

## ABSTRACT

EGI FERNANDES BANGUN (03082180066)

### **PREDICTION OF DIABETES MELLITUS BY USING NAIVE BAYES CLASSIFIER METHOD**

*(xvi + 47 pages : 10 images; 6 table; 4 appendix )*

*The Basic Health Research (Riskesdas) conducted in 2018 collected data on people with Diabetes Mellitus in the population aged 15 years. The criteria for Diabetes Mellitus in Basic Health Research (Riskesdas) 2018 refers to the consensus of the Indonesian Endocrinology Association (PERKENI) which follows the criteria of the American Diabetes Association (ADA). The results of Riskesdas 2018 stated that the prevalence of Diabetes Mellitus in Indonesia based on a doctor's diagnosis at the age of 15 years was 2%. This figure shows an increase compared to the prevalence of Diabetes Mellitus in 2013 which was 1.5%. Data mining can be used because it can make predictions for Diabetes Mellitus in patients by utilizing mathematical or statistical methods so that Diabetes Mellitus results can be predicted correctly and accurately. Naive Bayes Classifier is a classification method rooted in Bayes theorem. In this study, the dataset used was data from women aged over 21 years. The raw data for this study consisted of 8 columns and 768 rows. The data has 8 attributes consisting of Pregancies, Glucose, Blood Pressure, Skin Thickness, Insulin, BMI (Body Mass Index), Diabetes Pedigree Function, Age and Outcome. From the research results, the accuracy value obtained by using the Naïve Bayes method to classify Diabetes Mellitus is 78%. And for the results of the calculation of precision for those affected with Diabetes Mellitus is 81% and those who are not Diabetes Mellitus is 72%. And for recall for those affected with Diabetes Mellitus is 87% and 61% for those who are not Diabetes Mellitus. Meanwhile, the f1-score for those with diabetes mellitus was 84% and 66 for those without diabetes mellitus.*

*Keywords: Diabetes Mellitus, Naïve Bayes Classifier, Data Mining*

## ABSTRAK

EGI FERNANDES BANGUN (03082180066)

### **MEMPREDIKSI PENYAKIT DIABETES MELITUS MENGGUNAKAN METODE *NAIVE BAYES CLASSIFIER***

(xvi + 49 halaman: 10 gambar; 6 tabel; 4 lampiran)

Riset Kesehatan Dasar (Riskesdas) yang dilakukan pada tahun 2018 melakukan pengumpulan data penderita Diabetes Melitus pada penduduk berumur  $\geq 15$  tahun. Kriteria Diabetes Melitus pada Riset Kesehatan Dasar (Riskesdas) 2018 mengacu pada konsensus Perkumpulan Endokrinologi Indonesia (PERKENI) yang mengikuti kriteria *American Diabetes Association* (ADA). Hasil Riskesdas 2018 menyatakan bahwa prevalensi Diabetes Melitus di Indonesia berdasarkan diagnosis dokter pada umur  $\geq 15$  tahun sebesar 2%. Angka ini menunjukkan peningkatan dibandingkan prevalensi Diabetes Melitus pada tahun 2013 sebesar 1,5%. Data mining dapat digunakan karena dapat membuat hasil prediksi penyakit Diabetes Melitus pada penderita dengan memanfaatkan metode matematika atau statistika sehingga hasil Diabetes Melitus dapat diprediksi dengan benar dan akurat. *Naive Bayes Classifier* merupakan sebuah metode klasifikasi yang berakar pada teorema Bayes. Pada penelitian ini, *dataset* yang digunakan ialah data dari perempuan yang berusia diatas 21 tahun. Data mentah untuk penelitian ini terdiri atas 8 kolom dan 768 baris. Data tersebut mempunyai atribut sebanyak 8 yang terdiri atas *Pregancies, Glucose, Blood Pressure, Skin Thickness, Insulin, BMI (Body Mass Index), Diabetes Pedigree Function, Age* dan *Outcome*. Dari hasil penelitian Nilai keakuratan (*accuracy*) yang didapatkan dengan menggunakan metode Naive Bayes untuk mengklasifikasikan penyakit Diabetes Melitus sebesar 78%. Dan untuk hasil dari perhitungan *precision* untuk yang terkena Diabetes Melitus adalah 81% dan yang tidak Diabetes Melitus adalah 72%. Dan untuk *recall* untuk yang terkena Diabetes Melitus adalah 87% dan 61% untuk yang tidak Diabetes Melitus. Sedangkan untuk *f1-score* yang terkena Diabetes Melitus adalah 84% dan 66 untuk yang tidak Diabetes Melitus.

**Kata Kunci:** Diabetes Mellitus, *Naive Bayes Classifier*, *Data Mining*