OVERSAMPLING USING THE SMOTE METHOD AND MACHINE LEARNING RANDOM FOREST TO HANDLING DATA IMBALANCE IN INTRUSION DETECTION SYSTEMS

Name   : Grace Putra Benyamin Habeahan

Id Number   : 6103221520

Supervisor   : Eko Prayitno, M. Kom

ABSTRACT

Network security has become a critical issue in the digital era, as the threat of cyberattacks increases. This study discusses the application of the Random Forest algorithm and the SMOTE technique to address data imbalance in an intrusion detection system using the CICIDS2017 dataset. The process begins with data preprocessing, normalization, and model training. The SMOTE-based model achieved an accuracy of 99.86%, with a precision of 0.9949, a recall of 0.9967, and an f1-score of 0.9958. The SMOTE technique has been shown to improve the model's performance in detecting attacks compared to the non-SMOTE model.

Keywords: Network Security, Intrusion Detection, SMOTE, Random Forest, CICIDS2017