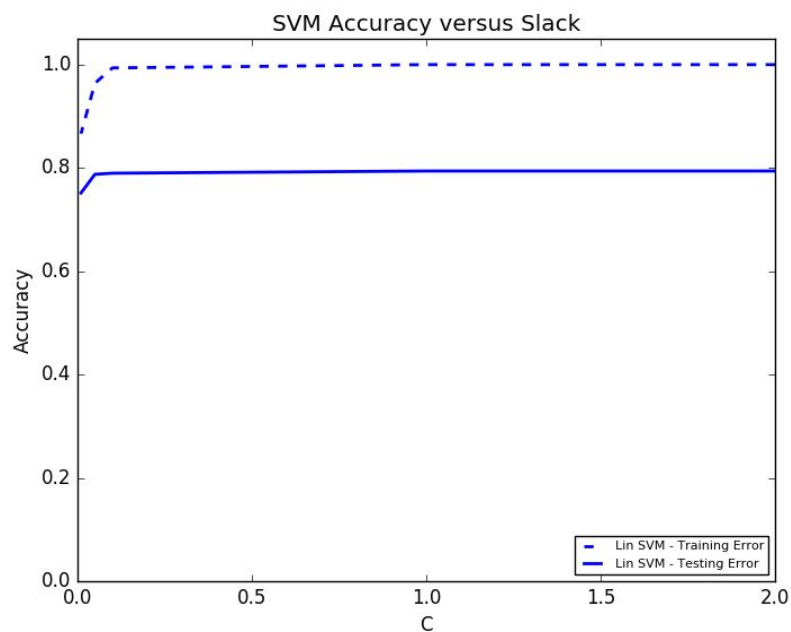


1. Linear SVM

- see appendix
- The optimal parameter appears to be $C=0.1$ or 0.05 . A low parameter will not allow enough slack and won't come up with a good fit of the data. However, one that is too high will allow too many margin violations.
- Neither models seem to over/underfit. Both maintain a fairly steady difference between testing and training accuracy.

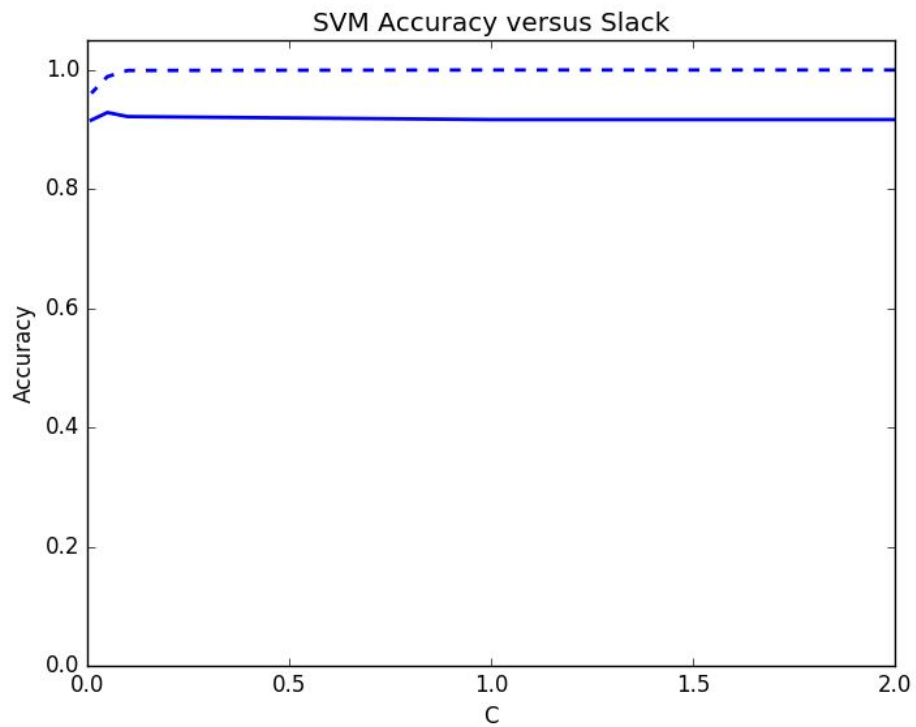
MNIST



a. MNIST DATA

Hyper-Parameter Value	Training Accuracy	Test Accuracy
2	1.0	0.7876857749469215
1	1.0	0.7876857749469215
0.1	0.9936305732484076	0.7940552016985138
0.05	0.9639065817409767	0.7813163481953291
0.01	0.8662420382165605	0.7515923566878981

20NG

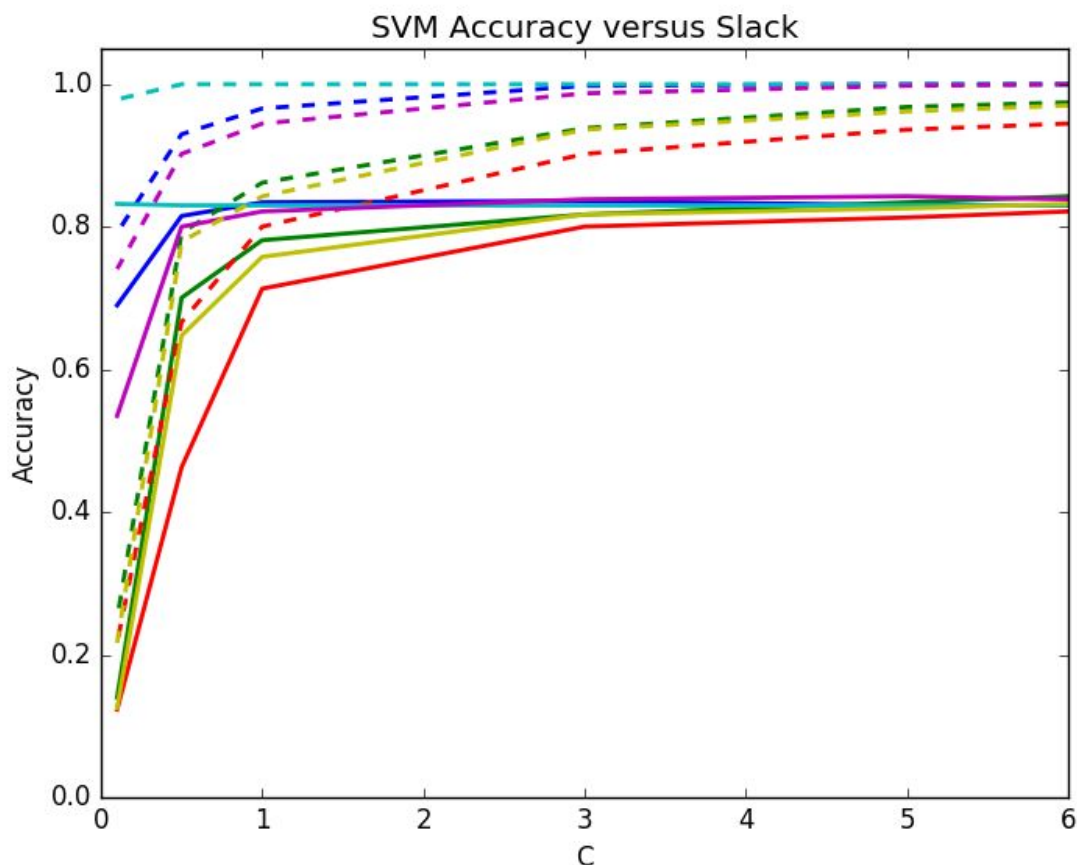


Hyper-Parameter Value	Training Accuracy	Test Accuracy
2	[1.0	0.9167374681393373]
1	[1.0	0.9167374681393373]
0.1	[0.9991503823279524	0.9218351741716228]
0.05	[0.9889549702633815	0.9286321155480034]
0.01	[0.9609175870858114	0.9158878504672897]

2. Polynomial SVM

- a. see below
- b. On the MNIST data, the best performing model has a degree of 2, coefficient of 2, and slack of 6. Looking at the performance of the various models, the most important feature seemed to be the slack; those models with a very low slack performed consistently badly, regardless of the degree and coefficient. The best model for the 20NG set had a degree of 2, coefficient of 10, and slack of 6. The performance on this dataset seemed to be more variant; while some testing scores were much higher than what was achieved on the other dataset,
- c. Lower slack parameters caused very badly fitting scores. When all three parameters were low, both models performed terribly, demonstrating underfitting. When only the slack was low, the model seemed to overfit. This makes sense because the model is trying to match the training dataset as exactly as possible without allowing many errors, which will not allow for generalization.

MNIST



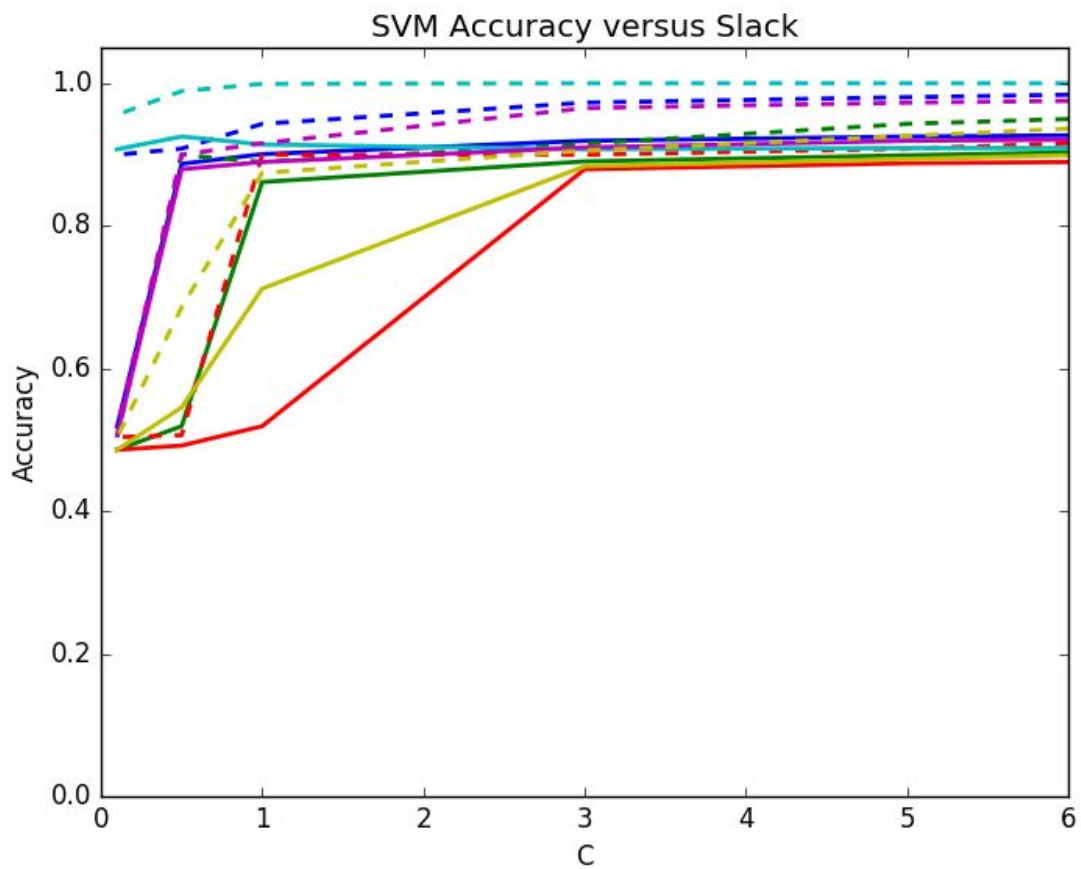
<i>Polynomial Degree</i>	<i>Coefficient</i>	<i>Slack</i>	Training	Testing
2	10	6	[1.0	0.8301486199575372]
2	10	5	[1.0	0.8301486199575372]
2	10	3	[0.9978768577494692	0.8365180467091295]
2	10	1	[0.9660297239915074	0.8343949044585988]
2	10	0.5	[0.9299363057324841	0.8152866242038217]
2	10	0.1	[0.7919320594479831	0.6900212314225053]
<u>2</u>	<u>2</u>	<u>6</u>	<u>[0.9745222929936306</u>	<u>0.8428874734607219]</u>
2	2	5	[0.9681528662420382	0.8343949044585988]
2	2	3	[0.9384288747346072	0.8174097664543525]
2	2	1	[0.861995753715499	0.7813163481953291]
2	2	0.5	[0.7919320594479831	0.7006369426751592]

2	2	0.1	[0.2505307855 626327	0.14012738853 503184]
2	1	6	[0.9447983014 861996	0.82165605095 5414]
2	1	5	[0.9363057324 840764	0.81316348195 32909]
2	1	3	[0.9023354564 755839	0.80042462845 01062]
2	1	1	[0.8004246284 501062	0.71337579617 83439]
2	1	0.5	[0.6666666666 666666	0.46284501061 571126]
2	1	0.1	[0.2165605095 5414013	0.12314225053 078556]
3	10	6	[1.0	0.83014861995 75372]
3	10	5	[1.0	0.83014861995 75372]
3	10	3	[1.0	0.83014861995 75372]
3	10	1	[1.0	0.83014861995 75372]
3	10	0.5	[1.0	0.83014861995

				75372]
3	10	0.1	[0.9787685774 946921	0.83227176220 8068]
3	2	6	[1.0	0.83864118895 96603]
3	2	5	[0.9978768577 494692	0.84288747346 07219]
3	2	3	[0.9872611464 968153	0.83864118895 96603]
3	2	1	[0.9447983014 861996	0.82165605095 5414]
3	2	0.5	[0.9023354564 755839	0.80042462845 01062]
3	2	0.1	[0.7409766454 352441	0.53503184713 3758]
3	1	6	[0.9702760084 92569	0.83227176220 8068]
3	1	5	[0.9617834394 904459	0.82590233545 64756]
3	1	3	[0.9363057324 840764	0.81740976645 43525]
3	1	1	[0.8428874734 607219	0.75796178343 94905]
3	1	0.5	[0.7791932059	

			447984	0.64755838641 18896]
3	1	0.1	[0.2165605095 5414013	0.12526539278 131635]

20NG



<i>Polynomial Degree</i>	<i>Coefficient</i>	<i>Slack</i>	Training	Testing
<u>2</u>	<u>10</u>	<u>6</u>	<u>[0.9838572642</u> <u>31096</u>	<u>0.92693288020</u> <u>39083]</u>
2	10	5	[0.9804587935 429057	0.92523364485 98131]

2	10	3	[0.9728122344 944775	0.91928632115 548]
2	10	1	[0.9430756159 728122	0.90059473237 04333]
2	10	0.5	[0.9082412914 188616	0.88700084961 7672]
2	10	0.1	[0.8997451146 983857	0.51826677994 90229]
2	2	6	[0.9498725573 491928	0.90314358538 6576]
2	2	5	[0.9430756159 728122	0.89889549702 63382]
2	2	3	[0.9150382327 952421	0.89039932030 58623]
2	2	1	[0.8895497026 338148	0.86151231945 62447]
2	2	0.5	[0.8997451146 983857	0.51911639762 10705]
2	2	0.1	[0.5038232795 242141	0.48598130841 121495]
2	1	6	[0.9158878504 672897	0.88954970263 38148]
2	1	5	[0.9082412914 188616	0.88785046728

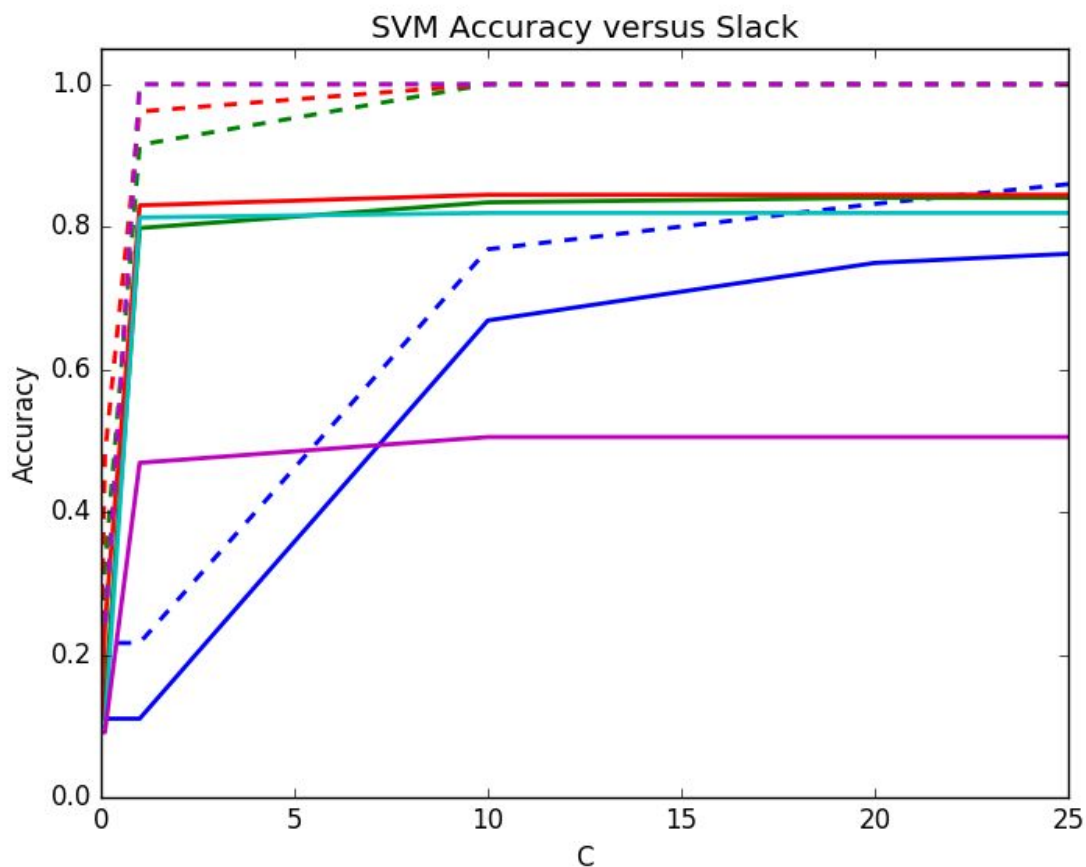
				97196]
2	1	3	[0.8997451146 983857	0.87935429056 92439]
2	1	1	[0.8997451146 983857	0.51911639762 10705]
2	1	0.5	[0.5063721325 403568	0.49192863211 5548]
2	1	0.1	[0.5038232795 242141	0.48598130841 121495]
3	10	6	[1.0	0.90909090909 09091]
3	10	5	[1.0	0.90909090909 09091]
3	10	3	[1.0	0.90739167374 6814]
3	10	1	[0.9991503823 279524	0.91418861512 31946]
3	10	0.5	[0.9889549702 633815	0.92523364485 98131]
3	10	0.1	[0.9558198810 535259	0.90739167374 6814]
3	2	6	[0.9753610875 106202	0.92098555649 95751]
3	2	5	[0.9728122344	

			944775	0.91928632115 548]
3	2	3	[0.9651656754 460493	0.90994052676 29567]
3	2	1	[0.9158878504 672897	0.88954970263 38148]
3	2	0.5	[0.8997451146 983857	0.87935429056 92439]
3	2	0.1	[0.5157179269 328802	0.50637213254 03568]
3	1	6	[0.9362786745 964317	0.89889549702 63382]
3	1	5	[0.9260832625 318607	0.89379779099 40527]
3	1	3	[0.9056924384 027187	0.88360237892 94817]
3	1	1	[0.8742565845 369583	0.71197960917 58709]
3	1	0.5	[0.6856414613 423959	0.54545454545 45454]
3	1	0.1	[0.5038232795 242141	0.48598130841 121495]

3. RBF

- See below
- For the MNIST data, the best model had a gamma of 0.01 and a slack of 25. The best for the 20NG set had a gamma of 0.005 and a slack of 10.
- This model didn't perform as well on either dataset. The model with a gamma of 0.1 and a slack of 25 (purple in the graph) overfit badly in both models. The model Any model with a low slack tended to do badly, some improving with a higher slack.

MNIST



- MNIST

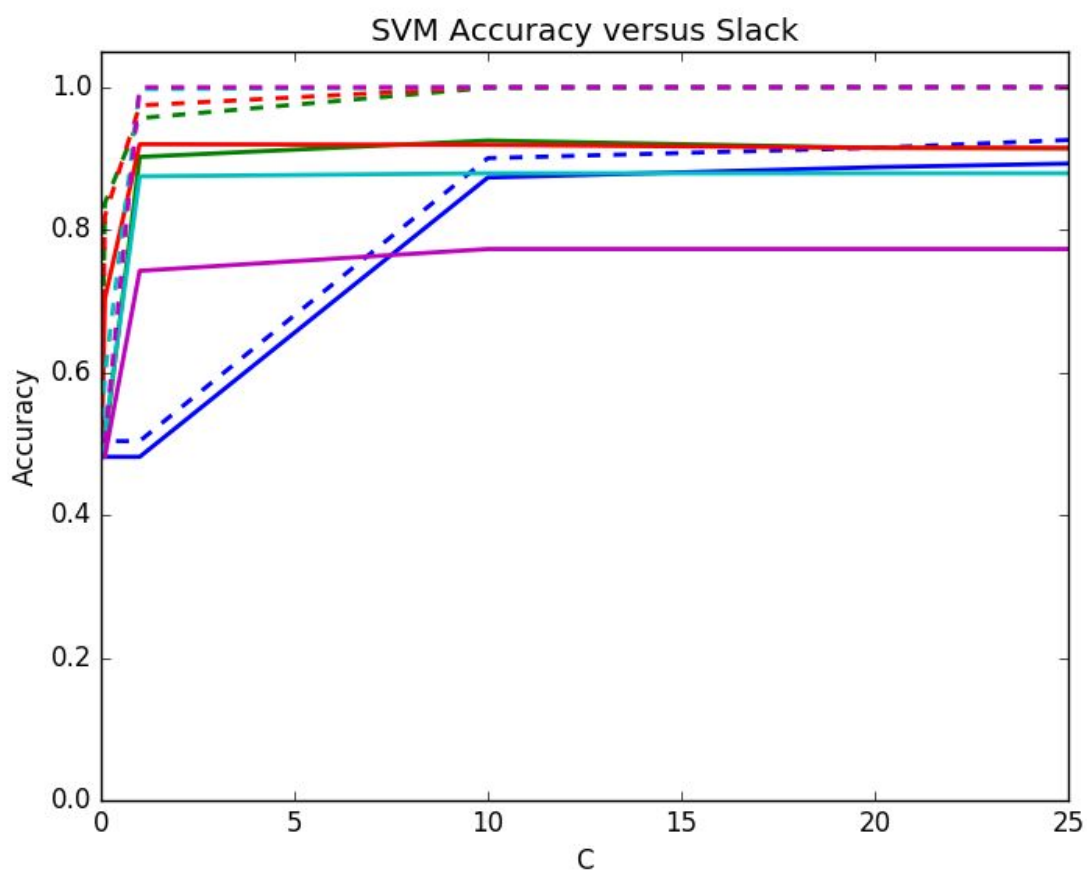
Gamma	Slack	Training Accuracy	Testing Accuracy
0.0001	25	[0.8598726114649	

		682	0.7622080679405521]
	20	[0.832271762208068	0.7494692144373672]
	10	[0.7685774946921444	0.6687898089171974]
	1	[0.21656050955414013	0.11040339702760085]
	0.1	[0.21656050955414013	0.11040339702760085]
	0.01	[0.21656050955414013	0.11040339702760085]
0.005	25	[1.0	0.8407643312101911]
	20	[1.0	0.8407643312101911]
	10	[1.0	0.8343949044585988]
	1	[0.9150743099787686	0.7983014861995754]
	0.1	[0.3205944798301486	0.13800424628450106]
	0.01	[0.21656050955414013	0.11040339702760085]

<u>0.01</u>	<u>25</u>	<u>[1.0</u>	<u>0.84501061571125</u> <u>27]</u>
	20	[1.0	0.84501061571125 27]
	10	[1.0	0.84501061571125 27]
	1	[0.9617834394904 459	0.83014861995753 72]
	0.1	[0.4798301486199 5754	0.23991507430997 877]
	0.01	[0.2229299363057 325	0.11040339702760 085]
0.05	25	[1.0	0.81953290870488 33]
	20	[1.0	0.81953290870488 33]
	10	[1.0	0.81953290870488 33]
	1	[1.0	0.81316348195329 09]
	0.1	[0.2314225053078 5563	0.09978768577494 693]
	0.01	[0.2292993630573 2485	0.09978768577494

			693]
0.1	25	[1.0	0.505307855626327]
	20	[1.0	0.505307855626327]
	10	[1.0	0.505307855626327]
	1	[1.0	0.4692144373673036]
	0.1	[0.22929936305732485	0.09129511677282377]
	0.01	[0.22929936305732485	0.09129511677282377]

20NG



Gamma	Slack	Training Accuracy	Testing Accuracy
0.0001	25	[0.9260832625318607	0.892948173322005]
	20	[0.9150382327952421	0.8878504672897196]
	10	[0.9005947323704333	0.8734069668649108]
	1	[0.5038232795242141	0.4817332200509771]

	0.1	[0.5038232795242 141	0.48173322005097 71]
	0.01	[0.5038232795242 141	0.48173322005097 71]
0.005	25	[1.0	0.91418861512319 46]
	20	[1.0	0.91503823279524 21]
<u>0.005</u>	<u>10</u>	<u>[0.9991503823279 524</u>	<u>0.92523364485981 31]</u>
	1	[0.9566694987255 735	0.90229396771452 84]
	0.1	[0.8377230246389 125	0.50042480883602 38]
	0.01	[0.5038232795242 141	0.48173322005097 71]
0.01	25	[1.0	0.91503823279524 21]
	20	[1.0	0.91503823279524 21]
	10	[1.0	0.91928632115548]
	1	[0.9745114698385 726	0.92013593882752

			76]
	0.1	[0.8190314358538 657	0.70348343245539 5]
	0.01	[0.5038232795242 141	0.48173322005097 71]
0.05	25	[1.0	0.87935429056924 39]
	20	[1.0	0.87935429056924 39]
	10	[1.0	0.87935429056924 39]
	1	[0.9974511469838 573	0.87510620220900 59]
	0.1	[0.5964316057774 002	0.51401869158878 5]
	0.01	[0.5038232795242 141	0.48173322005097 71]
0.1	25	[1.0	0.77315208156329 65]
	20	[1.0	0.77315208156329 65]
	10	[1.0	0.77315208156329 65]
	1	[1.0	

			0.7425658453695837]
	0.1	[0.5038232795242141	0.4817332200509771]
	0.01	[0.5038232795242141	0.4817332200509771]