

# Stroke Prediction using Machine Learning

Yusra Erlangga Putra<sup>1</sup>, Sheryl Anastasya<sup>2</sup>, Resky Auliyah Kartini Askin<sup>3</sup>, Ivan Betrandi<sup>4</sup>, Amaliah Diah<sup>5</sup>  
Departemen Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Hasanuddin, Makassar, Indonesia

<sup>1</sup>yusraerlangg@gmail.com, <sup>2</sup>jane.doe@email.com

\*Corresponding author

## Abstrak

Stroke merupakan penyebab utama kecacatan jangka panjang dan kematian di seluruh dunia, dengan risiko yang meningkat seiring bertambahnya usia serta adanya faktor risiko seperti hipertensi dan diabetes. Tujuan penelitian ini adalah mengembangkan model prediksi dini dan akurat untuk stroke guna memungkinkan intervensi perawatan kesehatan preventif yang efektif. Studi ini menggunakan algoritma Random Forest dan Support Vector Machine (SVM). Karena ketidakseimbangan pada dataset, teknik Synthetic Minority Oversampling Technique (SMOTE) diterapkan untuk meningkatkan representasi data minoritas. Model dioptimalkan melalui penyetelan hiperparameter menggunakan Bayesian Optimization. Evaluasi model dilakukan dengan metrik akurasi, precision, recall, dan F1-score, dengan validasi silang untuk memastikan keandalan pada data yang belum terlihat. Hasil eksperimen menunjukkan bahwa algoritma SVM dengan SMOTE dan optimasi Bayesian mencapai akurasi tertinggi sebesar 0,726404. Temuan ini menunjukkan bahwa model pembelajaran mesin yang dioptimalkan dapat memberikan kontribusi signifikan dalam prediksi dini stroke dan mendukung pengambilan keputusan dalam sistem perawatan kesehatan preventif.

**Kata kunci:** Prediksi Stroke, pembelajaran mesin, Random Forest, Support Vector Machine, SMOTE, Bayesian Optimization

## 1. Introduction

### 1.1 Backgorund

### 1.2 Literature Review

### 1.3 Research Rationale

### 1.4 Research Questions and Objectives

## 2. Research Methods

### 2.1 Formulas

$$E = mc^2 \quad (1)$$

## 3. Results and Discussion

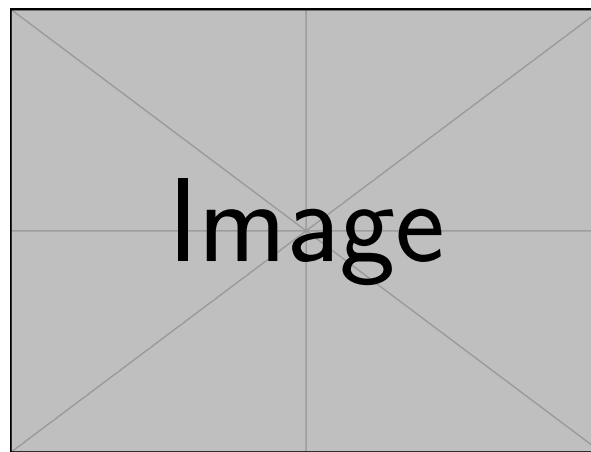


Figure 2: A single-column figure example.

## 4. Conclusion

## Acknowledgements

## References

[1]

[2]

[3]

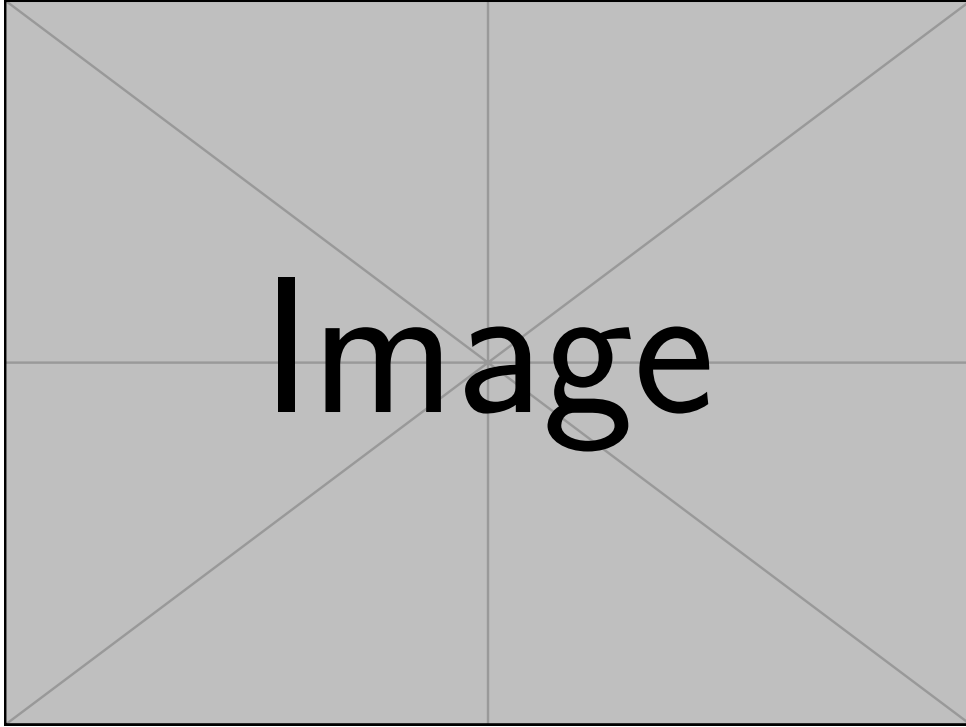


Figure 1: Description of your figure. Make sure to reference it in text using Fig. 1.

Table 1: Dataset characteristics for stroke prediction

Characteristic	Value	Percentage
Total Samples	5110	100%
Stroke Cases	249	4.87%
Non-stroke Cases	4861	95.13%