K-nearest Neighbours Classification (knn)

KNN also known as K-nearest neighbour is a [supervised and pattern classification learning algorithm](https://ranvir.xyz/blog/how-to-evaluate-your-machine-learning-model-like-a-pro-metrics/#supervised-learning-and-classification-problems) which helps us find which class the new input(test value) belongs to when k nearest neighbours are chosen and distance is calculated between them. It attempts to estimate the conditional distribution of Y given X, and classify a given observation(test value) to the class with highest estimated probability.

It first identifies the k points in the training data that are closest to the test value and calculates the distance between all those categories. The test value will belong to the category whose distance is the least.

Supervised machine learning algorithms are used to solve classification or regression problems.

The KNN Algorithm simply do:

* **Step-1:** Select the number K of the neighbours
* **Step-2:** Calculate the Euclidean distance of **K number of neighbours**
* **Step-3:** Take the K nearest neighbours as per the calculated Euclidean distance.
* **Step-4:** Among these k neighbours, count the number of the data points in each category.
* **Step-5:** Assign the new data points to that category for which the number of the neighbour is maximum.
* **Step-6:** Our model is ready.

What is the Advantage of KNN Algorithm:

* It is simple to implement.
* It is robust to the noisy training data
* It can be more effective if the training data is large.

KNN can be helpful with disease predictions. For instance, with KNN can predicting someone is having diabetes or not, final label can be 1person have diabetes or 0 person hasn’t got diabetes.

Or it can be useful with prediction of pictures. For instance with enough data test KNN can predict picture if its cat or not.