Dr Mahalingam College of Engineering and technology,Pollachi

Department of ECE

Course: 19ECCN1702/MACHINE LEARNING

Question Bank

UNIT 1 -INTRODUCTION

S.No	Questions	Marks
1	Compare traditional learning and Machine learning	**2**
2	Differentiate Batch vs online learning	2
3	Differentiate Instance vs Model based learning	2
4	Distinguish supervised and unsupervised learning	2
5	Compare Variance and Bias	2
6	Define Machine Learning	2
7	Compare overfitting vs underfiiting	2
8	Brief on batch algorithm.	2
9	Brief on e on online learning system	2
10	List the four main challenges in Machine Learning	2
11	Find out whether a mail is spam or ham, is supervised or unsupervised learning? Also tell whether it is a regression or classification problem.	2
12	Justify the need for separate test data set.	2
13	Give any 4 applications of ML	2
14	Give the commonly used supervised algorithms	2
15	List the commonly used unsupervised algorithms	2
16	Mention cost function for a single variable input and output system.	2
17	Compare classification and regression tasks in supervised algorithms	7
18	Distinguish the working procedure of Batch learning and online learning in detail.	8
19	Explain the various types of machine learning system with suitable example for each kind of learning	15
20	Describe about Testing and Validating the model in detail with suitable example	15

21	Elaborate scuss in detail about the challenges in machine learning	15
22	Brief the 3Vs that decides whether a problem should opt online ot batch algorithm.	8
23	Discuss various types of unsupervised and semi-supervised algorithms with suitalble examples	8
24	Explain in detail with a neat sketch, denoting the transformation from traditional approach to the machine learning approaches that helps humans to learn	15
25	Explain overfitting and underfitting of data	7
26	Find the suitable type of ML algorithm for detecting the debit card fraud based on location. Justify the answer.	7
27	Identify the suitable type of ML algorithm for Predcting the covid status of a patient from the data. Justify the answer.	8
28	Brief on reinforcement algorithm	7

Dr Mahalingam College of Engineering and technology,Pollachi

Department of ECE

Course: 19ECCN1702/MACHINE LEARNING

Question Bank

UNIT 2

S.No	Questions	Marks
1	Brief the steps to be followed in End-to-End Machine Learning Project	2
2	A Model should learn from the data and be able to predict the median housing price in any district, given all the other metrics. Suggest the suitable type of learning algorithm.	2
3	Compare different sampling methods.	2
4	Define Hypothesis	2
5	define in what ratio the dataset set has to be split into training dataset and testing dataset	2
6	Distinguish head( ) and info( ) functions.	2
7	Give few open source data set repository.	2
8	Give the essential libraries of python required for machine learning	2
9	Illustrate the construction of a Pipeline for Investment analysis model	2
10	List the steps involved in problem framing process of ML	2
11	state the formula for RMSE, and describe the variables in formula.	2
12	what is Histogram	2
13	what is MAE.	2
14	what is pipeline	2
15	Write the purpose of describe( ) function.	2
16	Compare the two feature scaling methods	7
17	The values of independent variable x and dependent value y are given below x y 0 2 1 3 2 5 3 4 4 6 Find the least square regression line y=ax+b. Estimate the value of y when x is 10	7
18	Mention how the data is cleaned for machine learning algorithm	8
19	How to improve the sampling method.	7
20	Explain the purpose of grid search in fine tuning the model.	8
21	Explain the process of measuring accuracy using Cross-Validation	8
22	Discuss on Experimenting with Attribute Combinations in data	7
23	Brief on confusion matrix with a simple example.	8
24	Brief the steps to be followed in End-to-End Machine Learning Project in detail	15
25	Apply suitable python functions to prepare the data for ML algorithms and explain the steps in detail.	15
26	Explain in detail about the evaluation metrics of machine learning in detail	15
27	i) For the dataset set collected on a population as shown in Table.1, the problem statement is to predict what range car a person will buy. Identify whether this is a regression or classification problem supervised / unsupervised / Semi supervised learning Multivariate/univariate Online or offline learning Instance based or model based learning	15

Table 1

Gender	Profession	Income level	Marital status	No. of children	Owner ship of four wheeler *
Male	Yes	Low	Y	2	0
Male	No	Medium	N	0	1
Female	Yes	High	N	0	2
Female	Yes	Low	Y	1	0
Male	No	Low	Y	2	0
Female	No	Medium	Y	2	1
Male	Yes	High	N	0	3

0- owns no car
1- owns car value less than 5 lakhs
2- owns car value between 5 -10 lakhs
3- owns car value between 10-20 lakhs

Dr Mahalingam College of Engineering and technology,Pollachi

Department of ECE

Course: 19ECCN1702/MACHINE LEARNING

Question Bank

UNIT 3

<table>
<tr><th>S.No</th><th>Questions</th><th>Marks</th></tr>
<tr><td>1</td><td>Compare Linear regression and Logistic regression</td><td>2</td></tr>
<tr><td>2</td><td>Give any 2 problems where Linear regression can be used appropriately</td><td>2</td></tr>
<tr><td>3</td><td>List any 2 problems where logistic regression can be used appropriately</td><td>2</td></tr>
<tr><td>4</td><td>Give a short note on Eigen vectors</td><td>2</td></tr>
<tr><td>5</td><td>Define Maximum likelihood estimation</td><td>2</td></tr>
<tr><td>6</td><td>Give a short note on MDL</td><td>2</td></tr>
<tr><td>7</td><td>List the types of gradient descent</td><td>2</td></tr>
<tr><td>8</td><td>Give the formula followed in Naïve Bayes theorem</td><td>2</td></tr>
<tr><td>9</td><td>Give a brief note on Maximum likelihood estimation</td><td>7</td></tr>
<tr><td>10</td><td>Compare the 3 different gradient descent functions</td><td>7</td></tr>
<tr><td>11</td><td>Compare Linear and Logistic Regression in detail with suitable examples</td><td>8</td></tr>
<tr><td>12</td><td>Medical researchers want to know how exercise and weight impact the probability of having a heart attack
The response variable in the model will be heart attack and it has two potential outcomes
A heart attack occurs
A heart attack does not occur
The results of the model will tell researchers exactly how changes in exercise and weight affect the probability that a given individual has a heart attack
Which is the most suitable model to produce the expected results, Justify</td><td>8</td></tr>
<tr><td>13</td><td>Apply Naïve Bayes algorithm to the following car theft problem to classify a Red Domestic SUV is getting stolen or not.
<table>
<tr><th>Example No</th><th>Color</th><th>Type</th><th>Origin</th><th>Stolen?</th></tr>
<tr><td>1.</td><td>Red</td><td>Sports</td><td>Domestic</td><td>Yes</td></tr>
<tr><td>2.</td><td>Red</td><td>Sports</td><td>Domestic</td><td>No</td></tr>
<tr><td>3.</td><td>Red</td><td>Sports</td><td>Domestic</td><td>Yes</td></tr>
<tr><td>4.</td><td>Yellow</td><td>Sports</td><td>Domestic</td><td>No</td></tr>
<tr><td>5.</td><td>Yellow</td><td>Sports</td><td>Imported</td><td>Yes</td></tr>
<tr><td>6.</td><td>Yellow</td><td>SUV</td><td>Imported</td><td>No</td></tr>
<tr><td>7.</td><td>Yellow</td><td>SUV</td><td>Imported</td><td>Yes</td></tr>
<tr><td>8.</td><td>Yellow</td><td>SUV</td><td>Domestic</td><td>No</td></tr>
<tr><td>9.</td><td>Red</td><td>SUV</td><td>Imported</td><td>No</td></tr>
<tr><td>10.</td><td>Red</td><td>Sports</td><td>Imported</td><td>Yes</td></tr>
</table></td><td>15</td></tr>
</table>

14		15

S.No	Outlook	Temperature	Humidity	Windy	Play Golf
0	Rainy	Hot	High	FALSE	No
1	Rainy	Hot	High	TRUE	No
2	Overcast	Hot	High	FALSE	Yes
3	Sunny	Mild	High	FALSE	Yes
4	Sunny	Cool	Normal	FALSE	Yes
5	Sunny	Cool	Normal	TRUE	No
6	Overcast	Cool	Normal	TRUE	Yes
7	Rainy	Mild	High	FALSE	No
8	Rainy	Cool	Normal	FALSE	Yes
9	Sunny	Mild	Normal	FALSE	Yes
10	Rainy	Mild	Normal	TRUE	Yes
11	Overcast	Mild	High	TRUE	Yes
12	Overcast	Hot	Normal	FALSE	Yes
13	Sunny	Mild	High	TRUE	No

Calculate the probability and predict the classification for the test data with the training data table using Naïve Bayes classifier method. Training data is the above table. Test data: today = (Sunny, Hot, Normal, False)

Unit	Question	Mark
3	How does the decision trees are making decisions	2
3	Infer on Ensemble of Decision tress	2
3	When do you stop splitting Decision Tree?	2
3	Interpret on Kernel trick	2
3	Define "entropy" and "information gain" in the context of decision tree construction.	2
3	How does "support vectors" in SVMs are important in defining the decision boundary	2
3	Define the leaf node in decision tree	2
3	Discuss the advantages of combining multiple decision trees to form an ensemble, such as Random Forests.	8
3	Discuss on Gradient boosted regression trees with its advantageous	8
3	Compare the various supervised learning algorithms with its merits and demerits	7
3	Define Support Vector Machines (SVMs) and discuss their use in both classification and regression tasks.	7
3	Build a Decision Tree for Supervised machine learning problem and explore on Ensemble of decision trees.	15
3	Illustrate the SVM optimization problem and the role of kernels in extending SVMs to non-linear problems.	15
4	State the applications of clustering technique	2
4	Mention the demerits of K means Clustering algorithm	2
4	List the methods to merge clusters in Agglomerative Clustering	2
4	Identify the parameters required to do clustering in DBSCAN algorithm	2
4	Distinguish between content based filtering and collaborative filtering	2
4	Draw the dendrogram diagram for bottom to top approach in Agglomerative clustering	2
4	List the parameters used in DBSCAN algorithm	2
4	List the different scales of k means clustering algorithm	2
4	Brief about the challenges in unsupervised learning	2
4	List the steps involved in k means clustering algorithm	2
4	How can the choice of the number of clusters (k) impact K-Means results?	2
4	How does the DBSCAN classify data points into different groups?	2

4	Identify the fundamental idea behind Gaussian Mixture Models	2
4	Draw the architecture of Gaussian mixture Model	2
4	Differentiate soft clustering and hard clustering	2
4	Compare various clustering methods in unsupervised learning	8
4	Discuss the time complexity of agglomerative clustering and strategies for optimizing its computational efficiency, especially for large datasets.	8
4	Explain the concept of core points, border points, and noise points in DBSCAN. How does DBSCAN handle clusters of varying densities?	8
4	Explain the concept of hybrid recommendation systems that combine collaborative filtering and content filtering.	8
4	Brief on Content based filtering in unsupervised learning	7
4	Brief on collaborative filtering unsupervised learning	7
4	Discuss the challenges associated with parameter estimation in GMMs.	7
4	Discuss the precision-recall trade-off and how it impacts the choice of threshold in binary classification models.	7
4	Illustrate the principle of K means clustering algorithm for unsupervised machine learning problems with its demerits and merits	15
4	Elucidate on DBSCAN algorithm for large dataset Machine learning problems.	15
4	Infer on Gaussian mixture model as an Expectation-Maximization (EM) Algorithm	15
4	Illustrate the principle of Agglomerative clustering algorithm for unsupervised machine learning problems with its demerits and merits	15
5	Define Artificial Neural Networks	2
5	Give the signum activation function expression and draw it.	2
5	Brief on threshold logic unit with diagram	2
5	Write the expression of Hyperbolic tangent function and draw it.	2
5	Articulate the significance of filters used in convolutional layer	2
5	What is zero padding in convolution?	2
5	Describe softmax activation function.	2
5	Give four hyper parameters used for tuning Neural Network	2
5	What is meant by stride in CNN?	2
5	List a few CNN architectures	2
5	Compare Biological Neuron and Artifcilal Neuron	2
5	List the popular activation functions in neural network	2

5	Draw an ANN to perform OR operation	2
5	State the features of AlexNet in CNN	2
5	State the features of GoogLeNet in CNN	2
5	Elaborate on various activation functions of neural networks	8
5	Solve the X-OR classification problem using MLP	8
5	Discuss the architecture of Regression MLP	8
5	Discuss the architecture of Classification MLP	8
5	Brief on the architecure of Multilayer Perceptron	7
5	Elaborate on back propagation algorithm to train MLP	7
5	Elucidate on pooling layers in convolution neural network	7
5	Explain about any 4 CNN architectures in detail	7
5	**Demonstrate multilayer perceptron and back propagation training algorithm in detail.**	15
5	**Discuss convolutional neural network architecture with necessary diagrams in detail**	15
5	**Discuss in detail Regression and Classification Multilayer Perceptron architectures in detail**	15
5	Demonstrate the evolution of Artificial Neural Network in align with biological neural network	15