CSIS

Coursework Specification

Assignment Details	
Module Name	Applied Machine Learning
Module Leader	Dr Paul Yoo
Set by	Dr Paul Yoo
Piece number	1 of 1
Assignment Type	Individual Practical Project
Contribution	30%
Hand-in point	VLE – Turnitin
Due Date[1]	Friday, 15 January 2021 – a copy of your report must be uploaded to the VLE by 1600. A copy of your code must be submitted to your group TA. See below for details. https://www.dcs.bbk.ac.uk/intranet/index.php/Coursework_Deadlines_Autumn_2020
Marks will be Returned on	1600 on Monday, 15 February 2021

[1] If you anticipate an issue with meeting the deadline due to exceptional circumstances, or have missed the deadline, please see http://www.bbk.ac.uk/management/current-students/mitigating-circumstances for details on the extension, deferral and suspension processes.

Assessment Requirements

Cyber Threat Detection Competition using Machine Learning

Software to detect cyber threats and protects a computer system (e.g. IoT, AV and UAV) from various cyber-attacks. The task of cyber threat detector learning is to build a predictive model (i.e. a machine-learning classifier) capable of distinguishing between “intrusive” traffic, called threats or attacks, and “good” normal traffic.

The Aegean WiFi Intrusion/threat Dataset (AWID) project was prepared and managed by George Mason University and University of the Aegean. The objective was to survey and evaluate research in cyber threat detection. A standard set of data to be audited, which includes real traces of both normal and intrusive 802.11 traffic with a wide variety of cyber threats/intrusions simulated in a physical lab which realistically emulates a typical SOHO infrastructure, was provided.

Details of the dataset can be found in the page below.

http://icsdweb.aegean.gr/awid/draft-Intrusion-Detection-in-802-11-Networks-Empirical-Evaluation-of-Threats-and-a-Public-Dataset.pdf

AWID dataset categorises the attacks according to the methodology of execution. Attacks that have similar patterns of expression fall under one of the groups:

a. injection attacks,
b. flooding attacks,
c. impersonation attacks,
d. passive attacks.

We will use the reduced CLS portion of the AWID dataset because it serves as a useful starting point for preliminary research, and affords a baseline against which the new built models can be compared with the state-of-the-art Deep Feature Extraction and Selection (D-FES) method[2], and other methods using the same reduced CLS portion[3,4,5].

[2]ME Aminanto, R Choi, HC Tanuwidjaja, PD Yoo and K Kim (2018) Deep abstraction and weighted feature selection for Wi-Fi impersonation detection, IEEE Transactions on Information Forensics and Security, 13(3), 621–636.
[3]ME Aminanto and K Kim (2017) Detecting impersonation attack in WiFi networks using deep learning approach, Information Security Applications 17th International Workshop. Jeju Island, South Korea, 25-27 August 2016, 136–147.
[4]C Kolias, G Kambourakis, A Stavrou and S Gritzalis (2016a) Intrusion detection in 802.11 networks : empirical evaluation of threats and a public dataset, IEEE Communication Surveys and Tutorials, 18(1), 184–208.
[5]Lee SJ, Yoo PD, Asyhari AT, Jhi Y, Chermak L, Yeun CY, Taha K. IMPACT: Impersonation attack detection via edge computing using deep autoencoder and feature abstraction. IEEE Access. 2020 Apr 2;8:65520-9.

Although flooding and injection attack signatures are also available within the AWID-CLS dataset, impersonation attacks were our focus as Hirte, Honeypot and EvilTwin impersonation attacks have previously been identified as the most severe threats to a network[5] and have been the focus of earlier research[6,7]. As a result, detecting impersonation attacks as the focus of our work will allow us to directly compare the performance of our model to others, a key weakness within the current body of machine-learning-based IoT IDS research.

A complete listing of the set of features defined for the connection records, the relevant papers, the dataset description is given in the link below.

http://icsdweb.aegean.gr/awid/

The dataset for your project is available on the module page of the VLE – please do not use the one in the above link.

Aim of the Assessment

The aim of this assessment is to provide a hands-on, practical, assessment of your machine learning skills for cyber threat detection application.

Description of Task to be Completed

Your task is to build a predictive model (i.e. a machine learning classifier) capable of distinguishing between “intrusive” traffic, called threats/intrusions or attacks, and “good” normal traffic. There is a research element in this coursework (e.g. creating additional features).

- **Downloading dataset.** The dataset for this project is available on VLE. You need to use `train_imperson_without4n7_balanced_data.csv` for training and *`test_imperson_without4n7_balanced_data.csv`* for testing. The first row of each dataset gives variable numbers (this may need to be removed). The original dataset

---

[5]C Kolias, G Kambourakis, A Stavrou and S Gritzalis (2016a) Intrusion detection in 802.11 networks : empirical evaluation of threats and a public dataset, IEEE Communication Surveys and Tutorials, 18(1), 184–208.
[6]ME Aminanto, R Choi, HC Tanuwidjaja, PD Yoo and K Kim (2018) Deep abstraction and weighted feature selection for Wi-Fi impersonation detection, IEEE Transactions on Information Forensics and Security, 13(3), 621–636.
[7]Parker L, Yoo P, Asyhari T, Chermak L, Jhi Y, Taha K. DEMISe: interpretable deep extraction and mutual information selection techniques for IoT intrusion detection. InARES'19 Proceedings of the 14th International Conference on Availability, Reliability and Security 2019 Aug 26. ACM.

has 154 input variables and 1 target variable however two of the input variables numbered 4 and 7 (*frame.time_epoch* and *frame.time_relative)* have been removed from both datasets as they provide temporal information which may cause unfair prediction. The training set has 97044 observations while testing set has 40158 observations.

- **Constructing and Selecting Features.**
    - Create additional features using a representation learner (e.g. Autoencoder, VAEs, SAEs, GAN etc). The latent space representation, simply a representation of compressed data by a representation leaner, may contain important information needed to represent original data point.
    - Combine the additional features with original dataset. The combined dataset should have 154 (original features) + additional features. For example, if you created 10 features using SAE, the new combined dataset must have 164 input features + 1 target.
    - Apply feature selection techniques (e.g. filter, wrapper and embedded) and see if any of the additional features created by a representation learner are selected.
- **Building ML algorithms.**
    - Select candidate algorithms. Discuss the selection strategies for the candidate algorithms.
    - Finding the best configuration for these hyperparameters in such a high dimensional space is not a trivial challenge. Consider the model design components (e.g. no of layers, no of units per layer, loss function, activations, optimisers, dropout layer etc) as well as the hyperparameters (e.g. learning rate, dropout rate, batch size etc). Perform model-specific optimisations and iteratively debug model as complexity is added. Discuss the selection strategies for searching for the best configuration (e.g. trial and error, grid search, random search, Bayesian optimisation etc).
- **Evaluating model and analysing the results.**
    - Evaluate the classification performance (e.g. accuracy, detection rate, false alarm, type II error, MCC and TBM (time has taken to build model) and TTM (time has taken to test model) – go beyond these measures if necessary) of the selected models on the test data and interpret the results.
    - Discuss general model trade-offs (accuracy vs speed vs interpretability) of the chosen models considering a particular application (e.g. IoT, autonomous vehicle, etc) and propose two models (e.g. features selected and ML classifiers) and provide justification of choice.

Deliverables Required and Submission Information

You must produce a report of 2,000 words (±10%). The cover page must show your name, student number, the wordcount of your report (excluding appendices), module title, assignment type and assignment title. There should be some substantial tables and figures (make a good use of appendix) that help to cram all your information into the word count.

The report must be presented using either Arial 10 point or Times New Roman 11 point font for the main body of the text and 1.5 line spacing. Pages must have a minimum of 2.54 (1 inch) margins, i.e. MS Word 'normal' margins. IEEE referencing must be used, for guidance see: https://ieeeauthorcenter.ieee.org/wp-content/uploads/IEEE-Reference-Guide.pdf

The page number should be in the footer. You are strongly recommended to upload it in PDF format where possible, especially if including tables or figures. The submitted file title must be in the following format:

Student Number SurnameInitial ModuleName Assignment type e.g. 654321 BloggsJ AML GPP. Your code must also be submitted along the report (email your group TA).

Estimated Time to Complete

There will be time that is allocated for working on your project in Week 9. However, it is your responsibility to allocate an appropriate amount of time to this piece of work.

Marking Scheme

Marks will be awarded in the following areas	%Weighting	Marking Descriptors					
		Excellent 80-100%	Very Good 70-79%	Good 60-69%	Satisfactory 50-59%	Poor 40-49%	Very Poor 0-39%
Constructing and Selecting features	30%	Strong justifications. Full consideration of various techniques in three different categories. Strong evidence of wide reading and research.	Very good justifications for selection. Application of feature selection techniques are effective showing insight and creativity. Evidence of wider reading and research.	Good justifications. Appreciation of principles. Evidence of consideration of alternatives and judgement in decision. Evidence of wider reading and research.	Some weaknesses in the process and can apply these reasonably well with fair justification.	The process are weak and little evidence is provided but just workable. Justification for decisions is limited.	Little or no attempt to apply appropriate feature selection techniques.
Building ML algorithms	30%	Full consideration has been given to options for algorithms. Entirely appropriate choices made and expertly applied. Provision of baselines. Strong selection strategies.	Very good selection of algorithms and procedures - competently applied. Justification for selection is sound. Very good selection strategies. Evidence of wider reading.	Good judgement in selection and application of algorithm and procedures. Provision of baseline. Evidence of wider reading.	Reasonable selection and application of algorithm and procedures.	Poor but acceptable selection. Decisions may be based on student's convenience.	Appropriate analysis or judgement severely lacking or not provided.
Evaluating model and analyzing the results	30%	Strong justification on evaluation methods. Selected models are fairly evaluated. All evaluation measures are correctly calculated. Evidence of wider reading and choice of extra measures. Quality of interpretation is very high and is based on several insightfully chosen sources.	Good justification on evaluation methods. Selected models are fairly evaluated. All evaluation measures are correctly calculated. Evidence of wider reading. Interpretation is very good and is based on comparison with some well-chosen sources, but some points could be developed.	Selected evaluation methods are appropriate to the objectives. Good justification for selection. All evaluations are provided. Interpretation is good, and based on comparison of some relevant sources, but the sources used could be extended.	Reasonable justification for selection showing some awareness of appropriate principles. Some calculations are incorrect or unfairly evaluated. Interpretation is reasonable and very few sources are used.	Some of the selected method are appropriate, with limited justification. Some calculations are incorrect or unfairly evaluated. Little interpretation on the experimental results. Lacks depth. Poor results.	An insufficient awareness of principles with very weak or no justification. Incorrect calculations. Unfair evaluation. No or trivial interpretation on the results.

Marks will be awarded in the following areas	%Weighting	Marking Descriptors					
		Excellent 80-100%	Very Good 70-79%	Good 60-69%	Satisfactory 50-59%	Poor 40-49%	Very Poor 0-39%
Report Documentation	10%	Organisation of work is of a very high standard, likely to be highly stimulating, and at the limits of what may be expected at postgraduate level. Work is of a standard publishable in a refereed journal.	Documentation is very well ordered, concise and coherent. Excellent use of appendices and illustrations.	Organisation of work is likely to show few mistakes/limitations. Very good use of appendices and illustrations.	There is an overall structure evident but does not offer strong flow and progression. Appendices and illustrations mainly used well.	Structure of work is weak or inconsistent. Only the main points are logically organised/linked. Quite good use of appendices and illustrations.	Unstructured and/or incoherent. Illustrations are very poorly presented. Appendices are very poorly presented.