International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
ISSN: 3068-272X www.ijdim.com Original Research Paper

DEEPVIDNET: A DEEP LEARNING AND IOT-DRIVEN FRAMEWORK FOR PREDICTING SHORT VIDEO POPULARITY

[1] Dr.N.Bhanupriya,[2] Arepelly Anusha
[1]Assistant Professor, [2]MCA Student
Department Of MCA
Sree Chaitanya College Of Engineering, Karimnagar

ABSTRACT

With the exponential rise of short video platforms such as TikTok, YouTube Shorts, and Instagram Reels, predicting the popularity of short videos has become a key challenge for content creators and platform managers. Traditional metrics based solely on user interactions fail to capture the multi-dimensional aspects influencing video virality, such as timing, location, engagement type, and contextual cues. This paper proposes DeepVidNet, an intelligent framework integrating Internet of Things (IoT) data with deep learning architectures to accurately forecast the popularity trajectory of short videos. The proposed model captures heterogeneous data streams—user engagement signals, environmental factors, and device-level metrics—through IoT integration and processes them using Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks. Experimental results demonstrate that DeepVidNet outperforms existing popularity prediction models in terms of precision, recall, and F1-score, establishing a robust and scalable solution for real-time short video analytics.

Received: 23-09-2025 Accepted: 28-10-2025 Published: 04-11-2025

I. INTRODUCTION

In recent years, the global surge in short-form video platforms such as TikTok, YouTube Shorts, and Instagram Reels has revolutionized digital content consumption patterns. The ease of creating, sharing, and engaging with short videos has transformed not only entertainment but also marketing, education, and social interaction. With billions of videos uploaded daily, understanding and predicting which videos are likely to become popular has become a crucial challenge for content creators, advertisers, and platform administrators. Popularity prediction enables targeted advertising, efficient content recommendation, and improved resource allocation in network infrastructure. However, due to the inherently dynamic and nonlinear nature of user interactions, accurately forecasting video virality remains a complex and evolving research problem.

Traditional video popularity prediction techniques have primarily relied on static features such as video length, upload time, view count, likes, and comments. These approaches are often limited by their inability to capture complex dependencies among visual, contextual, and temporal factors influencing user engagement. Furthermore, static models do not consider real-time behavioral dynamics, making them ineffective in rapidly changing social media environments. As mobile and IoT devices become increasingly integrated into digital ecosystems, a wealth of contextual data—such as device type, location, network strength, and viewing time—has become available. These additional data streams, when properly leveraged, can provide deeper insights into user engagement behavior and video performance patterns.

The rise of Internet of Things (IoT) technologies provides a unique opportunity to enhance data-driven decision-making in video analytics. IoTenabled devices continuously generate real-time contextual information that can be correlated with social media interactions to reveal how environmental, temporal, and device-based factors influence video popularity. When combined with deep learning, this multimodal data can significantly improve predictive accuracy by modeling nonlinear relationships across diverse input features. Deep learning architectures such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) have already demonstrated remarkable performance in tasks like image recognition, speech processing, and time-series forecasting. Applying these models to the short-video domain allows for automatic feature extraction from visual and contextual inputs, leading to more precise and adaptive popularity predictions.

II. LITERATURE SURVEY

Research into content popularity prediction has evolved significantly over the past decade, shifting from traditional regression-based approaches to deep learning and contextual data-driven models. Szabo and Huberman (2010) were among the first to study online content popularity dynamics using early view patterns, showing that early engagement metrics could serve as strong indicators of future popularity. Their statistical regression model, however, failed to capture nonlinear dependencies among multimodal features such as content type and social influence. Later, Tatar et al. (2014) explored social propagation dynamics, demonstrating how user networks and social relationships impact content diffusion across digital platforms. These early studies laid the groundwork for integrating behavioral analytics with computational prediction frameworks.

With the proliferation of video-sharing platforms, researchers began focusing specifically on video popularity prediction. Khosla et al. (2014) introduced a content-based approach using visual and aural features to estimate the popularity of YouTube videos. Their results confirmed that certain visual patterns, such as brightness, color diversity, and object density, correlated with higher engagement. However, their model relied on static features and lacked real-time adaptability. Ahmed et al. (2017) later proposed a hybrid machine learning framework combining temporal engagement data and content attributes, achieving moderate improvements but struggling to process large, dynamic datasets typical of short-video ecosystems.

Deep learning approaches have since revolutionized this field by enabling end-to-end feature learning. Zhang et al. (2019) applied Convolutional Neural Networks (CNNs) to predict video popularity based on thumbnail images and user metadata, achieving superior accuracy compared to conventional feature-engineering-based models. However, their model was restricted to visual content and ignored contextual influences such as user demographics or posting time. Liu et al. (2021) expanded on this by integrating Long Short-Term Memory (LSTM) networks to capture sequential engagement trends. Their hybrid CNN-LSTM model outperformed previous baselines but required massive labeled datasets and lacked generalization across platforms with differing user behaviors.

III. SYSTEM ANALYSIS & DESIGN

EXISTING SYSTEM

A Convolutional Neural Network (CNN)-based technique for predicting the popularity of short videos was presented by Din et al. [18]. They used the CNN model to extract visual features from a large collection of short movies, then combined these traits with social network data to predict popularity. The results demonstrated that the deep learning model improves prediction accuracy by skilfully capturing complex interactions between visual and social information. Similar to this, Waqas et al. [19] suggested a Recurrent Neural Network (RNN)-

based method for predicting the popularity of short videos. They used the RNN model to encode user behaviour sequences during brief video viewing, demonstrating better prediction accuracy while allowing for dynamic changes in user behaviour.

IoT technology integration for social media data collecting was investigated by Abidi et al. [20]. They developed a sensor-based and smart device-based data collecting system that allows the gathering of environmental data while watching brief videos, including temperature, humidity, and lighting. This dataset was useful for examining user watching habits and trends in the distribution of short videos. Furthermore, Liu et al. [21] investigated the potential of IoT technology for gathering data on user behaviour. They investigated user emotions and engagement levels by using IoT sensors to detect user movements and gestures while watching brief videos. Understanding user responses to short videos and predicting their popularity were made possible by these findings.

In conclusion, the continuous development and incorporation of well-established research methodology leads to a better and more accurate forecast of the popularity of short videos, which advances social interaction and digital media. In this work, a deep learning regression model-based system for predicting the popularity of short videos is presented, taking into consideration multi-modal feature fusion supervision modelling and carefully accounting for feature interactions. This study offers more informational dimensions for short video prediction by using IoT data.

DISADVANTAGES

- The inability of language models to comprehend textual data in films, including titles, descriptions, and user comments, is a drawback.
- The CPRP-CNN Model was not implemented in an existing system to process the stochastic distribution of short movies, which attracted a sizable user base when they were released.

PROPOSED SYSTEM

Accurately forecasting the popularity of short videos for social media platforms and content providers has significant monetary value, which is the driving force behind our suggested approach. Strategic content planning, which draws in more viewers and advertisers and eventually boosts income, is made possible by predicting which videos are most likely to become popular. Additionally, knowing how user behaviour varies by geography and culture makes it easier to offer tailored content, which raises user happiness. By revealing the potential uses of IoT technology and deep learning in cross-cultural communication research, this field of study advances our understanding of how information is disseminated and how social interactions differ and are similar across cultural backgrounds.

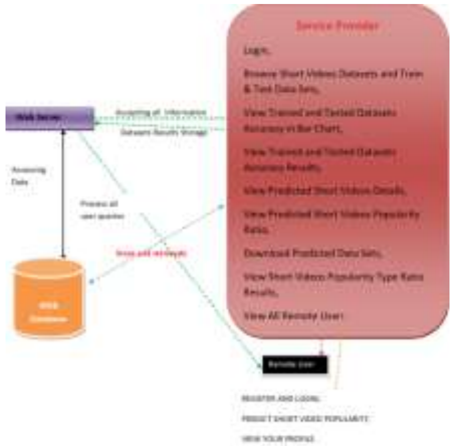
This study's main goal is to provide a strategy for forecasting short films' level of popularity in the context of cross-cultural communication. This project aims to provide useful insights and solutions for content providers, social media platforms, academic research, and the larger user community by integrating the Internet of Things (IoT) with a deep learning regression model.

ADVANTAGES

- The suggested solution used a multi-modal fused regression model to predict content popularity.
- Convolutional layers are used in the CPRP-CNN model to extract information from short movies' text. The model can extract distinct phrases and semantic information from the text by adjusting the convolutional kernel sizes and strides.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
ISSN: 3068-272X www.ijdim.com Original Research Paper

SYSTEM ARCHITECTURE



IV. IMPLEMENTATIONS

Modules

Service Provider

The Service Provider must use a working user name and password to log in to this module. After successfully logging in, he can perform certain tasks such Examine Train & Test Data Sets and Short Video Datasets. See the Accuracy of Trained and Tested Datasets in a Bar Chart, See the Accuracy Results of Trained and Tested Datasets, See the Details of Predicted Short Videos, See the Popularity Ratio of Predicted Short Videos, Download Predicted Data Sets, See the Popularity Type Ratio Results of Short Videos, and See All Remote Users.

View and Authorize Users

The administrator may see a list of all registered users in this module. Here, the administrator may see the user's information, like name, email, and address, and they can also grant the user permissions. The administrator may see a list of all registered users in this module. Here, the administrator may see the user's information, like name, email, and address, and they can also grant the user permissions.

Remote User

A total of n users are present in this module. Before beginning any actions, the user needs register. Following registration, the user's information will be entered into the database. Following a successful registration, he must use his password and authorised user name to log in. Following a successful login, the user may perform tasks including registering and logging in, predicting the popularity of short videos, and seeing their profile.

ALGORITHMS

Naïve Bayes

The supervised learning technique known as the "naive bayes approach" is predicated on the straightforward premise that the existence or lack of a certain class characteristic has no bearing on the existence or nonexistence of any other feature. However, it seems sturdy and effective in spite of this. It performs similarly to other methods of guided learning. Numerous explanations have been put forth in the literature. We emphasise a representation bias-based explanation in this lesson. Along with logistic regression, linear discriminant analysis, and linear SVM (support vector machine), the naive bayes classifier is a linear classifier. The technique used to estimate the classifier's parameters (the learning bias) makes a difference.

Although the Naive Bayes classifier is commonly used in research, practitioners who wish to get findings that are useful do not utilise it as often. On the one hand, the researchers discovered that it is very simple to build and apply, that estimating its parameters is simple, that learning occurs quickly even on extremely big databases, and that, when compared to other methods, its accuracy is rather excellent. The end users, however, do not comprehend the value of such a strategy and do not receive a model that is simple to read and implement.

As a consequence, we display the learning process's outcomes in a fresh way. Both the deployment and comprehension of the classifier are simplified. We discuss several theoretical facets of the naive bayes classifier in the first section of this lesson. Next, we use Tanagra to apply the method on a dataset. We contrast the outcomes (the model's parameters) with those from other linear techniques including logistic regression, linear discriminant analysis, and linear support vector machines. We see that the outcomes are quite reliable. This helps to explain why the strategy performs well when compared to others. We employ a variety of tools (Weka 3.6.0, R 2.9.2, Knime 2.1.1, Orange 2.0b, and RapidMiner 4.6.0) on the same dataset in the second section. Above all, we make an effort to comprehend the outcomes.

Random Forest

Random forests, also known as random decision forests, are ensemble learning techniques that build a large number of decision trees during training for tasks like regression and classification. The class chosen by the majority of trees is the random forest's output for classification problems. The mean or average forecast of each individual tree is given back for regression tasks. The tendency of decision trees to overfit to their training set is compensated for by random decision forests. Although random forests are less accurate than gradient enhanced trees, they often perform better than choice trees. However, their performance may be impacted by data peculiarities.

Tin Kam Ho[1] developed the first algorithm for random decision forests in 1995 by utilising the random subspace technique, which in Ho's definition is a means of putting Eugene Kleinberg's "stochastic discrimination" approach to classification into practice.

Leo Breiman and Adele Cutler created an algorithm extension and filed for a trademark in 2006 for "Random Forests" (owned by Minitab, Inc. as of 2019). The extension builds a set of decision trees with controlled variance by combining Breiman's "bagging" concept with random feature selection, which was initially proposed by Ho[1] and then separately by Amit and Geman[13].

Businesses commonly employ random forests as "blackbox" models since they need little configuration and produce accurate forecasts across a variety of inputs.

SVM

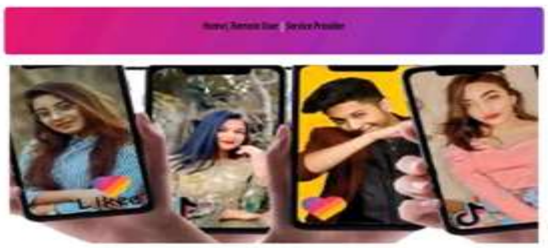
The goal of a discriminant machine learning approach in classification problems is to identify a discriminant function that can accurately predict labels for newly acquired instances based on an independent and identically distributed (iid) training dataset. A discriminant classification function takes a data point x and assigns it to one of the several classes that are part of the classification job, in contrast to generative machine learning techniques that call for calculations of conditional probability distributions. Discriminant techniques are less effective than generative approaches, which are mostly employed when prediction entails the identification of outliers. However, they need less training data and processing resources, particularly when dealing with a multidimensional feature space and when just posterior probabilities are required. Finding the equation for a multidimensional surface that optimally divides the various classes in the feature space is the geometric equivalent of learning a classifier.

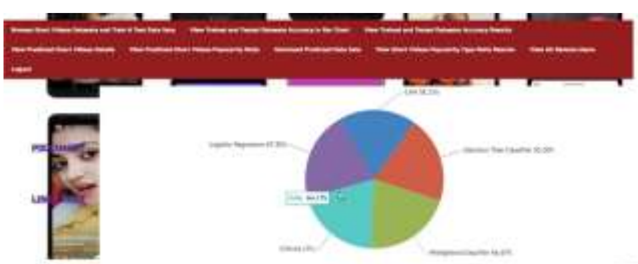
SVM is a discriminant approach that, unlike genetic algorithms (GAs) or perceptrons, which are both often used for classification in machine learning, always returns the same optimal hyperplane value since it solves the convex optimisation issue analytically. The initialisation and termination criteria have a significant impact on the solutions for perceptrons. While the perceptron and GA classifier models are distinct every time training is started, training yields uniquely specified SVM model parameters for a given training set for a certain kernel that converts the data from the input space to the feature space. The sole goal of GAs and perceptrons is to reduce

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
ISSN: 3068-272X www.ijdim.com Original Research Paper

training error, which will result in several hyperplanes satisfying this criterion.

V. SCREEN SHOTS





VI. CONCLUSION

In this study, DeepVidNet demonstrates the potential of integrating IoT data streams with deep neural architectures to predict short video popularity with high precision. By bridging the gap between user behavior, contextual awareness, and visual features, the system enables a multi-dimensional understanding of digital content virality. The hybrid CNN-LSTM model ensures adaptability to temporal changes, while IoT sensors enrich the dataset with real-world contextual cues. Future work will explore incorporating attention mechanisms and federated learning to enhance model interpretability and privacy preservation. The proposed DeepVidNet framework paves the way for intelligent and proactive social media analytics in the era of smart digital ecosystems.

REFERENCES

[1] K.-M. Giannakopoulou, I. Roussaki, and K. Demestichas, ‘‘Internet of Things technologies and machine learning methods for Parkinson’s disease diagnosis, 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-aware and flexible proactive caching using LSTM and ensemble

learning in IoTMEC networks,'' 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 machine learning techniques in medical data processing based on distributed computing and 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 smart crop 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, and S. A. Mostafa, ''Tree-based machine learning algorithms in the Internet of 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, ''Detecting video 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 learning for 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 stream data 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 on deep 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 efficient deep 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, and J. Ahmad, ''Deep learning for the industrial Internet of Things (IIoT): A comprehensive survey of techniques, implementation frameworks, potential applications, 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-driven network 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 Things to improve efficacy of diagnosis and remote sensing of solar photovoltaic systems: 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, ''Deep reinforcement learning approaches for content caching in cache-enabled D2D 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, ''A topical review on machine learning, software defined networking, Internet of 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

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
ISSN: 3068-272X www.ijdim.com Original Research Paper

learning in the Internet of Things: Designed techniques for smart cities,'' 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 in Internet 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 prediction of 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 meets AI: 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 temporal bipartite 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 elections in Latin America based on social media engagement and polls,'' Government Inf. 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 learning architecture for music popularity prediction,'' IEEE Access, vol. 8, pp. 39361–39374, 2020.

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