

Regression - NN





K – Nearest Neighbours(Regression)

- It assumes that similar things exist in close proximity.

Algorithm:

- * Step 1: Choose the no. K of neighbours
- * Step 2: Take the K nearest neighbours of the new data points by Euclidean distance
- * Step 3: Among K Neighbours, take the average of these Neighbours
- * Step 4: Assign this average value to the new point

Rishi Bansal



K-NN(Regression)

- Dataset

ID	Height	Age	Weight
1	5	45	77
2	5.11	26	47
3	5.6	30	55
4	5.9	34	59
5	4.8	40	72
6	5.8	36	60
7	5.3	19	40
8	5.8	28	60
9	5.5	23	45
10	5.6	32	58
11	5.5	38	?

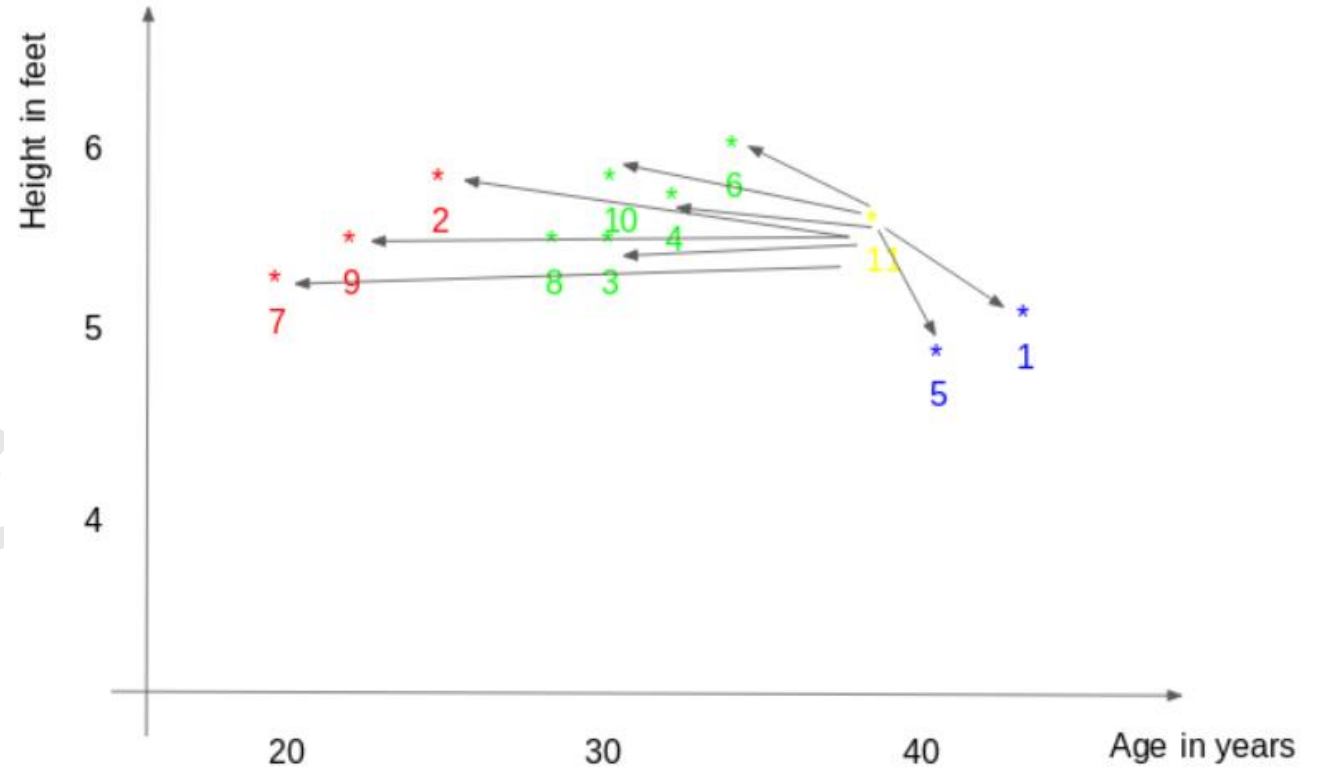
Ris

Analytics Vidya



K-NN(Regression)

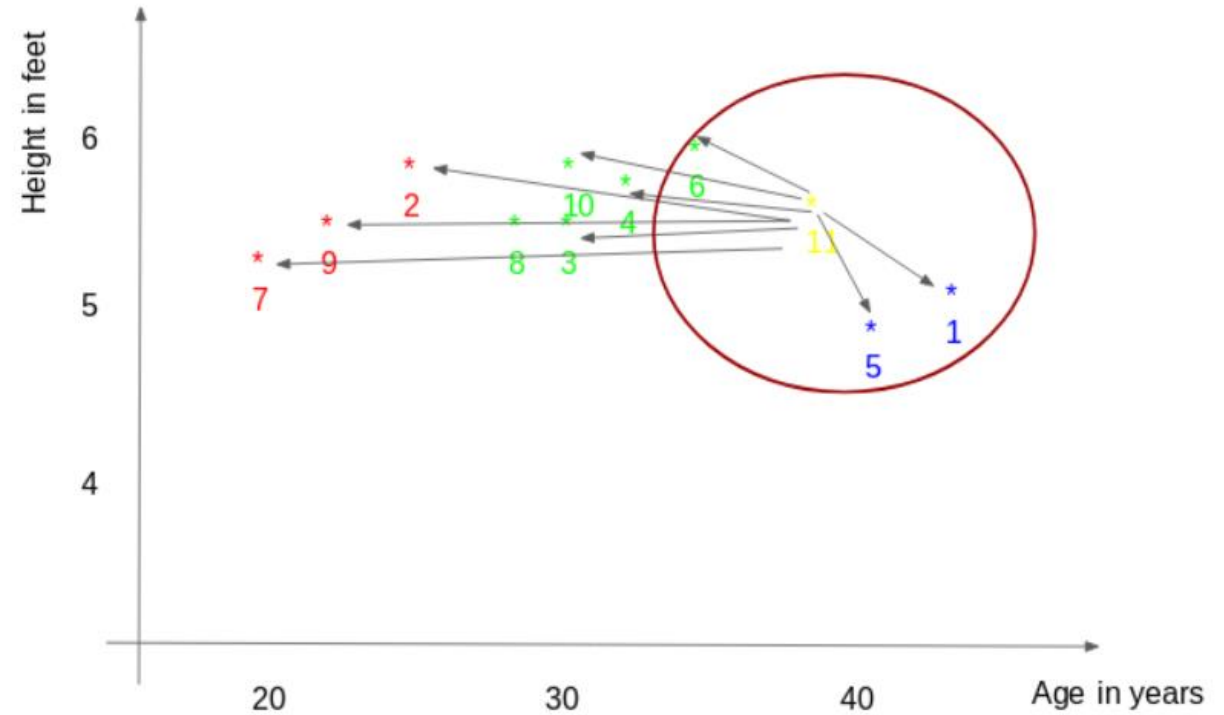
- Take the K nearest neighbours of the new data points by Euclidean distance



Analytics Vidya

K-NN(Regression)

- The average of these points is the prediction for the new point.
- $(77 + 72 + 60)/2 = 69.66$ kg



Analytics Vidya



Scikit Datasets

- https://scikit-learn.org/stable/datasets/toy_dataset.html
- `load_boston(R)`
- `load_iris(C)`
- `Load_diabetes(R)`

Rishi Bansal



Kaggle Dataset

- <https://www.kaggle.com/farhanmd29/50-startups>

Rishi Bansal





Application - KNN

- Calculate profit of a company
- House Price Prediction
- Economic forecasting
- Data compression

Rishi Bansal

