IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 12 July 2024, accepted 7 August 2024, date of publication 16 August 2024, date of current version 28 August 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3445281

RESEARCH ARTICLE

Cross-Sector Application of Machine Learning in Telecommunications: Enhancing Customer Retention Through Comparative Analysis of Ensemble Methods

MOHAMMAD AFZAL[1], (Member, IEEE), SHANUR RAHMAN[2], (Member, IEEE), DISHA SINGH[3], AND ALE IMRAN[4], (Member, IEEE)

[1]Department of Mechanical Engineering, Aligarh Muslim University, Aligarh, Uttar Pradesh 202002, India
[2]Department of Computer Engineering, Aligarh Muslim University, Aligarh, Uttar Pradesh 202002, India
[3]Department of Civil Engineering, Aligarh Muslim University, Aligarh, Uttar Pradesh 202002, India
[4]Zakir Husain College of Engineering and Technology, Aligarh Muslim University, Aligarh, Uttar Pradesh 202002, India

Corresponding author: Shanur Rahman (shanurrahman02@gmail.com)

ABSTRACT This work investigates the Supply chain evolution over diverse industries and forecasts a telecom churn using a publicly available dataset, besides offering a thorough technique for assessing and projecting customer attrition. The study begins with a thorough literature search that examines the growth of AI in supply chain management (SCM) across several industries, including healthcare and e-commerce, in order to give a comprehensive background and contextualize the telecom sector's accomplishments with reference to AI. The methodology includes applying different machine learning models to predict customer turnover by meticulously pre-processing the data and conducting exploratory data analysis (EDA). In order to handle missing values, a hybrid technique using K-Nearest Neighbors (KNN) Imputer for numerical features and Simple Imputer for categorical variables was used during the data preprocessing stage, which involves deleting duplicate entries and unnecessary columns. Through informative statistics and graphics, Exploratory Data Analysis (EDA) revealed important elements like rival offerings and the kind of internet service, thereby delivering insightful information regarding what causes churn. During the investigation, various machine learning models, such as Decision Tree, Random Forest, K-Neighbors, and XGBoost classifiers, were used. With an accuracy of 98.25%, Random Forest shows improved performance over the Decision Tree model, which had an accuracy of 98.02%. These models were tested using accuracy, precision, recall, F1-score, and AUC-ROC. This approach emphasizes how important predictive analytics is to comprehending the dynamics of customer turnover and further lays the groundwork for tactical interventions meant to improve customer satisfaction and retention in the telecom industry.

INDEX TERMS Artificial intelligence, machine learning, supply chain management, electronics commerce, healthcare, telecommunication.

I. INTRODUCTION

In today's competitive corporate landscape, understanding and anticipating customer behavior, particularly churn, is critical for long-term growth and profitability, particularly in industries such as telecoms. The rate at which customers break off their relationship with a firm is known as customer churn, and it presents both possibilities and substantial obstacles for organizations looking to improve their customer retention strategy. In addition to reducing income loss, efficient customer churn management promotes enduring client happiness and loyalty.

In order to predict and reduce attrition, the telecoms industry which is marked by intense competition and quick

The associate editor coordinating the review of this manuscript and approving it for publication was Tariq Umer.

technological advancements heavily depends on data-driven insights. The way organizations analyze massive volumes of data to effectively estimate customer behavior has been changed by advances in machine learning and artificial intelligence (AI). By using past data, telecom businesses can now forecast when and why customers are likely to churn, enabling proactive client retention initiatives [1], [2].

This study uses a telecom dataset that is available to the public to investigate a structured methodology for assessing and forecasting customer attrition. In order to guarantee data integrity and relevance, the study starts with thorough data pre-treatment [3], [4]. Next, intelligent exploratory data analysis (EDA) is used to find patterns and trends driving churn decisions. The research then uses a variety of machine learning models, such as Random Forest, K-Neighbors, Decision Trees, and XGBoost classifiers, to predict customer attrition with a high degree of accuracy [5].

The practical ramifications of this research for telecommunication firms looking to optimize client retention tactics through predictive analytics make it noteworthy. This study helps the telecom sector adopt a customer-centric mindset and improve operational efficiencies by identifying important churn reasons and assessing the effectiveness of several machine learning algorithms [6].

The goal of this investigation is to give telecom stakeholders practical advice on how to use AI and machine learning to prevent customer attrition and foster sustainable growth and competitive advantage in a changing market.

II. LITERATURE REVIEW

Under the literature review heading, we analyze the algorithms utilized in supply chain management within the e-Commerce, healthcare, and telecommunications sectors over the past two decades. We also examine how the telecommunications sector intersects with e-Commerce and healthcare. This review aims to uncover trends, advancements, and the impacts of algorithmic applications within these sectors, emphasizing their roles in enhancing operational efficiencies, supporting decision-making processes, and addressing sector-specific challenges.

A. AI IN E-COMMERCE

This study focuses on leveraging machine learning (ML) for supply chain optimization [7] across various industries, including e-Commerce. The e-commerce sector witnessed a significant focus on leveraging machine learning (ML) for supply chain optimization. Pioneering studies highlighted ML's superiority in demand prediction, utilizing support vector machines and neural networks [8]. Beyond demand forecasting, the era showcased the multifaceted applications of ML in addressing various e-Commerce challenges. Studies explored fraud detection, ensuring secure transactions, and mitigating risks [9] while intelligent Quality of Service brokering aimed at enhancing the overall customer experience [10]. Additionally, agent-based inventory planning and dynamic pricing strategies were investigated, showcasing the versatility of ML in optimizing inventory levels and refining pricing strategies in the rapidly evolving landscape of e-Commerce [11], [12]. This period set the stage for a broader integration of ML techniques into the e-commerce supply chain, emphasizing its potential to tackle diverse operational challenges. In the field of e-Commerce logistics optimization, a pivotal study [13] led by researchers Liu and Zhang introduces an ant colony optimization model tailored for addressing last-mile distribution challenges in rural areas. This government-subsidized research not only aims to maximize the profitability of logistics enterprises operating in rural e-Commerce logistics (RECL) but also emphasizes the pressing need for innovative solutions. The work of Liu and Zhang demonstrates the superior efficiency and effectiveness of the proposed ant colony optimization approach through a comprehensive comparison with alternative methods. In conjunction, other substantial contributions in e-Commerce research are evident in a study by Niu et al. [14], which delves into predictive analytics of online shoppers' behavior on Walmart.com, utilizing modern machine-learning techniques to forecast customer purchase conversion. Furthermore, the work of Khrais et al. [15] explores the impact of Artificial Intelligence on shaping consumer demand in e-Commerce, emphasizing product personalization and customization while addressing ethical concerns through Explainable Artificial Intelligence (XAI). Additionally, a research paper led by Baryannis et al. [16] provides a thorough assessment of AI applications in supply chain risk management, encompassing various techniques such as Petri nets, modular reasoning systems, and case-based reinforcement learning. Lastly, the study led by Bulsara et al. underscores potential blockchain applications in e-commerce, emphasizing the impact on payment processing, security, supply chain management, smart contracts, and ethical practices, with the involvement of major companies in implementation [17]. Machine learning's role in e-commerce supply chain management has rapidly evolved. Studies focus on recommender systems, emphasizing personalized recommendations. Algorithms like FP-Growth and Apriori are employed for transactional data analysis [18]. Table 1 offers a condensed representation of the literature review analysis. Another study introduces DBSL and DHEL for parcel loss prediction, leveraging Explainable AI. These advances underscore the continuous refinement of machine learning models in optimizing e-commerce supply chains, utilizing collaborative filtering, reinforcement learning, deep learning, and explainable AI to address specific challenges. Results from these studies support ongoing sophistication in machine learning applications within the e-Commerce sector [19].

B. AI IN HEALTHCARE

In the healthcare industry, the literature review spanning highlighted key developments with Nilay Shah's seminal work scrutinizing the pharmaceutical supply chain

TABLE 1. E-commerce literature review.

S.No.	Title	Publisher	Published Year	Summary
1	Parcel loss prediction in last-mile delivery: deep and non-deep approaches with insights from Explainable AI	arXiv	2023	The paper proposes novel machine learning approaches, DBSL and DHEL, for parcel loss prediction in last-mile delivery, achieving effective results.[19]
2	E-commerce platform based on Machine Learning Recommendation System	IEEE	2021	The paper emphasizes e-commerce recommender systems, using FP-Growth and Apriori algo for association rule mining and personalized product recommendations.[18]
3	Blockchain technology for e-commerce industry	Science and Engineering Research Support Society	2020	The paper explores blockchain's impact on e-commerce, citing benefits like lower costs, enhanced security, and transparent supply chain management.[17]
4	Route Optimization for Last-Mile Distribution of Rural E-Commerce Logistics Based on Ant Colony Optimization	IEEE	2020	The paper introduces an ant colony optimization model for efficient last-mile distribution in rural e-commerce logistics, outperforming other algorithms.[13]
5	Role of Artificial Intelligence in Shaping Consumer Demand in E-Commerce	MDPI	2020	The study examines AI's role in shaping consumer demand in e-commerce, addressing personalization, ethical concerns, and employing data analysis strategies.[15]
6	Decision Support Systems and Artificial Intelligence in Supply Chain Risk Management	Springer	2019	The paper assesses AI applications in supply chain risk management, discussing techniques like Petri nets, reinforcement learning, and various algorithms for SCRM tasks.[16]
7	Predictive analytics of E-commerce search behavior for conversion	Springer	2017	The paper employs logistic regression and random forest algorithms to predict e-commerce conversion based on Walmart.com search behavior, emphasizing comprehensive analysis and insights into consumer decisions.[14]
8	A Dynamic Pricing Approach in E-Commerce Based on Multiple Purchase Attributes	Springer	2010	The paper introduces a dynamic pricing approach in e-commerce, utilizing a feed-forward neural network to optimize seller profit based on multiple purchase attributes.[12]
9	An intelligent Quality of Service brokering model for e-commerce	Elsevier	2010	The paper introduces a fuzzy logic-based Quality of Service brokering model for e-commerce, addressing resource optimization and enhancing user-perceived QoS.[10]
10	Application of machine learning techniques for supply chain demand forecasting	Elsevier	2008	The paper explores machine learning techniques for demand forecasting in e-commerce supply chains, emphasizing improved accuracy and inventory management capabilities.[8]
11	Survey of fraud detection techniques	IEEE	2004	The article surveys contemporary fraud detection techniques, analyzing approaches to mitigate financial damage from various fraud types.[9]
12	Agents with Genders for Inventory Planning in E-Management	Springer	2001	The paper presents an agent-based model with genetic algorithms for inventory planning, introducing agent gender to enhance cooperation.[11]

landscape [20]. In healthcare supply chain management, Ford and Scanlon's 2007 paper delved into the application of supply chain management principles, emphasizing challenges like cost containment and information asymmetry while advocating for more integrated strategies considering contextual local market conditions and stakeholder interests [21]. S. Pala-niappan and R. Awang present an Intelligent Heart Disease Prediction System (IHDPS) that uses Decision Trees, Naïve Bayes, and Neural Network data mining approaches, revealing Naïve Bayes as the most accurate (86.53%) in predicting heart disease, followed by Neural Network (85.68%) and Decision Trees (80.4%). The review evaluates algorithms, attributes, and proposes future advancements [22]. The healthcare industry saw significant developments in the adoption of emerging technologies. The healthcare industry witnesses enriched insights through impactful research papers. One study [23] navigates the RFID technology adoption, amalgamating UTAUT with demographic variables. Notably, neural network analysis highlights its superior predictive prowess. Another paper [24] explores blockchain in pharmaceuticals, emphasizing traceability and B2B smart contracts. The Drug Supply Chain

TABLE 2. Healthcare literature review.

S.No.	Title	Publisher	Published Year	Summary
1	Digging DEEP: Futuristic building blocks of omni-channel healthcare supply chains resiliency using machine learning approach	Elsevier	2023	The paper explores omni-channel healthcare supply chain resiliency factors, utilizing K-means clustering for insights, proposing research directions, and acknowledging analytical limitations.[28]
2	Artificial Intelligence Digital Enablers in Facilitating Demand Forecasting of Biopharmaceutical Supply Chains	IJRAR	2022	The paper explores AI tools in biopharmaceutical demand forecasting, referencing external sources for metrics, highlighting data quality's pivotal role in accuracy.[27]
3	Ethical, legal, and social considerations of AI-based medical decision-support tools: A scoping review	Elsevier	2022	The paper scrutinizes ELSI aspects of AI in medical decision support, emphasizing proactive Ethics by Design and addressing transformative potential and challenges.[26]
4	Enhancing Efficiency in Healthcare Supply Chains: Leveraging Machine Learning for Optimized Operations	IJFMR	2021	The paper optimizes healthcare supply chain with ML classifiers, revealing Random Forest's superiority in multiple categories, emphasizing data refinement and validation.[4]
5	IoT-Enabled Healthcare: Benefits, Issues and Challenges	ICFNDS	2020	The paper explores IoT's transformative potential in healthcare, detailing benefits, architectural considerations, and addressing security challenges, offering a comprehensive survey.[3]
6	A Blockchain and Machine Learning-Based Drug Supply Chain Management and Recommendation System for Smart Pharmaceutical Industry	MDPI	2020	The DSCMR system integrates blockchain and machine learning for drug traceability and personalized recommendations, showcasing efficacy in a pharmaceutical case study.[25]
7	Towards a Blockchain Based Traceability Process: A Case Study from Pharma Industry	Springer	2019	The paper explores a pharmaceutical company's adoption of a blockchain-based serialization system to enhance traceability and comply with EU regulations.[24]
8	Predicting RFID adoption in healthcare supply chain from the perspectives of users	Elsevier	2015	The paper innovatively predicts RFID adoption in healthcare by merging UTAUT with demographics, using neural networks for enhanced predictive accuracy.[23]
9	Mayo clinical Text Analysis and Knowledge Extraction System (cTAKES): architecture, component evaluation and applications	American Medical Informatics Association	2010	The paper explores cTAKES, an open-source clinical text analysis tool, detailing its architecture, functionalities, and performance metrics, highlighting its competitive edge.[35]
10	Intelligent heart disease prediction system using data mining techniques	IEEE	2008	The paper presents an Intelligent Heart Disease Prediction System using data mining, with Na´ve Bayes leading in accuracy at 86.53%.[22]
11	Promise and problems with supply chain management approaches to health care purchasing	Health Care Management Review	2007	The paper explores challenges and limitations in applying supply chain management to healthcare, advocating for integrated, context-specific strategies and comprehensive research.[21]

Management and Recommendation system (DSCMR) [25] merges blockchain and machine learning for personalized medication recommendations, showcasing efficacy through comparisons. Additionally, IoT's transformable potential in healthcare [3] is explored, focusing on patient experiences, drug delivery, and cost-efficiency. These papers collectively contribute to understanding technology adoption, traceability solutions, and the evolving landscape of healthcare technologies. Research by Roy et al. [4] optimizes healthcare supply chains using machine learning, achieving 87% accuracy in areas such as inspection outcomes and transportation modes. Another study [26] delves into ethical, legal, and social considerations of AI-based medical decision-support tools, emphasizing patient safety and algorithmic transparency. Shashi and Manish explores artificial intelligence digital enablers for demand forecasting in biopharmaceutical supply chains, acknowledging complexity and referencing external metrics [27]. Kumar et al. focus on omni-channel healthcare

TABLE 3. Telecommunication literature review.

S.No.	Title	Publisher	Published Year	Summary
1	Federated machine learning for privacy preserving, collective supply chain risk prediction	Taylor & Francis	2023	The paper proposes federated learning to facilitate collective risk prediction in supply chains without data leakage, benefiting organizations with inadequate datasets.[1]
2	Blockchained supply chain management based on IoT tracking and machine learning	Springer	2022	6G IoT is crucial for modern supply chains. Blockchain ensures collaboration, while a multi-head attention-based GRU model improves inbound logistics prediction accuracy.[39]
3	Machine learning: Best way to sustain the supply chain in the era of industry 4.0	MaterialsToday	2021	The rise of Industry 4.0 emphasizes digital integration in manufacturing, highlighting the importance of IT-enabled systems and machine learning for future industry advancements.[38]
4	Simulation-based optimization of a stochastic supply chain considering supplier disruption: Agent-based modeling and reinforcement learning	Scientia Iranica	2019	The paper explores uncertainties in supply chain management with a multi-period stochastic model, employing agent-based simulation and Q-learning optimization, showing superior efficiency.[36]
5	BadNets: Identifying Vulnerabilities in the Machine Learning Model Supply Chain	IEEE	2017	The paper highlights security risks in outsourced deep learning training, presenting backdoored neural networks capable of behaving maliciously on specific inputs.[31]
6	Customer churn prediction in telecommunications	Expert Systems with Applications	2012	A study introduces enhanced features for predicting land-line customer churn, employing seven techniques. Results show improved effectiveness, with C4.5 and SVM notably successful.[37]
7	Artificial intelligence applications in the telecommunications industry	Expert Systems with Applications	2007	The study investigates AI applications in telecommunications, with network management as a key focus. Expert systems and machine learning are widely used, with promise for distributed AI.[6]
8	A reinforcement learning approach for dynamic supplier selection	IEEE	2007	The paper addresses supplier selection in manufacturing, employing an auction framework where suppliers compete based on previous actions, utilizing reinforcement learning and fictitious play.[30]
9	Applying data mining to telecom churn management	Elsevier	2006	The paper investigates churn prediction strategies for mobile operators in Taiwan following deregulation, identifying decision tree and neural network methods that are effective utilising varied subscriber data.[29]

supply chain resiliency, utilizing K-means clustering to propose Industry 5.0 integration strategies [28]. Although accuracy isn't quantified, these studies contribute valuable insights specifically to the healthcare sector from 2020 to the current year. Consult Table 2 for a comprehensive outline of the literature review.

C. AI IN TELECOMMUNICATION

Taiwan's 1997 deregulation of wireless telecom services increased competition among mobile operators, making churn control critical. This study done by Hung et al. [29] analyzes the effectiveness of decision trees and neural networks in predicting subscriber churn using call detail records, service logs, and consumer demographics. Table 3 offers a condensed representation of the literature review analysis. The integration of AI technology in telecommunications over the last decade has led to significant developments, especially in network management. Expert systems and machine learning are common, with distributed AI and machine learning emerging as new areas of research. Current applications of AI in Configuration Management and Maintenance (CMM) tackle industry-specific challenges, while its use in service product management (SPM) requires further research [6]. The reasearch done by Kim et al. [30] investigate supplier selection in manufacturing using recurrent games, incorporating Fictitious Play (FP) and Reinforcement Learning (RL) for adaptive decision-making. This methodology enhances computer efficiency in recurrent order patterns and contributes to understanding procurement decision-making processes. A study done by Delen [5] on customer churn prediction in telecom evaluates seven prediction methods and introduces new features such as call information and demographic profiles. By testing logistic regression, decision trees, and other algorithms on a dataset of 827,124 Irish telecom users, the study emphasizes customizing modeling methodologies and addressing challenges related to high-dimensional data in Naive Bayes modeling. In recent years, there has been a greater focus on managing supply chain uncertainty. A study done by Zhong et al. [2] simulates a multi-period stochastic supply chain with supplier disruptions and demand unpredictability, using agent-based modeling and Q-learning for optimization. The study underscores the importance of innovative decision-making in complex supply chain scenarios. Another study done by Gu and others, [31] examines the security implications of deep learning techniques, focusing on ''BadNets,'' which exhibit deceptive behavior with specific inputs. Examples like a backdoored handwritten digit classifier and a manipulated street sign classifier illustrate the persistence of these backdoors, emphasizing the need for further research into detection and inspection approaches. Liu et al. [32] Reviewed the impact of modern economy, technology on industries, emphasizing the importance of IT-enabled systems in defining Industry 4.0. Digital transformation interconnects operations and detailed digital mapping, highlighting machine learning's critical role in modern industrial processes and providing insights for navigating industrial changes. Similarly, another review done by Tufano et al. [33] explores the influence of IT-enabled systems on Industry 4.0, discussing current opinions, issues, and contributions to future industrial landscapes. It underscores the symbiotic relationship between technology and industry, highlighting machine learning's role in modernizing industrial processes. Some other study like Supply chain risk prediction using AI [34] done by Ali and others, often assumes autonomous functioning within interconnected supply chains, which presents challenges for organizations with limited datasets. Federated learning offers a solution for collective risk prediction while maintaining data privacy. Another research work like Supply Chain risk prediction using AI demonstrates its effectiveness, though factors like unbalanced training data and algorithm selection significantly impact performance. This study enhances the understanding of individual and group learning paradigms in supply chain risk prediction, guiding future research in this emerging field.

D. RELATION OF E-COMMERCE AND HEALTHCARE SECTOR WITH TELECOMMUNICATION SECTOR

1) E-COMMERCE WITH TELECOMMUNICATION

The e-Commerce sector is profoundly impacted by the Telecom industry, which provides the essential infrastructure and technologies for online retail to flourish. High-speed internet connectivity is fundamental for conducting online transactions, supporting customer interactions, and executing digital marketing strategies. The widespread adoption of mobile internet and the proliferation of smartphones, driven by telecom advancements, have transformed e-Commerce through mobile commerce (m-commerce), enabling consumers to shop from anywhere at any time, significantly expanding market reach. Telecom technologies enhance logistics and supply chain management by enabling real-time tracking, efficient inventory management, and improved communication among stakeholders, ensuring timely deliveries and better customer satisfaction. In addition, advanced telecom networks support innovative customer service solutions, such as live chat support, video consultations, and personalized marketing, creating seamless and engaging shopping experiences. Moreover, telecom services enable cloud-based platforms that e-Commerce companies rely on for scalable and flexible operations. In summary, the telecom industry's contributions are indispensable to the operational success, scalability, and customer-centric advancements of the e-Commerce sector, driving its growth and innovation.

2) HEALTHCARE WITH TELECOMMUNICATION

The telecom industry is revolutionizing the healthcare sector by enabling a range of advanced services that significantly enhance patient care and operational efficiency. Tele-medicine, supported by high-speed internet and robust telecom networks, facilitates remote consultations, diagnostics, and treatment, which is especially beneficial for individuals in rural or under served regions with limited access to healthcare facilities. The integration of IoT (Internet of Things) in medical devices allows continuous monitoring of patients' health metrics, such as heart rate, blood pressure, and glucose levels, transmitting this data in real-time to healthcare providers for prompt intervention. Secure and efficient data transfer enabled by telecom networks is crucial for maintaining comprehensive electronic health records (EHRs), supporting accurate diagnostics, personalized treatment plans, and facilitating medical research. Moreover, the coordination of emergency services heavily relies on effective communication systems, ensuring rapid and precise mobilization during critical situations. Tele-health platforms and mobile health applications leverage telecom advancements, allowing patients to access health information, schedule appointments, and receive follow-up care conveniently. Overall, the telecom industry's infrastructure and innovations play a pivotal role in improving the accessibility, quality, and efficiency of healthcare delivery.

III. RESEARCH GAP

Despite increasing research on AI and machine learning applications in supply chain management, including e-Commerce, healthcare, and telecom, significant gaps remain. The literature frequently overlooks the possible advantages of cross-sector integration in favor of applications that are sector-specific. Furthermore, real-time data processing and the scalability of these models in live contexts are not given as much attention as they should, even though historical data is frequently employed for model training and analysis. It is not well known how blockchain technology and artificial intelligence may work together to improve data security and transparency. Moreover, the social, legal, and ethical ramifications of using AI are not fully explored. Additionally, the literature frequently uses classic machine learning models and pays little attention to more sophisticated methods like reinforcement learning and deep learning, which may provide better results and insights. Closing these gaps may result in supply chain AI systems that are more complete and efficient.

IV. METHODOLOGY

This part conveys the methodology of the analysis and the prediction and forecasting of the Telecom Customer Churn Prediction dataset that is readily available publicly on kaggle platform. This dataset was taken because it perfectly aligns with our task and it has potential to help, with which it can give insightful visualization results and portray the significance on how forecasting has significant effects that can evolve Supply chain management.

A. DATA PRE-PROCESSING AND ANALYSIS

Data pre-processing is a pivotal stage, ensuring the dataset's integrity and preparing it for insightful analyses. Initially, the telecom customer churn dataset underwent a meticulous review, leading to the removal of extraneous columns like 'Zip Code,' 'Latitude,' 'Longitude,' and 'City' shown in Table 4. Simultaneously, duplicated entries were identified and eliminated to maintain data fidelity.

Addressing missing values is crucial, and a hybrid approach was adopted. The KNN Imputer was employed for numerical features like 'Avg Monthly GB Download' and 'Avg Monthly Long Distance Charges,' utilizing the k-nearest neighbors algorithm for accurate imputation. Categorical variables such as 'Churn Reason' and 'Churn Category' underwent imputation using the Simple Imputer with a 'most frequent' strategy. The resulting copy dataset is now devoid of missing values and irrelevant columns, establishing a robust foundation for analysis. Figure 1 shows the correlation of the data features with each other that are present in Table 4.

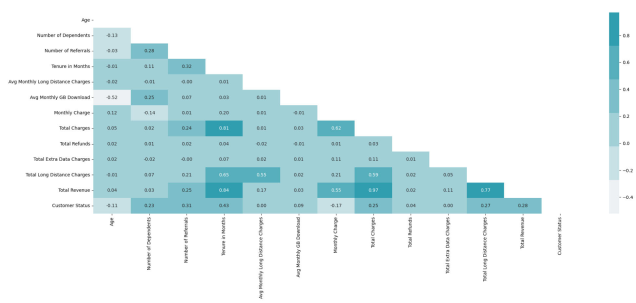

FIGURE 1. Custom diverging colormap of the dataset.

Exploratory Data Analysis (EDA) is the lens through which patterns and insights within the pre-processed dataset are unveiled. Descriptive statistics provide a quantitative overview, offering insights into central tendencies and feature dispersions. Visualizations, such as the histogram in Figure 2 of 'Churn Reason,' expose that the primary driver of customer

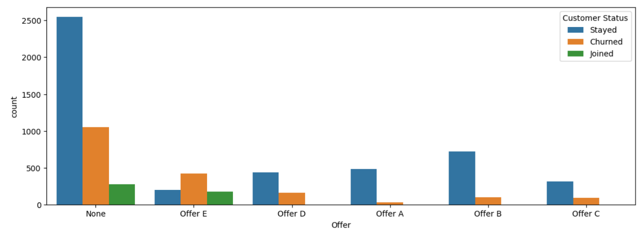

FIGURE 2. Churn prediction and offer graphs.

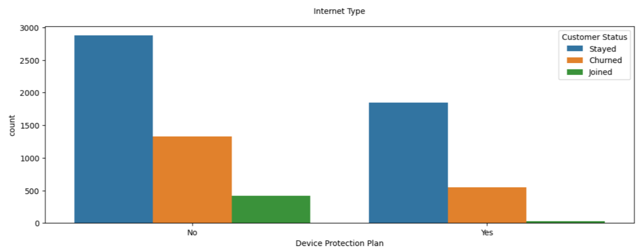

FIGURE 3. Count vs Device protection plan graphs.

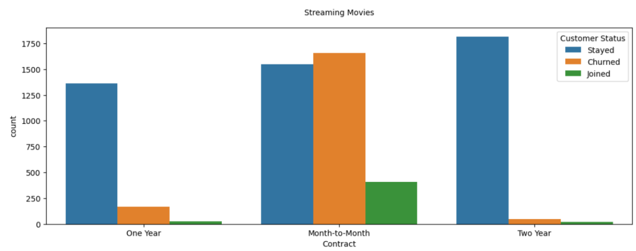

FIGURE 4. High churn rate because of month-to-month contract.

churn is competitors offering superior devices. Categorical variables like 'Type of Offer' and 'Internet Service' are scrutinized through count plots, illuminating the impact of specific factors on customer status. Concurrently, numerical features undergo histogram analysis, revealing patterns associated with customer churn. These findings, coupled with the earlier pre-processing steps, empower stakeholders to make informed decisions.

The exploration revealed noteworthy findings. The histogram of 'Churn Reason' underscores that competitor offerings in Figure 3, particularly superior devices, are a substantial driver of customer churn. Categorical variable analyses highlight the influence of factors like 'Type of Offer' and 'Internet Service', providing actionable insights for targeted strategies. Furthermore, numerical feature histograms shed light on patterns associated with customer churn, with monthly contract subscriptions emerging as a significant contributing factor as shown in Figure 4.

The combination of meticulous pre-processing and insightful EDA not only ensures data quality but also provides a nuanced understanding of customer churn dynamics in the telecommunications sector. Some other data visualization that helps us determine the cause of churn prediction are Month-to-month contracts as shown in Figure 4 and the

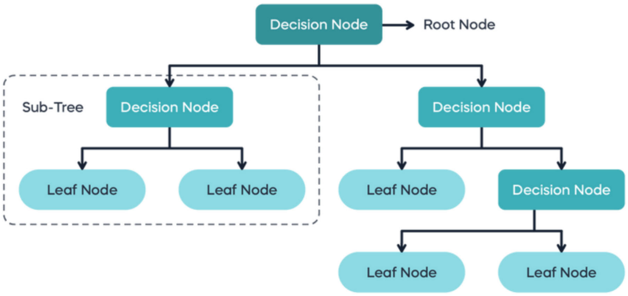

FIGURE 5. Architecture of decision tree classifier.

absence of premium tech assistance are significant factors that impact the prediction of churn. Because they offer less commitment than long-term contracts, month-to-month agreements frequently result in greater churn rates. Furthermore, clients who don't receive first-rate tech help could feel undervalued and end up canceling their agreements. These factors have a big impact on customer retention since they show how crucial supporting and adaptable service offerings are to satisfying customers' needs and lowering attrition rates.

These findings pave the way for strategic interventions aimed at customer retention and satisfaction, setting the stage for informed decision-making in the industry.

B. MODEL SELECTION AND TRAINING

1) DECISION TREE CLASSIFIER

The decision tree is a popular supervised learning algorithm that may be used for regression and classification applications. The algorithm as seen in Figure 5, is a potent machine learning algorithm that creates a hierarchical tree structure by recursively dividing the input space according to characteristics, simulating decision-making processes.

The algorithm optimizes for maximum information gain or Gini impurity reduction at each node by choosing the most informative feature to split the data into. By applying ideas from statistics and information theory, the method seeks to reduce uncertainty and improve prediction accuracy. There are various kinds of uncertainties that happen in supply chain management and this is where decision trees help analyze and make decisions in uncertain situations [40]. Using Decision trees can bring substantial value to supply chain management by giving critical foresight in a complicated and dynamic industry such as delivery route optimization, supplier selection based on pricing and reliability, and inventory management based on sales patterns or seasonal capacity.

Mathematically, Decision tree classifier is written as:

$$\begin{aligned} \text{Entropy (S)} &= -\sum P(I) \times \log_2(P(I)) \\ \text{Information Gain (S, A)} &= \text{Entropy (S)} \\ &\quad -\sum P(S|A) \times \text{Entropy (S | A)} \end{aligned} \tag{1}$$

TABLE 4. Features of dataset.

	No Unique	Missing Value	Null Value	Duplicated	Dtype
Customer ID	7043	0	0	0	object
Gender	2	0	0	0	object
Age	62	0	0	0	int64
Married	2	0	0	0	object
Number of Dependents	10	0	0	0	int64
City	1106	0	0	0	object
Zip Code	1626	0	0	0	int64
Latitude	1626	0	0	0	float64
Longitude	1625	0	0	0	float64
Number of Referrals	12	0	0	0	int64
Tenure in Months	72	0	0	0	int64
Offer	6	0	0	0	object
Phone Service	2	0	0	0	object
Avg Monthly Long Distance Charges	3583	682	682	0	float64
Multiple Lines	2	682	682	0	object
Internet Service	2	0	0	0	object
Internet Type	3	1526	1526	0	object
Avg Monthly GB Download	49	1526	1526	0	float64
Online Security	2	1526	1526	0	object
Online Backup	2	1526	1526	0	object
Device Protection Plan	2	1526	1526	0	object
Premium Tech Support	2	1526	1526	0	object
Streaming TV	2	1526	1526	0	object
Streaming Movies	2	1526	1526	0	object
Streaming Music	2	1526	1526	0	object
Unlimited Data	2	1526	1526	0	object
Contract	3	0	0	0	object
Paperless Billing	2	0	0	0	object
Payment Method	3	0	0	0	object
Monthly Charge	1591	0	0	0	float64
Total Charges	6540	0	0	0	float64
Total Refunds	500	0	0	0	float64
Total Extra Data Charges	16	0	0	0	int64
Total Long Distance Charges	6068	0	0	0	float64
Total Revenue	6975	0	0	0	float64
Customer Status	3	0	0	0	object
Churn Category	5	5174	5174	0	object
Churn Reason	20	5174	5174	0	object

Here we have trained our decision tree model using the Scikit learn library which uses the Decision tree ID3 model in general. The accuracy is 98.02%. The Precision and recall of the classifier is 0.98 respectively.

2) RANDOM FOREST CLASSIFIER

Several decision trees are used in the Random Forest algorithm, an ensemble learning method, to improve prediction accuracy and robustness. Based on scientific and mathematical concepts, it applies randomization to feature selection and makes use of the bagging (bootstrap aggregating) notion. In Figure 6, we have a representation of Random forest classifier in which a random subset of attributes is considered at each split after sampling a portion of training

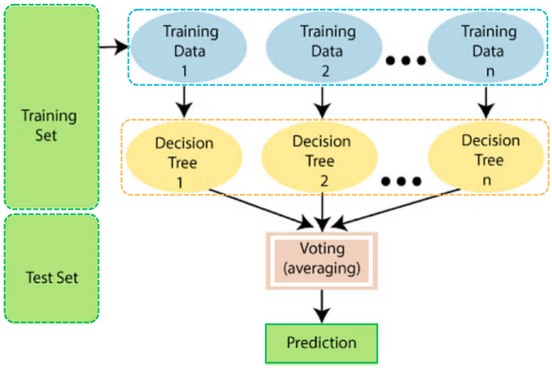

FIGURE 6. Architecture of random forest classifier.

data with a replacement for each tree in the forest. A majority vote or weighted average of the individual tree projections yields the final prediction. This ensemble method encourages variation among component trees, which reduces over-fitting and enhances generalization [41]. The algorithm utilizes statistical principles mathematically, using ideas such as voting and averaging to produce a more reliable and accurate model. Since it really works as an upgrade of Decision trees, and work great on issues like Churn prediction, Demand forecasting and Customer segmentation, this model showed the higher accuracy in our problem statement.

Mathematically it is written as:

$$\mathbf{R}Ff\,i_1 = \frac{\sum_{j \in \text{all trees}} \text{norm}(fi_{ij})}{T} \tag{2}$$

where, RFfi sub(i) = i, a featured calculated from all trees in the Random forest models normfi sub(ij) = Normalised feature significance for i in tree j T = total number of trees

This algorithm was applied to our processed dataset and it gave a much better result and the accuracy of our model is 98.25%. The Precision and recall of the classifier are 0.98 respectively.

3) K-NEIGHBORS CLASSIFIER

A non-parametric, instance-based learning strategy with a strong scientific and mathematical basis is the k-Nearest Neighbours (KNN) method. Based on the average value or majority class of their K-nearest neighbors in the feature space, it predicts and classify the data items. According to the local similarity principle, which holds that similar instances cluster together, the technique works. The architecture of KNN is seen below in Figure 7. KNN is a useful tool in machine learning because of its simplicity and flexibility to adapt to a variety of data distributions, yet large datasets may require careful consideration of their computational efficiency. Every decision made by KNN classifier is really dependent on the data variables and through this we can cluster-out the customers who are going to cancel the subscription [42]. After training the model, the accuracy is 91.21% which is comparatively low than other models. The Precision and recall of the classifier are 0.91 respectively.

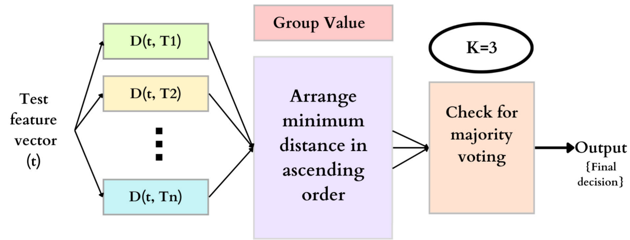

FIGURE 7. Architecture of k-neighbors classifier.

4) XGB CLASSIFIER

Extreme Gradient Boosting, or XGBoost, is a sophisticated ensemble learning method that is well-known for its quickness, effectiveness, and precision in predictions. With XGBoost, gradient descent optimization, and boosting principles are combined in a rigorously scientific and mathematical manner. Here is a great representation of the architecture and working of the XGB classifier in Figure 8. Through the sequential construction of a regularised objective function that consists of a loss term and a complexity term, it typically creates a collection of weak learners in the form of decision trees. The approach utilizes a second-order Taylor expansion to estimate the loss function, improving accuracy and gracefully managing over-fitting through regularisation. For efficiency, it also makes use of a distributed and parallelized computing environment. The secret to XGBoost's success is its versatility in handling different kinds of data, resolving imbalances, and producing comprehensible feature importance ratings [43]. Mathematically we can visualize XGBoost classifier as:

$$\text{Obj}(\theta) = \sum_{i=1}^{N} \ell(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \tag{3}$$

where, l: Loss function that calculates the discrepancy between the anticipated value and the actual target value

Omega(f k):The regularization term is applied to every tree in the ensemble to assist prevent overfitting.

It is effective at classification and regression problems and is frequently employed in real-world applications and machine-learning contests. This model was chosen because the data was imbalanced and XGBoost works really well and in optimized way. The accuracy of the XGB Model is 95.60%. The Precision and recall of the classifier are 0.91 respectively.

C. RESULTS

In the data analysis of the customer churn prediction, all the variables and their relations to one another are used as important feature for the prediction. The thorough Exploratory Data Analysis (EDA) provides important new information about the dynamics of customer turnover in the telecommunication industry. Important factors impacting churn are identified using descriptive statistics and visualizations like count plots and histograms. Most notably, the 'Churn Reason' histogram emphasizes competitors with

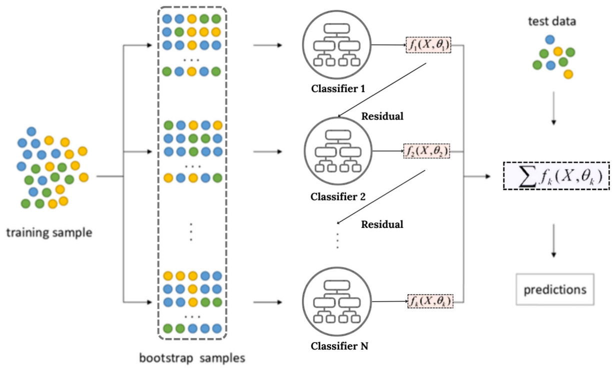

FIGURE 8. Architecture of XGB classifier.

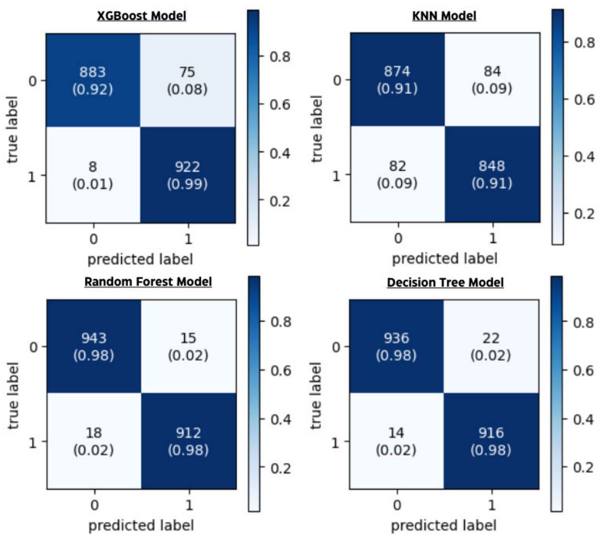

FIGURE 9. Confusion matrix of machine learning models.

better devices as the main motivator. The impact of category factors on customer retention can be revealed through analysis, such as 'Type of Offer' and 'Internet Service'. Moreover, month-to-month contracts and the absence of premium tech assistance are shown to be major causes of churn in numerical feature histograms. These results highlight how crucial accommodating and helpful service options are to raising client retention and satisfaction. When used in conjunction with careful pre-processing, the EDA guarantees strong data quality and provides useful insights. Through comprehension of these dynamics, interested parties can develop focused churn mitigation methods, which in turn can facilitate well-informed decision-making and strategic interventions meant to reinforce customer satisfaction and loyalty in the cut-throat telecommunication market.

Results recorded in Table 5 show that the Random Forest classifier has the accuracy of 98.25% in the customer churn prediction followed by Decision Tree and XG-Boost classifier with 98.09% and 95.60%.

Due to a number of significant benefits, the random forest classifier here frequently performs better than other

TABLE 5. Accuracy of the models.

Models	Training Accuracy	Testing Accuracy	F1-Score
Decision Tree	1.0000	0.9809	0.99
Random Forest	1.0000	0.9825	0.98
K-Neighbor	0.9553	0.9121	0.91
XGBoost	0.9567	0.9560	0.99

classifier models now in use. First of all, because it is non-parametric, it can handle both skewed and non-ordinal categorical data since it does not assume anything about the data's underlying distribution. Because the random forest uses an ensemble learning approach, which aggregates the predictions of numerous decision trees, it can also handle a large number of independent predictor variables without over-fitting. When compared to other classifiers, this strong methodology typically yields better predicted accuracy and generalization, particularly in complex datasets with high dimensionality and intricate feature interactions.

The confusion matrices for machine learning models are shown in Figure 9, which also shows how well these models predict customer turnover in the telecommunications sectors. These studies demonstrate how well ensemble learning methods capture the dynamics of client attrition. We pre-processed the data and fixed imbalances in the dataset before training multiple models to forecast client retention. The variables in the data, along with the choices and patterns found by the machine learning models, are the only factors that go into the outcomes and forecasts. In the customer churn dataset, both categorical and numerical variables show underlying patterns that are well-captured by XGBoost and Random Forests. The least amount of performance fluctuation indicates that both models are equally good at spotting the patterns linked to customer attrition. The Random Forest classifier showed the best accuracy of all the models tested. Adjusting the hyperparameters to attain the best accuracy is the reason for any performance discrepancies.

V. CONCLUSION & FUTURE SCOPE

In conclusion, this study emphasizes the importance of machine learning in understanding and reducing customer churn in the telecommunication business. Our detailed literature research, which looked at the evolution of AI in supply chain management (SCM) across industries such as healthcare and e-Commerce, helped to contextualize our findings. This review examined how advances in AI have altered other industries, providing useful insights into probable future improvements in telecom. The investigation began with meticulous data preparation, which included addressing missing values, encoding category variables, and normalizing numerical features. This was followed by exploratory data analysis (EDA) to pinpoint major churn causes such as competitor products, service categories, and customer support interactions.

Predictive models were built using a variety of classification methods such as Decision Tree, Random Forest, K-Nearest Neighbors (KNN), and Extreme Gradient Boosting (XGBoost). To measure their performance, these models were examined using accuracy, precision, recall, F1-score, and AUC-ROC. Random forest classifier performed the best, demonstrating its ability to capture complicated correlations in data.

The established models enable telecom operators to identify at-risk customers early on, allowing them to execute targeted retention efforts such as personalized offers and higher service quality. This proactive approach helps to reduce turnover and increase client loyalty.

The study also emphasizes the importance of continuous model improvement and the integration of real-time data in order to keep up with market developments. Future study could look at advanced approaches like deep learning and consumer feedback to gain a more complete picture of customer behavior. This study establishes a solid foundation for boosting client retention in the telecommunication industry.

REFERENCES

[1] G. Zheng, L. Kong, and A. Brintrup, "Federated machine learning for privacy preserving, collective supply chain risk prediction," *Int. J. Prod. Res.*, vol. 61, no. 23, pp. 8115–8132, Dec. 2023, doi: 10.1080/00207543.2022.2164628.

[2] R. Y. Zhong, G. Q. Huang, S. Lan, Q. Y. Dai, X. Chen, and T. Zhang, "A big data approach for logistics trajectory discovery from RFID-enabled production data," *Int. J. Prod. Econ.*, vol. 165, pp. 260–272, Jul. 2015, doi: 10.1016/j.ijpe.2015.02.014.

[3] B. Al-Shargabi and S. Abuarqoub, "IoT-enabled healthcare: Benefits, issues and challenges," in *Proc. 4th Int. Conf. Future Netw. Distrib. Syst. (ICFNDS)*, Nov. 2020, pp. 1–5.

[4] S. Roy and M. Mitra, "Enhancing efficiency in healthcare supply chains: Leveraging machine learning for optimized operations," *Int. J. Multidisciplinary Res.*, vol. 3, no. 2, pp. 1–8, Nov. 2021.

[5] D. Delen, "A comparative analysis of machine learning techniques for student retention management," *Decis. Support Syst.*, vol. 49, no. 4, pp. 498–506, Nov. 2010, doi: 10.1016/j.dss.2010.06.003.

[6] J. Qi, F. Wu, L. Li, and H. Shu, "Artificial intelligence applications in the telecommunications industry," *Expert Syst.*, vol. 24, no. 4, pp. 271–291, Sep. 2007, doi: 10.1111/j.1468-0394.2007.00433.x.

[7] J. V. Blackhurst, K. P. Scheibe, and D. J. Johnson, "Supplier risk assessment and monitoring for the automotive industry," *Int. J. Phys. Distrib. Logistics Manage.*, vol. 38, no. 2, pp. 143–165, Mar. 2008, doi: 10.1108/09600030810861215.

[8] R. Carbonneau, K. Laframboise, and R. Vahidov, "Application of machine learning techniques for supply chain demand forecasting," *Eur. J. Oper. Res.*, vol. 184, no. 3, pp. 1140–1154, Feb. 2008, doi: 10.1016/j.ejor.2006.12.004.

[9] Y. Kou, C.-T. Lu, S. Sirwongwattana, and Y.-P. Huang, "Survey of fraud detection techniques," in *Proc. IEEE Int. Conf. Netw., Sens. Control*, vol. 2, Mar. 2004, pp. 749–754, doi: 10.1109/ICNSC.2004.1297040.

[10] A. O. Ajayi, G. A. Aderounmu, H. A. Soriyan, and A. David, "An intelligent quality of service brokering model for e-commerce," *Expert Syst. Appl.*, vol. 37, no. 1, pp. 816–823, Jan. 2010, doi: 10.1016/j.eswa.2009.05.103.

[11] H. H. Pham and V.-H. Nguyen, "Agents with genders for inventory planning in E-management," in *Proc. Conf. Can. Soc. Comput. Stud. Intell.*, 2001, pp. 267–276, doi: 10.1007/3-540-45153-6_26.

[12] T. K. Ghose and T. T. Tran, "A dynamic pricing approach in e-commerce based on multiple purchase attributes," in *Artificial Intelligence (Canadian AI 2010)* (Lecture Notes in Computer Science), vol. 6085, A. Farzindar and V. Kešelj, Eds., Berlin, Germany: Springer, 2010, doi: 10.1007/978-3-642-13059-5_13.

[13] W. Liu, "Route optimization for last-mile distribution of rural e-commerce logistics based on ant colony optimization," *IEEE Access*, vol. 8, pp. 12179–12187, 2020, doi: 10.1109/ACCESS.2020.2964328.

[14] X. Niu, C. Li, and X. Yu. (2017). *Predictive Analytics of e-commerce Search Behavior for Conversion*. [Online]. Available: https://core.ac.uk/download/pdf/301371722.pdf

[15] L. T. Khrais, "Role of artificial intelligence in shaping consumer demand in e-commerce," *Future Internet*, vol. 12, no. 12, p. 226, Dec. 2020, doi: 10.3390/fi12120226.

[16] G. Baryannis, S. Dani, S. Validi, and G. Antoniou, "Decision support systems and artificial intelligence in supply chain risk management," in *Revisiting Supply Chain Risk* (Springer Series in Supply Chain Management), vol. 7. Cham, Switzerland: Springer, 2019, doi: 10.1007/978-3-030-03813-7_4.

[17] H. P. Bulsara and P. S. Vaghela, "Blockchain technology for e-commerce industry," *Int. J. Adv. Sci. Technol.*, vol. 29, no. 5, pp. 3793–3798, 2020.

[18] M. Tahir, R. N. Enam, and S. M. Nabeel Mustafa, "E-commerce platform based on machine learning recommendation system," in *Proc. 6th Int. Multi-Topic ICT Conf. (IMTIC)*, Nov. 2021, pp. 1–4, doi: 10.1109/IMTIC53841.2021.9719822.

[19] J. de Leeuw, Z. Bukhsh, and Y. Zhang, "Parcel loss prediction in last-mile delivery: Deep and non-deep approaches with insights from explainable AI," 2023, *arXiv:2310.16602*.

[20] N. Shah, "Pharmaceutical supply chains: Key issues and strategies for optimisation," *Comput. Chem. Eng.*, vol. 28, nos. 6–7, pp. 929–941, Jun. 2004, doi: 10.1016/j.compchemeng.2003.09.022.

[21] E. W. Ford and D. P. Scanlon, "Promise and problems with supply chain management approaches to health care purchasing," *Health Care Manage Rev.*, vol. 32, no. 3, pp. 192–202, 2007, doi: 10.5465/AMBPP.2006.27161719.

[22] S. Palaniappan and R. Awang, "Intelligent heart disease prediction system using data mining techniques," in *Proc. IEEE/ACS Int. Conf. Comput. Syst. Appl.*, Apr. 2008, pp. 108–115, doi: 10.1109/AICCSA.2008.4493524.

[23] A. Y.-L. Chong, M. J. Liu, J. Luo, and O. Keng-Boon, "Predicting RFID adoption in healthcare supply chain from the perspectives of users," *Int. J. Prod. Econ.*, vol. 159, pp. 66–75, Jan. 2015, doi: 10.1016/j.ijpe.2014.09.034.

[24] F. Chiacchio, D. D'Urso, L. Compagno, M. Chiarenza, and L. Velardita, "Towards a blockchain based traceability process: A case study from pharma industry," in *Advances in Production Management Systems. Production Management for the Factory of the Future* (IFIP Advances in Information and Communication Technology), vol. 566, F. Ameri, K. Stecke, G. von Cieminski, and D. Kiritsis, Eds., Cham, Switzerland: Springer, 2019, doi: 10.1007/978-3-030-30000-5_56.

[25] K. Abbas, M. Afaq, T. Ahmed Khan, and W.-C. Song, "A blockchain and machine learning-based drug supply chain management and recommendation system for smart pharmaceutical industry," *Electronics*, vol. 9, no. 5, p. 852, May 2020, doi: 10.3390/electronics9050852.

[26] A. Čartolovni, A. Tomičić, and E. L. Mosler, "Ethical, legal, and social considerations of AI-based medical decision-support tools: A scoping review," *Int. J. Med. Informat.*, vol. 161, May 2022, Art. no. 104738, doi: 10.1016/j.ijmedinf.2022.104738.

[27] S. Manish, "Artificial intelligence digital enablers in facilitating demand forecasting of biopharmaceutical supply chains," *Int. J. Res. Anal. Rev.*, vol. 9, no. 2, pp. 1–5, 2022.

[28] A. Kumar, F. Naz, S. Luthra, R. Vashistha, V. Kumar, J. A. Garza-Reyes, and D. Chhabra, "Digging DEEP: Futuristic building blocks of omni-channel healthcare supply chains resiliency using machine learning approach," *J. Bus. Res.*, vol. 162, Jul. 2023, Art. no. 113903, doi: 10.1016/j.jbusres.2023.113903.

[29] S.-Y. Hung, D. C. Yen, and H.-Y. Wang, "Applying data mining to telecom churn management," *Expert Syst. Appl.*, vol. 31, no. 3, pp. 515–524, Oct. 2006, doi: 10.1016/j.eswa.2005.09.080.

[30] T. Kim, R. U. Bilsel, and S. R. T. Kumara, "A reinforcement learning approach for dynamic supplier selection," in *Proc. IEEE Int. Conf. Service Operations Logistics, Informat.*, Aug. 2007, pp. 1–6, doi: 10.1109/SOLI.2007.4383959.

[31] T. Gu, B. Dolan-Gavitt, and S. Garg, "BadNets: Identifying vulnerabilities in the machine learning model supply chain," 2017, *arXiv:1708.06733*.

[32] C. Liu, Y. Feng, D. Lin, L. Wu, and M. Guo, "Iot based laundry services: An application of big data analytics, intelligent logistics management, and machine learning techniques," *Int. J. Prod. Res.*, vol. 58, no. 17, pp. 5113–5131, Sep. 2020, doi: 10.1080/00207543.2019.1677961.

[33] A. Tufano, R. Accorsi, and R. Manzini, "Machine learning methods to improve the operations of 3PL logistics," *Proc. Manuf.*, vol. 42, pp. 62–69, Jan. 2020, doi: 10.1016/j.promfg.2020.02.023.

[34] N. Ali, T. M. Ghazal, A. Ahmed, S. Abbas, M. A. Khan, H. M. Alzoubi, U. Farooq, M. Ahmad, and M. A. Khan, "Fusion-based supply chain collaboration using machine learning techniques," *Intell. Autom. Soft Comput.*, vol. 31, no. 3, pp. 1671–1687, 2022, doi: 10.32604/iasc.2022.019892.
[35] G. K. Savova, J. J. Masanz, P. V. Ogren, J. Zheng, S. Sohn, K. C. Kipper-Schuler, and C. G. Chute, "Mayo clinical text analysis and knowledge extraction system (cTAKES): Architecture, component evaluation and applications," *J. Amer. Med. Inform. Assoc.*, vol. 17, no. 5, pp. 507–513, Sep. 2010, doi: 10.1136/jamia.2009.001560.
[36] A. Aghaie and M. H. Heidary, "Simulation based optimization of a stochastic supply chain considering supplier disruption: An agent-based modeling and reinforcement learning," *Scientia Iranica*, vol. 26, no. 6, pp. 3780–3795, Jul. 2018, doi: 10.24200/sci.2018.20789.
[37] B. Huang, M. T. Kechadi, and B. Buckley, "Customer churn prediction in telecommunications," *Expert Syst. Appl.*, vol. 39, no. 1, pp. 1414–1425, Jan. 2012, doi: 10.1016/j.eswa.2011.08.024.
[38] D. Nagar, S. Raghav, A. Bhardwaj, R. Kumar, P. Lata Singh, and R. Sindhwani, "Machine learning: Best way to sustain the supply chain in the era of Industry 4.0," *Mater. Today, Proc.*, vol. 47, pp. 3676–3682, Jan. 2021, doi: 10.1016/j.matpr.2021.01.267.
[39] Z. Dong, W. Liang, Y. Liang, W. Gao, and Y. Lu, "Blockchained supply chain management based on IoT tracking and machine learning," *EURASIP J. Wireless Commun. Netw.*, vol. 2022, no. 1, p. 127, Dec. 2022, doi: 10.1186/s13638-022-02209-0.
[40] D. Zhang, Y. Tang, and X. Yan, "Supply chain risk management of badminton supplies company using decision tree model assisted by fuzzy comprehensive evaluation," *Expert Syst.*, vol. 41, no. 5, May 2024, Art. no. e13275, doi: 10.1111/exsy.13275.
[41] M. R. Ali, S. M. A. Nipu, and S. A. Khan, "A decision support system for classifying supplier selection criteria using machine learning and random forest approach," *Decis. Anal. J.*, vol. 7, Jun. 2023, Art. no. 100238, doi: 10.1016/j.dajour.2023.100238.
[42] J. Zou and H. Li, "Precise marketing of e-commerce products based on KNN algorithm," *Comput. Intell. Neurosci.*, vol. 2022, Aug. 2022, Art. no. 4966439, doi: 10.1155/2022/4966439.
[43] X. Li and Z. Li, "A hybrid prediction model for e-commerce customer churn based on logistic regression and extreme gradient boosting algorithm," *Ingénierie des Systèmes d'Inf.*, vol. 24, no. 5, pp. 525–530, Nov. 2019, doi: 10.18280/isi.240510.
[44] S. L. Zhuang. *Telecom Customer Churn Prediction*. [Online]. Available: http://kaggle.com/shilongzhuang/telecom-customer-churn-by-maven-analytics

MOHAMMAD AFZAL (Member, IEEE) received the B.Tech. degree in mechanical engineering from the Department of Mechanical Engineering, Aligarh Muslim University, Aligarh, India. He is currently with Siemens Energy. He has published three research papers, two in journals and one in an international conference, related to data science and machine learning. His research interests include data science and AI drives, his active participation in projects and hackathons, aiming for innovation. Across various domains, he seeks to leverage these technologies for meaningful advancements, contributing to progress and societal betterment through data-driven methodologies and artificial intelligence solutions.

SHANUR RAHMAN (Member, IEEE) received the B.Tech. degree in computer engineering from the Department of Computer Engineering, Aligarh Muslim University, India. He is currently with MapleSoftware. He has cultivated a portfolio driven by a profound motivation toward machine learning and data science. His accomplishments include a research paper presented at a national conference in the domain of AI, ML, and the IoT. Through diverse experiences in national competitions, internships, and different hackathons in the field of ML, AI, and the IoT, he has honed his skills. With a relentless dedication to pushing the boundaries of technology, particularly in trustworthy machine learning, the IoT, and bioinformatics, he exemplifies promise within the research community. Notably, he was honored with the Best Paper Award for his outstanding research.

DISHA SINGH received the B.Tech. degree in civil engineering from the Department of Civil Engineering, Aligarh Muslim University, Aligarh, India. She is currently interning at a startup dealing with AI in healthcare research topics that provides help to radiologists all around the world. She has published two research papers in international conferences on the frontiers of AI and computer vision and gained substantial experience through national and international internships, contests, and hackathons. With a deep-rooted passion for trustworthy machine learning, computer vision, and AI in healthcare, her dedication to excellence and willingness to push the boundaries of technology makes her a promising member of the research community. She was awarded the Best Paper Award for one of her research articles.

ALE IMRAN (Member, IEEE) received the B.Tech. and M.Tech. degrees (Hons.) in electronics engineering from Aligarh Muslim University (AMU), Aligarh, India, in 2007 and 2010, respectively. He is currently an Assistant Professor with the Z. H. College of Engineering and Technology, Aligarh Muslim University. He is the author or co-author of over 40 research papers published in reputed journals and conference proceedings. His current research interests include technology and RE entrepreneurship and the theme of this work. He is also a member of The Institution of Electronics and Telecommunication Engineers (IETE) and Indian Society of Technical Education (ISTE). He has won best paper awards at prestigious international conferences held at IIT Madras and NTU Singapore.