Netflix or AMC: Predicting Release Strategies in the Age of Options

Lavada Blanton
Oklahoma State University
MS in Business Analytics and Data Science

AMERICAS 2021 DISCOVERY SUMMIT ONLINE

Netflix or AMC: Predicting Release Strategies in the Age of Options

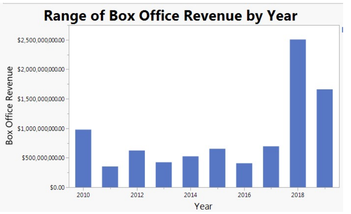
Lavada Blanton
Oklahoma State University

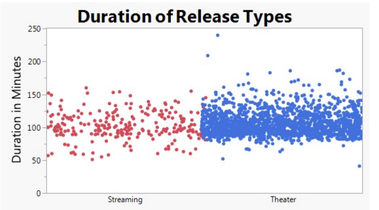
Introduction

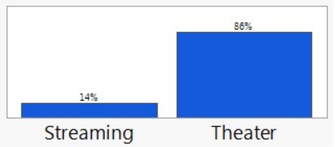
- COVID-19 caused production companies to decide the risks involved in the traditional movie release in theaters.
- Streaming services such as, Netflix, Amazon Prime Video, and HBO Max, were once frowned upon by critics but now have become a respected key player in the film industry.
- This project looks at key indicators in box office success such as, IMDB score, gross revenue, and critic reviews, to evaluate the best “Movie Mix” to be distributed either in theaters or on a streaming service.

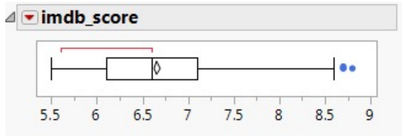
Objectives & Data Description

- The objective of this project is to predict whether movies should be distributed through streaming services or box offices.
- Success was measured by an IMDB score of 5.5 or higher.
- Sampled 1913 movie titles released between 2010 and 2021.
- Various attributes such as Genre, Premiere Date, Duration, Budget, Theatrical Box Office Revenue and Content Rating were used to determine the best “Movie Mix” on the respective release types.









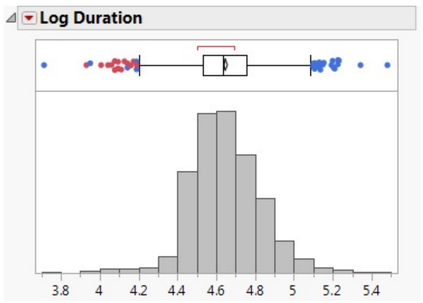
Approach

Preprocessing and Data Collection

- Gathered data from IMDB, Rotten Tomatoes, and the-numbers.com.
- Filtered movies released before 2010 and created a Covid flag for movies released between March 2019 and March 2021.

Pre Transformation :

Summary Statistics	
Mean	105.99214
Std Dev	18.800186
Skewness	1.0262885
Kurtosis	2.9569335

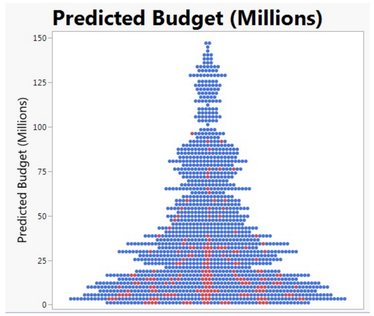


Post Transformation :

Summary Statistics	
Mean	4.6484346
Std Dev	0.1718342
Skewness	0.1417944
Kurtosis	1.6047742

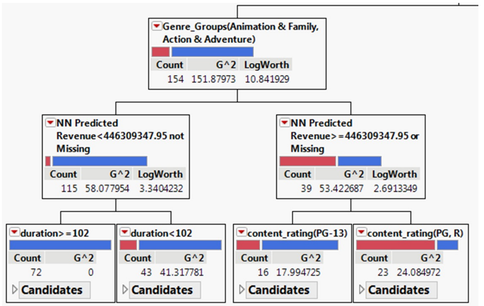
Box Office Predictions

- Predicted worldwide box office revenue for streaming movies.
- Produced by a neural network with sample of theatrical movies.



Decision Tree

- Filtered out dataset to only movies that had an IMDB rating of 5.5 or greater.
- This reduced the sample to around 1400 movies.
- With this sample, created a decision tree with a target of release type (Streaming or Theatrical)



Use Case

- In 2019, the global film industry was worth around $136 billion. This means there is lots of risk and reward involved in every aspect of how a movie is made.
- These ‘Movie Mixes’ could be used by executives in production studios, streaming services, and theaters to make informed decisions about future movies.
- Below is a list of the top 5 highest grossing movies of all time.

Movie Title	IMDB Score	Worldwide Box Office	Year
Avatar	7.8	$2.84 Billion	2009
Avengers: Endgame	8.4	$2.79 Billion	2019
Titanic	7.8	$2.21 Billion	1997
Star Wars: Episode VII The Force Awakens	7.8	$2.06 Billion	2015
Avengers: Infinity War	8.4	$2.04 Billion	2018

AMERICAS 2021 DISCOVERY SUMMIT ONLINE

Netflix or AMC: Predicting Release Strategies in the Age of Options

Lavada Blanton
Oklahoma State University

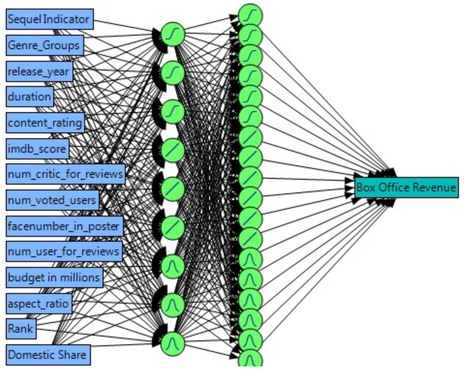
Models

Neural Network

- Neural Network was created to predict worldwide box office sales in streaming movies.

Model
NTanH(6)NLinear(6)NGaussian(6)NTanH2(3)NLinear2(3)NGaussian2(3)

Training		Validation	
gross		gross	
Measures	Value	Measures	Value
RSquare	0.90234	RSquare	0.8738627
RMSE	26959627	RMSE	34918071
Mean Abs Dev	16343294	Mean Abs Dev	18760160
-LogLikelihood	9486.7402	-LogLikelihood	2423.5815
SSE	3.721e+17	SSE	1.573e+17
Sum Freq	512	Sum Freq	129



Stepwise Variable Selection

Source	LogWorth	PValue
Genre_Groups	62.111	0.00000
content_rating	25.306	0.00000
Sequel Indicator	23.561	0.00000
imdb_score	20.802	0.00000
Quarter[premiere]	4.550	0.00003

Summary of Fit

RSquare	0.884957
RSquare Adj	0.881582
Root Mean Square Error	19.58894
Mean of Response	56.2693
Observations (or Sum Wgts)	738

Analysis of Variance

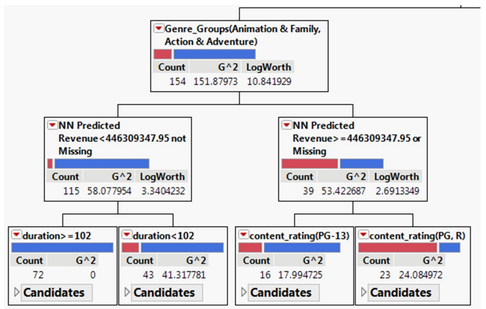
Source	DF	Sum of Squares	Mean Square	F Ratio
Model	21	2113466.4	100641	262.2733
Error	716	274748.3	384	Prob > F
C. Total	737	2388214.7		<.0001*

Bootstrap Forest

- Bootstrap Forest was used to create final recommendations.

Column Contributions

Term	Number of Splits	G^2	Portion
Log Predicted Gross NN	83	278.404483	0.4875
Genre_Groups	82	114.870973	0.2011
Log Duration	85	91.771621	0.1607
Predicted Budget (Millions)	44	33.0263734	0.0578
Quarter[premiere]	46	27.7471136	0.0486
content_rating	34	24.1660201	0.0423
num_voted_users	9	1.10970612	0.0019
director_facebook_likes	0	0	0.0000
actor_1_facebook_likes	0	0	0.0000
Sqrt Budget (Millions)	0	0	0.0000



Results

Streaming

Cheap Feel-Good Comedies

- Average Duration: 91 minutes
- Genre: 59% Comedy, 31% Drama
- Content Rating: 80% R
- Average Budget: $2.1 Million
- Release in 2nd Quarter

Theatrical

Biographical Reenactments on a Budget

- Average Duration: 97 minutes
- Genre: 41% Biography, 35% Adventure
- Content Rating: 92% PG
- Average Budget: $8.6 Million
- Release in 1st or 3rd Quarter

Big Budget Thrills

- Average Duration: 147 minutes
- Genre: 50% Adventure, 50% Horror
- Content Rating: 50% R, 50% PG-13
- Average Budget: $146.5 Million
- Release in 4th Quarter

Adventures for All

- Average Duration: 126 minutes
- Genre: 73% Adventure
- Content Rating: 71% PG-13
- Average Budget: $10.3 Million
- Release in 2nd or 4th Quarter