Temporal Externalities of Attention on TV Advertising

ABSTRACT. In this project, I investigate temporal externalities of attention on TV Ads. I use a dataset from TVision, which is unique in that it contains both 'traditional' features (viewing histories, user and item characteristics), as well as the attention ratio of viewers on Ads. I find that viewers' attention is persistent over a short period of time. I then apply the knowledge to incorporate temporal features such as previous Ad and previous attention to the traditional latent factor techniques, resulting highly effective advertising recommendations.

1 Introduction

Online and TV advertising have largely adopted videos as a major source of advertisement. The prerequisite of effective advertising is the engagement of viewers. More engagement increases the likelihood of purchases in the future. In practice, audiences are usually exposed to not only one ad, but a sequence of ads. This implies there might be an intertemporal externality effect of ads on viewer's engagement level. Understanding how one ad affects viewers' attention to subsequent ads is crucial to maximize the effectiveness of ads.

In this project, I take advantage of a novel dataset that directly measures views' attention to TV ads at home. I find that viewer's attention is persistent during a short period of time. Stimulating Ads have positive externalties on views' engagement for subsequent ads. Then I apply the insight of temporal externalities to predict the attention of a particular viewer to a particular Ad. I adopt the Latent Factor Model (LFM) and the Factorizing Personalized Markov Chain (FPMC) in the recommender system to make predictions. I find that incorporating the previous Ad and engagement improve the model performance by 38.1% in terms of the MSE on the test set. It also boosts the speed of convergence, increasing computational efficiency.

2 Related Literature

Mostly related to this project is the works on audience externalities. Gomes et al. (2009) [1] showed audience externalities to be economically and statistically significant in search advertising. Wilbur, Xu Kempe (2013) [2] studied the audience externalities in TV advertising. They observed that ads can cause viewers to switch channels which reduces the availability of audiences to subsequent ads. They proposed the Audience Value Maximization Algorithm, using the audience presence data to get the advertisers and audience value. McGranaghan, Liaukonyte and Wilbur (2022) [3] explored the unique data from TVision and found that recreational product ads preserve audience tuning and presence, and the attention helps predict brand search lift after ads. This project focuses on externalities on viewers' attention. While viewers presence serves as a proxy for engagement, there are huge differences between them. In the presence and attention data, I observed that 70% of the audiences didn't pay attention to the ads. Hence, TV advertisers can boost the effectiveness of ads by maximizing total engagements instead of the number of viewings.

The project is broadly related to the temporal behavior on advertising. Haugtvedt (1990) [4] shows that Ads with high personal relevance gain more attention. Joo, Liu and Wilbur (2020) [5] found that there is a divergence in consumer liking and wanting in response to sequence of ads. The message liking was highest early in the sequence, whereas wanting of the promoted item was highest late in the sequence. McGranaghan, Liaukonyte and Wilbur (2022) [3] showed that the attention decreases in the first two time slots after a program but does not always decrease after the third slot. This might suggest that attention to a sequence of ads is not always decreasing over time, and personalized content or other ad features can bring the audiences back to the ads.

I use the Latent Factor and Personalized Markov Chain models to make predictions on the attention level of Ads. The Latent Factor Model performs recommendation by projecting users and items into some low-dimensional space. The Personalized Markov Chain Model incorporates temporal factors. The most related literature on recommendation systems is the one class recommendation and rating predictions using supervised models. Linden et al. (2003) [6] showed that amazon uses item-to-item collaborative filtering techniques to recommend items to targeted customers. McAuley et al. (2013) [7] Modeled the visual evolution of fashion trends with one-class collaborative filtering. McAuley et al. (2019) [8] developed models to generate personalized recipes from historical user preferences. This study mostly related to these previous works, but I apply the techniques in a different setting. I develop models to recommend personalized Ads based on the historical user preferences.

Table 1. TV Advertising Data in a Month Period

Users	3397
Total Ads	830,174
Total Brands	3043
Total Industries	170
Effective Attention	267,0633

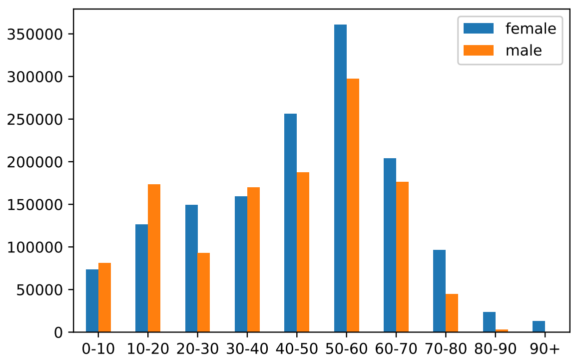

Fig. 1. Age Groups by Gender

3 Data Description

The viewer tuning, presence and attention data I used are provided by TVision, an audience measurement company which currently operates in three major US metropolitan areas: Boston, Chicago and Dallas. The company installs its hardware and trains its software to detect each household member. It enables us to identify which channel a household is watching and whether each detected viewer is paying attention to the TV. The panel data consists of 3397 individuals, 682 households and 830,174 airing advertising over one month (Dec 2016 - Jan 2017). Basic statistics are shown in Table 1.

I also show the top five most popular ads categories with regard to industries and brands, as well as the most popular programs. From Table 2, I can tell that roughly about 13% of the data consists of users viewing NFL football games. The most ads being watched are from TV Networks and Auto Makers, followed by Department Stores.

The above are some statistics about the ads features. In Figure 1 and Figure 9, I also show the user features, such as age, gender and education. The age follows a normal distribution. This dataset has slightly more females than males in general. The users are mostly high school and college graduates or with associate degree.

Among the 830,174 advertisements broadcasting on the channels, I find the average attention (engagement level) for each viewer is around 0.077. This is the ratio of average attention time divided by average advertising duration. While it implies that on average the attention rate of viewers on a particular ad is only about 7.7 percent of the time, the distribution of the attention ratio is very sparse across users. This

Table 2. Popular Industry, Brand and Program

Industry	Count	Percentage
TV Networks	289,185	10.82%
Auto Makers	117,841	4.41%
Department Stores	92,439	3.46%
Movies	65,275	2.44%
Quick Serve	63,423	2.37%
Brands	**Count**	**Percentage**
CBS	3,1453	1.17%
Walmart	2,3105	0.86%
NBC	22,903	0.85%
ABC	22,049	0.82%
Geico	19,856	0.74%
Program	**Count**	**Percentage**
NFL Football	358,265	13.41%
ABC World News Tonight	32,504	1.22%
Law & Order	30,601	1.15%
College Football	27,732	1.04%
The Big Bang Theory	27,096	1.01%

suggests a significant demand for personalized recommendation. Typically, higher attention is from watching more popular programs, the data implies an inter-temporal spillover effect: the previous ad being watched has a significant influence on the attention level of the next ads, suggesting a need to modify or personalize advertising sequences toward a particular user.

4 Exploratory Analyses

My focus is on the factors that drive the magnitude of the attention ratio. I asked some questions relating the the characteristics of Ads and viewers to the attention ratio:

1. Do people give more attention to popular ads in particular industries and brands?
2. Do people of different ages, gender and income level have different tastes in ads?

4.1 Attention Heterogeneity in Ads

Clearly, popular advertising brands, industry and programs in general have higher levels of attention than the sample mean 0.077. The Ads from CBS achieves the highest attention rate 10.56%, while the overall engagement on ads is the highest during people watching NFL Football games.

Table 5 in the Appendix shows more details. The result

Table 3. Attention to the Most Popular Industry, Brand and Program

Ad Title	Attention Ratio	Age Group
Future of Football	0.067	0-10
Star	0.073	10-20
Holiday 2016: The Delivery	0.098	20-30
Family of Products	0.113	30-40
Perfect Pancakes	0.103	40-50
Official Trunk of the NFL	0.115	50-60
Alexa Moments: Mascot Keys	0.125	60+
Ad Title	**Attention Ratio**	**Gender**
Family of Products	0.113	Male
Official Trunk of the NFL	0.118	Female
Ad Title	**Attention Ratio**	**Income Group**
Holiday 2016: The Delivery	0.098	$100K+
Pizza Hut: Singing Snowman	0.072	$75-99k
Show Your Pet You Care	0.125	$60-74k
Match Game	0.093	$50-59k
Chevrolet TV Spot	0.053	$20-49k
T-Mobile TV Spot	0.113	$20k-

shows that the characteristics of ads have a strong influence to the attention ratio, thus I should consider them as predictive features.

4.2 Attention Heterogeneity in Users

Speaking of user differences, as I show in Table 3, different age groups like to watch different ads. For kids under ten years old, the most popular ad is 'Future of Football: Eye in the Sky'. I believe this is because the kids were watching what their parents watched. People in the forties like to watch 'Perfect Pancakes' Featuring Jeremy Rabe, while the fifties like to watch 'Official Truck of the NFL'. The elders, in the nineties, mostly females, favorite local ads. Overall, more popular ads gain higher attention. The most popular ads for female and male are 'Official Truck of the NFL' and 'Family of Products: Testimonials', respectively. The levels of attention are identical, with females slightly more concentrated. In terms of the most popular ads among different income groups, it is interesting that higher income people like to watch ads related to high-tech and holiday.

Table 3 shows more detailed results, indicating that the characteristics of households and users do have an influence to the attention ratio as well. I should also include them as predictive features.

4.3 Temporal Attention

Other than the features of ads and users, the most important relationship I have discovered in the data is the temporal characteristics of attention. I plot the attention level over time, in terms of calendar date, weekdays and daily hours and find some interesting pattern. Followed the exploratory analysis below, I are confident that there is a significant time trending effect on the engagement level of the viewers. Thus, I would like to put it in the predictive model.

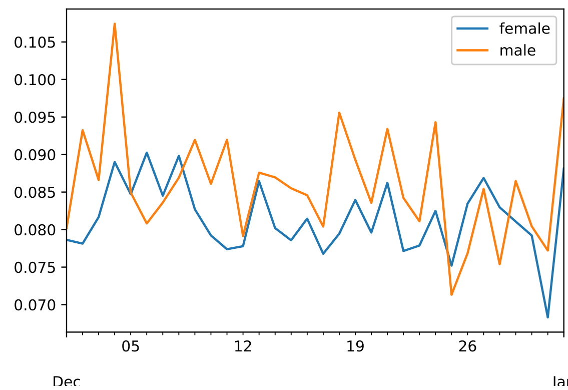

Fig. 2. Average Attention Over Time by Gender

4.3.1 Attention over Calendar Date

Clearly, in Figure 10 (Appendix), there is a seasonality pattern for the engagement level of viewers. I suspect the peaks and valleys stem from weekday effects or maybe from the broadcasting of popular programs, like NFL football games. On average, female viewers attain lower engagement levels than male viewers, but the fluctuation of female engagement level is less than that of males. As shown in Figure 2.

4.3.2 Attention over Weekdays

In Figure 3, the aggregated engagement level increases from Wednesday to Sunday, and reaches the peaks on Saturday and Sunday, then decreases significantly on Monday. This reflects that there are more audiences on weekends than on weekdays. The average engagement level reaches the peaks on Sunday and Tuesday.

Figure 4 shows an interesting pattern that there is the lowest average engagement level on Saturdays. It implies that while there are more audiences on Saturday than Weekdays, their engagement level on the advertisements is actually the lowest. This might because on Saturdays, there are significantly more ads to watch than the other days. The increased amount of advertisements to be watched resulted in distracting the attention of the audiences. The peak on Sunday can be interpreted partly by NFL football games, which is the largest proportion of TV programs in the dataset, and 54% of the NFL broadcasts are on Sundays.

4.3.3 Attention over 24 Hours

The aggregated (Appendix: Figure 11) and average (Figure 5) engagement levels have an identical trend in the twenty four hours time window. It gradually increases from 10 a.m. in the morning and reaches the peak at 10 p.m. to midnight.

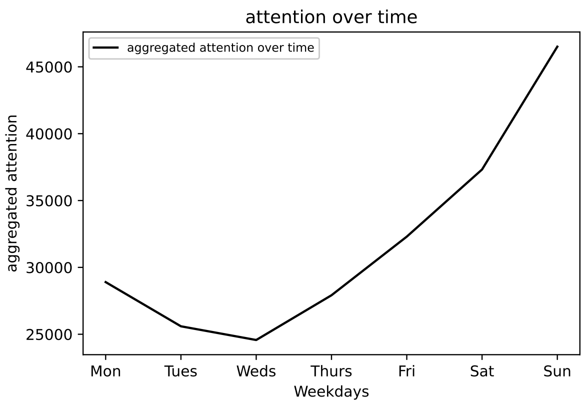

Fig. 3. Aggregated Attention over weekdays

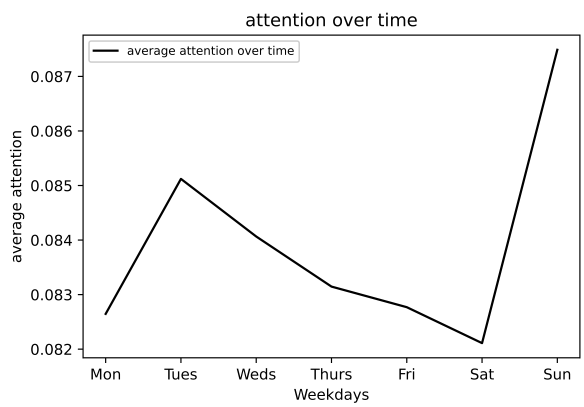

Fig. 4. Average Attention Over Weekdays

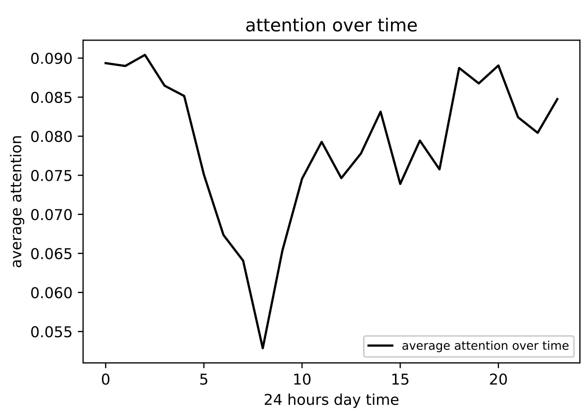

Fig. 5. Average Attention Daily

5 Descriptive Models

After exploring the Ads, viewer and temporal features and how they different across attention ratio, I conduct two regressions and a Spearman ranking correlation test to answer the questions like:

1. Is the previous attention persistent?
2. If yes, how will it influence the next engagement level?
3. Is there a heterogeneous effect of ads across different industries on the engagement level?

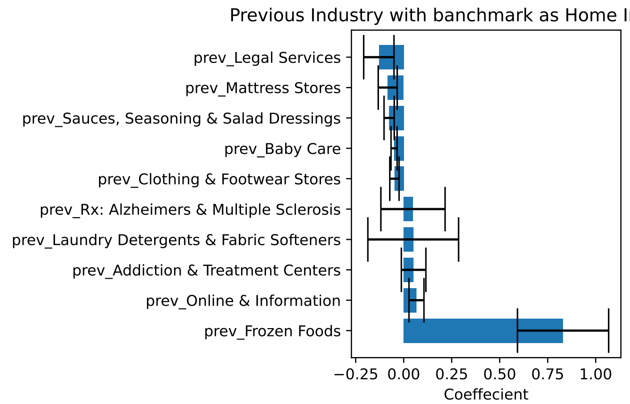

Fig. 6. Previous Industry with benchmark as Home Improvement

4. what is the engagement correlation between previous ads' category and the next ads' category?

5.1 Persistence of Attention

I define the previous attention ration as an average over a previous 10 minutes window. I regress the previous attention ratio $a-1$ on the current attention ratio a and controlled for viewer specific effect (viewer characteristics i):

$$Y_{a,i} = \beta_1 X_{a-1,i} + \varepsilon$$

The result shows that the coefficient is 0.5216 with t-value 387.834, meaning a large and statistically significant positive correlation between the previous engagement level and the current engagement level. This implies that the attention is persistent over a short period of time. See Figure 12 in Appendix.

5.2 Ads Category Affect Attention

I define the previous industry as the last industry of Ad being watched over a previous 10 minutes window. I regress the current industry s and previous industry $s-1$ on the current attention ratio, controlled for viewer specific effect. The industry are dummy variables.

$$Y_{a,i} = \beta_2 X_{s,i} + \beta_3 X_{s-1,i} + \varepsilon$$

I mostly care about the coefficient of previous industries. The top and bottom five industries are shown in Figure 6. The Frozen Foods category has a large, positive and statistically significant effect on the attention ratio for the next Ad, while the Legal Services category has the most negative and statistically significant effect for the next Ad. This shows that the ads categories have temporal influences.

5.3 Temporal Effect of Ads Category

I find the Spearman ranking correlation between

$$\beta_2, \beta_3$$

is 0.6453 with p-value 3.6779e-21. This implies that there is a statistically significant positive correlation between the rankings of the coefficients of previous and current Ads category with respect to the attention ratio. In conclusion, industries that have a large effect on contemporary engagement also have large externalties on the next period engagement.

6 Prediction Tasks

I provide some methods of predicting the attention of audiences via Latent Factor Models. To start, a data sampling technique is used to create the training and test sets. I then learn user preferences, time trends, and the impact of item features on users' attention ratio with Latent Factor Models. These relationships determine the values of the model parameters via training, ensuring the model can predict the level of audiences' attention. When given a datum containing information about users, ads, time, and temporal features, the model can generate an estimated value of attention ratio. The ability of models' predictions are evaluated by computing the MSE metric between the predicted value and the real attention ratio.

6.1 Data Sampling

The training data of a Latent Factor Model is a list of vectors $(u, i, \ldots, r)$. The very first step is to reassign an ID to each user (audience) and item (advertisement). As the original user ID and item ID are out of order in the raw data, reassigning IDs to each of them enables the model to simulate the non-linear relationship with efficiency. The data consist of multiple primitive features about audiences and ads. To improve the precision of the model, besides these basic features, I incorporated the features of time, the household information, which had proved to be important in the exploratory analyses. To be specific, I created new features such as previous ad ID (ad that has been watched), as well as the previous attention ratio. The reason is that I observed a strong persistence in attention ratio. The ID reassigning technique can also be applied to these features because these influences are non-linear as well.

6.2 The Latent Factor Model (LFM)

I first try to solve the problem by adopting collaborative filtering methods. The term collaborative filtering means generating predictions for a user by gathering preferences from many users (collaborating). The Latent Factor Model I discussed is a type of collaborative filtering model. It relates some observable (manifest) variables to a set of latent (unobservable) variables. The model projects users and items to a joint latent factor space where it assumes that users' position represented by these latent factors determines the manifest variables like ratings. When given an input, the latent factors collaboratively filter (predict) the outcome by looking into a similar group of users' preferences. These latent factors enables us to predict the outcome of manifest variables.

$$\hat{R}(u,i) = \alpha + \beta_i + \beta_u + \gamma_i * \gamma_u$$

I used this basic Latent Factor Model as the benchmark. It only takes two variants, which are the ID of users and items, serving . It is selected to serve as the baseline of the recommendation system. as it exploits primitive features of the data to extract latent factors, which helps to predict an estimated score of the test sample. The philosophy behind this model is that the various interactions between user and data item contain some high-order, non-linear information. This information is represented by $\gamma_i * \gamma_u$, which mainly describes the connection between certain users and certain items. The result of test samples, which is the attention ratio, is believed to be affected by this connection in a non-linear way and some other user and item biases. In this case, despite the higher-level relationship between user and items, the result is also determined by the advertisement itself, users' viewing habits, and the average attention ratio.

6.3 Factorizing Personalized Markov Chain (FPMC)

The definition of the Markov Chain is the outcome of a current event depends on the state attained in the previous event. In the field of recommendation systems, the concept is used to describe certain scenarios, where users' interactions are usually affected by their previous one. The Markov Chain model is applicable here as audiences usually watch ads consecutively, and the previous ad can affect their attention to the next ad. A boring previous ad might make users less focused, resulting a poor attention to the next ad, while good ads might make audiences pay more attention to what comes after. I use the Factorizing Personalized Markov Chain Model to combine two styles of modeling. This model adds the previously consumed item as a latent factor, in a similar way as users and items. The predicted rating, therefore, turns out to be a function of u, i, j, denoting user, ad, and previous ad.

$$\hat{R}(u,i,j) = \alpha + \beta_i + \beta_u + \gamma_i * \gamma_u + \gamma_{i'} * \gamma_j$$

The second model considerate not only the previously watched advertisement but also the user preferences. Note that the feature vectors of items(ads) are represented by two different matrices in this formula. The separate representation of i and j, is the key to modeling the Markov process, so two feature matrices are denoted to store them separately. By summing up these modifications together, the model predicts the result based on a comprehensive background. Therefore, the model usually generates more accurate results when compared to LFM.

6.4 FPMC with Time (FPMC-T)

Despite using the classical FPMC model, I seek to incorporate the time features in the model. The argument is based on the observation that the time affects the level of attention for most audiences. In the data, the audiences tend to pay more attention to ads broadcast on Tuesday and Sunday and get more involved in late night hours. Therefore, I propose a

Factorizing Personalized Markov Chain with Time.

$$\hat{R}(u,i) = \alpha + \beta_i + \beta_u + \beta_t + \gamma_i * \gamma_u + \gamma_{i'} * \gamma_j + \gamma_{i''} * \gamma_t$$

The FPMCT (Factorizing Personalized Markov Chain with Time) Model is an upgraded version of the FPMC model. Despite considering users, items. This model adds some features based on the time when the advertisement is broadcast. The assumption is that people's engagement level on the identical advertisement can vary over time. The influence of the broadcast time affects the result in a similar manner as the users' preference does. Therefore, a bias factor and a product of two feature matrices are added to the FPMC formula, turning into the formula of FPMCT. In the exploratory analysis, I have observed a seasonality pattern in the data, thus I expected that including time variate would give us more predictive power.

6.5 FPMC with Previous Ratio (FPMC-R)

Besides the temporal feature of Ads, I also consider to include the temporal feature of viewer's attention.The prediction formula is as follows:

$$\hat{R}(u,i) = \alpha + \beta_i + \beta_u + \beta_t + \gamma_i * \gamma_u + \gamma_{i'} * \gamma_j + \gamma_{u'} * R_j$$

My motivation is to model the effect of the previous persistence of the attention level. I spot multiple records where users pay a high level of attention to consecutive ads played during a certain period of time period. The function, therefore, contains a linear item of $_{u'} * R_j$, the interactions between user features and the attention ratio of the previous commercial ad watched. With this new function, I can capture user specific persistence of attention.

6.6 FPMC with Time and Previous Ratio (FPMC-TR)

Here I proposed a unique model to exploit some linear features of an inter-temporal relationship in the data. The model includes both time and previous attention ratio as the interaction terms. The prediction formula is as follows:

$$\hat{R}(u,i) = \alpha + \beta_i + \beta_u + \beta_t + \gamma_i * \gamma_u + \gamma_{i'} * \gamma_j + \gamma_{i''} * \gamma_t + \gamma_{u'} * R_j$$

7 Model Results

The attention is very sparse as compared to the presence. 70% of the sample have zero attention. To gain a higher predicting power, I only consider advertisements with attention ratio larger than zero, using an item set of size 883859. Using the sampling method described in section 5.1, I create item data samples of size 4200k, which I partition into 60%, 20%, 20% as training, validation and test sets. I select the model that performs best on the validation set and report its performance on the test set.

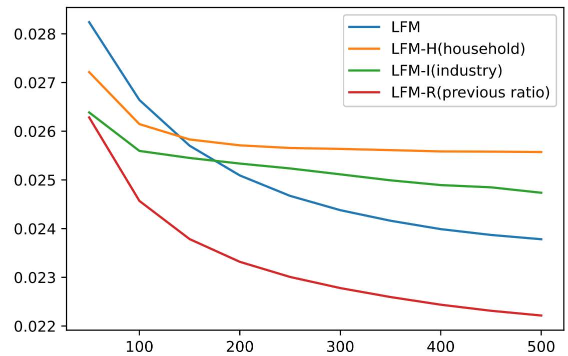

Fig. 7. Objective and Training Steps

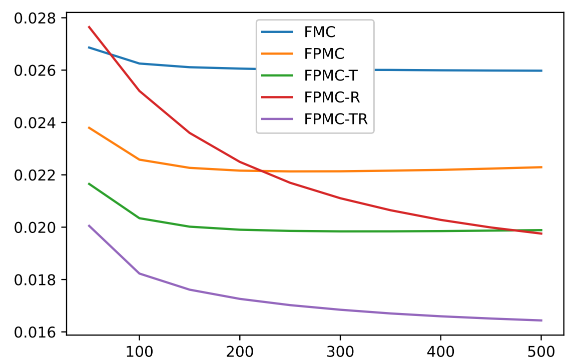

Fig. 8. Objective and Training Steps

7.1 LFM results

As the Figure 7 shows, adding features of household and industry enhanced the ability of predicting attention ratio. Notably, the previous ratio is most predictive to the next attention ratio. The LFM-R model improves it's training performance by 12.5% and the model converges at a faster speed as compared to the LFM model.

I calculate the MSE of each LFMs in Table 4. When compared to the basic LFM, the LFM-R model has some improvement of MSE on the training set. Given other models' performance always suffer a loss on the validation set, the improvement of the LFM-R's performance on the validation set is tremendous.

7.2 FPMC results

Given that the LFM-R performs best among the LFMs. I are confident that the previous interactions enhance the ability of predicting attention ratio. The advertisement watching process is consecutive, where the next are highly dependent on their previous interactions. Thus I adopt the new Markov Chain model with previous Ads and previous attention levels. With information on the previous consumed item, the FPMC converges 3 times faster as compared to the LFM. The FPMC model with time and previous ratios converges even faster, as shown in Figure 8.

Table 4. Model MSE

Model	MSE (training)	MSE(validation)
LFM	0.04596	0.05345
LFM-H	0.05347	0.05371
LFM-I	0.05234	0.05318
LFM-R	0.04348	0.04378
FMC	0.04658	0.06066
FPMC	0.03831	0.05261
FPMC-T	0.03148	0.05294
FPMC-TR	0.0247	0.04161

Table 5. Model MSE on Test

Model	MSE (training)	MSE(Test)
LFM	0.04499	0.05274
FPMC-TR	0.02288	0.04009

In terms of the FPMC-T model, I have tested on two types of time. The weekdays and the calendar date (unix time). I find that the FPMC-T (weekday) has MSE 0.05488 in validation set. It works better than FPMC model, but doesn't perform better than the FPMC-T(unix time). My primitive insight into this result is that the weekday captures the time trend but does not contain a richer information as the unix time does. There is a loss of information during this procedure. So I decide to use the calendar date to capture finer time trend in the model.

In Table 4, The MSE on validation set have gradually improved with addition to the previous Ad, time trend and previous attention. The predicting performance of FPMC-TR have improved by 45.8% as compared to the FMC model, and improved by 28.4% as compared to the LFM.

I finally adjust the learning rate and apply the FPMC-TR model to the test set and compare it with the benchmark LFM model. I find that the MSE of FPMC-TR on test is 0.04009 while the MSE of FMC on the test set is 0.05274, meaning that the FPMC-TR model excels the benchmark by 31.8%, Table 5.

8 Conclusions

The effectiveness of TV Advertising is an important and classical topic in the marketing field. The attention ratio is a finer and better metric in evaluating the effectiveness of Ads. I use the unique data from TVision to study what factors influence the attention ratio of audiences. I find that the characteristics of Ads and viewer preferences will impact the engagement level. Besides, I also find there is a temporal relationship between the previous engagement level and the next engagement level.

Based on those findings, I apply the Latent Factor model and the Factorizing Personalized Markov Chain model to make prediction on attention ratio. In general, adding user specific feature (like household information), and item specific feature (like industry categories) improve model predictions. I find the temporal features help a lot in enhancing the predicting power. The FPMC model outperforms LFM, while the FPMC-T and FPMC-TR works better than FPMC. The FPMC-TR model outperforms the benchmark LFM model by 31.8%, suggesting a strong time effect and intertemporal externalities on attention.

References

[1] Gomes, R., Immorlica, N., and Markakis, E., 2009. "Externalities in keyword auctions: An empirical and theoretical assessment". In International workshop on internet and network economics, Springer, pp. 172–183.

[2] Wilbur, K. C., Xu, L., and Kempe, D., 2013. "Correcting audience externalities in television advertising". *Marketing Science,* **32**(6), pp. 892–912.

[3] Gordon, B. R., Jerath, K., Katona, Z., Narayanan, S., Shin, J., and Wilbur, K. C., 2021. "Inefficiencies in digital advertising markets". *Journal of Marketing,* **85**(1), pp. 7–25.

[4] Haugtvedt, C. P., and Strathman, A. J., 1990. "Situational product relevance and attitude persistence". *ACR North American Advances*.

[5] Joo, M., Liu, W., and Wilbur, K. C., 2020. "Divergent temporal courses for liking versus wanting in response to persuasion.". *Emotion,* **20**(2), p. 261.

[6] Linden, G., Smith, B., and York, J., 2003. "Amazon. com recommendations: Item-to-item collaborative filtering". *IEEE Internet computing,* **7**(1), pp. 76–80.

[7] He, R., and McAuley, J., 2016. "Ups and downs: Modeling the visual evolution of fashion trends with one-class collaborative filtering". In proceedings of the 25th international conference on world wide web, pp. 507–517.

[8] Majumder, B. P., Li, S., Ni, J., and McAuley, J., 2019. "Generating personalized recipes from historical user preferences". *arXiv preprint arXiv:1909.00105*.

Appendix

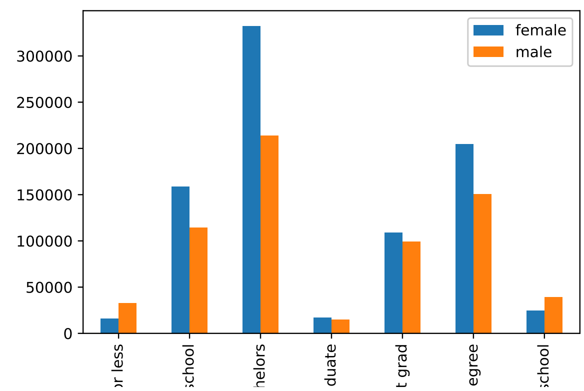

Fig. 9. Education Level by Gender

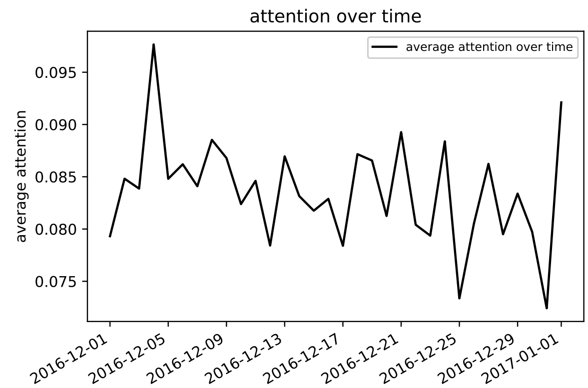

Fig. 10. Average Attention over Timer

Table 6. Attention to the Most Popular Industry, Brand and Program

Industry	Attention Ratio
TV Networks	0.089
Auto Makers	0.086
Department Stores	0.085
Movies	0.085
Quick Serve	0.086
Brands	**Attention Ratio**
CBS	0.105
Walmart	0.086
NBC	0.096
ABC	0.089
Geico	0.082
Program	**Attention Ratio**
NFL Football	0.097
ABC World News Tonight	0.081
Law & Order	0.070
College Football	0.080
The Big Bang Theory	0.085

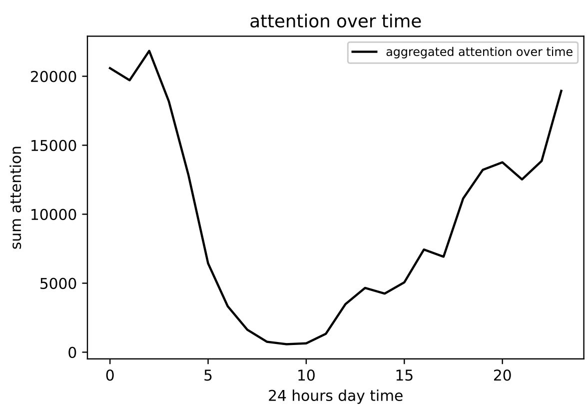

Fig. 11. Aggregated Attention Daily

```
                            OLS Regression Results
==============================================================================
Dep. Variable:               demean_y   R-squared:                       0.270
Model:                            OLS   Adj. R-squared:                  0.270
Method:                 Least Squares   F-statistic:                 1.504e+05
Date:                Sat, 20 Nov 2021   Prob (F-statistic):               0.00
Time:                        12:39:09   Log-Likelihood:                 71976.
No. Observations:              407460   AIC:                        -1.439e+05
Df Residuals:                  407458   BIC:                        -1.439e+05
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept   1.935e-16      0.000   6.09e-13      1.000      -0.001       0.001
demean_x       0.5216      0.001    387.834      0.000       0.519       0.524
==============================================================================
Omnibus:                    53144.712   Durbin-Watson:                   2.108
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            99109.057
Skew:                           0.847   Prob(JB):                         0.00
Kurtosis:                       4.723   Cond. No.                         4.23
==============================================================================
```

Fig. 12. Regression

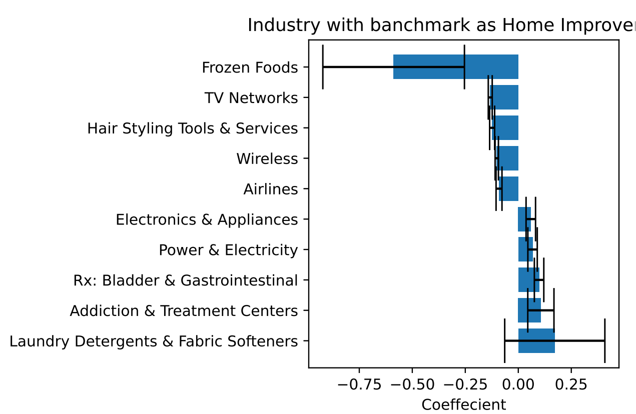

Fig. 13. Education Level by Gender