International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

MICRO-DECISION IMPACT ANALYSIS USING USER CLICK STREAM DATA

[1]K. BINDU PRIYA, [2]Y VANAJA, [3]G PRAVEEN KUMAR, [4]K SANJANA, [5]V NAGASAI

[1]Assistant Professor, [2345]Students, Department of Computer Science and Engineering, SRI VASAVI INSTITUTE OF ENGINEERING & TECHNOLOGY, NANDAMURU, ANDHRA PRADESH

ABSTRACT

The rapid growth of e-commerce and web-based applications has increased the importance of understanding detailed user behavior during online interactions. Every action performed by a user on a digital platform, such as clicking, scrolling, hovering, or pausing on a page, represents a micro-decision that contributes to the overall outcome of a browsing session. Traditional web analytics systems typically focus on high-level metrics such as page views, bounce rates, and session duration, which often fail to explain the underlying behavioral patterns that influence user conversion. This research presents a system for Micro-Decision Impact Analysis using User Click Stream Data to capture and analyze fine-grained interaction events generated during user sessions. The proposed system integrates a lightweight JavaScript tracking mechanism embedded in the client application to record interaction signals including clicks, scroll behavior, hover duration, zoom interactions, and dwell time. These events are transmitted to a FastAPI-based backend server and stored in a structured MySQL database. After session completion, feature engineering techniques are applied to generate session-level behavioral metrics such as hesitation score, scroll velocity, hover duration, zoom interaction frequency, and review reading time. A machine learning model based on the XGBoost algorithm is then trained using both publicly available datasets and real-time session data to predict user conversion probability. To improve transparency and interpretability, SHAP (SHapley Additive exPlanations) is used to quantify the contribution of each feature to the prediction outcome. The system also supports automatic model retraining when new session data accumulates and provides role-based dashboards and an analytics chatbot to deliver actionable insights for improving user experience and optimizing conversion strategies.

I INTRODUCTION

The rapid advancement of digital technologies and the expansion of online services have significantly transformed the way users interact with web platforms. E-commerce websites, digital marketplaces, and service portals generate massive amounts of behavioral data that reflect how users navigate and interact with digital interfaces. Understanding these behavioral patterns is essential for improving user experience, increasing customer engagement, and enhancing conversion rates. Clickstream data represents a detailed record of the sequence of actions performed by users while browsing a website, including page visits, clicks, scrolling behavior, and interaction durations. These events provide valuable insights into the decision-making processes of users during their online sessions. Traditional web analytics tools mainly focus on aggregated statistics such as page views, session duration, and bounce rates, which provide only a limited understanding of user behavior [1]. Researchers have emphasized the importance of analyzing detailed interaction patterns to uncover deeper insights into how users interact with digital interfaces [2]. Clickstream analysis has been widely used to understand navigation behavior and identify patterns associated with user engagement [3]. Studies in web usage mining highlight that detailed behavioral signals can reveal hidden relationships between interface design and user actions [4]. Data mining techniques have been applied to analyze browsing patterns and discover meaningful behavioral insights from large web datasets [5]. With the increasing availability of large-scale data collection systems, researchers have started exploring more sophisticated analytical techniques to study user interactions and predict behavioral outcomes [6].

Machine learning has become an important tool for analyzing clickstream data and modeling user behavior in digital environments. Various predictive models have been developed to estimate the likelihood of user actions such as purchases, registrations, or session abandonment based on browsing patterns and interaction features [7]. Algorithms such as decision trees, logistic regression, and support vector machines have been applied to classify user sessions and predict conversion outcomes [8]. Recent advancements in gradient boosting algorithms have further improved the predictive capabilities of machine learning models for behavioral analytics [9]. Feature engineering techniques are used to transform raw clickstream data into meaningful behavioral indicators that represent user engagement and decision-making processes [10]. Metrics such as hesitation time, scrolling speed, hover duration, and interaction frequency provide valuable insights into how users evaluate content before making decisions [11]. Modern web technologies allow the integration of lightweight event tracking scripts that capture real-time interaction data from user sessions [12]. These systems enable organizations to collect detailed behavioral information without affecting website performance [13]. Furthermore, explainable artificial intelligence methods have been introduced to interpret machine learning predictions and identify the key factors influencing user decisions [14]. Techniques such as SHAP analysis help quantify the contribution of individual features to model predictions and improve transparency in decision-making systems [15]. As a result, micro-decision impact analysis has emerged as an important approach for understanding user behavior and improving digital platform performance [16].

II LITERATURE SURVEY

User behavior analysis in web environments has been widely studied in the fields of web mining, data analytics, and human–computer interaction. Early research in web usage mining focused on analyzing server logs to identify navigation patterns and frequently visited pages within websites [17]. These studies primarily relied on session statistics and page transition sequences to understand user browsing behavior. Data mining techniques such as clustering and association rule mining were commonly used to identify patterns in user navigation paths [18]. Although these approaches provided useful insights into website usage patterns, they lacked the ability to capture detailed user interactions occurring within individual pages. Later studies introduced clickstream analytics frameworks that collect detailed event-level data such as clicks, scrolling behavior, and mouse movements [19]. These systems provided a deeper understanding of user engagement by analyzing how users interact with different interface components. Researchers have shown that analyzing detailed interaction signals can significantly improve the accuracy of user behavior prediction models [20]. Predictive analytics techniques have been widely used to identify patterns that indicate purchase intent or session abandonment [21]. Logistic regression and decision tree models have been used to predict user conversion outcomes based on browsing behavior and interaction features [22]. However, these traditional models often struggle to capture complex relationships between multiple behavioral variables.

Recent research has focused on applying advanced machine learning algorithms to analyze large-scale clickstream datasets and extract meaningful behavioral insights. Gradient boosting models such as XGBoost have demonstrated superior performance in predicting user actions due to their ability to handle nonlinear relationships and complex feature interactions [23]. Researchers have also explored feature engineering techniques that transform raw interaction data into behavioral metrics such as engagement level, hesitation score, and exploration depth [24]. These features help capture subtle user behaviors that influence decision-making processes. Deep learning models have also been applied to sequential clickstream data to identify temporal patterns in browsing behavior [25]. Recurrent neural networks and sequence modeling techniques have been used to analyze the order and timing of user interactions within sessions [26]. In addition to predictive modeling, explainable artificial intelligence has become an important area of research in behavioral analytics. Methods such as SHAP and LIME provide interpretable explanations for machine learning predictions and help identify the factors that influence user decisions [27]. These techniques enable analysts to understand why certain behavioral features lead to specific outcomes and provide actionable insights for improving website design and user experience [28]. Modern analytics platforms also integrate visualization dashboards and conversational interfaces that allow stakeholders to explore behavioral insights through interactive reports and natural language queries [29]. Overall, the literature highlights the growing importance of combining clickstream data collection, machine learning analysis, and explainable AI techniques to understand user behavior in digital environments [30].

III METHODOLOGY

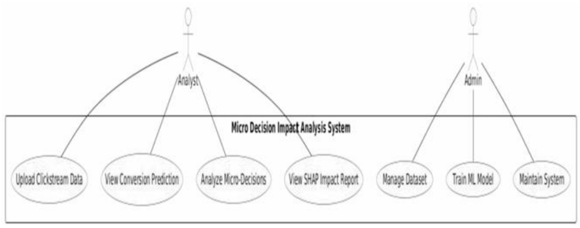
The methodology for the proposed Micro-Decision Impact Analysis system consists of multiple stages including data collection, preprocessing, feature engineering, machine learning modeling, and interpretability analysis. The process begins with the collection of user interaction data through a lightweight JavaScript tracking script embedded in the client-side application. This tracker captures various interaction events during a browsing session, including mouse clicks, scroll movements, hover durations, zoom interactions, and dwell time on specific content sections. Each interaction event is recorded with a timestamp and a unique session identifier to maintain session continuity. The captured events are transmitted to a backend server implemented using the FastAPI framework through RESTful API endpoints. The backend validates incoming data and stores the interaction logs in a MySQL database where session information, event types, and timestamps are organized in structured tables. After session completion, preprocessing techniques are applied to clean the data, remove incomplete records, and organize events into

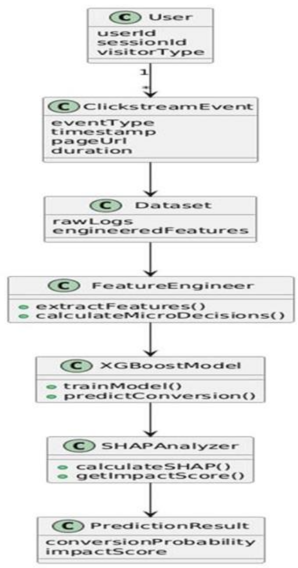
International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

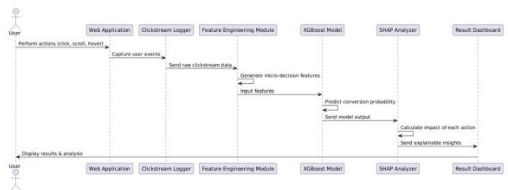
session-level datasets. The next stage involves feature engineering, where raw interaction logs are transformed into behavioral metrics that represent micro-decision indicators. These features include hesitation score, scroll velocity, hover duration, zoom interaction count, and review reading time. The engineered features are combined with labeled session outcomes indicating whether a user completed a conversion or abandoned the session. A machine learning model based on the XGBoost algorithm is then trained using both publicly available datasets and collected session data. The dataset is divided into training and testing subsets to evaluate model performance. After training, SHAP analysis is applied to interpret the model predictions by quantifying the contribution of each feature to the predicted conversion probability. The system also includes an automated retraining mechanism that updates the model when sufficient new session data becomes available.

IV SYSTEM DESIGN

The system design for the Micro-Decision Impact Analysis platform follows a modular architecture that integrates client-side event tracking, backend processing, data storage, machine learning analytics, and visualization components. The first component of the system is the client-side tracking module implemented using a lightweight JavaScript script embedded within the web application. This module monitors various user interactions including clicks, scrolling behavior, hover events, zoom actions, and page navigation activities. Each interaction event is recorded along with metadata such as timestamp, page identifier, session ID, and interaction type. The tracking module operates asynchronously to ensure that event logging does not affect the performance of the web application. Once the events are captured, they are transmitted securely to the backend server through RESTful API calls. The backend component of the system is implemented using the FastAPI framework, which handles event ingestion, validation, and data storage. FastAPI provides high performance and scalability for handling large volumes of interaction data generated by multiple users simultaneously. The collected clickstream data is stored in a MySQL database where structured tables maintain session information, event details, and interaction metrics. This storage architecture ensures efficient data retrieval for analytics and model training processes.







The analytics and machine learning layer forms the core intelligence component of the system. After session completion, stored interaction data is processed through feature engineering modules that convert raw event logs into meaningful behavioral metrics. These metrics represent micro-decision indicators such as hesitation score, scroll velocity, hover duration, zoom interaction frequency, and review reading time. The processed dataset is then used to train a predictive model based on the XGBoost algorithm. The model estimates the probability of user conversion for each browsing session. To improve transparency and interpretability, the system integrates SHAP analysis to explain model predictions and identify the most influential behavioral features affecting the outcome. The system also includes a model retraining module that periodically updates the machine learning model when sufficient new session data becomes available. Additionally, the platform provides role-based dashboards for different user personas including Admin, Analyst, and Machine Learning Engineer. These dashboards present visual analytics such as feature importance charts, session flow diagrams, and conversion prediction statistics. An integrated analytics chatbot enables users to ask natural language questions regarding behavioral patterns, feature impacts, and user experience recommendations derived from SHAP insights. This architecture ensures efficient data collection, scalable processing, and actionable insight generation for improving digital platform performance.

V PROPOSED SYSTEM

The proposed system introduces an intelligent framework for analyzing micro-level user interactions in digital platforms to better understand decision-making behavior and improve conversion outcomes. Traditional web analytics systems rely mainly on aggregated metrics such as page views, bounce rates, and session duration, which provide limited insights into the actual behavioral factors influencing user decisions. In contrast, the proposed system focuses on capturing detailed interaction signals that represent individual micro-decisions performed by users while navigating a website. These micro-decisions include actions such as clicking on specific elements, hovering over product descriptions, scrolling through content sections, zooming into images, and pausing on certain interface components. Each of these actions reflects a level of interest or hesitation that contributes to the overall decision-making process of the user. By capturing and analyzing these interaction events, the system provides a deeper understanding of how users evaluate content and navigate digital interfaces before completing or abandoning a transaction. The proposed framework includes a lightweight JavaScript tracker that records interaction events in real time without affecting page performance. These events are transmitted to a scalable backend infrastructure where they are stored, processed, and analyzed to identify behavioral patterns associated with successful conversions.

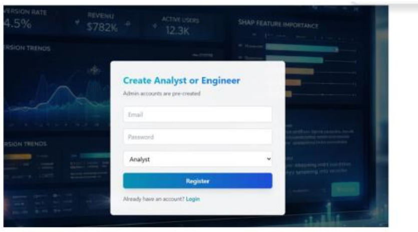
Another key feature of the proposed system is the integration of advanced machine learning techniques for predictive behavioral analysis. After collecting interaction data, feature engineering methods are applied to generate meaningful behavioral metrics such as hesitation score, interaction frequency, scroll velocity, and content engagement level. These features serve as inputs to a predictive model based on the XGBoost algorithm, which estimates the probability that a user session will result in a conversion. The model

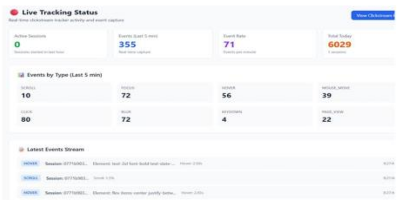
International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

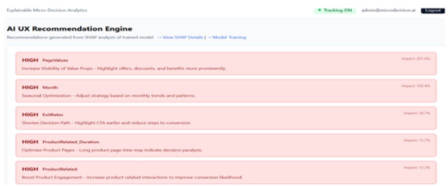
analyzes the relationships between behavioral signals and session outcomes to identify patterns that indicate purchase intent or abandonment risk. To ensure transparency and trust in the predictions generated by the model, the system integrates SHAP (SHapley Additive exPlanations) analysis. SHAP provides detailed explanations of how each feature contributes to the final prediction outcome, allowing analysts to understand the factors influencing user decisions. The system also includes automated model retraining capabilities that update the predictive model as new session data becomes available. In addition, role-based dashboards and an analytics chatbot provide intuitive interfaces for exploring behavioral insights and generating user experience recommendations. By combining real-time interaction tracking, machine learning prediction, and explainable AI techniques, the proposed system provides a comprehensive solution for micro-decision impact analysis in digital environments.

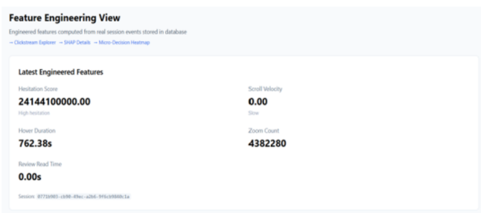
VI RESULTS & DISCUSSION

The experimental evaluation of the proposed system demonstrates its effectiveness in predicting user conversion outcomes using micro-level interaction data. The XGBoost model trained on engineered behavioral features achieved higher prediction accuracy compared to baseline models that relied only on traditional session metrics. Features such as hesitation score, hover duration, and scroll velocity were found to have significant influence on conversion predictions. SHAP analysis provided detailed explanations of the model predictions and identified the most influential features contributing to user decisions. The analysis revealed that sessions with higher engagement levels and longer content exploration durations were more likely to result in successful conversions. Additionally, the visualization dashboards allowed analysts to observe behavioral trends and identify common drop-off points within user sessions. The integrated analytics chatbot further enhanced usability by enabling stakeholders to query behavioral insights using natural language. These results confirm that micro-decision analysis provides deeper insights into user behavior compared to conventional web analytics methods.









VII CONCLUSION

Micro-Decision Impact Analysis using User Click Stream Data provides a powerful framework for understanding detailed user behavior in digital platforms. Traditional web analytics approaches rely mainly on aggregated metrics that fail to capture the subtle interaction patterns influencing user decisions. The system proposed in this study addresses this limitation by capturing fine-grained interaction events such as clicks, scrolling behavior, hover durations, and zoom interactions that represent individual micro-decisions made by users during browsing sessions. These interaction signals are collected through a lightweight tracking mechanism and processed using a scalable backend infrastructure for storage and analysis. Feature engineering techniques transform raw clickstream data into meaningful behavioral metrics that reflect user engagement and decision-making processes. A predictive machine learning model based on the XGBoost algorithm is used to estimate the likelihood of user conversion based on these behavioral indicators. The integration of SHAP explainability techniques enhances transparency by identifying the contribution of each feature to the model's predictions. The system also supports automated model retraining and provides role-based dashboards and conversational analytics tools to deliver actionable insights. Experimental results demonstrate that micro-interaction features significantly improve prediction accuracy compared to traditional analytics metrics. Overall, the proposed system provides a comprehensive solution for analyzing user behavior, optimizing user experience, and improving conversion rates in digital environments.

REFERENCES

1. Aggarwal, C. (2015). Data mining: The textbook. Springer.
2. Bishop, C. (2006). Pattern recognition and machine learning. Springer.
3. Bucklin, R., & Sismeiro, C. (2009). Clickstream data analysis. Marketing Science.
4. Chen, M., Mao, S., & Liu, Y. (2014). Big data analytics. Mobile Networks and Applications.
5. Cooley, R., Mobasher, B., & Srivastava, J. (1999). Data preparation for web usage mining. Knowledge and Information Systems.
6. Davenport, T., & Harris, J. (2007). Competing on analytics. Harvard Business School Press.
7. Domingos, P. (2012). A few useful things to know about machine learning. Communications of the ACM.
8. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.
9. Friedman, J. (2001). Gradient boosting machine. Annals of Statistics.
10. Han, J., Kamber, M., & Pei, J. (2012). Data mining: Concepts and techniques. Morgan Kaufmann.
11. Hastie, T., Tibshirani, R., & Friedman, J. (2009). Elements of statistical learning.
12. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation.
13. Jansen, B. (2017). Understanding user behavior on the web.

14. Ke, G., et al. (2017). LightGBM framework. NeurIPS.

15. Kimball, R., & Ross, M. (2013). Data warehouse toolkit.

16. Lundberg, S., & Lee, S. (2017). SHAP explanations. NeurIPS.

17. Marr, B. (2016). Big data in practice.

18. Molnar, C. (2020). Interpretable machine learning.

19. Mobasher, B. (2007). Web personalization. Adaptive Web.

20. Montgomery, A. et al. (2004). Modeling online browsing behavior. Management Science.

21. Pirolli, P. (2007). Information foraging theory.

22. Provost, F., & Fawcett, T. (2013). Data science for business.

23. Chen, T., & Guestrin, C. (2016). XGBoost. KDD Conference.

24. Russell, S., & Norvig, P. (2021). Artificial intelligence: A modern approach.

25. Shmueli, G., Bruce, P., & Patel, N. (2019). Data mining for business analytics.

26. Srivastava, J. et al. (2000). Web usage mining. ACM SIGKDD Explorations.

27. Tan, P., Steinbach, M., & Kumar, V. (2019). Introduction to data mining.

28. Eckerson, W. (2010). Performance dashboards.

29. Knaflic, C. (2015). Storytelling with data.

30. Provost, F., & Murray, S. (2019). Data science for business.

31. Mahesh Ganji. (2025). Enhancing Oracle Cloud HR Reporting Through AI-Driven Automation. Journal of Science & Technology, 10(6), 28–36. https://doi.org/10.46243/jst.2025.v10.i06.pp28-36

32. Todupunuri, A. (2025). THE ROLE OF AGENTIC AI AND GENERATIVE AI IN TRANSFORMING MODERN BANKING SERVICES. American Journal of AI Cyber Computing Management, 5(3), 85–93. https://doi.org/10.64751/ajaccm.2025.v5.n3.pp85-93

33. Todupunuri, A. . (2024). Artificial Intelligence Ethics: Investigating Ethical Frameworks, Bias Mitigation, and Transparency in AI Systems to Ensure Responsible Deployment and Use of AI Technologies. International Journal of Innovative Research in Science,Engineering and Technology, 13(09), 1–14. https://doi.org/10.15680/ijirset.2024.1309002

34. Sushma Babburi. (2025). Token-Based Data Accounting System For Transparent Model Training And Cost Allocation. American Journal of AI Cyber Computing Management, 5(4), 463–474. https://doi.org/10.64751/ajaccm.2025.v5.n4.pp463-474

35. Snigdha Gaddam. (2025). SOFTWARE STACK PREPARED FOR AI TRANSITIONING FROM MODULES TO MODELS. American Journal of AI Cyber Computing Management, 5(4), 451–462. https://doi.org/10.64751/ajaccm.2025.v5.n4.pp451-462

36. Gaddam, S. INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING.

37. Bajarang Bhagwat, V. (2023). Optimizing Payroll to General Ledger Reconciliation: Identifying Discrepancies and Enhancing Financial Accuracy. JOURNAL OF ADVANCE AND FUTURE RESEARCH, 1(4). https://doi.org/10.56975/jaafr.v1i4.501636

38. Srinivasa Kalyan Immadi. (2025). Harnessing Artificial Intelligence In Oracle Hcm: Revolutionising Workforce Management With Automation And Predictive Analytics. International Journal of Data Science and IoT Management System, 4(4), 7–13. https://doi.org/10.64751/ijdim.2025.v4.n4.pp7-13

39. S. M. K. P. (2025). Cryptography in iOS: A Study of Secure Data Storage and Communication Techniques. International Journal on Science and Technology, 16(1). https://doi.org/10.71097/ijsat.v16.i1.1403

40. Suhasnadh Reddy Veluru, Sai Teja Erukude, and Viswa Chaitanya Marella. 2025. Multimodal Detection of Fake Reviews using BERT and ResNet-50. In 2025 4th International Conference on Innovative Mechanisms for Industry Applications (ICIMIA). IEEE, 877–882.

41. Cyril, H. P. (2025). Event-Driven Provisioning Architectures For Modern Telecom Networks: Overcoming Legacy Limitations And Enabling Autonomous 6g Operations. International Journal of Advanced Research in Computer Science, 16(6), 75–82. https://doi.org/10.26483/ijarcs.v16i6.7389

42. Jay Bharat Mehta. (2025). AUTONOMOUS PATCH VALIDATION FOR ZERO-DAY EXPLOITS IN ENTERPRISE CLOUDS. International Journal of Applied Mathematics, 38(4s), 1270–1285. https://doi.org/10.12732/ijam.v38i4s.685

43. Reddy, S. K. (2025). Hyperpersonalization driven by AI is expected to be at the Lead in shaping the future of loyalty rewards. Journal of Emerging Technologies and Innovative Research.

44. Reddy, S. K. R. (2021). Strengthening the Security of Loyalty Reward Systems: An In-Depth Analysis of Emerging Cyber Threats and Protection Mechanisms. Journal of Computational Analysis and Applications, 29(6).

45. Poojari, R. (2026). Privacy-Preserving Generative AI in Healthcare Systems Using Federated Learning Approaches. International Journal of Data Science and IoT Management System, 5(1), 78-88.

46. Uday Kumar Kalae. (2025). AN AUTOMATED SYSTEM FOR MANAGING HIGH-AVAILABILITY CLOUD INFRASTRUCTURE THROUGH INFRASTRUCTURE-ASCODE (IAC) PRACTICES. American Journal of AI Cyber Computing Management, 5(2), 42–50. https://doi.org/10.64751/ajaccm.2025.v5.n2.pp42-50

47. Saikumar, B. (2024). Optimizing Crew Scheduling and Absence Management using Microservices: Enhancing Reliability and Efficiency in Crew Management Systems. International Journal of Enhanced Research in Management & Computer Applications, 13(11), 50–55. https://doi.org/10.55948/ijermca.2024.0116

48. Saikumar, B. (2023). Enhancing Client Engagement through AI-Driven Real-Time

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

Reporting and Automated Alerts. International Journal of Enhanced Research in Science, Technology & Engineering, 12(11), 111–117. https://doi.org/10.55948/ijerste.2023.1115