Q1.

For the first question I used a script that read my blogdata from assignment 8, blogdata1.txt, and it printed out the cosine distance of the 500 dimension using knnestimate() to find the nearest neighbors for http://f-measure.blogpost.com/ and http://ws-dl.blogpost.com/ with the restraints k=(1,2,5,10,20), shown below:

K =	1	2	5	10	20
f-measure	44	26.5	16.4	24.7	33.2
ws-dl	50	33.5	42.4	35.7	45.55

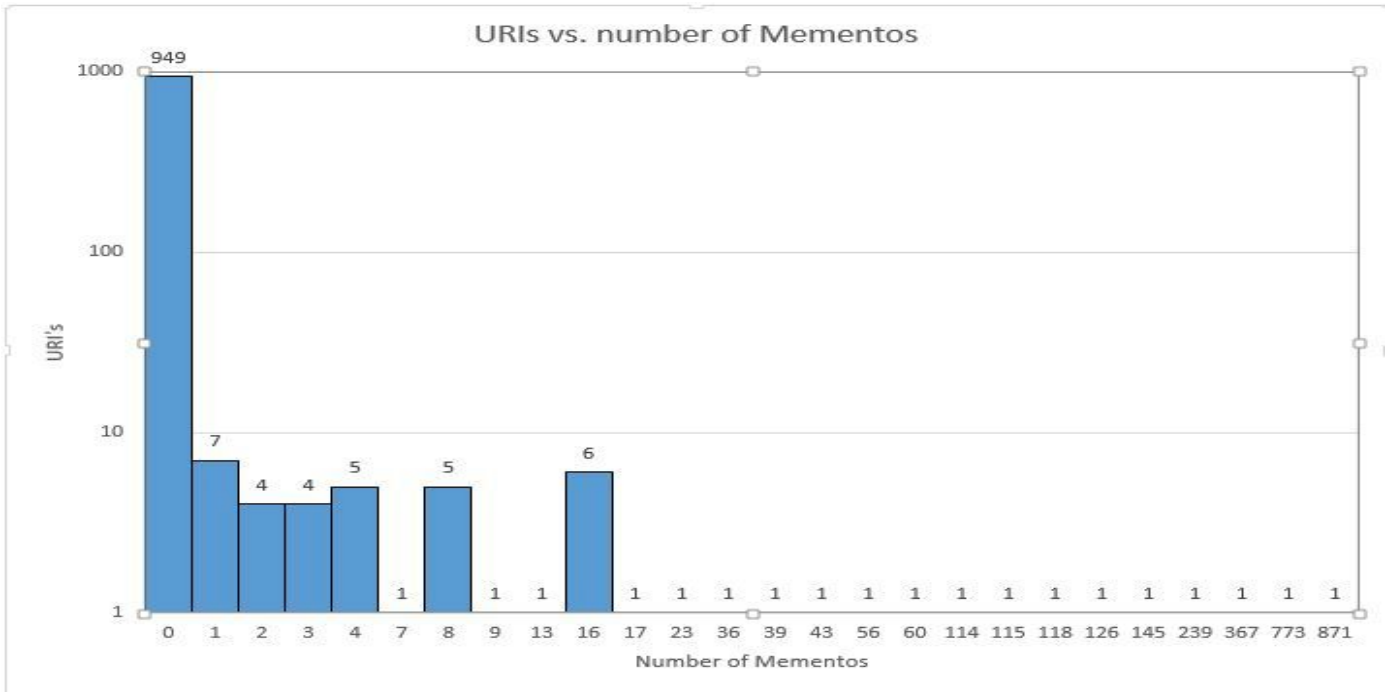
Q2.

For this question for rerunning A9Q2 but using LIBSVM. Since I had 5 categories I had to run the program 5 times in order to get all the necessary data. The script I used allowed me to manually assign the first 50 classifications to its true values while it guess on the rest of them They were all ran under both their true state and their opposite state. Based off the 10-fold cross-validation Music has the highest mean for being correct for each attempt.

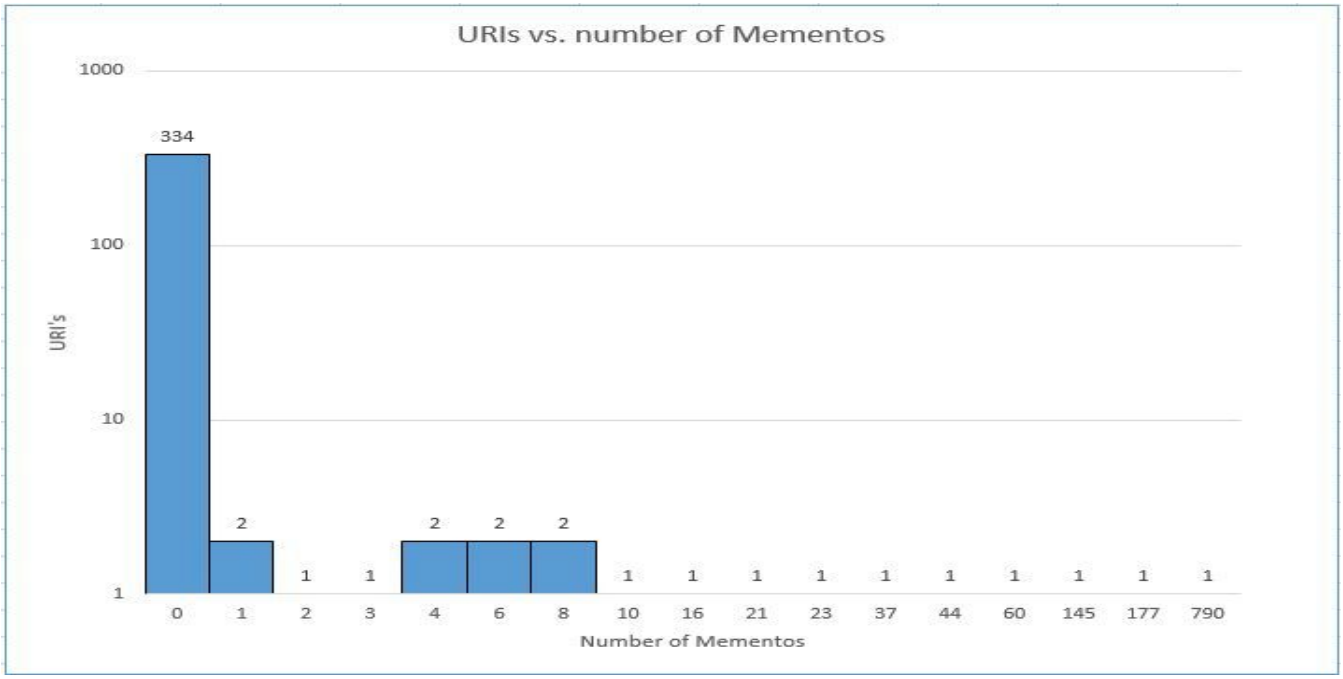
Category	**Try 1**	**Try 2**	**Try 3**	**Try 4**	**Try 5**	**Try 6**	**Try 7**	**Try 8**	**Try 9**	**Try 10**	**Mean**
2-Music	0.9	0.9	0.9	1	1	1	1	1	1	1	0.97
3-Movie	0.8	0.9	0.9	1	1	1	1	1	1	0.9	0.95
5-Other	0.8	0.8	0.8	0.9	0.9	1	1	0.9	1	0.9	0.90
1-Book	0.8	0.8	0.9	0.9	0.8	1	1	0.9	1	0.8	0.89
4-Show	0.7	0.8	0.8	0.9	0.8	0.9	0.9	0.8	0.8	0.7	0.81

EC1.

For this question I found my twitter links from assignment 2 and created a new histogram to compare it to the one from Assignment 2 Question 2. Over the past few months my chart has changed quite a bit based off the information given. The Histograms are shown below for comparison



Histogram from A2



Histogram from A10