MACHINE LEARNING INTERVIEW QUESTIONS

ASKED IN REAL DATA SCIENCE INTERVIEWS

-Nitin Mukesh

Commonly Asked Topics

- **Logistic regression**
- **KNN**
- **SVM**
- **Decision Tree**
- **Random Forest**
- **Bagging**
- **Boosting**
- **LDA & QDA**
- **Naive Bayes classifier**
- **Classification Metrics**
- **Overfitting and underfitting**

- Difference between logistic and SVM. **(Accenture)**
- How to deal with overfitting in Logistic Regression? **(Delhivery)**
- If the accuracy of the model is 95%, is it good? **(ANZ)**
- Does standardization have any impact on the performance of the model in the case of a random forest? **(Wells Fargo)**
- How to handle imbalanced data in classification? Which metric to use in this case? Why is accuracy not a good measure in this case?**(FCS Limited)**
- Explain the Logistic regression model. **(ANZ, Miko.ai, IPSOS)**
- Suppose you have 200 variables, will you consider all the variables to build a model? **(ANZ)**
- Difference between random forest and logistic regression? **(ANZ)**
- What is a random forest and how it differs from Adaboost? **(Meru cabs)**
- Explain bagging in Machine learning? **(TCS innovation lab)**
- Explain F score, precision, recall, ROC curve. **(Meru cabs)**
- When can you say, KNN model is overfitted? **(Delhivery)**

- Write the loss function of logistic and linear regression. **(Delhivery)**
- How is the prediction made in the random forest? **(FCS Limited)**
- How is XGBoost different from gradient boosting? **(Wells Fargo)**
- What will happen if we use MSE in logistic regression? **(Delhivery)**
- If you have categorical features in your data, then how will you find the Nearest Neighbor using KNN? **(ANZ)**
- Explain the working of the Support Vector Machine Algorithm? **(ICICI Lombard, Cognizant)**
- What is the kernel trick in SVM? Why is it computationally efficient? **(FCS Limited)**
- Explain the complete working of the decision tree algorithm. **(WNS Global, TCS innovation lab)**
- What is a confusion matrix? **(ANZ)**
- Difference between Linear regression vs logistic regression. **(Wells Fargo)**
- What is underfitting and overfitting for example? **(Accenture)**
- What are the steps to build a model from scratch? **(ANZ)**
- Why random forests are called random? **(FCS Limited)**

- Can Ridge or Lasso be used for NLP? **(Miko.ai, NLP)**
- What is the difference between Bagging and Boosting? **(Miko.ai, TCS Innovation Lab)**
- Differentiate between SVM and Random Forest? **(Miko.ai)**
- Which is better, Random forest or decision tree? **(TCS Innovation Lab)**
- What is the Ginni index and entropy in the decision tree? **(Cognizant)**
- What do you mean by deviance in Logistic regression? **(ANZ)**
- What is a convex function? **(Amazon)**
- What are the criteria to evaluate a Logistic regression model? **(MMA IPSOS)**