พวงมาลัยสั่นขณะขับ เป็นต้น

6. การเปลี่ยนแปลงสภาพรถ เช่น การเพิ่มแรงเครื่องทำให้ผู้ขับขี่เกิดความคะนองและขับรถเร็ว การแปลงสภาพรถตามความพอใจ โดยไม่คำนึงถึงสภาพรถที่ได้รับการออกแบบมา

สาเหตุจากสภาพถนน และสภาพแสงสว่าง

1. บริเวณที่เกิดอุบัติเหตุ ได้แก่ ทางแยก ทางโค้ง ทางตรง ทางเบี่ยง สะพาน ทางลาดชันหรือเนินเขา ทางเชื่อมโยงทางแยก เป็นต้น ซึ่งบริเวณที่มักเกิดเหตุบ่อยที่สุด คือ ทางตรง โดยสภาพเส้นที่ดีเรียบ มักทำให้ผู้ขับขี่ขาดความระมัดระวัง และขับรถด้วยความเร็วสูง นอกจากนี้จะพบว่าถนน 3 ช่องทางจะเกิดอุบัติเหตุมากกว่าถนน 2 ช่องทาง และถนน 4 ช่องทาง และถนนสี่แยกจะอันตรายกว่าสามแยก

2. สภาพถนนที่เป็นหลุมเป็นบ่อ มีโคลนตม มีเครื่องกีดขวางมากๆ หรือถนนที่แคบ ถนนที่ลื่น มีส่วนทำให้เกิดอุบัติเหตุขึ้นได้

3. สภาพแสงสว่างบนถนน เช่น แสงสว่างที่ส่องจากรถคันที่สวนมา โดยการเปิดไฟสูงและมีความสว่างสูง ทำให้ตามัวมองไม่ชัดเจน บนท้องถนนมืดไม่มีไฟฟ้า ไม่มีแสงสว่าง ทำให้มองไม่เห็นทางหรือมองไกลไม่ได้ ย่อมเป็นอันตรายต่อการขับรถ อย่างไรก็ตามแสงสว่างในเวลากลางวัน หรือความสว่างของถนนก็มักทำให้เกิดอุบัติเหตุสูงกว่าเวลากลางคืน แต่ความรุนแรงจะเกิดในเวลากลางคืนมากกว่า

สาเหตุจากดินฟ้าอากาศ

1. ฝนตกหนัก น้ำท่วม ทำให้ถนนเป็นหลุมเป็นบ่อ เป็นหล่มโคลน ถนนลื่น ทำให้รถตกถนน พลิกคว่ำการเกิดพายุหรือหมอกลงจัด ทำให้มีควันปกคลุมมองไม่เห็นทาง เกิดอุบัติเหตุได้ง่าย พายุหิมะในต่างประเทศอาจมีพายุหิมะ ทำให้ถนนลื่นมองไม่เห็นทาง สภาพดินฟ้าอากาศที่ดีอุบัติเหตุมักเกิดจากสภาพดินฟ้าอากาศที่ดีเสมอ ทั้งนี้เพราะผู้ขับขี่ขับรถด้วยความเร็วสูง และขาดความระมัดระวังอันตราย

สาเหตุจากกฎหมาย มีส่วนเกี่ยวข้องกับการเกิดอุบัติเหตุ ดังนี้

1. การขาดการเผยแพร่ประชาสัมพันธ์ให้ประชาชนทุกคนทราบกฎ ระเบียบข้อบังคับ และบทลงโทษ ในการฝ่าฝืนกฎต่างๆ ทำให้ประชาชนขาดจิตสำนัก และฝ่าฝืนกฎระเบียบต่างๆ ซึ่งมีผลให้เกิดอุบัติเหตุได้

2. บทลงโทษหรือค่าปรับยังไม่เหมาะสม ทำให้มีการฝ่าฝืนกฎจราจรอยู่เสมอ

3. การที่กฎหมายมิได้กำหนดเพศ อายุสูงสุดของผู้ขับขี่ รวมทั้งการศึกษาขั้นต่ำของผู้ขับขี่ยานพาหนะ ถึงแม้ว่าผู้ขับขี่จะสอบผ่าน และได้รับใบอนุญาตขับขี่มาแล้ว ก็อาจทำผิดกฎจราจรและทำให้เกิดอุบัติเหตุได้

4. การขาดการกวดขัน จับกุม หรือยังไม่จริงจังหรือเข้มงวดในการพิจารณาดำเนินคดีหรือจับกุม ผู้กระทำผิด เป็นสาเหตุให้ขับรถหรือใช้รถใช้ถนนอย่างเสรี ตามอำเภอใจ ซึ่งมักทำให้เกิดอุบัติเหตุ

2.2.2 การป้องกันอุบัติเหตุในการจราจรทางบก [12]

อุบัติเหตุในการจราจรทางบก สามารถป้องกันได้ดังต่อไปนี้

1. การป้องกันด้านบุคคล การป้องกันอุบัติเหตุในการจราจรทางบกด้านบุคคลนั้น ควรพิจารณาในเรื่องสุขภาพ การศึกษา และความปลอดภัยในการขับขี่ การโดยสาร และการเดินเท้า ซึ่งมีวิธีการป้องกันดังต่อไปนี้เรื่องสุขภาพ ผู้ขับขี่รถ ผู้โดยสารและผู้เดินเท้า ควรมีสภาพร่างกายและจิตใจที่แข็งแรงสมบูรณ์ และเป็นปกติอยู่เสมอ ทั้งในช่วงก่อนเดินทาง ขณะเดินทาง และหลังการเดินทาง สำหรับผู้ขับขี่รถ จะต้องรักษาสุขภาพให้แข็งแรงอยู่เสมอ ส่วนสภาพทางจิตใจนั้น ผู้ขับขี่ควรคุมอารมณ์ และจิตใจไว้ได้ มีสติสัมปชัญญะเสมอ ในการขับรถ หากมีอาการอ่อนเพลีย ง่วงนอน หรือมีความวิตกกังวลใจ ตื่นเต้น กระวนกระวายใจ มีอารมณ์เสียเกิดขึ้นบ่อยๆ มีความเครียด มีโรคทางจิตทางประสาท ก็ไม่ควรขับรถ เพราะจะขาดสมาธิในการขับรถ และอาจเกิดอุบัติเหตุขึ้นได้ง่าย

สำหรับผู้โดยสารยานพาหนะ และผู้เดินเท้า ก็ควรรักษาสุขภาพให้ดีทั้งทางกายและจิตใจ เพราะหากมีความผิดปกติของร่างกาย และจิตใจ ก็จะทำให้ประสบอุบัติเหตุได้ง่ายเช่นเดียวกัน

2. การศึกษา การป้องกันอุบัติเหตุที่สำคัญประการหนึ่งคือ การศึกษาหาความรู้และการถ่ายทอด หรือให้ความรู้แก่ทุกคน ในเรื่องความปลอดภัยในการจราจร การแนะนำประชาชนใช้รถใช้ถนน ให้รู้จักระมัดระวังในการเดินทาง ขณะสภาพดินฟ้าอากาศผิดปกติสำหรับผู้ขับขี่รถจะต้องเรียนรู้ เรื่องเกี่ยวกับตัวรถ สภาพการใช้งาน เรียนรู้วิธีการขับขี่ เส้นทางการเดินทาง เรียนรู้มารยาท และกฎการจราจรด้วย

3. ความปลอดภัยในการขับขี่ยวดยานพาหนะ การขับขี่ยวดยานพาหนะไม่ว่าจะเป็นรถยนต์ รถจักรยาน รถจักรยานยนต์ รถบรรทุก หรือรถประจำทางก็ตามผู้ขับขี่จะต้องปฏิบัติตนเพื่อความปลอดภัยดังนี้

1) ต้องได้รับใบอนุญาตขับรถก่อนการใช้รถ(ผู้ขับขี่รถยนต์ส่วนบุคคลและรถจักรยานยนต์ ต้องมีอายุ ไม่ต่ำกว่า 18 ปีบริบูรณ์ ผู้ขับขี่รถยนต์สาธารณะต้องมีอายุ 25 ปีบริบูรณ์ และผู้ขับขี่รถจักรยานต้องมีอายุ 13 ปีบริบูรณ์) ซึ่งผ่านการทดสอบจากเจ้าหน้าที่ขนส่ง

2) การตรวจสภาพของรถทุกครั้งก่อนที่จะนำออกไปใช้ควรตรวจสอบให้เรียบร้อย

3) ควรวางแผนการขับรถ วางแผนการใช้เส้นทาง วางแผนขับรถ ตรวจสอบหรือกำหนดล่วงหน้าถึงจุดจอดรถ จุดจอดพัก จุดเติมน้ำมัน หากต้องเดินทางระยะไกล หรือบริเวณทางด่วน

4) แต่งกายให้รัดกุม และใช้สีที่มองเห็นได้ชัดเจน โดยเฉพาะผู้ที่ขับขี่รถจักรยาน และรถจักรยานยนต์

5) สวมใส่เครื่องป้องกันอันตราย เช่น สวมหมวกนิรภัยทุกครั้งที่ขับขี่ หรือโดยสารรถจักรยานยนต์ ต้องรัดเข็มขัดนิรภัยทุกครั้งที่ขับขี่ หรือโดยสาร

6) ขับรถตามกฎจราจรอย่างเคร่งครัด

7) ขับรถความระมัดระวังเสมอ

8) ควรขับรถอย่างมีมารยาท มีน้ำใจ สุภาพ สุขุม และรู้จักให้อภัยเมื่อ มีการผิดพลาดเกิดขึ้น

9) ผู้ที่ขับขี่ยวดยานพาหนะไม่ควรใช้ยาเสพติดต่างๆ รวมทั้งไม่ควรดื่มสุราหรือของมึนเมาต่างๆ เมื่อจะขับรถ

10) ทุกครั้งที่ขับขี่ยานพาหนะ ผู้ขับขี่ต้องมีสติมั่นคง ไม่ตกใจง่าย สามารถควบคุมสติได้ดี ซึ่งจะทำให้ตัดสินใจ และเลือกใช้วิธีแก้สถานการณ์เฉพาะหน้าได้ถูกต้อง

4. ความปลอดภัยในการโดยสารยานพาหนะ เพื่อความปลอดภัย ผู้โดยสารยานพาหนะทางบกควรระมัดระวังในการเดินทาง

ความปลอดภัยในการเดินทางเท้าควรปฏิบัติตนเพื่อความปลอดภัยดังนี้

1) การเดินถนน ถนนที่มีทางเท้าให้เดินบนทางเข้า โดยเดินชิดด้านซ้ายมือ ไม่เดินใกล้ทางรถโดยหันหลังให้รถที่กำลังแล่นมา และก่อนที่จะก้าวเดินไปทางรถต้องมองซ้าย มองขวาทั้งสองด้านก่อนเสมอ

2) การข้ามถนน เพื่อความปลอดภัยในการข้ามถนน ตรงช่องทางข้าม (ทางม้าลาย สะพานลอย หรืออุโมงค์) การข้ามถนนตรงช่องทาง ที่ควบคุมด้วยสัญญาณไฟจราจร และการข้ามถนน โดยไม่มีสัญญาณอะไรเลย

2.2.3 ระดับความรุนแรงของอุบัติเหตุบนท้องถนน

ระดับความรุนแรงของอุบัติเหตุบนท้องถนน สามารถแบ่งได้เป็น 3 ระดับ ซึ่งประกอบด้วย เสียชีวิต บาดเจ็บสาหัส และบาดเจ็บเล็กน้อย ดังรายละเอียดต่อไปนี้

เสียชีวิต หมายถึง ผู้ป่วยที่ได้รับอุบัติเหตุเข้ารักษาที่โรงพยาบาล และเสียชีวิตภายใน 30 วันหลังจากเข้ารับการรักษา หรือเสียชีวิตในที่เกิดเหตุ

บาดเจ็บสาหัส หมายถึง ผู้ป่วยที่ได้รับอุบัติเหตุและแพทย์ให้นอนพักรักษาในโรงพยาบาล เช่น การเสียเลือดมาก หรือกระดูกหักมากกว่าหนึ่งแห่ง หรือสมองได้รับการกระทบกระเทือน หรือได้รับบาดเจ็บเป็นแผลฉกรรจ์ หรืออวัยวะของร่างกายฉีกขาดถึงขั้นพิการ

บาดเจ็บเล็กน้อย หมายถึง ผู้ป่วยที่ได้รับอุบัติเหตุ และได้รับการรักษาที่ห้องฉุกเฉินแล้วแพทย์อนุญาตให้กลับบ้าน เช่น ศีรษะแตกแต่สมองไม่ได้รับการกระทบกระเทือน หรือได้รับบาดเจ็บเป็นแผลเพียงเล็กน้อยเมื่อได้รับการรักษาก็สามารถกลับบ้านได้ หรือมีการร้องขอความช่วยเหลือด้วยความตกใจหรือการเสียสติไปชั่วขณะ

จากสภาพถนนในจังหวัดขอนแก่น ซึ่งเป็นจังหวัดที่มีการเกิดอุบัติเหตุบนท้องถนนอันดับต้นของภาคตะวันออกเฉียงเหนือ เนื่องจากจังหวัดขอนแก่นมีขนาดพื้นที่ค่อนข้างใหญ่และการขนส่งสินค้าเส้นทางนี้ เพิ่มมากขึ้นตามลำดับ จากเหตุผลดังกล่าวผู้วิจัยได้เลือกจังหวัดขอนแก่นเป็นพื้นที่สร้างแบบจำลองการพยากรณ์การเกิดอุบัติเหตุบนท้องถนน เพราะเป็นพื้นที่ที่ครอบคลุมและมีความเสี่ยงต่อการเกิดอุบัติเหตุบนท้องถนน ซึ่งเป็นภาพรวมของประเทศ

2.3 เหมืองข้อมูล

คำว่า “เหมืองข้อมูล” มีนักวิจัยได้ให้นิยามไว้หลายความหมายดังนี้ ขจรศักดิ์ ศรีอ่อน [13] ได้ให้ความหมายเหมืองข้อมูล ไว้ว่า เป็นกระบวนการของการกลั่นกรองสารสนเทศ (Information) ที่ซ่อนอยู่ในฐานข้อมูลใหญ่ เพื่อทำนายแนวโน้มและพฤติกรรม โดยอาศัยข้อมูลในอดีต โดยค้นหาแบบรูปความสัมพันธ์ และความรู้ใหม่อื่นๆ งานวิจัยถัดมาของวาทินี นุ้ยเพียร [14] ได้นิยามศัพท์ของเหมืองข้อมูล (Data Mining) ไว้ว่า เป็นกระบวนการกลั่นกรองข้อมูลจากฐานข้อมูลขนาดใหญ่ โดยมองที่ความสัมพันธ์ของข้อมูลแนวโน้มของข้อมูลต่างๆ เพื่อให้สามารถนำข้อมูลที่กลั่นกรองได้มาใช้ประโยชน์ และเป็นข้อมูลสนับสนุนในการตัดสินใจในเรื่องต่างๆ ในขณะที่สุมาลัย นุชิต [15] ได้ให้ความหมายเหมืองข้อมูล ไว้ว่า เป็นการค้นหาความสัมพันธ์และรูปแบบทั้งหมด ซึ่งมีอยู่จริงในฐานข้อมูล แต่ได้ถูกซ่อนไว้ภายในข้อมูลจำนวนมาก การทำเหมืองข้อมูลจะทำการสำรวจและวิเคราะห์อย่างอัตโนมัติหรือกึ่งอัตโนมัติ ในปริมาณข้อมูลจำนวนมากให้อยู่ในรูปแบบที่เต็มไปด้วย

ความหมาย และอยู่ในรูปของกฎโดยความสัมพันธ์เหล่านี้แสดงให้เห็นถึงความรู้ต่างๆ ที่มีประโยชน์ในฐานข้อมูล ส่วนงานวิจัยของ Dong-xiao [16] ได้ให้ความหมายเหมืองข้อมูล ไว้ว่า เป็นกระบวนการค้นหารูปแบบที่มีความสัมพันธ์ โดยอาศัยการรู้จำแบบ วิธีการทางสถิติ ทางคณิตศาสตร์ และเทคโนโลยี เพื่อให้รูปแบบของข้อมูลมีประสิทธิภาพ และงานวิจัยของ Poel [17] ได้ให้ความหมายเหมืองข้อมูล ไว้ว่า เป็นวิธีการสกัดความรู้ที่ซ่อนอยู่ในฐานข้อมูลขนาดใหญ่หรือวิเคราะห์ข้อมูลที่ได้ถูกออกแบบมาเพื่อใช้ในกระบวนการตัดสินใจ

ดังนั้นจึงสรุปได้ว่า เหมืองข้อมูล คือ กระบวนการค้นหาวิธีการสร้างแบบจำลอง และความสัมพันธ์ที่ซ่อนอยู่ในชุดข้อมูลจำนวนมหาศาลโดยอัตโนมัติ ซึ่งใช้ขั้นตอนวิธีการทางสถิติ การเรียนรู้ของเครื่อง และการรู้จำแบบ ปัจจุบันการทำเหมืองข้อมูลได้เข้ามามีบทบาทในการวิเคราะห์ข้อมูลการเกิดโรคต่างๆมากมาย โดยทำการรวบรวมข้อมูลการเกิดโรคจากหน่วยงานต่างๆ ทั้งภาครัฐและภาคเอกชน และนำข้อมูลทั้งหมดมาวิเคราะห์เพื่อให้ได้แบบจำลองสำหรับการพยากรณ์การเกิดโรค ซึ่งการทำเหมืองข้อมูลในการสร้างแบบจำลอง สามารถแบ่งได้เป็น 2 ประเภทใหญ่ๆ ประกอบด้วย

1. การสร้างแบบจำลองเพื่อการทำนาย (Predictive Data Mining) หรือเรียกว่า (Supervised Learning) คือ การนำข้อมูลในอดีตมาสร้างแบบจำลองต้นแบบเพื่อทำการทำนายอนาคต โดยมีการใช้ข้อมูลสำหรับสอน (Training Data) ซึ่งทุกข้อมูลจะมีคุณสมบัติ เป็นค่าที่ใช้ในการทำนายผลของข้อมูล อัลกอริทึมประเภทนี้จะมุ่งเน้นในการแบ่งแยกข้อมูลออกเป็นกลุ่มตามค่าคุณสมบัติของข้อมูล ซึ่งถ้าค่าคุณสมบัติของข้อมูลมีค่าไม่ต่อเนื่อง จะเรียกกระบวนการที่ใช้แบ่งแยกว่าการจำแนกประเภท (Classification) ถ้าค่าคุณสมบัติของข้อมูลมีค่าต่อเนื่อง จะเรียกกระบวนการที่ใช้ว่าการถดถอย (Regression) หรือการทำนาย (Predictive)

2. การสร้างแบบจำลองเพื่อการบรรยาย (Descriptive Data Mining) หรือเรียกว่า (Unsupervised learning) คือ การนำข้อมูลที่มีอยู่มาดูและศึกษาเพื่อหากฎความสัมพันธ์ต่างๆ (Association Rules) หรือหาการจัดกลุ่มข้อมูล (Clustering Analysis) ซึ่งไม่ได้มีจุดมุ่งหมายเพื่อการทำนาย

ในการทำเหมืองข้อมูลมี 3 ขั้นตอนดังต่อไปนี้ [18]

1. ขั้นตอนก่อนการทำเหมืองข้อมูล

1.1 ศึกษาและวิเคราะห์ข้อมูล เป็นการรวบรวมข้อมูลต่างๆ ที่เกี่ยวข้อง เพื่อใช้ในการวิเคราะห์ด้วยเทคนิคการทำเหมืองข้อมูล ต้องทำความเข้าใจถึงแหล่งที่มาและการจัดเก็บของข้อมูล โดยต้องเป็นข้อมูลที่ได้มาจากแหล่งข้อมูลที่ถูกต้องน่าเชื่อถือ และเป็นข้อมูลที่เหมาะสมมีรายละเอียดเพียงพอต่อการนำไปใช้ในการวิเคราะห์

1.2 การกลั่นกรองข้อมูล (Data Cleaning) เนื่องจากข้อมูลการเกิดอุบัติเหตุที่ได้จากโปรแกรม Excel และได้นำข้อมูลเข้าสู่โปรแกรม ผู้วิจัยอาจกรอกข้อมูลซ้ำกัน หรือมีข้อมูลบางส่วนขาดหายไป ดังนั้นจึงต้องมีการลบข้อมูลซ้ำกันออก และจัดเรียงข้อมูลให้ถูกต้องให้เหมาะสมกับงานวิจัยนี้

1.3 การแปลงข้อมูล (Data Transformation) โดยนำจำนวนครั้งของการเกิดอุบัติเหตุโดยแยกตามเพศ ยานพาหนะ มูลค่าทรัพย์สิน ความเสียหายที่เกิดขึ้นกับบุคคล เดือน และปี เป็นการแทนค่าปัจจัย เพื่อสร้างแบบจำลองให้มีประสิทธิภาพในการจำแนกข้อมูล จึงทำการแปลงข้อมูล

2. การสร้างแบบจำลอง เป็นกระบวนการค้นหารูปแบบของข้อมูลเพื่อใช้ในการพยากรณ์โดยใช้เทคนิคต่างๆ ที่อยู่ในการทำเหมืองข้อมูล และ Machine Learning

3. การวัดประสิทธิภาพ

2.4 อนุกรมเวลา

คำว่า “อนุกรมเวลา” มีนักวิจัยได้ให้นิยามไว้หลายความหมายดังนี้ รศ.ดร.นพ.พงษ์เทพ วิวรรณธนะเดชและคณะ [7] ได้ให้ความหมายของคำว่า อนุกรมเวลา ไว้ว่า เซตของข้อมูลเชิงปริมาณที่จัดเก็บในช่วงเวลาหนึ่ง ข้อมูลอนุกรมเวลาอาจะอยู่ในลักษณะที่เป็นรายปี รายไตรมาส หรือรายเดือนก็ได้ ทั้งนี้ขึ้นอยู่กับความเหมาะสมในการนำไปใช้ประโยชน์ งานวิจัยถัดมา ทรงศิริ แต้สมบัติ [8] ได้ให้ความหมาย อนุกรมเวลา ไว้ว่า การพยากรณที่อยูภายใตขอสมมุติฐานวารูปแบบของขอมูลในอดีตยังคงเกิดขึ้นตอไปในอนาคต หรืออาจกลาวไดวาลักษณะของการเปลี่ยนแปลงรูปแบบของขอมูลหรือตัวแปรที่เราสนใจศึกษา เชน ปริมาณความตองการ หรืออุปสงคในชวงเวลาที่ผานมา ซึ่งจะเรียกวาตัวแปรตาม (Dependent variable) จะทำใหสามารถคาดการณหรือทำนายไดวาในอนาคตลักษณะของขอมูลก็ควรจะอยูในรูปแบบเชนนั้นตอไป และ Wei,William W.S ได้ให้ความหมายอนุกรมเวลาไว้ว่า เซตของข้อมูลเชิงปริมาณที่จัดเก็บในช่วงเวลาหนึ่ง ตัวอย่างเช่น ดัชนีตลาดหลักทรัพย์ในแต่ละวันเมื่อปิดทำการซื้อขายในแต่ละวัน รายได้ประชาชาติ (GNP) รายไตรมาส รายรับในแต่ละปีของบริษัทแห่งหนึ่ง เป็นต้น

ดังนั้นจึงสรุปได้ว่า อนุกรมเวลา หมายถึง ข้อมูลหรือค่าสังเกตที่เปลี่ยนแปลงไปตามลำดับเวลาที่เกิดขึ้นอย่างต่อเนื่อง ข้อมูลเหล่านี้ถูกเก็บรวบรวม ณ ช่วงเวลาต่างๆ เช่น รายวัน รายสัปดาห์ รายเดือน รายไตรมาส หรือรายปี

การวิเคราะห์อนุกรมเวลา (Time Series Analysis) เป็นการพยากรณ์ที่อาศัยข้อมูลในอดีตมาพิจารณา ว่าลักษณะการเปลี่ยนแปลงของข้อมูล เมื่อเวลาเปลี่ยนไปมีลักษณะอย่างไร มีการเคลื่อนไหวมากน้อยเพียงใด โดยมีข้อสมมุติว่าการเคลื่อนไหวของข้อมูลในอนาคตจะไม่แตกต่างกับอดีต

ข้อมูลอนุกรมเวลา (Time Series data) [19] คือ ชุดของข้อมูลที่เก็บรวบรวมตามระยะเวลาเป็นช่วง ๆ อย่างต่อเนื่องกัน เช่น ข้อมูลยอดขายสินค้าที่เก็บรวบรวมต่อเนื่องกันไปเป็นระยะเวลาหลาย ๆ เดือน ข้อมูลรายได้ประชาชาติปีต่าง ๆ ที่เก็บรวบรวมต่อเนื่องกันไปเป็นระยะเวลาหลาย ๆ ปี เป็นต้น ข้อมูลอนุกรมเวลาอาจอยู่ในลักษณะที่เป็นข้อมูลรายปี รายไตรมาส หรือรายเดือนก็ได้ ทั้งนี้ขึ้นอยู่กับความเหมาะสมในการนำไปใช้ประโยชน์ เนื่องจากข้อมูลทางธุรกิจมีการเปลี่ยนแปลงตลอดเวลา ผู้นำทางธุรกิจหรือองค์กรต้องหาวิธีพัฒนาต่าง ๆ ที่สามารถนำไปใช้ประกอบการตัดสินใจวางแผน เกี่ยวกับผลที่เกิดจากความเปลี่ยนแปลงในการดำเนินการอย่างใดอย่างหนึ่ง ดังนั้นการวิเคราะห์อนุกรมเวลาจึงเข้ามามีบทบาทช่วยในการตัดสินใจ เทคนิคอย่างหนึ่งที่ใช้ช่วยในการควบคุมการดำเนินการในปัจจุบันและในการวางแผนความต้องการในอนาคต คือ การพยากรณ์ (forecasting) ซึ่งการพยากรณ์นั้นทำได้หลายวิธี แต่ละวิธีต่างมีเป้าหมายเดียวกัน คือ ทำนายเหตุการณ์ในอนาคต

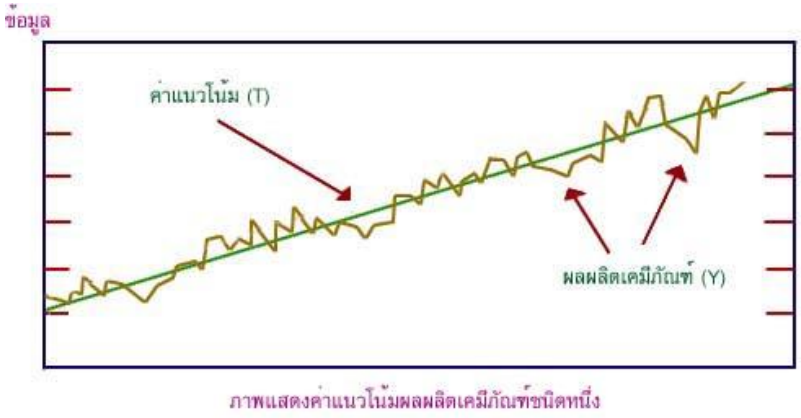
2.4.1 ลักษณะข้อมูลที่จะนำมาวิเคราะห์อนุกรมเวลา

1. เป็นข้อมูลที่รวบรวมมานานพอสมควร
2. เป็นข้อมูลที่มีความสัมพันธ์กับเวลา
3. ระยะเวลาที่ใช้ในการเก็บรวบรวมข้อมูลจะต้องเท่ากัน
4. หากข้อมูลที่เก็บรวบรวมมาได้ ถูกกระทบกระเทือนเนื่องจากช่วงเวลาที่ไม่เท่ากันจะต้องปรับแก้ข้อมูลก่อนนำมาวิเคราะห์อนุกรมเวลา

2.4.2 ส่วนประกอบของอนุกรมเวลา

ในการวิเคราะห์อนุกรมเวลา ผู้วิเคราะห์จะแยกองค์ประกอบต่าง ๆ ที่ประกอบกันขึ้นเป็นอนุกรมเวลา โดยจะมีการเปลี่ยนแปลงไปตามอิทธิพลต่าง ๆ เช่น การเปลี่ยนแปลงการผลิต เทคโนโลยี สภาวะอากาศ เป็นต้น ในการหาคุณลักษณะของอนุกรมเวลาเราสามารถใช้แบบจำลองได้หลายแบบ แบบจำลองที่ใช้โดยนักเศรษฐศาสตร์แบบหนึ่ง คือ แบบจำลองแบบคลาสสิก (classical model) เป็นการอธิบายถึงองค์ประกอบของการแปรผันของอนุกรมเวลา 4 ส่วน ดังนี้

1. ค่าแนวโน้ม (Secular trend) แทนด้วย Tt เป็นการเปลี่ยนแปลงข้อมูลมีลักษณะราบเรียบ แนวโน้ม อาจมีลักษณะเป็นเส้นตรงหรือเส้นโค้งในทางเพิ่มขึ้นหรือลดลง ค่าแนวโน้มของข้อมูลเป็นการเคลื่อนไหวในช่วงระยะเวลาที่ค่อนข้างนานพอสมควร ควรเป็นข้อมูลรายปีและควรมีข้อมูลอย่างน้อย 15 ปี ซึ่งจะแสดงทิศทางของอนุกรมเวลา ตัวอย่าง กราฟแสดงค่าแนวโน้มของผลผลิตเคมีภัณฑ์ชนิดหนึ่ง ดังภาพประกอบที่ 2.3

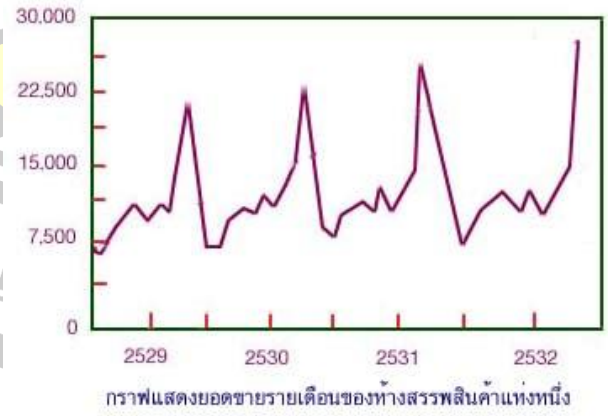


ภาพประกอบที่ 2.3 กราฟแสดงค่าแนวโน้มของผลผลิตเคมีภัณฑ์ชนิดหนึ่ง

ที่มา : บริษัท 168 เอคดูเคชั่น จำกัด [19]

จากกราฟ Y แทนข้อมูลอนุกรมเวลาของผลผลิตเคมีภัณฑ์ชนิดหนึ่ง ในช่วงเวลา 15 ปี ค่าแนวโน้มแทนด้วยกราฟเส้นตรง ซึ่งซ้อนอยู่บนเส้นกราฟของ Y เส้นกราฟแสดงแนวโน้มนอกจากจะมีลักษณะเป็นเส้นตรงดังรูปแล้ว อาจมีลักษณะเป็นเส้นโค้ง เช่น เส้นโค้งเอ็กโพเนนเชียลหรือพาราโบลา

2. การเปลี่ยนแปลงหรือความแปรผันตามฤดูกาล (Seasonal Variation) แทนด้วย St เป็นการเปลี่ยนแปลงข้อมูลมีลักษณะการเพิ่มขึ้น หรือลดลงในลักษณะเดียวกันของรอบระยะเวลาหนึ่งที่แน่นอน เรียกว่า การเปลี่ยนแปลงตามฤดูกาล หน่วยของระยะเวลาสำหรับข้อมูลอาจเป็นรายชั่วโมง รายวัน รายสัปดาห์ รายเดือน รายไตรมาส สำหรับข้อมูลรายปีไม่มีการแปรผันตามฤดูกาล การเปลี่ยนแปลงตามฤดูกาลนั้นกำหนดระยะเวลาการเกิดซ้ำในรอบหนึ่ง ๆ ได้ค่อนข้างแน่นอน ตัวอย่างเช่น ยอดขายรายเดือนของห้างสรรพสินค้าแห่งหนึ่ง ดังภาพประกอบที่ 2.4

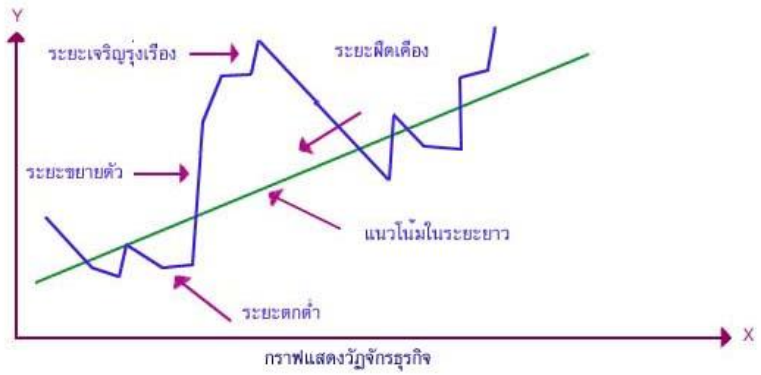


ภาพประกอบที่ 2.4 กราฟแสดงยอดขายรายเดือนของห้างสรรพสินค้าแห่งหนึ่ง

ที่มา : บริษัท 168 เอคดูเคชั่น จำกัด [19]

จากกราฟ จะเห็นว่ายอดขายของห้างสูงประมาณเดือนธันวาคมของทุกปี ซึ่งเป็นเทศกาลคริสต์มาสและปีใหม่ ประชาชนจึงมีการจับจ่ายใช่สอยมาก ส่วนในราวเดือนพฤษภาคมของทุกปี ยอดขายจะต่ำกว่าในเดือนอื่น ๆ ที่เป็นเช่นนี้เพราะเป็นช่วงเปิดภาคเรียน ประชาชนต้องเตรียมเงินไว้สำหรับค่าใช้จ่ายในการศึกษาของบุตรหลาน

3. การเปลี่ยนแปลงหรือความแปรผันตามวัฏจักร (Cyclical Variation) แทนด้วย Ct การเปลี่ยนแปลงตามวัฏจักร มีการเปลี่ยนแปลงเคลื่อนไหวในลักษณะซ้ำ ๆ กันและจะมีลักษณะคล้ายคลึงกับการเปลี่ยนแปลงตามฤดูกาล จะต่างกันก็ตรงที่การเปลี่ยนแปลงตามวัฏจักรแต่ละรอบจะใช้ระยะเวลาที่นานกว่า คือ ตั้งแต่ 5 ปีขึ้นไป ข้อมูลที่มีการเปลี่ยนแปลงตามวัฏจักรในทางธุรกิจเรียกว่า "วัฏจักรธุรกิจ" (BusinessCyclical) โดยทั่วไปประกอบด้วย ระยะเจริญรุ่งเรือง (prosperity) ระยะฝืดเคือง (recession) ระยะตกต่ำ (depression) และระยะขยายตัว (recovery) ดังภาพประกอบที่ 2.5



ภาพประกอบที่ 2.5 กราฟการเกิดระยะต่าง ๆที่เกิดอย่างต่อเนื่องเป็นวัฏจักร
ที่มา : บริษัท 168 เอคดูเคชั่น จำกัด [19]

4. การเปลี่ยนแปลงหรือความแปรผัน เนื่องจากเหตุการณ์ผิดปกติ (Irregularl Variation) แทนด้วย I_t เป็นการเปลี่ยนแปลงของข้อมูลอนุกรมเวลาที่เกิดจากเหตุการณ์ที่เราไม่สามารถคาดการณ์ได้ล่วงหน้า เช่น การเกิดไฟไหม้ในโรงงาน การเกิดอุทกภัย การนัดหยุดงานของคนงาน แผ่นดินไหว เป็นต้น ซึ่งเหตุการณ์เหล่านี้เป็นสิ่งที่เกิดขึ้นโดยบังเอิญไม่คาดคิดมาก่อน เป็นการเปลี่ยนแปลงที่เป็นเชิงสุ่ม(random variation) เพราะไม่ได้อยู่ภายใต้เงื่อนไขที่เรากำหนด

จากองค์ประกอบของอนุกรมเวลาทั้ง 4 อย่าง คือ T S C และ I ในข้อมูลอนุกรมชุดหนึ่ง ๆ ไม่จำเป็นต้องครบองค์ประกอบข้างต้นก็ได้ ทั้งนี้ ขึ้นอยู่กับชนิดข้อมูลของเรา

2.4.3 รูปแบบของอนุกรมเวลา

ปัจจัยทั้ง 4 ข้างต้น ถ้า Y แทนข้อมูลอนุกรมเวลาชุดหนึ่ง ๆ เราสามารถกำหนดแบบจำลองได้ 2 แบบ ดังนี้

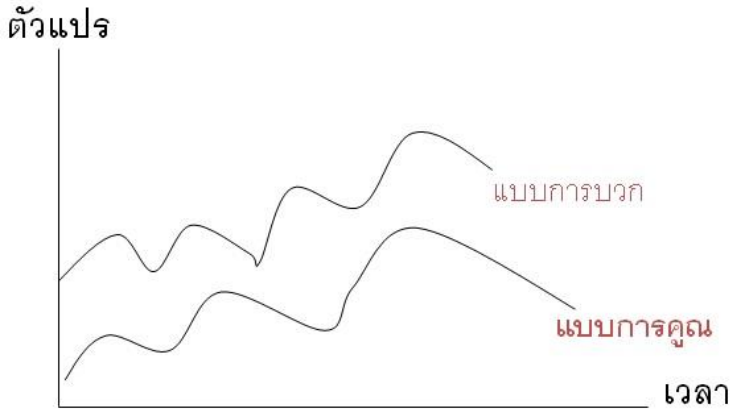
1. แบบจำลองผลบวก (Additive model) ถือว่าข้อมูลในแต่ละอนุกรมเวลาประกอบด้วยผลบวกขององค์ประกอบทั้ง 4 อย่าง

$$Y_t = T_t + S_t + C_t + I_t \qquad (2.1)$$

2. แบบจำลองผลคูณ (Multiplicative model) ถือว่าข้อมูลในแต่ละอนุกรมเวลาประกอบด้วยผลคูณขององค์ประกอบทั้ง 4 อย่าง

$$Y_t = T_t * S_t * C_t * I_t \qquad (2.2)$$

แบบแผนอนุกรมเวลา ดังภาพประกอบที่ 2.6



ภาพประกอบที่ 2.6 แบบแผนอนุกรมเวลา

ในงานวิจัยนี้จะทำการนำข้อมูลอนุกรมเวลาเพื่อทำการศึกษาถึงรูปแบบของการเกิดอุบัติเหตุ ซึ่งอาจเกิดจากการเปลี่ยนแปลงหรือความแปรผัน เนื่องจากเหตุการณ์ผิดปกติ

การพยากรณ์ (Forecasting) [20] หมายถึงการคาดการณ์ (Predict) เกี่ยวกับแนวโน้มของสิ่งที่สนใจที่จะเกิดขึ้นในอนาคต เพื่อใช้เป็นสารสนเทศ (Information) ประกอบการตัดสินใจ ซึ่งการพยากรณ์จะต้องดำเนินการ เป็นส่วนแรกก่อนที่จะวางแผน หรือ การเตรียมการที่จะเริ่มทำอะไรเพื่อความถูกต้อง และ แม่นยำในการตัดสินใจ ดังนั้น ในการดำเนินธุรกิจภายใต้ความไม่แน่นอนจำเป็นที่จะต้องทราบถึงความเป็นไปในอนาคต โดยอาศัยเทคนิคการพยากรณ์ต่าง ๆ เท่าที่จำเป็น [21, 22]

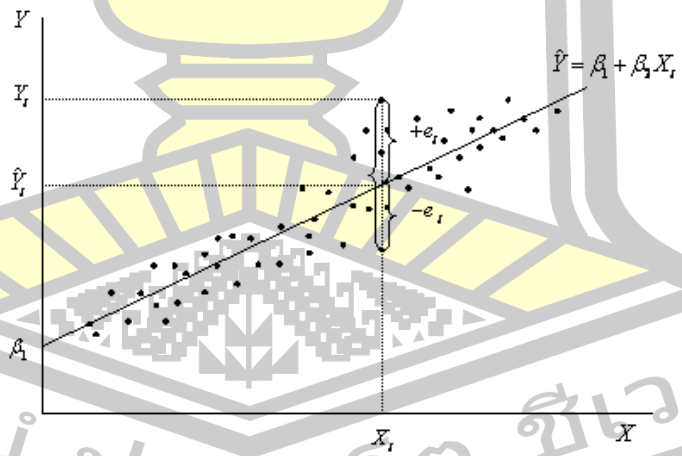
2.5 เทคนิคการพยากรณ์แบบอนุกรมเวลา

เทคนิคในเหมืองข้อมูลได้ถูกนำมาใช้ในการสร้างแบบจำลองกันอย่างแพร่หลายและมีประสิทธิภาพ โดยผู้ทำการวิจัยได้ศึกษาและค้นคว้าเกี่ยวกับเทคนิคการพยากรณ์จากงานวิจัยและ

แหล่งข้อมูลต่าง โดยเทคนิคที่ใช้สำหรับการพยากรณ์แบบอนุกรมเวลาได้แก่ การถดถอยเชิงเส้น (linear regression: LR) โครงข่ายประสาทเทียม (Artificial Neural Networks: ANN), Support Vector Machine for Regression (SVMR), Sequential Minimal Optimization for Regression (SMOreg) และ Gussian Process (GP) ซึ่งสามารถอธิบายได้ดังต่อไปนี้

2.5.1 วิธีการพยากรณ์โดยใช้การวิเคราะห์เชิงเส้น (Linear regression analysis)

การถดถอยเชิงเส้น (linear regression: LR) แบ่งออกเป็น สองชนิดคือ Linear Regression, Non-linear Regression การวิเคราะห์ Linear Regression แบ่งออกเป็น Simple Linear Regression (SLR), Multiple Linear Regression (MLR) การวิเคราะห์การถดถอยอย่างง่าย (Simple Regression Analysis) เป็นการศึกษาความสัมพันธ์ระหว่างตัวแปร 2 ตัว (ในที่นี้คือตัวแปร X และ Y) ที่มีความสัมพันธ์กันในลักษณะเชิงเส้น(Linear) โดยมีสมการถดถอย คือสมการเส้นตรง ซึ่ง และ เป็นพารามิเตอร์ที่ไม่ทราบค่าจึงจะต้องประมาณค่าโดยใช้ข้อมูลจากตัวอย่างโดยวิธีที่นิยมใช้ในการประมาณค่าของ และ ก็คือวิธี กำลังสองน้อยที่สุด (The Least Squares Method) ซึ่งจะแทนค่าของ และ ด้วยค่า a และ b โดยที่ a ก็คือค่าคงที่ (Constant) เป็นค่าที่เส้นกราฟถดถอยตัดกับแกน Y ส่วน b เป็นความชัน (Slope) ของเส้นกราฟซึ่งแสดงถึงอัตราการเปลี่ยนแปลงของ Y เมื่อ X เปลี่ยนแปลงเรียกส่วนนี้ว่าสัมประสิทธิ์การถดถอย (Regression Coefficient) หรือสัมประสิทธิ์การพยากรณ์สมการที่ได้คือ $Y = a + bX$



ภาพประกอบที่ 2.7 การทำงานของ Linear Regression

ข้อดีของ Linear Regression คือ

สามารถวิเคราะห์การถดถอยจึงเป็นการศึกษาหาอิทธิพลของตัวแปร 2 ตัวแปร

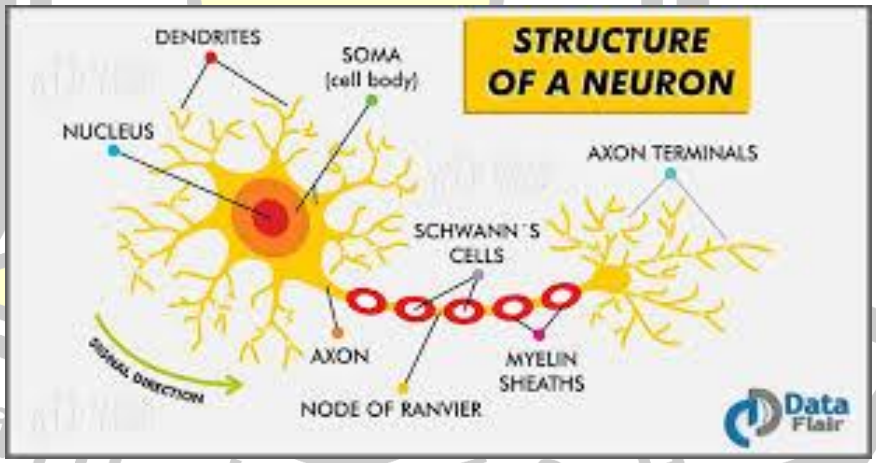
สามารถดูว่าหาก X เพิ่มขึ้น 1 หน่วย Y จะเพิ่มขึ้นหรือลดลงเป็นเท่าไร

พหุน ปณ ฑิโต ชีเว

สามารถนำไปสร้างเป็นสมการทำนาย สำหรับทำนายผลในอนาคต

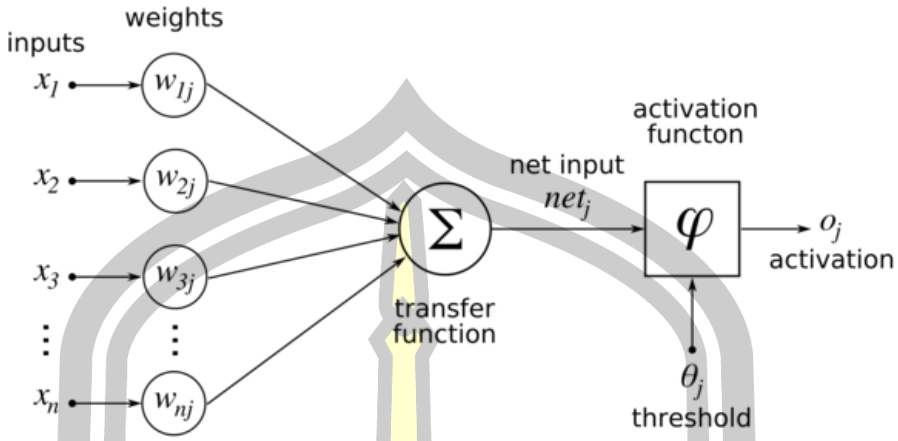
สามารถพิจารณาได้ว่าตัวแปรทั้ง 2 มีความสัมพันธ์กันหรือไม่

2.5.2 โครงข่ายประสาทเทียม (Artificial Neural Networks: ANN) เป็นเทคนิคของการทำเหมืองข้อมูล เป็นโมเดลทางคณิตศาสตร์ โดยการคำนวณแบบ (Connectionist) โดยทำงานของโครงข่ายประสาทเทียม (Artificial Neural Networks) เหมือนกับการทำงานของเครือข่ายประสาทในสมองมนุษย์ ซึ่งมีความสามารถในการเรียนรู้ การจดจำ รูปแบบ (Pattern Recognition) และการสร้างความรู้ใหม่ (Knowledge Extraction) แนวคิดนี้ได้มาจากการศึกษาโครงข่ายไฟฟ้าชีวภาพ (Bioelectric Network) ในสมอง ซึ่งประกอบด้วย เซลล์ประสาท หรือ “นิวรอน” (Neurons) และ “จุดประสานประสาท” (Synapses) แต่ละเซลล์ประสาทประกอบด้วยปลายในการรับกระแสประสาท เรียกว่า “เดนไดรท์” (Dendrite) ซึ่งเป็น input และปลายในการส่งกระแสประสาทเรียกว่า “แอคซอน” (Axon) ซึ่งเป็นเหมือน output ของเซลล์การทำงานของ Neural Networks คือ การนำ Input เข้าไปใน Network โดยเอา Input มาคูณกับ Weight ของแต่ละขาผลที่ได้จาก Input ทุกขาของ Neuron จะเอามารวมกันต่อมา เอาค่าที่ได้มาเทียบกับ Threshold ที่กำหนดไว้ถ้าผลรวมมีค่ามากกว่า Threshold แล้ว Neuron ก็จะส่ง Output ออกไป Output นี้จะถูกส่งไปยัง Input ของ Neuron ที่เชื่อมกันใน Network ถ้าค่าน้อยกว่า Threshold ก็จะไม่เกิด Output สมการทางคณิตศาสตร์ที่ได้ คือ



ภาพประกอบที่ 2.8 แสดง Model ของ Neuron ในสมองมนุษย์

ที่มา : Data Flair Artificial Neural Network (ANN) in Machine Learning [23]



ภาพประกอบที่ 2.9 รูปแบบการทำงานของ ANN แบบ Multi-Layer Prpceptron

ที่มา : https://en.wikibooks.org/wiki/Artificial_Neural_Networks/Activation_Functions

ข้อดี มีไว้สำหรับกรณีที่ปัญหามีความซับซ้อน

2.5.3 Support Vector Machine Regression (SVMR) ซัพพอร์ทเวกเตอร์แมชชีนแบบจำแนกประเภทข้อมูลนั้นสามารถแบ่งข้อมูลเป็นประเภท+หรือ– ได้ โดยใช้ +1แทนข้อมูลประเภท+ และใช้ตัวเลข -1แทนข้อมูลประเภท– ส่วนงานวิจัยนี้เราต้องการนำ ซัพพอร์ทเวกเตอร์แมชชีนมาใช้ในการทำนายค่าตัวเลขที่เป็นจำนวนจริง ซึ่งทำได้ด้วยการใช้ฟังก์ชันสูญเสียแบบ **ε**-insensitive (**ε**-insensitive loss function) เข้ามาช่วยในการสร้าง ( ) k f x ที่ทุกๆข้อมูลการฝึกxมีความเบี่ยงเบน **ε**จากค่า y ที่เป็นเป้าหมายตามเงื่อนไข • + – ≤ **ε** i i (w x b) y และ y – (w• x + b) ≤ **ε** i iสำหรับ i = 1,□,l ซัพพอร์ตเวกเตอร์แมชชีนแบบนี้เรียกว่าการถดถอยด้วยซัพพอร์ตเวกเตอร์แมชชีนหรือ SVMR แบบระยะขอบแข็ง(hard margin)เงื่อนไขเพิ่มเติม คือ

$$\bullet + - \leq + + i\ i\ i\ (w\ x\ b)\ y\ \varepsilon\ \xi$$

$$- \bullet + \leq + - i\ i\ i\ y\ (w\ x\ b)\ \varepsilon\ \xi$$

$$+ \geq 0$$

$$i\ \xi$$

และ – ≥ 0

$$i\ \xi$$

สำหรับ i = 1,...,l []

ξ เป็นตัวปรับข้อมูลที่ผิดพลาด เรียกว่า ตัวแปรหย่อน(Slack variable) การทำงานของSVMR

ข้อดี เป็นการวิเคราะห์ข้อมูลอนุกรมเวลาเพื่อการพยากรณ์ที่เกิดความผิดพลาดต่ำสุด

พหุน ปณ ศโต ชีเว

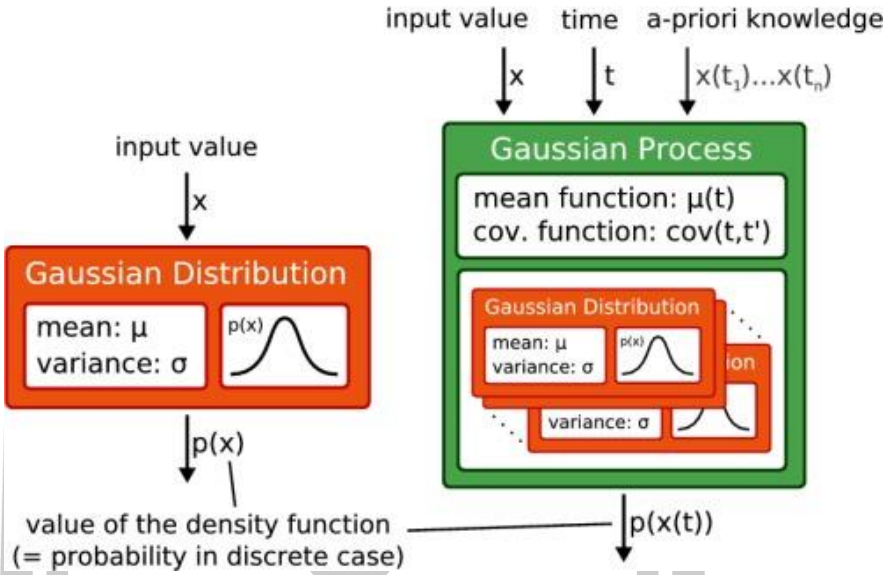
A Library for Support Vector Machines (LibSVM) [24] เป็นไลบรารี่สำหรับ SVM ที่พัฒนาโดย Chih-Chung Chang และ Chih-Jen Lin ซึ่ง LibSVM นี้จะช่วยให้การวิเคราะห์SVM นั้นง่ายขึ้น โดย LibSVM ได้รองรับซัพพอร์ตเวกเตอร์แบบการแบ่งกลุ่ม (Support Vector Classification: SVC) แบบการวิเคราะห์แบบรีเกรสชัน (Support Vector Regression: SVMR) Sven และ Rohit [6] ได้สร้างแบบจำลองเพื่อการพยากรณ์ข้อมูลอนุกรมเวลาในรูปแบบฤดูกาลจำนวน 8 ชุดข้อมูลด้วยโครงข่ายประสาทเทียม โดยทดลองกับโมเดลโครงข่ายประสาทเทียมแบบต่างๆ แสดงให้เห็นว่าโครงข่ายประสาทเทียมมีประสิทธิภาพในการแก้ปัญหาการพยากรณ์อนุกรมเวลาได้เป็นอย่างดี

2.5.4 Sequential Minimal Optimization for Regression (SMOreg) ถูกพัฒนาจาก SMO อัลกอลิทึม บนพื้นฐานของ SVM (Support Vector Machine) วิธีนี้ใช้การแทนค่าที่ข้อมูลสูญหายและแปลงข้อมูลคุณลักษณะเชิงกลุ่ม (nominal) ให้เป็นข้อมูลไบนารี่ (Binary) นอกจากนี้ยังทำให้ข้อมูลคุณลักษณะทุกค่าอยู่ในรูปแบบมาตรฐาน (Normalized)

ข้อดี มีคุณสมบัติใช้งานกับ non-linear ได้อย่างมีประสิทธิภาพ SMOreg ยังช่วยในการจัดการโครงสร้างของ Model และลดความแปรปรวนของข้อมูลทำให้ข้อมูลมีความน่าเชื่อถือในการพยากรณ์

2.5.5 Gussian Process (GP) เป็นเทคนิคที่นำเอาระบบ Stochastic Process ซึ่ง ปัจจุบันถูกยอมรับแล้วว่าเป็นเครื่องมือที่มีประสิทธิภาพในการแก้ปัญหาประเภท Regression, Classification, และ Decision ใน Machine learning สามารถทำงานได้ดีถึงแม้ว่ามี Training Data น้อยและมีประสิทธิภาพและ Convergence rate ดีกว่า ARMA, NN และ SVR ตามลำดับ[25] ซึ่งมีงานวิจัยของทัศนัย พลอยสุวรรณ และคณะ นำเสนอการพยากรณ์หาค่าความต้องการไฟฟ้าสูงสุด (Peak load) ระยะยาวระหว่างปี 2011-2012 โดยใช้ GaussianProcess ซึ่งผลการพยากรณ์ที่ดีกว่าเมื่อเปรียบเทียบกับกับการใช้โครงข่ายประสาทเทียม (Neural Networks)

พหุน ปณฺฑิโต ชีเว



ภาพประกอบที่ 2.10 หลักการทำงานของ Guassian Process

ข้อดี มีประสิทธิภาพและสามารถทำงานได้ดี มี Training Data น้อย มีการเบี่ยงเบนข้อมูลน้อย

3.การวัดประสิทธิภาพของแบบจำลอง

ในการวัดประสิทธิภาพการทำงานของแบบจำลองเพื่อการพยากรณ์ในข้อมูลอนุกรมเวลาจำเป็นต้องแบ่งข้อมูลออกเป็น 2 ชุด คือ ข้อมูลการเรียนรู้ (Training data) และ (2) ชุดข้อมูลการทดสอบ (Testing data) โดยใช้หลักการ Sliding Windows

Sliding window เป็นการตรวจสอบเป็นพิเศษสามารถใช้สำหรับการคาดการณ์ชุดที่จุดเวลาที่มีการเข้ารหัสเป็นตัวอย่าง จะใช้หน้าต่างบางตัวอย่างสำหรับการเรียนรู้ และใช้หน้าต่าง อีกส่วนหนึ่งสำหรับการทดสอบ หน้าต่างจะถูกย้ายข้ามชุดตัวอย่างและการวัดประสิทธิภาพการทำงานที่นำมาเฉลี่ยหลังจากนั้น พารามิเตอร์ ระบุว่าตัวอย่างทั้งหมดในอดีตควรจะใช้สำหรับการฝึกอบรม (แทนที่จะเป็นเพียงหน้าต่างปัจจุบัน) เช่น มีข้อมูลการเกิดอุบัติเหตุตั้งแต่เดือน มกราคม 2554 ถึง ธันวาคม 2556 นำไปสร้างแบบจำลอง และนำข้อมูลในเดือนมกราคม 2557 ถึง มิถุนายน 2560 มาทำการทดสอบ เป็นต้น

ส่วนการวัดค่าความคาดเคลื่อนของการพยากรณ์ ทำให้ทราบว่าค่าพยากรณ์ที่ได้จากวิธีดังกล่าวเป็นค่าการพยากรณ์ที่เหมาะสมและใกล้เคียงกับสภาพความเป็นจริงมากที่สุด ซึ่งจะใช้ดัชนีการประเมิน 2 ตัว คือ Mean Absolute Error และ Root Mean Square Error

1) ค่าเฉลี่ยความคลาดเคลื่อนสมบูรณ์ (Mean Absolute Error: MAE) ค่าเฉลี่ยความคลาดเคลื่อนสมบูรณ์ เป็นวิธีการวัดค่าความแตกต่างระหว่างค่าจริงและค่าที่ประมาณได้จาก

แบบจำลอง หากค่า MAE มีค่าน้อย แสดงว่าแบบจำลองสามารถประมาณค่าได้ใกล้เคียงกับค่าจริง ดังนั้นหากค่านี้มีค่าเท่ากับศูนย์แล้วจะหมายความว่า ไม่เกิดความคลาดเคลื่อนในแบบจำลองนี้เลย ค่า MAE สามารถคำนวณได้ดังสมการ

$$MAE = \frac{1}{n}\sum_{i-1}^{n}\left|\hat{x}_i - x_i\right|$$

เมื่อ n คือ จำนวนตัวแปร

$\hat{x}_i$ คือ ค่าพยากรณ์

x_i คือ ค่าจริง

2) ค่าความคลาดเคลื่อนรากกำลังสองเฉลี่ย (Root Mean Square Error: RMSE) ความคลาดเคลื่อนรากกำลังสองเฉลี่ย คือ ค่าเฉลี่ยส่วนเบี่ยงเบนมาตรฐาน (RMSD) เป็น วิธีการวัดความคลาดเคลื่อนจากค่าที่พยากรณ์จากแบบจำลองกับค่าจริงที่เกิดขึ้น หากค่า RMSE มีค่าน้อย แสดงว่าแบบจำลองสามารถประมาณค่าได้ใกล้เคียงกับ ค่าจริง ดังนั้นหากค่านี้มีค่าเท่ากับศูนย์แล้ว จะหมายความว่า ไม่เกิดความ คลาดเคลื่อนในแบบจำลอง ค่า RMSE สามารถคำนวณได้ดังสมการ

$$RMSE = \sqrt{\frac{\sum_{t-1}^{T}\left(y_i - \hat{y}_i\right)^2}{n-1}}$$

เมื่อ n คือ จำนวนตัวแปร

$\hat{x}_i$ คือ ค่าพยากรณ์

x_i คือ ค่าจริง

2.6 งานวิจัยที่เกี่ยวข้องกับการเกิดอุบัติเหตุบนท้องถนน

1. Song bai และคณะ [26] ได้ทำการศึกษาเรื่องงานวิจัยการเกิดอุบัติเหตุบนท้องถนน โดยใช้เทคนิคกฎการหาความสัมพันธ์หลายมิติ มีวัตถุประสงค์วิเคราะห์หาปัจจัยที่ก่อให้เกิดอุบัติเหตุบนทางด่วนในประเทศจีน ทำการศึกษาและเก็บรวบรวมข้อมูลการเกิดอุบัติเหตุบนทางด่วน โดยการประยุกต์ใช้เทคนิคการทำเหมืองข้อมูล ประเภทกฎการหาความสัมพันธ์หลายมิติ ได้ถูกนำไปใช้ในการพัฒนาระบบ จากการศึกษาพบว่า สามารถพัฒนาระบบอัจฉริยะเพื่อช่วยตัดสินใจในขณะเกิดอุบัติเหตุบนท้องถนนแผนกขนส่งได้

2. พิณรัตน์ นุชโพธิ์ [27] ได้ทำการศึกษาเรื่องการศึกษาอุบัติเหตุการจราจรทางถนนในประเทศไทย โดยใช้กฎการหาความสัมพันธ์ของการทำเหมืองข้อมูล มีวัตถุประสงค์เพื่อวิเคราะห์หาโอกาสที่ทำให้เกิดอุบัติเหตุการจราจรทางถนนในประเทศไทย โดยทำการศึกษาปัจจัยที่เกี่ยวข้องกับการเกิดอุบัติเหตุที่ผ่านมาแล้วนำมาวิเคราะห์หาสาเหตุที่ทำให้มีโอกาสเกิดอุบัติเหตุ ซึ่งมีทั้งหมด 99 แอตทริบิวต์ มีการจัดกลุ่มข้อมูลให้มีความสมบูรณ์เพื่อลดการกระจายของข้อมูล และทำการสุ่มค่า

ความน่าจะเป็นที่น้อยที่สุด ซึ่งค่าที่พบที่เหมาะสม คือ 0.1 เพื่อคำนวณหาโอกาสที่เกิดอุบัติเหตุในสถานการณ์ต่างๆ จากผลการศึกษาทำให้ทราบความสัมพันธ์กันของปัจจัยระดับความรุนแรง และเพศ พบว่า เพศชายมีระดับความรุนแรงของการเกิดอุบัติเหตุทำให้บาดเจ็บสาหัสที่มากกว่าเพศหญิง

3. Li-Yen Chang และ Hsiu-Wen Wang [28] ได้ทำการศึกษาเรื่องการวิเคราะห์ระดับความรุนแรงของการบาดเจ็บบนท้องถนน ด้วยโปรแกรมที่ไม่ใช่ตัวแปร โดยใช้เทคนิคการจัดหมวดหมู่ต้นไม้ มีวัตถุประสงค์เพื่อหาปัจจัยที่ก่อให้เกิดอุบัติเหตุบนท้องถนน ศึกษาและรวบรวมข้อมูลการเกิดอุบัติเหตุของเมืองไทเปประเทศไต้หวัน ในปี 2001 โดยการประยุกต์ใช้เทคนิคเหมืองข้อมูลด้วยวิธี Classification and Regression Tree(CART) หาความสัมพันธ์ระดับความรุนแรงของการเกิดอุบัติเหตุการจราจรบนท้องถนน จากการศึกษาพบว่า ปัจจัยด้านสิ่งแวดล้อม มนุษย์และยานพาหนะเป็นปัจจัยเสี่ยงที่ก่อให้เกิดอุบัติเหตุบนท้องถนนได้

4. Mats Ramstedt [29] ได้ทำการศึกษาเพื่อประเมินความสัมพันธ์ระหว่างการบริโภคเครื่องดื่มแอลกอฮอล์และอุบัติเหตุร้ายแรงในประเทศสหรัฐอเมริกาเพื่อเปรียบเทียบกับผลการวิจัยจากยุโรปและแคนาดาโดยใช้ข้อมูลประจำปีเกี่ยวกับอุบัติเหตุร้ายแรงแบ่งตามเพศและอายุ เพื่อวิเคราะห์เกี่ยวข้องกับการบริโภคเครื่องดื่มแอลกอฮอล์ของปี 1950-2002 โดยใช้เทคนิค Box-เจนกินส์ สำหรับการวิเคราะห์อนุกรมเวลา

ผลการทดลองสรุปได้ว่าประชากรบริโภคเครื่องดื่มแอลกอฮอล์ในวัยรุ่นชายกลุ่มอายุ 15-34 และกลุ่ม 35-49 มีอัตราการเกิดอุบัติเหตุที่ร้ายแรง ในทางตอนเหนือของยุโรป แคนาดา และยุโรปใต้ มีความรุนแรงมากในช่วงอายุ 15-34 ทางยุโรปใต้ อุบัติเหตุที่เกิดขึ้นจากการดื่มแอลกอฮอล์ ทางตอนเหนือของยุโรปคิดเป็น 45% ยุโรปกลาง 40% ยุโรปใต้ 35% และสหรัฐอเมริกา *39%* คิดจากสัดส่วนประชากร 100,000 คน ของอุบัติเหตุร้ายแรงทั้งหมดที่มีเครื่องดื่มแอลกอฮอล์

5. วราพร งามสุข [5] ได้ทำการศึกษาเพื่อเปรียบเทียบการพยากรณ์อนุกรมเวลาโดยวิธีแบบฉบับและวิธีบอกซ์-เจนกินส์ โดยใช้ข้อมูลสถิติจำนวนการเกิดอุบัติเหตุในประเทศไทยโดยใช้ข้อมูลรายเดือนระหว่างเดือนมกราคม พ.ศ. 2545 ถึง เดือนธันวาคม พ.ศ. 2554 รวมทั้งสิ้น 72 เดือน จากการศึกษาพบว่าผลการพยากรณ์โดยวัดค่าประสิทธิภาพโดยใช้ (RMSE) เท่ากับ 229.5816 ครั้ง และตัวแบบพยากรณ์ที่เหมาะสมสำหรับ การพยากรณ์โดยวิธีบอกซ์-เจนกินส์คือ ARIMA 0,1,3 , 2,1,0 โดยค่าประสิทธิภาพที่ได้ (RMSE) เท่ากับ 51.5996 ดังนั้นการพยากรณ์ข้อมูลอนุกรมเวลาโดยวิธีบอกซ์-เจนกินส์ จึงมีความเหมาะสมกับข้อมูลชุดดังกล่าว

6. พิณรัตน์ นุชโพธิ์ และคณะ [30] ได้ทำการศึกษาความสัมพันธ์ปริมาณอุบัติเหตุในรูปอนุกรมเวลาแบบแยกส่วนประกอบ ทำนายปริมาณอุบัติเหตุบนถนนหลวง ในรูปอนุกรมเวลาแบบแยกส่วนประกอบ (Decomposition Approach) โดยรวบรวมฐานข้อมูลจำนวนอุบัติเหตุที่เกิดขึ้น จากสำนักอำนวยความปลอดภัย กรมทางหลวงและสำนักงานตำรวจแห่งชาติ ตั้งแต่ปี พ.ศ. 2538 ถึง พ.ศ.

2546 มาทำการวิเคราะห์ในการสร้างรูปแบบจำลอง การพยากรณ์ปริมาณอุบัติเหตุในรูปแบบอนุกรมเวลาแบบแยกส่วนประกอบ การทดสอบแผนการกระจายของข้อมูลในรูปแบบอนุกรมเวลาทั้งแบบเชิงเส้นตรง และแบบไม่เชิงเส้นตรงเพื่อตรวจสอบดูแนวโน้มการเกิดอุบัติเหตุในอนุกรมเวลา

ผลแสดงการพยากรณ์จำนวนอุบัติเหตุบนทางหลวงช่วงระหว่างปี พ.ศ. 2538 ถึง พ.ศ. 2546 จะเห็นได้ว่า ปริมาณอุบัติเหตุบนทางหลวงแผ่นดินของประเทศไทยมีแนวโน้มตามรูปแบบพาราโบล่าตามส่วนประกอบการวิเคราะห์ที่ได้จากตัวข้อมูลปริมาณอุบัติเหตุตามแนวโน้มของเวลา และอิทธิพลอันเนื่องมาจากเทศกาลต่าง ๆ ที่มีส่วนเกี่ยวข้องต่อผู้ใช้ และมีช่วงที่เกิดอุบัติเหตุสูงสุดประจำปีประมาณช่วงเดือนเมษายน และเดือนมกราคม

7. จันทนี วงศ์พานิช [31] ได้ศึกษาเปรียบเทียบวิธีพยากรณ์เดี่ยวกับ วิธีพยากรณ์รวม โดยใช้ข้อมูลอนุกรมเวลาของจำนวนคดีอุบัติเหตุจราจรทางบกที่เพิ่มขึ้นใน กรุงเทพมหานคร วิธีพยากรณ์เดี่ยวมี 3 วิธี ได้แก่ วิธีกำลังสองน้อยที่สุดที่มีแนวโน้มลักษณะ เอ็กซ์โพเนนเชียล วิธีโฮลท์-วินเทอร์ และวิธีบ๊อกซ์-เจนกินส์ สำหรับวิธีพยากรณ์รวมได้ใช้ วิธีเฉลี่ยถ่วงน้ำหนักของเบทส์ แกรงเจอร์ และนิวโบลด์ เป็นเทคนิคในการรวมวิธีพยากรณ์ เดี่ยว 3 คู่เข้าด้วยกัน

ผลการศึกษาพบว่า โดยเฉลี่ยแล้ววิธีพยากรณ์รวมสามารถพยากรณ์ข้อมูลอุบัติเหตุจราจรที่เกิดขึ้นล่วงหน้าในกรุงเทพมหานคร ได้แม่นยำกว่าวิธีพยากรณ์เดี่ยว วิธีพยากรณ์เดี่ยววิธีกำลังสองน้อยที่สุดเหมาะสมสำหรับพยากรณ์ข้อมูลที่จะเกิดในอนาคตในช่วงสั้นๆ 3 เดือน เช่นเดียวกับวิธีบ๊อกซ์-เจนกินส์ เหมาะสม สำหรับพยากรณ์ข้อมูลที่เกิดขึ้นในอนาคตในช่วงสั้นๆ 3 เดือน โดยใช้ข้อมูลที่เกิดขึ้นในอดีต 60 เดือน สำหรับวิธีโฮลท์-วินเทอร์เหมาะสมสำหรับพยากรณ์ข้อมูลที่จะเกิดในอนาคตในช่วง 6 ถึง 12 เดือน วิธีพยากรณ์รวมที่เหมาะสมได้แก่ วิธีโฮลท์-วินเทอร์รวมกับวิธีบ๊อกซ์-เจนกินส์ ซึ่งสามารถพยากรณ์ข้อมูลที่จะเกิดในอนาคต ได้ทั้งในระยะสั้นและระยะยาว โดยใช้ข้อมูลในอดีต 60 เดือน ผลพยากรณ์ของวิธีนี้มีความคลาดเคลื่อนต่ำ

8. เกษม ชูจารุกุล และคณะ [32] ได้นำเสนอการวิเคราะห์ข้อมูลเบื้องต้นทางด้านความปลอดภัยทางถนน โดยอาศัยการคาดการณ์การเกิดอุบัติเหตุจากแบบจำลองทางสถิติ ซึ่งผลจากการศึกษาพบว่าสายทางบนทางหลวงในประเทศไทยที่อยู่ในความรับผิดชอบของกรมทางหลวง แบบทวินามลบจะมีความเหมาะสมและถูกต้องมากกว่าแบบจำลองแบบปัวซอง โดยตัวแปรที่พบว่ามีนัยสาคัญทางสถิติประกอบด้วย ความยาวช่วงสายทาง ปริมาณจราจร จำนวนช่องจราจร สายทางที่มีสะพานข้าม และลักษณะสายทางที่เป็นแนวราบ แนวดิ่ง ทางแยก และทางเชื่อม จุดประสงค์เพื่อคาดการณ์จำนวนอุบัติเหตุที่จะเกิดขึ้นในแต่ละช่วงสายทาง

ผลลัพธ์จากการวิเคราะห์เปรียบเทียบผลกระทบที่เกิดขึ้นจากแบบจำลองอุบัติเหตุ ทำให้ทราบถึงผลกระทบของปัจจัยด้านเรขาคณิตของถนน และปัจจัยที่เกี่ยวข้องกับการจราจรต่อการเกิดอุบัติเหตุ ไปประยุกต์ใช้คาดการณ์ความรุนแรงในการเกิดอุบัติเหตุได้ โดยแบ่งการพิจารณาตามระดับ

ความรุนแรง นอกจากนี้ยังสามารถพัฒนาวิธีการศึกษาโดยใช้แบบจาลองทางสถิติประเภทอื่นๆ ที่อาจจะมีความสอดคล้องกับข้อมูลที่แตกต่างกันในแต่ละพื้นที่ เพื่อเป็นการบูรณการให้แบบจำลองการคาดการณ์จำนวนอุบัติเหตุที่จะเกิดขึ้นในประเทศไทยมีประสิทธิภาพ อันจะเป็นประโยชน์ในการแก้ไขปัญหาอุบัติเหตุทางถนนต่อไป

9. Chong [33] ได้ทำการศึกษาเรื่องการวิเคราะห์การเกิดอุบัติเหตุบนท้องถนน โดยใช้เทคนิคการเรียนรู้เชิงกลไก มีวัตถุประสงค์เพื่อวิเคราะห์สาเหตุการเกิดอุบัติเหตุบนท้องถนนและสาเหตุการบาดเจ็บที่เกิดขึ้น โดยข้อมูลการเกิดอุบัติเหตุบนท้องถนนได้จากตัวอย่างระบบยานยนต์แห่งชาติ และตัวอย่างระบบประมาณการทั่วไป ระหว่างปี พ.ศ 2538 ถึง พ.ศ. 2543 จากผลการทดลองพบว่าเทคนิคต้นไม้ตัดสินใจร่วมกับเทคนิคโครงข่ายประสาทเทียม ให้ความถูกต้องที่ดีที่สุดคิดเป็นร้อยละ 90

10. อมรภัทร์ และคณะ[34] ได้สร้างแบบจำลองทำนายความเสี่ยงการเกิดอุบัติเหตุบนท้องถนนโดยใช้เทคนิคฤงจำแนก พบว่าระหว่างปี พ.ศ.2552 ถึง พ.ศ. 2556 จากการทดลองพบว่าเทคนิคฤงจำแนกสามารถสร้างแบบจำลองที่มีประสิทธิภาพมีค่าความแม่นยำมากกว่าแบบจำลองจำแนกเบย์ และแบบจำลองเครื่องมือสนับสนุนเส้นสมมุติ ยิ่งไปกว่านั้นยังพบอีกว่าเทคนิคฤงจำแนกสามารถเพิ่มค่าความแม่นยำให้กับแบบจาลองเครื่องมือสนับสนุน จากเดิมคิดเป็นร้อยละ 0.86

2.7 งานวิจัยที่เกี่ยวข้องกับเทคนิคอนุกรมเวลา

1. เศรษฐชัย ชัยสนิทและ รัชฎาวรรณ นิ่มนวล [35] ได้ศึกษาเรื่อง โปรแกรมพยากรณ์ความต้องการผลไม้ไทยเพื่อการส่งออก ด้วยเทคนิคการถดถอยเชิงเส้น โดยใช้เทคนิคการถดถอยเชิงเส้นเพื่อประเมินประสิทธิภาพของโปรแกรมพยากรณ์ที่สร้าง เครื่องมือที่ใช้ในการศึกษาได้แก่ เทคนิคการถดถอยเชิงเส้น (Linear Regression) ข้อมูลที่ใช้ในการศึกษาคือ ข้อมูลสถิติการส่งออกขององค์การตลาดเพื่อการเกษตรกระทรวงเกษตรและสหกรณ์ ตั้งแต่ปี พ.ศ. 2547 ถึง พ.ศ. 2552 ผลที่ได้จากงานวิจัยนี้พบว่าโปรแกรมพยากรณ์ ความต้องการผลไม้ไทยเพื่อการส่งออกโดยใช้เทคนิคการถดถอยเชิงเส้น (Linear Regression) มีค่าเปอร์เซ็นต์ความผิดพลาดสมบูรณ์เฉลี่ย (MAPE) น้อยกว่า15% ซึ่งสรุปได้ว่าโปรแกรมพยากรณ์ที่สร้างขึ้นนี้มีประสิทธิภาพในการพยากรณ์ เป็นไปตามสมมติฐานที่ตั้งไว้อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.01

2. สักการะ จู่ซู่และธนพลเจน สุทธิเวชกุล [36] ได้ทำการเปรียบเทียบเทคนิคการพยากรณ์แบบจำลองการถดถอยและแบบจำลองโครงข่ายประสาทเทียม โดยใช้ข้อมูลปริมาณการยืมหนังสือห้องสมุดโรงเรียนนาบอนระหว่างปี พ.ศ. 2549 ถึง พ.ศ. 2552 ผลการวิจัยพบว่า แบบจำลองโครงข่ายประสาทเทียม มีค่าพยากรณ์ปริมาณการยืมหนังสือใกล้เคียงกับค่าปริมาณการยืมที่เกิดขึ้น

จริงมากกว่าแบบจำลองการถดถอย และค่าความแม่นยำของแบบจำลองโครงข่ายประสาทเทียม มีความแม่นยำสูงกว่าแบบจำลองการถดถอย

3. ทวี ชัยพิมลผลินและทวีศักดิ์ วังไพศาล [24] ได้ศึกษาเรื่อง แบบจำลองโครงข่ายประสาทเทียมสำหรับการพยากรณ์น้ำท่วมในลุ่มน้ำมูลตอนล่าง โดยงานวิจัยนี้ทำการศึกษาเทคนิคการคัดเลือกข้อมูล 5 วิธี เพื่อให้แบบจำลองโครงข่ายประสาทเทียมพยากรณ์ระดับน้ำ 24 และ 48 ชั่วโมงล่วงหน้า โดยข้อมูลนำเข้าเป็นข้อมูลระดับน้ำรายชั่วโมงและยังมีการทดสอบกระบวนการเรียนรู้ของโครงข่ายประสาทเทียม ซึ่งจากผลการวิจัยพบว่าเทคนิคการคัดเลือกข้อมูลนำเข้าวิธี C และ CS เป็นเทคนิคที่คัดเลือกข้อมูลที่ดีที่สุด จำนวนโหนดในชั้นซ่อนเร้นที่เหมาะสมคือ 1 โหนดและกระบวนการเรียนรู้ทั้ง 2 ประเภทให้ผลการพยากรณ์ที่ใกล้เคียงกันโดยมีประสิทธิภาพในการพยากรณ์ค่อนข้างดีโดยมีค่าความคลาดเคลื่อนประมาณ 8.7 และ 12.7 เซนติเมตร หรือ ค่า CE มีค่า 0.99 และ 0.98 สำหรับการพยากรณ์ 24 และ 48 ชั่วโมงล่วงหน้า ตามลำดับ

4. อภิญญา และประสิทธิ์ พยัคฆพงษ์ [37] ได้ทำการศึกษา เพื่อเปรียบเทียบวิธีการพยากรณ์ราคาถั่วเหลืองและปาล์มน้ำมันโดยใช้วิธีของบ๊อกซ์-เจนกินส์และโครงข่ายประสาทเทียม และหารูปแบบที่เหมาะสมในการพยากรณ์ โดยใช้ข้อมูลรายเดือนตั้งแต่เดือนตุลาคม 2540 ถึง กันยายน 2551 ใช้ค่าผลบวกกำลังสองของความคลาดเคลื่อน SSE ในการวัดประสิทธิภาพของค่าพยากรณ์เป็นเกณฑ์ในการเลือกวิธีพยากรณ์ สรุปได้วิธีที่เหมาะสมกับการพยากรณ์ราคาถั่วเหลืองคือ โครงข่ายประสาทเทียม และการพยากรณ์ราคาปาล์มน้ำมันคือวิธีของบ๊อกซ์-เจนกินส์ โดยมีรูปแบบคือ ARIMA(0,1,1)(1,0,1)12

5. พงษ์ศิริ ศิริพานิช [38] ได้ทำการศึกษา การพยากรณ์อนุกรมเวลาแบบผสมระหว่าง ARIMA และเครือข่ายประสาทเทียม ในการพยากรณ์หุ้นในตลาดหลักทรัพย์โดยใช้ข้อมูลของบริษัท ปตท.จำกัด(มหาชน)และธนาคารกรุงเทพ จำกัด พบว่าการพยากรณ์ข้อมูลราคาปิดของหุ้น PTT ด้วยตัวแบบผสมระหว่าง ARIMA และANN สามารถให้ค่าพยากรณ์ในอนาคตระยะสั้นได้ถูกต้องมากกว่าวิธีตัวแบบ ARIMA และตัวแบบ ANN แต่สำหรับการพยากรณ์ค่าระยะยาว ตัวแบบ ANN ให้ค่าพยากรณ์ได้ถูกต้องมากที่สุด แต่ในการพยากรณ์ข้อมูลราคาปิดของหุ้น BBL พบว่าตัวแบบ ANN ให้ค่าพยากรณ์ได้ถูกต้องมากที่สุดทั้งในระยะสั้นและระยะยาว

6. อรรคพล เซี่ยงโหล [39] ได้ทำการศึกษาการเปรียบเทียบความแม่นยำในการพยากรณ์รูปแบบอนุกรมเวลากับเทคนิคการพยากรณ์แบบเป็นเหตุเป็นผล กรณีศึกษาบริษัทผู้ผลิตรถกระบะแห่งหนึ่งในประเทศไทย งานวิจัยนี้เป็นการศึกษาการพยากรณ์ปริมาณยอดขายของบริษัทผู้ผลิตรถกระบะแห่งหนึ่งในประเทศไทยเพื่อทำการวางแผนทางธุรกิจของทางบริษัทและปรับปรุงข้อมูลที่ใช้ในการวางแผนทางธุรกิจให้ดีขึ้น ซึ่งเวลาในการทำวิจัยนี้ใช้ข้อมูลปริมาณการขายโดยรวบรวมข้อมูลตั้งแต่ พ.ศ. 2549 ถึง พ.ศ. 2552 จากนั้นผู้วิจัยทำการเลือกรูปแบบเทคนิคและวิธีการที่เหมาะสมกับลักษณะ

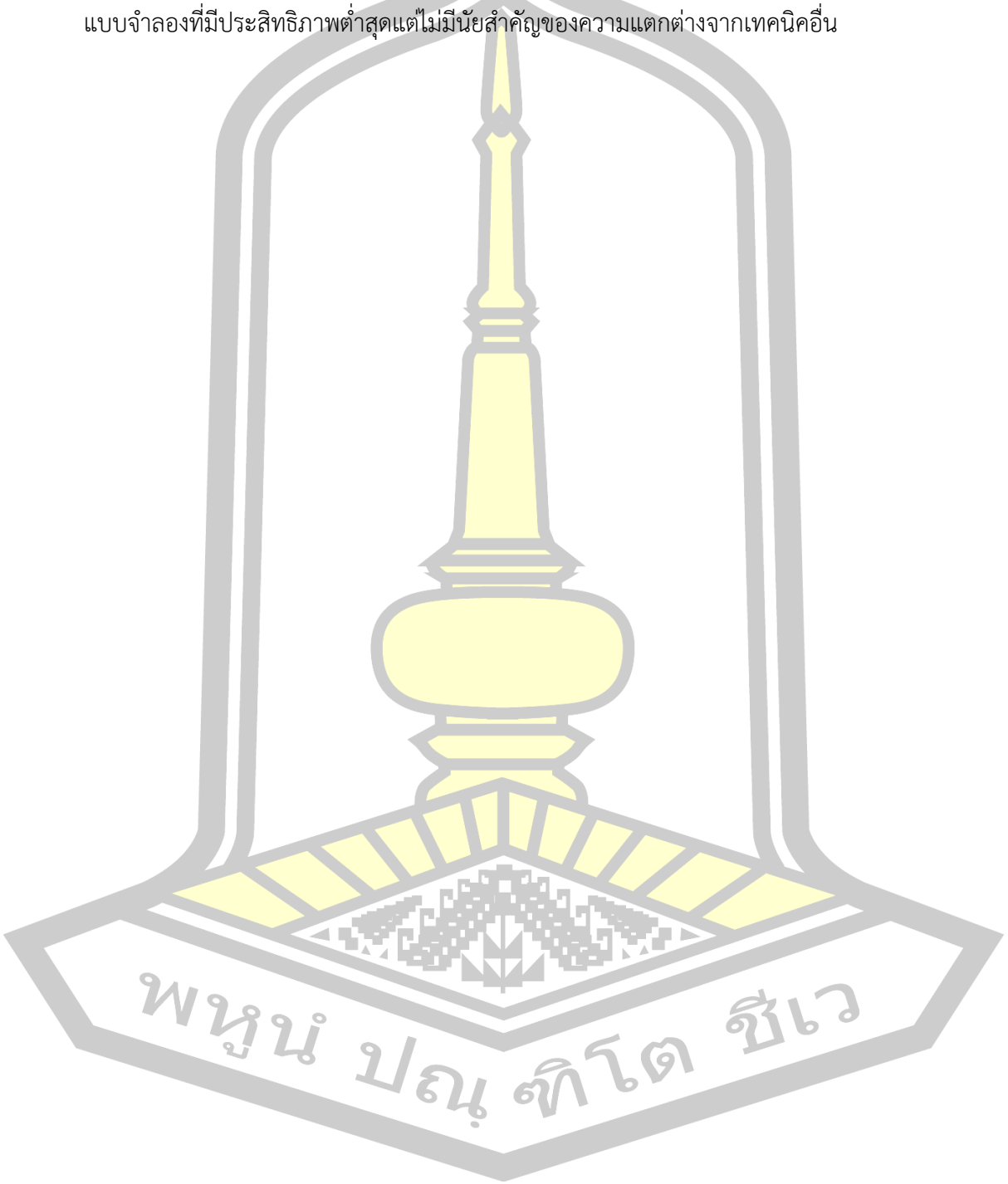
ข้อมูลโดยพิจารณาเปรียบเทียบความคลาดเคลื่อนของวิธีการพยากรณ์การหาค่าเฉลี่ยความผิดพลาดร้อยละสัมบูรณ์ (Mean absolute percentage error,MAPE ผลที่ได้จากการวิจัยพบว่า วิธีปรับเรียบแบบเอกซ์โพเนนเชียลให้ค่าคลาดเคลื่อนน้อยที่สุด

7. ทัศนัย พลอยสุวรรณ [40] ได้ศึกษา การพยากรณ์หาค่าความต้องการไฟฟ้าสูงสุด(Peak Load) ระหว่าง ค.ศ. 2011 ถึง ค.ศ. 2012 ของการไฟฟ้าฝ่ายผลิตแห่งประเทศไทย โดยใช้กระบวนการเกาส์เซี่ยน (Gussain Process) ซึ่งใช้ข้อมูลระหว่างปี ค.ศ. 2000 ถึง ค.ศ. 2011 เป็นข้อมูลฝึกสอน โยประกอบด้วย 4 ตัวแปร ได้แก่ เวลาเป็นรายเดือน,ค่า Peak load เป็นรายเดือน,ค่าผลิตภัณฑ์มวลรวมในประเทศ (GDP) และ ผลิตภัณฑ์ประชาชาติ (GNP) ผลการทดลองพบว่า การพยากรณ์โดยใช้กระบวนการเกาเซี่ยนให้ค่าความผิดพลาดที่น้อยกว่าและให้ประสิทธิภาพที่ดีกว่า การใช้โครงข่ายประสาทเทียม (Neural Networks)

8. พิณสุดา สิงไชยและ พงษ์ศักดิ์ กีรติวินทกร [41] ได้ศึกษาเรื่อง การพยากรณ์ความต้องการพลังงานไฟฟ้าแบบระยะสั้นโดยเทคนิค SVM-R ร่วมกับฟังก์ชันเรเดียลเบซิส (Radial BasisFunction: RBF) ซึ่งเป็นฟังก์ชันเคอร์แนล (Kernel Function)ใน SVM ในงานวิจัยจะใช้ข้อมูลปริมาณการใช้พลังงานไฟฟ้าที่เกิดจากการใช้งานจริงเป็นราย 1 นาทีในรูปแบบของอนุกรมเวลา โดยแบ่งออกเป็นกลุ่มโรงงานอุตสาหกรรม, กลุ่มอาคารพาณิชย์ และกลุ่มบ้านพักอาศัยโดยเลือกใช้ LibSVM ซึ่งเป็นไลบรารี่ SVM มาช่วยในการพยากรณ์ความต้องการพลังงานไฟฟ้าด้วยเทคนิค SVM-R นั้นง่ายขึ้นได้ทำการศึกษา การพยากรณ์ความต้องการพลังงานไฟฟ้ารายา 1 นาที ของกลุ่มโรงงานอุตสาหกรรม กลุ่มอาคารพาณิชย์ และกลุ่มบ้านพักอาศัยด้วยเทคนิค SVM-R พบว่าการพยากรณ์ของกลุ่มอาคารพาณิชย์จะมีค่าใกล้เคียงกับปริมาณการใช้พลังงานไฟฟ้าจริงมากที่สุด เมื่อเทียบค่าของ R2 ของกลุ่มโรงงานอุตสาหกรรมและกลุ่มบ้านพัก

9. Choubey และคณะ [31] ได้ศึกษาแบบจำลอง SMOreg ซึ่งเป็นเทคนิคหนึ่งในตระกูลเทคนิคเครื่องเวกเตอร์การสนับสนุน (Support Vector Machine: SVM) เพื่อพยากรณ์ปริมาณน้ำที่ไหลบ่า และปริมาณน้ำฝน ช่วงละสี่เดือนของแม่น้ำ Narmada จังหวัด Madhya Pradesh ประเทศอินเดีย ข้อมูลที่นำมาใช้จากปี ค.ศ. 1975 ถึง ค.ศ. 2010 ในการทดลองคณะวิจัยได้โดยแบ่งข้อมูลออกเป็นสองชุด คือ ชุดสอน (Training) และชุดทดสอบ (Testing) โดย ชุดสอนเป็นข้อมูลจากปี ค.ศ. 1975 ถึง ค.ศ. 2000 ส่วนชุดทดสอบใช้ข้อมูลจากปี ค.ศ. 2001 ถึง ค.ศ. 2010 ในการวัดประสิทธิภาพการทำงานของแบบจำลองเขาได้ ใช้ Mean Absolute error (MAE), Root Mean Square Error (RMSE), Relative Absolute Error (RAE) และ Root Relative Square Error (RRSE) ผลการการทดสอบพบว่า แบบจำลอง SMOreg มี มีค่า RMSE เท่ากับ 2.3731 ค่า RAE เท่ากับ 65.28% และค่า RRSE เท่ากับ 62.491% ซึ่งเป็นค่าที่เปรียบเทียบกับค่าจริง จะเห็นว่าความแตกต่างระหว่างค่าที่พยากรณ์และค่าจริงมีความแตกต่างกันมาก แต่ไม่มีนัยสำคัญ

10. Zainal NA และ Mustaffa [42] ได้นำเอาเทคนิค Artificial Neural Network มาใช้ในสร้างแบบจำลองเพื่อการพยากรณ์ราคาทองคำ โดยใช้ให้ค่า RMSE และ MAE ในการแสดงประสิทธิภาพของแบบจำลอง ผลปรากฏว่าเทคนิค Artificial Neural Network สามารถสร้างแบบจำลองที่มีประสิทธิภาพต่ำสุดแต่ไม่มีนัยสำคัญของความแตกต่างจากเทคนิคอื่น



บทที่ 3
วิธีการดำเนินวิจัย

งานวิจัยนี้เพื่อศึกษาค้นคว้าและการสร้างแบบจำลองเพื่อพยากรณ์การเกิดอุบัติเหตุบนท้องถนนโดยใช้เทคนิคอนุกรมเวลา กรณีศึกษาของจังหวัดขอนแก่น ดังนั้น ในบทนี้แสดงถึงวิธีดำเนินการวิจัย ประกอบด้วย การศึกษาและวิเคราะห์ข้อมูล การเตรียมข้อมูล การสร้างแบบจำลอง และการวัดประสิทธิภาพของแบบจำลอง ดังต่อไปนี้

3.1 การศึกษาและวิเคราะห์ข้อมูล

การเตรียมข้อมูลของงานวิจัยนี้ข้อมูลได้จากงานจราจรสถานีตำรวจภูธรจังหวัดขอนแก่น โดยใช้ข้อมูลย้อนหลังตั้งแต่ เดือนมกราคม 2554 ถึง มิถุนายน 2560 จำนวน 7 ปี รวมทั้งหมด 78 เดือน การเตรียมข้อมูล เพื่อให้ได้รูปแบบการทำเหมืองข้อมูล ซึ่งทำการแปลงข้อมูลเพื่อให้เข้าสู่โปรแกรม Weka ที่จะมีการจัดเก็บข้อมูลการเกิดอุบัติเหตุบนท้องถนนของอำเภอเมืองขอนแก่น โดยแยกเป็นรายเดือน ที่มีข้อมูลการเกิดอุบัติเหตุจากยานพาหนะ มูลค่าความเสียหายที่ขึ้น ปริมาณการบาดเจ็บของเพศชายและหญิง โดยข้อมูลทั้งหมดของจัดเก็บข้อมูลในรูปแบบกระดาษ ดังนั้นผู้วิจัยได้จัดทำการแบ่งค่าข้อมูลเป็น Attribute โดยแต่ละ Attribute จะมีค่าข้อมูลของแต่ละข้อมูลที่ได้จากงานจราจรสถานีตำรวจภูธรจังหวัดขอนแก่นตั้งแต่ปี 2554 ดังภาพประกอบที่ 3.1

(cgac9010) จำนวนคดีอุบัติเหตุจราจรทางบกจำแนกตามรายละเอียดและรายเดือน ตำรวจภูธรจังหวัดขอนแก่น ตำรวจภูธรภาค 4 หน้าที่ 1

มกราคม 2554 - ธันวาคม 2554 วันที่ 31 ต.ค. 2556

เดือน	รับแจ้ง (ราย)	คนเดินเท้า	รถจักรยาน	รถสามล้อ	รถจักรยานยนต์	รถสามล้อเครื่อง	รถยนต์นั่ง	รถโดยสารเล็ก(ตู้)	รถบรรทุกเล็ก (ปิกอัพ)	รถโดยสารขนาดใหญ่	รถบรรทุก 6ล้อ	รถบรรทุก 10ล้อ	รถอีแต๋น	รถแท็กซี่	รถอื่นๆ	มูลค่าทรัพย์สินเสียหายรวม (บาท)	ความเสียหายที่เกิดขึ้นกับบุคคล						ผู้ต้องหา				
																	ตาย		บาดเจ็บสาหัส		บาดเจ็บเล็กน้อย		จับกุม		หลบหนี		ไม่รู้ตัว
																	ชาย	หญิง	ชาย	หญิง	ชาย	หญิง	ชาย	หญิง	ชาย	หญิง	
มกราคม	47	7	0	0	10	0	5	1	9	2	3	1	0	0	2	2,206,440	26	8	14	7	20	13	36	6	1	0	3
กุมภาพันธ์	41	5	1	0	14	0	6	0	0	0	0	1	0	0	2	631,030	16	2	5	2	9	6	26	5	0	0	0
มีนาคม	43	2	2	0	18	0	3	0	6	0	4	2	0	1	4	617,000	15	7	1	2	10	5	26	4	4	1	0
เมษายน	43	5	3	0	12	0	1	0	6	0	1	0	0	0	2	1,261,600	22	4	4	4	8	1	29	2	0	0	0
พฤษภาคม	45	3	0	0	9	0	3	1	6	0	2	0	0	0	1	616,000	10	4	5	3	6	2	23	5	0	1	1
มิถุนายน	43	2	0	0	10	0	2	0	5	1	0	1	0	0	2	634,500	16	4	7	8	3	5	24	5	1	0	1
กรกฎาคม	22	1	0	0	9	0	3	0	8	0	1	0	0	0	1	505,750	6	2	4	3	5	5	13	1	1	0	1
สิงหาคม	20	0	0	0	9	0	3	1	3	0	1	0	0	0	0	57,800	7	2	3	3	3	1	12	0	1	0	1
กันยายน	23	1	0	0	9	0	1	0	7	1	0	0	0	0	1	107,800	9	2	5	0	1	0	21	1	0	0	0
ตุลาคม	37	2	0	0	8	0	1	0	3	1	4	0	0	0	1	268,600	15	7	1	2	15	6	25	2	0	0	0
พฤศจิกายน	38	5	0	0	16	0	1	1	10	0	3	1	0	0	1	296,500	15	6	10	1	9	4	27	5	1	0	0
ธันวาคม	49	2	2	0	15	0	8	0	8	0	0	0	0	0	3	1,206,340	20	4	3	2	12	2	23	2	1	0	0
รวม	451	35	8	0	139	0	37	4	71	5	19	6	0	1	20	8,409,360	177	52	62	37	101	50	285	38	10	2	7

ภาพประกอบที่ 3.1 ตัวอย่างข้อมูลการเกิดอุบัติเหตุบนท้องของจังหวัดขอนแก่น

Month	Pedes	Bicycl	Tricyc	Motor	Tuk tu	Private	Van	Picku	Bus	Truck (	Heavy	Farm	Taxi	Other	Property	M	F	D_M	D_F	S_M	S_F	H_M	H_F
1	7	0	0	10	0	5	1	9	2	3	1	0	0	2	2206440	60	28	26	8	14	7	20	13
2	5	1	0	14	0	6	0	0	0	0	1	0	0	2	631030	30	10	16	2	5	2	9	6
3	2	2	0	18	0	3	0	6	0	4	2	0	1	4	617000	26	14	15	7	1	2	10	5
4	5	3	0	12	0	1	0	6	0	1	0	0	0	2	1261600	34	9	22	4	4	4	8	1
5	3	0	0	9	0	3	1	6	0	2	0	0	0	1	616000	21	9	10	4	5	3	6	2
6	2	0	0	10	0	2	0	5	1	0	1	0	0	2	634500	26	17	16	4	7	8	3	5
7	1	0	0	9	0	3	0	8	0	1	0	0	0	1	505750	15	10	6	2	4	3	5	5
8	0	0	0	9	0	3	1	3	0	1	0	0	0	0	57800	13	6	7	2	3	3	3	1
9	1	0	0	9	0	1	0	7	1	0	0	0	0	1	107800	15	2	9	2	5	0	1	0
10	2	0	0	8	0	1	0	3	1	4	0	0	0	1	268600	31	15	15	7	1	2	15	6
11	5	0	0	16	0	1	1	10	0	3	1	0	0	1	296500	34	11	15	6	10	1	9	4
12	2	2	0	15	0	8	0	8	0	0	0	0	0	3	1206340	35	8	20	4	3	2	12	2
13	5	0	0	11	0	3	0	2	3	0	0	0	0	2	859050	20	6	11	4	4	0	5	2
14	0	0	0	11	0	8	0	3	0	0	1	0	0	0	85000	20	10	11	7	6	3	3	0
15	2	0	0	14	0	1	1	6	0	1	1	0	1	1	345700	26	17	15	8	3	3	8	6
16	5	0	0	6	0	0	0	4	1	0	0	0	0	3	324000	24	24	15	16	3	4	6	4
17	2	0	0	8	0	3	0	1	0	1	1	0	0	1	177200	12	6	3	3	4	2	5	1
18	3	1	0	11	0	2	0	3	0	0	1	0	0	5	140510	18	10	13	6	2	1	3	3
19	2	0	0	14	0	5	0	7	0	1	0	0	0	0	0	20	6	14	3	3	1	3	2
20	3	0	0	3	0	1	0	3	2	0	1	0	0	1	341500	18	11	13	2	2	2	3	7
21	1	0	0	4	1	3	0	1	1	1	2	0	0	0	177221	8	2	7	1	1	1	0	0
22	4	0	0	5	0	2	0	4	0	0	0	0	0	0	0	18	8	11	4	3	3	4	1
23	4	0	0	5	0	2	0	4	0	0	0	0	0	0	0	10	4	6	2	2	1	2	1

ภาพประกอบที่ 3.2 Data transformation

จากภาพประกอบที่ 3.2 คือ ข้อมูลจากงานจราจรสถานีตำรวจภูธรจังหวัดขอนแก่น ได้จัดเก็บข้อมูลปริมาณการเกิดอุบัติเหตุเป็นรายเดือนในแต่ละปี จากนั้นผู้วิจัยจึงได้รวบรวมข้อมูลแต่ละเดือนเข้าด้วยกัน โดยแบ่งข้อมูล แต่ละเดือนแบ่งตามการประสบเหตุ เพศ การบาดเจ็บ และมูลค่าความเสียหาย ในขั้นการแปลงข้อมูล (Data transformation)

3.2 การทำกระบวนการก่อนการสร้างแบบจำลอง

การทำกระบวนการก่อนการสร้างแบบจำลอง โดยการแปลงข้อมูล (Data transformation) จากข้อมูลที่ได้จากงานจราจรสถานีตำรวจภูธรจังหวัดขอนแก่น ดังภาพประกอบที่ 3.2 จัดเก็บข้อมูลในรูปแบบ Excel ซึ่งประกอบด้วยความสูญเสียที่เกิดขึ้นจากยานพาหนะ มูลค่าความเสียหายที่ขึ้น เพศ ความสูญเสียที่เกิดขึ้น โดยแบ่งข้อมูลเป็นตัวแปรตามชื่อของดังตารางที่ 3.1 ในงานวิจัยนี้ข้อมูลประกอบด้วยตัวแปรทั้งหมด 25 ตัวแปร และสามารถแบ่งตามความเสียหายที่เกิดขึ้นได้ โดยแยกตัวแปรเป็นทั้งหมด 5 ประเภท คือ

ตารางที่ 3.1 คำอธิบายตัวแปรและชนิดของตัวแปร

ชื่อตัวแปร	คำอธิบาย	ชนิดของตัวแปร
Month	เดือนที่เกิดอุบัติเหตุ	Number
ยานพาหนะ		
Pedestrian	จำนวนของคนเดินเท้าที่เกิดอุบัติเหตุ	Number
Bicycle	จำนวนรถจักรยานที่เกิดอุบัติเหตุ	Number
Tricycle	จำนวนรถสามล้อที่เกิดอุบัติเหตุ	Number
Motorcycle	จำนวนรถมอเตอร์ไซค์ที่เกิดอุบัติเหตุ	Number

ตารางที่ 3.1 คำอธิบายตัวแปรและชนิดของตัวแปร (ต่อ)

ชื่อตัวแปร	คำอธิบาย	ชนิดของตัวแปร
Tuk tuk	จำนวนรถสามล้อเครื่องที่เกิดอุบัติเหตุ	Number
Private car	จำนวนรถยนต์ที่เกิดอุบัติเหตุ	Number
Van	จำนวนรถตู้ที่เกิดอุบัติเหตุ	Number
Pickuptruck	จำนวนรถบรรทุกเล็กที่เกิดอุบัติเหตุ	Number
Bus	จำนวนรถบัสที่เกิดอุบัติเหตุ	Number
Truck (6 wheel)	จำนวนรถบรรทุก 6 ล้อที่เกิดอุบัติเหตุ	Number
Heavy truck	จำนวนรถบรรทุกขนาดใหญ่ที่เกิดอุบัติเหตุ	Number
Farm vehicle	จำนวนรถการเกษตรที่เกิดอุบัติเหตุ	Number
Taxi	จำนวนรถแท็กซี่ที่เกิดอุบัติเหตุ	Number
Other	จำนวนรถอื่นๆที่เกิดอุบัติเหตุ	Number
ความเสียหายที่เกิดขึ้น		
Value (Baht)	มูลค่าความเสียหายที่เกิดขึ้น (บาท)	Number
เพศ		
M	จำนวนของผู้ประสบอุบัติเหตุเพศชาย	Number
F	จำนวนของผู้ประสบอุบัติเหตุเพศหญิง	Number
ความสูญเสีย		
D_M	จำนวนของผู้เสียชีวิตเพศชาย	Number
S_M	จำนวนของผู้บาดเจ็บสาหัสเพศชาย	Number
H_M	จำนวนของผู้บาดเจ็บเล็กน้อยเพศชาย	Number
D_F	จำนวนของผู้เสียชีวิตเพศหญิง	Number
S_F	จำนวนของผู้บาดเจ็บสาหัสเพศหญิง	Number
H_F	จำนวนของผู้บาดเจ็บเล็กน้อยเพศหญิง	Number

ในการแบ่งตัวแปรจากการสูญเสียทั้งการเสียชีวิต บาดเจ็บสาหัส และบาดเจ็บเล็กน้อย โดยใช้ค่าจริงที่เกิดขึ้น

3.3 การสร้างแบบจำลอง

ในการสร้างแบบจำลองให้ได้แบบจำลองที่มีประสิทธิภาพดีที่สุด การสร้างแบบจำลองเพื่อพยากรณ์การเกิดอุบัติเหตุบนท้องถนนโดยใช้เทคนิคอนุกรมเวลา จะใช้เทคนิคในการทำเหมืองข้อมูลร่วมกับด้วยการ วิเคราะห์อนุกรมเวลา (Time series analysis) ด้วยเทคนิคการพยากรณ์เชิงปริมาณ ข้อมูลที่นำมาใช้การพยากรณ์การเกิดอุบัติเหตุบนท้องถนน โดยประกอบด้วย เพศ ยานพาหนะ มูลค่าทรัพย์สินที่เสียหาย ความเสียหายที่เกิดขึ้นกับบุคคล เดือน และปีที่เกิดอุบัติเหตุ โดยนำข้อมูลจาก

สถานีตำรวจจังหวัดขอนแก่น ใช้ข้อมูลตั้งแต่เดือนมกราคม 2554 ถึง มิถุนายน 2560 จำนวน 7 ปี รวมทั้งหมด 78 เดือน จากนั้นนำเข้า โปรแกรม WEKA 3.9.1 สร้างแบบจำลอง 5 แบบเพื่อใช้เปรียบเทียบแบบจำลองเพื่อพยากรณ์การเกิดอุบัติเหตุบนท้องถนนโดยใช้ แบบจำลองดังนี้

1. การถดถอยเชิงเส้น (linear regression: LR)
2. โครงข่ายประสาทเทียม (Artificial Neural Networks: ANN แบบ MLP)
3. Support Vector Machine for Regression (SVMR)
4. Sequential Minimal Optimization for Regression (SMOreg)
5. Gussian Process (GP)

3.3.1 การทำงานของแบบจำลองการถดถอยเชิงเส้น (linear regression: LR) การทำงานของแบบจำลอง LR จะรับค่า x คือ ข้อมูลคูณกับ b โดยค่า b คือความชัน (Slope) และบวกกับค่า a โดยที่ a ก็คือค่าคงที่ (Constant) เป็นค่าที่เส้นกราฟถดถอยตัดกับแกน Y ส่วนของเส้นกราฟซึ่งแสดงถึงอัตราการเปลี่ยนแปลงของ Y เมื่อ X เปลี่ยนแปลงเรียกส่วนนี้ว่าสัมประสิทธิ์การถดถอย (Regression Coefficient) หรือสัมประสิทธิ์การพยากรณ์สมการที่ได้คือ

$$Y = a + bX \qquad (3.1)$$

ในงานวิจัยนี้ผู้วิจัยได้ใช้แบบจำลอง ในการสร้าง แบบจำลองเพื่อวัด ประสิทธิภาพการทำงานแบบจำลองเพื่อ พยากรณ์การเกิดอุบัติเหตุบนท้องถนนโดยใช้แบบจำลองการถดถอยเชิงเส้น (linear regression: LR) ในงานวิจัยนี้ผู้วิจัยได้ใช้โปรแกรม Weka 3.9.1 ในการใช้งานของโปรแกรมผู้วิจัยได้ตั้งค่าการใช้งาน ของแบบจำลองการถดถอยเชิงเส้น (linear regression: LR) ในโปรแกรม Waka 3.9.1 เป็นค่าเริ่มต้น

พารามิเตอร์ของเทคนิคการถดถอยเชิงเส้น Linear regression โดยคลาสของเทคนิคการถดถอยเชิงเส้นสาหรับการพยากรณ์ และมีเมนูการตั้งค่าต่อไปนี้

1. attributeSelectionMethod คือการเลือกวิธีการในการทางานของเทคนิค Linear regression ในเมนูนี้เลือกเป็นค่าเริ่มต้นคือ M5 method
2. batchsize คือ จำนวนตัวอย่างการพยากรณ์ที่ต้องการประมวลผลเบื้องต้น
3. debug คือ รายงานการแก้ไขจุดบกพร่องของข้อมูล
4. doNotCheckCapabilities คือ ถ้ามีการตั้งค่าความสามารถในการจาแนกประเภทจะไม่ได้รับการตรวจสอบก่อนสร้างตัวจำแนกประเภท ใช้ด้วยความระมัดระวัง
5. eliminate colinear attributes คือ การจำกัดคุณสมบัติในแนวเดียวกัน
6. minimal คือ ค่าที่ต่าที่สุด
7. numDecimalPlaces คือ จำนวนตำแหน่งทศนิยมที่จะใช้สาหรับผลลัพธ์ของตัวเลขในโมเดล

8. outputAdditionalStats คือ ค่าเบี่ยงเบนมาตรฐานของค่าสัมประสิทธิ์ และค่าสถิติ (t-test)

9. ridge คือ ความคุ้มค่าสูงสุดของพารามิเตอร์

ค่าตัวแปรของพารามิเตอร์ที่นำมาใช้ในงานวิจัย

1. attributeSelectionMethod ใช้ M5 method
2. batchsize จำนวนตัวอย่างพยากรณ์เริ่มต้นที่ 100
3. numDecimalPlaces ใช้ทศนิยม 4 ตำแหน่ง
4. ridge ใช้ค่า 1.0E8

3.3.2 การทำงานของแบบจำลองโครงข่ายประสาทเทียม (Artificial Neural Networks: ANN) การทำงานของ Neural networks คือเมื่อมี input เข้ามายัง network ก็เอา input มาคูณกับ weight ของแต่ละขา ผลที่ได้จาก input ทุก ๆ ขาของ neuron จะเอามารวมกันแล้วก็เอามาเทียบกับ threshold ที่กำหนดไว้ ถ้าผลรวมมีค่ามากกว่า threshold แล้ว neuron ก็จะส่ง output ออกไป output นี้ก็จะถูกส่งไปยัง input ของ neuron อื่น ๆ ที่เชื่อมกันใน network ถ้าค่าน้อยกว่า threshold ก็จะไม่เกิด output เขียนออกมาได้ดังนี้

if (sum(input * weight) > threshold) then output

สมการของแบบจำลองโครงข่ายประสาทเทียมดังนี้

$$n = \sum_{i=1}^{z} x_i w_i + b \qquad (3.2)$$

โดยที่ n คือ ผลรวมที่ได้จากฟังก์ชันผลรวม

xi คือค่าข้อมูลเข้า w

i คือ ค่าน้ำหนักของนิวรอน

z คือ จำนวนนิวรอนของชั้นที่เข้า b คือ ค่าความโน้มเอียง

ลักษณะของโครงข่ายประสาทเทียมสามารถแบ่งได้ 2 แบบ คือ 1) โครงข่ายประสาทเทียมแบบ ชั้นเดียว (single layer) ซึ่งจะมีเพียงชั้นสัญญาณประสาทขาเข้า และชั้นสัญญาณประสาทขาออก เท่านั้น เช่น โครงข่ายเพอเซบตรอนอย่างง่าย (simple perceptron) และโครงข่ายโฮปฟิลด์ (hopfield networks) และ 2) โครงข่ายประสาทเทียมแบบหลายชั้น (multilayer) ซึ่งมีลักษณะเช่นเดียวกับโครงข่ายประสาทเทียมแบบชั้นเดียว แต่จะมีชั้นแอบแฝง (hidden) เพิ่มขึ้น โดยอยู่ส่วนกลางระหว่างชั้นนำข้อมูลป้อนเข้าและชั้นส่งขอมูลออก ทั้งนี้ชั้นแอบแฝงอาจมีมากกว่า 1 ชั้น ในงานวิจัยนี้ผู้วิจัยได้ใช้แบบจำลอง Multi-Layer Perceptron: MLP

ในการสร้างแบบจำลองในการสร้างแบบจำลองเพื่อวัดประสิทธิภาพ การทำงานแบบจำลองเพื่อพยากรณ์การเกิดอุบัติเหตุบนท้องถนนของอำเภอเมืองขอนแก่น โดยใช้แบบจำลองโครงข่ายประสาทเทียม (Artificial Neural Networks: ANN แบบ MLP) ในงานวิจัยนี้ผู้วิจัยได้ใช้โปรแกรม Weka 3.9.1 ในการใช้งานของโปรแกรมผู้วิจัยได้ตั้งค่าการใช้งาน ของแบบจำลองโครงข่ายประสาทเทียม (Artificial Neural Networks: ANN แบบ MLP) ในโปรแกรม Waka 3.9.1 เป็นค่าเริ่มต้น

การแสดงถึงพารามิเตอร์ของเทคนิค โครงข่ายประสาทเทียมแบบเปอร์เซ็ปตรอนหลายชั้น Multi-Layer Perceptron: (MLP) ซึ่งเป็นการจำแนกประเภทที่ใช้ในการแพร่กระจายแล้วย้อนกลับไปยังตัวอย่างการแยกประเภทดังต่อไปนี้

1. GUI คือ หน้าจอติดต่อกับผู้ใช้งาน
2. autoBuild คือ เพิ่มการเชื่อมต่อการประมวลผลในโครงข่ายประสาทเทียม
3. batchsize คือ ตัวเลขที่ต้องการของตัวอย่างในการประมวลผลถ้าเป็นการพยากรณ์จะอยู่ในช่วงดำเนินการ
4. debug คือ รายงานการแก้ไขจุดบกพร่องของข้อมูล
5. decay คือ สาเหตุของการเรียนรู้นี้มีอัตราที่ลดลง
6. doNotCheckCapabilities คือ ถ้ามีการตั้งค่าความสามารถในการจาแนกประเภทจะไม่ได้รับการตรวจสอบก่อนสร้างตัวจำแนกประเภท ใช้ด้วยความระมัดระวัง
7. hiddenlayears คือ กำหนดการประมวลผลของโครงข่ายประสาทเทียม
8. learningRate คือ จำนวนน้ำหนักที่เปลี่ยนไป
9. momentum คือ การประยุกต์การเคลื่อนที่ของน้ำหนักอยู่ในระหว่างการปรับปรุง
10. normalizeTobinaryFillter คือ ก่อนจะประมวลผลกับการกลั่นกรอง
11. normalizeAttributes คือ ทำคุณลักษณะให้เป็นปกติ
12. numDecimalPlaces คือ ที่อยู่ของจานวนทศนิยมสาหรับรูปแบบจานวนการส่งออก
13. reset คือ การอนุญาตของเครือข่ายนี้เพื่อตั้งค่าการเรียนรู้ที่ต่ากว่า
14. seed คือ ผลที่ได้ของผู้ใช้ส่งไปสุ่มจานวนใหม่
15. trainingtime คือ จานวนตัวเลขของการเทรนข้อมูล
16. validationSetsize คือ ขนาดร้อยละของชุดตรวจสอบ
17. validationThreshold คือ การหยุดการตรวจสอบข้อมูลการทดสอบที่ผ่านการทดสอบมาแล้ว

ค่าตัวแปรของพารามิเตอร์ที่นำมาใช้ในงานวิจัย

1. hiddenlayears ค่าโครงข่ายประสาทเทียมเริ่มต้นเป็น
2. learningRate ใช้ค่าน้ำหนักเปลี่ยนทีละ 0.3
3. momentum ใช้ค่าถ่วงน้ำหนัก 0.2
4. normalizeTobinaryFillter ใช้ค่าเป็นTrue
5. normalizeNumericClass ใช้ค่าเป็น True
6. numDecimalPlaces True
7. validationThreshold ใช้ค่า 20

3.3.3 Support Vector Machine for Regression (SVMR) เป็นเทคนิคการจำแนกประเภทของข้อมูลที่มีพื้นฐานมาจากทฤษฎีการเรียนรู้จากสถิติ คล้ายเทคนิคโครงข่ายประสาทเทียม แต่ SVM ใช้หลักการลดค่าความเสี่ยงเชิงโครงสร้างให้ต่ำสุด (Structural Risk Minimization) เพื่อลดค่าความผิดพลาดของการทำนาย (Minimized Error) พร้อมกับเพิ่มระยะการแบ่งแยกให้มากที่สุด (Maximized Margin) หลักการของ SVM คือการหาสัมประสิทธิ์ของสมการเพื่อสร้างเส้นแบ่งแยกประเภทข้อมูล ทำโดยการเลือกเส้นหรือระนาบแบ่งแยกประเภทข้อมูลที่เหมาะสมที่สุด

ในงานวิจัยนี้ผู้วิจัยได้ใช้แบบจำลอง ในการสร้างแบบจำลองเพื่อวัด ประสิทธิภาพการทำงานแบบจำลองเพื่อ พยากรณ์การเกิดอุบัติเหตุบนท้องถนนของอำเภอเมืองขอนแก่น โดยใช้แบบจำลอง Support Vector Machine Regression (SVMR) ในการใช้งานของโปรแกรมผู้วิจัยได้ตั้งค่าการใช้งาน

การแสดงถึงพารามิเตอร์ของเทคนิค Support Vector Machine for Regression (SVMR) ซึ่งเป็นการจำแนกประเภทที่ใช้ในการแพร่กระจายแล้วย้อนกลับไปยังตัวอย่างการแยกประเภทดังต่อไปนี้

1. SVMType คือ ประเภทของเทคนิค SVM
2. batchsize คือ ตัวเลขที่ต้องการของตัวอย่างในการประมวลผลถ้าเป็นการพยากรณ์จะอยู่ในช่วงดำเนินการ
3. cascheSize คือ ขนาดความจุมีหน่วยเป็น Megabyte
4. coef0 คือ ค่าสัมประสิทธิ์ที่ใช้
5. cost คือ ค่าตัวแปร C สำหรับ C-SVC ,epsilon-SVR, and nu-SVR
6. debug คือ รายงานการแก้ไขจุดบกพร่องของข้อมูล
7. degree คือ ระดับ
8. doNotCheckCapabilities คือ ถ้ามีการตั้งค่าความสามารถในการจำแนกประเภทจะไม่ได้รับการตรวจสอบก่อนสร้างตัวจำแนกประเภท ใช้ด้วยความระมัดระวัง

9. doNotReplaceMissingValues คือ หยุดการทางานอัตโนมัติแทนที่ค่าที่หายไป

10. asp คือ เกณฑ์ที่สิ้นสุด

11. gamma คือ แกมมาที่ต้องใช้

12. KernalType คือ ประเภทของ Kernal

13. loss คือ epsilon สาหรับการสูญเสียฟังก์ชันใน epsilon

14. normalize คือ การทาให้ข้อมูลเป็นปกติ

15. nu คือ ค่าของ nu สาหรับ nu-SCV, one-class SVM และ nu-SVR

16. numDecimalPlaces คือ ที่อยู่ของจานวนทศนิยมสาหรับรูปแบบจำนวนการส่งออก

17. probabilityEstimates คือ การสร้างความน่าจะเป็นหรือประมาณการแทนด้วย -1/+1 สำหรับการจัดหมวดหมู่ของปัญหาที่เกิดขึ้น

18. seed คือ ผลที่ได้ของผู้ใช้ส่งไปสุ่มจำนวนใหม่

19. shrinking คือ ลดการแก้ปัญหา

20. weights คือ น้าหนักที่ใช้สาหรับเรียนรู้รายการที่ว่างเปล่า (ซึ่งเป็นค่าเริ่มต้นของการใช้งาน)

ค่าตัวแปรของพารามิเตอร์ที่นำมาใช้ในงานวิจัย

1. SVMType ใช้ค่า nu-SVR(regression)
2. cost ใช้ค่า 1.0
3. weights ใช้ค่าเริ่มต้นคือค่าว่าง

3.3.4 Sequential Minimal Optomization for Regression(SMOreg) เป็นอัลกอริทึมสำหรับทำการสอน SVM ในกรณีปัญหาพหุนามกำลังสองใหญ่ถูกแตกออกมาเป็นชุดปัญหาขนาดเล็กซึ่งจะสามารถถูกแก้ได้ด้วย SMO และสามารถจัดการกับชุดข้อมูลที่ทำการ Training ที่มีชุดข้อมูลขนาดใหญ่โดยใช้เวลาน้อยลง ปัญหาคู่ (Dual problem) สามารถถูกแก้ไขปัญหาแทนที่ของ L_P โดย L_D เป็นผลลัพธ์โดยการแทนที่ L_P ซึ่งแสดงได้ดังสมการ

$$L_d = \sum_i a_i - \frac{1}{2}\sum_i a_i \, a_j \, y_i y_j x_i \cdot x_j \qquad (3.3)$$

ในงานวิจัยนี้ผู้วิจัยได้ใช้แบบจำลอง ในการสร้างแบบจำลองเพื่อวัด ประสิทธิภาพการทำงานแบบจำลองเพื่อ พยากรณ์การเกิดอุบัติเหตุบนท้องถนนของอำเภอเมืองขอนแก่น โดยใช้

แบบจำลอง Sequential Minimal Optimization for Regression (SMOreg) ในการใช้งานของโปรแกรมผู้วิจัยได้ตั้งค่าการใช้งานเริ่มต้น

การแสดงพารามิเตอร์ของเทคนิค Sequential Minimal Optimization for Regression (SMOreg) ซึ่งเทคนิค (SMOreg) มีเครื่องมือเวกเตอร์สาหรับสนับสนุนเทคนิคการถดถอยเชิงเส้นมีค่าเริ่มต้นการทางานดังต่อไปนี้

1. batchsize คือ ตัวเลขที่ต้องการของตัวอย่างในการประมวลผลถ้าเป็นการพยากรณ์จะอยู่ในช่วงดาเนินการ

2. C คือ ระดับความซับซ้อนของพารามิเตอร์ C

3. debug คือ รายงานการแก้ไขจุดบกพร่องของข้อมูล

4. doNotCheckCapabilities คือ ถ้ามีการตั้งค่าความสามารถในการจำแนกประเภทจะไม่ได้รับการตรวจสอบก่อนสร้างตัวจาแนกประเภท ใช้ด้วยความระมัดระวัง

5. filterType คือ การกำหนดข้อมูลอย่างไรให้มีการเปลี่ยนแปลง

6. Kernel คือ ประเภทของ kernel ที่ใช้

7. numDecimalPlaces คือ ที่อยู่ของจำนวนทศนิยมสำหรับรูปแบบจำนวนการส่งออก

8. regOptimizer คือ ขั้นตอนวิธีกระบวนการเรียนรู้

ค่าตัวแปรของพารามิเตอร์ที่นำมาใช้ในงานวิจัย

1. C ใช้ค่าความซับซ้อนของที่ 1.0

2. Kernel ใช้ PolyKernel –E1.0-C250007

3. numDecimalPlaces ใช้ค่าทศนิยม 2 ตำแหน่ง

4. regOptimizer ใช้ค่าการเรียนรู้ที่ RegSMOimproved

3.3.5 Guassian Process (GP)

ในงานวิจัยนี้ผู้วิจัยได้ใช้แบบจำลอง ในการสร้างแบบจำลองเพื่อวัด ประสิทธิภาพการทำงานแบบจำลองเพื่อ พยากรณ์การเกิดอุบัติเหตุบนท้องถนนของอำเภอเมืองขอนแก่น โดยใช้แบบจำลอง Guassian Process (GP) ในงานวิจัยนี้ผู้วิจัยได้ใช้โปรแกรม Weka 3.9.1 ในการใช้งานของโปรแกรมผู้วิจัยได้ตั้งค่าการใช้งาน ของแบบจำลอง Guassian Process (GP) ในโปรแกรม Waka 3.9.1 เป็นค่าเริ่มต้น

การแสดงพารามิเตอร์ของเทคนิค Guassian Process ซึ่งเทคนิค (GP) มีเครื่องมือเวกเตอร์สำหรับสนับสนุนเทคนิคการถดถอยเชิงเส้นมีค่าเริ่มต้นการทำงานดังต่อไปนี้

1. batchsize คือ ตัวเลขที่ต้องการของตัวอย่างในการประมวลผลถ้าเป็นการพยากรณ์จะอยู่ในช่วงดำเนินการ

2. debug คือ รายงานการแก้ไขจุดบกพร่องของข้อมูล

3. doNotCheckCapabilities คือ ถ้ามีการตั้งค่าความสามารถในการจาแนกประเภทจะไม่ได้รับการตรวจสอบก่อนสร้างตัวจาแนกประเภท ใช้ด้วยความระมัดระวัง

4. filterType คือ การกำหนดข้อมูลอย่างไรให้มีการเปลี่ยนแปลง

5. Kernel คือ ประเภทของ kernel ที่ใช้

6. noise คือ ข้อมูลจริงที่นำมาใช้

7. numDecimalPlaces คือ ที่อยู่ของจำนวนทศนิยมสาหรับรูปแบบจำนวนการส่งออก

8. seed คือ ผลที่ได้ของผู้ใช้ส่งไปสุ่มจำนวนใหม่

ค่าตัวแปรของพารามิเตอร์ที่นำมาใช้ในงานวิจัย

1. filterType ใช้ค่า Normalize Training data
2. Kernel ใช้ PolyKernel –E1.0-C250007
3. noise ใช้ค่า 1.0 จากข้อมูลจริง

3.4 การวัดประสิทธิภาพ

การวิเคราะห์ประสิทธิภาพแบบจำลองเพื่อพยากรณ์การเกิดอุบัติเหตุบนท้องถนน ในการวัดประสิทธิภาพของแบบจำลองในงานวิจัยนี้ผู้วิจัยได้ใช้หลักการหน้าต่างเคลื่อนที่ (Sliding Window) คือ หลักการแบ่งข้อมูลออกเป็น ชุดข้อมูลที่ใช้ในการฝึกสอน หรือการเรียนรู้ (Training data) และชุดข้อมูลที่ใช้ในการทดสอบ (Testing data) รอบแรกข้อมูล 3 ปีแรก (พ.ศ. 2554 ถึง พ.ศ. 2556) เป็นชุดข้อมูลฝึกสอน และให้ 6 เดือนเป็นชุดข้อมูลทดสอบ (เดือนมกราคม ถึง มิถุนายน พ.ศ. 2560) แล้วจึงทำการเคลื่อนหน้าต่างข้อมูลไปรอบละ 6 เดือนรวมทั้งสิ้น 7 รอบ

มีการวัดค่าประสิทธิภาพของแบบจำลอง ด้วยค่าเฉลี่ยความคลาดเคลื่อนสัมบูรณ์ (Mean Absolute error: MAE) และค่าความคลาดเคลื่อนเฉลี่ยกำลังสอง (Root Mean Square Error: RMSE) ในการแสดงประสิทธิภาพในการทำพยากรณ์ของแต่ละเทคนิค แล้วทำการเลือกเทคนิคที่เหมาะสมกับข้อมูล มาทำการพยากรณ์การเกิดอุบัติเหตุในอนาคตอีก 3 ปี ซึ่งส่งผลให้การพยากรณ์มีความแม่นยำสูงสุด ดังภาพประกอบที่ 3.3

ชุดข้อมูลฝึกสอน	ชุดข้อมูลทดสอบ	
ม.ค. 54 – ธ.ค. 56	ม.ค. 57- มิ.ย. 57	รอบที่ 1

....	ชุดข้อมูลฝึกสอน	ชุดข้อมูลทดสอบ	
ม.ค.54	ก.ค. 54 – มิ.ย. 57	ก.ค. 57- ธ.ค. 57	รอบที่ 2

....	ชุดข้อมูลฝึกสอน	ชุดข้อมูลทดสอบ	
ก.ค. 54	ม.ค. 55 – ธ.ค. 57	ม.ค. 58- มิ.ย. 58	รอบที่ 3

....	ชุดข้อมูลฝึกสอน	ชุดข้อมูลทดสอบ	
ม.ค. 55	ก.ค. 55 – มิ.ย. 58	ก.ค. 57- ธ.ค. 57	รอบที่ 4

....	ชุดข้อมูลฝึกสอน	ชุดข้อมูลทดสอบ	
ก.ค. 55	ก.ค. 55 – ธ.ค. 58	ก.ค. 58- ธ.ค. 58	รอบที่ 5

....	ชุดข้อมูลฝึกสอน	ชุดข้อมูลทดสอบ	
ม.ค. 56	ก.ค. 56 – มิ.ย.59	ก.ค.59- ธ.ค.59	รอบที่ 6

....	ชุดข้อมูลฝึกสอน	ชุดข้อมูลทดสอบ	
ก.ค. 56	ม.ค. 57– ธ.ค.59	ม.ค. 60- มิ.ย.60	รอบที่ 7

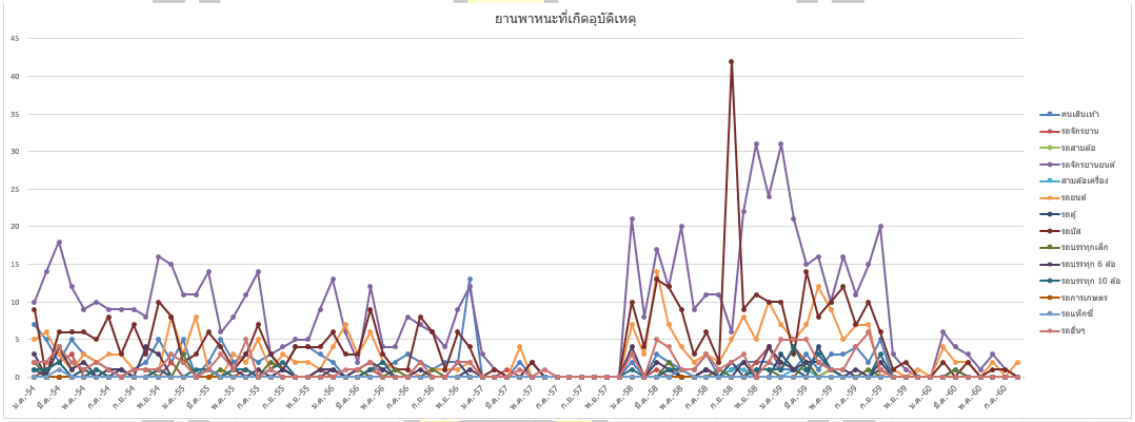
ภาพประกอบที่ 3.3 ตัวอย่างหลักการ Sliding Windows ที่ใช้ในงานวิจัย จำนวน 7 รอบ พ.ศ. 2560

บทที่ 4
ผลการศึกษา

ในบทนี้ผู้วิจัยจะได้นำเสนอ การวิเคราะห์ข้อมูลเกิดอุบัติเหตุจริง ต่อมาทำการเปรียบเทียบประสิทธิภาพของแบบจำลอง นำเสนอผลการพยากรณ์ที่ได้ และนำเสนอระบบการแสดงผลการพยากรณ์

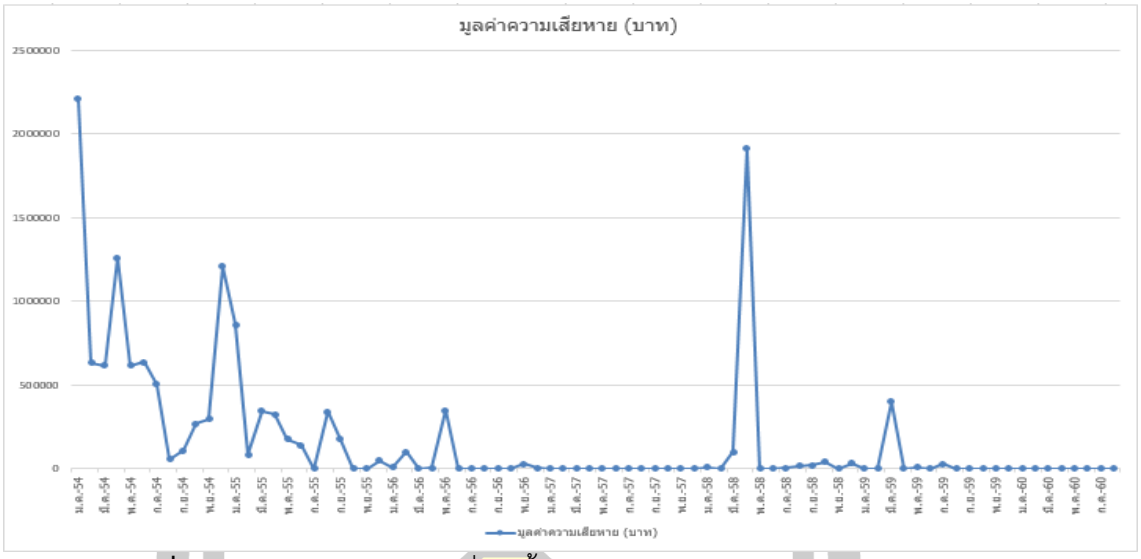
4.1 การวิเคราะห์ข้อมูลข้อมูลการเกิดอุบัติเหตุ

ในการวิเคราะห์ข้อมูลจากการเกิดของอุบัติเหตุระหว่างเดือนมกราคม พ.ศ. 2554 ถึง เดือนมิถุนายน พ.ศ. 2560 ในพื้นที่อำเภอเมือง จังหวัดขอนแก่น ในแต่ละปีเพื่อให้เป็นการเพื่อเข้าใจถึงปริมาณจำนวนอุบัติเหตุที่เกิดโดยแบ่งตาม ประเภทยานพาหนะ เพศ มูลค่าความเสียหาย และความสูญเสียที่เกิดขึ้น



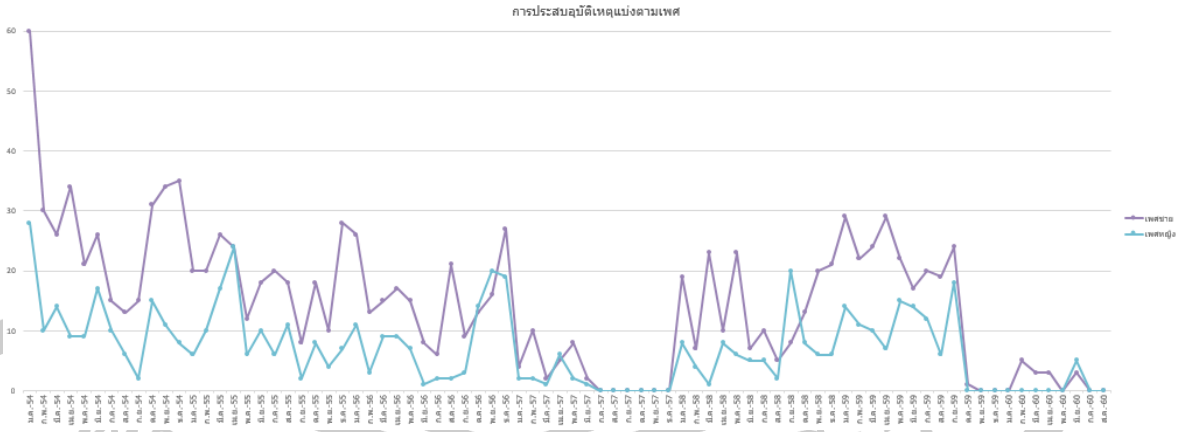
ภาพประกอบที่ 4.1 จำนวนอุบัติเหตุที่เกิดจากยานพาหนะ จากมกราคม พ.ศ. 2554 ถึง มิถุนายน พ.ศ. 2560

จากภาพประกอบที่ 4.1 จากกราฟจะพบช่วงเดือนมกราคมและเมษายนที่เป็นช่วงเทศกาลจะมีการประสบอุบัติเหตุจากรถจักรยานยนต์มีจำนวนสูงที่สุดที่ประสบอุบัติเหตุ รองลงมาเป็นรถบัสและอันดับที่ 3 เป็นรถยนต์ ส่วนยานพาหนะอื่นๆมีเพียงเล็กน้อย แต่ช่วงเวลาที่เกิดจะเป็นช่วงปีใหม่และเทศกาลสงกรานต์ที่มีเยอะที่สุด จะเห็นได้ว่าลักษณะของกราฟเป็นการฝันแปรตามฤดูกาล (Seasonal variation) เป็นการเปลี่ยนแปลงตามอนุกรมเวลาที่เกิดขึ้นซ้ำๆ กันในช่วงเวลาหนึ่ง



ภาพประกอบที่ 4.2 มูลค่าความเสียหายที่เกิดขึ้น จากมกราคม พ.ศ. 2554 ถึง มิถุนายน พ.ศ. 2560

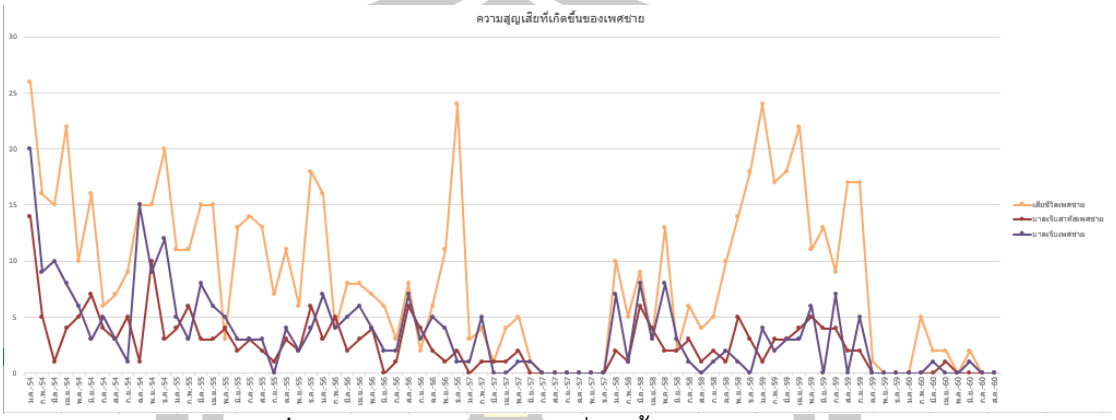
จากภาพประกอบที่ 4.2 จากกราฟจะพบว่าช่วงเดือนมกราคมและเมษายนที่เป็นช่วงเทศกาล จะมีการเกิดอุบัติเหตุและมูลค่าความสูญเสียที่มากที่สุดในรอบปี มากกว่าเดือนอื่นๆที่มีเพียงเล็กน้อย เท่านั้นจะเห็นได้ว่าลักษณะของกราฟเป็นการฝันแปรตามฤดูกาล (Seasonal variation) เป็นการเปลี่ยนแปลงตามอนุกรมเวลาที่เกิดขึ้นซ้ำๆ กันในช่วงเวลาหนึ่ง



ภาพประกอบที่ 4.3 จำนวนของเพศหญิง และชายที่เกิดอุบัติเหตุ
มกราคม พ.ศ. 2554 ถึง มิถุนายน พ.ศ. 2560

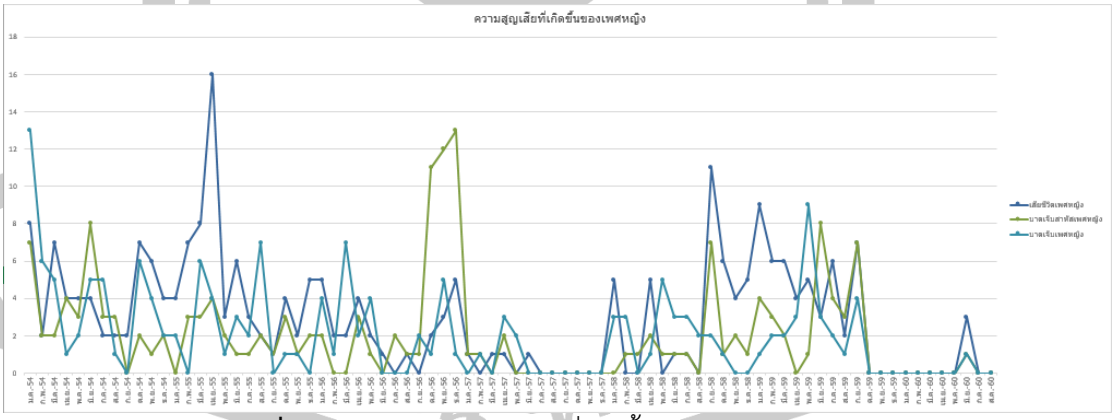
จากภาพประกอบที่ 4.3 จากกราฟจะพบว่าในช่วงเดือนมกราคมและเมษายนของทุกปีที่เป็นช่วงเทศกาล ทั้งเพศหญิงและชาย จะมีการประสบอุบัติเหตุโดยเพศหญิงจะประสบอุบัติเหตุมากกว่าเพศชาย แต่ก็จะมีบางช่วงที่เพศชายที่การประสบอุบัติเหตุใกล้เคียงกับเพศหญิง จะเห็นได้ว่าลักษณะ

ของกราฟเป็นการฝันแปรตามฤดูกาล (Seasonal variation) เป็นการเปลี่ยนแปลงตามอนุกรมเวลาที่เกิดขึ้นซ้ำๆ กันในช่วงเวลาหนึ่ง



ภาพประกอบที่ 4.4 จำนวนความสูญเสียที่เกิดขึ้นของเพศชาย จากเกิดอุบัติเหตุ มกราคม พ.ศ. 2554 ถึง มิถุนายน พ.ศ. 2560

จากภาพประกอบที่ 4.4 จากกราฟจะว่าการสูญเสียที่เกิดขึ้นกับเพศหญิงนั้น เกิดการเสียชีวิตของเพศชายมีจำนวนสูงที่สุดประสบอุบัติเหตุ รองลงมาเป็นการบาดเจ็บสาหัสและการบาดเจ็บเล็กน้อย พบว่า การเกิดอุบัติเหตุมีแนวโน้มแปรผันที่ผิดปกติ (Irregular variation) เป็นการเปลี่ยนแปลงของข้อมูลที่เกิดจากปัจจัยอื่น ๆ นอกเหนือจากค่าแนวโน้ม การเปลี่ยนแปลงตามฤดูกาลหรือการเปลี่ยนแปลงตามวัฏจักร



ภาพประกอบที่ 4.5 จำนวนความสูญเสียที่เกิดขึ้นของเพศหญิง จากเกิดอุบัติเหตุ มกราคม พ.ศ. 2554 ถึง มิถุนายน พ.ศ. 2560

จากภาพประกอบที่ 4.5 จากกราฟจะว่าการสูญเสียที่เกิดขึ้นกับเพศหญิงนั้น เกิดการเสียชีวิตของเพศหญิงมีจำนวนสูงที่สุดที่ประสบอุบัติเหตุ รองลงมาเป็นการบาดเจ็บสาหัสและการบาดเจ็บเล็กน้อย พบว่า การเกิดอุบัติเหตุมีแนวโน้มแปรผันที่ผิดปกติ (Irregular variation) เป็นการเปลี่ยนแปลงของข้อมูลที่เกิดจากปัจจัยอื่น ๆ นอกเหนือจากค่าแนวโน้ม การเปลี่ยนแปลงตามฤดูกาลหรือการเปลี่ยนแปลงตามวัฏจักร

ดังนั้นปริมาณอุบัติเหตุของยานพาหนะ มูลค่าความเสียหาย และเพศ ที่เกิดขึ้น ในเขตอำเภอเมือง จังหวัดขอนแก่น จากเดือนมกราคม พ.ศ. 2554 ถึง มิถุนายน พ.ศ. 2560 พบว่า การเกิดอุบัติเหตุมีแนวโน้มแปรผันตามฤดูกาล โดยเฉพาะในช่วงเทศกาล จะเห็นได้ว่า มีการประสบอุบัติเหตุมากที่สุด ส่วนจำนวนการสูญเสียที่เกิดขึ้นนั้นแนวโน้มแปรผันที่ผิดปกติซึ่งเป็นการเปลี่ยนแปลงของข้อมูลที่เกิดจากปัจจัยอื่น ๆ นอกเหนือจากค่าแนวโน้ม การเปลี่ยนแปลงตามฤดูกาลหรือการเปลี่ยนแปลงตามวัฏจักร

4.2 การเปรียบเทียบการทำงานของแบบจำลอง

ในการศึกษาประสิทธิภาพของแบบจำลองนี้ผู้วิจัยได้เปรียบเทียบ วัดประสิทธิภาพของแบบจำลองด้วย MAE และ RMSE ผลลัพธ์ที่ได้ดังนี้

4.2.1 การเปรียบเทียบค่าความผิดพลาดกำลังสองเฉลี่ย

ค่าความผิดพลาดกำลังสองเฉลี่ย (Mean Absolute Error: MAE) เป็นค่าที่ใช้วัดความถูกต้องของการพยากรณ์ที่นิยมใช้กันอย่างแพร่หลายจากการทดลองสามารถแสดงผลการเปรียบเทียบของค่า MAE ของ 5 แบบจำลอง LR, ANN, SMOreg, SVMR และ GP ในการสร้างแบบจำลองปริมาณการเกิดของอุบัติเหตุ จากยานพาหนะและการสูญเสียที่เกิดขึ้น 22 ชนิดตัวแปร ได้แก่คนเดินเท้า,รถจักรยาน, รถสามล้อ, รถจักรยานยนต์, สามล้อเครื่อง ,รถยนต์นั่ง, รถโดยสารเล็ก (ตู้), รถบรรทุกเล็ก(ปิ๊กอัพ), รถบัส, รถบรรทุก 6 ล้อ, รถบรรทุก 10 ล้อ, รถอีแต๋น, รถแท็กซี่, รถอื่นๆ, มูลค่าทรัพย์สิน(บาท), ชาย, หญิง, เสียชีวิตชาย, เสียชีวิตหญิง, บาดเจ็บสาหัสชาย, บาดเจ็บสาหัสหญิง, บาดเจ็บชาย, บาดเจ็บหญิง, ด้วยหลักการ Sliding Windows ดังตารางที่ 4.1, 4.2 และ 4.3 ตามลำดับ

พหูน ปณ ฑิโต ชีเว

ตารางที่ 4.1 เปรียบเทียบค่า MAE ของ LR, ANN, SMOreg, SVMR และ GP โดยแบ่งตามยานพาหนะ

MODEL	LR	ANN	SMOreg	SVMR	GP
Pedestrian	3.85	4.59	2.76	1.61	3.77
Bicycle	5.03	1.86	3.77	0.31	1.69
Tricycle	0.12	0.31	0.00	0.03	0.19
Motorcycle	22.93	15.02	12.69	7.31	11.80
Tuk tuk	0.88	0.52	0.14	0.08	0.22
Private car	4.68	6.58	5.94	3.04	4.25
Van	2.80	1.72	0.97	0.40	0.99
Pickup truck	12.81	13.70	11.46	4.28	8.23
Bus	0.87	1.37	0.76	0.30	0.47
Truck (6 wheel)	1.01	1.89	1.25	0.61	1.08
Heavy truck	1.43	1.47	1.55	0.50	1.19
Farm vehicle	0.02	0.20	0.03	0.03	0.81
Taxi	0.09	0.11	0.04	0.06	0.10
other	3.05	2.65	3.37	1.27	2.07
Average	4.25	3.71	3.20	1.42	2.63
S.D.	6.32	4.85	4.15	2.11	3.44

ในการสร้างแบบจำลองปริมาณการเกิดของอุบัติเหตุ จากยานพาหนะ ได้แก่ คนเดินเท้า, รถจักรยาน, รถสามล้อ, รถจักรยานยนต์, สามล้อเครื่อง ,รถยนต์นั่ง, รถโดยสารเล็ก(ตู้), รถบรรทุกเล็ก (ปิ๊กอัพ), รถบัส, รถบรรทุก 6ล้อ, รถบรรทุก 10 ล้อ, รถอีแต๋น, รถแท็กซี่ ,รถอื่นๆ ด้วยหลักการ Sliding Windows จากตารางจะเห็นได้ว่าค่าเฉลี่ย (Average) ของค่า MAE ในส่วนของ Testing แบบจำลอง SVMR มีค่าเฉลี่ย (Average) มีค่า MAE 1.42 และ ค่า S.D. ของค่า MAE อยู่ที่ 2.11 ตามลำดับในขณะที่ แบบจำลอง คือ LR, ANN, SMOreg และ GP มีค่า MAE ในส่วนของ Testing มีค่าเฉลี่ย (Average) 4.25, 3.71, 3.20 และ 2.63 ตามลำดับ มีค่า S.D. ของค่า MAE ของแบบจำลอง LR, ANN, SMOreg และ GP 6.32, 4.85, 4.15 และ 3.44 ตามลำดับ ปรากฏว่าเทคนิค SVMR โดยใช้ให้ค่า MAE ที่ต่ำที่สุด ซึ่งแสดงให้เห็นว่าประสิทธิภาพของแบบจำลองที่สร้างจากเทคนิค SVMR มีค่าแม่นยำมากที่สุด

ตารางที่ 4.2 ตารางเปรียบเทียบค่า MAE ของ LR, ANN, SMOreg, SVMR และ GP โดยแบ่งตามเพศ

MODEL	LR	ANN	SMOreg	SVMR	GP
M	141.53	31.10	30.86	11.23	20.36
F	69.56	213.22	17.81	6.75	11.90
Average	105.54	122.16	24.34	8.99	16.13
S.D.	50.89	128.78	9.23	3.17	5.98

ในการสร้างแบบจำลองปริมาณการเกิดของอุบัติเหตุ แบ่งตามเพศ ได้แก่ ชาย, หญิง ด้วยหลักการ Sliding Windows จากตารางจะเห็นได้ว่าค่าเฉลี่ย (Average) ของค่า MAE ในส่วนของ Testing แบบจำลอง SVMR มีค่าเฉลี่ย (Average) มีค่า MAE 8.99 และ ค่า S.D. ของค่า MAE อยู่ที่ 3.17 ตามลำดับในขณะที่ แบบจำลอง คือ LR, ANN, SMOreg และ GP มีค่า MAE ในส่วนของ Testing มีค่าเฉลี่ย (Average) 105.54, 122.16 ,24.34 และ 16.13 ตามลำดับ มีค่า S.D. ของค่า MAE ของแบบจำลอง LR, ANN, SMOreg และ GP 50.89,128.78, 9.23 และ 5.98 ตามลำดับ ปรากฏว่าเทคนิค SVMR โดยใช้ให้ค่า MAE ที่ต่ำที่สุด ซึ่งแสดงให้เห็นว่าประสิทธิภาพของแบบจำลองที่สร้างจากเทคนิค SVMR มีค่าแม่นยำมากที่สุด

ตารางที่ 4.3 ตารางเปรียบเทียบค่า MAE ของ LR, ANN, SMOreg, SVMR และ GP โดยแบ่งตามการสูญเสียที่เกิดขึ้น

MODEL	LR	ANN	SMOreg	SVMR	GP
D_M	47.20	18.32	19.52	7.21	10.62
S_M	5.98	6.72	8.29	2.25	4.56
H_M	47.66	4.88	5.89	3.00	2.52
D_F	25.35	142.54	9.22	2.66	5.82
S_F	28.72	9.95	20.33	2.32	10.48
H_F	8.89	3.58	3.46	1.67	1.93
Average	27.30	31.00	11.12	3.19	5.99
S.D.	17.94	54.90	7.12	2.02	3.80

ในการสร้างแบบจำลองปริมาณการเกิดของอุบัติเหตุ แบ่งตามความสูญเสียที่เกิดขึ้น คือ เสียชีวิตเพศชาย, บาดเจ็บสาหัสเพศชาย,บาดเจ็บเพศชาย เสียชีวิตเพศหญิง,บาดเจ็บสาหัสเพศหญิง และบาดเจ็บเพศหญิง ด้วยหลักการ Sliding Windows จากตารางจะเห็นได้ว่าค่าเฉลี่ย (Average) ของค่า MAE ในส่วนของ Testing แบบจำลอง SVMR มีค่าเฉลี่ย (Average) มีค่า MAE 3.19 และค่า S.D. ของค่า MAE อยู่ที่ 2.02 ตามลำดับในขณะที่ แบบจำลอง คือ LR, ANN, SMOreg และ GP มีค่า MAE ในส่วนของ Testing มีค่าเฉลี่ย (Average) 27.30, 31.00, 11.12 และ 5.99 ตามลำดับ มีค่า S.D. ของค่า MAE ของแบบจำลอง LR, ANN, SMOreg และ GP 17.94 ,54.90, 7.12 และ 3.80 ตามลำดับ ปรากฏว่าเทคนิค SVMR โดยใช้ให้ค่า MAE ที่ต่ำที่สุด ซึ่งแสดงให้เห็นว่าประสิทธิภาพของแบบจำลองที่สร้างจากเทคนิค SVMR มีค่าแม่นยำมากที่สุด

4.2.2 การเปรียบเทียบค่าความคลาดเคลื่อนเฉลี่ยกำลังสอง

รากที่สองของค่าความคลาดเคลื่อนเฉลี่ยกำลังสอง (Root Mean Square Error: RMSE) เป็นค่าที่ใช้วัดความถูกต้องโดยเฉลี่ยของการพยากรณ์ จากการทดลองสามารถแสดงผลการเปรียบเทียบของค่า RMSE ของ 5 แบบจำลอง LR, ANN, SMOreg, SVMR และ GP ในการสร้างแบบจำลองปริมาณการเกิดของอุบัติเหตุ จากยานพาหนะและการสูญเสียที่เกิดขึ้น 22 ชนิดตัวแปร ได้แก่ คนเดินเท้า,รถจักรยาน, รถสามล้อ, รถจักรยานยนต์, สามล้อเครื่อง, รถยนต์นั่ง, รถโดยสารเล็ก (ตู้), รถบรรทุกเล็ก(ปิ๊กอัพ), รถบัส, รถบรรทุก 6 ล้อ, รถบรรทุก 10 ล้อ, รถอีแต๋น, รถแท็กซี่ ,รถอื่นๆ, มูลค่าทรัพย์สิน(บาท), ชาย, หญิง, เสียชีวิตเพศชาย, บาดเจ็บสาหัสเพศชาย,บาดเจ็บเพศชาย เสียชีวิตเพศหญิง,บาดเจ็บสาหัสเพศหญิง และบาดเจ็บเพศหญิง ด้วยหลักการ Sliding Windows ได้ดังตารางที่ 4.4 ,4.5 และ 4.6 ตามลำดับ

ตารางที่ 4.4 ตารางเปรียบเทียบค่า RMSE ของ ของ LR, ANN, SMOreg, SVMR และ GP แบ่งตามยานพาหนะ

MODEL	LR	ANN	SMOreg	SVMR	GP
Pedestrian	4.14	5.13	3.02	1.71	3.97
Bicycle	5.45	2.00	4.04	0.37	1.86
Tricycle	0.17	0.38	0.00	0.03	0.19
Motorcycle	23.63	16.07	13.90	7.64	12.51
Tuk tuk	0.98	0.55	0.18	0.10	0.24
Private car	4.92	7.67	6.37	3.32	4.54
Van	3.26	1.97	1.21	0.52	1.21

ตารางที่ 4.4 ตารางเปรียบเทียบค่า RMSE ของ ของ LR, ANN, SMOreg, SVMR และ GP แบ่งตามยานพาหนะ(ต่อ)

MODEL	LR	ANN	SMOreg	SVMR	GP
Pickup truck	14.82	15.38	12.08	4.59	9.54
Bus	1.10	1.61	0.93	0.37	0.57
Truck (6 wheel)	1.10	2.09	1.40	0.67	1.19
Heavy truck	1.62	1.57	1.76	0.63	1.31
Farm vehicle	0.02	0.24	0.03	0.03	0.81
Taxi	0.12	0.14	0.06	0.09	0.13
other	3.34	2.96	3.61	1.53	2.29
Average	4.62	4.13	3.47	1.54	2.88
S.D.	6.67	5.32	4.45	2.21	3.74

ในการสร้างแบบจำลองปริมาณการเกิดของอุบัติเหตุ จากยานพาหนะ ได้แก่ คนเดินเท้า, รถจักรยาน, รถสามล้อ, รถจักรยานยนต์, สามล้อเครื่อง ,รถยนต์นั่ง, รถโดยสารเล็ก(ตู้), รถบรรทุกเล็ก (ปิ๊กอัพ), รถบัส, รถบรรทุก 6 ล้อ, รถบรรทุก 10 ล้อ, รถอีแต๋น, รถแท็กซี่ ,รถอื่นๆ ด้วยหลักการ Sliding Windows จากตารางจะเห็นได้ว่าค่าเฉลี่ย (Average) ของค่า RMSE ในส่วนของ Testing แบบจำลอง SVMR มีค่าเฉลี่ย (Average) มีค่า RMSE 1.54 และ ค่า S.D. ของค่า RMSE อยู่ที่ 2.21 ตามลำดับในขณะที่ แบบจำลอง คือ LR, ANN, SMOreg และ GP มีค่า RMSE ในส่วนของ Testing มีค่าเฉลี่ย (Average) 4.62, 4.13, 3.47 และ 2.88 ตามลำดับ มีค่า S.D. ของค่า RMSE ของแบบจำลอง LR, ANN, SMOreg และ GP 6.67, 5.32, 4.45 และ 3.74 ตามลำดับ ปรากฏว่าเทคนิค SVMR โดยใช้ให้ค่า RMSE ที่ต่ำที่สุด ซึ่งแสดงให้เห็นว่าประสิทธิภาพของแบบจำลองที่สร้างจากเทคนิค SVMR มีค่าแม่นยำมากที่สุด

พหุน ปณฺฑิโต ชีเว

ตารางที่ 4.5 ตารางเปรียบเทียบค่า RMSE ของ ของ LR, ANN, SMOreg, SVMR และ GP แบ่งตามเพศ

MODEL	LR	ANN	SMOreg	SVMR	GP
M	156.37	32.82	33.23	11.49	21.66
F	76.83	216.85	19.84	6.97	12.73
Average	116.60	124.83	26.54	9.23	17.20
S.D.	56.24	130.13	9.47	3.20	6.32

ในการสร้างแบบจำลองปริมาณการเกิดของอุบัติเหตุ แบ่งตามเพศ ได้แก่ ชาย, หญิง ด้วยหลักการ Sliding Windows จากตารางจะเห็นได้ว่าค่าเฉลี่ย (Average) ของค่า RMSE ในส่วนของ Testing แบบจำลอง SVMR มีค่าเฉลี่ย (Average) มีค่า RMSE 9.23 และ ค่า S.D. ของค่า RMSE อยู่ที่ 3.20 ตามลำดับในขณะที่ แบบจำลอง คือ LR, ANN, SMOreg และ GP มีค่า RMSE ในส่วนของ Testing มีค่าเฉลี่ย (Average) 116.60, 124.83 ,26.54 และ 12.73 ตามลำดับ มีค่า S.D. ของค่า RMSE ของแบบจำลอง LR, ANN, SMOreg และ GP 56.24, 130.13, 9.47 และ 6.32 ตามลำดับ ปรากฏว่าเทคนิค SVMR โดยใช้ให้ค่า MAE ที่ต่ำที่สุด ซึ่งแสดงให้เห็นว่าประสิทธิภาพของแบบจำลองที่สร้างจากเทคนิค SVMR มีค่าแม่นยำมากที่สุด

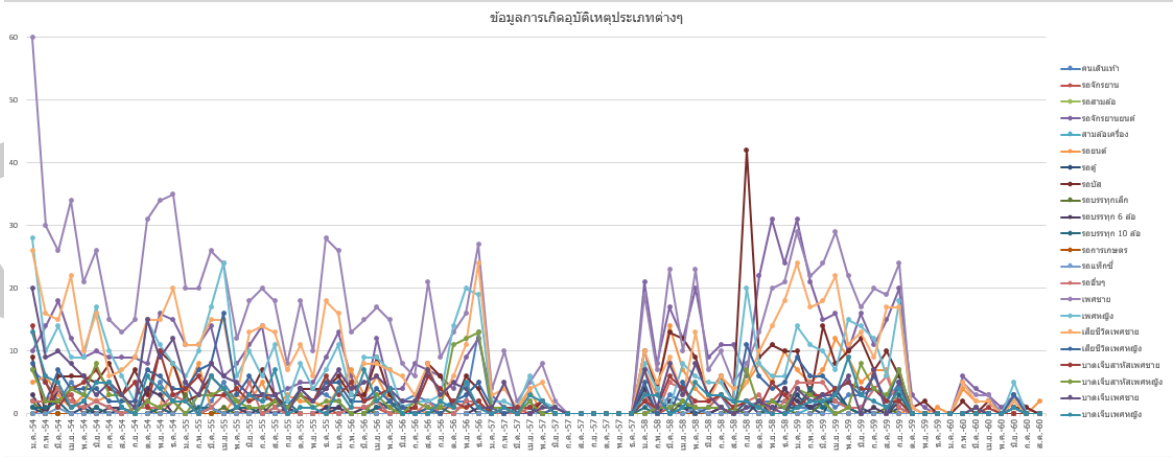
ตารางที่ 4.6 ตารางเปรียบเทียบค่า RMSE ของ ของ LR, ANN, SMOreg, SVMR และ GP แบ่งตามการสูญเสียที่เกิดขึ้น

MODEL	LR	ANN	SMOreg	SVMR	GP
D_M	51.56	19.32	21.12	7.49	11.23
D_F	6.28	151.49	9.89	2.80	6.28
S_M	6.42	7.41	9.07	2.30	4.98
S_F	30.68	10.73	21.16	2.58	11.14
H_M	49.32	5.54	6.57	3.15	2.73
H_F	9.96	4.05	4.03	1.93	2.20
Average	25.70	33.09	11.97	3.37	6.43
S.D.	21.21	58.26	7.39	2.06	3.97

ในการสร้างแบบจำลองปริมาณการเกิดของอุบัติเหตุ แบ่งตามความสูญเสียที่เกิดขึ้น คือ เสียชีวิตเพศชาย, บาดเจ็บสาหัสเพศชาย,บาดเจ็บเพศชาย เสียชีวิตเพศหญิง,บาดเจ็บสาหัสเพศหญิง และบาดเจ็บเพศหญิง ด้วยหลักการ Sliding Windows จากตารางจะเห็นได้ว่าค่าเฉลี่ย (Average) ของค่า RMSE ในส่วนของ Testing แบบจำลอง SVMR มีค่าเฉลี่ย (Average) มีค่า RMSE 3.37 และค่า S.D. ของค่า RMSE อยู่ที่ 2.06 ตามลำดับในขณะที่ แบบจำลอง คือ LR, ANN, SMOreg และ GP มีค่า RMSE ในส่วนของ Testing มีค่าเฉลี่ย (Average) 25.70, 33.09 ,11.97 และ 6.43 ตามลำดับ มีค่า S.D. ของค่า RMSE ของแบบจำลอง LR, ANN, SMOreg และ GP 21.21 ,58.63, 7.39 และ 3.97 ตามลำดับ ปรากฏว่าเทคนิค SVMR โดยใช้ให้ค่า RMSE ที่ต่ำที่สุด ซึ่งแสดงให้เห็นว่าประสิทธิภาพของแบบจำลองที่สร้างจากเทคนิค SVMR มีค่าแม่นยำมากที่สุด

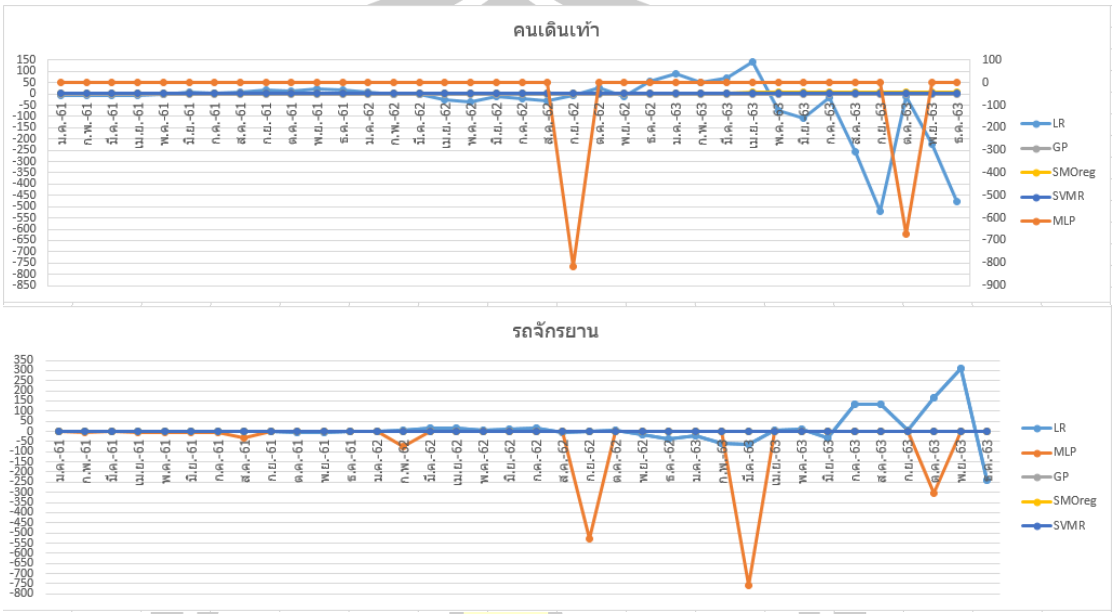
4.3 ผลการพยากรณ์การเกิดอุบัติเหตุบนท้องถนน ของจังหวัดขอนแก่น

ในงานวิจัยนี้ผู้วิจัยได้นำแบบจำลอง SVMR มาใช้ในการวิเคราะห์ปริมาณการเกิดอุบัติเหตุบนท้องถนนของอำเภอเมืองขอนแก่น จากข้อมูลสถานีงานจราจรสถานีตำรวจภูธรจังหวัดขอนแก่น จำนวน 22 ตัวแปร ได้แก่ คนเดินเท้า, รถจักรยาน, รถสามล้อ, รถจักรยานยนต์, สามล้อเครื่อง , รถยนต์, รถตู้, รถบรรทุกเล็ก, รถบัส, รถบรรทุก 6 ล้อ, รถบรรทุก 10 ล้อ, รถการเกษตร, รถแท็กซี่, เพศชาย, เพศหญิง, เสียชีวิตเพศชาย, บาดเจ็บสาหัสเพศชาย, บาดเจ็บเพศชาย เสียชีวิตเพศหญิง, บาดเจ็บสาหัสเพศหญิง และบาดเจ็บเพศหญิง โดยใช้ข้อมูลใน (วันที่ 1 เมษายน พ.ศ. 2555 ถึง วันที่ 30 มีนาคม พ.ศ. 2558) เป็นชุดข้อมูลฝึกสอน (Training) ด้วยหลักการ Sliding Windows และใช้ข้อมูล (วันที่ 1 เมษายน พ.ศ. 2558 ถึง วันที่ 30 มีนาคม พ.ศ. 2559) เป็นชุดข้อมูลทดสอบ (Testing) สามารถแสดงได้ดังภาพประกอบที่ 4.6

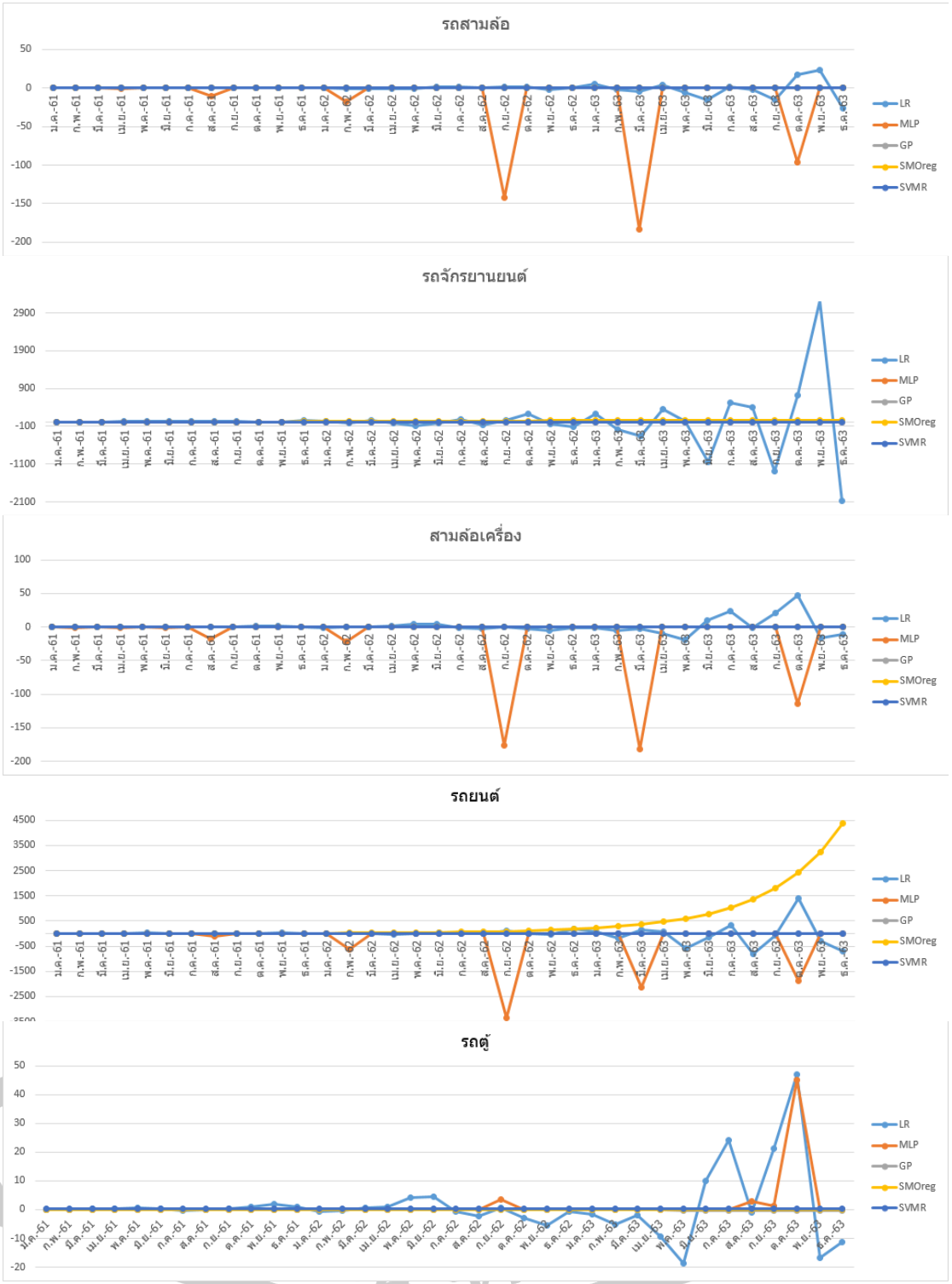


ภาพประกอบที่ 4.6 ข้อมูลการเกิดอุบัติเหตุประเภทต่างๆ

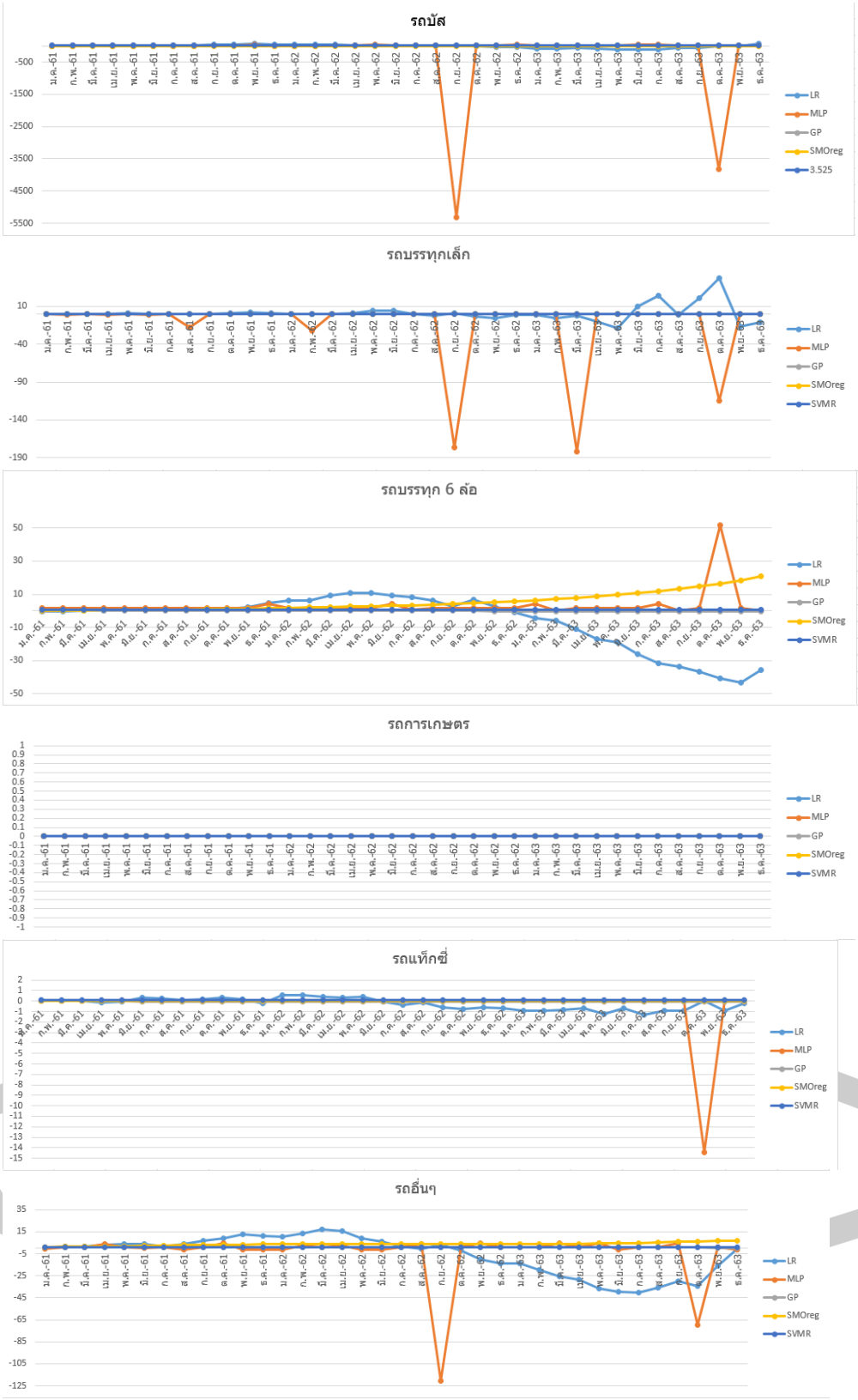
4.3.1 ผลการพยากรณ์การเกิดอุบัติเหตุบนท้องถนน ในช่วง 3 ปี โดยเปรียบเทียบจากทุกเทคนิค ได้ดังกราฟ



ภาพประกอบที่ 4.7 ผลการพยากรณ์การเกิดอุบัติเหตุบนท้องถนน จากประเภทยานพาหนะ



ภาพประกอบที่ 4.8 ผลการพยากรณ์การเกิดอุบัติเหตุบนท้องถนน จากประเภทยานพาหนะ

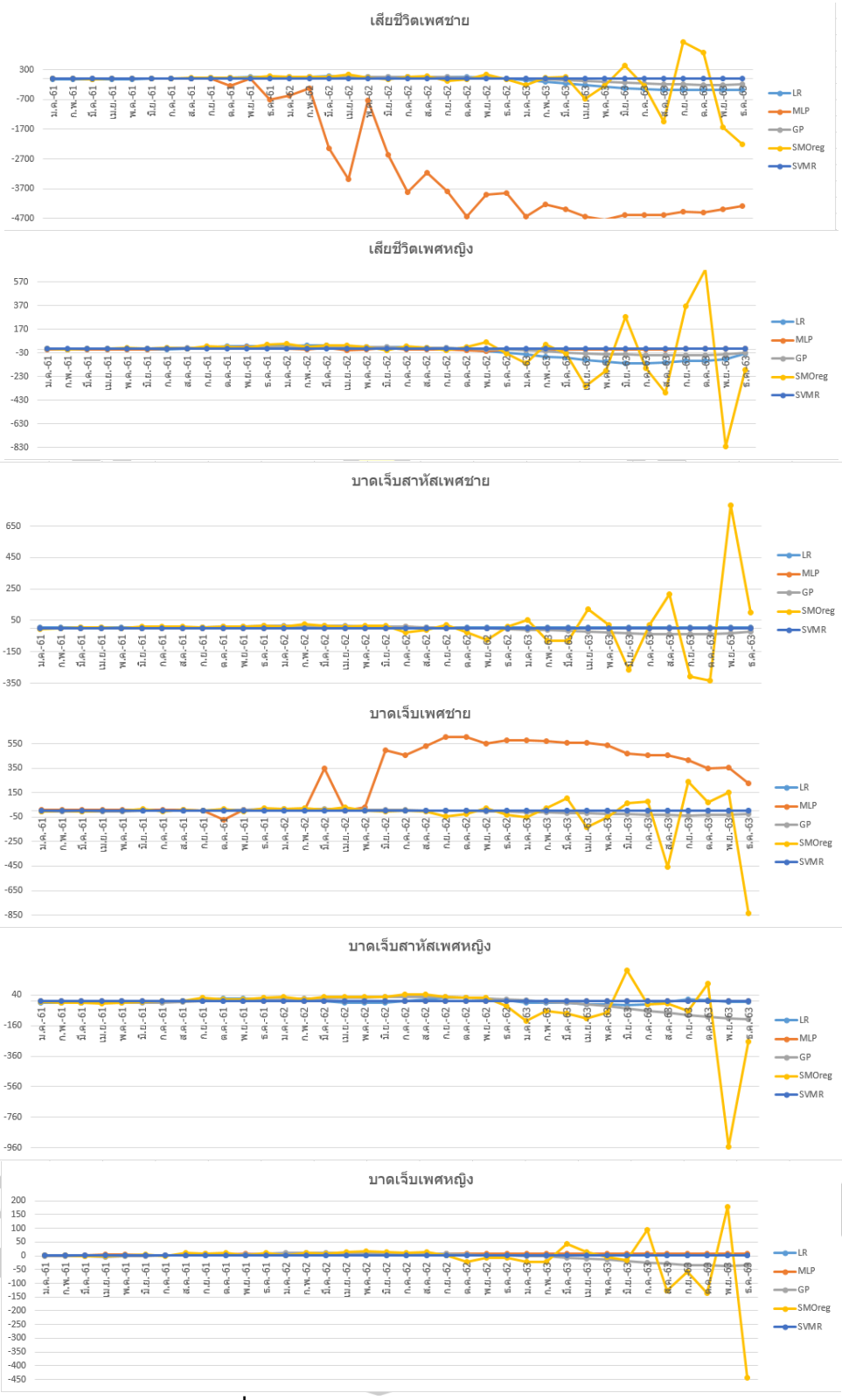


ภาพประกอบที่ 4.9 ผลการพยากรณ์การเกิดอุบัติเหตุบนท้องถนน จากประเภทยานพาหนะ

จากภาพประกอบที่ 4.7 ,4.8 และ 4.9 ที่พบว่าเทคนิคที่พยากรณ์ออกมาเป็นค่าติดลบซึ่งหมายถึงความที่มีความคาดเคลื่อนของการพยากรณ์ที่ได้ แต่มีเทคนิค SVMR ที่สามารถพยากรณ์ได้ใกล้เคียงค่าจริงมากที่สุดคือ มีค่าเข้าใกล้ 0 มากที่สุด ในการพยากรณ์กับการเกิดอุบัติเหตุโดยแบ่งตามยานพาหนะ ตั้งแต่ปี พ.ศ. 2561 ถึง พ.ศ. 2563

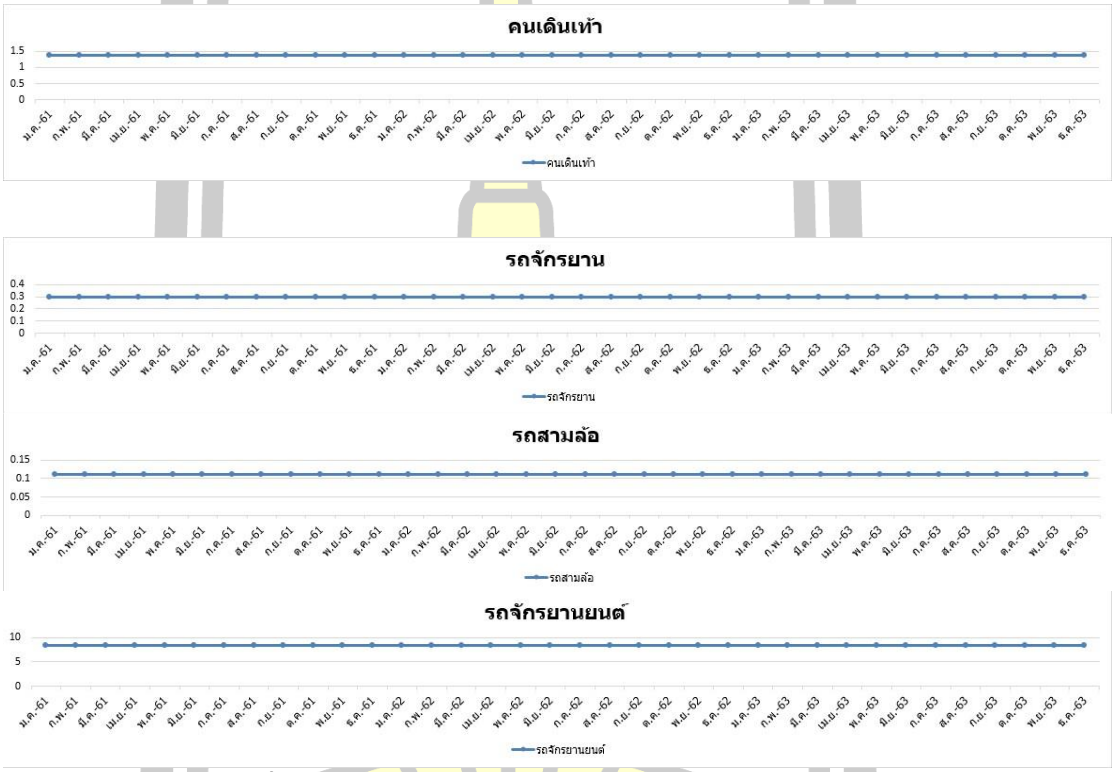
ภาพประกอบที่ 4.10 ผลการพยากรณ์การเกิดอุบัติเหตุบนท้องถนน โดยแบ่งตามเพศ

จากภาพประกอบที่ 4.10 ที่พบว่าเทคนิคที่พยากรณ์ออกมาเป็นค่าติดลบซึ่งหมายถึงความที่มีความคาดเคลื่อนของการพยากรณ์ที่ได้ ทั้งเพศชายและเพศหญิง พบว่า ด้วย 5 เทคนิค ที่นำมาใช้ลักษณะเส้นกราฟของเทคนิค SMOreg มีความแปรปรวนและค่าติดลบมากที่สุด ส่วนเทคนิค SVMR มีค่าใกล้เคียงค่าที่เกิดขึ้นจริงมากที่สุด ในการพยากรณ์กับการเกิดอุบัติเหตุโดยแบ่งตามเพศ ตั้งแต่ปี พ.ศ. 2561 ถึง พ.ศ. 2563



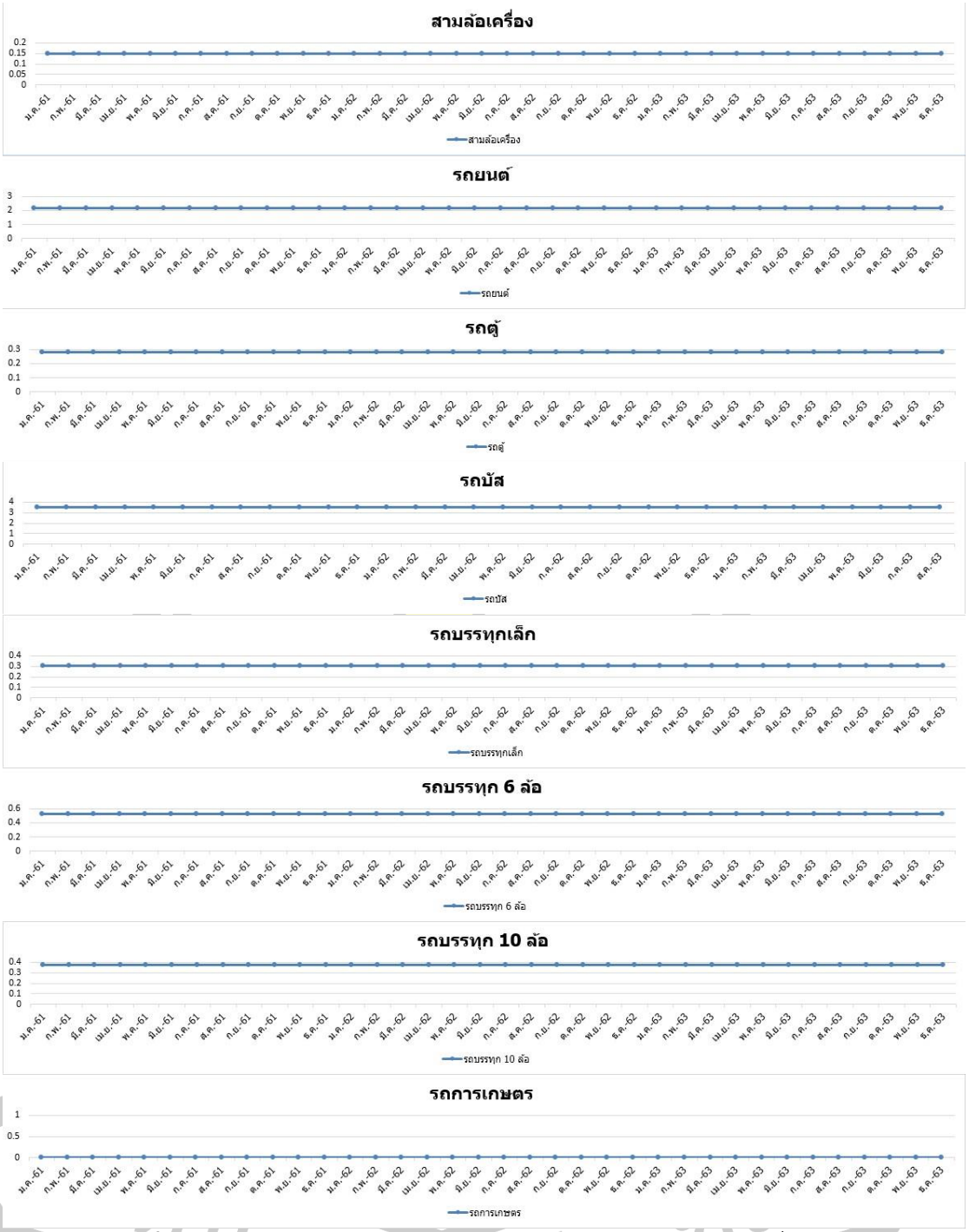
ภาพประกอบที่ 4.11 ผลการพยากรณ์การเกิดอุบัติเหตุบนท้องถนน
โดยแบ่งตามความสูญเสียที่เกิดขึ้น

จากภาพประกอบที่ 4.11 ที่พบว่าเทคนิคที่พยากรณ์ออกมาเป็นค่าติดลบ ซึ่งหมายถึงความที่มีความคาดเคลื่อนของการพยากรณ์ที่ได้ ทั้งเพศชายและเพศหญิง พบว่า ด้วย 5 เทคนิค ที่นำมาใช้ลักษณะเส้นกราฟของเทคนิค SMOreg มีความแปรปรวนและค่าติดลบมากที่สุด ส่วนเทคนิค SVMR มีค่าใกล้เคียงค่าที่เกิดขึ้นจริงมากที่สุด ในการพยากรณ์กับการเกิดอุบัติเหตุโดยแบ่งตามการสูญเสียที่เกิดตั้งแต่ปี พ.ศ. 2561 ถึง พ.ศ. 2563

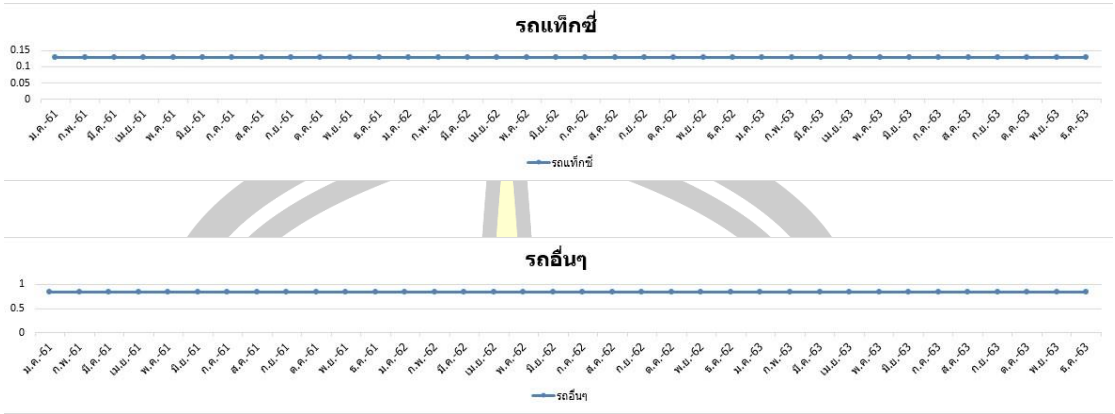


ภาพประกอบที่ 4.12 ผลการพยากรณ์รายเดือน การเกิดอุบัติเหตุของ คนเดินเท้า , รถจักรยาน,รถสามล้อ และรถจักรยานยนต์

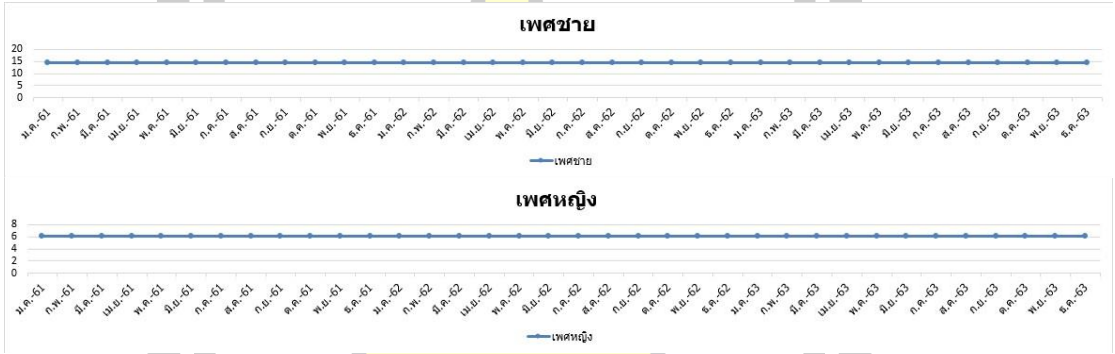
พหุน ปณ ฑิโต ชีเว



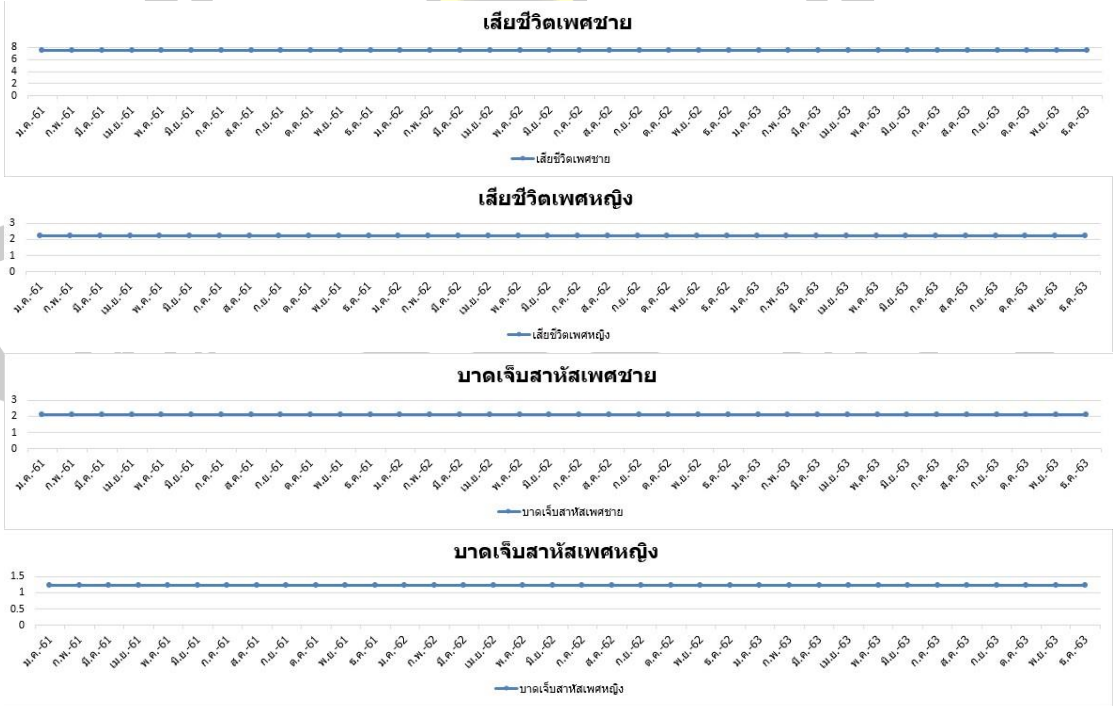
ภาพประกอบที่ 4.13 ผลการพยากรณ์รายเดือน การเกิดอุบัติเหตุของ สามล้อเครื่อง, รถยนต์, รถตู้, รถบัส, รถบรรทุกเล็ก, รถบรรทุก 6 ล้อ, รถบรรทุก10 ล้อ, และรถการเกษตร

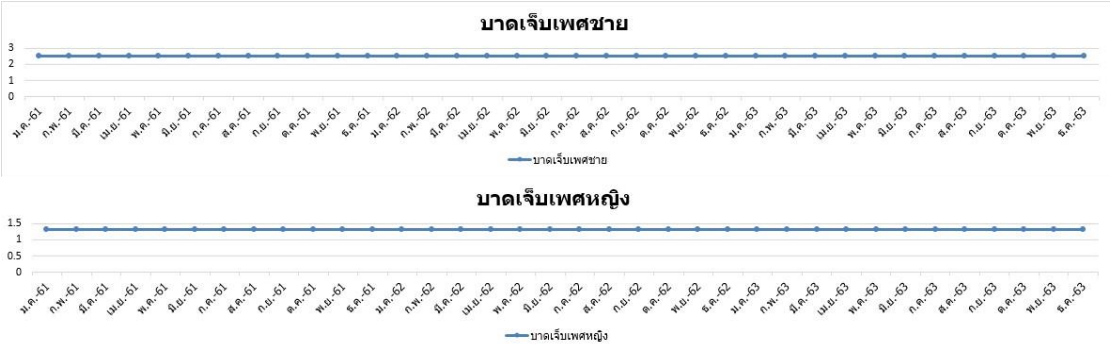


ภาพประกอบที่ 4.14 ผลการพยากรณ์รายเดือน การเกิดอุบัติเหตุของรถแท็กซี่ และ รถอื่นๆ



ภาพประกอบที่ 4.15 ผลการพยากรณ์รายเดือน การสูญเสียเพศชายและหญิง

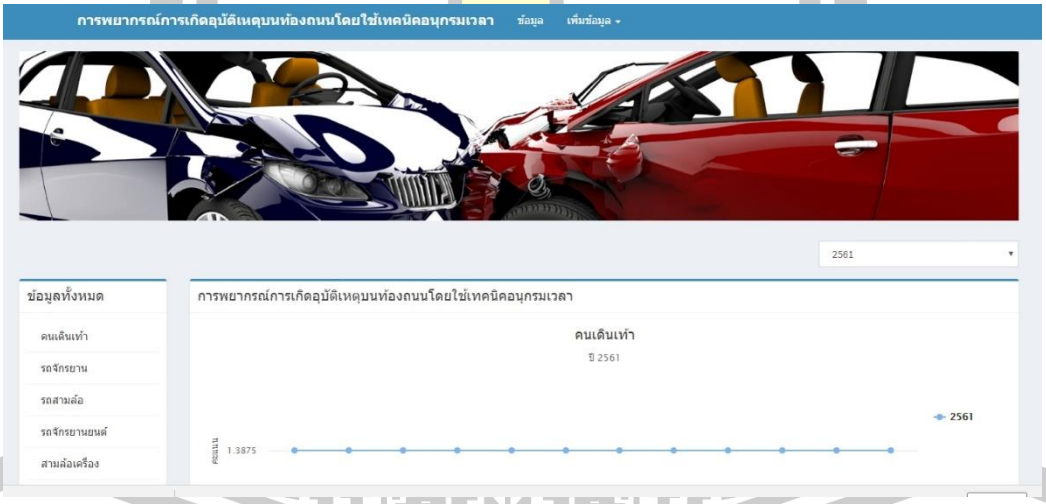




ภาพประกอบที่ 4.16 ผลการพยากรณ์รายเดือน การเสียชีวิตเพศชายและหญิง การบาดเจ็บสาหัสเพศชายและหญิง, บาดเจ็บเพศชายและหญิง

จากภาพประกอบที่ 4.12- 4.16 ผลที่ได้จากการพยากรณ์ที่ผ่านมาผู้วิจัยได้นำแบบจำลองจากเทคนิค SVMR มาพยากรณ์ของเดือนมกราคม 2561-ธันวาคม 2563 จะเห็นได้ว่าเทคนิค SVMR โดยใช้ให้ค่า RMSE และ MAE ในการแสดงประสิทธิภาพของแบบจำลอง สามารถสร้างแบบจำลองที่สร้างจากเทคนิค SVMR ที่มีประสิทธิภาพให้ผลการพยากรณ์ที่ใกล้เคียงค่าจริงมากที่สุด

4.4 การแสดงผลเว็บไซต์ของการพยากรณ์การเกิดอุบัติเหตุล่วงหน้า



ภาพประกอบที่ 4.17 การแสดงผลการพยากรณ์การเกิดอุบัติเหตุล่วงหน้า

หลังจากที่ได้ค่าพยากรณ์แล้ว ผู้วิจัย จึงได้นำผลที่ได้มาแสดงผลลงใน Website เพื่อที่จะได้นำข้อมูลเหล่านั้นมาประกอบก่อนการเดินทาง

บทที่ 5
สรุป วิจารณ์ และข้อเสนอแนะ

การศึกษาครั้งนี้ วัตถุประสงค์ดังนี้

1. เพื่อสร้างแบบจำลองในการพยากรณ์การเกิดอุบัติเหตุบนท้องถนนด้วยเทคนิคอนุกรมเวลาและเปรียบเทียบประสิทธิภาพแบบจำลอง

2. เพื่อแสดงผลของการพยากรณ์การเกิดอุบัติเหตุล่วงหน้า

สามารถสรุปผลการศึกษา อภิปรายผล รวมทั้งข้อเสนอแนะ และแนวมางในการวิจัย ได้ดังต่อไปนี้

5.1 สรุป

จากการทดลอง ผู้วิจัยได้ศึกษาแบบจำลองที่เหมาะสมสำหรับการพยากรณ์การเกิดอุบัติเหตุทางถนน โดยใช้เทคนิคอนุกรมเวลา โดยใช้ 5 เทคนิค คือ LR, ANN, SMOreg, SVMR และ GP ผลการเปรียบเทียบเทคนิคการพยากรณ์การเกิดอุบัติเหตุ โดยใช้ด้วยค่าเฉลี่ยความคลาดเคลื่อนสัมบูรณ์ (Mean Absolute error: MAE) และค่าความคลาดเคลื่อนเฉลี่ยกำลังสอง (Root Mean Square Error: RMSE) พบว่า เทคนิค SVMR มีค่าเฉลี่ยความคลาดเคลื่อนสัมบูรณ์ (MAE) และค่าความคลาดเคลื่อนเฉลี่ยกำลังสอง (RMSE) ดังนี้

1. จากตารางที่ 4.1 ค่าเฉลี่ยของ MAE ของแบบจำลอง SVMR ของการเกิดอุบัติเหตุของยานพาหนะ มีค่า MAE 1.42 และ ค่า S.D. อยู่ที่ 2.11 ในขณะที่แบบจำลอง คือ LR, ANN, SMOreg และ GP มีค่า MAE มีค่าเฉลี่ยอยู่ที่ 4.25, 3.71, 3.20 และ 2.63 ตามลำดับ มีค่า S.D. ของค่า MAE ของแบบจำลอง LR, ANN, SMOreg และ GP 6.32, 4.85, 4.15 และ 3.44 ตามลำดับ

2. จากตารางที่ 4.2 ค่าเฉลี่ยของ MAE ของแบบจำลอง SVMR การพยากรณ์การเกิดอุบัติเหตุทางถนน จากการแบ่งเพศ มีค่า MAE 8.99 และ ค่า S.D. อยู่ที่ 3.17 ในขณะที่แบบจำลอง คือ LR, ANN, SMOreg และ GP มีค่า MAE มีค่าเฉลี่ยอยู่ที่ 105.54, 122.16 ,24.34 และ 16.13 ตามลำดับ มีค่า S.D. ของค่า MAE ของแบบจำลอง LR, ANN, SMOreg และ GP 50.89, 128.78, 9.23 และ 5.98 ตามลำดับ

3. จากตารางที่ 4.3 ค่าเฉลี่ยของ MAE ของแบบจำลอง SVMR การพยากรณ์การเกิดอุบัติเหตุทางถนน จากความสูญเสียที่เกิดขึ้นกับร่างกาย มีค่า MAE 3.19 และ ค่า S.D. อยู่ที่ 2.02 ในขณะที่ แบบจำลอง คือ LR, ANN, SMOreg และ GP มีค่า MAE มีค่าเฉลี่ย 27.30, 31.00, 11.12

และ 5.99 ตามลำดับ มีค่า S.D. ของค่า MAE ของแบบจำลอง LR, ANN, SMOreg และ GP 17.94 , 54.90, 7.12 และ 3.80 ตามลำดับ

4. จากตารางที่ 4.4 ค่าเฉลี่ยของ RMSE ของแบบจำลอง SVMR การพยากรณ์การเกิดอุบัติเหตุทางถนน ของการเกิดอุบัติเหตุของยานพาหนะ มีค่า RMSE 1.54 และ ค่า S.D. อยู่ที่ 2.21 ในขณะที่ แบบจำลอง คือ LR, ANN, SMOreg และ GP มีค่า MAE มีค่าเฉลี่ยอยู่ที่ 4.62, 4.13, 3.47 และ 2.88 ตามลำดับ มีค่า S.D. ของแบบจำลอง LR, ANN, SMOreg และ GP 6.67, 5.32, 4.45 และ 3.74 ตามลำดับ

5. จากตารางที่ 4.5 ค่าเฉลี่ยของ RMSE ของแบบจำลอง SVMR จากการพยากรณ์การเกิดอุบัติเหตุทางถนน จากแบ่งเพศ มีค่า RMSE 9.23 และ ค่า S.D. อยู่ที่ 3.20 ตามลำดับ ในขณะที่ แบบจำลอง คือ LR, ANN, SMOreg และ GP มีค่า RMSE มีค่าเฉลี่ย 116.60, 124.83 ,26.54 และ 12.73 ตามลำดับ มีค่า S.D. ของค่า RMSE ของแบบจำลอง LR, ANN, SMOreg และ GP 56.24, 130.13, 9.47 และ 6.32 ตามลำดับ

6. จากตารางที่ 4.6 ค่าเฉลี่ยของ RMSE ของแบบจำลอง SVMR จากการพยากรณ์การเกิดอุบัติเหตุทางถนน จากความสูญเสียที่เกิดขึ้นกับร่างกาย มีค่า RMSE 3.37 และ ค่า S.D. อยู่ที่ 2.06 ตามลำดับ ในขณะที่ แบบจำลอง คือ LR, ANN, SMOreg และ GP มีค่า RMSE มีค่าเฉลี่ย (Average) 25.70, 33.09 ,11.97 และ 6.43 ตามลำดับ มีค่า S.D. ของค่า RMSE ของแบบจำลอง LR, ANN, SMOreg และ GP 21.21 ,58.63, 7.39 และ 3.97 ตามลำดับ

7. ผลการพยากรณ์ที่ได้ดังกล่าว สามารถนำผลที่ได้มาแสดงให้บนเว็บไซต์แก่ผู้คนที่สนใจประกอบการตัดสินใจเพื่อการเดินทาง ใช้รถใช้ถนนต่อไป

จากผลการการเปรียบเทียบประสิทธิภาพ ที่ได้กล่าวข้างต้นนี้แสดงให้เห็นว่าแนวโน้มของการพยากรณ์ส่วนใหญ่มีทิศทางไปทางเทคนิค (SVMR) เป็นส่วนมากรองลงมาคือ เทคนิค (GP) ซึ่งมี จึงสรุปได้ว่าเทคนิค Support Vector Machine for Regression: (SVMR) มีความเหมาะสมแก่การพยากรณ์การเกิดอุบัติเหตุทางถนน มากที่สุด

5.2 วิจารณ์

สรุปได้ว่า SVMR มีความเหมาะสมในการพยากรณ์การเกิดอุบัติเหตุทางถนน โดยใช้เทคนิคอนุกรมเวลา และมีความเหมาะสมในการพัฒนาแบบจำลองเพื่อพยากรณ์การเกิดอุบัติเหตุบนท้องถนนและเพื่อใช้ในการวางแผนในการเดินทางและได้หลีกเลี่ยงการเกิดอุบัติเหตุที่จะเกิดขึ้น ซึ่งสอดคล้องกับงานวิจัยของ พิณสุดา สิงไชย [41] ที่นำเอาเทคนิค SVMR มาใช้ในงานวิจัยการการพยากรณ์ความต้องการพลังงานไฟฟ้าราย 1 นาที ด้วยเทคนิค SVM-R พบว่าการพยากรณ์ของกลุ่มอาคารพาณิชย์จะมีค่าใกล้เคียงกับปริมาณการใช้พลังงานไฟฟ้าจริงมากที่สุด

5.3 ข้อเสนอแนะ

งานวิจัยนี้ได้มีการศึกษาเทคนิคการพยากรณ์จำนวน 5 เทคนิค โดยเทคนิคที่ทดสอบแล้วได้ค่าการพยากรณ์ที่คลาดเคลื่อนน้อยที่สุด ในเอกสารฉบับนี้ หากมีการศึกษาเทคนิควิธีอื่นเข้ามาเปรียบเทียบเพิ่มเติม ทำให้ได้ค่าการพยากรณ์ที่ต่างกันในแต่ละเทคนิค โดยการนำข้อมูลการเกิดอุบัติเหตุโดยใช้ข้อมูลย้อนหลัง 7 ปี โดยแยกข้อมูลทดสอบเป็นรายเดือน ซึ่งสามารถเพิ่มตัวแปรด้านอื่นๆ เช่น ความหนาแน่นของการจราจร แสงสว่างของถนน และสามารถประยุกต์กับงานวิจัยที่มีข้อมูลใกล้เคียงกันก็สามารถนำไปปรับใช้ให้เหมาะสมกับการพยากรณ์เพื่อให้แบบจำลองมีประสิทธิภาพมากที่สุดในการในไปใช้งานทางด้านอื่นๆต่อไป

พหูน ปณฺฑิโต ชีเว

บรรณานุกรม

บรรณานุกรม

[1] ศูนย์ข้อมูลอุบัติเหตุ. รายงานสถิติการใช้สิทธิ พรบ.รายจังหวัด. สืบค้นเมื่อ: 30 ธันวาคม 2560, สืบค้นจาก http://rvpreport.rvpeservice.com/viewrsc.aspx?report=0486&session=16.

[2] ศูนย์วิจัยอุบัติเหตุแห่งประเทศไทย. โครงการวิจัย การใช้ความเร็วในการขับขี่ที่ปลอดภัย สืบค้นเมื่อ: 29 มิถุนายน 2560, สืบค้นจาก http://www.tarc.ait.ac.th/th/speed.php.

[3] ศูนย์วิจัยอุบัติเหตุแห่งประเทศไทย. การพัฒนาแบบจำลองการเกิดอุบัติเหตุ. สืบค้นเมื่อ : 17 กรกฎาคม 2560, สืบค้นจาก http://www.tarc.ait.ac.th/th/index.php.

[4] ศูนย์วิจัยอุบัติเหตุแห่งประเทศไทย. การพัฒนาแบบจำลองการเกิดอุบัติเหตุ. สืบค้นเมื่อ : 26 กันยายน 2560, สืบค้นจาก http://bhs.doh.go.th/files/Project/TACR/ Accident%20Prediction%20Model_th.pdf.

[5] วราพร งามสุข. "การพยากรณ์อนุกรมเวลาโดยการเปรียบเทียบวิธีแบบฉบับและวิธีบอกซ์-เจนกินส์ กรณีศึกษาจำนวนการเกิดอุบัติเหตุในประเทศไทย". ปัญหาพิเศษปริญญาวิทยาศาสตรบัณฑิต สาขาวิชาสถิติ, มหาวิทยาลัยบูรพา, 2555.

[6] Sven F. Crone, R. D., "Forecasting Seasonal Time Series with Neural Networks: A Sensitivity Analysis of Architecture Parameters," *2007 International Joint Conference on Neural Networks 12-17 Aug. 2007 Orlando, FL, USA*, IEEE Trans. Neural Networks, 2099-2104, 2007.

[7] ชมรมนักวิทยุสมัครเล่นแก่นทอง จังหวัดขอนแก่น.สัญลักษณ์เมืองดอกคูณเสียงแคน. สืบค้นเมื่อ : 21 เมษายน 2561, สืบค้นจาก http://www.kaentong.com/index.php?topic= 8604.

[8] บริษัท ไททัวร์อินโฟร์ จำกัด. จังหวัดขอนแก่น. สืบค้นเมื่อ : 30 ธันวาคม 2560, สืบค้นจาก http://www.thai-tour.com/thai-tour/northeast/khonken/data/area.htm.

[9] มหาวิทยาลัยรามคำแหง สาขาวิทยบริการ จังหวัดขอนแก่น. สืบค้นเมื่อ : 11 ธันวาคม 2560. แผนที่ขอนแก่น. สืบค้นจาก http://www.kk.ru.ac.th/mapkhon.htm

[10] สำนักงานอำนวยความสะดวกปลอดภัย กรมทางหลวง. ผู้ประสบภัย. สืบค้นเมื่อ : 30 ธันวาคม 2560, สืบค้นจาก http://bhs.doh.go.th/statistic/casualty.

[11] สารานุกรมไทยสำหรับเยาวชน. อุบัติเหตุเนื่องจากการจราจร. สืบค้นเมื่อ : 29 เมษายน 2560, สืบค้นจาก http://kanchanapisek.or.th/kp6/sub/book/ book.php? book=8&chap=7&page=t8-7-infodetail04.html.

[12] สถาบันการพละศึกษา วิทยาเตสุพรรณบุรี. การป้องกันอุบัติเหตุในการจราจรทางบก. สืบค้นเมื่อ : 20 เมษายน 2561, สืบค้นจาก http://www.ipesp.ac.th/learning/supitcha/html/E4-2-2.html.

[13] ขจรศักดิ์ ศรีอ่อน. การทำนายสาเหตุของเหตุการณ์กระแสไฟฟ้าขัดข้อง โดยใช้เทคนิคการทำเหมืองข้อมูลในระบบจำหน่ายของการไฟฟ้าส่วนภูมิภาค เขต 1 ภาคกลาง. วิทยานิพนธ์ปริญญาวิศวกรรมศาสตรมหาบัณฑิต, มหาวิทยาลัยเกษตรศาสตร์, 2552.

[14] วาทินี นุ้ยเพียร และคณะ. การเปรียบเทียบประสิทธิภาพและวิเคราะห์การจำแนกข้อมูลโดยใช้โครงข่ายประสาทเทียม ซัพพอร์ตเวกเตอร์แมชชีน นาอีฟเบย์ และเคเนียรเรสต์เนเบอร์. วิทยานิพนธ์ปริญญาเทคโนโลยีสารสนเทศมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ, มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2553.

[15] สุมาลัย นุชิต และคณะ. การวิเคราะห์ปัจจัยที่มีอิทธิผลต่อการเพิ่มลดของจำนวนนักศึกษา กรณีศึกษา วิทยาลัยการอาชีพวังสะพุง. วิทยานิพนธ์ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์, มหาวิทยาลัยขอนแก่น, 2553.

[16] N, D.-x., and W, Y.-l., "Support Vector Machines Based on Data Mining Technology in Power Load Forecasting," *International Conference on Wireless Communications, Networking and Mobile Computing,* pp. 5373 - 5376, 2007.

[17] Poel, DVd, D'Haen, J., and Thorleuchter, D., "Predicting customer profitability during acquisition : Finding the optimalcombination of data source and data mining technique," *Expert Systems with Applications,* 40, 2007-2012, 2013.

[18] Jiawei Han, and Micheline Kamber, "Data mining: Concepts and techniques," Morgan Kaufmann, Elsevier Science, San Francisco, 2006. http://myweb.sabanciuniv.edu/rdehkharghani/files/2016/02/The-Morgan-Kaufmann-Series-in-Data-Management-Systems-Jiawei-Han-Micheline-Kamber-Jian-Pei-Data-Mining.-Concepts-and-Techniques-3rd-Edition-Morgan-Kaufmann-2011.pdf.

[19] บริษัท 168 เอคดูเคชั่น จำกัด. (2561). อนุกรมเวลา. สืบค้นเมื่อ : 21 เมษายน 2561, สืบค้นจาก http://119.46.166.126/self_all/selfaccess12/m6/713/math6_2page2_8.php.

[20] วิชัย แหวนเพชร. การวางแผนและควบคุมการผลิต, พิมพ์ครั้งที่ 5. กรุงเทพฯ: ธรรกมลการพิมพ์, 2551.

[21] วิชัย สุรเชิดเกียรติ. การพัฒนาคอมพิวเตอร์ซอฟต์แวร์ เพื่อการพยากรณ์ข้อมูลอนุกรมเวลาแยกองค์ประกอบ. กรุงเทพฯ: สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ, 2548.

[22] Ruekkasaem, L., "Demand Forecasting for Production Planning :A Case Study of Cleanroom Apparel," *Parichart Journal, Thaksin University,* 28, 290 -304, 2015.

[23] Data Flair. (2018). *Artificial Neural Network (ANN) in Machine Learning*. Available: https://data-flair.training/blogs/artificial-neural-network/.

[24] Chih-Chung Chang, and Chih-Jen Lin, "A Library for Support Vector Machines " *ACM Transactions on Intelligent Systems and Technology,* 2, 2011

[25] C. E. Rasmussen, and C. K. I. Williams, "Gaussian processes for machine learning," *Cambridge, Mass.: MIT Press,* 2006.

[26] Song-bai H, Ya-jun W, Yue-kun S, Wen-wei G, Qiang C, and Ya-qin A, "The Research of Multidimensional Association Rule in Traffic Accidents," *Wireless Communications, Networking and Mobile Computing,* pp. 1-4, 2008.

[27] พิณรัตน์ นุชโพธิ์. การศึกษาอุบัติเหตุการจราจรทางถนนในประเทศไทยโดยใช้กฎการหาความสัมพันธ์ของการทำเหมืองข้อมูล. วิทยานิพนธ์ปริญญาครุศาสตร์อุตสาหกรรม สาขาวิชาคอมพิวเตอร์และเทคโนโลยีสารสนเทศ, มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี, 2549.

[28] Li-Yen Chang, and Hsiu-Wen Wang, "Analysis of traffic injury severity : An application of non-parametric classification tree techniques," *Journal Accident Analysis and Prevention,* 38, 1019–1027, 2006.

[29] Mats Ramstedt, "Alcohol and fatal accidents in the United States A time series analysis for 1950–2002," *Accident Analysis and Prevention,* 40, 1273-1281, 2008.

[30] พิณรัตน์ นุชโพธิ์ และคณะ. ความสัมพันธ์ปริมาณอุบัติเหตุในรูปอนุกรมเวลาแบบแยกส่วนประกอบ. วิทยานิพนธ์ปริญญาวิทยาศาสตรบัณฑิต สาขาวิชาสถิติ, มหาวิทยาลัยบูรพา, 2555.

[31] จันทนี วงศ์พานิช. การเปรียบเทียบเทคนิคทางสถิติโดยวิธีพยากรณ์เดี่ยวและวิธีพยากรณ์รวม : กรณีศึกษา ข้อมูลอุบัติเหตุจราจรในเขตกรุงเทพมหานคร. วิทยานิพนธ์ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาชีวสถิติ, มหาวิทยาลัยมหิดล, 2543.

[32] เกษม ชูจารุกุล, สมชาย วิกิจไพศาล และ วิศณุ ทรัพย์สมพล. "การพัฒนาแบบจำลองอุบัติเหตุสำหรับถนนทางหลวงในประเทศไทย". การประชุมวิชาการวิศวกรรมโยธาแห่งชาติ ครั้งที่ 19, pp. 2418-2422, 2557.

[33] Chong, M., Abraham, A., and Paprzycki, M., "Traffic Accident Analysis Using Machine Learning Paradigms," *Informatica,* 29, 89-98, 2005.

[34] อมรภัทร์ หาญโคกกรวด. "แบบจำลองทำนายความเสี่ยงการเกิดอุบัติเหตุบนท้องถนนโดยใช้เทคนิคถุงจำแนก". การประชุมวิชาการมหาวิทยาลัยมหาสารคามวิจัย ครั้ง ที่ 9; pp. 254-261, 2556.

[35] เศรษฐชัย ชัยสนิท และรัชฎาวรรณ นิ่มนวล. โปรแกรมพยากรณ์ ความต้องการผลไม้ไทยเพื่อการส่งออก ด้วยเทคนิคการถดถอยเชิงเส้น. วารสารวิทยาศาสตร์เกษตร, 40(3), 253-256, 2552.

[36] สักการะ จู่ซู่ และ ธนพลเจน สุทธิเวชกุล. "เปรียบเทียบการพยากรณ์ปริมาณการยืมหนังสือห้องสมุดโรงเรียนนาบอน ด้วยการวิเคราะห์การถดถอยและโครงข่ายประสาทเทียม". The 6th National Conference on Computing and Information Technology (NCCIT2010) vol. 1. pp. 110-115, 2010.

[37] อภิญญา หิรัญวงษ์ และประสิทธิ์ พยัคฆพงษ์. "การพยากรณ์ราคาพืชน้ำมันโดยวิธีของบ๊อกซ์-เจนกินส์และโครงข่ายประสาทเทียม," 47*th Kasetsart University Annual Conference,* pp.196-202, 2552.

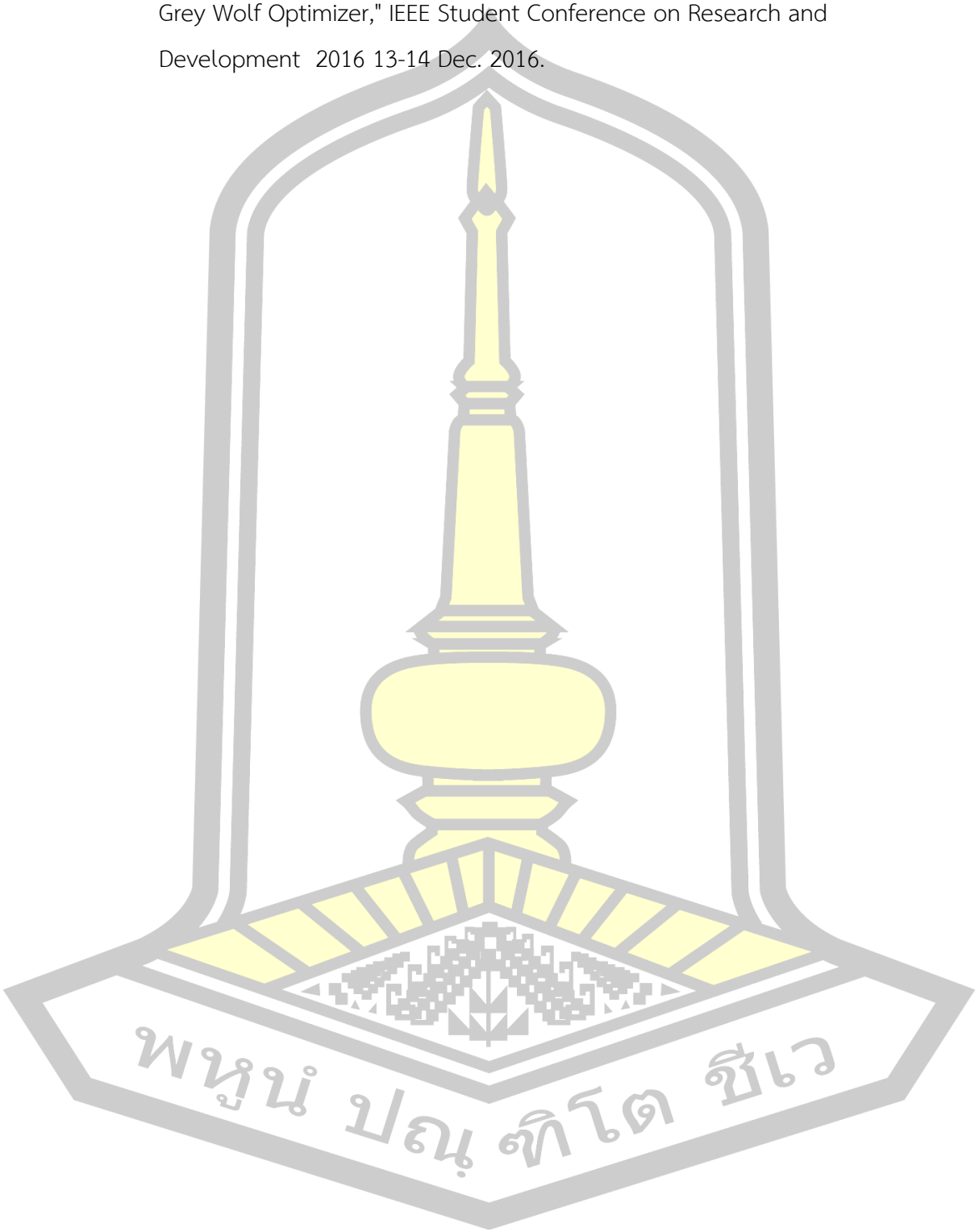
[38] พงษ์ศิริ ศิริพานิช. การพยากรณ์อนุกรมเวลาด้วยตัวแบบผสม ARIMA และเครือข่ายประสาทเทียม. วิทยานิพนธ์ปริญญาวิทยาศาสตรบัณฑิต สาขาวิชาคณิตศาสตร์และเทคโนโลยีสารสนเทศ, มหาวิทยาลัยศิลปากร, 2550.

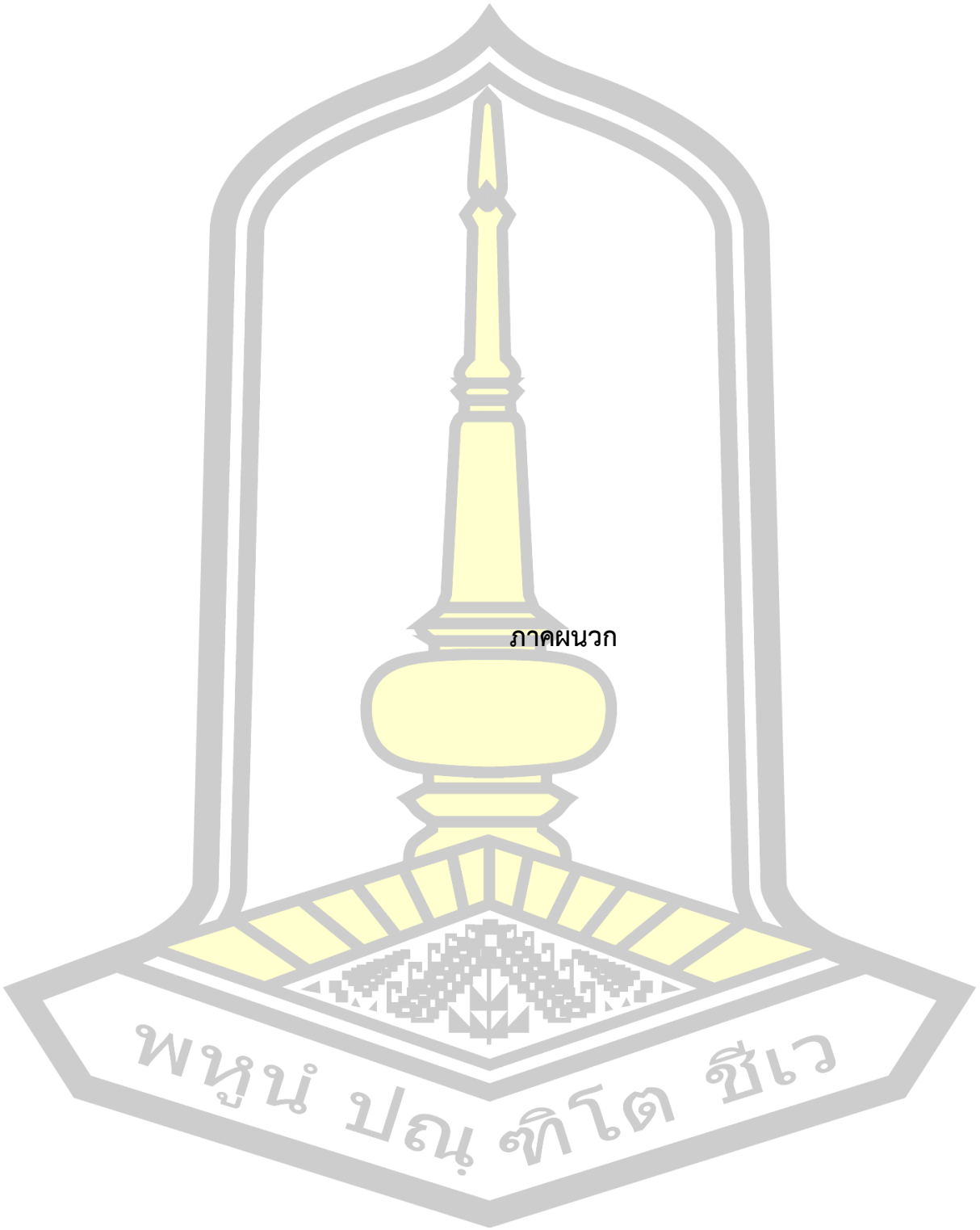
[39] อรรคพล เชี่ยงโหล. การเปรียบเทียบความแม่นยำในการพยากรณ์รูปแบบอนุกรมเวลากับเทคนิคการพยากรณ์แบบเป็นเหตุเป็นผล กรณีศึกษา บริษัทผู้ผลิตรถกระบะแห่งหนึ่งในประเทศไทย. การค้นคว้าอิสระปริญญาวิทยาศาสตรมหาบัณฑิต สาขาการจัดการโลจิสติกส์, มหาวิทยาลัยเกษตรศาสตร์ วิทยาเขตศรีราชา, 2552.

[40] ทัศนัย พลอยสุวรรณ และธนวันต์ ยลพันธ์. การพยากรณ์ค่าความต้องการใช้ไฟฟ้าสูงสุดของการไฟฟ้าฝ่ายผลิตแห่งประเทศไทยโดยกระบวนการเกาส์เซี่ยน. การประชุมทางวิชาการทางวิศวกรรมไฟฟ้า ครั้งที่ 36; 2556.

[41] พิณสุดา สิงไชย และพงษ์ศักดิ์ กีรติวินทกร. ความต้องการพลังงานไฟฟ้าสำหรับศูนย์จัดการความต้องการพลังงานไฟฟ้าของประเทศไทย. วารสารเทคโนโลยีสารสนเทศ, 10, 32-42, 2557.

[42] Zainal NA and Mustaffa Z. "Developing a gold price predictive analysis using Grey Wolf Optimizer," IEEE Student Conference on Research and Development 2016 13-14 Dec. 2016.





ภาคผนวก

การเปรียบเทียบประสิทธิภาพของแบบจำลองการเกิดอุบัติเหตุทางถนน โดยใช้เทคนิคอนุกรมเวลา

Performance Comparison of the Road Occurrence Accidents Prediction Models using Time Series Techniques

ปทิตญา บุญรักษา[1*] และ จารี ทองคำ [2]

นิสิต สาขาเทคโนโลยีสารสนเทศ คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม[1]

อาจารย์ที่ปรึกษา กลุ่มสารสนเทศเชิงประยุกต์ สาขาเทคโนโลยีสารสนเทศ คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม[2]

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อ ศึกษาและเปรียบเทียบประสิทธิภาพของแบบจำลองในการพยากรณ์จำนวนผู้ประสบอุบัติเหตุบนท้องถนน ตั้งแต่ เดือนมกราคมพ.ศ.2554 ถึง เดือนธันวาคม 2559 ของจังหวัดขอนแก่น โดยใช้ 5 เทคนิคที่มีประสิทธิภาพในการสร้างแบบจำลองคือ Linear regression (LR), Artificial Neural Network (ANN), Sequential Minimal Optimization for Regression (SMOreg), Support Vector Machine Regression (SVR) และ Gussian Process (GP) มาใช้ คณะผู้วิจัยได้ใช้หลักการ Sliding Window ในการแบ่งกลุ่มข้อมูลเป็นชุดข้อมูลเรียนรู้ และชุดข้อมูลทดสอบ เพื่อวัดประสิทธิภาพการพยากรณ์ของแบบจำลองด้วยค่า เพื่อวัดประสิทธิภาพการพยากรณ์ของแบบจำลองด้วยค่า mean absolute error (MAE) และ root mean square error (RMSE)

ผลการวิจัยพบว่า เทคนิค SVM มีประสิทธิภาพในการสร้างแบบจำลองในการพยากรณ์จำนวนผู้ประสบอุบัติเหตุบนท้องถนนที่มีค่า ซึ่งเป็นค่าความผิดพลาดต่ำสุดเมื่อเปรียบเทียบกับแบบจำลองที่สร้างจากเทคนิค LR, ANN, SMOreg และ Gussian Process

คำสำคัญ: การพยากรณ์, อุบัติเหตุ, อุบัติเหตุบนท้องถนน, อนุกรมเวลา,เหมืองข้อมูล

ABSTRACT

The purposes of the research were to study and comparison of effective models for forecasting number of road accidents. The data were collected from January 2010 to December 2016 at Khon Kaen province. Five techniques were used including Linear regression (LR), Artificial Neural Network (ANN), Sequential Minimal Optimization for Regression (SMOreg), Support Vector Machine Regression (SVR) and Gussian Process (GP). The Sliding Window method was utilized to segment data into learning and testing sets. The predictive efficiency of the models were measured by Mean Absolute Error (MAE) and Root Mean Square Error (RMSE).

The research finding showed that SVM techniques is effective in predicting the number of road accident victims. This is the lowest error value when compared with LR, ANN, SMOreg and Gussian Process models.

Keywords: Accident, Road Accidents, Time Series, Data Mining

บทนำ

การเกิดอุบัติเหตุจราจรทางบกเป็นปัญหาที่สำคัญปัญหาหนึ่งของโลกและมีแนวโน้มความรุนแรงของการเกิดอุบัติเหตุจราจรทางบกเพิ่มมากขึ้น ซึ่งไม่เพียงจะก่อให้เกิดผลเสียต่อร่างกาย ชีวิต และทรัพย์สินแต่ยังสร้างความเสียหายทางเศรษฐกิจของประเทศ สำนักงานอำนวยความปลอดภัย กรมทางหลวง รายงานว่ามีผู้ประสบอุบัติเหตุ 211,499 ครั้ง ทั่วทั้งประเทศ ในปี พ.ศ. 2560 ซึ่งมีจำนวนผู้บาดเจ็บ 240,899 ราย ทุพพลภาพ 711 และเสียชีวิต 7,074 ราย โดยเฉพาะจังหวัดขอนแก่นเป็นเมืองที่ขนส่งสินค้าเส้นใหญ่นี้เพิ่มมากขึ้นมีความเสี่ยงการเกิดอุบัติเหตุบนท้องถนน ในปี พ.ศ. 2560 มีผู้ประสบอุบัติเหตุเกิดขึ้น 3,364 ครั้ง เป็นจำนวน 4,811 ราย มีจำนวนผู้บาดเจ็บ 3,667 ราย ทุพพลภาพ 12 ราย และเสียชีวิต 152 ราย[1]

ข้อมูลอุบัติเหตุจราจรทางบก เป็นข้อมูลอนุกรมเวลา ซึ่งเป็นชุดของข้อมูลที่เก็บรวบรวมตามระยะเวลารายวันต่อเนื่องกัน ข้อมูลดังกล่าวสามารถวิเคราะห์ได้ด้วยการวิเคราะห์อนุกรมเวลา (Time Series Analysis) [2] ซึ่งเป็นการพยากรณ์ที่อาศัยข้อมูลในอดีตมาพยากรณ์ การเปลี่ยนแปลงของข้อมูล เมื่อเวลาเปลี่ยนการเกิดอุบัติเหตุจะมีแนวโน้มสูงขึ้นหรือลดลง เพื่อให้ทางกลมทางบก โรงพยาบาล ได้นำไปช่วยในการจัดเตรียมบุคลากรทางการแพทย์ เครื่องมือแพทย์ และอื่นๆ ซึ่งการการวิเคราะห์อนุกรมเวลาโดยทั่วไปใช้การวิเคราะห์ถดถอย (Linear Regression)

ปัจจุบันเทคนิคในเหมืองข้อมูลเข้ามามีบทบาทในการวิเคราะห์ข้อมูลอนุกรมเวลา เช่น เทคนิค Linear regression (LR), Artificial Neural Network (ANN), Sequential Minimal Optimization for Regression (SMOreg), Support Vector Machine Regression (SVR) และ Gussian Process (GP) ผู้วิจัยสนใจศึกษาการประยุกต์เทคนิคการพยากรณ์ที่เหมาะสมการวิเคราะห์อนุกรมเวลาในการพยากรณ์พื้นที่จังหวัดขอนแก่นเพื่อเปรียบเทียบสมรรถนะของแบบจำลองการเกิดอุบัติเหตุจากอุบัติเหตุทางถนนในพื้นที่จังหวัดขอนแก่น คณะผู้วิจัยได้ใช้หลักการ Sliding Window ในการแบ่งกลุ่มข้อมูลเป็นชุดข้อมูลเรียนรู้ และชุดข้อมูลทดสอบ และวัดประสิทธิภาพการพยากรณ์ของแบบจำลอง เพื่อหาเทคนิคการพยากรณ์ที่มีความแม่นยำมากที่สุดมาใช้ในการพัฒนาแบบจำลอง

1. วัตถุประสงค์การวิจัย

1. เพื่อศึกษาเทคนิคอนุกรมเวลาที่เหมาะสมในการเปรียบเทียบสมรรถนะของแบบจำลองการเกิดอุบัติเหตุทางถนน

2. เพื่อเปรียบเทียบประสิทธิภาพผลการพยากรณ์การเกิดอุบัติเหตุบนท้องถนน

2. เอกสารและงานวิจัยที่เกี่ยวข้อง

2.1 เอกสารที่เกี่ยวข้อง

ข้อมูลอนุกรมเวลา(Time Series data) คือ ชุดของข้อมูลที่เก็บรวบรวมตามระยะเวลาเป็นช่วง ๆ อย่างต่อเนื่องกัน เช่น ข้อมูลยอดขายสินค้าที่เก็บรวบรวมต่อเนื่องกันไปเป็นระยะเวลาหลาย ๆ เดือน ข้อมูลรายได้ประชาชาติปีต่าง ๆ ที่เก็บรวบรวมต่อเนื่องกันไปเป็นระยะเวลาหลาย ๆ ปี เป็นต้น ข้อมูลอนุกรมเวลาอาจอยู่ใน

ลักษณะที่เป็นข้อมูลรายปี รายไตรมาส หรือรายเดือนก็ได้ ทั้งนี้ขึ้นอยู่กับความเหมาะสมในการนำไปใช้ประโยชน์ เนื่องจากข้อมูลทางธุรกิจมีการเปลี่ยนแปลงตลอดเวลา ผู้นำทางธุรกิจหรือองค์กรต้องหาวิธีพัฒนาต่าง ๆ ที่สามารถนำไปใช้ประกอบการตัดสินใจวางแผน เกี่ยวกับผลที่เกิดจากความเปลี่ยนแปลงในการดำเนินการอย่างใดอย่างหนึ่ง ดังนั้นการวิเคราะห์อนุกรมเวลาจึงเข้ามามีบทบาทช่วยในการตัดสินใจ เทคนิคอย่างหนึ่งที่ใช้ช่วยในการควบคุมการดำเนินการในปัจจุบันและในการวางแผนความต้องการในอนาคต คือ การพยากรณ์ (forecasting) ซึ่งการพยากรณ์นั้นทำได้หลายวิธี แต่ละวิธีต่างมีเป้าหมายเดียวกัน คือ ทำนายเหตุการณ์ในอนาคต

เทคนิคการวิเคราะห์การถดถอย (Linear Regression) เป็นเทคนิคที่ใช้ในการวิเคราะห์การถดถอย Regression Analysis แบ่งออกเป็น สองชนิดคือ Linear Regression, Non-linear Regression การวิเคราะห์ Linear Regression แบ่งออกเป็น Simple Linear Regression (SLR), Multiple Linear Regression (MLR) การวิเคราะห์การถดถอยอย่างง่าย (Simple Regression Analysis) เป็นการศึกษาความสัมพันธ์ระหว่างตัวแปร 2 ตัว (ในที่นี้คือตัวแปร X และ Y) ที่มีความสัมพันธ์กันในลักษณะเชิงเส้น (Linear) เป็นวิธีการทางสถิติอย่างหนึ่ง ที่ใช้ในการตรวจสอบลักษณะของความสัมพันธ์ระหว่างตัวแปร ตั้งแต่ 2 ตัวขึ้นไป โดยแบ่งเป็นตัวแปรอิสระและตัวแปรตาม ซึ่งวราพร งามสุข [5] ได้ทำการศึกษาเพื่อเปรียบเทียบการพยากรณ์อนุกรมเวลาด้วยวิธีการถดถอย โดยใช้ข้อมูลสถิติจำนวนการเกิดอุบัติเหตุในประเทศไทย ระหว่างเดือนมกราคม พ.ศ. 2545 ถึง เดือนธันวาคม พ.ศ. 2554 พบว่าค่าประสิทธิภาพที่ได้ RMSE เท่ากับ 51.59% ในงานวิจัยนี้ได้ใช้ Linear Regression (LR)

เทคนิค Artificial Neural Networks [3] เป็นศาสตร์แขนงหนึ่งทางด้านปัญญาประดิษฐ์ (Artificial Intelligence : AI) ที่สามารถนำไปประยุกต์ใช้กับงานหลายด้านได้อย่างมีประสิทธิภาพ หลักการสำคัญของโครงข่ายประสาทเทียม คือ ความพยายามที่จะลอกเลียนแบบการทำงานของเซลล์ประสาทในสมองมนุษย์เพื่อทำงานได้อย่างมีประสิทธิภาพ ลักษณะทั่วไปของโครงข่ายประสาทเทียม คือ การที่โหนด (node) ต่าง ๆ จำลองมาจากไซแนป (synapse) ของเซลล์ประสาทระหว่าง เดนไดรท์ (dendrite) และแอกซอน (axon) โดยมีฟังก์ชันเป็นตัวกำหนดสัญญาณส่งออก (activation function or transfer function) นั่นเอง ลักษณะของโครงข่ายประสาทเทียมสามารถแบ่งได้ 2 แบบ คือ 1) โครงข่ายประสาทเทียมแบบ ชั้นเดียว (single layer) ซึ่งจะมีเพียงชั้นสัญญาณประสาทขาเข้า และชั้นสัญญาณประสาทขาออก เท่านั้น เช่น โครงข่ายเพอเซบตรอนอย่างง่าย (simple perceptron) และโครงข่ายโฮปฟิลด์ (hopfield networks) เป็นต้น และ 2) โครงข่ายประสาทเทียมแบบหลายชั้น (multilayer) ซึ่งมีลักษณะเช่นเดียวกับโครงข่ายประสาทเทียมแบบชั้นเดียว แต่จะมีชั้นแอบแฝง (hidden) เพิ่มขึ้นโดยอยู่ส่วนกลางระหว่างชั้นนำข้อมูลป้อนเข้าและชั้นส่งขอมูลออก ทั้งนี้ชั้นแอบแฝงอาจมีมากกว่า 1 ชั้น ในงานวิจัยนี้ได้ใช้ multilayer ในการสร้างแบบจำลอง

เทคนิค Sequential Minimal Optimization for Regression (SMOreg) ถูกพัฒนาจาก SMO อัลกอลิทึม บนพื้นฐานของ SVM (Support Vector Machine) วิธีนี้ใช้การแทนค่าที่ข้อมูลสูญหายและแปลงข้อมูลคุณลักษณะเชิงกลุ่ม (nominal) ให้เป็นข้อมูลไบนารี (Binary)นอกจากนี้ยังทำให้ข้อมูลคุณลักษณะทุกค่าอยู่ในรูปแบบมาตรฐาน (Normalized) SMOreg ยังมีคุณสมบัติใช้งานกับ non-linear ได้อย่างมีประสิทธิภาพ SMOreg ยังช่วยในการจัดการโครงสร้างของ Model และลดความเสี่ยงของข้อมูลทำให้ข้อมูลมีความน่าเชื่อถือในการพยากรณ์

เทคนิค Support Vector Machine (SVM) เป็นเทคนิคการจำแนกประเภทของข้อมูลที่มีพื้นฐานมาจากทฤษฎีการเรียนรู้จากสถิติ คล้ายเทคนิคโครงข่ายประสาทเทียม แต่ SVM ใช้หลักการลดค่าความเสี่ยงเชิงโครงสร้างให้ต่ำสุด (Structural Risk Minimization) เพื่อลดค่าความผิดพลาดของการทำนาย (Minimized Error) พร้อมกับเพิ่มระยะการแบ่งแยกให้มากที่สุด (Maximized Margin)หลักการของ SVM คือการหาสัมประสิทธิ์ของสมการเพื่อสร้างเส้นแบ่งแยกประเภทข้อมูล ทำโดยการเลือกเส้นหรือระนาบแบ่งแยกประเภทข้อมูลที่เหมาะสมที่สุด A Library for Support Vector Machines (LibSVM) [4] เป็นไลบรารี่สำหรับ SVM ที่พัฒนาโดย Chih-Chung Chang และ Chih-Jen Lin ซึ่ง LibSVM นี้จะช่วยให้การวิเคราะห์SVM นั้นง่ายขึ้น โดย LibSVM ได้รองรับซัพพอร์ตเวกเตอร์แบบการแบ่งกลุ่ม (Support Vector Classification: SVC) แบบการวิเคราะห์แบบรีเกรสชัน (Support Vector

Regression: SVR) Sven และ Rohit [5] ได้สร้างแบบจำลองเพื่อการพยากรณ์ข้อมูลอนุกรมเวลาในรูปแบบฤดูกาล จำนวน 8 ชุดข้อมูลด้วยโครงข่ายประสาทเทียม โดยทดลองกับโมเดลโครงข่ายประสาทเทียมแบบต่างๆ แสดงให้เห็นว่าโครงข่ายประสาทเทียมมีประสิทธิภาพในการแก้ปัญหาการพยากรณ์อนุกรมเวลาได้เป็นอย่างดี

เทคนิค Gussian Precesses (GP) [6]เป็นเทคนิคที่นำเอาระบบ Stochastic Process ซึ่ง ปัจจุบันถูกยอมรับแล้วว่าเป็นเครื่องมือที่มีประสิทธิภาพในการแก้ปัญหาประเภท Regression, Classification, และ Decision ใน Machine learningสามารถทำงานได้ดีถึงแม้ว่ามี Training Data น้อยและมีประสิทธิภาพและ Convergence rate ดีกว่า ARMA, NN และ SVR ตามลำดับ[7] ซึ่งมีงานวิจัยของทัศนัย พลอยสุวรรณและคณะ[8] นำเสนอการพยากรณ์หาค่าความต้องการไฟฟ้าสูงสุด (Peak load) ระยะยาวระหว่างปี 2011-2012โดยใช้ GaussianProcess ผลการพยากรณ์พบว่าให้ค่าความผิดพลาดน้อยและมีประสิทธิภาพดีกว่าที่ดีกว่าเมื่อเปรียบเทียบกับกับการใช้โครงข่ายประสาทเทียม (Neural Networks)

2.2 งานวิจัยที่เกี่ยวข้อง

รศ.ดร.นพ.พงษ์เทพ วิวรรณธนะเดชและคณะ[9] ได้ให้ความหมายของคำว่า อนุกรมเวลา ไว้ว่า เซตของข้อมูลเชิงปริมาณที่จัดเก็บในช่วงเวลาหนึ่ง ข้อมูลอนุกรมเวลาอาจจะอยู่ในลักษณะที่เป็นรายปี รายไตรมาส หรือรายเดือนก็ได้ ทั้งนี้ขึ้นอยู่กับความเหมาะสมในการนำไปใช้ประโยชน์ การวิเคราะห์อนุกรมเวลา[10] เป็นกระบวนการของการใช้เทคนิคทางสถิติเพื่อสร้างแบบจำลองและอธิบายชุดข้อมูลจุดขึ้นอยู่กับเวลา การคาดการณ์ชุดอนุกรมเวลาเป็นกระบวนการของการใช้แบบจำลองเพื่อสร้างการคาดการณ์ สำหรับเหตุการณ์ในอนาคตจากเหตุการณ์ในอดีตที่ผ่านมา ข้อมูลชุดเวลามีลำดับเวลาตามธรรมชาติซึ่งแต่ละข้อมูลเป็นตัวอย่างที่มีแนวคิดที่จะเรียนรู้และลำดับข้อมูล

Zainal และMustaffa[11] ได้นำเอาเทคนิค Artificial Neural Network มาใช้ในสร้างแบบจำลองเพื่อการพยากรณ์ราคาทองคำ โดยใช้ให้ค่า RMSE และ MAE ในการแสดงประสิทธิภาพของแบบจำลอง ผลปรากฏว่าเทคนิค Artificial Neural Network สามารถสร้างแบบจำลองที่มีประสิทธิภาพต่ำสุดแต่ไม่มีนัยสำคัญของความแตกต่างจากเทคนิคอื่น

Yang, Zhai, Xu และ Han[12] ได้ใช้ SMO ในสร้างแบบจำลองเพื่อการพยากรณ์จำนวนของจุดบอดบนดวงอาทิตย์ ของ Wolfer ผลปรากฏว่า เทคนิค SMO ดีกว่า SVM และ Neuaral Network

วิธีดำเนินการวิจัย

1. เครื่องมือการวิจัย

1.1 โปรแกรมที่ใช้ในการวิเคราะห์ค่าสถิติการพยากรณ์ ใช้โปรแกรม WEKA 3.9.1

1.2 เทคนิคที่นำมาพยากรณ์ จำนวน 5 เทคนิค ได้แก่ ได้แก่ เทคนิควิธีการวิเคราะห์การถดถอย (Regression Analysis) วิธีและโครงข่ายประสาทเทียม (Artificial Neural Network: ANN) เทคนิคSequential Minimal Optimization for Regression (SMOreg) เทคนิคซัพพอร์ตเวคเตอร์แมชชีนส์ (SVM) และเทคนิค Gussian Process (GP)

2.ขั้นตอนการดำเนินการวิจัยประกอบด้วย 4 ขั้นตอน ดังนี้

2.1 เก็บรวบรวมข้อมูลการเกิดของอุบัติเหตุระหว่าง เดือนมกราคมพ.ศ.2554 ถึง เดือนมิถุนายน 2560 ในพื้นที่จังหวัดขอนแก่น

2.2 จัดเตรียมข้อมูล โดยจัดเรียงข้อมูลตามช่วงเวลารายเดือนตามปัจจัยที่เกี่ยวข้องโดยข้อมูลได้ถูกเตรียมข้อมูลใน Excel

2.3 สร้างแบบจำลอง โดยการทำการทดสอบด้วยเทคนิค Linear regression (LR), Artificial Neural Network (ANN), Sequential Minimal Optimization for Regression (SMOreg), Support Vector Machine Regression (SVR) และ Gussian Process (GP)

2.4 วัดประสิทธิภาพ ด้วยการเปรียบเทียบค่าเฉลี่ยความคาดเคลื่อนสมบูรณ์และค่าความคลาดเคลื่อนเฉลี่ยกำลังสอง ของทั้ง 5 เทคนิค

3. ข้อมูลที่ใช้ในการวิจัย

ในงานวิจัยฉบับนี้ใช้ข้อมูลการเกิดของอุบัติเหตุรายเดือนในพื้นที่จังหวัดขอนแก่น โดยใช้ข้อมูล ย้อนหลังตั้งแต่ เดือนมกราคม2554-มิถุนายน 2560 จำนวน 7 ปี รวมทั้งหมด 78 เดือน

4.สถิติที่ใช้ในการวิจัย

การวิเคราะห์ประสิทธิภาพของแบบจำลองในงานวิจัยนี้ผู้วิจัยได้ใช้หลักการหน้าต่างเคลื่อนที่ (Sliding Window) คือ หลักการแบ่งข้อมูลออกเป็น ชุดข้อมูลที่ใช้ในการฝึกสอน หรือการเรียนรู้ (Training data) และ ชุดข้อมูลที่ใช้ในการทดสอบ (Testing data) รอบแรกข้อมูล 3 ปีแรก (พ.ศ. 2554 - พ.ศ. 2556) เป็นชุดข้อมูลฝึกสอน และให้หกเดือนเป็นชุดข้อมูลทดสอบ (เดือนมกราคม-มิถุนายน พ.ศ. 2555) แล้วจึงทำการเคลื่อนหน้าต่างข้อมูลไปรอบละ 6 เดือนรวมทั้งสิ้น 7 รอบ แล้วจึงมีการวัดค่าประสิทธิภาพของแบบจำลอง ด้วยค่าเฉลี่ยความคลาดเคลื่อนสัมบูรณ์ (Mean Absolute error: MAE) และค่าความคลาดเคลื่อนเฉลี่ยกำลังสอง (Root Mean Square Error: RMSE)

ผลการวิจัย

ผลการวิจัยในฉบับนี้ ผู้วิจัยได้ทำการศึกษาข้อมูล แล้วจึงใช้โปรแกรม WEKA 3.9.1 ในการทำการทดลองและวัดค่าประสิทธิภาพของแบบจำลองการเกิดอุบัติเหตุทางถนน ด้วยเทคนิคอนุกรมเวลา 5 เทคนิคได้แก่

1.1 เทคนิคการถดถอยเชิงเส้น (Linear Regression: LR)

1.2 เทคนิคโครงข่ายประสาทเทียม (Artificial Neural Network: ANN)

1.3 เทคนิค Sequential Minimal Optimization for Regression (SMOreg)

1.4 เทคนิคซัพพอร์ตเวคเตอร์แมชชีนส์ (SVR)

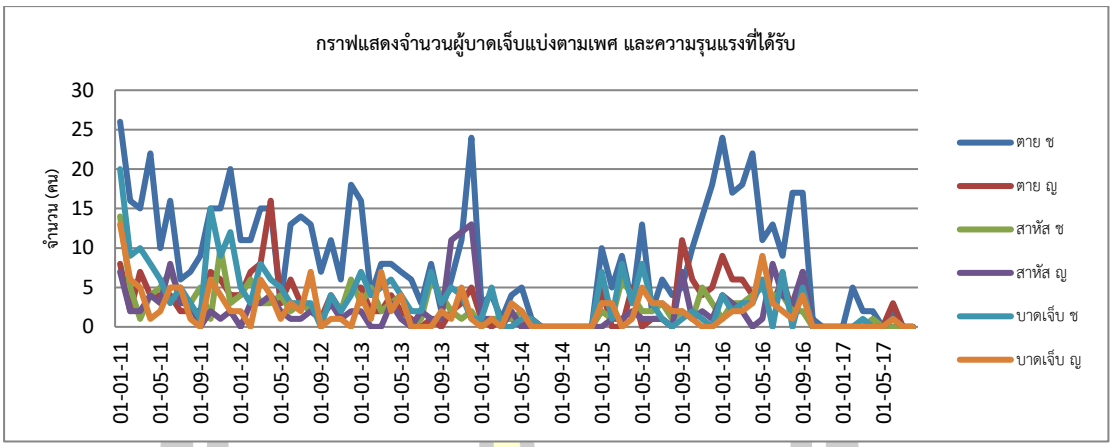
1.5 เทคนิค GussianProcess (GP)

โดยการเปรียบเทียบค่าเฉลี่ยความคลาดเคลื่อนสัมบูรณ์ (Mean Absolute error: MAE) และค่าความคลาดเคลื่อนเฉลี่ยกำลังสอง (Root Mean Square Error: RMSE)

1. ผลการศึกษาข้อมูลการเกิดอุบัติเหตุ

จากการศึกษาการเกิดอุบัติเหตุการเกิดของอุบัติเหตุระหว่างเดือน มกราคม พ.ศ. 2554 ถึง เดือนมิถุนายน 2560 ในพื้นที่จังหวัดขอนแก่น สามารถแสดงได้ดังภาพที่ 1.1

พหุน ปณ ฑิโต ชีเว

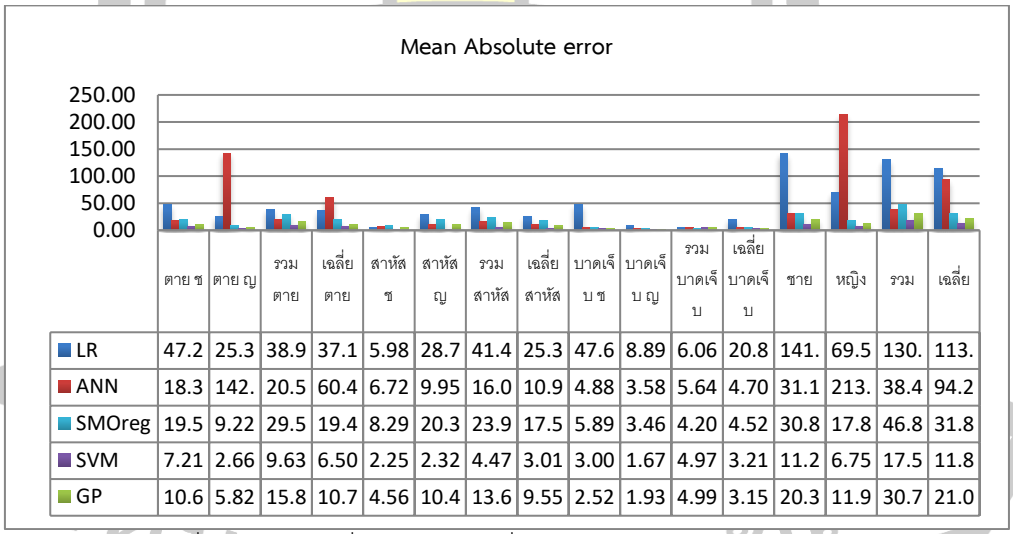


ภาพที่ 1 ข้อมูลผู้ประสบเหตุ แบ่งเพศ และระดับความรุนแรงที่เกิดขึ้น

จากภาพที่ 1 พบว่าข้อมูลอยู่ในลักษณะไม่เส้นตรง (Non-Linear) จากการเกิดอุบัติเหตุในแต่ละเดือนโดยแยกตามเพศและความรุนแรง

2.ผลการเปรียบเทียบค่าเฉลี่ยความคลาดเคลื่อนสัมบูรณ์ของแบบจำลองการเกิดอุบัติเหตุทางถนน ของจังหวัดขอนแก่นโดยใช้เทคนิคอนุกรมเวลา

การเปรียบเทียบค่าเฉลี่ยความคลาดเคลื่อนสัมบูรณ์ (MAE) ซึ่งเป็นค่าความแตกต่างโดยเฉลี่ยระหว่างการเกิดอุบัติเหตุทางถนนจากการพยากรณ์ และเกิดอุบัติเหตุทางถนนจริง สามารถแสดงได้ภาพที่ 2



	ตาย ช	ตาย ญ	รวมตาย	เฉลี่ยตาย	สาหัส ช	สาหัส ญ	รวมสาหัส	เฉลี่ยสาหัส	บาดเจ็บ ช	บาดเจ็บ ญ	รวมบาดเจ็บ	เฉลี่ยบาดเจ็บ	ชาย	หญิง	รวม	เฉลี่ย
LR	47.2	25.3	38.9	37.1	5.98	28.7	41.4	25.3	47.6	8.89	6.06	20.8	141.	69.5	130.	113.
ANN	18.3	142.	20.5	60.4	6.72	9.95	16.0	10.9	4.88	3.58	5.64	4.70	31.1	213.	38.4	94.2
SMOreg	19.5	9.22	29.5	19.4	8.29	20.3	23.9	17.5	5.89	3.46	4.20	4.52	30.8	17.8	46.8	31.8
SVM	7.21	2.66	9.63	6.50	2.25	2.32	4.47	3.01	3.00	1.67	4.97	3.21	11.2	6.75	17.5	11.8
GP	10.6	5.82	15.8	10.7	4.56	10.4	13.6	9.55	2.52	1.93	4.99	3.15	20.3	11.9	30.7	21.0

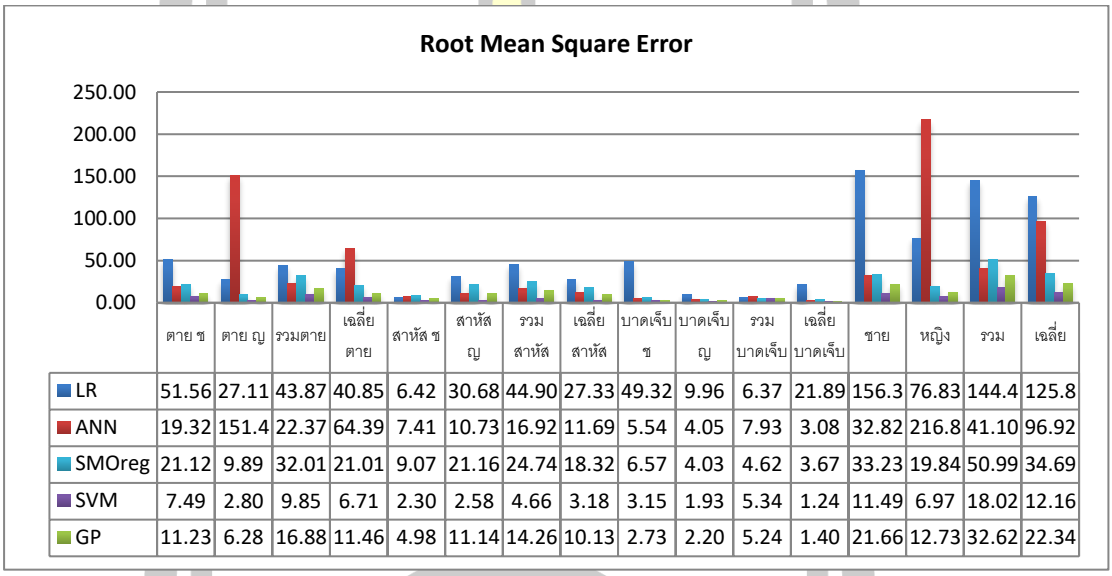
ภาพที่ 2 แสดงค่าเฉลี่ยความคลาดเคลื่อนสัมบูรณ์ (MAE) ของแต่ละเทคนิค

จากภาพที่ 2 จะพบว่าค่าเฉลี่ยความคาดเคลื่อนสมบูรณ์ (Mean Absolute error: MAE) เป็นค่าที่ใช้วัดความถูกต้องของการพยากรณ์ที่นิยมใช้กันอย่างแพร่หลาย จากการทดลองสามารถแสดงผลการเปรียบเทียบของค่า MAE ของ 5 แบบจำลอง LR, ANN, SMOreg, SVM และ GP ในหัวข้อนี้ได้ทำการเปรียบเทียบประสิทธิภาพของแบบจำลองเพื่อพยากรณ์จำนวนผู้บาดเจ็บ บาดเจ็บสาหัส และเสียชีวิตจากการเกิดอุบัติเหตุบนท้องถนน ของเพศชาย (M) เพศหญิง (F) และรวมทั้งชายและหญิง ดัง ภาพที่ 1 ในการพยากรณ์จำนวนผู้บาดเจ็บจากการเกิดอุบัติเหตุบนท้องถนนของเพศชาย (M) เพศหญิง (F) และรวมทั้งชายและหญิง พบว่า SVM สามารถพยากรณ์แล้วให้

ค่า MAE ต่ำสุดเพียง 3.00 ในการพยากรณ์การบาดเจ็บของชาย 1.67 การบาดเจ็บของเพศหญิง และ 4.97 การบาดเจ็บโดยรวม โดยมีค่าเฉลี่ยรวมเป็น 3.21 มีค่า 2.25 ในการพยากรณ์การบาดเจ็บสาหัสเพศชาย 2.32 บาดเจ็บสาหัสเพศหญิง 4.47 บาดเจ็บสาหัสรวม โดยมีค่าเฉลี่ยรวมเป็น 3.01 และ 7.21 การพยากรณ์การเสียชีวิตเพศชาย 2.66 ในการเสียชีวิตเพศหญิง 9.62 การเสียชีวิตรวมทั้งชายและหญิง โดยมีค่าเฉลี่ยรวมเป็น 6.50

3.ผลการเปรียบเทียบค่าความคลาดเคลื่อนเฉลี่ยกำลังสอง ของแบบจำลองการเกิดอุบัติเหตุทางถนน ของจังหวัดขอนแก่นโดยใช้เทคนิคอนุกรมเวลา

การเปรียบเทียบค่าความคลาดเคลื่อนเฉลี่ยกำลังสอง ซึ่งเป็นการเปรียบเทียบความแตกต่างระหว่างค่าพยากรณ์และค่าจริง ของการเกิดอุบัติเหตุทางถนนเฉลี่ยกำลังสอง ซึ่งถ้าค่าที่ได้เท่ากับ 0 แสดงว่าค่าพยากรณ์ตรงกับค่าความเป็นจริงที่สุดซึ่งหมายถึงดีที่สุด จากการทดลองสามารถแสดงผลการทดลองได้ดังภาพที่ 3



	ตาย ช	ตาย ญ	รวมตาย	เฉลี่ยตาย	สาหัส ช	สาหัส ญ	รวมสาหัส	เฉลี่ยสาหัส	บาดเจ็บ ช	บาดเจ็บ ญ	รวมบาดเจ็บ	เฉลี่ยบาดเจ็บ	ชาย	หญิง	รวม	เฉลี่ย
LR	51.56	27.11	43.87	40.85	6.42	30.68	44.90	27.33	49.32	9.96	6.37	21.89	156.3	76.83	144.4	125.8
ANN	19.32	151.4	22.37	64.39	7.41	10.73	16.92	11.69	5.54	4.05	7.93	3.08	32.82	216.8	41.10	96.92
SMOreg	21.12	9.89	32.01	21.01	9.07	21.16	24.74	18.32	6.57	4.03	4.62	3.67	33.23	19.84	50.99	34.69
SVM	7.49	2.80	9.85	6.71	2.30	2.58	4.66	3.18	3.15	1.93	5.34	1.24	11.49	6.97	18.02	12.16
GP	11.23	6.28	16.88	11.46	4.98	11.14	14.26	10.13	2.73	2.20	5.24	1.40	21.66	12.73	32.62	22.34

ภาพที่ 3 ความคลาดเคลื่อนเฉลี่ยกำลังสอง (RMSE) ของแต่ละเทคนิค

ค่าความคลาดเคลื่อนเฉลี่ยกำลังสอง (RMSE) เป็นวิธีการวัดความคลาดเคลื่อนจากค่าที่พยากรณ์จากแบบจำลองกับค่าจริงที่เกิดขึ้น ในหัวข้อนี้ได้ทำการเปรียบเทียบประสิทธิภาพของแบบจำลองเพื่อพยากรณ์จำนวนผู้บาดเจ็บ บาดเจ็บสาหัส และเสียชีวิตจากการเกิดอุบัติเหตุบนท้องถนนโดยใช้ 5 เทคนิค แสดงค่าความคลาดเคลื่อนเฉลี่ยกำลังสอง (RMSE) ซึ่งเป็นค่าที่ใช้วัดประสิทธิภาพของ 5 แบบจำลองรวมถึง LR, ANN, SMOreg, SVM และ GP ในหัวข้อนี้ได้ทำการเปรียบเทียบประสิทธิภาพของแบบจำลองเพื่อพยากรณ์จำนวนผู้บาดเจ็บบาดเจ็บสาหัส การเสียชีวิตจากการเกิดอุบัติเหตุบนท้องถนน ของเพศชาย (M) เพศหญิง (F) และรวมทั้งชายและหญิง ดังภาพที่ 3 ในการพยากรณ์จำนวนผู้บาดเจ็บจากการเกิดอุบัติเหตุบนท้องถนนของเพศชาย (M) เพศหญิง (F) และรวมทั้งชายและหญิง พบว่า SVM สามารถพยากรณ์แล้วให้ค่า RMSE ต่ำสุดเพียง 3.15 ในการพยากรณ์การบาดเจ็บของชาย 1.93 การบาดเจ็บของเพศหญิง และ 5.34 การบาดเจ็บโดยรวม โดยมีค่าเฉลี่ยรวมเป็น 1.24 มีค่า 2.30 ในการพยากรณ์การบาดเจ็บสาหัสเพศชาย 2.58 บาดเจ็บสาหัสเพศหญิง 4.66 บาดเจ็บสาหัสรวม โดยมีค่าเฉลี่ยบาดเจ็บรวมเป็น 3.18 และ 7.49 ในการพยากรณ์การเสียชีวิตเพศชาย 2.80 เสียชีวิตเพศหญิง 9.85 การเสียชีวิตรวมทั้งชายและหญิง โดยมีค่าเฉลี่ยรวมเป็น 6.71

อภิปรายผลการวิจัย

ผู้วิจัยได้ศึกษาแบบจำลองที่เหมาะสมสำหรับการพยากรณ์การเกิดอุบัติเหตุทางถนน โดยใช้เทคนิคอนุกรมเวลา โดยใช้ 5 เทคนิค ผลการเปรียบเทียบเทคนิคการพยากรณ์การเกิดอุบัติเหตุโดยใช้ด้วยค่าเฉลี่ยความคลาดเคลื่อนสัมบูรณ์ (Mean Absolute error: MAE) และค่าความคลาดเคลื่อนเฉลี่ยกำลังสอง (Root Mean Square Error: RMSE) พบว่า ส่วนเทคนิคLR มีค่าของการบาดเจ็บเฉลี่ย 20.87,21.89 บาดเจ็บสาหัสเฉลี่ย 25.39, 27.33 และการเสียชีวิตเฉลี่ยอยู่ที่ 37.18, 40.85,เทคนิค ANNค่าของการบาดเจ็บเฉลี่ย 4.70, 3.08 บาดเจ็บสาหัสเฉลี่ย 10.91, 11.69 และการเสียชีวิตเฉลี่ยอยู่ที่ 60.47, 64.39, เทคนิค SMOreg ค่าของการบาดเจ็บเฉลี่ย 4.52 ,3.67 บาดเจ็บสาหัสเฉลี่ย 17.52 , 18.32 และการเสียชีวิตเฉลี่ยอยู่ที่ 19.45, 21.01, เทคนิค SVM มีค่าของการบาดเจ็บเฉลี่ย 3.21, 1.24 บาดเจ็บสาหัสเฉลี่ย 3.01 , 3.18 และการเสียชีวิตเฉลี่ยอยู่ที่ 6.50, 6.71 และ เทคนิค GP ค่าของการบาดเจ็บเฉลี่ย 3.15 , 1.40 บาดเจ็บสาหัสเฉลี่ย 9.55, 10.13 และการเสียชีวิตเฉลี่ยอยู่ที่ 10.75, 11.46, วิธีการสรุปผลการวิเคราะห์พบว่า เทคนิค SVM ค่าเฉลี่ยความคลาดเคลื่อนสัมบูรณ์ (Mean Absolute error: MAE) และค่าความคลาดเคลื่อนเฉลี่ยกำลังสอง เป็นวิธีที่มีค่าความคลาดเคลื่อนต่ำที่สุด มีความเหมาะสมในการพัฒนาแบบจำลองเพื่อพยากรณ์การเกิดอุบัติเหตุบนท้องถนนโดยใช้เทคนิคอนุกรมเวลา และ เพื่อใช้ในการวางแผนในการเดินทางและได้หลีกเลี่ยงการเกิดอุบัติเหตุ

ข้อเสนอแนะ

งานวิจัยนี้ได้มีการศึกษาเทคนิคการพยากรณ์จำนวน 5 เทคนิค โดยเทคนิคที่ทดสอบแล้วได้ค่าการพยากรณ์ที่คลาดเคลื่อนน้อยที่สุดในเอกสารฉบับนี้ หากมีการศึกษาเทคนิควิธีอื่นเข้ามาเปรียบเทียบเพิ่มเติม ทำให้ได้ค่าการพยากรณ์ที่ต่างกันในแต่ละเทคนิค โดยการนำข้อมูลการเกิดอุบัติเหตุโดยใช้ข้อมูลย้อนหลัง 7 ปี โดยแยกข้อมูลทดสอบเป็นรายเดือน ในงานวิจัยที่มีข้อมูลใกล้เคียงกันก็สามารถนำไปปรับใช้ให้เหมาะสมกับการพยากรณ์

เอกสารอ้างอิง

[1] ศูนย์ข้อมูลอุบัติเหตุ. (2560). รายงานสถิติการใช้สิทธิ พรบ. รายจังหวัด. สืบค้นเมื่อ 23 กันยายน 2560, จาก http://rvpreport.rvpeservice.com/viewrsc.aspx?report=0486&session=16

[2] ศูนย์วิจัยอุบัติเหตุแห่งประเทศไทย. (2552). การพัฒนาแบบจำลองการเกิดอุบัติเหตุ. สืบค้นเมื่อ 26 กันยายน 2560, จาก http://bhs.doh.go.th/files/Project/TACR/Accident %20Prediction %20Model_th.pdf

[3] วราพร งามสุข. (2555). การพยากรณ์อนุกรมเวลาโดยการเปรียบเทียบวิธีแบบฉบับและวิธีบอกซ์-เจนกินส์ กรณีศึกษาจำนวนการเกิดอุบัติเหตุในประเทศไทย (ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์), มหาวิทยาลัยบูรพา.

[4] ธนาวุฒิ ประกอบผล. (2552). โครงข่ายประสาทเทียม Artificial Neural Networks (Vol. 12).

[5] Chih-Chung Chang และand Chih-Jen Lin. (April 2011). A Library for Support Vector Machines (LibSVM). ACM Transactions on Intelligent Systems and Technology (TIST), 2(3).

[6] Sven F. Crone และRohit Dhawan. (2007). Forecasting Seasonal Time Series with Neural Networks: A Sensitivity Analysis of Architecture Parameters. IEEE Trans.Neural Networks(IJCNN), 2099-2104.

[7] Amine Bermark Sofiane Brahim-Belhouari. (2004). Gaussian process for nonstationary time series prediction. Computational Statics & Data Analysis 47(47), 705-712.

[8] C. E. Rasmussen and C. K. I. Williams. (2006). Gaussian processes for machine learning. Cambridge, Mass.: MIT Press.

[9] ทัศนัย พลอยสุวรรณ และธนวันต์ ยลพัธน์. (2556). การพยากรณ์ค่าความต้องการใช้ไฟฟ้าสูงสุดของการไฟฟ้าฝ่ายผลิตแห่งประเทศไทยโดยกระบวนการเกาส์เซี่ยน. การประชุมและนำเสนอผลงานวิชาการทางวิศวกรรมไฟฟ้า ครั้งที่ 36.

[10] พงษ์เทพ วิวรรณธนะเดช และคณะแพทย์ศาสตร์. (2550). การวิเคราะห์อนุกรมเวลา(Time Series Analysis)และระบบสารสนเทศทางภูมิศาสตร์(GIS)ในการพัฒนาองค์ความรู้ด้านสิ่งแวดล้อมและสุขภาพ. ศูนย์ประสานข้อมูลปัญหาหมอกควันในพื้นที่ภาคเหนือภาควิชาเวชศาตร์ชุมชน คณะแพทย์ศาตร์ มหาวิทยาลัยเชียงใหม่, pp.1-35.

[11] Adaptavist Theme. (2560). Time Series Analysis and Forecasting with Weka. Retrieved 23 September 2017, from http://wiki.pentaho.com/display/DATAMINING/Time+Series+Analysis+and+Forecasting+with+Weka

[12] Mustaffa Z Zainal NA. (2016). Developing a gold price predictive analysis using Grey Wolf Optimizer. IEEE Student Conference on Research and Development 2016.

[13] Jin-Fang Yang; Yong-Jie Zhai; Da-Ping Xu. (2007). SMO Algorithm Applied in Time Series Model Building and Forecast. Pu Han 2007 International Conference on Machine Learning and Cybernetics, 4, 2395 – 2400.

พหูน ปณ ฑิโต ชีเว

ประวัติผู้เขียน

ชื่อ	นางสาวปทิตญา บุญรักษา
วันเกิด	วันที่ 15 กันยายน พ.ศ. 2527
สถานที่เกิด	อำเภอด่านซ้าย จังหวัดเลย
สถานที่อยู่ปัจจุบัน	บ้านเลขที่ 106 หมู่ 8 ตำบลด่านซ้าย อำเภอด่านซ้าย จังหวัดเลย รหัสไปรษณีย์ 42120
ตำแหน่งหน้าที่การงาน	Administrator
สถานที่ทำงานปัจจุบัน	บริษัท เอ็นทีที ดาต้า (ประเทศไทย) จำกัด เลขที่ 89 อาคารคอสโม ออฟฟิศ พาร์ค ชั้น 7 ตำบลบ้านใหม่ อำเภอปากเกร็ด จังหวัดนนทบุรี รหัสไปรษณีย์ 11120
ประวัติการศึกษา	พ.ศ. 2545 มัธยมศึกษาตอนปลาย โรงเรียนศรีสองรักษ์วิทยา ตำบลด่านซ้าย อำเภอด่านซ้าย จังหวัดเลย พ.ศ. 2550 ปริญญาวิทยาศาสตรบัณฑิต (วท.บ.) สาขาวิชาสื่อนฤมิต มหาวิทยาลัยมหาสารคาม พ.ศ. 2561 ปริญญาวิทยาศาสตรมหาบัณฑิต (วท.ม.) สาขาวิชาเทคโนโลยีสารสนเทศ มหาวิทยาลัยมหาสารคาม

พหูน ปณฺฑิโต ชีเว