Comparison of Suicidal Content Analysis Algorithm

Dr. S.Manju[1], S.Dheerej[2]
[1]Associate Professor, Dept of MCA
[2]Dept of MCA
[1, 2] PSG College of Arts and Science, Coimbatore, Tamilnadu, India

Abstract- *Virtual entertainment stages have changed conventional specialized strategies by permitting clients around the world to impart immediately, transparently, and often. Individuals utilize virtual entertainment to offer their viewpoint and share their own accounts and battles. Pessimistic sentiments that express difficulty, considerations of death, what's more, self-hurt are boundless in virtual entertainment, particularly among youthful ages. Consequently, utilizing web-based entertainment to distinguish and recognize self-destructive ideation will assist with giving legitimate intercession that will ultimately discourage others from self-hurting and ending it all and forestall the spread of self-destructive ideations via virtual entertainment. Many examinations have been done to recognize self-destructive ideation and ways of behaving in friendly media. This paper presents a complete outline of momentum research endeavours to identify self-destructive ideation utilizing AI calculations via online entertainment. This survey of examinations exploring the possibility of online entertainment use for self-destructive ideation identification is planned to work with additional exploration in the field and will be a useful asset for specialists participating in self-destructive text grouping.*

Keywords- suicide ideation,word embedding, machine learning; deep learning, Text classification,Social media

I. INTRODUCTION

A great many people routinely utilize web-based entertainment, for example, discussion boards, writing for blog sites, and social organizing stages, with 3.96 billion individuals effectively using the web[1]. Facebook, Twitter, Snapchat, and other virtual entertainment organizing locales permit clients to share material and cooperate with others.Numerous clients like to use virtual entertainment organizations to share their contemplations and feelings, and their day to day encounters, issues, and issues. Self-destructive ideation, demise, and self-hurting considerations are among the most generally examined topics via virtual entertainment.Even though people who are aware that they are experiencing self-destructive considerations may not receive the appropriate treatment for a variety of reasons, it is critical to identify individuals who are contemplating suicide or who may have suicidal thoughts through their tweets or blogs. Early detection of suicidal individuals has the potential to save many lives. Consequently, utilizing a suicidal detection system may be beneficial to numerous individuals and have a significant impact on their treatment.

The examinations talked about in this paper took a gander at online entertainment content to consequently identify self-destructive ideation and conduct. This article intends to give a complete outline of continuous examination via virtual entertainment stages that utilize AI to distinguish self-destructive ideation.

Based on their use, a few tasks and datasets are discussed and summarized. Researchers who want to develop applications that make use of text classification techniques or suicidal text classification should read this article. Additionally, to aid future field research and investigate the viability of utilizing social media to detect suicidal ideation. In this study, the terms "suicidal ideation," "suicidal thoughts," and "suicidality" will be used interchangeably.

II. METHODOLOGIES IN SUICIDAL CONTENT ANALYSIS

A. TEXT PREPROCESSING / FEATURES

Text pre-processing refers to the process of cleaning and transforming raw text data before it is used for further analysis or natural language processing (NLP) tasks. we extracted several features including Latent Dirichlet Allocation (LDA), Tri-gram,Temporal features, Emoticons, Sentiment Analysis, Tf-Idf, Bi-Gram, Textual Features, Linguistic Features (LIWC), Syntactic Features (POS), Statistical Features, GloVe, Bag of Words.
Textual features refer to specific attributes or characteristics of text data that are used for text analysis tasks such as text classification, sentiment analysis, and topic modelling.

LATENT DIRICHLET ALLOCATION (LDA):

A statistical model called Latent Dirichlet Allocation (LDA) is used in topic modelling, which is a way to automatically identify topics in a collection of documents. The model assumes that each topic is a probability distribution over words and that each document is a mixture of a small number of topics. As a generative model, LDA attempts to infer the hidden variables (the topics and their distribution in each document) that are responsible for generating the observed data (the words in the documents) by assuming a process by which documents are created.

TRI-GRAM:

A trigram, also known as a 3-gram, is a contiguous sequence of three tokens (usually words) in a text document. In natural language processing, trigrams are commonly used to extract features from text for various tasks such as language modelling, information retrieval, and sentiment analysis.

BI-GRAM:

A bigram, also known as a 2-gram, is a sequence of two adjacent tokens (usually words) in a text document. In natural language processing, bigrams are commonly used to extract features from text for various tasks such as language modelling, information retrieval, and sentiment analysis.

EMOJIS/EMOTICONS SENTIMENT:

Users use emojis to express their feelings through simple images and nonverbal elements. Emojis can help us distinguish between content with positive and negative sentiments and understand the emotion of a text or tweet. Numerous brand-new emoticons that can be categorized as either positive, negative, or neutral are typically included in user-tweeted messages.

SENTIMENT ANALYSIS:

Feeling investigation is a characteristic language handling strategy used to consequently distinguish and separate abstract data, like sentiments and feelings, from message information. The objective of feeling investigation is to decide if a piece of message communicates a good, pessimistic, or nonpartisan opinion towards a particular point or element. It involves a number of steps, including feature extraction, sentiment classification, and text pre-processing. Pre-processing the text data by removing stop words, punctuation, and other noise that doesn't help convey emotion is the first step. The next step is to extract pertinent features from the text, such as the use of intensifiers, the frequency of positive and negative words, and the presence of emoticons or emojis.

TF-IDF:

A statistical measure called term frequency-inverse document frequency (TF-IDF) is used in text mining and information retrieval to determine how important a word is in relation to an entire corpus of documents. The number of times a term (word) appears in a document is called the term frequency (TF), and the inverse document frequency (IDF) is a measure of the term's rarity across the corpus of documents. A weight that indicates a term's importance in a document is the product of these two values, TF-IDF.

LINGUISTIC FEATURES (LIWC):

LIWC stands for "Linguistic Inquiry and Word Count", and it is a software program that analyses written or spoken language in order to identify various linguistic features. LIWC is based on a set of language categories that have been shown to be relevant to understanding psychological processes and personality traits, which includes Personal Pronouns, Positive and negative emotion, cognitive process, social process, articles, function words, punctuation.

SYNTACTIC FEATURES (POS):

Syntactic features are useful information in natural language processing tasks. For our self-destructive ideation recognition model, we removed grammatical features (POS) [2] as elements to catch the comparable linguistic properties in client posts. Things, action words, participles, articles, pronouns, intensifiers, and conjunctions are all included in typical POS labels. Additionally, POS subgroups were identified as providing additional insight into the posts' syntactic properties. After each post had been parsed and tagged, the number of each category in the title and body text was simply counted.

STATISTICAL FEATURES:

The length of client created posts differs, and texts contain a few measurable highlights. Short, straightforward sentences are used in some posts, while lengthy paragraphs and intricate sentences are used in others.

The following statistical characteristics were gathered following segmentation and tokenization:

I. The title's total number of words, symbols, and characters
II. The total number of words, symbols, and characters in the sentences and paragraphs that make up the text body.

GLOBAL VECTOR(GloVe):

GloVe (Global Vectors) is a word embedding technique that was developed by researchers at Stanford University. Like Word2Vec, GloVe is a neural network-based approach that generates vector representations of words in a high-dimensional space. It is a count-based word embedding method that leverages the co-occurrence statistics of words to create dense word embeddings. GloVe embeddings have been shown to capture both syntactic and semantic relationships between words. However, GloVe uses a different training objective and weighting scheme compared to Word2Vec.

BAG OF WORDS:

Bag of Words (BoW) is a simple technique used for text representation in natural language processing. The approach involves representing text data as a bag, or a set, of its words, ignoring their order and syntax. This means that the words in a document are counted and then converted into numerical values that can be used for various text analysis tasks.The BoW technique involves tokenization, counting, and vectorization.

B. CLASSIFICATION MODELS

Suicidal content analysis is a complex problem that requires careful consideration of the language used, the context, and the potential consequences of false positive or false negative predictions. In general, it's important to approach this problem with caution and to seek the input of mental health professionals.

The accuracy of a classification model refers to the proportion of correct predictions that the model makes on a dataset. The accuracy of a model is often measured using metrics such as accuracy and F1 score. Some models may achieve high accuracy on one dataset but perform poorly on another, so it's important to evaluate the performance of the model on a held-out test set in order to get a more accurate estimate of its accuracy.

LOGISTIC REGRESSION:

Logistic Regression is a well-known AI calculation utilized for parallel grouping issues. It belongs to the family of supervised learning algorithms, where the goal is to predict a binary output variable based on one or more input variables, also known as features.The algorithm works by modelling the probability of the output variable belonging to the positive class given the input features. This is achieved by fitting a logistic function to the input data using a set of training examples with known output labels.

Random Forest:

Random Forest is a popular machine learning algorithm used for classification and regression tasks. A troupe technique joins numerous choice trees to make expectations. In an irregular backwoods, numerous choice trees are made utilizing various subsets of the preparation information and highlights. This random sampling of data and features helps to reduce overfitting and improve the accuracy of the model.

Support Vector Machine (SVM):

An artificial intelligence calculation known as Support Vector Machine (SVM) is used to examine order and relapse. It is a supervised learning algorithm that determines the optimal decision boundary between the various data classes.A kernel function is used to map the data points in SVM to a higher-dimensional space. The best hyperplane for separating the various data classes can be found with the assistance of this transformation. As the decision boundary, the hyperplane with the greatest margin between classes is chosen.

XGBoost:

Regression and classification problems are common applications of the popular machine learning algorithm known as XGBoost (Extreme Gradient Boosting). It is an outfit learning strategy that consolidates numerous frail models to make a more grounded model. XGBoost works by training decision trees iteratively to fix mistakes made by previous models. The calculation allots higher loads to the misclassified relevant pieces of information and lower loads to the accurately arranged data of interest, which is known as slope supporting.

In addition to gradient boosting, XGBoost includes several key features that improve its performance, such as regularization, handling missing values, and parallel processing. The algorithm also uses a customized loss function that balances the trade-off between model accuracy and computational efficiency.

Multilayer Perceptron (MLP):

A type of artificial neural network known as the Multilayer Perceptron (MLP) is made up of multiple layers of connected nodes or neurons. For classification and regression tasks, it is a well-known and effective machine learning algorithm.An input layer, one or more hidden layers, and an output layer make up the MLP.Each layer consists of a set of neurons that perform computations on the input signals and pass the output to the next layer. The connections between the neurons are defined by a set of weights, which are learned during the training process using a method such as backpropagation.

The neurons in the MLP's hidden layers introduce non-linearity into the model by employing a non-linear activation function like the Rectified Linear Unit (ReLU) function or the sigmoid function. The MLP can model intricate connections between the input features and the output variable thanks to the activation function.

Naive Bayes:

Credulous Bayes is utilized for arrangement errands which depends on Bayes' hypothesis, which depicts the likelihood of an occasion happening given earlier information or proof.

Because Naive Bayes assumes that the input features are independent of one another, the probability of another feature being present does not change if one feature is present or not. The algorithm can function effectively even with high-dimensional data thanks to this assumption, which makes the calculations simpler.

Using Bayes' theorem, the probability of each class is calculated using the input features in the Naive Bayes algorithm:

P(C) is the prior probability of class C, P(xi|C) is the probability of feature xi given class C, P(xi|C) is the probability of feature xi given class C, and P(x1, x2,..., xn) is the probability of the input features. P(C|x1, x2,..., xn) = P(C) * P(x1|C) * P(x2|C) *...

III. COMPARATIVE ANALYSIS

S.no	Year	Author	TEXT PREPROCESSING / FEATURES	Proposed Algorithm	Result	
					F1 score	Accuracy
1	2022	Moumita Chatterjee, Piyush Kumar, Poulomi Samanta, Dhrubasish Sarkar[3]	LDA + Trigram + TF-IDF + Tweet Statistics + Temporal + Emoticons + Sentiment Analysis	Logistic Regression	0.81	87%
			LDA + Trigram + TF-IDF + Sentiment Analysis + Emoticons	Logistic Regression	0.83	85%
			LDA	Logistic Regression	0.85	86%
			Trigram + TF-IDF	Random Forest (RF)	0.75	77%
			Temporal features	Support Vector Machine (SVM)	0.75	74%
			Sentiment Analysis	XGBoost	0.74	75%
2	2018	Yevhen Tyshchenko[4]	GloVe	Convolutional Neural Network (CNN)	0.84	78%
			Bag of Words	Support Vector Machine (SVM)	0.82	81%
			Bag of Words	Random Forest (RF)	0.8	79%
3	2019	Michael M. Tadesse, Hongfei Lin , Bo Xu , And Liang Yang[5]	Bi-Gram	Support Vector Machine (SVM)	0.8	80%
			LIWC+LDA+bigram	Multilayer Perceptron (MLP)	0.93	91%
4	2018	Shaoxiong Ji, Celina Ping Yu, Sai-fu Fung, Shirui Pan, and Guodong Long[6]	Statistics + topic + TF-IDF + POS + LIWC	Random Forest (RF)	0.96	96%
			Statistics + topic + TF-IDF + POS + LIWC	XGBoost	0.95	95%

5	2022	Theyazn H. H. Aldhyani, Saleh Nagi Alsubari, Ali Saleh Alshebami , Hasan Alkahtani and Zeyad A. T.Ahmed[7]	Textual Features	CNN–BiLSTM	0.95	95%
			Textual Features	XGBoost	0.91	91.30 %
			LIWC	CNN–BiLSTM	0.84	84.50 %
			LIWC	XGBoost	0.86	86.90 %
6	2019	Michael Mesfin Tadesse, Hongfei Lin, Bo Xu and Liang Yang[8]	Word2vec	LSTM-CNN	0.93	93.80 %
			Statistics + TF-IDF + Bag of Words	XGBoost	0.83	88.30 %
			Statistics + TF-IDF + Bag of Words	Naive Bayes	0.81	82.50 %

IV. CONCLUSION

The utilization of virtual entertainment stages to communicate encounters and sentiments has opened new roads for breaking down and diagnosing self-destructive ideation and other mental problems. Social media early detection of suicidal ideation reduces suicide, provides an automatic and comprehensive screening for suicidal tendencies, and stops the spread of suicidal content. This overview investigates existing strategies for identifying self-destructive ideation in web-based entertainment utilizing AI techniques. The usefulness and practicality of utilizing social media platforms like Twitter, Reddit, and Weibo to detect suicidal ideation have been demonstrated by a significant amount of research. Most of the research has focused on techniques for detecting suicidal ideation in widely spoken languages like English, with Indian receiving less attention. More research is needed to find suicidal ideation in India because of the growing number of people who use social media.

REFERENCES

[1] G. Astoveza, R. J. P. Obias, R. J. L. Palcon, R. L. Rodriguez, B. S. Fabito, and M. V. Octaviano, “Suicidal behavior detection on twitter using neural network,” in TENCON 2018 - 2018 IEEE Region 10 Conference, 2018, pp. 0657–0662.

[2] A. Voutilainen, “Part-of-speech tagging,” in The Oxford Handbook of Computational Linguistics, pp. 219–232, Oxford University Press, 2003.

[3] Moumita Chatterjee A, Piyush Kumar B, Poulomi Samanta B , Dhrubasish Sarkar B, “Suicide ideation detection from online social media: A multi-modal feature based technique”, International Journal of Information Management Data Insights 2(2) (2022) 100103

[4] Yevhen Tyshchenko” Depression and anxiety detection from blog posts data”, Nature Precis. Sci., Inst. Computer Sci., Univ. Tartu, Tartu, Estonia in 2018

[5] Michael M. Tadesse, Hongfei Lin, Bo Xu, And Liang Yan, “Detection of Depression-Related Posts in Reddit Social Media Forum”,IEEE Access, (2019), 44883-44893, 7.

[6] Shaoxiong Ji, Celina Ping Yu, Sai-fu Fung, Shirui Pan, Guodong Long,” Supervised learning for suicidal ideation detection in online user content”,Complexity, (2018), 2018.

[7] Theyazn H. H. Aldhyani, Saleh Nagi Alsubari, Ali Saleh Alshebami, Hasan Alkahtani and Zeyad A. T. Ahmed,” Detecting and Analyzing Suicidal Ideation on Social Media Using Deep Learning and Machine Learning Models”,International Journal of Environmental Research and Public Health, (2022), 19(19).

[8] Michael Mesfin Tadesse, Hongfei Lin, Bo Xu and Liang Yang, “Detection of Suicide Ideation in Social Media Forums Using Deep Learning”, Algorithms, (2020), 13(1).