

**Twitter Sentiment Analysis mengenai Universitas Dian Nuswantoro
Semarang menggunakan
Algoritma K-Nearest Neighbor**

DANNY OKA RATMANA

(Pembimbing : Aisyatul Karima, S.Kom, MCS)

Teknik Informatika - S1, FIK, Universitas Dian Nuswantoro

www.dinus.ac.id

Email : 111201207151@mhs.dinus.ac.id

ABSTRAK

Twitter merupakan salah satu jejaring sosial yang cukup populer di kalangan masyarakat sekarang ini. Banyak dari mereka yang cukup aktif di media sosial ini untuk sekedar menuliskan kicauan pribadinya maupun mendapatkan beragam informasi dari pengguna twitter lainnya. Pemberitaan mengenai Universitas Dian Nuswantoro Semarang (UDINUS) dari berbagai media pun juga kerap dituliskan dalam akun twitter milik mereka. Pada penelitian kali ini dilakukan Sentiment Analysis untuk mengetahui bagaimana opini masyarakat terhadap UDINUS. Setelah dilakukan implementasi Sentiment Analysis tentang Universitas Dian Nuswantoro Semarang dengan data yang diperoleh dari Twitter sejumlah 1230 data, berdasarkan intensitas term frequency dapat dilihat bahwa kemunculan kata (term) "Semarang" menjadi peringkat pertama pada topik terkait Universitas Dian Nuswantoro. Muncul pula beberapa term lain seperti "mahasiswi", "foto", "bugil", "juara", "mahasiswa", "wisuda", "kampus" dan lain sebagainya. Terbukti Algoritma K-Nearest Neighbor mampu mengklasifikasikan dengan baik data tweets yang telah diperoleh kedalam dua kelas yaitu positif dan negatif. Hal ini ditunjukkan dari tingginya nilai evaluasi yang diperoleh dari perhitungan Accuracy, Recall, Precision serta F-measure. Dimana nilai Accuracy tertinggi diperoleh pada percobaan k=5 dengan metode pelabelan menggunakan cluster yakni sebesar 98,85%. Kemudian untuk perhitungan Precision pada kelas negatif relatif rendah namun untuk kelas positif relatif tinggi. Untuk nilai Recall perhitungan antara kelas positif dan negatif sama tingginya, dengan nilai terendah adalah 86,36% pada kelas negatif menggunakan metode pelabelan manual, kemudian nilai perhitungan dari F-Measure tertinggi didapat pada perhitungan k=5 pada kelas negatif dengan metode pelabelan cluster yakni sebesar 99,34%.

Kata Kunci : Twitter Sentiment Analysis, Information Retrieval, K-Nearest Neighbor

Twitter Sentiment Analysis of Dian Nuswantoro University Semarang using K-Nearest Neighbor Algorithm

DANNY OKA RATMANA

(Lecturer : Aisyatul Karima, S.Kom, MCS)

*Bachelor of Informatics Engineering - S1, Faculty of Computer
Science, DINUS University*

www.dinus.ac.id

Email : 111201207151@mhs.dinus.ac.id

ABSTRACT

Twitter is a social network that is quite popular among the people today. Many of them are quite active in this social media for simply write personal tweets or get an information from other twitter users. News about Dian Nuswantoro University Semarang (UDINUS) from various media were also often written in their own twitter accounts. In this research Sentiment Analysis conducted to determine how public opinion against UDINUS. After Sentiment Analysis of Dian Nuswantoro University Semarang with 1230 of Twitter data implemented, according the intensity of term frequency can be seen that occurrence of the word (term) "Semarang" ranked first on topics related Dian Nuswantoro University. There were also some other term such as "mahasiswa", "foto", "bugil", "juara", "mahasiswa", "wisuda", "kampus" and so forth. Proven K-Nearest Neighbor algorithm is able to classify the tweets well into two classes: positive and negative. It can be seen from the high evaluation value that obtained from the calculation of Accuracy, Recall, Precision and F-measure. Where highest Accuracy value obtained in experiments k=5 that using cluster labeling method which is equal 98,85%. Then for Precision calculation on negative class relatively low but for a positive class is relatively high. For Recall calculation between positive and negative classes have same grade level with the lowest score is 86,36% on negative class that using manual labeling method, then the highest values of F-Measure calculation obtained on k=5 for negative class that using cluster labeling method which is equal 99,34%.

Keyword : Twitter Sentiment Analysis, Information Retrieval, K-Nearest Neighbor