



Cosm Medical - Personal Pelvic Health

Interview Follow-up Test

Please read carefully and use Python to write scripts for the following questions:

- 1- Load iris data in Scikit-learn package and save into pandas
- 2- Describe the data statistically using panda
- 3- Write a function to detect and handle the missing, noise and outliers in the data
- 4- Write a function to study the statistical relationship between the features and between the features and the target.
 - a. Function takes data pandas as input and,
 - b. Return csv file with relationship statistics
 - c. Use Seaborn package to visualize the relationships (using your preferred method)
- 5- Use the Scikit-learn package to predict the target as follows:
 - a. Normalize the data with the appropriate normalizer (if needed)
 - b. Split the data into training and testing
 - c. Select two appropriate machine learning techniques to train the data
 - d. Test the test data and evaluate the results using the appropriate metrics
 - e. Save the results in csv file with appropriate headers
 - f. Visualize the results

Please note:

- 1- You will need to do some research and self-learning to understand to answer some of the questions.
- 2- You are not expected to finish all the test, however, self-learning, attempts, and research will be much appreciated.

- 3- There are decisions to be made based on the nature of data, you have to figure that out yourself as this is part of the test
- 4- Please reply with a py file for the code and a document describing your work
- 5- Return by: Monday, July 6, 2020 at 9:00 am Toronto time