Introduction

Yuqing Zhang

2021/8/16

Netflix is the world's leading streaming entertainment service. Research has shown that Netflix users spend an average of 17.8 minutes per day searching for something to watch on its platform. In this project, our main objective is to minimize the browsing time of Netflix. The average browsing time is our metric of interest.

We have dealt with several possible relevant factors:

- Tile Size (the ratio of a tile's height to the overall screen height)
- Match Score (A prediction of how much you will enjoy watching the show or movie)
- Preview Length (The duration of a show or movie's preview)
- Preview Type (The type of autoplay preview: teaser/trailer or actual content)

After inputting some reasonable values of the above factors, we can obtain simulated data generated by a response surface simulator. For a continuous response like our response variable, the response surface experiment requires investigating and characterizing response surfaces of the expected value of the response.

$$\mathrm{E}[Y] = f\left(x_1, x_2, \ldots, x_{K'}\right)$$

Our main goal of response surface methodology is to build a second-order model to optimize the response, where Response Surface Methodology has been used. We hope to approximate $f\left(x_1, x_2, \ldots, x_{K'}\right)$ by

$$\eta = \beta_0 + \sum_{j=1}^{K'} \beta_j x_j + \sum_{j<l} \beta_{jl} x_j x_l + \sum_{j=1}^{K'} \beta_{jj} x_j^2$$

We hope to identify the coordinate of the factors that optimize the response, which in this case is to minimize the browse time.

To avoid any waste of sources, the first thing we must do is to do factor screening. Usually, we do a 2-level factorial experiment. After doing the regression, we can identify the non-significant factors which we can ignore in the following steps.

To make sure we are exploring in the region meeting the requirement of convexity, we must check the curvature, by comparing the center point condition with the factorial points conditions. Only the presence of quadratic curvature can tell that you are in the vicinity of the optimum. If not, we need to use the method of steepest ascent/descent to move from our original region towards the vicinity of the optimum.

After reaching the vicinity of the optimum, we can finally fit the full second-order model to do the response surface experiment. We can identify the stationary point as the optimum. Then the optimum value yields the optimal response value.