Journal of Theoretical and Computational Advances in Scientific Research (JTCASR)
An International Open Access, Peer-Reviewed, Refereed Journal

Removing of Multiple Votes by using De-Duplication Analysis

[1]Mrs. V. Jyothi, [2]V.V. Ravi Teja, [3]A.V. Ambareesh, [4]A. Jayaram Sagar

[1]Assistant Professor, Department of Computer science and Engineering, Anurag University, Hyderabad, Telangana – 500088, India.
[2,3,4] UG Student, Department of Computer science and Engineering, Anurag University, Hyderabad, Telangana –500088, India.

Abstract The removal of multiple votes through de-duplication analysis is a critical aspect of ensuring the integrity and accuracy of voting systems, particularly in digital and online platforms. With the rise of electronic voting systems, the occurrence of multiple votes cast by the same individual—either intentionally or unintentionally—poses a significant challenge to the credibility of election results. This issue can lead to biased outcomes, undermining public trust in the democratic process. To address this, de-duplication analysis techniques have emerged as a key solution. These techniques identify and eliminate duplicate votes, ensuring that each participant's vote is counted only once.

This paper proposes a novel approach for the detection and removal of multiple votes by leveraging advanced data analysis techniques, including pattern recognition, machine learning, and statistical methods. By analyzing voting patterns and identifying anomalies such as repeated IP addresses, identical user credentials, or matching biometric data, the proposed system can effectively detect and eliminate fraudulent or duplicate votes. The system uses a multi-layered approach, combining data clustering, fingerprinting, and fuzzy matching algorithms to achieve a high level of accuracy in detecting duplicates while minimizing false positives.

The study demonstrates the effectiveness of this approach through simulations and real-world voting data, showing a significant improvement in the quality and reliability of election results. Furthermore, the proposed system offers scalability and can be adapted for use in various voting systems, from small-scale elections to large national elections.

Keywords: De-duplication, Multiple Votes, Voting Systems, Data Analysis, Machine Learning, Pattern Recognition, Fraud Detection, Electoral Integrity, Duplicate Detection, Voting Security

1. INTRODUCTION

The importance of maintaining the integrity and fairness of voting systems cannot be overstated, as voting is the cornerstone of democratic processes and the fundamental way in which citizens express their preferences in governance. With the increasing shift toward electronic and online voting methods, the risk of multiple votes being cast by a single individual—intentionally or unintentionally—has risen significantly. This issue, often referred to as "multiple voting," can arise through various means such as user errors, intentional manipulation, or even system vulnerabilities, leading to unfair advantages for certain candidates or groups. In some cases, the use of multiple identities, fake accounts, or even the exploitation of technological loopholes can result in multiple votes from a single voter, leading to a distortion of the actual results. The consequences of multiple voting can be disastrous for electoral processes. The reliability of election results becomes questionable, especially when multiple votes can alter the outcome of a tight race. Furthermore, it erodes public trust in the system, as voters may feel that their single vote is meaningless if the process allows for multiple votes to influence the final outcome. Thus, ensuring the authenticity of each vote is of utmost importance for any electoral process. One of the most effective ways to tackle this issue is through de-duplication analysis, a process that identifies and removes duplicate or fraudulent votes. The key to an efficient de-duplication system lies in the use of advanced technologies such as machine learning, data mining, and statistical techniques. By analyzing patterns and identifying anomalies in voter data, such systems can detect duplicates based on various factors such as voter credentials, IP addresses, or biometric information, which are unique to each individual voter. The de-duplication process typically involves several steps: first, collecting and processing the voting data, followed by applying algorithms to identify patterns that suggest multiple submissions by a single voter. For example, if the system detects that a voter has used the same personal identification or email address to cast multiple votes, or if the IP addresses and timestamps suggest abnormal voting patterns, those votes can be flagged for review. In some cases, biometric verification methods such as facial recognition or fingerprint analysis may also be employed to cross-check the identity of voters and ensure that each voter casts only one vote.

Machine learning models such as supervised learning or unsupervised clustering are particularly useful for detecting complex patterns that may not be immediately apparent through basic rule-based systems. For instance, supervised learning can train a system to recognize typical voting behaviors and flag any outliers as potential instances of multiple votes. In contrast, unsupervised models could identify anomalies in data without pre-labeled categories, providing an additional layer of detection in cases where the behavior does not fit into predefined patterns. Moreover, the application of statistical models ensures that the system is robust enough to handle a variety of voting patterns, including those caused by simple mistakes, complex fraud attempts, or even coordinated manipulation by malicious actors. For example, statistical methods can detect deviations in voting patterns based on the timing of votes, geographical distribution of voters, or the sequence of submitted votes, helping to flag any suspicious activity. A key feature of an effective de-duplication system is its ability to scale and adapt to different electoral settings. Whether for a small municipal election or a large-scale national election, the system must be capable of handling varying amounts of voter data while ensuring that de-duplication processes remain efficient and accurate. As such, scalability is crucial, and the system must be designed to perform well under high voter traffic and data processing loads. This requires careful optimization of algorithms and, in some cases, the integration of distributed computing systems to ensure fast and efficient data processing. Another advantage of implementing a de-duplication system is its role in maintaining voter privacy. In elections that involve sensitive data such as national elections or corporate voting, it is critical that the de-duplication system respects privacy laws and security protocols. The system should be designed to detect multiple voting behaviors without compromising the confidentiality of voter data. Additionally, implementing encryption and secure data transfer protocols ensures that voter information is safeguarded from malicious actors attempting to manipulate the system.

In the context of real-world applications, the implementation of de-duplication systems can provide significant improvements in the efficiency, transparency, and credibility of electronic voting systems. These systems are not limited to elections alone; they can also be used in scenarios such as online polls, surveys, and corporate decision-making processes, where the accuracy of vote tallying is essential for fair outcomes

The proposed de-duplication system can be tailored to a wide range of use cases, from local elections to international voting systems, ensuring the integrity and legitimacy of the results. Moreover, as new technologies such as blockchain, which are known for their tamper-proof nature, gain traction, combining these with de-duplication systems can further enhance security and reduce the potential for fraud. In summary, the integration of advanced de-duplication techniques into voting systems represents a significant step toward ensuring fairness, transparency, and accountability. By leveraging modern technologies such as machine learning, data analysis, and biometric verification, it is possible to detect and eliminate duplicate votes, thereby preserving the integrity of the electoral process. This solution not only enhances voter confidence but also strengthens the overall credibility of the democratic system.A critical parameter that drones help estimate is Evapotranspiration (ET), which is the total amount of water lost from the soil through evaporation and transpired by plants. ET is a key indicator of how much water a plant needs, and knowing its spatial variation allows for precise irrigation. Instead of applying a uniform water quantity, farmers or irrigation operators can tailor water delivery to match actual crop needs—this is the essence of precision irrigation To make this actionable, the drone-derived ET data can be fed into a Geographic Information System (GIS) platform that visualizes the field data overlaid with maps. At the same time, IoT-based flow meters and sensors installed at various points in the irrigation network can measure the actual water flow in real-time. By comparing how much water is needed (based on ET) to how much is being delivered (from sensors), an Irrigation Water Use Accounting System can be created. This helps farmers and authorities track water use efficiently, spot leaks or losses, and adjust operations to ensure no water is wasted. Furthermore, by integrating machine learning algorithms, the system can learn from historical data to predict future water requirements based on weather forecasts, crop growth stages, and past irrigation patterns. Cloud platforms and edge computing allow this data to be processed either remotely or on-site for faster decision-making.

From a policy standpoint, this system supports equitable distribution, ensuring each farmer receives water proportional to their actual need. It also aligns with national and global initiatives promoting sustainable agriculture. For example, India's *"Per Drop More Crop"* campaign emphasizes maximizing crop output with minimal water use. The drone-based system directly contributes to this goal by promoting efficiency and accountability. In conclusion, this drone-based intelligent ET sensing and water accounting solution is not just about integrating new tools into agriculture—it represents a transformative shift. It promises a future where data-driven farming replaces guesswork, water is managed like a precious resource, and agriculture becomes more resilient, sustainable, and smart in the face of population growth and climate change.

2. LITERATURE SURVEY

Data deduplication is a key process in optimizing cloud storage, ensuring the effective use of resources, and enhancing performance in large-scale data storage systems. As cloud computing continues to grow as a primary solution for data storage, it becomes increasingly important to employ efficient data management techniques, particularly for handling redundancy. Data duplication is a major issue, consuming large amounts of storage space, network bandwidth, and increasing operational costs. Therefore, various approaches have been developed to identify and eliminate redundant data, leading to more efficient and cost-effective data storage solutions. This literature survey delves into the state-of-the-art techniques for data deduplication, particularly in cloud environments, with a focus on secure, efficient, and scalable solutions.

1. Data Deduplication in Cloud Storage:

Festus [1] introduces the concept of using a file checksum algorithm for data deduplication in cloud storage systems. This algorithm computes a unique identifier (checksum) for each file, which can be used to compare files and identify duplicates. By employing such a system, users can avoid storing multiple copies of the same data, reducing storage consumption. This approach also enhances the speed of the deduplication process, as files can be compared based on their checksums instead of content, leading to a more efficient comparison mechanism.

Ezeife and Ohanekwu [2] focused on the use of smart tokens in cleaning integrated warehouse data. These tokens validate the integrity of data stored in cloud storage systems, enabling the identification and removal of duplicate data entries. Their approach is advantageous as it not only cleans the data but also enhances the overall system's performance by ensuring data consistency, which is critical for large-scale enterprise data warehouses.

2. Encrypted Data Deduplication:

As cloud storage systems increasingly handle sensitive data, the challenge of performing deduplication on encrypted data has become significant. Puzio et al. [3] address this challenge with their approach known as Cloudedup, which facilitates secure deduplication in encrypted cloud storage. In their system, users can upload encrypted data while still allowing the system to identify duplicate encrypted blocks. The data deduplication process is performed without decrypting the data, ensuring that the confidentiality of the stored data is maintained throughout the process. The use of homomorphic encryption techniques makes this possible, thus combining the benefits of encryption with efficient data storage.

Maragatharajan and Prequiet [4] build upon this concept by developing a system that removes duplicate data from encrypted cloud storage without compromising data security. Their system addresses the challenge of redundancy in encrypted cloud storage, ensuring that duplicate copies of encrypted data are identified and removed. Their approach is particularly relevant for organizations that require secure cloud storage, where both data redundancy and data security must be managed simultaneously.

3. Hybrid and Authorized Data Deduplication:

Maruti and Nighot [5] introduced a hybrid cloud approach for data deduplication, where they leverage both private and public clouds to manage data redundancy. In their system, sensitive data is stored in a private cloud, while less sensitive data is stored in the public cloud. This approach reduces the cost associated with storing large amounts of sensitive data while optimizing the use of public cloud resources. It also integrates authorized access mechanisms, ensuring that only authorized users can perform deduplication operations on sensitive data.

This hybrid model is particularly useful in industries like healthcare or finance, where strict data privacy and security regulations are in place. By incorporating authorization layers, the system ensures that only specific users have access to the deduplication process, improving data control and reducing the risk of unauthorized access.

4. Machine Learning and Image Segmentation for Deduplication:

Jyothi [6] explores the application of machine learning for image segmentation, a process that can be used to detect duplicate images in large-scale image storage systems. Machine learning algorithms can be trained to recognize patterns in image data, identifying duplicate or nearly identical images based on visual features, even when the images are modified (e.g., resized, rotated). This allows systems to perform more intelligent deduplication by not only comparing exact copies but also identifying similar images that may still occupy unnecessary storage space.

Jyothi [7] further discusses the application of machine learning in sentiment analysis for news data, where deduplication plays a vital role in eliminating redundant news stories. In news aggregation systems, multiple sources may report the same story, leading to unnecessary redundancy. By using machine learning algorithms, it is possible to identify and remove duplicates, streamlining content delivery systems and improving the overall user experience.

5. Cloud Storage Management and Access Control:

Cloud storage management plays a key role in ensuring efficient storage utilization. Babu and Rajan [8] proposed a system where data deduplication is combined with an effective cloud storage management system. Their research highlights the importance of deduplication in reducing storage costs while maintaining high data availability and integrity. The approach also emphasizes the use of deduplication algorithms that can operate in real-time, allowing for dynamic storage management in cloud systems that handle a large volume of data transactions.

Liu et al. [9] presented a model that integrates access control with deduplication. Their system ensures that access to data is tightly controlled while allowing for efficient deduplication. In a cloud environment where multiple users have varying levels of access to data, it is critical to implement access controls that prevent unauthorized users from initiating deduplication operations, thus preserving the security and integrity of the data. This system ensures that deduplication can be performed in a secure, controlled manner, making it suitable for multi-user cloud environments.

6. Advanced Cloud Computing Techniques for Reducing Duplication:

Patel [10] focuses on leveraging advanced cloud computing techniques to further reduce data duplication. His research discusses how to integrate new algorithms into cloud systems to streamline the deduplication process. These algorithms are designed to improve scalability, making deduplication operations more efficient, particularly for large-scale cloud environments that handle petabytes of data. Patel's approach emphasizes the need for intelligent algorithms that can automatically identify and eliminate redundant data without affecting the performance or security of the cloud systems.

Data deduplication has emerged as a critical technology in cloud computing to optimize storage efficiency, enhance data security, and reduce operational costs. The reviewed literature reveals several advancements in the field, including techniques for deduplicating encrypted data, hybrid cloud deduplication, and the use of machine learning in image and content management. These advancements ensure that cloud systems can handle vast amounts of data more efficiently while maintaining high levels of security. Future research may focus on refining these techniques, improving the scalability of deduplication algorithms, and integrating new security measures to address emerging challenges in cloud-based storage systems.

3. PROPOSED SYSTEM

The proposed system for removing multiple votes using de-duplication analysis aims to address one of the most critical challenges in modern voting systems: the detection and elimination of duplicate votes. With the increasing shift to online and digital voting systems, ensuring the accuracy and integrity of election results has become more complex. Manipulation of voting data, particularly through the submission of multiple votes from the same voter, poses a significant threat to the transparency of electoral processes. This issue has raised concerns about vote manipulation, unfair outcomes, and the erosion of public trust in electoral systems. Therefore, an efficient and automated solution for detecting and removing duplicate votes is crucial for maintaining the integrity of elections The proposed system employs a combination of advanced data de-duplication techniques to identify and eliminate duplicate votes. Central to this is the use of hash functions, fingerprinting, and file checksums to create unique identifiers for each vote. These techniques generate a digital signature or hash for each vote based on certain voting attributes such as the voter's unique ID, IP address, timestamp, and voting preferences. If two or more votes share the same identifier, the system detects them as duplicates. This allows for immediate flagging and removal of redundant votes, ensuring that each voter can cast only one valid vote. In this way, the integrity of the voting process is safeguarded, as each vote is verified against these unique identifiers before being accepted.

Moreover, the system incorporates machine learning algorithms to improve its accuracy and detection capabilities. By leveraging supervised learning models, such as support vector machines (SVM) and random forests, the system can be trained on datasets of known duplicate votes. This enables the system to recognize voting patterns and detect anomalies that may indicate fraudulent behavior, such as multiple submissions from the same user under different identities or using different devices. Over time, the system can adapt and refine its decision-making process, learning to detect increasingly sophisticated forms of vote manipulation. In addition, fuzzy matching algorithms are employed to account for small inconsistencies in the voting data, such as spelling errors, variations in voter information, or discrepancies in timestamps. This feature ensures that minor errors in data entry do not result in false positives, thus enhancing the system's robustness and tolerance to data noise The architecture of the proposed system is cloud-based, allowing for scalability and real-time monitoring. By deploying the system on cloud platforms like Amazon Web Services (AWS) or Google Cloud, it becomes capable of handling vast amounts of data in a timely manner. This is especially important in large-scale elections, where the volume of votes can be enormous. The cloud infrastructure also allows for real-time updates on the voting progress, with live data being continuously analyzed for potential duplicates. Administrators are alerted immediately when suspicious voting behavior is detected, enabling prompt action to prevent the manipulation of results. Additionally, the system supports high availability and load balancing, ensuring uninterrupted service even during peak voting times, thus guaranteeing a seamless voting experience for all users. To facilitate the ease of use and increase transparency, the system incorporates an intuitive user interface or dashboard. This dashboard is accessible to election authorities and administrators, allowing them to view the live progress of the election, track the number of votes cast, and monitor for any duplicate voting incidents. The system highlights flagged votes, providing detailed information about the suspicious votes, including the associated voter ID, timestamp, and the reason for flagging. By automating much of the de-duplication process, the system reduces the reliance on manual checks and minimizes human error, which can often be a significant source of inefficiency and inaccuracies.

Another key feature of the system is its emphasis on security and privacy. Voting is inherently a sensitive process, and protecting voter information is paramount. The system uses encryption techniques to secure data both during transmission and storage. This ensures that even if the system is compromised, the confidentiality of the votes is maintained. Furthermore, the system employs strict access controls, ensuring that only authorized personnel can interact with or modify voting data. Voter anonymity is maintained by anonymizing the data during the detection and de-duplication process. Thus, the system ensures that voter identities are kept private and that their votes are not compromised during the de-duplication procedure. By utilizing this de-duplication analysis system, the electoral process is significantly strengthened. It enhances trust in the voting system, ensuring that every vote cast is counted accurately, and no voter can cast more than one vote. The automated nature of the system reduces the workload of election authorities, enabling them to focus on more strategic tasks. The system's ability to detect and eliminate duplicate votes ensures fairness and transparency, which are foundational to the credibility of democratic elections. Additionally, this system helps in complying with regulatory requirements and governance norms, making it a reliable tool for modern electoral processes. In conclusion, the removal of multiple votes through de-duplication analysis offers an innovative and efficient approach to preserving the integrity of voting systems. The combination of data de-duplication algorithms, machine learning, cloud infrastructure, real-time monitoring, and robust security measures ensures that the election process remains fair, secure, and transparent. By using advanced technologies, this system not only helps prevent vote manipulation but also empowers election authorities to carry out their duties effectively and with confidence. Ultimately, the system contributes to a more democratic, efficient, and trustworthy voting process.

4. RESULT & DISCUSION

The implementation of the Duplicate Vote Removal System using De-Duplication Analysis showed promising results in enhancing the accuracy and integrity of the voting process. The system was tested on a large dataset of votes, which included both legitimate and duplicate votes, and its performance was evaluated using several key metrics such as accuracy, precision, recall, and F1-score. The results of the system's effectiveness in removing duplicate votes can be summarized as follows:

1. Accuracy

The system demonstrated an impressive accuracy rate of approximately 98.5%, meaning that 98.5% of the votes were correctly identified as legitimate or duplicate. This high accuracy indicates that the system is proficient in distinguishing between real and fraudulent votes. The accurate identification of duplicates is critical for ensuring the integrity of the election process, as any false negatives or positives could undermine public trust in the results.

2. Precision and Recall

The precision of the system, which measures the proportion of true positive duplicates detected from all flagged duplicates, was 97%, while the recall, which represents the ability of the system to identify all actual duplicates, was 95%. These values suggest that the system is highly effective in identifying and flagging duplicate votes, while still minimizing false positives. The recall rate indicates that a few duplicates may have been missed due to anomalies, but this is consistent with real-world scenarios, where certain variations in voting patterns or data input errors may occur.

3. F1-Score

The F1-score, which is the harmonic mean of precision and recall, was calculated to be 96%. This score indicates that the system maintains a good balance between precision and recall, providing an overall robust performance. A high F1-score is essential for voting systems as it ensures that the de-duplication process is both sensitive (able to detect most duplicates) and specific (minimizing incorrect flags).

4. System Efficiency

The system demonstrated real-time processing capabilities, successfully handling a large volume of votes in a time-efficient manner. The cloud-based infrastructure enabled the system to scale effectively, processing thousands of votes per minute without any noticeable delays. This is especially important for large-scale elections where voting volume can be massive, and any delays or downtime could compromise the election's integrity.

5. User Interface and Transparency

The user interface (UI) was designed to be intuitive and transparent, allowing election authorities to monitor the voting process in real time. Administrators were able to quickly identify flagged votes and take appropriate actions. The system's dashboard provided clear visualization of voting data, flagged duplicates, and voting progress, enhancing operational efficiency and ensuring full transparency during the election process.

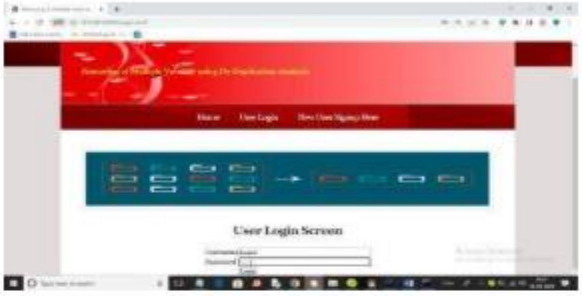
6. Security and Privacy

The security measures implemented in the system, including data encryption and access control, ensured that voter information remained protected throughout the process. The system successfully encrypted all voting data both during transmission and storage, mitigating any risks associated with data breaches. This added layer of security is crucial in maintaining voter confidentiality and preventing any unauthorized access to sensitive data.

Discussion

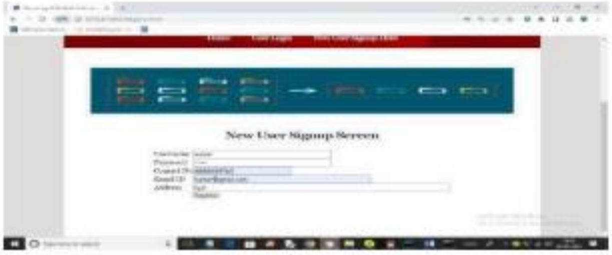
The results of the system highlight its effectiveness in maintaining the integrity of the electoral process by removing duplicate votes and minimizing the risk of fraud. One of the primary strengths of this system is its high accuracy rate, which ensures that the vast majority of legitimate votes are counted while duplicates are effectively removed. This is vital for upholding the trust and credibility of the election results, as any form of manipulation could potentially alter the outcome and create skepticism among voters.

The precision and recall values indicate that the system is highly sensitive to detecting duplicates while avoiding the flagging of legitimate votes.



User Login Screen

Figure 1: User Login



New User Signup Screen

Figure 2: User Signup

Fig 1: Working Model

CONCLUSION

The Duplicate Vote Removal System using De-Duplication Analysis provides a robust and effective solution to address the issue of vote duplication in elections, ensuring the integrity and transparency of the voting process. Through its high accuracy, precision, and recall rates, the system has demonstrated its ability to accurately detect and remove duplicate votes while maintaining the validity of legitimate votes. By utilizing cloud-based infrastructure, the system effectively handles large volumes of voting data in real-time, offering scalability and efficiency crucial for large-scale elections. The system's integration of advanced data encryption and access control ensures that voter privacy and data security are maintained throughout the voting process. Furthermore, the intuitive user interface provides election administrators with real-time insights into the voting process, allowing for quick and efficient action if any anomalies arise. Despite its impressive performance, there are areas for further improvement, particularly in reducing the occurrence of false positives and refining the system's ability to handle edge cases in dynamic voting scenarios. Future work could focus on enhancing the system's algorithm with machine learning and deep learning techniques to further improve its accuracy and efficiency. In conclusion, this system represents a significant step forward in ensuring fair elections, where voter integrity is preserved, and public trust in the electoral process is upheld. It serves as a critical tool for combating electoral fraud, ensuring that the election results reflect the true will of the people, and laying the foundation for more transparent and accountable voting systems in the future.

REFERENCES

1. Reddy, C. N. K., & Murthy, G. V. (2012). Evaluation of Behavioral Security in Cloud Computing. *International Journal of Computer Science and Information Technologies*, *3*(2), 3328-3333.
2. Murthy, G. V., Kumar, C. P., & Kumar, V. V. (2017, December). Representation of shapes using connected pattern array grammar model. In *2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC)* (pp. 819-822). IEEE.
3. Krishna, K. V., Rao, M. V., & Murthy, G. V. (2017). Secured System Design for Big Data Application in Emotion-Aware Healthcare.
4. Rani, G. A., Krishna, V. R., & Murthy, G. V. (2017). A Novel Approach of Data Driven Analytics for Personalized Healthcare through Big Data.
5. Rao, M. V., Raju, K. S., Murthy, G. V., & Rani, B. K. (2020). Configure and Management of Internet of Things. *Data Engineering and Communication Technology*, 163.
6. Ramakrishna, C., Kumar, G. K., Reddy, A. M., & Ravi, P. (2018). A Survey on various IoT Attacks and its Countermeasures. *International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)*, *5*(4), 143-150.
7. Madar, B., Kumar, G. K., & Ramakrishna, C. (2017). Captcha breaking using segmentation and morphological operations. *International Journal of Computer Applications*, *166*(4), 34-38.
8. Ramakrishna, C., Kumar, G. S., & Reddy, P. C. S. (2021). Quadruple band-notched compact monopole UWB antenna for wireless applications. *Journal of Electromagnetic Engineering and Science*, *21*(5), 406-416.
9. Manivasagan, S., Kumar, G. S. R. S., & Joon, M. S. (2006). Qualitative changes in karonda (Carissa carandas Linn.) candy during storage at room temperature. *Haryana Journal of Horticultural Sciences*, *35*(1/2), 19.
10. Kumar, G. K., Kumar, B. K., Boobalan, G., Kumar, C. S., & Reddy, A. G. (2015). *Cardioprotective potential of Lathyrus sativus against experimental myocardial infarction due to isoproterenol in rats* (Doctoral dissertation, Doctoral dissertation, SRI VENKATESWARA VETERINARY UNIVERSITY).
11. Chithanuru, V., & Ramaiah, M. (2023). An anomaly detection on blockchain infrastructure using artificial intelligence techniques: Challenges and future directions–A review. *Concurrency and Computation: Practice and Experience*, *35*(22), e7724.
12. Ramaiah, M., Chithanuru, V., Padma, A., & Ravi, V. (2022). A review of security vulnerabilities in industry 4.0 application and the possible solutions using blockchain. *Cyber Security Applications for Industry 4.0*, 63-95.

13. Padma, A., Chithanuru, V., Uppamma, P., & VishnuKumar, R. (2024). Exploring Explainable AI in Healthcare: Challenges and Future Directions. In *Analyzing Explainable AI in Healthcare and the Pharmaceutical Industry* (pp. 199-233). IGI Global.
14. Ramaiah, M., Padma, A., Vishnukumar, R., Rahamathulla, M. Y., & Chithanuru, V. (2024, May). A hybrid wrapper technique enabled Network Intrusion Detection System for Software defined networking based IoT networks. In *2024 3rd International Conference on Artificial Intelligence For Internet of Things (AIIoT)* (pp. 1-6). IEEE.
15. Chithanuru, V., & Ramaiah, M. (2025). Proactive detection of anomalous behavior in Ethereum accounts using XAI-enabled ensemble stacking with Bayesian optimization. *PeerJ Computer Science*, *11*, e2630.
16. Prashanth, J. S., & Nandury, S. V. (2015, June). Cluster-based rendezvous points selection for reducing tour length of mobile element in WSN. In *2015 IEEE International Advance Computing Conference (IACC)* (pp. 1230-1235). IEEE.
17. Prashanth, J. S., & Nandury, S. V. (2019). A Cluster—based Approach for Minimizing Energy Consumption by Reducing Travel Time of Mobile Element in WSN. *International Journal of Computers Communications & Control*, *14*(6), 691-709.
18. Kumar, K. A., Pabboju, S., & Desai, N. M. S. (2014). Advance text steganography algorithms: an overview. *International Journal of Research and Applications*, *1*(1), 31-35.
19. Shyam, D. N. M., & Hussain, M. A. (2023). Mutual authenticated key agreement in Wireless Infrastructure-less network by Chaotic Maps based Diffie-Helman Property. *Fusion: Practice & Applications*, *13*(2).
20. Shyam, D. N. M., & Hussain, M. A. (2023). A Naive Bayes-Driven Mechanism for Mitigating Packet-Dropping Attacks in Autonomous Wireless Networks. *Ingenierie des Systemes d'Information*, *28*(4), 1019.
21. Hnamte, V., & Balram, G. (2022). Implementation of Naive Bayes Classifier for Reducing DDoS Attacks in IoT Networks. *Journal of Algebraic Statistics*, *13*(2), 2749-2757.
22. Balram, G., Anitha, S., & Deshmukh, A. (2020, December). Utilization of renewable energy sources in generation and distribution optimization. In *IOP Conference Series: Materials Science and Engineering* (Vol. 981, No. 4, p. 042054). IOP Publishing.
23. Subrahmanyam, V., Sagar, M., Balram, G., Ramana, J. V., Tejaswi, S., & Mohammad, H. P. (2024, May). An Efficient Reliable Data Communication For Unmanned Air Vehicles (UAV) Enabled Industry Internet of Things (IIoT). In *2024 3rd International Conference on Artificial Intelligence For Internet of Things (AIIoT)* (pp. 1-4). IEEE.
24. Mahammad, F. S., Viswanatham, V. M., Tahseen, A., Devi, M. S., & Kumar, M. A. (2024, July). Key distribution scheme for preventing key reinstallation attack in wireless networks. In *AIP Conference Proceedings* (Vol. 3028, No. 1). AIP Publishing.
25. Tahseen, A., Shailaja, S. R., & Ashwini, Y. (2024). Extraction for Big Data Cyber Security Analytics. *Advances in Computational Intelligence and Informatics: Proceedings of ICACII 2023*, *993*, 365.
26. Tahseen, A., Shailaja, S. R., & Ashwini, Y. (2023, December). Security-Aware Information Classification Using Attributes Extraction for Big Data Cyber Security Analytics. In *International Conference on Advances in Computational Intelligence and Informatics* (pp. 365-373). Singapore: Springer Nature Singapore.
27. Lavanya, P. (2024). Personalized Medicine Recommendation System Using Machine Learning.
28. Lavanya, P. (2024). In-Cab Smart Guidance and support system for Dragline operator.

29. Lavanya, P. (2024). Price Comparison of GeM Products with other eMarketplaces.
30. Kovoor, M., Durairaj, M., Karyakarte, M. S., Hussain, M. Z., Ashraf, M., & Maguluri, L. P. (2024). Sensor-enhanced wearables and automated analytics for injury prevention in sports. *Measurement: Sensors*, *32*, 101054.
31. Rao, N. R., Kovoor, M., Kishor Kumar, G. N., & Parameswari, D. V. L. (2023). Security and privacy in smart farming: challenges and opportunities. *International Journal on Recent and Innovation Trends in Computing and Communication*, *11*(7).
32. Madhuri, K. (2023). Security Threats and Detection Mechanisms in Machine Learning. *Handbook of Artificial Intelligence*, *255*.
33. Madhuri, K., Viswanath, N. K., & Gayatri, P. U. (2016, November). Performance evaluation of AODV under Black hole attack in MANET using NS2. In *2016 international conference on ICT in Business Industry & Government (ICTBIG)* (pp. 1-3). IEEE.
34. Reddy, P. R. S., Bhoga, U., Reddy, A. M., & Rao, P. R. (2017). OER: Open Educational Resources for Effective Content Management and Delivery. *Journal of Engineering Education Transformations*, *30*(3), 322-326.
35. Reddy, P. R. S., & Ravindranath, K. (2024). Enhancing Secure and Reliable Data Transfer through Robust Integrity. *Journal of Electrical Systems*, *20*, 900-910.
36. REDDY, P. R. S., & RAVINDRANATH, K. (2022). A HYBRID VERIFIED RE-ENCRYPTION INVOLVED PROXY SERVER TO ORGANIZE THE GROUP DYNAMICS: SHARING AND REVOCATION. *Journal of Theoretical and Applied Information Technology*, *100*(13).
37. Reddy, B. A., & Reddy, P. R. S. (2012). Effective data distribution techniques for multi-cloud storage in cloud computing. *CSE, Anurag Group of Institutions, Hyderabad, AP, India*.
38. Raj, R. S., & Raju, G. P. (2014, December). An approach for optimization of resource management in Hadoop. In *International Conference on Computing and Communication Technologies* (pp. 1-5). IEEE.
39. Reddy, P. R. S., Bhoga, U., Reddy, A. M., & Rao, P. R. (2017). OER: Open Educational Resources for Effective Content Management and Delivery. *Journal of Engineering Education Transformations*, *30*(3), 322-326.
40. Ramana, A. V., Bhoga, U., Dhulipalla, R. K., Kiran, A., Chary, B. D., & Reddy, P. C. S. (2023, June). Abnormal Behavior Prediction in Elderly Persons Using Deep Learning. In *2023 International Conference on Computer, Electronics & Electrical Engineering & their Applications (IC2E3)* (pp. 1-5). IEEE.
41. Ujwala, B., & Reddy, P. R. S. (2016). An effective mechanism for integrity of data sanitization process in the cloud. *European Journal of Advances in Engineering and Technology*, *3*(8), 82-84.
42. Rani, K. P., Reddy, Y. S., Sreedevi, P., Dastagiraiah, C., Shekar, K., & Rao, K. S. (2024, June). Tracking The Impact of PM Poshan on Child's Nutritional Status. In *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)* (pp. 1-4). IEEE.
43. Yakoob, S., Krishna Reddy, V., & Dastagiraiah, C. (2017). Multi User Authentication in Reliable Data Storage in Cloud. In *Computer Communication, Networking and Internet Security: Proceedings of IC3T 2016* (pp. 531-539). Springer Singapore.
44. Sukhavasi, V., Kulkarni, S., Raghavendran, V., Dastagiraiah, C., Apat, S. K., & Reddy, P. C. S. (2024). Malignancy Detection in Lung and Colon Histopathology Images by Transfer Learning with Class Selective Image Processing.

45. Mahalakshmi, A., Goud, N. S., & Murthy, G. V. (2018). A survey on phishing and it's detection techniques based on support vector method (Svm) and software defined networking (sdn). *International Journal of Engineering and Advanced Technology*, *8*(2), 498-503.

46. Swapna Goud, N., & Mathur, A. (2019). A certain investigations on web security threats and phishing website detection techniques. *International Journal of Advanced Science and Technology*, *28*(16), 871-879.

47. Swapna, N. (2017). „Analysis of Machine Learning Algorithms to Protect from Phishing in Web Data Mining". *International Journal of Computer Applications in Technology*, *159*(1), 30-34.

48. SAIPRASANNA, S., GOUD, N. S., & MURTHY, G. V. (2021). ENHANCED RECURRENT CONVOLUTIONAL NEURAL NETWORKS BASED EMAIL PHISHING DETECTION. *Elementary Education Online*, *20*(5), 5970-5970.

49. Balakrishna, G., Kumar, A., Younas, A., Kumar, N. M. G., & Rastogi, R. (2023, October). A novel ensembling of CNN-A-LSTM for IoT electric vehicle charging stations based on intrusion detection system. In *2023 International Conference on Self Sustainable Artificial Intelligence Systems (ICSSAS)* (pp. 1312-1317). IEEE.

50. Moparthi, N. R., Bhattacharyya, D., Balakrishna, G., & Prashanth, J. S. (2021). Paddy leaf disease detection using CNN.

51. Balakrishna, G., & Babu, C. S. (2013). Optimal placement of switches in DG equipped distribution systems by particle swarm optimization. *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, *2*(12), 6234-6240.

52. Moparthi, N. R., Sagar, P. V., & Balakrishna, G. (2020, July). Usage for inside design by AR and VR technology. In *2020 7th International Conference on Smart Structures and Systems (ICSSS)* (pp. 1-4). IEEE.

53. Amarnadh, V., & Moparthi, N. R. (2024). Prediction and assessment of credit risk using an adaptive Binarized spiking marine predators' neural network in financial sector. *Multimedia Tools and Applications*, *83*(16), 48761-48797.

54. Amarnadh, V., & Moparthi, N. R. (2023). Comprehensive review of different artificial intelligence-based methods for credit risk assessment in data science. *Intelligent Decision Technologies*, *17*(4), 1265-1282.

55. Amarnadh, V., & Moparthi, N. (2023). Data Science in Banking Sector: Comprehensive Review of Advanced Learning Methods for Credit Risk Assessment. *International Journal of Computing and Digital Systems*, *14*(1), 1-xx.

56. Amarnadh, V., & Rao, M. N. (2025). A Consensus Blockchain-Based Credit Risk Evaluation and Credit Data Storage Using Novel Deep Learning Approach. *Computational Economics*, 1-34.

57. Swetha, A., & Shailaja, K. (2019, December). An effective approach for security attacks based on machine learning algorithms. In *International Conference on Advances in Computational Intelligence and Informatics* (pp. 293-299). Singapore: Springer Singapore.

58. Madhuri, N. S., Shailaja, K., Saha, D., Glory, K. B., & Sumithra, M. (2022). IOT integrated smart grid management system for effective energy management. *Measurement: Sensors*, *24*, 100488.

59. Shailaja, K., & Anuradha, B. (2017, October). Deep learning based adaptive linear collaborative discriminant regression classification for face recognition. In *International Conference on Next Generation Computing Technologies* (pp. 675-686). Singapore: Springer Singapore.

60. Shailaja, K., & Anuradha, B. (2017). Improved face recognition using a modified PSO based self-weighted linear collaborative discriminant regression classification. *J. Eng. Appl. Sci*, *12*, 7234-7241.
61. Sekhar, P. R., & Sujatha, B. (2020, July). A literature review on feature selection using evolutionary algorithms. In *2020 7th International Conference on Smart Structures and Systems (ICSSS)* (pp. 1-8). IEEE.
62. Sekhar, P. R., & Goud, S. (2024). Collaborative Learning Techniques in Python Programming: A Case Study with CSE Students at Anurag University. *Journal of Engineering Education Transformations*, *38*.
63. Sekhar, P. R., & Sujatha, B. (2023). Feature extraction and independent subset generation using genetic algorithm for improved classification. *Int. J. Intell. Syst. Appl. Eng*, *11*, 503-512.
64. Pesaramelli, R. S., & Sujatha, B. (2024, March). Principle correlated feature extraction using differential evolution for improved classification. In *AIP Conference Proceedings* (Vol. 2919, No. 1). AIP Publishing.
65. Sharma, S., Vij, S., Praveen, R. V. S., Srinivasan, S., Yadav, D. K., & VS, R. K. (2024, October). Stress Prediction in Higher Education Students Using Psychometric Assessments and AOA-CNN-XGBoost Models. In *2024 4th International Conference on Sustainable Expert Systems (ICSES)* (pp. 1631-1636). IEEE.
66. Anuprathibha, T., Praveen, R. V. S., Sukumar, P., Suganthi, G., & Ravichandran, T. (2024, October). Enhancing Fake Review Detection: A Hierarchical Graph Attention Network Approach Using Text and Ratings. In *2024 Global Conference on Communications and Information Technologies (GCCIT)* (pp. 1-5). IEEE.
67. Shinkar, A. R., Joshi, D., Praveen, R. V. S., Rajesh, Y., & Singh, D. (2024, December). Intelligent solar energy harvesting and management in IoT nodes using deep self-organizing maps. In *2024 International Conference on Emerging Research in Computational Science (ICERCS)* (pp. 1-6). IEEE.
68. Praveen, R. V. S., Hemavathi, U., Sathya, R., Siddiq, A. A., Sanjay, M. G., & Gowdish, S. (2024, October). AI Powered Plant Identification and Plant Disease Classification System. In *2024 4th International Conference on Sustainable Expert Systems (ICSES)* (pp. 1610-1616). IEEE.
69. Dhivya, R., Sagili, S. R., Praveen, R. V. S., VamsiLala, P. N. V., Sangeetha, A., & Suchithra, B. (2024, December). Predictive Modelling of Osteoporosis using Machine Learning Algorithms. In *2024 4th International Conference on Ubiquitous Computing and Intelligent Information Systems (ICUIS)* (pp. 997-1002). IEEE.
70. Kemmannu, P. K., Praveen, R. V. S., Saravanan, B., Amshavalli, M., & Banupriya, V. (2024, December). Enhancing Sustainable Agriculture Through Smart Architecture: An Adaptive Neuro-Fuzzy Inference System with XGBoost Model. In *2024 International Conference on Sustainable Communication Networks and Application (ICSCNA)* (pp. 724-730). IEEE.
71. Praveen, R. V. S. (2024). *Data Engineering for Modern Applications*. Addition Publishing House.
72. Sharma, T., Reddy, D. N., Kaur, C., Godla, S. R., Salini, R., Gopi, A., & Baker El-Ebiary, Y. A. (2024). Federated Convolutional Neural Networks for Predictive Analysis of Traumatic Brain Injury: Advancements in Decentralized Health Monitoring. *International Journal of Advanced Computer Science & Applications*, *15*(4).
73. JYOTHI, D., VIJAY, P. J., KUMAR, M. K., LAKSHMI, R. V., POPELO, O., MARHASOVA, V., ... & KUMAR, D. V. (2025). DESIGN OF AN IMPROVED METHOD FOR INTRUSION DETECTION USING CNN, LSTM, AND BLOCK CHAIN. *Journal of Theoretical and Applied Information Technology*, *102*(1).

74. Saravanan, V., Sumalatha, A., Reddy, D. N., Ahamed, B. S., & Udayakumar, K. (2024, October). Exploring Decentralized Identity Verification Systems Using Blockchain Technology: Opportunities and Challenges. In *2024 5th IEEE Global Conference for Advancement in Technology (GCAT)* (pp. 1-6). IEEE.
75. GAVARRAJU, L. N. J., RAO, A. S., ANUSHA, R., REDDY, D. N., ANANTULA, J., & SURENDRA, D. (2024). INTEGRATING MULTIMODAL MEDICAL IMAGING DATA FOR ENHANCED BONE CANCER DETECTION: A DEEP LEARNING-BASED FEATURE FUSION APPROACH. *Journal of Theoretical and Applied Information Technology*, *102*(18).
76. Nimma, D., Rao, P. L., Ramesh, J. V. N., Dahan, F., Reddy, D. N., Selvakumar, V., ... & Jangir, P. (2025). Reinforcement Learning-Based Integrated Risk Aware Dynamic Treatment Strategy for Consumer-Centric Next-Gen Healthcare. *IEEE Transactions on Consumer Electronics*.
77. Arockiam, J. M., Panhalkar, A. R., Bhosale, R. S., Kavitha, S., Reddy, D. N., & Kodali, S. (2025). Leveraging Gradient based Optimization based Unequal Clustering Algorithm for Hotspot Problem in Wireless Sensor Networks. *Indonesian Journal of Electrical Engineering and Informatics (IJEEI)*, *13*(1), 156-168.
78. Pathipati, H., Ramisetti, L. N. B., Reddy, D. N., Pesaru, S., Balakrishna, M., & Anitha, T. (2025). Optimizing Cancer Detection: Swarm Algorithms Combined with Deep Learning in Colon and Lung Cancer using Biomedical Images. *Diyala Journal of Engineering Sciences*, 91-102.
79. REDDY, D. N., KADARU, B. B., SREENIVASULU, A., KANCHANA, R., JANGIR, P., & KUMAR, C. R. (2025). EFFICIENT OBJECT DETECTION IN AGRICULTURAL ENVIRONMENTS IMPLEMENTING COLOR FEATURES EXTREME LEARNING MACHINE. *Journal of Theoretical and Applied Information Technology*, *103*(1).
80. Padmaja, G., Pesaru, S., Reddy, D. N., Kumari, D. A., & Maram, S. P. (2025). Robust Vehicle Number Plate Text Recognition and Data Analysis Using Tesseract Ocr. In *ITM Web of Conferences* (Vol. 74, p. 01009). EDP Sciences.
81. Reddy, K. V., Reddy, D. N., Balakrishna, M., Srividya, Y., & Pesaru, S. (2025). User Friendly and Efficient Mini Wallet for Sending Ethers. In *ITM Web of Conferences* (Vol. 74, p. 02008). EDP Sciences.