SLA
EASY WAY TO IT JOB

Share on your Social Media

Top 25 Machine Learning Interview Questions and Answers

Published On: April 1, 2022

Machine Learning Interview Questions and Answers

Machine Learning Process is the new-age technology adopted by companies that are striving hard for making information and services more accessible to people. Machine Learning along with Artificial Intelligence is the cause of growing improvement in the industries like finance, banking, retail, manufacturing, healthcare, and businesses. Here are the popular and frequently asked "*Machine Learning Interview Questions and Answers*" for clearing the technical rounds successfully.

Related Courses at SLA

→ **Machine Learning Training in Chennai**

→ **Machine Learning Online Training**

Related Posts

MEAN Stack Interview

1. What are the various types of machine learning processes?

There are three major types of machine learning as follows

- Supervised Learning is a model used to make predictions or decisions based on past or labeled data. Labeled data are the sets of data that are given tags or labels for making data meaningful
- Unsupervised Learning is the model that doesn't have labeled data and is used to identify patterns, anomalies, and relationships in the input data.
- Reinforcement Learning is the model that is based on the rewards received for its previous actions and operations. If the environment is given with a target. Every time the agent takes some action for reaching the target along with feedback. If it is positive feedback, the action will be moved forward. If it is negative feedback, the agent will go away from a goal.

2. Describe overfitting and how it can be avoided?

Overfitting is the situation that occurs when a model learns the training set very well to take up random fluctuations in the training data as concepts.

It impacts the model's ability to generate to doesn't apply to new data. When a model is given the training data, it returns 100% accuracy.

Technically a slight loss but if we use the test data, there will be errors and low efficiency. It can be avoided in multiple ways as follows.

- Regularization involves a cost term for the features with the objective function

- Making a simple model with lesser variables and parameters
- Cross-validation methods like k-folds
- Regularization techniques like LASSO can also be used to penalize the fault parameters.

3. What is the difference between Data Mining and Machine Learning?

Data Mining is the process in which the structured data tries to abstract knowledge or insights from unknown patterns.

During the process, machine learning algorithms are used.

Machine Learning defines the study, design, and development of the algorithms that provide the ability to the processes to learn without being explicitly programmed.

4. How Machine Learning differs from Deep Learning?

Machine Learning is about algorithms that are used to parse data, learn from the data, and then implement whatever they have learned to make informed decisions.

While Deep Learning is a part of machine learning that is inspired by the structure of the human brain for feature detection.

5. What are the various algorithm methods in Machine Learning?

There are many types of algorithm methods in Machine Learning and they have supervised learning, semi-supervised learning, unsupervised learning, transduction, and reinforcement learning.

6. What are the five popular algorithms used in the Machine Learning process?

The most used algorithms for machine learning processes are decision trees, probabilistic networks, neural networks, support vector machines, and

nearest neighbor algorithms.

7. What is ensemble learning?

In some models, classifiers are made and combined to solve a particular computational program is known as ensemble learning.

The ensemble learning methods are also referred to as committee-based learning or learning multiple classifier systems.

It is used to train various hypotheses to fix the same issue but one of the most suitable examples is the random forest trees where several decision trees are used to predict outcomes.

They are used to improve the classification, prediction, function approximation, etc of a model.

8. Define model selection in Machine Learning

Model Selection is the process of selecting models among diverse mathematical models that are used to define the same data.

It is implemented in the fields of statistics, machine learning, and data mining

9. What are the three stages of developing hypotheses or models in machine learning?

Following are the three major stages for building hypotheses or modes

- Model Building is used to choose a suitable algorithm for the model to train according to the needs of the given problem.
- Applying the Model that is responsible for checking the accuracy of the model through the test data

- Model Testing is the process to perform the required changes after testing and applying the final model.

10. Define 'Training Set' and 'Training Test'

The set of data used to discover the potential predictive relationship is known as 'Training Set'. It is the example that is given to the learner.

Training Set is used to test the accuracy of the hypothesis generated by the learner.

It is the set of instances held back from the learner. Hence, the training set is differing from the test set.

11. Describe the common ways to handle the missing data in a dataset

While working with data handling, missing data is one of the common factors and it is considered as one of the greatest challenges faced by data analytics.

It can be imputed the missing values using various ways. To handle the missing data in datasets the following methods will be used

- Deleting the rows
- Replacing with mean or median or mode
- Predicting the missing values
- Assigning a unique category
- Implementing algorithms to support missing values

12. What is ILP?

ILP is Inductive Logic Programming that is part of machine learning which uses logic programming.

It aims at searching patterns in data that can be used to build predictive models. In this process, the logic programs are assumed as a hypothesis.

13. Describe the important steps that involve in Machine Learning Process

One must follow several necessary steps to achieve a good working model for performing the machine learning process.

They include parameter tuning, data preparation, data collection, training the model, model evaluation, and prediction.

Machine Learning Engineer Salary

14. Define Precision and Recall

Precision and Recall are the measures that are used in the information retrieval domain for measuring how well an information retrieval system reclaims the related data as requested by the user.

Precision will be said as positive predictive value and it is the fraction of relevant instances among the received instances.

While Recall is the fraction of relevant instances that have been extracted over the total amount of relevant instances.

It is also known as sensitivity.

15. What are the functions of supervised learning?

In the Supervised Learning Process, there are some functions namely classification, speech recognition, regression, predict time series, and annotate strings.

16. List out the functions of unsupervised learning

Unsupervised learning is based on the following functions

- Finding clusters of the data
- Finding low-dimensional representations of the data
- Finding interesting directions of the data
- Finding novel observations or database

cleaning
- Finding interesting coordinates and correlations

17. What is Genetic Programming?

Genetic Programming is a subset of machine learning that implements an algorithm along with random mutation, a fitness function, crossover, and multiple generations of evolution that resolve user-defined tasks.

The genetic programming model is based on testing and selecting the best option out of a set of results.

18. Describe SVM in Machine Learning

SVM is the short form of Support Vector Machine and they have supervised learning models that are associated learning algorithms for analyzing the data used for classification and regression analysis.

It can handle combined binary classifiers and modified binary to incorporate multiclass learning.

19. Which is more important – Model Accuracy or Model Performance?

Model Accuracy is a subset of model performance. The accuracy of the model is directly proportional to the performance of the model.

The better performance of the model will be more accurate predictions.

20. What is bagging and boosting?

Bagging is the process of ensemble learning that is used to improve unstable estimation or classification schemes.

Boosting is the method that is used to sequentially reduce the bias of the combined models.

21. Define Cluster Sampling

Cluster Sampling is the process of randomly

selecting intact groups within a defined population that shares similar characteristics.

A cluster sample is a probability where each sampling unit is a collection or cluster of elements.

For instance, if we are clustering the total number of managers is a set of companies, the managers will represent elements and companies represent clusters.

22. What are the two main components of the Bayesian Logic Program?

The Bayesian Logic Program contains the following two main components

- Logical that contains a set of Bayesian Clauses that capture the qualitative structure of the domain
- Quantitative is used to encode the quantitative information of the domain.

23. Where can we implement supervised machine learning in modern businesses?

Supervised Machine Learning Processes can be applied in the following sectors

- Email Spam Detection: Here the training model will be used in historical data that contains emails categorized as spam or not spam. This labeled information is fed as input to the model
- Healthcare Diagnosis: The model is used to provide images regarding the diseases, a model will be trained to detect if a person is suffering from the disease or not
- Sentiment Analysis: The model is used to process algorithms to mine documents to determine whether they are positive, neutral, or negative in sentiment.
- Fraud Detection: The model is used to identify suspicious patterns to detect instances of possible fraud.

24. What are the two techniques that are followed in unsupervised machine learning?

Following are the two techniques used in unsupervised machine learning

- Clustering: It involves data to be divided into subsets. The subsets are also known as clusters that contain data that are similar to each other. Various clusters reveal different details about the objects, unlike regression or classification.
- Association: It involves identifying the patterns of associations between variables or items. For instance, an e-commerce website suggests items that are to buy based on their previous purchases or search histories or spending habits or wish list items, or purchase habits, etc.

25. What is a Decision Tree in Machine Learning?

Decision Trees are used in the supervised machine learning process where the data is continuously split as per certain parameters.

It develops classification or regression models as same as a tree structure with datasets that are broken up into smaller subsets while building the decision tree.

The tree will be defined by two entities namely decision nodes and leaves. The leaves are the decisions or the results, and the decision nodes are where the data is split.

Decision Trees are used to manage both categorical and numerical data.

Machine Learning Training

Conclusion

We will update **Machine Learning Interview Questions and Answers** as per the upgrades of technologies and tools involved in machine learning processes. Learn the Best **Machine Learning Training in Chennai** at Softlogic as we offer comprehensive machine learning concepts with practical understanding.

Share on your Social Media

Softlogic Academy

Softlogic Systems

KK Nagar [Corporate Office]

No.10, PT Rajan Salai, K.K. Nagar, Chennai – 600 078.
Landmark: Karnataka Bank Building
Phone: +91 86818 84318
Email: enquiry@softlogicsys.in
Map: Google Maps Link

OMR

No. E1-A10, RTS Food Street
92, Rajiv Gandhi Salai (OMR),
Navalur, Chennai - 600 130.
Landmark: Adj. to AGS Cinemas
Phone: +91 89256 88858
Email: info@softlogicsys.in
Map: Google Maps Link

Courses

Navigation

About Us

Blog Posts

Careers

Contact

Placement Training

Corporate Training

Hire With Us

Job Seekers

SLA's Recently Placed Students

Reviews

Sitemap

Important Links

Disclaimer

Privacy Policy

Terms and Conditions

Social Media Links

Python

Software Testing

Full Stack Developer

Java

Power BI

Clinical SAS

Data Science

Embedded

Cloud Computing

Hardware and Networking

VBA Macros

Mobile App Development

DevOps

Review Sources

Google

Trustpilot

Glassdoor

Mouthshut

Sulekha

Justdial

Ambitionbox

Indeed

Software Suggest

Sitejabber