# INCLINATION ANALYSIS OF MARATHI LANGUAGE USING CLASSIFIER

Abuzar Siddiqui[1] Siddhesh Shirke[2] Sachin Singh[3] Vaishali Yeole[4]
1Engineering Student, 2Engineering Student, 3Engineering Student, 4Assistant Professor
1Computer Engineering Department,
1K. C. College of Engineering and Management Studies and Research, Mumbai, India

***Abstract:*** Inclination Analysis (SA) is trending and most talked about fields in data mining (DM) and natural language processing (NLP).This field has grown tremendously with the advent of internet. The Internet has provided a platform for people to express their views, emotions and Inclination towards products, people and life in general. Thus, the Internet is now a vast resource of opinion rich textual data. The goal of Inclination Analysis is to harness this data in order to obtain important information regarding public opinion, that would help make smarter business decisions, political campaigns and better product consumption. Inclination Analysis focuses on identifying whether a given piece of text is subjective or objective and if it is negative or positive.

***Index Terms*- multi-way Inclination analysis, hierarchical classifiers, support vector machine, decision tree, naive bayes**

## I. INTRODUCTION

Inclination analysis can be defined as a process that automates mining of attitudes, opinions, views and emotions from text, speech and database sources through Natural Language Processing (NLP). Inclination analysis involves classifying opinions in text into categories like "positive" or "negative" or "neutral". It's also referred as subjectivity analysis, opinion mining, and appraisal extraction.

The words sentiment & emotions are used interchangeably but there is difference between them.

- Emotions: is a state of mind. For eg: angry.
- Sentiment: behaviour influenced by the emotion. For eg: Writing a review.

## II. EXISTING SYSTEM

In recent years a lot of work has been done in the field of “Inclination analysis“ by number of researchers. In fact work in the field started since the beginning of the century. A lot of work is done for languages like English, Arabic, Hebrew but a reliable system for Marathi Language is not yet created.

## III. PROPOSED SYSTEM

To overcome the above mentioned drawbacks we will create a system that can extract the reviews information from Marathi paragraph using web mining technique and analyze this sentences using Inclination analysis.The system will use the concept of polarity calculation & will use Naive Bayes classifier to classify the reviews based on the calculated polarity values.

## IV. SYSTEMIMPLEMENTATION

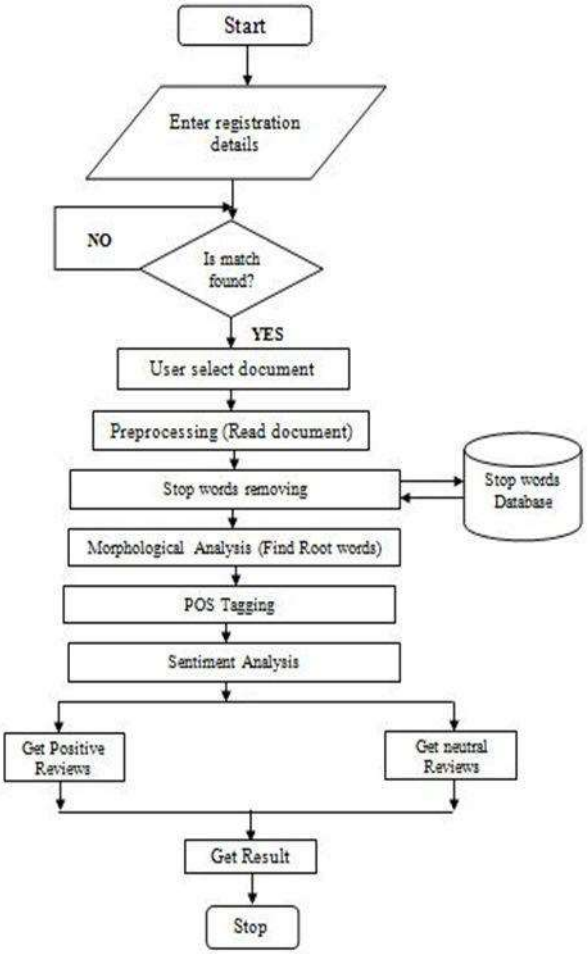


Fig.1. Overview of System

In this proposed system we can use Naive Bayes algorithm for analyze reviews and to tag a given review as positive or negative. The results can be used for various purposes such as guiding decisions to improve the system.

**Module Description:**

1. **Login**

In this module using username and password user login into system. In this login system authentication of user so only valid person login into the system.

2. **Data Collection**

In this user select one document from system and click on submit after submitting our system get preprocess. In this phase all data get from document and remove Stop words where Inclinationanalysis has to perform.

**3. Morphological Analysis**

In this step, the analyses of internal structure of words and the form and structure of organisms and their specific features from paragraph and find out Root words from the given paragraph.

**4. POS Tagging**

A Part-Of-Speech Tagger (POS Tagger) is step that reads text in paragraph and assigns parts of speech to each word (and other token), such as noun, verb, adjective, etc., although generally computational applications use more fine-grained POS tags like 'noun-plural'.

5. **InclinationAnalysis**

The comments using web mining technique from paragraph which can be analyzed using Naive Bayes algorithm and get result as positive and negative reviews.

6.**Output**

In this module we can display out final meaningful words from the given paragraph.

## VI. APPLICATION

- Applications that use Reviews from Websites
- Applications as a Sub-component Technology
- Applications in Business Intelligence
- Applications across Domains
- Applications in Smart Homes

## ADVANTAGES :

- Focuses exclusively on Marathi Language.
- Using this analysis we can easily identify drawbacks of our product to iteratively improve the product.

## LIMITATION :

- It cannot detect sarcasm in text.
- To overcome this limitation, we can use audio of the subject's opinion & perform stress analysis on it.

## VII. CONCLUSION

- Inclination Analysis will be a mandatory requirement eventually in all businesses.
- Naive Bayes classifier is easy to train & provides 61.79% accuracy.
- This problem is yet to find sufficient amount of interest in Marathi text processing & mining community.

## REFERENCES

[1] Muhammad Abdul-Mageed and Mona T Diab. Subjectivity and sentiment annotation of modern standard arabic newswire. In Proceedings of the 5th Linguistic Annotation Workshop, 2011.

[2] Muhammad Abdul-Mageed and Mona T Diab. Awatif: A multi-genre corpus for modern standard arabic subjectivity and sentiment analysis. In LREC, pages 3907–3914, 2012.

[3] Muhammad Abdul-Mageed, Mona T Diab, and Mohammed Korayem. Subjectivity and sentiment analysis of modern standard arabic. In ACL, 2011.

[4] Muhammad Abdul-Mageed et al. SAMAR: A system for subjectivity and sentiment analysis of arabic social media. In WASSA, 2012.

[5] Nawaf Abdulla et al. Arabic sentiment analysis: Corpus-based and lexicon-based. In AEECT.IEEE, 2013.

[6] Nawaf Abdulla et al. Automatic lexicon construction for arabic sentiment analysis. In FiCloud, 2014.

[7] Nawaf Abdulla et al. Towards improving the lexicon-based approach for arabic sentiment analysis. International Journal of Information Technology and Web Engineering (IJITWE), 9(3):55–70, 2014.

[8] Nawaf A Abdulla, Mahmoud Al-Ayyoub, and Mohammed NajiAlKabi. An extended analytical study of arabic sentiments. International Journal of Big Data Intelligence, 1(1):103–113, 2014.

[9] Mohammed Korayem, David Crandall, and Muhammad Abdul-Mageed. Subjectivity and sentiment analysis of arabic: A survey. In AboulEllaHassanien, Abdel-BadeehM. Salem, Rabie Ramadan, and Tai-hoon Kim, editors, Advanced Machine Learning Technologies and Applications, volume 322 of Communications in Computer and Information Science, pages 128–139. Springer Berlin Heidelberg, 2012.

[10] P.D. Turney," Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews," Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL), Philadelphia, pp. 417-424, July 2002