

ANALISIS PERBANDINGAN ALGORITMA NAIVE BAYES DAN ALGORITMA C4.5 UNTUK KLASIFIKASI MULTI DATA

¹Dwi Luki Arisandy (1310651061), ²Victor Wahanggara, S. Kom., M. Kom.,

³Yeni Dwi Rahayu, S. ST., M. Kom.

Jurusan Teknik Informatika Fakultas Teknik Universitas Muhammadiyah Jember

Email: lukiarie111088@gmail.com

ABSTRAK

Klasifikasi merupakan suatu pekerjaan menilai objek data untuk memasukkannya ke dalam kelas tertentu dari sejumlah kelas yang tersedia. Metode Naïve Bayes memiliki tingkat akurasi tinggi pada saat proses klasifikasi. Hal tersebut diungkapkan oleh beberapa pendapat. Selain Naïve Bayes, algoritma C4.5 juga memiliki tingkat akurasi yang tinggi. Penelitian yang dilakukan oleh Puji Astuti (2016) berjudul “Komparasi Penerapan Algoritma C45, Knn Dan Neural Network Dalam Proses Kelayakan Penerimaan Kredit Kendaraan Bermotor” akurasi tertinggi yang dihasilkan C4.5 yaitu 92,89% dibandingkan K-nn 77,78% dan Nn 91,1%. Dari kedua penelitian diatas menunjukkan bawa algoritma Naive Bayes dan algoritma C4.5 memiliki tingkat akurasi yang tinggi pada proses klasifikasi data. Pada penelitia yang saya lakukan dengan membandingkan kedua metode antara Naive Baye dan C4.5 menggunakan 10 data set dengan varian jumlah data record dan atribut yang berbeda menunjukkan hasil rata-rata akurasi untuk Algoritma C4.5 pada K=5 yaitu 79,28% dan Naive bayes 77,58%. Pada K=10 didapatkan sebuah hasil rata-rata akurasi C4.5 yaitu 79,38% dan Naive bayes 77%. Pada Use Training Set (UTS) didapatkan sebuah hasil rata-rata akurasi C4.5 yaitu 89,46% dan Naive bayes 79,39%. Dari hasil akurasi ketiga teknik pengujian tersebut didapatkan kesimpulan bahwa algoritma C4.5 lebih unggul untuk tingkat akurasi, namun untuk kecepatan proses klasifikasi Naive Bayes lebih unggul.

Kata kunci :algoritma Naïve Bayes, C4.5, Analisis, Klasifikasi, Perbandingan algoritma.

ANALISIS PERBANDINGAN ALGORITMA NAIVE BAYES DAN ALGORITMA C4.5 UNTUK KLASIFIKASI MULTI DATA

¹Dwi Luki Arisandy (1310651061), ²Victor Wahanggara, S. Kom., M. Kom.,

³Yeni Dwi Rahayu, S. ST., M. Kom.

Jurusan Teknik Informatika Fakultas Teknik Universitas Muhammadiyah Jember

Email: lukiarie111088@gmail.com

ABSTRACT

Classification is a work of assessing a data object to include it in a certain class of available classes. Naïve Bayes method has a high degree of accuracy during the classification process. This is expressed by some opinions. In addition to Naïve Bayes, the C4.5 algorithm also has a high degree of accuracy. Research conducted by Puji Astuti (2016) titled "Comparison of Application of Algorithm C4.5, Knn And Neural Network In The Process Of Motor Vehicle Acceptability Eligibility" highest accuracy produced C4.5 is 92.89% compared to K-nn 77,78% and Nn 91.1%. From both the above research shows that Naive Bayes algorithm and C4.5 algorithm have a high degree of accuracy in the process of data classification. In a study I conducted by comparing the two methods between Naive Baye and C4.5 using 10 data sets with variants of different data records and attributes showed the average accuracy for Algorithm C4.5 at K = 5 ie 79.28% and Naive Bayes 77.58%. At K = 10 obtained an average yield accuracy of C4.5 that is 79.38% and Naive bayes 77%. In Use Training Set (UTS) got an average result of accuracy of C4.5 that is 89,46% and Naive bayes 79,39%. From the results of the accuracy of these three testing techniques it can be concluded that the C4.5 algorithm is superior to the accuracy level, but for the speed of the classification process Naive Bayes is superior.

Keywords: *Naïve Bayes algorithm, C4.5, Analysis, Classification, Comparison algorithm.*