CLASSIFYING THE GENRE OF TYPICAL BRAZILIAN MUSICS BY LYRICS

Patrick Guimarães[1]
Lukas Iepsen[2]

Abstract: We use many machine learning and deep learning models for the task of genre classification using lyrics. Utilising our own dataset with lyrics of typical genres in Brazil divided in 7 classes, we apply some models used in machine learning and deep learning classification tasks. We explore the performance of usual models for text classification using a input in portuguese language. We also compare the use of RNN and classic machine learning approaches for text classification, exploring the most used methods in the field.

1 INTRODUCTION

Problems that involves text classification are Natural Language Processing (NLP) problems. In the problem presented by this work the goal is to classify the music genre by the lyric, a problem that was also explored by other related works (TSAPTSINOS, 2017) (MAYER; NEUMAYER; RAUBER, 2008). The traditional approaches for text classification problem use n-gram, SVM or Machine Learning models. With the increased computational power, Deep Learning techniques can be used in this task too (YOUNG et al., 2018).

2 DATASET

The music lyrics was acquired from the website https://letras.mus.br, where we collected around of 6000 lyrics, divided into 7 Brazilian typical music genres. The Dataset distribution can be seen on figure 1.

3 MODEL CONFIGURATION

The best results were achieved using a lstm word-level. The model uses a maximum of 200 words per segment and a maximum of 100 segments. The LSTM applies a sigmoid activation function. We also used dropout and gradient clipping. We dropout at one layer with probability p = 0.5 and gradients are clipped at a maximum norm of 1. For the loss and optimizer, we used, categorical cross-entropy and RMSprop with a learning rate of 0.01, respectively. In the output, a softmax function is used. We ran the model for 10 epochs for better visualization, but after 4th epoch the validation accuracy tends to decrease. We applied different batch sizes in the experiments, we opted for a mini-batch of 16 as it shows the best results. The dataset was divided by 70% for training, 10% for validation and 20% for testing.

[1] UFPEL, pgdvargas@inf.ufpel.edu.br
[2] UFPEL, lbiepsen@inf.ufpel.edu.br

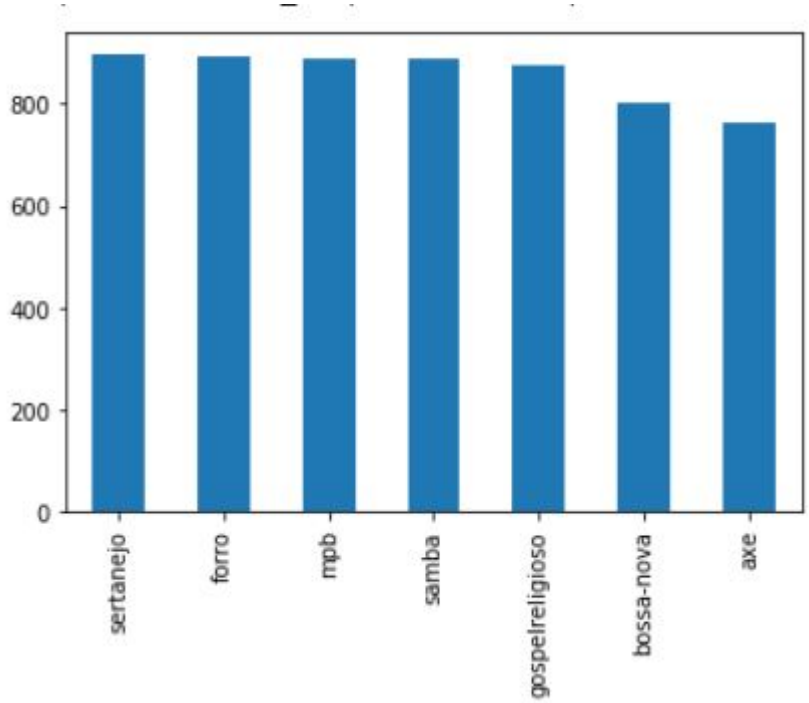

Figure 1 – Dataset genres distribution.

4 RESULTS

We applied 6 different models to the problem in our experiments, then being Long Short-Term Memory (LSTM), FastText, XGBoost, Random Forest, Decision Tree and Multilayer Perceptron (MLP). In Table 1 the accuracy of said models is presented, as one can see the best result is achieved by the LSTM with a 50% accuracy on the testing dataset.

Table 1 – Accuracy of Portuguese music genre classification.

Model	**Accuracy**
LSTM	50%
FastText	49%
XGBoost	48%
Random Forest	45%
Decision Tree	38%
MLP	15%

Figure 2 shows the accuracy and loss in the LSTM model. The LSTM model, which takes into account word order and tries to implement a memory of these words tends to over fit in the training dataset after the 5th epoch.

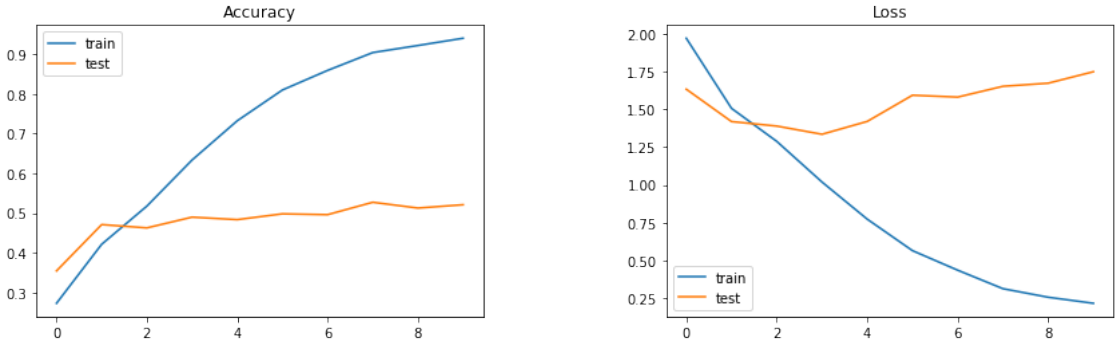

Figure 2 – Accuracy and Loss for the LSTM model on testing dataset

The confusion matrix for the testing dataset can be seen on figure 3. We also present the most frequent words on two of the lyrics styles from our dataset on figure 4. We can see that even though the two styles are completed different musically, they share many words in their lyrics. So is no surprise that the model is confused as to their classification.

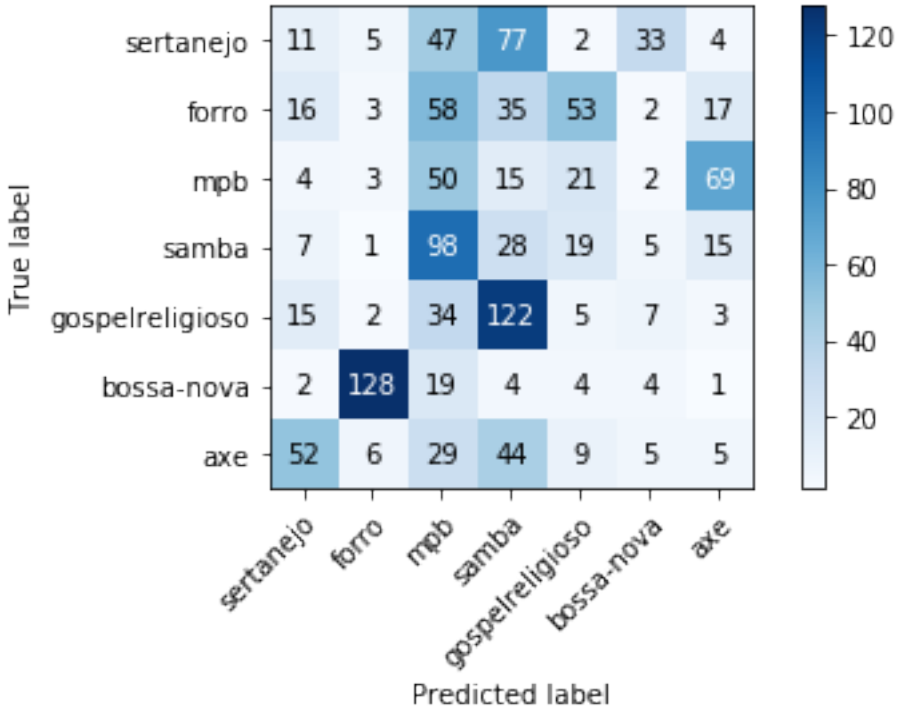

Figure 3 – Confusion Matrix for the testing.

word	number_of_times
amor	1220
coração	560
quero	443
gente	434
Você	407
Me	405
vida	388
tudo	376
dia	333
tá	331

word	number_of_times
amor	851
vida	351
ser	295
tão	246
sei	245
bem	243
coração	243
gente	239
dia	236
Você	234

Figure 4 – Most frequent word from forró and bossa-nova.

5 CONCLUSION

Genre classification is presented as a hard task. Due to some of the styles sharing many words it is unclear if whether a person would be able to distinguish between the lyrics of these genres. To produce a better classifier we must take into account more than just the words. A combination of audio and lyrics could be applied for bigger accuracy. Our future work is to increase our sample by getting a more lyrics and to explore more of the similarity of the classes at the moment of classification.

BIBLIOGRAPHY

MAYER, R.; NEUMAYER, R.; RAUBER, A. Combination of audio and lyrics features for genre classification in digital audio collections. In: ACM. **Proceedings of the 16th ACM international conference on Multimedia**. [S.l.], 2008. p. 159–168.

TSAPTSINOS, A. Lyrics-based music genre classification using a hierarchical attention network. **arXiv preprint arXiv:1707.04678**, 2017.

YOUNG, T. et al. Recent trends in deep learning based natural language processing. **ieee Computational intelligenCe magazine**, IEEE, v. 13, n. 3, p. 55–75, 2018.