e-ISSN:2582-7219

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

Volume 7, Issue 4, April 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.521

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 4, April 2024 |

| DOI:10.15680/IJMRSET.2024.0704009 |

Web User Behavior Analysis using CNN Algorithm

Dr.J.Jeyaboopathiraja[1], Shahil Gupta A[2]

[1]Assistant Professor, PG & Research Department of Computer Science, Sri Ramakrishna College of Arts & Science, Coimbatore, India

[2]UG Student, PG & Research Department of Computer Science, Sri Ramakrishna College of Arts & Science, Coimbatore, India

ABSTRACT: Understanding user behavior on websites or web applications is pivotal for businesses striving to establish a strong online presence and effectively engage with their target audience. In this paper, we propose a comprehensive analysis of web user behavior, leveraging a multifaceted approach that encompasses data collection, preprocessing, analysis, and interpretation. Through the utilization of advanced analytics techniques and machine learning algorithms, we aim to uncover valuable insights into user interactions, preferences, and engagement patterns. By analyzing various dimensions of user behavior such as session duration, page views, conversion rates, and social shares, we seek to identify trends, correlations, and anomalies that can inform strategic decision-making processes. Additionally, we explore the impact of contextual factors such as content categories, device types, and referral sources on user behavior to gain a holistic understanding of the user journey. The integration of data visualization techniques enables us to communicate complex findings effectively and facilitate actionable recommendations for website optimization, content personalization, and targeted marketing initiatives. Ultimately, our goal is to empower businesses with actionable insights derived from a deep understanding of web user behavior, thereby enabling them to enhance user experience, drive customer satisfaction, and achieve their business objectives in the digital landscape.

KEYWORDS: Cloud, storage, encryption, security, decryption

I. INTRODUCTION

In today's digital age, understanding user behavior on websites and web applications has become imperative for businesses aiming to stay competitive and meet the evolving needs of their customers. The internet serves as a vast platform for interaction, commerce, and information dissemination, with user behavior acting as a key driver of success in various online endeavors. Analyzing web user behavior provides invaluable insights into how users navigate websites, engage with content, and interact with features, enabling businesses to tailor their offerings and strategies accordingly. Traditional approaches to web user behavior analysis have often relied on simplistic metrics such as page views and click-through rates. While informative to some extent, these metrics fail to capture the intricacies of user interactions and preferences in a holistic manner. With the advent of advanced machine learning techniques, particularly deep learning, there exists an opportunity to delve deeper into understanding web user behavior through more sophisticated models.

Huge numbers of documents are available on the web, so this makes a great problem to find the one which is appropriate. Many search techniques were available, among that, full-text search is a famous one and it is given by most of the services like Google and it is helpful for gathering the documents which have been viewed previously and so we know the appropriate keyword to search. Though it is useful, we have some drawback in this approach also, since it is not feasible to identify the not-yet seen appropriate document for a particular topic. Following details explain why text-based search isn't suitable for searching the appropriate document. Ambiguity of natural communication: synonyms, homographs and modulation of words will undermine the procedures; it sees the exploration technique in the form of sequence of characters only. High-level, vague concepts: high-level, ambiguity is a determined abstract idea such as "Kosovo conflict", "Industrial Revolution" or the "Iraq War", these were frequently not declared certainly in the appropriate records; so existing search engines can't identify such details.

Ontologies create the fundamental infrastructure of the Semantic Web. As "ontology" we conceive any ceremonies with a well-established mathematical understanding, which is able to denote the sub-concept taxonomy, idea examples and user-defined connection among ideas. Those ceremonies permits many refined likeliness of the set of information than classical thesauri. They denote information on the semantic level, i.e., they have semantic individuals (ideas, relationships and examples) rather than effortless words, which avoids the revealed "noise" from the associations. Additionally, they permit defining the semantic connections among the entities, and also to amass the

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 4, April 2024 |

| DOI:10.15680/IJMRSET.2024.0704009 |

entrenched particulars and axioms regarding the information field. This extra turn of phrase authority permits the recognition of the strength circumstance of the particular relations. Ontologies will resolve these issues of full-text search theoretically. Ontologies work along with the keyword sourced search; this is one of the inspirations of the Semantic Web (SW). The semantic web aims to construct the expansion of the existing web where the details provided are with well-established context, allows a person to work together. The Information Retrieval in the Semantic web (IRIS) gives an inference engine to do the interpretation with regulations across a computer science prose field ontology to extract the absolute idea. At the time of interaction among the users in the system, semantic based, broader and narrower ideas (are suggested to help user browsing and refining the queries.

II. RELATED WORKS

Understanding web user behavior is essential for businesses striving to establish a strong online presence and effectively engage with their target audience. In today's digital age, where websites and web applications serve as primary touch points for user interaction, analyzing how users navigate, consume content, and interact with digital interfaces is crucial for optimizing user experience and driving business success. Therefore, research in web user behavior encompasses a broad range of disciplines, including human-computer interaction, web analytics, digital marketing, and data science. By integrating insights from these fields, researchers seek to unravel the complexities of user behavior and develop strategies to enhance user engagement, satisfaction, and conversion rates. User navigation patterns and information-seeking behavior play a pivotal role in shaping the user experience on websites and web applications. Researchers have extensively studied how users navigate through web pages, interact with hyperlinks, and search for information, drawing inspiration from theories such as information foraging and cognitive psychology. These studies have revealed valuable insights into factors influencing user decision-making, such as the perceived value of information, the cost of navigation, and the influence of external cues. By understanding these patterns, designers and developers can optimize website layouts, navigation structures, and content organization to facilitate seamless information retrieval and enhance user satisfaction.

Machine learning techniques offer powerful tools for analyzing and predicting user behavior on websites and web applications. Researchers leverage techniques such as clustering, classification, and predictive modeling to extract actionable insights from user data. For instance, machine learning algorithms can segment users based on their behavior, predict user preferences, and personalize content recommendations. Additionally, researchers use anomaly detection algorithms to identify unusual patterns or deviations in user behavior, such as fraudulent activities or bot traffic. By harnessing the power of machine learning, businesses can gain deeper insights into user behavior, improve targeting and personalization efforts, and drive better business outcomes. Despite the advancements in web user behavior research, researchers face several challenges and limitations in collecting, analyzing, and interpreting user data. One of the primary challenges is data privacy and compliance with regulations such as GDPR, which restrict the collection and use of personal data for research purposes. Additionally, researchers must contend with issues such as sample bias, data quality, and the dynamic nature of online interactions, which can affect the validity and reliability of research findings. Furthermore, the interdisciplinary nature of web user behavior research requires researchers to integrate insights from multiple disciplines, including psychology, sociology, computer science, and marketing, posing challenges in terms of collaboration and knowledge integration

III. METHODOLOGY

The proposed system for web user analysis aims to revolutionize how businesses understand and leverage user behavior data to enhance their online platforms. Leveraging cutting-edge deep learning techniques alongside traditional web analytics methodologies, this system offers a holistic approach to analyzing user interactions. By integrating recurrent neural networks (RNNs) and convolutional neural networks (CNNs), the system can capture intricate patterns in user behavior, providing deeper insights into navigation patterns, content preferences, and engagement levels. This hybrid approach allows businesses to extract high-level features from raw user data, enabling more accurate and nuanced analysis. Furthermore, advanced data preprocessing techniques ensure the reliability and integrity of the analysis results, allowing businesses to make data-driven decisions with confidence.

Gradient Boosting Machines (GBMs) are powerful machine learning algorithms used for both regression and classification tasks. GBMs belong to the ensemble learning family, which involves combining multiple individual models to create a stronger overall model. What sets GBMs apart is their ability to build additive models sequentially, with each new model focusing on the errors (residuals) of the previous ones, hence the term "gradient boosting." The fundamental principle behind GBMs is to iteratively fit weak learners, typically decision trees, to the residuals of the previous predictions, thereby reducing the errors at each step. Each subsequent tree is trained to predict the residuals of the combined ensemble, gradually improving the overall model's performance. This sequential approach allows GBMs

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 4, April 2024 |

| DOI:10.15680/IJMRSET.2024.0704009 |

to learn complex patterns in the data, making them highly effective in capturing nonlinear relationships and interactions between variables.

This work uses MCPSOs to achieve User Preferences (UPs) list based on PSOs which are bio-inspired stochastic population and whose modus operandi is based global optimizations. PSOs refer to the population as a swarm and individuals are particles. Every particle moves around the search space with a varying speed, looking for possible regions. This work transforms every particle into a particle agent which *r* gathers information on browsed/recorded histories, log files, searching times, and registration data. Every particle agent with the letter I encompasses 4 main characteristics.

- The particle agent's position in 'd' dimensional space $\vec{x}_i = (x_{i1}, x_{i2}, \dots x_{id})$.
- The best user preference position discovered so far in the search process $\vec{p}_i = (p_{i1}, p_{i2}, \dots p_{id})$.
- The best site identified in neighbourhood so far $\vec{n}_i = (n_{i1}, n_{i2}, \dots n_{id})$ and
- velocity $\vec{v}_i = (v_{i1}, v_{i2}, \dots v_{id})$.

Each particle agent's velocity and position are updated in iterations based on Equations (1) and (2):

$$\vec{v}_i(t+1) = \vec{v}_i(t) + c_1.r_1.(\vec{p}_{best} - \vec{x}_i) + c_2.r_2.(\vec{g}_{best} - \vec{x}_i) \quad (1)$$

$$\vec{x}_i(t+1) = \vec{x}_i(t) + \vec{v}_i(t+1) \quad (2)$$

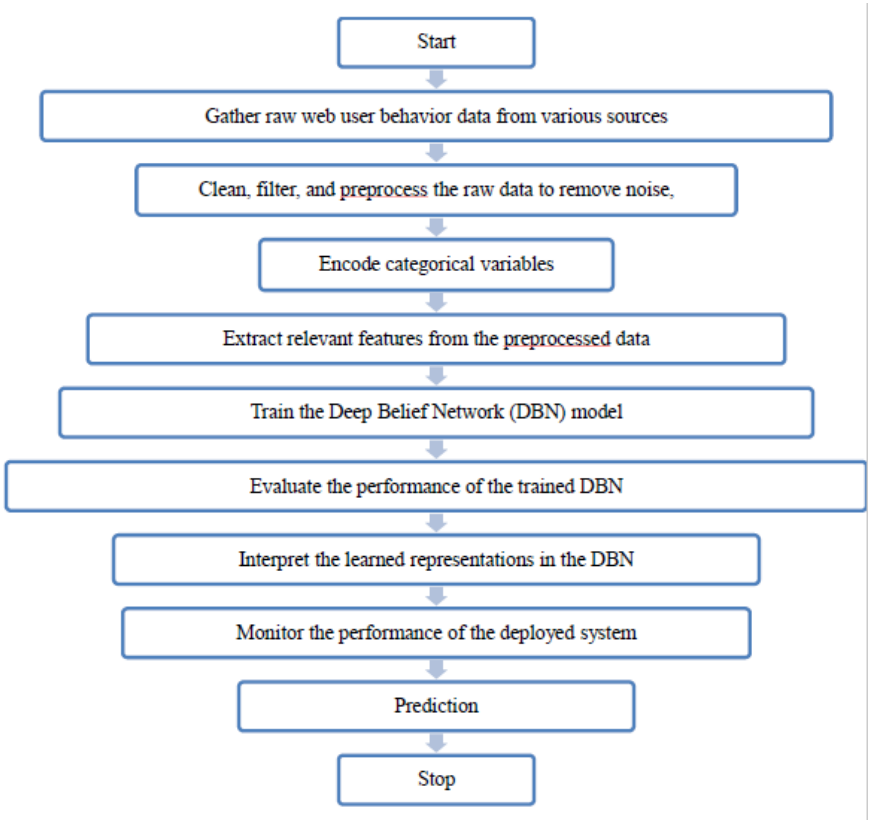

Fig-1 System Architecture

The implementation of the proposed web user analysis system involves several key steps to ensure its effectiveness and reliability. Firstly, it requires gathering and preprocessing relevant user interaction data from web servers or tracking tools. This involves cleaning the data, handling missing values, and transforming it into a suitable format for analysis. Next, the system integrates deep learning models, such as recurrent neural networks (RNNs) and

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 4, April 2024 |

| DOI:10.15680/IJMRSET.2024.0704009 |

convolutional neural networks (CNNs), using frameworks like TensorFlow or PyTorch. These models are trained on the preprocessed data to learn patterns in user behavior, such as navigation paths, content preferences, and engagement levels. Additionally, traditional web analytics methodologies, including data visualization and statistical analysis, are implemented to provide insights into user interactions. Interactive dashboards and customizable reports are developed to visualize key metrics and trends, making it easier for stakeholders to interpret and act upon the analysis results. Real-time analytics capabilities are also implemented to monitor user behavior continuously and detect anomalies or emerging trends promptly. Furthermore, the system incorporates experimentation and optimization features, allowing businesses to conduct A/B testing and iterative improvement processes. This involves deploying changes to websites or marketing strategies and evaluating their impact on user behavior through controlled experiments. Predictive modeling techniques, such as gradient boosting machines (GBMs) or random forests, are employed to forecast future user behavior trends, enabling businesses to anticipate customer needs and tailor their strategies accordingly.

Deep Belief Networks (DBNs) serve as the foundational element within the proposed system, offering a sophisticated framework for analyzing web user behavior. These deep learning models excel in autonomously learning hierarchical representations of data, eliminating the necessity for manual feature engineering. By leveraging DBNs, the system can adeptly capture intricate patterns and relationships inherent within web user behavior data, facilitating a deeper understanding of user interactions on websites and web applications. This capability to automatically extract meaningful features empowers the system to uncover nuanced insights that might otherwise remain obscured, enabling businesses to make informed decisions and optimize their online platforms effectively. Through the utilization of DBNs, the proposed system transcends traditional analytics methodologies, offering a robust and scalable solution for comprehensively analyzing and interpreting web user behavior.

IV. RESULTS AND DISCUSSION

The proposed architecture is executed by utilizing the python software. The test data web-documents are gathered from various domains such as Food, Education, News and Health-care. These documents are stored in the WSEs like Semantic search engine and then semantic information is extracted by using Ontology. The experimental process starts with the semantic extraction module in order to drew-out the significant ideas of the domains and it is translated into semantic patterns.

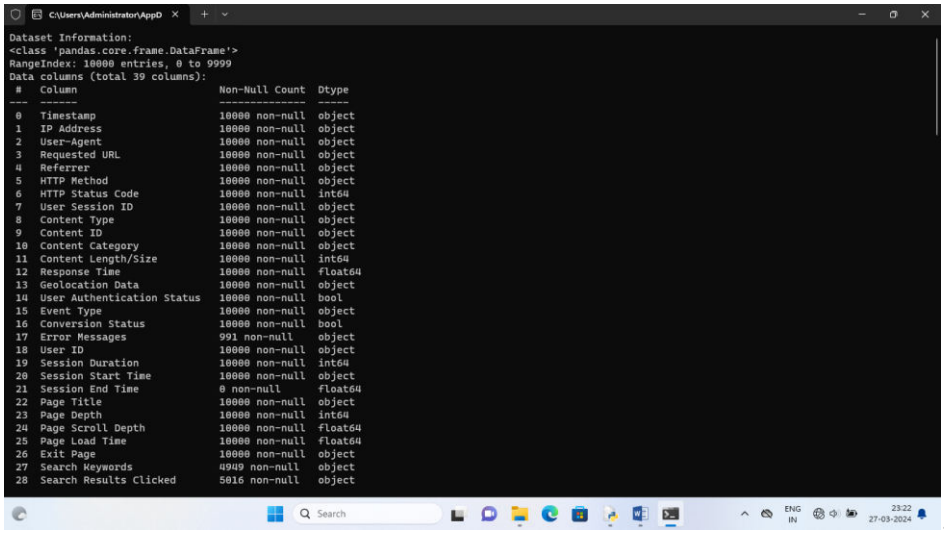

Fig-2 Attributes

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 4, April 2024 |

| DOI:10.15680/IJMRSET.2024.0704009 |

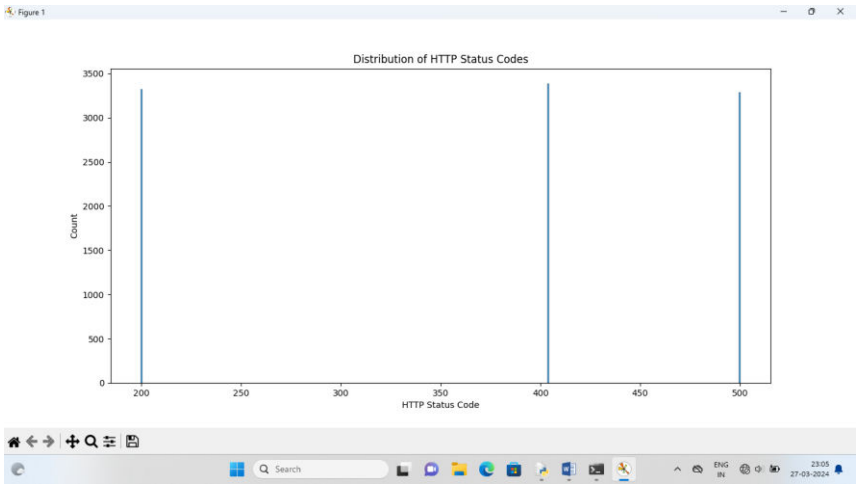

Fig-3 Http Cookies

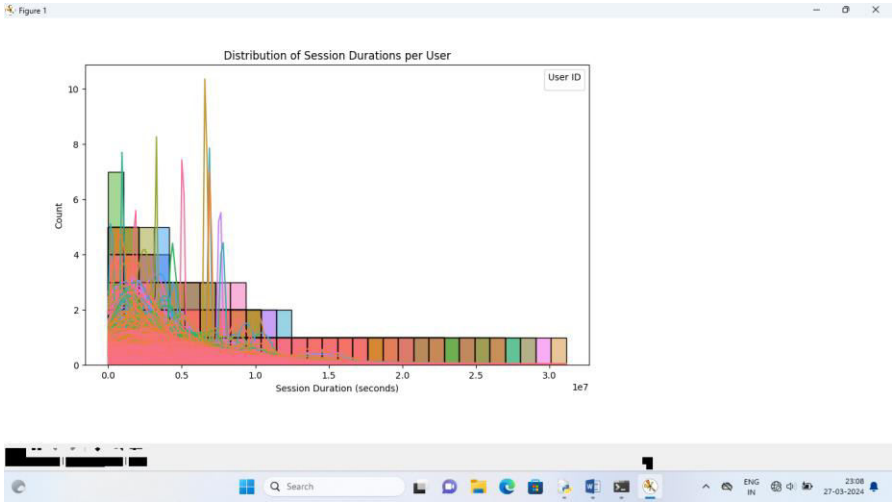

Fig-4 Distribution of Session Duration

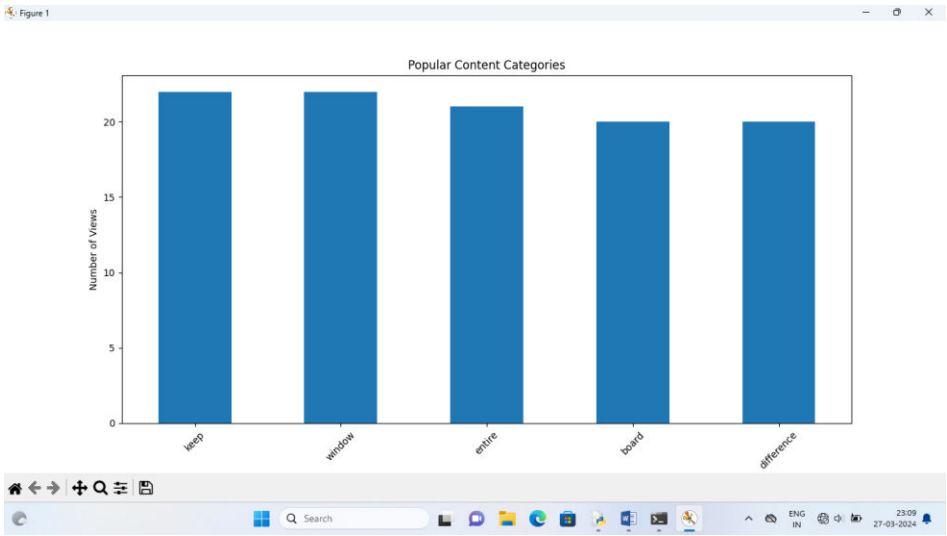

Fig-4 Popular Content categories

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 4, April 2024 |

| DOI:10.15680/IJMRSET.2024.0704009 |

V. CONCLUSION

In conclusion, the development of the web user behavior analysis system represents a significant step forward in leveraging data-driven approaches to understand and optimize user interactions on websites or web applications. Through the integration of advanced techniques such as Deep Belief Networks (DBNs) for analyzing complex patterns in user behavior data, the system enables businesses to gain valuable insights into user preferences, trends, and engagement metrics. By accurately capturing, preprocessing, and analyzing web user data, the system empowers stakeholders to make informed decisions regarding website design, content optimization, and marketing strategies. The performance testing and evaluation of machine learning models ensure the system's reliability, scalability, and effectiveness in handling diverse user traffic scenarios. Furthermore, the intuitive user interface facilitates easy access to analysis results, enabling stakeholders to derive actionable insights and drive continuous improvement in user experience and business performance. Overall, the web user behavior analysis system serves as a valuable tool for businesses seeking to enhance user engagement, optimize digital strategies, and achieve their objectives in an increasingly competitive online landscape.

REFERENCES

1. Molinillo, S., Aguilar-Illescas, R., Anaya-Sánchez, R. and Liébana-Cabanillas, F., 2021. Social commerce website design, perceived value and loyalty behavior intentions: The moderating roles of gender, age and frequency of use. Journal of Retailing and Consumer Services, 63, p.102404.
2. Zhu, Y., Xi, D., Song, B., Zhuang, F., Chen, S., Gu, X. and He, Q., 2020, April. Modeling users' behavior sequences with hierarchical explainable network for cross-domain fraud detection. In Proceedings of The Web Conference 2020 (pp. 928-938).
3. Anwyl-Irvine, A., Dalmaijer, E.S., Hodges, N. and Evershed, J.K., 2021. Realistic precision and accuracy of online experiment platforms, web browsers, and devices. Behavior research methods, 53(4), pp.1407-1425.
4. Chuklin, A., Markov, I. and De Rijke, M., 2022. Click models for web search. Springer Nature.
5. Ullah, M.U., Hassan, A., Asif, M., Farooq, M.S., Saleem, M. and Ullah, U., 2022. Intelligent intrusion detection system for apache web server empowered with machine learning approaches. International Journal of Computational and Innovative Sciences, 1(1), pp.21-27.
6. Dolega, L., Rowe, F. and Branagan, E., 2021. Going digital? The impact of social media marketing on retail website traffic, orders and sales. Journal of Retailing and Consumer Services, 60, p.102501.
7. Li, Y., Li, X. and Cai, J., 2021. How attachment affects user stickiness on live streaming platforms: A socio-technical approach perspective. Journal of Retailing and Consumer Services, 60, p.102478.
8. Duvaud, S., Gabella, C., Lisacek, F., Stockinger, H., Ioannidis, V. and Durinx, C., 2021. Expasy, the Swiss Bioinformatics Resource Portal, as designed by its users. Nucleic Acids Research, 49(W1), pp.W216-W227.
9. Ling, E.C., Tussyadiah, I., Tuomi, A., Stienmetz, J. and Ioannou, A., 2021. Factors influencing users' adoption and use of conversational agents: A systematic review. Psychology & marketing, 38(7), pp.1031-1051.
10. Zwilling, M., Klien, G., Lesjak, D., Wiechetek, Ł., Cetin, F. and Basim, H.N., 2022. Cyber security awareness, knowledge and behavior: A comparative study. Journal of Computer Information Systems, 62(1), pp.82-97.
11. Lorenz-Spreen, P., Lewandowsky, S., Sunstein, C.R. and Hertwig, R., 2020. How behavioural sciences can promote truth, autonomy and democratic discourse online. Nature human behaviour, 4(11), pp.1102-1109.
12. Wijaya, I.G.N.S., Triandini, E., Kabnani, E.T.G. and Arifin, S., 2021. E-commerce website service quality and customer loyalty using WebQual 4.0 with importance performances analysis, and structural equation model: An empirical study in shopee. Register, 7(2), pp.107- 124.
13. Shahbaznezhad, H., Dolan, R. and Rashidirad, M., 2021. The role of social media content format and platform in users' engagement behavior. Journal of Interactive Marketing, 53(1), pp.47-65.
14. Meel, P. and Vishwakarma, D.K., 2020. Fake news, rumor, information pollution in social media and web: A contemporary survey of state-of-the-arts, challenges and opportunities. Expert Systems with Applications, 153, p.112986.
15. Busalim, A.H. and Ghabban, F., 2021. Customer engagement behaviour on social commerce platforms: An empirical study. Technology in society, 64, p.101437

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com