Important interview

How can outliers be treated in a dataset?
When should normalization be used, and when is standardization more appropriate?
Why might RMSE be considered less ideal for evaluating regression models?
What are some popular machine learning algorithms, and what loss functions are commonly associated with them?
Why is the logistic function often unsuitable for multiclass classification problems?
What methods can be used to make a time series dataset stationary?
What are different weight initialization techniques used in neural networks?
When is it more appropriate to use MAE versus MSE for evaluating regression models?
Under what circumstances should Gini impurity be used instead of entropy for decision tree algorithms?
How can multicollinearity be addressed in a dataset?
What strategies can be employed to handle data security in machine learning projects?
Which machine learning algorithms are generally unaffected by imbalanced datasets?
What are different encoding techniques for categorical data?
What are the different types of models used in various machine learning algorithms?
What is Hamming distance and in what contexts is it used?
What are two methods to handle limited RAM issues in deep learning and machine learning?
What is the difference between linearly separable and non-linearly separable data?
What are common distance metrics used in machine learning and data analysis, and what types of data are they suitable for?
How does Adaboost differ from XGBoost?
How can categorical data be prepared for distance-based algorithms like Euclidean distance?
Which clustering algorithm is typically faster: K-means, Hierarchical Clustering, or DBSCAN?
Why are RNNs preferred over CNNs for NLP tasks?
What defines a standard normal distribution?
What is the relationship between sample mean and population mean?

What are edge devices in the context of machine learning and data science?
Explain Chebyshev's inequality and its significance in statistics.
What is the Out-of-Bag (OOB) score in Random Forest, and how is it used?
How can the relationships between three categorical columns be visualized? What types of plots are best suited for this purpose?
Why is model explainability important, and what techniques can be used for local interpretation of individual predictions?
What are the advantages of using k-means++ for the initialization parameter in the K-means algorithm?
What are the maximum values for entropy and Gini impurity in binary classification?
What are the different types of ensemble techniques, and which algorithms are used within each type?
If a model has low bias and high variance, what does this indicate and how can it be addressed?
Can you provide two examples of datasets that exhibit negative skewness?
How is the number of features determined in Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA)?
What is the advantage of DBSCAN over K-means clustering?
How can the optimal number of clusters be validated in clustering analysis?
What is mini-batch K-means, and how does it differ from traditional K-means?
What are the different types of ensemble techniques, and which algorithms can be used within each type?
What are class weights, and in which situations are they used?
Why is KNN not commonly used in industry despite its simplicity?
Can you explain the differences between Ball Tree, KD Tree, Auto, and Brute methods for nearest neighbor searches, and their respective use cases?
How can a normal distribution be converted to a standard normal distribution?
What methods can be used to address data shortage issues?
How can we determine if we have sufficient data or if more data is needed for a deep learning project?
What are the key assumptions of linear regression?
How should one decide whether to use a machine learning model or a deep learning model for a given problem?
What role do activation functions play in deep learning?
Is class imbalance an issue for classification problems, regression

problems, or both?
What is the difference between categorical and object data types in pandas?
How does deep learning handle outliers, and what impact can they have on model performance?
What are some statistical methods for identifying outliers in a dataset?
Can the R^2 value of a regression model be negative, and if so, what does it indicate?
What is the difference between epochs and iterations in the context of training a deep learning model?
What regularization techniques are commonly used in Convolutional Neural Networks (CNNs)?
How should the value of k be selected when using the K-Nearest Neighbors (KNN) algorithm?
Do outliers affect the performance of K-Nearest Neighbors (KNN), and if so, how?
What is Bessel's correction, and why is it used in statistical calculations?
What are the skewness and kurtosis values for a standard normal distribution?
Difference between confusion matrix and classification report?
What is the significance of p-value?
How regularly must an algorithm be updated?
Describe Markov chains?
Difference between an error and a residual error
Difference between Point Estimates and Confidence Interval
Which is faster, python list or Numpy arrays, and why?
What are decorators in python?
What is the difference between a parametric and a non-parametric test?
When not to use PCA for dimensionality reduction?
What is the difference between catboost and XGboost?
How can you convert a numerical variable to a categorical variable and when can it be useful?
What are generalized linear models?
What is the difference between ridge and lasso regression? How do they differ in terms of their approach to model selection and regularization?
Why is SVM called a large margin classifier?
What will happen if we increase the number of neighbors in KNN?
What is the difference between extra trees and random forests?
What is the problem with using label encoding for nominal data?

What can be an appropriate encoding technique when you have hundreds of categorical values in a column?
What is the significance of C in SVM?
How do c and gamma affect overfitting in SVM?
How do you choose the number of models to use in a Boosting or Bagging ensemble?
How does the Naive Bayes algorithm compare to other supervised learning algorithms?
Can you explain the concept of the “kernel trick” and its application in Support Vector Machines (SVMs)?