

## **ABSTRACT**

### **OPTIMIZING HYPERPARAMETER TUNING USING GRIDSEARCHCV FOR THE RANDOM FOREST AND SUPPORT VECTOR MACHINE (SVM) METHODS IN CLASSIFYING THE HUMAN DEVELOPMENT INDEX STATUS IN INDONESIA IN 2022**

**By**

**NURUL HIDAYAH ITSNAINI**

Random Forest and Support Vector Machine (SVM) methods are among the techniques often used in classification. In an effort to build an optimal classification model, determining the right parameters or known as hyperparameter tuning, is a major challenge that can affect the performance of a model, and Grid Search Cross-Validation (GridSearchCV) is one method that can be used to obtain the optimal hyperparameters. This study aims to optimize hyperparameter tuning using GridSearchCV to improve classification accuracy of the Human Development Index (HDI) status, and compare the performance of Random Forest and SVM models. The analysis results show that the use of GridSearchCV is effective in improving classification accuracy of both methods. In the Random Forest method with a data split of 80% training data and 20% testing data, there was an increase in accuracy from 97,31% to 98,38%. Similarly, the SVM method with linear kernel with a data split of 70% training data and 30% testing data, which achieved an accuracy value from 98,2% to 99,28%, making SVM a better method compared to Random Forest in classifying HDI data.

**Keywords:** Classification, Random Forest, SVM, Hyperparameter Tuning, GridSearchCV, Human Development Index

## ABSTRAK

### OPTIMASI *HYPERPARAMETER TUNING* MENGGUNAKAN *GRIDSEARCHCV* PADA METODE *RANDOM FOREST* DAN *SUPPORT VECTOR MACHINE* (SVM) UNTUK KLASIFIKASI STATUS INDEKS PEMBANGUNAN MANUSIA DI INDONESIA TAHUN 2022

Oleh

NURUL HIDAYAH ITSNAINI

Metode *Random Forest* dan *Support Vector Machine* (SVM) merupakan salah satu teknik yang sering digunakan dalam pengklasifikasian. Dalam upaya membangun model klasifikasi yang optimal, penentuan parameter yang tepat, atau dikenal sebagai *hyperparameter tuning*, menjadi tantangan utama yang dapat mempengaruhi kinerja suatu model, dan metode *Grid Search Cross-Validation* (*GridSearchCV*) merupakan salah satu metode yang dapat dipilih untuk memperoleh *hyperparameter* yang optimal. Penelitian ini bertujuan untuk mengoptimalkan *hyperparameter tuning* menggunakan *GridSearchCV* guna meningkatkan akurasi klasifikasi terhadap status Indeks Pembangunan Manusia (IPM), serta membandingkan kinerja model *Random Forest* dan SVM. Hasil analisis menunjukkan bahwa penggunaan *GridSearchCV* efektif dalam meningkatkan akurasi klasifikasi kedua metode. Pada metode *Random Forest* dengan *split* data 80% data *training* dan 20% data *testing*, terjadi peningkatan akurasi dari 97,31% menjadi 98,38%. Begitu pula pada metode SVM dengan kernel linear dengan *split* data 70% data *training* dan 30% data *testing*, yang mencapai nilai akurasi dari 98,2% menjadi 99,28%, menjadikan SVM sebagai metode yang lebih baik dibandingkan dengan *Random Forest* dalam mengklasifikasikan data IPM.

**Kata Kunci:** Klasifikasi, *Random Forest*, SVM, *Hyperparameter Tuning*, *GridSearchCV*, Indeks Pembangunan Manusia