iJRASET
International Journal For Research in Applied Science and Engineering Technology



# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

**Volume:** 9 **Issue:** IX **Month of publication:** September 2021

**DOI:** https://doi.org/10.22214/ijraset.2021.38256

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429*
*Volume 9 Issue IX Sep 2021- Available at www.ijraset.com*

# Smart System to Detect Adult Content and Child Pornography on Web

Shubham Kumar
*Computer Science and Engineering, SRM Institute of Science and Technology*

***Abstract:*** ***Adult contents on the internet is very common today but it has become a serious concern now because of many reasons such as the proliferation of free to view adult websites has made it easier for individuals of any age to gain access to explicit content. Children increasingly use mobile devices such as smartphones to access the internet and these adult content can have bad impact on their mind. Excessive exposure to these contents can lead to addiction which can have very adverse effect on their mind and their heath.so i came up with an idea to reduce this activity on internet saving people and childrens in particular. In this paper we will discuss a pipeline of a system developed which consists of three modules. The first one is scanning the heading and subheading of the web page to identify if the page is toxic or not. For this task we have used the SOTA model. The second module finds if the page contains a video and identifies it as adult video or not through video caption. we have used LSTM[1] network for this classification task and the last module is CV module which is the most important part of the project. It is an age detection module which detects the age of the people inside the video. The objective was to block the video if the age of any people doing activity exceeds 18. These modules when passed to the pipeline will forbid internet users to watch any kind of adult content specifically which involves a child. It's important to stop this because this has a very negative impact on our society and it is ruining our culture.***

***Index Terms: Deep Learning, Natural Language Processing, Computer vision, Transformers, LSTM***

## I. INTRODUCTION

These days adult contents on the internet are increasing which may cause harm to our system, mental health. Sexual images and pornography are easily accessible on websites today for streaming. Anyone with internet connection can be exposed to this even without warning. As of now there is no standard of age verification necessary before such contents can be consumed by users, however few sites take some precautions to prevent children from entering into this trap but their process is not very much effective. The problem does not stop here,people also get access to uncensored adult contents which is claimed to be uploaded on internet by false means. The ease with which youths can access the uncensored adult contents is a concern for our soceity. After noticing such problems,i thought of putting some effort to limit this. we will discuss Methodologies, Experiments and Improvement scope in the future regarding the system developed to stop this. Three machine learning models have been developed to deal with this problem.

System architecture flowchart

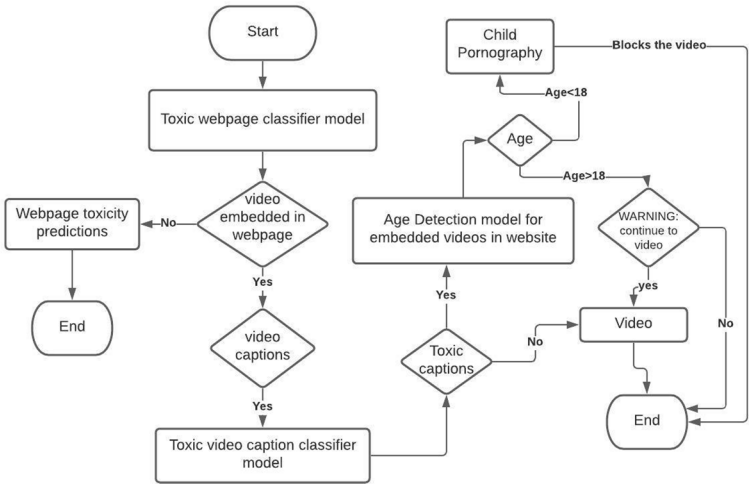



Figure 1

©IJRASET: All Rights are Reserved

*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429*
*Volume 9 Issue IX Sep 2021- Available at www.ijraset.com*

## II. ML MODELS

### A. *Classifying the Text Toxic Content in Page WITH BERT*

This is the first step in our system design. we trained a classification model to determine the contents of a webpage. After various experiments we selected distilbert[3] model over others.

### B. *Classifying the Category of the Video Through Video Captions*

This is the second model in the process. It is activated as soon as it is confirmed that the web page opened contains videos and captions on those videos. This model reads the caption, verifies the source and predicts if any video is likely to have adult contents which is unsafe.

### C. *Age Detection in the Video*

This is the third step in this process. This model is activated after a web page is loaded completely and some videos are found on the websites. It waits for the system to extract the video captions and details. This model was trained for detecting the age of the people in the video and blocking the same in few seconds after the video starts if the age of any person doing activity exceeds 18 for restricting people to watch any **child porn**.

## III. METHODOLOGIES, EXPERIMENTS AND RESULTS

### A. *Module 1 (Web page scanning)*

In this module we will scan the web page for any toxic text contents available in the page and warn the user about the same for a confirmation to enter into that page.The text language is supported in 104 language but the best performance was observed on english texts is because of the fact that training was done mostly on english data.For this task my benchmark data was a kaggle dataset namely Jigsaw Multilingual Toxic classification dataset from a kaggle competition.It contains 2,23,549 sentences and among which i selected 2,00,000 sentences for analysis and training for toxicity classification after data preprocessing step.The performance was calculated on several simple machine learning models using tf-idf vectorizer as follows:

| | F1-score | Accuracy |
|---|---|---|
| 1. Logistic Regression | 0.58 | 0.93 |
| 2. SVM | 0.61 | 0.91 |
| 3. Decision-Tree | 0.68 | 0.94 |
| 4. LSTM | 0.87 | 0.98 |

The performance measured was good with these models but not optimum to put into production.so we went up on this scale and tried SOTA models on our task.After experiments we achieved a convincing result with distilbert-multilingual-base model i.e an accuracy and f1-score of 0.87. We used the same data as mentioned above for the data preparation in the case of distilbert. For training we developed a custom model by fine-tuning distilbert[3] with the architecture mentioned in Figure 2 .The model is able to do cross-lingual learning in some of the languages which are similar to english.

Model Architecture (Module 1)

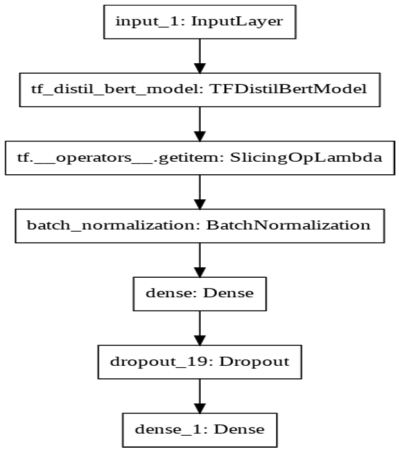



Figure 2

The model is responsible for scanning the heading,sub-heading and other important parts of the text to verify the page to be safe and toxic free.

©IJRASET: All Rights are Reserved

*B. Module 2 (Caption detection for any adult video content)*

In this second module of ML pipeline i tried to capture the captions of the video on the web page. This module is optional and is activated only if caption is present for the video block. The motive behind the development of this module was to block the adult videos like porngraphic videos as soon as the page opens not letting the internet users to even put their hands on those videos. The data was scraped from different adult websites and twitter was annotated manually by myself. During data preparation all the required data preprocessing parts were performed in the process of cleaning such as removing the stopwords, normalizing using lemmatization and singularization of the text and some other required fixes using regex. After data collection and data preprocessing. Now comes the part of developing ML model architecture and training it until a good accuracy and loss metric is observed.

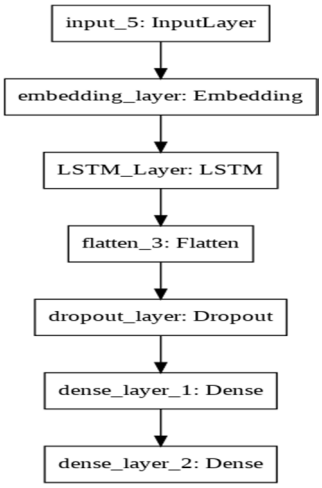



Figure 3

LSTM Based Model Architecture (Module 2)

In this task, LSTM[1] and GRU[4] are somewhat similar so I decided to go with LSTM one and perform the training process. Further I tried Bidirectional LSTM and achieved the accuracy of 99.4 % on the training test and 99% on the validation set.The model is saved and used in the backend for caption detection.

*C. Module 3 (Age Detection in Videos)*

This is the 3rd module of the project; here we execute our age detection algorithm only if caption detection module gives positive outcome from the captions extracted. While video starts to play, In the first few seconds it detects the age of the people in the video and restrict the video if any person in the video is below the age of 18. Intention was to ban child pornography here and it has been implemented successfully. We have used transfer learning in this module. We implemented this module with pre-trained weights and opencv framework. The model predicts the classes as range of ages and it is used to calculate the ages below 18 in videos. As soon as the video starts, the trained model detects the age of the people in video and marks their age in few seconds ,if the age is below 18 then the video is blocked immediately

## REFERENCES

[1] Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long Short-Term Memory. Neural Comput. 9, 8 (November 15, 1997), 1735–1780. DOI: https://doi.org/10.1162/neco.1997.9.8.1735

[2] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Vol. 1. 4171--4186. https://arxiv.org/abs/1810.04805

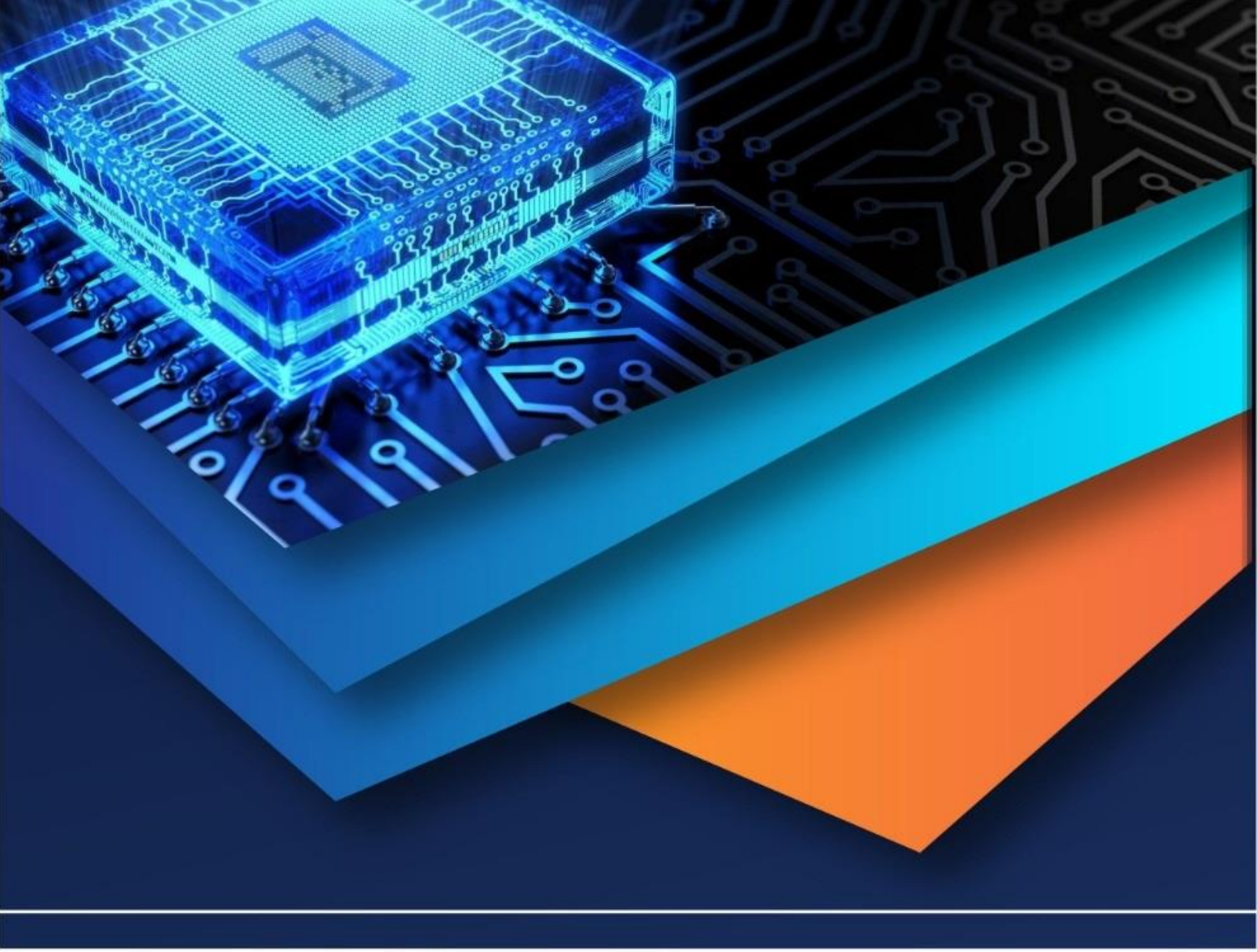
[3] Victor Sanh, Lysandre Debut, Julien Chaumond, and Thomas Wolf. 2019. DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. 2019. ArXiv abs/1910.01108. https://arxiv.org/abs/1910.01108

[4] Junyoung Chung, Caglar Gulcehre,Kyung Hyun Cho,Yoshua Bengio.Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling.https://arxiv.org/abs/1412.3555

## AUTHORS

Author – Shubham Kumar, Bachelor of Engineering, SRM Institute of Science and Technology and shubhamkr2091@gmail.com.

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)