prelim_results

Nicholas Mitchell

January 23, 2016

Contents

1 Preliminary Results

Using several measures of accuracy and error, results are displayed for the same gradient boosting model (i.e. using the same input parameters) for each of the four data sets earlier defined, for each of the the lag values one to five.

These results all stem from the LOCF imputation method.

2 Two Models

1. Model 1: mstop = 5000, nu = 0.005

   (a) Table 1

Lag	Subset	Pred. Acc. (%)	MAE	MSE	RMSE	mstop avg.	mstop SD
1	Macro	81.10	2.98 (-3)	1.67 (-5)	4.09 (-3)	1369	1052
1	SA_{all}	67.07	5.02 (-3)	4.59 (-5)	6.78 (-3)	414	469
1	SA_{avg}	69.81	4.84 (-3)	4.36 (-5)	6.60 (-3)	398	379
1	Mix	82.62	2.94 (-3)	1.64 (-5)	4.05 (-3)	870	733
2	Macro	80.31	2.93 (-3)	1.65 (-5)	4.06 (-3)	1171	933
2	SA_{all}	63.97	5.01 (-3)	4.54 (-5)	6.74 (-3)	387	516
2	SA_{avg}	69.62	4.84 (-3)	4.31 (-5)	6.57 (-5)	358	359
2	Mix	82.90	2.94 (-3)	1.68 (-5)	4.10 (-3)	715	490
3	Macro	81.04	2.92 (-3)	1.62 (-5)	4.02 (-3)	1003	808
3	SA_{all}	64.83	5.01 (-3)	4.55 (-5)	6.75 (-3)	363	433
3	SA_{avg}	70.03	4.85 (-3)	4.27 (-5)	6.54 (-3)	349	378
3	Mix	83.33	2.92 (-3)	1.65 (-5)	4.06 (-3)	602	334
4	Macro	80.70	2.92 (-3)	1.62 (-5)	4.03 (-3)	871	689
4	SA_{all}	62.63	5.07 (-3)	4.58 (-5)	6.77 (-3)	351	537
4	SA_{avg}	70.14	4.89 (-3)	4.30 (-5)	6.55 (-3)	408	555
4	Mix	82.70	2.93 (-3)	1.67 (-5)	4.08 (-3)	543	244
5	Macro	81.90	2.90 (-3)	1.60 (-5)	4.00 (-3)	814	615
5	SA_{all}	63.19	5.00 (-3)	4.57 (-5)	6.76 (-3)	303	414
5	SA_{avg}	68.10	4.85 (-3)	4.25 (-5)	6.52 (-3)	370	444
5	Mix	82.67	2.92 (-3)	1.67 (-5)	4.09 (-3)	518	214

2. Model 2: mstop = 2000, nu = 0.05

   (a) Table 2

Lag	Subset	Pred. Acc. (%)	MAE	MSE	RMSE	mstop avg.	mstop SD
1	Macro	80.95	2.99 (-3)	1.67 (-5)	4.09 (-3)	151	192
1	SA_{all}	66.92	5.02 (-3)	4.59 (-5)	6.77 (-3)	41	46
1	SA_{avg}	69.51	4.83 (-3)	4.35 (-5)	6.59 (-3)	39	37
1	Mix	82.62	2.93 (-3)	1.64 (-5)	4.05 (-3)	91	116
2	Macro	80.46	2.93 (-3)	1.64 (-5)	4.05 (-3)	119	128
2	SA_{all}	64.73	5.01 (-3)	4.54 (-5)	6.74 (-3)	42	117
2	SA_{avg}	69.62	4.85 (-3)	4.32 (-5)	6.57 (-3)	37	62
2	Mix	82.75	2.95 (-3)	1.69 (-5)	4.11 (-3)	70	52
3	Macro	80.73	2.93 (-3)	1.62 (-5)	4.03 (-3)	104	129
3	SA_{all}	65.29	5.01 (-3)	4.54 (-5)	6.74 (-3)	38	87
3	SA_{avg}	70.03	4.86 (-3)	4.28 (-5)	6.54 (-3)	35	42
3	Mix	83.64	2.91 (-3)	1.64 (-5)	4.05 (-3)	59	35
4	Macro	80.55	2.92 (-3)	1.62 (-5)	4.02 (-3)	91	130
4	SA_{all}	63.40	5.07 (-3)	4.58 (-5)	6.77 (-3)	38	106
4	SA_{avg}	69.98	4.90 (-3)	4.31 (-5)	6.56 (-3)	42	66
4	Mix	82.54	2.92 (-3)	1.66 (-5)	4.07 (-3)	53	25
5	Macro	81.60	2.90 (-3)	1.61 (-5)	4.01 (-3)	81	74
5	SA_{all}	63.96	5.01 (-3)	4.58 (-5)	6.77 (-3)	35	110
5	SA_{avg}	67.94	4.85 (-3)	4.26 (-5)	6.52 (-3)	37	46
5	Mix	82.67	2.93 (-3)	1.68 (-5)	4.10 (-3)	50	19

3. Initial inspection

We define a measurement of error and success of a prediction as the sign accuracy. This is a simply test with binary response. Did we predict the sign of the market return correctly? If yes, we record '1', of not, '0'. The column Predictive accuracy is then defined as the percentage of correct predictions measured by the sign accuracy. The numerical descrepancy between the two values is not taken into consideration.

Inspecting first the predictive accuracy in Table 1 for the subsets Macro, SA_{all} and SA_{avg}, we see that every model unequivocally gains some traction on the data, each model performing better than a random selection producing 50% accuracy for a binary response.

We notice further that the traditional factors grouped in the Macro data subset offer the greatest level accuracy, being consistently over 80%. This is a results that one would come to expect. The two variations of the sentiment analysis data perform less convincingly, nonetheless they consistently return a net positive contribution, beating 50% in every case.

Now we have established the Macro data set as the superior, we can compare it to the Mix data set. Excitingly, this data set outperforms the Macro data set across the board - on average by **1.83 %** for Model 1 and **1.99 %** for Model 2.

Another feature to be noticed from both Table 1 and Table 2 is that the Mix data set not only exhibited a greater predictive accuracy, it also managed to do so in less iterations during the gradient boosting. This 'faster descent' can be explained by several features of the data set. First of all in comparison to the

A final important conclusion to be drawn from these results, is that the errors associated with the predictions of the Macro and the Mix subsets are very similar. [Sum up the number that are smaller/larger].

Now turning our attention to the last two columns of the tables, we can compare the critical model parameter: **mstop**. Model 1 has a finer granularity and so makes the gradient descent at a slower pace. This means higher computational cost due to more iterations being required to reach the local minima during the gradient descent. This is reflected in the column **mstop avg**, with many cases averaging an mstop value of more than 500.

Comparing this to the average mstop values obtained when using the same data, but with a larger shrinkage value (i.e. a higher granularity), we see that using a shrinkage value ten-fold smaller produces values almost linearly scalable, the mstp values being approximately ten times smaller.