



COMSATS University Islamabad, Lahore Campus

Assignment 1 - Data Understanding and Pre-processing using Python and Pandas

Course Name:	Machine Learning	Course Code:	CSC354	Credit Hours:	3(3,0)
Course Instructor/s:	Dr. Rao Muhammad Adeel Nawab	Program Name	BS (CS)		
Semester :		Batch :	Spring 20212		
Submission Deadline:	Sunday, March 20, 2022 at 23:59	Maximum Marks:	20		

Instructions:

- اللہ آپ کو دیکھ رہا ہے
- اگر آپ نقل کر کے پاس ہو گئے تو اللہ تعالیٰ کے ہاں فیصل ہو جائیں گے (اور اللہ سے محروم ہو جائیں گے)
- اگر آپ نے نقل نہ کی اور فیصل ہو گئے تو اللہ تعالیٰ کے ہاں پاس ہو جائیں گے (اور اللہ کو پا جائیں گے)

Assignment - Aims

- Recall the Equation
 - Life = Technical Skills (15%) + Human Engineering (85%)
- Technical Skills
 - The main aim of this assignment is to understand the basics of Python Programming Language and Pandas Library, so that you should become familiar with the basic techniques needed to read, understand and pre-process different types of data for building Machine Learning Models

- To introduce how we can quickly and easily learn any complex Concept using the best learning methodology i.e.,
 - DO IT YOURSELF
- Human Engineering
 - To introduce how to achieve excellence in Human Engineering, by practicing three main pillars of Human Engineering
 1. Truth
 2. Honesty
 3. Justice
 - To give confidence to students that

. اگر دنیا میں کسی نے کوئی کام کیا ہے تو آپ بھی کر سکتے ہیں
 - To introduce how we can develop the habit to
 - Stop Complaining! Stop Criticizing! Start Contributing from Today 😊

Assignment - Learning Outcomes

- After reading, understanding, analyzing and doing this Assignment, In Sha Allah, you will learn:
 - Technical Skills
 - Main Concepts related to Python Programming Language and Pandas Library to read, understand and pre-process different types of data for building various useful Machine Learning Models
 - How to quickly and easily learn any concept using the best learning methodology i.e.,
 - DO IT YOURSELF
 - Human Engineering
 - Never ever compromise on three main pillars of Human Engineering
 - Truth
 - Honesty
 - Justice

- **Learn the art to become an authority in any Field of Life by learning how to systematically perform any Real-world Task using a Template-based Approach**

Assignment - Pre-requisite

- **To completely and correctly understand, and complete this Assignment before Deadline, you will need**
 - **Purity in Intention**
 - **To Put 100% Effort with Sincerity on Daily Basis**
 - **To Email Your Daily Work Progress with Honesty 😊**
 - **Python 3.7.4.final.0**
 - **Anaconda 1.9.12**
 - **Pandas 0.25.1**

Main Tasks / Sub-tasks

Note that this Assignment is planned, designed, and developed using a Half-Cooked Approach i.e., some work is done by the instructor and the remaining will be done by you. Since the Assignment is to read, understand and pre-process different types of data using Python and Pandas. Following the Half-Cooked Approach, you are provided with an initial tutorial on Python and Pandas (from Instructor). Your job is to carry out the Tasks given below and answer the associated questions.

Tasks

Your main job is to improve the Overall Quality of the Initial Tutorial by performing the following Tasks.

Important Note: You should do this Assignment using Jupyter Notebook.

- **Task 1**
 - **Task 1.1**
 - **For each Concept in the Initial Tutorial (for e.g., Variables, String, List, Tuple, Dictionary etc.), write down**
 - **Definition**
 - **Purpose**
 - **Importance**
 - **Applications**
 - **Strengths**
 - **Weaknesses**
 - **Suitable to Use**
 - **Given and Task**
 - **Input and Output**
 - **At least 3 – 5 Examples (Not Given in the Initial Tutorial)**
 - **At least 1 Example to Solve Real-world Problems**
 - **Note that you can add Concepts in your Assignment Solution that are not mentioned in the Initial Tutorial 😊**
 - **Task 1.2**
 - **At the end of each Concept, add**
 - **TODO Tasks**
 - **Your Turn Tasks**
- **Task 2**
 - **Task 2.1**
 - **Improve the Initial Tutorial, so that it reflects the phenomenon of**
 - **Completeness and Correctness**
 - **Task 2.2**
 - **Improve the Initial Tutorial to make it self-explanatory**

- **Task 3**
 - **Task 3.1**
 - **Improve Formatting of the Initial Tutorial to Improve the Look and Feel of the Tutorial**
- **Very Important Note**
 - **Recall the Equation**
 - **Life = Technical Skills (15%) + Human Engineering (85%)**
 - **Please Note that whenever you describe any Concept, it should be Teaching both Technical Skills and Human Engineering**
 - **Consider the following Two Examples to Print a Message (String) on the Output Screen.**
 - **Examples**
 - **Example 1**
 - **print("Hello World!")**
 - **Example 2**
 - **print("Allah is One. Allah created us and He has control over everything:")**
 - **Analysis**
 - **Analysis – Example 1**
 - **Only Teaching us Technical Skills**
 - **Analysis – Example 2**
 - **Teaching us Technical Skills + Human Engineering**
 - **Conclusion**
 - **We can easily Learn and Teach both Technical Skills + Human Engineering through our Assignment Solution**

Supporting Material for Each Task / Sub-task

- **Supporting Material**
 - **Lecture to Install and Setup Python and Anaconda Framework**
 - **To start working on this Assignment, you must have first installed Python and setup a suitable IDE (Integrated Development Environment)**

- The following link contains detailed Lectures on installation of Python and Anaconda Framework (i.e., IDE)
 - <https://ilmoirfan.com/introduction-to-python/>
- Initial Tutorial
 - For Initial Tutorial, see Python and Pandas - A Tutorial.ipynb File

Assignment Marks

- The following Table shows Marks for each Task / Sub-task

Task	Marks
Task 01 – Task 1.1	50
Task 01 – Task 1.2	20
Task 02 – Task 2.1	10
Task 02 – Task 2.2	10
Task 03 – Task 3.1	10
Total Marks	100

Important Note

- In CSC354 – Machine Learning Course, the weightage of this Assignment is of 20 Marks

Guidelines for Assignment Submission

Upload your Tutorial (Assignment Solution) on Google classroom in PDF Format (along with Jupyter Notebook File and any Dataset(s) that you have used in your Tutorial).

- Guidelines for Assignment Submission
 - The Assignment 1 is due by Sunday, March 20, 2022, at 23:59.

- **Always Remember! Little Efforts Daily Will Make You The Greatest 😊**
 - If you spend 2 – 3 Hours Daily (5 – 6 Days a Week), In Sha Allah, You will be able to complete this Assignment before the Deadline
- **Tip**
 - **To quickly and easily complete this Assignment on Time**
 - **First Focus on Accuracy, Then Speed 😊**
- **Penalty for Late Submission:**
 - **Note that there will be deduction of 5% Marks (per day) on late submission. Note that if you submit the Assignment Three Days after the Deadline it will be considered un-submitted and you will be given 0 Marks.**

Policy for Plagiarism

Institutional Policy will be followed for punishment if a student is found guilty of plagiarism.