Capstone Project Concept Note and Implementation Plan

Project Title: AI-Driven Hybrid Cyber Defensive System for Intelligent Malware Detection and Threat Insight

Team Members

[Chimdessa Tesfaye]
[Getachewu Getu]
[Meseret]
[Rediet]

Concept Note

1. Project Overview

This project aims to develop an AI-driven hybrid cyber defensive system that leverages machine learning (ML), deep learning (DL), and explainable AI (XAI) to detect and mitigate intelligent malware and cyber threats in real-time. The system addresses the growing sophistication of cyber threats, which traditional signature-based security tools often fail to counter. By integrating proactive threat detection, behavioural analysis, and transparent decision-making, the project contributes to Sustainable Development Goal (SDG) 9 (Industry, Innovation, and Infrastructure) by enhancing the resilience of digital infrastructure against cyberattacks.

2. Objectives

- Develop a hybrid AI-based system integrating ML, DL, and XAI for malware detection.
- Process unstructured threat intelligence using NLP for real-time insights.
- Design an interpretable system with automated response capabilities.
- Evaluate the effectiveness of different AI approaches in cybersecurity applications.

3. Background

Modern cyber threats, such as polymorphic malware and advanced persistent threats (APTs), evade traditional detection methods. Existing solutions like signature-based systems lack adaptability, while AI-driven approaches often suffer from opacity. This project bridges these gaps by combining advanced AI techniques with explainability, ensuring both accuracy and transparency in threat detection.

4. Methodology

The project will employ:

- **ML/DL Models**: Random Forest, XGBoost, LSTM, and CNN for malware classification.
- **NLP**: BERT and SpaCy for processing threat intelligence from unstructured sources.
- **XAI Frameworks**: SHAP and LIME to provide interpretable model outputs.
- **Hybrid Analysis**: Combining static (code features) and dynamic (behavioural) analysis for robust detection.

5. Architecture Design Diagram

Key Components:

1. **Data Ingestion Layer**: Collects malware samples, network logs, and threat feeds.
2. **Preprocessing Module**: Cleans and normalizes data for model input.
3. **AI Engine**: Runs ML/DL models for threat detection.
4. **XAI Module**: Generates explanations for model decisions.
5. **Response System**: Triggers automated countermeasures based on threat severity.

6. Data Sources

The system will use:

- **Kaggle Malware Dataset** (CSV/PE files) for labeled training data.
- **GitHub Malware Feeds** (raw binaries) for real-world samples.
- **Network Traffic Logs** (PCAP files) for anomaly detection. Data preprocessing includes imputation, feature scaling, and resampling to address class imbalance.

7. Literature Review

Existing research (e.g., Hasan & Kabir, 2021; Prity et al., 2024) highlights the limitations of traditional methods and the promise of AI in cybersecurity. This project extends their work by integrating XAI for transparency and hybrid models for improved accuracy, as demonstrated in Venkatraman et al. (2019).

Implementation Plan

1. Technology Stack

- **Programming Languages**: Python (Pandas, NumPy, Scikit-learn).
- **Libraries/Frameworks**: TensorFlow, Keras, PyTorch (DL), SpaCy, BERT (NLP), SHAP/LIME (XAI).
- **Tools**: Jupyter Notebooks, Git, Docker (containerization).
- **Hardware**: GPU-enabled servers for model training.

2. Timeline (Gantt Chart)

Task	Week 1	Week 2	Week 3	Week 4
Data Collection & Preprocessing	✓			
Model Development		✓	✓	
Training & Evaluation			✓	✓
Deployment & Testing				✓

Task Distribution:

- **Member 1**: Data preparation and preprocessing.
- **Member 2**: ML model development.
- **Member 3**: DL/NLP integration and
- **Member 4**: XAI implementation.

3. Milestones

1. Completion of data preprocessing pipeline (Week 1).
2. Successful training of hybrid ML/DL models (Week 3).
3. Deployment of MVP with ≥90% detection accuracy (Week 4).

4. Challenges and Mitigations

- **Data Quality**: Use robust imputation and outlier detection.
- **Model Performance**: Employ ensemble learning and hyperparameter tuning.
- **Technical Constraints**: Optimize models for GPU acceleration.

5. Ethical Considerations

- **Privacy**: Anonymize sensitive data in logs.
- **Bias**: Ensure diverse malware samples to prevent model skew.
- **Transparency**: Use XAI to maintain accountability in automated decisions.

6. References

[1] Hasan & Kabir, 2021. [2] Prity et al., 2024. [3] Venkatraman et al., 2019. (Full references in document).