

ABSTRAK

Nama : Lintang Annurya Dewa
NIM : 41519010054
Program Studi : Teknik Informatika
Judul Laporan : Implementasi Metode *Random Over-Under Sampling* Pada Algoritma *Random Forest Classifier* Dalam Proses Analisis Sentimen Pada Website Steam (Studi Kasus: Game Cyberpunk 2077)
Pembimbing : Siti Maesaroh, S.Kom., M.T.I.

Penelitian ini membahas tentang implementasi metode Random Over-Under Sampling pada algoritma Random Forest Classifier dalam proses analisis sentimen pada website Steam, dengan studi kasus pada game Cyberpunk 2077. Game ini mengalami masalah signifikan setelah perilisannya, terutama di konsol PS4 dan Xbox One, yang menyebabkan banyak pengguna memberikan ulasan negatif terkait dengan masalah teknis. Dataset yang digunakan diperoleh melalui proses scraping dari website Steam, berisi 106101 data ulasan sentiment positif dan negatif tentang game Cyberpunk 2077. Penelitian ini menunjukkan bahwa implementasi metode Random Over-Under Sampling pada algoritma Random Forest Classifier memberikan peningkatan kinerja dalam analisis sentimen pada website Steam, khususnya dalam konteks game Cyberpunk 2077. Pengujian dilakukan dengan dua metode rasio, yaitu 80:20 dan 70:30, dan keduanya menghasilkan hasil yang serupa dalam memprediksi sentimen ulasan game. Kedua model tersebut memiliki tingkat akurasi sekitar 87,8% dengan tingkat kesalahan klasifikasi sekitar 12,1% hingga 12,2%. Skor F1 dan ROC-AUC juga menunjukkan hasil yang serupa, dengan nilai sekitar 0,74 hingga 0,70. Namun, implementasi metode Over-Under Sampling memberikan peningkatan kinerja yang signifikan pada model. Tingkat akurasi meningkat menjadi 93,560%, dengan tingkat kesalahan klasifikasi yang lebih rendah, yaitu 6,439%. Selain itu, skor F1 yang tinggi (0,89857) dan ROC-AUC yang tinggi (0,93334) menunjukkan bahwa model mampu dengan baik dalam mengklasifikasikan sentimen ulasan game Cyberpunk 2077.

Kata kunci:

Analisis Sentimen, TF-IDF, Random Forest, Oversampling, Undersampling

ABSTRACT

Name : Lintang Annurya Dewa
NIM : 41519010054
Study Program : Teknik Informatika
Title Thesis : Implementation of Random Over-Under Sampling Method on Random Forest Classifier Algorithm for Sentiment Analysis on Steam Website (Case Study: Cyberpunk 2077 Game)
Consellor : Siti Maesaroh, S.Kom., M.T.I.

This research discusses the implementation of the Random Over-Under Sampling method on the Random Forest Classifier algorithm in the sentiment analysis process on the Steam website, with a case study on the game Cyberpunk 2077. This game experienced significant issues after its release, especially on the PS4 and Xbox One consoles, which resulted in many users giving negative reviews related to technical problems. The dataset used was obtained through scraping from the Steam website, containing 106101 data of positive and negative sentiment reviews about the game Cyberpunk 2077. This study shows that the implementation of the Random Over-Under Sampling method on the Random Forest Classifier algorithm improves the performance of sentiment analysis on the Steam website, particularly in the context of the game Cyberpunk 2077. Testing was conducted using two ratio methods, namely 80:20 and 70:30, and both produced similar results in predicting the sentiment of game reviews. Both models achieved an accuracy rate of around 87.8% with a classification error rate of approximately 12.1% to 12.2%. The F1 score and ROC-AUC also showed similar results, with values ranging from 0.74 to 0.70. However, the implementation of the Over-Under Sampling method provided a significant performance improvement to the model. The accuracy rate increased to 93.560%, with a lower classification error rate of 6.439%. Additionally, the high F1 score (0.89857) and high ROC-AUC (0.93334) indicate that the model performs well in classifying the sentiment of Cyberpunk 2077 game reviews.

Key words:

Sentiment Analysis, TF-IDF, Random Forest, Oversampling, Undersampling