Trip Attraction Model using Social Network Data and Deep Learning

Wirach Hirun[1,*], Titaporn Pobutdee[2]

[1,2]Faculty of Science and Engineering, Kasetsart University

Chalermphrakiat Sakon Nakhon Province Campus

Received: 31 July 2018

Revised: 17 September 2018

Accepted: 7 November 2018

ABSTRACT

This research aims to develop a trip attraction model using online social network data. This research used the tourist data of the national park as a case study. The deep learning approach was employed for developing a model. The performance of the model was then compared against linear regression analysis method and artificial neural network method. The results indicated that check-in data could be used to create the trip attraction model. The model derived from deep learning approach provided the highest predictive accuracy. The mean percentage absolute error (MAPE) of deep learning method was 43.99, which was lower than those of the linear regression analysis method (386.48) and the artificial neural network method (88.97). As such, developing a model using social network data will greatly reduce the cost of data surveying.

Keywords: Trip attraction model, Transport model, Online social network data

* Corresponding Author; E-mail: wirach.hi@ku.th

แบบจำลองการดึงดูดการเดินทางโดยใช้ข้อมูลเครือข่ายสังคมออนไลน์และการเรียนรู้เชิงลึก

วิรัช หิรัญ[1,*], ฐิตาภรณ์ พ่อบุตรดี[2]

[1,2]คณะวิทยาศาสตร์และวิศวกรรมศาสตร์

มหาวิทยาลัยเกษตรศาสตร์ วิทยาเขตเฉลิมพระเกียรติ จังหวัดสกลนคร

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาแบบจำลองการดึงดูดการเดินทางโดยใช้ข้อมูลเครือข่ายสังคมออนไลน์ โดยงานวิจัยนี้ใช้ข้อมูลการเดินทางเข้าสู่พื้นที่อุทยานแห่งชาติเป็นกรณีศึกษา ทำการพัฒนาแบบจำลองโดยใช้วิธีการเรียนรู้เชิงลึก และเปรียบเทียบกับการวิเคราะห์ด้วยวิธีวิเคราะห์ความถดถอยและวิธีโครงข่ายประสาทเทียม ผลการวิจัยพบว่า ข้อมูลจากการเช็คอินสามารถนำมาสร้างแบบจำลองการดึงดูด การเดินทางได้ เมื่อเปรียบเทียบแบบจำลองทั้งสามแบบพบว่า แบบจำลองที่ได้จากวิธีการเรียนรู้เชิงลึกให้ความถูกต้องในการพยากรณ์สูงที่สุด โดยให้ค่าเฉลี่ยเปอร์เซ็นต์ของความคลาดเคลื่อนสัมบูรณ์เท่ากับ 43.99 ทั้งนี้ค่าที่ได้ต่ำกว่าวิธีวิเคราะห์ความถดถอยเชิงเส้นและวิธีโครงข่ายประสาทเทียมซึ่งมีค่าเท่ากับ 386.48 และ 88.97 ตามลำดับ ทั้งนี้การพัฒนาแบบจำลองโดยใช้ข้อมูลเครือข่ายสังคมออนไลน์จะช่วยลดค่าใช้จ่ายในการสำรวจข้อมูลลงได้อย่างมาก

คำสำคัญ : แบบจำลองการดึงดูดการเดินทาง แบบจำลองการขนส่ง ข้อมูลเครือข่ายสังคมออนไลน์

* ผู้ประสานงานหลัก; อีเมล์: wirach.hi@ku.th

ความเป็นมาและความสำคัญของปัญหา

แบบจำลองการดึงดูดการเดินทางเป็นแบบจำลองหนึ่งในแบบจำลองต่อเนื่อง 4 ขั้นตอน ซึ่งใช้ในงานวางแผนด้านการจราจรและขนส่ง รวมทั้งสามารถนำมาประยุกต์ใช้ในการวางแผนด้านการขนส่งสาธารณะ หรือใช้สำหรับการวางแผนการจัดการในพื้นที่ อย่างไรก็ตาม การสำรวจข้อมูลเพื่อจัดทำแบบจำลองการดึงดูดการเดินทางต้องใช้ค่าใช้จ่ายและเวลาค่อนข้างมาก ดังนั้นจึงมักประสบปัญหาในการดำเนินการโดยเฉพาะพื้นที่ซึ่งมีแรงดึงดูดการเดินทางไม่มากนัก เช่น สถานที่ท่องเที่ยวขนาดเล็ก จึงเป็นสาเหตุให้ขาดข้อมูลในการวางแผนจัดการด้านการจราจรและขนส่งที่มีประสิทธิภาพ

ในปัจจุบันเครือข่ายสังคมออนไลน์และโทรศัพท์สมาร์ทโฟนได้เข้ามามีบทบาทในการใช้ชีวิตประจำวันมากขึ้น ในปี พ.ศ.2559 พบว่าการใช้สื่อสังคมออนไลน์ในประเทศไทยมีการใช้เฟสบุ๊ค 41 ล้านยูสเซอร์ ส่วนแอฟพลิเคชั่นอื่นๆ ได้แก่ ทวิตเตอร์มีจำนวนผู้ใช้ 5.3 ล้านยูสเซอร์ อินสตาแกรมมีจำนวนผู้ใช้ 7.8 ล้านยูสเซอร์ และไลน์มีจำนวนผู้ใช้ 33 ล้านยูสเซอร์ การใช้งานสื่อสังคมออนไลน์นี้ได้เข้ามาเกี่ยวข้องกับการเดินทางไปยังสถานที่ต่างๆ โดยคนที่เดินทางไปยังสถานที่ต่างๆนิยมการเช็คอินหรือโพสข้อความเพื่อแสดงถึงการมาถึงสถานที่นั้นๆ ซึ่งข้อมูลดังกล่าวมีส่วนเกี่ยวข้องกับการเดินทางเข้าสู่พื้นที่ และคาดว่าจะสามารถนำมาใช้เป็นตัวแปรเพื่อใช้ในการสร้างแบบจำลองการดึงดูดการเดินทางได้ แต่เนื่องจากข้อมูลที่เกี่ยวข้องกับการเดินทาง มักเป็นข้อมูลที่มีความหยาบและยากต่อการสร้างสมการเพื่อหาความสัมพันธ์ ดังนั้นจึงจำเป็นต้องใช้วิธีการในการสร้างแบบจำลองที่มีประสิทธิภาพ

การเรียนรู้เชิงลึก (Deep learning)เป็นสาขาหนึ่งของการเรียนรู้ของเครื่อง (Machine learning) ซึ่งเป็นวิธีการที่พัฒนามาจากวิธีโครงข่ายประสาทเทียมให้มีประสิทธิภาพมากยิ่งขึ้น และถูกประยุกต์ใช้ในงานด้านการวิเคราะห์ข้อมูลขนาดใหญ่ (Big data) รวมถึงยังมีความสามารถในการวิเคราะห์ความสัมพันธ์ของตัวแปรอิสระและตัวแปรตาม โดยการเรียนรู้เชิงลึกถูกใช้ในการวิเคราะห์ข้อมูลที่ซับซ้อนเพื่อหาความสัมพันธ์ อาทิเช่น การจำแนกภาพถ่าย

วัตถุประสงค์ของการวิจัย

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาแนวทางในการประยุกต์ใช้ข้อมูลเครือข่ายสังคมออนไลน์เพื่อใช้ในการสร้างแบบจำลองการดึงดูดการเดินทาง โดยการใช้ข้อมูลการเช็คอินในโปรแกรมเฟสบุ๊คของการเดินทางเข้าสู่สถานที่ท่องเที่ยวประเภทอุทยานแห่งชาติและใช้การเรียนรู้เชิงลึกในการประมวลผลข้อมูล โดยผลที่ได้จะเป็นแนวทางเพื่อนำประยุกต์ใช้ในพื้นที่อื่นๆ ต่อไป ทำให้สามารถสร้างแบบจำลองการดึงดูดการเดินทางเพื่อนำไปประยุกต์ใช้เป็นฐานข้อมูลการเดินทางเข้าสู่พื้นที่ต่างๆ โดยเฉพาะแหล่งท่องเที่ยวขนาดเล็กซึ่งจะเป็นประโยชน์ต่องานด้านการวางแผนการขนส่งและการจัดการสถานที่

เอกสารและงานวิจัยที่เกี่ยวข้อง

เนื่องจากในปัจจุบันการใช้งานเครือข่ายสังคมออนไลน์มีปริมาณเพิ่มขึ้นอย่างมาก ประกอบกับองค์ความรู้เรื่องการทำเหมืองข้อมูลมีการพัฒนาไปอย่างมากเช่นเดียวกัน ดังนั้นนักวิจัยด้านการขนส่งจึงได้พยายามที่จะพัฒนาวิธีการเพื่อจะใช้ประโยชน์จากการพัฒนาดังกล่าวในการรวบรวมข้อมูลด้านการขนส่งทดแทนวิธีการเดิมซึ่งต้องใช้ทรัพยากรสูง Hasan et al. (2013) ใช้ข้อมูลจากทวิตเตอร์ร่วมกับข้อมูลการเช็คอินจากแอพพลิเคชั่นอื่นเช่น Foursquare เพื่อศึกษารูปแบบการเดินทางในเมืองนิวยอร์ค ชิคาโก และลอสแองเจรีส โดยแบ่งรูปแบบกิจกรรมออกเป็น 7 กิจกรรม ประกอบด้วย เดินทางกลับบ้าน ไปทำงาน รับประทานอาหาร บันเทิง พักผ่อนหย่อนใจ และซื้อ

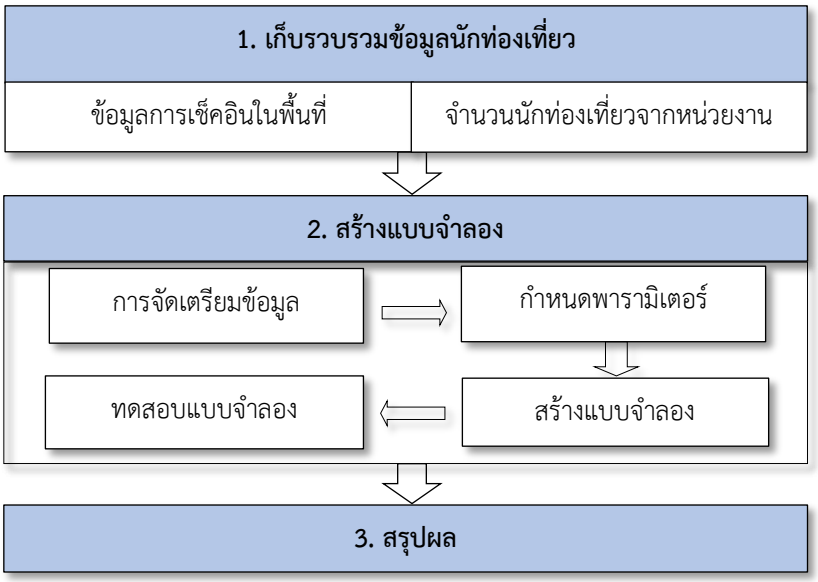
ของ Yang et al. (2014) ใช้ข้อมูลจากแอพพลิเคชั่น Foursquare เพื่อสร้างแบบจำลองการดึงดูดการเดินทางในเมืองชิคาโก Hu & Jin (2015) ใช้ข้อมูลจากแอพพลิเคชั่น Foursquare เพื่อสร้างแบบจำลองการดึงดูดการเดินทางในเมืองออสติน รัฐเท็กซัส โดยการใช้ข้อมูลการเช็คอินในสถานที่ต่างๆ จำนวน 124,611 ครั้ง แบบจำลองถูกพัฒนาขึ้นในแต่ละช่วงเวลาของวันและทำการเปรียบเทียบข้อมูลที่ได้จากแบบจำลองและข้อมูลการสำรวจจากภาคสนาม Maghrebi et al. (2015) ใช้ข้อมูลจากทวิตเตอร์ในการศึกษาข้อมูลการเดินทางในพื้นที่เมืองซิดนีย์ โดยการใช้การแปลผลจากข้อความเพื่อระบุถึงประเภทของกิจกรรมของผู้เดินทางในจุดต่างๆ ของเมืองในแต่ละช่วงเวลา ซึ่งสามารถนำไปใช้การศึกษาการเดินทางได้

กรอบแนวคิดในการวิจัย

จากวัตถุประสงค์ของงานวิจัยซึ่งต้องการพัฒนาแนวทางในการประยุกต์ใช้ข้อมูลเครือข่ายสังคมออนไลน์สำหรับใช้สร้างแบบจำลองการดึงดูดการเดินทางและลักษณะของข้อมูลการเดินทางซึ่งมักมีการกระจายตัวค่อนข้างมาก ดังนั้นจึงจำเป็นต้องใช้เครื่องมือในการประมวลผลและวิเคราะห์ข้อมูลที่มีประสิทธิภาพ แต่เดิมวิธีการวิเคราะห์ความถดถอยเป็นวิธีการที่นิยมใช้ในการอธิบายความสัมพันธ์ระหว่างตัวแปรที่ต้องการศึกษา ซึ่งสามารถแสดงความสัมพันธ์ในรูปแบบของสมการทางคณิตศาสตร์ อย่างไรก็ตาม ในบางกรณี วิธีดังกล่าวให้ความแม่นยำในการพยากรณ์ที่ค่อนข้างต่ำ รวมทั้งปัญหาด้านข้อมูลที่ไม่สอดคล้องกับเงื่อนไขของการวิเคราะห์ความถดถอย ในปัจจุบันการเรียนรู้ของเครื่อง (Machine learning) อาทิ วิธีโครงข่ายประสาทเทียมและการเรียนรู้เชิงลึก มีการพัฒนาไปอย่างมาก รวมทั้งได้รับการยอมรับว่ามีประสิทธิภาพสูงในการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรที่ต้องการศึกษา อย่างไรก็ตามวิธีการโครงข่ายประสาทเทียมและ การเรียนรู้เชิงลึกไม่สามารถแสดงความสัมพันธ์ทางคณิตศาสตร์ระหว่างตัวแปรที่ต้องการศึกษาได้ ดังนั้นงานวิจัยนี้จึงทำการเปรียบเทียบแบบจำลองที่สร้างด้วยวิธีโครงข่ายประสาทเทียมและการเรียนรู้เชิงลึกกับแบบจำลองที่สร้างจากวิธีวิเคราะห์ความถดถอย เพื่อให้บรรลุวัตถุประสงค์ดังกล่าว

วิธีดำเนินการวิจัย

งานวิจัยนี้จึงแบ่งขั้นตอนการวิจัยออกเป็น 3 ส่วน ดังแสดงในภาพที่ 1



ภาพที่ 1 ขั้นตอนการวิจัย

ส่วนที่ 1 ส่วนของการเก็บรวบรวมข้อมูล ประกอบด้วย การเก็บข้อมูลการเช็คอินของผู้ที่เช็คอินในพื้นที่ท่องเที่ยวในความรับผิดชอบของ สำนักอุทยานแห่งชาติ กรมอุทยานแห่งชาติ สัตว์ป่าและพันธุ์พืช จำนวน 154 แห่ง โดยทำการเก็บข้อมูลทุก 3 วัน ระหว่างเดือนพฤศจิกายน 2560 – มกราคม 2561 และทำการรวบรวมข้อมูลข้อมูลนักท่องเที่ยวที่เดินทางไปท่องเที่ยวยังแหล่งท่องเที่ยวจากระบบฐานข้อมูลเพื่อการตัดสินใจในการบริหารจัดการอุทยานแห่งชาติ

ส่วนที่ 2 การสร้างแบบจำลองโดยงานวิจัยนี้เลือกใช้วิธีในการสร้างแบบจำลอง 3 วิธี ได้แก่ วิธีการวิเคราะห์ความถดถอย วิธีโครงข่ายประสาทเทียม และวิธีการเรียนรู้เชิงลึก โดยมีขั้นตอนการดำเนินงานดังต่อไปนี้

ขั้นตอนที่ 1 เตรียมข้อมูลที่ได้จากการเช็คอิน โดยแบ่งข้อมูลออกเป็น 2 ส่วน ในสัดส่วน 70: 30 ส่วนแรกใช้สำหรับสร้างแบบจำลอง และส่วนที่สองใช้สำหรับทดสอบแบบจำลอง

ขั้นตอนที่ 2 กำหนดรูปแบบของแบบจำลอง และค่าพารามิเตอร์ของแบบจำลอง

ขั้นตอนที่ 3 สร้างแบบจำลองจากข้อมูลชุดที่ 1

ขั้นตอนที่ 4 ทดสอบแบบจำลองโดยใช้ข้อมูลชุดที่ 2 โดยการประยุกต์ใช้ค่าเฉลี่ยเปอร์เซ็นต์ของความคลาดเคลื่อนสัมบูรณ์ (Mean absolute percent error: MAPE) เป็นตัววัดประสิทธิภาพของแบบจำลอง ดังแสดงดังต่อไปนี้

$$MAPE = \frac{\sum_{i=1}^{n} \left| \frac{e_i}{z_i} \right|}{N} x100 \qquad (1)$$

โดยที่

e_i = ความแตกต่างของค่าที่ได้จากแบบจำลองและค่าจริง

Z_i = ค่าจริง

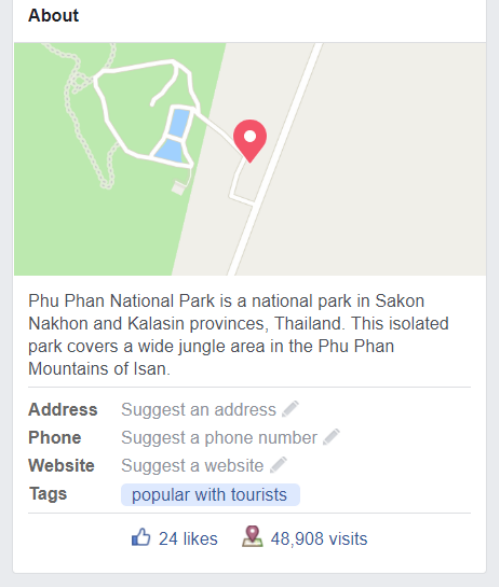
N = จำนวนข้อมูลทั้งหมดที่นำมาเปรียบเทียบ

ส่วนที่ 3 เป็นการสรุปผลการวิจัยเพื่อพิจารณาถึงความเหมาะสมในการนำข้อมูลจากการเช็คอินมาใช้ในการสร้างแบบจำลองการดึงดูดการเดินทาง

ผลการวิจัย

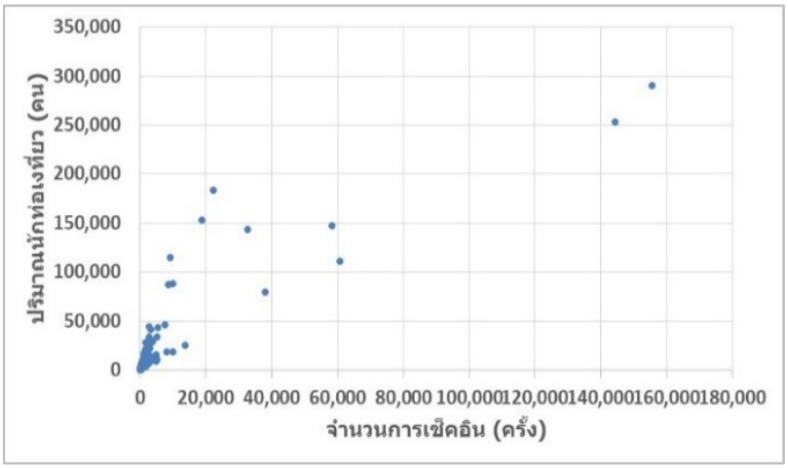
ผลการเก็บรวบรวมข้อมูลจำนวนการเช็คอินของผู้ใช้งานเฟสบุ๊คในสถานที่ท่องเที่ยวในความดูแลของสำนักอุทยานแห่งชาติ กรมอุทยานแห่งชาติ สัตว์ป่าและพันธุ์พืช โดยการบันทึกข้อมูลจำนวนคนที่เช็คอินในเพจดังกล่าว ดังแสดงในภาพที่ 2 พบว่าการสร้างเพจของสถานที่มีความหลากหลาย ตัวอย่างเช่น อุทยานแห่งชาติภูพาน มีการสร้างเพจชื่อ “อุทยานแห่งชาติภูพาน อ.เมือง จ.สกลนคร” และชื่อ “Phu Phan National Park” ซึ่งสองเพจนี้มีความแตกต่างกัน ซึ่งผู้ที่เดินทางไปยังอุทยานแห่งชาติภูพานอาจทำการเช็คอินในเพจใดเพจหนึ่งจากทั้งสองเพจ นอกจากนี้ในสถานที่ท่องเที่ยวเดียวกันอาจมีแหล่งท่องเที่ยวอื่นอีก เช่น ในพื้นที่อุทยานแห่งชาติภูพานมีสถานที่ท่องเที่ยวอื่นได้แก่ น้ำตกคำหอม และพระตำหนักภูพานราชนิเวศน์ ซึ่งในกรณีเช่นนี้ นักท่องเที่ยวซึ่งเดินทางไปยังอุทยานแห่งชาติภูพานและเดินทางไปยังน้ำตกคำหอมอาจทำการเช็คอินที่น้ำตกคำหอม โดยไม่เช็คอินที่อุทยานแห่งชาติภูพาน ดังนั้น

ในการบันทึกข้อมูลจำนวนการเช็คอินของสถานที่ท่องเที่ยวจึงทำการบันทึกข้อมูลของเพจที่ระบุถึงสถานที่ท่องเที่ยวนั้นๆ และสถานที่ท่องเที่ยวอื่นๆ ซึ่งตั้งอยู่ในพื้นที่อุทยานเดียวกัน



ภาพที่ 2 รายละเอียดข้อมูลของจำนวนเช็คอินในสถานที่

เมื่อเปรียบเทียบข้อมูลการเช็คอินและจำนวนนักท่องเที่ยวที่เดินทางไปยังสถานที่ท่องเที่ยว ดังแสดงในภาพที่ 3 โดยพบว่าค่าสัมประสิทธิ์สหสัมพันธ์ของข้อมูลมีค่าเท่ากับ 0.875531 ซึ่งแสดงให้เห็นว่าข้อมูลมีความสัมพันธ์กันค่อนข้างสูง



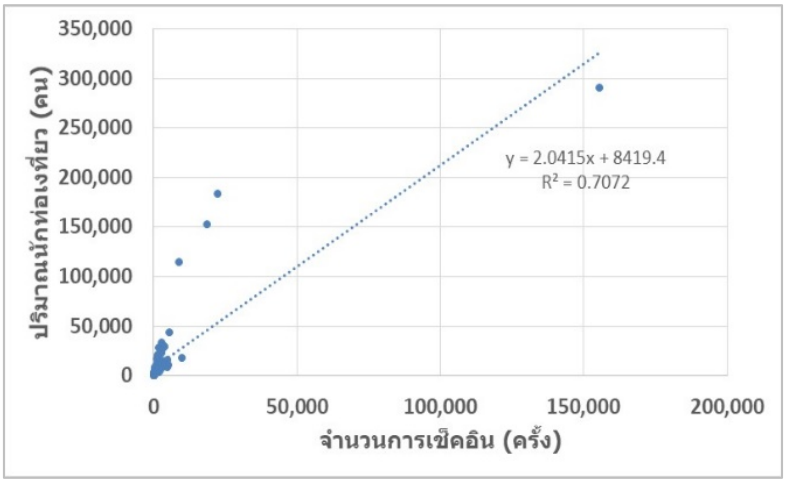
ภาพที่ 3 ความสัมพันธ์ระหว่างจำนวนนักท่องเที่ยวและจำนวนการเช็คอิน

งานวิจัยนี้ใช้การวิเคราะห์ความถดถอยเชิงเส้นอย่างง่ายโดยตัวแปรตาม ได้แก่ ปริมาณนักท่องเที่ยวในช่วงเวลา 1 เดือน และตัวแปรต้นคือจำนวนการเช็คอินในช่วงเวลา 1 เดือน ผลการสร้างแบบจำลองด้วยวิธีวิเคราะห์ความถดถอยได้รูปแบบความสัมพันธ์ดังต่อไปนี้

$$TNUM = 8419.364 + 2.041xCHK$$
$$(3.208) \qquad (13.812)$$

โดยที่ TNUM หมายถึง จำนวนนักท่องเที่ยว
CHK หมายถึง จำนวนการเช็คอิน

โดยแบบจำลองที่สร้างขึ้นมีค่า R^2 เท่ากับ 0.707 ซึ่งแสดงว่าแบบจำลองสามารถจำลองความสัมพันธ์ของตัวแปรอยู่ในเกณฑ์ค่อนข้างดี สำหรับค่าสถิติ t ของตัวแปรจำนวนการเช็คอินและค่าคงที่ของสมการ มีค่าเท่ากับ 13.812 และ 3.208 ตามลำดับ ซึ่งแสดงว่าค่าสัมประสิทธิ์ของจำนวนการเช็คอินและค่าคงที่ของสมการมีค่ามากกว่า 0 อย่างมีนัยสำคัญ ดังแสดงในภาพที่ 4



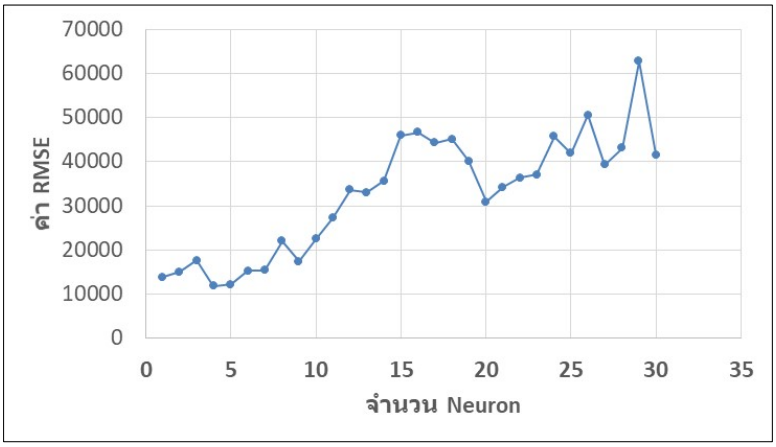
ภาพที่ 4 การหาความสัมพันธ์โดยวิธีวิเคราะห์ความถดถอยเชิงเส้น

โครงข่ายประสาทเทียมเป็นแบบจำลองทางคณิตศาสตร์ที่จำลองการทำงานของโครงข่ายประสาทของมนุษย์ เป็นวิธีการที่สามารถหาความสัมพันธ์ของข้อมูลป้อนเข้าและข้อมูลส่งออกได้ค่อนข้างแม่นยำ โดยไม่ต้องอาศัยความสัมพันธ์ทางคณิตศาสตร์ (Lomlek & Yongcharoen, 2014) นอกจากนี้ยังมีข้อดีคือ ไม่มีข้อตกลงเบื้องต้นเหมือนวิธีการทางสถิติทั่วไป ไม่สนใจว่าข้อมูลจะมีการแจกแจงแบบใด และสามารถวิเคราะห์ได้เมื่อมีค่านอกกลุ่ม (Sangarun & Sangarun, 2015) โครงข่ายประสาทเทียมทำการจำลองการทำงานของเซลล์ประสาทของมนุษย์ด้วยการกำหนดให้นิวรอนซึ่งเป็นหน่วยประมวลผลข้อมูลสามารถรับข้อมูลป้อนเข้า (input) และส่งออกข้อมูลส่งออก (output) รวมทั้งเชื่อมโยงกับนิวรอนตัวอื่นเป็นโครงข่าย ซึ่งนิวรอนแต่ละตัวสามารถกำหนดค่าน้ำหนัก (weight) ของข้อมูลป้อนเข้าและมีค่าจุดกระตุ้น (threshold) เพื่อใช้กำหนดน้ำหนักรวมของข้อมูลป้อนเข้าสำหรับใช้ในการส่งค่าข้อมูลส่งออกไปเป็นข้อมูลป้อนเข้าของนิวรอนตัวอื่นในโครงข่ายได้ ซึ่งค่าน้ำหนัก (weight) นี้เปรียบเสมือนความจำระยะยาวของโครงข่ายประสาทเทียมและเป็นส่วนสำคัญเพื่อให้คอมพิวเตอร์รู้จำสิ่งที่กำหนดให้คอมพิวเตอร์เรียนรู้

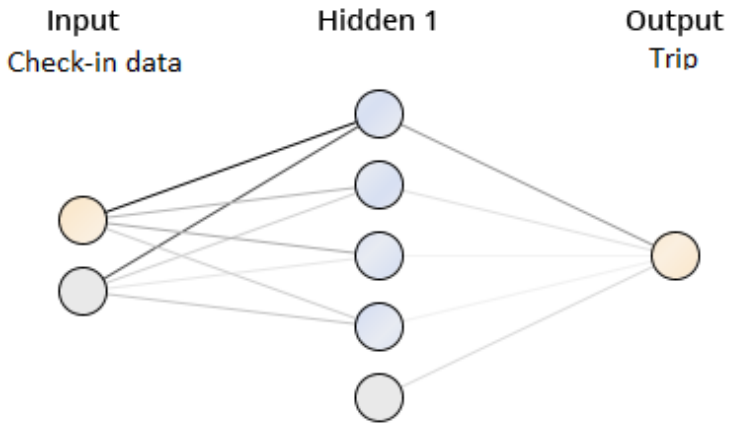
โครงข่ายประสาทเทียม มีการพัฒนาขึ้นหลายรูปแบบแต่แบบที่ได้รับการยอมรับอย่างแพร่หลายคือโครงข่ายประสาทเทียมแบบแพร่ย้อนกลับ (Back Propagation Neural Network) ซึ่งประกอบด้วยชั้นของข้อมูลเบื้องต้น 3 ชั้นเรียงกันเป็นชั้นๆ ได้แก่ ชั้นอินพุต (input layer) ชั้นซ่อน (hidden layer) และชั้นเอาท์พุต (output layer) การเรียนรู้ของโครงข่ายประสาทเทียมแบบแพร่ย้อนกลับประกอบด้วย 3 ขั้นตอน คือ 1) ขั้นตอนการคำนวณไป

ข้างหน้า (feed forward) เมื่อโครงข่ายได้รับข้อมูลป้อนเข้าสำหรับสอนให้แก่โครงข่ายในแต่ละครั้งแล้ว โครงข่ายจะคำนวณค่าจากชั้นอินพุตไปยังชั้นซ่อนและไปสู่ชั้นเอาท์พุต 2) ขั้นตอนการคำนวณและการแทนค่ากลับของผลรวมของความผิดพลาดของผลลัพธ์กับค่าเป้าหมาย โดยการนำค่าข้อมูลส่งออก และนำไปเปรียบเทียบกับค่าเป้าหมาย จากนั้นจึงทำการคำนวณหาค่าความผิดพลาดซึ่งค่าความผิดพลาดนี้จะถูกส่งกลับเข้าสู่โครงข่ายโดยใช้อัลกอริทึมแบบแพร่ย้อนกลับเพื่อใช้ในการปรับปรุงค่าน้ำหนักของโครงข่าย 3) ขั้นตอนการปรับค่าน้ำหนัก (weight) ไปเรื่อยๆ ขั้นตอนนี้จะขึ้นอยู่กับขั้นตอนการคำนวณไปข้างหน้าเพียงขั้นตอนเดียว โดยจะนำค่าน้ำหนักที่ได้จากการเรียนรู้ไปใช้ในการพยากรณ์ต่อไป (Buket & Seresangtakul, 2013)

งานวิจัยนี้เลือกใช้แบบจำลองโครงข่ายประสาทเทียมแบบแพร่ย้อนกลับ โดยกำหนดจำนวนรอบเท่ากับ 700 ค่าอัตราการเรียนรู้ในชั้นซ่อนและชั้นแสดงผลเท่ากับ 0.43 และค่าโมเมนตัมเท่ากับ 0.2 ส่วนจำนวนนิวรอนในชั้นซ่อนซึ่งเป็นส่วนสำคัญต่อประสิทธิภาพของแบบจำลองได้จากการทดลองเปลี่ยนจำนวนนิวรอนระหว่าง 1-30 และทำการเลือกค่านิวรอนที่เหมาะสมจากค่ารากที่สองของค่าเฉลี่ยความผิดพลาดกำลังสอง (RMSE) ดังแสดงในภาพที่ 5 โดยพบว่าจำนวนนิวรอนที่เหมาะสมมีค่าเท่ากับ 4 จากนั้นจึงทำการสร้างแบบจำลองโครงข่ายประสาทเทียม ดังแสดงในภาพที่ 6



ภาพที่ 5 การหาค่าจำนวนนิวรอนที่เหมาะสมสำหรับแบบจำลองโครงข่ายประสาทเทียม



ภาพที่ 6 โครงข่ายประสาทเทียมของแบบจำลอง

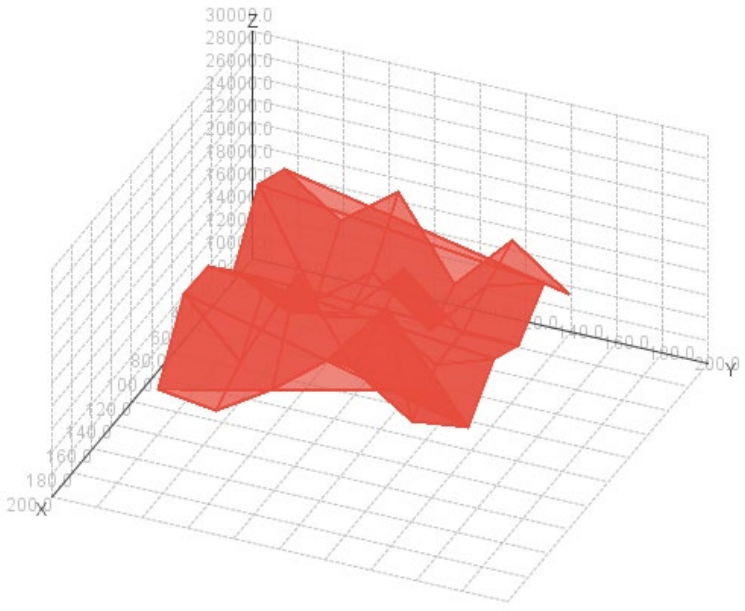
โครงข่ายประสาทเทียมที่สร้างขึ้นมีค่ารากที่สองของค่าเฉลี่ยความผิดพลาดกำลังสองเท่ากับ 16819.358 ส่วนค่าน้ำหนักและค่าจุดกระตุ้นของแต่ละนิวรอน ดังแสดงในตารางที่ 1

ตารางที่ 1 รายละเอียดของแบบจำลองโครงข่ายประสาทเทียม

โหนด	ค่าน้ำหนัก	ค่าจุดกระตุ้น
1	8.052	5.758
2	2.273	-1.464
3	2.549	-0.758
4	1.514	-1.770

การเรียนรู้เชิงลึกเป็นการพัฒนาขึ้นจากโครงข่ายประสาทเทียม โดยการกำหนดจำนวนชั้นซ่อนให้มีได้มากกว่า 1 ชั้น และใช้ Rectified Linear Unit (ReLU) เป็นฟังก์ชั่นกระตุ้น การพัฒนาดังกล่าวทำให้การเรียนรู้เชิงลึกสามารถวิเคราะห์ปัญหาที่ซับซ้อนได้มีประสิทธิภาพมากยิ่งขึ้น เช่น การวิเคราะห์ภาพถ่ายหรือ การสั่งการด้วยเสียง สำหรับการศึกษาด้านการขนส่ง ได้มีนำการเรียนรู้เชิงลึกมาประยุกต์ใช้ในด้านต่างๆ (Ke et al. (2017) ทำการศึกษาความต้องการการเดินทางในช่วงเวลาสั้นของบริการ on-demand ride service ในประเทศจีน โดยใช้การเรียนรู้เชิงลึกชนิด fusion convolutional long short-term memory network (FCL-Net) ผลการวิจัยพบว่าวิธีการเรียนรู้เชิงลึกให้ค่ารากที่สองของค่าเฉลี่ยความผิดพลาดกำลังสอง (RMSE) น้อยกว่าวิธีโครงข่ายประสาทเทียมร้อยละ 48.3 นอกจากนี้ Liu & Chen (2017) ได้ทำการศึกษาเพื่อคาดการณ์จำนวนผู้โดยสารที่ใช้บริการรถโดยสารด่วนพิเศษ (BRT) ในเมืองเซี่ยเหมิน โดยพบว่าการประยุกต์ใช้วิธีการเรียนรู้เชิงลึกประเภท Hybrid model (SAE-DNN) ให้ผลการคาดการณ์จำนวนผู้โดยสารได้แม่นยำกว่าวิธีการอื่นๆ

งานวิจัยนี้ประยุกต์ใช้การเรียนรู้เชิงลึกโดยการกำหนดค่า epochs เท่ากับ 7 epsilon เท่ากับ $1x\ 10^{-8}$ และ rho เท่ากับ 0.99 ส่วนการดำเนินการเพื่อหาโครงสร้างที่เหมาะสมยังไม่มีวิธีการดำเนินการที่แน่นอน ดังนั้นงานวิจัยนี้จึงทดลองทำการเปลี่ยนแปลงค่า hidden layer sizes ระหว่าง 25 และ 150 ดังแสดงในภาพที่ 7



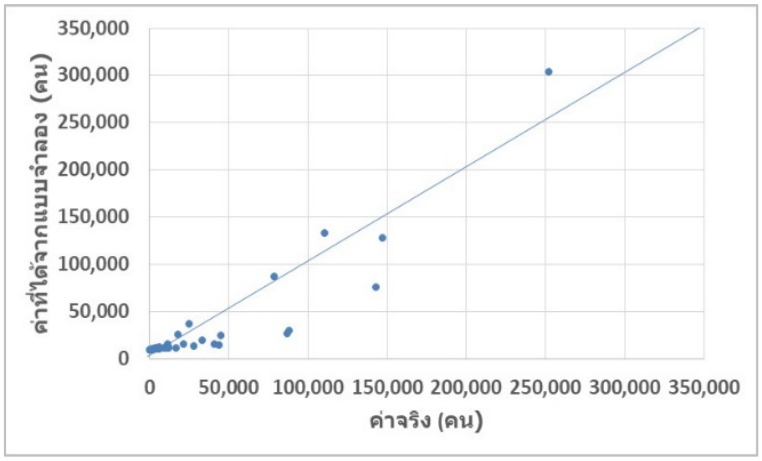
ภาพที่ 7 การหาค่าจำนวนนิวรอนที่เหมาะสมกับสำหรับแบบจำลองการเรียนรู้เชิงลึก

โดยพบว่าแบบจำลองที่มีโครงสร้างของชั้นซ่อนจำนวน 2 ชั้น มีค่า hidden layer sizes เท่ากับ 50 และ 100 ให้ค่ารากที่สองของค่าเฉลี่ยความผิดพลาดกำลังสองมีค่าน้อยสุด โดยมีค่าเท่ากับ 13443.4 และแบบจำลองที่พัฒนาขึ้นมีค่า R^2 เท่ากับ 0.825

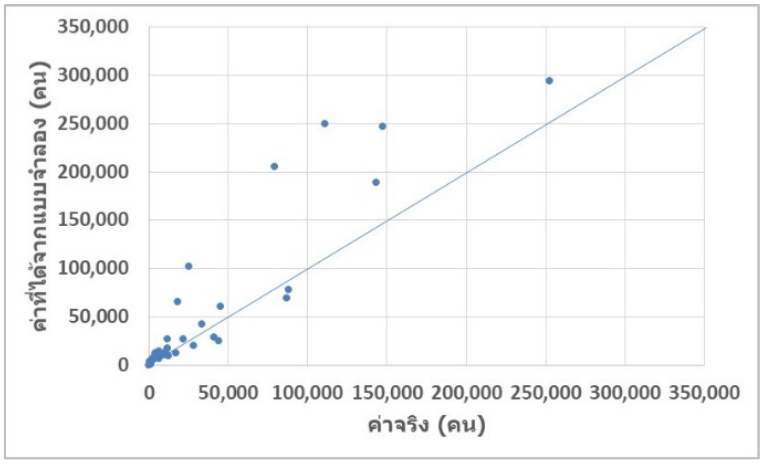
อภิปรายผล

งานวิจัยนี้ทำการพัฒนาแบบจำลองการดึงดูดการเดินทางของอุทยานแห่งชาติในประเทศไทย โดยใช้วิธีการเรียนรู้เชิงลึก และเปรียบเทียบกับวิธีความถดถอยเชิงเส้น ซึ่งเป็นวิธีการวิเคราะห์ทางสถิติแบบดั่งเดิมและวิธีโครงข่ายประสาทเทียม ซึ่งมีการประยุกต์ใช้ในการวิเคราะห์ข้อมูลที่ซับซ้อนที่ได้รับการยอมรับมากขึ้นในปัจจุบัน ผลการวัดประสิทธิภาพของแบบจำลอง ดังแสดงในภาพที่ 8-10 และการใช้ค่าเฉลี่ยเปอร์เซ็นต์ของความคลาดเคลื่อนสัมบูรณ์ดังแสดงในตารางที่ 2 พบว่าแบบจำลองที่พัฒนาขึ้นด้วยวิธีการเรียนรู้เชิงลึกให้ประสิทธิภาพในการพยากรณ์สูงที่สุด อย่างไรก็ตามประสิทธิภาพที่ได้จากแบบจำลองที่พัฒนาขึ้นอาจยังไม่สูงมากนัก ทั้งนี้เมื่อเปรียบเทียบกับงานวิจัยวิจัยที่ผ่านมาพบว่างานวิจัยของ Ke et al. (2017) ได้ค่าเฉลี่ยเปอร์เซ็นต์ของความคลาดเคลื่อนสัมบูรณ์จากวิธีโครงข่ายประสาทเทียม และวิธีการเรียนรู้เชิงลึกเท่ากับ 27.95 และ 21.79 ตามลำดับ โดยงานวิจัยดังกล่าวใช้ข้อมูลในการวิเคราะห์จำนวน 1,000,000 ข้อมูล ส่วนงานวิจัยของ Liu & Chen (2017) พบว่าแบบจำลองที่พัฒนาขึ้นให้ค่าเฉลี่ยเปอร์เซ็นต์ของความคลาดเคลื่อนสัมบูรณ์ จากวิธีการเรียนรู้เชิงลึกระหว่าง 10.02 – 22.27 โดยงานวิจัยดังกล่าวใช้ข้อมูลการเดินทางเข้า-ออกของผู้โดยสาร จำนวน 11,996,975 ข้อมูลในการวิเคราะห์

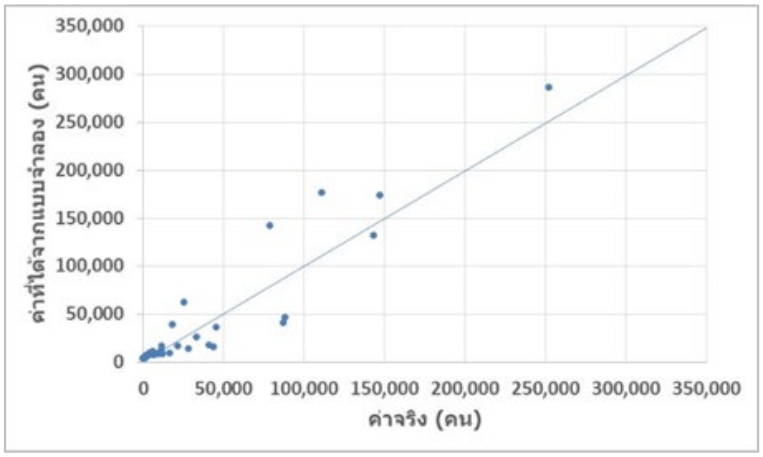
เมื่อพิจารณาจำนวนข้อมูลของงานวิจัยทั้งสองดังกล่าวและงานวิจัยนี้ พบว่า มีความแตกต่างกันค่อนข้างมาก ดังนั้นหากทำการรวบรวมข้อมูลให้มีปริมาณมากขึ้นจะเป็นแนวทางหนึ่งที่จะเพิ่มประสิทธิภาพของแบบจำลองได้ นอกจากนี้ เนื่องจากแบบจำลองนี้พัฒนาขึ้นจากข้อมูลการเดินทางไปยังอุทยานแห่งชาติ ดังนั้นการนำไปใช้ในสถานที่ดึงดูดการเดินทางประเภทอื่นซึ่งคุณลักษณะที่แตกต่างออกไป จำเป็นต้องมีการทดสอบและปรับเทียบแบบจำลองก่อนการนำไปใช้งาน



ภาพที่ 8 การเปรียบเทียบค่าที่ได้จากแบบจำลองความถดถอยเชิงเส้นและค่าจริง



ภาพที่ 9 การเปรียบเทียบค่าที่ได้จากแบบจำลองโครงข่ายประสาทเทียมและค่าจริง



ภาพที่ 10 การเปรียบเทียบค่าที่ได้จากแบบจำลองการเรียนรู้เชิงลึกและค่าจริง

ตารางที่ 2 ประสิทธิภาพของแบบจำลองที่พัฒนาขึ้น

แบบจำลอง	ค่าเฉลี่ยเปอร์เซ็นต์ของความคลาดเคลื่อนสัมบูรณ์
ความถดถอยเชิงเส้น	386.48
โครงข่ายประสาทเทียม	88.97
การเรียนรู้เชิงลึก	43.99

การพัฒนาแบบจำลองการดึงดูดการเดินทางที่ได้จากงานวิจัยนี้เป็นต้นแบบในการพัฒนาวิธีการสร้างแบบจำลองการดึงดูดการเดินทางด้วยวิธีการใหม่ ซึ่งจะช่วยลดค่าใช้จ่ายในการสำรวจข้อมูลและสามารถปรับข้อมูลให้ทันสมัยได้สะดวกมากขึ้น การดำเนินการดังกล่าวจะเป็นประโยชน์ในการวิเคราะห์ทางด้านการขนส่งหรือการวิเคราะห์การเข้าถึงพื้นที่ต่างๆ ให้มีประสิทธิภาพและประหยัดค่าใช้จ่ายมากขึ้น

ข้อเสนอแนะ

1.ข้อเสนอแนะในการนำผลวิจัยไปใช้

(1) งานวิจัยนี้ดำเนินการเฉพาะพื้นที่อุทยานแห่งชาติ ดังนั้นหากต้องการใช้ข้อมูลเพื่อการพัฒนาแบบจำลองของแหล่งท่องเที่ยวหรือสถานที่อื่นควรทำการศึกษาเพิ่มเติม

(2) เนื่องจากประสิทธิภาพของแบบจำลองที่ได้จากงานวิจัยนี้อาจไม่สูงมากนัก เนื่องจากจำนวนข้อมูลที่ใช้ในการวิเคราะห์ ดังนั้นการนำไปใช้ต้องพิจารณาถึงความคาดเคลื่อนที่อาจเกิดขึ้น

2.ข้อเสนอแนะในการวิจัยครั้งต่อไป

(1) ควรใช้จำนวนข้อมูลให้มีปริมาณมากขึ้นซึ่งจะส่งผลต่อประสิทธิภาพของแบบจำลอง

เอกสารอ้างอิง (References)

Buket , A. and Seresangtakul, P. (2013). Sugarcane Production Forecasting Model of the Northeastern by Artificial Neural Network. *KKU Science Journal*, 41(1), 213-225. (in Thai)

Hasan, S., Zhan, X. and Ukkusuri., S.V. (2013). Understanding Urban Human Activity and Mobility Patterns Using Large-Scale Location-based Data from Online Social Media. *2nd ACM SIGKDD international workshop on urban computing*, 11 August 2013 at Chicago, Illinois, USA, 1-8.

Hu, W.N. and Jin, P.J. (2015). Dynamic Trip Attraction Estimation with Location Based Social Network Data Balancing between Time of Day Variations and Zonal Differences. *International Workshop on Spatiotemporal Computing*, 13–15 July 2015 at Fairfax, Virginia, USA, 193-198.

Ke, J., Zheng H., Yang, H., and Chen, X. (2017). Short-Term Forecasting of Passenger Demand under On-Demand Ride Services: A Spatio-Temporal Deep Learning Approach. *Transportation Research Part C*, 85, 591-608.

Liu, L. and Chen, R-C. (2017). A Novel Passenger Flow Prediction Model Using Deep Learning Methods. *Transportation Research Part C : Emerging Technologies*, 84, 74-91.

Lomlek, P. and Yongcharoen, W. (2014). Application of Neural Network for the Performance Prediction of Absorption Chiller. *Journal of Energy Research*, 11(2), 67-78. (in Thai)

Maghrebi, M., Abbasi, A., Rashidi, T.H., and Waller, S.T. (2015). Complementing Travel Diary Surveys with Twitter Data: Application of Text Mining Techniques on Activity Location, Type and Time. *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, 15-18 September 2015 at Las Palmas, Gran Canaria, Spain, 208-213.

Sangarun, T. and Sangarun, A. (2015). Classification Techniques of Affecting Factor on Resigning of Hotel's Employee by Artificial Neural Network. *Southeast Bangkok Journal*, 1(2), 96-112. (in Thai)

Yang, F., Jin, P.J., Cheng, Y., and Ran, B. (2015). Origin-Destination Estimation for Non Commuting Trips Using Location-Based Social Networking Data. *International Journal of Sustainable Transportation*, 9(8), 551-564.