



Mining of Movie Box Office and Movie Review Topics Using Social Network Big Data

Yinchang Chen¹ and Zhe Dai^{2*}

¹School of Media and Law, NingboTech University, Ningbo, China, ²Research Institute of Theatre Film and Television, Communication University of Zhejiang, Hangzhou, China

OPEN ACCESS

Edited by:

Sang-Bing Tsai,
Wuyi University, China

Reviewed by:

Shaun Yuxiang Wang,
Johns Hopkins University,
United States
Xuan Chen,
Dongbang Culture University,
South Korea

*Correspondence:

Zhe Dai
dz0226@126.com

Specialty section:

This article was submitted to
Organizational Psychology,
a section of the journal
Frontiers in Psychology

Received: 24 March 2022

Accepted: 03 May 2022

Published: 26 May 2022

Citation:

Chen Y and Dai Z (2022) Mining of
Movie Box Office and Movie Review
Topics Using Social Network Big
Data.
Front. Psychol. 13:903380.
doi: 10.3389/fpsyg.2022.903380

In order to solve the problems of high investment and low box office losses in the film industry, this study analyzes the topic of film box office and film and television reviews based on social network big data. Firstly, the factors that affect the box office of the movie are analyzed. Secondly, continuous and discrete feature parts, text parts, and fusion parts are merged. The box office prediction model of mixed features using deep learning is established, and the movie box office is predicted. Finally, compared with other algorithms and models, the box office prediction model of mixed features using deep learning is verified. The results show that compared with other models, the prediction accuracy of the mixed feature movie box office prediction model using depthwise separable convolution (DSC)-Transformer is higher than that of other algorithm models. Its optimal mean square error (MSE) value is 0.6549, and the optimal mean absolute error (MAE) value is 0.1706. The constructed model predicts the box office of nine movies, and the error between the predicted value and the true value is about 10%. Therefore, the established movie box office prediction model has a good effect. This study can predict movies' box office to reduce investment risk, so it is of great significance to movie investors and the social economy.

Keywords: movie box office, deep learning, big data, social network, box office prediction

INTRODUCTION

With the popularization of social networks, more and more users obtain various information through social network platforms, including news, entertainment, education, and orders. Meanwhile, users can also send comments and opinions through platform. Especially Weibo, as a very popular network platform, can be created and utilized. The timely extraction of valuable information from these data has become the research content of various industries (Palomba, 2020). Due to the development of social economy, people's living standards are constantly improving. They are beginning to pay more attention to spiritual pursuits, and entertainment consumption is becoming more and more important in their lives. Movies account for a large part of entertainment consumption. Electrophoresis not only enriches people's leisure life and spiritual world but also is an important means of cultural exchanges between countries (Galvão and Henriques, 2018; Ru et al., 2018).

At present, big data and cloud computing have become the most basic data forms and scientific computing methods in the current society. They not only save a lot of time and effort for users but also provide decision-making support for industries and enterprises. Big data are imported into the film industry from many aspects and plays an important role in film investment and financing, content creation, publicity and distribution, cinema terminals, and derivatives development. With the rapid development of the film industry, film production is also increasing. Movies are a high-risk industry, movie box office. It is not only affected by factors, such as movie plot, movie quality, screening time, movie scale, etc., but also by previous publicity, word of mouth, leading actors, and directors. Therefore, by predicting the box office of the movie, it is possible to adjust the initial investment, such as movie promotion, which can reduce the investment (Ahmed et al., 2020; Bogaert et al., 2021). Predicting the box office revenue of a movie before it is shown on the big screen has become an emerging demand. Forecasting movie box office can provide information for stock market investment decision-making, advertising company's promotion strategy design, and movie theater broadcast density. This task is very challenging, and it is affected by many complex factors (Gaenssle et al., 2018; Franses, 2021; Kang, 2021). In the current era of big data, deep learning technology has a certain role in promoting massive data analysis. Wang et al. (2020) conducted a strategic investigation on the factors affecting the box office of movies and proposed a new framework composed of a series of feature learning models and prediction and ranking models. They used big data to model these factors to predict movie box office revenue. There are two specific learning feature models: (1) the new dynamic heterogeneous network embedding model can simultaneously learn the potential ideas of participants, directors, and companies; meanwhile, it can also capture their partnership and (2) the deep neural network-based model aims to reveal advanced representations of movie quality from the trailer. Using the learned features, train a mutually enhanced prediction and ranking model to obtain box office prediction results. Finally, the framework is applied to the Chinese film market. Comprehensive performance evaluation is performed using real data. The results show that the extracted knowledge and prediction results have good performance. There are many factors that affect the box office of a movie. The life cycle of a movie includes production, distribution, and exhibition factors. The most important costs are related to production factors. However, the title of a movie is one of the relatively inexpensive production factors that studios can use (Kim et al., 2018; Liu and Xie, 2019). Bae and Kim (2019) examined an informative movie title, that is, a movie title containing movie genre or storyline information. A 5-year analysis of the Korean film market shows that: information-rich movie titles have a positive impact on the box office revenue of under-advertised movies. Among them, the pre-release publicity activities are measured by the amount of media exposure before the release. From the perspective of selected features, the feature factors proposed by different scholars are also different. There are also different views on

the influence of the same feature factors on the movie box office. This shows that the factors affecting movies' box office and their mechanisms are very complex. Predicting movie box office is very difficult. When deep learning models are used to make predictions, they often only use simple BP neural networks and LSTM models, which cannot learn various features well.

One of the things that makes a movie good or bad is its box office. A high box office indicates that the movie is good, and investors will also consider the risks and benefits of the movie when investing in the movie. Therefore, if the model can accurately predict the box office of a movie, it can reduce the huge losses caused by investment risks to a certain extent and adjust reasonable shooting, production, publicity, and distribution strategies to maximize the return on investment. In order to solve this problem, a mixed feature movie box office prediction model based on deep learning is established to predict movie box office. It can reduce investment risk, which is significant to film investors and the social economy. The research idea is based on the problem that the encoding part of the existing Transformer model uses a fully connected feedforward neural network, which is computationally expensive and difficult to extract local invariant features. An improved depthwise separable convolution (DSC)-Transformer model is proposed. The depthwise separable convolution replaces the fully connected feedforward neural network in the encoding part and combined with the bidirectional encoder representation from transformers (BERT) model and the depthwise separable convolution, a hybrid feature prediction model is formed to predict and analyze the movie box office.

MOVIE BOX OFFICE PREDICTION ALGORITHM AND MODEL DESIGN

Analysis of Factors Affecting Movie Box Office

There are many factors that affect the box office of a movie. In movie box office prediction, some influencing factors are usually selected as variables. Not all factors can be used to predict the box office of a movie. For example, because the information is kept confidential, the producer will not disclose it online before the movie is released. There are also some factors such as search popularity, Internet word-of-mouth, and movie ratings that cannot be predicted. However, the characteristics of different characteristic data are different, so the data will also have different processing methods (Vujić and Zhang, 2018; Bao et al., 2021; Liao and Huang, 2021).

1. Director, actor, and screenwriter. The director plays an important role in the box office of a movie and determines the style, standard and quality of a movie. Actors are a key factor at the box office. An actor's acting skills and influence will affect the box office. Screenwriting is also critical. This directly determines the direction and logic of the entire plot. If the plot of a movie is chaotic and unclear, it will directly affect the reputation of the movie. But these

three factors cannot be directly used as variables to predict, so they need to be converted, as in Eq. (1):

$$p_i = \frac{\sum_{j=1}^m b_j}{m} \quad (1)$$

p is the filmmaker (director, editor, and actor), m is the m movies directed, written, and participated in by the filmmaker, i is the filmmaker's serial number, j is the j -th movie, and b_j is the j -th movie box office.

2. Production and distribution companies. The quality and publicity of the film depends on the production company and the distribution company. A good production company can guarantee the quality of movies, such as Huayi Brothers Media Group, Wanda Media Co. Ltd., August First Film Studio, etc. The greater the publicity of the movie by the distribution company, the more attention it will receive from the public, which will also have a certain impact on the box office.
3. Movie type. Everyone's ability to appreciate movies is different. In the field of film and television, film and television appreciation are usually divided into cognitive appreciation, critical appreciation, and ablative appreciation. After investigation, the box office of science fiction, comedy, and action genres will be higher. The type of movie will also have a certain impact on the box office.

In addition, movie online word-of-mouth will also have an impact on movie box office. In the field of professional film

and television appreciation, the plot and conflict of the film, the lens and the picture, the relationship between the characters in the film, the performance of the actors, the dubbing of the film and the music are more considered. Under normal circumstances, many ordinary viewers will post movie word-of-mouth reviews in online communities. The audience will comment using their viewing experience (Dewani et al., 2021). Most online film reviews contain spoilers. This spoiler information is information to resolve the uncertainty of the plot in advance. Ryoo et al. (2021) used the box office revenue and online word-of-mouth data of movies released in the United States to study the consequences of spoiler comments. In order to capture the degree of information in the spoiler review text to reduce the uncertainty of the drawing, a spoiler strength measurement is proposed. Related topic models are used and measured.

Improved DSC-Transformer

In 2017, Google proposed the Transformer model. This model can be used to extract features. It can solve the problem of slow training of recurrent neural network (RNN). The Transformer model structure is shown in **Figure 1**.

The Transformer model structure is divided into two parts: encoding and decoding. Among them, the encoding part includes six encoders, and the decoding part also includes six decoders. **Figure 2** shows the structure of the encoder and decoder.

In **Figure 2**, the encoding part of the Transformer model uses a fully connected feedforward neural network. Fully-connected feedforward neural networks have problems such as many calculation parameters and difficulty in extracting local invariant features. Therefore, the model needs to

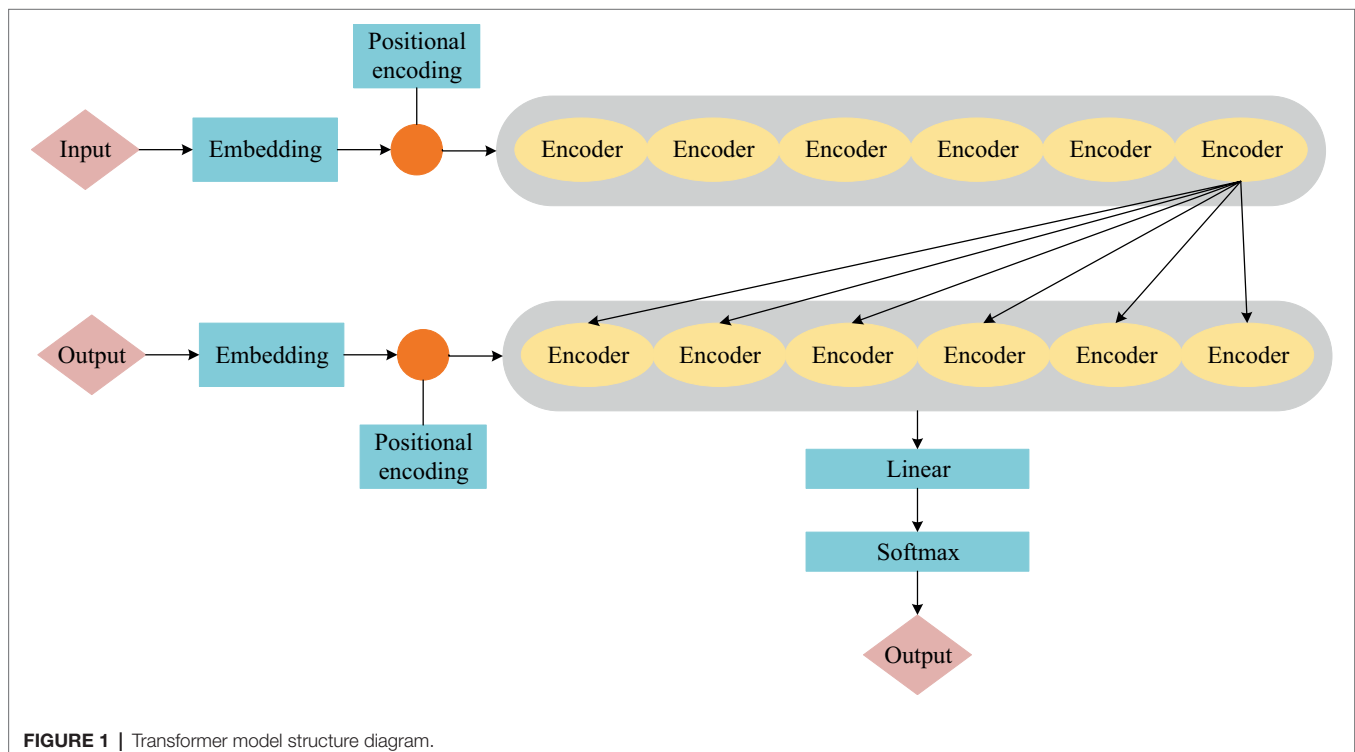


FIGURE 1 | Transformer model structure diagram.

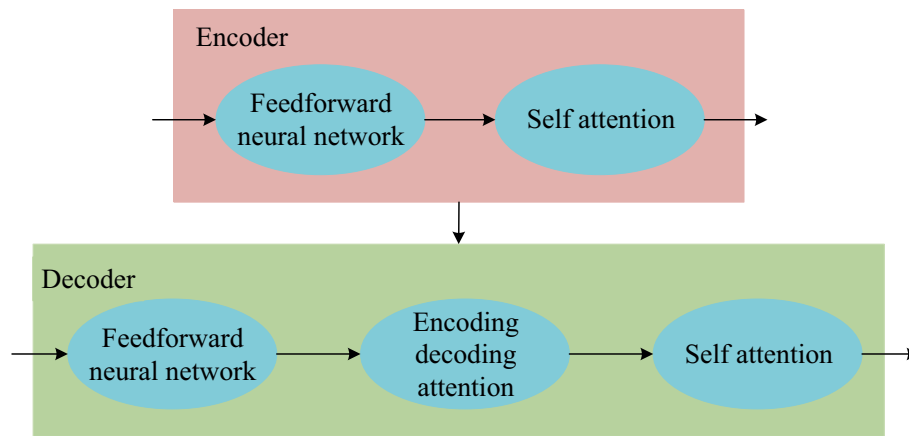


FIGURE 2 | Partial structure of encoder and decoder.

be improved (Schwaller et al., 2019; Mousavi et al., 2020). DSC is used to replace the fully connected feedforward neural network. DSC is a model with few parameters and low computational cost for feature extraction. The model is composed of depthwise convolution (DW) and pointwise convolution (PW; Bai et al., 2018; Liu et al., 2019). The process comparison between DSC and standard convolution is shown in **Figure 3**.

Assume that the side length of the convolution kernel is D_K . The number of input and output channels are M and N , respectively. The side length of the input feature map is D_F . The calculation amount N_{std} of standard convolution can be expressed as Eq. (2):

$$N_{std} = D_K \cdot D_K \cdot M \cdot D_F \cdot D_F \quad (2)$$

When using DSC calculation, filter the DW part, and perform channel conversion on the PW part. Their sizes are $(D_K, D_K, 1)$ and $(1, 1, M)$ respectively. The numbers are M and N , respectively. These two quantities are calculated as Eqs. (3) and (4):

$$N_{depthwise} = D_K \cdot D_K \cdot M \cdot D_F \cdot D_F \quad (3)$$

$$N_{pointwise} = M \cdot N \cdot D_F \cdot D_F \quad (4)$$

The total calculation amount can be expressed as Eq. (5):

$$N_{ds} = N_{depthwise} + N_{pointwise} \quad (5)$$

The ratio of the calculated amount of DSC and standard convolution is obtained, as in Eq. (6):

$$\frac{N_{ds}}{N_{std}} = \frac{1}{N} + \frac{1}{D_K^2} \quad (6)$$

DSC greatly reduces the amount of calculation and effectively reduces the consumption of computing resources. Therefore, using DSC to improve the Transformer model can achieve good calculation efficiency and improve accuracy (Kamal et al., 2019; Shang et al., 2020). The data are extracted with features, and then input into the DSC-Transformer model. The self-attention layer processes the input value to obtain a weighted feature vector. This vector and the original input form a skip connection structure. Then, the output of the previous layer is taken as the input into the DSC layer, and the final output of the DSC-Transformer model can be obtained after a series of summation and normalization processing. The structure of the DSC-Transformer model is shown in **Figure 4**.

BERT Model

BERT is a two-stage model. The first stage is to use the bidirectional language model for pre-training. The second stage uses a fine-tuning mode to solve specific downstream tasks (Cai et al., 2020; Mozafari et al., 2020; Liu et al., 2021; Yu et al., 2021). The BERT model uses a large amount of data for pre-training and combines the advantages of various models. Therefore, this model has achieved good results in many natural language processing tasks. The BERT model structure is shown in **Figure 5**. The first task is to randomly select some words to predict when inputting a sentence and then replace them with a special symbol. After that, the model learns the words that should be filled in these places according to the given labels. The second task adds a sentence-level continuity prediction task to the bidirectional language model, that is, predicting whether the two texts input to BERT is continuous texts. This task is introduced to allow the model to learn the relationship between consecutive text segments.

Design of Box Office Prediction Model for Mixed-Feature Movies

The predictive features used include text features, continuous features, and discrete features. Therefore, these features need

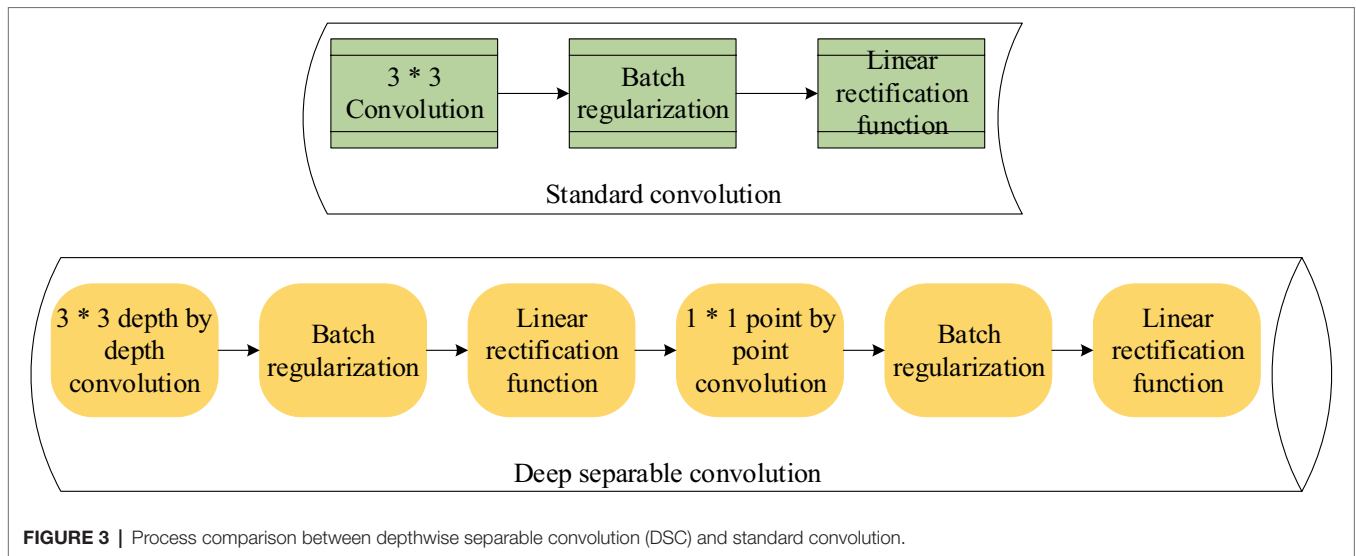


FIGURE 3 | Process comparison between depthwise separable convolution (DSC) and standard convolution.

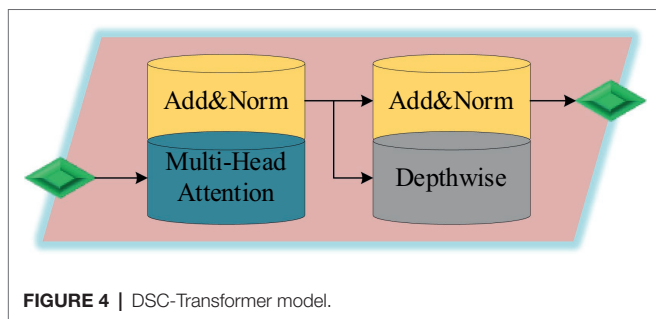


FIGURE 4 | DSC-Transformer model.

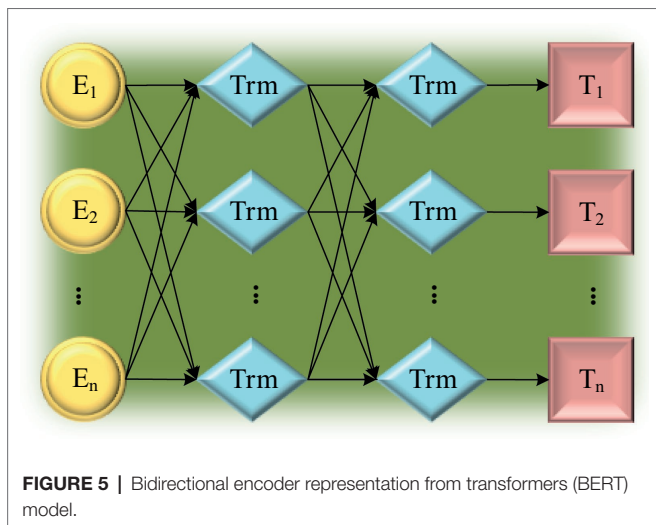


FIGURE 5 | Bidirectional encoder representation from transformers (BERT) model.

to be handled separately. Among them, the ideas of factorization machines (FM) and field-aware factorization machines (FFM) are used to process the discrete and continuous feature data. Firstly, the continuous and discrete features are, respectively, mapped to the first fully connected layer. The characteristics of the respective fields are constituted. These features are

regarded as a learner and are all integrated into the second fully connected layer. A more powerful learner is formed.

For the text part, the BERT model is used to extract features. Two parts of the result are obtained. Part of the result is the high-level text features of the serialized output. This part of the result is learned using a two-layer DSC-Transformer. The sequence feature of the text is obtained. However, another part of the result is the overall characteristics of the text obtained from the BERT model. Finally, the learning results of the two parts are spliced together and sent as input to the final fusion part. After two DSCs, the prediction result will be obtained. Continuous and discrete feature parts, text parts, and fusion parts are merged. The final mixed feature movie box office prediction model is composed.

Data Acquisition and Processing Feature Data Selection

Movie feature data are mainly from the piaofang.maoyan.com and movie.douban.com between 2017 and 2020. The selected feature information is shown in Table 1.

Data Acquisition

Selenium-based crawler technology is used to obtain data. Firstly, Chrome Driver is initialized. Then, find and locate elements step by step according to id, tagName, className, etc., until the obtained movie feature value data can be located. For some values that are too small, they can be directly removed, and the data that meet the conditions are stored. A total of more than 40,000 pieces of data have been obtained. These data are stored in Redis. List in Redis is used as a data structure, numbered as key. Each feature value is stored as a value, and these data are preprocessed and saved in csv format.

Loss Function

Usually, the loss function is used to evaluate the calculation method of the model. In deep learning, the loss function is

TABLE 1 | The selected feature information.

	piaofang.maoyan.com page	movie.douban.com page
Feature information	Screening dates, working days and holidays, weekends, the number of days the movie has been released, the cumulative box office so far, the box office of the movie today, the proportion of the box office that day, average attendance, attendance, etc.	Movie name, introduction, director, starring role, genre, Douban rating, number of ratings, number of film reviews, number of short reviews, film length, release time, etc.

used to quantify the difference in probability distribution between the predicted value of the model's output and the actual label value. Different loss functions are applied in different occasions. The loss functions used are MSE and MAE. MSE measures the average of the square of the difference between the model predicted value and the actual value, while MAE is the average of the sum of the absolute values of the difference between the model predicted value and the actual value. Assuming real data $y = (y_1, y_2, \dots, y_n)$, the fitted data are $\hat{y} = (\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n)$. The calculation of MSE and MAE is shown in Eqs. (7) and (8):

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2 \tag{7}$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \tag{8}$$

In general, the receptive field of the structure of one layer is limited, and it is difficult to integrate features of different depths. Multilayer structures are prone to overfitting. Therefore, the use of DW-Transformer models with different layers to learn high-order text features is compared, and the prediction results of the mixed-feature movie box office prediction model are compared. Since neural network models are often complex, computation requires a lot of computing resources and time. Therefore, the problem of efficiency has become an important problem that has plagued the field of deep learning for a long time. Hyperparameter tuning is very important. It can optimize the network model and improve the calculation speed and prediction accuracy. Additionally, the prediction accuracy of the model with different hyperparameters is tested. LSTM and Bi-LSTM are network structures that can process time-series data. Transformer models can be used to extract features. These three models are machine learning commonly used in movie box office predictions. Therefore, the established DSC-Transformer models are compared with them. In addition,

the established model is compared with the prediction results of the machine learning method LightGBM. The established prediction model is used to predict the box office of some movies.

RESULTS AND DISCUSSION

Comparison of DSC-Transformer Structure of Different Layers

Different layers of DSC-Transformer models are used to learn high-level text features. In this case, the prediction results of the entire mixed feature movie box office prediction model are shown in Figures 6, 7.

Figure 6 shows that when a layer of DSC-Transformer is used to learn high-level text features, the corresponding MSE value is the best at 27 iterations, and its value is 0.6878. When using a two-layer DSC-Transformer to learn high-level text features, the corresponding MSE value is the best at 23 iterations, and its value is 0.6549. When using a three-layer DSC-Transformer to learn high-level text features, the corresponding MSE value is the best at 25 iterations, and its value is 0.7281. When using four-layer DSC-Transformer to learn high-level text features, the corresponding MSE value is the best when iterates 32 times, and its value is 0.6848. When using a five-layer DSC-Transformer to learn high-level text features, the corresponding MSE value is the best at 31 iterations, and its value is 0.6643. When the six-layer DSC-Transformer is used to learn high-level text features, the corresponding MSE value is the best at 35 iterations, and its value is 0.6941.

Figure 7 shows that when a one-layer DSC-Transformer is used to learn high-level text features, the corresponding MAE value is the best at 47 iterations, and its value is 0.1877. When using a two-layer DSC-Transformer to learn high-level text

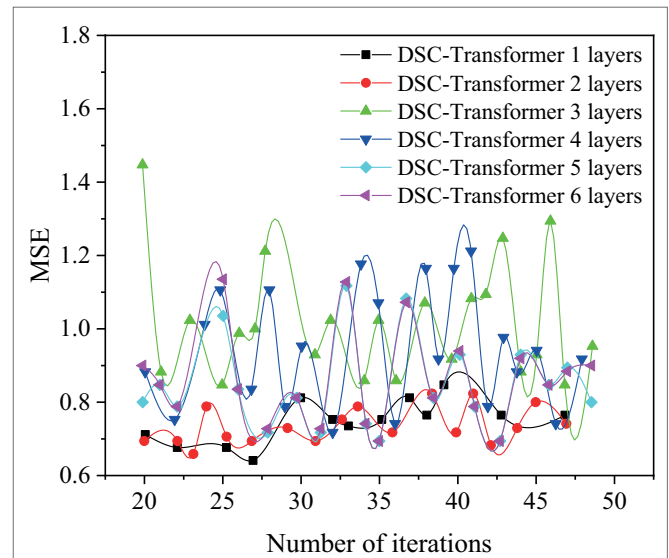
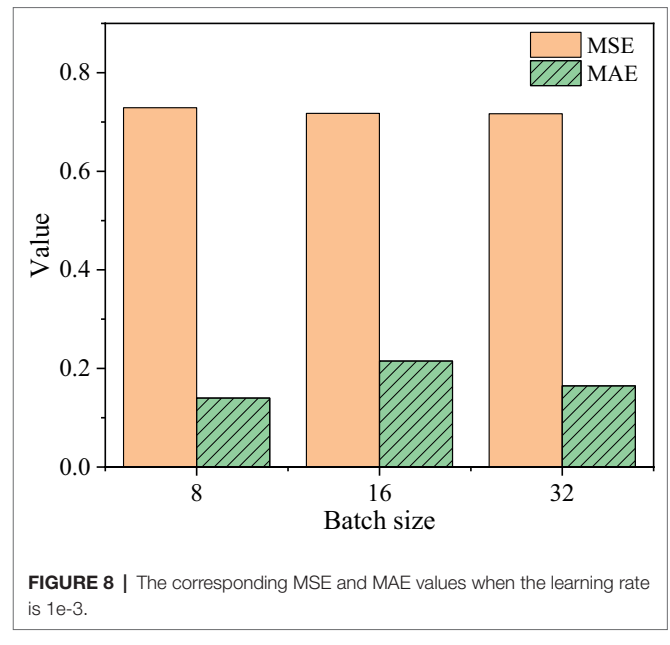
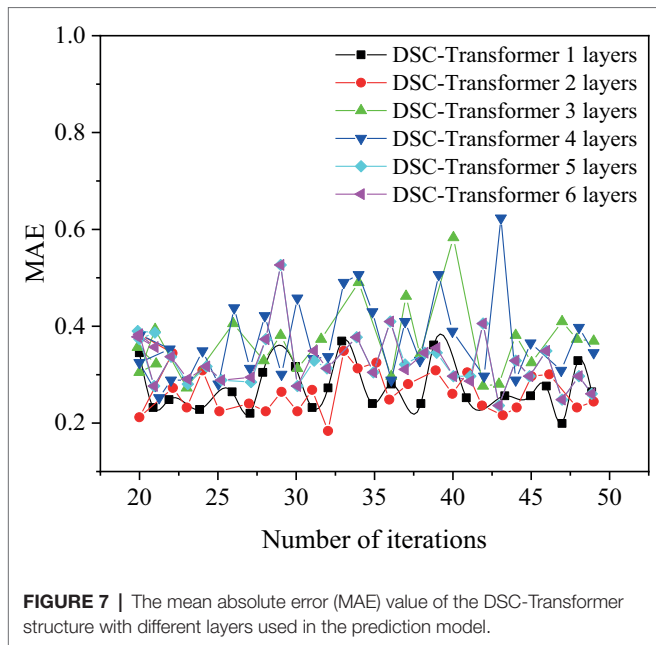
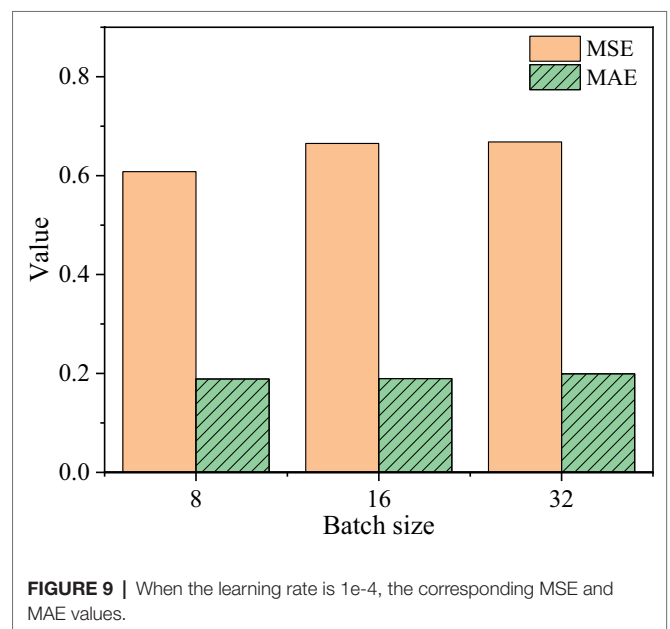


FIGURE 6 | The mean square error (MSE) value of the DSC-Transformer structure with different layers used in the prediction model.



features, the corresponding MAE value is the best at 32 iterations, and its value is 0.1705. When using the three-layer DSC-Transformer to learn high-level text features, the corresponding MAE value is the best at 42 iterations, and its value is 0.2671. When using a four-layer DSC-Transformer to learn high-level text features, the corresponding MAE value is the best at 21 iterations, and its value is 0.2381. When using a five-layer DSC-Transformer to learn high-level text features, the corresponding MAE value is the best at 44 iterations, and its value is 0.2316. When the six-layer DSC-Transformer is used to learn high-level text features, the corresponding MAE value is the best at 43 iterations, and its value is 0.2362. When using the two-layer DSC-Transformer to learn high-level text features, the MSE and MAE values are both optimal, so the two-layer DSC-Transformer is the optimal design of the model.



Hyperparameter Related Experiments

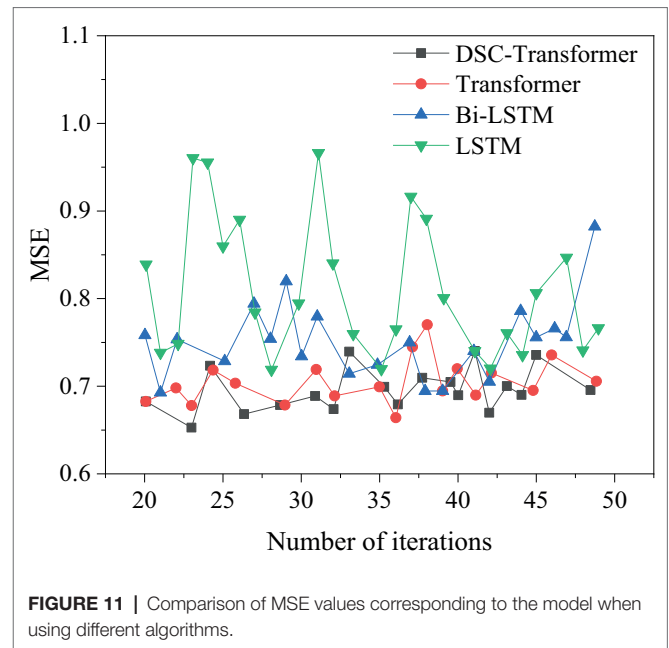
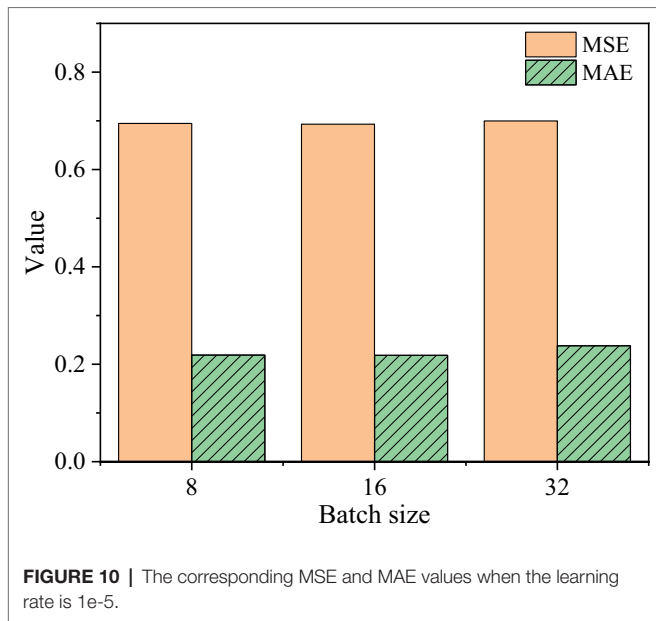
In order to study the influence of the two parameters of learning rate and batch size on the prediction results of the two-layer DSC-Transformer structure learning high-level text feature model, the learning rate is selected as $1e-3$, $1e-4$, and $1e-5$, and the batch size is 8, 16, and 32 samples for experiment. The specific results are shown in **Figures 8–10**.

Figure 8 shows that when the learning rate is $1e-3$ and the batch size is 8, the corresponding MSE value is 0.729 and the MAE value is 0.14. When the batch size is 16, the corresponding MSE value is 0.7176 and the MAE value is 0.2152. When the batch size is 32, the corresponding MSE value is 0.7168 and the MAE value is 0.1646.

Figure 9 shows that when the learning rate is $1e-4$ and the batch size is 8, the corresponding MSE value is 0.6081 and the MAE value is 0.1889. When the batch size is 16, the corresponding MSE value is 0.665 and the MAE value is 0.1896.

When the batch size is 32, the corresponding MSE value is 0.6682 and the MAE value is 0.1993.

Figure 10 shows that when the learning rate is $1e-5$ and the batch size is 8, the MSE value is 0.6945 and the MAE value is 0.219. When the batch size is 16, the corresponding MSE value is 0.693 and the MAE value is 0.2185. When the batch size is 32, the corresponding MSE value is 0.6998, and the MAE value is 0.2379. It can be concluded that: when the learning rate is $1e-4$ and the batch size is 8, the MSE value of the mixed-feature movie box office prediction model is optimal; when the learning rate is $1e-3$ and the batch size is 8, the box office of the mixed-feature movie is 8 The MAE value of the prediction model is optimal.



Comparison of Results Using Different Algorithms

Different algorithm models are used to learn high-level text features, and the corresponding prediction results of the mixed feature movie box office prediction model are shown in Figures 11, 12.

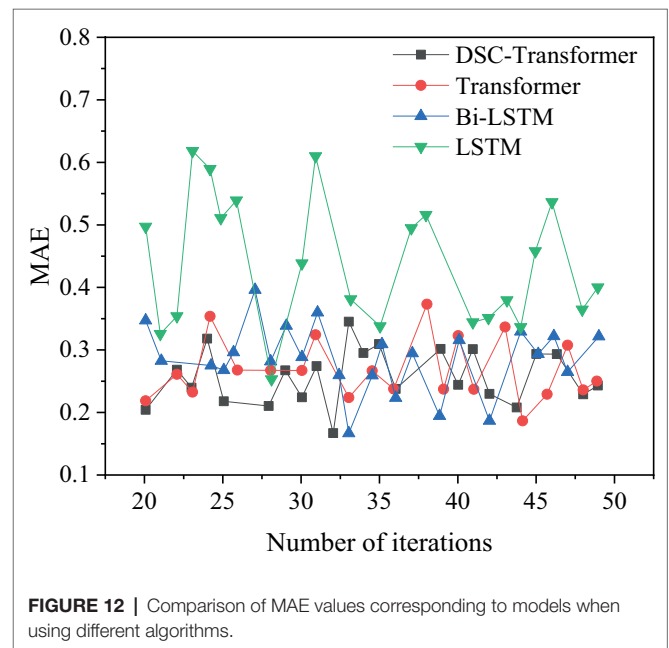
Figure 11 shows that among the four different algorithm models, the optimal MSE value of the model using the DSC-Transformer is 0.6549. The optimal MSE value of the model used by Transformer is 0.6560. The optimal MSE value of the model using Bi-LSTM is 0.6843. The optimal MSE value of the model used by LSTM is 0.7064, and the MSE value of the model used by DSC-Transformer has achieved the best results. The prediction accuracy of the established DSC-Transformer model is higher than that of other algorithm models.

Figure 12 shows that among the four different algorithm models, the optimal MAE value of the model using the DSC-Transformer is 0.1706. The optimal MAE value of the model using Transformer is 0.1866. The optimal MAE value of the model using Bi-LSTM is 0.1758. The optimal MAE value of the model used by the LSTM is 0.3075, and the MAE value of the model used by the DSC-Transformer has achieved the best results. This shows that the prediction accuracy of the mixed feature movie box office prediction model using the DSC-Transformer model is higher than that of other algorithm models.

The box office prediction model of mixed features and LightGBM are used to predict the box office. The comparison results are shown in Figures 13, 14.

Figure 13 shows that when the number of iterations of the mixed feature movie box office prediction model is 23, the MSE value is optimal, and its value is 0.6625. When the number of iterations of LightGBM is 44, the MSE value is optimal, and its value is 0.7489. The mixed-feature movie box office prediction model is 13.04% higher than the prediction result of LightGBM.

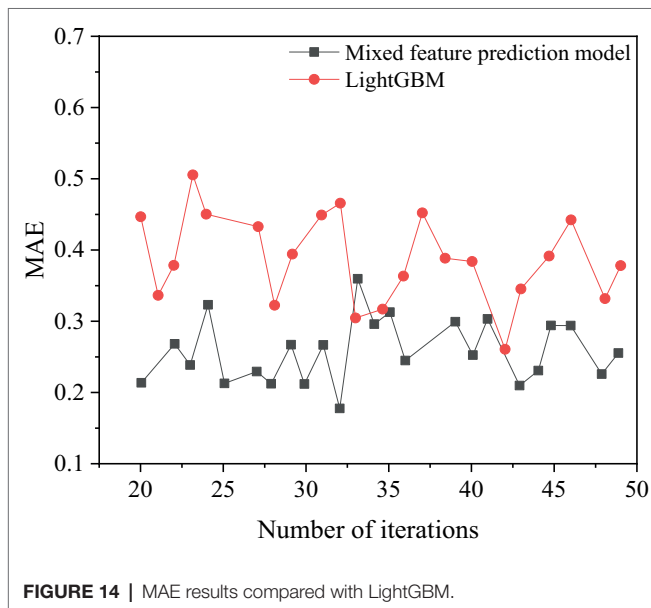
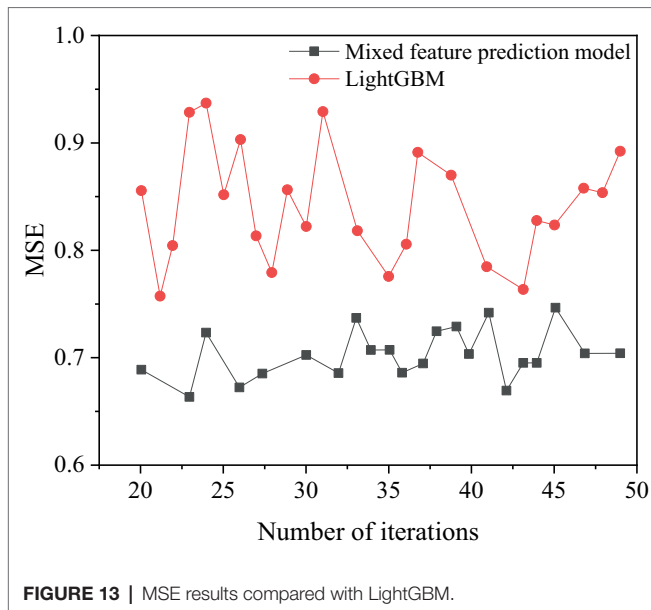
Figure 14 shows that when the number of iterations of the mixed feature movie box office prediction model is 32,



the MAE value is optimal, which is 0.1798. When the number of iterations of LightGBM is 43, the MAE value is optimal, and its value is 0.2581. The movie box office prediction model is 43.55% higher than the prediction result of LightGBM. In summary, the mixed feature movie box office prediction model proposed has higher prediction accuracy than LightGBM and can achieve good prediction results.

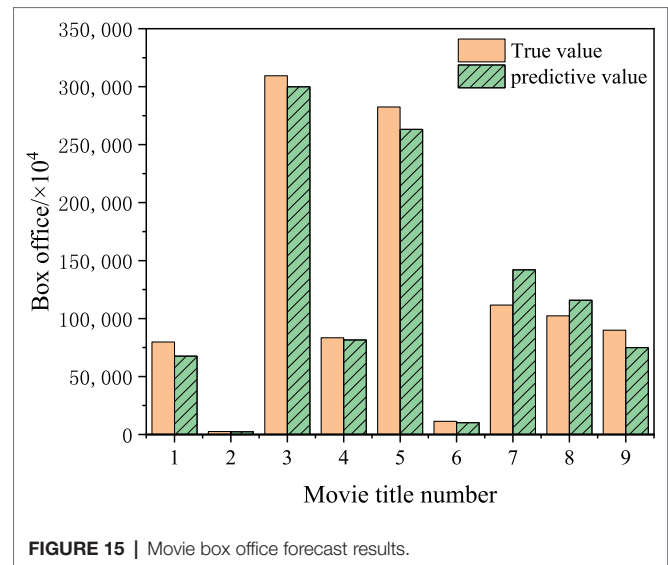
Box Office Forecast

In 2019, action and crime movies *P Storm* released, drama, suspense, crime movies *Hunt Down*, in 2020, war movies *the Eight Hundred* released, sports movies *Leap*, patriotic movies



My People My Homeland, romance movie *the Story of Xi Bao*, war movie *the Sacrifice*, action, crime movie *Shock Wave 2*, drama genre *A Little Red Flower*. These nine movies are numbered 1–9, respectively. The proposed mixed-feature movie box office prediction model is used to predict the box office of these nine movies. The results are shown in **Figure 15**.

Figure 15 shows that the error between the predicted value and the true value of *4P Storm* is 15.3%. The error between the predicted value and the true value of *Hunt Down* is 10.5%. The error between the predicted value and the true value of *the Eight Hundred* is 3.1%. The error between the predicted value and the true value of *Leap* is 2.2%. The error between the predicted value and the true value of *My People My Homeland* is 6.8%. The error between the predicted



value and the true value of *the Story of Xi Bao* is 9.8%. The error between the predicted value and the true value of *the Sacrifice* is 27.2%. The error between the predicted value and the true value of *Shock Wave 2* is 13.2%. The error between the predicted value and the true value of *A Little Red Flower* is 16.7%. Among them, *Leap* has the best prediction effect. *The Sacrifice* had the worst prediction effect. The error between the predicted value and the true value of other films is about 10%, which shows that the movie box office prediction model established has a good predictive effect.

CONCLUSION

In recent years, the film industry has developed rapidly. The film production process itself is complex, involves many links, and takes a long time, making the film industry increasingly risky. Based on social network big data, it is established to analyze the movie box office and film and television review topics. Firstly, the factors that affect the movie box office are analyzed. Secondly, continuous and discrete feature parts, text parts, and fusion parts are merged. A mixed-feature movie box office prediction model based on deep learning is established to predict movie box office. Finally, the mixed feature movie box office prediction model based on deep learning is verified compared with other algorithms and models. When the learning rate is $1e-4$, and the batch size is 8, the MSE value of the mixed-feature movie box office prediction model is optimal. The experimental results of hyperparameter correlation show that when the learning rate is $1e-3$ and the batch size is 8, the MAE value of the mixed feature movie box office prediction model reaches the optimum. Compared with different algorithms, the prediction accuracy of the proposed mixed feature movie box office prediction model is higher than that of other algorithms, and it can achieve a good prediction effect. The

proposed model can predict the box office of movies, thereby reducing investment risks. It is of great significance to film investors and social economy. There are still some shortcomings, such as marketing costs, production costs and other data are more difficult to obtain. Therefore, it is not very comprehensive when choosing the factors affecting the box office of a movie. In future research, some new influencing factors need to be added.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, and further inquiries can be directed to the corresponding author.

REFERENCES

- Ahmed, U., Waqas, H., and Afzal, M. T. (2020). Pre-production box-office success quotient forecasting. *Soft. Comput.* 24, 6635–6653. doi: 10.1007/s00500-019-04303-w
- Bae, G., and Kim, H. (2019). The impact of movie titles on box office success. *J. Bus. Res.* 103, 100–109. doi: 10.1016/j.jbusres.2019.06.023
- Bai, L., Zhao, Y., and Huang, X. (2018). A CNN accelerator on FPGA using depthwise separable convolution. *IEEE Trans. Circuits Syst. II: Express Br.* 65, 1415–1419. doi: 10.1109/TCSII.2018.2865896
- Bao, T., Kim, H., and Chang, B. H. (2021). An exploratory study on influencing factors of film equity crowdfunding success: based on Chinese movie crowdfunding. *The Journal of the Korea Contents Association* 21, 1–14. doi: 10.5392/JKCA.2021.21.02.001
- Bogaert, M., Ballings, M., Van den Poel, D. and Oztekin, A. (2021). Box office sales and social media: a cross-platform comparison of predictive ability and mechanisms. *Decis. Support. Syst.* 147:113517. doi: 10.1016/j.dss.2021.113517
- Cai, L., Song, Y., Liu, T., and Zhang, K. (2020). A hybrid BERT model that incorporates label semantics via adjustive attention for multi-label text classification. *IEEE Access* 8, 152183–152192. doi: 10.1109/ACCESS.2020.3017382
- Dewani, P. P., Malhan, M., and Mathur, S. (2021). Box office collection of sequel movies: exploring brand extension effect. *J. Glob. Bus. Adv.* 14, 288–311. doi: 10.1504/JGBA.2021.116719
- Franses, P. H. (2021). Modeling box office revenues of motion pictures. *Technol. Forecast. Soc. Chang.* 169:120812. doi: 10.1016/j.techfore.2021.120812
- Gaenssle, S., Budzinski, O., and Astakhova, D. (2018). Conquering the box office: factors influencing success of international movies in Russia. *Rev. Netw. Econ.* 17, 245–266. doi: 10.1515/rne-2019-0017
- Galvão, M., and Henriques, R. (2018). Forecasting movie box office profitability. *Journal of Information Systems Engineering & Management* 3, 1–9. doi: 10.20897/jisem/2658
- Kamal, K., Yin, Z., Wu, M., et al. (2019). Depthwise separable convolution architectures for plant disease classification. *Comput. Electron. Agric.* 165:104948. doi: 10.1016/j.compag.2019.104948
- Kang, D. (2021). Box-office forecasting in Korea using search trend data: a modified generalized bass diffusion model. *Electron. Commer. Res.* 21, 41–72. doi: 10.1007/s10660-020-09456-7
- Kim, Y., Kang, M., and Jeong, S. R. (2018). Text mining and sentiment analysis for predicting box office success. *KSII Trans. Internet Inf. Syst.* 12, 4090–4102. doi: 10.3837/tiis.2018.08.030
- Liao, L., and Huang, T. (2021). The effect of different social media marketing channels and events on movie box office: An elaboration likelihood model perspective. *Inf. Manag.* 58:103481. doi: 10.1016/j.im.2021.103481
- Liu, Z., Lang, L., Li, L., Zhao, Y., and Shi, L. (2021). Evolutionary game analysis on the recycling strategy of household medical device enterprises under government dynamic rewards and punishments. *Math. Biosci. Eng.* 18, 6434–6451. doi: 10.3934/mbe.2021320
- Liu, Y., and Xie, T. (2019). Machine learning versus econometrics: prediction of box office. *Appl. Econ. Lett.* 26, 124–130. doi: 10.1080/13504851.2018.1441499

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work, and approved it for publication.

FUNDING

This work was supported by the National Social Science Fund of China of the Youth Project A Study of the Overseas Communication of Chinese Film and Television to Shape National Image (no. 20cxw010) and the National Ethnic Affairs Commission of the People's Republic of China Project Political Security Governance in Minority Areas from the Perspective of Overall Security (no. 2021-GMB-009).

- Liu, B., Zou, D., Feng, L., Feng, S., Fu, P., and Li, J. (2019). An fpga-based cnn accelerator integrating depthwise separable convolution. *Electronics* 8, 281. doi: 10.3390/electronics8030281
- Mousavi, S. M., Ellsworth, W. L., Zhu, W., Chuang, L. Y., and Beroza, G. C. (2020). Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. *Nat. Commun.* 11(1), –12. doi: 10.1038/s41467-020-17591-w
- Mozafari, M., Farahbakhsh, R., and Crespi, N. (2020). Hate speech detection and racial bias mitigation in social media based on BERT model. *PLoS One* 15:e0237861. doi: 10.1371/journal.pone.0237861
- Palomba, A. (2020). Consumer personality and lifestyles at the box office and beyond: how demographics, lifestyles and personalities predict movie consumption. *J. Retail. Consum. Serv.* 55:102083. doi: 10.1016/j.jretconser.2020.102083
- Ru, Y., Li, B., Liu, J., and Chai, J. (2018). An effective daily box office prediction model based on deep neural networks. *Cogn. Syst. Res.* 52, 182–191. doi: 10.1016/j.cogsys.2018.06.018
- Ryoo, J. H., Wang, X., and Lu, S. (2021). Do spoilers really spoil? Using topic modeling to measure the effect of spoiler reviews on box office revenue. *J. Mark.* 85, 70–88. doi: 10.1177/0022242920937703
- Schwaller, P., Laino, T., Gaudin, T., Bolgar, P., Hunter, C. A., Bekas, C., et al. (2019). Molecular transformer: a model for uncertainty-calibrated chemical reaction prediction. *ACS Cent. Sci.* 5, 1572–1583. doi: 10.1021/acscentsci.9b00576
- Shang, R., He, J., Wang, J., Xu, K., Jiao, L., and Stolkin, R. (2020). Dense connection and depthwise separable convolution based CNN for polarimetric SAR image classification. *Knowl.-Based Syst.* 194:105542. doi: 10.1016/j.knsys.2020.105542
- Vujić, S., and Zhang, X. (2018). Does twitter chatter matter? Online reviews and box office revenues. *Appl. Econ.* 50, 3702–3717. doi: 10.1080/00036846.2018.1436148
- Wang, Z., Zhang, J., Ji, S., Meng, C., Li, T., and Zheng, Y. (2020). Predicting and ranking box office revenue of movies based on big data. *Inf. Fusion.* 60, 25–40. doi: 10.1016/j.inffus.2020.02.002
- Yu, H., Zhao, Y., Liu, Z., Liu, W., Zhang, S., Wang, F., et al. (2021). Research on the financing income of supply chains based on an E-commerce platform. *Technol. Forecast. Soc. Change* 169:120820. doi: 10.1016/j.techfore.2021.120820

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Chen and Dai. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.