A SURVEY OF EXPLAINABLE RECOMMENDER SYSTEMS

Alfrin Saldanha and Hyoil Han

School of Information Technology, Illinois State University, Normal, USA

ABSTRACT

The rapid advancement of Artificial Intelligence has driven the adoption of machine learning technologies across diverse domains, with recommender systems playing a pivotal role in delivering personalized suggestions. However, as user-centric applications become increasingly sophisticated, providing recommendations without clear explanations is no longer adequate. Explainable recommendation systems bridge this gap by enhancing transparency, user understanding, and trust through interpretable and contextually relevant explanations. These systems strive to balance high recommendation accuracy with the clarity of their explanations. This paper examines state-of-the-art models and methodologies in explainable recommendation systems, focusing on their computational underpinnings, evaluation metrics, and practical outcomes. We analyze the strengths and limitations of existing approaches and discuss opportunities for integrating innovative techniques and emerging technologies. Our study aims to advance the development of more effective, explainable recommendation systems adaptable to diverse application domains, aligning with the interdisciplinary focus of computational science.

KEYWORDS

Explainable Recommender Systems, Artificial Intelligence,Knowledge Mining, Virus, Machine Learning

1. INTRODUCTION

Over the past decade, technology has transformed significantly, driven by the widespread availability of affordable internet. This global connectivity has brought users and systems closer together, fueling advancements in e-commerce, entertainment, social media, medicine, and transportation. The resulting surge in information about products, users, and systems has increased reliance on these technologies, making our lives more convenient. Among these innovations, recommender systems have emerged as a solution to the growing demand for personalization and convenience.

Advancements in algorithms have expanded the application of recommender systems across various domains, leveraging data from users, systems, and networks to suggest new items. By analyzing user preferences, these systems help users navigate vast choices to make better decisions. Despite their benefits, recommender systems face a critical challenge: the lack of explainability. They have evolved beyond merely providing recommendations to supporting users

David C. Wyld et al. (Eds): CSITAI, MATHCS, ITCAU, COMSAP, ACINT - 2024
pp. 159-171, 2024. - CS & IT - CSCP 2024 DOI: 10.5121/csit.2024.142512

in decision-making, but understanding how these systems generate explanations is key to building trust.

Like traditional recommender systems, explainable ones rely on diverse input features for their underlying models or algorithms. With the rise in online activity, abundant data can now be collected from sources such as reviews, comments, tags, social networks, user profiles, images, and interactions. The first step involves gathering this input data, followed by extracting the relevant aspects or features required for the system's specific use cases.

This paper explores various techniques employed by existing systems to extract features essential for generating explainable recommendations. While the extracted features may often resemble those used in traditional approaches, their selection largely depends on the desired level of interpretability and the specific application. Ensuring efficient interpretability is critical for modern recommendation systems, as it plays a key role in capturing and enhancing user trust. To achieve this, it is vital to understand how recommendations can be effectively translated into explanations. Notably, users tend to trust recommendations more when they are communicated in natural language, highlighting the importance of providing clear and human-like explanations.

However, achieving natural language explanations often involves balancing a trade-off between prediction accuracy and explanation interpretability. It is crucial to identify what best meets the needs of the target audience and to strike an optimal balance between these two objectives. Given users' reliance on recommendations, it is imperative that these systems be transparent, interpretable, trustworthy, and explainable. The applications of explainable recommendation systems are expanding beyond the e-commerce domain. For instance, in the medical field, the accuracy of generated explanations is essential for supporting healthcare professionals in making informed decisions. When system-generated recommendations influence critical decisions, such as a doctor's treatment plan, the system must provide clear reasoning to justify its suggestions.

The subsections are organized as follows: Section 2 provides an overview of the background of explainable recommender systems. Section 3 outlines the classification of various approaches to explainable recommender systems, with detailed explanations of each category in the corresponding subsections. Finally, the conclusion is presented.

2. BACKGROUND

The definition of a good explanation remains subjective and depends largely on the specific goals of the recommender system. Tintarev et al. [1] examine various types of explanations for recommendations, emphasizing their role in enhancing user trust, loyalty, and the likelihood of persuading users to purchase an item. They argue that accuracy metrics alone are insufficient to evaluate the effectiveness of recommender systems. Instead, they identify seven key properties that explanations should possess: Transparency, Scrutability, Trustworthiness, Effectiveness, Persuasiveness, Efficiency, and Satisfaction. Tintarev et al. [1] also highlight the importance of how explanations are presented, asserting that their design and display significantly influence their impact. By employing a user-centered design approach, they identify essential features of effective explanations and emphasize that measuring their persuasiveness in influencing user decisions is a strong indicator of their efficiency.

Explanations within recommender systems serve to clarify generated recommendations, thereby improving user satisfaction, trust, and loyalty. In a subsequent study, Tintarev et al. [2] explore strategies for designing and evaluating recommendation systems, providing specific guidelines to enhance explainability. These guidelines outline various criteria and metrics for assessing the efficiency of recommendations and their accompanying explanations. The study also reviews diverse methods for presenting explanations in ways that align with user preferences and needs. Furthermore, Tintarev et al. [2] categorize explanation styles tailored to different systems and reaffirm the importance of the seven essential criteria: Transparency, Scrutability, Trust, Effectiveness, Persuasiveness, Efficiency, and Satisfaction.

According to Tintarev et al. [3], explanations play a crucial role in enhancing user experience, building trust, and persuading users to purchase more items. They emphasize identifying factors that can assist users in making better decisions. Research using the MovieLens dataset highlights potential biases in experiments where positive ratings displayed as histograms may lead to overestimated results. Additionally, the generated explanations often tend to be more persuasive than genuinely effective. Tintarev et al. [3] also note that explanations based on movie features, such as actors, are not always personalized, as the significance of these features can vary among users. This variability makes it challenging to tailor explanations to individual preferences. To address this, they leverage item features in recommendations to help users better understand the relationships between items. By presenting personalized information, Tintarev et al. [3] aim to support users in making more informed decisions.

Recommender systems play a significant role in enhancing movie recommendations by improving quality, which, in turn, builds user trust and provides greater value. Nanou et al. [4] emphasize the importance of presentation in shaping users' perceptions of a recommender system's quality. Their study explores how different recommendation modalities influence system persuasiveness and user satisfaction. To investigate these factors, they compare various movie recommendation interfaces, beginning with a preliminary survey to identify a diverse demographic sample for their study. After defining the target group, they conduct experiments to analyze how recommendation modalities and the organization of information impact user trust and persuasion.

Nanou et al. [4] use the movieSTAR prototype to assess multiple experimental setups. In the first experiment, they evaluate two recommendation interfaces: a top-N item list and the same recommendations organized by genre. The second experiment compares three interfaces: text-only information in the top-N list, a combination of text and images for the same recommendations, and text with video trailer support. These experiments provide insights into how presentation formats affect user engagement and trust in movie recommendation systems.

Explanations are vital to recommender systems, as they demonstrate the system's efficiency and enhance user trust. While recommender systems should aim to meet these criteria, achieving a balance often involves trade-offs. Al et al. [5] highlight the significant role that the format and visualization of explanations play in influencing decision-making, which is the process of selecting an optimal action from multiple alternatives to achieve a specific goal.

Human-Computer Interaction (HCI) underscores the interdependence between humans and computers, emphasizing the importance of explanation interfaces. These interfaces consist of three key components: explanation, presentation, and interaction, which collectively help users understand why a specific recommendation was made. Visual representations, such as graphs, are particularly effective, as humans tend to process visual information more efficiently than textual data. However, while incorporating visual techniques can enhance user understanding, designers must be cautious not to overload the presentation with excessive information. Too much detail can lead to clutter, diminishing the effectiveness of the visualization and complicating the user's decision-making process.

3. Classification Based on Techniques Employed in Explainable Recommender Systems

This section provides a classification of explainable recommender systems according to the techniques they employ. We categorized these systems into nine distinct approaches: attention-based, GRU-based, rule-based, template-based, path-based, factorization-based, LIME-based, aspect-based, and hybrid-based. Each category is defined by the specific techniques utilized to generate explanations, offering a structured framework for understanding the diverse methodologies applied in explainable recommendation systems.

3.1. Attention-Based Approach

The attention mechanism has become a widely adopted concept in machine learning, particularly for its ability to enhance interpretability in recommender systems. As the demand for explainable recommendations grows, attention mechanisms have been increasingly utilized to make recommendations more transparent. Wang et al. [6] employ a gradient-boosting decision tree (GDBT) technique to extract cross features, which are then used to generate explanations for users. Additionally, Wang et al. [5] leverage an attention mechanism to assign personalized weights to cross features, identifying and selecting the most important ones to serve as explanations.

Barkan et al. [7] introduce a novel approach called Attentive Multi-Persona Collaborative Filtering (AMPCF), which learns user personas and uses them as a basis for providing explanations in recommendation lists. This method dynamically adjusts the weight of each persona to create user representations and generate interpretable recommendations based on their importance. Similarly, Chen et al. [8] focus on capturing visual preferences, combining these with reviews and regional features through an attention mechanism. This technique assigns attention weights to highlight specific regions in images, serving as explanations for recommendations. Seo et al. [9] further refine this approach by combining global and local attention to better capture user-item properties and semantic meanings, creating a more interpretable representation of users and items.

Chen et al. [10] emphasize the value of reviews in helping users make informed decisions and propose a neural attention mechanism to improve the explainability and performance of recommender systems. This mechanism assigns weights to individual reviews, prioritizing the most informative ones to enhance recommendation transparency. Wang et al. [11] take this further by implementing agents to generate sentence-level explanations through attention-based selection and GRU models, boosting the efficiency of these explanations. Dong et al. [12] design the Asymmetrical Hierarchical Network with Attentive Interactions (AHN) framework, which learns attention weights at both the review and sentence levels, further enriching interpretability.

High-ranking attention weights provide valuable insights that enhance the clarity and interpretability of recommendations. Chen et al. [13] use a personalized attention mechanism to assign weights, determine sentence importance, and explain the rating process. Yu et al. [14] generate explanations for recommendations made by either item-based or collaborative filtering systems, utilizing weighted members to create explanations for a group of items or users, improving the overall transparency of the recommendation process.

3.2. GRU-Based Approach

Gated Recurrent Units (GRUs) are a type of Recurrent Neural Network (RNN) that utilize a gating mechanism and require less memory compared to Long Short-Term Memory (LSTM) networks. Over the years, the use of GRUs in recommender systems has gained significant traction. Chen et al. [15] apply GRUs to generate word sequences that serve as textual explanations for recommendations. The GRU maps user and item latent factors, learns the hidden states, and produces word sequences that explain the recommendations.

Lin et al. [16] utilize GRUs to generate comments from visual features. By integrating a cross-modality attention mechanism, the GRU processes hidden states and converts visual features into meaningful text for comment generation. Similarly, Lu et al. [18] leverage GRUs to generate explanations from user and item textual features. Textual feature vectors are input into the GRU, which then generates word tokens to provide explanations.

Zhao et al. [19] implement GRUs to calculate word sequence probabilities and generate natural language explanations for song recommendations. They enhance this process by using N-gram models to identify optimal word combinations and refine sentence grammar through part-of-speech (POS) tagging. Li et al. [20] take a different approach, employing GRUs within their proposed modeling framework to produce abstractive tips from latent factors. GRUs generate word sequences for these tips, and the beam search algorithm selects the best sequences with the highest log-likelihood for use as explanations.

3.3. Rule-Based Approach

Rule-based methods leverage logical associations to generate explanations. Peake et al. [21] utilize approximate matrix factorization in combination with global association rules to produce explainable recommendations. The rules corresponding to the filtered Top-N recommendations serve as the basis for generating explanations for the matrix factorization outputs. Tsukuda et al. [22] employ the HyPER [23] framework, which leverages probabilistic soft logic (PSL) to recommend items based on predefined rules. The HyPER framework learns these rules and generates explanations tailored to various styles influenced by personal, social, and item-specific factors.

Samih et al. [24] model user-item relationships through extraction rules developed using the PSL framework. Natural language explanations are created by aligning these rules with three distinct explanation styles. Ma et al. [25] propose an approach where rule features are used to generate explanations; items are assigned weights based on associated rules, and the rule with the highest score is selected as the explanation for the recommendation. Kouki et al. [26] develop a system that generates natural language explanations by using ground rules as input for a translation system, which outputs human-readable sentences. Similarly, Tsukuda et al. [27] implement the HyPER framework, using PSL and template-based rules to provide interpretable recommendations.

3.4. Template-Based Approach

Template-based systems use predefined structures to generate explanations. By combining feature-opinion pairs or employing case-based reasoning, these approaches provide clear, structured insights into recommendations. They are particularly effective in presenting user-specific or item-specific details in natural language.Zhang et al. [28] leverage explicit features to create personalized recommendations by designing both template-based and word cloud-based explanations. Matched feature-opinion pairs are combined with predefined templates to generate clear and user-friendly explanations for the recommended items. Balog et al. [29] focus on providing explanations rooted in user preferences rather than specific item recommendations. Their approach uses templates to produce textual representations, which are presented to users in natural language.

McSherry et al. [30] propose a dynamic approach to generating explanations using templates. Their top-case method generates explanations by matching user queries with relevant case attributes. Similarly, Tran et al. [31] employ templates to structure explanations, designing them around various strategies such as social choice-based preference aggregation, preference aggregation combined with decision history, and preference aggregation integrated with future decision planning.

3.5. Path-Based Approach

Path-based systems play a crucial role in enhancing the interpretability of explainable recommendation systems. The relationship between users and items contains rich semantic information that can be extracted through paths. Various researchers have explored different methodologies to leverage this extracted information and incorporate it into their models to improve explanation quality.

Huang et al. [32] develop a sequential recommender system that accounts for users' dynamic interests to provide accurate and interpretable recommendations. They extract semantic paths between user-item pairs from a knowledge graph and utilize these paths to generate path-level explanations. Fu et al. [34] focus on capturing user-item paths and employ a fairness-aware path re-ranking method to produce explainable recommendations. Similarly, Xian et al. [35] train an agent to identify reasoning paths that not only recommend items but also serve as explanations. The extracted information from their path-searching method enhances the interpretability of recommendations.

Zhang et al. [33] generate explanations for recommended items by identifying the factors that most significantly contribute to meta-path generation. To refine these explanations, they use a beam search algorithm on these factors to identify specific paths and provide detailed, meaningful explanations

3.6. Factorization-Based Approach

Factorization methods utilize latent representations to model user preferences and explanations simultaneously. Techniques such as matrix factorization and tensor decomposition are employed to integrate features and produce interpretable predictions. Abdollahi et al. [36] propose neighbor-style explanations, which can be either user-based or item-based. They represent explainability through a bipartite graph and use a matrix factorization (MF) approach to jointly learn latent vectors and generate explanations.

Liu et al. [37] develop interpretable recommendations using a probabilistic factorization model, analyzing users' historical data and addressing missing not-at-random (MNAR) scenarios. By focusing on the most influential historical data, they provide more explainable recommendations. Similarly, Zhang et al. [28] integrate constructed features into a factorization model to generate both predictions and explanations. Their approach employs templates based on users' key features to create interpretable outputs.

Wang et al. [38] extend this concept by constructing a two- or three-way tensor to generate explainable recommendations within a latent space using a joint factorization technique. This method projects features and opinionated phrases to produce clear and meaningful explanations. Cheng et al. [39] focus on learning user preferences across various aspects and incorporate these into their proposed Aspect-Aware Latent Factor Model (MLFM), which generates explainable predictions for user ratings.

3.7. LIME-Based Approach

Local Interpretable Model-Agnostic Explanations (LIME) is a versatile method used to generate explanations by analyzing the importance of features in predictive models. Its applications span various domains, from text-based recommendation systems to visual recognition tasks, enhancing interpretability across different modalities. Singh et al. [40] apply the LIME approach to explain a classifier's output. They use a deep relevance matching model to create query documents, calculate relevance scores, and convert these scores into a probability distribution to generate explanations. Similarly, Khan et al. [41] utilize LIME to explain malaria-infected red blood cell (RBC) images. By analyzing weighted, perturbed image pixels, LIME effectively identifies infected RBCs and provides meaningful explanations for its predictions.

3.8. Aspect-Based Approach

Aspect-based systems focus on user and item attributes to provide explanations for recommendations. By constructing preference and quality matrices, these approaches generate explanations aligned with user priorities, thereby enhancing decision-making. Hou et al. [42] utilize User Aspect Preference (UAP) and Item Aspect Quality (IAQ) data to produce explainable recommendations, offering insights into why users select particular items. They create UAP and IAQ matrices, which are integrated into the item rating matrix to facilitate aspect-level explainability.

Luo et al. [43] address users' dynamic and personalized preferences to generate aspect-level explanations. They employ a Long Short-Term Memory (LSTM) unit and build user aspect preference vectors and item aspect quality vectors using an encoder-decoder-based network, enabling more detailed and personalized explanations. Zhao et al. [44] take a different approach, applying a bidirectional recurrent neural network (Bi-RNN) combined with Gated Recurrent Units (GRU), an exact pairwise matching technique, and the tf-idf method to calculate the final score for answer candidates, further enhancing the interpretability of recommendations.

.

3.9. Hybrid Approaches

Hybrid systems combine multiple techniques to enhance interpretability, such as integrating neural networks with probabilistic models or knowledge graphs to enable dynamic and context-aware explanations. Zhao et al. [44] utilize a soft pairwise matching technique to identify associations between internal and external words, improving the interpretability of PQA-based recommender systems through re-weighted internal and external word pairs. Tan et al. [45] focus on learning user and item preference distributions alongside latent factors to produce interpretable predictions, leveraging extracted latent topic features for clarity. Zhang et al. [46] construct product profiles by calculating sentiment polarity scores for matched features based on feature frequency and integrated product reviews, aiding users in decision-making and enhancing system interpretability.

Tao et al. [47] extract features from unstructured data using deep sequential modeling and a recurrent semantic memory unit (RSMU) to capture semantic information, generating interpretable results. Zanker et al. [48] propose a knowledge-based reasoning framework represented by a directed acyclic graph (DAG), where paths in the graph provide explanations, with nodes containing arguments. Park et al. [49] use high predicted ratings as explanation sources, supplemented by similar sequence entities derived from node similarities and extracted node pairs to improve interpretability.

Sun et al. [50] design an encoder-decoder architecture for review generation, where the encoder transforms attributes into vectors, and the decoder employs an LSTM unit to retrieve word embeddings. Costa et al. [51] develop an LSTM RNN-based model to generate reviews by combining review scores and item ratings, learning grammar through LSTM cells to produce clear and concise explanations. Baral et al. [52] represent POI-aspect relations using bipartite graphs, generating explanations through bipartite cores, shingles, and ranking-based methods. Zhao et al. [53] propose the SAR model for POI recommendations, identifying top aspects matching user preferences to generate explanations.

Lee et al. [54] evaluate stories based on frequency and aggressiveness, constructing a character composition matrix to generate explanations by analyzing group proximity and importance. Cheng et al. [55] introduce the ERRA model (Explainable Recommendation by Personalized Review Retrieval and Aspect Learning), which incorporates a retrieval enhancement mechanism to extract additional training data information, improving explanation accuracy. The model includes an aspect enhancement component to identify top-n user-relevant aspects, refining representations for more personalized and persuasive explanations. Experiments on three datasets confirm ERRA's superior performance. Zhao et al. [56] propose the MMI (Maximizing Mutual Information) framework, a flexible, model-agnostic method to align generated explanations with predicted ratings or key item attributes. By leveraging mutual information (MI) as an alignment metric and incorporating a neural MI estimator for reinforcement learning, the framework fine-tunes explanation models. Experiments show that MI improves alignment and enhances user decision support, outperforming existing methods. User studies further validate that MI-optimized explanations are more effective and user-preferred due to their superior alignment.

Explainable recommender systems can be divided into two categories based on how explanations are generated: model-intrinsic and model-agnostic systems. Model-intrinsic systems integrate the explanation process within the recommendation model itself, leveraging inherently interpretable frameworks. In contrast, model-agnostic systems generate recommendations independently and then attach explanations as an additional layer. Most current recommender systems belong to the model-agnostic category.

4. CONCLUSIONS

We conducted an in-depth analysis and classification of existing models and approaches used in explainable recommender systems, highlighting their evaluation methodologies, outcomes, and the strengths and limitations of the techniques employed. Additionally, we explored various strategies for developing novel methods or combining established approaches with emerging technologies to improve the performance of explainable recommender systems across different domains. In future research, we plan to focus on further advancing model-intrinsic explainable recommender systems.

REFERENCES

[1] N. Tintarev et. al., “Explanations of recommendations,” in Proceedings of the 2007 ACM Conference on Recommender Systems, ser. RecSys ’07. New York, NY, USA: Association for Computing Machinery, 2007, p. 203–206. [Online]. Available: https://doi.org/10.1145/1297231.1297275

[2] N. Tintarev and J. Masthoff, “Designing and evaluating explanations for recommender systems,” in Recommender Systems Handbook, 2011.

[3] Tintarev, Nava and Masthoff, Judith, “Effective explanations of recommendations: User-centered design,” in Proceedings of the 2007 ACM Conference on Recommender Systems, ser. RecSys ’07. New York, NY, USA: Association for Computing Machinery, 2007, p. 153–156. [Online]. Available: https://doi.org/10.1145/1297231.1297259

[4] T. Nanou, G. Lekakos, and K. Fouskas, “The effects of recommendations’ presentation on persuasion and satisfaction in a movie recommender system,” Multimedia Syst., vol. 16, no. 4–5, p. 219–230, Aug. 2010. [Online]. Available: https://doi.org/10.1007/s00530-010- 0190-0

[5] M. Z. Al-Taie and S. Kadry, “Visualization of explanations in recom- mender systems,” Journal of Advanced Management Science Vol, vol. 2, no. 2, pp. 140–144, 2014

[6] X. Wang, X. He, F. Feng, L. Nie, and T.-S. Chua, “Tem: Tree- enhanced embedding model for explainable recommendation,” in Proceedings of the 2018 World Wide Web Conference, ser. WWW ’18. Republic and Canton of Geneva, CHE: International World Wide Web Conferences Steering Committee, 2018, p. 1543–1552. [Online]. Available: https://doi.org/10.1145/3178876.3186066

[7] O. Barkan, Y. Fuchs, A. Caciularu, and N. Koenigstein, “Explainable recommendations via attentive multi-persona collaborative filtering,” in Fourteenth ACM Conference on Recommender Systems, ser. RecSys ’20. New York, NY, USA: Association for Computing Machinery, 2020, p. 468–473. [Online]. Available: https://doi.org/10.1145/3383313.3412226

[8] X. Chen, H. Chen, H. Xu, Y. Zhang, Y. Cao, Z. Qin, and H. Zha, “Personalized fashion recommendation with visual explanations based on multimodal attention network: Towards visually explainable recommendation,” in Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval, ser. SIGIR’19. New York, NY, USA: Association for Computing Machinery, 2019, p. 765–774. [Online]. Available https://doi.org/10.1145/3331184.3331254

[9] S. Seo, J. Huang, H. Yang, and Y. Liu, “Interpretable convolutional neural networks with dual local and global attention for review rating prediction,” in Proceedings of the Eleventh ACM Conference on Recom- mender Systems, 2017, pp. 297–305.

[10] C. Chen, M. Zhang, Y. Liu, and S. Ma, “Neural attentional rating regression with review-level explanations,” in Proceedings of the 2018 World Wide Web Conference, ser. WWW ’18. Republic and Canton of Geneva, CHE: International World Wide Web Conferences Steering Committee, 2018, p. 1583–1592. [Online]. Available: https://doi.org/10.1145/3178876.3186070

[11] X. Wang, Y. Chen, J. Yang, L. Wu, Z. Wu, and X. Xie, “A reinforcement learning framework for explainable recommendation,” in 2018 IEEE International Conference on Data Mining (ICDM), 2018, pp. 587–596.

[12] X. Dong, J. Ni, W. Cheng, Z. Chen, B. Zong, D. Song, Y. Liu, H. Chen, and G. de Melo, “Asymmetrical hierarchical networks with attentive interactions for interpretable review-based recommendation,” 2019.

[13] X. Chen, Y. Zhang, and Z. Qin, “Dynamic explainable recommendation based on neural attentive models,” in Proceedings of the AAAI Confer- ence on Artificial Intelligence, vol. 33, 2019, pp. 53–60.

[14] C. Yu, L. V. S. Lakshmanan, and S. Amer-Yahia, “Recommendation diversification using explanations,” in 2009 IEEE 25th International Conference on Data Engineering, 2009, pp. 1299–1302.

[15] Z. Chen, X. Wang, X. Xie, T. Wu, G. Bu, Y. Wang, and E. Chen, “Co- attentive multi-task learning for explainable recommendation.” in IJCAI, 2019, pp. 2137–2143.

[16] Y. Lin, P. Ren, Z. Chen, Z. Ren, J. Ma, and M. de Rijke, “Explainable outfit recommendation with joint outfit matching and comment genera- tion,” IEEE Transactions on Knowledge and Data Engineering, vol. 32, no. 8, pp. 1502–1516, 2020.

[17] H. Chen, X. Chen, S. Shi, and Y. Zhang, “Generate natural language ex- planations for recommendation,” in Proceedings of SIGIR’19 Workshop on ExplainAble Recommendation and Search. ACM, 2019.

[18] Y. Lu, R. Dong, and B. Smyth, “Why i like it: Multi-task learning for recommendation and explanation,” in Proceedings of the 12th ACM Conference on Recommender Systems, ser. RecSys ’18. New York, NY, USA: Association for Computing Machinery, 2018, p. 4–12. [Online]. Available : https://doi.org/10.1145/3240323.3240365

[19] G. Zhao, H. Fu, R. Song, T. Sakai, Z. Chen, X. Xie, and X. Qian, “Personalized reason generation for explainable song recommendation,” ACM Trans. Intell. Syst. Technol., vol. 10, no. 4, Jul. 2019. [Online] . Available: https://doi.org/10.1145/3337967

[20] P. Li, Z. Wang, Z. Ren, L. Bing, and W. Lam, “Neural rating regression with abstractive tips generation for recommendation,” Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval, Aug 2017. [Online]. Available: http://dx.doi.org/10.1145/3077136.3080822

[21] G. Peake and J. Wang, “Explanation mining: Post hoc interpretability of latent factor models for recommendation systems,” in Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery amp; Data Mining, ser. KDD ’18. New York, NY, USA: Association for Computing Machinery, 2018, p. 2060–2069. [Online]. Available: https://doi.org/10.1145/3219819.3220072

[22] K. Tsukuda and M. Goto, “Explainable recommendation for repeat consumption,” in Fourteenth ACM Conference on Recommender Systems, ser. RecSys ’20. New York, NY, USA: Association for Computing Machinery, 2020, p. 462–467. [Online]. Available: https://doi.org/10.1145/3383313.3412230

[23] P. Kouki, S. Fakhraei, J. Foulds, M. Eirinaki, and L. Getoor, “Hyper: A flexible and extensible probabilistic framework for hybrid recommender systems,” 09 2015.

[24] A. Samih, A. Adadi, and M. Berrada, “Towards a knowledge-based explainable recommender systems,” in Proceedings of the 4th International Conference on Big Data and Internet of Things, ser. BDIoT’19. New York, NY, USA: Association for Computing Machinery, 2019. [Online]. Available: https://doi.org/10.1145/3372938.3372959

[25] W. Ma, M. Zhang, Y. Cao, W. Jin, C. Wang, Y. Liu, S. Ma, and X. Ren, “Jointly learning explainable rules for recommendation with knowledge graph,” in The World Wide Web Conference, 2019, pp. 1210– 1221.

[26] P. Kouki, J. Schaffer, J. Pujara, J. O’Donovan, and L. Getoor, “Personalized explanations for hybrid recommender systems,” in Proceedings of the 24th International Conference on Intelligent User Interfaces, ser. IUI ’19. New York, NY, USA: Association for Computing Machinery, 2019, p. 379–390. [Online]. Available: https://doi.org/10.1145/3301275.3302306

[27] K. Tsukuda and M. Goto, “Dualdiv: Diversifying items and explanation styles in an explainable hybrid recommendation,” in Proceedings of the 13th ACM Conference on Recommender Systems, RecSys ’19. New York, NY, USA: Association for Computing Machinery, 2019, p. 398–402. [Online]. Available: https://doi.org/10.1145/3298689.3347063

[28] Y. Zhang, G. Lai, M. Zhang, Y. Zhang, Y. Liu, and S. Ma, "Explicit factor models for explainable recommendation based on phrase-level sentiment analysis," in Proceedings of the 37th International ACM SIGIR Conference on Research amp; Development in Information Retrieval, ser. SIGIR '14. New York, NY, USA: Association for Computing Machinery, 2014, p. 83–92. [Online]. Available: https://doi.org/10.1145/2600428.2609579

[29] K. Balog, F. Radlinski, and S. Arakelyan, "Transparent, scrutable and explainable user models for personalized recommendation," in Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval, ser. SIGIR'19. New York, NY, USA: Association for Computing Machinery, 2019, p. 265–274. [Online]. Available: https://doi.org/10.1145/3331184.3331211

[30] D. McSherry, "Explanation in recommender systems," Artificial Intelligence Review, vol. 24, no. 2, pp. 179–197, 2005.

[31] T. N. T. Tran, M. Atas, A. Felfernig, V. M. Le, R. Samer, and M. Stettinger, "Towards social choice-based explanations in group recommender systems," in Proceedings of the 27th ACM Conference on User Modeling, Adaptation, and Personalization, UMAP '19. New York, NY, USA Association for Computing Machinery, 2019, p. 13–21. [Online]. Available: https://doi.org/10.1145/3320435.3320437

[32] X. Huang, Q. Fang, S. Qian, J. Sang, Y. Li, and C. Xu, "Explainable interaction-driven user modeling over knowledge graph for a sequential recommendation," in Proceedings of the 27th ACM International Conference on Multimedia, ser. MM '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 548–556. [Online]. Available: https://doi.org/10.1145/3343031.3350893

[33] Y. Zhang, X. Xu, H. Zhou, and Y. Zhang, "Distilling structured knowledge into embeddings for explainable and accurate recommendation," WSDM '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 735–743. [Online]. Available: https://doi.org/10.1145/3336191.3371790

[34] Z. Fu, Y. Xian, R. Gao, J. Zhao, Q. Huang, Y. Ge, S. Xu, S. Geng, C. Shah, Y. Zhang, and G. de Melo, "Fairness-aware explainable recommendation over knowledge graphs," in Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval, ser. SIGIR '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 69–78. [Online]. Available: https://doi.org/10.1145/3397271.3401051

[35] Y. Xian, Z. Fu, S. Muthukrishnan, G. de Melo, an ○ Y. Zhang, "Reinforcement knowledge graph reasoning for explainable recommendation," SIGIR'19. New York, NY, USA: Association for Computing Machinery, 2019, p. 285–294. [Online]. Available: https://doi.org/10.1145/3331184.3331203

[36] B. Abdollahi and O. Nasraoui, "Using explainability for constrained matrix factorization," in Proceedings of the Eleventh ACM Conference on Recommender Systems, RecSys '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 79–83. [Online]. Available: https://doi.org/10.1145/3109859.3109913

[37] H. Liu, J. Wen, L. Jing, J. Yu, X. Zhang, and M. Zhang, "In2rec: Influence-based interpretable recommendation,". CIKM '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 1803–1812. [Online]. Available: https://doi.org/10.1145/3357384.3358017

[38] N. Wang, H. Wang, Y. Jia, and Y. Yin, "Explainable recommendation via multi-task learning in opinionated text data," The 41st International ACM SIGIR Conference on Research Development in Information Retrieval, Jun 2018. [Online]. Available: http://dx.doi.org/10.1145/3209978.3210010

[39] Z. Cheng, X. Chang, L. Zhu, R. C. Kanjirathinkal, and M. Kankanhalli, "Mmalfm," ACM Transactions on Information Systems, vol. 37, no. 2, p. 1–28, Mar 2019. [Online]. Available: http://dx.doi.org/10.1145/3291060

[40] J. Singh and A. Anand, "Exs: Explainable search using local model agnostic interpretability," in Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining, WSDM '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 770–773. [Online]. Available: https://doi.org/10.1145/3289600.3290620

[41] A. Khan, K. D. Gupta, D. Venugopal, and N. Kumar, "Cidmp: Completely interpretable detection of malaria parasite in red blood cells using lower-dimensional feature space," 2020 International Joint Conference on Neural Networks (IJCNN), Jul 2020. [Online]. Available: http://dx.doi.org/10.1109/IJCNN48605.2020.9206885

[42] Y. Hou, N. Yang, Y. Wu, and P. S. Yu, "Explainable recommendation with fusion of aspect information," World Wide Web, vol. 22, no. 1, p. 221–240, Jan. 2019. [Online]. Available: https://doi.org/10.1007/s11280-018-0558-1

[43] H. Luo, N. Yang, and P. S. Yu, "Hybrid deep embedding for recommendations with dynamic aspectlevel explanations," 2020. [Online]. Available: https://arxiv.org/pdf/2001.10341

[44] J. Zhao, Z. Guan, and H. Sun, "Riker: Mining rich keyword representations for interpretable product question answering," in Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery amp; Data Mining, KDD '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 1389–1398. [Online]. Available: https://doi.org/10.1145/3292500.3330985

[45] Y. Tan, M. Zhang, Y. Liu, and S. Ma, "Rating-boosted latent topics: Understanding users and items with ratings and reviews," in IJCAI, 2016.

[46] Y. Zhang, "Incorporating phrase-level sentiment analysis on textual reviews for personalized recommendation," in Proceedings of the Eighth ACM International Conference on Web Search and Data Mining, WSDM '15. New York, NY, USA: Association for Computing Machinery, 2015, p. 435–440. [Online]. Available: https://doi-org.libproxy.lib.ilstu.edu/10.1145/2684822.2697033

[47] Z. Tao, S. Li, Z. Wang, C. Fang, L. Yang, H. Zhao, and Y. Fu, "Log2intent: Towards interpretable user modeling via recurrent semantics memory unit," KDD '19. New York, NY, USA: Association for Computing Machinery, 2019, p. 1055–1063. [Online]. Available: https://doi.org/10.1145/3292500.3330889

[48] M. Zanker and D. Ninaus, "Knowledgeable explanations for recommender systems," in 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, vol. 1, 2010, pp. 657–660.

[49] H. Park, H. Jeon, J. Kim, B. Ahn, and U. Kang, "Uniwalk: Explainable and accurate recommendation for rating and network data," 2017.

[50] P. Sun, L. Wu, K. Zhang, Y. Fu, R. Hong, and M. Wang, "Dual learning for explainable recommendation: Towards unifying user preference prediction and review generation," in Proceedings of The Web Conference 2020, WWW '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 837–847. [Online]. Available: https://doi.org/10.1145/3366423.3380164

[51] F. Costa, S. Ouyang, P. Dolog, and A. Lawlor, "Automatic generation of natural language explanations," in Proceedings of the 23rd International Conference on Intelligent User Interfaces Companion, ser. IUI '18 Companion. New York, NY, USA: Association for Computing Machinery, 2018. [Online]. Available: https://doi.org/10.1145/3180308.3180366

[52] R. Baral, X. Zhu, S. S. Iyengar, and T. Li, "Reel: Review the aware explanation of location recommendation," in Proceedings of The 26th Conference on User Modeling, Adaptation, and Personalization, UMAP '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 23–32. [Online]. Available: https://doi.org/10.1145/3209219.3209237

[53] K. Zhao, G. Cong, Q. Yuan, and K. Q. Zhu, “Sar: A sentiment-aspect- region model for user preference analysis in geo-tagged reviews,” in 2015 IEEE 31st International Conference on Data Engineering. IEEE, 2015, pp. 675–686.

[54] O.-J. Lee and J. Jung, “Explainable movie recommendation systems by using story-based similarity,” The Workshop on Explainable Smart Systems (ExSS 2018), held in conjunction with the 23rd International Conference on Intelligent User Interfaces (IUI), 2018.

[55] Hao Cheng, Shuo Wang, Wensheng Lu, Wei Zhang, Mingyang Zhou, Kezhong Lu, Hao Liao, “Explainable Recommendation with Personalized Review Retrieval and Aspect Learning Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics, pages 51–64, July 9-14, 2023.

[56] Yurou Zhao, Yiding Sun, Ruidong Han, Fei Jiang, Lu Guan, Xiang Li, Wei Lin, Weizhi Ma, Jiaxin Mao “Aligning Explanations for Recommendation with Rating and Feature via Maximizing Mutual Information,” Proceedings of the 33rd ACM International Conference on Information and Knowledge Management (CIKM '24), Pages 3374 - 3383, Oct. 21-25, 2024.

AUTHORS

Alfrin Saldanha was a graduate student at the School of Information Technology at Illinois State University while he was working on the project for the paper.

Hyoil Han is an Associate Professor of Computer Science at the School of Information Technology at Illinois State University. She earned a B.S. in Electrical Engineering from Korea University, an M.S. in Electrical Engineering from the Korea Advanced Institute of Science and Technology, and a Ph.D. in Computer Science and Engineering from the University of Texas at Arlington. Her professional experience includes roles as a Technical Staff Member at Samsung Electronics and Korea Telecom. Dr. Han has published over 60 technical papers, with her research focusing on machine learning, natural language processing, big data management, and AI applications in security. She is also a member of the ACM.