SENTIMENT ANALYSIS OF DIFFERENT LANGUAGES

CONTENTS

ABSTRACT

- Sentiment Analysis (SA) has been a core interest in the field of text mining research, dealing with computational processing of sentiments, views, and subjective nature of the text.
- In this work, we focus on mining sentiments of 5 languages and analyzing them.
- This research portrays a complete analysis specifically in the field of Hindi, Malayalam, Kannada, Marathi and Gujarati Language.
- It discusses the importance of techniques based on different aspects such as their impact on the issues of Sentiment Analysis, levels of analysis, basically it gives a comprehensive overview of the majority of work done in NLP.

LITERATURE SURVEY

S.No	TOPIC	APPROACH	AUTHOR	DESCRIPTION
1.	Sentiment Analysis of Indian Languages using Convolutional Neural Networks.	The method used is CNN for sentiment analysis of Bengali and Telugu movie data.	Shalini K, Aravind Ravikumar, Vineetha R C, Aravinda Reddy , Anand Kumar M and Soman K.P	It is a text mining technique that effectively measures the inclination of opinions and aids in analyzing the subjective information from the given context.
2.	Sentiment Analysis of Movie Reviews Using Heterogeneous Features.	Naive Bayes and Linear Support Vector Machine (LSVM) is used to build the system model.	Rachana Bandana	Here used different large and small data sets for different preprocessing techniques, heterogeneous features and classification algorithms and observe various results.

LITERATURE SURVEY

S.No	TOPIC	APPROACH	AUTHOR	DESCRIPTION
3.	Sentiment Analysis of IMDb Movie Reviews Using Long Short-Term Memory.	LSTM classifier based on the RNN algorithm is used for analyzing sentiments of the IMDb movie reviews	Saeed Mian Qaisar	While dealing with the IMDb intended database portion, the confusion matrix is generated to assess the performance accuracy of the LSTM classifier. Confusion matrix provides a summary of correct and incorrect predictions
4.	Sentiment Analysis of Movie Reviews in Hindi Language using Machine Learning.	The model is implemented over Random Forest and SVM for Sentiment Analysis	Charu Nanda, Mohit Dua and Garima Nanda	The proposed approach would classify the reviews in Hindi language into two classes: Positive and negative and will also evaluate the performance of by working on the algorithm through different metrics.

LITERATURE SURVEY

S.No	TOPIC	APPROACH	AUTHOR	DESCRIPTION
5.	Sentiment analysis using Telugu Senti-WordNet.	The Implementation is done using Telugu Senti-WordNet.	Reddy Naidu, Santosh Kumar Bharti, Korra Sathya Babu and Ramesh Kumar Mohapatra	The proposed system for sentiment analysis has attained an accuracy of 74% for subjectivity classification and 81% for sentiment classification in the domain of news data. This paper exploits the available Telugu Senti-WordNet to perform sentiment analysis for Telugu e-Newspapers sentences.

INTRODUCTION

- Sentiment analysis is the process of using natural language processing, text analysis, and statistics to analyze Users sentiment to determine whether data is positive, negative or neutral.
- In this work, we focus on mining sentiments and analyzing them for 5 languages.
- This model will help you to derive polarity precision.
- For example: “I really like the new design of your website!” → Positive.

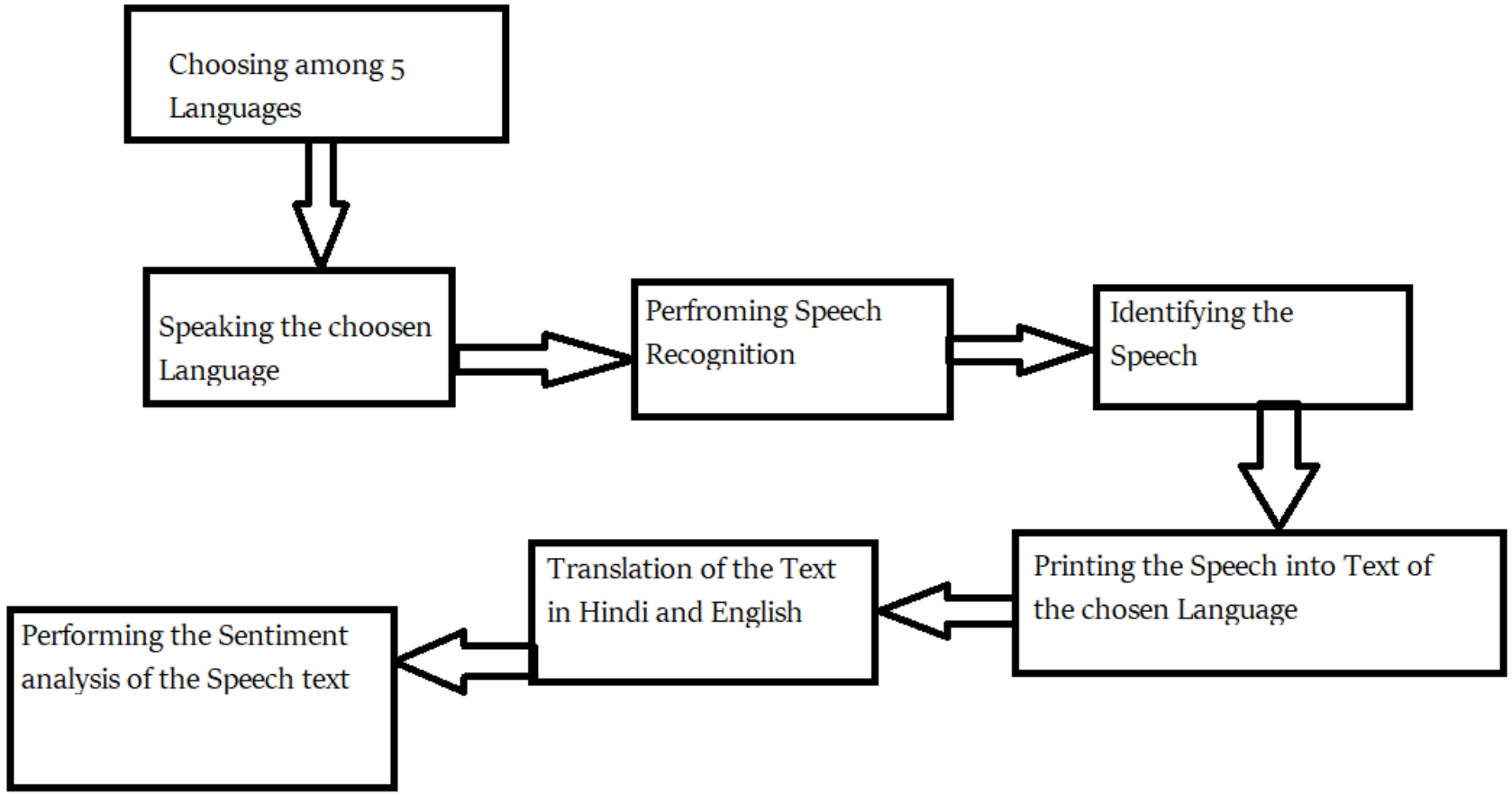
METHODOLOGY

- Apart from the existing system for Sentiment Analysis here Speech Recognition is being used in the proposed system in order to improve the Real time efficiency.
- Speech recognition is a machine's ability to listen to spoken words and identify them.
- Natural Language Processing(NLP) is used which makes it possible for computers to read text, hear speech, interpret it, measure sentiment and determine which parts are important.
- Vader Sentiment is used which is basically a rule-based sentiment analysis tool that is specifically attuned to sentiments expressed and works well on texts from other domains.

METHODOLOGY

- With the help of Speech Recognition the system is recognizing the user's speech and there is a pause threshold which represents the minimum length of silence (in seconds) that will register as the end of a phrase.
- Google Speech Recognition API is used to perform speech recognition on a Audio data instance.
- Google Translation library uses the Google Translate Ajax API to detect and translate the text.
- Sentiment polarity for an element defines the orientation of the expressed sentiment.
- Then Polarity Score is calculated, Polarity is float which lies in the range of [-1,1] where 1 means positive statement and -1 means a negative statement.

METHODOLOGY



RESULTS AND DISCUSSION

- In this System We are analysing the sentiments of speech in 5 different languages .
- 1 is for Hindi , 2 is for Malayalam, 3 is for Marathi, 4 is for Kannada, 5 is for Gujarati
- Select a language from the 5 different languages.
- The input speech from the user in a chosen language will be recognised, then it will be converted into the text of the particular chosen language by the user.

.

RESULTS AND DISCUSSION

- Then the text of that language is translated into English and Hindi.
- The sentimental score of the text is analysed with the help of Sentiment Intensity Analyzer.
- Finally the sentiment of the sentence will be shown.
- The sentiments can be positive, negative or neutral.

RESULTS AND DISCUSSION

```
choose the language
1.Hindi
2.Malayalam
3.Marathi
4.Kannada
5.Gujarati2
Listening
Recognizing
the query is printed=' സുഖമാണോ '
Translated(src=ml, dest=en, text=How are you, pronunciation=None, extra_data="{'translat...")
Translated(src=ml, dest=hi, text=क्या हाल है, pronunciation=kya haal hai, extra_data="{'translat...")

Translated Sentence= Translated(src=ml, dest=en, text=How are you, pronunciation=None, extra_data='
Dictionary= {'neg': 0.0, 'neu': 1.0, 'pos': 0.0, 'compound': 0.0}
It is a Neutral Sentence

Process finished with exit code 0
```

CONCLUSION

- The sentiment analysis of Indian languages using NLP has gained high importance in almost every fields such as entertainment industry, business, forensics, etc.
- Enormous amount of user-generated content in a native language such as Hindi is available, which needs to be analyzed to generate information.
- This research work is created by collecting the speeches from the users in 5 different languages by using Speech Recognition and NLP.
- Translating the speech of a particular language into Hindi and English and then annotate into positive, negative, and neutral classes.

REFERENCES

- Kush Shrivastava and Shishir Kumar "A Sentiment Analysis System for the HindiLanguage by Integrating Gated Recurrent Unit with Genetic Algorithm." The International Arab Journal of Information Technology, Vol. 6, November 2020
- Shalini K, Aravind Ravikumar, Vineetha R C, Aravinda Reddy , Anand Kumar M and Soman K.P "Sentiment Analysis of Indian Languages using Convolutional Neural Networks." 2018 International Conference on Computer Communication and Informatics (ICCCI -2018), July 11-12 2019.
- Rachana Bandana "Sentiment Analysis of Movie Reviews Using Heterogeneous Features." 2018 2nd International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech), May 04-05 ,2018.

REFERENCES

- Saeed Mian Qaisar "Sentiment Analysis of IMDb Movie Reviews Using Long Short-Term Memory." 2020 2nd International Conference on Computer and Information Sciences (ICCIS), October 13-15 2020.
- Charu Nanda, Mohit Dua and Garima Nanda "Sentiment Analysis of Movie Reviews in Hindi Language using Machine Learning." 2018 International Conference on Communication and Signal Processing (ICCSP), April 03-05 2018.
- Reddy Naidu, Santosh Kumar Bharti, Korra Sathya Babu and Ramesh Kumar Mohapatra "Sentiment analysis using Telugu SentiWordNet." 2020 IEEE Conference on Computer Applications(ICCA), February 27-28 2020.

THANK YOU

ANY QUESTIONS ?