CS/ECEME532 Period 11 Activity

Preamble

```
In [1]: %matplotlib notebook
        import numpy as np, pandas
        import numpy.ma as ma
        import matplotlib.pyplot as plt, matplotlib.gridspec as gridspec
        from scipy.io import loadmat # read a .mat data
        from sklearn.metrics import *
        from sklearn.linear_model import *
        from sklearn.preprocessing import *
        %load_ext rpy2.ipython
        %r install.packages("ROCR")
        %r install.packages("Rtsne")
        %r,%r_out=%rmagic(verbose=1) %load_ext ipy_table
```

A means may some 'random' components to it. You will get different results depending on your luck. Even when you run an identical code, you will see some different results from your peers. So, - we need the following line of code to start with.

```
In [2]: np.random.seed(1)
```

Indeed, one may be tempted to try so many random seeds until you get a good performance!

Don't do that. :). Some published ML are suffering from 'reproducibility crisis' partially due to this. See these for more details: http://www.unsuite.com/1756-0500-https://www.nature.com/articles/d41586-019-00857-9 https://www.wired.com/story/artificial-intelligence-confronts-reproducibility-crisis/

And see the following figure from the attached paper:

HyfCTwoval (n = 2000)
Fraction Average Error
Fraction Average Error
Fraction Average Error