e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Conference:01/March-2023 Impact Factor- 7.868 www.irjmets.com
National Conference on Trending Technology for Achieving Sustainable Development Goals
NCTTASDG 2023 Organized by Shri Shankarprasad Agnihotri College of Engineering, Wardha

IMPLEMENTATION OF AI-BASED SOCIAL MEDIA FORVULGAR CONTENT DETECTORANDREMOVER

Aarti Burghate[*1], Poonam Bramhane[*2], Pranjali Kanhekar[*3], Ruchika Parate[*4], Anjali dongre[*5], Pranjal kuhikar[*6]

[*1]Assistant Professor, Department of Information Technology NIT Nagpur

[*2,3,4,5,6]Undergraduate Scholars, Department of Information Technology NIT Nagpur

ABSTRACT

The rapid growth of social media has made it a primary source of communication and information for millions ofpeople. However, this has also led to an increase in the amount of vulgar and inappropriate content appearing onthese platforms. To address this issue, this project proposes the development of an AI-based solution for detectingandremovingvulgarcontentfromuser-generatedcontentinsocialmediaplatforms.The solution uses machine learning algorithms, specifically deep neural networks, to accurately detect vulgarcontentinrealtime.Alargedatasetofvulgarandnonvulgartextwascollectedandpreprocessedtotrainthemodel ,ensuring its effectiveness in detecting vulgar content. The model was fine-tuned using transfer learning to improveitsperformance.The trained model is integrated into the backend of the social media platform, where it automaticallyflagsandremovesanyvulgarcontentdetectedinusergeneratedcontent.Thesolutionisdesignedtobeeff icientandscalable,allowingittohandlethelargevolumeofcontentgeneratedonsocialmediaplatforms.Thesolutionis continuouslymonitoredandupdatedtoensureitsaccuracyandeffectiveness.Thisincludesretraining the model with additional data to improve its performance and adding new features to handle emergingtrendsinvulgarcontent.Theimplementationofthissolutionpromotesasaferandmoreinclusiveonlineenvir onmentforusers, whilealsofollowingethicaland legalguidelinesforonlinecontent moderation.

Keywords: Instagram, Social Network, VulgarContent

I. INTRODUCTION

Social media means the communication between human directly or indirectly on the electronics device throughinternet. The most popular social media platforms are Facebook, WhatsApp, Twitter, Instagram, LinkedIn etc. Andsocial media has become very much a part of our life that we started socializing on social mediavirtuallyinsteadofphysically.Socialmediahasbothpositiveandnegativeeffects.So,itisuptotheuseranduptotherule ofthatcountryin what way the social media should be used. Natural Language Processing is a branch of artificial intelligencebetweencomputersandhumansusingthenaturallanguage.Duringthelastfewyears,therehasbeenaninc reasingbody of research understanding hateful language in fields including Natural Language Processing (NLP), ArtificialIntelligence. Itis verymuchimportanttounderstandthatthisbehaviorcannotonlyimmenselyaffectthe lifeofan individual oagroupbutcouldbesuicidalinsomecasesadverselyhamperingthementalhealthofthevictim.

II. INSTAGRAM

Instagram is a free image and video sharing application available of iPhone and Android. People can upload imageor videos to our service and share them with their followers or with a select group of friends. They can alsoview,commentandlikepostssharedbytheirfriendsonInstagramallowsuserstoeditanduploadphotosandshortvi deosthrough a mobile application. Users can add a subtitle to each of their posts and use hashtags and location-basedgeotagstoguide thesepostsandmake themsocialeby otherusers withinthe app...

III. METHODOLOGY

ThiswebapplicationwascreatedusingthePythonDjangoFramework.We'llincludeaCNNalgorithmwithartificialinte lligence.Theabilityofartificialintelligencetoclosethegapbetweenhumanandcomputerskillshasbeengrowingdram atically. Both professionals and amateurs focus on many facets of the field to achieve great results. The field ofcomputervisionisoneofseveralsuchdisciplines.Inrecenttimes,CNNshaveexcelledinanumberofNLPtasks.OurCN Nsimulationismotivatedby(Kim,2014).Weprovideabrand-newdeeplearning-basedmethodforautomaticallyrecognizing.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Conference:01/March-2023 Impact Factor- 7.868 www.irjmets.com
National Conference on Trending Technology for Achieving Sustainable Development Goals
NCTTASDG 2023 Organized by Shri Shankarprasad Agnihotri College of Engineering, Wardha

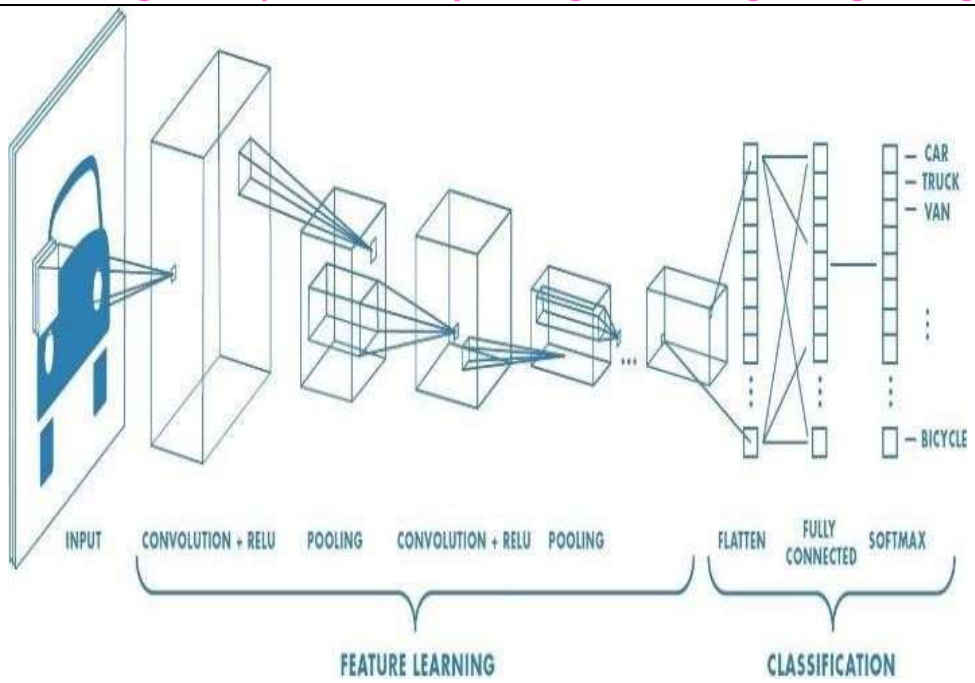

Figure1:CNNAlgorithm

IV. MODELINGANDANALYSIS

In this section, we will describe in detail the text classification model based on CNN used in this study. The flowchart of the model, as shown in Figure 2, including three modules: text representation, classifier training, andperformanceevaluation.

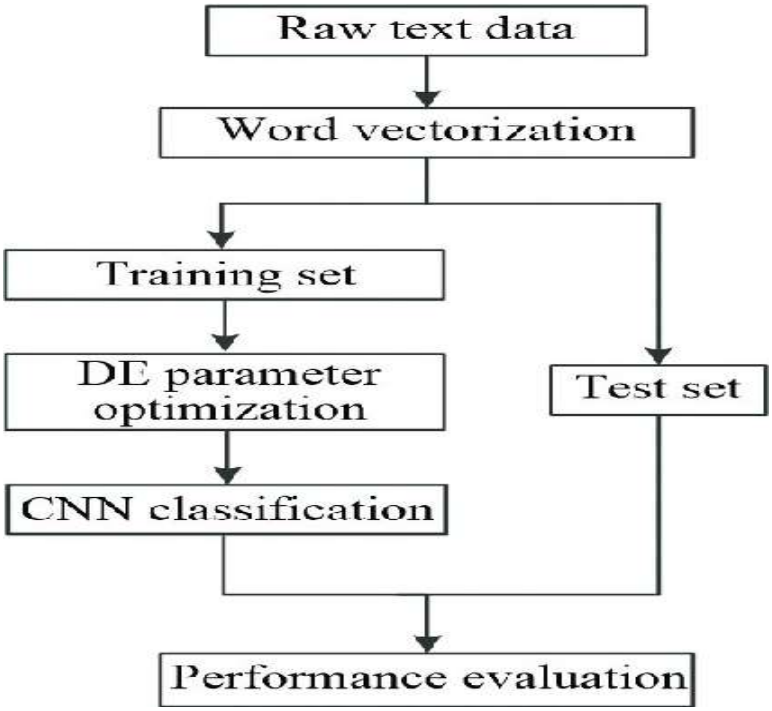

Figure2:FlowChatofCNN Algorithm

In the past few decades, Deep Learning has proved to be a very powerful tool because of its ability to handle largeamounts of data. The interest to use hidden layers has surpassed traditional techniques, especially in patternrecognition.One ofthemostpopulardeepneuralnetworksisConvolutionalNeuralNetworksindeeplearning.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Conference:01/March-2023 Impact Factor- 7.868 www.irjmets.com
National Conference on Trending Technology for Achieving Sustainable Development Goals
NCTTASDG 2023 Organized by Shri Shankarprasad Agnihotri College of Engineering, Wardha

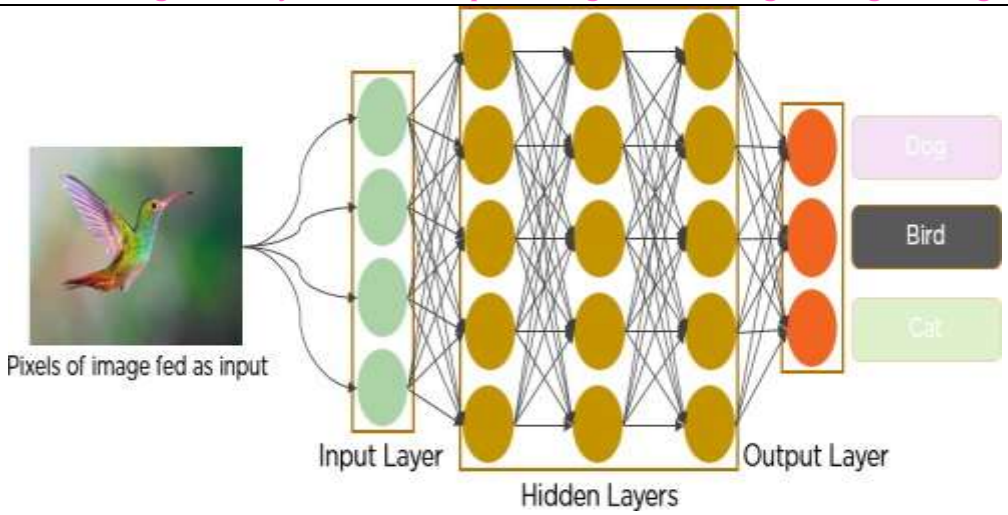

Figure3:Modelingand Analysis

V. RESULTSANDDISCUSSION

SearchSystem/Message/Post/Notification/Settingwillbeworkingproperly.

Notifications: - Notifications means it gives some information about what is going on, on our website or someupdates.Whenpeoplefollow youandlike orcommentyourpostthatshowin thenotification.

Search:-

Hearyoucanfindthepeoplebytheirnamesandtheir IDssoyoucanmakethemyouronlinefriends.Youcansearchonsocial mediausingkeywordsand youcan findphotosandvideos,hashtags,andtags.

Message: - A message is a communication or statement conveyedfrom one person to another.An interactiveexchange of messages from a conversation. A message may be delivered by various means, including couriertelegraphy,carrier pigeon andelectronicbus.

Post: - People can upload photos or videos to our service and share them with their followers or with a selectgroup offriends.Theycanalsoview,comment andlikeposts sharedby theirfriendsonInstagram.

Algorithm willoperate all thesettings:-TheNodesarecompletelylinkedtoallnodesinthelayerbelowthemina networkthat is fully connected. This generates a sophisticated model to investigate all CNN connections betweennodes.The intricacy,however, comes ata greatcostwhentraining thenetwork.

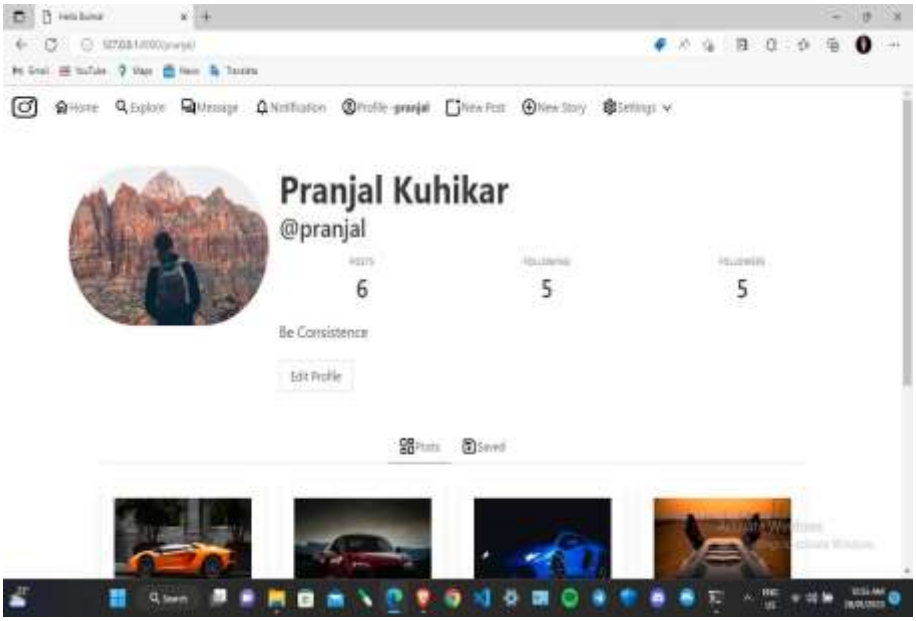

Figure4: Profile

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Conference:01/March-2023 Impact Factor- 7.868 www.irjmets.com
National Conference on Trending Technology for Achieving Sustainable Development Goals
NCTTASDG 2023 Organized by Shri Shankarprasad Agnihotri College of Engineering, Wardha

VI. CONCLUSION

We can make our social media web application that provides all facilities like Instagram. A social media app likeInstagram interacts with the attention of the public, check out this post. In it, we provide a quick and easy-to-follow step-by-step guide on how to make a social media app that can easily get a lot of engaged followers andeffectivelyspreadyourbrand'simageonline. An Instagram-like social media app is an application that enables users to freely produce, assess, and consumecontent delivered asa mixture of text, sound, videos, and pictures at any time of the dayor night. No doubt, thesocial media landscape has changed drastically in recent years. Social media has evolved into a direct- to-consumer format wherepeople are used to presentingandreceivingpersonalizedcontent. the implementation of an AI-based social media platform for detecting and removing vulgar content has thepotential to create a safer and more comfortable online community for users. By utilizing natural languageprocessingandmachinelearningtechniques,theplatformcanautomaticallyanalyzeusergeneratedcontent andflag any posts or comments that contain offensive language or imagery. The platform also includes user-facingtools for reporting and removing offensive content, as well as mechanisms for penalizing or banning users whorepeatedlyviolatecommunityguidelines.

While the development of such a system requires a team of experts in various fields, including NLP, machinelearning, software development, and UI/UX design, the benefits of creating a safer and more positive onlinecommunity make it a worthwhile endeavor. It is also important to keep in mind the importance of data privacyandsecurityregulationswhenworkingwithuser-generatedcontent.

VII. REFERENCES

[1] Aghababaei, S., Makrehchi, M., 2017. Mining social media content for crime pre-diction. In: Proceedings - 2016IEEE/WIC/ACMInternationalConferenceonWebIntelligence.IEEE,pp.526e531.https://doi.org/10.1109/WI.2016.0089 .WI2016.

[2] Al-Room, K., et al., 2021. 'Drone forensics: a case study of digital forensic investigations conducted oncommon drone models', international journal of digital crime and forensics. IGI Global 13 (1), 1e25.https://doi.org/10.4018/IJDCF.2021010101.

[3] Anand,M.,Eswari,R.,2019.Classificationofabusivecommentsinsocialmediausingdeeplearning.In:Proceedingsofthe3rdInternationalConferenceonComputingMethodologiesandCommunication,ICCMC2019.InstituteofElectricalandElectronicsEngineersInc.,pp.974e977.https://doi.org/10.1109/ICCMC.2019.8819734

[4] Arshad, H., Jantan, A., Omolara, E., 2019. Evidence collectionand forensicson socialnetworks:researchchallengesanddirections. Digit. Invest. 28,126e138.https://doi.org/10.1016/j.diin.2019.02.001.ElsevierLtd.