



วารสารวิชาการ วิศวกรรมศาสตร์ ม.อบ. UBU Engineering Journal

บทความวิจัย

การประมาณความยาวของแถวคอยแบบแปรผันตามเวลาบริเวณทางแยกสัญญาณไฟจราจรโดยใช้ข้อมูลแถบสีจาก Google Maps

Estimating time-varying queue length at signalized intersection by using color-coded traffic on Google Maps

พงศ์พัฒน์ จอดนอก* รัฐพล ภูบุบผาพันธ์

สาขาวิชาวิศวกรรมขนส่ง สำนักวิชาวิศวกรรมศาสตร์ มหาวิทยาลัยเทคโนโลยีสุรนารี อำเภอเมือง จังหวัดนครราชสีมา 30000

Pongpat Jodnok* Rattaphol Pueboobpaphan

School of Transportation Engineering, Institute of Engineering, Suranaree University of Technology Nakhon Ratchasima 30000

* Corresponding author.

E-mail: m6040321@sg.sut.ac.th; Telephone: 063 8816102

วันที่รับบทความ 19 มิถุนายน 2563; วันที่แก้ไขบทความ ครั้งที่ 1 5 สิงหาคม 2563; วันที่แก้ไขบทความ ครั้งที่ 2 9 กันยายน 2563

วันที่ตอบรับบทความ 24 พฤศจิกายน 2563

บทคัดย่อ

ปัญหาการจราจรติดขัดบริเวณทางแยกไฟสัญญาณจราจรเป็นปัญหาที่พบเจอได้บ่อยในเขตเมือง บ่อยครั้งพบว่ารถฉุกเฉินเช่นรถพยาบาลต้องติดอยู่ในแถวคอยบริเวณทางแยกสัญญาณไฟ ส่งผลให้การเดินทางรับส่งผู้ป่วยเป็นไปอย่างล่าช้าและอาจเป็นสาเหตุของการสูญเสียตามมา งานวิจัยฉบับนี้มีเป้าหมายในการพัฒนาระบบไฟสัญญาณจราจรอัจฉริยะเพื่อให้สิทธิพิเศษในการผ่านทางแก่รถฉุกเฉิน โดยมุ่งเน้นพัฒนาวิธีการในการตรวจวัดความยาวของแถวคอยแบบแปรผันตามเวลาบริเวณทางแยกที่มีไฟสัญญาณจราจรซึ่งเป็นข้อมูลสำคัญส่วนหนึ่งในการดำเนินการของระบบ อย่างไรก็ตาม งานวิจัยในอดีตส่วนใหญ่ใช้วิธีการวิเคราะห์ความยาวแถวคอยด้วยอุปกรณ์ตรวจวัดจราจร แต่ทางแยกส่วนใหญ่ในประเทศไทยยังไม่มีการติดตั้งและใช้งานอุปกรณ์ดังกล่าว ดังนั้นงานวิจัยนี้จึงนำเสนอแนวทางใหม่เพื่อใช้ในกรณีที่บริเวณทางแยกไม่มีอุปกรณ์ตรวจวัดจราจร โดยอาศัยข้อมูลแถบสีจาก Google Maps ในการประมาณค่าระยะความยาวแถวคอยผ่านวิธีการที่แตกต่างกัน 3 วิธีได้แก่วิธีการถดถอยเชิงเส้นแบบพหุคูณ, อัลกอริทึมการพยากรณ์แบบการสุ่มป่าไม้ (Random Forest) และวิธีการใช้ค่าเฉลี่ยความยาวแถวคอย จากการศึกษาพบว่าระยะความยาวของแถบสีที่แสดงบน Google Maps สามารถประมาณค่าแนวโน้มความยาวแถวคอยที่เกิดขึ้นจริงได้ดีในระดับหนึ่งและมีความสัมพันธ์กันมากขึ้นเมื่ออยู่ในช่วงเวลาเร่งด่วนที่มีการจราจรหนาแน่น โดยพบว่าวิธี Random Forest สามารถพยากรณ์ระยะความยาวแถวคอยได้ดีที่สุดเมื่อเปรียบเทียบกับวิธีอื่น

คำสำคัญ

ระยะความยาวแถวคอย แผนที่ของกูเกิล อัลกอริทึมการพยากรณ์แบบการสุ่มป่าไม้ ทางแยกไฟสัญญาณจราจร

Abstract

Traffic jam, at a signalized intersection, is a common problem in urban areas that emergency vehicles such as ambulances are usually stuck in a queue. Thus, emergency medical transportation may be impeded leading to unnecessary damage or death. This research aims to develop an intelligent traffic signal system that prioritizes emergency vehicles. The main focus is developing a method for measuring the time-varying queue length at a signalized intersection, which is an integral part of the system's operation. In the past, most of research analyzed the queue lengths by using traffic detectors, but they have rarely been installed and used in Thailand. In this paper, we present a new approach to

estimate the queue length that does not need a traffic detector; instead, it utilized color-coded traffic data from Google Maps to predict the queue length based on three different methods: which are Multiple Linear Regression, Random Forest, and The Mean method. The study found that the length of the colored bars displayed on Google Maps can, to some extent, provide a good approximation of the actual queue length trend, and are more correlated with two methods especially in rush hours with heavy traffic. The Random Forest, among three methods, yields the best performance in predicting the queue length when compared to the other two methods.

Keywords

queue length; google maps; random forest; signalized intersection

1. คำนำ

สภาพการจราจรติดขัดบนท้องถนนในปัจจุบันอาจเป็นปัญหาที่หลายคนเคยพบเจอ เช่นเดียวกับรถฉุกเฉิน (Emergency Vehicle) ที่ต้องเดินทางเพื่อรับส่งผู้ป่วยที่ต้องการการรักษาแบบเร่งด่วนแต่ต้องติดอยู่บริเวณทางแยกไฟสัญญาณจราจรเป็นเวลานาน แผนหลักการแพทย์ฉุกเฉินแห่งชาติ ฉบับ 3.1 พ.ศ. 2562 – 2564 มีการคาดการณ์ประมาณจำนวนผู้ป่วยฉุกเฉินเสียชีวิตนอกโรงพยาบาล ปีละประมาณ 60,000 คน หากประเทศไทยมีระบบการแพทย์ฉุกเฉินที่มีประสิทธิภาพจะสามารถช่วยรักษาชีวิตผู้เจ็บป่วยฉุกเฉินดังกล่าวได้ถึงประมาณร้อยละ 15–20 หรือประมาณปีละ 9,000–12,000 คน [1] ทำให้ทราบถึงความสำคัญว่าความล่าช้าในการให้ความช่วยเหลืออาจส่งผลกระทบต่อชีวิตของผู้ป่วยได้ ดังนั้นงานวิจัยฉบับนี้จึงมีความสนใจในการศึกษาหาแนวทางของการลดความล่าช้าบริเวณทางแยกไฟสัญญาณจราจรของรถฉุกเฉิน ซึ่งในอนาคตจำเป็นต้องมีเทคโนโลยีที่เข้ามาช่วยเพิ่มประสิทธิภาพการทำงานของตู้ควบคุมไฟสัญญาณจราจร เมื่อทราบว่ารถฉุกเฉินกำลังเดินทางเข้ามาใกล้ทางแยกแล้วให้สามารถปรับเปลี่ยนไฟสัญญาณจราจรโดยอัตโนมัติเพื่อเปิดสัญญาณไฟเขียวปล่อยให้รถที่หยุดรออยู่ก่อนหน้าจนเกิดเป็นแถวคอยบริเวณทางแยกสามารถเดินทางผ่านไปเพื่อรองรับการเดินทางของรถฉุกเฉินที่กำลังจะมาถึงให้สามารถเดินทางผ่านไปได้อย่างไม่ต้องหยุดรถ ในการดำเนินการดังกล่าวจำเป็นต้องทราบถึงระยะความยาวแถวคอยบริเวณทางแยกในขณะที่ยังทราบถึงรถฉุกเฉินกำลังจะมาถึง จึงเป็นที่มาการศึกษาของงานวิจัยฉบับนี้ที่ต้องการจะประมาณระยะความยาวแถวคอยบริเวณทางแยกไฟสัญญาณจราจร และจากการศึกษางานวิจัยก่อนหน้านั้นมีวิธีการประมาณระยะความยาวแถวคอยที่เป็นที่นิยมคือ การประมาณระยะแถวคอยโดยใช้ข้อมูลจาก

อุปกรณ์ตรวจวัดจราจร (Detector) [2-5] แต่อย่างไรก็ตามพื้นที่ส่วนใหญ่ในประเทศไทยยังไม่มีติดตั้ง Detector ที่บริเวณทางแยก

ดังนั้นจึงต้องหาแนวทางที่สามารถใช้ในการประมาณระยะความยาวแถวคอยได้อย่างมีประสิทธิภาพและใช้งานเข้าได้ง่าย โดยงานวิจัยนี้มีสมมติฐานว่าระยะความยาวและลักษณะสีของแถบสีที่แสดงบน Google Maps มีความสัมพันธ์และสามารถนำมาใช้ในการประมาณระยะความยาวแถวคอยผ่านการสร้างแบบจำลองได้ และเพื่อให้แบบจำลองมีความแม่นยำในการพยากรณ์มากขึ้นจึงต้องมีการศึกษาในหลายช่วงเวลาและมีความหลากหลายของสภาพการจราจรที่แตกต่างกันด้วย

2. ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

2.1 วิธีการประมาณระยะความยาวแถวคอยในอดีต

ระยะความยาวแถวคอยมีความสำคัญต่อการวัดประสิทธิภาพของไฟสัญญาณจราจรทั้งในเรื่องของความล่าช้าและการหยุดคอยไฟสัญญาณจราจรของยานพาหนะ [6-8] ซึ่งการประมาณระยะความยาวแถวคอยสามารถทำได้หลายวิธีขึ้นอยู่กับความพร้อมและความเหมาะสมของพื้นที่นั้นๆ ในงานศึกษาของ [9] ได้เสนอการประมาณระยะความยาวแถวคอยแบบ Real Time บริเวณทางแยกไฟสัญญาณจราจรโดยใช้เวลาในการเดินทางจาก Mobile Sensors ร่วมกับ LWR Shockwave Theory โดยทฤษฎีนี้ถูกพัฒนาโดย [2,3] และการวิเคราะห์ความยาวแถวคอยโดยวิธีการนี้ จะต้องมีการพิจารณาการเปลี่ยนแปลงรูปแบบของการเดินทางและความล่าช้าของยานพาหนะบริเวณทางแยก เช่น ความไม่ต่อเนื่อง (Discontinuities) และ ความไม่ราบรื่น (Non-Smoothness) และงานของ [10] ได้เสนอการประมาณระยะความยาวแถวคอยโดยการใช้ข้อมูลจากเครื่องตรวจจับ RFID โดยใช้

ประโยชน์จากความล่าช้าของการรอกิวของยานพาหนะแต่ละคัน จากงานวิจัยที่กล่าวมาข้างต้นส่วนใหญ่จะใช้ Detector หรือ Sensor ในการศึกษาทำให้พบปัญหาที่ระยะแถวคอยยาวเกินกว่าจุดติดตั้ง Detector จึงไม่สามารถประมาณระยะความยาวแถวคอยได้อย่างแม่นยำ ต่อมาพบว่ามีการศึกษาเพื่อยืนยันว่าการประยุกต์ใช้ Shockwave Theory ร่วมกับ Loop Detector สามารถประมาณความยาวแถวคอยบริเวณทางแยกได้แม่นยำมากขึ้นโดย [5] วิธีการเหล่านี้จำเป็นต้องมีการติดตั้ง Sensor หรือ Detector ซึ่งในประเทศไทยยังไม่มีติดตั้งอย่างแพร่หลายทำให้เป็นอุปสรรคในการใช้วิธีการข้างต้น

2.2 ข้อมูลจาก Google Maps

การแสดงผลจราจรแบบเรียลไทม์ของ Google Maps ในประเทศไทยมีความแม่นยำมากขึ้นเนื่องจากทาง Google นำข้อมูลจากผู้ใช้งาน Android และ iOS ที่เปิดใช้งาน Google Maps และยินยอมแชร์ตำแหน่งขณะกำลังใช้งาน ทำให้สมาร์ตโฟนจะส่งข้อมูลแบบไม่ระบุตัวตน (Anonymous Data) กลับไปให้ทาง Google ว่าโทรศัพท์กำลังเคลื่อนที่ด้วยความเร็วเท่าใด แล้วนำไปเปรียบเทียบกับโทรศัพท์เครื่องอื่น ๆ ที่อยู่บนถนนเดียวกับผู้ใช้งานและบริเวณรอบๆ เพื่อตรวจสอบความถูกต้อง แล้วจะประมวลผลข้อมูลเหล่านี้จนเห็นภาพรวมของการจราจรทั้งหมด แล้วแสดงผลในรูปแบบเส้นแถบสีดังแสดงในรูปที่ 1 [11]



รูปที่ 1 แสดงแถบสีบน Google Maps

2.3 โปรแกรมและแบบจำลองการพยากรณ์

โปรแกรมที่ใช้ในงานวิจัยฉบับนี้ได้แก่โปรแกรม R ซึ่งเป็น Open Source Software ที่นิยมใช้คำนวณทางสถิติ เนื่องจากโปรแกรม R ค่อนข้างมีความยืดหยุ่นในการใช้งานจึงสามารถใช้เพื่อสร้างแบบจำลองในการพยากรณ์ค่าต่าง ๆ ได้ เช่น การสร้างแบบจำลองการถดถอยเชิงเส้น (Linear Regression

Model) และ การสร้างแบบจำลอง Random Forest Model เป็นต้น ซึ่งทั้งสองวิธีการที่กล่าวมาข้างต้นใช้ในการพยากรณ์ค่าระยะความยาวแถวคอยของการศึกษาในครั้งนี้

2.3.1 Linear Regression Model

การวิเคราะห์การถดถอยเป็นวิธีที่ใช้เทคนิคทางสถิติในการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรอิสระ (Independent Variable) กับตัวแปรตาม (Dependent Variable) ซึ่งเป็นตัวแปรที่เปลี่ยนแปลงไปตามตัวแปรอิสระ ในกรณีที่ความสัมพันธ์ของสัมประสิทธิ์ในสมการการถดถอยเป็นแบบเชิงเส้นจะเรียกวิธีนี้ว่าการวิเคราะห์การถดถอยเชิงเส้น โดยการวิเคราะห์การถดถอยเชิงเส้นพหุคูณ (Multiple Linear Regression Analysis) เป็นการวิเคราะห์ความสัมพันธ์ที่มีตัวแปรตาม 1 ตัว และตัวแปรอิสระตั้งแต่ 2 ตัวขึ้นไป ซึ่งสมการการถดถอยเชิงเส้นพหุคูณสามารถแสดงได้ดังสมการที่ (1)

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \varepsilon \quad (1)$$

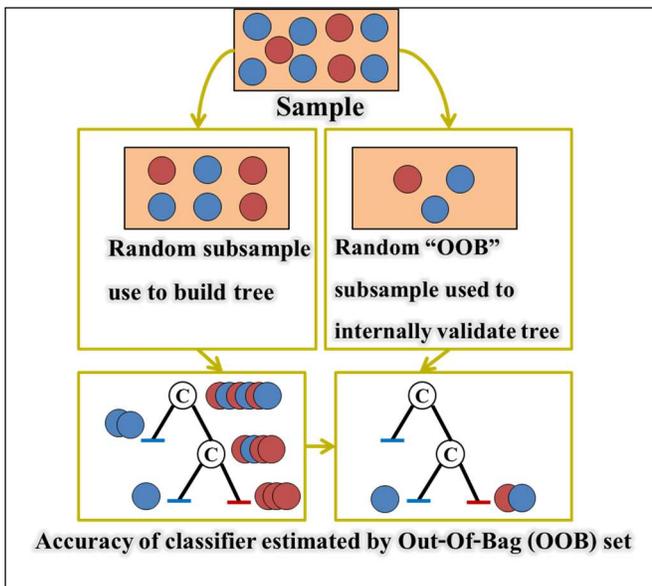
โดยที่ x_1, x_2, \dots, x_n คือ ตัวแปรอิสระ; $\beta_0, \beta_1, \dots, \beta_n$ คือ สัมประสิทธิ์ของสมการการถดถอย การประมาณค่าของสัมประสิทธิ์ในสมการถดถอยเชิงเส้นพหุคูณนั้นนิยมใช้วิธีการกำลังสองน้อยที่สุด การวิเคราะห์การถดถอยมีสมมติฐานพื้นฐาน 4 ข้อ คือ

- 1) ประชากรของค่าเฉลี่ยความผิดพลาดมีค่าเฉลี่ยคือศูนย์
- 2) ประชากรของค่าความผิดพลาดมีความแปรปรวนคงที่
- 3) ประชากรของค่าความผิดพลาดมีการแจกแจงแบบปกติ
- 4) ค่าความผิดพลาดมีความเป็นอิสระต่อกัน

แต่การวิเคราะห์การถดถอยเชิงเส้นก็มีข้อจำกัดหลายประการ เช่น ไม่เหมาะกับการพยากรณ์ในกรณีที่ความสัมพันธ์ไม่เป็นเชิงเส้น (Non-linear Relationships) และใช้สำหรับการพยากรณ์ที่มีผลลัพธ์ที่เป็นตัวเลขเท่านั้น นอกจากนี้ผู้ที่พยากรณ์โดยใช้วิธีนี้มักละเลยและมองข้ามในการตรวจสอบว่าข้อมูลที่มีอยู่นั้นเป็นไปตามข้อสมมติฐานพื้นฐาน 4 ข้อที่กล่าวถึงข้างต้นของการวิเคราะห์การถดถอยหรือไม่ ซึ่งอาจส่งผลให้ผลการพยากรณ์ที่ได้ผิดพลาดไปมากและมีความน่าเชื่อถือลดลง [12]

2.3.2 Random Forest Model

เทคนิค Random Forest [13] เป็นเทคนิคการสุ่มเลือกข้อมูลและคุณลักษณะ Decision Tree ซึ่งถูกสร้างจากการนำข้อมูลไปสุ่มเลือกตัวอย่างแบบเลือกแล้วใส่กลับ แล้วนำมาสร้างเป็น Tree ซึ่งจะมีตัวอย่างส่วนหนึ่งที่จะไม่ถูกเลือก ข้อมูลส่วนนี้เรียกว่า Out-of-Bag (OOB) จะถูกนำมาใช้ในการทดสอบ Decision Tree ซึ่งเรียกวิธีดังกล่าวว่า Bagging ผลลัพธ์ที่ได้จะอิสระจาก Decision Tree ในแต่ละต้นถูกนำมาคิดเป็นผลการโหวต ผลโหวตที่มากที่สุดจะใช้ระบุสถานะของคลาสดังรูปที่ 2 เทคนิค Random Forest ไม่จำเป็นต้องมีข้อมูลทดสอบเพื่อประมาณความผิดพลาดเพราะข้อมูล OOB นั้นถูกนำมาใช้ทดสอบ Decision Tree แล้วนั่นเอง โดย Decision Tree แต่ละแบบจำลองใน Random Forest ถือว่าเป็น Weak Learner คือ Model ที่ยังไม่มี ความแม่นยำมาก แต่เมื่อนำเอาแต่ละ Decision Tree มาทำการพยากรณ์ร่วมกันก็จะได้แบบจำลองรวมที่มีประสิทธิภาพ และแม่นยำมากกว่า Decision Tree ที่ทำการพยากรณ์แบบเดี่ยว ๆ โดย Random Forest ใช้ได้ทั้งกับปัญหา Classification และ Regression



รูปที่ 2 การทำงานของหนึ่ง Decision Tree ใน Random Forest [14]

3. การสำรวจและเก็บข้อมูล

งานวิจัยนี้ต้องการหาความสัมพันธ์ระหว่างความยาวและสีของแถบสีบน Google Maps กับระยะเวลาความยาวแฉกคยจริง

ดังนั้นจึงต้องมีการสำรวจและเก็บข้อมูลทั้งลักษณะการเปลี่ยนแปลงของระยะความยาวและสีของแถบสีบน Google Maps พร้อมกับบันทึกระยะเวลาความยาวแฉกคยจริงในพื้นที่ศึกษาไปพร้อมกันด้วย ซึ่งพื้นที่ที่ศึกษาคือบริเวณทางแยกไฟสัญญาณจราจรบนถนนเส้นหลักกลางเมืองนครราชสีมาใกล้กับห้างสรรพสินค้าเทสโก้โลตัส ซึ่งเป็นบริเวณที่มีการจราจรหนาแน่นและเกิดปัญหาจราจรติดขัดอยู่เสมอ โดยทิศทางที่ศึกษาจะเป็นทิศทางเข้าเมืองนครราชสีมาดังแสดงในรูปที่ 3

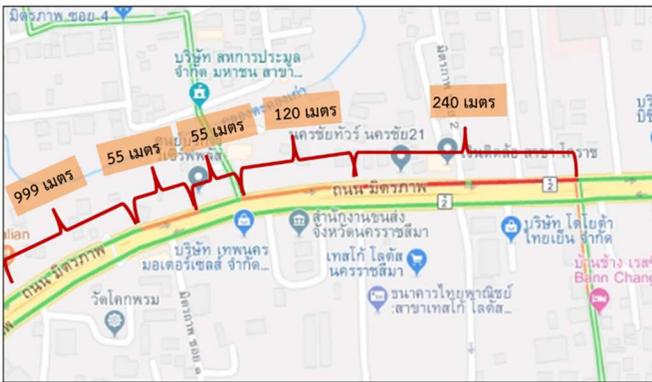


รูปที่ 3 บริเวณพื้นที่ที่ทำการศึกษา

3.1 ข้อมูลจาก Google Maps

การเก็บข้อมูลจาก Google Maps จะใช้การแคปเจอร์ภาพหน้าจอคอมพิวเตอร์ (Screenshot) ที่เปิดใช้งาน Google Maps บริเวณที่ทำการสำรวจข้อมูล (ช่วงเวลาเดียวกันกับการเก็บข้อมูลระยะเวลาความยาวแฉกคยจริง ณ จุดสำรวจ) โดยใช้โปรแกรม Auto Screen Capture ที่เป็นตัวช่วยในการบันทึกภาพหน้าจอแบบอัตโนมัติโดยที่จะบันทึกทุก ๆ 1 นาที และในขณะเดียวกันบนหน้าเว็บไซต์ของ Google Maps ก็จะถูกตั้งระบบให้มีการ Refresh หน้าเว็บไซต์ตลอดเวลาโดยใช้โปรแกรม Auto Refresh เพื่อให้เห็นถึงการเปลี่ยนแปลงลักษณะของแถบสีอย่างต่อเนื่อง จากการสังเกตพบว่า ระยะความยาวแถบสีบน Google Maps จะแสดงความยาวเป็นช่วง ๆ แต่ละช่วงจะมีความยาวอยู่ระหว่าง 110-140 เมตร และจะแสดงลักษณะของสีตามความเร็วเฉลี่ยของรถที่อยู่บริเวณนั้น ในกรณีของพื้นที่ศึกษาซึ่งไม่มีทางแยกอยู่ก่อนหน้าพบว่าแถบสีเขียวจะเป็นแถบสีสุดท้ายเสมอ และยังพบอีกว่าความยาวของแถบสีที่อยู่ก่อนหน้าแถบสีเขียวจะสามารถแปรผันเพิ่มขึ้นหรือลดลงครั้งละประมาณ 25-35 เมตร รูปที่ 4 แสดงตัวอย่างแถบ

สีบนถนนซึ่งเป็นพื้นที่ศึกษาซึ่งแสดงช่วงความยาวออกเป็น 4 ช่วง ได้แก่ ช่วงที่ 1 มีความยาว 240 เมตร แสดงแถบสีแดง หมายถึงรถติดแต่สามารถเคลื่อนที่ได้ ช่วงที่ 2 มีความยาว 120 เมตร แสดงแถบสีส้มหมายถึงรถสามารถเคลื่อนที่ได้แบบช้า ๆ ช่วงที่ 3 มีความยาว 55 เมตร แสดงแถบสีเขียวหมายถึงถนนโล่งรถเคลื่อนที่ได้ปกติ และช่วงที่ 4 แสดงแถบสีส้มหมายถึงรถสามารถเคลื่อนที่ได้แบบช้า ๆ โดยส่วนใหญ่พบว่าแถบสีเขียว ซึ่งเป็นแถบสีสุดท้ายจะยาวต่อเนื่องไปอีกเป็นระยะทางไกล แต่เนื่องจากไม่สามารถมองเห็นได้ภายในขอบเขตของจอคอมพิวเตอร์จึงระบุแถบสีที่ 5 ให้เป็นแถบสีเขียวและให้ข้อมูลเป็นตัวเลข 999 เพื่อแสดงถึงระยะทางของแถบสีที่ค่อนข้างยาวต่อเนื่องไปจนสิ้นขอบจอคอมพิวเตอร์ ตัวอย่างของการกรอกข้อมูลระยะความยาวจากการสำรวจของ Google Maps แสดงดังตารางที่ 1



รูปที่ 4 แสดงความยาวแถบสีบน Google Maps

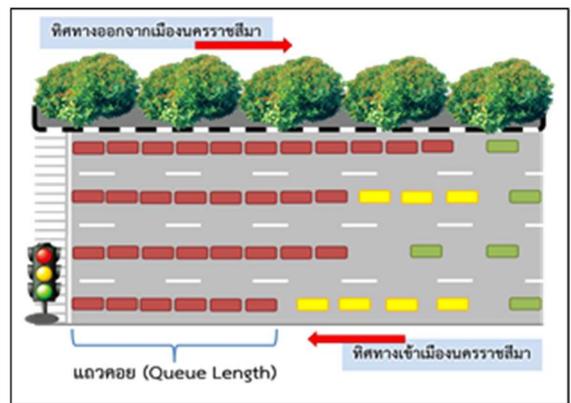
ตารางที่ 1 ตารางแสดงตัวอย่างข้อมูลจาก Google Maps

Items	Color	Length
1st color from stop line	Red	240 Meter
2nd color from stop line	Orange	120 Meter
3rd color from stop line	Green	55 Meter
4th color from stop line	Orange	55 Meter
5th color from stop line	Green	999 Meter

3.2 การเก็บข้อมูลระยะความยาวแถวคอยจริง

การเก็บระยะความยาวแถวคอยจริงบนพื้นที่ศึกษาได้จากการสังเกตของผู้สำรวจโดยจะมีการวัดระยะจากเส้นหยุดทุก ๆ 20 เมตร และทำเครื่องหมายเพื่อบอกระยะความยาวบริเวณ

ไหล่ทางหรือบนพื้นที่ฟุตบอลเพื่อให้ง่ายต่อการบันทึกข้อมูล พร้อมทั้งจัดทำแผนที่บริเวณพื้นที่ศึกษาโดยระบุระยะความยาวจากเส้นหยุดไปยังจุดสังเกตต่าง ๆ เช่น ตึก/อาคาร เสาไฟ ป้ายโฆษณา และสิ่งก่อสร้างอื่นที่มองเห็นได้ง่ายในระยะไกลเพื่อใช้เป็นตัวช่วยในการเก็บระยะความยาวแถวคอยจริงด้วย อย่างไรก็ตาม ระยะความยาวแถวคอยของแต่ละช่องจราจรอาจจะไม่เท่ากันแต่โดยส่วนใหญ่แล้วจะมีค่าใกล้เคียงกันดังแสดงในรูปที่ 5 งานวิจัยนี้จึงเลือกใช้ค่าเฉลี่ยของระยะความยาวแถวคอยของถนนทั้ง 4 ช่องจราจร เพื่อเป็นตัวแทนของระยะความยาวแถวคอยที่จะนำไปใช้สำหรับเป็นตัวแปรตามในแบบจำลอง



รูปที่ 5 แนวคิดการเก็บระยะความยาวแถวคอยของถนน 4 ช่องจราจร

หลังจากนั้นจะทำการบันทึกระยะเวลาแถวคอยจริงทุก ๆ 1 นาที เช่นเดียวกับการ Capture หน้าจอคอมพิวเตอร์เพื่อให้เวลาตรงกันเนื่องจากทำการบันทึกระยะเวลาทุก ๆ 1 นาทีทำให้บางครั้งอาจจะอยู่ในช่วงเวลาไฟเขียวที่รถสามารถเคลื่อนที่ได้โดยไม่เกิดระยะความยาวแถวคอยก็จะทำการระบุว่าไม่มีแถวคอยหรือมีระยะความยาวแถวคอยเป็น 0 นั่นเอง หลังจากนั้นจะนำระยะความยาวแถวคอยจริงนี้มาใช้เป็นข้อมูลตัวแปรตามสำหรับนำเข้าไปในแบบจำลองเพื่อการเรียนรู้และปรับแก้จนได้แบบจำลองที่สามารถพยากรณ์ได้อย่างมีประสิทธิภาพแล้วแบ่งข้อมูลความยาวแถวคอยจริงบางส่วนเพื่อใช้ในการเปรียบเทียบกับระยะความยาวแถวคอยที่ได้จากการพยากรณ์ของแบบจำลองเพื่อการปรับปรุงแบบจำลองจนสามารถประมาณระยะความยาวแถวคอยได้ถูกต้องมากยิ่งขึ้นอีกด้วย

การสำรวจและเก็บข้อมูลเริ่มตั้งแต่วันที่ 07.00 น. ถึง 19.00 น. เพื่อให้ครอบคลุมเวลาในช่วงโมงเร่งด่วนที่มีปริมาณ

จราจรหนาแน่น และช่วงเวลาไม่เร่งด่วนที่ยานพาหนะสามารถเคลื่อนที่ได้ปกติ ซึ่งได้ทำการสำรวจเช่นเดียวกันนี้ทั้งในวันธรรมดาจำนวน 3 วัน คือ วันพุธที่ 4 เมษายน 2562 วันจันทร์ที่ 15 กรกฎาคม 2562 และวันอังคารที่ 23 กรกฎาคม 2562 และวันหยุดเสาร์อาทิตย์จำนวน 2 วัน คือ วันอาทิตย์ที่ 28 เมษายน 2562 และวันอาทิตย์ที่ 14 กรกฎาคม 2562 เพื่อการศึกษาปัจจัยต่าง ๆ ที่เกี่ยวข้องและมีความหลากหลายของช่วงเวลาในการเรียนรู้ของแบบจำลองการพยากรณ์

3.3 ข้อมูลและตัวแปร

ข้อมูลที่ใช้เพื่อนำเข้าในแบบจำลอง (Input Data) ได้แก่ ข้อมูลระยะความยาวแถบสีบน Google Maps ทุก ๆ 1 นาที ซึ่งจะเป็นตัวแปรอิสระ (X) ถูกแบ่งออกเป็น 24 ตัวแปร ดังแสดงในตารางที่ 2 โดยงานวิจัยนี้พิจารณาแถบสีที่แสดงบน Google Maps จำนวน 6 แถบสี ซึ่งเป็นจำนวนแถบสีที่อยู่ติดกันที่มากที่สุดบนถนนเส้นที่พิจารณาที่สามารถมองเห็นได้จากจอคอมพิวเตอร์ โดยจำนวนแถบสีอาจจะมากหรือน้อยกว่านี้ได้ขึ้นอยู่กับการย่อหรือขยายภาพหน้าจอ อย่างไรก็ตามในงานวิจัยนี้พยายามย่อภาพหน้าจอของ Google Maps ให้ได้ระยะทางจากทางแยกไปให้ไกลที่สุดโดยที่ยังสามารถมองเห็นรายละเอียดของแถบสีได้ และในหนึ่งแถบสีจะแสดงสีโดยขึ้นอยู่กับสภาพจราจรซึ่งอาจเป็นสีใดสีหนึ่งจาก 4 สี คือ สีน้ำตาลแดง สีแดง สีส้ม และสีเขียว ในการแสดงผลของแต่ละแถบสีจะแสดงเพียงครั้งละ 1 สีเท่านั้น ดังนั้นจึงระบุตัวแปรอิสระโดยใช้ตัวเลขแรกหมายถึงแถบสีที่เท่าไรและเลขตัวที่สองหมายถึงสีอะไร เช่น สมมติว่าบนถนนที่กำลังพิจารณาโดยเริ่มนับจากเส้นหยุดไปจนถึงปลายของถนนมีแถบสีที่มองเห็นบนหน้าจอเพียง 3 แถบสี แถบแรกเป็นสีแดง แถบที่สองเป็นสีเขียว และแถบที่สามเป็นสีส้ม ก็จะได้ตัวแปรเป็น X12, X24, X33 ตามลำดับ ส่วนตัวแปรตาม (Y) คือค่าความยาวแถวคอยจริงบริเวณทางแยก ซึ่งอยู่ในประเภทข้อมูลแบบต่อเนื่อง (Continuous Data) ทั้งตัวแปร X และ Y

นอกจากตัวแปรอิสระ (X) และตัวแปรตาม (Y) ที่ใช้นำเข้าแบบจำลองเพื่อให้หาความสัมพันธ์ระหว่างตัวแปรแล้วนั้นยังมีตัวแปรอีกบางส่วนที่ใช้เพื่อแบ่งกรณีในการศึกษา ได้แก่ วันในสัปดาห์ (Day_type), ช่วงเวลา (Peak/Off-Peak) ดังแสดงในตารางที่ 3

ตารางที่ 2 แสดงตัวแปรอิสระ (X) ระยะความยาวของแถบสีจาก Google Maps

	สีน้ำตาลแดง	สีแดง	สีส้ม	สีเขียว
แถบที่ 1	X11	X12	X13	X14
แถบที่ 2	X21	X22	X23	X24
แถบที่ 3	X31	X32	X33	X34
แถบที่ 4	X41	X42	X43	X44
แถบที่ 5	X51	X52	X53	X54
แถบที่ 6	X61	X62	X63	X64

ตารางที่ 3 แสดงตัวแปรนำเข้าแบบจำลองเพื่อใช้แบ่งกรณีศึกษา

ตัวแปรอื่นๆ	
Day_Type	Weekday (1) /Weekend (2)
Peak/Off-Peak	เร่งด่วน (1)/ไม่เร่งด่วน (2)

4. การวิเคราะห์ผล

4.1 การวิเคราะห์ลักษณะเบื้องต้นของข้อมูล

ข้อมูลจากการสำรวจของทิศทางขาเข้าเมืองนครราชสีมาจำนวนทั้งหมด 2,399 ชุดข้อมูล เนื่องจากงานวิจัยนี้ให้ความสำคัญกับระยะความยาวแถวคอยเป็นหลักซึ่งความยาวของแถวคอยนั้นก็ขึ้นอยู่กับรอบสัญญาณไฟจราจรจึงทำการพิจารณาข้อมูลใน 2 ลักษณะซึ่งในการแบ่งกรณีในการสร้างแบบจำลองก็แบ่งตามลักษณะนี้ด้วย ได้แก่ 1) ข้อมูลระยะความยาวแถวคอยในช่วงสัญญาณทั้งไฟเขียวและแดง (Green & Red) 2) ข้อมูลระยะความยาวแถวคอยเฉพาะในช่วงสัญญาณไฟแดง (Red)

ผลการคำนวณค่าทางสถิติเบื้องต้นเพื่อพิจารณาลักษณะของข้อมูลพบว่าการกระจายตัวของข้อมูลค่อนข้างมากเนื่องจากมีค่าความแปรปรวน (Variance) ค่อนข้างสูงและเมื่อพิจารณาเฉพาะข้อมูลช่วงเวลาเร่งด่วนจะเห็นได้ว่าการกระจายตัวของข้อมูลลดน้อยลงเป็นผลมาจากช่วงเวลาเร่งด่วนจะมีปริมาณจราจรสูงทำให้แถวคอยส่วนใหญ่มีค่าความยาวใกล้เคียงกันมากขึ้นต่างจากข้อมูลช่วงเวลาทั้งหมดที่บางเวลาแถวคอยสั้นมาก (หรือมีค่าเป็น 0 อยู่ในช่วงเวลาสัญญาณไฟ

เขียว) และบางช่วงมีระยะแถวคอยยาวมากการกระจายตัวของข้อมูลจึงมากและทำให้มีความแปรปรวนสูงตามมา ส่วนค่าความเบ้ของข้อมูล (Skewness) ที่มีค่ามากกว่า 0 ส่งผลให้

ทราบว่าคุณค่าส่วนใหญ่เบ้ไปทางขวาแสดงว่าระยะความยาวแถวคอยส่วนใหญ่มีค่ามากกว่าค่าเฉลี่ยของข้อมูลทั้งหมด

ตารางที่ 4 ค่าสถิติเบื้องต้นของระยะความยาวแถวคอยจริง

แถวคอย (Queue Length)	Traffic Lights	No. of cases	Mean (meter)	Variance	Skewness	Kurtosis
ระยะแถวคอยทิศทางเข้าเมืองทั้งหมด	Green & Red	2,399	128.60	12,153.61	0.87	0.19
	Red	1,647	106.54	9,512.30	1.18	1.20
ระยะแถวคอยทิศทางเข้าเมืองเฉพาะช่วงเร่งด่วน	Green & Red	364	119.97	10,081.38	1.09	1.33
	Red	262	112.25	9,409.68	1.31	2.06
ระยะแถวคอยทิศทางเข้าเมืองเฉพาะช่วงไม่เร่งด่วน	Green & Red	2,035	130.10	12,513.85	0.83	0.03
	Red	1,385	105.46	9,531.203	1.15	1.03

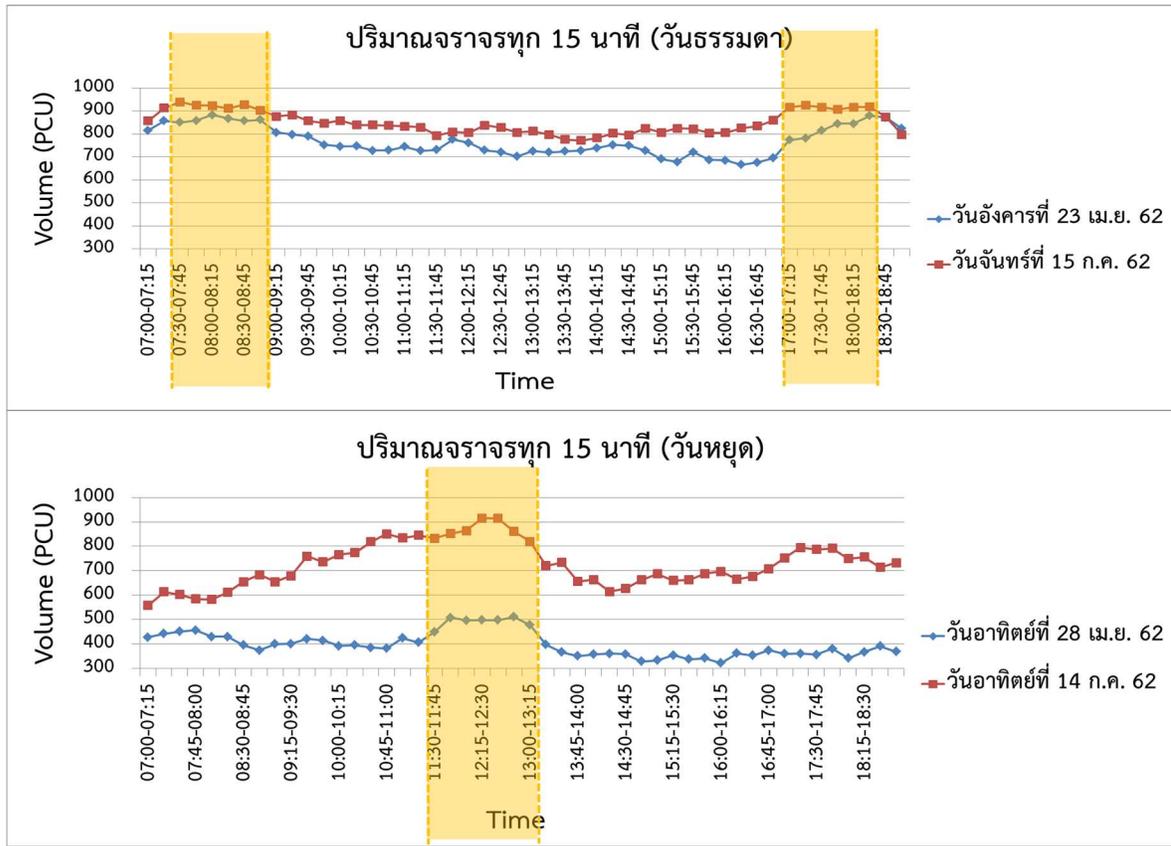
แต่เมื่อพิจารณาเปรียบเทียบข้อมูลระยะแถวคอยระหว่างในช่วง Green & Red กับ Red พบว่าในช่วงสัญญาณเฉพาะไฟแดง (Red) มีความแปรปรวนน้อยกว่าช่วงสัญญาณทั้งไฟเขียวและแดง (Green & Red) และในส่วนของความเบ้ก็มีค่าเพิ่มขึ้นแสดงให้เห็นว่าระยะความยาวแถวคอยจริงที่มีค่าเป็น 0 ส่วนใหญ่ได้ถูกตัดออกไปแล้ว แต่ก็ยังมีบางส่วนที่แสดงค่าระยะความยาวแถวคอยเป็น 0 เนื่องจากบางครั้งที่ทำการสำรวจข้อมูลอยู่ในช่วงสัญญาณไฟแดงแต่ไม่มีรถเคลื่อนที่เข้ามาบริเวณทางแยก พอติดกับช่วงเวลาที่ต้องบันทึกข้อมูลเมื่อครบทุก 1 นาที ก็จะส่งผลให้มีค่าความแปรปรวนของข้อมูลโดยภาพรวมทุกกรณีมีค่าค่อนข้างสูงดังแสดงในตารางที่ 4

4.2 การแบ่งประเภทข้อมูล

งานวิจัยนี้มีข้อสันนิษฐานว่าหากสภาพจราจรมีความหนาแน่นอาจจะทำให้ข้อมูลแถบสีจาก Google Maps มีความสัมพันธ์เพิ่มขึ้นกับระยะความยาวของแถวคอยจริง ดังนั้นในการศึกษาจึงได้ทำการแบ่งกรณี โดยกรณีแรกนั้นจะใช้ชุดข้อมูลทั้งหมด (Case1_All) โดยไม่มีการแยกว่าช่วงใดเป็นช่วงเร่งด่วนหรือไม่เร่งด่วน ซึ่งจะทำให้ได้แบบจำลองเพียงตัวเดียวสำหรับใช้พยากรณ์ทั้งช่วงเร่งด่วนและไม่เร่งด่วน กรณีที่สองจะเป็นการแยกข้อมูลตามสภาพจราจร ได้แก่ Case2.1_Peak

คือข้อมูลช่วงเวลาเร่งด่วน และ Case2.2_Off-Peak คือข้อมูลช่วงเวลาไม่เร่งด่วน ซึ่งจะทำให้ได้แบบจำลองแยกสำหรับแต่ละช่วงเวลา อย่างไรก็ตาม เพื่อให้สามารถเปรียบเทียบผลกันได้โดยตรง ผลการพยากรณ์ความยาวแถวคอยจากแบบจำลองที่แยกตามช่วงเวลาจะถูกนำมารวมกัน (Case2.1_Peak ร่วมกับ Case2.2_Off-Peak) ก่อนที่จะนำไปเปรียบเทียบกับผลการพยากรณ์จากกรณีที่ใช้ข้อมูลทั้งหมด (Case1_All)

การแบ่งข้อมูลช่วงเวลาเร่งด่วน (Peak) และไม่เร่งด่วน (Off-Peak) จะใช้ข้อมูลปริมาณจราจรเป็นตัวช่วยในการแยกช่วงเวลา ซึ่งจะให้ผลต่างกันระหว่างวันธรรมดาและวันหยุด จากการวิเคราะห์พบว่าในวันธรรมดามีปริมาณจราจรสะสมมากกว่าวันหยุดและในวันธรรมดามีช่วงเวลาเร่งด่วนเข้าอยู่ในช่วง 07:30 – 09:00 น. เป็นช่วงเวลาในการเดินทางของคนเพื่อไปทำกิจกรรมต่าง ๆ เช่น เดินทางไปทำงาน เดินทางไปส่งบุตรหลานที่โรงเรียน เป็นต้น และมีเวลาเร่งด่วนเย็นอยู่ในช่วง 17:00 – 18:30 น. ซึ่งก็เป็นเวลาเดินทางกลับจากการทำกิจกรรมต่าง ๆ นั่นเอง และในวันหยุดมีเวลาเร่งด่วนเพียงช่วงเดียวซึ่งอยู่ในช่วงเวลากลางวันคือเวลา 11:45 – 13:15 น. ดังแสดงในรูปที่ 6



รูปที่ 6 แสดงการแบ่งช่วงเวลาเร่งด่วนโดยใช้ปริมาณจราจรแยกวันธรรมดาและวันหยุด

4.3 การสร้างแบบจำลองการพยากรณ์

งานวิจัยนี้ได้เลือกวิธี Linear Regression มาพิจารณา เนื่องจากผลลัพธ์สามารถแสดงเป็นสมการและทำให้สามารถอธิบายความสัมพันธ์ระหว่างตัวแปรได้อย่างชัดเจน นอกจากนี้ยังได้เลือกวิธี Random Forest ซึ่งเป็นอีกหนึ่งวิธีที่ง่ายและได้รับความนิยม [15] และอีกหนึ่งวิธีเป็นการใช้ค่าเฉลี่ย (Mean) ของความยาวแถวคอยที่สำรวจได้ในอดีตมาใช้ในการพยากรณ์ความยาวแถวคอยในอนาคตโดยแยกการหาค่าเฉลี่ยสำหรับการพยากรณ์ตามกรณี ได้แก่ ช่วงเร่งด่วนหรือไม่เร่งด่วน และ วันธรรมดาหรือวันหยุด ซึ่งสามารถใช้วิธีค่าเฉลี่ยเป็นบรรทัดฐานในการเปรียบเทียบกับวิธีอื่น ๆ เพื่อแสดงให้เห็นถึงความสามารถของแบบจำลองวิธีอื่นในการตรวจจับการแปรผันของความยาวแถวคอยเมื่อเทียบกับการใช้ค่าเฉลี่ยตามช่วงเวลา

การสร้างแบบจำลองนั้น ตัวแปรอิสระ (X) จะเป็นข้อมูลระยะความยาวแถบสีจาก Google Maps และตัวแปรตาม (Y) เป็นระยะความยาวแถวคอยจริง โดยจะมีการพิจารณาครอบคลุม 2 กรณีย่อย กรณีแรกจะนำตัวแปรอิสระทั้งหมด

จากหกแถบสี (X11-X64) มาพิจารณาสร้างแบบจำลอง และกรณีที่สองจะทำการพิจารณาตัวแปรอิสระเฉพาะแถบสีที่ 1-3 (X11-X34) ซึ่งเป็นแถบสีสามแถบแรกที่อยู่ติดกันบนทางแยกที่คาดว่าจะมีความสัมพันธ์กับความยาวแถวคอยมากที่สุด

ในภาพรวมของการสร้างแบบจำลองทั้งหมด จะมีจำนวนกรณีแบบจำลองย่อยทั้งสิ้น 24 กรณี โดยลำดับแรกจะพิจารณาจากการแบ่งชุดข้อมูลตามช่วงเร่งด่วนหรือช่วงไม่เร่งด่วน (Case2.1_Peak รวมกับ Case2.2_Off-Peak) หรือไม่มีการแบ่ง (Case1_All) และการพิจารณาตัวแปรว่าใช้ทั้ง 6 แถบสี (ตัวแปรทั้งหมด) หรือใช้เพียง 3 แถบสีแรก (ตัวแปรสามแถบแรก) ลักษณะการแบ่งดังกล่าวสามารถแบ่งประเภทของแบบจำลองได้ 4 กรณีดังนี้

- 1) Model_Type_1 คือ Case1_All ตัวแปรทั้งหมด
- 2) Model_Type_2 คือ Case2.1_Peak รวมกับ Case2.2_Off-Peak ตัวแปรทั้งหมด
- 3) Model_Type_3 คือ Case1_All ตัวแปรสามแถบแรก

4) Model_Type_4 คือ Case2.1_Peak ร่วมกับ Case2.2_Off-Peak ตัวแปรสามแถบแรก

ลำดับที่สองจะพิจารณาจากช่วงสัญญาณไฟจราจร ซึ่งถูกแบ่งออกเป็น 2 กรณี คือ 1) กรณีที่พิจารณารวมช่วงสัญญาณทั้งไฟเขียวและแดง (Green & Red) และ 2) กรณีที่พิจารณาเฉพาะช่วงสัญญาณไฟแดง (Red) และการแบ่งกรณีลำดับสุดท้ายคือการแบ่งตามวิธีการศึกษา คือ 1) Linear Regression Model (LRM) 2) Random Forest Model (RFM) และ 3) วิธีการใช้ค่าเฉลี่ย (Mean) ดังนั้นรวมทั้งสิ้นจะมี 24 กรณี (4x2x3) เพื่อใช้ในการวิเคราะห์ผลในลำดับต่อไป

นอกจากนั้นชุดข้อมูลจากทุก Case จะถูกแบ่งข้อมูลของทั้งวันทำงานและวันหยุดออกเป็น 2 กลุ่ม กลุ่มแรกจำนวน 90% ของข้อมูลสำหรับการ Train และอีก 10% สำหรับการ Test ซึ่งข้อมูลถูกแบ่งออกจากกันชัดเจนไม่ได้ใช้ร่วมกัน โดยข้อมูลของทั้ง Train และ Test จะมีทั้งช่วงวันธรรมดาและวันหยุดเหมือนกัน การทำดังกล่าวทำให้แบบจำลองผ่านการ Train ครอบคลุมสภาพจราจรจากทั้งวันหยุดและวันธรรมดารวมถึงช่วงเร่งด่วนและไม่เร่งด่วน และการ Test ก็ครอบคลุมสภาพจราจรที่มีความหลากหลายเช่นกัน โดยทุกวิธีการ LRM, RFM และวิธี Mean จะอาศัยกระบวนการ Cross-Validation ในการแบ่งข้อมูลชุดแรกออกเป็นหลายส่วน เช่น 5-Fold Cross-Validation จากการศึกษาส่วนใหญ่การทำ Cross-Validation จะถูกแบ่งเป็น 5 หรือ 10 หรือ 100 ส่วน [15] โดยการแบ่งส่วนที่มากขึ้นจะทำให้เปรียบเทียบมีการพิจารณาทดสอบกับชุดข้อมูลที่ต่างกันหลายชุด ซึ่งผลการทดสอบจะสะท้อนความสามารถในการพยากรณ์ของแบบจำลองได้ดียิ่งขึ้น อย่างไรก็ตามการแบ่งส่วนที่มากเกินไปจะทำให้แต่ละส่วนมีชุดข้อมูลน้อยและทำให้ใช้เวลาในการประมวลผลนานขึ้น ซึ่งยังไม่มีข้อแนะนำที่ชัดเจนว่าควรใช้จำนวนส่วนเท่าใด การวิจัยในครั้งนี้จึงเลือกจำนวน 5 ส่วน ซึ่งไม่มากหรือน้อยจนเกินไป

การสร้างแบบจำลองโดยวิธีการของ Linear Regression Model (LRM) ใช้การเลือกตัวแปรโดยวิธีลดตัวแปร (Backward Elimination) เป็นวิธีการคัดเลือกตัวแปรอิสระออกจากสมการที่ละตัวแปร โดยเริ่มจากการสร้างสมการถดถอยที่นำตัวแปรอิสระทุกตัวเข้าสู่สมการ แล้วจึงคัดเลือกตัวแปรอิสระออกทีละตัวโดยตัวแปรอิสระที่มีความสัมพันธ์กับตัว

แปรตามน้อยที่สุดและไม่มีความสำคัญจะถูกตัดออกไปเรื่อย ๆ จนกระทั่งเหลือแค่ตัวแปรที่มีนัยสำคัญเท่านั้น

ส่วนแบบจำลองจากวิธีการ Random Forest Model (RFM) เรียกใช้ Packages ที่มีชื่อว่า Caret โดยใช้ Method เป็นของ Ranger ซึ่งเป็น Algorithms ที่เอาไว้ออกใช้กับ RFM โดยเฉพาะ วิธี RFM จาก Ranger มี hyperparameter ที่ต้องการปรับจูนค่า (Tuning) ทั้งสิ้น 2 ตัวได้แก่ Mtry และ Min.Node.Size ในการศึกษานี้มีการกำหนดช่วงของจำนวนตัวแปรที่อาจจะแบ่งได้ในแต่ละโหนด (Mtry) ไว้ระหว่าง 2-10 ตัวแปร และขนาดโหนดที่น้อยที่สุด (Min.Node.Size) อยู่ในช่วง 5-10 โหนด ซึ่งการกำหนดช่วงดังกล่าวนี้มีความครอบคลุมค่าที่สมเหตุสมผลของพารามิเตอร์แต่ละตัวสำหรับการศึกษานี้ โดย Caret Package มีความสามารถในการปรับจูน (Tuning) เพื่อหาค่า Hyperparameter ที่มีค่าเหมาะสมจากช่วงที่กำหนดได้โดยอัตโนมัติ

การวิเคราะห์การเปรียบเทียบระหว่างค่าระยะความยาวแถวคอยจริงซึ่งได้จากข้อมูล Testing ที่แบ่งไว้ 10% กับค่าที่ได้จากการพยากรณ์ของแต่ละแบบจำลองซึ่งได้แก่ Linear Regression Model (LRM) และ Random Forest Model (RFM) และวิธีการเปรียบเทียบกับค่าเฉลี่ย (Mean) หลังจากนั้นจึงจะพิจารณาประสิทธิภาพในการพยากรณ์ของแบบจำลองจากราคาค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error, RMSE) ได้ผลลัพธ์แสดงดังตารางที่ 5

จากตารางข้างต้นพบว่าแบบจำลองที่อยู่ในช่วงสัญญาณเฉพาะไฟแดงมีค่า RMSE น้อยกว่าแบบจำลองที่อยู่ในช่วงทั้งสัญญาณไฟเขียวและแดงซึ่งเป็นไปในทางเดียวกันทั้ง 3 วิธีการของ LRM, RFM และ Mean แสดงให้เห็นว่าระยะความยาวแถวคอย Google Maps จะสามารถพยากรณ์ค่าความยาวแถวคอยได้ดีขึ้นหากพิจารณาเฉพาะช่วงเวลาไฟแดง ดังนั้นจึงจะทำการเปรียบเทียบกรณีอื่น ๆ ที่อยู่ในช่วงสัญญาณไฟแดงเท่านั้น และเมื่อพิจารณาแต่ละวิธีการที่ใช้สร้างแบบจำลองพบว่าการใช้วิธีการ RFM มีประสิทธิภาพดีที่สุด เนื่องจากมีการวิเคราะห์ในรูปแบบการสร้างป่าไม้ที่มีการวิเคราะห์ค่อนข้างละเอียด รองลงมาคือ วิธีการ LRM และอันดับสุดท้ายคือวิธีการการเปรียบเทียบกับค่า Mean ที่มีค่า RMSE มากที่สุด

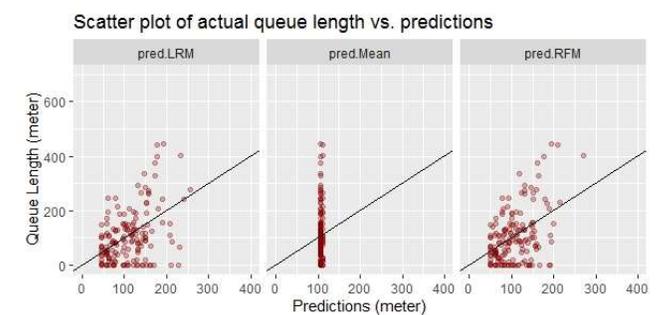
ตารางที่ 5 ประสิทธิภาพการพยากรณ์ของแบบจำลอง

	Traffic Lights	RMSE of LRM (m)	RMSE of RFM (m)	RMSE of Mean (m)
Model_Type_1 Case1_All ตัวแปรทั้งหมด	Green & Red	95.4	95.8	113.0
	Red	87.8	85.2	94.9
Model_Type_2 Case2.1_Peak รวมกับ Case2.2_Off- Peak ตัวแปร ทั้งหมด	Green & Red	93.8	97.2	113.0
	Red	90.8	86.8	94.9
Model_Type_3 Case1_All ตัว แปรสามแถบ แรก	Green & Red	96.8	97.9	104.0
	Red	87.4	85.3	101.0
Model_Type_4 Case2.1_Peak รวมกับ Case2.2_Off- Peak ตัวแปร สามแถบแรก	Green & Red	96.7	97.1	104.0
	Red	85.1	83.2	101.0

เมื่อพิจารณาเปรียบเทียบกันระหว่างแบบจำลองประเภทที่ 1 และ 2 (Model_Type_1 vs. Model_Type_2) ซึ่งเป็นการใช้ข้อมูลจากตัวแปรทั้งหมดและพิจารณาเฉพาะช่วงเวลาไฟแดงอย่างเดียว พบว่าแบบจำลองประเภทที่ 1 (Model_Type_1) มีค่า RMSE น้อยกว่าแบบจำลองประเภทที่ 2 (Model_Type_2) โดยให้ผลลัพธ์เหมือนกันทั้งวิธีการ LRM (87.8 meter เทียบกับ 90.8 meter) และ RFM (85.2 meter เทียบกับ 86.8 meter) ซึ่งได้ผลแตกต่างไปจากที่คาดหวังไว้

อย่างไรก็ตามเมื่อมีการตัดตัวแปรนำเข้าไปให้เหลือเฉพาะตัวแปรจากสามแถบแรก (Model_Type_3 และ Model_Type_4) พบว่าแบบจำลอง Model_Type_4 มีค่า RMSE ลดลงจากแบบจำลอง Model_Type_2 อย่างเห็นได้ชัด และยังมีค่า RMSE น้อยกว่าแบบจำลอง Model_Type_3 ทั้งวิธีการ LRM (87.4 meter เทียบกับ 85.1 meter) และ RFM (85.3 meter เทียบกับ 83.2 meter) เนื่องจากมีการตัดตัว

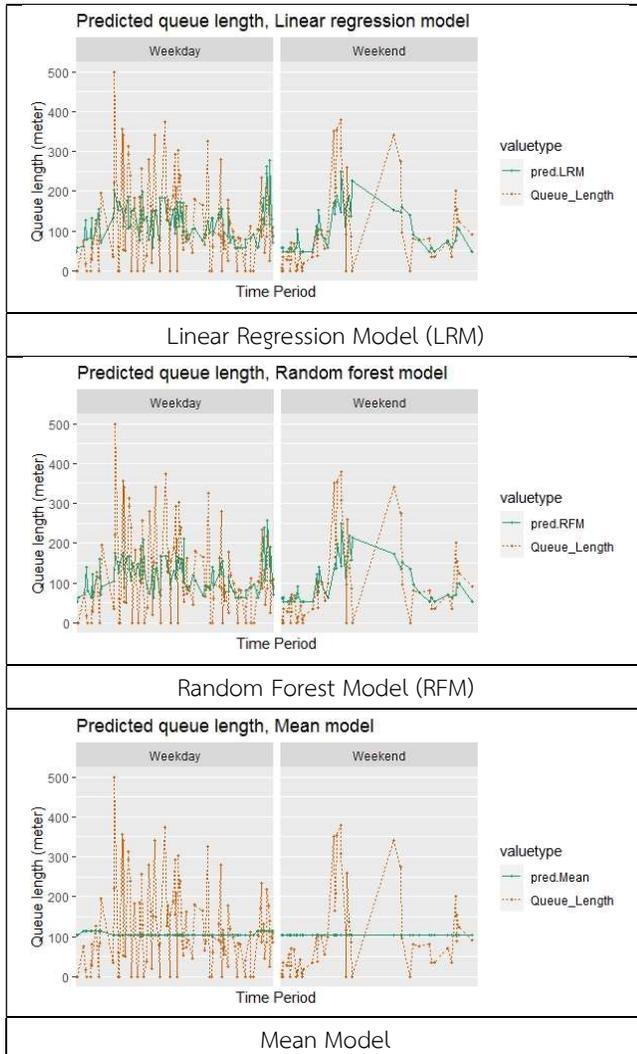
แปร X ที่ส่งผลกับตัวแปร Y น้อยออกไป โดยการตัดตัวแปร X ตั้งแต่แถบสีที่ 4-6 (X41-X64) ออกไปทำให้การวิเคราะห์แยกกรณีตามช่วงเวลาเร่งด่วนและไม่เร่งด่วนให้ผลการศึกษาที่มีค่าคลาดเคลื่อนน้อยลง ส่วนการวิเคราะห์โดยวิธีการใช้ค่า Mean ยังให้ค่าความคลาดเคลื่อนมากที่สุดเช่นเคย ส่งผลให้แบบจำลองที่ให้ผลการพยากรณ์ที่ดีที่สุดคือ แบบจำลองที่ 4 ที่มีการแยกกรณีตามช่วงเวลาเร่งด่วนและมีข้อมูลอยู่ในช่วงเวลาเฉพาะสัญญาณไฟแดงเท่านั้นและใช้วิธีการ RFM ในการสร้างแบบจำลองโดยให้ค่าความคลาดเคลื่อนน้อยที่สุดอยู่ที่ 83.2 เมตร



รูปที่ 7 แสดงผลการพยากรณ์ของแบบจำลองที่ 4 แยกตามช่วงเวลาและอยู่ในช่วงสัญญาณเฉพาะไฟแดง

ผลการพยากรณ์ดังกล่าวสามารถนำมาสร้างแผนภาพการกระจาย (Scatter Plot) ระหว่างค่าระยะความยาวแถวคอยจริง (Queue-Length) กับค่าระยะแถวคอยที่ได้จากการพยากรณ์ (Predictions) ดังแสดงในรูปที่ 7 จะเห็นได้ว่าแผนภาพของวิธีการ RFM มีความสัมพันธ์กันในเชิงบวกมากกว่าวิธีอื่น ๆ แต่ระดับของความสัมพันธ์ก็ยังไม่สูงมากนัก โดยช่วงที่ค่าระยะความยาวแถวคอยจริงมีค่ามากขึ้น ค่าที่ได้จากการพยากรณ์ก็จะให้ค่ามากด้วยเช่นกัน ซึ่งจะเห็นความสัมพันธ์ดังที่กล่าวมาข้างต้นจากแบบจำลองของ LRM และ RFM ส่วนแบบจำลองที่ได้จากการหาค่าเฉลี่ย (Mean) จะให้กราฟเป็นเส้นตรงในระนาบเดียวกันซึ่งจะเป็นความสัมพันธ์ที่ไม่ดีเท่ากับแบบจำลองทั้งสองที่กล่าวมา เมื่อนำค่าจริงของระยะความยาวแถวคอยจากชุดข้อมูล test มาสร้างแผนภาพตามเวลาเทียบกับผลการพยากรณ์ดังแสดงในรูปที่ 8 จะพบว่าแบบจำลอง RFM และ LRM มีความสามารถในการตรวจจับ

แนวโน้มเฉลี่ยของการแปรผันของแถวคอยได้ดีในระดับหนึ่ง โดยมีแนวโน้มเพิ่มและลดลงตามค่าที่สำรวจได้จริง อย่างไรก็ตามทั้งสองวิธียังไม่สามารถตรวจจับช่วงที่แถวคอยมีค่าเพิ่มขึ้นสูงหรือลดลงต่ำมาก ๆ ได้ดีเท่าที่ควร



รูปที่ 8 แสดงความสัมพันธ์ระหว่างระยะแถวคอยจริงกับค่าที่พยากรณ์ได้ตามช่วงเวลา

ถึงแม้แบบจำลองจะสามารถตรวจจับแนวโน้มการแปรผันในภาพรวมใหญ่ได้ แต่ก็ยังไม่สามารถสะท้อนการแปรผันแบบละเอียดออกมาได้ ทั้งนี้เนื่องจากการเก็บข้อมูลความยาวแถวคอยจะเก็บทุก ๆ 1 นาที ซึ่งอาจจะตรงกับช่วงที่เริ่มสัญญาณไฟแดงที่มีแถวคอยสั้นมากหรือช่วงเวลาไฟแดงได้สักระยะหนึ่งแล้วทำให้มีแถวคอยยาวมาก ดังนั้นความยาวแถวคอยจริงที่สำรวจได้จึงอาจมีค่าค่อนข้างเหวี่ยงขึ้นลง แต่ถ้ามองในภาพรวมก็สามารถมองเห็นแนวโน้มการเพิ่มขึ้นหรือ

ลดลงของความยาวแถวคอยในภาพใหญ่ได้ ซึ่งอาจจะยังพบว่ามีความคลาดเคลื่อนอยู่บ้าง ดังจะเห็นได้ในรูปที่ 7 และ 8 สาเหตุอีกส่วนหนึ่งน่าจะมาจากลักษณะของข้อมูลความยาวแถวคอยจาก Google Maps โดยเฉพาะแถวคอยแรก ๆ ที่อยู่ติดกับทางแยกซึ่งมักจะแสดงความยาวเป็นช่วง ๆ เพิ่มขึ้นช่วงละประมาณ 110-140 เมตร ซึ่งถือว่ามีความละเอียดของข้อมูลค่อนข้างต่ำ ดังนั้นการนำข้อมูลจากแถวคอยดังกล่าวซึ่งมีความละเอียดต่ำไปใช้เป็นตัวแปรอิสระในการสร้างแบบจำลองในการพยากรณ์ระยะความยาวแถวคอยที่มีการแปรผันในระดับเมตรตลอดเวลานั้นก็อาจจะมีขีดจำกัดในด้านของความแม่นยำ อย่างไรก็ตาม ถึงแม้จะมีข้อจำกัดด้านความละเอียดของข้อมูลดังกล่าวแต่แบบจำลอง RFM ก็ยังสามารถตรวจจับการแปรผันได้ดีระดับหนึ่งโดยเฉพาะกรณีที่มีการแยกระหว่างช่วงเร่งด่วนและไม่เร่งด่วน และยังให้ผลการพยากรณ์ที่แม่นยำและสามารถเชื่อถือได้มากกว่าการใช้ค่าเฉลี่ยโดยตรง นอกจากนี้หากในอนาคตทาง Google Maps มีการปรับความละเอียดของความยาวแถวคอยโดยเฉพาะแถวคอยแรก ๆ ที่อยู่ติดกับทางแยก ก็อาจจะทำให้แนวทางที่ใช้ในการศึกษาได้ผลการพยากรณ์ที่ดีขึ้นก็เป็นได้

5. สรุปผลการศึกษา

จากการศึกษาการประมาณระยะความยาวแถวคอยโดยใช้ Google Maps พบว่าตัวแปรอิสระ (X) ที่ได้จากระยะความยาวแถวคอยของ Google Maps มีความสามารถในการพยากรณ์ค่าแนวโน้มของตัวแปรตาม (Y) ซึ่งได้แก่ค่าระยะความยาวแถวคอยจริงได้ดีในระดับหนึ่ง โดยแบบจำลองที่ให้ผลการพยากรณ์ที่ดีที่สุดได้แก่วิธี Random Forest Model (RFM) โดยการศึกษาครั้งนี้ใช้ข้อมูลทิศทางการเข้าตัวเมืองนครราชสีมาพบว่าแบบจำลองจะสามารถพยากรณ์ค่าได้ดีเมื่อมีการแยกช่วงเวลาตามเวลาเร่งด่วนและไม่เร่งด่วนในการทดสอบแบบจำลองแล้วค่อยนำมาวิเคราะห์รวมกัน ตัวแปรอิสระ (X) ที่มีความสัมพันธ์หรือส่งผลกับตัวแปรตาม (Y) คือ ตัวแปรอิสระที่มาจากสามแถวคอยแรก ได้แก่ ตัวแปร X11-X34 และอยู่ในช่วงเวลาสัญญาณไฟแดงเท่านั้น ซึ่งในอนาคตจะสามารถใช้รูปแบบของแบบจำลองจากการศึกษานี้เพื่อสร้างแบบจำลองของบริเวณทางแยกและทิศทางอื่น ๆ ได้ และเมื่อทราบถึงความยาวของแถวคอยก็จะสามารถคำนวณเวลาในการจัดการ

กับแถวคอยได้ เมื่อมีรถฉุกเฉินเดินทางเข้ามาในบริเวณทางแยกไฟสัญญาณจราจรก็จะเคลื่อนผ่านไปโดยไม่ต้องหยุดรถ ซึ่งจะเป็นแนวทางของการศึกษาในลำดับต่อไป

6. ข้อเสนอแนะ

ควรนำแบบจำลองไปทดสอบกับทางแยกอื่นที่อาจมีปัจจัยที่แตกต่างกันในการศึกษา ซึ่งจะเป็นการทดสอบความสามารถของแบบจำลองในด้านของ Transferability และควรสร้างแบบจำลองโดยวิธีการที่หลากหลายมากขึ้น เช่น วิธีการ Support Vector Machine (SVM) และ K-Nearest Neighbors (KNN) ซึ่งอาจจะให้ผลการพิจารณาที่แตกต่างและเป็นประโยชน์ในการศึกษา และการศึกษาครั้งต่อไปควรมีการถ่ายรูปหรืออัดวิดีโอของสถานที่จริงตลอดทุก Sample Point เพื่อบอกได้ว่าปลายแถวของแถวคอยอยู่ตำแหน่งใดจะได้อ้างอิงกับผลการพยากรณ์ ในภายหลังเพื่อองค์ความรู้ใหม่ และควรทำการวิเคราะห์ในหลากหลายกรณีศึกษาที่มีปัจจัยแตกต่างกัน เช่น กรณีทางแยกที่มีหรือไม่มีทางแยกก่อนหน้าอยู่ในบริเวณใกล้เคียง ช่วงเวลาเร่งด่วนไม่เร่งด่วน แยกวันทำงานวันหยุด และกรณีที่แยกแถวคอยในช่วงเวลาสัญญาณไฟเขียวที่มีความยาวเป็นศูนย์ออก ซึ่งจะช่วยให้เข้าใจถึงปัจจัยที่ส่งผลต่อประสิทธิภาพในการพยากรณ์ความยาวแถวคอยจากแถบสีได้ดียิ่งขึ้น

กิตติกรรมประกาศ

การวิจัยครั้งนี้ได้รับการสนับสนุนจากทุนกิตติบัณฑิตมหาวิทยาลัยเทคโนโลยีสุรนารี ขอขอบคุณหน่วยงานตำรวจนครราชสีมาที่อำนวยความสะดวกในการเก็บข้อมูล

เอกสารอ้างอิง

- [1] อูรา สุวรรณรักษ์ (บรรณาธิการ). *แผนหลักการแพทย์ฉุกเฉินแห่งชาติ ฉบับที่ 3.1 พ.ศ. (2562-2564)*. นนทบุรี: อัลทิเมท พรินติ้ง; 2561.
- [2] Lighthill MJ, Whitham GB. On kinematic waves II. A theory of traffic flow on long crowded roads. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*. 1955;229(1178): 317–345.

- [3] Richards PJ. Shock waves on the highway. In: Reilly, W. *Highway Capacity Manual 2000*. 4th ed. Washington, DC: National Research Council; 1956. p. 42–51.
- [4] Stephanopoulos G, Michalopoulos PG. Modelling and analysis of traffic queue dynamics at signalized intersections. *Transportation Research Part A*. 1979;13(5): 295-307.
- [5] Liu HX, Wu X, Ma W, Hu H. Real-time queue length estimation for congested signalized intersections. *Transportation Research Part C*. 2009;17: 412-427.
- [6] Webster FV, Cobbe BM. *Traffic Signals*. London: Her Majesty's Stationery Office; 1966.
- [7] Cronje WB. Analysis of existing formulas for delay, overflow, and stops. *Transportation Research*. 1983;905: 89–93.
- [8] Balke K, Charara H, Parker R. *Development of a Traffic Signal Performance Measurement System (TSPMS)*. Texas (Austin): Department of Transportation Research and Technology Implementation Office 2005.
- [9] Ban XJ, Hao P, Sun Z. Real time queue length estimation for signalized intersections using travel times from mobile sensors. *Transportation Research Part C*. 2011;19(6): 1133-1156.
- [10] Wu A, Yang X. Real-time queue length estimation of signalized intersections based on RFID data. *Procedia- Social and Behavioral Sciences*. 2013;96: 1477-1484.
- [11] อภิรักษ์ บุตรละ. *การประยุกต์ใช้ Google Map ในการพัฒนาระบบการคำนวณค่ารถ taxi ในเขตพื้นที่กรุงเทพมหานครและปริมณฑล*. วิทยานิพนธ์วิทยาศาสตรมหาบัณฑิต บัณฑิตวิทยาลัย มหาวิทยาลัยกรุงเทพ 2553.
- [12] นันทชัย กานตานันทะ. การพยากรณ์ด้วยวิธีการพยากรณ์เชิงสาเหตุ. *วารสารวิศวกรรมศาสตร์ (ISSN: 1906-3636)*. 2555;4(1): 35-48.
- [13] พัฒนพงษ์ ตลรัตน์. การเปรียบเทียบประสิทธิภาพของแบบจำลองในการพยากรณ์ความสำเร็จการศึกษาของ

นักเรียนระดับประกาศนียบัตรวิชาชีพ. *วารสาร
วิทยาศาสตร์และเทคโนโลยี*. 2560;37(3): 380-388.

- [14] Nicholas S. *Biostar Bioinformatics Explained*.
Available from: <https://www.biostars.org/p/86981/> [Accessed: 04th June 2020].
- [15] ชนิษฐา ดีสุบิน. การพัฒนาตัวแบบการพยากรณ์ความ
ถนัดทางการเรียนตามทฤษฎี 4 MAT โดยการวิเคราะห์
ด้วยวิธีต้นไม้. *วารสารนวัตกรรมการศึกษา*. 2560;2(1):
43-58.