POLITECNICO DI TORINO

Dipartimento di Automatica e Informatica

PhD in Computer and Control Engineering

XXX cycle

Supervisor

Prof. Maurizio MORISIO

Text-based Sentiment Analysis and Music Emotion Recognition

PhD Candidate: *Erion ÇANO*

1. Introduction

The Internet of today is full of crowd opinions about products, songs, movies etc. Their influence on our daily decision making is so high that:

- 40M reviews are added every day on Yelp
- 92 % of people read online reviews before watching, listening or buying an item
- 1-3 negative reviews about an item convince people not to watch listen or buy it
- 99 % of unhappy customers leave a review

Sentiment Analysis of these text reviews is essential for advertising companies.

2. Objectives

The main objective of our work is to boost up sentiment category recognition in emotional texts like song lyrics, movie reviews, item reviews etc. Specifically, we address the following issues:

- Creating large and public datasets of polarized lyrics
- Investigating the role and quality of word embeddings on sentiment analysis
- Designing a deep neural network architecture to boost up classification accuracy of lyrics and reviews.

3. MoodyLyrics Datasets

ML, ML4Q and MLPN are 3 big, polarized and public datasets of song lyrics we constructed. We crowdsourced Last.fm emotion user tags the most frequent of which are shown in Figure 1.

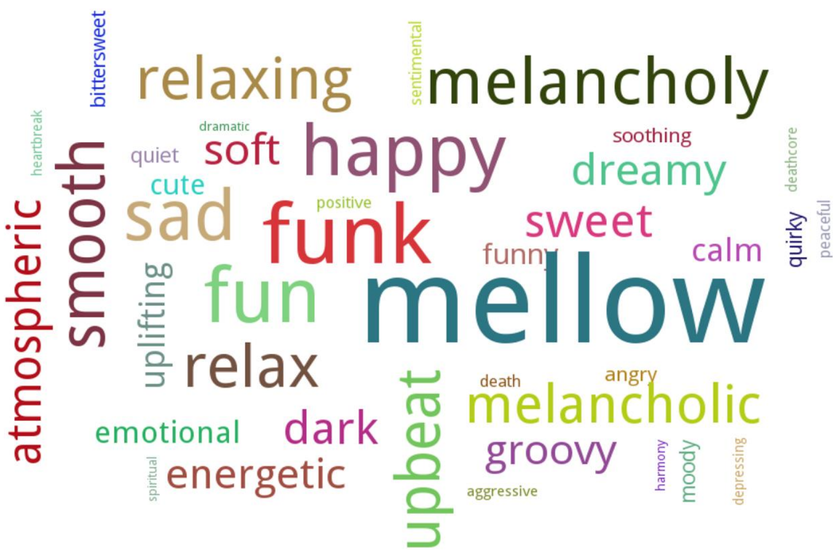

Figure 1: Most frequent Last.fm emotion tags

As emotion representation model we use that of Russell which categorizes songs as *happy, angry, sad* or *relaxed.* You can find the 3 datasets at http://softeng.polito.it/erion/

4. NgramCNN Architecture

To boost up classification accuracy of positive and negative lyrics or reviews we constructed NgramCNN, a deep neural network architecture depicted in Figure 2. Our experimental work in [2] about the quality of word embeddings revealed that corpus size has an important influence on word representation. For this reason we use weights of GoogleNews corpus in the embedding layer. The rest of layers are convolutions with several kernels of size 1, 2 or 3.

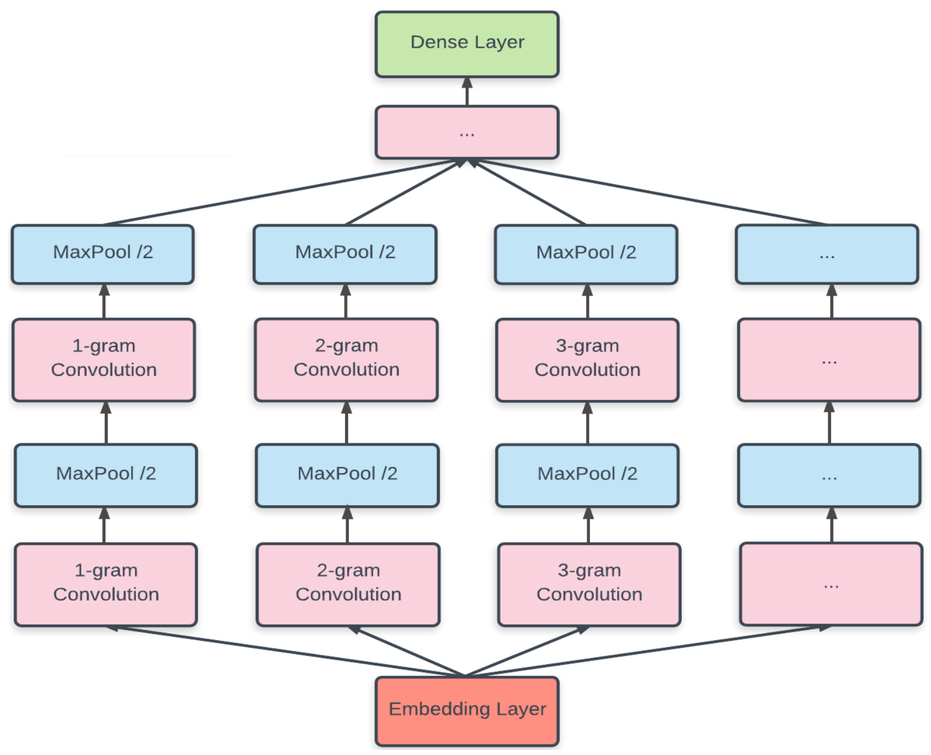

Figure 2: Schematic view of NgramCNN

Each convolution layer is followed by a max pooling layer with pool size 2. At the end, a dense layer serves as classifier. The fundamental idea of this architecture is the use of generic deep features with a simple classifier.

Alg\Data	MLPN	MovieRev
Opt. LR	73.1	89.4
Opt. SCV	72.7	88.5
LSTM-CNN	70.9	84.9
SingleCNN	72.4	89.9
NgramCNN	**75.6**	**91.2**

Table 1: Evaluation of NgramCNN

5. Evaluation

We used MoodyLyrcsPN and movie review dataset to evaluate NgramCNN architecture. The results are presented in Table 1. As you can see, NgramCNN performs at least 2.5 % better than the similar algorithms on song lyrics. On movie reviews it gives about 0.5 % higher accuracy than current state of the art.

6. Conclusions

- Crowdsourcing user tags from online portals is a viable and effective method for constructing experimental datasets.
- Word embeddings provide a dense representation of words that is highly convenient for sentiment analysis tasks.
- Deeper neural networks with convolution and pooling layers provide better sentiment analysis accuracy than shallow models or simple off-the-shelf linear classifiers.

7. References

1. Çano E.; Morisio M.: Music Mood Dataset Creation Based on Last.fm Tags. In: Fourth International Conference on Artificial Intelligence and Applications, pp. 15-26, Vol. 7, Vienna, Austria
2. Çano E.; Morisio M.: Quality of Word Embeddings on Sentiment Analysis Tasks. In: NLDB 2017, Springer, pp. 332-338, Belgium.