

IMPLEMENTASI ALGORITMA *EXTRA TREES* UNTUK KLASIFIKASI CUACA PROVINSI DKI JAKARTA DENGAN *OVERSAMPLING SMOTE*

RAIHAN KEMMY RACHMANSYAH

ABSTRAK

Cuaca merupakan keadaan udara pada suatu wilayah tertentu dengan jangka waktu yang terbatas. Cuaca di Indonesia khususnya di Provinsi DKI Jakarta sangat tidak menentu serta sulit untuk diprediksi. Cuaca yang sulit diprediksi membuat banyak aktivitas warga Provinsi DKI Jakarta terganggu sehingga diperlukan sebuah ilmu teknologi yang diimplementasikan untuk mengklasifikasikan sebuah cuaca. Maka dari itu, penelitian ini menerapkan metode *Machine Learning* untuk klasifikasi cuaca Provinsi DKI Jakarta menggunakan algoritma *Extra Trees* dengan metode *oversampling SMOTE*. Data yang digunakan pada penelitian ini adalah data Prakiraan Cuaca Provinsi DKI Jakarta pada tahun 2017 hingga 2018 yang diperoleh dari situs <https://data.jakarta.go.id/>. Data yang diperoleh memiliki distribusi data yang tidak seimbang sehingga perlu diseimbangkan terlebih dahulu menggunakan metode *resampling* data menggunakan *Synthetic Minority Oversampling Technique* atau *SMOTE*. Berdasarkan hasil penelitian, metode *oversampling SMOTE* tidak mempengaruhi hasil evaluasi pada klasifikasi cuaca Provinsi DKI Jakarta menjadi lebih baik. Hasil evaluasi terbaik didapatkan oleh model menggunakan algoritma *Extra Trees* tanpa metode *oversampling SMOTE* pada rasio 80% data latih dan 20% data uji dengan nilai akurasi sebesar 79,8%, *precision* 63,1%, dan *recall* 56,1%.

Kata Kunci: Klasifikasi, *Extra Trees*, *Oversampling SMOTE*, Cuaca Provinsi DKI Jakarta

**IMPLEMENTATION OF THE EXTRA TREES ALGORITHM FOR
WEATHER CLASSIFICATION OF DKI JAKARTA PROVINCE WITH
OVERSAMPLING SMOTE**

RAIHAN KEMMY RACHMANSYAH

ABSTRACT

Weather is the state of the air in a certain area for a limited period of time. The weather in Indonesia, especially in DKI Jakarta Province, is very erratic and difficult to predict. Weather that is difficult to predict disrupts many activities of DKI Jakarta Province residents, so a technological science is needed to classify weather. Therefore, this research applies the Machine Learning method for weather classification in DKI Jakarta Province using the Extra Trees algorithm with the SMOTE oversampling method. The data used in this research is DKI Jakarta Province Weather Forecast data from 2017 to 2018, obtained from the site <https://data.jakarta.go.id/>. The data obtained has an unbalanced data distribution, so it needs to be balanced first using the data resampling method using the Synthetic Minority Oversampling Technique, or SMOTE. Based on the research results, the SMOTE oversampling method does not affect the evaluation results on the weather classification of DKI Jakarta Province for the better. The best evaluation results were obtained by the model using the Extra Trees algorithm without the SMOTE oversampling method at a ratio of 80% training data and 20% test data, with an accuracy value of 79,8%, precision of 63,1%, and recall of 56,1%.

Keywords: *Classification, Extra Trees, Oversampling SMOTE, Weather of DKI Jakarta Province*