# ITC 6230-A1 Deep Learning

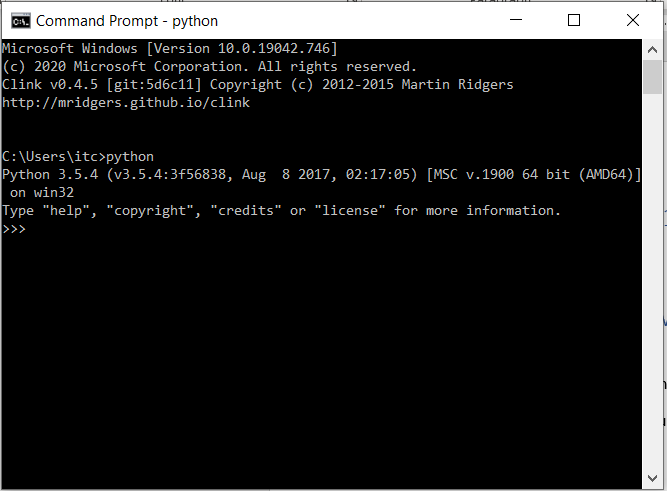
## Winter 2024

## 1st LAB AND HOMEWORK ASSIGNMENT: Working with RIPPERk in Python

Due: Jan. 27, 2024, 23:59

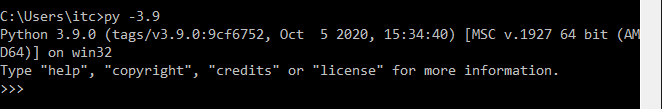
### 1. Installing Python on your machine (mandatory step)

Open a command prompt on your machine, and type “python” to see if you have python installed.



If you don’t get a “>>>” prompt from the Python interpreter, then either you don’t have Python installed or it’s not in your PATH.

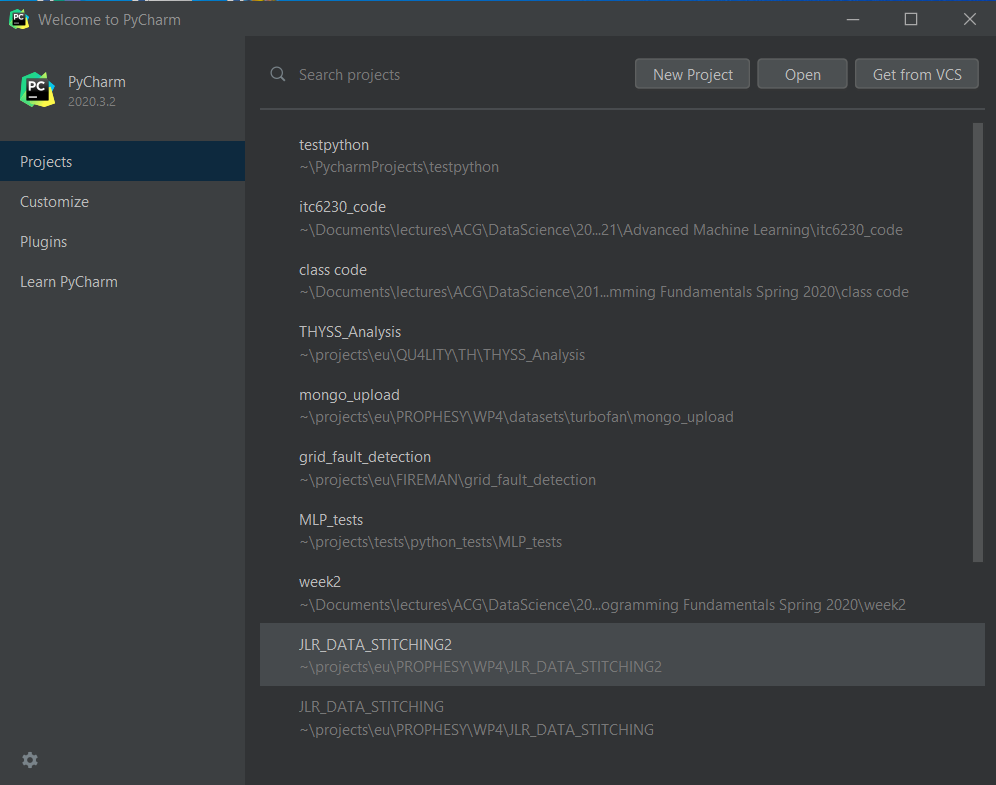
In any case, make sure you have Python 3.7+ installed on your machine (best if you have 3.9+)



### 2. Installing PyCharm IDE on your machine (optional but highly recommended step).

If you don’t have the PyCharm IDE installed on your machine, it is highly recommended that you install it (ACG students are entitled to the Professional version while they have a valid ACG email address)

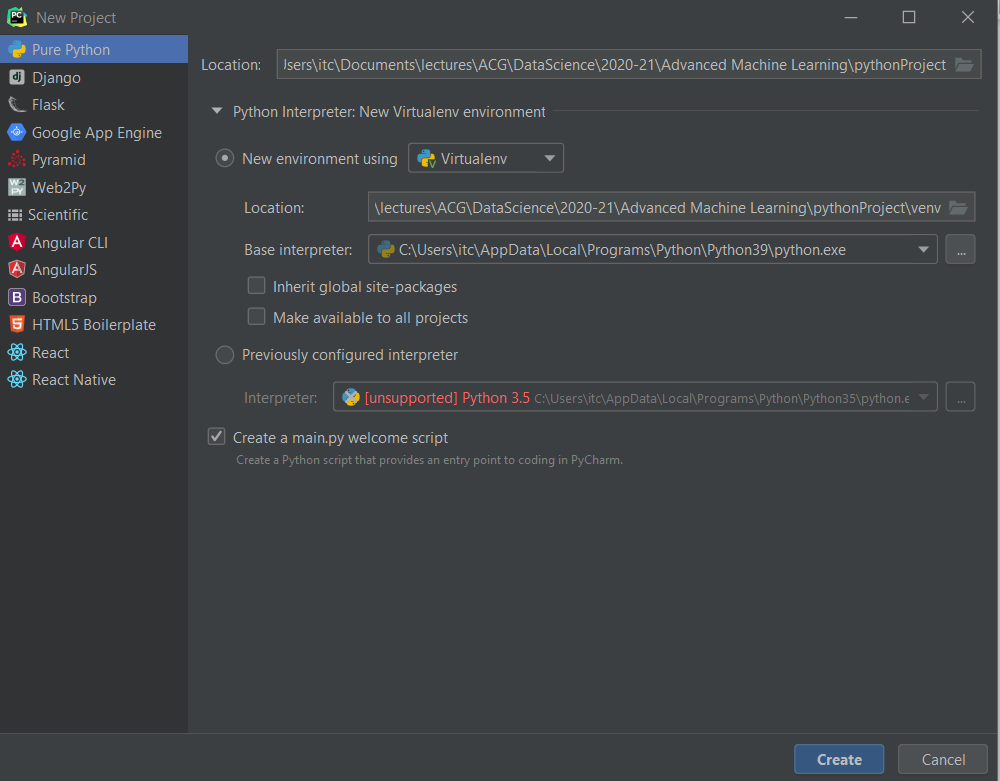
Once you have installed PyCharm, every time you fire up your IDE, a screen like the following will appear (unless you exited without closing at least one project before):



The first time you fire up the IDE the screen will look different since there aren’t going to be any Python projects that the IDE knows about on your machine.

### 3. Using PyCharm for the first time (optional step)

Click the “New Project” button, and then from the next dialogue:

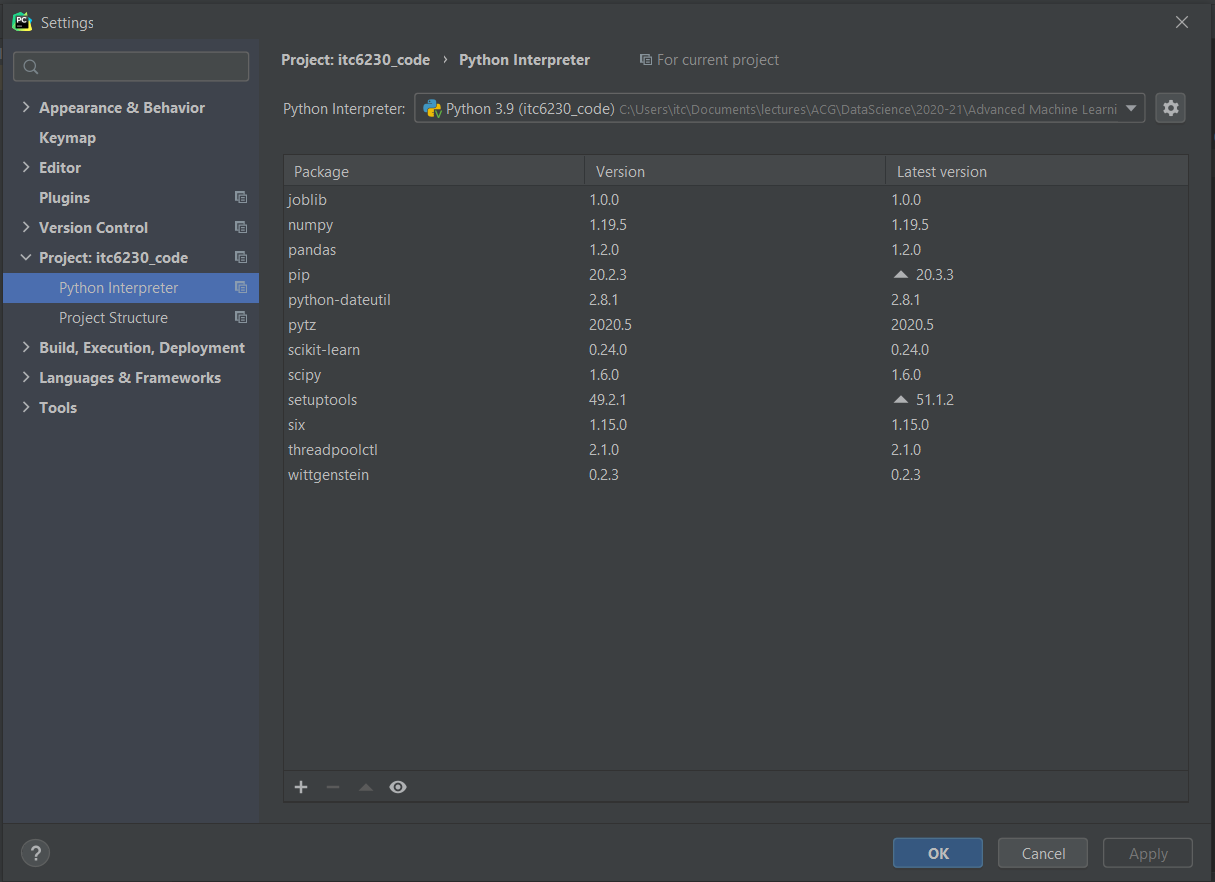


Choose “Pure Python” from the menu on the left, and enter values for the location (the directory where your project’s code and data will be saved). Make sure that the “New environment using “ is set to “Virtualenv”, and that there is a valid value for the “Base interpreter” choice.

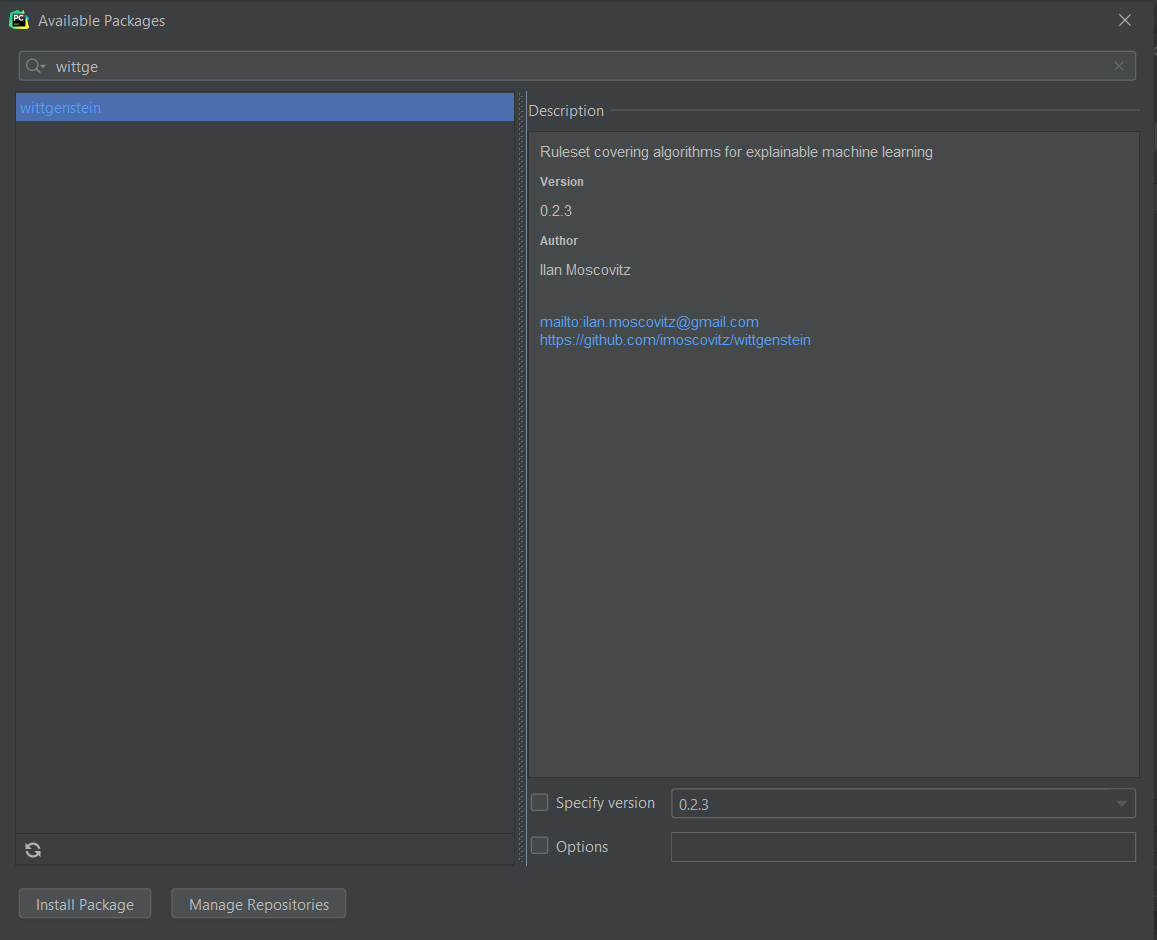
Click the “Create” button on the bottom right and wait until the IDE has finished working on the creation of your new project.

### 4. Working with the “Wittgenstein” Python package for rule-learning:

If you use PyCharm, go to File 🡪Settings 🡪Project <the name of your project> 🡪 Python interpreter to see what packages are already installed in your environment.



Since this is a new project, you will only have two base packages installed. Click the “+” button on the lower end of the dialogue to install more packages. Search for “wittgenstein”, and select it from the results list, as shown below:



Click on “Install Package” button on the lower left corner of the dialogue, and wait until the package is installed.

### HOW TO INSTALL Wittgenstein WITH PIP:

If you don’t use PyCharm, you can always install any package that’s on the PyPi index via the pip package manager:

pip install wittgenstein

### Working with Wittgenstein:

STEP a.

In the same way that you installed the wittgenstein package, install the

* pandas and
* sklearn

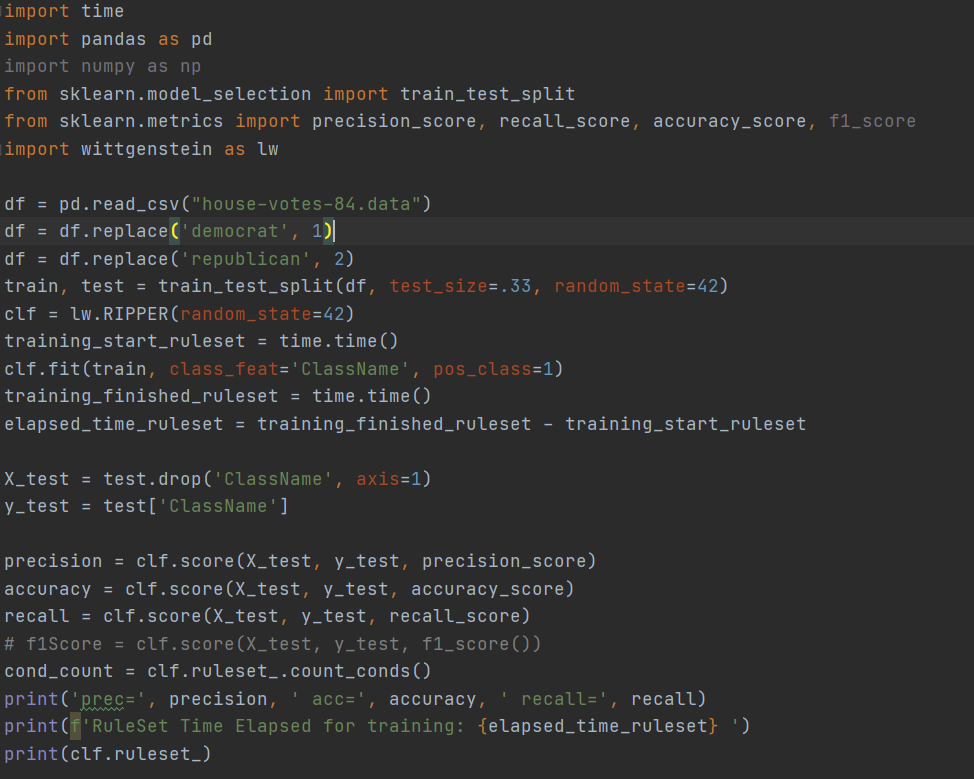
packages in your environment.

STEP b.

Download from BB the file “house-votes-84.data” (it’s a CSV file) and store it in the root directory of your current project.

STEP c.

In a new python file called witt\_test.py (in PyCharm just create a new Python file by right-clicking on your project in the project explorer pane on the left, and choose “New” and then “Python File”), type the following code:



Right-click the window where you wrote your code and choose “Run witt\_test” (or click the green Play button on the top right of the IDE screen).

# HOMEWORK #1 QUESTIONS:

1. (30 pts) In STEP c above, what is the output of the above code?

2. (40 pts total)

a. (15 pts) By carefully looking at the ruleset the model has learned, and given that the “?” value SHOULD represent a “Don’t Know” value on how a congressman/woman voted, what can you say about how the package handles “?” values?

b. (25 pts) Use pandas to replace the “Don’t Know” values with np.nan (Not-A-Number in the numpy library). Re-run the fit method of the Wittgenstein package RIPPER classifier. What accuracy do you obtain now?

3. (30 pts) Modify the house-votes-84.data file so that it contains no missing values. Change every missing value in a cell by the majority of the values for that column among the congressmen/women of the same party (i.e. assume that the congressman voted the same way as most of the other members of their own party). How does this change the accuracy or precision of the model when it is trained on the modified data?