1. Machine learning is  
a) The autonomous acquisition of knowledge through the use of computer programs  
b) The autonomous acquisition of knowledge through the use of manual programs  
c) The selective acquisition of knowledge through the use of computer programs  
d) The selective acquisition of knowledge through the use of manual programs  
View Answer

Answer: a  
Explanation: Machine learning is the autonomous acquisition of knowledge through the use of computer programs.

2. Factors which affect the performance of learner system does not include  
a) Representation scheme used  
b) Training scenario  
c) Type of feedback  
d) Good data structures  
View Answer

Answer: d  
Explanation: Factors which affect the performance of learner system does not include good data structures.

3. Different learning methods does not include  
a) Memorization  
b) Analogy  
c) Deduction  
d) Introduction  
View Answer

Answer: d  
Explanation: Different learning methods does not include introduction.

4. In language understanding, the levels of knowledge that does not include  
a) Phonological  
b) Syntactic  
c) Empirical  
d) Logical  
View Answer

Answer: c  
Explanation: In language understanding, the levels of knowledge that does not include empirical knowledge.

5. A model of language consists of the categories which does not include  
a) Language units  
b) Role structure of units  
c) System constraints  
d) Structural units  
View Answer

Answer: d  
Explanation: A model of language consists of the categories which does not include structural units.

6. What is a top-down parser?  
a) Begins by hypothesizing a sentence (the symbol S) and successively predicting lower level constituents until individual preterminal symbols are written  
b) Begins by hypothesizing a sentence (the symbol S) and successively predicting upper level constituents until individual preterminal symbols are written  
c) Begins by hypothesizing lower level constituents and successively predicting a sentence (the symbol S)  
d) Begins by hypothesizing upper level constituents and successively predicting a sentence (the symbol S)  
View Answer

Answer: a  
Explanation: A top-down parser begins by hypothesizing a sentence (the symbol S) and successively predicting lower level constituents until individual preterminal symbols are written.

7. Among the following which is not a horn clause?  
a) p  
b) Øp V q  
c) p → q  
d) p → Øq  
View Answer

Answer: d  
Explanation: p → Øq is not a horn clause.

8. The action ‘STACK(A, B)’ of a robot arm specify to  
a) Place block B on Block A  
b) Place blocks A, B on the table in that order  
c) Place blocks B, A on the table in that order  
d) Place block A on block B  
View Answer

Answer: d  
Explanation: The action ‘STACK(A,B)’ of a robot arm specify to Place block A on block B.

1. What are the three essential components of a learning system? Give a definition of

each. Give an example of each, including equations where necessary. (1 mark)

A. Model, gradient descent, learning algorithm

B. Error function, model, learning algorithm

C. Accuracy, Sensitivity, Specificity

D. Model, error function, cost function

2. The error function most suited for gradient descent using logistic regression is

A. The entropy function

B. The squared error

C. The cross-entropy function

D. The number of mistakes

3. After SVM learning, each Lagrange multiplier ai takes either zero or non-zero value.

What does it indicate in each situation?

A. A non-zero ai indicates the datapoint i is a support vector, meaning it touches the

margin boundary.

B. A non-zero ai indicates that the learning has not yet converged to a global minimum.

C. A zero ai indicates that the datapoint i has become a support vector datapoint, on

the margin.

D. A zero ai indicates that the learning process has identified support for vector i.

Page 2 of 4

COMP61011

4. A “Bayesian Network” is most accurately described as

A. A special case of a neural network that makes use of Bayes’ Theorem.

B. The network variant of Bayes Theorem, assuming independent features.

C. A probabilistic model of which Naive Bayes is a special case.

D. A network of probabilistic learning functions, connected by Bayes Rule.

5. The SVM optimization problem is given by L = 1

2wTw􀀀åiai(yi f (xi)􀀀1).

What is the purpose of the åiai(yi f (xi)􀀀1) part?

A. To optimise the value of k in the nearest neighbour model.

B. To constrain the optimisation to ensure correct classifications.

C. To maximise the margins.

D. To minimise the margins.

6. An appropriate learning algorithm for the SVM is

A. Quadratic programming of soft margins

B. Quadratic programming via Gradient Descent

C. Gradient Descent with Lagrange Multiplier constraints

D. Quadratic programming via Sequential Minimal Optimization

7. Assuming log base 2, the entropy of a binary feature with p(x = 1) = 0:75 is

A. 0.1875

B. 0.8113

C. 0.1887

D. 2.41

Page 3 of 4 [PTO]

COMP61011

8. Boosting is said to be a good classifier because

A. It creates all ensemble members in parallel, so their diversity can be boosted.

B. It attempts to minimise the margin distribution.

C. It attempts to maximise the margins on the training data

D. None of the above

9. What does it mean to perform a data bootstrap?

A. To sample M features with replacement from the total M.

B. To sample M features without replacement from the total M.

C. To sample N examples with replacement from the total N.

D. To sample N examples without replacement from the total N.

10. A doctor can run a test for the horrible disease Examophobia. The test has two possible

outcomes: positive and negative. It is known that among all students, if Examophobia

is present, the test comes out positive 80% of the time, and negative 20% of the time.

If Examophobia is not present, the test comes out positive 1% of the time, negative

99%. Among the general student population, Examophobia is known to occur in 35%

of all students. A student enters the clinic and tests positive for the disease. What is

the probability they really have Examophobia?

A. 0.987

B. 0.977

C. 0.2

D. 0.35

1. A \_\_\_\_\_\_\_\_\_ is a decision support tool that uses a tree-like graph or model of decisions and their possible consequences, including chance event outcomes, resource costs, and utility.  
a) Decision tree  
b) Graphs  
c) Trees  
d) Neural Networks  
View Answer

Answer: a  
Explanation: Refer the definition of Decision tree.

2. Decision Tree is a display of an algorithm.  
a) True  
b) False  
View Answer

Answer: a  
Explanation: None.

3. Decision Tree is  
a) Flow-Chart  
b) Structure in which internal node represents test on an attribute, each branch represents outcome of test and each leaf node represents class label  
c) Flow-Chart & Structure in which internal node represents test on an attribute, each branch represents outcome of test and each leaf node represents class label  
d) None of the mentioned  
View Answer

Answer: c  
Explanation: Refer the definition of Decision tree.

4. Decision Trees can be used for Classification Tasks.  
a) True  
b) False  
View Answer

Answer: a  
Explanation: None.

5. Choose from the following that are Decision Tree nodes  
a) Decision Nodes  
b) End Nodes  
c) Chance Nodes  
d) All of the mentioned  
View Answer

Answer: d  
Explanation: None.

6. Decision Nodes are represented by \_\_\_\_\_\_\_\_\_\_\_\_  
a) Disks  
b) Squares  
c) Circles  
d) Triangles  
View Answer

Answer: b  
Explanation: None.

7. Chance Nodes are represented by,  
a) Disks  
b) Squares  
c) Circles  
d) Triangles  
View Answer

Answer: c  
Explanation: None.

8. End Nodes are represented by \_\_\_\_\_\_\_\_\_\_  
a) Disks  
b) Squares  
c) Circles  
d) Triangles  
View Answer

Answer: d  
Explanation: None.

9. Following are the advantage/s of Decision Trees. Choose that apply.  
a) Possible Scenarios can be added  
b) Use a white box model, If given result is provided by a model  
c) Worst, best and expected values can be determined for different scenarios  
d) All of the mentioned  
View Answer

Answer: d  
Explanation: None.