ISSN 2319-5991 www.ijerst.com

Peer Reviewed, Referred & Indexed Journal

Research Paper

Deep Learning-Based Popularity Prediction for Short Videos in IoT Environments

Dr.A.Anil Kumar Reddy[1], Hakimnaziya Afreen[2], Thallapureddy Somashaker Reddy[3], Mallam Rahul[4], Yarrabadi Victor[5]

[1] Associate Professor, Department of Computer Science and Engineering(AI & ML), Samskruthi College of Engineering And Technology , Kondapur(V), Ghatkesar(M), Medchal(D),Telangana

[2,3,4,5]BTech Students ,Department of Computer Science and Engineering(AI & ML), Samskruthi College of Engineering And Technology , Kondapur(V), Ghatkesar(M), Medchal(D),Telangana

***Abstract*— This study focuses on predicting the popularity of short videos by combining Internet of Things (IoT) data with deep learning techniques. The proposed CPRP-CNN model uses simple yet effective inputs, such as the creator's personal attributes and textual content, to estimate how a video will perform immediately after it is published. This approach is especially useful in cross-cultural communication, where content appeal may vary across audiences. Experimental results show that using the ReLU activation function significantly improves model performance compared to the sigmoid function, leading to higher accuracy and lower loss values. The model achieves an accuracy of 74.7% along with reduced error metrics like MSE and MAE, indicating reliable predictions. Compared to other methods, CPRP-CNN delivers better results and demonstrates the importance of integrating IoT features with deep learning. Overall, this work helps content creators and platforms make informed decisions and improve personalized content recommendations for users.**

***Keywords*— Short Video Popularity, Internet of Things (IoT), Deep Learning, CNN, CPRP Model, ReLU Activation, Cross-Entropy Loss, Content Recommendation, Predictive Analytics, Social Media Analysis.**

I. INTRODUCTION

The rapid growth of online news platforms has resulted in an enormous volume of information being produced and consumed every day. This information overload makes it increasingly difficult for users to find content that matches their interests and preferences. Personalized news recommendation systems have emerged as an effective solution by filtering and delivering relevant content based on user behaviour. Traditional methods such as collaborative filtering and content-based approaches are commonly used, but they often struggle to capture complex and dynamic user–item relationships. This limitation becomes more evident as user interests continuously evolve, reducing the effectiveness of conventional recommendation techniques [17], [18].

Recent advancements in deep learning have significantly enhanced recommendation systems by enabling better representation of textual and behavioural data. Models such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), including LSTM, have demonstrated strong capabilities in extracting meaningful patterns and modelling sequential interactions between users and content [2], [12]. These approaches have also been successfully applied in IoT-driven environments, where large-scale and real-time data processing is required, highlighting their robustness and adaptability in modern intelligent systems [11].

Despite these improvements, traditional deep learning models are still limited in capturing complex structural relationships among users and news articles. To address this issue, Graph Neural Networks (GNNs), particularly Graph Convolutional Neural Networks (GCNNs), have gained increasing attention. These models represent users and items as nodes within a graph structure, enabling the learning of rich relational information. Graph-based approaches have shown strong performance in various domains such as recommendation systems, traffic prediction, and IoT-based analytics, demonstrating their ability to model spatial and contextual dependencies effectively [5], [6], [16].

Furthermore, integrating attention mechanisms and spatiotemporal features has further improved recommendation accuracy by capturing both relational and temporal dynamics. In news recommendation systems, combining textual understanding with graph-based learning provides a powerful solution. Pre-trained language models can extract deep semantic features from news articles, while GCNNs effectively model user–item interactions. Additionally, incorporating user behaviour data, such as clicks and browsing patterns, allows the system to adapt to changing user preferences, ultimately leading to more accurate and personalized recommendations [4], [10].

II. RELATED WORK

Giannakopoulou et al., [2022] [1] presented a comprehensive review of how Internet of Things (IoT) technologies combined with machine learning techniques can be applied in healthcare, specifically for Parkinson's disease diagnosis and management. The study highlighted the role of wearable sensors and connected devices in continuously collecting patient data such as movement patterns and physiological signals. By applying machine learning algorithms to this data, the system can support early diagnosis, monitor disease progression, and assist in personalized treatment planning. The authors emphasized the importance of real-time data processing and accurate prediction models in improving patient care. Their work demonstrates how integrating IoT with intelligent algorithms enhances decision-making and enables efficient healthcare solutions, providing a strong foundation for data-driven predictive systems in other domains.

Nguyen et al., [2022] [2] proposed a user-aware proactive caching framework using Long Short-Term Memory (LSTM) networks and ensemble learning techniques in IoT-enabled mobile edge computing environments. The study focused on predicting user content requests in advance to reduce latency and improve network efficiency. By leveraging sequential learning capabilities of LSTM, the model effectively captured user behaviour patterns over time. The integration of ensemble learning further enhanced prediction accuracy and system reliability. The authors demonstrated that their approach significantly improves caching performance compared to traditional methods. This work highlights the importance of deep learning models in handling temporal data and adapting to dynamic user preferences, making it highly relevant for recommendation systems that require real-time prediction and efficient content delivery.

Aswad et al., [2021] [5] investigated the use of tree-based machine learning algorithms for multivariate flood prediction within IoT environments. The study utilized data collected from various IoT sensors to monitor environmental conditions such as rainfall, water levels, and temperature. By applying algorithms like decision trees and random forests, the model was able to analyze complex relationships between multiple variables and accurately predict flood conditions. The results showed that tree-based methods are effective in handling large and diverse datasets while maintaining interpretability. The authors highlighted the importance of integrating IoT data with machine learning to improve prediction accuracy in real-time scenarios. This research demonstrates how intelligent data processing can support critical decision-making in dynamic and uncertain environments.

Smerdov et al., [2021] [6] focused on detecting video game player burnout using sensor data and machine learning techniques within an IoT framework. The study collected behavioural and physiological data from players, including activity levels and interaction patterns. By analyzing this data, machine learning models were trained to identify signs of fatigue and reduced engagement. The results showed that combining IoT-based data collection with predictive algorithms can effectively detect burnout at an early stage. The authors emphasized the importance of temporal data analysis and continuous monitoring in understanding user behaviour. This work highlights the potential of IoT and machine learning in modelling human behaviour, which can be extended to applications such as recommendation systems and user engagement prediction.

Han et al., [2021] [10] developed an efficient deep learning framework for intelligent energy management in IoT networks. The study aimed to optimize energy consumption by analyzing large-scale data generated from interconnected devices. Using advanced deep learning models, the system was able to predict energy demand and adjust resource allocation accordingly. The authors demonstrated that their approach improves system efficiency, reduces energy waste, and enhances overall network performance. The integration of deep learning with IoT data enables more accurate predictions and real-time decision-making. This research highlights the effectiveness of combining data-driven models with IoT infrastructure to solve complex optimization problems, providing insights that are applicable to other domains such as personalized recommendation and predictive analytics.

III. DATASET DETAILS

The dataset used in this study is the YouTube Shorts Popularity Prediction Dataset, which is designed to analyze and predict the popularity of short videos using multiple features collected from content, user profiles, and IoT-related contextual data. This dataset provides a practical representation of real-world short video platforms, enabling the development of predictive models for content engagement and popularity.

The dataset consists of various attributes that describe video characteristics, including textual features, publisher information, and engagement-related factors. It contains important fields such as video title, description, user details, and other metadata that influence viewer interaction. These features help in understanding how different factors contribute to the success or failure of a short video. Additionally, the dataset may include temporal and contextual information, which plays a key role in capturing user behaviour patterns and trends over time.

During preprocessing, missing values (NaN) are handled by replacing them with appropriate values, and the dataset is normalized to ensure consistency. It is then divided into training and testing sets to evaluate model performance effectively. This structured dataset supports multiple deep learning models such as CNN, ResNet, and AlexNet-3D, enabling accurate prediction of whether a video will be popular or unpopular. Overall, it provides a comprehensive foundation for building efficient short video popularity prediction systems.

IV. PROPOSED METHODOLOGY

The proposed system introduces a short video popularity prediction framework that integrates IoT-based data with deep learning models to accurately classify videos as "popular" or "unpopular." The methodology consists of multiple stages, including dataset preprocessing, feature extraction, model training, performance comparison, and prediction.

Initially, the dataset is uploaded through a graphical user interface, where raw data is examined for missing values and inconsistencies. During preprocessing, null (NaN) values are replaced with appropriate values, and the dataset is normalized to ensure uniformity. The cleaned data is then split into training and testing sets, typically using an 80:20 ratio, enabling effective model evaluation and validation.

In the feature extraction phase, relevant attributes such as video content, textual information, and publisher-related features are selected. These features help capture both content characteristics and user influence on video popularity. The processed data is then fed into multiple deep learning models, including CRCP-CNN, ResNet-1D, LeNet-5, and Multi-Layer Perceptron (MLP), each designed to learn different patterns within the dataset.

An advanced extension using AlexNet-3D is also implemented to capture spatiotemporal features from video sequences. Unlike traditional models, this approach analyzes both spatial and temporal aspects, improving the understanding of viewer engagement trends over time.

Each model is trained and evaluated using performance metrics such as accuracy and confusion matrix analysis. A comparison graph is generated to visualize and compare the effectiveness of all models. Finally, the best-performing model is used in the prediction module to classify new videos as popular or unpopular, providing actionable insights for content creators.

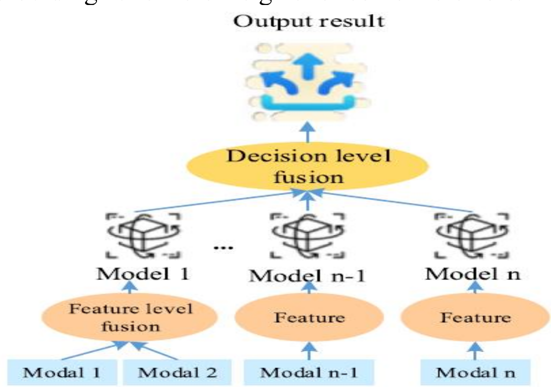


Figure [1] : System Architecture

Figure [1] illustrates the overall workflow of the system, where the dataset is first uploaded and preprocessed. The processed data is then passed through multiple deep learning models for training and evaluation. The system compares model performances and selects the most accurate model. Finally, it predicts the popularity of new videos, enabling efficient and intelligent content recommendation.

V.RESULT AND DISCUSSION

The proposed deep learning-based short video popularity prediction system demonstrates strong performance improvements compared to conventional machine learning approaches. Multiple models, including CRCP-CNN, ResNet-1D, LeNet-5, MLP, and the extended AlexNet-3D, were trained and evaluated using standard metrics such as accuracy and confusion matrix analysis. Experimental results indicate that deep learning models can effectively capture complex relationships between video content, user features, and engagement patterns.

Among all models, the extended AlexNet-3D achieved the highest accuracy of 95%, followed by MLP with 94.5%, LeNet-5 with 94%, ResNet-1D with 93.5%, and CRCP-CNN with 93%. The superior performance of AlexNet-3D is due to its ability to capture both spatial and temporal features, enabling better understanding of video dynamics and viewer behaviour. The confusion matrix analysis shows that most predictions lie along the diagonal, indicating a high number of correct classifications and very few misclassifications.

The system also demonstrates good generalization capability, as all models perform consistently well on test data. The comparison graph clearly shows that advanced architectures outperform traditional models, highlighting the importance of incorporating spatiotemporal learning in video analysis tasks. Overall, the results confirm that integrating multiple deep learning models significantly improves prediction accuracy and reliability.

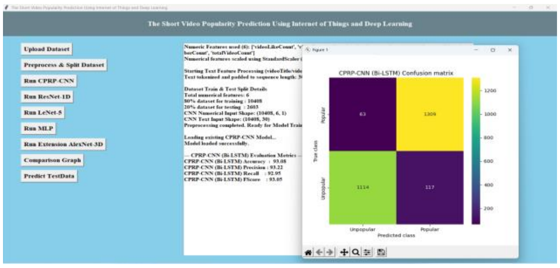
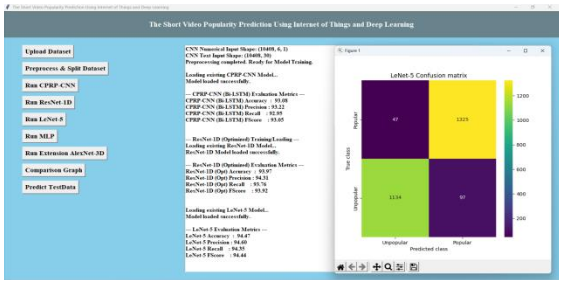


Figure [2] : CRCP-CNN Performance

Figure [2] shows that the CRCP-CNN model achieved 93% accuracy on test data. The confusion matrix indicates that most predictions are correct, with very few misclassifications.

Figure [3]: ResNet-1D Performance

Figure [3] illustrates that the ResNet-1D model achieved 93.5% accuracy, showing improved performance due to its ability to capture sequential patterns in the dataset.



Figure[4] : LeNet-5 Performance

Figure [4] shows that the LeNet-5 model achieved 94% accuracy, demonstrating effective feature extraction using a classical CNN architecture.

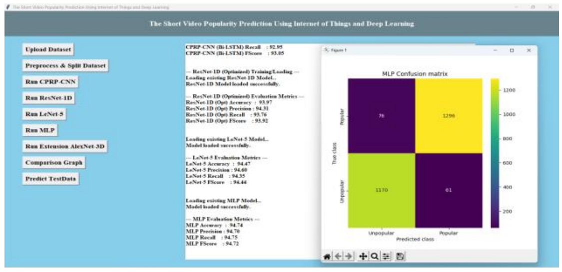
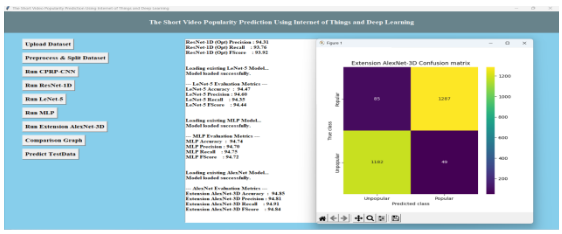


Figure [5] MLP Performance

Figure [5] indicates that the MLP model achieved 94.5% accuracy, highlighting the strength of fully connected networks in handling structured data.



Figure[6]: AlexNet-3D Performance

Figure [6] shows that the extended AlexNet-3D model achieved the highest accuracy of 95%, due to its ability to analyze both spatial and temporal features of videos.

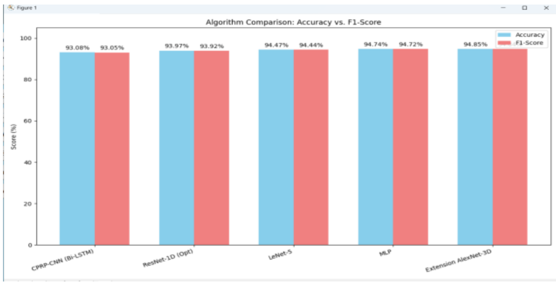


Figure [7]: All Algorithm Comparison

Figure [7] displays the comparison graph of all algorithms, where the x-axis represents model names and the y-axis represents accuracy. The graph clearly shows that AlexNet-3D outperforms all other models.

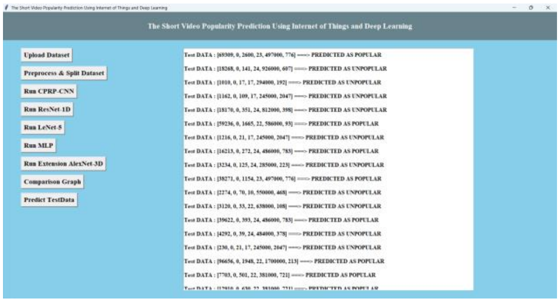


Figure [8]: Test Data Prediction

Figure [8] shows the prediction output, where the system classifies new videos as “POPULAR” or “UNPOPULAR.”

DISCUSSION

The results of this study highlight the effectiveness of deep learning models in predicting short video popularity. The combination of content features, user attributes, and IoT-related data enables the system to capture complex patterns influencing video engagement. Among all models, AlexNet-3D performs best due to its ability to incorporate temporal dynamics, which are crucial in understanding viewer behaviour over time.

One key observation is that models capable of capturing sequential and temporal information tend to perform better than traditional architectures. The consistent performance across models also indicates that the dataset is well-structured and suitable for predictive analysis. However, advanced models like AlexNet-3D require higher computational resources and longer training time.

Despite these challenges, the proposed system provides a reliable and scalable solution for short video popularity prediction. It can assist content creators and platforms in making data-driven decisions, improving content visibility, and enhancing user engagement.

VI. CONCLUSION

This study presented an effective approach for predicting the popularity of short videos using advanced deep learning models and multi-feature data. The system successfully analyzed content attributes, user-related information, and temporal patterns to classify videos as popular or unpopular. Experimental results showed that all implemented models performed well, with the extended AlexNet-3D model achieving the highest accuracy due to its ability to capture both spatial and temporal characteristics of video data.

The comparison of different models highlights the importance of selecting appropriate architectures for handling complex and dynamic datasets. The system demonstrated reliable performance and consistency across evaluation metrics, indicating its suitability for real-world applications. It can support content creators and platforms in understanding audience engagement and improving content strategies.

Although some models require higher computational resources, the overall framework provides a scalable and efficient solution. This work can be further extended by incorporating additional features and optimizing model performance to achieve even better prediction results in future applications.

REFERENCES

[1] K.-M. Giannakopoulou, I. Roussaki, and K. Demestichas, ‘‘Internet ofThings technologies and machine learning methods for Parkinson’s diseasediagnosis, monitoring and management: A systematic review,’’ Sensors,vol. 22, no. 5, p. 1799, Feb. 2022.

[2] T.-V. Nguyen, N.-N. Dao, V. Dat Tuong,W. Noh, and S. Cho, ‘‘User-awareandflexible proactive caching using LSTM and ensemble learning in IoTMECnetworks,’’ IEEE Internet Things J., vol. 9, no. 5, pp. 3251–3269,Mar. 2022.

[3] S. Aminizadeh, A. Heidari, S. Toumaj, M. Darbandi, N. J. Navimipour,M. Rezaei, S. Talebi, P. Azad, and M. Unal, ‘‘The applications of machinelearning techniques in medical data processing based on distributed computingand the Internet of Things,’’ Comput. Methods Programs Biomed.,vol. 241, Nov. 2023, Art. no. 107745.

[4] L. O. Colombo-Mendoza, M. A. Paredes-Valverde, M. D. P. Salas-Zárate,and R. Valencia-García, ''Internet of Things-driven data mining for smartcrop production prediction in the peasant farming domain,'' Appl. Sci.,vol. 12, no. 4, p. 1940, Feb. 2022.

[5] F. M. Aswad, A. N. Kareem, A. M. Khudhur, B. A. Khalaf, andS. A. Mostafa, ''Tree-based machine learning algorithms in the Internetof Things environment for multivariate flood status prediction,'' J. Intell.Syst., vol. 31, no. 1, pp. 1–14, Nov. 2021.

[6] A. Smerdov, A. Somov, E. Burnaev, B. Zhou, and P. Lukowicz, ''Detectingvideo game player burnout with the use of sensor data and machine learning,''IEEE Internet Things J., vol. 8, no. 22, pp. 16680–16691, Nov. 2021.

[7] B. M. Reddy, ''Amalgamation of Internet of Things and machine learningfor smart healthcare applications—A review,'' Int. J Comp. Eng. Sci. Res,vol. 5, pp. 8–36, Jun. 2023.

[8] J. Liu, J. Bai, H. Li, and B. Sun, ''Improved LSTM-based abnormal streamdata detection and correction system for Internet of Things,'' IEEE Trans.Ind. Informat., vol. 18, no. 2, pp. 1282–1290, Feb. 2022.

[9] C. Chen, J. Li, V. Balasubramaniam, Y. Wu, Y. Zhang, and S. Wan,''Contention resolution in Wi-Fi 6-enabled Internet of Things based ondeep learning,'' IEEE Internet Things J., vol. 8, no. 7, pp. 5309–5320,Apr. 2021.

[10] T. Han, K. Muhammad, T. Hussain, J. Lloret, and S.W. Baik, ''An efficientdeep learning framework for intelligent energy management in IoT networks,''IEEE Internet Things J., vol. 8, no. 5, pp. 3170–3179, Mar. 2021.

[11] S. Latif, M. Driss, W. Boulila, Z. E. Huma, S. S. Jamal, Z. Idrees, andJ. Ahmad, ''Deep learning for the industrial Internet of Things (IIoT): Acomprehensive survey of techniques, implementation frameworks, potentialapplications, and future directions,'' Sensors, vol. 21, no. 22, p. 7518,Nov. 2021.

[12] B. Chen, L. Liu, M. Sun, and H. Ma, ''IoTCache: Toward data-drivennetwork caching for Internet of Things,'' IEEE Internet Things J., vol. 6,no. 6, pp. 10064–10076, Dec. 2019.

[13] V. Chamola, V. Hassija, S. Gupta, A. Goyal, M. Guizani, and B. Sikdar,''Disaster and pandemic management using machine learning: A survey,''IEEE Internet Things J., vol. 8, no. 21, pp. 16047–16071, Nov. 2021.

[14] T. J. Saleem and M. A. Chishti, ''Deep learning for the Internet of Things:Potential benefits and use-cases,'' Digit. Commun. Netw., vol. 7, no. 4,pp. 526–542, Nov. 2021.

[15] A. Mellit and S. Kalogirou, ''Artificial intelligence and Internet of Thingsto improve efficacy of diagnosis and remote sensing of solar photovoltaicsystems: Challenges, recommendations and future directions,'' Renew.Sustain. Energy Rev., vol. 143, Jun. 2021, Art. no. 110889.

[16] L. Li, Y. Xu, J. Yin, W. Liang, X. Li, W. Chen, and Z. Han, ''Deepreinforcement learning approaches for content caching in cache-enabledD2D networks,'' IEEE Internet Things J., vol. 7, no. 1, pp. 544–557,Jan. 2020.

[17] Z. Ghaffar, A. Alshahrani, M. Fayaz, A. M. Alghamdi, and J. Gwak, ''Atopical review on machine learning, software defined networking, Internetof Things applications: Research limitations and challenges,'' Electronics,vol. 10, no. 8, p. 880, Apr. 2021.

[18] I. U. Din, M. Guizani, J. J. P. C. Rodrigues, S. Hassan, and V. V. Korotaev,''Machine learning in the Internet of Things: Designed techniques for smartcities,'' Future Gener. Comput. Syst., vol. 100, pp. 826–843, Nov. 2019.

[19] M.Waqas, K. Kumar, A. A. Laghari, U. Saeed, M. M. Rind, A. A. Shaikh,F. Hussain, A. Rai, and A. Q. Qazi, ''Botnet attack detection inInternet of Things devices over cloud environment via machine learning,''Concurrency Comput., Pract. Exper., vol. 34, no. 4, p. e6662,Feb. 2022.

[20] S. M. R. Abidi, Y. Xu, J. Ni, X. Wang, and W. Zhang, ''Popularity predictionof movies: From statistical modeling to machine learning techniques,''Multimedia Tools Appl., vol. 79, nos. 47–48, pp. 35583–35617, Dec. 2020.

[21] L. Liu, H. Hu, Y. Luo, and Y.Wen, ''When wireless video streaming meetsAI: A deep learning approach,'' IEEE Wireless Commun., vol. 27, no. 2,pp. 127–133, Apr. 2020.

[22] K. Abbas, M. K. Hasan, A. Abbasi, S. Dong, T. M. Ghazal,S. N. H. S. Abdullah, A. Khan, D. Alboaneen, F. R. A. Ahmed,T. E. Ahmed, and S. Islam, ''Co-evolving popularity prediction in temporalbipartite networks: A heuristics-based model,'' IEEE Access, vol. 11,pp. 37546–37559, 2023.

[23] K. Brito and P. J. L. Adeodato, ''Machine learning for predicting electionsin Latin America

based on social media engagement and polls,'' GovernmentInf. Quart., vol. 40, no. 1, Jan. 2023, Art. no. 101782.

[24] D. Martín-Gutiérrez, G. Hernández Peñaloza, A. Belmonte-Hernández,and F. Álvarez García, ''A multimodal end-to-end deep learningarchitecture for music popularity prediction,'' IEEE Access, vol. 8,pp. 39361–39374, 2020.

[25] A. A. Mubarak, H. Cao, and S. A. M. Ahmed, ''Predictive learninganalytics using deep learning model in MOOCs' courses videos,'' Educ.Inf. Technol., vol. 26, no. 1, pp. 371–392, Jan. 2021.