

# CTMSCS SI Lab : AI

## ODD Paper

### 1. Network Login Attempt Analysis

#### **Scenario**

A cybersecurity analyst is monitoring **login attempts per hour** on a critical server to identify unusual behavior.

#### **Tasks**

Write a Python program using **NumPy, Pandas, and Matplotlib** to perform the following:

a) Use NumPy to generate a random array of size 10 representing the **number of login attempts per hour** and compute:

- Mean login attempts
- Standard deviation (to understand traffic variability)

b) Load a CSV file `login_attempts.csv` using Pandas and display:

- First 5 rows
- Column names
- Summary statistics

*(Sample columns: Hour, Successful\_Logins, Failed\_Logins)*

c) Plot a **bar chart** showing **Failed\_Logins per hour** to visually identify suspicious spikes.

► **Output:** High variance or spikes may indicate brute-force attacks.

### 2. Firewall Alert Trend Monitoring

#### **Scenario**

A SOC (Security Operations Centre) receives daily firewall alerts and wants to analyze alert patterns.

#### **Tasks**

a) Generate a NumPy array of size 10 representing **daily firewall alert counts** and calculate mean and standard deviation.

b) Load `firewall_logs.csv` and display:

- First 5 rows

- Column names
- Summary statistics

*(Sample columns: Date, Blocked\_IPs, Allowed\_IPs, Alerts\_Count)*

c) Plot a **line chart** showing **Alerts\_Count** over days.

→ **OUTPUT:** Sudden increases may indicate scanning or DDoS attempts.

### 3. Using a regression dataset, perform the following tasks:

a) **Dataset Selection** : California Housing dataset (sklearn)

Online dataset: <https://www.kaggle.com/datasets/camnugent/california-housing-prices>

b) **Data Preprocessing**

- Handle missing values
- Feature scaling
- Train-test split

c) **Feature Selection** (Apply any one techniques:)

- Correlation analysis
- SelectKBest (f\_regression)

d) **Regression Algorithms Implementation**

- Linear Regression
- Decision Tree Regressor
- Random Forest Regressor

e) **Model Evaluation**

- Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), R<sup>2</sup> Score

f) **Performance Comparison:** Compare all models and justify the best-performing regression model.