

Delivery Report

Predicting Movie Revenue

Domain Impact

The target stakeholders for this project are movie production companies, these companies are producing many different movies at a time in a market that is constantly growing.

As seen by the graphs on the right, the movie industry has already been steadily growing for more than 40 years.

However, as the movie industry grows in scope and revenue, so does the potential financial risk taken by large production studios.

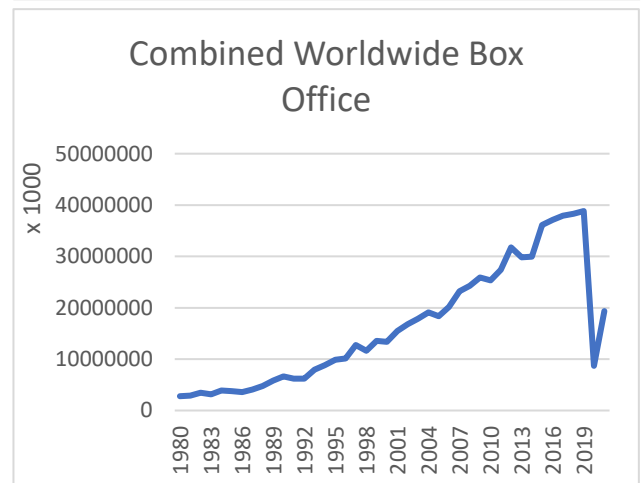
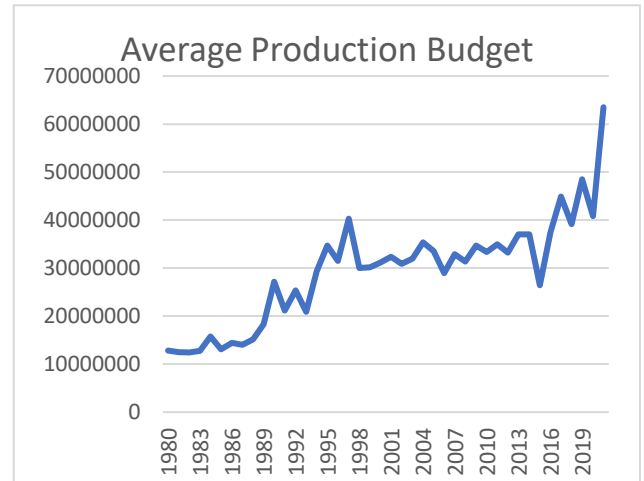
This risk has never been as clear as lately, just by looking at the numbers it is clearly visible that this ever-growing industry took a massive hit recently since COVID-19 forcibly changed the market space.

The financial risks with producing movies will only continue to grow, for this reason it is important to create risk mitigation tools that can be systematically introduced into the cycle of a movie's development.

Introduction

This project's goal is to show the viability of a risk reducing system powered by artificial intelligence, that over time can decrease the financial risks taken and improve the overall performance of the movie industry.

This project is focused on creating an AI model that can predict a movie's revenue before it's release, this makes the model's predictions a perfect asset to be used when considering if a movie should be made, or what the expected profitability of a movie will be.



Data supplied by the-numbers.com

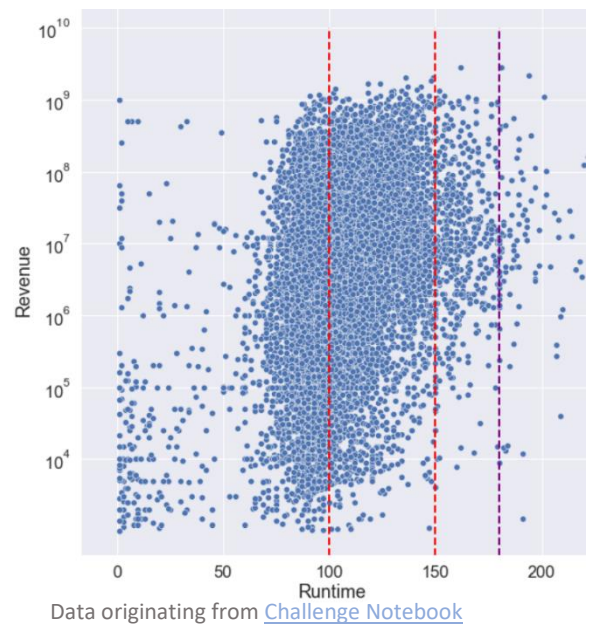
Findings

There were many different discoveries made during this project that can indicate a movie's revenue, most of these have been applied inside of the model to work as a prediction feature.

However, not all features work as well for predicting revenue, one of these was runtime.

"It may not be scientifically proven, but sometimes it feels like movies are, indeed, getting longer." -[Variety.com](#)

There is a good reason why movies are being pushed to be longer recently. As shown in the scatterplot on the right, movies see a base revenue increase between the 100–150 minute mark with very little risk, while runtime can be pushed all the way up to 180 min with a higher but still low amount of risk with an overall higher average revenue.



This already shows one aspect of a movie's risk to revenue that can be reduced right now, while no advanced systems or AI models are needed.

However, there are many more of these features: Budget, Cast/Crew size, Average Actor/Director Popularity, etc. All these aspects of a movie are used to create a model that can give accurate revenue predictions before any movie trailer has even been released.

Showcase

The model used in this project has been trained using many different data features, used cross-validation to create accurate and trusted performance results and has been extensively tuned to create the best model possible using the Support Vector Regression algorithm to predict the revenue of movies.

The final model used has a variance score (R^2) of 0.61, this means that 60% of the variation when predicting revenue can be explained by the model, leaving only 40% of the total revenue a movie earns as a risk.

This model can reduce the profitability risk of a movie from 100% down to only 40% using nothing but publicly available data.

Recommendation

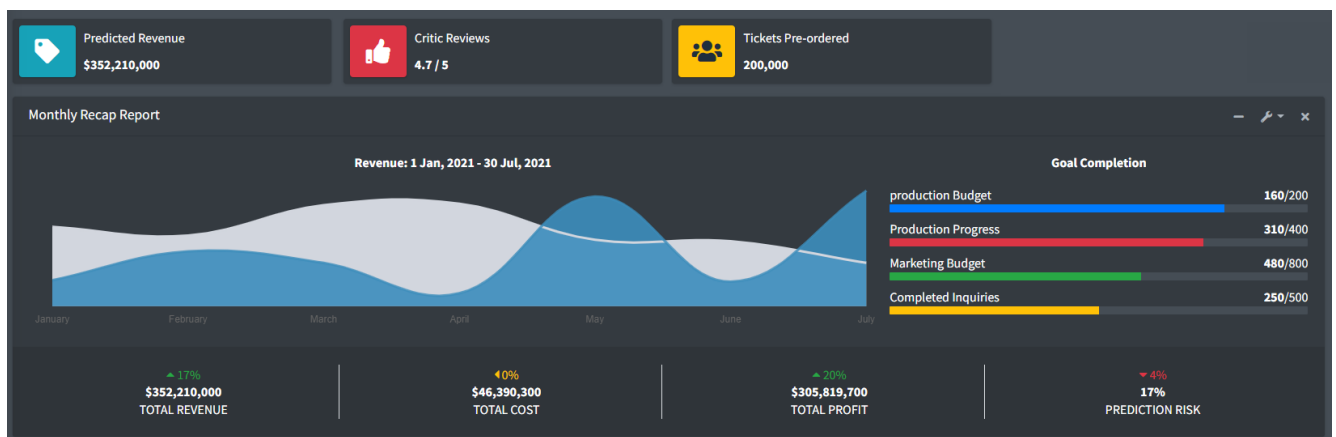
As described in the [Showcase](#), the revenue prediction model can already predict a movie's revenue before its release with fairly good accuracy, however I think it could be even better. This model uses only publicly available data, this means that no marketing strategies or even precise budget values could be used.

My recommendation would be to allow this model to be integrated internally into already existing monitoring/administrative systems and allow it to predict with also privately kept data such as the marketing budget, I think this would have a positive impact on the model's performance and allows it to be used to its fullest potential.

As for integrating this model into the development cycle of movies, predicting the revenue can allow for many more KPI's to be created and used as indicators of how a movie is going to perform, or as a consideration when starting new projects.

In this context it can be directly integrated into already existing dashboards to supply visual feedback.

A possible implementation of these predictions into a dashboard is supplied below.



Design template from [adminLTE 3](#)