

1 Algorithm

Algorithm 1: SVM Baseline

Input: Files (Files train and test), tp (Tunned params), cv (k-Folds Cross Validation), metric (Metric), df (Default paramns LBD)

```

1 forall fold ∈ Files do
2   x_train, y_train, x_test, y_test = fold;
3   model = GridSearch(svn.SVC(df), tp, cv, metric);
4   model.fit(x_train, y_train);
5   y_pred = model.predict(x_test);
6   fl_macro.add(fl_score(y_test, y_pred, "fl_macro"));
7   fl_micro.add(fl_score(y_test, y_pred, "fl_micro"));
8 return mean(fl_macro), sd(fl_macro), mean(fl_micro), sd(fl_micro);

```

O algoritmo SVM tem complexidade $O(n_samples^3 \times n_features)^1$. Utilizando o *TFIDF* do dataset foi possível obter a informação de samples (linhas) e features (colunas), a Tabela 1 apresenta o impacto da entrada no algoritmo SVM. Essas informações foram obtidas do fold 0, neste caso será necessário calcular o valor para os demais folds (4).

Dataset	Samples	Features	SVM
Stanford Tweets	286	1333	31183743448
20NG	15071	98230	$3,362562409 \times 10^{17}$
ACM	19914	48919	$3,86799071 \times 10^{17}$

Table 1: Datasets.

2 Experiments

Methods	Macro F1	Micro F1	Time (s)
Metalazy artigo	83.81 (7.52)	83.86 (7.51)	
SVM Kernel linear	79.01 (5.14)	79.09 (5.12)	1.17
SVM Kernel rbf	79.84 (5.37)	79.94 (5.38)	1.31
Metalazy com SVM linear	79.59 (2.76)	79.67 (2.79)	65.19
Metalazy com SVM rbf	80.10 (2.77)	80.23 (2.79)	65.91

Table 2: Results for dataset **Stanford Tweets**.

¹<https://scikit-learn.org/stable/modules/svm.html#complexity>

Methods	Macro F1	Micro F1	Time (s)
Metalazy artigo	90.49 (0.46)	90.75 (0.44)	
SVM Kernel linear	80.06 (0.80)	79.85 (0.85)	2597.41
SVM Kernel rbf	84.94 (0.65)	84.76 (0.54)	3854.49
Metalazy com SVM linear	()	()	
Metalazy com SVM rbf	86.60 (0.64)	86.90 (0.65)	21424

Table 3: Results for dataset **20NG**.

Methods	Macro F1	Micro F1	Time (s)
Metalazy artigo	36.46 (1.01)	67.68 (0.21)	
SVM Kernel linear	31.05 (1.88)	66.24 (0.64)	1704.74
SVM Kernel rbf	31.19 (1.93)	66.51 (0.61)	2244.36
Metalazy com SVM linear	()	()	
Metalazy com SVM rbf	34.78 (2.73)	76.67 (0.91)	107886

Table 4: Results for dataset **REUT**.

Methods	Macro F1	Micro F1	Time (s)
Metalazy artigo	67.66 (2.03)	77.50 (0.67)	
SVM Kernel linear	63.59 (1.20)	73.66 (0.62)	1888.14
SVM Kernel rbf	64.37 (1.35)	74.94 (0.54)	2700.52
Metalazy com SVM linear	()	()	
Metalazy com SVM rbf	61.31 (1.26)	75.74 (0.59)	47233

Table 5: Results for dataset **ACM**.