

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Kesimpulan yang diperoleh dari hasil Klasifikasi Cuaca Berdasarkan Faktor Penentu Hujan dengan Menggunakan Algoritma *K-Nearest Neighbor* yaitu sebagai berikut:

1. Algoritma *K-Nearest Neighbor* dapat digunakan untuk mengklasifikasikan cuaca dengan perhitungan jarak dengan *Euclidean Distance*. Data yang digunakan diambil dari website yaitu <http://dataonline.bmkg.go.id/> dan mengambil data pada Kabupaten Tuban dengan rentang waktu 01 Januari 2022 hingga 31 Januari 2023. Dataset cuaca harian di Kabupaten Tuban yang telah dikumpulkan terdapat 396 *record* yang terdiri dari sembilan atribut penentu hujan dan satu atribut label. Setelah itu, dilakukan transformasi data dan validasi data menggunakan split data dengan ratio 90:10. Sebanyak 256 data digunakan sebagai data latih dan sebanyak 40 data digunakan sebagai data uji. Setelah itu, dilakukan perhitungan jarak menggunakan *Euclidean Distance*. Untuk menemukan nilai *k* yang terbaik dilakukan beberapa percobaan dengan mengubah nilai *k* sebanyak 3 kali yaitu 3,5 dan 7. Adapun pada nilai *k* sebesar 3 menghasilkan *accuracy* sebesar 70%. Pada nilai *k* sebesar 5 menghasilkan *accuracy* sebesar 82,5%. pada nilai *k* sebesar 7 menghasilkan *accuracy* sebesar 77,5%.
2. Evaluasi algoritma *K-Nearest Neighbor* pada klasifikasi cuaca berbasis penentu hujan menggunakan confusion matrix, dengan percobaan mengubah nilai *k* sebanyak tiga kali. Adapun pada nilai *k* sebesar 3 menghasilkan nilai *precision* sebesar 0,89, *recall* sebesar 0,90 dan *accuracy* sebesar 70%. pada nilai *k* sebesar 5 menghasilkan nilai *precision* sebesar 0,89, *recall* 0,91 dan *accuracy* sebesar 82,5%. pada nilai *k* sebesar 7 menghasilkan nilai *precision* sebesar 0,83, *recall* sebesar 0,90 dan *accuracy* sebesar 77,5%.

5.2 Saran

Dalam penelitian ini pastinya memiliki kekurangan yang selanjutnya diperlukan perbaikan. Oleh karenanya, penulis menyarankan agar dapat menjadi bahan pertimbangan selanjutnya:

1. Pada penelitian ini untuk pengukuran jarak terdekatnya menggunakan *euclidean distance*, pengukuran jarak pada algoritma *K-Nearest Neighbor* sangat dipengaruhi oleh pengukuran jarak yang digunakan karena sangat berpengaruh pada akurasi. Sehingga untuk penelitian selanjutnya perlu adanya pengkajian yang lebih lanjut dengan menggunakan pengukuran jarak lain seperti *Minkowski Distance*, *Chebychev Distance*, *Manhattan Distance* untuk membandingkan keefektifan penggunaan fungsi jarak pada *K-Nearest Neighbor* dengan asumsi penggunaan nilai k 3, 5, dan 7.
2. Pada algoritma *K-Nearest Neighbor* antar *variable* saling memiliki keterkaitan, maka pada penelitian selanjutnya untuk melakukan pengkajian lebih lanjut mengenai faktor yang paling mempengaruhi hujan, sehingga tahap penyeleksian fitur bisa maksimal dan dapat meningkatkan nilai akurasi.
3. Mengembangkan sistem dengan menggunakan algoritma lainnya seperti *Naïve Bayes*, *Decision Tree*, dan lain-lain.
4. Menambahkan fitur seperti import excel untuk mempermudah dalam memasukkan data ke dalam sistem.

UNUGIRI