COMPUTATIONAL RESEARCH PROGRESS IN APPLIED SCIENCE & ENGINEERING (CRPASE)

CRPASE: TRANSACTIONS OF
ELECTRICAL, ELECTRONIC AND COMPUTER ENGINEERING

Journal homepage: http://www.crpase.com

CRPASE: Transactions of Electrical, Electronic and Computer Engineering, Vol. 06(03), 132-140, September 2020 ISSN 2423-4591

Research Article

News Recommendation Algorithm Based on Deep Learning

Shaik Sohail Ahmed*, Usama Hanif

University of Electronic Science and Technology of China

Keywords

Recommendar System,
Deep learning,
News Recommendations,
Neural Networks.

Abstract

Reading news articles online from different news sources has become a widely Popular around the globe. With a wide variety of news, articles available users are easily overwhelmed by information that they interested in. News recommender systems are one of the best approaches that help the user to find their interesting articles to read. The news recommender system presents the user the articles they are interested in rather than presenting them in order of their occurrence.In our work, we present our research on developing a news recommendation system with the help of the IT Home News App. We build the recommendation model from the two directions of user interest transfer and situational factors intersection where we consider the user interests and geo factors. The final prediction result of the whole model is made by combining the prediction of the two models. By combining two models, the whole recommendation model can obtain generalization ability and memory ability at the same time, which can better predict the news that users are interested in at the next time point and recommend the article to the users.

1. Introduction

In recent years, there is a vast craze for deep learning technology. At the same time, there is a massive growth in applications of neural network [1-6], optimizations and machine learning [7-43], computer technology, digital platforms and computer vision. Nowadays people tend to migrate to a digital platform to survive with the basic needs which brought a massive change in people but also brought problems such as information overload. Since a huge amount of data is available on the web but it becomes difficult to extract exact information required for the user. The content platform based on the Internet has gradually met people's needs in education, culture, entertainment, and other aspects. The main channels for people to obtain information have changed from TV, newspapers to Internet applications. Information is easy to be fission and widely spread on the Internet. Through the mobile Internet, we can know the world. Among them, Tencent News and other news applications have a super large user scale, which has become the head news platform.

It is precise because of the explosion of internet content development, network data surge, humans inevitably came to the era of big data[1]. The biggest characteristic of the big data era is massive redundant data, also called information overload. The amount of content that each user can accept at a given time is limited, and the growth of the amount of data makes it difficult for users to obtain interesting or matching content. For internet content producers, content distribution platforms, and users, it is important that they get the right information.

Because of this background, the recommendation system [2][3] came into being. In recent years, Internet head content distributors gradually find out the problem of information overload, and provide personalized recommendation as a basic service to users. Personalized recommendation system is a processing platform based on the analysis and mining of massive basic data. Through historical and original information, it can recommend content to users, realize content distribution in front of thousands of people, and increase the efficiency of content distribution. The main problem of the recommendation system is to model the user's interest and select the content with the highest score from

* Corresponding Author: ShaikSohail Ahmed
E-mail address: shaiksohail098@gmail.com

Received: 12 June 2020; Revised: 22 July 2020; Accepted: 28 July 2020

Please cite this article as: Sh. S. Ahmed, U. Hanif, News Recommendation Algorithm Based on Deep Learning, Computational Research Progress in Applied Science & Engineering, CRPASE: Transactions of Electrical, Electronic and Computer Engineering 6 (2020) 132–140.

the candidate set to recommend to the user. If users have a high similarity to the current recommendation content, they can click, purchase, and other behaviors; otherwise, users may ignore the recommendation and do not have any behaviors. Through this series of feedback on the ground, a personalized recommendation system cannot be short-term evolution, tend to improve. The fact of the mobile Internet proves that the personalized recommendation module has made a remarkable effect on the business-related collection. Tik-Tok, Amazon,Spotify, and so on, the personalized recommendation has become an indispensable service, which can give users more accurate content, enhance the daily average user market and retention rate, and also create a better user experience for users.

2. Literature Review

This section summarizes the traditional news recommendation research and selects the recommendation model suitable for situational factors. Finally, it introduces and summarizes the deep learning model applied in this research work.

Firstly, it defines the problem of news recommendation. News recommendation mainly starts from the user entity and news entity and builds recommendation model based on the user's own characteristics the features of the news in the user's browsing records and other situational factors mines the user's interest through the recommendation model, judges the news that the user may be interested in, and finally generates recommendation list. Therefore, the problem of news recommendation can be divided into two parts: user and news entity feature extraction and news recommendation model construction.

In the feature extraction part the user entity can be directly represented by the user's own attributes, while the news entity is mainly the text content, so it needs to use natural language processing related methods to digitize the text and then extract the features. In the next part of this chapter the text representation method will be studied and according to the advantages and disadvantages of the method combined with the news text content the text representation method will be selected. At the same time, this study will also use the method of deep learning to extract news features.In the construction of news recommendation model, this study puts forward new ideas in the application of situational factors and user interest transfer. The user's interest in news is related to the user's own situational attributes, and can also be reflected by the user's historical news browsing records. In the past research process usually under the premise that the user's history is missing or the target user is a new user, the basic information and situational factor information of the user are used to measure the similarity between the user and other users, and then the relevant interests of other users are used to build the recommendation model, such as collaborative filtering model and content-based recommendation model in the cold start situation Next. In this study, we hope to combine the user's personal information, situational factors, and user's historical behavior records to establish a model, through the feature fusion between different entities to make full use of the information contained in the feature variables, so as to achieve more accurate recommendation results.

3. Selection of Text Representation Method

In the past, most researches on text feature representation mainly focused on the topic of the article and the frequency of words It is a commonly used aspect of text feature representation, but it only considers the relative probability of the occurrence of words, and does not involve the position of words in the article, the sequence relevance between words and the actual meaning of words. In our work we choose to use Bert text vector to construct news text features. Word vector is a feature representation method of words in high-dimensional space, which can show the close degree of semantic relationship with other words through the distance between words. In general, the spatial dimension of word vector is 300-500 Among dimensions, the higher the dimension of word vector is, the more hidden variables it contains, and the higher the ability to describe words. The construction of news features through word vectors can show the article information in a higher dimension. By using attention model and special training method, the Bert model can better characterize the characteristics of the article. Moreover, it is a pre training model. It uses the large-scale computing resources and expectations of Google to obtain a very good training model. When it is used, only fine tuning is needed, and the optimal results are obtained in many open tasks.

4. Deep Learning Based Recommendations

Recommender system are now considered as a outbreak to overcome the data overloading and many researchers are trying to develop highly reliable recommendation techniques to the users. Generally, Recommender system is classified based on what information is and how it is visualized. However, it is mainly classified into three categories[2][16]

- Content-based recommendations: The user will be recommended items similar to the ones the user preferred in the past. In other words, the system normally compares the content of the user to the content that user shown interest in past.
- Collaborative recommendations: The user will be recommended items that people with similar tastes and preferences liked in the past. In this, system tries to find the similarity between the user and item. These kind of systems are wildly used in recommending movies, books music etc., to user.
- Hybrid approaches: These methods combine collaborative and content-based methods to improve the accuracy of the recommendation systems.

Deep learning based recommendation systems are classified into two types. They are Recommendation *with Neural Building Blocks* and *Recommendation with Deep Hybrid Models* as shown in figure 1. The deep neural networks architectures such as Multi-layer perceptron, Auto encoders, CNNs, RNNs, RBM, NADE, AM, AN etc., – In the first type, a type of NN architecture is used based on the requirement, and in the second type a combination of the NNs are used within the recommendation model. In this

section we will introduce different recommendations techniques that are used in deep learning tasks.

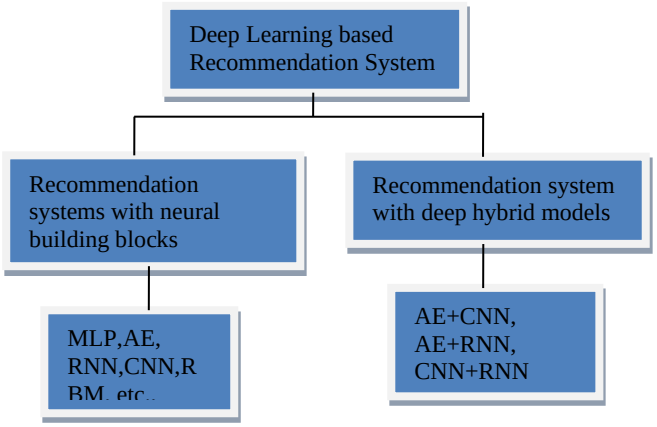

Figure 1. Classification of deep neural networks based Recommendations

5. News Recommendation Model Based on Deep learning

In recent years, with the development of deep learning theory and technology, some domestic scholars have introduced deep learning model into recommendation field. Xuejian Wang et al. Established a dynamic attention depth model and applied it to news recommendation, and achieved excellent results. Dynamic attention can mine users' news interest points and users' attention degree according to their browsing news history records, and match news list to recommend interested news for users according to their interest points and attention degree. This is a good example of the application of dynamic attention model to the field of news recommendation. This research applies the model originally applied to natural language processing to the field of news recommendation, and has achieved good results.

The main reason why the problem of news recommendation is difficult to solve lies in the complex structure of the data under the news scene, including the text data of the news itself, the user comments and browsing behavior data of the news, as well as the attribute data and situation data of the user itself. Because of such complex data features, it is difficult for the traditional recommendation model to comprehensively consider these multi-source data and integrate them to mine the user's interest through the association between features, which is usually not comprehensive for the user's feature mining. For this problem, Cheng et al. Proposed a deep and wide learning model for mobile app recommendation by using multi-source heterogeneous data such as user characteristics, context characteristics and project characteristics. The model has a high level of memory and generalization ability at the same time. Zheng The main idea of deep cooperative neural network is to use two parallel network models to learn the hidden features between users and projects. One is to build user preferences through user comments, the other is to build project features through project comments Build interaction layer on neural network to predict users' project rating and effectively improve the quality of recommendation. In the model of using neural network, researchers combine the characteristics of users, news and user behavior data through neural network model, which can achieve better recommendation effect than traditional recommendation model.

In foreign countries, many scholars have transferred the neural network model commonly used in the field of natural language processing to the field of recommendation algorithm for some improvement, and achieved good results. V Kumar proposes a deep neural network architecture based on two-level method. The news is expressed by embedding words. Using the articles read by users to generate user characteristics, and using deep neural network to train and find the corresponding relationship between news and users, remarkable results have been achieved. Shumpei Okura et al.[2] Introduced RNN into the field of news recommendation, expressed the news in terms of embedding words, used RNN to construct a self noise reduction encoder to embed the news words into RNN network, matched the output results with the news, and obtained the recommended news list. After testing, this method performed well in the actual data set of Yahoo. Ali elkahky et al. Use deep learning method to map users and projects to potential space, in which the similarity between users and their preferred projects is maximized. By introducing multi view deep learning model to expand the model, we can learn the characteristics of projects from different fields and users. In the practice of windows apps, it is proved that this aspect can achieve excellent cross domain recommendation effect.

6. Introducing Variables and Symbols

In this section, we first study the feature extraction of news and users, and then propose the theoretical framework of news recommendation model based on the features of news and users, as well as the defects of existing news recommendation research. Then we build the recommendation model from the two directions of user interest transfer and situational factors intersection, and finally integrate the two models to get the final news recommendation model as shown in figure 2.

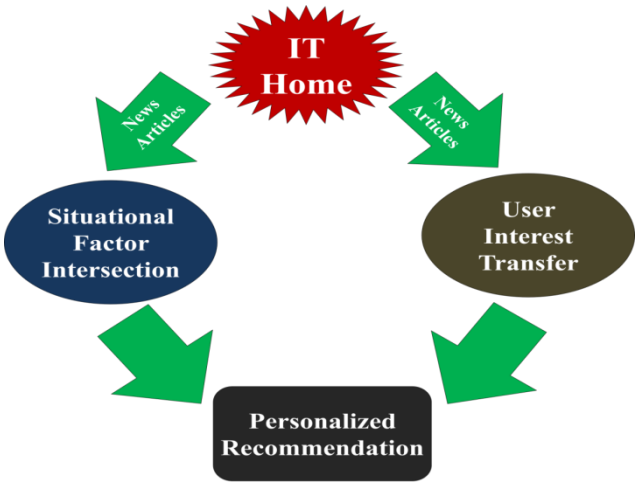

Figure 2. Our proposed model for News recommendation system.

7. News Feature Extraction

We use the Bert pre training model to train the word vector. We use 28w news texts collected from it home app

as corpus for training. The dimension of the word vector is set to 100. Through the selection of the fixed word synonyms, we can see that the word vector can measure the similarity between words very accurately.Text content needs to be preprocessed before establishing text features. First, we need to cut the text content words, to stop words and punctuation, so that the words contained in the text are mostly keywords, with actual meaning. In the process of preprocessing, keep the relative position of words in the article, so as to keep the word sequence relationship, so that the key meaning of the text will not deviate, and the meaning contained in the word sequence will not disappear due to the disorder of the word sequence. In the process of preprocessing, we can do our best Can retain the original meaning of more articles.

In order to ensure that the amount of data in the actual calculation process is not too large and the structure of data is not too complex, it needs to be based on the actual computing power of the hardware facilities selects top-N words in the document and retains the original order as the characteristics of the article. Due to the limitation of hardware computing power, the first 80 key words used in news text are used to construct article features,and the article is represented by vectors.

In this paper, long short memory neural network and convolution neural network are used as feature extraction methods. The original use of convolutional neural network is to extract image features from the pixel distribution of the image. In this research environment, the pixel matrix of image and the word vector matrix of text have many similarities. In a large number of text research, many researchers have transferred the convolutional neural network to the field of text mining. Convolution neural network is to extract and combine the local key features of news gradually through convolution, pooling and other steps, and finally get the final feature vector of the text through splicing. Long and short memory neural network can extract time series features, and because of its special gating structure, it can avoid gradient disappearance. The process of text feature extraction of convolutional neural network is shown in the following figure.

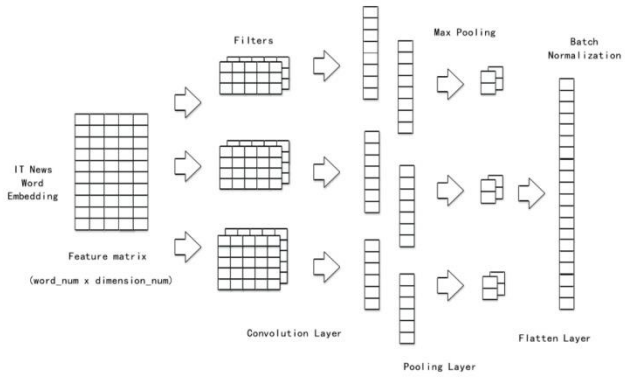

Figure 3. Text feature extraction of convolutional neural network.

8. User Space Chracteristics

At this stage, we need to acquire the spatial characteristics of users and transform them into the way that the model can be used expression. On the IT home news APP, when users comment, they will leave their city granularity geographic location in the comment information, and the enumeration value of China's geographic location data is huge. Due to the limited hardware conditions of this study, it is impossible to train a large number of user data at one time, and the geographic location data with city granularity will show strong sparsity. Therefore, the geographic location data with province granularity is used to reduce sparsity with coarse-grained data. This study does not consider the impact of population migration, only uses the province with the most frequent user comments as the province of the user, and uses one hot code to represent the user's city.

8.1. User Time Chracteristics

The comment data also contains the time of the user's comments, which represents the time when the user browses the news users may browse different types or topics of news at different times of the day and their interests will be the time period of days is different. Therefore, the time of the day will be processed in sections. According to people's normal work and rest. The time is divided into 6:00 ~ 11:00, 11:00 ~ 14:00, 14:00 ~ 17:00, 17:00 ~ 21:00, 21:00 ~ 24:00,

There are 6 time periods from 24:00 to 6:00. Map user comment time to each time period, using one hot code the comment time of the user.

At the same time, the extraction of comment time is also used to calculate the time when users comment the news from the current news recommendation time poor. In today's information age, users receive more and more information every day. Over time, the interest points may fade away gradually, so the time difference between users' comments on the news and the current news recommendation time is an important indicator to measure the weight of interest features. At the same time, the model also needs to consider that users' long-term interests do not change with time. Therefore, the model based on time dimension needs to consider both the long-term interests of users and the short-term interests that only appear in a certain period of time. In addition, it also needs to consider the possibility of other interests, so that the model has better generalization ability.

The user's preference can be expressed by the device type. The recommended object of this paper is it news the type of mobile phone users use can directly reflect the user's preference for a brand, or for a technology product, a function has a special preference. In more times, the devices used by users can more reflect the brand preference of users. For this reason, a special column will be opened in the IT home news APP according to the user's mobile phone brand to display the brand related news. This is a direct reflection of the user's use of the device and the brand that the user is interested in. At the same time, it can be further expanded according to the device type, including the price range of the mobile phone, which can reflect the user's social class to a certain extent. The technology highlights of mobile phone can reflect the user's preference for a certain technology field and so on. Through the expansion of external information, users' characteristics can become more abundant. However, due to the limitation of data conditions, only the type of

equipment brand is considered in this study, and no further expansion is made.

In the research of this paper, we hope to get users' implicit interest preferences through the intersection of mobile phone types and other characteristics of users. When users comment on news, they will directly mark the type of mobile phone in the comments and display the form is mobile phone brand - mobile phone model - mobile phone color. Due to the large number of mobile phone models and colors, it will cause too many variables Sparse, the memory ability of the model is reduced, and too many model parameters are difficult to train. So we only take the brand of mobile phone the one hot code is also used for variable conversion.

9. Combination of News and User Chracteristics

In this study, the feature variables are divided into two types: one is the feature variables that change with time, including users news browsing records at various time stages, as well as feature variables that do not change with time, including user's personal information and situational factors. The user's geographical location, the user's use of mobile phone brand, and the time period when the user comments on news are relatively stable variables, which will not change greatly on the time line. Therefore, the user's mobile phone brand, user's comment time period, and the user's geographical location features are combined to form a new feature vector. The feature variables, which are composed of user's personal information and user's situational factors, are the input features of the model. After the above information about users themselves, situational factors, as well as the extraction process of users' news history.

10. User Interest Migration

Interest migration refers to the change of user's interest in a period of time. User's interest is not immutable. This research is mainly based on the user's comments on news. Although users have a positive, negative or neutral attitude towards the news comments, users' comments on the news show that users have a certain interest in the news, which can be reflected to a certain extent by the text content of the news. In the theoretical model of user interest transfer in this study, the user comment news text is used as the basis for the user's interest in the news on the time slice. Through the news that the user comments on for a continuous period of time, the current news interest of the user is reflected, as well as the changes of the user's interest in the period of time. In this paper, the user interest transfer model will learn the user's interest in news and the change of interest to predict the user's interest in news in the next stage.

The goal of news recommendation theory model based on interest transfer is to combine multiple continuous time slices the interest reflected by this feature above and the change of interest are modeled for interest learning and fitting to predict the next under the time slice, users are interested in news. The news recommendation model based on interest migration is shown in Figure 3.

The design idea of news recommendation based on interest transfer is to extract features dynamically in time line set up a prediction model, fit the change of user interest, accurately predict the current interest of users, based on the existing news features, the user's interest in other news is speculated, so that the recommendation model can get enough generalization ability.

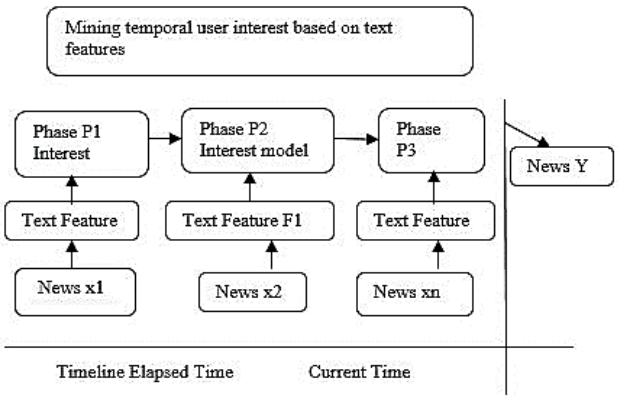

Figure 4. News recommendation model based on interest transfer.

11. Situational Factors and User Interest

The trend of users' interest over time is studied in depth, hoping to pass the trend of news features commented by users on the time line is deeply explored, and the user's long-term and short-term interests are considered to predict the user's interest at the next time point, and then the recommendation list is provided. Based on the LSTM model, this paper constructs a long-term and short-term memory model to learn user's interest in news in time series, including long-term interest and short-term interest. Considering the overall pattern of interest and the changing trend of users' interest, we can predict the news that users are interested in the next time period.

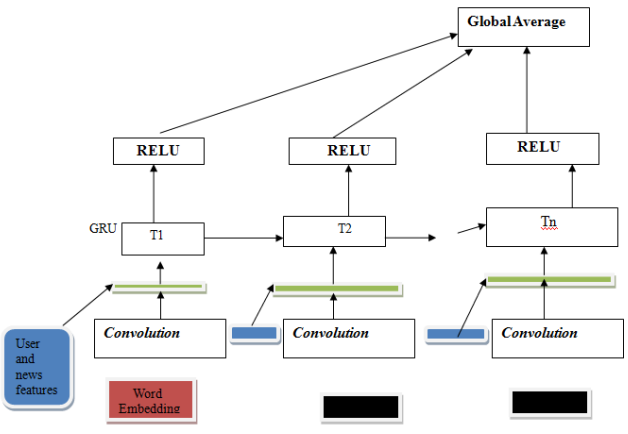

Figure 5. News recommendation model based on user interest migration.

In previous studies, situational factors were mostly used when a user lacked history or was a cold start user. The combination of situational factors and user personal information can reflect the similarity between users in a coarse granularity. In most previous studies, this similarity is used to find coarse-grained similar users to solve the cold

start problem of new users. It can be seen that the use of situational factors is relatively single, most of which are the linear relationship of situational factors, without considering the application of higher dimensions of situational factors.

But in the real environment, user interest is not only affected by a single attribute value, but also by the cross influence of multiple attribute values. Generally, the cross combination of discrete features can describe users more stereoscopically, and can also be extended on the original basis to describe users at a finer granularity. In the past research, only using a single situational factor can only divide into large categories of users, and the users in this large group may be different in most other features, so it is difficult to carry out research under the same conditions as the target users. However, the emergence of cross features increases the dimension of user groups, reduces the number of users in large groups, and screens out more individuals unrelated to the target users, which makes the relationship between users closer and the preference of interest more consistent. In the industry, the prediction model of user interest is often constructed by linear combination of attribute variables, without considering the effect of cross between features on user interest. In this study, feature cross is introduced into user interest model to describe user attributes in a higher dimension from the perspective of feature combination.

In this paper, we use the geographical location of users, the time window of users comments, the types of users devices and the records of users browsing news to cross feature, and explore the role of context cross feature in the news recommendation model. Prime features are shown in the Table 1.

Table 1. Prime features

USER_ID	News _Id	Location	Time Period	DEVICE	HAS INTEREST
F3da6fa7	20220554	Sichuan	20:00 ~ 24:00	VIVO	1
Sm33ff42	20246055	Beijing	10:00 ~ 15:00	APPLE	1
Ab7733k	29467557	Hubei	12:00 ~ 16:00	HUAWEI	0

Allwe mainly infer whether we are interested in news by situation cross tuple. At the same time, in order to mine the role of hidden variables in context cross, reduce the sparsity of data, and reduce the complexity of model parameter training, in this study, feature cross will be factorized, and hidden variables will be used to mine users' implicit interest preferences in news recommendation from a higher dimension. Feature combination can divide users into more detailed user groups, while fine-grained The users themselves have higher level of similarity, so the inclusion of hidden variables can divide the user groups from a higher dimension, and obtain the user's interest preferences more accurately.

After the establishment of the theoretical model of news recommendation based on interest transfer and the model of news recommendation based on situation intersection, the final prediction result of the whole model is made by combining the prediction of the two models. The model of news recommendation based on interest transfer can judge the news of interest by predicting the news interest of users, and the model of news recommendation based on user characteristics intersection can change Through the combination of the two models, the whole recommendation model can obtain generalization ability and memory ability at the same time, which can better predict the news that users are interested in at the next time point.

Based on the situational cross factorization machine news recommendation model, the main considerations are user information and situational factors to include, the geographical location of the user, the type of mobile phone the user uses, and the time period of each user comment. First of all, the variables are discretized.

The user comments are divided into 6 time periods: 6:00 ~ 11:00, 11:00 ~ 14:00, 14:00 ~ 17:00, 17:00 ~ 21:00, 21:00 ~ 24:00, 24:00 ~ 6:00. The geographical location is divided by province, the mobile phone type is divided by brand, and one hot is used In the coding process, three variables are combined as feature vectors. At the same time, the model will transform the features, increase the interactive features, and improve the nonlinearity of the model. The nonlinear feature is composed of the product of feature variables. In the transformation of features, based on the joint features, new features are generated through the interaction between features, and nonlinear modules are added to the model.

In general, discrete features are encoded by the unique heat coding method. Then, the number of dimensions of a feature is equal to the enumeration number of the discrete features. The number of parameters that need to be learned in this layer model is huge, and the cases where x_(i)and x_j are not zero are very few. This paper uses factorization machine to solve this problem. The auxiliary variable $V,V\in\mathbb{R}$ x is introduced into the model to represent W. Through the factorization machine, the complexity of the model is changed to O(kn), which greatly reduces the computational complexity of the original polynomial model. In parameter optimization, SGD is used to minimize the error between the predicted value and the real value.

The factorizer can be used as a hidden layer in neural network, and it can be trained by back propagation method together with neural network model. At the same time, in order to increase the nonlinearity of magic and prevent the model from over fitting too early, the activation function can be added to the output of factorizer. The situation cross factorization machine is composed of four layers. In the model, the vectors after the original features and joint features are firstly used as the input of the model. Because the feature vectors are very sparse, the embedding method is used to transform the high-dimensional sparse vectors into the low-dimensional dense vectors, and then the neural network training is carried out.

After the sparse joint feature and transformed feature are added into the embedding layer, dense vectors are obtained and hidden layer is constructed by the factorization

machine, and the RELU function is used as the activation function of the hidden layer.If the spatiotemporal feature model is trained separately, Softmax is used as the activation function for output layer.

12. Integrating News Dynamic Characteristics

LinUCB algorithm is one of the most famous algorithms in contextual bandit. It pioneers the introduction of context information into Bandit algorithm and has been regarded as a benchmark algorithm m by many academic researches. Its basic assumption is that there is a linear relationship between the expected return of each action and the characteristics of users, and it uses ridge regression to fit, and then uses the idea of UCB to recommend news to users.

After every recommendation, LinUCB will update the 0 vector of the recommended news according to the current user's context vector, user's feedback and article ID, so as to constantly optimize the accuracy of the 0 vector. LinUCB will update its model parameters every time it gets feedback from users, so it is an online algorithm. The algorithm can adjust its strategy according to the feedback of users, so it can reduce the recommendation priority of non popular news and improve the recommendation priority of popular news in time.

13. Experimental Results

The research object of this paper is IT home app, which was established on May 15, 2011. It home app is a leading real-time it information and digital product website in the industry. It home quickly selects Pan Technology News, shares real-time IT industry trends and digital product information following the trend, and provides high-quality PC and mobile technology articles and rich system application beautification resources. It is very representative in science and technology news app.

The data used in this study includes two parts: the first part is the history of users' browsing news behavior, news content users have commented on, comment content, comment floor, comment time point, support number, objection number. The second part is the user's personal characteristics, user ID, user nickname, user registration time, user level, user mobile phone type. The third part is situational factors, including geographical location, user comment window after the establishment of the model, the recommended method commonly used in industry is selected as the baseline algorithm.

The dataset for the experiment are the comments on the IT news App which they will leave their city granularity geographic location in the comment information.. Due to the limited hardware conditions of this study, it is impossible to train a large number of user data at one time, and the geographic location data with city granularity will show strong sparsity. Therefore, the geographic location data with province granularity is used to reduce sparsity with coarse-grained data. This study does not consider the impact of population migration, only uses the province with the most frequent user comments as the province of the user, and uses one hot code to represent the user's city.

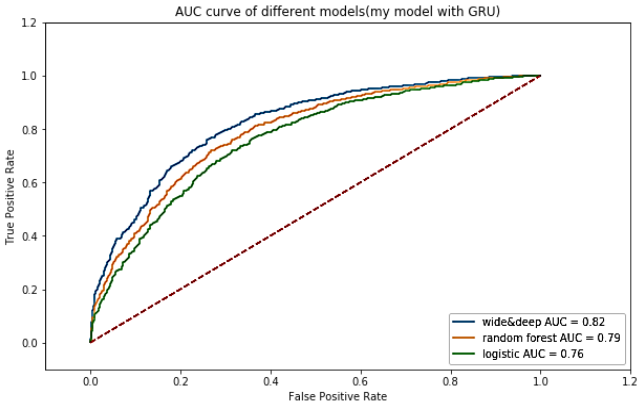

Figure 6. AUC Curve of different models

In our work, experiments are plotted with different models and compared with our model. In practical, performance measurement is an essential task. So when it comes to classification problems,we can count on AUC – Roc Curve. The x-axis represent the true positive rate and y-axis represent false positive rate

TPR/Recall/Sensitivity =TP/(TP+FN)

In this part, we conduct experiments by comparing our personalized news recommendation algorithm with the traditional filtering method. Initially, experiment compares the results of improved wide and deep algorithm with the traditional machine learning methods.

Initially, from the below plot the wide and deep with GRU (indicated in blue),random forest (indicated in purple), and logistic (indicated in green).Moreover the above results shows machine learning models are used to predit the best model and compared the performances. In prediction the AUC of the wide and deep model is 0.82 when compared to other models random forest and logistic are 0.76 and 0.79. hence our model best fit for recommending news articles to users. To extract the userinterest from the news we used different neural networks and compared the performance.

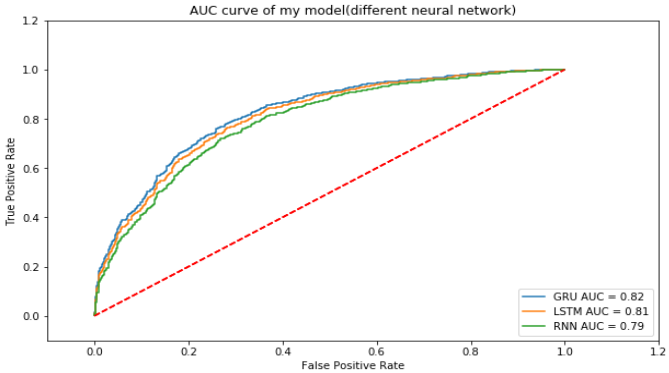

Figure 7. AUC Curve of the proposed model

In the figures we compared our model with different neural networks such as RNN and LSTM . from the plotted graph the GRU auc is indicated with blue line whereas LSTN

and RNN are indicated as purple and green lines.Although it is seen that the RNN(green line) performs the worst as it is always lower than all the others, it is not the problem of the approach, but it could be that the descriptions for the category are simply not good enough to highlight the common characters of categories (to increase the true positives) or to make them well distinguishable (to reduce the false positives). Additional features of the categories may be helpful.

The LSTM (purple line)beats the RNN model at the final dashboard ,which makes more encouraging to perform the best fit model to achive better personalized recommendation to users.

The improvements are seen with our model GRU (blue line) when compared with different neural networks. Therefore, if we want to provide better news recommendation to user we need to go with our model to provide the best match to user.The below figure shows that the comparison of different neural networks such as GRU, LSTM,RNN . GRU Auc is 0.82 when compared to Lstm and RNN plotted as 0.81 and 0.79.

14. Conclusions

The improvements are seen with our model GRU (blue line) when compared with different neural networks. Therefore, if we want to provide better news recommendation to user we need to go with our model to provide the best match to user.The below figure shows that the comparison of different neural networks such as GRU, LSTM,RNN . GRU Auc is 0.82 when compared to Lstm and RNN plotted as 0.81 and 0.79.

The Approach we used in this paper can be improved in several ways to get the better news recommending articles to uses. We considered two different dimensions (user interest transfer and situational factors) to implement the hybrid approach. The accuracy of our news recommender system can be improved by considering another dimension “location”, which could be implemented to recommend news articles to the user based on his location. Also, our users provided explicit feedback about the categories in which they were interested. The recommender system could be improved by using the users interests based on their reading habits. We should explore the effectiveness of the system and compare wide range of news articles to get better results.

References

[1] L. R. Ricci, , B. Shapira, Recommender Systems Handbook. 2011.

[2] S. Okura, Y. Tagami, S. Ono, and A. Tajima, “Embedding-based News Recommendation for Millions of Users,” pp. 1933–1942, 2017.

[3] S. Agarwal, Data mining: Data mining concepts and techniques. 2014.

[4] N. A. Golilarz, N. Robert, J. Addeh, A. Salehpour, Translation invariant wavelet based noise reduction using a new smooth nonlinear improved thresholding function, Computational Research Progress in Applied Science & Engineering 3 (2017) 104–108.

[5] N. A. Golilarz, N. Robert, J. Addeh, Survey of image de-noising using wavelet transform combined with thresholding functions, Computational Research Progress in Applied Science & Engineering 3 (2017) 132–135.

[6] N. A. Golilarz, H. Demirel, Thresholding neural network (TNN) based noise reduction with a new improved thresholding function, Computational Research Progress in Applied Science & Engineering 3 (2017) 81–84.

[7] A. Addeh, A. Khormali, N. A. Golilarz, Control chart pattern recognition using RBF neural network with new training algorithm and pracrical features, ISA Transactions 79 (2018) 202–216.

[8] N. A. Golilarz, A. Addeh, H. Gao, L. Ali, A. M. Roshandeh, H. M. Munir, R. Khan, A new automatic method for control chart patterns recognition based on ConvNet and harris hawks meta heuristic optimization algorithm, IEEE Access 7 (2019) 149398–149405.

[9] L. Ali, C. Zhu, N. A. Golilarz**,** A. Javeed, M. Zhou, Y Liu, Reliable Parkinson’s Disease Detection by Analyzing Handwritten Drawings: Construction of an Unbiased Cascaded Learning System based on Feature Selection and Adaptive Boosting Model, IEEE Access 7 (2019) 116480–116489.

[10] L. Ali, A. Niamat, J. A. Khan, N. A. Golilarz, X. Xingzhong, A. Noor, R. Nour, S. A. Chan Bukhari, An Optimized Stacked Support Vector Machines Based Expert System for the Effective Prediction of Heart Disease, IEEE Access 7 (2019) 54007–54014.

[11] F. Rahmani, D. Quispe,T. Agarwal, M. Barzegaran, Speed Control of Brushless DC Motor by DC-DC Boost and Buck Converters Using GaN and SiC Transistorsfor ImplementingtheElectric Vehicles , Computational Research Progress in Applied Science & Engineering 6 (2020) 70–75.

[12] M. Khatibi, F. Rahmani ,T. Agarwal, Comparative Analysis of Power System Model Reduction, Computational Research Progress in Applied Science & Engineering 6 (2020) 46–51.

[13] S. K. Dwivedi and C. Arya, A survey of news recommendation approaches, Proc. 2016 Int. Conf. ICT Business, Ind. Gov. ICTBIG 2016, (2017).

[14] Y. Xiao, P. Ai, C. H. Hsu, H. Wang, and X. Jiao, “Time-ordered collaborative filtering for news recommendation,” China Commun 12 (2015) 53–62.

[15] M. An, F. Wu, C. Wu, K. Zhang, Z. Liu, and X. Xie, Neural News Recommendation with Long- and Short-term User Representations (2019) 336–345.

[16] S. Zhang, L. Yao, A. Sun, and Y. Tay, Deep Learning based Recommender System: A Survey and New Perspectives 1 (2017) 1–35.

[17] M. Zihayat, A. Ayanso, X. Zhao, H. Davoudi, and A. An, A utility-based news recommendation system, Decis. Support Syst. 117 (2019)14–27.

[18] G. De Souza Pereira Moreira, F. Ferreira, and A. M. Da Cunha, News Session-Based Recommendations using Deep Neural Networks, ACM Int. Conf. Proceeding Ser., (2018) 15–23.

[19] G. Yuyun and Z. Qi, Hashtag recommendation using attention-based convolutional neural network, IJCAI Int. Jt. Conf. Artif. Intell., vol. 2016–Janua (2016) 2782–2788.

[20] S. Zhang, Y. Tay, L. Yao, and A. Sun, Next Item Recommendation with Self-Attention, 2018.

[21] H. Faris, A. A. Heidari, A. Z. Ala’M, M. Mafarja, I. Aljarah, M. Eshtay, S. Mirjalili, Time-Varying Hierarchical Chains of Salps with Random Weight Networks for Feature Selection, Expert Systems with Applications 140 (2020) 112898.

[22] A. A. Heidari, R. A. Abbaspour, H. Chen, Efficient boosted grey wolf optimizers for global search and kernel extreme learning machine training, Applied Soft Computing 81 (2019) 105521.

[23] H. Chen, C. Yang, A. A. Heidari, X. Zhao, An efficient double adaptive random spare reinforced whale optimization

algorithm, Expert Systems with Applications 154 (2020) 113018.

[24] I. Aljarah, M. Mafarja, A. A. Heidari, H. Faris, S. Mirjalili, Clustering analysis using a novel locality-informed grey wolf-inspired clustering approach, Knowledge and Information Systems 62 (2019) 507–539.

[25] H. Chen, A. A. Heidari, X. Zhao, L. Zhang, H. Chen, Advanced orthogonal learning-driven multi-swarm sine cosine optimization: Framework and case studies, Expert Systems with Applications 144 (2020) 113113.

[26] Z. Xu, Z. Hu, A. A. Heidari, M. Wang, X. Zhao, H. Chen, X. Cai, Orthogonally-designed adapted grasshopper optimization: A comprehensive analysis, Expert Systems with Applications 150 (2020) 113282.

[27] H. M. Ridha, A. A. Heidari, M. Wang, H. Chen, Boosted mutation-based Harris hawks optimizer for parameters identification of single-diode solar cell models, Energy Conversion and Management 209 (2020) 112660.

[28] H. Zhang, A. A. Heidari, M. Wang, L. Zhang, H. Chen, C. Li, Orthogonal Nelder-Mead moth flame method for parameters identification of photovoltaic modules, Energy Conversion and Management 211 (2020) 112764.

[29] H. Chen, A. A. Heidari, H. Chen, M. Wang, Z. Pan, A. H. Gandomi, Multi-population differential evolution-assisted Harris hawks optimization: Framework and case studies, Future Generation Computer Systems 111 (2020) 175–198.

[30] A. Abbassi, R. Abbassi, A. A. Heidari, D. Oliva, H. Chen, A. Habib, M. Jemli, M. Wang, Parameters identification of photovoltaic cell models using enhanced exploratory salp chains-based approach, Energy 198 (2020) 117333.

[31] E. R-Esparza, L. A. Zanella-Calzada, D. Oliva, A. A. Heidari, D. Zaldivar, M. Pérez-Cisneros, L. K. Fong, An Efficient Harris Hawks-inspired Image Segmentation Method, Expert Systems with Applications 155 (2020) 113428.

[32] L. Ali, S. Khan, N. A. Golilarz, Y. Imrana, I. Qasim, A. Noor, R. Nour, A Feature-Driven Decision Support System for Heart Failure Prediction Based on χ 2 Statistical Model and Gaussian Naive Bayes, Computational and Mathematical Methods in Medicine (2019) 1–8 .

[33] R. Khan, X. Zhang, R. Kumar, A. Sharif, N. A. Golilarz, M. Alazab, An adaptive multi-layer botnet detection technique using machine learning classifiers, Applied Sciences 9 (2019).

[34] L. Ali, I. Wajahat, N. A. Golilarz, F. Keshtkar, and S. A. C. Bukhari, Lda-ga–svm: improved hepatocellular carcinoma prediction throughdimensionality reduction and genetically optimized support vector machine, Neural Computing and Applications (2020) 1–10.

[35] I. Bargegol, M. Nikookar, R.V. Nezafat, E.J. Lashkami, A.M. Roshandeh, Timing optimization of signalized intersections using shockwave theory by genetic algorithm, Computational Research Progress in Applied Science & Engineering 01 (2015) 160–167.

[36] N. A. Golilarz, H. Gao, H. Demirel, Satellite Image De-noising with Harris Hawks Meta Heuristic Optimization Algorithm and Improved Adaptive Generalized Gaussian Distribution Threshold Function, IEEE Access 7 (2019) 57459–57468.

[37] H. M. Munir, R. Ghannam, H. Li, T. Younas, N. A. Golilarz, M. Hassan, A. Siddique, Control of Distributed Generators and Direct Harmonic Voltage Controlled Active Power Filters for Accurate Current Sharing and Power Quality Improvement in Islanded Microgrids, Inventions 4 (2019).

[38] L. Ali, C. Zhu, M. Zhou, Y. Liu, Early diagnosis of parkinson's disease from multiple voice recordings by simultaneous sample and feature selection, Expert Systems with Applications 137 (2019) 22–28.

[39] L. Ali, C. Zhu, Z. Zhang, Y. Liu, Automated detection of parkinson's disease based on multiple types of sustained phonation' using linear discriminant analysis and genetically optimized neural network, IEEE Journal of Translational Engineering in Health and Medicine 7 (2019) 1–10.

[40] N. A. Golilarz, H. Demirel, Image de-noising using un-decimated wavelet transform (UWT) with soft thresholding technique. In: Proceedings of the IEEE 9th International Conference on Computational Intelligence and Communication Networks (CICN), Girne, Cyprus, September 16–17 (2017).

[41] N. A. Golilarz, H. Demirel, Thresholding neural network (TNN) with smooth sigmoid based shrinkage (SSBS) function for image de-noising. In: Proceedings of the IEEE 9th International Conference on Computational Intelligence and Communication Networks (CICN), Girne, Cyprus, September 16–17 (2017).

[42] N. A. Golilarz, H. Gao, W. Ali, M. Shahid, Hyper-spectral remote sensing image de-noising with three dimensional wavelet transform utilizing smooth nonlinear soft thresholding function. In: Proceedings of the IEEE15th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP), Chengdu, China, December 14–16 (2018).

[43] N. A. Golilarz, H. Demirel, H. Gao, Adaptive generalized Gaussian distribution oriented thresholding function for image de-noising, International Journal of Advanced Computer Science and Applications (IJACSA) 10 (2019) 10–15.