A Machine and Deep Learning Framework for Robust Health Insurance Fraud Detection and Prevention

Suhag Pandya
Independent Researcher
spandya5886@ucumberlands.edu

Abstract*: Healthcare fraud is the deliberate submission of false information or the fabrication of facts in order to get entitlement payments. As a result, it wastes healthcare funds and raises healthcare expenses. For both insurance firms and consumers, predicting health insurance prices is an essential undertaking. The purpose of this research is to examine the feasibility of using ML models for accurate identification of medical insurance fraud. Using a dataset with more than 1300 entries and important characteristics such charges, smoking status, geography, BMI, age, sex, and children, this research investigates the use of ANN for strong health insurance fraud detection. Traditional models like Ridge, Lasso, and XGBoost fared poorly when compared to the ANN, which achieved a R² of 92.72 and a low RMSE of 0.27, according to error measures utilised to assess a model's performance. A findings show that an ANN is good at identifying fraudulent health insurance claims, which bodes well for its future use in better fraud prevention systems. Limitations include the dataset's small size and limited features, suggesting that future studies should expand the dataset and explore more advanced ML techniques for further optimisation*

Keywords: Healthcare, Insurance Fraud, Machine Learning, Fraud Detection, Predictive Modeling, Claims Processing