International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Recommendation System for E-Learning Resources by Neural Networks A Preview

Nency Sharma [a], ***Dr. Anand Rajavat*** [b]

[a] *M. Tech Scholar, Department of Computer Science Engineering, Shri Vaishnav Vidyapeeth Vishwavidyalaya, Indore, India*
[b] *Director, Shri Vaishnav Vidyapeeth Vishwavidyalaya, Indore, India*

A B S T R A C T

With several learning resources available online, a recommendation system can recommend appropriate resources based on the student's need. Automatic multimedia learning resources recommendation has become an increasingly relevant problem: it allows students to discover new learning resources that match their needs, and enables the e-learning system to target the learning resources to the right students. Automatic multimedia learning resources recommendation has become an increasingly relevant problem, it allows students to discover new learning resources that match their tastes, and enables the e-learning system to target the learning resources to the right students. The task of recommendation algorithms in e-learning systems is to give student a personalized and suitable learning service With the increase in e-learning applications and its allied benefits, optimizing e-learning resources based on the requirements of the users has become a challenging task due to the large number of resources. Different individuals prefer different learning resources especially based on their age, country, neighborhood and requirements. There is no specific way to tailor make the recommendation system which can help users get to the best resources within the least amount of time. However, collaborative learning has come up as an effective technique in machine learning based approaches for the design of recommendation systems.

Keywords: **e-learning, learning resource recommendation system, ensemble neural network, collaborative learning,**

1. Introduction

E-learning has come to the forefront today due to the access to online experts, large number of resources and convenience of the targeted audience. Today, online learning resources are gaining huge popularity such as:

- Coursera
- Udemy
- NPTEL
- MIT.OCW etc
- Byjus etc.

With several learning resources available online, a recommendation system can recommend appropriate resources based on the student's need. Automatic multimedia learning resources recommendation has become an increasingly relevant problem: it allows students to discover new learning resources that match their needs, and enables the e-learning system to target the learning resources to the right students. Automatic multimedia learning resources recommendation has become an increasingly relevant problem, it allows students to discover new learning resources that match their tastes, and enables the e-learning system to target the learning resources to the right students. The task of recommendation algorithms in e-learning systems is to give student a personalized and suitable learning service [1]. The learning resources are more and more diversified recent years, and it could be audio, video, pictures, text, and so on. In the past decades, a multitude of recommendation algorithms has been developed. They can be divided into two groups: history data-based recommendation (HDBR) methods and content-based recommendation (CBR) methods. The HDBR methods have been widely researched for recommendation systems. These methods only rely on the user's history data without requiring the details of such resources. Collaborative Filtering (CF) is one of the most distinguished approaches [2]. CF methods can be classified into two types:

1. neighborhood-based method
2. model-based methods.

Model-based methods use L2 norm to normalize the solution [3]. HDBR methods require extensive historical data, which is difficult to obtain from the e-learning system and to achieve a decent performance, and they always suffer from the "cold start" problem. Thus, CBR methods may be a better choice for learning resources recommendation in e-learning systems. However, CBR methods characterize each user and item. Today learning resources come in various formats: audio, video, pictures, text, etc, and there is a huge amount of implicit information in learning resources that can be difficult to obtain, such as knowledge point, complexity, and prepared knowledge. Hence, content-based recommendation algorithm for learning resources is difficult to construct. The salient features of the approach employing collaborative learning and ensemble neural networks is described in the subsequent sections. Therefore, recommendation systems are modelled as systems applying the following approaches.

i. Techniques based on statistical learning

ii. Techniques based on time-series approach

iii. Techniques using Artificial Intelligence and Machine Learning

2. Literature Review

Shu et al. in [1] proposed a system that was primarily aimed towards the batch mode of active learning that is a sub-class of machine learning. In this technique the authors made an attempt to combine the positive attributes of active reinforced leaing with deep neural learning for the purpose of movie rating prediction. The authors also explored the mutual similarity among the features selected for training the neural network designed.

Moreno et al. in [2] proposed a technique that was based on the construction of a system that found the parameters in a family of arcs. The authors investigated the power of the opinion based emotional arcs in the prediction problems. The authors proposed that more audio-visual data could have a better well rounded prediction paradigm. The opinion or sentiment data was collected from sources such as Twitter and Reddit

Kohavi et al. in [3] proposed a recommendation system that was used for based on the support vector machine (SVM). The system was enhanced using the multi-class or multi-dimensional Support Vector Machine approach with splitting the hyper-plane of the SVM classifier. It was shown that the proposed system outperformed existing systems.

Mao et al. in [4] proposed a technique based on deep neural networks and deep learning in conjugation with tag extraction approach. The tags that justifiably affect the performance of the movie were chosen. It was shown that the proposed system could evaluate tags better than existing tag extraction systems and hence would yield better accuracy.

Hosburg et al. in [5] used the classification format after using the polarity of views about the concerned recommendation. The approach used Random Forests algorithm can prove to be effective in time series prediction problems. The major lacking aspect seen in this approach was the lack of data pre-processing so as to make the system learn additional parameters in prediction with higher accuracy. The final results did have a graphical representation making it more palatable. The approach used a very similar polarity count mechanism trying to extract polarized meanings form review sentences.

Wang et al. in [6]presented an approach that used the semantic expressions of Vietnamese comments regarding the data to be analyzed.The authors designed a methodology used in this study considering the short-term data as well as the day of week as inputs. The authors were intelligent enough to execute data structuring in a way so as to train the prediction system in the most effective manner.

Si et al. in [7]. addressed the challenge of box office prediction using a forward traversal learning rule with day classification. It is commonly used as the neural network needs to be fed using previous samples. The previous point is effective since the neural network may try to find out the correlation between the present-day rating and the day of the week.

Zhao et al. in [8] investigated the problem of prediction problems using the gray line artificial neural networks (GMANN). The approach was shown to perform better than the conventional ANN structure. It should be noted that there is no ruleto find the number of layers in the hidden layer of the number of neurons. Thus an exhaustive search needs to be carried out.

Pyo et al. in [9] present a technique that embodies the social network and extracted sentiment cycle. One could not say that the approach is very innovative; rather one would say that the approach is quite exhaustive and it tries to provide an optimized approach for time series prediction. They have tried out various maneuvered versions of the approach to obtain the optimized approach. It was however found that the proposed system obtained a relatively high accuracy of 88%.

Kassak et al. in [10] have utilized the techniques of machine learning and text mining analysis for prediction problems. The approach is exhaustive since the authors have tried out an exhaustive set of training functions. Another thing that they have done is the fact that they have also altered the configuration of the designed ANN to find the optimized one.

Zhao et al. in [11] proposed a technique to predict sales and recommend the relevant data. They rather focused on weighted sums of movies which had a significant impact on the data sets. They made a comparative analysis of the LM, NN and the SVM approaches. The proposed system results challenge the weak form of the Efficient Market Hypothesis (EMH) by demonstrating much improved and better predictions, compared to other approaches.

Polatidis et al.in [12] used the approach of supervised learning for designing recommendation systems. The paper made a comparative study of the work with a Probabilistic Neural Network (PNN) approach that tries to compute the cost function for prediction and uses the concept of conditional probability.

Koren et al.in [13]adopted the approach of RBF centers along with the ANN structure. The authors present an Artificial Neural Network (ANN) approach to predict box office success, particularly with respect to the forecast of their trend movements up or down. They used the Immune Principle which had no clear demarcation in case the values of the predicted movies did not fluctuate.

Salah et al. in [14] used the concept of cross domain datasets for training the ANN structure designed. The authors use different AI tools v.i.z. the Multi-layer Perceptron (MLP), the Convolutional Neural Networks (CNN), and the Long Short-Term Memory (LSTM) recurrent neural networks techniques. It was found that the proposed approach put for an optimized network paradigm for box office success prediction.

Pan et al. in [15] proposed a technique using appropriate time window and prediction horizon has been chosen. This makes the prediction much more accurate. Scaling and normalization of data has been done. This has been done since all features do not affect the outcome in the same way. 80% of the data has been used for training and 20% of the data has been used for testing. The authors use 10% of training set for hyper parameter optimization.

Bell et al. in [16] present a new data mining approach for the forecasting problems of web applications. The proposed approach focused on the subset of actual data used, the architecture of neural network used and the learning algorithm to be employed were chosen according to the sub-set data.

Ignatov et al. in [17]gives presented a comparative analysis of the different neural network architectures for the prediction of time series box office prediction. The problem being addressed in the paper is the upward or downward movement of the ratings with respect to time. It was shown that inclusion of semantic analysis proved critical in increasing accuracy.

Claster et al. in [18] suggested a technique for analyzing multi-dimensional sentiments for classification of data. The authors investigated the power of the opinion based emotional arcs in the prediction problems. The authors proposed that more audio-visual data could have a better well rounded prediction paradigm.

Zhai et al. in [19] proposed times series predictive models based on the use of use different AI tools v.i.z. the Multi-layer Perceptron (MLP), the Convolutional Neural Networks (CNN), and the Long Short-Term Memory (LSTM) recurrent neural networks techniques. The authors emphasize upon the importance of choosing the correct input features, along with their preprocessing, for the specific learning algorithm one wants to use.

Fikir et al. in [20] proposed a technique based on collaborative filtering that could be used for movie rating prediction. The main focus was on the reduction of computational cost and increasing the accuracy of prediction.

3. Forecasting and Recommendation Framework

A holistic framework for recommendations is shown in the figure below.

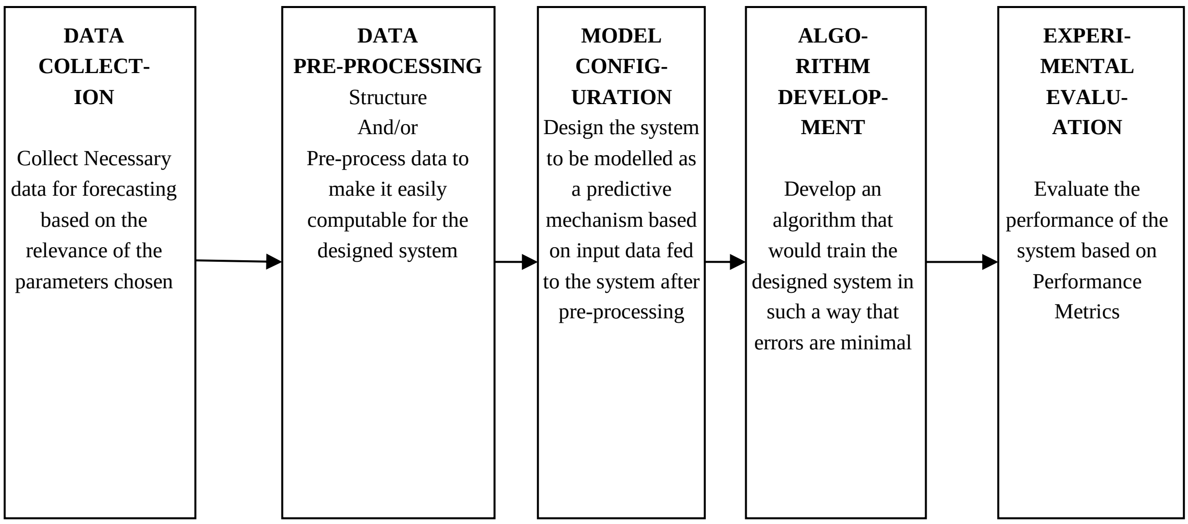

Fig. 1: Framework for Predictive-Recommendation

4. Proposed Methodology

4.1 Introduction to Artificial Neural Networks

Artificial Neural Networks (ANN) are a practical way of implementing artificial intelligence (AI). The design of ANN relies on the fundamental attributes of the human brain of the following:

1) Structure that takes data in a parallel manner.

2) The competence of the human brain to learn and subsequently adapt.

The self-organizing capability of the human brain is also tried to be incorporated in the ANN structure.

Considering that data comes in the form of parallel streams to the brain, it subsequently gets summed up and all data paths create weights or experiences, the mathematical model of the neural network can be given by:

$Y = \sum_{i=1}^{n} X_i . W_i \ + \ \theta_i$ (4.1)

HereX_i represents the data coming through a path, W_i represents the weight created by the data element X, θ is the bias.

ANN process data coming in the neurons of the input layer, then forwarding them to the hidden layer and finally yielding the result in the output layer[10].

The weights of the system depend on the input training data and it subsequently decides the synaptic connections and the weights[11] [12].

4.2 Collaborative Filtering for Recommendation Systems:

Training of ANN is of major importance before it can be used to predict the outcome of the data inputs. Neural Networks can be used for a variety of different purposes such as pattern recognition in large and complex data pattern sets wherein the computation of parameters would be extremely daunting for conventional statistical techniques. The weights or the equivalents of experiences are evaluated and updated based on the data patterns which are fed to the neural networks for training. The framework of collaborative learning consists of three major parts: the generation of a population of classifier heads in the training graph, the formulation of the learning objective, and optimization.

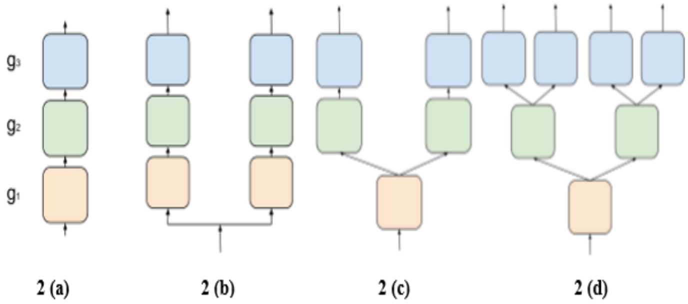

Fig 2 (a) Target Network (b) Multiple Instances (c) Simple ILR Sharing (d) Hierarchical ILRs haring

The figure above depicts the sub-categories of training. Similar to auxiliary training [8], we add several new classifier heads into the original network graph during training time. At inference time, only the original network is kept and all added parts are discarded. Unlike auxiliary training, each classifier head here has an identical network to the original one in terms of graph structure. This approach leads to advantages over auxiliary training in terms of engineering effort minimization. First, it does not require to design additional networks for the auxiliary classifiers. Second, the structure symmetry for all heads does not require additional different weights associated with loss functions to well balance injected backpropagation error flows, because an equal weight for each head's objective is optimal for training. Mathematically:

If the target network to be trained in figure 2(a) is given by:

$z = g(x, \theta)$ (4.2)

Here,g is determined by the graph architecture, θ represents the network parameters.The term g can also be represented as the cascade of the following sub-nets, given mathematically by:

$g(x, \theta) = g3(g2(g1(x1, \theta 1), \theta 2), \theta 3)$(4.3)

The cascade of the network is often termed as Ensemble Neural Network (ENN).Here,

$\theta = [\theta 1, \theta 2, \theta 3]$ (4.4)

In general, it is observed that that the training memory size is roughly proportional to the number of layers/operations. With the multi-instance pattern, the number of parameters in the whole training graph is proportional to the number of heads. Obviously, ILR sharing can proportionally reduce the memory consumption and speed up training, compared to multiple instances without sharing.Back propagation's popularity stems from the fact that:

1) Its stable

2) Its fast.

The mathematical treatment of the weight updating rule can be given by:

Let $\Delta\boldsymbol{\omega}$be the amount by which the weight is updated in every iteration. Then $\Delta\boldsymbol{\omega}$ is mathematically computed as:

$\Delta\omega = J^T J + \mu I^{-1} J^T e$ (4.5)

where $\boldsymbol{\omega}$ the weight vector,**I** is the identity matrix, $\boldsymbol{\mu}$ is the combination coefficient,

$$J = \begin{bmatrix} \frac{\partial e_{11}}{\partial \omega_1} & \frac{\partial e_{11}}{\partial \omega_2} & \cdots & \frac{\partial e_{11}}{\partial \omega_N} \\ \frac{\partial e_{12}}{\partial \omega_1} & \frac{\partial e_{12}}{\partial \omega_2} & \cdots & \frac{\partial e_{12}}{\partial \omega_N} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial e_{1M}}{\partial \omega_1} & \frac{\partial e_{1M}}{\partial \omega_2} & \cdots & \frac{\partial e_{1M}}{\partial \omega_N} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial e_{Q1}}{\partial \omega_1} & \frac{\partial e_{Q1}}{\partial \omega_2} & \cdots & \frac{\partial e_{Q1}}{\partial \omega_N} \\ \frac{\partial e_{Q1}}{\partial \omega_1} & \frac{\partial e_{Q1}}{\partial \omega_2} & \cdots & \frac{\partial e_{Q1}}{\partial \omega_N} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial e_{QR}}{\partial \omega_1} & \frac{\partial e_{QR}}{\partial \omega_2} & \cdots & \frac{\partial e_{QR}}{\partial \omega_N} \end{bmatrix} \quad (4.6)$$

Where **Q** is the number of training patterns, **R** is the number of outputs, **N** is the number of weights.Elements in error vector **e** are calculated by

$$e = \begin{bmatrix} e_{11} \\ e_{12} \\ \cdots \\ e_{1R} \\ \cdots \\ e_{Q1} \\ e_{Q2} \\ \cdots \\ e_{QR} \end{bmatrix}$$

$$e_{qr} = d_{qr} - 0_{qr} \qquad (4.7)$$

Where

$d_{qr}$and $0_{qr}$are the desired output and actual outputrespectively and training pattern is q.

4.3 Resilient Back Propagation

The Back propagation (BP) neural network is the most popular among all the neural network applications. It has the advantages of yielding high classification accuracy. However, practical applications are difficult to be satisfied because of the problems of slow learning and the likelihood of being trapped into a local minimum especially when the size of the network is large. These problems are due to the fact that the learning of BP neural network is mechanical and elementary. Many researchers have worked to overcome these problems, especially the local convergence [9].

The purpose of the resilient propagation (RPROP) training algorithm is to eliminate the limitations of these magnitudes of the partial derivatives. Only the sign of the derivative can determine the direction of the weight update; the magnitude of the derivative has no effect on the weight update. Another most difficult aspect of the back propagation learning was picking the correct training parameters. It also has the nice property that it requires only a modest increase in memory requirements. Additionally, resilient propagation is considerably more efficient than back propagation.

Resilient propagation, in short, RPROP is one of the fastest training algorithms available. The RPROP algorithm just refers to the direction of the gradient. In back propagation the change in weight is calculated with the magnitude of the partial derivative:

$\Delta w_{i,j}(t) = \alpha x_i(t).\delta_j(t)$(4.15)

Here,

$\boldsymbol{\alpha}$is the learning rate

$\boldsymbol{x_i}$is the propagating to ith neuron at time ‘t’

$\boldsymbol{\delta_j}$is the corresponding error gradient

5. Conclusion

With the increase in e-learning applications and its allied benefits, optimizing e-learning resources based on the requirements of the users has become a challenging task due to the large number of resources. In this paper, a collaborative learning-based approach has been proposed for ensemble neural networks (ENNs) to design a recommendation system for learning resources.

It can be concluded from the previous discussions that with more candidates opting for e-learning applications due to its specific advantages, it has become necessary to design an optimized recommendation system for learning resources. One of the major challenges is however the plethora of resources to choose from. The proposed work embodies the use of the age, country and personal interests of the users to design the recommendation system.

References

[1] Jiangbo Shu, Xiaoxuan Shen, Hai Liu, Baolin Yi, Zhaoli Zhang, “A content-based recommendation algorithm for learning resources”, Springer (2019)

[2]Moreno, M.N., Segrera, S., López, V.F., Muñoz, M.D., Sánchez, Á.L.: Web mining based framework for solving usual problems in recommender systems. A case study for movies′ recommendation. Neurocomputing 176, 72-80 (2018)

[3] Kohavi, R.: A study of cross-validation and bootstrap for accuracy estimation and model selection. In: proceeding International Joint Conference on Artificial Intelligence, pp 1137–1143 (2017)

[4] Mao, K., Chen, G., Hu, Y., Zhang, L.: Music recommendation using graph based quality model. Signal Process. 120, 806–813 (2016)

[5] Horsburgh, B., Craw, S., Massie, S.: Learning pseudo-tags to augment sparse tagging in hybrid music recommender systems. Artif. Intell. 219, 25–39 (2015)

[6] Wang, Y., Shang, W.: Personalized news recommendation based on consumers' click behavior. In: Fuzzy Systems and Knowledge Discovery (FSKD), 2015 12th International Conference on, 2015. IEEE, pp 634–638

[7] Shi, B., Ifrim, G.: Hurley N Learning-to-Rank for Real-Time High-Precision Hashtag Recommendation for Streaming News. In: Proceedings of the 25th International Conference on World Wide Web, 2016. International World Wide Web Conferences Steering Committee, pp 1191–1202

[8] Zhao, X., Yuan, J., Wang, M., Li, G., Hong, R., Li, Z., Chua, T.-S.: Video recommendation over multiple information sources. Multimed. System 19(1), 3–15 (2013)

[9] Pyo, S., Kim, E., Kim, M.: Automatic and personalized recom- mendation of TV program contents using sequential pattern mining for smart TV user interaction. Multimed. System 19(6), 527–542 (2013)

[10] Kaššák, O., Kompan, M., Bieliková, M.: Personalized hybrid recommendation for group of users: top-N multimedia recom- mender. Inf. Process. Manag. 52(3), 459–477 (2016)

[11] Zhao, W., Wu, R., Liu, H.: Paper recommendation based on the knowledge gap between a researcher's background knowledge and research target. Inf. Process. Manag. 52(5), 976–988 (2016)

[12] Polatidis, N., Georgiadis, C.K.: A multi-level collaborative fil-

[13] Koren, Y., Bell, R.: Advances in collaborative filtering. In: Recommender systems handbook, pp 77–118. Springer, Hei- delberg (2015)

[14] Salah, A., Rogovschi, N., Nadif, M.: A dynamic collaborative filtering system via a weighted clustering approach. Neuro- computing 175, 206–215 (2016)

[15] Pan, W., Xia, S., Liu, Z., Peng, X., Ming, Z.: Mixed factori- zation for collaborative recommendation with heterogeneous explicit feedbacks. Inf. Sci. 332, 84–93 (2016)

[16] Bell, R.M., Koren, Y.: Scalable collaborative filtering with jointly derived neighborhood interpolation weights. In: Data Mining, 2007. ICDM 2007. Seventh IEEE International Con- ference on, 2007. IEEE, pp 43–52.

[17]Ignatov, D.I., Nikolenko, S.I., Abaev, T., Poelmans, J.: Online recommender system for radio station hosting based on infor- mation fusion and adaptive tag-aware profiling. Expert System Appl. 55, 546–558 (2016)

[18] William B. Claster and DINH Quoc Hung, SubanaShanmuganathan, "Unsupervised Artificial Neural Nets for Modeling Movie Sentiment", 2010 Second International Conference on Computational Intelligence, Communication Systems and Networks, IEEE 2010.

[19] FangwenZhai, Qinghua Wen, Zehong Yang, Yixu Song, "Hybrid Forecasting Model Research on Stock Data Mining", IEEE 2010

[20] O. Bora Fikir, Ilker O. Yaz, TanselÖzyer, "A movie rating prediction algorithm based on collaborative filtering" IEEE 2010.