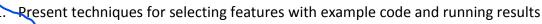
INTELLIGENT SYSTEM DEVELOPMENT

EXERCISE 3

Request

- Students write the answers and results of code & running program in file doc
- Due date: 1 week

SELECT FEATURES



- https://towardsdatascience.com/feature-selection-techniques-in-machine-learning-with-python-f24e7da3f36e
- b. https://www.kdnuggets.com/2021/06/feature-selection-overview.html
- c. https://www.geeksforgeeks.org/feature-selection-techniques-in-machine-learning/
- d. https://www.kaggle.com/code/piyushagni5/feature-selection-techniques-in-machine-learning
- e. https://www.analyticsvidhya.com/blog/2020/10/feature-selection-techniques-in-machine-learning/
- f. https://www.javatpoint.com/feature-selection-techniques-in-machine-learning
- g. https://trainindata.medium.com/feature-selection-for-machine-learning-a-comprehensive-overview-bd571db5dd2d
- Student use data from link of classifying prices of mobile phone and execute 3 techniques you present in Question 1 (refer to link 1.a)
- https://www.kaggle.com/datasets/iabhishekofficial/mobile-price-classification Present 3 techniques with data from Chap 12

CLEANING DATA

- 1. Cleaning data: steps with python. Present steps and examples
 - a. https://www.analyticsvidhya.com/blog/2021/06/how-to-clean-data-in-python-for-machine-learning/
 - using data: https://www.kaggle.com/c/sberbank-russian-housing-market/data .
 - b. https://www.tutorialspoint.com/python data science/python data cleansing.htm
- 2. Demo steps of cleaning the following data. Record code and running results on file doc
 - a. https://www.w3schools.com/python/pandas/pandas_cleaning.asp
 - b. https://realpython.com/python-data-cleaning-numpy-pandas/

Here are the datasets that we will be using in **b**:

- <u>BL-Flickr-Images-Book.csv</u> A CSV file containing information about books from the British Library
- <u>university_towns.txt</u> A text file containing names of college towns in every US state
- <u>olympics.csv</u> A CSV file summarizing the participation of all countries in the Summer and Winter Olympics

You can download the datasets from Real Python's GitHub repository