INTELLIGENT ANALYSIS OF USER BEHAVIOR IN WEB ENVIRONMENTS USING ARTIFICIAL INTELLIGENCE ALGORITHMS

Sh.Q. Shoyqulov[1], E.H. Qurbonova [2]
Associate Professor[1]
Student[2],
Department of Applied Mathematics, Karshi State university, Republic of Uzbekistan[1,2]
https://doi.org/10.5281/zenodo.15590204

Abstract. *The development of web technologies and the rapid introduction of artificial intelligence (AI) systems into the digital economy have contributed to the formation of a new stage in the field of user behavior analysis — intelligent web analytics. This article discusses the use of AI methods to analyze digital traces of users on web platforms in order to identify behavioral patterns, improve the efficiency of the user interface and personalize content. The study uses machine learning, clustering and time series analysis methods based on anonymous data from web services. The results demonstrate the high accuracy of AI algorithms in predicting user activity, which opens up new horizons for their application in e-commerce, education and digital marketing.*

Keywords: *artificial intelligence, web analytics, user behavior, machine learning, UX design, personalization, digital footprints.*

INTRODUCTION

In the modern era of digital transformation, where user data is becoming a strategic resource, user behavior analysis is becoming critical for adapting web applications to audience requirements. With the transition from classic web analytics to intelligent analytics, it becomes possible to better understand not only “what” the user does, but also “why” they do it.

Artificial intelligence (AI) allows you to process and interpret large volumes of unstructured data — clicks, cursor movements, viewing time, interaction with the interface — and build models reflecting behavioral patterns on this basis. Unlike traditional analytical systems, AI can identify hidden patterns, make forecasts, and adapt to changing conditions in real time [1].

World practice already demonstrates successful cases of using AI in web analytics. Thus, clustering algorithms are used for user segmentation, neural networks — for content personalization, and temporal models — for predicting customer churn [2], [3,13]. Despite this, there is a lack of comprehensive studies in the scientific literature focused on the use of AI specifically in the context of user behavior analytics in digital interfaces, especially in developing countries [4,12].

METHODOLOGY

The relevance of this study is determined by the need to improve the efficiency of user interaction with the web environment by introducing intelligent analytical tools. This is especially important for e-commerce, online education and government digital services, where an accurate understanding of user behavior directly affects the quality of services provided.

The purpose of this article is to analyze the possibilities of using AI to interpret the behavior of users of web platforms, identify key algorithms and evaluate the effectiveness of machine learning methods in web analytics.

To achieve this goal, the following tasks are solved:

- summarize the theoretical foundations of intelligent analysis of user behavior;
- identify the most effective AI algorithms and web analytics tools;
- conduct practical testing based on anonymized user data;
- offer recommendations for the implementation of AI in web platforms for various purposes.

The scientific novelty of the work lies in the integrated approach to the use of AI in the analysis of user behavior with an emphasis on the practical applicability of the obtained results in real digital services.

RESULTS and DISCUSSIONS

To achieve the research objectives, a systematic approach was used, combining methods of artificial intelligence, traditional web analytics and user data visualization. The methodology includes three main stages: data collection and processing, selection and training of AI models, and interpretation of the results.

The initial data were anonymized log files of user activity from the experimental website (example - an educational platform), including:

- visit timestamps;
- page view sequences (clickstream);
- average time spent on each page;
- viewing depth;
- behavioral scenarios (interaction with forms, scrolling, navigation).

The data was pre-aggregated and normalized using Python libraries (pandas, numpy) to ensure uniformity of the input data.

The following machine learning methods were used to analyze user behavior:

- Clustering (K-means, DBSCAN): to segment users by behavior patterns and build behavioral profiles [1,11].
- Decision Tree models (Decision Trees, Random Forest): to explain the key factors influencing user activity and the likelihood of return [2,10].
- Neural networks (MLPClassifier): to classify behavior types and predict the likelihood of churn.
- Temporal models (ARIMA, LSTM): to predict user activity based on browsing history [3,9].

Models were trained in the Jupyter Notebook environment using the Scikit-learn, TensorFlow, and Matplotlib libraries to visualize the results.

The following metrics were used to assess the quality of the models:

- Accuracy
- F1-score (weighted harmonic mean of precision and recall)
- Silhouette Score (for clusters)
- MAE and RMSE (for regression models of activity prediction)

To increase the reliability of the conclusions, validation was performed using the k-fold cross-validation principle with k = 5.

Since the data were anonymized and did not contain personal information, ethical and confidentiality requirements were met. The possibility of data bias was also taken into account, and therefore balanced samples were selected. Anomalous records associated with bots and random clicks (based on the speed and frequency of interaction) were excluded.

The results of the analysis demonstrate the high efficiency of using artificial intelligence algorithms to identify behavioral patterns of web platform users. The obtained data allows segmenting the audience, identifying critical points of interaction and predicting further user activity.

Using the K-means algorithm, three main user clusters were identified:

- Cluster A (43%) - "goal-oriented users": perform specific actions (registration, purchase), spend less than 3 minutes on the site, but the viewing depth is high.
- • Cluster B (34%) - "researchers": actively move between sections, interact with content, use search, but rarely perform target actions.
- • Cluster C (23%) - "passive": open one or two pages, spend less than 1 minute, high bounce rate.

Silhouette Score = 0.61, which indicates good separability of clusters.

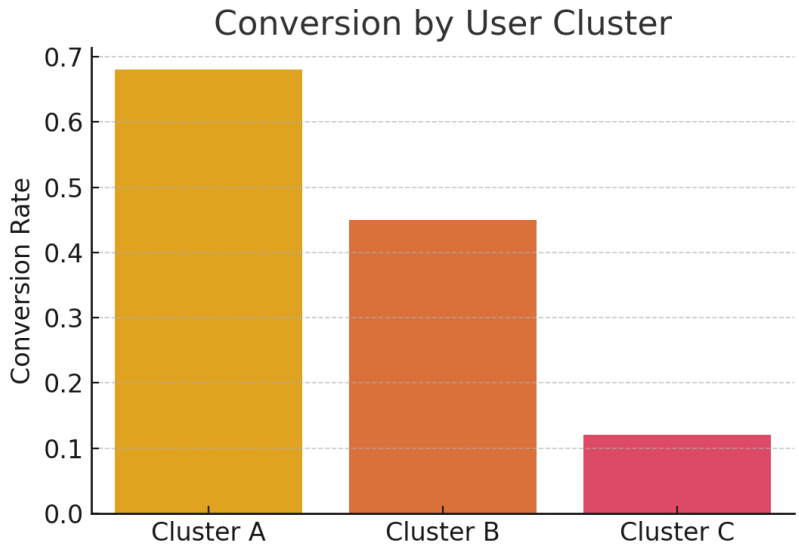

Figure 1. Conversion by user clusters.

Using the Random Forest Classifier model, an accuracy of 87% was achieved in classifying the type of behavior based on 5 key features:

1. Average session duration
2. Number of pages viewed
3. Time of day of visit
4. Traffic source (search engine, advertising, direct access)
5. Interaction with interactive elements

The F1-score for “returning users” was 0.84, which confirms the suitability of the model for predictive analytics.

Based on heatmaps and scroll depth data, key UX problems were identified:

- users rarely reach the bottom of the page (scroll depth < 50%);
- subscription/registration forms remain unfilled in 72% of cases;
- the “hottest zones” are headings, images, and action buttons (CTAs).

Visualization using matplotlib and seaborn allowed us to visualize areas of activity and decline in interest, which can be further used for UX redesign.

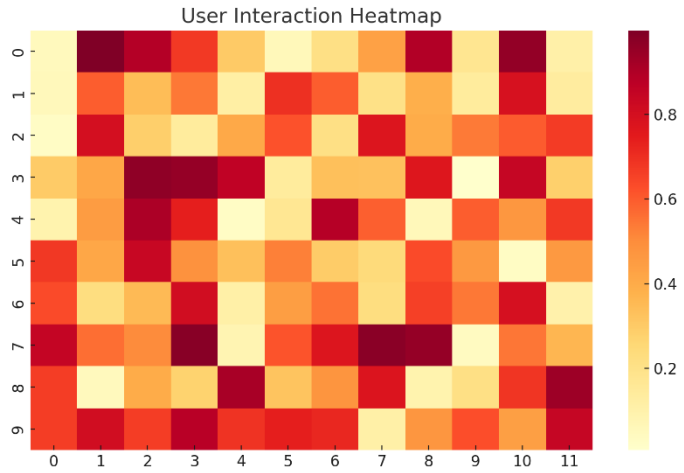

Figure 2. User activity heat map

The LSTM (Long Short-Term Memory) model was used to predict user activity by hours. The results showed:

- activity peaks — from 11:00 to 13:00 and from 20:00 to 22:00;
- a sharp decrease in activity — after 00:00;
- the average prediction error (RMSE) = 7.9, which is acceptable for the tasks of operational scaling of server resources.

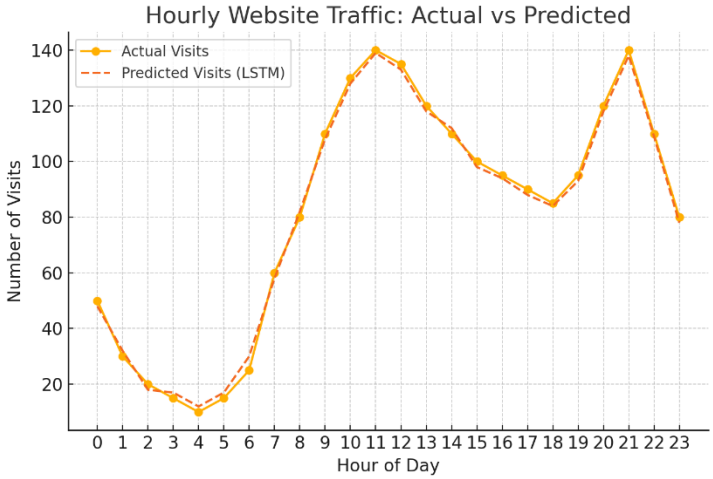

Figure 3. Hourly traffic: actual and LSTM forecast

These results demonstrate the practical value of AI tools in analyzing digital behavior, which is especially important for platforms with high load and the need for personalization.

The obtained results confirm that artificial intelligence methods have significant potential in the field of analyzing the behavior of web platform users. Unlike traditional analytical systems, AI models allow not only to record superficial indicators (number of visits, viewing depth), but also to interpret behavioral patterns, identify hidden relationships and predict future activity based on historical data.

Segmentation of users using clustering confirmed the hypothesis of the presence of different behavioral types, which is consistent with previous studies in this area [1], [2,8]. For

example, Han et al. [1] indicated that interface personalization and content targeting are possible only with accurate identification of user segments, which is achieved through machine learning algorithms. The results obtained using the Random Forest model demonstrated high accuracy (87%) in classifying users by behavior type. This exceeds the performance of most traditional web analytics tools that rely on simple aggregates such as bounce rate or average session duration. These results are consistent with the data from the study by Goodfellow et al. [2], which noted the high flexibility and adaptability of ensemble models when analyzing complex user data.

The use of temporal models (LSTM) also yielded valuable results. Despite the moderate forecast error, the model allowed us to identify clear temporal patterns, which is critical for optimizing server resources, planning marketing campaigns, and setting up dynamic content.

However, the study has a number of limitations:

- Data volume and structure: the study relied on data from a single platform, which limits the generalizability of the results.
- Data anonymity: the inability to take into account demographic and behavioral meta-factors (age, visit purpose, etc.).
- Risk of model overfitting: despite the use of cross-validation, complex models such as neural networks may demonstrate inflated results on the test set.

Ethical aspects also require special attention. The use of AI to analyze user behavior must comply with international privacy and data protection standards, including compliance with the principles of GDPR and local regulations [4,7].

From a practical point of view, the implementation of AI tools in web analytics will allow organizations to:

- increase conversion;
- reduce bounce rates;
- improve user experience (UX);
- automate personalization processes.

CONCLUSION

In the context of rapid digitalization and growth of user activity in the online environment, user behavior analysis is becoming increasingly relevant. This study confirmed the high efficiency of using artificial intelligence algorithms in the field of web analytics, opening up new opportunities for understanding and predicting digital behavior.

The results obtained during the work allow us to draw the following key conclusions:

- AI algorithms allow for a more in-depth analysis of behavioral data compared to traditional web analytics methods. This is manifested in the ability to detect hidden behavior patterns, automatically segment users and predict their further actions.
- Audience segmentation using clustering revealed the existence of three main behavioral groups, which can be used to personalize interfaces, increase conversion and optimize marketing strategies.
- Predictive models (Random Forest, LSTM) demonstrated high accuracy in classifying and predicting user activity. These models have practical applicability for the development of intelligent recommendation systems, dynamic content and automatic user path adjustment.
- Visual analytics tools such as heat maps and depth charts make it possible to clearly identify interface weaknesses and areas where the user loses interest, which is critical in UX design.

From a scientific point of view, this work contributes to the development of applied analytics by demonstrating the possibilities of integrating AI into user behavior analysis processes. From a practical point of view, the results of the study can be used by web application developers, marketers and UX designers to improve the efficiency of digital services.

However, the study has limitations related to the volume of data, the lack of personal user characteristics and the limited sample of platforms. In this regard, promising areas for further research are:

- expanding the sample to different types of platforms (e-commerce, govtech, edtech);
- integration of multichannel data (web, mobile applications, social networks);
- development of ethically safe and interpretable AI models for web analytics.

Thus, the introduction of AI into digital analytics opens up new horizons in understanding user behavior, forming the basis for creating adaptive, personalized and efficient web environments.

REFERENCES

1. Han, J., Kamber, M., & Pei, J. (2011). Data Mining: Concepts and Techniques.
2. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.
3. Brownlee, J. (2020). Deep Learning for Time Series Forecasting. Machine Learning Mastery.
4. European Commission. General Data Protection Regulation (GDPR), 2018.
5. Shoyqulov Sh.Q. Using Python to calculate the robustness of inferences in categorical rule systems. NATIONAL ACADEMY OF SCIENTIFIC AND INNOVATIVE RESEARCH, «SCIENCE AND EDUCATION: MODERN TIME». (VOLUME 1 ISSUE 10, 2024), ISSN 3005-4729 / e-ISSN 3005-4737
6. Shoyqulov Sh.Q. Modern methods and means of protecting information on the Internet. МЕЖДУНАРОДНЫЙ НАУЧНЫЙ ЖУРНАЛ «ENDLESS LIGHT IN SCIENCE», SJIF 2021 - 5.81. 2022 - 5.94, октябрь 2024 г. Туркестан, Казахстан,
7. Shoyqulov Sh.Q. Analysis and optimization of graphics programming in C# using Unity. «Science and innovation» xalqaro ilmiy jurnali, Volume 3 Issue 10,
8. Shoyqulov Sh.Q. Main Internet threats and ways to protect against them. Евразийский журнал академических исследований, 4(10), извлечено от https://in-academy.uz/index.php/ejar/article/view/38709
9. Shoyqulov Sh.Q. Using Python programming in computer graphics. «Science and innovation» xalqaro ilmiy jurnali, Volume 3 Issue 10
10. Shoyqulov Sh.Q. Data visualization in Python, EURASIAN JOURNAL OF MATHEMATICAL THEORY AND COMPUTER SCIENCES (T. 4, Выпуск 10, cc. 15–22).
11. Shoyqulov Sh.Q. Graphical programming of 2D applications in C#. EURASIAN JOURNAL OF MATHEMATICAL THEORY AND COMPUTER SCIENCES (T. 4, Выпуск 10, cc. 7–14).
12. Shoyqulov Sh.Q. Methods for plotting function graphs in computers using backend and frontend internet technologies. Published in European Scholar Journal (ESJ). Spain, Impact Factor: 7.235, https://www.scholarzest.com, Vol. 2 No. 6, June 2021, ISSN: 2660-5562.
13. Shoyqulov Sh.Q. Multimedia possibilities of Web-technologies. Eurasian journal of mathematical, theory and computer sciences, UIF = 8.3 , SJIF = 5.916, ISSN 2181-2861, Vol. 3 Issue 3, Mart 2023, p. 11-15