

**SISTEM KLASIFIKASI SERANGAN SQL *INJECTION* &  
XSS PADA RAMA *REPOSITORY* DENGAN METODE  
*SUPPORT VECTOR MACHINE* (SVM)**

**TUGAS AKHIR**

**Diajukan Untuk Melengkapi Salah Satu Syarat  
Memperoleh Gelar Sarjana Komputer**



**OLEH:**

**TIA HERMITA**

**09011281722051**

**JURUSAN SISTEM KOMPUTER  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS SRIWIJAYA**

**2021**

**LEMBAR PENGESAHAN**

**SISTEM KLASIFIKASI SERANGAN SQL INJECTION &  
XSS PADA RAMA REPOSITORY DENGAN METODE  
SUPPORT VECTOR MACHINE (SVM)**

**SKRIPSI**

**Diajukan Untuk Melengkapi Salah Satu Syarat  
Memperoleh Gelar Sarjana Komputer**

**Oleh:**

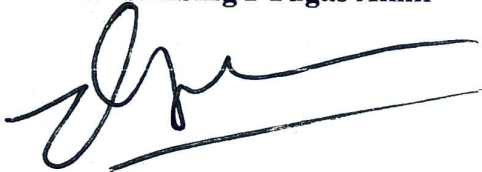
**TIA HERMITA**

**09011281722051**

**Indralaya, 28 Juli 2021**

**Mengetahui,**

**Pembimbing I Tugas Akhir**



**Deris Stiawan, M.T., Ph.D. IPU.**  
**NIP. 197806172006041002**

**Pembimbing II Tugas Akhir**



**Ali Bardadi, S.SI., M.Kom.**  
**NIP. 198806292019031007**

**Ketua Jurusan Sistem Komputer**



**Dr. Ir. H. Sukemi, M.T.**  
**NIP. 196612032006041001**

## HALAMAN PERSETUJUAN

Telah diuji dan lulus pada :

Hari : Kamis

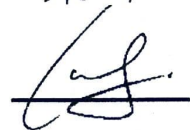
Tanggal : 15 Juli 2021

Tim Penguji :


1. Ketua Sidang : Sarmayanta Sembiring, M.T



2. Sekretaris Sidang : Iman Saladin B. Azhar, M.MSI



3. Penguji Sidang : Ahmad Heryanto, M.T



4. Pembimbing I : Deris Stiawan, M.T., Ph.D., IPU



5. Pembimbing II : Ali Bardadi, S.SI., M.Kom.

Mengetahui,

Ketua Jurusan Sistem Komputer



Dr. Ir. H. Sukemi, M.T.

NIP. 196612032006041001

## HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini:

Nama : Tia Hermita  
NIM : 09011281722051  
Judul : Sistem Klasifikasi Serangan SQL *Injection* & XSS Pada RAMA  
*Repository* Dengan Metode *Support Vector Machine* (SVM)

Hasil Penyecekan *Software iThenticate / Turnitin*: 4%

Menyatakan bahwa laporan tugas akhir saya merupakan hasil karya sendiri dan bukan hasil penjiplakan atau plagiat. Apabila ditemukan unsur penjiplakan atau plagiat dalam laporan tugas akhir ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya.

Demikian, pernyataan ini saya buat dalam keadaan sadar dan tanpa paksaan dari siapapun.



Indralaya, Juli 2021



**Tia Hermita**  
**NIM. 09011281722051**

## HALAMAN PERSEMBAHAN

*“Hai orang-orang yang beriman, jadikanlah sabar dan shalat sebagai penolongmu”*

*(QS. Al-Baqarah : 153)*

*“Dan aku belum pernah kecewa dalam berdoa kepada-Mu, ya Tuhanku”*

*(QS. Maryam : 4)*

*“Kalau terlalu berat, just take some rest, don't ever give up”*

*(Kim Taehyung)*

*“If you want to give up, remember why you started”*

*(Park Jimin)*

*“Tugas Akhir ini kupersembahkan untuk kedua orang tua ku, kedua saudari ku, ketiga kucing ku, dan keluarga besar yang selalu menyertakan do'a dan memberikan semangat kepadaku”*

*“Lebih dari siapa pun, terima kasih kepada diri ini (Tia Hermita, S.Kom)”*

## KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakatuh

Puji dan syukur Alhamdulillah selalu penulis panjatkan atas kehadiran Allah Subhanahu Wata'ala yang telah melimpahkan rahmat serta karunia-Nya, sehingga penulis dapat menyelesaikan Proposal Tugas Akhir ini yang berjudul **“Sistem Klasifikasi Serangan SQL Injection & XSS pada RAMA Repository Dengan Metode Support Vector Machine (SVM)”**.

Pada kesempatan ini, penulis ingin mengucapkan rasa syukur kepada Allah Subhanahu Wata'ala dan terima kasih kepada beberapa pihak antara lain:

- Orang tua penulis, Ayahanda Sukurman dan Ibunda Suriwati serta kakak dan adik penulis yaitu Arianti Yulanda S.Kep., Ns. dan Anita Kristi yang selalu memberikan do'a, kekuatan dan semangat kepada penulis.
- Bapak Jaidan Jauhari, S.Pd., M.T., selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
- Bapak Dr. Ir. H. Sukemi, M.T., selaku Ketua Jurusan Sistem Komputer Fakultas Ilmu Komputer Universitas Sriwijaya.
- Bapak Deris Stiawan, M.T., Ph.D., IPU., selaku Dosen Pembimbing I Tugas Akhir di Jurusan Sistem Komputer Fakultas Ilmu Komputer Universitas Sriwijaya.
- Bapak Ali Bardadi, S.SI., M.Kom., selaku Dosen Pembimbing II Tugas Akhir di Jurusan Sistem Komputer Fakultas Ilmu Komputer Universitas Sriwijaya.
- Bapak Firdaus, M.Kom., selaku Pembimbing Akademik.
- Mbak Nurul Afifah M.Kom., yang telah membantu memberi arahan dalam mengerjakan tugas akhir.
- Mbak Renny Virgasari selaku admin Jurusan Sistem Komputer yang telah membantu penulis dalam hal administrasi.
- Amartya Bimantara, Lisa Melinda, Febi Rusmiati dan Aulia Melynda Putri selaku Tim Riset RAMA Repository Universitas Sriwijaya.

- Meutia Zamieyus, Ahmad Afidin, Nuzula Rahma Safitri, Agung Setiawan, Abdi Bimantara S.Kom, M. Taufiq Qurahman dan teman-teman seperjuangan grup Riset Comnets lainnya.
- Ghina, Xossya, Anom, Harun selaku sohib kental di SK.
- Muezza, Gula, dan Namu yang lebih dari sekadar kucing peliharaan yang menghilangkan stress serta menjadi teman begadang.
- Kakak tingkat dan teman-teman seperjuangan Sistem Komputer 2017.
- Seluruh pihak yang tidak dapat penulis sebutkan satu persatu.
- Almamater.

Penulis menyadari bahwa laporan ini masih sangat jauh dari kata sempurna. Untuk itu kritik dan saran yang membangun sangatlah diharapkan penulis. Akhir kata penulis berharap, semoga proposal tugas akhir ini bermanfaat dan berguna bagi khalayak.

Wassalamu'alaikum Warahmatullahi Wabarakatuh.

Indralaya, Juli 2021  
Penulis

# ***Classification System of SQL Injection & XSS on RAMA Repository Using Support Vector Machine (SVM)***

**Tia Hermita (09011281722051)**

*Dept. of Computer Engineering, Faculty of Computer Science, Sriwijaya  
University*

*email: [tiahermita29@gmail.com](mailto:tiahermita29@gmail.com)*

## ***ABSTRACT***

*RAMA repository is a repository for exposure to national research results in the form of student project final assignments (Diploma), thesis (S1), theses (S2), dissertations (S3) and other research reports. This makes the RAMA repository vulnerable to online database attacks including SQL injection and XSS. SQL injection attack is an online database attack in the first place, where the attacker modifies the SQL query sent by the user and injects malicious SQL queries into it. While XSS attacks are carried out by stealing web browser resources such as cookies from user accounts in order to access sensitive information on web applications. This study proposes the Correlation Feature Selection (CFS) algorithm to select the best features and the Support Vector Machine (SVM) algorithm to classify SQL and XSS injection attacks on RAMA repository data using 4 kernels, namely linear kernel, RBF, polynomial and sigmoid. Based on the test results, the RBF kernel produces the best performance in classifying with an accuracy rate of 99.35%, precision 99.33%, recall 99.35% and f1-score 99.34%.*

***Keywords:*** *Classification, RAMA repository, SQL Injection, XSS, Correlation-based Feature Selection (CFS), Support Vector Machine (SVM)*



# **Sistem Klasifikasi Serangan SQL *Injection* & XSS pada RAMA Repository Dengan Metode *Support Vector Machine* (SVM)**

**Tia Hermita (09011281722051)**

Jurusan Sistem Komputer, Fakultas Ilmu Komputer, Universitas Sriwijaya

email: [tiahermita29@gmail.com](mailto:tiahermita29@gmail.com)

## **ABSTRAK**

RAMA *repository* merupakan tempat penyimpanan paparan hasil penelitian nasional baik berbentuk tugas akhir proyek mahasiswa (Diploma), skripsi (S1), tesis (S2), disertasi (S3) maupun laporan penelitian lainnya. Hal ini membuat RAMA *repository* rentan terhadap serangan database online diantaranya injeksi SQL dan XSS. Serangan injeksi SQL merupakan serangan database online diurutan pertama, dimana penyerang mengubah kueri SQL yang dikirim oleh pengguna dan menyuntikkan kueri SQL berbahaya ke dalamnya. Sedangkan serangan XSS dilakukan dengan mencuri sumber daya web browser seperti cookie dari akun pengguna agar dapat mengakses informasi sensitive pada aplikasi web. Penelitian ini mengusulkan algoritma *Correlation Feature Selection* (CFS) untuk memilih fitur terbaik dan algoritma *Support Vector Machine* (SVM) untuk melakukan pengklasifikasian serangan injeksi SQL dan XSS pada data RAMA *repository* dengan menggunakan 4 kernel yaitu kernel *linear*, RBF, *polynomial* dan *sigmoid*. Berdasarkan hasil pengujian, kernel RBF menghasilkan performa paling baik dalam melakukan pengklasifikasian dengan tingkat akurasi sebesar 99.35%, presisi 99.33%, recall 99.35% dan f1-score 99.34%.

**Kata Kunci:** Klasifikasi, RAMA *repository*, SQL *Injection*, XSS, *Correlation-based Feature Selection* (CFS), *Support Vector Machine* (SVM)

## DAFTAR ISI

	<b>Halaman</b>
<b>HALAMAN JUDUL</b> .....	<b>i</b>
<b>LEMBAR PENGESAHAN</b> .....	<b>ii</b>
<b>HALAMAN PERSETUJUAN</b> .....	<b>iii</b>
<b>HALAMAN PERNYATAAN</b> .....	<b>iv</b>
<b>HALAMAN PERSEMBAHAN</b> .....	<b>v</b>
<b>KATA PENGANTAR</b> .....	<b>vi</b>
<b>ABSTRACT</b> .....	<b>viii</b>
<b>ABSTRAK</b> .....	<b>ix</b>
<b>DAFTAR ISI</b> .....	<b>x</b>
<b>DAFTAR GAMBAR</b> .....	<b>xiii</b>
<b>DAFTAR TABEL</b> .....	<b>xiv</b>
<b>DAFTAR LAMPIRAN</b> .....	<b>xv</b>
<b>BAB I     PENDAHULUAN</b>	
1.1   Latar Belakang .....	1
1.2   Perumusan Masalah .....	3
1.3   Batasan Masalah .....	3
1.4   Tujuan .....	3
1.5   Manfaat .....	4
1.6   Metodologi Penelitian .....	4
1.7   Sistematika Penulisan .....	5
<b>BAB II    TINJAUAN PUSTAKA</b>	
2.1   Pendahuluan .....	6
2.2 <i>SQL Injection</i> .....	8
2.2.1   Jenis Serangan <i>SQL Injection</i> .....	9
2.3 <i>Cross-Site Scripting</i> .....	9
2.4   Dataset <i>RAMA Repository</i> .....	10
2.5   Seleksi Fitur CFS .....	10
2.6 <i>Synthetic Minority Oversampling Technique (SMOTE)</i> .....	11

2.7	<i>Support Vector Machine (SVM)</i> .....	11
2.7.1	Kernel Linear .....	12
2.7.2	Kernel RBF .....	12
2.7.3	Kernel Polynomial .....	13
2.7.4	Kernel Sigmoid .....	13
2.8	<i>Confusion Matrix</i> .....	13

### **BAB III METODOLOGI PENELITIAN**

3.1	Pendahuluan .....	16
3.2	Kerangka Kerja Penelitian .....	16
3.3	Kerangka Kerja Metodologi Penelitian .....	18
3.4	Kebutuhan Perangkat .....	19
3.5	Persiapan Data .....	19
3.6	Konversi Data .....	20
3.7	Seleksi fitur Menggunakan <i>Correlation Feature Selection</i> .....	22
3.8	Klasifikasi dengan Algoritma SVM .....	23
3.9	Validasi Hasil .....	24

### **BAB IV HASIL DAN ANALISA**

4.1	Pendahuluan .....	26
4.2	Hasil Konversi Dataset .....	26
4.3	Seleksi Fitur CFS .....	31
4.4	Validasi Hasil .....	32
4.4.1	Validasi Hasil Skenario 1 .....	32
4.4.2	Validasi Hasil Skenario 2 .....	36
4.4.3	Validasi Hasil Skenario 3 .....	37
4.4.4	Validasi Hasil Skenario 4 .....	39
4.5	Validasi BACC dan MCC .....	41
4.5.1	Validasi BACC dan MCC Kernel Linear .....	41
4.5.2	Validasi BACC dan MCC Kernel RBF .....	42
4.5.3	Validasi BACC dan MCC Kernel Polynomial .....	43
4.5.4	Validasi BACC dan MCC Kernel Sigmoid .....	43

4.6	Analisa Hasil Perbandingan Fitur Seleksi .....	43
4.7	Analisa Hasil Perbandingan Algoritma SVM .....	44
<b>BAB V</b>	<b>KESIMPULAN</b>	
5.1	Kesimpulan .....	46
5.2	Saran .....	47
<b>DAFTAR PUSTAKA</b>	.....	<b>48</b>
<b>LAMPIRAN</b>	.....	<b>52</b>

## DAFTAR GAMBAR

	<b>Halaman</b>
<b>Gambar 2.1</b> Konsep Kernel <i>Trick</i> .....	12
<b>Gambar 2.2</b> <i>Confusion Matrix Multi-Class</i> .....	13
<b>Gambar 3.1</b> Kerangka Kerja Penelitian Keseluruhan .....	17
<b>Gambar 3.2</b> Kerangka Kerja Metodologi Penelitian .....	18
<b>Gambar 3.3</b> Flowchart Konversi Data .....	20
<b>Gambar 3.4</b> Proses Pemilihan Database dan Tabel yang Akan Diekstrak .....	21
<b>Gambar 3.5</b> Pemilihan Path dan Type File Hasil Konversi .....	21
<b>Gambar 3.6</b> Tampilan Fitur Dataset pada Google Colab .....	22
<b>Gambar 3.7</b> Seleksi Fitur Menggunakan Algoritma CFS .....	23
<b>Gambar 3.8</b> Flowchart Klasifikasi SVM .....	24
<b>Gambar 4.1</b> Dataset Sebelum di Konversi .....	26
<b>Gambar 4.2</b> Dataset Setelah di Konversi .....	27
<b>Gambar 4.3</b> Tautology Queries .....	27
<b>Gambar 4.4</b> Union Query .....	27
<b>Gambar 4.5</b> Piggy-based Query .....	28
<b>Gambar 4.6</b> Timming Attack .....	28
<b>Gambar 4.7</b> Order by .....	28
<b>Gambar 4.8</b> Count .....	29
<b>Gambar 4.9</b> Removed .....	29
<b>Gambar 4.10</b> Javascript .....	29
<b>Gambar 4.11</b> Alert .....	29
<b>Gambar 4.12</b> Persentase Jumlah Data Sebelum di <i>Balance</i> .....	30
<b>Gambar 4.13</b> Persentase Jumlah Data Setelah di <i>Balance</i> .....	30
<b>Gambar 4.14</b> Hasil Seleksi Fitur Dataset .....	31
<b>Gambar 4.15</b> Kurva ROCAUC Training Testing Kernel Linear .....	35
<b>Gambar 4.16</b> Kurva ROCAUC Training Testing Kernel RBF .....	37
<b>Gambar 4.17</b> Kurva ROCAUC Training Testing Kernel Polynomial .....	38
<b>Gambar 4.18</b> Kurva ROCAUC Training Testing Kernel Sigmoid .....	40
<b>Gambar 4.19</b> Perbandingan Performa Fitur Seleksi .....	44
<b>Gambar 4.20</b> Hasil Perbandingan Kernel Algoritma SVM .....	45

## DAFTAR TABEL

	<b>Halaman</b>
<b>Tabel 2.1</b> Penelitian Terkait Mengenai SQL dan XSS .....	7
<b>Tabel 2.2</b> Perbandingan Dengan Penelitian Sebelumnya .....	8
<b>Tabel 3.1</b> Spesifikasi Perangkat Keras .....	19
<b>Tabel 3.2</b> Spesifikasi Perangkat Lunak .....	19
<b>Tabel 3.3</b> Jumlah Dataset Sebelum di <i>Balance</i> .....	19
<b>Tabel 3.4</b> Jumlah Dataset Setelah di <i>Balance</i> .....	20
<b>Tabel 3.5</b> Hasil Konversi Dataset menggunakan <i>tools</i> MySQL Workbench ....	21
<b>Tabel 3.6</b> Hyperparameter Pengklasifikasian SVM .....	25
<b>Tabel 4.1</b> Atribut yang Digunakan Dalam Penelitian .....	32
<b>Tabel 4.2</b> <i>Confusion Matrix</i> Kernel Linear .....	32
<b>Tabel 4.3</b> Validasi Fine Tuning Kernel Linear .....	35
<b>Tabel 4.4</b> Nilai <i>confusion matrix</i> Fine Tuning Kernel Linear .....	36
<b>Tabel 4.5</b> <i>Confusion Matrix</i> Kernel RBF .....	36
<b>Tabel 4.6</b> Validasi Fine Tuning Kernel RBF .....	36
<b>Tabel 4.7</b> Nilai <i>confusion matrix</i> Fine Tuning Kernel RBF .....	37
<b>Tabel 4.8</b> <i>Confusion Matrix</i> Kernel Polynomial .....	38
<b>Tabel 4.9</b> Validasi Fine Tuning Kernel Polynomial .....	38
<b>Tabel 4.10</b> Nilai <i>confusion matrix</i> Fine Tuning Kernel Poynomial .....	39
<b>Tabel 4.11</b> <i>Confusion Matrix</i> Kernel Sigmoid .....	39
<b>Tabel 4.12</b> Validasi Fine Tuning Kernel Sigmoid .....	40
<b>Tabel 4.13</b> Nilai <i>confusion matrix</i> Fine Tuning Kernel Sigmoid .....	41
<b>Tabel 4.14</b> Hasil Validasi BACC dan MCC Kernel Linear .....	42
<b>Tabel 4.15</b> Hasil Validasi BACC dan MCC Kernel RBF .....	42
<b>Tabel 4.16</b> Hasil Validasi BACC dan MCC Kernel Polynomial .....	43
<b>Tabel 4.17</b> Hasil Validasi BACC dan MCC Kernel Sigmoid .....	43
<b>Tabel 4.18</b> Hasil Perbandingan Performa Fitur Seleksi .....	44

## DAFTAR LAMPIRAN

	<b>Halaman</b>
HASIL TURNITIN .....	52
HASIL SULIET .....	53

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Serangan injeksi SQL dan *Cross-Site Scripting* (XSS) merupakan kerentanan umum serta menjadi ancaman utama bagi aplikasi web berbasis data [1] dan merupakan metode paling umum dibalik pelanggaran keamanan data [2] dengan bentuk serangan injeksi dimana penyerang memasukkan perintah SQL kedalam parameter input untuk mengubah eksekusi kueri SQL di server [3]. Dalam sebuah perintah SQL, injeksi SQL dan XSS memiliki makna disetiap karakternya sehingga diperlukannya pengecekan terhadap karakter-karakter inputan yang masuk dari user kemudian dilakukan pengkonversian inputan agar menjadi inputan yang aman.

Kerentanan dapat dihasilkan secara tidak sengaja oleh pengembang perangkat lunak selama fase pengembangan [4]. *SQL Injection Attacks* (SQLIA) terjadi saat penyerang dapat memasukkan serangkaian pernyataan SQL yang berbahaya kedalam “kueri” dalam bentuk web, nama domain, atau permintaan halaman sehingga mengakibatkan cacat injeksi SQL, NoSQL, *OS Injection*, dan *LDAP Injection* melalui manipulasi data input pengguna untuk dieksekusi oleh database *back-end* [5].

*Cross-Site Scripting* (XSS) mengacu pada eksploitasi kerentanan sisi klien dengan injeksi skrip klien yang berbahaya [6] yang memungkinkan penyerang untuk menginjeksi kode skrip (biasanya) *JavaScript* yang dijalankan di browser dan dilihat oleh pengguna lain [7], dimana penyerang atau *attacker* akan mencuri sumber daya web browser seperti *cookie* atau *credentials* dari akun pengguna untuk mengakses informasi sensitif pada aplikasi web dan memasukkan konten berbahaya kedalam halaman web yang disajikan oleh aplikasi web terpercaya [8].

RAMA *repository* merupakan tempat penyimpanan paparan hasil eksperimen nasional baik berbentuk tugas akhir proyek mahasiswa (diploma), skripsi (S1), tesis (S2), disertasi (S3) maupun laporan penelitian yang dilakukan dosen atau peneliti yang bukan penerbitan jurnal, simposium, atau buku. Dimana RAMA *repository* ini tersambung pada ANJANI (Anjungan Integritas Akademik)



guna mendeteksi plagiasi nasional yang mengintegrasikan data dengan GARUDA dan SINTA [9].

Pada penelitian sebelumnya [10] menggunakan algoritma *Term Frequency Inverse Document Frequency* (TFIDF) untuk melakukan vektorisasi teks pada pernyataan SQL berdasarkan karakter sensitif dan kata kunci, sehingga dapat menghasilkan nilai fitur dengan perbedaan yang jelas antara vektor teks pernyataan tipe serangan dan pernyataan tipe normal. Dimana pada penelitian ini, algoritma TFIDF ini dikombinasikan dengan metode SVM, KNN, dan DT untuk mengetahui hasil performa terbaik dari gabungan algoritma dan metode yang digunakan. Dimana dari pengujian didapatkan hasil bahwa algoritma TFIDF + SVM menghasilkan tingkat akurasi paling baik yaitu 99.08% dibandingkan dengan gabungan algoritma TDIDF dengan kedua metode lainnya.

Pada penelitian [11] menggunakan algoritma *correlation feature selection* (CFS) untuk menyeleksi fitur apa saja yang memiliki korelasi kuat dengan kelas target tetapi tidak berkorelasi dengan atribut lain. Saat algoritma CFS belum diterapkan, hasil prediksi CKD (*Chronic Kidney Disease*) dari algoritma SVM menghasilkan tingkat akurasi sebesar 63.75%, sedangkan ketika algoritma CFS diterapkan sebagai fitur seleksi yang kemudian dikombinasikan dengan algoritma SVM untuk melakukan prediksi, maka didapatkan peningkatan performa akurasi yang lebih baik yaitu sebesar 88.75%.

Pada penelitian [12] menggunakan algoritma Decision Tree, Adaptive Boosting (AdaBoost) dan Random Forest untuk melakukan pengklasifikasian yang dikombinasikan dengan algoritma SMOTE untuk menyeimbangkan data penelitian. Algoritma SMOTE diterapkan dengan mengambil sampel yang berlebihan untuk mengatasi distribusi data ke kelas minoritas. Tingkat akurasi yang dihasilkan dari kombinasi algoritma Decision Tree dan SMOTE sebesar 92.8%, untuk kombinasi algoritma AdaBoost dan SMOTE sebesar 93.2%, dan kombinasi algoritma Random Forest dan SMOTE sebesar 94.6%.

Berdasarkan penjelasan diatas, penulis akan membahas mengenai klasifikasi SQL injection dan XSS pada RAMA *repository* menggunakan algoritma *correlation feature selection* (CFS) untuk menyeleksi fitur dan algoritma *Support Vector Machine* (SVM).

## 1.2 Perumusan Masalah

Perumusan masalah dalam penelitian ini, yaitu:

1. Bagaimana penerapan algoritma *correlation-based feature selection* (CFS) terhadap kinerja sistem klasifikasi pada RAMA repository?
2. Bagaimana cara mengklasifikasi paket data normal dan paket data serangan injeksi SQL & XSS pada RAMA repository
3. Bagaimana pengaruh tiap-tiap kernel terhadap performa tingkat akurasi, presisi, recall dan f1-score pada algoritma SVM?

## 1.3 Batasan Masalah

Batasan masalah dalam penelitian ini, yaitu:

1. Dataset yang digunakan pada riset ini bersumber dari website resmi RAMA repository
2. Data yang diujikan antara lain data normal dan data serangan injeksi SQL dan XSS
3. Seleksi fitur dilakukan menggunakan algoritma *correlation feature selection* (CFS)
4. Algoritma yang digunakan untuk mengklasifikasikan serangan injeksi SQL dan XSS pada RAMA repository ialah algoritma *Support Vector Machine* (SVM)
5. Pada riset ini tidak membahas sistem pencegahan serangan injeksi SQL dan XSS pada RAMA repository

## 1.4 Tujuan

Tujuan dari penelitian ini, ialah:

1. Mengetahui pengaruh algoritma *correlation-based feature selection* (CFS) terhadap kinerja sistem klasifikasi pada RAMA repository
2. Menerapkan algoritma SVM untuk mengklasifikasi data normal dan data serangan injeksi SQL & XSS pada RAMA repository
3. Mengukur tingkat performa akurasi, presisi, recall dan f1-score terhadap kinerja pengklasifikasian algoritma SVM

## 1.5 Manfaat

Manfaat dari penelitian ini, yaitu:

1. Algoritma *correlation-based feature selection* (CFS) sangat berpengaruh dalam peningkatan kinerja sistem klasifikasi serangan injeksi SQL dan XSS pada RAMA repository
2. Dapat mengklasifikasi data normal dan data serangan injeksi SQL dan XSS pada RAMA repository dengan menggunakan algoritma SVM
3. Dapat memvalidasi tiap-tiap kernel dalam meningkatkan performa akurasi, presisi, recall, dan f1-score

## 1.6 Metodologi Penelitian

Metodologi dalam penelitian ini melalui beberapa tahapan yaitu:

1. Tahapan Pertama (Study Pustaka / Literatur)  
Tahapan ini dilakukan dengan pencarian dan pengumpulan referensi berupa literature dari berbagai sumber ilmiah seperti *journal*, *paper*, buku dan internet mengenai topik penelitian yang dilakukan.
2. Tahapan Kedua (Pengolahan Data)  
Tahapan ini membahas proses konversi data atau merubah format data yang telah diperoleh yaitu dari format *.sql* menjadi file dengan format *.csv* agar dapat dibaca oleh mesin. Dan ditahap ini juga dilakukan proses seleksi fitur yang akan digunakan dengan menggunakan algoritma *correlation-based feature selection* (CFS).
3. Tahapan Ketiga (Klasifikasi)  
Tahapan ini akan dilakukan proses pengklasifikasian data normal dan data serangan injeksi SQL & *cross-site scripting* dengan menggunakan algoritma SVM.
4. Tahapan Keempat (Analisa)  
Tahapan ini akan menganalisis hasil dari eksperimen yang telah dilaksanakan dengan tujuan untuk dapat melihat tingkat performa dari penelitian yang telah dilakukan.

## 1.7 Sistematika Penulisan

Sistematika penulisan selama penyusunan tugas akhir ini, ialah:

### **BAB I. PENDAHULUAN**

Bab ini berisi tujuan dasar terkait subjek riset baik itu dari latar belakang, perumusan masalah, batasan masalah, tujuan dan manfaat dari riset yang dilakukan.

### **BAB II. TINJAUAN PUSTAKA**

Bab ini berisi beberapa *literature review* dari penelitian terkait topik klasifikasi injeksi SQL dan XSS menggunakan algoritma SVM.

### **BAB III. METODOLOGI PENELITIAN**

Bab ini memuat metode penelitian yang dilakukan. Penjelasan meliputi mulai dari tahap mempersiapkan data, penerapan algoritma *correlation-based feature selection* dan SVM.

### **BAB IV. ANALISA DAN PEMBAHASAN**

Bab ini memuat hasil dari eksperimen yang telah dilakukan serta analisis dari tiap data yang didapatkan dari hasil penelitian.

### **BAB V. KESIMPULAN DAN SARAN**

Bab ini menjelaskan kesimpulan yang didapat dari hasil riset yang dilakukan serta saran yang bisa digunakan untuk penelitian kedepannya.

## DAFTAR PUSTAKA

- [1] O. C. Abikoye, A. Abubakar, A. H. Dokoro, O. N. Akande, and A. A. Kayode, "A novel technique to prevent SQL injection and cross-site scripting attacks using Knuth-Morris-Pratt string match algorithm," *Eurasip J. Inf. Secur.*, vol. 2020, no. 1, pp. 47–54, 2016, doi: 10.1186/s13635-020-00113-y.
- [2] P. R. McWhirter, K. Kifayat, Q. Shi, and B. Askwith, "SQL Injection Attack classification through the feature extraction of SQL query strings using a Gap-Weighted String Subsequence Kernel," *J. Inf. Secur. Appl.*, vol. 40, pp. 199–216, 2018, doi: 10.1016/j.jisa.2018.04.001.
- [3] Q. Li, W. Li, J. Wang, and M. Cheng, "A SQL Injection Detection Method Based on Adaptive Deep Forest," *IEEE Access*, vol. 7, pp. 145385–145394, 2019, doi: 10.1109/ACCESS.2019.2944951.
- [4] K. Kamtuo and C. Soomlek, "Machine Learning for SQL Injection Prevention on Server-Side Scripting," 2016.
- [5] Z. S. Alwan and M. F. Younis, "Detection and Prevention of SQL Injection Attack: A Survey," *Int. J. Comput. Sci. Mob. Comput.*, vol. 6, no. 8, pp. 5–17, 2017, [Online]. Available: <https://www.ijcsmc.com/docs/papers/August2017/V6I8201701.pdf>.
- [6] C. Obimbo, K. Ali, and K. Mohamed, "Using IDS to prevent XSS Attacks," *Int'l Conf. Secur. Manag.*, pp. 233–239, 2017.
- [7] B. K. Ayeni, J. B. Sahalu, and K. R. Adeyanju, "Detecting Cross-Site Scripting in Web Applications Using Fuzzy Inference System," *J. Comput. Networks Commun.*, vol. 2018, 2018, doi: 10.1155/2018/8159548.
- [8] M. Ebady Manaa and R. Hussein, "Preventing Cross Site Scripting Attacks in Websites," 2016.
- [9] R Ristekdikti, "Panduan RAMA REPOSITORY," 2019.
- [10] Y. Li and B. Zhang, "Detection of SQL Injection Attacks Based on Improved TFIDF Algorithm," *J. Phys. Conf. Ser.*, vol. 1395, no. 1, 2019, doi: 10.1088/1742-6596/1395/1/012013.
- [11] D. Aprilianto, "SVM Optimization with Correlation Feature Selection Based

- Binary Particle Swarm Optimization for Diagnosis of Chronic Kidney Disease,” pp. 24–31, 2020.
- [12] G. Hu, T. Xi, F. Mohammed, and H. Miao, “Classification of wine quality with imbalanced data,” *Proc. IEEE Int. Conf. Ind. Technol.*, vol. 2016-May, pp. 1712–1717, 2016, doi: 10.1109/ICIT.2016.7475021.
- [13] M. Hasan, Z. Balbahaith, and M. Tarique, “Detection of SQL Injection Attacks: A Machine Learning Approach,” *2019 Int. Conf. Electr. Comput. Technol. Appl. ICECTA 2019*, 2019, doi: 10.1109/ICECTA48151.2019.8959617.
- [14] Solomon Ogbomon Uwagbole and L. F. William J. Buchanan, “Applied Machine Learning Predictive Analytics to SQL Injection Attack Detection and Prevention,” pp. 1087–1090, 2017.
- [15] A. Luo, W. Huang, and W. Fan, “A CNN-based Approach to the Detection of SQL Injection Attacks,” *Proc. - 18th IEEE/ACIS Int. Conf. Comput. Inf. Sci. ICIS 2019*, pp. 320–324, 2019, doi: 10.1109/ICIS46139.2019.8940196.
- [16] P. Tang, W. Qiu, Z. Huang, H. Lian, and G. Liu, “Detection of SQL injection based on artificial neural network,” *Knowledge-Based Syst.*, vol. 190, p. 105528, 2020, doi: 10.1016/j.knosys.2020.105528.
- [17] H. Zhang, J. Zhao, B. Zhao, X. Yan, H. Yuan, and F. Li, “SQL injection detection based on deep belief network,” Oct. 2019, doi: 10.1145/3331453.3361280.
- [18] X. Xie, C. Ren, Y. Fu, J. Xu, and J. Guo, “SQL Injection Detection for Web Applications Based on Elastic-Pooling CNN,” *IEEE Access*, vol. 7, pp. 151475–151481, 2019, doi: 10.1109/ACCESS.2019.2947527.
- [19] Y. Fang, J. Peng, L. Liu, and C. Huang, “WOVSQLI: Detection of SQL injection behaviors using word vector and LSTM,” in *ACM International Conference Proceeding Series*, Mar. 2018, pp. 170–174, doi: 10.1145/3199478.3199503.
- [20] D. Mitropoulos, P. Louridas, M. Polychronakis, and A. D. Keromytis, “Defending Against Web Application Attacks: Approaches, Challenges and Implications,” *IEEE Trans. Dependable Secur. Comput.*, vol. 16, no. 2, pp. 188–203, 2019, doi: 10.1109/TDSC.2017.2665620.

- [21] E. Pollack, "Dynamic SQL," *Dyn. SQL*, pp. 31–60, 2019, doi: 10.1007/978-1-4842-4318-3.
- [22] J. Abirami, R. Devakunchari, and C. Valliyammai, "A top web security vulnerability SQL injection attack - Survey," *ICoAC 2015 - 7th Int. Conf. Adv. Comput.*, 2016, doi: 10.1109/ICoAC.2015.7562806.
- [23] F. Tong and Z. Yan, "A hybrid approach of mobile malware detection in Android," *J. Parallel Distrib. Comput.*, vol. 103, pp. 22–31, 2017, doi: 10.1016/j.jpdc.2016.10.012.
- [24] Z. Xiao, Z. Zhou, W. Yang, and C. Deng, "An approach for SQL injection detection based on behavior and response analysis," *2017 9th IEEE Int. Conf. Commun. Softw. Networks, ICCSN 2017*, vol. 2017-Janua, pp. 1437–1442, 2017, doi: 10.1109/ICCSN.2017.8230346.
- [25] Z. S. Alwan and M. F. Younis, "Detection and Prevention of SQL Injection Attack: A Survey," *Int. J. Comput. Sci. Mob. Comput.*, 2017.
- [26] P. Raman, D. Howe, and A. Somayaji, "Jaspin: Javascript Based Anomaly Detection of Cross-Site Scripting Attacks," *Carlet. Univ.*, no. September, pp. 1–122, 2008.
- [27] F. A. Mereani, "The International Conference on Advanced Machine Learning Technologies and ... - Google Books," vol. 2, pp. 200–210, 2018, doi: 10.1007/978-3-319-74690-6.
- [28] M. A. Hall, "Correlation-based Feature Selection for Discrete and Numeric Class Machine Learning."
- [29] R. J. Palma-Mendoza, L. de-Marcos, D. Rodriguez, and A. Alonso-Betanzos, "Distributed correlation-based feature selection in spark," *Inf. Sci. (Ny)*, vol. 496, pp. 287–299, Sep. 2019, doi: 10.1016/j.ins.2018.10.052.
- [30] P. Yildirim, "Filter Based Feature Selection Methods for Prediction of Risks in Hepatitis Disease," *Int. J. Mach. Learn. Comput.*, vol. 5, no. 4, pp. 258–263, Aug. 2015, doi: 10.7763/ijmlc.2015.v5.517.
- [31] S. T. Jishan, R. I. Rashu, N. Haque, and R. M. Rahman, "Improving accuracy of students' final grade prediction model using optimal equal width binning and synthetic minority over-sampling technique," *Decis. Anal.*, vol. 2, no. 1, pp. 1–25, 2015, doi: 10.1186/s40165-014-0010-2.

- [32] S. Amarappa and S. V Sathyanarayana, “Data classification using Support vector Machine (SVM), a simplified approach.” [Online]. Available: [www.ijecse.org](http://www.ijecse.org).
- [33] A. M. Andrew, “ AN INTRODUCTION TO SUPPORT VECTOR MACHINES AND OTHER KERNEL-BASED LEARNING METHODS by Nello Christianini and John Shawe-Taylor, Cambridge University Press, Cambridge, 2000, xiii+189 pp., ISBN 0-521-78019-5 (Hbk, £27.50). ,” *Robotica*, vol. 18, no. 6, pp. 687–689, Nov. 2000, doi: 10.1017/s0263574700232827.
- [34] A. S. Nugroho, A. B. Witarto, and D. Handoko, “Application of Support Vector Machine in Bioinformatics,” *Proceeding Indones. Sci. Meet. Cent. Japan, December 20, 2003, Gifu-Japan*, pp. 842–847, 2011, doi: 10.1109/CCDC.2011.5968300.
- [35] M. Grandini, E. Bagli, and G. Visani, “Metrics for multi-class classification: An overview,” *arXiv*, pp. 1–17, 2020.