

Pendapat publik melalui media sosial dapat digunakan untuk meningkatkan kualitas aplikasi *telemedicine*. Analisis sentimen membantu mengklasifikasikan pendapat pengguna aplikasi tersebut. Penelitian sebelumnya menunjukkan bahwa penambahan kata pada kamus sentimen dan penggabungan metode kamus dengan *machine learning* dapat meningkatkan performa analisis sentimen. Namun, sebuah kamus sentimen saja belum cukup akurat dalam menentukan kelas sentimen kata. Oleh karena itu, penelitian ini menggabungkan kamus InSet (*Indonesia Sentiment Lexicon*) dan kamus Sentiwords pada metode kamus sentimen gabungan dan *Support Vector Machine* (SVM) untuk meningkatkan performa analisis sentimen *telemedicine*. Aplikasi *telemedicine* yang dianalisis adalah aplikasi Halodoc dan Alodokter.

Data yang sudah dilabeli oleh kamus sentimen dan TF-IDF menjadi data latih dan data uji untuk SVM. Evaluasi dilakukan dengan prosedur *repeated k-fold cross validation* dengan 10-fold dan pengulangan tiga kali. *Classification report* kemudian dilakukan pada tiap *fold* dan menghasilkan nilai *precision*, *recall*, *f1-score* dan akurasi. Perbandingan antara metode yang diusulkan dan metode yang menggunakan satu kamus sentimen dilakukan menggunakan uji t dengan nilai  $\alpha=0,05$ . Karena total pengujian sebanyak tiga kali maka dilakukan *Bonferroni Correction* sehingga nilai *alpha* setelah dikoreksi menjadi 0,016.

Hasil perbandingan menunjukkan bahwa kamus Sentiwords memiliki performa terendah di semua skenario evaluasi, sementara kamus InSet mendominasi dengan performa terbaik. Pada dokumen Halodoc kamus InSet memiliki nilai *precision* 85,92%, *recall* 86,24%, *f1-score* 85,76% dan akurasi 86,24%. Pada dokumen Alodokter, kamus InSet lebih unggul dalam nilai *recall* dan akurasi sebesar 84,28% dan kamus gabungan lebih unggul dalam nilai *precision* sebesar 83,90% dan *f1-score* 83,70%. Uji t menunjukkan bahwa performa metode kamus gabungan lebih unggul secara signifikan hanya dibandingkan dengan kamus Sentiwords, sedangkan dibandingkan dengan kamus InSet, performa kamus gabungan justru lebih rendah walaupun tidak secara signifikan. Hal ini disebabkan oleh penggabungan kamus yang belum mempertimbangkan bobot dari kata sinonim, sehingga terjadi inkonsistensi antara bobot suatu kata dengan bobot kata sinonimnya.

**Kata Kunci:** Analisis Sentimen, *Lexicon*, *Support Vector Machine*, *Indonesia Sentiment Lexicon*, SentiStrength\_id

## ABSTRACT

Public opinion through social media can enhance the quality of telemedicine applications. Sentiment analysis aids in classifying user opinions of these applications. Previous research has shown that adding lexicons and combining lexicon methods with machine learning can improve sentiment analysis performance. However, relying solely on a sentiment lexicon may not accurately determine the sentiment class of words. Therefore, this study combines the InSet (Indonesia Sentiment Lexicon) and Sentiwords in the hybrid lexicon-Support Vector Machine (SVM) method to enhance sentiment analysis performance in telemedicine. The analyzed telemedicine applications are Halodoc and Alodokter.

The labeled data, using sentiment lexicons and TF-IDF, are used as training and testing data for SVM. Evaluation is conducted using the repeated k-fold cross-validation procedure with 10 folds and three repetitions. Classification reports are generated for each fold, providing precision, recall, F1-score, and accuracy values. A comparison between the proposed method and the method using a single sentiment lexicon is performed using a t-test with  $\alpha=0.05$ . Due to three repetitions, Bonferroni correction is applied, resulting in a corrected alpha value of 0.016.

The comparison results indicate that Sentiwords performs the poorest in all evaluation scenarios, while InSet dominates with the best performance. For the Halodoc document, InSet achieves a precision of 85.92%, recall of 86.24%, F1-score of 85.76%, and accuracy of 86.24%. For the Alodokter document, InSet outperforms in recall and accuracy with values of 84.28%, while the hybrid lexicon performs better in precision with 83.90% and an F1-score of 83.70%. The t-test indicates that the performance of the hybrid lexicon method is significantly superior only compared to the Sentiwords, while compared to the InSet, the performance of the hybrid lexicon is lower, although not significantly. This is due to the merging of lexicons that did not consider the weights of synonymous words, resulting in inconsistency between the weight of a word and the weight of its synonymous words.

**Keywords:** Sentiment Analysis, Lexicon, Support Vector Machine, Indonesia Sentiment Lexicon, SentiStrength\_id