ORIGINAL PAPER

DOI: 10.26794/2587-5671-2025-29-6-243-268
UDC 338.467,339.13,330.131.5(045)
JEL C51, C53, L82, E22, G11

Forecasting the Financial Efficiency of Russian Cinema Using a Multifactor Ensemble Machine Learning Model Trained on Historical Data

A.V. Dozhdikov
Institute of Socio-Political Studies of the Russian Academy of Sciences, Moscow, Russian Federation

ABSTRACT

The object of the study is data on the distribution of Russian cinema films from July 2022 to September 2023. Specifically, it analyzes 185 films that were released during this period. The research tool consists of 26 and 146-factor machine learning models that have been pre-trained based on previous periods (from 2004 to July 2022, with 1,500 films). **The purpose** of the study is to demonstrate that machine learning models, trained on historical data, can accurately predict future data, which is especially important for funding programs aimed at developing national cinema in the Russian Federation and attracting private investment, in light of the departure of foreign film distributors from the film market. The study used **methods** to evaluate film projects based on their historical profitability using rental indicators and the characteristics of the creative teams involved in producing them. The emphasis is on ensemble models — AdaBoost, Bagging, ExtraTrees, GradientBoosting, RandomForest, Stacking, Voting, XGBoost, CatBoost. **The novelty** of this research lies in introducing of new sources into the scientific community and the potential for practical application of the developed methods for both public and private investors to evaluate film projects prior to the start of the production cycle. **Conclusions:** Conclusions: Based on the analysis of the quality metrics (accuracy, ROC AUC, and others) for a sample of 185 newly released films (through September 2023), we found that the drop in these metrics was not significant. This suggests that it is possible to use pre-trained models based on historical data to make predictions about fees and other rental outcomes. By analyzing the past work of the project director, screenwriters, cameramen, producers, artists, editor, composer and key actors of the project, estimated distribution data, and the amount of project funding, it is possible to make an accurate prediction about the success of a film. This will allow you to see the total fees, payback period, number of views, and viewer rating.

Keywords: financial efficiency; film box office forecast; box office data; national cinema; machine learning; ensemble models; classification; regression; CatBoost; XGBoost; government policy in cinema; export of film content

For citation: **Dozhdikov A.V. Forecasting the financial efficiency of Russian cinema using a multifactor ensemble machine learning model trained on historical data. *Finance: Theory and Practice.* 2025;29(6):243-268. DOI: 10.26794/2587-5671-2025-29-6-243-268**

INTRODUCTION

At the stage of development (pre-production), investors evaluate from several dozen scripts to several hundred and thousands in large studios, select key actors, form creative and film crews, and determine the budget. It is at this stage that it is advisable to consider various scenarios and combinations of creative and film crews before investing a large amount of money in the production of not even one film, but a carefully formed and calculated "portfolio" of film projects. An accurate income forecast at this early stage significantly reduces investment risk [1]. The huge initial investments associated with the film industry, both public and private, require investments to be based on reliable and reproducible forecasting methods [2] despite the fact that film production itself is riddled with subjectivity and uncertainty all over the world: every decision made can affect both the quality of the project and the financial side of the rental [3].

An accurate prediction of the success of the film is possible using the analysis of historical and current data. The success of a film project is determined by the combined contribution of the actors, director, composer, producer and other members of the creative team. Access to the biographical data of these specialists can be obtained from open sources such as the Internet Movie Database (IMDb) [4]. Rotten Tomatoes, as well as from social networks, which provide information about subscribers of these artists. In addition, you can view data on the number of views of projects in various databases and analytical materials, including trailers and other relevant materials [5].

It is possible to predict not only quantitative data on financial success and the number of views [6], but also the results of the Oscar Film Awards [7] and other "qualitative" events.

Predicting film distribution and box office receipts using data mining is developed in Hollywood and Indian Bollywood [8], as well as in Nigerian Nollywood [9]. The Asian trend is the active use of machine learning and neural networks. Most scientific publications on this topic are published in mainland China. Korea. Taiwan, India, Indonesia, Sri Lanka and other Asian countries.

Box office forecasting is usually carried out using machine learning algorithms [10], including the use of big data before the direct release of projects into wide distribution, because rental results directly affect investment decisions in the stock market, the development of advertising promotion strategies, and the planning of film screenings by cinema chains [11]. Nevertheless, it is also possible to predict the results at an early stage of production (pre-production), when filming of the project has not yet begun. Forecasting the profitability of films at an early stage of production can be useful for making decisions about investing [12], issuing stocks and bonds, and using other financial instruments.

In Russia, the industry of forecasting national film distribution is less developed than, for example, in China. However, the predictability of Russian cinema proved in [13], the presence of clear trends in payback, viewings, and audience ratings, depending on the characteristics and creative composition of film production teams, will make it possible to more successfully implement state policy in the field of art, culture, and creative industries development. All forecasts and conclusions can be made before the start of the most expensive filming process, and the basic data on the factors of "success", including the duration of the film project, genre orientation, age rating and other elements have already been determined by the author in his previous scientific work [14].

RESEARCH METHODS AND DATA SOURCES

Film distribution forecasting has gone through a long stage from statistical modeling in the early 2000s to machine learning methods in the second half of the 2010s [15]. As already noted, an accurate forecast of gross box office receipts is of great benefit for investment and management of the film industry, respectively, the use of different forecasting methods and a set of economic factors will allow to identify the most productive combinations [16].

Most studies use multifactorial or multimodal forecasting of movie ratings based on open data [17] with a number of factors ranging from several to several hundred. Both data from the local film market, such as Taiwan [18], and global databases such as IMDB [19] are used.

Both the rental data itself and its history, as well as related ones, such as search queries and trends, can be studied [20], and correlation analysis of big network data and movie time series data can be performed [21]. There are also textual techniques that include the use of language models to analyze the text of user comments on a future project, so it is possible to analyze the frequency of words in reviews [22] and search for combinations of them that predict success.

In this paper, the object of research is limited. It is the rental historical data that is studied, the success/failure of the film and a more complex classification are predicted, the number of views, fees, payback (fees/budget ratio) and audience rating.

As a rule, researchers compare different machine learning methods for efficiency and accuracy [23], since it is impossible to predict the results of a particular method in advance without a detailed analysis of the initial data set.

Predicting the success of a film using machine learning methods requires the use of various ways to increase accuracy [24]. In the present study, this issue is solved by adjusting machine learning hyper parameters, as well as using synthetic data to balance predicted classes.

As a rule, the study begins with linear regression, logistic regression, or the "support vector machine" [25]. This intermediate stage was presented in a previous scientific paper [14]. It should be noted that a similar approach was used in the analysis of the Turkish national cinema, including multiple regression analysis, ridge regression and lasso regression, tree-like methods such as "random forest". SVM and KNN, and the best results were predictably obtained based on the "random forest" method [26].

First of all, the researcher should be interested in the so-called "ensemble methods", which, when properly configured, give higher classification accuracy, allowing the use of multifactorial models such as XGBoost. LightGBM, and CatBoost, which were used by colleagues from China when analyzing film distribution taking into account data on the COVID-19 pandemic [27]. Ensemble methods based on decision trees are more efficient than methods using algorithms similar to k–NN [28]. Researchers from Korea [29] predict that stacking will become an important tool. This method combines all forecasting algorithms into one large meta-model.

The original dataset consists of 1685 entries about national films, starting with the film "72 meters" (ID_kinopoisk = 70952. release date — 12.02.2004) through "Concert canceled" (ID_kinopoisk = 5325618. release date — 07.09.2023). The data is collected by the author from open sources. The dataset was divided into 2 parts — 1.500 films for training and testing machine learning models and a control and validation dataset for 185 films in chronological order of release. Machine learning models did not have access to the latest dataset during the training period. The first dataset was divided into training and training samples in a 70/30 ratio with fixed separation parameters (random state = 42 for train_test_split). Categorical variables (genres) were encoded by LabelEncoder, and the gaps were filled in with median values.

Approximately 30% of Russian films that are released at the box office do not have open budgets. This significantly affects the accuracy of the research and distorts the results of regression models.

The hypothesis of the study is related to the fact that models trained and tested on data from previous periods (with a random distribution of projects for test and training samples) will show sufficiently high metrics and quality indicators even if a control dataset is used from these new projects in chronological order within a narrow time frame, taking into account objective changes in the rental structure, caused by the departure of foreign film distributors. Thus, pre-trained models will work in obviously worse conditions in terms of changing market factors and the "quality" of the data. If the decrease in quality metrics on the control dataset is insignificant compared to the test data, then the main thesis of the study will be proved. Multifactor models trained on historical data from previous periods are effective. In practice, by training the model on data from 2004–2024, we will be able to accurately predict the fate of all national projects in 2025, 2026, and other time periods.

PRELIMINARY RESULTS OF THE STUDY

For the pilot approach, a simplified version was used, a 26-factor model that included the following data: the week and month of the film's release, the number of screens, as well as items such as the film's budget (in millions of rubles), the film's age rating from 0+ to 18+, the total duration of the film in minutes, and the average fee/budget ratio by genre, average viewership rating of Kinopoisk by genre, average fees by genre, average number of views for genre, fees/budget ratio for project subgenre, average viewer rating for project subgenre, average fees for project subgenre, average views (number) for a subgenre of a project, average fees/budget ratio for a director's projects, average director's rating on Kinopoisk, average fees for a given director's projects, average views of a director's films, average fees/budget ratio for a screenwriter, average audience rating for a screenwriter, average fees for a screenwriter's projects, average number of screenwriter's project views, fees/budget ratio of the second screenwriter, average viewership rating of the second screenwriter, average number of fees of the second screenwriter's projects, the average number of views of the second scriptwriter's projects. Thus, the pilot dataset is limited to the general rental data and historical data of directors and screenwriters (in the case of novice directors and screenwriters. median values were used to prevent their discrimination).

As a rule, a binary classification can be used for this type of task: the project did not pay off at the box office, collected less than 2 of its budgets with class — 0, and the project paid off at the box office by collecting 2 or more of its budgets — 1. However, it is possible to predict a more complex project class with a slight loss of accuracy, for example: 0 — failed at the box office, 1 — did not pay off, 2 — paid off, 3 — made a profit of over 100% of the production budget. At the next stage of the regression, fees (box), views (views), fees/budget ratio (box_buget) and viewer rating (Kinopoisk_R) will be determined.

As a test, several simple ensemble machine learning models were used: AdaBoostClassifier. BaggingClassifier. ExtraTreesClassifier. GradientBoostingClassifier. RandomForestClassifier. StackingClassifier. VotingClassifier. HistGradientBoostingClassifier. For the four-class classification, the HistGradient-BoostingClassifier model showed the best results in terms of accuracy[1] = 0.8878 and ROC_AUC score[2] = 0.9611. With hyper parameters selected by the GridSearchCV method (learning_rate = 0.1. max_depth = 4. max_iter = 80), accuracy = 0.8918 and ROC_AUC score = 0.9653 were obtained. For more information, see *Fig. 1*.

Following the Chinese researchers, we note that XGBoosts and similar models are effective ensemble algorithms for solving this class of problems [30].

When translating class labels in the one-vs-all format. ROC_AUC scores were obtained for each class (*Fig. 2*).

Using the 26-factor model, we will determine the quantitative indicator. the audience rating of Kinopoisk (a dimensionless coefficient). Let's go through several variants of machine learning models: AdaBoostRegressor. HistGradientBoostingRegressor. BaggingRegressor. ExtraTreesRegressor. Gradient-BoostingRegressor. RandomForestRegressor. CatBoostRegressor. As estimated metrics for each model, we use the standard set of MSE (Mean Squared Error) — the mean squared error, R 2 (R-squared) — the coefficient of determination, which evaluates the predictive ability of the model, the proportion of explained variance relative to the total variance, and MAE — (Mean Absolute Error) — the average absolute error (*Table*).

The best results in this case were shown by the ExtraTreesRegressor ensemble model with selected hyperparameters (max_depth = 10. min_samples_split = 10. n_estimators = 250), its indicators were improved to MSE = 0.2356. MAE = 0.3010. R 2 = 0.8492). Thus, even a simple 26-factor model can provide quantitative predictions regarding viewership ratings.

In this paper, taking into account volume limitations, only general data is provided on the developed 26-factor and 146-factor regression models that predict the absolute values of fees, number of views, payback and viewer rating.

MULTIFACTORIAL (COMPLETE) CLASSIFICATION MODEL

To test the hypothesis, we used a more complex predictive model that took into account both rental data from previous periods and an expanded list of members of the creative team. 146 is a parametric model that includes the following indicators: the week and month of the film's release, the number of rental screens. the budget of the film (where known), age rating, duration, genres (up to 5 titles), historical data on the project director, screenwriters, cameraman,

[1] Accuracy — the number of correctly predicted classification results to the total number of predicted results, ranges from 0 to 1.

[2] The ROC curve (Receiver Operating Characteristic or Receiver Operating Characteristic) is a graph illustrating the performance of a classification model at all possible classification thresholds. The AUC (Area Under the ROC Curve) indicator is a measure that allows you to summarize the performance of a model with a single number by measuring the area under the ROC curve. The AUC ranges from 0 to 1, where a higher AUC value indicates higher model performance. An indicator of 0.5 corresponds to "random fortune telling".

Classification Report

	precision	recall	f1-score	support
0	0.9535	0.9750	0.9642	400
1	0.5641	0.5366	0.5500	41
2	0.6154	0.3333	0.4324	24
3	0.6579	0.7353	0.6944	34
Accuracy			0.8918	499
macro avg	0.6977	0.6451	0.6603	499
weighted avg	0.8851	0.8918	0.8862	499

Fig. 1. **Error Matrix and Classification Report of the 26-Factor Model**

Source: Compiled by the author.

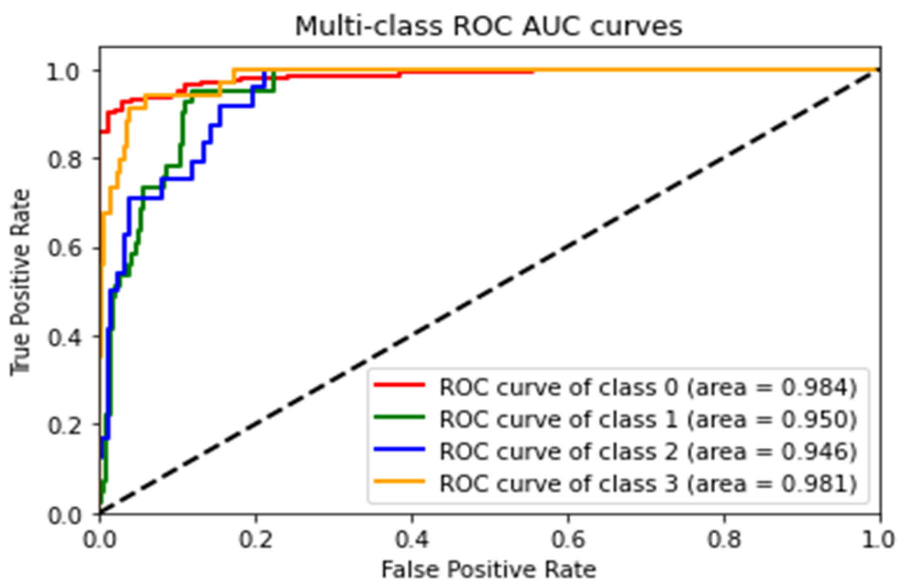


Fig. 2. **ROC_AUC_Score Curve for a Four-Class 26-Factor Model**

Source: Compiled by the author.

producers of the project, artist, editor, composer, main to the actors of the project. The list of historical data included average historical data on rental screens, budgets, fees, views. Kinopoisk rating, IMDb (if available), relative indicators — fees per screen, screen views, budget fees.

To prove the original statement, the total dataset of 1.685 films is divided into 3 parts. The machine learning model is being trained and tested on 1.500 records (the records in the samples are randomly distributed), and the generalizing ability of the model is being tested on 185, which are already arranged in chronological order from the moment of release in mid-2022 to September 2023.

The problem of class imbalance is already relevant for the two-class classification: the number of films

Table

Characteristics of Several Predictive Ensemble Regression Models for Kinopoisk Ratings

No.	Machine learning model	MSE	R 2	MAE
1	AdaBoostRegressor	0.385185	0.753436	0.457239
2	BaggingRegressor	0.283048	0.818815	0.332489
3	ExtraTreesRegressor	0.238982	0.847023	0.301738
4	GradientBoostingRegressor	0.270861	0.826616	0.338732
5	RandomForestRegressor	0.255361	0.836538	0.306668
6	HistGradientBoostingRegressor	0.294991	0.811170	0.351616
7	CatBoostRegressor	0.261589	0.832552	0.337352

Source: Compiled by the author.

that were successful at the box office (class = 1) is an order of magnitude less than unsuccessful (class = 0). Therefore, it is advisable to use synthetic data using the balanced-learn library. Thus, the total number of records from the main dataset of 1500 becomes 2670. The sample is divided into training and test samples in a ratio of 70 to 30 using train_test_split. By analogy with the 26-factor model, a machine learning algorithm has been selected that provides the best quality metrics in this configuration. It turned out to be a domestic development from the specialists of Yandex[3] CatBoostClassifier.

The model, trained using synthetic data on balanced classes, was wrong only in 12 cases, showing an accuracy of 0.9863 and a limiting ROC_AUC_ score = 0.9999 after 1000 iterations. In principle, this may indicate an "overfitting" model and, consequently, a low predictive ability. For verification, we will use the actual data on 185 films, to which the model, according to the conditions of hypothesis testing, did not have access either at the training stage or at the testing stage.

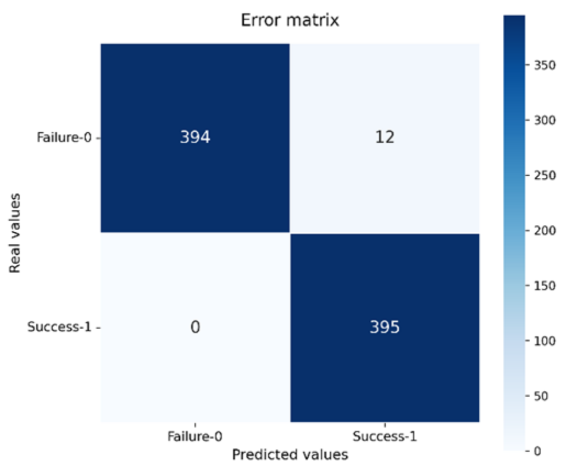
Despite the assumption that the model was retrained, the unbalanced validation data by class (170 unsuccessful projects. 15 that paid off at the box office) and the fact that they reflect a new historical period in the development of Russian cinema associated with the departure of foreign distributors, the classifying ability of the model remained extremely high. Accuracy = 0.9568. ROC_AUC_score = 0.9803.

Let's test the proposed approach on a four-class classification. The sample is also divided into 1.500 films for training and testing the model, and 185 for subsequent validation. A significant difference is that using the imblearn.over_sampling module of the RandomOverSampler function and synthetic data. 4 classes were balanced, and the total number of rows in the dataset out of 1.500 became 4843. The CatBoostClassifier model was used for classification in adaptation to 4 classes. The accuracy values of 0.9842 and ROC_AUC_score = 0.9997 were obtained on the test data.

Similarly, we use the new data to validate the model. Due to the specifics of the data, the results were slightly lower than the test results, which again do not exclude the possibility of using this approach to predict the data of new films at the pre-production stage. The accuracy of the four-class classification was 0.9189. ROC_AUC_ score = 0.9663.

When evaluating the precision, recall, f1-score, and ROC_AUC metrics, we note that the extreme classes (0 and 3) have the highest results, and the most errors are related to the identification of intermediate classes 1 and 2. This is presumably due to a lack of information on the project budget (up to 30%) and conditional payback criteria: two project budgets are rather a rule of thumb for determining the payback of a project.

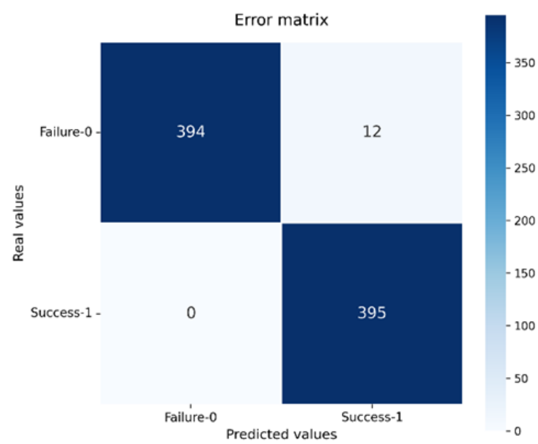
[3] CatBoost is a fast, scalable, and high — performance decision tree gradient enhancement library. It is used for ranking, classification, regression and other machine learning tasks. URL: https://github.com/catboost/ (accessed on 15.03.2024).



Classification Report:

	precision	recall	f1-score	support
0	1.0000	0.9729	0.9863	406
1	0.9729	1.0000	0.9863	395
accuracy			0.9863	801
macro avg	0.9865	0.9865	0.9863	801
weighted avg	0.9866	0.9863	0.9863	801

Fig. 3. **Error Matrix and Classification Report of the 146-Factor Model**
Source: Compiled by the author.



Classification Report:

	precision	recall	f1-score	support
0	0.9821	0.9706	0.9763	170
1	0.7059	0.8000	0.7500	15
accuracy			0.9568	185
macro avg	0.8440	0.8853	0.8632	185
weighted avg	0.9597	0.9568	0.9580	185

Fig. 4. **Error Matrix and Classification Report of the 146-Factor Model based on New Data**
Source: Compiled by the author.

Based on the results of the experiment, the hypothesis about the predictability of financial results for Russian cinema was confirmed using a machine learning model that was trained on data from previous years for new films, including those that were in the preliminary stages of production before the most expensive phase of filming began. Practical examples of each Russian film are given in Appendix A, which shows the predicted class and actual class, along with errors, for two-class and four-class classification.

A slight decrease in classification metrics may be due to: the "retraining" of the model, the use of synthetic data to equalize the imbalance by class, the chronologically "narrow" release period of 185 films, their presentation in chronological order, partly due to fundamental changes in the characteristics of the Russian film market due to the departure of foreign distributors. Each episode of these variants needs additional research. Nevertheless, taking into account the possible further training of predictive models based on new data, it is possible to quickly pre-evaluate the prospects of both an individual film project and a group of film projects in terms of their future financial and other rental results.

The continuation of the study is related to obtaining more accurate budget data (there is a shortage of budget data in 30% of Russian films. which

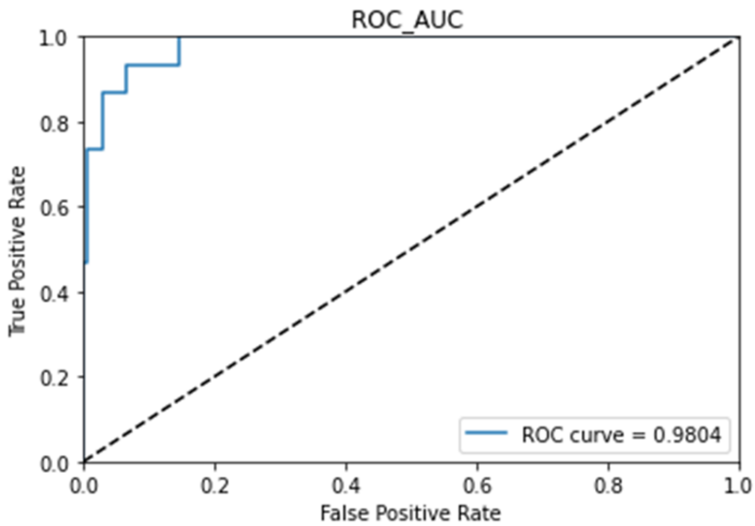


Fig. 5. **ROC_AUC_Score Curve for a Two-Class 146-Factor Model on New Data**

Source: Compiled by the author.

Classification Report:

	precision	recall	f1-score	support
0	1.0000	0.9390	0.9685	377
1	0.9491	1.0000	0.9739	354
2	1.0000	1.0000	1.0000	363
3	0.9890	1.0000	0.9945	360
accuracy			0.9842	1454
macro avg	0.9845	0.9847	0.9842	1454
weighted avg	0.9849	0.9842	0.9841	1454

Fig. 6. **Error Matrix and Classification Report of the 146-Factor Model (4 Classes)**

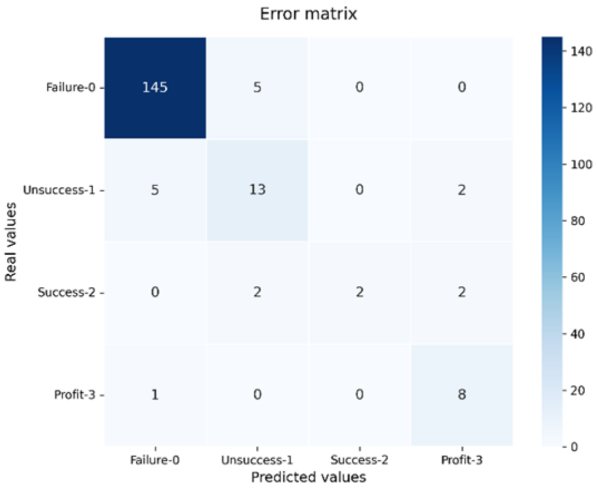
Source: Compiled by the author.

negatively affects the accuracy of forecasting). The low information openness of the accounting department of the Russian film industry negatively affects the ability to accurately predict fees based on regression models of machine learning. Nevertheless, there is a possibility of such forecasts.

185 films are worth 21.7 billion rubles of fees and 73.5 million views. The main factor is the Cheburashka[4] anomaly; the project has collected over 7 billion rubles and 23.5 million views. This film is an anomaly due to the release during the New Year period, the lack of competitors and an objective factor — a rare combination of the creative group (the director with the largest series of commercially successful works in history. actors and other members of the creative group with high rental rates).

Taking into account the exclusion of Cheburashka's results and the missing data for Kinopoisk's rating and budgets, we calculated the prediction accuracy (R 2 coefficient of determination) between actual and predicted values. The accuracy for Kinopoisk's

[4] See the publication: Dozhdikov A.V. Cheburashka's success can be repeated with the help of AI. And not only in the film industry! Plas. 21.03.2023. URL: https://plusworld.ru/journal/2023/plus-3–2023/uspekh-cheburashki-mozhno-povtorit-s-pomoshchyu-ii-i-ne-tolko-v-kinoindustrii/ (accessed on 15.03.2024).



Classification Report:

	precision	recall	f1-score	support
0	0.9732	0.9667	0.9699	150
1	0.6818	0.7500	0.7143	20
2	1.0000	0.3333	0.5000	6
3	0.6667	0.8889	0.7619	9
accuracy			0.9189	185
macro avg	0.8304	0.7347	0.7365	185
weighted avg	0.9276	0.9189	0.9169	185

Fig. 7. **Error Matrix and Classification Report of the 146-Factor Model (4 Classes) Based on New Data**
Source: Compiled by the author.

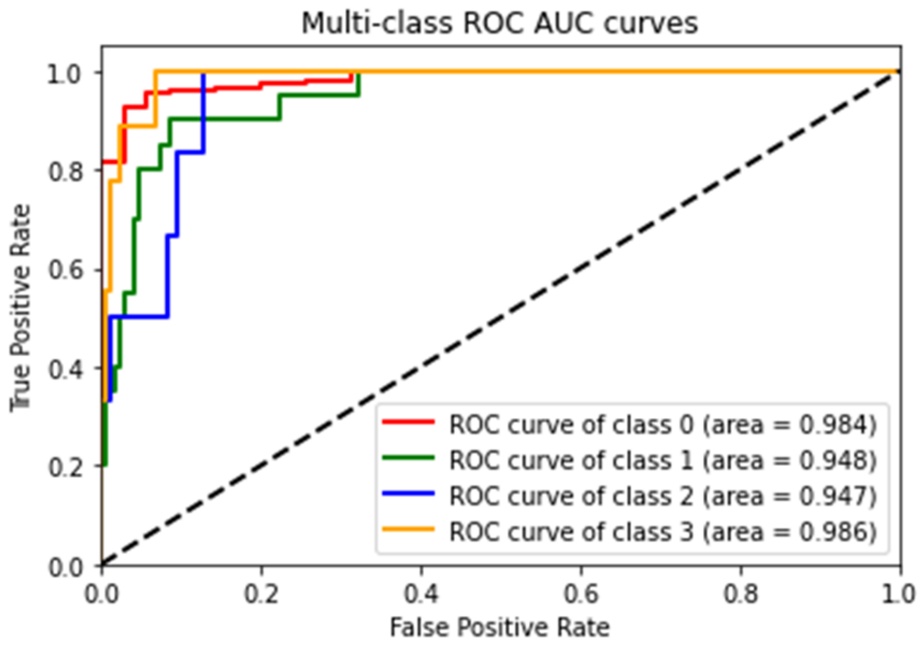


Fig. 8. **ROC_AUC_Score Curve for a Four-Class 146-Factor Model on New Data**
Source: Compiled by the author.

audience rating was 0.8392, for box office fees it was 0.9076 and for viewings it was 0.8797.

The average absolute error MAE (mean absolute error) for fees was 26.78 million rubles, for views — 108.23 thousand, for the rating of Kinopoisk — 0.5874 rating points. Thus, in terms of the average absolute error, this forecasting technique is not suitable for low-budget (usually copyrighted or documentary) films with a small number of screens at the box office in terms of fees and views. It is recommended to evaluate, first and foremost, medium-budget (80–300 million rubles) and large-budget (up to 850 million rubles) films. Projects with a large budget often do not pay off at the box office and do not recoup two of their budget due to the limited size of the Russian market.

For detailed information on the actual and predicted values, see Appendices A and B.

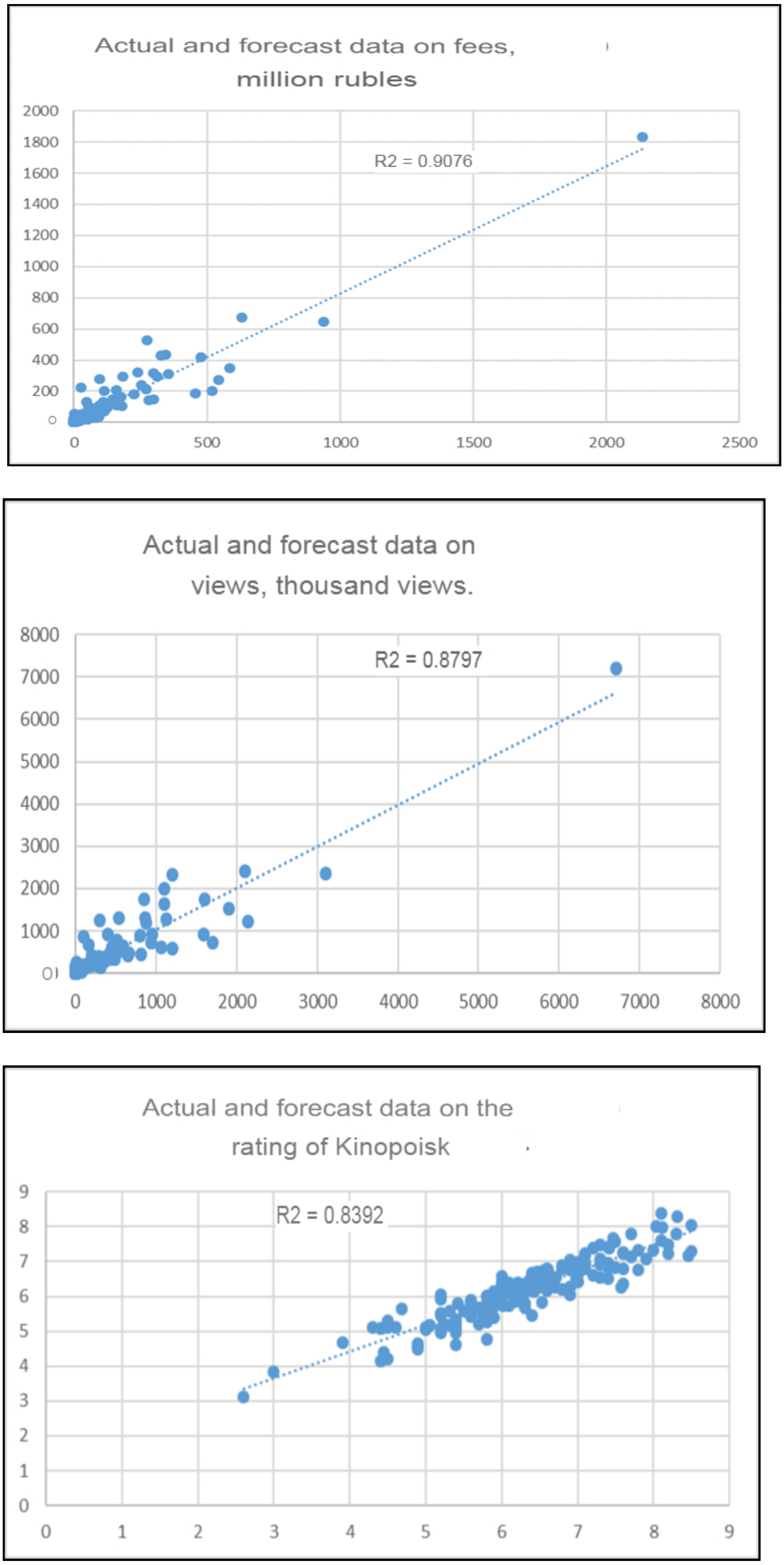


Fig. 9. **Discrepancy Between Actual and Predicted Values of Box Office, Views and Viewership Ratings**

Source: Compiled by the author.

Individual results — predicted and actual results may differ from each other for films with minimum and maximum scores. However, for the "portfolio of film projects", the discrepancy between the predicted indicators and the actual ones will be significantly smaller. The larger the number of projects evaluated, the smaller the difference between the total predicted and actual financial results.

RESEARCH DIRECTIONS

Further development of methods for predicting both audience ratings [31] and box office receipts [32] involves the use of neural networks, including "deep trust networks" — Deep Belief Network or DBN [33].

It is possible to combine approaches of classical machine learning and neural networks such as IFOA-GRNN [34]. A combination of classical machine learning methods such as "support vector machines", neural networks, and natural language processing methods combined with traditional historical data from IMDb, Rotten Tomatoes, Box Office Mojo, and Metacritic sources also seems promising [27].

A separate area involving the use of elements of "computer vision" is the use of graphic images, including film posters, analyzed using feature extraction for a convolution neural network. The use of deep learning algorithms will make it possible to create a new generation of decision support systems for film production, as well as promising "hybrid" human-machine decision-making systems based on human-machine interaction

CONCLUSIONS

Appendices A and B present the results of the classification of 185 motion pictures (two-class and four-class classifications) and the results of predicting quantitative data on 185 motion pictures. The forecast and actual values are indicated. We note the high accuracy of the classification results. The predictive ability of regression models (how much a given film will earn and how many viewers will watch it) is limited by the following factors:

1. 30% of the budgets of Russian film projects are unknown (hidden by producers from the public).

2. The final results are influenced by a number of additional parameters, in particular, the marketing budget, the characteristics of the advertising campaign, the release time of the film, which can be changed by an "administrative" decision, competition with other films, and the cannibalization of the target audience.

3. More or less accurate quantitative forecasts are possible for most medium-budget and large-budget films. The coincidence of projected results with actual results is least among blockbuster projects, the rolling fate of which includes an "administrative" resource and low-budget projects.

Taking into account these data and the further training of forecasting models based on the data of new periods, opportunities arise:

• evaluation of film projects at the stage of production pitching (the Cinema Foundation. the Ministry of Culture. private organizations), before the start of the shooting period; to obtain a preliminary forecast, the estimated rental characteristics of the film and information about the creative group are sufficient;

• forming a portfolio of film projects to attract private and public investment; 20–30 films in the "portfolio" per year assume stable profitability over the medium-term planning horizon;

• creation of long-term strategies for the development of national cinematography;

• development of regulatory strategic documents in the field of national cinematography support;

• selection of rental characteristics and characteristics of the creative team in order to achieve maximum effect in terms of audience reach and audience ratings in the production of national films and projects with priority government support.

REFERENCES

1. Paul C., Das P.K. Predicting movie revenue before committing significant investments. *Journal of Media Economics*. 2022;34(2):63–90. DOI: 10.1080/08997764.2022.2066108
2. Mbunge E., Fashoto S.G., Bimha H. Prediction of box-office success: A review of trends and machine learning computational models. *International Journal of Business Intelligence and Data Mining*. 2022;20(2):192–207. DOI: 10.1504/IJBIDM.2022.120825

3. Tantawichien J., Mizuyama H., Nonaka T. Designing a human computation game for enhancing early-phase movie box office prediction. In: Hamada R., et al. Neo-simulation and gaming toward active learning. Singapore: Springer; 2019:13–22. (Translational Systems Sciences. Vol. 18.). DOI: 10.1007/978–981–13–8039–6_2
4. Pirunthavi V., Vithusia R.P., Abishankar K., Ekanayake E.M., Yanusha M. Movie success and rating prediction using data mining algorithms. In: Int. res. conf. of Uva Wellassa University (IRCUWU-2020) "Sustainable business transition through information and knowledge dissemination". Badulla: Uva Wellassa University; 2020:175–176. URL: https://www.uwu.ac.lk/wp-content/uploads/2020/proceeding_of_ircuwu2020_v2.pdf
5. Chakraborty P., Rahman Z., Rahman S. Movie success prediction using historical and current data mining. *International Journal of Computer Applications*. 2019;178(47):1–5. DOI: 10.5120/ijca2019919415
6. Murschetz P.C., Bruneel C., Guy J.L., et al. Movie industry economics: How data analytics can help predict movies' financial success. *Nordic Journal of Media Management*. 2020;1(3):339–359. DOI: 10.5278/njmm.2597–0445.5871
7. Bruneel C., Guy J.-L., Haughton D., et al. Movie analytics and the future of film finance. Are Oscars and box office revenue predictable? In: Murschetz P., Teichmann R., Karmasin M., eds. Handbook of state aid for film. Media business and innovation. Cham: Springer; 2018:551–578. DOI: 10.1007/978–3–319–71716–6_30
8. Ruus R., Sharma R. Predicting movies' box office result — a large scale study across Hollywood and Bollywood. In: Cherifi H., Gaito S., Mendes J., Moro E., Rocha L., eds. Complex networks and their applications VIII. (COMPLEX NETWORKS 2019). Cham: Springer; 2020:982–994. (Studies in Computational Intelligence. Vol. 882). DOI: 10.1007/978–3–030–36683–4_78
9. Adekola O.D., Maitanmi S.O., Kasali F.A., et al. Movie success prediction using data mining. *British Journal of Computer, Networking and Information Technology*. 2021;4(2):22–30. DOI: 10.52589/BJCNIT-CQOCIREC
10. Namlı Ö., Ozcan T. Forecasting of box office revenue using machine learning algorithms. *Yönetim Bilişim Sistemleri Dergisi.* 2017;(3):130–143.
11. Wang Z., Zhang J., Ji S., et al. Predicting and ranking box office revenue of movies based on Big Data. *Information Fusion*. 2020;60:25–40. DOI: 10.1016/j.inffus.2020.02.002
12. Shahid M.H., Islam M.A. Investigation of time series-based genre popularity features for box office success prediction. *The Open Access Journal for Computer Science*. 2023;9: e1603. DOI: 10.7717/peerj-cs.1603
13. Dozhdikov A. Prediction of the results of movie release using machine learning. *Voprosy teoreticheskoi ekonomiki = Theoretical Economics.* 2023;(4):93–114. (In Russ.). DOI: 10.52342/2587–7666VTE_2023_4_93_114
14. Dozhdikov A.V. Determination of investment success and its factors for Russian cinema at the box office using machine learning. *Finance: Theory and Practice*. 2024;28(1):188–203. DOI: 10.26794/2587–5671–2024–28–1–188–203
15. Abidi S.M.R., Xu Y., Ni J., Wang X., Zhang W. Popularity prediction of movies: From statistical modeling to machine learning techniques. *Multimedia Tools and Applications*. 2020;79(4):35583–35617. DOI: 10.1007/s11042–019–08546–5
16. Li D., Liu Z.-P. Predicting box-office markets with machine learning methods. *Entropy*. 2022;24(5):711. DOI: 10.3390/e24050711
17. Qin M., Zhou Q., Chen W., Zhao L. MAMRP: Multi-modal data aware movie rating prediction. In: Yang X., et al. Advanced data mining and applications (ADMA 2023). Cham: Springer; 2023:660–675. (Lecture Notes in Computer Science. Vol. 14177). DOI: 10.1007/978–3–031–46664–9_44
18. Lu S.-H., Wang H.-J., Nguyen A.T. Machine learning applications on box-office revenue forecasting: The Taiwanese film market case study. In: Optimal transport statistics for economics and related topics. Cham: Springer; 2023:384–402. (Studies in Systems, Decision and Control. Vol. 483). DOI: 10.1007/978–3–031–35763–3_49
19. Gore M., Sheth A., Abbad S., Jain P., Mishra P. IMDB box office prediction using machine learning algorithms. *International Journal for Research in Applied Science and Engineering Technology*. 2022;10(5):2438–2442. DOI: 10.22214/ijraset.2022.42653
20. Kang D. Box-office forecasting in Korea using search trend data: A modified generalized Bass diffusion model. *Electronic Commerce Research*. 2021;21(3):41–72. DOI: 10.1007/s10660–020–09456–7
21. Li Na., Xia L. Correlation analysis of network Big Data and film time-series data based on machine learning algorithm. *Mathematical Problems in Engineering*. 2022. DOI: 10.1155/2022/4067554

22. Yoo B.-K., Kim S.-H. Movie box office prediction at the distribution stage using text mining of movie reviews. *The Korean Logistics Research Association*. 2023;33(1):95–105. DOI: 10.17825/klr.2023.33.1.95
23. Quader N., Gani O., Chaki D. Performance evaluation of seven machine learning classification techniques for movie box office success prediction. In: 3rd Int. conf. on electrical information and communication technology (EICT). (Khulna, December 7–9, 2017). New York, NY: IEEE; 2017:1–6. DOI: 10.1109/EICT.2017.8275242
24. Lee K., Park J., Kim I., Choi Y. Predicting movie success with machine learning techniques: Ways to improve accuracy. *Information Systems Frontiers*. 2018;20(3):577–588. DOI: 10.1007/s10796–016–9689-z
25. Hu Y.-H., Shiau W.-M., Shih S.-P., Chen C.-J. Considering online consumer reviews to predict movie box-office performance between the years 2009 and 2014 in the US. *The Electronic Library*. 2018;36(6):1010–1026. DOI: 10.1108/EL-02–2018–0040
26. Gürbüz A., Biçer E., Kaya T. Prediction of gross movie revenue in the Turkish box office using machine learning techniques. In: Kahraman C., et al., eds. Intelligent and fuzzy systems (INFUS 2022). Cham: Springer; 2022:86–92. (Lecture Notes in Networks and Systems. Vol. 505). DOI: 10.1007/978–3–031–09176–6_10
27. Ni Y., Dong F., Zou M., Li W. Movie box office prediction based on multi-model ensembles. *Information*. 2022;13(6):299. DOI: 10.3390/info13060299
28. Lee S., Bikash K.C., Choeh J.Y. Comparing performance of ensemble methods in predicting movie box office revenue. *Heliyon*. 2020;6(6): e04260. DOI: 10.1016/j.heliyon.2020.e04260
29. Park J.H., Lim C. Predicting movie audience with stacked generalization by combining machine-learning algorithms. *Communications for Statistical Applications and Methods*. 2021;28(3):217–232. DOI: 10.29220/CSAM.2021.28.3.217
30. Dutta S., Dasgupta K. A shallow approach to gradient boosting (XGBoosts) for prediction of the box office revenue of a movie. In: Proc. Int. conf. on innovations in software architecture and computational systems. Singapore: Springer; 2021:207–219. DOI: 10.1007/978–981–16–4301–9_16
31. Su Y., Zhang Y., Yan J. Neural network based movie rating prediction. In: Proc. Int. conf. on Big Data and computing (ICBDC'18). (Shenzhen, April 28–30, 2018). New York, NY: Association for Computing Machinery; 2018:33–37. DOI: 10.1145/3220199.3220204
32. Ru Y., Li B., Chai J. An effective daily box office prediction model based on deep neural networks. *Cognitive Systems Research*. 2018;52:182–191. DOI: 10.1016/j.cogsys.2018.06.018
33. Wang W., Xiu J., Yang Z., Liu C. A deep learning model for predicting movie box office based on deep belief network. In: Tan Y., Shi Y., Tang Q., eds. Advances in swarm intelligence (ICSI 2018). Cham: Springer; 2018:530–541. (Lecture Notes in Computer Science. Vol. 10942). DOI: 10.1007/978–3–319–93818–9_51
34. Lu W., Zhang X., Zhan X. Movie box office prediction based on IFOA-GRNN. *Discrete Dynamics in Nature and Society*. 2022:3690077. DOI: 10.1155/2022/3690077

ABOUT THE AUTHOR

Anton V. Dozhdikov — Cand. Sci. (Polit.), Senior Researcher at the UNESCO Chair, Institute of Socio-Political Research of the Federal Center of Theoretical and Applied Sociology of the Russian Academy of Sciences, Moscow, Russian Federation
http://orcid.org/0000-0002-1069-1648
antondnn@yandex.ru

Conflicts of Interest Statement: The author has no conflicts of interest to declare.

The article was submitted on 01.04.2024; revised on 01.05.2024 and accepted for publication on 22.05.2024. The authors read and approved the final version of the manuscript.

Translated by N.I. Sokolova

Appendix A

Results of Predictive Classification in Accordance with Actual Values

Title of the movie	2-class classification		4-class classification	
	Actual class	Predicted class	Actual class	Predicted class
Nakhimov naval graduates	0	0	0	0
Asphalt Sun	0	0	0	0
Divorce: The Second Film	0	0	0	0
Legends of the Eaglet	0	0	1	0
Archipelago	0	0	0	0
Night Mode: Film	0	0	0	0
Call Sign “Baron”	0	0	0	0
The Little Prince	0	0	0	0
Lena & Justice	0	0	1	0
Barbosky Team	0	0	0	0
Calendar Maya	0	0	1	1
Turbosaurs. Go Ahead!	0	0	0	0
Crazy Teachers	0	0	0	0
Where Our Home Is	0	0	0	0
Start Over	0	0	0	0
At the Very White Sea	0	0	0	0
The Tiber	0	0	0	0
Non-Childish House	0	0	0	0
Kat	0	0	0	0
The Swan Lake	0	0	0	0
The Little Red Riding Hood	0	0	0	0
Who’s there?	0	0	0	0
Parent	0	0	0	0
Express	0	0	1	1
It’s more fun together	0	0	0	0
The French Master	0	0	0	0
Distant Loved Ones	0	0	0	0
My Son’s Mother	0	0	0	0

Appendix A (continued)

Title of the movie	2-class classification		4-class classification	
	Actual class	Predicted class	Actual class	Predicted class
Goliath	0	0	0	0
Zina and Lech: Operation “Tail and Udder”	1	1	3	3
Heart of Parma	0	0	1	1
A Tale for Old	0	0	1	0
Love Stories	0	0	0	0
Ivan Semenov: The School Commotion	1	1	2	3
Additional Lesson	0	0	0	0
You Can’t Cry	0	0	0	0
Girl in a Million	0	0	0	0
The Formidable Dad	0	0	0	0
A Great Journey. Special Delivery	0	0	0	0
Liberia: Treasure Hunters	0	0	0	1
The Adventures of Little Baja	0	0	0	0
Summer Will Be Over Soon	0	0	0	0
Carpenter	0	0	0	0
Velga	0	0	0	0
Petropolis	0	0	0	0
Kunnei	1	1	3	3
Hosanna	0	0	0	0
Patroness	0	0	0	0
The Saint	0	0	0	0
I will live	0	0	0	0
The Golden Neighbors	0	0	0	0
The Lovers	0	1	1	1
Peter the Great. the last Tsar and the first Emperor	0	0	1	1
A Similar Person	0	0	0	0
The Stars will show me the Way.	0	0	0	0
The Wizards	0	0	0	0
Like a Man	0	0	0	0
Curse: The Dead Earth	0	0	0	0

Appendix A (continued)

Title of the movie	2-class classification		4-class classification	
	Actual class	Predicted class	Actual class	Predicted class
The Day Before	0	0	0	0
Zemun	0	0	0	0
Angels are Sweet	0	0	0	0
A Thousand Cheap Lighters	0	0	0	0
Chink: the Tailed Detective	0	0	0	0
Elephant	0	0	0	0
Jeanne	0	0	0	0
I'm Rewinding!	0	0	0	0
Grisha Subbotin	0	0	0	0
The Kingdom Against the Robbers	0	0	0	0
Honest Divorce 2	0	0	0	0
Winter Season	0	0	0	0
Smart Masha	0	0	0	0
The Sisters	0	0	0	0
Our Winter	0	0	0	0
Physics Department	0	0	0	0
F20	0	0	0	0
Emperor	0	0	0	0
Pelevin	0	0	0	0
Christmas Trees 9	1	1	3	3
By Touch	0	0	0	0
A Pool	0	0	0	0
Shadow. Take Gordey	0	0	0	0
The Nutcracker and the Magic Flute	0	0	0	0
Secret Santa	0	0	1	1
On Exhale	0	0	0	0
A Wedge Came Together on You…	0	0	0	0
Balabanov: The Bell Tower. Requiem	0	0	0	0
Nearby	0	0	0	0
Masha and the Bear in Cinema: 12 Months	0	0	0	0

Appendix A (continued)

Title of the movie	2-class classification		4-class classification	
	Actual class	Predicted class	Actual class	Predicted class
Naughty Boy 2	1	1	3	3
Exes: Happy End	0	0	0	0
Chuck & Huck: A Great Adventure	0	0	1	1
Mira	0	0	0	1
Christmas Trees and Needles	0	0	0	0
Cheburashka	1	1	3	3
Music Video Makers	0	0	0	0
Umka	0	0	0	0
Youth	0	0	0	0
Turbosaurs. Winter Adventures	0	0	0	0
The Path of the Dream	0	0	0	0
Camel's Arc	0	0	0	0
Lily and the Sea / Grand Marin	0	0	0	0
Vysotsky: Unknown Pages	0	0	0	0
Don't Bury Me Without Ivan	0	0	0	0
Dear Mom and Dad	0	0	0	0
What Men Talk About...	0	1	1	3
One Real Day	0	0	0	0
Liza	0	0	0	0
An Easy Introduction	0	0	0	0
The White Angel of the Tundra	0	1	1	1
Free Relationship	0	1	1	1
To Be	0	0	0	1
Snow. Sister. and Wolverine	0	0	0	0
The Righteous One	0	0	0	0
Naughty	1	1	3	3
The Snow Queen: Defrosting	0	0	0	0
This Love	0	0	0	0
Timbre	0	0	0	0
Women of Altai	0	0	0	0

Appendix A (continued)

Title of the movie	2-class classification		4-class classification	
	Actual class	Predicted class	Actual class	Predicted class
Bugun-Bylyr 2	1	1	2	2
Checago	0	0	0	0
Rabies	0	0	0	0
Dads Versus Moms	1	0	3	0
Zhenya and Vanya at the Edge of the World	0	0	0	0
Where Is the Border?	0	0	0	0
Nuremberg	0	0	1	1
A Healthy Person	0	0	0	0
The Unscrupulous in the Countryside	0	0	0	0
The Detector	0	0	0	0
She's gone crazy	1	1	2	3
Escort Girl	0	0	0	0
The Cold Race	0	0	0	0
The Holy Archipelago	0	0	0	0
Medea	0	0	0	0
Towards the Sun. Along the Rows of Corn	1	1	2	2
Almost Like Everyone Else	0	0	0	0
Sergius Against Evil Spirits: The Sabbath	0	0	0	0
The Puppeteer	0	0	0	0
Let's Talk About It	0	0	0	0
Cats of the Hermitage	0	0	1	0
Aita	1	1	3	3
Lovey-Dovey	0	0	0	1
My God! I Feel Your Approach!	0	0	0	0
A Strange House	0	0	0	0
Flood	0	0	0	0
I want to! I will!	0	0	0	0
Live broadcast	0	0	0	0
The waves	0	0	0	0
Doctor	0	0	0	0

Appendix A (continued)

Title of the movie	2-class classification		4-class classification	
	Actual class	Predicted class	Actual class	Predicted class
Zoskina gas station	0	0	0	0
Russian Cross	0	0	0	0
Challenge	0	1	1	3
Yaga and The Book of Spells	0	0	0	0
Amnesty	0	0	0	0
Treasures of Guerrilla Forest	0	0	0	0
Yura. the janitor	0	0	0	0
The lesser evil	0	0	0	0
Masquerade	0	0	0	0
14+ continues	0	0	1	0
Bulgakov. this world is mine!	0	0	0	0
Khitrova. The Sign of Four	0	0	0	0
Kaka. directly. another movie	1	1	3	3
Before dawn	0	0	1	1
Masha and The Bear in the movie. say “oh!”	0	0	1	1
Knots	0	0	0	0
Tin can head	0	0	0	0
Sun tastes like	0	0	0	0
Bullfinch	0	0	0	0
Through time	0	0	0	0
Mikulai	0	0	0	0
Maldives waiting	0	0	0	0
Breathing	0	0	0	1
Turbosaur. hello. siren!	0	0	0	0
Status	0	0	0	0
Johnny	0	0	0	0
Centaur	0	0	1	1
Business in Russia	0	0	0	0
To Palych!	1	0	2	1
Baba Yaga saves the world	1	0	2	1

Appendix A (continued)

Title of the movie	2-class classification		4-class classification	
	Actual class	Predicted class	Actual class	Predicted class
Witness	0	0	0	0
Like stars	0	0	0	0
Quest	0	0	0	0
Ruslan and Ludmila. more than a fairy tale	0	0	0	0
Chizhik-Pyzhik returns	0	0	0	0
Nina. girl and piano thieves	0	0	0	0
Concert canceled	0	0	0	0

Source: Compiled by the author.

Appendix B

Actual and Predicted Figures for Box Office Receipts, Audience Views, and "Kinopoisk" Ratings

No.	Name	ID	Actual box office receipts	Expected box office receipts	Actual views	Predicted views	Actual rating	Predicted rating
1	Nakhimov naval graduates Petropolis	1403069	65 000 000	75 190 862	264 000	342 669	5.24	5.2
2	Asphalt Sun	1392757	9 100 000	14 684 488	36 000	58 174	7.10	6.78
3	Divorce: The Second Film	5042126	6 900 000	14 253 677	26 000	56 278	6.43	6.15
4	Legends of the Eaglet	4517052	62 000 000	44 973 131	284 000	169 044	7.47	7.66
5	Archipelago	1256273	819 000	960 973	3 600	3 999	5.75	5.48
6	Night Mode: Film	5034524	73 000 000	29 859 533	26 000	75 715	5.99	6
7	Call Sign "Baron"	1395858	978 000	1 265 734	3 500	6 263	5.79	5.59
8	The Little Prince	572553	17000	134 147	105	614	8.03	8.02
9	Lena & Justice	4522328	82 000 000	49 468 637	30 000	119 982	6.55	6.72
10	Barbosky Team	4518842	97 000 000	113 138 269	448 000	616 840	8.11	7.99
11	Calendar Maya	4963617	115 000 000	70 163 944	483 000	320 873	7.40	6.94
12	Turbosaurs. Go Ahead!	5042646	17 000 000	27 931 832	95 000	104 736	7.60	6.36
13	Crazy Teachers	5042247	11 000 000	6 346 878	4 600	15 196	4.50	5.31
14	Where Our Home Is	4934473	289 000	362 769	1 700	1 708	8.31	8.3
15	Start Over	2000101	55 000 000	44 862 489	208 000	185 372	6.10	5.74

Appendix B (continued)

No.	Name	ID	Actual box office receipts	Expected box office receipts	Actual views	Predicted views	Actual rating	Predicted rating
16	At the Very White Sea	4624744	568 000	751 443	4 000	3 683	6.42	6.39
17	The Tiber	4518317	19 000 000	17 323 733	67 000	77 573	5.97	6.01
18	Non-Childish House	4680791	81 000 000	29 539 678	29 000	120 898	6.53	5.82
19	Kat	4680804	2 900 000	8 406 769	8 600	41 158	5.80	6.03
20	The Swan Lake	1282100	24 000	1 406 730	81	4 571	6.25	5.96
21	The Little Red Riding Hood	1445165	144 000 000	147 842 388	546 000	518 831	4.45	4.39
22	Who's there?	4999693	91 000 000	87 215 892	32 000	201 908	6.22	6.17
23	Parent	1399249	20 000 000	5 417 875	6 900	21 429	7.21	7.39
24	Express	2000127	81 000 000	37 800 832	27 000	108 361	6.37	6.47
25	It's more fun together	5075871	51 000 000	20 383 758	9 500	71 274	5.05	5.19
26	The French Master	4528694	6 400 000	2 321 512	11 000	9 450	5.32	5.59
27	Distant Loved Ones	4489530	24 000 000	23 899 852	90 000	73 357	7.42	7.38
28	My Son's Mother	4400160	9 200 000	28 682 771	32 000	149 687	5.86	5.89
29	Goliath	5047455	48 000	237 145	137	1 145	6.46	6.7
30	Zina and Lech: Operation "Tail and Udder"	5119760	15 000 000	11 095 392	46 000	48 230	Н/Д	4.31
31	Heart of Parma	1045585	937 000 000	648 048 356	3 100 000	2 345 388	6.72	6.55
32	A Tale for Old	4739575	3 800 000	4 201 228	13 000	18 707	6.00	6.21
33	Love Stories	984365	70 000 000	61 204 656	253 000	234 582	4.68	5.63
34	Ivan Semenov: The School Commotion	5021585	144 000 000	122 546 247	648 000	421 285	6.19	5.87
35	Additional Lesson	1451292	12 000 000	9 790 986	43 000	46 126	6.98	6.51
36	You Can't Cry	4322004	3 400 000	4 308 830	11 000	20 444	6.66	6.37
37	Girl in a Million	5117212	132 000	239 639	608	1 136	4.89	4.48
38	The Formidable Dad	1445164	158 000 000	108 267 371	662 000	478 651	6.01	5.74
39	A Great Journey. Special Delivery	1445162	180 000 000	104 317 294	817 000	442 602	7.57	6.25
40	Liberia: Treasure Hunters	1398953	238 000 000	323 131 523	860 000	1 305 321	6.22	6.38
41	The Adventures of Little Baja	4530500	3 700 000	5 332 417	18 000	27 764	Н/Д	5.21
42	Summer Will Be Over Soon	1443813	1 200 000	1 190 119	3 900	4 133	6.14	6.06
43	Carpenter	4381764	1 400 000	11 233 990	4 100	56 437	6.61	6.77
44	Velga	1344504	825 000	1 079 702	3 000	9 060	6.28	5.96
45	Petropolis	1399052	2 700 000	4 828 853	9 100	29 594	5.42	5.8

Appendix B (continued)

No.	Name	ID	Actual box office receipts	Expected box office receipts	Actual views	Predicted views	Actual rating	Predicted rating
46	Kunnei	5160746	6 300 000	4 484 640	22 000	23 461	Н/Д	5.25
47	Hosanna	4933779	88 000	857 851	518	8 312	8.50	7.29
48	Patroness	1043835	182 000	764 612	1 100	4 438	Н/Д	6.08
49	The Saint	4959314	172 000	932 162	971	8 734	8.46	7.17
50	I will live	4660300	20 000 000	30 428 381	73 000	113 036	6.90	6.87
51	The Golden Neighbors	1328028	2 400 000	3 599 552	9 100	13 557	4.40	4.16
52	The Lovers	4639557	176 000 000	163 037 230	586 000	624 977	5.90	6.05
53	Peter the Great. the last Tsar and the first Emperor	5116673	275 000 000	529 076 995	1 200 000	2 323 883	7.00	6.65
54	A Similar Person	5074909	554 000	1 759 313	1 700	8 795	6.00	6.42
55	The Stars will show me the Way.	1162851	3 400 000	11 182 553	15 000	56 452	7.30	6.91
56	The Wizards	1008397	71 000 000	80 602 121	289 000	392 378	5.20	6.05
57	Like a Man	4559710	27 000 000	28 027 786	87 000	110 736	7.10	7.24
58	Curse: The Dead Earth	4837786	1 200 000	7 014 696	4 200	29 224	5.20	5.93
59	The Day Before	4321128	841 000	999 946	3 000	4 658	5.40	5.07
60	Zemun	4388687	3 500 000	4 690 601	12 000	15 947	6.20	6.21
61	Angels are Sweet	4694772	196 000	421 344	868	1 957	6.30	6.37
62	A Thousand Cheap Lighters	5016804	1 100 000	1 573 941	3 800	6 038	6.30	6.28
63	Chink: the Tailed Detective	5088766	85 000 000	91 722 969	354 000	369 629	6.20	6.12
64	Elephant	5047298	20 000 000	21 931 224	68 000	90 386	6.60	6.48
65	Jeanne	900055	3 400 000	57 025 433	11 000	252 381	6.70	6.27
66	I'm Rewinding!	5126797	14 000 000	31 093 554	54 000	110 591	7.60	6.78
67	Grisha Subbotin	1410951	4 200 000	5 267 243	17 000	21 104	5.20	4.96
68	The Kingdom Against the Robbers	4853046	42 000 000	53 329 355	189 000	252 476	4.40	5.09
69	Honest Divorce 2	4856074	32 000 000	56 615 622	114 000	194 309	5.90	5.39
70	Winter Season	2000121	1 800 000	2 838 079	5 500	10 619	7.40	6.96
71	Smart Masha	4710763	15 000 000	25 819 110	57 000	79 271	5.90	6.14
72	The Sisters	4661294	22 000 000	35 564 788	75 000	121 605	5.60	5.71
73	Our Winter	5104424	4 600 000	5 732 696	14 000	25 268	6.80	6.9
74	Physics Department	5133122	259 000	18 388 610	1 200	148 222	5.90	5.7
75	F20	5074907	480 000	699 433	1 700	3 363	5.70	5.57

Appendix B (continued)

No.	Name	ID	Actual box office receipts	Expected box office receipts	Actual views	Predicted views	Actual rating	Predicted rating
76	Emperor	5066764	222 000	251 261	621	962	6.20	6.22
77	Pelevin	5195073	1 300 000	1 216 460	3 400	4 139	6.30	6.15
78	Christmas Trees 9	4958223	585 000 000	346 912 974	1 900 000	1 514 429	4.90	4.64
79	By Touch	4975736	29 000 000	38 459 281	100 000	138 538	5.70	5.67
80	A Pool	4913135	2 900 000	4 488 803	10 000	16 685	5.00	5.12
81	Shadow. Take Gordey	4625021	28 000 000	227 049 548	100 000	869 181	6.30	5.66
82	The Nutcracker and the Magic Flute	1435400	115 000 000	112 315 705	483 000	438 012	7.90	7.07
83	Secret Santa	4901797	84 000 000	64 549 229	297 000	290 577	5.70	5.22
84	On Exhale	4474120	927 000	1 064 510	3 300	4 363	6.00	6.04
85	A Wedge Came Together on You…	1392734	1 700 000	1 453 542	5 500	5 554	5.20	5.47
86	Balabanov: The Bell Tower. Requiem	5149154	2 400 000	3 414 803	7 000	24 239	8.10	7.6
87	Nearby	4819385	342 000	530 297	967	1 550	5.60	5.9
88	Masha and the Bear in Cinema: 12 Months	5189169	109 000 000	123 711 842	499 000	461 413	8.50	8.03
89	Naughty Boy 2	5023667	518 000 000	204 013 845	1 700 000	720 589	6.10	6.39
90	Exes: Happy End	5091614	183 000 000	297 062 688	539 000	1 310 025	6.00	6.04
91	Chuck & Huck: A Great Adventure	1346357	253 000 000	238 259 209	956 000	925 131	6.60	6.72
92	Mira	1227997	343 000 000	436 726 746	1 100 000	1 628 008	7.40	6.52
93	Christmas Trees and Needles	4694773	59 000 000	50 819 536	185 000	215 306	7.20	6.6
94	Cheburashka	4370148	7 020 407 894	1 630 133 131	23 510 559	5 909 126	7.30	6.54
95	Music Video Makers	4534033	22 000 000	18 306 527	65 000	100 910	6.00	6.15
96	Umka	5230220	9 900 000	17 223 971	50 000	91 658	6.00	6.35
97	Youth	5074901	3 400 000	6 256 382	11 000	25 064	6.00	6.57
98	Turbosaurs. Winter Adventures	5236779	21 000 000	23 833 876	108 000	110 362	6.00	6.21
99	The Path of the Dream	5079139	1 600 000	1 907 112	4 100	7 473	7.30	7.48
100	Camel’s Arc	5129254	47 000	143 501	206	655	Н/Д	5.99
101	Lily and the Sea / Grand Marin	4478721	1 200 000	2 773 415	3 400	8 571	6.20	6.31
102	Vysotsky: Unknown Pages	1101274	161 000	287 034	711	976	Н/Д	5.4
103	Don’t Bury Me Without Ivan	5237750	10 000 000	11 652 379	34 000	52 359	8.30	7.78

Appendix B (continued)

No.	Name	ID	Actual box office receipts	Expected box office receipts	Actual views	Predicted views	Actual rating	Predicted rating
104	Dear Mom and Dad	5249732	799 000	943 153	3 300	3 739	Н/Д	5.75
105	What Men Talk About...	5045951	299 000 000	316 029 793	857 000	1 758 233	6.50	6.54
106	One Real Day	5047485	4 500 000	11 423 414	16 000	56 322	6.50	6.26
107	Liza	1438922	2 600 000	3 189 975	11 000	13 730	5.80	5.76
108	An Easy Introduction	5245365	1 400 000	1 361 926	4 400	5 769	Н/Д	4.58
109	The White Angel of the Tundra	5254553	2 900 000	4 453 915	10 000	16 843	Н/Д	5.84
110	Free Relationship	4533254	158 000 000	209 632 007	510 000	776 981	6.90	7.04
111	To Be	4682864	123 000 000	114 574 015	436 000	489 336	5.50	5.58
112	Snow. Sister. and Wolverine	5129333	707 000	1 682 918	1 700	8 202	6.30	6.27
113	The Righteous One	4484927	630 000 000	672 236 093	2 100 000	2 418 854	7.60	7.22
114	Naughty	5189350	271 000 000	215 114 973	806 000	881 792	5.40	5.32
115	The Snow Queen: Defrosting	4756225	113 000 000	118 293 013	475 000	518 710	7.40	7.43
116	This Love	5246539	1 900 000	1 812 341	6 500	7 420	5.90	5.87
117	Timbre	5034460	135 000	216 136	480	957	5.90	6.01
118	Women of Altai	5253702	61 000	145 889	218	656	Н/Д	5.89
119	Bugun-Bylyr 2	5078980	7 000 000	6 200 515	22 000	27 076	Н/Д	5.22
120	Checago	5077317	16 000 000	13 522 122	55 000	54 290	6.60	6.78
121	Rabies	1346359	97 000 000	278 004 790	303 000	1 257 287	6.90	6.28
122	Dads Versus Moms	4937051	119 000 000	86 542 200	427 000	346 151	4.50	4.22
123	Zhenya and Vanya at the Edge of the World	5034461	29 000	143 549	100	713	6.50	6.44
124	Where Is the Border?	5129306	3 400 000	3 633 843	26 000	26 229	Н/Д	5.66
125	Nuremberg	905033	328 000 000	428 709 934	1 100 000	1 990 970	6.00	6.25
126	A Healthy Person	2000122	6 900 000	47 188 942	20 000	220 279	7.40	6.82
127	The Unscrupulous in the Countryside	5129252	312 000 000	294 513 410	876 000	1 198 275	6.50	6.34
128	The Detector	4694776	3 400 000	6 281 438	11 000	24 772	5.00	5.06
129	She's gone crazy	4686066	478 000 000	420 677 635	1 600 000	1 748 092	7.80	7.33
130	Escort Girl	5117304	55 000 000	50 478 520	161 000	223 909	4.60	5.12
131	The Cold Race	5030860	18 000	126 849	83	577	5.20	5.07
132	The Holy Archipelago	5101891	11 000 000	6 603 084	38 000	27 522	7.70	7.15
133	Medea	5089031	14 000 000	28 082 456	48 000	125 688	4.50	5.12

Appendix B (continued)

No.	Name	ID	Actual box office receipts	Expected box office receipts	Actual views	Predicted views	Actual rating	Predicted rating
134	Towards the Sun. Along the Rows of Corn	4729320	454 650 603	186 629 797	1 595 307	925 087	6.40	6.68
135	Almost Like Everyone Else	5078979	7 000 000	5 342 655	21 000	20 280	6.40	6.58
136	Sergius Against Evil Spirits: The Sabbath	5126940	13 779 921	49 570 326	46 744	124 553	6.90	6.04
137	The Puppeteer	4528911	12 000 000	13 600 474	48 026	62 152	5.90	5.77
138	Let's Talk About It	5246830	12 000	531 577	62	6 167	5.80	5.44
139	Cats of the Hermitage	4907586	282 172 228	144 228 031	1 200 747	579 042	5.20	5.52
140	Aita	5104425	25 782 268	11 745 092	77773	40 709	6.90	6.77
141	Lovey-Dovey	4715474	48 009 304	131 558 167	165 120	666 435	4.40	5.08
142	My God! I Feel Your Approach!	5266955	238 870	309 805	912	1 299	Н/Д	5.99
143	A Strange House	5140353	18 000 000	12 574 717	61 000	58 847	4.90	4.57
144	Flood	5001203	218 000	1 443 540	557	7 049	5.60	5.84
145	I want to! I will!	4493006	20 396 022	34 036 615	75 750	117 771	6.30	5.79
146	Live broadcast	4400163	881 000	27 460 151	2 634	153 902	6.80	6.2
147	The waves	5059142	82 000	1 711 084	245	10 090	5.40	4.61
148	Doctor	5002368	496 626	527 667	1 411	1 654	8.20	7.24
149	Zoskina gas station	5214412	992 542	22 542 776	3 347	116 461	5.40	5.2
150	Russian Cross	4676639	14 266 106	25 192 400	58 000	95 270	7.50	7.58
151	Challenge	4448519	2 135 633 446	1 831 151 743	6 713 830	7 190 055	7.60	7.27
152	Yaga and The Book of Spells	5194326	224 991 304	179 621 842	943 294	730 191	7.80	6.76
153	Amnesty	5255504	27 969 579	31 987 451	109 670	110 815	6.00	6.25
154	Treasures of Guerrilla Forest	5264352	4 046 002	4 661 213	21 659	17 202	7.10	6.96
155	Yura. the janitor	5134108	51 876 565	52 737 549	177 188	162 620	6.80	6.81
156	The lesser evil	4920455	155 872	26 033 180	501	111 844	7.00	6.99
157	Masquerade	1320002	249 900	803 808	731	2 761	5.80	4.77
158	14+ continues	4489519	93 459 371	32 870 284	315 983	132 375	4.30	5.11
159	Bulgakov. this world is mine!	5277469	46 970	148 747	218	663	7.70	7.79
160	Khitrova. The Sign of Four	5101673	105 619 225	66 927 173	361 396	281 500	7.00	6.41
161	Kaka. directly. another movie	5236765	355 513 600	310 246 710	1 127 420	1 287 789	2.60	3.13
162	Before dawn	5254168	11 819 913	6 411 726	39034	27405	6.10	6.35
163	Masha and The Bear in the movie, say "oh!"	5264991	100 179 923	105 054 457	497854	485 502	8.10	8.4

Appendix B (continued)

No.	Name	ID	Actual box office receipts	Expected box office receipts	Actual views	Predicted views	Actual rating	Predicted rating
164	Knots	5238327	162 086	256 413	559	1 008	5.80	5.81
165	Tin can head	4493055	10 888 045	12 459 348	40 940	47995	6.90	6.28
166	Sun tastes like	5230216	28 469 452	37 316 323	144 074	167 361	7.10	6.94
167	Bullfinch	4485219	56 233 634	54 201 899	185 902	229 464	7.10	7.17
168	Through time	1045993	54 010 222	104 828 315	200 792	429 832	5.40	5.29
169	Mikulai	5074908	2 377 184	2 156 652	8 107	9 878	6.60	6.6
170	Maldives waiting	5079082	57 138 515	53 508 509	189 574	209 671	6.40	5.47
171	Breathing	4715481	114 702 994	204 653 433	403 389	928 421	6.60	6.16
172	Turbosaur, hello, siren!	5278422	12 626 919	19 862 822	66 014	91 165	8.00	7.31
173	Status	1346399	173 817	242 248	562	951	3.90	4.69
174	Johnny	5307214	124 132	260 285	440	962	6.40	6.34
175	Centaur	5235968	123 095 316	95 594 774	400 593	405 597	7.50	6.83
176	Business in Russia	5135024	9 600	128 667	46	617	5.30	5.19
177	To Palych!	5235230	298 541 846	148 807 227	1 065 302	615 172	5.60	5.78
178	Baba Yaga saves the world	4536580	544 129 078	275 566 451	2 138 645	1 227 269	5.80	5.26
179	Witness	5332755	14 454 555	20 879 076	49 196	62 262	3.00	3.85
180	Like stars	5313842	1 972 989	3 639 148	6 340	9 097	5.80	5.97
181	Quest	5118213	12 899 066	13 947 447	55 977	62 621	5.60	5.41
182	Ruslan and Ludmila. more than a fairy tale	5356063	110 201 546	133 865 184	459 620	568 500	8.20	7.48
183	Chizhik-Pyzhik returns	5003770	40 760 444	38 860 073	182 266	167 487	7.30	7.07
184	Nina. girl and piano thieves	4667350	13 150 750	26 628 390	57799	94 541	5.40	4.97
185	Concert canceled	5325618	462 037	1 126 690	1 460	5 357	5.70	5.9
	Total for the period:		21702451281	15374212770	73465643	62363992		

Source: Compiled by the author.

Note: The data is presented as of December 1. 2023. The number of views and the amount of fees may change up, the viewer rating — up and down. An anomaly showing the work of the administrative resource, the selection of rental time and the exclusion of competing films is the film Cheburashka, which, under all other conditions, would have paid off at the box office and received good figures. but taking into account the release on New Year's Eve, in the absence of competition and an active advertising campaign. it showed a high level of fees.