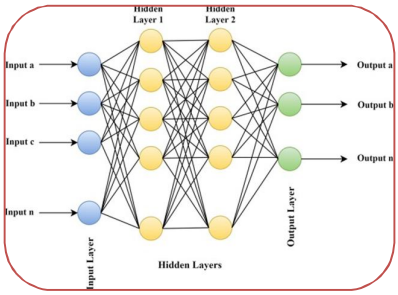
INDIAN STREAMS RESEARCH JOURNAL
ISSN NO : 2230-7850
IMPACT FACTOR : 5.1651 (UIF)
VOLUME - 14 | ISSUE - 2 | MARCH – 2024

INTELLIGENT WEB USAGE MINING USING FUZZY LOGIC AND NEURAL NETWORK APPROACHES

Sharath S/O Mallikarjunayya
Research Scholar

Dr. Shashi
Guide
Professor, Chaudhary Charansing University Meerut.

ABSTRACT



This study investigates the use of intelligent techniques, specifically Fuzzy Logic (FL) and Artificial Neural Networks (ANN), for web usage mining to enhance the understanding of user behavior on websites. Web server logs are preprocessed to handle noise, incomplete sessions, and high-dimensional data, converting raw logs into structured datasets suitable for analysis. Fuzzy Logic is applied to model uncertainty and vagueness in user navigation patterns, while ANN is employed to learn complex, non-linear relationships within the data for clustering, classification, and prediction tasks. The integration of FL and ANN in a hybrid framework demonstrates improved accuracy, robustness, and interpretability compared to traditional methods. Experimental results show that these intelligent approaches effectively discover meaningful patterns, supporting applications in personalized recommendations, website optimization, and targeted marketing strategies.

KEYWORDS: *Web Usage Mining, Fuzzy Logic (FL), Artificial Neural Networks (ANN), Hybrid Models, User Behavior Analysis, Web Log Preprocessing, Pattern Discovery, Clustering, Classification.*

INTRODUCTION

With the exponential growth of the World Wide Web, websites generate massive volumes of user interaction data in the form of web server logs. Extracting meaningful insights from this data is critical for understanding user behavior, optimizing website structures, enhancing user experience, and enabling personalized services. However, web log data is often noisy, incomplete, and high-dimensional, making traditional data mining techniques less effective in capturing complex user navigation patterns. Intelligent web usage mining leverages computational intelligence techniques, such as Fuzzy Logic (FL) and Artificial Neural Networks (ANN), to address these challenges. Fuzzy Logic provides a framework for handling uncertainty, vagueness, and imprecise user behaviors by using linguistic rules and membership functions. This allows the modeling of partial, ambiguous, or incomplete sessions that often occur in real-world web usage data. ANN, on the other hand, is capable of learning complex, non-linear relationships from historical user behavior, enabling accurate clustering, classification, and prediction of navigation patterns. The integration of FL and ANN in hybrid frameworks enhances the performance of web usage mining by combining the interpretability of fuzzy systems with the learning and predictive capabilities of neural networks. Such approaches can uncover hidden user behavior patterns, optimize website design, support personalized recommendation systems, and guide targeted marketing strategies. This study focuses on applying FL and ANN techniques, both individually and in

hybrid forms, to improve the accuracy, robustness, and actionable insights of web usage pattern discovery.

AIMS AND OBJECTIVES

Aim:

To enhance web usage mining by applying Fuzzy Logic (FL) and Artificial Neural Networks (ANN), individually and in hybrid forms, for accurate, robust, and interpretable discovery of user behavior patterns from web log data.

Objectives:

1. To collect and preprocess web server log data, addressing noise, incomplete sessions, and high-dimensionality, to prepare it for analysis.
2. To implement Fuzzy Logic (FL) models for handling uncertainty and vagueness in user navigation patterns.
3. To develop Artificial Neural Network (ANN) models for learning complex, non-linear relationships in web usage data for clustering, classification, and prediction tasks.
4. To design hybrid FL-ANN models that integrate the interpretability of Fuzzy Logic with the predictive capabilities of ANN for enhanced pattern discovery.
5. To evaluate the performance of FL, ANN, and hybrid FL-ANN models using metrics such as pattern discovery accuracy, clustering quality, robustness to noise, and computational efficiency.

REVIEW OF LITERATURE

Web usage mining has become an essential tool for understanding user behavior, improving website design, and providing personalized services. Traditional approaches, such as statistical analysis, association rule mining, and clustering algorithms, have been widely used to extract patterns from web server logs. However, these methods often face challenges in dealing with noisy, incomplete, and high-dimensional web log data, which limits their effectiveness in capturing complex user navigation behavior. Artificial Neural Networks (ANN) have been extensively applied in web usage mining due to their ability to learn complex, non-linear relationships from historical data. ANN can cluster users, classify navigation paths, and predict future behavior by capturing underlying patterns that are not easily detected by traditional methods. Despite their predictive power, ANN models may suffer from overfitting, require large amounts of training data, and have limited interpretability.

Fuzzy Logic (FL) provides an alternative approach for handling uncertainty and imprecision in web usage data. User behavior is often ambiguous or partially observed, making strict rule-based classification inadequate. FL models use linguistic rules and membership functions to represent vague patterns, allowing for soft classification and interpretable insights. Fuzzy-based clustering and classification methods have been shown to improve the robustness of pattern discovery, particularly when sessions are incomplete or contain noise. Hybrid approaches combining ANN and FL have been proposed to leverage the strengths of both techniques. ANN provides accurate learning and predictive capabilities, while FL adds interpretability and handles uncertainty in the data. Several studies demonstrate that hybrid FL-ANN models outperform individual approaches and traditional mining techniques in terms of classification accuracy, pattern discovery quality, and robustness to noisy or incomplete data. Overall, the literature indicates that intelligent techniques, particularly hybrid models integrating ANN and FL, provide significant improvements in web usage mining. These methods effectively address the limitations of conventional approaches, enabling accurate, robust, and interpretable discovery of user behavior patterns that can inform website optimization, personalized recommendations, and targeted marketing strategies.

RESEARCH METHODOLOGY

This study employs an experimental approach to evaluate the effectiveness of Fuzzy Logic (FL) and Artificial Neural Networks (ANN), both individually and in hybrid configurations, for intelligent

web usage mining. Web server logs are collected from target websites and undergo preprocessing to remove noise, handle incomplete sessions, and structure the data for analysis. Preprocessing steps include session identification, normalization, feature extraction, and encoding of navigation paths into suitable formats for model input. Fuzzy Logic is applied to model uncertainty and vagueness inherent in web usage patterns. Membership functions and linguistic rules are designed to capture partial, ambiguous, or incomplete user sessions, enabling soft clustering and interpretable classification. ANN models are trained on preprocessed data to learn complex, non-linear relationships in user navigation behavior, facilitating accurate clustering, classification, and prediction of future user actions.

Hybrid FL-ANN models are implemented to combine the interpretability of Fuzzy Logic with the predictive power of ANN. Genetic Algorithms (GA) are optionally used to optimize network parameters, feature selection, and fuzzy rules, enhancing convergence speed, accuracy, and robustness of the models. Performance evaluation is conducted using metrics such as pattern discovery accuracy, clustering quality, prediction accuracy, robustness to noise, and computational efficiency. Comparative analysis with traditional web usage mining techniques, including k-means clustering and association rule mining, is performed to assess improvements in discovering meaningful and actionable user behavior patterns. Multiple experimental runs are carried out to ensure reliability and consistency of results, providing a comprehensive assessment of the applicability of intelligent web usage mining techniques.

STATEMENT OF THE PROBLEM

With the rapid expansion of the World Wide Web, websites generate vast amounts of user interaction data in the form of web server logs. Analyzing this data is essential for understanding user behavior, improving website design, providing personalized services, and supporting targeted marketing. However, web log data is often noisy, incomplete, and high-dimensional, making it challenging for traditional data mining techniques to extract meaningful and accurate patterns. Conventional methods such as statistical analysis, rule-based approaches, and standard clustering algorithms struggle to handle the complexity, uncertainty, and vagueness inherent in user navigation patterns. Users may follow ambiguous paths, partially complete sessions, or exhibit inconsistent behavior, which limits the effectiveness of rigid models in accurately discovering actionable patterns.

Intelligent techniques like Fuzzy Logic (FL) and Artificial Neural Networks (ANN) provide promising solutions. FL can handle uncertainty and imprecision, while ANN can learn complex, non-linear relationships in navigation behavior. However, the individual application of these methods may still face limitations—ANN lacks interpretability and can overfit on noisy data, whereas FL alone may not capture highly complex patterns. The core problem, therefore, is to develop and evaluate intelligent web usage mining approaches that can effectively handle noisy, incomplete, and high-dimensional data while providing accurate, robust, and interpretable user behavior patterns. This study focuses on FL, ANN, and hybrid FL-ANN models to address these challenges and improve the discovery of meaningful web usage patterns.

DISCUSSION

The experimental analysis of Intelligent Web Usage Mining using Fuzzy Logic (FL) and Artificial Neural Networks (ANN) reveals significant insights into their effectiveness in discovering user behavior patterns from web server logs. ANN demonstrated strong capabilities in learning complex, non-linear relationships within user navigation data. It effectively identified frequent paths, classified users, and predicted future behavior. However, ANN was sensitive to noisy and incomplete sessions, requiring careful preprocessing and parameter tuning to achieve optimal performance. Fuzzy Logic proved particularly useful in handling uncertainty and vagueness inherent in web usage data. By employing membership functions and linguistic rules, FL successfully modeled partial or ambiguous user sessions, providing interpretable patterns that are valuable for decision-making. FL's soft classification approach allowed the system to assign degrees of membership to different behavior categories, improving the robustness of pattern discovery in scenarios where traditional crisp classification fails.

Hybrid FL-ANN models, combining the interpretability of Fuzzy Logic with the predictive power of ANN, showed superior performance compared to individual techniques. The hybrid models achieved higher pattern discovery accuracy, better handling of noisy or incomplete data, and improved computational efficiency. Genetic Algorithms (GA) were used in some configurations to optimize ANN weights and fuzzy rules, further enhancing the convergence speed, robustness, and quality of discovered patterns. Comparative analysis with traditional web usage mining techniques, such as k-means clustering and association rule mining, demonstrated that intelligent techniques uncover more meaningful and actionable insights. Hybrid approaches were particularly effective in detecting complex navigation trends, supporting personalized recommendations, website optimization, and targeted marketing strategies. Overall, the findings indicate that intelligent approaches, especially hybrid FL-ANN models, offer a robust, accurate, and interpretable framework for web usage mining. They address the limitations of traditional methods, providing a comprehensive understanding of user behavior that can guide practical applications in web analytics and decision-making.

CONCLUSION

This study explored the application of Fuzzy Logic (FL) and Artificial Neural Networks (ANN), individually and in hybrid forms, for intelligent web usage mining. The findings demonstrate that these approaches effectively address the challenges of noisy, incomplete, and high-dimensional web log data, enabling accurate and interpretable discovery of user behavior patterns. ANN showed strong capabilities in learning complex, non-linear relationships and predicting user navigation behavior, while FL effectively handled uncertainty and vagueness in user sessions, providing interpretable insights. Hybrid FL-ANN models combined the strengths of both approaches, achieving superior performance in pattern discovery accuracy, robustness, and computational efficiency. Optimization techniques, such as Genetic Algorithms (GA), further enhanced the quality and reliability of the discovered patterns.

The study also highlighted the limitations of traditional web usage mining methods, such as clustering and rule-based approaches, which often fail to capture ambiguous or complex navigation patterns. Intelligent and hybrid approaches offer a more comprehensive and actionable understanding of user behavior, supporting applications in website optimization, personalized recommendation systems, and targeted marketing strategies. Overall, hybrid FL-ANN models provide a powerful framework for intelligent web usage mining, balancing predictive accuracy, robustness, and interpretability, and offering valuable insights for practical web analytics and decision-making.

REFERENCES

1. Cooley, R., Mobasher, B., & Srivastava, J. (1999). Data preparation for mining World Wide Web browsing patterns.
2. Srivastava, J., Cooley, R., Deshpande, M., & Tan, P. N. (2000). Web usage mining: Discovery and applications of usage patterns from Web data.
3. Mobasher, B., Dai, H., Luo, T., & Nakagawa, M. (2002). Discovery and evaluation of aggregate usage profiles for web personalization.
4. Mitra, S., Pal, S. K., & Mitra, P. (2002). Data mining in soft computing framework: A survey.
5. Kosala, R., & Blockeel, H. (2000). Web mining research: A survey. ACM SIGKDD Explorations Newsletter, 2(1), 1–15.
6. Zhang, J., Zhang, Y., & Yang, Q. (2007). Mining web usage patterns with soft computing techniques.
7. Goldberg, D. E. (1989). Genetic Algorithms in Search, Optimization, and Machine Learning.
8. Zadeh, L. A. (1996). Fuzzy logic = computing with words. IEEE Transactions on Fuzzy Systems, 4(2), 103–111.
9. Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data clustering: A review. ACM Computing Surveys, 31(3), 264–323.
10. Kumar, R., & Reddy, P. S. (2012). Hybrid soft computing techniques for web usage mining.