

BAB 1

PENDAHULUAN

1.1. Latar Belakang

Penyakit stroke adalah sebuah kondisi dimana tersumbatnya aliran darah ke otak yang dapat menyebabkan kematian sel [1]. Menurut data World Health Organization (WHO), penyakit stroke adalah penyebab kematian kedua terbanyak sedunia, yaitu 11% dari total kematian. Dapat disimpulkan bahwasanya penyakit stroke adalah penyakit yang sangat mematikan bila tidak dapat ditangani secara cepat dan tepat [2]. Untuk mencegah terjadinya stroke diperlukan cara untuk memprediksi stroke melalui parameter medis. Dengan masuknya pengaruh teknologi kedalam dunia medis, penyakit stroke dapat diprediksi dengan *Machine Learning* berbasis algoritma klasifikasi, sehingga dengan parameter medis yang telah ditentukan, *model Machine Learning* yang telah dibuat dapat memprediksi apakah seseorang terkena penyakit stroke.

Penelitian sebelumnya yang melakukan prediksi penyakit telah banyak dilakukan, sebagai contoh penelitian [3] yaitu “*Prediction of Heart Disease Using Machine Learning*” dan penelitian [4] yaitu “*Long Short-Term Memory Recurrent Neural Network for Stroke Prediction*”. Penelitian [3] menyajikan model *Neural Network* dengan *Multi Level Perceptron* (MLP) yang dapat memprediksi penyakit jantung. *Output* yang dibuat pada penelitian [3] akan memberikan hasil prediksi dalam bentuk “YES/NO”. Penelitian [4] menyajikan model *Recurrent Neural Network* dengan *Long Short-Term Memory* (LSTM) yang dapat memprediksi penyakit stroke. *Output* dari penelitian [4] merupakan detail hasil model yang telah dibuat, seperti nilai *Precision*, *Recall*, *F1* dan *Accuracy*.

Berdasarkan masalah yang dijelaskan pada paragraf sebelumnya, untuk melakukan deteksi awal penyakit stroke, penulis menyarankan untuk menganalisa parameter medis yang dapat menyebabkan penyakit stroke dan pembuatan model *Machine Learning* yang dapat melakukan prediksi penyakit stroke. Dalam penelitian ini, akan dibuat beberapa algoritma klasifikasi seperti *Logistic Regression*, *Random Forest Classifier* dan *Extreme Boosting Classifier*. Algoritma dengan nilai *precision*, *recall*, *F1* dan *ROC-AUC* terbaik akan disimpulkan sebagai algoritma terbaik dalam kasus ini.

1.2. Rumusan Masalah

Berdasarkan uraian masalah yang telah ditemukan, maka rumusan masalah yang terdapat pada penelitian ini adalah:

1. Bagaimana cara mendapatkan faktor utama penyebab penyakit stroke guna mencegahnya sejak awal.
2. Bagaimana cara melakukan prediksi terhadap penyakit stroke berdasarkan parameter medis yang telah ada.

1.3. Batasan Masalah

Batasan masalah pada penelitian ini adalah :

1. Menggunakan algoritma klasifikasi dalam pembuatan model;
2. *Dataset* yang dipakai dalam penelitian ini berasal dari Kaggle.
3. Parameter yang dipakai dalam penelitian ini adalah faktor-faktor medis yang dapat menyebabkan seseorang terkena penyakit stroke.

1.4. Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk menganalisa parameter medis yang merupakan faktor penyebab penyakit stroke dengan *Exploratory Data Analysis* dan melakukan prediksi penyakit stroke dengan parameter medis tersebut.

1.5. Manfaat Penelitian

Manfaat yang terdapat pada penelitian ini adalah :

1. Memudahkan masyarakat untuk mendeteksi penyakit stroke lebih awal;
2. Memudahkan tenaga kesehatan dalam melakukan diagnosa awal penyakit stroke;
3. Membuka wawasan lebih terhadap dunia aviasi dan *machine learning*