**ABSTRACT**

In the rapidly evolving landscape of healthcare insurance, accurately predicting medical insurance premiums is crucial for insurers aiming to optimize pricing strategies and for policyholders seeking cost-effective coverage. This study explores the application of machine learning techniques to predict medical insurance premiums based on a comprehensive dataset comprising 25000 entries with 24 diverse features. The dataset encompasses a wide array of variables such as age, gender, and income, as well as lifestyle factors like smoking habits, exercise frequency, and alcohol consumption. Additionally, medical history elements such as pre-existing conditions, average glucose levels, and body mass index are incorporated to enhance the predictive power of the models.

The research methodology involves a meticulous approach to data preprocessing, including handling missing values and feature scaling, followed by exploratory data analysis to uncover correlations and patterns. Various machine learning algorithms are employed to construct predictive models. The performance of these models is rigorously evaluated using metrics such as RMSE, MAE, and R-squared, with hyperparameter tuning and cross-validation techniques applied to achieve optimal results.

The findings of this study highlight the significance of lifestyle and medical factors in determining insurance premiums, offering valuable insights into the risk assessment process. The predictive models demonstrate promising accuracy, showcasing the potential of machine learning in transforming insurance pricing mechanisms. Insurers can leverage these insights to develop more precise and personalized premium calculations, thereby enhancing customer satisfaction and retention. Furthermore, the models provide policyholders with a transparent understanding of how their health and lifestyle choices impact their insurance costs, empowering them to make informed decisions regarding their coverage.