

# Box Office Revenue Prediction

Nhan Nguyen

Computer Science

University of Colorado Boulder

Colorado United States

Nhan.N.Nguyen@colorado.edu

Melissa Pollich

Computer Science

University of Colorado Boulder

Colorado United States

Melissa.Pollich@colorado.edu

Benjamin Breyfogle

Computer Science

University of Colorado Boulder

Colorado United States

Benjamin.Breyfogle@colorado.edu

## PROBLEM STATEMENT/MOTIVATION

As of 2018, it has been reported that the global box office is worth \$41.7 billion. However, when the home entertainment revenue is included, the global film industry is worth over three times that at \$136 billion. Despite the massive profits, movie industry is one of the riskiest markets for investors due to its uncertainty and unpredictability. Once a movie fails to meet expectation, it can potentially place a stress on the financial status of the movie studio, and lead to the withdrawal of funds from the investors. Therefore, through analyzing the performance of movies on the market, we could find the correlation between attributes of a movie and predict the revenue on release.

## CCS CONCEPTS

• **Computing methodologies** → Retrieval model and ranking

## KEYWORDS

Revenue

### ACM Reference format:

N. Nguyen, M. Pollich and B. Breyfogle. 2018. Box Office Revenue Prediction. In *Proceedings of ACM Woodstock conference (WOODSTOCK'18)*. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/1234567890>

\*Article Title Footnote needs to be captured as Title Note

†Author Footnote to be captured as Author Note

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

WOODSTOCK'18, June, 2018, El Paso, Texas USA

## LITERATURE SURVEY

Galvão, M. and Henriques, R. (2018). Forecasting Movie Box Office Profitability. *Journal of Information Systems Engineering & Management*, 3(3), 22. <https://doi.org/10.20897/jisem/2658>

The revenue of a movie can be predicted by constructing a predictive model based on regression, decision trees or neural networks. The paper proposed 3 different variables to use: an interval variable for the value of profitability of individual movies, a categorical variable with multiple classes of profitability values, and a binary variable that shows whether the movie is profitable or not. Using these three variables can increase the accuracy of the prediction and help answer how the market behaves, what appeal to the audiences and the risk associated with movies production.

## PROPOSED WORK

Data regression

Data cleansing

## DATA SET

### Boxofficemojo Alltime Domestic Data

The data contains 16223 unique values of the lifetime gross, ranking, title, studio and production year of Hollywood movies. They were scrapped from BoxofficeMojo's listing and based on domestic gross.

<https://www.kaggle.com/eliasdabbas/boxofficemojo-alltime-domestic-data/version/3>

## **EVALUATION METHOD**

We are looking for a high percent of support and confidence between budget and studios, casts, titles.

## **TOOLS**

Python - Pandas

Github

## **MILESTONES**

Running codes by next week

Graphs of the data set by next weeks

Data integration

Conclusion

## **ACKNOWLEDGMENTS**

## **REFERENCES**

- [1] Galvão, M. and Henriques, R. (2018). Forecasting Movie Box Office Profitability. Journal of Information Systems Engineering & Management, 3(3), 22. <https://doi.org/10.20897/jisem/2658>