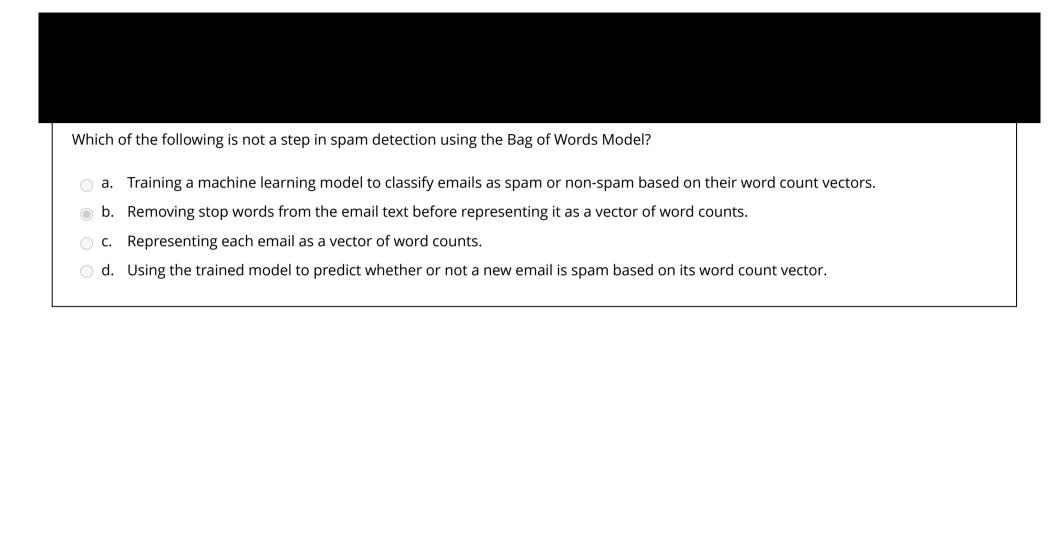
| Iniciada | sexta-feira, 3 de novembro de 2023 às 16:51 |
|-------------|--|
| Estado | Terminada |
| Terminada | sexta-feira, 3 de novembro de 2023 às 17:29 |
| Tempo gasto | 38 minutos 5 segundos |
| Nota | 16,08 num máximo de 20,00 (80,42 %) |

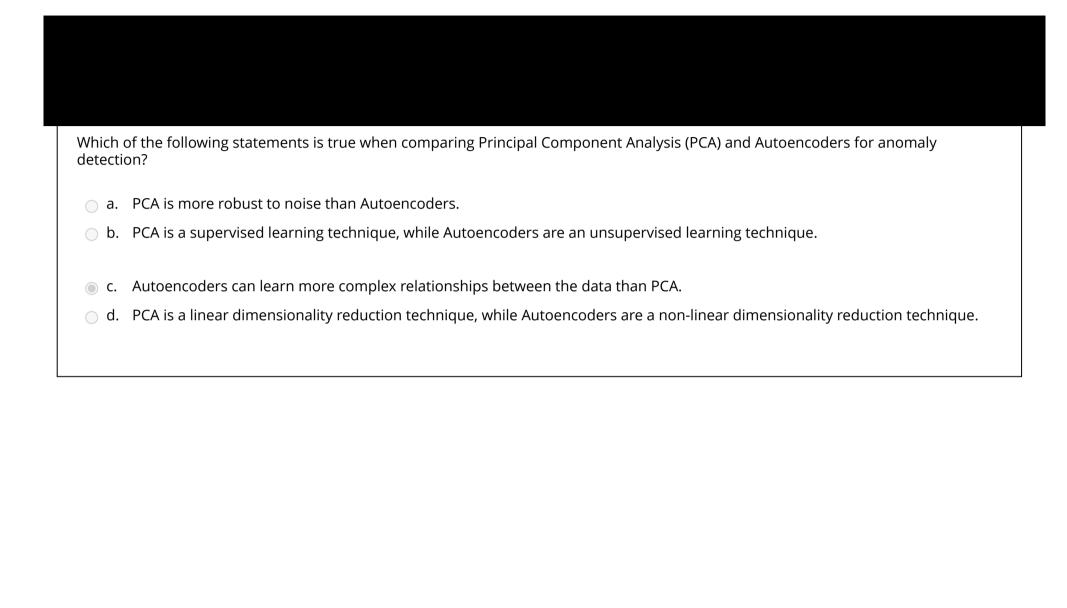


| Which | of the following is the key difference between a collective anomaly and a point anomaly in anomaly detection? |
|-------|--|
| _ a. | A collective anomaly is an anomaly that occurs in a single data stream, while a point anomaly is an anomaly that occurs across multiple data streams. |
|) b. | A collective anomaly is an anomaly that is caused by a single underlying event, while a point anomaly is an anomaly that is caused by multiple underlying events. |
| C. | A collective anomaly is a group of data points that deviate significantly from the expected behavior, while a point anomaly is a single data point that deviates significantly from the expected behavior. |
| _ d. | All of the above. |
| | |
| | |
| | |

| | Which of the following best describes the "curse of dimensionality" in the context of data analysis and machine learning? | |
|------------|--|--|
| | a. The difficulty of finding patterns in high-dimensional data. | |
| | ○ b. The requirement of more data to train machine learning models in high-dimensional spaces. | |
| | ⊚ c. All of the options | |
| | od. The tendency of the volume of data to increase exponentially as the number of dimensions increases. | |
| | | |
| | | |
| | Pergunta 4 | |
| Respondida | | |
| | Nota: 1,00 em 1,00 | |
| | | |
| | Which of the following types of clustering algorithms is most suitable for identifying clusters of varying shapes and sizes in a dataset with noisy data points? | |
| | ○ a. Spectral clustering | |
| | │ ○ b. Partitioning clustering | |

o. Hierarchical clustering

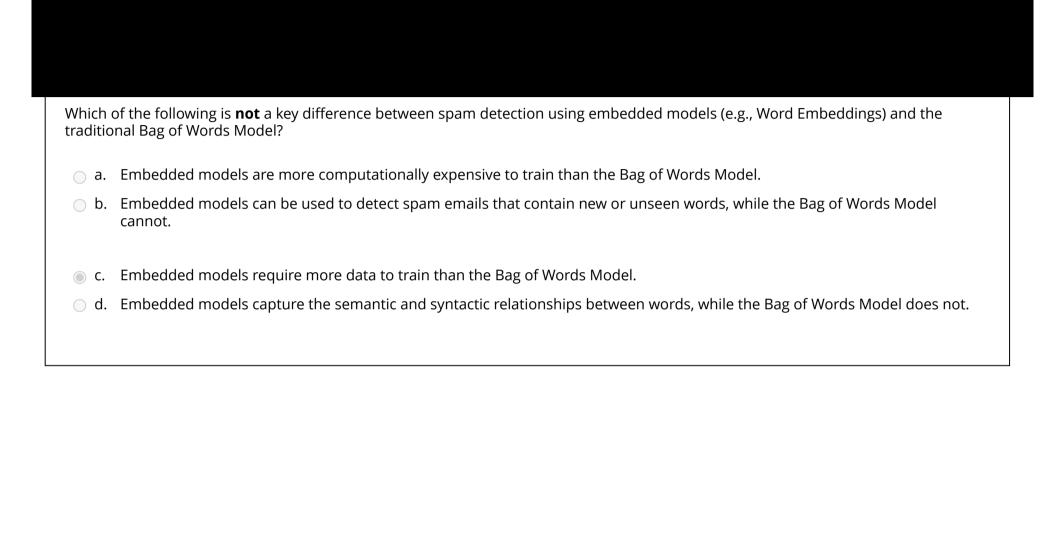
d. Density-based clustering



How does FastText represent words differently from the BOW model?

Para o BOW, um texto como um email ou um SMS não possui estrutura semântica ou sintactica, é apenas um agrupamento sem ordem de palavras, e sobre as quais apenas vai analisar a frequência com que ocorrem. Por causa disto, apesar de textos que possam ser de teor dramaticamente diferente, se partilharem palavras iguais em grande proporção, serão considerados como similares.

No FastText, as palavras vão ser entendidas como vectores, o que resulta em que a posição e organização das palavras relativamente umas as outras, contexto, resulte em mudanças significativas do vetor e padrões sejam mais facilmente reconheciveis.



Which of the following is **not** a common technique used to reduce the false positive rate of spam filters?
a. Training the spam filter on a large dataset of labeled emails.
b. Using machine learning algorithms to identify patterns in spam emails.
c. Using a whitelist to allow users to identify email addresses that should never be flagged as spam.
d. Using a blacklist to block email addresses from known spammers.

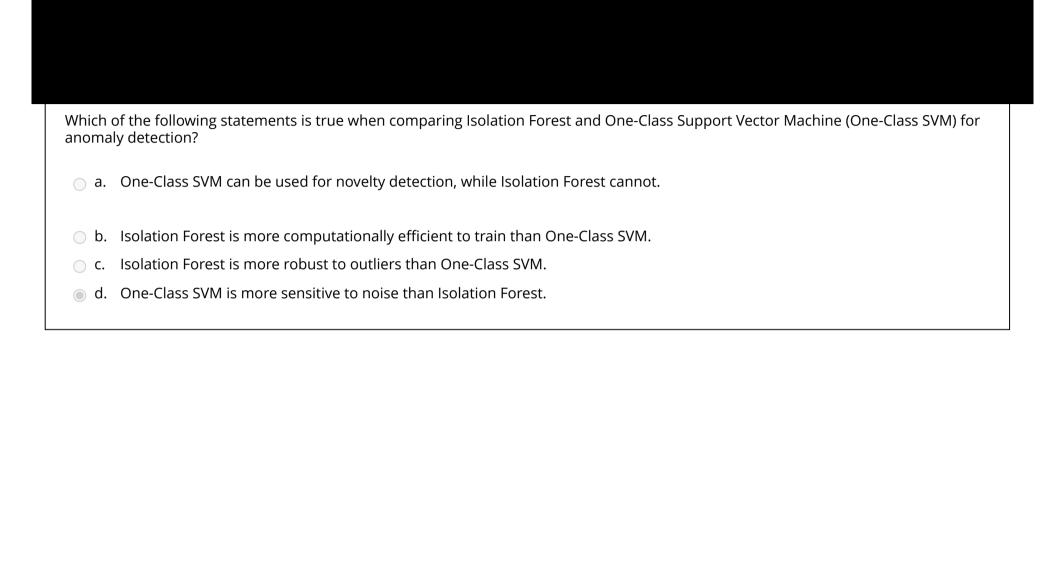
Which of the following is a key challenge in developing effective spam filters?
a. To identify and prevent phishing emails from compromising users' accounts.
b. All of the answers.
c. To detect and remove malware from attachments to incoming emails.
d. To block unsolicited emails from reaching users' inboxes.

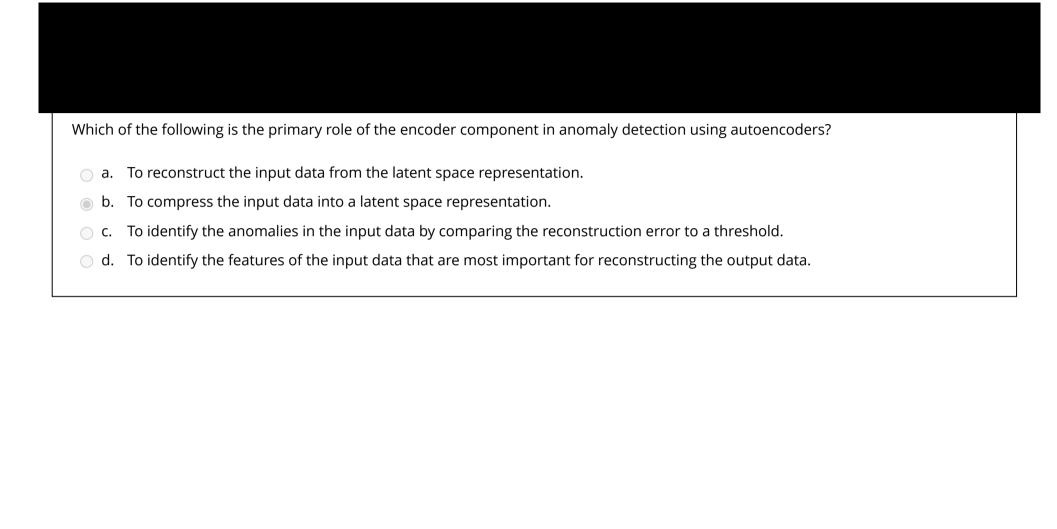
Respondida

Nota: 0,33 em 1,00

Which of the following types of anomaly detection primarily relies on identifying data points that are farthest from the cluster centers? Density-based clustering Hierarchical clustering Distance-based clustering d. Spectral clustering Pergunta 11 Respondida Nota: 1,00 em 1,00

Which of the following types of cybersecurity attacks is most likely to be detected using anomaly detection?
a. Known attacks with a well-defined signature.
b. Attacks that are carefully targeted at a specific organization or individual.
c. Novel attacks that have never been seen before.
d. Attacks that are disguised as legitimate activity.





How can clustering algorithms be used to identify anomalies? What are the main differences between the Isolation Forest approach?

Algoritmos de clustering são usados para agrupar dados e detetar quais pontos não encaixam ou se distanciam da maioria, estes podem representar anomalias. Mas para isso funcionar, temos de assumir que anomalias são relativamente raras em relação aos dados normais. Isto pode tornar o treino de modelos dificeis se existir um grande numero de anomalias em relação ao que é considerado normal.

No caso das Isolation Forests, este algoritimo isola as anomalias em relação a forma como os seus caminhos divergem, geralmente com caminhos mais curtos que o que talvez fosse experado de dados não anomalos.

| Which | of the following statements about machine learning (ML) and security is incorrect ? |
|----------------------|--|
| _ a. | ML can be used to develop security solutions that can detect and respond to threats more quickly and effectively than traditional methods. |
| ○ b. | ML can be used to automate security tasks, such as log analysis and threat detection. |
| _ c. | ML can be used to identify and exploit security vulnerabilities in software and systems. |
| d. | ML is a silver bullet that can solve all security problems. |
| | |
| | |
| | |
| | |

| nt of a Bayesian spam filter? |
|--|
| s the probability of an email being spam based on the words and phrases it contains. |
| nat the filter uses to learn to identify spam emails. |
| |
| |

| Which o | of the following statements about bias and variance in machine learning is incorrect ? |
|----------------------|---|
| a. | Variance is the tendency of a machine learning model to make different predictions on the same data, depending on how the model is trained. |
| ○ b. | Bias and variance are inversely proportional, meaning that reducing bias will increase variance, and vice versa. |
| C. | Bias and variance can be reduced by using a more complex model. |
| ○ d. | Bias is the tendency of a machine learning model to make predictions that are consistently different from the true values. |
| | |
| | |
| | |
| | |
| | |

Which of the following is **not** a goal of the ML process?

- a. To understand the underlying patterns in the data.
- b. To develop a model that can make accurate predictions on new data.
- oc. To build a model that is robust to changes in the data.
- d. To minimize the cost of training and deploying the model.

Manuais

<u>Página de suporte com</u> <u>manuais</u> <u>para docentes e estudantes</u> Suporte

<u>bud.ua.pt</u> Extensão: 22299 Telf: +351 234 370099 Outros sites

<u>Universidade de Aveiro</u> <u>Notícias UA</u> <u>Formulários</u>