
What movie do you want to watch next?

predicting movie success using Movies MARK

Everyone loves movies.

Streaming has reshaped cinema and the COVID-19 pandemic has left many of us wondering “what should we watch next?”

Using data-wrangling, programming, and machine learning skills, we plan to answer:

What makes movies successful?





Our Team

Of budding Data Scientists collaborated virtually across Zoom and Slack (as “the_clever_crew”) to bring you this fine work.

- **Maggie Allen**
Presentation
- **Andrew Malony**
GitHub + Graphs
- **Rose Baumann**
Database
- **Kathy Morrissey**
Machine Learning Model

—

But first we must ask
How do we define
Movie Success?



suc·cess

/sək'ses/

noun

1. the accomplishment of an aim or purpose.
"there is a thin line between success and failure"
2. **ARCHAIC**
the good or bad outcome of an undertaking.
"the good or ill success of their maritime enterprises"





suc·cess

/sək'ses/

noun

sadfasdfa

1. Popularity (Proprietary ratings, User Ratings)
2. Estimated Profitability (Revenue-Budget)
3. Awards



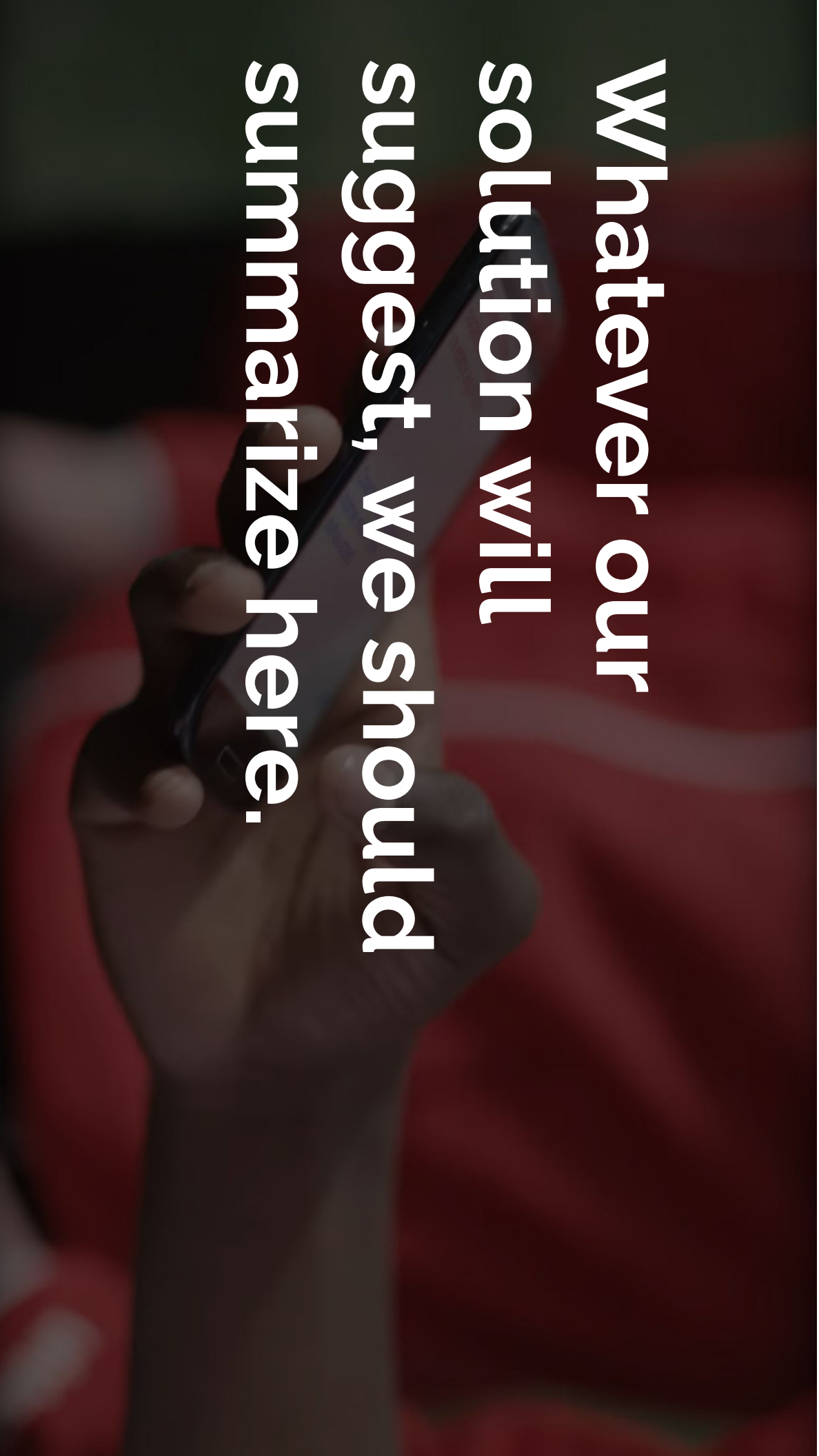


Meet Steel.

He recently had a baby so he has no time to watch a bunch of bad movies. When he finally has a free evening, he wants the first movie he streams to be one he's happy to talk about with his new baby boy.

Let's see what he should look for in his next popcorn night's entertainment...

**Whatever our
solution will
suggest, we should
summarize here.**



A man with long brown hair, wearing a plaid shirt, is sitting on a dark blue couch. He is holding a smartphone in his right hand and looking at it. In the background, a large television screen displays a movie scene with a person in a red shirt. The room is dimly lit, and there is a potted plant on the floor next to the couch.

Then, Steel discovered Movies MARK®

Now he can use our simple list of factors
to determine if this is a movie he would
be happy to watch..

Our Data

- Internet Movies DataBase (IMDb)
- The Movies DataBase (TMDb)
- Kaggle Movies DataSet

Source: <https://www.imdb.com>

<https://www.themoviedb.org/?language=en-US>

<https://www.kaggle.com/rounakbanik/the-movies-dataset>

All
Countries
1970+

Data Clean Up

Datatypes..

tbd

TBD

tbd

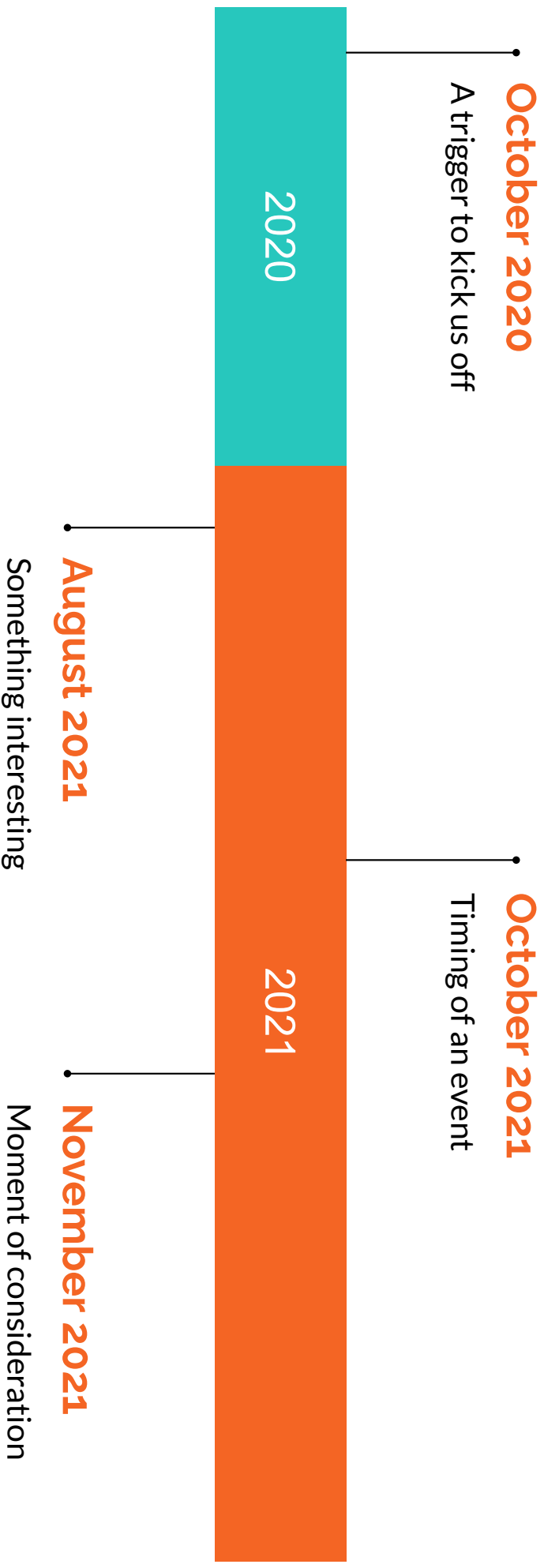
TBD

TBD

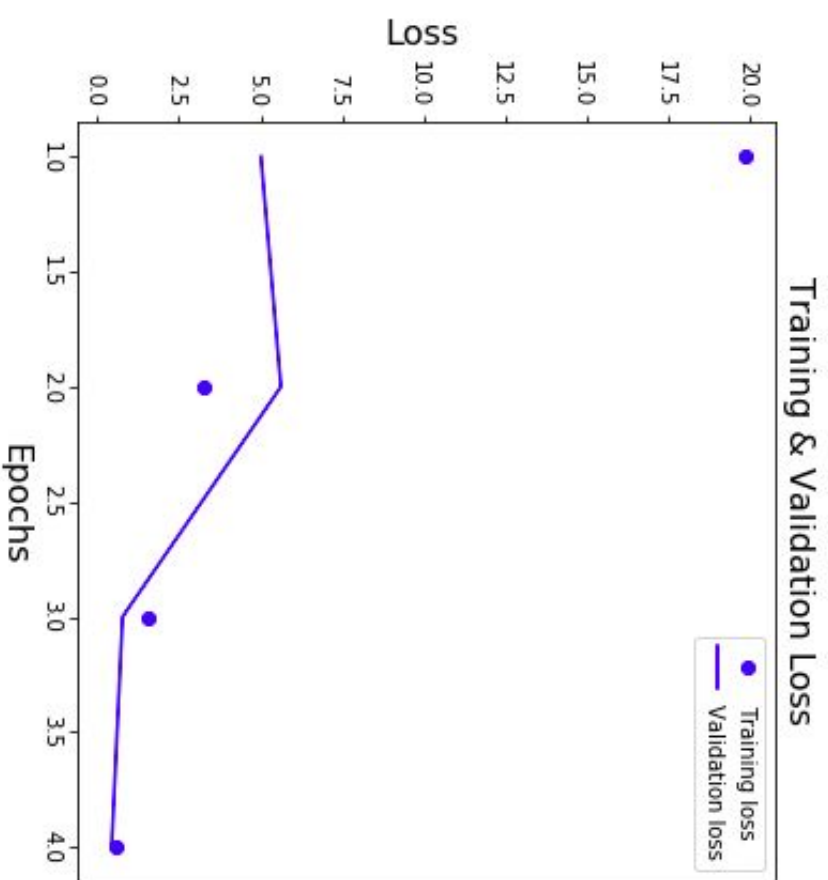
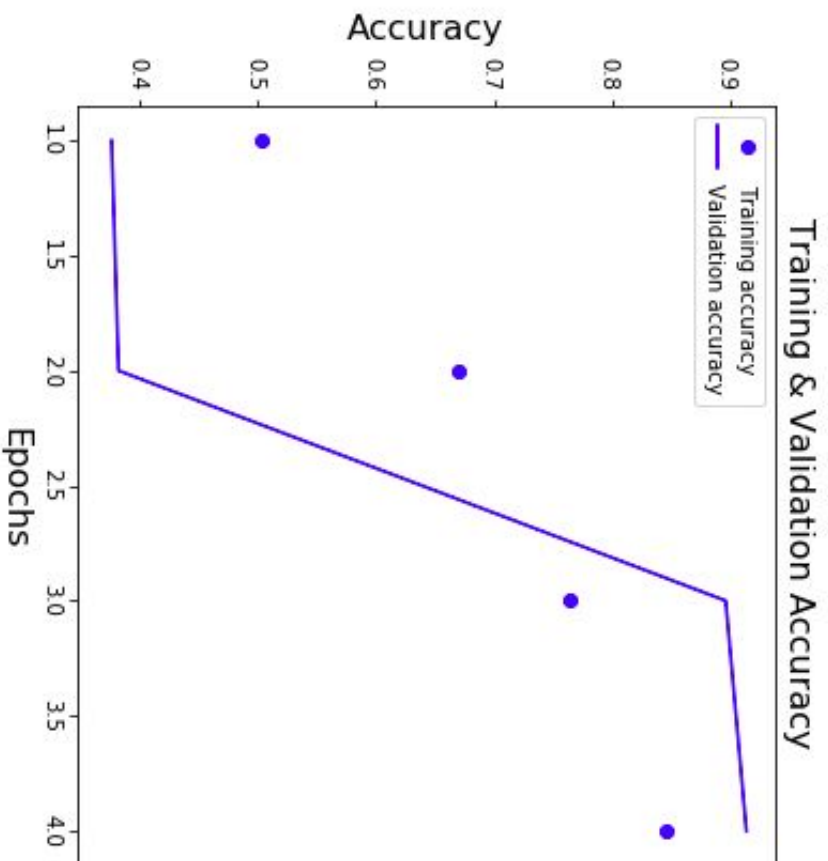
—

Data Source	Column Header	Drop decision
IMDb	actor	

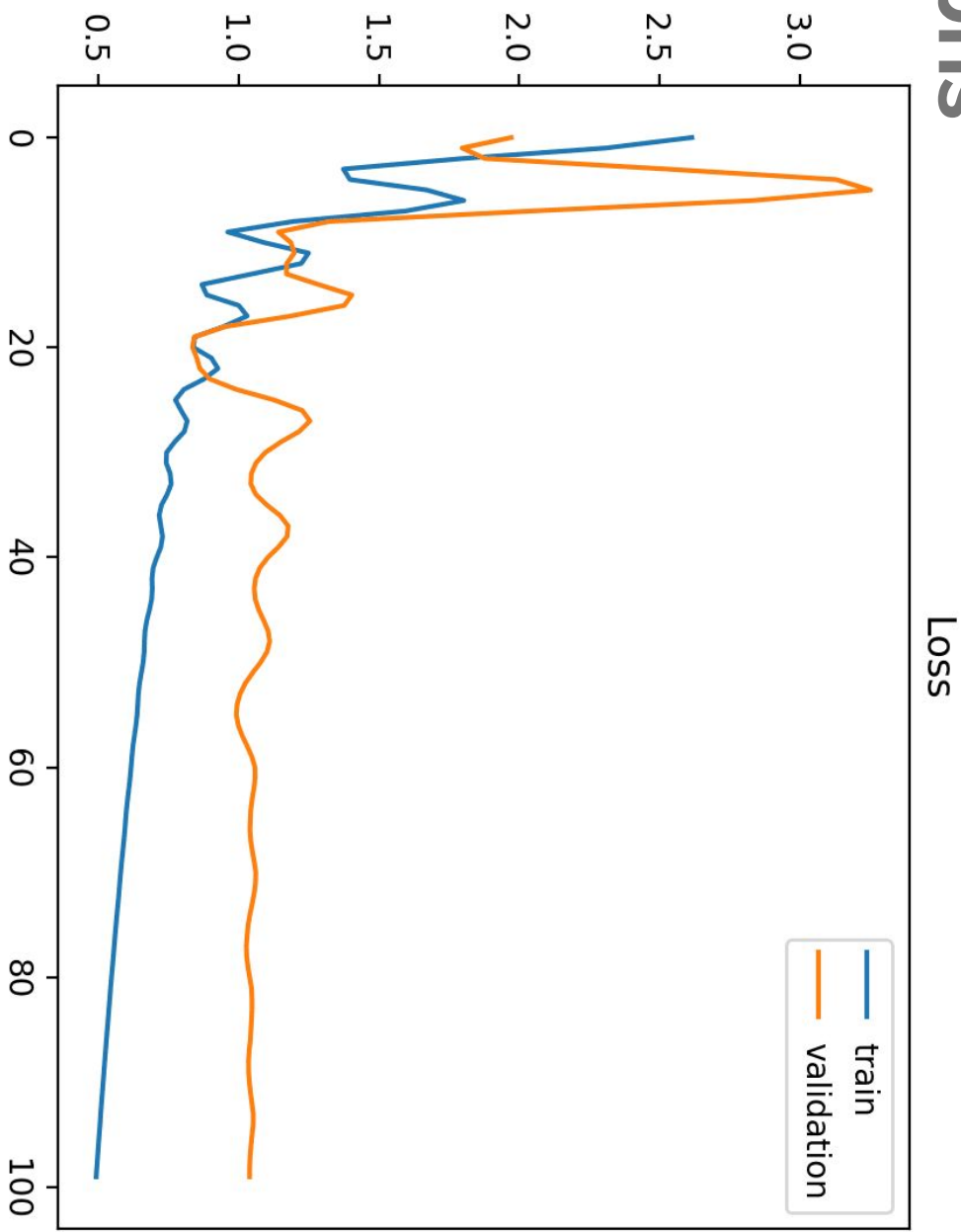
Pretty Graphs



Pretty Graphs



Pretty Grapghs



—
Clever saying to finish us off on a note with a touch of humor..

**Some sort of
summary bullet of
interest.**

Thank you.

