

Penerapan Adaptive Synthetic Sampling Approach dalam Menangani Ketidakseimbangan Kelas pada Dataset Wi-Fi Attacks = Application of Adaptive Synthetic Sampling Approach in Handling Class Imbalance in Wi-Fi Attacks Dataset

Karin Marshanda, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=9999920544966&lokasi=lokal>

Abstrak

Intrusion Detection System (IDS) merupakan sistem untuk mendeteksi serangan dalam jaringan, baik lokal maupun internet. Dalam melakukan deteksi penyalahgunaan atau deteksi anomali, beberapa peneliti telah menggunakan data mining untuk mengidentifikasi berbagai jenis intrusi, termasuk yang jarang terjadi. Namun, data mining rentan terhadap data imbalance (data tidak seimbang) yang dapat mengurangi efektivitas algoritma klasifikasi karena asumsi mayoritas classifier terhadap distribusi yang seimbang. Berdasarkan permasalahan tersebut, maka akan dilakukan penelitian terkait penanganan data imbalance menggunakan metode Adaptive Synthetic Sampling (ADASYN) dengan cara menghasilkan data sintetis pada kelas minoritas agar algoritma klasifikasi dapat bekerja lebih baik. Metode ADASYN efektif bekerja pada variabel prediksi berjumlah 2 kelas (binary class), namun dikarenakan penelitian ini berurusan dengan masalah multiclass, maka akan digunakan pendekatan One-Vs-One (OVO) untuk menyeimbangkan kelas. Keefektifan ADASYN akan dievaluasi melalui implementasinya pada dataset Wi-Fi attacks, yaitu Aegean Wi-Fi Intrusion Dataset (AWID2). Data sebelum dan setelah rebalancing dievaluasi dengan menggunakan metode klasifikasi seperti regresi logistik dan Support Vector Machine (SVM), untuk dibandingkan nilai precision, recall, spesifisitas, serta F1-score dari kedua dataset tersebut. Meskipun ADASYN hanya meningkatkan nilai precision dalam dataset Wi-Fi attacks, dengan menggunakan metode klasifikasi SVM kernel polynomial terbukti efektif dalam mendeteksi kelas serangan, meskipun performa metrik lainnya tidak mencapai tingkat yang sama.

.....An Intrusion Detection System (IDS) is a system designed to detect attacks within networks, both local and internet-based. In the realm of misuse detection or anomaly detection, researchers have utilized data mining to identify various types of intrusions, including those that occur infrequently. However, data mining is susceptible to data imbalance, which can reduce the effectiveness of classification algorithms due to their assumption of balanced distribution. To address this issue, research will focus on handling data imbalance using the Adaptive Synthetic Sampling (ADASYN) method, which generates synthetic data for the minority class to enhance the performance of classification algorithms. ADASYN is effective for predictive variables with binary class scenarios, but since this study deals with multiclass problems, an One-Vs-One (OVO) approach will be employed to balance the classes. The effectiveness of ADASYN will be evaluated by implementing it on the Wi-Fi attacks dataset, specifically the Aegean Wi-Fi Intrusion Dataset (AWID2). Data before and after rebalancing will be evaluated using classification methods such as logistic regression and Support Vector Machine (SVM). Metrics including precision, recall, specificity, and F1-score will be compared between the two datasets. Although ADASYN only improves precision values in the Wi-Fi attacks dataset, using SVM with a polynomial kernel has proven effective in detecting attack classes, although other metric performances did not reach the same level.