

Research on Precision Advertising Push Method based on Neural Network

Sheng Qi, Yanbin Long

Liaoning University of Science and Technology, China

Abstract

This project relates to an accurate advertising push method based on neural network, which belongs to the field of advertising technology. The method comprises the following steps: obtaining the sample vectors of each sample of potential users corresponding to the target advertisement; The sample vector is used to train the depth Q network, and the target depth Q network corresponding to the trained target advertisement is obtained; Obtain the target vector of each potential user corresponding to the target advertisement; According to the target vector and target depth Q network of potential users corresponding to the target advertisement, the target potential users to be pushed by the target advertisement at the next push time are obtained. This project can reduce the advertising expenses of enterprises, and also make the effect of enterprise advertising push better.

Keywords

Neural Network; Advertising; Precise Push.

1. Preface

At present, online advertising has become an important way of marketing and has attracted more and more enterprises' attention. Improper advertising investment will lead to waste of advertising expenses and decline of economic benefits of enterprises, so accurate advertising push is crucial; Accurate advertisement push mainly refers to targeted push of advertisements that users are interested in or useful.

The existing method of accurate advertising push is generally to conduct frequent immersion push for key users with potential purchase intention according to certain rules, and will not stop pushing until the marketing purpose is achieved or there is no effect for a long time. This method not only has poor advertising push effect, but also causes waste of advertising expenses of enterprises. In addition, it may also cause boredom of push users or make them feel that their privacy has been violated, so the effect of this advertising push is poor.

2. Project Scheme

This project provides an accurate advertising push method based on neural network, which is used to solve the problem of poor advertising push effect. The technical solution adopted is as follows: an implementation example of this project provides an accurate advertising push method based on neural network, which includes the following steps: obtaining the sample vectors of each sample of potential users corresponding to the target advertisement; The sample vector is used to train the depth Q network, and the target depth Q network corresponding to the trained target advertisement is obtained; Obtain the target vector of each potential user corresponding to the target advertisement; According to the target vector and target depth Q network of potential users corresponding to the target advertisement, the target potential users to be pushed by the target advertisement at the next push time are obtained.

Using the collaborative filtering algorithm, the potential users corresponding to the target advertisement can be obtained.

The method to obtain the sample vectors of each sample of potential users corresponding to the target advertisement includes: obtaining the sample historical time period; According to the preset first time length, the sample history time period is divided to obtain each sample history characteristic time period; Divide the sample history characteristic time period according to the preset second time length to obtain each sample history sub time period; Obtain the type tags of all the commodities that the potential users of each sample have browsed, clicked and purchased in each sample history sub period; According to the type tags of all commodities browsed, clicked and purchased by the potential users of each sample in each sample history sub period, the sample vector corresponding to the potential users of each sample in each sample history sub period is constructed; Any sample vector corresponding to any sample potential user corresponds to a sample history sub period.

The training process of target depth Q network includes: randomly selecting a sample vector and recording it as the first sample vector; Input the first sample vector into the depth Q network, and the agent corresponding to the depth Q network observes the state S_{1t} and executes the action a_{1t} to determine whether the action a_{1t} is the target advertising push for the sample potential users corresponding to the first sample vector. If so, the reward value r_{1t} and the new state S_{1t+1} will be obtained; According to the state S_{1t} , execution action a_{1t} , new state S_{1t+1} and r_{1t} , the transition vector T_1 is constructed, and the network updates the weight according to the transition vector T_1 to obtain the updated weight W_1 ; Then randomly select a sample vector and record it as the second sample vector; Input the second sample vector into the depth Q network with the weight of W_1 , and the agent observes the state S_{2t} and executes the action a_{2t} to determine whether the action a_{2t} is the target advertising push for the sample potential users corresponding to the second sample vector. If so, the reward value r_{2t} and the new state S_{2t+1} will be obtained; According to the state S_{2t} , execution action a_{2t} , new state S_{2t+1} and r_{2t} , the transfer vector T_2 is constructed; The network updates the weight according to the transfer vector T_2 to get the updated weight W_2 ; By analogy, stop network training until the training time reaches the preset training duration, and get the trained target depth Q network; The transfer vector is stored in the experience playback pool; Any sample vector corresponds to a sample potential user, a sample history sub period and a sample history feature period.

The methods to obtain reward values include: for the first sample vector: obtain the behavior states of the potential users of the sample corresponding to the first sample vector after the target advertisement is pushed. The behavior state types include the user's browsing behavior, clicking behavior, collection behavior, purchase attempt behavior, purchase behavior, evaluation behavior and return and exchange behavior after purchase; Judge whether the sample potential users corresponding to the first sample vector have click behavior. If there is no click behavior, obtain the total length of stay of the target advertisement on the central axis of the display screen of the sample potential users; The product of the total duration of stay and the corresponding weight is recorded as the first eigenvalue; Obtain the product of the advertising cost of a single push of the target advertisement and the corresponding weight, and record it as the second eigenvalue; The difference between the first characteristic value and the second characteristic value is recorded as the reward value after target advertisement push for the sample potential users corresponding to the first sample vector; If it is judged that the sample potential user corresponding to the first sample vector has clicked, the total time of browsing the target advertisement by the sample potential user after clicking and the total duration of the signal received by the input device during browsing the target advertisement are obtained; The ratio of the total duration to the total duration of browsing the target advertisement is recorded as the activity of the potential users of the sample in the sample

history sub period corresponding to the first sample vector;According to the behavior states of the sample potential users after the click behavior, the corresponding influence values of the behavior states of the sample potential users after the click behavior are obtained;The product of the total length of time that the sample potential users browse the target advertisement after clicking and the corresponding weight is recorded as the third characteristic value;The cumulative sum of the influence value corresponding to each behavior state multiplied by the corresponding weight is recorded as the fourth characteristic value;The product of the activity of the potential users of the sample and the corresponding weight in the sample history sub period corresponding to the first sample vector is recorded as the fifth eigenvalue;The sum of the third eigenvalue, the fourth eigenvalue and the fifth eigenvalue is recorded as the sixth eigenvalue;The difference between the sixth characteristic value and the second characteristic value is recorded as the reward value after target advertisement push for the sample potential users corresponding to the first sample vector.

According to the behavior states of the sample potential user after the click behavior occurs, the method to obtain the impact value corresponding to the behavior states of the sample potential user after the click behavior occurs, including: if the sample potential user has a favorite behavior after the click behavior occurs, the impact value corresponding to the favorite behavior is recorded as 1;If the potential users of this sample have attempted to purchase after clicking, the impact value corresponding to the attempted purchase behavior will be recorded as 1;If the potential users of this sample have purchase behavior after clicking behavior, the impact value corresponding to the purchase behavior will be recorded as 1;If the potential users of the sample have evaluation behavior after purchasing behavior, the corresponding evaluation grade value will be recorded as the impact value corresponding to the evaluation behavior;If the potential users of this sample have returned or replaced goods after purchasing, the impact value corresponding to the return or replacement behavior will be recorded as - 1.

The method of constructing experience playback pool includes: constructing holiday type experience playback pool, week type experience playback pool, hour type experience playback pool;The number of experience playback pools of holiday type is the same as the number of elements in holiday date series;The holiday date series is constructed by sorting the holiday dates corresponding to the sample historical time period in chronological order;The number of experience playback pools of week type is 7;The number of experience playback pools of hour type is the number of tag value types in the tag value sequence.The process of obtaining the tag value sequence is: sorting each sample history sub time period in each sample history characteristic time period in chronological order to obtain the sample history sub time period sequence corresponding to each sample history characteristic time period;The number of sample history sub time periods in each sample history sub time period series is the same;Mark each element in the sub time period series of each sample history, and construct the tag value series;The marking rule is: mark the first sample history sub period in each sample history sub period series as 1, mark the second sample history sub period in each sample history sub period series as 2, and so on, complete the marking, and build a tag value series according to the tag value.

The process of storing the transfer vector in the experience playback pool includes: for the transfer vector corresponding to any sample history sub time segment, construct the encoding vector corresponding to the transfer vector according to the sample history sub time segment corresponding to the transfer vector;The coding vector is three-dimensional, including holiday dimension, week dimension and hour dimension;When the agent completes an update of the neural network weight, it stores the transfer vector into the corresponding holiday type experience playback pool, week type experience playback pool and hour type experience playback pool according to the coding vector of the transfer vector;One transfer vector corresponds to three experience playback pools, including one holiday type experience

playback pool, one week type experience playback pool and one hour type experience playback pool.

The method for obtaining the encoding vector corresponding to the transfer vector includes: judging whether the date of the sample history characteristic time segment corresponding to the sample history sub time segment corresponding to the transfer vector is a holiday, if so, the order of the date corresponding to the sample history sub time segment in the holiday day period series is recorded as the value corresponding to the holiday dimension in the encoding vector corresponding to the transfer vector; Otherwise, the value corresponding to the holiday dimension in the code vector corresponding to the transfer vector will be recorded as 0; According to the date of the sample history characteristic time segment corresponding to the sample history sub time segment corresponding to the transfer vector, the week value corresponding to the sample history characteristic time segment corresponding to the sample history sub time segment is obtained, and the week value is recorded as the value corresponding to the week dimension in the coding vector corresponding to the transfer vector; Mark the tag value corresponding to the sample history sub time period corresponding to the transfer vector as the value corresponding to the hour dimension in the coding vector corresponding to the transfer vector.

3. Steps of Accurate Advertising Push Method

This embodiment provides an accurate advertising push method based on neural network, which is described in detail as follows:

This accurate advertising push method based on neural network includes the following steps:

Step S001, obtain the sample vectors of each sample of potential users corresponding to the target advertisement.

Because the user's identity and preference for the products advertised are perceptual, abstract, difficult to quantify, and highly random, a person's needs or preferences are different at different times of the week, or even the day, and the products advertised have their own suitable push time; The advertising push of refers to replacing the specified information module with advertising. This embodiment does not consider the specific means of advertising push; For example, advertising of tourist attractions before the May Day holiday can bring more economic benefits than advertising of tourist attractions after the holiday. This embodiment mainly analyzes the potential users corresponding to the advertisement. This embodiment does not analyze the preference of a single user as the starting point, but seeks to find when the currently analyzed commodity advertisement will recommend to the potential users. The probability of successful sale is greater, that is, it is more likely to bring benefits to select those potential users to push at a certain time. Therefore, this implementation uses the target depth Q network to analyze the target vector of each potential user corresponding to the target advertisement, and obtains each target potential user to push the target advertisement at the next push time; This method can not only reduce the cost of enterprise advertising, but also make the effect of enterprise advertising push better.

This embodiment only analyzes the advertisement corresponding to one commodity and records it as target advertisement; This embodiment only analyzes the target advertisement on one platform, which is recorded as the target platform; In general, the target advertisement is a card type information module; This embodiment only analyzes the potential users corresponding to the target advertisement, that is, the potential users corresponding to the target advertisement are obtained by using the collaborative filtering algorithm, and there is a certain relationship between the potential users and the target advertisement. For example, user y1 purchased the goods corresponding to the target advertisement, while user y1 also purchased commodity x1, while user y2 did not purchase the goods corresponding to the target

advertisement, However, if user y_2 purchases commodity x_1 , user y_2 is the potential user corresponding to the target advertisement. Since online advertisements are recorded by the server from pushing to user browsing, user clicks, user purchases, etc., the type tags of all commodities browsed, clicked, and purchased by potential users in the historical period can be obtained from the database; Since this embodiment mainly uses the deep Q network for analysis, and the network needs to train the sample set, so the sample history time period is obtained. This embodiment sets the sample history time period as 30 days; According to the preset first time length, the sample history time period is divided to obtain each sample history characteristic time period; Set the preset first time length to 1 day; Divide the sample history characteristic time period according to the preset second time length to obtain each sample history sub time period; Set the preset value of the second time length as 1 hour, that is, the time length of the sample history sub period is 1 hour; Obtain the type tags of all the commodities that the potential users of each sample have browsed, clicked and purchased in each sample history sub period; According to the type tags of all commodities browsed, clicked and purchased by potential users of each sample in each sample history sub period, the corresponding sample vectors of potential users of each sample are constructed; That is, any sample vector corresponding to any sample potential user corresponds to a sample history sub period.

Step S002, use the sample vector to train the depth Q network, and obtain the target depth Q network corresponding to the trained target advertisement.

In this embodiment, the sample vector is used to train the depth Q network, and the specific process of obtaining the target depth Q network corresponding to the trained target advertisement is as follows:

The structure of deep Q network includes input layer, convolution layer, full connection layer and output layer; The purpose of deep Q network training is to maximize cumulative rewards; Randomly select a sample vector and input it into the depth Q network, record it as the first sample vector, then observe the state S_{1t} and execute the action a_{1t} , judge whether the action a_{1t} is the target advertising push for the sample potential users corresponding to the first sample vector, if so, the reward value r_{1t} and the new state S_{1t+1} will be obtained; According to the state S_{1t} , execution action a_{1t} , new state S_{1t+1} and r_{1t} , the transition vector T_1 is constructed, and the network updates the weight according to the transition vector T_1 to obtain the updated weight W_1 . Then randomly select a sample vector and input it into the depth Q network with the weight of W_1 , record it as the second sample vector, observe the state S_{2t} and execute the action a_{2t} , judge whether the action a_{2t} is the target advertising push for the sample potential users corresponding to the second sample vector, if so, get the reward value r_{2t} and the new state S_{2t+1} , according to the state S_{2t} , execute the action a_{2t} . The new states S_{2t+1} and r_{2t} are constructed to obtain the transfer vector T_2 , and the network updates the weight according to the transfer vector T_2 to obtain the updated weight W_2 ; By analogy, stop the network training until the training time reaches the preset training duration, and get the trained target depth Q network, and the trained target depth Q network can get more accurate action strategies. The preset training duration needs to be set according to the training data set, for example, it can be set to 30 days; Any sample vector corresponds to a sample potential user, a sample history sub period and a sample history feature period.

4. The Specific Process of Calculating the Reward in the Embodiment

The specific process of calculating the reward is as follows: for the first sample vector: the specific process of calculating the reward value after the target advertising push for the sample potential users corresponding to the first sample vector is as follows: obtaining the behavior states of the sample potential users after the target advertising push, where the behavior state

types include the user's browsing behavior, clicking behavior, collection behavior, purchase attempt behavior, purchase behavior, evaluation behavior and return behavior after purchase; In general, after the advertisement is pushed to the user, the user will click to collect, try to purchase or purchase, and then evaluate or return the goods after purchase. Therefore, when the target advertisement is pushed to the sample potential users corresponding to the first sample vector, first judge whether the sample potential users have click behavior. If there is no click behavior, then obtain the total duration of the target advertisement on the central axis of the display screen of the sample potential users, The central axis of the display screen refers to a straight line passing through the central point of the display screen and perpendicular to the screen length and width; Calculate the product of the total duration of stay and the corresponding weight, and record it as the first eigenvalue; Obtain the product of the advertising cost of a single push of the target advertisement and the corresponding weight, and record it as the second eigenvalue; The difference between the first characteristic value and the second characteristic value is recorded as the reward value after target advertisement push for the sample potential users corresponding to the first sample vector.

If it is judged that the sample potential user has clicked, the total duration of browsing the target advertisement after clicking and the total duration of the signal received by the input device during browsing the target advertisement will be obtained. The total duration of the signal received by the input device can reflect the clicks, typing and other behaviors; The ratio of the total duration to the total duration of browsing the target advertisement is recorded as the activity of the potential users of the sample in the sample history sub period corresponding to the first sample vector; Obtain the behavior states of the sample potential users after they have clicked after the target advertisement is pushed to the sample potential users; According to the behavior states of the sample potential users after the click behavior, the corresponding influence values of the behavior states of the sample potential users after the click behavior are obtained; The product of the total length of time that the sample potential users browse the target advertisement after clicking and the corresponding weight is recorded as the third characteristic value; The cumulative sum of the influence value corresponding to each behavior state multiplied by the corresponding weight is recorded as the fourth characteristic value; The product of the activity of the potential users of the sample and the corresponding weight in the sample history sub period corresponding to the first sample vector is recorded as the fifth eigenvalue; The sum of the third eigenvalue, the fourth eigenvalue and the fifth eigenvalue is recorded as the sixth eigenvalue; The difference between the sixth characteristic value and the second characteristic value is recorded as the reward value after target advertisement push for the sample potential users corresponding to the first sample vector.

The specific process of obtaining the influence value corresponding to each behavior state is as follows:

If the potential user of this sample has a favorite behavior after clicking, the influence value corresponding to the favorite behavior will be recorded as 1; If the potential users of this sample do not have the favorite behavior after clicking, the influence value corresponding to the favorite behavior will be recorded as 0. If the potential users of this sample have attempted to purchase after clicking, the impact value corresponding to the attempted purchase behavior will be recorded as 1; If the potential users of this sample do not attempt to purchase after clicking, the impact value corresponding to the attempt to purchase will be recorded as 0. If the potential users of this sample have purchase behavior after clicking behavior, the impact value corresponding to the purchase behavior will be recorded as 1; If the potential users of the sample have no evaluation behavior after the purchase behavior, the impact value corresponding to the evaluation behavior is recorded as 0; If the potential users of this sample have evaluation behavior after purchasing behavior, the corresponding evaluation grade value will be recorded as the impact value corresponding to the evaluation behavior. The higher the

evaluation grade value, the better the user's evaluation. If the potential users of this sample have returned or replaced goods after purchasing, the corresponding impact value of the return or replacement behavior will be recorded as - 1; If the potential user of this sample does not return or replace the goods after the purchase, the impact value corresponding to the return or replacement will be recorded as 0.

The above transfer vectors are stored in the experience playback pool, including three types of experience playback pools, namely, holiday type experience playback pool, week type experience playback pool, and hour type experience playback pool; Obtain the holiday dates corresponding to the sample historical time period, and construct the holiday date series in chronological order; The number of holiday type experience playback pools is the number of elements in the holiday date series. One element in the holiday date series corresponds to one holiday type experience playback pool. The value range of the number of holiday type experience playback pools is [0,30]; The number of experience playback pools of week type is 7. One week value corresponds to one week type experience playback pool, for example, Monday corresponds to one week type experience playback pool; The number of hour type experience playback pools is the number of elements in the tag value sequence. One tag value corresponds to one hour type experience playback pool. The number of hour type experience playback pools is 24.

The process of obtaining the tag value sequence is: sorting each sample history sub time period in each sample history characteristic time period in chronological order to obtain the sample history sub time period sequence corresponding to each sample history characteristic time period; The number of sample history sub time periods in each sample history sub time period series is the same; Mark each element in the sub time period series of each sample history, and construct the tag value series; The marking rule is: mark the first sample history sub time segment in each sample history sub time segment series as 1, mark the second sample history sub time segment in each sample history sub time segment series as 2, and so on, complete the marking, and construct the tag value sequence $A1=\{1,2,\dots\}$ according to the tag value.

Then, the coding vector of the sample vector corresponding to each sample history sub period is constructed. The coding vector is three-dimensional, including holiday dimension, week dimension and hour dimension; The specific construction process is as follows:

For any sample history sub time segment, first judge whether the date of the sample history characteristic time segment corresponding to the sample history sub time segment is a holiday. If yes, record the order of the date corresponding to the sample history sub time segment in the holiday date series as the value corresponding to the holiday dimension in the code vector corresponding to the sample history sub time segment; Otherwise, the value corresponding to the holiday dimension in the code vector corresponding to the sample history sub period will be recorded as 0, and the date refers to the day in the year, month and day. Then, according to the date of the sample history characteristic time segment corresponding to the sample history sub time segment, the week value corresponding to the sample history characteristic time segment corresponding to the sample history sub time segment is obtained, and the week value is recorded as the value corresponding to the week dimension in the coding vector corresponding to the sample history sub time segment; For example, if the value of the week is Wednesday, the value of the week dimension in the coding vector corresponding to the sub time period of the sample history is Wednesday. Then the tag value corresponding to the sample history sub period is recorded as the value corresponding to the hour dimension in the coding vector corresponding to the sample history sub period.

Therefore, the coding vector corresponding to each sample history sub time segment can be obtained through the above process. Since each sample history sub time segment corresponds to a transfer vector, the coding vector corresponding to each sample history sub time segment is recorded as the coding vector of the transfer vector corresponding to each sample history

sub time segment; That is, when the agent completes an update of the neural network weight, it will simultaneously store the transfer vector into the experience playback pool of the corresponding holiday type, the experience playback pool of the corresponding week type, and the experience playback pool of the corresponding hour type according to the encoding vector of the transfer vector, and one transfer vector corresponds to three experience playback pools; That is, one transfer vector corresponds to one holiday type experience playback pool, one week type experience playback pool, and one hour type experience playback pool; For example, if the encoding vector corresponding to the transfer vector is $F=[i, j, k]$, the transfer vector will be stored in three experience playback pools at the same time, where i is the experience playback pool corresponding to the element order i in the holiday date series, j is the experience playback pool corresponding to the week value j , and k is the experience playback pool corresponding to the tag value k .

Because the experience playback pool in the deep Q network calls the transfer vectors of all time periods, it consumes computing resources. In addition, considering that the purchase desire of users in the same time period every day or every weekend will have some connection, for example, people have more time and precision to select goods on weekends, Therefore, this embodiment adaptively selects a multi-dimensional experience playback pool according to the time period, generates spatio-temporal dependency, and adds a priori condition to the training of the neural network.

The loss function of the depth Q network is the loss function of the TD algorithm, and the random gradient descent algorithm is used as the optimizer. In order to reduce the previous training time, the "behavior cloning" method is used to initialize the parameters of the neural network during training, and the cloning object is the behavior track of the current mainstream advertising recommendation algorithm.

Step S003, obtain the target vector of each potential user corresponding to the target advertisement; According to the target vector and target depth Q network of potential users corresponding to the target advertisement, the target potential users to be pushed by the target advertisement at the next push time are obtained.

Obtain the target vector of each potential user corresponding to the target advertisement according to the above method of obtaining the sample vector; Input the target vector of each potential intended user corresponding to the target advertisement into the trained target depth Q network to obtain the corresponding push probability value of each potential intended user; Each potential user whose push probability value is greater than the preset threshold is recorded as the target potential user, that is, the target advertisement will push each target potential user at the next push time. The preset threshold should be set according to the actual situation.

5. Innovation

In this project, the sample vector is used to train the depth Q network, and the target depth Q network corresponding to the trained target advertisement is obtained; According to the target vector and target depth Q network of potential users corresponding to the target advertisement, the target potential users to be pushed by the target advertisement at the next push time are obtained. This project can reduce the advertising expenses of enterprises, and also make the effect of enterprise advertising push better.

References

- [1] Influenza big data analysis based on neural network [J] Lu Min; Duan Jinbin; Ge Xianshun, Chinese Journal of Medical Library and Information, 2020 (03).

- [2] Dialogue system memory mechanism based on recurrent neural network [J]Yi Wei;He Jia;Zou Maoyang, Computer Engineering and Design, 2019 (11).
- [3] Digital nerve extends to great wisdom -- commercial application of scanning neural network technology [J]Zhu Yan. Internet Weekly, 1999 (17).
- [4] Session oriented demand awareness attention graph neural network recommendation model [J]Zheng Xiaoli;Wang Wei;Du Yuhui;Zhang Chuang. Computer Engineering and Application.
- [5] Small sample face recognition based on twin neural network [J]Wan Lizhi;Zhang Yunchu;Gezhedong; Wang Chao. Journal of Shandong Jianzhu University, 2022 (01).
- [6] Expert recommendation method of attention map neural network based on memory [J]Chen Zhuo;Zhu Miao;Du Junwei;Yuan Ximing. Journal of Hunan University (Natural Science Edition), 2022 (06).
- [7] Research and analysis of housing price prediction based on neural network [J]Zheng Weipeng;Si Yachao. Journal of Hebei Institute of Architecture and Engineering, 2022 (01).
- [8] Topic trend prediction method combining cyclic neural network and convolutional neural network [J]Liu Chen;Liu Chao, Industrial Control Computer, 2022 (08).
- [9] Click rate prediction model based on deep neural network [J]Liu Hongli;Wusen;Wei Guiying;Li Xin;Gao Xiaonan. Journal of Engineering Science, 2022 (11).
- [10] Second change makeup system based on deep neural network [J]Shen Junyu;Li Xuchen;Hu Chao;Sun Xiaoyu;Qiu Shuangfei;Nie Xiaowei;Le Optics. Computer Knowledge and Technology, 2022 (27).