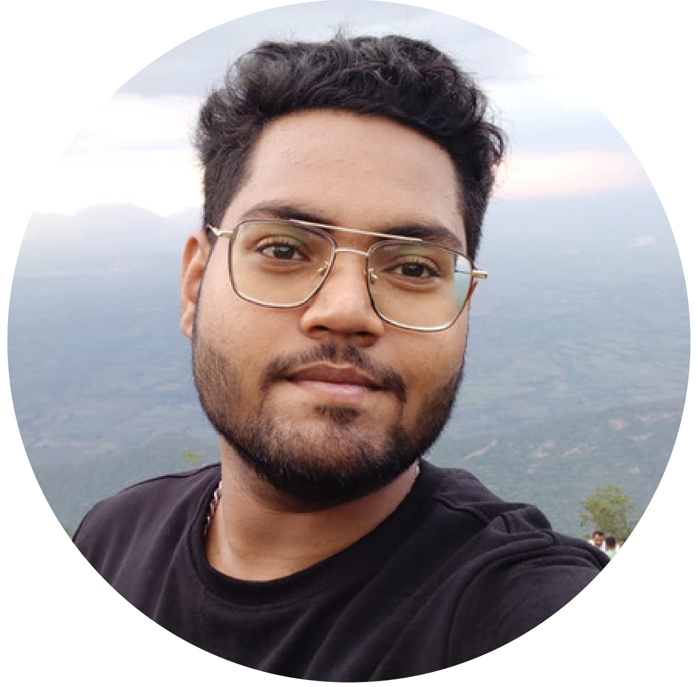
Indian Institute of
Information Technology
Kottayam

SQLIADS

SQL Injection Attack and Detection System using ML

Guided By- Dr. E. Silmabarasan

Harsh Gupta
(2020BCS0187)



Harsh Raj
(2020BCS0119)

Harish Vishwa
(2020BCS0117)

Likhitha
(2020BCS0087)

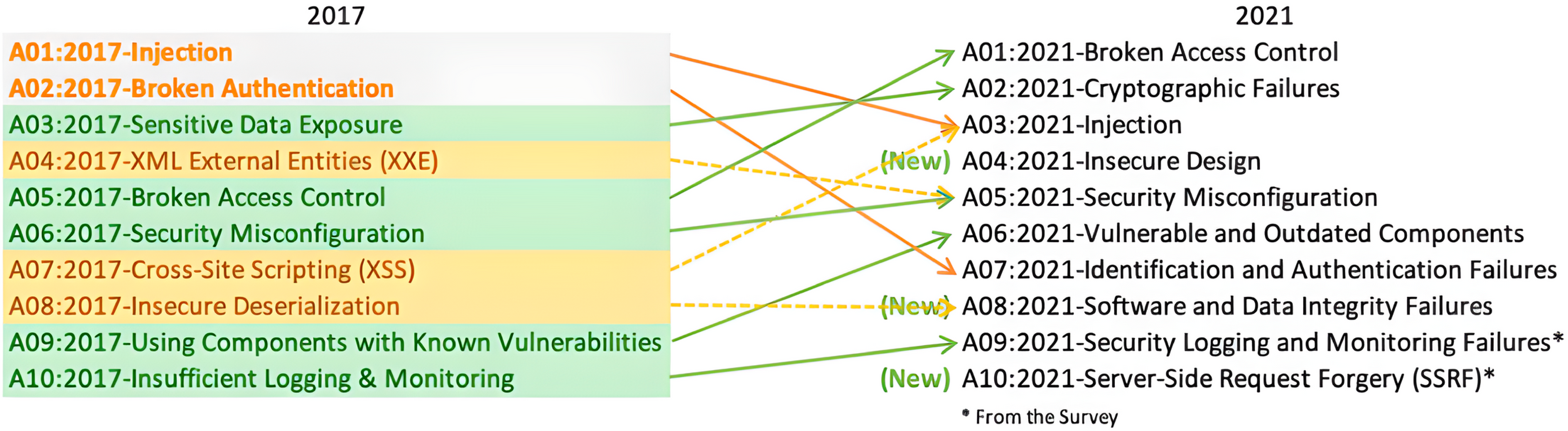
INTRODUCTION

Web applications are vital for tasks like online banking and shopping but they are usually susceptible to SQL injections.[1]

SQL injection is a type of attack where an attacker injects malicious code into a web application's database.[3]

This code can then be used to steal data, modify data, or even take control of the application.

In this study, we will investigate the different methods that can be used to detect SQL injection attacks and their limitations.



2017

A01:2017-Injection

A02:2017-Broken Authentication

A03:2017-Sensitive Data Exposure

A04:2017-XML External Entities (XXE)

A05:2017-Broken Access Control

A06:2017-Security Misconfiguration

A07:2017-Cross-Site Scripting (XSS)

A08:2017-Insecure Deserialization

A09:2017-Using Components with Known Vulnerabilities

A10:2017-Insufficient Logging & Monitoring

2021

A01:2021-Broken Access Control

A02:2021-Cryptographic Failures

A03:2021-Injection

(New) A04:2021-Insecure Design

A05:2021-Security Misconfiguration

A06:2021-Vulnerable and Outdated Components

A07:2021-Identification and Authentication Failures

(New) A08:2021-Software and Data Integrity Failures

A09:2021-Security Logging and Monitoring Failures*

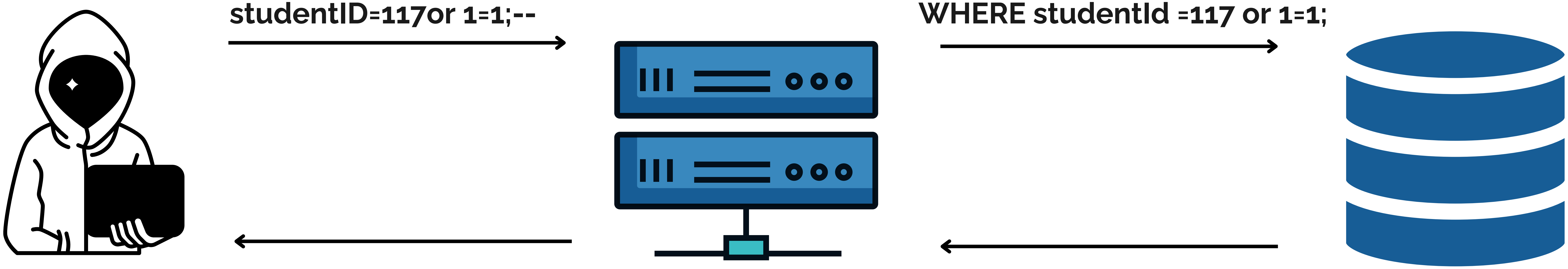
(New) A10:2021-Server-Side Request Forgery (SSRF)*

* From the Survey

*Image and data taken from OWASP

The Open Web Application Security Project (OWASP) ranks SQL injection as one of the top 10 web application security vulnerability. This is because it is a very effective attack that can be used to exploit a wide variety of web applications.

SQL Injection



www.students.com?
studentID=117or 1=1;--

SELECT * FROM students
WHERE studentId =117 or 1=1;

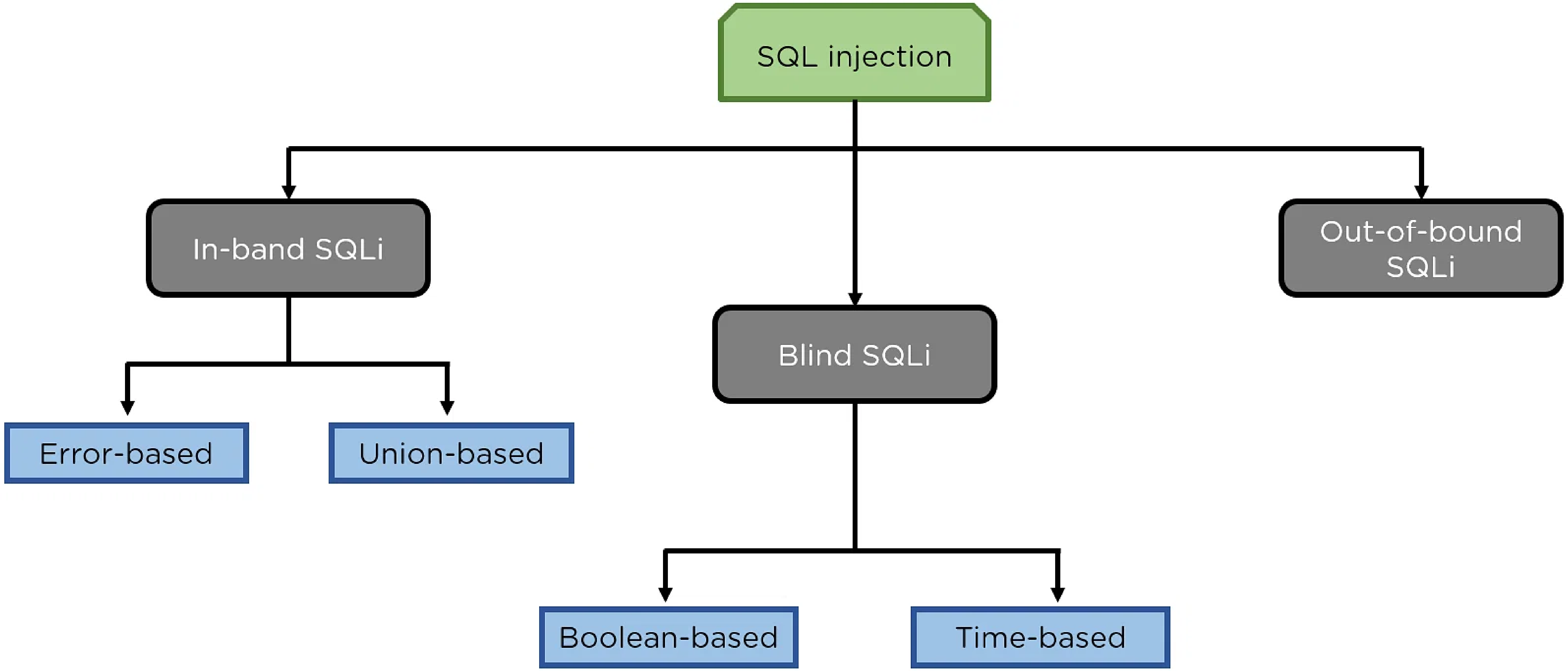
Attacker

Data of all students
is returned to Attacker

Web API Server

Return data for
all students

SQL Databse



Types of SQL Injection Attacks

Literature Review

Author/Title	Advantages	Limitations
Artificial Intelligence Techniques for SQL Injection Attack Detection (Irungu et al. (2023))	- **High accuracy:** The model achieves a high accuracy of 98.3605% for SVM, 96.296% for KNN, and 97.530% for Random Forest. - **Scalability:** The model is scalable, meaning that it can be used to detect SQLIA in large datasets. This is important because the number of SQL queries that are generated on a daily basis is constantly increasing.	- **Can be bypassed by sophisticated attackers:** The model is not perfect and can be bypassed by sophisticated attackers. However, it is still a valuable tool for detecting SQLIA. - **Not applicable to all types of SQLIA**: The model is not applicable to all types of SQLIA. For example, it is not able to detect blind SQLIA attacks.
A Survey on SQL Injection Attacks, Detection and Prevention (Hu et al. (2020))	**Comprehensive Review**: The paper promises a detailed review of various types of SQL injection attacks and detection techniques based on machine learning. This comprehensive approach provides readers with a thorough understanding of the subject matter.	**Lack of Emulation Analysis**: The paper mentions that the methods were not analyzed by emulating them. This means that the evaluation did not involve practical testing or simulation of real-world scenarios. Without emulation or practical testing, it's challenging to assess how these methods would perform in actual applications.

Literature Review

Author/Title	Advantages	Limitations
A Survey on SQL Injection Attack: Detection and Challenges	**Diverse Detection Methods**: The research explores various methods for detecting SQL Injection attacks, including static analysis, dynamic analysis, hybrid approaches, and the use Machine Learning (ML). This diversity allows for a comprehensive evaluation of detection techniques.	**Detection of Blind SQL Attacks:** Blind SQL Attack cannot be detected using Static and Dynamic Analysis.
SQL Injection Detection Using Machine Learning Techniques and Multiple Data Sources (Ross et al. (2018))	• **Multi-Source Data Analysis:** By collecting data from multiple sources the system enhances accuracy in detecting SQL injection attacks. This approach allows for a more comprehensive analysis. • **Alternative Algorithms:** The project explores alternative detection algorithms such as rule-based and decision tree algorithms, These algorithms are also faster in terms of model building and execution time.	• **Data Collection Complexity:** Collecting data from multiple sources can be complex and may require additional resources and configurations..

Motivation

- The increasing threat of SQL injection attacks has motivated researchers to develop more effective detection methods. Traditional methods of detection are often ineffective against new attacks.
- One promising machine learning approach is to use Artificial Neural Networks (ANNs). They promise improved pattern recognition and enhanced scalability compared to traditional methods.

Motivation

- In one study, researchers compared the performance ANNs and support vector machines (SVMs) for SQL injection detection. The results showed that ANNs had a higher accuracy, shorter testing time, and lower false positive rate than SVMs.

- The results of this study are encouraging, and they suggest that ML has the potential to fight against SQL injection attacks. However, more research is needed to make the model have more accuracy and lesser testing time.

Motivation

Model	Accuracy	Model Training time	Model Testing Time
SVM	94.025%	2min 35sec	1min 10sec
ANN	96.71%	46min 3sec	3.35sec

Webapp Dataset

Model	True Pos.	False Pos.	True Neg.	False Neg.
SVM	9134	329	9671	866
ANN	9501	158	9842	499

Confusion Matrix

Motivation

Model	Accuracy	Model Training time	Model Testing Time
SVM	95.19%	2min 11sec	1min 3sec
ANN	97.28%	41min 23sec	2.95sec

Dataiphy Dataset

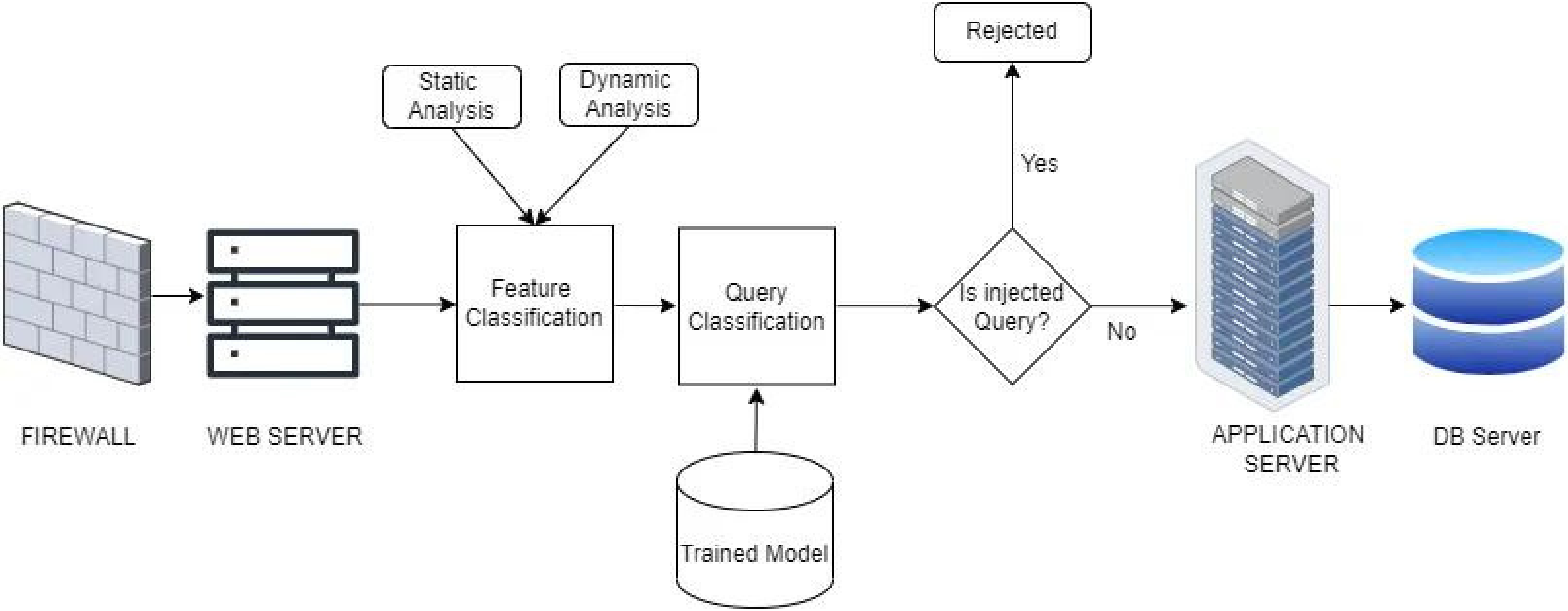
Model	True Pos.	False Pos.	True Neg.	False Neg.
SVM	9266	188	9812	774
ANN	9613	156	9844	387

Confusion Matrix

Problem Statement

- To detect and prevent SQL Injection attack.
- To analyze and conclude which ML model can be used to detect SQLI with the best precision.
- To find a trade-off balance between training efficiency and testing time.

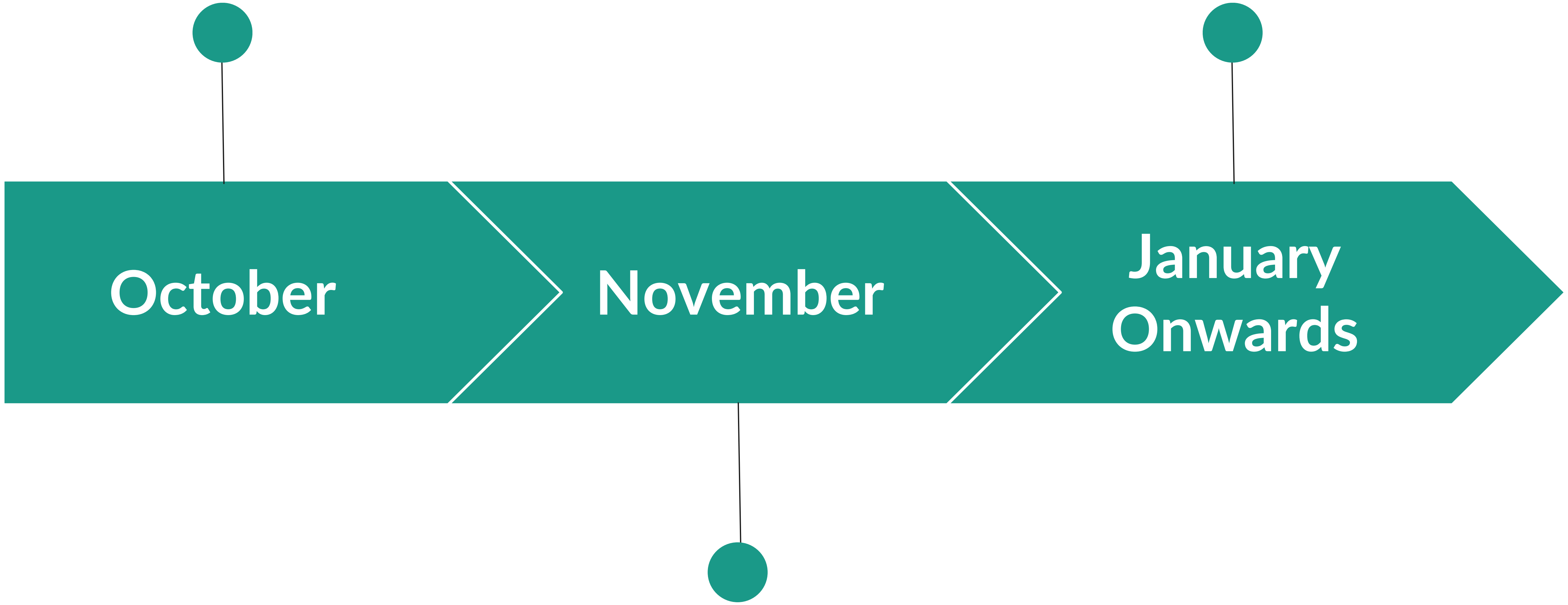
Architecture



Conclusion

We've acknowledged that SQL poses a significant threat to web platforms. To counteract these threats effectively, we're turning to machine learning models like SVM and ANN.

Our aim to make a model which has high accuracy, less training time and fast testing. ensuring that web applications remain fast and secure. To achieve this, we're working on creating a model with high accuracy that trains quickly and detects threats rapidly.



October

To find or generate a classified dataset which has both In-band and Blind SQLI queries.

November

Experiment with SVM and ANN models

January Onwards

Optimizing the result obtained in previous experiments

References

- *John Irungu, Steffi Graham, Anteneh Girma, and Thabet Kacem. 2023. Artificial Intelligence Techniques for SQL Injection Attack Detection. In Proceedings of the 2023 8th International Conference on Intelligent Information Technology (ICIIT '23). Association for Computing Machinery, New York, NY, USA, 38–45. https://doi.org/10.1145/3591569.3591576.*
- *Jianwei Hu, Wei Zhao, and Yanpeng Cui. 2020. A Survey on SQL Injection Attacks, Detection and Prevention. In Proceedings of the 2020 12th International Conference on Machine Learning and Computing (ICMLC '20). Association for Computing Machinery, New York, NY, USA, 483–488. https://doi.org/10.1145/3383972.3384028.*
- *P. Kumar and R. K. Pateriya, "A survey on SQL injection attacks, detection and prevention techniques," 2012 Third International Conference on Computing, Communication and Networking Technologies (ICCCNT'12), Coimbatore, India, 2012, pp. 1-5, doi: 10.1109/ICCCNT.2012.6396096.*
- *Ross, Kevin, "SQL Injection Detection Using Machine Learning Techniques and Multiple Data Sources" (2018). Master's Projects. 650. DOI: https://doi.org/10.31979/etd.zknb-4z36*
- *Kevin Zhang. 2019. A machine learning based approach to identify SQL injection vulnerabilities. In 2019 34th IEEE/ACM International Conference on Automated Software Engineering (ASE). IEEE, 1286–1288.*
- *Li Q, Wang F, Wang J, et al. LSTM-Based SQL Injection Detection Method for Intelligent Transportation System[J]. IEEE Transactions on Vehicular Technology, 2019, 68(5): 4182-4191*

Thank You!!