

KOMPARASI ALGORITMA KLASIFIKASI DALAM MEMPREDIKSI TIMBULNYA PENYAKIT STROKE

Fenna Kemala Hutabarat

Fakultas Teknologi dan Ilmu Komputer Universitas Prima Indonesia

ABSTRAK

Untuk mencegah terjadinya penyakit stroke diperlukan sebuah cara untuk memprediksi apakah seseorang sudah terkena stroke melalui parameter medis. Dengan masuknya pengaruh teknologi kedalam dunia medis, kini penyakit stroke dapat diprediksi dengan menggunakan metode *Data Science* yang bermula dari *Data Acquisition*, *Data Cleaning*, *Exploratory Data Analysis*, *Preprocessing* dan tahapan terakhir adalah *Model Building*. Berdasarkan model yang telah dibuat, disimpulkan bahwasanya algoritma dengan performa terbaik pada kasus ini adalah *XGBoost* dengan nilai *precision* 0.9, nilai *recall* sebesar 0.95, nilai *f1* sebesar 0.92 dan nilai *ROC-AUC* sebesar 0.978 setelah mendapat *cross validation* sebanyak 5 *folds*. Dengan hasil demikian, model yang telah dibuat dapat digunakan untuk melakukan prediksi secara *real time*.

Kata Kunci – *Machine Learning*, *Logistic Regression*, *Random Forest*, *XGBoost*, *Stroke*

Comparison of Classification Algorithm in Predicting Stroke Disease

Fenna Kemala Hutabarat

Faculty of Technology and Computer Science Universitas Prima Indonesia

ABSTRACT

To prevent stroke, there must be some way to predict whether someone has had a stroke through medical parameters. With the influence of technology in the medical world, stroke can now be predicted using the Data Science method, starting from Data Acquisition, Data Cleaning, Exploratory Data Analysis, Preprocessing, and the last stage is Model Building. Based on the model that has been made, it is concluded that the algorithm with the best performance, in this case, is XGBoost with a precision value of 0.9, a recall value of 0.95, and f1 value of 0.92 and a ROC-AUC value of 0.978 after getting five folds of cross-validation. With these results, the model created can be used to make predictions in real-time.

Keywords – Machine Learning, Logistic Regression, Random Forest, XGBoost, Stroke.
