Approaching Explainable Recommendations for Personalized Social Learning

The current stage of the educational platform ”WhoTeach”

Luca Marconi[1,2][0000−0002−0236−6159], Ricardo Anibal Matamoros Aragon[1,2][0000−0002−1957−2530], Italo Zoppis[1][0000−0001−7312−7123], Sara Manzoni[1][0000−0002−6406−536X], Giancarlo Mauri[1][0000−0003−3520−4022], and Francesco Epifania[2]

[1] Department of Computer science, University of Milano Bicocca, Milano, Italy
{l.marconi3,r.matamorosaragon}@campus.unimib.it
{italo.zoppis,sara.manzoni,giancarlo.mauri}@unimib.it
[2] Social Things srl, Milano, Italy
{luca.marconi,francesco.epifania,ricardo.matamoros}@socialthingum.com

Abstract. Learning and training processes are starting to be affected by the diffusion of Artificial Intelligence (AI) techniques and methods. AI can be variously exploited for supporting education, though especially deep learning (DL) models are normally suffering from some degree of opacity and lack of interpretability. Explainable AI (XAI) is aimed at creating a set of new AI techniques able to improve their output or decisions with more transparency and interpretability. Deep attentional mechanisms proved to be particularly effective for identifying relevant communities and relationships in any given input network that can be exploited with the aim of improving useful information to interpret the suggested decision process. In this paper we provide the first stages of our ongoing research project, aimed at significantly empowering the recommender system of the educational platform ”WhoTeach” by means of explainability, to help teachers or experts to create and manage high-quality courses for personalized learning.

The presented model is actually our first tentative to start to include explainability in the system. As shown, the model has strong potentialities to provide relevant recommendations. Moreover, it allows the possibility to implement effective techniques to completely reach explainability [3].

Keywords: Social Networks · WhoTeach · Social Recommendations · Graph Attention Networks.

---

1 Introduction

Nowadays, learning and training processes are starting to be affected by the diffusion of AI techniques and methods [1] [2]. However, in order to effectively and significantly improve education, researchers, teachers or experts need to exploit their full potential. In the educational field it could be particularly significant to understand the reasons behind models outcomes, especially when it comes to suggestions to create, manage or evaluate courses or didactic resources. In order to address this issue, explainable AI (XAI) could be crucial and determining in education, as in many other fields [6], being aimed at creating a set of new AI techniques able to make their own decisions more transparent and interpretable.

In this context, explainable AI in the field of Recommender Systems (XRS) is aimed at providing intuitive explanations for the suggestions and recommendations given by the algorithms [10]. Basically the community tries to address the problem of why certain recommendations are suggested by the applied models. At the same time, different attempts in current deep learning literature try to extend deep techniques to deal with social data, recommendations and explanations. The "attentional mechanism" was introduced for the first time in the deep learning community in order to allow the model to detect the most relevant information due to the attention weights [22] and has recently been successful for the resolution of a series of objectives [11].

Specifically, in literature explainable attentional models have been used in domains ranging from medical care [24] [25] to e-commerce and online purchase [26]. It is worth to notice that attention weights can have a role in starting to fostering explainability [27].

In this article, we provide the first stages of our ongoing research project, aimed at significantly empowering the RS of our educational platform "WhoTeach" [29] by the means of explainability. Specifically, we report our current positioning in the state of the art with the proposed model to extend the social engine of "WhoTeach" with a graph attentional mechanism aiming to provide social recommendations for the design of new didactic programs and courses. The presented model allows us to start to include explainability in the system.

We have started to define our positioning in the state-of-the-art according to three dimensions studied in the XAI literature [10] [16]: the model itself, the display style of the explanations we aim to provide and the social aspects of our potential XRS. The first and the second dimension considered come from the literature, while the third one is the result of the importance of the social feature and data in WT.

In particular:

– **Display style**: in order to optimize the user experience we are working to define the way explanations will be provided, as the different possibilities in literature show [14] [28].

- **XRS model**: as described, the present attentional model shows the potentialities to actually include explainability in the RS. Thus, in the next stages of the project we are going to both improve the present model and empirically evaluate other possible models, so as to integrate them and effectively include explainability in the RS.
- **Social dimension**: by the means of the social data in the platform from users (e.g. teachers, students, experts) we are going to perform further experimentation to assess the present situation and understand the way to empirically evaluate other models.

From the study of the state of the art, we then strive to inscribe our current work and its future steps in the XRS literature, so as to define our present positioning and prepare for future work and stages towards explainability.

2 WhoTeach

WhoTeach (WT) is a complete digital learning platform for supporting heterogenous learning ecosystems in their processes and activities, due to its numerous synchronous and asynchronous features and functionalities. WT is aimed at promoting the development of customized learning and training paths by aggregating and disseminating knowledge created and updated by experts. The platform is conceived as a Social Intelligent Learning Management System (SILMS) and it is structured around three components:

1. The Recommender System (RS), to help experts and teachers to quickly and effectively assemble high-quality contents into courses: thanks to an intelligent analysis of available material, it is aimed at suggesting teachers the best resources to include, in any format, according to teachers' needs or requirements.
2. The "Knowledge co-creation Social Platform", which is a technological infrastructure based on an integrated and highly interactive social network, endowed with many features to share information, thematic groups and discussion forums.
3. The content's repository where to upload contents from any course or training material, either proprietary or open. This serves as a basis for both the recommender system to elaborate materials and also users who want to create personalized courses.

3 Main Concepts and Definitions

A graph (annotated with $G = (V, E)$) is a theoretical object widely applied to model the complex set of relationships that typically characterize current networks. This object consists of a set of "entities" (vertices or nodes), V, and relationships between them, i.e. edges, E. In this paper, we use attributed graphs, i.e., particular graphs where each vertex $v \in V$ is labeled with a set of attribute

values. Moreover, given a vertex $v \in V$, we indicate with $\mathcal{N}(v) = \{u : \{v, u\} \in E\}$ the neighborhood of the vertex v.

Given a graph G, we use the corresponding adjacency matrix A to indicate whether two vertices v_i, v_j of G are connected by an edge, i.e., $(A)_{i,j} = 1$, if $\{v_i, v_j\} \in E$.

In order to summarize the relationships between vertices and capture relevant information in a graph, embedding (i.e., objects transformation to lower dimensional spaces) is typically applied [23]. This approach allows to use a rich set of analytical methods, offering to deep neural networks the capability of providing different levels of representation. Embedding can be performed at different level: for example, at the node level, at the graph level, or even through different mathematical strategies. Typically, the embedding is realized by fitting the (deep) network's parameters using standard gradient-based optimization. In particular, the following definitions can be useful [11].

Definition 1. *Given a graph $G = (V, E)$ with V as the set of vertices and E the set of edges, the objective of node embedding is to learn a function $f : V \to \mathcal{R}^k$ such that each vertex $i \in V$ is mapped to a k-dimensional vector, $\boldsymbol{h}$.*

Definition 2. *Given a set of graphs, $\mathcal{G}$, the objective of graph embedding is to learn a function $f : \mathcal{G} \to \mathcal{R}^k$ that maps an input graph $G \in \mathcal{G}$ to a low dimensional embedding vector, $\boldsymbol{h}$.*

4   GAT models

In our application, we use the attentional-based node embedding proposed in [12]. For a general definition of the notion of "attention", here we conveniently adapt the one reported in [11].

Definition 3. *Let A be an user/item relationship matrix, $G[A] = (V, E)$ the corresponding weighted graph, and $V = \{U, R\}$ the set of users U and items R, respectively. Given a pair of vertices $(u, r), u \in U, r \in R$, an attentional mechanism for G is a function $a : \mathcal{R}^n \times \mathcal{R}^n \to \mathcal{R}$ which computes coefficients $e_{u,r} = a\big(\boldsymbol{h}_u^{(l)}, \boldsymbol{h}_r^{(l)}\big)$ across the pairs of vertices, u, r, based on their feature representation $\boldsymbol{h}_u^{(l)}, \boldsymbol{h}_r^{(l)}$ at level l.*

Coefficients $e_{u,r}$ are considered as the importance of the vertex r's features to (user) u.

Following [12], we define a as a feed-forward neural network with a learnable (weight) vector of parameters $\boldsymbol{a}$ and nonlinear *LeakyReLU* activation function. In this way, we have

$$e_{u,r}^{(l)} = LeakyReLU\Big(\boldsymbol{a}^{(l)^T}\left[\mathbf{W}^{(l)}\boldsymbol{h}_u^{(l)}||\mathbf{W}^{(l)}\boldsymbol{h}_r^{(l)}\right]\Big). \tag{1}$$

where $\mathbf{W}$ is a learnable parameter matrix and $\mathbf{W}^{(l)}\boldsymbol{h}_u^{(l)}||\mathbf{W}^{(l)}\boldsymbol{h}_r^{(l)}$ is the concatenation of the embedded representation for the vertices u, r.

The coefficients $e_{u,r}$ can be normalized using, e.g., the *softmax* function

$$\alpha_{u,r}^{(l)} = \frac{\exp(e_{u,r}^{(l)})}{\sum_{k \in \mathcal{N}(u)} \exp(e_{u,k}^{(l)})}.$$

The mechanisms parameters, $\boldsymbol{a}$, are then updated with the others network's parameters accordingly to typical optimization algorithms. When only resources (items) around u are considered, the normalized (attention) coefficients $\alpha_{u,r}$ can be used to compute a combination of the resources $\boldsymbol{h}_r^{(l)}$ in $\mathcal{N}(u)$ as follows

$$\boldsymbol{h}_u^{(l+1)} = \sigma\Big(\sum_{r \in \mathcal{N}(u), r \in R} \alpha_{u,r}^{(l)} \mathbf{W}^{(l)} \boldsymbol{h}_r^{(l)} \Big) \tag{2}$$

where σ is non linear vector-valued function (sigmoid). With this formulation, Eq. 2 provides the next level embedding for user u scaled by the attention scores which, in turn, can be interpreted as the relevance of the resources used by the user u. Similarly to Eq. 2, the following quantity can be interpreted as the user scores who applied, in particular, the resource r.

$$\boldsymbol{h}_r^{(l+1)} = \sigma\Big(\sum_{u \in \mathcal{N}(r), u \in U} \alpha_{u,r}^{(l)} \mathbf{W}^{(l)} \boldsymbol{h}_u^{(l)} \Big) \tag{3}$$

In this way, the "GAT layer" returns for each pair $(u, r) \in U \times R$ the embedded representation $(\boldsymbol{h}_u^{(l+1)}, \boldsymbol{h}_r^{(l+1)})$. In our experiments we will consider only one level of embedding, i.e., $l = 1$.

Therefore, as previously described in the section 3, we introduce a novel kind of information representation $\boldsymbol{h}_u^{(l)}$ for users and $\boldsymbol{h}_r^{(l)}$ resources, allowing us to visualize either the user u or the resource r as the main element according to its neighborhood. Nevertheless, this representation is still not able to explain and justify the recommendations given to a specific user. Indeed, it provides the starting point to apply the attention mechanism, which introduces the possibility to give a weight $e_{u,r}^{(l)}$ to the most relevant information encoded in the embedded representation for both the user $\boldsymbol{h}_u^{(l+1)}$ and the resource $\boldsymbol{h}_r^{(l+1)}$. Then, the attention weights $e_{u,r}^{(l)}$ permit to improve the model performances, reducing the error for the recommendations.
Above all, they foster the possibility to explain why a given resource r is recommended to a specific user u. In particular, this approach for computing the attention weights $e_{u,r}^{(l)}$ is also applied in other works related to collaborative filtering RS [32]. Other works explore different display styles, as visual explanations [31]. In conclusion, the ability to highlight the most useful information to realize the recommendations allowed us to start to introduce explainability in the system.

5   Numerical experiments

Here we report a short review of the numerical experiments described in [29].
The experiments use an homogeneous set of data whose characteristics combine well with the requirements of the WhoTeach platform. These data come from the "Goodbooks" data-set (https://www.kaggle.com/zygmunt/goodbooks-10k), a large collection reporting up to 10K books and 1000000 ratings (from "1" to "5") assigned by 53400 readers.
The experiments aim to evaluate the capability of the attentional-based models to reduce error (loss function) between the reported and predicted preference scores.

The models was implemented using the Pytorch library (https://pytorch.org/), and then executed using different hyperparameters.At the present stage, the attention-based model was compared with alternative models: dot product model, element-wise product model (Hadamard product model), concatenation model. Performances were averaged on the number of folds (10 cross-validation).

A general better tendency to reduce the MSE loss is observed when attention layer with concatenation is applied as a base module for the considered stacked layer.

6   Conclusions and future work

In this work we have reported our work in progress for providing "WhoTeach" with an explainable recommender system, aimed to significantly empower its ability to help teachers or experts to create high-quality courses. It is totally clear that further improvements of the present XRS could significantly help users to better understand the reason why specific items are recommended.
At the present stage, we started to propose a model based on the attentional mechanisms, which allows to justify the chosen recommendations provided by the model by the means of the attention weights. This model is specifically focused on exploiting social information for educational services, thus extending the social engine of our educational platform "WhoTeach" to reinforce the AI engine. Finally, we have reported our present and further positioning in the state of the art of XAI, showing the potentialities of the present model and the next steps in our work. One of the most important further step will be the optimization of the computational complexity for the computation of the attention weights. Moreover, we will strive to improve the display style of the explanations provided to users, in order to consequently improve the user experience.

References

1. Holmes, R., Wayne, B., Author, M., Fadel, A. C.: Artificial intelligence in education. journal, Boston: Center for Curriculum Redesign, Location (2019). https://doi.org/10.1109/MIS.2016.93
2. Timms, M. J. : Letting artificial intelligence in education out of the box: educational cobots and smart classrooms. In: International Journal of Artificial Intelligence in Education (2016).https://doi.org/10.1007/s40593-016-0095-y
3. Dondi, R., Mauri, G. Zoppis, I. : Clique editing to support case versus control discrimination. In : Intelligent Decision Technologie (2016). https://doi.org/10.1016/j.tcs.2019.09.022
4. Zoppis, I. Dondi, R., Coppetti, D., Beltramo, A., Mauri, G. : Distributed Heuristics for Optimizing Cohesive Groups: A Support for Clinical Patient Engagement in Social Network Analysis. In : 2018 26th Euromicro International Conference on Parallel, Distributed and Network-based Processing (PDP) (2018). https://doi.org/10.1109/PDP44162.2018
5. Zoppis, I., Manzoni, S., Mauri, G. : A Computational Model for Promoting Targeted Communication and Supplying Social Explainable Recommendations. In : 2019 IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS) 2019. https://doi.org/10.1109/CBMS.2019.00090
6. Fox, M., Long, D., Magazzeni, D. : Explainable Planning (2017).
7. Bonhard, P., Sasse, M. A. : Knowing me, knowing you'—Using profiles and social networking to improve recommender systems. In : BT Technology Journal (2006). https://doi.org/10.1007/s10550-006-0080-3
8. Gupta, P., Goel, A., Lin, J., Sharma, A., Wang, D., Zadeh, R. : The who to follow service at twitter. In: Proceedings of the 22nd international conference on World Wide Web (2013). https://doi.org/10.1145/2488388.2488433
9. Zhou, X., Xu, Y., Li, Y., Josang, A., Cox, C. : The state-of-the-art in personalized recommender systems for social networking. In : Artificial Intelligence Review (2012). https://doi.org/10.1007/s10462-011-9222-1
10. Yongfeng, Z., Xu, C., : Explainable Recommendation: A Survey and New Perspectives (2018). https://doi.org/10.1561/1500000066
11. Lee, J., Boaz, R., Ryan, A., Kim, S., Ahmed, N., K., Koh, E. : Attention models in graphs: A survey. In : arXiv preprint arXiv:1807.07984 (2018). https://doi.org/10.1145/3363574
12. Velickovic, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., Bengio, Y. : Graph attention networks. In : arXiv preprint arXiv:1710.10903 (2017). https://doi.org/10.17863/CAM.48429
13. Sharma, A., Cosley, D. : Do Social Explanations Work?: Studying and Modeling the Effects of Social Explanations in Recommender Systems. In : Proceedings of the 22Nd International Conference on World Wide Web (2013). https://doi.org/10.1145/2488388.2488487
14. Pedreschi, D., Giannotti, F., Guidotti, R., Monreale, A., Ruggieri, S., Turini, F. : Meaningful explanations of Black Box AI decision systems. In: Proceedings of the AAAI Conference on Artificial Intelligence (2019). https://doi.org/10.1609/aaai.v33i01.33018001
15. Adadi, A., Berrada, M. : Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). In: IEEE Access (2018). https://doi.org/10.1109/ACCESS.2018.2870052

16. Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., Pedreschi, D. : A Survey of Methods for Explaining Black Box Models. In : ACM Comput. Surv. (2018). https://doi.org/10.1145/3236009

17. Apolloni, B. Bassis, S. Mesiti, M., Valtolina, S., Epifania, F. : A Rule Based Recommender System. In: Advances in Neural Networks (2016). https://doi.org/10.1007/9783319337470_9

18. Haekyu, P., Hyunsik, J., Junghwan, K., Beunguk, A., U, K.: UniWalk: Explainable and Accurate Recommendation for Rating and Network Data (2017).

19. Gori, M., Monfardini, G., Scarselli, F. : A new model for learning in graph domains. In: Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005. https://doi.org/10.1109/IJCNN.2005.1555942

20. Zoppis, I., Dondi, R., Manzoni, S., Mauri, G., Mar-coni, L., and Epifania, F. : Optimized social explanation for educational platforms (2019). https://doi.org/10.5220/0007749500850091

21. Scarselli, F., Gori, M., Tsoi, A. C., Hagenbuchner, M., Monfardini, G. : The graph neural network model (2008). https://doi.org/10.1109/TNN.2008.2005605

22. Bahdanau, D., Cho, K., Bengio. Y.: Neural machine translation by jointly learning to align and translate. In: arXiv preprint arXiv:1409.0473 (2014).

23. Goyal, P., Ferrara, E.: Graph embedding techniques, applications, and performance: A survey. In: Knowledge-Based Systems (2018). https://doi.org/10.1016/j.knosys.2018.03.022

24. Mullenbach, J., Wiegreffe, S., Duke, J., Sun, J., Eisenstein, J. : Explainable Prediction of Medical Codes from Clinical Text (2018).

25. Wang, N., Mingxuan, C., Subbalakshmi, K. P.: Explainable CNN-attention Networks (C-Attention Network) for Automated Detection of Alzheimer's Disease (2020).

26. Chen, C., Zhang, M., Liu, Y., Ma, S.: Neural Attentional Rating Regression with Review-Level Explanations, in : International World Wide Web Conferences Steering Committee (2018). https://doi.org/10.1145/3178876.3186070

27. Akash Kumar, M., Nema, P., Narasimhan, S., Mitesh M., Vasan Srinivasan, B., Ravindran, B.: Towards Transparent and Explainable Attention Models (2020).

28. Liu, P., Zhang, L., Atle Gulla, J.: Dynamic attention-based explainable recommendation with textual and visual fusion. In: Information Processing and Management (2019). https://doi.org/10.1016/j.ipm.2019.102099

29. Zoppis, I., Manzoni, S. Mauri, G., Matamoros, R., Marconi, L., Epifania, F.: Attentional Neural Mechanisms for Social Recommendations in Educational Platforms. In : Proceedings of the 12th International Conference on Computer Supported Education - Volume 1 CSEDU, (2020). https://doi.org/10.5220/0009568901110117

30. Dondi, R., Mauri, G., Zoppis, I.: On the tractability of finding disjoint clubs in a network. In: Theoretical Computer Science (2019).

31. Chen, X., Chen, H., Xu, H., Zhang, Y., Cao, Y., Q, Z., Zha, H. In: Personalized Fashion Recommendation with Visual Explanations Based on Multimodal Attention Network: Towards Visually Explainable Recommendation (2019). https://doi.org/10.1145/3331184.3331254

32. Chen, J., Zhuang, F., Hong, X., Ao, X., Xie, X., He, Q. In: Attention-Driven Factor Model for Explainable Personalized Recommendation (2018). https://doi.org/10.1145/3209978.3210083