Pergunta **1**

Resposta guardada

Nota: 1,00

In the context of spam detection using a Bayesian spam filter, what is the main principle behind its operation?

- a. It calculates the probability of an email being spam based on word frequencies.
- b. It relies on deep learning algorithms to detect spam patterns in emails.
- c. It uses image recognition techniques to identify spam emails.
- d. It analyzes the sender's reputation to classify emails as spam or not.

Limpar a minha escolha

Pergunta **2**

Resposta guardada

Nota: 1,00

In the context of anomaly detection, when comparing Isolation Forest and One-Class Support Vector Machine (One-Class SVM), which statement is true?

- a. Isolation Forest can efficiently handle high-dimensional data, while One-Class SVM may struggle with high-dimensional feature spaces.
- b. One-Class SVM is more interpretable and transparent in its anomaly detection results compared to Isolation Forest.
- c. Isolation Forest is a parametric method, while One-Class SVM is non-parametric.
- d. Isolation Forest is primarily suited for linear data, whereas One-Class SVM is more effective for non-linear data.

Limpar a minha escolha

Pergunta **3**

Resposta guardada

Nota: 1,00

What are the main phases of the ML process?

a. 1. Information extraction from Data
2. Knowledge extraction from Information
3. Knowledge Model training
4. Model Evaluation

b. 1. Data gathering and cleaning
2. Statistical models and evaluation

c. 1. Data gathering
2. Data Pre-processing
3. Model training
4. Model evaluation
5. Model deploy

d. 1. Model training and evaluation
2. Data Optimization and Visualization
3. Model deployment and maintenance

Limpar a minha escolha

Pergunta **4**

Resposta guardada

Nota: 1,00

What does ML mean?

a. A computer program that uses math-based models to learn.

b. Machine Learning.

c. A math field that explores optimization to learn.

d. The capability of a model to improve its own performance by using previous data and through trial and error

Limpar a minha escolha

Tempo restante 1:24:39

Pergunta **5**

Resposta guardada

Nota: 1,00

What are the main classes of models explored until now?

- a. Reinforcement Learning, Supervised and Unsupervised Learning and Blind Optimization
- b. Reinforcement Learning and Blind Optimization
- c. Semi-Supervised, Reinforcement Learning, Supervised and Unsupervised Learning and Blind Optimization
- d. Supervised and Unsupervised Learning

Limpar a minha escolha

Tempo restante 1:24:21

Pergunta 6

Por responder

Sem avaliação

Describe the limitations of the BOW (Bag of Words) model, and discuss how the Language model (e.g. FastText) can overcome them.

Pergunta 7

Resposta guardada

Nota: 1,00

Which type of clustering algorithm, among density-based, hierarchical, and partitioning, is most suitable for identifying clusters of varying shapes and sizes in a dataset with noisy data points?

- a. Density-based clustering
- b. Hierarchical clustering
- c. All three types of clustering are equally suitable for this purpose.
- d. Partitioning clustering

Limpar a minha escolha

Pergunta 8

Resposta guardada

Nota: 1,00

When comparing spam detection using embedded models (e.g., Word Embeddings) and the traditional Bag of Words Model, what is a key difference between the two approaches?

- a. The Bag of Words Model uses deep learning techniques, while embedded models use traditional statistical methods.
- b. Both approaches rely on sender reputation for classification.
- c. Embedded models consider the order of words, while the Bag of Words Model does not.
- d. Embedded models primarily focus on email header analysis.

Limpar a minha escolha

Pergunta 9

Resposta guardada

Nota: 1,00

In the context of anomaly detection, what distinguishes a "collective anomaly" from a "point anomaly"?

- a. Collective anomalies are contextual in nature, while point anomalies are independent of any context.
- b. Point anomalies refer to anomalies that are related to contextual factors, while collective anomalies are isolated deviations.
- c. Collective anomalies are single data points significantly deviating from the norm, while point anomalies involve multiple data points exhibiting a collective pattern.
- d. Point anomalies are rare events that occur in isolation, while collective anomalies are common patterns in the data.

Limpar a minha escolha

Pergunta 10

Resposta guardada

Nota: 1,00

In anomaly detection using autoencoders, what is the primary role of the encoder component?

- a. To generate synthetic anomalies.
- b. To perform feature extraction.
- c. To reconstruct the input data.
- d. To increase the dimensionality of the data.

Limpar a minha escolha

Tempo restante 1:24:03

Pergunta 11

Por responder

Nota: 1,00

When performing spam detection as a binary classification task, what are the two primary classes that emails are categorized into?

- a. Spam and ham
- b. Safe and dangerous
- c. Valuable and worthless
- d. Inbound and outbound

Pergunta 12

Por responder

Sem avaliação

Describe how an AutoEncoder can be used as an anomaly detection model (with dynamic thresholding).

Pergunta 13

Por responder

Nota: 1,00

Why is ML considered for Security and Privacy fields?

- a. Given the latest version of the Internet (IoT) the number of entities will increase at such a rate that human-maintained networks are not possible anymore. That is the need for AI/ML to optimize and protect our networks as autonomously as possible.
- b. It is a new trend on computers and should be explored.
- c. ML allows for faster solutions that will decrease latency while improves security.
- d. ML allows for smaller solutions with huge savings in energy consumption.

Pergunta 14

Por responder

Nota: 1,00

In the context of anomaly detection through clustering, which type of anomaly detection primarily relies on identifying data points that are farthest from the cluster centers?

a. Contextual anomaly detection

b. Point anomaly detection

c. Sequential anomaly detection

d. Collective anomaly detection

Pergunta 15

Por responder

Nota: 1,00

In the context of machine learning and model performance, what does bias and variance represent?

a. Bias and variance are statistical measures of the model's interpretability and transparency.

b. Bias is a measure of the model's ability to generalize to new data, while variance is a measure of its accuracy on the training data.

c. Bias measures the model's ability to fit the training data accurately, while variance measures its ability to generalize to new, unseen data.

d. Bias represents the model's overall complexity, while variance represents the dataset's diversity.

Tempo restante 1:22:55

Pergunta 16

Por responder

Nota: 1,00

When comparing clustering algorithms and Autoencoders for anomaly detection, which of the following statements is accurate?

a. Clustering algorithms and Autoencoders serve distinct purposes, and neither is superior for anomaly detection universally.

b. Clustering algorithms are mainly used for dimensionality reduction, while Autoencoders are better suited for identifying anomalies in unlabeled data.

c. Clustering methods are primarily unsupervised, while Autoencoders rely on labeled data for effective anomaly detection.

d. Autoencoders are inherently less scalable compared to clustering methods in handling large datasets.

Pergunta 17

Por responder

Nota: 1,00

When discussing email SPAM issues, what is the primary purpose of a spam filter?

a. To detect and filter out unsolicited and potentially harmful or irrelevant emails.

b. To automatically reply to every email received, verifying the sender's legitimacy.

c. To increase the visibility of promotional emails in the recipient's inbox.

d. To send all incoming emails to the recipient's inbox for review.

Pergunta 18

Por responder

Nota: 1,00

Which of the following best describes the "curse of dimensionality" in the context of data analysis and machine learning?

a. The curse of dimensionality is a problem related to overfitting in machine learning algorithms.

b. The curse of dimensionality refers to the high computational cost associated with increasing dataset sizes.

c. The curse of dimensionality is a statistical term that signifies the presence of outliers in high-dimensional datasets.

d. The curse of dimensionality is a term used to describe the difficulties of visualizing and analyzing data in high-dimensional spaces.

Pergunta 19

Por responder

Nota: 1,00

When comparing Principal Component Analysis (PCA) and Autoencoders for anomaly detection, which statement is true?

a. Autoencoders are computationally less efficient and not suitable for real-time anomaly detection tasks.

b. PCA is a linear technique, while Autoencoders can capture both linear and nonlinear patterns, making them more versatile for anomaly detection.

c. PCA is typically better at capturing complex patterns and nonlinear relationships in data.

d. Autoencoders are primarily used for dimensionality reduction and not suitable for anomaly detection.

Pergunta 20

Por responder

Nota: 1,00

In the ongoing fight against spam, which of the following is a common technique used to reduce the impact of spam emails and protect recipients?

a. Requiring all email senders to use CAPTCHA verification.

b. Adding a spam header to all incoming emails.

c. Implementing sender-only whitelists for all email servers.

d. Using AI and machine learning for content analysis.

Tempo restante 1:21:27

Pergunta 21

Por responder

Nota: 1,00

In the context of identifying cybersecurity attacks using anomaly detection, which type of attack is most likely to be detected?

- a. Attacks that exploit known vulnerabilities
- b. Attacks that mimic normal network traffic
- c. Well-known and widely recognized attacks
- d. Attacks that use advanced encryption techniques

Pergunta 22

Por responder

Nota: 1,00

In spam detection using the Bag of Words Model, which of the following best describes the process?

- a. It analyzes the entire email content as a continuous sequence of words.
- b. It classifies emails as spam or not based on sender reputation alone.
- c. It uses deep learning to analyze the email's structural elements.
- d. It counts the frequency of specific words and constructs a feature vector.