

ABSTRAK

Nama : Muhammad Rezanur Ardianto
Nim : 41520110043
Program Studi : Teknik Informatika
Judul Proposal Penelitian : Komparasi Algoritma Klasifikasi KNN, Naïve Bayes, dan Decision Tree Untuk Memprediksi Penyakit Diabetes Dengan Menggunakan RapidMiner Studio
Pembimbing : Rushendra, S.Kom, M.T

Penyakit diabetes merupakan ibu dari segala macam penyakit karena dampak dari penyakit diabetes adalah akan menimbulkan komplikasi terhadap penyakit mematikan lainnya seperti serangan jantung dan stroke yang setiap tahunnya mengakibatkan kematian 3,8 juta jiwa. Penyakit diabetes tidak dapat sembuh sepenuhnya namun jika telah terkenda diabetes dapat melakukan jaga pola makan, olahraga yang teratur, dan juga konsumsi obat. Oleh karena itu salah satu cara menghindari penyakit diabetes adalah dengan cara meghindari atau mencegah penyakit diabetes salah satu caranya adalah dengan memiliki informasi apa saja yang menyebabkan diabetes terjadi. Dengan data yang dimiliki dan kemampuan dunia IT yang dapat memprediksi sebuah kejadian diharapkan dapat membantu masyarakat untuk mendapatkan informasi yang berguna untuk mencegah terkenanya penyakit diabetes.

Kumpulan data ini didapat dari situs data dunia Kaggle yang berasal dari Neha Prerna Tigga dan Dr. Shruti Garg dari the Department of Computer Science and Engineering, sebanyak 952 data yang terdiri dari 17 variable dan 1 class target. Penelitian ini menggunakan model algoritma KNN, Naïve Bayes, dan Decision Tree, serta penentuan nilai K-fold cross validation yang bertujuan untuk memprediksi pengukuran diagnostik apakah seorang pasien menderita diabetes. Untuk mencapai tingkat akurasi terbaik, Pada data penyakit diabetes yang ditemukan terdiri dari Age, Gender, Family_Diabetes, highBP, PhysicallyActive, BMI, Smoking, Alcohol, Sleep, SoundSleep, RegularMedicine, JunkFood, Stress, BPLevel, Pregancies, Pdiabetes, UriationFreq, dan Diabetic sebagai Output. Berdasarkan hasil dari penelitian yang telah dilakukan didapatkan kesimpulan bahwa Algoritma KNN sangat cocok untuk memprediksi penyakit Diabetes karena memiliki nilai akurasi sebesar 96,13% % pada nilai K-fold cross validation bernilai 10 dan konstanta KNN bernilai 2.

Kata Kunci : Algoritma knn, Algortima Naïve Bayes, Decision Tree, RapidMiner Studio, Diabetes

ABSTRACT

Name : Muhammad Rezanur Ardianto
Nim : 41520110043
Study Program : Teknik Informatika
Title Research Proposal : Comparison of Classification Algorithms KNN, Naïve Bayes, and Decision Tree for Predicting Diabetes Using RapidMiner Studio

Counsellor : Rushendra, S.Kom, M.T

The disease of diabetes is considered the mother of all diseases because its impact can lead to complications with other deadly diseases such as heart attacks and strokes, resulting in 3.8 million deaths each year. Diabetes cannot be cured completely, but individuals with diabetes can manage it by maintaining a healthy diet, regular exercise, and medication. Therefore, one way to prevent diabetes is to avoid it or take preventive measures by having information about its causes. With the available data and the predictive capabilities of IT, it is hoped that this information can help the community prevent the occurrence of diabetes.

The data used in this study were obtained from the Kaggle world data site, provided by Neha Prerna Tigga and Dr. Shruti Garg from the Department of Computer Science and Engineering. The dataset consists of 952 records with 17 variables and 1 target class. The research employed the KNN, Naïve Bayes, and Decision Tree algorithm models, along with the determination of K-fold cross-validation values to predict diagnostic measurements indicating whether a patient is suffering from diabetes.

To achieve the highest level of accuracy, the data related to diabetes include Age, Gender, Family_Diabetes, highBP, PhysicallyActive, BMI, Smoking, Alcohol, Sleep, SoundSleep, RegularMedicine, JunkFood, Stress, BPLevel, Pregancies, Pdiabetes, UriationFreq, and Diabetic as the output. Based on the research results, it is concluded that the KNN algorithm is highly suitable for predicting diabetes, with an accuracy rate of 96,13%% at a K-fold cross-validation value of 10 and a constant KNN value of 2.

Key Word : Algoritma knn, Algoritma Naïve Bayes, Decision Tree, RapidMiner Studio, Diabetes