

Effectivess comparison report

Raphael Rodrigues Campos

January 17, 2016

Eu implementei o BROOF usando Extremely Randomized Trees no lugar da RF, gerando o algoritmo que chamei de BERT (Boosted Extremely Randomized Trees).

A própria ERT se sai melhor em alguns datasets do que a RF. Portanto, era de se esperar que a BERT se saísse um pouco melhor que o BROOF, como pode-se verificar no arquivo anexo.

O arquivo anexo possui uma tabela comparando todos os métodos rodados até agora.

Além da implementação do BERT, eu também implementei método de ensemble “Stacked Generalization” descrito em [1] David H. Wolpert, “Stacked Generalization”, Neural Networks, 5, 241–259, 1992.

O método comb1 na tabela é o stacking de 2 níveis para combinação dos métodos LazyNN_RF e BROOF. No nível do zero do stacking foram utilizados os classificadores LazyNN_RF e BROOF para gerar o conjunto de treino do nível 1. No nível 1 foi utilizado uma RF com 200 árvores.

Os resultados apresentados são promissores. Sobretudo quando se trata de métrica microf1, onde tivemos mais ganhos significativos.

Resultados

% latex table generated in R 3.2.4 by xtable 1.8-0 package % Tue Apr 26 11:38:20 2016

% latex table generated in R 3.2.4 by xtable 1.8-0 package % Tue Apr 26 11:38:43 2016

Legenda para os métodos:

- BERT: Boosted Extremely Randomized Trees
- LXT: Lazy Extremely Randomized Trees
- RF: Random Forest com 200 árvores
- RF1000: Random Forest com 1000 árvores
- XT: Extremely Randomized Trees com 200 árvores
- XT1000: Extremely Randomized Trees com 1000 árvores
- COMB1: Stacking (Lazy + BROOF)
- COMB2: Stacking (LXT + BERT)
- COMB3: Stacking (Lazy + BROOF + LXT + BERT)
- COMBSOTA: Stacking (KNN + RF + SVM + NB)

V1	V2	20NG	4UNI	ACM	REUTERS90	MEDLINE
SVM-L2	microF1	90.06 \pm 0.43	83.48 \pm 1.08	75.4 \pm 0.66	68.19 \pm 1.15	0 \pm 0
	macroF1	89.93 \pm 0.43	73.39 \pm 2.17	63.84 \pm 0.55	31.95 \pm 2.59	0 \pm 0
BERT	microF1	88.93 \pm 0.39	84.61 \pm 0.98	74.8 \pm 0.59	67.33 \pm 0.72	0 \pm 0
	macroF1	88.59 \pm 0.5	73.61 \pm 1.85	62.1 \pm 0.99	29.24 \pm 1.4	0 \pm 0
SVM-L1	microF1	89.8 \pm 0.4	78.23 \pm 1.49	75.31 \pm 0.74	68.25 \pm 1.2	0 \pm 0
	macroF1	89.59 \pm 0.43	67.47 \pm 3.01	62.33 \pm 1.76	31.37 \pm 2.22	0 \pm 0
SVM-MAX	microF1	88.35 \pm 0.37	81.36 \pm 1.01	73.82 \pm 0.78	67.6 \pm 1.1	0 \pm 0
	macroF1	88.3 \pm 0.38	68.01 \pm 2.39	62.55 \pm 1.53	31.73 \pm 3.13	0 \pm 0
BROOF	microF1	87.96 \pm 0.24	84.41 \pm 1.07	73.35 \pm 0.79	66.79 \pm 0.97	0 \pm 0
	macroF1	87.44 \pm 0.28	73.23 \pm 1.1	60.76 \pm 0.8	28.48 \pm 2.17	0 \pm 0
KNN	microF1	87.53 \pm 0.69	75.63 \pm 0.94	70.99 \pm 0.96	68.07 \pm 1.07	0 \pm 0
	macroF1	87.22 \pm 0.66	60.34 \pm 1.36	55.85 \pm 0.97	29.93 \pm 2.48	0 \pm 0
SVM-NONE	microF1	83.47 \pm 0.46	80.55 \pm 0.72	71.34 \pm 1.01	66.6 \pm 1.06	0 \pm 0
	macroF1	83.37 \pm 0.42	71.04 \pm 2.06	61.08 \pm 0.67	31.68 \pm 3.32	0 \pm 0
NB	microF1	88.99 \pm 0.54	62.63 \pm 1.7	73.54 \pm 0.71	65.32 \pm 1.13	82.92 \pm 0.14
	macroF1	88.68 \pm 0.55	51.38 \pm 3.19	58.03 \pm 0.85	27.86 \pm 0.79	63.8 \pm 0.43
RF	microF1	83.64 \pm 0.29	81.52 \pm 1	71.05 \pm 0.31	63.92 \pm 0.81	81.54 \pm 0.08
	macroF1	83.08 \pm 0.35	65.44 \pm 1.91	56.56 \pm 0.45	24.36 \pm 1.98	67.4 \pm 0.36
XT	microF1	85.94 \pm 0.23	81.66 \pm 1.03	71.94 \pm 0.66	64.33 \pm 0.86	81.48 \pm 0.11
	macroF1	85.57 \pm 0.22	65.44 \pm 2.41	57.4 \pm 1.13	24.47 \pm 2.22	67.34 \pm 0.29
LAZY	microF1	87.96 \pm 0.37	82.34 \pm 0.61	74.02 \pm 0.79	66.3 \pm 1.07	0 \pm 0
	macroF1	87.39 \pm 0.37	68.33 \pm 1.6	59.46 \pm 1.35	26.61 \pm 2.12	0 \pm 0
LXT	microF1	88.39 \pm 0.51	81.24 \pm 0.71	69.63 \pm 0.91	65.92 \pm 0.82	0 \pm 0
	macroF1	88.05 \pm 0.44	66.89 \pm 1.23	57.33 \pm 1.48	26.71 \pm 2.53	0 \pm 0

Table 1: Comparação entre todos os métodos

V1	V2	20NG	4UNI	ACM	REUTERS90
COMBALL	microF1	91.67 \pm 0.44	86.74 \pm 1.17	78.46 \pm 0.72	80.02 \pm 1.24
	macroF1	91.43 \pm 0.42	79.45 \pm 2.23	63.72 \pm 1.01	37.84 \pm 3.14
COMB3	microF1	90.63 \pm 0.57	86.79 \pm 0.86	77.34 \pm 0.6	79 \pm 1.14
	macroF1	90.4 \pm 0.57	79.63 \pm 1.91	62.91 \pm 0.92	33.93 \pm 2.97
COMB1	microF1	89.32 \pm 0.42	86.52 \pm 1.18	76.74 \pm 0.73	77.22 \pm 1.14
	macroF1	89.01 \pm 0.44	78.66 \pm 1.9	62.2 \pm 1.01	31.71 \pm 2.7
COMB2	microF1	90.2 \pm 0.51	86.54 \pm 1.06	76.88 \pm 0.55	78.25 \pm 1.17
	macroF1	89.95 \pm 0.52	79.41 \pm 1.63	62.66 \pm 0.81	32.86 \pm 2.23
COMBSOTA	microF1	90.65 \pm 0.4	83.79 \pm 1.3	77.9 \pm 0.73	74.41 \pm 1.21
	macroF1	90.41 \pm 0.4	74.19 \pm 2.13	63.15 \pm 0.76	28.18 \pm 1.58
SVM-L2	microF1	90.06 \pm 0.43	83.48 \pm 1.08	75.4 \pm 0.66	68.19 \pm 1.15
	macroF1	89.93 \pm 0.43	73.39 \pm 2.17	63.84 \pm 0.55	31.95 \pm 2.59
BERT	microF1	88.93 \pm 0.39	84.61 \pm 0.98	74.8 \pm 0.59	67.33 \pm 0.72
	macroF1	88.59 \pm 0.5	73.61 \pm 1.85	62.1 \pm 0.99	29.24 \pm 1.4

Table 2: Comparação entre todos os métodos