MULTIMODAL SENTIMENT ANALYSIS ON SONGS USING ENSEMBLE CLASSIFIERS

Draft of March 24, 2015 at 23 : 54

BY

ESTEBAN GOMEZ

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Bachelor of Science in Electrical and Computer Engineering
in the Undergraduate College of Engineering of the
University of Illinois at Urbana-Champaign, 2015

Urbana, Illinois

Adviser:

Minh N. Do

ABSTRACT

We consider the problem of performing sentiment analysis on songs by combining audio and lyrics in a large and varied dataset, using the Millions Song Dataset for audio features and the MusicXMatch dataset for lyric information.

The algorithms presented on this paper utilize ensemble classifiers as a method of fusing data vectors from different feature spaces. We find that a multimodal classification outperforms using only audio or only lyrics. This paper argues that utilizing signals from different spaces can account for inter-class inconsistencies and leverages class-specific performance. The experimental results show that multimodal classification not only improves overall classification, but is also more consistent across different classes.

Keywords: Signal Processing; Sentiment Analysis; Machine Learning; Feature Fusion; Multimodal Classification ; I NEED TO DISCUSS THIS PART

TABLE OF CONTENTS

1

INTRODUCTION

In recent years, music based services have been trying to make their applications more user-centered. One way to make sure that the user experience is foremost, is to provide services that match the user's current emotional state. This can be implemented by understanding the emotional content of specific songs and playlists.

Sentiment analysis is the portion of Music Informational Retrieval (MIR) where an algorithm recognizes the main emotions that that a song evokes. Emotions are subjective, so classifying media into distinct groups is a challenging problem. Most human subjects agree in broad strokes on emotional classifications. However it is not uncommon to find media inputs where there is no consensus, leading to inconsistencies in class groupings. As a result, sentiment analysis is challenging to incorporate into real-world applications because an algorithm's recognition accuracy might differ between classes.

The motivation behind this research is to find a method that accounts for these inter-class inconsistencies across a large dataset. During initial testing it became apparent that certain classifiers perform well on specific sentiments and fail to learn features that represent others. This research seeks to account for this difference by combining the classifiers to output classifications in a more consistent manner, thus improving the reliability of the overall system.

The dataset used for this paper is called the Million Song Dataset, and was compiled by Labrosa [I.E]. This dataset contains a million different songs represented by their pitch, loudness and timbre. The songs are also accompanied by plenty of metadata such as tags. Sentiment classification was obtained from these tags. If a sentiment was used to described that song, it is assumed that the song conveyed that sentiment. The lyric information was obtained from the musicXmatch dataset [II.F], which provides lyric information out of order in a bag-of-words format. Since there was no way to obtain semantic information from an unordered bag-of-words representation, this research did not focus on the impact of semantics on sentiment classification.

This document starts with a brief overview of previous work performed on this topic, followed by a description of the algorithms developed for the experiments, then a section containing results and corresponding analysis and closes with conclusions and suggestions for future work.

2

PREVIOUS WORK

There has been considerable amount of work done in the field of multimodal sentiment analysis. This section will briefly cover a portion of the relevant research that was considered during the development of the presented methodology. The relevant topics that were research for this paper were: Audio Sentiment Analysis, Text Sentiment Analysis and Multimodal Classification.

2.1 Audio Sentiment Analysis

The study of the relationship between emotional content and audio signals is a very mature field. Researchers have expanded on the success found in the speech recognition community while using Mel-Frequency Cepstral Components (MFCC) to explore their uses in music modeling [I.A]. MFCCs are currently a staple in audio processing and are commonly used in MIR applications such as genre classification [I.B], since it is a quantifiable method for comparing the timbral texture of songs. Timbre has been successfully used to classify the emotional content of songs [I.C]. It has also been used to generate songs that evoke particular emotions [I.D]. These vectors have been commonly classified using Support Vector Machines (SVM) and Naive Bayes classifiers. The dataset used for these experiments is the Million Song Dataset (MSDS) [I.E] since metadata provided is intended for MIR research. The audio features used were the MFCC-like timbre vectors provided by EchoNest in the MSDS.

2.2 Text Sentiment Analysis

Similarly the study of the relationship between text and emotional content is quite developed. From predicting Yelp ratings based on the sentiment expressed on the review [II.A] to extracting the emotional progression of major literary pieces [I.D]. There are many methods to represent and extract emotional information from texts. The Yelp experiment uses a statistical word vectors to capture word semantics and emotions as a probability. Other researchers have represented textual information in a bag-of-features framework and classified them using Naive Bayes, SVMs and Maximum Entropy classifiers to recognize positive or negative valances (II.B).

Capturing the semantic nuisances has also been an area of great interest, which is the study of how a given word might have different emotional value depending on its context. This level of analysis requires the creation of complex sentiment vectors that encode how meanings change based on semantics [II.C]. Similarly researchers have improved classification accuracy by preprocessing text [II.D], and use the cleaned data to capture emotional subtleties like the use of negation and modifiers to emotional words [II.E].

Although there is a great body of research on how to obtain rich sentiment vectors from the text, the goal of this paper is demonstrate the added advantage of a multimodal approach. As a result, the features used will be the bag-of-words vector provided by the musicXatch dataset [II.F].

2.3 Multimodal Classification

Multimodal classification is the task of using feature vectors from different spaces, for example text and audio, to reach a single classification. There are two main methods of combining the information from both vector spaces: feature fusion and classifier fusion [III.A].

Feature Fusion is the technique that takes signals from different feature spaces and joins them to train a single multimodal classifier. The standard fusion method is called series fusion, which consists on concatenating the vectors together and training the classifier on the union of both spaces. Several alternatives have been suggested to maintain the same amount of expressibility in the vector while keeping the vector space as small as possible. Instead of concatenating the vectors together, it is possible to join vectors in parallel [III.B] by making vectors from the linear combinations of a real-valued feature with another complex-valued feature. The benefit of the series fusion over the parallel method is that many diverse features can be fused together to obtain more robust data. As seen in the research by Liang et al., genre classification was improved by joining five different vectors all resulting from different preprocessing methods for text and audio vectors [III.C].

Classifier fusions consists on training an array of unimodal classifiers and using some function to consolidate the predictions [III.D]. This method seamlessly fuses features from very different spaces. Caridakis et al. combined facial expressions, body gestures and speech by having a classifier voting system where the class with most votes and higher probability was chosen amongst all the decisions [III.E]. The final decision-making process can be taken a step further by adding an additional classifier that learns from the decisions provided from classifier array [III.F]. The algorithms presented on this paper were largely based on this last approach.

3

MOTIVATION

Multimodal classification has been successful in improving the accuracy of classification [III.G, III.H]. Some of the previous work either ran the experiments on highly homogenous datasets where all the music was in the same language, belonged to the same genre or were carefully classified by a single subject thus eliminating class inconsistencies.

4

EXPERIMENT

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Donec diam dui, posuere non sodales in, consectetur quis magna. Proin dapibus nibh sed turpis sollicitudin, et iaculis massa cursus. Sed interdum metus non arcu interdum hendrerit. Cras eu semper nisi. Etiam hendrerit quam quis ipsum pulvinar, eget commodo augue mattis. Morbi vehicula et ligula in lobortis. Vivamus mattis pharetra tellus eget suscipit. Ut non sem eget diam pharetra efficitur a lacinia quam. Sed ac tempor enim. Nulla facilisi. Ut augue est, sodales eu turpis ut, viverra fringilla sem. Ut sed molestie ex, at rutrum mauris. Donec dignissim dapibus posuere.

Quisque posuere nunc nulla, eget porttitor ligula blandit at. Duis tellus leo, rutrum a massa volutpat, hendrerit euismod nunc. Pellentesque auctor, risus a posuere convallis, velit nulla molestie quam, sit amet viverra sem nulla fermentum nisl. Nulla vel tellus interdum, suscipit diam ac, bibendum erat. Sed a risus ac massa ullamcorper placerat ut ut justo. Nullam faucibus id mi vel vehicula. Nam imperdiet leo ex, pellentesque blandit tellus egestas sed. Nunc pellentesque auctor lacinia. Proin gravida tellus quam. Aliquam erat volutpat. Phasellus hendrerit erat non erat cursus egestas. Ut nec sem quis leo viverra ultricies at nec dolor. Vestibulum ac dictum urna.

Nulla facilisi. Cras vulputate ante ac nisl elementum gravida. Nam quis pellentesque enim. Suspendisse gravida semper turpis at mattis. Aliquam in condimentum nulla, sed molestie enim. Mauris mattis, tortor ut porttitor malesuada, neque urna cursus dui, id sollicitudin nunc velit ut quam. In in turpis at massa auctor egestas at sed neque. Quisque luctus diam in nibh fringilla, vitae iaculis lorem eleifend.

Fusce efficitur odio sed magna fermentum congue. In hac habitasse platea dictumst. Pellentesque rhoncus, massa a viverra dapibus, nibh orci euismod dui, eget commodo velit nulla sed lorem. Aenean sed enim eu massa pulvinar venenatis quis vitae odio. Mauris lorem purus, fermentum nec quam gravida, maximus aliquam risus. Mauris ornare justo enim, ut pretium enim lacinia non. Morbi quis pulvinar nibh, in tristique diam. Pellentesque auctor nec diam sed imperdiet. Etiam nec maximus magna. Sed pretium justo in nulla efficitur dictum. Donec maximus nisi vitae est vulputate blandit. Maecenas eget nisi ac odio dignissim pretium. Mauris mollis molestie sapien, eget facilisis urna laoreet ac. Nulla non vulputate risus. Integer sed lacinia tortor. Sed fermentum velit sit amet aliquet efficitur.

Donec vehicula arcu odio, eget pellentesque diam mollis nec. Nam mollis enim eu ultricies sodales. In molestie luctus accumsan. Etiam ut tellus sed elit accumsan dignissim auctor sit amet nulla. Sed augue odio, aliquet a massa et, euismod tempor neque. Ut vitae fringilla lectus. Vestibulum ex lacus, molestie sed viverra non, venenatis sed nulla. Curabitur rutrum enim sapien, et lobortis sapien scelerisque a. Ut faucibus, erat non convallis blandit, dui lorem sollicitudin mauris, sit amet dictum mi enim ac tortor. Integer placerat vehicula tellus vitae egestas.

5

RESULTS

Mauris diam sapien, consequat non velit at, viverra pretium est. Mauris quis leo est. Praesent dictum posuere accumsan. Suspendisse feugiat metus quis risus ultrices, lobortis aliquam purus faucibus. Curabitur viverra tempus massa vitae blandit. Etiam sodales facilisis ullamcorper. Morbi porttitor consectetur est, sit amet vestibulum turpis euismod sagittis. Morbi nec odio nulla. Nunc mollis eros ut velit aliquam mollis. Pellentesque feugiat suscipit nibh, sed consequat eros venenatis at. Quisque nec risus justo. Sed mauris sapien, tincidunt ut dictum a, accumsan faucibus augue. Suspendisse sollicitudin congue arcu, efficitur vulputate nisi semper vitae. Morbi porttitor sodales lorem ac venenatis. Cras vitae tellus ultrices erat ultrices congue posuere ut velit. Quisque id tortor vel arcu efficitur posuere.

Sed molestie nibh lacinia rutrum mattis. Praesent dignissim eget turpis in porta. Fusce ac enim at augue dignissim rhoncus ac quis elit. Etiam accumsan, magna vitae tristique tincidunt, lectus turpis feugiat magna, sit amet consectetur ante risus non erat. In tincidunt ultrices nisl vel dignissim. Phasellus ipsum leo, feugiat et imperdiet nec, cursus et ex. Duis vehicula, sapien ut congue scelerisque, augue leo dictum ex, eget ultricies urna nisi et diam. Aliquam ullamcorper, tellus aliquet bibendum ullamcorper, neque nisl euismod diam, at facilisis ex nisl in leo. Mauris auctor sodales lorem id ultrices. Nullam mauris ligula, sollicitudin at leo et, eleifend tempus ex. Suspendisse nec blandit risus. Suspendisse sollicitudin ipsum vitae orci pulvinar, convallis ullamcorper orci molestie.

Vestibulum egestas dictum eros eu posuere. Vestibulum quis tincidunt sem. Etiam in facilisis dui. Cras viverra magna eget fringilla semper. Etiam ut orci fringilla, faucibus sem ac, facilisis dolor. Nunc molestie, neque id varius fermentum, ipsum sem tempus neque, quis commodo dui tortor placerat nulla. Morbi posuere nisl vitae posuere varius. Phasellus ac accumsan lectus. Nunc dictum fermentum vehicula. Morbi fermentum, odio id sollicitudin sodales, magna nibh vulputate tortor, eget ultricies lacus sem a sem. Quisque pellentesque id libero ac efficitur. Nulla aliquet facilisis venenatis.

Aliquam blandit cursus faucibus. Integer semper est a laoreet tincidunt. Suspendisse potenti. Quisque eget ullamcorper ligula, porta ullamcorper neque. Suspendisse vel justo arcu. Nam pretium turpis id lacinia ultrices. Phasellus ullamcorper lorem est, quis condimentum sem molestie sit amet. Etiam pellentesque sodales ante, vitae rhoncus nisl ornare non. Etiam nec ante laoreet, maximus lectus ac, tempus orci. Proin efficitur ultricies purus, vel volutpat purus ultrices sed. Fusce quis est ex. Nunc suscipit urna et arcu mattis, id vestibulum purus placerat.

Nulla non interdum nulla, vel convallis neque. Praesent eu sapien id tellus convallis bibendum. Integer quis justo et est lacinia consectetur. Praesent nec tellus vitae lacus convallis egestas. Pellentesque aliquam semper quam ut tempor. Sed posuere condimentum arcu, et vulputate nibh sodales ac. Aliquam imperdiet in nulla ac vulputate. Aenean nibh risus, vulputate non risus quis, dignissim pellentesque velit. Donec mi neque, aliquet ut sapien id, lobortis elementum ipsum. Nulla mattis accumsan eros, convallis faucibus ante congue vitae. Nunc pretium ultrices libero. Maecenas eu felis massa.

Donec in enim eget lacus facilisis suscipit. Vestibulum elementum lorem vitae felis euismod tincidunt. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Etiam nec eleifend tellus, ut venenatis metus. Aenean sit amet nisl pellentesque, iaculis purus vel, ornare lorem. Maecenas efficitur porttitor nibh, non malesuada nunc hendrerit ac. Nunc condimentum, orci at porta aliquam, urna purus rutrum mauris, at mattis lorem erat id enim. Nam auctor lacinia neque. Nullam gravida erat sollicitudin, placerat ante a, tincidunt lorem. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In iaculis ipsum nec velit imperdiet, faucibus commodo ipsum tempus.

Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Nullam luctus vulputate turpis sit amet gravida. Curabitur id lobortis est, quis pellentesque orci. Cras quis finibus mauris. Curabitur congue elit massa, a pulvinar arcu hendrerit non. Aliquam erat volutpat. Aliquam erat volutpat. In tempor leo ac ex rhoncus, ut facilisis lorem ornare.

6

ANALYSIS

Mauris diam sapien, consequat non velit at, viverra pretium est. Mauris quis leo est. Praesent dictum posuere accumsan. Suspendisse feugiat metus quis risus ultrices, lobortis aliquam purus faucibus. Curabitur viverra tempus massa vitae blandit. Etiam sodales facilisis ullamcorper. Morbi porttitor consectetur est, sit amet vestibulum turpis euismod sagittis. Morbi nec odio nulla. Nunc mollis eros ut velit aliquam mollis. Pellentesque feugiat suscipit nibh, sed consequat eros venenatis at. Quisque nec risus justo. Sed mauris sapien, tincidunt ut dictum a, accumsan faucibus augue. Suspendisse sollicitudin congue arcu, efficitur vulputate nisi semper vitae. Morbi porttitor sodales lorem ac venenatis. Cras vitae tellus ultrices erat ultrices congue posuere ut velit. Quisque id tortor vel arcu efficitur posuere.

Sed molestie nibh lacinia rutrum mattis. Praesent dignissim eget turpis in porta. Fusce ac enim at augue dignissim rhoncus ac quis elit. Etiam accumsan, magna vitae tristique tincidunt, lectus turpis feugiat magna, sit amet consectetur ante risus non erat. In tincidunt ultrices nisl vel dignissim. Phasellus ipsum leo, feugiat et imperdiet nec, cursus et ex. Duis vehicula, sapien ut congue scelerisque, augue leo dictum ex, eget ultricies urna nisi et diam. Aliquam ullamcorper, tellus aliquet bibendum ullamcorper, neque nisl euismod diam, at facilisis ex nisl in leo. Mauris auctor sodales lorem id ultrices. Nullam mauris ligula, sollicitudin at leo et, eleifend tempus ex. Suspendisse nec blandit risus. Suspendisse sollicitudin ipsum vitae orci pulvinar, convallis ullamcorper orci molestie.

Vestibulum egestas dictum eros eu posuere. Vestibulum quis tincidunt sem. Etiam in facilisis dui. Cras viverra magna eget fringilla semper. Etiam ut orci fringilla, faucibus sem ac, facilisis dolor. Nunc molestie, neque id varius fermentum, ipsum sem tempus neque, quis commodo dui tortor placerat nulla. Morbi posuere nisl vitae posuere varius. Phasellus ac accumsan lectus. Nunc dictum fermentum vehicula. Morbi fermentum, odio id sollicitudin sodales, magna nibh vulputate tortor, eget ultricies lacus sem a sem. Quisque pellentesque id libero ac efficitur. Nulla aliquet facilisis venenatis.

7

CONCLUSIONS

Nulla non interdum nulla, vel convallis neque. Praesent eu sapien id tellus convallis bibendum. Integer quis justo et est lacinia consectetur. Praesent nec tellus vitae lacus convallis egestas. Pellentesque aliquam semper quam ut tempor. Sed posuere condimentum arcu, et vulputate nibh sodales ac. Aliquam imperdiet in nulla ac vulputate. Aenean nibh risus, vulputate non risus quis, dignissim pellentesque velit. Donec mi neque, aliquet ut sapien id, lobortis elementum ipsum. Nulla mattis accumsan eros, convallis faucibus ante congue vitae. Nunc pretium ultrices libero. Maecenas eu felis massa.

Donec in enim eget lacus facilisis suscipit. Vestibulum elementum lorem vitae felis euismod tincidunt. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Etiam nec eleifend tellus, ut venenatis metus. Aenean sit amet nisl pellentesque, iaculis purus vel, ornare lorem. Maecenas efficitur porttitor nibh, non malesuada nunc hendrerit ac. Nunc condimentum, orci at porta aliquam, urna purus rutrum mauris, at mattis lorem erat id enim. Nam auctor lacinia neque. Nullam gravida erat sollicitudin, placerat ante a, tincidunt lorem. Lorem ipsum dolor sit amet, consectetur adipiscing elit. In iaculis ipsum nec velit imperdiet, faucibus commodo ipsum tempus.

Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Nullam luctus vulputate turpis sit amet gravida. Curabitur id lobortis est, quis pellentesque orci. Cras quis finibus mauris. Curabitur congue elit massa, a pulvinar arcu hendrerit non. Aliquam erat volutpat. Aliquam erat volutpat. In tempor leo ac ex rhoncus, ut facilisis lorem ornare.

8

FUTURE WORK

Keywords: Signal Processing; Sentiment Analysis; Machine Learning; Feature Fusion; Multimodal Classification ; I NEED TO DISCUSS THIS PART