

KLASIFIKASI PENYAKIT DIABETES MENGGUNAKAN ALGORITMA *NAÏVE BAYES*

Edwin Arizal Mandalika¹, Moh. Dasuki², Reni Umilasari³

Program Studi Teknik Informatika, Fakultas Teknik

Universitas Muhammadiyah Jember

Email: eamandalika@gmail.com¹, moh.dasuki22@unmuhjember.ac.id²,
reni.umilasari@unmuhjember.ac.id³

ABSTRAK

Diabetes adalah penyakit kronis yang menimbulkan gangguan metabolisme dan dikenali melalui kadar gula darah yang tinggi. Dengan memanfaatkan teknologi informasi dan komunikasi maka saat ini proses diagnosis secara dini penyakit diabetes dapat dilakukan dengan cara mengolah *dataset* rekam medis orang yang memiliki penyakit diabetes dengan parameter gejala umum dan tidak umum yang ditemukan pada fase awal sampai fase diagnosa dan hasil dari pengelolaan data tersebut bisa digunakan guna membuat suatu sistem klasifikasi penyakit diabetes. Tujuan utama dari klasifikasi adalah untuk mengidentifikasi pola atau hubungan di antara data yang ada sehingga dapat dikategorikan ke dalam kelas yang tepat. Algoritma *Naïve Bayes* merupakan salah satu algoritma yang biasa diterapkan dalam klasifikasi untuk menentukan kelas data. Algoritma ini melakukan klasifikasi probabilistik yang digunakan untuk menentukan kelas suatu item berdasarkan beberapa Atribut yang melekat pada item tersebut. Data yang digunakan yaitu dataset penyakit diabetes yang berasal dari Kaggle berjumlah 70.692 data meliputi 35.346 record orang yang terkena diabetes dan 35.346 record orang tidak terkena diabetes. Hasil dari penelitian dengan menggunakan metode *Naïve Bayes* dengan *k-fold cross validation* menunjukkan hasil yang cukup baik dengan akurasi tertinggi sebesar 73%, presisi 73% dan recall 72% pada data uji.

Kata Kunci: Diabetes, *Naïve Bayes*, Klasifikasi

CLASSIFICATION OF DIABETES USING THE NAÏVE BAYES ALGORITHM

Edwin Arizal Mandalika¹, Moh. Dasuki², Reni Umilasari³

Informatic Engineering Study Program, Faculty of Engineering

University of Muhammadiyah Jember

Email: eamandalika@gmail.com¹, moh.dasuki22@unmuhjember.ac.id²,
reni.umilasari@unmuhjember.ac.id³

ABSTRACT

Diabetes is a chronic disease that causes metabolic disorders and is known through high blood sugar levels. By utilizing technology and communication, currently the process of early diagnosis of diabetes can be carried out by processing a dataset of medical records of people who have diabetes with parameters of common and uncommon symptoms found in the initial phase until the diagnosis phase and the results of managing this data can be used. to create a classification system for diabetes. The main goal of classification is to identify patterns or relationships between existing data so that it can be broken down into appropriate classes. The Naïve Bayes algorithm is an algorithm that is usually applied in classification to determine data classes. This algorithm performs a probabilistic classification which is used to determine the class of an item based on several attributes attached to the item. The data used is a diabetes dataset originating from Kaggle totaling 70,692 data including 35,346 records of people who have diabetes and 35,346 records of people who do not have diabetes. The results of research using the Naïve Bayes method with k-fold cross validation show quite good results with the highest accuracy of 73%, precision 73% and recall 72% on test data.

Keywords: *Diabetes, Naïve Bayes, Classification*