

ABSTRAK

CLUSTERING DAN KLASIFIKASI DATA CUACA KOTA CILACAP MENGUNAKAN K-MEANS DAN RANDOM FOREST

Fadil Danu Rahman

Pengamatan dan analisis data cuaca merupakan aspek penting dalam memahami kondisi atmosfer di suatu wilayah. Penelitian ini bertujuan untuk menganalisis data cuaca berbasis BMKG dari Stasiun Meteorologi Tunggul Wulung Cilacap menggunakan metode *K-Means clustering* dan algoritma *Random Forest*. Data cuaca dari tahun 1975 hingga 2023 diambil untuk mengidentifikasi pola dan karakteristik unik dalam kondisi atmosfer.

Metode *K-Means clustering* digunakan untuk membentuk *cluster*, yang kemudian digunakan sebagai dasar untuk klasifikasi kondisi cuaca dengan algoritma *Random Forest*. Melalui penggunaan algoritma *Random Forest*, model klasifikasi berhasil memprediksi kondisi cuaca dengan tingkat akurasi yang memuaskan. Meskipun demikian, penurunan performa pada rentang tahun 2018-2023 menunjukkan adanya tantangan dalam memodelkan pola cuaca yang kompleks.

Analisis menggunakan metode *Elbow* dan *Silhouette* menunjukkan jumlah *cluster* optimal dan evaluasi kualitas pengelompokan. Implikasi temuan ini diharapkan dapat memberikan manfaat dalam pemahaman dan prakiraan cuaca yang lebih akurat, dengan potensi dampak positif pada berbagai sektor, seperti pertanian dan transportasi. Dengan memadukan teknik *clustering* dan klasifikasi, penelitian ini membuka peluang untuk pengembangan lebih lanjut dalam analisis cuaca berbasis data.

Kata kunci : Cuaca, *K-Means clustering*, *Random Forest*

ABSTRACT

CLUSTERING AND CLASSIFICATION OF CILACAP CITY WEATHER DATA USING K-MEANS AND RANDOM FOREST

Fadil Danu Rahman

Observing and analyzing weather data is crucial in understanding atmospheric conditions in a particular region. This study aims to analyze BMKG-based weather data from the Tunggul Wulung Cilacap Meteorological Station using the K-Means clustering method and Random Forest algorithm. Weather data from 1975 to 2023 were collected to identify unique patterns and characteristics in atmospheric conditions.

The K-Means clustering method was utilized to form clusters, which were then used as a basis for weather condition classification using the Random Forest algorithm. Through the implementation of the Random Forest algorithm, the classification model successfully predicted weather conditions with satisfactory accuracy. However, a decrease in performance during the 2018-2023 period indicates challenges in modeling complex weather patterns.

Analysis using the Elbow and Silhouette methods revealed the optimal number of clusters and evaluated the quality of clustering. The implications of these findings are expected to provide benefits in understanding and forecasting more accurate weather conditions, with potential positive impacts on various sectors such as agriculture and transportation. By integrating clustering and classification techniques, this research opens opportunities for further development in data-driven weather analysis.

Keywords: Weather, K-Means clustering, Random Forest