

ABSTRAK

Indonesia merupakan wilayah yang hampir seluruhnya memiliki iklim tropis. Sebagai daerah yang kebanyakan memiliki iklim tropis, Indonesia mengalami variasi suhu yang sempit namun variasi curah hujan yang beragam. Curah hujan di Indonesia memiliki tingkat keberagaman yang signifikan. Untuk mencegah bahaya yang ditimbulkan oleh curah hujan tinggi, seperti banjir dan tanah longsor. Selain persiapan terhadap bencana, informasi mengenai curah hujan juga bermanfaat dalam bidang pertanian, transportasi, dan industri. Dengan implementasi klasifikasi *data mining*, akan membantu proses prediksi curah hujan di Indonesia. Penelitian ini menggunakan data iklim harian di Indonesia dan mengadopsi algoritma random forest sebagai metode klasifikasi. Algoritma tersebut dipilih karena dapat menghasilkan model klasifikasi yang akurat dan stabil tanpa memerlukan penyesuaian parameter yang kompleks. Selain itu, metode naïve bayes juga digunakan karena kemudahan implementasinya dan pemodelan probabilitas sederhana yang dapat diterapkan pada berbagai jenis data klasifikasi. Berdasarkan hasil penelitian, dapat disimpulkan bahwa algoritma random forest memiliki performa dan akurasi yang lebih baik daripada algoritma naïve bayes dalam proses klasifikasi dataset iklim di Indonesia. Algoritma random forest mencapai akurasi sebesar 86.55%, sementara algoritma naïve bayes hanya mencapai akurasi sebesar 36.61%. Hasil penelitian ini diharapkan dapat berguna sebagai referensi untuk studi literatur penelitian selanjutnya dan digunakan untuk memantau curah hujan harian di Indonesia guna mencegah bencana alam.

Kata kunci— *curah hujan, data mining, klasifikasi, naïve bayes, random forest*