**1. What is your definition of clustering? What are a few clustering algorithms you might think of?**

**Ans:** Clustering is a type of unsupervised learning where data points are grouped together based on their similarities. Some clustering algorithms include K-Means, Hierarchical Clustering, DBSCAN (Density-Based Spatial Clustering of Applications with Noise), and Gaussian Mixture Models (GMM).

**2. What are some of the most popular clustering algorithm applications?**

**Ans:** Some popular applications of clustering algorithms include customer segmentation in marketing, image segmentation in computer vision, document clustering in natural language processing, anomaly detection in cybersecurity, and gene expression analysis in bioinformatics.

**3. When using K-Means, describe two strategies for selecting the appropriate number of clusters.**

**Ans:** Elbow method: Plotting the within-cluster sum of squares against the number of clusters and selecting the point where the rate of decrease sharply changes, resembling an elbow.

Silhouette analysis: Computing the silhouette score for different numbers of clusters and choosing the number that maximizes the silhouette score, indicating the quality of the clustering.

**4. What is mark propagation and how does it work? Why would you do it, and how would you do it?**

**Ans:** Mark propagation is a semi-supervised learning technique where labels are propagated from labeled data points to unlabeled data points based on the similarity between data points. It helps in the propagation of known information to unknown data points, potentially aiding in the creation of a more comprehensive labeled dataset for training models.

**5. Provide two examples of clustering algorithms that can handle large datasets. And two that look for high-density areas?**

**Ans:** Large datasets: K-Means with mini-batch variations, and hierarchical clustering with optimized implementations.

High-density areas: DBSCAN (Density-Based Spatial Clustering of Applications with Noise), and OPTICS (Ordering Points To Identify the Clustering Structure).

**6. Can you think of a scenario in which constructive learning will be advantageous? How can you go about putting it into action?**

**Ans:** Constructive learning can be advantageous in scenarios where the learner needs to progressively build its understanding from limited initial knowledge. It can be implemented by starting with a small subset of data, incrementally adding more data points, and continuously updating the model's understanding to adapt to new information.

**7. How do you tell the difference between anomaly and novelty detection?**

**Ans:** Anomaly detection focuses on identifying instances that significantly deviate from the norm, while novelty detection aims to identify new instances that differ from the training data. Anomaly detection often deals with outliers or unexpected events, whereas novelty detection focuses on detecting previously unseen patterns.

**8. What is a Gaussian mixture, and how does it work? What are some of the things you can do about it?**

**Ans:** Gaussian Mixture is a probabilistic model that assumes all data points are generated from a mixture of several Gaussian distributions. It works by estimating the parameters of the Gaussian components to fit the data distribution. It can be used for density estimation, clustering, and generating new samples from the learned distribution.

**9. When using a Gaussian mixture model, can you name two techniques for determining the correct number of clusters?**

**Ans:** Bayesian Information Criterion (BIC): Penalizes complex models to prevent overfitting and aids in selecting the appropriate number of clusters.

Akaike Information Criterion (AIC): A measure of the relative quality of a statistical model that helps in selecting the number of components in the Gaussian mixture model based on the goodness of fit and complexity of the model.