SIDDHARTH INSTITUTE OF ENGINEERING & TECHNOLOGY: PUTTUR
(AUTONOMOUS)
Siddharth Nagar, Narayanavanam Road – 517583
BIT BANK (OBJECTIVE)

Subject with Code: Advanced Machine Learning(20CS0918)   **Course & Branch:** B.Tech – CSM
Year & Sem: III-B.Tech & II-Sem   **Regulation :** R20

UNIT-1
INTRODUCTION TO MACHINE LEARNING

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	The term machine learning was coined in which year?			[ ]
	A) 1958	**B) 1959**	C) 1960	D) 1961
3	The father of machine learning is _____ .			[ ]
	A) **Geoffrey Everest Hinton**	B) Geoffrey Hill	C) Geoffrey Chaucer	D) None of the above
4	What is a well-posed problem in the context of intelligent machines?.			[ ]
	A)**well-posed problem is a task or challenge that is clearly defined, has a unique solution, and is feasible to solve using computational methods.**	B) It is a task or challenge that is not clearly defined, has a unique solution, and is feasible to solve using computational methods.	C) Both A and B	D) None of the above
5	What are the three traits that well-posed learning problem has			[ ]
	A) **Task(T), Performance Measure (P), Experience(E)**	**B) Train (T), Test (T), Performance (P)**	C) Task (T), Split(S), Experience(E) **D) none**	
6	Application of Machine learning is _____.			[ ]
	A) Email filtering	B) sentimental analysis	C) face recognition	**D) All of the above**
7	Which among the following algorithms are used in Machine learning?			[ ]
	A) Naive Bayes	B) Support Vector Machine	C) K-Nearest Neighbors	**D) All of the above**
8	What is the application of machine learning methods to a large database called?			[ ]

	A) Big data computing	B) Internet of things	C) **Data mining**	D) Artificial intelligence	
9	In advanced machine learning, data can be represented in various forms such as_____, _____, and__.				
	A) Tabular, Textual and Structured	B) **Structural, Pre defined, semi data**	C) **Stíuctuíed, Semi-stíuctuíed and Unstíuctuíed data**	D) all of the above	[ ]
10	Machine learning approaches can be traditionally categorized into _____ categories.				[ ]
	A) 3	B) 4	C) 7	D) 9	
11	What is a typical application of machine learning in natural language processing?				[ ]
	A) Sentiment analysis	B) Fraud detection	C) object detection	D) Disease diagnosis	
12	Which step of the machine learning pipeline involves splitting the dataset intotraining and testing subsets?				[ ]
	A) Data Preprocessing	B) Model Selection	C) **Model Evaluation**	D) Model Deployment	
13	_____ is the machine learning algorithms that can be used with labeled data.				[ ]
	A) Regression algorithms	B) Clustering algorithm	C) Association algorithm	D) All of the above	
14	What is unstructured data?				[ ]
	A) Data that is well-defined and easy to search	B) Data that is well-defined and easy to search	C) **Data that is qualitative and represented in chats, videos, and audio.**	D) Data that is encrypted and secured.	
15	Which of the following is a fundamental concept in linear algebra used in machinelearning?				[ ]
	A) Eigenvalues.	B) Singular value decomposition	C) **Matrix multiplication**	D) Determinants.	
16	_____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	
17	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	
18	_____ 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	
19	_____ 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	
20	Which Language is Best for Machine Learning?				[ ]
	A) C	B) Java	C) **Python**	D) HTML	
21	The supervised learning problems can be grouped as_____.				[ ]
	A) Regression problems	B) Classification problem	C) **Both A and B**	D) None of the above	
22	The unsupervised learning problems can be grouped as _____.				[ ]
	A) Clustering	B) Association	C) **Both A and B**	D) None of the above	

23	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**	
24	The term machine learning was coined by _____.				[ ]
	A) James Goslin	**B) Arthur Samuel**	C) Guido van Rossum	D) None of the above	
25	Machine Learning can automate many tasks, especially the ones that only humans can perform with their innate intelligence.				[ ]
	A) True	**B) False**	C) Both A) and B)		
26	Features of Machine Learning are_____.				[ ]
	A) Automation	B) Improved customer experience	C) Business intelligence	**D) All of the above**	
27	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	
28	Machine learning is a subset of				[ ]
	A) Deep Learning	**B) Artificial Intelligence**	C) Data Learning	D) None of the above	
29	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	
30	.... 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	
31	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	
32	What is the main purpose of using linear algebra in machine learning? A) To visualize data B) To represent data in a structured form C) To perform calculations efficiently D) To analyze data qualitatively				[ ]
33	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	
34	Which data mining technique is used for finding patterns in data. sets?				[ ]
	A) **Association rule mining**	B)regression analysiss	C) Decision Trees	D) Data Preprocessing	
35	Which of the following is an example of a data mining application? A) Predicting stock market trends				[ ]

	B) Sorting files in a folder C) Formatting text in a document D) Sending emails				
36	What is the output of training process in machine learning? A) Null B) Accuracy C) **Machine learning model** D) Machine learning algorithm				[ ]

37	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

38	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**	

39	Hyperparameter tuning involves A)Selecting the features with the highest importance B) **Optimizing the parameters of the machine learning algorithm** C) Transforming the input features into a higher-dimensional space D) Evaluating the model's performance on unseen data	[ ]

40.	Data used to build a data mining model.				[ ]
	A) training data	B) validation data	C) test data	D) hidden data	

UNIT-II
SUPERVISED LEARNING

1	A_____ is a supervised learning algorithm that is used for classification and regression modelling.				[	]
	A) Univariate Tree	**B) Decision Tree**	C Multivariate Tree	D) All		
2	The data consists of only one variable is known as _____				[	]
	A) Bivariate Data	B) Multivariate Data	C) Both A &B	**D) Univariate Data**		
3	What is the main **pur**pose of Multiv**a**riate Data?				[	]
	A) to study relationship among them	B) to explain	C) to describe	D) none of the above		
4	Which of the following regularization techniques can help prevent overfitting in linear basis function models?				[	]
	A) Lasso regression	B) Ridge Regression	C) Elastic net regression	**D) all the above**		
5	Which t**e**chnique is used to compr**e**ss the data in machi**n**e learning?				[	]
	A) Decision	B) Bivariate	**C) Pruning**	D) all the above		
6	Which of the following is a common type of basis function used in linear basis function models?				[	]
	A) Gaussian basis function	B) Sigmoid basis function	C) Polynomial basis function	**D) All the above**		
7	Which of the following techniques is commonly used to train linear basis function models?				[	]
	A) **Gradient descent**	B) K-means clustering	C) PCA	D) Decision Trees		
8	Which of the following is NOT a supervised learning algorithm?				[	]
	A) Decision tress	B) SVM	C) Linear regression	**D)k-means clustering**		
9	Which evaluation metric can be used to assess the bias and variance of a model in the bias-variance decomposition?				[	]
	A) Accuracy	**B) Mean Squared Error (MSE)**	C)AUC	D) F1 score		
10	which theory is a statistical approach that provides a framework for making decisions under uncertainty.				[	]
	A) Bayesian Decision	B) Bayes Estimator	C)Both A&B	D) None		
11	Baye's theorem is mainly depends upon _____				[	]
	A) Empirical probability	B) Classical probability	**C) Conditional probability**	D) None		
12	p(a/b)=(p(b/a)*p(a))/p(b) . what meant by p(a/b):				[	]
	A) Likelihood	B) prior probability	C) marginal	**D) hypothesis**		
13	p(a/b)=(p(b/a)*p(a))/p(b) . what meant by p(b/a):				[	]
	A) Likelihood	B) prior probability	C) marginal	D) hypothesis		
14	p(a/b)=(p(b/a)*p(a))/p(b) . what meant by p(a):				[	]
	A) Likelihood	**B) prior probability**	C) marginal	D) hypothesis		

15	p(a/b)=(p(b/a)*p(a))/p(b) . what meant by p(b):				[ ]
	A) Likelihood	B) prior probability	**C) marginal**	D) hypothesis	
16	_____ is a type of machine learning where the algorithm learns to map input data to known data.				[ ]
	A) Supervised Learning	B) Unsupervised Learning	C) Semi-Supervised Learning	D)Reinforcement Learning	
17	Full form of **MLE** in the machine learning.				[ ]
	A) Method likelihood Elimination	B) Minimum Likelihood Elimination	**C) maximum Likelihood Estimation**	D) Maximum Likelihood Elimination	
18	What is the role of Maximum Likelihood Estimation in ML.				[ ]
	A) To increase the complexity	**B) To determine the values for parameters of the model**	C) To determine the parameters	D) To decrease the complexity	
19	Application of machine learning methods to large database is called_____				[ ]
	A) Data Mining	B) Artificial Intelligence	C) big data computing	D) Internet of things	
20	Full form of LDA:				[ ]
	A) Learning discriminate Analysis	B) Least Discriminate Analysis	**C) Linear Discriminant Analysis**	D) Labelled Discriminate analysis	
21	In Bayesian linear regression, what term represents the uncertainty or variability in our estimates of the regression coefficients				[ ]
	A) Posteriar Distribution	B) Prior distribution	C) Likelihood function	D) None of the above	
22	Which of the following statements is true regarding bias and variance in machine learning models?				[ ]
	A) High bias models tend to overfit the training data	**B)) High bias models are less sensitive to noise in the training data.**	C) High variance models tend to underfit the training data.	D) High variance models have lower generalization error.	
23	Naïve bayes undergoes in which method.				[ ]
	A) parametric	B) Non-parametric	C) Both A & B	D) None	
24	Which term in the bias-variance decomposition represents the error caused by the model's sensitivity to small fluctuations in the training data?				[ ]
	A) Bias	B) Noise	C) MSE	**D) Variance**	
25	What is the equation of a simple linear regression model?				[ ]
	A) Y=mx+b	B) Y=e$\beta 0$+$\beta 1X$	**C) Y=$\beta 0$+$\beta 1X$**	D) y=1+e−($\beta 0$+$\beta 1X$)1	
26	How many types of gradient descent are there_____				[ ]
	A) 4	**B) 3**	C)2	D)1	
27	logistic regression is used for _____?				[ ]
	A) classification	B) Regression	C) clustering	D) All of these	
28	what the cost function of logistic regression?				[ ]

	A) sigmoid function	B) Logistic function	**C) Both A & B**	D) None	
29	What are general limitations of back propagation rule?				[ ]
	A) local minima	B) slow convergence	C) Scaling	**D) All the above**	
30	Which evaluation metric is commonly used for assessing the performance of linear classifiers?				[ ]
	A) Accuracy	B)F1 score	C) AUC-ROC	D) MSE	
31	In which rule the error is propagated from output node to input node or hidden node in machine learning?				[ ]
	A) Back Propagation	B) MLP	C) LDA	D) None	
32	If machine learning model output involves target variable then that model is called as				[ ]
	A) Descriptive model	**B) Predictive model representation**	C)Reinforcement learning	D) All of the above	
33	Which probabilistic discriminative model is commonly used for binary classification tasks?				[ ]
	A) **Logistic Regression**	B) DT	C)KNN	D) Naïve Bayes	
34	Data used to build a data mining model.				[ ]
	A) Training data	B) Validation data	C) Test data	D) Hidden data	
35	Following are the types of supervised learning				[ ]
	A) Classification	B) Regression	C) Subgroup discovery	**D) All of the above**	
36	The output of training process in machine learning is				[ ]
	A) Machine learning model	B) Machine learning algorithm	C) Null	D) Accuracy	
37	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	
38	You are given reviews of few net-flix series marked as positive, negative and neutral. Classifying reviews of a new net-flix series is an example of				[ ]
	A) supervised learning	B) unsupervised learning	C) semi-supervised learning	D) reinforcement learning	
39	What is the purpose of using regularization techniques such as Ridge Regression or Lasso Regression in Multiple Linear Regression?				[ ]
	A) To reduce the number of independent variables in the model.	B) To increase the complexity of the model and improve prediction accuracy.	C) To enforce sparsity in the regression coefficient	**D) To address multicollinearity and prevent overfitting.**	
40	Which method is commonly used to estimate parameters in probabilistic generative models?				[ ]
	A) Maximum Likelihood Estimation (MLE)	B) Gradient Descent	C) PCA	**D) Expectation-Maximization (EM) algorithm**	

UNIT-III
UNSUPERVISED LEARNING

1	Which learning algorithm is the training of a machine using information neither that is neither classified nor labeled and allowing the algorithm to act on that information without guidance?			[ ]
	A) Supervised	**B)Unsupervised**	C) Both A & B	D) None

2	In which algorithm is the machine is restricted to find the hidden structure in unlabeled data by itself? A) Supervised Learning B) Reinforcement Learning C) Unsupervised learning D) All of the above	[ ]

3	Unsupervised learning is mainly classified into how many categories of algorithm.			[ ]
	A) 2	B) 4	C) 3	D) 5

4	Unsupervised learning algorithm is mainly deal with _____.			[ ]
	A) Labeled	**B) Unlabeled**	C) Both a& b	D) Only A

5	How many types of clusters are there in unsupervised learning algorithm?			[ ]
	A) 2	B) 4	C) 6	**D) 5**

6	K-means cluster is present in which algorithm?	[ ]
	A) **Unsupervised learning** B) Supervised learning C) Reinforcement learning Active learning	

7	why **K-Means** is a Unsupervised Algorithm?	[ ]
	A) it is labelled B) it is unlabelled C) it's a type of cluster D) **Both A & B**	

8	The data which is required for both Supervised and Unsupervised learning algorithm. A) Active learning B) Hidden data C) **Input data** D) all of the above	[ ]

9	What's the Accurate of Unsupervised learning algorithm? A) Less Accurate B) Moderate C) High Accurate D) Very high accurate	[ ]

10	For which algorithm the no. of class is not required?	[ ]
	A) **unsupervised learning** B) supervised learning C) Reinforcement Learning active learning	

11	In data Analysis which analysis is used by Unsupervised learning algorithm? 1. Uses offline Analysis 2. Uses real-time Analysis	[ ]

	A) only 1	B) both 1 & 2	**C) only 2**	D) None	
12	The data which is not required for Unsupervised learning algorithm?				[ ]
	A) Output data	B) Hidden data	C) Input data	D) All of the above	
13	Desired output is not given for _____ **A) Unsupervised Learning** B) Supervised learning C) Reinforcement Learning D) Active Learning				[ ]
14	Unsupervised learning algorithm is also known as_____				[ ]
	A) supervised	B) Classification	C) active	**D) Clustering**	
15	Dimensionality reduction can be easily accomplished by using _____ learning algorithm.				[ ]
	A) unsupervised learning	B) supervised learning	C) Reinforcement Learning	D) Active learning	
16	What's the level of cost for Unsupervised learning algorithm?				[ ]
	A) Very High Cost	B) High Cost	C) Medium Cost	**D) Low Cost**	
17	What are the Disadvantages of Unsupervised Learning Algorithm? i) Difficult to measure accuracy. ii) Effectiveness due to lack of predefined answers during training. iii) The results often have lesser accuracy. Iv) Lack of Guidance.				[ ]
	A) Only I & ii	B) Only i, iii, iv	C) Only ii & iv	**D) All of the above**	
18	In **K-Means** clustering . what is mean by cluster?				[ ]
	A) No. of class	**B) No. of clusters**	C) No. objectives	D) None of the above	
19	Lack of Guidance is mostly present in which algorithm?				[ ]
	A) Reinforcement Learning	B) supervised learning	**C) unsupervised learning**	D) Active learning	
20	Fuzzy is mostly present in which type?				[ ]
	A) Clustering	B) Regression	C) classification	D) Association	
21	What is full form of **KNN.**				[ ]
	A) K not-near Neighbors	B) K never Near	C) K NeighborNot	**D) K Nearest Neighbors**	
22	The of finding hidden structure in unlabelled data is called_____.				[ ]
	A) reinforcement learning **B) unsupervised learning** C) supervised learning D) active learning				
23	**PCA** is				[ ]
	A) Forward feature Selection				

	B) Backward Feature Selection **C) Feature Selection** D) All of the above				
24	What is full form of **PCA.** A) **Principal Component Analysis** B) Principal Content Analysis C) Proper Content Analysis D) Proper Component Analysis				[ ]
25	Identification of cancer cells and Customer segmentation are the application of _____.				[ ]
	A) Classification	B) Classification	C) Association	**D) Clustering**	
26	what is 2nd step in K-Means clustering Algorithm.				[ ]
	A)Select the number k to decide the no. of clusters B) **Select random k points or clusters** C) Assign each data point to their closest centroid D) Calculate the variables				
27	Which algorithm is defined as the combination of various unsupervised ML algorithms and is used to determine the local maximum likelihood estimation?				[ ]
	A) K-means	B) KNN	C) Fuzzy	**D)Expectation-Maximization**	
28	_____ variable mixture models are used to investigate heterogeneous populations consisting of two or more clusters of objects.				[ ]
	A) Local	B) Global	**C) Latent**	D)All of the above	
29	Unsupervised learning is used to solve_____.				[ ]
	A) Clustering	B) Association	C) Dimensionality	**D) All of the above**	
30	_____ clustering is a learning frame work using a special object function.				[ ]
	A) Reinforcement Learning	**B) unsupervised learning**	C) supervised learning	D) Active learning	
31	Which clustering is a method of cluster analysis that seeks to build a hierarchy of cluster.				[ ]
	A) Hierarchy clustering	B) Partitioning	C) Distribution Model- Based	D) Fuzzy	
32	Strategies for hierarchical clustering generally fall into two categories: they are_____ and____.				[ ]
	A) Regression & classification	B) K-means & KNN	**C)Agglomerative & Divisive**	D) K-Means & Divisive	
33	Which of the following is a disadvantage of hierarchical clustering? A) It requires the user to specify the number of clusters. B) It can be computationally expensive for large datasets. C) It always produces the same clustering result for a given dataset. D) None of the above				[ ]
34	What is the difference between agglomerative and divisive hierarchical clustering? A) Agglomerative clustering starts with one cluster and divides it into smaller clusters, while divisive clustering starts with all data points in one cluster and merges them into larger clusters				[ ]

	B) **Agglomerative clustering starts with all data points in one cluster and merges them into larger clusters, while divisive clustering starts with one cluster and divides it into smaller clusters.** C) There is no difference between agglomerative and divisive clustering. D) None of the Above	

35	Which of the following is NOT a common application of mixture of latent variable models?			[ ]
	A) Clustering	B) Anomaly detection	C) Dimensionality reduction	**D) Binary classification**

36	Which of the following is not a component of a Mixture of Latent Variable Model?			[ ]
	A) Latent variables	B) Mixing coefficients	**C) Observations**	D) Model parameters

37	What is the main purpose of the EM algorithm? A) **To maximize the likelihood function of a dataset.** B) To minimize the error of a classification model C) To find the best parameters for a regression model D) To reduce the complexity of a dataset.	[ ]

38	In the E-step of the EM algorithm, what is computed? A) The likelihood function of the data given the parameters. B) **The posterior probability of the latent variables given the data and the parameters.** C) The gradient of the likelihood function with respect to the parameters. D) The Hessian matrix of the likelihood function	[ ]

39	In the M-step of the EM algorithm, what is optimized? A) The posterior probability of the latent variables. B) **The likelihood function of the data given the parameters.** C) The gradient of the likelihood function with respect to the parameters. D) The Hessian matrix of the likelihood function	[ ]

40	In which type of problems is the EM algorithm commonly used?				[ ]
	A) **Clustering.**	B) Regression.	C) Classification.	D) Feature selection.	

UNIT-IV

NONPARAMETRIC METHODS AND DIMENSIONALITY REDUCTION

1	What is non-parametric density estimation? A) Estimating density using fixed parameters **B) Estimating density without assuming a fixed functional form** C) Estimating density using a parametric model D) Estimating density with a large number of parameters	[ ]
2	Which of the following is a non-parametric method for density estimation? A) Gaussian Mixture Models (GMM) B) **Kernel Density Estimation (KDE**) C) Maximum Likelihood Estimation (MLE) D) Naive Bayes	[ ]
3	What is the main idea behind the K-nearest neighbor (KNN) algorithm? A) It is a supervised learning algorithm used for classification and regression tasks. B) It is an unsupervised learning algorithm used for clustering. **C) It predicts the class of an unknown sample based on the classes of its k nearest neighbors.** D) It calculates the mean value of the target variable for the k nearest neighbors.	[ ]
4	What is the significance of the parameter k in KNN? A) It represents the number of features in the dataset. B) It determines the distance metric used for calculating distances between data points. **C) It specifies the number of nearest neighbors to consider for classification.** D) It controls the regularization strength of the algorithm.	[ ]
5	In KNN, how is the distance between two data points typically measured? **a) Euclidean distance** b) Manhattan distance c) Minkowski distance d) Hamming distance	[ ]
6	What is the training time complexity of the KNN algorithm? a) O(1) b) O(log n) **c) O(n)** d) O(n^2)	[ ]
7	Is KNN considered a parametric or non-parametric machine learning algorithm? a) Parametric **b) Non-parametric** c) Both parametric and non-parametric d) None of the above	[ ]
9	Which of the following is a non-parametric classification algorithm? a) Logistic Regression **b) Decision Trees** c) Naive Bayes	[ ]

	d) Support Vector Machines		
10	Which of the following is a non-parametric classification algorithm? a) Logistic Regression **b) Decision Trees** c) Naive Bayes d) Support Vector Machines	[	]
11	Which non-parametric classification algorithm uses a set of rules to make decisions based on feature conditions? **a) Decision Trees** b) Logistic Regression c) K-means clustering d) Support Vector Machines	[	]
12	Can non-parametric classification algorithms handle high-dimensional data well? a) Yes, they are designed to handle high-dimensional data efficiently. b) No, they struggle with high-dimensional data due to the curse of dimensionality. **c) It depends on the specific algorithm and the dataset characteristics.** d) They perform better with low-dimensional data.	[	]
13	What is the condensed nearest neighbor (CNN) algorithm in machine learning? a) A clustering algorithm used for feature selection. **b) A supervised learning algorithm used for classification.** c) An unsupervised learning algorithm used for dimensionality reduction. d) A semi-supervised learning algorithm used for anomaly detection	[	]
14	What is the main goal of the condensed nearest neighbor (CNN) algorithm? a) To find the nearest neighbors of a given data point. **b) To select a subset of representative data points from the original dataset.** c) To identify outliers in the dataset. d) To cluster similar data points together.	[	]
15	Which of the following is a variation of the condensed nearest neighbor (CNN) algorithm? a) K-nearest neighbor (KNN) b) Support vector machines (SVM) c) Naive Bayes **d) Edited nearest neighbor (ENN)**	[	]
16	Which of the following is a common approach for subset selection? a) Principal Component Analysis (PCA) **b) Lasso regularization** c) Naive Bayes classification d) K-means clustering	[	]
17	Which of the following is a greedy search algorithm commonly used for subset selection? **a) Forward selection** b) Backward elimination c) Recursive feature elimination d) Ridge regression	[	]
18	What is temporal difference learning? A. A type of reinforcement learning B. A type of unsupervised learning C. A type of supervised learning	[	]

	D. A type of semi-supervised learning
19	What is the purpose of regularization in subset selection? **a) To penalize the inclusion of additional features in the model.** b) To remove outliers from the dataset. c) To handle missing values in the features. d) To handle class imbalance in the target variable.
20	What is the main goal of subset selection in machine learning? a) To reduce the dimensionality of the dataset. b) To improve the interpretability of the model. **c) To enhance the model's predictive performance.** d) To remove outliers from the dataset.
21	What is Principal Component Analysis (PCA) in machine learning? a) **It is a dimensionality reduction technique.** b) It is a classification algorithm. c) It is an unsupervised learning technique for clustering. d) It is a feature selection method.
22	How does PCA achieve dimensionality reduction? a) By selecting a subset of the original features. **b) By projecting the data onto a lower-dimensional subspace.** c) By removing outliers from the dataset. d) By generating synthetic features based on the existing ones.
23	Which of the following is a key concept in PCA? **a) Covariance matrix** b) Cluster centroids c) Decision boundaries d) Logistic regression coefficients
24	What is the relationship between the number of principal components and the dimensionality of the data? a) They are always equal. b) The number of principal components is always higher than the dimensionality. **c) The number of principal components is always lower than the dimensionality.** d) It depends on the explained variance ratio and the desired level of dimensionality reduction
25	What is the main goal of Factor Analysis? a) To improve the interpretability of the data. b) To maximize the predictive performance of the model. c) To identify the most important features in the dataset. **d) To uncover underlying latent factors that explain the correlations among variables.**
26	Which of the following is a key concept in Factor Analysis? a) Principal components b) Cluster centroids c) Decision boundaries **d) Latent factors**
27	What is the difference between exploratory factor analysis (EFA) and confirmatory factor analysis (CFA)? a) EFA is an unsupervised learning technique, while CFA is a supervised learning technique. **b) EFA aims to explore the underlying structure of the data, while CFA tests a pre-defined**

	hypothesis. c) EFA uses the correlation matrix of the variables, while CFA uses the covariance matrix. d) EFA can handle both continuous and categorical variables, while CFA can only handle continuous variables.	
28	What does the factor loading represent in Factor Analysis? **a) The strength of the relationship between variables and factors.** b) The importance of each variable in the dataset. c) The percentage of variance explained by each latent factor. d) The accuracy of the factor analysis model.	[ ]
29	What is the relationship between the number of latent factors and the dimensionality of the data? a) They are always equal. b) The number of latent factors is always higher than the dimensionality. c) The number of latent factors is always lower than the dimensionality. **d) It depends on the amount of variance explained and the desired level of dimensionality reduction**.	[ ]
30	What is the input required for Multidimensional Scaling (MDS)? a) The original high-dimensional feature vectors. **b) The pairwise distances or dissimilarities between samples.** c) The class labels of the samples. d) The number of dimensions to reduce to.	[ ]
31	Which of the following types of Multidimensional Scaling (MDS) preserves the exact pairwise distances in the lower-dimensional space? a**) Metric MDS** b) Non-metric MDS c) Classical MDS d) Kernelized MDS	[ ]
32	What does Stress represent in Multidimensional Scaling (MDS)? a) The degree of clustering in the data. b) The proportion of variance explained by the reduced dimensions. c) **The discrepancy between the original pairwise distances and the distances in the lower-dimensional space**. d) The quality of the model's predictions.	[ ]
33	What is Linear Discriminant Analysis (LDA) in machine learning? a) It is a dimensionality reduction technique. b) It is a clustering algorithm. **c) It is a classification algorithm.** d) It is a feature selection method.	[ ]
34	Which of the following is true about Linear Discriminant Analysis (LDA)? a) LDA assumes that the class labels are independent of each other. b) LDA can only be applied to binary classification problems. **c) LDA assumes that the class-conditional distributions are normally distributed.** d) LDA cannot handle categorical features.	[ ]
35	How does LDA handle dimensionality reduction? a) By selecting a subset of the original features. **b) By projecting the data onto a lower-dimensional subspace.** c) By removing outliers from the dataset.	[ ]

	d) LDA does not perform dimensionality reduction.	
36	What is the purpose of Fisher's criterion in LDA? **a) To find the optimal projection direction that maximizes class separability.** b) To measure the correlation between features and class labels. c) To handle class imbalance in the dataset. d) To handle missing values in the features.	[ ]
37	What is the relationship between LDA and logistic regression? a) LDA is a variant of logistic regression. b) LDA is a non-linear extension of logistic regression. c) LDA and logistic regression are completely unrelated. **d) LDA can be used as a preprocessing step for logistic regression.**	[ ]
38	What are the main reasons for applying dimensionality reduction? a) To decrease the computational complexity of the model. b) To visualize high-dimensional data. c) To remove noise and redundant information. **d) All of the above.**	[ ]
39	Which of the following is an unsupervised dimensionality reduction technique? a) Linear Regression **b) Principal Component Analysis (PCA)** c) Support Vector Machines (SVM) d) Naive Bayes Classifier	[ ]
40	Which of the following is a linear dimensionality reduction technique? a) t-SNE b) Random Forest **c) LDA (Linear Discriminant Analysis)** d) DBSCAN	[ ]

UNIT-V

REINFORCEMENT LEARNING

1	What is reinforcement learning? a) A type of unsupervised learning b) A type of supervised learning c) A type of semi-supervised learning **d) A type of machine learning that involves learning through trial and error and rewards**	[	]
2	What is a reward in reinforcement learning? a) A punishment for an incorrect action **b) A positive reinforcement for a correct action** c) A negative reinforcement for an incorrect action d) None of the above	[	]
3	What is the purpose of an agent in reinforcement learning? a) To provide feedback to the system b) To provide a reward signal to the system **c) To interact with the environment and learn from it** d) None of the above	[	]
4	Which of the following is a type of reinforcement learning algorithm? a) Decision tree b) K-means **c) Q-learning** d) None of the above	[	]
5	_____ Which of the following is not a component of a reinforcement learning system? a) Environment b) Agent c) Reward **d) Decision tree**	[	]
6	In reinforcement learning, what is exploration? **a) Trying out different actions to learn about the environment** b) Trying out the same action repeatedly c) Avoiding exploration and only taking known actions d) None of the above	[	]
7	In reinforcement learning, what is exploitation? a) Trying out different actions to learn about the environment b) Trying out the same action repeatedly **c) Taking known actions to maximize reward** d) None of the above	[	]
8	_____ Which of the following is an example of a reinforcement learning problem? a) Predicting the price of a stock b) Identifying images of cats and dogs c) Playing a game of chess **d) Teaching a robot to navigate a maze**	[	]
9	What is the goal of a reinforcement learning algorithm? **a) To learn the best action to take in any situation** b) To learn the best model to use for prediction c) To learn the best features to use for classification d) None of the above	[	]
10	What is the difference between on-policy and off-policy reinforcement learning? **a) On-policy algorithms learn from the current policy, while off-policy algorithms learn from a different policy** b) Off-policy algorithms learn from the current policy, while on-policy algorithms learn from a different policy	[	]

	c) On-policy algorithms are more efficient than off-policy algorithms d) None of the above	
11	What is the discount factor in reinforcement learning? a) A factor that reduces the reward over time b) A factor that increases the reward over time **c) A factor that determines the weight of future rewards** d) None of the above	[ ]
12	What is the difference between value-based and policy-based reinforcement learning? **a) Value-based algorithms learn a value function, while policy-based algorithms learn a policy function** b) Policy-based algorithms learn a value function, while value-based algorithms learn a policy function c) Value-based algorithms are more efficient than policy-based algorithms d) None of the above	[ ]
13	What is the difference between model-based and model-free reinforcement learning? **a) Model-based algorithms use a model of the environment, while model-free algorithms do not** b) Model-free algorithms use a model of the environment, while model-based algorithms do not c) Model-based algorithms are more efficient than model-free algorithms d) None of the above	[ ]
14	What is the single case k-armed bandit problem? **a) A problem in reinforcement learning where an agent must learn to choose actions that maximize a reward signal** b) A problem in supervised learning where the goal is to predict a continuous output variable c) A problem in unsupervised learning where the goal is to discover patterns in the data d) A problem in deep learning where the goal is to learn a hierarchical representation of the input data	[ ]
15	What is the goal of the single case k-armed bandit problem? **a) To maximize the cumulative reward obtained by the agent over a sequence of actions** b) To minimize the error between the predicted and true values of a continuous output variable c) To discover meaningful patterns in the data without any prior knowledge d) To learn a hierarchical representation of the input data using multiple layers of nonlinear transformations	[ ]
16	What is the k-armed bandit problem? a) A variant of the single case k-armed bandit problem where there are k possible actions an agent can take **b) A variant of the multi-armed bandit problem where there are k possible actions an agent can take** c) A variant of the single case k-armed bandit problem where there are multiple episodes of the problem d) A variant of the multi-armed bandit problem where there are multiple episodes of the problem	[ ]
17		
	In the single case k-armed bandit problem, what is the reward function? **a) A function that maps each action taken by the agent to a reward value** b) A function that maps each state of the environment to a reward value c) A function that maps each observation made by the agent to a reward value	[ ]

	d) A function that maps each episode of the problem to a reward value		
18	In the single case k-armed bandit problem, what is the action-value function? **a) A function that maps each action to the expected reward for that action** b) A function that maps each state to the expected reward for all actions in that state c) A function that maps each observation to the expected reward for all actions in that observation d) A function that maps each episode to the expected reward for all actions taken in that episode	[	]
19	What is the greedy action selection strategy? **a) A strategy where the agent always chooses the action with the highest estimated action-value** b) A strategy where the agent randomly chooses an action from the set of all possible actions c) A strategy where the agent chooses the action with the highest expected reward based on the current state d) A strategy where the agent chooses the action with the highest variance in reward values	[	]
20	What is the primary goal of RL? a. To classify data b. To predict outcomes **c. To find optimal actions in an environment** d. To minimize errors in a model	[	]
21	What is the difference between an agent and an environment in RL? a. An agent is a physical robot, while an environment is a virtual world. **b. An agent interacts with the environment, while the environment is passive.** c. An agent controls the environment, while the environment controls the agent. d. There is no difference between an agent and an environment in RL.	[	]
22	What is a policy in RL? a. A set of rules that govern how an agent behaves in an environment. b. The sequence of states an agent encounters during a learning episode. c. The expected value of rewards an agent receives for taking a particular action in a particular state. **d. The probability distribution over actions given a particular state.**	[	]
23	What is the difference between on-policy and off-policy RL? a. On-policy RL updates the policy based on actions taken by another agent, while offpolicy RL updates the policy based on actions taken by the current agent. **b. On-policy RL updates the policy based on the current policy, while off-policy RL updates the policy based on a different policy**. c. On-policy RL updates the policy based on the value function, while off-policy RL updates the policy based on the state-action value function. d. There is no difference between on-policy and off-policy RL.	[	]
24	What is the Bellman equation in RL? a. An equation that calculates the optimal policy for a given environment. b. An equation that calculates the expected value of rewards for a given state and action. **c. An equation that calculates the expected value of rewards for a given state and the best action to take.** d. An equation that calculates the expected value of rewards for a given state and all possible actions.	[	]
25	What is the Q-value in RL? **a. The expected value of rewards for a given state and action.** b. The probability distribution over actions given a particular state.	[	]

	c. The value of the policy in a particular state. d. The value of the state in a particular action.	
26	What is the difference between exploration and exploitation in RL? a. Exploration is the process of selecting the best action in a given state, while exploitation is the process of trying out different actions. **b. Exploration is the process of trying out different actions, while exploitation is the process of selecting the best action in a given state.** c. Exploration and exploitation are the same thing. d. Exploration is the process of changing the policy, while exploitation is the process of evaluating the policy	[ ]
27	_____ What is model-based learning? a. A type of learning where models are used to generate predictions b. A type of learning where data is used to generate models **c. A type of learning where models and data are used to generate predictions** d. A type of learning where models are used to generate data	[ ]
28	What is a model? a. A set of rules used to make predictions b. A set of data used to make predictions **c. A mathematical representation of a system or process** d. A set of features used to make predictions	[ ]
29	What is the purpose of model-based learning? a. To generate models **b. To generate predictions** c. To generate data d. To generate features	[ ]
30	Which of the following is an example of model-based learning? **a. Decision trees** b. k-means clustering c. Random forests d. Support vector machines	[ ]
31	Which of the following is not a type of model-based learning? A. Supervised learning B. Unsupervised learning C. Reinforcement learning **D. Deep learning**	[ ]
32	Which of the following is an example of supervised learning? A. k-means clustering B. Decision trees **C. Support vector machines** D. Apriori algorithm	[ ]
33	Which of the following is an example of unsupervised learning? A. Linear regression B. Naive Bayes C. K-nearest neighbors **D. Hierarchical clustering**	[ ]
34	What is the difference between supervised and unsupervised learning?	[ ]

	A. In supervised learning, the output variable is known. In unsupervised learning, the output variable is unknown. B. In supervised learning, the input variable is known. In unsupervised learning, the input variable is unknown. C. In supervised learning, the input and output variables are known. In unsupervised learning, only the input variable is known. D. In supervised learning, the input and output variables are unknown. In unsupervised learning, only the input variable is known.	
35	Which of the following is an example of reinforcement learning? A. Linear regression B. Naive Bayes **C. Q-learning** D. Hierarchical clustering	[ ]
36	What is the goal of reinforcement learning? A. To learn a mapping between input and output variables B. To cluster similar data points together **C. To optimize a reward function** D. To generate new data	[ ]
37	Which of the following is an example of a model-based algorithm? **A. Logistic regression** B. Random forests C. K-nearest neighbors D. Gradient boosting	[ ]
38	What is logistic regression used for? **A. Classification** B. Regression C. Clustering D. Dimensionality reduction	[ ]
39	Which of the following is not a type of regression algorithm? A. Linear regression B. Logistic regression **C. K-means clustering** D. Polynomial regression	[ ]
40	Which of the following is not a type of classification algorithm? A. Decision trees **B. k-means clustering** C. Naive Bayes D. Support vector machines	[ ]
41	Which of the following is a limitation of model-based learning? **A. It requires a lot of data** B. It is computationally expensive C. It can only handle small datasets D. It cannot handle missing data	[ ]
42	What is temporal difference learning? **A. A type of reinforcement learning** B. A type of unsupervised learning	[ ]

C. A type of supervised learning
D. A type of semi-supervised learning

43 What is the main objective of temporal difference learning?
A. To maximize rewards　　　　　　　　　　　　　　　　　　　　[　　]
B. To minimize loss
C. To maximize accuracy
D. To minimize error

44 What is the update rule for the value function in temporal difference learning?
A. $V(s) = V(s) + \alpha*(R + \gamma\ V(s') - V(s))$　　　　　　　　　[　　]
B. $V(s) = V(s) + \alpha(R - V(s))$
C. $V(s) = V(s) + \alpha*(\gamma\ V(s') - V(s))$
D. $V(s) = V(s) + \alpha(R + V(s'))$

45 What is generalization in machine learning?
a. The ability of a model to perfectly fit the training data.　　　　　　　[　　]
b. The ability of a model to perform well on new, unseen data
c. The ability of a model to memorize the training data.
d. The ability of a model to perform well on the training data.

46 Which of the following is an example of a high-variance model?
a. A linear regression model.　　　　　　　　　　　　　　　　　　[　　]
b. A decision tree model with a large number of nodes.
c. A support vector machine model.
d. A logistic regression model.

47 Which of the following is a common type of cross-validation?
a. Hold-out validation.　　　　　　　　　　　　　　　　　　　　[　　]
b. Randomized validation.
c. Stratified validation.
d. All

48 Which of the following is true about partially observable environments?
a) The agent has complete information about the environment　　　　　　[　　]
b) The agent has incomplete information about the environment
c) The agent has no information about the environment
d) None of the above

49 Which of the following is an example of a partially observable environment?
a) Chess
b) Tic-Tac-Toe　　　　　　　　　　　　　　　　　　　　　　　　[　　]
c) Rock-Paper-Scissors
d) Poker

50 What is a belief state in POMDPs?
a) A state of complete knowledge about the environment
b) A state of incomplete knowledge about the environment　　　　　　　[　　]
c) A state where the agent has no information about the environment
d) None of the above