

A MACHINE LEARNING ALGORITHM-BASED SURVEY OF CLICK BAIT DETECTION IN SOCIAL MEDIA

S.S. Senthil Priya, S. Manju Priya*

Abstract

The task of detecting clickbait involves locating the deceptive form of advertising known as clickbait, it uses hyperlink text or a thumbnail link to draw users' attention and encourage them to read, view, or listen to the linked piece of online content. Clickbait is frequently sensationalised or misleading. This article focuses about the overview of clickbait detection in social media platforms using deep learning and machine learning algorithms.

Keywords: click bait, social media, machine learning.

I. INTRODUCTION

In the age of instant gratification, social media platforms became the primary means of interpersonal communication. As user opinions can be easily posted on social media sites like Twitter, Facebook, YouTube, and others, a lot of false and subpar audio-visual content is dispersed globally. Some examples of clickbait phrases are: “You'll Never Believe This”, “How to Achieve Results Using This”, “One Weird Trick” ,“They Didn't Know”. Clickbait is an overblown headline that encourages readers to click on a link to an article, image, or video. Instead of providing objective facts, clickbait headlines usually prey on people's emotions and interests. In general, clickbait is intentionally misleading or astonishing information that is produced solely to increase website page views. Publishers rarely employ clickbait to increase visitors to their website in order to increase advertising revenue, clicks, or conversions. Language used in clickbait is more emotive than informative. An example for clickbait is described in the below figure1. Instead of

searching a relevant content simply attract towards a thumbnail view or the catchy phrase of a video or post. This encouraged people to read a lot of erroneous and deceptive material



Figure 1: Example of Clickbait [8]

In 2016 [1], Clickbait detection were done by creating a corpus from various social media platforms like Facebook, reddit and twitter. By using Convolutional Neural Networks model achieves high performance in identification of click baits.

In 2017[2], Clickbait detection were done by analyse various click bait posts and headlines from the social media platform and filtered many features. In total of 331 features and to avoid overfitting we keep 180 features. The top 60 features are chosen to reduce run time and lessening noise interference.

In 2017[3], Clickbait detection were done by examine various social media posts in clicakbait challenge. The authors found that clickbait post titles are statistically significant shorter than legitimate post titles. The amount of formal English words in the given content can be counted,

¹Department of Computer Science

²Department of Computer Applications

Karpagam Academy of Higher Education, Coimbatore, Tamil Nadu, India which is another good method for spotting clickbait.

*Corresponding Author

In 2020[4], False news Detection were found in Reddit with deep learning framework. Using our Part of Speech Analysis Module, the authors concentrate on linguistic analysis during the knowledge discovery phase and the structure of clickbait headlines. Classification requires the use of extended short-term memory for decision-making.

In 2021[5], A unified approach using cognitive evidence in Youtube videos, a clickbait detection were done. The first stage is to extract the audio from the videos, which is then converted into textual data and used to extract features based on the video content. Second, features are collected from the comments based on human replies and reactions to the posted content[6,7]. Finally, features based on user profiles are retrieved. Finally, the classifier performs admirably after receiving all these features.[8,9,10].

II. COMPARISON OF EXISTING ALGORITHMS

[1] Clickbait detection were done by creating a corpus from the social media in Convolutional Neural Networks and provides the accuracy of 0.90 along with the precision of 0.85 and Recall of 0.88

[2] Clickbait detection were done with the machine learning algorithm of Random Forest Regression with accuracy of 0.82 along with MSE of 0.035 and F1 Score of 0.61

[3] With XGBoost's machine learning classifier, clickbait was detected with accuracy of 0.812, precision of 0.819, and recall of 0.906

[4] False news Detection were found in Reddit with deep learning framework and provides a accuracy of 0.97 along with a precision of 0.96, Recall of 0.98 and F! Score of 0.973

[5] A unified approach using cognitive evidence in Youtube videos, a clickbait detection were done and found J48 clasifier produces an accuracy of 98.89%

The comparisons of the existing algorithms were discussed in the below table 1 in that the machine learning algorithms of J48 Produces the best accuracy[11,12].

Paper & Year	Social media & Features	Accuracy	Methodology
[1] 2016	Reddit, Twitter & facebook-headlines	0.90	Convolutional Neural Network Model
[2] 2017	Clickbait challenge-Posts and headlines	0.82	Random Forest Regression
[3] 2017	Clickbait challenge-Posts	0.81	XGBoost
[4] 2020	Reddit-fake news detection	0.97	POSAM, LSTM
[5] 2021	Youtube-videos	0.98	J48

Table 1: Comparisons of existing algorithms

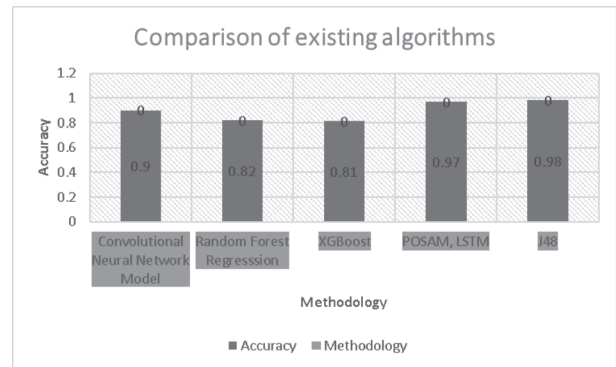


Figure 2: Chart representation for comparison of existing algorithm

III. CONCLUSION

In this study, the survey on clickbait detection is analyzed with the various social media platforms and features were discussed. The accuracy from various research were compared along with the methodology used. The machine learning algorithms of J48 classifier produces the best accuracy of 0.98.

In the future the clickbait will be detected in the Social media platforms using deep learning algorithms.

REFERENCES

1. A. Agrawal, "Clickbait detection using deep learning," 2016 2nd International Conference on Next Generation Computing Technologies (NGCT), 2016, pp. 268-272, doi: 10.1109/NGCT.2016.7877426.
2. Cao, Xinyue, Le, Thai, Jason, Zhang "Machine Learning Based Detection of Clickbait Posts in Social Media" in social media submission challenge 2017
3. Elyashar, Aviad and Bendahan, Jorge and Puzis, Rami "Detecting Clickbait in Online Social Media: You Won't Believe How We Did It"(2017),The Salmon Clickbait Detector at the Clickbait Challenge 2017
4. A deep learning framework for clickbait detection on social area network using natural language cues Bilal Naem¹ Aymen Khan¹ Mirza Omer Beg¹ Hasan Mujtaba¹ January 2020 © Springer Nature Singapore Pte Ltd. 2020
5. Varshney D, Vishwakarma DK. A unified approach for detection of Clickbait videos on YouTube using cognitive evidences. Appl Intell (Dordr). 2021;51(7):4214-4235. doi: 10.1007/s10489-020-02057-9. Epub 2021 Jan 2. PMID: 34764575; PMCID: PMC7778503.
6. Tobias Blanke & Tommaso Venturini, 2022."A network view on reliability: using machine learning to understand how we assess news websites," Journal of Computational Social Science, Springer, vol. 5(1), pages 69-88, May.
- 7.<https://ideas.repec.org/a/spr/jcsosc/v5y2022i1d10.1007_s42001-021-00116-w.html>
8. Muhammad Saad Javed & Hammad Majeed & Hasan Mujtaba & Mirza Omer Beg, 2021."Fake reviews classification using deep learning ensemble of shallow convolutions," Journal of Computational Social Science, Springer, vol. 4(2), pages 883-902, November.<https://ideas.repec.org/a/spr/jcsosc/v4y2021i2d10.1007_s42001-021-00114-y.html>
9. Anna Ruelens, 2022."Analyzing user-generated content using natural language processing: a case study of public satisfaction with healthcare systems," Journal of Computational Social Science, Springer, vol. 5(1), pages 731 - 749, May.<https://ideas.repec.org/a/spr/jcsosc/v5y2022i1d10.1007_s42001-021-00148-2.html>
10. Clickbait Detection using Multiple Categorization Techniques Abinash Pujahari¹ and Dilip Singh Sisodia² ^{1,2}Department of Computer Science and Engineering National Institute of Technology Raipur, India
11. Daoud, D. M., and S. Abou El-Seoud. "An Effective Approach for Clickbait Detection Based on Supervised Machine Learning Technique". International Journal of Online and Biomedical Engineering (iJOE), vol. 15, no. 03, Feb. 2019, pp. pp. 21 - 32, doi:10.3991/ijoe.v15i03.9843.
12. <https://hookagency.com/blog/clickbait-titles/>