Artificial Intelligence for Security, 2025/26

Group Assignment

(5/11/2025)

In this assignment you will analyse recent and large **cyber-security datasets**. Details are as follows:

- You must complete the assignment in **groups of 3 (minimum) or 4 (maximum)** students.
- Note that we encourage students to work together as a group as much as possible, rather than trying to divide up the tasks between the group members.
- The assignment involves creating a Python **notebook** to analyse the data as described below.
- Your notebook should be **self-explanatory**, with clear descriptions of the analysis performed and the conclusions drawn.
- The **assignment is due** on **Thursday the 18th of December** at 23:00 (via WeBeep).
- Each group will also need to briefly (in 10 minutes) **present their project.** – Please don't prepare presentation slides as we want to see your notebook.
- Presentations will take place during the practical session on **Friday the 19th of December**, from 14:15 to 19:15. All team members should be present in person (or online if previously agreed with the lecturer).

The assignment will be marked based on:

1. the appropriate and correct use of the methods applied,
2. the depth and insightfulness of the analysis,
3. the clarity of the description in the notebook, and
4. the quality of the presentation.

The assignment makes use of **large publicly available datasets containing REAL cybersecurity threats** listed below. Each group should **choose ONE** of the following datasets to work on. A **maximum of 3 groups** can work on the same dataset, so please register your interest to work on one them in this file (first come first served):

- https://docs.google.com/document/d/1GJQ21QER4UD6vcHKocyodfLr8DZVoCBTkLCAzhzZoUA/edit?usp=sharing
- If you are looking to form a group, you can add your name to the table at the bottom of the file.

Note that some of the **datasets can be very large** and thus you may need to **sample a random subset** of the data in order to perform certain types of analysis.

1. Dataset of **Network activity of IoT devices**
   https://www.stratosphereips.org/datasets-iot23

2. Dataset of **Industrial Control System (ICS) Cyber Attacks** – Dataset 1: Power systems dataset
   https://sites.google.com/a/uah.edu/tommy-morris-uah/ics-data-sets

3. Dataset of **Bitcoin transactions**
   https://github.com/git-disl/EllipticPlusPlus

4. Dataset of **Android malware** apks
   https://data.mendeley.com/datasets/rvjptkrc34/2

5. Dataset of **Jailbreak attacks on LLMs**
   https://huggingface.co/datasets/qualifire/prompt-injections-benchmark/viewer/default/test?row=85&views%5B%5D=test

6. Dataset of **Intrusions in Unmanned Aerial Vehicle (UAV) networks**
   https://zenodo.org/records/15125851

Tasks:

The assignment requires you to visualize various aspects of the data and build supervised and unsupervised models with it. The tasks listed below are **only suggestions** for some types of analysis that you could perform. Note that you need not perform all of the tasks, and that some of the tasks **may not be applicable** for certain datasets. Moreover, you are strongly encouraged to **be creative** and **investigate new tasks** that are not listed below. Remember to document and discuss all your insights in your notebook!

A) Compute summary statistics:

Answer questions like:

- How many rows and how many columns are there in the data?
- What are the names and datatypes in each column?
- Do any columns contain categorical values? If so, what values can they take? Which value is most frequent? For numeric columns calculate summary statistics (min/max, mean/median, standard deviation, etc.). What do these values tell you?
- Is there a target variable that needs to be predicted? How frequent/balanced are its values?

B) Visualise the data:

Visualise the distributions contained in the various columns:

- Create bar plots and histograms for various features. Are there any obvious outliers?
- If there is a categorical target variable, plot summary statistics mentioned above for different values of the target variable. Are there differences in the statistics across the groups? What does that mean?
- Plot a heatmap of the correlations. Which features are correlated with one another? Do the correlations make sense?

C) Perform supervised learning:

Train various classifiers using subsets of the features available:

- Do any the features contain missing values. If so, how should you handle them?
- Which features are important for the classification?
- Which classifiers give the best performance?
- What evaluation methods are most appropriate for this problem?

D) Perform unsupervised learning:

Run clustering algorithms on the data:

- Try to plot the clusters in 2 dimensions by either selecting pairs of attributes or by performing dimensionality reduction.
- Do the clusters look reasonable?
- Can you understand the groups in some way? Do the clusters correspond to intuitive subsets of the data? Could they be useful for defining features for a classifier?
- Try running the clustering separately for different values of the target variable. Does that produce more reasonable groups?

E) Perform anomaly detection:

Run univariate and multivariate anomaly detection techniques:

- Do the results tell you anything interesting about the data?
- Can you interpret the outliers or are they noise?

F) POTENTIAL further investigations:

- Make use of some deep learning techniques to process the dataset
- Look for other related datasets online and compare performance on those datasets.