УДК 004.89

Lyvytska D. O., student of the 4th year of specialty 113 «Applied mathematics»
Kovenko V. A., 1st year master's student of VNTU, specialty 126 «Information systems and technologies»
Sichko T.V., Ph.D., Associate Professor, Associate Professor of the Department of Information Technologies

OVERVIEW OF A NEURAL NETWORK BASED SEQUENCE MODELING APPROACH TO TACKLE A PROBLEM OF PERSONALIZED RECOMMENDATIONS IN STREAMING SERVICES

Vasyl' Stus Donetsk National University, Vinnytsia
Vinnytsia National Technical University, Vinnytsia

Nowadays recommendation systems is a key component of any streaming service holding either UGC (user generated content) or VOD (video on demand) content, which allows to optimize search through the catalog and increase user engagement. Among different types of recommendation systems, personalized one is the most crucial in both helping overwhelmed users to choose the next item to watch and lifting the time spent on a platform. Choosing the right approach to tackle a problem of personalized recommendations is a tricky question that involves many factors. In this work, a sequence modeling approach to tackle a problem of personalized recommendations is reviewed. The general tips and advice about data preparation and algorithm's modeling parts are provided.

Clearly understanding user preferences and recommending most suitable items to the end user are the essential tasks of recommendation systems. Such tasks are tightly coupled with the business values and KPIs. For instance Netflix states that the combined effect of personalization and recommendations save them more than $1B per year, whereas the YouTube's CPO, Neal Mohan, reports that over 70 percent of the time spent on YouTube is spent watching recommended videos. The choice of the type of recommendation system along with the actual algorithm is crucial as it influences both the infrastructure of the overall solution (technical part) and the set of possible features (business part). Content-based recommendation systems usually utilize content information stored in metadata to derive a similarity matrix and provide the end user with a possibility to see related content to the one watched [1]. The actual problem with such systems is that they do not consider user behavior when making the decision. On the other hand, user-based recommendation systems consider the similarity between users based on watched items. Both approaches, while working in practice, do not provide personalization to the needed extent. Recently, the usage of deep learning algorithms for tackling recommendation problems became popular [2, 3]. In this work, an overview of a sequence modeling approach for solving a problem of personalized recommendations is reviewed. The work provides a general overview of using such an approach in terms of data preparation and modeling.

Recommendation metrics are important to measure systems adequacy towards a specific data split. For simplicity, the following two metrics are considered.

MAP (Mean Average Precision) is a metric which is dependent both on the order and the relevance of recommendations. MAP is working with binary output (relevant/non-relevant), which ideally suits the case of recommendations from implicit feedback. It's calculated by taking an average over AP (Average Precision) for all the users:

$$AP@K \; = \frac{1}{K}\sum_{i=0}^{K} + \frac{H_i}{i+1},$$

where K- number of watched and predicted items, H- a list of cumulative hits.

The main logic behind MAP is that the further the relevant item is situated in the recommendation list, the lower the score is and vice versa.

Coverage is a metric, which calculates the portion of content covered by the model's prediction at specific length. This metric is simple and efficient for the calculation of model’s personalization. The logic behind it is that when the coverage is too big, the model is simply giving random recommendations, while when it’s too low the model is recommending the same items to all the users, meaning that no personalization exists. The best value of coverage is different for different applications. Coverage is calculated as the ratio of the number of unique items predicted by the model to all users to the number of unique items in the dataset:

$$\boldsymbol{COV@10} \; = \frac{\boldsymbol{N(P_{unique})}}{\boldsymbol{N(I_{unique})}},$$

where $N(P_{unique})$- number of unique items recommended by the model,
$N(I_{unique})$- number of unique items in the dataset

One thing which is mandatory for sequence modeling approaches is specific data preparation. In the recommendation domain, the actual user interactions can be modeled as a sequence. For simplicity of explanation, it will be assumed that feedback is just an implicit watch event. The actual user watching history can be considered as a sequence that should be modeled. The task is then stated as the following: recommend top-N new items to watch given that the user watched M items in the past. A very similar approach presented by Hidasi et al. suggests making recommendations based on user sessions, rather than on user history [4]. In practice, both approaches work nicely, while the latter gives an opportunity to create a bigger dataset for training. To prepare the dataset, one should split the sequences in the same manner it’s done for LM (language modeling) tasks in NLP, thus creating X sequences of item identifiers of length M-1 and Y ones of length 1. While the user history can be of variable length, it’s suggested to set the global maximal length for it (both because of computational and logical reasons), which can be derived from data analysis (median number of items watched in one user session or median length of user history during a particular period of time are possible options). Overpopulated items which are the product of promotions, local and global trends can bias the system, resulting in low coverage. To tackle this problem, it’s recommended to downsample too popular items on the level of sequence preparation or solve this problem through specific mechanics of the algorithm (imbalanced class weights could be an example). Finally, to solve a cold start problem, one can apply heuristic to pad training sequences with the most popular item, thus a user with no watching history will first get general most popular recommendations and during further interactions the model will adapt and suggest personalized items.

For modeling user behavior, the LSTM model is suggested. LSTM (Long short term memory) is a network from the family of RNNs and is usually used for sequence modeling. Having both short-term and long-term mechanisms inside LSTM allows it to overcome a problem of vanishing gradients, stabilizing the training and granting the possibility to work with longer sequences of data. Item identifiers are mapped to vectors through the embedding layer, allowing the model to learn hidden characteristics of them in a better way. Recent advances in NLP techniques have brought a new mechanism called attention to weight specific representation in the manner beneficial to minimization of loss function. Attention is also suggested for the usage in our scenario, both because of improved accuracy and a possibility to partially explain recommendations. The last layer is a classification one with softmax activation (figure 1).

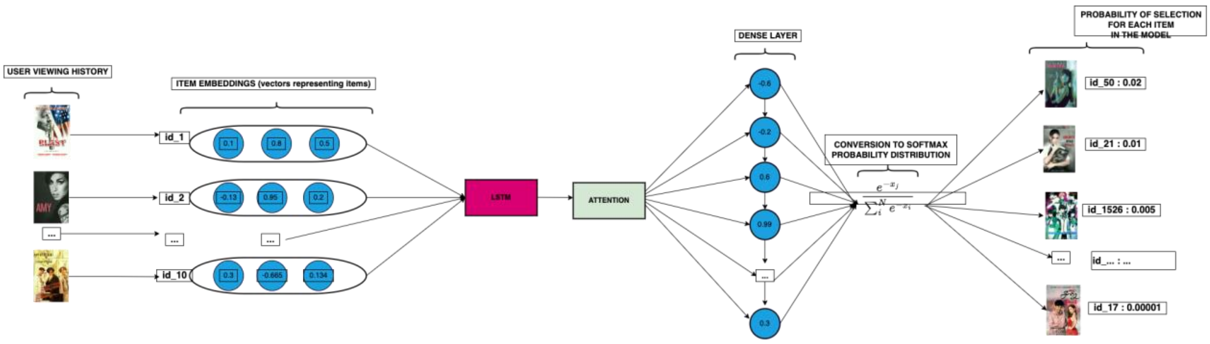

Figure 1 - proposed architecture for sequence modeling

The model is trained via gradient descent by minimizing cross-entropy loss.

Training a deep neural network can be difficult in terms of choosing right hyper-parameters, thus a bayesian hyper-parameters optimization framework is suggested for usage. Best hyper-parameters are found with respect to maximization of MAP score on validation subset of data.

General tips for two crucial steps of the machine learning pipeline related to sequence modeling approach in the recommendation domain, namely data preparation and modeling, are discussed. Such an approach grants better personalization due to modeling of user behavior, isn’t dependent on the number of users in the system and is closely related to the task of LM (language modeling), which allows to use recent advances of the NLP sphere with zero to little changes. The suggested architecture is extendable and can be augmented with additional features of items and users.

References

1. Kovenko V., Bogach I., Baraban M. ITEM-BASED COLLABORATIVE FILTERING BASED ON NLP TECHNIQUES, ITKI, vol. 51, no. 2, pp. 17–22, Oct. 2021.

2. He, X., Liao, L., Zhang, H., Nie, L., Hu, X., & Chua, T.-S. (2017). Neural collaborative filtering. Proceedings of the 26th international conference on world wide web. pp. 173–182.

3. Sedhain, S., Menon, A. K., Sanner, S., & Xie, L. (2015). Autorec: autoencoders meet collaborative filtering. Proceedings of the 24th International Conference on World Wide Web. pp. 111–112.

4. Hidasi B., Karatzoglou A., Baltrunas L., Tikk D. Session-based Recommendations with Recurrent Neural Networks , cite arxiv:1511.06939 Comment: Camera ready version (17th February, 2016) Affiliation update (29th March, 2016) .

УДК 004.08

Мазур Ю.О., Діденко М.М,
студентки 4 курсу
спеціальності 125
«Кібербезпека»
Нескородєва Т. В., д.т.н.,
доцент, завідувач кафедри
інформаційних технологій

РОЗРОБКА ІТ ПРОДУКТУ ДЛЯ ЛЮДЕЙ З ВАДАМИ СЛУХУ ТА МОВЛЕННЯ

Донецький національний університет імені Василя Стуса, м. Вінниця

За даними міністерства соціальної політики України близько для 200 тисяч українців жестова мова є рідною. Для понад 40 тисяч осіб жестова мова – це єдина можливість комунікувати та взаємодіяти між собою [1]. Мало хто з нас задумується над серйозністю даної проблеми, та тим, як людям з вадами важко робити різні повсякденні речі, про які ми навіть не задумуємось: завести нового друга, запитати дорогу, попросити певний товар в магазині. Кожного дня люди з вадами слуху та мовлення стикаються кожного дня з однією серйозною проблемою –комунікуація з оточуючими. Зовсім малий відсоток людей, які не мають даної проблеми, знають можу жестів і можуть нею спілкуватись. На ринку ІТ продуктів існує чимало програм, що допомагають з перекладами слів на різні мови, але майже немає застосунків, що допоможе зрозуміти жестову мову. Аби допомогти людям з вадами слуху та голосу вільно спілкуватись з навколишнім середовище, було розроблено ідею створення ІТ продукту Deaf and Dumb, що у