ABSTRACT

USING DEEP BELIEF NETWORKS WITH CONDITIONAL TABULAR GENERATIVE ADVERSARIAL NETWORK TO DETECT NETWORK INTRUSION

By

KENNISKIU FORTINO KURNIAWAN

20/457772/PA/19810

The digital network serves as a fundamental element of contemporary society, linking people globally and catering to various requirements. Despite its advantages, there are significant risks posed by security breaches and data leaks. Defenders bear the burden of a precision-driven role, contrasting with attackers who only require one successful breach. Machine learning, particularly Deep Belief Networks (DBN) in Network Intrusion Detection Systems (NIDS), demonstrates effectiveness but grapples with imbalanced datasets. Wheelus et al. (2018) utilized the Synthetic Minority Over-sampling Technique (SMOTE) to address this issue but encountered constraints. An alternative approach, Conditional Tabular Generative Adversarial Network (CTGAN), suggested by Xu et al. (2019), shows promise in overcoming these limitations.

This study emphasizes the crucial necessity of tackling imbalanced datasets to unlock the complete potential of DBN in strengthening network defense against dynamic cybersecurity threats through the application of CTGAN. To prove that CTGAN improves the work of DBN, this research compares both the models of DBN, one that is fed the oversampled dataset, and one that is fed the raw dataset.

It is revealed that implementing CTGAN enhances DBN, resulting in an accuracy of 99 percent, as measured by the Area Under the Curve for Precision-Recall (AUC-PR), whereas the DBN using the original dataset achieved only 95 percent accuracy in detecting the intrusion class. The higher AUC-PR signifies that when the model predicts a positive class, it is more likely to be correct (indicating high precision), and it also captures a substantial portion of the actual positive instances (demonstrating high recall).

ABSTRAK

MENGGUNAKAN JARINGAN KEYAKINAN DALAM (DEEP BELIEF NETWORKS) DENGAN JARINGAN ADVERSARIAL GENERATIF TABULAR KONDISIONAL UNTUK DETEKSI INTRUSI JARINGAN

By
KENNISKIU FORTINO KURNIAWAN
20/457772/PA/19810

Jaringan digital berfungsi sebagai elemen mendasar dalam masyarakat kontemporer, menghubungkan orang secara global dan memenuhi berbagai kebutuhan. Meskipun memiliki keuntungan, ada risiko signifikan yang ditimbulkan oleh pelanggaran keamanan dan kebocoran data. Para pembela membawa beban peran yang didorong oleh presisi, berbeda dengan para penyerang yang hanya memerlukan satu pelanggaran yang berhasil. Pembelajaran mesin, khususnya Jaringan Keyakinan Mendalam (DBN) dalam Sistem Deteksi Intrusi Jaringan (NIDS), menunjukkan efektivitas tetapi menghadapi masalah dengan dataset yang tidak seimbang. Wheelus dkk. (2018) menggunakan Teknik Oversampling Minoritas Sintetis (SMOTE) untuk mengatasi masalah ini tetapi mengalami kendala. Pendekatan alternatif, Jaringan Generatif Adversarial Tabular Bersyarat (CTGAN), yang diusulkan oleh Xu dkk. (2019), menunjukkan potensi dalam mengatasi batasan-batasan ini.

Penelitian ini menekankan kebutuhan mendesak untuk menangani dataset yang tidak seimbang guna membuka potensi penuh DBN dalam memperkuat pertahanan jaringan terhadap ancaman keamanan Siber dinamis melalui aplikasi CTGAN. Untuk membuktikan bahwa CTGAN meningkatkan kinerja DBN, penelitian ini membandingkan kedua model DBN, satu yang diberi dataset yang dioversampling, dan satu yang diberi dataset mentah.

Diketahui bahwa penerapan CTGAN meningkatkan DBN, menghasilkan akurasi 99 persen, seperti diukur oleh Area Under the Curve untuk Presisi-Recall (AUC-PR), sedangkan DBN yang menggunakan dataset asli hanya mencapai akurasi 95 persen dalam mendeteksi kelas

intrusi. AUC-PR yang lebih tinggi menunjukkan bahwa ketika model memprediksi kelas positif, kemungkinan besar prediksi tersebut benar (menunjukkan presisi tinggi), dan juga menangkap sebagian besar instance positif sebenarnya (menunjukkan recall tinggi).