Phase 1 – Executive Summary

Netflix is the world's leading streaming entertainment service founded in 1997, headquartered in Los Gatos, California. People can enjoy TV series, documentaries and feature films in various genres and languages. As of July 2021, Netflix has 209 million paid members in more than 190 countries. We will start an experiment inspired by Netflix, which includes a hypothetical question and a web-based response surface simulator.

The goal of the experiment is to minimize the average time for Netflix users to browse the homepage, and thus improve the user viscosity. Therefore, MOI of this experiment is the average browsing time. The response variable is the browsing spent by Netflix users on the homepage. We find four factors that affect the average browsing time: Tile Size (denoted as Tile.Size), Match Score (Match.Score), Preview Length (Prev.Length), and Preview Type (Prev.Type).

In order to study how to reduce the average browsing time of the Netflix homepage, we divide the experiment into three phases.

Phase 1 is factor screening. The goal of Phase 1 is to find the significant factors, and remove the insignificant ones. After a 2^k factorial experiment, the significant factors include Preview Length, Match Score, and interaction of Preview Length : Match Score. Tile Size is not significant to minimize the average browsing time.

Phase 2 is to find the optimal regions using the method of steepest descent. The central point is [85,65]. For Preview Length, the optimal area is [75,95]. For Match Score, it is [55,75].

Phase 3 is to find the optimal point by conducting central component design and using a second order response surface model. The optimal point of Preview Length is located at around 37. The optimal point of Match Score is around 89. The predicted value of the average browsing time is about 8 minutes. The estimated 95% confidence interval for the expected browsing time is [4.08, 11.24].

Phase 2 – Introduction

As a top streaming entertainment company, Netflix has attracted many users who love movies and TV shows. The company is also committed to making its content more attractive to movie-lovers, such as reducing the time users spend browsing the homepage as much as possible, so that users can quickly find and enjoy the movies they are interested in. Therefore, this experiment is designed to minimize the average browsing time in order to retain users.

There are four factors that are influencing the average browsing time, including Tile Size, Match Score, Preview Length and Preview Type. We will find the significant factors in Phase 1 through factor screening, and control the relatively insignificant factors to complete subsequent experiments. We will use the method of steepest descent (MSD) in Phase 2 to locate the optimal regions, and find the relationship between significant factors and the average browsing time. In Phase 3, we will introduce axial factors and conduct the response surface methodology (RSM) to optimize the response.

RSM is used to explore the relationship between the explanatory variables and one or more response variables. The primary purpose of RSM is response optimization. It designs sequential experimentations to obtain an optimal response.

In our Netflix experiment, Phase 2 uses MSD to locate the optimal regions of the factors, and then conduct a curvature test. In Phase 2, information gained from the last step/experiments can help to inform the future steps/experiments. We conduct 9 experiments/steps to locate the optimal response area. We observe changes of the average browsing time in the MOI plot, and successfully find the optimal region in the fifth experiment/step, where the center point is close to [85,65]. As axial factors are added, we will use RSM to find the optimal point.

Phase 3 – Factor Screening

The purpose of Phase 1 is to find the significant factors among the four design factors in the Netflix experiment. We will explore the significant factors and conduct subsequent experiments, and set those insignificant factors as default values.

We are not able to run all experiments with n = 100. Thus we will use 2k factorial experiment design in Phase 1, and investigate 3 design factors in 8 conditions. The result of the 2k factorial experiment design is more accurate than the 2k-p fractional factorial experiment design which only requires 4 conditions.

Prev.Length	Match.Score	Tile.Size	Prev.Type
100	80	0.1	TT
120	80	0.1	TT
100	100	0.1	TT
120	100	0.1	TT
100	80	0.3	TT

120	80	0.3	TT
100	100	0.3	TT
120	100	0.3	TT

Table 1: Input data for factor screening

After the factor screening design, only 2 factors have significantly small p-values that are <2e-16: Preview Length and Match Score. Therefore, for the purpose of minimizing the average browsing time, Preview Length and Match Score are important to the result. The estimated beta value for Preview Length is 0.94606, while for Match Score is 0.96639. In the follow-up experiments, we will set Tile Size as default value 0.2 and Preview Type as TT.

In conclusion, we use 2k factorial experiment design in Phase 1 and successfully find the significant factors: Preview Length and Match Score. We will use the data of Preview Length and Match Score, and conduct the method of steepest descent (MSD) and curvature test in Phase 2.

Phase 4 – Method of Steepest Descent

In Phase 2, we use the method of steepest descent (MSD) and curvature test in order to locate the optimal region that will minimize the average browsing time.

After factor screening in Phase 1, we find that Preview Length and Match Score are significant to the result. Therefore, we input the following data to conduct MSD:

Prev.Length	Match.Score	Tile.Size	Prev.Type
100	80	0.2	TT
120	80	0.2	TT
100	100	0.2	TT
120	100	0.2	TT
100	80	0.2	TT
120	80	0.2	TT
100	100	0.2	TT
120	100	0.2	TT

Table 2: input data for MSD

We then start a series of experiments to calculate the average browsing time, and see which step will minimize the average browsing time. We set the step size of the Preview Length equal to 5 seconds, and repeat the process of xi+1=xi-λg (*where* = 5s/prev.length, g is the transpose matrix of prev.length *and* match.score) until we find the lowest point of the average browsing time in Step 5, which is [85,65].

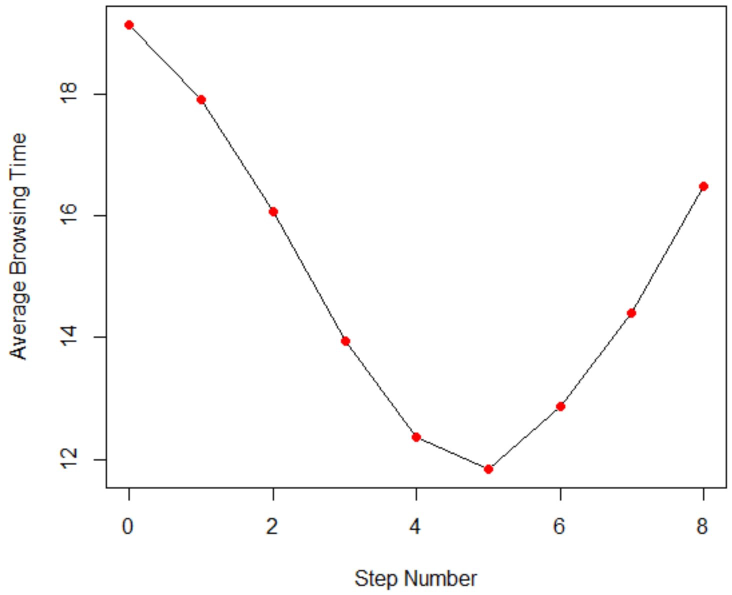

Table 3: changes in the average browsing time when using MSD

We repeat 9 steps with 9 conditions to find the optimal region. Conditions tested from Step 0 to Step 8 are as followed:

Prev.Length	Match.Score	Tile.Size	Prev.Type
110	90	0.2	TT
105	85	0.2	TT
100	80	0.2	TT
95	75	0.2	TT
90	70	0.2	TT
85	65	0.2	TT
80	60	0.2	TT
75	55	0.2	TT
70	50	0.2	TT

Table 4: List of conditions from step 0 and step 8 in MSD

Then we consider the optimal region for Preview Length is [75,95], and the optimal region for Match Score is [55,75]. That is to say, Preview Length has a high level of 95 and a low level of 75. Match Score has a high level of 55 and a low level of 75. We conduct a test for curvature that contains the central point condition and two-level factorial conditions to see whether these intervals contain the optimal points. This is a second order model that contains the main effects of Preview Length and Match Score, as well as the two-factor interaction.

Prev.Length	Match.Score	Tile.Size	Prev.Type
85	65	0.2	TT
75	55	0.2	TT
75	75	0.2	TT
95	55	0.2	TT
95	75	0.2	TT

Table 5: List of conditions in the curvature test

The result shows an extremely small p-value for [85,65], which is less than 2e-16. Therefore, we conclude that the optimal region tested includes the optimal point. The next experiment in Phase 3 is to optimize the response.

Phase 5 – Response Optimization

In Phase 3, we conduct a central component design (CCD) and fit the second order response surface model to identify the location of the optimal point.

The data we entered contains a central point condition, 4 two-level factorial conditions, and 4 axial conditions. The value a = 1 is chose to balance both practical and statistical concerns.

Prev.Length	Match.Score	Tile.Size	Prev.Type
85	65	0.2	TT
75	55	0.2	TT
75	75	0.2	TT
95	55	0.2	TT
95	75	0.2	TT
75	65	0.2	TT
95	65	0.2	TT
85	55	0.2	TT
85	75	0.2	TT

Table 6: List of conditions for CCD

The goal of CCD is to fit a full second response surface model. The p-values for Preview Length, Match Score, I(Prev.Length^2), I(Match.Score^2) and the interaction are all extremely small, which means that the full second-order response surface model has a good fitness to the conditions. The R-squared is 0.7093, which reveals that 70.93% of the data fit the model well. Then we visualize this second-order response surface model, and find the optimal value.

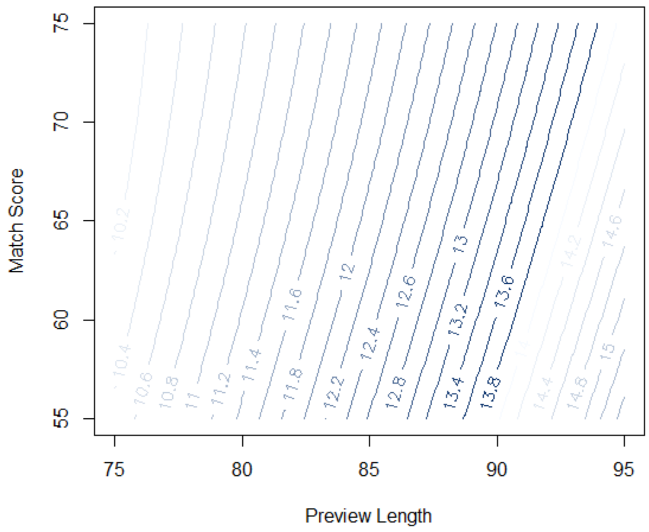

Table 7: The 2D contour plot that visualizes the response surface

The value of predicted browsing time is 7.662668. The optimal value for Preview Length is 36.91301, while the optimal value for Match Score is 88.84779. The 95% confidence interval to minimize the expected browsing time is [4.08,11.24], meaning that there is a probability of 95% that the location of the optimum is in this interval.