COMP8325 Applications of Artificial Intelligence for Cyber Security

Group Project Description
Total Marks: 20
Weighting: 20%

MACQUARIE
University

DEADLINE:

- Group Formation Due: Friday, 5 April, 11:55 PM (End of Week 7).
- Presentation: In person, Tuesday and Wednesday, 28-29 May, 10:00 AM to 12:00 PM (During Week 13 lecture and workshop sessions).
- Presentation Slides + Final Report + Source Code: Friday, 31 May, 05:00 PM (End of Week 13).

LEARNING OUTCOME

By completing this assignment, you should demonstrate your ability to:

- Understand and detect abnormal patterns in a variety of real-world datasets for cyber security
- Perform data pre-processing, data exploration, and feature selection on various types and volumes of data for different cyber security applications
- Perform machine learning training and evaluation for cyber security applications
- Understand the security threats to machine learning systems deployed for cyber security
- Communicate professionally in written and oral form to a range of audiences

SUBMISSION

This is a group project (of at most 6 students). Submission is required per group. Each group needs to submit three types of files:

(1) Source code: Three Jupyter Notebook files (i.e., with the extension, `.ipynb`) with the implementation and analysis of the Data Analysis Task.
   - See Section 1.1 for details of this task
   - Since this is a collaborative project, you can use tools such as Github which allows your team to make collaborative edits. See for example: https://reproducible-science-curriculum.github.io/sharing-RR-Jupyter/01-sharing-github/

(2) Final report: A document summarising the Research Task.
   - See Section 1.2 for details of this task.

(3) Presentation slides on the Research Task.
   - The slides themselves will be presented in person by the group.
   - The presentation time is 12 minutes per group

All submissions are through iLearn via Turnitin submission links. At the beginning of each file, please write down your group name.

MARKING CRITERIA

Table 1 and Table 2 show, respectively, a comprehensive marking rubric for the Data Analysis and Research tracks of the group project. Here is a brief version of the criteria that will be used for evaluating the group project's:

- The efficacy of organization, presentation, and readability demonstrated in the Jupyter notebooks and the final report.
- The production of high-quality, easily comprehensible source code.
- Demonstrated critical thinking skills and a profound comprehension of the findings derived from the data analysis task.
- Execution of a well-conducted and sufficiently deep literature review, showcasing an understanding of related literature.
- Effective presentation skills exhibited by the team, indicative of professional communication capabilities and seamless team coordination.

LATE SUBMISSIONS

- Late submissions follow the policies specified in the Unit Guideline.
- If you have a legitimate reason for submitting late, discuss this with the convenor before the assignment due date.

1. Project Details

You are required to work in a group of at most five students to

(1) Conduct *anomaly detection* on real-world data sets in various cyber security applications (see Section 1.1), and
(2) Review recent research on the topics related to machine learning and security (see Section 1.2).

1.1. **Data Analysis Task (10 marks).** You are required to conduct *anomaly detection* on three data sets listed below. For each dataset apply at least *two* anomaly detection algorithms. Unlike the assignments, we have not performed any pre-processing on these data sets.

(1) **Network Intrusion Detection.** This is a dataset of raw network packets. See https://research.unsw.edu.au/projects/unsw-nb15-dataset for a detailed description of the dataset. On the website, download the datasets:
   - `HERE -> UNSW-NB14-CSV Files -> NUSW-NB15_features.csv`
   - `HERE -> UNSW-NB14-CSV Files -> UNSW-NB15_1.csv`

   The first one is the feature names file, and the second one is the data. The packets are labeled normal and attack packets. You can train and evaluate the accuracy of the anomaly detection algorithms on an appropriate split between train and test sets.
(2) **Malicious URL Detection.** This is a dataset of malicious URLs such as spam, phishing, and exploits. See http://www.sysnet.ucsd.edu/projects/url/ for a detailed description of the dataset. On the website, download the datasets:
   - `URL Data Set (Matlab)`

Once you download this dataset you should have a feature names file and the data. The URLs are labeled benign and malicious. You can train and evaluate the accuracy of the anomaly detection algorithms on an appropriate split between train and test sets.

(3) **Probing/Port Scan Detection.** This is a dataset of probing attacks (port scans) performed through various tools. See https://github.com/gubertoli/ProbingDataset for a detailed description of the dataset. On the website, download the datasets:

- `mawilab -> data -> normal_dataset.csv`
- `testbed -> data -> malicious_dataset.csv`

The first dataset consists of normal traffic, while the second dataset comprises attack traffic. Both datasets contain feature names as the first row. You should merge the datasets into one and then train and evaluate the accuracy of the anomaly detection algorithms on an appropriate split between train and test sets.

Deliverables

Three Jupyter notebooks, one for each dataset. Each notebook should describe:

- Data pruning/preprocessing
- Model training
- Model evaluation using appropriate performance metrics
- A descriptive analysis and comparison of the relative performance of the anomaly detection algorithms

1.2. Research Task (10 marks).

- You are required to conduct a literature review on a topic related to security/privacy attacks on machine learning. Please see the list of topics below.
- Find and study at least five research-based publications (journal articles, conference proceedings, or books) related to your chosen topic.
- Document your findings as follows: (a) a brief introduction to the attack and its potential impact, (b) a literature review on the evolution of the attack, and (c) a literature review on mitigation strategies.
- You can utilize Google Scholar or the university online library to search for related research publications. Additionally, you should consider the quality of the publication venue when selecting articles. To further assist you in conducting the literature review, please refer to 'HowToDoResearch.pdf' for details.

Topics

Choose from one of the following topics.

- Adversarial machine learning
- Membership inference attacks in machine learning
- Model stealing attacks in machine learning
- Poisoning attacks in machine learning

See the excellent blog post by Elie Bursztein on an overview of these topics: https://elie.net/blog/ai/attacks-against-machine-learning-an-overview/. The blog post itself contains some links that you can use to start your review of your chosen topic.

Deliverables

Deliverables per group for this task are:

- A final report of at least two pages introducing the attack, its variations, and mitigation strategies.
- Presentation slides about your report.
- A 12-minute in-class presentation. Each group member has to present at least 2 minutes per person.

TABLE 1. Marking rubric for Data Analysis Tasks.

Marks	**Marking Rubric for Task 1**
0.15	Accomplishment of all required tasks
0.15	Well-organised, presented, and readable (Jupyter notebook file)
0.6	Discussion of the selected ML pipeline/algorithms.
0.25	The quality of source code, especially the ease of using the code to perform prediction/testing on reserved data.
0.35	Loading Data
0.5	Pre-processing
0.5	Training
0.5	Evaluation lacks discussion
3	Subtotal
Marks	**Marking Rubric for Task 2**
0.15	Accomplishment of all required tasks
0.15	Well-organised, presented, and readable (Jupyter notebook file)
0.75	Discussion of the selected ML pipeline/algorithms.
0.25	The quality of source code, especially the ease of using the code to perform prediction/testing on reserved data.
0.35	Loading Data
0.6	Pre-processing
0.5	Training
0.75	Evaluation lacks discussion
3.5	Subtotal
Marks	**Marking Rubric for Task 3**
0.15	Accomplishment of all required tasks
0.15	Well-organised, presented, and readable (Jupyter notebook file)
0.75	Discussion of the selected ML pipeline/algorithms.
0.25	The quality of source code, especially the ease of using the code to perform prediction/testing on reserved data.
0.35	Loading Data
0.6	Pre-processing
0.5	Training
0.75	Evaluation lacks discussion
3.5	Subtotal

TABLE 2. Marking rubric for Research Track: Report and Presentation.

Marks	**Marking Rubric for Report**
1	The relevancy of content
1	Analysis of arguments and techniques in literature
1	Discussion of the pros and cons of research/literature
1	Organization of the content and arguments
0.75	Presentation of the content and formal writing
0.25	References
5	Subtotal
Marks	**Marking Rubric for Report**
2	Relevant content discussing the topic
1	Professional communication
1	Coordination among group members
1	Visual ads and presentation
0.75	Presentation of the content and formal writing
5	Subtotal
10	**Grant total**