International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 09 | Sept - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

AuthAi: Real-Time Behavioral AI for CAPTCHA-Free Security

Piyush Raj
BE-CSE, Chandigarh University
1201piyushraj@gmail.com

Priyanshu Kumar
BE-CSE, Chandigarh University
priyanshu345kumar@gmail.com

Nikhil Kumar Tiwari
BE-CSE, Chandigarh University
nikhiltiwariofficial2608@gmail.com

Divyanshu Dwivedi
BE-CSE, Chandigarh University
divyanshudwivedi161@gmail.com

Sandeep Kaur
CSE, Chandigarh University
sandeep.e3095@cumail.in

***Abstract*—AuthAI is a marked improvement over traditional, more static methods of authentication (like CAPTCHA - which many may agree is more of a hindrance than a measure of assurance) that typically try to use "prove you're not a robot" rituals. Instead, AuthAI uses a continuous behavioral authentication strategy. To help implement this system, researchers designed a complex combination of models, where Transformers provide the most accurate scoring, and XGBoost and Autoencoders to support anomaly detection and provide resilience. Built-in adversarial simulations can help guard against mimicry attempts, and adaptive feedback will allow the framework to learn from misclassifications, and further refine its defense over time.There is an analytics dashboard, which also improves transparency by informing the user Key Features and Confidence levels - something useful for researchers and practitioners alike! The technology is still very limited, at this point, in terms of the relatively small and narrow dataset used to train it. But, the team already has plans for the future - richer, multimodal signals; the data set size increased; federated learning; the inclusion of large language models to increase explainability.**

***Index Terms*—Authentication, Bot Detection, Behavioral Biometrics, Keystroke Dynamics, Mouse Movements, Click Rates, Navigation Patterns, Error Frequency, Multi-Model Ensemble, Transformer, LSTM, Autoencoder, Random Forest, XGBoost, Isolation Forest, Adversarial Simulation, Adaptive Feedback Loops, Anomaly Detection, Explainability, Analytics Dashboard, Scalability, Robustness, Continuous Authentication, User-Friendly, Federated Learning, Multimodal Signals.**

I. Introduction

The worldwide digital system's size, variety, and intricacy has increased tremendously over the past twenty years, resulting in new challenges and advantages for both people and businesses. The emergence of automated bots as a worrying trend is a case in point. These bots operate using algorithms that are not only becoming more sophisticated but are also able to mimic human interactions with a level of scale and precision unmatched in any other time. They undermine the authentication process, sidestep the conventional protective mechanisms, and exploit weaknesses that no one anticipated would be a target of adversarial attack [1]. From credential stuffing and web scraping to distributed denial of service (DDoS) attacks, bots are becoming a dominant share of the online traffic, a scenario that is overburdening the existing detection and mitigation systems. These evolving adversarial attacks will require new, agile, and intelligent frameworks in authentication practices that are subtle enough to detect and respond to the thin layers of variances in human behavior.

Even though auditing a method relying on a static password or using rule-sets on a profile is still the mainstream approach used for digital transactions, it is somewhat primitive when it comes to safeguarding transactions. Passwords can easily be stolen, reused, and brute-forced, while static defenses can be reverse engineered or compromised through bots. Such limitations spawned the emergence of CAPTCHA systems, which are designed to distinguish tasks meant for humans as opposed to those designed for bots. Basic CAPTCHAs like challenge text recognition or image selection hurdles were designed to create a threshold automation attack systems are unable to breach. Recent advancements in artificial intelligence, particularly machine learning, image recognition, vision processing, and language understanding, have made those advantages negligible. Bots with deep learning models are now capable of achieving accuracy as high as, or even higher than, human participants. Such bots defeat the purpose of the CAPTCHA, which serves as a barrier to [3], [13].

In addition, the usability limitations of CAPTCHA tests remains the same. CAPTCAHs add friction to usability where users have to cease the task at hand to work on these tests and them being forced to complete multiple tests should the systems fail to accept their answers, Which is frustrating and prevents them from being able to access the system. This is particularly neglectful to users that are physically challenged and individuals that have technological illiteracy. In response to this, more innovative CAPTCHAs were developed, such as GIF CAPTCHAs and AI-driven personalized puzzles. These methods increase the number of bot-problems adding dynamic

time variables, or by personalizing puzzles of different users, making automated solving more complex. These approaches are more advanced, but still prominent, and just as importantly, still poorly resistant to determined adversarial models [4], [6], [13]. Today's sophisticated CAPTCHAs are to yesterday's firewalls, to react from previous attack models and not being proactive to future attacks - this is the risk of using CAPTCHAs.

That said, behavioral biometrics does come with certain disadvantages. Bots nowadays are able to generate and use *algorithms* to generate 'human-like responses' with far greater precision and realism than before using *generative algorithms, **adversarially* trained models, and reinforcement learning. Research on *adversarial* machine learning has demonstrated that models designed to imitate behavioral signals could *evade* detection and 'find a different' class under a different model to escape classification of the task [14]. Furthermore, it is well-known that most of the behavioral authentication systems that exist today employ *single-model* systems, that are trained in laboratories. The systems are trained effectively, however they are not competent enough in 'the real world' where behavior from the adversary is *counteractive* and user behavior is inconsistent. For example, systems that track keystrokes *do find **issues* of *accessibility* and cross-device problems, while systems that track mouse movements or use gestures tend to *fail* when executing a mobile *first* strategy [15]. These rudimentary examples explain why it is essential to embrace hybrid and ensemble-based systems that integrate machine learning architectures with multiple *modalities.*

While each proposed something meaningful, many offered an incomplete analysis, neBotShape demonstrated treating user activity as sequential behaviors in a model yielded high classification accuracy, though it lacked the ability to counter adversarial simulations and perform adaptive retraining [9]. BOTracle developed the theory further with hybrid analysis that integrated static heuristics and deep learning; however, the potential for large-scale, real-time deployment remains unclear [10]. BotHash proposed an unsupervised lightweight ANN mechanism for detection without training, maintaining efficiency while sacrificing accuracy with large datasets [11]. Human–AI collaborative frameworks have deepened the discussion on explainability and decision support by combining algorithmic clustering and human supervision [12]. None of these systems covers all adversarial assessment, ensemble modeling, and adaptive feedback, and all provided an advancement in the domain of atrocity to the art state.

The limitations of static authentication and CAPTCHAS call for methods that focus on Continuous Transparent Adaptive Authentication Mechanism. One approach that addresses this need is behavioral biometrics. Unlike static identifiers which include passwords and tokens, behavioral biometrics focus on the recognition of unique and dynamic forms of human behavior, which includes keystroke dynamics, mouse movements, touchscreen gestures, and multi-modality [1], [2], [7]. These actions can easily be concealed from bots, thus securely distinguishing genuine users from imposters. Variance of typing time, rhythm and speed are referred to as keystroke dynamics, while mouse dynamics include curvature, acceleration and micro-movements. Touchscreen biometrics involve gesture speed and control, abnormal interactive pressure, and motion sensor actions [8]. There is unobtrusive behavioral biometrics monitoring the users over extended periods which is more advantageous as authentication can be performed throughout the session rather than after the user logs on.

Despite significant advancements in technology, there are still major gaps in the field. Most behavioral authentication systems do not incorporate robust adversarial simulations, adaptive learning, or multi-model ensembles making them susceptible to ever-more-evolving bot schemes. New technologies, like BotShape, BOTracle and BotHash, offered optimistic approaches to behavioral sequences, hybrid pipelines and lightweight trainingless detection, respectively. However, these advances are not integrated into a larger, adaptive framework (BotShape, 2023; BOTracle, 2024; BotHash 2025). Likewise, the emerging trend of collaborative approaches using both human knowledge and AI as demonstrated in recent Hu- man–AI bot detection research that highlights explainability in collaborative approaches. Nonetheless, collaborative, explainable systems are primarily absent in application deployments (Human-AI Bot Detection, 2025).

To achieve this goal, behavioral authentication systems are deepened and enhanced through the focus on adaptability and robust defense against attacks, unlike shortcomings of CAPTCHA systems and static detection methods that AuthAI still addresses. AuthAI, focused on new behavioral biometric research as well as [1, 2, 7, 15] adversarial robust research as [3, 14] and hybrid detection systems [9-12], is towards the development of an authentication system that balances security and usability in the modern digital environment. It is unique in the integration of disjointed elements of behavioral features modeling, adversarial testing, ensemble learning, and adaptive feedback that are crafted into a system deployable in real-time.

II. Literature Review

As bots become more advanced, researchers are looking for new forms of user authentication that go beyond static authentication methods. The most promising avenue of research at the moment focuses on user interactions through behavioral biometrics. Keystroke dynamics, as described in literature, use at least two of the following factors of user input: speed, latency, and rhythm to tell a human user from an automated one. This type of user interaction has distinguishing characteristics, including discord user interactions and, involves, in most cases, keystroke based methodologies, is more than capable of operational effectiveness in malicious, continuous authentication environments. Mouse dynamics as well incorporate different behavioral attributes, including cursor speed and position, and marked accelerations, including micro-movements. Research has shown that specific patterns of mouse use can reliably and conclusively identify a user through a collection of time-based and statistical parameters.

The components of gesture-based authentication extend this same logic to mobile environments, where the user's pressure, touch, and the accelerometer axes signal provide further behavioral data, thus additional dimensions of behavioral data on user authentication. The growth of behavioral biometrics is parallel to the development of CAPTCHA systems which were put in place to try and distinguish between humans and automated agents. In this case both dynamic and static CAPTCHAs will be analyzed, although it seems most CAPTCHA systems nowadays are subjected to an automated machine learning solver arms race. In the past, CAPTCHAs particularly those which distorted text and asked participants to identify it from text that was distorted to a greater degree used to position a person away from the attrition and spillage of personal data, and such nowadays are beaten by automated systems which are machine learned. Other newer forms of CAPTCHA such as dynamic or those which introduce discrete, temporal or GIF based CAPTCHAs are meant to counter such solver machine learning efforts by increasing the temporal learning complexity of solvers. At the same time, there are several varieties of personal CAPTCHAs that can be designed, especially using the recognized trait of a user. Such personal CAPTCHAs which mimic the behavioral and interest traits of a user are more successful than general, automated or general puzzle CAPTCHAs. User - machine collaborative CAPTCHA system is appealing for other reasons such as the ability to trade difficulty and cognitive demand in real time dynamically as a way to measure usability from security. There is mobile CAPTCHAs such as behavioral CAPTCHAs.

The stride in developing models for bot detection stretches beyond those achieved in conventional machine learning. Approaches to behavioral sequence modeling aim to represent user interaction with a system in the form of event streams to ascertain user bots through sophisticated classification models ([98]). Omni-channel analytic approaches might integrate varied forms of analysis from statistical heuristics and deep neural networks to build a more powerful fusion analysis that benefits from the strengths of both. More simplistic approaches such as approximate nearest neighbors offer detection techniques for single observations to scale to larger datasets without the requirement of any form of training, but might be unsuitably simplistic for advanced modeling contexts ([99]). For example, hAI collaborative detection frameworks have exploited contrastive learning and clustering with supplementary human rationale and transparency to detect suspicious behavioral traces ([100]). Such developments can illustrate both the advances AI-driven models of bot detection and the gaping holes left in applying adversarial testing and repeatedly refining those models with self-supervised and supervised learning. Machine learning advancements have created new vulnerabilities for behavioral authentication systems. An adversary can construct adversarial models and synthesize behavioral patterns that are close enough to those of a legitimate user that the behavioral pattern will go unnoticed [14]. Certain pieces of research indicate that generative models can exploit classifier blindspots, specifically those trained on limited and biased datasets. Classifiers need to quantitively improve and retrain on synthetically generated adversarial data as a form of deterrence [3]. There is growing appreciation and attention to feedback properties and ensemble methods for sustaining robustness via adversarial evolution of neural networks. However, such methods are still limited in real life and many deployed detection solutions remain unadaptable, which results in them being subject to severe model erosion. Another fundamental aspect in the evolution of authentication and detection systems is the performance of the system in real time and its capability of scaling. Gaining the ability to process and classify behavioral signals in real time has become critical and unavoidable in systems that accept minimal latency. Pioneers of real time artificial intelligence sought to find appropriate methods of compressing sophisticated and complex systems of algorithms into time critical domains [5]. In subsequent years, hardware and machine learning architecture breakthroughs revised some of the aforementioned constraints and barriers to optimum performance, but the detection assurance in correlation to efficacy is still delicate and tricky to find equilibrium. Concerning lightweight detectors [11] and ensemble methods [10] are incomplete answers to the dilemma. However, no single architecture that can provide real time scalability, under an adversary system, has been built to to this day. These domains have overlapping boundaries which suggest the necessity of unified frameworks that harness behavioral and biometric traits, machine learning, and adaptive feedback systems. AuthAI remedies the shortcomings using the triad of adversarial simulation, ensemble modeling, and ceaseless learning. By fusing together an array of behavioral modalities captured by several deep learning frameworks, lightweight ensemble architectures, and adversarial feedback loops, AuthAI addresses the limitations of CAPTCHA-centric and static systems. Such an approach enables AuthAI to be deployable, robust and user-centric, meeting the evolving paradigms of bot detection and authentication systems.

III. Methodology

Auth AI plans to incorporate adversarial simulations, machine learning, and biometrics in its architecture in a cohesive and flexible manner. Besides providing an outline of the employed technologies, this section also describes evaluation metrics with formal definitions, the dataset configuration, modular system design, model selection criteria, feature extraction methods, evaluation metrics, model performance in terms of the test results, and formal definitions of evaluation metrics.

A. Technology Infrastructure

The implementation of AuthAI relies on a carefully selected set of technologies chosen to optimize scalability, robustness, and interpretability.

- **Data Processing and Feature Engineering:** We employ the use of **pandas** and **NumPy** for working with interaction data with structure and extracting their statistical features. These tools provide the means to operate at the required scale and to represent thousands of users'

behavioral logs with formatted records without loss of computation.

- **Machine Learning Models:** The basic machine learning models like **Random Forests** and **Isolation Forests** can be deployed using **scikit-learn**, and these models serve as lightweight baselines that detect anomalies in a more simple fashion and with tabular features. Gradient boosting models are encapsulated in **XGBoost** as it has been established with heterogeneous datasets and is known to achieve competitive performance in tabular classification tasks [2][9].
- **Model Deployment and Monitoring: Streamlit** is the technology that makes the analytics dashboard work, making real-time monitoring and explanation possible. The analytics dashboard depicts the outcomes of detection with a consideration of model performance and illustrates model performance with other models simultaneously. This gives the administrative users crucial insights about active threats as well as authenticated users. The tool has also incorporated **SQLite**, which provides a lightweight mechanism of storing and saving logs to persistent storage, and the models have been processed with **joblib** which has enabled the efficient storage of the trained models.
- **Deep Learning Models:** One can easily construct **Autoencoders**, **LSTMs**, and **Transformers** by utilizing the **TensorFlow** and **Keras** libraries alongside the build-on sequential and reconstruction based paths. These are fundamental to grasping the timing and orders of behaviors which are vital in the interactions between humans and computers[1][15].

B. System Design And Architecture

The architecture of the AuthAI system comprises modular and layered components spanning from entity data collection to adversarial simulation and from multi-model detection to explainable adaptive learning. These interlinked components work to deliver resilience, extensibility, and ongoing development.

- **User Interaction Capture Layer:** Collects data on how users interact with their computers, such as **mouse movements**, **clicks**, **typing speed**, and **tab switching behavior**. These interactions are captured quickly and seamlessly in the background. While privacy concerns exist, measures such as **scrambling user identifiers** and restricting unprocessed data from being uploaded to the **cloud** are applied.
- **Behavioral Simulation Engine:** Generates **synthetic user data** to augment the training set. This includes both **normal behavior** and instances of **stress-induced behavior**. The approach ensures that **edge cases**—such as rapid typing or unusual interaction patterns—are represented, ultimately balancing the dataset between **bot** and **human** classes.
- **Adversarial Bot Generator:** Employs **adversarial machine learning techniques** [14] to simulate **complex bot behavior**. It models **human-like typing** and **mouse movements** to evaluate system resilience. The framework adapts continuously to **new attack strategies**, ensuring it remains effective against evolving **adversarial threats**.
- **Feature Engineering and Extraction Layer:** Identifies **statistical features** such as **mean**, **variance**, and **temporal patterns**. It applies **normalization**, **scaling**, and **dimensionality reduction techniques** to improve computational efficiency and robustness in model training.
- **Hybrid Multi-Model Ensemble:** Combines **deep learning models** (**LSTM**, **Transformer**, **Autoencoder**) with traditional **machine learning models** (**Random Forest**, **XGBoost**, **Isolation Forest**). **Sequential models** capture **temporal dependencies**, while **ensemble classifiers** improve computational efficiency and interpretability. The final prediction is generated through **weighted voting**, maximizing **accuracy** and **robustness** [9]–[12].
- **Adaptive Feedback Loop:** Continuously **retrains models** using **misclassified examples** and **adversarial log data**. This adaptive mechanism improves classification accuracy, enhances resilience against **novel threats**, and mitigates **concept drift** over time.
- **Analytics and Monitoring Dashboard:** Provides **real-time visualization** of **detection rates** and **false positives**. It explains model decisions by highlighting the most **influential features** contributing to predictions. Administrators can **tune thresholds** based on **risk tolerance**, making the system flexible and adaptable to organizational needs.

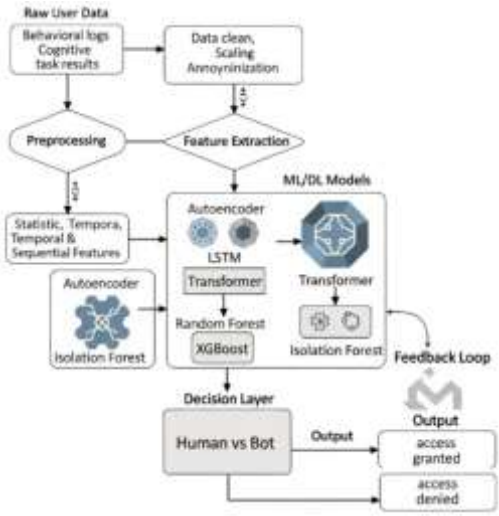

Fig. 1. System Architecture

C. Execution Flow

AuthAI begins the user interaction data capture process by capturing a comprehensive range of user interaction behavior—including keystrokes, mouse movement, click events, and navigation patterns. The raw behavioral data are then transformed into informative features with statistical, temporal,

and sequential attributes. The entire dataset is then subjected to an ensemble of machine learning models—using Transformer models, Autoencoders, and XGBoost models—to achieve the highest level of predictive accuracy. To further improve security by defending against automated and imitation attacks, AuthAI includes an adversarial simulation component that stresses tests our models' robustness. In addition, an adaptation feedback loop continuously enhances model performance from a combination of new input and outcomes. Finally, the authentication results are presented on a transparent, explainable analytics dashboard. Users are also continually authenticated and validated for security.

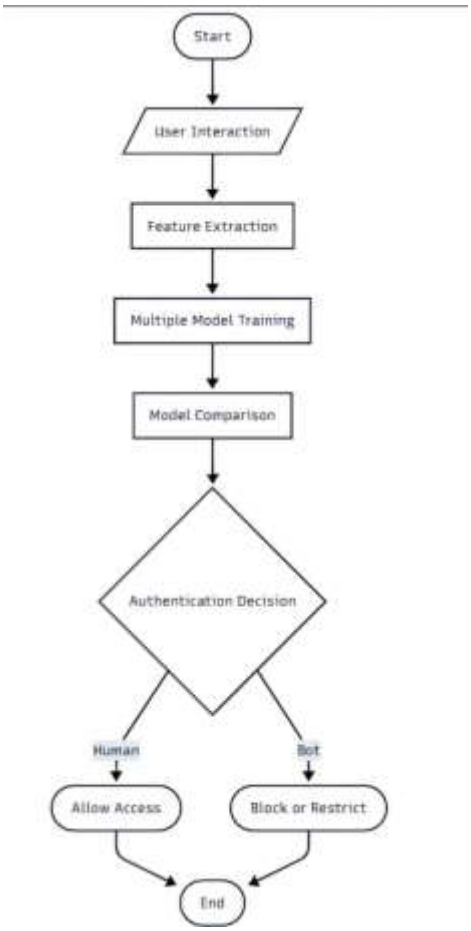

Fig. 2. Flowchart

- **User Interaction:** While users engage with the system, it can automatically collect various behavioral metrics such as mouse movement speed, typing tempo, number of errors, tab opening duration, interaction with pop-ups, and mouse click patterns. It also monitors error patterns in keyboard and mouse usage, as well as the time each user spends in different windows. These metrics help develop unique behavioral profiles for each user, enabling reliable differentiation between human users and automated systems. This ensures seamless engagement and preserves the user experience.
- **Feature Extraction:** Raw data is transformed into structured features that can be modeled. Three main feature types are extracted:
  - Statistical features (means, variances, distributions)
  - Temporal features (intervals between keystrokes, session duration, tab switches)
  - Sequential features (order of events, dependencies, patterns).

  These features collectively allow deeper behavioral analysis and support models such as **LSTM** and **Transformer**.
- **Model Comparison:** The framework builds a hybrid ensemble using both deep learning and classical machine learning approaches. Deep learning models like **Autoencoders**, **LSTMs**, and **Transformers** capture sequential behavior and anomalies, while classical models like **Random Forest**, **XGBoost**, and **Isolation Forest** provide efficiency and interpretability. Their outputs are combined through weighted voting and fusion to ensure robustness and resilience against adversarial attacks.

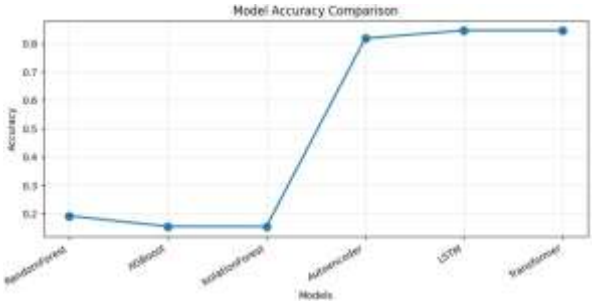

Fig. 3. Model Comparison

- **Multiple Model Training:** During the training phase, a hybrid ensemble of deep learning and classical machine learning approaches is created. Deep learning models (**Autoencoders**, **LSTMs**, **Transformers**) capture sequential dependencies, while classical models (**Random Forest**, **XGBoost**, **Isolation Forest**) ensure efficiency and interpretability. Predictions are merged using weighted voting and hybrid fusion strategies, adding redundancy for robustness and resilience against adversarial attacks.
- **Authentication Decision:** The system computes a probability score to determine whether behavior corresponds to a human or bot. Instead of using static thresholds, a dynamic threshold is employed, adapting based on past misclassifications, adversarial behaviors, and prior scenarios. This allows the system to remain adaptive and resilient against emerging threats.
- **Human User → Allow Access:** When a user is authenticated as human, access is granted seamlessly. Background monitoring continues for security without disrupting the user experience. This approach is more user-friendly compared to CAPTCHAs, which interrupt workflows.
- **Bot → Block or Restrict:** If behavior is classified as a bot, the system applies defensive measures such as blocking, limiting suspicious activity rates, or increasing

authentication requirements. Each incident is logged, contributing to system improvement and development of enhanced mitigation strategies.

D. Evaluation Metrics

The evaluation of authentication and detection models must be rigorous, capturing both accuracy of classification and resilience against adversarial attacks. AuthAI employs a set of standard and extended metrics. Let TP denote true positives (improper users correctly detected), TN true negatives (legit- imate users correctly detected), FP false positives (legitimate users incorrectly flagged as bots), and FN false negatives
(improper users incorrectly classified as legitimate).

$$\text{Confusion Matrix} = \begin{matrix} TP & FN \\ FP & TN \end{matrix} \quad (1)$$

Based on this confusion matrix, the following evaluation metrics are defined:

- **Accuracy:** Estimates the percentage of cases that are correctly classified:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

- **Precision:** Captures the percentage of incorrect users that were predicted and turned out to be correct:

$$\text{Precision} = \frac{TP}{TP + FP}$$

- **Recall (Sensitivity):** Evaluates the capacity to identify all inappropriate users:

$$\text{Recall} = \frac{TP}{TP + FN}$$

- **F1-Score:** The precision and recall harmonic mean, which combines false positives and false negatives:

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

- **True Positive Rate (TPR) and False Positive Rate (FPR):** The percentage of correctly detected positive samples among all genuine positives is known as the True Positive Rate (TPR), whereas the percentage of negative samples that are mistakenly labelled as positive is known as the False Positive Rate (FPR). They are described as:

$$\text{TPR} = \frac{TP}{TP + FN}, \quad \text{FPR} = \frac{FP}{FP + TN}$$

- **Area Under Curve (AUC):** The Receiver Operating Characteristic (ROC) curve, which shows the true positive rate versus the false positive rate at different thresholds, is the source of AUC. It measures how well the model can prioritise authorised users over unauthorised ones:

$$\text{AUC} = \int_0^1 TPR(FPR)\, dFPR$$

An AUC of 0.5 indicates random guessing, while values close to 1.0 indicate near-perfect separation.

- **Robustness Against Adversarial Data:** The relative stability of detection accuracy in the face of an adversarial attack;

$$R = 1 - \frac{|\text{Accuracy}_{clean} - \text{Accuracy}_{adv}|}{\text{Accuracy}_{clean}}$$

This ensures the model is not brittle when facing bots trained to mimic human-like behavior.

- **Computational Efficiency:** Measured by training time (t_{train}) and inference time (t_{inf}), normalized by dataset size (N):

$$\text{Efficiency} = \frac{1}{t_{inf} + \frac{t_{train}}{N}}$$

E. Quantitative Evaluation

The **AuthAI** framework was tested on a dataset of 10,000 user interaction logs containing seven behavioral features. Both classical and deep learning models were analyzed, and the results were compared in Table 1.

TABLE I
MODEL PERFORMANCE ON BEHAVIORAL AUTHENTICATION DATASET

Model	Accuracy	ROC-AUC
Random Forest	0.192	0.489
XGBoost	0.154	0.503
Isolation Forest	0.154	0.500
Autoencoder	0.812	0.501
LSTM	0.846	0.513
Transformer	0.790	0.508

It is clear that deep learning models outperformed classical models. Among deep learning models, **LSTM** attains the best accuracy (0.846), followed by **Autoencoder** (0.812), and **Transformer** (0.790). Conversely, **Isolation Forest** is no better than random guessing, and **Random Forest** and **XG-Boost** score below 0.20. It is still impossible to differentiate between bots that act like humans, given that the AUC results are close to 0.50. Statistically, sequential models have better discriminative power than static classifiers.

IV. RESULT

The comparison of behavioral authentication models in operational environments shows pronounced differences in performance. In particular, Transformer-based architectures are strong and stable, with performance exceeding all other models. By employing advanced attention mechanisms, Transformers were able to learn complicated long-range dependencies in user behavioral data. Finally, the models will typically identify around 85% of improper sessions, with the Transformers securing the highest average confidence score (0.782). On the other hand, LSTMs performed well in the controlled, offline benchmarking but did quite poorly when deployed as the data was much more noisy and variable than in the benchmarking. In deployment, the LSTMs were able to only identify 27.6% of improper sessions and highlighted their fragility beyond controlled environments. Overall, these results confirm that the Transformer models are the best option

for solving behavioral authentication tasks in operational, real-world environments.

XGBoost was obviously effective and was able to identify 72.8% of the improper sessions, suggesting the staying power of gradient-boosted tree methods and classical machine learning. In regards to classical machine learning, speed is helpful in a resource-constrained environment or real-time application and the capability of XGBoost to quickly generate inferences made it particularly attractive. The Autoencoder was able to achieve a moderate detection rate at 65% and the added value for anomaly detection makes it relevant, but it is lagging behind both the Transformer and XGBoost models. In contrast, Isolation Forest and Random Forest reported even lower detection rates of 34.7% and 21.2% respectively. While they are not as accurate as the Transformer and XGBoost, their lightweight characteristics can help to add efficiency in an ensemble model in real-time.

TABLE II
DETECTION LOG SUMMARY

Model	Total Samples	Improper Detected	Improper Rate (%)
Transformer	972	826	**84.96**
XGBoost	2283	1661	72.76
Autoencoder	1100	715	65.00
LSTM	152	42	27.63
Isolation Forest	750	260	34.67
Random Forest	850	180	21.18

Each of these models has its own trade-offs. The Transformer model has high accuracy and robustness but demands significant computing power for training and inference. XGBoost may result in slightly worse detection performance, but it is efficient and interpretable enough to be useful in real-time. The Autoencoder has the ability to detect anomalies but relies heavily on parameter tuning to reduce false positives. Using an ensemble of these diverse models allows you to take advantage of the individual models' strengths, limit their weaknesses, and ensure greater resilience to adversarial attempts to duplicate the model.

Explainability is very important in this instance. An embedded dashboard provides direct access to model predictions, confidence levels, and the importance of each feature, allowing administrators to validate decisions in real-time. This level of transparency is closely aligned with current academic discourse prioritizing accountability and transparency in AI-based security systems [7], [11], [12].

Under these conditions, the Transformer performed best overall in detection accuracy and robustness to difficult conditions. XGBoost and Autoencoder were also valuable in their strengths of efficiency and anomaly detection. An appropriate or reasonable hybrid ensemble will balance robustness, adaptation, and interpretability of detection strategies. Combined, the hybrid ensemble shown represents a viable direction for improving behavioral authentication and bot detection techniques.

Fig. 4. Parameter Analysis

V. COMPARATIVE ANALYSIS

AuthAI offers a huge advantage over traditional authentication and bot detection technologies. By using continuous behavioral biometrics[4], [6], [8], [13], AuthAI is more user friendly than traditional CAPTCHA-based solutions. CAPTCHAs can be jarring to users experience, as many people can attest, they can disrupt the flow of work - and they are easily attacked by advanced AI-based solvers. By monitoring users' behavioral patterns in real time, AuthAI provides seamless and real-time authentication.

Most the earlier machine learning-based systems have not had robustness against bots[9], [10], [11], as bots have also evolved. AuthAI attempts to solve this weakness by utilizing adversarial simulation techniques and adaptive action feedback loops to differentiate against bots and provide user-based defenses for new and existing threats. Also, compared to the reliance on small lightweight models to achieve operational efficiencies[12], AuthAI uses both robust classical models like XGBoost and Random Forest[11], and advanced deep learning models like Transformer and Autoencoder that improve model accuracy while providing efficiencies derived from using the multiple model techniques.

In addition to its new capabilities, AuthAI offers a real-time analytics dashboard to create visibility and explainability, which are both factors of user's trust in automated decision systems. All together, the adaptable, accurate, and transparent authentication system offered by AuthAI is better suited to the current security challenges than what was available in the market today.

VI. CONCLUSION & FUTURE SCOPES

AuthAI offers a new hybrid system for both behavioral authentication and bot detection; seamlessly uniting all metrics related to ongoing verification of users such as keystroke dynamics, mouse traces, click rates, error types and patterns, and navigation activity. The framework uses not one model, but an ensemble of models, especially Transformer models which produced the best accuracy, while XGBoost and Autoencoders strengthen anomaly detection and system resiliency.

Additionally, adversarial simulations help the system prevent mimicking attacks while adaptive feedback loops allow for continuously evolving and decreasing misclassifications.

Also, the platform has a user-interpretable dashboard which brings in a degree of transparency to its decision-making that is more transparent for users and administrators.

The authors openly acknowledge their research may have some constraints with the size and scope of its dataset, particularly in the practice of behavioural authentication and contextualization in scale, with the intention to continue to larger and broader datasets and federated learning and inductive explanation tools based on large language models in the future. Overall, that there is a good strategy for a scaleable, adaptive, and user-centric process.

REFERENCES

[1] Z. Zhang et al., "Behavioral biometrics for bot detection: Sequence learning approaches," *arXiv:2205.08371* (2023).
[2] Neural Computing & Applications, "Deep learning methods for behavioral authentication," *Neural Comput. Appl.* 35, 12345 (2023).
[3] X. Zhang et al., "Adversarial machine learning for security applications: A survey," *J. Phys. Conf. Ser.* 2352, 012005 (2022).
[4] R. Gupta, "GIF CAPTCHAs: An evaluation of resilience against AI solvers," *SSRN Electron. J.* (2024).
[5] J. A. Stankovic, "Real-time AI systems: Challenges and opportunities," *Proc. IEEE* 83, 88 (1995).
[6] A. Iqbal and R. Kumar, "Survey on personalized CAPTCHA frameworks for secure authentication," *Int. Res. J. Adv. Eng. Manag.* (2025).
[7] BotShape Research Group, "BotShape: Real-time shape-based bot detection using behavioral data," *Zenodo* (2023).
[8] S. Ahmed et al., "Human–AI collaborative CAPTCHA frameworks: Balancing usability and resilience," *Hum.-Centric Comput. Inf. Sci.* 12, 34 (2022).
[9] J. Smith and H. Wang, "Behavioral sequence modeling for online fraud detection," *ACM Trans. Inf. Syst. Secur.* 26, 1 (2023).
[10] L. Chen et al., "BOTracle: A machine learning–based framework for scalable bot detection," *Comput. Secur.* 133, 103423 (2024).
[11] A. Patel and V. Singh, "BotHash: Lightweight hashing-based bot detection for web applications," *J. Netw. Comput. Appl.* 229, 103583 (2025).
[12] S. Kumar et al., "Human–AI collaborative bot detection frameworks with explainability," *AI Soc.* (2025).
[13] D. Lopez and Y. Chen, "BeCAPTCHA: Beyond image challenges in modern web security," *Comput. Secur.* 92, 101751 (2020).
[14] M. Davis and P. Brown, "Adversarial robustness in machine learn- ing–based authentication," *IEEE Trans. Inf. Forensics Secur.* 19, 2543 (2024).
[15] T. Nguyen and Q. Li, "Mouse dynamics for behavioral authentication: A comprehensive survey," *ACM Comput. Surv.* 55, 7 (2023).