Development and Comparison of Forecasting Models for Gross Written Premiums of Life Insurance Companies in Thailand Using Data Mining Technique
การพัฒนาและเปรียบเทียบตัวแบบการพยากรณ์เบี้ยประกันภัยรับรวมของบริษัทประกันชีวิตในประเทศไทยโดยใช้เทคนิคเหมืองข้อมูล

Anupong Sukprasert[1,*], Oranan Samrual[2], and Ubonwan Wisetrat[3]
อนุพงศ์ สุขประเสริฐ[1,*], อรนันท์ สำรวล[2], และ อุบลวรรณ วิเศษรัตน[3]

Received: 15 March 2025;
Revised: 21 April 2025;
Accepted: 22 April 2025;
Published: 24 April 2025;

Abstract

Accurately forecasting gross written premiums (GWP) of life insurance companies is vital for strategic planning, risk management, and policy formulation in Thailand's insurance industry. Traditional statistical models commonly used in prior research often struggle to capture complex non-linear relationships in financial data. To address this limitation, the objective of this study is to develop and compare predictive models for GWP using data mining techniques. This study is among the first to evaluate model performance using a comprehensive monthly dataset covering a 12-year period from January 2012 to December 2023, totaling 144 records. The Cross-Industry Standard Process for Data Mining (CRISP-DM) methodology was adopted as the analytical framework. Four machine learning algorithms were implemented: Decision Tree, Random Forest, Support Vector Machine (SVM), and Neural Network. The empirical results reveal that the Neural Network model outperformed all other techniques, achieving a Mean Squared Error (MSE) of 1,174,870.83, a Root Mean Square Error (RMSE) of 1,083.91, a Mean Absolute Percentage Error (MAPE) of 2.53%, and a coefficient of determination (R^2) of 97.60%. These findings confirm the superior predictive accuracy of the Neural Network model. The study contributes to the advancement of premium forecasting by offering a robust, data-driven approach that supports strategic decision-making, enhances risk management practices, and improves the precision of premium calculation systems for life insurance companies in Thailand.

Keywords: Gross Written Premium, Data Mining, Insurance Risk Management, Economic Forecasting, Neural Networks

[1,*] Lecturer, Dr., Program in Business Computer, Mahasarakham Business School, Mahasarakham University, MahaSarakham 44150 Thailand; อาจารย์ ดร. สาขาคอมพิวเตอร์ธุรกิจ คณะการบัญชีและการจัดการ มหาวิทยาลัยมหาสารคาม จังหวัดมหาสารคาม 44150 ประเทศไทย; Email: anupong.s@acc.msu.ac.th

[2] Student, Program in Business Computer, Mahasarakham Business School, Mahasarakham University, MahaSarakham 44150 Thailand; นิสิตสาขาคอมพิวเตอร์ธุรกิจ คณะการบัญชีและการจัดการ มหาวิทยาลัยมหาสารคาม จังหวัดมหาสารคาม 44150 ประเทศไทย; Email: 64010912516@msu.ac.th

[3] Student, Program in Business Computer, Mahasarakham Business School, Mahasarakham University, MahaSarakham 44150 Thailand; นิสิตสาขาคอมพิวเตอร์ธุรกิจ คณะการบัญชีและการจัดการ มหาวิทยาลัยมหาสารคาม จังหวัดมหาสารคาม 44150 ประเทศไทย; Email: 64010912518@msu.ac.th

*Corresponding authors: Anupong Sukprasert (anupong.s@acc.msu.ac.th)

Citation: Sukprasert, A., Samrual, O., & Wisetrat, U. (2025). Development and Comparison of Forecasting Models for Gross Written Premiums of Life Insurance Companies in Thailand Using Data Mining Technique. *Journal of Computer and Creative Technology, 3*(1), 143-155. https://doi.org/10.14456/jcct.2025.12.

บทคัดย่อ

การพยากรณ์เบี้ยประกันภัยรับรวมของบริษัทประกันชีวิตอย่างแม่นยำ มีความสำคัญต่อการวางแผนกลยุทธ์ การบริหารความเสี่ยง และการกำหนดนโยบายในอุตสาหกรรมประกันชีวิตของประเทศไทย งานวิจัยที่ผ่านมาโดยมากมักใช้แบบจำลองทางสถิติดั้งเดิมซึ่งมีข้อจำกัดในการจัดการกับข้อมูลที่มีความสัมพันธ์แบบไม่เชิงเส้นอย่างซับซ้อน งานวิจัยนี้จึงมีวัตถุประสงค์เพื่อพัฒนาและเปรียบเทียบตัวแบบการพยากรณ์เบี้ยประกันภัยรับรวม โดยประยุกต์ใช้เทคนิคการทำเหมืองข้อมูล โดยใช้ชุดข้อมูลรายเดือนเป็นระยะเวลา 12 ปี ตั้งแต่เดือนมกราคม พ.ศ. 2555 ถึงเดือนธันวาคม พ.ศ. 2566 รวมจำนวนทั้งสิ้น 144 รายการ และวิเคราะห์ข้อมูลตามกระบวนการมาตรฐานการทำเหมืองข้อมูล (CRISP-DM) ตัวแบบที่ใช้ในการเปรียบเทียบประกอบด้วย 4 เทคนิค ได้แก่ เทคนิคต้นไม้ตัดสินใจ เทคนิคต้นไม้ป่าสุ่ม เทคนิคซัพพอร์ตเวกเตอร์แมชชีน และเทคนิคโครงข่ายประสาทเทียม ผลการวิเคราะห์แสดงให้เห็นว่า เทคนิคโครงข่ายประสาทเทียมมีความแม่นยำสูงที่สุด โดยมีค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (MSE) เท่ากับ 1,174,870.83 ค่ารากที่สองของค่าคลาดเคลื่อนกำลังสองเฉลี่ย (RMSE) เท่ากับ 1,083.91 ค่าเปอร์เซ็นต์ค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ย (MAPE) เท่ากับ 2.53% และค่าสัมประสิทธิ์การตัดสินใจ (R^2) เท่ากับ 97.60% ซึ่งสะท้อนถึงประสิทธิภาพในการพยากรณ์ที่เหนือกว่าเทคนิคอื่น ๆ ผลการศึกษานี้สามารถนำไปประยุกต์ใช้ในการพัฒนาระบบการพยากรณ์เบี้ยประกันภัยรับรวม เพื่อสนับสนุนการวางแผนเชิงกลยุทธ์ การบริหารความเสี่ยง และเพิ่มความแม่นยำของระบบการคำนวณเบี้ยประกันภัยของบริษัทประกันชีวิตในประเทศไทย

คำสำคัญ: เบี้ยประกันภัยรับรวม, การทำเหมืองข้อมูล, การบริหารความเสี่ยงในธุรกิจประกันภัย, การพยากรณ์ทางเศรษฐกิจ, โครงข่ายประสาทเทียม

1. บทนำ (Introduction)

ในปัจจุบัน การเปลี่ยนแปลงของสภาพเศรษฐกิจและสังคมส่งผลให้การประกันชีวิตมีบทบาทสำคัญมากขึ้น เนื่องจากหากเกิดการสูญเสียบุคคลในครอบครัว ย่อมส่งผลกระทบต่อสถานะทางการเงินของครอบครัวนั้นโดยตรง ดังนั้น การประกันชีวิตจึงเป็นกลไกที่ช่วยบรรเทาความเสี่ยงด้านภาระทางการเงินของครอบครัวผู้สูญเสียได้ในระดับหนึ่ง ทั้งนี้ สมาคมประกันชีวิตไทย (Thai Life Assurance Association, 2024e) ได้รายงานว่าในช่วงครึ่งแรกของปี 2567 (มกราคม – มิถุนายน) มีเบี้ยประกันภัยรับรวมอยู่ที่ 311,413.63 ล้านบาท ซึ่งเติบโตเพิ่มขึ้นร้อยละ 3.80 เมื่อเทียบกับช่วงเวลาเดียวกันของปี 2566 ที่มีเบี้ยประกันภัยรับรวม 300,005 ล้านบาท และมีอัตราการเติบโตอยู่ที่ร้อยละ 3.78 นอกจากนี้ เบี้ยประกันภัยรับรายใหม่มีมูลค่า 88,332.86 ล้านบาท โดยมีอัตราการเติบโตเพิ่มขึ้นร้อยละ 1.76 ขณะที่เบี้ยประกันภัยรับปีต่อไปอยู่ที่ 223,080.77 ล้านบาท ซึ่งมีอัตราการเติบโตอยู่ที่ร้อยละ 4.63 และมีอัตราความคงอยู่ของกรมธรรม์อยู่ที่ร้อยละ 83

ในธุรกิจประกันชีวิต ความแม่นยำในการพยากรณ์เบี้ยประกันภัยรับรวมมีความสำคัญต่อการวางแผนกลยุทธ์ การบริหารความเสี่ยง และการกำหนดนโยบายของบริษัทประกันชีวิต อย่างไรก็ตาม งานวิจัยที่ผ่านมาโดยมากมักใช้แบบจำลองทางสถิติดั้งเดิม เช่น การวิเคราะห์อนุกรมเวลา (Time Series Analysis) และการถดถอย (Regression Analysis) ซึ่งมีข้อจำกัดในการจัดการกับข้อมูลที่ซับซ้อนหรือไม่เป็นเชิงเส้น ในช่วงที่ผ่านมา เทคโนโลยีการทำเหมืองข้อมูล (Data Mining) และการเรียนรู้ของเครื่อง (Machine Learning) ได้กลายเป็นแนวทางที่มีประสิทธิภาพในการวิเคราะห์ข้อมูลเชิงลึก โดยเฉพาะในการคาดการณ์พฤติกรรมผู้บริโภค แนวโน้มตลาด และข้อมูลทางการเงิน ซึ่งช่วยเสริมความแม่นยำในการวางแผนกลยุทธ์และการบริหารความเสี่ยงของธุรกิจประกันชีวิต (Varadarajan & Kakumanu, 2024)

จากการตรวจสอบวรรณกรรมที่เกี่ยวกับการพยากรณ์เบี้ยประกันภัยรับรวมของบริษัทประกันชีวิตในประเทศไทย พบว่ายังมีการนำเทคนิค Machine Learning มาใช้โดยตรงไม่มากนัก ขณะเดียวกัน งานวิจัยของ Huadsri & Phaphan (2024) ได้ทำการศึกษาพัฒนาโมเดลพยากรณ์เบี้ยประกันชีวิตโดยใช้ข้อมูลรายเดือนตั้งแต่ปี พ.ศ. 2546 – พ.ศ. 2565 และ

Citation: Sukprasert, A., Samrual, O., & Wisetrat, U. (2025). Development and Comparison of Forecasting Models for Gross Written Premiums of Life Insurance Companies in Thailand Using Data Mining Technique. *Journal of Computer and Creative Technology, 3*(1), 143-155. https://doi.org/10.14456/jcct.2025.12.

เปรียบเทียบประสิทธิภาพของโมเดล SARIMAX กับเทคนิค Machine Learning อื่น ๆ เช่น Random Forest และ Support Vector Regression (SVR) โดยพบว่า SVR มีความแม่นยำสูงสุดในการพยากรณ์เบี้ยประกันชีวิต นอกจากนี้ ความคลาดเคลื่อนในการพยากรณ์เบี้ยประกันภัยรับรวมสามารถส่งผลกระทบต่อประสิทธิภาพทางการเงินของบริษัทประกันชีวิต โดยอาจนำไปสู่การตั้งสำรองและการกำหนดราคาเบี้ยประกันที่ไม่สอดคล้องกับความเป็นจริง ซึ่งอาจกระทบต่อเสถียรภาพทางการเงินในระยะยาวของธุรกิจ

ด้วยเหตุนี้ งานวิจัยนี้จึงมีแนวคิดในการพัฒนาตัวแบบสำหรับพยากรณ์เบี้ยประกันภัยรับรวม โดยอาศัยเทคนิคการทำเหมืองข้อมูล เพื่อให้ได้แบบจำลองที่สามารถคาดการณ์แนวโน้มการเปลี่ยนแปลงของเบี้ยประกันภัยรับรวมได้อย่างแม่นยำ ซึ่งสามารถนำไปใช้เป็นแนวทางในการวางแผนกลยุทธ์และบริหารจัดการธุรกิจประกันชีวิตในประเทศไทยให้มีประสิทธิภาพสูงสุดในระยะยาว

2. วัตถุประสงค์งานวิจัย (Research Objectives)

1. เพื่อสร้างตัวแบบสำหรับพยากรณ์เบี้ยประกันภัยรับรวม ของบริษัทประกันชีวิตในประเทศไทย
2. เพื่อเปรียบเทียบประสิทธิภาพของตัวแบบสำหรับพยากรณ์เบี้ยประกันภัยรับรวม ของบริษัทประกันชีวิตในประเทศไทย

3. การทบทวนวรรณกรรมและทฤษฎีที่เกี่ยวข้อง (Literature Review)

เบี้ยประกันภัยรับรวม (Gross Written Premium: GWP) ถือเป็นแหล่งรายได้หลักของบริษัทประกันชีวิต โดยเฉพาะกรมธรรม์ประเภทสามัญที่มุ่งเน้นการคุ้มครองระยะยาวและมีมูลค่าเอาประกันภัยสูง เหมาะสำหรับผู้มีรายได้ปานกลางถึงสูง (Thai Reinsurance Public Company Limited, 2020) การทำความเข้าใจปัจจัยที่ส่งผลต่อ GWP จึงมีความสำคัญต่อการวางแผนกลยุทธ์และการบริหารความเสี่ยงของบริษัทประกันชีวิต

งานวิจัยในอดีตเกี่ยวกับการพยากรณ์เบี้ยประกันภัยรับรวมของบริษัทประกันชีวิตในประเทศไทยส่วนใหญ่ยังคงใช้แบบจำลองทางสถิติดั้งเดิม เช่น อนุกรมเวลาและการถดถอยเชิงเส้น ซึ่งมีข้อจำกัดในการประมวลผลข้อมูลที่มีความซับซ้อนหรือไม่เป็นเชิงเส้น (Hyndman & Athanasopoulos, 2021) จากการสำรวจช่องว่างของวรรณกรรม พบว่ายังไม่มีงานวิจัยใดที่วิเคราะห์ปัจจัยเศรษฐกิจมหภาคและตัวชี้วัดทางการเงินจำนวน 12 ตัวแปรอย่างครบถ้วนในการพยากรณ์เบี้ยประกันภัยรับรวม ได้แก่ วันที่ ดัชนีราคาผู้บริโภค ดัชนีความเชื่อมั่นทางธุรกิจ ปริมาณเงินความหมายกว้าง อัตราดอกเบี้ยเงินฝากประจำ ดัชนีราคาตลาดหลักทรัพย์ อัตราผลตอบแทนพันธบัตรรัฐบาล อัตราการหมุนของสินทรัพย์ อัตราผลตอบแทนต่อสินทรัพย์ อัตราเบี้ยประกันรับสุทธิ และอัตราส่วนหนี้สินต่อส่วนของผู้ถือหุ้น

นอกจากนี้ Phaphan & Puttamat (2023) ศึกษาการใช้โมเดลต้นไม้ตัดสินใจ (Decision Tree) ในการพยากรณ์เบี้ยประกันภัยรับรวมจากข้อมูลในช่วงปี พ.ศ. 2560 ถึงปี พ.ศ. 2565 ซึ่งมีข้อจำกัดด้านช่วงเวลาและเทคนิคที่ใช้ โดยยังไม่ได้เปรียบเทียบกับเทคนิคการเรียนรู้ของเครื่อง (Machine Learning) อื่น ๆ ที่มีศักยภาพสูงกว่า

ดังนั้น งานวิจัยนี้จึงมุ่งวิเคราะห์ปัจจัยเศรษฐกิจมหภาคและตัวชี้วัดทางการเงินที่ส่งผลต่อเบี้ยประกันภัยรับรวม พร้อมเปรียบเทียบประสิทธิภาพของโมเดล ได้แก่ เทคนิคต้นไม้ตัดสินใจ (Decision Tree) เทคนิคต้นไม้ป่าสุ่ม (Random Forest) เทคนิคซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine: SVM) และเทคนิคโครงข่ายประสาทเทียม (Neural Network) ภายใต้กรอบกระบวนการวิเคราะห์ข้อมูลตามมาตรฐานในการทำเหมืองข้อมูล (CRISP-DM) โดยใช้ข้อมูลรายเดือนในช่วงเวลา 12 ปี ซึ่งนับเป็นการประยุกต์ใช้เทคนิคเหมืองข้อมูลกับปัจจัยเศรษฐกิจมหภาคของไทยอย่างเป็นระบบ เพื่อสนับสนุนการตัดสินใจเชิงกลยุทธ์ในภาคธุรกิจประกันชีวิต

4. วิธีดำเนินงานวิจัย (Research Methodology)

งานวิจัยนี้ดำเนินการวิเคราะห์ข้อมูลตามกระบวนการมาตรฐานของการทำเหมืองข้อมูล (Cross - Industry Standard Process for Data Mining: CRISP - DM) ซึ่งเป็นแนวทางที่ได้รับความนิยมและยอมรับอย่างแพร่หลายในงานวิเคราะห์ข้อมูลเชิงธุรกิจ (Sukprasert, 2023) โดยกระบวนการ CRISP-DM ประกอบด้วย 6 ขั้นตอนหลัก ได้แก่ (1) การทำความเข้าใจเกี่ยวกับปัญหา (Problem Understanding) (2) การทำความเข้าใจข้อมูล (Data Understanding) (3) การ

Citation:
Sukprasert, A., Samrual, O., & Wisetrat, U. (2025). Development and Comparison of Forecasting Models for Gross Written Premiums of Life Insurance Companies in Thailand Using Data Mining Technique. *Journal of Computer and Creative Technology, 3*(1), 143-155. https://doi.org/10.14456/jcct.2025.12.

เตรียมข้อมูล (Data Preparation) (4) การสร้างตัวแบบ (Modeling) (5) การประเมินผลตัวแบบ (Evaluation) และ (6) การนำตัวแบบไปประยุกต์ใช้งาน (Deployment)

ในการวิจัยนี้ กระบวนการ CRISP-DM ถูกนำมาใช้เพื่อสร้างและเปรียบเทียบประสิทธิภาพของตัวแบบการพยากรณ์เบี้ยประกันภัยรับรวมของบริษัทประกันชีวิตในประเทศไทยอย่างเป็นระบบ โดยมีการนำเทคนิคการเรียนรู้ของเครื่อง (Machine Learning) ที่เหมาะสมมาใช้ในขั้นตอนการสร้างตัวแบบ และดำเนินการประเมินผลด้วยเกณฑ์มาตรฐานทางสถิติ เพื่อเลือกตัวแบบที่มีประสิทธิภาพสูงสุด รายละเอียดของกระบวนการวิจัยตลอดทั้ง 6 ขั้นตอน ดังนี้

4.1 การทำความเข้าใจเกี่ยวกับปัญหา (Problem Understanding)

จากการศึกษาข้อมูลเกี่ยวกับเบี้ยประกันภัยรับรวม พบว่าในปี พ.ศ. 2567 สมาคมประกันชีวิตไทยได้ประมาณการว่าอัตราการเติบโตของธุรกิจประกันชีวิตจะอยู่ในช่วงร้อยละ 2.0 – 4.0 พร้อมกันนี้ สมาคมฯ ยังได้กำหนดนโยบายที่ส่งเสริมให้บริษัทประกันชีวิตดำเนินธุรกิจโดยให้ความสำคัญกับการบริหารจัดการความเสี่ยงอย่างรอบด้าน ทั้งในช่วงก่อนการรับประกันภัย (Underwriting) และหลังจากเริ่มความคุ้มครอง ตลอดจนการรักษาสถานะทางการเงินให้มั่นคง ผ่านการดำรงเงินกองทุนที่มีอัตราส่วนความเพียงพอของเงินกองทุนตามความเสี่ยง (Capital Adequacy Ratio: CAR Ratio) สูงกว่าระดับขั้นต่ำที่กำหนดโดยหน่วยงานกำกับดูแล (Supervisory CAR) ทั้งนี้ เพื่อสร้างความเชื่อมั่นแก่ผู้เอาประกันภัย และเพื่อให้บริษัทประกันชีวิตสามารถให้ความคุ้มครองได้อย่างต่อเนื่องตลอดอายุสัญญา (Bangkok Biz News, 2024)

จากบริบทดังกล่าว งานวิจัยนี้จึงมีความสนใจในการประยุกต์ใช้เทคนิคการทำเหมืองข้อมูล (Data Mining) เพื่อพัฒนาตัวแบบสำหรับการพยากรณ์เบี้ยประกันภัยรับรวมของบริษัทประกันชีวิตในประเทศไทย ซึ่งจะช่วยให้สามารถวางแผนเชิงกลยุทธ์ได้อย่างมีประสิทธิภาพ เพิ่มขีดความสามารถในการแข่งขัน และสามารถประเมินความเสี่ยงที่อาจเกิดขึ้นได้อย่างรอบด้าน ทั้งในมิติของเศรษฐกิจมหภาคและการแข่งขันภายในอุตสาหกรรมประกันชีวิต

4.2 การทำความเข้าใจเกี่ยวกับข้อมูล (Data Understanding)

งานวิจัยนี้ได้ศึกษาและรวบรวมข้อมูลที่เกี่ยวข้องเพื่อนำมาใช้ในการพัฒนาตัวแบบพยากรณ์เบี้ยประกันภัยรับรวมโดยประยุกต์ใช้เทคนิคการทำเหมืองข้อมูล (Data Mining) ซึ่งในการศึกษาครั้งนี้ได้พิจารณาตัวแปรต้นทั้งหมดจำนวน 12 ตัวแปร โดยรวบรวมจากแหล่งข้อมูลที่เชื่อถือได้ ได้แก่ ดัชนีราคาผู้บริโภค (Consumer Price Index: CPI) จากสำนักงานนโยบายและยุทธศาสตร์การค้า (Trade Policy and Strategy Office, 2024) ดัชนีความเชื่อมั่นทางธุรกิจ ปริมาณเงินความหมายกว้าง (M2) และอัตราดอกเบี้ยเงินฝากประจำ จากธนาคารแห่งประเทศไทย จากธนาคารแห่งประเทศไทย (Bank of Thailand, 2024a; Bank of Thailand, 2024b; Bank of Thailand, 2024c) ดัชนีราคาตลาดหลักทรัพย์ (SET Index, 2024) อัตราผลตอบแทนพันธบัตรรัฐบาล จากเว็บไซต์ Investing (2024) อัตราการหมุนเวียนของสินทรัพย์ (Asset Turnover) (Thai Life Assurance Association, 2024a) อัตราผลตอบแทนต่อสินทรัพย์ (Return on Assets) (Thai Life Assurance Association, 2024f) อัตราเบี้ยประกันรับสุทธิ (Net Premium Ratio) (Thai Life Assurance Association, 2024d) และอัตราส่วนหนี้สินต่อส่วนของผู้ถือหุ้น (Debt-to-Equity Ratio) (Thai Life Assurance Association, 2024b) รวมถึงตัวแปรตาม คือ เบี้ยประกันภัยรับรวมประเภทสามัญ จากสมาคมประกันชีวิตไทย (Thai Life Assurance Association, 2024c)

ข้อมูลทั้งหมดถูกจัดเก็บในรูปแบบของข้อมูลรายเดือนเป็นระยะเวลา 12 ปี ครอบคลุมระยะเวลาตั้งแต่เดือนมกราคม พ.ศ. 2555 ถึงเดือนธันวาคม พ.ศ. 2566 รวมทั้งสิ้น 144 แถว (Row) และ 12 คอลัมน์ (Column) ซึ่งจัดเก็บไว้ในรูปแบบไฟล์ Excel ภายใต้ชื่อชุดข้อมูล “Gross Written Premium” เพื่อใช้ในการวิเคราะห์และพัฒนาแบบจำลองเชิงพยากรณ์อย่างเป็นระบบ

4.3 การเตรียมข้อมูล (Data Preparation)

การเตรียมข้อมูลเป็นกระบวนการสำคัญที่ช่วยจัดการกับข้อมูลดิบ (Raw Data) เพื่อให้มีความถูกต้อง น่าเชื่อถือ และมีคุณภาพเพียงพอสำหรับการนำไปใช้ในการวิเคราะห์ โดยเฉพาะอย่างยิ่งในการพัฒนาตัวแบบการเรียนรู้ของเครื่องที่ต้องอาศัยข้อมูลที่สะอาดและสอดคล้องกัน งานวิจัยนี้ดำเนินการเตรียมข้อมูลก่อนการวิเคราะห์โดยแบ่งออกเป็น 3 ขั้นตอนหลัก ดังนี้

1. เนื่องจากข้อมูลตัวแปรเศรษฐกิจและการเงินที่ใช้ในการวิเคราะห์ถูกรวบรวมจากหลายแหล่ง ดังนั้นจึงดำเนินการบูรณาการข้อมูล (Data Integration) โดยนำข้อมูลทั้งหมดมารวมกันให้อยู่ในรูปแบบเดียวกันในไฟล์ Excel เพื่อให้พร้อมต่อการนำไปวิเคราะห์ในลำดับถัดไป ข้อมูลดังกล่าวอยู่ในรูปแบบของข้อมูลรายเดือน ครอบคลุมระยะเวลา 12 ปี

Citation: Sukprasert, A., Samrual, O., & Wisetrat, U. (2025). Development and Comparison of Forecasting Models for Gross Written Premiums of Life Insurance Companies in Thailand Using Data Mining Technique. *Journal of Computer and Creative Technology, 3*(1), 143-155. https://doi.org/10.14456/jcct.2025.12.

ตั้งแต่เดือนมกราคม พ.ศ. 2555 ถึงเดือนธันวาคม พ.ศ. 2566 รวมทั้งสิ้น 144 แถว (Row) และ 12 คอลัมน์ (Column) ประกอบด้วยตัวแปรเศรษฐกิจ 11 ตัวแปร และตัวแปรตาม คือ เบี้ยประกันภัยรับรวม (Gross Written Premium) โดยชุดข้อมูลนี้ถูกจัดเก็บภายใต้ชื่อไฟล์ว่า “Gross Written Premium” ดังแสดงใน Table 1.

2. เพื่อให้ข้อมูลมีคุณภาพเหมาะสมต่อการวิเคราะห์ จากนั้นได้ดำเนินการทำความสะอาดข้อมูลด้วยการตรวจสอบค่าที่ขาดหาย (Missing Values) และค่าผิดปกติ (Outliers) ในเบื้องต้น โดยใช้ฟังก์ชัน ISBLANK() ใน Excel ร่วมกับการตรวจสอบด้วย Boxplot ผลการตรวจสอบพบว่า ข้อมูลไม่มีค่าที่ขาดหายและไม่พบ Outliers ที่ส่งผลกระทบต่อการวิเคราะห์ จึงสามารถนำข้อมูลเข้าสู่กระบวนการวิเคราะห์ได้ทันทีโดยไม่ต้องใช้วิธีแทนค่าหรือการตัดค่าผิดปกติออก หลังจากนั้น ได้มีการกำหนดชนิดของข้อมูล (Data Type) ให้สอดคล้องกับลักษณะของตัวแปรแต่ละตัว โดยคอลัมน์วันที่ (Date) ถูกกำหนดเป็นชนิดข้อมูลแบบวันเวลา (Date Format) เพื่อรองรับการวิเคราะห์เชิงเวลา ตัวแปรที่เป็นค่าจำนวนเต็ม เช่น ปริมาณเงินความหมายกว้าง (M2) ถูกกำหนดเป็น Integer ขณะที่ตัวแปรเชิงเศรษฐกิจอื่น ๆ รวมถึงเบี้ยประกันภัยรับรวม (GWP) ถูกกำหนดเป็นชนิดข้อมูลแบบเลขจำนวนจริง (Real/Float) เพื่อให้รองรับการวิเคราะห์ด้วยอัลกอริธึมการเรียนรู้ของเครื่องได้อย่างเหมาะสมและแม่นยำ

3. ในการวิเคราะห์ข้อมูล ได้กำหนดบทบาทของตัวแปรแต่ละตัวให้มีหน้าที่เฉพาะอย่างเหมาะสมกับโครงสร้างของการพัฒนาแบบจำลองการพยากรณ์ โดยแบ่งออกเป็น 3 ประเภท ได้แก่ ตัวระบุ (Identifier), ตัวแปรตาม (Dependent Variable) และตัวแปรต้น (Independent Variables)

แอตทริบิวต์ “วันที่” (Date) ถูกกำหนดให้ทำหน้าที่เป็น ตัวระบุ (Identifier) ซึ่งใช้สำหรับจัดเรียงลำดับข้อมูลตามลำดับเวลา โดยไม่ถูกนำเข้าเป็นส่วนหนึ่งของการสร้างแบบจำลองโดยตรง เพื่อป้องกันการแทรกแซงเชิงโครงสร้างในกระบวนการเรียนรู้ของอัลกอริธึม

แอตทริบิวต์ “เบี้ยประกันภัยรับรวม” (Gross Written Premium: GWP) ทำหน้าที่เป็น ตัวแปรตาม (Dependent Variable หรือ Label) ซึ่งเป็นเป้าหมายของการพยากรณ์ในงานวิจัยนี้ และใช้เป็นตัวชี้วัดในการประเมินความแม่นยำของแบบจำลอง

ส่วนแอตทริบิวต์ที่เหลืออีก 10 ตัวแปร ซึ่งประกอบด้วยดัชนีเศรษฐกิจมหภาคและตัวชี้วัดทางการเงิน เช่น CPI, M2, อัตราดอกเบี้ย, ดัชนี SET, และอัตราผลตอบแทนทางการเงินของบริษัทประกันภัย ถูกกำหนดให้เป็น ตัวแปรต้น (Independent Variables หรือ Predictors) ที่ใช้ในการวิเคราะห์เพื่อพยากรณ์ค่าของ GWP โดยผ่านอัลกอริธึมของเทคนิคการเรียนรู้ของเครื่องที่ใช้ในการศึกษา

Table 1. Data on Factors Affecting Gross Written Premiums

Name	Description	Data Type	Role
Date	Date	Date	Identifier
CPI	Consumer Price Index	Real	Independent Variables
BCI	Business Confidence Index	Real	Independent Variables
M2	Broad money supply (1+2)	Integer	Independent Variables
Interest rates	Fixed Deposit interest rates	Real	Independent Variables
SE Price Index	Stock Exchange of Thailand Price Index	Real	Independent Variables
GBY	Government bond yields	Real	Independent Variables
ATR	Asset turnover rate	Real	Independent Variables
ROA	Return on assets	Real	Independent Variables
NIP	Net insurance premium received	Real	Independent Variables
DTER	Debt to equity ratio	Real	Independent Variables
GWP	Gross Written Premium	Real	Dependent Variable

Citation:
Sukprasert, A., Samrual, O., & Wisetrat, U. (2025). Development and Comparison of Forecasting Models for Gross Written Premiums of Life Insurance Companies in Thailand Using Data Mining Technique. *Journal of Computer and Creative Technology, 3*(1), 143-155. https://doi.org/10.14456/jcct.2025.12.

4.4 การสร้างแบบจำลอง (Modeling)

ภายหลังจากกระบวนการเตรียมข้อมูลเสร็จสิ้น ต่อมาได้ดำเนินการสร้างแบบจำลองเพื่อพยากรณ์เบี้ยประกันภัยรับรวมของบริษัทประกันชีวิตในประเทศไทย โดยอ้างอิงจากวรรณกรรมที่เกี่ยวข้อง พบว่า Hangkaew et al. (2020) ได้ศึกษาการทำนายการเลือกประเภทประกันภัยของลูกค้า โดยประยุกต์ใช้เทคนิคการทำเหมืองข้อมูล เช่น การถดถอยลอจิสติกพหุ (Multinomial Logistic Regression), การวิเคราะห์จำแนกกลุ่ม (Discriminant Analysis), ต้นไม้ตัดสินใจ (Decision Tree) และโครงข่ายประสาทเทียม (Neural Network) นอกจากนี้ Phaphan & Puttamat (2023) ได้ศึกษาเปรียบเทียบแบบจำลองการพยากรณ์เบี้ยประกันภัยรับรวม โดยใช้เทคนิค Decision Tree Regression, Random Forest Regression, Support Vector Machine Regression (SVR) และ Polynomial Regression

จากวรรณกรรมที่เกี่ยวข้องดังกล่าว งานวิจัยนี้จึงเลือกใช้เทคนิคพื้นฐานจำนวน 4 เทคนิค ซึ่งได้รับความนิยมในงานพยากรณ์เชิงธุรกิจ ได้แก่ ต้นไม้ตัดสินใจ (Decision Tree) ต้นไม้ป่าสุ่ม (Random Forest) ซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine: SVM) และโครงข่ายประสาทเทียม (Neural Network) ทั้งนี้ เทคนิคทั้ง 4 มีคุณสมบัติที่สามารถจัดการกับข้อมูลที่มีความไม่เป็นเชิงเส้น มีความสามารถในการเรียนรู้แบบจำแนกและถดถอย และสามารถปรับใช้กับข้อมูลที่มีหลายมิติได้อย่างยืดหยุ่น (Zhou, 2021; Chowdhury et al., 2024) การเปรียบเทียบประสิทธิภาพของแต่ละเทคนิคจึงช่วยให้สามารถเลือกแบบจำลองที่แม่นยำและเหมาะสมที่สุดเพื่อพัฒนาแบบจำลองสำหรับการพยากรณ์เบี้ยประกันภัยรับรวมของบริษัทประกันชีวิตในประเทศไทย โดยรายละเอียดของแต่ละแบบจำลองแสดงดังต่อไปนี้

1. เทคนิคต้นไม้ตัดสินใจ (Decision Tree) เป็นวิธีที่เข้าใจง่ายและได้รับความนิยม มีการเรียนรู้แบบมีผู้สอน ใช้แก้ปัญหาการจำแนกประเภท (Classification) และการถดถอย (Regression) โดยส่วนใหญ่นิยมใช้ในการจำแนกประเภท ต้นไม้ตัดสินใจประกอบด้วย 3 ส่วนหลักคือ ปม (Node) กิ่ง (Branch) และใบ (Leaf) ปมราก (Root Node) เป็นปมแรกในการตัดสินใจ ผลลัพธ์จะแทนด้วยใบ และกิ่งจะแสดงค่าของแต่ละคุณลักษณะ (Kittibawornkul et al., 2022) เทคนิคต้นไม้ตัดสินใจทำงานโดยการแบ่งข้อมูลออกเป็นกลุ่มย่อยอย่างเป็นลำดับชั้น (Hierarchical Splitting) โดยอาศัยค่าเกณฑ์ เช่น Mean Squared Error (MSE) เป็นตัวชี้วัดในการตัดสินใจแบ่งโหนดแต่ละขั้น เพื่อให้ค่าคลาดเคลื่อนภายในกลุ่มย่อยลดลงมากที่สุด เป้าหมายของแบบจำลองคือการสร้างกิ่งและใบที่สามารถทำนายค่าของตัวแปรตามได้แม่นยำที่สุด

2. เทคนิคต้นไม้ป่าสุ่ม (Random Forest) เป็นอัลกอริธึมแบบ Ensemble Learning ที่สร้าง Decision Trees หลายต้นขึ้นมาโดยใช้การสุ่มข้อมูลและสุ่มตัวแปร (Bootstrap Sampling และ Feature Randomness) แบบจำลองจะทำนายค่าด้วยการเฉลี่ยผลลัพธ์จากต้นไม้ทั้งหมดเพื่อลดความผันผวนและความเอนเอียง ช่วยเพิ่มความแม่นยำในการพยากรณ์และลด Overfitting ที่อาจเกิดขึ้นจากการใช้ Decision Tree เพียงต้นเดียว แม้ว่าจะเป็นเทคนิคต้นไม้ตัดสินใจ (Decision Tree) เหมือนกันแต่ข้อมูลและแอตทริบิวต์ที่ใช้ในการสร้างโมเดลต่างกันก็ทำให้โมเดลที่สร้างขึ้นมาได้มีลักษณะที่ต่างกัน (Wanon & Muangsan, 2021)

3. เทคนิคซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine: SVM) เป็นแบบจำลองการเรียนรู้ของเครื่องแบบมีผู้สอน (Supervised Learning) ที่นิยมใช้สำหรับการจำแนกประเภท (Classification) โดยมีหลักการทำงานคือการสร้างเส้นแบ่งเขต (Decision Boundary) หรือเส้นจำแนกประเภท (Classifier) ที่เกิดขึ้น (Sanguansat, 2019) สำหรับกรณีของปัญหาการพยากรณ์เชิงปริมาณ (Regression) อัลกอริธึม SVM ถูกปรับใช้ในรูปแบบที่เรียกว่า Support Vector Regression (SVR) ซึ่งมีหลักการพยายามหาฟังก์ชันที่สามารถประมาณค่าความสัมพันธ์ระหว่างตัวแปรต้นและตัวแปรตามได้แม่นยำ โดยไม่จำเป็นต้องครอบคลุมจุดข้อมูลทั้งหมด แต่จะให้ความสำคัญกับข้อมูลที่อยู่นอกช่วงขอบเขตของค่าความคลาดเคลื่อนที่ยอมรับได้ (เรียกว่า Epsilon Margin) นอกจากนี้ยังสามารถประยุกต์ใช้ Kernel Trick เพื่อเปลี่ยนข้อมูลที่ไม่สามารถแยกเส้นตรงได้ในพิกัดเดิม ให้กลายเป็นเส้นตรงในมิติที่สูงขึ้น (Higher-dimensional Feature Space) ซึ่งทำให้สามารถเรียนรู้ความสัมพันธ์เชิงซับซ้อนและไม่เป็นเชิงเส้นได้อย่างมีประสิทธิภาพ จึงเหมาะอย่างยิ่งสำหรับการพยากรณ์ข้อมูลทางเศรษฐกิจและการเงินที่มีลักษณะไม่เป็นเชิงเส้น

4. เทคนิคโครงข่ายประสาทเทียม (Neural Network) เป็นอัลกอริธึมที่จำลองการทำงานของเซลล์ประสาททางชีววิทยา โดยโครงสร้างของเครือข่ายประกอบด้วยเลเยอร์นำเข้า เลเยอร์ซ่อน และเลเยอร์ส่งออก ซึ่งเชื่อมต่อกันด้วยค่าถ่วงน้ำหนัก (Weights) ที่สามารถปรับเปลี่ยนได้ระหว่างการเรียนรู้ สำหรับโครงข่ายแบบเพอร์เซปตรอนหลายชั้น (Multilayer Perceptron: MLP) การประมวลผลในแต่ละเลเยอร์จะผ่านฟังก์ชันกระตุ้น เช่น ReLU หรือ Sigmoid และปรับค่าถ่วง

Citation: Sukprasert, A., Samrual, O., & Wisetrat, U. (2025). Development and Comparison of Forecasting Models for Gross Written Premiums of Life Insurance Companies in Thailand Using Data Mining Technique. *Journal of Computer and Creative Technology, 3*(1), 143-155. https://doi.org/10.14456/jcct.2025.12.

น้ำหนักด้วยวิธี Backpropagation เพื่อให้ผลลัพธ์เข้าใกล้ค่าจริงมากที่สุด เทคนิคนี้สามารถจัดการกับข้อมูลที่ไม่เป็นเชิงเส้นและมีความซับซ้อนได้ดี จึงเหมาะสำหรับปัญหาการพยากรณ์ (Sangsri & Chaiwuttisak, 2021)

4.5 การประเมินผล (Evaluation))

งานวิจัยนี้ได้ดำเนินการแบ่งชุดข้อมูลออกเป็น 2 ส่วนหลัก ได้แก่ ชุดข้อมูลสำหรับการฝึกสอนแบบจำลอง (Training Set) และชุดข้อมูลสำหรับการทดสอบ (Testing Set) โดยใช้เทคนิคการตรวจสอบความถูกต้องแบบไขว้ (Cross Validation) เพื่อเพิ่มความน่าเชื่อถือของการประเมินผล ซึ่งในงานวิจัยนี้ใช้วิธีการแบ่งข้อมูลออกเป็น 10 ส่วน หรือ 10-Fold Cross Validation โดยจะทำการฝึกแบบจำลอง 10 รอบ โดยในแต่ละรอบจะสุ่มแบ่งข้อมูลออกเป็น 9 ส่วนเพื่อฝึก และอีก 1 ส่วนเพื่อทดสอบ แล้วหมุนเวียนจนครบทั้ง 10 รอบก่อนจะนำผลเฉลี่ยมาใช้ในการประเมินแบบจำลอง เพื่อให้ผลการประเมินมีความเสถียรและลดอคติจากการสุ่มแบ่งข้อมูล สำหรับเกณฑ์ที่ใช้ในการประเมินประสิทธิภาพของตัวแบบประกอบด้วยตัวชี้วัดหลักจำนวน 5 ตัว ได้แก่ ค่ารากที่สองของค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE) ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Squared Error: MSE) สัมประสิทธิ์การตัดสินใจ (Coefficient of Determination: R-squared หรือ R^2) ค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ยในรูปเปอร์เซ็นต์ (Mean Absolute Percentage Error: MAPE) และค่าสัมประสิทธิ์สหสัมพันธ์ (Correlation Coefficient) การเลือกใช้เกณฑ์ทั้ง 5 ตัวชี้วัดดังกล่าว มีวัตถุประสงค์เพื่อให้การประเมินประสิทธิภาพของแบบจำลองมีความครอบคลุมทั้งในด้านระดับความแม่นยำ (Accuracy) ความสัมพันธ์เชิงเส้นระหว่างค่าพยากรณ์กับค่าจริง (Relationship) และความคลาดเคลื่อนเชิงสัมพัทธ์ (Relative Error) ซึ่งช่วยสะท้อนความสามารถของแบบจำลองในการพยากรณ์ข้อมูลเชิงปริมาณได้อย่างรอบด้าน RMSE และ MSE เป็นตัวชี้วัดที่ใช้วัดค่าความคลาดเคลื่อนเชิงสัมบูรณ์ โดย RMSE ให้ค่าน้ำหนักกับค่าคลาดเคลื่อนที่มากกว่าค่าเฉลี่ย ในขณะที่ MSE แสดงค่าความผิดพลาดในรูปกำลังสองซึ่งนิยมใช้ในการเปรียบเทียบระหว่างแบบจำลอง R-squared ใช้วัดความสามารถของแบบจำลองในการอธิบายความแปรปรวนของข้อมูลได้มากน้อยเพียงใด ยิ่งค่า R^2 เข้าใกล้ 1 ยิ่งสะท้อนว่าแบบจำลองมีประสิทธิภาพสูง ส่วน MAPE เป็นตัวชี้วัดที่นิยมใช้ในบริบทที่ต้องการพิจารณาความคลาดเคลื่อนในเชิงสัดส่วน ทำให้สามารถเปรียบเทียบแบบจำลองข้ามหน่วยวัดได้อย่างมีประสิทธิภาพ สุดท้าย Correlation Coefficient เป็นค่าที่สะท้อนความสัมพันธ์เชิงเส้นระหว่างค่าที่แบบจำลองพยากรณ์ได้กับค่าจริง โดยยิ่งมีค่าใกล้ 1 มากเท่าใด ยิ่งสะท้อนว่าแบบจำลองมีแนวโน้มในการพยากรณ์ได้ใกล้เคียงกับความเป็นจริง โดยกระบวนการประเมินผลทั้งหมดดำเนินการผ่านโปรแกรม RapidMiner Studio Version 10.3 ซึ่งรองรับทั้งการวิเคราะห์แบบจำลอง การประเมินผลด้วยตัวชี้วัดทางสถิติ และการเปรียบเทียบประสิทธิภาพของแบบจำลองอย่างเป็นระบบ (Sukprasert, 2023)

1. ค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Squared Error: RMSE) เป็นการวัดความคลาดเคลื่อนระหว่างค่าจริงและค่าพยากรณ์ โดยคล้ายกับ MSE แต่ RMSE เพิ่มน้ำหนักให้กับความคลาดเคลื่อนขนาดใหญ่ด้วยการหาค่ารากที่สองของค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง ค่า RMSE สามารถเขียนอธิบายในรูปสมการที่ (1) ดังนี้

$$\mathrm{RMSE}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2} \quad \text{... (1)}$$

โดยที่ n คือ จำนวนตัวอย่างที่ได้ทั้งหมดที่ใช้ในการศึกษา
y_i คือ ค่าจริง (Actual Data)
$\hat{y}_i$ คือ ค่าพยากรณ์ (Predictive value)

2. ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Squared Error: MSE) เป็นการวัดค่าความคลาดเคลื่อน โดยการนำค่าความคลาดเคลื่อนมายกกำลังแล้วนำไปหาค่าเฉลี่ย ในการวัดค่าความแม่นยำจากวิธีการนี้ยิ่งค่าที่ได้มีค่าน้อยแสดงว่าโมเดลที่ได้นั้นจะมีความแม่นยำมาก ค่า MSE สามารถเขียนอธิบายในรูปสมการที่ (2) ดังนี้

$$\mathrm{MSE}=\frac{1}{n}\sum_{i=0}^{n}(y_i-\widehat{y_i})^2 \quad \text{... (2)}$$

โดยที่ n คือ จำนวนตัวอย่างที่ได้ทั้งหมดที่ใช้ในการศึกษา
y_i คือ ค่าจริง (Actual Data)
$\hat{y}_i$ คือ ค่าพยากรณ์ (Predictive value)

Citation:
Sukprasert, A., Samrual, O., & Wisetrat, U. (2025). Development and Comparison of Forecasting Models for Gross Written Premiums of Life Insurance Companies in Thailand Using Data Mining Technique. *Journal of Computer and Creative Technology, 3*(1), 143-155. https://doi.org/10.14456/jcct.2025.12.

3. สัมประสิทธิ์การตัดสินใจ (Coefficient of Discrimination) หรือ R-squared (R^2) คือ ตัวบ่งชี้ว่าโมเดลการประมาณค่าข้อมูล มีความเหมาะสมกับข้อมูลมากน้อยเพียงใด ซึ่งจะแสดงถึงขอบเขตความสามารถในการคาดเดาการแปรผันของตัวแปรตามของโมเดล ค่า R-squared (R^2) สามารถเขียนอธิบายในรูปสมการที่ (3) ดังนี้

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \quad \text{... (3)}$$

โดยที่ n คือ จำนวนตัวอย่างที่ได้ทั้งหมดที่ใช้ในการศึกษา
y_i คือ ค่าจริง (Actual Data)
$\hat{y}_i$ คือ ค่าพยากรณ์ (Predictive value)
$\bar{y}$ คือ ค่าเฉลี่ยผลลัพธ์ที่สังเกตได้

4. เปอร์เซ็นค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ย (Mean Absolute Percentage Error: MAPE) เป็นวิธีการหาค่าเปอร์เซ็นเฉลี่ยของความแตกต่างสมบูรณ์ระหว่างค่าจริงและค่าพยากรณ์ โดยเป็นการแสดงความถูกต้องของการพยากรณ์ในรูปของร้อยละของค่าคลาดเคลื่อน ซึ่งการเป็นการแปลงข้อมูลให้อยู่ในรูปแบบมาตรฐานเพื่อให้เข้ากับขนาดของค่าสังเกต ซึ่งจะช่วยลดความเบ้ลง ค่า MAPE สามารถเขียนอธิบายในรูปสมการที่ (4) ดังนี้

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\frac{|y_i - \hat{y}_i|}{|y_i|} \times 100\% \quad \text{... (4)}$$

โดยที่ n คือ จำนวนตัวอย่างที่ได้ทั้งหมดที่ใช้ในการศึกษา
y_i คือ ค่าจริง (Actual Data)
$\hat{y}_i$ คือ ค่าพยากรณ์ (Predictive value)

5. ค่าสัมประสิทธิ์สหสัมพันธ์ (Correlation Coefficient) นิยมใช้สัญลักษณ์ r แทนสัมประสิทธิ์สหสัมพันธ์ของกลุ่มตัวอย่าง และ ρ แทนสัมประสิทธิ์สหสัมพันธ์ของประชากร โดยมีค่าระหว่าง -1 ถึง 1 พร้อมเครื่องหมาย + หรือ - เพื่อบ่งบอกทิศทางของความสัมพันธ์ ส่วนความหมายของขนาดความสัมพันธ์สามารถตีความหมายได้ ดังนี้ (Kreuger & Neuman, 2005)

ค่าสัมประสิทธิ์สหสัมพันธ์อยู่ระหว่าง	0.80 - 1.00	หมายาความว่า มีระดับความสัมพันธ์ สูงมาก
ค่าสัมประสิทธิ์สหสัมพันธ์อยู่ระหว่าง	0.60 - 0.79	หมายาความว่า มีระดับความสัมพันธ์ สูง
ค่าสัมประสิทธิ์สหสัมพันธ์อยู่ระหว่าง	0.40 - 0.59	หมายาความว่า มีระดับความสัมพันธ์ ปานกลาง
ค่าสัมประสิทธิ์สหสัมพันธ์อยู่ระหว่าง	0.20 - 0.39	หมายาความว่า มีระดับความสัมพันธ์ ต่ำ
ค่าสัมประสิทธิ์สหสัมพันธ์อยู่ระหว่าง	0.00 - 0.19	หมายาความว่า มีระดับความสัมพันธ์ ต่ำมาก

4.6 การนำไปใช้งาน (Deployment)

เมื่อดำเนินการวิเคราะห์ข้อมูลครบทั้ง 5 ขั้นตอนของกระบวนการ CRISP - DM แล้ว จึงจะสามารถระบุเทคนิคที่มีความเหมาะสมที่สุดในการพยากรณ์เบี้ยประกันภัยรับรวมของบริษัทประกันชีวิตในประเทศไทย ซึ่งผลการพยากรณ์นี้มีความแม่นยำเพียงพอที่จะใช้สนับสนุนการวางแผนเชิงกลยุทธ์ของบริษัทประกันชีวิต ช่วยให้สามารถคาดการณ์แนวโน้มการเปลี่ยนแปลงในอนาคตได้อย่างมีประสิทธิภาพ

นอกจากนี้ ผลการพยากรณ์ยังสามารถนำไปประยุกต์ใช้ในการบริหารความเสี่ยงทางการเงิน ปรับปรุงกลยุทธ์ด้านการตลาด และออกแบบผลิตภัณฑ์ประกันชีวิตให้สอดคล้องกับความต้องการของตลาดในแต่ละช่วงเวลา สำหรับหน่วยงานกำกับดูแล อาทิ สำนักงานคณะกรรมการกำกับและส่งเสริมการประกอบธุรกิจประกันภัย (คปภ.) ผลการศึกษานี้สามารถใช้เป็นข้อมูลเชิงประจักษ์เพื่อกำหนดนโยบายควบคุมและส่งเสริมอุตสาหกรรมประกันชีวิตอย่างเป็นระบบ ทั้งในด้านการกำหนดอัตราเบี้ยประกัน การวางหลักเกณฑ์ด้านเงินกองทุน และการประเมินเสถียรภาพทางการเงินของบริษัทประกัน

ในภาพรวม ผลการศึกษานี้มีศักยภาพในการเพิ่มประสิทธิภาพการบริหารจัดการของบริษัทประกันชีวิตไทยให้สามารถแข่งขันได้อย่างยั่งยืนในระยะยาว และรองรับความผันผวนทางเศรษฐกิจและสังคมที่ส่งผลต่ออุตสาหกรรมประกันภัยได้อย่างเหมาะสมและทันท่วงที

Citation: Sukprasert, A., Samrual, O., & Wisetrat, U. (2025). Development and Comparison of Forecasting Models for Gross Written Premiums of Life Insurance Companies in Thailand Using Data Mining Technique. *Journal of Computer and Creative Technology, 3*(1), 143-155. https://doi.org/10.14456/jcct.2025.12.

5. ผลการวิจัย (Results)

5.1 ผลการสร้างตัวแบบสำหรับพยากรณ์เบี้ยประกันภัยรับรวมของบริษัทประกันชีวิตในประเทศไทย

ก่อนการสร้างแบบจำลองพยากรณ์เบี้ยประกันภัยรับรวม จะต้องทำการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรต้นด้านเศรษฐกิจและการเงิน กับตัวแปรตาม ได้แก่ เบี้ยประกันภัยรับรวม (Gross Written Premium: GWP) เพื่อศึกษาความเชื่อมโยงเชิงสถิติในเบื้องต้น และใช้ประกอบการคัดเลือกตัวแปรที่เหมาะสมสำหรับการพยากรณ์ ดัง Table 2.

Table 2. Illustrates the relationship between variables.

Variable	CPI	BCI	Interest rates	M2	SE Price Index	GBY	ATR	ROA	NIP	DTER	GWP
CPI	1	-0.027	-0.777	0.873	0.122	-0.234	-0.682	-0.380	-0.333	-0.490	0.511
BCI	-0.027	1	0.298	-0.287	0.273	0.420	0.133	-0.172	0.200	0.265	-0.048
Interest rates	-0.777	0.298	1	-0.913	-0.031	0.625	0.544	0.119	0.222	0.697	-0.631
M2	0.873	-0.287	-0.913	1	0.069	-0.612	-0.737	-0.312	-0.311	-0.583	0.654
SE Price Index	0.122	0.273	-0.031	0.069	1	-0.045	0.151	-0.493	0.150	-0.321	0.224
GBY	-0.234	0.420	0.625	-0.612	-0.045	1	0.265	0.003	0.074	0.515	-0.535
ATR	-0.682	0.133	0.544	-0.737	0.151	0.265	1	0.255	0.493	-0.058	-0.260
ROA	-0.380	-0.172	0.119	-0.312	-0.493	0.003	0.255	1	0.037	0.091	-0.262
NIP	-0.333	0.200	0.222	-0.311	0.150	0.074	0.493	0.037	1	-0.050	0.481
DTER	-0.490	0.265	0.697	-0.583	-0.321	0.515	-0.058	0.091	-0.050	1	-0.589
GWP	0.511	-0.048	-0.631	0.654	0.224	-0.535	-0.260	-0.262	0.481	-0.589	1

จาก Table 2. แสดงค่าสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรต้นกับตัวแปรตาม โดยพบว่า อัตราดอกเบี้ยเงินฝากประจำ (Interest Rates) และ ปริมาณเงินความหมายกว้าง (M2) มีความสัมพันธ์กับ GWP ในระดับสูง โดยมีค่าสัมประสิทธิ์สหสัมพันธ์เท่ากับ -0.631 และ 0.654 ตามลำดับ ซึ่งสะท้อนว่าอัตราดอกเบี้ยเงินฝากประจำมีความสัมพันธ์ในเชิงลบกับเบี้ยประกันภัยรับรวม ขณะที่ M2 มีความสัมพันธ์ในเชิงบวก สำหรับตัวแปรอื่น ๆ ได้แก่ ดัชนีราคาผู้บริโภค (CPI), อัตราผลตอบแทนพันธบัตรรัฐบาล (GBY), อัตราเบี้ยประกันรับสุทธิ (NIP) และ อัตราส่วนหนี้สินต่อส่วนของผู้ถือหุ้น (DTER) มีความสัมพันธ์กับ GWP ในระดับปานกลาง โดยมีค่าสัมประสิทธิ์สหสัมพันธ์เท่ากับ 0.511, -0.535, 0.481 และ -0.589 ตามลำดับ ขณะที่ ดัชนีราคาตลาดหลักทรัพย์ (SE Price Index), อัตราการหมุนของสินทรัพย์ (ATR) และ อัตราผลตอบแทนต่อสินทรัพย์ (ROA) มีความสัมพันธ์กับ GWP ในระดับต่ำ โดยมีค่าสัมประสิทธิ์สหสัมพันธ์เท่ากับ 0.224, -0.260 และ -0.262 ตามลำดับ นอกจากนี้ ดัชนีความเชื่อมั่นทางธุรกิจ (BCI) มีค่าสัมประสิทธิ์สหสัมพันธ์เพียง -0.048 แสดงให้เห็นถึงความสัมพันธ์ในระดับต่ำมาก หรือแทบไม่มีความสัมพันธ์กับ GWP อย่างชัดเจน

5.2 ผลการเปรียบเทียบประสิทธิภาพของตัวแบบสำหรับพยากรณ์เบี้ยประกันภัยรับรวมของบริษัทประกันชีวิตในประเทศไทย

จากผลการวิเคราะห์ดังกล่าว สามารถสรุปได้ว่าตัวแปรทางเศรษฐกิจมีระดับความสัมพันธ์กับเบี้ยประกันภัยรับรวมที่แตกต่างกัน ซึ่งข้อมูลเชิงสถิตินี้สามารถนำมาใช้ประกอบการคัดเลือกตัวแปรสำหรับการสร้างแบบจำลองพยากรณ์ และสนับสนุนการวางแผนเชิงกลยุทธ์ในอุตสาหกรรมประกันชีวิตได้อย่างมีประสิทธิภาพ ต่อจากนั้นได้นำข้อมูลที่ผ่านการวิเคราะห์มาสร้างตัวแบบสำหรับการพยากรณ์เบี้ยประกันภัยรับรวมของบริษัทประกันชีวิตในประเทศไทย พร้อมทั้งประเมินประสิทธิภาพของแบบจำลองโดยเปรียบเทียบผลลัพธ์จากทั้ง 4 เทคนิคที่เลือกใช้ พบว่า เทคนิคโครงข่ายประสาทเทียม ให้ค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย ต่ำที่สุด เท่ากับ 1,083.91, ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยเท่ากับ 1,174,870.83 ค่าสัมประสิทธิ์การตัดสินใจ เท่ากับ 97.60% เปอร์เซ็นค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ย 2.53% รองลงมาคือ เทคนิคต้นไม้ตัดสินใจ ให้ค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย เท่ากับ 2,287.84 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย เท่ากับ 5,234,195.04 ค่าสัมประสิทธิ์การตัดสินใจ เท่ากับ 89.10% เปอร์เซ็นค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ย 5.46% เทคนิคต้นไม้ป่าสุ่ม ให้ค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย เท่ากับ

Citation:
Sukprasert, A., Samrual, O., & Wisetrat, U. (2025). Development and Comparison of Forecasting Models for Gross Written Premiums of Life Insurance Companies in Thailand Using Data Mining Technique. *Journal of Computer and Creative Technology, 3*(1), 143-155. https://doi.org/10.14456/jcct.2025.12.

2,334.63 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย เท่ากับ 5,450,515.70 ค่าสัมประสิทธิ์การตัดสินใจ เท่ากับ 89.60% เปอร์เซ็นค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ย 5.57% และเทคนิคที่ให้ค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ยสูงที่สุด คือ เทคนิคซัพพอร์ตเวกเตอร์แมชชีน ให้ค่ารากที่สองของค่าความคลาดเคลื่อนกำลังสองเฉลี่ย เท่ากับ 6,867.84 ค่าความคลาดเคลื่อนกำลังสองเฉลี่ย เท่ากับ 47,167,186.24 ค่าสัมประสิทธิ์การตัดสินใจ เท่ากับ 61.50% เปอร์เซ็นค่าคลาดเคลื่อนสัมบูรณ์เฉลี่ย 18.49% ตามลำดับ ดังนั้นเทคนิคที่มีความเหมาะสมที่สุดสำหรับการพยากรณ์เบี้ยประกันภัยรับรวม คือ เทคนิคโครงข่ายประสาทเทียม ดัง Table 3.

Table 3. Comparison of the Performance of Predictive Models.

Regression Techniques	Regression Performance			
	RMSE	MSE	R^2	MAPE
Decision Tree	2,287.84	5,234,195.04	89.10%	5.46%
Random Forest	2,334.63	5,450,515.70	89.60%	5.57%
Support Vector Machine	6,867.84	47,167,186.24	61.50%	18.49%
Neural Network*	1,083.91	1,174,870.83	97.60%	2.53%

* An appropriate technique for developing a forecasting model for Gross Written Premiums of Life Insurance Companies in Thailand

จากนั้นได้นำค่าจริงและค่าพยากรณ์ของเบี้ยประกันภัยรับรวมของแต่ละเทคนิคมาสร้างเป็นแผนภูมิแท่งเพื่อเปรียบเทียบประสิทธิภาพของตัวแบบที่ได้จากเทคนิคโครงข่ายประสาทเทียม ดังแสดงใน Figure 1.

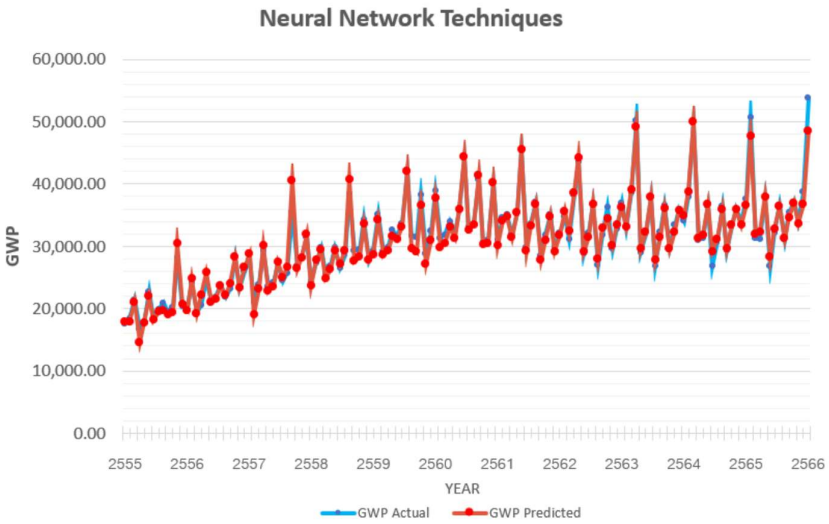

Figure 1. Line graph showing the trend of Gross Written Premiums (GWP), actual values, and forecasted values using the Neural Network technique.

6. สรุปและอภิปรายผลการวิจัย (Conclusion and Discussion)

งานวิจัยฉบับนี้มีวัตถุประสงค์เพื่อเปรียบเทียบประสิทธิภาพของแบบจำลองการพยากรณ์เบี้ยประกันภัยรับรวมของบริษัทประกันชีวิตในประเทศไทย โดยประยุกต์ใช้เทคนิคการทำเหมืองข้อมูล (Data Mining Techniques) ผ่านกระบวนการพยากรณ์เชิงปริมาณ (Regression Models) จำนวน 4 เทคนิค ได้แก่ เทคนิคต้นไม้ตัดสินใจ (Decision Tree) เทคนิคต้นไม้ป่าสุ่ม (Random Forest) เทคนิคซัพพอร์ตเวกเตอร์แมชชีน (Support Vector Machine: SVM) และเทคนิคโครงข่ายประสาทเทียม (Neural Network) การประเมินประสิทธิภาพของแบบจำลองดำเนินการโดยใช้เกณฑ์หลัก 5

Citation: Sukprasert, A., Samrual, O., & Wisetrat, U. (2025). Development and Comparison of Forecasting Models for Gross Written Premiums of Life Insurance Companies in Thailand Using Data Mining Technique. *Journal of Computer and Creative Technology, 3*(1), 143-155. https://doi.org/10.14456/jcct.2025.12.

ตัวชี้วัด ได้แก่ ค่ารากที่สองของค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Root Mean Square Error: RMSE) ค่าคลาดเคลื่อนกำลังสองเฉลี่ย (Mean Squared Error: MSE) สัมประสิทธิ์การตัดสินใจ (R-squared: R^2) ค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ยในรูปเปอร์เซ็นต์ (Mean Absolute Percentage Error: MAPE) และค่าสัมประสิทธิ์สหสัมพันธ์ (Correlation Coefficient) เพื่อให้ได้ผลการเปรียบเทียบที่ครอบคลุมทั้งด้านระดับความแม่นยำ ความสัมพันธ์เชิงเส้น และความคลาดเคลื่อนเชิงสัมพัทธ์ ผลการวิเคราะห์พบว่า แบบจำลองที่ใช้เทคนิคโครงข่ายประสาทเทียม (Neural Network) ให้ผลการพยากรณ์ที่แม่นยำที่สุด โดยใช้โครงข่ายที่มีชั้นซ่อน (Hidden Layers) จำนวน 6 ชั้น ได้ค่า MSE เท่ากับ 1,174,870.83 ค่า R^2 เท่ากับ 97.60% ค่า RMSE ต่ำที่สุดที่ 1,083.91 และค่า MAPE เท่ากับ 2.53% ซึ่งแสดงให้เห็นถึงความสามารถของโครงข่ายประสาทเทียมในการเรียนรู้ความสัมพันธ์ที่ไม่เป็นเชิงเส้นและสามารถจัดการกับข้อมูลที่มีความซับซ้อนได้อย่างมีประสิทธิภาพ อย่างไรก็ตาม แบบจำลองโครงข่ายประสาทเทียมยังมีข้อจำกัดที่ควรพิจารณา ได้แก่ ความซับซ้อนในการตีความผลลัพธ์ เนื่องจากมีโครงสร้างหลายชั้นและมีพารามิเตอร์จำนวนมากที่ต้องปรับแต่ง (Hyperparameter Tuning) รวมถึงความต้องการใช้ทรัพยากรในการประมวลผลที่สูง และความเสี่ยงต่อการเกิด Overfitting หากไม่มีการเตรียมข้อมูลและควบคุมพารามิเตอร์อย่างเหมาะสม เมื่อเปรียบเทียบกับผลการศึกษาของ Phaphan & Puttamat (2023) ซึ่งใช้ข้อมูลรายเดือนระหว่างปี พ.ศ. 2560 – 2565 (72 เดือน) และพบว่าแบบจำลอง Decision Tree Regression ให้ผลการพยากรณ์ดีที่สุด โดยมีค่า RMSE เท่ากับ 1,654.00 และค่า MAPE เท่ากับ 2.93% พบว่าผลการวิจัยฉบับนี้ให้ผลลัพธ์ที่แม่นยำกว่า อันเนื่องมาจากการใช้ชุดข้อมูลที่ครอบคลุมระยะเวลาที่ยาวกว่า (ปี พ.ศ. 2555 – 2566 รวม 144 เดือน) ประกอบกับการคัดเลือกตัวแปรต้นที่ครอบคลุมมากขึ้น และกระบวนการปรับแต่งพารามิเตอร์ของแต่ละแบบจำลองที่มีความละเอียดรอบคอบมากขึ้น

ผลการศึกษานี้สะท้อนให้เห็นถึงศักยภาพของการนำปัญญาประดิษฐ์ (Artificial Intelligence: AI) โดยเฉพาะโครงข่ายประสาทเทียม มาใช้เป็นเครื่องมือในการวิเคราะห์และพยากรณ์ข้อมูลในอุตสาหกรรมประกันภัย ซึ่งสามารถประยุกต์ใช้เพื่อสนับสนุนการบริหารจัดการเบี้ยประกันภัยอย่างมีประสิทธิภาพ และช่วยลดความไม่แน่นอนทางการเงินของบริษัทประกันชีวิต นอกจากนี้ แบบจำลองพยากรณ์ที่มีความแม่นยำสามารถนำไปใช้ประกอบการกำหนดนโยบายด้านการบริหารความเสี่ยงทางการเงิน ภายใต้กรอบมาตรฐานการดำรงเงินกองทุนตามระดับความเสี่ยง (Risk-Based Capital: RBC) ของสำนักงานคณะกรรมการกำกับและส่งเสริมการประกอบธุรกิจประกันภัย (คปภ.) ได้อย่างเป็นรูปธรรม โดยเฉพาะการนำแบบจำลองดังกล่าวไปประยุกต์ใช้ในกระบวนการ Stress Testing, Scenario Analysis และการวางแผนเงินกองทุนสำรอง เพื่อให้สอดคล้องกับภาวะเศรษฐกิจที่เปลี่ยนแปลงอย่างรวดเร็ว

7. ข้อเสนอแนะงานวิจัย (Recommendation)

1. การศึกษานี้สามารถต่อยอดโดยการเพิ่มตัวแปรต้นที่เกี่ยวข้องเพิ่มเติม เพื่อให้การวิเคราะห์ปัจจัยที่ส่งผลต่อเบี้ยประกันภัยรับรวมของบริษัทประกันชีวิตในประเทศไทยมีความครอบคลุมมากยิ่งขึ้น อีกทั้งการวิเคราะห์ความสัมพันธ์ของปัจจัยที่มีนัยสำคัญ ยังสามารถนำไปสู่การพัฒนาแบบจำลองพยากรณ์ที่มีความแม่นยำและประสิทธิภาพที่สูงขึ้นได้

2. ด้านเทคนิคการพยากรณ์ ควรพิจารณานำเทคนิคขั้นสูงด้าน Deep Learning มาประยุกต์ใช้ โดยเฉพาะโมเดลที่เหมาะสมกับข้อมูลอนุกรมเวลาและความสัมพันธ์ที่ไม่เป็นเชิงเส้น เช่น RNN และ LSTM มาใช้ในการวิเคราะห์ รวมถึงการพัฒนาแบบจำลองแบบผสมผสานระหว่าง ARIMA กับ LSTM หรือ Neural Network เพื่อเพิ่มความแม่นยำในการพยากรณ์ในบริบทของข้อมูลที่มีความซับซ้อน

3 ควรมีการศึกษาผลกระทบของ เหตุการณ์เฉพาะเจาะจง เช่น สถานการณ์ทางเศรษฐกิจ โครงสร้างประชากรที่เปลี่ยนแปลง (เช่น สังคมผู้สูงอายุ) หรือพฤติกรรมผู้บริโภคในแต่ละยุค ซึ่งอาจมีผลต่อความแม่นยำของโมเดล และช่วยให้สามารถปรับปรุงแบบจำลองให้มีความยืดหยุ่นและตอบสนองต่อบริบทที่เปลี่ยนแปลงได้ดียิ่งขึ้น

8. กิตติกรรมประกาศ

งานวิจัยฉบับนี้สำเร็จลุล่วงไปได้ด้วยดี ด้วยความอนุเคราะห์และการสนับสนุนจากหน่วยงานต่าง ๆ ดังต่อไปนี้ สมาคมประกันชีวิตไทย กระทรวงพาณิชย์ ธนาคารแห่งประเทศไทย Investing และ SET Index ที่ได้ให้ข้อมูลอันเป็นประโยชน์และสอดคล้องกับวัตถุประสงค์ของการวิจัยอย่างครบถ้วน ขอแสดงความขอบคุณเป็นอย่างสูงสำหรับข้อมูลอันมีค่าที่นำมาใช้ในการวิเคราะห์ในครั้งนี้ นอกจากนี้ขอขอบคุณ คณะการบัญชีและการจัดการ มหาวิทยาลัยมหาสารคาม ที่ให้

Citation:
Sukprasert, A., Samrual, O., & Wisetrat, U. (2025). Development and Comparison of Forecasting Models for Gross Written Premiums of Life Insurance Companies in Thailand Using Data Mining Technique. *Journal of Computer and Creative Technology, 3*(1), 143-155. https://doi.org/10.14456/jcct.2025.12.

การสนับสนุนในการดำเนินงานวิจัย รวมถึงผู้ที่ได้ให้คำปรึกษาและข้อเสนอแนะตลอดระยะเวลาการทำวิจัย ตลอดจนคณาจารย์ทุกท่านที่มีส่วนเกี่ยวข้อง

9. เอกสารอ้างอิง (References)

Bangkok Biz News. (2024). *Life Insurance Business in the First Half of 2024: Total Premiums Received Reach 310 Billion Baht, Up 3.8%*. https://www.bangkokbiznews.com/finance/investment/1138229. (In Thai)

Bank of Thailand. (2024a). *Broad Money Supply*. https://app.bot.or.th/BTWS_STAT/statistics/BOTWEBSTAT.aspx?reportID=7&language=TH. (In Thai)

Bank of Thailand. (2024b). *Business Confidence Index*. https://app.bot.or.th/BTWS_STAT/statistics/BOTWEBSTAT.aspx?reportID=405&language=TH. (In Thai)

Bank of Thailand. (2024c). *Fixed Deposit Interest Rates*. Retrieved August 3, 2024, from https://www.bot.or.th/th/statistics/interest-rate.html. (In Thai)

Chowdhury, R. S., Islam, M. A., Yusuf, D. H. B. M., Amin, M. B., Hassan, M. S., Barua, S., & Abdullah, M. (2024). Implications of Artificial Intelligence (AI) and Machine Learning-based Fintech for the Financial Assets Related Traditional Investment Theories. *Journal of Infrastructure, Policy and Development*, *8*(12), 7415. https://doi.org/10.24294/jipd.v8i12.7415.

Hangkaew, C., Sinsomboonthong, J., & Supapakorn, T. (2020). Prediction of choosing types of insurance for the customers. *Huachiew Chalermprakiet Science and Technology Journal, 6*(2), 81–95. (In Thai)

Huadsri, S., & Phaphan, W. (2024). The Development of Forecasting Models for Life Insurance Data by Employing Time-series Analysis and Machine Learning Technique. *WSEAS Transactions on Mathematics*, *23*, 196–205. https://doi.org/10.37394/23206.2024.23.23.

Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and Practice* (3rd ed.). OTexts.

Investing. (2024). *Thailand 10-Years Government Bond Yield*. https://th.investing.com/rates-bonds/thailand-10-year-bond-yield-historical-data. (In Thai)

Kittibawornkul, W., Nonsiri, S., & Kamin, P. (2022). The Comparison of Machine Learning Algorithms for Predictive Maintenance of Aircraft Engine. *APHEIT Journal (Science and Technology), 11*(1), 1–14. (In Thai)

Kreuger, L. W., & Neuman, W. L. (2005). *Social Work Research Methods: Qualitative and Quantitative Appro aches With Research Navigator*. Pearson.

Phaphan, W., & Puttamat, W. (2023). The Comparison of Forecasting Models for Total Premiums of Life Insurance Companies in Thailand. *Huachiew Chalermprakiet Science and Technology Journal, 9*(2), 64-74. (In Thai)

Sangsri, P., & Chaiwuttisak, P. (2021). Hybrid of Neural Network and Markov Chain Method for Predicting PM2.5 Concentrations. *Thai Science and Technology Journal*, *29*(5), 752–765. (In Thai)

Sanguansat, P. (2019). *Artificial Intelligence with Machine Learning*. IDC Premier. (In Thai)

SET Index. (2024). Stock Exchange of Thailand Price Index. https://www.set.or.th/static/mktstat/Table_Index.xls?001. (In Thai)

Sukprasert, A. (2023). *Data Mining with RapidMiner Studio* (5th ed.). Mahasarakham University. (In Thai)

Thai Life Assurance Association. (2024a). *Asset Turnover Rate*. https://www.tlaa.org/download.php?cid=26&cname=&show=0. (In Thai)

Citation: Sukprasert, A., Samrual, O., & Wisetrat, U. (2025). Development and Comparison of Forecasting Models for Gross Written Premiums of Life Insurance Companies in Thailand Using Data Mining Technique. *Journal of Computer and Creative Technology, 3*(1), 143-155. https://doi.org/10.14456/jcct.2025.12.

Thai Life Assurance Association. (2024b). *Debt to Equity Ratio*. https://www.tlaa.org/download.php?cid=26&cname=&show=0. (In Thai)

Thai Life Assurance Association. (2024c). *Gross Written Premium*. https://www.tlaa.org/page_statistics_new.php?cid=79&cname. (In Thai)

Thai Life Assurance Association. (2024d). *Net Insurance Premium Received*. https://www.tlaa.org/download.php?cid=26&cname=&show=0. (In Thai)

Thai Life Assurance Association. (2024e). *Overview of the Life Insurance Business in the First Half of 2024*. https://www.tlaa.org/page_bx.php?cid=23&cname=&cno=1805. (In Thai)

Thai Life Assurance Association. (2024f). *Return on Assets (ROA)*. https://www.tlaa.org/download.php?cid=26&cname=&show=0. (In Thai)

Thai Reinsurance Public Company Limited. (2020). *Annual Report 2020*. https://thre.listedcompany.com/misc/ar/20210322-thre-ar-2020-en.pdf.

Trade Policy and Strategy Office. (2024). *General Consumer Price Index Analysis Report in August 2024*. https://index.tpso.go.th/cpi/index-analysis-report/1. (In Thai)

Varadarajan, V., & Kakumanu, V. K. (2024). Evaluation of Risk Level Assessment Strategies in Life Insurance: A Review of the Literature. *Journal of Autonomous Intelligence, 7*(5), 1147. https://doi.org/10.32629/jai.v7i5.1147.

Wanon, S., & Muangsan, R. (2021). Improving Prediction Models of Student Business Career Using Sampling Techniques for Learning in Multi-class Imbalance Data Set. *Journal of Chaiyaphum Review*, *4*(1), 39–49. (In Thai)

Zhou, Z. H. (2021). *Machine Learning*. Springer Singapore. https://doi.org/10.1007/978-981-15-1967-3.

Citation:
Sukprasert, A., Samrual, O., & Wisetrat, U. (2025). Development and Comparison of Forecasting Models for Gross Written Premiums of Life Insurance Companies in Thailand Using Data Mining Technique. *Journal of Computer and Creative Technology, 3*(1), 143-155. https://doi.org/10.14456/jcct.2025.12.