SIT719 Security and Privacy Issues in Analytics

Pass Task 3.1: Collection of scikit-learn commands for machine learning

Overview

The next-generation cyber-security defense mechanisms are based on advanced analytics like machine learning and artificial intelligence. Machine learning can help to recognize the attack and misbehaviors from any dataset. *scikit-learn* is a widely used machine learning library which has a large collection of supervised and unsupervised machine learning algorithms implemented.

During this week, you have been demonstrated how scikit-learn can be used for machine learning using a 'payment fraud dataset.' If you are interested to learn more about scikit-learn, please see the below tutorial.
https://scikit-learn.org/stable/tutorial/basic/tutorial.html

The scikit-learn user guide can be obtained from the link below:
https://scikit-learn.org/stable/user_guide.html

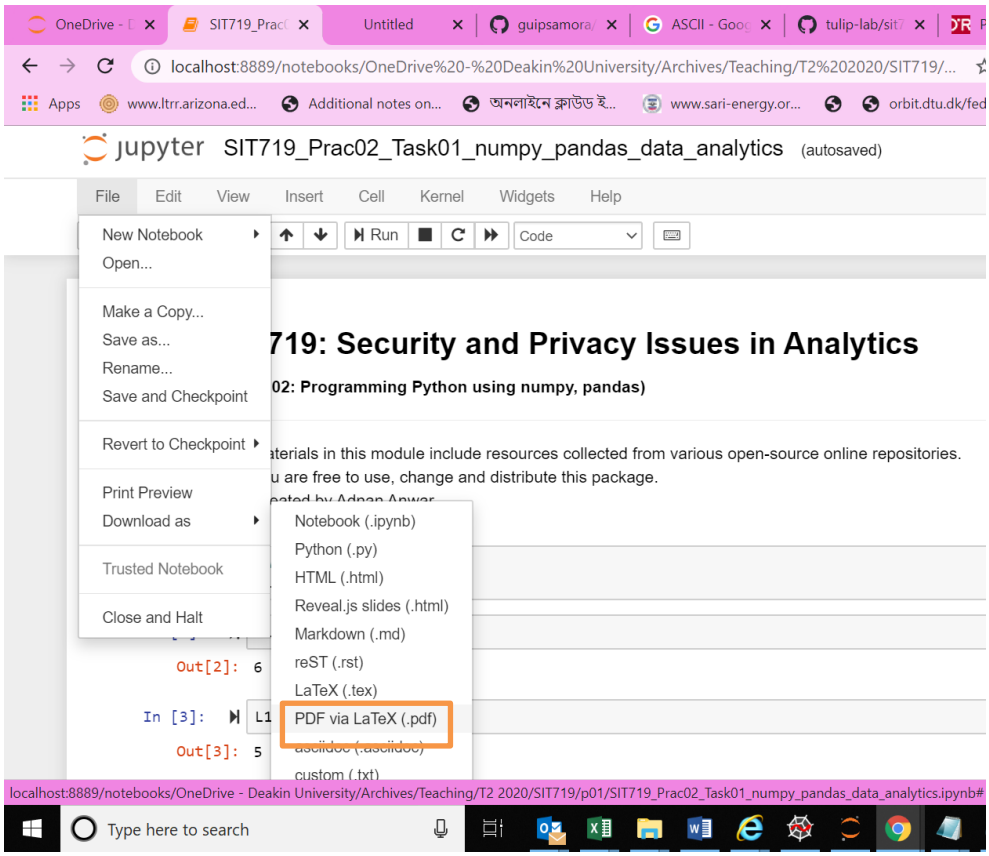
This is a pass task, so you must ensure that you attempt this task to pass the unit.

Task Description

Instructions: In this task, you need to write a collection of scikit-learn commands for machine learning in a notebook file. For each command, add a simple explanation (e.g., the purpose or what it does) and then example(s). There is no limit on how many you need to write. However, you must try to choose a good selection of functions/commands so that someone who has no prior knowledge can have a good understanding after reading your notebook with scikit learn commands/examples. Run the notebook and then save the notebook file as a PDF. You can create the notebook either in jupyter or using google colab, up to you. (I have added two screenshots regarding how to save in a PDF- one for Jupyter notebook, and the other from google colab). You can also follow any other steps to save/print the notebook as a PDF. **YOU JUST NEED TO SUBMIT THE FINAL PDF** of the notebook into the OnTrack system.

Please note, it is a graded task and there is no right or wrong answer. Your tutor/marker will assign you some marks based on the quality of the selection of the commands, how well you have explained the commands with examples and finally the aesthetic look of your submission and reflection of the quality of your work from the tutor's judgement.

Jupyter: (File >> Download As >> PDF via Latex) Alternative: you can also print as pdf.



Google Colab: (File >> Print >> Save as PDF)