The K-Nearest Neighbor ( KNN ) algorithm is a supervised learning method used for classification and regression. It is a non-parametric , Instance based algorithm , meaning it does not make assumptions about the underlying data distribution. Instead , it makes predictions based on similarity between data points.

KNN working

* Step 1 :
* **Choose the number of neighbors (K):** K represents the number of closest data points (neighbors) used to make predictions.
* A small “k” (e.g. 1 or 3) may cause overfitting (too sensitive to noise)
* A large “k” (e.g. 10 or 15) may cause under fitting (to general and ignoring patterns).
* Step 2 :
* **Find the nearest neighbor**: In order to find the nearest neighbor, we measure the distance between the input data point and all other points in the dataset. Common distance metrics include
* Euclidean Distance ( most common )
* Manhattan Distance
* **Calculate distance** : formula for Euclidean distance ( between two ✌ points A(x1,y1) and B(x2,y2)
* Step 3:
* **Find the nearest neighbors**: After calculating distance, the algorithm selects the “k” closest data points.
* Step 4:
* **For classification**: The most common class (majority vote) among the k neighbors is chosen.
* **For Regression:** The average, value of the K neighbors are taken.

