

ABSTRAK

PRAKIRAAN CUACA KOTA BANDAR LAMPUNG DENGAN MENGUNAKAN *RANDOM FOREST* DAN *C4.5*

Oleh

RAHMA FERIKHA SHAUMI

Prakiraan cuaca merupakan informasi yang penting bagi berbagai instansi maupun masyarakat luas. Prakiraan cuaca biasanya digunakan untuk kepentingan berbagai sektor seperti transportasi, pariwisata perkebunan dan lain-lain. Penelitian ini bertujuan untuk membuat model baru mengenai prakiraan cuaca dengan menggunakan algoritma random forest dan C4.5 dengan menggunakan aplikasi WEKA. Dataset menggunakan data Stasiun Meteorologi Maritim Panjang dengan banyaknya data 365 hari dengan 6 atribut.

Dataset yang digunakan terlebih dahulu melalui *preprocessing* data, *replace missing values*, *training* data dan penerapan kedua algoritma hingga menghasilkan prediksi yang kemudian dievaluasi dengan MAE dan RMSE. Hasil dari penelitian ini yaitu algoritma *C.4.5* lebih baik pada uji coba dengan *cross-validation folds* 5, 10 dan 15 dibandingkan dengan algoritma *random forest*, yaitu dengan nilai MAE 0.0284, 0.0277, dan 0.0285 serta nilai RMSE 0.1408, 0.1392 dan 0.1414. Sedangkan hasil dari *random forest* untuk nilai MAE adalah 0.1131, 0.112 dan 0.1119 serta nilai RMSE 0.216, 0.2146 dan 0.2146. Kemudian algoritma *C4.5* juga lebih baik pada uji coba *percentage split* 25%, 66% dan 90% dibandingkan algoritma *random forest* yaitu dengan nilai MAE 0.0613, 0.0329 dan 0.0234 serta nilai RMSE 0.2287, 0.1463 dan 0.102. Sedangkan hasil dari *random forest* untuk nilai MAE adalah 0.111, 0.1202 dan 0.1295 serta nilai RMSE 0.2218, 0.2428 dan 0.2523.

Kata kunci: prakiraan cuaca, *random forest*, *C4.5*, WEKA.

ABSTRACT

WEATHER FORECAST FOR BANDAR LAMPUNG CITY USING RANDOM FOREST AND C4.5

Oleh

RAHMA FERIK SHAUMI

Weather forecasts are essential information for various agencies and the wider community. Weather forecasts are usually used to benefit multiple sectors such as transportation, tourism, plantations, and others. This study aims to create a new model regarding weather forecasting using the random forest and C4.5 algorithms using the WEKA application. The dataset uses data from the Long Maritime Meteorological Station with 365 days of data with six attributes.

The dataset used first goes through data preprocessing, replacing missing values, training data, and applying the two algorithms to produce predictions which are then evaluated with MAE and RMSE. The results of this study are that the C4.5 algorithm is better in trials with cross-validation folds 5, 10, and 15 compared to the random forest algorithm, with MAE values of 0.0284, 0.0277, and 0.0285 and RMSE values of 0.1408, 0.1392 and 0.1414. While the results of the random forest for the MAE values are 0.1131, 0.112, and 0.1119, and the RMSE values are 0.216, 0.2146, and 0.2146. Then the C4.5 algorithm is also better in the percentage split trials of 25%, 66%, and 90% compared to the random forest algorithm, with MAE values of 0.0613, 0.0329, and 0.0234 and RMSE values of 0.2287, 0.1463 and 0.102. While the results of the random forest for the MAE values are 0.111, 0.1202, and 0.1295, and the RMSE values are 0.2218, 0.2428, and 0.2523.

Keywords: weather forecast, random forest, C4.5, WEKA.