ABSTRACT

This study aims to design a system capable of predicting disease risk levels using the Naive Bayes algorithm and questionnaire data. Respondent data was collected via Google Forms, covering variables such as age, lifestyle, stress levels, sleep quality, physical activity, and family health history. The data underwent preprocessing and was converted into a tabular format before being divided into 80 training data points and 20 testing data points. The Naive Bayes algorithm was used to classify disease risk into low, moderate, and high categories. Test results showed that the model could generate predictions with a high level of accuracy. The trained model was then implemented in a web-based system so that users could easily and quickly determine their disease risk.

Keywords: Naive Bayes, Disease Risk Prediction.