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 implementacáo 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 combinacáo 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.

```
## [1] "~/Documents/Master Degree/Master Project/Implementation/LazyNN_RF/release/results/results_grid/
## [1] "~/Documents/Master Degree/Master Project/Implementation/LazyNN_RF/release/results/results_grid/
## [1] "~/Documents/Master Degree/Master Project/Implementation/LazyNN_RF/release/results/results_grid/
## [1] "~/Documents/Master Degree/Master Project/Implementation/LazyNN_RF/release/results/results_grid/
## [1] "~/Documents/Master Degree/Master Project/Implementation/LazyNN_RF/release/results/results_grid/
## [1] "~/Documents/Master Degree/Master Project/Implementation/LazyNN_RF/release/results/results_grid/
## [1] "~/Documents/Master Degree/Master Project/Implementation/LazyNN_RF/release/results/results_grid/
## [1] "~/Documents/Master Degree/Master Project/Implementation/LazyNN_RF/release/results/results_grid/
## [1] "~/Documents/Master Degree/Master Project/Implementation/LazyNN_RF/release/results/results_grid/
## [1] "~/Documents/Master Degree/Master Project/Implementation/LazyNN_RF/release/results/results_grid/
## [1] "~/Documents/Master Degree/Master Project/Implementation/LazyNN_RF/release/results/results_grid/
## [1] "~/Documents/Master Degree/Master Project/Implementation/LazyNN_RF/release/results/results_grid/
## [1] "~/Documents/Master Degree/Master Project/Implementation/LazyNN_RF/release/results/results_grid/
## [1] "~/Documents/Master Degree/Master Project/Implementation/LazyNN_RF/release/results/results_grid/
## [1] "~/Documents/Master Degree/Master Project/Implementation/LazyNN_RF/release/results/results_grid/
## [1] "~/Documents/Master Degree/Master Project/Implementation/LazyNN_RF/release/results/results_grid/
## [1] "~/Documents/Master Degree/Master Project/Implementation/LazyNN_RF/release/results/results_grid/
## [1] "~/Documents/Master Degree/Master Project/Implementation/LazyNN_RF/release/results/results_grid/
## [1] "~/Documents/Master Degree/Master Project/Implementation/LazyNN_RF/release/results/results_grid/
## [1] "~/Documents/Master Degree/Master Project/Implementation/LazyNN_RF/release/results/results_grid/
## [1] "~/Documents/Master Degree/Master Project/Implementation/LazyNN_RF/release/results/results_grid/
## [1] "~/Documents/Master Degree/Master Project/Implementation/LazyNN_RF/release/results/results_grid/
## [1] "~/Documents/Master Degree/Master Project/Implementation/LazyNN_RF/release/results/results_grid/
## [1] "~/Documents/Master Degree/Master Project/Implementation/LazyNN_RF/release/results/results_grid/
## [1] "~/Documents/Master Degree/Master Project/Implementation/LazyNN_RF/release/results/results_grid/
## [1] "~/Documents/Master Degree/Master Project/Implementation/LazyNN_RF/release/results/results_grid/
## [1] "~/Documents/Master Degree/Master Project/Implementation/LazyNN_RF/release/results/results_grid/
## [1] "~/Documents/Master Degree/Master Project/Implementation/LazyNN_RF/release/results/results_grid/
## [1] "~/Documents/Master Degree/Master Project/Implementation/LazyNN_RF/release/results/results_grid/
## [1] "~/Documents/Master Degree/Master Project/Implementation/LazyNN_RF/release/results/results_grid/
## [1] "~/Documents/Master Degree/Master Project/Implementation/LazyNN_RF/release/results/results_grid/
## [1] "~/Documents/Master Degree/Master Project/Implementation/LazyNN_RF/release/results/results_grid/
## [1] "~/Documents/Master Degree/Master Project/Implementation/LazyNN_RF/release/results/results_grid/
## [1] "~/Documents/Master Degree/Master Project/Implementation/LazyNN_RF/release/results/results_grid/
## [1] "~/Documents/Master Degree/Master Project/Implementation/LazyNN_RF/release/results/results_grid/
```

Resultados

% latex table generated in R 3.2.4 by xtable 1.8-0 package % Sun Apr 10 16:32:36 2016

V1	V2	20NG	4UNI	ACM	REUTERS90
BERT	microF1	82.12 ± 0.51	**84.61 ± 0.98**	**74.8 ± 0.59**	**67.33 ± 0.72**
	macroF1	81.86 ± 0.54	**73.61 ± 1.85**	**62.1 ± 0.99**	**29.24 ± 1.4**
BROOF	microF1	86.77 ± 0.39	**84.41 ± 1.07**	**73.35 ± 0.79**	**66.79 ± 0.97**
	macroF1	86.25 ± 0.49	**73.23 ± 1.1**	**60.76 ± 0.8**	**28.48 ± 2.17**
KNN	microF1	87.53 ± 0.69	75.63 ± 0.94	70.99 ± 0.96	**68.07 ± 1.07**
	macroF1	87.22 ± 0.66	60.34 ± 1.36	55.85 ± 0.97	**29.93 ± 2.48**
LAZY	microF1	87.96 ± 0.37	82.34 ± 0.61	**74.02 ± 0.79**	**66.3 ± 1.07**
	macroF1	87.39 ± 0.37	68.33 ± 1.6	59.46 ± 1.35	26.61 ± 2.12
LXT	microF1	**88.39 ± 0.51**	81.24 ± 0.71	69.63 ± 0.91	65.92 ± 0.82
	macroF1	**88.05 ± 0.44**	66.89 ± 1.23	57.33 ± 1.48	26.71 ± 2.53
NB	microF1	**88.99 ± 0.54**	62.63 ± 1.7	**73.54 ± 0.71**	65.32 ± 1.13
	macroF1	**88.68 ± 0.55**	51.38 ± 3.19	58.03 ± 0.85	**27.86 ± 0.79**
RF	microF1	83.64 ± 0.29	81.52 ± 1	71.05 ± 0.31	63.92 ± 0.81
	macroF1	83.08 ± 0.35	65.44 ± 1.91	56.56 ± 0.45	24.36 ± 1.98
SVM	microF1	**88.35 ± 0.37**	81.36 ± 1.01	**73.82 ± 0.78**	**67.6 ± 1.1**
	macroF1	**88.3 ± 0.38**	68.01 ± 2.39	**62.55 ± 1.53**	**31.73 ± 3.13**
XT	microF1	0 ± 0	81.66 ± 1.03	71.94 ± 0.66	64.33 ± 0.86
	macroF1	0 ± 0	65.44 ± 2.41	57.4 ± 1.13	24.47 ± 2.22

Table 1: Comparaćão entre todos os métodos

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)