P-ISSN : 2580-4316
E-ISSN : 2654-8054 https://doi.org/10.37817/ikraith-informatika.v10i2

Implementasi *Middleware* Deteksi *SQL Injection* Berbasis *Multinomial Naïve Bayes* dan Validasi Ketahanan Terhadap SQLMap

[1]Malik Syafi'i, [2]Arip Solehudin, [3]Purwantoro
[123]Informatika, Universitas Singaperbangsa Karawang, Karawang

E-mail: [1]maliksyafii426@gmail.com, [2]arip.solehudin@fik.unsika.ac.id,
[3]purwantoro.masbro@staff.unsika.ac.id

ABSTRAK

Penelitian ini bertujuan mengembangkan middleware deteksi dan pemblokiran otomatis serangan *SQL Injection* (SQLi) menggunakan algoritma *Multinomial Naïve Bayes* (MNB) yang diintegrasikan pada aplikasi web berbasis Flask. Berbeda dengan studi sebelumnya yang berfokus pada akurasi model statis atau perbandingan kinerja sistem, penelitian ini menekankan pada penerapan metodologi *Knowledge Discovery in Databases* (KDD) secara penuh untuk membangun model pertahanan aktif serta validasi empiris ketahanannya terhadap alat eksploitasi otomatis. Model dilatih menggunakan dataset publik RbSQLi melalui tahapan seleksi, pra-pemrosesan teks berbasis *Regex Abstraction*, transformasi TF-IDF, hingga evaluasi internal. Hasil pengujian menunjukkan bahwa model mencapai akurasi, presisi, *recall*, dan F1-Score sebesar 100% pada data uji terisolasi. Validasi operasional menggunakan *black-box penetration testing* dengan SQLMap membuktikan bahwa middleware berhasil memblokir seluruh vektor serangan utama (*Error-based, Union-based, Boolean-blind, Time-blind,* dan *Stacked Queries*), sehingga mengubah status parameter target dari rentan (*injectable*) menjadi aman (*not injectable*). Temuan ini menegaskan bahwa pendekatan KDD dengan algoritma MNB mampu menghasilkan lapisan keamanan aplikatif yang efektif dalam menetralisir ancaman SQLi otomatis tanpa memerlukan arsitektur komputasi yang kompleks.

Kata kunci : ***SQL Injection, Multinomial Naïve Bayes, Knowledge Discovery in Databases, Middleware, SQLMap.***

ABSTRACT

This research aims to develop an automated detection and blocking middleware for SQL Injection (SQLi) attacks using the Multinomial Naïve Bayes (MNB) algorithm integrated into a Flask-based web application. Unlike previous studies focusing on static model accuracy or system performance comparisons, this study emphasizes the full application of the Knowledge Discovery in Databases (KDD) methodology to build an active defense model and empirically validate its resilience against automated exploitation tools. The model was trained using the public RbSQLi dataset through stages of selection, Regex Abstraction-based text preprocessing, TF-IDF transformation, and internal evaluation. Test results show that the model achieved 100% accuracy, precision, recall, and F1-Score on isolated test data. Operational validation using black-box penetration testing with SQLMap proved that the middleware successfully blocked all major attack vectors (Error-based, Union-based, Boolean-blind, Time-blind, and Stacked Queries), thereby changing the target parameter status from vulnerable (injectable) to safe (not injectable). These findings confirm that the KDD approach with the MNB algorithm can produce an effective application security layer in neutralizing automated SQLi threats without requiring complex computational architectures.

Keyword : ***SQL Injection, Multinomial Naïve Bayes, Knowledge Discovery in Databases, Middleware, SQLMap.***

https://doi.org/10.37817/ikraith-informatika.v10i2
P-ISSN : 2580-4316
E-ISSN : 2654-8054

1. PENDAHULUAN

Peningkatan adopsi teknologi digital memperluas permukaan serangan (*attack surface*) aplikasi web, menjadikan *SQL Injection* (SQLi) sebagai salah satu ancaman paling kritis akibat kesederhanaan eksekusinya namun berdampak fatal pada integritas dan kerahasiaan data (Ahmad & Karim, 2021; Madya et al., 2025). Insiden global seperti pencurian data oleh kelompok ResumeLooters dan kebocoran data instansi pemerintah di Indonesia membuktikan bahwa validasi input konvensional sering kali gagal menghadapi variasi payload modern yang dihasilkan alat otomatisasi seperti SQLMap (Group-IB, 2024; Antara News, 2024).

Meskipun berbagai penelitian terdahulu melaporkan akurasi tinggi menggunakan arsitektur *deep learning* seperti CNN-BiLSTM (Gandhi et al., 2021) atau RNN Autoencoder (Alghawazi et al., 2023), terdapat kesenjangan antara performa akademis dan kelayakan implementasi praktis. Mayoritas studi hanya mengevaluasi model secara pasif pada dataset tanpa mengintegrasikannya sebagai komponen pertahanan aktif dalam pipeline aplikasi web nyata (Mahmood, 2025). Selain itu, kompleksitas komputasi model canggih sering kali menjadi hambatan operasional, padahal algoritma dengan kompleksitas rendah seperti Naïve Bayes terbukti kompetitif untuk klasifikasi teks keamanan siber (Arnap & Kusrini, 2024; Konyrbaev et al., 2024).

Penelitian ini mengisi celah tersebut dengan menerapkan metodologi *Knowledge Discovery in Databases* (KDD) secara komprehensif sebagai kerangka kerja utama, bukan sekadar sebagai sub-tahapan pendukung. Fokus penelitian diarahkan pada pembangunan middleware berbasis *Multinomial Naïve Bayes* yang divalidasi ketahanannya secara langsung terhadap simulasi serangan otomatis SQLMap. Kontribusi utama artikel ini adalah demonstrasi empiris efektivitas pipeline KDD dalam menghasilkan model keamanan ringan yang mampu mengubah status kerentanan sistem secara deterministik pada lingkungan aplikasi web berbasis Flask.

2. LANDASAN TEORI

2.1 SQL Injection dan Alat Otomatisasi

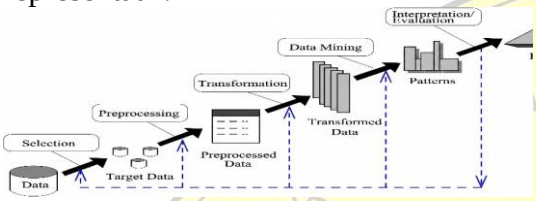
SQLi memanfaatkan kelemahan validasi input untuk memanipulasi kueri basis data secara dinamis. Serangan ini dikategorikan menjadi *In-Band* (Error-based, Union-based), *Blind/Inferential* (Boolean-based, Time-based), dan *Out-of-Band* (EC-Council, 2020; Mehmood & Ijaz, 2024). Dalam pengujian keamanan modern, alat seperti SQLMap menjadi standar validasi karena kemampuannya mengotomatisasi enumerasi dan eksploitasi berbagai teknik injeksi, termasuk penggunaan *tamper scripts* untuk mengelabui mekanisme filtrasi (Maulana & Subardono, 2025). Keberhasilan middleware keamanan diukur dari kemampuannya menahan agregasi serangan otomatis ini hingga parameter dinyatakan *not injectable*.

2.2 Knowledge Discovery in Databases (KDD)

KDD merupakan pendekatan sistematis untuk mengekstraksi pola berharga dari data masif yang terdiri dari tahapan *Data Selection, Data*

P-ISSN : 2580-4316
E-ISSN : 2654-8054 https://doi.org/10.37817/ikraith-informatika.v10i2

Cleaning, Data Transformation, Data Mining, serta *Interpretation and Evaluation* (Karnila et al., 2022). Dalam konteks keamanan siber, KDD menjamin bahwa model deteksi dibangun dari data yang berkualitas, tertransformasi dengan benar, dan dievaluasi secara valid sebelum diimplementasikan, sehingga mengurangi risiko *false positive* akibat data yang kotor atau tidak representatif.



Gambar 1. Tahapan *Knowledge Discovery in Databases*

2.3 Multinomial Naïve Bayes dan Pra-pemrosesan Teks

Multinomial Naïve Bayes (MNB) mengoperasionalkan Teorema Bayes dengan asumsi independensi kondisional antar fitur, menjadikannya efisien untuk klasifikasi teks berbasis frekuensi kata (Wabang et al., 2022). Efektivitas MNB sangat bergantung pada pra-pemrosesan teks, khususnya tokenisasi berbasis *Regular Expression* (Regex) dan ekstraksi fitur TF-IDF yang memberikan bobot lebih tinggi pada token khas serangan SQLi (Gerliandeva et al., 2024; Petrus et al., 2023). Kombinasi Regex abstraction dan TF-IDF memungkinkan model fokus pada pola sintaksis logika serangan alih-alih nilai literal yang bervariasi.

3. METODOLOGI

Penelitian ini menerapkan metodologi KDD secara penuh sebagai kerangka kerja operasional yang terdiri dari lima tahapan sistematis:

Data Selection: Dataset RbSQLi dari Mendeley Data diseleksi berdasarkan relevansi variabel muatan kueri dan label kategorikal eksplisit.

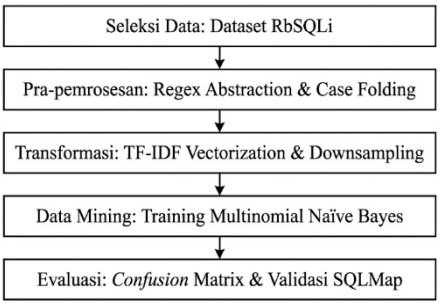
Data Cleaning & Preprocessing: Dilakukan penanganan *missing values*, penghapusan duplikat, *case folding*, serta abstraksi nilai literal string dan angka menjadi token standar menggunakan Regex untuk mereduksi variabilitas. Label ditransformasi ke format biner (0: benign, 1: SQLi) dan kelas diseimbangkan melalui *downsampling* proporsional rasio 1:1.

Data Transformation: Ekstraksi fitur dilakukan menggunakan TF-IDF *word n-gram* (1,3) dengan batas 5.000 fitur. Proses *fitting* vektorizer hanya dilakukan pada data latih untuk mencegah *data leakage*.

Data Mining: Model MNB dilatih menggunakan *pipeline* terintegrasi dan divalidasi stabilitasnya melalui *5-Fold Stratified Cross-Validation*. Model terbaik diserialisasi ke format .pkl untuk deployment.

Interpretation & Evaluation: Evaluasi dilakukan dalam dua tahap. Pertama, evaluasi internal menggunakan *Confusion Matrix* (Akurasi, Presisi, Recall, F1-Score) pada data uji independen. Kedua, evaluasi eksternal melalui *black-box testing* menggunakan SQLMap pada aplikasi Flask yang telah terintegrasi middleware, untuk memvalidasi perubahan status parameter dari *injectable* menjadi *not injectable*.

https://doi.org/10.37817/ikraith-informatika.v10i2
P-ISSN : 2580-4316
E-ISSN : 2654-8054



Gambar 2. Alur KDD yang di Adaptasi

4. HASIL DAN PEMBAHASAN

4.1 Hasil Preparasi Data dan Pelatihan Model

Proses KDD berhasil mereduksi dataset awal yang tidak seimbang menjadi 162.864 sampel terstruktur dengan distribusi kelas 1:1. Tahap abstraksi Regex terbukti krusial dalam menyatukan variasi payload numerik dan string menjadi pola sintaksis umum, sehingga mempercepat konvergensi model MNB. Dataset kemudian dibagi menjadi 80% data latih (130.291 sampel) dan 20% data uji (32.573 sampel).

Tabel 1. Hasil Penyeimbangan Kelas Dinamis

Jenis Serangan	Jumlah Sampel
Sintaks Normal	81432
Error-Based	13572
Time-Based	13572
Meta_based	13572
Union-Based	13572
Boolean-based	13572
Stackqueries_based	13572
Total	162.864

4.2 Evaluasi Internal Model

Berdasarkan pengujian pada subset data uji independen, model MNB menunjukkan performa klasifikasi yang optimal sebagaimana disajikan pada Tabel 2.

Tabel 2. Confusion Matrix Model Naïve Bayes pada Data Pengujian

	Prediksi: Serangan (1)	**Prediksi: Normal (0)**
Aktual: Serangan (1)	16.287 (TP)	0 (FN)
Aktual: Normal (0)	0 (FP)	16.286 (TN)

Model mencapai Akurasi 100%, Presisi 100%, Recall 100%, dan F1-Score 100%. Pencapaian ini didorong oleh efektivitas kombinasi Regex abstraction dan TF-IDF n-gram yang berhasil mengisolasi fitur sintaksis kueri secara kontras. Meskipun hasil ini merepresentasikan batas atas (*upper bound*) pada dataset sintetis yang teratur, hal ini menjadi prasyarat mutlak sebelum model diuji pada skenario serangan dunia nyata yang lebih dinamis. Analisis kualitatif juga mencatat potensi *False Positive* pada input bahasa alami yang mengandung substring SQL, yang menjadi catatan untuk pengembangan dataset masa depan.

4.3 Validasi Ketahanan Terhadap SQLMap

Model yang telah tervalidasi diimplementasikan sebagai middleware pada aplikasi Flask. Middleware dikonfigurasi untuk mencegat permintaan HTTP GET/POST, melakukan inferensi real-time, dan mengembalikan respons HTTP 403 Forbidden jika terdeteksi sebagai serangan.

Validasi eksternal dilakukan menggunakan SQLMap terhadap aplikasi terproteksi. Hasil pengujian menunjukkan bahwa middleware berhasil memblokir seluruh teknik eksploitasi yang diujikan. Status parameter target berubah secara definitif dari *injectable* (pada kondisi

P-ISSN : 2580-4316
E-ISSN : 2654-8054 https://doi.org/10.37817/ikraith-informatika.v10i2

tanpa proteksi) menjadi *not injectable* saat middleware aktif. Rekapitulasi hasil validasi disajikan pada Tabel 3.

Tabel 3. Hasil Validasi Ketahanan Middleware Terhadap SQLMap

Tipe Eksploitasi	Status Middleware	Keterangan Efek Middleware
Boolean-based blind	Berhasil Diblokir	Request ditolak
Error-based	Berhasil Diblokir	Request ditolak
Stacked queries	Berhasil Diblokir	Request ditolak
Time-based blind	Berhasil Diblokir	Request ditolak
UNION query	Berhasil Diblokir	Request ditolak
Ekstraksi DB (--dbs)	Berhasil Diblokir	Log: Not Injectable

Temuan ini membuktikan bahwa model MNB yang dibangun melalui proses KDD yang ketat tidak hanya akurat secara statistik di atas kertas, tetapi juga tangguh secara operasional dalam menghadapi variasi payload otomatis SQLMap. Kemampuan middleware untuk memutus koneksi sebelum kueri mencapai basis data mengonfirmasi efektivitasnya sebagai lapisan pertahanan proaktif.

Gambar 3. Hasil pengujian dengan SQLmap

5. KESIMPULAN

Penerapan metodologi KDD secara penuh dalam membangun middleware deteksi SQLi berbasis Multinomial Naïve Bayes terbukti menghasilkan sistem pertahanan yang efektif. Model mencapai metrik evaluasi internal yang sempurna (100% Akurasi, Presisi, Recall, F1-Score) berkat pra-pemrosesan Regex abstraction dan TF-IDF yang tepat. Lebih penting lagi, validasi eksternal menggunakan SQLMap mengonfirmasi bahwa middleware berhasil memblokir seluruh vektor serangan utama dan mengubah status kerentanan sistem menjadi aman (not injectable). Penelitian ini mendemonstrasikan bahwa pendekatan KDD dengan algoritma ringan dapat menjembatani kesenjangan antara akurasi akademis dan kelayakan implementasi keamanan aplikatif yang tangguh terhadap alat eksploitasi otomatis.

6. UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih yang sebesar-besarnya kepada Fakultas Ilmu Komputer Universitas Singaperbangsa Karawang atas dukungan fasilitas dan lingkungan akademik yang kondusif selama pelaksanaan penelitian ini. Apresiasi khusus disampaikan kepada Bapak Arip Solehudin, M.Kom. dan Bapak Purwantoro, M.Kom. selaku dosen

pembimbing yang telah memberikan arahan, bimbingan teknis, serta masukan berharga sejak tahap perancangan hingga penyusunan naskah. Penulis juga berterima kasih kepada seluruh dosen Program Studi Informatika yang telah membekali penulis dengan landasan keilmuan yang kuat, serta kepada rekan-rekan seperjuangan dan semua pihak yang telah memberikan dukungan moral maupun material sehingga artikel ilmiah ini dapat terselesaikan dengan baik.

DAFTAR PUSTAKA

Ahmad, K., & Karim, M. (2021). A Method to Prevent SQL Injection Attack using an Improved Parameterized Stored Procedure. In *IJACSA) International Journal of Advanced Computer Science and Applications* (Vol. 12, Number 6). www.ijacsa.thesai.org

Alghawazi, M., Alghazzawi, D., & Alarifi, S. (2023). Deep Learning Architecture for Detecting SQL Injection Attacks Based on RNN Autoencoder Model. *Mathematics*, *11*(15). https://doi.org/10.3390/math11153286

Arnap, A., & Kusrini. (2024). Enhancing SQL Injection Attack Detection Using Naïve Bayes and SMOTE Method on Imbalanced Datasets. *Journal of Artificial Intelligence and Engineering Applications*, *4*(1), 2808–4519. https://ioinformatic.org/

EC-Council. (2020). *Ethical hacking and countermeasures* (Version 11). EC-Council.

Gandhi, N., Patel, J., Sisodiya, R., Doshi, N., & Mishra, S. (2021). A CNN-BiLSTM based Approach for Detection of SQL Injection Attacks. *Proceedings of 2nd IEEE International Conference on Computational Intelligence and Knowledge Economy, ICCIKE 2021*, 378–383. https://doi.org/10.1109/ICCIKE51210.2021.9410675

Gerliandeva, A., Chrisnanto, Y. H., & Ashaury, H. (2024). Optimasi Klasifikasi Sentimen pada Komentar Online menggunakan Multinomial Naïve Bayes dan Ekstraksi Fitur TF-IDF serta N-grams. *Jurnal Pekommas*, *9*(2), 260–272. https://doi.org/10.56873/jpkm.v9i2.5585

Group-IB. (2024, February 6). Stealing your career: Group-IB uncovers ResumeLooters' data thefts. https://www.group-ib.com/media-center/press-releases/resumelooters/

Karnila, S., Rizkyandi, A., Kurniawan, R., & Nurjoko. (2022). Market Basket Analysis on Transaction Data Using the Apriori Algorithm. *Journal of Physics: Conference Series*.

Konyrbaev, N., Nikitenko, Y., Shtanko, V., Lakhno, V., Baishemirov, Z., Ibadulla, S., Galymzhankyzy, A., & Myrzabek, E. (2024). EVALUATION AND OPTIMIZATION OF THE NAIVE BAYES ALGORITHM FOR INTRUSION DETECTION SYSTEMS USING THE USB-IDS-1 DATASET. *Eastern-European Journal of Enterprise Technologies*, *6*(2(132)), 74–82. https://doi.org/10.15587/1729-4061.2024.317471

Madya, A. D., Purnomo, R., & Nurfiyah. (2025). Analisis Kerentanan Keamanan Website Menggunakan Open Web Application Security Project (Owasp) Top-10 Studi Kasus (web.bnpb.go.id). *Indonesian Journal of Education And Computer Science*, *3*(2), 51–65. https://doi.org/10.60076/indotech.v3i2.1411

Mahmood, S. S. (2025). SQL Injection Detection Using Machine Learning and Explainability. *Journal of Internet Services and Information Security*, *15*(2), 309–324. https://doi.org/10.58346/JISIS.2025.I2.022

Maulana, D., & Subardono, A. (2025). Analisis Efektivitas Tools SQLMap, Havij, dan Ghauri dalam Melakukan Serangan SQL Injection pada Website. *Journal of Internet and Software Engineering*, *6*(1).

Mehmood, M., & Ijaz, A. (2024). A Systematic Literature Review on SQL Injection Attacks. *NUML International Journal of Engineering and Computing*, *2*(2). https://doi.org/10.52015/nijec.v2i2.50

Petrus, J., Ermatita, Sukemi, & Erwin. (2023). A Novel Approach: Tokenization Framework based on Sentence Structure in Indonesian Language. In *IJACSA) International Journal of Advanced Computer Science and Applications* (Vol. 14, Number 2). https://doi.org/10.14569/IJACSA.2023.0140264

Wabang, K., Nurhayati, O. D., & Farikhin. (2022). Application of The Naïve Bayes Classifier Algorithm to Classify Community Complaints. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, *6*, 872–876. https://doi.org/10.29207/resti.v6i5.4498