Categorization of Text into Appropriate Sentiment for Automatic Synthesis of Expressive Speech

Sayeda Swaleha Peerzade, Prof. Ramesh Bhat

Department Of Computer Science

PESIT, Bangalore

Visvesvaraya Technological University

e-mail ID: swalehapeerzade@gmail.com

Abstract—**The attention of researchers is directed towards the generation of expressive speech. Hence the text which would enter the text-to-speech system must be such that, the sentiment which it encompasses is to be known. So that this sentiment analyzed text can be then used to synthesize expressive speech. This paper concentrates on identifying and categorizing the sentiment present in text, which is the input to the text-to-speech system, which is the preliminary step for the production of expressive speech. The sentence is preprocessed and classification is done using the classifiers and sentiment tagged sentence is passed as input to the text-to-speech system and text-to-speech system selects the voice based on the sentiment and converts text to expressive speech. The twitter corpus is used as the dataset for training and testing the classifier because of its affect based expressiveness.**

Keywords- Sentiment Analysis; Text-To-Speech System; Natural Language Processing; Expressive Speech; Opinion Mining

I. INTRODUCTION

Sentiment analysis is the procedure of distinguishing and categorizing the sentiment present in a piece of text, to figure out if the people's opinion towards a specific topic is positive or negative. Sentiment Analysis is categorized into three levels as follows: Sentiment Analysis of a document, which takes into account the whole document, detects the sentiment present in the entire document and converges it into the single sentiment. Sentiment Analysis of sentence recognizes each sentence and gives sentiment for it. Sentiment Analysis of Aspects identifies different features of entities and gives them the sentiment. The attention of researchers is directed towards the generation of expressive speech. Hence the text which would enter the text-to-speech system must be such that, the sentiment which it encompasses is to be known. So that this sentiment analyzed text can be then used to synthesize expressive speech. This paper concentrates on recognizing and categorizing the sentiment present in text, which is the input to the text-to-speech system, which is the preliminary step for the production of expressive speech. Couple of speech specialists they considered domain based expressiveness for example, the advertising scenario, which causes problem when new domains are discovered, then Few others made use of affect based expressiveness for example, story-telling. It is additionally used as a part of conversational interfaces, where distinct topics are secured under a sole discussion.

Presently there are numerous difficulties in making an interpretation of human affect into explicit representations. One way is to consider the scientific classifications of affective states like sentence level and emotion level.

Couple of researchers tried categorizing emotions as anger and happiness. While few others directly considered the six emotions happy, sad, anger, fear, disgusted and surprised called the BigSix. But the directly expression of these emotions is not usual in conversational interfaces like twitter etc. Hence we need some complete distinction as to positive and negative. This is regarded as the sentiment analysis problem.

Neutral sentiment at the same level of the hierarchy can also be considered. The text-to-speech system applications make use of this neutral sentiment to generate usual messages that are not positive or negative e.g, when reporting objective information such as a description. The proposed project categorizes input plain text into appropriate sentiment i.e. positive, negative or neutral for automatic synthesis of expressive speech at sentence level.

II. PROBLEM STATEMENT

The prediction of affect in text is a topic that is mainly related to NLP, but it has also attracted the attention of speech researchers. In the TTS scenario, the granularity of the text under analysis is usually determined to be the sentence, as sentences are sensibly short textual representations with a rich expressive content and natural expressive variations can be considered within the same paragraph. Sentiment Analysis solutions of NLP work with compilation of long texts and are not analyzed at the sentence level. Sentiment analysis methods must be adapted to the TTS environment. Few heuristic approaches tackled this short text setting by effectively weighting the lexicon and then spotting keywords in the sentences. But this is not that efficient. So machine learning approaches need to be used to directly learn from the previous example sentences.

The sentiment analysis framework needs to consider both the diversity in the nature of the features

extracted from the text and the diversity in the learning principles of the classifiers, and must select the most effective system.

III. RELATED WORK

1. Ahmed Hassan , Walaa Medhat , Hoda Korashy "Sentiment analysis algorithms and applications: A survey" Ain Shams Engineering Journal (2014),.

Sentiment Analysis (SA) or Opinion Mining (OM) is the computational study of people's opinions, attitudes and emotions toward an entity. . The three main classification levels in SA: document-level, sentence-level, and aspect-level SA. Document-level SA aims to classify an opinion document as expressing a positive or negative opinion or sentiment. It considers the whole document a basic information unit i.e. talking about one topic. Sentence-level SA aims to classify sentiment expressed in each sentence. Aspect-level SA aims to classify the sentiment with respect to the specific aspects of entities.

2. Cecilia Ovesdotter Alm Dan Roth, Richard Sproat "Emotions from text: machine learning for text-based emotion prediction"

In addition to information, text contains attitudinal, and more specifically, emotional content. The text-based emotion prediction problem empirically, using supervised machine learning with the SNoW learning architecture is used. The goal is to classify the emotional affinity of sentences in the narrative domain of children's fairy tales, for subsequent usage in appropriate expressive rendering of text-to-speech synthesis.

3. Surabhi Thorat, "Opinion Mining and Sentiment Analysis- Its Tools and Challenges" Marathwada Institute of Technology(E), Aurangabad(M.S), India

Opinion mining is a combination of information retrieval and computational linguistic techniques deals with the opinions expressed in a document. Opinion Mining or Sentiment Analysis is the field to extract the opinionated services are largely available on internet. The data given by reviewers are collected from the e-commerce websites like www.amazon.com for product reviews. A large number of user-generated reviews are available on the Internet. The reviewers data used in most of the sentiment classification studies are collected from the e-commerce websites like www.amazon.com for product reviews.

**4. Jurafsky and Martin, "Speech and Language Processing"**

Words are traditionally grouped into equivalence classes called parts of speech (POS), Word Classes, Morphological Classes or lexical tags. The part of speech for a word gives a significant amount of information about the word and its neighbors. This is clearly true for major categories, (verb versus noun), but is also true for the many finer distinctions. For example, these tagsets distinguish between possessive pronouns (my, your, his, her, its) and personal pronouns (I, you, he, me). Knowing whether a word is a possessive pronoun or a personal pronoun can tell us what words are likely to occur in its vicinity

(possessive pronouns are likely to be followed by a noun, personal pronouns by a verb). This can be useful in a language model for speech recognition.

5. A. Trilla, F. Alías, and I. Lozano , "Text classification of domain-styled text and sentiment-styled text for expressive speech synthesis," in Proc. VI Jornadas en Tecnología del Habla (FALA2010), Nov. 2010, pp. 75–78.

In the context of text processing for Text-to-Speech (TTS) synthesis, this work aims to automatically direct the expressiveness in speech through tagging the input text appropriately. Since the nature of text presents different characteristics according to whether it is domain-dependent (related to its topics) or sentiment-dependent, it is studied how these traits influence the identification of expressiveness in text.

IV. SYSTEM ARCHITECTURE
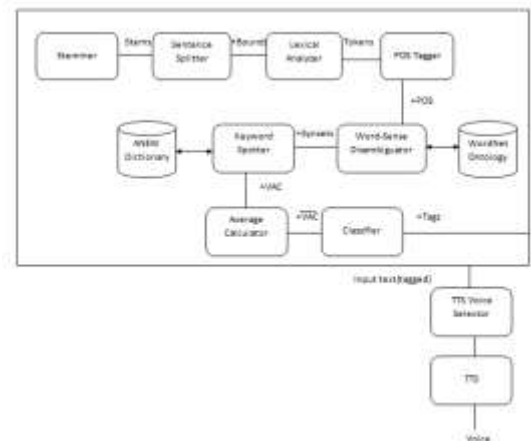
V. DESIGN STEPS

- **Stemming**

In this step, the suffixes are removed to index the base word which would share the common meaning. This is accomplished using porter stemming algorithm

- **Sentence splitting**

In this step, the input sentences are segmented on encountering the boundaries, which are called as sentence delimiters such as periods, exclamation and question marks. This is accomplished using regex.

- **Stop word removal**

In this step, the text is converted into tokens. The function words or stop words like an, the etc are removed.

- **Parts-Of-Speech tagging**

It determines the function of nouns, verbs and adjectives within the sentence using OpenNLP tagger

- **Word-Sense Disambiguation**

It determines the meaning of the word with the relevance according to the context. This is WordNet ontology is used for this purpose. The senses of the words are scored. And the synonyms of the words are retrieved to increase feature space.

- **Keyword spotting**

Using ANEW dictionary of affect the dimensions like Valence, Activation and Control are provided. POS tags and stems of the emotional words are considered. This is known as circumflex.

- **Average calculation**

It computes the averaged emotional dimensions for the text under analysis.

- **Classification**

The appropriate sentiment is predicted using the classifiers, which makes use of the features that are taken from the text under analysis. The recognition and creation of model is done using the bag of words.

Multinomial Naive Bayes Classifier : It is a probabilistic classification method which assumes conditional dependence among the features and builds a language model. In reality, the textual data the assumption does not hold, since it is oversimplified model its probabilistic estimates are of low quality, in spite of this, its decisions of classification are surprisingly good. The MNB provides good efficiency i.e the time performance is optimal and also good accuracy, hence research involving text classification and sentiment analysis considers Naive Bayes Classifier as a baseline.

Multinomial Logistic Regression : Probabilistic discriminative approach that fits a set of exponential functions via the Maximum A Posteriori estimation .MLR obeys the maximum entropy principle, therefore it does not make any further assumption beyond what is directly observed in the training data. Moreover, it makes no assumptions about the relationships among the features, and so might potentially be more effective when conditional independence assumptions are not met , also overcoming the sparseness problem. MLR has been used for SA in TTS and TC environments
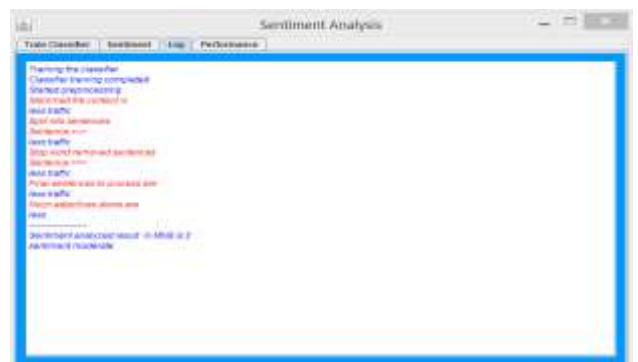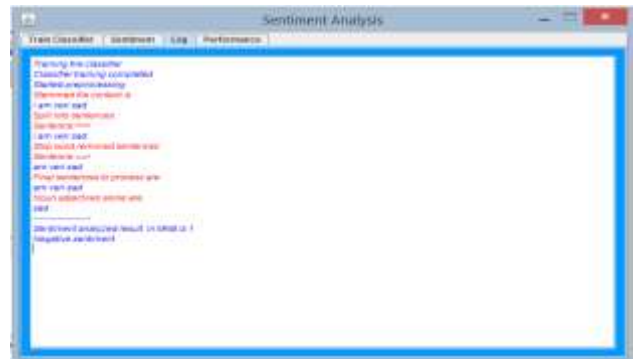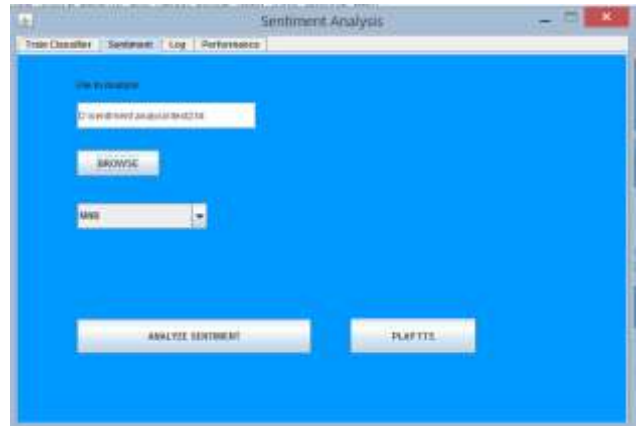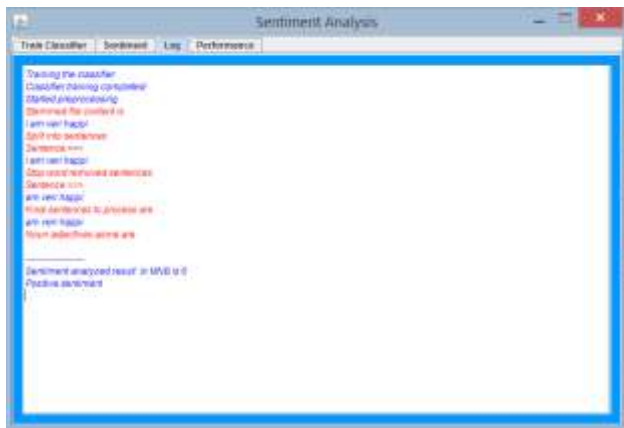
- **TTSVoiceSelector:**

This module will select the voice to represent it whether it is positive, negative are neutral sentiments.

- **TTS:**

Once classified the sentiments , the related voice will be played as the output.

VI. SNAPSHOTS

VII. CONCLUSION

The identification of affect in text is a complex problem that has many facets to consider. The proposed project performs an exhaustive and comprehensive study to tackle a particular three class sentiment analysis problem, at the sentence level, framed by a TTS scenario and without using additional textual data. This is one of the first attempts to adapt conventional SA methods to the TTS

VIII. REFERENCES

1. Alexandre Trilla and Francesc Alías, "Sentence Based Sentiment Analysis For Expressive Text-To-Speech" IEEE Transactions On Audio, Speech, And Language Processing, Vol. 21, No. 2, February 2013

2. Cecilia Ovesdotter, Alm Dan Roth, Richard Sproat, "Emotions from text: machine learning for text-based emotion prediction"

3. Walaa Medhat , Ahmed Hassan, Hoda Korashy, "Sentiment analysis algorithms and applications: A survey" Ain Shams Engineering Journal (2014),

4. Surabhi Thorat, "Opinion Mining and Sentiment Analysis-Its Tools and Challenges"

5. F. Alías, X. Sevillano, J. C. Socoró, and X. Gonzalvo, "Towards high-quality next generation text-to-speech synthesis: A multidomain approach by automatic domain classification," IEEE Trans. Audio, Speech, Lang. Process., vol. 16, no. 7, pp. 1340–1354, Sep. 2008.

6. J. Bellegarda, "A data-driven affective analysis framework toward naturally expressive speech synthesis," IEEE Trans. Audio, Speech, Lang. Process., vol. 19, no. 5, pp. 1113–1122, Jul. 2011.

7. Jurafsky and Martin, "Speech and Language Processing"

8. Zheng-Jun Zha, Jianxing Yu, Meng Wang, and Tat-Seng Chua, "Product Aspect Ranking and Its Applications"

9. Danushka Bollegala,, David Weir and John Carroll, "Cross-Domain Sentiment Classification using a Sentiment Sensitive Thesaurus"

10. A. Trilla, F. Alías, and I. Lozano , "Text classification of domain-styled text and sentiment-styled text for expressive speech synthesis," in Proc. VI Jornadas en Tecnología del Habla , pp 75–78, Nov. 2010

11. Jiawei Han and Micheline Kamber, "Data Mining: Concepts and Techniques"