

# ESTIMATING BOX OFFICE RATINGS AND REVENUE BASED ON TWITTER SENTIMENT ANALYSIS



By Ramakant, Kshitij, Vijay, Varun, Vivekananda

Under the guidance of Dr. Steffen Heber

NC STATE UNIVERSITY

## Purpose and Motivation

- Predicting the success and opening weekend box office collections using Tweets tweeted before the movie release.
- Tweets give the actual mood of the audience on a movie before its release while the results of IMDB or similar sources are pertained to critics and not the general audience.

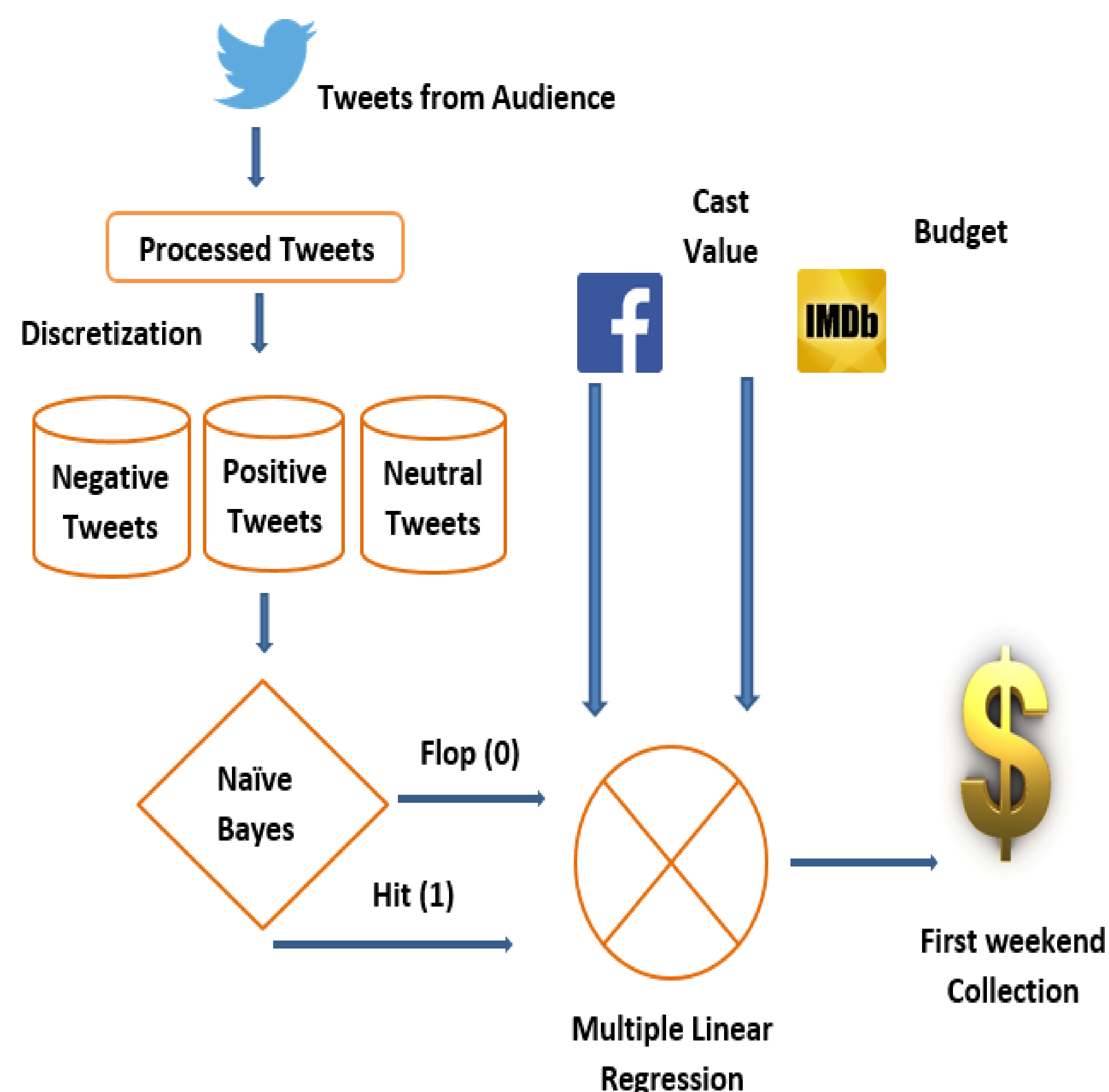
## Why Naive Bayes?

- Simple Model
- Fast to train and classify
- Efficient and popular for text based classification

## Why Multiple Linear Regression?

- This model was used because all the explanatory variables have a linear effect on the estimated collections.

## Methodology

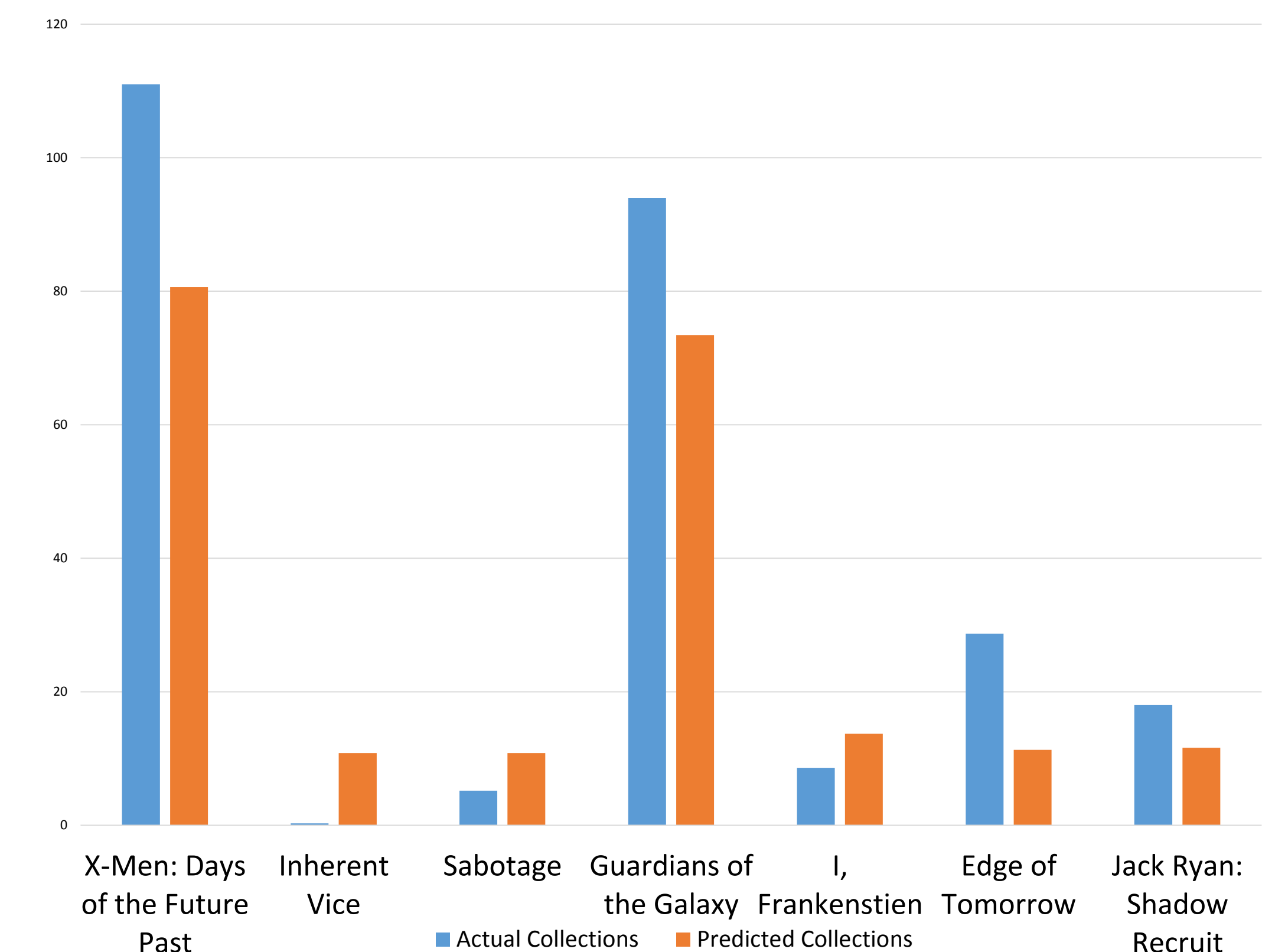


## Deviations from Proposal

- 2 bins were used to classify the movies' success.

## Results

### Actual vs Predicted Collections (in Million USD)



### Actual vs Predicted Hit/Flop

