

## INTISARI

### **PENGARUH PANJANG KALIMAT PADA ANALISIS SENTIMEN MENGUNAKAN ALGORITMA *SUPPORT VECTOR MACHINE* DAN *CONVOLUTIONAL NEURAL NETWORK***

Oleh

Agung Pambudi  
17/418614/PPA/05398

Penelitian terkait analisis sentimen dalam beberapa tahun terakhir sudah banyak dilakukan, tetapi jarang yang memperhatikan pengaruh panjang kalimat dari dataset terhadap performansi metode yang digunakan. Oleh karena itu, penelitian ini mencoba melakukan analisis sentimen terkait pengaruh panjang kalimat pada komentar mahasiswa training ICT terhadap algoritma *Support Vector Machine* (SVM) dan Convolutional Neural Network (CNN). SVM merupakan metode dalam *machine learning* yang sering digunakan untuk menyelesaikan permasalahan analisis sentimen karena memiliki kelebihan nilai akurasi serta toleransi yang tinggi terhadap atribut yang tidak relevan. Di samping itu, terdapat metode *deep learning* yang juga dapat digunakan untuk melakukan klasifikasi teks, salah satunya yaitu CNN. Dalam analisis sentimen, metode CNN yang memiliki lapisan konvolusi dapat digunakan untuk mengekstrak informasi dengan potongan-potongan teks yang lebih besar. Metode CNN juga membutuhkan lebih sedikit koneksi dan parameter-parameter sehingga lebih mudah untuk dilakukan pelatihan.

Dari hasil pengujian yang dilakukan diperoleh kesimpulan bahwa panjang kalimat pada dataset akan mempengaruhi performansi metode SVM dan CNN jika dikombinasikan dengan model Word2vec. Sedangkan untuk metode SVM+TFIDF, performansi tidak begitu dipengaruhi oleh panjang kalimat, namun metode ini memiliki waktu proses yang paling cepat jika dibandingkan metode lainnya. Adapun metode CNN yang dikombinasikan dengan ekstraksi fitur Word2vec dan hyperparameter dengan ukuran filter sebesar (3,4,5), dropout sebesar 0,5, ukuran batch sebesar 50, epoch sebanyak 20, dan ukuran dimensi vektor representasi kata sebesar 300, menghasilkan rata-rata performansi terbaik untuk seluruh kategori dataset dalam penelitian ini. Nilai performansi tersebut diantaranya akurasi sebesar 93,23%, presisi sebesar 94,63%, recall sebesar 95,93%, dan f1-score sebesar 95,27% dan waktu proses 15,28 detik.

**Kata Kunci:** Analisis Sentimen, SVM, CNN

## **ABSTRACT**

### **EFFECT OF SENTENCE LENGTH IN SENTIMENT ANALYSIS USING SUPPORT VECTOR MACHINE AND CONVOLUTIONAL NEURAL NETWORK ALGORITHM**

By

Agung Pambudi  
17/418614/PPA/05398

Research related to sentiment analysis in recent years has been widely done. However, researchers rarely pay attention to the effect of sentence length from the dataset on the performance of the method used. Therefore, this study tries to analyze the sentiment related to the effect of sentence length using ICT training student comments on the Support Vector Machine (SVM) algorithm and Convolutional Neural Network (CNN). SVM is an algorithm in machine learning that is so familiar to solve sentiment analysis problems because it can obtain high accuracy and tolerance for irrelevant attributes. On the other hand, there are deep learning methods that can also be used to classify text, which is CNN. In sentiment analysis, the CNN algorithm with a convolutional layer can extract information with larger pieces of text. The CNN method also requires fewer connections and parameters that will make it easier to train.

The test results concluded that the length of sentences on the dataset would affect the SVM and CNN methods' performance when combined with the Word2vec model. As for the SVM+TFIDF method, performance is not so influenced by sentence length, but this method has the fastest process time than other methods. The CNN method combined with word2vec and hyperparameter feature extraction with filter size (3,4,5), dropout of 0.5, batch size of 50, the epoch of 20, and the word representation vector dimension of 300, provide the best performance average for all dataset categories in this study. Those performance values include accuracy of 93.23%, precision of 94.63%, recall of 95.93%, f1-score of 95.27%, and processing time of 15.28 seconds.

**Keywords: Sentiment analysis, SVM, CNN**