Top Machine Learning Interview Questions and Answers for 2017

01 Jul 2017

According to a list released by the popular job portal Indeed.com on 30 fastest growing jobs in technology-

- Data science and machine learning jobs dominated the list of top tech jobs.
- Data scientist job postings saw an increase of 135% while machine learning engineer job postings saw an increase of 191% in 2017.
- 3 out of the top 10 tech job positions went to AI and data related positions, with machine learning jobs scoring a strong second place in the list.
- More than 10% of jobs in UK this year have been tech jobs demanding data science, machine learning and AI skills.

With the demand for machine learning engineers and data scientists outstripping the supply, organizations are finding it difficult to hire skilled talent and so are prospective candidates for machine learning jobs finding it difficult to crack a machine learning interview. Machine learning is a broad field and there are no specific machine learning interview questions that are likely to be asked during a machine learning engineer job interview because the machine learning interview questions asked will focus on the open job position the employer is trying to fill. For instance, if you consider a machine learning engineer job role for finance vs. a robotics job, both of them will be completely different in terms of data, architecture and the responsibilities involved. Machine learning engineer job role for robotics will require a candidate to focus working on Neural Networks based architecture while the machine learning tasks for finance will focus working more on Linear and Logistic regression algorithms.

A machine learning interview is definitely not a pop quiz and one must know what to expect going in. In our earlier posts, we have discussed about the different kind of big data interview questions and data science interview questions that are likely to be asked in an analytic job related interview. The main focus of this blog post is to give prospective machine teaching engineers a knick-knack of the kind of machine learning interview questions that are likely to be asked in a machine learning job interview.

Related Posts

Most of the machine learning interview questions listed in this article do not have a single good answer as the employer is looking to assess candidate's grasp on a particular topic instead of finding the right answer to the question.

Machine Learning Interview Questions and Answers

1) What is the difference between inductive machine learning and deductive machine learning?

In inductive machine learning, the model learns by examples from a set of observed instances to draw a generalized conclusion whereas in deductive learning the model first draws the conclusion and then the conclusion is drawn. Let's understand this with an example, for instance, if you have to explain to a kid that playing with fire can cause burns. There are two ways you can explain this to kids, you can show them training examples of various fire accidents or images with burnt people and label them as "Hazardous". In this case the

kid will learn with the help of examples and not play with fire. This is referred to as Inductive machine learning. The other way is to let your kid play with fire and wait to see what happens. If the kid gets a burn they will learn not to play with fire and whenever they come across fire, they will avoid going near it. This is referred to as deductive learning.

2) How will you know which machine learning algorithm to choose for your classification problem?

If accuracy is a major concern for you when deciding on a machine learning algorithm then the best way to go about it is test a couple of different ones (by trying different parameters within each algorithm) and choose the best one by cross-validation. A general rule of thumb to choose a good enough machine learning algorithm for your classification problem is based on how large your training set is. If the training set is small then using low variance/high bias classifiers like Naïve Bayes is advantageous over high variance/low bias classifiers like k-nearest neighbour algorithms as it might overfit the model. High variance/low bias classifiers tend to win when the training set grows in size.

3) Why is Naïve Bayes machine learning algorithm naïve?

Naïve Bayes machine learning algorithm is considered Naïve because the assumptions the algorithm makes are virtually impossible to find in real-life data. Conditional probability is calculated as a pure product of individual probabilities of components. This means that the algorithm assumes the presence or absence of a specific feature of a class is not related to the presence or absence of any other feature (absolute independence of features), given the class variable. For instance, a fruit may be considered to be a banana if it is yellow, long and about 5 inches in length. However, if these features depend on each other or are based on the existence of other features, a naïve Bayes classifier will assume all these properties to contribute independently to the probability that this fruit is a banana. Assuming that all features in a given dataset are equally important and independent rarely exists in the real-world scenario.

If you would like more information about Data Science Training, click the Request Info. button on top of this page.

4) How will you explain machine learning in to a layperson?

Machine learning is all about making decisions based on previous experience with a task with the intent of improving its performance. There are multiple examples that can be given to explain machine learning to a layperson –

- Imagine a curious kid who sticks his palm
- You have observed from your connections that obese people often tend to get heart diseases thus you make the decision that you will try to remain thin otherwise you might suffer from a heart disease. You have observed a ton of data and come up with a general rule of classification.
- You are playing blackjack and based on the sequence of cards you see, you decide whether to hit or to stay. In this case based on the previous information you have and by looking at what happens, you make a decision quickly.

5) List out some important methods of reducing dimensionality.

- Combine features with feature engineering.
- Use some form of algorithmic dimensionality reduction like ICA or PCA.
- Remove collinear features to reduce dimensionality.

6) You are given a dataset where the number of variables (p) is greater than the number of observations (n) (p>n). Which is the best technique to use and why?

When the number of variables is greater than the number of observations, it represents a high dimensional dataset. In such cases, it is not possible to calculate a unique least square coefficient estimate. Penalized regression methods like LARS, Lasso or Ridge seem work well under these circumstances as they tend to shrink the coefficients to reduce variance. Whenever the least square estimates have higher variance, Ridge regression technique seems to work best.

7) "People who bought this, also bought…." recommendations on Amazon are a result of which machine learning algorithm?

Recommender systems usually implement the collaborative filtering machine learning algorithm that considers user behaviour for recommending products to users. Collaborative filtering machine learning algorithms exploit the behaviour of users and products through ratings, reviews, transaction history, browsing history, selection and purchase information.

.8) Name some feature extraction techniques used for dimensionality reduction.

- Independent Component Analysis
- Principal Component Analysis
- Kernel Based Principal Component Analysis

9) List some use cases where classification machine learning algorithms can be used.

- Natural language processing (Best example for this is Spoken Language Understanding)
- Market Segmentation
- Text Categorization (Spam Filtering)
- Bioinformatics (Classifying proteins according to their function)
- Fraud Detection
- Face detection

10) What kind of problems does regularization solve?

Regularization is used to address overfitting problems as it penalizes the loss function by adding a multiple of an L1 (LASSO) or an L2 (Ridge) norm of your weights vector w.

11) How much data will you allocate for your training, validation and test sets?

There is no to the point answer to this question but there needs to be a balance/equilibrium when allocating data for training, validation and test sets.

If you make the training set too small, then the actual model parameters might have high variance. Also, if the test set is too small, there are chances of unreliable estimation of model performance. A general thumb rule to follow is to use 80: 20 train/test spilt. After this the training set can be further split into validation sets.

12) Which one would you prefer to choose – model accuracy or model performance?

Model accuracy is just a subset of model performance but is not the be-all and end-all of model performance. This question is asked to test your knowledge on how well you can make a perfect balance between model accuracy and model performance.

13) What is the most frequent metric to assess model accuracy for classification problems?

Percent Correct Classification (PCC) measures the overall accuracy irrespective of the kind of errors that are made, all errors are considered to have same weight.

14) When will you use classification over regression?

Classification is about identifying group membership while regression technique involves predicting a response. Both techniques are related to prediction, where classification predicts the belonging to a class whereas regression predicts the value from a continuous set. Classification technique is preferred over regression when the results of the model need to return the belongingness of data points in a dataset to specific explicit categories. (For instance, when you want to find out whether a name is male or female instead of just finding it how correlated they are with male and female names.

15) Why is Manhattan distance not used in kNN machine learning algorithm to calculate the distance between nearest neighbours?

Manhattan distance has restrictions on dimensions and calculates the distance either vertically or horizontally. Euclidean distance is better option in kNN to calculate the distance between nearest neighbours because the data points can be represented in any space without any dimension restriction.

Role Specific Open Ended Machine Learning Interview Questions

1) Given a particular machine learning model, what type of problems does it solve, what are the assumptions the model makes about the data and why it is best fit for a particular kind of problem?

 Suggested Reading to answer the above question - Machine Learning Algorithms

2) Is the given machine learning model prove to over fitting? If so, what can you do about this?

3) What type of data does the machine learning model handle –categorical, numerical, etc.?

4) How interpretable is the given machine learning model?

5) What will you do if training results in very low accuracy?

6) Does the developed machine learning model have convergence problems?

7) Which is your favourite machine learning algorithm? Why it is your favourite and how will you explain about that machine learning algorithm to a layperson?

8) What approach will you follow to suggest followers on Twitter?

9) How will generate related searches on Bing?

10) Which tools and environments have you used to train and assess machine learning models?

11) If you claim in your resume that you know about recommender systems, then you might be asked to explain about PLSI and SVD models in detail.

12) How will you apply machine learning to images?

Machine Learning Interview Questions asked at Top Tech Companies

Machine Learning Interview Questions asked at Amazon

1) How will you weigh 9 marbles 3 times on a balance scale to find the heaviest one?

2) Why is gradient checking important?

3) What is loss function in a Neural Network?

4) Which one is better – random weight assignment or assigning same weights to the units in the hidden layer?

5) How will you design a spam filter?

6) Explain the difference between MLE and MAP inference.

7) Given a number, how will you find the closest number in a series of floating point data?

8) What is boosting?

Machine Learning Interview Questions asked at Baidu

1) What are the reasons for gradient descent to converge slow or not converge in various machine learning algorithms?

2) Given an objective function, calculate the range of its learning rate.

3) If the gradient descent does not converge, what could be the problem?

4) How will you check for a valid binary search tree?

Machine Learning Interview Questions asked at Spotify

1) Explain BFS (Breadth First Search algorithm)

2) How will you tell if a song in our catalogue is a duplicate or not?

3) What is your favorite machine learning algorithm and why?

4) Given the song list and metadata, disambiguate artists having same names.

5) How will you sample a stream of data to match the distribution with real data?

Machine Learning Interview Questions asked at Capital One

1) Given two years of transaction history, what features will you use to predict the credit risk?

2) Differentiate between gradient boosted tree and random forest machine learning algorithm.

3) How will you use existing features to add new features?

4) Considering that you have 100 data points and you have to predict the gender of a customer. What are the difficulties that could arise?

5) How will you set the threshold for credit card fraud detection model?

A machine learning interview is a compound process and the final result of the interview is determined by multiple factors and not just by looking at the number of right answers given by the candidate. If you really want that machine learning job, it's going to take time and dedication as you practice multiple ways to answer the above listed machine learning interview questions, but hopefully it is the enjoyable kind. You will learn a lot and get a good deal of knowledge preparing for your next machine learning interview with the help of these questions.