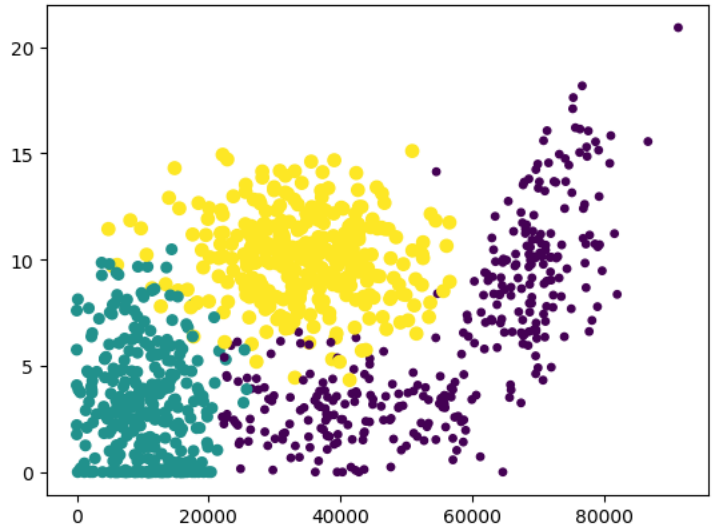
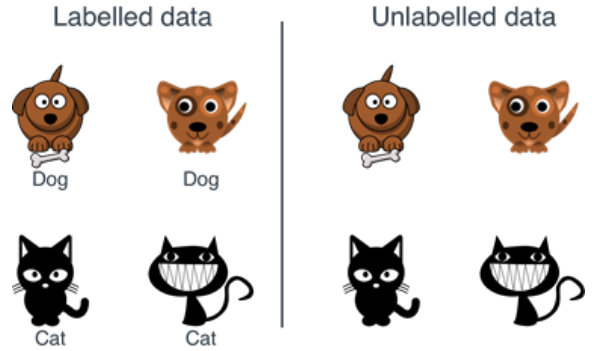
Siddhardhan

K-Nearest Neighbors (KNN) - intuition

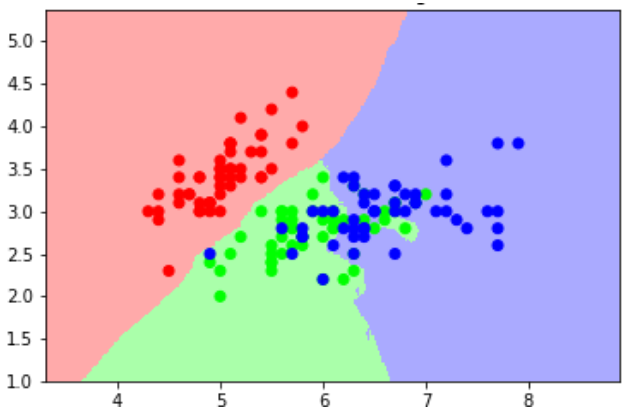


K-Nearest Neighbors

About K-Nearest Neighbors:

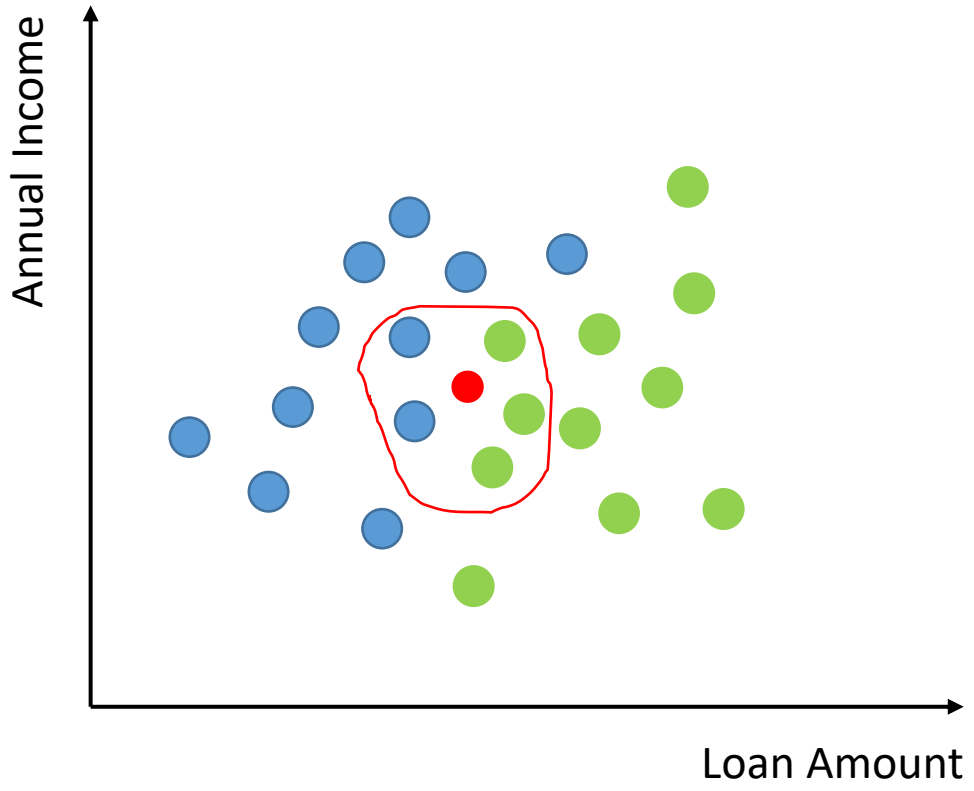


1. Supervised Learning Model
2. Used for both Classification & Regression
3. Can be used for non-linear data
4. K - Neighbors



K-Nearest Neighbors

Classification Problem:



K = 5

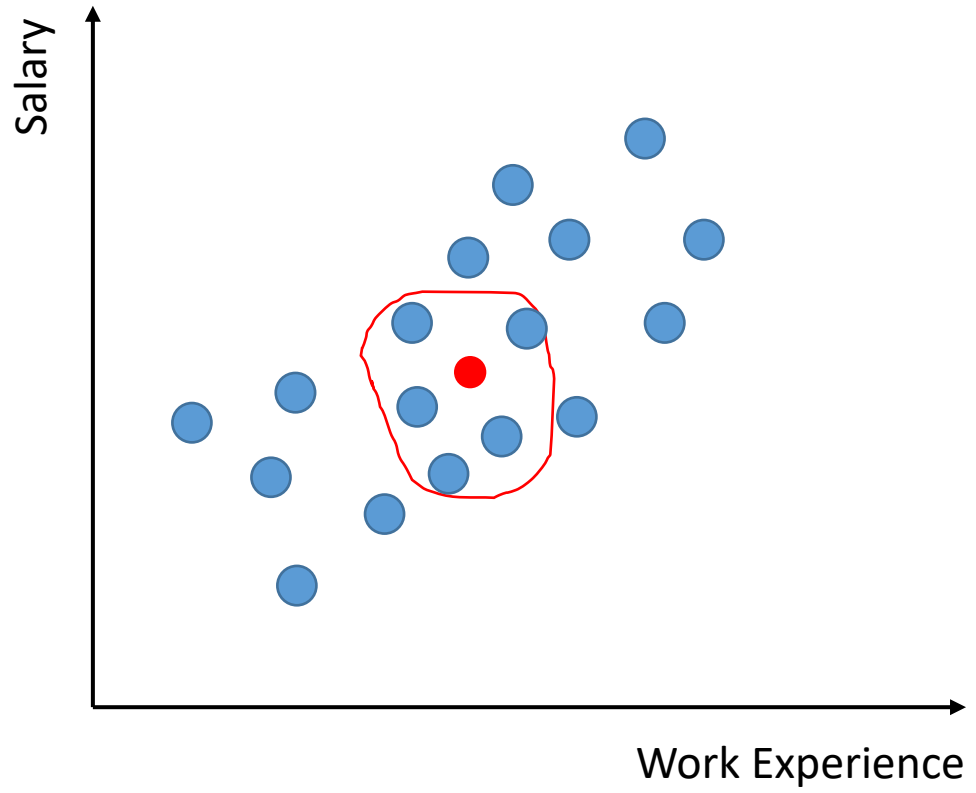
- Didn't repay on time
- Repaid on time
- May not repay the loan on time

To Measure the distance between the data points:

- Euclidean Distance
- Manhattan Distance

K-Nearest Neighbors

Regression Problem:



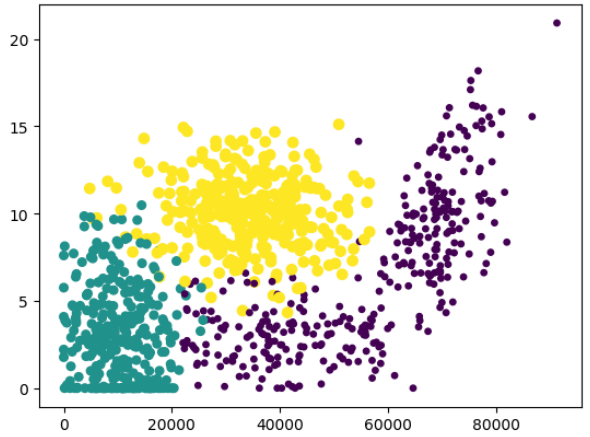
K = 5

Salary of the person can be calculated as the mean of 5 nearest neighbors.

K-Nearest Neighbors

Advantages:

1. Works well with smaller datasets with less number of features
2. Can be used for both Classification & Regression
3. Easy to implement for Multi-class classification problems
4. Different distance criteria can be used
   (eg: Euclidean Distance, Manhattan Distance)



Disadvantages:

1. Choosing optimum “K” value
2. Less efficient with high dimensional data.
3. Doesn’t perform well on imbalanced dataset
4. Sensitive to Outliers