ICIEICE - 2021

Product Recommendation based on user Intention Analysis

S. Dhaarani[1], Dr. V. Latha Jothi[2]
[1,2]Velalar college of Engineering and Technology, Tamilnadu, India

***Abstract*— The development of internet and e-commerce websites has led to the digitalization of data. Digital data in massive numbers leaves traces of what customers see, what they read, their involvement, and behavior which can be mined for learning experiences. Predictive analytics in data utilization, statistical algorithms and machine-learning techniques were used to identify possible trends, events and behaviors based on historical data. There is a countless variety of research so far with the end goal to predict customer intention of items. Several techniques have been accessible in data mining and supervised learning for predicting customer intention, for example, Recurrent Neural Networks, Session-based algorithms, etc... In this paper, the different strategies that are available and used for forecasting customer prediction have been discussed.**

***Index Terms*— Data Mining, Methodologies, Supervised Learning, Prediction, Recommendation.**

I. INTRODUCTION

When an online store is set up to permit businesses to buy from other businesses, the proceeding is called business-to-business (B2B) online shopping. Online shopping comes under e-business which allows customers to buy stuffs from a trader directly over the Internet via websites. Clients track the products that they get attracted to either by exploring the page of the trader directly or by surfing among different vendors using a shopping search engine, which shows the identical product's obtainability and price range at different e-retailers. A classic online store allows the client to have a glance at the trader’s wide range of products and assistance, view snapshots of the products, along with catalog about the product parameters, quality and prices.

Online stores typically authorize clients to make use of the "search" features to find peculiar models or items. The visitor should have rights to access the Internet and a genuine way of making the payment with an aim of completing a transaction. Physical items will be shipped to the customer by the vendor.

Due to today’s adaptation from exploring physical stores to online purchases, predicting customer behavior in the theme of e-commerce is taking importance. It can increase customer fulfillment and sales, emerging in high rise of conversion rates and competitive advantage, by enabling a more customized shopping process. This study reviews contrasting techniques and papers to anticipate the purchase idea of a customer in an e-commerce website.

Grenze ID: 02.ICIEICE.2021.1.17

II. LITERATURE SURVEY

Here, a concise survey on several noteworthy researchers is explained.

A. Synergies of data mining and multiple attribute decision making

Mohammad Hasan Aghdaieet.al., proposed a novel combined outlook for supplier clustering and cluster evaluation and selection with integrating data mining and MADM methods. Suppliers are grouped using a data mining tool called Two-stage cluster analysis. The SWARA method was implied to weigh the features for cluster analysis. The outcome of SWARA was used as weighted inputs for VIKOR which then ranks the clusters from best to worst. A real case history was picked up to convey the performance and application of the model.

B. Predicting Shopping Behavior with Mixture of RNNs

Arthur Torth et.al., contemplates real web interactions from a US e-commerce branch of Rakuten to speculating three possible outcomes: purchase, abandoned shopping basket and browsing-only. To timely detect the outcomes, hints left behind by customers that are concealed into clickstream data and logged during each session are considered. Clickstream data is handled to investigate with mixtures of high-order Markov Chain Models (MCMs) and mixtures of Recurrent Neural Networks (RNNs) that make use of Long Short-Term Memory (LSTM) architecture. Then each model is compared and contrasted at different lengths and reports on precision, recall and F-measures are made. It is demonstrated that LSTM RNNs generalize better and with less data than high-order Markov chain models. When sequences are trimmed to 75% of their length, a comparatively small feature set predicts purchase with an F-measure of 0.80 and browsing-only with an F-measure of 0.98. Along with distinguishing between browsing-only and cart-interaction sessions, this paper can also specifically discriminate between cart abandonment and purchase sessions.

C. Recurrent Neural Networks with Top-k Gains for Session-based Recommendations

BalázsHidasi et.al., instigated novel ranking loss functions along with an enhanced sampling strategy that has contributed impressive top-k gains for RNNs for session-based recommendations. It is also conceivable that these approaches could also supply similar benefits in the region of Natural Language Processing domain that shares significant similarities to the recommendation domain in connection with machine learning (e.g. ranking, retrieval) and data structure (e.g. sparse large input and output space). The enhanced performance of these losses over alternatives, along with further logics and refinements explained, allows for an entire improvement of up to 35% concerning MRR and Recall, 20% on past session-based RNN results and up to 53% over classical collaborative filtering approaches.

D. Recurrent Recommender Networks

The depiction of heuristic recommender systems is a recognized and well-studied thesis. A mainstream strategy is to review problems of the structure introduced in Netflix contest. Given a series of tuples comprising of users, movies, timestamps and ratings, the plan is to acquire ratings for alternative combinations of the first three attributes (user, movie, time). Performance is then sustained by the deviation of the prediction from the attested rating. Recurrent Recommender Networks (RRN) is put forward to speculate future behavioral trajectories. This is attained by supplying both customers and movies with a Long Short-Term Memory (LSTM) autoregressive model that traps dynamics, besides more traditional low-rank factorization. On multiple real-world datasets, this method functions well.

E. Session-based Item Recommendation in E-Commerce On Short-Term Intents, Reminders, Trends, and Discounts

Dietmar Jannach et.al., came up with a better understanding of features that can make e-commerce suggestions successful in practice. Recommendations including the products that are previously recognized, fast-moving and that are given reduction at present can be very useful. Resting on those factors, an innovative algorithm is designed that associates a neighborhood-based scheme with a deep neural network to speculate the related items for a prescribed shopping zone. The data file is taken from Zalando, a well-known e-commerce site that has an annual ledger of client interactions. The activities were taken into account for nearly 4,00,000 variety of shop items. This work not only shows more effective outcomes in offline trials but also a perceptible raise with respect to business metric of online retailing.

F. Classifying and Recommending Using Gradient Boosted Machines And Vector Space Models

Humphrey Sheil et.al proposed a twin tasking approach of user classification and content ranking in e-commerce done using a non-restrictive dataset. Gradient Boosted Machine (GBM) is a straightforward idea to train as it ceaselessly extends an ensemble of classification and regression trees (CART) to foster judgment on unseen data. New trees are continually added throughout the training to stronger the objective function and for correcting the flaws made by the initial trees. The features are calculated and are saved in LIBSVM (label: value) format which is used by GBM to mould a forest of CART trees. In all the experiments done, GBM performed compatibly well that the outcome attained depends greatly on feature engineering than algorithm improvements.

G. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling

Junyoung Chung et.al., compares different recurrent units in Recurrent Neural Networks (RNNs). Chiefly, the priority is made on more enlightened units that carry out a gating mechanism, namely long short-term memory (LSTM) unit and a newly proposed Gated Recurrent Unit (GRU). The analysis is done on recurrent units on the venture of polyphonic music modeling and speech signal modeling. The inspection disclosed that these ultra-modern recurrent units are in fact one step ahead of more classical recurrent units such as tanh units. Depending on the inspection, it is wind up that by using a stable collection of parameters for all frameworks on some datasets, GRU can outperform LSTM units in the matter of merging in CPU time and in parameter updates and logical reasoning. To understand how better a gated unit helps to learn and to separate the contribution of each component, more in-depth analysis must be implemented in the future.

H. Evaluation of Session-based Recommendation Algorithms

Malte Ludewiget.al., compared a collection of new and computationally tangled algorithms for session-based recommendation with the delicate approaches using a mixture of data files and ranking metrics. This research covers simple heuristics as baseline methods, nearest-neighbor techniques, recurrent neural networks, and factorization-based methods. The chief inputs to all procedures are a training set of past customer sessions, where each session contains a set of consecutively ordered steps of a given type. The experimental analysis shows that recurrent neural network turns out to be comparatively best in case of prediction accuracy.

I. Personalizing Session-based Recommendations with Hierarchical Recurrent Neural Networks

Massimo Quadrana et.al., proposed a smooth and continuous way to customize RNN models with cross-section data transfer and comes out with a hierarchical RNN model that relays end evolves latent hidden states of the RNNs across user sessions. Two datasets were used for conducting experiments. The suggested HRNNs model notably outplays both state-of-the-art session-based RNNs and basic personalization strategies for session-based recommendation. While this seems like a fair-minded approach, this analytical approach does not give over satisfactory outcomes. If the item usage and user features are also included, then the paper could take along a better result than that is achieved now.

J. New Item Consumption Prediction Using Deep Learning

Michael Shekasta et.al., proposed a content-based algorithm named Purchase Intent Session-based (PISA) algorithm is employed to weigh up the purchase intention for cold start session-based scenarios. This perspective hires deep learning techniques for modeling the content in addition to purchase intent guessing. Though content-based approaches go wrong while executing uneven data files, this technique handles such situations. Experiments were conducted showing that associating PISA with baseline in a non-cold start framework additionally enhances performance. The inquiry demonstrates that PISA outplays a notable deep learning baseline when new products are launched.

III. METHODOLOGY

Many research scholars have served their endeavors to foresee the predictions of purchasing a product on an e-commerce website. They have utilized plenty of strategies in the field of computer technology and also in economics intending to pick up a bit of this unstable data and make an extraordinary fortune out of prediction in e-commerce websites. There are several models used to forecast the prediction of customer intention and some were explained in Table 1.

TABLE I. SEVERAL MODELS FOR PREDICTING ONLINE SHOPPERS' PURCHASE INTENTION

Ref	Methods	Dataset	Advantage	Disadvantage
[1]	Synergies between data mining and multiple attribute decision making.	Data from Leading automobile producers in Iran.	The proposed SWARA-VIKOR integrated approach can be regarded as another effective technique for supplier clustering and ranking.	A new synergy is needed to integrate DM with MADM models.
[2]	High-order Markov Chain Models (MCMs) and mixtures of Recurrent Neural Networks (RNNs)	Web interactions from a US e-commerce branch of Rakuten	RNNs generalize better and with less data.	Decision strategies and features are not taken into account.
[3]	Top-k gains for RNNs for session-based recommendations.	RSC15, VIDEO, VIDXL and CLASS.	A class of loss functions together with sampling strategy have provided impressive top-k gains for session-based recommendations.	A high capacity of data is required. It consumes huge time and reliability is dependent on the details provided.
[4]	Recurrent Recommender Networks	Netflix contest	Offers excellent prediction accuracy besides being very compact.	For temporal and causal aspects, these approaches are lacking.
[5]	Neighbourhood-based scheme with a deep neural network	Zalando, a well-known e-commerce site	Reminders led to a measurable increase on the subject of business metric on the e-commerce site.	Better methods are still needed to automatically assess the user's current intent and motivation to visit the site.
[6]	Gradient Boosted Machine (GBM)	Open dataset from the ACM RecSys 2015.	GBM functioned consistently well as a robust classifier.	More detailed investigation into the word embedding similarity features is not carried out effectively.
[7]	Gated Recurrent Unit	Three polyphonic music datasets along with two internal datasets provided by Ubisoft	GRU performs well both in merging CPU time, parameter updates and logical reasoning.	More in-depth analysis is needed.
[8]	Simple heuristics as baseline methods, nearest-neighbour techniques, recurrent neural networks, factorization-based methods.	A training set of past customer sessions.	Prediction accuracy is much effective in simpler methods.	Additional research is required concerning to the development of sophisticated models.
[9]	Hierarchical recurrent neural networks	XING RecSys Challenge 2016 and dataset from a YouTube-like video-on-demand web site	Resultsshow large improvements in the session-only RNNs.	Study on usage of items and user features would have improved session-based recommendation even further.
[10]	Purchase Intent Session-based (PISA) algorithm	Dominant e-commerce site.	This technique is highly effective when associated with standard user-item recommendation systems.	Less effective in case of small datasets.

IV. CONCLUSIONS

This survey paper concludes that prediction in purchase intention might be maximized in the course of contrasting procedures and techniques, but every method has its own advantages and limitations. A wide number of methods have been discussed for purchase intent of shoppers' and it is possible to produce a new and cross method to project the behavior prediction or to anticipate the economic condition of a company. But it's also essential to design as per the framework by which the network can be maximized with the exactness and the performance with less computational complication. Other factors and techniques may be considered for precision and efficiency to foretell the intention of the customers.

REFERENCES

[1] Aghdaie, Mohammad Hasan, Sarfaraz HashemkhaniZolfani, and EdmundasKazimierasZavadskas (2014), “Synergies of data mining and multiple attribute decision making.” Procedia-Social and Behavioral Sciences 110: 767-776.

[2] Arthur Toth, Louis Tan, Giuseppe Di Fabbrizio, and Ankur Datta (2017), “Predicting Shopping Behavior with Mixture of RNNs”, In ACM SIGIR Forum.

[3] BalazsHidasi and Alexandros Karatzoglou (2017), “Recurrent neural networks with top-k gains for session-based recommendations.” arXiv preprint arXiv:1706.03847.

[4] Chao-Yuan Wu, Amr Ahmed, Alex Beutel, Alexander J. Smola, and How Jing (2017) “Recurrent Recommender Networks”. in Proceedings of the Tenth ACM International Conference on Web Search and Data Mining (WSDM ’17). ACM, New York, NY, USA, 495–503.

[5] Dietmar Jannach, Malte Ludewig, and Lukas Lerche (2017) “Session-based item recommendation in e-commerce: on short-term intents, reminders, trends and discounts.” User Modeling and User-Adapted Interaction 27, 3-5, 351–392.

[6] Humphrey Sheil and Omer Rana (2017), “Classifying and Recommending Using Gradient Boosted Machines and Vector Space Models”. In Advances in Computational Intelligence Systems. UKCI 2017., Zhang Q Chao F., Schockaert S. (Ed.),

[7] Junyoung Chung, CaglarGulcehre, KyungHyun Cho, and YoshuaBengio (2014), “Empirical evaluation of gated recurrent neural networks on sequence modeling.” arXiv preprint arXiv:1412.3555.

[8] Malte Ludewig and Dietmar Jannach (2018), “Evaluation of Session-based Recommendation Algorithms”, arXiv preprint arXiv:1803.09587.

[9] Massimo Quadrana, Alexandros Karatzoglou, BalázsHidasi, Paolo Cremonesi (2015), “Personalizing Session-based recommendations with recurrent neural networks.” arXiv preprint arXiv:1511.06939.

[10] Michael Shekasta, et al., (2019), “New Item Consumption Prediction Using Deep Learning.” arXiv preprint arXiv:1905.01686.