**Offering Marketing Cost to Movie Producers**

Paper No. 1, Oct 7, 2018

Ali Taheri

University of the Pacific

“Hollywood is the land of hunch and the wild guess.” (Litman & Ahn)

Hollywood is one of three big industries in California, and giant companies are doing their best to maximize their profits from their investments. Many factors would affect the revenue of a movie, and there isn’t any analytic form for finding that, so we need to build some models for predicting the required parameters of movies.

The main concern is planning the full path that starts from a concept and would continue with writing script, drawing the storyboard, funding, cast and crew, locations, and production.

After production, they must decide on the plan for marketing and release the title on selected theaters.

Producers are eager to know how much they must invest in marketing to reach their desired value of revenue for their feature films. The purpose of this paper is to offer an advertisement plan for the movie producers.

Authors of "Predicting consumer behavior with Web search," identified variables of search volume, budget, number of screens and HSX for each movie to predict revenue as their outcome. It is clear that there isn’t any search for the movie before the advertisement, so we cannot consider search volume in our analyses as a predictor variable.

For making an offer to producers, we could try to make a model for the cost of marketing based on budget, revenue, HSX and rating of actors and/or directors. With these predictors, ‏we can build models from existing databases of IMDB.

For better results, we could select movies with desirable revenue based on some discussion with the producer. It means producer could mention minimum acceptable value for the ratio of revenue to budget and we would select movies from the IMDB database based on that threshold. We use hold-out cross-validation (30%) technique for the training of our models.

Models could be linear, weighted linear and polynomial. If the mean error rate is high and If we see partial linearity in the graph of revenue vs budget, we could apply some clustering techniques and then make different models for those clusters. For example, we could see some linearity in two regions of movies with revenue lower than 100 million USD and higher than that value, then we could consider two different linear equations for them.

Even If our models are not OK yet, we could think about other predictors same to the genre. It means we could make a specific model for each genre and reach a more precise model for each of them.

After finding a suitable model for predicting the cost of marketing for feature films, we could then choose a target revenue from based on producer’s expectations, then estimate the required marketing budget for each movie. There is a rule of thumb that is to spend the maximum of 50 percent of the rest of the production costs (pre-production, filming, and post-production) for marketing and advertisement. This way we can have an upper limit on marketing cost.

An assumption we made in our model is the choice of financially successful movie titles. Preparing a logistic regression model would be helpful to have an estimate for the chance of flop for movies. We could adjust the cost of marketing and advertisement base on this probability (e.g., decreasing the estimated cost of marketing from our model by chance of flop).