Advanced Clickbait Detection System Using Cognitive Evidences and Ensemble Learning

Atharva Chandwadkar[1] Prof. Neeta Deshpande[2]
[1,2]Department of Computer Engineering
[1,2]Gokhale Education Society's R. H. Sapat College of Engineering, Management Studies and Research, India

Abstract— Clickbait is a form of fake content that is deliberately designed to attract the user's attention and be intriguing enough to follow links to read, view, and listen to attached content. The purpose of the trailer is to harness curiosity by providing information in short statements. However, the claims made are not enough to click on linked content to invite users to relevant pages, play with human psychology, and satisfy their curiosity without compromising the user experience. To address this issue, we have developed a Clickbait Video Detector (CVD) scheme.

Keywords: Clickbait Detection, Cognitive Evidence, Dataset, User Profiling, YouTube

I. Introduction

A. Related Technologies

1.1 Multi-layer classifier framework consisting of six basic models (K-nearest neighbour method, support vector machine, XGBoost, naïve bays, logistic regression, multi-layer perceptron) and a meta-classifier (random forest). It has been. The developed clickbait detection model achieved 92.89% accuracy on the new BollyBAIT dataset and 95.38% accuracy on the deceptive video dataset (MVD)[1]. In addition, the above classifier does not use meta-characteristics or other statistics that depend on the user's interaction with the video (likes, followers, number of comments, etc.) for classification, so potential clickbait video. Can be used to rank before them uploaded.

Clickbait is a form of fake content that is deliberately designed to attract the user's attention and be intriguing enough to follow links to read, view, and listen to attached content. The purpose of the trailer is to harness curiosity by providing information in short statements. However, the claims made are not enough to click on linked content to invite users to relevant pages, play with human psychology, and satisfy their curiosity without compromising the user experience. To address this issue, we have developed a Clickbait Video Detector (CVD) scheme [2].

In an era of instant satisfaction, people communicate with each other primarily through social media platforms. Social media platforms such as Twitter, Facebook, and YouTube make it easy to post user feedback, exposing many misleading and unreliable multimedia content that is widely distributed through popular social media platforms. Among all other platforms, YouTube is one of the leaders in video sharing, with billions of users covering nearly one-third of the Internet population and billions of views per day. I'm plagued by clickbait videos that have achieved and faithfully represent what I'm referring to Clickbait is intentionally designed to get the user's attention, follow links, and be interested in viewing, viewing, or listening to attached content. In 1994, George Lowenstein defined Clickbait as a "data-deficient curiosity theory." According to this definition, he described Clickbait as "a knowledge gap theory of curiosity that uses human psychology to trick users into displaying content that is not faithful to what they claim and ultimately degrades the user experience. Was defined. "Non-clickbait can be defined as content that presents important news and faithfully conveys the same photo of the alleged content to the viewer." This article details the clickbait video detection mechanism on online social media platforms. Clickbait video detection is a smart job because it automatically analyses the video content using video clickbait detection frameworks/tools/plugins, similarly in the future it can also be used as a smart warning system that can help automatically report the credibility of the video. content to users. Figures 1 and 2 show an instance of clickbait and a video without clickbait. Since clickbait detection and prevention in videos may be a largely unexplored area, the first objective of the research conducted is to develop a clickbait prevention and detection model (CPDM) for YouTube videos that support only (i) video content, i.e. video and associated audio, (ii) Video Title, and (iii) Thumbnail. The title and thumbnail of the video are used to help the viewer decide whether to click on the video and ideally should accurately represent the content of the video. The model will determine if they need to be exaggerated, manipulated and engineered to entice users to click on them, or if they are "clickbait". Besides, we aim to utilize the information and features extracted from the descriptors and the video to build a model that can determine if the latter fulfils the promise of content suggested by the former. This contrasts with the previously developed models for clickbait detection on YouTube which are heavily dependent on meta-features such as comments, number of likes etc. for the video accumulated over a considerable period of time

A defining characteristic of clickbait is misrepresentation in the enticement presented to the user to manipulate them to click onto a link. While there is no universally agreed-upon definition of clickbait, Merriam-Webster defines clickbait as "something designed to make readers want to click on a hyperlink, especially when the link leads to content of dubious value or interest." Dictionary.com states that clickbait is "a sensationalized headline or piece of text on the Internet designed to entice people to follow a link to an article on another web page." BuzzFeed editor Ben Smith states that his publication doesn't do clickbait, using a strict definition of clickbait as a headline that is dishonest about the content of the article. Smith notes that Buzzfeed headlines such as "A 5-Year-Old Girl Raised Enough Money To Take Her Father Who Has Terminal Cancer To Disney World" deliver exactly what the headline promises. The fact that the headline is written to be eye-catching is irrelevant in Smith's view since the headline accurately describes the article. Facebook, while trying to reduce the amount of clickbait shown to users, defined the term as a headline that encourages users to click, but doesn't tell them what they will see. However, this definition excludes a lot of content that is commonly considered clickbait. A more general definition of is a headline that is intentionally over-promised and missing. Articles related to such headlines are often not original, either simply repeating the headline or copying the content from a more authentic news source. The term Clickbait is sometimes used in articles that are offensive to people. In such cases, the article is not clickbait in terms of a legitimate definition of the term. If you have two videos with similar titles but different content, the speech is converted to text and then compared to the title to see how faithfully the speech represents the title. At the same time, we also addressed the issue when commenting the uploader does not allow comments from viewers, so the base functionality is not fetched. In this case, relying solely on a particular feature set is not effective. We have introduced another tip to deal with this case. H. A reliable source of information that allows you to predict the reliability of your video. A detailed description is given in a later section. The proposed work makes the following major contributions to the Clickbait detection task: & the proposed work will greatly contribute to providing a new methodology for capturing and annotating various categories of YouTube videos. Since there are very few datasets published in this area, we will create a self-generated dataset [MVD] for Clickbait and Non-Clickbait videos. & As far as we know, our work consists of providing three sets of cognitive evidence for Clickbait's detection task, specifying the desired action required for each specific possible detection case. & This treatise initially contains a measure of audio title similarity that was not included in previous studies. Some of the previous studies use video thumbnails and titles that are not effective if the two videos [Clickbait and Non-Clickbait have the same title and probably have the same thumbnail. In that case, it is important to do clickbait. It's discriminatory, and non-Clickbait video comments support the claim. Title Figure 2 Example other than Clickbait 4216D[3]. Vashney and D.K. Vishwakarma Clickbait eins Therefore, voice is also analyzed to address this issue. This presents an important sign of discrimination. & In some scenarios, the uploader does not allow comments from viewers. In this case, the comment part does not contribute and some important functionality for reliable prediction is lost. To address this issue, we have also introduced other evidence to collect important clues from trusted web links when there is no user response. & Examining the performance of the model using different classifiers, comparative analysis showed that the proposed method was superior to other prior arts in the same dataset.

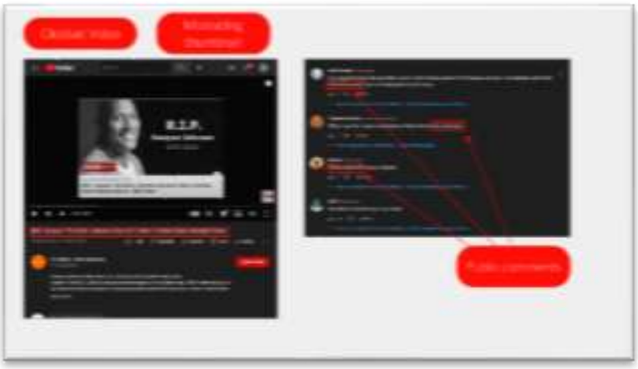

Fig. 1: Clickbait video from YouTube

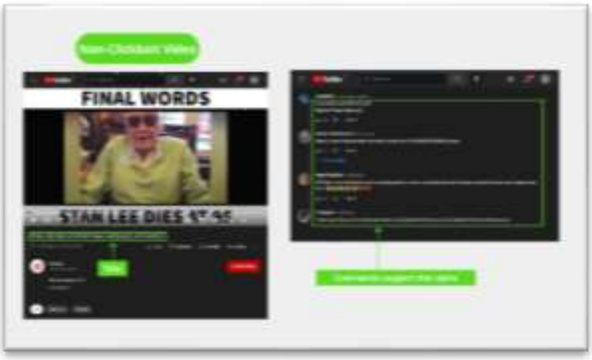

Fig. 2: Non - clickbait video from Youtube

YouTube videos often include eye-catching descriptions and interesting thumbnails designed to increase views and thereby increase revenue for the person who posted the video. This creates an incentive for people to post clickbait videos where the content may be significantly different from the title, description or thumbnail. In fact, users are tricked into clicking on clickbait videos. In this research, we consider the challenging problem of YouTube video clickbait detection. We experiment with several state-of-the-art machine learning techniques using various text functions.

II. LITERATURE REVIEW

As the name suggests, image forgery detection is all about trying to detect malicious agents information that is conveyed through images. In 2018, Zhang et al. [1] developed a "fauxography" detector that could detect images that are misleading on social media platforms.

Palod et al. [3] passed pre-trained Word2Vec comment embedding through an The LSTM network generated a "fake" vector and achieved an F-score of 0.82.

Shang et al. [4] proposed a model that included network feature extraction, metadata feature extraction, and language feature extraction for clickbait detection in Youtube videos. Network feature extraction used the comments in the videos and extracted semantic features. They relied on the extraction of linguistic features about embedding documents for comments using Doc2Vec and also used a metadata module. In 2019, Reddy et al. [4] implemented a model using word embedding and trained on a support vector machine (SVM). In [4] Dong et al. have proposed an "attentive model with deep similarity awareness" that focuses on the relationship between titles that are misleading and target content. This method was quite different from traditional feature engineering and seemed to work reasonably well. In [8] Setlur considered a partially supervised trust network together with a an attention-based gated network. Based on a small labeled data set, this method provided promising results. In many of the above approaches, only the textual information given in is used name and description along with metadata properties were considered consideration during model training. An exception is the work in [25], where the authors also used comments to extract features. This is also worth noting all used Word2Vec, BERT and Doc2Vec embedding layers the aforementioned implementations. In this research, we experiment with multiple embedding layers, including BERT, DistilBERT and Word2Vec. BERT has been shown to be effective in previous research due to its deep contextualizing nature [14]. Combination of multiple models, known as ensemble learning, produced interesting results in [28] and we also consider ensemble models in the form of random forest classifiers.

A. Natural Language Processing

Natural Language Processing (NLP) is a machine function for processing and understanding human speech. It is used to solve many real-world problems, such as: B. Machine translation, answering questions, predicting words. Figure 2 shows the timeline of recent advances in NLP. In the early 1990s, I trained NLP algorithms using a statistical and probabilistic approach. However, with the advent of the Internet, the amount of data has increased significantly, and such algorithms have become inadequate. In 2001, Bengio etc. I experimented with a feedforward neural network. Later, recurrent neural networks (RNNs) and long short-term memory (LSTM) models were introduced. As of 2012, techniques such as Latent Semantic Indexing (LSI), Latent Semantics Analysis (LSA), and Support Vector Machine (SVM) have become widespread in the NLP field. Part-of-sale (POS) tagging is a commonly used approach.2013, Tomaset.Al Introduced Word2Vec, which is used to generate vector representations of words. These embeddings are taken from relatively simple neural network weights, and vectors can capture important semantic information based on the cosine distance between Word2Vec embeddings. Launched in 2014, the Global Vector for Word Representation (GloVe) is an attempt to try the best combination of LSA, LSI, and Word2Vec. It is based on the appearance of words throughout the corpus. CNNs and LSTMs are becoming more popular for NLP-related tasks. This is because such models can effectively capture sequential information. LSTMs are highly specialized types of RNNs that mitigate the gradient problems that occur with plain vanilla RNNs. The Gated Recurrent Unit (GRU) is a variant of the LSTM introduced in 2014 that is lighter because it requires less parameters to be trained. Sutskeveretal. proposed a sequence-by-sequence learning approach using an encoder-decoder architecture. In fact, such an encoder-decoder model seems to be the leading language modelling framework for today's NLP tasks. The concept of attention mechanism was developed by Bandana et al. In 2015, we overcame the fixed vector length limitation of the input set of the inter-sequence model. Attention provides information about the importance of phrases during the decision-making process. The formula for the sigmoid function:

$$\sigma(\omega^{T}x + b) = \frac{1}{1 + e^{-(\omega^{T}x+b)}}$$

This gives a value in the range 0 to 1 and can be interpreted as a probability. Clickbait detection problems can be treated as a kind of binomial logistic regression with output of either 0 or 1.

B. Random Forest

Random Forest is based on a simple decision tree. A large group of decision trees work together as an ensemble. Each tree is trained with a subset of data and functions. This is a process called bootstrap aggregation or bagging. When bagging, the data in each tree is randomly selected on your behalf. The final prediction of Random Forest can be obtained via a simple voting scheme. The random overall structure alleviates the tendency of individual decision trees to over fit training data. The important hyper parameters of Random Forest are n-estimator, n-job, maximum function, minimum sample leaf. The n-estimator's parameter represents the number of trees to be built. Adding a tree usually comes at the expense of computational time, but improves performance. The max-features parameter is the number of features required to split at a particular node. The n-jobs parameter is the number of processors running in parallel.

C. Multilayer Perceptron

Multilayer Perceptron (MLP) is a basic type of feedforward neural network that contains an input and output layers and at least one hidden layer. The output layer of the MLP can be used for prediction or classification. Next, I will briefly explain the regularization and activation functions. For more information on this topic and related topics, see. Neural network models tend to be over fitted. Overfitting models are very effective in classifying training data, but they are less accurate in predicting test data. In fact, the model "remembers" the training data rather than learning from it. A convenient way to avoid overfitting is to use dropouts. The dropout ignores a certain number of nodes during the various training steps. This simple approach activates the atrophied node in the learning process. The activation function is used to determine the output of a node in a neural network. There are several types of activation functions, such as tanh, sigmoid, ReLU, and leaky ReLU. In this study, we experimented with ReLU and Tanh.

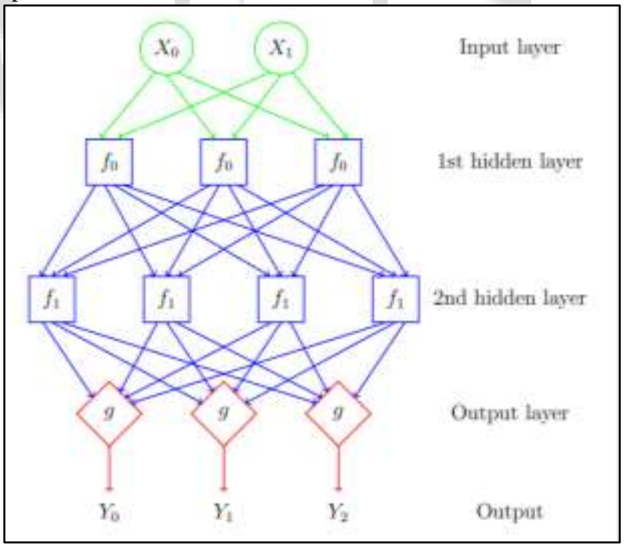

Fig. 3: Multilayer Perceptron

III. PROPOSED METHODOLOGY

Due to the small number of datasets available, one of the main contributions to this task is the creation of datasets. That's why Clickbait and non-Clickbait video datasets were developed by using YouTube REST Data API v3 to collect a variety of videos. Details include video content (title, likes, dislikes, views, etc.), comments, and channel details (subscribers, subscription dates, videos, views). Few datasets are available in the field of deceptive video detection, which leads to the goal of creating a generalized dataset that contains different categories. We have collected 987 videos (474 Clickbait and 513 Non-Clickbait) from a list of 16 most popular video categories 4 defined by YouTube. To collect clickbait videos.

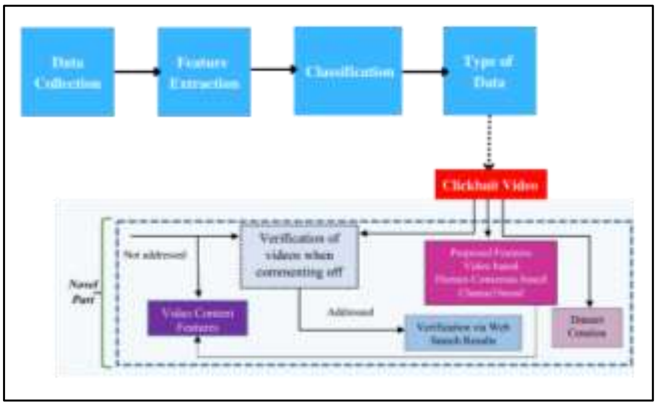

Fig. 4: Proposed architecture click bait classification

I manually crawled and annotated each of the 474 videos. We analysed several channels that posted Clickbait prominently and their follow-up channels to encourage users to access the video. Some of the channels 5 you post are interested in allowing users to access the links to enhance the impression of your video. YouTube has good checking algorithms and blocking mechanisms to detect malicious videos, and most videos are still active and will not be deleted. The study analyzed that most clickbaits and hoaxes were posted about the death of a celebrity, which was later false and degraded the user experience. Without proper validation, these published messages also suppress general sentiment. For this reason, Clickbait video detection is one of the most important areas of research. We use a viral video strategy to reduce the time it takes to collect fake videos.

SR.NO	Category	No.	CLASS
1.	Auto & Vehicles	5	Clickbait
2.	Auto & Vehicles	4	Non- Clickbait
3.	Comedy	12	Clickbait
4.	Comedy	43	Non- Clickbait
5.	Education	6	Clickbait
6.	Education	2	Non- Clickbait
7.	Entertainment	241	Clickbait
8.	Entertainment	227	Non- Clickbait
9.	Film & Animation	28	Clickbait
10.	Film & Animation	35	Non- Clickbait
11.	Gaming	8	Clickbait
12.	Gaming	23	Non- Clickbait
13.	Music	6	Clickbait
14.	Music	82	Non- Clickbait
15.	News & Politics	26	Clickbait
16.	News & Politics	1	Non- Clickbait
17.	Non- Profit & Activism	6	Clickbait
18.	People & Blogs	132	Non- Clickbait
19.	People & Blogs	6	Clickbait
20.	Pets & Animals	1	Non- Clickbait
21.	Pets & Animals	3	Clickbait
22.	Science & Technology	61	Non- Clickbait
23.	Science & Technology	2	Clickbait
24.	Sports	22	Non- Clickbait
25.	Sports	4	Clickbait
26	Travel & Events	4	Non- Clickbait
		474 Clickbait	
		513 non-Clickbait	

4 Cognitive Evidence of Clickbait Detection This section provides a problem definition for the Clickbait detection task and describes the Clickbait Video Detector, a proposed method (CVD) to address the problem. Defines a clickbait detection task for a particular video set. The system needs to determine which videos report clickbait and do not faithfully represent the events associated with them. The ultimate goal of flagging Clickbait videos is to warn users that the video content does not meet the claims it represents and help prevent the spread of false content. This document considered the following three detection cases listed in Table 2 below. Some evidence is needed to justify and validate the above case and warn users to believe and disseminate false information and think twice. If these three cases are identified, the video may be clickbait and may not faithfully represent the associated event. Formally, the task takes a series of video IDs VID = VID1, VID2, VID3. Using VIDN as input, the classifier must determine whether each of these video VID is clickbait or non-clickbait by assigning the label Y = {C, R}. Therefore, formulate the task as a binary classification problem. The performance of this problem is analysed and evaluated by calculating various performance measurements such as target class (clickbait) fit, recall, and F1 score. There are three types of cognitive evidence that you have been considered for the detection of clickbait. Each of these evidence gives a significant contribution in finding the clues in predicting a video as clickbait or not.

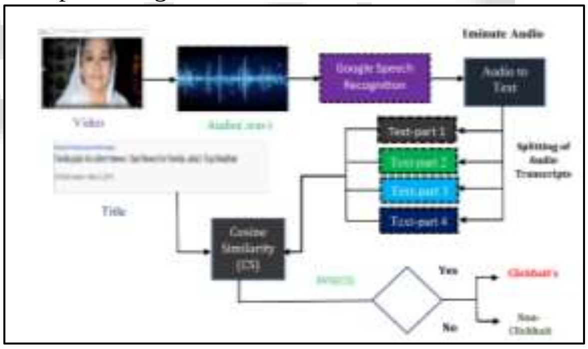

Fig. 5: The process of retrieving evidence 1

Clickbait Video Detection This section describes a proposed method, CVD (Clickbait Video Detector), to address previously formulated issues. This technique consists of three main pieces of evidence obtained using three functional components based on video, human consensus, and user profiles. The first feature component is used to extract video-related features (for example, language title similarity, count, dislike count, etc. Video content-based features: Video content-based features, The first component of the proposed method. This component is responsible for extracting video content-based features such as language title similarity, number of likes, number of dislikes, proportion of dislikes, etc. Language and title similarity is one of the important functions and plays a major role. Role in the preservation of evidence 1. The audio title similarity is similar to the audio text associated with the video title. This identifies whether the particular allegations associated with the video faithfully portray the associated event. We extracted the audio from each video to determine how faithfully the video represented the claim. The Google Speech to Text API is used for the speech part and later converted to text. We applied cosine similarity between the text and title extracted from the audio portion of the video and measured the similarity between them. Google's Speech-to-Text API was used for speech recognition. There are three main ways to perform speech-to-text speech recognition (synchronous, asynchronous, and streaming recognition) 7. We used the synchronous recognition method because we need to process less than a minute of data and the sync recognition request is limited to. Audio data with a duration of 1 minute or less. A detailed description of how the complete process takes place is shown in FIG. This shows Proof1's audio title similarity search process. The first step is to serve the video as input and convert it to audio format. Speech is converted to text format using the Google Speech API. Audio is partially processed to achieve better similarity results. For each video, a 1-minute segment was analysed. This is because within this period you will have enough information to predict that the video is fake or reliable. The 1-minute voice transcript is divided into 4 parts, each 15 seconds. The four texts included in Figure 6 are intended to break a minute of audio transcript into smaller sub pieces to see how similar the titles / claims are to the content. some malicious intent. Figure 7 shows the process of retrieving Evidence 2. Here multilingual content has also been addressed. If any of the content is found to be in a different language, it is translated into English text using google translator, then further be used for analysis. As a result, we have also addressed the multilingual content. A total of 6 Human consensus-based features are extracted. The Human-consensus based features are shown in Table 4. These features are evaluated to represent the statistical characteristics of the responses of the viewers. Experiments: In this study, we experimented with several techniques, including several language modelling techniques. We used Word2Vec, BERT, and Distil BERT to embed words. The architecture of each model is also shown. We used grid search to train and build the model to get the optimal set of parameters. This section briefly describes each model, and the next section shows the results of each experiment in a vector representation of the words in the dataset. The logistic regression model is trained with these embedding's, along with additional features, specifically the number of comments, likes, dislikes, and channel subscriptions. Experiment II: Random Forest with Word2Vec In this experiment, a random forest classifier is trained with Word2Vec embedding. Again, I used the Word2Vec model provided by Gensim. n What are the tested values for the estimator! 10, 20, 30, 50, 100. The set of input functions is the number of comments, the number of likes, the number of dislikes, the title such as the number, the description, and the metadata function. Of a channel subscription. Experiment III: MLP using Word2Vec Again, we use the Word2Vec model provided by Gensim. The embedded title and description are concatenated with the video's metadata attributes and sent to the MLP for classification. The stack size is 10 for 40 epochs. 12 The activation functions used are ReLU and Sigmoid. Figure 11 in the appendix shows the overall architecture of the model. Notice that we use two input embedded layers for the text data (that is, the title and description) and concatenate them. After this step, the output of the high density layer is flattened and concatenated with the input of the metadata function. Finally, the fully connected layers are used for classification. Experiment IV: DropOut, batch normalization, MLP using Word2Vec This experiment is an optimization of the previous experiment. This model uses an additional high density layer with batch normalization and a dropout rate of 0.5. We used a parametrically modified linear unit (PReLU) as the activation function. The stack size will be 10 again at 40 epochs. Figure 12 in the appendix shows the overall architecture of the model. In this model, the output of an embedded layer of text data is concatenated, followed by a fully connected high-density layer, stack normalization, and activation. This output is eventually flattened and concatenated with the metadata feature. 5 Experiment Results (Show Table) Recall that in Experiment I, we used logistic regression with Word2Vec embedding for feature titles and descriptions along with metadata features. In this case, using only the title as input gives 52% accuracy, and using all these features gives 70% accuracy.

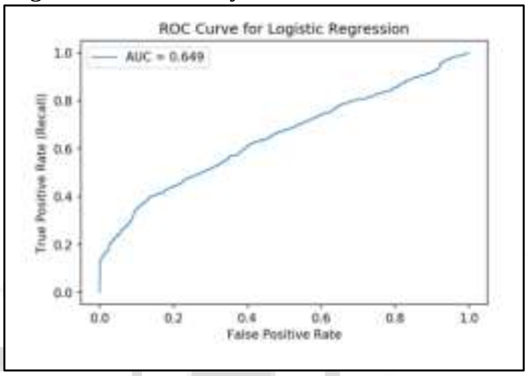

Fig. 6: ROC curve for logistic regression

Experiment II uses a random forest classifier based on title, description, likes, dislikes, comments, and subscriptions. I used Word2Vec embedding for the title and description. I trained this model with multiple input sets. The first set of inputs contains only title and metadata functions. The last input set contained all the features. Not surprisingly, the accuracy improves as features are added. Table 3 shows 80.1% accuracy and accuracy for the model with the first set of input features. H. Titles of favorite and disliked counts and two metadata attributes. Table 4 shows a report of this experiment using the title and metadata as characteristics. The accuracy of this experiment is 92.5%. The report shows the model's fit and recall when classifying Clickbait and non-Clickbait videos. If you classify videos other than Clickbait, your model will perform slightly better. Figure 8 shows the ROC curve for a random forest model. Here, the input characteristics are the title, description, and metadata characteristics.

Class	Precision	Recall	F-Score	Support
NonClickbait	0.81	0.80	0.81	1275
Clickbait	0.80	0.81	0.80	1182
Accuracy	----	-----	0.80	2475
Macro avg	0.80	0.80	0.80	2475
Weight avg	0.80	0.80	0.80	2475

Table 3: Random Forest based on title and likes/dislikes counts

Class	Precision	Recall	F-Score	Support
NonClickbait	0.93	0.93	0.93	1275
Clickbait	0.91	0.92	0.92	1182
Accuracy	----	-----	0.93	2475
Macro avg	0.93	0.93	0.93	2475
Weight avg	0.93	0.93	0.93	2475

Table 4: Random Forest based on title and metadata features

In experiment III a simple MLP is used for classification, based on Word2Vec embeddings for title and description that are concatenated with metadata features. In this case, the test accuracy is observed to fluctuate during the training process, but the best average accuracy achieved is better than 91%. Figure 9 (a) shows the accuracy for this experiment over the 30 training epochs.

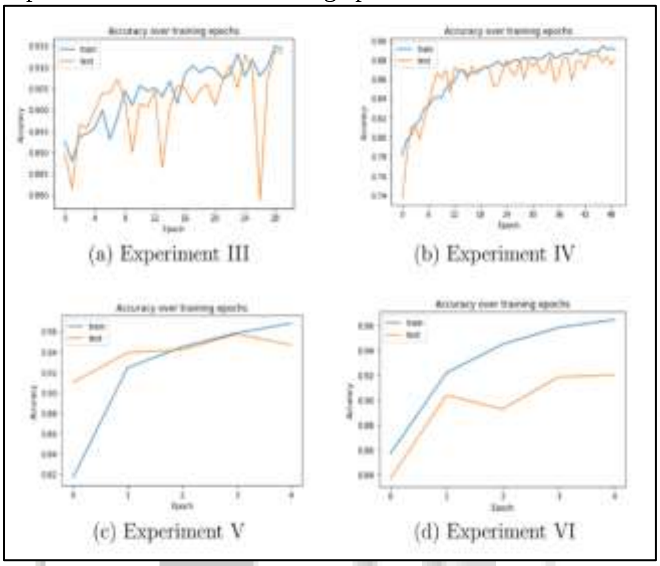

Fig. 7: Accuracy over training epochs for experiments III through VI

Experiment IV uses a modified MLP with PReLU as the batch normalization and activation function. In this case, the training is more stable, but the accuracy is slightly worse than in Experiment III, as shown in Figure 9 (b). Experiment V used a BERT-based transfer learning model for word embedding. This experiment with BERT gives 94.5% accuracy. In this experiment, the length of the input sequence is fixed at 180 characters. Figure 9 (c) shows a plot of the accuracy and training epochs of both the training set and the validation set. Keep in mind that the number of epochs is small compared to the other models under consideration due to the long training time required. In Experimental VI, we used a lighter variant of the BERT model, Distil BERT, for word embedding. The accuracy achieved is about 92%. This model trains much faster than BERT, but the accuracy achieved with BERT is slightly better than Distil BERT. Table 5 shows the fit and recall of Experimental VIs, and Figure 9 (d) shows the fit of training and testing across the epoch. Figure 10 summarizes the results of the six experiments in terms of accuracy (up to two decimal places). Note that the bar chart in Figure 10 uses "MLP Plus" to show the MLP model with dropouts and batch normalization. Also, from left to right, the bars represent Experiments I to VI, respectively.

IV. CONCLUSION AND FUTURE WORK

In this paper, we develop the CVD scheme to detect clickbait video. The scheme leverages on three components for learning three sets of latent features based on User Profiling, Video Content, and Human Consensus that further be used to retrieve three sets of cognitive evidence, as an innovative idea for the detection of clickbait videos on YouTube. The set of features are given as an input to machine learning model and performance has been analyzed by considering all set of features and each feature independently by employing a different set of the classifier, and it has been observed that J48 outperforms all other with an accuracy of 98.89% by applying all set of features using cross-validation technique, while 97.47% using percentage split technique on the self-generated dataset [MVD]. The proposed method also performs well on the FVC, FVC-2018 dataset, and outperforms the state-of-the-art with an improved Recall and F score. From the analysis, it has been observed that non-clickbait's video channel has more number of subscribers with respect to registration age as compare to clickbait's video channel. The results show that the average cosine similarity for non-clickbait videos is too low for the average cosine similarity for clickbait videos. This is probably because the fake video may repeat the same sentence in the title over and over again to extend the length of the video. Redundant and inaccurate content. More features may be added for future work. For example, a video transcript can contain useful information. For example, the content of Clickbait videos is often quite different from the title, so the "gap" between the transcript and the title can provide important insights. You can also take into account the network structure of comments and replies that represent semantic features and attributes. In this study, we experimented with BERT, Word2Vec, and Distil BERT for word embedding. Embedding DocToVec can also be considered for future work. You can use the Random Forest classifier to explore other ensemble techniques, including the XGBoost.

REFERENCES

[1] Guan Jiexiong, Research and Implementation of Object Scout Recognition Algorithm Based on Computer Vision, Shanghai Normal University Master of Philosophy, 2019.11.

[2] Wang Rui, Embedded Image Recognition System and Optimization Based on Convolutional Neural Network, Master Thesis,2019.6.

[3] Zhang Yuting, Zhu Minling, Zheng Zhiwei, MonocularBased on the Analysis and Realization of Binocular Vision Mapping Process, Academic trends and new perspectives, 2019.

[4] Mao Mengna, Research on Target Recognition and Grab Technology of Binocular Vision Robot, Master Thesis, 2019.6.4.

[5] Lian Zifeng, Research on Image Recognition Algorithms Based on Deep Neural Networks, PhD thesis,2017.3.12.

[6] Thomas P (2017) Clickbait identification using neural networks. arXiv preprint arXiv:1710.08721. https://arxiv.org/pdf/1710. 08721.pdf

[7] Zannettou S, Chatzis S, Papadamou K, & Sirivianos M (2018) The good, the bad and the bait: Detecting and characterizing clickbait on youtube. In IEEE Security and Privacy Workshops, pp 63–69. https://doi.org/10.1109/SPW.2018.00018
[8] Heejung Jwa, Dongsuk Oh, Kinam Park, Jang Mook Kang, and Heuiseok Lim. exBAKE: Automatic fake news detection model based on bidirectional encoder representations from transformers (BERT). Applied Sciences, 9(19), 2019.
[9] Srijan Kumar, Robert West, and Jure Leskovec. Disinformation on the Web: Impact, characteristics, and detection of Wikipedia hoaxes. In Proceedings of the 25th International Conference on World Wide Web, pages 591–602, 2016.
[10] Martin Potthast, Tim Gollub, Matthias Hagen, and Benno Stein. The clickbait challenge 2017: Towards a regression model for clickbait strength. https: //arxiv.org/abs/1812.10847.

