УДК 09.03.04

**НЕЙРОСЕТЕВАЯ СИСТЕМА ПРОГНОЗИРОВАНИЯ РЕНТАБЕЛЬНОСТИ КИНОБИЗНЕСА[[1]](#footnote-1)**

**Чепоков Елизар Сергеевич**

Пермский государственный национальный исследовательский университет,

ВШЭ. 614060, Россия, г. Пермь, ул. Бульвар Гагарина, 37а, eschepokov@mail.ru

В статье представлено описание разработки нейросетевой системы для прогнозирования рентабельности кинобизнеса, а также описывается анализ предыдущих исследований в этой области и существующих решений. Система позволяет предсказать кассовые сборы фильма и шанс окупаемости. С помощью разработанной интеллектуальной системы проведено исследование предметной области, выявлены закономерности, имеющие практической значение.

**Ключевые слова**: искусственный интеллект, нейросетевые технологии, прогнозирование, кинобизнес, фильмы.

**Введение.** Производство кинофильмов, было и остается одним из наиболее рискованных видов предпринимательства, особенно во времена ограничений на посещение общественных мест. Затрачивая огромные бюджеты в несколько сотен миллионов долларов на производство фильма кинокомпании ожидают соответствующие доходы от показа фильма в кинотеатрах и продаж на стриминговых сервисах, но инвесторы отмечают, что предсказать финансовый успех фильма крайне сложно даже для опытных участников рынка. Система прогнозирования способна снизить риски неудачных инвестиций в кинопроизводство на ранних этапах создания кинокартины, а также спрогнозировать целесообразность выпуска кинокартины на экраны кинотеатров с сопутствующими затратами на рекламу и аренду кинозалов.

Одними из первых исследователей, применивших в кинобизнесе метод экономико-математического моделирования, были J. Prag и J. Casavant, которые в 1994 опубликовали статью с сообщением о создании регрессионной модели на основе выборки из 625 американских фильмов [3]. Рассматриваемый ими набор входных переменных включал производственный бюджет, критические обзоры, наличие звезд, наличие франшизы, наличие премий, жанр и рейтинг. Аппарат нейронных сетей для прогнозирования кассовых сборов фильмов был впервые применен в 2002 году американскими учеными R. Sharda и D. Delen. В 2006 году эти же авторы построили модели на основе логистической регрессии, дискриминантного анализа, классификационного и регрессионного дерева, а также нейронной сети, показавшей наилучший результат [5]. В 2010 году они повторили прогнозировании и смогли достичь лучшего результата включив деревья решений и более полную выборку фильмов [6]. Эффективность использования искусственного интеллекта при оценке кассовых сборов фильмов так же была обоснована в работах российских и зарубежных исследователей [1, 7, 8, 11, 12].

Данное исследование является расширением вышеупомянутых работ по созданию динамической системы оценки успеха фильмов. Основная цель настоящей работы заключается в создании множества на основе фильмов, вышедших в прокат, а также создании и обучении нейросети на основе собранных данных. Конечным результатом должна быть нейросетевая система, способная прогнозировать кассовые сборы фильмов с погрешностью не более 30 %.

При построении нейросетевой модели в качестве входных данных было решено выбрать параметры, представленные в таблице 1.

Таблица 1. Входные параметры нейронной сети

|  |  |  |
| --- | --- | --- |
| Номер | Название | Расшифровка, комментарии |
| X1 | Возрастное ограничение фильма | Целое число  (закодированы в числа от 1 до 5  0+: 1; 6+: 2; 12+: 3; 16+: 4; 18+: 5) |
| X2 | Длительность фильма в минутах | Целое число |
| X3 | Сезон выхода фильма | Целое число (Закодированы в числа от 1 до 4: 1-весна, 2-лето, 3-осень, 1-зима) |
| X4 | Выход фильма в период высокой посещаемости кинотеатров | Шкала 0/1, описывает вышел ли фильм в период летних или зимних каникул или иных длинных праздников |
| X5 | Наличие у режиссеров престижных наград | Шкала 0/1, описывает наличие у режиссера наличие таких наград как «Оскар», «Золотой глобус», SAAG и др. |
| X6 | Наличие у сценаристов престижных наград | Шкала 0/1, описывает наличие у сценаристов наличие таких наград как «Оскар», «BAFTA» и др. |
| X7 | Наличие у 3-х звезды на главных ролях престижных наград | Шкала 0/1, описывает наличие у каждого из 3х звезд на главных ролях наличие таких наград как «Оскар», «Золотой глобус», SAAG и др. |
| X8 | Количество оскаров у съемочной группы | Целое число, суммарное количество полученных премий «Оскар» у звезд, режиссера и сценаристов |
| X9 | Основной жанр фильма | Целое число (Закодированы в числа от 1 до 25: 1-Action, 2-Adventure, 3-Drama и т.д.) |
| X10 | Является ли фильм частью франшизы | Шкала 0/1, описывает является ли фильм продолжением франшизы |
| X11 | Бюджет фильма | Положительное число (в долларах США) |

Выходной параметр (Y1) – Кассовые сборы фильма (сгруппированные по множествам $1-2,5 млн, $2,5-5 млн и т. д.).

Обучающее множество было собрано с помощью программы-парсера с интернет-ресурса IMDB [2]. Всего были собраны данные о 1000 фильмах из рейтинга топ-500 и низ-500, после чего была произведена очистка данных от выбросов, пустых и нестандартных данных, таким образом в итоговое множество вошли 632 примера. Данное множество было разделено на обучающее и тестирующее в соотношении 80% к 20%.

Проектирование, обучение, тестирование нейронной сети выполнялись с помощью программы «Нейросимулятор 5» [9] по методике [10]. После оптимизации, спроектированная нейронная сеть представляла собой персептрон, который имеет одиннадцать входных нейронов, один выходной и один скрытый слой с тремя нейронами.

Средняя ошибка тестирования составила 21.7%, что можно считать приемлемым результатом. На рисунке 1 представлена гистограмма, демонстрирующая разницу между фактическими и прогнозируемыми нейросетью кассовыми сборами, среди случайно отобранных 35 фильмов из тестирующего множества.

Рисунок 1. Результат тестирования нейронной сети

Оценка значимости входных параметров была выполнена с помощью программы «Нейросимулятор 5», результат отображен на рисунке 2.

Как видно из рисунка 2, наиболее значимым являются бюджет фильма. На втором и третьем месте по значимости являются количество оскаров у съемочной группы и Наличие наград у сценаристов.

Рисунок 2. Значимость входных параметров нейросетевой модели

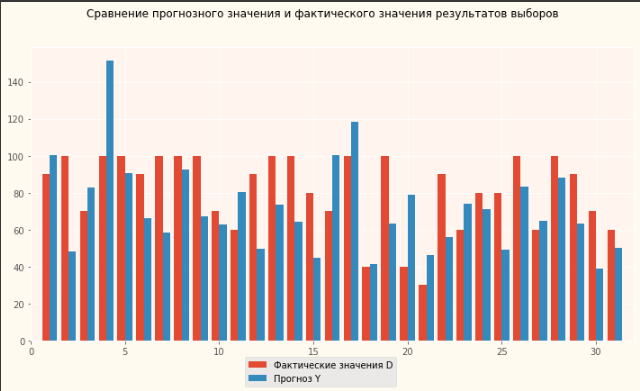
Далее был разработан нейросетевой симулятор с применением языка программирования «Python», с использованием библиотеки «Keras». Средняя ошибка тестирования с помощью разработанной нейросети составила 24.2%. На рисунке 3 представлена гистограмма, демонстрирующая разницу между фактическими и прогнозируемыми нейросетью кассовыми сборами.

Рисунок 3. Результат тестирования нейронной сети

Далее было проведено исследование кинорынка. Исследование производилось с помощью метода «замораживания» [12], суть которого заключается в варьировании значения одного параметра и фиксировании значений всех других параметров. Данный метод позволяет выявить влияние исследуемого параметра на выходной. Для этого были выбраны 3 случайных фильма. Результаты представлены в виде графика на риунке 4.

Рисунок 4. Зависимость кассовых сборов от бюджета фильма

Опираясь на данные результаты, можно с уверенностью сказать, что бюджет влияет на итог кассовых сборов, но данное влияние просматривается по-разному для каждого фильма, в зависимости от совокупности побочных входных данных. Полученные результаты исследования не противоречат действительности, что подтверждает, что спроектированную нейронную сеть можно считать пригодной для прогнозирования кассовых сборов фильмов.

**Заключение.** В ходе проделанного исследования выявлены наиболее значимые критерии, влияющие на успех фильма. Так же продемонстрированы способы использования созданной модели для получения различных статистических данных способствующих повышению качества фильма.

Практическая ценность исследования основывается на необходимости решения важных задач, таких как, научное обоснование и создание эффективно функционирующего механизма прогнозирования коммерческого потенциала кинопроекта, принятия рациональных управленческих решений о целесообразности его реализации, определение направлений и принципов эффективного управления кинематографическим бизнес-процессом, в частности, на ранних стадиях создания кинофильмов.

Библиографический список

1. Eliashberg J., Sawhney M.S. Modeling Goes to Hollywood: Predicting Individual Differences in Movie Enjoyment. Management Science, 1994, vol. 40, iss. 9, pp. 1151–1173.
2. Internet Movie Database [Электронный ресурс], 2022, URL: <https://www.imdb.com/> (дата обращения: 20.03.2022).
3. Prag J., Casavant J. An empirical study of the determinants of revenues and marketing expenditures in the motion picture industry // Journal of Cultural Economics, 1994, vol. 18(3), pp. 217–235.
4. Litman B.R. Predicting Success of Theatrical Movies: An Empirical Study. Journal of Popular Culture, 1983, vol. 16, no. 9, pp. 159–175. doi: 10.1111/j.0022-3840.1983.1604\_159.x 5.
5. Sharda R., Delen D. Predicting box-office success of motion pictures with neural networks // Expert Systems with Applications, 2006, vol. 30, pp. 243–254.
6. Sharda R., Delen D. Predicting the financial success of Hollywood movies using an information fusion approach // Industrial Engeneering Journal, 2010, vol. 21, pp. 30–38.
7. Yasnitskii L.N., Beloborodova N.O., Medvedeva E.Yu. The method for forecasting box-office grosses of movies with neural network // Digest Finance. 2017. Т. 22. № 3 (243). С. 298-309.
8. Yasnitsky, L.N., Mitrofanov, I.A., Immis, M.V. Intelligent System for Prediction Box Office of the Film // Lecture Notes in Networks and Systems. 2020. Vol. 78. Pp. 18-25. <https://doi.org/10.1007/978-3-030-22493-6_3>.
9. Черепанов Ф.М., Ясницкий Л.Н. Нейросимулятор 4.0. Свидетельство о регистрации программы для ЭВМ RUS 2014612546. Заявка № 2014610341 от 15.01.2014.
10. Ясницкий Л.Н. Введение в искусственный интеллект. М.: Издательский центр «Академия», 2005. – 176 с
11. Ясницкий Л.Н., Белобородова Н. О., Медведева Е.Ю. Методика нейросетевого прогнозирования кассовых сборов кинофильмов // Финансовая аналитика: проблемы и решения. 2017. Т. 10. № 4 (334). С. 449-463.
12. Ясницкий Л.Н., Плотников Д.И. Экономико-математическая нейросетевая модель для оптимизации финансовых затрат в кинобизнесе // Фундаментальные исследования. 2016. № 11-2. С. 339-342.

**DEVELOPMENT OF MVP MACHINE LEARNING SYSTEM FOR PREDICTING THE PROFITABILITY OF THE FILM BUSINESS**

***Chepokov Elizar S.***

Perm State University

Str. Gagarin Boulevard, 37a, Perm, Russia, 614060, [eschepokov@mail.ru](mailto:eschepokov@mail.ru)

The article describes the development of a neural network system for predicting the profitability of the film business, and also describes the analysis of previous research in this area and existing solutions. The system allows you to predict the box office of the film and the chance of payback. With the help of the developed intellectual system, a study of the subject area was carried out, patterns of practical importance were identified.

**Keywords:** artificial intelligence, neural network technologies, forecasting, film-making industry, films.

1. © Чепоков Е.С., 2022

   [↑](#footnote-ref-1)