

ABSTRACT

Weather forecasting is important for the life of the wider community. Therefore, the accuracy of weather forecasting must be high. Research related to weather forecasting has been carried out several times by various methods such as naïve bayes and C4.5. However, there are still some shortcomings such as the small number of datasets and categories used, so that it will have a big effect on accuracy. Based on this, the weather forecast optimization is carried out using the ensemble method on the Naïve Bayes algorithm and C4.5 with five categories. The dataset used is weather data observed from BMKG Bandung for 10 years.

The results of the study show that the use of the dataset from BMKG Bandung does not get maximum results because the value of each parameter has almost the same value in each category, making the pattern difficult to recognize and classify correctly. The results of classification and optimization show the accuracy of naïve Bayes is 49.45%, while for C4.5 is 41.24%. The results of the accuracy after optimization are not very significant. The increase in accuracy that occurs in the bagging process is 0.31% for naïve Bayes and 4.52% for C4.5. Then, the increase in accuracy in the boosting process is -2.69% for naïve Bayes and -2.42% for C4.5. Based on the experiments, the highest accuracy was obtained by the Naïve Bayes bagging model. Based on the description of the experimental results of this study, it is proposed to include forecasters who are experienced in weather forecasting research, so that the research results have good accuracy and can also be implemented directly in the weather forecasting process.

Keyword: C4.5, ensemble method, naïve bayes, weather forecasting.

ABSTRAK

Peramalan cuaca merupakan hal yang penting bagi kehidupan masyarakat luas. Oleh karena itu, akurasi dari peramalan cuaca haruslah tinggi. Penelitian terkait prakiraan cuaca telah beberap kali dilakukan dengan berbagai metode seperti naïve bayes dan C4.5. Namun, masih terdapat beberapa kekurangan seperti sedikitnya dataset dan kategori yang digunakan, sehingga akan berpengaruh besar pada akurasi. Berdasarkan hal itu maka dilakukan optimasi prakiraan cuaca menggunakan metode ensemble pada algoritma naïve bayes dan C4.5 dengan lima kategori. *Dataset* yang digunakan merupakan data cuaca hasil pengamatan dari BMKG Bandung selama 10 tahun.

Hasil dari penelitian menunjukkan penggunaan *dataset* dari BMKG Bandung tidak mendapatkan hasil yang maksimal karena nilai dari setiap parameter memiliki nilai yang hampir sama pada setiap kategori, sehingga membuat pola sulit dikenal dan diklasifikasikan dengan benar. Hasil klasifikasi dan optimasi menunjukkan akurasi dari *naïve bayes* adalah 49,45%, sedangkan untuk C4.5 adalah 41,24%. Hasil akurasi setelah dioptimasi tidak terlalu signifikan. Kenaikan akurasi yang terjadi pada proses *bagging* adalah 0,31% untuk *naïve bayes* dan 4,52% untuk C4.5. Lalu, kenaikan akurasi pada proses *boosting* adalah -2,69% untuk *naïve bayes* dan -2,42% untuk C4.5. Berdasarkan eksperimen yang dilakukan akurasi tertinggi didapatkan oleh model *bagging Naïve Bayes*. Berdasarkan uraian dari hasil eksperimen penelitian ini, diusulkan untuk mengikut sertakan *forecaster* yang sudah berpengalaman dalam penelitian peramalan cuaca, sehingga hasil penelitian

memiliki akurasi yang baik dan juga dapat diimplementasikan secara langsung dalam proses peramalan cuaca.

Kaca Kunci: C4.5, Metode *Ensemble*, Naïve Bayes, Peramalan cuaca.