

Company Name - Fractal

Role - Data Scientist

Q. Difference between array and list

A. The main difference between these two data types is the operation you can perform on them. Lists are containers for elements having differing data types but arrays are used as containers for elements of the same data type.

Q. Which is faster dictionary or list for look up

A. Dictionary is faster because you used a better algorithm. The reason is because a dictionary is a lookup, while a list is an iteration. Dictionary uses a hash lookup, while your list requires walking through the list until it finds the result from beginning to the result each time.

Q. How much time SVM takes to complete if 1 iteration takes 10sec for 1st class.

And there are 4 classes.

A. It would take $4 \times 10 = 40$ seconds to train one-vs-all method one to one.

Q. Kernels in svm, there difference?

A. Kernel Function in SVM is a method used to take data as input and transform into the required form of processing data.

Gaussian Kernel Radial Basis Function (RBF) : It is used to perform transformation, when there is no prior knowledge about data and it uses radial basis method to improve the transformation.

Sigmoid Kernel: this function is equivalent to a two-layer, perceptron model of neural network, which is used as activation function for artificial neurons.

Polynomial Kernel: It represents the similarity of vectors in training set of data in a feature space over polynomials of the original variables used in kernel.

Linear Kernel: used when data is linearly separable.

For detailed answer, look over [this](#) link.

Q. What is the difference between squared error and absolute error?

A. Squared error is the squared difference between the predicted values and the actual value. **Absolute Error is the** difference between the measured value and true value. The squared error is differentiable everywhere, whereas the absolute error is not (its derivative is undefined at 0). This makes the squared error more susceptible to mathematical optimization strategies.

Q. Under what aspects Naive Bayes is bad?

A. In Naive Bayes, all predictors (or traits) are assumed to be independent, which is rarely the case in real life. This limits the algorithm's usability in real-world scenarios. The 'zero-frequency problem' occurs when an algorithm assigns zero probability to a categorical variable whose category in the test data set was not present in the training dataset. To get over this problem, you should employ a smoothing approach. You shouldn't take its probability outputs seriously because its estimations can be off in some instances.

Q. How will you tackle an exploding gradient problem?

A. The problem of exploding gradients can be solved by reducing the number of layers in the network. Using a smaller batch size when training the network may also be beneficial. Long Short-Term Memory (LSTM) memory units and maybe similar gated-type neuron structures can be used to reduce exploding gradients. When you have an exploding gradient problem, you can use gradient clipping. First, we choose a threshold value, and if the value produced by the gradient function is greater than this threshold, we change it to something else. It may also be beneficial to use suitable weight initialization techniques.

Q. How will you prevent overfitting when creating a statistical model?

A. Cross-validation is an effective tool for avoiding overfitting. The aim is to create many micro train-test splits using your initial training data. These divisions can be used to fine-tune your model. More data can help algorithms recognize the signal more accurately. Also, make a list of relevant variables and terms that you'll probably use in your own model.

Q. Plots to evaluate models?

A. Residual plots, validation curve, gain and lift chart, kolmogorov smirnov chart.

Q. Algorithms which solve overfitting problem?

A. PCA, Ridge regression, L1/L2 regularization etc.

Q. What is Confounding variables?

A. Confounding variable is a variable that is not included in an experiment, yet affects the relationship between the two variables in an experiment.

Q. How can you convert timestamps to date time in MySQL?

A. We can convert the timestamp to date time with the help of `FROM_UNIXTIME ()` function.

Q. How does random Random Forest works? How is it different from decision trees?

A. A decision tree combines some decisions, whereas a random forest combines several decision trees. Thus, it is a long process, yet slow. Whereas, a decision tree is fast and operates easily on large data sets, especially the linear one. The random forest model needs rigorous training.

Q. Different activation function?

A. Binary Step Function, Linear Activation Function, Sigmoid/Logistic Activation Function, Tanh Function (Hyperbolic Tangent), ReLU Activation Function

Q. TF/IDF vectorization?

A. Tf-Idf vectorization gives a numerical representation of words entirely dependent on the nature and number of documents being considered. The same words will have different vector representations in another corpus.

NOTE:

If you want to learn Data Science from scratch and become interview ready, check out our popular Data science course.

<https://ai.cloudymml.com/Learn-Data-Science-from-Scratch>

