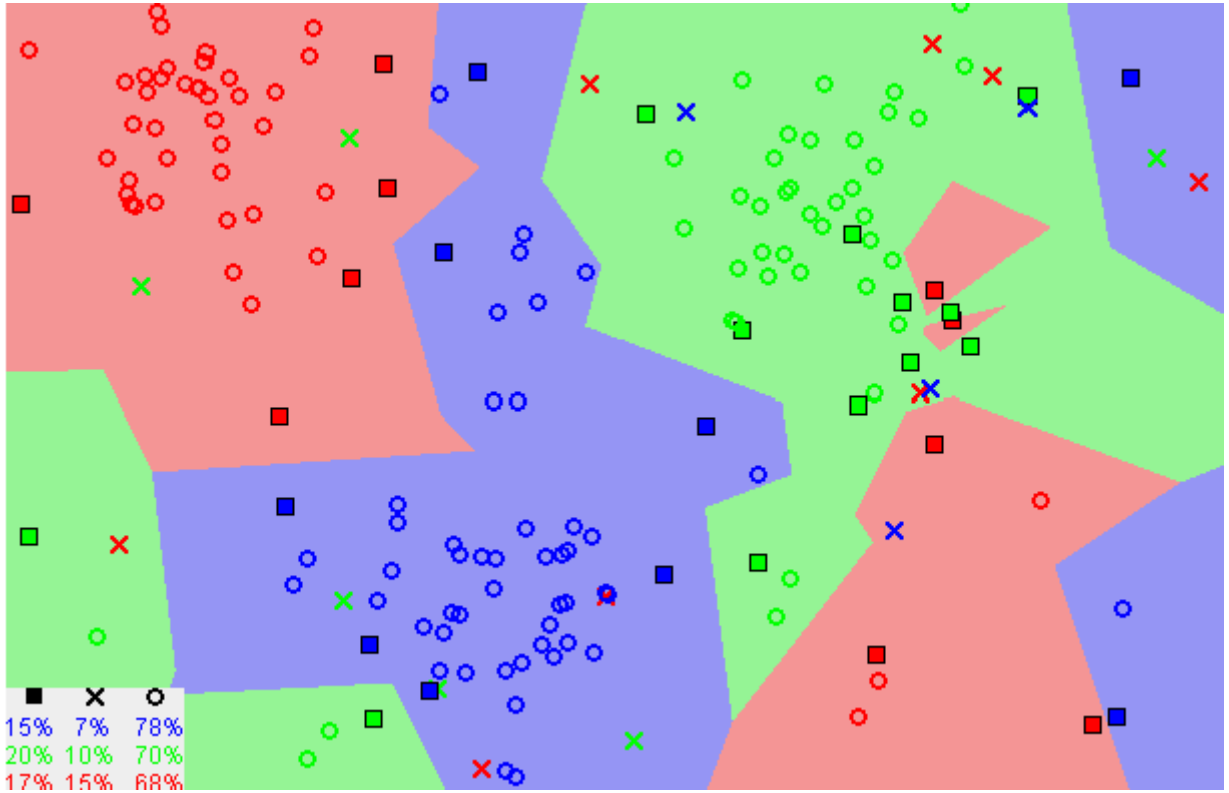
KNN Algorithm

By : L Vinay Rajiv Reddy

What is KNN Algorithm ?

- Simple, Easy-to-implement, Supervised Algorithm
- Classification & Regression problems
- “Similar things are near to each other”



Steps:

1. **Load** the data
2. **Initialize K** to your chosen number of neighbors
3. For each example in the data **3.1 Calculate the distance** between the query example and the current example from the data.
   3.2 Add the distance and the index of the example to an ordered collection
4. **Sort** the ordered collection of distances and indices from smallest to largest (in ascending order) by the distances
5. **Pick** the first K entries from the sorted collection
6. Get the **labels** of the selected K entries
7. If **regression**, return the mean of the K labels
8. If **classification**, return the mode of the K labels

Note:

As k decreases to 1, system becomes less stable.

Advantages:

→ Simple,

→ Algorithm is versatile,

→ No need to build a separate model.

Dis advantages:

→ Algorithm gets significantly slower as the number of examples and/or predictors/independent variables increase.

Application:

- Recommender System.


KNN BE LIKE
"Show me your friends, and I'll tell you who you are."