

ABSTRAK

Shalsabilillah Naufal. Diagnosa Penyakit Ginjal Berdasarkan Data Pasien ICU Dengan Algoritma CatBoost Skripsi, Program Studi Teknik Informatika, Fakultas Teknologi Informasi. Universitas Yarsi, Jakarta 2021.

Penyakit Ginjal Kronis (PGK) merupakan jenis penyakit ginjal yang terjadi hilangnya fungsi ginjal secara bertahap. Beberapa kondisi, seperti diabetes dan tekanan darah tinggi, menjadi penyebab terjadinya PGK. Dalam jangka Panjang, kondisi-kondisi ini bisa menyebabkan kerusakan pada ginjal sehingga fungsi ginjal menurun. Prevalensi Penyakit Ginjal Kronis (PGK) meningkat seiring meningkatnya jumlah penduduk usia lanjut dan kejadian penyakit diabetes melitus serta hipertensi. Sekitar 1 dari 10 populasi global mengalami Penyakit Ginjal Kronis (PGK) pada stadium tertentu. Hasil systematic review dan metaanalysis yang dilakukan oleh Hilll et al, 2016, mendapatkan prevalensi global PGK sebesar 13.4%. Menurut hasil Global Burden of Disease tahun 2010, PGK merupakan penyebab kematian peringkat ke-27 didunia tahun 1990 dan meningkat menjadi urutan ke-18 pada tahun 2010. Sedangkan di Indonesia, perawatan penyakit ginjal merupakan ranking kedua pembiayaan terbesar dari BPJS Kesehatan setelah penyakit jantung. Oleh sebab itu, penulis memanfaatkan menggunakan *machine learning* untuk mendiagnosa Penyakit Ginjal. Penulis melakukan eksperimen dengan menggunakan data RS. Mortality pada pasien yang dirawat di Intensive Care Units (ICU), data tersebut merupakan database (MIMIC-III) perawatan kritis yang tersedia untuk umum. Pada penelitian ini, penulis mengimplementasikan algoritma CatBost untuk mendeteksi Penyakit Ginjal Kronis (PGK) karena algoritma ini mampu berjalan dengan baik dan mampu memberikan akurasi yang bagus dalam data kategori. Penulis menggunakan teknik *data cleaning* dan *fill missing value*. Penulis membandingkan performa prediksi dengan algoritma *Random Forest*, *Decision Tree Classifier*, dan *CatBoost*. Dengan metode yang dilakukan algoritma CatBoost mendapatkan AUC score tertinggi sebesar 92.21%, sedangkan untuk Random Forest sebesar 89.28%, dan Decision Tree sebesar 70.23%.

Kata Kunci : Machine Learning, Data Cleaning, Fill Missing Value, CatBost.