JGTAI Journal of Global Trends in Artificial Intelligence

Volume 1, No. 1, June 2025-2026
Available online at: www.jgtai.com

A Review of Emerging Research Trends based on Machine Learning Models for Insurance Premium Prediction

Mr. Sachin Manekar
Assistant Professor
Department of Computer Sciences and Applications)
Mandsaur University, Mandsaur
Sachin.manekar@meu.edu.in

***Abstract*—Insurance premium prediction is a fundamental task in the insurance industry, as it enables insurers to accurately assess risk, determine fair pricing, and maintain financial sustainability. The relations that exist, perhaps in the majority of insurance industry data, are mostly nonlinear, and hard to accurately model with traditional actuarial methods. Therefore, machine learning (ML) techniques have been developed as efficient alternatives in order to provide improvement of the accuracy of premium prediction by using data-driven risk assessment. This survey provides an up-to-date overview of the latest developments in the field of insurance premium prediction based on ML algorithms, including the popular methods Support Vector Regression (SVR), Random Forest (RF), Linear Regression (LR), Gradient Boosting Machine (GBM), eXtreme Gradient Boosting (XGBoost), and Long Short-Term Memory (LSTM). Moreover, the paper discusses the research trends, including Explainable Artificial Intelligence (XAI), Large Language Models (LLM), telematics-based pricing, AI-guided personalized insurance, and MLOps, which leverage these advancements to improve transparency, automation, and decision support in premium prediction. Additionally, the survey highlights some of the challenges faced, including data quality, model interpretability, security and regulatory compliance, and summarizes the literature of the past few years and proposes future research directions. In this review, researchers and practitioners get a comprehensive overview of existing methodologies, technological advances, and open challenges in machine learning for insurance premium prediction.**

Keywords—Insurance Premium Prediction, Machine Learning, Risk Assessment, Explainable Artificial Intelligence (XAI), Large Language Models (LLMs).

I. INTRODUCTION

Insurance is in a dangerous and unpredictable world. The types and amounts of risk to which individuals, families, businesses, and properties are subject can vary greatly. These risks include the potential for physical harm, financial loss, and even death [1]. The most important things in life are having a life and being healthy. While it is true that most hazards can't be completely eliminated, the financial industry has come up with a plethora of solutions to help people and businesses mitigate these risks through the use of capital. An insurance policy that reduces or eliminates the financial impact of potential losses [2]. The importance of insurance companies being exact in measuring the amount covered by policies and the insurance charges that must be paid is directly related to the value of insurance in people's lives.

The idea of insurance has been around since the Middle Ages. Communities have always looked for ways to combine their resources so that members can weather unexpected losses [3]. Premiums are the amount a policyholder pays to transfer risk to an insurer, and this can have significant implications for an insurer's difference between profit and loss, insolvency and viability as well as the customers' price to purchase risk transfer channel in a specific risk-adjusted context. Premium calculations based on these methods typically cannot incorporate the increasing volume, complexity, and heterogeneity of datasets associated with modern insurance. The policy as a whole is affected by the exclusion of any component from the computation of amounts. Accuracy in carrying out these duties is, thus, of top priority. Insurers use experts in this field because it is where human error is most likely to occur. When determining the insurance premium, they also make use of several tools. To ensure that each prospective policyholder receives the appropriate coverage, accurate prediction is a critical procedure in the insurance sector. Insurance companies can better adjust premiums and make auto insurance affordable for more drivers when they have better prediction capabilities.

ML allows computers to learn and gain experience to comprehend the complexity of a problem, which allows them to execute tasks [4][5]. In most cases, computers can figure out how to do things that are really challenging to program. Using ML techniques, insurance companies can quickly categorize potential clients according to their risk level [6]. ML also helps insurance businesses with other things, such as creating effective marketing campaigns and setting fair premium prices for their products [7]. ML algorithms excel at modeling complicated, non-linear relationships in massive datasets without relying on overly conservative statistical assumptions.

A. *Structure of Paper*

This paper is structured as follows for the rest of it. Section II explains how insurance premiums are predicted, while Section III delves into ML methods. Section IV reviews emerging research trends and challenges, followed by the literature review, research gap, and the conclusion with future research directions

II. Insurance Premium Prediction: Concepts and Foundations

The fundamental reason for the existence of insurance is to offer financial protection against various risks; this mechanism serves the inherent human need for security. They are able to accomplish this by exchanging the risk they are taking on for a monetary premium. Nevertheless, this is a complex process because the insurance premium varies according to the seriousness of the risk or because each policyholder pays a different premium. Basically, insurance is a way for people to transfer risk. The insurer agrees to pay a predetermined sum to the policyholder in the case that a specified event takes place in return for the policyholder paying a premium. two types of insurance policies: life and non-life [8]. Not only that, but it also helps with financial planning and provides much-needed support to families. Insurance against non-life events shields policyholders from financial collapse in the event of an accident or illness. It follows that non-life insurances include travel, health, vehicle, and property policies.

A. Role of Insurance Premium

A policyholder pays a certain amount to an insurance company each month in exchange for the promise of financial protection in the event of a loss. A fund is established by collecting premiums from various policyholders. This fund is then utilized to compensate insured individuals in the event that covered events, like accidents, illnesses, or property damage, take place. Optimal pricing strategies allow insurance businesses to better assess risks, reduce financial losses, and increase operational efficiency. At the same time, policyholders get premiums that are more reflective of their actual risk levels. These are some types of insurance:

1) Health Insurance

Medical care, including hospitalization, diagnostics, treatment, medication, and preventative services, can be expensive. Fortunately, health insurance can help pay these costs [9][10]. Policyholders pay a premium to the insurer, which assumes the financial risk associated with eligible healthcare expenses

2) Life Insurance

An insurance policy that pays out a certain amount to heirs upon the policyholder's death [11]. Term and permanent policies are the two most common kinds of life insurance. Protect loved ones financially for a set amount of time with term life insurance. Those planning to leave an inheritance to loved ones after death may find that permanent life insurance, with its more extensive coverage, is the way to go. Pick from a variety of permanent policies, such as burial, variable universal, survivorship, whole, and universal life insurance.

3) Auto insurance

The insurance provider review driving record, including infractions, parking tickets, license suspensions, and accidents, when applying for vehicle insurance [12]. A motorist's premium lower if their record is free of accidents and infractions compared to a driver whose record is filled with these things [13].

B. Insurance Premium Calculation

The process of determining an insurance premium is not uniform, but there are certain commonalities. The underwriting fee, marginal costs, policy administration, and a buffer zone should all be included. What this means is that the information the policyholder provides to rating engines is what ultimately determines the cost of the insurance premium. The total cost of the insurance includes both the premium and any other expenses related to the policy, such as the cost of a line of coverage.

C. Insurance Premium Ratemaking

The term "ratemaking" describes this procedure well. Insurance firms determine the prices that policyholders pay for products or individual policies through a process called ratemaking. The insurers perform the process of their ratemaking to decide the premiums that are reasonable, limited within the framework of basic necessities and within the limits of their willingness to accept risk. This procedure consists of two parts: a priori ratemaking and a posteriori ratemaking [14].

III. Machine Learning for Insurance Premium Prediction

Data and algorithms are used in ML, a branch of AI and computer science, to mimic human learning [15]. These algorithms can help discover important features from data mining by providing data classification or prediction through statistical analysis. The results of these insights, when applied properly, can significantly affect critical growth metrics in companies and applications. The study shows that only two factors, age and smoking, have a significant impact on the amount of insurance. Smoking is the most important. However, certain factors such as location and climate are important. The results of the research show that age and smoking status are the primary factors that influence insurance amounts, and smoking has the greatest effect.

A. Machine Learning Models for Insurance Premium Prediction

The insurance premium forecasting ML models employed by leveraging historical data and risk-related attributes to enhance pricing accuracy and support data-driven decision-making.

1) Support Vector Regression (SVR)

Traditional ML methods, such as support vector machines (SVMs), use appropriate kernel functions and parameters to do regression and classification assessments. A linear, sigmoid, polynomial, or Gaussian function could be one of these. For continuous data prediction, SVR—a subset of SVM techniques—is useful. The purpose of SVR's hyperplane construction is to preserve small sample points from the hyperplane.

2) Random Forest (RF)

The RF ensemble is a kind of ML that makes use of several decision trees. Due to its high anti-interference capacity, RF is used by many academics to conduct regression analysis studies. The basic idea behind RF for regression analysis is to randomly select a subset of characteristics and data from the healthcare cost data in order to construct each decision tree. In order to generate a prediction, the trained RF uses the test set as input and averages the outcomes of all the decision trees.

3) Long Short-Term Memory (LSTM)

Sequential data processing is where LSTM neural networks shine. As an RNN model optimization, it shares RNN's high-quality features. The LSTM is able to accomplish forgetting and remembering while preserving data because it incorporates input gates, forgetting gates, and output gates into its structure. This allows it to overcome the limitations of RNNs' single memory overlay technique.

4) Linear Regression

Linear regression is the simplest supervised ML approach that aims for a continuous output. Linear Regression is not meant to classify data into separate groups, but rather to generate predictions within a continuous range.

5) Gradient Boosting Machine (GBM)

The GBM class of ML algorithms is strong and has proven effective in many real-world scenarios. GBM's ability to handle complex data interactions well and its excellent predicted performance have earned it a lot of respect. A lot of people use it for things like Kaggle tournaments and when they need to make precise predictions in the real world.

6) eXtreme Gradient Boosting (XGBoost) Model

The XGBoost model was an improved version of the original gradient boosting model, developed into one that was more efficient and scalable. The purpose of XGBoost is to use the boosting approach to increase processing speed and model performance [16]

IV. Machine Learning for Insurance Premium Prediction

Data and algorithms are used in ML, a branch of AI and computer science, to mimic human learning [15]. These algorithms can help discover important features from data mining by providing data classification or prediction through statistical analysis. The results of these insights, when applied properly, can significantly affect critical growth metrics in companies and applications. The study shows that only two factors, age and smoking, have a significant impact on the amount of insurance. Smoking is the most important. However, certain factors such as location and climate are important. The results of the research show that age and smoking status are the primary factors that influence insurance amounts, and smoking has the greatest effect.

A. Machine Learning Models for Insurance Premium Prediction

The insurance premium forecasting ML models employed by leveraging historical data and risk-related attributes to enhance pricing accuracy and support data-driven decision-making.

1) Support Vector Regression (SVR)

Traditional ML methods, such as support vector machines (SVMs), use appropriate kernel functions and parameters to do regression and classification assessments. A linear, sigmoid, polynomial, or Gaussian function could be one of these. For continuous data prediction, SVR—a subset of SVM techniques—is useful. The purpose of SVR's hyperplane construction is to preserve small sample points from the hyperplane.

2) Random Forest (RF)

The RF ensemble is a kind of ML that makes use of several decision trees. Due to its high anti-interference capacity, RF is used by many academics to conduct regression analysis studies. The basic idea behind RF for regression analysis is to randomly select a subset of characteristics and data from the healthcare cost data in order to construct each decision tree. In order to generate a prediction, the trained RF uses the test set as input and averages the outcomes of all the decision trees.

3) Long Short-Term Memory (LSTM)

Sequential data processing is where LSTM neural networks shine. As an RNN model optimization, it shares RNN's high-quality features. The LSTM is able to accomplish forgetting and remembering while preserving data because it incorporates input gates, forgetting gates, and output gates into its structure. This allows it to overcome the limitations of RNNs' single memory overlay technique.

4) Linear Regression

Linear regression is the simplest supervised ML approach that aims for a continuous output. Linear Regression is not meant to classify data into separate groups, but rather to generate predictions within a continuous range.

5) Gradient Boosting Machine (GBM)

The GBM class of ML algorithms is strong and has proven effective in many real-world scenarios. GBM's ability to handle complex data interactions well and its excellent predicted performance have earned it a lot of respect. A lot of people use it for things like Kaggle tournaments and when they need to make precise predictions in the real world.

6) eXtreme Gradient Boosting (XGBoost) Model

The XGBoost model was an improved version of the original gradient boosting model, developed into one that was more efficient and scalable. The purpose of XGBoost is to use the boosting approach to increase processing speed and model performance [16]

V. Literature Review

Recent research on the prediction of insurance premiums using machine learning is summarized, along with the methods used, major results, and limitations of the current approaches. It provides an overview of the current state of the art in intelligent insurance pricing research and outlines potential future advances in this area.

Hasiboyic Hasic and Tanovic (2026) Motor third-party liability insurance is required in most countries, however, it is very hard to keep customers of a single insurance company. They offer a hybrid churn prediction system combining accuracy and interpretability. They apply ensemble learning approaches, in which they use CatBoost, which is trained with a real Auto Liability dataset. The main contribution of this work is the integration of the Shapley Additive explanations (SHAP) method for deconstructing model predictions. This approach provides global insight into key churn drivers – such as premium fluctuations and vehicle age – and local explanations at the individual policy level [17].

Mishra et al. (2026) used five ML models to more accurately calculate the premium prices, considering factors such as physical health, personal details, and environmental conditions. The models used in this study are LR, DT, KNN, RF, and XGBoost. The data used in this study consists of information from 1,338 people, including their age, BMI, smoking habits, their location, and air and water quality. The Linear Regression model was used as the basis, and the score was [18]

Mahaveerakannan and Arunraj (2025) compared the suggested optimization method to RF and Gradient Boosting, it performed better than the others, leading to a 4.3% increase in accuracy and a lower RMSE of 2510.23. When comparing different methods for estimating insurance premiums, ensemble-based models clearly come out on top. Their suggested optimization approach improves prediction performance by enhancing feature selection and hyperparameters. Two major indicators include current smoking status and BMI [19].

Jaishree, Kokila and Hegde (2025) provide a machine learning system that can predict health insurance costs using a user's age, gender, BMI, family size, smoking status, and geography as well as other lifestyle and demographic variables. Multiple regression models, including XGBoost, LR, DT, RF, and Gradient Boosting, were trained using preprocessed data from Kaggle. Applying standard performance evaluation criteria yielded a result of 0.9548. That users to gain access to the system [20].

Hema Ambiha and Subithra (2024) investigations into vehicle insurance claims considerably lessening the need for human labor and monetary losses. The use of efficient classifiers is critical in establishing the veracity of a claim. The first and most important stage in detecting false claims is accident prediction. Several decision tree classifiers, such as Simple CART, SVM, and LMT, are employed by the WEKA tool to analyze collisions and claims. When comparing the accuracy of these classifiers' execution, Random Forest stands out as the most effective in forecasting the occurrence of accidents and producing the best results [21].

Kumar et al. (2024) examine and compares four approaches to pure premium prediction: GLM, AGLM, XGBoost, and neural networks. The study analyzes and contrasts the models' strengths and shortcomings in property and casualty insurance pricing and assesses their quantitative and qualitative performance on a test set. ML offers obvious benefits in actuarial pricing, notwithstanding these challenges. Utilizes vast amounts of data and sophisticated analytics to enhance risk profiles, pricing precision, profitability, and market competitiveness. To achieve these outcomes, however, a holistic strategy that integrates the technical complexity of the ML solutions with insurance's regulatory and operational constraints is required [22]

Ahmad. Agarwal and Ansari (2023) utilize ML techniques such as RF and LR, this work proposes an appropriate approach to generate results with a high degree of accuracy and a low relative error. Along with this, it shows how to create a subset of data specifically for testing and training ML systems. The predicted value is compared to the true value of the simulated data to determine the usefulness of the suggested method. By estimating the total risk of health care and medical system expenses, insurance companies can establish reliable financial mechanisms, like payroll taxes or monthly premiums, to cover the medical benefit agreements outlined in insurance policies [23].

Bora et al. (2022) include two ML algorithms: Multiple LR and RF. The expected results are then explained using XAI. In the first place, they lay out the basics using model-specific methods built on the Interpret library from Microsoft. By utilizing domain experts' knowledge to assess the most impactful aspects and contribute their insights, these methods can aid in validating the prediction models [24]

Table I provides a summary and comparison of current research on insurance premium prediction using ML approaches. It gives a quick overview of the recent developments and research trends in this field, with a summary of the approaches, main results, benefits, drawbacks, and future recommendations of each study.

TABLE I. COMPARATIVE ANALYSIS OF RECENT STUDIES ON INSURANCE PREMIUM PREDICTION

Reference	Approach	Findings	Advantages	Challenges	Recommendations
Hasiboyic Hasic and Tanovic (2026)	CatBoost with SMOTE and SHAP for interpretable insurance prediction	CatBoost achieved higher AUC-ROC than Logistic Regression, while SHAP provided both global and local explanations of predictions.	High predictive accuracy, effective handling of class imbalance, and improved model interpretability.	Focuses on customer churn rather than direct premium prediction; computational cost of XAI methods.	Extend the framework to insurance premium prediction and incorporate real-time explainable decision support.
Mishra et al. (2026)	LR, DT, KNN, RF, and XGBoost using health, demographic, and environmental features	Ensemble learning models produced more accurate insurance premium predictions than baseline Linear Regression.	Incorporates multiple risk factors such as BMI, smoking, air quality, and water quality to improve prediction accuracy.	Limited dataset (1,338 records) and restricted validation across diverse populations.	Expand the dataset and evaluate deep learning and Explainable AI techniques for robust premium prediction.
Mahaveerakannan and Arunraj (2025)	Metaheuristic optimization with Random Forest and Gradient Boosting	Optimized ensemble models achieved R^2 = 92.5% and RMSE = 2510.23, outperforming conventional ML models.	Improved feature selection, optimized hyperparameters, and enhanced prediction performance.	Higher computational complexity and dependence on optimization techniques.	Investigate lightweight optimization methods and validate the model on larger insurance datasets.
Jaishree, Kokila and Hegde (2025)	LR, DT, RF, Gradient Boosting, and XGBoost on a medical insurance dataset	Ensemble models achieved high prediction accuracy, with an overall score of 0.9548.	Accurate premium estimation using demographic and lifestyle features; practical implementation.	Limited to a Kaggle dataset with restricted real-world validation.	Test the models on real-world insurance datasets and incorporate additional customer risk factors.
Hema Ambiha and Subithra (2024)	Simple CART, SVM, LMT, and Random Forest using WEKA	Random Forest achieved the highest accuracy in accident prediction, while SVM performed best for fraud detection.	Reduces manual effort and improves insurance claim analysis.	Focuses mainly on fraud detection rather than premium prediction.	Integrates fraud detection and premium prediction into a unified insurance analytics framework.
Kumar et al. (2024)	Comparative evaluation of GLM, AGLM, XGBoost, and Neural Networks	XGBoost and Neural Networks outperformed traditional actuarial models in premium prediction accuracy.	Better risk assessment, pricing precision, and business competitiveness.	Limited model interpretability and regulatory concerns.	Develop interpretable and regulation-compliant machine learning models for actuarial pricing.

Ahmad, Agarwal and Ansari (2023)	LR and RF for insurance premium prediction	Random Forest achieved higher accuracy and lower prediction error than Linear Regression.	Simple implementation and effective premium prediction.	Limited dataset and evaluation in a single insurance domain.	Validate the approach on larger datasets and compare with advanced ensemble models.
Bora et al. (2022)	Multiple Linear Regression, Random Forest, and XAI (LIME, SHAP)	Explainable AI successfully interpreted premium prediction models and highlighted influential features.	Improves transparency, trust, and model interpretability.	Limited to traditional ML models and health insurance data.	Integrate advanced ensemble and deep learning models with scalable XAI techniques.

Research Gap: Despite the success of ML in predicting insurance premiums, there are still some gaps in the research. Existing literature mainly concentrates on prediction accuracy in the traditional and ensemble models, and has less focus on interpretability, fairness, and transparency. In addition, many approaches are based on benchmark datasets, limiting usefulness for real-world insurance scenarios. The use of emerging technologies, like Explainable XAI, LLMs, telematics, and MLOps, is still lacking in integration in a unified prediction framework. Finally, there is a need for further research to develop sustainable, interpretable, privacy-preserving models using diverse real-world data sets, ensuring fairness, security, and compliance with regulations.

VI. Conclusion and future work

In the insurance industry, ML has emerged as a transformative technology, providing tools for accurate risk assessment, personalized pricing, and data-driven decision-making. This survey explored the basic principles of insurance premium prediction and investigated some popular ML methods such as SVR, RF, LR, GBM, XGBoost, and LSTM. The survey also highlighted new technologies, such as telematics-based insurance, AI-powered personalized pricing, XAI, LLMs, and MLOps, which have helped to boost the accuracy of predictions, model explanation, and operational efficiency. In addition, the following challenges were addressed, including data quality, privacy, security, class imbalance, model interpretability and regulatory compliance. The reviewed papers indicate that ensemble learning methods are also superior to traditional statistical methods and that the XAI techniques further enhance the transparency and reliability of the premium prediction. Overall, ML has great promise in developing smart, effective, and customer-oriented insurance pricing strategies.

Future work should aim to implement scalable, robust, and interpretable ML models that can handle big data, heterogeneous data, and real-time insurance information. Incorporating ground-breaking technologies such as XAI, LLMs, telematics, IoT, and MLOps into single premium prediction systems can improve the precision, explainability, and efficiency of deployment. Furthermore, approaches that protect the privacy of customer information, such as secure data sharing protocols and federated learning, should be explored to achieve a balance between model accuracy and sensitive customer data protection. Further research is needed on fairness, algorithmic bias, data imbalance, regulatory compliance, and other factors to guarantee trustworthy and ethical insurance applications. Finally, the broad validation using diverse real-world insurance datasets and insurance domains will improve generalizability and applicability of next-generation ML-based insurance premium prediction systems.

References

[1] S. Mukherjee, "Performance Evaluation of Deep Learning Models for Early Detection of Infectious Diseases in Healthcare Systems," in *SoutheastCon 2026*, IEEE, Feb. 2026, pp. 1–6. doi: 10.1109/SoutheastCon63549.2026.11475977.

[2] M. Polam *et al.*, "Predictive Modeling for Property Insurance Premium Estimation Using Machine Learning Algorithms," *Int. J. Recent Innov. Acad. Res.*, vol. 9, no. 3, pp. 55–66, 2025.

[3] T. K. Hira, "Financial Sustainability and Personal Finance Education," in *Handbook of Consumer Finance Research: Second Edition*, 2016. doi: 10.1007/978-3-319-28887-1_29.

[4] C. S. Pareek, "Testing Ethical AI in Life Insurance: Ensuring Fairness, Transparency, and Accountability in Automated Decisions," *Int. J. Futur. Eng. Innov.*, vol. 2, pp. 43–50, 2025, doi: 10.54660/IJFEI.2025.2.2.43-50.

[5] V. Rengaraj, Y. Lnu, Y. Patel, and S. Vinod, "Decoding-Enhanced Bidirectional Encoder Representations with Attention-Based Pooling Mechanism for Insurance Fraud Detection," in *2026 3rd International Conference on Research Methodologies in Knowledge Management, Artificial Intelligence and Telecommunication Engineering (RMKMATE)*, IEEE, Apr. 2026, pp. 1–9. doi: 10.1109/RMKMATE69073.2026.11518574.

[6] K. K. Mohammed, "Leadership Practices of Data Engineering for AI and Machine Learning," *Int. J. Sci. Res. Eng. Trends*, vol. 12, no. 1, 2026, doi: 10.5281/zenodo.18677279.

[7] R. Sahai *et al.*, "Insurance Risk Prediction Using Machine Learning," in *Lecture Notes on Data Engineering and Communications Technologies*, 2023, pp. 419–433. doi: 10.1007/978-981-99-0741-0_30.

[8] N. A. I. and M. R. Muhammad Mehedi Masud, "A conceptual framework for purchase intention of sustainable life insurance: a comprehensive review," *Int. J. Innov. Sustain. Dev.*, vol. 12, 2020.

[9] A. Warrier, "Machine Learning-Enhanced Data Quality Validation for Electronic Health Record Integration," *Int. J. Sci. Technol.*, vol. 11, no. 1, pp. 1–8, Feb. 2020, doi: 10.71097/IJSAT.v11.i1.8387.

[10] M. Mittal, "AI in Healthcare: The Next Frontier In Medical Diagnosis," *Int. J. Inf. Technol. Manag. Inf. Syst.*, vol. 16, no. 2, pp. 1448–1455, Mar. 2025, doi: 10.34218/IJITMIS_16_02_091.

[11] P. C. Shekhar, "From Traditional to Transformational: Leveraging Digital Twins for Advanced Testing in Life Insurance," *Int. Sci. J. Eng. Manag.*, vol. 2, no. 8, pp. 1–8, AUG, 2023, doi: 10.55041/ISJEM0419.

[12] P. S, S. K. Anumula, N. Malik, G. Mohanavel, P. R. S. Rani, and D. Vijayanandh, "AI-driven Autonomous Vehicles: Navigating the Future of Transportation," in *2025 2nd International Conference on New Frontiers in Communication, Automation, Management and Security (ICCAMS)*, IEEE, Jul. 2025, pp. 1–6. doi: 10.1109/ICCAMS65118.2025.11233238.

[13] H. K. Yadav, "Design and Implement Machine Learning Based Predictive Analytics for Road Safety and Accident Prevention," in *2026 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI)*, Gwalior, India: IEEE, 2026, pp. 1–6, March. doi: 10.1109/IATMSI68868.2026.11465965.

[14] M. M. Abbas, "Risk assessment in insurance: contemporary methods and techniques applied at the Iraqi al-tadhamun insurance company," *Int. J. Financ. Manag. Econ.*, 2025, doi: 10.33545/26179210.2025.v8.i2.570.

[15] R. Lingam, "Integrating Trustworthiness Into the AI Lifecycle (Trustworthiness)," in *AI Safety and Preventing Harm in AI Systems*, IGI Global Scientific Publishing, 2026, pp. 213–248. doi: 10.4018/979-8-3373-6935-8.ch008.

[16] A. Sharma and A. Bansal, "A Review on Medical Insurance Costs

Using Advances in Machine Learning and Explainable AI Predictive Models," *Int. J. Sci. Dev. Res.*, vol. 10, no. 7, 2025.

[17] A. Ć. Hasibović, M. Hasić, and A. Tanović, "Interpretable Churn Prediction in Motor Liability Insurance: A Comparative Analysis of Ensemble Methods and SHAP Explanations," in *2026 25th International Symposium INFOTEH-JAHORINA (INFOTEH)*, 2026, pp. 1–6. doi: 10.1109/INFOTEH68759.2026.11477665.

[18] K. Mishra, K. Mishra, L. Kumar, K. Kumar, and A. Kumar, "Enhancing the Forecasting of Medical Insurance Premiums Through the Use of Multiple Methods," in *2026 International Conference on Intelligent Systems in Engineering, Secured Systems and Cybersecurity (ICISESSC)*, 2026, pp. 1022–1026. doi: 10.1109/ICISESSC68634.2026.11542626.

[19] R. Mahaveerakannan and M. Arunraj, "Enhancing Healthcare Insurance Cost Prediction with a Metaheuristic Optimization Approach: Comparative Analysis of Machine Learning Models," in *Proceedings of the 2025 3rd International Conference on Inventive Computing and Informatics, ICICI 2025*, 2025, pp. 1225–1232. doi: 10.1109/ICICI65870.2025.11069726.

[20] D. G. Jaishree, R. Kokila, and D. Hegde, "Medical Insurance Cost Forecasting Using Machine Learning Techniques," in *2025 International Conference on Intelligent Communication Networks and Computational Techniques (ICICNCT)*, 2025, pp. 1–5. doi: 10.1109/ICICNCT66124.2025.11232894.

[21] A. Hema Ambiha and V. Subithra, "Predicting Sham Claims in Vehicle Insurance using Machine Learning Techniques," in *2024 2nd International Conference on Self Sustainable Artificial Intelligence Systems (ICSSAS)*, 2024, pp. 673–681. doi: 10.1109/ICSSAS64001.2024.10760834.

[22] R. Kumar, M. Rakhra, D. Prashar, S. Upadhyay, L. Mrsic, and A. A. Khan, "A Machine Learning and Ratemaking Evaluation of Four Auto Insurance Pure Premium Modeling Algorithms," in *2024 28th International Computer Science and Engineering Conference (ICSEC)*, 2024, pp. 1–6. doi: 10.1109/ICSEC62781.2024.10770731.

[23] S. Ahmad, A. Agarwal, and H. Ansari, "Prediction of Insurance Premium using Machine Learning with an Adaptive Approach," in *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, 2023, pp. 1–5. doi: 10.1109/ICCCNT56998.2023.10307009.

[24] A. Bora, R. Sah, A. Singh, D. Sharma, and R. K. Ranjan, "Interpretation of machine learning models using XAI - A study on health insurance dataset," in *2022 10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, 2022, pp. 1–6. doi: 10.1109/ICRITO56286.2022.9964649.