A Formal Concept Analysis-based Method and A Graph Visualisation Approach to Enrich Explanations-by-examples

Marta Caro-Martínez, José L. Jorro-Aragoneses, Belén Díaz-Agudo, Juan A. Recio-García

***Abstract*—As it happens in many artificial intelligence models nowadays, complex recommender systems algorithms are considered black-boxes, so they include methods to explain to users the reasons for a recommendation. Moreover, many times, these recommendations are explanations-by-examples, due to the suitability of these types of explanations when recommending new items. Though, users may not understand these types of explanations since they lack of additional information that clarify the reasons behind the recommendation. In this paper, we propose a novel post-hoc method to generate explanations in black-box recommender systems enriching the aforementioned explanations-by-examples.**

***Index Terms*—XAI, Explanations-by-examples, FCA, Graph-based Visualisation**

I. Introduction

Recommender systems emerged to assist users in discovering the most suitable products, particularly when faced with a vast array of options and uncertainty about which one to select. There are many different strategies for implementing recommendation systems varying in their algorithms, complexity, effectiveness, interpretability, and suitability for different types of data and user preferences. In recent years there has been an important improvement in machine learning algorithms, also applied to recommender systems. Despite the increasing precision of these recommendation algorithms, the systems that use them are considered black-box approaches: these approaches are not interpretable by final users [8].

Recently, an active research area to resolve the interpretability and explainability problem is eXplainable Artificial Intelligence (XAI). XAI provides different methods to increase human understanding of the results provided by systems based on artificial intelligence techniques [1]. In the context of recommender systems, explanations improve the understanding of the user who receives a recommendation. XAI methods can increase the utility of recommender systems because they can improve users' satisfaction or users' confidence, which also enhances the user experience [3].

There are already different approaches for generating explanations for a recommender system. One of the most commonly used is the explanation-by-example methodology. Explanations using examples are easy to understand for a user because they are more intuitive than other common explanations like trace-based explanation. However, in some cases, the examples are not enough for users to understand a recommendation and they could need an extra component like a connection between the recommendation and the examples [7].

In this paper, we propose a methodology to enrich explanations-by-examples for black-box recommender systems. Commonly, explanations-by-examples only offer the examples as a way to explain a recommendation. Our approach enriches those examples using complementary knowledge to connect them with the recommendation and generate a final explanation. We use Formal Concept Analysis (FCA) [2] to get the most useful links between the recommended item and all explanation examples. Moreover, we use a graph-based visualisation to show the explanation to the final users. Our visualisation method creates an interactive experience for the user, who can watch the recommendation together with the examples and their links in a more appealing fashion. This work is a summary of a paper submitted and accepted to the 2nd World Conference on eXplainable Artificial Intelligence, for more details, see the paper to be published [4].

II. Linking items and attributes to selected explanation examples

Our methodology's starting point is the set of examples that are personalised to explain a recommendation for a target user. The issue here is that, at first sight, the target user might not understand the common features between the recommendation and the examples. Examples, on their own, may not be totally useful to understand a recommendation [5]. Moreover, not all the examples will be interesting for users if those do not share interesting attributes for the user with the recommended item. Then, one approach to clarify why the examples are useful to understand the recommendations is to add what the shared attributes between them and the recommendations are considering only the important features for the target user. Therefore, the main objective of this proposed approach is to get the examples and the shared attributes that are useful to explain the recommendation. In our proposed methodology, we solve this task using FCA.

FCA [2] formalises concepts as basic units of human thinking by analysing data in an object-attribute form. With FCA, we build a conceptual lattice relating our recommended item, our explanation examples, and their common attributes. The lattice contains internal nodes that represent the attributes,

Marta Caro-Martínez, Belén Díaz-Agudo and Juan A. Recio García are with Facultad de Informática, Universidad Complutense de Madrid, Spain. E-mails: {martcaro,belend,jareciog}@ucm.es

José L. Jorro-Aragoneses is with EPS Politécnica, Universidad Autónoma de Madrid, Spain. E-mail: jose.jorro@uam.es

[1*] Corresponding author: Marta Caro-Martínez.

Partially supported by Spanish Thematic Network on Recommender Systems (Action RED2022-134302-T funded by MCIN/AEI/10.13039/501100011033)

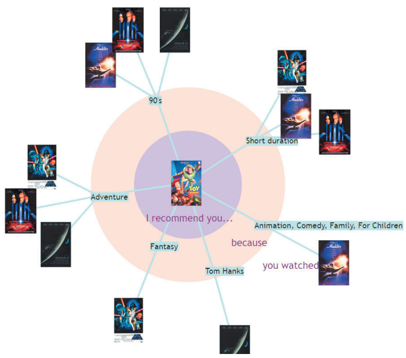

Fig. 1. An example of the proposed visualisation approach.

while the last level of the lattice represents the items. The parents of the leaf are the attributes of that item. Then, we can classify the recommended item in the lattice to infer the attributes shared by that item and the examples obtained in the previous step. Next, we can traverse the lattice from the top to the bottom to find the attributes that relate the items. However, some of these shared attributes may be too obvious for the target user. We have also looked into this problem [6]. Users prefer explanations where there are as many explanatory items as possible, especially if the lattice is small. Besides, users prefer explanations where the attributes are the most specific ones, i.e., the ones in the deepest lattice levels. Following these findings, from the initial set of explanation examples, we select only the example items that share direct parent nodes with the recommended item.

III. Visualising explanations through interactive graphs

Once we select the explanation examples, and we get the helpful shared attributes between them and the recommendation, we need to show all this information to the target user. We wanted to retain all the knowledge provided by the lattice, but displaying the explanation through an interesting and useful interface that allows users to interact with the explanation. After an object-guided iterative process to design our interface, we came up with an interactive graph-based visualisation.

An example of our visualisation is presented in Figure 1. The explanations are represented by a graph with three concentric circular areas. From the inside out, it displays the recommended item, the shared attributes and the explanatory examples according to the substructure selected from the lattice. A key feature of our visualisation approach is the interaction that allows the user to adapt the explanation to her own thoughts and preferences, so she can understand why a recommendation is suitable (or not) for her being satisfied during the process. A user can remove the attributes that are not relevant according to her beliefs about the explanation. The explanation system also includes functionality for undoing this action if the user is not happy with the resulting visual explanation.

IV. Conclusions

In this research work, we propose a new methodology to enrich explanations-by-examples in black-box recommender systems. These types of explanations are very well-known in the literature and widely used in many applications and platforms. However, those examples may not be as useful as they seem, since users may not remember those items or do not see the similarities between them and the recommendations. Then, these examples could not be enough for understanding the recommendations completely. To complement the explanation examples, we used Formal Concept Analysis to select the best examples for each user and to get the remarkable attributes that the recommendations and the examples share. To show to the users all these explanations, we employed a graph-based visualisation that allows users to interact with the explanation.

References

[1] Barredo Arrieta, A., et al.: Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible ai. Information Fusion **58**, 82 – 115 (2020)

[2] Birkhoff, G.: Lattice theory, vol. 25. American Mathematical Soc., USA (1940)

[3] Caro-Martínez, M., Jiménez-Díaz, G., Recio-García, J.A.: Conceptual modeling of explainable recommender systems: an ontological formalization to guide their design and development. Journal of Artificial Intelligence Research **71**, 557–589 (2021)

[4] Caro-Martinez, M., Jorro-Aragoneses, J., Diaz-Agudo, B., Recio-Garcia, J.A.: Graph-based interface for explanations by examples in recommender systems: A user study (in press). In: World Conference on Explainable Artificial Intelligence. Springer (2024)

[5] Jorro-Aragoneses, J., Caro-Martinez, M., Recio-Garcia, J.A., Diaz-Agudo, B., Jimenez-Diaz, G.: Personalized Case-Based Explanation of Matrix Factorization Recommendations. In: Case-Based Reasoning Research and Development: 27th International Conference, ICCBR 2019, Otzenhausen, Germany, September 8–12, 2019, Proceedings 27. pp. 140–154. Springer International Publishing, Germany (2019)

[6] Jorro-Aragoneses, J.L., Caro-Martínez, M., Díaz-Agudo, B., Recio-García, J.A.: A user-centric evaluation to generate case-based explanations using formal concept analysis. In: Case-Based Reasoning Research and Development: 28th International Conference, ICCBR 2020, Salamanca, Spain, June 8–12, 2020, Proceedings. pp. 195–210. Springer International Publishing, Spain (2020)

[7] Kenny, E.M., Ford, C., Quinn, M., Keane, M.T.: Explaining black-box classifiers using post-hoc explanations-by-example: The effect of explanations and error-rates in xai user studies. Artificial Intelligence **294**, 103459 (2021)

[8] Portugal, I., Alencar, P., Cowan, D.: The use of machine learning algorithms in recommender systems: A systematic review. Expert Systems with Applications **97**, 205–227 (2018)