KNN Interview Questions:

1.What are the key hyperparameters in KNN?

2.What distance metrics can be used in KNN?

**1. Key Hyperparameters in KNN**

K-Nearest Neighbors (KNN) is a simple yet powerful algorithm that classifies data points based on the majority vote of their nearest neighbors. The key hyperparameter in KNN is:

* **K:** This parameter determines the number of nearest neighbors to consider when classifying a data point. A larger value of K can reduce the impact of noise but may make the model less sensitive to local patterns. A smaller value of K can be more sensitive to noise but may capture finer-grained patterns.

**2. Distance Metrics in KNN**

The choice of distance metric significantly impacts the performance of KNN. Here are some common distance metrics used in KNN:

* **Euclidean Distance:**
  + The most commonly used distance metric.
  + Calculates the straight-line distance between two points in Euclidean space.
  + Formula:
  + d(x, y) = sqrt((x1 - y1)^2 + (x2 - y2)^2 + ... + (xn - yn)^2)
* **Manhattan Distance:**
  + Calculates the sum of the absolute differences between the corresponding coordinates of two points.
  + Formula:
  + d(x, y) = |x1 - y1| + |x2 - y2| + ... + |xn - yn|
* **Minkowski Distance:**
  + A generalization of Euclidean and Manhattan distances.
  + Formula:
  + d(x, y) = (|x1 - y1|^p + |x2 - y2|^p + ... + |xn - yn|^p)^(1/p)
    - When p = 1, it's Manhattan distance.
    - When p = 2, it's Euclidean distance.

The choice of distance metric depends on the specific characteristics of the data and the desired properties of the KNN model. For example, Euclidean distance is often used for continuous numerical features, while Manhattan distance can be more robust to outliers.