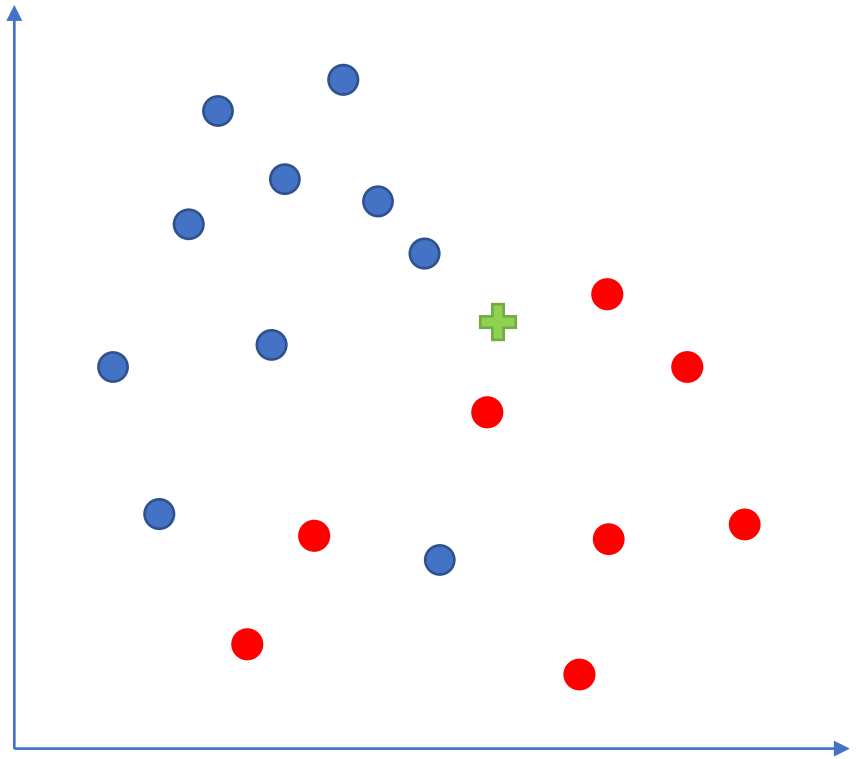
Basic Classification with K-NN

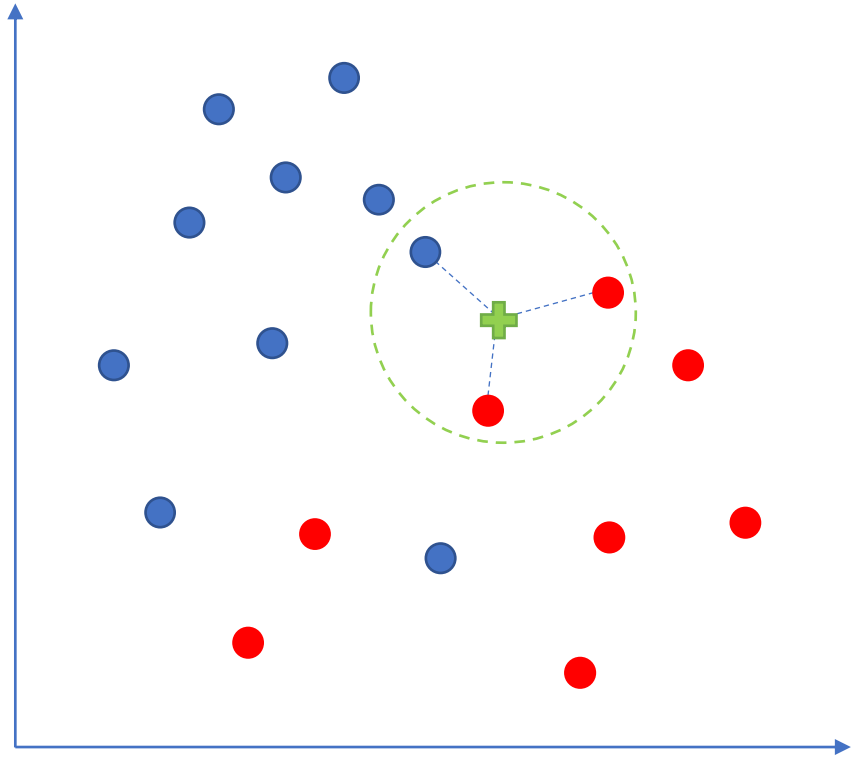
By: Nur Rohman Widiyanto

Institut Teknologi Sepuluh Nopember
IEEE Student Branch

How the K-NN Works



: *is the data to classify*



$k = 3$

Basic K-NN

Steps

- Prepare the Data **@Notebook** #1
  3_1_k-nearest_neighbor_basic.ipynb →
- Calculate distance between two points **@Notebook** #2
  3_1_k-nearest_neighbor_basic.ipynb →

$$D(x, y) = \left(\sum_{i=1}^{n} |x_i - y_i| \right)^{\frac{1}{p}}$$

- *x*, *y*: Data point
- *p*: Minkowski Distance (P = 1 atau P=2)

Steps

- **Get distance** between the **Test Point** and **all known data points @Notebook** #3 3_1_k-nearest_neighbor_basic.ipynb →
- **Sort the distance** to find the Points that closest to the **Test Point @Notebook** #4 3_1_k-nearest_neighbor_basic.ipynb →
- Use **Class Labels** of those closest points to **predict** the label of the **Test Point @Notebook** #5 3_1_k-nearest_neighbor_basic.ipynb →
- See the Classification Result **@Notebook** #6 3_1_k-nearest_neighbor_basic.ipynb →

K-NN from Scratch

Institut Teknologi Sepuluh Nopember
IEEE Student Branch

Steps

- Prepare the Data **@Notebook** #1
  3_2_k-nearest_neighbor_from_scratch.ipynb →
- Calculate distance between two points **@Notebook** #2
  3_2_k-nearest_neighbor_from_scratch.ipynb →

$$D(x, y) = \left(\sum_{i=1}^{n} |x_i - y_i| \right)^{\frac{1}{p}}$$

- *x*, *y*: Data point
- *p*: Minkowski Distance

Steps

- **Get distance** between the **Test Point** and **all known data points @Notebook** #3 3_2_k-nearest_neighbor_from_scratch.ipynb →
  - **The Test Point is Multiple (Test Points)**
  - **Sort the distance** to find the Points that closest to the **Test Point @Notebook** #3 3_2_k-nearest_neighbor_from_scratch.ipynb →
  - Use **Class Labels** of those closest points to **predict** the label of the **Test Point @Notebook** #3 3_2_k-nearest_neighbor_from_scratch.ipynb →
  - See the Multiple Classification Result **@Notebook** #3 3_2_k-nearest_neighbor_from_scratch.ipynb →

Steps

- Calculate the accuracy **@Notebook** #4 3_2_k-nearest_neighbor_from_scratch.ipynb →
- Calculate the accuracy in every *k* **@Notebook** #5 3_2_k-nearest_neighbor_from_scratch.ipynb →

K-NN with Scikit-learn

Institut Teknologi Sepuluh Nopember
IEEE Student Branch

Steps

- Prepare the Data **@Notebook** #1 3_3_k-nearest_neighbor_from_scratch.ipynb →
- **Calculate the distance** between the **Test Point** and **all known data points @Notebook** #2 3_3_k-nearest_neighbor_with_scikit.ipynb →
- Calculate the accuracy **@Notebook** #4 3_3_k-nearest_neighbor_from_scratch.ipynb →
- Calculate the accuracy in every *k* **@Notebook** #5 3_3_k-nearest_neighbor_from_scratch.ipynb →