

**KLASIFIKASI SENTIMEN DATA TIDAK SEIMBANG MENGGUNAKAN  
ALGORITMA SMOTE & *K-NEAREST NEIGHBOR* PADA ULASAN PENGGUNA  
APLIKASI MAXIM**

**Alessandro Samuel Tamada**

**Abstrak**

Maxim adalah aplikasi ojek *online* yang dipakai banyak masyarakat Indonesia, sebagai pendatang baru namun mampu bersaing dengan ojek *online* lainnya yang terlebih dahulu, dengan memberi banyak diskon dan harga yang lebih murah, pada awalnya banyak kontroversi mengenai aplikasi ini karena pelayanannya yang kurang baik, namun dari penelitian ini didapatkan bahwa banyak penilaian yang baik dari pada yang buruk. Penelitian ini mengambil data dari ulasan aplikasi Maxim dan diolah sehingga dapat diklasifikasi dengan algoritma KNN. Namun dari data yang didapatkan juga terdapat selisih yang cukup jauh antara data berlabel negatif dan data berlabel positif, karena itu harus menggunakan algoritma SMOTE agar mendapatkan label data yang seimbang. Hasil dari penelitian ini menggunakan data yang tidak seimbang dengan data positif sebesar 842 dan data negatif 158, didapatkan nilai tertinggi pada nilai  $K = 2$ , yaitu akurasi sebesar 0.875, presisi 0.877, *f1-score* 0.932, spesifisitas 0.111, dan sensitivitas 0.994. Sedangkan penelitian kedua dilakukan *oversampling* dengan teknik SMOTE agar data dapat seimbang yang kemudian data positif menjadi 842 dan data negatif juga 842, hasilnya tertinggi didapatkan pada nilai  $K = 1$  dan 2 dengan nilai yang sama yaitu akurasi 0.985, presisi 1, *f1-score* 0.986, spesifisitas 1 dan sensitivitas 0.973. Penelitian dengan data yang sudah melalui proses SMOTE ini mendapatkan hasil yang sangat baik dapat dilihat dari nilai spesifisitas dan sensitivitas yang memiliki selisih tipis, menunjukkan bahwa data penelitian yang digunakan seimbang. Sehingga dapat disimpulkan bahwa data yang baik untuk penelitian adalah data yang seimbang.

**Kata kunci :** Maxim, KNN, SMOTE.

**UNBALANCED DATA SENTIMENT CLASSIFICATION USING SMOTE & K-  
NEAREST NEIGHBOR ALGORITHM ON USER REVIEWS OF THE MAXIM  
APPLICATION**

**Alessandro Samuel Tamada**

**Abstract**

*Maxim is an online motorbike taxi application that is used by many Indonesian people, as a newcomer but able to compete with other online motorbike taxis that were earlier, by providing lots of discounts and cheaper prices, at first there was a lot of controversy regarding this application because the service was not good, however From this research it was found that there were more good assessments than bad. This research takes data from reviews of the Maxim application and processes it so that it can be classified using the KNN algorithm. However, from the data obtained there is also quite a large difference between data labeled as negative and data labeled as positive, therefore you have to use the SMOTE algorithm to get balanced data labels. The results of this study used unbalanced data with positive data of 842 and negative data of 158. The highest value was obtained at  $K = 2$ , namely accuracy of 0.875, precision of 0.877, f1-score of 0.932, specificity of 0.111, and sensitivity of 0.994. while the second research carried out oversampling using the SMOTE technique so that the data could be balanced, then positive data became 842 and negative data also 842, the highest results were obtained at values  $K = 1$  and 2 with the same values, namely accuracy 0.985, precision 1, f1-score 0.986, specificity 1 and sensitivity 0.973. Research with data that has gone through the SMOTE process has obtained very good results which can be seen from the specificity and sensitivity values which have a slight difference, indicating that the research data used is balanced. So it can be concluded that good data for research is balanced data.*

**Keywords :** Maxim, KNN, SMOTE.