

INTISARI

PERBANDINGAN KINERJA *NAÏVE BAYES CLASSIFIER* TERHADAP PENGGUNAAN *SOCIAL CONTEXT FEATURES* DALAM KLASIFIKASI INFORMASI PALSU DI MEDIA SOSIAL

Oleh :

Dino Dwi Jayanto

18/433772/PPA/05587

Keberadaan informasi palsu berpotensi memberikan dampak yang berbahaya karena masyarakat dapat terpengaruh informasi palsu dan mungkin mengambil tindakan atas informasi yang sesungguhnya tidak kredibel. Terlebih lagi dengan adanya *internet* dan media sosial penyebaran informasi palsu menjadi sangat mudah dan cepat. Konten informasi palsu yang beredar memang sengaja ditulis sedemikian rupa untuk menyesatkan pembacanya, sehingga klasifikasi informasi palsu yang hanya memperhatikan konten menjadi tidak terlalu dapat diandalkan.

Penyebaran informasi pada ekosistem media sosial melibatkan dimensi konten, dimensi sosial, dan dimensi temporal. Keberadaan dimensi sosial dan dimensi temporal sebagai *social context features* dipandang mampu meningkatkan kinerja klasifikasi informasi palsu pada ekosistem media sosial. Hal ini perlu dibuktikan dengan melakukan pengujian terhadap kinerja beberapa model klasifikasi, baik untuk model yang hanya memperhatikan fitur konten, maupun model yang menggunakan fitur konten beserta *social context features*.

Berdasarkan hasil pengujian dengan *K-Fold Cross Validation* diperoleh rata-rata akurasi *Naïve Bayes* sebesar 79,09% untuk klasifikasi yang hanya menggunakan dimensi konten, dan diperoleh rata-rata akurasi *Naïve Bayes Classifier* sebesar 91,07% untuk klasifikasi yang melibatkan dimensi konten, dimensi sosial dan temporal. Hal ini menunjukkan bahwa classifier yang melibatkan dimensi konten, sosial, dan temporal sebagai *social context features* memberikan kinerja klasifikasi yang lebih baik dibanding classifier yang hanya melibatkan dimensi konten saja. Pelibatan *social context features* memberikan peningkatan kinerja sebesar 11,98 %.

Kata Kunci: Klasifikasi Informasi Palsu, *Social Context Features*, Naive Bayes

ABSTRACT

THE COMPARISON OF NAÏVE BAYES CLASSIFIER PERFORMANCE ON THE USE OF SOCIAL CONTEXT FEATURES IN FAKE INFORMATION CLASSIFICATION ON SOCIAL MEDIA

By :

Dino Dwi Jayanto

18/433772/PPA/05587

The existence of false information has the potential to have a dangerous impact because the public can be affected by false information and may take action on information that is actually not credible. Moreover, with the internet and social media spreading false information is very easy and fast. The false information content in circulation is intentionally written in such a way as to mislead readers, so that the classification of false information that only pays attention to the content is not very reliable.

Dissemination of information on the social media ecosystem involves content dimensions, social dimensions, and temporal dimensions. The existence of social dimensions and temporal dimensions as social context features is seen to be able to improve the performance of the classification of false information in the social media ecosystem. This needs to be proven by testing the performance of several classification models, both for models that only pay attention to content features, and models that use content features along with social context features.

Based on the test results using K-Fold Cross Validaiton, the accuracy mean of Naïve Bayes Classifier is 79,09% for classifications that only use the content dimension, and the accuracy of the Naïve Bayes Classifier is 91,07% for classifications involving the content, social and temporal dimensions. This results shows that the classifier that involves the content, social, and temporal dimensions as social context features provides better classification performance than the classifier which only involves the content dimension. Involvement of social context features provides an increase in performance by 11,98%.

Keywords: Fake News Classification, *Social Context Features*, Naive Bayes