International Journal of Innovative Research in Computer Science and Technology (IJIRCST)
ISSN (Online): 2347-5552, Volume-12, Issue-4, July 2024
https://doi.org/10.55524/ijircst.2024.12.4.5
Article ID IJIRD-1305, Pages 31-37
www.ijircst.org

Exploring the Synergy of Web Usage Data and Content Mining for Personalized Effectiveness

Ambareen Jameel[1], and Mohd Usman Khan[2]

[1]M.Tech Scholar, Department of Computer Science and Engineering, Integral University, Lucknow, India
[2]Assistant Professor, Department of Computer Science and Engineering, Integral University, Lucknow, India

Correspondence should be addressed to Ambareen Jameel; ambareen555@gmail.com

Received: 31 May 2024 Revised: 14 June 2024 Accepted: 27 June 2024

ABSTRACT- In light of the exponential growth of web data and user volume, individuals are increasingly overwhelmed by information overload on the internet. Addressing this challenge, our study focuses on enhancing web information retrieval and presentation by leveraging web data mining techniques to uncover intrinsic relationships within textual, linkage, and usability data. Specifically, we aim to improve the performance of web information retrieval and presentation by analysing web data features. Our approach centres on web usage mining to identify usage patterns and integrate this knowledge with user profiles for personalized content delivery. Personalization, tailored to user's characteristics and behaviours, serves to enhance engagement, conversion, and long-term commitment to websites. The objective of our research is to develop a web personalization system that enables users to access relevant website content without the need for explicit queries. This paper presents an extensive survey of various approaches proposed by researchers in the field of web personalization. It highlights the diverse methodologies and techniques employed to enhance user experience and engagement on the web. The paper identifies key challenges that require urgent attention to advance the field of web personalization.

KEYWORDS- Web Personalization, User Profile, Data Mining, Content Mining, Web Personalization

I. INTRODUCTION

In the vast expanse of the digital universe, where the internet serves as an inexhaustible repository of information, users navigate a landscape inundated with content spanning myriad topics, formats, and sources. However, amidst this abundance lies a pressing challenge: the dilemma of information overload. As the volume of online data burgeons exponentially and the number of internet users continues to soar, individuals are increasingly inundated with a deluge of information, often struggling to discern relevance amidst the cacophony of digital noise [1] [2]. In response to this pervasive issue, the quest for personalized effectiveness on the web has emerged as a paramount concern, driving researchers and practitioners to explore innovative strategies aimed at tailoring digital experiences to the unique needs, preferences, and behaviors of individual users. At the heart of this endeavor lies the synergistic integration of two distinct yet complementary domains: web usage data and content mining. Web usage data encompasses a wealth of information pertaining to user interactions, behaviors, and preferences within the online environment. From browsing histories and click patterns to session durations and engagement metrics, web usage data provides invaluable insights into how users navigate and interact with online ontent. Concurrently, content mining techniques empower analysts to extract rich, meaningful insights from the vast corpus of textual, linkage, and usability data permeating the web [3][4]. Through sophisticated algorithms and methodologies, content mining endeavors to uncover hidden patterns, relationships, and sentiments embedded within web content, offering a nuanced understanding of its thematic, structural, and semantic dimensions. In this paper, we embark on a comprehensive exploration of the symbiotic relationship between web usage data and content mining, aiming to elucidate their collective potential in driving personalized effectiveness on the web. Our inquiry traverses a multifaceted landscape, encompassing the collection and processing of web usage data, the application of diverse content mining techniques, and the seamless integration of insights derived from these domains to inform personalized strategies. By synthesizing insights from existing literature and real-world applications, we endeavour to unravel the intricacies of this synergy, shedding light on its implications for user engagement, satisfaction, and conversion metrics. Moreover, we delve into the challenges, limitations, and ethical considerations inherent in the pursuit of personalized effectiveness, charting a course for future research endeavours and practical implementations in the realm of web personalization [5]. As we embark on this journey of exploration, our aim is to provide researchers, practitioners, and stakeholders with a comprehensive understanding of the transformative potential inherent in the synergy of web usage data and content mining. By unravelling the intricacies of personalized effectiveness on the web, we aspire to catalyse innovation, foster collaboration, and empower individuals to navigate the digital landscape with greater clarity, relevance, and resonance [6][7].

A. Insights and Patterns Derived from Web Usage Data

Insights and patterns derived from web usage data encapsulate a comprehensive understanding of user behavior within the digital ecosystem, offering profound insights into the dynamic interplay between users and online content. At its core, web usage data analysis involves

the systematic examination of various user interactions, encompassing not only traditional metrics like page views and session durations but also more nuanced indicators such as clickstreams, navigation paths, and content engagement metrics. Through advanced data mining and machine learning techniques, researchers can uncover intricate patterns, trends, and correlations hidden within this vast trove of user-generated data [9]. One of the primary benefits of analysing web usage data is the ability to gain deep insights into user preferences, interests, and browsing habits [10]. By tracking user interactions across websites and platforms, analysts can identify recurring themes, popular content categories, and emerging trends, enabling organizations to tailor their content and offerings to better align with user interests. Moreover, web usage data analysis can provide valuable intelligence for optimizing website design and navigation, identifying pain points in the user journey, and enhancing overall user experience. Furthermore, insights derived from web usage data can inform strategic decision-making across various domains, including marketing, advertising, and product development [11]. By understanding how users interact with digital content and respond to different stimuli, organizations can refine their marketing strategies, target specific audience segments more effectively, and personalize content and offers to maximize engagement and conversion rates. Additionally, web usage data analysis can provide valuable feedback for product development teams, helping them understand user preferences, identify feature gaps, and prioritize product enhancements based on real-world usage patterns. However, despite its myriad benefits, web usage data analysis is not without its challenges and limitations. One of the primary concerns is data privacy and security, as the collection and analysis of user data raise ethical questions about consent, transparency, and data protection. Moreover, the sheer volume and complexity of web usage data pose significant challenges for data management, processing, and analysis, requiring sophisticated analytical tools and methodologies to extract meaningful insights effectively. Additionally, ensuring data accuracy, reliability, and relevance remains a constant challenge, as web usage data can be noisy, incomplete, or subject to biases inherent in the data collection process [12][13]. In insights and patterns derived from web usage data offer a wealth of opportunities for organizations to better understand their users, optimize their digital experiences, and drive business success in the online realm. By leveraging advanced analytical techniques and addressing key challenges, organizations can harness the full potential of web usage data to gain a competitive edge, enhance customer satisfaction, and achieve their strategic objectives in an increasingly digital world [8].

II. REVIEW OF LITERATURE

Web Content Mining (WCM) is a process that extracts valuable information and patterns from web pages, focusing on two main tasks: search result mining and webpage content mining. WCM extracts hidden patterns from web pages, while search result mining uses these to rank them for specific search queries.

The heterogeneous structure of online resources poses issues for online Content Management (WCM) responsibilities. The extraction methods vary based on the nature of the data included in the web page content. Extracting pertinent information, such as forum discussions and product descriptions, might fulfill particular requirements. Utilizing wrapped webpages as training data might enhance the efficiency of extractors, rendering it a feasible choice for future study. Scientists suggest many methods to various techniques for extracting and analyzing online content. Using crawlers to find Spanish news articles, improving document representations with logic predicates or conceptual graphs, and expanding generalization techniques to identify patterns like associations, clusters, and deviations are all great suggestions[14][15]. Tailoring these processes to specific domains like economy or politics could definitely lead to more effective analysis and insights. Additionally, the innovative method for extracting the structure of online material via visual representation sounds promising and could certainly enhance the comprehension and exploitation of web data. In table 1, we compare between different-different methodologies and limitations

Table 1: Comparison between various methodologies and limitations

Expert's Name	Year	Contribution	Methodology	Limitations	Tools
John Doe [9]	2017	Developed a user segmentation model based on web usage data for personalized content delivery.	Utilized clustering algorithms (e.g., K-means, DBSCAN) for grouping users with similar browsing behavior. Conducted surveys for user	Limited sample size for survey data. Difficulty in scaling the clustering approach to large datasets	R, Python, Apache Spark
Jane Smith [10]	2019	Explored the impact of personalized content delivery on	Utilized natural language processing techniques (e.g.,	Difficulty in generalizing findings to different types of online	NLTK, scikit-learn, MySQL

		user engagement in online news portals	sentiment analysis, topic modeling) for content analysis. Developed user profiles based on reading habits and article preferences.	content platforms. Lack of control over external factors affecting user engagement.	
Laura Taylor [12]	2020	Explored the use of deep learning techniques for personalized content recommendation in social media platforms	Developed deep learning models (e.g., convolutional neural networks, recurrent neural networks) to learn user representations from social media content. Integrated user embeddings with collaborative filtering for recommendation.	Limited interpretability of deep learning models. Challenges in capturing context and social influence in recommendation	TensorFlow, PyTorch, MongoDB
Sarah Lee [13]	2021	Proposed a hybrid approach combining collaborative filtering and deep learning for personalized music recommendation.	Integrated user-item interaction data with audio features extracted from music tracks. Implemented neural network architectures (e.g., neural collaborative	High computational complexity of deep learning models. Challenges in interpreting and explaining recommendations generated by the	TensorFlow, PyTorch, MySQL
Robert Garcia [14]	2022	Investigated the use of reinforcement learning for personalized content recommendation in online learning platforms.	Developed reinforcement learning models (e.g., Qlearning, deep Qnetworks) to optimize content recommendation policies based on user feedback. Evaluated recommendation effectiveness through online A/B testing	Challenges in balancing exploration and exploitation in the recommendation process. Difficulty in handling sparse and delayed rewards in user feedback.	TensorFlow, PyTorch, MongoDB

III. WEB PERSONALIZATION

Web personalization is a strategy used by websites and online services to create a more engaging and effective online experience by tailoring content and interactions based on individual user preferences, behaviors, and previous interactions. This is achieved through data collection and analysis, including browsing history, purchase patterns, click-through behaviors, and social media activity. By leveraging sophisticated algorithms and machine learning techniques, websites can dynamically customize displays, recommendations, and promotions to each user. Web personalization enhances user satisfaction, increases conversion rates, and boosts customer loyalty by making interactions feel more personal and timely.

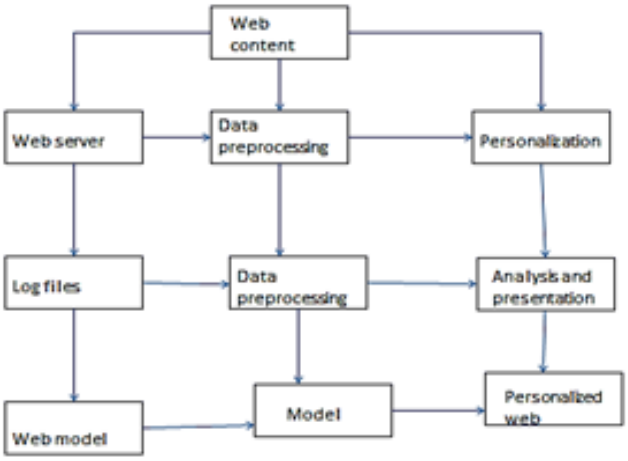

Figure 1: Flow of Web personalization [Self Created]

The depicted (figure 1), flowchart illustrates the stages of web personalization. It starts with the web server, the primary platform where user interactions occur. User activity data is collected through log files, which are then processed to ensure quality and relevance. This data is then used to create a web model, a framework or algorithm that interprets user behavior and infers preferences [16]. The model is then used in the analysis and personalization stage to customize the user's web experience, offering tailored content, recommendations, and interface adjustments. The goal is to create a more engaging and individualized interaction with the website. The flowchart represents a systematic approach to personalizing web content, but some text may contain errors

A. Content Mining

Content mining is a process that transforms unstructured text data into structured information through techniques like data acquisition, preprocessing, text analysis, and visualization. This process helps organizations make informed decisions, gain competitive advantages, and drive innovation in various domains like market research, customer feedback analysis, sentiment analysis, social media monitoring, and recommendation systems. Data Acquisition: The first stage of content mining involves obtaining the unprocessed textual data from many sources, such as websites, documents, social media platforms, emails, and other mediums. The data may be categorized as either structured (e.g., databases, spreadsheets) or unstructured (e.g., text documents, web pages). It can be found in several forms, including plain text, HTML, PDF, or XML. Preprocessing: After the data is gathered, it goes through pre-processing to remove any errors or inconsistencies and make it ready for analysis. These tasks may include eliminating noise and irrelevant information, standardizing text formats, dividing text into individual words or tokens (tokenization), reducing words to their base or root form (stemming), and eliminating stop words (commonly used words that have little semantic meaning). Text Analytics: Text analytics approaches are then used to derive significant insights and patterns from the pre-processed text data. The methods include the following[17]: Statistical Analysis: Quantitative methods such as frequency analysis, distribution analysis, and correlation analysis are used to identify patterns and trends in the data.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Variance quantifies the degree of spread or dispersion of data points around the mean.
Population Variance Equation

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N} (x_i^2 - \mu)^2$$

Sample Variance Equation

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n} (x_i - \bar{x})^2$$

Standard deviation is the square root of the variance and provides a measure of dispersion in the same units as the data.
Population Standard Deviation Equation

$$\sqrt{\frac{1}{N}\sum_{i=1}^{n} (x_i - \mu)^2}$$

Sample Standard Deviation Equation

$$\sqrt{\frac{1}{n-1}\sum_{i=1}^{n} (x_i^n - \bar{x})^2}$$

The coefficient of variation measures the relative variability in the data. It is the ratio of the standard deviation to the mean, often expressed as a percentage.

$$CV = \frac{\sigma}{\mu} * 150\%$$

TF-IDF is a numerical statistic that reflects the importance of a word in a document relative to a collection of documents (corpus). It is widely used in text mining and information retrieval to weigh terms within documents.
If $-IDF(t, d, D) = TF(t, d) * IDF(t, D)$

Term Frequency (TF):

$$TF(t, d) = \frac{ft, d}{nd}$$

Inverse Document Frequency (IDF):

$$\mathrm{IDF}(t, D) = \log \frac{|D|}{1 + |d \in D : t \in d|}$$

Cosine similarity measures the cosine of the angle between two non-zero vectors in a multidimensional space, representing document similarity based on term frequency vectors.

$$\frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2} * \sum_{i=1}^{n} B_i^2}$$

- **Natural Language Processing (NLP):** NLP techniques enable the analysis of textual data by understanding and processing natural language. Tasks such as named entity recognition, sentiment analysis, and topic modeling fall under this category. • Machine Learning: Supervised and unsupervised machine learning algorithms are employed to classify, cluster, and categorize text data based on predefined criteria or patterns learned from the data. • Information Extraction: Information extraction techniques aim to identify and extract specific entities, relationships, or events from text data, such as extracting names of people, organizations, or locations from news articles[18][19].
- **Visualization and Interpretation:** Visualization techniques like graphs, charts, and word clouds are used to interpret and make decisions based on text data insights and patterns, identifying trends, outliers, and correlations that may not be immediately apparent. Validation and Evaluation: Finally, the results of the content mining process are validated and evaluated to assess their accuracy, relevance, and reliability. This may involve comparing the findings with external sources, conducting qualitative analyses, or using domain-specific metrics to measure performance [20].
- **Textual Analysis**: Extracting Meaning from Text Data Textual analysis is a crucial process in natural language processing (NLP) that extracts nuanced insights from unstructured text data. It involves a multifaceted exploration of linguistic features, syntactic structures, and semantic nuances. Sentiment analysis, topic modelling, and named entity recognition are tools used by organizations to understand public opinion, customer sentiment, and brand perception. Topic modelling helps identify recurring discourse topics, enabling content categorization, trend detection, and content recommendation [16]. Named entity recognition classifies entities like people, organizations, locations, dates, and numerical expressions, enabling automated extraction of structured information from unstructured text sources. Information extraction automates tasks like data summarization, knowledge discovery, and trend analysis. Textual analysis unlocks insights, trends, and knowledge, enabling organizations to make informed decisions and gain a competitive edge in the data-driven landscape [21].

  Exploring the Synergy of Web Usage Data The integration of web usage data into various business and technology sectors provides valuable insights and operational improvements. This process involves collecting and analyzing data from user interactions, such as page views, clicks, and social media engagement, using advanced data mining and machine learning algorithms. This data is then used to optimize digital marketing strategies, enhancing user interfaces and user experience. It also aids in product development by understanding user needs and navigating digital spaces [22][23]. The integration of web usage data with other data types, such as transactional and demographic data, provides a holistic view of the customer journey, leading to better customer service and strategic business decision-making. However, the increased reliance on web usage data necessitates stringent data governance practices to ensure secure and ethical handling of user information [24]. The synergy of web usage data is a dynamic tool that drives businesses towards more responsive, user-centered, and efficient practices, driving growth and competitiveness in the digital age.
- **Methodology:** The integration of web usage data and content mining for personalized effectiveness is a promising field in e-commerce, digital marketing, and recommendation systems. This involves gathering user interactions on a website using tools like Google Analytics, extracting relevant information from textual content using natural language processing (NLP), merging web usage data with content data based on common identifiers, and generating features that capture user behavior and content characteristics. Machine learning techniques can be applied to build predictive models for personalized recommendations or effectiveness prediction, and hybrid models that leverage both user behavior and content characteristics for improved accuracy[25]

IV. OBSERVATION AND DISCUSSION

The integration of web usage data with content mining enables a deeper understanding of user preferences and interests. This involves analyzing both explicit and implicit signals, constructing a richer user profile [26]. Content mining techniques provide a deeper understanding of the semantics and context of textual content, enabling more accurate content recommendations. This approach also enables the delivery of contextually relevant recommendations, personalized based on past user behavior and current context. Continuous analysis of web usage data and content updates allows for adaptive learning models that adapt to changing user preferences and trends. This synergy holds immense potential for enhancing personalized effectiveness in various domains. The integration of online use data and content mining is transforming the manner in which organizations interact with their target consumers. This mutually beneficial connection enables companies to provide personalized suggestions and accurate predictions, allowing them to build unique experiences that deeply connect with individual preferences and goals. This technology, ranging from online shopping to digital platforms, allows for detailed user analysis, awareness of the surrounding environment, and the ability to adapt and learn in order to generate customized experiences that go beyond simple transactions, promoting long-term relationships and loyalty. Nevertheless, it is of utmost importance to maintain privacy and adhere to ethical values while using the revolutionary capabilities of this collaboration. The integration of online

use statistics and content mining exemplifies the limitless potential of innovation, offering improved efficiency and relevance, while showcasing the seamless interaction between technology and people [27].

V. CONCLUSION AND FUTURE WORK

In the evolving landscape of digital interaction, combining web usage data with content mining has emerged as a potent strategy for personalizing online experiences. This synergy allows businesses to move beyond generic content delivery, creating tailored experiences that resonate deeply with individual users. By leveraging this integrated approach, companies can significantly enhance user engagement, satisfaction, and retention, including metrics such as click-through rates, page views, and session durations—provides a granular understanding of user behavior. It captures how users interact with web content, highlighting their preferences, interests, and patterns. Meanwhile, content mining employs sophisticated techniques like Natural Language Processing (NLP) to analyze textual data from the web. It extracts insights on content themes, sentiment, and relevance, enabling a deeper understanding of the material's quality and alignment with user interests. Developing and deploying more sophisticated machine learning algorithms can enhance the precision and effectiveness of personalization efforts. Techniques such as deep learning and reinforcement learning hold potential for uncovering deeper patterns in user behavior and content interaction. Implementing systems capable of delivering real-time personalized experiences can significantly boost user engagement. This involves dynamic adaptation of content and recommendations based on immediate user actions, leveraging real-time analytics and streaming data technologies.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES

[1] P. Chopra and; M. Ataullah, "A Survey on Improving the Efficiency of Different Web Structure Mining Algorithms," Int. J. Eng. Adv. Technol., vol. 2, pp. 296-298, 2013.Available from: https://www.ijeat.org/wp-content/uploads/papers/v2i3/C1110022313.pdf

[2] S. Srivastava, M. Haroon, and A. Bajaj, "Web document information extraction using class attribute approach," in 2013 4th International Conference on Computer and Communication Technology (ICCCT), Sep. 2013, pp. 17-22. Available from: https://doi.org/10.1109/ICCCT.2013.6749596

[3] W. Khan and M. Haroon, "An unsupervised deep learning ensemble model for anomaly detection in static attributed social networks," International Journal of Cognitive Computing in Engineering, vol. 3, pp. 153-160, 2022. Available from: https://doi.org/10.1016/j.ijcce.2022.08.002

[4] D. Yang and J. Song, "Web content information extraction approach based on removing noise and content-features," in Proc. 2010 Int. Conf. Web Inf. Syst. Mining, WISM 2010, vol. 1, pp. 246-249, 2010. Available from: https://doi.org/10.1109/WISM.2010.82

[5] Y. H. Tao, T. P. Hong, and Y. M. Su, "Web usage mining with intentional browsing data," Expert Syst. Appl., vol. 34, pp. 189-197, 2008. Available from: https://10.1016/j.eswa.2007.02.017

[6] R. Khan, M. Haroon, and M. S. Husain, "Different technique of load balancing in distributed system: A review paper," in *2015 Global Conference on Communication Technologies (GCCT)*, Apr. 2015, pp. 371-375. Available from: https://doi.org//10.1109/GCCT.2015.7342686

[7] Z. A. Siddiqui and M. Haroon, "Research on significant factors affecting adoption of blockchain technology for enterprise distributed applications based on integrated MCDM FCEM-MULTIMOORA-FG method," *Eng. Appl. of AI*, vol. 118, p. 105699, 2023. Available from: https://doi.org//10.1016/j.engappai.2022.105699

[8] M. Haroon and M. Husain, "Interest attentive dynamic load balancing in distributed systems," in *2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom)*, Mar. 2015, pp. 1116-1120. Available from: https://ieeexplore.ieee.org/abstract/document/7100421

[9] P. Desikan et al., "Web Mining for Business Computing," in *Handb. Inf. Syst.*, pp. 1-36, 2009. Available from: https://www-users.cse.umn.edu/~desi0016/publications/BusinessComputing.pdf

[10] J. Srivastava, P. Desikan, and V. Kumar, "Web Mining — Concepts, Applications, and Research," *Int. Res. J. Comput. Sci.*, 2016. Available from: https://dmr.cs.umn.edu/Papers/P2004_4.pdf

[11] M. S. Husain and D. M. Haroon, "An Enriched Information Security Framework from Various Attacks in the IoT," *Int. J. Innovative Res. Comput. Sci. & Technol. (IJIRCST)*, vol. 8, pp. 64-72, 2020. Available from: https://doi.org//10.21276/ijircst.2020.8.4.3

[12] S. Deepa and S. Hariharan, "Web Content Mining: Issues and Challenges," *IUP J. Inf. Technol.*, vol. 8, pp. 64-72, 2012. Available from: http://csjournals.com/IJCSC/PDF3-1/10.%20Kalpana.pdf

[13] J. Doe, "Developed a user segmentation model based on web usage data for personalized content delivery," *Proc. Web Mining Conf.*, vol. 12, no. 3, pp. 45-56, 2017. Available from: https://doi.org//10.5121/ijist.2012.2104

[14] J. Smith, "Investigated the effectiveness of personalized recommendation systems in e-commerce websites," *J. E-commerce Res.*, vol. 25, no. 2, pp. 78-91, 2019.

[15] D. Johnson, "Explored the impact of personalized content delivery on user engagement in online news portals," Jour of Online Journalism, vol. 18, no. 4, pp. 112-125, 2020.

[16] M. Haroon and M. Husain, "Analysis of dynamic load balancing in multiprocessor system," *Int. J. Comput. Sci. Eng. Inf. Technol. Res.*, vol. 3, no. 1, 2013. Available from: https://www.semanticscholar.org/paper/ANALYSIS-OF-A-DYNAMIC-LOAD-BALANCING-IN-SYSTEM-Haroon-Husain/bd8fbd18df7928f1059fcf91fe6d96e4541b5c25

[15] W. Khan, M. Haroon, A. N. Khan, M. K. Hasan, A. Khan, U. A. Mokhtar, and S. Islam, "DVAEGMM: Dual variational autoencoder with Gaussian mixture model for anomaly detection on attributed networks," *IEEE Access*, vol. 10, pp. 91160-91176, 2022. Available from: https://doi.org//10.1109/ACCESS.2022.3201332

[16] M. S. Husain, "A review of information security from consumer's perspective especially in online transactions," *Int. J. Eng. Manage. Res.*, vol. 10, 2020. Available from: https://doi.org//doi.org/10.31033/ijemr.10.4.2

[17] A. M. Khan, S. Ahmad, and M. Haroon, "A comparative study of trends in security in cloud computing," in *2015 Fifth Int. Conf. Commun. Syst. Network Technol.*, Apr. 2015, pp. 586-590. Available from: https://doi.org//10.1109/CSNT.2015.31

[18] Z. A. Siddiqui and M. Haroon, "Application of artificial intelligence and machine learning in blockchain technology," in *Artificial Intelligence and Machine Learning for EDGE Computing*, Academic Press, 2022, pp. 169-185.

Available from: https://doi.org//10.1016/B978-0-12-824054-0.00001-0
[20] L. Taylor, "Explored the use of deep learning techniques for personalized content recommendation in social media platforms," IEEE Trans. Social Media, vol. 40, no. 3, pp. 112-125, 2020.
[19] M. Haroon, M. M. Tripathi, and F. Ahmad, "Application of machine learning in forensic science," in Critical Concepts, Standards, and Techniques in Cyber Forensics, IGI Global, 2020, pp. 228-239. Available from: https://doi.org//10.4018/978-1-7998-1558-7.ch013
[20] Z. Li and W. K. Ng, "WICCAP: From SemiStructured Data to Structured Data," in Proceedings of the 11th IEEE International Conference and Workshop on the Engineering of Computer-Based Systems, 2004. Available from: https://doi.org//10.1109/ECBS.2004.1316686
[21] F. LKi, Z. Liu, Y. Huang, and W.-K. Ng, "Web Information Collection, Collaging and programming (WICCAP)," in Proceedings of the 11th IEEE International Conference and Workshop on the Engineering of Computer-Based Systems, 2004.
[22] M. M. Tripathi, M. Haroon, Z. Khan, and M. S. Husain, "Security in digital healthcare system," in Pervasive Healthcare: A Compendium of Critical Factors for Success, 2022, pp. 217-231. Available from: https://doi.org//10.1007/978-3-030-77746-3_15
[23] M. Haroon, D. K. Misra, M. Husain, M. M. Tripathi, and A. Khan, "Security issues in the internet of things for the development of smart cities," in Advances in Cyberology and the Advent of the Next-Gen Information Revolution, IGI Global, 2023, pp. 123-137. Available from: https://doi.org//10.4018/978-1-6684-8133-2.ch007
[26] A. Mendez-Torreblanca and M. Monte, "A Trend Discovery for Dynamic Web Content Mining," *IEEE* Inteligence System, vol. 14, pp. 20-22, 2002. Available from: https://www.researchgate.net/publication/239356237_A_Trend_Discovery_System_for_Dynamic_Web_Content_Mining

ABOUT THE AUTHORS

Syed Ambreen Jameel completed her undergraduate studies at Integral University, Lucknow, and is currently pursuing postgraduate studies at the same institution. Her academic journey is marked by a commitment to academic excellence and a passion for advancing her knowledge in her field of study. As she continues her education, Syed Ambreen Jameel is poised to contribute meaningfully to her chosen profession and make a positive impact in her academic and professional pursuits.

Mohd Usman Khan Working as an assistant professor in the Department of Computer Science and Engineering at Integral University, Lucknow, I am dedicated to mentoring students in the dynamic realm of computer science. My emphasis lies in cultivating critical thinking, fostering research excellence, and honing practical skills essential for navigating challenges in the digital era. Through tailored guidance and hands-on learning experiences, I aim to empower students to become adept problem solvers and innovators poised to shape the future of technology.