

PERBANDINGAN ALGORITMA NAÏVE BAYES DAN DECISION TREE PADA ANALISIS SENTIMEN MEDIA SOSIAL TERHADAP MASKAPAI PENERBANGAN

Nama Mahasiswa : Nurul Nabila
NIM : 3011710047
Pembimbing : Lailatul Hidayah, S.Kom.,M.S.

ABSTRAK

Pesatnya perkembangan teknologi informasi saat ini membawa kehidupan manusia di era yang tak terbatas, salah satunya adalah media sosial. Twitter salah satu media sosial yang digunakan penggunanya untuk memberikan opini sebagai salah satu bentuk respon publik terhadap suatu hal. Industri kerap memanfaatkannya untuk mengumpulkan informasi di Twitter, salah satunya adalah industri penerbangan. Penelitian ini menganalisa sentimen masyarakat terhadap maskapai penerbangan dengan menggunakan sentimen positif, negatif dan netral. Penelitian ini menggunakan algoritma *Naïve Bayes* dan *Decision Tree*. Dalam pengujian algoritma juga menggunakan *grid search* untuk mencari parameter terbaik sehingga dapat meningkatkan performa algoritma *Decision Tree*. Dari proses pengujian hasil akurasi pada algoritma *Naïve Bayes* sebesar 87.0%, pada algoritma *Decision Tree Grid Search* sebesar 95.5%.

Parameter yang digunakan algoritma *Decision Tree* menggunakan *Grid Search* yaitu '*splitter*', '*max_features*', '*criterion*', '*max_dept*', dengan nilai '*max_depth*' 10–100 dan penentuan *best parameter* untuk *splitter* berdasarkan '*best*' dan '*random*', penentuan '*max_features*' berdasarkan '*auto*', '*sqrt*', '*log2*', penentuan '*criterion*' berdasarkan '*gini*', '*entropy*'. Sehingga dapat disimpulkan bahwa algoritma *Decision Tree* dengan *Grid Search* merupakan algoritma yang lebih baik dalam penelitian ini karena mendapatkan hasil performa yang lebih baik.

Kata kunci: Maskapai Penerbangan, *Naïve Bayes*, *Decision Tree*, *grid search*, Klasifikasi, Media Sosial, Twitter, Sentimen.

COMPARISON OF NAÏVE BAYES AND DECISION TREE ALGORITHM ON SOCIAL MEDIA SENTIMENT ANALYSIS OF AIRLINE

Student Name : Nurul Nabila
Identity Number : 3011710047
Supervisor : Lailatul Hidayah, S.Kom., M.S.

ABSTRACT

The rapid development of information technology today brings human life in an infinite era, one of which is social media. Twitter is one of the social media used by users to provide opinions as a form of public response to something. Industries often use it to collect information on Twitter, one of which is the aviation industry. This study analyzes public sentiment towards airlines using positive, negative and neutral sentiments. This research uses Naïve Bayes algorithm and Decision Tree. In testing the algorithm, it also uses a grid search to find the best parameters so that it can improve the performance of the Decision Tree algorithm. From the testing process, the accuracy of the Naïve Bayes algorithm is 87.0%, the Decision Tree Grid Search algorithm is 95.5%.

The parameters used by the Decision Tree algorithm using a Grid Search are 'splitter', 'max_features', 'criterion', 'max_dept', with a value of 'max_depth' 10–100 and the determination of the best parameter for the splitter based on 'best' and 'random'. 'max_features' based on 'auto', 'sqrt', 'log2', 'criterion' determination based on 'gini', 'entropy'. So it can be concluded that the Decision Tree algorithm with Grid Search is a better algorithm in this study because it gets better performance results.

Keywords: *Airline, Naïve Bayes, Decision Tree, grid search, Classification, Social Media, Twitter, Sentiment*