



Predicting Box Office Success: Do Critical Reviews Really Matter?

Lopamudra Pal
lpal@cs.umass.edu

Ly(Harriet) Bui
bui23L@mtholyoke.edu

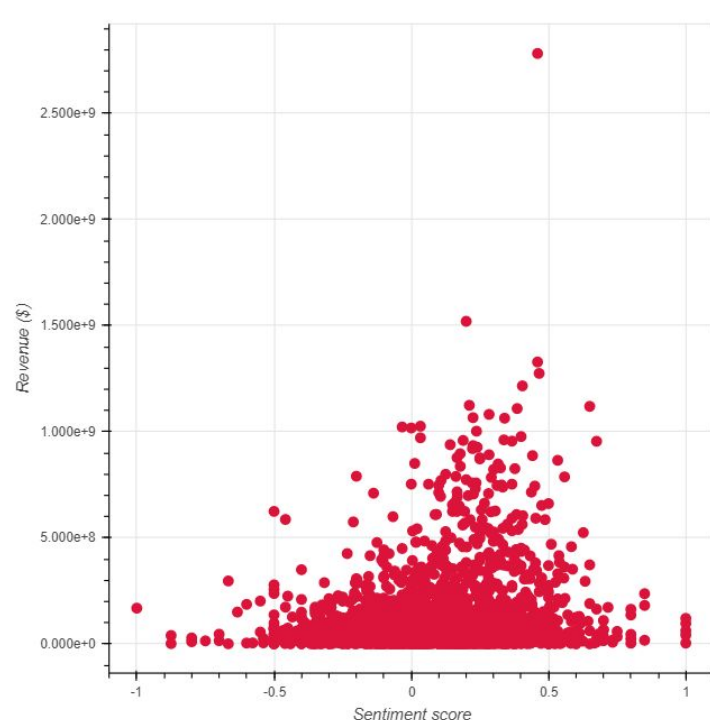
Rishi Mody
rmody@cs.umass.edu

APPROACH: Analysed the impact of movie review sentiments on revenues

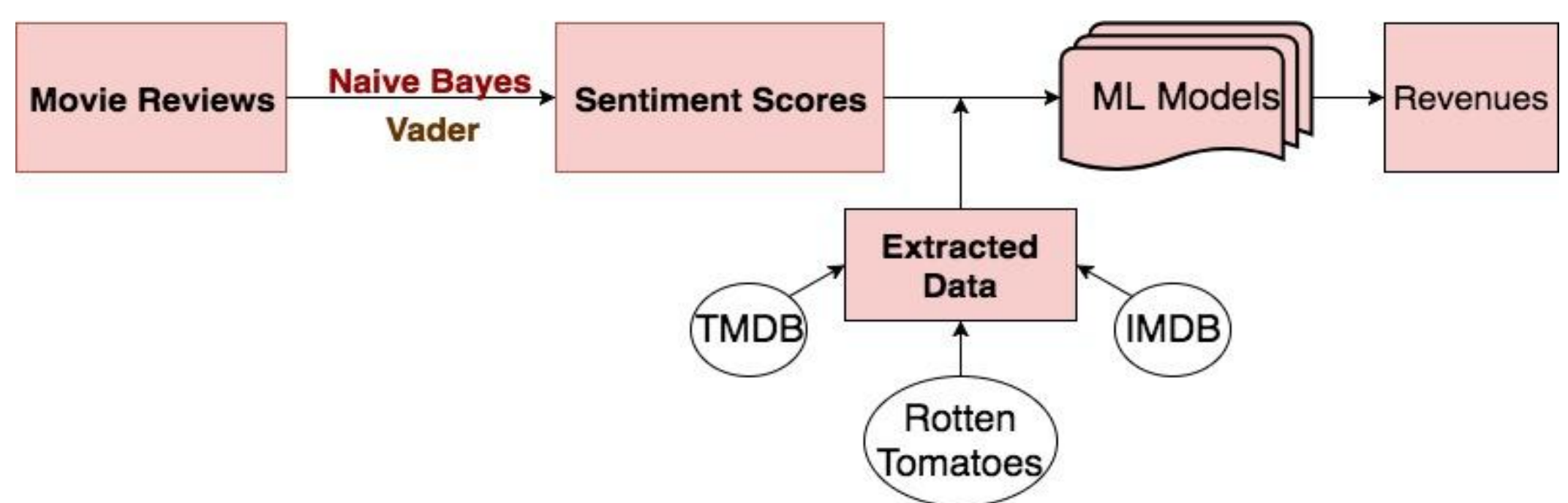
DATASET

1. Sentiment Analysis : Analyzed the Large Movie Review Dataset (Stanford) containing movie metadata and critic reviews.
2. Revenue Prediction: Extracted additional data from TheMovieDataBase (TMDB) and IMDB and rotten tomatoes.

SENTIMENT SCORES vs REVENUES



FLOWCHART



VADER - Valence Aware Dictionary Sentiment Reasoner

- Used for text sentiment analysis that is sensitive to both polarity (positive/negative) and intensity (strength) of emotion.
- Normalization used by Hutto - $\frac{x}{\sqrt{x^2 + \alpha}}$ where x is sum of word sentiment scores and $\alpha = 15$

NAIVE BAYES- Baseline Algorithm for Sentiment Analysis

Independence assumption between features, requires a small amount of training data

$$\text{Sentiment score} = P(y_{\text{positive}} | \mathbf{w}_d) = \frac{P(y_{\text{positive}})P(\mathbf{w}_d | y_d)}{P(\mathbf{w}_d)} - P(y_{\text{negative}} | \mathbf{w}_d) = \frac{P(y_{\text{negative}})P(\mathbf{w}_d | y_d)}{P(\mathbf{w}_d)}$$

RESULTS

- *Fun, Energetic, and Altogether Wonderful*: 0.9717
- *"I am sorry, but in the end, the story is very weak, and this angers me"*: -0.9977

RMSE (DT Regressor) without sentiment score: 111.33

RMSE (DT Regressor) with sentiment score: 91.92

REVENUE PREDICTION ACCURACY



CONCLUSION & FUTURE WORK

- From our experiments we conclude that our correlation between the sentiment scores and movie revenues is not very strong.
- Explore other algorithms to generate sentiment scores and compare results.
- Expect better performance with more data.