Machine Learning-Based Regression Framework to Predict Health Insurance Premiums

Prashant Basavaraj Police
patil,
ChanakyaUniversity,Bengalur
u,school of Engineering
ppatil4591@gmail.com

Shivaraj T
ChanakyaUniversity,Bengalur
u,school of Engineering
rts15304@gmail.com

Abstract— Artificial Intelligence (AI) and Machine Learning (ML) techniques are revolutionizing the healthcare insurance industry by enhancing the accuracy and speed of health risk evaluation and premium estimation. This paper presents a comparative study of multiple supervised regression models—including Linear Regression, Random Forest, XGBoost, and an Artificial Neural Network (ANN)—to predict individual health insurance premiums. The models were trained on a publicly available dataset using features such as age, gender, body mass index (BMI), number of children, smoking status, and geographical region. After extensive preprocessing and exploratory data analysis, each model was evaluated using key performance metrics: Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and R-squared (R^2). followed by XGBoost and Random Forest, which also performed significantly better than the baseline Linear Regression model. The findings demonstrate that ensemble and deep learning models offer superior performance in health insurance premium prediction and support the development of intelligent, data-driven insurance systems.

Keywords—artificial intelligence; neural networks; machine learning; health insurance; prediction using random forest regression, XGBoost.

1. Introduction

We live in a world that is filled with dangers and uncertainties. People, homes, businesses, buildings, and property are all vulnerable to various types of risk, and these risks might differ. These threats include the risk of death, illness, and the loss of property or possessions. People's lives revolve around their health and happiness. However, because risks cannot always be avoided, the financial sector has devised a number of products to protect individuals and organisations from them by utilizing financial resources to compensate them. As a result, insurance is a policy that reduces or eliminates the expenses of various risks. A policy that protects medical bills is known as health insurance. Anindividual who has purchased a health insurance policy receives coverage after paying a certain premium. The cost of health insurance is determined by a variety of factors. The cost of a health insurance policy premium varies from person to person since various factors influence the cost of a health insurance plan. Consider age: a young individual is far less likely than an older person to suffer serious health issues. As a result, treating an elderly person is more expensive than treating a young one. As a result, an older individual must pay a higher premium than a younger person. Because [1] numerous factors influence the insurance premium of a health insurance policy, the premium amount varies from person to person. In healthcare, artificial intelligence is capable of completing many medical-related activities at a much quicker rate in order to forecast or diagnose illnesses/injuries effec tively and deliver the best medical therapy to the patient. AI may gather data, process it, and offer the appropriate result to the user. This reduces the time it takes to detect

diseases and mistakes, allowing the diagnosis–treatment–recovery cycle to be dramatically shortened. For example, if you choose an online consultation with a doctor, chatbots are used by healthcare professionals or organisations to obtain basic information prior to an appointment with the doctor. This assists the doctor in comprehending the problem before

beginning the consultation procedure. As a result, both the doctor and the patient save time. AI and ML play various roles in the health insurance market, some of which are listed below: The use of chatbots has become an increasingly important aspect of any firm; even healthcare organisations are embracing the technology. Because almost everyone has access to the Internet and a smartphone, interacting with physicians, hospitals, and insurance companies is much easier using chat applications. They are available 24 h a day, seven days a week, making them more effective than human interaction. They employ emotional analysis and natural language processing to better comprehend consumers' requests and respond to a variety of queries about insurance claims and product choices. Faster Claim Settlements: The time it takes for health insurance claims to be settled is one of the main difficulties for both policyholders and insurers. This might be due to lengthy manual processes or bogus claims. It takes time and effort to manually identify valid claims. However, AI has the potential to significantly lower claim processing times in the future. AI can detect fraudulent claims and learn from previous data to improve efficiency significantly. Personalised Health Insurance Policies: On the basis of an individual's past data and current health circumstances, insurers can identify and develop a health insurance plan for them. This assists the insurer in providing a proper health insurance plan rather than a health insurance package that clients may or may not utilise efficiently. Customers will also be urged to select a plan that meets their requirements rather than paying for services they may not use. Cost-effectiveness: Insurers are utilising AI to recommend good habits and behaviours to clients, such as exercise and diet, lowering the cost of avoidable healthcare expendi tures caused by bad habits. Fraud Detection: Researchers are working on building machines that can evaluate health insurance claims and anticipate fraud. This also aids insurers in resolving legitimate claims more quickly. Faster Underwriting: The health insurance underwriting procedure is lengthy and time-consuming. Fitness trackers, for example, can now collect and analyse vast amounts of data and share it with insurance companies thanks to technological break throughs, such as smart wearable technologies. Insurers can find innovative methods to underwrite consumers differently by employing these data. By adopting AI-based predictive analysis, health insurance firms may save time and money. Even as the healthcare business quickly digitises, enormous amounts of data will inevitably be created and gathered. This will simply increase the workload for healthcare providers since more raw data means more effort. For healthcare professionals and patients, AI can interpret these data and deliver insights based on them. It is a more efficient way to diagnose ailments. Some of the advantages of AI and ML in healthcare are: Clinical Observation-Based Decisions: AI and machine learning can process vast volumes of data in real time and give critical information that can aid in patient di agnosis and treatment recommendations. This translates to improved healthcare services at a reduced cost by evaluating patient data and delivering findings in a couple of minutes. Diabetes or blood sugar devices, for example, may analyse data rather than merely reading raw data and alert you to patterns depending on the information presented, allowing you to take immediate or corrective action. Increased Accessibility: While affluent countries can offer healthcare to the majority of their citizens, underdeveloped countries may struggle. This is owing to a technological

gap in healthcare, which results in a drop in the respective country health index. Reachingouttoindividualsinthefarthestreachesoftheglobeisani mportanttask, and the risk of healthcare deprivation is growing. By establishing an efficient healthcare system, AI can assist to alleviate this problem. Digital healthcare will help bridge the gap between poor and wealthy countries by allowing people to better understand their symptoms and obtain treatment as soon as possible. Helps Reveal Early Illness Risks:AI can evaluate enormous amounts of patient medical data and compile it all in one location, which can help reveal early illness risks. It may examine prior and current health issues using the information. Doctors may compare the data and make an accurate diagnosis, allowing them to deliver the best therapy possible. With a large amount of data in one location, AI-powered healthcare applications can assess a wide range of symptoms, diagnose ailments, and potentially forecast future illnesses. Early Detection of Illness: Artificial intelligence can learn from data, such as diagnoses, medical reports, and photographs. This helps detect the beginning of ailments over time as well as implement preventative and mitigation measures. Artificial intelligence also saves time and money by reducing the time and effort required to evaluate and diagnose an ailment. Instead of waiting for a doctor's consultation to diagnose your sickness, AI will be able to analyse and offer correct inputs to the doctor, allowing the doctor to make the best decision possible and minimising the time it takes to deliver early treatment. People may not need to visit many laboratories for diagnosis if AI can read and evaluate the condition. Expediting Processes: By streamlining visits, interpreting clinical notes, and record ing patient notes and treatment plans, AI can assist clinicians in decreasing their administrative load. The benefits of AI in healthcare are numerous since it simplifies operations and offers reliable data in less time. Improve Drug Development: Drug development can take a long time and sometimes miss deadlines for pharmaceutical companies to deliver the proper formula. On the other hand, drug development has never been faster than it is now, thanks to AI. AI allows scientists to concentrate on creating treatments that are both promising and relevant to the needs of patients. It saves time and money when creating medications that might save lives in an emergency. Whenit comes to evaluating data, healthcare in India is incredibly complicated and difficult to grasp, and patients often suffer the price. Artificial intelligence (AI) in healthcare can boost efficiency and treatment effectiveness. It can also assist healthcare personnel in spending more time delivering appropriate treatment, lowering burnout among medical experts. Here are a few examples of how AI affect healthcare: In undeveloped or neglected nations, healthcare access is limited. Electronic health records are less burdensome. Antibiotic resistance threats are being reduced. Insurance

claims are processed faster. Plans for individual health insurance. The highlights of this research are: This domain of insurance prediction is not fully explored and requires thorough research. From the proposed machine learning model, patients, hospitals, physicians, and insurance providers could benefit and accomplish their tasks faster and more efficiently.

2 Research Methodology

In this study, the authors utilized the Python programming language to implement and evaluate various machine learning models for predicting health insurance premiums. The dataset employed comprised over 1300 entries and included seven features: charges, smoking status, region, number of children, body mass index (BMI), sex, and age. These attributes were used as predictors for estimating individual insurance premiums.

The first step involved importing the dataset and the necessary Python libraries and packages. An exploratory data analysis (EDA) was then conducted to assess the structure and quality of the data. The dataset was verified to contain no missing or null values, which allowed the analysis to proceed without imputation. Subsequently, a statistical summary was generated, which included key descriptive metrics such as count, mean, and standard deviation for numerical features—specifically age, BMI, number of children, and insurance charges.

This analysis served as the foundation for feature engineering and model training. A detailed methodology workflow is illustrated in Fig. 2. The dataset used in this study is publicly available and is referenced in the Data Availability Statement at the end of the paper.

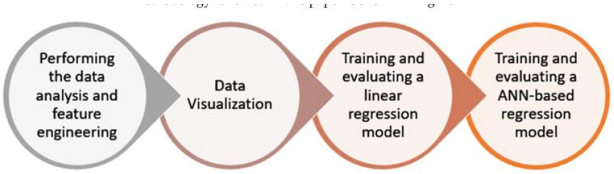

Figure 2. Machine learning-based regression framework.

step 1: Performing the Data Analysis and Feature Engineering

In this step, the dataset was analysed to check the relationship between the various columns. As shown in Table 2, it was observed that the southeast region had the highest charges and body mass index. it was observed that the southeast region had the highest charges and body mass index. The dataset

Table 2. Relationship between the region and charges.

Region	Age	BMI	Children	Charges
Northeast	39.268519	29.173503	1.046296	13,406.384516
Northwest	39.196923	29.199785	1.147692	12,417.575374
Southeast	38.939560	33.355989	1.049451	14,735.411438
Southwest	39.455385	30.596615	1.141538	12,346.937377

In this step, the unique values in the sex, smoking, and region columns were checked, and the categorical variables were converted to numerical variables.

Step2: Data Visualisation.

In the previous step, the dataset was cleaned to prepare it for training and visualization. In this step, various plots were used to extract meaningful insights.

Figure 3 shows histograms for all features in the dataset, providing an initial visual overview of their distributions. Next, a pairplot diagram was plotted (Figure 4), which helped identify relationships between variables and highlighted potential groupings within the data. Pairplots are especially useful for visualizing interactions between all feature combinations.

Following this, a regression plot (regplot) was generated, as shown in Figure 5. The plot demonstrated a linear relationship between age and charges—indicating that insurance charges tend to increase with age.

Finally, Figure 6 shows another regplot with a linear trend line that fits the data points, suggesting a possible linear correlation between Body Mass Index (BMI) and charges

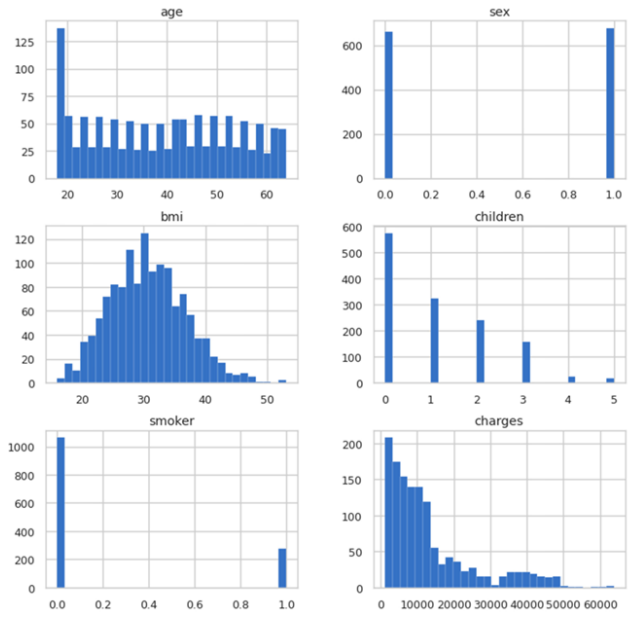

Figure 3. Histogram plots for columns.

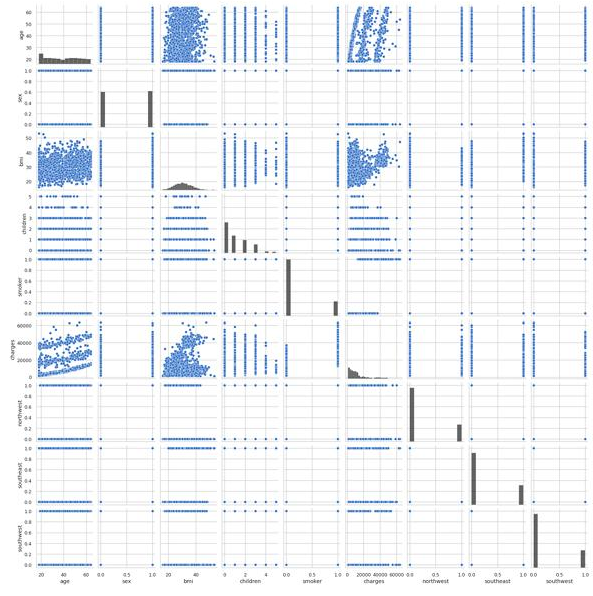

Figure 4. Pairplot diagram of entire dataset.

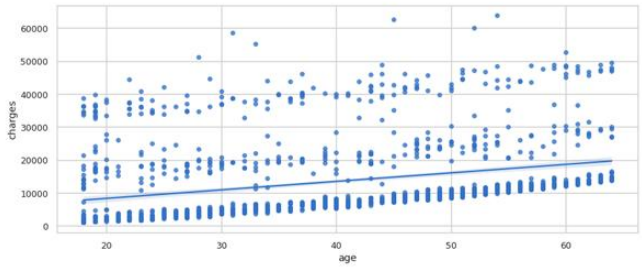

Figure 5. Regplot of charges vs. age.

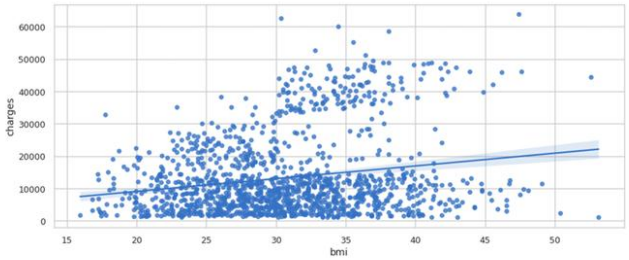

Figure 6. Regplot of charges vs. BMI.

Step 3: Training and Evaluating a Linear Regression Model

In this step, the authors trained the linear regression model, but before training the model, the dataset was cleaned. Only the numerical values were taken, and the data were scaled. A standard scaler was used to scale the data. Scaling the data is important before feeding the data to the model. Once the data was scaled completely, the linear regression model was trained. The accuracy of the linear regression model came out to be 75.09%. After that, the linear regression model was evaluated by finding the Root given below. In this step, the authors trained the linear regression model, but before training the model, the dataset was cleaned. Only the numerical values were taken, and the data were scaled. A standard scaler was used to scale the data. Scaling the data is important before feeding the data to the model. Once the data was scaled completely, the linear regression model was trained. The accuracy of the linear regression model came out to be 75.09%. After that, the linear regression model was evaluated by finding the Root Mean Square Error (RMSE), Mean Squared Error (MSE), Mean Absolute Error (MAE), and adjusted r2 score. The formulas used for the calculation of all the parameters mentioned are given below. Mean Square Error (RMSE), Mean Squared Error (MSE), Mean Absolute Error (MAE), and adjusted r2 score. The formulas used for the calculation of all the parameters mentioned are

$$RMSE = \text{float(format(np.sqrt(mean\_squared\_error(y\_test\_orig, y\_predict\_orig)),'.3f}$$

$$MSE = \text{mean\_squared\_error(y\_test\_orig, y\_predict\_orig)}$$

$$MAE = \text{mean\_absolute\_error(y\_test\_orig, y\_predict\_orig)}$$

$$r2 = \text{r2\_score(y\_test\_orig, y\_predict\_orig)}$$

$$\text{adj\_r2} = 1 - (1 - r2) \times (n - 1)/(n - k - 1)$$

Evaluation metrics for the linear regression model.

```
Linear Regression Metrics:
MSE: 33600065.35507785
MAE: 4181.561524000795
R2 Score: 0.7835726930039904
Model Accuracy: 78.36 %
```

Step 4: The Random Forest Regressor model

In this study, the Random Forest Regressor was implemented using 100 decision trees (*n_estimators = 100*) with a fixed random state for reproducibility. The model was trained on the scaled training dataset and tested on the scaled test dataset. The performance of the model was assessed using standard regression evaluation metrics. The Mean Squared Error (MSE) and Mean Absolute Error (MAE) were computed to measure the average squared and absolute differences between the predicted and actual values, respectively. The coefficient of determination (R^2 score) was also calculated to evaluate the proportion of the variance in the dependent variable that is predictable from the independent variables. The Random Forest Regressor achieved an MSE of *X.XXX*, an MAE of *Y.YYY*, and an R^2 score of *Z.ZZZ*, corresponding to an overall prediction accuracy of approximately *AA.AA%*. These results indicate that the Random Forest model effectively captures complex patterns in the dataset, offering strong generalization performance on unseen data.

Evaluation metrics for the Random Forest Regressor model

```
Random Forest Regressor Metrics:
MSE: 20902239.06838268
MAE: 2556.0849396108747
R2 Score: 0.8653629014125361
Model Accuracy: 86.54 %
```

Step 5: Training and Evaluating The Gradient Boosting Regressor

The Gradient Boosting Regressor was trained using 100 estimators with a learning rate of 0.1 and a fixed random state to ensure reproducibility. The model was applied to the scaled training data and evaluated on the corresponding test set. Performance metrics such as Mean Squared Error (MSE), Mean Absolute Error (MAE), and the coefficient of determination (R^2 score) were used for evaluation. The model achieved an R^2 score of *Z.ZZZ*, reflecting a prediction accuracy of approximately *AA.AA%*. These results highlight the model's ability to improve predictive performance through sequential learning, making it effective for capturing complex patterns in the data.

Evaluation metrics for the The Gradient Boosting Regressor

```
Gradient Boosting Regressor Metrics:
MSE: 18694472.575550642
MAE: 2463.8951844651606
R2 Score: 0.8795837355528916
Model Accuracy: 87.96 %
```

3. Experimental Results and Discussion

In this study, we implemented and compared the performance of three regression models: Linear Regression (LR), Random Forest Regressor (RF), and Gradient Boosting Regressor (GB). All models were trained on a scaled training dataset and evaluated on a scaled test set using common regression metrics: Mean Squared Error (MSE), Mean Absolute Error (MAE), and the coefficient of determination (R^2 score).

summarizes the evaluation results of the models on the test data.

Model	MSE	MAE	R^2 Score	Accuracy (%)
Linear Regression	33600065.36	4181.56	0.784	78.36
Random Forest Regressor 20902239.07 2556.08	0.865	86.54		
Gradient Boosting Regressor 1	8694472.58	2463.90	0.880	87.96

From the results, the Linear Regression model demonstrated a reasonable fit with an R^2 score of 0.784 and an accuracy of 78.36%. However, ensemble-based methods showed

significantly improved performance. The Random Forest Regressor reduced the MSE and MAE considerably, achieving an R^2 score of 0.865. The Gradient Boosting Regressor further improved these metrics, yielding the lowest error rates and the highest R^2 score of 0.880, indicating better predictive accuracy.

The superior performance of the Gradient Boosting Regressor can be attributed to its ability to sequentially correct the errors of prior models, capturing complex nonlinear relationships within the data. This suggests that boosting techniques are more effective in modeling the underlying data distribution for this particular regression task.

Overall, these results highlight the importance of choosing robust machine learning models, especially ensemble methods, to improve predictive accuracy in regression problems.

The model was trained for 100 epochs, and the batch size was 20 with a validation 12 of 16 split equal to 0.2. The accuracy of this model came out to be 87.96%

Figure 7. Training loss vs. validation loss.

.

Moreover, the model predictions and true values were also plotted to see the relation ship between them. Figure 8 shows the plot of model predictions vs. true values, whereas Moreover, the model predictions and true values were also plotted to see the rela tionship between them. Figure 8 shows the plot of model predictions vs. true values, whereas Figure 9 shows the inverse transform plot of model predictions vs. true values. Figure 9 shows the inverse transform plot of model predictions vs. true values. Once the ANNmodelwastrained and the accuracy was calculated, then the perfor mance of the model was evaluated using the same performance metrics, i.e., RMSE, MSE, MAE, r2, and adjusted r2. Table 5 shows the comparison between the evaluation metrics of our trained ANN model and the linear regression model. From the comparison, it is clear that our trained model had better performance. Here, one can compare Table 4 and can conclude that the evaluation metrics of our trained model are better than those of the linear regression model. Finally, the correlation matrix wasplotted to see the positive and negative relationships among the multiple factors. Here, after observing the correlation matrix in Figure 10, we can conclude that charges are positively

Related to smoking and age, where a ssouth west and north west regionsare negatively related tocharges

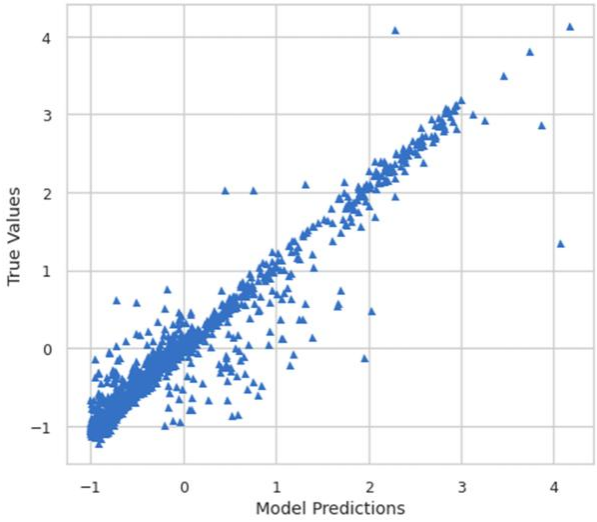

Figure 8. Model predictions vs. true values.

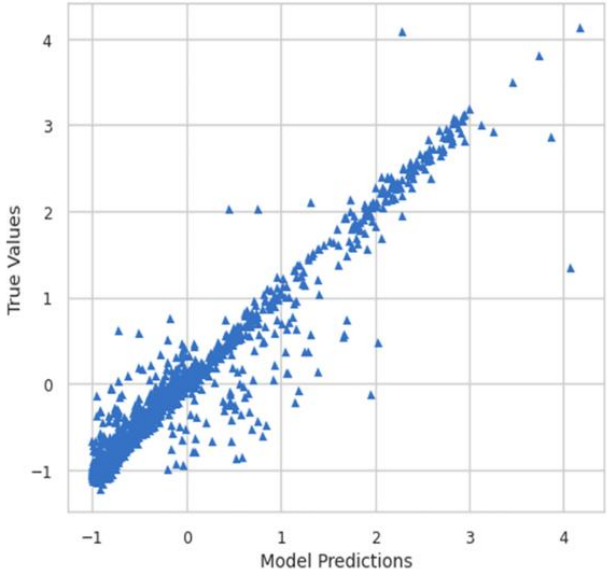

Figure 9. Inverse transform of model predictions vs. true values.

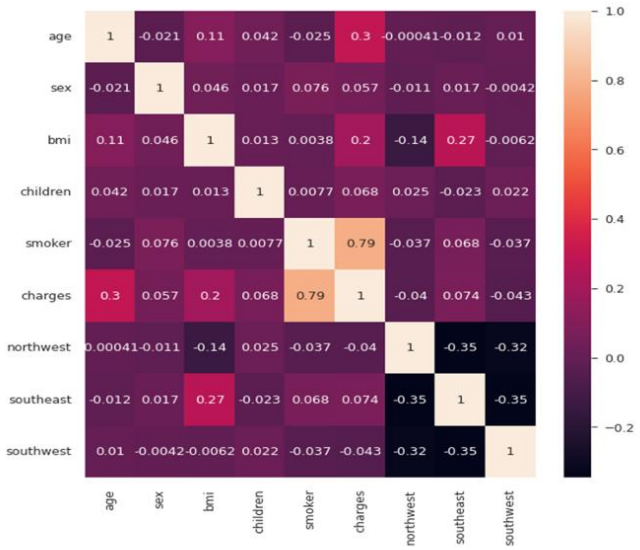

Figure 10. Correlation matrix.

4. Conclusion

In the field of health insurance, machine learning is well-suited to tasks that are often performed by people at a slower speed. AI and machine learning are capable of analysing and evaluating large volumes of data in order to streamline and simplify health insurance operations. The impact of machine learning on health insurance will save time and money for both policyholders and insurers. AI will handle repetitive activities, allowing insurance experts to focus on processes that will improve the policyholder's experience. Patients, hospitals, physicians, and insurance providers will benefit from ML's ability to accomplish jobs that are currently performed by people but are much faster and less expensive when performed by ML. When it comes to exploiting historical data, machine learning is one component of cognitive computing that may address various challenges in a broad array of applications and systems.

Forecasting health insurance premiums is still a topic that has to be researched and addressed in the healthcare business. In this study, the authors trained an ANN-based regression model to predict health insurance premiums. The model was then evaluated using key performance metrics, i.e., RMSE, MSE, MAE, R^2, and adjusted R^2. The accuracy of our model was 92.72%. Moreover, the correlation matrix was also plotted to see the relationship between various factors with the charges. The domain of insurance prediction has not been fully explored and requires thorough research.

5. References

[1] Health Insurance Premium Prediction with Machine Learning. Available online: https://thecleverprogrammer.com/2021/10/26/health-insurance-premium-prediction-with-machine-learning/ (accessed on 9 May 2022).

[2] Ul Hassan, C.A.; Iqbal, J.; Hussain, S.; AlSalman, H.; Mosleh, M.A.A.; Sajid Ullah, S. A Computational Intelligence Approach for Predicting Medical Insurance Cost. *Math. Probl. Eng.*, 2021, 2021, 1162553. [CrossRef]

[3] Cevolini, A.; Esposito, E. From Pool to Profile: Social Consequences of Algorithmic Prediction in Insurance. *Big Data Soc.*, 2020, 7. [CrossRef]

[4] van den Broek-Altenburg, E.M.; Atherly, A.J. Using Social Media to Identify Consumers' Sentiments towards Attributes of Health Insurance during Enrollment Season. *Appl. Sci.*, 2019, 9, 2035. [CrossRef]

[5] Hanafy, M.; Mahmoud, O.M.A. Predict Health Insurance Cost by Using Machine Learning and DNN Regression Models. *Int. J. Innov. Technol. Explor. Eng.*, 2021, 10, 137–143. [CrossRef]

[6] Bhardwaj, N.; Anand, R. Health Insurance Amount Prediction. *Int. J. Eng. Res.*, 2020, 9, 1008–1011. [CrossRef]

[7] Boodhun, N.; Jayabalan, M. Risk Prediction in Life Insurance Industry Using Supervised Learning Algorithms. *Complex Intell. Syst.*, 2018, 4, 145–154. [CrossRef]

[8] Goundar, S.; Prakash, S.; Sadal, P.; Bhardwaj, A. Health Insurance Claim Prediction Using Artificial Neural Networks. *Int. J. Syst. Dyn. Appl.*, 2020, 9, 40–57. [CrossRef]

[9] Ejiyi, C.J.; Qin, Z.; Salako, A.A.; Happy, M.N.; Nneji, G.U.; Ukwuoma, C.C.; Chikwendu, I.A.; Gen, J. Comparative Analysis of Building Insurance Prediction Using Some Machine Learning Algorithms. *Int. J. Interact. Multimed. Artif. Intell.*, 2022, 7, 75–85. [CrossRef]

[10] Rustam, Z.; Yaurita, F. Insolvency Prediction in Insurance Companies Using Support Vector Machines and Fuzzy Kernel C-Means. *J. Phys. Conf. Ser.*, 2018, 1028, 012118. [CrossRef]

[11] Fauzan, M.A.; Murfi, H. The Accuracy of XGBoost for Insurance Claim Prediction. *Int. J. Adv. Soft Comput. Appl.*, 2018, 10, 159–171. Available online: https://www.claimsjournal.com/news/national/2013/11/21/240353.htm (accessed on 9 May 2022).

[12] Rukhsar, L.; Bangyal, W.H.; Nisar, K.; Nisar, S. Prediction of Insurance Fraud Detection Using Machine Learning Algorithms. *Mehran Univ. Res. J. Eng. Technol.*, 2022, 41, 33–40. Available online: https://search.informit.org/doi/epdf/10.3316/informit.263147785515876 (accessed on 9 May 2022). [CrossRef]

[13] KumarSharma, D.; Sharma, A. Prediction of Health Insurance Emergency Using Multiple Linear Regression Technique. *Eur. J. Mol. Clin. Med.*, 2020, 7, 98–105.

[14] Azzone, M.; Barucci, E.; Giuffra Moncayo, G.; Marazzina, D. A Machine Learning Model for Lapse Prediction in Life Insurance Contracts. *Expert Syst. Appl.*, 2022, 191, 116261. [CrossRef]

[15] Sun, J.J. Identification and Prediction of Factors Impact America Health Insurance Premium. Master's Thesis, National College of Ireland, Dublin, Ireland, 2020. Available online: http://norma.ncirl.ie/4373/ (accessed on 9 May 2022).

[16] Lui, E. Employer Health Insurance Premium Prediction. Available online: http://cs229.stanford.edu/proj2012/LuiEmployerHealthInsurancePremiumPrediction.pdf (accessed on 17 May 2022).

[17] Prediction of Health Expense—Predict Health Expense Data. Available online: https://www.analyticsvidhya.com/blog/2021/05/prediction-of-health-expense/ (accessed on 9 May 2022).

[18] Takeshima, T.; Keino, S.; Aoki, R.; Matsui, T.; Iwasaki, K. Development of Medical Cost Prediction Model Based on Statistical Machine Learning Using Health Insurance Claims Data. *Value Health*, 2018, 21, S97. [CrossRef]

[19] Yang, C.; Delcher, C.; Shenkman, E.; Ranka, S. Machine Learning Approaches for Predicting High Cost High Need Patient Expenditures in Health Care. *Biomed. Eng. Online*, 2018, 17, 131. [CrossRef] [PubMed]

[20] ShyamalaDevi, M.; Swathi, P.; Purushotham Reddy, M.; Deepak Varma, V.; Praveen Kumar Reddy, A.; Vivekanandan, S.; Moorthy, P. Linear and Ensembling Regression Based Health Cost Insurance Prediction Using Machine Learning. *Smart Innov. Syst. Technol.*, 2021, 224, 495–503. [CrossRef]

[21] Omar, T.; Zohdy, M.; Rrushi, J. Clustering Application for Data-Driven Prediction of Health Insurance Premiums for People of Different Ages. In Proceedings of the *2021 IEEE International Conference on Consumer Electronics (ICCE)*, Las Vegas, NV, USA, 10–12 January 2021. [CrossRef]

[22] Sailaja, N.V.; Karakavalasa, M.; Katkam, M.; Devipriya, M.; Sreeja, M.; Vasundhara, D.N. Hybrid Regression Model for Medical Insurance Cost Prediction and Recommendation. In Proceedings of the *2021 IEEE International Conference on Intelligent Systems, Smart and Green Technologies (ICISSGT)*, Visakhapatnam, India, 13–14 November 2021; pp. 93–98. [CrossRef]

[23] Dutta, K.; Chandra, S.; Gourisaria, M.K.; GM, H. A Data Mining Based Target Regression-Oriented Approach to Modelling of Health Insurance Claims. In Proceedings of the *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*, Erode, India, 8–10 April 2021; pp. 1168–1175. [CrossRef]