

## ABSTRAK

Nama : Muhammad Yan Silmi Kaaffah  
NIM : 41819010104  
Nama : Muhammad Bima Aditya  
NIM : 41819010040  
Pembimbing TA : Sarwati Rahayu, ST., MMSI  
Judul : DATA MINING UNTUK PREDIKSI PENYAKIT  
STROKE PADA SEMUA USIA  
MENGUNAKAN METODE K-NEAREST  
NEIGHBOUR

Stroke adalah masalah kesehatan ketika aliran darah ke otak terputus atau terhenti sama sekali, menyebabkan jaringan otak kekurangan oksigen dan nutrisi. Stroke merupakan penyakit dengan jumlah kematian tertinggi kedua, terutama karena kecacatan. Dari Organisasi Kesehatan Dunia saja, 70% stroke disebabkan secara global dalam beberapa tahun terakhir, Machine Learning (ML) telah membentuk ruang potensial di bidang medis. Model pembelajaran mesin dapat diterapkan data mining adalah aktivitas yang mengumpulkan data historis dan menggunakannya untuk menemukan pola reguler dalam data dalam jumlah besar. Data yang diteliti merupakan hasil pemeriksaan terhadap 5110 orang. Hasil akurasi yang tertinggi pada penelitian ini yang menggunakan algoritma klasifikasi K-Nearest Neighbor, nilai accuracy sistem yang diperoleh oleh K-Nearest Neighbor sebesar 97,3%. Didapati 1672 data uji yang benar diprediksi sebagai orang yang tidak terkena stroke (True Positive), 1623 data uji yang diprediksi sebagai orang stroke tidak terkena stroke (True Negative), 14 data uji yang diprediksi sebagai orang stroke yang dimana merupakan orang tidak stroke (False Negative), dan 79 data uji yang diprediksi sebagai orang tidak stroke yang dimana orang tersebut merupakan orang stroke (False Positive).

Kata kunci:

Stroke, prediksi, *K-Nearest Neighbour*, *dataset*

## ABSTRACT

Name : Muhammad Yan Silmi Kaaffah  
Student Number : 41819010104  
Name : Muhammad Bima Aditya  
Student Number : 41819010040  
Counsellor : Sarwati Rahayu, ST., MMSI  
Title : DATA MINING UNTUK PREDIKSI PENYAKIT  
STROKE PADA SEMUA USIA  
MENGUNAKAN METODE K-NEAREST  
NEIGHBOUR

Stroke is a health problem when blood flow to the brain is interrupted or completely stopped, causing brain tissue to lack oxygen and nutrients. Stroke is a disease with the second highest number of deaths, mainly due to disability. From the World Health Organization alone, 70% of strokes are caused globally in recent years, Machine Learning (ML) has formed a potential space in the medical field. Machine learning models can be applied to data mining is an activity that collects historical data and uses it to find regular patterns in large amounts of data. The data studied is the result of examining 5110 people. The highest accuracy results in this study used the K-Nearest Neighbor classification algorithm, the system accuracy value obtained by K-Nearest Neighbor was 97.3%. It was found that 1672 test data were correctly predicted as people who were not affected by a stroke (True Positive), 1623 test data predicted as stroke people not having a stroke (True Negative), 14 test data predicted as stroke people which were people who did not have a stroke (False Negative), and 79 test data predicted as people did not have a stroke where the person was stroke people (False Positive).

*Keywords:*

*Stroke, Predict, K-Nearest-Neighbour, Dataset*