

INTISARI

ANALISIS SENTIMEN KOMENTAR YOUTUBE MENGGUNAKAN ALGORITME NAÏVE BAYES, DECISION TREE, DAN K-NEAREST NEIGHBOR

Oleh

Achmad Riyadi

14/364212/PA/15946

Analisis sentimen digunakan untuk menentukan sikap dari seseorang terhadap suatu produk, topik, ulasan, komentar pada sosial media. Untuk toko daring atau pasar daring, ulasan dari pengguna sebelumnya dapat memainkan peran penting dalam proses pengambilan keputusan pelanggan, yang menunjukkan bahwa prediksi sentimen yang akurat dari ulasan tertentu dapat meningkatkan potensi keuntungan.

Metode klasifikasi yang digunakan pada penelitian ini adalah algoritme *naïve bayes*, *k-nearest neighbors*, dan *decision tree*. *Naive Bayes* digunakan untuk menghitung nilai kemungkinan dari kelas - kelas yang ada dan ini merupakan isu besar tentang bagaimana mengklasifikasikan data mentah secara rasional untuk meminimalisir suatu resiko. *K-Nearest Neighbors* (KNN) merupakan algoritma pembelajaran *supervised* dimana hasil klasifikasi dari data baru yang masuk tergantung dengan nilai paling banyak dari para tetangganya. *Decision Tree* memiliki keuntungan antara lain mudah dan cepat, menggunakan memori yang sedikit, bisa menangani data yang memiliki *noisy*, memiliki akurasi yang tinggi, efektif terhadap data yang besar.

Hasil dari penelitian ini adalah algoritme yang memiliki nilai akurasi paling tinggi adalah *decision tree* dengan nilai 98.66%. Algoritme yang memiliki nilai presisi paling tinggi adalah *decision tree* dengan nilai 98.91% dimana berarti *decision tree* memiliki tingkat *false positive* yang rendah. Algoritme yang memiliki nilai *recall* paling tinggi adalah *naïve bayes* dengan nilai 99.36% dimana berarti *naïve bayes* memiliki tingkat *false negative* yang rendah. Algoritme *decision tree* lebih unggul diantara algoritme *naïve bayes* dan algoritme *k-nearest neighbors* dalam melakukan analisis sentimen.

Kata Kunci: Analisis Sentimen, Algoritme *Naïve Bayes*, Algoritme *K-Nearest Neighbors*, Algoritme *Decision Tree*.

ABSTRACT

SENTIMENT ANALYSIS OF YOUTUBE COMMENTS USING NAÏVE BAYES, DECISION TREE, AND K-NEAREST NEIGHBOR ALGORITHM

By

Achmad Riyadi

14/364212/PA/15946

Sentiment analysis is used to determine a person's attitude towards a product, topic, review, comment on social media. For online stores or online market, former reviews can play an important role in customer's decision making process, which indicates that an accurate prediction of sentiment from a certain review can increase potential profit.

The classification method used in this study is the naïve bayes algorithm, k-nearest neighbor, and decision tree. Naïve Bayes is used to calculate the probability value of the existing classes and this is a big issue about how to classify raw data rationally to minimize a risk. K-Nearest Neighbors (KNN) is a supervised learning algorithm where the classification results of the new incoming data depend on the most values of its neighbors. Decision Tree have advantages such as being easy and fast, using little memory, being able to handle noisy data, having high accuracy, being effective with large data.

The result of this research is the algorithm that has the highest accuracy value is the decision tree with a value of 98.66%. The algorithm that has the highest precision value is the decision tree with a value of 98.91% which means that the decision tree has a low false positive rate. The algorithm that has the highest recall value is Naive Bayes with a value of 99.36%, which means that Naive Bayes has a low false negative rate. The decision tree algorithm is superior to the naïve Bayes algorithm and the k-nearest neighbors algorithm in performing sentiment analysis.

Keywords: Sentiment Analysis, Naïve Bayes Algorithm, K-Nearest Neighbors Algorithm, Decision Tree Algorithm