Algorithm Understanding
Is SVM (Support Vector Machine) a supervised or unsupervised learning algorithm?
Why is SVM such a powerful classification method?
What are 3 disadvantages of SVMs?

SVMs are supervised learning models with learning algorithms for data classification and regression analysis. For example, SVM models take features and target values and based on these information, they can give a prediction (target value) of a (set of) feature(s).
Some of the biggest benefits of using SVMs are:
_Perform well when we don't know much about the data
_Non-linear kernels can be used.
-Produce a clear margin of separation
-Effective in high dimensional spaces
-Can be used in cases where the number of features are higher than the number of samples
-Use a subset of training data points in the decision function and therefore is memory efficient
,however, SVMs have their own shortcomings such as:
-Choosing an optimal kernel and parameters can be expensive
-Doesn't perform well with large data sets (The training time increases greatly)
-Doesn't do a good job with noisy datasets, ex: overlapping target classes
-Doesn't provide probability estimates directly, these estimates are calculated over expensive cross-validations.

Interview Readiness
What is the time complexity of SVM?
What is it for Logistic Regression?

The time complexity of SVM is heavily dependent on SVM type and the used kernel. The time complexity of SVM is generally between O(n^2) and O(n^3) with "n" being the number of training instances.
The time complexity of Logistic Regression is O(d), in which "d" is the vector size of "w" and "w" is the vector of weights in logistic regression.

Interview Readiness

Explain feature importance for the Random Forest algorithm? When examining feature importance, what is Gini impurity or information gain?

In random forest, feature importance is calculated by the decrease in node impurity, weighted by the probability of reaching that specific node. The probability of the node can be calculated by the number of samples that reach the node, divided by the total number of samples. When the values are higher, the feature is more important.

. Gini : $\text{Gini}(E) = 1 - \sum_{j=1}^{c} p_j^2$

. Entropy : $H(E) = -\sum_{j=1}^{c} p_j \log p_j$

Gini impurity or information gain provide us with metrics that can help us build a better model.
Gini impurity is a measure of node purity (or impurity). It is a measure of how often a randomly chosen variable will be misclassified. Gini impurity calculate each feature importance as the sum over of splits across all trees that include the feature., proportional to the number of samples splits.
Information gain is the difference between uncertainty of the starting node and weighted impurity of the two child node. Information gain helps us deciding which feature should be used to split the data. In another words, it measures the reduction of surprise when splitting a dataset.

Interview Readiness
SHAP (SHapley Additive exPlanations) is a game theoretic approach to explain the output of any machine learning model, what is it and how does it work?

SHAP is a mathematical method to explain the prediction of ML models. This model can be used to explain the prediction of any ML model by calculating the contribution of each feature to the prediction. This model connects optimal credit allocation with local explanations using Shapley values from game theory and their related extensions.
SHAp quantifies the contributions that each player (feature) brings to the game (prediction). This model quantify the contribution of each feature by calculating the deviation of the predicted outcome when a feature or a set of features are absent in the calculations. For example, to explain an image, pixels can be grouped to super pixels and then we can distribute the predictions among these super pixels.