Fact Finder

Saurabh Chaudhary, Kavita Kamtekar, Suraj Salunke, Sai Tata, Manthan Kale,

Maheswari Vidyadharani , Gayatri Kulkarni, Rushabh Shingala

Seidenberg School of Computer Science and Information Systems

Pace University, New York, NY, USA

Abstract— **Fact Finder is a web app designed to combat online misinformation by differentiating between fake and real news. Additionally, it can create a summary of the news. For anyone concerned about online misinformation and facing time constraints for reading news articles. Who uploaded the news article the fact finder web app is a web app that uses machine learning and deep learning algorithms that can detect misinformation unlike believing any information on the internet. Our application verifies the authenticity of information and saves time by creating summaries of articles."Fact Finder" can provide users with reliable, accurate, and up-to-date information on various topics. Additionally, it offers tools or resources for users to verify the accuracy of claims they encounter online.**

***Keywords*— Machine Learning, TensorFlow, News Articles, Logistic Regression Model Firebase, Machine Learning, Fake News, Real News, Firebase, scikit-learn.**

INTRODUCTION

Misinformation and the proliferation of fake news are big problems with wide-reaching effects, influencing public opinions, politics, and the trust people have in society. Social media makes it easy for misleading or completely false information to spread quickly, which can cause panic or lead people to make poor decisions. Fact Finder is always improving to keep up with the latest tricks used in misinformation. It gives users the tools to tell fact from fiction, helping create a better-informed public. The app uses the latest in data mining, natural language processing, and information retrieval to automatically spot and correct wrong information. It also uses strong fact-checking methods that double-check information with several reliable sources, which makes its processes clear and trustworthy. Fact Finder also focuses a lot on teaching and media smarts. It has interactive modules

LITERATURE REVIEW

Fake new are easy to spread and there are very little tools available to connect this problem. In today world were their is lot of content available on the internet it is important to know which information is true and which is not. AI is useful but it can also be used to spread fack new and also with digital content it has become very easy for the people to create a fake content. This content could manipulate the voting decision as well. AI model can also be used to detect the fact news and this where logistic regress comes into play.

Methodology

We mixed them all up to create a diverse set for training our tools. Each article got a tag — '1' for fake and '0' for real. To clean up the mess and make the text neat for analysis, we used some Python magic (like the re and porterStemer libraries) to strip out unwanted characters, punctuations, and common words that don't add much info, plus we chopped words down to their roots.
Turning Text into Numbers and Teaching the Model.
Next, we turned all that cleaned-up text into numbers that reflect how important each word is in the context of the whole bunch of articles. This is done using a technique called TF-IDF vectorization. We split all these number-fied articles into training and testing groups. We used a logistic regression model — it’s simple but good at sorting things into two groups (real or fake).

The model is trained using logistic regression here we have used logistic regression is used to train the model and the model was trained with the data set of true and fake cvs file. the data was clean and then given as an input to the model. the model which is using logistic regression analysed the text and then it was trained. after train we have to test the model for its accuracy if the accuracy of the model is greater then 85 percent is can be sad that that the model has performed well. so for this train data is split is into two section one section has 80 percent of data and another has 20 percent of data. the model was trained with 80 percent of data and then it will be tested on 20 percent of the data. Model showed the accuracy of more than 90 percent this indicates that

we have achieved our goal.

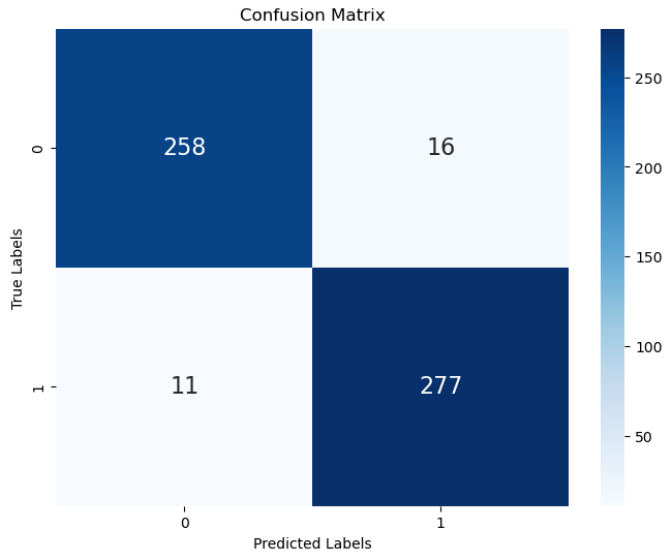
Summary generation

summary generation ai model is also used for creating a summary of the news this model is different form the one which is used to detect fake news this model is based on Neural network. Neural network is a type of network which contains three things first is input layer second is middle layer and the third is output layer. the input has the input of the data and next is the middle layer where there are activation functions and lastly we have out of the layer which is the result of the training. middle layer can have multiple layer this used to increase the accuracy of the model. the library which is used here the TensorFlow this is the library which is most popular library.

EVALUATION METRICS

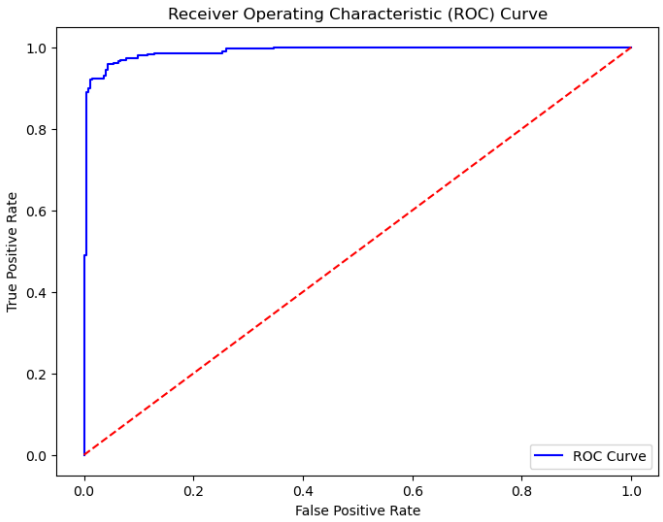
With help of the accuracy we can find out how well the model is performing. If the accuracy is more then the is less false positive and integrity of the of the news is more.

RESULTS



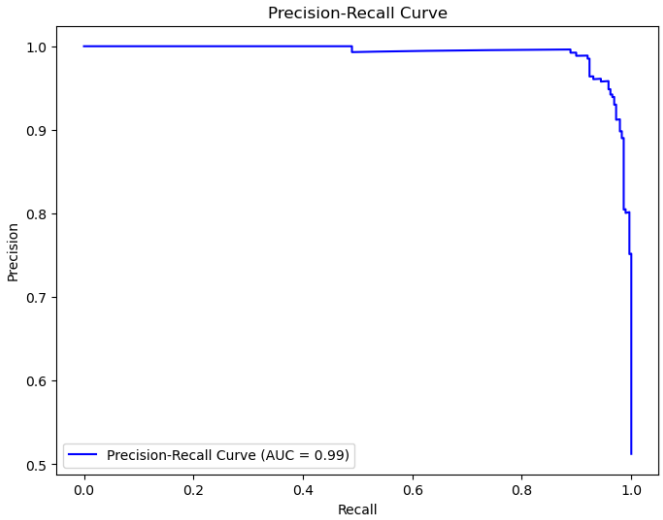
ROC CURVE:

It about true positive and true negative and the greater true positive and less false positive is good sign.



PRECISION-RECALL CURVE:

.with the combination of precision and recall the model is about to give balance output.



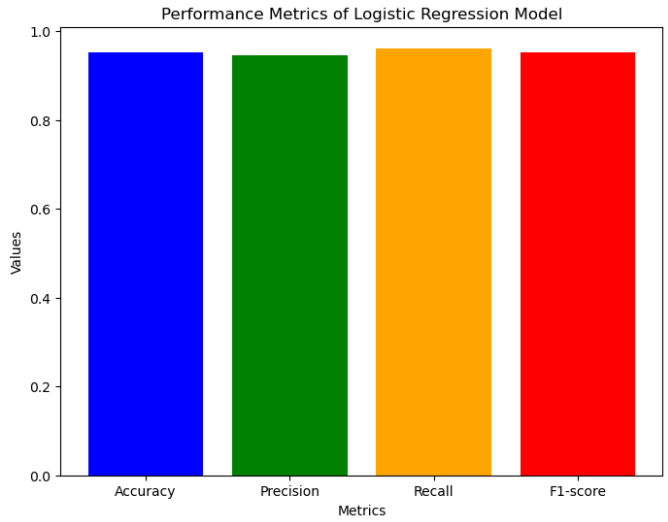
PERFORMANCE METRICS:

Accuracy Score:0.95 is the way our model has generate as the total result

Precision Score:0.94 accuracy is good for predication

Recall Score: 0.96 means it is detecting true news in this percen

F1 Score: it is average of recall and precision. 0.94 is good for our model



Media Literacy and Education

Understanding how crucial it is for people to be sharp when evaluating information, a bunch of studies have looked into ways to boost media smarts. There are even courses designed to help spot dodgy misinformation tricks, judge whether a source looks legit, and check out if what's being claimed is true. FactFinder is all in on this – it's got educational tools and pushes media literacy as a key way to fight against fake news. Plus, there's interesting stuff being done with digital literacy and even learning through games, which might help shape how FactFinder teaches these skills. Cross-Sector Collaborations: Misinformation is complex and it needs a team effort. That's why FactFinder is useful working together with users from tech companies, news outlets, universities, and government agencies. This teamwork is important for making FactFinder's work better. There are also better projects like the Trust Study and the Credibility Coalition where experts from different areas come together to set standards, share what works best, and make sure information is shared openly, responsibly and accurately. These partnerships could be a great model for how FactFinder teams up with others. As there is also fake information and it used to spread false info and it keeps changing, researchers need to keep upping their game. Looking ahead, there's potential in using better tech like deep learning and analyzing different kinds of media to spot the fake news better. It's also worth looking into how combining human smarts with AI in checking facts could work, and understanding the social conflicts that help misinformation spread.

III. CURRENT SOLUTIONS

our objective is first user should be able get the news result that our model should tell the user this news is true or fake. second our model should be able to generate summary from the model we have trained for the same. and the lastly we have to give user to post and chat with other user.

VI. PRODUCT RESULTS

So far, the accuracy of our model is around 84% which is good enough considering that the model is learning and improving every day. The model is very good at detecting fake news and the real news that are clear and perfectly giving the output result whether it's actual news or fake news.. We have doubled our test data and every day try to get in as much data as possible so that the model performance improves and accuracy. Although we have come both on the front-end part of the FactFinder website as well as the backend part. The UI/UX of the FactFinder website looks great and has incorporated features like User Profile Creation, Verify News, Login Page, and About FactFinder details and About us section.

Fig: FactFinder Home Page

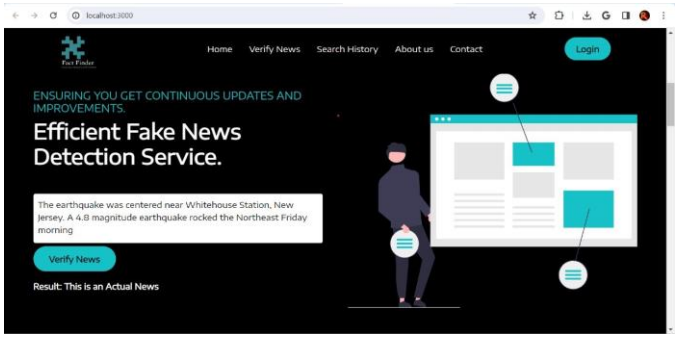


Fig: Verify Actual News

.

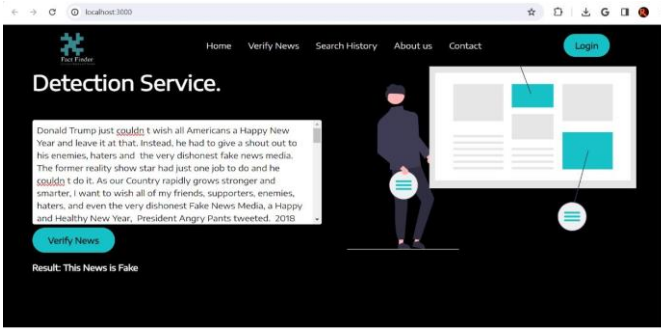


Fig: Verify Fake News

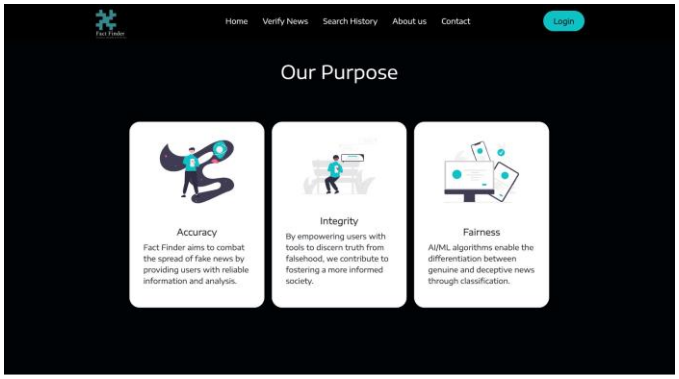


Fig: About us Tab

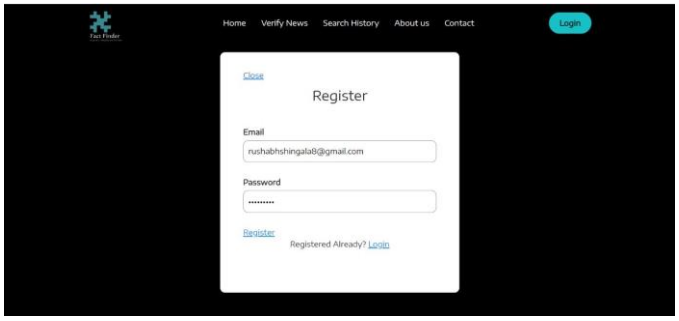


Fig: Registration Tab

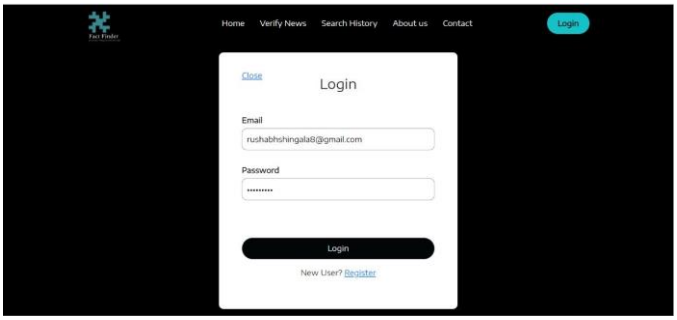


Fig: Login Tab

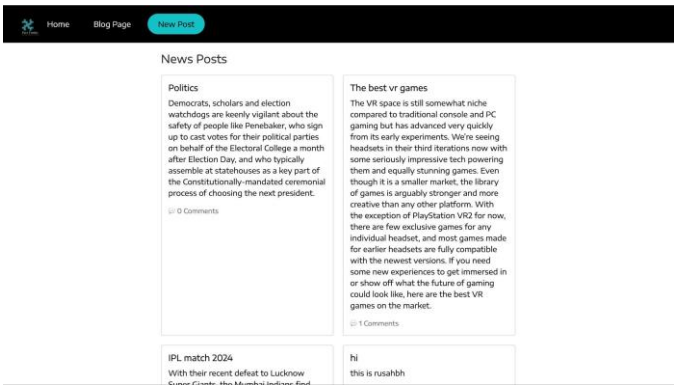


Fig: Post Blog

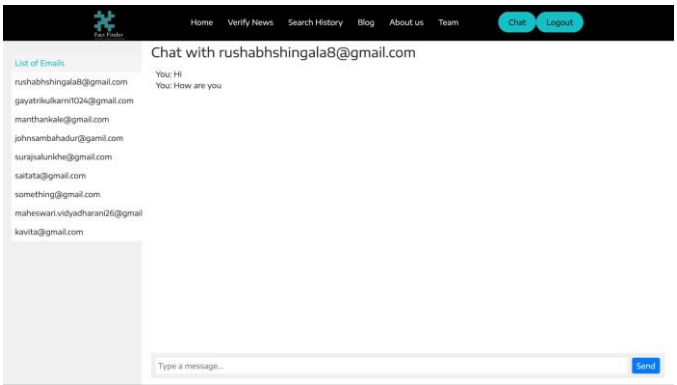


Fig: Chat Section

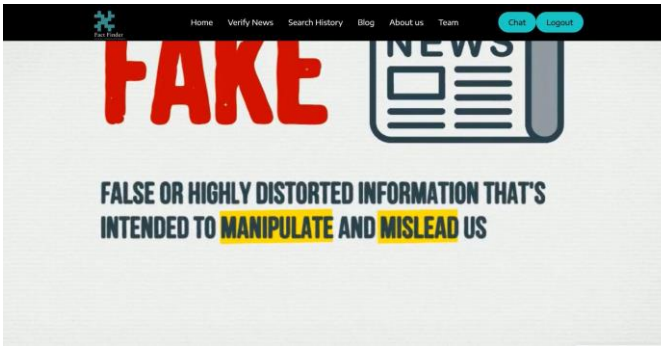


Fig After Login page.

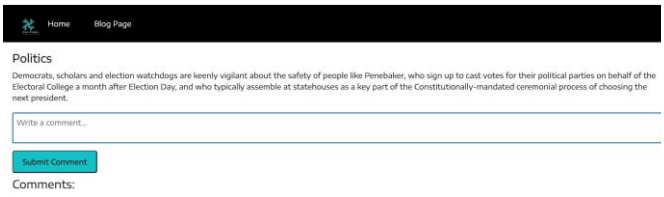


Fig Comment

I. CONCLUSIONS

The version of this template is V2. Most of the formatting instructions in this document have been compiled by Causal Productions from the IEEE LaTeX style files. Causal Productions offers both A4 templates and US Letter templates for LaTeX and Microsoft Word. The LaTeX templates depend on the official IEEEtran.cls and IEEEtran.bst files, whereas the Microsoft Word templates are self-contained. Causal Productions has used its best efforts to ensure that the templates have the same appearance.

Causal Productions permits the distribution and revision of these templates on the condition that Causal Productions is credited in the revised template as follows: "original version of this template was provided by courtesy of Causal Productions (www.causalproductions.com)".

ACKNOWLEDGMENT

REFERENCES

[1] (2002) The IEEE website. [Online]. Available: http://www.ieee.org/
[2] M. Shell. (2002) IEEEtran homepage on CTAN. [Online]. Available: http://www.ctan.org/tex-archive/macros/latex/contrib/supported/IEEEtran/
[3] FLEXChip Signal Processor (MC68175/D), Motorola, 1996.
[4] "PDCA12-70 data sheet," Opto Speed SA, Mezzovico, Switzerland.
[5] A. Karnik, "Performance of TCP congestion control with rate feedback: TCP/ABR and rate adaptive TCP/IP," M. Eng. thesis, Indian Institute of Science, Bangalore, India, Jan. 1999.
[6] J. Padhye, V. Firoiu, and D. Towsley, "A stochastic model of TCP Reno congestion avoidance and control," Univ. of Massachusetts, Amherst, MA, CMPSCI Tech. Rep. 99-02, 1999.
[7] Sharma, K., Qian, F., Jiang, H., Ruchansky, N., Zhang, M., & Liu, Y. (2019). Combating fake news: A survey on identification and mitigation techniques. ACM Transactions on Intelligent Systems and Technology (TIST), 10(3), 1-42.
[8] Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification, IEEE Std. 802.11, 1997.
[9] Khanam Z., Alkhaldi S. "An Intelligent Recommendation Engine for Selecting the University for Graduate Courses in KSA: SARS Student Admission Recommender System." In: Smys S., Bestak R., Rocha Á. (eds) Inventive Computation Technologies. ICICIT 2019. Lecture Notes in Networks and Systems, vol 98. Springer, Cham. 2019.
[10] Zhang, Jiawei, Bowen Dong, and S. Yu Philip. "Fakedetector: Effective fake news detection with deep diffusive neural network." 2020 IEEE 36th International Conference on Data Engineering (ICDE). IEEE, 2020.
[11] Cardoso Durier da Silva, F., Vieira, R., & Garcia, A. C. (2019, January). Can machines learn to detect fake news? a survey focused on social media. In Proceedings of the 52nd Hawaii International Conference on System Sciences.
[12] Bovet, Alexandre, and Hernán A. Makse. "Influence of fake news in Twitter during the 2016 US presidential election." Nature communications 10.1 (2019): 1-14. The science of fake news.
[13] Veronica Perez-Rosas et al. Available at: https://www.researchgate.net/publication/319255985_Automatic_Detection_of_Fake_News August, 2017.
[14] Prabhjot Kaur et al. " Hybrid Text Classification Method for Fake News Detection." International Journal of Engineering and Advanced Technology (IJEAT) , 2388-2392. 2019.
[15] I. Traore et al. "Detection of Online Fake News Using N-Gram Analysis and Machine Learning Techniques." International Conference on Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments (pp. 127–138). Springer International Publishing . 2017.