Research on Library Personalized Recommendation System Based on Restricted Boltzmann Machine

Xin-Yan WANG* and He HUANGa

Intelligent Science & Information Engineering College, Xi'an Peihua University, Xi'an, China

*36404843@qq.com, a2643304165@qq.com

Keywords: In-depth learning; Library; Personalized recommendation system.

Abstract. At present, most recommendation systems in libraries are content-based and collaborative filtering-based, but these recommendation systems ignore the deep-seated characteristics of readers' personalized information. In order to improve the function of Library recommendation system, this paper proposes a library recommendation system based on restricted Boltzmann machine and collaborative filtering algorithm, and simulates the performance of the algorithm. The results show that the proposed algorithm has good application effect.

Introduction

The core value of libraries is to realize the interaction between readers and librarians. The service concept should be reader-centered. Individualized service technology should be used to realize the different information needs of readers and to actively push the resources most suitable for users to readers, so as to improve the utilization rate of Library resources. Personalized information recommendation belongs to the category of information filtering, and it is also an important research direction of data mining. The earliest collaborative filtering system is GroupLens system. Since then, personalized recommendation has become an important and independent research direction.

With the development of personalized recommendation technology, its application in Library recommendation has gradually become a research hotspot. At present, most recommendation systems in libraries are content-based recommendation technology and collaborative filtering-based recommendation technology. But these recommendation systems only realize the statistics and analysis of the reader's mass data, ignoring the deep-seated characteristics of the reader's personalized information.

Deep learning is mainly used in the analysis and prediction of data. It realizes the learning and modeling, classification and prediction of massive data by machine. The results are more accurate than manual experience, and the application is very simple. In-depth learning includes Recurrent Neural Network, Convolutional Neural Network and Restricted Boltzmann Machine. Therefore, in order to improve the function of Library personality recommendation system, this study proposes a collaborative filtering library personality recommendation system based on Restricted Boltzmann Machine (RBM).

Recommendation System Framework Based On Restricted Boltzmann Machine

The recommendation process of recommendation system includes three aspects: obtaining feedback data from users; selecting appropriate recommendation algorithm model according to practical problems, which is the core of recommendation system; and obtaining recommendation results and feeding back to users. These three steps can produce more accurate and personalized recommendation results.

Traditional book recommendation mainly uses collaborative filtering algorithm, but the recommendation quality, efficiency and accuracy of this algorithm are often low, and the changes of readers' interest in books are ignored. In recent years, many scholars at home and abroad have studied recommendation system based on in-depth learning technology, such as recommendation
technology based on cyclic neural network, multi-layer perceptron. Recommendation technology, recommendation system based on restricted Boltzmann machine (RBM), etc. Therefore, this study proposes a Book Recommendation Algorithm Based on deep learning model (restricted Boltzmann machine).

**Books Recommendation Based on Restricted Boltzmann Machine and Collaborative Filtering**

The book recommendation algorithm based on RBM and collaborative filtering mainly includes two aspects: the construction of reader's book interest model, the evaluation and recommendation of book rating.

**Construction of Reader's Book Interest Model**

The construction of reader's book interest model is based on the historical information of reader's book borrowing stored in the library, including book type, renewal or not, borrowing cycle and the time node of book borrowing. The reader's interest in a book is defined as $I_{nt}$. See formula (1).

$$I_{nt} = \frac{t - \alpha}{\beta - \alpha}, 0 \leq I_{nt} \leq 1$$

In formula (1), alpha and beta respectively represent the minimum borrowing time of books in all the books borrowed by readers and the maximum borrowing time stipulated by the library; $T$ represents the relative borrowing time of readers, and the calculation formula is as follows:

$$t = \begin{cases} 
\alpha, t' \leq \alpha \\
t', \alpha \leq t' \leq \beta \\
\beta + 1 - \frac{t'}{\beta}, \beta < t' \leq 2\beta \\
\alpha, t' > 2\beta 
\end{cases}$$

Assuming that the library has $O$ books, $N$ represents the number of readers, $X$ represents the number of readers who have not borrowed the book, and 1-5 is the reader's book interest score, expressed in K. The reader's interest score $P$ of a book can be converted to a score of 1 to 5 according to formula (3).

$$\text{Score}_i = \begin{cases} 
1, \ P \in [0,0.2) \\
2, \ P \in [0.2,0.4) \\
3, \ P \in [0.4,0.6) \\
4, \ P \in [0.6,0.8) \\
5, \ P \in [0.8,1.0) 
\end{cases}$$

RBM is a kind of random field based on energy function, which includes the weight matrix $W$ and the visible layer $V$ between the hidden layer $h$, the hidden layer and the visible layer. Assuming that the bias of the hidden layer $h$ is $A_j$ and the bias of the visible layer $V$ is $b_i$, this paper proposes an item-based real value RBM (IR-RBM) model based on the traditional RBM. The IR-RBM model is established for book recommendation projects. The number of hidden layer units $F$ is set, the number of visible layer units $M$ is set, and the number of visible layer nodes is an integer between 0 and K. Among them, 0 means that the reader has not borrowed the book, and the rating of borrowed reader is expressed by 1-5 real value. In the IR_RBM model, the visible layer $V$ is replaced by the matrix of $O^*K$ (reader's book interest score column matrix), where K is the result of K binary scoring, K=5. In Figure 2, Missing represents the reader's non-borrowed book interest score, without probability calculation, the energy function of IR-RBM can be found in formula (5).
\[
E(v, h|\theta) = \frac{1}{2} \sum_{i=1}^{M} v_i^2 - \sum_{i} W_{ij}v_i - \sum_{i} v_i b_i - \sum_{j} a_j h_j
\]  
(4)

Formula (4): \(AJ\) is the offset of the jth node of the hidden layer, the initial value is set to 0; \(Bi\) is the offset of the ith unit of the visible layer; and \(Wij\) is the connection weight between the ith visible unit and the jth node of the hidden layer. Initialized as a random number satisfying the normal distribution \(N(0, 0.01)\). \(\theta\) is a parameter of RBM, which is composed of \(\{W, a, b\}\). According to the energy formula, when the scoring data is input into the model, the activation probability \(p\) of the j hidden layer node is 1.

\[
p(h_j = 1|v, \theta) = \sigma(a_j + \sum_i v_i W_{ij})
\]  
(5)

Formula (5): \(\sigma\) represents the activation function. \(\sigma(x) = 1/1 + \exp(-x)\).

When the state of the hidden layer is determined, the value of the ith element is calculated according to formula (6).

\[
v_i = b_i + \sum_j h_j W_{ij}
\]  
(6)

Parametric updating is achieved by using the method of contrast divergence. The criteria for updating parameters are shown in Formula 7. \(E\) stands for learning rate, and the initial value is set to 0.001.

\[
\Delta W_{ij} = e(\langle v_i h_j \rangle_{data} - \langle v_i h_j \rangle_{recon})
\]  
(7)

**Evaluation and Recommendation of Book Scoring**

After establishing the IR-RBM algorithm model, it is necessary to score books according to readers’ interest in borrowing books, and recommend books with higher scores to readers.

Score valuation:
1) According to the interest model, the ratings of reader \(u\) on visible units in IR-RBM model are obtained.
2) Computation of J nodes of all hidden layers.

\[
\hat{\rho}_j = \rho(j = 1) = \text{sigmoid}(\sum_{i} \sum_{v_i} v_i^k, v_i^k) = \frac{1}{1 + \exp(-\sum_{i} \sum_{v_i} v_i^k, v_i^k)}\]

(8)

3) For \(k = 1,\ldots, K\), calculation

\[
P(v_i^k = 1|\hat{\rho}) = \frac{\exp(\sum_{j=1}^{F} \hat{\rho}_j W_{ij}^k)}{\sum_{k}^{K} \exp(\sum_{j=1}^{F} \hat{\rho}_j W_{ij}^k)}
\]  
(9)

4) Take \(k = 1,\ldots\). The expectation of all estimates of \(K\) is the final score of the book.

\[
R(u, i) = \sum_{k=1}^{K} P(v_i^k = 1|\hat{\rho})^k
\]  
(10)

Through collaborative filtering Top_N algorithm calculation, the book with the highest reader rating is recommended. Using the trained IR-RBM to predict the score of \(Ru\) and \(i\), the top_N algorithm of collaborative filtering is used to calculate the N books with the highest ratings of the readers and recommend these books.
Simulation Experiment Analysis

Setting up of Experimental Environment

In order to verify the validity of the algorithm proposed above, simulation experiments are carried out, in which Python Programming and MATLAB simulation are used to implement the algorithm. Taking the real data of book borrowing in a university library as an example, this paper chooses 3000 user data, 12,000 book borrowing records and 6,000 book data as test and training sets after extracting and modifying the data. The 12,000 records of the book lending behavior data set are randomly divided into five parts. One part of the data set is randomly selected as the test set of the experiment and the other four parts are the training set.

Evaluation Parameter

The evaluation index of this paper is accuracy. Let X be the training set of this system, i.e. 12,000 book borrowing records, of which T(u) is the favorite book set of user u; based on the borrowing records of readers, the list of books recommended by user u is made up of an algorithmic model, whose length is N and recorded as R(u); and the calculation of book recommendation accuracy is shown in formula (11).

\[
\text{Precision} = \frac{\sum_{u \in U} |R(u) \cap T(u)|}{\sum_{u \in U} |R(u)|}
\]

(11)

Based on the training set, the user book interest model is constructed, and then the personalized recommendation of books is realized by using collaborative filtering recommendation algorithm and the data in the training test set based on the restricted Boltzmann machine and collaborative filtering recommendation algorithm model respectively. Five experiments have been carried out, that is, five groups of data take turns to do the test set. The results show that the proposed algorithm has a higher accuracy rate of book recommendation.

Conclusion

With the development of computer technology, the functions of library personalized recommendation system will be more perfect, and the accuracy of recommendation will be higher and higher. Compared with collaborative filtering recommendation algorithm, the proposed model of book recommendation algorithm based on restricted Boltzmann machine and collaborative filtering has improved the accuracy of book recommendation significantly. I believe that with the development of Science and technology, the accuracy of book recommendation will be improved significantly. It will be higher and higher to serve the readers better.

Acknowledgement

This research was financially supported by the Xi'an Peihua University's school-level scientific research project in 2019, the topic is the Design and research of intelligent Garbage classification management system based on Internet of things, the number is PHKT19068.

References


