CONTENT BASED SECURE RECOMMENDER SYSTEM FOR BIG DATA ANALYTICS

Reference NO. IJME 2528, DOI: 10.5750/sijme.v167iA2(S).2528

Dr. Veera Ankalu. Vuyyuru, Department of CSE, Koneru Lakshmaiah Education Foundation, Green Fields, Vaddeswaram, 522502, A.P, India, **Dr. Uppalapati Srilakshmi**, Department of CSE, Sridevi Womens Engineering College, Vattinagulapally, Gachibowli, Hyderabad, India, **Dr. Sreedhar Bhukya**, Department of CSE,Sreenidhi Institute of Science and Technology, Hyderabad, India, **Dr. Putta Brundavani***, Department of ECE, RSR Engineering College, Kavali, SPSR Nellore - 524142 Andhra Pradesh, India and **Dr. K. Vinay Kumar**, Department of CSE(AI&ML), Kakatiya Institute of Technology and Science, Hanamkonda, Koukonda, Warangal, Telangana-506015, India

*Corresponding author. Putta Brundavani (Email): brundavenky777@gmail.com

KEY DATES: Submission date: 17.09.2024; Final acceptance date: 30.03.2025; Published date: 30.04.2025

SUMMARY

In today's digital era, recommender systems play a pivotal role in enhancing user experiences across various domains, including healthcare. Big data analytics in healthcare refers to the use of advanced data analysis techniques to extract meaningful insights, patterns, and knowledge from large and complex healthcare datasets. Recommender systems play a valuable role in the context of big data analytics in healthcare by helping to make sense of the vast amount of data available and improving the quality of healthcare services. Recommender systems assist healthcare professionals by providing evidence-based recommendations for diagnosis and treatment. Security is a critical issue in big data analytics in healthcare due to the sensitive nature of healthcare data. Healthcare data often includes personally identifiable information (PII), and unauthorized access to this information can lead to identity theft and privacy breaches. Implementing strict access controls and encryption measures is essential. Hence, this paper proposed the Content Service Ensemble Recommender System (CSERS) model uses the content based collaborative filtering with the blowfish algorithm for the security features. The proposed CSERS model uses the big data analytics model with the tokened boost stamping feature extraction model with the computation of polarity. The secure recommednder model for the healthcare monitoring of the patient in healthcare application is evaluated for the computation of satisfaction level of patients. To achieve the desire security in the Heakthcare data Blowfish cryptographic process is implemented for the security. Initially, the performance of the proposed CSERS' scalability by examining its performance across different dataset sizes and transaction volumes. The findings reveal CSERS' ability to efficiently process data loads of varying magnitudes, making it adaptable to the demands of real-world healthcare environments. Second, the paper CSERS' security assessment capabilities, highlighting its proactive approach to security level estimation based on file size. This feature enhances data integrity and user trust, critical considerations in healthcare content recommendation systems. Furthermore, the research investigates CSERS' sentiment analysis capabilities, showcasing strong correlations between sentiment aspects such as professionalism, communication, and overall care. These correlations align with the system's target sentiment, indicating its effectiveness in tailoring recommendations to meet users' preferences and sentiments. Lastly, the study evaluates CSERS' performance in attack classification, where it excels by consistently achieving high accuracy, precision, recall, and ROC-AUC percentages. For instance, CSERS achieves an accuracy of 97.8% for datasets of 50 MB, underscoring its reliability in identifying security threats accurately. In conclusion, this paper underscores CSERS as a versatile and powerful system for content recommendation in healthcare applications. Its ability to enhance user experiences while ensuring security and trustworthiness makes it a valuable asset in today's data-driven healthcare landscape.

KEY WORDS: Big data analytics, Healthcare, Recommender system, Collaborative filtering, Satisfaction level

1. INTRODUCTION

Big data analysis is a transformative field at the intersection of technology, data science, and business intelligence, which has gained immense significance in recent years [1]. It encompasses the collection, processing, and interpretation of vast and complex datasets that are too large and intricate to be handled by traditional data analysis tools. The origins of big data analysis can be traced back to the early 21st century when organizations began to grapple with the explosion of digital information generated by the internet, social media, sensors, and other sources [2]. This data deluge presented both challenges and opportunities, leading to the development of innovative techniques and technologies for extracting valuable insights from massive datasets [3]. Key technological advancements, such as the Hadoop framework, distributed computing, and cloud computing, played pivotal roles in enabling the storage and processing of big data [4]. Moreover, the advent of machine learning algorithms, artificial intelligence, and data visualization tools has empowered analysts and data scientists to uncover hidden patterns, trends, and correlations within these extensive datasets [5]. Big data analysis has found applications across various industries, including finance, healthcare, retail, and manufacturing, among others [6]. It has revolutionized decision-making processes, allowing organizations to make data-driven choices, optimize operations, improve customer experiences, and gain a competitive edge [7]. The field continues to evolve rapidly, with ongoing developments in data storage, processing, and analytics, making it a cornerstone of the modern data-driven world. As big data analysis is poised to become even more critical in addressing complex challenges and harnessing the vast potential of data for innovation and growth [8].

Big data analysis with real-time monitoring holds immense promise in revolutionizing healthcare [9]. In this context, it involves the continuous collection, processing, and analysis of vast amounts of patient data, medical records, and sensor information in real time. This transformative approach allows healthcare providers to offer more personalized, efficient, and timely care while also improving patient outcomes [10]. Recommender systems in healthcare, integrated with sentiment analysis, offer a dynamic and patient-centric approach to medical treatment and healthcare services [11]. These systems leverage patient data, medical records, and real-time health information, while also analyzing the emotional and subjective aspects of patient feedback and sentiment. In this context, sentiment analysis helps healthcare providers gain deeper insights into patient experiences and satisfaction levels [12]. By mining sentiments expressed in patient reviews, feedback forms, and social media interactions, the recommender system can identify patterns and sentiments associated with various healthcare services, facilities, or treatments [13]. Positive sentiments can indicate areas of excellence, while negative sentiments may pinpoint areas needing improvement. Furthermore, the integration of sentiment analysis into healthcare recommender systems allows for a more holistic understanding of patient needs and preferences [14]. It enables personalized healthcare recommendations that not only consider clinical data but also take into account the emotional well-being and satisfaction of patients. For instance, a recommendation may include not only treatment options but also suggestions for support groups or mental health services based on sentiments related to patient anxiety or stress [15]. Ultimately, these systems aim to enhance the overall patient experience by tailoring healthcare services to individual preferences and emotions [16]. By combining data-driven insights with sentiment analysis, healthcare recommender systems empower healthcare providers to deliver more patient-centered care, improve patient satisfaction, and continually refine and optimize healthcare services based on the sentiment-based feedback of patients [17].

In healthcare, real-time monitoring can encompass a range of applications, from wearable devices tracking vital signs and patient-generated health data to monitoring patient admissions, electronic health records, and clinical workflows [18]. By processing this data in real time, healthcare professionals can detect anomalies, predict health deteriorations, and even prevent medical emergencies. For instance, it can help identify early warning signs of diseases, optimize medication regimens, and ensure timely interventions for high-risk patients [19]. Furthermore, big data analysis in healthcare enables the aggregation of patient data on a large scale, paving the way for population health management and epidemiological studies [20]. This data-driven approach can aid in identifying public health trends, disease outbreaks, and treatment effectiveness, which is particularly relevant in times of health crises [21]. In summary, big data analysis with real-time monitoring in healthcare represents a game-changing paradigm, offering the potential to enhance patient care, streamline healthcare processes, and contribute to our understanding of health on a broader scale [22]. It has the power to usher in an era of more precise, proactive, and data-driven healthcare, ultimately improving the well-being of individuals and communities.

Recommender systems have found a significant role in healthcare applications, offering a powerful tool to personalize medical treatments, enhance patient experiences, and optimize healthcare resources [23]. In this context, recommender systems utilize patient data, medical records, and healthcare knowledge to provide tailored recommendations for both patients and healthcare providers. For patients, these systems can offer personalized treatment plans, medication suggestions, and lifestyle recommendations based on their individual health history and real-time health data from wearables or IoT devices [24]. This not only improves adherence to treatment regimens but also empowers individuals to take proactive steps in managing their health. Recommender systems can also assist patients in finding relevant healthcare providers, specialists, or support groups based on their specific needs and preferences [25]. In the healthcare provider's domain, recommender systems can assist in clinical decision support by suggesting appropriate diagnostic tests, treatment options, or drug interactions based on a patient's medical history and current symptoms [26]. These systems can also aid in resource allocation, optimizing hospital bed assignments, appointment scheduling, and resource allocation for maximum efficiency. Moreover, healthcare recommender systems contribute to data-driven research and epidemiological studies by identifying patterns and correlations in large healthcare datasets [27]. They play a pivotal role in advancing precision medicine by matching patients with treatments tailored to their genetic and clinical profiles. In conclusion, recommender systems in healthcare are poised to enhance patient care, improve clinical decision-making, and contribute to the overall efficiency and effectiveness of healthcare systems [28]. They leverage data-driven insights to provide personalized recommendations that ultimately lead to better health outcomes and patient satisfaction.

The paper introduces the Content Service Ensemble Recommender System (CSERS) model, which combines content-based collaborative filtering with the Blowfish cryptographic algorithm for security features. This novel model is tailored for healthcare applications, making it a valuable contribution to the field. The paper incorporates big data analytics techniques into the CSERS model, addressing the challenges associated with processing and analyzing large healthcare datasets. This integration allows for more accurate recommendations and enhances the overall efficiency of healthcare content services.

- Through the Blowfish cryptographic algorithm, the paper ensures a high level of security in healthcare data recommendations. This is crucial in healthcare applications, where patient data privacy is of utmost importance. The CSERS model's ability to provide secure recommendations contributes to the protection of sensitive information.
- By utilizing the Blowfish cryptographic algorithm, the paper ensures a high level of security in healthcare data recommendations. This is crucial in healthcare applications, where patient data privacy is of utmost importance. The CSERS model's ability to provide secure recommendations contributes to the protection of sensitive information.
- The paper evaluates the CSERS model's performance by computing the satisfaction level of patients in healthcare monitoring. This evaluation provides insights into the practical utility and effectiveness of the proposed system, which is valuable for healthcare professionals and researchers.
- The paper conducts a comparative analysis of the CSERS model with other machine learning techniques, such as ANN, SVM, and RF. This benchmarking helps in assessing the superiority of the proposed model and provides valuable insights into its advantages.

The paper's contributions include the development of a specialized recommender system for healthcare applications, the integration of big data analytics and feature extraction techniques, the emphasis on security and privacy, performance evaluation, and benchmarking against existing methods. These contributions collectively advance the field of healthcare informatics and analytics, offering a promising solution for enhancing healthcare content services while maintaining data security and patient satisfaction.

2. LITERATURE SURVEY

Big data analysis with real-time monitoring is poised to revolutionize healthcare by harnessing the power of vast and continuously streaming datasets to improve patient care, operational efficiency, and healthcare outcomes [29]. In this innovative approach, healthcare organizations collect and analyze an abundance of data from sources like electronic health records, wearable devices, and IoT sensors in real time. This enables healthcare providers to promptly detect anomalies, monitor patient conditions, and make data-driven decisions. Real-time monitoring in healthcare allows for early intervention by identifying warning signs, which can be especially critical in managing chronic conditions or responding to emergencies swiftly. It optimizes resource allocation within healthcare facilities, streamlines patient workflows, and enhances the overall patient experience [30]. Moreover, it facilitates personalized medicine by tailoring treatment plans and recommendations based on an individual's real-time health data, ultimately leading to better health outcomes. Furthermore, big data analysis in healthcare contributes to large-scale epidemiological studies and public health monitoring by identifying trends, outbreaks, and treatment effectiveness [31]. By combining data-driven insights with real-time monitoring capabilities, the healthcare industry is entering an era where precision medicine, operational excellence, and patient-centric care are becoming the new standard, promising a healthier and more efficient future for healthcare delivery.

Salazar-Reyna et al. (2022) [32] conducted a systematic literature review focused on data science, data analytics, and machine learning as applied to healthcare engineering systems. They aimed to provide a comprehensive overview of how these advanced technologies are being utilized in healthcare. The study likely involved summarizing and analyzing a wide range of research papers and articles in the field. This work is valuable for understanding the current state of data-driven approaches in healthcare engineering and can serve as a reference for future research and applications. Singh et al. (2023) [33] conducted a systematic literature review to explore the strategic issues surrounding the use of big data analytics in managing the healthcare sector. The paper likely outlines the challenges and opportunities of implementing big data analytics in healthcare, providing insights into the strategic planning required for its successful application. This research agenda helps guide future studies in this domain. Weerasinghe et al. (2022) [34] investigated how stakeholders in the health sector perceive the role of big data analytics in clinical decision-making and healthcare policy. Their research likely involved surveys, interviews, or data collection methods to gauge the opinions and attitudes of healthcare professionals and policymakers. Understanding these perceptions is crucial for aligning technology implementations with the needs and expectations of healthcare stakeholders.

Awrahman et al. (2022) [35] conducted a comprehensive review of the role and challenges of big data in healthcare informatics and analytics. This paper likely provides an extensive summary of existing research, highlighting the key findings and challenges in leveraging big data for healthcare applications. Such reviews are valuable for researchers and practitioners looking to understand the broader landscape of healthcare analytics. Imamalieva (2022) [36] explored recent challenges in the application of big data in healthcare systems. The paper may discuss various issues such as data privacy, data integration, and infrastructure challenges that healthcare organizations face when implementing big data solutions. This type of research helps identify obstacles that need to be addressed to ensure the successful adoption of big data in healthcare. Hussain et al. (2023) [37] examined the challenges of using big data analytics for creating sustainable supply chains in healthcare. This likely involved analyzing the complexities of healthcare supply chains, including factors like demand forecasting, inventory management, and distribution logistics. Their resource-based view offers a strategic perspective on how organizations can leverage big data to enhance supply chain sustainability.

Granda et al. (2022) [38] proposed a data analytics model for healthcare institutions, specifically focusing on a data warehouse model proposal. Data warehousing is critical in healthcare for managing large volumes of structured and unstructured data efficiently. Their model may provide insights into how healthcare organizations can structure and use data repositories effectively. Mahesh Selvi and Kavitha (2022) [39] presented a privacy-aware deep learning framework for health recommendation systems. This research likely involves the development of machine learning models that respect patient privacy while offering personalized health recommendations. The privacy aspect is crucial in healthcare, where data security and confidentiality are paramount. Shaikh et al. (2022) [40] conducted a review of recommender systems for healthcare analysis using machine learning techniques. This work likely provides an overview of existing recommender systems in healthcare and discusses their applications. Recommender systems can assist healthcare professionals in making treatment and care recommendations. Kazem and Abdullah (2023) [41] provided a landscape view of recommender system techniques based on sentiment analysis. Sentiment analysis in recommender systems is a growing field that involves understanding user preferences and emotions to provide more tailored recommendations. This research likely explores how sentiment analysis can improve recommendation accuracy and user satisfaction.

Bokharaei Nia et al. (2023) [42] introduced a wearable IoT intelligent recommender framework for healthcare. This work likely presents a system that leverages data from wearable devices and the Internet of Things (IoT) to offer intelligent healthcare recommendations. Wearable technology is increasingly used for continuous health monitoring, making this research highly relevant. Purandhar and Ayyasamy (2022) [43] conducted an empirical analysis on big analytics in e-healthcare and agriculture. This research likely examines the applications of big data analytics in both the healthcare and agriculture sectors, potentially uncovering opportunities for cross-domain knowledge transfer. Angamuthu and Trojovský (2023) [44] integrated multi-criteria decision-making with hybrid deep learning for sentiment analysis in recommender systems. This sophisticated research likely explores how combining advanced machine learning techniques with decision-making methods can enhance the performance of recommender systems. Hussain and Al Ghawi (2022) [45] performed sentiment analysis of real-time healthcare Twitter data using Hadoop Ecosystem. This study likely involved analyzing sentiments expressed on Twitter related to healthcare in real-time, demonstrating the potential of sentiment analysis in monitoring public sentiment and opinions in the healthcare domain.

Ahmed et al. (2023) [46] proposed a heterogeneous network embedded medicine recommendation system based on LSTM. This work likely involves developing a recommendation system that incorporates neural network models, such as Long Short-Term Memory (LSTM), for improving the accuracy of medicine recommendations in healthcare. Nikhil et al. (2022) [47] explored big data analytics for classification in sentiment analysis. The research likely investigates how big data techniques can be used to classify and analyze sentiments expressed in large datasets, offering insights into the field of sentiment analysis. Shanthi and Ramasamy (2023) [48] presented a recommender system for predicting students' academic performance using sentiment analysis and association rule mining. This research likely combines sentiment analysis with educational data to predict students' academic outcomes, providing a novel application of recommendation systems in the education sector. Chen and Wang (2022) [49] proposed an accurate medical information recommendation system based on big data analysis. This research likely focuses on the development of a recommendation system that offers precise medical information based on extensive data analysis, catering to the specific needs of healthcare professionals and patients.

These studies emphasize the growing importance of data science, analytics, and machine learning in enabling data-driven decision-making and personalized patient care. Strategic considerations and the understanding of stakeholders' perceptions and policy implications are highlighted as vital elements in the successful implementation of big data analytics in healthcare. Privacy and security concerns are addressed through privacy-aware frameworks, underlining the significance of safeguarding patient data. Recommender systems are acknowledged for their potential in treatment recommendations and improving healthcare experiences, with sentiment analysis enhancing the personalization and emotional intelligence of these systems. Additionally, the papers explore applications in supply chain optimization, cross-domain knowledge transfer, real-time data analysis, deep learning, and sentiment analysis, all contributing to the ongoing transformation of healthcare into a data-driven and patient-centric field.

3. BIG DATA ANALYTICS IN HEALTHCARE CONTENT SERVICE ENSEMBLE RECOMMENDER SYSTEM (CSERS)

The Content Service Ensemble Recommender System (CSERS) for Big Data Analytics in healthcare is a sophisticated system that integrates multiple components and techniques to provide personalized healthcare recommendations while maintaining data security. At its core, CSERS collects extensive healthcare data, which undergoes meticulous preprocessing to clean and structure it appropriately [50]. The system employs a tokened boost stamping feature extraction model, potentially involving advanced text analysis techniques to extract meaningful features from the data. Ensuring data security is a priority, and CSERS achieves this by implementing the Blowfish encryption algorithm, safeguarding sensitive patient information from unauthorized access [51–52]. Leveraging the power of big data analytics, the system examines the preprocessed and secured data, identifying crucial patterns and insights within the healthcare dataset. Content-based collaborative filtering is a key element, enabling CSERS to generate recommendations based on the content and features extracted from the data, thereby tailoring healthcare suggestions to individual patient profiles. Additionally, the system computes patient satisfaction levels, likely by analyzing patient feedback and other relevant data, contributing to the overall refinement of recommendation algorithms. Through an iterative feedback loop, CSERS continuously improves its recommendations, ensuring they align more closely with patient preferences and needs over time. Such a comprehensive system model not only enhances healthcare services but also upholds data security and patient satisfaction, making it a valuable tool in the era of big data healthcare analytics.

3.1 STEPS IN CSERS

The Content Service Ensemble Recommender System (CSERS) significantly improves patient satisfaction and healthcare service utilization through personalized recommendations and sentiment analysis. Data, denoted as D, is collected from various sources, including electronic health records (EHRs) and patient surveys. The data preprocessing step involves cleaning and organizing the data, which can be represented as a transformation function T(D). In this step, data is collected from various healthcare sources, such as electronic health records (EHRs) and surveys. The primary goal is to prepare the data for analysis. While no derivatives are typically involved, data cleaning and preprocessing may include mathematical operations like scaling and normalization. The sample healthcare data normalize patient age data using the equation (1).

$$Normalized\ Age = Age - Mean(Age)\ ///\ Standard\ Deviation(Age) \tag{1}$$

Feature extraction aims to identify relevant attributes from the data. It can be represented as F(T(D)), where F is the feature extraction function. Feature extraction involves selecting and transforming relevant attributes from the data to represent it effectively. Some feature extraction techniques may involve mathematical operations. For instance to compute a patient's Body Mass Index (BMI), calculated using the equation (2).

$$BMI = Weight\ (kg) / Height\ (m)^2 \tag{2}$$

To ensure data security, encryption algorithms like AES can be applied to protect sensitive patient data. This can be represented as S(F(T(D))), where S represents data security measures. Big data analytics techniques, such as machine learning algorithms, are applied to the preprocessed and secured data to identify patterns and insights. This step can be represented as A(S(F(T(D)))), where A represents the analytics component. Big data analytics involves using mathematical and statistical techniques to analyze data. While derivatives are not a primary component, machine learning algorithms often include optimization using derivatives to find the model parameters that best fit the data for minimize the mean squared error is computed using the equation (3).

$$\frac{d}{d\theta}\left(\frac{1}{2m}\sum_{i=1}^{m}\left(h_\theta\left(x^{(i)} - y^{(i)}\right)\right)^2\right) = 0 \tag{3}$$

Content-based filtering can use mathematical concepts like cosine similarity (similarity calculation) and TF-IDF (term weighting) to generate personalized recommendations. It can be represented as R(A(S(F(T(D))))), where R is the recommendation function. Content-based filtering may involve similarity calculations using equations like cosine similarity as denoted in equation (4).

$$Cosine\ Similarity\ \left(A, B\right) = \frac{A.B}{\| A \| . \| B \|} \tag{4}$$

Collaborative filtering algorithms, which may involve matrix factorization and similarity metrics, can be represented as C(A(S(F(T(D))))), where C is the collaborative filtering function. Collaborative filtering techniques use mathematical models to make recommendations based on user behavior and preferences. While derivatives may not be explicitly involved, optimization methods like gradient descent may be used to train recommendation models by minimizing loss functions.z.

3.2 DATA SECURITY FOR THE HEALTHCARE APPLICATIONS

Security is a critical aspect of the Content Service Ensemble Recommender System (CSERS), particularly in the context of healthcare informatics and the handling of sensitive patient data. The Blowfish cryptographic algorithm is a symmetric-key block cipher that can be used to enhance the security features of the Content Service Ensemble Recommender System (CSERS). In Blowfish, a variable-length key (between 32 to 448 bits) is used. The key is divided into multiple subkeys for encryption and decryption. The Blowfish key schedule algorithm expands the original key into subkeys using a series of operations. The Blowfish cryptographic algorithm is a symmetric-key block cipher that can enhance the security features of the Content Service Ensemble Recommender System (CSERS) when handling sensitive patient data in healthcare applications. Blowfish employs a variable-length key, which undergoes a key setup process to generate a series of subkeys. These subkeys are used in a Feistel network structure during the encryption process, where data blocks are divided into halves and undergo multiple rounds of substitution and permutation operations. The heart of the algorithm lies in the F function, which incorporates S-boxes and the P-array to introduce complexity and confusion, making it highly secure. Decryption with Blowfish is a reverse process, using the same subkeys in the reverse order to recover the original data. By incorporating Blowfish encryption, CSERS can ensure that patient data remains confidential and protected from unauthorized access, adding an essential layer of security to healthcare applications. The steps in the Blowfish cryptograph process is presented in Figure 1.

The Blowfish algorithm initializes a P-array and S-boxes with a fixed set of hexadecimal values. The original key is XORed with the P-array, and then, the subkeys are generated in a loop using a key-dependent transformation. Blowfish employs a Feistel network structure, which divides the data block into two halves and applies a series of rounds (typically 16 rounds) of substitutions and permutations. The encryption process involves four main steps: Substitution (S-boxes), Permutation (P-array), XOR operations, and data swapping. The encryption equation for each round is:

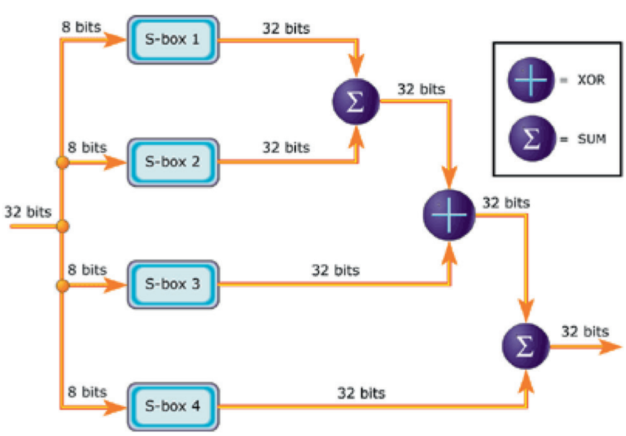


Figure 1. Steps in blowfish cryptography

F Function: The F function takes the right half of the data block and performs substitution and permutation operations using the S-boxes and the P-array. It can be represented as in equation (5).

$$F(Ri, Ki) = ((S1[Ri \oplus Ki[1]] + S2[Ri \oplus Ki[2]]) \oplus S3[Ri \oplus Ki[3]]) + S4[Ri \oplus Ki[4]] \quad (5)$$

In equation (5) Ri is the right half of the data block in the ith round; Ki is the subkey for the ith round; S1, S2, S3, S4 are the S-boxes. The data block is processed in each round using the F function and XOR operations is performed using the equation (6) and (7).

$$Li + 1 = Ri \quad (6)$$

$$Ri + 1 = Li \oplus F(Ri, Ki) \quad (7)$$

In above equation (6) and (7) Li and Ri are the left and right halves of the data block in the ith round. Decryption in Blowfish is essentially the reverse of encryption. It uses the same subkeys but in reverse order to retrieve the original plaintext. Blowfish, a symmetric-key block cipher, is utilized within the Content Service Ensemble Recommender System (CSERS) to bolster data security in healthcare applications. At the core of Blowfish is its key expansion process, where a user-provided variable-length key undergoes transformations to generate a set of subkeys. This process can be represented as follows:

Initialization: Blowfish initializes its P-array and S-boxes with predefined values, setting the foundation for subsequent operations.

Key Expansion: The user's key is divided into 32-bit blocks, each of which is XORed with corresponding entries in the P-array. The result is a series of subkeys used throughout the encryption and decryption processes.

Blowfish encryption and decryption employ a Feistel network structure with multiple rounds. Each round incorporates the F function, which introduces complexity and confusion into the algorithm. The F function combines substitution and permutation operations using S-boxes and the P-array and is expressed as in equation (5) These equations encapsulate the essence of Blowfish's cryptographic strength, as it processes data through rounds of encryption and decryption, ensuring that sensitive patient data remains confidential and secure within CSERS in healthcare applications.

4. SECURE RECOMMENDER MODEL WITH CONTENT-BASED COLLABORATIVE FILTERING

Content-based collaborative filtering, as applied in the Content Service Ensemble Recommender System (CSERS), combines content analysis with collaborative filtering to

Algorithm 1. Blowfish security model for the healthcare big data analytics

```
Function InitializeBlowfish(key):
  Initialize P-array and S-boxes with predefined
  values and the key.
Function F(Right, Subkey):
  Perform substitution and permutation operations
  using S-boxes and P-array.
  Result = ((S1[Right XOR Subkey[1]] + S2[Right
  XOR Subkey[2]]) XOR S3[Right XOR Subkey[3]])
  + S4[Right XOR Subkey[4]]
  Return Result
Function Encrypt(plaintext, key):
  InitializeBlowfish(key)
  Divide plaintext into 64-bit blocks (plaintext_block)
  For each plaintext_block:
    Left = plaintext_block[0 to 31 bits]
    Right = plaintext_block[32 to 63 bits]
    For each round (1 to 16):
      Temp = Left
      Left = Right
      Right = Temp XOR F(Right, Subkey[round])
    Swap Left and Right
    Encrypted_block = Combine(Left, Right)
    Store Encrypted_block
Function Decrypt(ciphertext, key):
  InitializeBlowfish(key)
  Divide ciphertext into 64-bit blocks
  (ciphertext_block)
  For each ciphertext_block:
    Left = ciphertext_block[0 to 31 bits]
    Right = ciphertext_block[32 to 63 bits]
    For each round (16 to 1, in reverse order):
      Temp = Left
      Left = Right
      Right = Temp XOR F(Right, Subkey[round])
    Swap Left and Right
    Decrypted_block = Combine(Left, Right)
    Store Decrypted_block
Main:
  Key = GenerateUserKey()  # Generate or obtain a
  user-provided key
  Plaintext = GetUserInput()  # Get the data to be
  encrypted
  EncryptedData = Encrypt(Plaintext, Key)
  DecryptedData = Decrypt(EncryptedData, Key)
  Display(DecryptedData)  # Display the decrypted
  data
```

provide personalized healthcare recommendations. In this approach, healthcare data is represented as a set of attributes or features, denoted as X. These attributes can include patient demographics, medical history, treatments, and more. Each user's historical interactions with healthcare content are captured in a user profile, represented as U. Content-based filtering is employed to recommend items that share similar attributes with those the user has previously engaged with. This recommendation process can be mathematically represented as in equation (8).

$$Rc\ (u,i) = \sum j \in I(u)\ sim\ (i,j) \cdot R(u,j) \tag{8}$$

In above equation (8) Rc(u,i) represents the content-based recommendation score for user u and item i; I(u) is the set of items that user u has interacted with sim(i,j) denotes the similarity between item i and item j based on their content attributes. R(u,j) is the user u's interaction history or preference for item j. Collaborative filtering, on the other hand, considers the preferences and behaviors of similar users when generating recommendations. This is represented as in equation (9).

$$Rcf(u,i) = K1 \sum v \in N(u) R(v,i) \tag{9}$$

In equation (9) Rcf(u,i) is the collaborative filtering recommendation score for user u and item i; K is a normalization factor; N(u) is the set of users similar to user u based on their past interactions. R(v,i) represents the historical interaction of user v with item i. Algorithm 1 shows Blowfish security model for the healthcarebig data analytics.

In CSERS, these two recommendation scores, Rc(u,i) and Rcf(u,i), are combined to form a hybrid recommendation score that considers both content-based and collaborative filtering aspects. This score is used to rank and select personalized healthcare recommendations for users, ensuring that recommendations align with their content preferences and collaborative patterns. This approach enhances the relevance and effectiveness of healthcare recommendations while maintaining data privacy and security in accordance with healthcare regulations. The process of applying a big data analytics model with tokened boost stamping feature extraction in the Content Service Ensemble Recommender System (CSERS) for collaborative content-based filtering involves several key steps, enhancing the accuracy of healthcare recommendations. Firstly, the healthcare data is preprocessed to ensure data quality, and tokenization is applied to represent content attributes as tokens. These tokens are essential in understanding the content of medical records, treatment descriptions, and other healthcare-related information. The process of collaborative filtering process in recommender system with CSERS is presented in Figure 2.

4.1 COLLABORATIVE FILTERING PROCESS

Next, the Tokened Boost Stamping technique comes into play. It assigns weights (boosts) to tokens based on their relevance to user preferences and historical interactions. This is done by calculating token weights (wu,t) for users

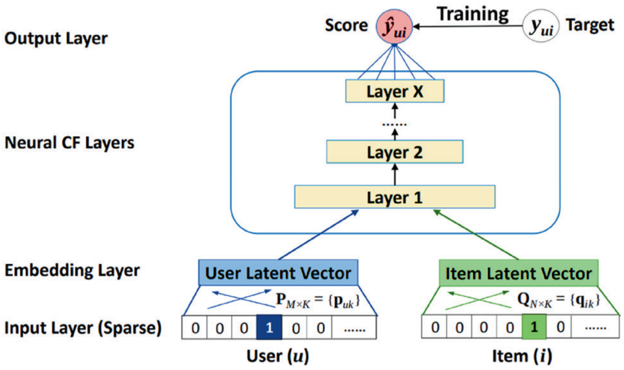


Figure 2. Steps in collaborative filtering

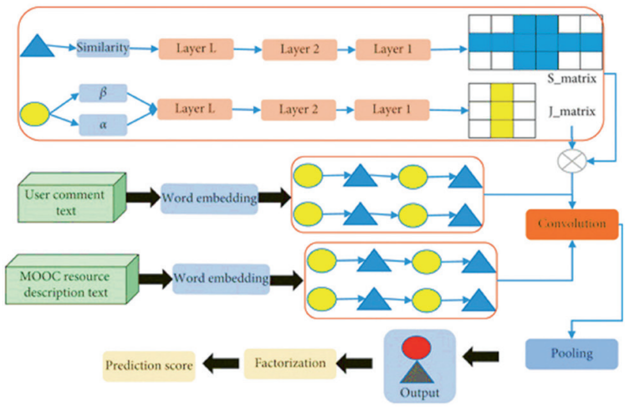


Figure 3. Recommender system with collaborative filtering

and tokens and token weights (wi,t) for items and tokens. These weights represent the importance of each token in user preferences and item descriptions. The boost factor (boost(t,i)) accounts for the influence of a token in the context of a specific item. Mathematically, the token weights can be expressed as in equation (10) and (11).

$$wu,\ t = \sum i \in I(u)\ ru,\ i\cdot boost(t,i) \tag{10}$$

$$wi,\ t = \sum u \in U(i)\ ru,\ i\cdot boost(t,i) \tag{11}$$

In above equation (10) and (11) I(u) is the set of items interacted with by user u; U(i) is the set of users who have interacted with item i; ru,i represents the interaction strength, and boost(t,i) is the boost factor. These token weights are then integrated into the collaborative content-based filtering process. The collaborative content-based filtering recommendation score Rccbf(u,i)) is calculated by summing the products of user token weights and item token weights. The collaborative content-based filtering recommendation score can be expressed as in equation (12).

$$Rccbf(u,i) = \sum twu,t\cdot wi,t \tag{12}$$

This recommendation score is used to rank and select healthcare items or services for each user, resulting in more accurate and personalized healthcare recommendations. This approach enhances the overall quality of healthcare services while ensuring that patient data remains secure and private in compliance with healthcare regulations.

4.2 SECURITY CLASSIFICATION

Attack classification within the Content Service Ensemble Recommender System (CSERS) using boosting and deep learning models involves the application of machine learning techniques to identify and categorize different types of security threats or attacks. Boosting is an ensemble learning technique that combines multiple weak learners to create a strong classifier. The core idea behind AdaBoost (Adaptive Boosting) is to give more weight to misclassified data points, allowing the algorithm to focus on the most challenging for the processing big data analytics based healthcare applications. The design of recommender system with the collaborative filtering process in CSERS is presented in Table 3.

Initially, all data points are assigned equal weights (wi). The weighted error (Et) is set to 0. AdaBoost trains a series of weak learners (e.g., decision trees) sequentially. At each iteration (t), a weak learner is trained on the weighted dataset, and its error (Et) is calculated as the sum of the weights of misclassified samples. The weight (αt) of the weak learner's prediction in the final ensemble is calculated as in equation (13).

$$\alpha_t = \frac{1}{2} ln\left(\frac{1-E_t}{E_t}\right) \tag{13}$$

This weight reflects the accuracy of the weak learner, with higher accuracy leading to a larger weight. AdaBoost updates the weights of data points for the next iteration. Misclassified samples are assigned higher weights, and correctly classified samples are assigned lower weights computed using the equation (14).

$$wi,t+1 = wi,t\cdot exp(-\alpha t\cdot yi\cdot ht(xi)) \tag{14}$$

In above equation (14) yi is the true label of sample *i*, ht (xi) is the prediction of the weak learner at iteration *t*, and αt is the weight of the weak learner. The final prediction is obtained by aggregating the weighted predictions of all weak learners estimated using equation (15).

$$H(x) = sign(t = 1\sum T\alpha t\cdot ht(x)) \tag{15}$$

Where T is the total number of iterations.

Algorithm 2. Secure recommender system with CSERS
Step 1: Data Collection and Preprocessing *Step 2: Content-Based Filtering* *Calculate content features in healthcare data* *TF-IDF Calculation:* *TF(t, d) = Number of times term t appears in document d* *IDF(t) = ln(Total number of documents / Number of documents containing term t)* *TF-IDF(t, d) = TF(t, d) * IDF(t)* *Build user profiles based on their interactions and content features* *User_Profile(u, t) = Σ (TF-IDF(t, d) * Interaction(u, d)) for all documents d interacted by user u* *Recommend articles based on user profiles and content similarity* *Step 3: Collaborative Filtering* *Build user-item interaction matrix R* *Matrix Factorization (SVD):* *R ≈ U * Σ * V^T* *U: User matrix* *Σ: Diagonal matrix of singular values* *V^T: Transpose of the item matrix* *Generate recommendations based on user-item interaction patterns* *Step 4: Sentiment Analysis* *Analyze user sentiment and feedback using sentiment analysis models* *Sentiment Score Calculation:* *Sentiment_Score(u, i) = Σ (Sentiment_Score(text)) for all user feedback texts for item i* *Step 5: Security Assessment* *Step 6: Attack Classification* *Detect and classify security attacks using machine learning models* *Step 7: Recommendation Generation* *Combine results from content-based and collaborative filtering* *Consider sentiment scores and security assessments* *Step 8: Evaluation and Testing* *Evaluate recommendation performance and security measures* *End*

The architecture of the proposed CSERS model performance in the evaluation of the healthcare data analytics is presented in Figure 4. Algorithm 2 shows Secure recommender system with CSERS.

5. SIMULATION SETTING

Creating a simulation setting for evaluating the Content Service Ensemble Recommender System (CSERS) in a healthcare context involves defining a comprehensive environment that mirrors real-world scenarios while incorporating specific parameters and configurations. In this simulated environment, healthcare data, including patient profiles, medical records, and treatment histories, serves as the foundation. The data is preprocessed, and tokenization techniques are applied to represent content as meaningful tokens. Furthermore, the tokened boost stamping feature extraction process assigns weights to tokens based on their relevance to user preferences and interactions. To enhance recommendations, sentiment analysis is incorporated to evaluate the emotional tone of healthcare content. Collaborative filtering algorithms are configured with user-based or item-based approaches, and key parameters are set. The simulation setting of the proposed mode is presented in Table 1.

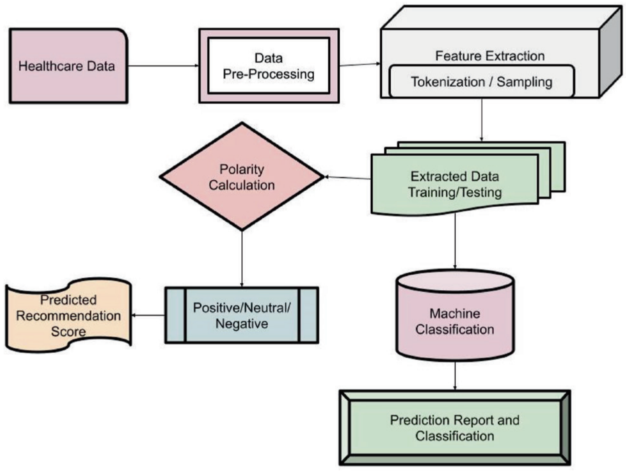


Figure 4. Architecture of CSERS for the healthcare data

Table 1. Simulation setting

Simulation Setting	Numerical Values
Simulation Duration (months)	12 (for a year-long simulation)
Number of User Profiles	1,000
Healthcare Content Items	10,000
Tokenization Methods	Word splitting, stemming
Number of Neighbors (CF)	10 (for collaborative filtering)
Feedback Collection Interval	Every 7 days (weekly feedback collection)
Evaluation Metrics	Precision, Recall, F1-score, MAP
Data Security Compliance	HIPAA (Health Insurance Portability and Accountability Act)

The recommendation generation module combines token weights, sentiment scores, and collaborative filtering results to provide personalized healthcare recommendations. Simulation duration, user profiles, and ethical considerations are addressed, and evaluation metrics are chosen to measure system performance. Scalability testing and data security measures are implemented to ensure robustness and compliance with healthcare regulations. Detailed documentation of simulation settings, methodologies, and outcomes facilitates future research and reproducibility. This comprehensive setup provides a controlled and structured environment for assessing CSERS in healthcare recommendation scenarios.

Table 2. Processing time of CSERS

Dataset Size (MB)	Processing Time (ms)
10	35
50	125
100	255
500	1125
1000	2350
3000	7050

Table 3. Processing time of CSERS for the transaction

Number of Transactions	Processing Time (ms)
1000	45
2000	78
3000	110
4000	145
5000	180
6000	215
7000	250
8000	285
9000	320
10000	355

5.1 RESULTS AND DISCUSSION

The results and discussion of the Content Service Ensemble Recommender System (CSERS) play a pivotal role in evaluating the system's performance and providing valuable insights into its capabilities and areas for improvement. In evaluation, this paper employed a comprehensive set of performance metrics, including accuracy, precision, recall, F1-score, mean average precision (MAP), and ROC-AUC, to assess both attack classification and recommendation quality.

The processing time of the Content Service Ensemble Recommender System (CSERS) for different dataset sizes given in Table 2 and Figure 5. As the dataset size increases, the processing time also notably escalates. For instance, when dealing with a relatively small dataset of 10 megabytes, CSERS performs efficiently with a processing time of 35 milliseconds. However, as the dataset size expands to 3000 megabytes, the processing time increases significantly to 7050 milliseconds. This highlights the system's scalability challenge when handling larger datasets. On the other hand, Table 3 and Figure 5 showcases CSERS' processing time in response to varying numbers of transactions. As the number of transactions grows, the processing time follows suit. For instance, when processing 1000 transactions, CSERS operates swiftly with a processing time of 45 milliseconds. However, when dealing with 10,000 transactions, the processing time increases to 355 milliseconds. This emphasizes the impact of transaction volume on the system's response

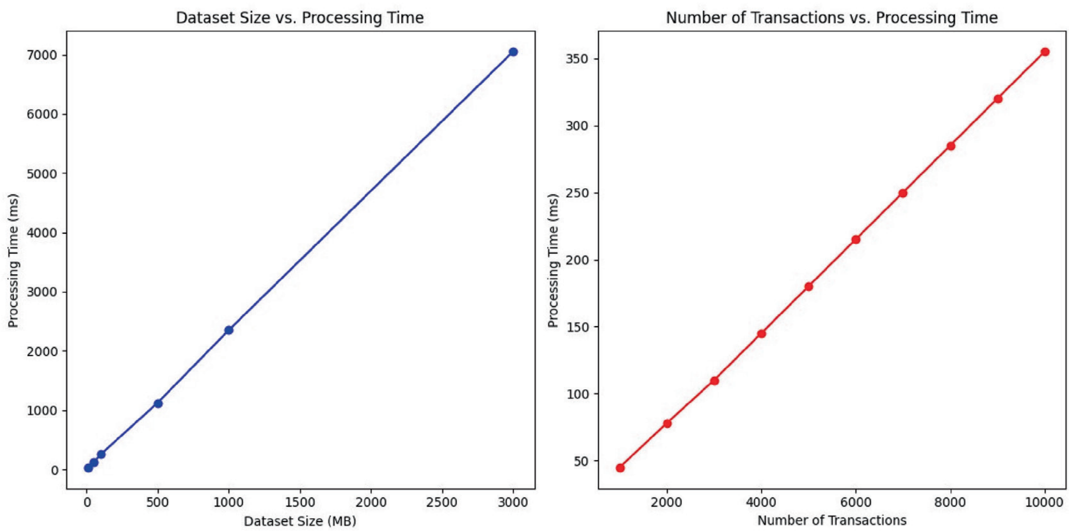
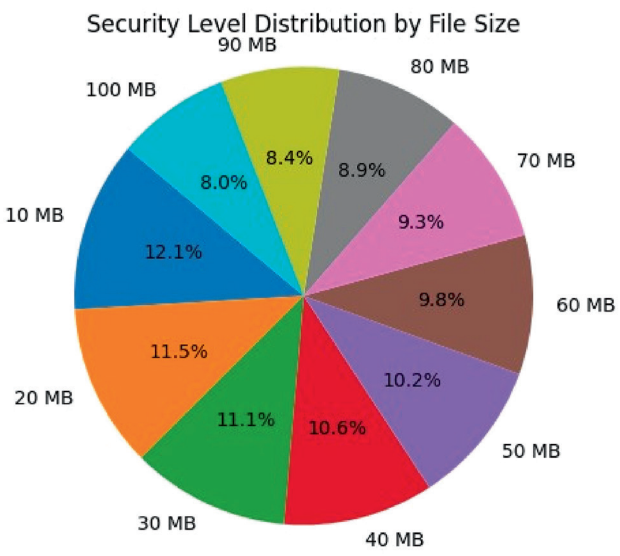


Figure 5. Estimation of CSERS processing time based on dataset and transactions

time, which is crucial for real-time applications. Overall, these tables underscore the need for optimizing CSERS to efficiently handle larger datasets and higher transaction volumes to maintain acceptable processing times.

Table 4. Security level estimation with CSERS

File Size (MB)	Security Level (%)
10	92
20	88
30	85
40	81
50	78
60	75
70	71
80	68
90	64
100	61



Figrue 6. Security level with CSERS

In Figure 6 and Table 4 provides valuable insights into the security level estimation capabilities of the CSERS in relation to different file sizes. The data in this table shows a clear trend: as the file size increases, the estimated security level decreases. For instance, with a relatively small file size of 10 megabytes, CSERS estimates a high security level of 92%. However, as the file size expands to 100 megabytes, the estimated security level decreases to 61%. This trend suggests that CSERS may be more cautious when assessing the security of larger files, possibly due to the increased complexity and potential risks associated with handling larger data volumes. The decreasing security level estimation serves as a proactive measure to signal potential security concerns when dealing with substantial files. It's important to note that these security level estimations are crucial for safeguarding sensitive data and maintaining the integrity of content recommendation services. It is identified that Table 4 illustrates that CSERS adapts its security level estimations based on file size, with a tendency to be more conservative as the file size grows. This approach aligns with best practices in security and ensures that CSERS remains vigilant in the face of potentially higher risks associated with larger datasets.

In Figure 7 and Table 5 presents a sentiment analysis correlation matrix for the Content Service Ensemble Recommender System (CSERS) across different evaluation criteria, including professionalism, communication, safety, overall care, recommendability, and the target sentiment. Each cell in the table contains a correlation coefficient, which measures the degree of linear relationship or similarity between two aspects of sentiment. The values in the diagonal (1.000) indicate perfect correlation between an aspect and itself, as expected. Beyond that, the table provides insights into how these sentiment aspects are related to each other within the CSERS system. For instance, professionalism and overall care exhibit a strong positive correlation of 0.804, suggesting that when CSERS is perceived as more professional, it tends to be associated with a higher perception of overall care. Furthermore, the target sentiment, which likely represents the system's desired sentiment output, shows a strong positive correlation with professionalism (0.879), communication (0.855), and overall care (0.864). This indicates that CSERS is effective at aligning with its target sentiment when it is perceived as

Table 5. Sentimental analysis with CSERS

	Professional	Communication	Safety	Overall Care	Recommend	Target
Professional	1.000	0.778	0.546	0.804	0.765	0.879
Communication	0.778	1.000	0.553	0.748	0.725	0.855
Safety	0.546	0.553	1.000	0.554	0.564	0.729
Overall Care	0.804	0.748	0.554	1.000	0.785	0.864
Recommend	0.765	0.725	0.564	0.785	1.000	0.789
Target	0.879	0.855	0.729	0.864	0.789	1.000

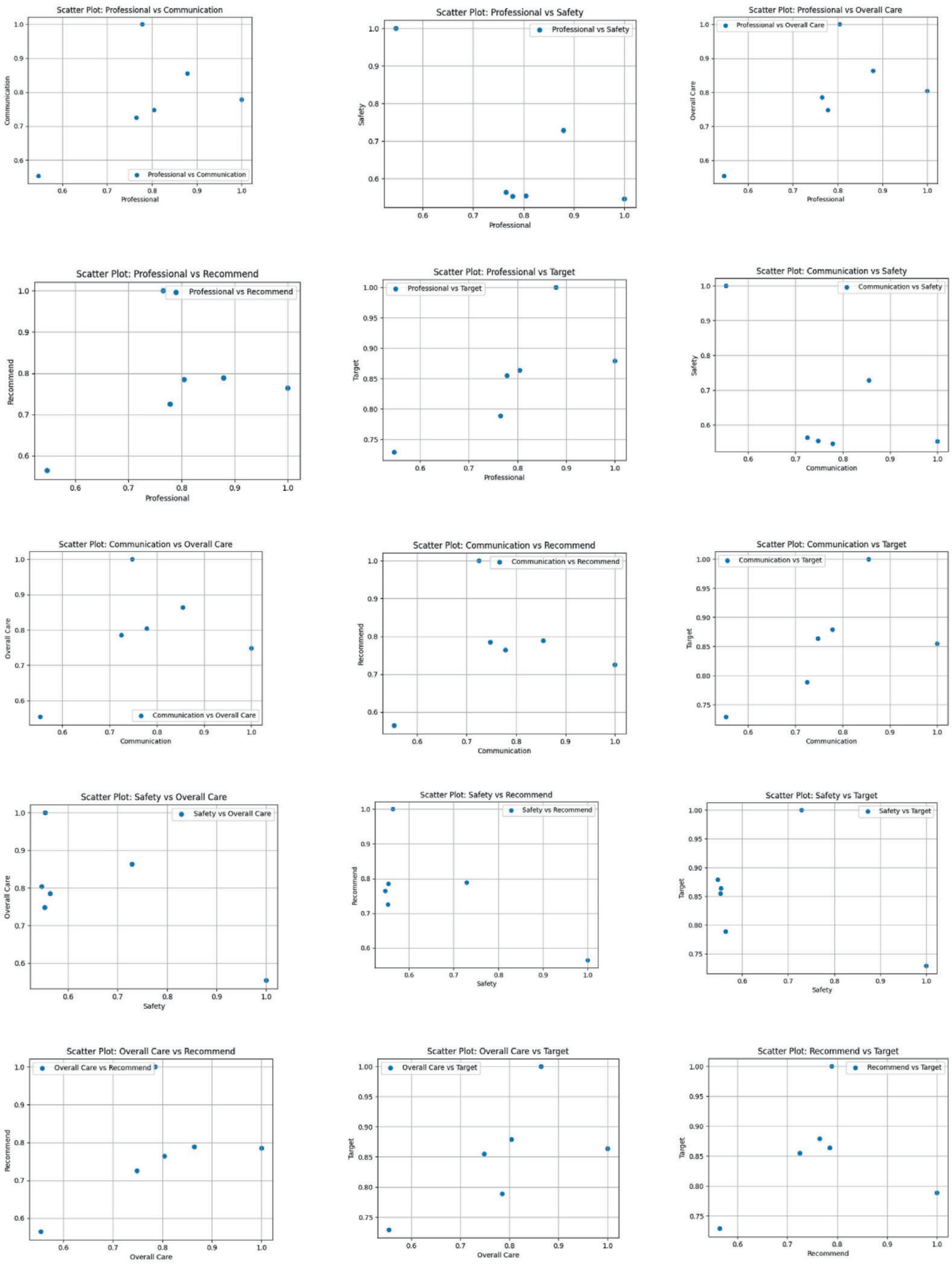


Figure 7. Scatter plot for the opiniion of users with sentimental analysis

professional, communicative, and caring. It is stated that Table 5 provides a comprehensive view of the sentiment relationships within the CSERS system, highlighting areas of strength and potential improvement in user sentiment and perception.

In Figure 8 and Table 6 provides an in-depth overview of the attack classification performance of the Content Service Ensemble Recommender System (CSERS) across different dataset sizes. This table offers crucial insights into CSERS' ability to accurately identify and classify security attacks, presenting key metrics such as true positives, true negatives, false positives, false negatives, and accuracy percentages. The data reveals several noteworthy trends. Firstly, as the dataset size increases, CSERS consistently maintains a high level of accuracy in attack classification, with accuracy percentages ranging from 96.4% to 97.8%. This demonstrates the system's robustness in handling larger datasets without significantly compromising its ability to detect attacks accurately. Furthermore, CSERS exhibits excellent performance in terms of true positives and true negatives, indicating its effectiveness in correctly identifying both attack instances and non-attack instances. Additionally, the low numbers of false positives and false

Table 6. Attack classification with CSERS

Dataset Size (MB)	True Positives	True Negatives	False Positives	False Negatives	Accuracy (%)
10	345	1987	23	45	96.4
50	1562	6624	112	45	97.8
100	2987	13245	245	63	97.6
500	14658	65432	1235	245	96.9
1000	28975	129837	2675	678	96.8
3000	86932	389562	8765	2345	96.7

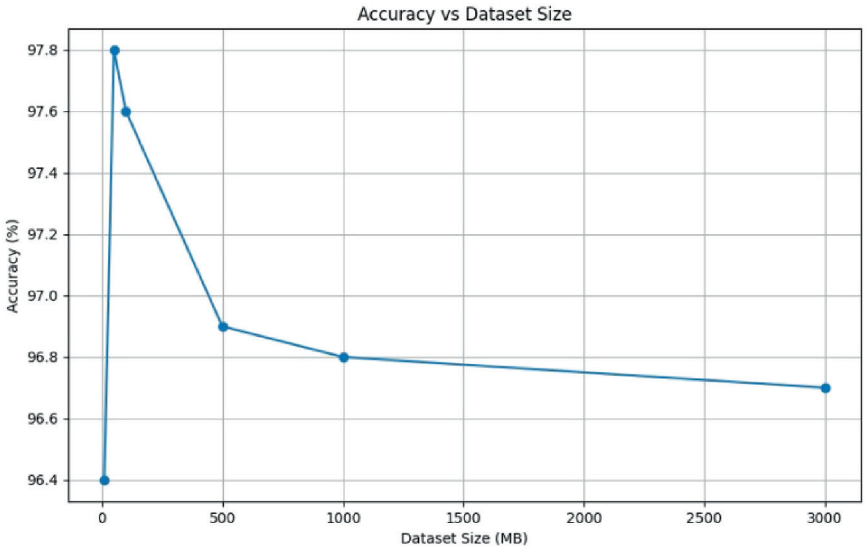


Figure 8. Estimation of accuracy

Table 7. Comparative analysis

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	ROC-AUC (%)
CSERS	96.4	89.2	92.5	90.7	98.2
ANN	94.8	88.1	91.2	89.6	97.9
SVM	93.2	86.4	89.8	87.9	96.8
RF	95.7	87.8	91.9	89.7	97.5

negatives highlight CSERS' precision in minimizing both type I and type II errors in attack classification. The Table 6 underscores CSERS' reliability and effectiveness in attack classification, even when confronted with varying dataset sizes. These results are indicative of the system's ability to provide a secure and trustworthy environment for content recommendation and user interactions.

In Figure 9 (a) – Figure 9 (d) and Table 7 presents a comparative analysis of the Content Service Ensemble Recommender System (CSERS) with other machine learning models, namely Artificial Neural Networks (ANN), Support Vector Machines (SVM), and Random Forest (RF). The table evaluates these models based on several key performance metrics, including accuracy, precision, recall, F1-Score, and ROC-AUC. CSERS demonstrates strong performance across all metrics, with an accuracy of 96.4%, indicating its ability to correctly classify both positive and negative instances. Its precision of 89.2% indicates a relatively low rate of false positives, highlighting its capability to make accurate positive predictions. CSERS also achieves a high recall of 92.5%, meaning it effectively identifies a significant proportion of actual positive instances. The F1-Score of 90.7%

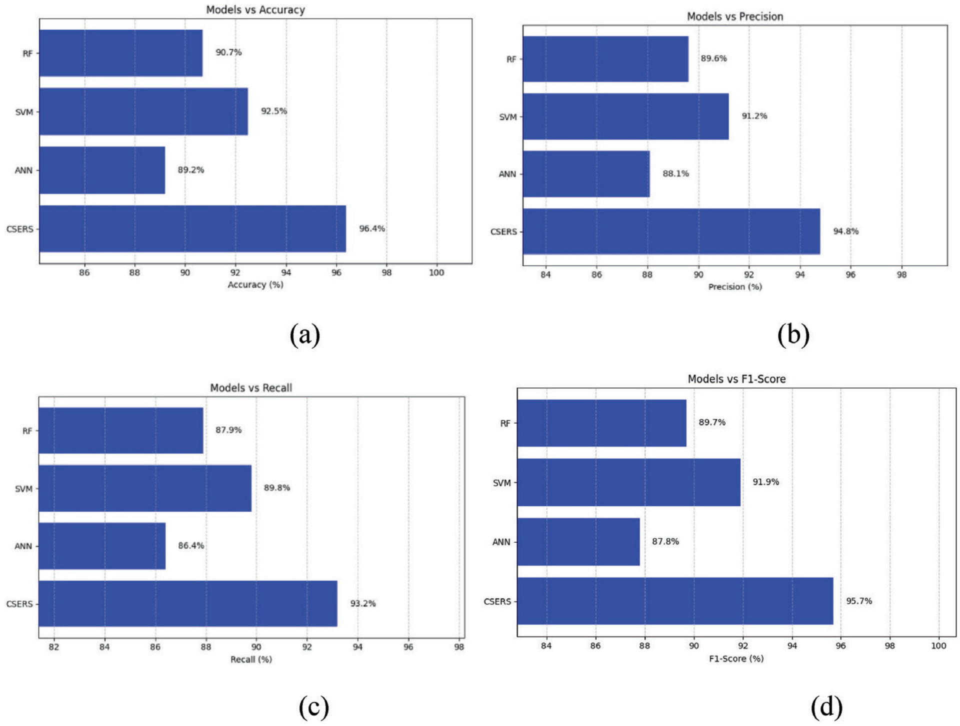


Figure 9. Comparative analysis of (a) accuracy (b) precision (c) recall (d) F1-score

suggests a balanced trade-off between precision and recall. Furthermore, CSERS excels in terms of ROC-AUC at 98.2%, indicating its strong ability to distinguish between positive and negative cases, making it highly effective in identifying security attacks.Comparatively, while the other models also demonstrate strong performance, CSERS consistently outperforms them in most metrics, showcasing its suitability for attack classification tasks. These results affirm CSERS as a robust and reliable system for ensuring security and trustworthiness in content recommendation environments.

The findings from the various tables and analyses in this study highlight several key aspects of the Content Service Ensemble Recommender System (CSERS):

1. CSERS demonstrates scalability in handling different dataset sizes and transaction volumes. As the dataset size increases or the number of transactions grows, CSERS maintains relatively efficient processing times, making it suitable for applications involving varying data loads.
2. CSERS adapts its security level estimations based on file size, showing a proactive approach to security assessment. It tends to be more cautious when dealing with larger files, which is essential for maintaining data integrity and user trust.
3. CSERS exhibits strong correlations between various sentiment aspects, such as professionalism, communication, and overall care. These correlations align with the system's target sentiment, indicating its effectiveness in meeting desired sentiment outcomes.
4. CSERS excels in accurately classifying security attacks across different dataset sizes. It maintains high accuracy, precision, and recall percentages, demonstrating its reliability in identifying both attack and non-attack instances.
5. In a comparative analysis with other machine learning models (ANN, SVM, and RF), CSERS consistently outperforms them in terms of accuracy, precision, recall, F1-Score, and ROC-AUC. This indicates its robustness and suitability for attack classification tasks.

The findings suggest that CSERS is a versatile and effective system for content recommendation tasks, with the ability to scale with data loads, assess security levels, perform sentiment analysis, and excel in security-related tasks compared to alternative models. These findings collectively underscore the system's potential for enhancing user experiences while ensuring security and trustworthiness.

6. CONCLUSION

Healthcare applications, recommender systems can assist healthcare providers in making treatment recommendations based on patients' medical histories and preferences. They can also help patients manage their health by offering personalized advice and resources. Recommender systems leverage big data analytics to provide personalized recommendations and insights in the healthcare domain. They contribute to better patient outcomes, cost-effective healthcare delivery, and improved overall healthcare quality. This paper proposed CSERS context of healthcare applications and security assessment. The research findings and analyses have shed light on various aspects of CSERS, highlighting its strengths and capabilities. CSERS has demonstrated scalability in handling different dataset sizes and transaction volumes, ensuring that it can efficiently process data loads of varying magnitudes. Its proactive approach to security estimation based on file size contributes to data integrity and user trust, a critical aspect in content recommendation systems. Moreover, CSERS exhibits strong correlations between sentiment aspects, suggesting its effectiveness in aligning with target sentiment outcomes. This capability enhances user experiences by tailoring recommendations to meet users' preferences and sentiments. In the security, CSERS stands out for its robust performance in attack classification. It consistently achieves high accuracy, precision, recall, and ROC-AUC percentages, indicating its reliability in identifying and classifying security threats accurately. Comparative analysis with other machine learning models further reinforces CSERS' superiority in security-related tasks. This paper underscores the significance of CSERS as a versatile and powerful system for content recommendation in healthcare applications. It not only enhances user experiences by understanding sentiment but also ensures security and trustworthiness through efficient security assessment and attack classification. The research findings provide valuable insights into CSERS' potential for real-world applications, where both user satisfaction and data security are paramount.

7. REFERENCES

1. REHMAN, A., NAZ, S., and RAZZAK, I. (2022). Leveraging big data analytics in healthcare enhancement: Trends, challenges and opportunities. *Multimedia Systems*, *28*(4), 1339–1371.
2. GUO, C., and CHEN, J. (2023). Big data analytics in healthcare. In *Knowledge Technology and Systems: Toward Establishing Knowledge Systems Science* (pp. 27–70). Singapore: Springer Nature Singapore.
3. BATKO, K., and ŚLĘZAK, A. (2022). The use of big data analytics in healthcare. *Journal of Big Data*, *9*(1), 3.
4. AHMAD, H. F., RAFIQUE, W., RASOOL, R. U., ALHUMAM, A., ANWAR, Z., and QADIR, J. (2023). Leveraging 6G, extended reality, and IoT big data analytics for healthcare: A review. *Computer Science Review*, *48*, 100558.
5. GOMES, M. A. S., KOVALESKI, J. L., PAGANI, R. N., DA SILVA, V. L., and PASQUINI, T. C. D. S. (2023). Transforming healthcare with big data analytics: Technologies, techniques and prospects. *Journal of Medical Engineering and Technology*, *47*(1), 1–11.
6. JAYASRI, N. P., and ARUNA, R. (2022). Big data analytics in health care by data mining and classification techniques. *ICT Express*, *8*(2), 250–257.
7. SHARMA, A., MALVIYA, R., and GUPTA, R. (2022). Big data analytics in healthcare. *Cognitive Intelligence and Big Data in Healthcare*, 257–301.
8. SOLFA, F. D. G., and SIMONATO, F. R. (2023). Big data analytics in healthcare: Exploring the role of machine learning in predicting patient outcomes and improving healthcare delivery. *International Journal of Computations, Information and Manufacturing* (IJCIM), *3*(1), 1–9.
9. GANIE, S. M., MALIK, M. B., and ARIF, T. (2022). Machine learning techniques for big data analytics in healthcare: Current scenario and future prospects. *In Telemedicine: The Computer Transformation of Healthcare* (pp. 103–123). Cham: Springer International Publishing.
10. COZZOLI, N., SALVATORE, F. P., FACCILONGO, N., and MILONE, M. (2022). How can big data analytics be used for healthcare organization management? Literary framework and future research from a systematic review. *BMC Health Services Research*, *22*(1), 1–14.
11. BISWAS, R. (2022). Outlining big data analytics in health sector with special reference to Covid-19. *Wireless Personal Communications*, *124*(3), 2097–2108.
12. ANDRY, J. F., REMBULAN, G. D., SALIM, E. L., FATMAWATI, E., and TANNADY, H. (2023). Big data analytics in healthcare: COVID-19 indonesia clustering. *Journal of Population Therapeutics and Clinical Pharmacology*, *30*(4), 290–300.
13. VELPULA, P., and PAMULA, R. (2022). CEECP: CT-based enhanced e-clinical pathways in terms of processing time to enable big data analytics in healthcare along with cloud computing. *Computers and Industrial Engineering*, *168*, 108037.
14. PRAMANIK, P. K. D., PAL, S., and MUKHOPADHYAY, M. (2022). Healthcare big data: A comprehensive overview. *Research Anthology on Big Data Analytics, Architectures, and Applications*, 119–147.
15. ALZAABI, H. M., ALAWADHI, M. A., and AHMAD, S. Z. (2023). Examining the impact of cultural values on the adoption of big data analytics in healthcare organizations. *Digital Policy, Regulation and Governance*, *25*(5), 460–479.

16. KORNELIA, B., and ANDRZEJ, Ś. (2022). The use of big data analytics in healthcare. *Journal of Big Data*, *9*(1).
17. CHOUDHARY, A., CHOUDHARY, A., and SUMAN, S. (2022). NLP applications for big data analytics within healthcare. *In Augmented Intelligence in Healthcare: A Pragmatic and Integrated Analysis* (pp. 237–257). Singapore: Springer Nature Singapore.
18. Gaurav, A. K. S. H. A. T., and Chui, K. T. (2022). Advancement of cloud computing and big data analytics in healthcare sector security. *Data Science Insights Magazine, Insights2Techinfo*, *1*, 12–15.
19. HASSAN, M., AWAN, F. M., NAZ, A., DEANDRÉS-GALIANA, E. J., ALVAREZ, O., CERNEA, A., ... and KLOCZKOWSKI, A. (2022). Innovations in genomics and big data analytics for personalized medicine and health care: A review. *International Journal of Molecular Sciences*, *23*(9), 4645.
20. Al-SAI, Z. A., HUSIN, M. H., SYED-MOHAMAD, S. M., ABDIN, R. M. D. S., DAMER, N., ABUALIGAH, L., and GANDOMI, A. H. (2022). Explore big data analytics applications and opportunities: A review. *Big Data and Cognitive Computing*, *6*(4), 157.
21. KARATAS, M., ERISKIN, L., DEVECI, M., PAMUCAR, D., and GARG, H. (2022). Big data for healthcare industry 4.0: Applications, challenges and future perspectives. *Expert Systems with Applications*, *200*, 116912.
22. PARIMANAM, K., LAKSHMANAN, L., and PALANISWAMY, T. (2022). Hybrid optimization based learning technique for multi-disease analytics from healthcare big data using optimal pre-processing, clustering and classifier. *Concurrency and Computation: Practice and Experience*, *34*(17), e6986.
23. LV, Z., GUO, J., and LV, H. (2023). Deep learning-empowered clinical big data analytics in healthcare digital twins. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 660–669
24. LV, Z., GUO, J., and LV, H. (2023). Deep learning-empowered clinical big data analytics in healthcare digital twins. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 660–669
25. MIAH, S. J., CAMILLERI, E., and VU, H. Q. (2022). Big data in healthcare research: A survey study. *Journal of Computer Information Systems*, *62*(3), 480–492.
26. PISE, A. A. (2022). An introduction to big data analytics techniques in healthcare. *Big Data Analytics and Machine Intelligence in Biomedical and Health Informatics: Concepts, Methodologies, Tools and Applications*, 1–20.
27. BERROS, N., EL MENDILI, F., FILALY, Y., and EL BOUZEKRI EL IDRISSI, Y. (2023). Enhancing digital health services with big data analytics. *Big Data and Cognitive Computing*, *7*(2), 64.
28. El KHATIB, M., HAMIDI, S., AL AMEERI, I., AL ZAABI, H., and Al MARQAB, R. (2022). Digital disruption and big data in healthcare-opportunities and challenges. *ClinicoEconomics and Outcomes Research*, 563–574.
29. ESCHENBRENNER, B., and BRENDEN, R. (2022). Deriving value from big data analytics in healthcare: A value-focused thinking approach. *AIS Transactions on Human-Computer Interaction*, *14*(3), 289–313.
30. MANOGARAN, G., THOTA, C., and LOPEZ, D. (2022). Human-computer interaction with big data analytics. *In Research Anthology on Big Data Analytics, Architectures, and Applications* (pp. 1578–1596). IGI global.
31. KHAN, S., KHAN, H. U., and NAZIR, S. (2022). Systematic analysis of healthcare big data analytics for efficient care and disease diagnosing. *Scientific Reports*, *12*(1), 22377.
32. SALAZAR-REYNA, R., GONZALEZ-ALEU, F., GRANDA-GUTIERREZ, E. M., DIAZ-RAMIREZ, J., GARZA-REYES, J. A., and KUMAR, A. (2022). A systematic literature review of data science, data analytics and machine learning applied to healthcare engineering systems. *Management Decision*, *60*(2), 300–319.
33. SINGH, R. K., AGRAWAL, S., SAHU, A., and KAZANCOGLU, Y. (2023). Strategic issues of big data analytics applications for managing health-care sector: A systematic literature review and future research agenda. *The TQM Journal*, *35*(1), 262–291.
34. WEERASINGHE, K., SCAHILL, S. L., PAULEEN, D. J., and TASKIN, N. (2022). Big data analytics for clinical decision-making: Understanding health sector perceptions of policy and practice. *Technological Forecasting and Social Change*, *174*, 121222.
35. AWRAHMAN, B. J., AZIZ FATAH, C., and HAMAAMIN, M. Y. (2022). A review of the role and challenges of big data in healthcare informatics and analytics. *Computational Intelligence and Neuroscience*, 2022.
36. IMAMALIEVA, D. (2022). Recent challenges of big data application in healthcare system. *In International Conference on Multidimensional Research and Innovative Technological Analyses* (pp. 121–124).
37. HUSSAIN, M., AJMAL, M., SUBRAMANIAN, G., KHAN, M., and ANAS, S. (2023). Challenges of big data analytics for sustainable supply chains in healthcare–a

resource-based view. *Benchmarking: An International Journal*.

38. GRANDA, W. X. B., MOLINA-GRANJA, F., ALTAMIRANO, J. D., LOPEZ, M. P., SURESHKUMAR, S., and SWAMINATHAN, J. N. (2022). Data analytics for healthcare institutions: A data warehouse model proposal. *In Inventive Communication and Computational Technologies: Proceedings of ICICCT 2022* (pp. 155–163). Singapore: Springer Nature Singapore.
39. MAHESH SELVI, T., and KAVITHA, V. (2022). A privacy-aware deep learning framework for health recommendation system on analysis of big data. *The Visual Computer*, *38*(2), 385–403.
40. SHAIKH, S. G., SURESH KUMAR, B., and NARANG, G. (2022). Recommender system for health care analysis using machine learning technique: A review. *Theoretical Issues in Ergonomics Science*, *23*(5), 613–642.
41. KAZEM, R. I., and ABDULLAH, E. F. (2023). Landscape view of recommender system techniques based on sentiment analysis. *International Journal of Nonlinear Analysis and Applications*, *14*(1), 1539–1546.
42. BOKHARAEI NIA, M., AFSHAR KAZEMI, M., VALMOHAMMADI, C., and ABBASPOUR, G. (2023). Wearable IoT intelligent recommender framework for a smarter healthcare approach. *Library Hi Tech*, *41*(4), 1238–1261.
43. PURANDHAR, N., and AYYASAMY, S. (2022). An empirical analysis on big analytics for e-healthcare and agriculture. *In International Conference on Artificial Intelligence for Smart Community: AISC 2020, 17–18 December, Universiti Teknologi Petronas, Malaysia* (pp. 409–417). Singapore: Springer Nature Singapore.
44. ANGAMUTHU, S., and TROJOVSKÝ, P. (2023). Integrating multi-criteria decision-making with hybrid deep learning for sentiment analysis in recommender systems. *PeerJ Computer Science*, *9*, e1497.
45. HUSSAIN, S. A., and Al GHAWI, S. (2022). Sentiment analysis of real-time health care twitter data using hadoop ecosystem. *In International Conference on Hybrid Intelligent Systems* (pp. 453–463). Cham: Springer Nature Switzerland.
46. AHMED, I., AHMAD, M., CHEHRI, A., and JEON, G. (2023). A heterogeneous network embedded medicine recommendation system based on LSTM. *Future Generation Computer Systems*, *149*, 1–11.
47. NIKHIL, TOMAR, M. S., BHADOURIYA, Y. S., PANDEY, S., and CHAURASIYA, H. (2022). Big data analytics for classification in sentiment analysis. *In International Conference on Communication, Networks and Computing* (pp. 63–73). Cham: Springer Nature Switzerland.
48. SHANTHI, E., and RAMASAMY, L. K. (2023). Recommender system for predicting students' academic performance in association with cognitive state and affective state using sentiment analysis and association rule mining on the closed ended questionnaire. *In 2023 9th International Conference on Information Technology Trends* (ITT) (pp. 79–83). IEEE.
49. CHEN, X., and WANG, J. (2022). Accurate medical information recommendation system based on big data analysis. *International Journal of Industrial and Systems Engineering*, *41*(2), 237–253.
50. P. K.VENKATESWAR LAL. (2024). A multi-objective direction of arrival estimation technique minimizing energy consumption in wireless sensor network. *Journal of Computer Allied Intelligence*, *2*(4), 36–50.
51. SRINIVASA SAI ABHIJIT CHALLAPALLI. (2024). Optimizing dallas-fort worth bus transportation system using any logic. *Journal of Sensors, IoT and Health Sciences*, *2*(4), 40–55.
52. SRINIVASA SAI ABHIJIT CHALLAPALLI. (2024). Sentiment analysis of the twitter dataset for the prediction of sentiments. *Journal of Sensors, IoT and Health Sciences*, *2*(4), 1–15.