

ABSTRAK

Indonesia memiliki iklim tropis yang dominan karenanya, Indonesia mengalami variasi suhu yang terbatas, namun memiliki variasi curah hujan yang beragam. Variabilitas curah hujan juga tidak lepas dengan dampak yang diberikannya pada berbagai aspek kehidupan manusia dan aktivitas bisnis. Oleh karena itu informasi curah hujan merupakan aspek penting dalam pengambilan keputusan. Namun, tentunya perlu tahapan dan metode untuk melakukan proses analisa. Maka dari itu penelitian ini bertujuan untuk mencari metode terbaik antara C4.5 dan *K-Nearest Neighbors* yang termasuk algoritma pada *data mining* untuk mengklasifikasi data curah hujan. Kedua algoritma digunakan untuk membangun model klasifikasi berdasarkan atribut atribut yang relevan. Kemudian, model-model tersebut diuji dan dievaluasi menggunakan berbagai metrik seperti Akurasi, *Precision*, *Recall* dan *F1-Score*. Dalam penelitian ini juga menerapkan *Tuning Hyperparameter* dengan metode *RandomizeSearchCV* untuk mendapatkan parameter terbaik yang dapat menghasilkan nilai akurasi yang maksimal. Hasil penelitian menunjukkan bahwa kedua algoritma memiliki kinerja yang baik dalam klasifikasi curah hujan. Jika berdasarkan nilai akurasi yang didapat dengan parameter *default* dari kedua algoritma, C4.5 memiliki nilai akurasi yang lebih tinggi sebesar 81.42%, sementara *K-Nearest Neighbors* hanya sebesar 78.10%. Namun setelah menggunakan parameter terbaik hasil dari penerapan *Tuning Hyperparameter RandomizedSearchCV*, perubahan nilai akurasi yang cukup signifikan terjadi pada *K-Nearest Neighbors* yang didapati sebesar 83.37%, sementara C4.5 bertambah menjadi 82.56%.

Kata Kunci : curah hujan, iklim Indonesia, klasifikasi, algoritma C4.5, K-Nearest Neighbor, data mining.