Iran. J. Insur. Res., **(*): *-*, ****** ***

Iranian Journal of Insurance Research
(IJIR)
Homepage: https://ijir.irc.ac.ir/?lang=en

Meta-models: A Novel Approach for Valuation and Risk Management of Large Variable Annuity Portfolios

Morteza Hosseingholi pour[1], Saghar Heidari[2,*]

[1] Department of Actuarial Science, Faculty of Mathematical Sciences, Shahid Beheshti University, Tehran, Iran.

[2] Department of Actuarial Science, Faculty of Mathematical Sciences, Shahid Beheshti University, Tehran, Iran.

ARTICLE INFO

***KEYWORDS*:**
Intelligent Models
Insurance Large Portfolio
Machine Learning
Meta Models
Risk Management
Variable Annuities

ABSTRACT

BACKGROUND AND OBJECTIVES: Variable annuities have become a popular tool in retirement planning due to their attractive guarantees. These insurance products offer a reliable, steady income stream, not only aiding in financial stability during retirement but also enhancing mental well-being and overall quality of life by enhancing financial security and reducing anxiety. Variable annuities, offered by insurance companies, are tax-deferred retirement products designed to help individuals navigate the challenges of preserving the long-term value of their assets amid economic volatility, inflation, and uncertainty. However, the risk management of these products, which requires precise, up-to-date, and frequent valuation, poses a significant computational challenge for insurance companies, especially in large portfolios. Today, common and traditional methods such as Monte Carlo simulation, are often inefficient for risk management of these products due to their time-consuming nature. In recent years, with the emergence of modern technologies such as meta-modeling methods as a solution to these challenges, they have been proposed, but unfortunately, limited research has focused on increasing the adaptability and intelligence of these secondary models when facing changing data. In this regard, to address this challenge, the main objective of this study is to improve the adaptability and intelligence of meta-models by developing an intelligent meta-model based on artificial intelligence that can value large variable annuity portfolios more quickly and accurately. As a result, this would enable better responsiveness to ongoing market developments and portfolio structure changes.

METHODS: In this study, to address data changes, we need to improve the adaptability and intelligence of meta-models, which will consequently improve their responsiveness to ongoing developments in the market and portfolio structure. This improvement is achieved through the development of methods for intelligently determining the number of representative contracts, optimally selecting a subset of contracts, and choosing the most accurate underlying valuation models. For this purpose, a suite of models and methods will be evaluated to assess their effectiveness and suitability in this context. Therefore, the methodology of this research involves the design and implementation of an intelligent meta-model, whose performance is evaluated using five well-known machine learning models from among machine learning models and according to the evaluation criteria commonly used in these methods. The main emphasis of the proposed approach in this study is on improving model accuracy through innovative approaches in data preprocessing and effective sampling, enabling the models so that the chosen models can generalize complex patterns to new data rather than merely memorizing them. These evaluations are conducted using standard metrics, and ultimately, the integration of these elements into an intelligent system aims to provide a final validation of the potential of artificial intelligence in this field.

FINDINGS: The results of this study confirm the effectiveness of artificial intelligence in enhancing the meta-models for valuing variable annuities. The final meta-model obtained from the proposed approach demonstrated highly accurate performance according to standard metrics, indicating the model's ability to cope with the valuation complexities of large insurance portfolios. The key factors of this success include the high importance of data preprocessing, targeted sampling, and the selection of efficient machine learning models as representative models. The results indicate that although increasing the size of the dataset for the meta-model may not necessarily improve prediction accuracy and efficiency, the use of optimization algorithms for selecting and tuning the model's parameters can significantly improve performance. Therefore, the findings emphasize the important role of parameter optimization over merely increasing the training data size in enhancing the final model's accuracy and efficiency in prediction.

CONCLUSION: This research focuses on intelligent enhancement of meta-models through artificial intelligence to transform variable annuity portfolio management. The main finding of this study is a

significant improvement in the quality and accuracy of the obtained results in valuation and risk management of the large portfolios of variable annuities in insurance industry. For this, by integrating precise and reliable methodologies into the proposed meta-model structure, enabling insurance companies to manage their portfolios of variable annuities with greater confidence and provide better service to policyholders. These results not only enhance the credibility and significant rolls of novel approach of artificial intelligence applications in the insurance industry but also offer practical insights for companies to develop stronger and more adaptable valuation and risk management models in the face of the financial complexities inherent in variable annuities.

Iran. J. Insur. Res., **(*): *-*, ****** ***

Iranian Journal of Insurance Research
(IJIR)
Homepage: https://ijir.irc.ac.ir/?lang=fa

متا مدل‌ها: روشی نوین برای ارزش‌گذاری و مدیریت ریسک پرتفوهای بزرگ مستمری‌های متغیر

مرتضی حسینقلی پور[1]، ساغر حیدری[2,*]

[1] *گروه بیمسنجی، دانشکده علوم ریاضی، دانشگاه شهید بهشتی، تهران، ایران.*

[2] *هیات علمی گروه بیمسنجی، دانشکده علوم ریاضی، دانشگاه شهید بهشتی، تهران، ایران.*

چکیده

پیشینه و اهداف: مستمری‌های متغیر به دلیل ضمانت‌های جذابشان، به ابزاری محبوب در برنامه‌ریزی طرح های بازنشستگی تبدیل شده‌اند. با این حال، مدیریت ریسک مرتبط با این محصولات، چالش محاسباتی قابل توجهی را برای شرکت‌های بیمه، به‌ویژه در پرتفوی‌های بزرگ، ایجاد می‌کند. در روش سنتی شبیه‌سازی مونت کارلو، به دلیل زمان‌بر بودن، ناکارآمد هستند. روش‌های مدل‌سازی متا به عنوان راه‌حلی برای این چالش‌ها مطرح شده‌اند، اما متاسفانه تحقیقات اندکی بر افزایش سازگاری و هوشمندی این مدل‌های ثانویه انجام شده است. برای مقابله با این چالش هدف اصلی این پژوهش بهبود قابلیت انطباق و هوش مدل‌های متا از طریق توسعه یک متا مدل هوشمند مبتنی بر هوش مصنوعی است تا بتواند ارزش‌گذاری پورتفوهای مستمری متغیر را با سرعت و دقت بیشتری انجام داده و در نتیجه قادر به پاسخگویی بهتر به تحولات مستمر بازار و ساختار پرتفوی است.

روش‌شناسی: افزایش دقت و کارایی از طریق توسعه روش‌هایی برای تعیین هوشمندانه تعداد قراردادهای نماینده، انتخاب بهینه زیرمجموعه‌ای از قراردادها و انتخاب دقیق‌ترین مدل‌های ارزش‌گذاری زیربنایی محقق می‌شود. برای این منظور مجموعه‌ای از مدل‌ها و روش‌ها را برای ارزیابی اثربخشی و تناسب آن‌ها در این زمینه ارزیابی خواهیم کرد. بنابراین روش‌شناسی این پژوهش شامل طراحی و اجرای یک متا مدل هوشمند است که عملکرد آن بر اساس پنج مدل معروف از میان مدل های یادگیری ماشین انجام می شود. تأکید اصلی رویکرد

کلمات کلیدی:
پرتفوهای بزرگ بیمه
متا مدل‌ها
مستمری‌های متغیر
مدل‌های هوشمند
مدیریت ریسک
یادگیری ماشین

پیشنهادی بر بهبود دقت مدل در پیش‌پردازش داده‌ها و نمونه‌گیری مؤثر می‌باشد تا مدل‌های انتخابی بتوانند الگوهای پیچیده را به جای حفظ کردن به‌درستی بر روی داده های جدید تعمیم دهند. این ارزیابی‌ها با معیارهای استاندارد انجام شده و در نهایت، یکپارچه‌سازی این عناصر در یک سیستم هوشمند برای اعتبارسنجی نهایی پتانسیل هوش مصنوعی در این حوزه انجام می‌شود.

یافته‌ها: یافته این پژوهش، کارایی هوش مصنوعی را در ارتقای متا مدل‌های ارزش‌گذاری مستمری‌های متغیر ثابت کرد. متا مدل نهایی ، عملکردی بسیار دقیق را با معیارهای استاندارد ثبت کرد که نشان‌دهنده توانایی مدل در مقابله با پیچیدگی‌های ارزش‌گذاری پورتفوهای بزرگ بیمه است. عوامل کلیدی این موفقیت شامل اهمیت بالای پیش‌پردازش داده‌ها، نمونه‌گیری هدفمند، و انتخاب مدل‌های یادگیری ماشین کارآمد به عنوان مدل نماینده بود. نتایج بدست آمده نشان می‌دهند که اگرچه افزایش اندازه مجموعه داده برای متا مدل ممکن است منجر به بهبود دقت و کارایی مدل پیش‌بینی نشود، اما استفاده از الگوریتم‌های بهینه‌سازی برای انتخاب و تنظیم پارامترهای مدل می‌تواند عملکرد مدل را به صورت چشمگیری بهبود بخشد.

نتیجه‌گیری: نتیجه نهایی این پژوهش بهبود چشمگیر کیفیت و دقت نتایج بدست آمده در ارزش‌گذاری و مدیریت ریسک در صنعت بیمه است. برای این منظور با ادغام رویکردهای دقیق و قابل اعتماد در ساختار متا مدل ها، شرکت‌های بیمه قادر خواهند بود تا پرتفوی‌های بزرگ حاصل از مستمری های متغیر خود را با اطمینان، دقت و سرعت بیشتری مدیریت کرده و در نتیجه بتوانند خدمات بهتری به بیمه‌گذاران ارائه دهند. این نتایج نه تنها اعتبار و لزوم استفاده از هوش مصنوعی را در صنعت بیمه افزایش می‌دهد، بلکه بینش‌های عملی را نیز در این راستا فراهم می‌کند تا شرکت‌ها بتوانند از آن برای توسعه مدل‌های ارزش‌گذاری و مدیریت ریسک قوی‌تر و سازگارتر استفاده کنند.

مقدمه

در دنیای امروز، رسیدگی به چالش‌های مالی و اقتصادی مرتبط با بیمه‌های مستمری[1] از اهمیت حیاتی برخوردار است، زیرا این محصولات می‌توانند حمایت‌های ارزشمندی را در دوران سالمندی و بازنشستگی فراهم آورند. مستمری‌ها یک جریان درآمدی پایدار و قابل اتکا را تضمین می‌کنند که نه تنها به ثبات مالی در دوران بازنشستگی کمک می‌کنند، بلکه با افزایش امنیت مالی و کاهش اضطراب، آرامش روانی و کیفیت کلی زندگی را بهبود می‌بخشند.

با توجه به شکاف موجود میان رشد درآمد بازنشستگان و نرخ تورم اقتصادی، فشارهای قابل توجهی بر این قشر جامعه وارد می‌شود. هزینه‌های زندگی به طور مداوم در حال افزایش است و این امر سبب شده است که طرح‌های مستمری سنتی، در تامین نیازهای بازنشستگان در طول دوران بازنشستگی ناتوان باشند. به همین دلیل، محصولات جدیدی نظیر مستمری‌های متغیر (VAs)[2]به بازار معرفی شده‌اند. VAs، که توسط شرکت‌های بیمه ارائه می‌شوند، محصولات بازنشستگی با امکان تعویق مالیاتی هستند که به افراد کمک می‌کنند تا در میان نوسان‌های بازار، تورم و عدم قطعیت‌های اقتصادی، ارزش بلندمدت دارایی‌های خود را حفظ نمایند. این محصولات به بیمه‌گذاران اجازه می‌دهند تا حق بیمه‌های خود را در صندوق‌های سرمایه‌گذاری مشخصی تخصیص دهند و در نتیجه در برابر تورم و چالش‌های اقتصادی محافظت شوند. با وجود پتانسیل این محصولات، این سرمایه‌گذاری‌ها نیز ممکن است دچار رکود شوند که می‌تواند احساس اضطراب و ناامنی را برانگیزد. برای مقابله با این وضعیت، شرکت‌های بیمه، تضمین‌هایی را در این مستمری‌ها تعبیه کرده‌اند که سرمایه‌گذاری بیمه‌گذاران را در برابر نوسان های بازار محافظت می‌کند. . بر اساس مزایای پرداختی، انواع متفاوتی از تضمین ها در محصولات بیمه ای وجود دارند، از جمله آن ها می توان به صندوق‌های مجزا[3] ، صندوق‌های سرمایه‌گذاری

[1] Annuity Insurances
[2] Variable Annuities
[3] Segregated Funds

تضمین‌شده، بیمه عمر متصل به واحد سرمایه‌گذاری[4] یا بیمه عمر مرتبط با سهام (Gan G., 2013) اشاره نمود. این لایه اضافی از حمایت، به افراد کمک می‌کند تا با اطمینان بیشتری در مستمری های متغیر سرمایه‌گذاری کنند و حس امنیت مالی و آرامش خاطر خود را بازیابند.

علیرغم اهمیت این محصولات بیمه‌ای، ارزش‌گذاری کارآمد و مدیریت ریسک پرتفوی‌های بزرگ مستمری متغیر، به دلیل پیچیدگی‌های محاسباتی، به طور کافی مورد توجه قرار نگرفته است. بیشتر مطالعات اخیر بر ارزیابی تضمین‌های تعبیه‌شده در قراردادهای انفرادی VA تمرکز دارند. اما، روش های ارزش‌گذاری و مدیریت ریسک طراحی‌شده برای قراردادهای انفرادی به راحتی قابل اعمال برای پرتفوی‌های بزرگ VA نیستند. از سوی دیگر، تضمین‌های موجود در VAs اغلب می‌توانند به عنوان اختیار مالی[5] برای بیمه‌گذار تلقی شوند. بنابراین، جای تعجب نیست که مفاهیم حاصل از معاملات اختیار معامله در قیمت‌گذاری و مدیریت ریسک این تضمین‌ها به کار گرفته شده‌اند. تضمین‌های تعبیه‌شده تعهدات مالی هستند که روش‌های سنتی بیمه‌گری به تنهایی قادر به مدیریت کامل آن‌ها نیستند. برای مدیریت ریسک‌های مالی مرتبط با این تضمین‌ها، شرکت‌های بیمه به پوشش ریسک پویا[6] به عنوان یک استراتژی کلیدی روی آورده‌اند. جنبه مهم پوشش ریسک پویا، کمی‌سازی ریسک است که شامل تعیین ارزش منصفانه بازار[7] این تضمین‌ها می‌باشد.

به دلیل پیچیدگی ذاتی این تضمین‌ها، ارزش‌گذاری آن‌ها نمی‌تواند به صورت صریح یا از طریق یک رابطه صریح انجام شود. برای غلبه بر این چالش، شرکت‌های بیمه معمولاً به شبیه‌سازی‌های مونت کارلو[8] متکی هستند تا ارزش منصفانه بازار این تضمین‌ها را تعیین کنند. در حالی که شبیه‌سازی‌های مونت کارلو انعطاف‌پذیری قابل توجهی ارائه می‌دهند و می‌توانند طیف وسیعی از تضمین‌ها را ارزش‌گذاری کنند، اما از نظر محاسباتی بسیار محدود بوده و نیازمند زمان و هزینه قابل توجهی هستند، به ویژه برای پرتفوی‌های بزرگ این مساله بیشتر مورد توجه است. علاوه بر این، یک برنامه پوشش ریسک پویا باید نه تنها دقیق، بلکه سریع نیز باشد تا بتواند به طور مؤثر پرتفوی VA را مدیریت کند. اگر محاسبات به سرعت انجام نشوند، پارامترهای پوشش ریسک مبتنی بر شرایط اولیه بازار ممکن است تا پایان محاسبات به طور قابل توجهی منحرف شوند، به ویژه زمانی که بازار حرکات سریعی را تجربه می‌کند. در نتیجه، اگرچه روش‌های شبیه‌سازی در مدیریت ریسک ضروری هستند، اما چالش‌های قابل توجهی از نظر پیچیدگی محاسباتی و هزینه ایجاد می‌کنند.

در پژوهش های اخیر روش های متا مدل به عنوان جایگزینی برای روش های شبیه سازی مطرح شدند. مفهوم بنیادی متامدل به معنای ایجاد یک مدل بر اساس یک زیرمجموعه از داده‌هاست که به‌طور مؤثر ویژگی‌های کل مجموعه داده را نمایان می‌کند. هدف این روش، کاهش قابل توجه حجم داده‌هایی است که باید مستقیما از طریق شبیه‌سازی مورد بررسی قرار گیرند. متامدل این امکان را می‌دهد تا فرآیند تحلیل و تصمیم‌گیری را تسهیل کرده و سرعت بخشیم. اما یکی از مشکلات اصلی متامدل‌های موجود، عدم هوشمندی آن‌هاست. به گونه‌ای که زمانی که داده‌ها تغییر می‌کنند یا پس از گذشت زمان، که الگوی داده‌ها دچار تغییر می‌شود، این متامدل‌ها نیاز به بروزرسانی دارند. از آنجا که این متامدل‌ها هوشمند نیستند، بروزرسانی لازم انجام نمی‌شود و به همین دلیل کارایی خود را از دست می‌دهند.

در سال‌های اخیر، تحقیقات متعددی در زمینه ارزش‌گذاری پرتفوی‌های مستمری متغیر انجام شده است که اکثر این مطالعات از متامدل‌ها به عنوان ابزار اصلی استفاده کرده‌اند. یکی از نقاط ضعف این تحقیقات، عدم قابلیت بروزرسانی خودکار متامدل‌های پیشنهادی در صورت بروز تغییرات قابل توجه در پرتفوی می باشد، زیرا روش‌های ارزش‌گذاری و مدل‌های پیش‌بینی‌کننده به همراه مجموعه داده‌های نماینده، در متامدل تغییر نمی‌کنند. همچنین، از آنجایی که در بلندمدت، رفتار و الگوی موجود در پرتفوی ممکن است تغییر کند، لذا ضرورت همگام‌سازی مدل‌ها و بروزرسانی متامدل با شرایط جدید احساس می‌شود.

در این راستا، هدف این پژوهش، بررسی هوشمندسازی متامدل‌ها بر پایه هوش‌مصنوعی است. برای این هدف با بهره‌گیری از الگوریتم‌های پیشرفته و روش‌های یادگیری ماشین، متامدل‌های توسعه یافته را طراحی کرده به طوری که این مدل ها قادر خواهند بود تا به‌طور خودکار خود را با تغییرات داده‌ها و شرایط بازار سازگار کنند و بدین ترتیب، کیفیت و دقت پیش‌بینی‌ها و مدیریت ریسک به میزان قابل توجهی افزایش خواهد یافت. این

[4] Unit-Linked Life Insurance
[5] Financial Options
[6] Dynamic Risk Hedging
[7] Fair Market Value
[8] Monte Carlo Simulations

نوآوری می‌تواند به شرکت‌های بیمه کمک کند تا با اطمینان بیشتری به مدیریت پرتفوهای مستمری متغیر خود بپردازند و در نتیجه، خدمات بهتری را به بیمه‌گذاران ارائه دهند.

مبانی نظری پژوهش

یک VA اساساً یک مستمری معوق با دو فاز کلیدی است: فاز انباشت[۹] و فاز پرداخت[۱۰]. در طول فاز انباشت، بیمه‌گذار حق بیمه‌های خود را در مجموعه‌ای از صندوق‌های سرمایه‌گذاری پیشنهادی توسط بیمه‌گر سرمایه‌گذاری می‌کند. این صندوق‌ها معمولاً در اوراق قرضه، صندوق‌های بازار پول، سهام و سایر محصولات مالی تخصیص می‌یابند. بیمه‌گران معمولاً تنوع گسترده‌ای از حساب‌های فرعی[۱۱] را برای تطابق با سطوح مختلف ریسک‌پذیری، از سرمایه‌گذاران محافظه‌کار تا ریسک‌پذیرتر، ارائه می‌دهند. ویژگی متغیر بودن مستمری به این معناست که ارزش حساب می‌تواند بر اساس عملکرد سرمایه‌گذاری‌های انتخاب شده نوسان کند، که پتانسیل رشد قابل توجهی دارد اما همزمان ریسک زیان را نیز به همراه دارد. پس از فاز انباشت، معمولاً در زمان بازنشستگی یا تاریخ مشخص شده دیگر، بیمه‌گر وارد فاز پرداخت می‌شود. در این فاز، بیمه‌گر مزایا را به صورت یکجا یا از طریق مجموعه‌ای از پرداخت‌ها که در قرارداد مشخص شده است، بازپرداخت می‌کند. مبلغ دریافتی می‌تواند بر اساس ارزش حساب در زمان پرداخت و ساختار صندوق‌های مستمری متفاوت باشد.

خرید یک مستمری متغیر مشابه سرمایه‌گذاری در صندوق‌های سرمایه‌گذاری مشترک است، زیرا بیمه‌گذاران می‌توانند وجوه خود را در یک یا چند گزینه سرمایه‌گذاری ارائه شده توسط صادرکننده تخصیص دهند. یک از مزیت های کلیدی مستمری های متغیر، تضمین‌های تعبیه‌شده است که سرمایه بیمه‌گذاران را در برابر نوسانات بازار محافظت می‌کند. این تضمین‌ها معمولاً به دو دسته اصلی تقسیم می‌شوند: حداقل مزایای فوت تضمین‌شده (GMDB)[۱۲] و حداقل مزایای زندگی تضمین‌شده (GMLB)[۱۳]. یک قرارداد GMDB تضمین می‌کند که مبلغ مشخصی در صورت فوت بیمه‌گذار در طول مدت قرارداد به ذی‌نفعان پرداخت شود، در حالی که GMLB حداقل پرداخت را از درآمد سرمایه اصلی یا سرمایه اولیه تضمین می‌کند، مشروط بر اینکه بیمه‌گذار تا زمان سررسید قرارداد زنده بماند.

بر اساس پژوهش (Bauer, 2008)، مزایای فوت در انواع مختلفی ارائه می‌شوند:

- مزایای فوت بازگشت حق بیمه[۱۴]: این ساده‌ترین شکل مزایای فوت است. در این حالت، پس از فوت بیمه‌گذار، ذی‌نفع مبلغی معادل بیشترین مقدار میان موجودی حساب در زمان فوت و کل حق بیمه‌های پرداخت شده توسط بیمه‌گذار را دریافت می‌کند.
- مزایای فوت با ترکیب سالانه[۱۵]: در این نوع قرارداد، مزایای فوت سالانه با یک نرخ بهره مشخص افزایش می‌یابد. شرکت بیمه تضمین می‌کند که بدون دریافت حق بیمه اضافی و مستقل از عملکرد بازار، مبلغ مزایای فوت را سالانه افزایش دهد. این ویژگی تضمین می‌کند که ذی‌نفعان مزایای فوت فزاینده دریافت خواهند کرد که اثرات تورم را در بلندمدت جبران می‌کند.
- مزایای فوت بر اساس بالاترین ارزش سالانه[۱۶]: در این تضمین، مبلغ مزایای فوت هر سال بر اساس بیشترین مقدار بین ارزش حساب یا مزایای فوت تضمین‌شده تنظیم می‌شود. اگر ارزش حساب از مبلغ مزایای فوت تضمین‌شده بیشتر شود، مبلغ مزایای فوت تا سطح ارزش حساب افزایش می‌یابد. این سازوکار تضمین می‌کند که ذی‌نفعان همیشه بیشترین مبلغ ممکن را دریافت خواهند کرد.

بر اساس مطالعات (Hardy, 2003)، (Shevchenko, 2016) و (Xu W. C., 2018)، چندین نوع GMLB وجود دارد که هر کدام سطوح مختلفی از حفاظت را ارائه می‌دهند:

[9] Accumulation
[10] Payout
[11] Sub-accounts
[12] Guarantee Minimum Death Benefit (GMDB)
[13] Guarantee Minimum Life Benefit (GMLB)
[14] Return of Premium Death Benefit (RPDB)
[15] Annual Compounding Death Benefit (ACDB)
[16] Annual Ratchet Death Benefit (ARDB)

- مزایای انباشت حداقل تضمین‌شده[17]: این محصول به بیمه‌گذاران اجازه می‌دهد تا سیاست خود را در تاریخ سررسید تمدید کنند. بیمه‌گر تضمین می‌کند که اگر مزایای تضمین‌شده در سررسید از ارزش صندوق بیشتر باشد، ارزش صندوق برای تطابق با مزایای تضمین‌شده افزایش یافته و قرارداد تمدید می‌شود. برعکس، اگر مزایای تضمین‌شده کمتر از ارزش صندوق باشد، قرارداد تمدید شده و مزایای تضمین‌شده با ارزش صندوق مجدداً تنظیم می‌گردد.
- مزایای درآمد حداقل تضمین‌شده[18]: این محصول تضمین می‌کند که بیمه‌گذاران بتوانند مبلغ انباشت‌شده در طول مدت قرارداد را با نرخ تضمین‌شده به یک مستمری تبدیل کنند. بیمه‌گذاران، صرف‌نظر از نوسانات بازار، حداقل پرداخت ماهانه تضمین‌شده‌ای را دریافت می‌کنند که درآمد دوران بازنشستگی آن‌ها را تضمین می‌نماید.
- مزایای سررسید حداقل تضمین‌شده[19]: این محصول پرداخت مبلغ مشخص در تاریخ سررسید قرارداد به بیمه‌گذار را تضمین می‌کند.
- مزایای برداشت حداقل تضمین‌شده[20]: این گزینه به بیمه‌گذاران اجازه می‌دهد تا نرخ تضمین‌شده‌ای را از حساب خود در طول مدت قرارداد برداشت کنند تا زمانی که سرمایه اولیه به طور کامل بازیابی شود.

افزودن این تضمین‌ها در VAs، افراد بسیاری را به سرمایه‌گذاری در این محصولات ترغیب کرده و منجر به افزایش چشمگیر فروش آن‌ها شده است. در نتیجه، شرکت‌های بیمه شاهد رشد قابل توجهی در پرتفوی‌های مستمری متغیر خود بوده‌اند. با این حال، این گسترش، چالش‌های جدیدی، به‌ویژه در زمینه‌های ارزش‌گذاری پرتفوی و مدیریت ریسک، به همراه داشته است.

مروری بر پیشینه پژوهش

محبوبیت VAs در ابتدا در ایالات متحده از دهه ۱۹۷۰ افزایش یافت، اگرچه ریشه توسعه آن‌ها حتی به دوران قدیمی‌تری بازمی‌گردد. در سال ۱۹۵۲، صندوق بازنشستگی دانشگاهی[21] محصولی با ویژگی‌های مستمری های متغیر معرفی کرد که به طور خاص برای محافظت از درآمد بازنشستگی معلمان در برابر تورم طراحی شده بود. این نوآوری که در ابتدا به عنوان یک قرارداد مستمری گروهی ارائه شد، اساس طیف گسترده‌تری از VAs موجود برای عموم را پایه‌ریزی کرد و پیشرفت‌های بیشتری را در صنعت بیمه رقم زد.

مفهوم مدل‌سازی متا اولین بار توسط کلینجان (Kleijnen, 1975) در زمینه مدل‌های شبیه‌سازی معرفی شد و مروری جامع بر روش، کاربردها و مزایای آن انجام شد. در سال‌های اخیر، روش‌های مدل‌سازی متا[22] (فرا مدل) به عنوان راه‌حل‌های مؤثر برای مقابله با چالش‌های محاسباتی ارزش‌گذاری پرتفوی‌های بزرگ VA ظهور کرده‌اند. حجازی در مقاله (Hejazi S. A., 2016) به بررسی استفاده از شبکه‌های عصبی برای ارزیابی سریع و کارآمد پرتفوهای بزرگ مستمری‌های متغیر پرداخت و نشان داد که این روش می‌تواند دقت ارزیابی را افزایش داده و زمان محاسبات را کاهش دهد. در نتیجه استفاده از تکنیک‌های یادگیری عمیق به متخصصان بیمه و مالی در تصمیم‌گیری‌های سریع‌تر و بهینه‌تر کمک می‌کند. پس از آن مقاله (Xu W. C., 2018) از روش‌های تطبیق لحظه ای به کمک روش‌های یادگیری ماشین استفاده می‌کند تا نوسانات و ریسک‌های مرتبط با این نوع پرتفوی‌ها را بهتر تحلیل کند. نتایج نشان می‌دهند که این روش‌ها می‌توانند تصمیم‌گیری‌های بهتری در مدیریت ریسک و بهینه‌سازی استراتژی‌های سرمایه‌گذاری فراهم کنند. جوان در پژوهش خود (Quan, 2022) به بررسی مدل‌های مبتنی بر درخت برای قیمت‌گذاری مستمری‌های متغیر پرداخت. نویسنده در این مقاله به بهینه‌سازی پارامترها و تحلیل‌های تجربی مدل‌های درختی (مانند درختان تصمیم‌گیری و مدل‌های تصادفی) می‌پردازد تا دقت قیمت‌گذاری این نوع مستمری‌ها را افزایش دهد. نتایج تحقیق نشان می‌دهد که این مدل‌ها می‌توانند ابزارهای مؤثری برای ارزیابی و مدیریت ریسک در بازار مستمری‌های متغیر باشند. اخیرا لیم در مقاله (Lim, 2025) به بررسی روش‌های قیمت‌گذاری پرتفوهای مستمری‌های متغیر با استفاده از شبکه‌های عصبی با عرض محدود و نامحدود پرداخت. نویسنده با تحلیل و مقایسه عملکرد این دو نوع شبکه عصبی در ارزیابی دقیق‌تر ارزش پرتفوی‌های مستمری‌های متغیر نشان داد که استفاده از شبکه‌های عصبی نامحدود می‌تواند به بهبود دقت و کارایی در قیمت‌گذاری و مدیریت ریسک این محصولات کمک کند.

[17] Guarantee Minimum Accumulated Benefit (GMAB)
[18] Guarantee Minimum Income Benefit (GMIB)
[19] Guaranteed Minimum Maturity Benefit (GMMB)
[20] Guaranteed Minimum Withdrawal Benefit (GMWB)
[21] College Retirement Equities Fund
[22] Meta-Modeling

در مدل‌سازی متا مبتنی بر شبیه‌سازی، یک مدل متا، نمایش ساده‌شده‌ای از یک مدل شبیه‌سازی پیچیده‌تر است. ایده اصلی این است که مدلی با استفاده از زیرمجموعه‌ای از داده‌ها ساخته شود که به طور مؤثری ویژگی‌های مجموعه داده کلی را ثبت کند و بتواند جایگزین شبیه‌سازی کامل شود. با به‌کارگیری این رویکرد، نیاز به ارزش‌گذاری مستقیم تمام بیمه نامه‌ها از طریق شبیه‌سازی‌های مونت کارلو به شدت کاهش می‌یابد. از آنجایی که تنها بخش کوچکی از قراردادهای مستمری متغیر تحت شبیه‌سازی کامل مونت کارلو قرار می‌گیرند، و مدل پیش‌بینی‌کننده متا بسیار سریع‌تر از فرآیند شبیه‌سازی عمل می‌کند، مدل‌سازی متا یک فرصت تحول‌آفرین ارائه می‌دهد.

اولین بار گان در (Gan G., 2015) از متا مدلینگ ساده برای قیمت‌گذاری پرتفوهای بزرگ مستمری‌های متغیر استفاده نمود. هدف مقاله، توسعه مدل‌های شبیه سازی به مدل های متا است که می‌توانند به سادگی رفتار پیچیده پرتفوها را تقریب بزنند و به تحلیلگران کمک کنند تا بدون نیاز به شبیه‌سازی‌های زمان‌بر، به نتایج دقیق‌تری دست یابند. این روش می‌تواند زمان مورد نیاز برای ارزیابی را کاهش دهد و به سازمان‌ها امکان می‌دهد منابع را حفظ کرده و در عین حال قابلیت‌های مدیریت ریسک خود را تقویت نمایند. مقاله شامل جمع‌آوری داده‌های تاریخی و تحلیل ویژگی‌های مستمری‌ها است و از تکنیک‌های آماری و مدل‌سازی برای شبیه‌سازی رفتار این محصولات استفاده می‌کند. پس از آن جوان در (Quan, 2022) به بررسی کاربرد مدل‌های مبتنی بر درخت در ارزیابی مستمری‌های متغیر پرداخت. برای این منظور با تنظیم بهینه پارامترهای مدل‌های درختی، مانند درخت تصمیم و جنگل تصادفی، دقت پیش‌بینی مدل‌ها را افزایش داد. تحلیل تجربی نتایج نشان داد که روش‌های یادگیری ماشین می‌توانند دقت بیشتری نسبت به روش‌های سنتی شبیه سازی ارائه دهند.

اما یکی از مشکلات اصلی متامدل‌های موجود، عدم هوشمندی آن‌هاست. به گونه‌ای که با تغییر داده‌ها یا پس از گذشت زمان، که الگوی داده‌ها دچار تغییر می‌شود، این متامدل‌ها نیاز به بروزرسانی دارند. از آنجا که این متامدل‌ها هوشمند نیستند، بروزرسانی لازم در این مدل ها انجام نمی‌شود. در نتیجه بدون قابلیت سازگاری خودکار[23]، اغلب تنظیمات حیاتی در این مدل ها از دست می‌روند و اثربخشی مدل کاهش می‌یابد. بر این اساس با هدف توسعه رویکرد متا مدلینگ (Gan G., 2015) و روش های یادگیری ماشین (Quan, 2022)، این پژوهش بر ارتقاء هوشمندانه مدل‌های متا از طریق هوش مصنوعی تمرکز دارد تا مدیریت پرتفوهای بزرگ مستمری متغیر را متحول سازد. برای این منظور با استفاده از الگوریتم‌های یادگیری ماشین پیشرفته، الگوریتم مدل‌های متا هوشمند با قابلیت خود-تنظیمی را برای ارزش‌گذاری پرتفوهای بزرگ توسعه می‌دهیم. سپس الگوریتم متا مدل طراحی شده هوشمند را بر روی پرتفوی مستمری متغیر پیاده‌سازی و ارزیابی می کنیم. نتایج نشان می دهند که متا مدل توسعه یافته هوشمند قادر به پاسخگویی خودکار به شرایط بازار و الگوهای داده‌ای در حال تحول با سرعت و دقت بالا هستند.

این رویکرد مبتنی بر هوش مصنوعی، دقت پیش‌بینی ارزش تضمین‌ها را به طور قابل توجهی بهبود بخشیده و همزمان اثربخشی مدیریت ریسک و کارایی محاسباتی را در سناریوهای پرتفوی بزرگ تقویت می‌کند. چارچوب مدل‌سازی متا هوشمند پیشنهادی، شرکت‌های بیمه را قادر می‌سازد تا به ارزش‌گذاری‌های پرتفوی قابل اطمینان‌تری دست یابند که به طور پویا با نوسانات بازار سازگار می‌شوند. در نتیجه، بیمه‌گران توانایی بهتری برای مدیریت پرتفوی‌های بزرگ مستمری متغیر خود و ارائه امنیت مالی برتر به بیمه‌گذاران کسب می‌کنند.

روش شناسی پژوهش

متامدل ها در واقع جایگزینی ساده برای یک مدل پیچیده می باشند که قادرند رفتار مدل پیچیده، زمان‌بر یا پرهزینه اصلی را با دقت قابل قبول و سرعت بسیار بیشتر تقریب بزنند. این روش در حوزه‌هایی مانند بیمسنجی، مهندسی مالی و مدیریت ریسک کاربرد دارد به‌ویژه زمانی که مدل اصلی نیازمند اجرای کند شبیه‌سازی‌های سنگین مونت‌کارلو است. به عنوان مثال در محاسبه ذخایر و ارزش قراردادها، قیمت گذاری و تحلیل تغییر پارامترها بسیار موثر است. متا مدل ها بعد از آموزش، قادر خواهند بود تا به جای مدل اصلی هزاران سناریو را در کسری از ثانیه اجرا کنند. در صنعت بیمه، این روش در کنار کاهش هزینه محاسباتی، می‌تواند سرعت تحلیل‌ها را چندین برابر کرده و در نتیجه تصمیم‌گیری را سریع‌تر و دقیق‌تر کند.

چارچوب الگوریتم متا مدل‌سازی برای ارزش‌گذاری تضمین‌ها شامل چهار گام کلیدی است: انتخاب بیمه‌نامه‌های نماینده، ارزش‌گذاری این بیمه‌نامه‌ها، توسعه متا مدل، و استفاده از متا مدل برای پیش‌بینی. در ابتدا، مجموعه‌ای کنترل‌شده از بیمه‌نامه‌های مستمری متغیر به عنوان نماینده انتخاب می‌شود تا ویژگی‌ها و رفتارهای متنوع کل پرتفوی را منعکس کند. سپس، شبیه‌سازی‌های مونت کارلو برای محاسبه ارزش منصفانه بازار هر بیمه‌نامه نماینده به کار گرفته شده و یک مجموعه داده ارزش‌گذاری شده ایجاد می‌گردد. در مرحله بعد، یک مدل پیش‌بینی (متا مدل) بر روی این مجموعه

[23] Automatic Adaptation

داده توسعه و آموزش داده می‌شود. در نهایت، از متا مدل برای پیش‌بینی ارزش منصفانه بازار مابقی بیمه‌نامه‌های مستمری متغیر در پرتفوی استفاده می‌شود.

بر این اساس این روش یک الگوریتم متا مدل هوشمند دو مرحله‌ای را معرفی می‌کند که برای بهبود کارایی و دقت در پردازش داده‌ها و مدل‌سازی اطلاعات طراحی شده است:

- مرحله اول: این مرحله بر انتخاب مدل‌های پیش‌بینی، انتخاب بهترین روش نمونه‌گیری، و تعیین اندازه مجموعه داده نماینده تمرکز دارد.
- مرحله دوم: با استفاده از روش نمونه‌گیری و اندازه نمونه انتخاب‌شده از مرحله اول، یک مجموعه داده نماینده ایجاد شده و مدل‌های پیش‌بینی آموزش داده می‌شوند.

مرحله اول مستلزم هزینه‌های محاسباتی بالایی است که کاربرد آن را محدود به تطبیق‌پذیری مدل هوشمند با تغییرات چشمگیر می‌کند. بنابراین، این مرحله تنها زمانی اجرا می‌شود که تغییرات قابل توجهی در مجموعه داده رخ دهد یا در فواصل زمانی مشخص نیاز به اجرای آن وجود داشته باشد. در مقابل، مرحله دوم به طور مداوم اجرا می‌شود تا ارزش‌گذاری‌های منظم و به‌موقع تضمین گردد. در این فرآیند، متا مدل هوشمند لزوم انجام شبیه‌سازی مونت کارلو بر روی تمام قراردادها را حذف می‌کند. در عوض، تنها بر روی بیمه‌نامه‌های شناسایی‌شده از طریق روش‌های نمونه‌گیری کارآمد تمرکز کرده و شبیه‌سازی‌ها را منحصراً بر روی بیمه‌نامه‌های منتخب اعمال می‌کند. علاوه بر این، برای ارزش‌گذاری تنها از مدل‌های پیش‌بینی انتخاب شده در مرحله اول استفاده می‌شود. این رویکرد نه تنها دقت را بهبود می‌بخشد، بلکه فرآیند را نیز ساده‌تر نموده و ارزش‌گذاری پرتفوی‌های مستمری متغیر را به طور قابل توجهی تسریع می‌بخشد. لازم به ذکر است اگر چه مرحله اول زمان بر و هزینه بر است ولی همچنان در مقایسه با روش شبیه سازی مونت کارلو سریعتر عمل می کند. بنابراین روش به صورت خودکار در صورت نیاز با اجرای مراحل دوگانه فوق تطبیق پذیری با داده های جدید و یا تغییرات با گذشت زمان را انجام می دهد.

بنابراین در صورت تغییرات ناچیز در داده ها تنها مرحله دوم اجرا می شود. فقط در صورت تغییرات چشمگیر و قابل توجه در داده ها و یا در بازه های زمانی بزرگ، مرحله اول اجرا می شود. لازم به ذکر است اگر چه مرحله اول زمان بر و هزینه بر است ولی همچنان در مقایسه با روش شبیه سازی مونت کارلو سریعتر عمل می کند. بنابراین روش به صورت خودکار در صورت نیاز با اجرای مراحل دوگانه فوق تطبیق پذیری با داده های جدید و یا تغییرات با گذشت زمان را انجام می دهد.

ایجاد مدل پیش‌بینی، تعیین اندازه نمونه، و انتخاب روش نمونه‌گیری مناسب برای مجموعه داده نماینده

توسعه یک مدل پیش‌بینی مؤثر شامل چندین مرحله کلیدی است که هر کدام برای اطمینان از دقت و کارایی ضروری هستند. این فرآیند با ایجاد یک مجموعه داده آزمایشی آغاز می‌شود، و به دنبال آن آماده‌سازی مجموعه‌های داده آموزشی صورت می‌گیرد. سپس داده‌ها استانداردسازی شده و از روش‌های کاهش بعد استفاده می‌شود. در این مرحله شبیه‌سازی‌های مونت کارلو انجام و مدل‌های پیش‌بینی با استفاده از داده‌های حاصل آموزش داده می‌شوند. این مدل‌ها ارزیابی شده و مدل‌هایی با بهترین عملکرد برای استفاده‌های بعدی انتخاب می‌شوند. در مرحله بعد، اندازه نمونه و روش نمونه‌گیری مناسب به منظور تضمین عملکرد بهتر مدل تعیین می‌شود.

مجموعه داده و انتخاب مجموعه‌های آموزش و آزمایش

مجموعه داده مورد استفاده در این پژوهش شامل ۱۹۰,۰۰۰ بیمه‌نامه مستمری متغیر است که اغلب به عنوان یک پرتفوی مصنوعی یا مجموعه داده ترکیبی شناخته می‌شود. این مجموعه داده توسط گان (Gan G., 2017) تولید شده است. برای دیدن جزئیات بیشتر در رابطه با نحوه تولید مجموعه داده ها می‌توانید به(Gan G., 2017) مراجعه کنید.

برای ایجاد مجموعه داده مصنوعی، از مقادیر متغیرها در جدول۱ استفاده شده است.

جدول۱ . مقادیر متغیرهای مورد استفاده برای ایجاد پرتفوی مصنوعی

Table 1. Variable Values used for Generating the Synthetic Portfolio

برای ایجاد یک سبد ترکیبی از قراردادهای مستمری متغیر، ۱۹ محصول مستمری با عناوین اختصار (برای اطلاعات بیشتر (Gan G., 2017) را ببینید)، به همراه توزیع جنسیت هر محصول که در جدول ۲ نشان داده شده است، انتخاب می شوند.

جدول۲ . توزیع جنسیت بر اساس نوع محصول مستمری

Table 2. Gender Distribution by Annuity Product Type

متغیر	مقدار
تاریخ تولد بیمه‌گذار	[1/1/1980 , 1/1/1950]
تاریخ صدور بیمه‌نامه	[1/1/2014 , 1/1/2000]
تاریخ ارزیابی	1/6/2014
مدت قرارداد (سال)	[15 , 30]
ارزش اولیه حساب	[500000 , 50000]
درصد زن	40%
کارمزد صندوق	46 (bps)، 57، 55، 45، 38، 10، 80، 60، 50، 30 (به ترتیب برای صندوق ۱ تا ۱۰)
کارمزد M&E	200 bps

مستمری / جنسیت	ABRP	ABRU	ABSU	DBAB	DBIB	DBMB	DBRP
زن	4068	3974	4054	3974	3948	4013	4002
مرد	5932	6026	5946	6026	6052	5987	5998
	DBRU	**DBSU**	**DBWB**	**IBRP**	**IBRU**	**IBSU**	**MBRP**
زن	3952	4038	4022	4007	4027	4007	3909
مرد	6048	5962	5978	5993	5973	5993	6091
	MBRU	**MBSU**	**WBRP**	**WBRU**	**WBSU**		
زن	3992	3980	3970	4076	3994		
مرد	6008	6020	6030	5924	6006		

متغیرهای موجود در پرتفوی مصنوعی از انواع محصولات مستمری در جدول ۳ مشخص شده اند.

جدول۳ . متغیرهای هر محصول مستمری

Table 3. Variables of Annuity Product

متغیر	توضیحات
fmv	ارزش بازار منصفانه
gender	جنسیت بیمه‌گذار
productType	نوع محصول
issueDate	تاریخ صدور بیمه‌نامه
matDate	تاریخ سررسید
ttm	زمان تا سررسید (سال)
birthDate	تاریخ تولد بیمه‌گذار
Age-insured	سن بیمه‌گذار (سال)
currentDate	تاریخ فعلی
baseFee	M&E کارمزد
riderFee	کارمزد الحاقیه
rollUpRate	نرخ جمع‌آوری
gbAmt	مزایای ضمانت شده
gmwbBalance	GMWB موجودی
wWithdrawalRate	نرخ برداشت تضمینی
withdrawal	برداشت تا این لحظه
FundValuei	i ارزش صندوق سرمایه‌گذاری
rollUp-status	وضعیت وجود ویژگی جمع‌آوری (۱ یا ۰)
wbWithdrawal-status	وضعیت وجود ویژگی برداشت (۱ یا ۰)

برای انتخاب داده‌های آزمایشی، از روش نمونه‌گیری طبقه‌بندی‌شده بدون جایگزینی استفاده می‌شود. این روش تضمین می‌کند که هر بیمه‌نامه تنها یک بار در مجموعه داده ظاهر شده و از انتخاب مجدد آن جلوگیری کرده و باعث می شود که مجموعه داده آزمایشی به‌اندازه کافی تنوع مجموعه داده را منعکس کند، بنابراین اعتبار ارزیابی مدل را افزایش می‌دهد. در مجموع ۲۰ بیمه‌نامه از هر محصول مستمری به عنوان داده آزمایش انتخاب شده که در نهایت ۳۸۰ بیمه‌نامه برای مجموعه داده آزمایش در نظر گرفته شد.

اطلاعات آماری ارائه شده در جدول ۴، بینش‌های ارزشمندی را در مورد مجموعه داده آزمایشی در مقایسه با کل پرتفوی نشان می‌دهد. قابل توجه است که میانگین ارزش gbAmt در مجموعه داده آزمایشی تقریباً ۳۰۲ هزار دلار است، در حالی که رقم مربوطه برای کل پرتفوی حدود ۳۱۲ هزار دلار است. علاوه بر این، مشاهده می‌گردد که ۲۵ درصد از مجموعه داده آزمایشی دارای gbAmt هستند که از ۱۶۲ هزار دلار تجاوز نمی‌کند، در حالی که این عدد برای کل پرتفوی ۱۷۹ هزار دلار است.
همچنین ۷۵ درصد از بیمه‌نامه‌ها در مجموعه داده آزمایشی مقدار gbAmt تا ۴۱۵ هزار دلار دارند که کمی کمتر از این مقدار برای کل پرتفوی (تقریباً ۴۲۵ هزار دلار) است. این ارقام نشان می‌دهند که به‌طور متوسط، بیمه‌نامه‌های موجود در مجموعه داده‌های آزمایشی مقادیر gbAmt کمتری نسبت به موارد موجود در پرتفوی دارند.

جدول۴ . آمار توصیفی برخی از متغیرها در مجموعه داده آزمایشی برحسب هزار دلار

Table 4. Descriptive Statistics of some Variables in the Test Dataset (in Thousand Dollars)

ttm	gbAmt	gmwbBalance	withdrawal	fmv	
15.30	302.36	38.29	18.83	93.45	mean
6.20	166.00	88.87	51.68	160.60	std
1.00	51.62	0	0	-28.05	min
11.00	162.39	0	0	4.00	25%
15.00	294.35	0	0	38.83	50%
19.25	415.68	0	0	96.01	75%
28.00	878.38	431.23	300.23	1073.08	max

از نظر مدت زمان تا سررسید، میانگین زمان تا سررسید کل پرتفوی حدود ۱۴/۵ سال است. در مقابل، بیمه‌نامه‌های موجود در مجموعه داده‌های آزمایشی دارای مدت زمان تا سررسید متوسط ۱۵/۳ سال هستند که نشان می‌دهد مجموعه داده آزمایشی در مقایسه با کل پرتفوی، مدت زمان تا سررسید کمی طولانی‌تر است.

میانگین withdrawal در مجموعه داده آزمایشی حدود ۱۹هزار دلار است، در حالی که میانگین withdrawal برای کل پرتفوی ۲۲ هزار دلار است که نشان‌دهنده اختلاف جزئی در میانگین withdrawal در بین بیمه‌نامه‌های موجود در مجموعه داده آزمایشی و پرتفوی کل است.

مشاهده می‌شود که پایین‌ترین مقدار gmwbBalance و withdrawal چه در مجموعه داده آزمایشی و چه در کل پرتفوی برابر صفر است، که این موضوع مربوط به بیمه‌نامه‌هایی است که دارای گزینه برداشت نیستند. به عبارت دقیق‌تر این مقادیر فقط در بعضی از انواع بیمه‌نامه مانند DBWB، WBRU ، WBRP وWBSU مقادیری بیشتر از صفر دارند.

به‌طور کلی، نزدیکی این اطلاعات بین مجموعه داده آزمایشی و پرتفوی نشان می‌دهد که مجموعه داده آزمایشی به عنوان نمونه‌ای مناسب از کل پرتفوی عمل می‌کند. این شباهت اطمینان را افزایش می‌دهد که اگر مدل‌های این پژوهش بتوانند مجموعه داده‌های آزمایشی را به‌طور دقیق ارزش‌گذاری کنند، احتمالاً زمانی که در کل پرتفوی اعمال می‌شوند، عملکرد خوبی خواهند داشت.

در ساخت یک مدل پیش‌بینی برای ارزش‌گذاری محصولات مستمری متغیر، کیفیت داده‌های آموزش نقش حیاتی ایفا می‌کند. اثربخشی مدل پیش‌بینی به شدت به داده‌هایی که با آن‌ها آموزش می‌بینند بستگی دارد، زیرا این داده‌ها اساس یادگیری مدل و پیش‌بینی‌های آتی را تشکیل می‌دهند. برای اطمینان از اندازه نمونه و روش نمونه‌گیری مناسب برای داده‌های آموزشی، یک رویکرد چند مرحله‌ای به کار گرفته می‌شود. بر اساس نتایج (Hejazi S. A., 2017)، (Gan G., 2017) ، (Gan, 2018) و (Quan, 2022) سه روش نمونه‌گیری برای تولید مجموعه داده آموزشی اعمال می‌شود: نمونه‌گیری تصادفی[۲۴]، نمونه‌گیری طبقه‌بندی‌شده[۲۵]، و نمونه‌گیری خوشه‌ای[۲۶]. در نتیجه، سه مجموعه داده متمایز ایجاد می‌شوند که به ترتیب با عناوین "TrainR"، "TrainS" و"TrainC" برچسب‌گذاری شده‌اند.

تعیین اندازه مناسب برای هر مجموعه داده مورد استفاده در فرآیند ارزش‌گذاری را می‌توان از طریق روش‌های تجربی یا با ارجاع به تحقیقات موجود در زمینه ارزش‌گذاری بیمه‌نامه‌های مستمری متغیر به دست آورد. مجموعه داده‌ای که بیش از حد بزرگ باشد می‌تواند منجر به زمان پردازش طولانی‌تر و هزینه‌های محاسباتی بالاتر شود که ممکن است با افزایش منافع بالقوه توجیه پذیر نباشد. بنابراین، ایجاد تعادل بین کارایی محاسباتی و کیفیت داده‌های آموزشی حیاتی است. در این تحقیق، اندازه‌های نمونه ۶۰۰، ۹۰۰ و ۱۲۰۰ بیمه‌نامه در نظر گرفته می‌شود. با استفاده از روش‌های نمونه‌گیری ذکر شده، داده‌ها از مجموعه داده اصلی استخراج می‌شوند و هر روش سه مجموعه داده متمایز تولید می‌کند (به عنوان مثال، TrainR1, TrainR2 , TrainR3برای نمونه‌گیری تصادفی).

از آنجا که سه روش نمونه‌گیری شامل نمونه‌گیری تصادفی، طبقه‌بندی‌شده و خوشه‌ای برای تولید مجموعه‌های داده متنوع مؤثر هستند، افزایش بیشتر تنوع مجموعه داده می‌تواند دامنه ارزیابی را گسترش داده و احتمال انتخاب اندازه نمونه و روش نمونه‌برداری مناسب را بهبود بخشد. در این

[24] Random Sampling
[25] Stratified Sampling
[26] Cluster Sampling

تحقیق، ایجاد تنوع در داده‌ها با ترکیب مجموعه داده‌های اصلی به روش‌های مختلف افزایش می‌یابد. به طور خاص، پس از انتخاب مجموعه داده‌های اولیه، مجموعه داده‌های جدیدی از طریق ترکیب‌های مختلف داده ها ایجاد می‌شوند. این ترکیب‌ها شامل داده‌های زیر است:

$$TrainR4 = TrainR1 \cup TrainR3,$$
$$TrainR5 = TrainR2 \cup TrainR3,$$
$$TrainS4 = TrainS1 \cup TrainS3,$$
$$TrainS5 = TrainS2 \cup TrainS3,$$
$$TrainC4 = TrainC1 \cup TrainC3,$$
$$TrainC5 = TrainC2 \cup TrainC3.$$

در مجموع، ۱۵ مجموعه داده آموزشی ایجاد می‌شود که شامل ۶۰۰، ۹۰۰، ۱,۲۰۰، ۱,۸۰۰ و ۲,۱۰۰ بیمه‌نامه هستند. یکی از دلایل کلیدی ترکیب مجموعه‌های داده، کاهش هزینه‌های محاسباتی است. به جای استفاده از شبیه‌سازی مونت کارلو برای قیمت‌گذاری بیمه‌نامه‌های جدید، یک مجموعه داده جدید با ترکیب بیمه‌نامه‌های قیمت‌گذاری شده قبلی تشکیل می‌شود. در مجموع، شبیه‌سازی مونت کارلو تنها برای قیمت‌گذاری ۸,۱۰۰ بیمه‌نامه برای مجموعه‌های داده آموزشی استفاده می‌شود. این رویکرد به کاهش هزینه محاسباتی کمک می‌کند و زمان و منابعی را که در غیر این صورت صرف شبیه‌سازی‌های گسترده می‌شد، صرفه‌جویی می‌نماید. با ایجاد مجموعه‌های داده متنوع از طریق ترکیب، این روش قابلیت اطمینان و سازگاری مدل‌های پیش‌بینی را افزایش می‌دهد و تضمین می‌کند که به کمک آن‌ها می‌توان طیف گسترده‌تری از سناریوها را با اثربخشی بیشتری مدیریت نمود.

آموزش مدل‌ها

در این تحقیق، از تابع StandardScaler در کتابخانه scikit-learn پایتون برای استانداردسازی مجموعه داده استفاده شده است. برای شناسایی مرتبط ترین متغیرها برای مدل، تابع SelectKBest با معیار mutual_info_regression به کار گرفته شده است که میزان ارتباط بین هر ویژگی و متغیر هدف در رگرسیون را اندازه‌گیری می‌کند. ویژگی‌هایی با امتیاز بالاتر از ۰/۱ انتخاب شدند، زیرا انتظار می‌رفت که سهم قابل توجهی در توسعه یک مدل پیش‌بینی قابل اطمینان داشته باشند. با توجه به تعداد نسبتاً کم متغیرهای باقی‌مانده، روش‌های استخراج ویژگی مانند تحلیل مؤلفه‌های اصلی غیرضروری تلقی شده و اعمال نشدند. ارزش‌گذاری تمام بیمه‌نامه‌ها برای هر دو مجموعه داده آموزشی و آزمایشی با استفاده از شبیه‌سازی‌های مونت کارلو انجام شد. مجموعه داده کامل، همانطور که در (Gan G., 2017) گزارش شده است، در حدود ۱۰۸/۳۱ساعت ارزش‌گذاری شد. با توجه به اینکه این مطالعه شامل ارزش‌گذاری حدود ۴/۷۶٪ از کل پرتفوی (تقریباً ۸,۴۸۰ بیمه‌نامه) است، تخمین زده می‌شود که زمان لازم برای ارزش‌گذاری این زیرمجموعه داده تقریباً ۳/۵۲ ساعت باشد.

چالش عمده در مدل‌سازی، وقوع بیش‌برازش است که در صورت هماهنگی بیش از حد مدل با داده‌های آموزشی اتفاق می‌افتد. در این حالت ماشین به جای یادگیری الگوهای کلی، ویژگی‌های خاص مجموعه آموزشی را به خاطر می‌سپارد. در حالی که مدل بر روی داده‌های آموزشی عملکرد خوبی دارد، در تعمیم دادن به داده‌های جدید و دیده نشده دچار مشکل می‌شود و قدرت پیش‌بینی آن را محدود می‌کند. برای مقابله با این مشکل، استفاده از یک مجموعه داده جداگانه برای انتخاب پارامتر، متمایز از مجموعه داده‌های آموزشی و آزمایشی، معمول است. با این حال، هدف این تحقیق توسعه مدل‌های متعدد برای یک تحلیل جامع است. تکیه بر یک مجموعه داده واحد برای تخمین پارامترها، تنوع مدل‌ها را محدود می‌کند، زیرا تعداد مجموعه‌های پارامتری ممکن را محدود می‌نماید.

چالش دیگر در این رویکرد، افزایش احتمالی هزینه‌های محاسباتی است، زیرا ایجاد مجموعه‌های داده جدید نیازمند شبیه‌سازی‌های مونت کارلو برای ارزیابی است. به دلیل این محدودیت‌ها، برای براورد پارامترها یک استراتژی کارآمدتر پیاده‌سازی خواهد شد که امکان توسعه طیف وسیع‌تری از مدل‌ها را فراهم می‌آورد. این استراتژی شامل استفاده همزمان از چندین مجموعه داده است که امکان براورد پارامتر را در مدل‌های مختلف در حین حفظ کارایی محاسباتی فراهم می‌کند. این رویکرد تضمین می‌کند که مدل‌ها در عین تنوع، قابل اطمینان باقی بمانند.

در ابتدا هر یک از مدل ها را به صورت $Model_m$ برای $m = 1, \dots, M$ نمایش می دهیم. در این پژوهش شش مدل $Model_m$ برای $m = 1, \dots, 6$ به ترتیب شامل مدل جنگل تصادفی(RF) [۲۷]، درخت تصمیم[28](DT)، K نزدیکترین همسایگی[۲۹] (KNN)، ماشین بردار گرادیان (SVM) [۳۰]، درخت تقویت‌یافته گرادیان(GBRT) [۳۱] و روش حافظه کوتاه مدت طولانی(LSTM) [۳۲] است. برای $m = 1, \dots, M$ ، پارامترهای هر مدل با استفاده از $K = 15$ مجموعه داده آموزشی تخمین زده می‌شوند. این امر منجر به تولید مجموعه‌ای با $M \times K$ مدل می‌شود که با $Model_{mk}$ نشان داده می‌شود. $Model_{mk}$ نشان‌دهنده مدل m -ام است که پارامترهای آن با استفاده از مجموعه داده آموزشی k-ام تخمین زده شده است.

پس از تخمین پارامترها، گام بعدی آموزش همه مدل‌های طراحی شده، $Model_{mk}$ با استفاده از $n = 1, \dots, N$ مجموعه داده آموزشی ساخته شده است. این فرآیند منجر به مجموعه‌ای متنوع از مدل‌ها می‌شود که هر یک بر روی زیرمجموعه‌های متمایزی از داده‌ها آموزش دیده‌اند. پس از آموزش، تعداد کل مدل‌ها برابر $M \times K \times N$ خواهد بود که با $Model_{mkn}$ نمایش داده می شود. $Model_{mkn}$ نشان‌دهنده مدل m-ام است که پارامترهای آن با استفاده از مجموعه داده آموزشی k-ام تخمین زده شده و با استفاده از مجموعه داده n-ام آموزش داده شده است.

به منظور براورد پارامترهای هر مدل، در جداول ۵ تا ۱۰ فهرستی جامع از پارامترها در کنار محدوده جستجوی مربوطه آنها دیده می‌شود که هدف، تعیین مقادیر بهینه برای هر مدل است. شایان ذکر است در صورتی که در جداول مذکور پارامتری وجود نداشته باشد، از مقدار پیش‌فرض آن پارامتر برای مدل استفاده خواهد شد. در طول این مرحله، از تمام مجموعه داده‌های آموزشی موجود استفاده می‌شود تا پارامترهای هر مدل تعیین‌شده به عنوان $Model_m$ برای $m = 1, \dots, M$ تخمین زده شود. در انتهای این مرحله، N مجموعه پارامتر (برابر با تعداد مجموعه داده آموزش) برای هر کدام از مدل‌های $Model_m$ نتیجه می‌شود.

جدول ۵. پارامترها و محدوده جستجو برای مدل RF

Table 5. Parameters and Search Ranges for RF Model

نام پارامتر	محدوده جستجو
n_estimators	50, 100, 300, 450, 500, 700
max_depth	2, 5, 10, 10, 20, 30
min_samples_split	2, 5
min_samples_leaf	1, 2

جدول ۶. پارامترها و محدوده جستجو برای مدل KNN

Table 6. Parameters and Search Ranges for KNN Model

نام پارامتر	محدوده جستجو
n_neighbors	2, 10, 20, 50
leaf_size	2, 5, 10
p	1, 2
weights	distance, uniform
metric	minkowski
algorithm	auto, ball_tree, kd_tree, brute

[27] Random Forest
[28] Decision Tree
[29] K-Nearest Neighbor
[30] Support Vector Machine
[31] Gradient Boosting Regression Trees
[32] Long Short Term Memory

جدول ۷. پارامترها و محدوده جستجو برای مدل GBRT

Table 7. Parameters and Search Ranges for GBRT Model

نام پارامتر	محدوده جستجو
n_estimators	10, 20, 50, 100, 200, 250, 300
max_depth	4, 8, 10
min_samples_split	2, 5
learning_rate	0.1, 0.01, 0.001

جدول ۸. پارامترها و محدوده جستجو برای مدل DT

Table 8. Parameters and Search Ranges for DT Model

نام پارامتر	محدوده جستجو
max_depth	2, 3, 5, 7, 10, 20, 30, 40, 50, 100
min_samples_split	2, 5, 7, 10
min_samples_leaf	1, 2, 5, 7, 10
max_features	sqrt, auto
criterion	absolute_error, friedman_mse

جدول ۹. پارامترها و محدوده جستجو برای مدل LSTM

Table 9. Parameters and Search Ranges for LSTM Model

نام پارامتر	محدوده جستجو
Units	50, 100, 150, 200
Number of Layers	3, 4, 5
Batch Size	16, 32, 64
Epochs	20, 50

جدول ۱۰. پارامترها و محدوده جستجو برای مدل SVR

Table 10. Parameters and Search Ranges for SVR Model

نام پارامتر	محدوده جستجو
C	0.1, 1
epsilon	0.01, 0.1, 0.5
max_iter	1, 10
kernel	poly, rbf, sigmoid
gamma	scale, auto
shrinking	True, False

در این تحقیق، از روش GridSearch برای براورد پارامترها و از معیار میانگین مربعات خطا (MSE)[۳۳] برای ارزیابی مدل استفاده می‌شود. جدول ۱۱ زمان‌های آموزش (بر حسب ثانیه) مدل‌های مختلف را ارائه می‌دهد و تفاوت‌های قابل توجه در پیچیدگی محاسباتی آن‌ها را برجسته می‌سازد. این تغییرات نشان می‌دهند که ویژگی‌های ذاتی هر مدل چگونه بر عملکرد و کارایی در مرحله آموزش تأثیر می‌گذارد. بخش قابل توجهی از پیچیدگی محاسباتی به پارامترهای خاص انتخاب شده برای هر مدل نسبت داده می‌شود. به عنوان مثال، در مدل جنگل تصادفی، یک پارامتر حیاتی تعداد درختان در جنگل است. با افزایش تعداد درختان، زمان آموزش نیز به دلیل محاسبات اضافی مورد نیاز برای ارزیابی هر درخت و تجمیع نتایج آن‌ها، افزایش می‌یابد. با بررسی زمان‌های آموزش، مشاهده می‌شود که مدل‌های KNN و DT زمان‌های آموزش نسبتاً کوتاهی، به ترتیب برابر با ۲/۰۰۵ ثانیه و ۱ ثانیه دارند. این زمان‌های آموزش سریع، سادگی این مدل‌ها را در مقایسه با همتایان پیچیده‌ترشان منعکس می‌کند. از سوی دیگر، مدل جنگل تصادفی به طور قابل توجهی زمان بیشتری برای آموزش نیاز دارد-۱۵۰/۸۶ ثانیه-که آن را در میان مدل‌های ارزیابی شده، در رتبه دوم از نظر مدت زمان آموزش قرار می‌دهد. برترین مدل از نظر زمانی، متعلق به مدل LSTM است که برابر با ۲۴۹۱/۸۶ ثانیه است. این زمان آموزش طولانی، نتیجه محاسبات سنگین در آموزش مدل‌های یادگیری عمیق است، به ویژه برای مدل‌هایی که برای پردازش داده‌های متوالی طراحی شده‌اند.

جدول ۱۱. زمان‌های آموزش (بر حسب ثانیه) برای مدل‌های مختلف

Table 11. Training Times for Different Models

مدل	زمان آموزش
RF	150.86
DT	1
KNN	2.005
SVR	86
GBRT	10.1
LSTM	2491.86

ارزیابی و انتخاب مدل‌های پیش‌بینی

حال علاقمندیم تا بدانیم کدام یک از مدل‌های مورد بحث در بخش قبلی برای اهداف این تحقیق مناسب‌تر هستند. صرف در نظر گرفتن یک سیستم رتبه‌بندی مبتنی بر معیارهای ارزیابی، برای تصمیم‌گیری آگاهانه کافی نیست. این محدودیت از آنجا ناشی می‌شود که روش‌های رتبه‌بندی اغلب در ثبت تفاوت‌های معنادار در عملکرد مدل‌ها ناتوان هستند و بنابراین یک رویکرد پیچیده‌تر برای توضیح و تفسیر مؤثر این تغییرات مورد نیاز است. علاوه بر این، هدف این تحقیق انتخاب یک مدل واحد به عنوان "بهترین مدل" نیست، بلکه هدف شناسایی مجموعه‌ای از مدل‌ها است که به طور جمعی عملکرد پیش‌بینی را بهبود بخشند. با انتخاب مجموعه‌ای متنوع از مدل‌ها، می‌توانیم از نقاط قوت هر یک از آن‌ها در قالب یک مدل برای بهبود کلی پیش‌بینی و مدیریت ریسک بهره‌مند شویم.

مدل‌ها بر اساس معیارهای ارزیابی به دو دسته مفید (مناسب) و غیرمفید (نامناسب) طبقه‌بندی می‌شوند. برای تسهیل این طبقه‌بندی، بازه‌های خاصی برای هر معیار تعریف شده است. این بازه‌ها از طریق تخصص و بینش متخصصان باتجربه در این حوزه تعیین می‌شوند. به عنوان مثال، سازمان‌هایی که تحمل ریسک کمتری دارند ممکن است محدوده‌های محافظه‌کارانه‌تری را برای معیارهای خود انتخاب کنند. در این مطالعه، بازه‌های

[33] Mean Squared Error

معیاری که در جدول ۱۲ ارائه شده‌اند، از نتایج (Quan, 2022) استخراج شده‌اند و برای معیارهای پرکاربرد مانند میانگین قدر مطلق خطا (MAE)[۳۴]، میانگین مربعات خطا (MSE)[۳۵]، میانگین خطا (ME)[۳۶] و ضریب (R^2)[۳۷] اعمال می‌شوند.

جدول ۱۲. محدوده معیارهای ارزیابی بر اساس نظر متخصصان

Table 12. Ranges of Evaluation Metrics based on Expertise

معیار	کران پایین	کران بالا
MAE	0	31.421
MSE	0	3278.5
ME	-2.213	2.213
R^2	84%	100%

پس از تعیین بازه‌ها برای معیارهای ارزیابی، مدلی که در داخل این بازه‌ها قرار گیرد به عنوان مدل **مفید** طبقه‌بندی می‌شود، در حالی که مدل‌هایی که خارج از این بازه‌ها قرار می‌گیرند، **غیرمفید** در نظر گرفته می‌شوند. از آنجایی که چندین معیار ارزیابی مورد استفاده قرار می‌گیرد، طبقه‌بندی یک مدل به عنوان مفید یا غیرمفید از طریق محاسبه مقدار میانگین این معیارها تعیین می‌شود. به طور خاص، برای هر مدل $m = 1, \dots, M$ و هر مجموعه داده $k = 1, \dots, K$ ، میانگین معیارهای ارزیابی مختلف به شرح زیر محاسبه می‌شود:

$$Mean_{MAE_{mk}} = \frac{1}{N}\sum_{n=1}^{N} MAE(Model_{mkn}),$$

$$Mean_{MSE_{mk}} = \frac{1}{N}\sum_{n=1}^{N} MSE(Model_{mkn}),$$

$$Mean_{ME_{mk}} = \frac{1}{N}\sum_{n=1}^{N} ME(Model_{mkn}),$$

$$Mean_{R^2{}_{mk}} = \frac{1}{N}\sum_{n=1}^{N} R^2(Model_{mkn}).$$

مقدار میانگین به ترکیب عملکرد مدل در سراسر معیارهای مختلف کمک می‌کند و ارزیابی جامع‌تری از مفید بودن آن را تسهیل می‌نماید.

در مرحله بعد، رتبه مدل $Model_{mk}$ که با $RANK_{Model_{mk}}$ نشان داده می‌شود به کمک رابطه زیر محاسبه می‌گردد:

$$RANK_{Model_{mk}} = R_{MAE}\left(Mean_{MAE_{mk}}\right) + R_{MSE}\left(Mean_{MSE_{mk}}\right) + R_{ME}(Mean_{ME_{mk}}) + R_{R^2}(Mean_{R^2{}_{mk}}),$$

که در آن داریم:

[34] Mean Absolute Error
[35] Mean Squared Error
[36] Mean Error
[37] R-squared

$$R_{MAE}\left(Mean_{MAE_{mk}}\right) = \{1, \quad Mean_{MAE_{mk}} \in [0, 31.421], 0, \quad Mean_{MAE_{mk}} \notin [0, 31.421],$$

$$R_{MSE}\left(Mean_{MSE_{mk}}\right) = \{1, \quad Mean_{MSE_{mk}} \in [0, 3278.5], 0, \quad Mean_{MSE_{mk}} \notin [0, 3278.5],$$

$$R_{ME}\left(Mean_{ME_{mk}}\right) = \{1, \quad Mean_{ME_{mk}} \in [-2.213, 2.213], 0, \quad Mean_{ME_{mk}} \notin [-2.213, 2.213],$$

$$R_{R^2}\left(Mean_{R^2{}_{mk}}\right) = \{1, \quad Mean_{R^2{}_{mk}} \in [0.84, 1], 0, \quad Mean_{R^2{}_{mk}} \notin [0.84, 1].$$

این امتیاز، عملکرد هر مدل را در تمام مجموعه‌های داده و معیارهای ارزیابی تجمیع کرده و یک رتبه‌بندی کلی از اثربخشی آن ارائه می‌دهد. اگر $RANK_{Model_{mk}} > \frac{n}{2}$ باشد، که در آن n تعداد معیارهای ارزیابی مورد استفاده برای سنجش عملکرد مدل است، مدل $Model_{mk}$ به عنوان یک مدل مفید برای ساخت مدل پیش‌بینی‌کننده طبقه‌بندی می‌شود. این بدان معناست که اگر بیش از نیمی از معیارهای ارزیابی نشان دهند که مدل عملکرد خوبی دارد، استفاده از آن مناسب تلقی می‌شود. در نتیجه، مدل‌های غیرمفید، یعنی آن‌هایی که رتبه‌بندی‌شان به این آستانه نرسد، از فرآیند مدل‌سازی حذف می‌شوند.

نتایج فرآیند ارزیابی و انتخاب مدل برای مدل‌های به کار رفته در جداول زیرارائه شده‌اند. در این جداول، در ستون با عنوان "P.DATASET" مشخص می‌کند که از کدام مجموعه داده برای براورد پارامترهای مدل استفاده شده است. ستون‌های برچسب‌گذاری شده MAE، ME،MSE و R^2 میانگین معیارهای ارزیابی را برای مدل‌هایی که با استفاده از تمام مجموعه داده‌های آموزشی موجود آموزش دیده‌اند، نمایش می‌دهند. مدل‌هایی که با علامت ستاره (*) مشخص شده‌اند، نشان می‌دهند که آستانه‌های امتیاز لازم را کسب کرده و بنابراین برای الزامات تحلیلی این تحقیق به عنوان مدل‌های مناسب طبقه‌بندی می‌شوند. جزئیات ارزیابی هر مدل را در بخش‌های بعدی بررسی خواهیم نمود.

تحلیل و ارزیابی مدل درخت تصمیم

جدول ۱۳ ارزیابی دقیقی از مدل‌های مختلف درخت تصمیم به کار رفته در این تحقیق ارائه می‌دهد. این جدول عملکرد مدل‌هایی را که با استفاده از مجموعه‌داده‌های مختلف آموزش دیده‌اند، نشان داده و نتایج ارزشمندی در مورد تأثیر روش‌های نمونه‌گیری مختلف بر دقت و کارایی مدل ارائه می‌دهد. در مجموع، هفت مدل به عنوان مدل‌های مناسب شناسایی شدند که در میان آن‌ها پنج مدل به بالاترین امتیاز ممکن دست یافتند. به طور خاص، مدل های درخت تصمیم که پارامترهای آن‌ها با استفاده از مجموعه‌داده‌های TrainC5، TrainR1، TrainR4، TrainR5، TrainS3 ، TrainS4 وTrainS5 براورد شده بودند، به عنوان مدل‌های مناسب شناخته شدند.

از دیگر مشاهدات کلیدی در جدول ۱۳ این است که مدل‌هایی که پارامترهایشان با استفاده از نمونه‌گیری خوشه‌ای براورد شده‌اند، نسبت به مدل‌هایی که با روش‌های نمونه‌گیری تصادفی و یا طبقه‌بندی شده براورد شده‌اند، عملکرد ضعیف‌تری داشتند. به طور خاص، مدل‌های مبتنی بر نمونه‌گیری تصادفی و طبقه‌بندی شده نتایج برتری را در سراسر معیارهای ارزیابی نشان دادند، که حاکی از آن است که انتخاب روش نمونه‌گیری تأثیر قابل توجهی بر دقت و قابلیت اطمینان مدل‌های درخت تصمیم دارد. هنگام تحلیل معیارهای ارزیابی، واضح است که اکثر مدل‌ها در یک محدوده قابل قبول قرار می‌گیرند که مناسب بودن آن‌ها را تأیید می‌کند. در واقع، اگر محدوده‌های قابل قبول معیارها کمی گسترش یابد، بخش بزرگ‌تری از مدل‌ها واجد شرایط مناسب بودن خواهند بود. این امر دقت و قابلیت اطمینان تمام مدل‌های درخت تصمیم ، به‌ویژه آن‌هایی که با نمونه‌گیری تصادفی و طبقه‌بندی شده آموزش دیده‌اند، را برجسته می‌سازد و آن‌ها را به ابزارهای قابل اطمینان برای ارزش‌گذاری قراردادهای مستمری متغیر تبدیل می‌کند. این موضوع بر اهمیت انتخاب روش‌های نمونه‌گیری مناسب برای بهبود عملکرد مدل تأکید دارد.

در میان مدل‌های ارزیابی شده، مدل درخت تصمیم که بهترین عملکرد را از خود نشان داد، مدلی بود که پارامترهای آن با استفاده از مجموعه داده TrainS5 تخمین زده شده بود. این مدل به MAE برابر با ۲۷/۱۳، MSE برابر با ۲۲۶۳/۵، R^2 برابر با ۰/۹۱، و ME برابر با ۰/۰۹- دست یافت. این معیارها نشان می‌دهند که این مدل به طور استثنایی دقیق و قابل اعتماد است و جایگاه برترین اجراکننده را در این تحلیل به خود اختصاص می‌دهد. در مقابل، مدل مبتنی بر مجموعه‌داده TrainC3 به طور قابل توجهی ضعیف‌تر از دیگر مدل‌ها عمل کرد. این امر به ویژه در مقدار معیار R^2 این مدل مشهود است که تنها ۰/۳۰ بود، که بسیار پایین‌تر از مقادیر R^2 حداقل ۰/۷۷ مشاهده شده برای دیگر مدل‌ها است. علاوه بر این، MAE برای مدل TrainC3 برابر با ۶۳/۹۳ بود. در حالی که بالاترین مقدار MAE در میان سایر مدل‌ها برابر ۳۷/۴۸ بود، این مغایرت‌ها بر تأثیر حیاتی کیفیت داده‌ها و روش‌های نمونه‌گیری مورد استفاده برای براورد پارامترها بر اثربخشی کلی مدل تأکید می‌کند.

جدول۱۳. میانگین و رتبه معیارهای ارزیابی برای مدل درخت تصمیم

Table 13. Average and Ranking of Evaluation Metrics for the DT Model

$RANK_{model}$	R_{R^2}	R_{ME}	R_{MSE}	R_{MAE}	R^2	ME	MSE	MAE	مجموعه داده
0	0	0	0	0	0.81	7.08	4825.39	35.08	TrainC1
2	1	0	0	1	0.87	2.34	3410.35	28.08	TrainC2
0	0	0	0	0	0.30	-7.55	17992.21	63.93	TrainC3
0	0	0	0	0	0.84	4.02	4240.93	32.74	TrainC4
3	1	1	0	1	0.86	1.46	3901.55	27.93	TrainC5*
3	1	0	1	1	0.89	-2.21	2770.54	29.97	TrainR1*
1	0	1	0	0	0.77	1.95	5865.57	36.19	TrainR2
1	0	1	0	0	0.82	1.70	4775.88	32.18	TrainR3
4	1	1	1	1	0.90	-1.79	2544.11	27.88	TrainR4*
3	1	0	1	1	0.89	-2.97	2894.57	28.82	TrainR5*
1	0	1	0	0	0.79	1.78	5428.20	37.48	TrainS1
2	1	1	0	0	0.85	1.49	3966.09	31.75	TrainS2
4	1	1	1	1	0.89	-0.93	2967.75	27.58	TrainS3*
4	1	1	1	1	0.89	0.75	2938.26	29.07	TrainS4*
4	1	1	1	1	0.91	-0.09	2263.51	27.13	TrainS5*

تحلیل و ارزیابی مدل جنگل تصادفی

جدول۱۴ نتایج ارزیابی مدل‌های مختلف جنگل تصادفی را ارائه می‌دهد. نتایج حاکی از آن است که ده مدل از میان سایر مدل‌ها به سطحی از دقت دست یافتند که برای طبقه‌بندی به عنوان مدل‌های پیش‌بینی‌کننده نهایی مناسب باشند. حائز اهمیت است که نمونه‌گیری طبقه‌بندی شده مؤثرترین رویکرد بود، زیرا تمام مدل‌هایی که پارامترهایشان با استفاده از این روش براورد شده‌اند، امتیازات لازم برای احراز صلاحیت را کسب کردند. به طور کلی، اکثریت مدل‌های جنگل تصادفی عملکرد خوبی از خود نشان دادند و هم سویی قوی مدل با اهداف تحقیق را به نمایش گذاشتند. حتی مدل‌هایی که واجد شرایط مناسب تشخیص داده نشدند نیز سطوح قابل قبولی از عملکرد را از خود نشان دادند.

هنگام مقایسه میانگین معیارهای ارزیابی مدل‌های RF با مدل‌های DT، آشکار است که مدل‌های RF به طور مداوم در چندین معیار کلیدی از همتایان DT خود پیشی گرفتند. این امر توانایی برتر مدل RF را در جذب و مدیریت الگوهای پیچیده داده‌ها نشان داده و در نتیجه برای مدل‌سازی جهت پیش‌بینی مؤثرتر می‌سازد. نتایج، قابلیت اطمینان و انعطاف‌پذیری مدل‌های RF را تأیید می‌کند و شواهد محکمی از مناسب بودن آن‌ها برای اهداف این تحقیق ارائه می‌دهد. توانایی آن‌ها در ارائه پیش‌بینی‌های قابل اعتماد و دقیق، پتانسیل آن‌ها را به عنوان ابزاری برتر برای تقویت قابلیت‌های پیش‌بینی و مدیریت ریسک تقویت می‌کند.

با این حال، ضروری است که تأکید شود هیچ یک از مدل‌های RF بر اساس معیار MSE امتیاز مورد نیاز را کسب نکردند. این محدودیت نشان می‌دهد که رویکرد RF، به ویژه از نظر MSE، با چالش‌هایی مواجه است و بنابراین، این مدل‌ها تحت این معیار خاص نمی‌توانند مناسب تلقی شوند. با این وجود، در فرآیند ارزیابی چندین معیار در نظر گرفته می‌شود تا مدل‌هایی که در سایر معیارها عملکرد خوبی داشتند، همچنان به عنوان مناسب طبقه‌بندی شوند. به جز مدلی که پارامترهای آن با استفاده از مجموعه داده TrainC3 تخمین زده شده بود، سایر مدل‌ها حداقل امتیاز ۲ را کسب کردند. این بدان معناست که این مدل‌ها حداقل دو معیار ارزیابی از میان سایر معیارها را برآورده کرده‌اند، که اعتبار و اثربخشی آن‌ها را برای مدل‌سازی پیش‌بینی در این مطالعه افزایش می‌دهد. با این حال، مجموعه داده TrainC3 به عنوان کم‌بازده‌ترین مدل مشخص شد و شکاف عملکرد قابل توجهی را در مقایسه با همتایان خود نشان داد. به عنوان مثال، در حالی که ضعیف‌ترین MAE در میان سایر مدل‌ها برابر ۲۷/۶۰ بود، مجموعه داده TrainC3 دارای MAE به طور قابل توجهی بالاتر از ۴۳/۳۸ بود که به عدم دقت آن در پیش‌بینی‌ها اشاره دارد. به طور مشابه، MSE برای سایر مدل‌ها، حتی در بدترین شرایط، حدود ۲۳۲۰ بود، در حالی که MSE مدل TrainC3 برابر با ۷۴۶۷ بود که نشان‌دهنده کاهش قابل توجه در عملکرد است. این مغایرت‌ها براهمیت کیفیت مجموعه داده و تأثیر روش‌های نمونه‌گیری بر عملکرد مدل‌های یادگیری ماشین تأکید می‌کند. مجموعه داده TrainC3، که با استفاده از نمونه‌گیری خوشه‌ای تولید شد، به طور قابل توجهی ضعیف‌تر از مدل‌هایی عمل کرد که از مجموعه‌داده‌های مشتق شده از نمونه‌گیری تصادفی یا طبقه‌بندی شده استفاده کردند.

جدول ۱۴. میانگین و رتبه معیارهای ارزیابی برای مدل جنگل تصادفی

Table 14. Average and Ranking of Evaluation Metrics for the RF Model

$RANK_{model}$	R_{R^2}	R_{ME}	R_{MSE}	R_{MAE}	R^2	ME	MSE	MAE	مجموعه داده
2	1	0	0	1	0.93	3.21	1871.86	27.60	TrainC1
3	1	1	0	1	0.95	0.061	1211.15	21.94	TrainC2*
0	0	0	0	0	0.71	9.18	7467.67	43.38	TrainC3
3	1	1	0	1	0.94	-0.45	1590.08	25.36	TrainC4*
3	1	1	0	1	0.95	-1.08	1204.33	21.34	TrainC5*
2	1	0	0	1	0.91	-3.49	3220.17	27.10	TrainR1
3	1	1	0	1	0.93	-1.27	1880.46	24.95	TrainR2*
3	1	1	0	1	0.94	-1.63	1598.86	32.72	TrainR3*
2	1	0	0	1	0.94	-2.53	1517.28	23.01	TrainR4
2	1	0	0	1	0.93	-2.47	1767.51	23.56	TrainR5
3	1	1	0	1	0.92	0.87	2028.88	26.89	TrainS1*
3	1	1	0	1	0.95	0.99	1379.29	22.57	TrainS2*
3	1	1	0	1	0.95	0.925	1288.99	21.74	TrainS3*
3	1	1	0	1	0.95	0.32	1305.84	22.02	TrainS4*
3	1	1	0	1	0.95	-0.03	1292.76	21.67	TrainS5*

تحلیل و ارزیابی مدل K نزدیک ترین همسایگی

جدول ۱۵ نتایج ارزیابی مدل‌های KNN را ارائه می‌دهد که به طور کلی حاکی از عملکرد بسیار مطلوب آن‌ها در برآورده کردن انتظارات است. مدل‌ها در اکثر معیارهای ارزیابی نتایج مطلوبی را نشان دادند و توانایی مدل KNN را در مدیریت مؤثر داده‌ها تایید نمودند. قابل توجه است که به استثنای مدلی که پارامترهای آن با استفاده از مجموعه داده TrainC3 براورد شده بود، سایر مدل‌های KNN امتیازات لازم را کسب کرده و برای اهداف این تحقیق مناسب طبقه‌بندی شدند. این امر قابلیت اطمینان مدل KNN را در امر پیش‌بینی، به ویژه هنگامی که مجموعه داده و روش‌های نمونه‌گیری مناسب استفاده می‌شوند، تقویت می‌کند.در مقایسه با مدل RF، تمام مدل‌های KNN از نظر معیار MSE عملکرد مطلوبی نشان دادند. در نتیجه، همه مدل‌ها، به استثنای یک مورد، امتیاز MSE قابل قبولی را کسب کرد. با این حال، توجه به این نکته حائز اهمیت است که تمام مدل‌های KNN دارای مقادیر منفی برای معیار ME (خطای میانگین) هستند. این امر نشان می‌دهد که این مدل‌ها تمایل دارند که مقادیر کم‌تر از مقادیر واقعی برای بیمه‌نامه‌ها پیش‌بینی‌کنند. علاوه بر این، اکثر مدل‌های KNN امتیاز مورد نیاز برای معیار ME را کسب نکردند.

جدول ۱۵. میانگین و رتبه معیارهای ارزیابی برای مدل KNN

Table 15. Average and Ranking of Evaluation Metrics for the KNN Model

$RANK_{model}$	R_{R^2}	R_{ME}	R_{MSE}	R_{MAE}	R^2	ME	MSE	MAE	مجموعه داده
3	1	0	1	1	0.91	-3.44	2127.71	27.60	TrainC1*
3	1	0	1	1	0.40	-2.54	1535.57	23.15	TrainC2*
0	0	0	0	0	0.64	-8.43	9064.96	49.96	TrainC3
3	1	0	1	1	0.92	-4.46	2035.82	26.50	TrainC4*
3	1	0	1	1	0.94	-3.66	1567.96	23.15	TrainC5*
3	1	0	1	1	0.88	-4.24	2888.21	30.03	TrainR1*
3	1	0	1	1	0.91	-4.46	2212.96	25.99	TrainR2*
4	1	1	1	1	0.93	-2.16	1826.70	24.47	TrainR3*
3	1	0	1	1	0.94	-2.72	1487.79	22.50	TrainR4*
3	1	0	1	1	0.93	-3.15	1823.31	23.31	TrainR5*
3	1	0	1	1	0.91	-3.40	2390.96	28.63	TrainS1*
3	1	0	1	1	0.94	-2.67	1619.19	24.07	TrainS2*
3	1	0	1	1	0.94	-2.33	1419.17	22.74	TrainS3*
4	1	1	1	1	0.93	-1.88	1663.45	23.71	TrainS4*
4	1	1	1	1	0.94	-1.71	1640.16	23.53	TrainS5*

تحلیل و ارزیابی مدل GBRT

جدول ۱۶ نتایج ارزیابی مدل‌های GBRT را ارائه می‌دهد. روش نمونه‌گیری طبقه‌بندی شده به طور مؤثری مجموعه داده‌های مناسبی را تشکیل داده است، به طوری که تمام مدل‌های GBRT که پارامترهای آن‌ها با استفاده از این مجموعه داده‌ها براورد شده، به عنوان مدل‌های مناسب طبقه‌بندی شدند. در مقابل، نمونه‌گیری خوشه‌ای نتایج به مراتب ضعیف‌تری را به همراه داشت. مدل‌های مشتق شده از این مجموعه داده خاص اغلب

نتوانستند معیارهای لازم را برآورده کنند و در نهایت به عنوان مدل‌های نامناسب طبقه‌بندی شدند. مشاهده می‌شود که قابل توجه‌ترین کاهش توسط مدلی که با مجموعه داده TrainC3مرتبط بود، نشان داده شد، که نتوانست هیچ امتیازی در معیارهای ارزیابی کسب کند. بر این اساس نتایج نشان می‌دهد اکثر مدل‌های مشتق شده از مجموعه‌داده‌های نمونه‌گیری طبقه‌بندی شده عملکرد استثنایی داشتند. شایان ذکر است که چهار مدل از پنج مدلی که با استفاده از این مجموعه‌داده‌ها توسعه یافته‌اند (با پارامترهای براورد شده از آن‌ها) با موفقیت حداکثر امتیاز را در معیارهای ارزیابی کسب کردند.

جدول ۱۶. میانگین و رتبه معیارهای ارزیابی برای مدل GBRT

Table 16. Average and Ranking of Evaluation Metrics for the GBRT Model

$RANK_{model}$	R_{R^2}	R_{ME}	R_{MSE}	R_{MAE}	R^2	ME	MSE	MAE	مجموعه داده
1	1	0	0	0	0.86	3.82	3584.07	34.03	TrainC1
4	1	1	1	1	0.91	0.75	2417.18	28.64	TrainC2*
0	0	0	0	0	0.55	-7.03	1147.79	54.07	TrainC3
2	1	1	0	0	0.87	2.16	3465.58	33.08	TrainC4
4	1	1	1	1	0.90	-0.25	2482.01	29.00	TrainC5*
1	1	0	0	0	0.87	-5.48	3288.52	31.63	TrainR1
2	1	1	0	0	0.86	0.72	3687.64	33.62	TrainR2
4	1	1	1	1	0.90	0.09	2676.92	29.74	TrainR3*
3	1	0	1	1	0.91	-3.17	2410.37	28.54	TrainR4*
3	1	0	1	1	0.88	-2.29	2901.18	29.55	TrainR5*
3	1	1	1	0	0.88	-0.04	3010.10	32.23	TrainS1*
4	1	1	1	1	0.90	1.93	2688.76	30.20	TrainS2*
4	1	1	1	1	0.90	-0.77	2590.62	29.34	TrainS3*
4	1	1	1	1	0.90	0.62	2484.23	29.02	TrainS4*
4	1	1	1	1	0.90	-0.05	2601.89	29.36	TrainS5*

تحلیل و ارزیابی مدل SVR

از جدول ۱۷ مشاهده می‌شود که هیچ یک از مدل‌های SVR اهداف خاص این پژوهش را برآورده نمی‌کنند. هیچ‌کدام از این مدل‌ها به امتیاز رضایت‌بخشی در معیارهای ارزیابی دست نیافتند، که نشان می‌دهد روش SVR با اهداف این مطالعه هم‌راستا نیست. هنگام مقایسه عملکرد مدل SVR با مدل‌های جایگزین، مانند KNN، RF،DT ، و GBRT، یک شکاف عملکردی قابل توجه آشکار می‌شود. معیارهای ارزیابی برای مدل‌های جایگزین به مراتب بهتر از مدل SVR است و اختلاف قابل ملاحظه‌ای را در اثربخشی می‌دهد. بنابراین، مدل SVR برای الزامات تحلیلی این پژوهش مناسب نیست و در مقایسه با سایر روش‌های مدل‌سازی، نتایج قابل اعتمادی ارائه نمی‌دهد.

جدول ۱۷. میانگین و رتبه معیارهای ارزیابی برای مدل SVR

Table 17. Average and Ranking of Evaluation Metrics for the SVR Model

$RANK_{model}$	R_{R^2}	R_{ME}	R_{MSE}	R_{MAE}	R^2	ME	MSE	MAE	مجموعه داده
0	0	0	0	0	-0.037	-46.83	26680.3	82.9	TrainC1
0	0	0	0	0	-0.03	-50.65	26528.3	81.3	TrainC2
0	0	0	0	0	-0.05	-40.19	27033.7	86.7	TrainC3
0	0	0	0	0	-0.25	-45.16	26359.3	82.3	TrainC4
0	0	0	0	0	-0.01	-48.70	26209.4	80.9	TrainC5
0	0	0	0	0	-0.07	-54.34	27442.2	82.9	TrainR1
0	0	0	0	0	-0.04	-52.51	26772.1	81.1	TrainR2
0	0	0	0	0	-0.04	-54.04	26811.5	80.6	TrainR3
0	0	0	0	0	-0.02	-49.86	26338.8	81.6	TrainR4
0	0	0	0	0	-0.02	-47.34	26338.8	82.4	TrainR5
0	0	0	0	0	-0.07	-55.51	26165.7	82.6	TrainS1
0	0	0	0	0	-0.03	-50.45	27630.4	82.4	TrainS2
0	0	0	0	0	-0.04	-53.27	26702.3	81.1	TrainS3
0	0	0	0	0	-0.02	50.06	26263.5	81.5	TrainS4
0	0	0	0	0	-0.01	-46.01	26005.2	82.4	TrainS5

تحلیل و ارزیابی مدل LSTM

بر اساس تحلیل عملکرد مدل LSTM در جدول ۱۸، علی‌رغم نیازهای محاسباتی بالای آن، نتایج رضایت‌بخش نبود. هیچ یک از مدل‌های LSTM به سطح عملکرد مورد نیاز برای اهداف این تحقیق دست نیافتند. این نتیجه نشان می‌دهد که اگرچه مدل‌های LSTM به طور کلی مدل‌های توانمندی هستند، اما نتوانستند دقت و قابلیت اطمینان لازم را برای مسئله خاص ما فراهم کنند.

هنگام مقایسه نتایج مدل‌هایLSTM با مدل‌های SVR، مشاهده می‌شود که مدل‌های LSTM به طور کلی در اکثر معیارهای ارزیابی عملکرد ضعیف‌تری دارند. تنها استثنا معیار ME است، که در آن برخی مدل‌های LSTM بهتر از همتایان SVR خود عمل کرده‌اند. با این حال، توجه به این مساله مهم است که این نقطه قوت مجزا در معیار ME، ضعف‌های کلی مدل‌های LSTM را جبران نمی‌کند. بررسی جامع نتایج به وضوح نشان‌دهنده عدم قابلیت اعتماد مدل‌های LSTM هنگام ارزیابی با مجموعه‌ای گسترده‌تر از معیارهای عملکرد است.

جدول ۱۸. میانگین و رتبه معیارهای ارزیابی برای مدل LSTM

Table 18. Average and Ranking of Evaluation Metrics for the LSTM Model

$\mathbf{RANK}_{model}$	$\mathbf{R}_{\mathbf{R^2}}$	$\mathbf{R}_{\mathbf{ME}}$	$\mathbf{R}_{\mathbf{MSE}}$	$\mathbf{R}_{\mathbf{MAE}}$	$\mathbf{R^2}$	**ME**	**MSE**	**MAE**	**مجموعه داده**
0	0	0	0	0	-0.03	4.64	26642	109.94	TrainC1
1	0	1	0	0	-0.05	1.86	27085	104.30	TrainC2
0	0	0	0	0	-0.02	-30.53	26326	90.23	TrainC3
0	0	0	0	0	-0.003	3.72	25890	102.63	TrainC4
0	0	0	0	0	-0.01	2.42	26199	102.57	TrainC5
0	0	0	0	0	-0.004	-11.21	25910	96.58	TrainR1
0	0	0	0	0	-0.005	-6.71	25939	97.26	TrainR2
0	0	0	0	0	-0.008	-4.12	26009	100.70	TrainR3
0	0	0	0	0	0	-4.16	25811	100.29	TrainR4
1	0	1	0	0	0.006	-2.15	25627	100.10	TrainR5
1	0	1	0	0	0.39	-0.94	15502	70.73	TrainS1
0	0	0	0	0	0.44	3.18	14267	68.22	TrainS2
0	0	0	0	0	0.45	4.11	13937	68.84	TrainS3
0	0	0	0	0	0.46	5.38	13684	66.38	TrainS4
0	0	0	0	0	0.47	5.33	13512	65.82	TrainS5

در نهایت در فرآیند توسعه مدل‌های پیش‌بینی، از مدل‌های مناسبی که در مراحل قبلی شناسایی شده‌اند، استفاده می‌شود. هر مدل پیش‌بینی خاص خود را تولید می‌کند و پیش‌بینی نهایی با میانگین‌گیری از این خروجی‌های متنوع به دست می‌آید. این رویکرد ترکیبی توانمندی و قابلیت اطمینان و اعتماد پیش‌بینی‌ها را افزایش می‌دهد. با ادغام بینش‌های حاصل از مدل‌های متعدد، طیف وسیع‌تری از الگوها و روندها در داده‌ها ثبت می‌شوند. این دیدگاه گسترده‌تر نه تنها دقت پیش‌بینی را بهبود می‌بخشد، بلکه از تصمیم‌گیری آگاهانه‌تر نیز پشتیبانی می‌کند. علاوه بر این، این استراتژی خطر اتکای بیش از حد به یک مدل واحد را کاهش می‌دهد و از این طریق یک چارچوب پیش‌بینی جامع‌تر و سازگارتر ایجاد می‌کند.

انتخاب بهترین روش نمونه‌گیری و اندازه نمونه برای مجموعه داده نماینده

انتخاب موثرترین مجموعه داده آموزشی و اندازه نمونه برای تمامی مدل‌های یادگیری ماشین بیان شده در بخش های قبلی، با محاسبه میانگین هر معیار ارزیابی در هر مدل آموزش‌دیده با مجموعه‌های داده مختلف آغاز می‌شود. این فرآیند برای هر مجموعه داده آموزشی که در نظر گرفته شده است، تکرار می‌شود. برای این منظور تعریف می کنیم:

$$Mean_{MAE}^{n} = \frac{1}{M \times K} \sum_{m=1}^{M} \sum_{k=1}^{K} MAE(Model_{mkn}),$$

$$Mean_{MSE}^{n} = \frac{1}{M \times K} \sum_{m=1}^{M} \sum_{k=1}^{K} MSE(Model_{mkn}),$$

$$Mean_{ME}^{n} = \frac{1}{M \times K} \sum_{m=1}^{M} \sum_{k=1}^{K} ME(Model_{mkn}),$$

$$Mean_{R^2}^{n} = \frac{1}{M \times K} \sum_{m=1}^{M} \sum_{k=1}^{K} R^2(Model_{mkn}),$$

که در آن، $Mean_{MAE}^{n}$ نشان‌دهنده میانگین معیار MSE برای مدل‌هایی است که با استفاده از مجموعه داده n -ام آموزش داده شده‌اند.

جدول ۱۹. میانگین معیارهای ارزیابی و رتبه بندی مجموعه داده ها بر اساس معیارهای ارزیابی

Table 19. Average Evaluation Metrics and Dataset Rankings based on Evaluation Metrics

مجموعه داده	MAE	MSE	ME	R^2	$RANK_{MAE}$	$RANK_{MSE}$	$RANK_{ME}$	$RANK_{R^2}$	میانگین نهایی
TrainC1	24.60	1828	-1.27	0.93	8	6	9	6	7.25
TrainC2	24.81	1851	-1.16	0.93	9	7	13	8	9.25
TrainC3	25.06	1874	-1.37	0.93	12	9	7	9	9.25
TrainC4*	24.00	1670	-1.43	0.94	1	1	5	1	2
TrainC5	24.31	1896	-1.32	0.93	5	10	8	11	8.50
TrainR1	24.99	1940	-1.70	0.93	11	11	2	10	8.50
TrainR2	24.34	1722	-1.50	0.93	6	5	3	5	4.75
TrainR3	24.17	1683	-1.26	0.94	3	3	11	3	5
TrainR4	24.23	1689	-1.45	0.93	4	4	4	4	4
TrainR5	31.02	1667	-1.41	0.94	2	2	6	2	3
TrainS1	35.53	3039	-0.49	0.88	13	13	14	13	13.25
TrainS2	35.95	4022	-0.45	0.84	15	15	15	15	15
TrainS3	35.17	3559	-2.52	0.85	14	14	1	14	10.75
TrainS4	24.97	2021	-1.19	0.92	10	12	12	12	11.5
TrainS5	24.41	1853	-1.27	0.93	7	8	10	7	8

در گام بعدی، مجموعه‌های داده بر اساس میانگین معیارهای ارزیابی خود رتبه‌بندی می‌شوند. در نتیجه، هر مجموعه داده بر اساس هر معیار ارزیابی، رتبه متمایزی کسب می‌کند. سپس، میانگین این رتبه‌ها برای هر مجموعه داده، محاسبه می‌گردد. مجموعه داده با کمترین میانگین نهایی به عنوان مناسب‌ترین مجموعه داده آموزشی انتخاب می‌شود. این مجموعه داده هم اندازه نمونه مناسب و هم روش نمونه‌گیری درست را تعیین می‌کند.

همانطور که در جدول ۱۹ نشان داده شده است، نمونه‌گیری طبقه‌بندی شده روش موثری برای انتخاب مجموعه‌های داده آموزشی نبود، زیرا مدل‌هایی که با این مجموعه داده‌ها آموزش دیدند، در مقایسه با سایرین، عملکرد ضعیفی از خود نشان دادند. به عنوان مثال، مدل‌های آموزش‌دیده با استفاده از مجموعه داده TrainS2 میانگین MAE تقریباً ۳۵/۹۵ را کسب کردند که منجر به پایین‌ترین رتبه در میان تمام مجموعه‌های داده بر اساس

این معیار شد. با این حال، توجه به این نکته مهم است که بیشتر مجموعه‌های داده عملکرد مشابهی را در معیارهای ارزیابی مختلف داشتند و همگی موفق به دستیابی به سطوح دقت قابل قبول شدند. در میان آن‌ها، TrainC4 به عنوان مناسب‌ترین مجموعه داده آموزشی انتخاب شد، زیرا پایین‌ترین میانگین معیار ارزیابی را داشت. این مجموعه داده در سه معیار ارزیابی متمایز رتبه اول را کسب کرد که نشان‌دهنده کیفیت بالای آن برای آموزش مدل است.

در نتیجه، روش نمونه‌گیری ایده‌آل، نمونه‌گیری خوشه‌ای تعیین شد. اندازه نمونه برای مجموعه داده آموزشی روی ۱۸۰۰ بیمه‌نامه تنظیم شد که به عنوان اندازه مجموعه داده نماینده برای متا مدل در نظر گرفته می‌شود.

ارزش‌گذاری پورتفوهای بیمه مستمری متغیر

برای ارزیابی عملکرد مرحله دوم الگوریتم متا مدل هوشمند، کل پورتفوی با استفاده از روش پیشنهادی ارزیابی می‌شود. از آنجایی که ارزش‌گذاری کل پورتفوی قبلاً انجام شده است، این امر امکان ارزیابی فوری نتایج و عملکرد روش را پس از اعمال آن فراهم می‌سازد. این فرآیند به شناسایی نقاط قوت و ضعف الگوریتم کمک کرده و مبنایی برای هرگونه تعدیل ضروری فراهم می‌کند. در نهایت، اطمینان حاصل می‌شود که روش پیشنهادی پورتفوی را به طور مؤثر و با نتایج دقیق و قابل اعتماد ارزش‌گذاری می‌کند.

در ابتدا، نمونه‌ای شامل ۱۸۰۰ بیمه‌نامه از کل پورتفوی با استفاده از نمونه‌گیری خوشه‌ای انتخاب می‌شود. سپس ارزش‌گذاری این نمونه از طریق شبیه‌سازی‌های مونت کارلو انجام می‌گیرد. در مرحله بعد، مدل‌های مناسب شناسایی شده، با استفاده از مجموعه داده نمونه آموزش داده می‌شوند و منجر به ایجاد مدل پیش‌بینی نهایی می‌گردند. این مدل شامل چندین مدل از جمله RF، DT، GBRT و KNN است. با ترکیب این مدل‌ها، از نقاط قوت فردی آن‌ها برای به حداکثر رساندن دقت پیش‌بینی استفاده می‌کنیم. در نهایت، مدل حاصل برای ارزش‌گذاری بخش باقی‌مانده پورتفوی به کار گرفته می‌شود.

در جدول ۲۰، عملکرد مدل پیش‌بینی نهایی با عملکرد مدل‌های تشکیل‌دهنده آن مقایسه شده است. نتایج به وضوح نشان می‌دهند که مدل نهایی، که پیش‌بینی‌ها را از چندین مدل مجزا جمع‌آوری می‌کند، به طور مداوم بهتر از هر یک از اجزای تشکیل‌دهنده خود عمل می‌کند. به طور خاص، هنگام استفاده از هر یک از مدل‌های تشکیل‌دهنده به صورت مستقل، دقت پیش‌بینی به طور قابل توجهی پایین‌تر از زمانی است که این مدل‌ها ترکیب می‌شوند. به عنوان مثال، مدل پیش‌بینی نهایی در مقایسه با سایر مدل‌ها، پایین‌ترین مقادیر MAE و MSE را کسب می‌کند. علاوه بر این، مقدار R^2 برابر با ۰/۹۴ نشان‌دهنده سطح بالایی از دقت پیش‌بینی است.

جدول ۲۰. ارزیابی مدل پیش‌بینی نهایی بر اساس عملکرد مدل‌های تشکیل‌دهنده

Table 20. Evaluation Metrics for the Final Predictive Model and Its Constituent Models

ME	R^2	MSE	MAE	مدل
-0.39	0.93	1959.32	24.10	**RF**
0.51	0.92	2292.72	26.36	**DT**
-0.31	0.93	1823.58	23.27	**KNN**
0.36	0.92	2292.41	28.41	**GBRT**
0.04	0.94	1824.60	23.96	**مدل نهایی**

بنابراین، متا مدل توسعه یافته در این مطالعه با موفقیت پورتفوهای بزرگ بیمه مستمری متغیر را با دقت بالا ارزش‌گذاری می‌کند. این امر نه تنها کارایی مدل پیش‌بینی نهایی را نشان می‌دهد، بلکه ارزش ترکیب چندین مدل را برای افزایش دقت پیش‌بینی برجسته می‌سازد.

در جدول ۲۱، ارزیابی مدل پیشنهادی در مقایسه با مدل‌های معرفی‌شده در مراجع (Gan G., 2017) و (Quan, 2022) ارائه شده است. لازم به ذکر است که این مقایسه تنها به ارزیابی خود مدل‌ها محدود نمی‌شود، زیرا این مدل‌ها از روش‌های پیشنهادی و نتایج نویسندگان آن مقالات نشات می‌گیرند. بنابراین، در واقع، روش‌شناسی مورد استفاده در این مطالعه و مطالعات دیگر مورد توجه است.

جدول ۲۱. مقایسه مدل پیشنهادی با مدل‌های معرفی‌شده قبلی به کمک معیارهای ارزیابی

Table 21. Evaluation Metrics for the Final Predictive Model and Comparative Models from Related Studies

Computational Time (Sec)	ME	R^2	MSE	MAE	مدل
0.13	-1.68	0.84	3278.5	31.42	**CART**
2.70	-2.21	0.92	1720.72	20.33	**Bagged trees**
4.69	-1.31	0.94	1214.90	19.34	**Gradient boosting**
0.25	-0.90	0.87	2754.85	26.54	**CIT**
1214.72	-1.60	0.89	2273.38	23.21	**CRF**
277.49	-0.81	0.86	3006.19	27.42	**Ordinary kriging**
23.44	-0.11	0.88	2554.42	27.77	**GB2**
32.07 (زمان کل پرتفوی)	0.04	0.94	1824.60	23.96	**مدل نهایی**

در رابطه با زمان مورد نیاز برای ارزش‌گذاری پورتفوی، مشاهده شد که فرآیند آموزش مدل‌ها و تولید پیش‌بینی‌ها و تمام مراحل چندگانه متا مدل ها برای قیمت گذاری کل پرتفوی بزرگ مستمری متغیر، ۳۲/۰۷ ثانیه به طول انجامید. این زمان شامل تمام مراحل ضروری برای آموزش و بهینه‌سازی مدل است که در مقایسه با زمان لازم برای ارزش‌گذاری پرتفوی بیمه با حدود ۸۴۸۰ بیمه نامه به روش شبیه سازی مونت کارلو، که در مجموع تقریباً ۳/۵۲ ساعت می باشد، همچنان قابل توجه است. با در نظر گرفتن تمام عوامل زمانی شامل شبیه سازی بیمه نامه های نماینده، این روش به عنوان یک راه حل سریع و کارآمد برای ارزش‌گذاری پرتفوی شناخته می‌شود که قادر است هم سرعت بالا و هم دقت بالا را ارائه دهد. از آنجا که زمان اجرای تمامی مدل ها در مقایسه با روش های شبیه سازی در کسری از ثانیه انجام می شود، در این پژوهش ارزیابی مجموعه مناسب انتخاب شده از روش های یادگیری ماشین بر اساس معیارهای دقت نتایج حاصله در روش پیشنهادی مدنظر می باشد.

مدل پیشنهادی این پژوهش دارای MAE برابر با ۲۳/۹۶ و MSE برابر با ۱۸۲۴/۶۰ است که هر دو نشان‌دهنده عملکرد بهتر در مقایسه با مدل‌های CART، CIT، CRF ، کریجینگ معمولی[۳۸] و GB2 هستند. علاوه بر این، مدل مورد مطالعه در این پژوهش دارای R^2 برابر با ۰/۹۴ است که بالاترین مقدار در میان مدل‌های مقایسه‌شده است. همچنین، مدل دارای ME برابر با ۰/۰۴ است که کمترین مقدار در میان مدل‌های بررسی‌شده بوده و نشان می دهد که مدل از نظر اریبی عملکرد بسیار خوبی دارد، در حالی که سایر مدل‌ها تمایل به کم‌برآوردی دارند. که این امر در کنار مقادیر پایین MAE وMSE دقت بالای مدل پیشنهادی در ارزش‌گذاری پورتفوی را تأیید می‌کند. لازم به ذکر است که اگر چه مدل های BT و GB دارای MAE وMSE کمتری هستند اما مقادیر ME به ترتیب برابر ۲/۲۱ – و ۱/۳۱ - حاکی از اریبی این دو مدل است.

جمع‌بندی و پیشنهادها

این پژوهش پتانسیل قابل توجه هوش مصنوعی را در تقویت عملکرد و قابلیت سازگاری متا مدل‌ها برای ارزش‌گذاری پورتفوهای بزرگ بیمه مستمری متغیر نشان می‌دهد. این پژوهش با پرداختن به محدودیت‌های متا مدل‌های سنتی، از جمله فقدان به‌روزرسانی خودکار در پاسخ به تغییرات پورتفوی، اهمیت ادغام الگوریتم‌های پیشرفته و روش‌های یادگیری ماشین را برای به‌روزرسانی متا مدل‌ها با شرایط جدید تاکید می‌کند. یافته‌ها این نکته را تایید می‌کنند که پیش‌پردازش داده‌ها، روش‌های نمونه‌گیری مؤثر و انتخاب مدل مناسب، کلید بهبود همزمان دقت و کارایی محاسباتی مدل‌های پیش‌بینی هستند.

[38] Ordinary kriging

علاوه بر این، نتایج نشان می‌دهند که اگرچه افزایش اندازه مجموعه داده برای متا مدل ممکن است منجر به بهبود دقت مدل پیش‌بینی نشود، اما استفاده از الگوریتم‌های بهینه‌سازی برای انتخاب و تنظیم پارامترهای مدل می‌تواند عملکرد مدل را بهبود بخشد. این کار نتایج ارزشمندی را برای شرکت‌های بیمه‌ای فراهم می‌کند که به دنبال مدیریت پورتفوهای بیمه مستمری متغیر خود با اطمینان بیشتر هستند و در نهایت، مدیریت ریسک بهتر و ارائه خدمات ارتقاء یافته به بیمه‌گذاران را تضمین می‌کند. با اتخاذ این رویکردهای نوآورانه، ارائه‌دهندگان بیمه می‌توانند به مدل‌های ارزش‌گذاری قوی‌تر و سازگارتر دست یابند که به خوبی برای ماهیت بازارهای مالی پویا مناسب هستند.

برای صنعت بیمه ایران، استفاده از روش‌های متامدلینگ و یادگیری ماشین می‌تواند مزیت رقابتی قابل‌توجهی ایجاد کند، اما این امر نیازمند رویکردی هدفمند و تدریجی است. در گام نخست، شرکت‌های بیمه باید بر بهبود کیفیت داده‌ها و یکپارچه‌سازی سامانه‌های اطلاعاتی تمرکز کنند؛ چرا که بدون داده‌های دقیق و قابل اعتماد، حتی پیشرفته‌ترین مدل‌ها نیز نتایج قابل اعتمادی ارائه نخواهند داد. در ادامه، بهره‌گیری از متامدل‌ها می‌تواند به کاهش چشمگیر زمان محاسبات بیمسنجی و امکان انجام سناریوسازی سریع در قیمت‌گذاری و مدیریت ریسک کمک کند. در نتیجه با قیمت گذاری به روز، دقیق و با سرعت می توان محصولات نوین وابسته به بازارهای سرمایه ایجاد و عرضه نمود.

.

تشکر و قدردانی

این پژوهش مستخرج از پایان نامه کارشناسی ارشد دانشجو، مرتضی حسینقلی پور به راهنمایی دکتر ساغر حیدری است. بدین وسیله از حمایت های دانشگاه شهید بهشتی قدردانی می شود.

سهم مشارکت هر نویسنده

تمام نویسندگان در انجام پژوهش مشارکت داشتند.

تعارض منافع

نویسندگان اعلام می‌کند که هیچ تضاد منافعی در مورد انتشار پژوهش ثبت شده وجود ندارد. علاوه بر این، موارد اخلاقی از جمله سرقت ادبی، رضایت آگاهانه، رفتار نادرست، جعل و/یا جعل داده ها، انتشار مضاعف و یا سوء رفتار به طور کامل توسط نویسندگان رعایت شده است.

فهرست منابع

Bauer, D. K. (2008). A universal pricing framework for guaranteed minimum benefits in variable annuities. ASTIN Bulletin: The Journal of the IAA, 38(2), 621-651.http://doi.org/10.1017/S0515036100015312

Chi, Y. a. (2012). Are flexible premium variable annuities underpriced? ASTIN Bulletin: The Journal of the IAA, 42(2), 559-574.http://doi.org/10.2143/AST.42.2.2182808

Gan, G. (2013). Application of data clustering and machine learning in variable annuity valuation. Insurance: Mathematics and Economics, 53(3), 795-801.http://doi.org/10.1016/j.insmatheco.2013.09.021

Gan, G. (2015). Application of Meta modeling to the valuation of large variable annuity portfolios. 2015 Winter Simulation Conference (WSC), 1103–1114. https://doi.org/10.1109/wsc.2015.7408237

Gan, G. (2018). Valuation of Large Variable Annuity Portfolios Using Linear Models with Interactions. Risks, 6(3), 71. https://doi.org/10.3390/risks6030071

Gan, G., & Valdez, E. A. (2017). Valuation of large variable annuity portfolios: Monte Carlo simulation and synthetic datasets. Dependence Modeling, 5(1), 354–374. https://doi.org/10.1515/demo-2017-0021

Gan, G., & Valdez, E. A. (2017). Regression Modeling for the Valuation of Large Variable Annuity Portfolios. North American Actuarial Journal, 22(1), 40–54. https://doi.org/10.1080/10920277.2017.1366863

Wilkie, D. (2003). Mary Hardy: Investment Guarantees: Modelling and risk management for equity-linked life insurance. John Wiley & Sons. ISBN 0-471-39290-1, 2003. ASTIN Bulletin, 33(2), 439–448. doi: 10.1017/S0515036100013568

Hejazi, S. A. (2016). A neural network approach to efficient valuation of large portfolios of variable annuities. Insurance: Mathematics and Economics, 70, 169-181.https://doi.org/10.1016/j.insmatheco.2016.06.013

Hejazi, S. A. (2017). A spatial interpolation framework for efficient valuation of large portfolios of variable annuities. arXiv preprint arXiv:1701.04134.https://doi.org/10.48550/arXiv.1701.04134

Kleijnen, J. P. (1975). A comment on Blanning's "Meta model for sensitivity analysis: the regression Meta model in simulation". Interfaces, 5(3), 21-23.

Lim, H. B. (2025). Valuation of variable annuity portfolios using finite and infinite width neural networks. Insurance: Mathematics and Economics, 120, 269-284.DOI: 10.1016/j.insmatheco.2024.12.005

Quan, Z. G. (2022). Tree-based models for variable annuity valuation: parameter tuning and empirical analysis. Annals of Actuarial Science, 16(1), 95-118.https://doi.org/10.1017/S1748499521000075

Shevchenko, P. V. (2016). A unified pricing of variable annuity guarantees under the optimal stochastic control framework. Risks, 4(3).

Xu, W. C. (2018). Moment matching machine learning methods for risk management of large variable annuity portfolios. Journal of Economic Dynamics and Control, 87, 1-20.DOI: 10.1016/j.jedc.2017.11.002

Xu, X. (2020). Variable annuity guaranteed benefits: An integrated study of financial modelling, actuarial valuation and deep learning. UNSW Sydney. https://doi.org/10.26190/unsworks/22295