ESP-JETA
ESP Journal of Engineering & Technology Advancements
ISSN: 2583-2646 / Volume 5 Issue 2 April 2025 / Page No: 17-23
Paper Id: JETA-V5I2P103 / Doi: 10.56472/25832646/JETA-V5I2P103

Orginal Article

The Role of Machine Learning in Big Data Analytics: Tools, Techniques, and Applications

Niravkumar Prajapati
Independent Researcher

Received Date: 07 January 2025 Revised Date: 12 February 2025 Accepted Date: 05 April 2025

Abstract: *Machine learning (ML) has shown to be a game-changer in big data analytics, allowing businesses to glean valuable insights from massive and intricate datasets. Highlighting the methods, tools, and applications that drive innovation and decision-making across sectors, this review article investigates the complementary nature regarding artificial intelligence and large analytics. Exploring the three cornerstones of under supervision, under supervision reinforcement education, and machine learning, and how each tackles issue like choosing characteristics, prioritization of data, and system adaptability. The paper reviews popular tools and frameworks, including TensorFlow, PyTorch, Hadoop, and Spark, supporting big data analytics. Furthermore, real-world applications in healthcare, finance, e-commerce, and smart cities are examined, showcasing how machine learning (ML) optimizes processes, predicts trends, and enhances user experiences. By identifying emerging trends and future research directions to really reap the rewards of enormous amounts of data, this study offers a thorough analysis of how machine learning contributes.*

Keywords: *Machine learning (ML), big data analytics, data pre-processing, supervised learning, unsupervised learning, reinforcement learning, TensorFlow, Spark, big data application, scalability, real-world applications, emerging trends.*

I. INTRODUCTION

One subfield of AI that is growing in importance is ML [1]. The goal of machine learning is to generate smart judgments and new information. A few ways to sort into. the three main types of ML algorithms are supervised, unsupervised, and semi-supervised. It is necessary to scale up ML when dealing with massive amounts of data. Tasks such as density estimation, grouping, classification, regression, etc., are some of the outputs of ML systems, which allows for another way of classifying ML. DT learning, association rule learning, ANN, clustering, SVM, BN, genetic algorithms, and other ML methodologies are all part of the toolbox.[2], etc. Machine learning is defined, its uses in Big Data discussed, and the opportunities, threats, and technological developments around machine learning in Big Data.

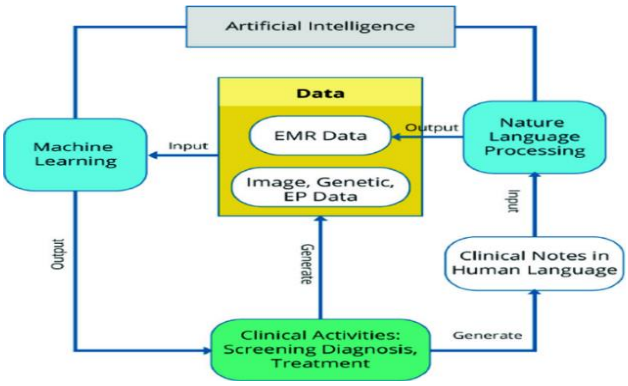

Figure 1: ML Algo for big Data Analytics

A number of procedures too tools for managing and evaluating both structured and unstructured data are developing quickly[3]. However, the majority of the tools demand a great deal of knowledge to apply and understand their concepts. This explains in detail the several open-source and corporate solutions that are now being utilized for big data analysis and learning[2]. The most popular systems, including IBM, HPE, SAPRANA, Microsoft Azure, and Oracle, are first briefly reviewed. Additionally, two open-source tools Spark MLlib and H2O have received a lot of attention.

Big data analytics may benefit from a number of interesting ML approaches, for example, learning representations [4], active instruction, simultaneous and dispersed knowledge, DL, kernel-based learning, transfer learning, and so on. In addition, new and complicated algorithms built on ML approaches are required for big data analytics to efficiently and accurately handle data in real time.

AI includes the subsets of ML and DL[5], play a vital part in utilising big data analysis's capabilities. By making it easier to collect, analyze, and understand large and complex datasets, these advanced approaches promote innovation and increase productivity across a wide range of industries.

A. Structure of the paper

The document is organized into the following sections Section II discusses methods for improving cloud security. The integration of web applications with machine learning is described in depth in Section III. Limitations and difficulties are covered in Section IV.Section V provides a review of the literature and indicates research gaps, and Section VI provides suggestions for conclusions and more study.

II. MACHINE LEARNING IN BIG DATA ANALYTICS

The study of ML explicitly focuses on the characteristics, functionality, and theory of learning algorithms and systems[6]. This field incorporates ideas from many other scientific, technological, and mathematical fields, making it incredibly interdisciplinary. These fields include AI, information theory, statistics, cognitive science, optimization theory, and optimal control. ML has been used in practically every scientific field because of its extensive variety of applications, this has profoundly impacted the fields of science and society at large.

A. Methods of Machine Learning and Big Data

The two main branches of ML are classification and regression. When the class attribute is discrete, it is called classification. When it is continuous, it is called regression. Classification techniques include k-NN, DT learning, and the NB classifier[7][8]. Regression techniques include LR and linear regression. Instances of unsupervised learning are grouped into comparable objects.

There are three main types of clustering:

- Supervised clustering: In relation to particular classes, it finds clusters with high probability densities.
- Unsupervised clustering: Two of its specialized goal functions, K-means and hierarchical clustering, take a similarity/dissimilarity metric as input and optimize intracluster similarity while minimizing intercluster similarity.
- Semi-supervised clustering: It improves clustering by using more guiding/adjusting domain information. Target variables for some of the data or pairwise restrictions between the observations might be examples of this domain knowledge.

DT uses feature values to categories data. Using a top-down greedy method, DTs are built iteratively from training data[2]. SVMs are binary classifiers that identify linear classifiers in higher dimensional feature spaces that are mapped from the original data space.

A Structure for massive amounts of data ML. The ML component is the backbone of large data ML frameworks; it interacts with user, domain, big data, and system components[9]. There are reciprocal relationships for example, Participation from the user's topic knowledge, personal preferences, and remarks on the component's usability may be utilized to enhance the learning experience. The learning component takes large data as its input and returns big data as its output[10], and by integrating educational objectives into the process for arriving at decisions, the domain may act as both a learning guide and a setting in which to apply taught models;The effectiveness and correct functioning of learning algorithms are impacted by the system architecture. Concurrently meeting the learning need could result in the system architectural being co-designed. Next, we display each component separately [11].

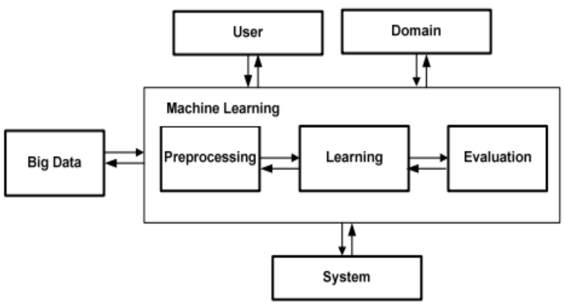

Figure1: A framework of machine learning on big data

Data Preprocessing Opportunities and Challenges Today, there are several obstacles to using big data's potential, from designing lower-layer processing systems to higher-layer analytical tools, as well as a number of unresolved issues in scientific research. Some of these issues are brought on by the nature of large data, some by the models and techniques employed for its analysis at the moment, and some by the shortcomings of the data processing systems in use nowadays[12].

- Data Complexity: Big data's diverse types, structures, and patterns increase computational challenges, Traditional methods struggle with big data due to its complexity and lack of understanding of data distribution and relationships, A complexity theory of big data is needed to simplify representation, guide model design, and improve computation.
- Computational Complexity: Big data's scale and speed exceed the capabilities of traditional computing methods[13], New paradigms and scalable algorithms are required, focusing on distributed, streaming, and data-centric approaches, Emphasis on reduction-based and sampling methods is critical for efficient computation.
- System Complexity**:** Big data systems must manage large volumes, diverse types, and real-time needs efficiently. Energy-efficient architectures and frameworks are vital, balancing throughput, accuracy, and resource optimization. Research should focus on performance evaluation, distributed systems, and iterative development to create efficient big data platforms[10].

III. TOOLS FOR MACHINE LEARNING IN BIG DATA ANALYTICS

The large information increases the complexity of data cleansing because of the sheer number and variety of sources[14]. Since AI-driven technologies that use ML algorithms to automatically identify and correct data anomalies are becoming more and more popular, techniques like data deduplication, outlier detection, and imputation for missing values are often used[15].

A. Frameworks and Platforms

Apache Spark and other modern technologies offer strong frameworks for performing complex operations on large datasets[16]. Reducing dataset dimensions while maintaining critical properties is one-way data reduction strategies try to deal with the massive amount of big data. One approach is to decrease the number of dimensions[5], SVD and PCA are two well-known dimensionality reduction techniques for feature selection and data compression. Particularly, autoencoders, a kind of neural network[17], have become useful instruments for efficient data reduction in large data settings.

Software such as Apache Beam, Flume, Samza, Ignite, and Pulsar assist data processing [18]. With Apache Atlas, data governance is controlled, guaranteeing adherence to regulations and efficient metadata administration.

B. Visualization Tools

Apache Superset and Zeppelin provide interactive analytics interfaces for visualization and exploration[18]. In the realm of frameworks for artificial intelligence and ML, such as Scikit-learn, Apache MX Net, TensorFlow, PyTorch, and Keras, are essential for creating complex models. For data analysis and visualization, data scientists depend on libraries like Pandas, NumPy, Matplotlib, Seaborn, and Bokeh[3][19].

Figure 2 : Big Data analytics tools

IV. TECHNIQUES USED IN MACHINE LEARNING FOR BIG DATA ANALYTICS

Several potential ML approaches may be applied for large data analytics, include representation learning[4], Artificial intelligence, active learning, DL, kernel-based learning, TL, DP. In addition[6], Modern, cutting-edge ML-based algorithms are essential for big data analytics since they permit aimed at the efficient in addition accurate processing of data in real-time[20][21].

Various Techniques

- The open-source framework that is widely used for creating scalability machine learning techniques is Apache Mahout. Apache Hadoop serves as the processing platform for many of its applications. Classification, clustering, and collaborative filtering methods are all part of it (recommender engines).
- An open source cluster computing technology called Apache Spark serves as the basis for the MLlib distributed machine learning platform[22]. It operates on Hadoop clusters and data and is 100 times quicker than MapReduce.

- The open-source RHadoop R For data analysis and statistics computation, an electronic infrastructure is utilized. RHadoop is a solution that uses Hadoop to connect to R.
- The distributed streaming processing method is used by the ML framework SAMOA Apache. It offers a variety of distributed streaming algorithms appropriate for typical information extracting and ML applications, with the value categorization [23], Grouping, programming abstractions, regression, and developing new methods.
- A software environment called WEKA offers algorithms for predictive modeling and data analysis.

ML is an analytics technique that teaches computers to learn from their mistakes and successes in the same way that humans do[24]. With ML algorithms, computational approaches allow them to "learn" knowledge through data instead than using a formula that has already been established [25]. The algorithms' performance adapts to increase in response to new learning data[6].

V. APPLICATIONS OF MACHINE LEARNING IN BIG DATA ANALYTICS

In a variety of fields, and related fields such as computer vision, voice processing, and NLP[26], ML techniques have made a huge difference in the fields of health, the IoT, and neuroscience[27]. ML is becoming increasingly popular in the time of large information. Big data is both a difficulty and an opportunity for machine learning algorithms to learn more about human behaviour and a variety of commercial applications. On one hand[10], Big data gives ML algorithms access to previously unheard-of levels of information, allowing them to identify patterns and create predictive models.

A. Time Series Analysis

DW-Ba-Stacking integrates ridge regression and weighted error coefficients for better traffic flow predictions. ARMA-GM-GABP combines ARMA, grey system theory, and optimized neural networks for accurate CBCFI predictions. CEEMD-MultiRocket enhances time series classification accuracy with reduced computational load. LST-GCN combines LSTM and GCN for improved spatiotemporal traffic predictions.

B. Evolutionary Computation

MSMA-SVM optimizes postgraduate employment predictions using a multi-population slime mould algorithm. RCCBOA improves SVM with a hybrid butterfly optimization strategy[28]. DECCWOA-KELM-FS predicts talent stability in higher education with high accuracy.

C. Pattern Recognition

Stock announcement events are classified with trigger words and co-occurrence analysis. UMBGN enhances multi-behavior recommendation systems. CNN-based defense mechanisms protect clinical text classification from adversarial attacks. PRI-MFC improves hierarchical clustering and fault diagnosis.

D. Computer Vision

YOLOv5 is enhanced with Dense Block and Ghost convolution for better vehicle detection[29]. Tomato leaf disease detection and maturity detection are improved using CNNs and optimized YOLOv5[30]. Attention mechanisms in YOLOv4 help detect abnormal pilot behavior [31].

E. Revolutionizing Image and Video Analysis

The field of image and video analysis has seen a radical change with the introduction of DL, especially with the use of CNN [5] DL models are used in the healthcare industry to analyze medical imaging data[32], which helps with the early identification and detection of conditions like diabetes mellitus and cancer[33].

VI. LITERATURE REVIEW

In this section, provides the previous research on Big Data Analytics and the Function of ML: Tools, Techniques, and Applications.

Das et al. (2021) This article offers a comprehensive look at the current state of ML in relation to biggest information. We go over the difficulties encountered in dealing with Out of the five, four Vs of big data. ML is promising when large-scale data processing is considered, as it has the capability of learning from the previous data and incorporating those findings on new incoming data. Although its applications in big data are limitless, it faces a vast number of challenges[2].

Srinivas Murri, Manoj Bhoyar, and Guru Prasad Selvarajan (2024). The intersection between ML and big data analytics is explained in this chapter, along with a variety of supervised and unsupervised machine learning approaches. With appropriate use cases, a number of social uses of big data—including those in health care, social analysis, finance, and security—are examined. ML algorithms come in two varieties: supervised and unsupervised. One method of supervised ML is SVM[3].

Table1: Comparative Analysis of Research Studies on Big Data Analytics for Tools, Techniques and Applications for Machine Learning

Reference	Challenge	Methodology	Key Findings	Limitations & Future Work
Das et al., (2021)	Managing ophisticated processing and large-scale data	Approaches that integrate large information with machine learning	Through the use of prior knowledge, ML facilitates the processing of massive data sets. Problems with processing efficiency and scalability plague it	Future work on making models more scalable and on processing data in real-time should be prioritised for the future.
Srinivas Murri, Manoj Bhoyar, Guru Prasad Selvarajan, (2024)	Diverse applications and data complexity	A variety of social domains have benefited from supervised and unsupervised machine learning approaches	Social applications like as healthcare, banking, and security make good use of a variety of ML methods. Key to supervised learning is support vector machines (SVMs)	Further research on applying machine learning to unstructured data and more specific social applications.
Mittal & Sangwan, (2019)	Obstacles in using conventional ML methods to huge data	Explores a range of options, including deep learning, dimensionality reduction, and parallel processing	Solutions such as parallel processing and GPUs enhance big data analytics by reducing processing time and handling high-dimensional data.	Need to explore more adaptive solutions for real-time big data applications and scalability.
Alfred, (2016)	Explanation of machine learning concepts and their use in large data	Detailed explanation of terms like AI, data mining, and business intelligence	Machine learning's ability to connect with BI and data mining makes it an indispensable component of big data analytics for decision-making	Further research required on specific industries and expanding the range of algorithms.
Khoshaba et al., (2022)	Combining and contrasting several machine learning frameworks	Comparative analysis of machine learning implementations in Apache Spark and Apache Mahout	Apache Spark and Mahout offer diverse implementations that impact big data accuracy and quality.	Future work should focus on optimizing algorithms for heterogeneous data sources and improving scalability.
Rathore et al., (2021)	The use of large data and AI-ML to implement Digital Twins (DTs) in industrial applications.	Electronic bibliographic databases and patents from several disciplines are systematically reviewed, Discovering the instruments for progress and presenting an AI-enhanced reference architecture that is powered by big data.	Featured cutting-edge implementations already in progress; Offered a blueprint for systems that can be enhanced using DT	Need for optimized DT frameworks; Integration challenges with real-time IoT systems; Scalability and security concerns in DT implementations.

Mittal & Sangwan (2019) The several issues that might arise while utilizing classical MLT for Big Data Analytics have been covered in this study, along with potential fixes. The following are some of the possible answers to the questions raised by our survey on the state of big data analytics GPUs, map reduce tasks, deep learning, online learning, incremental learning, dimensionality reduction approaches, and parallel processing[34].

Alfred, (2016). In this study, the growing importance of ML for large information coherent is discussed. The first part of the article explains what machine learning is and how it relates to other terms like artificial intelligence, data science, statistics, data analytics, data mining, business intelligence, and knowledge discovery[35].

Khoshaba et al. (2022) The usage of algorithms and their substantial influence on BigData are covered and shown in this article, along with a comparison of Apache Spark and Apache Mahout's implementations. With the aid of AI and ML, a vast array of diverse, heterogeneous, and even divergent data sources have been incorporated into computer science research concepts, yielding exceptional accuracy and data quality outcomes[36].

Rathore et al. (2021) The article highlights the most cutting-edge implementations of DTs or systems based on DTs for various industrial uses, highlighting the importance of AI and ML and large data in creating solutions. Our systematic review was based on multidisciplinary electronic bibliographic databases and also took into account recent patents in the field. Tools for development that can facilitate various phases of digital twinning were also discovered. We also went to the trouble of creating a reference architecture for developers to use when developing DT-enabled systems that are powered by AI and driven by big data[37]

Table I summarizes machine learning's role in big data, highlighting challenges, methodologies, key findings, and future research directions.

V. CONCLUSION AND FUTURE WORK

Machine Learning has been a game-changer for big data analytics, providing potential answers in fields as diverse as healthcare, banking, and social science. The research reviewed highlights both the potential and challenges inherent in the use of ML on massive datasets. While ML's capacity to learn from past data and adapt to new information greatly improves data processing, it does have certain limits when it comes to scalability, real-time processing, and optimising algorithms. Various methodologies, methods that have recently evolved to tackle these issues include DL, dimensionality reduction, and parallel processing. Furthermore, the use of frameworks like Apache Spark and Mahout further demonstrates the diverse approaches being explored to improve the accuracy and efficacy of ML in big data contexts. Improving ML models' flexibility, especially in real-time situations, and honing strategies to manage the expanding complexity and amount of large information should be the primary goals of future research. As the field evolves, the collaboration between big data technologies and ML will continue to drive innovation and unlock new possibilities for data-driven decision-making.

VI. REFERENCES

[1] L. Wang and C. Alexander, “Machine Learning in Big Data,” Int. J. Math. Eng. Manag. Sci., vol. 1, pp. 52–61, 2016, doi: 10.33889/IJMEMS.2016.1.2-006.

[2] N. Das, S. Sagnika, and B. S. Prasad Mishra, “7 Role of Machine Learning in Big Data,” in Applications of Machine Learning in Big-Data Analytics and Cloud Computing, 2021, pp. 139–163.

[3] M. M. Srinivas Murri, Manoj Bhoyar, Guru Prasad Selvarajan, “Transforming Decision-Making with Big Data Analytics: Advanced Approaches to Real-Time Insights, Predictive Modeling, and Scalable Data Integration,” vol. 16, no. 5, 2024.

[4] A. H. Gandomi, F. Chen, and L. Abualigah, “Machine Learning Technologies for Big Data Analytics,” Electronics, vol. 11, no. 3, 2022, doi: 10.3390/electronics11030421.

[5] N. Rane, M. Paramesha, S. Choudhary, and J. Rane, “Machine Learning and Deep Learning for Big Data Analytics: A Review of Methods and Applications,” Jul. 2024, doi: 10.5281/zenodo.12271006.

[6] J. Qiu, Q. Wu, G. Ding, Y. Xu, and S. Feng, “A survey of machine learning for big data processing,” 2016. doi: 10.1186/s13634-016-0355-x.

[7] D. S. Vinayak Pillai, “Techniques for Processing and Analyzing Large Data Sets Using Big Data Analytics,” 2024.

[8] A. K. Aftab Arif, Muhammad Ismaeel Khan, “An overview of cyber threats generated by AI,” Int. J. Multidiscip. Sci. Arts, vol. 3, no. 4, pp. 67–76, 2024.

[9] D. Rao, “Systematic Analysis of threats, Machine Learning solutions and Challenges for Securing IoT environment,” J. Cybersecurity Inf. Manag., vol. 14, no. 2, pp. 367–382, 2024.

[10] L. Zhou, S. Pan, J. Wang, and A. V. Vasilakos, “Machine learning on big data: Opportunities and challenges,” Neurocomputing, vol. 237, pp. 350–361, 2017, doi: 10.1016/j.neucom.2017.01.026.

[11] M. Shah, P. Shah, and S. Patil, “Secure and Efficient Fraud Detection Using Federated Learning and Distributed Search Databases,” in 2025 IEEE 4th International Conference on AI in Cybersecurity (ICAIC), IEEE, Feb. 2025, pp. 1–6. doi: 10.1109/ICAIC63015.2025.10849280.

[12] S. Arora and S. R. Thota, “Using Artificial Intelligence with Big Data Analytics for Targeted Marketing Campaigns,” no. June, 2024, doi: 10.48175/IJARSCT-18967.

[13] S. Murri, “Data Security Environments Challenges and Solutions in Big Data,” Int. J. Curr. Eng. Technol., vol. 12, no. 6, pp. 565–574, 2022.

[14] S. R. Thota, S. Arora, and S. Gupta, “Hybrid Machine Learning Models for Predictive Maintenance in Cloud-Based Infrastructure for SaaS Applications,” in 2024 International Conference on Data Science and Network Security (ICDSNS), IEEE, Jul. 2024, pp. 1–6. doi: 10.1109/ICDSNS62112.2024.10691295.

[15] V. Pillai, "Enhancing the transparency of data and ml models using explainable AI (XAI)," World J. Adv. Eng. Technol. Sci., pp. 397–406, 2024.
[16] V. Prajapati, "Role of Identity and Access Management in Zero Trust Architecture for Cloud Security : Challenges and Solutions," pp. 6–18, 2025, doi: 10.48175/IJARSCT-23902.
[17] K. Rajchandar, M. Ramesh, A. Tyagi, S. Prabhu, D. S. Babu, and A. Roniboss, "Edge Computing in Network-based Systems: Enhancing Latency-Sensitive Applications," in 2024 7th International Conference on Contemporary Computing and Informatics (IC3I), IEEE, Sep. 2024, pp. 462–467. doi: 10.1109/IC3I61595.2024.10828607.
[18] Y. Kumar, J. Marchena, A. H. Awlla, J. J. Li, and H. B. Abdalla, "The AI-Powered Evolution of Big Data," Appl. Sci., vol. 14, no. 22, 2024, doi: 10.3390/app142210176.
[19] K. M. R. Seetharaman, "Internet of Things (IoT) Applications in SAP: A Survey of Trends, Challenges, and Opportunities," Int. J. Adv. Res. Sci. Commun. Technol., vol. 3, no. 2, 2021, doi: DOI: 10.48175/IJARSCT-6268B.
[20] A. Gogineni, "Multi-Cloud Deployment with Kubernetes: Challenges, Strategies, and Performance Optimization," Int. Sci. J. Eng. Manag., vol. 1, no. 2, 2022.
[21] S. P. M Shah, "AI/ML Techniques for Real-Time Fraud Detection," DZone, 2025.
[22] D. D. Rao1 et al., "Strategizing IoT Network Layer Security Through Advanced Intrusion Detection Systems and AI-Driven Threat Analysis," J. Intell. Syst. Internet Things, vol. 12, no. 2, 2024, doi: https://doi.org/10.54216/JISIoT.120215.
[23] S. Patel, "INTEGRATING MACHINE LEARNING TECHNIQUES FOR BIG DATA ANALYTICS," Int. J. Adv. Comput. Res., vol. 8, pp. 2760–2763, Jun. 2017, doi: 10.26483/ijarcs.v8i5.4165.
[24] Godavari Modalavalasa, "The Role of DevOps in Streamlining Software Delivery: Key Practices for Seamless CI/CD," Int. J. Adv. Res. Sci. Commun. Technol., vol. 1, no. 12, pp. 258–267, Jan. 2021, doi: 10.48175/IJARSCT-8978C.
[25] S. Chatterjee, "Integrating Identity and Access Management for Critical Infrastructure : Ensuring Compliance and Security in Utility Systems," vol. 8, no. 2, pp. 1–8, 2022.
[26] M. Gopalsamy and K. B. Dastageer, "The Role of Ethical Hacking and AI in Proactive Cyber Defense : Current Approaches and Future Perspectives," Int. J. Innov. Sci. Res. Technol., vol. 10, no. 2, 2025, doi: 10.5281/zenodo.14916984.
[27] M. S. Akaash Vishal Hazarika, "EXPLORING FAULT TOLERANCE STRATEGIES IN BIG DATA INFRASTRUCTURES AND THEIR IMPACT ON PROCESSING EFFICIENCY," Int. J. Comput. Sci. Inf. Technol., vol. 16, no. 6, pp. 35–40, 2024.
[28] H. A. Selmy, H. K. Mohamed, and W. Medhat, "Big data analytics deep learning techniques and applications: A survey," Inf. Syst., vol. 120, p. 102318, 2024, doi: https://doi.org/10.1016/j.is.2023.102318.
[29] M. S. Akaash Vishal Hazarika, "Blockchain-based Distributed AI Models: Trust in AI model sharing," Int. J. Sci. Res. Arch., vol. 13, no. 2, pp. 3493–3498, 2024.
[30] S. B. Shah, "Machine Learning for Cyber Threat Detection and Prevention in Critical Infrastructure," J. Glob. Res. Electron. Commun., vol. 2, no. 2, pp. 1–7, 2025, doi: https://jgrec.info/index.php/jgrec.
[31] T. Li, W. Deng, and J. Wu, "Advanced Machine Learning Applications in Big Data Analytics," Electronics, vol. 12, no. 13, 2023, doi: 10.3390/electronics12132940.
[32] S. R. Sagili, C. Goswami, V. C. Bharathi, S. Ananthi, K. Rani, and R. Sathya, "Identification of Diabetic Retinopathy by Transfer Learning Based Retinal Images," in 2024 9th International Conference on Communication and Electronics Systems (ICCES), IEEE, Dec. 2024, pp. 1149–1154. doi: 10.1109/ICCES63552.2024.10859381.
[33] S. S. S. Neeli, "Heart Disease Prediction For A Cloud-Based Smart Healthcare Monitoring System Using Gans And Ant Colony Optimization," Int. J. Med. Public Heal., vol. 14, no. 4, p. 11, 2024.
[34] S. Mittal and O. P. Sangwan, "Big Data Analytics using Machine Learning Techniques," in 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence), 2019, pp. 203–207. doi: 10.1109/CONFLUENCE.2019.8776614.
[35] R. Alfred, "The rise of machine learning for big data analytics," in 2016 2nd International Conference on Science in Information Technology (ICSITech), 2016, p. 1. doi: 10.1109/ICSITech.2016.7852593.
[36] F. Khoshaba, S. Kareem, H. Awla, and C. Mohammed, "Machine learning algorithms in Bigdata Analysis and its applications: A review," in 2022 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), 2022, pp. 1–8. doi: 10.1109/HORA55278.2022.9799848.
[37] M. M. Rathore, S. A. Shah, D. Shukla, E. Bentafat, and S. Bakiras, "The Role of AI, Machine Learning, and Big Data in Digital Twinning: A Systematic Literature Review, Challenges, and Opportunities," IEEE Access, vol. 9, pp. 32030–32052, 2021, doi: 10.1109/ACCESS.2021.3060863..