

Health Insurance Cost Prediction and Analysis

Jeyakumar D¹, Bharath Teja S², Lokesh Reddy Y³, Ruthik P⁴

Head of the Department, Computer Science and Engineering¹

Students, Computer Science and Engineering^{2,3,4}

Dhanalakshmi College of Engineering, Chennai, India

Abstract: *The Indian authorities spends 1.5 percent annually GDP on public health, which is significantly lower than GDP other countries. On the other hand, global public health spending hand, has almost doubled in line with inflation in the last two decade, reaching \$8.5 trillion in 2019, or 9.8%globallyGDP. Transnational multi-private sectors provide around 60%. comprehensive medical treatment and 70% outpatient care, Which price sufferers astronomically excessive fees. Insurance data has increased dramatically in recent times decades and carriers now have access to it. Health insurance system is exploring predictive modelling to support its business operations and services. Computer algorithms and the machine Learning (ML) is used to study and analyse past insurance data and predict new output values based on customer trends behaviour, insurance contracts and data-driven business decisions, and support in formulating new schemes. In addition to it Machine Learning found huge and potential application in insurance industry. It develops real-time insurance costs a price prediction system called ML Health Insurance Prediction System using ML algorithms to help insurance companies on the market easily and quickly determination of premium values and thus limit health expenses. The proposed model includes a demonstration of Random Forest Regression to predict insurance costs and assess model results. In proposed model, achieved a forest regression model better results with R-squared value of 0.80 compared to all other models.*

Keywords: Insurance

REFERENCES

- [1]. "National Health Accounts," National Health Systems Resource Centre. [Online]. Available:<https://nhsrindia.org/national-health-accounts> records care: an introductory review," International Journal for Quality in Health Care, vol. 23, no. 3, pp. 331–341, 2011.
- [2]. Bertsimas, M. V. Bjarnad'ottir, M. A. Kane, J. C. Kryder, R. Pandey, S. Vempala, and G. Wang, "Algorithmic prediction of health-care costs," Operations Research, vol. 56, no. 6, pp. 1382–1392, 2008.
- [3]. Stucki, O. "Predicting the customer churn with machine learning methods: case: private insurance customer data" Master's dissertation, LUT University, Lappeenranta, Finland, 2019.
- [4]. Sterne, J. A., White, I. R., Carlin, J. B., Spratt, M., Royston, P., Kenward, Carpenter, J. R. (2009). Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. Bmj, 338L.
- [5]. H. Demirtas, "Flexible Imputation of Missing Data", J. Stat. Soft., vol. 85, no. 4, pp. 1–5, Jul. 2018. Available: DOI: 10.18637/jss.v085.b04 .
- [6]. H. Goldstein, W. Browne and J. Rasbash, "Multilevel modelling of medical data," Statistics in Medicine, John Wiley and Sons, vol. 21, no. 21, pp. 3291–3315, 2002.
- [7]. T. Han, A. Siddique, K. Khayat, J. Huang and A. Kumar, "An ensemble machine learning approach for prediction and optimization of modulus of elasticity of recycled aggregate concrete," Construction and Building Materials, vol. 244, pp. 118–271, 2020.
- [8]. X. Zhu, C. Ying, J. Wang, J. Li, X. Lai et al., "Ensemble of ML-kNN for classification algorithm recommendation," Knowledge-Based Systems, vol. 106, pp. 933, 2021.
- [9]. G. Reddy, S. Bhattacharya, S. Ramakrishnan, C. L. Chowdhary, S. Hakak et al., "An ensemble-based machine learning model for diabetic retinopathy classification," in 2020 Int. Conf. on Emerging Trends in Information Technology and Engineering, IC-ETITE, VIT Vellore, IEEE, pp. 1–6, 2020.

- [10]. Douglas C Montgomery, Elizabeth A Peck and G Geoffrey Vining, "Introduction to linear regression analysis", John Wiley & Sons, vol. 821, 2012.
- [11]. Tian Jinyu, Zhao Xin et al., "Apply multiple linear regression model to predict the audit opinion," in 2009 ISECS International Colloquium on Computing, Communication, Control, and Management, IEEE, pp.1-6, 2009.
- [12]. Ostertagova et al., "Modelling using Polynomial Regression", "Procedia Engineering", vol. 48, pp. 500-506, 2012.
- [13]. Donald W. Marquardt, Ronald D. Snee et al., "Ridge Regression in Practice", "The American Statistician", vol. 29, pp – 3-20, 2012.
- [14]. V. Roth, "The generalised LASSO", "IEEE Transactions on Neural Networks", vol. 15, pp – 16 28, 2004.