A Sample of 52 MCQs. Most are downloaded from public websites, from basic ML courses. They give a flavor of what the MCQs will be like, though QF634 will have several tougher questions with some calculations.

Source: https://www.interviewbit.com/machine-learning-mcq/

Q1. Among the following option identify the one which is not a type of learning
(Clue: semi-supervised learning combines training data sets without labels, e.g., for K-means clustering, and some labelled data included for help to identify K. Reinforcement learning typically extends prediction of next advertising strategy given CTR responses being rewarding/profitable in new neighborhoods vs. punishment/loss in poor CTR in other neighborhoods – reward vs. punishment may be extensions of supervised learning loss function.)

A. Semi unsupervised learning
B. Supervised learning
C. Reinforcement learning
D. Unsupervised learning

Ans: A

Q2. Identify the type of learning in which labeled training data is used.

A. Semi unsupervised learning
B. Supervised learning
C. Reinforcement learning
D. Unsupervised learning

Ans. B

Q3. Among the following identify the one in which dimensionality reduction reduces.

A. Performance
B. Entropy
C. Stochastics
D. Collinearity

Ans. D

Q4. Which of the following machine learning algorithm is based upon the idea of bagging?

A. Decision tree
B. Random-forest
C. Classification
D. Regression

Ans. B

Q5. Choose a disadvantage of decision trees among the following.
(Clue: Factor analysis is a method to reduce number of explanatory variables into a few factors where the factors can explain almost as much as the original features.)

A. Decision trees are robust to outliers
B. Factor analysis
C. Decision trees are prone to overfit
D. All of the above

Ans. C

Q6. What is the term known as on which the machine learning algorithms build a model based on sample data?

A. Data training
B. Training data
C. Transfer data
D. None of the above

Ans B

Q7. Machine learning is a subset of which of the following.
(Clue: Deep learning is sometimes used to describe NN with many layers.)

A. Artificial intelligence
B. Deep learning
C. Data learning
D. None of the above

Ans. A

Q8. Which of the following are common classes of problems in machine learning?

A. Regression
B. Classification
C. Clustering
D. All of the above

Ans. D

Q9. Among the following options identify the one which is false regarding regression.

A. It is used for the prediction
B. It is used for interpretation
C. It relates inputs to outputs
D. It discovers casual relationships

Ans. D

Q10. Identify the successful applications of ML.

A. Learning to classify new astronomical structures
B. Learning to recognize spoken words
C. Learning to drive an autonomous vehicle
D. All of the above

Ans. D

Q11. Identify the incorrect numerical functions in the various function representation of machine learning.

A. Case-based
B. Support vector machines
C. Linear regression
D. Neural network

Ans. A

Q12. Choose the general limitations of the backpropagation rule among the following.

A. Slow convergence
B. Scaling (increasing size of NN, width and depth)
C. Local minima problem (reaching local minima instead of global at end of iterations)
D. All of the above

Ans. D

Q13. Choose the most widely used metrics (delete “mattresses”) and tools to assess the classification models. (Clue: cost-sensitivity refers to consideration of computing time resource)

A. The area under the ROC curve
B. Confusion matrix
C. Cost-sensitive accuracy
D. All of the above

Ans. D

Q14. Which of the following is not a supervised learning

A. PCA
B. Naive Bayesian
C. Linear regression
D. Decision tree

Ans. A

https://letsfindcourse.com/data-science/machine-learning-mcq-questions-and-answers

Q15. To find the minimum or the maximum of a function, we set the gradient to zero because:

A. The value of the gradient at extrema of a function is always zero
B. Depends on the type of problem
C. Both A and B
D. None of the above

Ans. A

Q16. How do you handle missing or corrupted data in a dataset?

A. Drop missing rows or columns
B. Replace missing values with mean/median/mode
C. Assign a unique category to missing values
D. All of the above

Ans. D

Q17. When performing regression or classification, which of the following is the correct way to preprocess the data?

A. Normalize the data -> PCA -> training
B. PCA -> normalize PCA output -> training
C. Normalize the data -> PCA -> normalize PCA output -> training
D. None of the above

Ans. A

Q18. Which of the following is a reasonable way to select the number of principal components "k"?

A. Choose k to be the smallest value so that at least 99% of the variance (delete “variance”) is retained.
B. Choose k to be 99% of m (k = 0.99*m, rounded to the nearest integer).
C. Choose k to be the largest value so that 99% of the variance is retained.
D. Use the elbow method.

Ans. A

https://www.onlineinterviewquestions.com/machine-learning-mcq/

Q19. Logistic regression is a ……. regression technique that is used to model data having a …… outcome.

A. Linear, binary
B. Linear, numeric
C. Nonlinear, binary
D. Nonlinear, numeric

Ans. C

Q20. Regression trees are often used to model which data?
(Clue: nonlinear refers to nonlinear data where data elements are not lined up neatly as in a hyperplane)

A. Linear
B. Nonlinear
C. Categorical
D. None of the above

Ans. B

Q21. What is called the average squared difference between classifier predicted output and actual output?

A. Mean relative error
B. Mean squared error
C. Mean absolute error
D. Root mean squared error

Ans. B

Q22. ….. are the best machine learning method.

A. Fast
B. Accuracy
C. Scalable
D. All of the above

Ans. D

Q23. What is the output of training process in machine learning?

A. Null
B. Accuracy
C. Machine learning model
D. Machine learning algorithm

Ans. C

Q24. …. Is the approach of basic algorithm for decision tree induction.
(Clue: induction = building)

A. Greedy
B. Top Down
C. Procedural
D. Step by step

Ans. A

Q25. What is the way to ensemble multiple classifications or regression?
(Clue: refers to multiple classifiers and regression methods)

A. Bagging
B. Blending
C. Boosting
D. Stacking

Ans. D

Q26. What is the most common issue when using Machine Learning?
(Clue: common = 'critical' here)

A. Poor Data Quality
B. Lack of skilled resources
C. Inadequate infrastructure
D. None of the above

Ans. A

https://www.freetimelearning.com/online-quiz/machine-learning-quiz.php

Q27. What is Machine learning?
(Clue: selective = human, autonomous = automated, knowledge = assessing outcome)

A. The autonomous acquisition of knowledge through the use of manual programs

B. The selective acquisition of knowledge through the use of manual programs

C. The autonomous acquisition of knowledge through the use of computer programs

D. The selective acquisition of knowledge through the use of computer programs

Ans. C

Q28. Machine learning is a subset of

A. Data Learning
B. Deep Learning
C. Artificial Intelligence
D. None of the above

Ans. C

Q29. The prior goal of unsupervised learning model is to determine the _____ .

A. Discuss
B. Regression task
C. Data patterns
D. Classification task

Ans. C

https://study.sagepub.com/easterbysmith7e/student-resources/chapter-11/multiple-choice-questions

Q30. Which of the following is NOT true about variable (feature) importance scores?

A. there are different ways of measuring feature importance, for example by using mean decrease accuracy scores

B. show the size and direction of the effect of a predictor on an outcome

C. quantify the unique contribution of a given feature (predictor)

D. pertains mostly to supervised machine learning

Ans. B

Q31. What is the purpose of cross-validation?

A. to re-formulate a research hypothesis

B. to find the optimal number of clusters in unsupervised learning

C. to validate the performance of a machine learning algorithm by resampling the original dataset

D. to apply different thresholds of statistical significance

Ans. C

Q32. Supervised learning is _____.

A. an algorithm with a defined outcome variable (label)

B. an algorithm without labelling (a defined outcome variable)

C. when a researcher decides the order in which predictors will be included in the model

D. an algorithm that supervises itself without human interference

Ans. A

https://www.dailyrecruitment.in/machine-learning-mcq-questions-and-answer-pdf-download/

Q33. PCA is_____

A. backward feature selection
B. forward feature selection
C. feature extraction
D. None of these

Ans. C

Q34. Supervised learning and unsupervised clustering both require which is correct according to the statement.

A. input attribute
B. hidden attribute
C. output attribute
D. categorical attribute

Ans. A

Q35. A feature F1 can take certain value: A, B, C, D, E, & F and represents grade of students from a college. Here feature type is_____
(Clue: Categorical variables represent types of data which may be divided into groups. Examples of categorical variables are race, sex, age group, and educational – one group is not larger or smaller than the other. Ordinal data is a kind of qualitative data that groups variables into ordered categories with a natural order or rank. Nominal data are used to label variables without any quantitative value, e.g. gender, as distinct from cardinal data or more generally continuous real numbers.)

A. ordinal
B. nominal
C. categorical
D. Boolean

Ans. A

Q36. Which of the following is a good test dataset characteristic?

A. is representative of the dataset as a whole
B. large enough to yield meaningful results
C. All of above
D. None of above

Ans. C

Q37. Which of the following techniques would perform better for reducing dimensions of a data set?

A. removing columns which have high variance in data
B. removing columns which have too many missing value
C. removing columns with dissimilar data trends
D. None of the above

Ans. B

Q38. You are given reviews of few Netflix series marked as positive, negative and neutral. Classifying reviews of a new netflix series is an example of_____

A. unsupervised learning
B. semi supervised learning
C. supervised learning
D. reinforcement learning

Ans. C

Q39. Application of machine learning methods to large databases is called_____

A. big data computing
B. artificial intelligence
C. data mining
D. internet of things

Ans. C

Q40. Of the Following Examples, Which would you address using an supervised learning Algorithm?

A. given a set of news articles found on the web, group them into set of articles about the same story
B. given email labeled as spam or not spam, learn a spam filter
C. given a database of customer data, automatically discover market segments and group customers into different market segments
D. find the patterns in market basket analysis

Ans. B

Q41. If machine learning model output doesn't involve (delete "involves") target variable then that model is called as_____

A. predictive model
B. descriptive model
C. reinforcement learning
D. all of the above

Ans. B

Q42. Following are the descriptive models_____

A. classification
B. clustering
C. association rule
D. Both 1 and 2

Ans. D

Q43. Some telecommunication company wants to segment their customers into distinct groups ,this is an example of_____

A. supervised learning
B. unsupervised learning
C. data extraction
D. reinforcement learning

Ans. B

Q44. In multiclass classification number of classes must be_____

A. equals to two
B. less than two
C. greater than two
D. None

Ans. C

Q45. Which of the following can best (delete 'only') be used when training data are linearly separable?

A. linear logistic regression
B. linear soft margin svm
C. linear hard-margin svm
D. the centroid method

Ans. C

Q46. The (critical) effectiveness of an SVM depends upon_____

A. kernel parameters
B. selection of kernel
C. soft margin parameter
D. All of the above

Ans. B

Q47. What is the purpose of the Kernel Trick?

A. to transform the problem from regression to classification
B. to transform the problem from supervised to unsupervised learning.
C. to transform the data from nonlinearly separable to linearly separable
D. All of above

Ans. C

Q48. Which of the following evaluation metrics cannot be applied in case of logistic regression output to compare with target?

A. accuracy
B. auc-roc
C. logloss
D. mean-squared-error

Ans. D

QF634 examples

Q49. What are three key financial statements from which accounting ratios can be derived for predicting company performances?

A. Income, Investment, Balance Sheet
B. Income, Profit and Loss, Cashflow
C. Balance Sheet, Profit and Loss, Loan Account
D. Cashflow, Balance Sheet, Income

Ans. D

Q50. In a 1-NN algorithm, the training points are (1, 2, 0.5), (1, -0.5, 1.5), and (0, -1, 0). The first coordinate is value of label Y. Y=1 or Y=0. The second and third coordinates are features X1 and X2 values. A test point has second and third coordinates (0.5, 0.5), what is its predicted Y?

A. 0
B. 1
C. $\sqrt{2}$
D. None of the above

Ans. B

Q51. What is overfitting in ML?

A. extremely high training accuracy
B. extremely high test accuracy
C. model with too many features
D. high training accuracy but relatively low validation or test accuracy

Ans. D

Q52. Which of the following is the best feature to use in trying to predict crude oil spot price 3-month into the future? The latter are the labels. A 3-month crude oil forward contract (ignore the type of crude) is a contract to buy/sell crude oil at 3 months into the future. A 1-month crude oil forward contract is a contract to buy/sell crude oil at 1 month into the future. A 3-month crude oil call option is a contract to exercise for buyer at 3 months in the future when the crude oil price exceeds the strike price.

A. lagged 3-month crude oil spot price
B. 1-month forward crude oil price
C. 3-month forward crude oil price
D. 3-month crude oil call price

Ans. C