|  |  |  |  |
| --- | --- | --- | --- |
| **Project Title:** | **Intelligent Network Intrusion Detection Using Machine Learning** | **Project ID:** |  |
| **Team Members** |  | Date of Request: |  |
|  |  | **Executive Sponsor(s):** | N/A |
|  |  | **Project Type** | N/A |
| **Faculty Sponsor:** |  | **Research Track** |

**Introduction**

Provide a brief overview of your project and explain why your area of research is important. (150 words Max)

|  |
| --- |
| In today’s increasingly digital world, cyber threats pose significant risks to the confidentiality, integrity, and availability of network systems. This project focuses on developing an AI-powered Intrusion Detection System (IDS) that can monitor, analyze, and detect suspicious network activity in real time. Unlike traditional IDS solutions that rely on fixed signatures and predefined rules, this system uses machine learning and deep learning techniques to identify both known and emerging threats. By learning from network behavior and detecting anomalies, the AI-based IDS offers greater accuracy, faster response times, and adaptability to evolving cyberattacks. This research is crucial as it addresses the limitations of conventional security systems and contributes to building smarter, more resilient cybersecurity infrastructures. |

**Problem Definition:**

What problem do you intend to solve in this research? (200 words max)

|  |
| --- |
| Problem Statement:  As cyber threats continue to evolve in complexity and scale, traditional Intrusion Detection Systems (IDSs) that rely on static rules and known attack signatures struggle to detect new, unknown, or subtle attacks. These systems are limited in adaptability and often produce a high rate of false positives, which can overwhelm security teams and lead to delayed or missed threat responses. There is a critical need for an intelligent, adaptive solution that can analyze network traffic in real time and detect both known and emerging threats with high accuracy.  Evidence that the Problem Exists:  Numerous cybersecurity reports show a steady rise in zero-day attacks, advanced persistent threats (APTs), and sophisticated malware that bypass traditional detection systems. Additionally, incidents such as data breaches and ransomware attacks highlight the inability of legacy IDS tools to cope with modern attack vectors. Studies reveal that false positives account for up to 90% of alerts in some systems, reducing efficiency and increasing response times.  Benefits of Solving the Problem:  Developing an AI-based IDS will provide proactive and dynamic threat detection, reduce false alarms, and enhance real-time response. This leads to more robust network security, better resource utilization, and improved protection for sensitive data and infrastructure. |

**Methodology and Experience**

Describe the methodology you intend to use to solve the stated problem. This includes data collection and data analysis. Briefly describe your previous experience with this methodology. If you don't have previous experience, outline your plan to learn this methodology. (200 words max)

|  |
| --- |
| Methodology  This project employs a straightforward yet effective machine learning approach to detect network intrusions. A supervised learning model, such as Decision Tree or Random Forest, will be used due to their strong classification capabilities and ease of implementation. The focus is on building a reliable system that can distinguish between normal and malicious network traffic.  Data Collection  The project will utilize publicly available datasets such as NSL-KDD or CICIDS2017. These datasets provide pre-labeled examples of network activity, which include both normal behavior and a variety of attack types. This allows for efficient training and testing without the need for custom data collection.  Data Analysis  Preprocessing techniques like normalization and feature selection will be applied to prepare the data. The selected machine learning model will be trained and tested on this data, and its performance will be evaluated using metrics such as accuracy, precision, recall, and confusion matrix to ensure reliable results.  Previous Experience with Methodology or Plans to Learn  We have foundational knowledge of Python and Scikit-learn, which are suitable tools for this project. I plan to strengthen my understanding through hands-on practice and trusted online resources to ensure successful implementation of the chosen methodology. |

**Research Questions and/or Hypotheses:**

Please state you research questions. If you have a hypotheses, also state it here. (150 words max).

|  |
| --- |
| **Research Questions:**   1. How effectively can machine learning algorithms detect both known and unknown cyber threats in network traffic? 2. What are the most accurate features for distinguishing between normal and malicious network activity in available datasets? 3. How does the performance of simpler models (e.g., Decision Trees) compare to more complex models (e.g., deep learning) in detecting intrusions in network traffic?   **Hypothesis:** Machine learning-based Intrusion Detection Systems (IDS) can outperform traditional signature-based IDS by accurately detecting both known and unknown attack patterns with minimal false positives, particularly when using simpler models such as Decision Trees or Random Forests. |

**Literature Review and Related Work (150 words max)**

*Briefly summarize related literature in this area and how your work differs from what has already been done.*

|  |
| --- |
| Many existing Intrusion Detection Systems (IDS) use signature-based methods, which rely on known attack patterns. These systems have trouble detecting new or advanced attacks. Recent research has shown that machine learning (ML) methods, like decision trees, support vector machines (SVM), and deep learning, can improve detection by identifying new attack types. Studies by **Amor et al. (2020)** and **Sharma et al. (2019)** have shown that ML-based IDS can reduce false positives compared to traditional methods.  Many current methods are complex or focus on detecting specific types of attacks. This project stands out by using simpler models like Decision Trees and Random Forests, which are efficient and effective. The goal is to provide an easy-to-use and accurate solution that works with publicly available datasets (NSL-KDD, CICIDS2017) for detecting network threats in real time. |
|  |

**Deliverables**

What are the deliverables for your project?

|  |
| --- |
| **Intrusion Detection System (IDS):** A fully functional AI-powered IDS that can monitor and analyze network traffic to detect potential threats in real time using machine learning algorithms.  **Model Code:** The source code used to implement and train the machine learning models (e.g., Decision Trees, Random Forests) for intrusion detection.  **Data Preprocessing Scripts:** Scripts for cleaning, normalizing, and preparing the datasets (e.g., NSL-KDD, CICIDS2017) for use in training and testing the models.  **Evaluation Report:** A report detailing the performance of the IDS, including metrics such as accuracy, precision, recall, and false positive rates, along with a comparison of different models tested. |

## Project Team/Roles:

*Who needs to be involved as members of the project team; and what will their roles be on the team? (Please indicate their Gannon student ID number.)*

| Who | User ID | Role on Project Team |
| --- | --- | --- |
|  |  | 1.Design and implement the machine learning models (e.g., Decision Trees, Random Forests) for intrusion detection.  2. Handle data preprocessing, model training, and evaluation of performance metrics (accuracy, precision, recall). |
|  |  | 1. Develop the overall system architecture, integrating the machine learning models into the IDS.  2. Write the source code, prepare documentation, and ensure proper functioning of the IDS with real-time monitoring and alert generation. |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |