ABSTRAK

ANALISIS PERBANDINGAN SENTIMEN PENGGUNA *TWITTER* TERHADAP LAYANAN *INDIHOME* MENGGUNAKAN METODE KLASIFIKASI

Oleh

Della Puspitasari

9882405219121009

Internet sangat dibutuhkan untuk kehidupan sehari-hari, dimana di indonesia sudah banyak penyedia layanan internet salah satunya *indihome*. Analisis sentimen sendiri bertujuan untuk mengelompokan suatu teks kedalam kelas Negative, Positive dan Netral. Pada platform *twitter* banyak sekali ulasan mengenai *provider* internet salah satunya *indihome*, karena adanya pelayanan yang kurang baik atau untuk sekedar mengapresiasi pelayanan yang diberikan. Berdasarkan perhitungan hasil yang didapatkan 71,1% negative, 21,1% positive dan 7,7% netral. Data yang diperoleh tidak seimbang maka dari itu proses pengklasifikasian di bantu menggunakan *Smote*. Hasil perbandingan ke empat metode yang digunakan ialah *Support Vector Machine* dengan tingkat *accuracy* 89% tingkat AUC 89% jika menggunakan *smote* mendapatkan *accuracy* 93% dan tingkat AUC 97% dengan data *training* 80% dan *testing* 20%, *Naïve Bayes* dengan tingkat accuracy 86% AUC 95% jika menggunakan *smote* mendapatkan accuracy 89% AUC 89% dengan data *training* 80% dan *testing* 20% , *Random forest* dengan tingkat accuracy 82% AUC 91% jika menggunakan *smote* mendapatkan accuracy 85% AUC 91% dengan data *training* 70% dan *testing* 30%, *Decision tree* dengan tingkat accuracy 78% AUC 70% dengan data *training* 50% dan *testing* 50% jika menggunakan *smote* mendapatkan accuracy 81% AUC 71%. Dari keseluruhan perbandingan *accuracy* paling tinggi tanpa *smote* ataupun menggunakan *smote* ialah *Support Vector Machine*.

Kata kunci : *Indihome, Smote, Support Vector Machine, Naïve Bayes, Random forest, Decision tree*

ABSTRACT

COMPARATIVE ANALYSIS OF TWITTER USER SENTIMENT TOWARDS INDIHOME SERVICES USING CLASSIFICATION METHODS

Oleh

Della Puspitasari

9882405219121009

The internet is highly needed for daily life, where in Indonesia there are already many internet service providers, one of which is Indihome. Sentiment analysis itself aims to classify a text into Negative, Positive, and Neutral classes. On the Twitter platform, there are numerous reviews about internet service providers, including Indihome, either due to unsatisfactory service or simply to appreciate the services provided. Based on the calculation results, 71.1% are negative, 21.1% are positive, and 7.7% are neutral. Due to the imbalanced data, the classification process is assisted using SMOTE. The comparison results of the four methods used are as follows: Support Vector Machine: Accuracy 89%, AUC 89%. When using SMOTE, it achieves an accuracy of 93% and AUC of 97% with 80% data for training and 20% for testing. Naïve Bayes: Accuracy 86%, AUC 95%. With SMOTE, it achieves an accuracy of 89% and AUC of 95% with 80% data for training and 20% for testing. Random Forest: Accuracy 82%, AUC 91%. With SMOTE, it achieves an accuracy of 85% and AUC of 91% with 70% data for training and 30% for testing. Decision Tree: Accuracy 78%, AUC 70%. With SMOTE, it achieves an accuracy of 79% and AUC of 71% with 50% data for training and 50% for testing . Overall, the highest accuracy in both cases, with or without SMOTE, is achieved by the Support Vector Machine.

Keyword : *Indihome*, Smote, SVM, Naïve Bayes, *Random forest*, *Decision tree*