**CSE 519 -- Data Science**

**Quant Shop Video Review Form**

Name: Rohan Vaish Email: rvaish@cs.stonybrook.edu

ID Number 111447435

Video episode title: Modelling the Movies Episode number (1-8): 2

Of the eight episodes, how many have you previously watched?  (0-7): 3

Give the month and year you started your graduate studies at Stony Brook: August 2017

Did you learn anything by watching it?  (Y/N): Y  Describe stuff you did learn below (if Y):

This was the most interesting episode in my opinion due to my personal interest in Hollywood movies and Golden globe awards. Predicting the gross earning, Christmas day earning and the Golden globe winners had to be tough but fun. Also I learnt that predicting the rating and nominations is still an achievable task, but deciding the winner is almost not possible due to the fact that star power and history can bias the nominations, but wining an award hugely depends on the actual performance.

What did you learn from looking at their project report?

They extracted movie ratings, award information from IMDB, average critics rating, average user ratings from RottenTomatoes, metascore from MetaCritic and number of theatres from BoxOfficeMojo to form their dataset. They drew up a correlation matrix based on factors like previous awards, nominations with 35 more such parameters and their chances of winning to predict how these factors affect their future chances. They considered 2 baseline models, “average by genre by movie gross” and “linear regression by movie gross” (different for high grossing and low grossing movies) and went on to select the latter due to less relative absolute error. After much correct predictions, their model proved to be quite robust.

Rate program quality from 1 (terrible) to 10 (amazing):9 Was it worth watching (Y/N)? Y

What could we do to make the videos better, and more educational?

1. More visual depiction of the data sets.
2. More insight about how to choose the right data sets and models.
3. Incorporating the project reports in the videos in short

|  |
| --- |
|  |