Untitled

October 22, 2017

In [1]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt

%matplotlib inline
```

In [2]:
```
iris_log = pd.read_csv("Iris.log", sep="\t")
sonar_log = pd.read_csv("Sonar.log", sep="\t")
weighting_log = pd.read_csv("Weighting.log", sep="\t")
```

In [3]:
```
iris_log.head()
```

Out[3]:

	Distance	k	Performance
0	EuclideanDistance	1.0	0.930556
1	EuclideanDistance	2.0	0.948052
2	EuclideanDistance	3.0	0.945205
3	EuclideanDistance	4.0	0.933333
4	EuclideanDistance	5.0	0.949367

In [4]:
```
sonar_log.head()
```

Out[4]:

	Distance	k	Performance
0	EuclideanDistance	1.0	0.864078
1	EuclideanDistance	2.0	0.805825
2	EuclideanDistance	3.0	0.811321
3	EuclideanDistance	4.0	0.830000
4	EuclideanDistance	5.0	0.800000

In [5]:
```
weighting_log.head()
```

Out[5]:

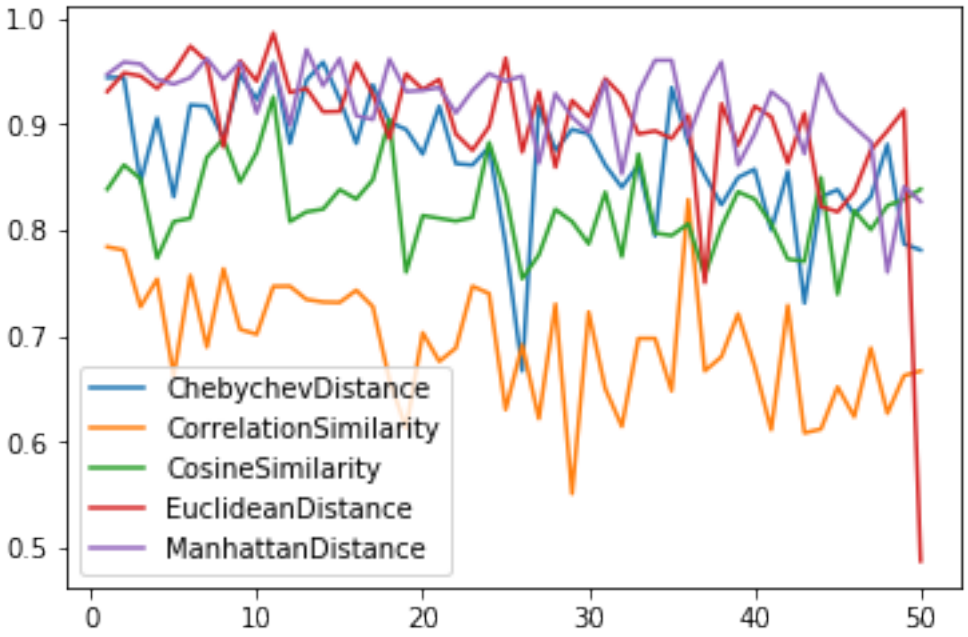
	Distance	k	Performance
0	EuclideanDistance	1.0	0.848980
1	EuclideanDistance	2.0	0.851852
2	EuclideanDistance	3.0	0.836066
3	EuclideanDistance	4.0	0.833992
4	EuclideanDistance	5.0	0.851852

In [6]:
```
dfs = {'Iris': iris_log,
       'Sonar': sonar_log,
       'Weighting': weighting_log}
```

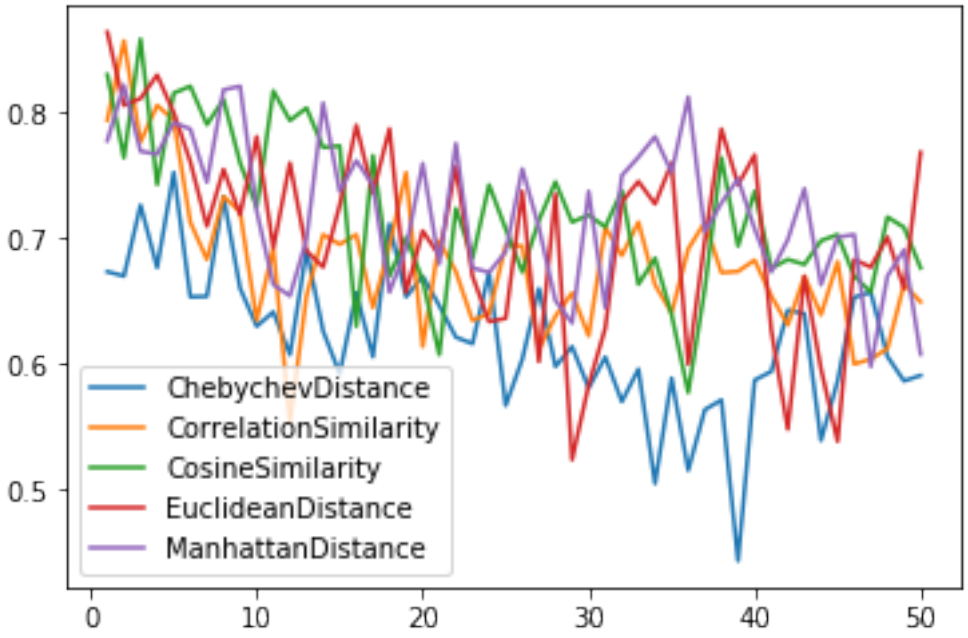
```
In [7]: for df_name, df in dfs.items():
            print('*'*50)
            print("{df}".format(df=df_name))
            for key, group in df.groupby("Distance"):
                print(key, '\t', "Max Perf:", str(group.Performance.max()), "\twith k =", str(gr
                plt.plot(group.k, group.Performance, label=key)

            plt.legend(loc='best')
            plt.show()
```

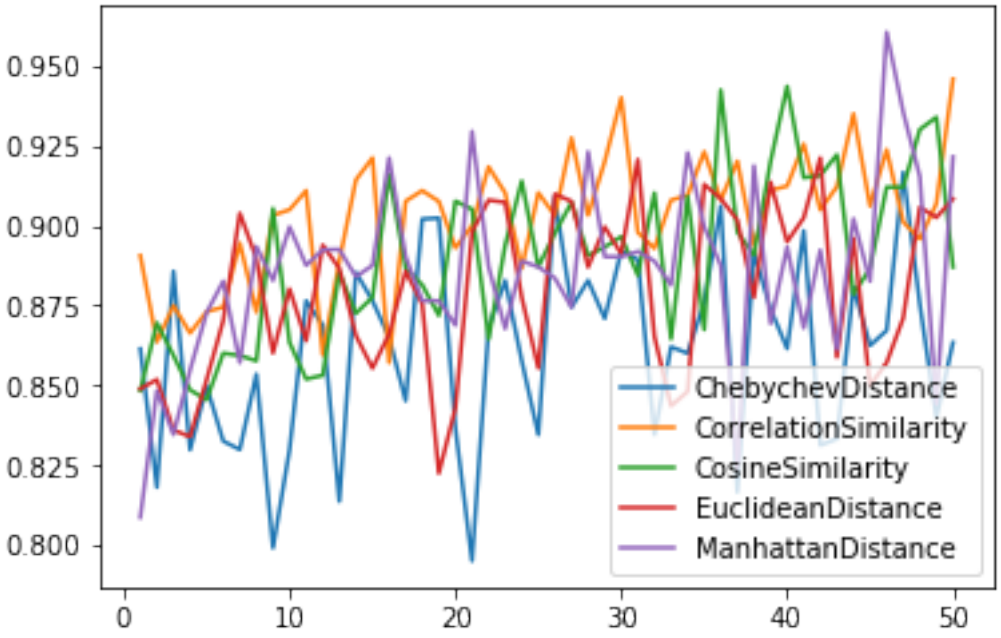
```
**************************************************
Iris
ChebychevDistance           Max Perf: 0.958333333333         with k = 14.0
CorrelationSimilarity           Max Perf: 0.828571428571         with k = 36.0
CosineSimilarity           Max Perf: 0.925373134328         with k = 11.0
EuclideanDistance           Max Perf: 0.985714285714         with k = 11.0
ManhattanDistance           Max Perf: 0.970149253731         with k = 13.0
```




```
**************************************************
Sonar
ChebychevDistance           Max Perf: 0.752475247525         with k = 5.0
CorrelationSimilarity           Max Perf: 0.857142857143         with k = 2.0
CosineSimilarity           Max Perf: 0.858585858586         with k = 3.0
EuclideanDistance           Max Perf: 0.864077669903         with k = 1.0
ManhattanDistance           Max Perf: 0.821782178218         with k = 2.0
```




```
**************************************************
Weighting
ChebychevDistance          Max Perf: 0.916666666667         with k = 47.0
CorrelationSimilarity          Max Perf: 0.945833333333         with k = 50.0
CosineSimilarity          Max Perf: 0.943548387097         with k = 40.0
EuclideanDistance          Max Perf: 0.921161825726         with k = 42.0
ManhattanDistance          Max Perf: 0.9604743083         with k = 46.0
```


0.950
0.925
0.900
0.875
0.850
0.825
0.800
0
10
20
30
40
50
ChebychevDistance
CorrelationSimilarity
CosineSimilarity
EuclideanDistance
ManhattanDistance