Q5 Name a parametric method and non-para method with an explanation on the major diff between them.

(and form)

- Parametric: Models with fixed finite set of parameters (μ and σ^2 of Gaussian)
- Non-Para: Parameters complexity of the models grows with the data
- Parametric Method: MLE, MAP, Non-Para Method: K-nn, Histograms
- MLE setup: N independent samples $\{x_1, \ldots, x_n\}$ from distribution P_X, we wish to estimate $P_{X|\Theta}(x|\theta)$, θ fixed parameters of the model chosen.
  Parameters as random variables ← we estimate. Assume Gaus distribution with unknown μ and σ^2 ($\theta = \begin{bmatrix}\mu \\ \sigma^2\end{bmatrix}$). we want to fit Gauss to data by estimating Θ.
- MLE is: $\hat{\theta}_{MLE} = \underset{\theta \in \Theta}{\operatorname{argmax}} \prod_{n=1}^{N} P_{X|\theta}(x_n|\theta)$
- K-nn: Prediction by majority vote (classification) or average (regression) of k closes training examples
  - fix number of and very interval: $\hat{P}_X(x) = \frac{k}{2Nd_k(x)}$ where $d_k(x)$ is the distance to k_{th} nearest point x_n
  - The bin around x is implicitly chosen so that it contains k neighbor
- Bayesian Classification for k-nn: $\hat{P}_{C|X}(c|x) = \frac{k_c \text{ (counts of neighbor from class)}}{k \text{ (number of neighbor)}}$
- Test point x which we want to classify select class c with the highest count among the k nearest neighbor to x.