

การส่งเสริมการทำประกันภัยรถยนต์ภาคสมัครใจโดยเทคนิคการเรียนรู้ของเครื่อง

PROMOTING VOLUNTARY MOTOR INSURANCE THROUGH MACHINE LEARNING TECHNIQUES

พจมาลย์ สุภรณ์ไพจิตร¹

ดร.ธนภัทร มั่งคะจิตร²

บทคัดย่อ

สารนิพนธ์ฉบับนี้มีวัตถุประสงค์เพื่อส่งเสริมการทำประกันภัยรถยนต์ภาคสมัครใจโดยใช้เทคนิคการเรียนรู้ของเครื่อง (Machine Learning Techniques) ในการจัดลำดับความสำคัญของกลุ่มจังหวัดตามรูปแบบการทำประกันภัยและการเรียกร้องค่าสินไหมทดแทนที่คล้ายคลึงกัน ด้วยเทคนิคการวิเคราะห์การจัดกลุ่มข้อมูล (Clustering) ร่วมกับการพยากรณ์ข้อมูลอนุกรมเวลา (Time-Series Forecasting) ของข้อมูลการรับประกันภัยและการเรียกร้องสินไหมทดแทนประกันภัยรถยนต์ภาคสมัครใจ โดยใช้ปัจจัยต่างๆ ดังนี้ (1) จำนวนรถจดทะเบียนสะสม (2) สัดส่วนรถไม่ทำประกันภัย (3) อัตราการเคลมต่อกรมธรรม์ (4) มูลค่าเคลมเฉลี่ยต่อครั้ง ผลการจัดกลุ่มสามารถแบ่งจังหวัดได้ 8 กลุ่ม และสามารถจัดลำดับความสำคัญกลุ่มจังหวัดที่เป็นเป้าหมายเพื่อส่งเสริมการทำประกันภัยได้ 2 กลุ่ม ได้แก่ กลุ่มลำดับความสำคัญที่ 1 เป็นกลุ่มจังหวัดที่มีปริมาณรถพอประมาณ และส่วนใหญ่ไม่ทำประกันภัย การเกิดอุบัติเหตุแต่ครั้งมีมูลค่าความเสียหายสูง กลุ่มลำดับความสำคัญที่ 2 เป็นกลุ่มจังหวัดที่มีปริมาณรถมาก และมากกว่าครึ่งไม่มีประกันภัย แต่ความเสี่ยงภัยในพื้นที่ยังต่ำ ซึ่งเมื่อพิจารณาร่วมกับผลการพยากรณ์แนวโน้มการเติบโตของข้อมูลสามารถใช้เป็นแนวทางวางแผนงานในการส่งเสริมการทำประกันภัยรถยนต์ภาคสมัครใจในแต่ละกลุ่มได้อย่างเหมาะสม

คำสำคัญ: ประกันภัยรถยนต์ภาคสมัครใจ, การส่งเสริมการทำประกันภัย, การเรียนรู้ด้วยเครื่อง,
การวิเคราะห์การจัดกลุ่มข้อมูล, การพยากรณ์ข้อมูลอนุกรมเวลา

¹นักศึกษาลัทธิสุตรวิศกรรมศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมข้อมูลขนาดใหญ่

²อาจารย์ที่ปรึกษาสารนิพนธ์

ABSTRACT

This thesis aims to promote voluntary motor insurance using machine learning techniques prioritize provincial clusters based on similar insurance patterns and claims applying Clustering combined with Time-Series Forecasting. The study utilizes data on voluntary car insurance and insurance claims, incorporating various factors: (1) the total number of registered vehicles, (2) the proportion of uninsured vehicles, (3) the claim rate per policy, and (4) the average claim value per incident. The clustering results divided the provinces into eight groups and identified two priority groups for promoting insurance. Priority1 consists of provinces with a moderate number of vehicles, a high proportion of uninsured vehicles, and a high average claim amount per incident. Priority2 consists of provinces with a large number of vehicles, over half of which are uninsured, but with lower risk levels. Considering the forecasted growth trends, these insights can guide the planning and promotion of motor insurance in each group effectively.

Keywords: Voluntary Motor Insurance, Promoting Insurance, Machine Learning Techniques, Clustering, Time-Series Forecasting

บทนำ

อุบัติเหตุจราจรทางบกเป็นหนึ่งในปัญหาสาธารณะที่นำมาซึ่งความสูญเสียในชีวิตและทรัพย์สินของผู้คนและมีแนวโน้มสูงขึ้นอย่างต่อเนื่อง การทำประกันภัยรถยนต์จึงมีความสำคัญในการช่วยจัดการความเสี่ยงและบรรเทาความเสียหายที่เกิดขึ้นกับผู้ประสบภัยจากจรด โดยประเทศไทยมีการรับประกันภัยรถยนต์ แบ่งเป็น 2 ประเภทหลักคือ การประกันภัยรถยนต์ภาคบังคับ และการประกันภัยรถยนต์ภาคสมัครใจ โดยการประกันภัยรถยนต์ภาคบังคับ เป็นการประกันภัยที่เจ้าของรถแต่ละคันต้องจัดให้มีประกันภัยตามความคุ้มครองที่กำหนดไว้ใน พระราชบัญญัติคุ้มครองผู้ประสบภัยจากรถ พ.ศ. 2535 มีวัตถุประสงค์เพื่อคุ้มครองและเยียวยาแก่ผู้ประสบภัยจากการใช้รถที่ได้รับอันตรายต่อชีวิตและร่างกาย แต่อย่างไรก็ตามมูลค่าความเสียหายที่เกิดขึ้นอาจจะมากกว่าจำนวนเงินเอาประกันภัยจากการทำประกันภัยรถยนต์ภาคบังคับ ซึ่งผู้ประสบภัยจะต้องผิดชอบความเสียหายส่วนเกิน ดังนั้นจึงมีการประกันภัยรถยนต์ภาคสมัครใจเกิดขึ้นเพื่อคุ้มครองความเสียหายที่เกินจากความรับผิดชอบของประกันรถยนต์ภาคบังคับ โดยครอบคลุมประเภทของภัยและจำนวนเงินเอาประกันภัยที่มากขึ้น รวมถึงการคุ้มครองความเสียหายที่เกิดขึ้นกับตัวรถ

จากสถิติของสำนักงานคณะกรรมการกำกับและส่งเสริมการประกอบธุรกิจประกันภัย (สำนักงาน คปภ.) พบว่า มูลค่าเบี้ยประกันวินาศภัยสะสม ในปี 2566 มาจากประกันภัยรถยนต์ภาคสมัครใจ 50% หรือคิดเป็นมูลค่า 141,060,322 บาท จะเห็นได้ว่า ประกันภัยรถยนต์เป็นธุรกิจที่มีความสำคัญอย่างยิ่ง

ต่อการเติบโตของอุตสาหกรรมประกันวินาศภัย และจากข้อมูลสถิติประกันภัยรถยนต์ภาคสมัครใจ (สำนักงาน คปภ.) และข้อมูลรถจดทะเบียน (กรมการขนส่งทางบก) ปี 2563-2566 พบว่า สัดส่วนการทำประกันภัยรถยนต์ภาคสมัครใจเฉลี่ยอยู่ที่ 27% โดยมีจำนวนรถที่ไม่มีประกันภัยรถยนต์ภาคสมัครใจในปี 2566 อีกจำนวน 32,193,422 คัน จากข้อมูลนี้แสดงให้เห็นว่า ตลาดการประกันภัยรถยนต์ภาคสมัครใจยังมีโอกาสเติบโตอีกมาก ดังนั้นเพื่อสนับสนุนการดำเนินงานของหน่วยงานภาครัฐในการส่งเสริมการทำประกันภัยรถยนต์ภาคสมัครใจ ผู้วิจัยจึงเสนอแนวทางเพื่อส่งเสริมการทำประกันภัยรถยนต์ภาคสมัครใจโดยใช้เทคนิคการเรียนรู้ของเครื่อง (Machine Learning Techniques) ในการจัดลำดับความสำคัญของกลุ่มจังหวัดตามรูปแบบการทำประกันภัยและการเรียกร้องค่าสินไหมทดแทนที่คล้ายคลึงกัน ด้วยเทคนิคการวิเคราะห์การจัดกลุ่มข้อมูล (Clustering) ร่วมกับการพยากรณ์ข้อมูลอนุกรมเวลา (Time-Series Forecasting) ของข้อมูลการรับประกันภัยและการเรียกร้องสินไหมทดแทนประกันภัยรถยนต์ภาคสมัครใจ เพื่อสนับสนุนการวางแผนงานของหน่วยงานที่เกี่ยวข้อง

วัตถุประสงค์ของงานวิจัย

เพื่อส่งเสริมการทำประกันภัยรถยนต์ภาคสมัครใจ โดยสามารถ

1. จัดลำดับความสำคัญของกลุ่มจังหวัดเป้าหมาย ตามรูปแบบการทำประกันภัยและการเรียกร้องค่าสินไหมทดแทนคล้ายคลึงกัน
2. พยากรณ์แนวโน้มการทำประกันภัยและการเรียกร้องค่าสินไหมทดแทนรายจังหวัด เพื่อสนับสนุนการวางแผนงานของหน่วยงานที่เกี่ยวข้อง

ขอบเขตของงานวิจัย

พื้นที่ศึกษาคือ จังหวัดของประเทศไทย โดยใช้ข้อมูลการทำประกันภัยรถยนต์ภาคสมัครใจ และข้อมูลการเรียกร้องสินไหมทดแทนประกันภัยรถยนต์ภาคสมัคร ตั้งแต่ปี 2562 ถึง เดือนมิถุนายน 2567 จากศูนย์กลางข้อมูลเปิดภาครัฐ ของสำนักงานคณะกรรมการกำกับและส่งเสริมการประกอบธุรกิจประกันภัย และข้อมูลรถจดทะเบียนสะสม ตั้งแต่ปี 2562 ถึง เดือนมิถุนายน 2567 ของกรมการขนส่งทางบก

แนวคิด ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

1. ความหมายของการประกันภัย และประเภทการประกันภัยรถยนต์

สมาคมประกันวินาศภัย (2564) ได้ให้ความหมายของการประกันภัย หมายถึง การที่บุคคลฝ่ายหนึ่งเรียกว่า "ผู้รับประกันภัย" ทำหน้าที่เป็นหลักประกันให้แก่บุคคลอีกฝ่ายหนึ่งเรียกว่า "ผู้เอาประกันภัย" ในกรณีที่ผู้เอาประกันภัยได้รับความเดือดร้อนทางการเงินจากเหตุการณ์ที่อาจเกิดขึ้นในอนาคต โดยผู้รับ

ประกันภัยสัญญาว่าจะจ่ายเงินค่าสินไหมทดแทนให้ตามจำนวนและเงื่อนไขที่ตกลงกันไว้ หรือจัดการให้ทรัพย์สินที่เอาประกันภัยนั้นกลับคืนสู่สภาพเดิมก่อนเกิดความเสียหาย โดยผู้เอาประกันภัยจะต้องจ่ายเบี้ยประกันภัยตามที่ตกลงกัน

การเอาประกันภัยเป็นวิธีการหนึ่งในการจัดการกับความเสี่ยงต่าง ๆ ที่เราเผชิญในชีวิตประจำวัน โดยการโอนความเสี่ยงไปยังผู้รับประกันภัย หากเกิดวินาศภัยขึ้น ผู้เอาประกันภัยจะได้รับการชดเชยค่าสินไหมทดแทนตามความเสียหายที่เกิดขึ้นจริง

ทั้งนี้การบริหารความเสี่ยงด้วยการทำประกันภัย ควรพิจารณาปัจจัยดังต่อไปนี้

- 1) ความรุนแรงของความสูญเสีย (Loss Severity) พิจารณาจากมูลค่าหรือผลกระทบ
- 2) ความถี่ของความสูญเสีย (Loss Frequency) พิจารณาจากความถี่ในการเกิด

2. ทฤษฎีทางการตลาดที่เกี่ยวข้อง

งานวิจัยนี้ได้ศึกษาทฤษฎีทางการตลาด ที่เกี่ยวข้องกับการแบ่งส่วนตลาด (Segmentation) การกำหนดตลาดเป้าหมาย (Targeting) มาใช้ในการ พิจารณาปัจจัย เพื่อแบ่งกลุ่มและกำหนดจังหวัดเป้าหมาย

การแบ่งส่วนตลาด และการกำหนดตลาดเป้าหมาย เป็นขั้นตอนตามกลยุทธ์ทางการตลาดที่เรียกว่า STP Marketing (Segmentation, Targeting, Positioning) ซึ่งเป็นแนวคิดที่ช่วยให้องค์กรสามารถกำหนดกลยุทธ์การตลาดที่มีประสิทธิภาพมากขึ้น โดยการแบ่งลูกค้าออกเป็นกลุ่ม ๆ และเลือกกลุ่มที่เหมาะสมที่สุดสำหรับผลิตภัณฑ์หรือบริการของตน แล้ววางตำแหน่งผลิตภัณฑ์หรือบริการนั้นให้ตรงกับความต้องการและความคาดหวังของกลุ่มลูกค้าเหล่านั้น

การแบ่งส่วนตลาด (Marketing Segmentation) หมายถึงการแบ่งกลุ่มลูกค้าและความต้องการของลูกค้าออกเป็นส่วนย่อยๆ โดยจัดให้ลูกค้าที่มีความต้องการคล้ายกันอยู่ในกลุ่มเดียวกัน เพื่อที่จะเลือกเป็นตลาดเป้าหมาย หรือกลุ่มเป้าหมายต่อไป ซึ่งโดยทั่วไปจะใช้เกณฑ์ ดังนี้

1) เกณฑ์ภูมิศาสตร์ (Geographic Segmentation) คือ การแบ่งส่วนตลาดตามสถานที่ เช่น ทวีป ภูมิภาค จังหวัด เขตเมืองหรือเขตชนบท เป็นต้น

2) เกณฑ์ประชากรศาสตร์ (Demographic Segmentation) คือ การแบ่งส่วนตลาดตามลักษณะประชากรของลูกค้า เช่น อายุ เพศ ขนาดครอบครัว เป็นต้น ซึ่งเกณฑ์นี้นิยมใช้ในการแบ่งส่วนตลาดมากที่สุด

3) เกณฑ์จิตวิทยา (Psychographic Segmentation) คือ การแบ่งส่วนตลาดตามปัจจัยจิตวิทยา เช่น ชื่นทางสังคม รูปแบบการดำรงชีวิต (Life Style) และบุคลิกภาพ (Personality) เป็นต้น โดยการแบ่งส่วนตลาดตามเกณฑ์นี้มีความซับซ้อนและต้องการการวิจัยเชิงลึกเกี่ยวกับพฤติกรรมและความต้องการของลูกค้า

4) เกณฑ์พฤติกรรมศาสตร์ (Behavior Segmentation) คือ การแบ่งส่วนตลาดตามพฤติกรรมการบริโภค เช่น พฤติกรรมการซื้อ ความภักดีต่อแบรนด์ โอกาสในการใช้งาน เป็นต้น

อย่างไรก็ตามในการแบ่งส่วนตลาดอาจใช้เกณฑ์หลายอย่างร่วมกันเพื่อให้สามารถแบ่งกลุ่มลูกค้าได้อย่างชัดเจนและมีประสิทธิภาพ

การเลือกตลาดเป้าหมาย (Market Targeting) หมายถึง การเลือกส่วนตลาดหนึ่งหรือหลายส่วนหลังจากที่แบ่งส่วนตลาดแล้ว ซึ่งบริษัทจะพิจารณาเลือกส่วนตลาดที่มีโอกาสและศักยภาพสูงสุดในการตอบสนองความต้องการของลูกค้าที่ โดยพิจารณาจาก

- 1) ขนาดของตลาด (Market Size) เป็นปัจจัยสำคัญที่ช่วยประเมินความน่าสนใจในการลงทุน
- 2) อัตราการเติบโตของตลาด (Market Growth Rate) เป็นตัวบ่งชี้ถึงโอกาสในอนาคตของตลาด
- 3) ความสามารถของธุรกิจในการตอบสนองความต้องการ (Business Capability) เช่น ความสามารถในการผลิตสินค้าและบริการ รวมถึงทรัพยากรที่มีอยู่และความสามารถในการแข่งขัน
- 4) สภาพการแข่งขันและความน่าสนใจของตลาด (Market Competition and Attractiveness) เพื่อให้ทราบถึงจำนวนคู่แข่งและความรุนแรงทางการแข่งขันในตลาด

3. การจัดกลุ่ม (Cluster analysis)

เป็นหนึ่งในเทคนิคการทำเหมืองข้อมูล เพื่อหาความรู้ที่ซ่อนอยู่ภายในข้อมูลหมวดหมู่ เพื่อให้สามารถอธิบายลักษณะของข้อมูลหรือการแบ่งกลุ่มข้อมูลที่น่าสนใจ (Kassambara, 2017) การจัดกลุ่มนี้เป็นกระบวนการเรียนรู้แบบไม่มีผู้สอน (Unsupervised learning) ซึ่งไม่มีกลุ่มที่กำหนดไว้ล่วงหน้า ในการวิจัยนี้ใช้วิธีการแบ่งกลุ่มแบบ K-Means clustering ซึ่งเป็นวิธีที่นิยมใช้งานทั้งในงานวิจัยและอุตสาหกรรมเพราะสามารถนำผลลัพธ์ไปประยุกต์ใช้ในด้านกลยุทธ์การตลาดอย่างมีประสิทธิภาพ (Kakkar and Parashar, 2014) ทั้งนี้ในการกำหนดจำนวนกลุ่ม (K) ที่เหมาะสม มีหลายวิธีที่นิยมใช้ เช่น Elbow method, Information criterion approach, Silhouette method, และ Cross-validation โดยวิธี Elbow และ Silhouette เป็นที่นิยมเนื่องจากสามารถใช้ระบุจำนวนคลัสเตอร์ที่เหมาะสมได้ง่าย (Kodinariya & Makwana, 2013; Habib, 2021) โดยในงานวิจัยนี้จะใช้วิธี Elbow method เพื่อระบุจำนวนกลุ่ม

4. การพยากรณ์แบบอนุกรมเวลา (Time-Series Forecasting)

เป็นเทคนิคที่ใช้ในการคาดการณ์ข้อมูลในอนาคตโดยอาศัยข้อมูลที่เกิดขึ้นในอดีต ซึ่งเป็นการจัดการกับข้อมูลที่มีการเรียงตามลำดับเวลา โดยเทคนิคที่ใช้ในการพยากรณ์แบบอนุกรมเวลามีหลายวิธี ทั้งแบบจำลอง ARIMA (Autoregressive Integrated Moving Average) การใช้แบบจำลองเชิงโครงข่ายประสาท (Neural Network Models) เช่น Recurrent Neural Networks (RNN) และ Long Short-Term Memory (LSTM) รวมถึงเทคนิคแบบผสม (Hybrid Models) เป็นการผสมผสานระหว่างแบบจำลองหลาย ๆ แบบ เพื่อเพิ่มประสิทธิภาพในการพยากรณ์ เช่น การรวม ARIMA กับ Neural Network โดยงานวิจัยนี้ใช้ Prophet ซึ่งพัฒนาโดย Facebook สำหรับการพยากรณ์อนุกรมเวลา ซึ่งถูกออกแบบมาให้ใช้งานง่ายและมีความยืดหยุ่นสูง สามารถจัดการกับข้อมูลที่มีความไม่สม่ำเสมอ มีค่าผิดปกติ และมีการขาดหายของข้อมูลได้

Prophet ใช้การรวมกันของแบบจำลองเชิงเส้นและไม่เชิงเส้นในการพยากรณ์ โดยส่วนประกอบหลักของแบบจำลองได้แก่ แนวโน้ม (Trend) ซึ่งใช้ piecewise linear หรือ logistic growth curve เพื่อจับลักษณะของแนวโน้มในระยะยาว ฤดูกาล (Seasonality) ซึ่งใช้ Fourier series ในการจับลักษณะของฤดูกาล และวันหยุด (Holiday) ที่สามารถรวมข้อมูลเกี่ยวกับวันหยุดหรือเหตุการณ์สำคัญที่มีผลต่อข้อมูล การใช้งาน Prophet เริ่มจากการเตรียมข้อมูล สร้างและฝึกแบบจำลอง ทำการพยากรณ์ และแสดงผลการพยากรณ์

การประเมินผลการพยากรณ์มักใช้ตัวชี้วัด เช่น Mean Absolute Error (MAE), Mean Squared Error (MSE) และ Root Mean Squared Error (RMSE) เพื่อวัดความแม่นยำของการพยากรณ์ การพยากรณ์แบบอนุกรมเวลาเป็นเครื่องมือสำคัญในการวางแผนและตัดสินใจทางธุรกิจและวิทยาศาสตร์ ซึ่งการเลือกใช้เทคนิคที่เหมาะสมกับลักษณะของข้อมูลจะช่วยให้ได้ผลลัพธ์ที่แม่นยำและเป็นประโยชน์มากขึ้น

5. งานวิจัยที่เกี่ยวข้อง

5.1 Multidimensional Analysis of Insurance Companies' Underwriting Choices: Case Studies of Yonghe Street and Cote de Nuits

งานวิจัยนี้พัฒนาโมเดลประเมินความเสี่ยงภัยพิบัติที่ส่งผลต่อการตัดสินใจรับประกันทรัพย์สินของบริษัทประกันภัย โดยใช้ดัชนีที่เกี่ยวข้องจากประเทศ ตุรกี ฝรั่งเศส อินเดีย ออสเตรเลีย และกาบอง ซึ่งดัชนีเหล่านี้ครอบคลุมจำนวนประชากรที่ได้รับผลกระทบ จำนวนผู้เสียชีวิต ความเสียหายต่อสิ่งปลูกสร้าง และผลกระทบทางเศรษฐกิจ ข้อมูลเหล่านี้ถูกนำมาพยากรณ์ล่วงหน้า 5 ปีด้วยโมเดล Long Short-Term Memory (LSTM) เพื่อคาดการณ์แนวโน้มภัยพิบัติในอนาคต จากนั้นทำการวิเคราะห์น้ำหนักของดัชนีแต่ละตัวด้วยวิธีการถ่วงน้ำหนักแบบเอนโทรปี (Entropy Weight Method) และจัดกลุ่มความเสี่ยงด้วย K-means Clustering โดยแบ่งพื้นที่ตามระดับความเสี่ยงภัยพิบัติ คือ ประชากร สิ่งปลูกสร้าง และเศรษฐกิจ ได้ทั้งหมด 8 กลุ่ม ซึ่งบริษัทประกันภัยจะไม่รับประกันในพื้นที่ที่มีความเสี่ยงสูงสุด

โมเดลนี้ถูกนำไปประยุกต์ใช้กับพื้นที่ถนนหย่งเหอในจีนและ Cote de Nuits ในฝรั่งเศส ผลการวิจัยแสดงให้เห็นว่าการแบ่งกลุ่มช่วยให้สามารถระบุพื้นที่ที่มีความเสี่ยงสูงสุดได้อย่างชัดเจน ซึ่งพื้นที่เหล่านี้จะถือว่าไม่เหมาะสมสำหรับการรับประกันภัย ในขณะที่พื้นที่ที่มีความเสี่ยงต่ำกว่าจะถูกพิจารณารับประกันได้ตามความเหมาะสม การพิจารณาความเสี่ยงที่เกิดขึ้นจริงในแต่ละพื้นที่ช่วยเพิ่มความปลอดภัยในการดำเนินธุรกิจและลดความเสี่ยงจากการจ่ายค่าสินไหมทดแทนในกรณีที่เกิดภัยพิบัติได้อย่างมีประสิทธิภาพ

5.2 Time Series Forecasting using Machine Learning

งานวิจัยที่ศึกษาเกี่ยวกับวิธีการพยากรณ์ข้อมูลอนุกรมเวลา โดยเน้นไปที่การวิเคราะห์และเปรียบเทียบประสิทธิภาพของโมเดล XGBoost, ARIMA, Long-Short Term Memory (LSTM) และ Facebook Prophet เพื่อหาว่าโมเดลใดเหมาะสมที่สุดในการพยากรณ์ราคาสินทรัพย์ที่มีความผันผวนและ

แนวโน้มที่แตกต่างกัน ข้อมูลที่ใช้ในงานวิจัยประกอบด้วยราคาของสินทรัพย์สามประเภท ได้แก่ Bitcoin ที่มีความผันผวนสูง, S&P 500 ที่มีแนวโน้มคงที่, และพันธบัตรรัฐบาลสหรัฐฯ อายุ 10 ปีที่มีความผันผวนต่ำกว่า ข้อมูลเหล่านี้ถูกใช้เพื่อทดสอบประสิทธิภาพของโมเดลต่าง ๆ โดยงานวิจัยได้ทำการปรับแต่งโมเดลเหล่านี้ให้เหมาะสมกับข้อมูลที่ใช้ทดสอบ และเปรียบเทียบผลการพยากรณ์โดยใช้ค่าความคลาดเคลื่อน เช่น Mean Squared Error (MSE) และ R-squared ผลการวิจัยพบว่าโมเดล Facebook Prophet มีประสิทธิภาพสูงสุดในการพยากรณ์ราคาของสินทรัพย์ที่มีความผันผวนสูง เช่น Bitcoin เนื่องจากสามารถปรับตัวตามการเปลี่ยนแปลงของข้อมูลได้ดีและสามารถปรับพารามิเตอร์ได้ตามความต้องการของผู้ใช้ ส่วนโมเดล LSTM ก็ทำงานได้ดีเช่นกัน โดยเฉพาะกับสินทรัพย์ที่มีแนวโน้มคงที่และความผันผวนต่ำ เช่น S&P 500 และพันธบัตรรัฐบาลสหรัฐฯ อายุ 10 ปี ในขณะที่โมเดล XGBoost และ ARIMA มีข้อจำกัดในการพยากรณ์สินทรัพย์ที่มีความผันผวนสูง แต่ยังคงมีประสิทธิภาพในสินทรัพย์ที่มีแนวโน้มคงที่มากกว่า

5.3 Forecasting at Scale

งานวิจัยนี้นำเสนอโมเดลการพยากรณ์อนุกรมเวลาที่ชื่อว่า Prophet ซึ่งถูกออกแบบมาให้ใช้งานได้ง่ายและปรับแต่งได้ตามความต้องการ เหมาะสำหรับข้อมูลธุรกิจที่มีความซับซ้อนและหลากหลาย โดยโมเดลนี้สามารถจัดการกับข้อมูลที่มีแนวโน้ม (Trend) การเปลี่ยนแปลงตามฤดูกาล (Seasonality) และผลกระทบจากเหตุการณ์พิเศษหรือวันหยุด (Holidays and Events) ได้อย่างมีประสิทธิภาพ โดยใช้หลักการวิเคราะห์การถดถอย (Regression Model) มีการแบ่งองค์ประกอบของข้อมูลออกเป็น 3 ส่วนหลัก ได้แก่ แนวโน้ม (Trend) ฤดูกาล (Seasonality) และผลกระทบจากวันหยุดหรือเหตุการณ์พิเศษ (Holidays and Events) ทำให้ผู้ใช้สามารถเข้าใจและปรับปรุงโมเดลได้อย่างง่ายดาย แม้ว่าจะมีการเปลี่ยนแปลงแนวโน้มที่ไม่คาดคิด โมเดลนี้ยังถูกออกแบบให้สามารถจัดการกับข้อมูลที่ไม่สม่ำเสมอ (Irregular) หรือมีความไม่สมบูรณ์ได้ดี นอกจากนี้ยังสามารถปรับแต่งพารามิเตอร์ได้อย่างง่ายดายโดยไม่จำเป็นต้องมีความรู้เชิงลึกทางสถิติมากนัก การวิจัยนี้ใช้ข้อมูลจาก Facebook เช่น ข้อมูลที่เกี่ยวข้องกับการสร้างและเข้าร่วมกิจกรรมบนแพลตฟอร์ม Facebook มาใช้ในการทดสอบประสิทธิภาพของโมเดลและเปรียบเทียบประสิทธิภาพกับโมเดลอื่นๆ ได้แก่ ARIMA , Exponential Smoothing , Seasonal Naive และ TBATS ผลการเปรียบเทียบกับโมเดลเหล่านี้ทำให้เห็นว่า Prophet มีความสามารถในการทำนายที่แม่นยำกว่า โดยเฉพาะในกรณีที่ข้อมูลมีความซับซ้อน เช่น มีการเปลี่ยนแปลงตามฤดูกาลหลายแบบหรือมีแนวโน้มที่ไม่คงที่ นอกจากนี้ งานวิจัยยังเสนอแนวทาง "Analyst-in-the-Loop" ซึ่งผสมผสานระหว่างการพยากรณ์อัตโนมัติและการปรับแต่งโดยผู้เชี่ยวชาญ เพื่อเพิ่มความยืดหยุ่นและความเชื่อถือได้ของการพยากรณ์ ทำให้ระบบการพยากรณ์สามารถปรับเปลี่ยนตามความต้องการและมีการตรวจสอบประสิทธิภาพอย่างสม่ำเสมอ ซึ่งช่วยให้สามารถจัดการกับการพยากรณ์ในระดับใหญ่ได้อย่างมีประสิทธิภาพ

ระเบียบวิธีวิจัย

สารนิพนธ์นี้เป็นการศึกษาเพื่อหาแนวทางในการส่งเสริมการทำประกันภัยรถยนต์ภาคสมัครใจ โดยมีขั้นตอนการดำเนินการต่อไปนี้

1. การทำความเข้าใจธุรกิจ

จากการวิเคราะห์ข้อมูลรถจดทะเบียนสะสมและการทำประกันภัยรถยนต์ภาคสมัครใจระหว่างปี 2562 ถึงไตรมาส 2 ปี 2567 พบว่า 51% ของรถจดทะเบียนเป็นรถจักรยานยนต์ส่วนบุคคล และ 44% เป็นรถยนต์ส่วนบุคคล โดยมีสัดส่วนการทำประกันภัยในกลุ่มรถยนต์อยู่ที่ 47% และกลุ่มรถจักรยานยนต์อยู่ที่ 2% แม้ว่าทั้งสองกลุ่มควรเป็นเป้าหมายในการส่งเสริมการทำประกันภัย แต่เนื่องจากข้อมูลด้านการประกันภัยของรถจักรยานยนต์มีน้อย งานวิจัยนี้จึงวิเคราะห์เฉพาะกลุ่มรถยนต์ส่วนบุคคล

และจากการศึกษาทฤษฎีทางการตลาดและการปรึกษาผู้เชี่ยวชาญด้านธุรกิจ (Domain Expert) สามารถกำหนดเกณฑ์ในการแบ่งกลุ่มจังหวัดเป้าหมาย ดังนี้

- 1) จำนวนรถจดทะเบียนสะสม แสดงขนาดของตลาด
- 2) สัดส่วนรถที่ไม่มีประกันภัย แสดงโอกาสในการส่งเสริมการทำประกันภัย
- 3) อัตราการเคลมต่อกรมธรรม์ แสดงความถี่ในการเกิดเหตุ
- 4) มูลค่าเคลมเฉลี่ยต่อครั้ง แสดงความรุนแรงของเหตุการณ์

2. การรวบรวมและทำความเข้าใจข้อมูล

งานวิจัยนี้รวบรวมข้อมูลจากสองแหล่งข้อมูลที่สำคัญคือ ข้อมูลประกันภัยรถยนต์ภาคสมัครใจ และการเรียกร้องสินไหมทดแทนจากศูนย์ข้อมูลเปิดภาครัฐ ของสำนักงาน คปภ. และข้อมูลการจดทะเบียนสะสมจากกรมการขนส่งทางบก อย่างไรก็ตาม พบว่ามีการจัดประเภทของรถยนต์ที่ไม่สอดคล้องกัน จึงจำเป็นต้องให้ผู้เชี่ยวชาญจัดทำตารางประเภทรถให้เป็นมาตรฐานเดียวกันเพื่อใช้ในการวิเคราะห์ต่อไป

3. การเตรียมข้อมูล

ในการวิจัยนี้ ได้เริ่มจากการหาผลรวมของจำนวนรถจดทะเบียนสะสม จำนวนกรมธรรม์ จำนวนการเคลม และมูลค่าเคลม โดยแยกตามรหัสจังหวัด รหัสประเภทรถ และวันที่ของข้อมูล จากนั้นทำการปรับรหัสจังหวัดเป็นชื่อจังหวัด และปรับประเภทรถให้เป็นมาตรฐานเดียวกัน โดยอ้างอิงจากตารางประเภทรถมาตรฐานที่จัดทำไว้ก่อนหน้านี้ หลังจากนั้นได้ตรวจสอบคุณภาพข้อมูลและทำความสะอาดข้อมูล เพื่อให้มั่นใจว่าข้อมูลมีความถูกต้องและครบถ้วน แล้วจึงทำการเชื่อมโยงข้อมูลประกันภัยรถยนต์ภาคสมัครใจ ข้อมูลการเรียกร้องสินไหมทดแทน และข้อมูลการจดทะเบียน ให้เป็นชุดข้อมูลเดียวกัน พร้อมสร้างตัวแปรสำคัญสำหรับใช้ในการวิเคราะห์ ได้แก่

- 1) จำนวนรถจดทะเบียนสะสม คำนวณจากข้อมูลรถจดทะเบียนสะสม
- 2) สัดส่วนรถไม่มีประกันภัย = (จำนวนรถจดทะเบียนสะสม – จำนวนกรมธรรม์) / จำนวนรถ

จดทะเบียนสะสม

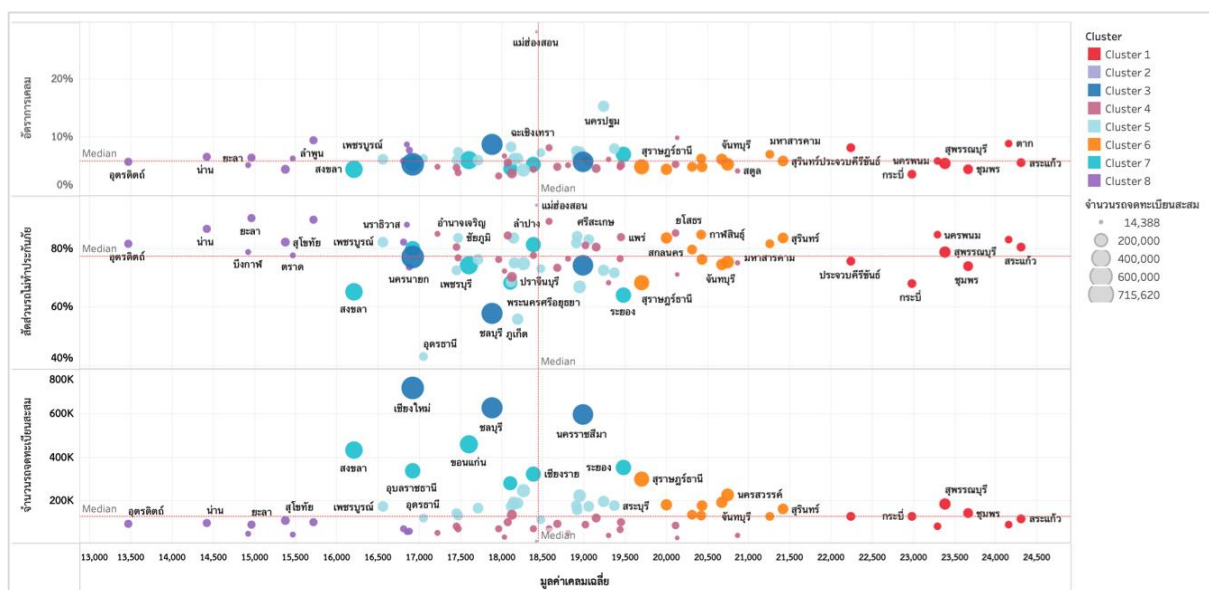
3) อัตราการเคลมต่อกรมธรรม์ = จำนวนการเคลม / จำนวนกรมธรรม์

4) มูลค่าเคลมเฉลี่ยต่อครั้ง = มูลค่าเคลมรวม / จำนวนการเคลม

จากนั้นได้วิเคราะห์ค่าผิดปกติ (Outlier) โดยใช้ Z-score ในการหาค่ามาตรฐานของตัวแปรเหล่านี้ จังหวัดที่มีค่า Z-score เกิน 3 หรือ ต่ำกว่า -3 จะถูกพิจารณาว่าเป็นค่าผิดปกติ ซึ่งในการวิเคราะห์ครั้งนี้ได้ตัดข้อมูลของกรุงเทพมหานคร สมุทรปราการ สมุทรสาคร ปทุมธานี และนนทบุรี ออกเพื่อให้ผลการวิเคราะห์มีความแม่นยำยิ่งขึ้น

4. การสร้างโมเดล และการวัดประสิทธิภาพ

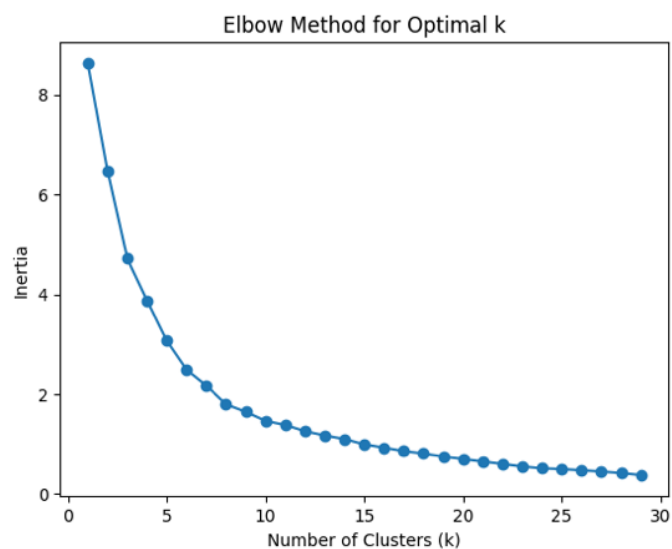
4.1 จัดลำดับความสำคัญของกลุ่มจังหวัดเพื่อส่งเสริมการทำประกันภัย โดยใช้วิธี K-means Clustering ในการแบ่งกลุ่มจังหวัดตามตัวแปรดังนี้ จำนวนรถจดทะเบียนสะสม สัดส่วนรถไม่มีประกันภัย อัตราการเคลมต่อกรมธรรม์ และมูลค่าเคลมเฉลี่ยต่อครั้ง จากข้อมูลไตรมาส 2 ปี 2567 กำหนดค่า K = 8 โดยหาค่า K ที่เหมาะสมด้วย Elbow Method ดังภาพที่ 1 ซึ่งได้ผลการวิเคราะห์กลุ่มจังหวัด ดังนี้



ภาพที่ 1 กราฟแสดงการแบ่งกลุ่มจังหวัดที่ได้จากการวิเคราะห์ K-Means Clustering

ตารางที่ 1 ผลการวิเคราะห์แบ่งกลุ่มจังหวัดด้วยวิธี K-means Clustering

Cluster	จำนวนจังหวัด	ลักษณะกลุ่ม
Cluster 1	7	รถไม่มีประกันภัยเหาะ มูลค่าเคลมสูง
Cluster 2	1	รถน้อย ส่วนใหญ่ไม่มีประกันภัย เคลมบ่อย
Cluster 3	3	รถเหาะ สัดส่วนรถไม่มีประกันภัยยังต่ำกว่าค่าเฉลี่ย
Cluster 4	19	รถน้อย ส่วนใหญ่ไม่มีประกันภัย มูลค่าเคลมสูง
Cluster 5	17	รถเหาะ สัดส่วนรถไม่มีประกันภัยยังต่ำกว่าค่าเฉลี่ย
Cluster 6	9	รถเหาะ รถไม่มีประกันภัยเหาะ มูลค่าเคลมสูง
Cluster 7	6	รถเหาะ สัดส่วนรถไม่มีประกันภัยยังต่ำกว่าค่าเฉลี่ย
Cluster 8	10	รถน้อย ส่วนใหญ่ไม่มีประกันภัย



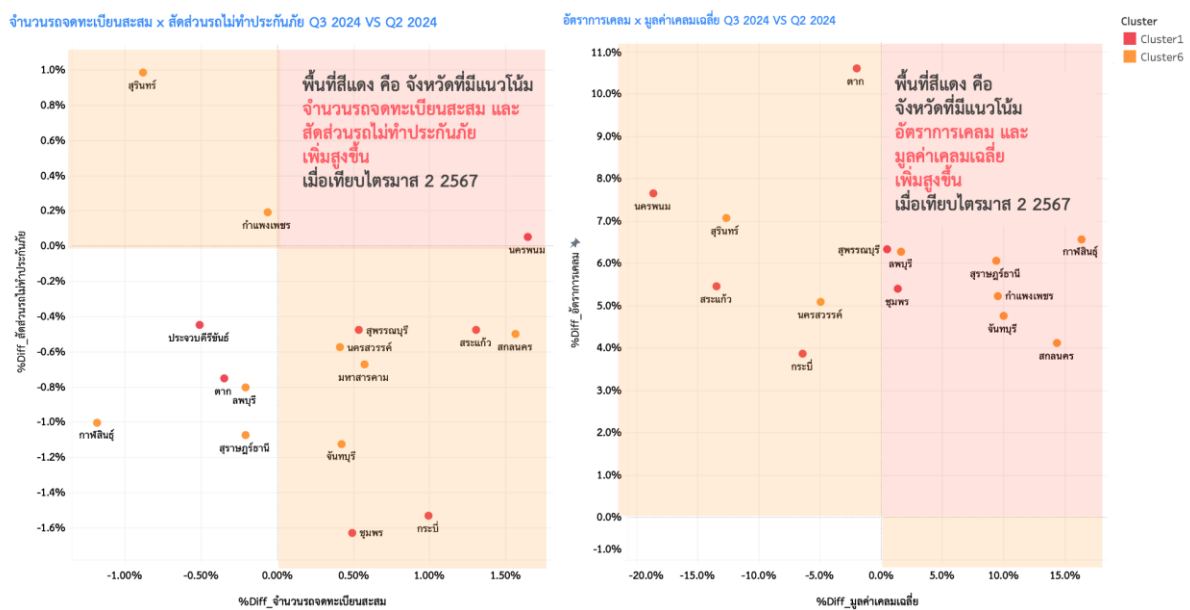
ภาพที่ 2 กราฟ Elbow Method ในการหาค่า Optimal cluster number (K)

4.2 สร้างแบบจำลองการพยากรณ์แนวโน้มจำนวนรถจดทะเบียนสะสม จำนวนกรมธรรม์ จำนวนการเคลม และ มูลค่าเคลม ในไตรมาส 3 ปี 2567 ถึงไตรมาส 2 ปี 2568 จากข้อมูลอดีตย้อนหลังตั้งแต่ไตรมาส 1 ปี 2562 ถึงไตรมาส 2 ปี 2567 โดยใช้โมเดล Prophet ผลการวัดประสิทธิภาพความถูกต้องของโมเดล ดังตารางที่ 2

ตารางที่ 2 ผลการวัดประสิทธิภาพของโมเดล Prophet

Feature	Mean Absolute Percentage Error
จำนวนรถจดทะเบียนสะสม	0.007
จำนวนกรมธรรม์	0.002
จำนวนการเคลม	0.070
มูลค่าเคลม	0.096

ค่าที่ได้จากการพยากรณ์นำไปคำนวณตัวแปร จำนวนรถจดทะเบียนสะสม สัดส่วนรถไม่มีประกันภัย อัตราการเคลมต่อกรมธรรม์ และมูลค่าเคลมเฉลี่ยต่อครั้ง จากนั้นนำไปคำนวณเปอร์เซ็นต์ความแตกต่างระหว่างค่าในไตรมาส 3 ปี 2567 ถึงไตรมาส 2 ปี 2568 กับค่าของไตรมาส 2 ปี 2567 เพื่อแสดงการเติบโตของขนาดตลาด โอกาส และความเสี่ยง ของแต่ละจังหวัด



ภาพที่ 3 แสดงผลการเปรียบเทียบค่าที่ได้จากการพยากรณ์แนวโน้ม กับค่าของไตรมาส 2 ปี 2567

สรุปผลการวิจัย และข้อเสนอแนะ

สรุปผลการวิจัย

1. จากการวิเคราะห์ธุรกิจ พบว่ามี 4 ปัจจัยสำคัญในการจัดลำดับความสำคัญของกลุ่มจังหวัดเป้าหมายสำหรับการส่งเสริมการทำความรู้จัก ได้แก่ (1) จำนวนรถจดทะเบียนสะสม (2) สัดส่วนรถไม่มีประกันภัย (3) อัตราการเคลมต่อกรมธรรม์ (4) มูลค่าเคลมเฉลี่ยต่อครั้ง

2. การจัดลำดับความสำคัญของกลุ่มจังหวัดด้วยเทคนิคการวิเคราะห์จัดกลุ่ม โดยใช้ข้อมูลไตรมาส 2 ปี 2567 สามารถแบ่งกลุ่มจังหวัดได้ 8 กลุ่ม และจัดลำดับความสำคัญได้ดังนี้

Priority 1 กลุ่มจังหวัดที่มีปริมาณรถพอประมาณ และส่วนใหญ่ไม่ทำประกันภัย การเกิดอุบัติเหตุแต่ครั้งมีมูลค่าความเสียหายสูง จำนวน 16 จังหวัด

Priority2 กลุ่มจังหวัดที่มีปริมาณรถมาก และมากกว่าครึ่งไม่มีประกันภัย ความเสี่ยงภัยในพื้นที่ต่ำ จำนวน 26 จังหวัด

3. การพยากรณ์แนวโน้มข้อมูลอนาคต และนำมาเทียบกับข้อมูลปัจจุบัน ทำให้เห็นแนวโน้มการเติบโตของขนาดตลาด โอกาส และความเสี่ยงซึ่งเมื่อพิจารณาเกี่ยวกับการจัดลำดับความสำคัญของกลุ่มจังหวัด สามารถใช้เป็นแนวทางวางแผนส่งเสริมการทำความรู้จักในแต่ละกลุ่มได้อย่างเหมาะสม

ข้อเสนอแนะ

1. เปลี่ยนแปลงการวิเคราะห์ข้อมูลจากระดับไตรมาสเป็นรายเดือน เพื่อให้ข้อมูลมีความทันสมัย และสามารถระบุแนวโน้มการเปลี่ยนแปลงที่เกิดขึ้นได้อย่างชัดเจนยิ่งขึ้น

2. นำข้อมูลยี่ห้อ รุ่นรถ มาใช้ในการวิเคราะห์จัดกลุ่มจังหวัดเพื่อช่วยในการแบ่งแยกกรณีที่กลุ่มเป้าหมายได้ละเอียดและเจาะจงมากขึ้น

บรรณานุกรม

ภาษาไทย

ณดา ทับทิมจัญญ์. (2559). การตลาดสำหรับสถาบันการศึกษา. วารสารปัญญาวิวัฒน์. 8(1), 288-302

พิมพ์ิสา จันทะ, กรวิทย์ ออกผล, Rodolphe Perrin, สืบสกุล คุรุรัตน์. (2566). การพยากรณ์ปริมาณเที่ยวบินของประเทศไทยหลังการระบาดของโรคโควิด 19 โดยใช้ FBprophet และ Neural-Prophet. วารสารวิจัยและนวัตกรรมท้องถิ่น, 18(1), 35-46

ศูนย์ข้อมูลกลางด้านการบาดเจ็บ กรมควบคุมโรค กระทรวงสาธารณสุข. ข้อมูลผู้บาดเจ็บเสียชีวิตจากอุบัติเหตุทางถนน (43 แฟ้ม : ผู้ป่วยนอก และผู้ป่วยใน). https://dip.ddc.moph.go.th/new/บริการ/hdc_opd_ipd01

สมาคมประกันวินาศภัยไทย. (2564). คู่มือประกันวินาศภัยไทย (พิมพ์ครั้งที่ 1). กรุงเทพฯ: สำนักพิมพ์
จุฬาลงกรณ์มหาวิทยาลัย.

สำนักงานคณะกรรมการกำกับและส่งเสริมการประกอบธุรกิจประกันภัย (คปภ.). ข้อมูลสถิติประกันภัย.
<https://www.oic.or.th/th>

วัฒนา ชยธวัช. (2565). การจัดกลุ่มจังหวัดในประเทศไทยด้วยปัจจัยประชากรศาสตร์และระดับการศึกษา ของ
โค วิ ด - 1 9 . วิ ท ย า ศ า ส ต ร์ แ ลະ เ ท ค โ น โ ล ยี สู่ ชุม ช น , 1 (1) , 1 – 1 8 .
<https://doi.org/10.57260/stc.2023.450>

ภาษาต่างประเทศ

Dutt, A. (2023). *Time Series Forecasting using Machine Learning*. Retrieved from
<https://www.udemy.com/course/time-series-analysis/>

Habib A.B. (2021). *Elbow Method vs Silhouette Co-efficient in Determining the Number of Clusters*.
Retrieved February,21,2020, from <https://medium.com/mlearning-ai/elbow-method-vs-silhouette-co-efficient-in-determining-the-number-of-clusters-33baff2fbee>.

Kakkar, P., Parashar, A. (2024). Comparison of Different Clustering Algorithms using WEKA Tool.
Engineering and Science.

Kassambara, A. (2017). Practical guide to cluster analysis in R: *Unsupervised machine learning (Vol. 1)*.
Sthda.

Kodinariya, T. M., Makwana, P. R. (2013). Review on determining number of Cluster in K-Means
Clustering. *International Journal*, 1(6), 90-95.

Taylor, S. J., Letham, B. (2017). *Forecasting at scale* (No. e3190v2). PeerJ Inc.
<https://doi.org/10.7287/peerj.preprints.3190v2>

Zheng, Yu. (2024). Multidimensional Analysis of Insurance Companies' Underwriting Choices: Case
Studies of Yonghe Street and Cote de Nuits. *Academic Journal of Business & Management*,
6(7). <https://doi.org/10.25236/AJBM.2024.060723>