

- ❖ Machine learning is the Science of Getting computers to learn without being explicitly programmed. Machine learning works on a simple concept that is understanding with experiences.
- ❖ The primary aim of machine learning is to allow computers to learn automatically without human interaction.
- ❖ **Scope of machine learning**
 1. Machine learning in education
 2. Machine learning in search engine
 3. Machine learning in digital marketing
 4. Machine learning in Healthcare
 5. Spam protector
 6. Traffic alert
 7. Social media
 8. Google Translate
- ❖ **Limitation of machine learning**
 1. Accuracy depends on training and learning which is not always available.
 2. It requires large data sets to learn about various topics which might be time taking and require various resources
 3. Good performance cannot always be guaranteed.
 4. A Mason needs to have heterogeneity in the data set to learn meaningful Insight.
- ❖ **Types of machine learning**

1. Supervised learning:

in supervised learning, given training explain examples of Input and corresponding output, the machine can predict outputs for new inputs in supervised learning, we train the images with respect to data that is well labeled and with the correct output.

2. Unsupervised learning:

Unsupervised learning deals with the unlabeled data
No training data set is provided which means, no training will be given to the machine. Therefore it must work on its own to discover the required information.

The machine is trained with unlabelled data.

❖ **Linear regression**

1. Linear regression is a machine learning algorithm based on supervised learning
2. It is the easiest and most popular machine learning algorithm
3. It is used for predictive analysis
4. It makes prediction for continuous variable size price of a product or house of salary
5. Regression models target prediction values based upon their independent variables.

❖ **Artificial Neural Network (ANN)**

1. An artificial neural network is a computational nonlinear model that is inspired by the brain.
2. ANN can perform tasks like classification, prediction, decision-making, visualization, and others just by considering examples.
3. It consists of a large collection of artificial neurons of the processing element which operates in parallel
4. ANNs Are capable of learning, which takes place by altering values.

1. Among the following option identify the one which is not a type of learning

1. Semi unsupervised learning
2. Supervised learning
3. Reinforcement learning
4. Unsupervised learning

2. Identify the kind of learning algorithm for “facial identities for facial expressions”.

1. Prediction
2. Recognition patterns
3. Recognizing anomalies
4. Generating patterns

3. Identify the model which is trained with data in only a single batch.

1. Offline learning
2. Batch learning
3. Both A and B
4. None

4. What is the application of machine learning methods to a large database called?

1. Big data computing
2. Internet of things
3. Data mining
4. Artificial intelligence

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

1. Semi unsupervised learning
2. Supervised learning
3. Reinforcement learning
4. Unsupervised learning

6. Identify whether true or false: In PCA the number of input dimensions is equal to principal components.

1. True
2. False

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

1. Performance
2. Entropy
3. Stochastics
4. Collinearity

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

1. Decision tree
2. Random-forest
3. Classification
4. Regression

9. Choose a disadvantage of decision trees among the following.

1. Decision trees are robust to outliers
2. Factor analysis
3. Decision trees are prone to overfit
4. All of the above

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

1. Data training
2. Training data
3. Transfer data
4. None of the above

11. Machine learning is a subset of which of the following.

1. Artificial intelligence
2. Deep learning
3. Data learning
4. None of the above

12. Which of the following machine learning techniques helps in detecting the outliers in data?

1. Classification
2. Clustering
3. Anomaly detection
4. All of the above

13. The father of machine learning is _____

1. Geoffrey Everest Hinton
2. Geoffrey Hill
3. Geoffrey Chaucer
4. None of the above

14. The most significant phase in genetic algorithm is _____

1. Mutation
2. Selection
3. Fitness function
4. Crossover

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

1. Regression
2. Classification
3. Clustering
4. All of the above

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

1. It is used for the prediction
2. It is used for interpretation
3. It relates inputs to outputs
4. It discovers causal relationships

17. Identify the successful applications of ML.

1. Learning to classify new astronomical structures
2. Learning to recognize spoken words
3. Learning to drive an autonomous vehicle
4. All of the above

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

1. Case-based
2. Support vector machines
3. Linear regression
4. Neural network

19. FIND-S algorithm ignores?

1. Positive
2. Negative
3. Both
4. None

20. Select the correct definition of neuro software.

1. It is software used by neurosurgeons
2. It is software used to analyze neurons
3. It is a powerful and easy neural network
4. None of the above

21. Choose whether the following statement is true or false: The backpropagation law is also known as the generalized Delta rule.

1. True
2. False

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

1. Slow convergence
2. Scaling
3. Local minima problem
4. All of the above

23. Analysis of ML algorithm needs

1. Statistical learning theory
2. Computational learning theory
3. Both A and B
4. None of the above

24. Choose the most widely used metric and tools to assess the classification models.

1. The area under the ROC curve
2. Confusion matrix
3. Cost-sensitive accuracy
4. All of the above

25. Full form of PAC is _____

1. Probably Approx Cost
2. Probably Approximate Correct
3. Probability Approx Communication
4. Probably Approximate Computation

26. Choose that following statement is true or false: True error is defined over the entire instance space, and not just over training data

1. True
2. False

27. Choose the options below of which the area CLT is comprised of.

1. Mistake bound
2. Sample complexity
3. Computational complexity
4. All of the above

28. Choose the instance-based learner.

1. Eager learner
2. Lazy learner
3. Both A and B are correct
4. None of the above

29. Identify the difficulties with the k-nearest neighbor algorithm.

1. Curse of dimensionality
2. Calculate the distance of the test case from all training cases
3. Both A and B
4. None of the above

30. The total types of the layer in radial basis function neural networks is _____

- A. 1
- B. 2
- C. 3
- D. 4

31. Which of the following is an application of CBR?

1. Diagnosis
2. Design
3. Planning
4. All of the above

32. Choose the correct advantages of CBR.

1. Fast to train
2. A local approx is found for each test case
3. Knowledge is in a form understandable by human
4. All of the above

33. Machine learning as various Search and Optimisation algorithms. Identify among the following which is not evolutionary computation.

1. Genetic algorithm
2. Genetic programming
3. Neuroevolution
4. Perceptron

34. Choose whether the following statement is true or false: Artificial intelligence is the process that allows a computer to learn and make decisions like humans.

1. True
2. False

35. Which of the following is not machine learning disciplines?

1. Information theory
2. Optimisation + control
3. Physics
4. Neuro statistics

36. What does K stand for in K mean algorithm?

1. Number of clusters
2. Number of data
3. Number of attributes
4. Number of iterations

37. Choose whether true or false: Decision tree cannot be used for clustering

1. True
2. False

38. Identify the clustering method which takes care of variance in data

1. Decision tree
2. Gaussian mixture model
3. K means
4. All of the above

39. Which of the following is not a supervised learning

1. PCA
2. Naive Bayesian
3. Linear regression
4. Decision tree

40. What is unsupervised learning?

1. Number of groups may be known
2. Features of group explicitly stated
3. Neither feature nor number of groups is known
4. None of the above

41. Which of the following is not a machine learning algorithm?

1. SVM
2. SVG
3. Random forest
4. None of the above

42. What is true about Machine Learning?

1. The main focus of ML is to allow computer systems to learn from experience without being explicitly programmed or human intervention.
2. ML is a type of artificial intelligence that extracts patterns out of raw data by using an algorithm or method.
3. Machine Learning (ML) is the field of computer science.
4. All of the above

43. Which of the following is not machine learning?

1. Artificial intelligence
2. Rule-based inference
3. Both A and B
4. None of the above

44. Identify the method which is used for trainControl resampling.

1. svm
2. repeatedcv
3. Bag32
4. None of the above

45. Among the following option identify the one which is used to create the most common graph types.

1. plot
2. quickplot
3. qplot
4. All of the above

Machine Learning MCQ Questions and Answers

1) What is machine learning ?

- A. Machine learning is the science of getting computers to act without being explicitly programmed.
- B. Machine Learning is a Form of AI that Enables a System to Learn from Data.
- C. Both A and B
- D. None of the above

2) Machine learning is an application of _____.

- A. Blockchain
- B. Artificial Intelligence
- C. Both A and B
- D. None of the above

3) Application of Machine learning is _____.

- A. email filtering
- B. sentimental analysis
- C. face recognition
- D. All of the above

4) The term machine learning was coined in which year?

- A. 1958
- B. 1959
- C. 1960
- D. 1961

5) Machine learning approaches can be traditionally categorized into _____ categories.

- A. 3
- B. 4
- C. 7
- D. 9

6) The categories in which Machine learning approaches can be traditionally categorized are _____.

- A. Supervised learning
- B. Unsupervised learning
- C. Reinforcement learning
- D. All of the above

7) _____ is the machine learning algorithms that can be used with labeled data.

- A. Regression algorithms
- B. Clustering algorithms
- C. Association algorithms
- D. All of the above

8) _____ is the machine learning algorithms that can be used with unlabeled data.

- A. Regression algorithms
- B. Clustering algorithms
- C. Instance-based algorithms
- D. All of the above

9) The Real-world machine learning use cases are _____.

- A. Digital assistants
- B. Chatbots
- C. Fraud detection
- D. All of the above

10) Which among the following algorithms are used in Machine learning?

- A. Naive Bayes
- B. Support Vector Machines
- C. K-Nearest Neighbors
- D. All of the above

11) _____ are the techniques of keyword normalization

- A. Lemmatization
- B. Stemming
- C. Both A and B
- D. None of the above

12) Replace missing values with mean/median/mode helps to handle missing or corrupted data in a dataset. True/False?

- A. True
- B. False

13) _____ is a disadvantage of decision trees?

- A. Decision trees are robust to outliers
- B. Decision trees are prone to be overfit
- C. Both A and B
- D. None of the above

14) _____ is a part of machine learning that works with neural networks.

- A. Artificial intelligence
- B. Deep learning
- C. Both A and B
- D. None of the above

15) Overfitting is a type of modelling error which results in the failure to predict future observations effectively or fit additional data in the existing model. Yes/No?

- A. Yes
- B. No
- C. May be
- D. Can't say

16) _____ is used as an input to the machine learning model for training and prediction purposes.

- A. Feature
- **B.Feature Vector**
- C.Both A and B
- D.None of the above

17) _____ is the scenario when the model fails to decipher the underlying trend in the input data.

- A. Overfitting
- **B.Underfitting**
- C.Both A and B
- D.None of the above

18) Which Language is Best for Machine Learning?

- A. C
- B.Java
- **C.Python**
- D.HTML

19) The supervised learning problems can be grouped as _____.

- A. Regression problems
- B.Classification problems
- **C.Both A and B**
- D.None of the above

20) The unsupervised learning problems can be grouped as _____.

- A. Clustering
- B.Association
- **C.Both A and B**
- D.None of the above

21) Automatic Speech Recognition systems find a wide variety of applications in the _____ domains.

- A. Medical Assistance
- B.Industrial Robotics
- C.Defence & Aviation
- D.All of the above

22) The term machine learning was coined by _____.

- A. James Gosling
- B.Arthur Samuel
- C.Guido van Rossum
- D.None of the above

23) Machine Learning can automate many tasks, especially the ones that only humans can perform with their innate intelligence.

- A. True
- B.False

24) Features of Machine Learning are _____.

- A. Automation
- B.Improved customer experience
- C.Business intelligence
- D.All of the above

25) Which machine learning models are trained to make a series of decisions based on the rewards and feedback they receive for their actions?

- A. Supervised learning
- B.Unsupervised learning
- C Reinforcement learning
- D.All of the above

1) Regression discovers causal relationships.

- A. True
- B. False

2) Missing data items are with Bayes classifier.

- A. Ignored
- B. Treated as equal compares
- C. Treated as unequal compares
- D. Replaced with a default value.

3) A model of language consists of the categories, does not include

- A. Language units
- B. Structural units
- C. System constraints
- D. Role structure of units

4) What is the output of training process in machine learning?

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

Next Question

5) Supervised learning and unsupervised clustering both require at least one

- A. Input attribute
- B. Output attribute
- C. Hidden attribute
- D. Categorical attribute

6) Machine learning is a subset of

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

7) What is the most common issue when using Machine Learning?

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

8) Real-Time decisions, Game AI, Learning Tasks, Skill acquisition, and Robot Navigation are applications of

- A. Reinforcement Learning

- B. Supervised Learning: Classification
- C. Unsupervised Learning: Regression
- D. None of the above

9) The Bayes rule can be used in

- A. Solving queries
- B. Increasing complexity
- C. Decreasing complexity
- D. Answering probabilistic query

10) Which one in the following is not Machine Learning disciplines?

- A. Physics

- B. Information Theory

- C. Neurostatistics

- D. Optimization Control

11) What are the three types of Machine Learning?

- A. Supervised Learning

- B. Unsupervised Learning

- C. Reinforcement Learning

- D. All of the above

12) 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

13) algorithms enable the computers to learn from data, and even improve themselves, without being explicitly programmed.

- A. Deep Learning
- **B.Machine Learning**
- C.Artificial Intelligence
- D.None of the above

14) What is the most significant phase in a genetic algorithm?

- A. Selection
- **B.Mutation**
- **C.Crossover**
- D.Fitness function

15) In Machine learning the module that must solve the given performance task is known as

- A. Critic
- B.Generalizer
- **C.Performance system**
- D.All of these

16) Which of the following is not a supervised learning?

- **A. PCA**
- B.Naive Bayesian
- C.Linear Regression
- D.Decision Tree Answer

17) Machine Learning is a field of AI consisting of learning algorithms that

- A. At executing some task
- B.Over time with experience
- C.Improve their performance
- **D.All of the above**

18) Which methods are used for the calibration in Supervised Learning?

- A. Platt Calibration
- B.Isotonic Regression
- **C.Both Platt Calibration & Isotonic Regression**
- D.None of the above

19) What is the disadvantage of decision trees?

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

20) 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

1. Application of machine learning methods to large databases is called

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

[discuss](#)

A. data mining.

2. If machine learning model output involves target variable then that model is called as

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

[discuss](#)

B. predictive model

3. In what type of learning labelled training data is used

- A. unsupervised learning
- B. supervised learning
- C. reinforcement learning
- D. active learning

4. In following type of feature selection method we start with empty feature set

- A. forward feature selection
- B. backword feature selection
- C. both a and b??
- D. none of the above

[discuss](#)

A.forward feature selection

5. In PCA the number of input dimensiona are equal to principal components

- A. true
- B. false

[discuss](#)

A.true

6. PCA can be used for projecting and visualizing data in lower dimensions.

- A. true
- B. false

[discuss](#)

A.true

7. Which of the following is the best machine learning method?

- A. scalable
- B. accuracy
- C. fast
- D. all of the above

[discuss](#)

D.all of the above

8. What characterize unlabeled examples in machine learning

- A. there is no prior knowledge
- B. there is no confusing knowledge
- C. there is prior knowledge
- D. there is plenty of confusing knowledge

[discuss](#)

D.there is plenty of confusing knowledge

9. What does dimensionality reduction reduce?

- A. stochastics
- B. collinerity
- C. performance
- D. entropy

[discuss](#)

B.collinerity

10. Data used to build a data mining model.

- A. training data
- B. validation data
- C. test data
- D. hidden data

[discuss](#)

A.training data

11. The problem of finding hidden structure in unlabeled data is called...

- A. supervised learning
- B. unsupervised learning
- C. reinforcement learning
- D. none of the above

[discuss](#)

B. unsupervised learning

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

- A. given email labeled as spam or not spam, learn a spam filter
- B. given a set of news articles found on the web, group them into set of articles about the same story.
- 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

[discuss](#)

A.given email labeled as spam or not spam, learn a spam filter

13. Dimensionality Reduction Algorithms are one of the possible ways to reduce the computation time required to build a model

- A. true
- B. false

[discuss](#)

A.true

14. 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. supervised learning
- B. unsupervised learning
- C. semisupervised learning
- D. reinforcement learning

[discuss](#)

A.supervised learning

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

- A. large enough to yield meaningful results
- B. is representative of the dataset as a whole
- C. both a and b
- D. none of the above

[discuss](#)

C.both a and b

16. Following are the types of supervised learning

- A. classification
- B. regression
- C. subgroup discovery
- D. all of the above

[discuss](#)

D.all of the above

17. Type of matrix decomposition model is

- A. descriptive model
- B. predictive model
- C. logical model
- D. none of the above

[discuss](#)

A.descriptive model

18. Following is powerful distance metrics used by Geometric model

- A. euclidean distance
- B. manhattan distance
- C. both a and b??
- D. square distance

[discuss](#)

C.both a and b??

19. The output of training process in machine learning is

- A. machine learning model
- B. machine learning algorithm
- C. null
- D. accuracy

[discuss](#)

A.machine learning model

20. 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

- A. nominal
- B. ordinal
- C. categorical
- D. boolean

[discuss](#)

B.ordinal

21. PCA is

- A. forward feature selection
- B. backword feature selection
- C. feature extraction
- D. all of the above

[discuss](#)

C.feature extraction

22. Dimensionality reduction algorithms are one of the possible ways to reduce the computation time required to build a model.

- A. true
- B. false

[discuss](#)

A.true

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

- A. removing columns which have too many missing values
- B. removing columns which have high variance in data
- C. removing columns with dissimilar data trends
- D. none of these

[discuss](#)

A.removing columns which have too many missing values

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

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

[discuss](#)

- C. input attribute.

25. What characterize is hyperplane in geometrical model of machine learning?

- A. a plane with 1 dimensional fewer than number of input attributes
- B. a plane with 2 dimensional fewer than number of input attributes
- C. a plane with 1 dimensional more than number of input attributes
- D. a plane with 2 dimensional more than number of input attributes

[discuss](#)

- B.a plane with 2 dimensional fewer than number of input attributes

1. What is true about Machine Learning?

- A. Machine Learning (ML) is that field of computer science
- B. ML is a type of artificial intelligence that extract patterns out of raw data by using an algorithm or method.
- C. The main focus of ML is to allow computer systems learn from experience without being explicitly programmed or human intervention.
- D. All of the above

Ans : D

Explanation: All statement are true about Machine Learning.

2. ML is a field of AI consisting of learning algorithms that?

- A. Improve their performance
- B. At executing some task
- C. Over time with experience
- D. All of the above

Ans : D

Explanation: ML is a field of AI consisting of learning algorithms that : Improve their performance (P), At executing some task (T), Over time with experience (E).

3. $p \rightarrow 0q$ is not a?

- A. hack clause
- B. horn clause
- C. structural clause
- D. system clause

Ans : B

Explanation: $p \rightarrow 0q$ is not a horn clause.

4. The action _____ of a robot arm specify to Place block A on block B.

- A. STACK(A,B)
- B. LIST(A,B)
- C. QUEUE(A,B)
- D. ARRAY(A,B)

Ans : A

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

5. A _____ begins by hypothesizing a sentence (the symbol S) and successively predicting lower level constituents until individual preterminal symbols are written.

- A. bottom-up parser
- B. top parser
- C. top-down parser
- D. bottom parser

Ans : C

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

6. A model of language consists of the categories which does not include _____.

- A. System Unit
- B. structural units.
- C. data units
- D. empirical units

Ans : B

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

7. Different learning methods does not include?

- A. Introduction
- B. Analogy
- C. Deduction
- D. Memorization

Ans : A

Explanation: Different learning methods does not include the introduction.

8. The model will be trained with data in one single batch is known as ?

- A. Batch learning
- B. Offline learning
- C. Both A and B
- D. None of the above

Ans : C

Explanation: we have end-to-end Machine Learning systems in which we need to train the model in one go by using whole available training data. Such kind of learning method or algorithm is called Batch or Offline learning.

9. Which of the following are ML methods?

- A. based on human supervision
- B. supervised Learning
- C. semi-reinforcement Learning
- D. All of the above

Ans : A

Explanation: The following are various ML methods based on some broad categories : Based on human supervision, Unsupervised Learning, Semi-supervised Learning and Reinforcement Learning

10. In Model based learning methods, an iterative process takes place on the ML models that are built based on various model parameters, called ?

- A. mini-batches
- B. optimizedparameters
- C. hyperparameters
- D. superparameters

Ans : C

Explanation: In Model based learning methods, an iterative process takes place on the ML models that are built based on various model parameters, called hyperparameters.

11. Which of the following is a widely used and effective machine learning algorithm based on the idea of bagging?

- A. Decision Tree
- B. Regression
- C. Classification
- D. Random Forest

Ans : D

Explanation: The Radom Forest algorithm builds an ensemble of Decision Trees, mostly trained with the bagging method.

12. 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

Explanation: The gradient of a multivariable function at a maximum point will be the zero vector of the function, which is the single greatest value that the function can achieve.

13. Which of the following is a disadvantage of decision trees?

- A. Factor analysis
- B. Decision trees are robust to outliers
- C. Decision trees are prone to be overfit
- D. None of the above

Ans : C

Explanation: Allowing a decision tree to split to a granular degree makes decision trees prone to learning every point extremely well to the point of perfect classification that is overfitting.

14. 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

Explanation: All of the above techniques are different ways of imputing the missing values.

15. 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

Explanation: You need to always normalize the data first. If not, PCA or other techniques that are used to reduce dimensions will give different results.

16. Which of the following statements about regularization is not correct?

- A. Using too large a value of lambda can cause your hypothesis to underfit the data.
- B. Using too large a value of lambda can cause your hypothesis to overfit the data
- C. Using a very large value of lambda cannot hurt the performance of your hypothesis.
- D. None of the above

Ans : D

Explanation: A large value results in a large regularization penalty and therefore, a strong preference for simpler models, which can underfit the data.

17. Which of the following techniques can not be used for normalization in text mining?

- A. Stemming
- B. Lemmatization
- C. Stop Word Removal
- D. None of the above

Ans : C

Explanation: Lemmatization and stemming are the techniques of keyword normalization.

18. In which of the following cases will K-means clustering fail to give good results?

- 1) Data points with outliers
 - 2) Data points with different densities
 - 3) Data points with nonconvex shapes
-
- A. 1 and 2
 - B. 2 and 3
 - C. 1 and 3
 - D. All of the above

Ans : D

Explanation: K-means clustering algorithm fails to give good results when the data contains outliers, the density spread of data points across the data space is different, and the data points follow nonconvex shapes.

19. 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 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

Explanation: This will maintain the structure of the data and also reduce its dimension.

20. What is a sentence parser typically used for?

- A. It is used to parse sentences to check if they are utf-8 compliant.
- B. It is used to parse sentences to derive their most likely syntax tree structures.
- C. It is used to parse sentences to assign POS tags to all tokens.
- D. It is used to check if sentences can be parsed into meaningful tokens.

Ans : B

Explanation: Sentence parsers analyze a sentence and automatically build a syntax tree.

01. What is Machine learning?

- 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

02. What is true about Machine Learning?

- A. Machine Learning (ML) is the field of computer science
- B. ML is a type of artificial intelligence that extract patterns out of raw data by using an algorithm or method
- C. The main focus of ML is to allow computer systems learn from experience without being explicitly programmed or human intervention
- D. All of the above

03. ML is a field of AI consisting of learning algorithms that?

- A. Improve their performance
- B. At executing some task
- C. Over time with experience
- D. All of the above

04. Different learning methods do not include?

- A. Memorization
- B. Analogy
- C. Introduction
- D. Deduction

05. Which of the following is a widely used and effective machine learning algorithm based on the idea of bagging?

- A. Decision Tree
- B. Random Forest
- C. Regression
- D. Classification

06. High entropy means that the partitions in classification are

- A. pure
- B. not pure
- C. useful
- D. useless

07. Which of the following are ML methods?

- A. Based on human supervision
- B. Supervised Learning
- C. Semi-reinforcement Learning
- D. All of the above

08. In language understanding, the levels of knowledge do not include?

- A. Phonological
- B. Syntactic
- C. Empirical
- D. Logical

09. A machine learning problem involves four attributes plus a class. The attributes have 3, 2, 2, and 2 possible values each. The class has 3 possible values. How many maximum possible different examples are there?

- A. 12
- B. 24
- C. 48
- D. 72

10. 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

11. 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

12. The most widely used metrics and tools to assess a classification model are:

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

13. A model of language consists of the categories which do not include?

- A. Language units
- B. Structural units
- C. Role structure of units
- D. System constraints

14. Suppose we would like to perform clustering on spatial data such as the geometrical locations of houses. We wish to produce clusters of many different sizes and shapes. Which of the following methods is the most appropriate?

- A. Decision Trees
- B. Model-based clustering
- C. K-means clustering
- D. Density-based clustering

15. Which of the following is a disadvantage of decision trees?

- A. Factor analysis
- B. Decision trees are robust to outliers
- C. Decision trees are prone to be overfit
- D. None of the above

16. Which of the following is true about Naive Bayes?

- A. Assumes that all the features in a dataset are equally important
- B. Assumes that all the features in a dataset are independent
- C. Both A and B
- D. None of the above options

17. Among the following which is not a horn clause?

- A. $p \rightarrow \emptyset q$
- B. p
- C. $p \rightarrow q$
- D. $\emptyset p \vee q$

18. Which of the following techniques can not be used for normalization in text mining?

- A. Stop Word Removal
- B. Stemming
- C. Lemmatization
- D. None of the above

19. 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 is retained
- B. Use the elbow method
- C. Choose k to be 99% of m ($k = 0.99*m$, rounded to the nearest integer)
- D. Choose k to be the largest value so that 99% of the variance is retained

20. In which of the following cases will K-means clustering fail to give good results?

- 1. Data points with outliers**
 - 2. Data points with different densities**
 - 3. Data points with nonconvex shapes**
- A. 1 & 2
B. 1, 2, & 3
C. 2 & 3
D. 1 & 3

1. What is machine learning?

- The selective acquisition of knowledge through the use of manual programs
- The selective acquisition of knowledge through the use of computer programs
- The autonomous acquisition of knowledge through the use of manual programs
- The autonomous acquisition of knowledge through the use of computer programs

2. Machine Learning is a field of AI consisting of learning algorithms that

- At executing some task
 - Over time with experience
 - Improve their performance
 - All of the above
-

3. is a widely used and effective machine learning algorithm based on the idea of bagging.

- Regression
 - Classification
 - Decision Tree
 - Random Forest
-

4. What is the disadvantage of decision trees?

- Factor analysis
 - Decision trees are robust to outliers
 - Decision trees are prone to be overfit
 - All of the above
-

5. How can you handle missing or corrupted data in a dataset?

- Drop missing rows or columns
- Assign a unique category to missing values
- Replace missing values with mean/median/mode
- All of the above

6. Which of the followings are most widely used metrics and tools to assess a classification model?

- Confusion matrix
- Cost-sensitive accuracy
- Area under the ROC curve
- All of the above

7. Machine learning algorithms build a model based on sample data, known as

- Training Data
- Transfer Data
- Data Training
- None of the above

8. Machine learning is a subset of

- Deep Learning
- Artificial Intelligence
- Data Learining
- None of the above

9. A Machine Learning technique that helps in detecting the outliers in data.

- Clustering
- Classification
- Anomaly Detection
- All of the above

10. Who is the father of Machine Learning?

- Geoffrey Hill
- Geoffrey Chaucer
- Geoffrey Everest Hinton
- None of the above

11. What is the most significant phase in a genetic algorithm?

- Selection
- Mutation
- Crossover
- Fitness function

12. Which one in the following is not Machine Learning disciplines?

- Physics
- Information Theory
- Neurostatistics
- Optimization Control

13. Machine Learning has various function representation, which of the following is not function of symbolic?

- Decision Trees
- Rules in proportional Logic
- Rules in first-order predicate logic
- Hidden-Markov Models (HMM)

14. algorithms enable the computers to learn from data, and even improve themselves, without being explicitly programmed.

- Deep Learning
- Machine Learning
- Artificial Intelligence
- None of the above

15. What are the three types of Machine Learning?

- Supervised Learning
- Unsupervised Learning
- Reinforcement Learning
- All of the above

16. Which of the following is not a supervised learning?

- PCA
- Naive Bayesian
- Linear Regression
- Decision Tree Answer

17. Real-Time decisions, Game AI, Learning Tasks, Skill acquisition, and Robot Navigation are applications of

- Reinforcement Learning
- Supervised Learning: Classification
- Unsupervised Learning: Regression
- None of the above

18. Which of the following is not numerical functions in the various function representation of Machine Learning?

- Case-based
- Neural Network
- Linear Regression
- Support Vector Machines

19. Common classes of problems in machine learning is

- Clustering
- Regression
- Classification
- All of the above

20. Which of the following clustering algorithm merges and splits nodes to help modify nonoptimal partitions?

- K-Means clustering
 - Conceptual clustering
 - Agglomerative clustering
 - All of the above
-

21. Missing data items are with Bayes classifier.

- Ignored
 - Treated as equal compares
 - Treated as unequal compares.
 - Replaced with a default value.
-

22. Which supervised learning technique can process both numeric and categorical input attributes?

- Bayes classifier
 - Linear regression
 - Logistic regression
 - None of the above
-

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

- Linear, binary
- Linear, numeric
- Nonlinear, binary
- Nonlinear, numeric

24. Regression trees are often used to model which data?

- Linear
- Nonlinear
- Categorical
- None of the above

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

- Mean relative error
- Mean squared error
- Mean absolute error
- Root mean squared error

26. Data used to optimize the parameter settings of a supervised learner model is called

- Test
- Training
- Validation
- None of the above

27. Bootstrapping allows us to choose the same training instance several times.

- True
- False

28. The average positive difference between computed and desired outcome values

- Mean positive error
- Mean absolute error
- Mean squared error
- Root mean squared error

-

29. Which of the following statement is true about prediction problems?

- The output attribute must be numeric.
 - The output attribute must be categorical
 - The resultant model is designed to determine future outcomes
 - The resultant model is designed to classify current behavior.
-

30. What is the another name for an output attribute?

- Predictive variable
 - Estimated variable
 - Dependent variable
 - Independent variable
-

31. Supervised learning and unsupervised clustering both require at least one

.....

- Input attribute
 - Output attribute
 - Hidden attribute
 - Categorical attribute
-

32. is not a machine learning algorithm.

- SVG
- SVM
- Random forest
- All of the above

33. Identify which is not machine learning disciplines?

- Physics
 - Information theory
 - Nuero Statistics
 - None of the above
-

34. What is the full form of PAC?

- Probably Approx Cost
 - Probably Approximate Correct
 - Probability Approx Communication
 - None of the above
-

35. Analysis of Machine Learning algorithm needs

- Statistical learning theory
- Computational learning theory
- Both Statistical & Computational learning theory
- None of the above

36. Choose the incorrect numerical functions in the various function representation of machine learning.

- Case-based
- Neural Network
- Linear regression
- All of true

37. What are successful applications of Machine Learning?

- Learning to recognize spoken words
 - Learning to drive an autonomous vehicle
 - Learning to classify new astronomical structures
 - All of the above
-

38. What is called the application of machine learning methods to large databases?

- Data mining
 - Internet of things
 - Artificial intelligence
 - None of the above
-

39. If machine learning model output involves target variable then that model is called as predictive model.

- True
 - False
-

40. are the best machine learning method.

- Fast
 - Accuracy
 - Scalable
 - All of the above
-

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

- Null
- Accuracy
- Machine learning model
- Machine learning algorithmp

42. A model of language consists of the categories, does not include

- Language units
- Structural units
- System constraints
- Role structure of units

43. Regression discovers causal relationships.

- True
- False

44. is the approach of basic algorithm for decision tree induction.

- Greedy
- Top Down
- Procedural
- Step by Step

45. What is the way to ensemble multiple classifications or regression?

- Bagging
- Blending
- Boosting
- Stacking

46. What is the most common issue when using Machine Learning?

- Poor Data Quality
- Lack of skilled resources
- Inadequate Infrastructure
- None of the above

47. In Machine learning the module that must solve the given performance task is known as

- Critic
 - Generalizer
 - Performance system
 - All of these
-

48. Which methods are used for the calibration in Supervised Learning?

- Platt Calibration
 - Isotonic Regression
 - Both Platt Calibration & Isotonic Regression
 - None of the above
-

49. How many types are available in machine learning?

- 2
 - 3
 - 4
 - 5
-

50. The Bayes rule can be used in

- Solving queries
- Increasing complexity
- Decreasing complexity
- Answering probabilistic query

1. What is Machine learning?

- 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

Answer: a

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

2. Which of the factors affect the performance of learner system does not include?

- a) Representation scheme used
- b) Training scenario
- c) Type of feedback
- d) Good data structures

Answer: d

Explanation: Factors that 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

Answer: d

Explanation: Different learning methods does not include the introduction.

4. In language understanding, the levels of knowledge that does not include?

- a) Phonological
- b) Syntactic
- c) Empirical
- d) Logical

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

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)

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) $\emptyset p \vee q$
- c) $p \rightarrow q$
- d) $p \rightarrow \emptyset q$

Answer: d

Explanation: $p \rightarrow \emptyset q$ 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

Answer: d

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

1. What is true about Machine Learning? 

- Machine Learning (ML) is that field of computer science
- ML is a type of artificial intelligence that extract patterns out of raw data by using an algorithm or method
- The main focus of ML is to allow computer systems learn from experience without being explicitly programmed or human intervention.
- All of the above

2. ML is a field of AI consisting of learning algorithms that? 

- Improve their performance
- At executing some task
- Over time with experience
- All of the above

3. The action _____ of a robot arm specify to Place block A on block B 

- STACK(A,B)
- LIST(A,B)
- QUEUE(A,B)
- ARRAY(A,B)

4. p ? 0q is not a? 

- hack clause
- horn clause
- structural clause
- system clause

5. A_____ begins by hypothesizing a sentence (the symbol S) and successively predicting lower level constituents until individual preterminal symbols are written. 

- bottom-up parser
- top parser
- top-down parser
- bottom parser

6. Different learning methods does not include? 

- Introduction
- Analogy
- Deduction
- Memorization

7. A model of language consists of the categories which does not include

-----.

- System Unit
- structural units.
- data units
- empirical units

8. Which of the following is a widely used and effective machine learning algorithm based on the idea of bagging?

- Decision Tree
- Regression
- Classification
- Random Forest

9. Which of the following are ML methods?

- based on human supervision
- supervised Learning
- semi-reinforcement Learning
- All of the above

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

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

11. The model will be trained with data in one single batch is known as ?

- Batch learning
- Offline learning
- Both A and B
- None of the above

12. In Model based learning methods, an iterative process takes place on the ML models that are built based on various model parameters, called ?

- mini-batches
- optimizedparameters
- hyperparameters
- superparameters

13. Which of the following statements about regularization is not correct?

- Using too large a value of lambda can cause your hypothesis to underfit the data.
- Using too large a value of lambda can cause your hypothesis to overfit the data
- Using a very large value of lambda cannot hurt the performance of your hypothesis
- None of the above

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

- Drop missing rows or columns
- Replace missing values with mean/median/mode
- Assign a unique category to missing values
- All of the above

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

- Normalize the data -> PCA -> training
- PCA -> normalize PCA output -> training
- Normalize the data -> PCA -> normalize PCA output -> training
- None of the above

16. Which of the following is a disadvantage of decision trees? 

- Factor analysis
- Decision trees are robust to outliers
- Decision trees are prone to be overfit
- None of the above

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

- Choose k to be the smallest value so that at least 99% of the variance is retained
- Choose k to be 99% of m ($k = 0.99*m$, rounded to the nearest integer).
- Choose k to be the largest value so that 99% of the variance is retained.
- Use the elbow method.

18. High entropy means that the partitions in classification are 

- pure
- not pure
- useful
- useless

19. What is a sentence parser typically used for? 

- It is used to parse sentences to check if they are utf-8 compliant.
- It is used to parse sentences to derive their most likely syntax tree structures.
- It is used to parse sentences to assign POS tags to all tokens.
- It is used to check if sentences can be parsed into meaningful tokens.

20. Which of the following techniques can not be used for normalization in text mining? 

- Stemming
- Lemmatization
- Stop Word Removal
- None of the above

21. Which of the following is NOT supervised learning? 

- PCA
- Decision Tree
- Linear Regression
- Naive Bayesian

22. Suppose we would like to perform clustering on spatial data such as the geometrical locations of houses. We wish to produce clusters of many different sizes and shapes. Which of the following methods is the most appropriate? 

- Decision Trees
- Density-based clustering
- Model-based clustering
- K-means clustering

23. What is the purpose of performing cross-validation? 

- To assess the predictive performance of the models
- To judge how the trained model performs outside the sample on test data
- both 1 and 2

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

- Drop missing rows or columns
- Replace missing values with mean/median/mode
- Assign a unique category to missing values
- All of the above -

MCQ Question of Machine learning

1. What is Machine Learning (ML)?
 - A. The autonomous acquisition of knowledge through the use of manual programs
 - B. The selective acquisition of knowledge through the use of computer programs
 - C. The selective acquisition of knowledge through the use of manual programs
 - D. The autonomous acquisition of knowledge through the use of computer programs

Correct option is D

2. Father of Machine Learning (ML)
 - A. Geoffrey Chaucer
 - B. Geoffrey Hill
 - C. Geoffrey Everest Hinton
 - D. None of the above

Correct option is C

3. Which is FALSE regarding regression?
 - A. It may be used for interpretation
 - B. It is used for prediction
 - C. It discovers causal relationships
 - D. It relates inputs to outputs

Correct option is C

4. Choose the correct option regarding machine learning (ML) and artificial intelligence (AI)
 - A. ML is a set of techniques that turns a dataset into a software
 - B. AI is a software that can emulate the human mind
 - C. ML is an alternate way of programming intelligent machines
 - D. All of the above

Correct option is D

5. Which of the factors affect the performance of the learner system does not include?
 - A. Good data structures
 - B. Representation scheme used
 - C. Training scenario
 - D. Type of feedback

Correct option is A

6. In general, to have a well-defined learning problem, we must identify which of the following
- The class of tasks
 - The measure of performance to be improved
 - The source of experience
 - All of the above

Correct option is D

7. Successful applications of ML
- Learning to recognize spoken words
 - Learning to drive an autonomous vehicle
 - Learning to classify new astronomical structures
 - Learning to play world-class backgammon
 - All of the above

Correct option is E

8. Which of the following does not include different learning methods
- Analogy
 - Introduction
 - Memorization
 - Deduction

Correct option is B

9. In language understanding, the levels of knowledge that does not include?
- Empirical
 - Logical
 - Phonological
 - Syntactic

Correct option is A

10. Designing a machine learning approach involves:-
- Choosing the type of training experience
 - Choosing the target function to be learned
 - Choosing a representation for the target function
 - Choosing a function approximation algorithm
 - All of the above

Correct option is E

11. Concept learning inferred a _____ valued function from training examples of its input and output.

- A. Decimal
- B. Hexadecimal
- C. Boolean
- D. All of the above

Correct option is C

12. Which of the following is not a supervised learning?

- A. Naive Bayesian
- B. PCA
- C. Linear Regression
- D. Decision Tree Answer

Correct option is B

13. What is Machine Learning?

- Artificial Intelligence
- Deep Learning
- Data Statistics
 - A. Only (i)
 - B. (i) and (ii)
 - C. All
 - D. None

Correct option is B

14. What kind of learning algorithm for “Facial identities or facial expressions”?

- A. Prediction
- B. Recognition Patterns
- C. Generating Patterns
- D. Recognizing Anomalies Answer

Correct option is B

15. Which of the following is not type of learning?

- A. Unsupervised Learning
- B. Supervised Learning
- C. Semi-unsupervised Learning
- D. Reinforcement Learning

Correct option is C

16. Real-Time decisions, Game AI, Learning Tasks, Skill Aquisition, and Robot Navigation are applications of which of the folowing

- A. Supervised Learning: Classification
- B. Reinforcement Learning
- C. Unsupervised Learning: Clustering
- D. Unsupervised Learning: Regression

Correct option is B

17. Targetted marketing, Recommended Systems, and Customer Segmentation are applications in which of the following

- A. Supervised Learning: Classification
- B. Unsupervised Learning: Clustering
- C. Unsupervised Learning: Regression
- D. Reinforcement Learning

Correct option is B

18. Fraud Detection, Image Classification, Diagnostic, and Customer Retention are applications in which of the following

- A. Unsupervised Learning: Regression
- B. Supervised Learning: Classification
- C. Unsupervised Learning: Clustering
- D. Reinforcement Learning

Correct option is B

19. Which of the following is not function of symbolic in the various function representation of Machine Learning?

- A. Rules in propotional Logic
- B. Hidden-Markov Models (HMM)
- C. Rules in first-order predicate logic
- D. Decision Trees

Correct option is B

20. Which of the following is not numerical functions in the various function representation of Machine Learning?

- A. Neural Network
- B. Support Vector Machines
- C. Case-based
- D. Linear Regression

Correct option is C

21. FIND-S Algorithm starts from the most specific hypothesis and generalize it by considering only

- A. Negative
- B. Positive
- C. Negative or Positive
- D. None of the above

Correct option is B

22. FIND-S algorithm ignores

- A. Negative
- B. Positive
- C. Both
- D. None of the above

Correct option is A

23. The Candidate-Elimination Algorithm represents the_.

- A. Solution Space
- B. Version Space
- C. Elimination Space
- D. All of the above

Correct option is B

24. Inductive learning is based on the knowledge that if something happens a lot it is likely to be generally

- A. True
- B. False Answer

Correct option is A

25. Inductive learning takes examples and generalizes rather than starting with _____

- A. Inductive
- B. Existing
- C. Deductive
- D. None of these

Correct option is B

26. A drawback of the FIND-S is that it assumes the consistency within the training set

- A. True
- B. False

Correct option is A

27.What strategies can help reduce overfitting in decision trees?

- Enforce a maximum depth for the tree
- Enforce a minimum number of samples in leaf nodes
- Pruning
- Make sure each leaf node is one pure class
 - A. All
 - B. (i), (ii) and (iii)
 - C. (i), (iii), (iv)
 - D. None

Correct option is B

28.Which of the following is a widely used and effective machine learning algorithm based on the idea of bagging?

- A. Decision Tree
- B. Random Forest
- C. Regression
- D. Classification

Correct option is B

29.To find the minimum or the maximum of a function, we set the gradient to zero because which of the following

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

Correct option is B

30.Which of the following is a disadvantage of decision trees?

- A. Decision trees are prone to be overfit
- B. Decision trees are robust to outliers
- C. Factor analysis
- D. None of the above

Correct option is A

31.What is perceptron?

- A. A single layer feed-forward neural network with pre-processing
- B. A neural network that contains feedback
- C. A double layer auto-associative neural network
- D. An auto-associative neural network

Correct option is A

32. Which of the following is true for neural networks?
- The training time depends on the size of the
 - Neural networks can be simulated on a conventional
 - Artificial neurons are identical in operation to biological
 - A. All
 - B. Only (ii)
 - C. (i) and (ii)
 - D. None

Correct option is C

33. What are the advantages of neural networks over conventional computers?
- They have the ability to learn by
 - They are more fault
 - They are more suited for real time operation due to their high „computational”
 - A. (i) and (ii)
 - B. (i) and (iii)
 - C. Only (i)
 - D. All
 - E. None

Correct option is D

34. What is Neuro software?
- A. It is software used by Neurosurgeon
 - B. Designed to aid experts in real world
 - C. It is powerful and easy neural network
 - D. A software used to analyze neurons

Correct option is C

35. Which is true for neural networks?
- A. Each node computes its weighted input
 - B. Node could be in excited state or non-excited state
 - C. It has set of nodes and connections
 - D. All of the above

Correct option is D

36.What is the objective of backpropagation algorithm?

- A. To develop learning algorithm for multilayer feedforward neural network, so that network can be trained to capture the mapping implicitly
- B. To develop learning algorithm for multilayer feedforward neural network
- C. To develop learning algorithm for single layer feedforward neural network
- D. All of the above

Correct option is A

37.Which of the following is true?

Single layer associative neural networks do not have the ability to:-

- Perform pattern recognition
- Find the parity of a picture
- Determine whether two or more shapes in a picture are connected or not
 - A. (ii) and (iii)
 - B. Only (ii)
 - C. All
 - D. None

Correct option is A

38.The backpropagation law is also known as generalized delta rule

- A. True
- B. False

Correct option is A

38.Which of the following is true?

- On average, neural networks have higher computational rates than conventional computers.
- Neural networks learn by
- Neural networks mimic the way the human brain
 - A. All
 - B. (ii) and (iii)
 - C. (i), (ii) and (iii)
 - D. None

Correct option is A

39.What is true regarding backpropagation rule?

- A. Error in output is propagated backwards only to determine weight updates
- B. There is no feedback of signal at nay stage
- C. It is also called generalized delta rule
- D. All of the above

Correct option is D

40.There is feedback in final stage of backpropagation

- A. True
- B. False

Correct option is B

41.An auto-associative network is

- A. A neural network that has only one loop
- B. A neural network that contains feedback
- C. A single layer feed-forward neural network with pre-processing
- D. A neural network that contains no loops

Correct option is B

42.A 3-input neuron has weights 1, 4 and 3. The transfer function is linear with the constant of proportionality being equal to 3. The inputs are 4, 8 and 5 respectively. What will be the output?

- A. 139
- B. 153
- C. 162
- D. 160

Correct option is B

43.What of the following is true regarding backpropagation rule?

- A. Hidden layers output is not all important, they are only meant for supporting input and output layers
- B. Actual output is determined by computing the outputs of units for each hidden layer
- C. It is a feedback neural network
- D. None of the above

Correct option is B

44.What is back propagation?

- A. It is another name given to the curvy function in the perceptron
- B. It is the transmission of error back through the network to allow weights to be adjusted so that the network can learn
- C. It is another name given to the curvy function in the perceptron
- D. None of the above

Correct option is B

45.The general limitations of back propagation rule is/are

- A. Scaling
- B. Slow convergence
- C. Local minima problem
- D. All of the above

Correct option is D

46.What is the meaning of generalized in statement “backpropagation is a generalized delta rule” ?

- A. Because delta is applied to only input and output layers, thus making it more simple and generalized
- B. It has no significance
- C. Because delta rule can be extended to hidden layer units
- D. None of the above

Correct option is C

47.Neural Networks are complex____functions with many parameter

- A. Linear
- B. Non linear
- C. Discrete
- D. Exponential

Correct option is A

48.The general tasks that are performed with backpropagation algorithm

- A. Pattern mapping
- B. Prediction
- C. Function approximation
- D. All of the above

Correct option is D

49. Backpropagation learning is based on the gradient descent along error surface.

- A. True
- B. False

Correct option is A

50. In backpropagation rule, how to stop the learning process?

- A. No heuristic criteria exist
- B. On basis of average gradient value
- C. There is convergence involved
- D. None of these

Correct option is B

51. Applications of NN (Neural Network)

- A. Risk management
- B. Data validation
- C. Sales forecasting
- D. All of the above

Correct option is D

52. The network that involves backward links from output to the input and hidden layers is known as

- A. Recurrent neural network
- B. Self organizing maps
- C. Perceptrons
- D. Single layered perceptron

Correct option is A

53. Decision Tree is a display of an Algorithm?

- A. True
- B. False

Correct option is A

54. Which of the following is/are the decision tree nodes?

- A. End Nodes
- B. Decision Nodes
- C. Chance Nodes
- D. All of the above

Correct option is D

55. End Nodes are represented by which of the following

- A. Solar street light
- B. Triangles
- C. Circles
- D. Squares

Correct option is B

56. Decision Nodes are represented by which of the following

- A. Solar street light
- B. Triangles
- C. Circles
- D. Squares

Correct option is D

57. Chance Nodes are represented by which of the following

- A. Solar street light
- B. Triangles
- C. Circles
- D. Squares

Correct option is C

58. Advantage of Decision Trees

- 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 above

Correct option is D

59. _____ terms are required for building a bayes model.

- A. 1
- B. 2
- C. 3
- D. 4

Correct option is C

60. Which of the following is the consequence between a node and its predecessors while creating bayesian network?

- A. Conditionally independent
- B. Functionally dependent
- C. Both Conditionally dependant & Dependant
- D. Dependent

Correct option is A

61. Why it is needed to make probabilistic systems feasible in the world?

- A. Feasibility
- B. Reliability
- C. Crucial robustness
- D. None of the above

Correct option is C

62. Bayes rule can be used for:-

- A. Solving queries
- B. Increasing complexity
- C. Answering probabilistic query
- D. Decreasing complexity

Correct option is C

63. _____ provides way and means of weighing up the desirability of goals and the likelihood of achieving

- A. Utility theory
- B. Decision theory
- C. Bayesian networks
- D. Probability theory

Correct option is A

64. Which of the following provided by the Bayesian Network?

- A. Complete description of the problem
- B. Partial description of the domain
- C. Complete description of the domain
- D. All of the above

Correct option is C

65. Probability provides a way of summarizing the _____ that comes from our laziness and

- A. Belief
- B. Uncertainty
- C. Joint probability distributions
- D. Randomness

Correct option is B

66. The entries in the full joint probability distribution can be calculated as

- A. Using variables
- B. Both Using variables & information
- C. Using information
- D. All of the above

Correct option is C

67. Causal chain (For example, Smoking cause cancer) gives rise to:-

- A. Conditionally Independence
- B. Conditionally Dependence
- C. Both
- D. None of the above

Correct option is A

68. The bayesian network can be used to answer any query by using:-

- A. Full distribution
- B. Joint distribution
- C. Partial distribution
- D. All of the above

Correct option is B

69. Bayesian networks allow compact specification of:-

- A. Joint probability distributions
- B. Belief
- C. Propositional logic statements
- D. All of the above

Correct option is A

70. The compactness of the bayesian network can be described by

- A. Fully structured
- B. Locally structured
- C. Partially structured
- D. All of the above

Correct option is B

71. The Expectation-Maximization Algorithm has been used to identify conserved domains in unaligned proteins only. State True or False.

- A. True
- B. False

Correct option is B

72. Which of the following is correct about the Naive Bayes?

- A. Assumes that all the features in a dataset are independent
- B. Assumes that all the features in a dataset are equally important
- C. Both
- D. All of the above

Correct option is C

73. Which of the following is false regarding EM Algorithm?

- A. The alignment provides an estimate of the base or amino acid composition of each column in the site
- B. The column-by-column composition of the site already available is used to estimate the probability of finding the site at any position in each of the sequences
- C. The row-by-column composition of the site already available is used to estimate the probability
- D. None of the above

Correct option is C

74. Naïve Bayes Algorithm is a learning algorithm.

- A. Supervised
- B. Reinforcement
- C. Unsupervised
- D. None of these

Correct option is A

75. EM algorithm includes two repeated steps, here the step 2 is_____.

- A. The normalization
- B. The maximization step
- C. The minimization step
- D. None of the above

Correct option is C

76. Examples of Naïve Bayes Algorithm is/are

- A. Spam filtration
- B. Sentimental analysis
- C. Classifying articles
- D. All of the above

Correct option is D

77. In the intermediate steps of “EM Algorithm”, the number of each

base in each column is determined and then converted to

- A. True
- B. False

Correct option is A

78. Naïve Bayes algorithm is based on and used for solving
classification problems.

- A. Bayes Theorem
- B. Candidate elimination algorithm
- C. EM algorithm
- D. None of the above

Correct option is A

79. Types of Naïve Bayes Model:

- A. Gaussian
- B. Multinomial
- C. Bernoulli
- D. All of the above

Correct option is D

80. Disadvantages of Naïve Bayes Classifier:

- A. Naive Bayes assumes that all features are independent or unrelated, so it cannot learn the relationship between
- B. It performs well in Multi-class predictions as compared to the other
- C. Naïve Bayes is one of the fast and easy ML algorithms to predict a class of
- D. It is the most popular choice for text classification problems.

Correct option is A

81. The benefit of Naïve Bayes:-

- A. Naïve Bayes is one of the fast and easy ML algorithms to predict a class of
- B. It is the most popular choice for text classification problems.
- C. It can be used for Binary as well as Multi-class
- D. All of the above

Correct option is D

82. In which of the following types of sampling the information is carried out under the opinion of an expert?

- A. Convenience sampling
- B. Judgement sampling
- C. Quota sampling
- D. Purposive sampling

Correct option is B

83. Full form of MDL?

- A. Minimum Description Length
- B. Maximum Description Length
- C. Minimum Domain Length
- D. None of these

Correct option is A

84. For the analysis of ML algorithms, we need

- A. Computational learning theory
- B. Statistical learning theory
- C. Both A & B
- D. None of these

Correct option is C

85. PAC stand for

- A. Probably Approximate Correct
- B. Probably Approx Correct
- C. Probably Approximate Computation
- D. Probably Approx Computation

Correct option is A

86. _____ hypothesis h with respect to target concept c and distribution D , is the probability that h will misclassify an instance drawn at random according to D .

- A. True Error
- B. Type 1 Error
- C. Type 2 Error
- D. None of these

Correct option is A

87. Statement: True error defined over entire instance space, not just training data

- A. True
- B. False

Correct option is A

88. What are the area CLT comprised of?

- A. Sample Complexity
- B. Computational Complexity
- C. Mistake Bound
- D. All of these

Correct option is D

88. What area of CLT tells "How many examples we need to find a good hypothesis?"?

- A. Sample Complexity
- B. Computational Complexity
- C. Mistake Bound
- D. None of these

Correct option is A

89.What area of CLT tells "How much computational power we need to find a good hypothesis ?"?

- A. Sample Complexity
- B. Computational Complexity
- C. Mistake Bound
- D. None of these

Correct option is B

90.What area of CLT tells "How many mistakes we will make before finding a good hypothesis ?"?

- A. Sample Complexity
- B. Computational Complexity
- C. Mistake Bound
- D. None of these

Correct option is C

91.(For question no. 9 and 10) Can we say that concept described by conjunctions of Boolean literals are PAC learnable?

- A. Yes
- B. No

Correct option is A

92.How large is the hypothesis space when we have n Boolean attributes?

- A. $|H| = 3^n$
- B. $|H| = 2^n$
- C. $|H| = 1^n$
- D. $|H| = 4^n$

Correct option is A

93.The VC dimension of hypothesis space H1 is larger than the VC dimension of hypothesis space H2. Which of the following can be inferred from this?

- A. The number of examples required for learning a hypothesis in H1 is larger than the number of examples required for H2
- B. The number of examples required for learning a hypothesis in H1 is smaller than the number of examples required for
- C. No relation to number of samples required for PAC learning.

Correct option is A

94. For a particular learning task, if the requirement of error parameter changes from 0.1 to 0.01. How many more samples will be required for PAC learning?

- A. Same
- B. 2 times
- C. 1000 times
- D. 10 times

Correct option is D

95. Computational complexity of classes of learning problems depends on which of the following?

- A. The size or complexity of the hypothesis space considered by learner
- B. The accuracy to which the target concept must be approximated
- C. The probability that the learner will output a successful hypothesis
- D. All of these

Correct option is D

96. The instance-based learner is a _____

- A. Lazy-learner
- B. Eager learner
- C. Can't say

Correct option is A

97. When to consider nearest neighbour algorithms?

- A. Instance map to point in k^n
- B. Not more than 20 attributes per instance
- C. Lots of training data
- D. None of these
- E. A, B & C

Correct option is E

98. What are the advantages of Nearest neighbour algo?

- A. Training is very fast
- B. Can learn complex target functions
- C. Don't lose information
- D. All of these

Correct option is D

99. What are the difficulties with k-nearest neighbour algo?

- A. Calculate the distance of the test case from all training cases
- B. Curse of dimensionality
- C. Both A & B
- D. None of these

Correct option is C

100. What if the target function is real valued in kNN algo?

- A. Calculate the mean of the k nearest neighbours
- B. Calculate the SD of the k nearest neighbour
- C. None of these

Correct option is A

101. What is/are true about Distance-weighted KNN?

- A. The weight of the neighbour is considered
- B. The distance of the neighbour is considered
- C. Both A & B
- D. None of these

Correct option is C

102. What is/are advantage(s) of Distance-weighted k-NN over k-NN?

- A. Robust to noisy training data
- B. Quite effective when a sufficient large set of training data is provided
- C. Both A & B
- D. None of these

Correct option is C

103. What is/are advantage(s) of Locally Weighted Regression?

- A. Pointwise approximation of complex target function
- B. Earlier data has no influence on the new ones
- C. Both A & B
- D. None of these

Correct option is C

104. The quality of the result depends on (LWR)
- A. Choice of the function
 - B. Choice of the kernel function K
 - C. Choice of the hypothesis space H
 - D. All of these

Correct option is D

105. How many types of layer in radial basis function neural networks?
- A. 3
 - B. 2
 - C. 1
 - D. 4

Correct option is A, Input layer, Hidden layer, and Output layer

106. The neurons in the hidden layer contains Gaussian transfer function whose output are _____ to the distance from the centre of the neuron.
- A. Directly
 - B. Inversely
 - C. equal
 - D. None of these

Correct option is B

107. PNN/GRNN networks have one neuron for each point in the training file, While RBF network have a variable number of neurons that is usually
- A. less than the number of training
 - B. greater than the number of training points
 - C. equal to the number of training points
 - D. None of these

Correct option is A

108. Which network is more accurate when the size of training set between small to medium?
- A. PNN/GRNN
 - B. RBF
 - C. K-means clustering
 - D. None of these

Correct option is A

109. What is/are true about RBF network?

- A. A kind of supervised learning
- B. Design of NN as curve fitting problem
- C. Use of multidimensional surface to interpolate the test data
- D. All of these

Correct option is D

110. Application of CBR

- A. Design
- B. Planning
- C. Diagnosis
- D. All of these

Correct option is A

111. What is/are advantages of CBR?

- A. A local approx. is found for each test case
- B. Knowledge is in a form understandable to human
- C. Fast to train
- D. All of these

Correct option is D

112. In k-NN algorithm, given a set of training examples and the value of $k < \text{size of training set } (n)$, the algorithm predicts the class of a test example to be the.
What is/are advantages of CBR?

- A. Least frequent class among the classes of k closest training
- B. Most frequent class among the classes of k closest training
- C. Class of the closest
- D. Most frequent class among the classes of the k farthest training examples.

Correct option is B

113. Which of the following statements is true about PCA?

- We must standardize the data before applying
- We should select the principal components which explain the highest variance
- We should select the principal components which explain the lowest variance
- We can use PCA for visualizing the data in lower dimensions
 - A. (i), (ii) and (iv).
 - B. (ii) and (iv)
 - C. (iii) and (iv)
 - D. (i) and (iii)

Correct option is A

114. Genetic algorithm is a

- A. Search technique used in computing to find true or approximate solution to optimization and search problem
- B. Sorting technique used in computing to find true or approximate solution to optimization and sort problem
- C. Both A & B
- D. None of these

Correct option is A

115. GA techniques are inspired by

- A. Evolutionary
- B. Cytology
- C. Anatomy
- D. Ecology

Correct option is A

116. When would the genetic algorithm terminate?

- A. Maximum number of generations has been produced
- B. Satisfactory fitness level has been reached for the
- C. Both A & B
- D. None of these

Correct option is C

117. The algorithm operates by iteratively updating a pool of hypotheses, called the

- A. Population
- B. Fitness
- C. None of these

Correct option is A

118. What is the correct representation of GA?

- A. GA(Fitness, Fitness_threshold, p)
- B. GA(Fitness, Fitness_threshold, p, r)
- C. GA(Fitness, Fitness_threshold, p, r, m)
- D. GA(Fitness, Fitness_threshold)

Correct option is C

119. Genetic operators includes

- A. Crossover
- B. Mutation
- C. Both A & B
- D. None of these

Correct option is C

120. Produces two new offspring from two parent string by copying selected bits from each parent is called

- A. Mutation
- B. Inheritance
- C. Crossover
- D. None of these

Correct option is C

121. Each schema the set of bit strings containing the indicated as

- A. 0s, 1s
- B. only 0s
- C. only 1s
- D. 0s, 1s, *s

Correct option is D

122. 0^*10 represents the set of bit strings that includes exactly (A) 0010, 0110

- A. 0010, 0010
- B. 0100, 0110
- C. 0100, 0010

Correct option is A

123. Correct (h) is the percent of all training examples correctly classified by hypothesis then Fitness function is equal to

- A. Fitness (h) = $(\text{correct} (h))^2$
- B. Fitness (h) = $(\text{correct} (h))^3$
- C. Fitness (h) = $(\text{correct} (h))$
- D. Fitness (h) = $(\text{correct} (h))^4$

Correct option is A

124. Statement: Genetic Programming individuals in the evolving population are computer programs rather than bit

- A. True
- B. False

Correct option is A

125. _____ evolution over many generations was directly influenced by the experiences of individual organisms during their lifetime

- A. Baldwin
- B. Lamarckian
- C. Bayes
- D. None of these

Correct option is B

126. Search through the hypothesis space cannot be characterized. Why?

- A. Hypotheses are created by crossover and mutation operators that allow radical changes between successive generations
- B. Hypotheses are not created by crossover and mutation
- C. None of these

Correct option is A

127. ILP stand for

- A. Inductive Logical programming
- B. Inductive Logic Programming
- C. Inductive Logical Program
- D. Inductive Logic Program

Correct option is B

128. What is/are the requirement for the Learn-One-Rule method?

- A. Input, accepts a set of +ve and -ve training examples.
- B. Output, delivers a single rule that covers many +ve examples and few -ve.
- C. Output rule has a high accuracy but not necessarily a high
- D. A & B
- E. A, B & C

Correct option is E

129. _____ is any predicate (or its negation) applied to any set of terms.

- A. Literal
- B. Null
- C. Clause
- D. None of these

Correct option is A

130. Ground literal is a literal that

- A. Contains only variables
- B. does not contains any functions
- C. does not contains any variables
- D. Contains only functions Answer

Correct option is C

131. _____ emphasizes learning feedback that evaluates the learner's performance without providing standards of correctness in the form of behavioural

- A. Reinforcement learning
- B. Supervised Learning
- C. None of these

Correct option is A

132. Features of Reinforcement learning

- A. Set of problem rather than set of techniques
- B. RL is training by reward and
- C. RL is learning from trial and error with the
- D. All of these

Correct option is D

133. Which type of feedback used by RL?
- A. Purely Instructive feedback
 - B. Purely Evaluative feedback
 - C. Both A & B
 - D. None of these

Correct option is B

134. What is/are the problem solving methods for RL?
- A. Dynamic programming
 - B. Monte Carlo Methods
 - C. Temporal-difference learning
 - D. All of these

Correct option is D

135. The FIND-S Algorithm
- A. Starts with starts from the most specific hypothesis
Answer
 - B. It considers negative examples
 - C. It considers both negative and positive
 - D. None of these Correct

136. The hypothesis space has a general-to-specific ordering of hypotheses, and the search can be efficiently organized by taking advantage of a naturally occurring structure over the hypothesis space

- A. TRUE
- B. FALSE

Correct option is A

137. The Version space is:

- A. The subset of all hypotheses is called the version space with respect to the hypothesis space H and the training examples D, because it contains all plausible versions of the target
- B. The version space consists of only specific
- C. None of these

Correct option is A

138. The Candidate-Elimination Algorithm
- A. The key idea in the Candidate-Elimination algorithm is to output a description of the set of all hypotheses consistent with the training
 - B. Candidate-Elimination algorithm computes the description of this set without explicitly enumerating all of its
 - C. This is accomplished by using the more-general-than partial ordering and maintaining a compact representation of the set of consistent
 - D. All of these

Correct option is D

139. Concept learning is basically acquiring the definition of a general category from given sample positive and negative training examples of the
- A. TRUE
 - B. FALSE

Correct option is A

140. The hypothesis h_1 is more-general-than hypothesis h_2 ($h_1 > h_2$) if and only if $h_1 \geq h_2$ is true and $h_2 \geq h_1$ is false. We also say h_2 is more-specific-than h_1
- A. The statement is true
 - B. The statement is false
 - C. We cannot
 - D. None of these

Correct option is A

141. The List-Then-Eliminate Algorithm
- A. The List-Then-Eliminate algorithm initializes the version space to contain all hypotheses in H , then eliminates any hypothesis found inconsistent with any training
 - B. The List-Then-Eliminate algorithm not initializes to the version
 - C. None of these Answer

Correct option is A

142. What will take place as the agent observes its interactions with the world?

- A. Learning
- B. Hearing
- C. Perceiving
- D. Speech

Correct option is A

143. Which modifies the performance element so that it makes better decision? Performance element

- A. Performance element
- B. Changing element
- C. Learning element
- D. None of the mentioned

Correct option is C

144. Any hypothesis found to approximate the target function well over a sufficiently large set of training examples will also approximate the target function well over other unobserved example is called:

- A. Inductive Learning Hypothesis
- B. Null Hypothesis
- C. Actual Hypothesis
- D. None of these

Correct option is A

145. Feature of ANN in which ANN creates its own organization or representation of information it receives during learning time is

- A. Adaptive Learning
- B. Self Organization
- C. What-If Analysis
- D. Supervised Learning

Correct option is B

146. How the decision tree reaches its decision?

- A. Single test
- B. Two test
- C. Sequence of test
- D. No test

Correct option is C

147. Which of the following is a disadvantage of decision trees?
- Factor analysis
 - Decision trees are robust to outliers
 - Decision trees are prone to be overfit
 - None of the above

Correct option is C

148. Tree/Rule based classification algorithms generate which rule to perform the classification.
- A. if-then.
 - B. then
 - C. do
 - D. Answer

Correct option is A

149. What is Gini Index?
- A. It is a type of index structure
 - B. It is a measure of purity
 - C. None of the options

Correct option is A

150. What is not a RNN in machine learning?
- A. One output to many inputs
 - B. Many inputs to a single output
 - C. RNNs for nonsequential input
 - D. Many inputs to many outputs

Correct option is A

151. Which of the following sentences are correct in reference to Information gain?
- A. It is biased towards multi-valued attributes
 - B. ID3 makes use of information gain
 - C. The approach used by ID3 is greedy
 - D. All of these

Correct option is D

152. A Neural Network can answer
- A. For Loop questions
 - B. what-if questions
 - C. IF-The-Else Analysis Questions
 - D. None of these Answer

Correct option is B

153. Artificial neural network used for
- A. Pattern Recognition
 - B. Classification
 - C. Clustering
 - D. All Answer

Correct option is D

154. Which of the following are the advantage/s of Decision Trees?
- 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

Correct option is D

155. What is the mathematical likelihood that something will occur?
- A. Classification
 - B. Probability
 - C. Naïve Bayes Classifier
 - D. None of the other

Correct option is C

- 156.What does the Bayesian network provides?
- A. Complete description of the domain
 - B. Partial description of the domain
 - C. Complete description of the problem
 - D. None of the mentioned

Correct option is C

157. Where does the Bayes rule can be used?
- A. Solving queries
 - B. Increasing complexity
 - C. Decreasing complexity
 - D. Answering probabilistic query

Correct option is D

158. How many terms are required for building a Bayes model?
- A. 2
 - B. 3
 - C. 4
 - D. 1

Correct option is B

159. What is needed to make probabilistic systems feasible in the world?
- A. Reliability
 - B. Crucial robustness
 - C. Feasibility
 - D. None of the mentioned

Correct option is B

160. It was shown that the Naive Bayesian method
- A. Can be much more accurate than the optimal Bayesian method
 - B. Is always worse off than the optimal Bayesian method
 - C. Can be almost optimal only when attributes are independent
 - D. Can be almost optimal when some attributes are dependent

Correct option is C

161. What is the consequence between a node and its predecessors while creating Bayesian network?
- A. Functionally dependent
 - B. Dependant
 - C. Conditionally independent
 - D. Both Conditionally dependant & Dependant

Correct option is C

162. How the compactness of the Bayesian network can be described?

- A. Locally structured
- B. Fully structured
- C. Partial structure
- D. All of the mentioned

Correct option is A

163. How the entries in the full joint probability distribution can be calculated?

- A. Using variables
- B. Using information
- C. Both Using variables & information
- D. None of the mentioned

Correct option is B

164. How the Bayesian network can be used to answer any query?

- A. Full distribution
- B. Joint distribution
- C. Partial distribution
- D. All of the mentioned

Correct option is B

165. Sample Complexity is

- A. The sample complexity is the number of training-samples that we need to supply to the algorithm, so that the function returned by the algorithm is within an arbitrarily small error of the best possible function, with probability arbitrarily close to 1
- B. How many training examples are needed for learner to converge to a successful hypothesis.
- C. All of these

Correct option is C

166. PAC stands for

- A. Probability Approximately Correct
- B. Probability Applied Correctly
- C. Partition Approximately Correct

Correct option is A

167. Which of the following will be true about k in k-NN in terms of variance

- A. When you increase the k the variance will increases
- B. When you decrease the k the variance will increases
- C. Can't say
- D. None of these

Correct option is B

168. Which of the following option is true about k-NN algorithm?

- A. It can be used for classification
- B. It can be used for regression
- C. It can be used in both classification and regression

Answer

Correct option is C

169. In k-NN it is very likely to overfit due to the curse of dimensionality. Which of the following option would you consider to handle such problem? 1). Dimensionality Reduction 2). Feature selection

- A. 1
- B. 2
- C. 1 and 2
- D. None of these

Correct option is C

170. When you find noise in data which of the following option would you consider in k- NN

- A. I will increase the value of k
- B. I will decrease the value of k
- C. Noise can not be dependent on value of k
- D. None of these

Correct option is A

171. Which of the following will be true about k in k-NN in terms of Bias?

- A. When you increase the k the bias will be increases
- B. When you decrease the k the bias will be increases
- C. Can't say
- D. None of these

Correct option is A

172. What is used to mitigate overfitting in a test set?
- A. Overfitting set
 - B. Training set
 - C. Validation dataset
 - D. Evaluation set

Correct option is C

173. A radial basis function is a
- A. Activation function
 - B. Weight
 - C. Learning rate
 - D. none

Correct option is A

174. Mistake Bound is
- A. How many training examples are needed for learner to converge to a successful hypothesis.
 - B. How much computational effort is needed for a learner to converge to a successful hypothesis
 - C. How many training examples will the learner misclassify before converging to a successful hypothesis
 - D. None of these

Correct option is C

175. All of the following are suitable problems for genetic algorithms EXCEPT
- A. dynamic process control
 - B. pattern recognition with complex patterns
 - C. simulation of biological models
 - D. simple optimization with few variables

Correct option is D

176. Adding more basis functions in a linear model... (Pick the most probably option)
- A. Decreases model bias
 - B. Decreases estimation bias
 - C. Decreases variance
 - D. Doesn't affect bias and variance

Correct option is A

177. Which of these are types of crossover
- A. Single point
 - B. Two point
 - C. Uniform
 - D. All of these

Correct option is D

178. A feature F1 can take certain value: A, B, C, D, E, & F and represents grade of students from a college. Which of the following statement is true in following case?
- A. Feature F1 is an example of nominal
 - B. Feature F1 is an example of ordinal
 - C. It doesn't belong to any of the above category.

Correct option is B

179. You observe the following while fitting a linear regression to the data: As you increase the amount of training data, the test error decreases and the training error increases. The train error is quite low (almost what you expect it to), while the test error is much higher than the train error. What do you think is the main reason behind this behaviour? Choose the most probable option.
- A. High variance
 - B. High model bias
 - C. High estimation bias
 - D. None of the above Answer

Correct option is C

180. Genetic algorithms are heuristic methods that do not guarantee an optimal solution to a problem
- A. TRUE
 - B. FALSE

Correct option is A

181. Which of the following statements about regularization is not correct?
- A. Using too large a value of lambda can cause your hypothesis to underfit the
 - B. Using too large a value of lambda can cause your hypothesis to overfit the
 - C. Using a very large value of lambda cannot hurt the performance of your hypothesis.
 - D. None of the above

Correct option is A

182. Consider the following: (a) Evolution (b) Selection (c) Reproduction (d) Mutation Which of the following are found in genetic algorithms?

- A. All
- B. a, b, c
- C. a, b
- D. b, d

Correct option is A

183. Genetic Algorithm are a part of

- A. Evolutionary Computing
- B. inspired by Darwin's theory about evolution – "survival of the fittest"
- C. are adaptive heuristic search algorithm based on the evolutionary ideas of natural selection and genetics
- D. All of the above

Correct option is D

184. Genetic algorithms belong to the family of methods in the

- A. artificial intelligence area
- B. optimization
- C. complete enumeration family of methods
- D. Non-computer based (human) solutions area

Correct option is A

185. For a two player chess game, the environment encompasses the opponent

- A. True
- B. False

Correct option is A

186. Which among the following is not a necessary feature of a reinforcement learning solution to a learning problem?

- A. exploration versus exploitation dilemma
- B. trial and error approach to learning
- C. learning based on rewards
- D. representation of the problem as a Markov Decision Process

Correct option is D

187. Which of the following sentence is FALSE regarding reinforcement learning

- A. It relates inputs to
- B. It is used for
- C. It may be used for
- D. It discovers causal relationships.

Correct option is D

188. The EM algorithm is guaranteed to never decrease the value of its objective function on any iteration

- A. TRUE
- B. FALSE Answer

Correct option is A

189. Consider the following modification to the tic-tac-toe game: at the end of game, a coin is tossed and the agent wins if a head appears regardless of whatever has happened in the game. Can reinforcement learning be used to learn an optimal policy of playing Tic-Tac-Toe in this case?

- A. Yes
- B. No

Correct option is B

190. Out of the two repeated steps in EM algorithm, the step 2 is

- A. the maximization step
- B. the minimization step
- C. the optimization step
- D. the normalization step

Correct option is A

191. Suppose the reinforcement learning player was greedy, that is, it always played the move that brought it to the position that it rated the best. Might it learn to play better, or worse, than a non greedy player?

- A. Worse
- B. Better

Correct option is B

192. A chess agent trained by using Reinforcement Learning can be trained by playing against a copy of the same
- A. True
 - B. False

Correct option is A

193. The EM iteration alternates between performing an expectation (E) step, which creates a function for the expectation of the log-likelihood evaluated using the current estimate for the parameters, and a maximization (M) step, which computes parameters maximizing the expected log-likelihood found on the E
- A. TRUE
 - B. FALSE

Correct option is A

194. Expectation–maximization (EM) algorithm is an
- A. Iterative
 - B. Incremental
 - C. None

Correct option is A

195. Feature need to be identified by using Well Posed Learning Problem:
- A. Class of tasks
 - B. Performance measure
 - C. Training experience
 - D. All of these

Correct option is D

196. A computer program that learns to play checkers might improve its performance as:
- A. Measured by its ability to win at the class of tasks involving playing checkers
 - B. Experience obtained by playing games against
 - C. Both a & b
 - D. None of these

Correct option is C

197. Learning symbolic representations of concepts known as:
- A. Artificial Intelligence
 - B. Machine Learning
 - C. Both a & b
 - D. None of these

Correct option is A

198. The field of study that gives computers the capability to learn without being explicitly programmed __
- A. Machine Learning
 - B. Artificial Intelligence
 - C. Deep Learning
 - D. Both a & b

Correct option is A

199. The autonomous acquisition of knowledge through the use of computer programs is called __
- A. Artificial Intelligence
 - B. Machine Learning
 - C. Deep learning
 - D. All of these

Correct option is B

200. Learning that enables massive quantities of data is known as
- A. Artificial Intelligence
 - B. Machine Learning
 - C. Deep learning
 - D. All of these

Correct option is B

201. A different learning method does not include
- A. Memorization
 - B. Analogy
 - C. Deduction
 - D. Introduction

Correct option is D

202. Types of learning used in machine
- A. Supervised
 - B. Unsupervised
 - C. Reinforcement
 - D. All of these

Correct option is D

203. A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience
- A. Supervised learning problem
 - B. Un Supervised learning problem
 - C. Well posed learning problem
 - D. All of these

Correct option is C

204. Which of the following is a widely used and effective machine learning algorithm based on the idea of bagging?
- A. Decision Tree
 - B. Regression
 - C. Classification
 - D. Random Forest

Correct option is D

205. How many types are available in machine learning?
- A. 1
 - B. 2
 - C. 3
 - D. 4

Correct option is C

205. A model can learn based on the rewards it received for its previous action is known as:
- A. Supervised learning
 - B. Unsupervised learning
 - C. Reinforcement learning
 - D. Concept learning

Correct option is C

206. A subset of machine learning that involves systems that think and learn like humans using artificial neural networks.

- A. Artificial Intelligence
- B. Machine Learning
- C. Deep Learning
- D. All of these

Correct option is C

207. A learning method in which a training data contains a small amount of labeled data and a large amount of unlabeled data is known as _____

- A. Supervised Learning
- B. Semi Supervised Learning
- C. Unsupervised Learning
- D. Reinforcement Learning

Correct option is C

208. Methods used for the calibration in Supervised Learning

- A. Platt Calibration
- B. Isotonic Regression
- C. All of these
- D. None of above

Correct option is C

209. The basic design issues for designing a learning

- A. Choosing the Training Experience
- B. Choosing the Target Function
- C. Choosing a Function Approximation Algorithm
- D. Estimating Training Values
- E. All of these

Correct option is E

210. In Machine learning the module that must solve the given performance task is known as:

- A. Critic
- B. Generalizer
- C. Performance system
- D. All of these

Correct option is C

211. A learning method that is used to solve a particular computational program, multiple models such as classifiers or experts are strategically generated and combined is called as __

- A. Supervised Learning
- B. Semi Supervised Learning
- C. Unsupervised Learning
- D. Reinforcement Learning
- E. Ensemble learning

Correct option is E

212. In a learning system the component that takes as input the current hypothesis (currently learned function) and outputs a new problem for the Performance System to explore.

- A. Critic
- B. Generalizer
- C. Performance system
- D. Experiment generator
- E. All of these

Correct option is D

213. Learning method that is used to improve the classification, prediction, function approximation etc of a model

- A. Supervised Learning
- B. Semi Supervised Learning
- C. Unsupervised Learning
- D. Reinforcement Learning
- E. Ensemble learning

Correct option is E

214. In a learning system the component that takes as input the history or trace of the game and produces as output a set of training examples of the target function is known as:

- A. Critic
- B. Generalizer
- C. Performance system
- D. All of these

Correct option is A

215. The most common issue when using ML is
- A. Lack of skilled resources
 - B. Inadequate Infrastructure
 - C. Poor Data Quality
 - D. None of these

Correct option is C

216. How to ensure that your model is not over fitting
- A. Cross validation
 - B. Regularization
 - C. All of these
 - D. None of these

Correct option is C

217. A way to ensemble multiple classifications or regression
- A. Stacking
 - B. Bagging
 - C. Blending
 - D. Boosting

Correct option is A

218. How well a model is going to generalize in new environment is known as
- A. Data Quality
 - B. Transparent
 - C. Implementation
 - D. None of these

Correct option is B

219. Common classes of problems in machine learning is _____
- A. Classification
 - B. Clustering
 - C. Regression
 - D. All of these

Correct option is D

220. Which of the following is a widely used and effective machine learning algorithm based on the idea of bagging?

- A. Decision Tree
- B. Regression
- C. Classification
- D. Random Forest

Correct option is D

221. Cost complexity pruning algorithm is used in?

- A. CART
- B. 5
- C. ID3
- D. All of

Correct option is A

222. Which one of these is not a tree based learner?

- A. CART
- B. 5
- C. ID3
- D. Bayesian Classifier

Correct option is D

223. Which one of these is a tree based learner?

- A. Rule based
- B. Bayesian Belief Network
- C. Bayesian classifier
- D. Random Forest

Correct option is D

224. What is the approach of basic algorithm for decision tree induction?

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

Correct option is A

225. Which of the following classifications would best suit the student performance classification systems?

- A. If-.then-analysis
- B. Market-basket analysis
- C. Regression analysis
- D. Cluster analysis

Correct option is A

226. What are two steps of tree pruning work?

- A. Pessimistic pruning and Optimistic pruning
- B. Post pruning and Pre pruning
- C. Cost complexity pruning and time complexity pruning
- D. None of these

Correct option is B

227. How will you counter over-fitting in decision tree?

- A. By pruning the longer rules
- B. By creating new rules
- C. Both By pruning the longer rules" and „ By creating new rules"
- D. None of Answer

Correct option is A

228. Which of the following sentences are true?

- A. In pre-pruning a tree is 'pruned' by halting its construction early
- B. A pruning set of class labeled tuples is used to estimate cost
- C. The best pruned tree is the one that minimizes the number of encoding
- D. All of these

Correct option is D

229. Which of the following is a disadvantage of decision trees?

- A. Factor analysis
- B. Decision trees are robust to outliers
- C. Decision trees are prone to be over fit
- D. None of the above

Correct option is C

230.In which of the following scenario a gain ratio is preferred over Information Gain?

- A. When a categorical variable has very large number of category
- B. When a categorical variable has very small number of category
- C. Number of categories is the not the reason
- D. None of these

Correct option is A

231. Major pruning techniques used in decision tree are

- A. Minimum error
- B. Smallest tree
- C. Both a & b
- D. None of these

Correct option is B

232.What does the central limit theorem state?

- A. If the sample size increases sampling distribution must approach normal distribution
- B. If the sample size decreases then the sample distribution must approach normal distribution.
- C. If the sample size increases then the sampling distributions much approach an exponential
- D. If the sample size decreases then the sampling distributions much approach an exponential

Correct option is A

233. The difference between the sample value expected and the estimates value of the parameter is called as?

- A. Bias
- B. Error
- C. Contradiction
- D. Difference

Correct option is A

234. In which of the following types of sampling the information is carried out under the opinion of an expert?

- A. Quota sampling
- B. Convenience sampling
- C. Purposive sampling
- D. Judgment sampling

Correct option is D

235. Which of the following is a subset of population?
- A. Distribution
 - B. Sample
 - C. Data
 - D. Set

Correct option is B

236. The sampling error is defined as?
- A. Difference between population and parameter
 - B. Difference between sample and parameter
 - C. Difference between population and sample
 - D. Difference between parameter and sample

Correct option is C

237. Machine learning is interested in the best hypothesis h from some space H , given observed training data D . Here best hypothesis means
- A. Most general hypothesis
 - B. Most probable hypothesis
 - C. Most specific hypothesis
 - D. None of these

Correct option is B

238. Practical difficulties with Bayesian Learning :
- A. Initial knowledge of many probabilities is required
 - B. No consistent hypothesis
 - C. Hypotheses make probabilistic predictions
 - D. None of these

Correct option is A

239. Bayes' theorem states that the relationship between the probability of the hypothesis before getting the evidence $P(H)$ and the probability of the hypothesis after getting the evidence $P(H|E)$ is
- A. $[P(E|H)P(H)] / P(E)$
 - B. $[P(E|H) P(E)] / P(H)$
 - C. $[P(E) P(H)] / P(E|H)$
 - D. None of these

Correct option is A

240. A doctor knows that Cold causes fever 50% of the time. Prior probability of any patient having cold is 1/50,000. Prior probability of any patient having fever is 1/20. If a patient has fever, what is the probability he/she has cold?

- A. $P(C/F) = 0.0003$
- B. $P(C/F) = 0.0004$
- C. $P(C/F) = 0.0002$
- D. $P(C/F) = 0.0045$

Correct option is C

241. Which of the following will be true about k in K-Nearest Neighbor in terms of Bias?

- A. When you increase the k the bias will be increases
- B. When you decrease the k the bias will be increases
- C. Can't say
- D. None of these

Correct option is A

242. When you find noise in data which of the following option would you consider in K- Nearest Neighbor?

- A. I will increase the value of k
- B. I will decrease the value of k
- C. Noise cannot be dependent on value of k
- D. None of these

Correct option is A

243. In K-Nearest Neighbor it is very likely to overfit due to the curse of dimensionality. Which of the following option would you consider to handle such problem?

- Dimensionality Reduction
- Feature selection
 - A. 1
 - B. 2
 - C. 1 and 2
 - D. None of these

Correct option is C

244. Radial basis functions is closely related to distance-weighted regression, but it is

- A. lazy learning
- B. eager learning
- C. concept learning
- D. none of these

Correct option is B

245. Radial basis function networks provide a global approximation to the target function, represented by _____ of many local kernel function.

- A. a series combination
- B. a linear combination
- C. a parallel combination
- D. a non linear combination

Correct option is B

246. The most significant phase in a genetic algorithm is

- A. Crossover
- B. Mutation
- C. Selection
- D. Fitness function

Correct option is A

247. The crossover operator produces two new offspring from

- A. Two parent strings, by copying selected bits from each parent
- B. One parent strings, by copying selected bits from selected parent
- C. Two parent strings, by copying selected bits from one parent
- D. None of these

Correct option is A

248. Mathematically characterize the evolution over time of the population within a GA based on the concept of

- A. Schema
- B. Crossover
- C. Don't care
- D. Fitness function

Correct option is A

249. In genetic algorithm process of selecting parents which mate and recombine to create off-springs for the next generation is known as:

- A. Tournament selection
- B. Rank selection
- C. Fitness sharing
- D. Parent selection

Correct option is D

250. Crossover operations are performed in genetic programming by replacing

- A. Randomly chosen sub tree of one parent program by a sub tree from the other parent program.
- B. Randomly chosen root node tree of one parent program by a sub tree from the other parent program
- C. Randomly chosen root node tree of one parent program by a root node tree from the other parent program
- D. None of these

Correct option is A

Unit-1

1. What is Machine Learning (ML)?

- (A) The autonomous acquisition of knowledge through the use of manual programs
- (B) The selective acquisition of knowledge through the use of computer programs
- (C) The selective acquisition of knowledge through the use of manual programs
- (D) The autonomous acquisition of knowledge through the use of computer programs

Answer

Correct option is D

2. Father of Machine Learning (ML)

- (A) Geoffrey Chaucer
- (B) Geoffrey Hill
- (C) Geoffrey Everest Hinton
- (D) None of the above

Answer

Correct option is C

3. Which is FALSE regarding regression?

- (A) It may be used for interpretation
- (B) It is used for prediction
- (C) It discovers causal relationships
- (D) It relates inputs to outputs

Answer

Correct option is C

4. Choose the correct option regarding machine learning (ML) and artificial intelligence (AI)

- (A) ML is a set of techniques that turns a dataset into a software
- (B) AI is a software that can emulate the human mind
- (C) ML is an alternate way of programming intelligent machines
- (D) All of the above

Answer

Correct option is D

5. Which of the factors affect the performance of the learner system does not include?

- (A) Good data structures
- (B) Representation scheme used
- (C) Training scenario
- (D) Type of feedback

Correct option is A

6. In general, to have a well-defined learning problem, we must identify which of the following

- (A) The class of tasks
- (B) The measure of performance to be improved
- (C) The source of experience
- (D) All of the above

Answer

Correct option is D

7. Successful applications of ML

- (A) Learning to recognize spoken words
- (B) Learning to drive an autonomous vehicle
- (C) Learning to classify new astronomical structures
- (D) Learning to play world-class backgammon
- (E) All of the above

Answer

Correct option is E

8. Which of the following does not include different learning methods

- (A) Analogy
- (B) Introduction
- (C) Memorization
- (D) Deduction

Answer

Correct option is B

9. In language understanding, the levels of knowledge that does not include?

- (A) Empirical
- (B) Logical
- (C) Phonological
- (D) Syntactic

Answer

Correct option is A

10. Designing a machine learning approach involves:-

- (A) Choosing the type of training experience
- (B) Choosing the target function to be learned
- (C) Choosing a representation for the target function
- (D) Choosing a function approximation algorithm
- (E) All of the above

Answer

Correct option is E

11. Concept learning inferred a _____ valued function from training examples of its input and output.

- (A) Decimal
- (B) Hexadecimal
- (C) Boolean
- (D) All of the above

Answer

Correct option is C

12. Which of the following is not a supervised learning?

- (A) Naive Bayesian
- (B) PCA
- (C) Linear Regression
- (D) Decision Tree

Answer

Correct option is B

13. What is Machine Learning?

- (i) Artificial Intelligence
 - (ii) Deep Learning
 - (iii) Data Statistics
- (A) Only (i)
 - (B) (i) and (ii)
 - (C) All
 - (D) None

Answer

Correct option is B

14. What kind of learning algorithm for "Facial identities or facial expressions"?

- (A) Prediction
- (B) Recognition Patterns
- (C) Generating Patterns
- (D) Recognizing Anomalies

Answer

Correct option is B

15. Which of the following is not type of learning?

- (A) Unsupervised Learning
- (B) Supervised Learning
- (C) Semi-unsupervised Learning
- (D) Reinforcement Learning

Answer

Correct option is C

16. Real-Time decisions, Game AI, Learning Tasks, Skill Aquisition, and Robot Navigation are applications of which of the folowing

(A) Supervised Learning: Classification

(B) Reinforcement Learning

(C) Unsupervised Learning: Clustering

(D) Unsupervised Learning: Regression

Answer

Correct option is B

17. Targetted marketing, Recommended Systems, and Customer Segmentation are applications in which of the following

(A) Supervised Learning: Classification

(B) Unsupervised Learning: Clustering

(C) Unsupervised Learning: Regression

(D) Reinforcement Learning

Answer

Correct option is B

18. Fraud Detection, Image Classification, Diagnostic, and Customer Retention are applications in which of the following

(A) Unsupervised Learning: Regression

(B) Supervised Learning: Classification

(C) Unsupervised Learning: Clustering

(D) Reinforcement Learning

Answer

Correct option is B

19. Which of the following is not function of symbolic in the various function representation of Machine Learning?

- (A) Rules in propositional Logic
- (B) Hidden-Markov Models (HMM)
- (C) Rules in first-order predicate logic
- (D) Decision Trees

Answer

Correct option is B

20. Which of the following is not numerical functions in the various function representation of Machine Learning?

- (A) Neural Network
- (B) Support Vector Machines
- (C) Case-based
- (D) Linear Regression

Answer

Correct option is C

21. FIND-S Algorithm starts from the most specific hypothesis and generalize it by considering only _____ examples.

- (A) Negative
- (B) Positive
- (C) Negative or Positive
- (D) None of the above

Answer

Correct option is B

22. FIND-S algorithm ignores _____ examples.

- (A) Negative
- (B) Positive
- (C) Both
- (D) None of the above

Answer

Correct option is A

23. The Candidate-Elimination Algorithm represents the _____.

- (A) Solution Space
- (B) Version Space
- (C) Elimination Space
- (D) All of the above

Answer

Correct option is B

24. Inductive learning is based on the knowledge that if something happens a lot it is likely to be generally.

- (A) True
- (B) False

Answer

Correct option is A

25. Inductive learning takes examples and generalizes rather than starting with _____ knowledge.

- (A) Inductive
- (B) Existing
- (C) Deductive
- (D) None of these

Answer

Correct option is B

26. A drawback of the FIND-S is that, it assumes the consistency within the training set.

- (A) True
- (B) False

Answer

Correct option is A

Unit-2

1. What strategies can help reduce overfitting in decision trees?

- (i) Enforce a maximum depth for the tree
 - (ii) Enforce a minimum number of samples in leaf nodes
 - (iii) Pruning
 - (iv) Make sure each leaf node is one pure class
- (A) All
 - (B) (i), (ii) and (iii)
 - (C) (i), (iii), (iv)
 - (D) None

Answer

Correct option is B

2. Which of the following is a widely used and effective machine learning algorithm based on the idea of bagging?

- (A) Decision Tree
- (B) Random Forest
- (C) Regression
- (D) Classification

Answer

Correct option is B

3. To find the minimum or the maximum of a function, we set the gradient to zero because which of the following

- (A) Depends on the type of problem
- (B) The value of the gradient at extrema of a function is always zero
- (C) Both (A) and (B)
- (D) None of these

Answer

Correct option is B

4. Which of the following is a disadvantage of decision trees?

- (A) Decision trees are prone to be overfit
- (B) Decision trees are robust to outliers
- (C) Factor analysis
- (D) None of the above

Answer

Correct option is A

5. What is perceptron?

- (A) A single layer feed-forward neural network with pre-processing
- (B) A neural network that contains feedback

(C) A double layer auto-associative neural network

(D) An auto-associative neural network

Answer

Correct option is A

6. Which of the following is true for neural networks?

- (i) The training time depends on the size of the network.
 - (ii) Neural networks can be simulated on a conventional computer.
 - (iii) Artificial neurons are identical in operation to biological ones.
- (A) All
(B) Only (ii)
(C) (i) and (ii)
(D) None

Answer

Correct option is C

7. What are the advantages of neural networks over conventional computers?

- (i) They have the ability to learn by example.
 - (ii) They are more fault tolerant.
 - (iii) They are more suited for real time operation due to their high ‘computational’ rates.
- (A) (i) and (ii)
(B) (i) and (iii)
(C) Only (i)
(D) All
(E) None

Answer

Correct option is D

8. What is Neuro software?

- (A) It is software used by Neurosurgeon
(B) Designed to aid experts in real world
(C) It is powerful and easy neural network
(D) A software used to analyze neurons

Answer

Correct option is C

9. Which is true for neural networks?

- (A) Each node computes its weighted input
(B) Node could be in excited state or non-excited state
(C) It has set of nodes and connections
(D) All of the above

Answer

Correct option is D

10. What is the objective of backpropagation algorithm?

- (A) To develop learning algorithm for multilayer feedforward neural network, so that
-

network can be trained to capture the mapping implicitly

- (B) To develop learning algorithm for multilayer feedforward neural network
(C) To develop learning algorithm for single layer feedforward neural network
(D) All of the above

Answer

Correct option is A

11. Which of the following is true?

Single layer associative neural networks do not have the ability to:-

- (i) Perform pattern recognition
(ii) Find the parity of a picture
(iii) Determine whether two or more shapes in a picture are connected or not
(A) (ii) and (iii)
(B) Only (ii)
(C) All
(D) None

Answer

Correct option is A

12. The backpropagation law is also known as generalized delta rule.

- (A) True

- (B) False

Answer

Correct option is A

13. Which of the following is true?

- (i) On average, neural networks have higher computational rates than conventional computers.

- (ii) Neural networks learn by example.

- (iii) Neural networks mimic the way the human brain works.

- (A) All

- (B) (ii) and (iii)

- (C) (i), (ii) and (iii)

- (D) None

Answer

Correct option is A

14. What is true regarding backpropagation rule?

- (A) Error in output is propagated backwards only to determine weight updates
- (B) There is no feedback of signal at nay stage
- (C) It is also called generalized delta rule
- (D) All of the above

Answer

Correct option is D

15. There is feedback in final stage of backpropagation algorithm.

(A) True

(B) False

Answer

Correct option is B

16. An auto-associative network is

- (A) A neural network that has only one loop
- (B) A neural network that contains feedback
- (C) A single layer feed-forward neural network with pre-processing
- (D) A neural network that contains no loops

Answer

Correct option is B

17. A 3-input neuron has weights 1, 4 and 3. The transfer function is linear with the constant of proportionality being equal to 3. The inputs are 4, 8 and 5 respectively.

What will be the output?

(A) 139

(B) 153

(C) 612

(D) 160

Answer

Correct option is B

18. What of the following is true regarding backpropagation rule?

- (A) Hidden layers output is not all important, they are only meant for supporting input and output layers
- (B) Actual output is determined by computing the outputs of units for each hidden layer
- (C) It is a feedback neural network
- (D) None of the above

Answer

Correct option is B

19. What is back propagation?

- (A) It is another name given to the curvy function in the perceptron
- (B) It is the transmission of error back through the network to allow weights to be adjusted so that the network can learn
- (C) It is another name given to the curvy function in the perceptron
- (D) None of the above

Answer

Correct option is B

20. The general limitations of back propagation rule is/are

- (A) Scaling
- (B) Slow convergence
- (C) Local minima problem
- (D) All of the above

Answer

Correct option is D

21. What is the meaning of generalized in statement “backpropagation is a generalized delta rule” ?

- (A) Because delta is applied to only input and output layers, thus making it more simple and generalized
- (B) It has no significance
- (C) Because delta rule can be extended to hidden layer units
- (D) None of the above

Answer

Correct option is C

22. Neural Networks are complex _____ functions with many parameters.

- (A) Linear
- (B) Non linear
- (C) Discrete
- (D) Exponential

Answer

Correct option is A

23. The general tasks that are performed with backpropagation algorithm

- (A) Pattern mapping
- (B) Prediction
- (C) Function approximation
- (D) All of the above

Answer

Correct option is D

24. Backpropagation learning is based on the gradient descent along error surface.

- (A) True
- (B) False

Answer

Correct option is A

25. In backpropagation rule, how to stop the learning process?

- (A) No heuristic criteria exist
- (B) On basis of average gradient value
- (C) There is convergence involved
- (D) None of these

Answer

Correct option is B

26. Applications of NN (Neural Network)

- (A) Risk management
- (B) Data validation
- (C) Sales forecasting
- (D) All of the above

Answer

Correct option is D

27. The network that involves backward links from output to the input and hidden layers is known as

- (A) Recurrent neural network
- (B) Self organizing maps
- (C) Perceptrons
- (D) Single layered perceptron

Answer

Correct option is A

28. Decision Tree is a display of an algorithm.

- (A) True
- (B) False

Answer

Correct option is A

29. Which of the following is/are the decision tree nodes?

- (A) End Nodes
- (B) Decision Nodes
- (C) Chance Nodes
- (D) All of the above

Answer

Correct option is D

30. End Nodes are represented by which of the following

- (A) Solar street light
- (B) Triangles
- (C) Circles
- (D) Squares

Answer

Correct option is B

31. Decision Nodes are represented by which of the following

- (A) Solar street light
- (B) Triangles
- (C) Circles
- (D) Squares

Answer

Correct option is D

32. Chance Nodes are represented by which of the following

- (A) Solar street light
- (B) Triangles
- (C) Circles
- (D) Squares

Answer

Correct option is C

33. Advantage of Decision Trees

- (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 above

Answer

Correct option is D

Unit-3

1. _____ terms are required for building a bayes model.

- (A) 1
- (B) 2
- (C) 3
- (D) 4

Answer

Correct option is C

2. Which of the following is the consequence between a node and its predecessors while creating bayesian network?

- (A) Conditionally independent
- (B) Functionally dependent
- (C) Both Conditionally dependant & Dependant
- (D) Dependent

Answer

Correct option is A

3. Why it is needed to make probabilistic systems feasible in the world?

- (A) Feasibility
- (B) Reliability
- (C) Crucial robustness
- (D) None of the above

Answer

Correct option is C

4. Bayes rule can be used for:-

- (A) Solving queries
- (B) Increasing complexity
- (C) Answering probabilistic query
- (D) Decreasing complexity

Answer

Correct option is C

5. _____ provides way and means of weighing up the desirability of goals and the likelihood of achieving them.

- (A) Utility theory
- (B) Decision theory
- (C) Bayesian networks
- (D) Probability theory

Answer

Correct option is A

6. Which of the following provided by the Bayesian Network?

- (A) Complete description of the problem
- (B) Partial description of the domain
- (C) Complete description of the domain
- (D) All of the above

Answer

Correct option is C

7. Probability provides a way of summarizing the _____ that comes from our laziness and ignorances.

- (A) Belief
- (B) Uncertainty
- (C) Joint probability distributions
- (D) Randomness

Answer

Correct option is B

8. The entries in the full joint probability distribution can be calculated as

- (A) Using variables
- (B) Both Using variables & information
- (C) Using information
- (D) All of the above

Answer

Correct option is C

9. Causal chain (For example, Smoking cause cancer) gives rise to:-

- (A) Conditionally Independence
- (B) Conditionally Dependence
- (C) Both
- (D) None of the above

Answer

Correct option is A

10. The bayesian network can be used to answer any query by using:-

- (A) Full distribution
- (B) Joint distribution
- (C) Partial distribution
- (D) All of the above

Answer

Correct option is B

11. Bayesian networks allow compact specification of:-

- (A) Joint probability distributions
- (B) Belief
- (C) Propositional logic statements
- (D) All of the above

Answer

Correct option is A

12. The compactness of the bayesian network can be described by
- (A) Fully structured
 - (B) Locally structured
 - (C) Partially structured
 - (D) All of the above

Answer

Correct option is B

13. The Expectation Maximization Algorithm has been used to identify conserved domains in unaligned proteins only. State True or False.

- (A) True
- (B) False

Answer

Correct option is B

14. Which of the following is correct about the Naive Bayes?

- (A) Assumes that all the features in a dataset are independent
- (B) Assumes that all the features in a dataset are equally important
- (C) Both
- (D) All of the above

Answer

Correct option is C

15. Which of the following is false regarding EM Algorithm?

- (A) The alignment provides an estimate of the base or amino acid composition of each column in the site
- (B) The column-by-column composition of the site already available is used to estimate the probability of finding the site at any position in each of the sequences
- (C) The row-by-column composition of the site already available is used to estimate the probability
- (D) None of the above

Answer

Correct option is C

16. Naïve Bayes Algorithm is a _____ learning algorithm.

- (A) Supervised
- (B) Reinforcement
- (C) Unsupervised
- (D) None of these

Answer

Correct option is A

17. EM algorithm includes two repeated steps, here the step 2 is _____.

- (A) The normalization
- (B) The maximization step
- (C) The minimization step
- (D) None of the above

Answer

Correct option is C

18. Examples of Naïve Bayes Algorithm is/are

- (A) Spam filtration
- (B) Sentimental analysis
- (C) Classifying articles
- (D) All of the above

Answer

Correct option is D

19. In the intermediate steps of "EM Algorithm", the number of each base in each column is determined and then converted to fractions.

- (A) True
- (B) False

Answer

Correct option is A

20. Naïve Bayes algorithm is based on _____ and used for solving classification problems.

- (A) Bayes Theorem
- (B) Candidate elimination algorithm
- (C) EM algorithm
- (D) None of the above

Answer

Correct option is A

21. Types of Naïve Bayes Model:

- (A) Gaussian
- (B) Multinomial
- (C) Bernoulli
- (D) All of the above

Answer

Correct option is D

22. Disadvantages of Naïve Bayes Classifier:

- (A) Naive Bayes assumes that all features are independent or unrelated, so it cannot learn the relationship between features.
- (B) It performs well in Multi-class predictions as compared to the other Algorithms.
- (C) Naïve Bayes is one of the fast and easy ML algorithms to predict a class of datasets.
- (D) It is the most popular choice for text classification problems.

Answer

Correct option is A

23. The benefit of Naïve Bayes:-

- (A) Naïve Bayes is one of the fast and easy ML algorithms to predict a class of datasets.
- (B) It is the most popular choice for text classification problems.
- (C) It can be used for Binary as well as Multi-class Classifications.
- (D) All of the above

Answer

Correct option is D

24. In which of the following types of sampling the information is carried out under the opinion of an expert?

- (A) Convenience sampling
- (B) Judgement sampling
- (C) Quota sampling
- (D) Purposive sampling

Answer

Correct option is B

25. Full form of MDL.

- (A) Minimum Description Length
- (B) Maximum Description Length
- (C) Minimum Domain Length
- (D) None of these

Answer

Correct option is A

Unit-4

1. For the analysis of ML algorithms, we need
- (A) Computational learning theory
 - (B) Statistical learning theory
 - (C) Both A & B
 - (D) None of these

Answer

Correct option is C

2. PAC stand for
- (A) Probably Approximate Correct
 - (B) Probably Approx Correct
 - (C) Probably Approximate Computation
 - (D) Probably Approx Computation

Answer

Correct option is A

3. _____ of hypothesis h with respect to target concept c and distribution D , is the probability that h will misclassify an instance drawn at random according to D
- (A) True Error
 - (B) Type 1 Error
 - (C) Type 2 Error
 - (D) None of these

Answer

Correct option is A

4. Statement: True error defined over entire instance space, not just training data
- (A) True
 - (B) False

Answer

Correct option is A

5. What are the area CLT comprised of?
- (A) Sample Complexity
 - (B) Computational Complexity
 - (C) Mistake Bound
 - (D) All of these

Correct option is D

6. What area of CLT tells “How many examples we need to find a good hypothesis ??”?

- (A) Sample Complexity
- (B) Computational Complexity
- (C) Mistake Bound
- (D) None of these

Answer

Correct option is A

7. What area of CLT tells “How much computational power we need to find a good hypothesis ??”?

- (A) Sample Complexity
- (B) Computational Complexity
- (C) Mistake Bound
- (D) None of these

Answer

Correct option is B

8. What area of CLT tells “How many mistakes we will make before finding a good hypothesis ??”?

- (A) Sample Complexity
- (B) Computational Complexity
- (C) Mistake Bound
- (D) None of these

Answer

Correct option is C

9. (For question no. 9 and 10) Can we say that concept described by conjunctions of Boolean literals are PAC learnable?

- (A) Yes
- (B) No

Answer

Correct option is A

10. How large is the hypothesis space when we have n Boolean attributes?

- (A) $|H| = 3^n$
- (B) $|H| = 2^n$
- (C) $|H| = 1^n$
- (D) $|H| = 4^n$

Answer

Correct option is A

11. The VC dimension of hypothesis space H1 is larger than the VC dimension of hypothesis space H2. Which of the following can be inferred from this?
- (A) The number of examples required for learning a hypothesis in H1 is larger than the number of examples required for H2
 - (B) The number of examples required for learning a hypothesis in H1 is smaller than the number of examples required for H2.
 - (C) No relation to number of samples required for PAC learning.

Answer

Correct option is A

12. For a particular learning task, if the requirement of error parameter changes from 0.1 to 0.01. How many more samples will be required for PAC learning?
- (A) Same
 - (B) 2 times
 - (C) 1000 times
 - (D) 10 times

Answer

Correct option is D

13. Computational complexity of classes of learning problems depends on which of the following?
- (A) The size or complexity of the hypothesis space considered by learner
 - (B) The accuracy to which the target concept must be approximated
 - (C) The probability that the learner will output a successful hypothesis
 - (D) All of these

Answer

Correct option is D

14. The instance-based learner is a _____
- (A) Lazy-learner
 - (B) Eager learner
 - (C) Can't say

Answer

Correct option is A

15. When to consider nearest neighbour algorithms?

- (A) Instance map to point in k^n
- (B) Not more than 20 attributes per instance

(C) Lots of training data

(D) None of these

(E) A, B & C

Answer

Correct option is E

16. What are the advantages of Nearest neighbour algo?

- (A) Training is very fast
- (B) Can learn complex target functions
- (C) Don't lose information
- (D) All of these

Answer

Correct option is D

17. What are the difficulties with k-nearest neighbour algo?

- (A) Calculate the distance of the test case from all training cases
- (B) Curse of dimensionality
- (C) Both A & B
- (D) None of these

Answer

Correct option is C

18. What if the target function is real valued in kNN algo?

- (A) Calculate the mean of the k nearest neighbours
- (B) Calculate the SD of the k nearest neighbour
- (C) None of these

Answer

Correct option is A

19. What is/are true about Distance-weighted KNN?
- (A) The weight of the neighbour is considered
 - (B) The distance of the neighbour is considered
 - (C) Both A & B
 - (D) None of these

Answer

Correct option is C

20. What is/are advantage(s) of Distance-weighted k-NN over k-NN?
- (A) Robust to noisy training data
 - (B) Quite effective when a sufficient large set of training data is provided

-
- (C) Both A & B
 - (D) None of these

Answer

Correct option is C

21. What is/are advantage(s) of Locally Weighted Regression?
- (A) Pointwise approximation of complex target function
 - (B) Earlier data has no influence on the new ones
 - (C) Both A & B
 - (D) None of these

Answer

Correct option is C

22. The quality of the result depends on (LWR)
- (A) Choice of the function
 - (B) Choice of the kernel function K
 - (C) Choice of the hypothesis space H
 - (D) All of these

Answer

Correct option is D

23. How many types of layer in radial basis function neural networks?

- (A) 3
- (B) 2
- (C) 1
- (D) 4

Answer

Correct option is A, Input layer, Hidden layer, and Output layer

24. The neurons in the hidden layer contains Gaussian transfer function whose output are _____ to the distance from the centre of the neuron.

- (A) Directly
- (B) Inversely
- (C) equal
- (D) None of these

Answer

Correct option is B

25. PNN/GRNN networks have one neuron for each point in the training file.

While RBF network have a variable number of neurons that is usually

- (A) less than the number of training points.

- (B) greater than the number of training points

- (C) equal to the number of training points

- (D) None of these

Answer

Correct option is A

26. Which network is more accurate when the size of training set between small to medium?

- (A) PNN/GRNN
- (B) RBF
- (C) K-means clustering
- (D) None of these

Answer

Correct option is A

27. What is/are true about RBF network?

- (A) A kind of supervised learning
- (B) Design of NN as curve fitting problem
- (C) Use of multidimensional surface to interpolate the test data
- (D) All of these

Answer

Correct option is D

28. Application of CBR

- (A) Design
- (B) Planning
- (C) Diagnosis
- (D) All of these

Answer

Correct option is A

29. What is/are advantages of CBR?

- (A) A local approx. is found for each test case
- (B) Knowledge is in a form understandable to human
- (C) Fast to train
- (D) All of these

Answer

Correct option is D

30 In k-NN algorithm, given a set of training examples and the value of $k < \text{size of training set } (n)$, the algorithm predicts the class of a test example to be the.

What is/are advantages of CBR?

- (A) Least frequent class among the classes of k closest training examples.
- (B) Most frequent class among the classes of k closest training examples.
- (C) Class of the closest point.
- (D) Most frequent class among the classes of the k farthest training examples.

Answer

Correct option is B

31. Which of the following statements is true about PCA?

- (i) We must standardize the data before applying PCA.
 - (ii) We should select the principal components which explain the highest variance
 - (iii) We should select the principal components which explain the lowest variance
 - (iv) We can use PCA for visualizing the data in lower dimensions
- (A) (i), (ii) and (iv).
 - (B) (ii) and (iv)
 - (C) (iii) and (iv)
 - (D) (i) and (iii)

Answer

Correct option is A

UNIT-5

1. Genetic algorithm is a

- (A) Search technique used in computing to find true or approximate solution to optimization and search problem
- (B) Sorting technique used in computing to find true or approximate solution to optimization and sort problem
- (C) Both A & B
- (D) None of these

Correct option is A

2. GA techniques are inspired by _____ biology.

- (A) Evolutionary
- (B) Cytology
- (C) Anatomy
- (D) Ecology

Correct option is A

3. When would the genetic algorithm terminate?

- (A) Maximum number of generations has been produced
- (B) Satisfactory fitness level has been reached for the population.
- (C) Both A & B
- (D) None of these

Answer

Correct option is C

4. The algorithm operates by iteratively updating a pool of hypotheses, called the

- (A) Population
- (B) Fitness
- (C) None of these

Answer

Correct option is A

5. What is the correct representation of GA?

- (A) GA(Fitness, Fitness_threshold, p)
- (B) GA(Fitness, Fitness_threshold, p, r)
- (C) GA(Fitness, Fitness_threshold, p, r, m)
- (D) GA(Fitness, Fitness_threshold)

Answer

Correct option is C

6. Genetic operators includes

- (A) Crossover
- (B) Mutation
- (C) Both A & B
- (D) None of these

Answer

Correct option is C

7. Produces two new offspring from two parent string by copying selected bits from each parent is called

- (A) Mutation
- (B) Inheritance
- (C) Crossover
- (D) None of these

Answer

Correct option is C

8. Each schema the set of bit strings containing the indicated as

- (A) 0s, 1s
- (B) only 0s
- (C) only 1s
- (D) 0s, 1s, *s

Answer

Correct option is D

9. 0^*10 represents the set of bit strings that includes exactly

- (A) 0010, 0110
- (B) 0010, 0010
- (C) 0100, 0110
- (D) 0100, 0010

Answer

Correct option is A

10. Correct (h) is the percent of all training examples correctly classified by hypothesis h . then Fitness function is equal to

- (A) Fitness (h) = $(\text{correct} (h))^2$
- (B) Fitness (h) = $(\text{correct} (h))^3$
- (C) Fitness (h) = $(\text{correct} (h))$
- (D) Fitness (h) = $(\text{correct} (h))^4$

Answer

Correct option is A

11. Statement: Genetic Programming individuals in the evolving population are computer programs rather than bit strings.

- (A) True
- (B) False

Answer

Correct option is A

12. _____ evolution over many generations was directly influenced by the experiences of individual organisms during their lifetime

- (A) Baldwin
- (B) Lamarckian
- (C) Bayes
- (D) None of these

Answer

Correct option is B

13. Search through the hypothesis space cannot be characterized. Why?

- (A) Hypotheses are created by crossover and mutation operators that allow radical changes between successive generations
- (B) Hypotheses are not created by crossover and mutation operators.
- (D) None of these

Answer

Correct option is A

14. ILP stand for

- (A) Inductive Logical programming
- (B) Inductive Logic Programming
- (C) Inductive Logical Program
- (D) Inductive Logic Program

Answer

Correct option is B

15. What is/are the requirement for the Learn-One-Rule method?

- (A) Input, accepts a set of +ve and -ve training examples.
- (B) Output, delivers a single rule that covers many +ve examples and few -ve.
- (C) Output rule has a high accuracy but not necessarily a high coverage.
- (D) A & B
- (E) A, B & C

Answer

Correct option is E

16. _____ is any predicate (or its negation) applied to any set of terms.

- (A) Literal
- (B) Null
- (C) Clause
- (D) None of these

Answer

Correct option is A

17. Ground literal is a literal that

- (A) Contains only variables
- (B) does not contains any functions
- (C) does not contains any variables
- (D) Contains only functions

Answer

Correct option is C

18. _____ emphasizes learning feedback that evaluates the learner's performance without providing standards of correctness in the form of behavioural targets.

- (A) Reinforcement learning
- (B) Supervised Learning
- (C) None of these

Answer

Correct option is A

19. Features of Reinforcement learning

- (A) Set of problem rather than set of techniques
- (B) RL is training by reward and punishments.
- (C) RL is learning from trial and error with the world.
- (D) All of these

Answer

Correct option is D

20. Which type of feedback used by RL?

- (A) Purely Instructive feedback
- (B) Purely Evaluative feedback
- (C) Both A & B
- (D) None of these

Answer

Correct option is B

21. What is/are the problem solving methods for RL?

- (A) Dynamic programming
- (B) Monte Carlo Methods
- (C) Temporal-difference learning
- (D) All of these

Answer

Correct option is D

