Scientific Research Publishing

Automated Actuarial Data Analytics-Based Inflation Adjusted Frequency Severity Loss Reserving Model

Brighton Mahohoho

Department of Mathematics & Computational Sciences, University of Zimbabwe, Harare, Zimbabwe
Email: brightonmahohoho.07@gmail.com

Abstract

In this paper, the Automated Actuarial Loss Reserving Model is developed and extended using machine learning. The traditional actuarial reserving techniques are no longer compatible with the increase in technological advancement currently at hand. As a result, the development of the alternative Artificial Intelligence Based Automated Actuarial Loss Reserving Methodology which captures diverse risk profiles for various policyholders through augmenting the Micro Finance services, Auto Insurance Services and Both Services lines of business on the same platform through the computation of the Comprehensive Automated Actuarial Loss Reserves (CAALR) has been implemented in this paper. The introduction of the four further types of actuarial loss reserves to those existing in the actuarial literature seems to significantly reduce lapse rates, reduce the reinsurance costs as well as expenses and outgo. As a matter of consequence, this helps to bring together a combination of new and existing policyholders in the insurance company. The frequency severity models have been extended in this paper using ten machine learning algorithms which ultimately leads to the derivation of the proposed machine learning-based actuarial loss reserving model which remarkably performed well when compared to the traditional chain ladder actuarial reserving method using simulated data.

Keywords

Automated Actuarial Loss Reserves, Artificial Intelligence, Micro Finance Loss Reserving, Auto Insurance Loss Reserving, Both Services Loss Reserving

1. Introduction

Artificial Intelligence (AI) has become an increasingly important tool for actuarial loss reserving and hence by integrating cognitive computing capabilities into their actuarial processes, the reserving actuarial teams have a new and powerful tool to better equip them to eschew repetitive tasks, like data cleaning, validating and loading, and analysis preparation. Furthermore, this allows actuaries to focus on more complex tasks that require human expertise and judgment according to [1]. On the same note, machine learning techniques can be employed in actuarial reserving to calculate claims reserves on individual claims data and also machine learning algorithms can be used in actuarial science functions such as rate making and reserving. Moreover, the use of machine learning methods in reserving can provide insights, improve accuracy and ostensibly reduce the time required for claim reserving [2]-[4].

1.1. Traditional Actuarial Loss Reserving Methods

One of the main traditional loss reserving methods is the Chain ladder method and this entire method is based on the assumption that the ratio of cumulative claims in successive years is constant [5]. Despite the method being simple and easy to implement, but it does not capture the risk characteristics of policyholders. The next in the line is the Bornhuetter-Ferguson method which is based on the assumption that the ratio of ultimate losses to incurred losses is constant again [6]. Moreover, the method is more complex than the chain ladder method, but it can be more accurate. In addition to that, the Bayesian methods are also considered to be part of traditional actuarial loss reserving methods. These methods are entirely based on the assumption that the parameters of the model are random variables with prior distributions. Further than this, these methods are relatively more flexible than other methods, however they are more computationally intensive. Also, the Stochastic models are also part of the traditional reserving methods. These types of loss reserving models are based on the assumption that the frequency and severity of claims are random variables. Stochastic models can be more accurate than other methods, but they can be more complex [7].

1.2. The Theory and Structure of the Traditional Chain Ladder Model

The Traditional Chain Ladder Model is a widely used actuarial technique for estimating future claims reserves in insurance. It falls under the broader category of claims reserving methods and is particularly effective when historical data is abundant and reasonably reliable.

The Chain Ladder method relies on the assumption of a stable claims development pattern over time. It operates under the principle that past claims experience can be used to predict future claims. The model involves developing a chain of ratios between successive periods' claims data, hence the name "Chain Ladder" presented by Table 1 below.

Table 1. Chain ladder method: general structure.

Accident Year	Development Year			
	1	2	$\cdots$	m
1	C_{11}	C_{12}	$\cdots$	C_{1m}
2	C_{21}	C_{22}	$\cdots$	C_{2m}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
n	C_{n1}	C_{n2}	$\cdots$	C_{nm}

Here are the key remarks:

Conservative Assumption: The model assumes that past claims experience is indicative of future claims behaviour, which might not always hold true, especially in rapidly changing markets. *Data Quality*: Accuracy of results heavily depends on the quality and reliability of historical claims data. Any anomalies or outliers can significantly impact the model's predictions. *Development Factors*: The method employs development factors to extrapolate future claims based on historical data. These factors capture the relationship between claims at different stages of development.

Theorem 1 *Bornhuetter-Ferguson Technique: This technique, often integrated with the Chain Ladder method, combines historical data with expected future developments to provide a more comprehensive estimate of reserves* [8].

Theorem 2 *Mack's Model: Mack developed a stochastic version of the Chain Ladder method, which acknowledges the uncertainty in future claims development. It incorporates statistical distributions to provide a range of potential outcomes.* [9]

The Chain Ladder algorithm can be summarized as follows [10]:

1) Calculate development factors for each period based on historical claims data.

2) Apply these factors to the latest available claims data to forecast future claims development.

3) Sum the forecasted claims to estimate total reserves.

Proposition 3 *The Chain Ladder method is robust when historical claims experience exhibits a consistent development pattern.*

The accuracy of Chain Ladder estimates diminishes when there are significant deviations in claims development trends.

1.3. Inflation Adjusted Frequency-Severity Approach in Machine Learning

This paper presents a general machine learning algorithm for implementing methods using the Inflation Adjusted frequency-severity approach. The Inflation Adjusted frequency-severity approach is crucial in actuarial science and risk management to account for the inflationary effects on frequency and severity of claims. This document details the algorithmic steps and theoretical foundations of this approach in the context of machine learning [11] and [12].

Algorithm

Algorithm 1 Inflation Adjusted Frequency-Severity Model

1: **Input:** Historical claims data $D = \{(x_i, y_i)\}_{i=1}^{n}$
2: **Input:** Inflation index $I = \{I_t\}_{t=1}^{T}$
3: Adjust claim amounts for inflation
4: **for** each claim i **do**
5:     $y_i' \leftarrow y_i \times \frac{I_{t_{current}}}{I_{t_i}}$ {Inflation adjustment}
6: **end for**
7: Split data into training and test sets
8: Train frequency model $f_{\text{freq}}(x)$ on adjusted data
9: Train severity model $f_{\text{sev}}(x)$ on adjusted data
10: Predict frequency $\hat{f}_{\text{freq}}(x_{new})$ and severity $\hat{f}_{\text{sev}}(x_{new})$ for new data x_{new}
11: **Output:** Predicted frequency and severity $\hat{f}_{\text{freq}}(x_{new})$, $\hat{f}_{\text{sev}}(x_{new})$

1.4. Theoretical Foundations

1.4.1. Proposition: Inflation Adjustment

Proposition 4 *Given a claim amount y_i at time t_i and the inflation index I_t, the inflation adjusted claim amount y_i' is given by:*

$$y_i' = y_i \times \frac{I_{t_{current}}}{I_{t_i}} \tag{1}$$

1.4.2. Theorem: Consistency of Frequency-Severity Estimation

Theorem 5 *Assuming that the inflation index I_t is accurately measured, the inflation-adjusted frequency-severity model produces consistent estimators for the underlying claim frequency and severity.*

Proof. Let $Y = \{y_i\}_{i=1}^{n}$ be the set of claim amounts and $I = \{I_t\}_{t=1}^{T}$ be the inflation index. Adjusted claims are $Y' = \left\{ y_i' = y_i \times \frac{I_{t_{current}}}{I_{t_i}} \right\}_{i=1}^{n}$. The frequency model $f_{\text{freq}}(x)$ is trained on Y'. By the properties of consistent estimators and assuming I_t is accurately measured, the estimators $\hat{f}_{\text{freq}}$ and $\hat{f}_{\text{sev}}$ are consistent for the true frequency and severity. $\square$

1.4.3. Claim: Improved Prediction Accuracy

Adjusting for inflation improves the accuracy of frequency and severity predictions in a machine learning model.

Justification. Adjusting for inflation removes the temporal variability in claim amounts due to economic factors, allowing the model to learn patterns related to the actual risk factors rather than inflationary effects. Empirical studies, such as those by [12], have shown improved prediction metrics when using inflation-adjusted data.

Incorporating inflation adjustments in the frequency-severity modelling process enhances the accuracy and reliability of predictions. The theoretical propositions and empirical claims support the effectiveness of this approach.

1.5. Rationale of the Study

In this paper a proposition to the Automated Actuarial Data Analytics Based Inflation Adjusted Frequency Severity Loss Reserving Model is implemented which is essentially a model that can be employed to predict the future losses of an insurance company. Moreover, the model takes into account the frequency and severity of claims and adjusted for inflation and thus in general the model can be used to estimate the amount of money that an insurance company needs to reserve for future claims. This model is computationally efficient and accurate than the rigid, inflexible and almost obsolete traditional actuarial loss reserving methods discussed above.

1.6. Merits of Machine Learning-Based Actuarial Loss Reserving Models over Traditional-Based Actuarial Loss Reserving Models

Machine learning-based actuarial loss reserving models have several advantages over traditional methods such as the chain ladder model and some of the advantages include: Improved accuracy where the machine learning algorithms can learn complex patterns in the data that may not be captured by traditional methods and this can lead to more accurate predictions of future losses. Flexibility is another advantage where machine learning algorithms can be more flexible than traditional methods, allowing for more complex models that can capture a wider range of patterns in the data [13]. Automation of actuarial tasks involved in actuarial loss reserving is quite implemented and on the other hand reducing the need for manual intervention [14]. Furthermore, the machine learning algorithms can process large amounts of data quickly, allowing for faster analysis and decision making. Ultimately, transparency is another merit where the machine learning algorithms can be more transparent than traditional methods, thus allowing actuaries to better understand how the model is making predictions [15].

On the other side of the coin, it is important to note that machine learning-based actuarial loss reserving models are not a panacea and thus they require large amounts of high-quality data and expertise in machine learning techniques to develop and maintain. In addition to that, they can be more computationally intensive than traditional methods, hence requiring more powerful hardware and longer processing times.

1.7. The Machine Learning Frequency Severity Approach

The frequency-severity machine learning model approach is used to predict the future losses of an insurance company and it takes into account the frequency and severity of claims and uses machine learning algorithms to make predictions [16]. Furthermore, the structure of the frequency-severity machine learning model approach can vary depending on the specific algorithm used. However, most models follow a similar structure: Data preparation being the first step in

building a frequency-severity machine learning model and this involves cleaning the data, transforming it into a format that can be used by the machine learning algorithm, and splitting it into training and testing data sets respectively. From there model selection becomes the second step which is entirely premised on the selection of an appropriate machine learning algorithm for the problem at hand [17]. In addition to that, there are so many different algorithms that can be used for frequency-severity modeling, including Decision Trees, Random Forests, Support Vector Machines and Artificial Neural Networks. The third step is the model training. Once an algorithm has been selected, the next step is to train the model on the training data set and this involves using the algorithm to learn patterns in the data that can be used to make predictions [18]. From there, model evaluation takes over. After the model has been trained, it is evaluated on the testing data set to see how well it performs and this involves comparing the predicted values to the actual values and calculating various metrics such as accuracy, precision, recall, F1-score, and or the ROC curve. Finally, the last step is the model deployment. Thus, once the model has been evaluated and found to be satisfactory, it can be deployed in production to make predictions on new data.

1.8. The Machine Learning Models Performance Evaluation

There are several ways to evaluate the performance of a machine learning model and here are some of the most common methods: Holdout method is one of the ways where we split the data set into two parts respectively the training and testing. From there we then train the model on the training data set and evaluate its performance on the testing data set. In addition to that, the holdout method is simple and easy to implement, but it has some limitations. For example, it may not be representative of the entire data set. Next method is the cross validation method. In this method, we divide the dataset into k-folds and use $k-1$ folds for training and the remaining fold for testing and we repeat this process k times, each time using a different fold for testing. Cross-validation is more reliable than the holdout method since it uses all the data for training and testing. Evaluation metrics method is another way of evaluating the machine learning models. With regards to this method there are several evaluation metrics available, such as accuracy, precision, recall, F1-score, or ROC curve and the choice of evaluation metric depends on the problem and goal.

Additionally, the experiment tracking is another useful way of evaluating the machine learning algorithm performance. This particular method is important to keep track of your experiment results, such as model parameters, scores, and errors and you can use tools like spreadsheets, notebooks, or dashboards to keep track of your experiments.

1.9. Novelty, Originality and Significance

A new model for estimating and predicting actuarial loss reserves whilst automating the micro finance services and auto insurance services on the same plat-

form for the general insurance sector is implemented. The study has come at a time where there is rapid growth in technological advancement and the use of Artificial Intelligence (AI) hence this has resulted in traditional loss reserving methods becoming obsolete and inappropriate to capture all risk characteristics and rating variables for determination and prediction of loss reserves both at micro and macro levels. In this study we have successfully automated auto insurance services and micro finance services on the same platform using Artificial Intelligence based actuarial data analytics through the extension of the frequency severity with conscription of inflation adjusted model to policyholder risk categories. Despite the policyholder being in any of the policies either having micro finance policy, Auto insurance policy or both our proposed model in this study works well despite the policyholder reserving category occupied by the policyholder. In a nutshell it is an improvement to our paper [1].

1.10. Contribution to the Body of Knowledge

This study makes significant contributions to the actuarial literature by addressing several limitations of traditional loss reserving methods and advancing the field through the application of modern artificial intelligence (AI) techniques. One of the key contributions of this paper is the introduction of comprehensive reserve categories. Traditional actuarial methods typically focus on a limited number of reserve categories, often leading to oversimplification and potential inaccuracies in loss estimation [19]. By defining six distinct reserve types—NYIC, IBNYR, RBNYS, RBCWP, RACBR, and RBNRS—the proposed model provides a more granular and accurate approach to loss reserving. This granularity helps in better capturing the nuances of different types of claims and their respective risk profiles [20]. Incorporating inflation adjustments directly into the frequency-severity framework is another significant advancement. Many existing models do not explicitly account for inflation, which can lead to under- or over-reserving in the face of changing economic conditions [21]. By adjusting for inflation, this model ensures that the reserves are more accurate and reflective of the true economic value of future claims, thus enhancing the reliability of financial planning and solvency assessments for insurers.

The use of AI for real-time actuarial data analytics represents a major leap forward in the automation and efficiency of the reserving process. Traditional methods are often manual and time-consuming, requiring significant actuarial judgment [22]. The proposed methodology leverages AI to automate the reserving process, reducing human error and improving the speed of reserve calculations. This automation allows actuaries to focus on higher-level strategic decision-making and enhances the overall responsiveness of the reserving process to new data and emerging trends [23]. The comparative analysis demonstrating the superior performance of the proposed model over traditional loss reserves further underscores its contribution to actuarial science. By employing machine learning techniques that account for inflation and various risk profiles, the mod-

el offers better predictive accuracy and adaptability [24]. This improved performance can lead to more effective risk management and pricing strategies, ultimately benefiting insurers and policyholders alike.

By incorporating reserves for future unknown claims (NYIC), incurred but not yet reported claims (IBNYR), reported but not yet settled claims (RBNYS), and other specific categories, the proposed model provides a holistic approach to risk management. This comprehensive framework ensures that all potential future liabilities are adequately accounted for, thereby enhancing the financial stability and resilience of insurance companies [25].

The practical implications of this research are profound. Insurance companies adopting this model can expect more accurate reserve estimates, reduced operational costs due to automation, and improved financial planning capabilities. Moreover, this paper lays the groundwork for future research in actuarial science, particularly in exploring the application of advanced AI techniques in other areas of insurance and risk management.

In short, this paper significantly enriches the actuarial literature by introducing a novel, AI-based, inflation-adjusted loss reserving model that addresses the limitations of traditional methods and offers substantial improvements in accuracy, efficiency, and comprehensiveness.

The introduction of the Comprehensive Automated Actuarial Loss Reserves (CAALR) methodology and the additional types of actuarial loss reserves in the paper serve to address several challenges and shortcomings present in existing actuarial practices:

Inadequate Provisioning for Future Unknown Claims: Traditional actuarial practices may not adequately account for future unknown claims, leading to underestimation of reserves. The CAALR methodology addresses this by introducing reserves such as NYIC (Not Yet Incurred Claims) and IBNYR (Incurred But Not Yet Reported), which allocate specific percentages of total reserves to cover these uncertainties. *Reporting Delays and Settlement Lags*: Reporting delays and settlement lags can distort the accuracy of loss reserves. The IBNYR and RBNYS reserves, which are allocated for claims incurred but not yet reported and reported but not yet settled, respectively, aim to mitigate the impact of these delays by setting aside appropriate reserves. *Reopened Claims and Reinsurance Needs*: Reopened claims and the need for reinsurance coverage pose additional challenges to traditional actuarial methods. Reserves like RACBR (Reopened and Closed But Reopened) and RBNRS (Reported But Needing Reinsurance) are introduced to address these specific scenarios, ensuring that adequate provisions are made for such occurrences. *Heterogeneous Risk Profiles*: Existing actuarial practices may not effectively capture the diverse risk profiles of policyholders across different lines of business. The CAALR methodology, augmented with Artificial Intelligence (AI) techniques, aims to overcome this limitation by integrating diverse risk profiles for various policyholders across Micro Finance, Auto Insurance, and other lines of business on a unified platform.

These innovations collectively enhance the accuracy and reliability of actuarial

loss reserves by better accounting for uncertainties, reducing reporting and settlement distortions, and accommodating diverse risk profiles. By addressing these challenges, the CAALR methodology and additional reserves contribute to more robust and informed decision-making in insurance reserving practices.

1.11. Review of Methods

Traditional actuarial techniques for loss reserving have been the cornerstone of the insurance industry for decades. These methods include the chain-ladder method, Bornhuetter-Ferguson method, and the expected loss ratio method. Despite their widespread use, these techniques have notable limitations.

Traditional methods often rely on several assumptions, such as constant loss development patterns and homogeneous risk profiles, which may not hold true in practice [19]. In addition to that, these methods can be inflexible in handling varying types of claims and dynamic changes in the underlying risk environment [20]. Traditional reserving methods are typically manual and labor-intensive, requiring significant actuarial judgment and time to adjust for various factors [22]. The ability to predict future claims accurately is often constrained due to the linear nature of traditional models, which might not capture the complex, non-linear relationships present in insurance data [21].

Recent advancements in data science and machine learning have introduced new methodologies for loss reserving that address many limitations of traditional techniques. Machine learning (ML) approaches offer several advantages. ML models leverage large datasets to uncover patterns and relationships that are not immediately apparent, improving predictive accuracy [26]. These models can adapt to new data and changing environments more easily than traditional methods [25]. ML techniques can automate the reserving process, reducing the time and effort required for manual adjustments (Kuo, *et al.*, 2018). ML models, such as neural networks and decision trees, can capture non-linear relationships between variables, providing more robust predictions [24].

Several studies have explored the application of machine learning in loss reserving. [26] applied generalized linear models (GLMs) and other machine learning algorithms to predict outstanding claims, demonstrating improved accuracy over traditional methods. [25] examined the use of Bayesian neural networks for reserve risk, showing that these models provide more reliable estimates and better handle uncertainty. [23] utilized gradient boosting machines (GBMs) to forecast claims development, highlighting their ability to manage complex data structures and interactions. [24] explored the use of deep learning techniques, including convolutional neural networks (CNNs), for loss reserving, finding that these models outperform classical actuarial methods in terms of predictive performance.

[2] pointed out that actuarial reserving techniques have evolved from the application of algorithms, like the chain-ladder method, to stochastic models of claims development, and, more recently, have been enhanced by the application

of machine learning techniques. Moreover the authors revisited the traditional reserving techniques within the framework of supervised learning to select optimal reserving models and showed that the use of optimal techniques can lead to more accurate reserves and investigate the circumstances under which different scoring metrics should be used.

[27], while traditional actuarial reserving methods assume that development patterns are stable over time, changes are often observed in practice. This paper explores the reasons for these changes and surveys the most relevant literature on methods that address the changes in development patterns. Finally, the paper suggests possible research for further improvements in reserving techniques.

[28], this paper argues that all reserving methods based on claims triangulations (the "triangle trick"), no matter how sophisticated the subsequent processing of the information contained in the triangle is, are inherently inadequate to accurately model the distribution of reserves, although they may be good enough to produce a point estimate of such reserves. The reason is that the triangle representation involves the compression (and ultimately the loss) of crucial information about the individual losses, which comes back to haunt us when we try to extract detailed information on the distribution of incurred but not reported (IBNR) and reported but not settled (RBNS) losses.

[29], loss reserves are typically one of the largest liabilities on an insurer's balance sheet since they can have a significant impact on profits as well as the insurer's solvency. The Chain Ladder model is an outstanding actuarial reserving technique that has been applied over the years to estimate Incurred But Not Reported claims. This project aims to provide the most accurate estimates possible for the calculation and prediction of reserve claim amounts in the context of corporate health insurance. For this, the Chain Ladder approach is compared with machine learning algorithms such as the Support Vector Machine (SVM), the Random Forest (RF), the Extreme Gradient Boosting (XGBoost) and Neural Networks (NN).

[30], Run-off triangles present usual instruments for claims reserve predictions. The paper suggests a relatively simple method of such predictions based on the Holt-Winters recursive formulas modified for missing data. The technique explicitly calculates the corresponding prediction error grounded in the state space modeling and evaluates the claims reserving risk in this way. Empirical data examples enable us to compare the suggested approach with results published by other authors.

[31], this paper explores the tuning and results of two-part models on rich datasets provided through the Casualty Actuarial Society (CAS). These datasets include bodily injury (BI), property damage (PD) and collision (COLL) coverage, each documenting policy characteristics and claims across a four-year period. The datasets are explored, including summaries of all variables, then the methods for modelling are set forth. Models are tuned and the tuning results are displayed, after which we train the final models and seek to explain select predictions. Data were provided by a private insurance carrier to the CAS after anon-

ymizing the dataset. These data are available to actuarial researchers for well-defined research projects that have universal benefit to the insurance industry and the public.

In conclusion, while traditional actuarial methods have laid the groundwork for loss reserving, their limitations have become more apparent with the advent of complex and dynamic risk environments. Machine learning approaches offer significant improvements in terms of flexibility, accuracy, and efficiency. The proposed methodology enhances these advancements by introducing comprehensive reserve categories, incorporating inflation adjustments, and leveraging AI for real-time analytics, thereby providing a robust solution to modern actuarial challenges.

2. Methodology

[32] defines research methodology as a technique or strategy developed to give insights on the phenomenon of interest. This has been also implemented in the development of the Automated Actuarial Loss Reserving data analytics based model using the ten machine learning models which are the General Linear Model (GLM), Generalized Additive Model (GAM), Regression Trees (RPART), Random Forests (RF), Generalized Boosting Machines (GBM), Extreme Gradient Boosting Method (XGB), Least Angle Regression (LAR), Extreme Learning Machines (ELM), Robust Regression Method (RRM) and Artificial Neural Network (ANN).

2.1. Machine Learning Algorithms Used

Here is a detailed description of the machine learning algorithms used in the study, including their underlying principles, strengths, and limitations:

2.1.1. Generalized Linear Models (GLM)
GLMs extend the linear regression model to accommodate non-normally distributed response variables by specifying a link function and a probability distribution for the response variable. *Strengths*: Interpretable coefficients, ability to model various types of response distributions. *Limitations*: Assumes linearity between predictors and response, may not capture complex relationships.

2.1.2. Generalized Additive Models (GAM)
GAMs extend GLMs by allowing for non-linear relationships between predictors and response through the use of smooth functions. *Strengths*: Flexibility to model non-linear relationships, can capture complex interactions. *Limitations*: Interpretability can be challenging, and requires careful tuning of smoothing parameters.

2.1.3. Regression Trees (RPART)
RPART constructs a binary tree structure where each node represents a split based on predictor variables, aiming to partition the data into homogeneous segments. *Strengths*: Easy to interpret, can handle non-linear relationships and

interactions. *Limitations*: Prone to overfitting, lack of smoothness in predictions.

2.1.4. Random Forest (RANGER)

Random Forest is an ensemble learning method that constructs multiple decision trees and combines their predictions to improve accuracy and reduce overfitting. *Strengths*: Robust to overfitting, handles high-dimensional data well, provides feature importance. *Limitations*: Lack of interpretability for individual trees, computationally intensive for large datasets.

2.1.5. Generalized Boosting Machines (GBM)

GBM builds an ensemble of weak learners (typically decision trees) sequentially, with each new model focusing on the errors made by previous models. *Strengths*: High predictive accuracy, handles complex interactions, robust to outliers. *Limitations*: Sensitive to overfitting, tuning hyperparameters can be time-consuming.

2.1.6. Extreme Gradient Boosting (XGB)

XGB is an optimized implementation of gradient boosting, emphasizing computational efficiency and scalability. *Strengths*: High performance, supports parallelization, handles missing data. *Limitations*: Requires careful tuning of hyperparameters, and may be sensitive to noisy data.

2.1.7. Least Angle Regression (LAR)

LAR is a regression method that sequentially adds predictors with the highest correlation to the response, adjusting their coefficients along the way. *Strengths*: Efficient for high-dimensional data, provides a path of coefficient changes. *Limitations*: Assumes linear relationships, and may not handle multicollinearity well.

2.1.8. Extreme Learning Machines (ELM)

ELM is a feedforward neural network with a single hidden layer where the weights connecting the input and hidden layers are randomly assigned and fixed, and only the output layer weights are learned. *Strengths*: Fast training speed, good generalization performance. *Limitations*: Lack of interpretability, may require tuning of hidden units and activation functions.

2.1.9. Robust Regression Method (RRM)

RRM aims to fit a regression model that is less sensitive to outliers by minimizing the influence of data points with large residuals. *Strengths*: Robust to outliers, maintains performance with contaminated data. *Limitations*: Sacrifices some efficiency compared to ordinary regression when data is clean.

2.1.10. Artificial Neural Network (ANN)

ANN consists of interconnected nodes organized in layers, where information is processed through weighted connections and nonlinear activation functions. *Strengths*: Ability to model complex relationships, good for pattern recognition tasks. *Limitations*: Requires large amounts of data, prone to overfitting without proper regularization.

These algorithms were likely selected based on their suitability for the task of predicting actuarial loss reserves, considering factors such as the complexity of relationships between predictors and response variables, the presence of non-linearities and interactions, and the need for robustness to outliers and missing data. The choice of algorithms may also have been influenced by their availability in popular statistical software packages like R, as well as their established performance in similar predictive modeling tasks within the insurance industry. Additionally, the hyperparameters for each algorithm were likely tuned to optimize predictive performance while avoiding overfitting.

Data was explored in R before proceeding as follows.

2.2. The Process

This section describes how the proposed model in this paper has been developed commencing with traditional actuarial loss reserving method preferably the general Chain ladder model and then proceeded to the machine learning based automated actuarial loss reserving model. The proposed Automated Actuarial Data Analytics Based Inflation Adjusted Frequency Severity Loss Reserving model begins with the Automated Micro finance services actuarial loss reserving followed by the Automated Auto Insurance Services Actuarial Loss Reserving and finally the Both Services Automated Actuarial Loss Reserving. The model can be implemented by an insurance company or any financial institutions such as banks, discount houses, brokers and any other related financial house. The outline of the study methodology has been outlined below respectively.

2.3. Traditional Chain Ladder Actuarial Loss Reserving Method

The run off triangle loss reserving method has first been applied from the R package Chain ladder [33] and [34] for the micro finance loss cohort, auto insurance cohort, the Both Services and finally the comprehensive services combining the three services offered by an insurer. As mentioned earlier in the introductory part of this paper that the Chain ladder model as one of the traditional loss reserving model is far from reflecting the current risk profile characteristics and thus the following model is proposed as presented below.

2.4. The Structure of the Automated Actuarial Data Analytics Based Inflation Adjusted Frequency Severity Loss Reserving Model

The frequency, severity and inflation models are fitted as follows with regards to the defined proposed actuarial notation.

Let

• $Freq_{micro}$ be the Frequency model for Automated Micro Finance Services Loss Reserving Model

• $Freq_{auto}$ be the Frequency model for Automated Auto Insurance Services loss Reserving Model

• *Freq*$_{both}$ be the Frequency model for Automated Both Services Loss Reserving Model

• *Sev*$_{micro}$ be the Severity model for Automated Micro Finance Services Loss Reserving Model

• *Sev*$_{auto}$ be the Severity model for Automated Auto Insurance Services loss Reserving Model

• *Sev*$_{both}$ be the Severity model for Automated Both Services Loss Reserving Model

• *Inf*$_{micro}$ be the Inflation adjustment model for Automated Micro Finance Services Loss Reserving Model

• *Inf*$_{auto}$ be the Inflation adjustment model for Automated Auto Insurance Services loss Reserving Model

• *Inf*$_{both}$ be the Inflation adjustment model for Automated Both Services loss Reserving Model

The above models are then developed and defined as shown in the stages outlined below. In actuarial literature, the Traditional Loss reserving methods have only been centered around two main types of actuarial reserves which are the IBNYR (Incurred But Not Yet Reported) and RBNYS (Reported But Not Yet Settled). In order to implement full automation process for these two main functions respectively the micro finance services and the auto insurance services, four further types of actuarial reserves have been introduced to co-exist with these two main actuarial types of reserves which are NYIC (Not Yet Incurred Claims), RBCWP (Reported But Closed With Payments), RACBR (Reported And Closed But Reopened) and RBNR (Reported But Needs Reinsurance). These have been defined as follows by Table 2 below.

Table 2. Actuarial loss reserves, definitions and associated allocated weights.

Actuarial Loss Reserve Definition	Allocated Weight
NYIC (Not Yet Incurred Claims)	0.1
IBNYR (Incurred But Not Yet Reported)	0.3
RBNYS (Reported But Not Yet Settled)	0.2
RBCWP (Reported But Closed With Payments)	0.2
RACBR (Reported And Closed But Reopened)	0.1
RBNRS (Reported But Needs Reinsurance)	0.1

Furthermore, the allocations proposed have been distributed to these types of actuarial reserve weights (subjectively) in order to estimate the reserves themselves from the actual loss attained by an insurer/financial institution of interest with the following assumptions below respectively.

• NYIC (0.1) is a reserve introduced in this paper set aside for future unknown claims and its pegged at 10% of total actuarial reserves presented in the proposed

actuarial reserving model.

- IBNYR (0.3) is a reserve set aside for future unknown claims that have been incurred but not yet reported to the insurer and they have been pegged at 30% of total actuarial reserves presented in the proposed actuarial reserving model. It is quite clear to notice that this type of reserve is slightly greater than all the reserves presented in our model since this type of reserve is most dominant due to large quantity of unreported claim supported by the existence of reporting delay time/lag in the vicinity of the actuarial loss reserving.

- RBNYS (0.2) is a reserve set aside for future unknown claims that have been reported but not yet settled by the insurer and they have been pegged at 20% of total actuarial reserves presented in the actuarial reserving model. In addition to that it is quite essential to notice that this type of reserve is greater than NYIC, RBCWP, RACBR and RBNR since, it houses a large stake of reported but not yet paid/settled claims.

- RBCWP (0.2) is a reserve which caters for future unknown claims that have been reported but closed with payment by the insurer and they have been pegged at 20% of total actuarial reserves presented in the proposed actuarial reserving model.

Additionally, it is quite critical to notice that this type of reserve is slightly greater than NYIC, RACBR and RBNR since this category of reserve houses a large stake of reported and fully paid/settled claims. This type of reserve has been proposed in this study to ensure that both settlement and delay time are reduced by the use of artificial intelligence based actuarial data analytics employed in actuarial real time loss reserving.

- RACBR (0.1) is a reserve introduced in this paper to cater for future unknown claims that have been reported and closed but then at some point in future these claims are then reopened and then the insurer is then required to fully settle this reopened claim. As a result we have proposed to peg them at 10% of total actuarial reserves presented in the actuarial reserving model.

- RBNRS (0.1) is a reserve that caters for future unknown claims that have been reported but then requires reinsurance whether proportional or Excess of loss reinsurance even any form of reinsurance is welcome. This reserve exists in order to reduce both the quantum, severity and incidence of catastrophic claims by ensuring that these reinsurance based reserves are put in place to meet catastrophic claims. In short, this type of actuarial reserve is pegged at 10% of total actuarial reserves presented in the actuarial reserving model.

2.5. The Proposed Actuarial Loss Reserving Process

Figure 1 is presented below as the Actuarial Loss Reserving automated and augmented by Artificial Intelligence based actuarial data analytics procedures.

In a nutshell, Figure 1 shows the general proposed Actuarial real time based loss reserving which insurance companies or any related finance houses of interest can adopt respectively.
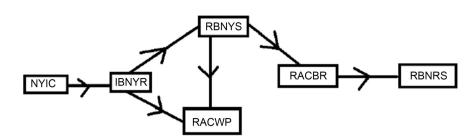
Figure 1. The proposed Actuarial loss reserving process.

2.6. Methodology for Automated Actuarial Loss Reserving

The frequency, severity and inflation models are fitted with respect to each policyholder category defined as: as follows with regards to Micro Finance policyholder category, Auto Insurance policyholder category and Both Services Policyholder category. Additionally, the model performance and evaluation is premised from which are the Root Mean Square Error (RSE), which is the main machine learning model metric. Finally the Comprehensive Services Automated Actuarial Loss Reserving Model estimates is coined by summing the predictions from the three main types of models outlined above with regards to the outlined policyholder category just mentioned now. Next from there, the model predictions for each of the mentioned three models are predicted on test data and are automated to give the Automated Actuarial Ultimate Loss (AAUL) per each policyholder in the test data set thus at micro level. Moreover the AAUL is obtained at macro level by summing the AAUL per each policyholder with regards to Micro Finance service category, Auto Insurance service category and Both Services category. From there, the predicted Total Ultimate Losses (TUL) with regards to Micro Finance services Model, Auto Insurance Services Model and Both Services Model are summed to give the Total Automated Actuarial Ultimate Losses which are both allocated and distributed across the proposed types of actuarial reserves is presented on **Table 2** above.

This is mathematically presented below as shown below.

$$\text{NYIC} = \text{Proposed Proportion for NYIC} \times \text{TUL} \tag{2}$$

$$\text{IBNYR} = \text{Proposed Proportion for IBNYR} \times \text{TUL}$$
$$\text{RBNYS} = \text{Proposed Proportion for RBNYS} \times \text{TUL}$$
$$\text{RBCWP} = \text{Proposed Proportion for RBCWP} \times \text{TUL} \tag{3}$$
$$\text{RACBR} = \text{Proposed Proportion for RACBR} \times \text{TUL}$$
$$\text{RBNRS} = \text{Proposed Proportion for RBNRS} \times \text{TUL}$$

The Case Reserves are allocated using the same basis for allocation and distribution into the main six types of actuarial loss reserves suggested in this paper, see **Table 2**.

2.7. Framework for Automated Actuarial Loss Reserving Model

In a nutshell below is **Figure 2** which shows the framework for the actuarial loss reserving model respectively.
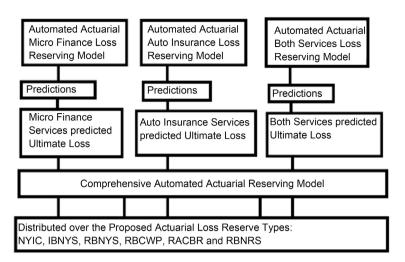
Figure 2. The proposed actuarial loss reserving structure.

The cumulative expected Ultimate Losses are added as shown by **Figure 3** below.
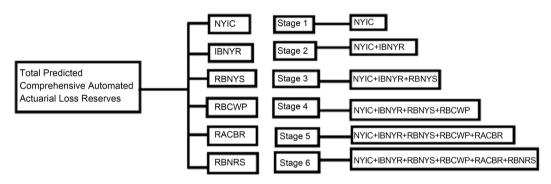

Figure 3. The proposed Actuarial loss reserving structure.

Figure 3 is complimented by the cumulative lower triangle presented on **Table 3** presented below.

Table 3. Lower triangle cumulative ultimate losses.

	Cumulative Addition of Types of proposed Actuarial Loss Reserves					
Stages	+	+	+	+	+	+
1	NYIC					
2	NYIC	IBNYR				
3	NYIC	IBNYR	RBNYS			
4	NYIC	IBNYR	RBNYS	RBCWP		
5	NYIC	IBNYR	RBNYS	RBCWP	RACBR	
6	NYIC	IBNYR	RBNYS	RBCWP	RACBR	RBNRS

Table 3 shows stages 1 up 6 where the predicted allocated ultimate loss for the presented actuarial loss reserve types are added cumulatively. These values are

then subtracted from the actual loss attained by the insurer to estimate the Automated Actuarial Loss Reserve with regards to each allocated actuarial loss reserve type and also paying special attention to the policyholder category; Micro Finance services category, Auto Insurance category and the Both Services category.

2.8. Novelty of the Proposed Methodology

The proposed Automated Actuarial Data Analytics Based Inflation Adjusted Frequency Severity Loss Reserving Model builds upon existing research by integrating multiple innovative elements. Unlike traditional and some ML models that focus on a limited set of reserve categories, this model introduces six distinct reserves (NYIC, IBNYR, RBNYS, RBCWP, RACBR, and RBNRS), providing a more granular approach to loss reserving. The model incorporates inflation adjustments directly into the frequency-severity framework, which is a significant enhancement over many existing ML models that do not explicitly account for inflation [21]. By utilizing artificial intelligence for real-time actuarial data analytics, the proposed model not only improves efficiency but also enhances the accuracy and responsiveness of the reserving process. The paper demonstrates the superior performance of the proposed model over traditional loss reserves by employing a machine learning approach that adjusts for inflation, showing better predictive power and adaptability.

3. Data

Simulated Auto Insurance Micro finance data consists of two sets the first set being the simulated Loss transactional data from year 2010 to year 2020. The second part of the data consists of the simulated machine learning data with both rating and risk characteristics presented by Table 2 above, and thus this data set is used to develop the Automated Actuarial Loss Reserving Model. This data with a sample size of 80,000 was used on a total of 69 variables. The simulated data has been hot encoded and came to 91 variables.

3.1. The Structure of the Simulated Traditional Chain Ladder Transactional Data

Below is how the simulated the aggregate loss data for the three main Policyholder categories has been carried out respectively the Micro finance policyholders, Auto Insurance policyholders and finally the Both services policyholders. Ultimately, the summations to the losses from these three main lines of business have been conducted to obtain the Comprehensive services aggregate loss which has been further used for applying the chain ladder model in each of these three main lines of business respectively.

Table 4 shows the simulated accident years and settlement period by years from 2010 to 2020 respectively. This is how the loss data has been simulated which then were finally converted to loss triangles.

Table 4. The structure of the simulated Traditional Chain Ladder transactional data.

Triangulation Loss Reserving Chain Ladder Model Transactional loss data simulation											
	Development period										
Accident Years	0	1	2	3	4	5	6	7	8	9	10
2010	2010	2010	2010	2010	2010	2010	2010	2010	2010	2010	2010
2011		2011	2011	2011	2011	2011	2011	2011	2011	2011	2011
2012			2012	2012	2012	2012	2012	2012	2012	2012	2012
2013				2013	2013	2013	2013	2013	2013	2013	2013
2014					2014	2014	2014	2014	2014	2014	2014
2015						2015	2015	2015	2015	2015	2015
2016							2016	2016	2016	2016	2016
2017								2017	2017	2017	2017
2018									2018	2018	2018
2019										2019	2019
2020											2020

Triangulation Loss Reserving Chain Ladder Model Transactional Loss Data Simulated Sample Size

Here is the Chain ladder loss data simulations with regards to presented sample sizes.

Table 5. Triangulation loss reserving chain ladder model transactional loss data simulated sample size.

Triangulation Loss Reserving Chain Ladder Model Transactional loss data simulated sample size											
	Development period										
Accident Years	0	1	2	3	4	5	6	7	8	9	10
2010	940	600	260	148	110	141	211	119	40	25	15
2011		550	310	135	480	170	312	125	25	50	18
2012			450	245	610	182	418	143	89	71	28
2013				230	720	235	121	131	109	80	98
2014					840	180	194	115	101	95	411
2015						190	140	118	247	188	325
2016							187	218	351	222	237
2017								200	540	210	113
2018									600	321	112
2019										485	181
2020											175

Table 5 shows the sample sizes for each accident year and associated settle-

ment period. After summing all the sample sizes presented a total of 160000 is recorded as the overall sample size with regards to the basic chain ladder model

3.2. The Structure of the Simulated Data for the Automated Actuarial Data Analytics Based Inflation Adjusted Frequency Severity Loss Reserving Model

The simulated data consisted of 70 variables from both continuous and factor variables, hence after hot encoding the data using R caret package. The total of 94 variables is attained. A sample size of 800,000 policyholders from the three main policyholder categories has been simulated. Afterwards the data partitioning rule of the 80:20 into the training data set and test data set followed respectively. Furthermore, the model has been trained using the training data set with a sample of 640,000 policyholders and also the test data set with a sample size of 16,000 has been used towards model testing, evaluation of the model predictions and finally comparing the obtained machine learning based loss reserving results with the traditional chain ladder model since they have the same sample size now.

Data, Associated Models and Variables

Table 2 shows the data used to develop the machine learning based automated actuarial loss reserving model. The frequency models, severity models and inflation adjustment models have been developed with regards to each of the three main policyholder categories present in these three main lines of business. The independent and dependent variables used for these models have been also constructed using both the rating and risk characteristics prior to each line of business accordingly.

3.3. Data Exploration and Analysis

[35] postulated that data exploration, or the search for features in data that may indicate deeper relationships among variables, relies heavily on visual methods because of the power of the human eye to detect structures.

Number of Policyholders in the Three Main Lines of Business

The number of policyholders in the three main lines of business has been presented below.

Table 6. Number of Policyholders in the three main lines of business.

Lines of Business and their associated number of policyholders		
Line of Business	Proportion	Number of Policyholders
Micro Finance Services	20%	16,000.00
Auto Insurance Services	30%	24,000.00
Both Services	50%	40,000.00
Total	100%	80,000.00

Table 6 shows that Both services line of business carries the greatest number of policyholders since they occupy 50% of the sample size, followed by Auto Insurance services policyholders who totalled 240,000 and the last line of business being Micro Finance services category with 160,000 policyholders being the least. This has been visually complimented by Figure 4 shown below.
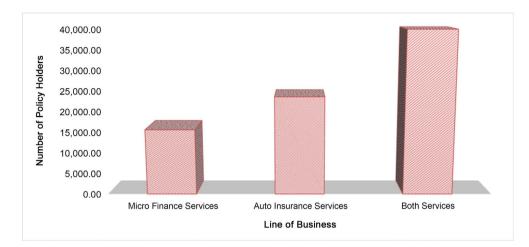

Figure 4. Number of Policyholders in the three main lines of business.

4. Main Results

This section describes results obtained from the methodology steps outlined in the methodology section 0.0.4. The first part shows results for the traditional chain ladder loss reserving method and the second part of the results shows the machine learning based automated actuarial loss reserving method. Finally the distribution and allocation of both estimated actuarial reserves and case reserves then follows before analysis of future trends for proposed types of actuarial loss reserves.

4.1. Traditional Actuarial Loss Reserving Method: Chain Ladder Model

The chain-ladder method is a way of estimating the amount of money that an insurance company needs to set aside for claims that have occurred but not yet been reported or paid and it uses historical data on how claims develop over time to project the future payments [33]. It is a traditional and widely used actuarial loss reserving technique.

4.1.1. Chain Ladder Development Factors for the Three Main Lines of Business

Most commonly as a first step, the age-to-age link ratios are calculated as the volume weighted average development ratios of a cumulative loss development triangle from one development period to the next $C_{ik}, ik = 1, \cdots, n$.

$$f_k = \frac{\sum_{i=1}^{n-k} C_{i,k+1}}{C_{i,k}} \tag{4}$$

These have been computed from the loss triangle data, see **Table A4** for Micro Finance Services loss data, **Table A5** for Auto Insurance Services loss data, **Table A6** for Both Services loss data and **Table A7** for Comprehensive Services loss data.

Table 7. Chain ladder development factors for the three main lines of business.

	Chain Ladder Development Factors			
	Micro Finance Services	Auto Insurance Services	Both Services	Comprehensive Services
1	1.733176	1.733212	1.732523	1.732897
2	1.297497	1.297562	1.297046	1.297318
3	1.188054	1.189741	1.187773	1.188489
4	1.148281	1.148149	1.147507	1.147892
5	1.131086	1.120407	1.120554	1.122840
6	1.097313	1.099558	1.098225	1.098463
7	1.050268	1.037893	1.038100	1.040731
8	1.017274	1.017159	1.017290	1.017243
9	1.009428	1.009102	1.009142	1.009192
10	1.006158	1.005835	1.005711	1.005851

Equation (4) leads to the chain ladder development factors presented on **Table 7**. These development factors have been computed with regards to the three main lines of business and finally the comprehensive services which is a combination of the three lines of businesses.
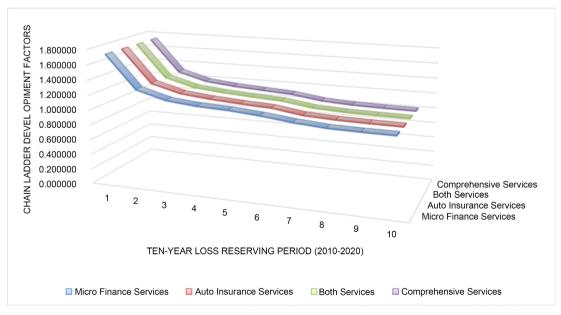

Figure 5. Chain Ladder development factors.

Figure 5 compliments results presented on **Table 7**. In short the development

factors across the three main lines of business are approximately the same and as a result the development factors too for Comprehensive Services also follows the similar pattern. From Figure 5 the development factors for the first three years is higher with the first year being the highest with a figure certainly above (1.000). In general the development factors decreased slowly until the tenth year. Finally, the Chain ladder method assumes that the development factors, which are the ratios of claims paid from one year to the next as shown above, are constant and can be used to predict the future payments [33].

4.1.2. Chain Ladder Ultimate Loss for the Three Main Lines of Business

The estimated ultimate loss is then obtained from the calculated loss development factors presented on Table 7.

Table 8. Chain ladder ultimate loss for the three main lines of business.

	Chain Ladder Total Ultimate Loss			
Year	Micro Ultimate Loss	Auto Ultimate Loss	Both Ultimate Loss	Comprehensive Ultimate Loss
2010	318.00	453.00	593.00	1364.00
2011	357.00	530.00	769.00	1656.00
2012	539.00	826.00	1134.00	2499.00
2013	4095.00	3050.00	4089.00	11,232.00
2014	8862.00	13,313.00	17,696.00	39,873.00
2015	7795.00	11,464.00	15,560.00	34,819.00
2016	6564.00	9351.00	12,554.00	24,699.00
2017	3409.00	5080.00	6949.00	15,439.00
2018	4042.00	6196.00	8164.00	18,404.00
2019	8733.00	12,929.00	17,440.00	39,103.00
2020	14,279.00	21,318.00	28,414.00	64,014.00

Table 8 shows that in general the ultimate losses for the three main lines of business increased from 2010 to 2020. which could mean and indicate a sign of generative business as more policyholders are joining the insurance company by taking policies from these three main lines of businesses. Figure 6 complements the results on Table 8 above.

The Comprehensive Services ultimate loss remains the highest since it is a summation of the total ultimate losses from the presented three main lines of business. Additionally, the both services line of business scooped the highest ultimate loss followed by Auto Insurance Services ultimate losses and the least being the Micro Insurance services line of business.

4.1.3. Chain Ladder Estimated Reserves

The chain ladder method leads to the determination of the following types of reserves with regards to the three main lines of business.
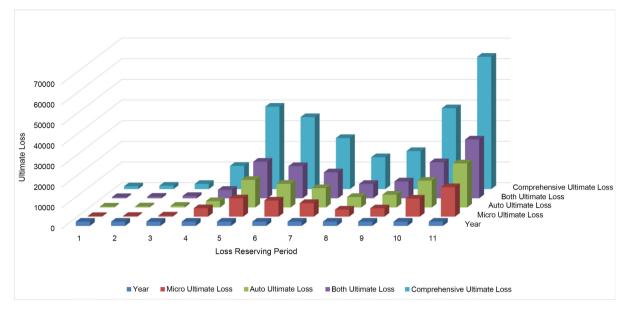
Figure 6. Chain ladder ultimate loss by lines of business.
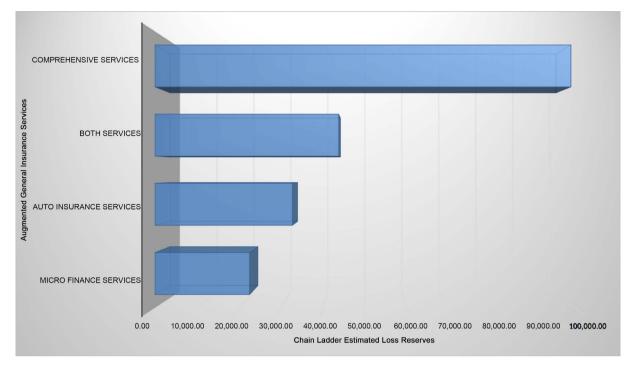
Table 9. Chain ladder estimated reserves.

	Chain Ladder Estimated Reserves			
	Micro Finance Services	Auto Insurance Services	Both Services	Comprehensive Services
Reserve	22,556.01	32,747.41	43,759.32	99,064.89

Table 9 indicates that Micro Finance services line of business attained a loss reserve estimate of (22,556.01) this was followed by Auto Insurance chain ladder loss reserve estimate of (32,747.41), on the same spot Both Services line of business attained the highest chain ladder loss reserve estimate of (43,759.32). Ultimately, Comprehensive Services achieved a chain ladder estimate of (99,064.89). This has also been expressed by Figure 7 shown below.

Figure 7 also offers a visual description of the chain ladder loss reserves for the three main lines of business outlined above. This validates the results presented on Table 9.

4.2. Machine Learning Based Automated Actuarial Loss Reserving Model

The chain ladder model does not capture the rating and risk factors in computation of the loss reserves, more over this traditional method is based on the loss development factors which are determined from historical data which turns to be outdated and primarily does not capture the current policyholder risk profile hence this turns to underestimation of loss reserves in general. In addition to that, the method is based on aggregate loss data which is both historical and cumulative in nature which is the opposite with the current digital era. As a result the application of triangulation method may seem to be appropriate as a starting point to estimate loss reserves but to be more precise and accurate machine learning methods which are Artificial intelligence based seems to be more

appropriate in both estimating predicting and analysing the future experience of loss reserves, since they capture all the rating and risk factors in the model.

Figure 7. Chain ladder augmented loss reserves.

4.2.1. Inflation Adjusted Frequency Severity Automated Actuarial Loss Reserving

Model for the three main augmented services in the general insurance. The Inflation Adjusted Frequency Severity Automated Actuarial Loss Reserving Model for the three main augmented services has been presented below with dependent and independent variables taken from **Table 2**.

Table 10. Inflation adjusted frequency severity automated actuarial loss reserving model for the three augmented services in the general insurance.

	Micro Finance Actuarial Loss Reserve Frequency Models						Auto Insurance Actuarial Loss Reserve Frequency Models						Both Services Actuarial Loss Reserve Frequency Models					
	Frequency Models		Severity Models		Inflation Models		Frequency Models		Severity Models		Inflation Models		Frequency Models		Severity Models		Inflation Models	
ML Model	Time (sec)	RMSE	Time (sec)	RMSE	Time (sec)	RMSE	Time (sec)	RMSE	Time (sec)	RMSE	Time (sec)	RMSE	Time (sec)	RMSE	Time (sec)	RMSE	Time (sec)	RMSE
GLM	1.42	1.96281	0.53	13.07809	0.58	0.28866	1.10	2.77271	1.69	19.97304	0.91	0.28863	0.48	1.98900	0.67	30.44805	0.55	0.28861
GAM	1.00	1.97552	0.65	13.06093	0.76	0.28914	1.02	2.78258	1.00	19.98818	0.87	0.28918	0.62	1.99607	0.69	30.48333	0.66	0.28915
RPART	0.91	1.96706	1.06	13.05895	0.89	0.28803	0.97	2.76338	1.00	20.00838	0.86	0.28803	0.86	2.00286	0.90	30.35886	0.91	0.28803
RANGER	155.22	1.96435	174.46	13.06100	250.26	0.29011	152.39	2.75244	834.70	19.93876	299.14	0.29070	72.44	1.99930	166.75	30.41722	197.39	0.29129
GBM	53.92	1.97027	40.02	13.09525	32.97	0.29009	106.06	2.75592	66.58	19.93758	55.64	0.28998	32.33	1.99996	29.68	30.36414	29.16	0.29001
XGB	6.43	1.97653	7.06	13.08003	8.44	0.28725	8.16	2.79038	7.95	20.00047	7.95	0.28734	6.85	2.00712	6.84	30.43097	6.93	0.28720
LAR	19.46	1.96305	18.46	13.07859	17.53	0.29027	19.66	2.74282	23.92	19.97357	19.59	0.29027	16.45	2.01450	16.86	30.43755	17.69	0.29027
ELM	0.78	1.96477	0.29	13.10956	0.25	0.29174	0.35	2.75028	0.43	20.09631	0.27	0.29379	0.28	2.01864	0.29	30.54861	0.25	0.29321
RRM	11.19	1.97172	9.34	13.16262	27.97	0.29214	11.35	2.77786	15.75	20.12218	36.94	0.29188	8.82	2.02479	7.69	30.49564	25.05	0.29231
ANN	7.14	1.97077	6.92	13.13190	10.20	0.28770	7.62	2.75445	8.33	19.97512	5.62	0.28761	7.70	1.99155	1.39	30.43648	4.35	0.28769

Table 10 shows the three main types of automated actuarial based loss regression models beginning with frequency models followed by severity models and finally inflation adjustment models across the three main lines of business. RMSE has been used as the main model performance and evaluation metric. From the three types of models, with special attention to the three main lines of business, the RMSE is approximately the same with regards to the three main types of models presented above. This is a positive sign for consistency, validity and reliability of results obtained from the presented 10 machine learning models. With regards to processing times for the three main types of regression models GLM, GAM and RPART were the fastest machine learning models to converge followed by ELM. XGB, RRM, ANN and LAR took exceptionally long to converge and give solutions for the three main types of models presented here. On the same spot, RF and GBM were the slowest machine learning models to converge and give solutions with regards to the three types of models presented across the three main lines of businesses.

4.2.2. Total Automated Actuarial Ultimate Loss Predictions for the Three Main Lines of Business

The predictions for frequency models, severity models and inflation adjustment models have been both automated and summed to give the Total Automated Actuarial Ultimate Loss Predictions based on the test data set with a sample size of 16,000 which tallies with the sample size used on traditional chain ladder method discussed on the previous section.

Table 11. Total automated actuarial ultimate loss predictions for the three main lines of business.

Total Automated Actuarial Ultimate Loss Predictions			
ML Methods	Micro Finance Services	Auto Insurance Services	Both Services
GLM	11,756.26	24,905.63	21,953.31
GAM	11,826.99	24,981.77	21,730.12
RPART	11,652.60	24,824.27	21,940.53
RANGER	12,040.11	25,055.17	22,331.23
GBM	11,864.68	24,812.98	21,980.00
XGB	11,731.08	24,918.66	21,935.16
LAR	11,794.26	24,790.51	22,345.15
ELM	11,831.94	24,774.00	22,261.64
RRM	10,793.25	23,554.02	20,442.17
ANN	11,689.83	24,693.88	21,775.60

By taking a closer look at Table 11, it is clear that across the ten machine learning models and with regards to the highlighted three main lines of business, the Micro finance services line of business achieved lowest values for the total ultimate losses followed by the Both services line of business. Finally the Auto

Insurance Services line of business achieved the highest values for the total ultimate losses. With regards to Micro Finance services, XGB received the highest score for the Micro Finance services total ultimate loss with (12,040.11), which came from the RANGER and on the same note RMM achieved the least score for Micro Finance services with (10,793.25). With regards to the Both services, RRM received the lowest score for total ultimate losses with (20,442.17) and also GBM scooped the highest score for the total ultimate loss with a value of (21,980.00). Similarly, with regards to the Auto Insurance Services line of business, RANGER received the highest score of (25,055.17) as total ultimate loss whilst the least score came from RRM with a value of (23,554.02).
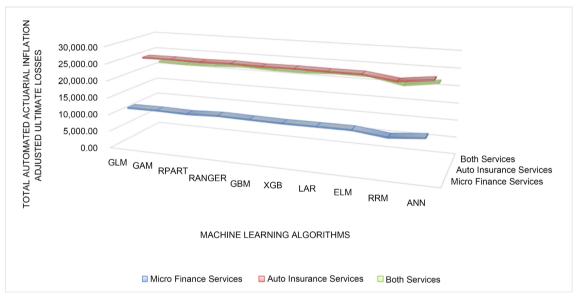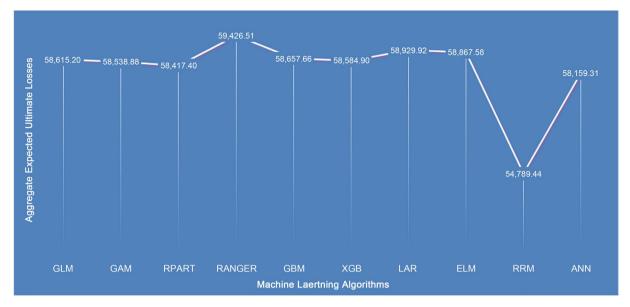
Figure 8. Total ultimate losses.

Figure 8 shows the Total automated actuarial inflation adjusted ultimate losses which too shows that the Both Services received highest values, followed by the Auto Insurance services and the least being the Micro Finance services.
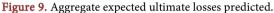
4.3. Criteria for Selecting the Best Machine Learning Model

Total automated actuarial inflation adjusted ultimate losses were totalled from **Table 10** per each machine learning learning algorithm with regards to the three main lines of business to obtain the aggregate expected ultimate losses under the Comprehensive Services category which is essentially the summation of the three main lines of business involved in the study. The greater the aggregate expected ultimate losses the better the machine learning models since there will be more actuarial loss reserves spared when actual loss is subtracted.

Table A8 from the appendix section, RANGER received a highest score for Aggregate expected ultimate loss (59,426.51) complimented with **Figure 9** where LAR attained the highest peak from the same figure. As a result RANGER is considered to be the best model for consideration for further actuarial evaluation

of loss reserves which consists of reserve estimation, allocation and distribution as well as future analysis of actuarial loss reserves.

Figure 9. Aggregate expected ultimate losses predicted.

4.3.1. Assumptions of the Automated Actuarial Loss Reserving Model

• The case reserves have been allocated and distributed using the same basis proposed on Table 12

• The actuarial loss reserves without case reserves and those without case reserves are considered for computing the net present value and accumulated value using the increasing interest rates, decreasing interest rates and constant interest rates over n future period of time.

• n can be number of days, number of weeks, number of months and or number of years

• The frequency, Severity and inflation rates are constant over n future period of time.

• The expenses and outgo are constant over n future period of time.

• Random Forest (RNAGER) being the best model machine learning model in the study has been used to develop the final Automated Actuarial Loss Reserving Model.

4.3.2. Proposed Actuarial Loss Reserve Allocations

Below is the proposed loss reserves allocations over the defined actuarial loss reserve types.

From Table 12, an allocation of (0.30) for IBNYR reserve has been proposed since this category houses the greatest number of claims before they are reported. From there RBNYS and RBCWP follow with each (0.1), these actuarial loss reserve types are not so larger than IBNYR reserve allocation since they each take part in settlement of already reported claims. NYIC, RACBR and RBNRS are the least with each allocated (0.1) in order to reduce the cost of reinsurance and

ceding costs.

Table 12. Proposed actuarial loss reserve allocations.

Reserve	Proposed Actuarial Loss Reserve Allocations
NYIC	10%
IBNYR	30%
RBNYS	20%
RBCWP	20%
RACBR	10%
RBNRS	10%

4.3.3. Allocation and Distribution of the Case Reserves

The case reserves have been allocated using the proportions presented on Table 12 also they have been classified according to the three main lines of businesses presented above.
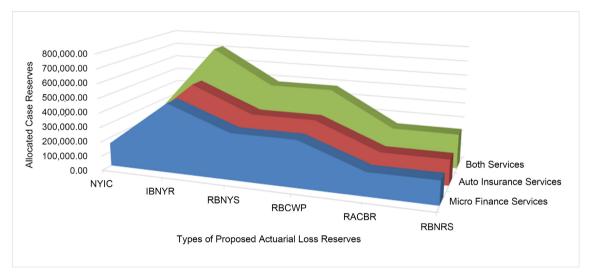

Figure 10. The allocation and distribution of the Case Reserves.

Figure 10 gives a visual expression for the case reserves allocated to proposed actuarial loss reserve types over the three main lines of businesses presented on Table A9 on the appendix section. The 3-dimensional graphs for the presented case reserves for the Micro Finance services, Auto Insurance services and Both services resemble the allocated proportions displayed on Table 12. In addition to that, Figure 10 shows that the Both services case reserves are the largest followed by Auto Insurance services case reserves and the least is Micro Finance services care reserves.

4.3.4. General Insurance Services Actual Losses Attained

The insurer experienced the actual losses from the various types of claims raised from the three main lines of business in this study.

Table 13. Augmented general insurance services actual losses attained general insurance services actual losses.

Loss Type	Total Micro Actual Loss	Total Auto Actual Loss	Total Both Actual Loss	Comprehensive Loss
No Claims	0.00	0.00	0.00	0.00
Incurred Claims	1594.967	2398.29	1602.73	5595.99
Open Reported Claims	3208.821	1606.10	3193.48	8008.40
Closed Reported Claims	4012.991	3193.83	2402.37	9609.19
Reopened Claims	2386.64	2391.14	3207.49	7985.27
Reinsurance Claims	1610.51	1582.08	2388.57	5581.16

Table 13 shows the general insurance services over the presented loss types. No claims recorded (0.00) values across the three main lines of business since they have not yet occurred and in our study is assumed to be covered by the proposed NYIC reserve category. Open Reported Claims and Reopened Claims comes second with fairly large values recorded for the three main lines of business. In addition to that the Open Reported claims are assumed to be covered by RBNYS reserves and Reopened claims are assumed to be covered by RACBR reserves. Furthermore, Incurred Claims, Closed Reported Claims and Reinsurance Claims attained highest values for the actual losses. Additionally, we make an assumption that Incurred Claims are covered by IBNYR reserves, Closed Reported Claims are covered by RBCWP reserves and Reinsurance Claims covered by RBNRS.
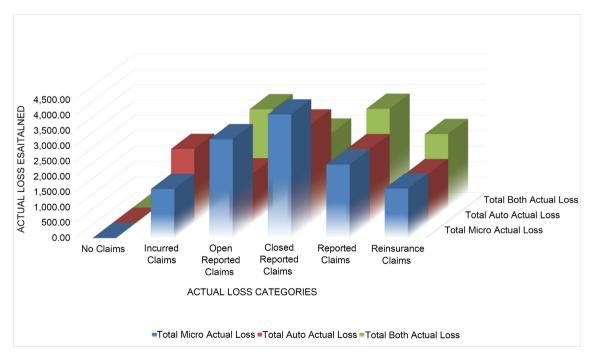

Figure 11. Augmented general insurance services actual losses attained.

Figure 11 reveals that Both Services line of business experienced a large chunk of Total actual losses followed by Micro Finance services line of business and the least Total actual losses came from the Auto Insurance services.

4.3.5. Predicted Ultimate Losses by Reserve Categories

The predicted ultimate losses have been presented below with regards to the three main lines of business and also the Comprehensive services which is basically an aggregation of all the three lines of business.
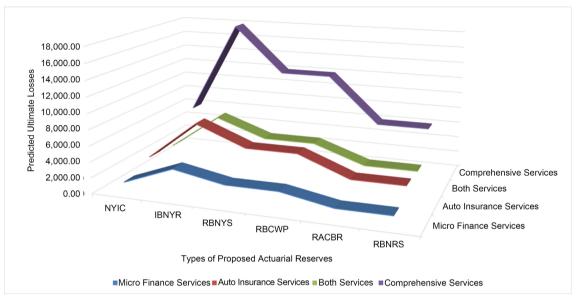

Figure 12. Predicted expected ultimate losses.

Figure 12 shows that the Both Services line of business scooped the largest stake of predicted ultimate losses followed by Auto Insurance services line of business and the last being the Micro Finance line of business. As a result, **Figure 12** compliments the results displayed on **Table A10**.

4.3.6. Cumulative Predicted Ultimate Losses by Reserve Categories

Below is the cumulative predicted ultimate losses by actuarial loss reserve categories.

Table 14. Cumulative predicted losses by reserve categories.

| Reserves | Cumulative Predicted Losses by Reserve categories | | | |
	Micro Finance Services	Auto Insurance Services	Both Services	Comprehensive Services
NYIC	1204.01	2505.52	2233.12	5942.65
IBNYR	8720.56	10,022.07	8932.49	23,770.60
RBNYS	13,731.60	15,033.10	13,398.74	35,655.91
RBCWP	18,742.63	20,044.14	17,864.98	47,541.21
RACBR	21,248.15	22,549.65	20,098.11	53,483.86
RBNRS	23,753.66	25,055.17	22,331.23	59,426.51
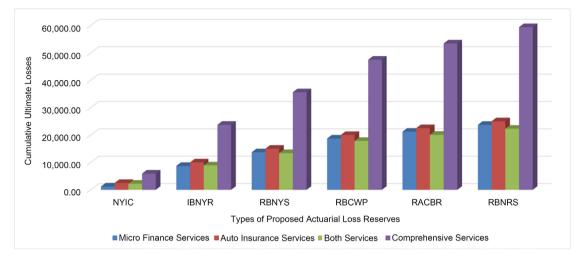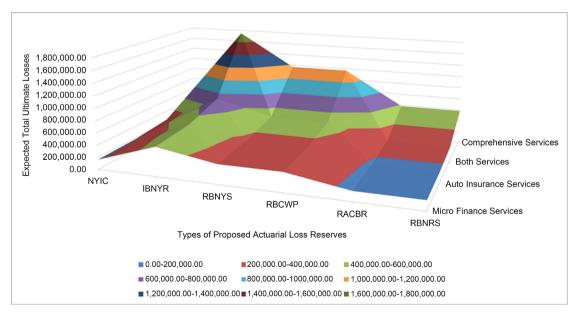
Table 14 is a result of the cumulative summation of the allocated actuarial loss reserve types presented on Table A10. This follows the presentation displayed on Figure 13.

Figure 13. Predicted expected ultimate losses.

Figure 13 reveals that out of the three main lines of business, Both Services recorded highest values for cumulative expected ultimate losses followed by Auto Insurance services line of business and the least being the Micro Finance services line of business. In addition to that, the comprehensive services as a representation of all the three services offered by the insurer indicate the highest values for the cumulative expected predicted losses.

4.3.7. Total Ultimate Losses by Reserve Categories

The Total Ultimate Losses by the proposed actuarial loss reserve categories have been presented below.

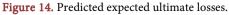**Figure 14.** Predicted expected ultimate losses.

Table A11 presents the recorded values for the Total Ultimate Losses by Reserve categories for the mentioned general insurance services displayed. Moreover, this has been validated by Figure 14.

4.3.8. Mathematical Computation of the Automated Actuarial Loss Reserves

The estimated Automated Actuarial Loss Reserves have been determined by the equation below.

$$AALR = CEUL - ALA \qquad (5)$$

where

- AALR Automated Actuarial Loss Reserves
- CEUL Cumulative Expected Ultimate Loss
- ALA Actual Loss Attained

As a result two scenarios namely computation of Automated Loss Reserves without case Reserves and Automated Loss Reserves with case Reserves have been considered in this paper as shown in the sections below.

4.3.9. Automated Actuarial Loss Reserves for Both without and with Case Reserves

Equation (5) leads to the determination of Automated Actuarial Loss Reserves without Case Reserves and with case reserves as shown below by Table 15 and Table 16.

Table 15. Estimated automated loss reserves without case reserves.

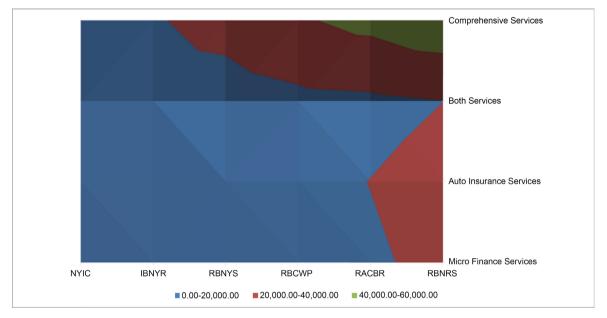| Estimated Automated Loss Reserves without case Reserves				
Reserves	Micro Finance Services	Auto Insurance Services	Both Services	Comprehensive Services
NYIC	1204.01	2505.52	2233.12	5942.65
IBNYR	7125.60	7623.78	7329.77	18,174.62
RBNYS	10,522.78	13,427.00	10,205.26	27,647.51
RBCWP	14,729.64	16,850.31	15,462.61	37,932.02
RACBR	18,861.50	20,158.52	16,890.62	45,498.59
RBNRS	22,143.15	23,473.09	19,942.66	53,845.35

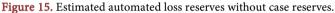Table 16. Estimated automated loss reserves with case reserves.

Estimated Automated Loss Reserves with case Reserves				
Reserves	Micro Finance Services	Auto Insurance Services	Both Services	Comprehensive Services
NYIC	161,275.11	178,652.72	242,255.62	582,183.45
IBNYR	487,338.90	536,065.38	723,404.01	1,746,897.02
RBNYS	330,664.98	365,721.40	485,435.34	1,180,129.11
RBCWP	334,871.84	369,144.71	488,300.80	1,190,413.62
RACBR	178,932.60	196,305.72	252,135.34	621,739.39
RBNRS	182,214.25	199,620.29	256,772.57	630,086.15

4.3.10. Estimated Automated Loss Reserves with Case Reserves

Both Table 15 and Table 16 indicate that all the calculated Actuarial loss reserves without and with Case Reserves with special attention to the three main lines of business as well and the Comprehensive services have yielded positive and large in magnitude proposed actuarial Loss reserves types.
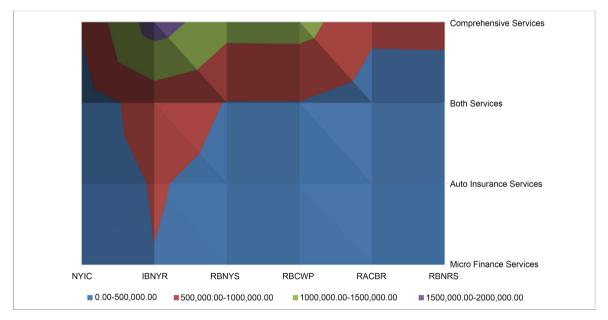

Figure 15. Estimated automated loss reserves without case reserves.

Figure 16. Estimated automated loss reserves with case reserves.

Figure 15 and Figure 16 reveal that across the proposed actuarial loss reserve types, Comprehensive Services sits on top with highest recorded values for Automated Actuarial Loss Reserves (AALR) with regards to both scenarios; without Case Reserves and with Case Reserves. The values for AALR are large and posi-

tive which indicates that my proposed model for estimation, prediction and evaluation of loss reserves is quite effective.

4.4. Fixed and Variable Interest Rates for Time Value of Computed Actuarial Types of Reserves

The fixed and Variable Interest rates for time value of computed actuarial types of reserves have been introduced here for evaluation of the loss reserves in the next 20 years.

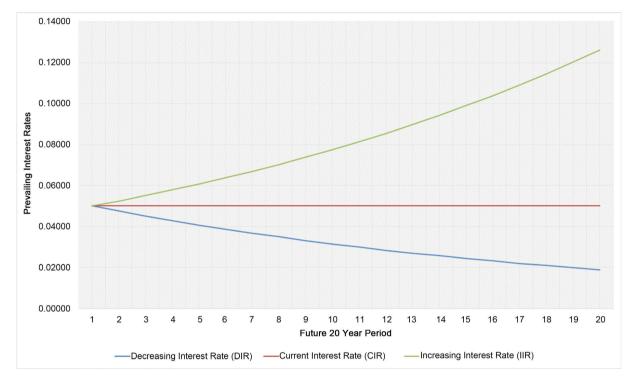**Figure 17.** Prevailing interest rates.

Figure 17 shows the fixed interest rate which is constant over the future 20-year period. The variable interest rates come from the decreasing interest rates which is shown on Figure 17 labeled in blue as it falls slowly from the fixed interest rate over the next 20-year time period. In addition to that, the increasing interest rate is the other type of variable interest rate which is indicated by an exponential rise over the next 20-year period.

4.5. Future Time Value of Comprehensive Automated Actuarial Loss Reserves

Since Figure 15 and Figure 16 indicate that both scenarios have yielded large quantum and positive loss reserves, in this study the Comprehensive Automated Actuarial Loss Reserves (CAALR) without reserves have been randomly selected for determination of the future time value of CAALR based on Net Present Values (NPVs) and Accumulated Values (ACVs) as shown in the subsections below.

4.5.1. Net Present Value for Comprehensive Automated Actuarial Loss Reserves

The Net Present Value for CAALR is computed using Equation (6) below.

$$\mathrm{NPV}\left(\mathrm{CAALR}\right) = \mathrm{CAALR}\left(1 + I_t\right)^{-n} \tag{6}$$

where NPV(CAALR) is the Net Present Value for CAALR without Case Reserves, I_t is fixed/variable interest rate at time t and n is the future period in years.

Figures 18-23 show that NPV for Comprehensive services where we derived the CAALR is positive and large after a long period of future 20 years thus presenting the insurer with large stake of reserves to cater for uncertain and future losses from the three main lines of business defined in this study.
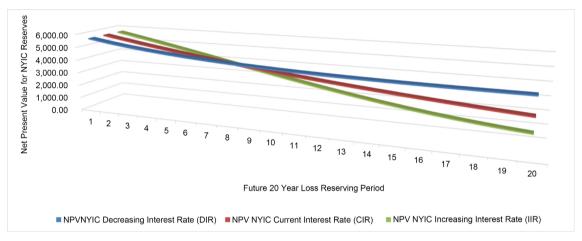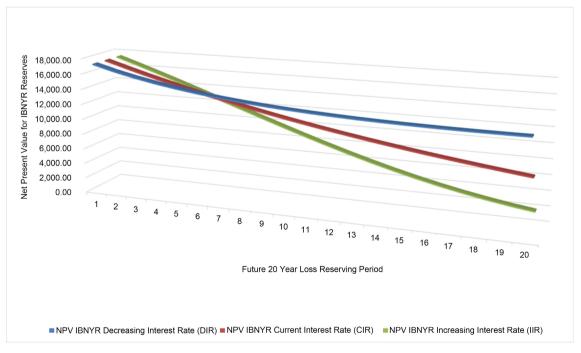

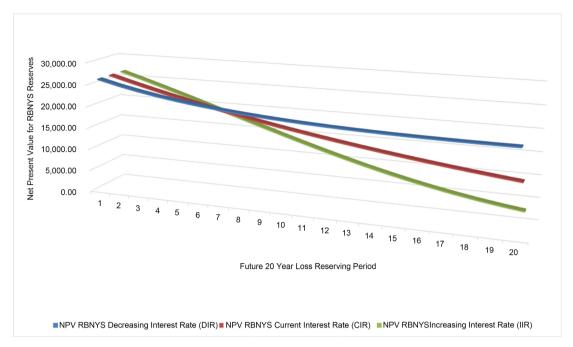**Figure 18.** NPV for NYIC reserves.

Figure 19. NPV for IBNYR reserves.

Figure 20. NPV for RBNYS reserves.

Figure 21. NPV for RBCWP reserves.

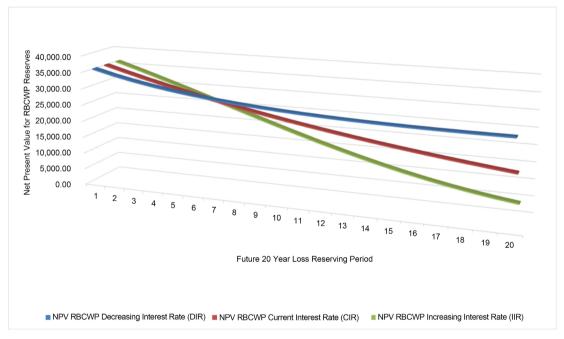### 4.5.2. Accumulated Value for Comprehensive Automated Actuarial Loss Reserves

The Accumulated Value for CAALR is computed using Equation (7) below.

$$\text{ACV}\left(\text{CAALR}\right) = \text{CAALR}\left(1 + I_t\right)^n \qquad (7)$$

where ACV (CAALR) is the Accumulated Value for CAALR without Case Reserves, I_t is fixed/variable interest rate at time t and n is the future period in years.
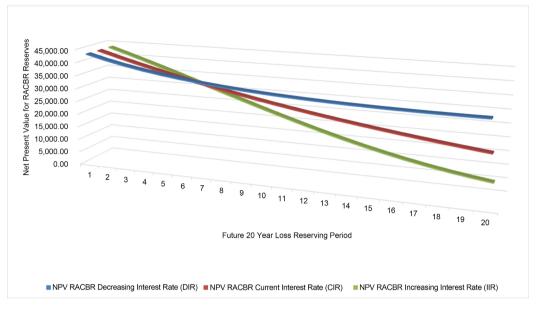
Figure 22. NPV for RACBR reserves.
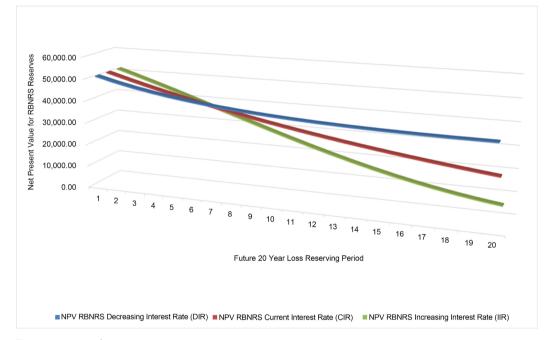
Figure 23. NPV for RBNRS reserves.

Figures 24-29 shows the ACV for Comprehensive services where we derived the CAALR's. Both the fixed and variable interest rates reveal that the CAALR are still positive and large in quantity after a long period of future 20-year period. This clearly shows that the insurer has inevitable capacity to generate more revenue from this model beyond future 20 years period.

4.6. Analysis of Future Claim Frequency, Claim Severity, Claim Inflation and Projected Ultimate Losses

After making an assumption that the Claim Frequency, Claim Severity and

Claim Inflation will remain constant over n, it is quite important to analyze the future behavior of these three main variables which are remarkably the backbone of our frequency, severity and inflation adjustment models. In addition to that it is also quite essential to articulate the behavior of future ultimate losses as well.
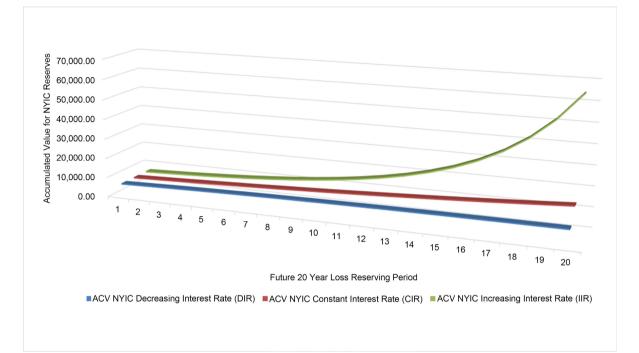
Figure 24. ACV for NYIC reserves.
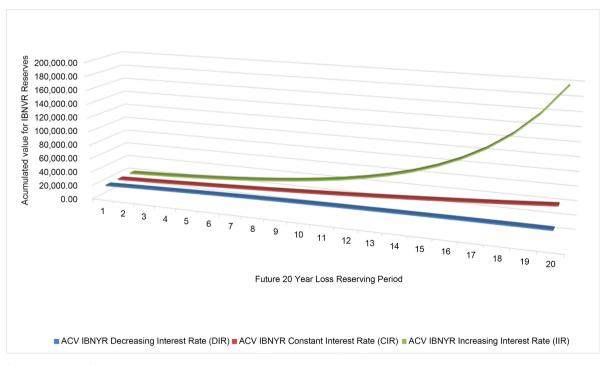
Figure 25. ACV for IBNYR reserves.
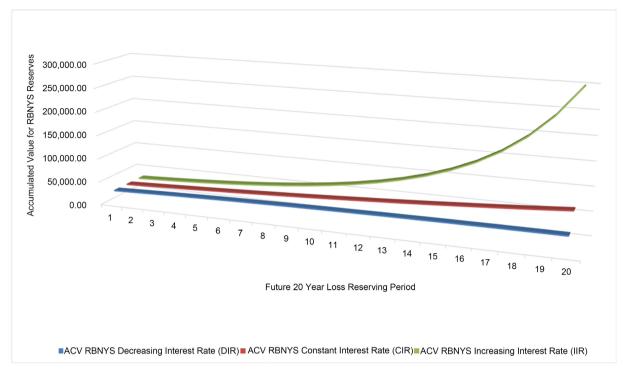
Figure 26. ACV for RBNYS reserves.
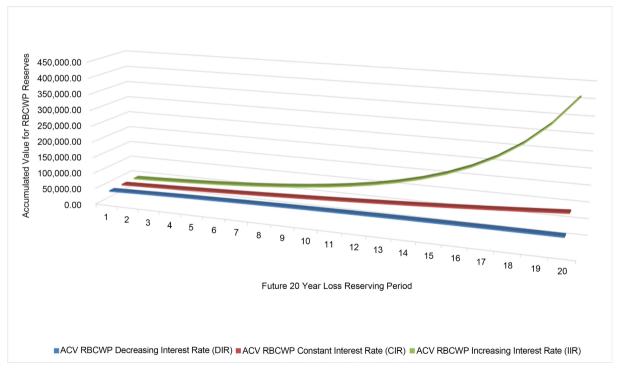
Figure 27. ACV for RBCWP reserves.

Figures 30-33 show the projected claim frequency, claim Severities, Claim Inflation and projected Ultimate Loss. In addition to that **Figure 30** shows that for the next ten years there seem to be the insurer needs to expect the greatest number of claims for the Auto Insurance services followed by Both services. It fol-

lows that the Micro Finance services are likely going to have small frequency claims in the next ten years. Moving next is Figure 31 which essentially indicates largest Severities to be anticipated by Both Insurance services followed by Auto Insurance services and the least being the Micro Finance services. Figure 32 shows the projected inflation index for the next ten years, which basically shows
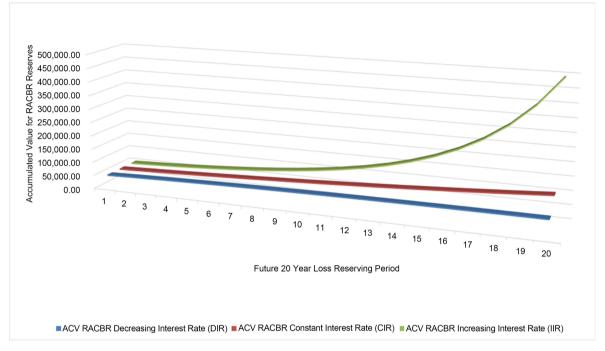
Figure 28. ACV for RACBR reserves.
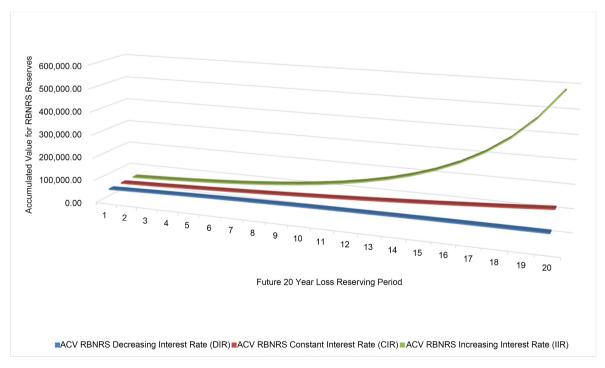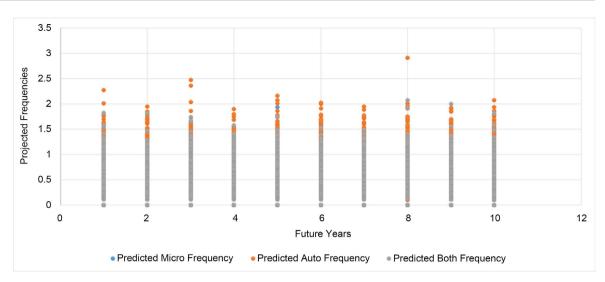
Figure 29. ACV for RBNRS reserves.

Figure 30. Projected frequencies.
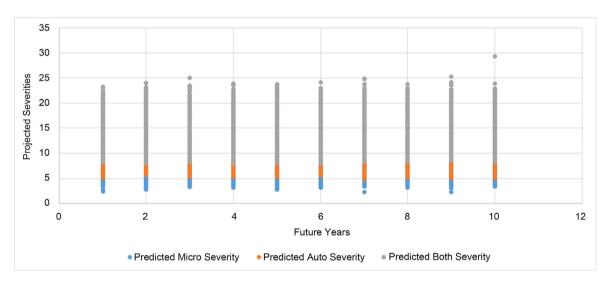
Figure 31. Projected severities.
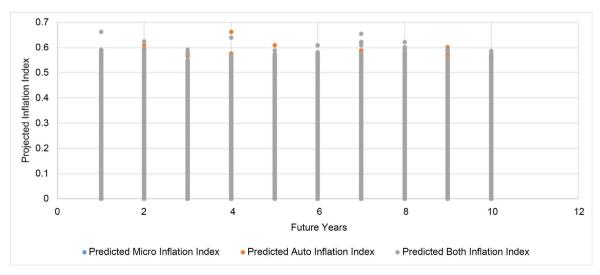
Figure 32. Projected inflation rates.
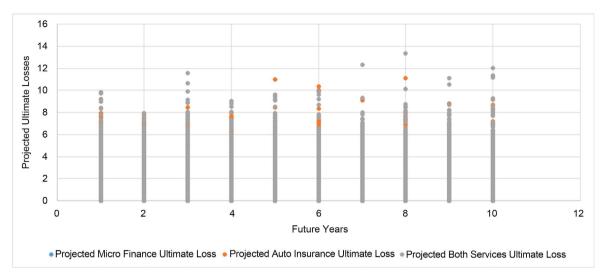
Figure 33. Projected ultimate losses.

that claim inflation for the Both services line of business will essentially be the highest followed by the Auto Insurance services line of business and the last being the Micro Finance services line of business. Additionally, **Figure 33** shows the projected ultimate losses expected from the three main lines of business accordingly. From the same figure, it is clear that Both Services will carry the largest stake of projected ultimate losses with the two other lines of business being quite small. This apparently shows that in the next ten years many people will join the Both Services policyholder category. Many policyholders in this line of business will come from the existing policyholders from the two other lines of businesses and the new policyholders preferring the Both services line of business.

Comprehensive Automated Actuarial Loss Reserves Scaling and Actual Loss Assessment

The Comprehensive Automated Actuarial Loss Reserves have been scaled down considering both scenarios, the first being the CAALR without Case Reserves and the second being the CAALR with Case Reserves at the following thresholds T_i 90%, 80%, 70%, 60% and 50% respectively as shown by **Table A12**.

From there the following Equation (8) is considered;

$$\text{CAALR}_{\text{scaled down}} = \text{CAALR}_{T_i} - \text{CAL} \tag{8}$$

where CAALR is the Comprehensive Automated Actuarial Loss Reserves derived from the Comprehensive Services, T_i is the scaled down threshold and CAL being the Comprehensive Actual Loss derived from the summation of all the actual losses experienced from the three main lines of business.

Table A12 indicates that the scaled down CAALR at all the thresholds still produced large and positive values across all the given thresholds. Furthermore, if the insurer is to reduce the current given CAALR to given thresholds from the 90% of the current CAALR up to 50% of the current CAALR, the insurer is still capable of retaining large and positive scores of the difference from CAL with

regards to both scenarios with and without Case Reserves. This ultimately serves as both sensitive and scenario tests towards model validation.

5. Innovation for the Developed Automated Actuarial Data Analytics Based Inflation Adjusted Frequency Severity Loss Reserving Model

In this section some discussions pertaining to the innovations brought by the developed model in this paper are presented as follows.

5.1. Comparisons between the Proposed Model and the Traditional Chain Ladder Method

The proposed model has better prediction accuracy of the outstanding liabilities emerging from the claims raised by the policyholders in the three main lines of business presented in this paper compared to traditional chain ladder models and they can also deal with structured and unstructured information, such as individual-level claim data, and outperform traditional chain ladder models. Additionally, they can capture complex patterns in the development of the claims and reduce estimation uncertainty [36]. Moreover, the proposed model requires minimal feature engineering and expert input, and can be automated to produce forecasts more frequently than manual workflows. In addition to that, the proposed model allows for joint modeling of paid losses and claims outstanding, and incorporation of heterogeneous inputs [37]. They display a general trend toward ever-increasing complexity and data-intensity, which can capture the dynamics of the insurance market [38].

5.2. Risk Mitigation and Reduction Model

The proposed model reduces the liquidity risk by creating sufficient cash to settle the comprehensive claims as they are incurred from both microfinance policy and car insurance policy. Due to the uncertainty of general insurer's liabilities in amount and timing this enables the insurer to maintain reasonable liquidity level to meet catastrophic events. As a result, the risk that changes in the value of the assets, or the liabilities are offset.

5.3. Adherence to IFRS 17 Regulations

The presented model supports all the accounting concepts such as the going concern concept, realization concept, accruals concept and so on. The autonomous augmentation of the presented three main lines of business in this study using artificial intelligence makes the insurance company continue to make foreseeable profits in future as shown by the positive NPVs and ACVs of the CAALR. As a result when given full implementation and careful revision, the model conforms to IFRS17 regulations [39] and [40].

5.4. Development of the Almost Zero-Based Reinsurance Model

The emergence and existence of the proposed actuarial loss reserve types and the

proposed allocation and distributions across the three main lines of business creates large stake of funds and a large pool of reserves within these lines of businesses, which ensures that there are enough funds set aside for catastrophic reserves. Moreover the insurer is capable of perceiving quick and much faster comprehensive claim settlement with no, little or minimum reinsurance cost. In short that essentially means that all forms of reinsurance such as facultative reinsurance, excess of loss reinsurance, proportional reinsurance, financial risk reinsurance are reduced and this then the proposed model may bring fronting and ceding to almost zero when carefully implemented [41].

5.5. Development of the Almost Zero Reporting-Settlement Delay Based Model

Since our model is AI-based, it ensures that once a claim is incurred and or any amount of money is requested or claimed from the three main lines of businesses, it is settled instantly and autonomously and hence this brings down the reporting delay and settlement delay to zero. Consequently, when the proposed model is given full implementation the role of broking and intermediary services may not be applicable which is both time and cost saving to the insurer [42].

5.6. Future Work

The model presented in this paper is a starting point towards Automated Actuarial Pricing and Underwriting using Artificial Intelligence. Moreover it marks the genesis of automating general insurance and life insurance on the same platform using Artificial Intelligence.

5.7. Discussion

In this paper, an Automated Actuarial Loss Reserving Model using ten machine learning models has been developed with regards to three main lines of business defined in this study. The Random Forest (RANGER) have obtained the highest score for Aggregate expected ultimate loss which essentially became the best model out of the ten used in the study. This best machine learning algorithm has been further conscripted towards mathematical computation of Automated Actuarial Loss reserves using both scenarios with and without case reserves. In both cases, after scaling the CAALR in presented thresholds, large and positive chunks of reserves after subtracting the comprehensive actual losses were obtained, which subsequently presents the insurer with large stake of funds available for meeting future liabilities and related claims.

6. Conclusion

The development of the Automated Actuarial Data Analytics Based Inflation Adjusted Frequency Severity Loss Reserving Model whilst automating autonomously microfinance services, car insurance services and Both Services using Artificial Intelligence can stand out to be one of the most useful achievements in

this current period when given full implementation. Eventually, this enables the insurance companies to improve and increase both the scope and scale of insurance processes whilst ensuring effectiveness, efficiency and economical ways of running the business. Given full-scale operation, the presented model in this paper is capable of reducing the high gearing levels by retaining large stakes of reserves at both micro and macro level reserving. Through this way, this model then coerces the insurer to make progressive and realizable revenues with quite large profits with almost zero reinsurance and adhering to IFRS17 regulations remarkably.

Data Availability

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

Acknowledgements

Special thanks go to Dr. Charles Chimedza, Dr. Florance Matarise, Dr. Sheunesu Munyira and other 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

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Richman, R. (2018) AI in Actuarial Science. *SSRN Electronic Journal*.

[2] Balona, C. and Richman, R. (2020) The Actuary and IBNR Techniques: A Machine Learning Approach. *SSRN Electronic Journal*.

[3] Panlilio, A., Canagaretna, B., Perkins, S., du Preez, V. and Lim, Z. (2018) Practical Application of Machine Learning within Actuarial Work. Institute and Faculty of Actuaries, London.

[4] Wüthrich, M.V. (2018) Machine Learning in Individual Claims Reserving. *Scandinavian Actuarial Journal*, **2018**, 465-480. https://doi.org/10.1080/03461238.2018.1428681

[5] Pentikäinen, T. and Rantala, J. (1992) A Simulation Procedure for Comparing Different Claims Reserving Methods. *ASTIN Bulletin*, **22**, 191-216. https://doi.org/10.2143/ast.22.2.2005115

[6] Radtke, M., Schmidt, K.D. and Schnaus, A. (2016) Handbook on Loss Reserving. Springer. https://doi.org/10.1007/978-3-319-30056-6

[7] Bjarnason, T. and Sjögren, M. (2014) Insurance Loss Reserving. LUP Student Papers, Lund University.

[8] Murphy, D. (1994) The Bornhuetter-Ferguson Principle. *North American Actuarial Journal*, **12**, 377-379.

[9] Mack, T. (1993) Distribution-Free Calculation of the Standard Error of Chain Ladder Reserve Estimates. *ASTIN Bulletin*, **23**, 213-225.

https://doi.org/10.2143/ast.23.2.2005092

[10] England, P.D. (2018) Actuarial Reserving: An Overview of Chain Ladder Method-ologies. *CAS Monograph Series*, **5**, 1-25.

[11] Bartley, J.W. and Boardman, C.M. (1990) The Relevance of Inflation Adjusted Ac-counting Data to the Prediction of Corporate Takeovers. *Journal of Business Finance & Accounting*, **17**, 53-72.
https://doi.org/10.1111/j.1468-5957.1990.tb00549.x

[12] Schwert, G.W. (1981) The Adjustment of Stock Prices to Information about Infla-tion. *The Journal of Finance*, **36**, 15-29.
https://doi.org/10.1111/j.1540-6261.1981.tb03531.x

[13] Hunsicker, C.F. (2023) On Machine Learning Approaches to Forecast Non-Life In-surers Loss Reserves. Master's Thesis, University of Twente.

[14] Blier-Wong, C., Cossette, H., Lamontagne, L. and Marceau, E. (2020) Machine Learning in Property and Casualty Insurance: A Review for Pricing and Reserving. *SSRN Electronic Journal*.

[15] Xiong, L., Manathunga, V., Luo, J., Dennison, N., Zhang, R. and Xiang, Z. (2023) Autoreserve: A Web-Based Tool for Personal Auto Insurance Loss Reserving with Classical and Machine Learning Methods. *Risks*, **11**, Article 131.
https://doi.org/10.3390/risks11070131

[16] Tuininga, F. (2022) A Machine Learning Approach for Modeling Frequency and Severity. Master's Thesis, University of Twente.

[17] Wilson, A.A., Nehme, A., Dhyani, A. and Mahbub, K. (2024) A Comparison of Generalised Linear Modelling with Machine Learning Approaches for Predicting Loss Cost in Motor Insurance. *Risks*, **12**, Article 62.
https://doi.org/10.3390/risks12040062

[18] Panjee, P. and Amornsawadwatana, S. (2024) A Generalized Linear Model and Machine Learning Approach for Predicting the Frequency and Severity of Cargo Insurance in Thailand's Border Trade Context. *Risks*, **12**, Article 25.
https://doi.org/10.3390/risks12020025

[19] Taylor, G. (2000) Loss Reserving: An Actuarial Perspective. Kluwer Academic Pub-lishers.

[20] England, P.D. and Verrall, R.J. (2002) Stochastic Claims Reserving in General In-surance. *British Actuarial Journal*, **8**, 443-518.
https://doi.org/10.1017/s1357321700003809

[21] Wüthrich, M.V. and Merz, M. (2008) Stochastic Claims Reserving Methods in In-surance. Wiley.

[22] Meyers, G. (2007) The Common Pitfalls of Loss Reserving. Casualty Actuarial Soci-ety, Arlington, 1-32.

[23] Kuo, W., Tsai, C.-F. and Chen, Y.-K. (2018) A Comparison of Machine Learning Methods for Insurance Claims Prediction. *Soft Computing*, **22**, 2769-2783.

[24] Richman, R. and Wüthrich, M.V. (2020) A Neural Network Extension of the Lee-Carter Model to Multiple Populations. *Annals of Actuarial Science*, **14**, 72-91.

[25] Gabrielli, A. and Wüthrich, M.V. (2018) Bayesian Neural Networks for Insurance Claims Reserving. *Risks*, **6**, Article 18.

[26] Antonio, K. and Plat, R. (2013) Micro-Level Stochastic Loss Reserving for General Insurance. *Scandinavian Actuarial Journal*, **2014**, 649-669.
https://doi.org/10.1080/03461238.2012.755938

[27] Clark, D.R. and Rangelova, D. (2015) Accident Year/Development Year Interactions. Casualty Actuarial Society. Arlington.

[28] Parodi, P. (2013) Triangle-Free Reserving. *British Actuarial Journal*, **19**, 168-218. https://doi.org/10.1017/s1357321713000093

[29] Catarina Ferreira de Jesus de, S. (2023) IBNR Techniques in Health Insurance: A Machine Learning Approach. Ph.D. Dissertation, Universidade NOVA de Lisboa.

[30] Cipra, T. and Hendrych, R. (2023) Holt-Winters Method for Run-off Triangles in Claims Reserving. *European Actuarial Journal*, **13**, 815-836. https://doi.org/10.1007/s13385-023-00348-2

[31] Matthews, S. and Hartman, B. (2022) Machine Learning in Ratemaking, an Application in Commercial Auto Insurance. *Risks*, **10**, Article 80. https://doi.org/10.3390/risks10040080

[32] Kothari, C.R. (2004) Research Methodology: Methods and Techniques. New Age International.

[33] Gesmann, M., Murphy, D., Zhang, Y.W., Carrato, A., Crupi, G., Dutang, C., *et al.* (2023) Statistical Methods and Models for Claims Reserving in General Insurance [R Package ChainLadder Version 0.2.11].

[34] Carrato, A., Concina, F., Gesmann, M., Murphy, D., Wuthrich, M. and Zhang, W. (2015) Claims Reserving with R: ChainLadder-0.2.2 Package Vignette.

[35] de Oliveira, M.C.F. and Levkowitz, H. (2003) From Visual Data Exploration to Visual Data Mining: A Survey. *IEEE Transactions on Visualization and Computer Graphics*, **9**, 378-394. https://doi.org/10.1109/tvcg.2003.1207445

[36] Kuo, K. (2019) Deeptriangle: A Deep Learning Approach to Loss Reserving. *Risks*, **7**, Article 97. https://doi.org/10.3390/risks7030097

[37] Zhang, Y., Wen, J., Yang, G., He, Z. and Wang, J. (2019) Path Loss Prediction Based on Machine Learning: Principle, Method, and Data Expansion. *Applied Sciences*, **9**, Article 1908. https://doi.org/10.3390/app9091908

[38] Taylor, G. (2019) Loss Reserving Models: Granular and Machine Learning Forms. *Risks*, **7**, Article 82. https://doi.org/10.3390/risks7030082

[39] Yousuf, W., Stansfield, J., Malde, K., Mirin, N., Walton, R., Thorpe, B., *et al.* (2021) The IFRS 17 Contractual Service Margin: A Life Insurance Perspective. *British Actuarial Journal*, **26**, e2. https://doi.org/10.1017/s1357321721000015

[40] Araiza Iturria, C.A., Godin, F. and Mailhot, M. (2021) Tweedie Double GLM Loss Triangles with Dependence within and across Business Lines. *European Actuarial Journal*, **11**, 619-653. https://doi.org/10.1007/s13385-021-00267-0

[41] Shiu, Y.M., Wang, C.F., Adams, A. and Shin, Y.C. (2012) On the Determinants of Derivative Hedging by Insurance Companies: Evidence from Taiwan Region. *Asian Economic and Financial Review*, **2**, 538.

[42] Crevecoeur, J., Antonio, K., Desmedt, S. and Masquelein, A. (2023) Bridging the Gap between Pricing and Reserving with an Occurrence and Development Model for Non-Life Insurance Claims. *ASTIN Bulletin*, **53**, 185-212. https://doi.org/10.1017/asb.2023.14

Appendix

Appendix A1. Data, Associated Models and Variables

Table A1. Data, associated models and variables.

Data	Model Types	Dependent Variables
Micro Finance Services	M1-Frequency model, M2-Severity model, M3-Inflation Adjustment model	Micro Number of Claims (M1), Micro Claim Amount (M2), Inflation Index (M3)
Auto Insurance Services	M1-Frequency model, M2-Severity model, M3-Inflation Adjustment model	Auto Number of Claims (M1), Auto Claim Amount (M2), Inflation Index (M3)
Both Services	M1-Frequency model, M2-Severity model, M3-Inflation Adjustment model	Both Number of Claims (M1), Both Claim Amount (M2), Inflation Index (M3)

Table A2. Lines of Business and associated independent variables.

Data	Independent Variables
Micro Finance Services	Age, Gender, Marital Status, Sales Channel, Credit Score, Residence, Occupation, Education, Employment Status, Amount Invested, Amount Requested, Micro claim category, Previous Loan history, Micro Loan Amount, monthly income,
Auto Insurance Services	Age, Gender, Marital Status, Sales Channel, Credit Score, Residence, Occupation, Education, Employment Status, Vehicle Size, Vehicle Type, Auto Claim Type, Vehicle Route, Driving Record, Vehicle Value
Both Services	Age, Gender, Marital Status, Sales Channel, Credit Score, Residence, Occupation, Education, Employment Status, Both Services Package, Both Services Claim Type, Both

Appendix A2. Machine Learning Algorithms, Associated R Packages and Hyper-Parameters

Table A3. 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
Generalized Boosting Machines (GBM)	gbm	n.trees: 100, interaction.depth: 3, n.minobsinnode: 10
Extreme Gradient Boosting (XGB)	xgboost	xgboost maximum depth: 3, number of rounds: 100
Least Angle Regression (LAR)	caret	Method: lars
Extreme Learning Machines (ELM)	elm	ELM-Type: regression, Hidden units: 20, activation function: sigmoid
Robust Regression Method (RRM)	robustbase	no hyper parameters used
Artificial Neural Network (ANN)	nnet	Size: 2, decay: 5e-4, maximum iterations: 200

Appendix A3. Micro Finance Services Triangle Loss Reserving Chain Ladder

Table A4. Micro finance services triangle loss reserving chain ladder.

MICRO FINANCE SERVICES TRIANGULATION LOSS RESERVING CHAIN LADDER METHOD										
Development period										
Accident Years	0	1	2	3	4	5	6	7	8	9	10
2010	18,674.98	11,923.36	5039.73	2923.81	2190.11	2747.68	4152.39	2307.47	763.69	531.03	315.62
2011	11,014.25	6247.69	2684.50	9544.14	3276.56	6228.35	2456.35	493.36	1054.23	352.59	
2012	8922.42	4412.05	12,223.26	3619.96	8419.61	2896.78	1825.81	1504.45	526.60		
2013	4575.25	14,258.03	4761.39	2472.26	2645.39	2212.79	1595.22	3935.92			
2014	16,972.47	3591.50	3818.61	2361.79	1943.69	3796.29	8109.31				
2015	3694.17	2800.28	2310.09	4828.53	3848.99	6500.04					
2016	3761.39	4360.01	6974.87	4421.60	4839.51						
2017	4072.46	10,792.96	4198.66	2189.02							
2018	12,013.95	6471.42	2184.66								
2019	9720.45	3637.31									
2020	3431.61										

Appendix A4. Auto Insurance Services Triangle Loss Reserving Chain Ladder

Table A5. Auto insurance services triangle loss reserving chain ladder.

AUTO INSURANCE SERVICES TRIANGULATION LOSS RESERVING CHAIN LADDER METHOD										
Development period										
Accident Years	0	1	2	3	4	5	6	7	8	9	10
2010	28,014.09	17,882.91	7413.13	4368.75	3347.79	4220.92	6369.84	3552.97	1221.29	760.47	450.15
2011	16,572.47	9205.82	4084.55	14,357.95	5095.28	9342.94	3763.57	775.49	1460.54	523.44	
2012	13,500.46	6440.85	18,179.61	5541.21	12,445.57	4263.67	2640.27	2051.24	808.88		
2013	6864.01	21,651.74	6960.85	3643.53	3998.27	3273.48	2384.15	2935.33			
2014	25,179.04	5362.74	5799.39	3472.70	2955.13	2828.36	12,346.77				
2015	5648.76	4277.38	3515.28	7425.06	5837.82	9669.30					
2016	5678.42	6512.07	10,567.46	6732.74	7039.38						
2017	5974.35	16,039.06	6249.64	3330.51							
2018	17,947.90	9670.44	3414.63								
2019	14,462.90	5491.15									
2020	5223.88										

Appendix A5. Both Services Triangle Loss Reserving Chain Ladder

Table A6. Both services triangle loss reserving chain ladder.

BOTH SERVICES TRIANGULATION LOSS RESERVING CHAIN LADDER METHOD										
Development period										
Accident Years	0	1	2	3	4	5	6	7	8	9	10
2010	37,621.43	24,214.60	9987.15	5808.40	4509.05	5596.08	8302.00	4683.55	1573.69	965.59	589.76
2011	22,236.10	12,433.56	5335.24	19,091.90	6811.94	12,426.13	5007.62	1031.05	2029.32	759.49	
2012	18,065.38	8586.35	24,664.53	7315.17	16,757.48	5712.27	3528.20	2887.54	1110.60		
2013	9304.65	28,890.47	9439.82	4724.91	5157.03	4405.93	3148.81	3935.92			
2014	33,842.41	7360.50	7694.76	4662.14	4074.20	3796.29	16,408.16				
2015	7698.26	5623.88	4642.92	9749.89	7583.30	13,136.81					
2016	7518.08	8816.27	14,108.57	9032.27	9458.99						
2017	8032.22	21,460.72	8303.34	4562.49							
2018	24,133.79	12,732.51	4512.99								
2019	19,325.94	7433.06									
2020	6989.89										

Appendix A6. Comprehensive Services Triangle Loss Reserving Chain Ladder

Table A7. Comprehensive services triangle loss reserving chain ladder.

COMPREHENSIVE TRIANGULATION LOSS RESERVING CHAIN LADDER METHOD										
Development period										
Accident Years	0	1	2	3	4	5	6	7	8	9	10
2010	84,310.50	54,020.87	22,440.01	13,100.97	10,046.95	12,564.67	18,824.24	10,543.99	3558.67	2257.09	1355.54
2011	49,822.82	27,887.08	12,104.29	42,993.99	15,183.79	27,997.42	11,227.54	2299.90	4544.08	1635.51	
2012	40,488.26	19,439.24	55,067.40	16,476.33	37,622.66	12,872.73	7994.28	6443.23	2446.08		
2013	20,743.90	64,800.24	21,162.06	10,840.69	11,800.70	9892.20	7128.18	10,807.17			
2014	75,993.92	16,314.73	17,312.76	10,496.63	8973.02	10,420.93	36,864.24				
2015	17,041.19	12,701.53	10,468.29	22,003.47	17,270.11	29,306.15					
2016	16,957.89	19,688.34	31,650.90	20,186.61	21,337.88						
2017	18,079.03	48,292.74	18,751.64	10,082.01							
2018	54,095.64	28,874.38	10,112.28								
2019	43,509.29	16,561.51									
2020	15,645.37										

Appendix A7. Total Expected Ultimate Losses from the Augmented Services Offered by the Insurer

Table A8. Total expected ultimate losses from the augmented services offered by the insurer.

	Total Automated Actuarial Ultimate Loss Predictions			
	Micro Finance Services	Auto Insurance Services	Both Services	Comprehensive Services
ML Methods	Total AALR predictions	Total AALR predictions	Total AALR predictions	Aggregate AALR predictions
GLM	11,756.26	24,905.63	21,953.31	58,615.20
GAM	11,826.99	24,981.77	21,730.12	58,538.88
RPART	11,652.60	24,824.27	21,940.53	58,417.40
RANGER	12,040.11	25,055.17	22,331.23	59,426.51
GBM	11,864.68	24,812.98	21,980.00	58,657.66
XGB	11,731.08	24,918.66	21,935.16	58,584.90
LAR	11,794.26	24,790.51	22,345.15	58,929.92
ELM	11,831.94	24,774.00	22,261.64	58,867.58
RRM	10,793.25	23,554.02	20,442.17	54,789.44
ANN	11,689.83	24,693.88	21,775.60	58,159.31

Appendix A8. Allocation and Distribution of the Case Reserves

Table A9. Allocated case reserves to augmented general insurance services.

	Allocation and distribution of the Case Reserves		
Reserves	Micro Finance Services	Auto Insurance Services	Both Services
NYIC	160,071.10	176,147.20	240,022.50
IBNYR	480,213.30	528,441.60	720,067.50
RBNYS	320,142.20	352,294.40	480,045.00
RBCWP	320,142.20	352,294.40	480,045.00
RACBR	160,071.10	176,147.20	240,022.50
RBNRS	160,071.10	176,147.20	240,022.50

Appendix A9. Predicted Ultimate Losses by Reserve Categories

Table A10. Predicted ultimate losses by reserve categories.

	Predicted Ultimate Losses by Reserve categories			
Reserves	Micro Finance Services	Auto Insurance Services	Both Services	Comprehensive Services
NYIC	1204.01	2505.52	2233.12	5942.65
IBNYR	3612.03	7516.55	6699.37	17,827.95
RBNYS	2408.02	5011.03	4466.25	11,885.30
RBCWP	2408.02	5011.03	4466.25	11,885.30
RACBR	1204.01	2505.52	2233.12	5942.65
RBNRS	1204.01	2505.52	2233.12	5942.65

Appendix A10. Total Ultimate Losses by Reserve Categories Including Case Reserves

Table A11. Total ultimate losses by reserve categories including case reserves.

	Total Ultimate Losses by Reserve categories			
Reserves	Micro Finance Services	Auto Insurance Services	Both Services	Compreehensive Services
NYIC	161,275.11	178,652.72	242,255.62	582,183.45
IBNYR	483,825.33	535,958.15	726,766.87	1,746,550.35
RBNYS	322,550.22	357,305.43	484,511.25	1,164,366.90
RBCWP	322,550.22	357,305.43	484,511.25	1,164,366.90
RACBR	161,275.11	178,652.72	242,255.62	582,183.45
RBNRS	161,275.11	178,652.72	242,255.62	582,183.45

Appendix A11. Comprehensive Automated Actuarial Loss Reserves Scaling and Actual Loss Assessment

Table A12. Comprehensive automated actuarial loss reserves scaling and actual loss assessment.

	Scaled down CAALR without Case Reserves						Scaled down CAALR without Case Reserves-Actual Comprehensive Loss				
Comprehensive Services	90%	80%	70%	60%	50%	Comprehensive Loss	90%	80%	70%	60%	50%
5942.65	5348.39	4754.12	4159.86	3565.59	2971.33	0.00	5348.39	4754.12	4159.86	3565.59	2971.33
18,174.62	16,357.16	14,539.69	12,722.23	10,904.77	9087.31	5595.99	10,761.17	8943.71	7126.24	5308.78	3491.32
27,647.51	24,882.75	22,118.00	19,353.25	16,588.50	13,823.75	8008.40	16,874.35	14,109.60	11,344.85	8580.10	5815.35
37,932.02	34,138.82	30,345.62	26,552.41	22,759.21	18,966.01	9609.19	24,529.63	20,736.43	16,943.23	13,150.03	9356.82
45,498.59	40,948.73	36,398.87	31,849.01	27,299.16	22,749.30	7985.27	32,963.47	28,413.61	23,863.75	19,313.89	14,764.03
53,845.35	48,460.82	43,076.28	37,691.75	32,307.21	26,922.68	5581.16	42,879.66	37,495.12	32,110.59	26,726.05	21,341.52

	Scaled down CAALR with Case Reserves						Scaled down CAALR with Case Reserves-Actual Comprehensive Loss				
Comprehensive Reserves	90%	80%	70%	60%	50%	Comprehensive Loss	90%	80%	70%	60%	50%
582,183.45	523,965.11	465,746.76	407,528.42	349,310.07	291,091.73	0.00	523,965.11	465,746.76	407,528.42	349,310.07	291,091.73
1,746,897.02	1,572,207.32	1,397,517.61	1,222,827.91	1,048,138.21	873,448.51	5595.99	1,566,611.33	1,391,921.63	1,217,231.92	1,042,542.22	867,852.52
1,180,129.11	1,062,116.19	944,103.28	826,090.37	708,077.46	590,064.55	8008.40	1,054,107.79	936,094.88	818,081.97	700,069.06	582,056.15
1,190,413.62	1,071,372.26	952,330.90	833,289.53	714,248.17	595,206.81	9609.19	1,061,763.07	942,721.71	823,680.35	704,638.99	585,597.62
621,739.39	559,565.45	497,391.51	435,217.57	373,043.64	310,869.70	7985.27	551,580.19	489,406.25	427,232.31	365,058.37	302,884.43
630,086.15	567,077.54	504,068.92	441,060.31	378,051.69	315,043.08	5581.16	561,496.38	498,487.76	435,479.15	372,470.53	309,461.92