

Artificial Intelligence-Based Automated Actuarial Pricing and Underwriting Model for the General Insurance Sector

Brighton Mahohoho¹, Charles Chimedza², Florance Matarise¹, Sheunesu Munyira¹

¹Department of Mathematics & Computational Sciences, University of Zimbabwe, Harare, Zimbabwe

²School of Statistics & Actuarial Science, University of Witwatersrand, Johannesburg, South Africa

Email: brightonmahohoho.07@gmail.com, Charles.Chimedza@wits.ac.za, matarise7@gmail.com, smunyira@science.uz.ac.zw

How to cite this paper: Mahohoho, B., Chimedza, C., Matarise, F. and Munyira, S. (2024) Artificial Intelligence-Based Automated Actuarial Pricing and Underwriting Model for the General Insurance Sector. *Open Journal of Statistics*, 14, 294-340. <https://doi.org/10.4236/ojs.2024.143014>

Received: May 22, 2024

Accepted: June 24, 2024

Published: June 27, 2024

Copyright © 2024 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

The Automated Actuarial Pricing and Underwriting Model has been enhanced and expanded through the implementation of Artificial Intelligence to automate three distinct actuarial functions: loss reserving, pricing, and underwriting. This model utilizes data analytics based on Artificial Intelligence to merge microfinance and car insurance services. Introducing and applying a no-claims bonus rate system, comprising base rates, variable rates, and final rates, to three key policyholder categories significantly reduces the occurrence and impact of claims while encouraging increased premium payments. We have enhanced frequency-severity models with eight machine learning algorithms and adjusted the Automated Actuarial Pricing and Underwriting Model for inflation, resulting in outstanding performance. Among the machine learning models utilized, the Random Forest (RANGER) achieved the highest Total Aggregate Comprehensive Automated Actuarial Loss Reserve Risk Pricing Balance (ACAALRRPB), establishing itself as the preferred model for developing Automated Actuarial Underwriting models tailored to specific policyholder categories.

Keywords

Artificial Intelligence, Automated Actuarial Loss Reserves, Automated Actuarial Risk Pricing, Automated Actuarial Underwriting

1. Introduction

The non-life insurance pricing consists of establishing a premium or a tariff paid

by the insured to the insurance company in exchange for the risk transfer. A usual way to obtain the insurance premium is to combine the conditional expectation of the claim frequency with the expected claim amount [1]. The role of the actuary in the pricing of general insurance business has received limited study in the past. This may have much to do with the fact that pricing has not been seen as the primary area in which actuaries operate, as demonstrated by the relatively limited attention devoted to this key area within the current actuarial training program [2]. Nevertheless, the past 20 years have seen a substantial development in the engagement of actuaries in pricing especially in markets where there are no regulatory requirements for their involvement. The growing awareness of risk as a factor in all industries, not solely insurance, has given greater impetus to these developments, yet little is understood about the actuarial role in pricing and to date little initial training time is devoted to this area of activity [2]. As a matter of consequence, our model given full implementation is most likely capable of attracting new customers, aging the insurer's underwriting capability, reducing expenses, improving the customer experience, improving risk selection and adding consistency. Insurance companies, like all businesses, operate in a social context. Within this context, however, insurance has a number of special features that distinguish it from other consumer services. Some of these features may lead to a perceived need for special regulation of insurance [3].

1.1. Actuarial Pricing

According to [4] the necessity of pricing for non-life insurance comes precisely in an attempt to combat the anti-selection phenomenon by dividing the insurance portfolio into sub-portfolios based on certain influence factors. As a result, every class will contain policyholders with similar risk profile that will pay the same reasonable premium. Nevertheless, a usual method for calculating the premium is to combine the conditional expectation of the claim frequency with the expected cost of claims, whilst considering the observable risk characteristics. Furthermore, the process of evaluating risks in order to determine the insurance premium is performed by the actuaries, which over time proposed and applied different statistical models. On the same note, [5] stated that the automobile insurance policy pricing relies on rating factors to assess the exposure to loss associated with an insurance policy. On the same note, these factors are used to separate the lower risk drivers and vehicles from the higher ones, and largely form the basis of what an individual is charged on an auto insurance policy.

The use of statistical learning models has been a common practice in actuarial science since the 1980s. The field quickly adopted linear models and generalized linear models for rate making and reserving. The statistics and computer science fields continued to develop more flexible models, outperforming linear models in several research fields. To our knowledge and given the sparse literature on the subject, the actuarial science community largely ignored these until the last few years.

1.2. Actuarial Underwriting for the General Insurance Sector

According to [6] underwriting in general insurance is a lengthy and detailed process that should be well-planned. In addition to that, various underwriting factors should be taken into consideration by the insurer before signing non-life insurance contracts. On the same note, the information regarding factors affecting insurance risks should be available to the underwriter however, sometimes the information provided for rating purposes is incomplete. Moreover, it may be difficult to obtain information as the insured may not always be willing to relinquish all the required information. As an example, people may be reluctant to provide the correct information if they know it may cause refusal of coverage. Eventually, it is important for underwriters to place great attention on the underwriting factors as it could greatly affect their decision about whether to accept a risk.

1.3. General Machine Learning Algorithm for Frequency-Severity Approach

The frequency-severity approach is widely used in actuarial science and insurance to model the number of claims (frequency) and the cost of claims (severity) separately. In this paper, we present a general machine learning algorithm for implementing such methods [7] and [8].

Algorithm

Algorithm 1 General Machine Learning Algorithm for Frequency-Severity Approach

- 1: **Input:** Dataset $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^N$, where x_i are features and y_i are targets.
 - 2: **Output:** Trained models for frequency and severity.
 - 3: Split the dataset \mathcal{D} into training set $\mathcal{D}_{\text{train}}$ and test set $\mathcal{D}_{\text{test}}$.
 - 4: **Frequency Modeling:**
 - 5: Define the frequency target variable y_{freq} as the number of claims.
 - 6: Train a model f_{freq} using $\mathcal{D}_{\text{train}}$ to predict y_{freq} .
 - 7: **Severity Modeling:**
 - 8: Define the severity target variable y_{sev} as the cost of claims given that a claim has occurred.
 - 9: Filter $\mathcal{D}_{\text{train}}$ to obtain only records with non-zero claims.
 - 10: Train a model f_{sev} using the filtered $\mathcal{D}_{\text{train}}$ to predict y_{sev} .
 - 11: **Prediction:**
 - 12: For each instance $x \in \mathcal{D}_{\text{test}}$:
 - 13: Predict the frequency $\hat{y}_{\text{freq}} = f_{\text{freq}}(x)$.
 - 14: Predict the severity $\hat{y}_{\text{sev}} = f_{\text{sev}}(x)$.
 - 15: Compute the total predicted cost $\hat{y} = \hat{y}_{\text{freq}} \times \hat{y}_{\text{sev}}$.
 - 16: **Evaluation:**
 - 17: Evaluate the performance of the frequency and severity models separately and jointly using appropriate metrics.
-

1.4. Mathematical Formulation

Frequency Model

Let X be the feature matrix and Y_{freq} be the frequency target variable.

$$\hat{Y}_{\text{freq}} = f_{\text{freq}}(X) \quad (1)$$

1.5. Severity Model

Let Y_{sev} be the severity target variable.

$$\hat{Y}_{\text{sev}} = f_{\text{sev}}(X) \quad (2)$$

Total Cost Prediction

The total cost prediction \hat{Y} is given by:

$$\hat{Y} = \hat{Y}_{\text{freq}} \times \hat{Y}_{\text{sev}} \quad (3)$$

1.6. Propositions and Theorems

Proposition 1 Given that f_{freq} and f_{sev} are unbiased estimators, the product $\hat{Y} = \hat{Y}_{\text{freq}} \times \hat{Y}_{\text{sev}}$ is an unbiased estimator of the total cost.

Proof. Let $\mathbb{E}[\cdot]$ denote the expectation operator. Since f_{freq} and f_{sev} are unbiased, we have:

$$\mathbb{E}[\hat{Y}_{\text{freq}}] = Y_{\text{freq}} \quad (4)$$

$$\mathbb{E}[\hat{Y}_{\text{sev}}] = Y_{\text{sev}} \quad (5)$$

Thus,

$$\mathbb{E}[\hat{Y}] = \mathbb{E}[\hat{Y}_{\text{freq}} \times \hat{Y}_{\text{sev}}] = \mathbb{E}[\hat{Y}_{\text{freq}}] \times \mathbb{E}[\hat{Y}_{\text{sev}}] = Y_{\text{freq}} \times Y_{\text{sev}} \quad (6)$$

Therefore, \hat{Y} is an unbiased estimator of Y . \square

The variance of the total cost prediction \hat{Y} is given by:

$$\text{Var}(\hat{Y}) = \text{Var}(\hat{Y}_{\text{freq}}) \cdot \mathbb{E}[\hat{Y}_{\text{sev}}]^2 + \mathbb{E}[\hat{Y}_{\text{freq}}]^2 \cdot \text{Var}(\hat{Y}_{\text{sev}}) + \text{Var}(\hat{Y}_{\text{freq}}) \cdot \text{Var}(\hat{Y}_{\text{sev}}) \quad (7)$$

Proof. Using the properties of variance and covariance, we have:

$$\text{Var}(\hat{Y}) = \text{Var}(\hat{Y}_{\text{freq}} \times \hat{Y}_{\text{sev}}) \quad (8)$$

$$\text{Var}(\hat{Y}) = \text{Var}(\hat{Y}_{\text{freq}}) \cdot \mathbb{E}[\hat{Y}_{\text{sev}}]^2 + \mathbb{E}[\hat{Y}_{\text{freq}}]^2 \cdot \text{Var}(\hat{Y}_{\text{sev}}) + \text{Var}(\hat{Y}_{\text{freq}}) \cdot \text{Var}(\hat{Y}_{\text{sev}}) \quad (9)$$

\square

Theorem 2 If the frequency and severity models are independently trained, the mean squared error (MSE) of the total cost prediction \hat{Y} is minimized when both models individually minimize their respective MSEs.

Proof. The MSE of \hat{Y} is given by:

$$\text{MSE}(\hat{Y}) = \mathbb{E}[(Y - \hat{Y})^2] \quad (10)$$

Expanding this, we get:

$$\text{MSE}(\hat{Y}) = \mathbb{E}[(Y_{\text{freq}} \times Y_{\text{sev}} - \hat{Y}_{\text{freq}} \times \hat{Y}_{\text{sev}})^2] \quad (11)$$

Assuming independence of the models, this simplifies to:

$$\text{MSE}(\hat{Y}) = \mathbb{E}[(Y_{\text{freq}} - \hat{Y}_{\text{freq}})^2] \cdot \mathbb{E}[(Y_{\text{sev}} - \hat{Y}_{\text{sev}})^2] \quad (12)$$

Thus, minimizing $\text{MSE}(\hat{Y})$ requires minimizing $\text{MSE}(\hat{Y}_{\text{freq}})$ and $\text{MSE}(\hat{Y}_{\text{sev}})$ individually. \square

This then presents a structured approach to implementing machine learning methods using the frequency-severity approach, including detailed algorithms, mathematical formulations, and theoretical foundations.

1.7. Novelty, Originality and Significance of the Study

Our model is an augmentation of microfinance services and car insurance services on a single platform from a comprehensive actuarial perspective whilst using Artificial Intelligence (AI). In addition to that our proposed model offers a real time insurance solution by also automation of three unique actuarial functions which are respectively Actuarial Loss Reserving (ALR), Actuarial Risk Pricing (ARP) and Actuarial Underwriting (AU) autonomously. Our models help insurance companies and other financial houses of interest with one stop real time insurance solution using the diverse policyholder risk characteristics. When given full implementation our model ensures that there are enough funds set aside for catastrophic events, bringing the reinsurance cost down and also reducing both the reporting and settlement delays to minimum whilst retaining sustainable customer satisfaction and hence brings continued business expansion to the delight of the general insurance business. The introduction to the bonus rating system across the policyholders as proposed in our study, also ensures that the policyholders get maximum stake according to their respective category variable and fixed bonus rates in the event of making a claim. This helps to reduce the prevalence, incidence and severity of making claims, hence giving benefit to the insurer or finance house of interest to save funds for paying any other outstanding claims, related expenses and induce more profitability to the business. Moreover proposed model also promotes policyholders to make savings and invest back in the economy of their respective countries which one way or the other improves both their standard of living and also improves the Gross Domestic Product of a country.

1.8. Contribution to Body of Knowledge

This paper makes several significant contributions to the field of actuarial science and insurance, particularly in the areas of pricing, underwriting, and loss reserving, through the integration of artificial intelligence and machine learning techniques.

The study introduces new terminologies to better categorize and manage premium payments and reserves: IBNYPP (Incurred But Not Yet Paid Premium): This concept refines the understanding of premium flows by accounting for premiums incurred but not yet paid, providing a more nuanced view of an insurer's premium income. PBNYSPP (Paid But Not Yet Settled Premium Payment): This term highlights premiums that are partially paid but not yet fully settled, which is critical for accurate financial reporting and reserve calculations. REOPP (Reopened Premium Payment) and REINSPP (Reinsured Premium Payment): These categories address premiums that are either reopened or require reinsurance, ensuring a comprehensive approach to premium management. These additions enhance the granularity and precision of actuarial models, enabling insurers to better manage their financial positions and risk exposures.

The study expands on traditional actuarial loss reserves by introducing:

REOPENED Reserve: To cover claims that were previously closed but have been reopened, addressing both known and unknown factors that could impact an insurer's liabilities. REINSURANCE Reserve: To manage catastrophic loss reserving, providing a buffer against extreme events that could significantly affect an insurer's solvency. These additions ensure that the actuarial loss reserving process is more robust and capable of accommodating various scenarios, thereby enhancing the insurer's risk management framework. The proposed Automated Actuarial Loss Reserving and Pricing (AALRP) model integrates these new terminologies and reserves into a cohesive framework. The model balances multiple types of reserves and premiums, providing a holistic view of the insurer's financial health: IBNYR + IBNYPP, RBNYS + PBNYSPP, REOPENED + REOPP and REINSURANCE + REINSPP. By combining these elements, the AALRP model offers a comprehensive approach to financial management in the insurance sector, accommodating both microfinance and car insurance services.

The development of the Automated Actuarial Underwriting (AAU) model, which proceeds in five stages, represents a significant advancement. This model evaluates the feasibility of underwriting over different periods, considering various reserves and premiums at each stage:

- Stage 1: Initial assessment using IBNYR and IBNYPP.
- Stage 2: Inclusion of RBNYS and PBNYSPP.
- Stage 3: Addition of REOPENED and REOPP.
- Stage 4: Consideration of REINSURANCE and REINSPP.
- Stage 5: Final evaluation with all previous elements plus total retained income.

This multi-stage approach ensures that underwriting decisions are well-founded and resilient, providing a structured method to assess long-term viability and profitability.

By utilizing eight machine learning algorithms, including Random Forest (RANGER), Generalized Linear Models (GLM), and Xtreme Gradient Boosting (XGB), this study demonstrates the efficacy of AI in improving actuarial practices: *Random Forest (RANGER)*: Identified as the most effective model for developing the AALRP model balances, highlighting its ability to handle complex datasets and improve predictive accuracy [9]. This integration of machine learning represents a significant shift from traditional methods, offering greater accuracy, efficiency, and adaptability in actuarial science.

Overall, this paper not only addresses the limitations of traditional actuarial techniques but also pioneers new methodologies and terminologies that significantly enhance the precision, robustness, and adaptability of actuarial models in the insurance industry. By doing so, it contributes to a more resilient and dynamic approach to insurance management in the face of evolving risks and data landscapes.

2. Review of Methods

Traditional actuarial techniques have long been employed in the insurance in-

dustry for pricing, underwriting, and loss reserving. These methods typically rely on historical data and well-established statistical models. However, they come with several limitations: Traditional actuarial models often depend heavily on assumptions about the distribution of data, which may not hold true in real-world scenarios [10]. These models are usually static, meaning they do not adapt well to rapidly changing environments or emerging risks [11]. Traditional methods struggle with the vast amounts of data generated in modern insurance practices and are not well-equipped to handle unstructured data [12]. The actuarial process often involves significant manual processing, making it time-consuming and prone to human error [13].

Recent advancements in machine learning (ML) have introduced more dynamic and adaptable methods for insurance practices, addressing some of the limitations of traditional techniques. These approaches offer several advantages: ML algorithms can process and learn from large datasets, providing more accurate and granular insights [14]. These models can adapt to new data and changing conditions, making them more resilient to evolving risk landscapes [15]. ML models automate much of the data processing and analysis, reducing the potential for human error and increasing efficiency [16].

GLMs have been widely used for predicting insurance losses. They offer flexibility and robustness in handling different types of data [17]. GBMs have been employed to enhance prediction accuracy by combining multiple weak prediction models [18]. [19] explored the use of neural networks for claims reserving, demonstrating their ability to capture complex non-linear relationships in the data. SVMs have been used for classification tasks in insurance pricing, offering high accuracy in identifying risk categories [20]. Random Forests have been applied to pricing models to improve prediction accuracy by aggregating multiple decision trees [9].

Logistic regression models have been used for underwriting decisions, particularly in assessing the probability of claim events [21]. XGBoost, a scalable tree boosting system, has been shown to enhance underwriting models by efficiently handling large datasets and improving prediction accuracy [22].

[23] came with his study on the Risk Premium Prediction of Car Damage Insurance using Artificial Neural Networks and Generalized Linear Models. The two methods were used in insurance pricing prediction and upon his study the Artificial Neural Networks proved to be more precise and accurate than the Generalized linear models.

[24] conducted the research on Data analytics for insurance loss modeling, telematics pricing and claims reserving. In his thesis credible innovations to the insurance loss modeling, pricing and claims reserving for the general automobile general insurance data were put in place however lack of the use and application of Artificial intelligence in his research resulted in his study entirely based on traditional statistical methods particularly the parametric methods.

[25] carried out a research on the Non-life insurance rate making techniques.

Their study was based on the customary actuarial distinction between the two main pricing techniques, namely the prior and the posterior rate making techniques. [26] carried out a study on Machine Learning and Traditional Methods Synergy in Non-Life reserving. This group of researchers looked at both traditional and machine based methods such as ANNs, Trees and Boosted Tweedie compound poisson model. In addition to that their research was entirely on general reserving and there was nothing discussed with regards to insurance pricing and underwriting as well.

[27] carried out a study on the Insurance Claim Analysis Using Machine Learning Algorithms. In his study, the author appreciated the role and application of deep learning methods towards claim modeling.

3. Methodology

This section describes the methodology for the derivation of the Automated Actuarial Pricing and Underwriting Model using eight machine learning methods namely, the Generalized Linear Models (GLMs), Generalized Additive Models (GAMs), Decision Trees (CART), Random Forests (RFs), Extreme Gradient Boosting Machines (XGBM), Least Angle Regression (LAR), Support Vector Machines (SVMs) and Artificial Neural Networks (ANNs).

The R packages and related hyper parameters used are presented on Appendix Table A1.

3.1. Methodology for the Development of the Artificial Intelligence Based Automated Actuarial Risk Pricing & Underwriting Models

The following descriptions navigate the development of the intended model in the study.

3.1.1. Automated Actuarial Loss Reserving Model

We set up the mentioned eight deep learning methods and within each method we first fitted three types of regression models respectively the Automated Actuarial Loss Reserving Frequency model with dependent variable: Comprehensive Number of Claims (Number of car insurance claims + Number of Requests) being the first, the second being Automated Actuarial Loss Reserving Severity Model with dependent variable: Comprehensive Claim Amount (Claims Incurred + Amount Requested) and lastly Automated Actuarial Loss Reserving Inflation model (Dependent variable: Inflation Index derived from the Consumer Price Index). After fitting these regressions simultaneously, we computed the predictions for each regression type and automated them to give the Inflation Adjusted Automated Actuarial Loss Reserves. From there we calculated the total Inflation Adjusted Automated Actuarial Risk Reserves for each machine learning method.

From there we used the predictions made on the test data set to derive another data set comprising of the Comprehensive Claim Amount from the original test

data, Inflation Adjusted AALR which now referred to as Automated Actuarial Loss Reserve Margin (AALRM). With these two variables, we calculated the Upper Automated Actuarial Loss Reserve Margin (UAALRM) by taking the sum of the two variables above and on the same note we obtained the Lower Automated Actuarial Loss Reserve Margin (LAALRM) by subtracting the Automated Actuarial Loss Reserve Margin (AALRM) from Comprehensive claim amount. Finally we obtained the Robust Automated Actuarial Loss Reserve Margin (RAALRM) by taking the average of Upper Automated Actuarial Loss Reserve Margin (UAALRM) and Lower Automated Actuarial Risk Margin (UAALRM). Using these three new variables we created a new reserve data set and partitioned into new train data set (80%) and new test data set (20%). Just after that, we fitted the final regression model for each machine learning model with Robust Automated Actuarial Loss Reserve Margin (UAALRM) as the dependant variable with other two variables which are the Upper Automated Actuarial Loss Reserve Margin (UAALRM) and Lower Automated Actuarial Loss Reserve Margin (LAALRM) as the independent variables. As a result, we derived this as the Final Automated Actuarial Loss Reserving Model (FAALRM) which we used to both estimate and make predictions on Robust Automated Actuarial Loss Reserve Margin (UAALRM). This can now be referred to as Automated Actuarial Loss Reserves (AALR).

3.1.2. Automated Actuarial Risk Pricing Model

We simultaneously fitted three regression models just as in the first stage respectively with the first model being the Automated Actuarial Pricing Frequency model with dependent variable: Comprehensive Number of Payments (Number of Investments + Number of Premium Payments), then the second regression model is the Automated Actuarial Pricing Severity Model with dependent variable: Comprehensive Payment Amount (Claims Paid + Microfinance Amount Paid) and the third regression model is the Automated Actuarial Pricing Inflation model with the dependent variable: Inflation Index derived as well from the Consumer Price Index (CPI). After fitting these regressions simultaneously, we computed the predictions for each regression type and automated them to give the Inflation Adjusted Automated Actuarial Risk Premiums. Afterwards we calculated the total Inflation Adjusted Automated Actuarial Risk Premiums for each machine learning method.

From there we used the predictions made on the test data set to derive another data set comprising of the Comprehensive Payment Amount, Inflation Adjusted Automated Actuarial Risk Premiums which we can now refer to as Automated Actuarial Risk Premium Margin (AARPM). With these two variables we calculated the Upper Automated Actuarial Risk Premium Margin (UAARPM) by taking the sum of the two variables, on the same note we obtained the Lower Automated Actuarial Risk Premium Margin (LAARPM) by subtracting the Automated Actuarial Risk Premium Margin (AARPM) from the Comprehensive Payment amount and finally we obtained the Robust Automated Actuarial Risk

Premium Margin (RAARPM) by taking the average of Upper Automated Actuarial Risk Premium Margin (UAARPM) and Lower Automated Actuarial Risk Premium Margin (LAARPM). Using these three main variables we created a new premium data set and partitioned it into new train data set (80%) and new test data set (20%).

Afterwards, we fitted the final regression model for each machine learning model with Robust Automated Actuarial Risk Premium Margin (RAARPM) as the dependent variable whilst the two other variables respectively the Upper Automated Actuarial Risk Premium Margin (UAARPM) and Lower Automated Actuarial Risk Premium Margin (LAARPM) being the independent variables. As a result we hence derived this as the Final Automated Actuarial Risk Premium Model (FAARPM). We then proceed to use this model to both estimate and make predictions on Robust Automated Actuarial Risk Premium Margin (UAARPM) which we can now refer to as Automated Actuarial Risk Premium (AARP).

3.1.3. Terminology and Assumptions for Automated Actuarial Loss Reserving Risk Pricing Models

We created and proposed new terminology concerning actuarial pricing respectively:

- IBNYPP (Incurred But Not Yet Paid Premium)—This refers to the premium that has been incurred by the policyholder but has not yet been paid to the insurer provided that the premiums are not fixed for both car insurance services and micro finance services since they depend on their frequency and severity of premium payments.
- PBNYSPP (Paid But Not Yet Settled Premium Payment)—These refer to the premiums partly paid to the insurance company but not yet fully settled in full, given that the premiums paid by policyholders in their categories vary according to the policyholder's frequency and severity of premium payments.
- REOPP (Reopened Premium Payment)—These refer to the premiums that were once paid continuously and suddenly stopped due to both certain and uncertain reasons. Now the new streams of premium payments to be rejuvenated by the existing or new policyholder with regards to the same policy are placed in this category.
- REINSPP (Re insured Premium Payment)—This refers to the long overdue, incredibly large unpaid premium which needs to be settled off either by the existing or new policyholder using reinsurance arrangement either internally (with the insurer) or externally (with any other third party of interest).

We also conscripted the existing two main types of actuarial loss reserves in the literature respectively IBNYR and RBNYS. We also introduced two further types of actuarial loss reserves respectively the REOPENED (Reopened Reserve) and the REINSURANCE (Reinsurance Reserve). In short let us define all the adopted and new types of reserves in the context of our automated model for car insurance services and micro finance services as follows.

- IBNYR (Incurred But Not Yet Reported)—these refer to the reserve allocated by the insurer to cater for Incurred claims which are not yet reported or known to the insurance company. These are considered for all the four main policyholder categories defined on **Table 2**.

- RBNYS (Reported But Not Yet Settled)—These refer to the reserves allocated by the insurer to cater for Reported But Not Yet Settled claims both from the micro finance services or car insurance services and both.

- REOPENED (Reopened Reserve)—These are actuarial reserves set aside to cater for reported claims that were once closed without payments or with partial payments for both known and unknown reasons, but have been reopened in future date and the insurer needs to meet these long standing claims whether from micro finance services or car insurance services or even both.

- REINSURANCE (Reinsurance Reserve)—These refer to the reserves allocated by the insurer to cater for catastrophic loss reserving either from from micro finance services or car insurance services or even both.

The terminology for Automated Actuarial Loss Reserving Model as well as Automated Risk Pricing Models have been recorded on **Table 1** below.

Table 1. Terminology proposed for the Automated Actuarial Loss Reserving & Premium pricing model.

Automated Actuarial Loss Reserve Terminology		Automated Actuarial Risk Pricing Terminology	
Type of Reserve	Definition	Type of Premium	Definition
IBNYR	Incurred But Not Yet Reported Reserve	IBNYPP	Incurred But Not Yet Reported Reserve
RBNYS	Reported But Not Yet Settled Reserve	PBNYSPP	Paid But Not Yet Settled Premium Payment
REOPENED	Reopened Reserve	REOPP	Reopened Premium Payment
REINSURANCE	Reinsurance Reserve	REINSPP	Reinsured Premium Payment

The following Policyholder categories have been suggested towards automation of car insurance services and micro finance services respectively.

Since we are automating car insurance services and microfinance banking services we have considered the four types of policyholder categories defined below.

Table 2 shows the policyholder categories and their associated definitions.

Table 2. Automated actuarial loss reserving risk pricing policyholder categories.

Automated Actuarial Loss Reserving Risk Pricing Policyholder Categories	
Category A	Policyholder with both Car Insurance policy & Microfinance policy
Category B	Policyholder with Micro finance policy only
Category C	Policyholder with Car Insurance policy only
Category D	Policyholder with no policy

3.2. Distribution of the Automated Actuarial Loss Reserves and Risk Premiums

The defined types of actuarial reserves and premiums are then distributed according to the portions shown below.

Table 3. Types of reserves and premiums and their associated weights.

Automated Actuarial Loss Reserving Distribution		Automated Actuarial Risk Premium Distribution	
IBNYR	80% of Total Robust AARP	IBNYPP	80% of Total Robust AARR
RBNYS	15% of Total Robust AARR	PBNYPP	15% of Total Robust AARP
REOPENED	4% of Total Robust AARR	REOPP	4% of Total Robust AARP
REINSURANCE	1% of Total Robust AARR	REINSPP	1% of Total Robust AARP

Table 3 shows the distribution of Total AARR (over the defined types of loss reserves) and also the distribution of Total AARP (over the defined types of risk premiums). From the table above each type of loss reserve has its corresponding type of risk premium as also shown by **Table 1**.

3.3. Allocations of the Automated Actuarial Loss Reserves and Risk Premiums over Policyholder Categories

Results of **Table 4** below shows the Policyholder Loss reserving/premium categories and their associated proportion assumed.

Table 4. Automated actuarial loss reserving and pricing assumptions.

	Policyholder Reserve allocation Categories					Policyholder Premium allocation Categories			
	Category A	Category B	Category C	Category D		Category A	Category B	Category C	Category D
IBNYR	50%	30%	20%	0%	IBNYPP	50%	30%	20%	0%
RBNYS	50%	30%	20%	0%	PBNYPP	50%	30%	20%	0%
REOPENED	50%	30%	20%	0%	REOPP	50%	30%	20%	0%
REINSURANCE	50%	30%	20%	0%	REINSPP	50%	30%	20%	0%

Table 4 shows that category A has large outlay with regards to both types of loss reserves and types of risk premiums respectively followed by Category B, then Category C and lastly Category D which is the reference policyholder category.

3.4. The Basis for Automated Actuarial Risk Premium Pricing

Table 5 below shows the types of loading and the associated actuary's estimated loadings these are used for the Automated Actuarial Risk Premium Pricing.

Table 5. Automated actuarial pricing loadings.

Type of Loading	Actuary's estimated value
Profit loading	5%
Solvency loading	4%
Expense	3%
Reinsurance	2%

The pricing loadings are multiplied together with each loading added to 1 in order to give the total premium loadings product which is further multiplied to the estimated AARP. On the same note we do nothing on AARR since the loadings are only for actuarial risk premium pricing.

Using all the assumptions stated in **Table 3** and **Table 4** above we compute the stated types of risk reserves and premiums respectively using the machine learning algorithms considered in this study. From there the autonomous argumentation of Loss Reserving and actuarial premium pricing commences and proceeds as shown by the next stage below.

3.5. The Framework for Automated Actuarial Loss Reserving & Pricing Model

Let us consider the framework below between types of loss reserves and risk premiums which is the foundation for Automated Actuarial Loss Reserving & Pricing model.

3.6. Computation of the Total Automated Actuarial Loss Reserving & Pricing Model Total Reserves and Premiums

Table 3 is the building block towards the simultaneous computation of total risk premiums and loss reserves and results are presented in Appendix and see **Table A2**. The results shown on this table have been arrived at by taking note of subsection 5 under **Table 6**.

Table 6. Derivation of automated actuarial loss reserving & pricing model balances for all policyholder categories.

Type of Reserve	Type of Premiums	Automated Actuarial Loss Reserving & Pricing Model Balances
IBNYR	IBNYPP	IBNYR + IBNYPP
RBNYS	PBNYSPP	RBNYS + PBNYSPP
REOPENED	REOPP	REOPENED + REOPP
REISNURANCE	REINSPP	REINSURANCE + REINSPP

3.7. Distribution of Automated Actuarial Loss Reserving & Pricing Model Balances among Policyholder Categories

At this stage the results from **Table A2** is further allocated among the four main

types of loss reserves and risk premiums using the Automated Actuarial Loss Reserving and Pricing Assumptions presented on **Table 4**. Referring to the Appendix section, we obtain further results respectively; the Automated Actuarial Loss Reserving & Pricing Total IBNYR Reserves and IBNYPP Premiums shown from **Table A3**, Automated Actuarial Loss Reserving & Pricing Total RBNYS Reserves and PBNYSP Premiums from **Table A4**, Automated Actuarial Loss Reserving & Pricing Total ROPENED Reserves and REOPP Premiums from **Table A5** and finally the Automated Actuarial Loss Reserving & Pricing Total REINSURANCE Reserves and REINSPP Premiums from **Table A6**.

3.8. Computation of the Comprehensive Automated Actuarial Loss Reserving & Pricing Balances

Here we sum the above calculated allocated Automated Actuarial Loss Reserving & Pricing Balances between Policyholder Categories from Subsection 6 above and obtain the Comprehensive Automated Actuarial Loss Reserving & Pricing Total Reserves and Premiums shown in the Appendix Section under **Table A7**.

3.9. Computation of the Aggregate Comprehensive Automated Actuarial Loss Reserving & Pricing Balances

At this stage we then proceed to sum the Comprehensive Automated Actuarial Loss Reserving & Pricing Balances determined from Subsection 6 above and hence we arrived at Aggregate Comprehensive Automated Actuarial Loss Reserving & Pricing Total Reserves and Premiums as reflected by **Table A8**.

Evaluation of No Claims Bonus Rates for Automated Actuarial Underwriting Model

Ultimately, we then use the policyholder Category No Claims Bonus rates in order to find the net present values and accumulated values for Aggregate Comprehensive Automated Actuarial Loss Reserving & Pricing Total Reserves and Premiums among the policyholder categories from **Table A8**. These bonus rates are used to navigate Automated Actuarial Underwriting modeling respectively.

Our model is an automation of car insurance services and micro finance services on the same platform with the aim to reducing the comprehensive number of claims by targeting reducing the number of claims (from Policyholders in categories A and B) and number of requests (from Policyholders in categories A and C) whilst simultaneously promoting increase in comprehensive number of payments to the insurer in form of increasing number of premium payments once the premium is due (from Policyholders in categories A and B) and also increasing the number of investments (from Policyholders in categories A and C). As a result, **Table 7** shows No Claims bonus rates for all policyholders in their diverse categories. It is quite interesting that during the Automated Actuarial Underwriting (AAU) modeling the base bonus rates are given to all policyholders as soon as they take any policy (A, B and C) from the insurer, also on the same platform the variable bonus rate is added on to the base bonus rate until we

arrive at the final bonus rate. According to our model the more the policyholder claims the variable bonus rate falls and the lesser the number of claims the better the variable bonus rate. From **Table 7** Policyholder category A has the largest maximum final bonus rate (5%) followed by Policyholder category B's maximum final bonus rate (4%), then Policyholder category C's maximum final bonus rate (3%) and the least being Policyholder category D's maximum final bonus rate (0%) since this category has no participant/active policyholders. In a nutshell these rates are applied on the derived Automated Actuarial Loss Reserving & Pricing Balances (AALRPB) for all policyholders in their respective categories taken from **Table 6**.

Table 7. No claims bonus rates for policyholders.

Policyholder Category No Claims Bonus Rates			
Category	Base bonus rates	Variable bonus rates	Final bonus rates
A	1%	4%	5%
B	1%	3%	4%
C	1%	2%	3%
D	0%	0%	0%

3.10. Automated Actuarial Underwriting Model

The following steps are considered in order to develop the Automated Actuarial Underwriting model.

3.11. Analysis of Surplus and Expenses

The major source of our expenses came from initial expenses, renewal expenses, taxes paid, underwriting costs and fees paid whilst our major source of income came from retained earnings. These incomes and expenses were totaled with respect from each mentioned category in the new test data set to match with the final regression estimating Automated Actuarial Loss Reserving & Pricing (AALRP) models **Table 19**. Soon afterwards we totaled all expenses to get the aggregate total for all the expenses and then distributed them between the policyholder categories A (50%), B (30%), C (20%) and D (0%) see **Table 4**. This enabled us to set up the actuarial underwriting model with respect to each category and also to maintain effectiveness, efficiency and economic distribution of expenses and outgo.

3.12. Calculation of Prevailing Interest Rates for Evaluation of the Automated Actuarial Underwriting Cash Flows

The following method was used for determining the prevailing interest rates respectively. We then simulated prevailing interest rates in R using uniform distribution. We assumed a period of 100 years which we took as n , in addition to that we considered the prevailing annual interest rates to vary between mini-

mum (1%) and maximum (5%). We then computed the average for the simulated prevailing annual interest rates which came to (3.05%). This is the final interest rate which we considered for determining the net present value for evaluation of net AALRPB of expenses and outgo respectively presented by **Table 8** below.

Table 8. Simulated Uniformly distributed prevailing interest rates.

n	100
rmin	1%
rmax	5%
Simulated average rate	3.05%

3.13. Development of Automated Actuarial Underwriting Models for Policyholder Categories

The Automated Actuarial Underwriting model is carried out within each Policyholder category starting with A, B and ending with C. The AAU model proceeds from the AALRP model balances, moreover AAU is evaluated in the following five stages by taking the first 100 years as the longest possible underwriting. Moreover, on the same note, we considered the same model for the first ten years which we proposed to the shortest possible time towards effective model testing towards the proposed actuarial underwriting method suggested in this study. Note that FBR is the Final Bonus Rate and E_t represents the Expenses and outgo at time t .

3.13.1. AAU Stage 1

$$(AAU)_t = \begin{cases} \sum (AALRPB_t \cdot (1 + FBR)^t - E_t) v^t > 0 & \text{AAU is possible} \\ \sum (AALRPB_t \cdot (1 + FBR)^t - E_t) v^t < 0 & \text{Go to next stage} \end{cases} \quad (13)$$

where $AALRPB_t = (IBNYR + IBNYPP)_t$ and $V^t = (1 + \text{average simulated annual prevailing rate})^{-t}$.

3.13.2. AAU Stage 2

$$(AAU)_t = \begin{cases} \sum (AALRPB_t \cdot (1 + FBR)^t - E_t) v^t > 0 & \text{AAU is possible} \\ \sum (AALRPB_t \cdot (1 + FBR)^t - E_t) v^t < 0 & \text{Go to next stage} \end{cases} \quad (14)$$

where $AALRPB_t = (IBNYR + IBNYPP)_t + (RBNYS + PBNYSPP)_t$.

3.13.3. AAU Stage 3

$$(AAU)_t = \begin{cases} \sum (AALRPB_t \cdot (1 + FBR)^t - E_t) v^t > 0 & \text{AAU is possible} \\ \sum (AALRPB_t \cdot (1 + FBR)^t - E_t) v^t < 0 & \text{Go to next stage} \end{cases} \quad (15)$$

where $AALRPB_t = (IBNYR + IBNYPP)_t + (RBNYS + PBNYSPP)_t + (REOPENED + REOPP)_t$.

3.13.4. AAU Stage 4

$$(AAU)_t = \begin{cases} \sum (AALRPB_t \cdot (1 + FBR)^t - E_t) v^t > 0 & \text{AAU is possible} \\ \sum (AALRPB_t \cdot (1 + FBR)^t - E_t) v^t < 0 & \text{Go to next stage} \end{cases} \quad (16)$$

where

$$AALRPB_t = (IBNYR + IBNYPP)_t + (RBNYS + PBNYSPP)_t + (REOPENED + REOPP)_t + (REINSURANCE + REINSPP)_t$$

3.13.5. AAU Stage 5

$$(AAU)_t = \sum (AALRPB_t \cdot (1 + FBR)^t - E_t) v^t > 0 \quad \text{AAU is possible} \quad (17)$$

where

$$AALRPB_t = (IBNYR + IBNYPP)_t + (RBNYS + PBNYSPP)_t + (REOPENED + REOPP)_t + (REINSURANCE + REINSPP)_t + \text{TOTAL RETAINED INCOME}_t$$

3.14. Novelty of the Methodology

The proposed methodology in this study introduces several novel concepts and builds upon existing machine learning techniques: Introducing terms such as IBNYPP (Incurred But Not Yet Paid Premium) and PBNYSPP (Paid But Not Yet Settled Premium Payment) to better categorize and manage different types of premiums. The study integrates new types of reserves like REOPENED and REINSURANCE reserves into the actuarial framework, expanding the traditional categories to better capture long-standing and catastrophic risks. The multi-stage Automated Actuarial Underwriting (AAU) model evaluates underwriting feasibility over multiple stages, considering various reserve and premium categories for more precise decision-making. The incorporation of an inflation-adjusted model accounts for economic changes, enhancing the accuracy of predictions over time. This methodology not only leverages the strengths of machine learning algorithms but also introduces a structured framework that comprehensively addresses the complexities of modern insurance practices.

4. Data

We used simulated Comprehensive General Car Insurance and Microfinance data from 1989 to 2022, a period of 33 years and a sample of 40,000 policyholders was considered for data analysis.

This data is divided into the following seven main parts namely, Policyholder Personal Data, Microfinance Policyholder Data, Policyholder Vehicle Data, Comprehensive Policyholder Claim Related Data, Comprehensive Policyholder Premium Payments Related Data and Policyholder External Data. From this data we simulated the 48 variables from the seven parts of data and hence we used it to develop the two mainline models which are the Automated Actuarial Loss

Reserving Model and the Automated Risk Pricing Models which we used to compute the Automated Actuarial Loss Reserving Risk Pricing Balances (AALRRPB) which we used for developing an Automated Actuarial Underwriting Model. Out of the 48 data variables major emphasis was placed on the four main principal/key variables (defined on the subsection 0.0.12 below) which we applied towards automating the car insurance services and the microfinance services on the same platform.

4.1. Principal Data Variable Exploratory Analysis

The following are defined principal data variables and how they have been derived.

- Comprehensive Claim Amount = (Claims Incurred + Amount Requested)—where Claim Incurred comes from car insurance services and Amount Requested comes from microfinance services.
- Comprehensive Number of Claims = (Number of Claims + Number of Requests)—where Number of Claims comes from car insurance services and Number of Requests comes from microfinance services.
- Comprehensive Payment Amount = (Current Premium + Amount Invested)—where the Current Premium comes from car insurance service policyholders and Amount Invested comes from microfinance services policyholders.
- Comprehensive Number of Payments = (Number of Investments + Number of Premium Payments)—where both the Number of Investments and Number of Requests come from policyholders affiliated to both or either car insurance services or microfinance services.

Further exploratory results for the key principal variables are shown in **Figures 1-4** below.

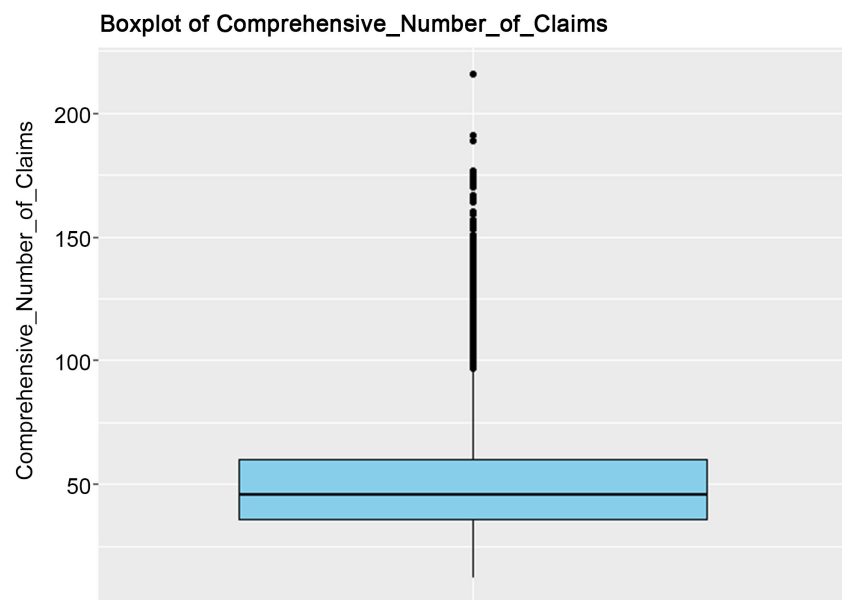


Figure 1. Box plot for Comprehensive Number of Claims.

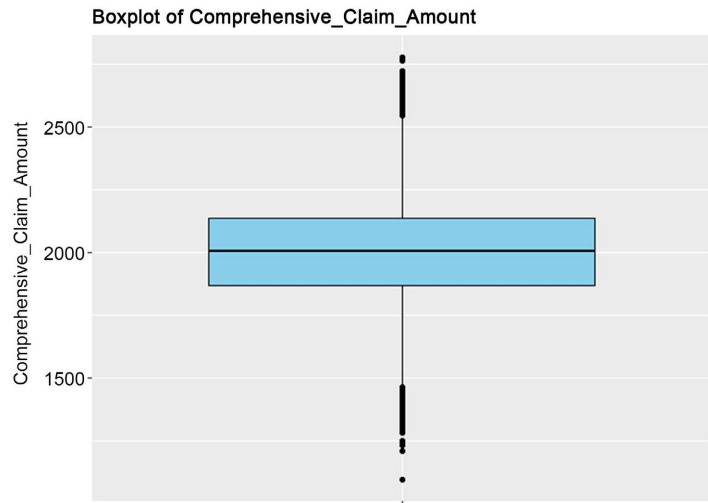


Figure 2. Box plot for Comprehensive Claim Amount.

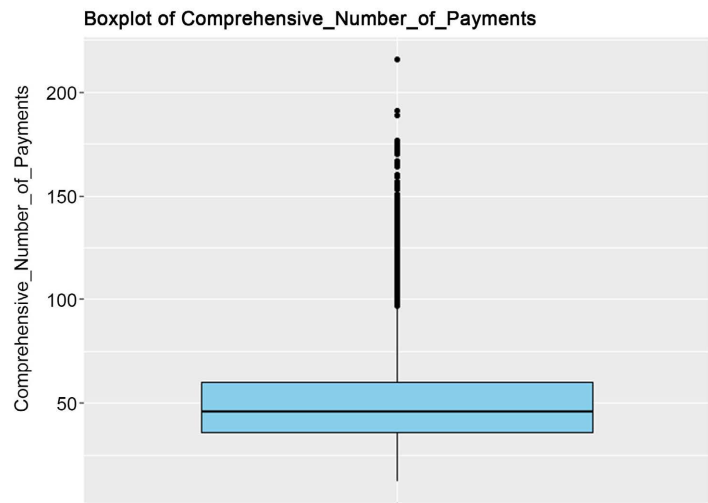


Figure 3. Box plot for Comprehensive Number of Premium Payments.

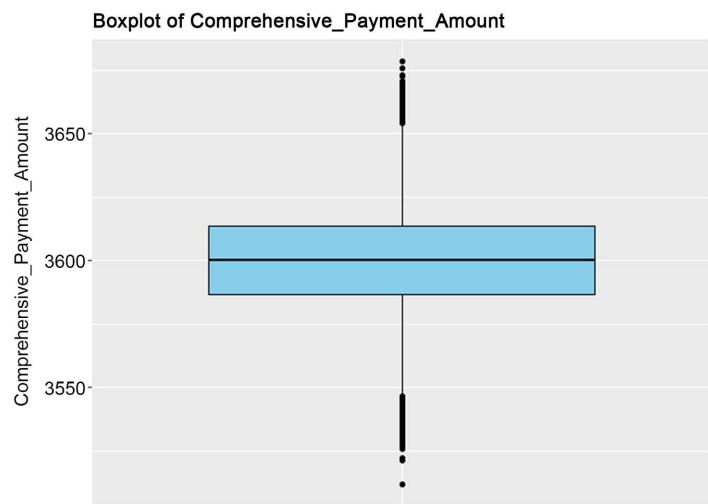


Figure 4. Box plot for Comprehensive Payment Amount.

Figures 1-4 reveal that the box in each plot represents the interquartile range (IQR), which contains the middle 50% of the data. The bottom and top edges of the box represent the first quartile (Q1) and third quartile (Q3) respectively. The length of the box (height in vertical box plots) indicates the spread of the middle 50% of the data. The line inside the box represents the median value of the data. It shows the central tendency of the distribution. The whiskers extend from the edges of the box to the smallest and largest observations within 1.5 times the IQR from the edges of the box. They indicate the range of the data, excluding potential outliers. Individual points beyond the whiskers represent potential outliers—data points that are significantly different from the rest of the data. Outliers may suggest unusual or extreme values in the simulated data set.

4.2. Correlation Analysis for the Principal Variables

Correlation analysis is a statistical technique used to measure the strength and direction of the relationship between two or more variables. It quantifies the degree to which changes in one variable are associated with changes in another variable [28]. The result of correlation analysis is a correlation coefficient ρ , which indicates the strength and direction of the relationship between variables. The ρ typically ranges between -1 and 1 . Furthermore, correlation coefficient of 1 indicates a perfect positive correlation, meaning that as one variable increases, the other variable also increases proportionally. A ρ of -1 indicates a perfect negative correlation, meaning that as one variable increases, the other variable decreases proportionally. A ρ of 0 indicates no correlation between the variables. There are different methods to calculate correlation coefficients, including Pearson correlation coefficient (for linear relationships), Spearman ρ (for monotonic relationships), and Kendall correlation coefficient (for ordinal relationships) [29]. The choice of method depends on the nature of the data and the relationship being investigated.

Figure 5 shows that there is perfect correlation among the principal data variables themselves as shown by the red colored box with value of (1.00) . On the same note the correlation between the other variables is nearly 0.00 which shows that there is independence and absolutely no dependency among the other variables. This clearly shows that the claims and payment characteristics of policyholders in their diverse categories are very random. As a result of this, there is need for insurance company to set aside readily available large stake of Automated Actuarial Loss Reserving Risk Pricing Balances (AALRRPB) to cater for uncertain claims which this study has provided an alternative solution to this problem. This is also validated by the correlation matrix presented on **Table 9**.

4.3. Factor Analysis (FA) and Principal Component Analysis (PCA) for Key Data Variables

Factor Analysis (FA) and Principal Component Analysis (PCA) are both multivariate statistical techniques used for data reduction and dimensionality reduction [30]. Factor Analysis is a statistical method used to identify underlying fac-

tors or latent variables that explain the correlations among observed variables. In addition to that, the primary goal of FA is to uncover the structure of relationships between observed variables and to understand the underlying constructs or factors that are responsible for these relationships.

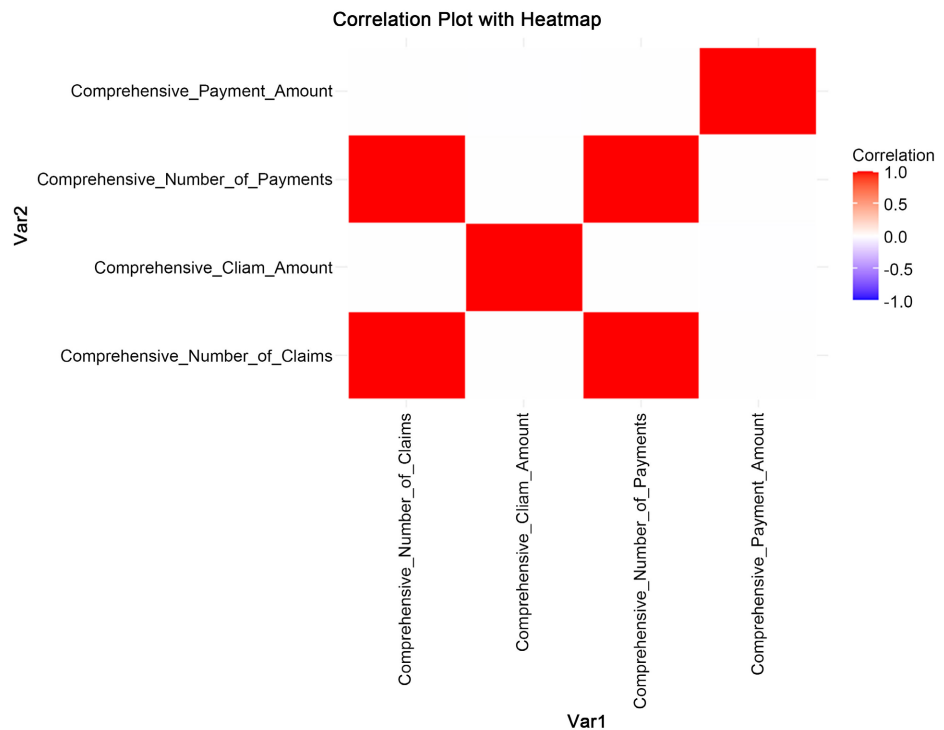


Figure 5. Correlation plot with heat map for the principal variables.

Table 9. Correlation matrix for the principal variables.

	Comprehensive Number of Claims	Comprehensive Claim Amount	Comprehensive Number of Payments	Comprehensive Payment Amount
Comprehensive Number of Claims	1.00	– 0.00	1.00	– 0.00
Comprehensive Claim Amount	– 0.00	1.00	– 0.00	– 0.01
Comprehensive Number of Payments	1.00	– 0.00	1.00	– 0.00
Comprehensive Payment Amount	– 0.00	– 0.01	– 0.00	1.00

FA assumes that observed variables are linear combinations of unobserved (latent) factors plus error terms. It seeks to explain the covariance between variables in terms of a smaller number of latent factors. Furthermore, FA allows for the interpretation of the data in terms of underlying constructs and can be used for data reduction and simplification. Factor loadings represent the correlations between observed variables and underlying factors, and they indicate how much

each variable contributes to each factor.

Principal Component Analysis is a technique used to transform a set of correlated variables into a set of uncorrelated variables called principal components [31]. The primary goal of PCA is to reduce the dimensionality of the data while preserving as much of the variability in the data as possible. PCA achieves dimensionality reduction by identifying a set of orthogonal axes (principal components) that capture the maximum variance in the data. Unlike Factor Analysis, PCA does not assume the presence of latent factors underlying the observed variables. Instead, it focuses on finding linear combinations of variables that explain the most variance in the data.

Moreover, PCA does not provide direct interpretation of underlying constructs or factors. Instead, it provides a data-driven representation of the structure of the data based on variance [32]. Principal components are ordered in terms of the amount of variance they explain, with the first component explaining the most variance and subsequent components explaining decreasing amounts of variance.

In short, while both FA and PCA are used for data reduction and dimensionality reduction, Factor Analysis focuses on uncovering underlying constructs or factors that explain correlations between observed variables, while PCA focuses on capturing the maximum variance in the data through orthogonal transformations of the variables. The choice between FA and PCA depends on the research question, assumptions about the data, and the objectives of the analysis [32].

Figure 6 below shows the importance of these key variables in our study.

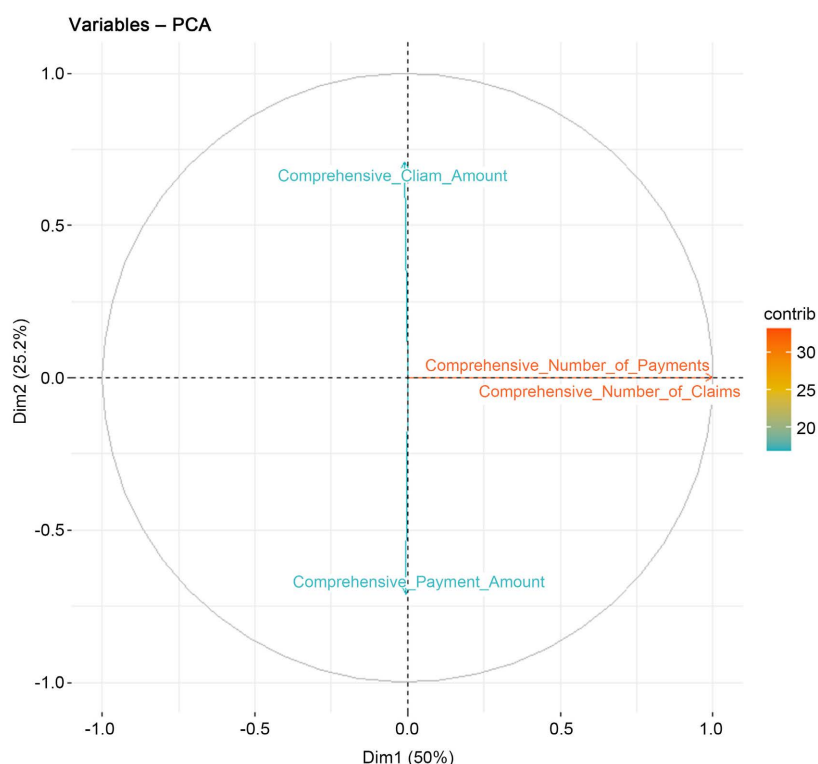


Figure 6. Factor analysis for the principal variables.

From **Figure 6**, there is evidence of the presence of unit root which suggests both weak and strong negative and positive influence of each key variable in the study. As a result these variables have been used to develop the Automated Actuarial Pricing and Underwriting Model respectively presented by the **Table 10** below.

Table 10. Eigenvalues for the principal variables.

Metrics	Dimension 1	Dimension 2	Dimension 3	Dimension 4
Variance	2.000	1.006	0.994	0.000
% of variance	50.001	25.160	24.839	0.000
Cumulative % of variance	50.001	75.161	100.000	100.000

The eigenvalues are all positive and greater than 1 which suggests that the four variables are very important in the study, see **Table 11** below.

Table 11. Standardized loadings (pattern matrix) based upon correlation matrix.

Metrics	MR1	MR2	MR4	MR3	h2	u2	com
Comprehensive Number of Claims	1	0.00	0	0	0.9975	0.0025	1
Comprehensive Claim Amount	0	- 0.08	0	0	0.0064	0.9936	1
Comprehensive Number of Payments	1	0.00	0	0	0.9975	0.0025	1
Comprehensive Payment Amount	0	0.08	0	0	0.0064	0.9936	1

The standardized loadings, often found in factor analysis, represent the correlations between observed variables (indicators) and latent factors, which are standardized to have a mean of zero and a standard deviation of one. These loadings indicate the strength and direction of the relationship between each observed variable and the underlying factor (s) being measured. The absolute value of the loading indicates the strength of the relationship between the variable and the factor. Higher absolute values (closer to 1) suggest a stronger relationship and this is presented by Comprehensive Number of Claims and Comprehensive Number of Payments. A positive loading indicates a positive correlation between the variable and the factor, meaning that as the variable increases, so does the underlying factor. Conversely, a negative loading indicates an inverse relationship: as the variable increases, the factor decreases.

In factor analysis, “SS Loadings” stands for “Sum of Squared Loadings.” It is a metric that represents the proportion of variance in the observed variables (indicators) that is accounted for by each latent factor. SS Loadings are calculated by summing the squares of the factor loadings for each variable on a specific factor. Proportion of Variance Explained: SS Loadings provide insight into how much of the total variance in each observed variable is explained by the under-

lying factors. In addition to that, we have higher SS Loadings indicate that a larger proportion of the variance in the variable is captured by the factor. **Table 12** indicates positive and large proportion variance and Cumulative Variance which ultimately leads to large Proportion Explained and Cumulative Proportion a clear indication that the four factors are important in the model. In short, the SS Loadings can be used as indicators of the goodness-of-fit of the factor model. The higher SS Loadings displayed suggest a better fit, as they indicate that the factors are accounting for a larger proportion of the variance in the data.

Table 12. SS Loadings metrics.

Metrics	MR1	MR2	MR4	MR3
SS loadings	2.00	0.01	0.0	0.0
Proportion Variance	0.50	0.00	0.0	0.0
Cumulative Variance	0.50	0.50	0.5	0.5
Proportion Explained	0.99	0.01	0.0	0.0
Cumulative Proportion	0.99	1.00	1.0	1.0

4.4. Data Preprocessing, Data Scaling and Data Partitioning

After loading the data in R, we both hot encoded the data using R caret package and preprocessed it using the min-max approach before partitioning it into training data set (80%) and test data set (20%).

5. Main Results

This section describes the results obtained with special attention to the methodology outlined in the previous section.

5.1. Automated Actuarial Loss Reserving Risk Pricing Frequency Models

The Automated Actuarial Loss Reserving Risk Pricing Frequency models are presented below.

Table 13 shows that SVM took affectionately long time to process the results for both Automated Loss Reserving and Automated Actuarial Risk Pricing frequency models. On the same spot both GLM, GAM and ANN converged faster for both models. In addition to that, Regression Trees (RPART), Random Forests (RANGER), XGB and LAR took much longer to process both the frequency models. All the machine learning models used to construct both types of models gave almost the same values with regards to model evaluation and performance metrics which are MSE, MAE and RMSE and this reflects both the reliability and validity of machine learning models.

5.2. Automated Actuarial Loss Reserving Risk Pricing Severity Models

The Automated Actuarial Loss Reserving Risk Pricing Severity Models are presented below.

Table 13. Automated actuarial loss reserving risk pricing frequency models.

Actuarial Loss Reserve Frequency Models					Actuarial Risk Pricing Frequency Models			
ML Model	Time (sec)	MSE	MAE	RMSE	Time (sec)	MSE	MAE	RMSE
GLM	1.34	2,861.6360	49.8888	53.4943	0.72	2,880.7160	49.9858	53.6723
GAM	1.16	2,863.4360	49.9662	53.5111	0.72	2,874.9590	50.0743	53.6186
RPART	2.35	2,853.3270	49.8003	53.4166	2.18	2,852.3730	49.9263	53.4076
RANGER	55.93	2,833.7720	49.7191	53.2332	63.06	2,869.8760	49.9212	53.5712
XGB	6.62	2,852.3530	49.9262	53.4074	6.90	2,829.1380	49.6351	53.1896
LAR	12.18	2,882.8090	50.1599	53.6918	16.05	2,866.5350	49.9787	53.5400
SVM	289.67	2,850.8520	49.7783	53.3934	381.22	2,865.5540	49.9669	53.5309
ANN	8.56	2,883.5730	50.2280	53.6989	8.30	2,853.1610	49.8162	53.4150

Table 14 reveals that the SVM machine learning model respectively took much longer time to give results for both Automated Loss Reserving Severity and Automated Risk Pricing Severity models. On the same note, GLM and GAM were the most fastest models that gave the results in the shortest possible time looking at both severity models. On the same note, Regression Trees (RPART), Random Forests (RANGER), XGB and LAR took much longer to process both severity models. From **Table 10**, all the machine learning models gave almost the same values for the machine learning model evaluation and performance metrics which indicates the reliability and validity of machine learning models.

Table 14. Automated actuarial loss reserving risk pricing severity models.

Actuarial Loss Reserve Severity Models					Actuarial Risk Pricing Severity Models			
ML Model	Time (sec)	MSE	MAE	RMSE	Time (sec)	MSE	MAE	RMSE
GLM	0.46	4,037,975.0000	1,999.3260	2,009.4710	0.44	12,957,822.0000	3,599.6420	3,599.6980
GAM	0.99	4,047,282.0000	2,001.2970	2,011.7860	0.83	12,957,248.0000	3,599.5630	3,599.6180
RPART	1.69	4,030,834.0000	1,997.4460	2,007.6940	1.64	12,958,334.0000	3,599.7110	3,599.7690
RANGER	269.12	4,047,592.0000	2,002.1210	2,011.8630	369.02	12,956,817.0000	3,599.5020	3,599.5580
XGB	6.73	4,053,955.0000	2,003.2360	2,013.4440	8.20	12,956,750.0000	3,599.4940	3,599.5490
LAR	15.24	4,049,226.0000	2,002.2220	2,012.2690	11.11	12,956,903.0000	3,599.5120	3,599.5700
SVM	264.64	4,050,622.0000	2,002.3030	2,012.6160	348.46	12,956,848.0000	3,599.5040	3,599.5620
ANN	9.11	4,047,632.0000	2,001.8680	2,011.8730	9.30	12,956,123.0000	3,599.4070	3,599.4620

5.3. Automated Actuarial Loss Reserving Risk Pricing Inflation Models

This model is aimed at both adjusting and stabilizing the frequency-severity models above with the zeal of coming up with inflation adjusted automated actuarial loss reserves risk premiums balances accordingly. Since we appreciate that the higher the rate of inflation the faster the rate of comprehensive claims as compared to comprehensive payments and the reverse being true [33], we con-

sidered running the two inflation models one for adjusting and stabilizing the Automated Actuarial Loss Reserving frequency and severity models and also the other doing the same task with respect to Automated Actuarial Risk Pricing frequency and severity models as shown below.

Table 15 shows that SVM was the slowest model to give the results for inflation adjusted frequency severity models for both models. This is well exhibited by their exceptionally large processing time compared to GLM, GAM, RPART and ANN whose processing time is quite faster. On the same note, RANGER and LAR took much longer processing time with respect to the two presented models. **Table 11** also shows that all the eight machine learning models gave approximately quite similar results for performance metrics which absolutely displays reliability and validity of results.

Table 15. Automated actuarial loss reserving risk pricing inflation models.

Actuarial Loss Reserve Inflation Models					Actuarial Risk Pricing Inflation Models			
ML Model	Time (sec)	MSE	MAE	RMSE	Time (sec)	MSE	MAE	RMSE
GLM	0.65	0.2630	0.5124	0.5129	0.57	0.2623	0.5117	0.5121
GAM	0.84	0.2621	0.5116	0.5120	0.87	0.2616	0.5111	0.5115
RPART	0.79	0.6858	0.8232	0.8281	0.76	0.2618	0.5113	0.5117
RANGER	62.91	0.2626	0.5120	0.5124	101.06	0.2621	0.5115	0.5119
XGB	6.82	0.2620	0.5113	0.5119	8.23	0.2628	0.5121	0.5126
LAR	31.19	0.2625	0.5120	0.5124	14.42	0.2623	0.5118	0.5122
SVM	1095.67	0.2637	0.5121	0.5135	958.20	0.2639	0.5123	0.5137
ANN	6.20	0.2623	0.5116	0.5121	6.89	0.2625	0.5118	0.5123

5.4. Total Automated Actuarial Inflation Adjusted Loss Reserves and Risk Premiums

Total Automated Actuarial Inflation Adjusted Loss Reserves and Risk Premiums is computed for each machine learning algorithm as shown below. The results are shown on **Table 16**.

Table 16. Total automated actuarial inflation adjusted risk reserves and premiums predictions.

Actuarial Loss Reserve Models		Actuarial Risk Pricing Models	
ML Model	Total Predictions	ML Model	Total Predictions
GLM	405.6083	GLM	391.9951
GAM	400.6047	GAM	395.6988
RPART	182.2968	RPART	397.2223
RANGER	402.8569	RANGER	395.4614
XGB	405.3866	XGB	393.1049
LAR	402.2832	LAR	396.5535
SVM	397.9315	SVM	397.2021
ANN	405.2760	ANN	398.4448

Table 16 shows that the Total Automated Actuarial Inflation Adjusted Risk Reserves and Premiums predictions are approximately close to each other since their model performance metrics revealed that they were also similar to each other.

5.5. Final Machine Learning Models for Estimating and Predicting the Robust Automated Actuarial Loss Reserves and Risk Premiums

Using the approach explained in our methodology sections subsection 0.0.4 and 0.0.5 we obtained the following final model for Automated Actuarial Loss Reserving Risk Pricing model.

Table 17 shows that RANGER, LAR, XGB and ANN algorithms took the high processing time across the two final automated models for Automated Actuarial Loss Reserving (AALR) and Automated Actuarial Risk Pricing (AARP). Moreover, the similar results obtained for predicted values for AALR and AARP reveal homogeneity, consistency, stability and reliability of results.

Table 17. Final Machine Learning models for estimating and predicting the Robust Automated Actuarial Loss Reserves and Risk Premiums.

Second Stage Actuarial Loss Reserve Models							Second Stage Actuarial Risk Pricing Models					
ML Model	Time (sec)	pred value	Max	Min	Range	RMSE	Time (sec)	pred value	Max	Min	Range	RMSE
GLM	0.02	2,002.8090	2,676.8750	1,353.8920	1,322.9830	0.0000	0.02	3,601.2530	3,665.8950	3,528.9690	136.9264	0.0000
GAM	0.00	1,998.4520	2,673.9730	1,306.0440	1,367.9280	0.0000	0.04	3,600.1060	3,656.9920	3,521.1380	135.8538	0.0000
RPART	0.00	1,954.2920	2,423.8510	1,583.0360	840.8148	37.7813	0.01	3,605.1970	3,641.4770	3,556.6680	84.8098	3.7731
RANGER	2.53	2,001.7290	2,666.5370	1,281.9200	1,384.6170	277.0213	9.56	3,599.3610	3,660.7060	3,540.9170	119.7890	0.0964
XGB	0.34	2,004.1210	2,713.3630	1,335.6060	1,377.7570	2.6967	0.39	3,600.4250	3,662.4330	3,537.0560	125.3770	0.2295
LAR	0.69	2,000.9420	2,610.6030	1,306.0440	1,304.5590	0.0000	0.65	3,599.6900	3,668.7050	3,531.7700	136.9346	0.0000
SVM	0.48	1,039.6070	2,020.6850	150.4887	1,870.1960	997.5602	0.52	1,829.1000	3,358.0400	417.0973	2,940.9420	1,821.3850
ANN	0.39	918.4288	1,857.0300	89.1541	1,767.8760	1,108.6850	0.33	1,835.7190	3,405.6240	365.6422	3,039.9820	1,826.9620

5.6. Distribution of Automated Actuarial Loss Reserves Risk Pricing Balances: Total Reserves and Total Premiums

Figure 7 below shows how the Automated Actuarial Loss Reserves Risk Pricing Balances (ALRRPB): Total Reserves and Total Premiums have been distributed among the types of loss reserves and also types of risk premiums, see the methodology subsection 6, as a result, the tables of results presented in the Appendix section shows the outcome, see **Table A2** with loss reserve-risk premium allocations taken from **Table 3**.

Figure 7 shows the AALRRPB (Total Reserves + Total Premiums) derived from the Automated Actuarial Loss Reserves and Automated Actuarial Risk

Premiums predicted from their final models as respectively shown on **Table 13**. From **Figure 7** RANGER scooped the largest stake for AALRRPB respectively.

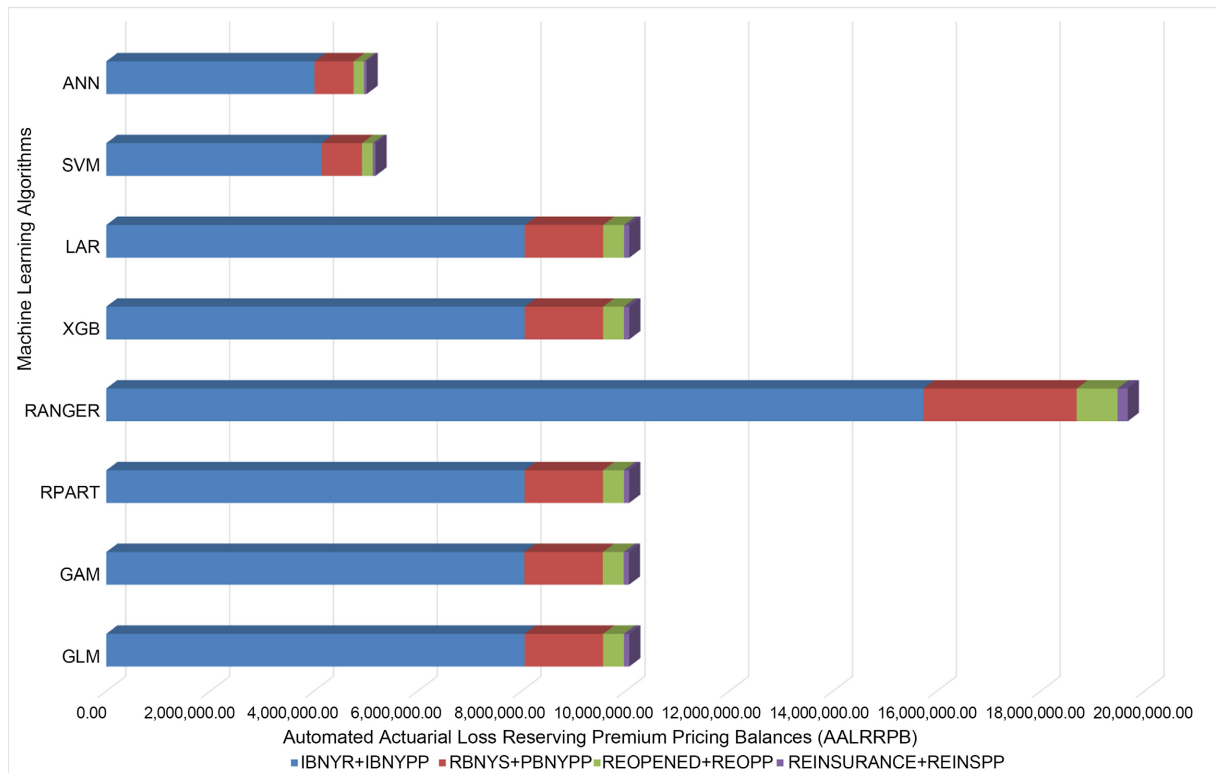


Figure 7. Automated actuarial loss reserves risk pricing balances (Total Reserves and Total Premiums).

5.6.1. Total Automated Actuarial Loss Reserves Risk Pricing Balances

Figures 8-11 below shows the allocated the Total Automated Actuarial Loss Reserves Risk Pricing Balances distributed across policyholder loss reserving-risk premium categories.

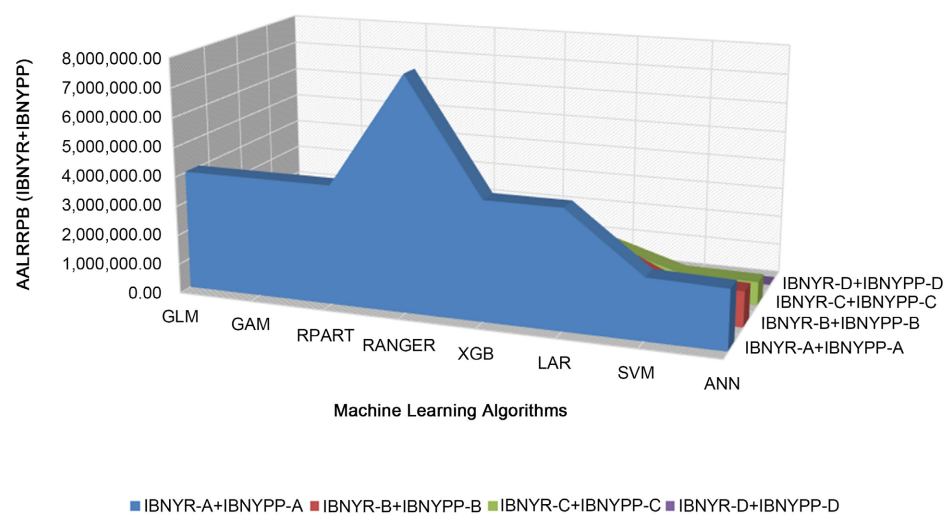


Figure 8. AALRRPB (IBNYR + IBNYPP).

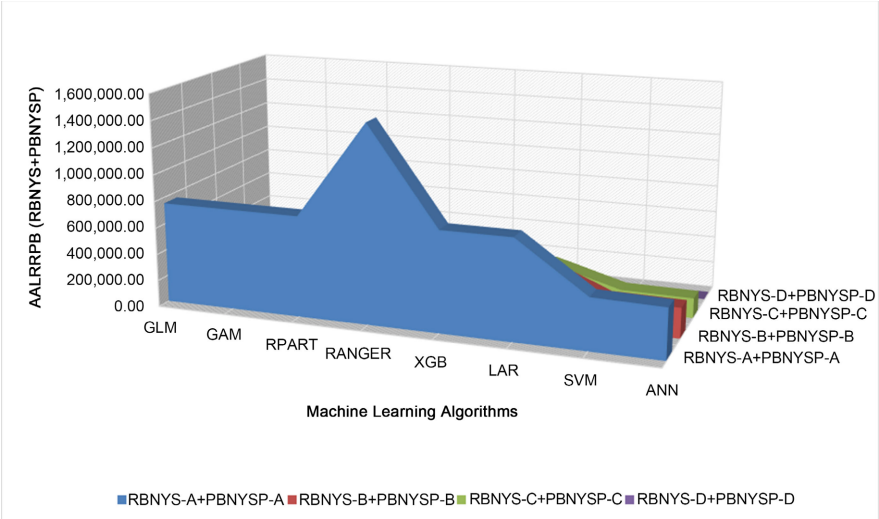


Figure 9. AALRRPB (RBNYS + PBNYSP).

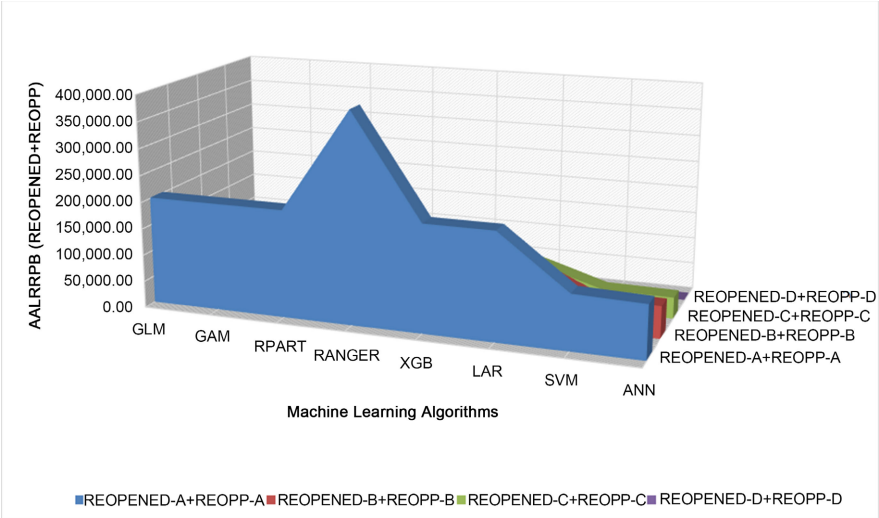


Figure 10. AALRRPB (REOPENED + REOPP).

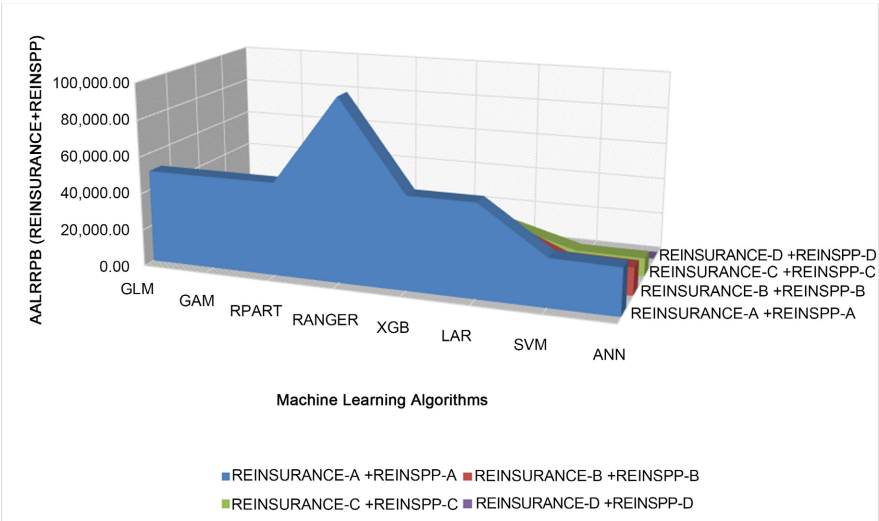


Figure 11. AALRRPB (REINSURANCE + REINSPP).

Figures 8-11 above show that the RANGER algorithm recorded a large stake of AALRRPB (IBNYR + IBNYPP), AALRRPB (RBNYS + PBNYSPP), AALRRPB (REOPENED + REOPP) and also AALRRPB (REINSURANCE + REINSPP). Across all the policyholder categories the AALRRPB (IBNYR + IBNYPP) enjoyed the largest portion, followed by AALRRPB (RBNYS + PBNYSPP), then AALRRPB (REOPENED + REOPP) and finally the AALRRPB (REINSURANCE + REINSPP) being the least. From **Figures 8-11**, it is clear that Policyholder category A enjoyed the largest portion of four main types of reserves and risk premiums followed by Category B, then Category C and lastly Category D. Nevertheless, Category D has attained (0%) across all the types of loss reserves and risk premiums, since there are no any active policyholders as also presented on **Table 4**. In reference to **Figures 8-11** large quantities of AALRRPB have been allocated to (IBNYR + IBNYPP) layer since it caters most for unreported comprehensive claims coming from both the microfinance policyholders and car insurance policyholders, as a result Category A which consists of policyholders with both policies has been allocated a larger portion. Consequently, the presence of other premium-reserve portions such as AALRRPB (RBNYS + PBNYSPP), AALRRPB (REOPENED + REOPP) and AALRRPB (REINSURANCE + REINSPP) in all policyholder categories reduces the reinsurance layers and costs to minimum levels if not zero, whilst promoting catastrophic reserving and comprehensive claim settlements in the shortest possible time with very minimum if at most zero reporting and settlement delays.

5.6.2. Comprehensive Automated Actuarial Loss Reserves Risk Pricing Balances (CAALRRPB)

The CAALRRPB have been obtained by summing the Comprehensive Automated Actuarial Loss Reserves (CAALR) and Comprehensive Automated Actuarial Risk Premiums (CAARP) as directed under methodology sub subsection 6 and results are also presented on the Appendix section, see **Table A7** whose results are shown by **Figure 12**.

Since Category A carries the largest stake on allocation and **Figure 12**. Comprehensive Automated Actuarial Loss Reserves Risk Pricing Balances distribution of reserves, see **Table 4** followed by Category B, then Category C and eventually Category D the reference. Clearly from **Figure 12**, RANGER recorded the highest value for AALRRPB (CAALR + CAARP) respectively.

5.6.3. Aggregate Comprehensive Automated Actuarial Loss Reserves Risk Pricing Balances (ACAALRRPB)

This is obtained by summing the calculated CAALRRPB with regards to methodology Subsection 6. **Table 7** shows the results which then are produced in **Figure 13**.

From **Figure 13**, the ACAALRRPB for RANGER recorded the highest peak which apparently makes it the best model for Automated Actuarial Underwriting process. This is so because by having large stake of ACAALRRPB set aside

that means the insurance company has the largest outlay of funds in form of ACAALRRPBs set aside for uncertain comprehensive claims from the main three policyholder categories defined in this study. This also prepares the insurer for large loss reserving respectively the catastrophic loss reserving as well as risk premium pricing.

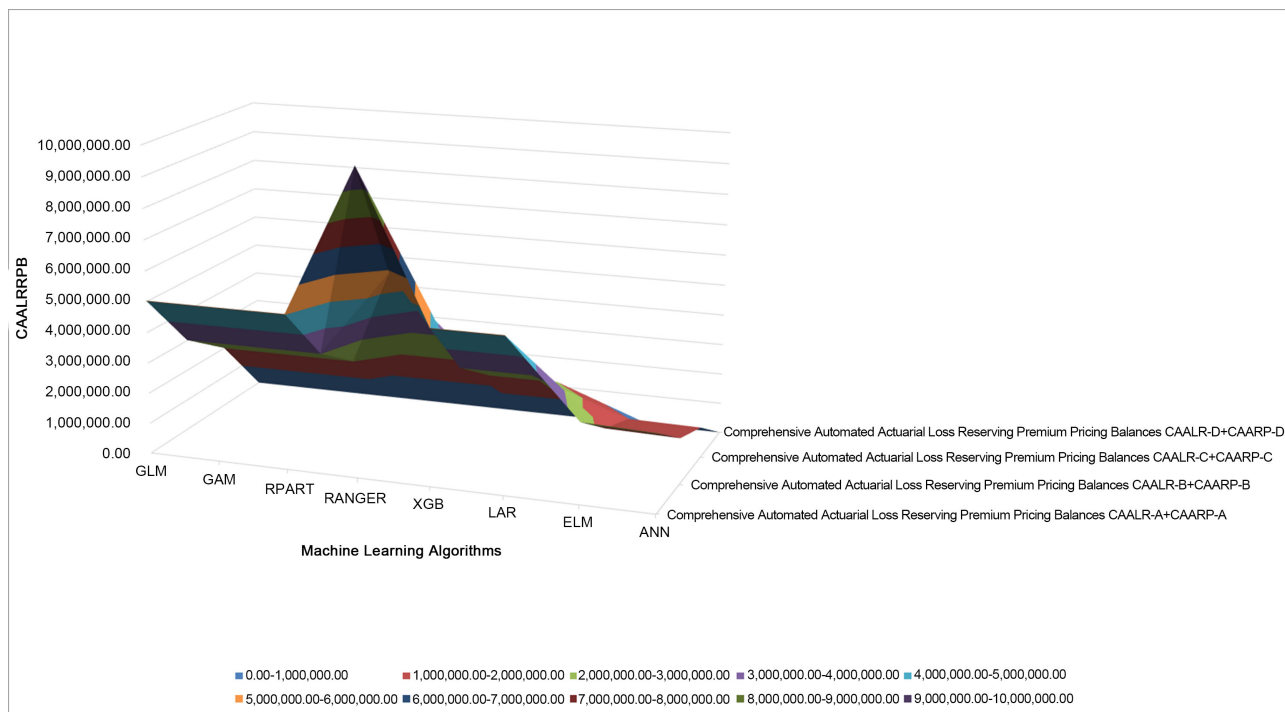


Figure 12. Comprehensive automated actuarial risk reserves and risk premiums.

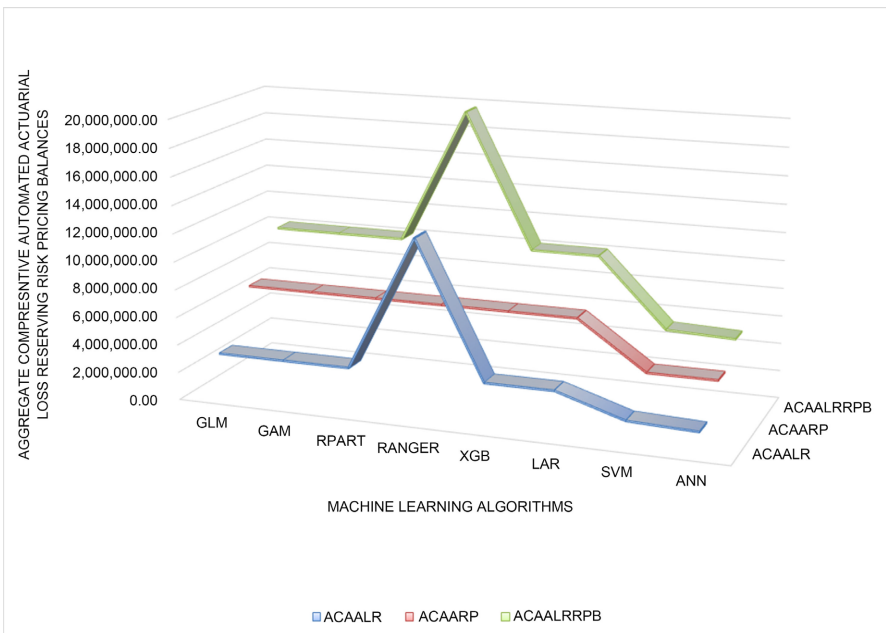


Figure 13. Aggregate comprehensive automated actuarial risk reserves and risk premiums.

5.6.4. Analysis of Surplus and Expenses for the Automated Actuarial Underwriting Model

This section commences the journey towards Automated Actuarial Underwriting with regards to **Table 18** below.

Table 18. Analysis of surplus and expenses.

Analysis of Surplus and Expenses			
Total Expenses & Outgo		Total Profits & Retained Income	
Total Claims Incurred	0.00	Total Retained Income	1,200,082,903.00
Total Claims Outstanding	0.00		
Total Amounts Requested	7,999,971.00		
Total Licenses Paid	80,223.06		
Total Tax Paid	160,084.90		
Total Initial Expenses	1,601,199.00		
Total Renewal Expenses	160,271.60		
Total Underwriting costs	1,201,529.00		
Total Expenses	11,203,278.56	Total Profits & Retained Income	1,200,082,903.00
Surplus	1,188,879,624.44		
	1,200,082,903.00		1,200,082,903.00

Table 18 shows evidence of positive surplus of (1,188,879,624.44) coming from retained earnings. This is a positive sign for cash generative business, it also reflects profitability and better liquidity levels since the total Retained income is able to cover all the expenses and outgo in this case. Moreover, the positive and large value for the surplus is a clear sign that there is credible business expansion as a result of automation of car insurance services and microfinance services simultaneously. On the same note high value for the surplus indicates more influx of income from comprehensive payments inform of investments and premium payments while on the same note there is credible evidence of decreasing comprehensive claims coming from car insurance policy claims and amount of money requested by policyholders in their diverse categories as indicated by recorded zero values for Total Claims incurred and Total Claims Outstanding. This is facilitated, by the fact that the policyholders in their categories seek to keep the frequency and severity of claims lower so that they benefit from a better rate of return on their stake credited with the maximum final bonus rate.

Table 19 shows the distribution of total expenses according to the proposed

allocations presented on **Table 4**. The total expenses and outgo of (11,203,278.56) were distributed according to their allocation shown in **Table 19**. As a matter of consequence, the Category A carries the largest figure for total expenses and outgo with (5,601,639.28), followed by Category B with total expenses and outgo totaling (3,360,983.57), Category C's total expenses and outgo comes to a figure of (2,240,655.71). Moreover Category D, had zero expenses and outgo since it is the reference category without any participant policyholder. These values of total expenses and outgo recorded by each policyholder category are employed towards the Automated Actuarial Underwriting modeling process with regards to each category.

Table 19. Distribution of total expenses according to policyholder categories.

Analysis of Expenses by Categories		
Category	Allocation	Expenses in each Category
A	50%	5,601,639.28
B	30%	3,360,983.57
C	20%	2,240,655.71
D	0%	0.00
Total Expenses	100%	11,203,278.56

5.7. Automated Actuarial Underwriting Model

This section shows results for the Automated Actuarial Underwriting from Automated Actuarial Loss Reserving Risk Pricing models and the following assumptions were used to develop the following models.

Assumptions of the Automated Actuarial Underwriting Model

- The moment a policyholder takes the policy he/she receives the base bonus rates shown on **Table 7**.
- The AALRRPB are compounded over n period of time to forecast their respective accumulated value using final bonus rates.
- n can be number of days, number of weeks, number of months and or number of years.
- The number of comprehensive payments is greater than the number of comprehensive claims.
- Expenses and outgo are deducted from the accumulated AALRRPB and the net balance is evaluated at prevailing rates shown on **Table 8**.
- The frequency, Severity and inflation rates are constant over n .
- The expenses and outgo are constant over n .
- Random Forest (RANGER) being the best model machine learning model in the study has been used to develop the final Automated Actuarial Underwriting Model, see **Figure 13**.
- Automated Actuarial Underwriting model is carried out for categories A, B

and C only since D is the reference category without any active policyholder.

The Automated Actuarial Underwriting process commences with AAUM Stage 1 0.0.8 followed by AAUM Stage 2 0.0.9, then AAUM Stage 3 0.0.10 and finally AAUM Stage 4 0.0.11. The ACAALRRPB were cumulatively added at each stage and thereafter remaining with net cash flows after taking away the respective expenses and outgo in the corresponding policyholder category, see **Table 19**. The net present value is then computed for the net cash flow balances using the prevailing interest rate shown on **Table 8**.

We tested our model with $n = 100$ years which is essentially the longest possible time the insurance company operating with the assumptions discussed above applying and hence the Automated Actuarial Underwriting results are shown below.

Table 20. Policyholder category based automated actuarial underwriting model.

Category A	AALRRPB	Cumulative AALRRPB	Net AAU Balance	Result	AAU Effect
AAU Stage 1	7,870,979.50	7,870,979.50	51,027,265.91	Large and positive	Possible
AAU Stage 2	1,475,808.50	9,346,788.00	23,121,306.98	Large and positive	Possible
AAU Stage 3	393,548.95	9,740,336.95	9,001,245.97	Large and positive	Possible
AAU Stage 4	98,387.23	9,838,724.18	210,023.34	Large and positive	Possible
Category B	AALRRPB	Cumulative AALRRPB	Net AAU Balance	Result	AAU Effect
AAU Stage 1	4,722,587.70	4,722,587.70	11,656,040.01	Large and positive	Possible
AAU Stage 2	885,485.10	5,608,072.80	5,175,804.38	Large and positive	Possible
AAU Stage 3	236,129.37	5,844,202.17	123,087.89	Large and positive	Possible
AAU Stage 4	59,032.34	5,903,234.51	5,736,637.66	Large and positive	Possible
Category C	AALRRPB	Cumulative AALRRPB	Net AAU Balance	Result	AAU Effect
AAU Stage 1	3,148,391.80	3,148,391.80	2,888,178.28	Large and positive	Possible
AAU Stage 2	590,323.40	3,738,715.20	74,255.64	Large and positive	Possible
AAU Stage 3	157,419.58	3,896,134.78	3,785,070.22	Large and positive	Possible
AAU Stage 4	39,354.89	3,935,489.67	84,009.33	Large and positive	Possible

Table 20 above shows the results for Automated Actuarial Underwriting models for Policyholder Category A, Policyholder Category B and Policyholder Category C. From these three models each model has four stage AAU process the first being stage 1 AAU see the equation 13, second Stage 2 AAU see equation 14, third Stage 3 AAU see equation 15, fourth Stage 4 AAU see equation 16 and finally Stage 5 AAU see equation 17. There is no fifth stage in the above models since all the total expenses distributed in the respective presented policyholder categories have been underwritten from stage 1 leaving off a large chunk of positive and large net AALRRPB cash flow balances and hence bringing

us to the fact that in all these stages Automated Actuarial Underwriting of all claims, expenses and outgo has been quite effective within the first 100 years of implementing this model. Taking a closer look at these three models the quantum of AALRRP balances is greatest in Stages 1, followed by Stage 2, then Stage 3 and finally Stage 4. The presence of large and positive AAU net present values across the three policyholder categories shows that a large chunk of ACAALRRPB have been spared for future growth and catastrophic reserving and thus reducing both the incidence, prevalence as well as the severity of reinsurance in its diverse forms. Large and positive AAU net present values are due to more comprehensive payments being made and reduced frequency and severity of comprehensive claims. This is a clear indication of cash generative and profitability business and hence this reflects that the insurance company has the capability to boost its financial performance, organizational productivity, reduced solvency margin in line with IFRS17 guidelines. As a result, automation of microfinance and car insurance services steers up the insurer's profitability and reduces the solvency risk as far as actuarial underwriting is concerned.

We also validated our model by forecasting the net present values for Automated Actuarial Underwriting Model using the earliest possible time, which is the first 10 years, as a matter to test our model and see if the model also works well within this period of time. The results are shown below.

In addition to that, **Figures 14–16** display a visual presentation of Automated Actuarial Underwriting for the first 10 years taken from **Table 21**. Furthermore, these figures show a positive trend in the predicted NPV Automated Actuarial Underwriting.

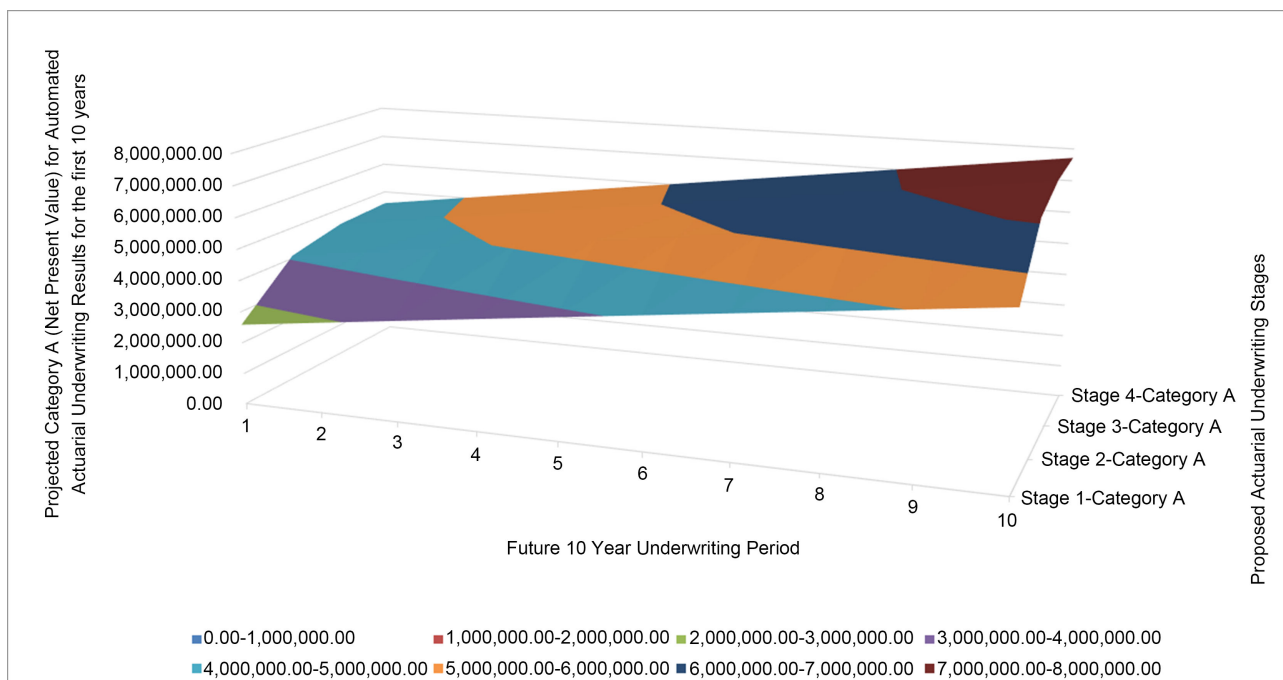


Figure 14. Predicted NPV for AAU for Category A.

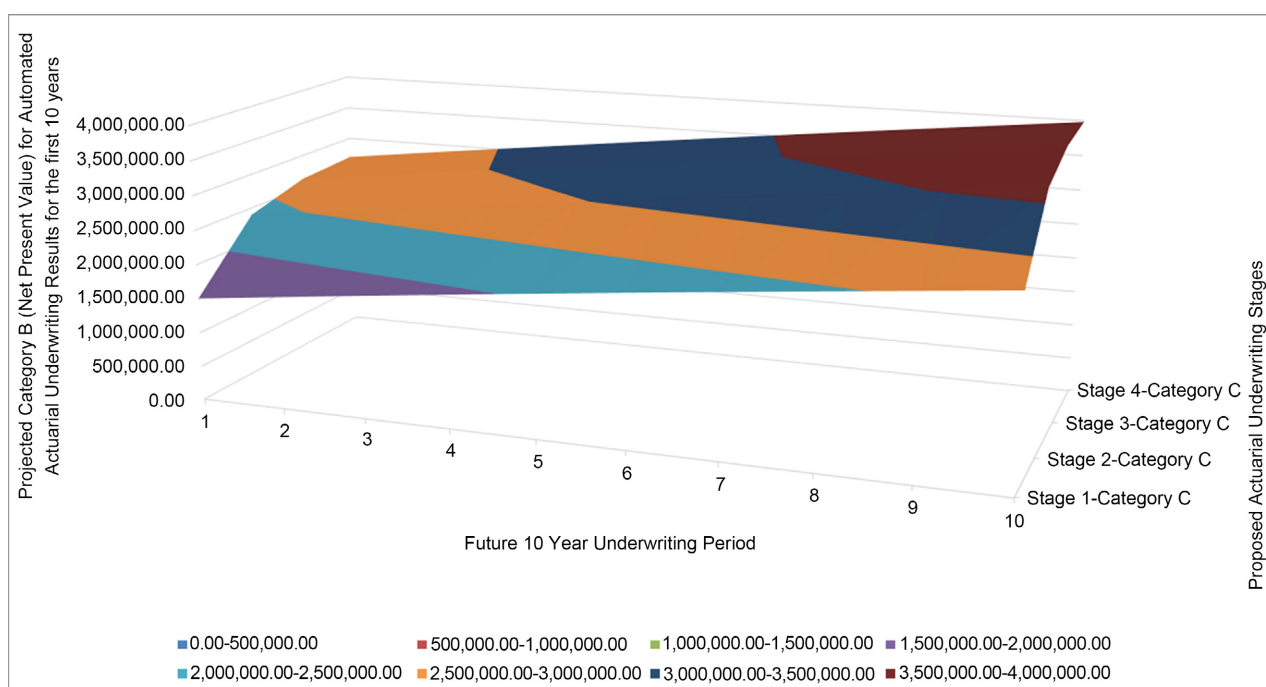


Figure 15. Predicted NPV for AAU for Category B.

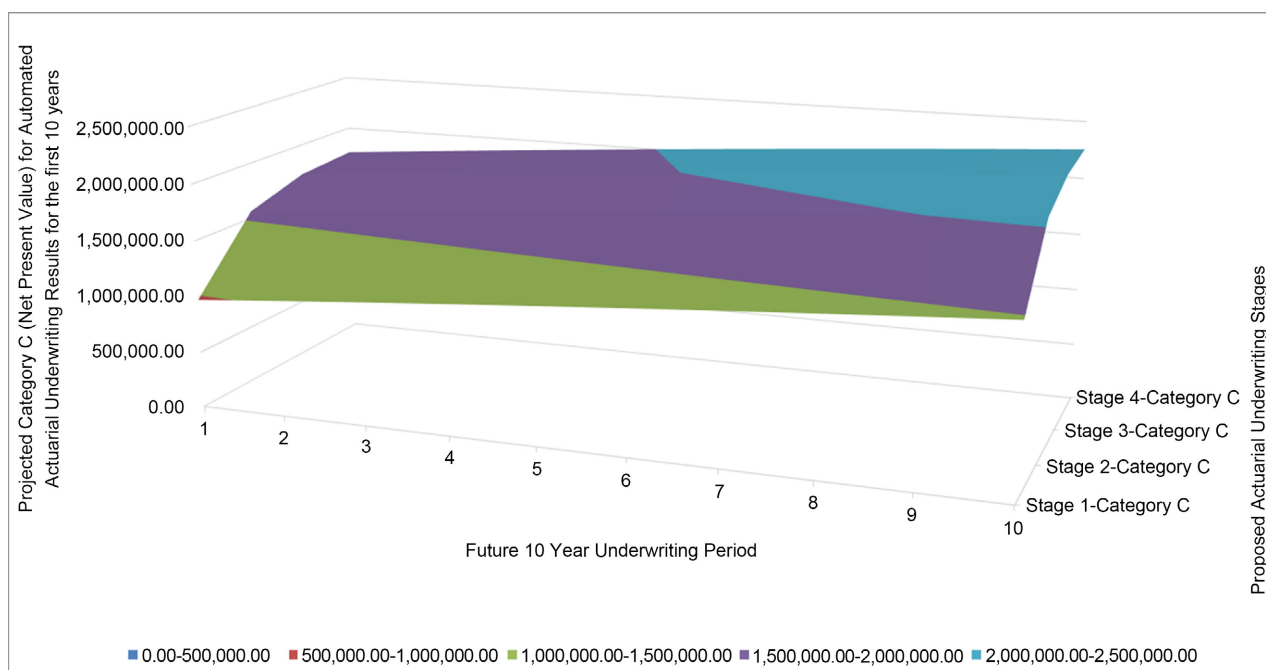


Figure 16. Predicted NPV for AAU for Category C.

Balances that indicate growth in the net ACAALRRPB cash flows after the underwriting of all expenses and outgo. Our proposed Automated Actuarial Underwriting model in four stages has immediately given large positive quantum of net AAU from the negative NPVs for AAU in Years 1 and 2 across the three policyholder categories. This has been possible just at AAU Stage 2 leaving off other stages with enormous amounts of money which the insurance company

can readily plough back into its future projects.

Table 21. Model validation by forecasting the NPV for automated actuarial underwriting model for the first 10-year underwriting period.

Policyholder Category	AALRRPB	Cumulative AALRRPB	Projected (Net Present Value) For Automated Actuarial Underwriting Results for the first 10 years									
			1	2	3	4	5	6	7	8	9	10
Stage 1-Category A	7,870,979.50	7,870,979.50	2,584,074.91	2,896,720.97	3,207,476.95	3,516,538.15	3,824,096.69	4,130,341.70	4,435,459.47	4,739,633.54	5,043,044.88	5,345,871.97
Stage 2-Category A	9,346,788.00	9,346,788.00	4,087,809.92	4,428,910.93	4,768,660.32	5,107,263.56	5,444,923.16	5,781,838.84	6,118,207.65	6,454,224.11	6,790,080.39	7,125,966.38
Stage 3-Category A	9,740,336.95	9,740,336.95	4,488,805.94	4,837,494.94	5,184,975.90	5,531,457.02	5,877,143.57	6,222,238.10	6,566,940.51	6,911,448.29	7,255,956.55	7,600,658.24
Stage 4-Category A	9,838,724.18	9,838,724.18	4,589,054.93	4,939,640.93	5,289,054.79	5,637,505.37	5,985,198.66	6,332,337.90	6,679,123.72	7,025,754.31	7,372,425.57	7,719,331.19
Stage 1-Category B	4,722,587.70	4,722,587.70	1,504,616.83	1,645,086.66	1,783,104.47	1,918,758.56	2,052,134.75	2,183,316.48	2,312,384.86	2,439,418.74	2,564,494.81	2,687,687.62
Stage 2-Category B	885,485.10	5,608,072.80	2,398,265.06	2,546,973.28	2,693,305.43	2,837,350.50	2,979,195.03	3,118,923.17	3,256,616.74	3,392,355.33	3,526,216.36	3,658,275.11
Stage 3-Category B	236,129.37	5,844,202.17	2,636,571.27	2,787,476.39	2,936,025.69	3,082,308.36	3,226,411.12	3,368,418.29	3,508,411.92	3,646,471.77	3,782,675.45	3,917,098.46
Stage 4-Category B	59,032.34	5,903,234.51	2,696,147.81	2,847,602.16	2,996,705.75	3,143,547.82	3,288,215.13	3,430,792.07	3,571,360.70	3,710,000.87	3,846,790.21	3,981,804.28
Stage 1-Category C	3,148,391.80	3,148,391.80	972,525.80	1,035,353.45	1,096,277.12	1,155,353.18	1,212,636.35	1,268,179.70	1,322,034.76	1,374,251.51	1,424,878.47	1,473,962.71
Stage 2-Category C	590,323.40	3,738,715.20	1,562,562.78	1,625,104.14	1,685,741.66	1,744,531.71	1,801,529.01	1,856,786.63	1,910,356.10	1,962,287.40	2,012,629.04	2,061,428.10
Stage 3-Category C	157,419.58	3,896,134.78	1,719,905.98	1,782,370.99	1,842,932.21	1,901,645.99	1,958,567.06	2,013,748.49	2,067,241.79	2,119,096.97	2,169,362.53	2,218,085.54
Stage 4-Category C	39,354.89	3,935,489.67	1,759,241.77	1,821,687.70	1,882,229.84	1,940,924.56	1,997,826.57	2,052,988.95	2,106,463.21	2,158,299.36	2,208,545.90	2,257,249.90

As a result, our model works well in both scenarios such as the longest possible time where $n = 100$ years and the earliest possible time where we validated our model with the first 10 years of operation. In both scenarios our model works well since there are large sums of net AAU model balances set aside for the future use by the insurance company.

5.8. Innovations for the Artificial Intelligence Based Automated Actuarial Pricing and Underwriting Model for the General Insurance Sector

This section shows the remarks concerning our study.

5.8.1. Risk Mitigation and Risk Minimization Model for Automated Actuarial Underwriting

Full implementation of our model mitigates and reduces the riskiness of failure by an insurance company since the various types of risks such as liquidity risks, currency risks, group risks just to mention a few are most likely to be reduced greatly. For example, the large and positive net AAU cash flows cash generated from the automated microfinance and car insurance services help a lot to diminish the liquidity risk. Moreover, these net large and positive AAU cash flows are then channeled towards underwriting of comprehensive claims, expenses and related outgo [34]-[36].

5.8.2. Adherence to IFRS 17 Regulations

Our model does support all the accounting concepts such as the money measurement concept, going concern concept, realization concept, prudence concept and accruals concept just to mention a few. The simultaneous and autonomous automation of microfinance services and car insurance services using Artificial Intelligence based methods ensures that the insurance 17 and 16. In a similar fashion our study conforms to IFRS17 standards [37] and [38].

5.8.3. Asset Liability Modelling (ALM)

Asset-liability management attempts to find the optimal investment strategy under uncertainty in both the asset and liability streams. The developed Automated Actuarial Pricing and Underwriting Model results in more comprehensive payments in form of amounts of money invested and the current premiums (thus boosting and increasing the asset base of the insurer) paid whilst decreasing the comprehensive claims, expenses and outgo (reducing the liability base of the insurer) [39]-[42].

5.8.4. No Broker and Other Intermediaries

Our model does not require the use of broker and other related insurance intermediaries since it deals directly with the policyholders in their respective categories and hence the broking costs are brought to zero. Furthermore, the policyholder is capable of knowing his/her premiums and also the amounts of money invested and requested from the AI-based model. Eventually, this reflects effectiveness and efficiency of our model [43]-[45].

5.8.5. Reduced Ceding Reinsurance

The setting up of the AALRRPB (IBNYR + IBNYPP): see **Table A3**, AALRRPB (RBNYS + PBNYPP): see **Table A4**, AALRRPB (REOPENED + REOPP): see **Table A5** and AALRRPB (REINSURANCE + REINSPP): see **Table A6** is meant to reduce drastically all forms of reinsurance by the insurer which are facultative reinsurance, excess of loss reinsurance, proportional reinsurance and catastrophic reinsurance. The augmentation of microfinance services and car insurance services on the same platform using Artificial Intelligence is a substantive step towards generating and guaranteeing large pool of funds to the insurer and thus reducing the quantum of reinsurance almost to zero since the insurance company is then capable of underwriting all its future claims with minimum or no reinsurance [46] and [47].

5.8.6. Reduced Underwriting Expenses and Related Outgo

The presence of the no-claims bonus rate system (see **Table 7**) across all the possible policyholder categories is a step towards these policyholders in their respective categories to reduce both the severity and frequency of comprehensive claims. Such a step turns to bring down the Underwriting expenses and outgo as one wishes to maximize his/her final bonus rate in order to earn a greater return [48]-[50].

6. Discussion

The primary objective of this paper was to introduce and develop an innovative framework for actuarial pricing and underwriting within the general insurance sector, utilizing artificial intelligence (AI) methodologies. The proposed model, referred to as the Automated Actuarial Loss Reserving and Pricing (AALRP) model, seeks to enhance the accuracy and efficiency of premium pricing and loss reserving through the integration of machine learning algorithms. By introducing new terminologies and categorizations, this paper aims to address current gaps in traditional actuarial practices. In response to the dynamic nature of insurance and microfinance services, we introduced four new types of premium-related concepts: IBNYPP (Incurred But Not Yet Paid Premium), PBNYSPP (Paid But Not Yet Settled Premium Payment), REOPP (Reopened Premium Payment), and REINSPP (Reinsured Premium Payment). These terminologies were designed to capture various stages of premium payment processes and their associated uncertainties. For instance, IBNYPP addresses premiums incurred but not yet paid, while REOPP and REINSPP handle premiums that require reactivation or reinsurance due to prolonged non-payment or large, overdue amounts. Similarly, we expanded the traditional actuarial loss reserves framework by adding REOPENED (Reopened Reserve) and REINSURANCE (Reinsurance Reserve) to the existing categories of IBNYR (Incurred But Not Yet Reported) and RBNYS (Reported But Not Yet Settled). These additions aim to better reflect the diverse scenarios encountered in modern insurance practices, such as reopening previously closed claims and addressing catastrophic losses through reinsurance. The development of the AALRP model balances is pivotal to this paper. We categorized policyholders into four distinct groups: those with both car insurance and microfinance policies (Category A), those with only microfinance policies (Category B), those with only car insurance policies (Category C), and those without any active policies (Category D). These categorizations allow for a tailored approach to premium pricing and loss reserving, considering the specific needs and risk profiles of each group. The core of the AALRP model involves the interaction between various types of reserves and premium payments. For example, the model calculates balances for different combinations of reserves and premium types (e.g., IBNYR + IBNYPP, RBNYS + PBNYSPP), providing a comprehensive view of the insurer's financial standing concerning different policyholder categories. The Automated Actuarial Underwriting (AAU) model, derived from the AALRP balances, was tested over a range of time periods, from 10 to 100 years, to determine its viability under varying conditions. The model's evaluation criteria were based on the net present value of the AALRP balances, considering expenses and outgoings over time. This iterative process ensures that the AAU model can adapt to changes in the economic environment, policyholder behavior, and other relevant factors. A significant innovation in this study is the application of eight machine learning algorithms—GLM, GAM, RPART, RANGER, XGB, LAR, SVM, and ANN—to develop and validate the

AALRP model. The use of Random Forest (RANGER) emerged as the most effective approach, yielding the highest Total Aggregate Comprehensive Automated Actuarial Loss Reserve Risk Pricing Balance (ACAALRRPB). This finding underscores the potential of AI and machine learning to revolutionize actuarial practices by providing more precise and dynamic pricing and reserving mechanisms. The proposed models and terminologies represent a significant step forward in the automation of actuarial tasks within the insurance sector. By leveraging AI, insurers can achieve greater accuracy in pricing and reserving, ultimately leading to more stable and fair premium structures. Future research should focus on refining these models through real-world testing and incorporating additional variables that may impact premium payments and loss reserves, such as economic indicators and demographic shifts. Moreover, expanding the application of these models to other types of insurance and financial services could provide further insights and improvements. The integration of more advanced AI techniques, such as deep learning, could also enhance the predictive power and adaptability of the models.

In conclusion, this paper presents a robust framework for the automation of actuarial pricing and underwriting, offering significant potential for improving the efficiency and accuracy of insurance operations. The introduction of new terminologies and the application of machine learning mark a progressive step towards modernizing actuarial science and meeting the evolving needs of the insurance industry.

7. Conclusions

This study introduced and proposed a comprehensive framework for automating actuarial pricing and underwriting models within the general insurance sector, particularly focusing on car insurance and microfinance services. The novel terminologies and models we developed, including IBNYPP (Incurred But Not Yet Paid Premium), PBNYSPP (Paid But Not Yet Settled Premium Payment), REOPP (Reopened Premium Payment), and REINSPP (Reinsured Premium Payment), address the complexities and variabilities in premium payment structures. These terminologies are essential for accurately categorizing and managing different premium statuses, thereby enhancing the precision of actuarial reserving and risk assessment. Our work also expanded on existing actuarial loss reserves, such as IBNYR (Incurred But Not Yet Reported) and RBNYS (Reported But Not Yet Settled), by introducing REOPENED (Reopened Reserve) and REINSURANCE (Reinsurance Reserve). These additions cater to reopened claims and catastrophic losses, ensuring a more robust and comprehensive reserve management system. This enhanced framework allows insurers to better manage their liabilities and improve financial stability. The development of the Automated Actuarial Loss Reserving and Pricing (AALRP) model balances provides a systematic approach to integrating various premium and reserve types. The AALRP model supports the automation of underwriting processes through

five distinct stages, leveraging the Final Bonus Rate (FBR) and prevailing economic conditions to assess underwriting viability over different time horizons. This multi-stage model ensures that the underwriting process remains dynamic and responsive to changing financial landscapes. We employed eight machine learning algorithms to evaluate the effectiveness of our proposed models, with Random Forest (RANGER) demonstrating superior performance. The RANGER model achieved the highest Total Aggregate Comprehensive Automated Actuarial Loss Reserve Risk Pricing Balance (ACAALRRPB), indicating its efficacy in predicting and managing actuarial risks. This finding underscores the potential of machine learning techniques in enhancing actuarial practices, offering a data-driven approach to risk management and decision-making.

In conclusion, our research presents a novel and practical approach to automating actuarial pricing and underwriting processes in the insurance sector. By integrating advanced actuarial concepts with machine learning algorithms, we provide a framework that can significantly improve the accuracy, efficiency, and responsiveness of insurance operations. Future work should focus on refining these models, exploring additional machine-learning techniques, and validating the framework across different insurance products and markets

Data Availability

The data was simulated in R and kept for ethical reasons.

Acknowledgements

Special thanks go to members of staff at the University of Zimbabwe through the Department of Mathematics & Computational Sciences for both academic, social and moral support.

Conflicts of Interest

There were not any conflicts of interest.

References

- [1] Mihaela, D. (2015) A Review of Theoretical Concepts and Empirical Literature of Non-Life Insurance Pricing. *Procedia Economics and Finance*, **20**, 157-162. [https://doi.org/10.1016/S2212-5671\(15\)00060-X](https://doi.org/10.1016/S2212-5671(15)00060-X)
- [2] Anderson, J.D., Bolton, C.G., Callan, G.L., Cross, M., Howard, S.K., Mitchell, G.R.J., Murphy, K.P., Rakow, J.C., Stirling, P.A. and Welsh, G.E. (2007) General Insurance Premium Rating—The Way Forward. Summary of the Recommendations of the General Insurance Premium Rating Working Party (GRIP). *British Actuarial Journal*, **13**, 637-644. <https://doi.org/10.1017/S1357321700001586>
- [3] Thomas, A.M. and Thomas, R.G. (1997) The Right to Underwrite? An Actuarial Perspective with a Difference. *Journal of Actuarial Practice*, **5**, 125-146.
- [4] Mihaela, D. (2015) Auto Insurance Premium Calculation Using Generalized Linear Models. *Procedia Economics and Finance*, **20**, 147-156. [https://doi.org/10.1016/S2212-5671\(15\)00059-3](https://doi.org/10.1016/S2212-5671(15)00059-3)

- [5] de Jong, C.H. (2021) Risk Classification and the Balance of Information in Insurance; an Alternative Interpretation of the Evidence. *Risk Management and Insurance Review*, **24**, 445-461. <https://doi.org/10.1111/rmir.12198>
- [6] Nicholas, Y., Raymond, L., Jyeh, O.M. and Liew, J.Y. (2019) Literature Review: Artificial Intelligence and Its Use in Actuarial Work. Actuarial Innovation & Technology, Society of Actuaries.
- [7] Tuininga, F. (2022) A Machine Learning Approach for Modeling Frequency and Severity. Master's Thesis, University of Twente.
- [8] Chaturvedi, K.K. and Singh, V.B. (2012) Determining Bug Severity Using Machine Learning Techniques. 2012 *CSI Sixth International Conference on Software Engineering (CONSEG)*. Indore, 5-7 September 2012, 1-6. <https://doi.org/10.1109/CONSEG.2012.6349519>
- [9] Breiman, L. (2001) Random Forests. *Machine Learning*, **45**, 5-32. <https://doi.org/10.1023/A:1010933404324>
- [10] Taylor, G.C. (2000) Loss Reserving: Past, Present and Future. *North American Actuarial Journal*, **4**, 1-17.
- [11] Meyers, G. (2015) Stochastic Loss Reserving Using Bayesian MCMC Models. Casualty Actuarial Society, Arlington.
- [12] Duncan, A., McPhail, P. and Siegel, A. (2021) Data Science for Actuaries. *British Actuarial Journal*, **26**.
- [13] Clark, D.R. (2005) LDF Curve-Fitting and Stochastic Loss Reserving: A Maximum Likelihood Approach. Casualty Actuarial Society Forum, 41-91.
- [14] Goldburd, M., Khare, A. and Tevet, D. (2016) Generalized Linear Models for Insurance Rating. Casualty Actuarial Society, Arlington.
- [15] Wuethrich, M.V. and Merz, M. (2020) Stochastic Claims Reserving Methods in Insurance. Wiley.
- [16] Antonio, K. and Plat, R. (2014) Micro-Level Stochastic Loss Reserving for General Insurance. *Scandinavian Actuarial Journal*, **2014**, 649-669. <https://doi.org/10.1080/03461238.2012.755938>
- [17] Haberman, S. and Renshaw, A.E. (1996) Generalized Linear Models and Actuarial Science. *Journal of the Royal Statistical Society. Series D*, **45**, 407-436. <https://doi.org/10.2307/2988543>
- [18] Friedman, J.H. (2001) Greedy Function Approximation: A Gradient Boosting Machine. *Annals of Statistics*, **29**, 1189-1232. <https://doi.org/10.2307/2988543>
- [19] Antonio, K., Bard, B. and Ohlsson, E. (2010) A Neural Network Approach to Loss Reserving. *ASTIN Bulletin: The Journal of the IAA*, **40**, 449-477.
- [20] Hastie, T., Tibshirani, R. and Friedman, J. (2009) The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer Science & Business Media.
- [21] Kuhn, M. and Johnson, K. (2013) Applied Predictive Modeling. Springer Science & Business Media. <https://doi.org/10.1007/978-1-4614-6849-3>
- [22] Chen, T. and Guestrin, C. (2016) XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, 13-17 August 2016, 785-794. <https://doi.org/10.1145/2939672.2939785>
- [23] Styruud, L. (2017) Risk Premium Prediction of Car Damage Insurance Using Artificial Neural Networks and Generalized Linear Models.
- [24] Verbelen, R. (2017) Data Analytics for Insurance Loss Modelling, Telematics Pric-

- ing and Claims Reserving. PhD Thesis, Ku Leuven.
- [25] El Kassimi, F. and Zahi, J. (2021) Non-Life Insurance Ratemaking Techniques : A Literature Review of the Classic Methods. *International Journal of Accounting, Finance, Auditing, Management and Economics*, **2**, 344-361.
 - [26] Jamal, S., Canto, S., Fernwood, R., Giancaterino, C., Lo-Renzoinvernizzi, M., Korzhynska, T., Martin, Z. and Shen, H. (2018) Machine Learning & Traditional Methods Synergy in Non-Life Reserving. Report of the ASTIN Working Party of the International Actuarial Association.
<https://www.actuaries.org/IAA/Documents/ASTIN/ASTINMLTMS%20ReportSIA MAL.pdf>
 - [27] Burri, R.D., Burri, R., Bojja, R.R. and Buruga, S.R. (2019) Insurance Claim Analysis Using Machine Learning Algorithms. *International Journal of Innovative Technology and Exploring Engineering*, **8**, 577-582.
<https://doi.org/10.35940/ijitee.F1118.0486S419>
 - [28] Filzmoser, P., Hron, K. and Templ, M. (2018) Correlation Analysis. In: Filzmoser, P., Hron, K. and Templ, M., Eds., *Applied Compositional Data Analysis*, Springer, 149-162. https://doi.org/10.1007/978-3-319-96422-5_8
 - [29] Cohen, J., Cohen, P., West, S.G. and Aiken, L.S. (2013) Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences. Routledge.
<https://doi.org/10.4324/9780203774441>
 - [30] Comrey, A.L. and Lee, H.B. (2013) A First Course in Factor Analysis. Psychology Press. <https://doi.org/10.4324/9781315827506>
 - [31] Jolliffe, I.T. (2002) Principal Component Analysis for Special Types of Data. In: Jolliffe, I.T., Ed., *Principal Component Analysis*, Springer, 338-372.
https://doi.org/10.1007/0-387-22440-8_13
 - [32] Williams, B., Onsmann, A. and Brown, T. (2010) Exploratory Factor Analysis: A Five-Step Guide for Novices. *Paramedicine*, **8**, 1-13.
<https://doi.org/10.33151/ajp.8.3.93>
 - [33] Oner, C. (2012) Inflation: Prices on the Rise. International Monetary Fund.
 - [34] Mkrtychev, S. and Enik, O. (2019) Automated Underwriting Control in a Regional Insurance Company. *Proceedings of the International Scientific Conference "Far East Con" (ISC FEC 2018)*, 291-293. <https://doi.org/10.2991/iscfec-18.2019.66>
 - [35] Boudreault, M. and Renaud, J.-F. (2019) Actuarial Finance: Derivatives, Quantitative Models and Risk Management. John Wiley & Sons.
<https://doi.org/10.1002/9781119526438>
 - [36] Francis, G. (2019) Enterprise Risk Management (ERM): Key Risks, Responses and Applications. 2019 *Enterprise Risk Management Symposium*, 2-3.
 - [37] Windsor, P., Yong, J. and Bell, M.C.-T. (2020) Accounting Standards and Insurer Solvency Assessment. International Monetary Fund.
<https://doi.org/10.2139/ssrn.3721187>
 - [38] Chiantera, M. (2023) Data Quality and Data Governance in Insurance Corporations Doctoral Dissertation, Politecnico di Torino.
 - [39] Grebeck, M. and Rachev, S. (2005) Stochastic Programming Methods in Asset-Liability Management. *Investment Management and Financial Innovations*, **2**, 82-90.
 - [40] Diehl, M. (2023) Large Insurance Portfolios-Simulation, Strategy Decisions, Asset-Liability Management. Doctoral Dissertation, Universitt Kaiserslautern-Landau.
 - [41] Chaim, R.M. (2010) Dynamic Asset and Liability Management. Pension Fund Risk

Management, 129.

- [42] Luckner, W.R., Abbott, M.C., Backus, J.E., Sergio Benedetti, F.S.A., Fcia, M., Bergman, D., Cox, S.H., SholomFeldblum, F.S.A., Gilbert, C.L., Liu, X.L. and Lui, V.Y. (2003) SOA Professional Actuarial Specialty Guide: Asset-Liability Management. Society of Actuaries.
- [43] Cummins, J.D. and Doherty, N.A. (2006) The Economics of Insurance Intermediaries. *Journal of Risk and Insurance*, **73**, 359-396.
<https://doi.org/10.1111/j.1539-6975.2006.00180.x>
- [44] Clemente, G.P. and Marano, P. (2020) The Broker Model for Peer-to-Peer Insurance: An Analysis of Its Value. *The Geneva Papers on Risk and Insurance-Issues and Practice*, **45**, 457-481. <https://doi.org/10.1057/s41288-020-00165-8>
- [45] Doherty, N.A. and Muermann, A. (2005) Insuring the Uninsurable: Brokers and Incomplete Insurance Contracts (No. 2005/24). CFS Working Paper.
- [46] Obalola, M.A. and Ukpong, M.S. (2022) Reinsurance and the Determinants of the Ceding Decision of Life Insurance Companies in Nigeria: An Empirical Analysis. *Journal of Corporate Governance, Insurance, and Risk Management*, **9**, 169-189.
<https://doi.org/10.51410/jcgirm.9.1.11>
- [47] Frees, E.W., Shi, P. and Valdez, E.A. (2009) Actuarial Applications of a Hierarchical Insurance Claims Model. *ASTIN Bulletin: The Journal of the IAA*, **39**, 165-197.
<https://doi.org/10.2143/AST.39.1.2038061>
- [48] Darcy, S.P. and Gorvett, R.W. (1998) A Comparison of Property/Casualty Insurance Financial Pricing Models. *Proceedings of the Casualty Actuarial Society*, **85**, 1-88.
- [49] Mourdoukoutas, F., Pantelous, A.A. and Taylor, G. (2023) Competitive Insurance Pricing Strategies for Multiple Lines of Business: A Game Theoretic Approach. *SSRN Electronic Journal*, 1-46.
- [50] Abbott, W.M., Clarke, T.G. and Treen, W.R. (1981) Some Financial Aspects of a General Insurance Company. *Journal of the Institute of Actuaries*, **108**, 119-209.
<https://doi.org/10.1017/S0020268100040725>

Appendix

Appendix A1. Machine Learning Algorithms, Associated R Packages and Hyper-Parameters

Table A1. Machine learning algorithms, associated R packages and Hyper-parameters.

Machine learning Algorithm	R packages used	Hyperparameters
Generalized Linear Models (GLM)	glm2	family distribution: Gaussian, link function: Identity
Generalized Additive Models (GAM)	gam	family distribution: Gaussian, link function: Identity
Regression Trees (RPART)	rpart	No hyperparameters used
Random Forest (RANGER)	ranger	number of trees: 500, Mtry: 8, Target node size: 5
Extreme Gradient Boosting (XGB)	xgboost	xgboost maximum depth: 3, number of rounds: 100
Least Angle Regression (LAR)	caret	Method: lars
Support Vector Machines (SVMM)	e10171	SVM-Type: eps-regression, SVM-Kernel: radial, cost: 1
Artificial Neural Network (ANN)	nnet	Size: 2, decay: 5e-4, maximum iterations: 200

Appendix A2. AALRRPB (Total Reserves + Total Premiums)

Table A2. Automated actuarial loss reserving risk pricing balances (Total Reserves + Total Premiums).

Automated Actuarial Loss Reserving Premium Pricing Balances (AALRRPB)				
ML Model	IBNYR + IBNYPP	RBNYS + PBNYPP	REOPENED + REOPP	REINSURANCE + REINSPP
GLM	8,059,676.00	1,511,189.80	402,983.80	100,745.95
GAM	8,054,664.00	1,510,249.70	402,733.20	100,683.31
RPART	8,058,505.00	1,510,969.20	402,925.20	100,731.30
RANGER	15,741,959.00	2,951,617.00	787,097.90	196,774.45
XGB	8,059,515.00	1,511,159.00	402,975.70	100,743.93
LAR	8,060,230.00	1,511,292.80	403,011.50	100,752.87
SVM	4,148,592.00	777,860.90	207,429.58	51,857.39
ANN	4,011,733.00	752,200.00	200,586.67	50,146.67

Appendix A3. AALRRPB (IBNYR + IBNYPP)

Table A3. Automated actuarial loss reserving risk pricing balances (IBNYR + IBNYPP).

Automated Actuarial Loss Reserving Premium Pricing Balances (AALRRPB)				
ML Model	IBNYR-A + IBNYPP-A	IBNYR-B + IBNYPP-B	IBNYR-C + IBNYPP-C	IBNYR-D + IBNYPP-D
GLM	4,029,838.00	2,417,902.80	1,611,935.20	0.00
GAM	4,027,332.00	2,416,399.20	1,610,932.80	0.00
RPART	4,029,252.50	2,417,551.50	1,611,701.00	0.00
RANGER	7,870,979.50	4,722,587.70	3,148,391.80	0.00
XGB	4,029,757.50	2,417,854.50	1,611,903.00	0.00
LAR	4,030,115.00	2,418,069.00	1,612,046.00	0.00
SVM	2,074,296.00	1,244,577.60	829,718.40	0.00
ANN	2,005,866.50	1,203,519.90	802,346.60	0.00

Appendix A4. AALRRPB (RBNYS + PBNYSP)

Table A4. Automated actuarial loss reserving premium pricing balances (RBNYS + PBNYSP).

Automated Actuarial Loss Reserving Premium Pricing Balances (AALRRPB)				
ML Model	RBNYS-A + PBNYSP-A	RBNYS-B + PBNYSP-B	RBNYS-C + PBNYSP-C	RBNYS-D + PBNYSP-D
GLM	755,594.90	453,356.94	302,237.96	0.00
GAM	755,124.85	453,074.91	302,049.94	0.00
RPART	755,484.60	453,290.76	302,193.84	0.00
RANGER	1,475,808.50	885,485.10	590,323.40	0.00
XGB	755,579.50	453,347.70	302,231.80	0.00
LAR	755,646.40	453,387.84	302,258.56	0.00
SVM	388,930.45	233,358.27	155,572.18	0.00
ANN	376,100.00	225,660.00	150,440.00	0.00

Appendix A5. AALRRPB (ROPNED + REOPP)

Table A5. Automated actuarial loss reserving premium pricing balances (ROPNED + REOPP).

Automated Actuarial Loss Reserving Premium Pricing Balances (AALRRPB)				
ML Model	ROPNED-A + REOPP-A	ROPNED-B + REOPP-B	ROPNED-C + REOPP-C	ROPNED-D + REOPP-D
GLM	201,491.90	120,895.14	80,596.76	0.00
GAM	201,366.60	120,819.96	80,546.64	0.00
RPART	201,462.60	120,877.56	80,585.04	0.00
RANGER	393,548.95	236,129.37	157,419.58	0.00
XGB	201,487.85	120,892.71	80,595.14	0.00
LAR	201,505.75	120,903.45	80,602.30	0.00
SVM	103,714.79	62,228.87	41,485.92	0.00
ANN	100,293.34	60,176.00	40,117.33	0.00

Appendix A6. AALRRPB (REINSURANCE + REINSPP)

Table A6. Automated actuarial loss reserving premium pricing balances (REINSURANCE + REINSPP).

Automated Actuarial Loss Reserving Premium Pricing Balances (AALRRPB)				
ML Model	REINSURANCE- A + REINSPP-A	REINSURANCE- B + REINSPP-B	REINSURANCE- C + REINSPP-C	REINSURANCE- D + REINSPP-D
GLM	50,372.98	30,223.79	20,149.19	0.00
GAM	50,341.66	30,204.99	20,136.66	0.00
RPART	50,365.65	30,219.39	20,146.26	0.00
RANGER	98,387.23	59,032.34	39,354.89	0.00
XGB	50,371.97	30,223.18	20,148.79	0.00
LAR	50,376.44	30,225.86	20,150.57	0.00
SVM	25,928.70	15,557.22	10,371.48	0.00
ANN	25,073.34	15,044.00	10,029.33	0.00

Appendix A7. Comprehensive Automated Actuarial Loss Reserving Risk Pricing Balances (CAALRRPB)

Table A7. Comprehensive automated actuarial loss reserving risk pricing balances.

Comprehensive Automated Actuarial Loss Reserving Premium Pricing Model				
ML Model	CAALR-A + CAARP-A	CAALR-B + CAARP-B	CAALR-C + CAARP-C	CAALR-D + CAARP-D
GLM	5,037,297.78	3,022,378.67	2,014,919.11	0.00
GAM	5,034,165.11	3,020,499.06	2,013,666.04	0.00
RPART	5,036,565.35	3,021,939.21	2,014,626.14	0.00
RANGER	9,838,724.18	5,903,234.51	3,935,489.67	0.00
XGB	5,037,196.82	3,022,318.09	2,014,878.73	0.00
LAR	5,037,643.59	3,022,586.15	2,015,057.43	0.00
ELM	2,592,869.94	1,555,721.96	1,037,147.97	0.00
ANN	2,507,333.17	1,504,399.90	1,002,933.27	0.00

Appendix A8. Aggregate Comprehensive Automated Actuarial Loss Reserving Risk Pricing Balances (ACAALRRPB)

Table A8. Aggregate comprehensive automated actuarial loss reserving risk pricing balances.

Aggregate Comprehensive Automated Actuarial Loss Reserving Risk Pricing Balances			
ML Model	ACAALR	ACAARP	ACAALRRPB
GLM	3,200,985.05	6,873,610.5000	10,074,595.55
GAM	3,195,537.58	6,872,792.6300	10,068,330.21
RPART	3,199,968.58	6,873,162.1200	10,073,130.70
RANGER	12,804,873.60	6,872,574.7500	19,677,448.35
XGB	3,200,773.63	6,873,620.0000	10,074,393.63
LAR	3,202,718.68	6,872,568.4900	10,075,287.17
SVM	1,674,997.35	3,510,742.5200	5,185,739.87
ANN	1,496,441.39	3,518,224.9500	5,014,666.34