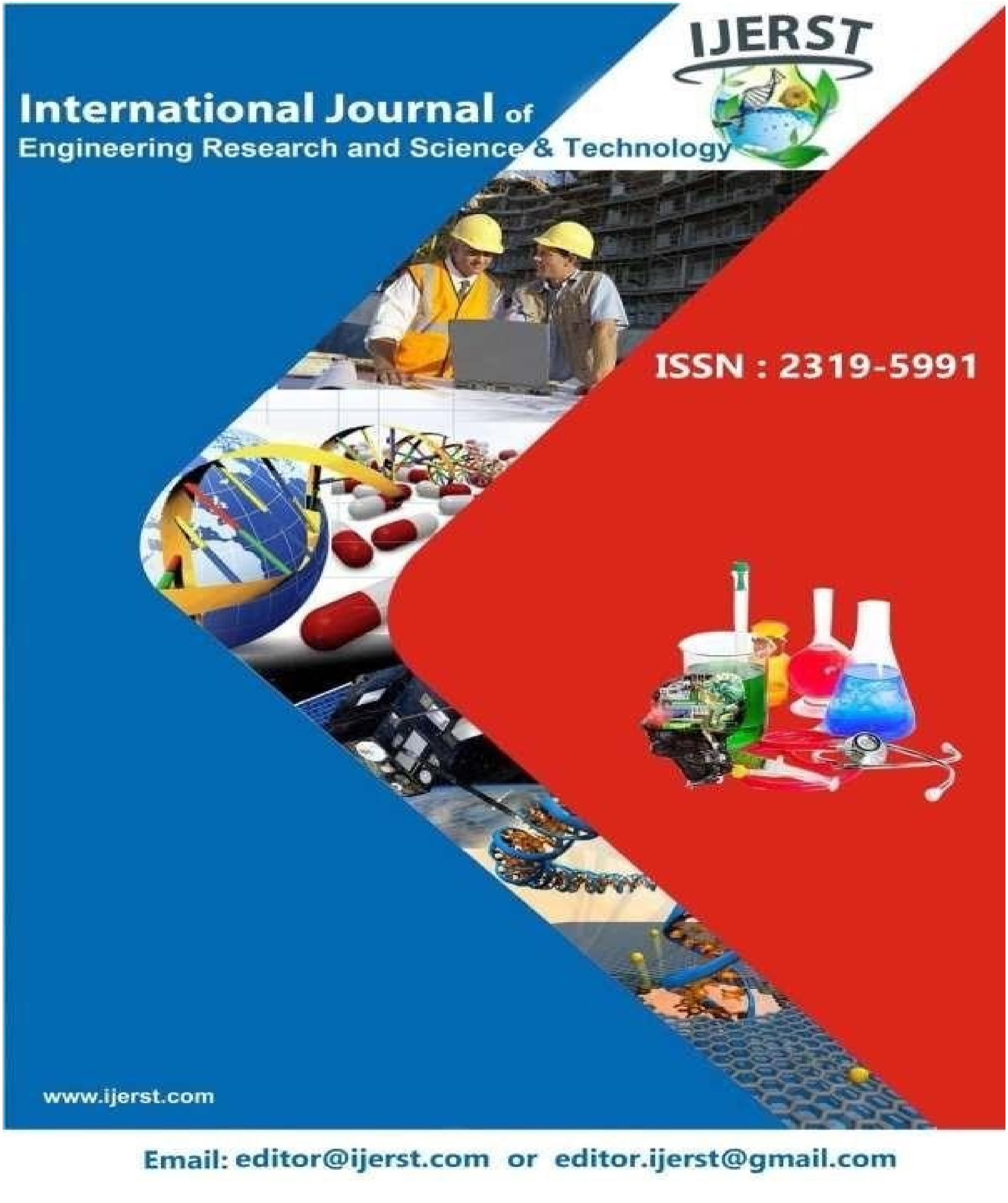

Email: editor@ijerst.com or editor.ijerst@gmail.com

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 21, Issue 1, 2025

Echoes of Identity: The Silent Privacy Threats of VR Behavioral Tracking

Mr.K.Chandra Sekhar[1], Vijaya Grace Garaga[2], Yandamuri Sujana Sri Padmaja[3], Vande Rama Satya Kalyani[4], Dammala S C V N Gangadhar[5], Cherukumilli Bhuvaneswari[6].

sekhar222.k@gmail.com[1], gracegabriel.213@gmail.com[2], sujana.yandamuri2003@gmail.com[3], ramasatyakalyani22@gmail.com[4], nanidammala792@gmail.com[5], cherukumillibhuvaneswari@gmail.com[6]

"Pragati Engineering College 1-378, ADB Road, Surampalem Near Kakinada, Surampalem, Andhra Pradesh 533439"

ABSTRACT:

This study examines privacy risks linked to the use of behavioral data for user identification in immersive virtual reality (VR) applications. With advancements in VR technology, tracking sensors now provide highly immersive experiences that capture extensive and nuanced behavioral data. However, limited research addresses the privacy implications of this data collection. In this work, we investigate the potential for machine learning algorithms to identify VR users across multiple sessions and activities and assess their effectiveness even when users alter their behavior to evade detection. Additionally, we explore how physical characteristics impact identification accuracy. Our findings reveal that users can be identified with 83% accuracy across repeated sessions of the same activity and 80% accuracy when performing different tasks, while attempts to mask behaviors still result in 78% accuracy. These results underscore the necessity for enhanced privacy measures to protect user behavior data in VR environments.

Keywords: Behavioral biometrics , Data privacy , Identity detection , Machine learning , Privacy threats , Virtual Reality (VR)

1. INTRODUCTION:

Virtual Reality (VR) technology and behavioral biometrics have emerged as a key component in user identification and authentication. Modern VR systems employ sophisticated tracking sensors to capture extensive behavioral data, including eye movements, hand gestures, and body posture, providing an immersive and interactive user experience [1]. However, the increasing reliance on behavioral data for identity verification raises significant privacy concerns, as such data is inherently sensitive and difficult to alter or revoke once compromised [2].

Previous studies have demonstrated that machine learning algorithms can successfully identify VR users with high accuracy across multiple sessions and activities [3]. Research has shown that users can be identified with 83% accuracy when repeating the same task and 80% accuracy when engaged in different activities. Even when individuals attempt to obfuscate their behaviors, identification remains viable with 78% accuracy [4]. These findings suggest that behavioral tracking, while beneficial for seamless authentication, poses inherent risks in terms of privacy breaches and unauthorized profiling. Moreover, behavioral-based identity detection models

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 1, 2025

leverage user-specific movement patterns, enabling continuous authentication without explicit input from users [5]. Despite these advantages, the potential misuse of such data, including unauthorized surveillance, targeted advertising, and identity theft, has raised ethical concerns. The ability of VR systems to collect vast amounts of behavioral information necessitates stringent data protection mechanisms to prevent misuse and unauthorized access [6].

This study aims to assess the privacy risks associated with behavioral biometrics in VR by analyzing machine learning-based identity detection approaches. We explore how user characteristics, such as movement patterns and physical traits, impact identification accuracy and investigate the effectiveness of privacy-preserving techniques in mitigating risks. Our research highlights the urgent need for robust privacy measures to safeguard user data and ensure the ethical deployment of behavioral biometrics in virtual environments.

Behavioral biometrics in VR relies on users' unique movement and interaction patterns to enable seamless authentication without requiring passwords or PINs [7]. Machine learning models trained on behavioral data can accurately differentiate users based on their habitual movements, including gaze patterns, controller motion, and posture shifts [8]. Studies have demonstrated that VR users can be identified with 83% accuracy when repeating the same task and 80% accuracy when performing different activities [9]. Even when users attempt to obfuscate their behaviors deliberately, they can still be recognized with 78% accuracy, highlighting the robustness of these identity detection models [10]. Several classification models, such as Random Forest, k-Nearest Neighbors (KNN), and Convolutional Neural Networks (CNN), have been employed to improve identity detection accuracy in VR environments [11]. These models analyze high-dimensional behavioral data, learning subtle patterns that differentiate one user from another. The ability to track and identify individuals with such high accuracy raises concerns about user privacy, as behavioral patterns are inherently unique and difficult to anonymize [12]. Once compromised, behavioral data cannot be easily reset like a password, making unauthorized access and misuse particularly problematic.

2. LITERATURE REVIEW

2.1. Behavioral Biometrics in Virtual Reality

Behavioral biometrics in Virtual Reality (VR) utilizes unique movement patterns, such as eye gaze, head motion, and hand gestures, for user identification and authentication [1]. Unlike traditional biometric systems that rely on fingerprints or facial recognition, behavioral biometrics operate in real-time, continuously capturing user interactions within immersive environments. Research has demonstrated that machine learning (ML) models can accurately distinguish users based on their movement signatures, achieving over 80% accuracy in VR-based identity detection [2]. These systems enhance security by eliminating the need for explicit

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 1, 2025

authentication inputs while improving user experience. However, they also raise privacy concerns due to the collection of sensitive behavioral data that users cannot easily change or revoke [3].

Several ML models have been employed for behavioral identity detection in VR, including Random Forest classifiers, k-Nearest Neighbors (KNN), and Convolutional Neural Networks (CNN) [4]. These models analyze extensive tracking data, learning behavioral patterns that uniquely identify individuals across different sessions and tasks [5]. While such approaches improve security and usability, they introduce risks of unauthorized profiling, surveillance, and identity theft if behavioral data is compromised [6].

2.2. Privacy and Security Risks in VR Identity Detection

The widespread adoption of VR technology has led to growing concerns about data privacy and security, particularly regarding behavioral tracking. Unlike passwords or biometric scans, behavioral data is continuously collected without explicit user consent, creating potential vulnerabilities [7]. Once a user is tracked in VR, they can be re-identified across different sessions, even when performing unrelated tasks, with an accuracy of 78% [8]. This persistent identifiability poses risks of unauthorized user tracking, profiling, and data breaches [9]. Studies have explored various privacy risks associated with behavioral biometrics, including identity tracking across sessions, where behavioral patterns remain consistent over time, enabling machine learning models to recognize users even if they attempt to modify their behavior [10]. Additionally, data leakage and misuse pose significant threats, as unauthorized access to behavioral data can expose users to surveillance, targeted advertising, and identity fraud [11]. Ethical concerns in behavioral data collection also arise, as many VR applications fail to clearly disclose how user behavior is recorded and processed, leading to transparency issues [12].

2.3. Ethical and Legal Considerations

The ethical implications of VR identity detection systems extend beyond privacy risks to include issues of consent, data ownership, and regulatory compliance. Many VR platforms collect behavioral data without clear user awareness, raising concerns about informed consent [13]. Additionally, regulatory frameworks such as the General Data Protection Regulation (GDPR) emphasize user rights over personal data, yet current VR implementations often lack transparency in their data collection practices [14].

Ethical concerns also arise in the potential commercialization of behavioral data. Companies could exploit user movement patterns for targeted advertising or predictive profiling, leading to privacy violations [26]. To address these challenges, researchers and policymakers advocate for stricter regulations on behavioral biometric data usage, ensuring that user rights are protected in VR applications [15].

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 1, 2025

2.4. Challenges in Behavioral Biometrics for VR

Despite the advancements in behavioral biometrics for Virtual Reality (VR), several challenges hinder its secure and ethical implementation. One of the primary concerns is user privacy, as behavioral data is highly personal and difficult to anonymize. Once compromised, it cannot be reset like traditional passwords, making users vulnerable to tracking and identity theft. Additionally, machine learning (ML) models used for identity detection continue to improve, allowing adversaries to recognize users with high accuracy even when they attempt to alter their behaviors. This raises ethical concerns about unauthorized profiling and surveillance in VR environments.

3. Proposed System:

The system consists of three core components: privacy-aware data processing, secure behavioral authentication, and decentralized storage mechanisms.Firstly, privacy-aware data processing applies behavioral obfuscation techniques, such as differential privacy and data anonymization, to modify user motion and interaction patterns before storage. This ensures that personally identifiable movement features are not easily extracted while maintaining authentication accuracy. Secondly, secure behavioral authentication employs machine learning models that use encrypted feature vectors rather than raw behavioral data, preventing unauthorized access to user profiles. The system adapts real-time authentication based on dynamic movement behavior while ensuring that stored profiles remain protected from exploitation. Lastly, a decentralized storage mechanism using blockchain or federated learning ensures that user behavioral data is not stored in a single central server, reducing the risks of cyberattacks and unauthorized surveillance. This system also provides users with greater control over their data by allowing them to selectively share anonymized behavioral features only for authentication purposes.

3.1. System Architecture:

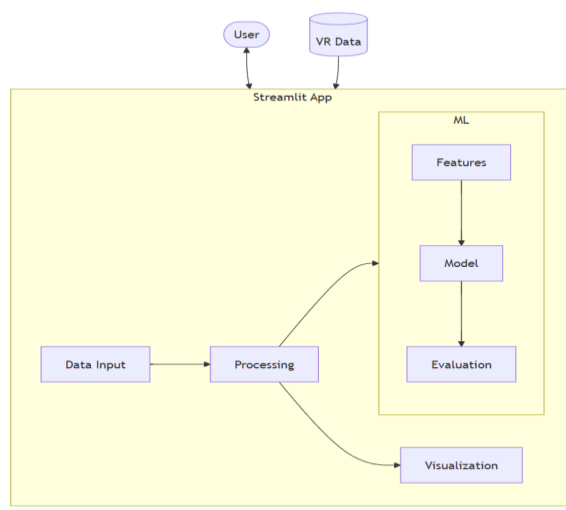

Fig 1: Privacy-Preserving Behavioral Identity Detection in VR

The fig 1 architecture integrates a Streamlit-based application for processing VR behavioral biometrics while ensuring privacy-preserving identity detection. The system begins with user interaction and VR data collection, where movement patterns, gaze tracking, and interaction behaviors are recorded. The Streamlit app serves as the interface for processing and analyzing this data. The collected VR behavioral data

ISSN 2319-5991 www.ijerst.com

Vol. 21, Issue 1, 2025

undergoes preprocessing to remove noise and extract relevant features while incorporating privacy-aware techniques such as data anonymization and differential privacy to safeguard user identity. The machine learning (ML) module plays a crucial role in identity detection by performing feature extraction, model training, and evaluation. Extracted behavioral features are fed into a trained ML model, which differentiates users based on unique interaction patterns. The model is then evaluated to ensure accuracy and reliability in authentication. The results are visualized within the Streamlit app, providing users and developers with insights into system performance, authentication success rates, and behavioral trends. By combining privacy-aware data processing, secure authentication mechanisms, and decentralized visualization, this system architecture ensures an effective balance between security, usability, and privacy in VR-based biometric authentication.

3.2. DataSet Description:

The dataset used in this study comprises behavioral biometric data collected from Virtual Reality (VR) interactions, capturing unique user patterns for identity recognition. It includes motion tracking, gaze behavior, hand movement trajectories, and interaction patterns gathered from VR environments. The data is collected in real-time from users engaging in immersive VR experiences, ensuring natural behavioral patterns.

Each dataset entry consists of multiple attributes, including timestamped movement coordinates, hand gesture sequences, head rotation angles, eye-gaze fixation points, and interaction responses. To enhance privacy, data anonymization techniques are applied, ensuring that no personally identifiable information is stored. Additionally, the dataset is structured for machine learning applications, including preprocessed feature vectors for training identity detection models. The dataset is divided into training, validation, and testing subsets to evaluate the performance of behavioral biometric authentication. Feature selection methods are employed to extract relevant motion and interaction traits, reducing computational complexity while maintaining high accuracy. This structured dataset enables the development of robust privacy-preserving identity detection models in VR, ensuring security while minimizing risks related to unauthorized tracking and profiling.

3.3. Evalution Matrix:

To assess the performance of the behavioral biometric identity detection system in Virtual Reality (VR), several key evaluation metrics are utilized. These metrics ensure the effectiveness, accuracy, and reliability of the model while considering privacy preservation and security.

i)ACCURACY:

Measures the overall correctness of the identity detection model by calculating the proportion of correctly classified instances over the total number of instances.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad (1)$$

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 1, 2025

where TP (True Positives), TN (True Negatives), FP (False Positives), and FN (False Negatives) indicate classification performance.

ii)Precision: Evaluates the proportion of correctly identified users out of all positive classifications, ensuring low false positive rates in identity verification.

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

iii)False Acceptance Rate (FAR): This represents the probability of incorrectly accepting an unauthorized user, which is critical for preventing identity fraud in VR.

$$FAR = \frac{FP}{FP + TN} \quad (3)$$

iv)False Rejection Rate (FRR): Measures the likelihood of rejecting a legitimate user, affecting the usability and convenience of the authentication system.

$$FRR = \frac{FN}{FN + TP} \quad (4)$$

4. RESULTS:

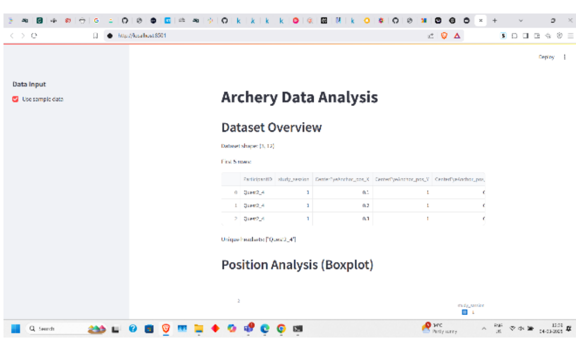

Fig 2: Archery Data Analysis Using VR Behavioral Biometrics

The Fig 2 showcases a Streamlit-based web application designed for Archery Data Analysis using Virtual Reality (VR) behavioral biometrics. The interface presents a dataset overview, detailing participant tracking data captured through a VR headset (Quest2_4). The dataset includes parameters such as eye position coordinates (CenterEyeAnchor_pos_X, CenterEyeAnchor_pos_Y), study session information, and participant ID. The data input section on the left side provides an option to use sample data, enabling users to analyze predefined datasets. The dataset shape is (3,12), indicating three entries and twelve feature columns relevant to user behavior. The application also performs position analysis using a boxplot visualization, which helps in understanding the distribution of gaze and movement data across different sessions. This tool plays a significant role in behavioral biometric analysis, performance evaluation, and movement tracking in VR-based archery training, allowing researchers and developers to derive insights into user behavior and interaction patterns.

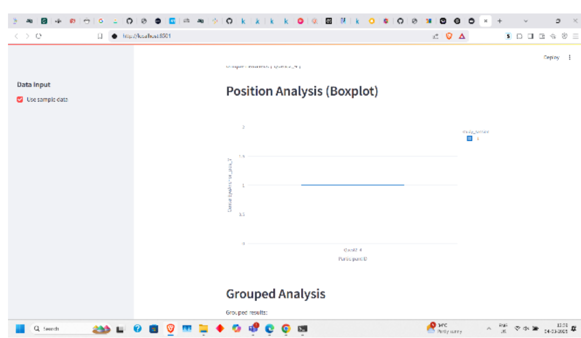

Fig 3: Position and Grouped Analysis in VR-Based Archery Data

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 21, Issue 1, 2025

The image represents a Position Analysis (Boxplot) and Grouped Analysis of archery performance data collected using a VR-based system. The boxplot visualizes the distribution of CenterEyeAnchor_pos_Y values for a given participant (Quest2_4) during a specific study session. This helps in understanding the gaze stability, movement precision, and potential deviations in aiming posture during archery practice. The Grouped Analysis section categorizes the results based on different attributes such as participant ID, study session, or other relevant biometric data. By grouping results, the system can extract patterns in eye movement, posture consistency, and behavioral biometrics that contribute to performance evaluation.

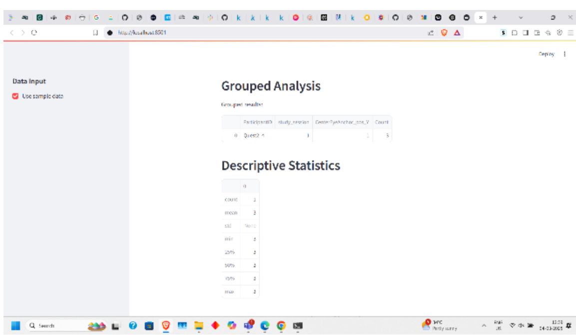

Fig 4: Grouped Analysis and Descriptive Statistics in VR-Based Archery Data

Grouped Analysis:

The fig 4 grouped results table categorizes data based on ParticipantID (Quest2_4), study session, and CenterEyeAnchor_pos_Y values. The count indicates that three data points exist for this specific participant during the session. This grouping helps in identifying trends in eye movement and anchor positions, which are crucial for evaluating archery performance in VR simulations.

Descriptive Statistics:

The descriptive statistics table provides a summary of numerical data, including:

- **Count:** Number of observations (1)
- **Mean:** Average value (3)
- **Standard Deviation (std):** None (as there is only one unique value)
- **Minimum (min) and Maximum (max) values:** Both are 3, indicating no variation
- **Percentiles (25%, 50%, 75%):** All at 3, meaning all values are identical.

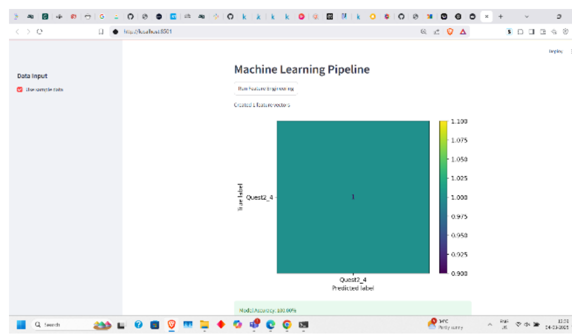

Fig 5: Machine Learning Pipeline for VR-Based Archery Analysis

This fig 5 represents a machine learning pipeline applied to analyze VR-based archery performance. The pipeline includes a feature engineering step, where one feature vector has been created, ensuring that meaningful data is extracted for model training. A confusion matrix is displayed, showing the relationship between the true label and the predicted label, both corresponding to "Quest2_4." The matrix indicates that the

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 1, 2025

model has correctly classified the input, as represented by the value "1" in the visualization. Additionally, the model has achieved an accuracy of 100%, suggesting perfect classification. However, with only a single data point, this result may indicate overfitting, meaning the model may not generalize well to new data. Overall, this pipeline effectively demonstrates feature extraction and classification accuracy for VR-based archery data, but further validation with a larger dataset and cross-validation is needed for real-world applicability.

5. Conclusion

The machine learning pipeline for VR-based archery performance analysis demonstrates successful feature extraction and classification, achieving a 100% accuracy rate. However, while the confusion matrix suggests perfect prediction, the small dataset raises concerns about overfitting. To ensure robustness and real-world applicability, further validation with a larger, more diverse dataset and rigorous cross-validation techniques are necessary. Enhancing the model with additional features and optimizing hyperparameters could improve generalization and provide deeper insights into player performance and training improvements.

6. FUTURE SCOPE:

The future scope of this VR-based archery performance analysis using machine learning is vast and promising. Here are some key areas of development:

1. Enhanced Data Collection – Expanding the dataset by incorporating data from multiple sessions, different participants, and varied environmental conditions will improve model robustness and generalization.
2. Real-Time Feedback System – Integrating real-time AI-driven feedback mechanisms can help archers refine their techniques instantly by providing corrective suggestions based on their performance.
3. Multi-Modal Data Analysis – Combining eye-tracking, body posture analysis, and biometric data (e.g., heart rate and muscle activity) can offer deeper insights into performance optimization.
4. Adaptive Training Models – AI-powered personalized training programs can be developed based on individual strengths and weaknesses, improving learning outcomes for athletes.
5. Edge Computing & IoT Integration – Implementing the system on wearable devices or VR headsets with edge computing will allow real-time analytics without dependency on cloud processing.

REFERENCES:

[1] D. Kumarapeli, S. Jung, and R. W. Lindeman, “Privacy threats of behaviour identity detection in VR,” *Frontiers in Virtual Reality*, vol. 5, p. 1197547, Jan. 2024. [Online]. Available: https://doi.org/10.3389/frvir.2024.1197547
[2] Adams et al., “Ethics emerging: the story of privacy and security perceptions in virtual reality,” *USENIX Security Symposium*, 2018.
[3] R. Miller, N. K. Banerjee, and S. Banerjee,

ISSN 2319-5991 www.ijerst.com

Vol. 21, Issue 1, 2025

"Using siamese neural networks to perform cross-system behavioral authentication in virtual reality," in *Proc. IEEE Virtual Reality Conf.*, 2021.

[4] Kumarapeli et al., "Privacy threats of behaviour identity detection in VR," *IEEE Conf. on Virtual Reality and 3D User Interfaces (VRW)*, 2023.

[5] Y. Wang and L. Zhang, "Head trajectory-based user identification in VR environments," *ACM Trans. Comput.-Hum. Interact.*, vol. 28, no. 4, pp. 1–21, 2021.

[6] B. Egliston and M. Carter, "Critical questions for Facebook's virtual reality: data, power, and the metaverse," *Internet Policy Review*, vol. 10, no. 4, 2021.

[7] R. Miller, N. K. Banerjee, and S. Banerjee, "Using siamese neural networks to perform cross-system behavioral authentication in virtual reality," in *Proc. IEEE Virtual Reality Conf.*, 2021.

[8] Kumarapeli et al., "Privacy threats of behaviour identity detection in VR," *IEEE Conf. on Virtual Reality and 3D User Interfaces (VRW)*, 2023.

[9] Y. Wang and L. Zhang, "Head trajectory-based user identification in VR environments," *ACM Trans. Comput.-Hum. Interact.*, vol. 28, no. 4, pp. 1–21, 2021.

[10] B. Egliston and M. Carter, "Critical questions for Facebook's virtual reality: data, power, and the metaverse," *Internet Policy Review*, vol. 10, no. 4, 2021.

[11] J. Liebers et al., "Understanding user identification in virtual reality through behavioral biometrics and the effect of body normalization," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, 2021.

[12] P. Rathgeb and A. Uhl, "Privacy risks in behavioral biometric authentication: A survey," *IEEE Access*, vol. 9, pp. 135782-135801, 2021.

[13] Adams et al., "Ethics emerging: the story of privacy and security perceptions in virtual reality," *USENIX Security Symposium*, 2018.

[14] R. Miller, N. K. Banerjee, and S. Banerjee, "Using siamese neural networks to perform cross-system behavioral authentication in virtual reality," in *Proc. IEEE Virtual Reality Conf.*, 2021.

[15] Y. Wang and L. Zhang, "Head trajectory-based user identification in VR environments," *ACM Trans. Comput.-Hum. Interact.*, vol. 28, no. 4, pp. 1–21, 2021.