Homework

In preparation For The Course

Below is a list of free online courses (if you subscribe as an auditor) and videos on Jupyter Notebook, Pycharm or Visual Studio Code IDE (one of the two is sufficient), basic Python object-oriented principles and Git .

Please note that Jupyter Notebook and IDE (Visual Studio Code or Pycharm) are both required for completing the final project. If you don't have a preference for an IDE my recommendation is Visual Studio Code since it has more features for data science. Jupyter Notebook and Pycharm can be installed through Anaconda by following this video. You can install Visual Studio Code by watching this video.

Regarding basic Python object-oriented principles since I will be teaching these only briefly in class it will be helpful for your understanding if you watch Module 5 of this course at least up to Constructors and Other Special Methods. Familiarity with basic object-oriented principles will enable you to understand the structure of the Classes in the Scikit-Learn library which we will be using all throughout the course.

Regarding Git it is not compulsory but only for enabling easier code writing in a group and the final project can be a good opportunity for you to start using it if you aren't already familiar with it. You can choose to watch this YouTube video Git and GitHub for Beginners - Crash Course, approximately 1 hour long or complete the object-oriented section in week 3 of this course by Meta.

Class 1

- Read Chapter 1 of the book Hands-On Machine Learning with Scikit-Learn, Keras and Tensorflow, O'Reilly Media, Inc. Aurelien Geron.
- Understanding the code in the Jupyter Notebook 01_the_machine_learning_landscape.ipynb in the Hands-On-ML-Code folder within the Python Code folder on Moodle.
- Complete the assignment at the end of chapter 1 in the book Hands-On Machine Learning with Scikit-Learn, Keras and Tensorflow, O'Reilly Media, Inc. Aurelien Geron. The assignment is comprised of 19 questions to test your understanding and the solutions are in the book Appendix.
- Read through the three EDAs (exploratory data analysis) Jupyter Notebooks/blog post which you can find in the EDA Tasks folder on Moodle. As I mentioned in class it's important you become familiar with the methodology in these EDAs so that you can implement similar analysis in your project.
    - 1st EDA: eda-cardio-vascular-disease.ipynb (Jupyter Notebook on Moodle)
    - 2nd EDA: eda-and-multi-class-classification.ipynb (Jupyter Notebook on Moodle)
    - 3rd EDA: How to perform Expletory Data Analysis with Seaborn (link to blog post)

Class 2

- Read Chapter 4 of the book Hands-On Machine Learning with Scikit-Learn, Keras and Tensorflow, O'Reilly Media, Inc. Aurelien Geron. From The bias variance Tradeoff up to Ridge Regression.
- Understanding the code in the Jupyter Notebook 04_training_linear_models.ipynb in the Hands-On-ML-Code folder within the Python Code folder on Moodle.