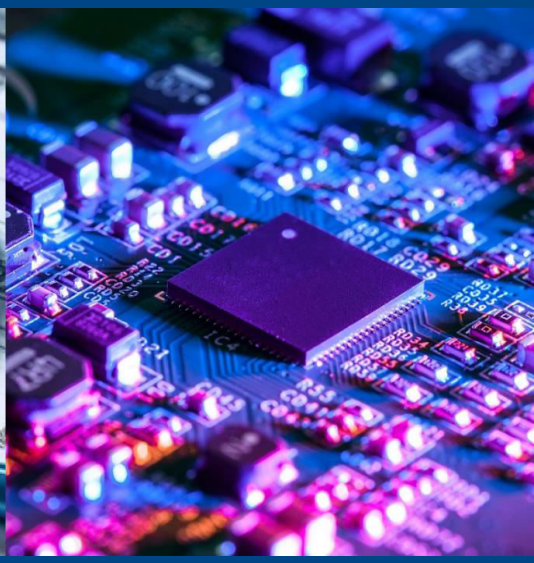
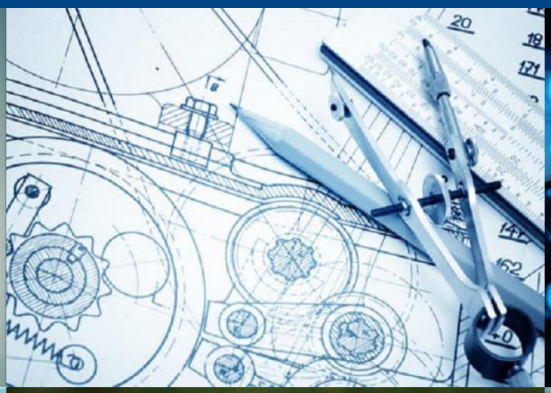
ISSN: 2582-7219

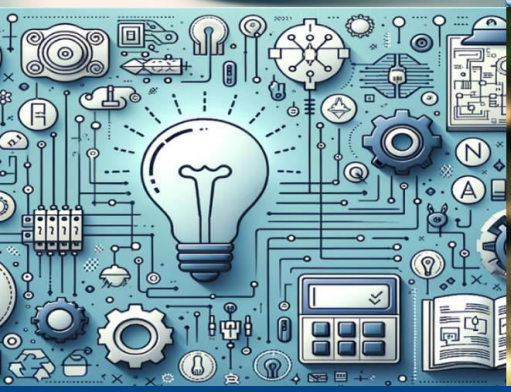




International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 9.864

Volume 9, Issue 6, June 2026

© 2026 IJMRSET | Volume 9, Issue 6, June 2026| DOI:10.15680/IJMRSET.2026.0906025

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 9.864| ESTD Year: 2018|

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

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Real Time Web Usage Analysis and Behavioral Tracking System

P. Stefina Angel [1], Mrs.A.Mary Joycy[2]

PG Scholar, Department of Master of Computer Applications, R.V.S. College of Engineering, Dindigul, Tamil Nadu, India[1]

Assistant Professor, Department of Master of Computer Applications, R.V.S. College of Engineering, Dindigul, Tamil Nadu, India[2]

ABSTRACT: The Real-Time Web Usage Analysis and Behavioral Tracking System addresses the critical shortcomings of traditional batch-processing analytics tools by delivering a low-latency, distributed framework capable of capturing and interpreting user behavior as it happens. By tracking micro-behaviors such as mouse movements, scroll patterns, keystroke dynamics, and rage clicks through a lightweight JavaScript tracker and processing them via Apache Kafka and Apache Flink, the system generates real-time behavioral insights with sub-second latency. Integrated machine learning modules enable accurate bot detection and fraud prevention, while a live WebSocket-powered dashboard provides dynamic heatmaps, session replays, and conversion analytics. Built with Privacy-by-Design principles and full GDPR and CCPA compliance, this system transforms raw user interaction data into actionable intelligence, empowering organizations to enhance user experience, strengthen security, and adopt a truly data-driven operational model.

I. INTRODUCTION

The Real-Time Web Usage Analysis and Behavioral Tracking System is an advanced, multi-tier platform designed to overcome the critical limitations of traditional web analytics tools, which rely on batch-processing models that introduce significant delays in data availability. Unlike conventional systems that only track high-level metrics such as page views and bounce rates, this system captures granular micro-behaviors — including mouse trajectories, scroll velocity, hover durations, keystroke dynamics, and rage clicks — providing a high-fidelity picture of user intent as it unfolds in real time. Built on a distributed, event-driven architecture comprising a lightweight JavaScript tracker, a high-throughput Apache Kafka ingestion pipeline, a real-time stream processing engine, and a live WebSocket-powered dashboard, the system processes behavioral data with sub-second latency.

It further integrates AI and machine learning modules for bot detection, fraud prevention, and friction identification, while adhering to Privacy-by-Design principles to ensure full compliance with GDPR and CCPA regulations. In essence, this system transforms raw user interaction data into actionable intelligence, empowering organizations to enhance user experience, strengthen security, and operate with a truly data-driven mindset.

II. EXISTING SYSTEM

Current web analytics platforms mainly depend on batch processing, causing delays in data analysis and limiting visibility to only high-level interactions such as page views and clicks while missing complete user behavior patterns. These systems use static identifiers like IP addresses and cookies, making them vulnerable to bots and spoofing techniques. Advanced tools such as Hotjar and FullStory provide deeper tracking but often increase browser load due to frequent DOM snapshots, negatively affecting performance, especially on mobile devices. Additionally, third-party hosting creates challenges related to data privacy, compliance, and control over data management. As a result, existing solutions lack real-time analytics, deep behavioral insights, strong security detection, and data ownership needed for modern digital environments.

© 2026 IJMRSET | Volume 9, Issue 6, June 2026| DOI:10.15680/IJMRSET.2026.0906025

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 9.864| ESTD Year: 2018|

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

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Disadvantages of Existing Systems

1. Existing systems use batch processing, which causes delays in data availability and limits real-time decision-making and response.
2. Traditional analytics track only basic events such as page views and clicks, while ignoring detailed user behavior and relying on identifiers like IP addresses and cookies that can be easily spoofed.
3. Heatmapping and session replay tools increase browser workload through continuous tracking, leading to slower performance, especially on mobile devices.
4. Third-party hosted platforms provide limited control over data privacy, storage, and compliance, increasing risks related to regulations and data security.

III. PROPOSED SYSTEM

The proposed **Real-Time Web Usage Analysis and Behavioral Tracking System** is developed to provide instant and intelligent analysis of user behavior through a Streaming First architecture. The system includes four layers: a lightweight JavaScript tracker for capturing user interactions, a data ingestion layer using Kafka or RabbitMQ, a real-time processing engine using Flink or Spark Streaming, and a visualization layer for analytics and storage. It monitors user activities such as mouse movement, scrolling, hovering, and typing patterns with minimal impact on browser performance. AI and machine learning are integrated to distinguish human users from automated bots in real time, improving security and reducing malicious activity. The system also follows a Privacy-by-Design approach with edge-level data masking to ensure GDPR and CCPA compliance. Overall, it transforms web interaction data into real-time insights to improve user experience, strengthen security, and support faster decision-making.

Advantages of Proposed System

1. The proposed system uses real-time stream processing to deliver behavioral insights with very low latency, allowing organizations to make faster decisions and respond instantly to user actions.
2. It captures detailed user behavior such as mouse movement, scrolling patterns, hover time, keystrokes, and interaction events, providing deeper understanding compared to traditional click-based analytics.
3. AI and machine learning techniques enable real-time identification of human users and detection of malicious bots, improving security and reducing risks such as automated attacks and data scraping.
4. The lightweight asynchronous tracking mechanism collects data efficiently without affecting browser performance, ensuring smooth user experience and faster page loading.

IV. SYSTEM ARCHITECTURE

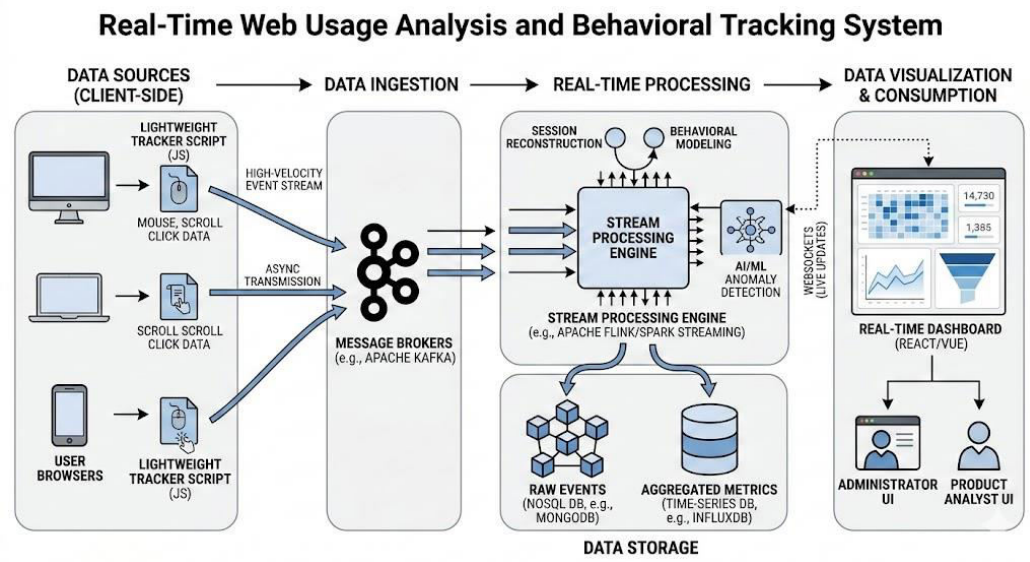


Fig No: 4.1

© 2026 IJMRSET | Volume 9, Issue 6, June 2026| DOI:10.15680/IJMRSET.2026.0906025

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 9.864| ESTD Year: 2018|

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

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

V. MODULE DESCRIPTION

1.Client-Side Data Acquisition Module (The Tracker)
2. Data Ingestion and Gateway Module
3. Real-Time Stream Processing Module
4. AI-Driven Anomaly Detection Module
5. Visualization and Administrative Module

1. Client-Side Data Acquisition Module (The Tracker)
This module collects user behavior data directly from the browser using a lightweight and asynchronous JavaScript tracker. It captures interactions such as mouse movement, clicks, scrolling, and typing patterns without affecting browser performance. Additional session details like browser type, device information, and URL are also collected. The captured data is compressed and sent efficiently using Beacon API or WebSockets to ensure smooth website performance.

2. DataIngestion and Gateway Module
This module acts as the entry point for behavioral data collection and ensures secure and reliable data transfer. It validates incoming requests and uses message brokers such as Apache Kafka or RabbitMQ to process large volumes of events efficiently. By separating data collection from processing, the system remains stable and available even during heavy traffic conditions.

3. Real-Time Stream Processing Module
This module processes raw user interaction data and converts it into meaningful insights. Using Apache Flink or Spark Streaming, it reconstructs user sessions, calculates engagement metrics, and identifies user behavior patterns. Complex Event Processing helps detect significant actions and generate real-time alerts for improving user experience and decision-making.

4. AI-Driven Anomaly Detection Module
This module uses artificial intelligence and machine learning to detect bots, suspicious activity, and usability issues in real time. It analyzes interaction characteristics such as mouse movement patterns and click behavior to distinguish humans from automated scripts. It also identifies friction points like rage clicks and dead clicks to improve interface quality and security.

6.Visualization and Administrative Module
This module presents processed behavioral insights through an interactive dashboard for administrators and analysts. Built using technologies such as React or Vue.js, it provides real-time updates through WebSockets, displays heatmaps, and generates reports on user engagement, conversion rates, session activity, and security events to support data-driven decisions.

VI. RESULTS

1. Admin Login Page

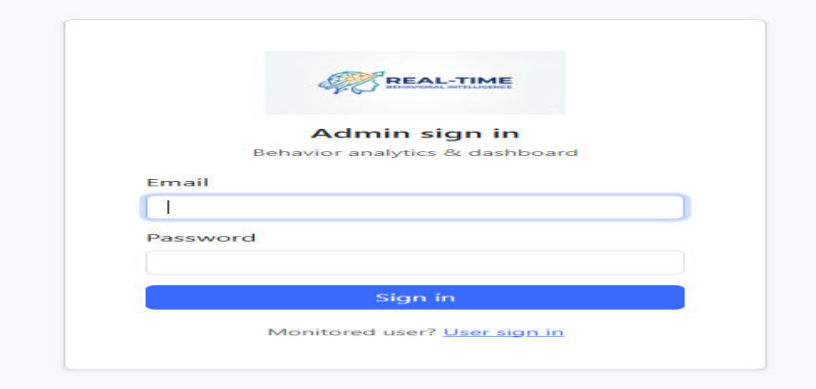


Fig No :6.1

© 2026 IJMRSET | Volume 9, Issue 6, June 2026| DOI:10.15680/IJMRSET.2026.0906025

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 9.864| ESTD Year: 2018|

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

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

2. Dashboard Page

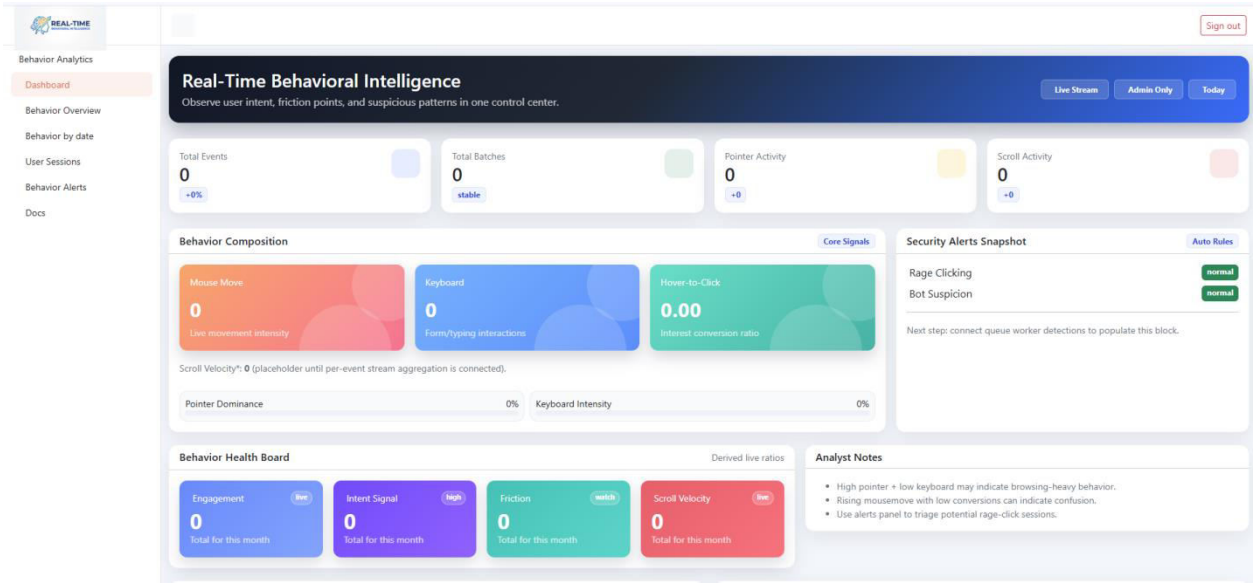


Fig No :6.2

3. User Behavior Data Page

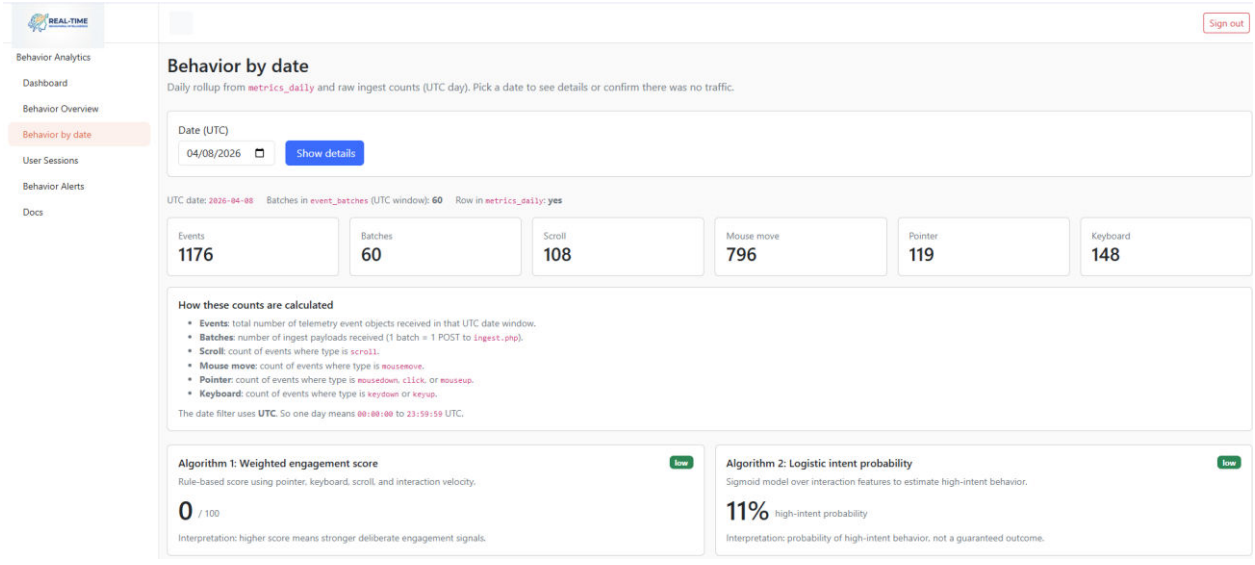


Fig No: 6.3

© 2026 IJMRSET | Volume 9, Issue 6, June 2026| DOI:10.15680/IJMRSET.2026.0906025

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 9.864| ESTD Year: 2018|

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

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

4. Behavior Alerts Page

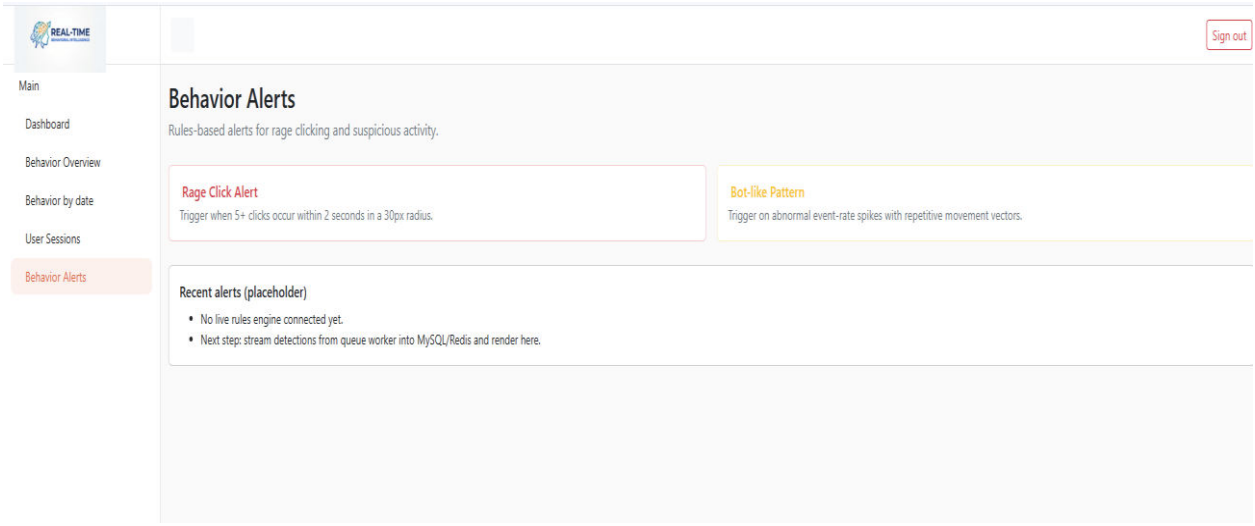


Fig No: 6.4

VII. CONCLUSION

The **Real-Time Web Usage Analysis and Behavioral Tracking System** provides an advanced solution to overcome the limitations of traditional web analytics by enabling instant and intelligent analysis of user behavior. Using a Streaming First architecture with a lightweight JavaScript tracker, Kafka-based data ingestion, real-time processing, and live dashboards, the system delivers fast and actionable insights with minimal delay. The integration of AI and machine learning improves security by detecting malicious bots and analyzingbehavioral patterns in real time. The system also captures detailed user interactions such as mouse movement, scrolling, hovering, and typing behavior to provide deeper understanding of user experience and support interface improvements. In addition, its Privacy-by-Design approach ensures secure data handling and compliance with GDPR and CCPA regulations. Overall, the system transforms raw web interaction data into valuable behavioral intelligence, helping organizations improve user experience, strengthen security, and support effective data-driven decision-making.

VIII. FUTURE ENHANCEMENT

The proposed **Real-Time Web Usage Analysis and Behavioral Tracking System** can be enhanced in the future by integrating advanced AI models such as LSTM and Transformer architectures to enable predictive behavioral analytics and proactive user engagement. The system can be extended to support mobile platforms on iOS and Android, creating a unified cross-platform behavioral tracking environment. Additional improvements may include computer vision for emotion detection, Natural Language Processing for analyzing user feedback and search intent, and edge computing to reduce latency and improve performance. Future versions can also support emerging interaction methods such as gaze tracking, gesture recognition, and voice-based analytics for AR and smart environments. Furthermore, adopting federated learning techniques can improve AI model accuracy while preserving user privacy by keeping behavioral data on local devices, making the system more intelligent, scalable, and privacy-focused for future digital ecosystems.

REFERENCES

1. Clifton, B. (2012). Advanced Web Metrics with Google Analytics. John Wiley & Sons.
2. Atterer, R., Wnuk, M., & Schmidt, A. (2006). Knowing the user's every move: User activity tracking for website usability evaluation and implicit interaction. Proceedings of the 15th International Conference on World Wide Web (WWW '06), 203–212.
3. Apache Software Foundation. (2023). Apache Kafka Documentation: A Distributed Streaming Platform. Retrieved

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 9.864| ESTD Year: 2018|

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

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

from https://kafka.apache.org/documentation/

4. Apache Software Foundation. (2023). Apache Flink: Stateful Computations over Data Streams. Retrieved from https://flink.apache.org/

5. Zheng, X., & Liu, M. (2011). An overview of online botnet detection techniques. International Journal of Computational Intelligence Systems, 4(2), 151–164.

6. European Parliament. (2016). General Data Protection Regulation (GDPR) — Regulation (EU) 2016/679. Official Journal of the European Union.

7. California State Legislature. (2018). California Consumer Privacy Act (CCPA) — AB-375 Privacy: Personal Information. California Legislative Information.

8. Sculley, D., Holt, G., Golovin, D., Davydov, E., Phillips, T., Ebner, D., & Dennison, D. (2015). Hidden technical debt in machine learning systems. Advances in Neural Information Processing Systems, 28, 2503–2511.

9. Marz, N., & Warren, J. (2015). Big Data: Principles and Best Practices of Scalable Real-Time Data Systems. Manning Publications.

10. Fielding, R. T. (2000). Architectural Styles and the Design of Network-based Software Architectures (Doctoral dissertation). University of California, Irvine.

11. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.

12. MongoDB Inc. (2023). MongoDB Documentation: The Most Popular Database for Modern Applications. Retrieved from https://www.mongodb.com/docs/

13. Hinton, G., &Salakhutdinov, R. (2006). Reducing the dimensionality of data with neural networks. Science, 313(5786), 504–507.

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