12th International Conference on Business Economics, Marketing & Management Research (BEMM-2025)
Proceedings Book Series -PBS-Vol 22, pp.47-52

Predictive Modeling for Health Insurance Reserving in Morocco: A Neural Network and Risk Profiling Perspective

Ayyoub SAOUDI[#1], Ghita HAJRAOUI[#2], Kenza HAJRAOUI[#3], Jamal ZAHI[#4]

1 University Hassan 1st, Faculty of Economics and Management, LM2CE, Settat, Morocco.

2 University Hassan II, ESSEM Business School, LARSEMO, Casablanca, Morocco

3 University Moulay Ismail,University School of Technology,LAREGS, Meknès, Morocco

4 University Hassan 1st, Faculty of Economics and Management, LM2CE, Settat, Morocco

Email 1 - a.saoudi@uhp.ac.ma

Email 2 - ghita.hajraoui@gmail.com

Email 3 - k.hajraoui@umi.ac.ma

Email 4 - zahi71@hotmail.com

***Abstract*—This research investigates the efficacy of Artificial Neural Networks (ANN) in improving health insurance reserving prediction within the context of Morocco. Employing meticulous methodologies encompassing comprehensive data collection, rigorous preprocessing, and ANN model development, the study demonstrates the capability of ANN to provide precise estimations of insurance reserves. Furthermore, through the utilization of risk class profiling techniques facilitated by decision trees, pivotal determinants affecting policyholders' risk levels, notably age, presence of chronic illnesses, and gender, are identified and analyzed. The findings underscore the significance of an integrated approach that amalgamates advanced modeling techniques with risk profiling methodologies for efficient health insurance reserving management. This study contributes to augmenting the financial robustness and longevity of insurance enterprises operating in the Moroccan health insurance domain while elevating the caliber of services extended to policyholders.**

***Keywords*—Artificial Neural Networks, Decision Trees, Health Insurance, Modeling techniques, Morocco, Reserving Prediction, Risk Profiling.**

I. INTRODUCTION

The Moroccan health insurance sector is integral to safeguarding the well-being of its citizens by offering crucial healthcare coverage. However, the effective management of this sector heavily relies on the accurate prediction of health insurance reserves.

In the past, traditional methods of health insurance reserving prediction have been employed, yet they often fall short in terms of accuracy and efficiency. Consequently, inaccuracies in reserve predictions can lead to financial instability within insurance companies, threatening their ability to meet future obligations and potentially undermining the entire healthcare system. [1]

ISSN: 2961-6611

Recognizing these challenges, there is a pressing need to explore advanced techniques that can improve the accuracy and reliability of health insurance reserving prediction. Artificial Neural Networks (ANN) present a promising solution in this regard that allows to analyze vast amounts of data and uncover hidden patterns that may elude traditional statistical methods.[2]

This study aims to unlock the potential of ANN technology to enhance the accuracy of health insurance reserving prediction specifically tailored to the Moroccan context.

II. Health Insurance Reserving

Health insurance plays a critical role in mitigating rising healthcare costs and ensuring equitable access to healthcare services. Accurate reserve estimation is fundamental for insurers to fulfill their obligations to policyholders and to optimize their risk management.

Traditionally, deterministic methods, such as the Chain Ladder technique [3], have been employed for technical reserve estimation in health insurance. While historically reliable, these methods struggle to capture the complexities of individual characteristics and the dynamic risk landscape. The growing volume of healthcare data renders conventional models inadequate in fully exploiting this information.

III. Artificial Neural Networks for EnhancedPerformance in claim Reserving prediction

Machine learning (ML) methods applied to individual reserving offer a dynamic solution for enhanced risk management by safeguarding solvency and adapting to evolving trends and changes in policyholder risk profiles. Pioneering work by P. Mulquiney [4] introduced neural networks to individual reserving, marking a significant advancement. However, new challenges have arisen, particularly regarding the static or dynamic nature of data. Accounting for the chronological evolution of certain variables becomes crucial for accurate claims payment modeling.

Artificial Neural Networks draw inspiration from the intricate neural structure of the human brain[5]. In the realm of reserving, they serve as predictive tools, enabling the estimation of diverse outcomes like projected yearly medical claims within insurance firms. ANN possess the ability to learn from past experiences and extrapolate insights, rendering them adept at forecasting forthcoming values using historical data and model outcomes. ANN demonstrates the potential to enhance precision and computational efficiency for forecasting tasks pertinent to reserving, particularly within sectors such as insurance. [5]

IV. Methodology

The estimation of individual reserving in health insurance using Artificial Neural Networks necessitates a systematic and comprehensive approach, commencing with meticulous data collection and extending to a rigorous evaluation of the resultant model's performance.

ISSN: 2961-6611

Our data originates from a health insurance company that has chosen to remain anonymous for confidentiality reasons. The dataset, comprising settlements from the year 2019, covers a variety of variables capturing diverse aspects of insured individuals. These variables contain demographic details such as gender, activity and age, as well as pertinent healthcare-related information, including medical procedure category, predisposition to long-term illness, average cost of medical procedures and the anticipated reimbursement amount for the insured.

To refine the dataset quality and optimize subsequent analysis, several preprocessing steps were precisely implemented. The entire process, from data acquisition to preprocessing and the actual implementation of the ANN model, was conducted using the Python programming language. Recognized for its versatility and the availability of comprehensive machine learning libraries, Python facilitated seamless data manipulation, model development, and precise performance assessment. Additionally, the CART decision tree was generated using RapidMiner, an integrated machine-learning platform. It encompasses components for hyperparameter tuning and automatic feature engineering, streamlining the model-building process and enhancing efficiency.

V. Results and Discussion

A. Analysis of Reserving Prediction via ANN Algorithm:

Initially, the entire dataset is partitioned into two distinct segments. The first segment constitutes the Initial Training Set, comprising 60% of the complete dataset. This subset is dedicated to the construction and preliminary training phases of the ANN model. Concurrently, the remaining 40% of the dataset forms the Temporary Set, which will be subsequently utilized for model validation and testing purposes.This partitioning into three distinct sets is essential to ensure the robustness and generalization of the model.

Once the data preprocessing is completed, the next step involves the construction and compilation of the ANN model. During this phase, the architecture of the neural network is defined, including the number of layers, the number of neurons in each layer, and the activation functions used.

A sequential model is instantiated using the Sequential module from Keras. Seven Dense layers are added to the model employing the ReLU activation function. The first layer specifies the input shape of the model based on the number of features in the training data.The 'adam' optimizer is employed for training the model. The loss function is defined as 'mean_squared_error' to quantify the discrepancy between predicted values and true values. Then, two metrics, 'mean_absolute_error' and 'RootMeanSquaredError', are specified to monitor the model's performance during training.

ISSN: 2961-6611

Our study delves into using ANN to enhance the precision of health insurance reserving prediction within the Moroccan context. The total sum of predictions amounts to 21,037,232 MAD. It represents the value of all predictions made by our model, indicating the overall financial forecast derived from our analysis.

Following the construction and compilation of the ANN model, we can now assess its performance on the unseen data from the Test and Validation Sets. Several metrics are employed to analyze the model's performance:

- Mean Squared Error (MSE): This metric, reported as 682550.75, quantifies the average squared difference between the predicted values and the actual values. A lower MSE signifies a better fit between the predictions and the ground truth.
- Mean Absolute Error (MAE): This metric, reported as 153.55, represents the average absolute difference between the predicted values and the actual values. It provides a simpler interpretation of the average prediction error.
- Coefficient of Determination (R^2): This metric, reported as 86%, indicates the proportion of variance in the target variable explained by the model.

Overall, based on the presented evaluation metrics, it can be concluded that the ANN model built for individual reserving demonstrates satisfactory performance and generalizes effectively to new data.

B. Profiling Policyholders Based on their Risk Level via Decision Trees

Using the results obtained from the provision estimation generated by the ANN model, we proceeded to establish risk profiles for policyholders across three distinct classes: R1, R2, and R3, ordered from least to most risky.

Decision trees were employed as a robust method to delineate these risk classes, leveraging the insights gleaned from the ANN predictions. By partitioning policyholders based on key risk indicators derived from the ANN estimations, decision trees offer a transparent and interpretable framework for profiling policyholders according to their relative risk levels.This profiling process enables insurance companies to stratify policyholders into risk categories, thereby facilitating targeted risk management strategies and personalized insurance offerings.

ISSN: 2961-6611

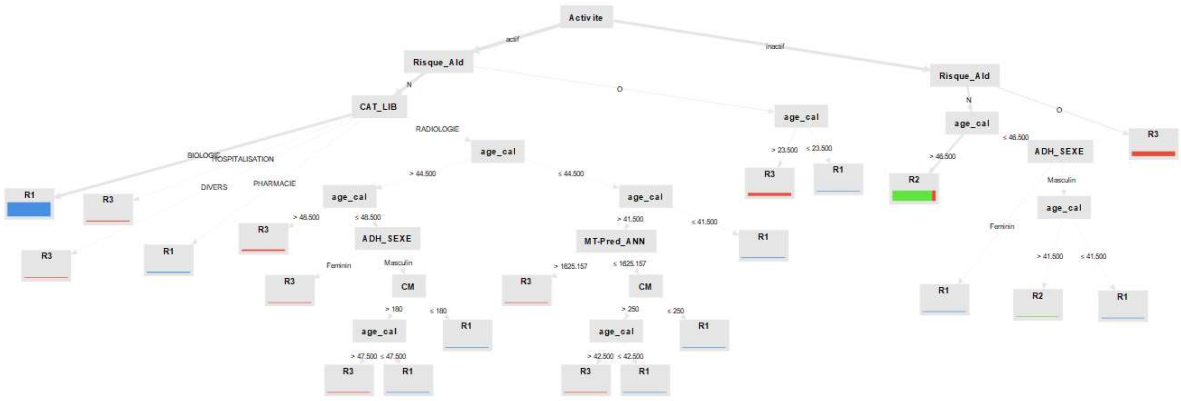

Fig. 1 Decision tree of policyholders' classification Source: Designed by the authors

Based on the profiling conducted using the CART decision tree algorithm, it becomes evident that the most salient selection variables include the age of the insured individual, wherein risk escalates with advancing age. Subsequently, the presence of a Long-Term Disease (Risque-ALD) emerges as the second variable, signifying that the existence of a chronic ailment amplifies the risk level of the insured individual. Following this, the gender variable is identified, succeeded by the activity status of the insured individual, encompassing all qualitative variables instrumental in determining the individual's risk level. Finally, the average cost, predicted amount, and category of healthcare procedure are integrated as significant determinants.

VI. CONCLUSION

In this study, we explored the use of artificial neural networks to improve the prediction of health insurance reserves in Morocco. Starting with a thorough analysis of relevant variables and a rigorous methodology ranging from data collection to the construction and evaluation of the ANN model, we demonstrated the ability of this approach to provide accurate estimates of insurance reserves. In addition, using decision tree-based risk class profiling techniques, we have identified the main factors influencing policyholders' risk levels, including age, presence of chronic illness and gender.Our results underline the importance of an integrated approach, combining advanced modeling techniques such as ANNs with risk profiling methods for effective health insurance reserving management.

REFERENCES

[1] LOPEZ, Olivier, MILHAUD, Xavier, et THÉROND, Pierre-E. A tree-based algorithm adapted to microlevel reserving and long development claims. ASTIN Bulletin: The Journal of the IAA, 2019, vol. 49, no 3, p. 741-762.

[2] GOUNDAR, Sam, PRAKASH, Suneet, SADAL, Pranil, et al. Health insurance claim prediction using artificial neural networks. International Journal of System Dynamics Applications (IJSDA), 2020, vol. 9, no 3, p. 40-57.

[3] E. Astesan, "Les réserves techniques des sociétés d'assurances contre les accidents d'automobiles," Librairie générale de droit et de jurisprudence, 1938.

ISSN: 2961-6611

[4] A. Saoudi, F. El Kassimi and J. Zahi, "Technical reserving in non-life insurance: A literature review of aggregated and individual methods, " Journal of Integrated Studies In Economics, Law, Technical Sciences & Communication, Vol (1), No (2) 2023.

[5] P. Mulquiney, "Artificial Neural Networks in Insurance Loss Reserving," in Joint Conference on Information Sciences, 2006.

[6] GROSSI, Enzo et BUSCEMA, Massimo. Introduction to artificial neural networks. European journal of gastroenterology & hepatology, 2007, vol. 19, no 12, p. 1046-1054.

[7] HAYKIN, Simon. Neural networks: a comprehensive foundation. Prentice Hall PTR, 1998.

[8] GÉRON, Aurélien. Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow. " O'Reilly Media, Inc.", 2022.

ISSN: 2961-6611