Celebrating 50 Years of Excellence

SYMBIOSIS INSTITUTE OF TECHNOLOGY, PUNE

TECHVORTEX 2.0

Team Name & Members:

The Qubits - Akhilesh T S

TELL US ABOUT YOURSELF

- Student of SASTRA University currently pursuing 4th year
- Secured 3rd place in TATA Communications Hackathon
- Won several competitions involving Data analysis and Machine Learning

Selected Problem Statement

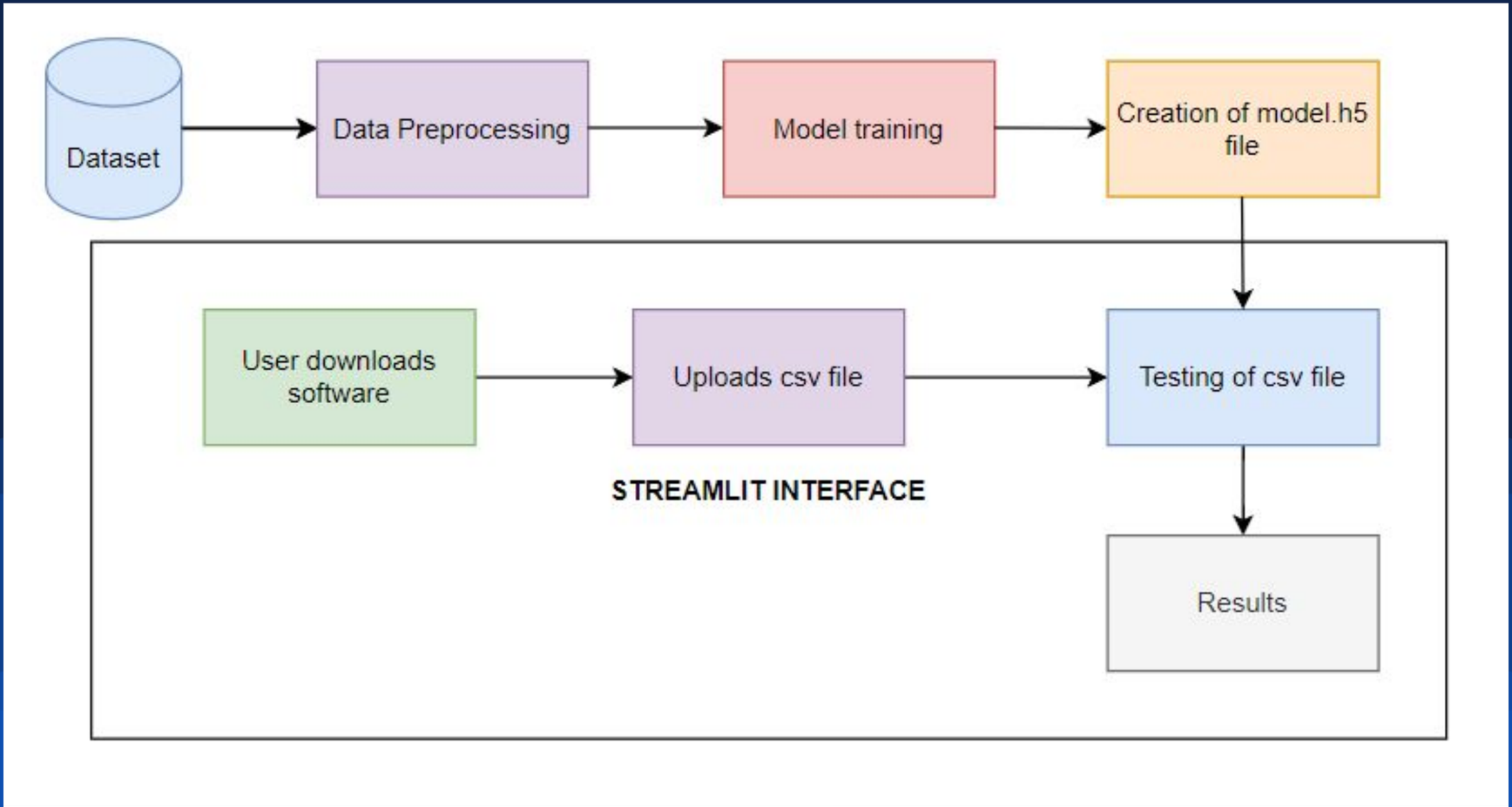
Domain Chosen: AI/ML and Cyber Security

- The increasing frequency of zero-day exploits and ransomware attacks highlights the need for advanced detection methods that can quickly adapt to new, unseen threats.
- Traditional security systems often fail to identify evolving threats in real time, making it essential to develop AI-driven systems capable of continuous, real-time network traffic monitoring.
- Capturing crucial details such as IP addresses, communication protocols, and metadata from network traffic is critical for detecting abnormal behavior indicative of malicious activity.

Tech Stack

- Streamlit: Web-based user interface framework for easy interaction and visualization.
- Pandas: Data manipulation and analysis for handling CSV files.
- Joblib: Model serialization/deserialization for loading the pre-trained ML model.
- Matplotlib: Data visualization to generate pie charts for attack statistics.
- Machine Learning (RandomForest Classifier: Pre-trained model (likely classification-based) used to predict intrusions from processed data.

Detailed Description of the Solution



How is your solution different?

- Real-Time Intrusion Detection: The system processes network traffic data from CSV files and detects intrusions in real-time using a machine learning model.
- Pre-trained models analyze uploaded data for abnormal patterns or malicious activities, classifying records as either normal or attack traffic.
- Provides clear insights into detected threats, including the percentage of attack traffic and specific samples of detected intrusions.
- Utilizes graphical visualizations (like pie charts) to help users quickly understand the proportion of normal traffic versus attacks.

Future Possible Enhancements

- Real-Time Data Streaming: Implement real-time data processing to detect and respond to intrusions as they occur.
- Advanced Data Visualization: Enhance visualizations with additional tools like heatmaps and time-series plots for deeper analysis of attack patterns.
- Model Performance Monitoring: Include features to track and display model performance metrics and support periodic model updates.
- User Authentication and Access Control: Introduce user authentication and role-based access control to secure the application and manage user permissions effectively.

Risks/ Challenges / Dependencies

- Data Quality and Integrity: Ensuring that the uploaded CSV files are correctly formatted, complete, and free from errors is crucial for accurate detection and analysis.
- Model Accuracy and Performance: The effectiveness of intrusion detection heavily depends on the quality and performance of the machine learning model. Regular updates and validation are needed to maintain accuracy.
- Real-Time Processing Challenges: Implementing real-time data streaming and analysis can be complex and resource-intensive, requiring robust infrastructure and efficient algorithms.
- Security and Privacy Concerns: Handling sensitive data and ensuring the application's security against potential attacks or unauthorized access is essential to protect user data and system integrity.

Acceptance Criteria Coverage

How many aspects of the problem statement have been covered?

- AI/ML
- Cyber Security

Result Analysis

- Random Forest Classifier used that predicts approximately 99% accuracy
- Scalable system with more different types of attacks
- F1-score, ROC AUC curve and other metrics generate good results with respect to the Intrusion Detection System

Key Contributions

- Interactive Web Interface: Provides a user-friendly interface via Streamlit for easy data upload, processing, and visualization, making intrusion detection accessible even to non-technical users.
- Automated Intrusion Detection: Utilizes a pre-trained machine learning model to automatically analyze network traffic and detect potential security threats, reducing the need for manual inspection.
- Real-Time Threat Visualization: Incorporates visual tools, such as pie charts, to clearly present the results of intrusion detection, helping users quickly understand the nature and extent of potential attacks.
- Customizable and Extensible Framework: Offers a modular and flexible framework that can be extended with real-time data processing, additional visualizations, and advanced model management, supporting future enhancements and adaptability.

Conclusion

- Effective Threat Detection: The implemented IDS successfully utilizes real-time network traffic monitoring and AI-driven algorithms to accurately detect and differentiate between benign and malicious activities, addressing the challenge of zero-day attacks and ransomware.

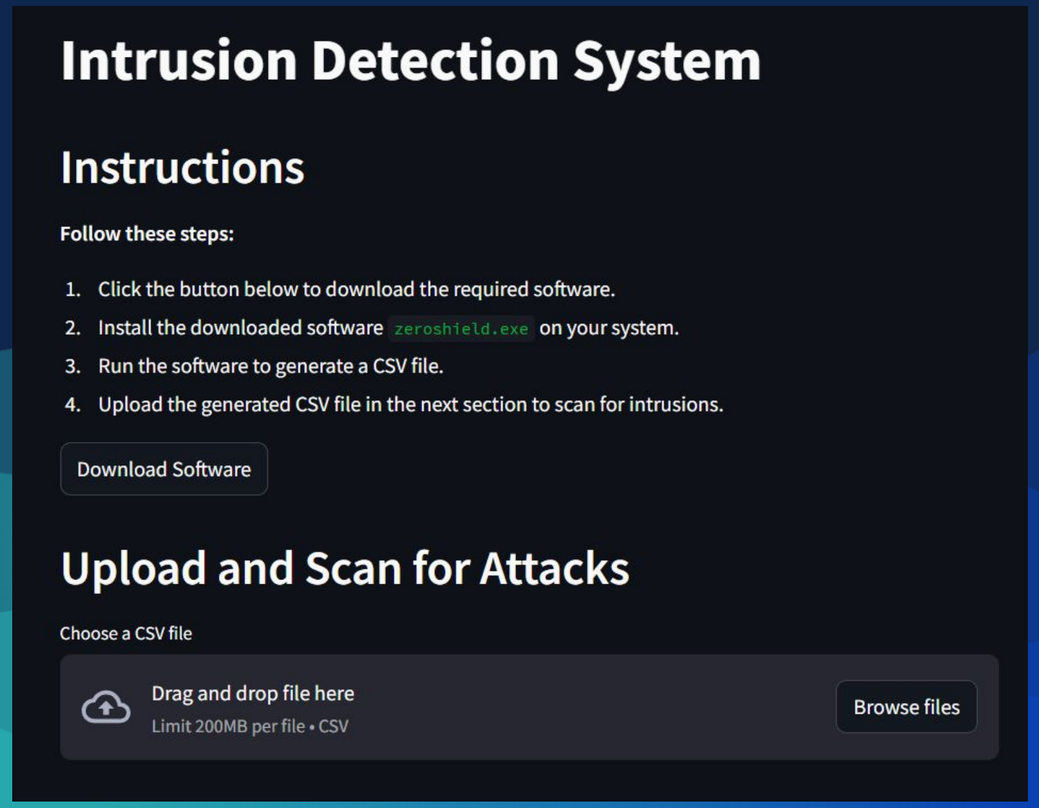
- Real-Time Performance: The system demonstrates efficient performance in detecting threats with low latency, ensuring timely identification and mitigation of potential security breaches.

- Visual and Metric Insights: The Streamlit interface provides clear visualizations and metrics, such as accuracy, precision, and ROC curves, facilitating a comprehensive understanding of the system's performance and areas for improvement.

- User Feedback Integration: By incorporating user feedback, the project ensures that the system remains adaptable and responsive to real-world needs, paving the way for continuous enhancement and effectiveness in evolving threat landscapes.

Anything Else ?

- **Github link:** https://github.com/akhilesh1709/Intrusion-detection-system



User Interface