Up-to-date Questions and Answers from authentic resources to improve knowledge and pass the exam at very first attempt. ----- Guaranteed.

KILL EXAMS

PASS

killexams.com

CertNexus

AIP-210

Certified Artificial Intelligence Practitioner (CAIP)

Question: 20

In a self-driving car company, ML engineers want to develop a model for dynamic pathing.

Which of following approaches would be optimal for this task?

A. Dijkstra Algorithm
B. Reinforcement learning
C. Supervised Learning.
D. Unsupervised Learning

Answer: B

Explanation:

Reinforcement learning is a type of machine learning that involves learning from trial and error based on rewards and penalties. Reinforcement learning can be used to develop models for dynamic pathing, which is the problem of finding an optimal path from one point to another in an uncertain and changing environment. Reinforcement learning can enable the model to adapt to new situations and learn from its own actions and feedback. For example, a self-driving car company can use reinforcement learning to train its model to navigate complex traffic scenarios and avoid collisions.

Question: 21

R-squared is a statistical measure that:

A. Combines precision and recall of a classifier into a single metric by taking their harmonic mean.
B. Expresses the extent to which two variables are linearly related.
C. Is the proportion of the variance for a dependent variable thafâ s explained by independent variables.
D. Represents the extent to which two random variables vary together.

Answer: C

Explanation:

R-squared is a statistical measure that indicates how well a regression model fits the data. R-squared is calculated by dividing the explained variance by the total variance. The explained variance is the amount of variation in the dependent variable that can be attributed to the independent variables. The total variance is the amount of variation in the dependent variable that can be observed in the data. R-squared ranges from 0 to 1, where 0 means no fit and 1 means perfect fit.

Question: 22

Which of the following equations best represent an LI norm?

A. |x| + |y|
B. |x|+|y|^2
C. |x|-|y|
D. |x|^2+|y|^2

Answer: A

Explanation:

An L1 norm is a measure of distance or magnitude that is defined as the sum of the absolute values of the components of a vector. For example, if x and y are two components of a vector, then the L1 norm of that vector is |x| + |y|. The L1 norm is also known as the Manhattan distance or the taxicab distance, as it represents the shortest path between two points in a grid-like city.

Question: 23

Which of the following statements are true regarding highly interpretable models? (Select two.)

A. They are usually binary classifiers.
B. They are usually easier to explain to business stakeholders.
C. They are usually referred to as "black box" models.
D. They are usually very good at solving non-linear problems.
E. They usually compromise on model accuracy for the sake of interpretability.

Answer: A,B,E

Explanation:

Highly interpretable models are models that can provide clear and intuitive explanations for their predictions, such as decision trees, linear regression, or logistic regression.

Some of the statements that are true regarding highly interpretable models are:

They are usually easier to explain to business stakeholders: Highly interpretable models can help communicate the logic and reasoning behind their predictions, which can increase trust and confidence among business stakeholders. For example, a decision tree can show how each feature contributes to a decision outcome, or a linear regression can show how each coefficient affects the dependent variable.

They usually compromise on model accuracy for the sake of interpretability: Highly interpretable models may not be able to capture complex or non-linear patterns in the data, which can reduce their accuracy and generalization. For example, a decision tree may overfit or underfit the data if it is too deep or too shallow, or a linear regression may not be able to model curved relationships between variables.

Question: 24

Which two of the following decrease technical debt in ML systems? (Select two.)

A. Boundary erosion
B. Design anti-patterns
C. Documentation readability
D. Model complexity
E. Refactoring

Answer: A,C,E

Explanation:

Technical debt is a metaphor that describes the implied cost of additional work or rework caused by choosing an easy or quick solution over a better but more complex solution. Technical debt can accumulate in ML systems due to various factors, such as changing requirements, outdated code, poor documentation, or lack of testing.

Some of the ways to decrease technical debt in ML systems are:

Documentation readability: Documentation readability refers to how easy it is to understand and use the documentation of an ML system. Documentation readability can help reduce technical debt by providing clear and consistent information about the systemâs design, functionality, performance, and maintenance. Documentation readability can also facilitate communication and collaboration among different stakeholders, such as developers, testers, users, and managers.

Refactoring: Refactoring is the process of improving the structure and quality of code without changing its functionality. Refactoring can help reduce technical debt by eliminating code smells, such as duplication, complexity, or inconsistency. Refactoring can also enhance the readability, maintainability, and extensibility of code.

Question: 25

Which of the following describes a neural network without an activation function?

A. A form of a linear regression
B. A form of a quantile regression
C. An unsupervised learning technique
D. A radial basis function kernel

Answer: A

Explanation:

A neural network without an activation function is equivalent to a form of a linear regression. A neural network is a computational model that consists of layers of interconnected nodes (neurons) that process inputs and produce outputs. An activation function is a function that determines the output of a neuron based on its input. An activation function can introduce non-linearity into a neural network, which allows it to model complex and non-linear relationships between inputs and outputs. Without an activation function, a neural network becomes a linear combination of inputs and weights, which is essentially a linear regression model.

Question: 26

The following confusion matrix is produced when a classifier is used to predict labels on a test dataset.

How precise is the classifier?

		Predicted	
		0	1
True	0	48 *true negatives*	8 *false positives*
	1	7 *false negatives*	37 *true positives*

A. 48/(48+37)
B. 37/(37+8)
C. 37/(37+7)
D. (48+37)/100

Answer: B

Explanation:

Precision is a measure of how well a classifier can avoid false positives (incorrectly predicted positive cases). Precision is calculated by dividing the number of true positives (correctly predicted positive cases) by the number of predicted positive cases (true positives and false positives). In this confusion matrix, the true positives are 37 and the false positives are 8, so the precision is 37/(37+8) = 0.822.

Question: 27

Given a feature set with rows that contain missing continuous values, and assuming the data is normally distributed, what is the best way to fill in these missing features?

A. Delete entire rows that contain any missing features.
B. Fill in missing features with random values for that feature in the training set.
C. Fill in missing features with the average of observed values for that feature in the entire dataset.
D. Delete entire columns that contain any missing features.

Answer: C

Explanation:

Missing values are a common problem in data analysis and machine learning, as they can affect the quality and reliability of the data and the model. There are various methods to deal with missing values, such as deleting, imputing, or ignoring them. One of the most common methods is imputing, which means replacing the missing values with some estimated values based on some criteria. For continuous variables, one of the simplest and most widely used imputation methods is to fill in the missing values with the mean (average) of the observed values for that variable in the entire dataset. This method can preserve the overall distribution and variance of the data, as well as avoid introducing bias or noise.

Question: 28

In addition to understanding model performance, what does continuous monitoring of bias and variance help ML engineers to do?

A. Detect hidden attacks
B. Prevent hidden attacks
C. Recover from hidden attacks
D. Respond to hidden attacks

Answer: B

Explanation:

Hidden attacks are malicious activities that aim to compromise or manipulate an ML system without being detected or noticed. Hidden attacks can target different stages of an ML workflow, such as data collection, model training, model deployment, or model monitoring. Some examples of hidden attacks are data poisoning, backdoor attacks, model stealing, or adversarial examples. Continuous monitoring of bias and variance can help ML engineers to prevent hidden attacks, as it can help them detect any anomalies or deviations in the data or the modelâs performance that may indicate a potential attack.

Question: 29

A company is developing a merchandise sales application The product team uses training data to teach the AI model predicting sales, and discovers emergent bias.

What caused the biased results?

A. The AI model was trained in winter and applied in summer.
B. The application was migrated from on-premise to a public cloud.
C. The team set flawed expectations when training the model.
D. The training data used was inaccurate.

Answer: A

Explanation:

Emergent bias is a type of bias that arises when an AI model encounters new or different data or scenarios that were not present or accounted for during its training or development. Emergent bias can cause the model to make inaccurate or unfair predictions or decisions, as it may not be able to generalize well to new situations or adapt to changing conditions. One possible cause of emergent bias is seasonality, which means that some variables or patterns in the data may vary depending on the time of year. For example, if an AI model for merchandise sales prediction was trained in winter and applied in summer, it may produce biased results due to differences in customer behavior, demand, or preferences.

Question: 30

You train a neural network model with two layers, each layer having four nodes, and realize that the model is underfit.

Which of the actions below will NOT work to fix this underfitting?

A. Add features to training data
B. Get more training data
C. Increase the complexity of the model
D. Train the model for more epochs

Answer: B

Explanation:

Underfitting is a problem that occurs when a model learns too little from the training data and fails to capture the underlying complexity or structure of the data. Underfitting can result from using insufficient or irrelevant features, a low complexity of the model, or a lack of training data. Underfitting can reduce the accuracy and generalization of the model, as it may produce oversimplified or inaccurate predictions.

Some of the ways to fix underfitting are:

Add features to training data: Adding more features or variables to the training data can help increase the information and diversity of the data, which can help the model learn more complex patterns and relationships.

Increase the complexity of the model: Increasing the complexity of the model can help increase its expressive power and flexibility, which can help it fit better to the data. For example, adding more layers or nodes to a neural network can increase its complexity.

Train the model for more epochs: Training the model for more epochs can help increase its learning ability and convergence, which can help it optimize its parameters and reduce its error.

Getting more training data will not work to fix underfitting, as it will not change the complexity or structure of the data or the model. Getting more training data may help with overfitting, which is when a model learns too much from the training data and fails to generalize well to new or unseen data.