CTMSCS SI Lab : AI

EVEN Paper

1. Malware Detection Log Analysis

Scenario
An antivirus system records the number of detected malware instances each day.
Tasks
a) Create a NumPy array representing malware detections over 10 days and compute mean and standard deviation.

b) Load malware_logs.csv and display:

- First 5 rows
- Column names
- Summary statistics

(Sample columns: Date, Malware_Detected, Files_Scanned)

c) Plot a bar chart of Malware_Detected vs Date.

➡ OUTPUT: Helps identify outbreak days or ineffective endpoint protection.

2. User Account Security Monitoring

Scenario
An organization monitors failed login attempts to detect compromised accounts.

Tasks
a) Generate a NumPy array of size 10 representing failed login attempts per user and calculate mean and standard deviation.

b) Load user_security_logs.csv and display:

- First 5 rows
- Column names
- Summary statistics

(Sample columns: User_ID, Failed_Attempts, Successful_Attempts)

c) Plot a bar chart of Failed_Attempts per user.

➡ Cyber Insight: Users with abnormal failed attempts may indicate credential stuffing**.**

3. Develop a detection model using Artificial Neural Networks (ANN).

Email Spam Detection Dataset

https://www.kaggle.com/datasets/uciml/sms-spam-collection-dataset

Tasks:

a) Preprocess dataset and perform feature scaling
b) Design an **ANN architecture**:

- Input layer
- Hidden layers
- Output layer

c) Train the ANN model
d) **Visualization (Compulsory):**

- Training vs validation accuracy/loss curves
- Confusion matrix (if classification)

e) Evaluate model using:

- Accuracy
- Precision, Recall
- ROC-AUC

f) Compare ANN performance with any **Three traditional ML classifiers** using both **metrics and plots**.