UIN SUSKA RIAU
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.
© Hak cipta milik UIN Suska Riau
State Islamic University of Sultan Syarif Kasim Riau
UIN SUSKA RIAU

BAB I
PENDAHULUAN

1.1 Latar Belakang

Teknologi pada masa yang akan datang, keamanan komputer dan jaringan akan menjadi infrastruktur yang sangat penting untuk sebagian besar skenario aplikasi, seperti perdagangan dan pemasaran online, arsip online dan transfer, dimana serangan jaringan cukup banyak. Dengan memeriksa berbagai data yang dikumpulkan dalam lalu lintas jaringan dapat berguna untuk mendeteksi serangan di komputer atau di jaringan dengan menggunakan teknik deteksi intrusi (Mallissery, Kolekar, & Ganiga, 2014).

Dalam beberapa tahun terakhir serangan pada jaringan semakin meningkat dan mengarah ke *Intrusion Detection System* (IDS) yang berubah menjadi komponen vital untuk mengamankan seluruh jaringan. Namun IDS juga memiliki banyak masalah sehubungan dengan perbaruan reguler, kurang mampu mendeteksi serangan yang tidak diketahui, tingkat kekhawatiran tinggi yang tidak dapat disesuaikan serta lebih banyak konsumsi sumber daya jaringan (Mallissery et al., 2014).

Intrusion Detection System (IDS) merupakan proses mekanisme keamanan yang digunakan untuk memonitor perilaku abnormal pada jaringan. IDS melakukan identifikasi dan menginformasikan tentang aktivitas pengguna normal atau tidak (Belavagi & Muniyal, 2016). IDS dapat menanggapi *anomaly* di berbagai kasus dengan mengambil tindakan seperti pemblokiran alamat IP atau pengguna dari mengakses jaringan (Bhattacharjee, 2017).

(Tavallaee, Stakhanova, & Ghorbani, 2010) melakukan penelitian dengan menganalisis beberapa penelitian tentang IDS dari tahun 2000-2008, dari hasil penelitiannya menyimpulkan bahwa 70% penelitian dilakukan menggunakan kumpulan dataset *Network Security Layer – Knowledge Discovery in Database* (NSL-KDD). Kemudian (Mahbod, 2009) juga mengatakan dataset NSL-KDD merupakan kumpulan data benchmark yang efektif sehingga membantu peneliti membandingkan metode-metode deteksi instrusi sebagai langkah awal membangun

UIN SUSKA RIAU
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.
© Hak cipta milik UIN Suska Riau
State Islamic University of Sultan Syarif Kasim Riau
UIN SUSKA RIAU

IDS. Beberapa penelitian telah dilakukan untuk menentukan kategori serangan berbasis klasifikasi.

Klasifikasi adalah proses pengelompokan objek ke dalam kelompok berdasarkan atribut umum, yang mana tugas klasifikasi dapat ditemukan di berbagai bidang seperti medis, keuangan dan pendidikan (Hamid, Ngu, & Alipiah, 2018). Berbagai algoritma klasifikasi diantaranya adalah *Learning Vector Quantization* (LVQ).

LVQ merupakan sebuah metode yang melakukan pembelajaran (*learning*) dengan *competitive layer* (lapisan kompetitif) yang terawasi (*supervised*). Lapisan kompetitif bisa secara otomatis belajar untuk mengklasifikasi *input* dari vektor-vektor. Jarak-jarak antar vektor *input* akan menentukan kelas-kelas yang akan dihasilkan dari lapisan kompetitif tersebut. Lapisan kompetitif akan meletakkan kedua vektor *input* pada kelas yang sama jika kedua vektor *input* tersebut mendekati sama (Kusumadewi, 2003).

Metode LVQ merupakan salah satu metode Jaringan Syaraf Tiruan yang dapat melakukan pengklasifikasian data dengan akurasi yang cukup akurat. Pada penelitian sebelumnya menunjukkan bahwa LVQ2.1 memiliki tingkat akurasi rata-rata 85% (Maruno, 1993), (Byrne, et al., 1994), (Reinhardt, el al., 1996), (Budianita, 2013). Metode LVQ bisa unggul lebih tinggi tingkat akurasi dari metode klasifikasi lainnya seperti K-NN (*K-Nearest Neighbor*) (Byrne, et al., 1994), dan LVQ2.1 lebih unggul dari versi sebelumnya, yaitu pengembangan dari metode LVQ1 (Budianita, 2013) dan memiliki *error rate* yang lebih kecil dibandingkan pendahulunya (Lee, 1993).

(Mukherjee & Sharma, 2012) mengatakan bahwa dataset NSL-KDD memiliki fitur yang tidak relevan dan berlebihan, sehingga mengakibatkan proses pendeteksian yang panjang dan menurunkan kinerja sistem. Dalam melakukan penerapan menggunakan teknik *machine learning* data yang berdimensi tinggi merupakan salah satu masalah, karena pada proses analisis akan memberikan dampak negatif (Guan, Yuan, Jin, & Lee, 2012) dalam (Wahyuni et al., 2016).

Untuk mengatasi data yang berdimensi tinggi tersebut, maka perlu dilakukan pengurangan terhadap fitur yang tidak relevan tersebut sehingga memberikan banyak keuntungan, seperti meningkatkan kinerja dan mengurangi

UIN SUSKA RIAU
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.
© Hak cipta milik UIN Suska Riau
State Islamic University of Sultan Syarif Kasim Riau

kompleksitas analisis data, serta menghindari *overfitting* (Guan et al., 2012) dalam (Wahyuni et al., 2016). Usaha untuk mengurangi fitur pada data yang berdimensi tinggi dapat dilakukan dengan menggunakan *feature selection*.

Feature selection merupakan teknik pemrosesan data untuk mengurangi jumlah fitur dan meningkatkan kinerja klasifikasi. Keunggulan menggunakan metode *feature selection* adalah meminimalkan pengetahuan yang hilang dan menghindari kesulitan dalam menciptakan fitur baru (Wang & Niu, 2016).

Beberapa penelitian mengenai penerapan *feature selection* diantaranya adalah (Dong & Wang, 2009), (Chuang, Yang, & Yang, 2009), (Mallissery et al., 2014) menerapkan *feature selection* yang berbeda-beda. Hasil penelitian mereka dapat disimpulkan bahwa dengan menerapkan *feature selection* dapat meningkatkan kecepatan proses klasifikasi tanpa menurunkan hasil klasifikasi, dan juga penggunaan *feature selection* tidak hanya meningkatkan performa klasifikasi, tetapi menurunkan beban komputasi klasifikasi. Salah satu metode *feature selection* adalah *Symmetrical Uncertainty* dan *Gain Ratio*.

Penelitian yang menerapkan *feature selection Symmetrical Uncertainty* dan *Gain Ratio* dilakukan oleh (Garg, 2014), yaitu menganalisa berbagai algoritma *feature selection* berbasis *filter* yang dievaluasi pada 10 metode klasifikasi. Metode *feature selection* yang digunakan adalah *Information Gain*, *Gain Ratio*, *Relief-F*, *One-R*, *Symmetrical Uncertainty*, *Chi-Square*. Dari hasil penelitian ini, kombinasi metode *feature selection Symmetrical Uncertainty* dan *Gain Ratio* memberikan nilai akurasi klasifikasi yang tinggi.

Berdasarkan pada hasil penelitian sebelumnya, maka penulis ingin melakukan penelitian dengan menerapkan *feature selection Symmetrical Uncertainty* dan *Gain Ratio* untuk klasifikasi dengan metode *Learning Vector Quantization* 2.1 (LVQ2.1) menggunakan dataset NSL-KDD.

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang masalah diatas, adapun rumusan masalah yang diangkat pada penelitian ini adalah bagaimana menerapkan *feature selection Gain Ratio* dan *Symmetrical Uncertainty* serta mengklasifikasikannya menggunakan metode *Learning Vector Quantization* 2.1 (LVQ2.1) pada dataset NSL-KDD.

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.
© Hakcipta milik UIN Suska Riau
State Islamic University of Sultan Syarif Kasim Riau
UIN SUSKA RIAU

1.3 Batasan Masalah

Adapun yang menjadi batasan masalah pada penelitian ini adalah sebagai berikut :

1. Kelas yang digunakan untuk klasifikasi berjumlah 4 yaitu *Denial of Service* (*DoS*), *User to Root* (*U2R*), *Remote to Local* (*R2L*) dan *Probe*.
2. Jumlah atribut yaitu 41. Atribut yang digunakan berdasarkan hasil dari *feature selection*.
3. Jumlah data yang digunakan yaitu 208 data (masing-masing kelas berjumlah 52 data).

1.4 Tujuan Penelitian

Tujuan penulis melakukan penelitian ini adalah untuk mengetahui akurasi yang diperoleh dengan menerapkan kombinasi *feature selection Gain Ratio* dan *Symmetrical Uncertainty* untuk klasifikasi metode *Learning Vector Quantization* 2.1 (LVQ2.1).

1.5 Sistematika Penulisan

Dalam laporan Tugas Akhir ini terdapat sistematika penulisan yang terdiri dari enam BAB, yaitu :

BAB I PENDAHULUAN

BAB I ini, mencakup hal-hal mengenai latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, serta sistematika penulisan laporan.

BAB II LANDASAN TEORI

BAB II menjelaskan landasan dalam penelitian yang berisikan mengenai teori-teori dan penelitian sebelumnya yang digunakan sebagai dasar dalam mengolah data penelitian.

BAB III METODOLOGI PENELITIAN

BAB III menjelaskan tahapan-tahapan yang akan dilaksanakan dalam proses penelitian. Tahapan tersebut meliputi perumusan masalah, studi pustaka, analisa tahapan KDD, perancangan dan implementasi aplikasi, pengujian aplikasi, serta kesimpulan dan saran.

UIN SUSKA RIAU
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.
© Hak cipta milik UIN Suska Riau
State Islamic University of Sultan Syarif Kasim Riau
UIN SUSKA RIAU

BAB IV ANALISA DAN PERANCANGAN

BAB IV menjelaskan tentang aplikasi yang akan dibangun berdasarkan analisa dengan *feature selection* dan metode *Learning Vector Quantization* 2.1, kemudian dilakukan perancangan aplikasi dari hasil analisa tersebut.

BAB V IMPLEMENTASI DAN PENGUJIAN

BAB V menjelaskan tentang penerapan dari hasil perancangan aplikasi yang sudah dirancang sebelumnya, meliputi implementasi *database*, metode yang digunakan serta *interface* aplikasi. Selanjutnya dilakukan pengujian terhadap aplikasi tersebut.

BAB VI PENUTUP

BAB VI menjelaskan tentang hasil penelitian yang telah dilakukan yang meliputi kesimpulan serta saran.