Things to study for data science interviews

Probability and statistics:

- Distributions and their nature (kurtosis , moments , deviations from normal, skewness) (normal distribution in depth)
- Central limit theorem
- Bayes theorem and some questions on it
- Hypothesis testing (critical value and p value)
- Kernel density estimation
- q-q plot (to check whether distribution is normal or not
- Box - cox and other transformations(log , exponential)
- Co-variance
- Pearson correlation coefficient
- Spearman ranking
- Confidence intervals(v important)
- Types of sampling (like random , stratified)
- Tests for similarity of two distributions (k-s test , permutation test etc)

A/B testing

Traditional A/B testing

- Using critical value hypothesis test
- Using p value hypothesis test
- Using Z test

Drawbacks of traditional A/B testing

Bayesian A/B testing

- Explore exploit dilemma
- Epsilon-greedy
- Optimised initial value theory

- Upper confidence bound theory
- Thompsons sampling (using conjugate priors)

Machine learning

Linear regression
Logistic
Decision trees
ensembles (bagging boosting)
Svm
Perceptron
Knn
Kmeans
Hierarchical clustering
Density based clustering

dl(not much imp)
backpropogation .
Activation functions
Weights initialization
optimisations
Cnn
Rnn (elman unit)
Vanish gradient problem
Lstm

reinforcement learning (markov decision chain models)