Getting Started with the public_baselines script + anaconda

Install Anaconda for Python v2 with defaults.. its huge. Make a project in PyCharm. Download the files from Kaggle: - a) 5 files we have in the *https://inclass.kaggle.com/c/ca-foscari-link-prediction/data]* - b) 1 file "public_baselines.py" the Italian guy uploaded in the forum

Change!: in the "public_baselines.py" from

```
Around line #58 with open("node_info.csv", "r") as f:
```

TO:

```
with open("node_information.csv", "r") as f:
```

Now in the PyCharm: Press Ctrl + Alt + S (or from menu File>Settings) In the windows that opens from left menu, go to Project: tab and choose the Project Interpreter link Here point to the Anaconda's Python 2 interpreter. Mine was in path "C:\Users\ 'username' \Anaconda2" but PyCharm found it automatically. Let the PyCharm finish any updates it makes and indexes and then RUN the script :) -->For my laptop, it took **15 mins and 10 secs**!!!

It finishes OK, but throughs a WARNING.. that is wants float64 instead of int ... I tried to get rid of this warning and seems that I got right, by making a small change at lines (around) #120 &

```
training_features = np.array([overlap_title, temp_diff,
comm_auth]).astype(np.float64).T
```

& #175

```
testing_features = np.array([overlap_title_test, temp_diff_test,
comm_auth_test]).astype(np.float64).T
```

For getting iGraph + cairo in windows:

- Go to: http://www.lfd.uci.edu/~gohlke/pythonlibs/ and download the x64 or x32 .whl (wheel) files for -- http://www.lfd.uci.edu/~gohlke/pythonlibs/#python-igraph **Check also the Python version** -- http://www.lfd.uci.edu/~gohlke/pythonlibs/#pycairo **Check also the Python version**
- You want to run this: (from: http://stackoverflow.com/questions/34113151/how-to-install-igraph-for-python-on-windows)

python -m pip install path/to/igraph.whl

- If you have many Python installed versions.. and want to install it to anaconda, go to the anaconda installed folder and run the command with python.exe -m pip install...