

การทำนายสภาพจราจรในเขตกรุงเทพฯ ชั้นใน ด้วยการเรียนรู้เชิงลึก Predicting Traffic Status of Inner Bangkok with Deep Learning

อมรเทพ ทองชีว (Amornthep Thongchiw)* ภัทราวดี มากมี (Pattrawadee Makmee)**
และ ยรรยงค์ พันธุ์สวัสดิ์ (Yunyong Punsawad)***

Received: September 6, 2022

Revised: October 18, 2022

Accepted: November 11, 2022

*ผู้นิพนธ์ประสานงาน: ภัทราวดี มากมี (Pattrawadee Makmee) อีเมล: pattrawadee@gmail.com

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพในการทำนายสภาพจราจรทางบก ในเขตกรุงเทพฯ ชั้นใน ระหว่างตัวแบบการเรียนรู้เชิงลึกที่ได้ปรับพารามิเตอร์ไว้อย่างเหมาะสม กับตัวแบบที่เป็นอัลกอริทึมทั่วไปในการดำเนินการวิจัยผู้วิจัยใช้โปรแกรมคอมพิวเตอร์ที่พัฒนาขึ้นเก็บข้อมูลสถานะการจราจรจากกูเกิลแมพ ในเขตกรุงเทพฯ ชั้นใน ต่อเนื่องทุกครึ่งชั่วโมง เป็นเวลา 2 เดือน ได้ข้อมูลภาพจราจร 1,844 ภาพ ผสมกับการแปลงข้อมูลเกี่ยวกับช่วงเวลา มาสร้างเป็นตัวแปรอิสระเพิ่มขึ้น ผลการวิจัยพบว่า ค่าร้อยละความถูกต้องการทำนาย (Accuracy) เมื่อทดสอบกับชุดข้อมูลทดสอบ (Test set) ในอัลกอริทึมต่าง ๆ คือ Decision Tree: 71%, Random Forest: 71%, K Nearest Neighbors: 73%, Naive Bayes: 63%, Gradient Boosting: 72% และตัวแบบการเรียนรู้เชิงลึก (Deep Learning): 75% โดยตัวแบบที่ได้ผลการทำนายสูงสุดคือ การเรียนรู้เชิงลึก การวิจัยนี้มีผลผลิตเป็นกรรมวิธีเชิงปฏิบัติในการเก็บข้อมูลจราจรจากภาพบนเว็บไซต์กูเกิลแมพ การจัดเตรียมข้อมูลภาพให้พร้อมสำหรับการเรียนรู้ด้วยเครื่อง (Machine Learning) และตัวแบบการทำนายสภาพจราจรด้วยการเรียนรู้เชิงลึก ซึ่งผลลัพธ์ของการทำนายคือ สภาพการจราจรทางบก ที่ทำนายล่วงหน้า 1 ชั่วโมง ณ จุดที่ศึกษา คือ บริเวณย่านถนนสาทร ระบุมการทำนายเป็นคลาส ตามระดับสีของกูเกิล แมพ ที่มีอยู่จำนวน 4 คลาส ได้แก่ สีเขียว หมายถึง สภาพการจราจรคล่องตัว, สีส้ม หมายถึง เคลื่อนตัวได้,

สีแดง หมายถึง การจราจรหนาแน่น และสีแดงเข้ม หมายถึง การจราจรติดขัดมาก โดยตัวแบบจะทำนายสภาพจราจรที่มีความน่าจะเป็นสูงสุดออกมา ผลการทำนายที่ได้ มีประโยชน์ต่อประชาชนผู้เดินทางหรือกิจการขนส่งโลจิสติกส์ ที่ใช้เส้นทางได้ทราบสภาพการจราจรล่วงหน้า รวมถึงผู้ที่สนใจศึกษาข้อมูลการจราจรด้วยการเรียนรู้ด้วยเครื่อง

คำสำคัญ: สภาพการจราจร การทำนาย การจำแนกประเภท การเรียนรู้ด้วยเครื่อง การเรียนรู้เชิงลึก

Abstract

The objectives of this research was to compare the efficiency of predicting road traffic status in the inner Bangkok area, between the properly parameterized deep learning models with different conventional algorithms. In process of research methodology, the researcher used a computer program developed to collect traffic situation images from Google Maps specific inner Bangkok area, every half hour for 2 months, 1,844 traffic images were obtained, combined with duration of the day to create more independent variables. The findings revealed that percentage of prediction accuracy when predicted with test data sets in different algorithms as follows, Decision Tree: 71%, Random Forest: 71%, K Nearest Neighbors: 73%, Naive Bayes: 63 %, Gradient Boosting: 72% and Deep Learning: 75%. This research proposed a practical

*สาขาวิชาการศึกษาและสถิติทางวิทยาการปัญญา วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา

*Graduate Student in Research and Statistics in Cognitive Science, College of Research Methodology and Cognitive Science, Burapha University.

**หน่วยวิจัยวิทยาการปัญญาและนวัตกรรม วิทยาลัยวิทยาการวิจัยและวิทยาการปัญญา มหาวิทยาลัยบูรพา

**Cognitive Science and Innovation Research Unit: CSIRU, College of Research Methodology and Cognitive Science, Burapha University.

***สำนักวิชาสารสนเทศศาสตร์ มหาวิทยาลัยวลัยลักษณ์

***School of Informatics, Walailak University.

method for collecting traffic data from images on Google Maps websites. Preprocessing data ready for machine learning and a model for predicting traffic status with deep learning. Which the result of the prediction is the road traffic condition predicted 1 hour in advance at the study point was the area of Sathorn Road. Specify a prediction as a class, according to the color scale of Google Maps, there are 4 classes, namely, green means active traffic, orange means moving, red means heavy traffic and dark red means traffic is congested very much. The model predicts the most probable traffic status. The results of the predictions are useful for people traveling or transportation logistics companies who use routes to know the traffic status in advance, including those interested in studying traffic data with machine learning.

Keywords: Traffic status, Prediction, Classification, Machine Learning, Deep Learning.

1. บทนำ

ในพื้นที่เขตเมืองต่าง ๆ ทั่วโลกได้มีการกล่าวถึงปัญหาจราจรติดขัดมาโดยตลอด สำหรับกรุงเทพฯ ข้อมูลจากศูนย์วิจัยกสิกรไทย ระบุว่าสถิติอัตราการใช้พลังงานเพิ่มขึ้นถึง 70% เมื่อเปรียบเทียบกับจากปี ค.ศ. 2010 ถึง ค.ศ. 2015 ทำให้เกิดความสูญเสียเพิ่มขึ้น หากใช้เวลาเดินทางนานขึ้น 35 นาที จะทำให้มีมูลค่าสูญเสียคิดเป็นเงินถึง 11,000 ล้านบาท หรือ 60 ล้านบาท/วัน การใช้เชื้อเพลิงเพิ่มขึ้นคิดเป็นเงิน 6,000 ล้านบาทต่อปี นอกจากนี้ผู้ที่ประสบกับปัญหาจราจรติดขัดจะเสียสุขภาพจิต และต้องปรับพฤติกรรมสูกับชีวิตที่เร่งรีบต้องใช้ชีวิตในกิจกรรมต่าง ๆ [1] จากปัญหาเหล่านี้มีนักวิจัยสนใจศึกษาเกี่ยวกับการทำนายสภาพการจราจรเพื่อคาดการณ์สภาพการจราจรล่วงหน้าและวางแผนการเดินทางให้เหมาะสมด้วยอัลกอริทึมสมัยใหม่ เช่น Neuron Network, k-NN, SVM และในปัจจุบันได้เริ่มมีการใช้การเรียนรู้เชิงลึกเพื่อค้นหารูปแบบความสัมพันธ์ที่ซ่อนอยู่ในโครงข่ายการจราจรหรือที่เรียกว่าความเกี่ยวข้องเชิงพื้นที่ (nonlinear spatial-temporal effects) [2]

อย่างไรก็ตามการศึกษาวิจัยเรื่องการเรียนรู้เชิงลึกกับข้อมูลแผนที่เพื่อหาความสัมพันธ์เชิงพื้นที่เพื่อทำนาย

สภาพการจราจรล่วงหน้ายังมีผู้ศึกษาไม่มาก และมีความท้าทายอยู่ที่การจัดเตรียมข้อมูลให้มีแบบแผนที่พร้อมใช้สำหรับทำนายแบบทันกาล (Real time) ผู้วิจัยจึงมีแนวคิดที่จะพัฒนาหลักการสร้างวิธีการทำนายสภาพการจราจรจากภาพกูเกิลแมพ (Google Maps) ด้วยการเรียนรู้เชิงลึก เพื่อศึกษาพื้นที่ที่เกี่ยวข้องกับจุดที่ติดขัดและสามารถทำนายสภาพการจราจรได้ดีขึ้นกว่าการใช้อัลกอริทึมพื้นฐานทั่วไป ด้วยเหตุผลของการใช้ข้อมูลจากกูเกิลแมพ นั้นเป็นระบบติดตามปัญหาการจราจรที่มีค่าใช้จ่ายถูกกว่าวิธีอื่น เมื่อการวิจัยเสร็จสิ้นยังจะเป็นประโยชน์ต่อผู้สัญจรในเขตกรุงเทพฯ และผู้ที่ทำหน้าที่ศึกษาปัญหาการจราจรติดขัดที่จะรู้สาเหตุของการติดขัดเป็นแบบโครงข่าย (Spatial Temporal) ตามสภาพความเป็นจริงแบบทันกาลต่อเนื่องได้ตลอดเวลา

กูเกิล แมพ เป็นผู้ให้บริการด้านข้อมูลเชิงพื้นที่และสภาพการจราจร ที่มีเทคโนโลยีขั้นสูง มีผู้ใช้จำนวนมากที่สุดในโลก กูเกิล แมพ เริ่มเปิดใช้อย่างเป็นทางการเมื่อวันที่ 8 กุมภาพันธ์ ค.ศ. 2005 ให้บริการในพื้นที่สหรัฐอเมริกา, อังกฤษ และแคนาดา และขยายพื้นที่ให้บริการเป็นทั่วโลก ในวันที่ 20 มิถุนายน ณ ปีเดียวกัน [3] หลักการทำงานของกูเกิล แมพ จะใช้อัลกอริทึมในการแปลงสัญญาณ GPS จากผู้ใช้โทรศัพท์มือถือมาคำนวณเป็นสภาพการจราจรและประมาณการเดินทางได้อย่างแม่นยำ [4]

ในด้านความแม่นยำและความเชื่อถือได้ของข้อมูลจากกูเกิล แมพ มีงานวิจัยที่ยืนยันถึงความถูกต้องและความแม่นยำในเรื่องระยะทางและเวลาการเดินทางของกูเกิล แมพ โดยทำการทดสอบด้วยการใส่อุปกรณ์จีพีเอสให้คนเดินเทียบกับลู่วิ่งไฟฟ้ามาตรฐานพบว่ามีความสัมพันธ์สหสัมพันธ์สเปียร์แมน (Spearman rank correlation coefficients) สูงถึง 0.928 ซึ่งผู้วิจัยได้สรุปว่ากูเกิลแมพ นั้นเป็นเครื่องมือที่สามารถใช้วัดระยะทางหรือทำการทดลองเชิงพื้นที่ได้อย่างมีประสิทธิภาพ มีความแม่นยำ และประหยัดค่าใช้จ่าย [5]

จากปัญหาด้านการจราจรทางบกและที่มาดังกล่าวผู้วิจัยจึงได้ดำเนินการวิจัยเรื่องการทำนายสภาพจราจรในเขตกรุงเทพฯ ชั้นใน ด้วยการเรียนรู้เชิงลึก มีวัตถุประสงค์การวิจัยเพื่อเปรียบเทียบประสิทธิภาพการทำนายสภาพจราจรทางบก ระหว่างตัวแบบการเรียนรู้เชิงลึกที่ได้รับการปรับพารามิเตอร์ไว้อย่างเหมาะสมกับอัลกอริทึมมาตรฐานทั่วไป เช่น ต้นไม้ตัดสินใจ (Decision Tree), แรนดอมฟอเรสต์

(Random Forest), เทคนิคเพื่อนบ้านใกล้สุด (K Nearest Neighbors), นาอีฟเบย์ (Naïve Bayes), กราเดียนท์บูตติ้ง (Gradient Boosting) โดยทำการเปรียบเทียบประสิทธิภาพการทำนายจากเกณฑ์ค่าความถูกต้อง (Accuracy), ค่าความเที่ยง (Precision), ค่ารีคอล (Recall) และเอฟวันสกอร์ (f1-score) สำหรับข้อค้นพบที่เกิดขึ้นจากงานวิจัยนี้ จะเป็นประโยชน์ต่อผู้ที่ทำหน้าที่แก้ไขปัญหการจราจรทางบกในเขตกรุงเทพฯ ชั้นใน ได้ทราบถึงความเกี่ยวข้องของของจุดปัญหาการจราจรติดขัดจากข้อมูลจริง (Data Driven) เพื่อเป็นข้อมูลประกอบการวางแผนแก้ปัญหาได้อย่างมีประสิทธิภาพต่อไป

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 ทฤษฎีที่เกี่ยวข้อง

วิธีการทำนายสภาพการจราจรทางบกจากภาพภูเกิล แมพ ด้วยการเรียนรู้เชิงลึก ประกอบขึ้นจากสองฐานคิดสำคัญคือ เทคนิคในการจัดการเรียนรู้เชิงลึกสำหรับข้อมูลภาพคอนโวลูชัน (Convolutional Neural Network: CNN) [6] และเทคนิคในการสุ่มข้อมูลคลาสไม่สมดุล (Random Over-Sampling Examples: ROSE) [7] เพื่อทำให้คลาสมีความสมดุลโดยเทคนิคนี้จะทำให้การเรียนรู้ของตัวแบบการทำนายสภาพการจราจรทำได้แม่นยำขึ้น

1. เทคนิคในการจัดการเรียนรู้เชิงลึกสำหรับข้อมูลภาพคอนโวลูชัน (CNN) จะประกอบด้วย 5 ชั้นพื้นฐานของการประมวลผลด้วยภาพ

1.1) กำหนดชั้นของข้อมูลนำเข้า (Input Layer) เป็นทางเข้าของข้อมูล

1.2) ชั้นคอนโวลูชัน (Convolutional Layers) ประกอบไปด้วยชุดอนุกรมของตัวกรอง และการรับรู้ข้อมูลภาพ เช่น ขอบภาพ (image edges), รูปแบบปกติ (regular patterns) และการเปลี่ยนแปลงของสี (color changes) ยิ่งข้อมูลภาพมีความซับซ้อนจำนวนสตรึมของข้อมูลก็จะเพิ่มขึ้นไปตามกัน

1.3) หน่วยปรับนำหน้าข้อมูล (Rectified Linear Unit: ReLU) จัดการฟังก์ชันเกี่ยวกับนำหน้าของอัลกอริทึมตามที่ Krizhevsky ศึกษาเกี่ยวกับฟังก์ชันไม่อิ่มตัว (non-saturating activation function)

$$f(x) = \max(0, x)$$

การฝึกสอนที่ใช้ CNN จะช่วยแก้ปัญหาขอบเขตการเดียน และช่วยลดเวลาในการประมวลผลได้

1.4) ชั้นพูลลิง (Pooling Layers) ชั้นตอนนี้จะช่วยลดขนาดของพื้นที่และจำนวนพารามิเตอร์เพื่อให้การค้นหารูปแบบทำได้ง่ายและลดการเข้ากับตัวแบบมากเกินไป (over fitting) การพูลลิงจะทำหลายครั้งเพื่อให้ CNN มีความคงตัว

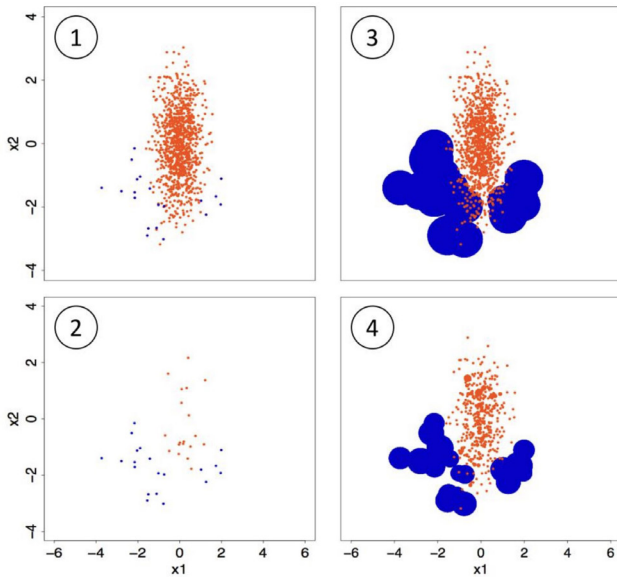
1.5) การรวมชั้นเข้ากันอย่างเต็มรูปแบบ (Fully connected Layers) เป็นขั้นตอนสุดท้ายในการรวมเน็ตเวิร์คและตัวแบบจัดจำแนกเข้าด้วยกันเพื่อสร้างตัวแบบใช้ทำนายเหตุการณ์ที่สนใจ

2. เทคนิคในการสุ่มข้อมูลคลาสไม่สมดุล

การสุ่มและสังเคราะห์ข้อมูลเพื่อให้จำนวนตัวแปรเป้าหมายมีความสมดุลและมีสารสนเทศเพียงพอก่อนเข้าขั้นตอนการฝึกสอนเพื่อสร้างตัวแบบนั้นมีอยู่หลายเทคนิค เช่น การสุ่มแบบสโมท (SMOTE: Synthetic Minority Over-sampling Technique) เป็นการสุ่มเลือกสมาชิกใกล้เคียง k ตัว ร่วมกับการใช้อัลกอริทึมที่ปรับสัดส่วนให้เหมาะสม (Adaptive Sampling) ซึ่งเป็นเทคนิคที่ได้ผลกับการแก้ปัญหาการจัดประเภทข้อมูลที่มีจำนวนคลาสไม่เท่ากันให้สามารถเพิ่มความถูกต้องในการทำนายได้ดีขึ้น [8] ในการวิจัยครั้งนี้ผู้วิจัยได้เลือกศึกษาเทคนิคการสุ่มตัวอย่างแบบโรส (ROSE: Random Over-Sampling Examples) ซึ่งเป็นไลบรารี (Library) ของโปรแกรม Python โดยการสุ่มแบบโรสจะทำการสังเคราะห์ข้อมูลขึ้นใหม่ตามขอบเขตความเป็นไปได้ของชุดข้อมูลพร้อมกับใช้อัลกอริทึมปรับสัดส่วนจำนวนคลาสจากคำนวณทวนซ้ำที่พิจารณาจากความถูกต้องจากการทำนายจนได้สัดส่วนที่เหมาะสมที่สุดวิธีการนี้จึงเหมาะสมและมีความมั่นใจได้กับการใช้ในทางปฏิบัติได้จริง เทคนิคการสุ่มแบบโรสถือเป็นอัลกอริทึมในการจัดเตรียมข้อมูลเพื่อเพิ่มความสามารถในการจำแนกประเภทให้มีความถูกต้องในคลาสส่วนน้อยให้เพิ่มมากยิ่งขึ้นโดยมีการทำงานตามภาพที่ 1 ดังนี้

จากภาพที่ 1 แสดงการสุ่มและสังเคราะห์ข้อมูลด้วยอัลกอริทึมโรส (ROSE) กล้องมุมบนซ้ายหมายเลขที่ 1) แสดงข้อมูลจริงที่มีความไม่สมดุลกันของจำนวนคลาสในกล้องมุมล่างซ้ายหมายเลขที่ 2) แสดงการสุ่มลดจำนวนให้เท่าคลาสที่น้อยกว่า ในกล้องมุมบนขวาหมายเลขที่ 3) แสดงการสังเคราะห์ข้อมูลเพิ่มจำนวนให้เท่าคลาสที่มากกว่า ด้วยการเรียนรู้จากข้อมูลที่อยู่ใกล้เคียง (Nearest Neighbors) และกล้องมุมล่างขวาหมายเลขที่ 4) เป็นการทำงานของอัลกอริทึมโรส (ROSE) ที่มีการปรับสัดส่วนของข้อมูลแต่ละคลาส

จากค่าการทำนายความถูกต้องและทำการปรับต่อไปจนได้จำนวนคลาสที่จะทำให้มีความถูกต้องในการทำนายสูงสุด



ภาพที่ 1 แสดงการสุ่มสังเคราะห์ของอัลกอริทึม Rose

2.2 งานวิจัยที่เกี่ยวข้อง

ตั้งแต่ทศวรรษที่ 20 (ค.ศ. 2000) เป็นต้นมา นักจัดการผังเมืองได้มีการพัฒนาวิธีการทั้งด้านทฤษฎีและวิธีการที่จะสร้างตัวแบบทำนายสภาพการจราจร โดยวิธีการในระยะแรกตั้งแต่ช่วงปี ค.ศ. 1950 เป็นต้นมาจะใช้วิธีการทางอนุกรมเวลาในระยะที่สองใช้อนุกรมเวลา ร่วมกับตัวกรองคอลแมน เพื่อให้ทำนายข้อมูลการจราจรที่อาจไม่เป็นระบบและมีรูปแบบที่ซับซ้อนรูปแบบที่แปลกแยก (Noise) ได้ดี และในยุคปัจจุบันนิยมใช้อัลกอริทึมสมัยใหม่หรือแมชชีนเลิร์นนิง สร้างตัวแบบทำนายสภาพการจราจรระยะสั้นหรือแบบทันกาล (Real time) ให้มีประสิทธิภาพเพิ่มขึ้น เพราะตัวแบบสมัยใหม่สามารถเรียนรู้รูปแบบข้อมูลได้หลากหลาย และไม่มีข้อจำกัดเฉพาะรูปแบบที่เป็นวัฏจักรซ้ำ ๆ (recurrent pattern) [9]

การใช้อัลกอริทึมแบบการเรียนรู้เชิงลึก (Deep neural networks: DNNs) ได้ถูกใช้กับการทำนายการเคลื่อนตัวและสภาพการจราจรจากฐานข้อมูลขนาดใหญ่ ซึ่งไม่เพียงแต่จะให้ประสิทธิภาพในการทำนายที่ดีกว่า แต่ยังเปิดเผยรูปแบบและลักษณะการเคลื่อนตัวที่เป็นเชิงพื้นที่ (spatial-temporal characteristics) ได้อีกด้วย ถือว่าเป็นวิธีการที่ทันสมัยและมีประสิทธิภาพสูงสุดในปัจจุบัน (state-of-the-art approaches) [10]

การศึกษาและทำนายจำนวนผู้เดินทางในเมืองหลวง (metro passenger flow) ด้วยการเรียนรู้เชิงลึก นั้นได้ขยายขอบเขตความรู้จากเดิมที่ใช้แค่ตัวแบบอนุกรมเวลาในปัจจุบันสามารถนำข้อเท็จจริงต่าง ๆ มาประกอบเข้าเป็นแบบที่มีข้อมูลที่หลากหลาย ทั้งเรื่องข้อมูลสภาพแวดล้อม ข้อมูลสภาพอากาศ ข้อมูลเกี่ยวกับรูปแบบการเคลื่อนตัวทางการจราจรและการปฏิบัติการของหน่วยงานต่าง ๆ ที่เกี่ยวข้อง มาผูกเป็นฐานข้อมูลขนาดใหญ่ เพื่อสร้างตัวแบบที่สะท้อนความเป็นจริง และฝึกกฎข้อจำกัดที่เคยมีอยู่ในฐานความรู้เรื่องระบบการขนส่ง [11]

การทำนายสภาพการจราจรถือเป็นส่วนสำคัญในระบบการจัดการจราจรอัจฉริยะ โดยได้มีการเริ่มใช้การเรียนรู้เชิงลึกเข้ามาใช้สร้างตัวแบบเพื่อเปิดเผยรูปแบบเชิงพื้นที่การจราจรในแต่ละช่วงเวลา (spatial-temporal traffic state) นำมาช่วยในการทำนายให้มีความแม่นยำขึ้น นอกจากนี้ยังสามารถขุดค้นเงื่อนไขของสภาพการจราจร (traffic flow mechanism) ซึ่งสามารถนำเอาเหตุการณ์และระบบในสถานการณ์ต่าง ๆ มารวมเข้าเป็นชุดข้อมูลขนาดใหญ่ใช้สร้างตัวแบบได้อีกด้วย [12]

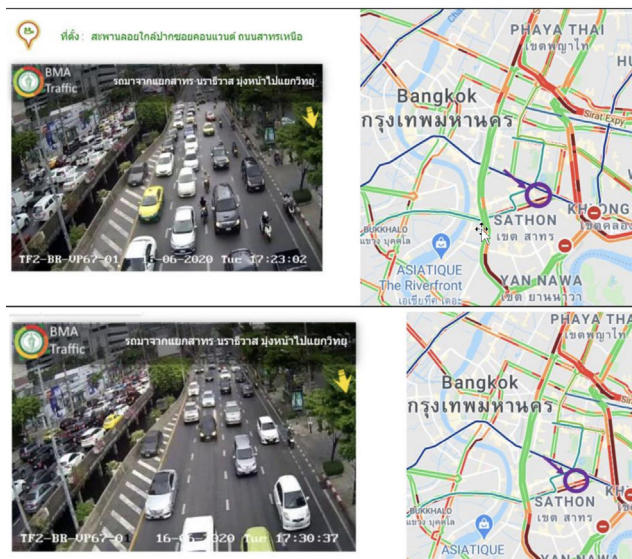
การเปลี่ยนแปลงของสภาพการจราจรอย่างรวดเร็ว เช่น การก่อสร้างหรือซ่อมปรับปรุงถนน การติดขัด การกระจายตัวของจำนวนรถยนต์ตามคาบเวลาเปลี่ยนแปลงไป สิ่งเหล่านี้ทำให้การทำนายสภาพการจราจรทำได้ยากเนื่องจากตัวแบบนั้นประกอบด้วยข้อมูลที่มีรูปแบบไม่คงตัว (non-stationary) ผู้วิจัยจึงเสนอวิธีการรวมตัวแบบ (ensemble) เพื่อแก้ปัญหาในเรื่องประสิทธิภาพการทำนาย โดยมี 3 ขั้นตอน คือ 1) แปลงข้อมูลที่เป็นตัวแบบปัญหาเชิงเส้นให้เป็นตัวแบบจัดจำแนก, 2) ใช้ต้นทุนสูญเสีย (loss functions) ร่วมกับการรวมตัวแบบ และ 3) ให้ตัวแบบเรียนรู้อีกครั้งโดยเลือกปรับน้ำหนักจากการทำนายผิดและผลโหวตของตัวแบบเพื่อให้ได้การตัดสินใจหลายระนาบ (decision hyperplane) ซึ่งการทดลองนี้พบว่าสามารถเพิ่มความถูกต้องในการทำนายได้ดีกว่าตัวแบบปริมาตร [13]

3. วิธีดำเนินการวิจัย

กลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้ เป็นภาพแสดงสภาพการจราจรที่ได้จากข้อมูลเว็บไซต์กูเกิลแมพ ในเขตกรุงเทพฯ จำนวน 1,844 ชุด ตามข้อแนะนำของ Figueroa et al. [14] ที่กำหนดจำนวนขั้นต่ำของกลุ่มตัวอย่าง 1,000 ตัวอย่าง

เพื่อให้ใช้ในการทำวิจัยด้วยการเรียนรู้ด้วยเครื่องได้ และผลการศึกษาของ Gaby Gorsky et al. [15] พบว่าจำนวนภาพ 1,000 ถึง 1,500 ภาพ เป็นจำนวนที่มากเพียงพอในการจำแนกคลาส ด้วยการเรียนรู้เชิงลึกได้อย่างแม่นยำ

ผู้วิจัยเก็บข้อมูลภาพจากกล้องวงจรปิด นำไปหาความสัมพันธ์ของระดับการติดขัดกับภาพภูเกิลแมพ ในพื้นที่เป้าหมาย 96 ภาพ (เขตกรุงเทพฯ ชั้นในย่านสาทร) เพื่อนำมาใช้ประเมินความถูกต้องของภาพภูเกิลแมพ โดยคำนวณตัวอย่างนี้ จากโปรแกรมสำเร็จรูป G*Power ซึ่งผู้วิจัยใช้ค่าขนาดอิทธิพล (Effect size) เป็นขนาดใหญ่มีค่าสัมประสิทธิ์สหสัมพันธ์ (Pearson's correlation coefficient) เกิน 0.5 ขึ้นไป เป็นเกณฑ์ในการวัดความสอดคล้องของเครื่องมือว่าเป็นไปในทิศทางเดียวกัน [16] โดยตัวอย่างข้อมูลภาพจากกล้องวงจรปิดและภาพจากภูเกิล แมพ ในช่วงเวลาเดียวกัน แสดงได้ตามภาพที่ 2



ภาพที่ 2 ตัวอย่างภาพจากกล้องวงจรปิดและภูเกิลแมพ

จากภาพที่ 2 ในภาพด้านบนเป็นตัวอย่างของสภาพการจราจรที่มีสีแดงเข้ม การจราจรติดขัดมาก ส่วนภาพด้านล่างเป็นตัวอย่างของสภาพการจราจรที่มีสีแดง การจราจรหนาแน่น ซึ่งผลที่ได้มีความสอดคล้องกัน โดยมีค่าสหสัมพันธ์เพียร์สันอยู่ที่ 0.96

เครื่องมือที่ใช้ในการวิจัยครั้งนี้คือ โปรแกรม Python เนื่องจากมีความพร้อมในการเก็บรวบรวมข้อมูล การสร้างตัวแบบการจัดจำแนกประเภท ตัวแบบการเรียนรู้เชิงลึก สามารถใช้ภาษา Python เป็นเครื่องมือหลักในการทำวิจัยได้อย่างสมบูรณ์

ส่วนเครื่องมือที่ใช้นำเสนอผลการทำนายและสามารถทำเป็นแอปพลิเคชันบนมือถือหรือเว็บไซต์ให้กลุ่มผู้ใช้งานจริง ผู้วิจัยได้ใช้โปรแกรม PowerBI ในการนำเสนอข้อมูล

การวิเคราะห์ข้อมูลด้วยโปรแกรมคอมพิวเตอร์สำเร็จรูป สถิติที่ใช้ ได้แก่ ร้อยละ ค่าเฉลี่ย ค่าเบี่ยงเบนมาตรฐาน การจัดอันดับ นำเสนอด้วยตาราง สำหรับตัวแบบการทำนายสภาพการจราจรด้วยการจำแนกประเภท เปรียบเทียบประสิทธิภาพการทำนายตามเกณฑ์ Accuracy, Precision, Recall และ f1-score และการอธิบายเพิ่มเติม

วิธีการดำเนินการวิจัย แบ่งออกเป็น 3 ระยะ ตามภาพที่ 3 และอธิบายขั้นตอนการดำเนินการวิจัยตามหัวข้อที่ 3.1 ถึง 3.3 ดังนี้

ระยะที่ 1: สร้างโปรแกรมเก็บข้อมูลภาพและกำหนดพารามิเตอร์ตัวแบบการเรียนรู้เชิงลึก

- 1.ศึกษาแนวคิดทฤษฎีที่เกี่ยวข้อง
- 2.พัฒนาโปรแกรมจัดเก็บข้อมูลภาพภูเกิลแมพ
- 3.ทดลองใช้กับพื้นที่เป้าหมาย /ปรับปรุงให้ใช้ได้จริง
- 4.จัดเตรียมข้อมูลสภาพการจราจรนำเข้าตัวแบบ
- 5.กำหนดพารามิเตอร์ตัวแบบการเรียนรู้เชิงลึก

ระยะที่ 2 : เปรียบเทียบประสิทธิภาพตัวแบบการเรียนรู้เชิงลึกกับการใช้อัลกอริทึมทั่วไป

วัดผลจากการทำนายสภาพการจราจรด้วยเกณฑ์มาตรฐานต่าง ๆ คือ Accuracy, True positive, False positive, Precision และ F-measure

ระยะที่ 3 : ศึกษาความถูกต้องในการทำนายสภาพการจราจรในเขตพื้นที่ กรุงเทพฯ ของตัวแบบการเรียนรู้เชิงลึก

ทำนายสภาพการจราจรล่วงหน้าด้วยตัวแบบการเรียนรู้เชิงลึก และรายงานความถูกต้องด้วยโปรแกรม Power BI

ภาพที่ 3 ขั้นตอนการวิจัยในภาพรวม

3.1 การสร้างโปรแกรมเก็บข้อมูลภาพและกำหนดพารามิเตอร์ตัวแบบการเรียนรู้เชิงลึก

เป็นกระบวนการศึกษาแนวคิดทฤษฎีที่เกี่ยวข้อง นำมาใช้พัฒนาโปรแกรมจัดเก็บข้อมูลสภาพจราจร ณ เขตกรุงเทพฯ ชั้นใน บริเวณถนนสาทรและบริเวณโดยรอบ จากเว็บไซต์กูเกิล แมพ จากนั้นทำการจัดเตรียมข้อมูลภาพให้เป็นข้อมูลแบบเมทริกซ์ และเก็บข้อมูลในรูปของตาราง โดยตัวแปรเป้าหมายระบุสถานะติดขัดให้กับชุดข้อมูลออกเป็น 4 ระดับตามสีของกูเกิล แมพ โดยสีเขียว, ส้ม, แดง, และแดงเข้ม และนำไปใช้การสร้างตัวแบบการเรียนรู้เชิงลึกทำนายสภาพจราจรทางบกล่วงหน้า ผลลัพธ์ในขั้นตอนสุดท้าย คือ ค่าพารามิเตอร์ของตัวแบบการเรียนรู้เชิงลึก ที่ผ่านการปรับจนเหมาะสมเพื่อนำไปใช้ทำนายสภาพการจราจรให้มีความถูกต้องมากที่สุด ในขั้นตอนต่อไป

3.2 การเปรียบเทียบประสิทธิภาพตัวแบบการเรียนรู้เชิงลึกกับการใช้อัลกอริทึมทั่วไป

ผู้วิจัยเปรียบเทียบประสิทธิภาพการทำนายของตัวแบบการเรียนรู้เชิงลึกที่ผ่านการปรับพารามิเตอร์จนเหมาะสมแล้วกับอัลกอริทึมทั่วไป โดยใช้พารามิเตอร์มาตรฐานในแต่ละอัลกอริทึม ตามตารางที่ 1 ดังนี้

ตารางที่ 1 ค่าพารามิเตอร์มาตรฐานของแต่ละอัลกอริทึม

อัลกอริทึม	ค่าพารามิเตอร์มาตรฐาน
Decision Tree	criterion='gini', min_samples_split=2, min_samples_leaf=1,max_features=No
Random Forest	n_estimators=100, criterion='gini', min_samples_split=2, min_samples_leaf=1, max_features='auto'
K Nearest Neighbors	n_neighbors=5, weights='uniform', leaf_size=30, metric='minkowski'
Naïve Bayes	alpha=1.0, fit_prior=True, class_prior=None
Gradient Boosting	loss='deviance', learning_rate=0.1, n_estimators=100, subsample=1.0, criterion='friedman_mse'

ทำการเปรียบเทียบประสิทธิภาพการทำนายสภาพการจราจรจากเกณฑ์วัดมาตรฐาน คือ Accuracy, Precision, Recall และ f1-score โดยคำนวณด้วยภาษา Python

3.3 การศึกษาข้อมูลเชิงพื้นที่ของสภาพการจราจรในเขตพื้นที่กรุงเทพฯ ชั้นใน

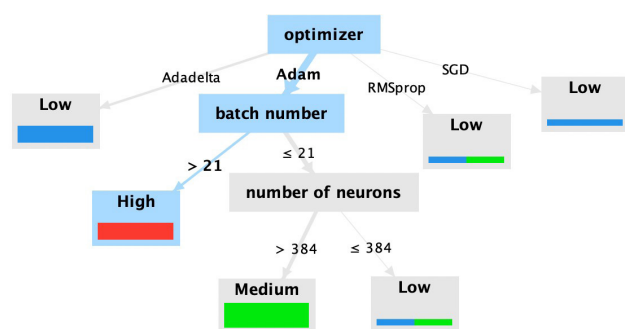
ผู้วิจัยได้ศึกษาข้อมูลเชิงพื้นที่ของสภาพการจราจรในเขตพื้นที่กรุงเทพฯ ด้วยหลักการที่พัฒนาขึ้น และทำนายสภาพการจราจรล่วงหน้าแบบทันเวลา (Real Time) ณ จุดที่ศึกษา คือ ย่านถนนพระราม 4 เพื่อทดสอบผลการทำนายในสถานการณ์จริงตามหลักการที่พัฒนาขึ้น

4. ผลการดำเนินงานวิจัย

ผลการวิเคราะห์ข้อมูล แบ่งเป็น 3 ส่วน คือ ส่วนที่ 1 ผลการจัดเตรียมข้อมูลภาพกูเกิล แมพ มาสร้างเป็นตัวแบบการเรียนรู้เชิงลึกทำนายสภาพการจราจรทางบก ส่วนที่ 2 ผลการเปรียบเทียบประสิทธิภาพตัวแบบทำนายที่พัฒนาขึ้นกับการใช้อัลกอริทึมมาตรฐาน และส่วนที่ 3 ผลการศึกษาความถูกต้องในการทำนายสภาพการจราจรในเขตพื้นที่กรุงเทพฯ ของตัวแบบการเรียนรู้เชิงลึก

4.1 ผลการพัฒนาตัวแบบการเรียนรู้เชิงลึกทำนายสภาพการจราจรทางบก

ผู้วิจัยได้กำหนดค่าพารามิเตอร์ของตัวแบบการเรียนรู้เชิงลึกที่เหมาะสมสำหรับข้อมูลสภาพการจราจรทางบก โดยเลือกจากผลการทดลองที่ดีที่สุด จากการทดลอง 27 ครั้ง และใช้โปรแกรม RapidMiner (เวอร์ชัน 9.1) ช่วยในการอธิบายการคัดเลือกพารามิเตอร์เป็นผังต้นไม้การตัดสินใจได้ตามภาพที่ 4



ภาพที่ 4 ผังการตัดสินใจการกำหนดค่าพารามิเตอร์

จากภาพที่ 4 ผู้วิจัยแบ่งคลาสระดับเป้าหมายความถูกต้องในการทำนายเป็น 3 ระดับ คือ ระดับต่ำ (Low) (แถบสีฟ้า) มีความถูกต้องต่ำกว่า 50%, ระดับกลาง (Medium) (แถบสีเขียว) มีความถูกต้องอยู่ระหว่าง 50% ถึง 80% และระดับสูง (High)

(แถบสีแดง) มีความถูกต้องมากกว่า 80% ผู้วิจัยเริ่มจากการทดลองที่ตัวปรับประสิทธิภาพ (Optimizer) พบว่าตัวปรับประสิทธิภาพอดัมส์ (Adam) ให้ผลที่ดีกว่าตัวปรับอื่นในขั้นตอนต่อมาจึงเลือกตัวปรับอดัมส์มาทำการปรับจำนวนชุดข้อมูลในแต่ละครั้งของการเรียนรู้ (Batch number) ที่ค่าต่าง ๆ คือ 4,5,6,12,18,24,30,36,42,48 พบว่าที่ค่าระหว่าง 30-48 ได้ผลความถูกต้องอยู่ในระดับสูง จากนั้นจึงทำการปรับจำนวนนิวรอน ที่ค่า 256,512,1024 พบว่า ค่า 1,024 ให้ผลดีที่สุด จากนั้นทำการปรับประเภทฟังก์ชันกระตุ้นแบบ linear, tanh และ relu พบว่า relu ให้ผลที่ดีที่สุด

ผู้วิจัยนำผลการทดลองที่ดีที่สุดมากำหนดค่าพารามิเตอร์ตัวแบบการเรียนรู้เชิงลึก สรุปได้ดังนี้ 1) จำนวนนิวรอนในเครือข่ายการเรียนรู้ (Hidden units) เท่ากับ 1024, 2) ประเภทของฟังก์ชันกระตุ้น (Activation function) เป็นแบบ relu, 3) ตัวปรับประสิทธิภาพ (Optimizer) เป็นแบบ Adam, 4) จำนวนรอบการเรียนรู้ (Training epochs) คือ 1,000 และ 5) จำนวนชุดข้อมูลในแต่ละครั้งของการเรียนรู้ (Batch number) คือ 30 ได้ผลลัพธ์ของความถูกต้องเฉลี่ย (Test Accuracy) เท่ากับ 82% และร้อยละส่วนเบี่ยงเบนมาตรฐาน (Accuracy standard deviations) เท่ากับ 2% ในการทำนายประเภทการจราจรด้วยวิธีการแบ่งข้อมูลแบบไขว้ 10 ชุด (10 folds evaluation) เพื่อใช้ข้อมูลเรียนรู้ให้ได้มากที่สุด

4.2 ผลการเปรียบเทียบประสิทธิภาพตัวแบบทำนายที่พัฒนาขึ้นกับการใช้อัลกอริทึมมาตรฐาน

ผลการวิจัยใช้การวิเคราะห์ ข้อมูลจากภาพสถานะทางการจราจรของเว็บไซต์กูเกิล ที่ได้มีการเก็บข้อมูลภาพแบบอัตโนมัติทุกครึ่งชั่วโมง ด้วยคำสั่งจากภาษา Python ที่ผู้วิจัยได้พัฒนาขึ้น ตั้งแต่วันที่ 11 กุมภาพันธ์ 2564 เวลา 16:00 น. ถึงวันที่ 25 มีนาคม 2564 เวลา 5:30 น. ณ บริเวณถนนสาทร จำนวน 1,844 ภาพ ภาพทั้งหมดถูกบันทึกเป็นไฟล์นามสกุลพีเอ็นจี (.png) แต่ละภาพมีขนาด 1,492 x 857 พิกเซล มีความละเอียด 72 พิกเซล/นิ้ว ผู้วิจัยได้เผยแพร่ไฟล์ภาพดังกล่าวทางเว็บไซต์ที่ shorturl.at/IqIRS (หรือติดต่อขอไฟล์ภาพจากผู้วิจัยได้ที่ amto1515@gmail.com)

เพื่อให้การทดสอบประสิทธิภาพการทำนายใกล้เคียงสภาพความเป็นจริง กล่าวคือ ตัวแบบจะต้องทำนายสภาพการจราจรที่ไม่เคยถูกฝึกสอนมาก่อนได้ การจัดเตรียมข้อมูล

จึงต้องแยกข้อมูลด้วยการสุ่มเป็น 2 กลุ่ม คือ ชุดข้อมูลฝึกสอน (Training Set) ใช้ในการสร้างตัวแบบ จำนวน 1,409 ตัวอย่าง คิดเป็นร้อยละ 80 กับ ชุดข้อมูลทดสอบ (Testing Set) ใช้ทดสอบผลการทำนายในชุดข้อมูลที่ไม่เคยพบมาก่อน จำนวน 353 ตัวอย่าง คิดเป็นร้อยละ 20 โดยมีจำนวนสภาพจราจร ตามข้อมูลในตารางที่ 2-3 ดังนี้

ตารางที่ 2 ข้อมูลภาพชุดข้อมูลฝึกสอน ถนนสาทรเหนือ

สภาพจราจร	จำนวนภาพ	ร้อยละ
แดงเข้ม: การจราจรติดขัดมาก	25	1.77
แดง: การจราจรหนาแน่น	81	5.75
ส้ม: การจราจรเคลื่อนตัวได้	243	17.25
เขียว: การจราจรคล่องตัว	1,060	75.23
รวม	1,409	100.00

ตารางที่ 3 ข้อมูลภาพชุดข้อมูลทดสอบ ถนนสาทรเหนือ

สภาพจราจร	จำนวนภาพ	ร้อยละ
แดงเข้ม: การจราจรติดขัดมาก	6	1.70
แดง: การจราจรหนาแน่น	19	5.38
ส้ม: การจราจรเคลื่อนตัวได้	66	18.70
เขียว: การจราจรคล่องตัว	262	74.22
รวม	353	100.00

ผลการเปรียบเทียบประสิทธิภาพตัวแบบการเรียนรู้เชิงลึกที่ได้ปรับพารามิเตอร์ไว้อย่างเหมาะสมกับการใช้อัลกอริทึมมาตรฐานต่าง ๆ ได้ผลตามตารางที่ 4

ตารางที่ 4 ค่าร้อยละเกณฑ์วัดประสิทธิภาพการทำนาย

อัลกอริทึม	Accuracy	Precision	Recall	f1-score
Deep Learning	75%	70%	71%	71%
Decision Tree	71%	71%	71%	71%
Gradient Boosting	72%	69%	72%	70%
Knn	73%	68%	73%	69%
Random Forest	71%	66%	71%	68%
Naïve Bayes	63%	64%	63%	63%

จากผลในตารางที่ 4 ตัวแบบที่มีผลความถูกต้องในการทำนาย (Accuracy) สูงสุด คือ ตัวแบบการเรียนรู้เชิงลึก มีค่าอยู่ที่ 75% สำหรับเกณฑ์ Precision ตัวแบบที่มีค่าสูงสุด คือ Decision Tree มีค่าอยู่ที่ 71% สำหรับเกณฑ์ Recall ตัวแบบที่มีค่าสูงสุดคือ Knn มีค่าอยู่ที่ 73% และเกณฑ์ f1-score ที่เป็นค่าเฉลี่ยถ่วงน้ำหนักระหว่างเกณฑ์ Precision และ Recall นั้นพบว่าตัวแบบที่ได้ค่าสูงสุด คือ ตัวแบบการเรียนรู้เชิงลึกและ Decision Tree ได้ค่าเท่ากันที่ 71% จากผลดังกล่าวสรุปได้ว่าตัวแบบการเรียนรู้เชิงลึกมีประสิทธิภาพการทำนายในเรื่องความถูกต้องและมีค่าเกณฑ์ f1-score สูงที่สุดเมื่อเทียบกับอัลกอริทึมมาตรฐานทั่วไป

4.3 ผลการศึกษาความถูกต้องในการทำนายสภาพการจราจรในเขตพื้นที่กรุงเทพฯ ของตัวแบบการเรียนรู้เชิงลึก

ในขั้นตอนนี้ได้ใช้ข้อมูลภาพชุดเดียวกับในหัวข้อ 4.2 แต่เปลี่ยนจุดเป้าหมายในการทำนายเป็น บริเวณถนนพระรามที่ 4 ใกล้เคียงจุดขึ้นลงทางพิเศษเฉลิมมหานคร เพื่อทดสอบความถูกต้องในการทำนายสภาพจราจรอีก 1 พื้นที่เนื่องด้วยข้อมูลภาพจากกูเกิล แมพ ที่เก็บรวบรวมอย่างเป็นระบบมีขอบเขตพื้นที่ครอบคลุมเขตกรุงเทพฯ ชั้นใน จึงมีความสะดวกในการดิงสารสนเทศของสภาพการจราจร ณ จุดเป้าหมายที่ต้องการศึกษาเพิ่มเติมได้สะดวก

จัดเตรียมข้อมูลด้วยการสุ่มเป็น 2 กลุ่ม คือ ชุดข้อมูลฝึกสอนใช้ในการสร้างตัวแบบจำนวน 1,409 ตัวอย่าง คิดเป็นร้อยละ 80 กับชุดข้อมูลทดสอบ ใช้ทดสอบผลการทำนายในชุดข้อมูลที่ไม่เคยพบมาก่อน จำนวน 353 ตัวอย่าง คิดเป็นร้อยละ 20 โดยมีจำนวนสภาพจราจร ตามข้อมูลในตารางที่ 5-6 ดังนี้

ตารางที่ 5 ข้อมูลภาพชุดข้อมูลฝึกสอน ถนนพระรามสี่

สภาพจราจร	จำนวนภาพ	ร้อยละ
แดงเข้ม: การจราจรติดขัดมาก	102	7.23
แดง: การจราจรหนาแน่น	86	6.10
ส้ม: การจราจรเคลื่อนตัวได้	123	8.73
เขียว: การจราจรคล่องตัว	1,098	77.92
รวม	1,409	100.00

ตารางที่ 6 ข้อมูลภาพชุดข้อมูลทดสอบ ถนนพระรามสี่

สภาพจราจร	จำนวนภาพ	ร้อยละ
แดงเข้ม: การจราจรติดขัดมาก	25	7.08
แดง: การจราจรหนาแน่น	28	7.09
ส้ม: การจราจรเคลื่อนตัวได้	18	5.09
เขียว: การจราจรคล่องตัว	282	79.87
รวม	353	100.00

ดำเนินการสร้างตัวแบบจากชุดข้อมูลฝึกสอนและนำมาทำนายผลกับทั้งชุดข้อมูลฝึกสอนเอง และชุดข้อมูลทดสอบเพื่อพิจารณาถึงความแตกต่างระหว่างความถูกต้องในการทำนายของชุดข้อมูลฝึกสอนกับข้อมูลทดสอบค่าความแตกต่างจากตารางที่ 7 คือผลต่างของความถูกต้องชุดฝึกสอนกับความถูกต้องชุดทดสอบ หากตัวแบบมีความถูกต้องในการทำนายกับข้อมูลฝึกสอนสูงมาก แต่เมื่อนำมาทำนายกับข้อมูลที่ไม่เคยพบมาก่อน คือ ข้อมูลทดสอบกลับได้ผลที่ไม่ดีเหตุการณ์ลักษณะนี้เรียกว่าโอเวอร์ฟิตติ้ง (Overfitting) ตัวแบบที่มีความต่างมากก็จะมีความเสี่ยงในการนำมาใช้จริง โดยผลการศึกษาสรุปได้ตามตารางที่ 7 ดังนี้

ตารางที่ 7 ค่าร้อยละความถูกต้องในการทำนาย

อัลกอริทึม	ความถูกต้องชุดฝึกสอน	ความถูกต้องชุดทดสอบ	ความแตกต่าง
Deep Learning	91%	82%	9%
Decision Tree	98%	79%	19%
Gradient Boosting	98%	82%	16%
Knn	92%	82%	10%
Random Forest	98%	82%	16%
Naïve Bayes	77%	78%	-1%

จากผลการศึกษาในตารางที่ 7 ตัวแบบการเรียนรู้เชิงลึกมีค่าโอเวอร์ฟิตต่ำสุดเป็นอันดับที่สองในบรรดาตัวแบบทั้งหมด โดยถึงแม้ว่า Naïve Bayes จะมีค่าความแตกต่างของความถูกต้องต่ำสุด คือ -1% แต่กลับมีความถูกต้องในการทำนายข้อมูลทดสอบที่ต่ำสุด คือ 78% หากตัด Naïve Bayes ออกตัวแบบการเรียนรู้เชิงลึกจะเป็นตัวแบบที่มีผลการทำนายถูกต้องสูงที่สุดที่ 82% และมีค่าความแตกต่างของความถูกต้องต่ำที่สุดที่ 9% แสดงถึงคุณสมบัติในการใช้งานจริงได้ดี

นอกจากนี้ผู้วิจัยได้นำผลการทำนายสภาพการจราจรทำเป็นรายงานอย่างง่ายด้วยโปรแกรม PowerBI แสดงตัวอย่างหน้าแอปพลิเคชันบนมือถือ ตามภาพที่ 5 ดังนี้



วันเวลา ณ ปัจจุบัน

ภาพที่ 5 แอปพลิเคชันทำนายสภาพการจราจรล่วงหน้า

5. อภิปรายผลการวิจัยและข้อเสนอแนะ

5.1 อภิปรายผลการวิจัย

5.1.1 การสร้างตัวแบบการเรียนรู้เชิงลึกทำนายสภาพการจราจรทางบก จากข้อมูลภาพ ภูเกิล แมพ ผลการปรับค่าพารามิเตอร์ที่ดีที่สุด จากการทดลองพารามิเตอร์หลัก เช่น จำนวนนิวรอนในเครือข่ายการเรียนรู้ เท่ากับ 1024 และจำนวนรอบการเรียนรู้ (Training epochs) คือ 1,000 ทั้งสองพารามิเตอร์นี้เป็นค่าที่มากที่สุด ดังนั้นจึงกล่าวได้ว่า หากใช้พารามิเตอร์ที่มีความซับซ้อนขึ้น

จะทำให้ผลการทำนายมีความแม่นยำได้ดีขึ้น และจากการทดลองในเรื่องประเภทของฟังก์ชันกระตุ้น ระหว่าง แบบ linear, tanh และ relu ผลการทดลองปรากฏว่าแบบ relu ให้ค่าการทำนายสูงที่สุด [6] สอดคล้องกับแนวคิดการใช้ข้อมูลภาพมาสร้างเป็นตัวแบบการเรียนรู้เชิงลึกตามที่ได้ผู้วิจัยได้ทบทวนวรรณกรรมมาที่ระบุถึงการใช้หน่วยปรับนำหนักข้อมูล (Rectified Linear Unit: ReLU) ที่มีความเหมาะสมกับลักษณะข้อมูลภาพ

5.1.2 ผลการเปรียบเทียบประสิทธิภาพการทำนายพบว่าการเรียนรู้เชิงลึกนั้นมีประสิทธิภาพสูงกว่าอัลกอริทึมมาตรฐานทั่วไป ใน 2 ด้านคือ ด้านที่ 1) จากผลการศึกษาในหัวข้อ 4.2 ผลประเมินประสิทธิภาพค่าความถูกต้องของการทำนายและค่า f1-score อยู่ในระดับสูงที่สุดเมื่อเทียบกับอัลกอริทึมทั่วไป และด้านที่ 2) จากผลการศึกษาในหัวข้อ 4.3 พบว่าตัวแบบการเรียนรู้เชิงลึกมีค่าโอเวอร์ฟิตต่ำ แสดงถึงตัวแบบที่มีความเสี่ยงต่ำในการทำนายผิดพลาด หากนำไปใช้ในสถานการณ์จริง

5.2 ข้อเสนอแนะ

จากการที่ผู้วิจัยใช้ข้อมูลจากเว็บไซต์กูเกิล แมพ เป็นหลัก จึงมีความเสี่ยงในเรื่องการพึ่งพิงเทคโนโลยีจากต่างประเทศเพียงอย่างเดียว แต่ด้วยหลักการที่พัฒนาขึ้นสามารถประยุกต์เข้าได้กับระบบเซ็นเซอร์ในการวัดปริมาณการจราจร หรือสามารถใช้ได้กับข้อมูลภาพแผนที่แสดงปริมาณการจราจรอื่น ๆ ที่ไม่ใช่เว็บไซต์กูเกิล แมพ ได้ด้วย ซึ่งผู้ที่สนใจสามารถนำหลักการดังกล่าวไปใช้กับข้อมูลแผนที่จราจรอื่น ๆ หรือใช้กับพื้นที่อื่น ๆ นอกเหนือจากเขตกรุงเทพฯ ชั้นใน ได้อีกด้วย ดังนั้นผู้วิจัยจึงเสนอแนะให้มีการพัฒนาระบบการสร้างและรายงานสถานการณ์จราจรขึ้นเป็นของประเทศไทย โดยเริ่มจากการขอรับทุนวิจัยจากรัฐบาลหรือสถาบันการศึกษาก็ตาม เพื่อให้ นักวิชาการของไทยได้พัฒนาระบบสารสนเทศทางการจราจรตั้งแต่ต้นน้ำถึงปลายน้ำได้เอง

6. เอกสารอ้างอิง

- [1] Kasikorn Research Center, *Impacts of Traffic Congestion on Bangkok's Economy and Life*. Available Online at <https://www.kasikornresearch.com/en/analysis/k-econ/economy/Pages/35760>, accessed on 20 November 2020.

- [2] N. G. Polson, and V. O. Sokolov. "Deep learning for short-term traffic flow prediction." *Transportation Research Part C: Emerging Technologies*, Vol. 79, pp. 1-17, June, 2017.
- [3] S.-Y. Yang, and C.-L. Hsu. "A location-based services and Google maps-based information master system for tour guiding." *Computers & Electrical Engineering*, Vol. 54, pp. 87-105, August, 2016.
- [4] M. Aghaabbasi, M. Moeinaddini, M. Z. Shah, and Z. Asadi-Shekari. "Addressing issues in the use of Google tools for assessing pedestrian built environments." *Journal of Transport Geography*, Vol. 73, pp. 185-198, December, 2018.
- [5] H. Khambati, K. Boles, and P. Jetty. "Google Maps offers a new way to evaluate claustrophobia." *Journal of Vascular Surgery*, Vol. 65, No. 5, pp. 1467-1472, May, 2017.
- [6] M.-J. Tsai, Y.-H. Tao, and I. Yuadi. "Deep learning for printed document source identification." *Signal Processing: Image Communication*, Vol. 70, pp. 184-198, February, 2019.
- [7] N. Lunardon, G. Menardi, and N. Torelli. "ROSE: A Package for Binary Imbalanced Learning." *R Journal*, Vol. 6, No. 1, June, 2014.
- [8] V. López, A. Fernández, S. García, V. Palade, and F. Herrera, "An insight into classification with imbalanced data: Empirical results and current trends on using data intrinsic characteristics." *Information Sciences*, Vol. 250, pp. 113-141, November, 2013.
- [9] Y. Zhao, H. Zhang, L. An, and Q. Liu. "Improving the approaches of traffic demand forecasting in the big data era." *Cities*, Vol. 82, pp. 19-26, December, 2018.
- [10] Y. Wu, H. Tan, L. Qin, B. Ran, and Z. Jiang, "A hybrid deep learning based traffic flow prediction method and its understanding." *Transportation Research Part C: Emerging Technologies*, Vol. 90, pp. 166-180, May, 2018.
- [11] Y. Liu, Z. Liu, and R. Jia. "DeepPF: A deep learning based architecture for metro passenger flow prediction." *Transportation Research Part C: Emerging Technologies*, Vol. 101, pp. 18-34, April, 2019.
- [12] J. Wang, R. Chen, and Z. He. "Traffic speed prediction for urban transportation network: A path based deep learning approach." *Transportation Research Part C: Emerging Technologies*, Vol. 100, pp. 372-385, March, 2019.
- [13] J. Xiao, Z. Xiao, D. Wang, J. Bai, V. Havaryimana, and F. Zeng, "Short-term traffic volume prediction by ensemble learning in concept drifting environments." *Knowledge-Based Systems*, Vol. 164, pp. 213-225, January, 2019.
- [14] R.L. Figueroa, Q. Zeng-Treitler, S. Kandula et al. "Predicting sample size required for classification performance." *BMC Med Inform Decis Mak*. Vol. 12, No. 8, February, 2012.
- [15] G. Gorsky, M. D. Ohman, M. Picheral, S. Gasparini, L. Stemmann, J.B. Romagnan, A. Cawood, Stéphane Pesant, C.G. Comas, and F. Prejger. "Digital zooplankton image analysis using the ZooScan integrated system," *Journal of Plankton Research*, Vol. 32, No. 3, pp. 285-303, March, 2010.
- [16] F. Faul, E. Erdfelder, A. Buchner, and A.-G. Lang. "Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses." *Behavior Research Methods*, Vol. 41, pp. 1149-1160, November, 2009.

