Movie Genre Detection from Subtitle using NLP

A novel approach to identify movie genre based on the emotions from subtitles

[1] Dr. Vinayak Bharadi, [2]Atharva Malji, [3]Atharva Mestri, [4]Yash Narvekar, [5]Amogh Anerao

[1]Associate Professor, Head of Department, [2,3,4,5]Student
Department of Information Technology
Finolex Academy of Management and Technology, Ratnagiri, India

Abstract: Movies have been an essential part of our entertainment since their inception. Owing to technological advancements and the increasing interests of business entities, the categorization of media products has been made conventional in the digital environment. In most cases, this is a multi-label scenario in which an object might be labeled with several categories. Most of the literature addresses the classification of movie genre as a mono-labeling task, generally based on audio-visual features. Our proposed system aims to detect movie genres based on the yet unused movie's subtitles which are a documented account of the movie's visual content and dialogues. The basic idea is to identify words that have a high frequency in a particular genre and use them as features for training the classification machine learning models. Alongside that, find emotions of the line for each subtitle and try to find a generalized emotion of the movie alongside it also find the profane and derogatory language and identify the occurrence of such instances in the movie to give it a rating. Here are using Google's T5 model for emotion detection.

Index Terms - **Emotion, Genre, Google T5, Profanity, Subtitles.**

I. INTRODUCTION

In today's world, movies are the best form of entertainment. Not only are they for entertainment, but they are also an important source of commercial, marketing, and educational benefits. With the growth of various technologies, online streaming of movies and TV shows has become popular, with Netflix, and Amazon Prime Video, there are multiple streaming platforms such as YouTube. Movies have different characteristics. Some of them are their type (for example, film, animation, documentary, flash), length, background music, emotional or tension quotient, etc. One of the essential characteristics of cinema is its genre. Several genres are prevalent such as sports, horror, romantic, thriller, crime, and fantasy. A movie's genre speaks about its general theme. A movie may belong to only one genre or can fall into several different genres. Subtitles of a movie provide a documented account of the movie's visual content and dialogues. Apart from being used just as a piece of information to understand the dialogues, we can use the subtitles to gather more insights about the movie. For successful detection of a movie's genre, it's critical to accurately identify the tone of the movie. Emotions play an important role in this identification. Identifying the emotions present in a movie and using them to determine the tone of a movie is a challenging task as there are many factors other than just text that are important to be taken into consideration such as audio-visual inputs.

II. LITERATURE REVIEW

Emotion detection is a key aspect when it comes to estimating an overall mood and tone of any movie. There are several studies that make use of audio-visual data inputs for emotion detection. In such techniques since it involves data rich sources making them easier to detect emotions. When it comes to emotion detection from text however, it becomes far more complex. However, unlike speech, where indications like tonal stress, facial expression, pitch, etc. are present, emotion recognition in text is a difficult undertaking because they are absent. [2] Long before the inception of computers, research on the definitions of emotion, its categories, and influences was crucial to deduce a person's emotional state in various scenarios. The problem of emotion identification from text is reduced to identifying relationships between particular input texts and the actual feelings that led the author to type or write in those specific ways.

[1] Intuitively, finding the relations usually relies on specific surface texts that are included in the input texts, and other deeper inferences. Once the relations can be determined, they can be generalized to predict other emotions from their articles, or even single sentences. This method can be used to identify the general tone of a movie based on the emotions that are present in a movie and then use it to predict it's genre.

In the research [3], the author developed a model for predicting the movie genre based on plot summaries. This article describes the implementation of several machine learning algorithms like Naive Bayes, Recurrent Neural Networks and Word2Vec + XGBoost for text classification and Probability threshold approach, K- binary transformation for genre selection to predict the movie genre based on summaries The experiment is performed with more than 250,000 movies which concluded that the Gated Recurrent Units (GRU) neural networks with probability threshold approach reaches the best outcome on the test sample.

III. SYSTEM MODEL

Emotion detection of a line means to classify a line into one of emotions present (joy, anger, fear, sadness, love, surprise). Google's T5 (Text-to-Text Transfer Transformer) is a Transformer that uses text-to-text approach
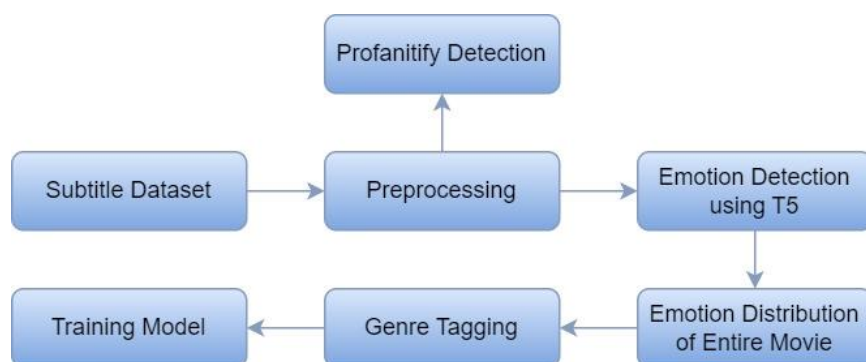

Fig. 1. Block diagram of Genre Detection, Profanity Detection and Emotion Detection from Subtitle.

Fig. 1. illustrates the block diagram of the proposed model where the genre detection model is trained from subtitle dataset that is generated using both emotion detection and labelling the movie's corresponding genre. Here the subtitle of a movie is given as input, which is further preprocessed to convert the subtitle file (.srt) into a pandas data-frame. This preprocessed data is passed to two nodes. First is the "Profanity Detection" node where the entire subtitle dataset is filtered against an already prepared word-bank for any profanity and derogatory words found. The other part is Emotion Detection where every line is passed to the emotion detection model to classify its emotion. Once the classification of all lines is done, we can use the distribution of different emotions spread across the movie's playtime as a classification vector. We then tag the movie with its respective genre and its corresponding emotion distribution to generate a dataset to train the Genre Detection Model.

IV. ALGORITHM

A. Emotion Detection

For detection of Emotion from text, Google's T5 model fine-tuned for emotion detection is used. It is a type of Transfer Learning, where a model is first pre-trained on a data-rich task before being fine-tuned on a downstream task. This model provides an easy way to extract emotion from text. Following are the steps:

1. Take input from the subtitle dataset.
2. Google's T5 Algorithm fine-tuned for emotion detection will detect the emotion of the text.
3. Store the text in the dataset for further processing.

B. Profanity Detection

For identification of profanity in a movie the subtitles are compared against an already created profanity word-bank. It detects all words that are present in the word-bank and that occurred in the subtitles to give us the frequency of profanity words.

V. IMPLEMENTATION

Step 1: Input Subtitle Dataset

Collecting subtitles of movies from various genre and to feed it for further processing. A standard subtitle file is downloaded. A standard subtitle file looks as follows:

1 00:00:51,052 --> 00:00:54,139 <i>This is Free City.</i>
2 00:00:55,807 --> 00:00:59,477 <i>Look at this guy. He's one of the sunglasses people.</i>
3 00:00:59,644 --> 00:01:02,731 <i>And the people who wear sunglasses are heroes.</i>
4 00:01:09,404 --> 00:01:12,824 <i>They have a devil-may-care attitude and they run this town.</i>

Step 2: **Data Preprocessing**

Data preprocessing task includes parsing the subtitle file to create a readable pandas dataframe with the text, start_time and end_time of every line. This timestamp can later be used for more detailed insights, however for the purpose of the current study we will not be using the timestamp. After preprocessing of the data, it looks as follows:

	0	1	2
0	This is Free City.	00:00:51.052000	00:00:54.139000
1	Look at this guy. He's one of the sunglasses p...	00:00:55.807000	00:00:59.477000
2	And the people who wear sunglasses are heroes.	00:00:59.644000	00:01:02.731000
3	They have a devil-may-care attitude and they r...	00:01:09.404000	00:01:12.824000
4	You are so hot.	00:01:16.244000	00:01:17.454000

Table 1. Processed Subtitles of a Movie with text and timestamp

Step 3: **Profanity Detection**

In this step the preprocessed subtitle data is compared against an already created word-bank. This then tells us about the number of profanity words in a movie.

Step 4: **Emotion Detection**

The Google T5 algorithm returns emotion for each subtitle line that is entered. And a dataset is generated from the detected emotions as shown below:

	text	start_time	end_time	emotion
0	This is Free City.	00:00:51.052000	00:00:54.139000	joy
1	Look at this guy. He's one of the sunglasses p...	00:00:55.807000	00:00:59.477000	joy
2	And the people who wear sunglasses are heroes.	00:00:59.644000	00:01:02.731000	joy
3	They have a devil-may-care attitude and they r...	00:01:09.404000	00:01:12.824000	anger
4	You are so hot.	00:01:16.244000	00:01:17.454000	love

Table 1. Subtitles of movie with their emotions

Step 5: **Emotion Distribution**

Once the emotions of each line of the movie's subtitle are detected we generate an emotion distribution chart to find the emotional trend of the movie. The following bar chart shows the distributions of emotions across the movie.
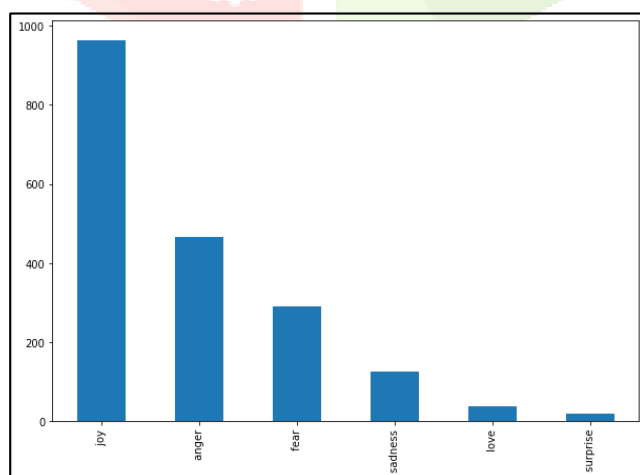

Fig. 2. Emotion distribution chart of all the

Step 6: **Genre Tagging**

Once the emotion distribution chart is ready, we tag that movie with it's respective genre, therefore creating a emotion to genre relationship. This process of tagging genre is carried out for multiple movies from multiple genres to build a dataset for training the model.

Step 7: **Model Training**

Final step is to train a model on a dataset that is created by the above process. For this purpose we will use multiple models such as Random Forest Regression, SVM or even CNN.

VI. CONCLUSION AND FUTURE SCOPE

Movies have been an integral part of our entertainment. Classifying their genre based only on the subtitles is a challenging task as most of the audio-visual details that support a dialogue are not present[2]. Our project is aimed at accurately identifying the movie's genre based on the subtitles by extracting emotions of lines. There are many systems present that make use of video or audio to identify a movie's genre whereas we make use of text-based emotion detection methods. Along with detecting the emotions we also identify any profanity present in the movie. The gained outcome from our project is able to detect emotion distribution of a movie and to subsequently identify its genre. Along with it, its also able to detect profanity present in the movie.

This model can be further expanded to make use of the present timestamp to more accurately find the time frame for which the movie had a profound emotion, which will then subsequently give temporal distribution of emotions. Along with it the profanity detection can be used to create a profanity rating system that is able to rate movies based on the profanity content present inside them.

References

[1] Kao, Edward & Chieh Chun, Liu & Yang, Ting-Hao & Hsieh, Chang-Tai & Soo, Von-Wun. (2009). Towards Text-based Emotion Detection: A Survey and Possible Improvements. Proceedings - 2009 International Conference on Information Management and Engineering, ICIME 2009. 70 - 74. 10.1109/ICIME.2009.113.

[2] Santosh Kumar Bharti, S Varadhaganapathy, Rajeev Kumar Gupta, Prashant Kumar Shukla, Mohamed Bouye, Simon Karanja Hingaa, Amena Mahmoud, "Text-Based Emotion Recognition Using Deep Learning Approach", Computational Intelligence and Neuroscience, vol. 2022, Article ID 2645381, 8 pages, 2022. https://doi.org/10.1155/2022/2645381

[3] Q. Hoang, "Predicting Movie Genres Based on Plot Summaries," arXiv:1801.04813 [cs, stat], Jan. 2018, arXiv: 1801.04813. [Online]. Available: http://arxiv.org/abs/1801.04813