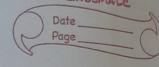# K - Nearest Neighbor (KNN)

KNN is a supervised ML algorithm. It is mostly used for classification, but can be used for regression also. It considers K-Nearest Neighbors (data points) to predict the class for a new data-point.

The algorithm's learning is -

- **Instance-based** -> Here the model does not learn weights from training data to predict output (as in model-based learning), but uses training instances to predict output for new data.

- **Lazy learning** -> The model does not learn anything during the training stage. It just stores the training data. At the prediction stage, ie The learning is done only at the prediction stage.

- **Non-parametric** -> In KNN, there is no pre-defined form of mapping function.

☑ **Pseudocode for KNN :-**

1. Find the optimal value of K

2. Calculate the distance of the new data-point from all the training data-points.

3. Sort these distances in ascending order.

4. Select the top k rows

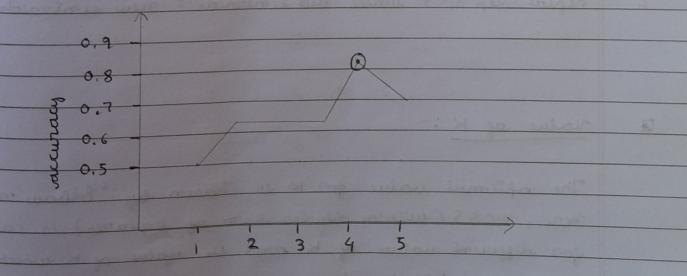5. Find the most frequent clan from these k rows and this will be the predicted clan. (For Classification)

||

Find the mean of k rows and this will be the predicted value. (For Regression)

Value of k :-

There is no particular way to determine best value of k.

It can be done by -
Selecting a range of values for k. Then fitting the model with each value of these values and finding their accuracy scores. The value of k which gives the highest accuracy is chosen.

K =

values of k