Automated Sentiment Analysis for Web Multimedia

Abstract

Over the last decade, the world has experienced rapid changes. Modern life has become, and the people of the world have to thank the immense contribution of internet technology to communication and information sharing. The Internet has emerged as a global encyclopedia of information. Any kind of information on any topic under the sun is available on the internet including videos and audios.

The proposed system is to automate the content identification using Machine Learning and Natural Language Processing. The scheme aims to extract the audio stream from any multimedia as the input. The system will automatically generate raw text useful for identification of the type of multimedia using natural language processing. Further using this data and machine learning, the system will label the content with the opinion of the speaker as output. It will show various sentiments along with their intensity.

The main application of this system is Google Pod-casts and YouTube Videos. The system will be able to flag the inappropriate content and inform officials about censorship automatically. Implementation is in the form of a web browser extension. The system will display relevant sentiments helping the users from click-baits.

Key-words: Sentiment Analysis, Natural Language Processing, Information Retrieval, Artificial Intelligence, Speech Recognition, Web Browser, User Experience.

Project Members:

Devashish Katoriya

Ashutosh Bhawsar

Ninad Kapadnis

Bhushan Shilawat

Project Guide:

Prof. A. V. Kolapkar