

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

PENERAPAN *SYMMETRICAL UNCERTAINTY* DAN *GAIN RATIO* UNTUK KLASIFIKASI SERANGAN JARINGAN MENGGUNAKAN METODE *MODIFIED K-NEAREST NEIGHBOR*

ARFAN EFLIADI

11251100412

Jurusan Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Sultan Syarif Kasim Riau

ABSTRAK

Dataset NSL-KDD merupakan set data terbaik dan menjadi titik acuan untuk menguji sistem kerja *Intrusion Detection System (IDS)*. Di dalam dataset ini terdapat 41 fitur, akan tetapi tidak semua fitur memberikan pengaruh pada saat melakukan intrusi serangan jaringan dan klasifikasi serangan hingga menyebabkan proses menjadi lebih lambat, oleh karena itu digunakan metode *feature selection symmetrical uncertainty* dan *gain ratio* untuk mengurangi fitur yang tidak relevan tersebut serta metode *Modified K-Nearest Neighbor (MK-NN)* untuk klasifikasi. Fitur yang digunakan pada penelitian ini berjumlah 11 fitur dengan nilai $K = 1, 3, 5, 7, 9, 11$ dengan *output* kelas yaitu *Normal*, *DoS*, *Probe*, *U2R* dan *R2L*. Hasil dari penelitian yang dilakukan diperoleh akurasi tertinggi yaitu 96.0784% dengan $K = 1$ dengan skenario perbandingan data 80%:20%. Jadi dapat disimpulkan bahwa dengan menggunakan kombinasi fitur seleksi *Symmetrical Uncertainty* dan *Gain Ratio* serta MK-NN dapat diterapkan untuk klasifikasi serangan jaringan dan menghasilkan nilai akurasi yang baik.

Kata kunci : data mining, dataset NSL-KDD, feature selection, modified k-nearest neighbor, klasifikasi



APPLICATION SYMMETRICAL UNCERTAINTY AND GAIN RATIO FOR ATTACK CLASIFICATION USE MODIFIED K-NEAREST NEIGHBOR

ARFAN EFLIADI
11251100412

Informatics Engineering
Faculty of Science and Technology
Sultan Syarif Kasim Islamic University of Riau

ABSTRACT

The NSL-KDD dataset is the best data set and is a reference point for the Intrusion Detection System (IDS) work system. In this dataset there are 41 features, but not all features have an influence when performing network attack intrusions and attack classifications to cause the process to be slower, therefore the method is used to feature a selection of symmetrical uncertainty and gain ratio to reduce the irrelevant features and methods. Modified K-Nearest Neighbor (MK-NN) for classification. The features used in this study amounted to 11 features with a K value of 1,3,5,7,9,11 with class outputs namely Normal, DoS, Probe, U2R and R2L. The results of the research conducted obtained the highest accuracy of 96.0784% with $K = 1$ with a data comparison scenario of 80%: 20%. So it can be concluded that using a combination of Symmetrical Uncertainty and Gain Ratio selection features and MK-NN can be applied to classifying network attacks and producing good accuracy values.

Keyword : classification, data mining, feature selection, modified k-nearst neighbor, NSL-KDD dataset.

UIN SUSKA RIAU

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.