

ABSTRAK

ANALISIS KETERKAITAN POLA MAKAN DAN GAYA HIDUP DALAM MEMPREDIKSI RISIKO PENYAKIT MENGGUNAKAN ALGORITMA APRIORI DAN *RANDOM FOREST*

Anisa Meilia

H1D021003

Pola makan dan gaya hidup merupakan faktor penting yang memengaruhi risiko penyakit kronis seperti jantung, obesitas, hipertensi, dan diabetes. Penelitian ini bertujuan untuk menganalisis keterkaitan antara pola makan dan gaya hidup terhadap risiko penyakit serta membangun sistem prediksi berbasis data. Metode yang digunakan adalah algoritma Apriori untuk menemukan aturan asosiasi dan algoritma *Random Forest* untuk membangun model prediksi. Dataset yang digunakan mencakup data demografi, data pola makan, data gaya hidup, dan risiko penyakit, yang diperoleh dari platform *Kaggle*. Hasil analisis menunjukkan bahwa algoritma Apriori mampu mengidentifikasi aturan keterkaitan yang cukup kuat dengan nilai *confidence* 30% antara variabel pola makan dan gaya hidup terhadap risiko penyakit. Sementara itu, model *Random Forest* menunjukkan performa akurasi tinggi, yaitu 93% untuk risiko hipertensi dan obesitas, 99% untuk risiko jantung, serta 94% untuk risiko diabetes dalam memprediksi masing-masing risiko penyakit. Penelitian ini juga menghasilkan aplikasi berbasis *website* menggunakan *framework* Streamlit yang menampilkan hasil analisis dan prediksi secara interaktif.

Kata kunci: Apriori, Gaya Hidup, Pola Makan, Prediksi, *Random Forest*, Risiko Penyakit.

ABSTRACT

ANALYZING THE RELATIONSHIP BETWEEN DIET AND LIFESTYLE IN PREDICTING DISEASE RISK USING APRIORI AND RANDOM FOREST ALGORITHMS

Anisa Meilia

H1D021003

Diet and lifestyle are important factors that influence the risk of chronic diseases such as heart disease, obesity, hypertension and diabetes. This research aims to analyze the relationship between diet and lifestyle to disease risk and build a data-based prediction system. The methods used are Apriori algorithm to find association rules and Random Forest algorithm to build prediction models. The dataset used includes demographic data, diet data, lifestyle data, and disease risk, obtained from the Kaggle platform. The analysis results show that the Apriori algorithm is able to identify a fairly strong association rule with a confidence value of 30% between diet and lifestyle variables and disease risk. Meanwhile, the Random Forest model showed high accuracy performance, namely 93% for hypertension and obesity risk, 99% for heart risk, and 94% for diabetes risk in predicting each disease risk. This research also produced a web-based application using the Streamlit framework that displays the results of analysis and predictions interactively.

Keywords: *Apriori, Diet, Disease Risk, Lifestyle, Prediction, Random Forest.*