Telecommunications package recommendation algorithm based on Deep forest

Yanhong Zhang, Meng Wang* and Yingfu Yu

Department of Aviation Foundation, Naval Aeronautical University, Yantai, China

*Corresponding author’s e-mail: wendao_2008@163.com

Abstract. In view of the wide variety of telecom packages and the difficulty of adapting to the needs of users, this paper introduces a recommendation model for telecom packages based on deep forests. This paper first analyzes the telecom package data, and then optimizes the deep forest according to its characteristics such as discrete, continuous attribute interleaving and high coupling characteristics, including the use of decision trees to discretize continuous features and design continuous window sliding mechanism. These methods can improve the ability of deep forest combination high coupling features. Finally, the model optimization measures were verified by detail experiments. The experimental results show that the optimized deep forest can be applied to the telecom package recommendation field. Compared with other shallow models and unoptimized deep forest models, the deep forest model has increased the F1 score by 5%; after adjusting the deep forest hyper parameters, the F1 score can be increased by 2%.

1. Introduction

In recent years, the growth rate of the telecommunications market has slowed down, and the competition for stock users has become fiercer. In order to better retain customers, telecom operators have launched a variety of telecom packages to meet the differentiated needs of users and attract users. Faced with a wide variety of package business, how to recommend the most suitable telecom package for users based on user consumption habits and preference data has become a very challenging task in the field of data analysis. In recent years, with the rise of big data and machine learning, a variety of new methods based on machine learning have been proposed [1, 2].

The mature deep learning model begins with the deep belief network proposed in the literature [3], which realizes the abstract expression of data information and the generation of high-level concepts by constructing a multi-layer feature extraction mechanism. At present, deep learning has been successful in the fields of image classification, speech recognition, network situational awareness and high-dimensional time series modeling [4]. At the same time, deep learning has been widely applied and researched in the field of personalized recommendation. The literature [5] proposed a collaborative filtering algorithm based on deep learning. The noise side automatic encoder is used to encode the project side information, and then the Person similarity score value with other items is calculated, and finally added with the timeSVD++ score as the final score value. The literature [6] designed a deep semantic model for information recommendation, using deep neural network to extract user features and project information into the same hidden space, and based on the cosine similarity between the user and project features to score regression prediction. It can be seen that the advantage of deep learning is that it can automatically extract features, avoid the inadequacy of artificial screening features, and improve the performance of predictive models.
In view of the advantages of deep learning and the shortcomings of current telecommunication recommendation algorithms, this paper studies the application of deep learning in telecom package recommendation, and proposes a telecom package recommendation algorithm based on deep forest. By constructing a deep forest based on dilated windows mechanism and other optimization methods, the model effectively realized the personalized matching between user needs and telecom packets.

2. Related Work
In recent years, the telecom package business recommendation system has gradually become a research hotspot. Multiple methods are proposed, such as K-means clustering, collaborative filtering and decision tree, et al. Literature [7] uses the decision tree to perform potential association mining on user attribute features, classifying customers into 6 categories and making personalized recommendations. In order to solve the telecom business recommendation task, K-means clustering algorithm and the Apriori algorithm are improved, and the C5.0 decision tree is used to screen the target customers [8]. The literature [9] proposes a random forest model based on the telecom operator's outbound recommendation system. The literature [10] calculates the user activity, establishes the three-dimensional user context of the user's external, internal and communication circles, and recommends the package business to the user. The literature [11] uses the Gaussian mixture model to model the user's call behavior and obtain the user's call mode. In this way, the user is recommended to personalize the service of idle call optimization.

These methods are all strong association features for mining user needs, and use the model to establish a mapping relationship between user requirements and telecom package characteristics. Although these methods can better realize the personalized recommendation task of the telecom package, there are still some shortcomings, which are mainly reflected in:
1) The need to select the data characteristics manually, and need to select and reconstruct the features before designing the model;
2) For shallow structures, models have limited ability to model complex problems.

3. Recommendation algorithm based on Deep forest

3.1. Deep forest structure
The literature [12] proposes a tree model based on deep learning structure. The model has the following three characteristics
1) Using a random forest to form a cascade forest, the initial features will be processed by a random forest and the probability vector of the decision result will be obtained, and then combined with the initial features and passed to the next random forest;
2) Using the multi-granularity scanning mechanism to obtain the temporal relationship of the initial features, the algorithm uses different sliding windows to scan the original data with unsynchronized numbers, and obtain initial features of different dimensions for the training of random forests;
3) The final classification result of the model is averaged for the probability result of the last layer of forest, and then the maximum value is taken to obtain the final result.

It has been verified by experiments that deep forest achieves the same scores as image classification, face recognition, sentiment classification and low-dimensional data. Compared with traditional deep learning algorithms, deep forest has the advantage of super-parameters. And the amount of computation of the model training, the main hyper parameter of deep forest is the number of forests per layer, and the number of decision trees in each forest. The hyper parameters to be configured are much smaller than the neural network; the main operation of deep forest is the feature information gain. The calculations are much lower than those of large-scale matrix operations and convolutions in deep neural networks, and there is no need to fine-tune the parameters. Aiming at the shortcomings of dropout, a stochastic constrained Boltzmann unit model design is proposed. Through the combination of visible unit visible elements and multi-model features, the data learning ability of the model group is improved and applied to Deep Boltzmann Machine.
3.2. Telecom package business data analysis
The telecom package business data has 26 dimensions, which correspond to whether the fixed mobile convergence package, whether it exceeds the package amount, the contract type, whether it is low-consumer users, gender, etc., which differs from the data collection of the experiment [4]

1) There are many types of features in the dataset, including discrete values and continuous values;
2) The feature dimensions of the data are small, and there are coupling relationships between different features, which need to be eliminated and reduced.

3.3. Deep forest optimization
According to the characteristics of telecom package business data, the improvement of deep forest is mainly manifested in the following two points.

1) Discretization of continuous variables through C4.5 decision tree;
   We used the C4.5 decision tree to discretize the continuous value features, as described below, sorting the samples that need to be processed from small to large according to the size of the continuous variables. Assuming that the attribute corresponds to a total of N attribute values, there are a total of N-1 possible candidate segmentation value points, and the value of each candidate segmentation threshold point is in the above-mentioned sorted attribute values.

2) Construction of feature fusion mechanism based on dilated windows.
   For the coupling relationship between different features in the dataset, such as 1_total_fee, 2_total_fee, 3_total_fee, etc., it needs to be culled and reduced. The traditional deep forest adopts a continuous sliding window to extract the dimension of the input data. This method takes into account the timing relationship between the features and has certain feature combination functions. However, this method lacks processing power for the combination of information between discontinuous features. The dilate convolution method proposed in the literature [12] is used to expand the sensing range of the sliding window, so as to improve the ability of the model to combine features. The dilated windows operation is shown in Fig. 1, wherein, a blank 1 indicates that dilated size is 1.

![Fig. 1. Dilated window sliding mechanism](image)

According to model optimization, deep forest training is as follows

1) Training data to obtain random training samples of different dimensions through multi-scale dilate windows;
2) Generate two different random forests based on the acquired training samples, one is a completely random forest and the other is a fixed random forest;
3) According to the results of random forest classification, the probability that the sample belongs to a certain class is added to the upper random forest training data as a new dimension of training data, and the random forest is continuously generated;
4) Continue to steps 2, 3 until the accuracy of the verification set is reduced, i.e. the training is over fitted.
Table 1. Experimental results of different models.

| Models              | F1-score /% |
|---------------------|-------------|
| Random forest       | 75.21       |
| C4.5 decision tree  | 70.32       |
| SVM                 | 74.11       |
| Deep forest         | 80.90       |
| Deep belief nets    | 81.20       |

Table 1 shows the experimental results of experiment 1. It can be seen from the table that the average value of F1 obtained by using deep forest and deep belief network is the highest, followed by random forest and C4.5 decision tree. This indicates that for shallow structure decision trees, the ability to deal with complex association problems is poor; random forests improve the ability of models to fit complex problems by constructing a higher-level decision mechanism; deep forests and deep belief networks as deep models. Compared with the shallow model, it has better feature learning ability. In view of the difficulty in setting the hyper parameters of the deep belief network and the time-consuming training, the deep forest has obvious advantages.

Fig. 2. Comparison of F1 scores between different deep forests

Fig. 2 shows the experimental results of experiment 2. As can be seen from the figure, the deep forests that adopt the hollow window sliding mechanism are mostly superior to the traditional deep
forests, especially the case where the window dilated size is 5. In view of the characteristics of the training data, for the five characteristics of the 1_total_fee, 2_total_fee, 3_total_fee, 4_total_fee, and 5_total_fee coupling mechanisms, the model has the highest F1 value when the maximum dilated size is 5. However, as the size of the cavity increases, the F1 value of the model appears to decrease. The reason is that the large size of the dilated will cause the feature dimension of the model training to be too small, resulting in the phenomenon of over-fitting. The same reason also leads to the phenomenon that the number of layers in the deep forest increases, and the F1 score does not increase and decrease.

4. Conclusion
In this paper, we use the deep forest as a model to predict user requirements for the personalized recommendation of telecom packages. In view of the characteristics of the training data, two methods such as continuous feature discretization and introduction of the dilated sliding window mechanism are used to optimize the deep forest. The experimental results show that the prediction accuracy of user matching packages is improved by constructing deep forests; the deep forest optimization method introduced helps the performance of deep forests. Future work will continue to optimize deep forests for different problem characteristics, and draw on deep forest design ideas, design other practical deep learning models and apply these models to more fields.

References
[1] Zhang, T. J ., Luo, X. G., Cai, L. Q., & Tang, J. F. (2014). Optimization model for mobile billing suite based on customer choice analysis. Systems Engineering-Theory & Practice, 34(2), 444-450.
[2] Shokouhyar, S., Dehkhodaei, A., & Amiri, B. (2021). A mixed-method approach for modelling customer-centric mobile phone reverse logistics: application of social media data. Journal of Modelling in Management, ahead-of-print(ahead-of-print).
[3] Nakano, T., Kasasaku, K., Minari, T., Satake, K., Yokoo, Y., & Yamanaka, M., et al. (2000). Geochemical characteristics of wet precipitation on the deep-forest, mountainous island of Yakushima, southern Japan: Sr isotopic signature of plant-derived Ca in rain.
[4] D Guo, Chen, H., Tang, L., Chen, Z., & Samui, P. (2021). Assessment of rockburst risk using multivariate adaptive regression splines and deep forest model. Acta Geotechnica, 1-23.
[5] Hafez, M. M., Redondo, R., Vilas, A. F., & HO Pazó. (2021). Multi-criteria recommendation systems to foster online grocery. Sensors, 21(11), 3747.
[6] Chen, Y., Guo, A., Chen, Q., Quan, B., & Hao, Z. (2021). Intelligent classification of antepartum cardiotocography model based on deep forest. Biomedical Signal Processing and Control, 67(2), 231-234.
[7] Zhu, G., Wang, Y., Cao, J., Bu, Z., & Liu, J. (2021). Neural attention travel package recommendation via exploiting long-term and short-term behaviors. Knowledge-Based Systems, 211, 511-513.
[8] Seiford, L. M., Cooper, W. W., & Zhu, J. (2000). A unified additive model approach for evaluating efficiency and congestion. Socio-Economic Planning Sciences, 34(1), 1-25.
[9] Chen, Z., Sun, X., & Shen, L. (2021). An effective tumor classification with deep forest and self-training. IEEE Access, (99), 1-1.
[10] Loto, R. T., Loto, C., Okeniyi, J., Agobe, E. I., & Opeoluwa, A. (2021). Gravimetric and data analysis of the protection performance of plant extracts on plain carbon steel in dilute h2so4 solution. Materials Science Forum, 1042, 79-87.
[11] Kim, D., Sang, W. K., Ji, H. L., Nam, K. M., & HJ Lee. (2021). A study on applicability of predictive maintenance and vibration data analysis by fast fourier transform into major equipment of submarine. Journal of the KNST, 4(1), 14-21.
[12] Zhou, W., Nie, L., Sun, F., Xu, X., & Zhang, Y. (2020). Automatic approach for fast processing and data analysis of seismic ahead-prospecting method: a case study in yunnan, china. Mathematical Problems in Engineering, 2020(12), 1-11.