

INTISARI

Implementasi dan perbandingan metode prapemprosesan pada analisis sentimen gubernur dki jakarta menggunakan metode support vector machine dan naive bayes

Oleh

Bagus Rianto

11/313590/PA/13710

Analisis sentimen merupakan ilmu lintas disiplin yang melingkupi Natural Language Processing(NLP), kecerdasan buatan, dan text mining. Penggunaan analisis sentimen biasa digunakan dalam melihat sentimen publik terhadap produk, tokoh terkenal, ataupun kejadian.

Pendekatan analisis sentimen sering dilakukan dengan machine learning dengan menggunakan klasifikasi yang muktahir seperti Support Vector Machine (SVM). Selain menggunakan SVM pendekatan lainnya juga banyak dilakukan, salah satunya Naive Bayes. Pendekatan - pendekatan tersebut sangat sering ditemukan namun sangat sedikit yang dilakukan dengan menggunakan bahasa Indonesia.

Penelitian ini menggunakan sumber data dari detik.com dan Kompas.com untuk menjadi pembelajaran model klasifikasi sentimen analisis. Dengan variasi-variasi proses prapemprosesan didapatkan model yang terbaik yaitu dengan menggunakan SVM dengan proses prapemprosesan cleansing, casefolding dan eliminasi KBBI. Dengan performa *precision* sebesar 65.61%, *recall* sebesar 65.36% dan F-measure 65.06%.

Kata-kata kunci : analisis sentimen, teks mining, opinion mining

ABSTRACT

Implementation and comparison of preprocessing methods on sentimen analysis of governor of jakarta using support vector machine and naive bayes

By

Bagus Rianto

11/313590/PA/13710

Sentiment analysis is an interdisciplinary science that surrounds Natural Language Processing (NLP), artificial intelligence, and text mining. The use of sentiment analysis commonly used in observing sentiment on products, famous figures, or events.

Sentiment analysis approach is often done by machine learning using the modern classification such as Support Vector Machine (SVM). In addition there are many other approaches, such as Naive Bayes. Those approaches are very common but just few have done using Indonesian language.

This research uses data from the online news portal detik.com and kompas.com to be training data to train against sentiment analysis model. Proprocess was customized by the variation of preprocessing function to find the best preprocessing step. The result is SVM with preprocessing function consist of cleansing, casefolding and elimination of KBBI. Measured performance for precision is 65.61%, recall is 65.36% and F-measure is 65.06%. Keywords : analysis sentiment, text mining, opinion min-

ing