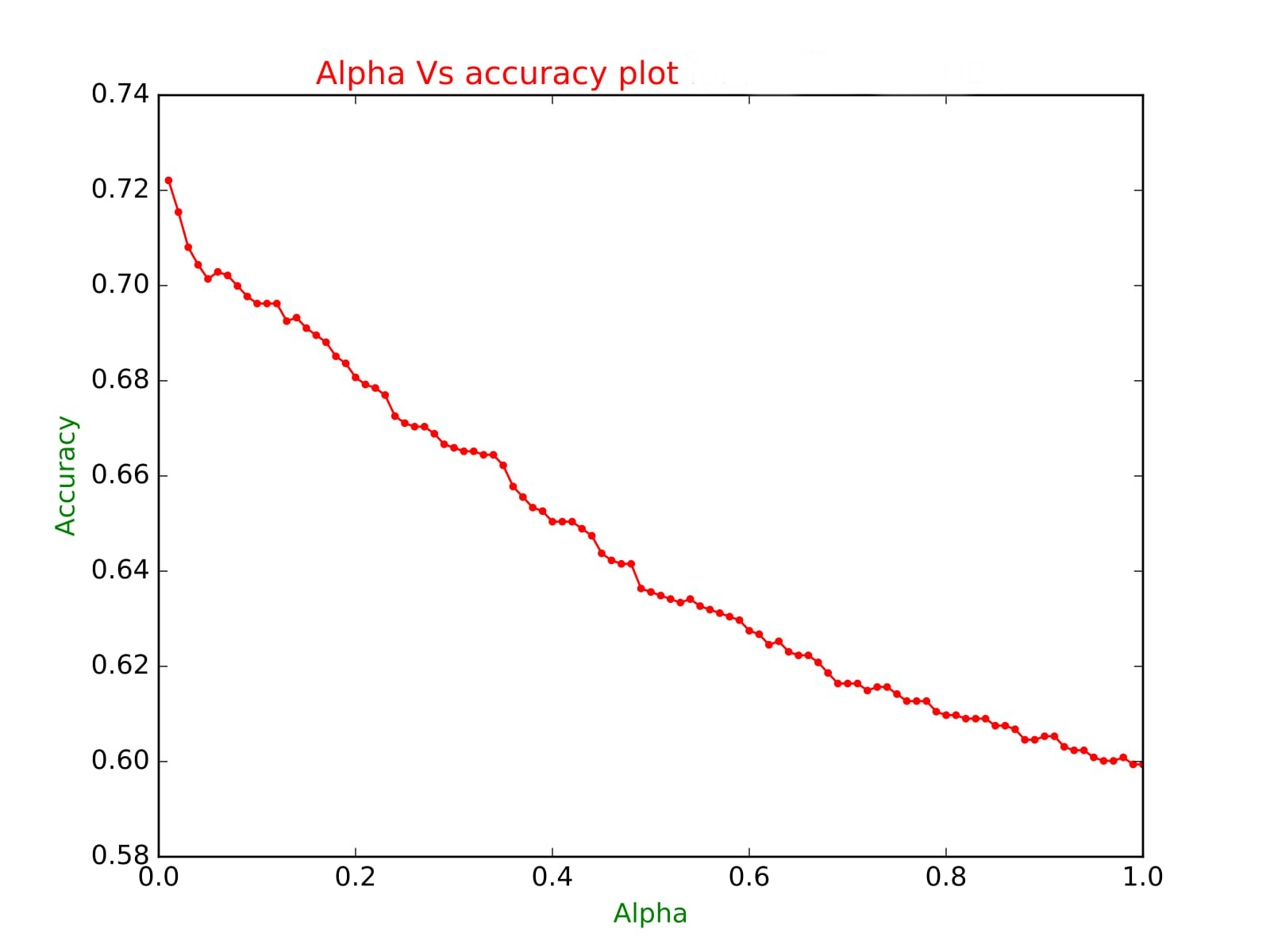
3(a) code provided as bayes.py

**3(b)**

**The below is the graph varying smoothing parameter alpha Vs the accuracy in range from 0 to 100%( .74 means 74%) for Bernoulli Naïve Bayes.**

****

**As we can see, the accuracy falls with increase in smoothing for Bernoulli naïve Bayes. The peak accuracy was observed at alpha = 0.01 and worst accuracy was observed at 1.0**

**Best accuracy at alpha = 0.01**

|  |  |  |
| --- | --- | --- |
| Alpha | Train + Predict Time | Accuracy in % |
| 0.01 | 180.972128868 | 72.20990392 |

**Worst accuracy at alpha = 1.0**

|  |  |  |
| --- | --- | --- |
| Alpha | Training Time | Accuracy in % |
| 1.0 | 182.027626038 | 59.9409 |

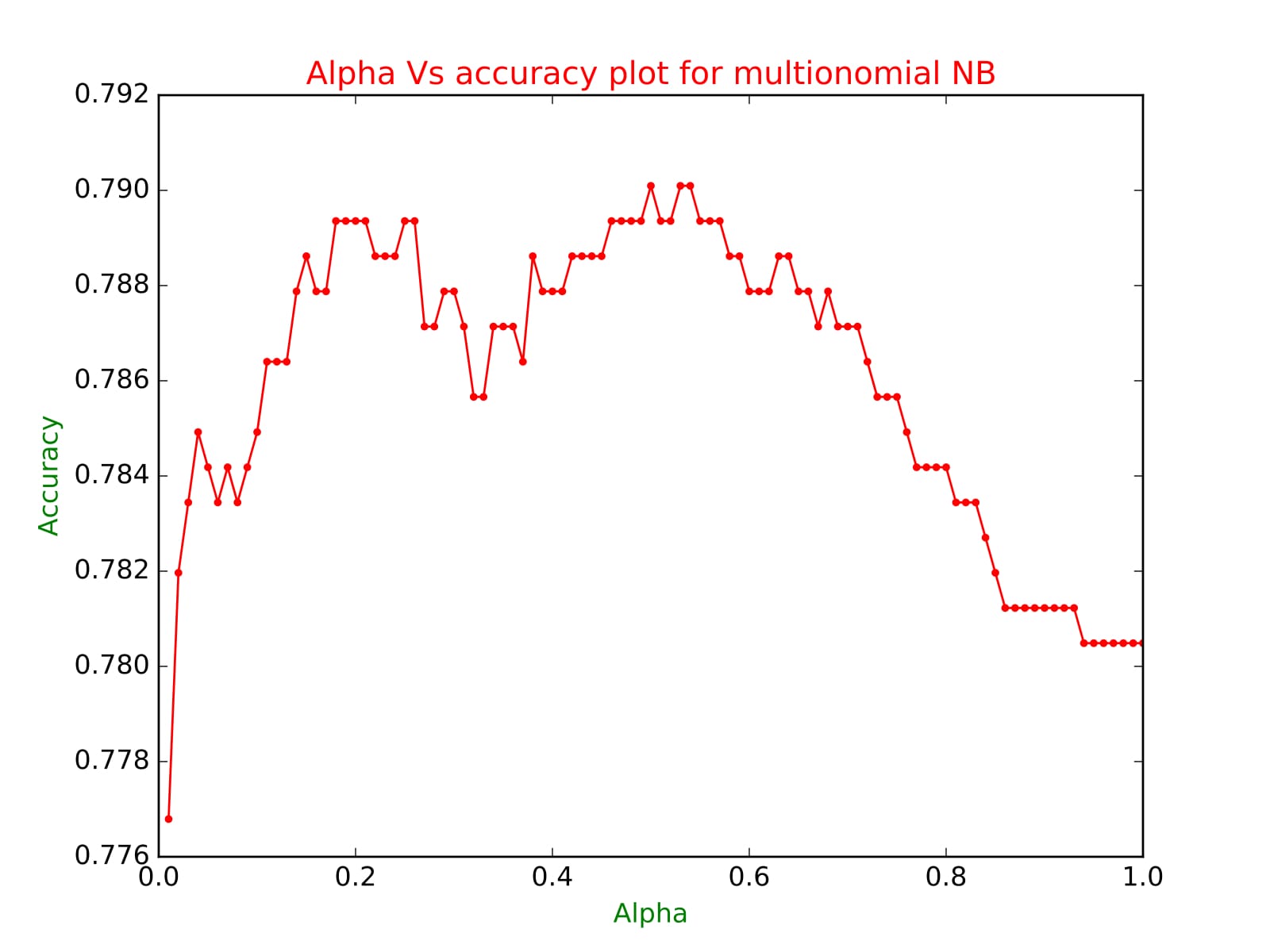
**More accuracy and test data are provided as a comparison with Multinomial**

**3(c):** The Multinomial Naïve Bayes was implemented using the python code submitted along with the assignment. The average train plus predict time was 100 secs for the algorithm. The smoothing parameter alpha was varied from 0.01 till 1.0 with regular increments of 0.01. The below graph was obtained as a result of the iterated train + predict routine below along with results from Bernoulli to compare.

|  |
| --- |
| for alpha in [float(j) / 100 for j in range(1, 101, 1)]:  print '--------------------------------------------------------------------------------------'  ta = time()  clf1 =MyMultinomialBayesClassifier(alpha)  clf1.train(X\_train,y\_train)  y\_pred1 = clf1.predict(X\_test)  acc.append(np.mean((y\_test-y\_pred1)==0))  alp.append(alpha)  tb = time()  print "training time: " + str(tb-ta) + " accuracy,alpha is: " + str(np.mean((y\_test-y\_pred1)==0)) +","+str(alpha) |

Output:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Multinomial Naïve Bayes | | | Bernoulli Naïve Bayes | | |
| Alpha | Train + Predict Time | Accuracy in % | Alpha | Train + Predict Time | Accuracy in % |
| 0.01 | 114.4258959 | 77.67923134 | 0.01 | 180.972128868 | 72.20990392 |
| 0.02 | 118.8346491 | 78.19660015 | 0.02 | 180.993458033 | 71.54471545 |
| 0.03 | 114.3443279 | 78.34441981 | 0.03 | 181.906514168 | 70.80561715 |
| 0.04 | 115.583375 | 78.49223947 | 0.04 | 181.214675903 | 70.436068 |
| 0.05 | 128.059237 | 78.41832964 | 0.05 | 181.357953072 | 70.14042868 |
| 0.06 | 122.699976 | 78.34441981 | 0.06 | 180.523658037 | 70.28824834 |
| 0.07 | 125.8528039 | 78.41832964 | 0.07 | 180.556339025 | 70.21433851 |
| 0.08 | 126.668664 | 78.34441981 | 0.08 | 180.881312132 | 69.99260902 |
| 0.09 | 129.751179 | 78.41832964 | 0.09 | 181.252819061 | 69.77087953 |
| 0.1 | 125.6159868 | 78.49223947 | 0.1 | 181.155525208 | 69.62305987 |
| 0.11 | 123.085779 | 78.64005913 | 0.11 | 200.572222948 | 69.62305987 |
| 0.12 | 121.728374 | 78.64005913 | 0.12 | 191.741470814 | 69.62305987 |
| 0.13 | 100.2698801 | 78.64005913 | 0.13 | 191.426566839 | 69.25351072 |
| 0.14 | 109.5300109 | 78.78787879 | 0.14 | 186.681269169 | 69.32742055 |
| 0.15 | 115.7458761 | 78.86178862 | 0.15 | 191.280533075 | 69.10569106 |
| 0.16 | 113.397073 | 78.78787879 | 0.16 | 187.281642914 | 68.9578714 |
| 0.17 | 103.0560992 | 78.78787879 | 0.17 | 190.664875031 | 68.81005174 |
| 0.18 | 100.0709479 | 78.93569845 | 0.18 | 191.539718866 | 68.51441242 |
| 0.19 | 98.784235 | 78.93569845 | 0.19 | 191.764376879 | 68.36659276 |
| 0.2 | 98.45435405 | 78.93569845 | 0.2 | 203.311460018 | 68.07095344 |
| 0.21 | 98.74279189 | 78.93569845 | 0.21 | 203.570017815 | 67.92313378 |
| 0.22 | 98.48586679 | 78.86178862 | 0.22 | 209.175596952 | 67.84922395 |
| 0.23 | 99.37589216 | 78.86178862 | 0.23 | 212.391278028 | 67.70140429 |
| 0.24 | 98.75725293 | 78.86178862 | 0.24 | 209.657526016 | 67.25794531 |
| 0.25 | 98.90837312 | 78.93569845 | 0.25 | 206.349309921 | 67.11012565 |
| 0.26 | 98.46985698 | 78.93569845 | 0.26 | 207.282773018 | 67.03621582 |
| 0.27 | 99.19650888 | 78.71396896 | 0.27 | 200.754622936 | 67.03621582 |
| 0.28 | 98.43934202 | 78.71396896 | 0.28 | 194.748994112 | 66.88839616 |
| 0.29 | 98.95235896 | 78.78787879 | 0.29 | 194.700381994 | 66.66666667 |
| 0.3 | 98.80279708 | 78.78787879 | 0.3 | 193.655354977 | 66.59275684 |
| 0.31 | 98.56015992 | 78.71396896 | 0.31 | 196.115650892 | 66.51884701 |
| 0.32 | 98.53452802 | 78.5661493 | 0.32 | 607.417876005 | 66.51884701 |
| 0.33 | 98.79165196 | 78.5661493 | 0.33 | 198.51395607 | 66.44493718 |
| 0.34 | 99.03140116 | 78.71396896 | 0.34 | 182.078330994 | 66.44493718 |
| 0.35 | 98.63639307 | 78.71396896 | 0.35 | 1741.851686 | 66.22320769 |
| 0.36 | 98.60356188 | 78.71396896 | 0.36 | 208.462009192 | 65.77974871 |
| 0.37 | 99.65364599 | 78.64005913 | 0.37 | 191.145642042 | 65.55801922 |
| 0.38 | 99.05669618 | 78.86178862 | 0.38 | 192.191334009 | 65.33628973 |
| 0.39 | 98.60373616 | 78.78787879 | 0.39 | 208.232443094 | 65.2623799 |
| 0.4 | 98.76156497 | 78.78787879 | 0.4 | 216.478410006 | 65.04065041 |
| 0.41 | 99.22933912 | 78.78787879 | 0.41 | 223.206805944 | 65.04065041 |
| 0.42 | 98.88934088 | 78.86178862 | 0.42 | 204.449005842 | 65.04065041 |
| 0.43 | 98.65378308 | 78.86178862 | 0.43 | 204.423100948 | 64.89283075 |
| 0.44 | 98.74117804 | 78.86178862 | 0.44 | 196.80291605 | 64.74501109 |
| 0.45 | 99.03860593 | 78.86178862 | 0.45 | 197.257192135 | 64.37546194 |
| 0.46 | 98.65555501 | 78.93569845 | 0.46 | 193.252833128 | 64.22764228 |
| 0.47 | 98.83235288 | 78.93569845 | 0.47 | 197.07911706 | 64.15373245 |
| 0.48 | 104.8705359 | 78.93569845 | 0.48 | 191.430484056 | 64.15373245 |
| 0.49 | 99.16663098 | 78.93569845 | 0.49 | 194.766113997 | 63.63636364 |
| 0.5 | 98.83667493 | 79.00960828 | 0.5 | 188.645231962 | 63.56245381 |
| 0.51 | 98.91985297 | 78.93569845 | 0.51 | 190.975454807 | 63.48854398 |
| 0.52 | 99.17009306 | 78.93569845 | 0.52 | 198.837553024 | 63.41463415 |
| 0.53 | 98.80243802 | 79.00960828 | 0.53 | 201.750118017 | 63.34072432 |
| 0.54 | 99.21981502 | 79.00960828 | 0.54 | 186.31755805 | 63.41463415 |
| 0.55 | 99.0750339 | 78.93569845 | 0.55 | 192.485630989 | 63.26681449 |
| 0.56 | 98.90470409 | 78.93569845 | 0.56 | 199.119551897 | 63.19290466 |
| 0.57 | 99.303334 | 78.93569845 | 0.57 | 193.173141003 | 63.11899483 |
| 0.58 | 98.97857904 | 78.86178862 | 0.58 | 190.716257811 | 63.045085 |
| 0.59 | 99.571347 | 78.86178862 | 0.59 | 191.579383135 | 62.97117517 |
| 0.6 | 99.31500506 | 78.78787879 | 0.6 | 200.226572037 | 62.74944568 |
| 0.61 | 99.47051311 | 78.78787879 | 0.61 | 202.397355795 | 62.67553585 |
| 0.62 | 99.21742797 | 78.78787879 | 0.62 | 198.145197153 | 62.45380636 |
| 0.63 | 99.1808331 | 78.86178862 | 0.63 | 201.759388208 | 62.52771619 |
| 0.64 | 99.22876 | 78.86178862 | 0.64 | 198.27939105 | 62.3059867 |
| 0.65 | 99.63462996 | 78.78787879 | 0.65 | 196.481271982 | 62.23207687 |
| 0.66 | 98.96145296 | 78.78787879 | 0.66 | 206.901082993 | 62.23207687 |
| 0.67 | 99.06072402 | 78.71396896 | 0.67 | 213.805708885 | 62.08425721 |
| 0.68 | 98.83556986 | 78.78787879 | 0.68 | 205.105123043 | 61.86252772 |
| 0.69 | 99.31913209 | 78.71396896 | 0.69 | 205.097599983 | 61.64079823 |
| 0.7 | 99.05826807 | 78.71396896 | 0.7 | 187.483760118 | 61.64079823 |
| 0.71 | 99.29860592 | 78.71396896 | 0.71 | 180.524951935 | 61.64079823 |
| 0.72 | 99.36991382 | 78.64005913 | 0.72 | 183.879939079 | 61.49297857 |
| 0.73 | 99.4142549 | 78.5661493 | 0.73 | 179.883658886 | 61.5668884 |
| 0.74 | 99.046206 | 78.5661493 | 0.74 | 182.215018988 | 61.5668884 |
| 0.75 | 99.77153397 | 78.5661493 | 0.75 | 183.879378796 | 61.41906874 |
| 0.76 | 99.30788589 | 78.49223947 | 0.76 | 182.213514805 | 61.27124908 |
| 0.77 | 99.20675707 | 78.41832964 | 0.77 | 179.467445135 | 61.27124908 |
| 0.78 | 99.32484007 | 78.41832964 | 0.78 | 183.681913137 | 61.27124908 |
| 0.79 | 99.24850202 | 78.41832964 | 0.79 | 180.937854052 | 61.04951959 |
| 0.8 | 99.32736897 | 78.41832964 | 0.8 | 183.96787405 | 60.97560976 |
| 0.81 | 99.52071095 | 78.34441981 | 0.81 | 181.619086027 | 60.97560976 |
| 0.82 | 99.23537612 | 78.34441981 | 0.82 | 183.536737919 | 60.90169993 |
| 0.83 | 98.96483397 | 78.34441981 | 0.83 | 180.616222143 | 60.90169993 |
| 0.84 | 99.05012012 | 78.27050998 | 0.84 | 184.489248991 | 60.90169993 |
| 0.85 | 99.10183787 | 78.19660015 | 0.85 | 180.347239017 | 60.75388027 |
| 0.86 | 99.60344195 | 78.12269032 | 0.86 | 182.432384968 | 60.75388027 |
| 0.87 | 99.639678 | 78.12269032 | 0.87 | 182.693991899 | 60.67997044 |
| 0.88 | 99.33443284 | 78.12269032 | 0.88 | 182.828011036 | 60.45824095 |
| 0.89 | 99.73030996 | 78.12269032 | 0.89 | 182.396009922 | 60.45824095 |
| 0.9 | 99.66559196 | 78.12269032 | 0.9 | 183.983263016 | 60.53215078 |
| 0.91 | 99.26899195 | 78.12269032 | 0.91 | 179.829323053 | 60.53215078 |
| 0.92 | 104.7267599 | 78.12269032 | 0.92 | 182.041614056 | 60.31042129 |
| 0.93 | 100.6318309 | 78.12269032 | 0.93 | 181.790756941 | 60.23651146 |
| 0.94 | 100.556495 | 78.04878049 | 0.94 | 183.343018055 | 60.23651146 |
| 0.95 | 100.845294 | 78.04878049 | 0.95 | 181.53332305 | 60.0886918 |
| 0.96 | 100.9083838 | 78.04878049 | 0.96 | 183.671306133 | 60.01478197 |
| 0.97 | 100.6127031 | 78.04878049 | 0.97 | 180.213320017 | 60.01478197 |
| 0.98 | 111.5329812 | 78.04878049 | 0.98 | 182.958544016 | 60.0886918 |
| 0.99 | 102.2297781 | 78.04878049 | 0.99 | 181.505044937 | 59.94087214 |
| 1 | 120.578788 | 78.04878049 | 1 | 182.027626038 | 59.9409 |



**The peak accuracy was obtained at alpha is 0.5 and accuracy was 79 % approximately which is an acceptable level of accuracy considering that the dimensions of the training and test data are very high.**

**(Refer to the excel snapshot above)**

**The Multinomial with non-binary achieved better results both in terms of accuracy and performance efficiency when compared to Bernoulli naïve Bayes with binary data**. As displayed in 3(b), the results of Bernoulli naïve Bayes; the accuracy and performance efficiency were both poor for Bernoulli NB as compared to Multinomial NB for any instance of smoothing parameter alpha. The performance efficiency of the Multinomial NB can be attributed to the fact that we do not consider absence of a feature which is not the case for Bernoulli NB where it needs to be considered.