

Dear Reviewer 1:

I will respond to each point by quoting each of your questions.

- Not found feature descriptions for SVM and XGBoost

The feature descriptions for both SVM and XGBoost are discussed in section 4.2 Advert Viewing Time Calculation, further shown in Table 3. The table, however, is complicated and a bit unreadable in its previous form and I have rewritten the methodology section to express better the feature vectors. However, to respond to your question directly, I will explain the feature vectors used in the first version below.

- Total Seconds Model: A single feature, the total number of seconds observed for the related advert.
- Day/Primetime Model: Each day is split into two features, Primetime and Non-Primetime. Then the same is added with Weekdays, Weekends, Holidays and Rest days. Total of 22 features.
  - Monday Primetime
  - Monday Non-Primetime
  - Tuesday Primetime
  - Tuesday Non-Primetime
  - Wednesday Primetime
  - Wednesday Non-Primetime
  - Thursday Primetime
  - Thursday Non-Primetime
  - Friday Primetime
  - Friday Non-Primetime
  - Saturday Primetime
  - Saturday Non-Primetime

- Sunday Primetime
- Sunday Non-Primetime
- Weekday Primetime
- Weekday Non-Primetime
- Weekend Primetime
- Weekend Non-Primetime
- Holidays Primetime
- Holidays Non-Primetime
- Rest Day Primetime
- Rest Day Non-Primetime

That is to say, in order to answer your final request to add Multi-Layer Perceptron and Logistic Regression, additional experiments were done, so this is merely explanatory of the experiments done in the first version of the manuscript. The feature vectors are now described in section 4.4. Input Data, and in subsections 4.4.1 Advert Viewing Time, and 4.4.2 Demographic Data. The detailed tables are in Appendix A and Tables A.18 and A.19.

- In table 1 and table 2, explanations of "Categories" are not found. - Therefore, readers can not experiment the same as this paper. - How did you divide data into categories? Hands? Rule base?

The contents of table 1 and 2 and the categorization of the data are explained in detail in section 4.1.2 Data Categorization of this paper already. After the revision and consequential reorganizing of sections, however, they are included in the renamed section 4.3.2. Prediction Target Data Categorization, and the re-numbered tables 3 and 4.

- In this paper, authors use K-fold cross-validation. What K value did you use in this research? - Why did you use K-fold cross-validation instead of training data, development data, test data?

- Please write hyper-parameters of SVM and XGBoost that were used for experiments.
- Additionally, how did you decide the hyper-parameters?

For our K-fold Cross Validation, we used a value of 5. This value is appropriate for the size of the data (3000 users and 36 usable products). We included this clarification in the new section 5.2. Experiment Parameters.

K-Fold Cross Validation was used in order to avoid leaving a static set of data that is unused in the training, avoiding possible bias in the testing sample. However, due to limitations on calculation processing power, the requested Neural Network experiments were done with simple training data and test data.

A more detailed description of the specific parameters was added in the new section 5.2. Experiment Parameters.

- I did not understand which feature is useful for predictions.

The main feature that was decided to be effective in this experiment was the total number of seconds of the advertisement being watched. In the new experiment design requested by another reviewer, however, demographic data is also included. This is explained further below, and in the revised section 4. Methodology, and the new section 4.1. Experiment Design Overview.

- Why do the authors show only results of category 2 and category 4?

The results from other categories produced similarly uninteresting results, and provided little to no material for discussion. Furthermore, since the purpose of advertisement is to increase sales, we thought appropriate to mainly discuss the results on customers that were purchasing, or that were possibly convinced to purchase because of the advert. However, in the revision of this paper we included all categories in section 5. Results due to the request of reviewers.

- Why did you use SVM and XGBoost?

The reason we used SVM and XGBoost is that they are considered well performing machines in the machine learning field for the size of the available data that we had. Other methodologies are generally thought perform better under the assumption that databases on a different order of magnitude are used. This explanation is also added to the manuscript in section 4.5 Prediction Models, and in section 8. Limitations.

The novel point of this study is using machine learning for TV commercial adverts and purchase behavior. Therefore, methodology concerning machine learning have not the novelty. Thus, authors should indicate which machine learning method is useful for this task. - Please add experiments using Multi-Layer Perceptron and Logistic Regression.

As asked, we have made new experiments using Logistic Regression. However, due to constraints of processing time for Neural Networks, we have not included them in the revised manuscript. The new design for the experiments (as asked by another reviewer) is described in detail in the new sections of 4. Methodology, as well as summarized in 4.1. Experiment Design Overview and in 5.1. Model Training. We also included a short summary below. After observing our results, however, in average, SVM outperformed XGBoost and Logistic Regression experiments. This is because, consistent with our original views when choosing the methodology, SVM is the most appropriate machine learning method when presented with the size and structure of our dataset. State of the art Neural Networks require datasets orders of magnitude much larger to perform their best. This was also added to a new section 8. Limitations, of our revised manuscript.

The experiments are described below:

- Product Based Models:
  - Advert Viewing Time
  - Advert Viewing Time, Demographics, Purchase Intention

- Demographics, Purchase intention

- User Based Models:

- Advert Viewing Time

- Advert Viewing Time, Demographics, Purchase Intention

- Demographics, Purchase intention

\* Of course, removing the Purchase Intention from the feature vector when it is the Prediction Target

With 3 methodologies:

- SVM

- XGBoost

- Logistic Regression

The models and features defined as follows:

- Advert Viewing Time:

- Weekday Configuration:

- \* Monday

- \* Tuesday

- \* Wednesday

- \* Thursday

- \* Friday

- \* Saturday

- \* Sunday

- Weekday and Time Slot Configuration:

- \* Monday Primetime

- \* Monday Non-Primetime

- \* Tuesday Primetime

- \* Tuesday Non-Primetime
- \* Wednesday Primetime
- \* Wednesday Non-Primetime
- \* Thursday Primetime
- \* Thursday Non-Primetime
- \* Friday Primetime
- \* Friday Non-Primetime
- \* Saturday Primetime
- \* Saturday Non-Primetime
- \* Sunday Primetime
- \* Sunday Non-Primetime

- Demographics:

- Age:

- \* 18 to 25 years old
    - \* 26 to 35 years old
    - \* 36 to 45 years old
    - \* 46 to 55 years old
    - \* 56 or older

- Sex:

- \* Male
    - \* Female

- Marital Status:

- \* Single
    - \* Married
    - \* Divorced or Widowed

- Parental status:

- \* Parent

- \* Not a Parent
- Income Bracket:
  - \* Not disclosed
  - \* No Income
  - \* Under 1,000,000 yen
  - \* From 1,000,000 yen to 2,000,000 yen
  - \* From 2,000,000 yen to 3,000,000 yen
  - \* From 3,000,000 yen to 4,000,000 yen
  - \* From 4,000,000 yen to 5,000,000 yen
  - \* From 5,000,000 yen to 6,000,000 yen
  - \* From 6,000,000 yen to 7,000,000 yen
  - \* From 7,000,000 yen to 10,000,000 yen
  - \* From 10,000,000 yen to 15,000,000 yen
  - \* From 15,000,000 yen to 20,000,000 yen
  - \* Over 20,000,000 yen
- Purchase Intention:
  - January Survey response
  - March Survey response

I hope this was helpful to answer your questions, and that you are satisfied with the additions we made to the paper.