K- Nearest Neighbors on Amazon Food Reviews

In [1]:
```
%matplotlib inline
import warnings
warnings.filterwarnings("ignore")

# General Packages
import os
import sqlite3
import pandas as pd
import numpy as np
import string
import re
import nltk
import datetime
import time

# Plotting Packages
import matplotlib.pyplot as plt
import seaborn as sns

# Packages for Tfidf
from sklearn.feature_extraction.text import TfidfTransformer
from sklearn.feature_extraction.text import TfidfVectorizer

# Packages for BOW (Bag of words)
from sklearn.feature_extraction.text import CountVectorizer
from sklearn.metrics import confusion_matrix
from sklearn import metrics
from sklearn.metrics import roc_curve, auc

# Packages for Text Preprocessing
from nltk.stem.porter import PorterStemmer
from nltk.corpus import stopwords
from nltk.stem.wordnet import WordNetLemmatizer

#Packages for Word2vec, Average Word2vec & Tf-Idf Weighted Word2Vec
from gensim.models import Word2Vec
from gensim.models import KeyedVectors
import pickle

#Packages for plotting Tsne plot
from sklearn.manifold import TSNE

from sklearn.cross_validation import train_test_split
from sklearn.neighbors import KNeighborsClassifier
from sklearn.metrics import accuracy_score
from sklearn.cross_validation import cross_val_score
from collections import Counter
from sklearn.metrics import accuracy_score
from sklearn import cross_validation
from sklearn.model_selection import TimeSeriesSplit
```

```
/usr/lib/python3/dist-packages/sklearn/cross_validation.py:44: DeprecationWarning: This module was deprec
ated in version 0.18 in favor of the model_selection module into which all the refactored classes and fun
ctions are moved. Also note that the interface of the new CV iterators are different from that of this mo
dule. This module will be removed in 0.20.
  "This module will be removed in 0.20.", DeprecationWarning)
```

Preprocessing Stage: Cleansed Stop Words, Punctuations & Html tags

In [2]:
```
#Connecting the Sqlite file after the Preprocessing Stage
#os.chdir('/Users/sujis/Downloads/AI')
con2 = sqlite3.connect('final.sqlite')
final= pd.read_sql_query(""" SELECT * FROM Reviews """, con2)
```

In [3]:
```
final.head(1)
```

Out[3]:

	index	Id	ProductId	UserId	ProfileName	HelpfulnessNumerator	HelpfulnessDenominator	Score	Time	Summary	
0	138706	150524	0006641040	ACITT7DI6IDDL	shari zychinski	0	0	positive	939340800	EVERY book is educational	n

In [4]:
```
# Due to memory constraints taking 2.5k positive & 2.5k negative reviews
final_dataset=pd.DataFrame()
positive_dataset=final.loc[final['Score'] == 'positive'].head(2500)
Negative_dataset=final.loc[final['Score'] == 'negative'].head(2500)
final_dataset=pd.concat([positive_dataset,Negative_dataset])
l1 = positive_dataset['Score']
l2 = Negative_dataset['Score']
label=pd.concat([l1,l2])
```

In [5]:
```
# Converting the Epoch Date format to normal date format
Time_P=[]
Time_N=[]
for each in positive_dataset['Time']:
    Time_P.append(time.strftime('%Y-%m-%d', time.localtime(each)))
for each in Negative_dataset['Time']:
    Time_N.append(time.strftime('%Y-%m-%d', time.localtime(each)))
Time=Time_P+Time_N
```

In [6]:
```
#BoW
count_vect = CountVectorizer() #in scikit-learn
final_counts = count_vect.fit_transform(final_dataset['CleanedText'].values)
print("the type of count vectorizer ",type(final_counts))
print("the shape of out text BOW vectorizer ",final_counts.get_shape())
print("the number of unique words ", final_counts.get_shape()[1])
```

```
the type of count vectorizer  <class 'scipy.sparse.csr.csr_matrix'>
the shape of out text BOW vectorizer  (5000, 10430)
the number of unique words  10430
```

In [8]:
```
# Converting Sparse matric to array
final_counts=final_counts.toarray()
```

In [9]:
```
#tsne
start_time_code = time.time()
model = TSNE(n_components=2, random_state=0, perplexity=30)
tsne_data = model.fit_transform(final_counts)
end_time_code = time.time()
print ("Running Time for code execution " + str(end_time_code - start_time_code) + " secs")
```

```
Running Time for code execution 592.2491195201874 secs
```

In [10]:
```
tsne_bow_data = np.vstack((tsne_data.T,label,Time)).T
```

In [11]:
```
tsne_bow_data.shape
```

Out[11]:
```
(5000, 4)
```

In [41]:
```
#Saving tsne data into seperate file for further usecases.
tsne_bow_df = pd.DataFrame(data=tsne_bow_data, columns=("X", "Y", "Label","Date"))
tsne_bow_df.to_csv('/home/venkatasujit272/tsne_bow_data.csv')
```

In [38]:
```
# Aliasing the Score field values into Binary format
def partition(x):
    if x == 'negative':
        return 0.
    return 1.

# Aliasing the Score to String Format (Positive for Score > 3 & Negative for Score  < 3)

actualScore = tsne_bow_df['Label']
positiveNegative = actualScore.map(partition)
tsne_bow_df['Label'] = positiveNegative
```

In [14]:
```
tsne_bow_df.head(1)
```

Out[14]:

	X	Y	Label	Date
0	3.95061	-6.3007	1.0	1999-10-08

In [15]:
```
#Sorting the Dataframe with Date to apply Timebased Split
tsne_bow_df.sort_values(by='Date',ascending=1)
```

Out[15]:

	X	Y	Label	Date
0	3.95061	-6.3007	1.0	1999-10-08
28	10.3196	0.287563	1.0	1999-10-25
377	6.70138	3.92468	1.0	1999-12-02
287	11.2809	0.770661	1.0	1999-12-06
376	6.7136	4.17597	1.0	2000-01-03
209	6.6723	3.98859	1.0	2000-01-09
2540	6.65033	3.90467	0.0	2000-01-19
198	5.98368	2.05558	1.0	2000-01-24
224	-6.36712	3.0824	1.0	2000-02-26
284	3.31828	3.42525	1.0	2000-06-03
378	3.31857	3.42551	1.0	2000-06-03
205	7.54684	2.14946	1.0	2000-06-23
206	7.4365	1.14201	1.0	2000-06-29
415	3.7623	1.40381	1.0	2000-07-31
684	8.15351	2.93929	1.0	2000-08-09
714	8.09218	2.87029	1.0	2000-08-09
212	8.90946	3.81539	1.0	2000-08-15
258	7.47962	4.37065	1.0	2000-10-03
691	5.38752	-6.31318	1.0	2000-12-05
315	-0.326053	-1.9004	1.0	2000-12-19
286	6.76072	4.30598	1.0	2000-12-30
692	8.29616	2.9847	1.0	2001-02-22
2547	11.1145	-1.07556	0.0	2001-06-11
2543	6.82782	4.33278	0.0	2001-08-08
233	5.08881	-0.01522	1.0	2001-09-24
701	8.19487	2.50405	1.0	2001-10-23
309	7.88662	-1.73058	1.0	2001-10-26
30	11.6625	0.0439728	1.0	2001-12-26
830	5.69966	-3.07542	1.0	2002-01-06
288	6.6525	4.14869	1.0	2002-02-04
...	...	...	...	...
232	9.56216	3.22136	1.0	2012-10-22
3856	7.94275	0.942424	0.0	2012-10-23
1792	-0.344878	2.78245	1.0	2012-10-23
813	3.68547	0.383336	1.0	2012-10-23
3649	4.61891	-2.95272	0.0	2012-10-23
2030	4.28381	4.84058	1.0	2012-10-23
1897	11.5943	-1.59847	1.0	2012-10-23
1086	7.93208	-0.563009	1.0	2012-10-23
3056	11.634	-2.49857	0.0	2012-10-23
1594	-0.34193	1.33567	1.0	2012-10-23
4556	7.18706	0.136966	0.0	2012-10-23
134	-0.500191	2.4304	1.0	2012-10-23
2717	2.57152	1.79645	0.0	2012-10-24
2705	6.26789	0.674835	0.0	2012-10-24
1295	7.01916	1.18204	1.0	2012-10-24
2391	8.2813	3.01715	1.0	2012-10-24
1603	1.27944	-5.16483	1.0	2012-10-25
3849	7.57325	-0.891235	0.0	2012-10-25
2715	3.31716	0.335818	0.0	2012-10-25
4115	7.10874	-5.62311	0.0	2012-10-25
1827	-1.43616	0.38397	1.0	2012-10-25
2978	5.04737	-1.58026	0.0	2012-10-26

	X	Y	Label	Date
135	2.42842	-1.21995	1.0	2012-10-26
782	7.05385	-1.49977	1.0	2012-10-26
136	-0.0579819	1.21796	1.0	2012-10-26
3569	4.14018	-0.925766	0.0	2012-10-26
1349	0.94107	-4.24592	1.0	2012-10-26
493	4.54338	1.3034	1.0	2012-10-26
4793	2.80861	-6.49147	0.0	2012-10-26
2592	11.4845	-4.11881	0.0	2012-10-26

5000 rows × 4 columns

In [16]:
```
# create design matrix X and target vector y
X = np.array(tsne_bow_df.iloc[:, 0:4]) # end index is exclusive
Y = np.array(tsne_bow_df['Label']) # showing you two ways of indexing a pandas df
```

In [17]:
```
#Time based split
tscv = TimeSeriesSplit(n_splits=5)
print(tscv)
```

```
TimeSeriesSplit(n_splits=5)
```

In [18]:
```
#Splitting the Dataset into Train set & Test set
for train_index, test_index in tscv.split(X):
   print("TRAIN:", train_index, "TEST:", test_index)
   X_tr, X_test = X[train_index], X[test_index]
   Y_tr, Y_test = Y[train_index], Y[test_index]
```

```
TRAIN: [  0   1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17
  18  19  20  21  22  23  24  25  26  27  28  29  30  31  32  33  34  35
  36  37  38  39  40  41  42  43  44  45  46  47  48  49  50  51  52  53
  54  55  56  57  58  59  60  61  62  63  64  65  66  67  68  69  70  71
  72  73  74  75  76  77  78  79  80  81  82  83  84  85  86  87  88  89
  90  91  92  93  94  95  96  97  98  99 100 101 102 103 104 105 106 107
 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125
 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143
 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161
 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179
 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197
 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215
 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233
 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251
 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269
 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287
 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305
 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323
 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341
 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359
 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377
 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395
 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413
 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431
 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449
 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467
 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485
 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503
 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521
 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539
 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557
 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575
 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593
 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611
 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629
 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647
 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665
 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683
 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701
 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719
 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737
 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755
 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773
 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791
 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809
 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827
 828 829 830 831 832 833 834] TEST: [ 835  836  837  838  839  840  841  842  843  844  845  846  847  84
8
  849  850  851  852  853  854  855  856  857  858  859  860  861  862
  863  864  865  866  867  868  869  870  871  872  873  874  875  876
  877  878  879  880  881  882  883  884  885  886  887  888  889  890
  891  892  893  894  895  896  897  898  899  900  901  902  903  904
  905  906  907  908  909  910  911  912  913  914  915  916  917  918
  919  920  921  922  923  924  925  926  927  928  929  930  931  932
  933  934  935  936  937  938  939  940  941  942  943  944  945  946
  947  948  949  950  951  952  953  954  955  956  957  958  959  960
  961  962  963  964  965  966  967  968  969  970  971  972  973  974
  975  976  977  978  979  980  981  982  983  984  985  986  987  988
  989  990  991  992  993  994  995  996  997  998  999 1000 1001 1002
 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016
 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030
 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044
 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058
 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072
 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086
 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100
 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114
 1115 1116 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128
 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142
 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156
 1157 1158 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170
 1171 1172 1173 1174 1175 1176 1177 1178 1179 1180 1181 1182 1183 1184
 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198
 1199 1200 1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212
 1213 1214 1215 1216 1217 1218 1219 1220 1221 1222 1223 1224 1225 1226
 1227 1228 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240
 1241 1242 1243 1244 1245 1246 1247 1248 1249 1250 1251 1252 1253 1254
 1255 1256 1257 1258 1259 1260 1261 1262 1263 1264 1265 1266 1267 1268
 1269 1270 1271 1272 1273 1274 1275 1276 1277 1278 1279 1280 1281 1282
 1283 1284 1285 1286 1287 1288 1289 1290 1291 1292 1293 1294 1295 1296
 1297 1298 1299 1300 1301 1302 1303 1304 1305 1306 1307 1308 1309 1310
 1311 1312 1313 1314 1315 1316 1317 1318 1319 1320 1321 1322 1323 1324
 1325 1326 1327 1328 1329 1330 1331 1332 1333 1334 1335 1336 1337 1338
 1339 1340 1341 1342 1343 1344 1345 1346 1347 1348 1349 1350 1351 1352
 1353 1354 1355 1356 1357 1358 1359 1360 1361 1362 1363 1364 1365 1366
 1367 1368 1369 1370 1371 1372 1373 1374 1375 1376 1377 1378 1379 1380
 1381 1382 1383 1384 1385 1386 1387 1388 1389 1390 1391 1392 1393 1394
```

```
 1395 1396 1397 1398 1399 1400 1401 1402 1403 1404 1405 1406 1407 1408
 1409 1410 1411 1412 1413 1414 1415 1416 1417 1418 1419 1420 1421 1422
 1423 1424 1425 1426 1427 1428 1429 1430 1431 1432 1433 1434 1435 1436
 1437 1438 1439 1440 1441 1442 1443 1444 1445 1446 1447 1448 1449 1450
 1451 1452 1453 1454 1455 1456 1457 1458 1459 1460 1461 1462 1463 1464
 1465 1466 1467 1468 1469 1470 1471 1472 1473 1474 1475 1476 1477 1478
 1479 1480 1481 1482 1483 1484 1485 1486 1487 1488 1489 1490 1491 1492
 1493 1494 1495 1496 1497 1498 1499 1500 1501 1502 1503 1504 1505 1506
 1507 1508 1509 1510 1511 1512 1513 1514 1515 1516 1517 1518 1519 1520
 1521 1522 1523 1524 1525 1526 1527 1528 1529 1530 1531 1532 1533 1534
 1535 1536 1537 1538 1539 1540 1541 1542 1543 1544 1545 1546 1547 1548
 1549 1550 1551 1552 1553 1554 1555 1556 1557 1558 1559 1560 1561 1562
 1563 1564 1565 1566 1567 1568 1569 1570 1571 1572 1573 1574 1575 1576
 1577 1578 1579 1580 1581 1582 1583 1584 1585 1586 1587 1588 1589 1590
 1591 1592 1593 1594 1595 1596 1597 1598 1599 1600 1601 1602 1603 1604
 1605 1606 1607 1608 1609 1610 1611 1612 1613 1614 1615 1616 1617 1618
 1619 1620 1621 1622 1623 1624 1625 1626 1627 1628 1629 1630 1631 1632
 1633 1634 1635 1636 1637 1638 1639 1640 1641 1642 1643 1644 1645 1646
 1647 1648 1649 1650 1651 1652 1653 1654 1655 1656 1657 1658 1659 1660
 1661 1662 1663 1664 1665 1666 1667]
TRAIN: [   0    1    2 ... 1665 1666 1667] TEST: [1668 1669 1670 1671 1672 1673 1674 1675 1676 1677 1678
1679 1680 1681
 1682 1683 1684 1685 1686 1687 1688 1689 1690 1691 1692 1693 1694 1695
 1696 1697 1698 1699 1700 1701 1702 1703 1704 1705 1706 1707 1708 1709
 1710 1711 1712 1713 1714 1715 1716 1717 1718 1719 1720 1721 1722 1723
 1724 1725 1726 1727 1728 1729 1730 1731 1732 1733 1734 1735 1736 1737
 1738 1739 1740 1741 1742 1743 1744 1745 1746 1747 1748 1749 1750 1751
 1752 1753 1754 1755 1756 1757 1758 1759 1760 1761 1762 1763 1764 1765
 1766 1767 1768 1769 1770 1771 1772 1773 1774 1775 1776 1777 1778 1779
 1780 1781 1782 1783 1784 1785 1786 1787 1788 1789 1790 1791 1792 1793
 1794 1795 1796 1797 1798 1799 1800 1801 1802 1803 1804 1805 1806 1807
 1808 1809 1810 1811 1812 1813 1814 1815 1816 1817 1818 1819 1820 1821
 1822 1823 1824 1825 1826 1827 1828 1829 1830 1831 1832 1833 1834 1835
 1836 1837 1838 1839 1840 1841 1842 1843 1844 1845 1846 1847 1848 1849
 1850 1851 1852 1853 1854 1855 1856 1857 1858 1859 1860 1861 1862 1863
 1864 1865 1866 1867 1868 1869 1870 1871 1872 1873 1874 1875 1876 1877
 1878 1879 1880 1881 1882 1883 1884 1885 1886 1887 1888 1889 1890 1891
 1892 1893 1894 1895 1896 1897 1898 1899 1900 1901 1902 1903 1904 1905
 1906 1907 1908 1909 1910 1911 1912 1913 1914 1915 1916 1917 1918 1919
 1920 1921 1922 1923 1924 1925 1926 1927 1928 1929 1930 1931 1932 1933
 1934 1935 1936 1937 1938 1939 1940 1941 1942 1943 1944 1945 1946 1947
 1948 1949 1950 1951 1952 1953 1954 1955 1956 1957 1958 1959 1960 1961
 1962 1963 1964 1965 1966 1967 1968 1969 1970 1971 1972 1973 1974 1975
 1976 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989
 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003
 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017
 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031
 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045
 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059
 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073
 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087
 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101
 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115
 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129
 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143
 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157
 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171
 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185
 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199
 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213
 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227
 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241
 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255
 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269
 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283
 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297
 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311
 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325
 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339
 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353
 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367
 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381
 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395
 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409
 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423
 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437
 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451
 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465
 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479
 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493
 2494 2495 2496 2497 2498 2499 2500]
TRAIN: [   0    1    2 ... 2498 2499 2500] TEST: [2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511
2512 2513 2514
 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528
 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542
 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556
 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570
 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584
```

```
 2585 2586 2587 2588 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598
 2599 2600 2601 2602 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612
 2613 2614 2615 2616 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626
 2627 2628 2629 2630 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640
 2641 2642 2643 2644 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654
 2655 2656 2657 2658 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668
 2669 2670 2671 2672 2673 2674 2675 2676 2677 2678 2679 2680 2681 2682
 2683 2684 2685 2686 2687 2688 2689 2690 2691 2692 2693 2694 2695 2696
 2697 2698 2699 2700 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710
 2711 2712 2713 2714 2715 2716 2717 2718 2719 2720 2721 2722 2723 2724
 2725 2726 2727 2728 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738
 2739 2740 2741 2742 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752
 2753 2754 2755 2756 2757 2758 2759 2760 2761 2762 2763 2764 2765 2766
 2767 2768 2769 2770 2771 2772 2773 2774 2775 2776 2777 2778 2779 2780
 2781 2782 2783 2784 2785 2786 2787 2788 2789 2790 2791 2792 2793 2794
 2795 2796 2797 2798 2799 2800 2801 2802 2803 2804 2805 2806 2807 2808
 2809 2810 2811 2812 2813 2814 2815 2816 2817 2818 2819 2820 2821 2822
 2823 2824 2825 2826 2827 2828 2829 2830 2831 2832 2833 2834 2835 2836
 2837 2838 2839 2840 2841 2842 2843 2844 2845 2846 2847 2848 2849 2850
 2851 2852 2853 2854 2855 2856 2857 2858 2859 2860 2861 2862 2863 2864
 2865 2866 2867 2868 2869 2870 2871 2872 2873 2874 2875 2876 2877 2878
 2879 2880 2881 2882 2883 2884 2885 2886 2887 2888 2889 2890 2891 2892
 2893 2894 2895 2896 2897 2898 2899 2900 2901 2902 2903 2904 2905 2906
 2907 2908 2909 2910 2911 2912 2913 2914 2915 2916 2917 2918 2919 2920
 2921 2922 2923 2924 2925 2926 2927 2928 2929 2930 2931 2932 2933 2934
 2935 2936 2937 2938 2939 2940 2941 2942 2943 2944 2945 2946 2947 2948
 2949 2950 2951 2952 2953 2954 2955 2956 2957 2958 2959 2960 2961 2962
 2963 2964 2965 2966 2967 2968 2969 2970 2971 2972 2973 2974 2975 2976
 2977 2978 2979 2980 2981 2982 2983 2984 2985 2986 2987 2988 2989 2990
 2991 2992 2993 2994 2995 2996 2997 2998 2999 3000 3001 3002 3003 3004
 3005 3006 3007 3008 3009 3010 3011 3012 3013 3014 3015 3016 3017 3018
 3019 3020 3021 3022 3023 3024 3025 3026 3027 3028 3029 3030 3031 3032
 3033 3034 3035 3036 3037 3038 3039 3040 3041 3042 3043 3044 3045 3046
 3047 3048 3049 3050 3051 3052 3053 3054 3055 3056 3057 3058 3059 3060
 3061 3062 3063 3064 3065 3066 3067 3068 3069 3070 3071 3072 3073 3074
 3075 3076 3077 3078 3079 3080 3081 3082 3083 3084 3085 3086 3087 3088
 3089 3090 3091 3092 3093 3094 3095 3096 3097 3098 3099 3100 3101 3102
 3103 3104 3105 3106 3107 3108 3109 3110 3111 3112 3113 3114 3115 3116
 3117 3118 3119 3120 3121 3122 3123 3124 3125 3126 3127 3128 3129 3130
 3131 3132 3133 3134 3135 3136 3137 3138 3139 3140 3141 3142 3143 3144
 3145 3146 3147 3148 3149 3150 3151 3152 3153 3154 3155 3156 3157 3158
 3159 3160 3161 3162 3163 3164 3165 3166 3167 3168 3169 3170 3171 3172
 3173 3174 3175 3176 3177 3178 3179 3180 3181 3182 3183 3184 3185 3186
 3187 3188 3189 3190 3191 3192 3193 3194 3195 3196 3197 3198 3199 3200
 3201 3202 3203 3204 3205 3206 3207 3208 3209 3210 3211 3212 3213 3214
 3215 3216 3217 3218 3219 3220 3221 3222 3223 3224 3225 3226 3227 3228
 3229 3230 3231 3232 3233 3234 3235 3236 3237 3238 3239 3240 3241 3242
 3243 3244 3245 3246 3247 3248 3249 3250 3251 3252 3253 3254 3255 3256
 3257 3258 3259 3260 3261 3262 3263 3264 3265 3266 3267 3268 3269 3270
 3271 3272 3273 3274 3275 3276 3277 3278 3279 3280 3281 3282 3283 3284
 3285 3286 3287 3288 3289 3290 3291 3292 3293 3294 3295 3296 3297 3298
 3299 3300 3301 3302 3303 3304 3305 3306 3307 3308 3309 3310 3311 3312
 3313 3314 3315 3316 3317 3318 3319 3320 3321 3322 3323 3324 3325 3326
 3327 3328 3329 3330 3331 3332 3333]
TRAIN: [   0    1    2 ... 3331 3332 3333] TEST: [3334 3335 3336 3337 3338 3339 3340 3341 3342 3343 3344
3345 3346 3347
 3348 3349 3350 3351 3352 3353 3354 3355 3356 3357 3358 3359 3360 3361
 3362 3363 3364 3365 3366 3367 3368 3369 3370 3371 3372 3373 3374 3375
 3376 3377 3378 3379 3380 3381 3382 3383 3384 3385 3386 3387 3388 3389
 3390 3391 3392 3393 3394 3395 3396 3397 3398 3399 3400 3401 3402 3403
 3404 3405 3406 3407 3408 3409 3410 3411 3412 3413 3414 3415 3416 3417
 3418 3419 3420 3421 3422 3423 3424 3425 3426 3427 3428 3429 3430 3431
 3432 3433 3434 3435 3436 3437 3438 3439 3440 3441 3442 3443 3444 3445
 3446 3447 3448 3449 3450 3451 3452 3453 3454 3455 3456 3457 3458 3459
 3460 3461 3462 3463 3464 3465 3466 3467 3468 3469 3470 3471 3472 3473
 3474 3475 3476 3477 3478 3479 3480 3481 3482 3483 3484 3485 3486 3487
 3488 3489 3490 3491 3492 3493 3494 3495 3496 3497 3498 3499 3500 3501
 3502 3503 3504 3505 3506 3507 3508 3509 3510 3511 3512 3513 3514 3515
 3516 3517 3518 3519 3520 3521 3522 3523 3524 3525 3526 3527 3528 3529
 3530 3531 3532 3533 3534 3535 3536 3537 3538 3539 3540 3541 3542 3543
 3544 3545 3546 3547 3548 3549 3550 3551 3552 3553 3554 3555 3556 3557
 3558 3559 3560 3561 3562 3563 3564 3565 3566 3567 3568 3569 3570 3571
 3572 3573 3574 3575 3576 3577 3578 3579 3580 3581 3582 3583 3584 3585
 3586 3587 3588 3589 3590 3591 3592 3593 3594 3595 3596 3597 3598 3599
 3600 3601 3602 3603 3604 3605 3606 3607 3608 3609 3610 3611 3612 3613
 3614 3615 3616 3617 3618 3619 3620 3621 3622 3623 3624 3625 3626 3627
 3628 3629 3630 3631 3632 3633 3634 3635 3636 3637 3638 3639 3640 3641
 3642 3643 3644 3645 3646 3647 3648 3649 3650 3651 3652 3653 3654 3655
 3656 3657 3658 3659 3660 3661 3662 3663 3664 3665 3666 3667 3668 3669
 3670 3671 3672 3673 3674 3675 3676 3677 3678 3679 3680 3681 3682 3683
 3684 3685 3686 3687 3688 3689 3690 3691 3692 3693 3694 3695 3696 3697
 3698 3699 3700 3701 3702 3703 3704 3705 3706 3707 3708 3709 3710 3711
 3712 3713 3714 3715 3716 3717 3718 3719 3720 3721 3722 3723 3724 3725
 3726 3727 3728 3729 3730 3731 3732 3733 3734 3735 3736 3737 3738 3739
 3740 3741 3742 3743 3744 3745 3746 3747 3748 3749 3750 3751 3752 3753
 3754 3755 3756 3757 3758 3759 3760 3761 3762 3763 3764 3765 3766 3767
 3768 3769 3770 3771 3772 3773 3774 3775 3776 3777 3778 3779 3780 3781
 3782 3783 3784 3785 3786 3787 3788 3789 3790 3791 3792 3793 3794 3795
```

```
 3796 3797 3798 3799 3800 3801 3802 3803 3804 3805 3806 3807 3808 3809
 3810 3811 3812 3813 3814 3815 3816 3817 3818 3819 3820 3821 3822 3823
 3824 3825 3826 3827 3828 3829 3830 3831 3832 3833 3834 3835 3836 3837
 3838 3839 3840 3841 3842 3843 3844 3845 3846 3847 3848 3849 3850 3851
 3852 3853 3854 3855 3856 3857 3858 3859 3860 3861 3862 3863 3864 3865
 3866 3867 3868 3869 3870 3871 3872 3873 3874 3875 3876 3877 3878 3879
 3880 3881 3882 3883 3884 3885 3886 3887 3888 3889 3890 3891 3892 3893
 3894 3895 3896 3897 3898 3899 3900 3901 3902 3903 3904 3905 3906 3907
 3908 3909 3910 3911 3912 3913 3914 3915 3916 3917 3918 3919 3920 3921
 3922 3923 3924 3925 3926 3927 3928 3929 3930 3931 3932 3933 3934 3935
 3936 3937 3938 3939 3940 3941 3942 3943 3944 3945 3946 3947 3948 3949
 3950 3951 3952 3953 3954 3955 3956 3957 3958 3959 3960 3961 3962 3963
 3964 3965 3966 3967 3968 3969 3970 3971 3972 3973 3974 3975 3976 3977
 3978 3979 3980 3981 3982 3983 3984 3985 3986 3987 3988 3989 3990 3991
 3992 3993 3994 3995 3996 3997 3998 3999 4000 4001 4002 4003 4004 4005
 4006 4007 4008 4009 4010 4011 4012 4013 4014 4015 4016 4017 4018 4019
 4020 4021 4022 4023 4024 4025 4026 4027 4028 4029 4030 4031 4032 4033
 4034 4035 4036 4037 4038 4039 4040 4041 4042 4043 4044 4045 4046 4047
 4048 4049 4050 4051 4052 4053 4054 4055 4056 4057 4058 4059 4060 4061
 4062 4063 4064 4065 4066 4067 4068 4069 4070 4071 4072 4073 4074 4075
 4076 4077 4078 4079 4080 4081 4082 4083 4084 4085 4086 4087 4088 4089
 4090 4091 4092 4093 4094 4095 4096 4097 4098 4099 4100 4101 4102 4103
 4104 4105 4106 4107 4108 4109 4110 4111 4112 4113 4114 4115 4116 4117
 4118 4119 4120 4121 4122 4123 4124 4125 4126 4127 4128 4129 4130 4131
 4132 4133 4134 4135 4136 4137 4138 4139 4140 4141 4142 4143 4144 4145
 4146 4147 4148 4149 4150 4151 4152 4153 4154 4155 4156 4157 4158 4159
 4160 4161 4162 4163 4164 4165 4166]
TRAIN: [   0    1    2 ... 4164 4165 4166] TEST: [4167 4168 4169 4170 4171 4172 4173 4174 4175 4176 4177
4178 4179 4180
 4181 4182 4183 4184 4185 4186 4187 4188 4189 4190 4191 4192 4193 4194
 4195 4196 4197 4198 4199 4200 4201 4202 4203 4204 4205 4206 4207 4208
 4209 4210 4211 4212 4213 4214 4215 4216 4217 4218 4219 4220 4221 4222
 4223 4224 4225 4226 4227 4228 4229 4230 4231 4232 4233 4234 4235 4236
 4237 4238 4239 4240 4241 4242 4243 4244 4245 4246 4247 4248 4249 4250
 4251 4252 4253 4254 4255 4256 4257 4258 4259 4260 4261 4262 4263 4264
 4265 4266 4267 4268 4269 4270 4271 4272 4273 4274 4275 4276 4277 4278
 4279 4280 4281 4282 4283 4284 4285 4286 4287 4288 4289 4290 4291 4292
 4293 4294 4295 4296 4297 4298 4299 4300 4301 4302 4303 4304 4305 4306
 4307 4308 4309 4310 4311 4312 4313 4314 4315 4316 4317 4318 4319 4320
 4321 4322 4323 4324 4325 4326 4327 4328 4329 4330 4331 4332 4333 4334
 4335 4336 4337 4338 4339 4340 4341 4342 4343 4344 4345 4346 4347 4348
 4349 4350 4351 4352 4353 4354 4355 4356 4357 4358 4359 4360 4361 4362
 4363 4364 4365 4366 4367 4368 4369 4370 4371 4372 4373 4374 4375 4376
 4377 4378 4379 4380 4381 4382 4383 4384 4385 4386 4387 4388 4389 4390
 4391 4392 4393 4394 4395 4396 4397 4398 4399 4400 4401 4402 4403 4404
 4405 4406 4407 4408 4409 4410 4411 4412 4413 4414 4415 4416 4417 4418
 4419 4420 4421 4422 4423 4424 4425 4426 4427 4428 4429 4430 4431 4432
 4433 4434 4435 4436 4437 4438 4439 4440 4441 4442 4443 4444 4445 4446
 4447 4448 4449 4450 4451 4452 4453 4454 4455 4456 4457 4458 4459 4460
 4461 4462 4463 4464 4465 4466 4467 4468 4469 4470 4471 4472 4473 4474
 4475 4476 4477 4478 4479 4480 4481 4482 4483 4484 4485 4486 4487 4488
 4489 4490 4491 4492 4493 4494 4495 4496 4497 4498 4499 4500 4501 4502
 4503 4504 4505 4506 4507 4508 4509 4510 4511 4512 4513 4514 4515 4516
 4517 4518 4519 4520 4521 4522 4523 4524 4525 4526 4527 4528 4529 4530
 4531 4532 4533 4534 4535 4536 4537 4538 4539 4540 4541 4542 4543 4544
 4545 4546 4547 4548 4549 4550 4551 4552 4553 4554 4555 4556 4557 4558
 4559 4560 4561 4562 4563 4564 4565 4566 4567 4568 4569 4570 4571 4572
 4573 4574 4575 4576 4577 4578 4579 4580 4581 4582 4583 4584 4585 4586
 4587 4588 4589 4590 4591 4592 4593 4594 4595 4596 4597 4598 4599 4600
 4601 4602 4603 4604 4605 4606 4607 4608 4609 4610 4611 4612 4613 4614
 4615 4616 4617 4618 4619 4620 4621 4622 4623 4624 4625 4626 4627 4628
 4629 4630 4631 4632 4633 4634 4635 4636 4637 4638 4639 4640 4641 4642
 4643 4644 4645 4646 4647 4648 4649 4650 4651 4652 4653 4654 4655 4656
 4657 4658 4659 4660 4661 4662 4663 4664 4665 4666 4667 4668 4669 4670
 4671 4672 4673 4674 4675 4676 4677 4678 4679 4680 4681 4682 4683 4684
 4685 4686 4687 4688 4689 4690 4691 4692 4693 4694 4695 4696 4697 4698
 4699 4700 4701 4702 4703 4704 4705 4706 4707 4708 4709 4710 4711 4712
 4713 4714 4715 4716 4717 4718 4719 4720 4721 4722 4723 4724 4725 4726
 4727 4728 4729 4730 4731 4732 4733 4734 4735 4736 4737 4738 4739 4740
 4741 4742 4743 4744 4745 4746 4747 4748 4749 4750 4751 4752 4753 4754
 4755 4756 4757 4758 4759 4760 4761 4762 4763 4764 4765 4766 4767 4768
 4769 4770 4771 4772 4773 4774 4775 4776 4777 4778 4779 4780 4781 4782
 4783 4784 4785 4786 4787 4788 4789 4790 4791 4792 4793 4794 4795 4796
 4797 4798 4799 4800 4801 4802 4803 4804 4805 4806 4807 4808 4809 4810
 4811 4812 4813 4814 4815 4816 4817 4818 4819 4820 4821 4822 4823 4824
 4825 4826 4827 4828 4829 4830 4831 4832 4833 4834 4835 4836 4837 4838
 4839 4840 4841 4842 4843 4844 4845 4846 4847 4848 4849 4850 4851 4852
 4853 4854 4855 4856 4857 4858 4859 4860 4861 4862 4863 4864 4865 4866
 4867 4868 4869 4870 4871 4872 4873 4874 4875 4876 4877 4878 4879 4880
 4881 4882 4883 4884 4885 4886 4887 4888 4889 4890 4891 4892 4893 4894
 4895 4896 4897 4898 4899 4900 4901 4902 4903 4904 4905 4906 4907 4908
 4909 4910 4911 4912 4913 4914 4915 4916 4917 4918 4919 4920 4921 4922
 4923 4924 4925 4926 4927 4928 4929 4930 4931 4932 4933 4934 4935 4936
 4937 4938 4939 4940 4941 4942 4943 4944 4945 4946 4947 4948 4949 4950
 4951 4952 4953 4954 4955 4956 4957 4958 4959 4960 4961 4962 4963 4964
 4965 4966 4967 4968 4969 4970 4971 4972 4973 4974 4975 4976 4977 4978
 4979 4980 4981 4982 4983 4984 4985 4986 4987 4988 4989 4990 4991 4992
 4993 4994 4995 4996 4997 4998 4999]
```

In [19]:
```
#Checking the Structures of the datasets
print(X_tr.shape)
print(X_test.shape)
print(Y_tr.shape)
print(Y_test.shape)
```

```
(4167, 4)
(833, 4)
(4167,)
(833,)
```

In [20]:
```
# Splitting the Train dataset into Cross Validation & Train Datasets
for train_index, test_index in tscv.split(X_tr):
   print("TRAIN:", train_index, "TEST:", test_index)
   X_train, X_cv = X[train_index], X[test_index]
   Y_train, Y_cv = Y[train_index], Y[test_index]
```

```
TRAIN: [  0   1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17
  18  19  20  21  22  23  24  25  26  27  28  29  30  31  32  33  34  35
  36  37  38  39  40  41  42  43  44  45  46  47  48  49  50  51  52  53
  54  55  56  57  58  59  60  61  62  63  64  65  66  67  68  69  70  71
  72  73  74  75  76  77  78  79  80  81  82  83  84  85  86  87  88  89
  90  91  92  93  94  95  96  97  98  99 100 101 102 103 104 105 106 107
 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125
 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143
 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161
 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179
 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197
 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215
 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233
 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251
 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269
 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287
 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305
 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323
 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341
 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359
 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377
 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395
 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413
 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431
 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449
 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467
 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485
 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503
 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521
 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539
 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557
 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575
 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593
 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611
 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629
 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647
 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665
 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683
 684 685 686 687 688 689 690 691 692 693 694 695 696] TEST: [ 697  698  699  700  701  702  703  704  705
  706  707  708  709  710
  711  712  713  714  715  716  717  718  719  720  721  722  723  724
  725  726  727  728  729  730  731  732  733  734  735  736  737  738
  739  740  741  742  743  744  745  746  747  748  749  750  751  752
  753  754  755  756  757  758  759  760  761  762  763  764  765  766
  767  768  769  770  771  772  773  774  775  776  777  778  779  780
  781  782  783  784  785  786  787  788  789  790  791  792  793  794
  795  796  797  798  799  800  801  802  803  804  805  806  807  808
  809  810  811  812  813  814  815  816  817  818  819  820  821  822
  823  824  825  826  827  828  829  830  831  832  833  834  835  836
  837  838  839  840  841  842  843  844  845  846  847  848  849  850
  851  852  853  854  855  856  857  858  859  860  861  862  863  864
  865  866  867  868  869  870  871  872  873  874  875  876  877  878
  879  880  881  882  883  884  885  886  887  888  889  890  891  892
  893  894  895  896  897  898  899  900  901  902  903  904  905  906
  907  908  909  910  911  912  913  914  915  916  917  918  919  920
  921  922  923  924  925  926  927  928  929  930  931  932  933  934
  935  936  937  938  939  940  941  942  943  944  945  946  947  948
  949  950  951  952  953  954  955  956  957  958  959  960  961  962
  963  964  965  966  967  968  969  970  971  972  973  974  975  976
  977  978  979  980  981  982  983  984  985  986  987  988  989  990
  991  992  993  994  995  996  997  998  999 1000 1001 1002 1003 1004
 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018
 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032
 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046
 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060
 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074
 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088
 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102
 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116
 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128 1129 1130
 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142 1143 1144
 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158
 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172
 1173 1174 1175 1176 1177 1178 1179 1180 1181 1182 1183 1184 1185 1186
 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198 1199 1200
 1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212 1213 1214
 1215 1216 1217 1218 1219 1220 1221 1222 1223 1224 1225 1226 1227 1228
 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240 1241 1242
 1243 1244 1245 1246 1247 1248 1249 1250 1251 1252 1253 1254 1255 1256
 1257 1258 1259 1260 1261 1262 1263 1264 1265 1266 1267 1268 1269 1270
 1271 1272 1273 1274 1275 1276 1277 1278 1279 1280 1281 1282 1283 1284
 1285 1286 1287 1288 1289 1290 1291 1292 1293 1294 1295 1296 1297 1298
 1299 1300 1301 1302 1303 1304 1305 1306 1307 1308 1309 1310 1311 1312
 1313 1314 1315 1316 1317 1318 1319 1320 1321 1322 1323 1324 1325 1326
 1327 1328 1329 1330 1331 1332 1333 1334 1335 1336 1337 1338 1339 1340
 1341 1342 1343 1344 1345 1346 1347 1348 1349 1350 1351 1352 1353 1354
 1355 1356 1357 1358 1359 1360 1361 1362 1363 1364 1365 1366 1367 1368
```

```
 1369 1370 1371 1372 1373 1374 1375 1376 1377 1378 1379 1380 1381 1382
 1383 1384 1385 1386 1387 1388 1389 1390]
TRAIN: [   0    1    2 ... 1388 1389 1390] TEST: [1391 1392 1393 1394 1395 1396 1397 1398 1399 1400 1401
1402 1403 1404
 1405 1406 1407 1408 1409 1410 1411 1412 1413 1414 1415 1416 1417 1418
 1419 1420 1421 1422 1423 1424 1425 1426 1427 1428 1429 1430 1431 1432
 1433 1434 1435 1436 1437 1438 1439 1440 1441 1442 1443 1444 1445 1446
 1447 1448 1449 1450 1451 1452 1453 1454 1455 1456 1457 1458 1459 1460
 1461 1462 1463 1464 1465 1466 1467 1468 1469 1470 1471 1472 1473 1474
 1475 1476 1477 1478 1479 1480 1481 1482 1483 1484 1485 1486 1487 1488
 1489 1490 1491 1492 1493 1494 1495 1496 1497 1498 1499 1500 1501 1502
 1503 1504 1505 1506 1507 1508 1509 1510 1511 1512 1513 1514 1515 1516
 1517 1518 1519 1520 1521 1522 1523 1524 1525 1526 1527 1528 1529 1530
 1531 1532 1533 1534 1535 1536 1537 1538 1539 1540 1541 1542 1543 1544
 1545 1546 1547 1548 1549 1550 1551 1552 1553 1554 1555 1556 1557 1558
 1559 1560 1561 1562 1563 1564 1565 1566 1567 1568 1569 1570 1571 1572
 1573 1574 1575 1576 1577 1578 1579 1580 1581 1582 1583 1584 1585 1586
 1587 1588 1589 1590 1591 1592 1593 1594 1595 1596 1597 1598 1599 1600
 1601 1602 1603 1604 1605 1606 1607 1608 1609 1610 1611 1612 1613 1614
 1615 1616 1617 1618 1619 1620 1621 1622 1623 1624 1625 1626 1627 1628
 1629 1630 1631 1632 1633 1634 1635 1636 1637 1638 1639 1640 1641 1642
 1643 1644 1645 1646 1647 1648 1649 1650 1651 1652 1653 1654 1655 1656
 1657 1658 1659 1660 1661 1662 1663 1664 1665 1666 1667 1668 1669 1670
 1671 1672 1673 1674 1675 1676 1677 1678 1679 1680 1681 1682 1683 1684
 1685 1686 1687 1688 1689 1690 1691 1692 1693 1694 1695 1696 1697 1698
 1699 1700 1701 1702 1703 1704 1705 1706 1707 1708 1709 1710 1711 1712
 1713 1714 1715 1716 1717 1718 1719 1720 1721 1722 1723 1724 1725 1726
 1727 1728 1729 1730 1731 1732 1733 1734 1735 1736 1737 1738 1739 1740
 1741 1742 1743 1744 1745 1746 1747 1748 1749 1750 1751 1752 1753 1754
 1755 1756 1757 1758 1759 1760 1761 1762 1763 1764 1765 1766 1767 1768
 1769 1770 1771 1772 1773 1774 1775 1776 1777 1778 1779 1780 1781 1782
 1783 1784 1785 1786 1787 1788 1789 1790 1791 1792 1793 1794 1795 1796
 1797 1798 1799 1800 1801 1802 1803 1804 1805 1806 1807 1808 1809 1810
 1811 1812 1813 1814 1815 1816 1817 1818 1819 1820 1821 1822 1823 1824
 1825 1826 1827 1828 1829 1830 1831 1832 1833 1834 1835 1836 1837 1838
 1839 1840 1841 1842 1843 1844 1845 1846 1847 1848 1849 1850 1851 1852
 1853 1854 1855 1856 1857 1858 1859 1860 1861 1862 1863 1864 1865 1866
 1867 1868 1869 1870 1871 1872 1873 1874 1875 1876 1877 1878 1879 1880
 1881 1882 1883 1884 1885 1886 1887 1888 1889 1890 1891 1892 1893 1894
 1895 1896 1897 1898 1899 1900 1901 1902 1903 1904 1905 1906 1907 1908
 1909 1910 1911 1912 1913 1914 1915 1916 1917 1918 1919 1920 1921 1922
 1923 1924 1925 1926 1927 1928 1929 1930 1931 1932 1933 1934 1935 1936
 1937 1938 1939 1940 1941 1942 1943 1944 1945 1946 1947 1948 1949 1950
 1951 1952 1953 1954 1955 1956 1957 1958 1959 1960 1961 1962 1963 1964
 1965 1966 1967 1968 1969 1970 1971 1972 1973 1974 1975 1976 1977 1978
 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991 1992
 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006
 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020
 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034
 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048
 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062
 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076
 2077 2078 2079 2080 2081 2082 2083 2084]
TRAIN: [   0    1    2 ... 2082 2083 2084] TEST: [2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095
2096 2097 2098
 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112
 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126
 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140
 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154
 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168
 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182
 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196
 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210
 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224
 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238
 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252
 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266
 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280
 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294
 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308
 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322
 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336
 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350
 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364
 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378
 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392
 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406
 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420
 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434
 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448
 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462
 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476
 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490
 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504
 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518
 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532
 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546
 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560
```

```
 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574
 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584 2585 2586 2587 2588
 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598 2599 2600 2601 2602
 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612 2613 2614 2615 2616
 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626 2627 2628 2629 2630
 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640 2641 2642 2643 2644
 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654 2655 2656 2657 2658
 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668 2669 2670 2671 2672
 2673 2674 2675 2676 2677 2678 2679 2680 2681 2682 2683 2684 2685 2686
 2687 2688 2689 2690 2691 2692 2693 2694 2695 2696 2697 2698 2699 2700
 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710 2711 2712 2713 2714
 2715 2716 2717 2718 2719 2720 2721 2722 2723 2724 2725 2726 2727 2728
 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738 2739 2740 2741 2742
 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752 2753 2754 2755 2756
 2757 2758 2759 2760 2761 2762 2763 2764 2765 2766 2767 2768 2769 2770
 2771 2772 2773 2774 2775 2776 2777 2778]
TRAIN: [   0    1    2 ... 2776 2777 2778] TEST: [2779 2780 2781 2782 2783 2784 2785 2786 2787 2788 2789
2790 2791 2792
 2793 2794 2795 2796 2797 2798 2799 2800 2801 2802 2803 2804 2805 2806
 2807 2808 2809 2810 2811 2812 2813 2814 2815 2816 2817 2818 2819 2820
 2821 2822 2823 2824 2825 2826 2827 2828 2829 2830 2831 2832 2833 2834
 2835 2836 2837 2838 2839 2840 2841 2842 2843 2844 2845 2846 2847 2848
 2849 2850 2851 2852 2853 2854 2855 2856 2857 2858 2859 2860 2861 2862
 2863 2864 2865 2866 2867 2868 2869 2870 2871 2872 2873 2874 2875 2876
 2877 2878 2879 2880 2881 2882 2883 2884 2885 2886 2887 2888 2889 2890
 2891 2892 2893 2894 2895 2896 2897 2898 2899 2900 2901 2902 2903 2904
 2905 2906 2907 2908 2909 2910 2911 2912 2913 2914 2915 2916 2917 2918
 2919 2920 2921 2922 2923 2924 2925 2926 2927 2928 2929 2930 2931 2932
 2933 2934 2935 2936 2937 2938 2939 2940 2941 2942 2943 2944 2945 2946
 2947 2948 2949 2950 2951 2952 2953 2954 2955 2956 2957 2958 2959 2960
 2961 2962 2963 2964 2965 2966 2967 2968 2969 2970 2971 2972 2973 2974
 2975 2976 2977 2978 2979 2980 2981 2982 2983 2984 2985 2986 2987 2988
 2989 2990 2991 2992 2993 2994 2995 2996 2997 2998 2999 3000 3001 3002
 3003 3004 3005 3006 3007 3008 3009 3010 3011 3012 3013 3014 3015 3016
 3017 3018 3019 3020 3021 3022 3023 3024 3025 3026 3027 3028 3029 3030
 3031 3032 3033 3034 3035 3036 3037 3038 3039 3040 3041 3042 3043 3044
 3045 3046 3047 3048 3049 3050 3051 3052 3053 3054 3055 3056 3057 3058
 3059 3060 3061 3062 3063 3064 3065 3066 3067 3068 3069 3070 3071 3072
 3073 3074 3075 3076 3077 3078 3079 3080 3081 3082 3083 3084 3085 3086
 3087 3088 3089 3090 3091 3092 3093 3094 3095 3096 3097 3098 3099 3100
 3101 3102 3103 3104 3105 3106 3107 3108 3109 3110 3111 3112 3113 3114
 3115 3116 3117 3118 3119 3120 3121 3122 3123 3124 3125 3126 3127 3128
 3129 3130 3131 3132 3133 3134 3135 3136 3137 3138 3139 3140 3141 3142
 3143 3144 3145 3146 3147 3148 3149 3150 3151 3152 3153 3154 3155 3156
 3157 3158 3159 3160 3161 3162 3163 3164 3165 3166 3167 3168 3169 3170
 3171 3172 3173 3174 3175 3176 3177 3178 3179 3180 3181 3182 3183 3184
 3185 3186 3187 3188 3189 3190 3191 3192 3193 3194 3195 3196 3197 3198
 3199 3200 3201 3202 3203 3204 3205 3206 3207 3208 3209 3210 3211 3212
 3213 3214 3215 3216 3217 3218 3219 3220 3221 3222 3223 3224 3225 3226
 3227 3228 3229 3230 3231 3232 3233 3234 3235 3236 3237 3238 3239 3240
 3241 3242 3243 3244 3245 3246 3247 3248 3249 3250 3251 3252 3253 3254
 3255 3256 3257 3258 3259 3260 3261 3262 3263 3264 3265 3266 3267 3268
 3269 3270 3271 3272 3273 3274 3275 3276 3277 3278 3279 3280 3281 3282
 3283 3284 3285 3286 3287 3288 3289 3290 3291 3292 3293 3294 3295 3296
 3297 3298 3299 3300 3301 3302 3303 3304 3305 3306 3307 3308 3309 3310
 3311 3312 3313 3314 3315 3316 3317 3318 3319 3320 3321 3322 3323 3324
 3325 3326 3327 3328 3329 3330 3331 3332 3333 3334 3335 3336 3337 3338
 3339 3340 3341 3342 3343 3344 3345 3346 3347 3348 3349 3350 3351 3352
 3353 3354 3355 3356 3357 3358 3359 3360 3361 3362 3363 3364 3365 3366
 3367 3368 3369 3370 3371 3372 3373 3374 3375 3376 3377 3378 3379 3380
 3381 3382 3383 3384 3385 3386 3387 3388 3389 3390 3391 3392 3393 3394
 3395 3396 3397 3398 3399 3400 3401 3402 3403 3404 3405 3406 3407 3408
 3409 3410 3411 3412 3413 3414 3415 3416 3417 3418 3419 3420 3421 3422
 3423 3424 3425 3426 3427 3428 3429 3430 3431 3432 3433 3434 3435 3436
 3437 3438 3439 3440 3441 3442 3443 3444 3445 3446 3447 3448 3449 3450
 3451 3452 3453 3454 3455 3456 3457 3458 3459 3460 3461 3462 3463 3464
 3465 3466 3467 3468 3469 3470 3471 3472]
TRAIN: [   0    1    2 ... 3470 3471 3472] TEST: [3473 3474 3475 3476 3477 3478 3479 3480 3481 3482 3483
3484 3485 3486
 3487 3488 3489 3490 3491 3492 3493 3494 3495 3496 3497 3498 3499 3500
 3501 3502 3503 3504 3505 3506 3507 3508 3509 3510 3511 3512 3513 3514
 3515 3516 3517 3518 3519 3520 3521 3522 3523 3524 3525 3526 3527 3528
 3529 3530 3531 3532 3533 3534 3535 3536 3537 3538 3539 3540 3541 3542
 3543 3544 3545 3546 3547 3548 3549 3550 3551 3552 3553 3554 3555 3556
 3557 3558 3559 3560 3561 3562 3563 3564 3565 3566 3567 3568 3569 3570
 3571 3572 3573 3574 3575 3576 3577 3578 3579 3580 3581 3582 3583 3584
 3585 3586 3587 3588 3589 3590 3591 3592 3593 3594 3595 3596 3597 3598
 3599 3600 3601 3602 3603 3604 3605 3606 3607 3608 3609 3610 3611 3612
 3613 3614 3615 3616 3617 3618 3619 3620 3621 3622 3623 3624 3625 3626
 3627 3628 3629 3630 3631 3632 3633 3634 3635 3636 3637 3638 3639 3640
 3641 3642 3643 3644 3645 3646 3647 3648 3649 3650 3651 3652 3653 3654
 3655 3656 3657 3658 3659 3660 3661 3662 3663 3664 3665 3666 3667 3668
 3669 3670 3671 3672 3673 3674 3675 3676 3677 3678 3679 3680 3681 3682
 3683 3684 3685 3686 3687 3688 3689 3690 3691 3692 3693 3694 3695 3696
 3697 3698 3699 3700 3701 3702 3703 3704 3705 3706 3707 3708 3709 3710
 3711 3712 3713 3714 3715 3716 3717 3718 3719 3720 3721 3722 3723 3724
 3725 3726 3727 3728 3729 3730 3731 3732 3733 3734 3735 3736 3737 3738
 3739 3740 3741 3742 3743 3744 3745 3746 3747 3748 3749 3750 3751 3752
```

```
 3753 3754 3755 3756 3757 3758 3759 3760 3761 3762 3763 3764 3765 3766
 3767 3768 3769 3770 3771 3772 3773 3774 3775 3776 3777 3778 3779 3780
 3781 3782 3783 3784 3785 3786 3787 3788 3789 3790 3791 3792 3793 3794
 3795 3796 3797 3798 3799 3800 3801 3802 3803 3804 3805 3806 3807 3808
 3809 3810 3811 3812 3813 3814 3815 3816 3817 3818 3819 3820 3821 3822
 3823 3824 3825 3826 3827 3828 3829 3830 3831 3832 3833 3834 3835 3836
 3837 3838 3839 3840 3841 3842 3843 3844 3845 3846 3847 3848 3849 3850
 3851 3852 3853 3854 3855 3856 3857 3858 3859 3860 3861 3862 3863 3864
 3865 3866 3867 3868 3869 3870 3871 3872 3873 3874 3875 3876 3877 3878
 3879 3880 3881 3882 3883 3884 3885 3886 3887 3888 3889 3890 3891 3892
 3893 3894 3895 3896 3897 3898 3899 3900 3901 3902 3903 3904 3905 3906
 3907 3908 3909 3910 3911 3912 3913 3914 3915 3916 3917 3918 3919 3920
 3921 3922 3923 3924 3925 3926 3927 3928 3929 3930 3931 3932 3933 3934
 3935 3936 3937 3938 3939 3940 3941 3942 3943 3944 3945 3946 3947 3948
 3949 3950 3951 3952 3953 3954 3955 3956 3957 3958 3959 3960 3961 3962
 3963 3964 3965 3966 3967 3968 3969 3970 3971 3972 3973 3974 3975 3976
 3977 3978 3979 3980 3981 3982 3983 3984 3985 3986 3987 3988 3989 3990
 3991 3992 3993 3994 3995 3996 3997 3998 3999 4000 4001 4002 4003 4004
 4005 4006 4007 4008 4009 4010 4011 4012 4013 4014 4015 4016 4017 4018
 4019 4020 4021 4022 4023 4024 4025 4026 4027 4028 4029 4030 4031 4032
 4033 4034 4035 4036 4037 4038 4039 4040 4041 4042 4043 4044 4045 4046
 4047 4048 4049 4050 4051 4052 4053 4054 4055 4056 4057 4058 4059 4060
 4061 4062 4063 4064 4065 4066 4067 4068 4069 4070 4071 4072 4073 4074
 4075 4076 4077 4078 4079 4080 4081 4082 4083 4084 4085 4086 4087 4088
 4089 4090 4091 4092 4093 4094 4095 4096 4097 4098 4099 4100 4101 4102
 4103 4104 4105 4106 4107 4108 4109 4110 4111 4112 4113 4114 4115 4116
 4117 4118 4119 4120 4121 4122 4123 4124 4125 4126 4127 4128 4129 4130
 4131 4132 4133 4134 4135 4136 4137 4138 4139 4140 4141 4142 4143 4144
 4145 4146 4147 4148 4149 4150 4151 4152 4153 4154 4155 4156 4157 4158
 4159 4160 4161 4162 4163 4164 4165 4166]
```

In [21]:
```
print(X_train.shape)
print(X_cv.shape)
print(X_test.shape)

print(Y_train.shape)
print(Y_cv.shape)
print(Y_test.shape)
```

```
(3473, 4)
(694, 4)
(833, 4)
(3473,)
(694,)
(833,)
```

In [22]:
```
# Removing the Date variable from the Data sets
new_X= []
for each in X_train:
    new_X.append(np.delete(each,3))
X_train = np.asarray(new_X)


new_X_cv= []
for each in X_cv:
    new_X_cv.append(np.delete(each,3))
X_cv = np.asarray(new_X_cv)

new_X_test= []
for each in X_test:
    new_X_test.append(np.delete(each,3))
X_test = np.asarray(new_X_test)
```

In [23]:
```
print(X_train.shape)
print(X_cv.shape)
print(X_test.shape)

print(Y_train.shape)
print(Y_cv.shape)
print(Y_test.shape)
```

```
(3473, 3)
(694, 3)
(833, 3)
(3473,)
(694,)
(833,)
```

In [24]:
```
# Finding the Test accuracy using Cross Validation Data Set
for i in range(1,30,2):
    # instantiate learning model (k = 30)
    knn = KNeighborsClassifier(n_neighbors=i)

    # fitting the model on crossvalidation train
    knn.fit(X_train, Y_train)

    # predict the response on the crossvalidation train
    pred = knn.predict(X_cv)

    # evaluate CV accuracy
    acc = accuracy_score(Y_cv, pred, normalize=True) * float(100)
    print('\nCV accuracy for k = %d is %d%%' % (i, acc))

knn = KNeighborsClassifier(1)
knn.fit(X_train,Y_train)
pred = knn.predict(X_test)
acc = accuracy_score(Y_test, pred, normalize=True) * float(100)
print('\n****Test accuracy for k = 1 is %d%%' % (acc))
```

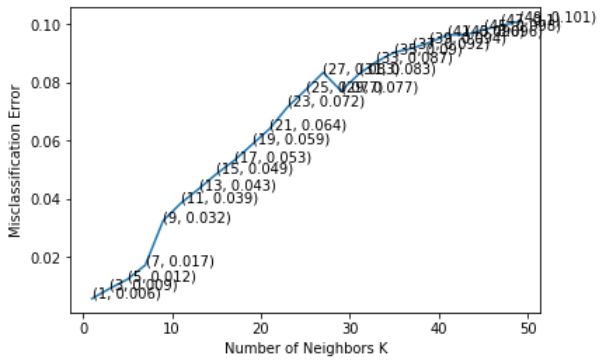
```
CV accuracy for k = 1 is 98%

CV accuracy for k = 3 is 98%

CV accuracy for k = 5 is 97%

CV accuracy for k = 7 is 96%

CV accuracy for k = 9 is 96%

CV accuracy for k = 11 is 96%

CV accuracy for k = 13 is 96%

CV accuracy for k = 15 is 95%

CV accuracy for k = 17 is 95%

CV accuracy for k = 19 is 94%

CV accuracy for k = 21 is 94%

CV accuracy for k = 23 is 93%

CV accuracy for k = 25 is 92%

CV accuracy for k = 27 is 89%

CV accuracy for k = 29 is 88%

****Test accuracy for k = 1 is 99%
```

In [25]:
```
# creating odd list of K for KNN
myList = list(range(0,50))
neighbors = list(filter(lambda x: x % 2 != 0, myList))

# empty list that will hold cv scores
cv_scores = []
```

In [27]:
```
for k in neighbors:
    knn = KNeighborsClassifier(n_neighbors=k)
    scores = cross_val_score(knn, X_train, Y_train, cv=10)
    cv_scores.append(scores.mean())
```

In [28]:
```
# changing to misclassification error
MSE = [1 - x for x in cv_scores]

# determining best k
optimal_k = neighbors[MSE.index(min(MSE))]
print('\nThe optimal number of neighbors is %d.' % optimal_k)

# plot misclassification error vs k
plt.plot(neighbors, MSE)

for xy in zip(neighbors, np.round(MSE,3)):
    plt.annotate('(%s, %s)' % xy, xy=xy, textcoords='data')

plt.xlabel('Number of Neighbors K')
plt.ylabel('Misclassification Error')
plt.show()

print("the misclassification error for each k value is : ", np.round(MSE,3))
```

```
The optimal number of neighbors is 1.
```




```
the misclassification error for each k value is :  [0.006 0.009 0.012 0.017 0.032 0.039 0.043 0.049 0.053
0.059 0.064 0.072
 0.077 0.083 0.077 0.083 0.087 0.09  0.092 0.094 0.096 0.096 0.098 0.1
 0.101]
```

In [29]:
```
# KNN with k = optimal_k

# instantiate learning model k = optimal_k
knn_optimal = KNeighborsClassifier(n_neighbors=optimal_k)

# fitting the model
knn_optimal.fit(X_train, Y_train)

# predict the response
pred = knn_optimal.predict(X_test)

# evaluate accuracy
acc = accuracy_score(Y_test, pred) * 100
print('\nThe accuracy of the knn classifier for k = %d is %f%%' % (optimal_k, acc))
```

```
The accuracy of the knn classifier for k = 1 is 99.639856%
```

Term Frequency– Inverse Document Frequency (TF-IDF)

In [165]:
```
# Configuring ngram range from unigram to bigram
tf_idf_vect = TfidfVectorizer(ngram_range=(1,2))
final_tf_idf = tf_idf_vect.fit_transform(final_dataset['CleanedText'].values)
print("the type of count vectorizer ",type(final_tf_idf))
print("the shape of out text TFIDF vectorizer ",final_tf_idf.get_shape())
print("the number of unique words including both unigrams and bigrams ", final_tf_idf.get_shape()[1])
```

```
the type of count vectorizer  <class 'scipy.sparse.csr.csr_matrix'>
the shape of out text TFIDF vectorizer  (5000, 151352)
the number of unique words including both unigrams and bigrams  151352
```

In [166]:
```
# Converting Sparse matric to array
final_tf_idf=final_tf_idf.toarray()
#tsne
start_time_code = time.time()
model = TSNE(n_components=2, random_state=0, perplexity=30)
tsne_data_tfidf = model.fit_transform(final_tf_idf)
end_time_code = time.time()
print ("Running Time for code execution " + str(end_time_code - start_time_code) + " secs")
```

```
Running Time for code execution 6420.043386220932 secs
```

In [167]:
```
tsne_data_tf = np.vstack((tsne_data_tfidf.T,label,Time)).T
```

In [168]:
```
#Saving tsne data into seperate file for further usecases.
tsne_tfidf_df = pd.DataFrame(data=tsne_data_tf, columns=("X", "Y", "Label","Date"))
tsne_tfidf_df.to_csv('/home/venkatasujit272/tsne_tfidf_data.csv')
```

In [169]:
```
# Aliasing the Score field values into Binary format
def partition(x):
    if x == 'negative':
        return 0.
    return 1.

# Aliasing the Score to String Format (Positive for Score > 3 & Negative for Score  < 3)

actualScore = tsne_tfidf_df['Label']
positiveNegative = actualScore.map(partition)
tsne_tfidf_df['Label'] = positiveNegative
```

In [170]:
```
tsne_tfidf_df.head(1)
```

Out[170]:

	X	Y	Label	Date
0	0.633459	-8.7365	1.0	1999-10-08

In [171]:
```
#Sorting the Dataframe with Date
tsne_tfidf_df.sort_values(by='Date',ascending=1)
```

Out[171]:

	X	Y	Label	Date
0	0.633459	-8.7365	1.0	1999-10-08
28	0.621319	-8.71731	1.0	1999-10-25
377	-1.79879	-10.9841	1.0	1999-12-02
287	-2.03936	-12.3553	1.0	1999-12-06
376	-1.86981	-11.1863	1.0	2000-01-03
209	-2.02703	-10.426	1.0	2000-01-09
2540	-1.19717	-10.54	0.0	2000-01-19
198	4.47318	-11.9677	1.0	2000-01-24
224	-2.31892	-12.8517	1.0	2000-02-26
284	0.119211	-10.6952	1.0	2000-06-03
378	0.132651	-10.7101	1.0	2000-06-03
205	0.971696	-7.40803	1.0	2000-06-23
206	1.12162	-7.44633	1.0	2000-06-29
415	10.4458	-7.40248	1.0	2000-07-31
684	0.929265	7.57625	1.0	2000-08-09
714	0.807889	7.65295	1.0	2000-08-09
212	-1.44218	-12.0895	1.0	2000-08-15
258	-1.72246	-12.3186	1.0	2000-10-03
691	0.6374	7.7362	1.0	2000-12-05
315	-2.57085	-11.4867	1.0	2000-12-19
286	-1.88425	-11.6942	1.0	2000-12-30
692	1.51776	7.20513	1.0	2001-02-22
2547	-2.17737	-11.5179	0.0	2001-06-11
2543	-1.21317	-11.2717	0.0	2001-08-08
233	-8.75209	-9.4265	1.0	2001-09-24
701	1.25664	7.48027	1.0	2001-10-23
309	-2.23182	-12.4742	1.0	2001-10-26
30	0.618278	-8.68635	1.0	2001-12-26
830	3.22437	10.3938	1.0	2002-01-06
288	-1.77231	-12.3715	1.0	2002-02-04
...	...	...	...	...
232	-2.24862	-10.3694	1.0	2012-10-22
3856	-1.18319	-2.79451	0.0	2012-10-23
1792	-7.34889	5.06849	1.0	2012-10-23
813	-8.16596	1.03918	1.0	2012-10-23
3649	3.53231	0.34179	0.0	2012-10-23
2030	5.86201	9.161	1.0	2012-10-23
1897	0.997734	4.99462	1.0	2012-10-23
1086	-2.20842	8.7914	1.0	2012-10-23
3056	3.91186	-4.51015	0.0	2012-10-23
1594	-5.56258	6.03292	1.0	2012-10-23
4556	5.0185	1.64046	0.0	2012-10-23
134	-5.32573	7.34865	1.0	2012-10-23
2717	-2.51862	8.37566	0.0	2012-10-24
2705	-3.72909	-1.20997	0.0	2012-10-24
1295	-4.48909	-1.0134	1.0	2012-10-24
2391	1.28051	7.21648	1.0	2012-10-24
1603	-12.9058	0.683586	1.0	2012-10-25
3849	-1.05837	2.06191	0.0	2012-10-25
2715	-6.49729	1.41066	0.0	2012-10-25
4115	1.82492	-7.29996	0.0	2012-10-25
1827	-6.37683	2.24747	1.0	2012-10-25
2978	12.1615	3.03545	0.0	2012-10-26

	X	Y	Label	Date
135	6.06442	-10.2238	1.0	2012-10-26
782	-5.81879	1.55783	1.0	2012-10-26
136	-6.23531	7.0223	1.0	2012-10-26
3569	12.9654	1.75552	0.0	2012-10-26
1349	-12.5389	0.879432	1.0	2012-10-26
493	11.2503	-9.80318	1.0	2012-10-26
4793	-2.08286	9.36516	0.0	2012-10-26
2592	10.135	-8.41314	0.0	2012-10-26

5000 rows × 4 columns

In [172]:
```
# create design matrix X and target vector y
X_tfidf = np.array(tsne_tfidf_df.iloc[:, 0:4]) # end index is exclusive
Y_tfidf = np.array(tsne_tfidf_df['Label']) # showing you two ways of indexing a pandas df
```

In [173]:
```
 # Timebasedf Split
tscv = TimeSeriesSplit(n_splits=5)
print(tscv)
```

```
TimeSeriesSplit(n_splits=5)
```

In [174]:
```
# Splitting the Dataset into Train & Test Datasets
for train_index, test_index in tscv.split(X_tfidf):
   print("TRAIN:", train_index, "TEST:", test_index)
   X_tfidf_tr, X_tfidf_test = X[train_index], X[test_index]
   Y_tfidf_tr, Y_tfidf_test = Y[train_index], Y[test_index]
```

```
TRAIN: [  0   1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17
  18  19  20  21  22  23  24  25  26  27  28  29  30  31  32  33  34  35
  36  37  38  39  40  41  42  43  44  45  46  47  48  49  50  51  52  53
  54  55  56  57  58  59  60  61  62  63  64  65  66  67  68  69  70  71
  72  73  74  75  76  77  78  79  80  81  82  83  84  85  86  87  88  89
  90  91  92  93  94  95  96  97  98  99 100 101 102 103 104 105 106 107
 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125
 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143
 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161
 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179
 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197
 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215
 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233
 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251
 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269
 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287
 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305
 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323
 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341
 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359
 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377
 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395
 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413
 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431
 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449
 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467
 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485
 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503
 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521
 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539
 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557
 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575
 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593
 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611
 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629
 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647
 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665
 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683
 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701
 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719
 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737
 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755
 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773
 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791
 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809
 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827
 828 829 830 831 832 833 834] TEST: [ 835  836  837  838  839  840  841  842  843  844  845  846  847  84
8
  849  850  851  852  853  854  855  856  857  858  859  860  861  862
  863  864  865  866  867  868  869  870  871  872  873  874  875  876
  877  878  879  880  881  882  883  884  885  886  887  888  889  890
  891  892  893  894  895  896  897  898  899  900  901  902  903  904
  905  906  907  908  909  910  911  912  913  914  915  916  917  918
  919  920  921  922  923  924  925  926  927  928  929  930  931  932
  933  934  935  936  937  938  939  940  941  942  943  944  945  946
  947  948  949  950  951  952  953  954  955  956  957  958  959  960
  961  962  963  964  965  966  967  968  969  970  971  972  973  974
  975  976  977  978  979  980  981  982  983  984  985  986  987  988
  989  990  991  992  993  994  995  996  997  998  999 1000 1001 1002
 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016
 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030
 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044
 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058
 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072
 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086
 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100
 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114
 1115 1116 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128
 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142
 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156
 1157 1158 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170
 1171 1172 1173 1174 1175 1176 1177 1178 1179 1180 1181 1182 1183 1184
 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198
 1199 1200 1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212
 1213 1214 1215 1216 1217 1218 1219 1220 1221 1222 1223 1224 1225 1226
 1227 1228 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240
 1241 1242 1243 1244 1245 1246 1247 1248 1249 1250 1251 1252 1253 1254
 1255 1256 1257 1258 1259 1260 1261 1262 1263 1264 1265 1266 1267 1268
 1269 1270 1271 1272 1273 1274 1275 1276 1277 1278 1279 1280 1281 1282
 1283 1284 1285 1286 1287 1288 1289 1290 1291 1292 1293 1294 1295 1296
 1297 1298 1299 1300 1301 1302 1303 1304 1305 1306 1307 1308 1309 1310
 1311 1312 1313 1314 1315 1316 1317 1318 1319 1320 1321 1322 1323 1324
 1325 1326 1327 1328 1329 1330 1331 1332 1333 1334 1335 1336 1337 1338
 1339 1340 1341 1342 1343 1344 1345 1346 1347 1348 1349 1350 1351 1352
 1353 1354 1355 1356 1357 1358 1359 1360 1361 1362 1363 1364 1365 1366
 1367 1368 1369 1370 1371 1372 1373 1374 1375 1376 1377 1378 1379 1380
 1381 1382 1383 1384 1385 1386 1387 1388 1389 1390 1391 1392 1393 1394
```

```
 1395 1396 1397 1398 1399 1400 1401 1402 1403 1404 1405 1406 1407 1408
 1409 1410 1411 1412 1413 1414 1415 1416 1417 1418 1419 1420 1421 1422
 1423 1424 1425 1426 1427 1428 1429 1430 1431 1432 1433 1434 1435 1436
 1437 1438 1439 1440 1441 1442 1443 1444 1445 1446 1447 1448 1449 1450
 1451 1452 1453 1454 1455 1456 1457 1458 1459 1460 1461 1462 1463 1464
 1465 1466 1467 1468 1469 1470 1471 1472 1473 1474 1475 1476 1477 1478
 1479 1480 1481 1482 1483 1484 1485 1486 1487 1488 1489 1490 1491 1492
 1493 1494 1495 1496 1497 1498 1499 1500 1501 1502 1503 1504 1505 1506
 1507 1508 1509 1510 1511 1512 1513 1514 1515 1516 1517 1518 1519 1520
 1521 1522 1523 1524 1525 1526 1527 1528 1529 1530 1531 1532 1533 1534
 1535 1536 1537 1538 1539 1540 1541 1542 1543 1544 1545 1546 1547 1548
 1549 1550 1551 1552 1553 1554 1555 1556 1557 1558 1559 1560 1561 1562
 1563 1564 1565 1566 1567 1568 1569 1570 1571 1572 1573 1574 1575 1576
 1577 1578 1579 1580 1581 1582 1583 1584 1585 1586 1587 1588 1589 1590
 1591 1592 1593 1594 1595 1596 1597 1598 1599 1600 1601 1602 1603 1604
 1605 1606 1607 1608 1609 1610 1611 1612 1613 1614 1615 1616 1617 1618
 1619 1620 1621 1622 1623 1624 1625 1626 1627 1628 1629 1630 1631 1632
 1633 1634 1635 1636 1637 1638 1639 1640 1641 1642 1643 1644 1645 1646
 1647 1648 1649 1650 1651 1652 1653 1654 1655 1656 1657 1658 1659 1660
 1661 1662 1663 1664 1665 1666 1667]
TRAIN: [   0    1    2 ... 1665 1666 1667] TEST: [1668 1669 1670 1671 1672 1673 1674 1675 1676 1677 1678
1679 1680 1681
 1682 1683 1684 1685 1686 1687 1688 1689 1690 1691 1692 1693 1694 1695
 1696 1697 1698 1699 1700 1701 1702 1703 1704 1705 1706 1707 1708 1709
 1710 1711 1712 1713 1714 1715 1716 1717 1718 1719 1720 1721 1722 1723
 1724 1725 1726 1727 1728 1729 1730 1731 1732 1733 1734 1735 1736 1737
 1738 1739 1740 1741 1742 1743 1744 1745 1746 1747 1748 1749 1750 1751
 1752 1753 1754 1755 1756 1757 1758 1759 1760 1761 1762 1763 1764 1765
 1766 1767 1768 1769 1770 1771 1772 1773 1774 1775 1776 1777 1778 1779
 1780 1781 1782 1783 1784 1785 1786 1787 1788 1789 1790 1791 1792 1793
 1794 1795 1796 1797 1798 1799 1800 1801 1802 1803 1804 1805 1806 1807
 1808 1809 1810 1811 1812 1813 1814 1815 1816 1817 1818 1819 1820 1821
 1822 1823 1824 1825 1826 1827 1828 1829 1830 1831 1832 1833 1834 1835
 1836 1837 1838 1839 1840 1841 1842 1843 1844 1845 1846 1847 1848 1849
 1850 1851 1852 1853 1854 1855 1856 1857 1858 1859 1860 1861 1862 1863
 1864 1865 1866 1867 1868 1869 1870 1871 1872 1873 1874 1875 1876 1877
 1878 1879 1880 1881 1882 1883 1884 1885 1886 1887 1888 1889 1890 1891
 1892 1893 1894 1895 1896 1897 1898 1899 1900 1901 1902 1903 1904 1905
 1906 1907 1908 1909 1910 1911 1912 1913 1914 1915 1916 1917 1918 1919
 1920 1921 1922 1923 1924 1925 1926 1927 1928 1929 1930 1931 1932 1933
 1934 1935 1936 1937 1938 1939 1940 1941 1942 1943 1944 1945 1946 1947
 1948 1949 1950 1951 1952 1953 1954 1955 1956 1957 1958 1959 1960 1961
 1962 1963 1964 1965 1966 1967 1968 1969 1970 1971 1972 1973 1974 1975
 1976 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989
 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003
 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017
 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031
 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045
 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059
 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073
 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087
 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101
 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115
 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129
 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143
 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157
 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171
 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185
 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199
 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213
 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227
 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241
 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255
 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269
 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283
 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297
 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311
 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325
 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339
 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353
 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367
 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381
 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395
 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409
 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423
 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437
 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451
 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465
 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479
 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493
 2494 2495 2496 2497 2498 2499 2500]
TRAIN: [   0    1    2 ... 2498 2499 2500] TEST: [2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511
2512 2513 2514
 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528
 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542
 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556
 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570
 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584
```

```
 2585 2586 2587 2588 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598
 2599 2600 2601 2602 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612
 2613 2614 2615 2616 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626
 2627 2628 2629 2630 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640
 2641 2642 2643 2644 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654
 2655 2656 2657 2658 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668
 2669 2670 2671 2672 2673 2674 2675 2676 2677 2678 2679 2680 2681 2682
 2683 2684 2685 2686 2687 2688 2689 2690 2691 2692 2693 2694 2695 2696
 2697 2698 2699 2700 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710
 2711 2712 2713 2714 2715 2716 2717 2718 2719 2720 2721 2722 2723 2724
 2725 2726 2727 2728 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738
 2739 2740 2741 2742 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752
 2753 2754 2755 2756 2757 2758 2759 2760 2761 2762 2763 2764 2765 2766
 2767 2768 2769 2770 2771 2772 2773 2774 2775 2776 2777 2778 2779 2780
 2781 2782 2783 2784 2785 2786 2787 2788 2789 2790 2791 2792 2793 2794
 2795 2796 2797 2798 2799 2800 2801 2802 2803 2804 2805 2806 2807 2808
 2809 2810 2811 2812 2813 2814 2815 2816 2817 2818 2819 2820 2821 2822
 2823 2824 2825 2826 2827 2828 2829 2830 2831 2832 2833 2834 2835 2836
 2837 2838 2839 2840 2841 2842 2843 2844 2845 2846 2847 2848 2849 2850
 2851 2852 2853 2854 2855 2856 2857 2858 2859 2860 2861 2862 2863 2864
 2865 2866 2867 2868 2869 2870 2871 2872 2873 2874 2875 2876 2877 2878
 2879 2880 2881 2882 2883 2884 2885 2886 2887 2888 2889 2890 2891 2892
 2893 2894 2895 2896 2897 2898 2899 2900 2901 2902 2903 2904 2905 2906
 2907 2908 2909 2910 2911 2912 2913 2914 2915 2916 2917 2918 2919 2920
 2921 2922 2923 2924 2925 2926 2927 2928 2929 2930 2931 2932 2933 2934
 2935 2936 2937 2938 2939 2940 2941 2942 2943 2944 2945 2946 2947 2948
 2949 2950 2951 2952 2953 2954 2955 2956 2957 2958 2959 2960 2961 2962
 2963 2964 2965 2966 2967 2968 2969 2970 2971 2972 2973 2974 2975 2976
 2977 2978 2979 2980 2981 2982 2983 2984 2985 2986 2987 2988 2989 2990
 2991 2992 2993 2994 2995 2996 2997 2998 2999 3000 3001 3002 3003 3004
 3005 3006 3007 3008 3009 3010 3011 3012 3013 3014 3015 3016 3017 3018
 3019 3020 3021 3022 3023 3024 3025 3026 3027 3028 3029 3030 3031 3032
 3033 3034 3035 3036 3037 3038 3039 3040 3041 3042 3043 3044 3045 3046
 3047 3048 3049 3050 3051 3052 3053 3054 3055 3056 3057 3058 3059 3060
 3061 3062 3063 3064 3065 3066 3067 3068 3069 3070 3071 3072 3073 3074
 3075 3076 3077 3078 3079 3080 3081 3082 3083 3084 3085 3086 3087 3088
 3089 3090 3091 3092 3093 3094 3095 3096 3097 3098 3099 3100 3101 3102
 3103 3104 3105 3106 3107 3108 3109 3110 3111 3112 3113 3114 3115 3116
 3117 3118 3119 3120 3121 3122 3123 3124 3125 3126 3127 3128 3129 3130
 3131 3132 3133 3134 3135 3136 3137 3138 3139 3140 3141 3142 3143 3144
 3145 3146 3147 3148 3149 3150 3151 3152 3153 3154 3155 3156 3157 3158
 3159 3160 3161 3162 3163 3164 3165 3166 3167 3168 3169 3170 3171 3172
 3173 3174 3175 3176 3177 3178 3179 3180 3181 3182 3183 3184 3185 3186
 3187 3188 3189 3190 3191 3192 3193 3194 3195 3196 3197 3198 3199 3200
 3201 3202 3203 3204 3205 3206 3207 3208 3209 3210 3211 3212 3213 3214
 3215 3216 3217 3218 3219 3220 3221 3222 3223 3224 3225 3226 3227 3228
 3229 3230 3231 3232 3233 3234 3235 3236 3237 3238 3239 3240 3241 3242
 3243 3244 3245 3246 3247 3248 3249 3250 3251 3252 3253 3254 3255 3256
 3257 3258 3259 3260 3261 3262 3263 3264 3265 3266 3267 3268 3269 3270
 3271 3272 3273 3274 3275 3276 3277 3278 3279 3280 3281 3282 3283 3284
 3285 3286 3287 3288 3289 3290 3291 3292 3293 3294 3295 3296 3297 3298
 3299 3300 3301 3302 3303 3304 3305 3306 3307 3308 3309 3310 3311 3312
 3313 3314 3315 3316 3317 3318 3319 3320 3321 3322 3323 3324 3325 3326
 3327 3328 3329 3330 3331 3332 3333]
TRAIN: [   0    1    2 ... 3331 3332 3333] TEST: [3334 3335 3336 3337 3338 3339 3340 3341 3342 3343 3344
3345 3346 3347
 3348 3349 3350 3351 3352 3353 3354 3355 3356 3357 3358 3359 3360 3361
 3362 3363 3364 3365 3366 3367 3368 3369 3370 3371 3372 3373 3374 3375
 3376 3377 3378 3379 3380 3381 3382 3383 3384 3385 3386 3387 3388 3389
 3390 3391 3392 3393 3394 3395 3396 3397 3398 3399 3400 3401 3402 3403
 3404 3405 3406 3407 3408 3409 3410 3411 3412 3413 3414 3415 3416 3417
 3418 3419 3420 3421 3422 3423 3424 3425 3426 3427 3428 3429 3430 3431
 3432 3433 3434 3435 3436 3437 3438 3439 3440 3441 3442 3443 3444 3445
 3446 3447 3448 3449 3450 3451 3452 3453 3454 3455 3456 3457 3458 3459
 3460 3461 3462 3463 3464 3465 3466 3467 3468 3469 3470 3471 3472 3473
 3474 3475 3476 3477 3478 3479 3480 3481 3482 3483 3484 3485 3486 3487
 3488 3489 3490 3491 3492 3493 3494 3495 3496 3497 3498 3499 3500 3501
 3502 3503 3504 3505 3506 3507 3508 3509 3510 3511 3512 3513 3514 3515
 3516 3517 3518 3519 3520 3521 3522 3523 3524 3525 3526 3527 3528 3529
 3530 3531 3532 3533 3534 3535 3536 3537 3538 3539 3540 3541 3542 3543
 3544 3545 3546 3547 3548 3549 3550 3551 3552 3553 3554 3555 3556 3557
 3558 3559 3560 3561 3562 3563 3564 3565 3566 3567 3568 3569 3570 3571
 3572 3573 3574 3575 3576 3577 3578 3579 3580 3581 3582 3583 3584 3585
 3586 3587 3588 3589 3590 3591 3592 3593 3594 3595 3596 3597 3598 3599
 3600 3601 3602 3603 3604 3605 3606 3607 3608 3609 3610 3611 3612 3613
 3614 3615 3616 3617 3618 3619 3620 3621 3622 3623 3624 3625 3626 3627
 3628 3629 3630 3631 3632 3633 3634 3635 3636 3637 3638 3639 3640 3641
 3642 3643 3644 3645 3646 3647 3648 3649 3650 3651 3652 3653 3654 3655
 3656 3657 3658 3659 3660 3661 3662 3663 3664 3665 3666 3667 3668 3669
 3670 3671 3672 3673 3674 3675 3676 3677 3678 3679 3680 3681 3682 3683
 3684 3685 3686 3687 3688 3689 3690 3691 3692 3693 3694 3695 3696 3697
 3698 3699 3700 3701 3702 3703 3704 3705 3706 3707 3708 3709 3710 3711
 3712 3713 3714 3715 3716 3717 3718 3719 3720 3721 3722 3723 3724 3725
 3726 3727 3728 3729 3730 3731 3732 3733 3734 3735 3736 3737 3738 3739
 3740 3741 3742 3743 3744 3745 3746 3747 3748 3749 3750 3751 3752 3753
 3754 3755 3756 3757 3758 3759 3760 3761 3762 3763 3764 3765 3766 3767
 3768 3769 3770 3771 3772 3773 3774 3775 3776 3777 3778 3779 3780 3781
 3782 3783 3784 3785 3786 3787 3788 3789 3790 3791 3792 3793 3794 3795
```

```
 3796 3797 3798 3799 3800 3801 3802 3803 3804 3805 3806 3807 3808 3809
 3810 3811 3812 3813 3814 3815 3816 3817 3818 3819 3820 3821 3822 3823
 3824 3825 3826 3827 3828 3829 3830 3831 3832 3833 3834 3835 3836 3837
 3838 3839 3840 3841 3842 3843 3844 3845 3846 3847 3848 3849 3850 3851
 3852 3853 3854 3855 3856 3857 3858 3859 3860 3861 3862 3863 3864 3865
 3866 3867 3868 3869 3870 3871 3872 3873 3874 3875 3876 3877 3878 3879
 3880 3881 3882 3883 3884 3885 3886 3887 3888 3889 3890 3891 3892 3893
 3894 3895 3896 3897 3898 3899 3900 3901 3902 3903 3904 3905 3906 3907
 3908 3909 3910 3911 3912 3913 3914 3915 3916 3917 3918 3919 3920 3921
 3922 3923 3924 3925 3926 3927 3928 3929 3930 3931 3932 3933 3934 3935
 3936 3937 3938 3939 3940 3941 3942 3943 3944 3945 3946 3947 3948 3949
 3950 3951 3952 3953 3954 3955 3956 3957 3958 3959 3960 3961 3962 3963
 3964 3965 3966 3967 3968 3969 3970 3971 3972 3973 3974 3975 3976 3977
 3978 3979 3980 3981 3982 3983 3984 3985 3986 3987 3988 3989 3990 3991
 3992 3993 3994 3995 3996 3997 3998 3999 4000 4001 4002 4003 4004 4005
 4006 4007 4008 4009 4010 4011 4012 4013 4014 4015 4016 4017 4018 4019
 4020 4021 4022 4023 4024 4025 4026 4027 4028 4029 4030 4031 4032 4033
 4034 4035 4036 4037 4038 4039 4040 4041 4042 4043 4044 4045 4046 4047
 4048 4049 4050 4051 4052 4053 4054 4055 4056 4057 4058 4059 4060 4061
 4062 4063 4064 4065 4066 4067 4068 4069 4070 4071 4072 4073 4074 4075
 4076 4077 4078 4079 4080 4081 4082 4083 4084 4085 4086 4087 4088 4089
 4090 4091 4092 4093 4094 4095 4096 4097 4098 4099 4100 4101 4102 4103
 4104 4105 4106 4107 4108 4109 4110 4111 4112 4113 4114 4115 4116 4117
 4118 4119 4120 4121 4122 4123 4124 4125 4126 4127 4128 4129 4130 4131
 4132 4133 4134 4135 4136 4137 4138 4139 4140 4141 4142 4143 4144 4145
 4146 4147 4148 4149 4150 4151 4152 4153 4154 4155 4156 4157 4158 4159
 4160 4161 4162 4163 4164 4165 4166]
TRAIN: [   0    1    2 ... 4164 4165 4166] TEST: [4167 4168 4169 4170 4171 4172 4173 4174 4175 4176 4177
4178 4179 4180
 4181 4182 4183 4184 4185 4186 4187 4188 4189 4190 4191 4192 4193 4194
 4195 4196 4197 4198 4199 4200 4201 4202 4203 4204 4205 4206 4207 4208
 4209 4210 4211 4212 4213 4214 4215 4216 4217 4218 4219 4220 4221 4222
 4223 4224 4225 4226 4227 4228 4229 4230 4231 4232 4233 4234 4235 4236
 4237 4238 4239 4240 4241 4242 4243 4244 4245 4246 4247 4248 4249 4250
 4251 4252 4253 4254 4255 4256 4257 4258 4259 4260 4261 4262 4263 4264
 4265 4266 4267 4268 4269 4270 4271 4272 4273 4274 4275 4276 4277 4278
 4279 4280 4281 4282 4283 4284 4285 4286 4287 4288 4289 4290 4291 4292
 4293 4294 4295 4296 4297 4298 4299 4300 4301 4302 4303 4304 4305 4306
 4307 4308 4309 4310 4311 4312 4313 4314 4315 4316 4317 4318 4319 4320
 4321 4322 4323 4324 4325 4326 4327 4328 4329 4330 4331 4332 4333 4334
 4335 4336 4337 4338 4339 4340 4341 4342 4343 4344 4345 4346 4347 4348
 4349 4350 4351 4352 4353 4354 4355 4356 4357 4358 4359 4360 4361 4362
 4363 4364 4365 4366 4367 4368 4369 4370 4371 4372 4373 4374 4375 4376
 4377 4378 4379 4380 4381 4382 4383 4384 4385 4386 4387 4388 4389 4390
 4391 4392 4393 4394 4395 4396 4397 4398 4399 4400 4401 4402 4403 4404
 4405 4406 4407 4408 4409 4410 4411 4412 4413 4414 4415 4416 4417 4418
 4419 4420 4421 4422 4423 4424 4425 4426 4427 4428 4429 4430 4431 4432
 4433 4434 4435 4436 4437 4438 4439 4440 4441 4442 4443 4444 4445 4446
 4447 4448 4449 4450 4451 4452 4453 4454 4455 4456 4457 4458 4459 4460
 4461 4462 4463 4464 4465 4466 4467 4468 4469 4470 4471 4472 4473 4474
 4475 4476 4477 4478 4479 4480 4481 4482 4483 4484 4485 4486 4487 4488
 4489 4490 4491 4492 4493 4494 4495 4496 4497 4498 4499 4500 4501 4502
 4503 4504 4505 4506 4507 4508 4509 4510 4511 4512 4513 4514 4515 4516
 4517 4518 4519 4520 4521 4522 4523 4524 4525 4526 4527 4528 4529 4530
 4531 4532 4533 4534 4535 4536 4537 4538 4539 4540 4541 4542 4543 4544
 4545 4546 4547 4548 4549 4550 4551 4552 4553 4554 4555 4556 4557 4558
 4559 4560 4561 4562 4563 4564 4565 4566 4567 4568 4569 4570 4571 4572
 4573 4574 4575 4576 4577 4578 4579 4580 4581 4582 4583 4584 4585 4586
 4587 4588 4589 4590 4591 4592 4593 4594 4595 4596 4597 4598 4599 4600
 4601 4602 4603 4604 4605 4606 4607 4608 4609 4610 4611 4612 4613 4614
 4615 4616 4617 4618 4619 4620 4621 4622 4623 4624 4625 4626 4627 4628
 4629 4630 4631 4632 4633 4634 4635 4636 4637 4638 4639 4640 4641 4642
 4643 4644 4645 4646 4647 4648 4649 4650 4651 4652 4653 4654 4655 4656
 4657 4658 4659 4660 4661 4662 4663 4664 4665 4666 4667 4668 4669 4670
 4671 4672 4673 4674 4675 4676 4677 4678 4679 4680 4681 4682 4683 4684
 4685 4686 4687 4688 4689 4690 4691 4692 4693 4694 4695 4696 4697 4698
 4699 4700 4701 4702 4703 4704 4705 4706 4707 4708 4709 4710 4711 4712
 4713 4714 4715 4716 4717 4718 4719 4720 4721 4722 4723 4724 4725 4726
 4727 4728 4729 4730 4731 4732 4733 4734 4735 4736 4737 4738 4739 4740
 4741 4742 4743 4744 4745 4746 4747 4748 4749 4750 4751 4752 4753 4754
 4755 4756 4757 4758 4759 4760 4761 4762 4763 4764 4765 4766 4767 4768
 4769 4770 4771 4772 4773 4774 4775 4776 4777 4778 4779 4780 4781 4782
 4783 4784 4785 4786 4787 4788 4789 4790 4791 4792 4793 4794 4795 4796
 4797 4798 4799 4800 4801 4802 4803 4804 4805 4806 4807 4808 4809 4810
 4811 4812 4813 4814 4815 4816 4817 4818 4819 4820 4821 4822 4823 4824
 4825 4826 4827 4828 4829 4830 4831 4832 4833 4834 4835 4836 4837 4838
 4839 4840 4841 4842 4843 4844 4845 4846 4847 4848 4849 4850 4851 4852
 4853 4854 4855 4856 4857 4858 4859 4860 4861 4862 4863 4864 4865 4866
 4867 4868 4869 4870 4871 4872 4873 4874 4875 4876 4877 4878 4879 4880
 4881 4882 4883 4884 4885 4886 4887 4888 4889 4890 4891 4892 4893 4894
 4895 4896 4897 4898 4899 4900 4901 4902 4903 4904 4905 4906 4907 4908
 4909 4910 4911 4912 4913 4914 4915 4916 4917 4918 4919 4920 4921 4922
 4923 4924 4925 4926 4927 4928 4929 4930 4931 4932 4933 4934 4935 4936
 4937 4938 4939 4940 4941 4942 4943 4944 4945 4946 4947 4948 4949 4950
 4951 4952 4953 4954 4955 4956 4957 4958 4959 4960 4961 4962 4963 4964
 4965 4966 4967 4968 4969 4970 4971 4972 4973 4974 4975 4976 4977 4978
 4979 4980 4981 4982 4983 4984 4985 4986 4987 4988 4989 4990 4991 4992
 4993 4994 4995 4996 4997 4998 4999]
```

In [175]:
```
print(X_tfidf_tr.shape)
print(X_tfidf_test.shape)
print(Y_tfidf_tr.shape)
print(X_tfidf_test.shape)
```

```
(4167, 4)
(833, 4)
(4167,)
(833, 4)
```

In [176]:
```
# Splitting the Train dataset into Train & Cross Validation Dataset
for train_index, test_index in tscv.split(X_tfidf_tr):
   print("TRAIN:", train_index, "TEST:", test_index)
   X_tfidf_train, X_tfidf_cv = X[train_index], X[test_index]
   Y_tfidf_train, Y_tfidf_cv = Y[train_index], Y[test_index]
```

```
TRAIN: [  0   1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17
  18  19  20  21  22  23  24  25  26  27  28  29  30  31  32  33  34  35
  36  37  38  39  40  41  42  43  44  45  46  47  48  49  50  51  52  53
  54  55  56  57  58  59  60  61  62  63  64  65  66  67  68  69  70  71
  72  73  74  75  76  77  78  79  80  81  82  83  84  85  86  87  88  89
  90  91  92  93  94  95  96  97  98  99 100 101 102 103 104 105 106 107
 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125
 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143
 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161
 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179
 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197
 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215
 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233
 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251
 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269
 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287
 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305
 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323
 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341
 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359
 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377
 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395
 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413
 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431
 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449
 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467
 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485
 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503
 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521
 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539
 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557
 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575
 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593
 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611
 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629
 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647
 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665
 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683
 684 685 686 687 688 689 690 691 692 693 694 695 696] TEST: [ 697  698  699  700  701  702  703  704  705
  706  707  708  709  710
  711  712  713  714  715  716  717  718  719  720  721  722  723  724
  725  726  727  728  729  730  731  732  733  734  735  736  737  738
  739  740  741  742  743  744  745  746  747  748  749  750  751  752
  753  754  755  756  757  758  759  760  761  762  763  764  765  766
  767  768  769  770  771  772  773  774  775  776  777  778  779  780
  781  782  783  784  785  786  787  788  789  790  791  792  793  794
  795  796  797  798  799  800  801  802  803  804  805  806  807  808
  809  810  811  812  813  814  815  816  817  818  819  820  821  822
  823  824  825  826  827  828  829  830  831  832  833  834  835  836
  837  838  839  840  841  842  843  844  845  846  847  848  849  850
  851  852  853  854  855  856  857  858  859  860  861  862  863  864
  865  866  867  868  869  870  871  872  873  874  875  876  877  878
  879  880  881  882  883  884  885  886  887  888  889  890  891  892
  893  894  895  896  897  898  899  900  901  902  903  904  905  906
  907  908  909  910  911  912  913  914  915  916  917  918  919  920
  921  922  923  924  925  926  927  928  929  930  931  932  933  934
  935  936  937  938  939  940  941  942  943  944  945  946  947  948
  949  950  951  952  953  954  955  956  957  958  959  960  961  962
  963  964  965  966  967  968  969  970  971  972  973  974  975  976
  977  978  979  980  981  982  983  984  985  986  987  988  989  990
  991  992  993  994  995  996  997  998  999 1000 1001 1002 1003 1004
 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018
 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032
 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046
 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060
 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074
 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088
 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102
 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116
 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128 1129 1130
 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142 1143 1144
 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158
 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172
 1173 1174 1175 1176 1177 1178 1179 1180 1181 1182 1183 1184 1185 1186
 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198 1199 1200
 1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212 1213 1214
 1215 1216 1217 1218 1219 1220 1221 1222 1223 1224 1225 1226 1227 1228
 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240 1241 1242
 1243 1244 1245 1246 1247 1248 1249 1250 1251 1252 1253 1254 1255 1256
 1257 1258 1259 1260 1261 1262 1263 1264 1265 1266 1267 1268 1269 1270
 1271 1272 1273 1274 1275 1276 1277 1278 1279 1280 1281 1282 1283 1284
 1285 1286 1287 1288 1289 1290 1291 1292 1293 1294 1295 1296 1297 1298
 1299 1300 1301 1302 1303 1304 1305 1306 1307 1308 1309 1310 1311 1312
 1313 1314 1315 1316 1317 1318 1319 1320 1321 1322 1323 1324 1325 1326
 1327 1328 1329 1330 1331 1332 1333 1334 1335 1336 1337 1338 1339 1340
 1341 1342 1343 1344 1345 1346 1347 1348 1349 1350 1351 1352 1353 1354
 1355 1356 1357 1358 1359 1360 1361 1362 1363 1364 1365 1366 1367 1368
```

```
 1369 1370 1371 1372 1373 1374 1375 1376 1377 1378 1379 1380 1381 1382
 1383 1384 1385 1386 1387 1388 1389 1390]
TRAIN: [   0    1    2 ... 1388 1389 1390] TEST: [1391 1392 1393 1394 1395 1396 1397 1398 1399 1400 1401
1402 1403 1404
 1405 1406 1407 1408 1409 1410 1411 1412 1413 1414 1415 1416 1417 1418
 1419 1420 1421 1422 1423 1424 1425 1426 1427 1428 1429 1430 1431 1432
 1433 1434 1435 1436 1437 1438 1439 1440 1441 1442 1443 1444 1445 1446
 1447 1448 1449 1450 1451 1452 1453 1454 1455 1456 1457 1458 1459 1460
 1461 1462 1463 1464 1465 1466 1467 1468 1469 1470 1471 1472 1473 1474
 1475 1476 1477 1478 1479 1480 1481 1482 1483 1484 1485 1486 1487 1488
 1489 1490 1491 1492 1493 1494 1495 1496 1497 1498 1499 1500 1501 1502
 1503 1504 1505 1506 1507 1508 1509 1510 1511 1512 1513 1514 1515 1516
 1517 1518 1519 1520 1521 1522 1523 1524 1525 1526 1527 1528 1529 1530
 1531 1532 1533 1534 1535 1536 1537 1538 1539 1540 1541 1542 1543 1544
 1545 1546 1547 1548 1549 1550 1551 1552 1553 1554 1555 1556 1557 1558
 1559 1560 1561 1562 1563 1564 1565 1566 1567 1568 1569 1570 1571 1572
 1573 1574 1575 1576 1577 1578 1579 1580 1581 1582 1583 1584 1585 1586
 1587 1588 1589 1590 1591 1592 1593 1594 1595 1596 1597 1598 1599 1600
 1601 1602 1603 1604 1605 1606 1607 1608 1609 1610 1611 1612 1613 1614
 1615 1616 1617 1618 1619 1620 1621 1622 1623 1624 1625 1626 1627 1628
 1629 1630 1631 1632 1633 1634 1635 1636 1637 1638 1639 1640 1641 1642
 1643 1644 1645 1646 1647 1648 1649 1650 1651 1652 1653 1654 1655 1656
 1657 1658 1659 1660 1661 1662 1663 1664 1665 1666 1667 1668 1669 1670
 1671 1672 1673 1674 1675 1676 1677 1678 1679 1680 1681 1682 1683 1684
 1685 1686 1687 1688 1689 1690 1691 1692 1693 1694 1695 1696 1697 1698
 1699 1700 1701 1702 1703 1704 1705 1706 1707 1708 1709 1710 1711 1712
 1713 1714 1715 1716 1717 1718 1719 1720 1721 1722 1723 1724 1725 1726
 1727 1728 1729 1730 1731 1732 1733 1734 1735 1736 1737 1738 1739 1740
 1741 1742 1743 1744 1745 1746 1747 1748 1749 1750 1751 1752 1753 1754
 1755 1756 1757 1758 1759 1760 1761 1762 1763 1764 1765 1766 1767 1768
 1769 1770 1771 1772 1773 1774 1775 1776 1777 1778 1779 1780 1781 1782
 1783 1784 1785 1786 1787 1788 1789 1790 1791 1792 1793 1794 1795 1796
 1797 1798 1799 1800 1801 1802 1803 1804 1805 1806 1807 1808 1809 1810
 1811 1812 1813 1814 1815 1816 1817 1818 1819 1820 1821 1822 1823 1824
 1825 1826 1827 1828 1829 1830 1831 1832 1833 1834 1835 1836 1837 1838
 1839 1840 1841 1842 1843 1844 1845 1846 1847 1848 1849 1850 1851 1852
 1853 1854 1855 1856 1857 1858 1859 1860 1861 1862 1863 1864 1865 1866
 1867 1868 1869 1870 1871 1872 1873 1874 1875 1876 1877 1878 1879 1880
 1881 1882 1883 1884 1885 1886 1887 1888 1889 1890 1891 1892 1893 1894
 1895 1896 1897 1898 1899 1900 1901 1902 1903 1904 1905 1906 1907 1908
 1909 1910 1911 1912 1913 1914 1915 1916 1917 1918 1919 1920 1921 1922
 1923 1924 1925 1926 1927 1928 1929 1930 1931 1932 1933 1934 1935 1936
 1937 1938 1939 1940 1941 1942 1943 1944 1945 1946 1947 1948 1949 1950
 1951 1952 1953 1954 1955 1956 1957 1958 1959 1960 1961 1962 1963 1964
 1965 1966 1967 1968 1969 1970 1971 1972 1973 1974 1975 1976 1977 1978
 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991 1992
 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006
 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020
 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034
 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048
 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062
 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076
 2077 2078 2079 2080 2081 2082 2083 2084]
TRAIN: [   0    1    2 ... 2082 2083 2084] TEST: [2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095
2096 2097 2098
 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112
 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126
 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140
 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154
 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168
 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182
 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196
 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210
 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224
 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238
 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252
 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266
 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280
 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294
 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308
 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322
 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336
 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350
 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364
 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378
 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392
 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406
 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420
 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434
 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448
 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462
 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476
 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490
 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504
 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518
 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532
 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546
 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560
```

```
 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574
 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584 2585 2586 2587 2588
 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598 2599 2600 2601 2602
 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612 2613 2614 2615 2616
 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626 2627 2628 2629 2630
 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640 2641 2642 2643 2644
 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654 2655 2656 2657 2658
 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668 2669 2670 2671 2672
 2673 2674 2675 2676 2677 2678 2679 2680 2681 2682 2683 2684 2685 2686
 2687 2688 2689 2690 2691 2692 2693 2694 2695 2696 2697 2698 2699 2700
 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710 2711 2712 2713 2714
 2715 2716 2717 2718 2719 2720 2721 2722 2723 2724 2725 2726 2727 2728
 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738 2739 2740 2741 2742
 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752 2753 2754 2755 2756
 2757 2758 2759 2760 2761 2762 2763 2764 2765 2766 2767 2768 2769 2770
 2771 2772 2773 2774 2775 2776 2777 2778]
TRAIN: [   0    1    2 ... 2776 2777 2778] TEST: [2779 2780 2781 2782 2783 2784 2785 2786 2787 2788 2789
2790 2791 2792
 2793 2794 2795 2796 2797 2798 2799 2800 2801 2802 2803 2804 2805 2806
 2807 2808 2809 2810 2811 2812 2813 2814 2815 2816 2817 2818 2819 2820
 2821 2822 2823 2824 2825 2826 2827 2828 2829 2830 2831 2832 2833 2834
 2835 2836 2837 2838 2839 2840 2841 2842 2843 2844 2845 2846 2847 2848
 2849 2850 2851 2852 2853 2854 2855 2856 2857 2858 2859 2860 2861 2862
 2863 2864 2865 2866 2867 2868 2869 2870 2871 2872 2873 2874 2875 2876
 2877 2878 2879 2880 2881 2882 2883 2884 2885 2886 2887 2888 2889 2890
 2891 2892 2893 2894 2895 2896 2897 2898 2899 2900 2901 2902 2903 2904
 2905 2906 2907 2908 2909 2910 2911 2912 2913 2914 2915 2916 2917 2918
 2919 2920 2921 2922 2923 2924 2925 2926 2927 2928 2929 2930 2931 2932
 2933 2934 2935 2936 2937 2938 2939 2940 2941 2942 2943 2944 2945 2946
 2947 2948 2949 2950 2951 2952 2953 2954 2955 2956 2957 2958 2959 2960
 2961 2962 2963 2964 2965 2966 2967 2968 2969 2970 2971 2972 2973 2974
 2975 2976 2977 2978 2979 2980 2981 2982 2983 2984 2985 2986 2987 2988
 2989 2990 2991 2992 2993 2994 2995 2996 2997 2998 2999 3000 3001 3002
 3003 3004 3005 3006 3007 3008 3009 3010 3011 3012 3013 3014 3015 3016
 3017 3018 3019 3020 3021 3022 3023 3024 3025 3026 3027 3028 3029 3030
 3031 3032 3033 3034 3035 3036 3037 3038 3039 3040 3041 3042 3043 3044
 3045 3046 3047 3048 3049 3050 3051 3052 3053 3054 3055 3056 3057 3058
 3059 3060 3061 3062 3063 3064 3065 3066 3067 3068 3069 3070 3071 3072
 3073 3074 3075 3076 3077 3078 3079 3080 3081 3082 3083 3084 3085 3086
 3087 3088 3089 3090 3091 3092 3093 3094 3095 3096 3097 3098 3099 3100
 3101 3102 3103 3104 3105 3106 3107 3108 3109 3110 3111 3112 3113 3114
 3115 3116 3117 3118 3119 3120 3121 3122 3123 3124 3125 3126 3127 3128
 3129 3130 3131 3132 3133 3134 3135 3136 3137 3138 3139 3140 3141 3142
 3143 3144 3145 3146 3147 3148 3149 3150 3151 3152 3153 3154 3155 3156
 3157 3158 3159 3160 3161 3162 3163 3164 3165 3166 3167 3168 3169 3170
 3171 3172 3173 3174 3175 3176 3177 3178 3179 3180 3181 3182 3183 3184
 3185 3186 3187 3188 3189 3190 3191 3192 3193 3194 3195 3196 3197 3198
 3199 3200 3201 3202 3203 3204 3205 3206 3207 3208 3209 3210 3211 3212
 3213 3214 3215 3216 3217 3218 3219 3220 3221 3222 3223 3224 3225 3226
 3227 3228 3229 3230 3231 3232 3233 3234 3235 3236 3237 3238 3239 3240
 3241 3242 3243 3244 3245 3246 3247 3248 3249 3250 3251 3252 3253 3254
 3255 3256 3257 3258 3259 3260 3261 3262 3263 3264 3265 3266 3267 3268
 3269 3270 3271 3272 3273 3274 3275 3276 3277 3278 3279 3280 3281 3282
 3283 3284 3285 3286 3287 3288 3289 3290 3291 3292 3293 3294 3295 3296
 3297 3298 3299 3300 3301 3302 3303 3304 3305 3306 3307 3308 3309 3310
 3311 3312 3313 3314 3315 3316 3317 3318 3319 3320 3321 3322 3323 3324
 3325 3326 3327 3328 3329 3330 3331 3332 3333 3334 3335 3336 3337 3338
 3339 3340 3341 3342 3343 3344 3345 3346 3347 3348 3349 3350 3351 3352
 3353 3354 3355 3356 3357 3358 3359 3360 3361 3362 3363 3364 3365 3366
 3367 3368 3369 3370 3371 3372 3373 3374 3375 3376 3377 3378 3379 3380
 3381 3382 3383 3384 3385 3386 3387 3388 3389 3390 3391 3392 3393 3394
 3395 3396 3397 3398 3399 3400 3401 3402 3403 3404 3405 3406 3407 3408
 3409 3410 3411 3412 3413 3414 3415 3416 3417 3418 3419 3420 3421 3422
 3423 3424 3425 3426 3427 3428 3429 3430 3431 3432 3433 3434 3435 3436
 3437 3438 3439 3440 3441 3442 3443 3444 3445 3446 3447 3448 3449 3450
 3451 3452 3453 3454 3455 3456 3457 3458 3459 3460 3461 3462 3463 3464
 3465 3466 3467 3468 3469 3470 3471 3472]
TRAIN: [   0    1    2 ... 3470 3471 3472] TEST: [3473 3474 3475 3476 3477 3478 3479 3480 3481 3482 3483
3484 3485 3486
 3487 3488 3489 3490 3491 3492 3493 3494 3495 3496 3497 3498 3499 3500
 3501 3502 3503 3504 3505 3506 3507 3508 3509 3510 3511 3512 3513 3514
 3515 3516 3517 3518 3519 3520 3521 3522 3523 3524 3525 3526 3527 3528
 3529 3530 3531 3532 3533 3534 3535 3536 3537 3538 3539 3540 3541 3542
 3543 3544 3545 3546 3547 3548 3549 3550 3551 3552 3553 3554 3555 3556
 3557 3558 3559 3560 3561 3562 3563 3564 3565 3566 3567 3568 3569 3570
 3571 3572 3573 3574 3575 3576 3577 3578 3579 3580 3581 3582 3583 3584
 3585 3586 3587 3588 3589 3590 3591 3592 3593 3594 3595 3596 3597 3598
 3599 3600 3601 3602 3603 3604 3605 3606 3607 3608 3609 3610 3611 3612
 3613 3614 3615 3616 3617 3618 3619 3620 3621 3622 3623 3624 3625 3626
 3627 3628 3629 3630 3631 3632 3633 3634 3635 3636 3637 3638 3639 3640
 3641 3642 3643 3644 3645 3646 3647 3648 3649 3650 3651 3652 3653 3654
 3655 3656 3657 3658 3659 3660 3661 3662 3663 3664 3665 3666 3667 3668
 3669 3670 3671 3672 3673 3674 3675 3676 3677 3678 3679 3680 3681 3682
 3683 3684 3685 3686 3687 3688 3689 3690 3691 3692 3693 3694 3695 3696
 3697 3698 3699 3700 3701 3702 3703 3704 3705 3706 3707 3708 3709 3710
 3711 3712 3713 3714 3715 3716 3717 3718 3719 3720 3721 3722 3723 3724
 3725 3726 3727 3728 3729 3730 3731 3732 3733 3734 3735 3736 3737 3738
 3739 3740 3741 3742 3743 3744 3745 3746 3747 3748 3749 3750 3751 3752
```

```
 3753 3754 3755 3756 3757 3758 3759 3760 3761 3762 3763 3764 3765 3766
 3767 3768 3769 3770 3771 3772 3773 3774 3775 3776 3777 3778 3779 3780
 3781 3782 3783 3784 3785 3786 3787 3788 3789 3790 3791 3792 3793 3794
 3795 3796 3797 3798 3799 3800 3801 3802 3803 3804 3805 3806 3807 3808
 3809 3810 3811 3812 3813 3814 3815 3816 3817 3818 3819 3820 3821 3822
 3823 3824 3825 3826 3827 3828 3829 3830 3831 3832 3833 3834 3835 3836
 3837 3838 3839 3840 3841 3842 3843 3844 3845 3846 3847 3848 3849 3850
 3851 3852 3853 3854 3855 3856 3857 3858 3859 3860 3861 3862 3863 3864
 3865 3866 3867 3868 3869 3870 3871 3872 3873 3874 3875 3876 3877 3878
 3879 3880 3881 3882 3883 3884 3885 3886 3887 3888 3889 3890 3891 3892
 3893 3894 3895 3896 3897 3898 3899 3900 3901 3902 3903 3904 3905 3906
 3907 3908 3909 3910 3911 3912 3913 3914 3915 3916 3917 3918 3919 3920
 3921 3922 3923 3924 3925 3926 3927 3928 3929 3930 3931 3932 3933 3934
 3935 3936 3937 3938 3939 3940 3941 3942 3943 3944 3945 3946 3947 3948
 3949 3950 3951 3952 3953 3954 3955 3956 3957 3958 3959 3960 3961 3962
 3963 3964 3965 3966 3967 3968 3969 3970 3971 3972 3973 3974 3975 3976
 3977 3978 3979 3980 3981 3982 3983 3984 3985 3986 3987 3988 3989 3990
 3991 3992 3993 3994 3995 3996 3997 3998 3999 4000 4001 4002 4003 4004
 4005 4006 4007 4008 4009 4010 4011 4012 4013 4014 4015 4016 4017 4018
 4019 4020 4021 4022 4023 4024 4025 4026 4027 4028 4029 4030 4031 4032
 4033 4034 4035 4036 4037 4038 4039 4040 4041 4042 4043 4044 4045 4046
 4047 4048 4049 4050 4051 4052 4053 4054 4055 4056 4057 4058 4059 4060
 4061 4062 4063 4064 4065 4066 4067 4068 4069 4070 4071 4072 4073 4074
 4075 4076 4077 4078 4079 4080 4081 4082 4083 4084 4085 4086 4087 4088
 4089 4090 4091 4092 4093 4094 4095 4096 4097 4098 4099 4100 4101 4102
 4103 4104 4105 4106 4107 4108 4109 4110 4111 4112 4113 4114 4115 4116
 4117 4118 4119 4120 4121 4122 4123 4124 4125 4126 4127 4128 4129 4130
 4131 4132 4133 4134 4135 4136 4137 4138 4139 4140 4141 4142 4143 4144
 4145 4146 4147 4148 4149 4150 4151 4152 4153 4154 4155 4156 4157 4158
 4159 4160 4161 4162 4163 4164 4165 4166]
```

In [177]:
```
print(X_tfidf_train.shape)
print(X_tfidf_cv.shape)
print(X_tfidf_test.shape)

print(Y_tfidf_train.shape)
print(Y_tfidf_cv.shape)
print(Y_tfidf_test.shape)
```

```
(3473, 4)
(694, 4)
(833, 4)
(3473,)
(694,)
(833,)
```

In [178]:
```
# Removing Date variable from all the datasets
tfidf_X= []
for each in X_tfidf_train:
    tfidf_X.append(np.delete(each,3))
X_tfidf_train = np.asarray(tfidf_X)


tfidf_X_cv= []
for each in X_tfidf_cv:
    tfidf_X_cv.append(np.delete(each,3))
X_tfidf_cv = np.asarray(tfidf_X_cv)

tfidf_X_test= []
for each in X_tfidf_test:
    tfidf_X_test.append(np.delete(each,3))
X_tfidf_test = np.asarray(tfidf_X_test)
```

In [179]:
```
print(X_tfidf_train.shape)
print(X_tfidf_cv.shape)
print(X_tfidf_test.shape)

print(Y_tfidf_train.shape)
print(Y_tfidf_cv.shape)
print(Y_tfidf_test.shape)
```

```
(3473, 3)
(694, 3)
(833, 3)
(3473,)
(694,)
(833,)
```

In [180]:
```
# Finding the Test accuracy using Cross Validation Data Set
for i in range(1,30,2):
    # instantiate learning model (k = 30)
    knn = KNeighborsClassifier(n_neighbors=i)

    # fitting the model on crossvalidation train
    knn.fit(X_tfidf_train, Y_tfidf_train)

    # predict the response on the crossvalidation train
    pred = knn.predict(X_tfidf_cv)

    # evaluate CV accuracy
    acc = accuracy_score(Y_tfidf_cv, pred, normalize=True) * float(100)
    print('\nCV accuracy for k = %d is %d%%' % (i, acc))

knn = KNeighborsClassifier(1)
knn.fit(X_tfidf_train,Y_tfidf_train)
pred = knn.predict(X_tfidf_test)
acc = accuracy_score(Y_tfidf_test, pred, normalize=True) * float(100)
print('\n****Test accuracy for k = 1 is %d%%' % (acc))
```

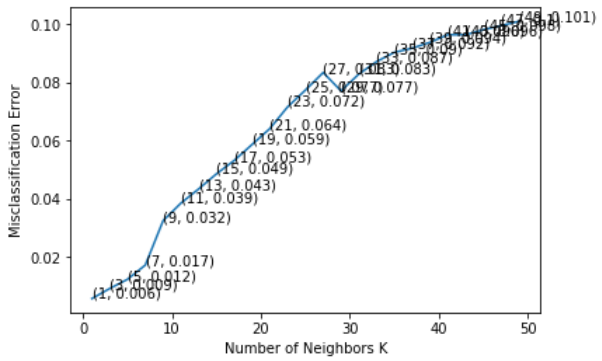
```
CV accuracy for k = 1 is 98%

CV accuracy for k = 3 is 98%

CV accuracy for k = 5 is 97%

CV accuracy for k = 7 is 96%

CV accuracy for k = 9 is 96%

CV accuracy for k = 11 is 96%

CV accuracy for k = 13 is 96%

CV accuracy for k = 15 is 95%

CV accuracy for k = 17 is 95%

CV accuracy for k = 19 is 94%

CV accuracy for k = 21 is 94%

CV accuracy for k = 23 is 93%

CV accuracy for k = 25 is 92%

CV accuracy for k = 27 is 89%

CV accuracy for k = 29 is 88%

****Test accuracy for k = 1 is 99%
```

In [181]:
```
# creating odd list of K for KNN
myList = list(range(0,50))
neighbors = list(filter(lambda x: x % 2 != 0, myList))

# empty list that will hold cv scores
cv_scores = []
```

In [182]:
```
for k in neighbors:
    knn = KNeighborsClassifier(n_neighbors=k)
    scores = cross_val_score(knn, X_tfidf_train, Y_tfidf_train, cv=10)
    cv_scores.append(scores.mean())
```

In [183]:
```
# changing to misclassification error
MSE = [1 - x for x in cv_scores]

# determining best k
optimal_k = neighbors[MSE.index(min(MSE))]
print('\nThe optimal number of neighbors is %d.' % optimal_k)

# plot misclassification error vs k
plt.plot(neighbors, MSE)

for xy in zip(neighbors, np.round(MSE,3)):
    plt.annotate('(%s, %s)' % xy, xy=xy, textcoords='data')

plt.xlabel('Number of Neighbors K')
plt.ylabel('Misclassification Error')
plt.show()

print("the misclassification error for each k value is : ", np.round(MSE,3))
```

```
The optimal number of neighbors is 1.
```




```
the misclassification error for each k value is :  [0.006 0.009 0.012 0.017 0.032 0.039 0.043 0.049 0.053
0.059 0.064 0.072
 0.077 0.083 0.077 0.083 0.087 0.09  0.092 0.094 0.096 0.096 0.098 0.1
 0.101]
```

In [184]:
```
# KNN with k = optimal_k

# instantiate learning model k = optimal_k
knn_optimal = KNeighborsClassifier(n_neighbors=optimal_k)

# fitting the model
knn_optimal.fit(X_tfidf_train, Y_tfidf_train)

# predict the response
pred = knn_optimal.predict(X_tfidf_test)

# evaluate accuracy
acc = accuracy_score(Y_tfidf_test, pred) * 100
print('\nThe accuracy of the knn classifier for k = %d is %f%%' % (optimal_k, acc))
```

```
The accuracy of the knn classifier for k = 1 is 99.639856%
```

Average Word2Vec (Avg W2V)

In [ ]:
```
# Word2Vec
#model = KeyedVectors.load_word2vec_format('GoogleNews-vectors-negative300.bin', binary=True)
```

In [76]:
```
# Train your own Word2Vec model using your own text corpus
i=0
list_of_sent=[]
for sent in final_dataset['CleanedText'].values:
    list_of_sent.append(sent.split())
```

In [77]:
```
# min_count = 5 considers only words that occured atleast 5 times
w2v_model=Word2Vec(list_of_sent,min_count=5,size=50, workers=4)
```

In [78]:
```
# Printing Sample words
w2v_words = list(w2v_model.wv.vocab)
print("number of words that occured minimum 5 times ",len(w2v_words))
print("sample words ", w2v_words[0:50])
```

```
number of words that occured minimum 5 times  3596
sample words  ['stare', 'thrip', 'made', 'teenag', 'mislead', 'support', 'televis', 'disc', 'decreas', 'p
eko', 'pound', 'preserv', 'die', 'canola', 'wash', 'difficulti', 'explain', 'relax', 'agar', 'guest', 'pe
rsian', 'primarili', 'king', 'second', 'indoor', 'beast', 'certifi', 'hundr', 'grab', 'overweight', 'imme
di', 'come', 'puke', 'here', 'desper', 'pic', 'prudhomm', 'ignor', 'angl', 'aint', 'certain', 'mental',
'reject', 'drop', 'scream', 'work', 'grow', 'evil', 'sort', 'labrador']
```

In [79]:
```
# average Word2Vec
# compute average word2vec for each review.
sent_vectors = []; # the avg-w2v for each sentence/review is stored in this list
for sent in list_of_sent: # for each review/sentence
    sent_vec = np.zeros(50) # as word vectors are of zero length
    cnt_words =0; # num of words with a valid vector in the sentence/review
    for word in sent: # for each word in a review/sentence
        if word in w2v_words:
            vec = w2v_model.wv[word]
            sent_vec += vec
            cnt_words += 1
    if cnt_words != 0:
        sent_vec /= cnt_words
    sent_vectors.append(sent_vec)
print(len(sent_vectors))
```

```
5000
```

In [80]:
```
#tsne
start_time_code = time.time()
model = TSNE(n_components=2, random_state=0, perplexity=30)
tsne_data_Aw2v = model.fit_transform(sent_vectors)
end_time_code = time.time()
print ("Running Time for code execution " + str(end_time_code - start_time_code) + " secs")
```

```
Running Time for code execution 119.21003150939941 secs
```

In [81]:
```
#Saving tsne data into seperate file for further usecases.
tsne_data_Aw2v = np.vstack((tsne_data_Aw2v.T,label,Time)).T
tsne_AW2V_df = pd.DataFrame(data=tsne_data_Aw2v, columns=("X", "Y", "Label","Date"))
tsne_AW2V_df.to_csv('/home/venkatasujit272/tsne_AW2V_data.csv')
```

In [106]:
```
# Aliasing the Score field values into Binary format
def partition(x):
    if x == 'negative':
        return 0.
    return 1.

# Aliasing the Score to String Format (Positive for Score > 3 & Negative for Score  < 3)

actualScore = tsne_AW2V_df['Label']
positiveNegative = actualScore.map(partition)
tsne_AW2V_df['Label'] = positiveNegative
```

In [107]:
```
tsne_AW2V_df.head(1)
```

Out[107]:

	X	Y	Label	Date
0	-9.86456	-6.63527	1.0	1999-10-08

In [108]:
```
#Sorting the Dataframe with Date to apply Timebased Split
tsne_AW2V_df.sort_values(by='Date',ascending=1)
```

Out[108]:

	X	Y	Label	Date
0	-9.86456	-6.63527	1.0	1999-10-08
28	-7.92173	-3.97298	1.0	1999-10-25
377	4.77067	-6.03357	1.0	1999-12-02
287	-10.1318	-6.53944	1.0	1999-12-06
376	10.7702	1.09368	1.0	2000-01-03
209	9.03657	-2.81984	1.0	2000-01-09
2540	6.11455	-6.58935	0.0	2000-01-19
198	7.51311	-9.64076	1.0	2000-01-24
224	-5.14856	-7.64139	1.0	2000-02-26
284	6.9695	0.245181	1.0	2000-06-03
378	7.04625	-0.410789	1.0	2000-06-03
205	-8.24021	-5.58602	1.0	2000-06-23
206	-9.00248	-5.94668	1.0	2000-06-29
415	8.8098	-8.75489	1.0	2000-07-31
684	8.90907	1.63241	1.0	2000-08-09
714	2.10214	6.8294	1.0	2000-08-09
212	8.88045	3.6079	1.0	2000-08-15
258	-10.0803	-6.65756	1.0	2000-10-03
691	5.75153	-7.13541	1.0	2000-12-05
315	10.6806	0.89765	1.0	2000-12-19
286	-8.364	-7.1144	1.0	2000-12-30
692	7.9073	4.42538	1.0	2001-02-22
2547	-9.44292	-5.99435	0.0	2001-06-11
2543	5.01425	-6.71079	0.0	2001-08-08
233	5.92343	-4.89261	1.0	2001-09-24
701	4.56288	5.71232	1.0	2001-10-23
309	5.0271	-6.83168	1.0	2001-10-26
30	0.247379	-6.5743	1.0	2001-12-26
830	-4.98638	-4.9586	1.0	2002-01-06
288	-9.87982	-6.59116	1.0	2002-02-04
...	...	...	...	...
232	4.12953	-6.38718	1.0	2012-10-22
3856	0.243018	-11.1007	0.0	2012-10-23
1792	-4.08604	-1.76852	1.0	2012-10-23
813	-7.30072	8.9931	1.0	2012-10-23
3649	2.88995	12.7322	0.0	2012-10-23
2030	1.38835	12.7569	1.0	2012-10-23
1897	5.90414	-6.32744	1.0	2012-10-23
1086	-1.18904	-0.262306	1.0	2012-10-23
3056	-8.98367	-5.92079	0.0	2012-10-23
1594	-4.89015	6.93228	1.0	2012-10-23
4556	-1.69037	-8.78941	0.0	2012-10-23
134	-4.31622	3.7928	1.0	2012-10-23
2717	-3.08961	0.0190142	0.0	2012-10-24
2705	-6.26781	-4.91724	0.0	2012-10-24
1295	0.71428	-3.04017	1.0	2012-10-24
2391	8.87518	2.60925	1.0	2012-10-24
1603	-5.54538	9.02892	1.0	2012-10-25
3849	-1.42926	-1.72346	0.0	2012-10-25
2715	-7.63241	8.99004	0.0	2012-10-25
4115	3.58862	-1.07547	0.0	2012-10-25
1827	-8.03734	0.590219	1.0	2012-10-25
2978	6.85339	-5.6653	0.0	2012-10-26

	X	Y	Label	Date
135	-4.22697	-4.25746	1.0	2012-10-26
782	-1.40514	-7.92723	1.0	2012-10-26
136	-4.24847	4.89457	1.0	2012-10-26
3569	4.41025	-3.19662	0.0	2012-10-26
1349	-8.73549	6.57709	1.0	2012-10-26
493	10.6767	-6.28655	1.0	2012-10-26
4793	-4.15521	4.98447	0.0	2012-10-26
2592	9.93938	-7.65937	0.0	2012-10-26

5000 rows × 4 columns

In [117]:
```
# create design matrix X and target vector y
X_AW2V = np.array(tsne_AW2V_df.iloc[:, 0:4]) # end index is exclusive
Y_AW2V = np.array(tsne_AW2V_df['Label']) # showing you two ways of indexing a pandas df
```

In [118]:
```
#Time based split
tscv = TimeSeriesSplit(n_splits=5)
print(tscv)
```

```
TimeSeriesSplit(n_splits=5)
```

In [125]:
```
#Splitting the Dataset into Train set & Test set
for train_index, test_index in tscv.split(X_AW2V):
   print("TRAIN:", train_index, "TEST:", test_index)
   X_AW2V_tr, X_AW2V_test = X[train_index], X[test_index]
   Y_AW2V_tr, Y_AW2V_test = Y[train_index], Y[test_index]
```

```
TRAIN: [  0   1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17
  18  19  20  21  22  23  24  25  26  27  28  29  30  31  32  33  34  35
  36  37  38  39  40  41  42  43  44  45  46  47  48  49  50  51  52  53
  54  55  56  57  58  59  60  61  62  63  64  65  66  67  68  69  70  71
  72  73  74  75  76  77  78  79  80  81  82  83  84  85  86  87  88  89
  90  91  92  93  94  95  96  97  98  99 100 101 102 103 104 105 106 107
 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125
 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143
 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161
 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179
 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197
 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215
 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233
 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251
 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269
 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287
 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305
 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323
 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341
 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359
 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377
 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395
 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413
 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431
 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449
 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467
 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485
 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503
 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521
 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539
 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557
 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575
 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593
 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611
 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629
 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647
 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665
 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683
 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701
 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719
 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737
 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755
 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773
 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791
 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809
 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827
 828 829 830 831 832 833 834] TEST: [ 835  836  837  838  839  840  841  842  843  844  845  846  847  84
8
  849  850  851  852  853  854  855  856  857  858  859  860  861  862
  863  864  865  866  867  868  869  870  871  872  873  874  875  876
  877  878  879  880  881  882  883  884  885  886  887  888  889  890
  891  892  893  894  895  896  897  898  899  900  901  902  903  904
  905  906  907  908  909  910  911  912  913  914  915  916  917  918
  919  920  921  922  923  924  925  926  927  928  929  930  931  932
  933  934  935  936  937  938  939  940  941  942  943  944  945  946
  947  948  949  950  951  952  953  954  955  956  957  958  959  960
  961  962  963  964  965  966  967  968  969  970  971  972  973  974
  975  976  977  978  979  980  981  982  983  984  985  986  987  988
  989  990  991  992  993  994  995  996  997  998  999 1000 1001 1002
 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016
 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030
 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044
 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058
 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072
 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086
 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100
 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114
 1115 1116 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128
 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142
 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156
 1157 1158 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170
 1171 1172 1173 1174 1175 1176 1177 1178 1179 1180 1181 1182 1183 1184
 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198
 1199 1200 1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212
 1213 1214 1215 1216 1217 1218 1219 1220 1221 1222 1223 1224 1225 1226
 1227 1228 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240
 1241 1242 1243 1244 1245 1246 1247 1248 1249 1250 1251 1252 1253 1254
 1255 1256 1257 1258 1259 1260 1261 1262 1263 1264 1265 1266 1267 1268
 1269 1270 1271 1272 1273 1274 1275 1276 1277 1278 1279 1280 1281 1282
 1283 1284 1285 1286 1287 1288 1289 1290 1291 1292 1293 1294 1295 1296
 1297 1298 1299 1300 1301 1302 1303 1304 1305 1306 1307 1308 1309 1310
 1311 1312 1313 1314 1315 1316 1317 1318 1319 1320 1321 1322 1323 1324
 1325 1326 1327 1328 1329 1330 1331 1332 1333 1334 1335 1336 1337 1338
 1339 1340 1341 1342 1343 1344 1345 1346 1347 1348 1349 1350 1351 1352
 1353 1354 1355 1356 1357 1358 1359 1360 1361 1362 1363 1364 1365 1366
 1367 1368 1369 1370 1371 1372 1373 1374 1375 1376 1377 1378 1379 1380
 1381 1382 1383 1384 1385 1386 1387 1388 1389 1390 1391 1392 1393 1394
```

```
 1395 1396 1397 1398 1399 1400 1401 1402 1403 1404 1405 1406 1407 1408
 1409 1410 1411 1412 1413 1414 1415 1416 1417 1418 1419 1420 1421 1422
 1423 1424 1425 1426 1427 1428 1429 1430 1431 1432 1433 1434 1435 1436
 1437 1438 1439 1440 1441 1442 1443 1444 1445 1446 1447 1448 1449 1450
 1451 1452 1453 1454 1455 1456 1457 1458 1459 1460 1461 1462 1463 1464
 1465 1466 1467 1468 1469 1470 1471 1472 1473 1474 1475 1476 1477 1478
 1479 1480 1481 1482 1483 1484 1485 1486 1487 1488 1489 1490 1491 1492
 1493 1494 1495 1496 1497 1498 1499 1500 1501 1502 1503 1504 1505 1506
 1507 1508 1509 1510 1511 1512 1513 1514 1515 1516 1517 1518 1519 1520
 1521 1522 1523 1524 1525 1526 1527 1528 1529 1530 1531 1532 1533 1534
 1535 1536 1537 1538 1539 1540 1541 1542 1543 1544 1545 1546 1547 1548
 1549 1550 1551 1552 1553 1554 1555 1556 1557 1558 1559 1560 1561 1562
 1563 1564 1565 1566 1567 1568 1569 1570 1571 1572 1573 1574 1575 1576
 1577 1578 1579 1580 1581 1582 1583 1584 1585 1586 1587 1588 1589 1590
 1591 1592 1593 1594 1595 1596 1597 1598 1599 1600 1601 1602 1603 1604
 1605 1606 1607 1608 1609 1610 1611 1612 1613 1614 1615 1616 1617 1618
 1619 1620 1621 1622 1623 1624 1625 1626 1627 1628 1629 1630 1631 1632
 1633 1634 1635 1636 1637 1638 1639 1640 1641 1642 1643 1644 1645 1646
 1647 1648 1649 1650 1651 1652 1653 1654 1655 1656 1657 1658 1659 1660
 1661 1662 1663 1664 1665 1666 1667]
TRAIN: [   0    1    2 ... 1665 1666 1667] TEST: [1668 1669 1670 1671 1672 1673 1674 1675 1676 1677 1678
1679 1680 1681
 1682 1683 1684 1685 1686 1687 1688 1689 1690 1691 1692 1693 1694 1695
 1696 1697 1698 1699 1700 1701 1702 1703 1704 1705 1706 1707 1708 1709
 1710 1711 1712 1713 1714 1715 1716 1717 1718 1719 1720 1721 1722 1723
 1724 1725 1726 1727 1728 1729 1730 1731 1732 1733 1734 1735 1736 1737
 1738 1739 1740 1741 1742 1743 1744 1745 1746 1747 1748 1749 1750 1751
 1752 1753 1754 1755 1756 1757 1758 1759 1760 1761 1762 1763 1764 1765
 1766 1767 1768 1769 1770 1771 1772 1773 1774 1775 1776 1777 1778 1779
 1780 1781 1782 1783 1784 1785 1786 1787 1788 1789 1790 1791 1792 1793
 1794 1795 1796 1797 1798 1799 1800 1801 1802 1803 1804 1805 1806 1807
 1808 1809 1810 1811 1812 1813 1814 1815 1816 1817 1818 1819 1820 1821
 1822 1823 1824 1825 1826 1827 1828 1829 1830 1831 1832 1833 1834 1835
 1836 1837 1838 1839 1840 1841 1842 1843 1844 1845 1846 1847 1848 1849
 1850 1851 1852 1853 1854 1855 1856 1857 1858 1859 1860 1861 1862 1863
 1864 1865 1866 1867 1868 1869 1870 1871 1872 1873 1874 1875 1876 1877
 1878 1879 1880 1881 1882 1883 1884 1885 1886 1887 1888 1889 1890 1891
 1892 1893 1894 1895 1896 1897 1898 1899 1900 1901 1902 1903 1904 1905
 1906 1907 1908 1909 1910 1911 1912 1913 1914 1915 1916 1917 1918 1919
 1920 1921 1922 1923 1924 1925 1926 1927 1928 1929 1930 1931 1932 1933
 1934 1935 1936 1937 1938 1939 1940 1941 1942 1943 1944 1945 1946 1947
 1948 1949 1950 1951 1952 1953 1954 1955 1956 1957 1958 1959 1960 1961
 1962 1963 1964 1965 1966 1967 1968 1969 1970 1971 1972 1973 1974 1975
 1976 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989
 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003
 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017
 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031
 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045
 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059
 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073
 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087
 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101
 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115
 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129
 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143
 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157
 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171
 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185
 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199
 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213
 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227
 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241
 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255
 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269
 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283
 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297
 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311
 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325
 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339
 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353
 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367
 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381
 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395
 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409
 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423
 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437
 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451
 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465
 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479
 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493
 2494 2495 2496 2497 2498 2499 2500]
TRAIN: [   0    1    2 ... 2498 2499 2500] TEST: [2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511
2512 2513 2514
 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528
 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542
 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556
 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570
 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584
```

```
 2585 2586 2587 2588 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598
 2599 2600 2601 2602 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612
 2613 2614 2615 2616 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626
 2627 2628 2629 2630 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640
 2641 2642 2643 2644 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654
 2655 2656 2657 2658 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668
 2669 2670 2671 2672 2673 2674 2675 2676 2677 2678 2679 2680 2681 2682
 2683 2684 2685 2686 2687 2688 2689 2690 2691 2692 2693 2694 2695 2696
 2697 2698 2699 2700 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710
 2711 2712 2713 2714 2715 2716 2717 2718 2719 2720 2721 2722 2723 2724
 2725 2726 2727 2728 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738
 2739 2740 2741 2742 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752
 2753 2754 2755 2756 2757 2758 2759 2760 2761 2762 2763 2764 2765 2766
 2767 2768 2769 2770 2771 2772 2773 2774 2775 2776 2777 2778 2779 2780
 2781 2782 2783 2784 2785 2786 2787 2788 2789 2790 2791 2792 2793 2794
 2795 2796 2797 2798 2799 2800 2801 2802 2803 2804 2805 2806 2807 2808
 2809 2810 2811 2812 2813 2814 2815 2816 2817 2818 2819 2820 2821 2822
 2823 2824 2825 2826 2827 2828 2829 2830 2831 2832 2833 2834 2835 2836
 2837 2838 2839 2840 2841 2842 2843 2844 2845 2846 2847 2848 2849 2850
 2851 2852 2853 2854 2855 2856 2857 2858 2859 2860 2861 2862 2863 2864
 2865 2866 2867 2868 2869 2870 2871 2872 2873 2874 2875 2876 2877 2878
 2879 2880 2881 2882 2883 2884 2885 2886 2887 2888 2889 2890 2891 2892
 2893 2894 2895 2896 2897 2898 2899 2900 2901 2902 2903 2904 2905 2906
 2907 2908 2909 2910 2911 2912 2913 2914 2915 2916 2917 2918 2919 2920
 2921 2922 2923 2924 2925 2926 2927 2928 2929 2930 2931 2932 2933 2934
 2935 2936 2937 2938 2939 2940 2941 2942 2943 2944 2945 2946 2947 2948
 2949 2950 2951 2952 2953 2954 2955 2956 2957 2958 2959 2960 2961 2962
 2963 2964 2965 2966 2967 2968 2969 2970 2971 2972 2973 2974 2975 2976
 2977 2978 2979 2980 2981 2982 2983 2984 2985 2986 2987 2988 2989 2990
 2991 2992 2993 2994 2995 2996 2997 2998 2999 3000 3001 3002 3003 3004
 3005 3006 3007 3008 3009 3010 3011 3012 3013 3014 3015 3016 3017 3018
 3019 3020 3021 3022 3023 3024 3025 3026 3027 3028 3029 3030 3031 3032
 3033 3034 3035 3036 3037 3038 3039 3040 3041 3042 3043 3044 3045 3046
 3047 3048 3049 3050 3051 3052 3053 3054 3055 3056 3057 3058 3059 3060
 3061 3062 3063 3064 3065 3066 3067 3068 3069 3070 3071 3072 3073 3074
 3075 3076 3077 3078 3079 3080 3081 3082 3083 3084 3085 3086 3087 3088
 3089 3090 3091 3092 3093 3094 3095 3096 3097 3098 3099 3100 3101 3102
 3103 3104 3105 3106 3107 3108 3109 3110 3111 3112 3113 3114 3115 3116
 3117 3118 3119 3120 3121 3122 3123 3124 3125 3126 3127 3128 3129 3130
 3131 3132 3133 3134 3135 3136 3137 3138 3139 3140 3141 3142 3143 3144
 3145 3146 3147 3148 3149 3150 3151 3152 3153 3154 3155 3156 3157 3158
 3159 3160 3161 3162 3163 3164 3165 3166 3167 3168 3169 3170 3171 3172
 3173 3174 3175 3176 3177 3178 3179 3180 3181 3182 3183 3184 3185 3186
 3187 3188 3189 3190 3191 3192 3193 3194 3195 3196 3197 3198 3199 3200
 3201 3202 3203 3204 3205 3206 3207 3208 3209 3210 3211 3212 3213 3214
 3215 3216 3217 3218 3219 3220 3221 3222 3223 3224 3225 3226 3227 3228
 3229 3230 3231 3232 3233 3234 3235 3236 3237 3238 3239 3240 3241 3242
 3243 3244 3245 3246 3247 3248 3249 3250 3251 3252 3253 3254 3255 3256
 3257 3258 3259 3260 3261 3262 3263 3264 3265 3266 3267 3268 3269 3270
 3271 3272 3273 3274 3275 3276 3277 3278 3279 3280 3281 3282 3283 3284
 3285 3286 3287 3288 3289 3290 3291 3292 3293 3294 3295 3296 3297 3298
 3299 3300 3301 3302 3303 3304 3305 3306 3307 3308 3309 3310 3311 3312
 3313 3314 3315 3316 3317 3318 3319 3320 3321 3322 3323 3324 3325 3326
 3327 3328 3329 3330 3331 3332 3333]
TRAIN: [   0    1    2 ... 3331 3332 3333] TEST: [3334 3335 3336 3337 3338 3339 3340 3341 3342 3343 3344
3345 3346 3347
 3348 3349 3350 3351 3352 3353 3354 3355 3356 3357 3358 3359 3360 3361
 3362 3363 3364 3365 3366 3367 3368 3369 3370 3371 3372 3373 3374 3375
 3376 3377 3378 3379 3380 3381 3382 3383 3384 3385 3386 3387 3388 3389
 3390 3391 3392 3393 3394 3395 3396 3397 3398 3399 3400 3401 3402 3403
 3404 3405 3406 3407 3408 3409 3410 3411 3412 3413 3414 3415 3416 3417
 3418 3419 3420 3421 3422 3423 3424 3425 3426 3427 3428 3429 3430 3431
 3432 3433 3434 3435 3436 3437 3438 3439 3440 3441 3442 3443 3444 3445
 3446 3447 3448 3449 3450 3451 3452 3453 3454 3455 3456 3457 3458 3459
 3460 3461 3462 3463 3464 3465 3466 3467 3468 3469 3470 3471 3472 3473
 3474 3475 3476 3477 3478 3479 3480 3481 3482 3483 3484 3485 3486 3487
 3488 3489 3490 3491 3492 3493 3494 3495 3496 3497 3498 3499 3500 3501
 3502 3503 3504 3505 3506 3507 3508 3509 3510 3511 3512 3513 3514 3515
 3516 3517 3518 3519 3520 3521 3522 3523 3524 3525 3526 3527 3528 3529
 3530 3531 3532 3533 3534 3535 3536 3537 3538 3539 3540 3541 3542 3543
 3544 3545 3546 3547 3548 3549 3550 3551 3552 3553 3554 3555 3556 3557
 3558 3559 3560 3561 3562 3563 3564 3565 3566 3567 3568 3569 3570 3571
 3572 3573 3574 3575 3576 3577 3578 3579 3580 3581 3582 3583 3584 3585
 3586 3587 3588 3589 3590 3591 3592 3593 3594 3595 3596 3597 3598 3599
 3600 3601 3602 3603 3604 3605 3606 3607 3608 3609 3610 3611 3612 3613
 3614 3615 3616 3617 3618 3619 3620 3621 3622 3623 3624 3625 3626 3627
 3628 3629 3630 3631 3632 3633 3634 3635 3636 3637 3638 3639 3640 3641
 3642 3643 3644 3645 3646 3647 3648 3649 3650 3651 3652 3653 3654 3655
 3656 3657 3658 3659 3660 3661 3662 3663 3664 3665 3666 3667 3668 3669
 3670 3671 3672 3673 3674 3675 3676 3677 3678 3679 3680 3681 3682 3683
 3684 3685 3686 3687 3688 3689 3690 3691 3692 3693 3694 3695 3696 3697
 3698 3699 3700 3701 3702 3703 3704 3705 3706 3707 3708 3709 3710 3711
 3712 3713 3714 3715 3716 3717 3718 3719 3720 3721 3722 3723 3724 3725
 3726 3727 3728 3729 3730 3731 3732 3733 3734 3735 3736 3737 3738 3739
 3740 3741 3742 3743 3744 3745 3746 3747 3748 3749 3750 3751 3752 3753
 3754 3755 3756 3757 3758 3759 3760 3761 3762 3763 3764 3765 3766 3767
 3768 3769 3770 3771 3772 3773 3774 3775 3776 3777 3778 3779 3780 3781
 3782 3783 3784 3785 3786 3787 3788 3789 3790 3791 3792 3793 3794 3795
```

```
 3796 3797 3798 3799 3800 3801 3802 3803 3804 3805 3806 3807 3808 3809
 3810 3811 3812 3813 3814 3815 3816 3817 3818 3819 3820 3821 3822 3823
 3824 3825 3826 3827 3828 3829 3830 3831 3832 3833 3834 3835 3836 3837
 3838 3839 3840 3841 3842 3843 3844 3845 3846 3847 3848 3849 3850 3851
 3852 3853 3854 3855 3856 3857 3858 3859 3860 3861 3862 3863 3864 3865
 3866 3867 3868 3869 3870 3871 3872 3873 3874 3875 3876 3877 3878 3879
 3880 3881 3882 3883 3884 3885 3886 3887 3888 3889 3890 3891 3892 3893
 3894 3895 3896 3897 3898 3899 3900 3901 3902 3903 3904 3905 3906 3907
 3908 3909 3910 3911 3912 3913 3914 3915 3916 3917 3918 3919 3920 3921
 3922 3923 3924 3925 3926 3927 3928 3929 3930 3931 3932 3933 3934 3935
 3936 3937 3938 3939 3940 3941 3942 3943 3944 3945 3946 3947 3948 3949
 3950 3951 3952 3953 3954 3955 3956 3957 3958 3959 3960 3961 3962 3963
 3964 3965 3966 3967 3968 3969 3970 3971 3972 3973 3974 3975 3976 3977
 3978 3979 3980 3981 3982 3983 3984 3985 3986 3987 3988 3989 3990 3991
 3992 3993 3994 3995 3996 3997 3998 3999 4000 4001 4002 4003 4004 4005
 4006 4007 4008 4009 4010 4011 4012 4013 4014 4015 4016 4017 4018 4019
 4020 4021 4022 4023 4024 4025 4026 4027 4028 4029 4030 4031 4032 4033
 4034 4035 4036 4037 4038 4039 4040 4041 4042 4043 4044 4045 4046 4047
 4048 4049 4050 4051 4052 4053 4054 4055 4056 4057 4058 4059 4060 4061
 4062 4063 4064 4065 4066 4067 4068 4069 4070 4071 4072 4073 4074 4075
 4076 4077 4078 4079 4080 4081 4082 4083 4084 4085 4086 4087 4088 4089
 4090 4091 4092 4093 4094 4095 4096 4097 4098 4099 4100 4101 4102 4103
 4104 4105 4106 4107 4108 4109 4110 4111 4112 4113 4114 4115 4116 4117
 4118 4119 4120 4121 4122 4123 4124 4125 4126 4127 4128 4129 4130 4131
 4132 4133 4134 4135 4136 4137 4138 4139 4140 4141 4142 4143 4144 4145
 4146 4147 4148 4149 4150 4151 4152 4153 4154 4155 4156 4157 4158 4159
 4160 4161 4162 4163 4164 4165 4166]
TRAIN: [   0    1    2 ... 4164 4165 4166] TEST: [4167 4168 4169 4170 4171 4172 4173 4174 4175 4176 4177
4178 4179 4180
 4181 4182 4183 4184 4185 4186 4187 4188 4189 4190 4191 4192 4193 4194
 4195 4196 4197 4198 4199 4200 4201 4202 4203 4204 4205 4206 4207 4208
 4209 4210 4211 4212 4213 4214 4215 4216 4217 4218 4219 4220 4221 4222
 4223 4224 4225 4226 4227 4228 4229 4230 4231 4232 4233 4234 4235 4236
 4237 4238 4239 4240 4241 4242 4243 4244 4245 4246 4247 4248 4249 4250
 4251 4252 4253 4254 4255 4256 4257 4258 4259 4260 4261 4262 4263 4264
 4265 4266 4267 4268 4269 4270 4271 4272 4273 4274 4275 4276 4277 4278
 4279 4280 4281 4282 4283 4284 4285 4286 4287 4288 4289 4290 4291 4292
 4293 4294 4295 4296 4297 4298 4299 4300 4301 4302 4303 4304 4305 4306
 4307 4308 4309 4310 4311 4312 4313 4314 4315 4316 4317 4318 4319 4320
 4321 4322 4323 4324 4325 4326 4327 4328 4329 4330 4331 4332 4333 4334
 4335 4336 4337 4338 4339 4340 4341 4342 4343 4344 4345 4346 4347 4348
 4349 4350 4351 4352 4353 4354 4355 4356 4357 4358 4359 4360 4361 4362
 4363 4364 4365 4366 4367 4368 4369 4370 4371 4372 4373 4374 4375 4376
 4377 4378 4379 4380 4381 4382 4383 4384 4385 4386 4387 4388 4389 4390
 4391 4392 4393 4394 4395 4396 4397 4398 4399 4400 4401 4402 4403 4404
 4405 4406 4407 4408 4409 4410 4411 4412 4413 4414 4415 4416 4417 4418
 4419 4420 4421 4422 4423 4424 4425 4426 4427 4428 4429 4430 4431 4432
 4433 4434 4435 4436 4437 4438 4439 4440 4441 4442 4443 4444 4445 4446
 4447 4448 4449 4450 4451 4452 4453 4454 4455 4456 4457 4458 4459 4460
 4461 4462 4463 4464 4465 4466 4467 4468 4469 4470 4471 4472 4473 4474
 4475 4476 4477 4478 4479 4480 4481 4482 4483 4484 4485 4486 4487 4488
 4489 4490 4491 4492 4493 4494 4495 4496 4497 4498 4499 4500 4501 4502
 4503 4504 4505 4506 4507 4508 4509 4510 4511 4512 4513 4514 4515 4516
 4517 4518 4519 4520 4521 4522 4523 4524 4525 4526 4527 4528 4529 4530
 4531 4532 4533 4534 4535 4536 4537 4538 4539 4540 4541 4542 4543 4544
 4545 4546 4547 4548 4549 4550 4551 4552 4553 4554 4555 4556 4557 4558
 4559 4560 4561 4562 4563 4564 4565 4566 4567 4568 4569 4570 4571 4572
 4573 4574 4575 4576 4577 4578 4579 4580 4581 4582 4583 4584 4585 4586
 4587 4588 4589 4590 4591 4592 4593 4594 4595 4596 4597 4598 4599 4600
 4601 4602 4603 4604 4605 4606 4607 4608 4609 4610 4611 4612 4613 4614
 4615 4616 4617 4618 4619 4620 4621 4622 4623 4624 4625 4626 4627 4628
 4629 4630 4631 4632 4633 4634 4635 4636 4637 4638 4639 4640 4641 4642
 4643 4644 4645 4646 4647 4648 4649 4650 4651 4652 4653 4654 4655 4656
 4657 4658 4659 4660 4661 4662 4663 4664 4665 4666 4667 4668 4669 4670
 4671 4672 4673 4674 4675 4676 4677 4678 4679 4680 4681 4682 4683 4684
 4685 4686 4687 4688 4689 4690 4691 4692 4693 4694 4695 4696 4697 4698
 4699 4700 4701 4702 4703 4704 4705 4706 4707 4708 4709 4710 4711 4712
 4713 4714 4715 4716 4717 4718 4719 4720 4721 4722 4723 4724 4725 4726
 4727 4728 4729 4730 4731 4732 4733 4734 4735 4736 4737 4738 4739 4740
 4741 4742 4743 4744 4745 4746 4747 4748 4749 4750 4751 4752 4753 4754
 4755 4756 4757 4758 4759 4760 4761 4762 4763 4764 4765 4766 4767 4768
 4769 4770 4771 4772 4773 4774 4775 4776 4777 4778 4779 4780 4781 4782
 4783 4784 4785 4786 4787 4788 4789 4790 4791 4792 4793 4794 4795 4796
 4797 4798 4799 4800 4801 4802 4803 4804 4805 4806 4807 4808 4809 4810
 4811 4812 4813 4814 4815 4816 4817 4818 4819 4820 4821 4822 4823 4824
 4825 4826 4827 4828 4829 4830 4831 4832 4833 4834 4835 4836 4837 4838
 4839 4840 4841 4842 4843 4844 4845 4846 4847 4848 4849 4850 4851 4852
 4853 4854 4855 4856 4857 4858 4859 4860 4861 4862 4863 4864 4865 4866
 4867 4868 4869 4870 4871 4872 4873 4874 4875 4876 4877 4878 4879 4880
 4881 4882 4883 4884 4885 4886 4887 4888 4889 4890 4891 4892 4893 4894
 4895 4896 4897 4898 4899 4900 4901 4902 4903 4904 4905 4906 4907 4908
 4909 4910 4911 4912 4913 4914 4915 4916 4917 4918 4919 4920 4921 4922
 4923 4924 4925 4926 4927 4928 4929 4930 4931 4932 4933 4934 4935 4936
 4937 4938 4939 4940 4941 4942 4943 4944 4945 4946 4947 4948 4949 4950
 4951 4952 4953 4954 4955 4956 4957 4958 4959 4960 4961 4962 4963 4964
 4965 4966 4967 4968 4969 4970 4971 4972 4973 4974 4975 4976 4977 4978
 4979 4980 4981 4982 4983 4984 4985 4986 4987 4988 4989 4990 4991 4992
 4993 4994 4995 4996 4997 4998 4999]
```

In [126]:
```
print(X_AW2V_tr.shape)
print(X_AW2V_test.shape)
print(Y_AW2V_tr.shape)
print(Y_AW2V_test.shape)
```

```
(4167, 4)
(833, 4)
(4167,)
(833,)
```

In [127]:
```
# Splitting the Train dataset into Cross Validation & Train Datasets
for train_index, test_index in tscv.split(X_AW2V_tr):
   print("TRAIN:", train_index, "TEST:", test_index)
   X_AW2V_train, X_AW2V_cv = X[train_index], X[test_index]
   Y_AW2V_train, Y_AW2V_cv = Y[train_index], Y[test_index]
```

```
TRAIN: [  0   1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17
  18  19  20  21  22  23  24  25  26  27  28  29  30  31  32  33  34  35
  36  37  38  39  40  41  42  43  44  45  46  47  48  49  50  51  52  53
  54  55  56  57  58  59  60  61  62  63  64  65  66  67  68  69  70  71
  72  73  74  75  76  77  78  79  80  81  82  83  84  85  86  87  88  89
  90  91  92  93  94  95  96  97  98  99 100 101 102 103 104 105 106 107
 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125
 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143
 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161
 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179
 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197
 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215
 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233
 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251
 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269
 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287
 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305
 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323
 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341
 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359
 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377
 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395
 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413
 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431
 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449
 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467
 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485
 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503
 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521
 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539
 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557
 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575
 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593
 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611
 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629
 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647
 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665
 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683
 684 685 686 687 688 689 690 691 692 693 694 695 696] TEST: [ 697  698  699  700  701  702  703  704  705
  706  707  708  709  710
  711  712  713  714  715  716  717  718  719  720  721  722  723  724
  725  726  727  728  729  730  731  732  733  734  735  736  737  738
  739  740  741  742  743  744  745  746  747  748  749  750  751  752
  753  754  755  756  757  758  759  760  761  762  763  764  765  766
  767  768  769  770  771  772  773  774  775  776  777  778  779  780
  781  782  783  784  785  786  787  788  789  790  791  792  793  794
  795  796  797  798  799  800  801  802  803  804  805  806  807  808
  809  810  811  812  813  814  815  816  817  818  819  820  821  822
  823  824  825  826  827  828  829  830  831  832  833  834  835  836
  837  838  839  840  841  842  843  844  845  846  847  848  849  850
  851  852  853  854  855  856  857  858  859  860  861  862  863  864
  865  866  867  868  869  870  871  872  873  874  875  876  877  878
  879  880  881  882  883  884  885  886  887  888  889  890  891  892
  893  894  895  896  897  898  899  900  901  902  903  904  905  906
  907  908  909  910  911  912  913  914  915  916  917  918  919  920
  921  922  923  924  925  926  927  928  929  930  931  932  933  934
  935  936  937  938  939  940  941  942  943  944  945  946  947  948
  949  950  951  952  953  954  955  956  957  958  959  960  961  962
  963  964  965  966  967  968  969  970  971  972  973  974  975  976
  977  978  979  980  981  982  983  984  985  986  987  988  989  990
  991  992  993  994  995  996  997  998  999 1000 1001 1002 1003 1004
 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018
 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032
 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046
 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060
 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074
 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088
 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102
 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116
 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128 1129 1130
 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142 1143 1144
 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158
 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172
 1173 1174 1175 1176 1177 1178 1179 1180 1181 1182 1183 1184 1185 1186
 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198 1199 1200
 1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212 1213 1214
 1215 1216 1217 1218 1219 1220 1221 1222 1223 1224 1225 1226 1227 1228
 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240 1241 1242
 1243 1244 1245 1246 1247 1248 1249 1250 1251 1252 1253 1254 1255 1256
 1257 1258 1259 1260 1261 1262 1263 1264 1265 1266 1267 1268 1269 1270
 1271 1272 1273 1274 1275 1276 1277 1278 1279 1280 1281 1282 1283 1284
 1285 1286 1287 1288 1289 1290 1291 1292 1293 1294 1295 1296 1297 1298
 1299 1300 1301 1302 1303 1304 1305 1306 1307 1308 1309 1310 1311 1312
 1313 1314 1315 1316 1317 1318 1319 1320 1321 1322 1323 1324 1325 1326
 1327 1328 1329 1330 1331 1332 1333 1334 1335 1336 1337 1338 1339 1340
 1341 1342 1343 1344 1345 1346 1347 1348 1349 1350 1351 1352 1353 1354
 1355 1356 1357 1358 1359 1360 1361 1362 1363 1364 1365 1366 1367 1368
```

```
 1369 1370 1371 1372 1373 1374 1375 1376 1377 1378 1379 1380 1381 1382
 1383 1384 1385 1386 1387 1388 1389 1390]
TRAIN: [   0    1    2 ... 1388 1389 1390] TEST: [1391 1392 1393 1394 1395 1396 1397 1398 1399 1400 1401
1402 1403 1404
 1405 1406 1407 1408 1409 1410 1411 1412 1413 1414 1415 1416 1417 1418
 1419 1420 1421 1422 1423 1424 1425 1426 1427 1428 1429 1430 1431 1432
 1433 1434 1435 1436 1437 1438 1439 1440 1441 1442 1443 1444 1445 1446
 1447 1448 1449 1450 1451 1452 1453 1454 1455 1456 1457 1458 1459 1460
 1461 1462 1463 1464 1465 1466 1467 1468 1469 1470 1471 1472 1473 1474
 1475 1476 1477 1478 1479 1480 1481 1482 1483 1484 1485 1486 1487 1488
 1489 1490 1491 1492 1493 1494 1495 1496 1497 1498 1499 1500 1501 1502
 1503 1504 1505 1506 1507 1508 1509 1510 1511 1512 1513 1514 1515 1516
 1517 1518 1519 1520 1521 1522 1523 1524 1525 1526 1527 1528 1529 1530
 1531 1532 1533 1534 1535 1536 1537 1538 1539 1540 1541 1542 1543 1544
 1545 1546 1547 1548 1549 1550 1551 1552 1553 1554 1555 1556 1557 1558
 1559 1560 1561 1562 1563 1564 1565 1566 1567 1568 1569 1570 1571 1572
 1573 1574 1575 1576 1577 1578 1579 1580 1581 1582 1583 1584 1585 1586
 1587 1588 1589 1590 1591 1592 1593 1594 1595 1596 1597 1598 1599 1600
 1601 1602 1603 1604 1605 1606 1607 1608 1609 1610 1611 1612 1613 1614
 1615 1616 1617 1618 1619 1620 1621 1622 1623 1624 1625 1626 1627 1628
 1629 1630 1631 1632 1633 1634 1635 1636 1637 1638 1639 1640 1641 1642
 1643 1644 1645 1646 1647 1648 1649 1650 1651 1652 1653 1654 1655 1656
 1657 1658 1659 1660 1661 1662 1663 1664 1665 1666 1667 1668 1669 1670
 1671 1672 1673 1674 1675 1676 1677 1678 1679 1680 1681 1682 1683 1684
 1685 1686 1687 1688 1689 1690 1691 1692 1693 1694 1695 1696 1697 1698
 1699 1700 1701 1702 1703 1704 1705 1706 1707 1708 1709 1710 1711 1712
 1713 1714 1715 1716 1717 1718 1719 1720 1721 1722 1723 1724 1725 1726
 1727 1728 1729 1730 1731 1732 1733 1734 1735 1736 1737 1738 1739 1740
 1741 1742 1743 1744 1745 1746 1747 1748 1749 1750 1751 1752 1753 1754
 1755 1756 1757 1758 1759 1760 1761 1762 1763 1764 1765 1766 1767 1768
 1769 1770 1771 1772 1773 1774 1775 1776 1777 1778 1779 1780 1781 1782
 1783 1784 1785 1786 1787 1788 1789 1790 1791 1792 1793 1794 1795 1796
 1797 1798 1799 1800 1801 1802 1803 1804 1805 1806 1807 1808 1809 1810
 1811 1812 1813 1814 1815 1816 1817 1818 1819 1820 1821 1822 1823 1824
 1825 1826 1827 1828 1829 1830 1831 1832 1833 1834 1835 1836 1837 1838
 1839 1840 1841 1842 1843 1844 1845 1846 1847 1848 1849 1850 1851 1852
 1853 1854 1855 1856 1857 1858 1859 1860 1861 1862 1863 1864 1865 1866
 1867 1868 1869 1870 1871 1872 1873 1874 1875 1876 1877 1878 1879 1880
 1881 1882 1883 1884 1885 1886 1887 1888 1889 1890 1891 1892 1893 1894
 1895 1896 1897 1898 1899 1900 1901 1902 1903 1904 1905 1906 1907 1908
 1909 1910 1911 1912 1913 1914 1915 1916 1917 1918 1919 1920 1921 1922
 1923 1924 1925 1926 1927 1928 1929 1930 1931 1932 1933 1934 1935 1936
 1937 1938 1939 1940 1941 1942 1943 1944 1945 1946 1947 1948 1949 1950
 1951 1952 1953 1954 1955 1956 1957 1958 1959 1960 1961 1962 1963 1964
 1965 1966 1967 1968 1969 1970 1971 1972 1973 1974 1975 1976 1977 1978
 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991 1992
 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006
 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020
 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034
 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048
 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062
 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076
 2077 2078 2079 2080 2081 2082 2083 2084]
TRAIN: [   0    1    2 ... 2082 2083 2084] TEST: [2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095
2096 2097 2098
 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112
 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126
 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140
 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154
 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168
 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182
 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196
 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210
 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224
 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238
 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252
 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266
 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280
 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294
 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308
 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322
 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336
 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350
 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364
 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378
 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392
 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406
 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420
 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434
 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448
 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462
 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476
 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490
 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504
 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518
 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532
 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546
 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560
```

```
 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574
 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584 2585 2586 2587 2588
 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598 2599 2600 2601 2602
 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612 2613 2614 2615 2616
 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626 2627 2628 2629 2630
 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640 2641 2642 2643 2644
 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654 2655 2656 2657 2658
 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668 2669 2670 2671 2672
 2673 2674 2675 2676 2677 2678 2679 2680 2681 2682 2683 2684 2685 2686
 2687 2688 2689 2690 2691 2692 2693 2694 2695 2696 2697 2698 2699 2700
 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710 2711 2712 2713 2714
 2715 2716 2717 2718 2719 2720 2721 2722 2723 2724 2725 2726 2727 2728
 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738 2739 2740 2741 2742
 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752 2753 2754 2755 2756
 2757 2758 2759 2760 2761 2762 2763 2764 2765 2766 2767 2768 2769 2770
 2771 2772 2773 2774 2775 2776 2777 2778]
TRAIN: [   0    1    2 ... 2776 2777 2778] TEST: [2779 2780 2781 2782 2783 2784 2785 2786 2787 2788 2789
2790 2791 2792
 2793 2794 2795 2796 2797 2798 2799 2800 2801 2802 2803 2804 2805 2806
 2807 2808 2809 2810 2811 2812 2813 2814 2815 2816 2817 2818 2819 2820
 2821 2822 2823 2824 2825 2826 2827 2828 2829 2830 2831 2832 2833 2834
 2835 2836 2837 2838 2839 2840 2841 2842 2843 2844 2845 2846 2847 2848
 2849 2850 2851 2852 2853 2854 2855 2856 2857 2858 2859 2860 2861 2862
 2863 2864 2865 2866 2867 2868 2869 2870 2871 2872 2873 2874 2875 2876
 2877 2878 2879 2880 2881 2882 2883 2884 2885 2886 2887 2888 2889 2890
 2891 2892 2893 2894 2895 2896 2897 2898 2899 2900 2901 2902 2903 2904
 2905 2906 2907 2908 2909 2910 2911 2912 2913 2914 2915 2916 2917 2918
 2919 2920 2921 2922 2923 2924 2925 2926 2927 2928 2929 2930 2931 2932
 2933 2934 2935 2936 2937 2938 2939 2940 2941 2942 2943 2944 2945 2946
 2947 2948 2949 2950 2951 2952 2953 2954 2955 2956 2957 2958 2959 2960
 2961 2962 2963 2964 2965 2966 2967 2968 2969 2970 2971 2972 2973 2974
 2975 2976 2977 2978 2979 2980 2981 2982 2983 2984 2985 2986 2987 2988
 2989 2990 2991 2992 2993 2994 2995 2996 2997 2998 2999 3000 3001 3002
 3003 3004 3005 3006 3007 3008 3009 3010 3011 3012 3013 3014 3015 3016
 3017 3018 3019 3020 3021 3022 3023 3024 3025 3026 3027 3028 3029 3030
 3031 3032 3033 3034 3035 3036 3037 3038 3039 3040 3041 3042 3043 3044
 3045 3046 3047 3048 3049 3050 3051 3052 3053 3054 3055 3056 3057 3058
 3059 3060 3061 3062 3063 3064 3065 3066 3067 3068 3069 3070 3071 3072
 3073 3074 3075 3076 3077 3078 3079 3080 3081 3082 3083 3084 3085 3086
 3087 3088 3089 3090 3091 3092 3093 3094 3095 3096 3097 3098 3099 3100
 3101 3102 3103 3104 3105 3106 3107 3108 3109 3110 3111 3112 3113 3114
 3115 3116 3117 3118 3119 3120 3121 3122 3123 3124 3125 3126 3127 3128
 3129 3130 3131 3132 3133 3134 3135 3136 3137 3138 3139 3140 3141 3142
 3143 3144 3145 3146 3147 3148 3149 3150 3151 3152 3153 3154 3155 3156
 3157 3158 3159 3160 3161 3162 3163 3164 3165 3166 3167 3168 3169 3170
 3171 3172 3173 3174 3175 3176 3177 3178 3179 3180 3181 3182 3183 3184
 3185 3186 3187 3188 3189 3190 3191 3192 3193 3194 3195 3196 3197 3198
 3199 3200 3201 3202 3203 3204 3205 3206 3207 3208 3209 3210 3211 3212
 3213 3214 3215 3216 3217 3218 3219 3220 3221 3222 3223 3224 3225 3226
 3227 3228 3229 3230 3231 3232 3233 3234 3235 3236 3237 3238 3239 3240
 3241 3242 3243 3244 3245 3246 3247 3248 3249 3250 3251 3252 3253 3254
 3255 3256 3257 3258 3259 3260 3261 3262 3263 3264 3265 3266 3267 3268
 3269 3270 3271 3272 3273 3274 3275 3276 3277 3278 3279 3280 3281 3282
 3283 3284 3285 3286 3287 3288 3289 3290 3291 3292 3293 3294 3295 3296
 3297 3298 3299 3300 3301 3302 3303 3304 3305 3306 3307 3308 3309 3310
 3311 3312 3313 3314 3315 3316 3317 3318 3319 3320 3321 3322 3323 3324
 3325 3326 3327 3328 3329 3330 3331 3332 3333 3334 3335 3336 3337 3338
 3339 3340 3341 3342 3343 3344 3345 3346 3347 3348 3349 3350 3351 3352
 3353 3354 3355 3356 3357 3358 3359 3360 3361 3362 3363 3364 3365 3366
 3367 3368 3369 3370 3371 3372 3373 3374 3375 3376 3377 3378 3379 3380
 3381 3382 3383 3384 3385 3386 3387 3388 3389 3390 3391 3392 3393 3394
 3395 3396 3397 3398 3399 3400 3401 3402 3403 3404 3405 3406 3407 3408
 3409 3410 3411 3412 3413 3414 3415 3416 3417 3418 3419 3420 3421 3422
 3423 3424 3425 3426 3427 3428 3429 3430 3431 3432 3433 3434 3435 3436
 3437 3438 3439 3440 3441 3442 3443 3444 3445 3446 3447 3448 3449 3450
 3451 3452 3453 3454 3455 3456 3457 3458 3459 3460 3461 3462 3463 3464
 3465 3466 3467 3468 3469 3470 3471 3472]
TRAIN: [   0    1    2 ... 3470 3471 3472] TEST: [3473 3474 3475 3476 3477 3478 3479 3480 3481 3482 3483
3484 3485 3486
 3487 3488 3489 3490 3491 3492 3493 3494 3495 3496 3497 3498 3499 3500
 3501 3502 3503 3504 3505 3506 3507 3508 3509 3510 3511 3512 3513 3514
 3515 3516 3517 3518 3519 3520 3521 3522 3523 3524 3525 3526 3527 3528
 3529 3530 3531 3532 3533 3534 3535 3536 3537 3538 3539 3540 3541 3542
 3543 3544 3545 3546 3547 3548 3549 3550 3551 3552 3553 3554 3555 3556
 3557 3558 3559 3560 3561 3562 3563 3564 3565 3566 3567 3568 3569 3570
 3571 3572 3573 3574 3575 3576 3577 3578 3579 3580 3581 3582 3583 3584
 3585 3586 3587 3588 3589 3590 3591 3592 3593 3594 3595 3596 3597 3598
 3599 3600 3601 3602 3603 3604 3605 3606 3607 3608 3609 3610 3611 3612
 3613 3614 3615 3616 3617 3618 3619 3620 3621 3622 3623 3624 3625 3626
 3627 3628 3629 3630 3631 3632 3633 3634 3635 3636 3637 3638 3639 3640
 3641 3642 3643 3644 3645 3646 3647 3648 3649 3650 3651 3652 3653 3654
 3655 3656 3657 3658 3659 3660 3661 3662 3663 3664 3665 3666 3667 3668
 3669 3670 3671 3672 3673 3674 3675 3676 3677 3678 3679 3680 3681 3682
 3683 3684 3685 3686 3687 3688 3689 3690 3691 3692 3693 3694 3695 3696
 3697 3698 3699 3700 3701 3702 3703 3704 3705 3706 3707 3708 3709 3710
 3711 3712 3713 3714 3715 3716 3717 3718 3719 3720 3721 3722 3723 3724
 3725 3726 3727 3728 3729 3730 3731 3732 3733 3734 3735 3736 3737 3738
 3739 3740 3741 3742 3743 3744 3745 3746 3747 3748 3749 3750 3751 3752
```

```
 3753 3754 3755 3756 3757 3758 3759 3760 3761 3762 3763 3764 3765 3766
 3767 3768 3769 3770 3771 3772 3773 3774 3775 3776 3777 3778 3779 3780
 3781 3782 3783 3784 3785 3786 3787 3788 3789 3790 3791 3792 3793 3794
 3795 3796 3797 3798 3799 3800 3801 3802 3803 3804 3805 3806 3807 3808
 3809 3810 3811 3812 3813 3814 3815 3816 3817 3818 3819 3820 3821 3822
 3823 3824 3825 3826 3827 3828 3829 3830 3831 3832 3833 3834 3835 3836
 3837 3838 3839 3840 3841 3842 3843 3844 3845 3846 3847 3848 3849 3850
 3851 3852 3853 3854 3855 3856 3857 3858 3859 3860 3861 3862 3863 3864
 3865 3866 3867 3868 3869 3870 3871 3872 3873 3874 3875 3876 3877 3878
 3879 3880 3881 3882 3883 3884 3885 3886 3887 3888 3889 3890 3891 3892
 3893 3894 3895 3896 3897 3898 3899 3900 3901 3902 3903 3904 3905 3906
 3907 3908 3909 3910 3911 3912 3913 3914 3915 3916 3917 3918 3919 3920
 3921 3922 3923 3924 3925 3926 3927 3928 3929 3930 3931 3932 3933 3934
 3935 3936 3937 3938 3939 3940 3941 3942 3943 3944 3945 3946 3947 3948
 3949 3950 3951 3952 3953 3954 3955 3956 3957 3958 3959 3960 3961 3962
 3963 3964 3965 3966 3967 3968 3969 3970 3971 3972 3973 3974 3975 3976
 3977 3978 3979 3980 3981 3982 3983 3984 3985 3986 3987 3988 3989 3990
 3991 3992 3993 3994 3995 3996 3997 3998 3999 4000 4001 4002 4003 4004
 4005 4006 4007 4008 4009 4010 4011 4012 4013 4014 4015 4016 4017 4018
 4019 4020 4021 4022 4023 4024 4025 4026 4027 4028 4029 4030 4031 4032
 4033 4034 4035 4036 4037 4038 4039 4040 4041 4042 4043 4044 4045 4046
 4047 4048 4049 4050 4051 4052 4053 4054 4055 4056 4057 4058 4059 4060
 4061 4062 4063 4064 4065 4066 4067 4068 4069 4070 4071 4072 4073 4074
 4075 4076 4077 4078 4079 4080 4081 4082 4083 4084 4085 4086 4087 4088
 4089 4090 4091 4092 4093 4094 4095 4096 4097 4098 4099 4100 4101 4102
 4103 4104 4105 4106 4107 4108 4109 4110 4111 4112 4113 4114 4115 4116
 4117 4118 4119 4120 4121 4122 4123 4124 4125 4126 4127 4128 4129 4130
 4131 4132 4133 4134 4135 4136 4137 4138 4139 4140 4141 4142 4143 4144
 4145 4146 4147 4148 4149 4150 4151 4152 4153 4154 4155 4156 4157 4158
 4159 4160 4161 4162 4163 4164 4165 4166]
```

In [128]:
```
print(X_AW2V_train.shape)
print(X_AW2V_cv.shape)
print(X_AW2V_test.shape)

print(Y_AW2V_train.shape)
print(Y_AW2V_cv.shape)
print(Y_AW2V_test.shape)
```

```
(3473, 4)
(694, 4)
(833, 4)
(3473,)
(694,)
(833,)
```

In [129]:
```
# Removing the Date variable from the Data sets
AW2V_X= []
for each in X_AW2V_train:
    AW2V_X.append(np.delete(each,3))
X_AW2V_train = np.asarray(AW2V_X)


AW2V_X_cv= []
for each in X_AW2V_cv:
    AW2V_X_cv.append(np.delete(each,3))
X_AW2V_cv = np.asarray(AW2V_X_cv)

AW2V_X_test= []
for each in X_AW2V_test:
    AW2V_X_test.append(np.delete(each,3))
X_AW2V_test = np.asarray(AW2V_X_test)
```

In [130]:
```
print(X_AW2V_train.shape)
print(X_AW2V_cv.shape)
print(X_AW2V_test.shape)

print(Y_AW2V_train.shape)
print(Y_AW2V_cv.shape)
print(Y_AW2V_test.shape)
```

```
(3473, 3)
(694, 3)
(833, 3)
(3473,)
(694,)
(833,)
```

In [131]:
```
# Finding the Test accuracy using Cross Validation Data Set
for i in range(1,30,2):
    # instantiate learning model (k = 30)
    knn = KNeighborsClassifier(n_neighbors=i)

    # fitting the model on crossvalidation train
    knn.fit(X_AW2V_train, Y_AW2V_train)

    # predict the response on the crossvalidation train
    pred = knn.predict(X_AW2V_cv)

    # evaluate CV accuracy
    acc = accuracy_score(Y_AW2V_cv, pred, normalize=True) * float(100)
    print('\nCV accuracy for k = %d is %d%%' % (i, acc))

knn = KNeighborsClassifier(1)
knn.fit(X_AW2V_train,Y_AW2V_train)
pred = knn.predict(X_AW2V_test)
acc = accuracy_score(Y_AW2V_test, pred, normalize=True) * float(100)
print('\n****Test accuracy for k = 1 is %d%%' % (acc))
```

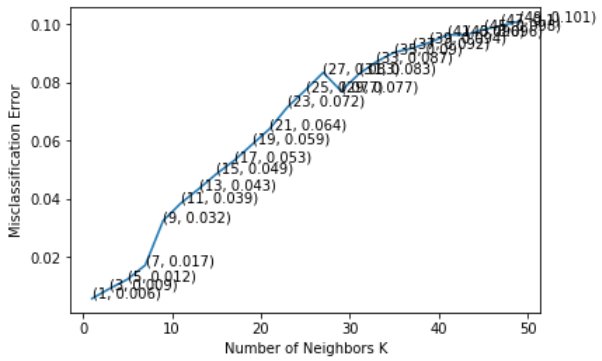
```
CV accuracy for k = 1 is 98%

CV accuracy for k = 3 is 98%

CV accuracy for k = 5 is 97%

CV accuracy for k = 7 is 96%

CV accuracy for k = 9 is 96%

CV accuracy for k = 11 is 96%

CV accuracy for k = 13 is 96%

CV accuracy for k = 15 is 95%

CV accuracy for k = 17 is 95%

CV accuracy for k = 19 is 94%

CV accuracy for k = 21 is 94%

CV accuracy for k = 23 is 93%

CV accuracy for k = 25 is 92%

CV accuracy for k = 27 is 89%

CV accuracy for k = 29 is 88%

****Test accuracy for k = 1 is 99%
```

In [135]:
```
# creating odd list of K for KNN
myList = list(range(0,50))
neighbors = list(filter(lambda x: x % 2 != 0, myList))

# empty list that will hold cv scores
cv_scores = []
```

In [136]:
```
for k in neighbors:
    knn = KNeighborsClassifier(n_neighbors=k)
    scores = cross_val_score(knn, X_AW2V_train, Y_AW2V_train, cv=10)
    cv_scores.append(scores.mean())
```

In [137]:
```
# changing to misclassification error
MSE = [1 - x for x in cv_scores]

# determining best k
optimal_k = neighbors[MSE.index(min(MSE))]
print('\nThe optimal number of neighbors is %d.' % optimal_k)

# plot misclassification error vs k
plt.plot(neighbors, MSE)

for xy in zip(neighbors, np.round(MSE,3)):
    plt.annotate('(%s, %s)' % xy, xy=xy, textcoords='data')

plt.xlabel('Number of Neighbors K')
plt.ylabel('Misclassification Error')
plt.show()

print("the misclassification error for each k value is : ", np.round(MSE,3))
```

```
The optimal number of neighbors is 1.
```




```
the misclassification error for each k value is :  [0.006 0.009 0.012 0.017 0.032 0.039 0.043 0.049 0.053
0.059 0.064 0.072
 0.077 0.083 0.077 0.083 0.087 0.09  0.092 0.094 0.096 0.096 0.098 0.1
 0.101]
```

In [138]:
```
# Finding Optimal K from a set of odd List of Integers
# instantiate learning model k = optimal_k
knn_optimal = KNeighborsClassifier(n_neighbors=optimal_k)

# fitting the model
knn_optimal.fit(X_AW2V_train, Y_AW2V_train)

# predict the response
pred = knn_optimal.predict(X_AW2V_test)

# evaluate accuracy
acc = accuracy_score(Y_AW2V_test, pred) * 100
print('\nThe accuracy of the knn classifier for k = %d is %f%%' % (optimal_k, acc))
```

```
The accuracy of the knn classifier for k = 1 is 99.639856%
```

TF-IDF - W2V

In [82]:
```
# TF-IDF weighted Word2Vec
tfidf_feat = tf_idf_vect.get_feature_names() # tfidf words/col-names
# final_tf_idf is the sparse matrix with row= sentence, col=word and cell_val = tfidf

tfidf_sent_vectors = []; # the tfidf-w2v for each sentence/review is stored in this list
row=0;
for sent in list_of_sent: # for each review/sentence
    sent_vec = np.zeros(50) # as word vectors are of zero length
    weight_sum =0; # num of words with a valid vector in the sentence/review
    for word in sent: # for each word in a review/sentence
        if word in w2v_words:
            vec = w2v_model.wv[word]
            # obtain the tf_idfidf of a word in a sentence/review
            tf_idf = final_tf_idf[row, tfidf_feat.index(word)]
            sent_vec += (vec * tf_idf)
            weight_sum += tf_idf
    if weight_sum != 0:
        sent_vec /= weight_sum
    tfidf_sent_vectors.append(sent_vec)
    row += 1
```

In [101]:
```
#tsne
start_time_code = time.time()
model = TSNE(n_components=2, random_state=0, perplexity=30)
tsne_data_tfidf_Aw2v = model.fit_transform(tfidf_sent_vectors)
end_time_code = time.time()
print ("Running Time for code execution " + str(end_time_code - start_time_code) + " secs")
```

```
Running Time for code execution 102.52370500564575 secs
```

In [139]:
```
tsne_data_tfidf_Aw2v.shape
```

Out[139]:
```
(5000, 2)
```

In [140]:
```
tsne_data_tf_Aw2v = np.vstack((tsne_data_tfidf_Aw2v.T,label,Time)).T
```

In [141]:
```
tsne_data_tf_Aw2v.shape
```

Out[141]:
```
(5000, 4)
```

In [142]:
```
#Saving tsne data into seperate file for further usecases.
tsne_Tf_AW2V_df = pd.DataFrame(data=tsne_data_tf_Aw2v, columns=("X", "Y", "Label","Date"))
tsne_Tf_AW2V_df.to_csv('/home/venkatasujit272/tsne_Tf_AW2V_data.csv')
```

In [143]:
```
# Aliasing the Score field values into Binary format
def partition(x):
    if x == 'negative':
        return 0.
    return 1.

# Aliasing the Score to String Format (Positive for Score > 3 & Negative for Score  < 3)

actualScore = tsne_Tf_AW2V_df['Label']
positiveNegative = actualScore.map(partition)
tsne_Tf_AW2V_df['Label'] = positiveNegative
```

In [144]:
```
tsne_Tf_AW2V_df.head(1)
```

Out[144]:

	X	Y	Label	Date
0	-2.5246	-10.7039	1.0	1999-10-08

In [145]:
```
#Sorting the Dataframe with Date to apply Timebased Split
tsne_Tf_AW2V_df.sort_values(by='Date',ascending=1)
```

Out[145]:

	X	Y	Label	Date
0	-2.5246	-10.7039	1.0	1999-10-08
28	-3.36253	-9.00849	1.0	1999-10-25
377	3.91971	-10.5973	1.0	1999-12-02
287	-4.60195	-11.4588	1.0	1999-12-06
376	1.9589	-2.21929	1.0	2000-01-03
209	2.61388	-2.75753	1.0	2000-01-09
2540	3.93277	-10.1787	0.0	2000-01-19
198	9.37026	-8.3316	1.0	2000-01-24
224	-7.84105	-2.76157	1.0	2000-02-26
284	8.18882	-3.9679	1.0	2000-06-03
378	7.18127	-4.67793	1.0	2000-06-03
205	-2.50795	-11.1343	1.0	2000-06-23
206	-4.70862	-11.6129	1.0	2000-06-29
415	11.3001	-5.85578	1.0	2000-07-31
684	6.40004	9.04671	1.0	2000-08-09
714	5.35139	10.0547	1.0	2000-08-09
212	1.94638	-2.10341	1.0	2000-08-15
258	3.55155	-10.6195	1.0	2000-10-03
691	1.09665	-6.80669	1.0	2000-12-05
315	1.88862	-2.23293	1.0	2000-12-19
286	1.61452	-11.2988	1.0	2000-12-30
692	8.08022	6.12519	1.0	2001-02-22
2547	-2.35202	-12.4363	0.0	2001-06-11
2543	3.7056	-10.4436	0.0	2001-08-08
233	1.99856	-10.2072	1.0	2001-09-24
701	7.51727	4.46549	1.0	2001-10-23
309	3.58549	-10.5966	1.0	2001-10-26
30	-2.37335	-6.28845	1.0	2001-12-26
830	9.13102	-3.75512	1.0	2002-01-06
288	-3.73461	-11.9175	1.0	2002-02-04
...	...	...	...	...
232	2.61009	-10.6303	1.0	2012-10-22
3856	-2.47455	-4.49752	0.0	2012-10-23
1792	-0.781262	-2.4867	1.0	2012-10-23
813	3.27652	4.65774	1.0	2012-10-23
3649	-0.606243	10.1302	0.0	2012-10-23
2030	11.3966	6.16946	1.0	2012-10-23
1897	8.59687	-4.20608	1.0	2012-10-23
1086	2.66058	0.475088	1.0	2012-10-23
3056	-0.870412	-12.1437	0.0	2012-10-23
1594	0.48707	6.28572	1.0	2012-10-23
4556	-4.63344	-5.06469	0.0	2012-10-23
134	-1.15171	7.04979	1.0	2012-10-23
2717	-0.138413	4.3474	0.0	2012-10-24
2705	-5.88004	0.477485	0.0	2012-10-24
1295	-2.43614	-2.72924	1.0	2012-10-24
2391	6.10425	9.29838	1.0	2012-10-24
1603	-2.66913	4.19046	1.0	2012-10-25
3849	3.78954	-0.0163171	0.0	2012-10-25
2715	1.8184	4.49809	0.0	2012-10-25
4115	-5.23026	-9.38369	0.0	2012-10-25
1827	-8.63288	6.96447	1.0	2012-10-25
2978	3.33108	-9.95436	0.0	2012-10-26

	X	Y	Label	Date
135	-2.33329	-9.65648	1.0	2012-10-26
782	-4.57583	-4.63614	1.0	2012-10-26
136	-5.42425	8.65321	1.0	2012-10-26
3569	-4.8137	-10.4036	0.0	2012-10-26
1349	-3.56628	5.03526	1.0	2012-10-26
493	12.6854	-1.58449	1.0	2012-10-26
4793	-2.07365	2.77399	0.0	2012-10-26
2592	9.60875	-8.42097	0.0	2012-10-26

5000 rows × 4 columns

In [146]:
```
# create design matrix X and target vector y
X_TW= np.array(tsne_Tf_AW2V_df.iloc[:, 0:4]) # end index is exclusive
Y_TW = np.array(tsne_Tf_AW2V_df['Label']) # showing you two ways of indexing a pandas df
```

In [147]:
```
tscv = TimeSeriesSplit(n_splits=5)
print(tscv)
```

```
TimeSeriesSplit(n_splits=5)
```

In [148]:
```
#Splitting the Dataset into Train set & Test set
for train_index, test_index in tscv.split(X_TW):
   print("TRAIN:", train_index, "TEST:", test_index)
   X_TW_tr, X_TW_test = X[train_index], X[test_index]
   Y_TW_tr, Y_TW_test = Y[train_index], Y[test_index]
```

```
TRAIN: [  0   1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17
  18  19  20  21  22  23  24  25  26  27  28  29  30  31  32  33  34  35
  36  37  38  39  40  41  42  43  44  45  46  47  48  49  50  51  52  53
  54  55  56  57  58  59  60  61  62  63  64  65  66  67  68  69  70  71
  72  73  74  75  76  77  78  79  80  81  82  83  84  85  86  87  88  89
  90  91  92  93  94  95  96  97  98  99 100 101 102 103 104 105 106 107
 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125
 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143
 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161
 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179
 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197
 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215
 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233
 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251
 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269
 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287
 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305
 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323
 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341
 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359
 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377
 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395
 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413
 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431
 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449
 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467
 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485
 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503
 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521
 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539
 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557
 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575
 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593
 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611
 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629
 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647
 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665
 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683
 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701
 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719
 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737
 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755
 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773
 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791
 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809
 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827
 828 829 830 831 832 833 834] TEST: [ 835  836  837  838  839  840  841  842  843  844  845  846  847  84
8
  849  850  851  852  853  854  855  856  857  858  859  860  861  862
  863  864  865  866  867  868  869  870  871  872  873  874  875  876
  877  878  879  880  881  882  883  884  885  886  887  888  889  890
  891  892  893  894  895  896  897  898  899  900  901  902  903  904
  905  906  907  908  909  910  911  912  913  914  915  916  917  918
  919  920  921  922  923  924  925  926  927  928  929  930  931  932
  933  934  935  936  937  938  939  940  941  942  943  944  945  946
  947  948  949  950  951  952  953  954  955  956  957  958  959  960
  961  962  963  964  965  966  967  968  969  970  971  972  973  974
  975  976  977  978  979  980  981  982  983  984  985  986  987  988
  989  990  991  992  993  994  995  996  997  998  999 1000 1001 1002
 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016
 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030
 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044
 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058
 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072
 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086
 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100
 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114
 1115 1116 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128
 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142
 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156
 1157 1158 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170
 1171 1172 1173 1174 1175 1176 1177 1178 1179 1180 1181 1182 1183 1184
 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198
 1199 1200 1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212
 1213 1214 1215 1216 1217 1218 1219 1220 1221 1222 1223 1224 1225 1226
 1227 1228 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240
 1241 1242 1243 1244 1245 1246 1247 1248 1249 1250 1251 1252 1253 1254
 1255 1256 1257 1258 1259 1260 1261 1262 1263 1264 1265 1266 1267 1268
 1269 1270 1271 1272 1273 1274 1275 1276 1277 1278 1279 1280 1281 1282
 1283 1284 1285 1286 1287 1288 1289 1290 1291 1292 1293 1294 1295 1296
 1297 1298 1299 1300 1301 1302 1303 1304 1305 1306 1307 1308 1309 1310
 1311 1312 1313 1314 1315 1316 1317 1318 1319 1320 1321 1322 1323 1324
 1325 1326 1327 1328 1329 1330 1331 1332 1333 1334 1335 1336 1337 1338
 1339 1340 1341 1342 1343 1344 1345 1346 1347 1348 1349 1350 1351 1352
 1353 1354 1355 1356 1357 1358 1359 1360 1361 1362 1363 1364 1365 1366
 1367 1368 1369 1370 1371 1372 1373 1374 1375 1376 1377 1378 1379 1380
 1381 1382 1383 1384 1385 1386 1387 1388 1389 1390 1391 1392 1393 1394
```

```
 1395 1396 1397 1398 1399 1400 1401 1402 1403 1404 1405 1406 1407 1408
 1409 1410 1411 1412 1413 1414 1415 1416 1417 1418 1419 1420 1421 1422
 1423 1424 1425 1426 1427 1428 1429 1430 1431 1432 1433 1434 1435 1436
 1437 1438 1439 1440 1441 1442 1443 1444 1445 1446 1447 1448 1449 1450
 1451 1452 1453 1454 1455 1456 1457 1458 1459 1460 1461 1462 1463 1464
 1465 1466 1467 1468 1469 1470 1471 1472 1473 1474 1475 1476 1477 1478
 1479 1480 1481 1482 1483 1484 1485 1486 1487 1488 1489 1490 1491 1492
 1493 1494 1495 1496 1497 1498 1499 1500 1501 1502 1503 1504 1505 1506
 1507 1508 1509 1510 1511 1512 1513 1514 1515 1516 1517 1518 1519 1520
 1521 1522 1523 1524 1525 1526 1527 1528 1529 1530 1531 1532 1533 1534
 1535 1536 1537 1538 1539 1540 1541 1542 1543 1544 1545 1546 1547 1548
 1549 1550 1551 1552 1553 1554 1555 1556 1557 1558 1559 1560 1561 1562
 1563 1564 1565 1566 1567 1568 1569 1570 1571 1572 1573 1574 1575 1576
 1577 1578 1579 1580 1581 1582 1583 1584 1585 1586 1587 1588 1589 1590
 1591 1592 1593 1594 1595 1596 1597 1598 1599 1600 1601 1602 1603 1604
 1605 1606 1607 1608 1609 1610 1611 1612 1613 1614 1615 1616 1617 1618
 1619 1620 1621 1622 1623 1624 1625 1626 1627 1628 1629 1630 1631 1632
 1633 1634 1635 1636 1637 1638 1639 1640 1641 1642 1643 1644 1645 1646
 1647 1648 1649 1650 1651 1652 1653 1654 1655 1656 1657 1658 1659 1660
 1661 1662 1663 1664 1665 1666 1667]
TRAIN: [   0    1    2 ... 1665 1666 1667] TEST: [1668 1669 1670 1671 1672 1673 1674 1675 1676 1677 1678
1679 1680 1681
 1682 1683 1684 1685 1686 1687 1688 1689 1690 1691 1692 1693 1694 1695
 1696 1697 1698 1699 1700 1701 1702 1703 1704 1705 1706 1707 1708 1709
 1710 1711 1712 1713 1714 1715 1716 1717 1718 1719 1720 1721 1722 1723
 1724 1725 1726 1727 1728 1729 1730 1731 1732 1733 1734 1735 1736 1737
 1738 1739 1740 1741 1742 1743 1744 1745 1746 1747 1748 1749 1750 1751
 1752 1753 1754 1755 1756 1757 1758 1759 1760 1761 1762 1763 1764 1765
 1766 1767 1768 1769 1770 1771 1772 1773 1774 1775 1776 1777 1778 1779
 1780 1781 1782 1783 1784 1785 1786 1787 1788 1789 1790 1791 1792 1793
 1794 1795 1796 1797 1798 1799 1800 1801 1802 1803 1804 1805 1806 1807
 1808 1809 1810 1811 1812 1813 1814 1815 1816 1817 1818 1819 1820 1821
 1822 1823 1824 1825 1826 1827 1828 1829 1830 1831 1832 1833 1834 1835
 1836 1837 1838 1839 1840 1841 1842 1843 1844 1845 1846 1847 1848 1849
 1850 1851 1852 1853 1854 1855 1856 1857 1858 1859 1860 1861 1862 1863
 1864 1865 1866 1867 1868 1869 1870 1871 1872 1873 1874 1875 1876 1877
 1878 1879 1880 1881 1882 1883 1884 1885 1886 1887 1888 1889 1890 1891
 1892 1893 1894 1895 1896 1897 1898 1899 1900 1901 1902 1903 1904 1905
 1906 1907 1908 1909 1910 1911 1912 1913 1914 1915 1916 1917 1918 1919
 1920 1921 1922 1923 1924 1925 1926 1927 1928 1929 1930 1931 1932 1933
 1934 1935 1936 1937 1938 1939 1940 1941 1942 1943 1944 1945 1946 1947
 1948 1949 1950 1951 1952 1953 1954 1955 1956 1957 1958 1959 1960 1961
 1962 1963 1964 1965 1966 1967 1968 1969 1970 1971 1972 1973 1974 1975
 1976 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989
 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003
 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017
 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031
 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045
 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059
 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073
 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087
 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101
 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115
 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129
 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143
 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157
 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171
 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185
 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199
 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213
 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227
 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241
 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255
 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269
 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283
 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297
 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311
 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325
 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339
 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353
 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367
 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381
 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395
 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409
 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423
 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437
 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451
 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465
 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479
 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493
 2494 2495 2496 2497 2498 2499 2500]
TRAIN: [   0    1    2 ... 2498 2499 2500] TEST: [2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511
2512 2513 2514
 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528
 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542
 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556
 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570
 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584
```

```
 2585 2586 2587 2588 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598
 2599 2600 2601 2602 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612
 2613 2614 2615 2616 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626
 2627 2628 2629 2630 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640
 2641 2642 2643 2644 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654
 2655 2656 2657 2658 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668
 2669 2670 2671 2672 2673 2674 2675 2676 2677 2678 2679 2680 2681 2682
 2683 2684 2685 2686 2687 2688 2689 2690 2691 2692 2693 2694 2695 2696
 2697 2698 2699 2700 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710
 2711 2712 2713 2714 2715 2716 2717 2718 2719 2720 2721 2722 2723 2724
 2725 2726 2727 2728 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738
 2739 2740 2741 2742 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752
 2753 2754 2755 2756 2757 2758 2759 2760 2761 2762 2763 2764 2765 2766
 2767 2768 2769 2770 2771 2772 2773 2774 2775 2776 2777 2778 2779 2780
 2781 2782 2783 2784 2785 2786 2787 2788 2789 2790 2791 2792 2793 2794
 2795 2796 2797 2798 2799 2800 2801 2802 2803 2804 2805 2806 2807 2808
 2809 2810 2811 2812 2813 2814 2815 2816 2817 2818 2819 2820 2821 2822
 2823 2824 2825 2826 2827 2828 2829 2830 2831 2832 2833 2834 2835 2836
 2837 2838 2839 2840 2841 2842 2843 2844 2845 2846 2847 2848 2849 2850
 2851 2852 2853 2854 2855 2856 2857 2858 2859 2860 2861 2862 2863 2864
 2865 2866 2867 2868 2869 2870 2871 2872 2873 2874 2875 2876 2877 2878
 2879 2880 2881 2882 2883 2884 2885 2886 2887 2888 2889 2890 2891 2892
 2893 2894 2895 2896 2897 2898 2899 2900 2901 2902 2903 2904 2905 2906
 2907 2908 2909 2910 2911 2912 2913 2914 2915 2916 2917 2918 2919 2920
 2921 2922 2923 2924 2925 2926 2927 2928 2929 2930 2931 2932 2933 2934
 2935 2936 2937 2938 2939 2940 2941 2942 2943 2944 2945 2946 2947 2948
 2949 2950 2951 2952 2953 2954 2955 2956 2957 2958 2959 2960 2961 2962
 2963 2964 2965 2966 2967 2968 2969 2970 2971 2972 2973 2974 2975 2976
 2977 2978 2979 2980 2981 2982 2983 2984 2985 2986 2987 2988 2989 2990
 2991 2992 2993 2994 2995 2996 2997 2998 2999 3000 3001 3002 3003 3004
 3005 3006 3007 3008 3009 3010 3011 3012 3013 3014 3015 3016 3017 3018
 3019 3020 3021 3022 3023 3024 3025 3026 3027 3028 3029 3030 3031 3032
 3033 3034 3035 3036 3037 3038 3039 3040 3041 3042 3043 3044 3045 3046
 3047 3048 3049 3050 3051 3052 3053 3054 3055 3056 3057 3058 3059 3060
 3061 3062 3063 3064 3065 3066 3067 3068 3069 3070 3071 3072 3073 3074
 3075 3076 3077 3078 3079 3080 3081 3082 3083 3084 3085 3086 3087 3088
 3089 3090 3091 3092 3093 3094 3095 3096 3097 3098 3099 3100 3101 3102
 3103 3104 3105 3106 3107 3108 3109 3110 3111 3112 3113 3114 3115 3116
 3117 3118 3119 3120 3121 3122 3123 3124 3125 3126 3127 3128 3129 3130
 3131 3132 3133 3134 3135 3136 3137 3138 3139 3140 3141 3142 3143 3144
 3145 3146 3147 3148 3149 3150 3151 3152 3153 3154 3155 3156 3157 3158
 3159 3160 3161 3162 3163 3164 3165 3166 3167 3168 3169 3170 3171 3172
 3173 3174 3175 3176 3177 3178 3179 3180 3181 3182 3183 3184 3185 3186
 3187 3188 3189 3190 3191 3192 3193 3194 3195 3196 3197 3198 3199 3200
 3201 3202 3203 3204 3205 3206 3207 3208 3209 3210 3211 3212 3213 3214
 3215 3216 3217 3218 3219 3220 3221 3222 3223 3224 3225 3226 3227 3228
 3229 3230 3231 3232 3233 3234 3235 3236 3237 3238 3239 3240 3241 3242
 3243 3244 3245 3246 3247 3248 3249 3250 3251 3252 3253 3254 3255 3256
 3257 3258 3259 3260 3261 3262 3263 3264 3265 3266 3267 3268 3269 3270
 3271 3272 3273 3274 3275 3276 3277 3278 3279 3280 3281 3282 3283 3284
 3285 3286 3287 3288 3289 3290 3291 3292 3293 3294 3295 3296 3297 3298
 3299 3300 3301 3302 3303 3304 3305 3306 3307 3308 3309 3310 3311 3312
 3313 3314 3315 3316 3317 3318 3319 3320 3321 3322 3323 3324 3325 3326
 3327 3328 3329 3330 3331 3332 3333]
TRAIN: [   0    1    2 ... 3331 3332 3333] TEST: [3334 3335 3336 3337 3338 3339 3340 3341 3342 3343 3344
3345 3346 3347
 3348 3349 3350 3351 3352 3353 3354 3355 3356 3357 3358 3359 3360 3361
 3362 3363 3364 3365 3366 3367 3368 3369 3370 3371 3372 3373 3374 3375
 3376 3377 3378 3379 3380 3381 3382 3383 3384 3385 3386 3387 3388 3389
 3390 3391 3392 3393 3394 3395 3396 3397 3398 3399 3400 3401 3402 3403
 3404 3405 3406 3407 3408 3409 3410 3411 3412 3413 3414 3415 3416 3417
 3418 3419 3420 3421 3422 3423 3424 3425 3426 3427 3428 3429 3430 3431
 3432 3433 3434 3435 3436 3437 3438 3439 3440 3441 3442 3443 3444 3445
 3446 3447 3448 3449 3450 3451 3452 3453 3454 3455 3456 3457 3458 3459
 3460 3461 3462 3463 3464 3465 3466 3467 3468 3469 3470 3471 3472 3473
 3474 3475 3476 3477 3478 3479 3480 3481 3482 3483 3484 3485 3486 3487
 3488 3489 3490 3491 3492 3493 3494 3495 3496 3497 3498 3499 3500 3501
 3502 3503 3504 3505 3506 3507 3508 3509 3510 3511 3512 3513 3514 3515
 3516 3517 3518 3519 3520 3521 3522 3523 3524 3525 3526 3527 3528 3529
 3530 3531 3532 3533 3534 3535 3536 3537 3538 3539 3540 3541 3542 3543
 3544 3545 3546 3547 3548 3549 3550 3551 3552 3553 3554 3555 3556 3557
 3558 3559 3560 3561 3562 3563 3564 3565 3566 3567 3568 3569 3570 3571
 3572 3573 3574 3575 3576 3577 3578 3579 3580 3581 3582 3583 3584 3585
 3586 3587 3588 3589 3590 3591 3592 3593 3594 3595 3596 3597 3598 3599
 3600 3601 3602 3603 3604 3605 3606 3607 3608 3609 3610 3611 3612 3613
 3614 3615 3616 3617 3618 3619 3620 3621 3622 3623 3624 3625 3626 3627
 3628 3629 3630 3631 3632 3633 3634 3635 3636 3637 3638 3639 3640 3641
 3642 3643 3644 3645 3646 3647 3648 3649 3650 3651 3652 3653 3654 3655
 3656 3657 3658 3659 3660 3661 3662 3663 3664 3665 3666 3667 3668 3669
 3670 3671 3672 3673 3674 3675 3676 3677 3678 3679 3680 3681 3682 3683
 3684 3685 3686 3687 3688 3689 3690 3691 3692 3693 3694 3695 3696 3697
 3698 3699 3700 3701 3702 3703 3704 3705 3706 3707 3708 3709 3710 3711
 3712 3713 3714 3715 3716 3717 3718 3719 3720 3721 3722 3723 3724 3725
 3726 3727 3728 3729 3730 3731 3732 3733 3734 3735 3736 3737 3738 3739
 3740 3741 3742 3743 3744 3745 3746 3747 3748 3749 3750 3751 3752 3753
 3754 3755 3756 3757 3758 3759 3760 3761 3762 3763 3764 3765 3766 3767
 3768 3769 3770 3771 3772 3773 3774 3775 3776 3777 3778 3779 3780 3781
 3782 3783 3784 3785 3786 3787 3788 3789 3790 3791 3792 3793 3794 3795
```

```
 3796 3797 3798 3799 3800 3801 3802 3803 3804 3805 3806 3807 3808 3809
 3810 3811 3812 3813 3814 3815 3816 3817 3818 3819 3820 3821 3822 3823
 3824 3825 3826 3827 3828 3829 3830 3831 3832 3833 3834 3835 3836 3837
 3838 3839 3840 3841 3842 3843 3844 3845 3846 3847 3848 3849 3850 3851
 3852 3853 3854 3855 3856 3857 3858 3859 3860 3861 3862 3863 3864 3865
 3866 3867 3868 3869 3870 3871 3872 3873 3874 3875 3876 3877 3878 3879
 3880 3881 3882 3883 3884 3885 3886 3887 3888 3889 3890 3891 3892 3893
 3894 3895 3896 3897 3898 3899 3900 3901 3902 3903 3904 3905 3906 3907
 3908 3909 3910 3911 3912 3913 3914 3915 3916 3917 3918 3919 3920 3921
 3922 3923 3924 3925 3926 3927 3928 3929 3930 3931 3932 3933 3934 3935
 3936 3937 3938 3939 3940 3941 3942 3943 3944 3945 3946 3947 3948 3949
 3950 3951 3952 3953 3954 3955 3956 3957 3958 3959 3960 3961 3962 3963
 3964 3965 3966 3967 3968 3969 3970 3971 3972 3973 3974 3975 3976 3977
 3978 3979 3980 3981 3982 3983 3984 3985 3986 3987 3988 3989 3990 3991
 3992 3993 3994 3995 3996 3997 3998 3999 4000 4001 4002 4003 4004 4005
 4006 4007 4008 4009 4010 4011 4012 4013 4014 4015 4016 4017 4018 4019
 4020 4021 4022 4023 4024 4025 4026 4027 4028 4029 4030 4031 4032 4033
 4034 4035 4036 4037 4038 4039 4040 4041 4042 4043 4044 4045 4046 4047
 4048 4049 4050 4051 4052 4053 4054 4055 4056 4057 4058 4059 4060 4061
 4062 4063 4064 4065 4066 4067 4068 4069 4070 4071 4072 4073 4074 4075
 4076 4077 4078 4079 4080 4081 4082 4083 4084 4085 4086 4087 4088 4089
 4090 4091 4092 4093 4094 4095 4096 4097 4098 4099 4100 4101 4102 4103
 4104 4105 4106 4107 4108 4109 4110 4111 4112 4113 4114 4115 4116 4117
 4118 4119 4120 4121 4122 4123 4124 4125 4126 4127 4128 4129 4130 4131
 4132 4133 4134 4135 4136 4137 4138 4139 4140 4141 4142 4143 4144 4145
 4146 4147 4148 4149 4150 4151 4152 4153 4154 4155 4156 4157 4158 4159
 4160 4161 4162 4163 4164 4165 4166]
TRAIN: [   0    1    2 ... 4164 4165 4166] TEST: [4167 4168 4169 4170 4171 4172 4173 4174 4175 4176 4177
4178 4179 4180
 4181 4182 4183 4184 4185 4186 4187 4188 4189 4190 4191 4192 4193 4194
 4195 4196 4197 4198 4199 4200 4201 4202 4203 4204 4205 4206 4207 4208
 4209 4210 4211 4212 4213 4214 4215 4216 4217 4218 4219 4220 4221 4222
 4223 4224 4225 4226 4227 4228 4229 4230 4231 4232 4233 4234 4235 4236
 4237 4238 4239 4240 4241 4242 4243 4244 4245 4246 4247 4248 4249 4250
 4251 4252 4253 4254 4255 4256 4257 4258 4259 4260 4261 4262 4263 4264
 4265 4266 4267 4268 4269 4270 4271 4272 4273 4274 4275 4276 4277 4278
 4279 4280 4281 4282 4283 4284 4285 4286 4287 4288 4289 4290 4291 4292
 4293 4294 4295 4296 4297 4298 4299 4300 4301 4302 4303 4304 4305 4306
 4307 4308 4309 4310 4311 4312 4313 4314 4315 4316 4317 4318 4319 4320
 4321 4322 4323 4324 4325 4326 4327 4328 4329 4330 4331 4332 4333 4334
 4335 4336 4337 4338 4339 4340 4341 4342 4343 4344 4345 4346 4347 4348
 4349 4350 4351 4352 4353 4354 4355 4356 4357 4358 4359 4360 4361 4362
 4363 4364 4365 4366 4367 4368 4369 4370 4371 4372 4373 4374 4375 4376
 4377 4378 4379 4380 4381 4382 4383 4384 4385 4386 4387 4388 4389 4390
 4391 4392 4393 4394 4395 4396 4397 4398 4399 4400 4401 4402 4403 4404
 4405 4406 4407 4408 4409 4410 4411 4412 4413 4414 4415 4416 4417 4418
 4419 4420 4421 4422 4423 4424 4425 4426 4427 4428 4429 4430 4431 4432
 4433 4434 4435 4436 4437 4438 4439 4440 4441 4442 4443 4444 4445 4446
 4447 4448 4449 4450 4451 4452 4453 4454 4455 4456 4457 4458 4459 4460
 4461 4462 4463 4464 4465 4466 4467 4468 4469 4470 4471 4472 4473 4474
 4475 4476 4477 4478 4479 4480 4481 4482 4483 4484 4485 4486 4487 4488
 4489 4490 4491 4492 4493 4494 4495 4496 4497 4498 4499 4500 4501 4502
 4503 4504 4505 4506 4507 4508 4509 4510 4511 4512 4513 4514 4515 4516
 4517 4518 4519 4520 4521 4522 4523 4524 4525 4526 4527 4528 4529 4530
 4531 4532 4533 4534 4535 4536 4537 4538 4539 4540 4541 4542 4543 4544
 4545 4546 4547 4548 4549 4550 4551 4552 4553 4554 4555 4556 4557 4558
 4559 4560 4561 4562 4563 4564 4565 4566 4567 4568 4569 4570 4571 4572
 4573 4574 4575 4576 4577 4578 4579 4580 4581 4582 4583 4584 4585 4586
 4587 4588 4589 4590 4591 4592 4593 4594 4595 4596 4597 4598 4599 4600
 4601 4602 4603 4604 4605 4606 4607 4608 4609 4610 4611 4612 4613 4614
 4615 4616 4617 4618 4619 4620 4621 4622 4623 4624 4625 4626 4627 4628
 4629 4630 4631 4632 4633 4634 4635 4636 4637 4638 4639 4640 4641 4642
 4643 4644 4645 4646 4647 4648 4649 4650 4651 4652 4653 4654 4655 4656
 4657 4658 4659 4660 4661 4662 4663 4664 4665 4666 4667 4668 4669 4670
 4671 4672 4673 4674 4675 4676 4677 4678 4679 4680 4681 4682 4683 4684
 4685 4686 4687 4688 4689 4690 4691 4692 4693 4694 4695 4696 4697 4698
 4699 4700 4701 4702 4703 4704 4705 4706 4707 4708 4709 4710 4711 4712
 4713 4714 4715 4716 4717 4718 4719 4720 4721 4722 4723 4724 4725 4726
 4727 4728 4729 4730 4731 4732 4733 4734 4735 4736 4737 4738 4739 4740
 4741 4742 4743 4744 4745 4746 4747 4748 4749 4750 4751 4752 4753 4754
 4755 4756 4757 4758 4759 4760 4761 4762 4763 4764 4765 4766 4767 4768
 4769 4770 4771 4772 4773 4774 4775 4776 4777 4778 4779 4780 4781 4782
 4783 4784 4785 4786 4787 4788 4789 4790 4791 4792 4793 4794 4795 4796
 4797 4798 4799 4800 4801 4802 4803 4804 4805 4806 4807 4808 4809 4810
 4811 4812 4813 4814 4815 4816 4817 4818 4819 4820 4821 4822 4823 4824
 4825 4826 4827 4828 4829 4830 4831 4832 4833 4834 4835 4836 4837 4838
 4839 4840 4841 4842 4843 4844 4845 4846 4847 4848 4849 4850 4851 4852
 4853 4854 4855 4856 4857 4858 4859 4860 4861 4862 4863 4864 4865 4866
 4867 4868 4869 4870 4871 4872 4873 4874 4875 4876 4877 4878 4879 4880
 4881 4882 4883 4884 4885 4886 4887 4888 4889 4890 4891 4892 4893 4894
 4895 4896 4897 4898 4899 4900 4901 4902 4903 4904 4905 4906 4907 4908
 4909 4910 4911 4912 4913 4914 4915 4916 4917 4918 4919 4920 4921 4922
 4923 4924 4925 4926 4927 4928 4929 4930 4931 4932 4933 4934 4935 4936
 4937 4938 4939 4940 4941 4942 4943 4944 4945 4946 4947 4948 4949 4950
 4951 4952 4953 4954 4955 4956 4957 4958 4959 4960 4961 4962 4963 4964
 4965 4966 4967 4968 4969 4970 4971 4972 4973 4974 4975 4976 4977 4978
 4979 4980 4981 4982 4983 4984 4985 4986 4987 4988 4989 4990 4991 4992
 4993 4994 4995 4996 4997 4998 4999]
```

In [149]:
```
print(X_TW_tr.shape)
print(X_TW_test.shape)
print(Y_TW_tr.shape)
print(Y_TW_test.shape)
```

```
(4167, 4)
(833, 4)
(4167,)
(833,)
```

In [150]:
```
# Splitting the Train dataset into Cross Validation & Train Datasets
for train_index, test_index in tscv.split(X_TW_tr):
   print("TRAIN:", train_index, "TEST:", test_index)
   X_TW_train, X_TW_cv = X[train_index], X[test_index]
   Y_TW_train, Y_TW_cv = Y[train_index], Y[test_index]
```

```
TRAIN: [  0   1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17
  18  19  20  21  22  23  24  25  26  27  28  29  30  31  32  33  34  35
  36  37  38  39  40  41  42  43  44  45  46  47  48  49  50  51  52  53
  54  55  56  57  58  59  60  61  62  63  64  65  66  67  68  69  70  71
  72  73  74  75  76  77  78  79  80  81  82  83  84  85  86  87  88  89
  90  91  92  93  94  95  96  97  98  99 100 101 102 103 104 105 106 107
 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125
 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143
 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161
 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179
 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197
 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215
 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233
 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251
 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269
 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287
 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305
 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323
 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341
 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359
 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377
 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395
 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413
 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431
 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449
 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467
 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485
 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503
 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521
 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539
 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557
 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575
 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593
 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611
 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629
 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647
 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665
 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683
 684 685 686 687 688 689 690 691 692 693 694 695 696] TEST: [ 697  698  699  700  701  702  703  704  705
  706  707  708  709  710
  711  712  713  714  715  716  717  718  719  720  721  722  723  724
  725  726  727  728  729  730  731  732  733  734  735  736  737  738
  739  740  741  742  743  744  745  746  747  748  749  750  751  752
  753  754  755  756  757  758  759  760  761  762  763  764  765  766
  767  768  769  770  771  772  773  774  775  776  777  778  779  780
  781  782  783  784  785  786  787  788  789  790  791  792  793  794
  795  796  797  798  799  800  801  802  803  804  805  806  807  808
  809  810  811  812  813  814  815  816  817  818  819  820  821  822
  823  824  825  826  827  828  829  830  831  832  833  834  835  836
  837  838  839  840  841  842  843  844  845  846  847  848  849  850
  851  852  853  854  855  856  857  858  859  860  861  862  863  864
  865  866  867  868  869  870  871  872  873  874  875  876  877  878
  879  880  881  882  883  884  885  886  887  888  889  890  891  892
  893  894  895  896  897  898  899  900  901  902  903  904  905  906
  907  908  909  910  911  912  913  914  915  916  917  918  919  920
  921  922  923  924  925  926  927  928  929  930  931  932  933  934
  935  936  937  938  939  940  941  942  943  944  945  946  947  948
  949  950  951  952  953  954  955  956  957  958  959  960  961  962
  963  964  965  966  967  968  969  970  971  972  973  974  975  976
  977  978  979  980  981  982  983  984  985  986  987  988  989  990
  991  992  993  994  995  996  997  998  999 1000 1001 1002 1003 1004
 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018
 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032
 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046
 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060
 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074
 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088
 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102
 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116
 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128 1129 1130
 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142 1143 1144
 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158
 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172
 1173 1174 1175 1176 1177 1178 1179 1180 1181 1182 1183 1184 1185 1186
 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198 1199 1200
 1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212 1213 1214
 1215 1216 1217 1218 1219 1220 1221 1222 1223 1224 1225 1226 1227 1228
 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240 1241 1242
 1243 1244 1245 1246 1247 1248 1249 1250 1251 1252 1253 1254 1255 1256
 1257 1258 1259 1260 1261 1262 1263 1264 1265 1266 1267 1268 1269 1270
 1271 1272 1273 1274 1275 1276 1277 1278 1279 1280 1281 1282 1283 1284
 1285 1286 1287 1288 1289 1290 1291 1292 1293 1294 1295 1296 1297 1298
 1299 1300 1301 1302 1303 1304 1305 1306 1307 1308 1309 1310 1311 1312
 1313 1314 1315 1316 1317 1318 1319 1320 1321 1322 1323 1324 1325 1326
 1327 1328 1329 1330 1331 1332 1333 1334 1335 1336 1337 1338 1339 1340
 1341 1342 1343 1344 1345 1346 1347 1348 1349 1350 1351 1352 1353 1354
 1355 1356 1357 1358 1359 1360 1361 1362 1363 1364 1365 1366 1367 1368
```

```
 1369 1370 1371 1372 1373 1374 1375 1376 1377 1378 1379 1380 1381 1382
 1383 1384 1385 1386 1387 1388 1389 1390]
TRAIN: [   0    1    2 ... 1388 1389 1390] TEST: [1391 1392 1393 1394 1395 1396 1397 1398 1399 1400 1401
1402 1403 1404
 1405 1406 1407 1408 1409 1410 1411 1412 1413 1414 1415 1416 1417 1418
 1419 1420 1421 1422 1423 1424 1425 1426 1427 1428 1429 1430 1431 1432
 1433 1434 1435 1436 1437 1438 1439 1440 1441 1442 1443 1444 1445 1446
 1447 1448 1449 1450 1451 1452 1453 1454 1455 1456 1457 1458 1459 1460
 1461 1462 1463 1464 1465 1466 1467 1468 1469 1470 1471 1472 1473 1474
 1475 1476 1477 1478 1479 1480 1481 1482 1483 1484 1485 1486 1487 1488
 1489 1490 1491 1492 1493 1494 1495 1496 1497 1498 1499 1500 1501 1502
 1503 1504 1505 1506 1507 1508 1509 1510 1511 1512 1513 1514 1515 1516
 1517 1518 1519 1520 1521 1522 1523 1524 1525 1526 1527 1528 1529 1530
 1531 1532 1533 1534 1535 1536 1537 1538 1539 1540 1541 1542 1543 1544
 1545 1546 1547 1548 1549 1550 1551 1552 1553 1554 1555 1556 1557 1558
 1559 1560 1561 1562 1563 1564 1565 1566 1567 1568 1569 1570 1571 1572
 1573 1574 1575 1576 1577 1578 1579 1580 1581 1582 1583 1584 1585 1586
 1587 1588 1589 1590 1591 1592 1593 1594 1595 1596 1597 1598 1599 1600
 1601 1602 1603 1604 1605 1606 1607 1608 1609 1610 1611 1612 1613 1614
 1615 1616 1617 1618 1619 1620 1621 1622 1623 1624 1625 1626 1627 1628
 1629 1630 1631 1632 1633 1634 1635 1636 1637 1638 1639 1640 1641 1642
 1643 1644 1645 1646 1647 1648 1649 1650 1651 1652 1653 1654 1655 1656
 1657 1658 1659 1660 1661 1662 1663 1664 1665 1666 1667 1668 1669 1670
 1671 1672 1673 1674 1675 1676 1677 1678 1679 1680 1681 1682 1683 1684
 1685 1686 1687 1688 1689 1690 1691 1692 1693 1694 1695 1696 1697 1698
 1699 1700 1701 1702 1703 1704 1705 1706 1707 1708 1709 1710 1711 1712
 1713 1714 1715 1716 1717 1718 1719 1720 1721 1722 1723 1724 1725 1726
 1727 1728 1729 1730 1731 1732 1733 1734 1735 1736 1737 1738 1739 1740
 1741 1742 1743 1744 1745 1746 1747 1748 1749 1750 1751 1752 1753 1754
 1755 1756 1757 1758 1759 1760 1761 1762 1763 1764 1765 1766 1767 1768
 1769 1770 1771 1772 1773 1774 1775 1776 1777 1778 1779 1780 1781 1782
 1783 1784 1785 1786 1787 1788 1789 1790 1791 1792 1793 1794 1795 1796
 1797 1798 1799 1800 1801 1802 1803 1804 1805 1806 1807 1808 1809 1810
 1811 1812 1813 1814 1815 1816 1817 1818 1819 1820 1821 1822 1823 1824
 1825 1826 1827 1828 1829 1830 1831 1832 1833 1834 1835 1836 1837 1838
 1839 1840 1841 1842 1843 1844 1845 1846 1847 1848 1849 1850 1851 1852
 1853 1854 1855 1856 1857 1858 1859 1860 1861 1862 1863 1864 1865 1866
 1867 1868 1869 1870 1871 1872 1873 1874 1875 1876 1877 1878 1879 1880
 1881 1882 1883 1884 1885 1886 1887 1888 1889 1890 1891 1892 1893 1894
 1895 1896 1897 1898 1899 1900 1901 1902 1903 1904 1905 1906 1907 1908
 1909 1910 1911 1912 1913 1914 1915 1916 1917 1918 1919 1920 1921 1922
 1923 1924 1925 1926 1927 1928 1929 1930 1931 1932 1933 1934 1935 1936
 1937 1938 1939 1940 1941 1942 1943 1944 1945 1946 1947 1948 1949 1950
 1951 1952 1953 1954 1955 1956 1957 1958 1959 1960 1961 1962 1963 1964
 1965 1966 1967 1968 1969 1970 1971 1972 1973 1974 1975 1976 1977 1978
 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991 1992
 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006
 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020
 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034
 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048
 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062
 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076
 2077 2078 2079 2080 2081 2082 2083 2084]
TRAIN: [   0    1    2 ... 2082 2083 2084] TEST: [2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095
2096 2097 2098
 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112
 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126
 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140
 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154
 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168
 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182
 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196
 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210
 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224
 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238
 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252
 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266
 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280
 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294
 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308
 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322
 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336
 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350
 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364
 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378
 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392
 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406
 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420
 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434
 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448
 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462
 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476
 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490
 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504
 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518
 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532
 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546
 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560
```

```
 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574
 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584 2585 2586 2587 2588
 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598 2599 2600 2601 2602
 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612 2613 2614 2615 2616
 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626 2627 2628 2629 2630
 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640 2641 2642 2643 2644
 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654 2655 2656 2657 2658
 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668 2669 2670 2671 2672
 2673 2674 2675 2676 2677 2678 2679 2680 2681 2682 2683 2684 2685 2686
 2687 2688 2689 2690 2691 2692 2693 2694 2695 2696 2697 2698 2699 2700
 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710 2711 2712 2713 2714
 2715 2716 2717 2718 2719 2720 2721 2722 2723 2724 2725 2726 2727 2728
 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738 2739 2740 2741 2742
 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752 2753 2754 2755 2756
 2757 2758 2759 2760 2761 2762 2763 2764 2765 2766 2767 2768 2769 2770
 2771 2772 2773 2774 2775 2776 2777 2778]
TRAIN: [   0    1    2 ... 2776 2777 2778] TEST: [2779 2780 2781 2782 2783 2784 2785 2786 2787 2788 2789
2790 2791 2792
 2793 2794 2795 2796 2797 2798 2799 2800 2801 2802 2803 2804 2805 2806
 2807 2808 2809 2810 2811 2812 2813 2814 2815 2816 2817 2818 2819 2820
 2821 2822 2823 2824 2825 2826 2827 2828 2829 2830 2831 2832 2833 2834
 2835 2836 2837 2838 2839 2840 2841 2842 2843 2844 2845 2846 2847 2848
 2849 2850 2851 2852 2853 2854 2855 2856 2857 2858 2859 2860 2861 2862
 2863 2864 2865 2866 2867 2868 2869 2870 2871 2872 2873 2874 2875 2876
 2877 2878 2879 2880 2881 2882 2883 2884 2885 2886 2887 2888 2889 2890
 2891 2892 2893 2894 2895 2896 2897 2898 2899 2900 2901 2902 2903 2904
 2905 2906 2907 2908 2909 2910 2911 2912 2913 2914 2915 2916 2917 2918
 2919 2920 2921 2922 2923 2924 2925 2926 2927 2928 2929 2930 2931 2932
 2933 2934 2935 2936 2937 2938 2939 2940 2941 2942 2943 2944 2945 2946
 2947 2948 2949 2950 2951 2952 2953 2954 2955 2956 2957 2958 2959 2960
 2961 2962 2963 2964 2965 2966 2967 2968 2969 2970 2971 2972 2973 2974
 2975 2976 2977 2978 2979 2980 2981 2982 2983 2984 2985 2986 2987 2988
 2989 2990 2991 2992 2993 2994 2995 2996 2997 2998 2999 3000 3001 3002
 3003 3004 3005 3006 3007 3008 3009 3010 3011 3012 3013 3014 3015 3016
 3017 3018 3019 3020 3021 3022 3023 3024 3025 3026 3027 3028 3029 3030
 3031 3032 3033 3034 3035 3036 3037 3038 3039 3040 3041 3042 3043 3044
 3045 3046 3047 3048 3049 3050 3051 3052 3053 3054 3055 3056 3057 3058
 3059 3060 3061 3062 3063 3064 3065 3066 3067 3068 3069 3070 3071 3072
 3073 3074 3075 3076 3077 3078 3079 3080 3081 3082 3083 3084 3085 3086
 3087 3088 3089 3090 3091 3092 3093 3094 3095 3096 3097 3098 3099 3100
 3101 3102 3103 3104 3105 3106 3107 3108 3109 3110 3111 3112 3113 3114
 3115 3116 3117 3118 3119 3120 3121 3122 3123 3124 3125 3126 3127 3128
 3129 3130 3131 3132 3133 3134 3135 3136 3137 3138 3139 3140 3141 3142
 3143 3144 3145 3146 3147 3148 3149 3150 3151 3152 3153 3154 3155 3156
 3157 3158 3159 3160 3161 3162 3163 3164 3165 3166 3167 3168 3169 3170
 3171 3172 3173 3174 3175 3176 3177 3178 3179 3180 3181 3182 3183 3184
 3185 3186 3187 3188 3189 3190 3191 3192 3193 3194 3195 3196 3197 3198
 3199 3200 3201 3202 3203 3204 3205 3206 3207 3208 3209 3210 3211 3212
 3213 3214 3215 3216 3217 3218 3219 3220 3221 3222 3223 3224 3225 3226
 3227 3228 3229 3230 3231 3232 3233 3234 3235 3236 3237 3238 3239 3240
 3241 3242 3243 3244 3245 3246 3247 3248 3249 3250 3251 3252 3253 3254
 3255 3256 3257 3258 3259 3260 3261 3262 3263 3264 3265 3266 3267 3268
 3269 3270 3271 3272 3273 3274 3275 3276 3277 3278 3279 3280 3281 3282
 3283 3284 3285 3286 3287 3288 3289 3290 3291 3292 3293 3294 3295 3296
 3297 3298 3299 3300 3301 3302 3303 3304 3305 3306 3307 3308 3309 3310
 3311 3312 3313 3314 3315 3316 3317 3318 3319 3320 3321 3322 3323 3324
 3325 3326 3327 3328 3329 3330 3331 3332 3333 3334 3335 3336 3337 3338
 3339 3340 3341 3342 3343 3344 3345 3346 3347 3348 3349 3350 3351 3352
 3353 3354 3355 3356 3357 3358 3359 3360 3361 3362 3363 3364 3365 3366
 3367 3368 3369 3370 3371 3372 3373 3374 3375 3376 3377 3378 3379 3380
 3381 3382 3383 3384 3385 3386 3387 3388 3389 3390 3391 3392 3393 3394
 3395 3396 3397 3398 3399 3400 3401 3402 3403 3404 3405 3406 3407 3408
 3409 3410 3411 3412 3413 3414 3415 3416 3417 3418 3419 3420 3421 3422
 3423 3424 3425 3426 3427 3428 3429 3430 3431 3432 3433 3434 3435 3436
 3437 3438 3439 3440 3441 3442 3443 3444 3445 3446 3447 3448 3449 3450
 3451 3452 3453 3454 3455 3456 3457 3458 3459 3460 3461 3462 3463 3464
 3465 3466 3467 3468 3469 3470 3471 3472]
TRAIN: [   0    1    2 ... 3470 3471 3472] TEST: [3473 3474 3475 3476 3477 3478 3479 3480 3481 3482 3483
3484 3485 3486
 3487 3488 3489 3490 3491 3492 3493 3494 3495 3496 3497 3498 3499 3500
 3501 3502 3503 3504 3505 3506 3507 3508 3509 3510 3511 3512 3513 3514
 3515 3516 3517 3518 3519 3520 3521 3522 3523 3524 3525 3526 3527 3528
 3529 3530 3531 3532 3533 3534 3535 3536 3537 3538 3539 3540 3541 3542
 3543 3544 3545 3546 3547 3548 3549 3550 3551 3552 3553 3554 3555 3556
 3557 3558 3559 3560 3561 3562 3563 3564 3565 3566 3567 3568 3569 3570
 3571 3572 3573 3574 3575 3576 3577 3578 3579 3580 3581 3582 3583 3584
 3585 3586 3587 3588 3589 3590 3591 3592 3593 3594 3595 3596 3597 3598
 3599 3600 3601 3602 3603 3604 3605 3606 3607 3608 3609 3610 3611 3612
 3613 3614 3615 3616 3617 3618 3619 3620 3621 3622 3623 3624 3625 3626
 3627 3628 3629 3630 3631 3632 3633 3634 3635 3636 3637 3638 3639 3640
 3641 3642 3643 3644 3645 3646 3647 3648 3649 3650 3651 3652 3653 3654
 3655 3656 3657 3658 3659 3660 3661 3662 3663 3664 3665 3666 3667 3668
 3669 3670 3671 3672 3673 3674 3675 3676 3677 3678 3679 3680 3681 3682
 3683 3684 3685 3686 3687 3688 3689 3690 3691 3692 3693 3694 3695 3696
 3697 3698 3699 3700 3701 3702 3703 3704 3705 3706 3707 3708 3709 3710
 3711 3712 3713 3714 3715 3716 3717 3718 3719 3720 3721 3722 3723 3724
 3725 3726 3727 3728 3729 3730 3731 3732 3733 3734 3735 3736 3737 3738
 3739 3740 3741 3742 3743 3744 3745 3746 3747 3748 3749 3750 3751 3752
```

```
 3753 3754 3755 3756 3757 3758 3759 3760 3761 3762 3763 3764 3765 3766
 3767 3768 3769 3770 3771 3772 3773 3774 3775 3776 3777 3778 3779 3780
 3781 3782 3783 3784 3785 3786 3787 3788 3789 3790 3791 3792 3793 3794
 3795 3796 3797 3798 3799 3800 3801 3802 3803 3804 3805 3806 3807 3808
 3809 3810 3811 3812 3813 3814 3815 3816 3817 3818 3819 3820 3821 3822
 3823 3824 3825 3826 3827 3828 3829 3830 3831 3832 3833 3834 3835 3836
 3837 3838 3839 3840 3841 3842 3843 3844 3845 3846 3847 3848 3849 3850
 3851 3852 3853 3854 3855 3856 3857 3858 3859 3860 3861 3862 3863 3864
 3865 3866 3867 3868 3869 3870 3871 3872 3873 3874 3875 3876 3877 3878
 3879 3880 3881 3882 3883 3884 3885 3886 3887 3888 3889 3890 3891 3892
 3893 3894 3895 3896 3897 3898 3899 3900 3901 3902 3903 3904 3905 3906
 3907 3908 3909 3910 3911 3912 3913 3914 3915 3916 3917 3918 3919 3920
 3921 3922 3923 3924 3925 3926 3927 3928 3929 3930 3931 3932 3933 3934
 3935 3936 3937 3938 3939 3940 3941 3942 3943 3944 3945 3946 3947 3948
 3949 3950 3951 3952 3953 3954 3955 3956 3957 3958 3959 3960 3961 3962
 3963 3964 3965 3966 3967 3968 3969 3970 3971 3972 3973 3974 3975 3976
 3977 3978 3979 3980 3981 3982 3983 3984 3985 3986 3987 3988 3989 3990
 3991 3992 3993 3994 3995 3996 3997 3998 3999 4000 4001 4002 4003 4004
 4005 4006 4007 4008 4009 4010 4011 4012 4013 4014 4015 4016 4017 4018
 4019 4020 4021 4022 4023 4024 4025 4026 4027 4028 4029 4030 4031 4032
 4033 4034 4035 4036 4037 4038 4039 4040 4041 4042 4043 4044 4045 4046
 4047 4048 4049 4050 4051 4052 4053 4054 4055 4056 4057 4058 4059 4060
 4061 4062 4063 4064 4065 4066 4067 4068 4069 4070 4071 4072 4073 4074
 4075 4076 4077 4078 4079 4080 4081 4082 4083 4084 4085 4086 4087 4088
 4089 4090 4091 4092 4093 4094 4095 4096 4097 4098 4099 4100 4101 4102
 4103 4104 4105 4106 4107 4108 4109 4110 4111 4112 4113 4114 4115 4116
 4117 4118 4119 4120 4121 4122 4123 4124 4125 4126 4127 4128 4129 4130
 4131 4132 4133 4134 4135 4136 4137 4138 4139 4140 4141 4142 4143 4144
 4145 4146 4147 4148 4149 4150 4151 4152 4153 4154 4155 4156 4157 4158
 4159 4160 4161 4162 4163 4164 4165 4166]
```

In [151]:
```
print(X_TW_train.shape)
print(X_TW_cv.shape)
print(X_TW_test.shape)

print(Y_TW_train.shape)
print(Y_TW_cv.shape)
print(Y_TW_test.shape)
```

```
(3473, 4)
(694, 4)
(833, 4)
(3473,)
(694,)
(833,)
```

In [152]:
```
# Removing the Date variable from the Data sets
TW_X= []
for each in X_TW_train:
    TW_X.append(np.delete(each,3))
X_TW_train = np.asarray(TW_X)


TW_X_cv= []
for each in X_TW_cv:
    TW_X_cv.append(np.delete(each,3))
X_TW_cv = np.asarray(TW_X_cv)

TW_X_test= []
for each in X_TW_test:
    TW_X_test.append(np.delete(each,3))
X_TW_test = np.asarray(TW_X_test)
```

In [153]:
```
print(X_TW_train.shape)
print(X_TW_cv.shape)
print(X_TW_test.shape)

print(Y_TW_train.shape)
print(Y_TW_cv.shape)
print(Y_TW_test.shape)
```

```
(3473, 3)
(694, 3)
(833, 3)
(3473,)
(694,)
(833,)
```

In [154]:
```
# Finding the Test accuracy using Cross Validation Data Set
for i in range(1,30,2):
    # instantiate learning model (k = 30)
    knn = KNeighborsClassifier(n_neighbors=i)

    # fitting the model on crossvalidation train
    knn.fit(X_TW_train, Y_TW_train)

    # predict the response on the crossvalidation train
    pred = knn.predict(X_TW_cv)

    # evaluate CV accuracy
    acc = accuracy_score(Y_TW_cv, pred, normalize=True) * float(100)
    print('\nCV accuracy for k = %d is %d%%' % (i, acc))

knn = KNeighborsClassifier(1)
knn.fit(X_TW_train,Y_TW_train)
pred = knn.predict(X_TW_test)
acc = accuracy_score(Y_TW_test, pred, normalize=True) * float(100)
print('\n****Test accuracy for k = 1 is %d%%' % (acc))
```

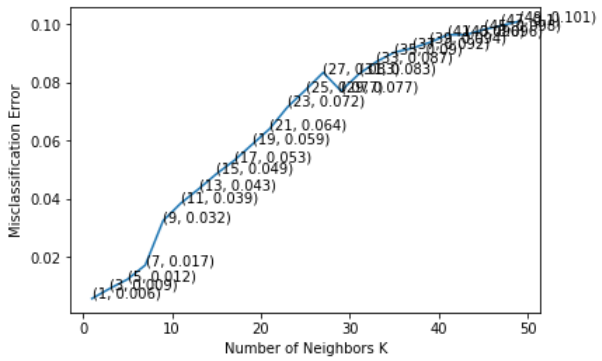
```
CV accuracy for k = 1 is 98%

CV accuracy for k = 3 is 98%

CV accuracy for k = 5 is 97%

CV accuracy for k = 7 is 96%

CV accuracy for k = 9 is 96%

CV accuracy for k = 11 is 96%

CV accuracy for k = 13 is 96%

CV accuracy for k = 15 is 95%

CV accuracy for k = 17 is 95%

CV accuracy for k = 19 is 94%

CV accuracy for k = 21 is 94%

CV accuracy for k = 23 is 93%

CV accuracy for k = 25 is 92%

CV accuracy for k = 27 is 89%

CV accuracy for k = 29 is 88%

****Test accuracy for k = 1 is 99%
```

In [155]:
```
# creating odd list of K for KNN
myList = list(range(0,50))
neighbors = list(filter(lambda x: x % 2 != 0, myList))

# empty list that will hold cv scores
cv_scores = []
```

In [156]:
```
for k in neighbors:
    knn = KNeighborsClassifier(n_neighbors=k)
    scores = cross_val_score(knn, X_TW_train, Y_TW_train, cv=10)
    cv_scores.append(scores.mean())
```

In [157]:
```
# changing to misclassification error
MSE = [1 - x for x in cv_scores]

# determining best k
optimal_k = neighbors[MSE.index(min(MSE))]
print('\nThe optimal number of neighbors is %d.' % optimal_k)

# plot misclassification error vs k
plt.plot(neighbors, MSE)

for xy in zip(neighbors, np.round(MSE,3)):
    plt.annotate('(%s, %s)' % xy, xy=xy, textcoords='data')

plt.xlabel('Number of Neighbors K')
plt.ylabel('Misclassification Error')
plt.show()

print("the misclassification error for each k value is : ", np.round(MSE,3))
```

```
The optimal number of neighbors is 1.
```




```
the misclassification error for each k value is :  [0.006 0.009 0.012 0.017 0.032 0.039 0.043 0.049 0.053
0.059 0.064 0.072
 0.077 0.083 0.077 0.083 0.087 0.09  0.092 0.094 0.096 0.096 0.098 0.1
 0.101]
```

In [158]:
```
# ============================== KNN with k = optimal_k ==============================================
# instantiate learning model k = optimal_k
knn_optimal = KNeighborsClassifier(n_neighbors=optimal_k)

# fitting the model
knn_optimal.fit(X_TW_train, Y_TW_train)

# predict the response
pred = knn_optimal.predict(X_TW_test)

# evaluate accuracy
acc = accuracy_score(Y_TW_test, pred) * 100
print('\nThe accuracy of the knn classifier for k = %d is %f%%' % (optimal_k, acc))
```

```
The accuracy of the knn classifier for k = 1 is 99.639856%
```

Conclusion: The Accuracy obtained by all different types of Doc2Vec conversion techniques (BOW, TF-IDF, AW2V, TF_IDF W2V) are same i.e. 99.6%