

Google Places for Sentiment Analysis

1. About

This repository contains all the resources used in the tests of the paper “Building a corpus from supermarket reviews in portuguese for document-level Sentiment Analysis”. The framework for data collection is yet to be finished and published, and will be public available here after the process been done.

In this repository are available the **linguistic resources** used in the paper and the **source codes** of the tests conducted.

For the tests, we use the Python language, as well as the sklearn package to run the machine learning algorithms. We also used spacy for NLP algorithms used in the preprocessing strategy.

As described in the paper, we test for Sentiment Analysis in portuguese:

- two lexical approaches;
 - using LIWC lexicon;
 - using a domain specif lexicon built specific for this paper;
- three different machine learning algorithms;
 - Naive Bayes;
 - SVM
 - Logistic Regression;
- three different functions in the Deep Learning Multi-layer Perceptron classifier;

Also, we test two different preprocessing approaches for the machine learning and deep learning algorithms (see the paper).

2. Instalation

In order to run the tests, you must have **python** installed in your machine. A bash script was also created to easily assist you in the configuration of the Python packages using pip. So, with you are running this on a Ubuntu based distribution of Linux, run the following commands:

```
sudo apt update
sudo apt install python3-pip
chmod +x configure.sh
./configure.sh
```

3. Citation

Will be available soon.