KNN

- Non – parametric supervised learning algorithm.

- Regression ✓
- Classification ✓

- **Algorithm**

Training set

X_1	X_2	X_3	...	X_p	Y
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.

Test point

X_1	X_2	X_3	...	X_p
.	.	.	.	.

1. **Find nearest neighbors:**

   For each test point, find K training data points that are closest in terms of Euclidean distance (or any other suitable metric)

2. **Assign label:**

   Get the labels corresponding to the K nearest neighbors. Call this vector of labels L.

   For classification problem, compute class probabilities by assigning the fraction of points in each class to the corresponding class.

For regression problem, the mean/median is the test label.

- Variation: Weight each nearest neighbor by its closeness and consider this weight while assigning the label to test data point.

- Model flexibility (and variance) decrease as K increases.