

In the Name of God



University of Tehran
College of Engineering
School of Electrical and Computer Engineering

Introduction to Data Science

General Guidelines

Spring 1403

Welcome to the 'Introduction to Data Science' course! In this document, we'll give you some instructions for the kick-off in the course, and also some suggestions, which can help you throughout your assignment.

Technical Instructions

Programming Language

The most common programming language used for data science is Python, although some may prefer other languages like R. Having a dynamic community and supporting rich packages, we'll use **Python 3** as the programming language. Other versions like **Python 2** may not work well with some packages needed for this course.

Coding Style, Environment and Packages

One of the important steps in a data science project, if not the most important one 😊, is the report. Therefore, we suggest you use [Jupyter Notebook/JupyterLab](#), so that you could markdown your report alongside your code.

You might think that coding style isn't important in data science. However, in both industry and academia, many successful notebooks are needed to be both reproducible and easy to read. We suggest you take a look at this [blog](#) and try your best to use it in your assignments.

We need to install many packages during this course, including Jupyter Notebook/Lab as your very first package 😊. It's better to use an isolated

python environment for this purpose, so that your primary python doesn't get involved with problems like version mismatch in packages. We suggest you use [virtualenv](#), which is an easy-to-use tool that creates isolated Python environments for Python libraries.

Report Instructions

As mentioned before, one the most important steps in a data science project is the report. You are going to be graded upon your analysis of the results and your explanation. Below, we'll demonstrate some quick tips, which can help you write better reports for your assignments.

You should show:

- The purpose of the assignment, e.g. web scraping
- A brief summary of what you're going to do
- An explanation for each dataset that is loaded in the notebook, why you're using it, is it for training process or test, etc
- An explanation for each step of the assignment
- An explanation for the result of each step, the quality of them and probable reasons for them
- The reason why you're using an algorithm, a measurement or any other thing among the other possible choices
- Your suggestions for the outcome of the assignment
- Your conclusion
- The references you used

Try to write as concisely as possible, but don't forget to add all the necessary information in your report. Hope you enjoy the course!