Machine Learning and AI Exam

Module: L6 – Artificial Intelligence and Machine Learning (ACCA6011)

Student: Anton Hall

ID: 1809131

Tutor: Gordan Dickers

Q1, a,

Anton Hall

Q1, a,

BMI Group = Medium.
Age Group = Lower
Gender = Male.

P (BMI = Medium | OSA) = 4/9
P (Age = lower | OSA) = 6/9
P (Gender = male | OSA) = 6/9.

OSA Yes = (4/9) x (6/9) x (6/9) = 0.198

P(BMI = Medium | No OSA) = 1/5
P(Age = lower | No OSA) = 1/5
P (Gender = male | No OSA) = 1/5

OSA No = (1/5) x (1/5) x (1/5) = 0.008

P(OSA | x) > P(NoOSA | x) = 0.198 > 0.008

Therefore the likelihood of OSA is Yes

Q1, b,

A way of dealing with very small probability numbers is to use the method of scientific notation. Therefore, the answer above for OSA no can be 8x10-3. The exponent is -3, making it 10 to the power of negative 3. When an exponent is negative, the solution is a number less than the origin or base number. To find our answer, we move the decimal to the left 3 times:

8.0 -> 0.008

Q2, a,

Informed Search: Informed Search algorithms uses knowledge for the searching process

Example: Graph Search

Uninformed Search: Uninformed search algorithms doesn't use knowledge for the searching process

Example: Depth First Search

Q2, b,

Dijkstra's algorithm is an algorithm for finding the shortest paths between nodes in a graph. It will find the shortest path if it exists, but it could get caught up in an infinitely long path. The original algorithm runs in runs in O(V^2) time but with a minimum priority queue it drops down to O(V+Elog V). Theses methods can be expensive in terms of time. Space complexity of Dijkstra's algorithm is O(V^2) where V denotes the number of vertices (or nodes) in the graph. The data structure used in Dijkstras algorithm is minimum priority queue.

Q2, c,

A* is similar to Dijkastra's but the evaluation function has been modified to include a heuristic:

f(n) = g(n) + h(n)

g(n) gives the path cost from the start node to node n - like Dijkastra. h(n) is the estimated cost of the cheapest path from n to goal.

Q2, d,

	CLOSED LIST (READ VERTICALLY)	**OPEN LIST (READ HORIZONTALLY)**		
1	A(0,A)	B(24,A)	G(9,A)	
2	B(24,A)	E(27.9,B)	G(9,A)	C(36.3,B)
3	E(27.9,B)	G(38.4,E)	F(36.3,E)	C(36.3,B)
4	G(38.4,E)	F(36.3,E)	H(65.4,G)	C(36.3,B)
5	F(36.3,E)	H(45.9,F)	C(36.3,B)	
6	H(45.9,F)	C(36.3,B)	D(55.2,H)	
7	C(36.3,B)	D(55.2,H)		
8	D(55.2,H)			

Q4, a,

Supervised Learning is a machine learning method that uses labelled datasets to train or supervise algorithms into classifying or predicting outcomes. It can be separated into two types of tasks:

- Classification and
- Regression

Unsupervised Learning is a machine learning method to analyse and cluster data using unlabelled datasets. They are used to discover patterns without the need for human intervention. Unsupervised learning models can be used for two main tasks:

- Clustering
- Identifying Classes

Q4, b,

Naïve Bayes is a supervised learning algorithm. To calculate naïve bayes a data set is needed which will need some labelling associated with it.

K- means clustering is a unsupervised learning algorithm. K-means algorithm only requires the data to be clustered it does not require the data to be labelled into classes.

Q4, c,

The goal of a perceptron is to make accurate classifications. To train a model to make accurate classifications perceptron synaptic weights need to be optimised for any specific classification task. They allow the output to trigger correctly based on all the instances of inputs, it's what allows it to learn. The presynaptic neurons are represented by their firing rates, x1, x2 and x3 each of which is multiplied by the synaptic weight w1, w2 and w3 and then summed together with the other inputs.

Q4, d,

Anton Hall

Q4, d,

$x1 = 0.8$ bias $= 0.3$

$x2 = -0.22$

$x3 = 0.12$

$y = (0.8 \times 0.1) + (-0.22 \times -0.4) + (0.12 \times 0.22) + 0.3$

$y = 0.08 + 0.088 + 0.0264 + 0.3$

$y = 0.4944$

$\phi(\max(0, y))$

$= 0.4944$

$\sigma(y) = \frac{1}{1+e^{-y}}$

$= \frac{1}{1+e^{-0.4944}}$

$= 0.621$

The constant gradient of ReLU result is faster learning.