Session-10 (KNN Theory)

c11eu_ML10(KNN)
Training Clarusway
Pear Deck - August 25, 2022 at 7:28PM

Part 1 - Summary

Use this space to summarize your thoughts on the lesson

Part 2 - Responses

Slide 1

Use this space to take notes:

Slide 2

SUMMARY of PREVIOUS CLASS

- Logistic Regression
- Classification: Hearing Test, Diabetes
- Classification Error Metrics
  - Model Predictions Results: TP, TN, FP, FN
  - Confusion Matrix
  - Classification Report: Accuracy, Recall, Precision, F1
  - ROC/AUC (Binary Classification-Detection)

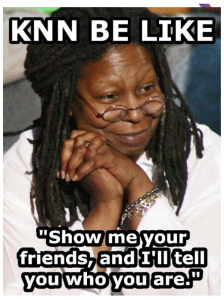
2

Use this space to take notes:

Slide 3

K Nearest Neighbors Theory

K Nearest Neighbors with Python

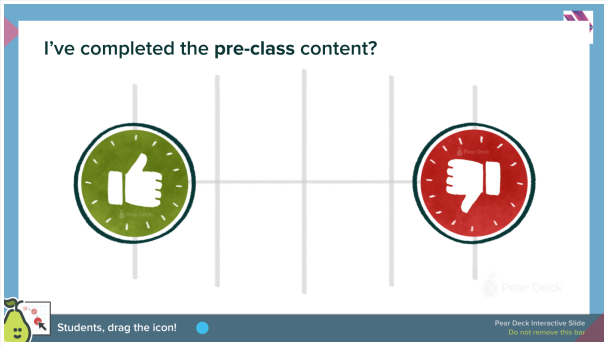


CLARUSWAY©

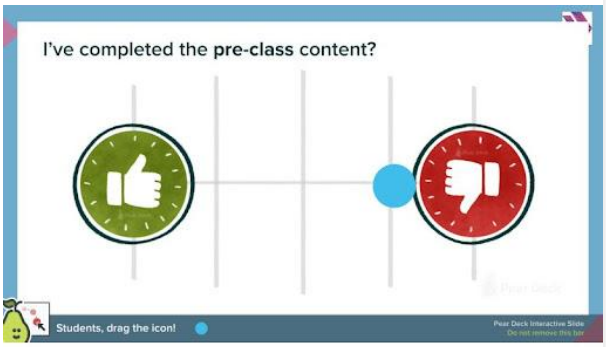
Use this space to take notes:

Slide 4	Your Response

Slide 4

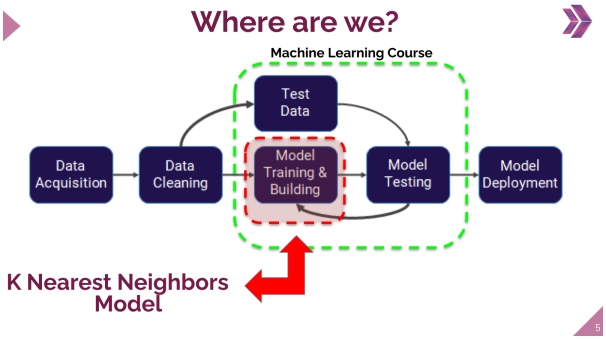


Your Response



Use this space to take notes:

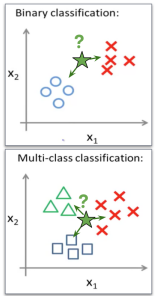
Slide 5



Use this space to take notes:

Slide 6

K Nearest Neighbors (KNN) Theory



k-Nearest Neighbors (KNN) is a simple, supervised **Classification algorithm.**

Some characteristics and usage areas of KNN classification:

- A lazy learner (no training),
- Non-linear and non-parametric
- Low dimensional datasets,
- Fault detection,
- Recommender systems etc.

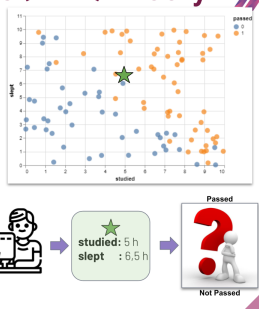
6

Use this space to take notes:

Slide 7

K Nearest Neighbors (KNN) Theory

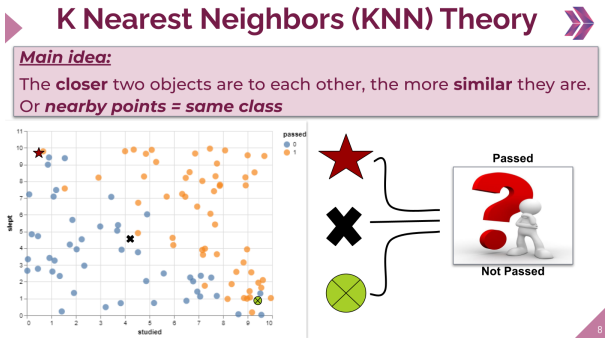
When to make a prediction for a new data point (★), it **finds the closest neighbor or neighbors** of the new data in the training set and makes a *classification according to the class of those neighbors.*



7

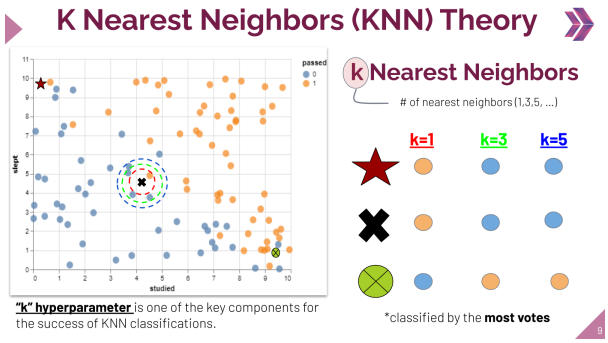
Use this space to take notes:

Slide 8



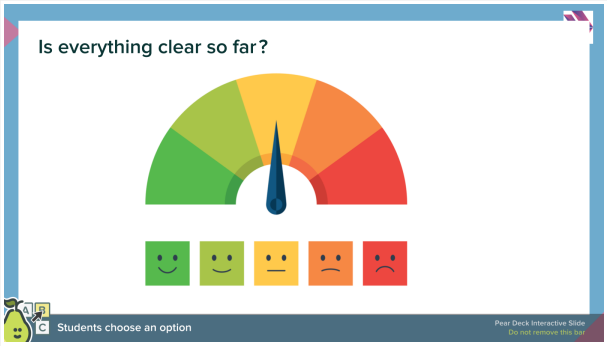
Use this space to take notes:

Slide 9

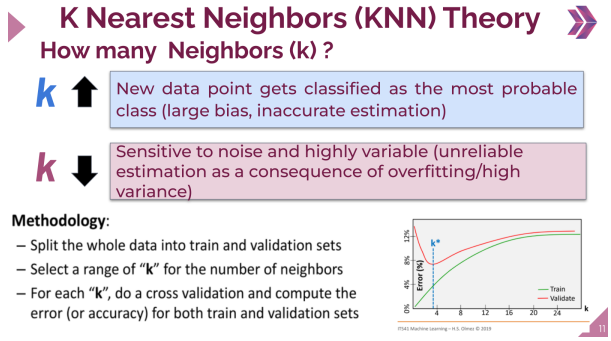


Use this space to take notes:

Slide 10	Your Response
	You Chose • **clear**

Slide 10	Your Response
	Other Choices • so so • confused

Use this space to take notes:

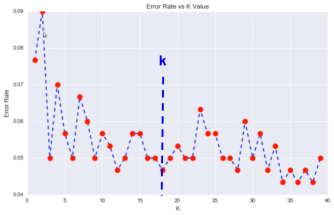
Slide 11


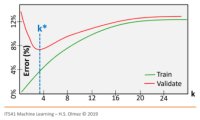
Use this space to take notes:

Slide 12

K Nearest Neighbors (KNN) Theory

Elbow for optimum "k"





12

Use this space to take notes:

Slide 13

K Nearest Neighbors (KNN) Theory

Distance between Neighbors ?

"weights" *parameter*? (default = "uniform")

We find nearest neighbors using the **distance function**, neighbors vote to predict class:

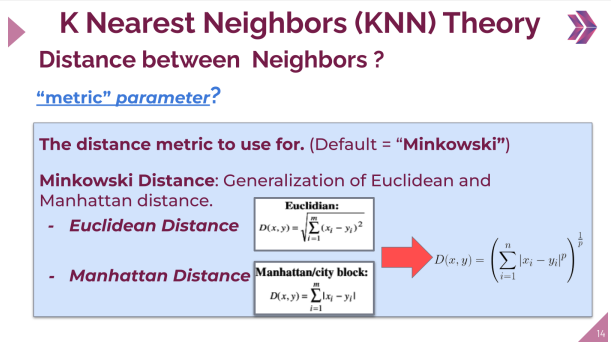
These are the **"weights" *parameter*** of KNN:

- **Uniform**: Majority voting
  (All points in each neighborhood are weighted equally.)
- **Distance:** Weighted majority voting
  (closer neighbors of a query point will have a greater influence than neighbors which are further away.)

13

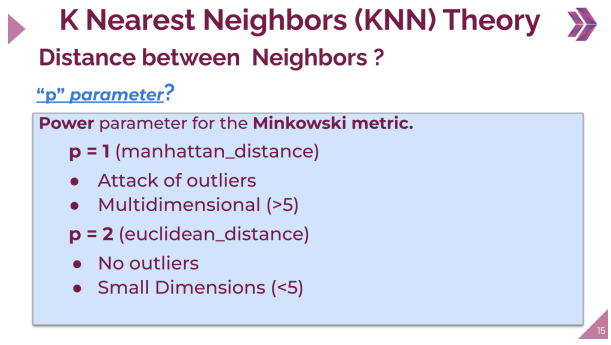
Use this space to take notes:

Slide 14



Use this space to take notes:

Slide 15

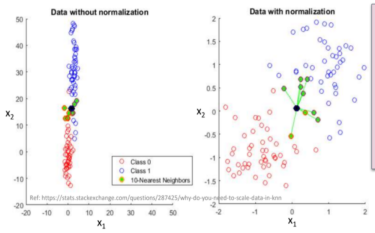


Use this space to take notes:

Slide 16

K Nearest Neighbors (KNN) Theory

Need Data Scaling ?



Because of the KNN algorithm **relies on distance** for classification, **normalizing** the training data can improve its accuracy dramatically.

16

Use this space to take notes:

Slide 17

K Nearest Neighbors (KNN) Theory

Pros & Cons

Pros:

- No assumptions about the data
- Simple to understand and easy to implement
- Need no training
- Flexible to distance selections
- Can be used both for Classification and Regression

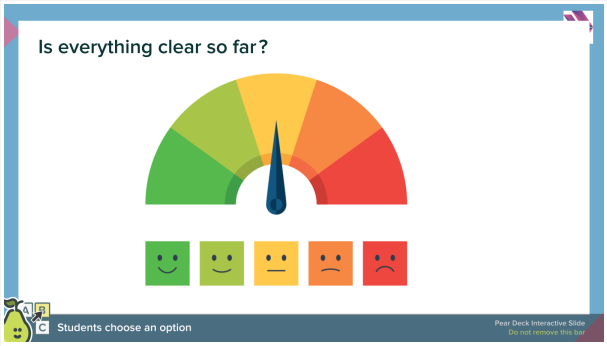
Cons:

- Not work well with high dimensional datasets
- Sensitive to outliers and imbalanced data (distances are affected)
- May overfit (low bias/high variance, choice of k is crucial)
- Needs scaling

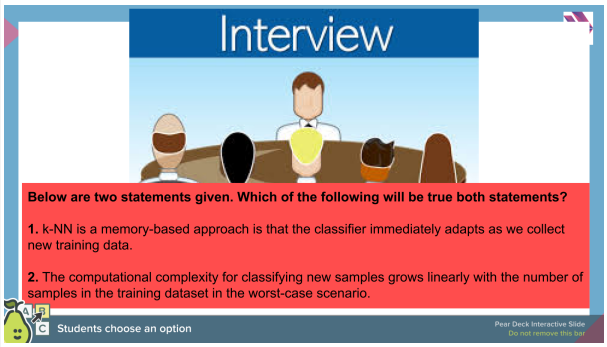
17

Use this space to take notes:

Slide 18	Your Response
	You Chose • **clear**

Slide 18	Your Response
	Other Choices • so so • confused

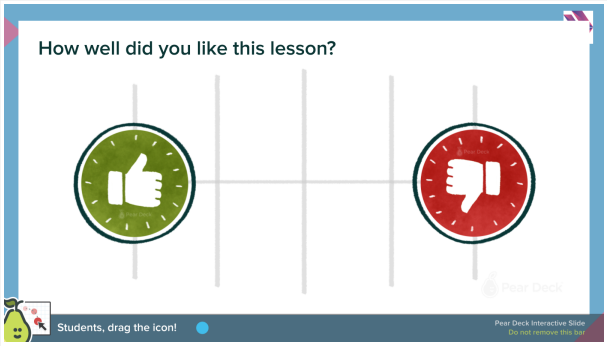
Use this space to take notes:

Slide 19	Your Response
	

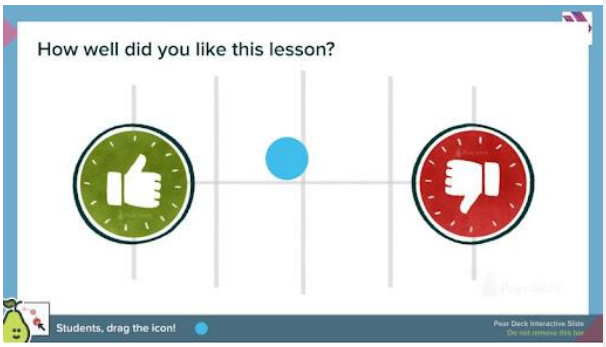
Use this space to take notes:

Slide 20	Your Response

Slide 20



Your Response



Use this space to take notes:

Slide 21

Use this space to take notes:

Slide 22

Use this space to take notes: