

TransactionSimilarityTask2

March 17, 2018

1 Transaction Similarity Task 2

```
In [39]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

```
In [40]: df = pd.read_csv('creditcard.csv')
```

1.1 Creating Sample Data Set

```
In [41]: s1 = df[df['Class']==0]
sampleClass1=s1[0:1000];

s2 = df[df['Class']==1]
sampleClass2=s2[0:5];

sample = sampleClass1.append(sampleClass2)
sample.shape
```

```
Out[41]: (1005, 31)
```

```
In [42]: from sklearn.metrics.pairwise import cosine_similarity
d = sample.as_matrix()
#cosine_similarity(d[0:1], d[1:2])
```

1.2 Util Functions

```
In [43]: #This function calculates Similarity by taking Txn id as parameter and gives top 10 s
def similarity(tId):
    lst=[]

    print('Source Transaction Id: ',tId," Class: ",d[tId][30])
    for i in range(1,len(d)):
        j=i+1
        lst.append(cosine_similarity(d[tId-1:tId], d[i:j])[0][0])

    dic = dict()
```

```

#creating dictionary with Key as Cosine Similarity and Value as Txn Id and Class
for i in range(0,len(lst)):
    dic[lst[i]]=(i,d[i][30])

#sorting to get top 10 results
lst.sort()

for i in range(10):
    print("Transaction id -> ",dic[lst[i]][0]," Class -> ",dic[lst[i]][1]," Similarities -> ",dic[lst[i]][2])

```

In [44]: sampleOfTransactions=[1,1002,1003,75,100,125,150,175,999];

```

for i in range(len(sampleOfTransactions)):
    print("\n")
    print("Transaction ",sampleOfTransactions[i]," Similarities-----")
    similarity(sampleOfTransactions[i]);

```

Transaction 1 Similarities-----

```

Source Transaction Id: 1 Class: 0.0
Transaction id -> 774 Class -> 0.0 Similarity: -4.089715805648982e-05
Transaction id -> 533 Class -> 0.0 Similarity: -7.434535628749325e-06
Transaction id -> 889 Class -> 0.0 Similarity: 7.350922185260839e-05
Transaction id -> 875 Class -> 0.0 Similarity: 7.551756231715775e-05
Transaction id -> 382 Class -> 0.0 Similarity: 7.903751010581647e-05
Transaction id -> 999 Class -> 0.0 Similarity: 8.142714822386527e-05
Transaction id -> 933 Class -> 0.0 Similarity: 8.2005564637028e-05
Transaction id -> 573 Class -> 0.0 Similarity: 8.362488265324227e-05
Transaction id -> 513 Class -> 0.0 Similarity: 0.000101594717993993
Transaction id -> 910 Class -> 0.0 Similarity: 0.00010547711504482537

```

Transaction 1002 Similarities-----

```

Source Transaction Id: 1002 Class: 1.0
Transaction id -> 0 Class -> 0.0 Similarity: 0.5281097473174975
Transaction id -> 999 Class -> 0.0 Similarity: 0.6655842613799403
Transaction id -> 382 Class -> 0.0 Similarity: 0.6656623372046245
Transaction id -> 513 Class -> 0.0 Similarity: 0.6656893848586873
Transaction id -> 533 Class -> 0.0 Similarity: 0.6657027473660231
Transaction id -> 774 Class -> 0.0 Similarity: 0.6657078316066102
Transaction id -> 532 Class -> 0.0 Similarity: 0.6657162167161307
Transaction id -> 933 Class -> 0.0 Similarity: 0.6657300276219009

```

Transaction id -> 573 Class -> 0.0 Similarity: 0.6657382208686583
Transaction id -> 910 Class -> 0.0 Similarity: 0.6657625393944645

Transaction 1003 Similarities-----

Source Transaction Id: 1003 Class: 1.0

Transaction id -> 0 Class -> 0.0 Similarity: 0.037827698175742726
Transaction id -> 1 Class -> 0.0 Similarity: 0.05632676812426587
Transaction id -> 2 Class -> 0.0 Similarity: 0.061742597086238124
Transaction id -> 50 Class -> 0.0 Similarity: 0.07929109388155606
Transaction id -> 163 Class -> 0.0 Similarity: 0.0805319378632574
Transaction id -> 3 Class -> 0.0 Similarity: 0.08208955597384802
Transaction id -> 88 Class -> 0.0 Similarity: 0.10513949088128065
Transaction id -> 19 Class -> 0.0 Similarity: 0.12232761856260588
Transaction id -> 7 Class -> 0.0 Similarity: 0.12814512943579326
Transaction id -> 11 Class -> 0.0 Similarity: 0.13532345695681952

Transaction 75 Similarities-----

Source Transaction Id: 75 Class: 0.0

Transaction id -> 0 Class -> 0.0 Similarity: 0.33442407478166597
Transaction id -> 1 Class -> 0.0 Similarity: 0.4874753048545898
Transaction id -> 2 Class -> 0.0 Similarity: 0.49160098533868024
Transaction id -> 50 Class -> 0.0 Similarity: 0.5071847247338102
Transaction id -> 163 Class -> 0.0 Similarity: 0.5081354296406299
Transaction id -> 3 Class -> 0.0 Similarity: 0.511630970303953
Transaction id -> 88 Class -> 0.0 Similarity: 0.5291394076884974
Transaction id -> 19 Class -> 0.0 Similarity: 0.5437432999757966
Transaction id -> 7 Class -> 0.0 Similarity: 0.5481131424089841
Transaction id -> 11 Class -> 0.0 Similarity: 0.5540570687770399

Transaction 100 Similarities-----

Source Transaction Id: 100 Class: 0.0

Transaction id -> 0 Class -> 0.0 Similarity: 0.17444660712481527
Transaction id -> 1 Class -> 0.0 Similarity: 0.23121807914943482
Transaction id -> 2 Class -> 0.0 Similarity: 0.2360103367198921
Transaction id -> 50 Class -> 0.0 Similarity: 0.25361318033449826
Transaction id -> 163 Class -> 0.0 Similarity: 0.2547915117820027
Transaction id -> 3 Class -> 0.0 Similarity: 0.2560379046736478
Transaction id -> 88 Class -> 0.0 Similarity: 0.2786400387703111
Transaction id -> 19 Class -> 0.0 Similarity: 0.2951291298497902
Transaction id -> 7 Class -> 0.0 Similarity: 0.29960460639997916
Transaction id -> 11 Class -> 0.0 Similarity: 0.3079514711345992

Transaction 125 Similarities-----

Source Transaction Id: 125 Class: 0.0
 Transaction id -> 0 Class -> 0.0 Similarity: 0.43760711479487663
 Transaction id -> 1 Class -> 0.0 Similarity: 0.629499418676218
 Transaction id -> 2 Class -> 0.0 Similarity: 0.6336103905983594
 Transaction id -> 50 Class -> 0.0 Similarity: 0.6471291200576046
 Transaction id -> 163 Class -> 0.0 Similarity: 0.648150047456263
 Transaction id -> 3 Class -> 0.0 Similarity: 0.6490763890653596
 Transaction id -> 88 Class -> 0.0 Similarity: 0.6667642192624242
 Transaction id -> 19 Class -> 0.0 Similarity: 0.6791265823573613
 Transaction id -> 7 Class -> 0.0 Similarity: 0.683184376076088
 Transaction id -> 11 Class -> 0.0 Similarity: 0.6884915576263109

Transaction 150 Similarities-----

Source Transaction Id: 150 Class: 0.0
 Transaction id -> 999 Class -> 0.0 Similarity: 0.27705067072597095
 Transaction id -> 774 Class -> 0.0 Similarity: 0.27710727615189007
 Transaction id -> 382 Class -> 0.0 Similarity: 0.277119861244254
 Transaction id -> 573 Class -> 0.0 Similarity: 0.2771285739385105
 Transaction id -> 513 Class -> 0.0 Similarity: 0.2771302379568234
 Transaction id -> 533 Class -> 0.0 Similarity: 0.2771311871305358
 Transaction id -> 933 Class -> 0.0 Similarity: 0.2771311964741862
 Transaction id -> 532 Class -> 0.0 Similarity: 0.2771519375189602
 Transaction id -> 910 Class -> 0.0 Similarity: 0.27716297966705294
 Transaction id -> 875 Class -> 0.0 Similarity: 0.2771747725235865

Transaction 175 Similarities-----

Source Transaction Id: 175 Class: 0.0
 Transaction id -> 513 Class -> 0.0 Similarity: 0.11856365007017164
 Transaction id -> 533 Class -> 0.0 Similarity: 0.11857555087467089
 Transaction id -> 999 Class -> 0.0 Similarity: 0.11858109945298119
 Transaction id -> 382 Class -> 0.0 Similarity: 0.11858540963778753
 Transaction id -> 774 Class -> 0.0 Similarity: 0.11858833981386777
 Transaction id -> 573 Class -> 0.0 Similarity: 0.1186024844542216
 Transaction id -> 933 Class -> 0.0 Similarity: 0.11860747809850233
 Transaction id -> 532 Class -> 0.0 Similarity: 0.11861678857522318
 Transaction id -> 889 Class -> 0.0 Similarity: 0.11864663179296663
 Transaction id -> 858 Class -> 0.0 Similarity: 0.1186479583362011

Transaction 999 Similarities-----

Source Transaction Id: 999 Class: 0.0
Transaction id -> 0 Class -> 0.0 Similarity: 0.0014292741736162472
Transaction id -> 1 Class -> 0.0 Similarity: 0.0061946556921491985
Transaction id -> 2 Class -> 0.0 Similarity: 0.011626619099284363
Transaction id -> 50 Class -> 0.0 Similarity: 0.02920337768772529
Transaction id -> 163 Class -> 0.0 Similarity: 0.030451437721352004
Transaction id -> 3 Class -> 0.0 Similarity: 0.03215026134128548
Transaction id -> 88 Class -> 0.0 Similarity: 0.055142739374035764
Transaction id -> 19 Class -> 0.0 Similarity: 0.07238190147479662
Transaction id -> 7 Class -> 0.0 Similarity: 0.07833278706679397
Transaction id -> 11 Class -> 0.0 Similarity: 0.08546955839297259