K Nearest Neighbors

Wednesday, September 1, 2021 6:25 PM

- Classification setting
  - $Y \to$ discrete, with cardinality $< \infty$, K
    $\{c_1, c_2 \dots c_K\}$
  - Error function
    $$\frac{1}{n}\sum_{i=1}^{n} I(y_i \neq \hat{y}_i)$$
    * Training & test are different!
    $$\therefore \text{Avg}\,(I(y_0 \neq \hat{y}_0))$$
  - Bayes classifier:
    - Given a $x_0 \to y_0$?
      $\hookrightarrow$ Max $P(Y = j \mid X = x_0)$

      conditional probability

    * Joint Dist: $P(Y, X) \to$ Conditional Dist $P(Y \mid X)$
  - Error Rate:
    $$1 - \max_j P(Y = j \mid X = x_0)$$
    Bayes Error Rate

* MAP: Maximum A Posteriori

- K-Nearest Neighbors
  - Non-parametric classifier

$$P(Y=j \mid X=x_0) = \frac{1}{K}\sum_{i \in N_0} I(y_i = j)$$

K closest points to x_0 in the training data (sample)

$K=4$
majority
max P

***Decision Boundary**

* Bias-Variance
↳ Regression:

$$\hat{f}(x_0) = \frac{1}{K}\sum_{x_i \in N_0} y_i$$

Is $f(X)$ linear?
- Yes: Linear Regression!
- No: Non-Parametric Model! → KNN Regression

What value of K?

- Small k - Most flexible $\Rightarrow$ High variance!
- Large k - Less flexible $\Rightarrow$ High bias!

*** Curse of Dimensionality**

- $x_1 \ldots x_p \quad p >> n$