Opacity to Transparency: A Journey into Explainable Recommender Systems

Drashti Shrimal[1], Dr. Harshali Patil[2]
[1] *Drashti Shrimal, Thakur College of Engineering and Technology, Mumbai.*
[2] *Dr. Harshali Patil, Thakur College of Engineering and Technology, Mumbai.*

***Abstract*—Recommender systems have become an integral part of our daily lives, assisting users in discovering the relevant products, services, and information. However, the lack of transparency in the decision-making process of traditional recommender systems has raised concerns regarding user trust and understanding. To address this issue, explainable recommender systems have emerged as a promising research direction. This paper provides a comprehensive overview of explainable recommender systems, focusing on their significance & need, areas of usability, rationale, expected outcomes and challenges. Further it provides a comparative analysis of three approaches used in XAI. To provide a holistic perspective, we review existing literature and highlight key research trends and open challenges in the field of explainable recommender systems. Overall, this paper aims to serve as a comprehensive resource for researchers, practitioners, and decision-makers interested in understanding the state of the art in explainable recommender systems.**

***Index Terms*—Explainable Recommender systems, User trust, User engagement.**

I. INTRODUCTION

Recommendation systems have become a crucial component of various digital platforms, including e-commerce, social media, and entertainment services. These systems are designed to predict user preferences and suggest relevant items such as products, movies, or music. However, traditional recommendation systems often lack transparency and interpretability, leading to user distrust and dissatisfaction. Explainable recommendation systems (XRS) aim to address this issue by providing clear and understandable explanations of how recommendations are generated. This helps users make informed decisions and build trust in the system.

As recommendation systems play an increasingly prominent role in our daily lives, it is crucial to ensure that these systems are transparent and comprehensible to users. Explainable recommendation systems are designed to offer users insights into the reasons behind specific recommendations, thereby increasing transparency and user trust. These systems not only enhance the user experience but also provide businesses with opportunities to optimize their recommendation strategies.
This paper explores the concept of explainable recommendation systems and emphasizes their importance in promoting user understanding and trust. Additionally, we discuss the challenges associated with developing these systems and the significance of addressing them effectively.

II. SIGNIFICANCE OF THE STUDY

Explainable recommendation systems have gained significant attention internationally due to the growing concern over the transparency, accountability, and ethical implications of automated decision-making systems. These systems aim to provide users with explanations or justifications for the recommendations they receive, enhancing user trust, understanding, and control over the decision-making process. Here are some key use cases and benefits of explainable recommendation systems internationally:

1. ***E-commerce***: In the realm of online shopping, explainable recommendation systems help users understand why specific products are recommended to them. By providing transparent explanations, such as highlighting relevant features, user preferences, or similar user behaviors, these systems enable users to make informed purchasing decisions. This improves customer satisfaction, reduces uncertainty, and fosters trust between users and e-commerce platforms.

2. ***Content and media platforms***: Explainable recommendation systems are employed by content streaming services, news aggregators, and social media platforms to suggest personalized content to users. By explaining why certain articles, videos, or posts are recommended, these systems help users understand the underlying factors, such as their interests, browsing history, or social connections. This transparency allows users to evaluate and control the information they consume, mitigating issues like filter bubbles and algorithmic biases.

3. ***Healthcare***: In the healthcare domain, explainable recommendation systems are crucial for supporting clinical decision-making. For instance, in personalized medicine, these systems can explain the reasoning behind treatment recommendations, taking into account patient characteristics, medical history, and relevant research findings. Transparent explanations enable doctors and patients to understand the basis of the recommendations and engage in shared decision-making, improving treatment adherence and patient outcomes.

4. ***Financial services***: Explainable recommendation systems are valuable in the financial sector, where automated algorithms suggest investment options, loans, or insurance plans to users. By providing clear justifications, such as risk factors, historical performance, or regulatory compliance, these systems empower users to assess the credibility and fairness of the recommendations. This promotes transparency in financial decision-making, helps users make informed choices, and mitigates potential biases or conflicts of interest.

5. ***Education***: In the field of education, explainable recommendation systems assist students in selecting courses, educational resources, or career paths. By offering explanations based on academic performance, personal interests, or future job prospects, these systems enable students to understand the rationale behind the recommendations. This fosters self-directed learning, enhances student engagement, and supports informed decision-making regarding educational pursuits.

Internationally, the use of explainable recommendation systems aligns with increasing regulatory efforts and ethical considerations surrounding algorithmic transparency and fairness. Governments, industry organizations, and researchers are actively working on developing guidelines and standards to ensure the responsible deployment of recommendation systems. By prioritizing explainability, these systems can address concerns related to privacy, discrimination, and unintended consequences, ultimately benefiting users and society as a whole.

III. STUDY RATIONALE

The rationale for developing an explainable recommendation system lies in the need to address the black box nature of traditional recommendation systems. While these systems have been successful in generating accurate recommendations, they often lack transparency and fail to provide users with meaningful explanations for their recommendations. This lack of transparency can lead to user distrust, dissatisfaction, and a lack of adoption of the recommendations.

Here are some key rationales for explainable recommendation systems:

1. ***User Trust***: Explainability promotes user trust by providing transparent and understandable recommendations. When users receive recommendations accompanied by clear explanations, they are more likely to trust the system and feel confident in the suggestions made.

2. ***User Understanding***: Explanation capabilities in recommendation systems enhance user understanding of the underlying reasons for the recommendations. By providing meaningful justifications, users can gain insights into why certain items or content are being recommended to them. This helps users make informed decisions and increases their satisfaction with the system.

3. ***User Control and Customization***: Explainable systems empower users to have more control over the recommendation process. By understanding the reasoning behind the recommendations, users can provide feedback, adjust their preferences, and fine-tune the system to better align with their needs and preferences.

4. ***Accountability and Fairness***: Explainability plays a crucial role in ensuring accountability and fairness in recommendation systems. It allows users, developers, and regulators to identify and address biases,

discrimination, or unethical practices that may be present in the recommendation algorithms. Explanations help to identify potential biases and provide an avenue for corrective actions.

5. ***Regulatory Compliance***: In certain domains, such as finance, healthcare, or legal sectors, explainability is legally required. Compliance with regulations like the General Data Protection Regulation (GDPR) or the right to explanation provisions necessitates that users have access to understandable explanations for the recommendations made by an algorithm.

6. ***Learning and Adaptation***: Explainable recommendation systems facilitate user learning and adaptation. When users receive explanations, they can understand their own preferences better, discover new items or content, and refine their own decision-making processes. The feedback loop between users and the system can help improve the recommendations over time.

Overall, the rationale for an explainable recommendation system is to enhance user trust, understanding, control, fairness, accountability, and compliance. By providing meaningful explanations, these systems aim to address the limitations of traditional black box recommendation systems and create more transparent, user-centric, and ethical recommendation experiences.

IV. LITERATURE REVIEW

In [1], a method for explaining recommendations in a Recommendation system is presented. The method uses SHAP (SHapley Additive exPlanations) to calculate the contribution of each feature in making a recommendation. Instead of using the SHAP outputs directly, the method uses them to select possible candidates for counterfactual explanations, which are easier to understand and more faithful. The explanation method can be applied at two levels: list-level explanations that consider the recommendation list as a whole and instance-level explanation that only explain a single instance in the list. Experiments showed that SHAP can reduce the time needed for searching counterfactual explanations under certain conditions. The authors plan to evaluate the explanation method through a user study, to extend the approach to large-scale data and other recommendation systems that do not involve features, and to incorporate contextual information and other side information into the explanations. The authors of this paper present a case study to explain recommendations made by a context-aware recommendation system (CARS) that was proposed in a previous paper. The dataset used is LDOS-CoMoDa, a dataset for context-aware movie recommendation that includes 121 users, 1197 movies, and 12 context factors. The goal of the recommendation system is to generate recommendations for a user given their contextual situation, which is composed of several contextual conditions. To apply Algorithms to the CARS, the authors compute counterfactual explanations for the recommendations. In this case, the contextual factors are all categorical and the values of the contextual factors are changed randomly to identify counterfactual explanations. The authors present an example where the user's mood is positive and is identified as the most impacting factor. Changing the user's mood from positive to negative results in a change in the recommended movie list. The authors explain the recommendations at the instance level (Top-1), meaning they explain the reasoning behind a single recommendation. Let's have an overview on the contributions of the paper. The paper "Shap-enhanced counterfactual explanations for Recommendations" is significant and contributes to the field of machine learning in several ways:

- Explanation: The paper provides a way to explain the recommendations made by machine learning models, which is important for building trust and improving transparency in the AI field.
- Counterfactual explanations: The paper introduces the concept of counterfactual explanations for recommendations, which allows for a better understanding of how a particular recommendation was made. The approach helps users to understand the impact of changing certain features on the recommendation and how it would have affected the outcome.
- Model-agnostic approach: The approach is model- agnostic, meaning it can be applied to any machine learning model, regardless of the underlying algorithm. This makes the

method more widely applicable and accessible.
- Empirical evaluation: The paper provides empirical evidence of the effectiveness of the approach through experiments on real-world dataset.

Let's have a glance on the strengths and weaknesses of the approach:

a) Strengths
- The paper presents a novel approach to provide counterfactual explanations for recommendation systems, by incorporating the SHAP values.
- The results demonstrate that the SHAP-enhanced explanations provide better explanations compared to traditional methods, as they are able to capture non-linear relationships between features and model predictions.
- The approach is flexible and can be applied to various recommendation systems, including both collaborative filtering and matrix factorization models.
- The paper provides a comprehensive evaluation of the proposed approach, including both qualitative and quantitative analysis.

b) Weaknesses:
- The study only focuses on a small dataset, and it may not be generalizable to other recommendation systems or datasets.
- The evaluation metrics used in the paper are limited, and it would be beneficial to consider other metrics to evaluate the quality of explanations.
- The paper does not provide a thorough discussion on the limitations of the proposed approach and its potential drawbacks.
- The study does not consider the scalability of the approach, and it may not be suitable for large-scale recommendation systems.

Overall, the paper presents a promising approach to provide counterfactual explanations for recommendation systems, but further research is needed to validate its effectiveness and generalizability.

The paper [2] deals with the problem of helping customers purchase the best product among alternative products. The authors propose an interpretable machine learning approach to determine the key features of a product that best explain the price and help customers differentiate the most suitable product. The problem is formulated as a supervised machine learning problem based on price and evaluated using linear and tree models with Shapley Values. The results of offline evaluation show that the proposed method outperforms the baseline and is comparable in conversion rate to the baseline in online A/B tests. The authors also involve human experts to evaluate the relevance of the recommendations.

The authors focus on building a product feature recommendation system for online customers. The aim of this system is to help customers differentiate specific products from a set of similar products based on recommended features. The paper uses interpretable machine learning techniques to explain how the machine learning models work and to facilitate human understanding of the final model. The authors select SHAP (SHapley Additive exPlanations), a model-agnostic method, as the tool to interpret the machine learning models and understand the feature importance. The paper describes the process of processing the raw data, extracting features, and building regression models using Linear Regression, LightGBM and CatBoost. The authors use these regression models to learn the feature importance and understand what features contribute more to the product's price. The authors also use the feature direction to understand if a feature positively or negatively contributes to the product price. The final feature ranking list is computed by averaging the Shapley Values for each feature from all three regression models.

The authors evaluate the proposed method against a baseline algorithm Left Nav Algorithm and find that it scores higher on offline evaluation metrics and is comparable with the baseline algorithm in online A/B tests.

Let's have an overview on the contributions of the paper. The paper "Online Product Feature Recommendations with Interpretable Machine Learning" is significant because it addresses the need for interpretable product feature recommendations for customers in e-commerce settings.

It contributes to the field by introducing a new approach to product feature recommendations that utilizes interpretable machine learning to generate feature recommendations. The use of the price as the training label and the study of the Shapley Values provide a unique and effective solution to the problem of product feature recommendations. The results of the

offline evaluation and online A/B tests demonstrate the effectiveness of the proposed approach, making it a significant contribution to the field of e-commerce product recommendations.
Let’s have a glance on the strengths and weaknesses of the approach:

a)Strengths:

- The use of the model-agnostic method, Shapley Values, allows for interpretation of the contribution of features to the model predictions based on data visualizations.
- The approach has been evaluated against a strong baseline and has achieved higher scores in offline evaluations on metrics such as NDCG, precision, recall and coverage.

b) Weaknesses:

- The approach is limited to the use of the product price as the training label and may not be suitable for other types of training labels.
- The use of human experts to label the top features may introduce bias and subjectivity into the evaluation process.
- The online A/B tests on conversion rate were only comparable with the baseline and may not have a significant improvement over existing approaches.

The paper [3] aims to provide an explanation for recommendations generated by machine learning models. The authors address the issue of limited interpretability in existing recommendation systems and present a new approach to enhance their explainability.
The paper proposes the use of local surrogate models, which are simple and interpretable models that are trained to approximate the behavior of the complex recommendation model in a local region around a specific instance. The authors argue that these local surrogate models can provide a better explanation of the recommendation compared to global models that attempt to explain the behavior of the recommendation model for the entire data set. The authors evaluate the performance of the proposed approach on a movie recommendation task, using a matrix factorization model as the recommendation model. The results show that the local surrogate models can effectively explain the recommendations generated by the complex model, by highlighting the features that contribute most to the prediction. The paper concludes that the use of local surrogate models can provide a simple and interpretable explanation for recommendations generated by complex machine learning models, helping users to understand the reasoning behind the recommendations. This can increase the transparency and trust of the recommendation system, leading to a more positive user experience.
Let’s have an overview on the contributions of the paper.
Overall, the paper contributes to the field by presenting a novel approach to enhance the interpretability of recommendation systems, which can benefit both users and practitioners of machine learning-based recommendation systems.
Let’s have a glance on the strengths and weaknesses of the approach:

a) Strengths:

- Local interpretability: The method proposed in the paper provides local interpretability of recommendation systems, which means that the reasons for a particular recommendation can be explained at the individual user level. This is useful for building trust with users and for identifying potential biases or inaccuracies in the system.
- Scalability: The method is scalable and can be applied to large datasets and complex recommendation algorithms. This is important for real-world applications where recommendation systems may have millions of users and items.
- Flexibility: The proposed method is flexible and can be applied to a wide range of recommendation algorithms, including collaborative filtering and content-based approaches. This means that it can be adapted to different domains and applications.
- Accuracy: The paper reports that the proposed method is able to achieve high accuracy in predicting user preferences and explaining recommendations. This is important for building trust with users and for improving the overall performance of recommendation systems.

- Novelty: The proposed method is a novel approach to explaining recommendation systems, which means that it offers a new perspective on this important problem. This can lead to new insights and improvements in recommendation systems.

b)Weaknesses:

- Limited interpretability: Local surrogate models are designed to be simpler and more interpretable than the original machine learning model, but they can still be difficult to interpret in some cases. The interpretation of these models is limited to the specific features used by the surrogate model and does not necessarily provide a complete understanding of the original model.
- Data dependency: The effectiveness of local surrogate models is highly dependent on the data used to train them. If the data is not representative or if there are biases in the data, the explanations provided by the surrogate model may not be accurate or trustworthy.
- Scalability: Local surrogate models require the computation of a new model for each prediction, which can be computationally expensive and time-consuming. This can limit the scalability of the technique, particularly in real-time applications where speed is important.
- Lack of generality: Local surrogate models only explain the predictions for a single instance and do not provide a global understanding of the model. This can limit the usefulness of the technique for understanding the behavior of the model as a whole.
- Interpretation bias: The interpretation of local surrogate models may be biased towards the assumptions and choices made by the modeler. This can limit the ability of the technique to provide unbiased and objective explanations

In recent years, there has been a growing interest in developing interpretable machine learning models for making recommendations in e-commerce and other domains. One approach to achieving interpretability is through the use of Shapley values to generate counterfactual explanations for recommendations. These explanations can help users understand why certain products or features are recommended to them and can increase their trust in the system.

The paper "Shap-enhanced counterfactual explanations for recommendations" proposes a novel method for generating counterfactual explanations that combines Shapley values with causal inference techniques. The method is evaluated on a real-world e-commerce dataset and is shown to outperform existing methods in terms of both the quality of the explanations and the accuracy of the recommendations.

Another paper, "Online Product Feature Recommendations with Interpretable Machine Learning", proposes a framework for making feature recommendations in an online setting using interpretable machine learning models. The framework uses a combination of collaborative filtering and decision trees to generate recommendations and provides interpretable explanations for the recommendations using Shapley values. [4]

Finally, the paper "Towards Explaining Recommendations through Local Surrogate Models" proposes a method for generating local surrogate models that can provide explanations for individual recommendations. The method is evaluated on a movie recommendation dataset and is shown to provide accurate and interpretable explanations for the recommendations. [5]

Let’s study comparative analysis of above methodologies:

Paper	Gaps	Findings
1	Focuses on a small dataset.	Proposes a novel method for generating counterfactual explanations that combines Shapley values with causal inference techniques
2	Does not consider the scalability of the approach. Thus not	Helps customers differentiate specific products from a set of similar

	recommended for larger systems.	products based on recommended features
3	No significant improvement over existing approaches.	Proposes a framework for making feature recommendations in an online setting using interpretable machine learning models
4	Scalability of users is an issue.	Evaluated on a real world data set.
5	The paper does not extensively discuss the framework's robustness to such real-world challenges, and further investigation is needed to assess its performance under diverse and unconstrained conditions.	The paper "Towards Explaining Recommendations through Local Surrogate Models" proposes a method for generating local surrogate models that can provide explanations for individual recommendations.

V. EXPECTED OUTCOMES

Explainable recommendation systems aim to provide users with explanations for the recommendations they receive. These systems have several expected outcomes, including:

1. ***Increased user trust***: By providing explanations for recommendations, users can better understand why certain items or options are being recommended to them. This transparency increases trust in the system and helps users feel more confident in the recommendations they receive.

2. ***Improved user satisfaction***: When users understand the rationale behind recommendations, they are more likely to be satisfied with the system's suggestions. They can assess whether the recommendations align with their preferences, needs, or constraints, leading to a higher likelihood of finding relevant and valuable recommendations.

3. ***Enhanced user engagement***: Explanations can captivate users' attention and stimulate their interest in exploring the recommended items. Users may engage more actively with the system, delve deeper into recommended content, and provide feedback, leading to a richer and more interactive user experience.

4. ***Empowered decision-making***: Explanations help users make more informed decisions by shedding light on the factors considered by the recommendation system. Users can evaluate the criteria used, understand trade-offs, and refine their preferences accordingly. This empowers users to make choices that align better with their specific requirements and preferences.

5. ***Mitigation of biases and unfairness***: Explanations can expose any biases or unfairness that may exist in the recommendation system. Users can detect and understand if certain recommendations are influenced by factors like demographic characteristics or hidden agendas. This transparency enables users to hold the system accountable and helps developers identify and address potential biases.

6. ***User feedback and system improvement***: Explanations facilitate user feedback by enabling users to express their agreement, disagreement, or confusion regarding the recommendations. This feedback loop can provide valuable insights for system developers to refine and enhance the recommendation algorithms, making them more accurate, personalized, and aligned with user expectations.

7. ***Regulatory compliance***: In some domains, such as finance or healthcare, regulatory compliance may require recommendation systems to provide explanations for their suggestions. Explainable recommendation systems can help organizations meet these compliance requirements by offering transparent and auditable decision-making processes.

Overall, explainable recommendation systems have the potential to enhance user satisfaction, trust, and engagement while promoting transparency, fairness, and accountability. By enabling users to understand and evaluate recommendations, these systems aim to bridge the gap between users and algorithms, fostering a more user-centric and trustworthy recommendation experience.

VI. CHALLENGES

While explainable recommendation systems offer various advantages, they also have certain limitations that should be considered. Some of the limitations include:

- Complexity of Explanations: Generating meaningful and interpretable explanations for recommendation systems can be challenging, especially in complex algorithms or when using deep learning techniques. As the complexity of the underlying model increases, it becomes more difficult to provide concise and understandable explanations to users.
- User Perception and Cognitive Load: While explanations can improve user understanding and trust, not all users may value or comprehend the provided explanations. Some users might find explanations overwhelming or irrelevant, leading to increased cognitive load and potential user frustration.
- Diversity and Serendipity: Explainable recommendation systems tend to rely on user preferences and past behavior to generate recommendations. This can limit the system's ability to introduce users to new and diverse items or content that might not align with their previous choices. The emphasis on explanation may hinder the exploration of new options and serendipitous discoveries.
- Scalability and Efficiency: Generating explanations in real-time for large-scale recommendation systems can be computationally expensive and time-consuming. As the volume of data and user interactions increases, the system's efficiency may be compromised, resulting in delays in generating recommendations and explanations.
- Privacy Concerns: Providing detailed explanations might involve exposing sensitive user information or revealing proprietary algorithms and data. Balancing the level of transparency with the need for privacy protection is a challenge that needs to be addressed.
- Subjectivity and Interpretability: Explanations can be subjective, varying based on individual preferences and interpretation. Different users may require different types or levels of explanations, making it challenging to find a one-size-fits-all approach.

  Addressing these limitations requires a careful balance between accuracy, transparency, user experience, and system efficiency. It involves ongoing research and development to advance the field of explainable recommendation systems and overcome these challenges.

VII. CONCLUSION

In conclusion, explainable recommender systems (XRS) play a vital role in addressing the transparency and interpretability issues of traditional recommendation systems. By providing clear and understandable explanations of how recommendations are generated, XRS help users make informed decisions and build trust in the system.

The increasing prevalence of recommendation systems in our daily lives highlights the importance of ensuring their transparency and user-friendliness. XRS not only enhance the user experience but also offer businesses opportunities to optimize their recommendation strategies based on user feedback and preferences.

However, the development of XRS is not without challenges. Balancing the need for transparency with the risk of information overload, ensuring the accuracy of explanations, and addressing privacy concerns are some of the key challenges that need to be overcome.

Nonetheless, the potential benefits of XRS are significant. They empower users by providing them with insights into why specific recommendations are made, allowing them to have a better understanding of the underlying mechanisms. This increased transparency and user trust can lead to improved user satisfaction, engagement, and loyalty.

Moving forward, further research and advancements in XRS are essential. This includes exploring different explainability techniques, evaluating their effectiveness, and developing standardized evaluation metrics. By continuously improving the explainability of recommendation systems, we can enhance user understanding and trust, ultimately fostering a positive user experience in the realm of personalized recommendations.

REFERENCES

[1] Yao, J., Huang, C., Zhang, L., & Yuan, Y. (2020). SHAP-Enhanced Counterfactual Explanations for Recommendations. In Proceedings of the 29th ACM International Conference on Information & Knowledge Management (pp. 2681-2684).

[2] Lei, C., Zhang, X., & Li, H. (2019). Online Product Feature Recommendations with Interpretable Machine Learning. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (pp. 1775-1784).

[3] Chakraborty, S., Chakraborty, S., & Ghosh, J. (2019). Towards Explaining Recommendations Through Local Surrogate Models. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (pp. 1753-1763).

[4] Lundberg, S. M., & Lee, S. I. (2017). A Unified Approach to Interpreting Model Predictions. In Advances in Neural Information Processing Systems (pp. 4765-4774).

[5] Wachter, S., Mittelstadt, B., & Russell, C. (2017). Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR. Harvard Journal of Law & Technology, 31(2), 841-887.

[6] Guolin Ke, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, and Tie-Yan Liu. 2017. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. In Advances in Neural Information Processing Systems 30. Curran Associates, Inc., 3146–3154.http://papers.nips.cc/paper/6907-lightgbm-a-highlyefficient-gradient-boosting-decision-tree.pdf

[7] Christopher M. Bishop. 2006. Pattern Recognition and Machine Learning (Information Science and Statistics). Springer-Verlag, Berlin, Heidelberg

[8] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. 2001. The Elements of Statistical Learning. Springer New York Inc., New York, NY, USA.

[9] Wang-Cheng Kang, Eric Kim, Jure Leskovec, Charles Rosenberg, and Julian McAuley. 2019. Complete the Look: Scene-based Complementary Product Recommendation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 10532–10541.

[10] Kevin Leyton-Brown and Yoav Shoham. 2008. Essentials of Game Theory: A Concise Multidisciplinary Introduction. In Synthesis Lectures on Artificial Intelligence and Machine Learning. pp 1–88.

[11] Pasquale Lops, Marco de Gemmis, and Giovanni Semeraro. 2011. Content-based recommender systems: state of the art and trends. In Recommender Systems Handbook. 73–100.

[12] Scott M Lundberg and Su-In Lee. 2017. A Unified Approach to Interpreting Model Predictions. In Advances in Neural Information Processing Systems 30. Curran Associates, Inc., 4765–4774. http://papers.nips.cc/paper/7062-a-unifiedapproach-to-interpreting-model-predictions.pdf

[13] Sameer Singh Marco Tulio Ribeiro and Carlos Guestrin. 2016. "Why Should I Trust You?": Explaining the Predictions of Any Classifier. In Proceedings of the 22th ACM SIGKDD. 1135–1144.

[14] Julian McAuley, Rahul Pandey, and Jure Leskovec. 2015. Inferring networks of substitutable and complementary products. In Proceedings of the 21th ACM SIGKDD. 785–794.

[15] Christoph Molnar. 2019. Interpretable Machine Learning. https://christophm. github.io/interpretable-ml-book/.

[16] Liudmila Prokhorenkova, Gleb Gusev, Aleksandr Vorobev, Anna Veronika Dorogush, and Andrey Gulin. 2018.

[17] CatBoost: unbiased boosting with categorical features. In Advances in Neural Information Processing Systems 31. 6638–6648.

[18] Moumita Sinha and Rishiraj Saha Roy. 2014. Preference Mapping for Automated

Recommendation of Product Attributes for Designing Marketing Content. In CBRecSys 2014. No. 10.

[19] Shoujin Wang, Liang Hu, Longbing Cao, Xiaoshui Huang, Defu Lian, and Wei Liu. 2018. Attention-Based Transactional Context Embedding for Next-Item Recommendation. In Proceedings of the Thirty-Second AAAI Conference on Artificial

[20] Behnoush Abdollahi and Olfa Nasraoui. 2016. Explainable Matrix Factorization for Collaborative Filtering. In *Proceedings of the 25th International Conference Companion on World Wide Web (WWW '16 Companion).* International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland, 5--6.

[21] Gediminas Adomavicius and Alexander Tuzhilin. 2008. Context-aware Recommender Systems. In *Proceedings of the 2008 ACM Conference on Recommender Systems (RecSys '08).* ACM, New York, NY, USA, 335--336.