

Detecting Emotional Scene of Videos from Subtitles

Guide: Prof. Amitabha Mukherjee

Utsav Sinha, 12775
Rajat Kumar Panda, 12545

CS365 Project Proposal
16th March 2015

Motivation

With a plethora of multimedia content being produced and widely available, it becomes more confusing and tiresome for a user to find the movie of his choice. After seeing the movie, one might want to see the highlights and the emotional scenes afterwards also. But the task becomes difficult since one can almost never recollect where in the movie were the scenes exactly and what the scenes actually were.

Taking motivation from this need, we propose to detect and report the emotional scenes of a video. This can be further expanded to automatically detect profanity in a video.

Introduction

Multimedia expresses emotional content in various ways the facial expression, dialogue, the way of speaking, the context, background scene, and music along with many others. In this project, we propose to use the dialogues of a video obtained from its subtitles to accomplish the task of detecting emotional scenes using Natural Language Processing (NLP) techniques.

The dialogues of a movie form an ideal source from which the sentiments of a video can be analysed. The words being spoken along with their semantic context determines would help in annotating the dialogue with its appropriate emotion.

Proposed Approach

We will use bag of words representations - WordNet [PYKJ11] and SentiWordNet [BES10] to get the synonym sets (synset) of words. This will form the basis of the clustering of emotion words. These clusters will then help to make the emotion dependency graph which will have various emotions such as hate, joy, fright etc as nodes and edges will be formed between nodes belonging to the same synset. This graph will result in creating emotion vectors for each emotion with a vector containing measures of degree of major sentiments like happiness, anger etc calculated using their conceptual distance from these major emotions.

We will use Porter Stemmer algorithm [Ste] to stem words occurring in various tenses and forms.

Training on pre-labelled dialogues will be done. This will involve learning the appropriate composition function that maps emotion vectors of phrases (constructed from emotion vector of individual words recursively) to the pre-labelled emotion of the sentence using Recursive Neural Networks [SPW⁺13]. With the increase in training data size, the mapping function will improve which would be put to test on test data. k-fold cross validation will be done to reduce error rate.

Another approach of changing the values of emotion vector of each emotion based on unsupervised learning from training corpus before feeding this input to neural net can also be used. This will remove the initial bias while statically constructing the emotion vector instead of changing values from learning.

We can then compute the emotion of a scene containing a group of sentences by using the sum (a function) of emotions of individual sentences.

Creating a labelled (with emotions) dataset from movie subtitles will also form a major goal of the project.

Dataset

The subtitles of movies can be obtained from [pod]. But subtitles with labelled emotions are not available which would form the ground truth against which our proposed model could be compared. This would be resolved by hand tagging the dialogues of some test data which would form the standard.

References

- [BES10] Stefano Baccianella, Andrea Esuli, and Fabrizio Sebastiani. Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In *LREC*, volume 10, pages 2200–2204, 2010.

- [pod] podnapisi. Subtitles datasource - <http://www.podnapisi.net>.
- [PYKJ11] Seung-Bo Park, Eunsoon Yoo, Hyunsik Kim, and Geun-Sik Jo. Automatic emotion annotation of movie dialogue using wordnet. In *Intelligent Information and Database Systems*, pages 130–139. Springer, 2011.
- [SPW⁺13] Richard Socher, Alex Perelygin, Jean Y Wu, Jason Chuang, Christopher D Manning, Andrew Y Ng, and Christopher Potts. Recursive deep models for semantic compositionality over a sentiment treebank. In *Proceedings of the conference on empirical methods in natural language processing (EMNLP)*, volume 1631, page 1642. Citeseer, 2013.
- [Ste] Porter Stemmer. Porter stemmer algorithm - <http://tartarus.org/martin/porterstemmer>.