Wrocław University of Science and Technology



Faculty of Electronics, Photonics and Microsystems PYTHON LABORATORY

Theme of class: Data Analysis and Machine Learning

Student: Hayrettin Aycetin (276807)

Date of class: 18.12.2023 15:15-16:55

Group No:3

Submission Date:07.01.2024

Lab assistant: Aleksander Kubeczek, Alicja Kwaśny GRADE:

Task 1

A close up of text

Description automatically generated Code:

A screenshot of a computer program

Description automatically generated

A screenshot of a computer

Description automatically generatedOutput:

Comments:

Firstly I imported all needed library for this list .Using read\_csv() function I read the dataset and assigned as df.I used head() function to read the first five entries of the dataframe.

This part was easy and straightforward.

Task 2

A close-up of a white background

Description automatically generatedCodes: Outputs:

A screenshot of a computer

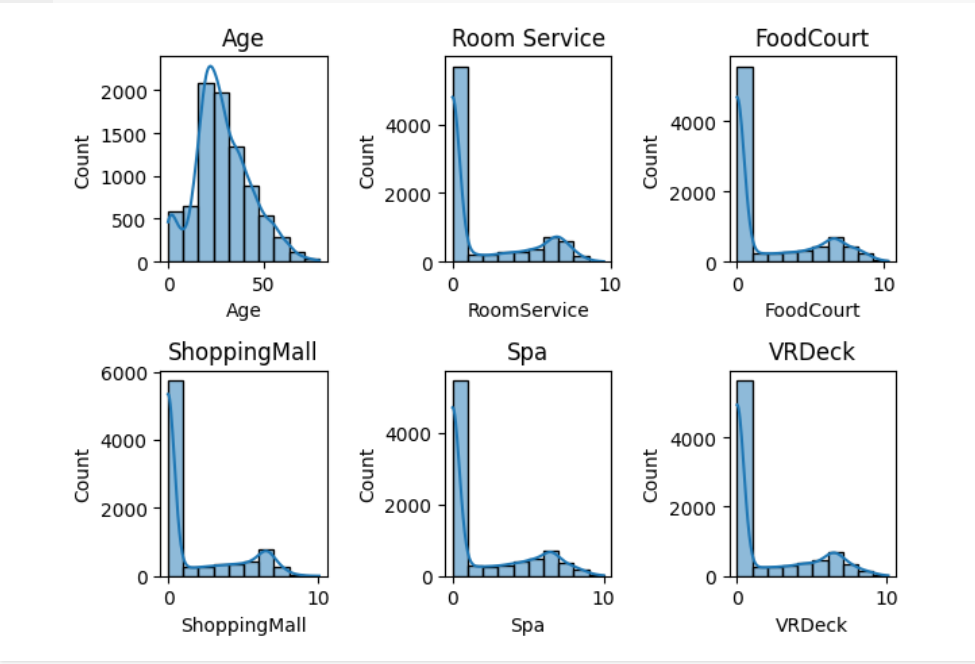
Description automatically generated

A close-up of a computer screen

Description automatically generated

A blue and white bar graph

Description automatically generated

A screenshot of a computer program

Description automatically generatedA close-up of a computer screen

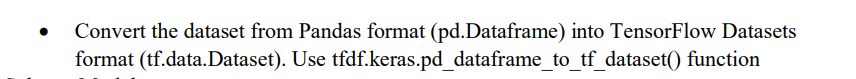
Description automatically generated

Comments:

Firstly I used describe() and info() functions to explore the dataset.

Later I made a bar graph from transported people’s value counts which is boolean.

For the last part I used subplots() function in pandas and set my subplots like an 2 rows 3 columns array.By using histplot() function and its parameters such as data,bins,kde and ax from seaborn library I made my plots.Bins controls the granularity of the histogram.Kde provides a smooth estimate of the distribution as a line on it and ax determines where the plots will be on the subplot array. The most important part was np.log1p(column): Applied the natural logarithm plus one (log(1 + x)) to the values in the column. This transformation is used for dealing with data that includes zero values, as it handles them more gracefully.

A screenshot of a computer code

Description automatically generatedTask 3

Codes: Outputs:

A screenshot of a computer

Description automatically generatedA close-up of a computer screen

Description automatically generated

A screenshot of a computer

Description automatically generatedA screenshot of a computer code

Description automatically generated

A black and blue text

Description automatically generated

A screenshot of a service

Description automatically generated

A computer code with many black text

Description automatically generated with medium confidence

No output

A screenshot of a computer code

Description automatically generated

No output

Comments:

This part was mainly about preparing and cleaning the dataset for tensorflow.Firstly I dropped name and passenger\_id colmns using drop() function and later checked missing values using given function.Later I replaced boolean values with their integer values such as True:1 and False:0 using replace() function.Then using fillna() function I filled empty values with zero.Later I splitted Cabin column using split() function to wanted columns and then dropped it.Lastly I split dataset to training and testing with given split\_dataset() function and using tfdf.keras.pd\_dataframe\_to\_tf\_dataset() function I converted pandas dataframes of testing and training to tensorflow dataframe and I used label = “Transported ” because that’s where we need to predict.

A close up of a text

Description automatically generatedTask 4

Code:

A black text on a white background

Description automatically generated

Output:



Comments:

In this part I just choose a model which is RandomForestModel as wanted in list and I assigned it to model.

Task 5

A black and blue text

Description automatically generated

A close-up of a website

Description automatically generated

Code:

A black text with red and blue text

Description automatically generated

A black and white text

Description automatically generated

A black and blue text

Description automatically generated

Outputs:

A close-up of a white background

Description automatically generated

A screenshot of a computer program

Description automatically generated

A screenshot of a computer program

Description automatically generated

Comments:

Firstly I created a random forest model in task 4 so I just continued here with the including a list of eval metrics given in the task and I used model.fit() to train the model.Later I used model.evaluate() to return a tuple of evaluation results.Lastly I used the given plotting to plot my model.

Task 6

A screenshot of a computer code

Description automatically generated

Code:

A black and green text

Description automatically generated

A screen shot of a computer code

Description automatically generated

Outputs:

A graph with a line

Description automatically generated

A white background with black text

Description automatically generated

Comments:

In this part mostly the whole code was given by instructor in the list I just changed rf to model.This part was a bit hard to understand but as long as I understood the first code snippet inspects and evaluates a machine learning model by using an inspector to check the model's evaluation and variable importances. The evaluation results, including various metrics, are printed, and the importance of a specific variable ("NUM\_AS\_ROOT") is retrieved.

Link of my Collab:

<https://colab.research.google.com/drive/13CCqkFQ5MkC5FO6O5WzAjdC3kwja8sLz?usp=sharing>

Conclusion:

In general it was an really interesting and fun lab.Also it was really helpful about data analysis and machine learning basics.