โครงงานนิสิตชั้นปีที่ 4

การทำนายความต้องการของจุดให้บริการรถเช่าใหม่จากข้อมูลของจุดให้บริการที่มีอยู่

Demand prediction for new car sharing station from existing station data

เสนอโดย

นายพีรณัฏฐ ประเสริฐพันธุ์ รหัสประจำตัว 613 03744 21

คณะกรรมการสอบโครงร่างโครงงานนิสิตชั้นปีที่ 4

ผศ.ดร.	พิศิษฎ์	จารุมณีโรจน์	อาจารย์ที่ปรึกษา
ผศ.ดร.	นันทชัย	กานตานันทะ	อาจารย์กรรมการ
ดร.	สรวิศ	วิฑูรทัศน์	กรรมการนอกมหาวิทยาลัย

ภาควิชาวิศวกรรมอุตสาหการ คณะวิศวกรรมศาสตร์

จุฬาลงกรณ์มหาวิทยาลัย ปีการศึกษา 2564

ลายเซ็นอาจารย์ที่ปรึกษา_____

บทคัดย่อ

ในปัจจุบันธุรกิจรถเช่าผ่านจุดให้บริการในตัวเมืองมีความแปลกใหม่ และเป็นที่น่าจับตามองของผู้ใช้บริการทั่ว ๆ ไปที่สามารถขับรถได้ และพักอาศัยอยู่ในตัวเมือง แต่ไม่มีรถยนต์ส่วนตัวเนื่องด้วยเหตุผลหลายประการ เช่น นิสิต/นักศึกษาจบใหม่ที่ยังไม่มีกำลังพอในการซื้อรถยนต์ส่วนตัว พื้นที่จอดรถส่วนตัวที่ว่างเหลืออยู่น้อย หรือมีความจำเป็นที่ต้องไป-กลับที่พัก แต่เนื่องจากรถยนต์ส่วนตัวไม่พร้อมใช้งาน เป็นต้น ธุรกิจดังกล่าวได้มองเห็นปัญหา และต้องการแก้ปัญหาดังกล่าวด้วยการเปิดจุดให้บริการเพื่อให้ผู้ใช้บริการเช่ารถในระยะเวลาที่ตกลงผ่านแอปพลิเคชันของบริษัทด้วยการจ่ายค่าบริการต่อครั้งการเช่า แต่เนื่องจากจำนวนจุดให้บริการที่กระจายอยู่ในเมืองนั้นยังไม่เพียงพอต่อการเข้าถึงของผู้ใช้งานได้อย่างทั่วถึง บริษัทจึงจำเป็นต้องพิจารณาที่เปิดจุดให้บริการใหม่อย่างถี่ถ้วนเพื่อความสามารถในการเข้าถึงกลุ่มลูกค้าที่มากขึ้น ทางผู้จัดทำจึงเลือกจัดทำโครงงานการทำนายความต้องการของจุดให้บริการรถเช่าใหม่จากข้อมูลของจุดให้บริการเดิม โดยตัวแบบทางคณิตศาสตร์ ด้วยการเปรียบเทียบจากค่าเฉลี่ยความผิดพลาดกำลังสอง (Root mean squared error : RMSE) และค่าเฉลี่ยความผิดพลาดสัมบูรณ์ (Mean squared error : MAE) ระหว่างตัวแบบความจำระยะสั้น ระยะยาว (Long-short term memory : LSTM) หรือหน่วยเวียนกลับแบบมีประตู (Gated recurrent unit : GRU) ประสานเข้ากับโครงข่ายประสาทการแปลงกราฟ (Graph convolutional neural network) พร้อมทั้งเปรียบเทียบวิธีการปรับปรุงตัวแบบด้วยการไล่ความชันแบบสุ่มด้วยโมเมนตัม (Stochastic Gradient Descent with momentum) การถ่ายถอดรากกำลังสองเฉลี่ย (Root Mean Square Propagation) และ การประมาณค่าช่วงเวลาปรับตัว (Adaptive Moment Estimation : Adam) เพื่อสร้างตัวแบบทำนายความต้องการของจุดให้บริการใหม่จากข้อมูลของจุดให้บริการที่มีอยู่ เพื่อพิจารณาผลกระทบ และเป็นส่วนช่วยในการตัดสินใจถึงประสิทธิภาพการใช้งานของการเปิดจุดให้บริการใหม่ต่อจุดให้บริการที่มีอยู่ โดยที่เมื่อทำการทดลองแล้วพบว่า ตัวแบบความจำระยะสั้น และระยะยาว (LSTM) และเสริมด้วยโครงข่ายประสาทการแปลงกราฟ (Graph convolutional neural network) ด้วยวิธีการปรับปรุง การประมาณค่าช่วงเวลาปรับตัว (Adaptive Moment Estimation) มีประสิทธิภาพในการทำนายสูงที่สุด เพราะความผิดพลาดในการทำนายน้อยที่สุดด้วยค่าเฉลี่ยความผิดพลาดกำลังสองเท่ากับ 4.3463 และค่าเฉลี่ยความผิดพลาดสัมบูรณ์เท่ากับ 2.5333 และการสร้างตัวแบบด้วยกระบวนการหน่วยเวียนกลับแบบมีประตูและเสริมด้วยโครงข่ายประสาทการแปลงกราฟ (GRU and Graph convolutional neural network) และวิธีการปรับปรุงการถ่ายถอดรากกำลังสองเฉลี่ย (Root Mean Square Propagation : RMSProp) มีประสิทธิภาพในการทำนายต่ำที่สุด เนื่องจากค่าความผิดพลาดในการทำนายมากที่สุดด้วยค่าเฉลี่ยความผิดพลาดกำลังสองเท่ากับ 9.0725 และค่าเฉลี่ยสัมบูรณ์เท่ากับ 4.8488 ซึ่งสรุปได้ว่าตัวแบบความจำระยะสั้น และระยะยาว (LSTM) และเสริมด้วยโครงข่ายประสาทการแปลงกราฟ (Graph convolutional neural network) ด้วยวิธีการปรับปรุงการประมาณค่าช่วงเวลาปรับตัว (Adaptive Moment Estimation) เป็นตัวแบบที่มีประสิทธิภาพในการทำนายหาความต้องการของจุดให้บริการใหม่ และจุดให้บริการที่อยู่ใกล้เคียง

สารบัญ

บทคัดย่อ 1

สารบัญ 2

บทที่ 1 บทนำ 4

1.1 ที่มาและความสำคัญ 4

1.2 วัตถุประสงค์ของโครงงาน 6

1.3 ขอบเขตของโครงงาน 7

1.4 ผลผลิตที่ได้รับ (Output) 7

1.5 ผลลัพธ์ที่ได้รับ (Outcome) 7

บทที่ 2 ทฤษฎีที่เกี่ยวข้อง 8

2.1 โครงข่ายประสาท (Neural network) 8

2.2 ฟังก์ชันการใช้งาน (Activation function) 14

2.3 ค่าสัมประสิทธิ์ความคล้ายคลึงโคไซน์ (Cosine similarity) 15

2.4 ฟังก์ชันการสูญหาย (Loss function) 15

2.5 ขั้นตอนวิธีการปรับปรุงจากค่าความชัน (Gradient-based optimization algorithm) 16

บทที่ 3 วิธีการดำเนินการวิจัย 21

3.1 การศึกษาระบบดั้งเดิมและวิธีการแก้ไขปัญหา 21

3.2 เก็บข้อมูล (Collect data) 22

3.3 ทำความสะอาดข้อมูล (Clean data) 24

3.4 สร้างและฝึกฝนตัวแบบ (Build and train model) 27

3.5 การปรับแต่งตัวแบบ (Model tuning) 28

3.6 การทำนาย (Prediction) 29

3.7 การสร้างให้เห็นภาพ (Visualization) 30

บทที่ 4 ผลการดำเนินงาน 31

4.1 การปรับแต่งตัวแบบ 31

4.2 การทำนายผลกระทบที่เกิดขึ้นบนจุดให้บริการที่มีอยู่ 34
บทที่ 5 สรุปผลการดำเนินงานการวิจัย 35
5.1 บทสรุป 35
5.2 อภิปรายผล 35
5.3 ข้อเสนอแนะ 35
รายงานอ้างอิง 36

บทที่ 1 บทนำ

1.1 ที่มาและความสำคัญ

1.1.1 ธุรกิจให้บริการรถเช่า (Car sharing)

ธุรกิจให้บริการรถเช่า (Car sharing) หมายถึง ธุรกิจการเช่ายืมรถยนต์ระยะสั้นออกแบบมาเพื่อสร้างความสะดวกสบายให้กับผู้ที่สามารถขับรถยนต์ส่วนบุคคล ซึ่งไม่มีความประสงค์ที่ต้องการใช้รถยนต์ส่วนบุคคล หรือต้องการใช้รถยนต์เพียงแค่ชั่วคราว ธุรกิจนี้เป็นการลดจำนวนผู้ใช้งานรถยนต์บนท้องถนน และสามารถเพิ่มการเข้าถึงรถยนต์ส่วนบุคคลได้ในราคาที่ไม่แพง (นันพงศ์ พันทวีศักดิ์, 2018) โดยอธิบายเกี่ยวกับรูปแบบการดำเนินธุรกิจให้บริการรถเช่าดังนี้

1. P2P (Peer to Peer) เป็นการจับคู่กันระหว่างเจ้าของรถยนต์ที่ส่วนบุคคลต้องการปล่อยเช่ารถยนต์กับผู้ที่ต้องการเช่าเพื่อก่อให้เกิดรายได้ระหว่างที่เจ้าของรถยนต์ไม่ได้ใช้งาน

2. B2C (Business to Customer) เป็นการพัฒนารูปแบบการเช่ารถยนต์มาจาก P2P เป็นการเช่ารถระยะสั้น หรือระยะยาวแก่ลูกค้า โดยรถยนต์ที่ให้เช่านั้นเป็นกรรมสิทธิ์ของบริษัท ไม่ใช่รถยนต์ส่วนตัว ซึ่งรูปแบบ B2C นี้มักมีจุดแข็งในการให้บริการในรูปแบบการเดินทางเที่ยวเดียว (one way trip) ลูกค้าสามารถใช้บริการได้จากจุดที่ลูกค้าสะดวกและไม่จำเป็นต้องส่งคืนรถที่จุดเดิม เนื่องจากมีการการจัดการรถยนต์ที่ดีและมีจุดจอดจำนวนมาก

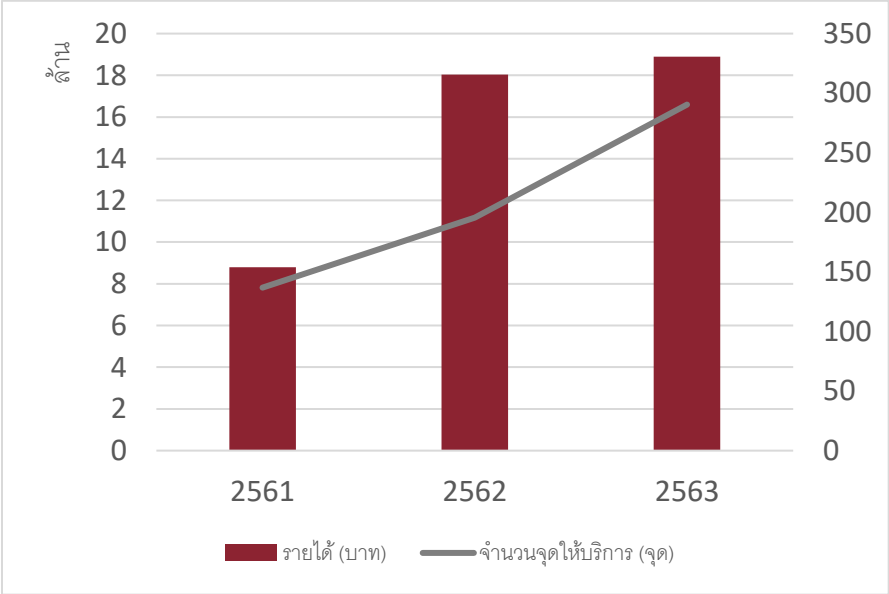
1.1.2 ธุรกิจให้บริการรถเช่าในประเทศไทย

ธุรกิจให้บริการรถเช่าในประเทศไทยพบข้อจำกัดที่มีอยู่มากในปัจจุบันในเรื่องการครอบคลุมของระบบขนส่งสาธารณะและปริมาณจุดจอดที่ยังไม่เพียงพอ ซึ่งไม่สามารถตอบสนองความต้องการกับการใช้งานของผู้บริโภคในประเทศไทยได้มาก แม้กรุงเทพมหานคร และปริมณฑล เป็นเมืองที่มีประชากรจำนวนมากและมีระบบขนส่งสาธารณะในหลากหลายรูปแบบ เช่น การขนส่งระบบราง BTS, MRT เรือด่วน และรถประจำทาง แต่ระบบขนส่งที่กล่าวมาข้างต้นไม่มีการเชื่องโยงเป็นเครือข่ายที่มากเพียงพอ และไม่ครอบคลุมทุกพื้นที่ จึงทำให้ผู้บริโภคเลี่ยงการเดินทางด้วยรถยนต์ส่วนตัวได้ยาก อีกทั้งยังไม่มีการกำหนดพื้นที่จอดรถสาธารณะในเขตชุมชนหรือย่านการค้า และอาคารจอดแล้วจรยังมีปริมาณน้อยไม่กระจายตัวไปยังบริเวณชานเมือง ส่งผลกระทบให้ธุรกิจให้บริการรถเช่าในรูปแบบ B2C ไม่สามารถให้บริการในรูปแบบการเดินทางเที่ยวเดียวได้ เนื่องจากจุดจอดที่ยังมีการขยายตัวที่ไม่มากเพียงพอให้เกิดการจูงใจผู้บริโภคมาใช้งานการบริการให้บริการรถเช่า

1.1.3 เกี่ยวกับบริษัทให้บริการรถเช่าที่ใช้เป็นกรณีศึกษา

บริษัทที่ใช้เป็นกรณีศึกษา ซึ่งเป็นบริษัทที่ดำเนินกิจการด้านการให้เช่ารถยนต์น้ำมัน และรถยนต์ไฟฟ้า รวมไปถึงสกู๊ตเตอร์ไฟฟ้าเป็นรายชั่วโมง หรือรายเดือน (ระยะยาว) ผ่านแอปพลิเคชันโทรศัพท์มือถือ โดยเป็นตัวกลางเชื่อมต่อเจ้าของรถยนต์ที่ไม่ได้ใช้งานเป็นประจำกับผู้ใช้งานเข้าด้วยกัน เพื่อให้ผู้ใช้งานที่ต้องการใช้รถยนต์สามารถเช่าได้ และเจ้าของรถยนต์ที่ไม่ได้ใช้งานเป็นประจำสามารถสร้างรายได้จากรถยนต์ของตนเองได้บริษัทที่ใช้เป็นกรณีศึกษา

บริษัทที่ใช้เป็นกรณีศึกษาเริ่มเปิดบริการเมื่อปี พ.ศ. 2559 บริษัทจึงมุ่งเน้นการเปิดให้บริการในกรุงเทพมหานครและปริมณฑล รวมถึงจังหวัดที่เป็นแหล่งท่องเที่ยว เพื่อให้ได้รับโอกาสที่ผู้บริโภคใช้บริการเช่ารถยนต์ของบริษัทมากที่สุด จากรูปที่ 1 แสดงถึงความสัมพันธ์ระหว่างรายได้ที่บริษัทได้รับผ่านการรับรู้ทางบัญชีและจำนวนจุดให้บริการที่บริษัทมีอยู่ว่าการเพิ่มจำนวนจุดให้บริการสามารถเพิ่มรายได้ให้แก่บริษัทได้ เนื่องจากการเปิดจุดให้บริการเพิ่มทำให้ลูกค้าสามารถเข้าถึงบริการของทางบริษัทได้มาก



รูปที่ 1 จำนวนจุดให้บริการ และรายได้จากกิจกรรมการเช่ารถยนต์รายชั่งโมงในปี พ.ศ. 2561 - 2563

1.1.4 ปัญหาที่เกิดขึ้น

จากสถานการณ์การแพร่กระจายโรคระบาดโรคติดเชื้อไวรัสโคโรนา 2019 เมื่อปลายปีพ.ศ. 2562 รัฐบาลประกาศใช้พระราชกำหนดฉุกเฉินส่งผลให้ประชาชนจำเป็นต้องลดการเดินทางออกจากที่พักอาศัยที่ไม่มีความจำเป็นลง เพื่อป้องกันการแพร่ระบาดของโรคระบาดดังกล่าว ซึ่งส่งผลต่อการเติบโตของยอดใช้บริการเช่ารถยนต์ของบริษัทอย่างมีนัยสำคัญ และในบางจุดที่ให้บริการไม่สามารถก่อให้เกิดรายได้ ทำให้บริษัทจำเป็นที่ต้องลดจุดที่ให้บริการลงเพื่อลดภาระค่าใช้จ่ายที่เกิดขึ้น

แม้ว่าการปิดตัวลงของจุดที่ให้บริการนั้นเป็นสิ่งที่ไม่สามารถหลีกเลี่ยงได้เนื่องจากสถานการณ์แพร่ระบาด แต่การสร้างจุดให้บริการขึ้นมาใหม่เป็นสิ่งที่มีความสำคัญเช่นกัน เพื่อให้เกิดกิจกรรมการเช่ารถยนต์ที่เพิ่มขึ้น ซึ่งแปรผันตรงกับการสร้างรายได้ให้กับบริษัท แต่ว่าถ้าหากว่าการจัดตั้งจุดให้บริการนั้น ๆ ไม่ได้ส่งผลในทางด้านดี เช่น การเปิดจุดบริการใหม่นั้นไม่ได้ก่อให้เกิดรายได้ที่เพียงพอ หรือไม่ได้ส่งเสริมกิจกรรมการเช่ารถซึ่งกันและกันกับจุดบริการที่อยู่ใกล้เคียง ทำให้บริษัทจำเป็นที่จำต้องแบกรับค่าใช้จ่ายที่เพิ่มมากขึ้นและต้องใช้เวลาเพิ่มเติมในการหาจุดให้บริการใหม่ โดยที่กระบวนการในการค้นหาและติดต่อทำสัญญาติดตั้งจุดให้บริการใหม่นั้นจะต้องใช้ระยะเวลาประมาณ 1-3 เดือน ดังตารางที่ 1 และการพิจารณาติดตั้งจุดบริการใหม่นั้นมาจากดุลพินิจของผู้ที่มีประสบการณ์โดยตรง หรือเป็นการใช้ผู้สังเกตการณ์ ซึ่งเป็นความเสี่ยงอย่างมากจากการพึ่งพาความสามารถจากบุคคลกลุ่มหนึ่ง เพราะหากบุคคลที่กล่าวไปข้างต้นไม่สามารถปฏิบัติหน้าที่ได้ เกษียณอายุ หรือหมดวาระการทำงานกับบริษัทแล้วทำให้บริษัทขาดความสามารถในการสร้างจุดให้บริการใหม่ที่มีประสิทธิภาพอย่างต่อเนื่องขึ้นมาได้

ตารางที่ 1 กระบวนการทำงาน และระยะเวลาในการดำเนินงานเพื่อติดตั้งจุดให้บริการใหม่

ลำดับที่	กระบวนการทำงาน	ระยะเวลา
1	รวบรวมหาพื้นที่ในการติดตั้งจุดให้บริการใหม่	1-2 สัปดาห์
2	คัดเลือกประชุมเพื่อตัดสินใจเลือกพื้นที่ในการติดตั้ง	4 วัน - 1 สัปดาห์
3	ยื่นข้อเสนอให้แก่ผู้บริหาร	3 วัน - 1 สัปดาห์
4	ยื่นข้อเสนอเพื่อขอเช่าพื้นที่ในการติดตั้งบริการ	1 สัปดาห์ - 2 เดือน
5	ช่างและวิศวกรเข้าสู่พื้นที่เพื่อทำการติดตั้งจุดให้บริการ	1-2 สัปดาห์

1.2 วัตถุประสงค์ของโครงงาน

เพื่อสร้างตัวแบบทางคณิตศาสตร์ในการทำนายจำนวนความต้องการใช้จุดให้บริการใหม่ และรับรู้ถึงผลกระทบต่อจุดให้บริการที่มีอยู่ ตลอดจนเป็นเครื่องมือช่วยในการตัดสินใจติดตั้งจุดให้บริการใหม่ของบริษัทกรณีศึกษาจากข้อมูลย้อนหลังของจุดให้บริการที่มีอยู่

1.3 ขอบเขตของโครงงาน

1. โครงงานนี้มุ่งเน้นศึกษาจุดที่ให้บริการที่ตั้งอยู่ในพื้นที่ของกรุงเทพมหานคร และปริมณฑลเท่านั้น

2. โครงงานนี้ศึกษาเฉพาะจำนวนกิจกรรมการจองรถยนต์จากจุดให้บริการที่ลูกค้าเช่ารถยนต์ผ่านแอปพลิเคชัน

3. โครงงานนี้ศึกษาเฉพาะการเช่ารถยนต์ผ่านแอปพลิเคชันของบริษัทที่ใช้เป็นกรณีศึกษาเท่านั้น

4. โครงงานนี้ศึกษาเฉพาะข้อมูลของเมื่อวันที่ 1 มกราคม พ.ศ. 2561 ถึง 31 กันยายน พ.ศ. 2564

1.4 ผลผลิตที่ได้รับ (Output)

ตัวแบบสามารถทำนายจำนวนความต้องการที่เกิดขึ้นในอนาคตได้อย่างแม่นยำ โดยการจำลองเปิดจุดให้บริการใหม่ สามารถทำนาย และแยกได้ว่าจำนวนความต้องการนั้นเกิดขึ้นมาจากการส่งเสริมซึ่งกันและกันของจุดให้บริการใกล้เคียง หรือเป็นการแย่งจำนวนความต้องการมาจากจุดให้บริการที่อยู่ใกล้เคียงกัน

1.5 ผลลัพธ์ที่ได้รับ (Outcome)

สถานที่ที่นำมาสร้างมาเป็นจุดให้บริการใหม่มาจากสถานที่ที่เป็นการเช่าตามระยะเวลาของสัญญาที่กำหนดเอาไว้ผ่านพนักงานที่ดูแลฝ่ายนี้โดยตรง การได้มาซึ่งสถานที่ที่ใช้เป็นจุดให้บริการใหม่ได้มาจากการสุ่ม หรือคัดเลือกมาจากบุคคลในบริษัทที่เล็งเห็นความเป็นไปได้ในการสร้างรายได้ให้แก่บริษัท โดยที่มักเป็นจุดที่มีผู้คนหนาแน่น จุดที่ติดกับรถไฟฟ้า หรือห้างสรรพสินค้า เป็นต้น ทางโครงงานนี้มีเป้าหมายที่ได้รับผลลัพธ์ดังนี้

1.5.1 ตัวแบบสามารถทำนายจำนวนการใช้งานจุดบริการของจุดให้บริการเดิม และจุดให้บริการใหม่หากมีการติดตั้งจุดบริการใหม่ขึ้นมา

1.5.2 ตัวแบบสามารถเป็นเครื่องมือช่วยการตัดสินใจถึงผลกระทบของการเปิดจุดให้บริการใหม่ได้ เนื่องจากผลลัพธ์การทำนายของตัวแบบสามารถบ่งชี้จำนวนความต้องการที่เพิ่มขึ้นหรือลดลงจากการติดตั้งจุดให้บริการใหม่ได้

บทที่ 2 ทฤษฎีที่เกี่ยวข้อง

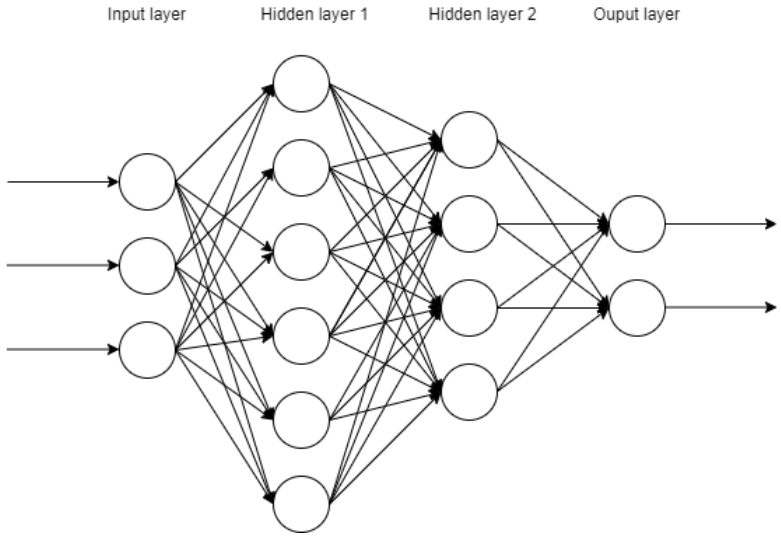
ทฤษฎีที่เกี่ยวข้องที่ใช้ในการสร้างตัวแบบการทำนายความต้องการของจุดให้บริการรถเช่าใหม่จากข้อมูลของจุดให้บริการที่มีอยู่ประกอบไปด้วยด้วย 4 ทฤษฎี ได้แก่

1. โครงข่ายประสาทเทียม (Neural network)
2. ค่าสัมประสิทธิ์ความคล้ายคลึงโคไซน์ (Cosine similarity)
3. ฟังก์ชันการสูญหาย (Loss function)
4. ขั้นตอนวิธีการปรับปรุงจากค่าความชัน (Gradient-based optimization algorithm)

2.1 โครงข่ายประสาท (Neural network)

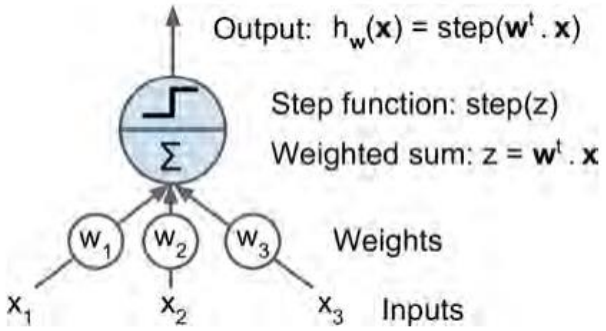
โครงข่ายประสาท (Neural network) หรือโครงข่ายประสาทเทียม (Artificial neural network) เป็นการคำนวณของระบบคอมพิวเตอร์ซึ่งได้รับแรงบันดาลใจมาจากระบบประสาทสมองดังรูปที่ 2 ประกอบด้วยองค์ประกอบการคำนวณอย่างง่ายที่เชื่อมต่อกันในรูปแบบต่าง ๆ เพื่อให้สอดคล้องกับการคำนวณเฉพาะ เช่น การจำแนกวัตถุ หรือการทำนายในอนาคต เป็นต้น โดยในโครงข่ายประสาทนั้นมีส่วนประกอบดังนี้

1) ชั้นผิวรับข้อมูล (Input layer) คือ ชั้นผิวที่มีหน้ารับข้อมูลนักพัฒนาเพื่อส่งต่อไปให้ชั้นผิวอื่น
2) ชั้นผิวซ่อน (Hidden layer) คือ ชั้นผิวที่นำข้อมูลที่ได้มาคำนวณ โดยอยู่ระหว่างชั้นผิวรับข้อมูล และชั้นผิวผลลัพธ์
3) ชั้นผิวผลลัพธ์ (Output layer) คือ ชั้นผิวที่สร้างผลลัพธ์จากข้อมูลที่ได้รับ



รูปที่ 2 สถาปัตยกรรมโครงข่ายประสาท

จากรูปที่ 2 เห็นได้ว่าสถาปัตยกรรมโครงข่ายประสาทลำเลียงข้อมูลผ่านจากชั้นผิวรับข้อมูลไปยังชั้นผิวซ่อน และได้ผลลัพธ์ผ่านชั้นผิวผลลัพธ์ วงกลมทุกวงนั้นถึงเชื่อมต่อเข้าหาซึ่งกันและกัน วงกลมแต่ละวงของสถาปัตยกรรมนี้ คือ ฟังก์ชันการใช้งาน (Activation function) และภายในลูกศรแต่ละลูกศรประกอบไปด้วยค่าน้ำหนักที่แตกต่างกัน โดยกระบวนการทำงานย่อยมีการทำงานดังรูปที่ 3

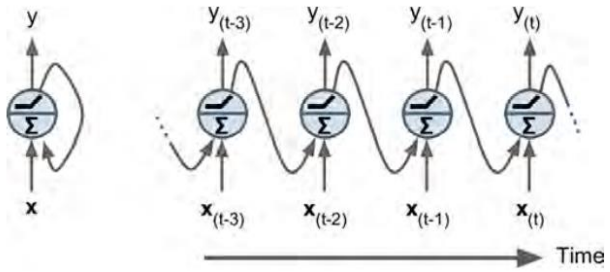


รูปที่ 3 กระบวนการทำงานย่อยในสถาปัตยกรรมโครงข่ายประสาท, โดย Géron, A. (2017). O'Reilly Media, Inc.

จากรูปที่ 3 เห็นว่าค่าผลลัพธ์นั้นสามารถคำนวณได้จากการนำผลรวมถ่วงน้ำหนักของข้อมูลใส่ในฟังก์ชันขั้นบันได (Step function) หรือเรียกว่า ฟังก์ชันการใช้งาน (Activation function) ซึ่งเป็นฟังก์ชันที่ไม่เป็นเส้นตรง (Non-linear function) ทำให้โครงข่ายประสาทสามารถคำนวณได้อย่างซับซ้อนขึ้น โดยที่ฟังก์ชันการใช้งานนั้นมีอยู่หลายฟังก์ชันด้วยกัน เช่น ฟังก์ชันการใช้งานเส้นตรงที่ถูกปรับแก้ (ReLU (Rectified linear unit) activation function) ฟังก์ชันการใช้งานเส้นตรงที่ถูกปรับแก้แบบรั่ว (Leaky ReLU activation function) ฟังก์ชันการใช้งานไฮเปอร์โบล่าแทนเจนต์ (Hyperbolic tangent activation function) เป็นต้น

2.1.1 ตัวแบบความจำระยะสั้น และระยะยาว (Long Short-term memory: LSTM)

ตัวแบบความจำระยะสั้น และระยะยาว (Long Short-term memory : LSTM) เป็นรูปแบบโครงข่ายประสาทซึ่งพัฒนามากจากโครงข่ายประสาทเทียมแบบวนกลับ (Recurrent neural network) โดยออกแบบเพื่อการใช้งานการประมวลผลรูปแบบข้อมูลที่เป็นลำดับ (Sequence) โดยที่ผลลัพธ์ก่อนหน้าถูกใช้เพื่อประมวลผลใหม่ในส่วนถัด ๆ ไป ดังรูปที่ 4 โดยใน 1 หน่วย (unit) ในแต่ละหน่วยเริ่มจากด้านซ้ายด้วยการป้อนข้อมูล x_1 และผลลัพธ์คือ y_1 จากนั้นผลลัพธ์จาก y_1 ถูกนำไปคิดประมวลผลกับหน่วยที่ 2, 3, 4, ... , n ตามลำดับ (Géron, n.d., 2017)



รูปที่ 4 รูปแบบการทำงานของ RNN อย่างง่าย, โดย Géron, A. (2017). O'Reilly Media, Inc.

จากรูปที่ 5 เห็นได้ว่า ในหน่วยของ LSTM นั้นแยกเวกเตอร์จากทางซ้ายออกเป็น 2 เวกเตอร์ คือ h(t) และ c(t) เราสามารถเข้าใจได้อย่างง่ายคือ h(t) แสดงถึงความจำระยะสั้น และ c(t) แสดงถึงความจำระยะยาว x(t) แสดงถึงข้อมูลป้อนเข้าหน่วย (input) และ y(t) แสดงถึงผลผลิต (output) ซึ่งอธิบายรูปแบบการทำงานดังต่อไปนี้

c(t) คือ ความจำระยะยาวที่ส่งถ่ายจากหน่วยด้านซ้ายไปด้านขวา เพื่อเป็นส่วนช่วยในการประมวลผลลัพธ์ในส่วนของหน่วยถัด ๆ ไป โดยที่ c(t-1) จากหน่วยก่อนหน้าต้องผ่านประตูการลืม (forget gate) เพื่อเป็นการทิ้งความทรงจำระยะยาวบางส่วน เพื่อไม่ให้ความจำระยะยาวทั้งหมดเข้าไปมีส่วนเกี่ยวข้องกับการประมวลผลในหน่วยนั้น ๆ และเพิ่มความทรงจำใหม่ผ่านกระกวนการเพิ่มความจำ โดยเข้ามาทางประตูการป้อน (input gate) หลังจากนั้นถูกส่งออกเป็น c(t) โดยไม่ทำการเปลี่ยนแปลงความทรงจำเพื่อให้ไปยังหน่วยถัดไป นอกจากนี้ ความทรงจำระยะยาวที่ผ่าน input gate แล้วถูกคัดลอก และถูกส่งผ่านไปยังฟังก์ชัน tanh จากนั้นถูกคัดกรองด้วยประตูผลผลิต (output gate) ซึ่งเป็นจุดที่สร้าง h(t) หรือก็คือความทรงจำระยะสั้นขึ้นมา

ในการคำนวณและประมวลผลนั้นนั้น ต้องได้รับข้อมูลมาจาก x(t) และความจำระยะสั้น h(t) ถูกส่งไปยัง 4 layers ต่าง ๆ ที่มีหน้าที่แตกต่างกันออกไป โดยที่ layer หลักก็คือ g(t) ซึ่งมีหน้าที่ในการวิเคราะห์ข้อมูลปัจจุบันกับความจำระยะสั้นเพื่อประมวลผลออกมาเป็น output y(t) และอีก 3 layers ที่เหลือคือ ตัวควบคุมประตู (gate controllers) มีค่าอยู่ระหว่าง 0 ถึง 1 ซึ่งคูณเข้ากับข้อมูลที่ไหลเข้ามายังประตูต่าง ๆ กล่าวคือมีหน้าที่ก็คือการเปิดปิดการไหลของข้อมูล โดย 0 หมายถึงการปิดประตูนั้น ๆ และ 1 หมายถึงการเปิดประตู

- ประตูป้อนข้อมูล (Input gate) ถูกควบคุมโดย i(t) ควบคุมว่าส่วนไหนของ g(t) ควรที่ถูกเพิ่มเข้าไปในความจำระยะยาว
- ประตูการลืม (Forget gate) ถูกควบคุมโดย f(t) ควบคุมในส่วนของความทรงจำระยะยาว
- ประตูผลผลิต (Output gate) ถูกควบคุมโดย o(t) ควบคุมว่าส่วนไหนของความจำระยะยาวควรที่ถูกอ่านและส่งออกมาเป็น output และความจำระยะสั้น h(t)

สมการในการคำนวณนั้นสามารถได้จากสมการที่ 1 – 6

$$i(t) = \sigma(W^i x_t + U^i h_{t-1} + b_i) \tag{1}$$

$$f(t) = \sigma(W^f x_t + U^f h_{t-1} + b_f) \tag{2}$$

$$o(t) = \sigma(W^o x_t + U^o h_{t-1} + b_o) \tag{3}$$

$$g(t) = \tanh\,(W x_t + U h_{t-1} + b) \tag{4}$$

$$c(t) = f(t) \otimes c(t-1) + i(t) \otimes g(t) \tag{5}$$

$$h(t) = o(t) \otimes \tanh\,(c(t)) \tag{6}$$

โดยที่ $i(t)$ คือ เวกเตอร์ประตูการป้อนข้อมูล

$f(t)$ คือ เวกเตอร์ประตูการลืม

$o(t)$ คือ เวกเตอร์ประตูผลผลิต

$g(t)$ คือ ค่าความทรงจำระยะสั้นที่เหลืออยู่

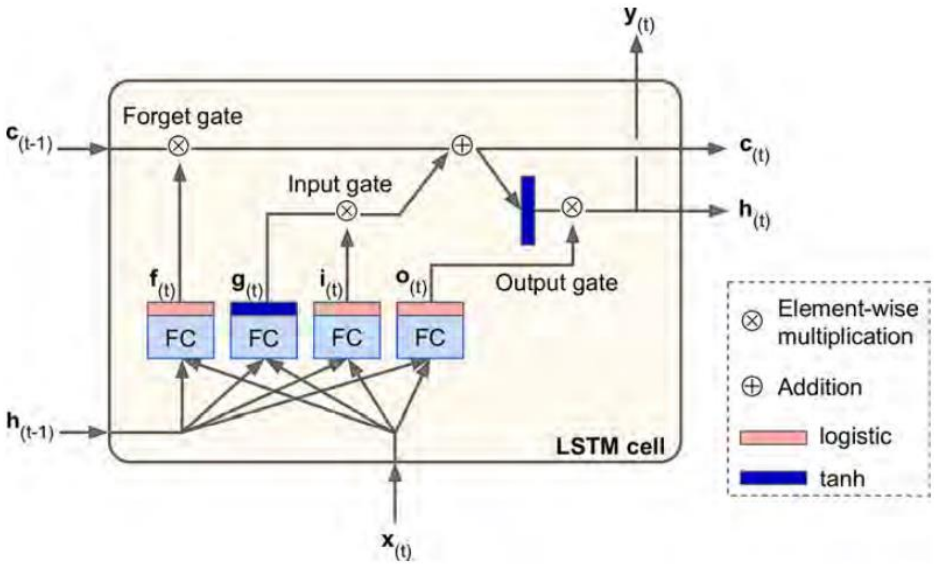
$c(t)$ คือ ค่าความทรงจำระยะยาว

$c(t-1)$ คือ ค่าความทรงจำระยะยาวก่อนหน้า

$h(t)$ คือ ค่าความทรงจำระยะสั้น

W^i, W^f, W^o, W คือ ค่าน้ำหนักของค่าป้อนเข้าในประตูป้อนข้อมูล ประตูการลืม ประตูส่งออกข้อมูล และความทรงจำที่เหลืออยู่ ตามลำดับ

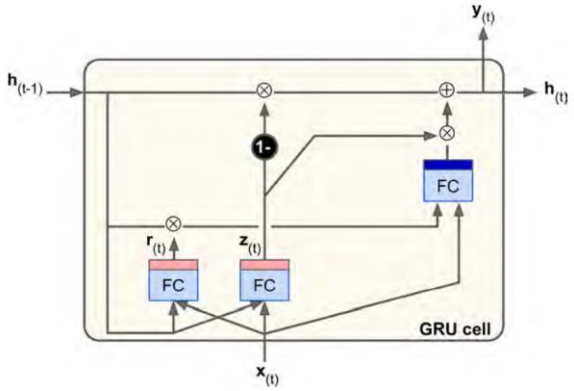
b_i, b_f, b_o, b คือ ค่าความลำเอียงที่เกิดขึ้นในประตูป้อนข้อมูล ประตูการลืม ประตูส่งออกข้อมูล และความทรงจำที่เหลืออยู่ ตามลำดับ โดยปกติค่า b_f มักถูกตั้งต้นด้วย 1 ในช่วงแรกของการฝึกฝนตัวแบบเพื่อป้องกันการลืมความทรงจำ



รูปที่ 5 ภายในของแต่ละหน่วยของ LSTM, โดย Géron, A. (2017). O'Reilly Media, Inc.

2.1.2 หน่วยเวียนกลับแบบมีประตู (Gated recurrent unit: GRU)

หน่วยเวียนกลับแบบมีประตู (Gated Recurrent Unit) คือหนึ่งในรูปแบบโครงข่ายประสาทซึ่งพัฒนามากจากโครงข่ายประสาทเทียมแบบวนกลับ โดยมีจุดประสงค์เพื่อใช้งานการประมวลผลรูปแบบข้อมูลที่เป็นลำดับเหมือนกับตัวแบบรูปแบบความจำระยะสั้น และระยะยาว (Long short-term memory) โดยที่มีประตูการลืม (Forget gate) แต่ไม่มีประตูผลผลิต (Output gate) ดังรูปที่ 6



รูปที่ 6 ภายในของแต่ละหน่วยของ GRU, โดย Géron, A. (2017). O'Reilly Media, Inc.

จากรูปที่ 6 เห็นได้ว่าหน่วยเวียนกลับแบบมีประตูนั้นรับค่าความทรงจำที่เกิดขึ้นจากหน่วยก่อนหน้าเพียงแค่เวกเตอร์เดียว คือ ความทรงจำระยะสั้น h(t) ซึ่งต่างจากตัวแบบรูปแบบความจำระยะสั้น และระยะยาวที่มีเวกเตอร์ที่หน่วยก่อนหน้าส่งไปยังหน่วยถัดไปด้วยกันทั้งหมด 2 เวกเตอร์ คือ ความทรงจำระยะสั้น และความทรงจำระยะยาว โดยที่หน่วยเวียนกลับแบบมีประตูมีตัวควบคุมประตู 2 ตัวดังนี้

- ประตูการปรับปรุง (Update gate) คือ ประตูที่มีส่วนช่วยให้ตัวแบบสามารถหาได้ว่าควรที่ส่งออกความทรงจำระยะสั้นไปยังหน่วยถัดไปเท่าใด โดยที่ใช้การปรับปรุงค่าความทรงจำระยะสั้นที่ได้มาจากหน่วยก่อนหน้า h(t-1) เพื่อให้ได้ความทรงจำระยะสั้น h(t) ที่ส่งต่อให้หน่วยถัดไปด้วยการคำนวณเวกเตอร์ประตูการปรับปรุง (update gate vector: z(t)) ซึ่งได้มาจากการคำนวณกันระหว่างค่าความทรงจำระยะสั้นก่อนหน้า h(t-1) และค่าข้อมูลป้อนเข้าหน่วย x(t) เพื่อใช้ในการปรับปรุงค่าความทรงจำระยะสั้นก่อนหน้าก่อนส่งออกไปยังหน่วยถัดไป ดังสมการที่ 7

$$z(t) = \sigma(W^z x_t + U^z h_{t-1}) \tag{7}$$

โดยที่ $\mathrm{z(t)}$ คือ เวกเตอร์ประตูการปรังปรุง

σ คือ ฟังก์ชัน sigmoid

W^z คือ ค่าน้ำหนักของค่าข้อมูลป้อนเข้าในเวกเตอร์ประตูการปรังปรุง

U^z คือ ค่าน้ำหนักของค่าความทรงจำระยะสั้นก่อนหน้าในเวกเตอร์ประตูการปรังปรุง

- ประตูการตั้งใหม่ (Reset gate) คือ ประตูที่มีส่วนช่วยให้ตัวแบบสามารถตัดสินใจได้ว่าควรที่ลืมความทรงจำระยะสั้นก่อนหน้ามากน้อยเท่าใดเพื่อใช้ในการคำนวณในหน่วยด้วยการคำนวณค่าเวกเตอร์ประตูการตั้งใหม่ (Reset gate vector: r(t)) ซึ่งได้มาจากการคำนวณกันระหว่างค่าความทรงจำระยะสั้นก่อนหน้า h(t-1) และค่าข้อมูลป้อนเข้าหน่วย x(t) ดังสมการที่ 8

$$r(t) = \sigma(W^r x_t + U^r h_{t-1}) \tag{8}$$

โดยที่ $\mathrm{r(t)}$ คือ เวกเตอร์ประตูการตั้งใหม่

W^r คือ ค่าน้ำหนักของค่าข้อมูลป้อนเข้าในเวกเตอร์ประตูการตั้งใหม่

U^r คือ ค่าน้ำหนักของค่าความทรงจำระยะสั้นก่อนหน้าในเวกเตอร์ประตูการตั้งใหม่

หลังจากที่ได้ค่าเวกเตอร์ประตูการปรังปรุงและเวกเตอร์ประตูการตั้งใหม่แล้ว หน่วยเวียนกลับแบบมีประตูนำค่าเวกเตอร์ทั้งสองมาคำนวณเพื่อที่ส่งค่าความทรงจำระยะสั้นใหม่ออกไปยังหน่วยถัดไป ดังสมการที่ 9 และสมการที่ 10)

$$h'(t) = \tanh\big(Wx_t + r \odot Uh(t-1)\big)) \tag{9}$$

$$h(t) = z \odot h(t-1) + (1-z) \odot h'(t) \tag{10}$$

โดยที่ $h(t)$ คือ ค่าความทรงจำระยะสั้นใหม่

$h'(t)$ คือ ค่าความทรงจำที่เหลืออยู่ หลังจากผ่านประตูการตั้งใหม่

$h(t-1)$ คือ ค่าความทรงจำระยะสั้นก่อนหน้า

$\odot$ คือ Hadamard product

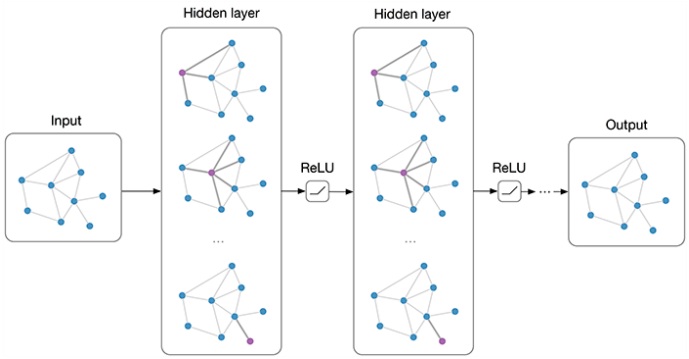
$tanh$ คือ ฟังก์ชัน Hyperbolic tangent

W คือ ค่าน้ำหนักของค่าข้อมูลป้อนเข้า

U คือ ค่าน้ำหนักของค่าค่าความทรงจำระยะสั้นก่อนหน้า

2.1.3 โครงข่ายประสาทการแปลงกราฟ (Graph convolutional neural network)

โครงข่ายประสาทการแปลงกราฟ (Graph convolutional neural network) เป็นรูปแบบหนึ่งของโครงข่ายประสาทเทียมที่ถูกออกแบบมาเพื่อให้ทำงานกับรูปแบบข้อมูลที่อยู่ในรูปของกราฟโดยตรง เพื่อการเข้ารหัสของโครงสร้างกราฟ (Graph structure) และคุณสมบัติของแต่ละจุด (Feature of nodes) ในกราฟโดยจำลองเสมือนการเห็นกราฟของมนุษย์ที่แบ่งเป็นพื้นที่ย่อย ๆ โดยการมองดังกล่าวเป็นการมองผ่านกราฟเมทริกซ์ประชิด (Adjacency matrix graph) เพื่อแปลงคุณสมบัติของแต่ละจุดให้เป็นหมวดหมู่อย่างชัดเจนในรูปแบบที่ต่างกันออกไปดังรูปที่ 7



รูปที่ 7 Graph convolutional neural network

ในการวิเคราะห์โครงข่ายประสาทการแปลงกราฟนั้นถูกแบ่งออกเป็นชั้นผิว (Layer) ซึ่งการส่งผ่านข้อมูลเพื่อที่ไปอีกชั้นผิวนั้น จำเป็นที่ต้องผ่านชั้นผิวซ่อน (Hidden layer) ซึ่งแต่ hidden layer นั้นคือฟังก์ชันทางคณิตศาสตร์มีค่าอยู่ระหว่างศูนย์ถึงหนึ่ง เรียกว่า ฟังก์ชันการใช้งาน (Activation function) โดยที่ภายในแต่ละชั้นผิวจะมีการคำนวณดังสมการที่ 11

$$H_{t-1}^{(l)} = \sigma\left(\sum_{A_{t-1}\in A_{t-1}} f(A_{t-1})H_{t-1}^{(l-1)}W_{t-1}^{(l-1)}\right) \tag{11}$$

โดยที่ H_{t-1}^{l}, $H_{t-1}^{(l-1)}$ คือ ชั้วผิวซ่อนในชั้นที่ l และ l-1 ณ เวลาที่ t-1 ตามลำดับ

σ คือ สมการที่ไม่เป็นเส้นตรง

$f(A_{t-1})$ คือ ฟังก์ชันบนกราฟเมทริกซ์ประชิด

$W_{t-1}^{(l-1)}$ คือ ค่าน้ำหนักที่ได้จากการฝึกฝนตัวแบบ

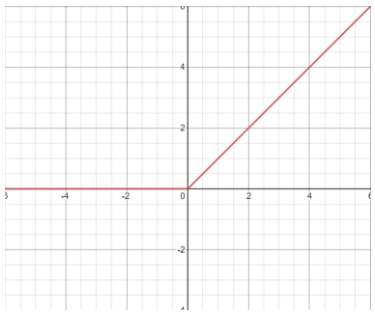
2.2 ฟังก์ชันการใช้งาน (Activation function)

ฟังก์ชันการใช้งาน (Activation function) คือ ฟังก์ชันที่รับผลรวมจากชั้นผิวก่อนหน้าหน้าเพื่อนำมาประมวลผลต่อแล้วพิจารณาวถึงความเหมาะสมในการส่งผลผลิตออกไปตามรูปแบบสมการที่ได้กำหนดเอาไว้

2.2.1 ฟังก์ชันการใช้งานเส้นตรงที่ถูกปรับแก้ (ReLU (Rectified Linear unit) activation function)

ฟังก์ชันการใช้งานเส้นตรงที่ถูกปรับแก้ (ReLU (Rectified Linear unit) activation function) เป็นฟังก์ชันที่มีส่วนของเส้นตรงอยู่ในฟังก์ชัน คือ ถ้าหากว่าค่า input ที่เราใส่เข้ามีค่ามามากกว่าศูนย์ ค่าความชันของฟังก์ชันมีค่าเท่ากับ 1 ดังรูปที่ 8 และสมการที่ 12 ซึ่งมีข้อดีคือทำให้การฝึกข้อมูลของตัวแบบนั้นเป็นไปได้อย่างรวดเร็ว

$$f(x) = \max(0, x) \begin{cases} 0 \; if \; x \leq 0 \\ x \; if \; x > 0 \end{cases} \tag{12}$$



รูปที่ 8 ReLU activation function

2.3 ค่าสัมประสิทธิ์ความคล้ายคลึงโคไซน์ (Cosine similarity)

ค่าความคล้ายคลึงโคโซน์ (Cosine similarity) คือการวัดค่าความคล้ายคลึงของ 2 เวกเตอร์ด้วยการใช้ค่าของเวกเตอร์ระหว่าง 2 ค่านั้นมาคำนวณหาค่าโคไซน์ ถ้าหากว่าค่าโคโซน์นั้นมีค่าเข้าใกล้หนึ่ง หมายความว่า เวกเตอร์ทั้งสองมีความคล้ายคลึงกัน ดังสมการที่ 13

$$similarity = \cos(\theta) = \frac{A \cdot B}{\|A\|\|B\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}} \tag{13}$$

โดยที่ n คือ จำนวนข้อมูลในเวกเตอร์

A คือ เวกเตอร์ A

B คือ เวกเตอร์ B

2.4 ฟังก์ชันการสูญหาย (Loss function)

ฟังก์ชันการสูญหาย (Loss function) ถูกใช้ในการหาค่าความคาดเคลื่อนระหว่างค่าของข้อมูลจริง (Actual value) กับ ค่าที่ถูกทำนาย (Predicted value) สำหรับเพิ่มประสิทธิภาพของค่าน้ำหนักที่อยู่ในตัวแบบโครงข่ายประสาท (Neural network model) และเพื่อเป็นตัวชี้วัดความแม่นยำของตัวแปรที่ถูกใส่เข้าไปในตัวแบบ โดยที่ฟังก์ชันการสูญหายถูกใช้ในบริบทที่แตกต่างกันตามเป้าหมายในการใช้งานของตัวแบบ เช่น การถดถอย (Regression) หรือการจัดหมวดหมู่ (Classification) เป็นต้น

2.4.1 ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Mean squared error)

ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Mean squared error) นั้นเป็นฟังก์ชันการสูญหายที่ถูกใช้กันอย่างมากสำหรับสมการถดถอย เป็นการวัดความสามารถของตัวแบบด้วยค่าความต่างกำลังสองระหว่างค่าข้อมูลจริง (Actual value) และค่าข้อมูลที่ทำนาย (Predicted value) หารด้วยจำนวนข้อมูลดังสมการที่ 14

$$MSE = \frac{1}{N}\sum_{i=0}^{N}(y_i - \hat{y}_i)^2 \tag{14}$$

โดยที่ N คือ จำนวนข้อมูล

y คือ ค่าข้อมูลจริง

$\hat{y}$ คือ ค่าข้อมูลที่ทำนาย

2.5 ขั้นตอนวิธีการปรับปรุงจากค่าความชัน (Gradient-based optimization algorithm)

ขั้นตอนวิธีการปรับปรุงจากค่าความชัน (Gradient-based optimization algorithm) คือขั้นตอนที่วนซ้ำที่ใช้ข้อมูลความชันของฟังก์ชันการสูญหายเพื่อปรับปรุงค่าตัวแปรต่าง ๆ ในแต่ละการวน

2.5.1 การไล่ความชันแบบสุ่มด้วยโมเมนตัม (Stochastic Gradient Descent with momentum)

การไล่ระดับความชัน (Gradient descent) เป็นหนึ่งในวิธีการปรับปรุงที่มีความแพร่หลายกันมากที่สุด ด้วยวิธีการที่ง่ายที่สุดที่ปรับปรุงคุณภาพของโครงข่ายประสาท โดยหลักการปรับค่าตัวแปรเป็นไปตามดังสมการที่ 15

$$\theta = \theta - \eta \cdot \nabla_{\theta} J(\theta) \tag{15}$$

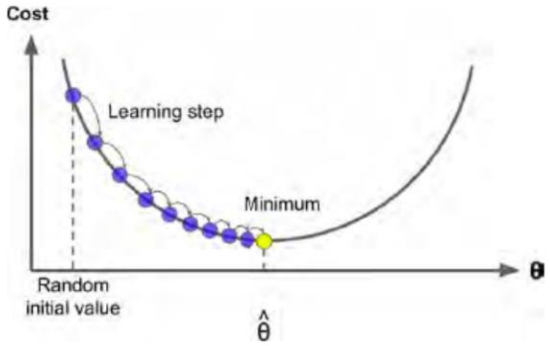
โดยที่ θ คือ ค่าตัวแปรของตัวแบบ (ค่าน้ำหนัก)

$J(\theta)$ คือ ฟังก์ชันสูญเสีย

$\nabla_{\theta} J(\theta)$ คือ ค่าความชันของฟังก์ชันของสูญเสีย

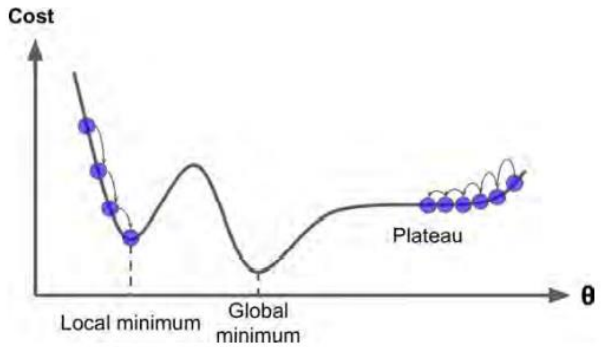
η คือ อัตราการเรียนรู้

จากสมการที่ 15 เห็นได้ว่าค่าตัวแปรของตัวแบบใหม่ (θ) สามารถหาได้จากค่าตัวแปรของตัวแบบเดิมลบกับค่าความชันของฟังก์ชันของสูญเสีย $(\nabla_{\theta} J(\theta))$ ที่อยู่ในทิศทางตรงกันข้ามของความชันเดิมเพื่อให้ความชันของฟังก์ชันของสูญเสียอยู่ในค่าที่ต่ำที่สุด (ค่าความชันมีค่าเข้าใกล้ศูนย์) ด้วยอัตราการเรียนรู้ η ซึ่งทำให้ค่าตัวแปรของตัวแบบ (θ) มีค่าที่เหมาะสมที่สุด กล่าวอีกนัยหนึ่งการไล่ระดับความชัน (Gradient descent) คือ การลงเขาอย่างช้าเพื่อให้ไปถึงจุดที่ต่ำที่สุดซึ่งมีความชันเข้าใกล้ศูนย์ ดังรูปที่ 9



รูปที่ 9 การไล่ระดับความชัน (Gradient descent), โดย Géron, A. (2017). O'Reilly Media, Inc.

ในทางปฏิบัติ การไล่ระดับความชัน (Gradient descent) นั้นมีข้อเสียเนื่องจากว่าการไล่ระดับความชันแบบปกตินั้นจำเป็นที่ต้องใช้ชุดข้อมูลฝึก (training dataset) ทั้งหมดในการคำนวณหาความชันซึ่งทำให้การคำนวณนั้นช้า และการไล่ระดับความชันนำทางได้ผิดพลาด เนื่องจากฟังก์ชันการสูญเสียมักมีจุดยอด (vertex) มากกว่า 1 จุด ซึ่งมีค่าความชันเข้าใกล้ศูนย์และเป็นจุดต่ำสุดเฉพาะที่ (local minimum) จึงเป็นเรื่องท้าทายในการไล่ระดับความชันแบบปกติได้ดังรูปที่ 10 จึงใช้การไล่ระดับความชันแบบสุ่ม (Stochastic Gradient Descent) เกิดขึ้นเพื่อแก้ไขปัญหา



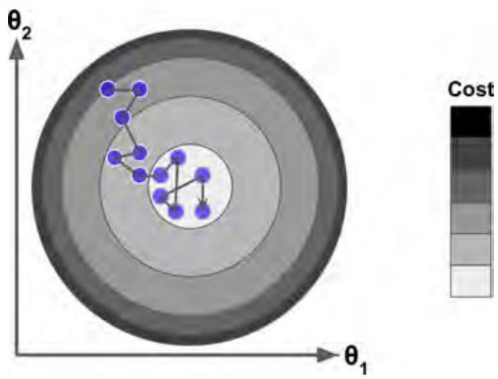
รูปที่ 10 ปัญหาของการไล่ระดับความชัน (Gradient descent), โดย Géron, A. (2017). O'Reilly Media, Inc.

การไล่ระดับความชันแบบสุ่ม (Stochastic Gradient descent) คือ การหยิบสุ่มข้อมูลมาบางส่วนและคำนวณหาค่าความชัน จึงทำให้วิธีนี้ใช้เวลาน้อยกว่า การไล่ระดับความชัน (Gradient descent) ลดความชันเพื่อไปหาจุดต่ำสุด ด้วยการไล่ระดับความชันแบบสุ่มนี้ค่าของฟังก์ชันการสูญเสียขึ้นและลงแบบไม่สม่ำเสมอ แต่ลดลงโดยเฉลี่ย ดังรูปที่ 11 เมื่อจบกระบวนการไล่ระดับ ค่าตัวแปรสุดท้ายมีค่าอยู่ในระดับที่ดี แต่ว่าไม่ได้อยู่ในระดับดีที่สุด (optimal) เนื่องจากวิธีการนี้เป็นการหยิบข้อมูลบางชุดมาแบบสุ่ม จึงทำให้วิธีการนี้มีโอกาสออกจากจุดต่ำสุดเฉพาะที่ได้ดีว่าการไล่ระดับความชัน (Gradient descent) แบบปกติ และในแต่ละการวนตัวแปรของตัวแบบถูกปรับปรุงดังสมการที่ 16

$$\theta = \theta - \eta \cdot \nabla_{\theta} J(\theta; x^{(i)}; y^{(i)}) \tag{16}$$

โดยที่ $J(\theta; x^{(i)}; y^{(i)})$ คือ ฟังก์ชันการสูญเสีย ที่เกิดจากการสุ่มค่า x และค่า y มาจากชุดข้อมูล

$\nabla_{\theta} J(\theta; x^{(i)}; y^{(i)})$ คือ ค่าความชันของฟังก์ชันการสูญเสียที่เกิดจากการสุ่มค่า x และค่า y มาจากชุดข้อมูล



รูปที่ 11 การไล่ระดับความชันแบบสุ่ม (Stochastic Gradient descent), โดย Géron, A. (2017). O'Reilly Media, Inc.

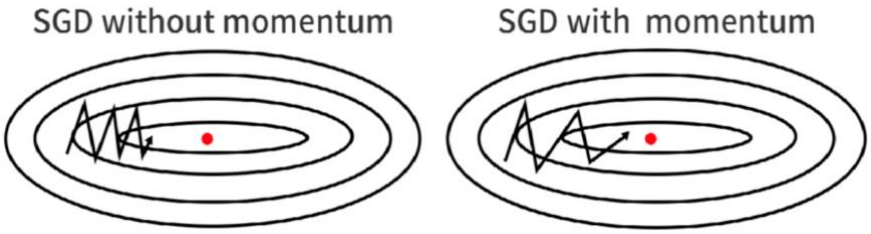
การไล่ความชันแบบสุ่มแบบปกติจะมีปัญหาด้านความเร็วในการหาจุดต่ำสุดของสมการสูญเสียซึ่งจะถูกแก้ไขด้วยวิธีการเสริมโมเมนตัม (Momentum) เนื่องจากการไล่ความชันแบบสุ่มด้วยโมเมนตัมจะช่วยให้กระบวนการไล่ความชันนั้นสามารถกระโดดไปยังจุดต่ำสุดได้ดังรูปที่ 12 และการปรับปรุงในแต่ละขั้นตอนของการไล่ความชันแบบสุ่มแบบมีโมเมนตัมจะต่างจากการไล่ความชันแบบสุ่มแบบปกติด้วยการเพิ่มตัวแปร γ เพื่อช่วยในการวนตัวแปรในการปรับปรุงดังสมการที่ 17 และสมการที่ 18

$$v_t = \gamma v_{t-1} + \eta \nabla_\theta J(\theta) \tag{17}$$

$$\theta = \theta - v_t \tag{18}$$

โดยที่ v_t คือ ค่าเฉลี่ยเอ็กซ์โพเนนเชียลของความชัน (Exponential average of gradients) ณ เวลา t

γ คือ สัมประสิทธิ์ของโมเมนตัมซึ่งเป็นตัวกำหนดสัดส่วนของค่าเฉลี่ยเอ็กซ์โพเนนเชียลของความชัน ณ เวลา t-1



รูปที่ 12 เปรียบเทียบระหว่างการไล่ความชันแบบสุ่มแบบที่ไม่มีโมเมนตัมและมีโมเมนตัม, โดย Juan Du, (2019). J. Phys

2.4.2 การถ่ายถอดรากกำลังสองเฉลี่ย (Root Mean Square Propagation)

การถ่ายถอดรากกำลังสองเฉลี่ย (Root Mean Square Propagation) หรือ RMSProp จะมีคุณสมบัติคล้ายคลึงกับการปรับปรุงการไล่ความชันแบบสุ่มด้วยโมเมนตัม แต่แตกต่างกันเนื่องจากจะใช้วิธีการปรับปรุงค่าตัวแปรด้วยค่าเฉลี่ยเคลื่อนที่ถ่วงน้ำหนักแบบทวีคูณกำลังสองของค่าความชัน (Exponential average of squares of gradients) โดยที่วิธีการนี้จะช่วยจำกัดความแปรผันในทิศที่ตั้งฉากกับทิศทางที่ไปสู่จุดต่ำสุดของสมการสูญเสีย ในขณะที่วิธีการปรับปรุงการปรับปรุงการไล่ความชันแบบสุ่มด้วยโมเมนตัมสนใจในการสร้างแรงที่มุ่งไปสู่ทิศต่ำสุดของสมการสูญเสียอย่างเดียว โดยในแต่ละการวนตัวแปรของตัวแบบถูกปรับปรุงดังสมการที่ 19 และสมการที่ 20

$$v_t = \rho v_{t-1} - (1-\rho) g_t^2 \tag{19}$$

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{v_t + \epsilon}} g_t \tag{20}$$

โดยที่ θ_t คือ ค่าตัวแปรของตัวแบบ (ค่าน้ำหนัก)

η คือ อัตราการเรียนรู้เริ่มต้น

ρ คือ ค่าน้ำหนักของแต่ละการปรับปรุงของค่าความชัน โดยปกติมีค่าเท่ากับ 0.9

v_t คือ ค่าเฉลี่ยเคลื่อนที่ถ่วงน้ำหนักแบบทวีคูณกำลังสองของค่าความชัน ณ เวลา t

g_t คือ ค่าความชันของค่าน้ำหนัก ณ เวลา t

ϵ คือ ค่าที่ทำให้แน่ใจว่าในสมการที่ 18 ไม่ถูกหารด้วย 0 โดยปกติมีค่าเท่ากับ 1e-10

2.4.3 การประมาณค่าช่วงเวลาปรับตัว (Adaptive Moment Estimation)

การประมาณค่าช่วงเวลาปรับตัว (Adaptive Moment Estimation) หรือ Adam เป็นการรวบรวมความคิดของการปรับปรุงโมเมนตัม (Momentum optimization) และการถ่ายถอดรากกำลังสองเฉลี่ย (Root Mean Square Propagation) โดยที่มีใช้ค่าเฉลี่ยการสลายตัวแบบทวีคูณของการระดับความชันก่อนหน้า (Exponentially decaying average of past gradients) เหมือนการปรับปรุงโมเมนตัม และใช้ค่าเฉลี่ยการสลายตัวแบบทวีคูณของความชันกำลังสองก่อนหน้า (exponentially decaying average of past squared gradients) เหมือนกับการถ่ายถอดรากกำลังสองเฉลี่ย

$$m_t = \beta_1 m_{t-1} - (1-\beta_1) g_t \tag{21}$$

$$v_t = \beta_2 v_{t-1} - (1-\beta_2) g_t^2 \tag{22}$$

โดยที่ v_t คือ ค่าเฉลี่ยเคลื่อนที่ถ่วงน้ำหนักแบบทวีคูณกำลังสองของค่าความชัน ณ เวลา t

m_t คือ ค่าเฉลี่ยเคลื่อนที่ถ่วงน้ำหนักของค่าความชัน ณ เวลา t

β_1 คือ ค่าน้ำหนักของแต่ละการปรับปรุงค่าความชัน โดยปกติมีค่าเท่ากับ 0.9

β_2 คือ ค่าน้ำหนักของแต่ละการปรับปรุงของค่าความชันกำลังสอง โดยปกติมีค่าเท่ากับ 0.999

ค่า m_t และค่า v_t คือค่าที่ถูกคำนวณมากจากสมการที่ 21 และ 22 ตามลำดับ โดยการรับมือค่าอคติของตัวแบบโครงข่ายประสาทด้วยการทำให้ค่าอคติดังกล่าวมีค่าเข้าใกล้ศูนย์เมื่อค่าอัตราการเสื่อมสลาย (decay rate) มีค่าที่เล็ก (β_1 และ β_2 มีค่าเข้าใกล้ 1) ดังสมการที่ 23 และ 24

$$\widehat{m}_t = \frac{m_t}{1-\beta_1^t} \tag{23}$$

$$\hat{v}_t = \frac{v_t}{1-\beta_2^t} \tag{24}$$

จากนั้นนำค่าที่ได้จากสมการที่ 23 และ 24 มาปรับปรุงค่าน้ำหนักดังสมการที่ 25

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\hat{v}_t+\epsilon}}\widehat{m}_t \tag{25}$$

บทที่ 3 วิธีการดำเนินการวิจัย

ขั้นตอนการดำเนินงานในการการทำนายความต้องการของจุดให้บริการใหม่จากข้อมูลของจุดให้บริการที่มีอยู่นั้นมีขั้นตอนการทำงานดังนี้

3.1 การศึกษาระบบดั้งเดิมและวิธีแก้ไขปัญหา

3.2 เก็บข้อมูล (Collect data)

3.3 ทำความสะอาดข้อมูล (Clean data)

3.4 สร้างตัวแบบ (Build model)

3.5 การปรับแต่งตัวแบบ (Model tuning)

3.6 การทำนาย (Prediction)

3.7 การสร้างให้เห็นภาพ (Visualization)

3.1 การศึกษาระบบดั้งเดิมและวิธีการแก้ไขปัญหา

3.1.1 ศึกษาระบบการดำเนินงานติดตั้งจุดให้บริการใหม่จากฝ่ายที่เกี่ยวข้อง

ศึกษาระบบการดำเนินงานติดตั้งจุดให้บริการใหม่จากฝ่ายที่เกี่ยวข้องกับการติดตั้งจุดให้บริการใหม่ พบว่าระบบที่ดำเนินการอยู่ในปัจจุบันนั้นมีปัญหาดังนี้

1. จำเป็นที่ต้องประชุมภายในฝ่าย และฝ่ายอื่นอย่างต่อเนื่องเพื่อกำหนดสถานที่ที่ติดตั้งในอนาคต
2. การคัดเลือกสถานที่ติดตั้งจุดให้บริการใหม่นั้นยังไม่สามารถทำนายผลที่เกิดขึ้นในอนาคตได้

3.1.2 ศึกษาภาษาเขียนโปรแกรม และหน่วยที่นำมาประกอบ

ด้วยการแก้ไขปัญหาทั้ง 2 ดังที่กล่าวไว้ในข้อ 3.1.1 นั้นจึงได้การคิดการแก้ไขปัญหาได้ดังนี้

การพัฒนาออกแบบตัวแบบเพื่อใช้ในการทำนายความต้องการที่เกิดขึ้นกับจุดให้บริการใหม่

การพัฒนาออกแบบตัวแบบเพื่อใช้ในการทำนายความต้องที่เกิดขึ้นกับจุดให้บริการใหม่ มีข้อดีเนื่องจากว่าเป็นการคำนวณที่เกิดจากการยึดข้อมูลย้อนหลัง จึงทำให้ไม่เกิดอคติในการคัดเลือกสถานที่ที่ติดตั้งในอนาคต และยังสามารถเป็นการระบุพื้นที่ในเหมาะสมที่สุดในการระบุพื้นที่ที่ติดตั้งจุดให้บริการใหม่ด้วยวิธีการวิศวกรรมย้อนกลับ โดยใช้ภาษาเขียนโปรแกรม Python แม้ว่าภาษา Python นั้นได้รับการเปรียบเทียบกับภาษา R ในด้านความเร็วในการประมวลผลของตัวเลขขนาดใหญ่ แต่เนื่องจากว่าภาษา Python นั้นมีการเขียนภาษาที่ชัดเจน สะอาดตา สามารถเรียนรู้เข้าใจได้เร็ว และมีกลุ่มนักพัฒนาหน่วยที่นำมาประกอบทางด้านวิทยาศาสตร์ข้อมูลอย่างต่อเนื่อง จึงเป็นภาษาเขียนโปรแกรมที่ได้เปรียบในระยะยาว

3.2 เก็บข้อมูล (Collect data)

ข้อมูลที่นำมาใช้ในโครงงานนี้ เก็บด้วยการบันทึกการเปิดใช้งานแอพพลิเคชันการเช่ารถยนต์ หรือการใช้บริการของลูกค้าผ่านแอพพลิเคชันของบริษัทกรณีศึกษา และเก็บข้อมูลเชิงพื้นที่จาก openstreetmap เป็นแผนที่ที่เปิดให้เข้าไปดึงข้อมูลออกมาใช้งานได้

3.2.1 เก็บข้อมูลการจองรถเช่าจากแอปพลิเคชัน

การเก็บข้อมูลการจองรถเช่าที่ได้จากแอปพลิเคชันนั้นได้ถูกพัฒนาโดยฝ่ายพัฒนาแอปพลิเคชันของบริษัท ซึ่งได้มีการเก็บข้อมูลไว้หลายส่วนด้วยกัน เช่น ระยะเวลาในการจอง ช่วงเวลาที่คืนรถ ระยะการเดินทาง เป็นต้น จึงทำให้เกิดการง่ายต่อผู้จัดทำในการสร้างตัวแบบเพื่อทำนายความต้องการของจุดให้บริการที่ตั้งใหม่

3.2.2 เก็บข้อมูลจาก openstreetmap

การเก็บข้อมูลจาก openstreetmap นั้นเริ่มจากการติดตั้งหน่วยที่นำมาประกอบ osmnx ลงบนเครื่องคอมพิวเตอร์เพื่อเป็นตัวกลางในการดึงข้อมูลจากเว็บไซต์ openstreetmap.org โดยการดึงนี้ดึงข้อมูลที่จำเป็นคือตำแหน่งและประเภทของจุดที่สนใจมาทั้งหมด 54,857 จุดโดยเก็บครอบคลุมจุดให้บริการ ดังตารางที่ 2

ตารางที่ 2 ตำแหน่งของจุดที่สนใจ (Points of interest)

ประเภท	ชื่อ	ละติจูด	ลองจิจูด
restaurant	Manna	13.7477023	100.5345542
school	Australian International School	13.7281689	100.5637277
restaurant	ฮาร์ดร็อค คาเฟ่ กรุงเทพ	13.7444202	100.5318788
restaurant	Zen	13.7281063	100.5331414
restaurant	Coyote	13.7279705	100.5332622
cafe	Starbucks Coffee	13.7275459	100.5334533
bar	Molly Malone's	13.7274447	100.5334988
cafe	Pola Pola	13.7272399	100.5335418
restaurant	Three on Convent	13.7271473	100.5341514
restaurant	The Papaya	13.7277058	100.5354816
cafe	Cafe Bangrak	13.7289158	100.5355833
cafe	Bell Cafe	13.6513812	100.4869036
restaurant	Smith’s Bar	13.7585554	100.4959363
hospital	โรงพยาบาลธนบุรี	13.7529428	100.480116
university	คณะวิศวกรรมศาสตร์	13.7365936	100.5330593

3.2.3 เก็บข้อมูลจำนวนรถยนต์ในแต่ละจุดให้บริการรายวันจากแอปพลิเคชัน

การเก็บข้อมูลจำนวนรถยนต์ในแต่ละจุดให้บริการรายวันจากแอปพลิเคชันนั้นถูกโดยฝ่ายพัฒนาแอปพลิเคชันของบริษัท ซึ่งข้อมูลดังกล่าวถูกเก็บไว้ในฐานข้อมูลของบริษัท โดยสามารถใช้งานข้อมูลด้วยการใช้ภาษา SQL เพื่อดึงออกมาจากฐานข้อมูล

3.2.4 เก็บข้อมูลการเปิด-ปิดจุดให้บริการ

การเก็บข้อมูลนี้ได้จากข้อมูลที่บันทึกไว้ภายบริษัท โดยเป็นข้อมูลที่บ่งบอกถึงเวลาที่ได้เปิดให้บริการและปิดให้บริการของจุดให้บริการต่าง ๆ โดยมีค่าเป็นศูนย์ เมื่อจุดให้บริการไม่เปิดให้บริการ แต่มีค่าเป็นหนึ่ง เมื่อจุดให้บริการเปิดให้บริการ ดังตารางที่ 3

ตารางที่ 3 ตัวอย่างข้อมูลการเปิด-ปิดจุดให้บริการ

เวลา	จุดให้บริการที่ 1	จุดให้บริการที่ 2	จุดให้บริการที่ 3
8/01/2562	0	1	1
9/01/2562	0	1	1
10/01/2562	0	1	1
11/01/2562	0	1	1
12/01/2562	0	1	1
13/01/2562	0	1	0
14/01/2562	0	1	0
15/01/2562	0	1	0
16/01/2562	1	1	0
17/01/2562	1	0	0
18/01/2562	1	0	1
19/01/2562	1	0	1

จากตารางที่ 3 เห็นได้ว่าจุดให้บริการที่ 1 ไม่เปิดให้บริการจนกระทั่งวันที่ 16 มกราคม พ.ศ. 2562 จึงเป็นวันแรกที่จุดให้บริการที่ 1 เปิดให้บริการ โดยที่จุดให้บริการที่ 2 เปิดตั้งแต่วันที่ 8 มกราคม พ.ศ. 2562 จนกระทั่งถึงวันที่ 17 มกราคม พ.ศ. 2562 จึงได้ปิดให้บริการ และจุดให้บริการที่ 3 เปิดตั้งแต่วันที่ 8 มกราคม พ.ศ. 2562 และปิดให้บริการวันที่ 13 มกราคม พ.ศ. 2562 โดยกลับมาเปิดให้บริการอีกครั้งในวันที่ 18 มกราคม พ.ศ. 2562

3.2.5 เก็บข้อมูลเหตุการณ์วันหยุดในประเทศไทย

การเก็บข้อมูลเหตุการณ์วันหยุดในประเทศไทยอ้างอิงจากวันหยุดตามประเพณีของสถาบันทางการเงินจากธนาคารแห่งประเทศไทย

3.3 ทำความสะอาดข้อมูล (Clean data)

การทำความสะอาดข้อมูล (Clean data) เริ่มจากการนำชุดข้อมูลดิบที่ได้จากการเก็บข้อมูลมาพิจารณาถึงส่วนที่จำเป็นต่อการใช้งานและไม่ใช้ออกมาจากกัน โดยเริ่มจากการนำข้อมูลมากำจัดข้อมูลส่วนเกินหรือข้อมูลซ้ำออกมาจากชุดข้อมูล เพื่อนำข้อมูลที่มีความสะอาดสูงมีความแปรปรวนต่ำ และข้อมูลสะท้อนถึงความเป็นจริงมากที่สุดสามารถนำไปพัฒนาตัวแบบได้อย่างมีประสิทธิภาพ โดยมีลำดับขั้นตอนการทำความสะอาดข้อมูลดังต่อไปนี้

1. การหาค่าระยะการกระจัดของแต่ละจุดให้บริการต่อจุดให้บริการอื่น

   การหาระยะการกระจัดของแต่ละจุดให้บริการต่อจุดให้บริการอื่นนั้นใช้ค่าละติจูด ลองจิจูดของแต่ละจุดให้บริการเข้ามาเปรียบเทียบกับเพื่อคำนวณหาระยะการกระจัดออกมาเป็นระยะกระจัด

2. การหาจำนวนจุดที่สนใจที่อยู่ใกล้เคียงกับจุดให้บริการ

   การหาจำนวนจุดสนใจที่อยู่ใกล้เคียงกับจุดให้บริการนั้นเริ่มจากการหาระยะห่างของจุดให้บริการกับจุดที่สนใจรายจุด และบันทึกเพื่อเก็บข้อมูลเอาไว้ โดยที่คำนวณระยะห่างระหว่างจุดใกล้เคียงกับจุดให้บริการนั้นอยู่ใกล้เคียงไม่เกิน 2 กิโลเมตร และแบ่งแยกประเภทของจุดสนใจนั้นเอาไว้ ดังตารางที่ 4

ตารางที่ 4 จำนวนจุดที่สนใจที่อยู่ใกล้เคียงกับจุดให้บริการ

ประเภท	จุดให้บริการที่ 1	จุดให้บริการที่ 2	จุดให้บริการที่ 3	จุดให้บริการที่ 4
Cafe	244	138	299	20
Police	18	9	19	1
Bus station	13	13	19	0
Bank	159	116	207	8
Clinic	19	2	10	0
School	57	32	68	16
Hospital	33	13	25	6
University	29	6	25	1
Restaurant	619	374	750	78
Marketplace	22	28	33	4
Fast food	100	38	97	1
Food court	31	5	29	1
College	3	2	3	0
Bar	159	34	149	2
pharmacy	38	15	38	16
Veterinary	1	0	1	3
Convenience	0	0	0	0

3. การหาความคล้ายคลึงโคไซน์เทียบกับของจุดสนใจแต่ละจุดให้บริการ (pair-wise station cosine similarity)

การคำนวณค่าความคล้ายคลึงโคไซน์ (cosine similarity) สามารถคำนวณด้วยหน่วยที่นำมาประกอบสำเร็จรูปจาก sci-kit learn ซึ่งได้มีการหาค่าความเหมือนของแต่ละจุดให้บริการจากจำนวนจุดที่สนใจด้วยการแบ่งแต่ละจุดให้บริการเป็นเวกเตอร์และคำนวณหาความคล้ายคลึงโคไซน์

4. การสร้างกราฟค่าน้ำหนักระยะการกระจัดของแต่ละจุดให้บริการต่อจุดให้บริการอื่น

การสร้างกราฟค่าน้ำหนักระยะการกระจัดของแต่ละจุดให้บริการต่อจุดให้บริการอื่นได้นำกฎข้อแรกทางภูมิศาสตร์ของโทเบลอ (Tobler's first law of geography) ว่า ทุกสิ่งล้วนสัมพันธ์กับสิ่งอื่น แต่สิ่งที่ใกล้กันมีความสัมพันธ์มากกว่าสิ่งที่อยู่ห่างออกไป (everything is related to everything else, but near things are more related than distant things.) (Di Chai, 2009) มาเป็นต้นแบบในการสร้างกราฟค่าน้ำหนักอยู่ในรูปแบบแมทริกซ์ประชิดดังสมการที่ 26 โดยที่ค่าน้ำหนักของกราฟนี้มีค่าเท่ากับส่วนกลับของระยะการกระจัด

$$A = \begin{pmatrix} 0 & \frac{1}{\text{ระยะการกระจัด}_{0,1}} & \cdots & \frac{1}{\text{ระยะการกระจัด}_{0,N-1}} \\ \frac{1}{\text{ระยะการกระจัด}_{1,0}} & 0 & \cdots & \frac{1}{\text{ระยะการกระจัด}_{1,N-1}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{\text{ระยะการกระจัด}_{N-1,0}} & \cdots & \cdots & 0 \end{pmatrix} \tag{26}$$

โดยที่ A คือ แมทริกซ์ประชิด

N คือ จำนวนจุดให้บริการ

5. การสร้างกราฟความคล้ายคลึงของแต่ละจุดให้บริการต่อจุดให้บริการอื่น

นำข้อมูลความคล้ายคลึงที่ได้จากการคำนวณมาสร้างเป็นกราฟโดยให้อยู่ในรูปแบบเมทริกซ์ประชิด ดังสมการที่ 27

$$A = \begin{pmatrix} 0 & r_{0,1} & \cdots & r_{0,N-1} \\ r_{1,0} & 0 & \vdots & r_{1,N-1} \\ \vdots & \cdots & \ddots & \vdots \\ r_{N-1,0} & \cdots & \cdots & 0 \end{pmatrix} \tag{27}$$

โดยที่ r คือ ค่าความคล้ายคลึง

6. การคำนวณกราฟผสม (Graph fusion)

กราฟผสม (Graph fusion) ถูกสร้างขึ้นด้วยการรวมกราฟค่าน้ำหนักระยะการกระจัดกับกราฟความคล้ายคลึงด้วยการทำให้กราฟทั้งสองเป็นมาตรฐาน (normalize) (Di Chai, 2009) โดยนำกราฟประชิดคูณกับเมทริกซ์ระดับแนวทแยง (Diagonal degree matrix) และบวกด้วยด้วยเมทริกซ์เอกลักษณ์ (Identity matrix) ดังสมการที่ 28 และสมการที่ 29

$$A' = D^{-1}A + I \tag{28}$$

$$D^{-1} = \begin{pmatrix} \sum_{j=0}^{N-1} A_{0,j} & 0 & \cdots & 0 \\ 0 & \sum_{j=0}^{N-1} A_{1,j} & \vdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sum_{j=0}^{N-1} A_{N-1,j} \end{pmatrix} \tag{29}$$

โดยที่ A' คือ เมทริกซ์ประชิดที่เป็นมาตรฐาน

D^{-1} คือ เมทริกซ์ระดับแนวทแยง

I คือ เมทริกซ์เอกลักษณ์

ค่าเมทริกซ์ประชิดที่เป็นมาตรฐานที่ได้จากสมการที่ 29 ถูกมาคำนวณต่อเพื่อให้ได้กราฟผสมด้วยการบวกกราฟเข้าด้วยกันทั้งหมดดังสมการที่ 30

$$\boldsymbol{F} = \sum_{\boldsymbol{i=1}}^{\boldsymbol{K}} \boldsymbol{A'_i} \tag{30}$$

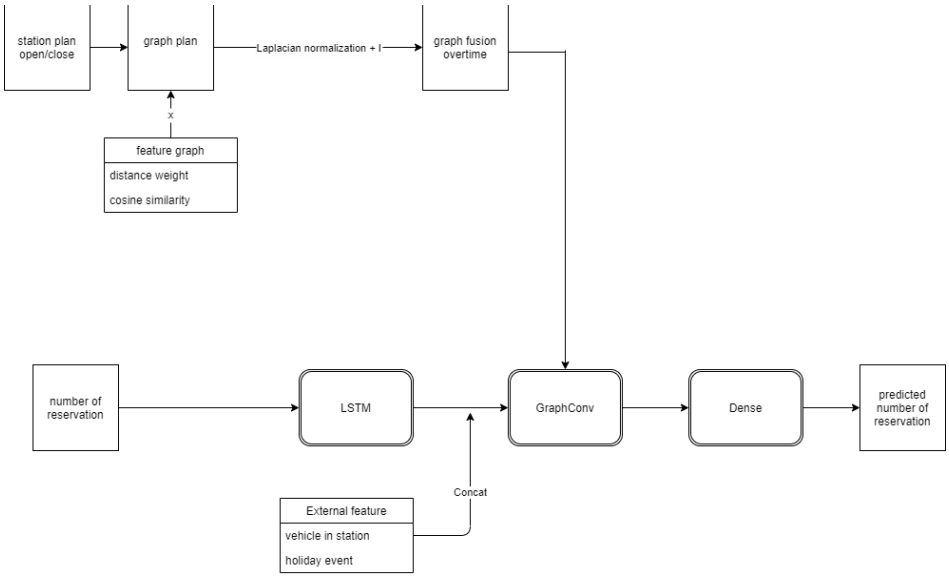
โดยที่ F คือ กราฟผสม

K คือ จำนวนเมทริกซ์ประชิดที่เป็นมาตรฐาน

3.4 สร้างและฝึกฝนตัวแบบ (Build and train model)

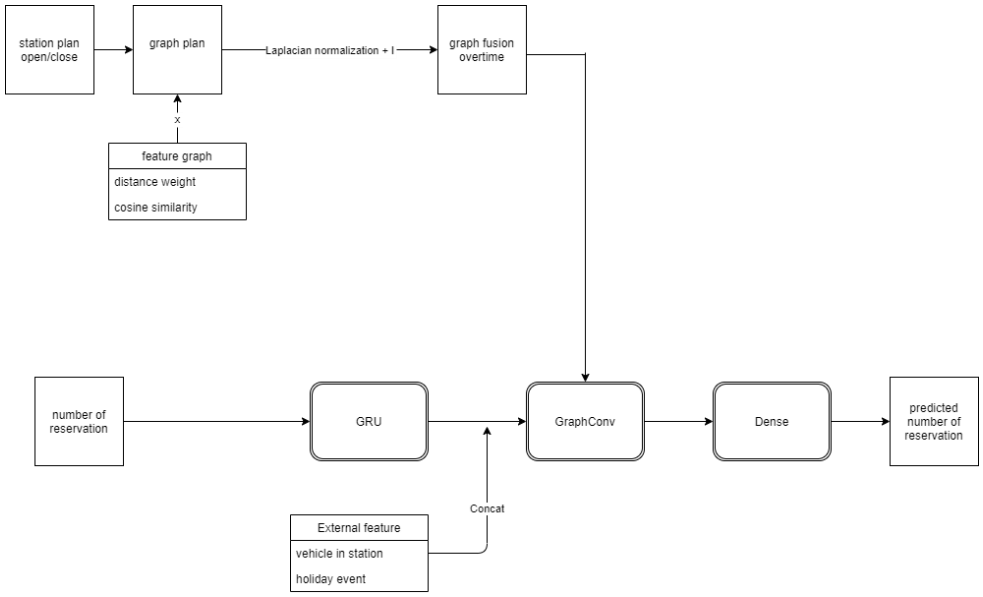
การสร้างตัวแบบ (Build model) เริ่มจากการพัฒนารูปแบบของตัวแบบจำนวน 2 ตัวแบบเพื่อใช้เปรียบเทียบประสิทธิภาพการทำนาย

1. กระบวนการสร้างตัวแบบระหว่างตัวแบบความจำระยะสั้น และระยะยาว และเสริมด้วยโครงข่ายประสาทการแปลงกราฟ (LSTM and Graph convolutional neural network model) (Luo, 2020) ด้วยข้อมูลที่ผ่านการทำความสะอาดแล้ว โดยจัดให้ข้อมูลจำนวนการใช้ในแต่ละจุดให้บริการรายวันเป็นข้อมูลป้อนเข้า (input data) ไปยังตัวชั้นผิวตัวแบบความจำระยะสั้น และระยะยาว โดยที่ผลลัพธ์ที่ออกมาถูกเชื่อม (Concatenate) ด้วยคุณสมบัติภายนอก (External feature) ที่ประกอบไปด้วยจำนวนรถยนต์ในจุดให้บริการและสถานะวันหยุด จากนั้นส่งค่าดังกล่าวไปยังชั้นผิวโครงข่ายประสาทการแปลงกราฟโดยมีกราฟผสมตามรายวัน (Graph fusion overtime) ที่ถูกคำนวณจากข้อมูลการเปิด-ปิดของจุดให้บริการในแต่ละวันคูณกับกราฟคุณสมบัติ (feature graph) เป็นกราฟกำกับ และผลลัพธ์ของชั้นผิวโครงข่ายประสาทการแปลงกราฟถูกส่งต่อไปยังชั้นผิวบีบแน่น (Dense layer) เพื่อให้ตัวแบบทำนายผลลัพธ์ออกมา ดังรูปที่ 13 โดยกำหนดให้ฟังก์ชันการสูญหายเป็นค่าคลาดเคลื่อนกำลังสองเฉลี่ย



รูปที่ 13 การสร้างตัวแบบด้วยกระบวนการ LSTM และ graph convolution neural network

2. กระบวนการสร้างตัวแบบระหว่างหน่วยเวียนกลับแบบมีประตูและเสริมด้วยโครงข่ายประสาทการแปลงกราฟ (GRU and Graph convolutional neural network model) โดยที่กระบวนการนี้ชั้นผิวหน่วยเวียนกลับแบบมีประตูแทนการใช้ชั้นผิวความจำระยะสั้น ระยะยาว และใช้ข้อมูลแบบเดียวกับกระบวนการแรก ดังรูปที่ 14



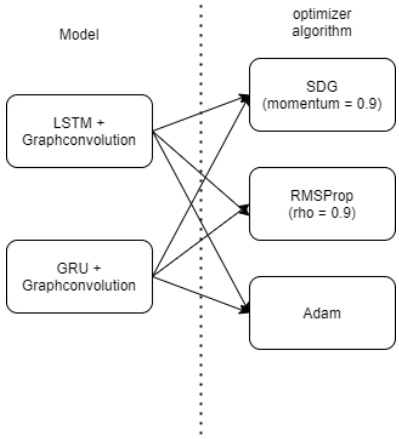
รูปที่ 14 การสร้างตัวแบบด้วยกระบวนการ GRU และ graph convolution neural network

เนื่องจากจำนวนการใช้งานของจุดให้บริการที่น้อย รูปแบบการเช่ารถยนต์ที่มีทั้งรูปแบบรายวันและระยะยาว ทำให้เป็นไปได้ยากในการสะท้อนความจริงออกมาผ่านข้อมูล จึงต้องมีสมมติฐานในการสะท้อนความจริงผ่านข้อมูลดังนี้

1. กำหนดรูปแบบการเช่ารถยนต์รายเดือนให้เป็นแบบรายวัน โดยกระจายรูปแบบการเช่าระยะให้อยู่ในรูปแบบการเช่ารายวัน โดยที่การเช่านี้เป็นการเช่าแบบรายวัน
2. เนื่องจากจำนวนการใช้งานที่น้อย จึงกำหนดให้รวบจำนวนการจองรายวันเป็นจำนวนการจองรายสัปดาห์
3. กำหนดให้ช่วงข้อมูลที่ใช้ดำเนินงานอยู่ในช่วงวันที่ 8 มกราคม พ.ศ. 2562 ถึง วันที่ 8 สิงหาคม พ.ศ. 2564
4. กำหนดจุดให้บริการที่ใช้เพื่อการฝึกฝนตัวแบบตั้งอยู่ที่กรุงเทพมหานครและปริมณฑลเท่านั้น
5. กำหนดให้สภาพอากาศไม่มีผลต่อเชิงลบ หรือเชิงบวกต่อของจำนวนการจอง

3.5 การปรับแต่งตัวแบบ (Model tuning)

การปรับแต่งตัวแบบ (Model tuning) เริ่มจากการปรับค่าตัวแปรหลังจากการสร้างตัวแบบทั้งสองด้วยการปรับปรุงค่าตัวแปรที่อยู่ในโมเดล คือ วิธีการปรับปรุง (optimization algorithm) ได้แก่ การไล่ความชันแบบสุ่มด้วยค่า โดยกำหนดให้ค่าโมเมนตัมมีค่าเท่ากับ 0.9 (Brownlee, 2019) และการถ่ายถอดรากกำลังสองเฉลี่ย โดยกำหนดให้ค่าน้ำหนักของแต่ละการปรับปรุงของค่าความชัน (ρ) มีค่าเท่ากับ 0.9 (Brownlee, 2019) และการประมาณค่าช่วงเวลาปรับตัว โดยกำหนดให้อัตราการเรียนรู้ (learning rate) ของทุกวิธีการปรับปรุงมีค่าเท่ากับ 0.001 (Luo, 2020) โดยสามารถเปรียบเทียบประสิทธิภาพของตัวแบบได้ทั้งหมด 6 กรณี ดังรูปที่ 15 โดยชุดข้อมูลที่ใช้ถูกแบ่งออกเป็น 2 แบบ คือ ชุดข้อมูลฝึกฝน (training dataset) ร้อยละ 80 และชุดข้อมูลทดสอบ (test dataset) ร้อยละ 20 เพื่อใช้วัดประสิทธิภาพของตัวแบบด้วยข้อมูลที่ตัวแบบไม่เคยเห็น



รูปที่ 15 การปรับแต่งตัวแบบ

3.5.1 การวัดประสิทธิภาพ

ในกรณีนี้การวัดประสิทธิภาพใช้ตัวชี้วัดทางสถิติ ได้แก่ รากที่สองของค่าเฉลี่ยความผิดพลาดกำลังสอง (Root mean square error :RMSE) และค่าเฉลี่ยความผิดพลาดสัมบูรณ์ (Mean absolute error :MAE) ดังสมการที่ 31 และสมการที่ 32

$$\boldsymbol{RMSE} = \sqrt{\frac{\mathbf{1}}{\boldsymbol{N}}\sum_{\boldsymbol{i=1}}^{\boldsymbol{N}}(\boldsymbol{y_i} - \boldsymbol{\hat{y}_i})^{\mathbf{2}}} \tag{31}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \tag{32}$$

โดยที่ N คือ จำนวนข้อมูล

y_i คือ ค่าของข้อมูลจริง (Actual value)

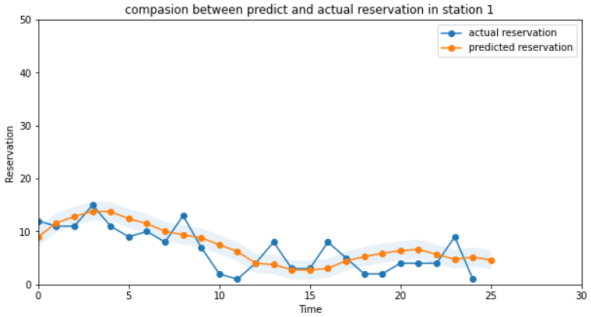
$\hat{y}_i$ คือ ค่าของข้อมูลที่ถูกทำนาย (Predicted value)

3.6 การทำนาย (Prediction)

การทำนายใช้ตัวแบบที่สมบูรณ์และมีประสิทธิภาพที่ดีที่สุดด้วยชุดข้อมูลที่เตรียมไว้เพื่อทำนายจำนวนการใช้งานของจุดบริการในระยะเวลาที่ 1 โดยที่ค่าของข้อมูลที่ถูกทำนายถูกจัดเก็บและถูกใช้เป็นชุดข้อมูลในการทำนายต่อในระยะเวลาที่ 2 การจัดเก็บแล้วการทำนายถูกทำซ้ำจนถึงระยะเวลาที่ n

3.7 การสร้างให้เห็นภาพ (Visualization)

การสร้างให้เห็นภาพการเปรียบเทียบระหว่างข้อมูลจำนวนการใช้งานที่เกิดขึ้นจริงกับจำนวนการใช้งานที่ถูกทำนายของจุดให้บริการที่ตั้งขึ้นใหม่ดังรูปที่ 16 ช่วยให้ง่ายต่อการทำความเข้าใจและสามารถเปรียบเทียบกับข้อมูลในอดีต โดยที่สามารถใช้สื่อสารกับเพื่อนร่วมงาน หรือบุคคลภายนอกได้ง่ายขึ้นด้วยการเปรียบเทียบระหว่างชุดข้อมูลที่ถูกทำนายและชุดข้อมูลจริง



รูปที่ 16 การสร้างให้เห็นภาพการเปรียบเทียบระหว่างข้อมูลจำนวนการใช้งานที่เกิดขึ้นจริงกับจำนวนการใช้งานที่ถูกทำนาย

บทที่ 4 ผลการดำเนินงาน

4.1 การปรับแต่งตัวแบบ

4.1.1 ผลลัพธ์โดยรวม

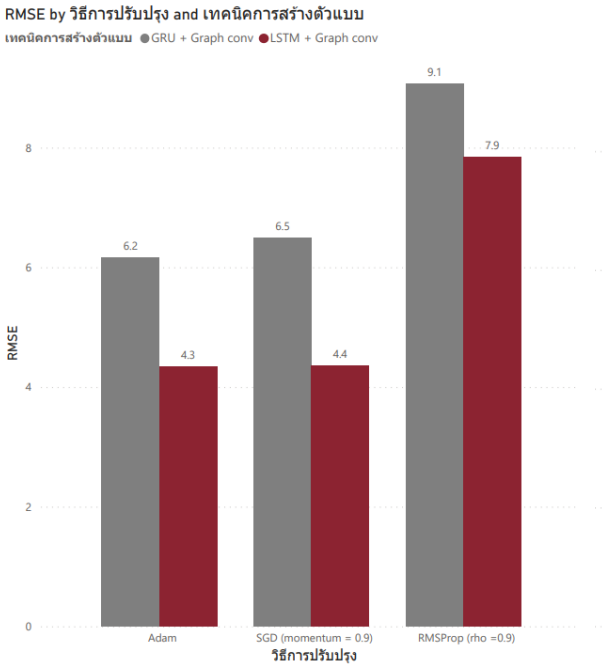
ในการปรับแต่งตัวแบบจะพิจารณาค่าเฉลี่ยของค่าเฉลี่ยความผิดพลาดกำลังสอง และค่าเฉลี่ยของค่าเฉลี่ยความผิดพลาดสัมบูรณ์ของการทำนาย 5 ครั้งเพื่อเป็นตัวบ่งชี้ประสิทธิภาพของตัวแบบจากตัวแบบทั้งหมด 6 กรณี จากตารางที่ 5 จะเห็นได้ว่าการสร้างตัวแบบด้วยตัวแบบความจำระยะสั้น และระยะยาว และเสริมด้วยโครงข่ายประสาทการแปลงกราฟ (LSTM and Graph convolutional neural network) ด้วยวิธีการปรับปรุงการประมาณค่าช่วงเวลาปรับตัว (Adaptive Moment Estimation) มีประสิทธิภาพในการทำนายสูงที่สุด เพราะความผิดพลาดในการทำนายน้อยที่สุดด้วยค่าเฉลี่ยความผิดพลาดกำลังสองเท่ากับ 4.3463 และค่าเฉลี่ยความผิดพลาดสัมบูรณ์เท่ากับ 2.5333 โดยที่การสร้างตัวแบบด้วยกระบวนการหน่วยเวียนกลับแบบมีประตูและเสริมด้วยโครงข่ายประสาทการแปลงกราฟ (GRU and Graph convolutional neural network) และวิธีการปรับปรุงการถ่ายถอดรากกำลังสองเฉลี่ย (Root Mean Square Propagation : RMSProp) มีประสิทธิภาพในการทำนายต่ำที่สุด เนื่องจากค่าความผิดพลาดในการทำนายมากที่สุดด้วยค่าเฉลี่ยความผิดพลาดกำลังสองเท่ากับ 9.0725 และค่าเฉลี่ยสัมบูรณ์เท่ากับ 4.8488

ตารางที่ 5 ผลลัพธ์จากการดำเนินงาน

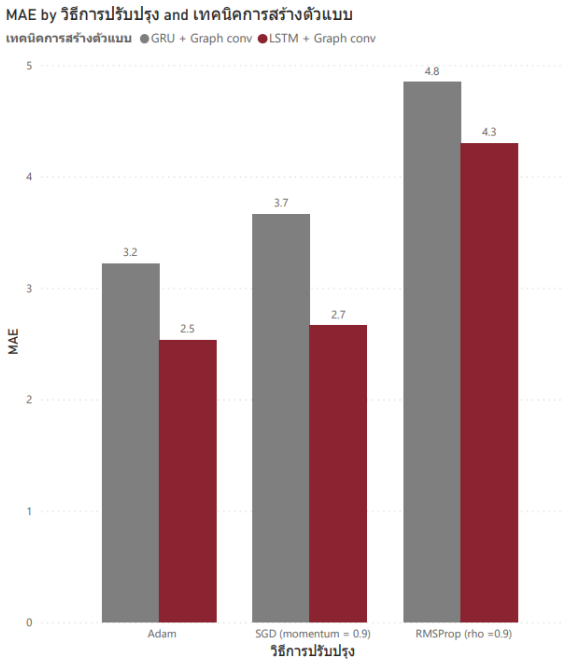
กรณีที่	เทคนิคการสร้างตัวแบบ	วิธีการปรับปรุง	RMSE	MAE
1	LSTM + Graph conv	SGD (momentum = 0.9)	4.3645	2.6625
2	LSTM + Graph conv	RMSProp (rho =0.9)	7.8517	4.3016
3	LSTM + Graph conv	Adam	4.3463	2.5333
4	GRU + Graph conv	SGD (momentum = 0.9)	6.4984	3.6656
5	GRU + Graph conv	RMSProp (rho =0.9)	9.0725	4.8488
6	GRU + Graph conv	Adam	6.1703	3.2180

4.1.2 เปรียบเทียบกระบวนการสร้างตัวแบบ

ในการปรับเปลี่ยนกระบวนการสร้างโดยเปรียบเทียบกับวิธีการปรับปรุงทั้ง 3 วิธีการปรับปรุง จะเห็นได้ว่าการใช้วิธีการสร้างตัวแบบด้วยตัวแบบความจำระยะสั้น และระยะยาว และเสริมด้วยโครงข่ายประสาทการแปลงกราฟ (LSTM and Graph convolutional neural network) จะมีค่าเฉลี่ยความผิดพลาดกำลังสองและค่าเฉลี่ยความผิดพลาดสัมบูรณ์ที่ลดลงจากการสร้างตัวแบบด้วยกระบวนการหน่วยเวียนกลับแบบมีประตูและเสริมด้วยโครงข่ายประสาทการแปลงกราฟ (GRU and Graph convolutional neural network) ในทุกวิธีการปรับปรุงดังรูปที่ 17 และรูปที่ 18



รูปที่ 17 เปรียบเทียบประสิทธิภาพในการทำนายของตัวแบบด้วยค่าความผิดพลาดกำลังสอง ตามกระบวนการสร้างตัวแบบ



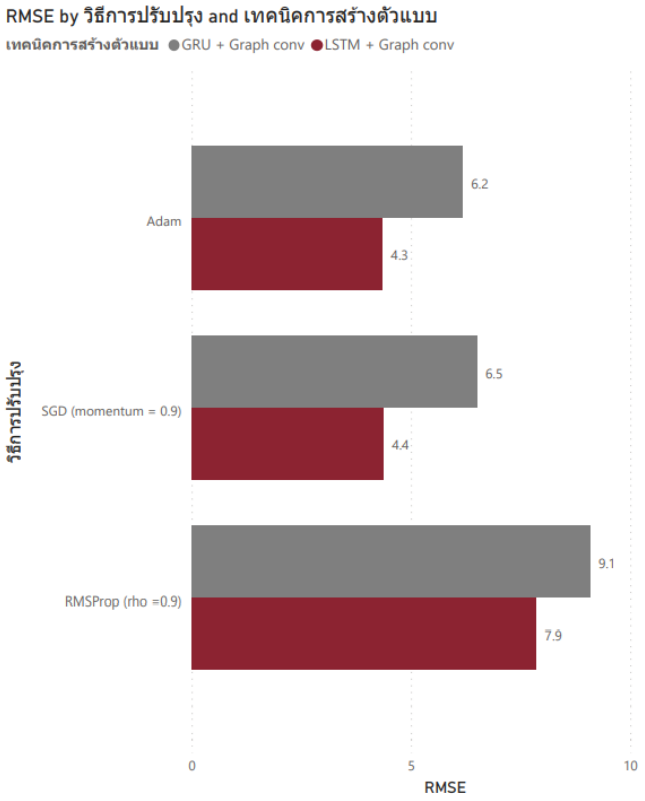
รูปที่ 18 เปรียบเทียบประสิทธิภาพในการทำนายของตัวแบบด้วยค่าเฉลี่ยความผิดพลาดสัมบูรณ์ ตามกระบวนการสร้างตัวแบบ

จากรูปที่ 17 และรูปที่ 18 จะเห็นได้ว่าตัวแบบที่สร้างด้วยกระบวนการหน่วยเวียนกลับแบบมีประตูและเสริมด้วยโครงข่ายประสาทการแปลงกราฟด้วยวิธีปรับปรุงการประมาณค่าช่วงเวลาปรับตัวจะมีประสิทธิภาพในการทำนายดีที่สุด โดยที่วิธีการปรับปรุงการไล่ความชันแบบสุ่มด้วยโมเมนตัม และการถ่ายถอดรากกำลังสองเฉลี่ยมีประสิทธิภาพในการทำนายที่ลดลงตามลำดับ ในทำนองเดียวกัน ตัวแบบที่สร้างด้วยกระบวนการตัวแบบความจำระยะสั้น และระยะยาว และเสริมด้วยโครงข่ายประสาทการแปลงกราฟด้วยวิธีปรับปรุงการประมาณค่าช่วงเวลาปรับตัวจะมีประสิทธิภาพในการทำนายดีที่สุด โดยที่วิธีการปรับปรุงการไล่ความชันแบบสุ่มด้วยโมเมนตัม และการถ่ายถอดรากกำลังสองเฉลี่ยมีประสิทธิภาพในการทำนายที่ลดลงตามลำดับ โดยเหตุผลที่การสร้างตัวแบบด้วยวิธีความจำระยะสั้น ระยะยาวสามารถทำนายได้แม่นยำกว่าหน่วยเวียนกลับแบบมีประตู เนื่องจากความจำระยะสั้น ระยะยาวนั้นมีการส่งต่อความทรงจำทั้งระยะสั้น และระยะยาวตัวแบบมีมิติในการฝึกฝนตัวแปรได้มากกว่า ในขณะที่กระบวนการหน่วยเวียนกลับแบบมีประตูนั้นมีเพียงแค่ความทรงจำระยะสั้นเท่านั้น จึงทำให้

กระบวนการสร้างด้วยหน่วยเวียนกลับแบบมีประตูไม่สามารถนำข้อมูลที่เป็นความทรงจำระยะยาวมาใช้เป็นเหตุผลประกอบการตัดสินใจสำหรับการสร้างฝึกฝนตัวแบบได้เท่าที่ควร

4.1.3 เปรียบเทียบวิธีการปรับปรุง

ในการปรับเปลี่ยนวิธีการปรับปรุงทั้ง 3 วิธีโดยเปรียบเทียบกับกระบวนการสร้างตัวแบบทั้ง 2 กระบวนการจะเห็นได้ว่าวิธีปรับปรุงการประมาณค่าช่วงเวลาปรับตัวจะมีค่าความผิดพลาดน้อยที่สุดในทุก ๆ กระบวนการสร้างตัวแบบ และการปรับปรุงการไล่ความชันแบบสุ่มด้วยโมเมนตัม และการถ่ายถอดรากกำลังสองเฉลี่ยมีประสิทธิภาพในการทำนายที่ลดลงตามลำดับ ดังรูปที่ 19



รูปที่ 19 เปรียบเทียบประสิทธิภาพในการทำนายของตัวแบบด้วยค่าความผิดพลาดกำลังสอง ตามวิธีการปรับปรุง

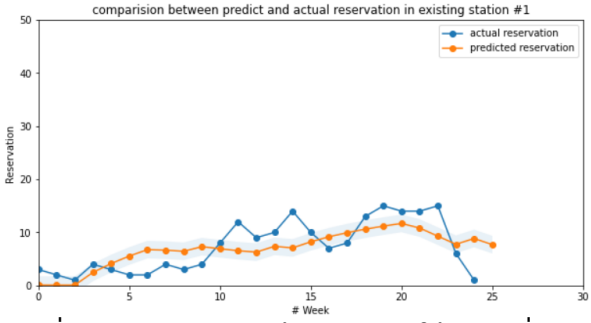
จากตารางที่ 6 จะเห็นได้ว่าวิธีการปรับปรุงการประมาณค่าช่วงเวลาปรับตัวมีค่าความผิดพลาดที่เกิดจากการทำนายใกล้เคียงกับวิธีการปรับปรุงการปรับปรุงการไล่ความชันแบบสุ่มด้วยโมเมนตัม โดยที่ตัวแบบที่สร้างโดยความจำระยะสั้น ระยะยาวและเสริมด้วยโครงข่ายประสาทการแปลงกราฟมีค่าเฉลี่ยความผิดพลาดกำลังต่างกันร้อยละ 0.42 และในทำนองเดียวกันตัวแบบที่สร้างโดยหน่วยเวียนกลับแบบมีประตูและเสริมด้วยโครงข่ายประสาทการแปลงกราฟมีค่าเฉลี่ยความผิดพลาดกำลังต่างกันร้อยละ 5.05 วิธีปรับปรุงการประมาณค่าช่วงเวลาปรับตัวมีความผิดพลาดน้อยที่สุด เพราะวิธีการดังกล่าวเป็นวิธีการที่ได้รับการปรับปรุงมาจากการปรับปรุงการไล่ความชันแบบสุ่มด้วยโมเมนตัม และการถ่ายถอดรากกำลังสองเฉลี่ย โดยเป็นการนำข้อดีของทั้งสองมาใช้ร่วมกัน ในขณะที่วิธีการปรับปรุงอื่นเป็นวิธีที่มีลักษณะเฉพาะตัวซึ่งมีข้อเสียที่แตกต่างกัน

ตารางที่ 6 เปรียบเทียบร้อยละความแตกต่างของค่าเฉลี่ยความผิดพลาดกำลังสองของทุกการทดลอง

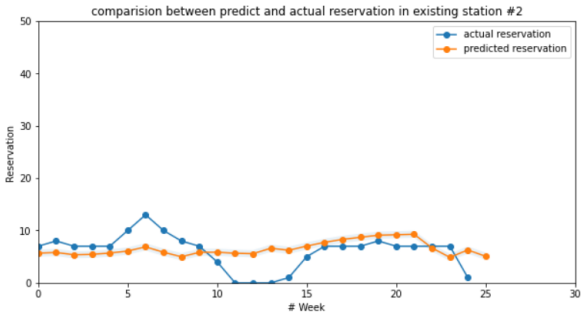
		LSTM and Graph conv			GRU and Graph conv		
		SGD (momentum = 0.9)	RMSProp (rho =0.9)	Adam	SGD (momentum = 0.9)	RMSProp (rho =0.9)	Adam
LSTM and Graph conv	SGD (momentum = 0.9)	0.00%	44.41%	0.42%	32.84%	51.89%	29.27%
	RMSProp (rho =0.9)	44.41%	0.00%	44.65%	17.24%	13.46%	21.41%
	Adam	0.42%	44.65%	0.00%	33.12%	52.09%	29.56%
GRU and Graph conv	SGD (momentum = 0.9)	32.84%	17.24%	33.12%	0.00%	28.37%	5.05%
	RMSProp (rho =0.9)	51.89%	13.46%	52.09%	28.37%	0.00%	31.99%
	Adam	29.27%	21.41%	29.56%	5.05%	31.99%	0.00%

4.2 การทำนายผลกระทบที่เกิดขึ้นบนจุดให้บริการที่มีอยู่

การทำนายผลกระทบที่เกิดขึ้นบนจุดให้บริการที่มีอยู่ด้วยการทำนายความต้องการบนจุดให้บริการที่มีอยู่จากตัวแบบความจำระยะสั้น และระยะยาว และเสริมด้วยโครงข่ายประสาทการแปลง ด้วยวิธีการปรับปรุงการประมาณค่าช่วงเวลาปรับตัว (Adaptive Moment Estimation) ซึ่งเป็นตัวแบบที่มีประสิทธิภาพในการทำนายบนจุดให้บริการใหม่ได้เหมาะสมที่สุดสามารถทำนายผลกระทบที่เกิดขึ้นจากบนจุดให้บริการที่มีอยู่ได้ดังรูปที่ 20 โดยที่มีค่าเฉลี่ยความผิดพลาดกำลังสองเท่ากับ 3.5756 และค่าเฉลี่ยความผิดพลาดสัมบูรณ์เท่ากับ 3.1142 และรูปที่ 21 โดยที่มีค่าเฉลี่ยความผิดพลาดกำลังสองเท่ากับ 3.3737 และค่าเฉลี่ยความผิดพลาดสัมบูรณ์เท่ากับ 2.8166



รูปที่ 20 การทำนายความต้องการบนจุดให้บริการที่อยู่หมายเลข 1



รูปที่ 21 การทำนายความต้องการบนจุดให้บริการที่อยู่หมายเลข 2

บทที่ 5 สรุปผลการดำเนินงานการวิจัย

5.1 บทสรุป

งานวิจัยนี้มีวัตถุประสงค์เพื่อสร้างตัวแบบทางคณิตศาสตร์ในการทำนายจำนวนความต้องการใช้จุดให้บริการใหม่และรับรู้ถึงผลกระทบต่อจุดให้บริการที่มีอยู่ พบว่าการสร้างตัวแบบด้วยตัวแบบความจำระยะสั้น และระยะยาว และเสริมด้วยโครงข่ายประสาทการแปลงกราฟด้วยวิธีปรับปรุงการประมาณค่าช่วงเวลาปรับตัวจะมีประสิทธิภาพในการทำนายมากที่สุดและค่าความผิดพลาดในการทำนายน้อยที่สุด ในขณะที่การสร้างตัวแบบด้วยหน่วยเวียนกลับแบบมีประตูและเสริมด้วยโครงข่ายประสาทการแปลงกราฟด้วยวิธีปรับปรุงการถ่ายถอดรากกำลังสองเฉลี่ยมีประสิทธิภาพในการทำนายน้อยที่สุด และค่าความผิดพลาดในการทำนายมากที่สุด โดยที่ตัวแบบที่มีวิธีการปรับปรุงการประมาณค่าช่วงเวลาปรับตัวกับวิธีการปรับปรุงการไล่ความชันแบบสุ่มด้วยโมเมนตัมมีประสิทธิภาพในการทำนายที่ใกล้เคียงกันและค่าความผิดพลาดในการทำนายใกล้เคียงกัน แต่วิธีการปรับปรุงการไล่ความชันแบบสุ่มด้วยโมเมนตัมจะมีค่าความผิดพลาดในการทำนายมากกว่าวิธีการปรับปรุงการประมาณค่าช่วงเวลาปรับตัวอยู่เล็กน้อย และสามารถทำนายความต้องการของจุดให้บริการที่มีอยู่ได้

5.2 อภิปรายผล

จากการวิจัยเพื่อการทำนายความต้องการของจุดให้บริการรถเช่าใหม่จากข้อมูลของจุดให้บริการที่มีอยู่ของบริษัทที่ใช้เป็นกรณีศึกษา โดยในแต่ละขั้นตอนจำเป็นต้องอาศัยความร่วมมือกันในหลายภาคส่วน ไม่ว่าจะเป็น ความร่วมมือของพนักงานในบริษัทกรณีศึกษา นักพัฒนาแอปพลิเคชัน นักออกแบบฐานข้อมูล และรวมถึงผู้วิจัย จากการศึกษาครั้งนี้สามารถแบ่งการอภิปรายออกเป็น 2 ประเด็น ดังนี้

1. ในการทำการทดสอบพบว่ามีความคาดเคลื่อนที่ไม่สามารถจัดการได้ เช่น อัตราการใช้งานรายวันของจุดให้บริการบางจุด อัตราส่วนของการจองรถยนต์ที่เป็นรายเดือนต่อการจองที่เป็นรายวัน ดังนั้นคณะผู้วิจัยจึงพยายามควบคุมความคลาดเคลื่อนที่สามารถควบคุมได้ให้ดีที่สุด คือ การตั้งสมมติฐานให้มีความเท่าเทียมกันในทุก ๆ จุดให้บริการ
2. ในการระดมสมองมีข้อจำกัด เนื่องจากมีการระบาดของโรคโควิด – 19 จึงต้องจึงต้องจัดการระดมสมองเป็นแบบออนไลน์ ทำให้อาจมีบรรยากาศ และการนำระดมสมองที่มีประสิทธิภาพน้อยกว่าแบบทั่วไป แต่การระดมสมองก็สามารถให้ผลที่ดีโดยคณะผู้วิจัยได้มีการสร้างปฏิสัมพันธ์ที่ดีกับพนักงานบริษัท และติดต่อกับพนักงานบริษัทที่เกี่ยวข้องทางออนไลน์อย่างใกล้ชิด เพื่อรายงานความคืบหน้าของการดำเนินงานอย่างต่อเนื่อง

5.3 ข้อเสนอแนะ

เนื่องจากโครงงานนี้มีประเด็นที่สามารถแก้ไขและพัฒนาเพื่อเป็นประโยชน์ต่อการศึกษาและพัฒนาความสามารถในการใช้งานเว็บไซต์ได้ในอนาคต จึงมีข้อเสนอแนะดังนี้

1. ขั้นตอนศึกษาพฤติกรรมการใช้งานของผู้ใช้งาน และการศึกษาข้อมูลเชิงลึก ควรมีจำนวนข้อมูลที่มากกว่านี้เพื่อตรวจสอบ ติดตาม ศึกษาได้มากขึ้น
2. เพื่อให้ตัวแบบทำนายได้แม่นยำควรมีจำนวนข้อมูลมากขึ้น เพื่อสร้างมาตรฐานที่ดีในการทำนาย และตัวแบบสามารถเรียนรู้การใช้งานในจุดบริการต่าง ๆ ได้มากขึ้น

รายงานอ้างอิง

นันทพงศ์ พันทวีศักดิ์. (2560). จูนเครื่องก่อนสตาร์ทบริการ car sharing สืบค้นจาก https://www.scbeic.com/th/detail/product/3429

Brownlee, J. (2019). How to Configure the Learning Rate When Training Deep Learning Neural Networks. https://machinelearningmastery.com/learning-rate-for-deep-learning-neural-networks

Dingil, A. E., Schweizer, J., Rupi, F., & Stasiskiene, Z. (2018). Road network extraction with OSMNx and SUMOPy

Du, J. (2019). The Frontier of SGD and Its Variants in Machine Learning.

Géron, A. (2017). Hands-On Machine Learning with Scikit-Learn & TensorFlow. O'Reilly Media, Inc., 401-480.

Huaxiu Yao, F. W., Jintao Ke, Xianfeng Tang,Yitian Jia, Siyu Lu, Pinghua Gong, Jieping Ye,Zhenhui Li. (2018). Deep Multi-View Spatial-Temporal Network for Taxi Demand Prediction.

Kim, Y., & Kim, S. (2021). Forecasting Charging Demand of Electric Vehicles Using Time-Series Models. Energies, 14(5). https://doi.org/10.3390/en14051487

Luo, M., Du, B., Klemmer, K., Zhu, H., Ferhatosmanoglu, H., & Wen, H. (2020). Data-driven forecast demand for fast expandsion car sharing. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 4(1), 1-21. https://doi.org/10.1145/3381005

Man Luo, H. W., Yi Luo, Bowen Du, Konstantin Klemmer, Hongming Zhu. (2019). Dynamic Demand Prediction for Expanding Electric Vehicle.

Wang, N., Jia, S., & Liu, Q. (2021). A user-based relocation model for one-way electric carsharing system based on micro demand prediction and multi-objective optimization. Journal of Cleaner Production, 296. https://doi.org/10.1016/j.jclepro.2021.126485

Wang, N., Jia, S., & Liu, Q. (2021). A user-based relocation model for one-way electric carsharing system based on micro demand prediction and multi-objective optimization. *Journal of Cleaner Production*, *296*

Xu Geng, Y. L., Leye Wang, Lingyu Zhang, Qiang Yang, Jieping Ye, Yan Liu. (2019). Spatiotemporal Multi-Graph Convolution Network for Ride-Hailing Demand Forecasting.

Yi Luo, Q. L., Hongming Zhu, Hongfei Fan, Tianyou Song, Chang Wu Yu, Bowen Du. (2019). Multistep Flow Prediction on Car-Sharing Systems: A Multi-Graph Convolutional Neural Network with Attention Mechanism.

Zhou, J., He, Z., Song, Y. N., Wang, H., Yang, X., Lian, W., & Dai, H.-N. (2020). Precious Metal Price Prediction Based on Deep Regularization Self-Attention Regression. IEEE Access, 8, 2178-2187. https://doi.org/10.1109/access.2019.2962202

Zhou, K., Wang, W., Hu, T., & Deng, K. (2020). Time Series Forecasting and Classification Models Based on Recurrent with Attention Mechanism and Generative Adversarial Networks. Sensors (Basel), 20(24). https://doi.org/10.3390/s20247211