Student Number:

The University of Melbourne
Semester 2 Assessment 2021

School of Computing and Information Systems
COMP90073 Security Analytics

Reading Time: 15 minutes.
Writing Time: 2 hours.
This paper has 18 pages including this cover page.
Common Content Papers: None

Authorised Materials: Lecture notes, books, computer, on-line material.

Instructions to Students:

- This paper counts for 60% of your final grade, and is worth 60 marks in total.

- There are 22 questions, with marks as indicated.

- Answer all the questions on the exam paper if possible, and then upload the completed exam paper containing your solutions. If you are unable to print the exam paper or electronically edit the exam paper, you may write on your own blank paper and then upload images of your written answers.

- You may upload your exam answers multiple times if you need to revise an answer at any time during the exam.

- You must not communicate with other students or seek assistance from anyone else taking whilst taking this exam, e.g. using messaging, chat rooms, email, telephone or face-to-face. Also, you must not assist anyone else taking the exam. You must not post answers to the questions or discussion of the questions online. Failure to comply with these instructions may be considered as academic misconduct.

- You are free to use the course materials and your laptop/PC in this exam but note that there is a 2-hour time window for the exam hence you should be mindful of the time spent using such resources.

- Answer the questions as clearly and precisely as you can.

- Your writing should be clear. Unreadable answers will be deemed wrong. Excessively long answers or irrelevant information may be penalised.

- For numerical methods, marks will be given for applying the correct method.

Library: This paper may not be reproduced or held by the Baillieu Library.

Section A: Short Answer Questions (Use your own words to provide a short explanation to each question) **[10 marks in total]**

1. What is the pattern for a typical UDP flood attack, and why? [1 marks]

   Answer:

2. What is the main difference between push-based and pull-based botnet propagation methods? [1 marks]

   Answer:

Assignment Project Exam Help

https://tutorcs.com

3. In which of the following attacks should **not** IP spoofing be used, and why?
   [1 marks]

WeChat: cstutorcs

   (a) TCP scan
   (b) DNS flood attack
   (c) Ping of death attack
   (d) DNS amplification attack

   Answer:

4. Given a data set with some missing feature values. What would be the best anomaly detection algorithm for this data? Justify your answer. [1 marks]

   Answer:

5. As a network manager how can you use contrast mining to identify anomalous patterns in network traffic? [1 marks]

Answer:

6. Why One-Class Support Vector Machine (OCSVM) is also called ν-SVM (what is the role of ν)? And how one should set it? [1 marks]

Answer:

7. Explain transferability in adversarial machine learning. [1 marks]

Answer:

8. Adversarial training is an effective defence method against adversarial attacks. How is it different from the normal training process? [1 marks]

Answer:

9. How is Deep Q-Network (DQN) different from classical Q-network? List three of them. [1 marks]

Answer:

10. In adversarial attacks against reinforcement learning agents, the attacker does not need to perturb the observed state at each time step. How should they decide whether to poison an observed state s_t or not? [1 marks]

Answer:

Section B: Method and calculation Questions   **[30 marks in total]**

11. You are a security expert working for ElecX Automobile Factory. Your responsibility is to secure factory's IT systems, in particular, Payroll, CRM (Customer Relationship Management System) and brochure hosting site.

    (a) How do you measure the confidentiality of information you need to protect, and how can it be applied to information in those three systems?   [2 marks]

    Answer:

    Assignment Project Exam Help

    https://tutorcs.com

    (b) "Segregation of duties" is a sample control for integrity in the CIA triad. Briefly explain its definition and the motivation behind it.

    WeChat: cstutorcs   [2 marks]

    Answer:

12. One recently disclosed critical vulnerability on Bankstr Fintech Group's online share trading platform allows an attacker to gain unauthorised access to customers share portfolio. Should it be exploited, this will cause Major impact to Bankstr Fintech Group. The detailed metrics and ratings of the exploit are tabled below.

Metrics	Rating
Skill (High skill level required → low or no skill required)	2
Ease of Access (very difficult to do → very simple to do)	5
Incentive (high incentive → Low incentive)	5
Resource (requires expensive or rare equipment → no resources required)	4

(a) What is the likelihood score?   [1 marks]

Answer:

(b) What is the risk level?   [1 marks]

Answer:

(c) What is the recommended action, and why? Choose the appropriate answer, and briefly explain your choice.

  i. Immediate action required to mitigate the risk or decide to not proceed

  ii. Action should be taken to compensate for the risk

  iii. Action should be taken to monitor the risk

                                                                [1 marks]

Answer:

13. The GameFest company designed a new version game, whose Intellectual Property is worth $1,500,000. The exposure factor is 70%, and the annualised rate of occurrence is 20%.

(a) What's the single loss expectancy?   [1 marks]

Answer:

(b) What's the annualised loss expectancy?   [1 marks]

Answer:

14. The table below shows a list of items, use FP-growth to identify frequent patterns with Min_sup=3. Your work should include FP-tree, Conditional pattern base, Conditional FP-tree, and Frequent patterns.   [3 marks]

TID	List of items
T100	{f, a, b, c, d, g, i, m, p}
T200	{a, b, c, k, l, m, s}
T300	{b, d, h, j, o, w}
T400	{b, c, k, m, p}
T500	{a, f, c, e, l, p, m, n}

Answer:

Assignment Project Exam Help

https://tutorcs.com

WeChat: cstutorcs

15. To maintain its efficiency, incremental LOF (iLOF) is required to delete historical samples, however, this impacts its performance if such samples re-occur. Discuss which phase of Memory Efficient Incremental LOF (MiLOF) addresses this issue? [1 marks]

Answer:

16. Which of the following is/are not true about DBSCAN clustering algorithm? Justify your answer. [2 marks]

    (a) For data points to be in a cluster, they must be in a distance threshold to a core point

    (b) It has strong assumptions for the distribution of data points in dataspace

    (c) It has substantially high time complexity of order $O(n^3)$

    (d) It does not require prior knowledge of the no. of desired clusters

    (e) It is robust to outliers

Answer:

17. Why in OCSVM we would like to maximise the distance of the decision boundary from the origin? And how this algorithm archives that? [1 marks]

Answer:

18. The loss function of Variational autoencoder (VAE) includes two parts, a reconstruction loss and a regulariser,

$$l(\theta, \phi) = E_{q_\phi(h|x)}[\log p_\theta(x|h) - D_{KL}(q_\phi(h|x)||p_\theta(h)))]$$

What is the role of each part? Provide an example scenario where one should use VAE over a basic autoencoder for anomaly detection, and discuss how VAE can generate better results. [2 marks]

Answer:

Assignment Project Exam Help

https://tutorcs.com

19. In Task 1 of Assignment 2, we asked you to train an anomaly detection algorithm on extracted features from network traffic, and gave you a training, a test, and a validation set. To address this task, one of your classmates, Sam, takes the following steps: cstutorcs

1. Sam starts by fitting a PCA (n_component = 4) on the validation set with all the 15 features (including stream ID), and calls the output model as "PCA_{fitted}".

2. Then, Sam applies the PCA to the validation set, denotes the reduced dataset (processed by PCA) as "$Data_{val\_PCA}$".

3. Afterwards, Sam trains OCSVM on $Data_{val\_PCA}$, and fine tunes the parameters to get the highest accuracy.

4. Finally, Sam extracts features from the training and test datasets by applying the PCA_{fitted} model, and applies OCSVM to both data sets.

Sam finds the False Positive (FP) rate is too high for the trained OCSVM model. Can you give some suggestions how effectively Sam can reduce the FP rate (while this might slightly affect the True Positive (TP) rate)? [2 marks]

Answer:

20. A binary linear Support Vector Machine (SVM) model (f) classifies input x using the following: $f(x) = w \cdot x + b$, *i.e.,* if $w \cdot x + b > 0$, x is classified into the positive class; otherwise, it is classified into the negative class. As demonstrated in Figure 1, in order to generate an adversarial sample x' against f for input x, one option is to perturb x in a direction orthogonal to the decision boundary hyperplane.
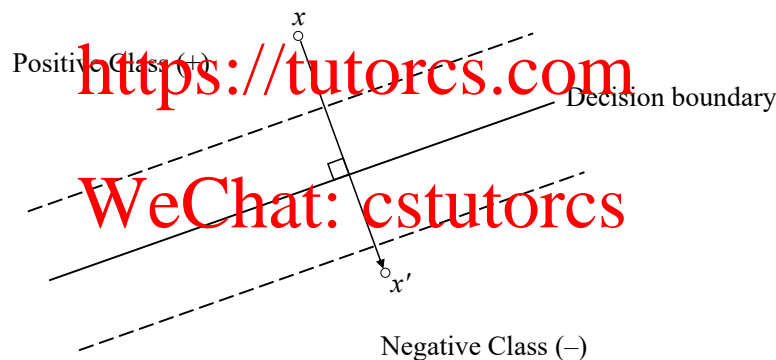

Figure 1: Generating an adversarial sample against a binary linear SVM classifier by moving the original input in a direction orthogonal to the decision boundary.

(a) If the attacker decides to follow a similar approach to the iterative gradient sign method, then in each iteration: $x' = clip_\epsilon(x - \alpha \cdot \underline{\hspace{1cm}})$, where α is the step size, and $clip_\epsilon$ is to make sure that x' is in the ϵ-neighbourhood of x, *i.e.,* in each dimension, the difference between the values of x and x' is not larger than ϵ. [1 marks]

Answer:

(b) If the attacker decides to replace the above gradient sign with normalised gradient, then in each iteration: $x' = clip_\epsilon(x - \alpha \cdot \underline{\hspace{2cm}})$.

[1 marks]

Answer:

(c) Suppose that $w = [3 \quad 4]$, $b = 1$, and $x = [x_1 \quad x_2]^T$, i.e., the input x is two dimensional. Generate an adversarial sample for point $(3, -1)$ using the method in the last step (b), where the gradient is normalised with the Euclidean norm—the square root of the sum of the squares of all elements. Specifically, the parameters are: (1) $\alpha = 1$, (2) $\epsilon = 1$. Note that both the intermediate and final results need to be clipped if necessary.

[4 marks]

Assignment Project Exam Help

Answer:

https://tutorcs.com

WeChat: cstutorcs

21. In standard adversarial training, an adversarial sample is created for each training instance, and all the adversarial variants are considered equally important. However, as shown in Figure 2, some training data are geometrically far away from(/close to) the decision boundary, and their adversarial samples are hard(/easy) to be misclassified. Therefore, to further improve the effectiveness of adversarial training, data should be treated differently: a larger(/smaller) weight should be assigned to the data point that has a smaller(/larger) distance to the decision boundary.
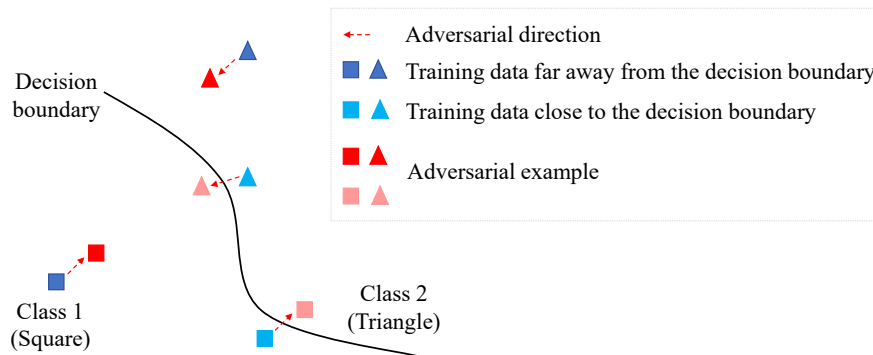
Figure 2: Different geometrical locations of adversarial examples.

To estimate the distance, one option is using the margin of the classifier (f): for a data point (x_i, y_i), its margin is the difference between the probabilities of the correct label y_i and the incorrect label, $margin(f, x_i, y_i) = p(f(x_i) = y_i) - p(f(x_i) \neq y_i) = 2 \cdot p(f(x_i) = y_i) - 1$ (to make it simple, here we only consider a binary classification problem). A large(/small) margin normally correspond to a data point far away from(/close) the decision boundary.

(a) Can you think of another way for estimating the distance between a data point and the decision boundary? (Hint: in Curriculum Adversarial Training, how is the attack strength measured?) [2 marks]

Answer:

(b) Suppose that the distance between (x_i, y_i) and the decision boundary is $d(x_i, y_i)$, how to redefine the objective function of adversarial training? (The objective function of the standard adversarial training is: $\min_\theta \mathbb{E}_{(x_i, y_i) \sim \mathcal{D}}[\max_{\delta_i} L(x_i + \delta_i, y_i; \theta)]$, where θ is the parameter of f, $\mathcal{D}$ is a certain distribution, δ_i is the adversarial perturbation for x_i, and L is the loss function) [2 marks]

Answer:

Assignment Project Exam Help

https://tutorcs.com

WeChat: cstutorcs

Section C: Application Question [**20 marks in total**]

In this section you are asked to demonstrate that you have gained a high-level understanding of the methods and algorithms covered in this subject, and can apply that understanding. These questions may require significantly more thought than those in Sections A and B, and should be attempted only after having completed the earlier sections.

22. R-Corp is an aircraft engine manufacturer. Marcus belongs to a hacking group, who is very interested in the latest turbofan engine developed by R-Corp. Marcus learned from LinkedIn that Dr. Zander is the current Chief Scientist of R-Corp. Marcus then crafted an email pretending from an acquaintance of Dr. Zander with a malware attached. Note that Marcus developed the malware by leveraging a recent Zero day vulnerability. Dr. Zander was lured to click on the malware in the email, which successfully exploited the targeted vulnerability on Dr. Zander's system, and then installed a backdoor. This gave Marcus the remote control of Dr. Zander's computer. After that, Marcus used a compromised server (C2 server) to send commands to maintain the control of Dr. Zander's computer. One night, Marcus started to upload key research documents from Dr. Zander's computer to his cloud storage folder.

Index	Name
CKC1	Reconnaissance
CKC2	Weaponization
CKC3	Delivery
CKC4	Exploitation
CKC5	Installation
CKC6	Command & Control (C2)
CKC7	Actions on Objectives

(a) Map the attack activities to Cyber Kill Chain (CKC) shown in the above table. For example, CKC1 – Marcus gathered information of Dr. Zander via LinkedIn. [3 marks]

CKC2 – _____

CKC3 – _____

CKC4 – _____

CKC5 – _____

CKC6 – _____

CKC7 – _____

(b) Suppose that R-Corp has deployed the following security systems:

– Gateway control: Web Proxy (Web Security System), Email Security System
– Network control: IPS (Intrusion Prevention System)
– Endpoint control: HIPS (Host based IPS)

Explain which attack step(s) can be potentially detected by each of these systems using Cyber Kill Chain? For example, *Email Security System: CKC3 — detect malware delivery via phishing email.*
Note that each system may detect multiple attacks. [4 marks]

Answer:

Assignment Project Exam Help

https://tutorcs.com

WeChat: cstutorcs

(c) After this security incident, R-Corp starts investing in security control technologies. You are working as a senior security consultant and supervising a new graduate Sam who is taking an internship at R-Corp. As the first task, Sam was asked to build an efficient anomaly-detection based IDS for a data set that has about 2000 features and 500 records. Given the description of the data, Sam decides to use a deep autoencoder, however, the performance of the model is very poor. Why do you think the autoencoder can't perform well on this data set? Among the different anomaly detection methods covered in the subject, what would be the best choice for this problem? Justify why your answer. [2 marks]

Answer:

(d) After choosing an appropriate model, Sam trains it and reports $Accuracy = 98\%$ on a validation set. What do you think of this results, and how well do you expect the model can identify anomalies (suspicious patterns) once deployed? [2 marks]

Answer:

(e) Sam manages to resolve the above issue, but still the model has difficulties identifying botnets in the network. You suspect that this is due to botnets behaviour that appear as collective anomalies. If that's the case, what algorithm should Sam uses? And how it should be implemented to fit with the description of the data? [2 marks]

Answer:

(f) The results has improved but still not satisfying. Now, given the connections between the nodes (i.e., adjacency matrix A) are available, you suggest Sam to use this information. Why such information can

improve the performance of the IDS?

Sam implements the random walk algorithm. However, despite the small size of data, the algorithm takes a long time to train. How would you suggest Sam to mitigate this issue? What is the intuition of your solution?   [2 marks]

Answer:

Now suppose that R-Corp has deployed a machine learning based system ($\mathcal{F}$) to detect malware: each software to be tested is represented by a d-dimensional binary feature vector $x = [x_1 \ x_2 \ ... \ x_i \ ... \ x_d]^T$, $x_i \in \{0, 1\}$, where $x_i = 1$ (bit 0) means that the software has (or does not have) the ith feature. The detection system calculates two scores for each vector: $\mathcal{F}_0(x)$ and $\mathcal{F}_1(x)$, which represent the probabilities of the corresponding software being benign and malicious, respectively. If $\mathcal{F}_0(x) > \mathcal{F}_1(x)$ (or $\mathcal{F}_0(x) \leq \mathcal{F}_1(x)$), the software is classified as benign (or malicious).

Suppose that the attacker only knows the above information about the malware detection system ($\mathcal{F}$), but does not know which features are selected, what machine learning algorithm is used, or the architecture and the parameters of the model, how should they design a gradient descent based black-box adversarial attack to bypass $\mathcal{F}$?   [5 marks]

Note: Explain in detail

 (g) How does black-box adversarial attack work?

 (h) How is the gradient descent method used in this specific attack (explain mathematically)?

 (i) Why can the attack still be effective even if the attacker does not have access to $\mathcal{F}$?

Answer:

Assignment Project Exam Help

https://tutorcs.com

END OF EXAM QUESTIONS

WeChat: cstutorcs

Extra space if needed to answer questions 1–22. If you write part of your answer here, please write the question number, and indicate at the corresponding question that you have used this space.

Assignment Project Exam Help

https://tutorcs.com

WeChat: cstutorcs

LAST PAGE OF EXAM