UNIVERSITY OF BIRMINGHAM

School of Computer Science

Week 9 Exercise Questions

Artificial Intelligence 2

2023

Artificial Intelligence 2

Exercise questions Week 9

Question 1

As a developer of a security equipment company, you are going to design an alarm that senses when an infra-red sensor gauge exceeds a given threshold. The infra-red sensor measures the infra-red temperature and the gauge measures the infra-red temperature obtained from the infra-red sensor. Consider the Boolean variables A (alarm sounds), F_a (alarm is faulty), F_g (gauge is faulty) and the G (gauge reading: normal and high) and T (actual infra-red temperature: normal and high).

(a) Draw a Bayesian network for this problem.

(b) Write down the joint probability distribution represented by this Bayesian network.

(c) How many parameters are required to describe this joint probability distribution? Show your working.

0 marks for question not valid - all questions must have the same number of marks

Question 2 (2022 AI2 Exam question)

As a machine learning expert for an AI cyber security company, your task is to design an automated network intrusion detection system. You have collected a large number of records of network activities. Each record includes the log information about network activity, such as protocol types, duration, number of failed logins, which are random variables, denoted as $\boldsymbol{X} = [X_1, X_2, \ldots, X_n]^\intercal$. Each record also includes a binary random variable Y called label that was labelled by cyber security experts as intrusions ($Y = 1$) or normal connections ($Y = 0$).

(a) Consider feature selection based on mutual information to reduce the number of independent variables.

(i) Explain to your colleague, who knows nothing about information theory, the concept of mutual information. **[2 marks]**

(ii) Explain the loop, i.e., lines 4-7 of the pseudocode in Table 1. Note $I(Y; X_i)$ is the mutual information between Y and X_i.

[3 marks]

1: **Initialisation:** Set $F \leftarrow \boldsymbol{X}$ and $S \leftarrow \emptyset$
2: $f_{\max} = \text{argmax}_{X_i \in \boldsymbol{X}} I(Y; X_i)$
3: Set $F \leftarrow F \setminus \{f_{\max}\}$ and $S \leftarrow f_{\max}$
4: **Repeat** until $|S| = K$:
5: $f_{\max} = \text{argmax}_{X_i \in F} I(Y; X_i) - \beta \sum_{X_s \in S} I(X_s; X_i)$
6: Set $F \leftarrow F \setminus \{f_{\max}\}$ and $S \leftarrow f_{\max} \cup S$
7: **End**
8: **End**

Table 1: Pseudocode of Mutual Information based Feature Selection Algorithm.

(b) After applying your feature selection algorithm, assume you selected four random variables as features, denoted as F_1, F_2, F_3, F_4. Based on these features, you now work with a cyber security expert to construct a Bayesian network to harness the domain knowledge of cyber security. The expert first divides intrusions into three cyber attacks, A_1, A_2, A_3, which are marginally independent from each other. The expert suggests the presence of the four features are used to find the most probable type of cyber attacks. The four features are conditionally dependent on the three types cyber attacks as follows: F_1 depends only on A_1, F_2 depends on A_1 and A_2. F_3 depends on A_1 and A_3, whereas F_4 depends only on A_3. We assume all these random variables are binary, i.e., they are either 1 (true) or 0 (false).

(i) Draw the Bayesian network according to the expert?s description. **[2 marks]**

(ii) Write down the joint probability distribution represented by this Bayesian network. **[3 marks]**

(iii) How many parameters are required to describe this joint probability distribution? Show your working. **[5 marks]**

(iv) Suppose in a record we observe F_2 is true, what does observing F_4 is true tell us? If we observe F_3 is true instead of F_2, what does observing F_4 is true tell us? **[5 marks]**