A Survey on Different Approaches used in Predicting Customer's Purchase Intent

Team: SSHOP_249

Name	**UID**	**Email-ID**
Heenal Doshi	004758927	doshi.heenal@gmail.com
Omkar Patil	904760474	omkar.9194@gmail.com
Pranav Sodhani	804591764	sodhanipranav@gmail.com
Shubham Mittal	104774903	mitshubh@gmail.com
Swati Arora	404758379	swati3124@gmail.com

ABSTRACT: With the advent of e-commerce applications in every sector, a large audience has shifted online to search and purchase products, thus outpacing traditional retailing. As such, it is important to look for consumer traits which might trigger a purchase action. This, in turn, can be put into use in designing personalized product recommendation services which has the potential to boost overall sales.

In this paper, we study various methods adopted in the literature to explore consumer signals which help in determining purchase intent. Our aim is to organize existing methods by categorizing existing approaches in different heads to present a systematic study.

I. INTRODUCTION

A diffusion of internet based e-commerce applications has not only changed consumer lifestyle but has also changed the way businesses are done to a large extent. With a competitive market and a drastically changing buyer-seller relationship, it is critical to understand consumer behaviors which can help in delivering personalized e-commerce services. Even a small change in conversion from a non-purchaser to a purchaser can lead to considerable increase in revenue [1]. At the same time, it is important to differentiate between groups of people with and without purchase intent.

Probably the earliest and most studied research in this direction has been Collaborative Filtering and Content based approaches. Collaborative Filtering tries to predict the purchase intent of target user based on known purchase history of similar users. Thus, is uses experiences and knowledge from a relatively large group of users known to have similar interests/tastes. Content based predictions, in contrast to Collaborative Filtering bases its prediction on information about target user and items only. That is, it doesn't exploit information from similar users like Collaborative Filtering. It is not uncommon to combine Collaborative and Content based filtering to get a better prediction. For example, Amazon uses both of them together to find recommendations for the user.

It has been observed that browsing patterns of users obtained using clickstream data can be used to predict the intention of purchase [10]. Purchasing intent has been well studied in the context of web search engines to identify the search goal of a user given her search history [11]. While most of the work examine user behavior on a single e-commerce platform [1, 5, 7], a few have performed a cross-platform analysis of how purchase intent varies with time [2]. Another direction of research work has involved exploring social media profiles and linguistic information from social media posts to determine the likelihood of purchase and thus developing a recommendation system based on it [3, 4].

Having mentioned various directions of approaches, the common theme we observed across various papers was, each technique has its own set of strengths and weaknesses and many researchers choose to combine techniques in different ways to achieve peak performance.

The work in our paper is organized as follows: In the following section, we try to summarize different approaches, we identified in literature, to predict purchase intent of the target user. We divide the approaches under four main heads; however, this division is only for our systematic study and not a standard followed throughout the literature. The division is according to the dataset used rather than classification models/ algorithms. Section III presents and discusses results obtained while using each of the approaches while Section IV draws a conclusion from all findings and suggests directions for future research.

II. METHODOLOGIES

1. Collaborative Filtering based approaches

Collaborative Filtering is a popular approach, in general used to predict the interest of the user by collecting preference or taste information from many users (collaborating). Recommendation systems identify users similar to the target user based on their purchase history and predict products that they have bought but the target user has not to have relatively higher purchase intent. However, in this approach, it is required that a target user has bought something before. Furthermore, it is difficult to recommend products that customers rarely buy e.g. car.

2. Content based approaches

Content specific to user such as demographics, past purchase history, queries typed by the user, product reviews, etc. can be used to determine new products the target user may be interested in buying. A lot of research has been done in the area of predicting purchase intent of customers by analyzing static patterns such as using attributes of items that a user has purchased in the past [8, 9]. In such works, temporal activity of a user is not monitored to see its effect on her likelihood of purchasing a product.

Although amazon.com's recommendation is considered a good example of collaborative filtering, it also includes recommendations calculated by keeping track of items purchased by the user and also manually set by the user [8]. On an e-commerce application, it is generally easier to fetch profile information about the users. Demographics such as gender, age, geographical information are often useful as a feature to help in distinguishing purchasers from non-purchasers. These features are static and form useful baseline since the information does not change frequently.

These features are generally used in combination with other methods rather than on their own.

3. Clickstream data based approaches

User's interaction with the e-commerce websites - browsing records have proven to be a very useful source in predicting the purchase intent of the user. A considerable amount of work has been dedicated to mining web traversal patterns [16].

Overall, the problem of calculating target user's purchase intent using clickstream has been formulated in various ways across literature. We discuss a few of them below,

A. Markov Models

A visitor's session-level sequence of clicks is viewed as navigation through a set of relevant pages to eventually make a purchasing decision. Markov models have been extensively used to capture temporal relationships between related entities. In the context of online shopping, Markov models were used by Lakshminarayan et al [14] for predicting levels of customer satisfaction. In [12], the authors integrate the pagerank method in a Markov model framework to provide web recommendations while in [13], Moshaber outlines methods to predict the next user-action proposed in the literature. Srivastava et al analyzed web data to explore usage patterns to understand requirements of web applications [15].

In [5], complex browsing pattern of users has been modeled via Markov chain theory to predict user's propensity to make a purchase. The goal is to identify prospects who will convert into purchasers based on their session-level page navigation. Firstly, the dataset is divided into 2 classes – buyers and non-buyers. 70% of the dataset is used for training while the remaining 30% is reserved for testing. The second step involves collecting URL sequences from the training set to create transition probability matrices for both the classes. The URL sequence in a given session is decomposed into pair-wise URLs (i -> j) and the probability pij forms the transition matrix. This is done for both the classes C1 and C2. When a new session from the test data consisting of url sequence U1, U2, U3, .., Un is presented to the model for evaluation, it computes the conditional probabilities for both the classes – P(U1, U2, …, Un | C1) and P(U1, U2, …, Un | C2) using the decomposition property of Markov chains. The class with greater conditional probability is assigned the session.

B. Bayesian Network Models

One disadvantage of using Markov models is that when trying to capture all possible sequences of user selection, the state space may become quite unimaginable. Bayesian network modeling has the potential to capture a consumer's behavior over time periods which can then be used to predict his/her purchase intent [1, 7]. In [7], the authors build a model in which a user's psychological state and preference are treated as 2 different latent variables with time-varying values. In the prediction process, when a new user comes in with a browsing history record, first the psychological states and preferences are determined. Then a ranked recommendation list is made based on the preferences in the previous period. Although the paper aims at making recommendations which converts a user to a buyer, the same model can be used to learn about the purchase intent of the user.

C. Task Completion

Categorization of user activity on an e-commerce application into a set of different tasks and then studying the relation between pattern of task completion and purchase intent has been one common feature selection process followed by authors in [1, 2, 5]. One approach is based on the idea that the prediction of online purchase, at the individual visitor level, can be improved by first grouping a given Web site's activities into a small number of discrete tasks, each of which must be completed for purchase conversion to take place (e.g., selection of product, placement into shopping cart, provision of shipping information, provision of billing information). Thus, rather than predict the buy/no buy outcome directly as a function of covariates, we predict it as the product of a chain of conditional probabilities each of which corresponds to the completion of a task required for purchase. These are manifested as, for example, the probability of product selection given site visit and the probability of shopping cart placement given product selection. Each conditional probability is, in turn, made a function of covariates capturing visitor's within-site browsing behavior, site usage, and idiosyncratic page exposures.

In [1], a task completion approach is used in a Bayesian framework to predict purchase intent of a consumer. Firstly, user activity is broken down into a set of tasks. Completion of each of the tasks is an indication of his/her purchase intent. In this work, it is assumed that a user's preferences might change as s function of his interaction with the website. To capture this, their model incorporates what a user does before completing each of the tasks such as browsing behavior, information search and processing.

The clickstream data can also be combined with static features: authors in [1] use the geographical location of users to build a random effects model in which visitors from the same locality are assumed to respond similarly and thus exhibit similar behavior while [2] uses demographic features – gender, geographic region and till date user activity on the platform.

4. External sources based approaches

A. Using Social Media

There have been efforts in understanding if user's social network profiles (like Facebook, Twitter) can be used to predict what the user would be interested in buying. This is especially useful now when it is not uncommon for users to connect from e-commerce websites to social networking sites. When a user connects her social networking profile to ecommerce websites, she agrees share basic information such as demographics, interests, likes etc. Authors of [17] suggest this information can be leveraged to predict the user's purchase behavior and even to solve the cold start problem. They analyze the correlations between social media profiles and purchase behaviors on e-commerce websites. They show that subset of FB features correlates with purchase behavior on eBay. Their results show a strong correlation between Facebook and Ebay categories thus Facebook categories can be highly predictive of Ebay categories. They formulate the problem as - given a social media user u and set of features derived from her social media account, produce a ranked list of categories that the user is most likely to buy from. They use features such as demographics, Facebook likes, categories and n-grams derived from liked pages.

There is also work in direction when such explicit connection between Facebook accounts and e-commerce accounts is not given. Authors of [18] use a game theoretical model called stable matching algorithm to identify related accounts across heterogenous networks: FB and Ebay. And then combine the social network information with online shopping information to predict behavior.

B. Content Discovery Applications:

Authors of [2] study and model how the purchase intent of a user varies over period of 28 days looking at the user’s activity on a Content discovery Application, Pinterest. It is a first of its kind study to use user’s long term behavior on external application to predict purchase. Firstly they make various observations about the behavior of purchaser as compared to non purchaser, like, purchaser tends to be more focused on specific content and performs more click-through and save actions. Also from the start purchaser is almost 40% more focused on his category of interest than a non-purchaser and so on.

They also observe a very evident change in purchaser’s behavior about 3 days before the day of purchase and this changed behavior goes on increasing all the time till the purchase. The authors use these meaningful insights to create features of the model such as: activity ratio, category interaction score etc. They also include temporal features to encode how the behavior changes across time. Finally they use Logistic Regression to predict purchase intent of the user across the 28 day period in experiment. They observe from their results that 3 days before the purchase day, it can be predicted with medium accuracy and the purchase intent goes on increasing upto the time of purchase.

Although we’ve tried to divide the methodologies under specific heads, in practice, common approach is to use these methodologies in combination, as hybrid approaches.

III. RESULTS

In this section, we list performance of few of the methods discussed in Section III. We emphasize that a straightforward comparison of these methods is not justified given that these operate on different datasets and use different performance metrics.

Figure 1 shows the performance of the markov model suggested in [5] when applied to the notebooks category. Different orders of Markov chains – 1, 2 and 3 were deployed and recall(left) and false positive rate(right) was plotted as a function of partial length of sequence in a session. It was apparent that order 1 markov chain is preferable due to its low complexity and ease of implementation

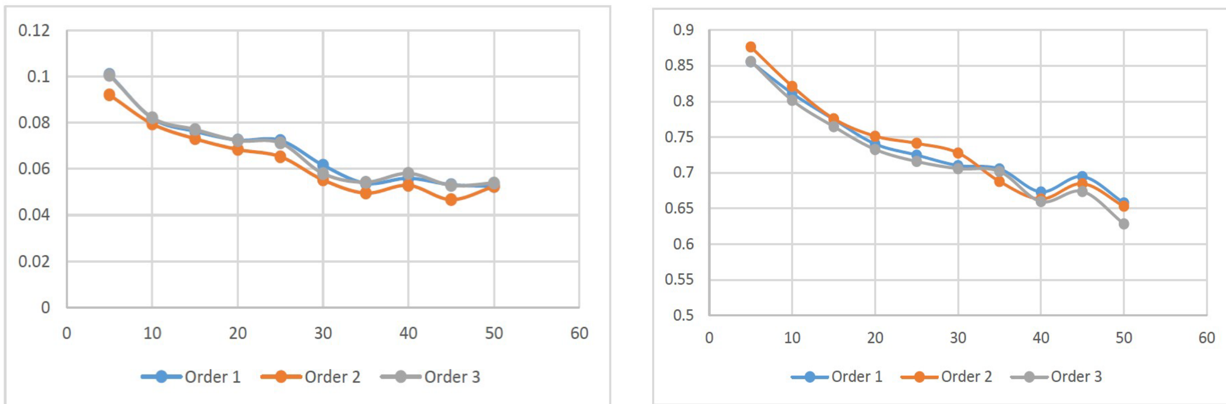

Figure 1: (left) Y-axis: Recall, X-axis: partial length of sequences and (right) Y-axis: False positive rate(FPR), X-axis: partial length of sequences

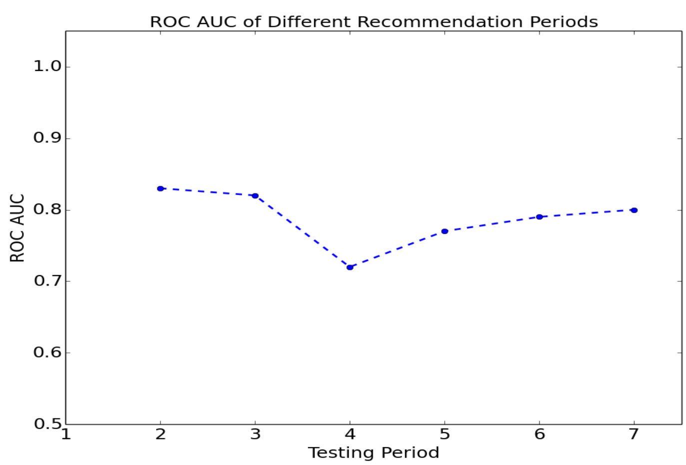

Figure 2: Receiver operating characteristics Area under Curve (ROC-AUC) as s function of the testing period (2-7) for method described in [7]

Figure 2 shows the performance for Bayesian network based Collaborative filtering for multi-period as described in [7]. We see that the ROC AUC is always above 0.7 thus establishing a good classification accuracy. These results confirm that period-related differences in purchaser’s browsing behavior exist which is efficiently captured by the model. Note that the results shown here indicate the performance of recommendation algorithm. As mentioned earlier, the same can be leveraged to predict user purchase intent.

Figure 3(a) shows the activity ratio defined in section III (C) as a function of the number of days left before a purchase is made. We see that as we approach the date of purchase, purchasers outperform non-purchasers in terms of activity.

Figure 3(b) shows the performance of the algorithm discussed in [2]. We see that the ROC AUC is constantly greater than 0.6 implying that a classification with sufficient accuracy can be made even before 28 days of purchase. As the date of purchase is approached, ROC-AUC shows a sudden peak implying that purchasers and non-purchasers show significant differences before a purchase is made.

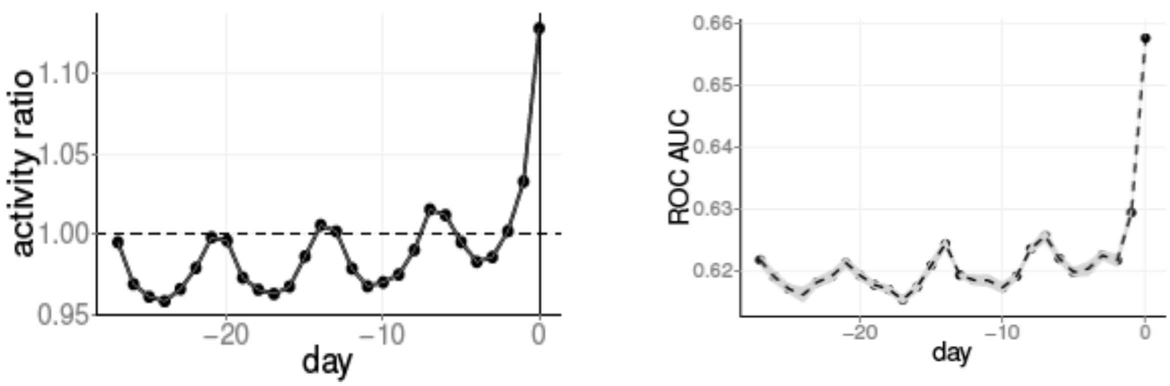

Figure 3: (a) Activity ratio as a function of the days remaining before a purchase is made and (b) Receiver Operating Characteristics Area Under Curve (ROC-AUC) as a function of the number of days left before a purchase

Feature Sets	P_1	P_2	P_3	P_4	P_5	$NDCG_1$	$NDCG_2$	$NDCG_3$	$NDCG_5$	$NDCG_7$
Baseline	0.668	0.547	0.513	0.454	0.451	0.668	0.694	0.709	0.701	0.680
D	0.670	0.593	0.565	0.534	0.504	0.670	0.728	0.735	0.721	0.710
F	0.708	0.652	0.621	0.572	0.549	0.708	0.761	0.765	0.749	0.736
L	0.706	0.647	0.613	0.568	0.538	0.706	0.759	0.761	0.748	0.733
N	0.705	0.636	0.605	0.563	0.533	0.705	0.757	0.760	0.745	0.732
$F + D$	0.715	0.649	0.623	0.575	0.553	0.715	0.766	0.770	0.765	0.753
$F + L$	0.718	**0.657**	0.625	0.576	0.555	0.718	0.770	0.775	0.768	0.755
$F + N$	0.717	0.655	0.623	0.578	0.552	0.717	0.769	0.776	0.766	0.752
$F + D + L$	0.723	0.653	**0.634**	**0.586**	0.559	0.723	0.775	0.782	**0.771**	0.756
$F + D + N$	0.722	**0.657**	0.624	0.577	0.558	0.721	0.773	0.780	0.770	0.758
$F + L + N$	0.729	0.656	0.629	0.581	0.563	0.729	0.780	0.778	0.763	0.750
$F{+}D{+}L{+}N$	**0.733**	0.655	0.628	0.582	**0.565**	**0.733**	**0.784**	**0.785**	0.770	**0.759**

D: Demographics; F: Facebook categories; L: Top Facebook likes; N: Top n-grams(n=1,2,3)

Figure 4: Results of using Logistic Regression with different feature families to predict purchase behavior. NDCG - Normalized discounted cumulative gain.

Figure 4 above shows the results obtained by authors in [17] by matching Facebook and Ebay accounts and using social media features. Overall they observe that using user's Facebook account information considerably improves predictability.

IV. CONCLUSION

From our study of the literature, we conclude that browsing behavior of the user forms a very important signal in predicting the purchase behavior. Information from other sources, especially social media can improve the prediction by a considerable amount, however this seems to be specific to domain under study.

Also the future trend seems to be towards using features from Internet of Things. With the advent of IOT a huge number of signals about the user can be obtained which were not available before. Hence we think that would further improve prediction.

REFERENCES

1. Sismeiro, C., & Bucklin, R. E. (2004). Modeling purchase behavior at an e-commerce web site: A task-completion approach. *Journal of marketing research*, *41*(3), 306-323.

2. Lo, C., Frankowski, D., & Leskovec, J. (2016, August). Understanding Behaviors that Lead to Purchasing: A Case Study of Pinterest. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 531-540). ACM.

3. Ding, X., Liu, T., Duan, J., & Nie, J. Y. (2015, January). Mining User Consumption Intention from Social Media Using Domain Adaptive Convolutional Neural Network. In *AAAI* (Vol. 15, pp. 2389-2395).

4. Gupta, V., Varshney, D., Jhamtani, H., Kedia, D., & Karwa, S. (2014, May). Identifying Purchase Intent from Social Posts. In *ICWSM*.

5. Lakshminarayan, C., Kosuru, R., & Hsu, M. (2016, April). Modeling complex clickstream data by stochastic models: Theory and methods. In *Proceedings of the 25th International Conference Companion on World Wide Web* (pp. 879-884). International World Wide Web Conferences Steering Committee.

6. Borges, J., & Levene, M. (2007). Evaluating variable-length markov chain models for analysis of user web navigation sessions. *IEEE Transactions on Knowledge and Data Engineering*, *19*(4).

7. Jiang, W., Wei, Q., & Chen, G. A Dynamic Bayesian Network based Collaborative Filtering Model for Multi-period Recommendation.

8. Pazzani, M. J., & Billsus, D. (2007). Content-based recommendation systems. In *The adaptive web* (pp. 325-341). Springer Berlin Heidelberg.

9. Basu, C., Hirsh, H., & Cohen, W. (1998, July). Recommendation as classification: Using social and content-based information in recommendation. In *Aaai/iaai* (pp. 714-720).

10. Moe, W. W. (2003). Buying, searching, or browsing: Differentiating between online shoppers using in-store navigational clickstream. *Journal of consumer psychology*, *13*(1-2), 29-39.

11. Dai, H. K., Zhao, L., Nie, Z., Wen, J. R., Wang, L., & Li, Y. (2006, May). Detecting online commercial intention (OCI). In *Proceedings of the 15th international conference on World Wide Web* (pp. 829-837). ACM.

12. Eirinaki, M., Vazirgiannis, M., & Kapogiannis, D. (2005, November). Web path recommendations based on page ranking and markov models. In *Proceedings of the 7th annual ACM international workshop on Web information and data management* (pp. 2-9). ACM.

13. Mobasher, B. (2007). Data mining for web personalization. In *The adaptive web* (pp. 90-135). Springer Berlin Heidelberg.

14. Lakshminarayan, C. K., Benson, A. E., & Yu, Q. (2013). *U.S. Patent No. 8,396,737*. Washington, DC: U.S. Patent and Trademark Office.

15. Srivastava, J., Cooley, R., Deshpande, M., & Tan, P. N. (2000). Web usage mining: Discovery and applications of usage patterns from web data. *Acm Sigkdd Explorations Newsletter*, *1*(2), 12-23.

16. Y. S. Lee, S. J. Yen and C. H. Wang, "Mining browsing and purchasing behaviors of web users," 2010 International Conference on Machine Learning and Cybernetics, Qingdao, 2010, pp. 2647-2652. doi: 10.1109/ICMLC.2010.5580913

17. Zhang, Y., and Pennacchiotti, M. Predicting purchase behaviors from social media. In WWW (2013), 1521–1532.

18. Wang, Y., Li, J., Liu, Q. et al. J Supercomput (2015) 71: 3320. doi:10.1007/s11227-015-1495-8