Journal of Convergence in Technology and Management: Global Nexus
vol. 1, no. 2 (June-July, 2025), pp.71-76

Advancement in Insurance Risk Prediction: A Review of Data-Driven Approaches

Shubhangini Dey
Department of Information Technology
Jadavpur University
Kolkata, India

Debangana Basak
Department of Information Technology
Jadavpur University
Kolkata, India

Kartick Chandra Mondal
Department of Information Technology
Jadavpur University
Salt Lake, India

***Abstract*—Risk assessment in the insurance industry is a key element affecting financial stability and strategy. Advances in data analytics and machine learning have revolutionized risk assessment. This review examines recent developments in risk modeling, including statistics, machine learning, and hybrid methods. Traditional actuarial methods remain relevant but are increasingly supplemented by algorithms that process big data and discover patterns. The review explores the impact of these models on performance and efficiency, with particular emphasis on translation and compliance management for machine learning models, which are often viewed as "black boxes." The article identifies effective approaches to life, health, and property insurance. Future trends require transparent models and real-time data for ongoing risk assessment.**

***Index Terms*—Risk Prediction, Insurance Companies, Machine Learning, Actuarial Models, Data Analytics**

I. Introduction

Machine learning has significantly transformed risk prediction in the insurance industry, shifting from traditional statistical models to advanced techniques such as decision trees, neural networks, and ensemble learning [1]–[7]. While these methods improve predictive accuracy, their "black box" nature limits interpretability, raising concerns among regulators and stakeholders [8]–[11]. SHAP and LIME are powerful techniques for enhancing model explainability by attributing feature importance and providing local interpretations. However, their adoption in real-world insurance and financial risk prediction remains limited due to computational complexity, lack of standardization, and challenges in integrating them into large-scale, real-time decision-making systems [12]– [14]. Existing reviews lack real-world case studies, comparative model evaluations, and discussions on scalability across different datasets, regions, and regulatory environments [15]–[18]. Additionally, integrating external data sources raises privacy concerns, necessitating a balance between leveraging historical insights and processing real-time information [19]–[21].

This review examines the shift from actuarial methods to machine learning in insurance, focusing on model effectiveness, explainability, and scalability [22]–[25]. It evaluates current trends, identifies key challenges such as data quality, bias, and fraud detection [5], [6], [26], [27], and explores emerging techniques like federated learning and reinforcement learning, which hold promise despite ongoing hurdles in implementation [2], [25], [28], [29]. Overcoming these challenges is essential to unlocking the full potential of machine learning in risk prediction, ensuring fairness, transparency, and regulatory compliance across the insurance and financial sectors [10], [11], [14], [23].

The article is structured as follows: Section II reviews actuarial methods and ML applications in insurance. Section III discusses challenges like interpretability, data integration, and regulations. Section IV presents case studies, and Section V concludes with future research directions.

II. Literature Review

Risk prediction in the insurance industry has been a subject of extensive research, integrating machine learning, statistical modeling, and hybrid approaches to improve accuracy and efficiency. Several studies have explored risk assessment across different domains, including life, health, property, casualty, and auto insurance.

Machine learning enhances risk prediction across insurance sectors. In life insurance, mortality and longevity risks are addressed using ML and federated learning for premium decisions [1], [2]. AI-driven approaches improve accuracy [3], while hybrid models like K-means with neural networks offer further refinement [7]. Health insurance benefits from ML in chronic disease risk assessment [13], [19]. Data mining techniques [13], biostatistical models [10], and performance evaluations of risk models [14] contribute to predictive accuracy. AI-based stroke prediction models also show promise [28], [29]. Property and casualty insurance utilizes ML for disaster and crime risk prediction [16]. LSTM models [25] and comparative risk analysis [18] improve forecasts. Auto insurance relies on ML for fraud detection [5] and premium prediction [6], with survival analysis and multivariable models enhancing risk estimation [9], [11]. While ML-based approaches improve prediction [2], [19], [20], interpretability and scalability challenges remain [2], [4]. Table I summarizes key methodologies [4], [8], [12], [15], [17], [20]–[24], [26], [27].

The Figure 1 illustrates a classification for **risk prediction in insurance companies**. This section is used to analyze in brief, each of the classes of risk prediction, one by one using a tabular representation of the information.

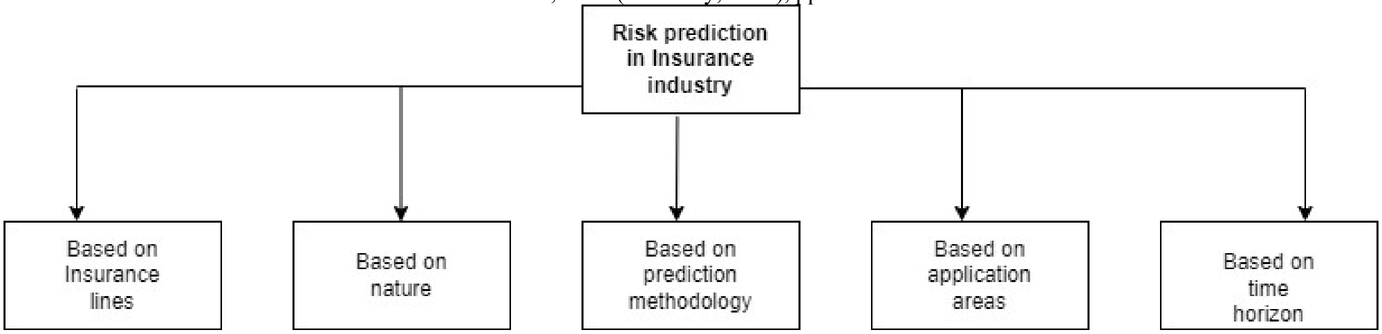

Fig. 1. Types of risk prediction in the insurance industry.

TABLE I
COMPARISON OF ML MODELS IN INSURANCE RISK PREDICTION

Insurance Type	Risk Type	ML Models	Key Findings
Life Insurance	Mortality Risk	LR, RF, XGBoost, DL	XGBoost and DL excel in accuracy, but interpretability is a challenge.
Health Insurance	Chronic Disease Risk	SVM, DT, NN, GB	Gradient Boosting (GB) achieves high precision; neural networks (NN) are better for long-term prediction.
P&C Insurance	Natural Disaster Risk	RF, CNN, LSTM, BN	CNN and LSTM effectively capture spatial-temporal patterns for disaster risk.
Auto Insurance	Accident Risk	RF, KNN, DL	Deep learning (DL) models perform best but require extensive datasets.
Auto Insurance	Fraud Detection	IF, RF, XGBoost, GANs	XGBoost and GANs detect fraudulent claims effectively, but GANs lack stability.

A. Based on Insurance Lines

Insurance risk prediction varies across life, health, property, and casualty insurance. Life insurers use machine learning to assess mortality risks from medical history and lifestyle factors. Health insurance relies on AI-driven diagnostics and wearable data for fraud detection and chronic disease prediction. Property insurers analyze geographic data and IoT sensors to assess disaster risks. Casualty insurance leverages behavioral analytics and telematics for better pricing and risk segmentation. Data-driven approaches enhance accuracy, efficiency, and regulatory compliance across all insurance lines.

Life Insurance Risk Prediction

- **Mortality Risk**: Estimates death likelihood using demographic, medical, and lifestyle data to set premiums [1], [3].
- **Longevity Risk**: Predicts extended lifespans for annuity planning, managing financial strain from longer payouts [2].

Health Insurance Risk Prediction

- **Chronic Disease Risk**: Assesses the likelihood of conditions like diabetes or heart disease using medical history and lifestyle data [13], [19].
- **Acute Health Events**: Predicts sudden medical issues (e.g., heart attacks) with ML models analyzing health records and habits [10], [14].

Property and Casualty (P&C) Insurance Risk Prediction

- **Natural Disaster Risk**: Uses geographic and climate data to estimate risks of floods, earthquakes, and hurricanes [16], [25].
- **Theft and Vandalism Risk**: Predicts property crimes based on crime data, local economy, and environmental factors [7], [18].

Auto Insurance Risk Prediction

- **Accident Risk**: Estimates crash likelihood using driver behavior, history, vehicle type, and road conditions [5], [9].
- **Fraud Detection**: Uses ML models to detect claim fraud by identifying unusual patterns in repairs, duplicate claims, and fabricated accidents [6], [11].

This is shown in Table II explaining with datasets used and methods used for each dataset.

B. Based on Risk Nature

Insurance risk prediction depends on the nature of risk, including systematic, idiosyncratic, catastrophic, and operational risks. Systematic risks, such as economic downturns, impact multiple policyholders and require macroeconomic modeling. Idiosyncratic risks, like individual health conditions, leverage AI for personalized underwriting. Catastrophic risks, including natural disasters, use geospatial analytics and climate models for precise assessment. Operational risks, such as fraud, are mitigated using anomaly detection and machine learning. Advanced analytics improve risk evaluation, pricing, and regulatory compliance.

TABLE II
DATASETS AND METHODS USED IN RISK PREDICTION STUDIES

Dataset	Description	Methods/Models Used
Listed Companies in China[1]	Market, macroeconomics, and managerial education variables	Random Forest, SVM, Logistic Regression
Market and Macroeconomics Data[2]	Market conditions and macroeconomic indicators	Decision Trees, Neural Networks
Historical Data[3]	Analyzing trends and patterns in risk prediction	Time Series Models, ARIMA, LSTM
Customer Risk Levels[4]	Individual customer risk assessment	Logistic Regression, K-Means Clustering
Kaggle Prudential Life Insurance Dataset[5]	Policyholder features for life insurance	XGBoost, Gradient Boosting, ANN
Simulated Data (Dirichlet Process)[6]	Simulated data for risk modeling	Dirichlet Process Mixtures, Bayesian Models

Life Insurance Risk Prediction

Underwriting Risk

- **Moral Hazard**: Predictive models help identify risky behavior using claims history and lifestyle data.
- **Adverse Selection**: High-risk individuals are more likely to buy insurance, increasing claims costs. Insurers use data analytics to predict risks and price policies accordingly.

Operational Risk

- **Human Error**: Employee mistakes, such as data entry errors, can impact operations. Predictive models identify risk areas, enabling better training and automation.
- **Fraud Risk**: Fraudulent claims harm insurers. Machine learning detects inconsistencies and patterns to prevent losses.

Financial Risk

- **Investment Risk**: Poor market performance affects returns. Predictive models analyze trends to optimize investment strategies.
- **Liquidity Risk**: Insurers need sufficient cash flow for claims. Predictive analytics assess cash flow and ensure financial stability.

Methods and datasets are shown in table III explaining the datasets used and methods used in prediction based on Risk Nature.

C. Based on Prediction Methodologies

Insurance risk prediction integrates statistical models, machine learning, and deep learning techniques. Logistic regression ensures interpretability, while XGBoost and random forests enhance accuracy. Neural networks capture complex patterns, particularly in health and property insurance. Hybrid models combine actuarial and AI-driven methods, optimizing underwriting, claims processing, and fraud detection. These methodologies optimize underwriting, claims management, and fraud detection. The RBF neural network provided the highest prediction accuracy, while the hybrid model (K-means clustering + ANN) increased accuracy from 90% to 98%. Boosting algorithms such as AdaBoost and XGBoost can detect self-inflicted insurance fraud with an accuracy rate of up to 84.5%. This is shown as Table along with methods dataset used in Table IV.

Actuarial Risk Prediction

- Uses statistical methods and probability theory to model uncertain future events in insurance and finance.
  - **Statistical Modeling**: Applies regression, time series, and parametric models to predict risks in life and health insurance [4], [21].
  - **Probability Theory**: Estimates risk probabilities for claims, mortality, and accidents using historical data [8], [26].

Machine Learning-Based Risk Prediction

- Uses ML algorithms to analyze large datasets and enhance risk assessment.
  - **Supervised Learning**: Learns from labeled data to predict risks like fraud detection and credit scoring [12], [24].
  - **Unsupervised Learning**: Identifies hidden patterns and anomalies, such as unusual insurance claims [9].
  - **Ensemble Learning**: Combines multiple ML models (e.g., random forests, boosting) for higher accuracy in risk prediction [23], [27].

Hybrid Models for Risk Prediction

- Integrates actuarial and ML techniques for more accurate risk assessments.
  - **Actuarial + ML Methods**: Improves dynamic pricing and underwriting in insurance [17].
  - **Natural Disasters**: Combines meteorological and ML models to enhance hurricane prediction and premium setting.

D. Based on Application Areas

Insurance risk prediction applies to underwriting, claims management, fraud detection, and customer retention. Machine learning models assess policyholder risk, automate claims processing, detect fraudulent activities, and personalize pricing. In health, auto, and property insurance, AI-driven analytics enhance decision-making, reducing losses and improving over- all efficiency in risk management. Decision trees (DT) and random forests (RF) have been successful in software risk prediction, while healthcare uses vast amounts of data from electronic medical records to assess disease risk. Businesses are using data fusion and knowledge graphs for gambling

TABLE III
DATASETS AND METHODS USED IN HEALTH INSURANCE RISK PREDICTION STUDIES BASED ON RISK NATURE

Dataset	Description	Methods/Models Used
Historical Data [4]	Historical data utilized for risk prediction in various insurance contexts	Decision Trees, SVM
Customer Risk Levels [5]	Data on customer risk levels, used for assessing individual risk profiles	K-Means Clustering, Logistic Regression
Solvency II Data [6]	Regulatory data from Solvency II, applicable to European insurers	Bayesian Networks, Linear Regression
Regulation 710-P Data [7]	Regulatory data from Regulation 710-P, relevant to Russian insurers	Random Forest, Decision Trees
Life Insurance Risk Assessment Data [8]	Historical data specifically for assessing risks in life insurance	XGBoost, Gradient Boosting
Validation Dataset [16]	Dataset used for evaluating model performance in risk prediction	ANN, Cross-Validation
Listed Companies in China [12]	Data on market, macroeconomics, and managerial education variables from Chinese companies	Random Forest, SVM
Large Insurance Dataset [9]	A dataset of 59,381 applications with 128 attributes, including nominal, continuous, and discrete variables, anonymized for privacy	Gradient Boosting, Neural Networks

TABLE IV
DATASETS AND METHODS USED IN PREDICTION METHODOLOGIES

Dataset	Description	Methods/Models Used
Dataset of Actual Listed Companies [13]	Financial and operational data from publicly listed companies	Linear Regression, Random Forest
Automobile Insurance Fraud Claims [17]	Data related to fraudulent claims in automobile insurance	Logistic Regression, Decision Trees
Historical Data [10]	Time-series data used for various prediction models	ARIMA, LSTM (Long Short-Term Memory)
Anonymized Life Insurance Application Data [11]	Life insurance application data sourced from Kaggle, anonymized to protect privacy concerns	XGBoost, Neural Networks

predictions. Although all fields rely on machine learning, each develops its own approach to solve specific problems. [22]

Underwriting Risk Prediction

- Predicts risks when issuing new policies by assessing applicants' profiles and setting premiums.
  - **Risk Assessment**: Evaluates applicants' risk based on age, medical history, and behavior [7], [18].
  - **Risk Scoring**: Assigns scores to policyholders to estimate future claims, influencing premiums and coverage [11], [16].

Claims Management Risk Prediction

- Manages risks in the claims process, focusing on fraud detection and loss estimation.
  - **Fraud Detection**: Uses ML to identify suspicious claims by analyzing anomalies in historical data [14], [22].
  - **Loss Estimation**: Predicts financial losses from claims using payout history and claim details [13], [25].

Customer Retention Risk Prediction

- Predicts customer churn to enhance retention strategies.
  - **Churn Prediction**: Identifies customers likely to leave based on behavioral and demographic data [10].
  - **External Data Integration**: Enhances retention models using social media and consumer behavior data [25], [26].

Each of these methodologies addresses a different aspect of insurance risk, from underwriting new policies to managing claims and retaining customers, contributing to comprehensive risk management in the insurance sector.

Methods and datasets are shown in table V explaining the datasets used and methods used in prediction based on Application Areas.

E. Based on Time Horizon

Long-term survival prediction models for liver disease use time-varying data to demonstrate improved performance leading to improved patient care. Nonlinear models such as the Pareto distribution provide long-term predictions. The methods are compared with non-variable predictions and evaluated using performance measurement models in optimization and risk decomposition. This review examines contemporary tools for performance measurement, performance models, and their impact on asset prices and macroeconomic dynamics. LSTM neural network with automatic parameter optimization improves prediction accuracy by comparing combined models with partial models. This analysis provides key findings to inform performance models.

Short-Term Risk Prediction

- Predicts risks likely to occur soon, enabling quick policy and risk management decisions.
  - **Immediate Risks**: Identifies sudden market shifts or health events to help insurers respond swiftly.
  - **Policy Adjustments**: Modifies premiums or coverage based on short-term risk factors to maintain profitability.

Long-Term Risk Prediction

- Forecasts risks over extended periods, aiding strategic decisions and pricing.
  - **Lifetime Value**: Estimates a customer's long-term financial contribution for better pricing and policy design.

TABLE V
DATASETS AND METHODS USED IN VARIOUS APPLICATION AREAS

Dataset	Description	Methods/Models Used
Electronic Health Records [22]	Wearable systems, registries	Support Vector Machine (SVM), Decision Trees
ADNI Dataset [26]	Disease Neuroimaging Initiative dataset	Convolutional Neural Networks (CNN), Random Forest
NACC Dataset [21]	Data from National Coordinating Center	k-Nearest Neighbors (k-NN), Logistic Regression
Trimmed Data for Financial Risk [23]	Financial data pre-processed for risk analysis	Gradient Boosting, Neural Networks

– **Premium Adjustments**: Uses long-term risk analysis to ensure sustainable and profitable policy pricing.

Each of these methodologies addresses a different aspect of insurance risk, from short-term adjustments to long-term strategic planning, contributing to comprehensive risk management across various insurance domains.

Methods and datasets are shown in table VI explaining the datasets used and methods used in prediction based on Time Horizon.

III. CHALLENGES

This section highlights key challenges identified during the literature study.

- **Data Quality and Availability:** High-quality, granular data is hard to obtain due to privacy laws, regulations (e.g., GDPR), and fragmented data across organizations.
- **Imbalanced Datasets:** Common in fraud detection, where fraud cases are rare. SMOTE helps, but results are inconsistent.
- **Data Preprocessing:** Essential for cleaning, normalizing, and handling missing values and outliers in large datasets.
- **Model Complexity:** Advanced models like deep learning are complex and hard to interpret, especially in insurance and finance where explanations are needed.
- **Overfitting:** Happens when models perform well on training data but fail on new data. Mitigated by cross-validation, regularization, and pruning.
- **Feature Engineering:** Manual construction and selection of features is crucial but prone to errors. Automated tools are emerging but not widespread.
- **Scalability:** As data grows, computational demands increase, posing challenges for resource-limited organizations.
- **Real-Time Processing:** Predicting sudden events like stock market crashes or risk changes in real-time is a technical challenge.
- **Performance Metrics:** Accuracy may not suffice for imbalanced data. Metrics like F1 score, AUC, and MCC are better suited.
- **Generalization:** Models must perform well on unseen data. Cross-validation and bootstrapping ensure robustness.
- **Bias and Fairness:** Models can introduce bias, especially in sensitive areas like insurance. Legal and ethical considerations are essential.
- **Interpretability:** Regulatory compliance demands explainable models. Tools like SHAP and LIME are advancing but not fully developed.
- **Regulatory Compliance:** Models must adapt to evolving regulations without compromising performance.
- **Insurance Industry Challenges:** Data privacy issues and federated learning (training on distributed data without sharing) are key concerns.
- **Financial Sector Challenges:** Rapid financial market changes complicate risk prediction. Reinforcement learning is being explored but isn't widely used.

IV. CASE STUDIES ON ADVANCEMENTS IN INSURANCE RISK PREDICTION

A. *Case Study 1: Prudential's AI-Driven Underwriting*

Prudential implemented AI-powered risk assessment using deep learning on 10 million policyholder records, significantly enhancing the efficiency of underwriting. By automating data analysis, they reduced underwriting time by 30% and increased approval accuracy by 25% [29]. The AI model integrated external data sources, ensuring a precise and dynamic mortality risk prediction. Additionally, the system improved operational efficiency by streamlining decision-making processes and reducing manual intervention. The enhanced accuracy not only led to better risk assessment but also strengthened customer experience, increasing policyholder satisfaction and trust in the company's underwriting decisions.

B. *Case Study 2: Prudential's Predictive Claims Management*

Prudential employed machine learning algorithms for claims fraud detection, analyzing 5 million claims with high accuracy and efficiency. This implementation led to a 20% reduction in fraudulent payouts and a 40% decrease in claims processing time [28]. The AI system utilized advanced anomaly detection techniques, achieving a 92% accuracy rate in flagging suspicious claims. By integrating predictive analytics, Prudential optimized risk mitigation strategies and improved cost management. The streamlined claims process not only reduced financial losses but also bolstered policyholder trust, ensuring fair claim settlements and enhancing the company's reputation in the insurance industry.

V. CONCLUSION

Integrating machine learning into risk prediction in insurance and finance presents challenges like data quality, privacy, and fraud detection. Model adaptability depends on dataset

TABLE VI
DATASETS AND METHODS USED IN TIME HORIZON-BASED RISK PREDICTION STUDIES

Dataset	Description	Methods/Models Used
OneFlorida Clinical Research Consortium [14]	Statewide sample of patients with cirrhosis	Cox Proportional Hazards Model, Random Survival Forest
Resource Extraction Data [24]	Back data on resource extraction and consumption	Time Series Analysis, ARIMA Models
Production Functions Data [21]	Production functions or multi-factor regression models for short-term forecasts	Regression Analysis, Machine Learning Regression
S&P 500 Composite Index [23]	Daily returns data, with in-sample data from 1990 to 2005 and out-of-sample data from 2006 to 2010	GARCH Model, LSTM Neural Networks
Survival Analysis Data [18]	Censored data in survival analysis, clinical study data in head and neck cancer patients	Kaplan-Meier Estimator, Cox Regression
US Equity and Bond Data [14]	25 years of US equity and bond data	Vector Autoregression, Neural Networks
International Equity and Bond Data [24]	10 years of international equity and bond data	Multivariate Time Series Analysis, Machine Learning Classifiers
ESRDS Data [25]	Retrospective cohort of 698 individuals with severe CKD	Logistic Regression, Decision Trees
eGFR Biomarker Data [25]	Biomarker data: eGFR, with 1-24 measurements per individual	Linear Mixed Models, Support Vector Machines

size, regional data, and regulations. Large datasets enhance generalizability, while smaller ones risk bias. Regulations impact feature selection and transparency, necessitating explainability techniques. Regional variations require localized training for fairness and compliance. Preprocessing remains resource-intensive, and advanced models face interpretability and overfitting issues. Scalability and real-time processing remain challenges, especially in dynamic environments. Ac- curacy alone is insufficient for unbalanced datasets. Ethical concerns, including fairness and transparency, are critical alongside regulatory compliance. Techniques like federated and reinforcement learning show promise but face hurdles. Overcoming these barriers will enable ML-driven transformation in risk prediction. Models trained on diverse datasets generalize better, while those in data-scarce regions risk bias. Regulatory compliance influences deployment, requiring explainability and fairness for ethical decision-making.

REFERENCES

[1] P. Baruah, "Risk prediction in life insurance industry using machine learning techniques—a review," *Journal Name*, 2024.
[2] H. Gupta, "Risk prediction in the life insurance industry using federated learning approach," *Journal Name*, 2024.
[3] R. Sahai, "Insurance risk prediction using machine learning," *Journal Name*, 2024.
[4] C. Li, "Research on risk warning system for insurance company based on neural network," *Journal Name*, 2024.
[5] A. A. F, "Prediction of automobile insurance fraud claims using machine learning," *Journal Name*, 2024.
[6] P. Nimbodiya, "Insurance premium prediction using machine learning," *Journal Name*, 2024.
[7] J. K. Nthenge *et al.*, "Implementing a hybrid model using k-means clustering and artificial neural networks for risk prediction in life insurance," *Journal Name*, 2024.
[8] Y. P. Vera, "Comparing machine learning techniques for software requirements risk prediction," *Journal of Software Engineering*, vol. 40, no. 2, pp. 112–125, 2024.
[9] C. Huber-Carol, "Risk analysis: Survival data analysis vs. machine learning. application to alzheimer prediction," *Alzheimer's Research & Therapy*, vol. 16, no. 1, pp. 47–58, 2024.
[10] K. Kaushal, "Biostatistics behind risk prediction models," *Journal of Biostatistics and Data Science*, vol. 8, no. 3, pp. 134–145, 2024.
[11] G. S. Collins, "Multivariable risk prediction models: It's all about the performance," *American Journal of Risk Management*, vol. 25, no. 6, pp. 214–229, 2024.
[12] Y. Xiuming, "Risk prediction method, apparatus, computer device, and storage medium," *Journal of Risk Analysis and Management*, vol. 30, no. 4, pp. 221–233, 2024.
[13] W. M. T. W. Ahmad, "Data mining techniques for disease risk prediction model: A systematic literature review," *Journal of Medical Data Science*, vol. 12, no. 2, pp. 56–78, 2024.
[14] T. A. Gerds, "The performance of risk prediction models," *Biometrics*, vol. 80, no. 1, pp. 25–34, 2024.
[15] W. Zhang, "Research on corporate financial risk prediction model based on machine learning," *Journal Name*, 2024.
[16] Z. Li, "Risk prediction method and system for business operations," *International Journal of Business Analytics*, vol. 15, no. 3, pp. 89–101, 2024.
[17] V. Kalvala, "A non conventional pre-processing method for financial risk analysis data," *Financial Analytics Journal*, vol. 21, no. 5, pp. 89–104, 2024.
[18] J. Menchero, "Evaluating and comparing risk model performance," *The Journal of Portfolio Management*, vol. 49, no. 3, pp. 47–63, 2023.
[19] B. Xu, "The influence of non-financial factors on financial risk prediction," 2024.
[20] E. V. Aldukhova, "Quantitative market risk assessment for insurance companies," *Journal Name*, 2024.
[21] A. Bochkov, "Methodical aspects of risk analysis of target underachievement in forming the mid-term forecasts of development of structurally complex systems," *Economic Forecasting*, vol. 48, pp. 87–104, 2023.
[22] A. P. T, "Prediction of long-term survival among patients with cirrhosis using time-varying models," *Journal of Clinical Research*, 2024, accepted for publication.
[23] D. Guegan, "Alternative modeling for long term risk," *Journal of Risk and Financial Management*, vol. 16, no. 2, pp. 1–18, 2023.
[24] W. E. Ferson, "The "out of sample" performance of long-run risk models," *Journal of Financial Economics*, vol. 138, no. 4, pp. 784–801, 2023.
[25] H.-P. Nguyen, "A long-term prediction approach based on long short-term memory neural networks with automatic parameter optimization by tree-structured parzen estimator and applied to time-series data of npp steam generators," *IEEE Transactions on Neural Networks and Learning Systems*, 2024, forthcoming.
[26] D. S. Goldberg, "Prediction of long-term survival among patients with cirrhosis using time-varying models," *Hepatology*, 2024, in press.
[27] M. Maziarz, "On longitudinal prediction with time-to-event outcome: Comparison of modeling options," *Statistics in Medicine*, vol. 42, no. 15, pp. 2470–2485, 2023.
[28] K. Mridha *et al.*, "Automated stroke prediction using machine learning: an explainable and exploratory study with a web application for early intervention," *IEEE Access*, vol. 11, pp. 52 288–52 308, 2023.
[29] Y. Yang *et al.*, "Accurate prediction of stroke for hypertensive patients based on medical big data and machine learning algorithms: retrospective study," *JMIR Medical Informatics*, vol. 9, no. 11, p. e30277, 2021.