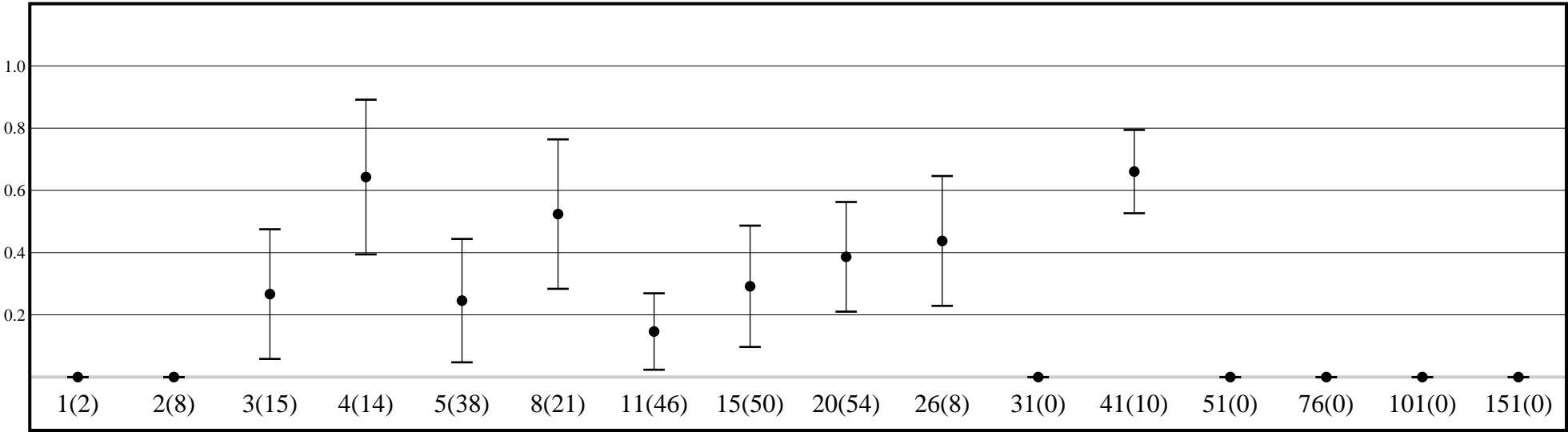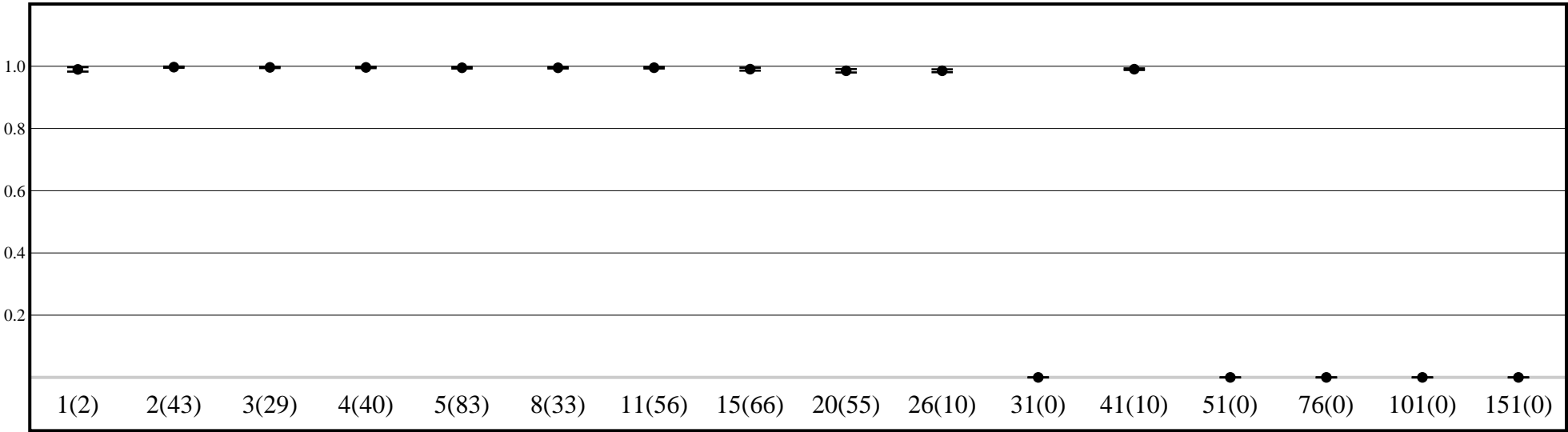| ID	H5GK5HK04
Comment	unwanteds 1000 times stronger than generic background (weighted n calc)
Date	Mon Mar 26 13:09:27 +0930 2007
SVN Version	http://svn.winnow.peerworks.org/branches/winnow-BUG312?rev=995
Total Classifications	109032
True Positive	751
False Negatives	1352
True Negative	106157
False Positive	772
Bias	1
Unknown Word Prob	0.5
Unknown Word Strength	0.45
Min Prob Strength	0.1
Max Discriminators	150

Test Results Categorized by Training Count

These graphs display the cross validation test results for a classifier arranged by the number of items a classifier was trained on for a given tag. The X axis displays the number of training items, where there is a gap, the label defines the lower bound of a range that extends to the next label. The number in brackets is the number of tags that fall into that 'bin'. The Y Axis is the mean true positive or true negative rate for the classifiers that fall in the 'bin'. Standard Deviation is shown as error bars on the chart.

True Positive Results
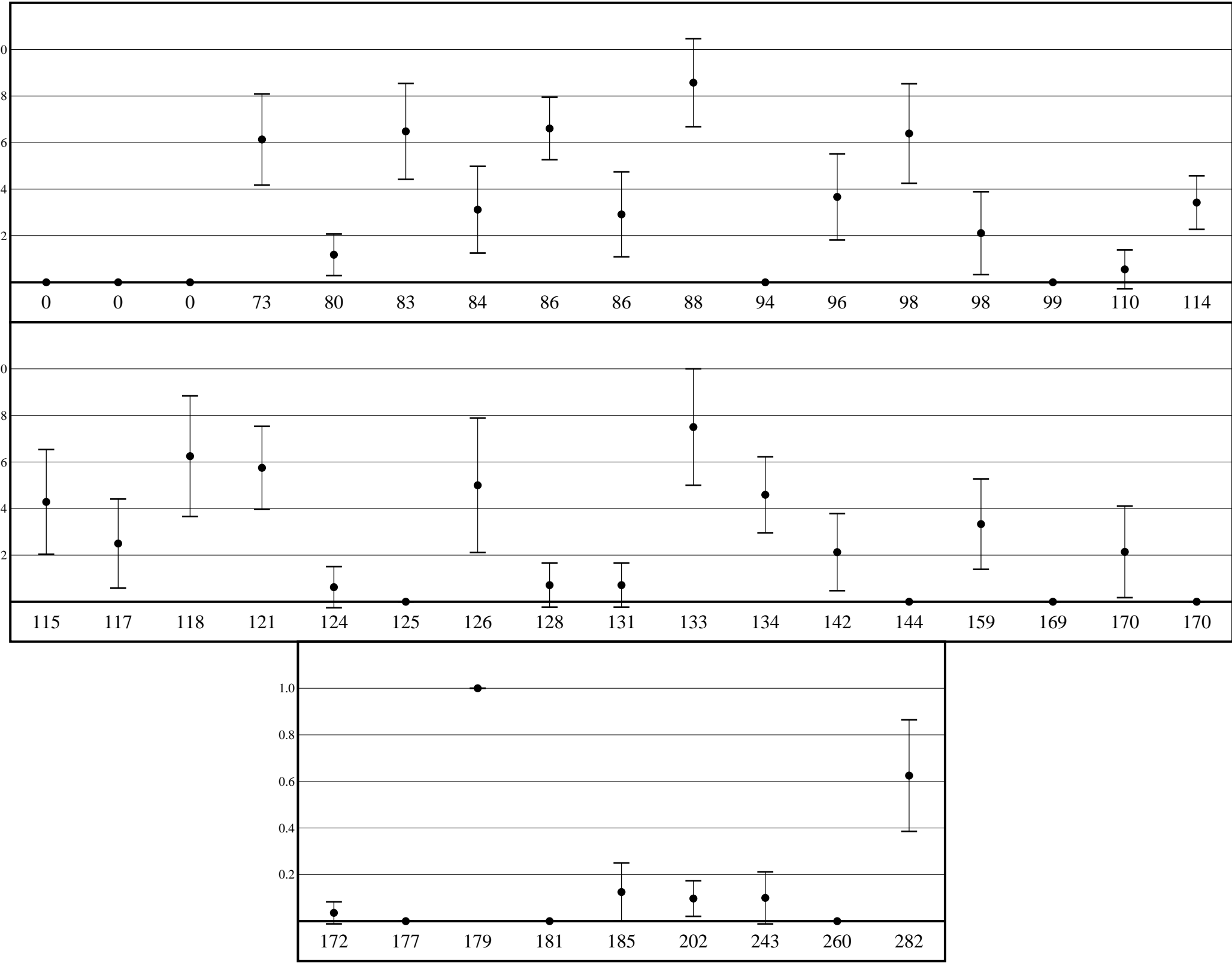

True Negative Results

Tags in bins

Bin	Tags
1..1	ande-publishing,
2..2	ande-spam, ande-language, ande-astronomy, ande-publishing, ande-endangered_species, ande-baseball, ande-cats, ande-*missing_content,
3..3	ande-ancient_world, ande-solar_energy, ande-language, ande-tags, ande-astronomy, ande-endangered_species, ande-baseball, ande-privacy, ande-open_source,
4..4	ande-ancient_world, ande-recipe, ande-solar_energy, ande-language, ande-tags, ande-endangered_species, ande-privacy, ande-open_source,
5..7	ande-the_media, ande-recipe, ande-philosophical, ande-humor, ande-curiosities, ande-solar_energy, ande-tags, ande-reading, ande-evolution, ande-antiscience, ande-nature_photography, ande-open_source,
8..10	ande-philosophical, ande-earth, ande-brain, ande-evolution, ande-antiscience, ande-book_review,
11..14	ande-writing, ande-online_culture, ande-biology, ande-education, ande-*character_encoding, ande-Vista, ande-earth, ande-health_&_medicine, ande-music, ande-collaboration, ande-design, ande-book_review,
15..19	ande-writing, ande-online_culture, ande-biology, ande-Google, ande-education, ande-Vista, ande-green, ande-health_&_medicine, ande-music, ande-collaboration, ande-food, ande-design,
20..25	ande-social_networking, ande-Google, ande-Internet, ande-green, ande-food, ande-Mac, ande-Web_2.0,
26..30	ande-social_networking, ande-Internet, ande-Web_2.0,
31..40	
41..50	ande-climate_change,
51..75	
76..100	
101..150	
151..1000	

Individual Tag Scores

Tag	Instances	Ave Item Size (Distinct Tokens)	Ave Distinct Tokens in Pool	TP	TN
ande-*missing_content	0	0	0	0.00 (0.00)	1.00 (0.00)
ande-cats	0	0	0	0.00 (0.00)	1.00 (0.00)
ande-spam	0	0	0	0.00 (0.00)	1.00 (0.00)
ande-publishing	2	181	181	0.00 (0.00)	0.99 (0.01)
ande-baseball	3	260	520	0.00 (0.00)	1.00 (0.00)
ande-astronomy	3	99	197	0.00 (0.00)	1.00 (0.00)
ande-ancient_world	4	133	400	0.75 (0.50)	1.00 (0.00)
ande-endangered_species	4	144	384	0.00 (0.00)	1.00 (0.00)
ande-privacy	4	170	510	0.00 (0.00)	1.00 (0.00)
ande-language	4	177	479	0.00 (0.00)	1.00 (0.00)
ande-solar_energy	5	282	1023	0.62 (0.48)	1.00 (0.00)
ande-tags	5	185	685	0.12 (0.25)	0.99 (0.01)
ande-open_source	5	126	464	0.50 (0.58)	0.99 (0.01)
ande-recipe	5	179	715	1.00 (0.00)	0.99 (0.01)
ande-reading	7	243	1366	0.10 (0.22)	1.00 (0.00)
ande-humor	7	125	708	0.00 (0.00)	1.00 (0.00)
ande-curiosities	7	94	550	0.00 (0.00)	0.99 (0.01)
ande-nature_photography	7	88	529	0.86 (0.38)	1.00 (0.00)
ande-the_media	7	169	949	0.00 (0.00)	0.99 (0.01)
ande-philosophical	8	170	1161	0.21 (0.39)	1.00 (0.00)
ande-antiscience	9	115	892	0.43 (0.45)	1.00 (0.00)
ande-evolution	10	98	815	0.64 (0.43)	0.99 (0.00)
ande-brain	10	118	1031	0.62 (0.52)	1.00 (0.00)
ande-earth	13	172	1920	0.04 (0.09)	0.99 (0.00)
ande-book_review	14	202	2362	0.10 (0.15)	1.00 (0.00)
ande-*character_encoding	14	131	1576	0.07 (0.19)	1.00 (0.00)
ande-writing	15	124	1638	0.06 (0.18)	1.00 (0.00)
ande-collaboration	16	117	1602	0.25 (0.38)	0.99 (0.00)
ande-health_&_medicine	16	86	1215	0.29 (0.36)	0.99 (0.01)
ande-music	17	110	1660	0.06 (0.17)	1.00 (0.01)
ande-design	17	159	2369	0.33 (0.39)	0.98 (0.01)
ande-online_culture	18	142	2284	0.21 (0.33)	0.99 (0.00)
ande-education	18	128	1975	0.07 (0.19)	1.00 (0.00)
ande-Vista	18	83	1339	0.65 (0.41)	0.98 (0.01)
ande-biology	19	98	1663	0.21 (0.36)	0.99 (0.00)
ande-food	23	96	1991	0.37 (0.37)	0.99 (0.01)
ande-green	24	134	2860	0.46 (0.33)	0.99 (0.01)
ande-Google	25	84	1906	0.31 (0.37)	0.99 (0.01)
ande-Mac	26	73	1707	0.61 (0.39)	0.97 (0.01)
ande-Internet	27	80	1925	0.12 (0.18)	1.00 (0.00)
ande-Web_2.0	28	121	3048	0.57 (0.36)	0.98 (0.01)
ande-social_networking	28	114	2845	0.34 (0.23)	0.98 (0.01)
ande-climate_change	48	86	3721	0.66 (0.27)	0.99 (0.01)
Average				0.29 (0.28)	0.99 (0.01)

True Positive Results by Average Distinct Tokens per Item

True Positive Results by Average Distinct Token Count in Pool