

k-Nearest Neighbours

Epoch IIT Hyderabad

Arin Aggarwal

MA22BTECH11006

kNN is a classifier algorithm which determines the class of a given point by observing its k-nearest neighbours' classes. The class with majority of nearest neighbour points is given to the point.

This is a very simple algorithm. It is a lazy algorithm as it needs to remember its training data instead of just having a standardized result to be used generally.

We need to find optimal value for k . A common value is $\frac{\sqrt{n}}{2}$ where n is the number of points in training dataset. Accuracy v/s k can be drawn to choose optimal value for k . Also, generally odd k value is chosen so that a tie between 2 classes does not arise. For ex. it may happen that there are 5-5 nearest neighbours for $k = 10$.