



NCAITC-2023

6th National Conference ((Mixed Mode) on

Advances in Information Technology and Computing

(Theme: Modern Trends in Information and Computing Technology)

April 27-28, 2023 (Thursday -Friday),

Organized by **DEPARTMENT OF INFORMATION TECHNOLOGY**
Sreenidhi Institute of Science and Technology (SNIST), Hyderabad,

Call for: Papers, Innovative Ideas & Project Presentations

National Conference on Advances in Information Technology and Computing (NCAITC) initiated in 2017 by Information Technology Department of Sreenidhi Institute of Science and Technology (SNIST) to provide an avenue for Students, Research Scholars, Faculty and Industry Practitioners to share related research ideas, approaches and methods in various current and upcoming areas of Computing, Information Technology & Communications.

The 6th National conference on Advances in Information Technology and Computing (Both Online and Physical) during April 27-28, 2023 invites papers, Innovative Ideas and Project Presentations in Poster from all UG, PG & RESEARCH SCHOLARS of any branch in the following areas. (Could be any field connected to information technology).

Tracks

- Artificial Intelligence, Machine Learning, and Deep Learning
- Block Chain and Applications
- Big Data and Data Science
- Networking and Information Security
- Distributed and Cloud Computing
- Multimedia Applications
- Smart & Intelligent Computing
- Health informatics
- Internet of Things Frameworks

Important Dates

Paper Submission: 10th April 2023; **Author Notification:** 17th April 2023

Camera Ready with revisions and registration: 23rd April 2023

Registration Fee: per paper: Rs 1000/- for Indians
USD 15/- for Foreigners

Fee payable to : Principal, Sreenidhi Institute of Science and Technology

Account Number: 62268594021 (Current Account)

IFSC Code : SBIN0021266

All accepted Papers after thorough revision will be invited for oral presentation published in conference proceedings and also in **International Journal of Human Computations and Intelligence (IJHCI) ISSN:2583-5696, Indexed in Google Scholar etc.,**
link: <https://www.milestoneresearch.in/JOURNALS/index.php/IJHCI>

All the Registered participants will be provided with

1. Soft copy of the proceedings; Working Lunch; Participation Certificate
2. Can attend Seminars/ webinars on April 27 -28 2023 by Eminent Working Academic and Industry Professionals
3. Best Paper Award /Best idea Award/Best Project Award Certificates

Chief Patron(s): Dr.K.T.Mahhe, Secretary, SNIST
Mr.K.Abhijit Rao, CEO, SNIST

Patron(s): Dr.C. Tomy, Director, SNIST
Dr.T.Ch.Siva Reddy, Principal SNIST

Convener: Dr.Sunil Bhutada, HoD-IT, SNIST

Co-Convener(s): Prof.P. Sreedhar, Assoc.Head-IT, , SNIST
Dr.Sheik Subhani, Assoc.Head-IT, , SNIST

Org. Secretary: Dr.N.Ch.S.N.Iyengar , Prof IT ,SNIST

Co.Org .Team Dr K Sreerama murthy, Asso.Prof IT ,SNIST
DR. Kranthi Kumar K, Asso.Prof IT ,SNIST

Conference Coordinators:

Mr.R.Kameshwar Reddy, Asst.Prof IT ,SNIST
Mr.K Vigneswara Reddy, Asst.Prof IT ,SNIST
Mr.P Mahesh Kumar, Asst.Prof IT, SNIST
Mrs.Jogu Saritha, Asst.Prof IT, SNIST
Ms.K.Nalini, Asst.Prof IT ,SNIST

For Details please contact

Dr.Sunil Bhutada, HoD-IT, SNIST; +91 9849822142
Dr.N.Ch.S.N.Iyengar , Prof IT ,SNIST; +91 9489343789

A deep learning semantic segmentation-based document classification method

Harshini R¹ . Nihar Gaokar¹ . Nagashree N²

¹School of Computer Science and Engineering, REVA University, Bengaluru, India.

²Department of Computer Science and Engineering, Sai Vidhya Institute of Technology, India

Received: 03 November 2022 / Revised: 26 November 2022 / Accepted: 06 December 2022

©Milestone Research Publications, Part of CLOCKSS archiving

Abstract – The introduction of the internet and internet-based existence has had a significant impact on people's lives. Additionally, mass media has an impact on everyday public life. Due to political power, many people take use of these rights to indulge in luxuries and elevate their social status. After COVID 19, People deliberately spread fake information through web-based social networking sites. This has an impact on how internet news sites originally operated and were intended. Therefore, we need some tools to automate the process and identify effective ways to classify it in order to stop the spread of such bad news. A computer vision based semantic segmentation method is proposed on a deep learning platform to perform multi-document analysis.

Index Terms – Deep learning, segmentation, document classification

I. INTRODUCTION

Online social networks have developed into a platform for individuals all over the world to obtain a wide range of news articles that are also well- liked. The medium via which news information is made available has changed significantly from the time of newspapers to contemporary internet. Fake news is low-quality content with purposefully false information that is disseminated by people or automated programs with malicious intent to manipulate messaging for snitching or political ends [1]. As a result, it is necessary to divide the news into two categories: legitimate news and fake news. An automated technology that can identify news articles is employed in place of the laborious and time-consuming manual tagging of news [2].

With respect to computer vision into consideration, many methods are in research towards document classification. There are multiple techniques in the literature such as logistic regression [3] and other statistical approaches [4] and certain ensemble methods. Out of these classifiers, they have a problem of overfitting with the dataset and all are discrete values which cannot be classified [5]. Many computer vision based semantic segmentation techniques are under research for efficient object recognition and classification [6,7,8]. With respect to document level analysis, semantic segmentation gives better comparison parameters. While training the dataset using deep learning aspect, semantic segmentation tries to capture a key with every pixel which gives guarantee towards better object recognition and classification [9,10].

II. PROPOSED METHODOLOGY

The proposed method has the following architecture as depicted in fig 1. Data Collection: The initial phase in the proposed work includes extraction of documents from the user's inputs which is in the form of emails, pdf, text files, documents etc. As there are multiple document with different modality that is they are not uniform, the set of documents are subjected to pre-processing. Pre-processing: During pre-processing, the

documents in different formats will be classified separately as text files, emails, pdfs etc. This should take place with a document classifiers. The method is shown below in fig 2.

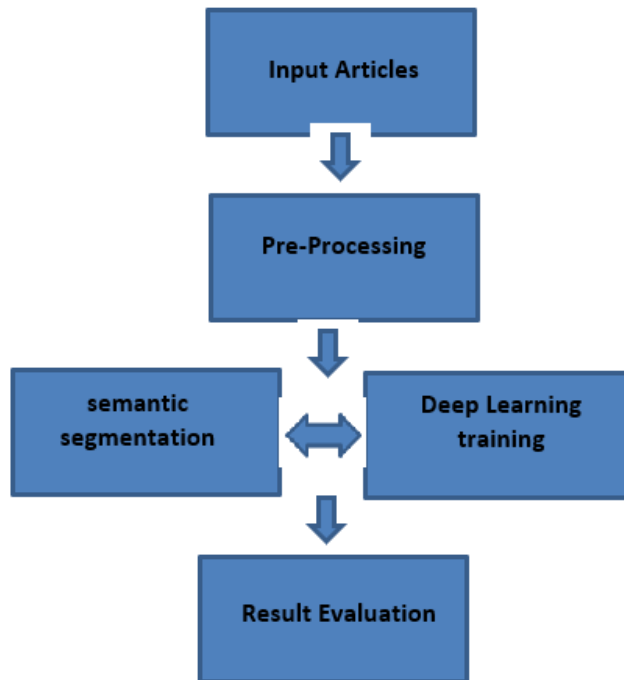


Fig 1: Proposed flow diagram



Fig 2: Multidocument classification (Pre-processing)

Semantic Segmentation: This phase involves segmenting the documents into its corresponding similarities. Feature descriptors from the document are extracted and thus based on the features idea segmentation is performed. It makes use of deep learning methodologies.

III. METHODS AND MATERIALS

The method is implemented in python where different cases of articles are extracted. As a part of pre-processing document analysis is carried out with feature vectors represented in python result as in fig 3 and 4

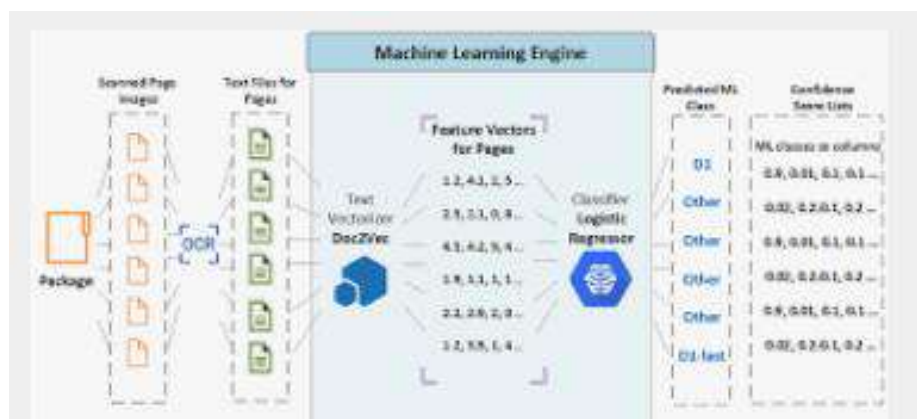
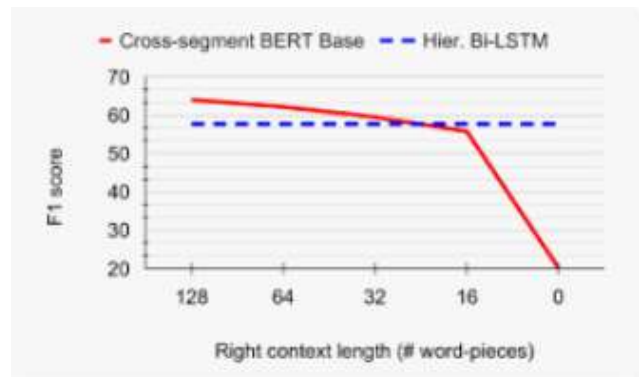


Fig 3: Feature vector representation in python



IV. CONCLUSION

One of the most crucial ways to inform people about events in and around our community is through the news. Therefore, those who create and disseminate it are responsible for stopping it from being genuine. In this study, supervised machine learning techniques and natural language processing are used to automatically classify news as either legitimate or fake.

REFERENCES

1. Debnath, K., & Kar, N. (2022, May). Email Spam Detection using Deep Learning Approach. In *2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON)* (Vol. 1, pp. 37-41). IEEE.
2. Junnarkar, A., Adhikari, S., Fagania, J., Chimurkar, P., & Karia, D. (2021, February). E-mail spam classification via machine learning and natural language processing. In *2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)* (pp. 693-699). IEEE.
3. Nagashree, N., Patil, P., Patil, S., & Kokatanur, M. (2021, June). Alpha beta pruned UNet-a modified unet framework to segment MRI brain image to analyse the effects of CNTNAP2 gene towards autism detection. In *2021 3rd International Conference on Computer Communication and the Internet (ICCCI)* (pp. 23-26). IEEE.
4. Cheng, Q., Xu, A., Li, X., & Ding, L. (2022, March). Adversarial Email Generation against Spam Detection Models through Feature Perturbation. In *2022 IEEE International Conference on Assured Autonomy (ICAA)* (pp. 83-92). IEEE.
5. Al-Shammari, N. K., Syed, T. H., & Syed, M. B. (2021). An Edge-IoT framework and prototype based on blockchain for smart healthcare applications. *Engineering, Technology & Applied Science Research*, *11*(4), 7326-7331.
6. Nagesh, N., Patil, P., Patil, S., & Kokatanur, M. (2022). An architectural framework for automatic detection of autism using deep convolution networks and genetic algorithm. *International Journal of Electrical & Computer Engineering* (2088-8708), *12*(2).
7. Ahmed, S. T., Sreedhar Kumar, S., Anusha, B., Bhumika, P., Gunashree, M., & Ishwarya, B. (2018, November). A Generalized Study on Data Mining and Clustering Algorithms. In *International Conference On Computational Vision and Bio Inspired Computing* (pp. 1121-1129). Springer, Cham.
8. Gibson, S., Issac, B., Zhang, L., & Jacob, S. M. (2020). Detecting spam email with machine learning optimized with bio-inspired metaheuristic algorithms. *IEEE Access*, *8*, 187914-187932.
9. Liu, X., Lu, H., & Nayak, A. (2021). A spam transformer model for SMS spam detection. *IEEE Access*, *9*, 80253-80263.
10. Ahmed, S. T., Kumar, V. V., Singh, K. K., Singh, A., Muthukumar, V., & Gupta, D. (2022). 6G enabled federated learning for secure IoMT resource recommendation and propagation analysis. *Computers and Electrical Engineering*, *102*, 108210.
11. Agboola, O. (2022). Spam Detection Using Machine Learning and Deep Learning.
12. Nagashree, N., Patil, P., Patil, S., & Kokatanur, M. (2022). InvCos curvature patch image registration technique for accurate segmentation of autistic brain images. In *Soft Computing and Signal Processing* (pp. 659-666). Springer, Singapore.