Featuristic: An interactive hybrid system for generating explainable recommendations – beyond system accuracy

Sidra Naveed
University of Duisburg-Essen
Duisburg, Germany
sidra.naveed@uni-due.de

Jürgen Ziegler
University of Duisburg-Essen
Duisburg, Germany
juergen.ziegler@uni-due.de

ABSTRACT
Hybrid recommender systems (RS) have shown to improve system accuracy by combining benefits from the collaborative filtering (CF) and content-based (CB) approaches. Recently, the increasing complexity of such algorithms has fueled a demand for researchers to focus more on the user-oriented aspects such as explainability, user interaction, and control mechanisms. Even in cases, where explanations are provided, the systems mostly fall short in explaining the connection between the recommended items and users' preferred features. Additionally, in most cases, rating or re-evaluating items is typically the only option for users to specify or manipulate their preferences. With the purpose to provide advanced explanations, we implemented a prototype system called *Featuristic*, by applying a hybrid approach that uses content-features in a CF approach and exploits feature-based similarities. Addressing important user-oriented aspects, we have integrated interactive mechanisms into the system to improve both preference elicitation and preference manipulation. Besides, we have integrated explanations for the recommendations into these interactive mechanisms. We evaluated our prototype system in two user studies to investigate the impact of the interactive explanations on the user-oriented aspects. The results showed that the *Featuristic System* with interactive explanations have significantly improved users' perception of the system in terms of the preference elicitation, explainability, and preference manipulation – compared to the systems that provide non-interactive explanations.

Author Keywords
Hybrid Recommender System; Explanations; Interactive Recommending; User Experience

CCS Concepts
•**Information systems** → **Recommender systems;**

INTRODUCTION

Recommender systems (RS) based on *Collaborative Filtering* (CF) or *Content-based* (CB), have been mainly focusing on improving the accuracy of predictions, by mostly using ratings provided by users for items. Recently, with the increasing complexity of RS algorithms, the user-oriented aspects have gained more attention from the research community. It has been shown that improving these aspects lead to a commensurate level of user satisfaction and user experience with the system [18, 33].

One of these aspects that may contribute to the actual user experience is the degree of control users have over the system and their preference profiles [16, 14, 18]. Yet, from a user's perspective, today's automated RS such as the ones used by Amazon [19] or Netflix [3], provide limited ways to influence the recommendation generation process. Usually, the only means to actively influence the results is by rating or re-rating single items, which raises the risk of users being stuck in a "filter bubble" [6, 29, 42, 28]. This effect makes it difficult for users to explore new areas of potential interest and to adapt their preferences towards the situational needs and goals [25].

Additionally, another problem can be seen in the general lack of explainability in most of the current RS which could negatively impact user's subjective system assessment and overall user experience. For instance, lack of explanations could result in the difficulty of understanding recommendations which maybe a hindrance for users to make their decisions [41, 22]. These aspects consequently negatively effect the overall user experience with the system [33]. Moreover, it is often unclear to users that how their expressed preferences actually correspond to the system's representation of the user model i.e. how manipulating the preference model affects the system's output [36, 46]. Hence, adding more interactivity to the system by letting users influence their recommendation processes and preference profiles is considered a possible solution in RS research to improve the system's explainability [16, 14, 18]. In this regard, only presenting users with the matching recommendations is not very supportive and it has been observed that users require additional information and interactive mechanisms to fully benefit from the system [43].

To address the limitations of state-of-the-art CF and CB approaches, limited hybrid approaches exist that focus on user-oriented aspects and user experience, beyond the algorithmic accuracy [6, 20, 30]. But such approaches are still limited in terms of providing explanations, as the connection between the recommended items and the user's preferences for item-features are not clearly explained to the user. Additionally, these systems rarely explore whether a combination of expla-

IntRS '20 - Joint Workshop on Interfaces and Human Decision Making for Recommender Systems, September 26, 2020, Virtual Event

nations with interaction tools, has a positive influence on the user-oriented aspects or not.

In this paper, we implemented an interactive hybrid system called *Featuristic* in the domain of digital cameras, that exploits content-features in a CF approach. The recommendations and corresponding explanations are generated based on users that are similar to the current user in terms of shared feature-based preferences. The implemented approach is inspired from the approach proposed in [26]. We exploited multiple data sources to provide explainable recommendations rather than relying only on item-ratings (CF approach) or item-features (CB approach). We further integrated these advanced explanations with interactive mechanisms with the purpose to improve the proposed prototype system with respect to three main user-oriented aspects: 1) Preference elicitation process 2) Explainability of recommendations and, 3) Preference manipulation of users. In this regard, we aim at addressing the following research question:

RQ: Does integration of the hybrid-style explanations with interaction tools improve the preference elicitation, explainability of recommendations, and preference manipulation for users – compared to a conventional filtering system with simple and non-interactive explanations?

To address the research question, we ran a user study in which we evaluated the *Featuristic System* with advanced interactive explanations, against conventional filtering approach with rather simple and non-interactive explanations. In a subsequent study we successfully evaluated the results of our first study, by isolating the affect of the underlying algorithms and only focusing on the affect of interactive explanations on the user-oriented aspects. For this purpose, we compared two versions of our prototype systems with or without interactive explanations.

RELATED WORK

Among other user-oriented factors, increasing the transparency of the RS has proved to improve the perceived recommendation quality, decision support, trust, overall satisfaction and higher acceptance of the recommendations [47, 33, 41, 22]. Several studies have investigated the aspect of transparency, by comparing different explanation styles [4], combining different explanation styles [38], considering factors like personalization [39, 40], tags [46], rankings [24], and natural language presentations [9]. However, the current RS often lacks in explaining to users; *how* a system generates recommendations and *why* it recommends certain items [35, 41].

In the context of CB approaches, for instance, item attributes can be used to textually explain the relevance of recommended items to the users' personal preferences, though it requires availability of content data. The most common example of such explanations is *Tagsplanation* where the recommended movies are explained based on the user's preferred tags, explaining how the movies are relevant to these users' preferred tags [46]. Billsus et al. [5] proposed a news RS where the explanations are presented by means of textual keywords.

In case of conventional CF approaches, users and items are represented through vectors containing the item-ratings. The algorithm tries to predict the missing ratings of the item, which have not been rated by the users, yet based on, for instance, the weighted average of the ratings provided by similar users (user-based CF) or of similar items (item-based CF). Explaining these predictions to users is sometimes, very complicated and might be difficult for users to understand. Herlocker et al. recognized this problem and compared 21 different explanation interfaces for CF, for getting an understanding of how users with similar tastes rated recommended items [13]. Their study indicated that users preferred rating histograms over other explanation styles. Numerous attempts have been made to increase the transparency of the RS through visual explanations such as; *flowcharts* [15], *Venn diagrams* [30], *graph-based* representations [45], *clustermaps* [45], *concentric circles* [17, 27], *paths among columns* [6], and *map visualizations* [23, 10]. Approaches such as, *PeerChooser* [27] and *SmallWorlds* [11] presented complex interactive visualizations with the aim to explain the output of CF: similar users are displayed by means of connected nodes, where the distance between the nodes reflects the similarity between two users.

Hybrid approaches have emerged to benefit from both CF and CB approaches when generating recommendations and its corresponding explanations [7]. Some of these approaches combined ratings with content features [37, 12], and others have additionally taken social data into account [6, 30, 44, 34]. However, these systems rarely focus on making the recommendation process more transparent and explainable. In cases where they attempt to provide explanations, these explanations are mostly presented visually. A prominent example is *Talk Explorer* [44] which uses cluster maps allowing the user to explore the connections of conference talks to user bookmarks, user tags, and social data. *SetFusion* [30] is the hybrid system which is based on *TalkExplorer* which uses Venn diagrams instead of cluster maps.

The aspects of user control and interactivity have also been integrated in the hybrid systems. A common example of such systems is *Tasteweights* [6] that exploit social, content, and expert data to provide interactive music recommendations. The system not only visually presents the relation between the user profile, data sources, and recommendations but it also allows the user to manipulate their recommendation process by changing weights associated with individual items and by expressing their relative trust for each context source. These interactions are dynamically reflected in the recommendation feedback in real time. In the same context, *MyMovieMixer* [21] is the hybrid approach that allow users to control their recommendation process. The system provides immediate feedback, highlighting the criteria used to generate the recommendations. *MoodPlay* [1] is an other example that combines content- and mood-based data for recommending music. Recommendations and an avatar representing the user profile is displayed in terms of visualization, enabling the user to understand why certain songs are recommended by means of the position in the latent space, presenting the relation to different moods, and allowing the user to influence the recommendation process by moving the avatar [1]. While these works have attempted to increase the transparency, user control, and interactive mechanisms, mostly including advanced visualizations, they usually fall short of explaining the connections between user preference profile in terms of item-features and the relevance of recommended items to this profile. Additionally, users are provided with limited mechanisms to modify their preference profile or manipulate their recommendation process – mostly in terms of rating or re-rating items. Current work aims to focus on the user-oriented aspects by combining the advanced hybrid-style explanations with interaction mechanisms.

A HYBRID SYSTEM BASED ON FEATURE-BASED SIMILARITY

Following steps are used to implement the hybrid approach and are briefly discussed here. 1) Creating feature-based profile of the current user 2) Creating other users' feature-based profiles – implicitly predicted from their item-ratings 3) Computing user-user similarities based on shared feature preferences 4) Generating recommendations and corresponding explanations from similar users' feature-based preferences.

Creating feature-based profile of the current user

In the first step, a feature-based profile of the current user is required to be used in a feature-based CF approach. For this purpose, first the user is required to select the feature-value and then must specify how important this value is for him/her in terms of five-point likert-scale (from "not important" = 0; "very important" = 1). For binary features e.g., *WLAN*, selecting a feature and giving it an importance scale, will add this feature in a user vector. In case of continuous features such as *Pixel Number*, the user can select any range-value and select the importance scale, which will be discussed in the section *"Similarity between users in terms of continuous features with range-value categories"*, specifying how these values will be mapped and saved in the user vector.

Additionally, we used the knowledge based data from a camera website[1] to identify the set of features which are important for the five most common photography modes i.e., sports, landscape, Filming, street, and portrait photography. The current user can explore any photography mode in terms of the pre-defined set of features associated with each mode. The current user can select one of the photography modes, with an option to exclude any feature from the features-set for that mode or can add the entire set of features directly into his/her preference profile as part of the mode. To increase the control over the system and to enable users to adjust their profile at any time, both the feature values and the importance scores can be adjusted.

Predicting feature-weights for users using ordinal regression model

The second step is to compute feature-based profiles of all other users by implicitly predicting from item-ratings. There are several techniques proposed in the literature to predict feature-weightings from item-ratings including *TF-IDF (Term Frequency- Inverse Document Frequency)* method and entropy-based feature-weighting method proposed in [8, 2]. On one hand, the *TF-IDF* does not provide satisfactory results as the item has mixed data type features. On the other hand, an entropy-based feature-weighting method is also limited in terms of computing the relevance between two continuous features with mutual information due to the problem of loss of information during the process of discretization in order to transform non-nominal to nominal data [48].

To overcome the limitation of entropy-based feature-weighting method, we applied ordinal regression model, which can predict an ordinal dependent variable (i.e. item-ratings in terms of five-point likert scale) given one or more categorical or continuous independent variables (i.e. item-features). The model is able to determine which of the features have a statistically significant effect on the item-ratings. The model allows to determine, how much of the variation in item-ratings can be explained by item-features and also, the relative contribution of each feature in explaining this variance. The steps applied for ordinal regression model are briefly described below.

Selecting specific features for the model

When constructing a regression model, it is important to identify the predictor variables (item-features) that contribute significantly to the model. To do so, correlation of the item-features with the item-ratings are computed on the overall ratings dataset, by applying Spearman's rank-order correlation. The top 15 features with highest significant correlations with the ratings are further considered for the model.

Predicting ratings from features

In the next step, *PLUM* procedure is used in *SPSS* to apply an ordinal regression model[2] . For each user in the dataset, the model was applied separately, taking only values into account which have a significant correlation with the user ratings.

Interpreting the output

For each user, we want to determine which features have a statistically significant effect on the item-ratings. For this purpose, parameter estimates table is used to interpret the results and identify the features and its values that have statistically significant effect in predicting the item-ratings, as well as the contribution of each feature-value in predicting this rating.

Computing user-user similarity based on feature-preferences

The feature-based profile explicitly created for the current user and implicitly computed using ordinal regression for all other users, are then used to identify peer users with similar taste in item-features as that of the current user. As the camera features are of mixed data type, categorical and continuous – separate measures have been considered, which take the data type into account when computing similarity between two users and is further explained below.

Similarity between users in terms of categorical features

To compute similarity between two users in terms of categorical feature-values and their corresponding weightings, we

[1]https://cameradecision.com/

[2]The technical details and steps applied in SPSS for PLUM procedure can be found in the link: https://statistics.laerd.com/spss-tutorials/ordinal-regression-using-spss-statistics-2.php

applied *Mean Squared Error (MSE)* which provides a quantitative score describing the degree of dissimilarity between two profiles.

Similarity between users in terms of continuous features with range-value categories

In case of continuous features with range-values, the traditional similarity measures fail to address the question that whether the partial presence of the range-value be treated as presence or absence of the feature or not? To address this issue, we computed similarity between two user vectors in terms of the continuous features with range-value categories, in a two step process.

1) Percentage similarity measure: For applying regression model, the continuous values are categorized into fixed predefined bins, where each binned category gets different weights for the respective user (section *"Predicting feature-weights for users using ordinal regression model"*). As the active user can select any customized range value that might not exactly correspond to these binned categories, we expressed the customized range selected by the active user, as a percentage at which it is expressed in each binned category. If the range-value is completely covered by a binned category, then it is assigned a value of 1, and 0 if it is not covered at all. For partially covered range value in a binned category, the percentage similarity is computed using one of the given formulas by matching each condition:

$$\begin{cases} if\, v_{min} < vi < vj < v_{max};\;\; i_{cu,f} = \frac{v_j - v_i}{v_{max} - v_{min}} * r_{cu,f} \\ elseif\, v_i < v_{min} < v_j < v_{max};\;\; i_{cu,f} = \frac{v_j - v_{min}}{v_{max} - v_{min}} * r_{cu,f} \\ elseif\, v_{min} < v_i < v_{max} < v_j;\;\; i_{cu,f} = \frac{v_{max} - v_i}{v_{max} - v_{min}} * r_{cu,f} \end{cases} \tag{1}$$

Here $[v_i, v_j]$ are the range values selected by the current user (cu) and $[v_{min}, v_{max}]$ are the minimum and maximum values of the binned category. To compute the importance weighting $i_{cu,f}$ of each binned category for the current user, we multiplied the computed percentage similarity with the current user's feature-specific weight for the selected range $r_{cu,f}$.

2) Applying MSE on percentage similarity : Once the current user's range values are mapped in terms of percentage at which it is expressed for each binned category, then the dissimilarity between current user and other user in terms of categories defined by range-values, is computed by applying MSE on these computed values.

Generating item recommendations

The final dissimilarity score between the active user and the other in terms of categorical and continuous features is computed by taking the average of the scores computed in section *"Computing user-user similarity based on feature-preferences"*. The 10 users with lowest MSE scores are considered for the recommendation process. From these similar users' profiles, the highest rated items are considered as potential list of recommendation. However, to filter out the items from this list, that not only matches the active user's feature preferences (user-item similarity) but also matches the feature requirements for the preferred photography mode (item-mode similarity), we applied post-filtering mechanisms in three-step process to generate a final list of recommendation.

Gower's similarity measure for categorical features

To compute similarities between the current user's preferred features and the potential items in terms of categorical features, we applied Gower's similarity measure that takes the type of variables into account. Details of the method can be found in [31]. Let the current user be defined by $cu = \{cu_f | f = 1, 2, ..., F\}$ and the item is defined by $item = \{item_f | f = 1, 2, ..., F\}$. The similarity between two profiles is computed using the Gower's similarity measure using the formula:

$$S_{(cu,item)} = \frac{\sum_{f=1}^{F} s_{(cu,item)_f} * \delta_{(cu,item)_f}}{\sum_{f=1}^{F} \delta_{(cu,item)_f}} \tag{2}$$

The similarity coefficient $\delta_{(cu,item)_f}$ determines whether the comparison can be made for the f-th feature between cu and item which is equal to 1 if comparison can be made between two objects for the feature f and 0 otherwise. $s_{(cu,item)_f}$ is the similarity coefficient that determines the contribution provided by the f-th feature between cu and item, where the way this coefficient is computed depends on the data type of features i.e., categorical and numeric. In case of categorical features i.e., nominal or ordinal, the coefficient gets a value 1 if both objects have observed the same state for the feature f and is 0 otherwise.

Linear modification of Gower's similarity measure for continuous features

The second step of the post-filtering process for item recommendations is to compute the similarity between the current user (cu) and the item in terms of continuous features, where the cu has a range-value and the item has one discrete value for the feature f. In this case, the Gower's coefficient of similarity $s_{(cu,item)_f}$ for the numeric feature fail to address the issue as it takes only one distinct value for each object [31].

To deal with this limitation, we proposed a linear modification of Gower's similarity coefficient $s_{(cu,item)_f}$ by computing a similarity score that is linearly decreasing with a feature-value's distance from the user's desired range if it is outside this range. The idea is to assign a similarity score to the feature of the item depending on how close the value is to the active user's selected range. Let v be the distinct value of feature f in an item, $[v_i, v_j]$ is the min and max values of range selected by an active user, and $[v_{min}, v_{max}]$ are the min and max value available in the dataset for the feature f. The linear function for Gower's similarity coefficient $s_{(cu,item)_f}$ is then computed using one of the given formulas by matching each condition:

$$\begin{cases} if\, v_{min} < v < v_i;\;\; s_{(cu,item)_f} = \frac{v - v_{min}}{v_i - v_{min}} \\ elseif\, v_i < v < v_j;\;\; s_{(cu,item)_f} = 1 \\ elseif\, v_j < v < v_{max};\;\; s_{(cu,item)_f} = (\frac{v_j}{v_{max} - v_j} + 1) + (\frac{-v}{v_{max} - v_j}) \end{cases} \tag{3}$$

The final user-item similarity score for current user's all selected features is then computed by putting the values of the respective similarity coefficient $s_{(cu,item)_f}$ for categorical and numeric features (computed in section *"Gower's similarity measure for categorical features"* and *"Linear modification of Gower's similarity measure for continuous features"*) and $\delta_{(cu,item)_f}$ in equation 3 and the top 10 items are then selected for recommendation.

FEATURISTIC: PROTOTYPE AND INTERACTION POSSIBILITIES

To implement the prototype system based on the method described in section 3, we collected our own explicit item-ratings data set. For this purpose, we conducted an online study on Amazon Mechanical Turk (AMT) [3] users by providing them with 60 digital cameras where each camera was described in terms of a list of 90-95 features extracted from a website with editorial product reviews [4]. Each participant was asked to evaluate at least 20 cameras in terms of five-star rating based on the available features, which resulted a total of 5765 ratings on 60 cameras by 150 users. The implemented prototype system called *Featuristic* is shown in Figure 1, which extends the conventional CF and CB approaches in terms of three main aspects as described below:

Preference Elicitation

Conventional CF or CB approaches, mostly elicit users' preferences for items in terms of rating or re-rating single items. The filtering process of such approaches often assumes that all features are equally important for users and does not take that aspect into account. In the *Featuristic System*, we elicit the new user's preferences for item-features by explicitly asking the user to select the preferred feature-values and indicates the importance of the feature-value using the importance slider (Figure 1a). This enables users to specify their preferences more precisely, especially in high-risk domains, e.g., digital cameras, where the features of items play a vital role in users' decision-making processes. The system further assists users in indicating their preferences more clearly especially, when users have limited domain knowledge, their preferences are not defined, or they are unaware of the context in which the camera can be used. This is done by providing users with an option to indicate their preferences for one of the five most common photography modes (Figure 1b). The system provides users with features-set along with the suggested values, explaining why these features with certain values are important for a particular mode (Figure 2a).

Explainable Recommendations

Current CF or CB approaches fail to explain the connection between recommended items and the user's preferences of item-features. This is addressed in the *Featuristic System* by showing a table that compares the features of each recommended item with the user's preferred features (Figure 1c).

Additionally, it is mostly unclear to users how their expressed feature preferences actually correspond to the system's representation of their preference models. Even in cases, when the explanations are provided, the rationale behind recommendations is mostly not explained to users. The *Featuristic System* visually explains how users are similar to the current user in terms of shared feature preferences (Figure 1d) and how recommendations are generated based on similar users' feature-based profiles.

Most of the current RS do not provide any insight into the distribution of the feature-values in the feature-space or even the availability of the offered items distributed over the feature-space. This might be useful for users to detect relevant features and to inform their own decision by thoroughly narrowing down the list of items based on the item-features. In *Featuristic System*, this aspect is integrated by showing the distribution of feature-values selected by similar users (Figure 1d). Then, the recommended items are mapped on top of this distribution (Figure 1e). This visually explains how the recommended items are generated from similar users' feature-based profiles, as most of the recommended items lie within most preferred feature-values by similar users.

As in the *Featuristic System*, users can indicate their preferences for one of the five photography modes – the approach also considers the features-set for the selected mode in computing similar users. For each item, the suitability score for each mode is computed and can be explored by clicking on the "suitability for other modes" which opens a bar chart in a pop-up window (Figure 2b). Clicking on any bar would expand the window with explanation of how the scores are computed in terms of one-to-one comparison of features of items with the required features of the mode.

Manipulation of Preferences

In most conventional CF approaches, the only way for users to indicate or modify their preferences is by (re)rating items. In case of the filtering systems, users can specify their preferences by selecting the desired value or value-range for a specific attribute of the items. In complex domains e.g., digital cameras where users mostly lack precise knowledge of the domain, providing explanations can be considered an important factor. On the other hand, providing interactivity and direct manipulation within an explanation might offer users a flexible and comprehensible way to manipulate their preferences.

In this respect, the *Featuristic System* integrates sliders (for continuous features) and toggle buttons (for binary features) with the explanations (Figure 1g), to facilitate the direct manipulation of preferences from the system provided explanations. The interactive explanations are further combined with recommendations – visually showing the location of the recommended items distributed over the feature-space (Figure 1e). The system allows the users to manipulate their preferences directly from the explanations, by either changing the feature-value or feature-rating – which results in dynamically updating recommendations.

EMPIRICAL EVALUATION

To investigate the impact of the explanation method developed when integrated with interaction mechanisms, on user oriented aspects, we designed a user study. Accordingly, we formulated the hypotheses with respect to user-oriented aspects focusing on, preference elicitation (H1), explainable recommendations

[3]https://www.mturk.com/
[4]https://www.test.de/

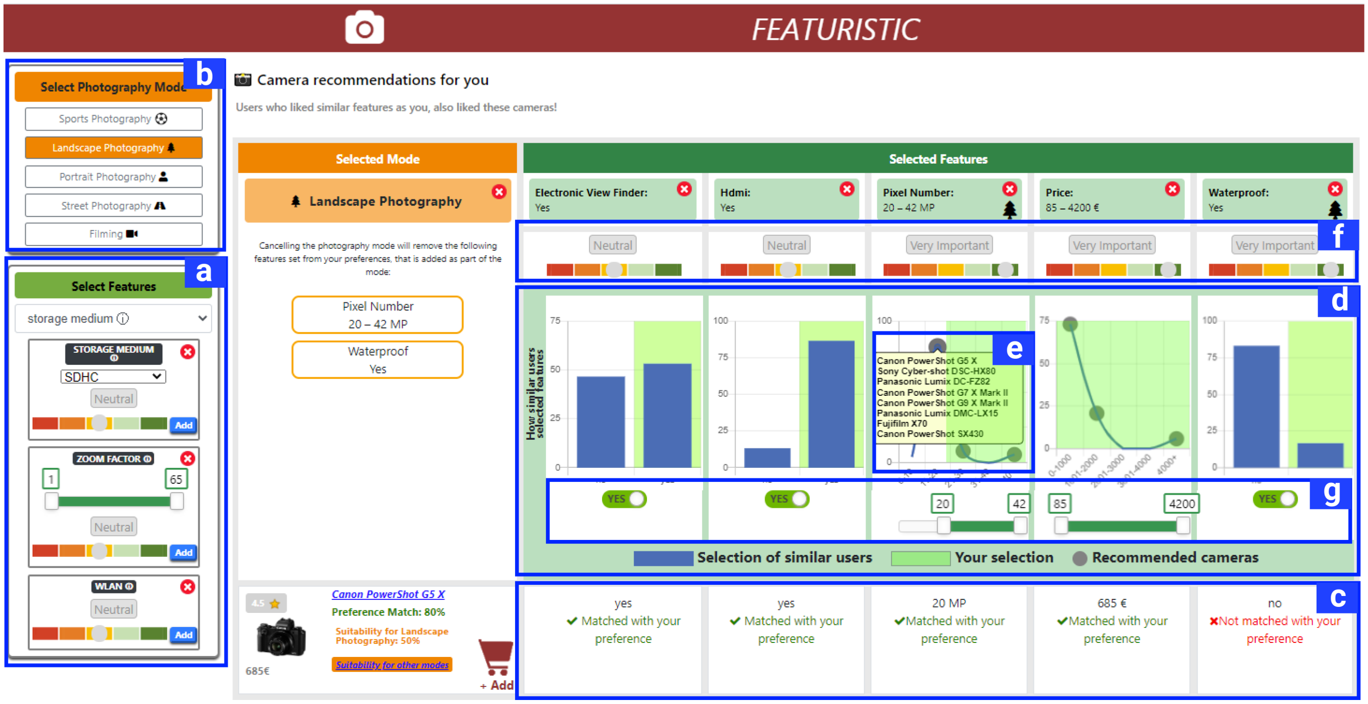

Figure 1: Screenshot of the Featuristic system. Filtering Area for selecting features (a) and choosing modes (b); One-to-one comparison of the recommended item with the user's selected feature-values (c); Graphical explanations showing the comparison of the current user's shared feature preferences with similar users (d); Recommended items mapped on top of the similar users' selection (e); Sliders to modify importance of feature-value (f); Sliders and toggle buttons to modify feature-value (g).

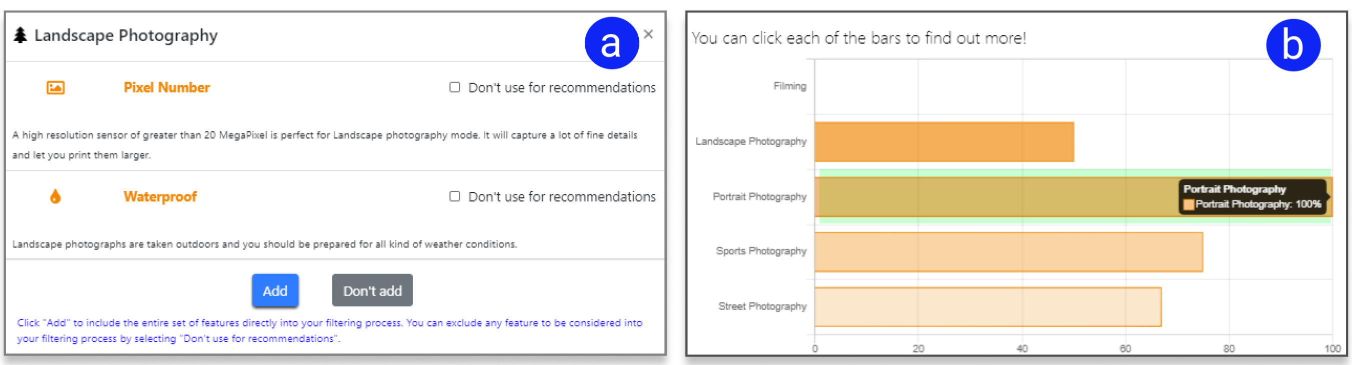

Figure 2: (a) shows the list of features along with an explanation of why these features are required for the photography mode, (b) shows the suitability scores of all modes for the recommended item based on the available features in the item.

(H2a and H2b), preference modification (H3a and H3b), and user experience (H4).

Hypotheses:
Integrating the feature-based CF style explanations with interaction tools when compare to a conventional filtering system, leads to:

- **H1:** More concrete preference elicitation
- **H2a:** Better explained recommendations
- **H2b:** More comprehensible recommendations
- **H3a:** More direct manipulation of user preferences
- **H3b:** More controllable manipulation of user preferences
- **H4:** An improved user experience

User Study 1
To address our hypotheses, we conducted an online crowd-sourced study via Prolific[5]. In this study, the *Featuristic System* that provides advanced interactive explanations is compared with the conventional *Filtering System* that only provides simple and non-interactive explanations.

[5] https://www.prolific.co/

Table 1: Self-Created items used for the constructs during the user study.

Construct	Self-Created Items
Preference Elicitation	- The system allows me to indicate my preferences for the camera-features efficiently. - The system allows me to indicate my preferences for the camera-features precisely. - The system allows me to specify how important the specific camera-features are to me.
Understandability	- The information provided for the recommended cameras is easy to understand. - Overall, I find it difficult to understand the information provided for the recommended cameras.
Decision Support	- The information provided helps me decide quickly. - Overall, I find it difficult to decide which camera to select.
Direct Manipulation	- Seeing other users feature-selection helps me in modifying my preferred features. - I am able to determine suitable feature-values for my selection. - I am confident in modifying my selected feature-values. - I am able to directly compare features present in given recommendations with features that other users have selected. - I am able to directly see the recommended cameras that lie within my feature selection.

Table 2: Mean values and standard deviations for the subjective system assessment of the two conditions. Significant differences are marked by *. Higher values (highlighted in bold) indicate better results.

	User Study 1 (df=54)					Follow-up User Study (df=36)				
Construct	**Featuristic**		**Conventional Filtering**			**Featuristic**		**Non-Interactive Featuristic**		
	M	*SD*	*M*	*SD*	*p*	*M*	*SD*	*M*	*SD*	*p*
Preference Elicitation (H1)	**3.90**	0.65	3.67	0.85	**.032***	3.80	0.85	**4.02**	0.64	**.006***
Transparency (H2a)	**4.26**	0.50	3.82	0.89	**.003***	3.89	0.80	3.89	0.85	>.999
Information Sufficiency (H2a)	**3.27**	0.87	2.94	0.85	**.019***	**4.02**	0.56	3.10	1.01	**<.001***
Understandability (H2b)	3.43	0.99	**3.64**	0.91	.069	**3.52**	0.95	3.48	1.04	.760
Decision Support (H2b)	3.10	1.03	**3.15**	1.04	.734	3.18	1.00	**3.33**	1.00	.260
Direct Manipulation (H3a)	**3.83**	.52	3.67	0.58	**.034***	**4.05**	0.47	3.76	0.60	**.016***
User Control (H3b)	3.93	.74	3.93	0.70	>.999	**4.40**	0.36	3.97	0.81	**.003***

Method

The study was conducted in a within-subject design, where participants were presented with two prototype systems in a counter-balanced order:

- **Featuristic System:** The interface design of the system is depicted in Figure 1. The interaction possibilities are further described in the section *"Featuristic: Prototype and Interaction possibilities"*.
- **Conventional Filtering System:** The system allowed participants to indicate preferences in terms of features by simply selecting feature-values. The system then generates recommendations and explanations only showing the comparison of recommended items with the participants' selected features and values (Figure 3A).

In each of the two resulting conditions, participants were provided with the same task scenario. In the system, they were first asked to indicate their preferences in terms of features according to the task scenario. The system then generates recommendations and corresponding explanations. Participants were required to explore the system recommendations and each of its presented explanations and functionality in order to understand the rationale behind the recommendations and its explanations and select camera(s) that matches their preferences according to the task scenario. After interacting with each system, they were then asked to evaluate the system by answering series of questions.

Participants and Questionnaire. A total of 55 Prolific users were recruited online (19 females) with age ranging from 18-54 years ($M = 28$, $SD = 8.7$). The study completion time was recorded approximately 15-20 minutes. To address our hypotheses, we mostly used the self-created items to evaluate both systems in terms of the above mentioned three aspects and are shown in Table 1. For *Preference Elicitation*, we used the self-created items. The aspect of *Explainable Recommen-*

dations was measured in terms of two sub-aspects i.e. *Explainability* (H2a) and *Comprehensibility* (H2b). For *Explainability*, we used the items related to *Transparency* and *Information Sufficiency* from [32]. For *Comprehensibility*, we used our self-created items related to *Understandability* and *Decision Support*. Furthermore, the aspect of *Preference Modification* was measured in terms of self-created items specifically related to the interactive mechanisms allowing the participants to directly manipulate their preferences (H3a). Additionally, we used items for *User Control* (H3b) taken from [32]. All questionnaire items were rated on a 1-5 Likert response scale.

Additionally, to test our fourth hypothesis, we used the short version of User Experience Questionnaire (UEQ) (7-point bipolar scale ranging from -3 to 3). For qualitative feedback, we provided open-ended questions asking the participants about their likes and dislikes for both systems in terms of the information provided on the interfaces.

Results

Hypothesis 1. To test our hypothesis, we conducted a one-way repeated measure ANOVA ($\alpha = 0.05$), revealing that *Featuristic* performed significantly better than the *Conventional Filtering System* for *Preference Elicitation*. Therefore, we can accept our H1, indicating that *Featuristic* leads to more concrete preference elicitation (Table 2).

Hypothesis 2a and 2b. To test H2a, which refers to the aspect of *Explainability* measured in terms of two sub-aspects i.e. *Transparency* and *Information Sufficiency*, we applied one-way repeated measure MANOVA ($\alpha = 0.05$). The results showed significant differences between two systems in terms of the two aggregated variables ($F(2,54) = 5.59$, $p < .006$, Wilk's $\lambda = 0.826$). Univariate test results further revealed that for both *Transparency* and *Information Sufficiency*, the *Featuristic* system significantly performed better than the Filtering system, indicating that the *Featuristic* leads to better explained recommendations. Therefore, we can accept H2a.

However, in terms of *Comprehensibility* (H2b) which is measured in terms of two sub-aspects i.e. *Understandability* and *Decision Support*, we found no significant differences between the two systems ($F(2,53) = 1.93$, $p < .15$, Wilk's $\lambda = 0.932$). Therefore, the Hypothesis 2b can not be accepted.

Hypothesis 3a and 3b. With respect to *Direct Manipulation of Preferences* (H3a), the result of one-way repeated measure ANOVA showed statistically significant difference between the two systems, where the Featuristic system performed significantly better than the filtering system as can be seen in Table 2. The result shows that the Featuristic system leads to more direct manipulation of user preferences, thus accepting our hypothesis 3a.

On the other hand, w.r.t. *User Control*, we found no significant difference between the two systems $F(1,54) = 0.00$, $p < 1.00$, Wilk's $\lambda = 1.00$, where surprisingly, both systems were perceived equally in terms of *User Control*. Therefore, the hypothesis 3b can not be accepted.

Hypothesis 4. To evaluate the systems with respect to the *User Experience*, we analyzed the different sub-scales of the UEQ, where we found no significant differences between the two systems. The Featuristic System received the following scores: 0.66 for pragmatic quality (Bad), 0.38 for Hedonic Quality (Bad), and 0.53 Overall (Bad). On the other hand, the Filtering System received the scores: 0.99 for Pragmatic Quality (Below average), 0.15 fro Hedonic Quality (Bad), and 0.58 Overall (Below Average). Yet, we can not accept this hypothesis.

Moreover, participants indicated their likes/dislikes for each system. When asked about the *Filtering System*, majority of participants liked the system because of its simple and clean design which is easy to understand and use the system and its functionalities. In comparison to the *Featuristic System*, some participants indicated their dislike about the *Filtering System* in terms of not being able to indicate the importance for the feature-values. For some participants the reason for not liking the *Filtering System* is because it does not show the graphs of features or does not include reviews from other people. On the other hand, when asked about the likes and dislikes for the *Featuristic System*, majority of participants liked the system because the system was clear, precise, intuitive, and innovative. Many participants liked the graph comparisons, where one participant indicated that *"The graphs feel like I have a more accurate decision"*, the other stated that: *"The graphs and the bar diagrams are innovative which is useful for more focused and serious buyers"*. Others also liked the option of selecting the importance of feature-values. Even though majority liked various functionalities of the *Featuristic System*, however, for some participants, the interface was quite complex with lots of information. One participant wrote that *"There is a lot of information for a novice"*. For some participants, the graphs were also difficult to understand.

Discussion.

The results show that the *Featuristic System* significantly improved the *Preference Elicitation* of users as compared to the *Filtering System* (H1). This might be due to the *Featuristic System's* ability, allowing users to not only select features and its values but also indicate the importance for each individual feature-value. This might have made the preference indication for users more precise and efficient as compared to conventional CF and CB systems, where it is mostly assumed that all features are equally important to users. This can also be reflected in participants' qualitative feedback. For example, one participant stated that *"I like specifying how important a feature was and not only if I wanted it or not"* and the other wrote *"I like being able to select how important a feature is with the sliding bar"*.

Additionally, we investigated the second main aspect of the *Featuristic System* i.e. *Explainable Recommendations* which is further measured in terms of two sub-aspects: *Explainability* (H2a) and *Comprehensibility* (H2b). With respect to *Explainability* (H2a), the *Featuristic System* is perceived significantly better than the *Filtering System*. This indicates, that the more advanced explanations in the *Featuristic System* made the recommendations more transparent and explainable for users which can be validated from the participants' qualitative feedback. One participant indicated that *"It gave you the information and segregation of data in an easy to read*

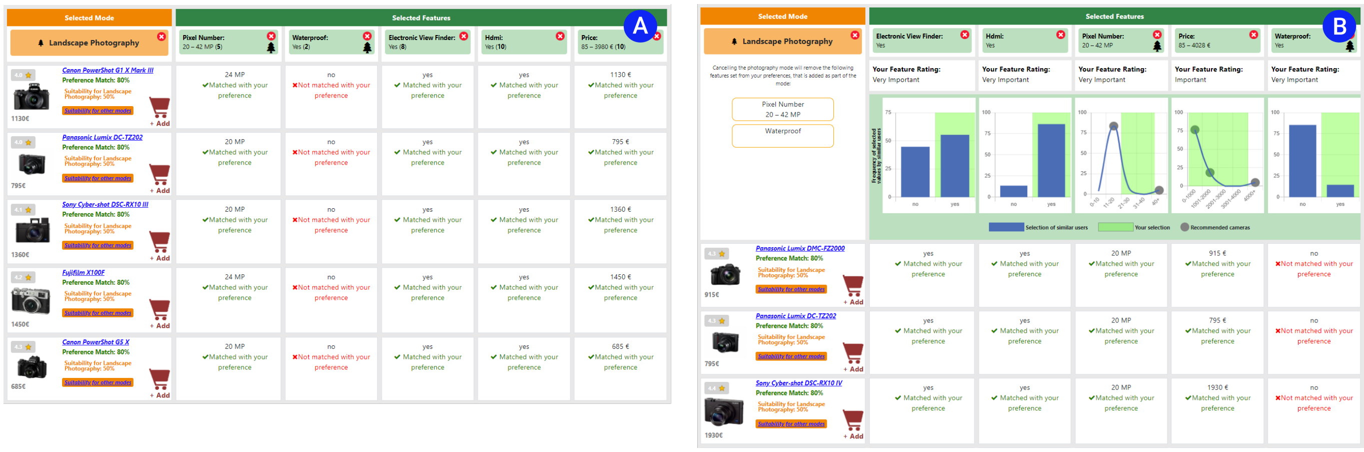

Figure 3: (A) Filtering System, (B) Featuristic System without interactive explanations.

(graphical) format.", where the other stated that *"You can see at a glance whether or not a specific camera has these features"*. For others it was useful to compare their choices with other users, where one participant wrote *"I love the fact that I had to compare my choices with recommendations of others"*. Another participant wrote *"I like the fact that the system brings other users' choice for me and also gave me detailed information about my search*. Additionally, we found no significant differences between two systems in terms of *Comprehensibility* (H2b) for aggregated variables, where the *Filtering System* showed slightly better results. This might be explained due to the fact that the two systems were quite different in terms of the functionalities and level of information provided. On one hand, the *Filtering System* provides rather simple and non-interactive explanations and on the other hand, the *Featuristic System* is more complex in terms of interactive functionalities and advanced explanations that it provided. Thus, making the *Filtering System* being perceived more comprehensible by users. This is also depicted in participants' qualitative feedback about the *Filtering System*, where they found the system much simpler, clean, and easy to understand as compare to the *Featuristic System*. For some of the participants, the *Featuristic System* provided too much information which is rather complex for them to comprehend.

With respect to *Direct Manipulation of Preferences* (H3a), the *Featuristic System* is perceived significantly better than the *Filtering System*, suggesting that integrating the interactive mechanisms with our explanations allowed users to directly manipulate their preferences through these explanations. Surprisingly, in terms of *User Control* (H3b), both systems are perceived of equal quality. As in the *Filtering System*, the only way provided to users to control the system's output is by selecting features or re-adjusting the feature-values. And it has been shown, that such user control mechanisms are easy to use compared to mechanisms that allow users to indicate the relative preferences [14] (e.g., feature-rating slider in *Featuristic*). In such cases, it is sometimes not clear to users if having the slider in the middle position has same meaning as having the slider at the maximum level. This might have made the interpretation of such control mechanisms complicated for users in the *Featuristic* and hence, not being perceived better by users compare to the *Filtering System*.

Additionally, for *User Experience* (H4), we found no significant differences between two systems. This might indicates that regardless of more advanced explanations with interactive mechanisms provided in *Featuristic* compared to the *Filtering System* with much simpler explanations, participants perceived both systems to be of similar quality in terms of the user experience. On the other hand, this might also be explained under the assumption that participants are different in terms of the domain knowledge and their ability to perceive and understand the system provided information and functionalities – as for some participants it might be easier to understand the information and its functionalities and for others too complicated. As stated by one of the participants about the *Featuristic System* that *"The information was easy to understand for me, but I can imagine less technical people would find information and graphics confusing."*

Follow-up User Study

To verify, that integrating the developed explanation method with interaction tools have positive impact on user-oriented aspects, which is independent of the types of underlying algorithms – we conducted a follow-up user study. In this study, we isolated the underlying algorithm by focusing only on the type of explanations provided. For this, we compared two versions of the *Featuristic System* that apply same underlying hybrid approach. The only difference is in terms of interactive and non-interactive explanations provided by the systems.

Method

The study was conducted via Prolific in a within-subject design and follows the same procedure and design as the first study described in section *"Featuristic: Prototype and Interaction possibilities"*. We again tested the same hypotheses described in section *"Hypotheses:"*, but this time, isolating the type of recommendation as the independent variable. We created two versions of the *Featuristic System*, described below:

- **Featuristic System:** The interface design and interaction its possibilities are described in *"User Study 1"* and shown in Figure 1.

- **Featuristic System without interactive explanations:** The prototype is similar to the one shown in the Figure 1. The only major difference is that the user is not provided with the functionality to modify or critique their selected feature-value or rating through graphical explanations of recommendation (See Figure 3B).

Participants and Questionnaire. A total of 37 Prolific users were recruited online (15 females) with age ranging from 18-50 years ($M = 24.86$, $SD = 6.9$). The study completion time was recorded approximately 15-20 minutes. To address our hypotheses, we used the same questionnaire items as in the first user study.

Results

To compare our two versions of Featuristic system, we applied one-way repeated measure MANOVA and the results can be seen in Table 2. With respect to *Preference Elicitation* (H1), the results showed significant difference, where the *Non-Interactive* version of the system is perceived significantly better than the *Interactive* version of the system. Therefore, we have to reject our H1.

For *Explainability* of recommendations (H2a) which is measured in terms of *Transparency* and *Information Sufficiency*, we found significant differences between two systems $F(2,35) = 16.30$, $p < .001$, Wilk's $\lambda = 0.518$, for aggregated variables. However, the result of univariate test showed significant difference only in terms of *Information Sufficiency*, where the *Interactive Featuristic* performed better. Overall, we can accept our H2a.

Regarding *Comprehensibility* (H2b), which is measured in terms of *Understandability* and *Decision Support*, we found no significant differences between two systems. Overall, we can not accept our H2b.

Additionally, in terms of *Direct Manipulation* and *User Control*, we again found significant differences between two systems, where the *Interactive Featuristic* performed significantly better than the *Non-interactive* system. Therefore, we can accept H3a and H3b. However, with respect to UEQ, we found no significant differences between the two systems, which leads to rejecting the H4.

Discussion

The results of the follow-up study showed, that in terms of *Explainability*, *User Control*, and *Direct Manipulation*, the *Interactive* version of *Featuristic* performed significantly better than the *Non-interactive version*. This clearly shows the positive impact of integrating interactive mechanisms with explanations, on these aspects. The results are similar to results of the first user study for most of the factors, where the *Interactive Featuristic* performed better. This verifies, that our advanced explanations showed positive impact on user-oriented aspects, independent of the underlying algorithms. The insignificant differences in terms of *Comprehensibility* and *User Experience*, might be due to the fact that both systems provided same functionalities and level of explanations. The only difference is with respect to the interactivity and non-interactivity of explanations. This might explain the reason for both systems being perceived equally in terms of *Comprehensibility* and *User Experience*. However, qualitative feedback showed that most of the participants like the interactive functionality of the *Featuristic System*. One participant stated that *" In my opinion, this system is more clear and clean than the other one. Although they look almost the same, I feel this one can be a bit more efficient. It is very helpful and intuitive".*

CONCLUSION AND OUTLOOK

In this paper, we showcased the possibility of integrating our proposed feature-based CF style explanations with interaction tools, through a prototype system called *Featuristic*. To study the impact from a user perspective in terms of *Preference Elicitation, Explainable Recommendations, Preference Manipulation, and User Experience*, we first compared our *Featuristic System* with the *Conventional Filtering System* that only provides simple and non-interactive explanations.The results showed that the *Featuristic System* is significantly perceived better than the *Conventional Filtering System* with respect to the aspects of *Preference Elicitation, Explainability, and Preference Manipulation*. However, we found no significant differences between the two systems in terms of the *User Experience* and *Comprehensibility*, which might be due to the complex structure of explanations and the system design, as stated by many participants in their qualitative feedback.

We further conducted a follow-up user study to verify, that the results from the first study are independent of the underlying algorithms. For this, we compared two versions of the *Featuristic System*, by isolating the types of underlying algorithms and only focusing on the type of explanations provided i.e., Interactive and Non-interactive explanations. The results showed that the Interactive version of *Featuristic* performed significantly better than the non-interactive version in terms of *Explainability, User Control, and Direct Manipulation*.

To summarize, the current work clearly showed the positive impact of integrating advanced explanations with interaction tools to improve the user-oriented aspects, especially in complex product domains. However, the current work has some limitation in terms of the complex system design which could further be simplified for improving the overall *User Experience*. Additionally, factors like user's cognitive effort and user experience with the product domain, might also impact the user perception of the system with respect to user-oriented aspects, and thus requires further investigation in future work.

REFERENCES

[1] Ivana Andjelkovic, Denis Parra, and John O'Donovan. 2016. Moodplay: Interactive mood-based music discovery and recommendation. In *Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization*. ACM, 275–279.

[2] Manuel J Barranco and Luis Martínez. 2010. A method for weighting multi-valued features in content-based filtering. In *International conference on industrial, engineering and other applications of applied intelligent systems*. Springer, 409–418.

[3] James Bennett, Stan Lanning, and others. 2007. The netflix prize. In *Proceedings of KDD cup and workshop*, Vol. 2007. New York, 35.

[4] Mustafa Bilgic and Raymond J Mooney. 2005. Explaining recommendations: Satisfaction vs. promotion. In *Beyond Personalization Workshop, IUI*, Vol. 5. 153.

[5] Daniel Billsus and Michael J Pazzani. 1999. A personal news agent that talks, learns and explains. In *Proceedings of the third annual conference on Autonomous Agents*. Citeseer, 268–275.

[6] Svetlin Bostandjiev, John O'Donovan, and Tobias Höllerer. 2012. TasteWeights: a visual interactive hybrid recommender system. In *Proceedings of the sixth ACM conference on Recommender systems*. ACM, 35–42.

[7] Robin Burke. 2007. Hybrid web recommender systems. In *The adaptive web*. Springer, 377–408.

[8] Jorge Castro, Rosa M. Rodriguez, and Manuel J. Barranco. 2014. Weighting of features in content-based filtering with entropy and dependence measures. *International journal of computational intelligence systems* 7, 1 (2014), 80–89.

[9] Shuo Chang, F Maxwell Harper, and Loren Gilbert Terveen. 2016. Crowd-based personalized natural language explanations for recommendations. In *Proceedings of the 10th ACM Conference on Recommender Systems*. ACM, 175–182.

[10] Emden Gansner, Yifan Hu, Stephen Kobourov, and Chris Volinsky. 2009. Putting recommendations on the map: visualizing clusters and relations. In *Proceedings of the third ACM conference on Recommender systems*. ACM, 345–348.

[11] Brynjar Gretarsson, John O'Donovan, Svetlin Bostandjiev, Christopher Hall, and Tobias Höllerer. 2010. Smallworlds: visualizing social recommendations. In *Computer Graphics Forum*, Vol. 29. Wiley Online Library, 833–842.

[12] Xiangnan He, Tao Chen, Min-Yen Kan, and Xiao Chen. 2015. Trirank: Review-aware explainable recommendation by modeling aspects. In *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*. 1661–1670.

[13] Jonathan L Herlocker, Joseph A Konstan, and John Riedl. 2000. Explaining collaborative filtering recommendations. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work*. ACM, 241–250.

[14] Dietmar Jannach, Sidra Naveed, and Michael Jugovac. 2016. User control in recommender systems: Overview and interaction challenges. In *International Conference on Electronic Commerce and Web Technologies*. Springer, 21–33.

[15] Yucheng Jin, Karsten Seipp, Erik Duval, and Katrien Verbert. 2016. Go with the flow: effects of transparency and user control on targeted advertising using flow charts. In *Proceedings of the International Working Conference on Advanced Visual Interfaces*. ACM, 68–75.

[16] Michael Jugovac and Dietmar Jannach. 2017. Interacting with recommenders—overview and research directions. *ACM Transactions on Interactive Intelligent Systems (TiiS)* 7, 3 (2017), 10.

[17] Antti Kangasrääsiö, Dorota Glowacka, and Samuel Kaski. 2015. Improving controllability and predictability of interactive recommendation interfaces for exploratory search. In *Proceedings of the 20th international conference on intelligent user interfaces*. ACM, 247–251.

[18] Joseph A Konstan and John Riedl. 2012. Recommender systems: from algorithms to user experience. *User modeling and user-adapted interaction* 22, 1-2 (2012), 101–123.

[19] Greg Linden, Brent Smith, and Jeremy York. 2003. Amazon. com recommendations: Item-to-item collaborative filtering. *IEEE Internet computing* 7, 1 (2003), 76–80.

[20] Benedikt Loepp, Tim Donkers, Timm Kleemann, and Jürgen Ziegler. 2019. Interactive recommending with tag-enhanced matrix factorization (TagMF). *International Journal of Human-Computer Studies* 121 (2019), 21–41.

[21] Benedikt Loepp, Katja Herrmanny, and Jürgen Ziegler. 2015. Blended recommending: Integrating interactive information filtering and algorithmic recommender techniques. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 975–984.

[22] Martijn Millecamp, Sidra Naveed, Katrien Verbert, and Jürgen Ziegler. 2019. To Explain or Not to Explain: the Effects of Personal Characteristics When Explaining Feature-based Recommendations in Different Domains. In *CEUR workshop proceedings*. CEUR.

[23] Afshin Moin. 2014. A unified approach to collaborative data visualization. In *Proceedings of the 29th Annual ACM Symposium on Applied Computing*. ACM, 280–286.

[24] Khalil Muhammad, Aonghus Lawlor, and Barry Smyth. 2016. On the use of opinionated explanations to rank and justify recommendations. In *The Twenty-Ninth International Flairs Conference*.

[25] Sayooran Nagulendra and Julita Vassileva. 2014. Understanding and controlling the filter bubble through interactive visualization: a user study. In *Proceedings of the 25th ACM conference on Hypertext and social media*. 107–115.

[26] Sidra Naveed and Jürgen Ziegler. 2019. Feature-Driven Interactive Recommendations and Explanations with Collaborative Filtering Approach.. In *ComplexRec@ RecSys*. 10–15.

[27] John O'Donovan, Barry Smyth, Brynjar Gretarsson, Svetlin Bostandjiev, and Tobias Höllerer. 2008. PeerChooser: visual interactive recommendation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1085–1088.

[28] Eli Pariser. 2011. *The filter bubble: What the Internet is hiding from you*. Penguin UK.

[29] Denis Parra and Peter Brusilovsky. 2015. User-controllable personalization: A case study with SetFusion. *International Journal of Human-Computer Studies* 78 (2015), 43–67.

[30] Denis Parra, Peter Brusilovsky, and Christoph Trattner. 2014. See what you want to see: visual user-driven approach for hybrid recommendation. In *Proceedings of the 19th international conference on Intelligent User Interfaces*. ACM, 235–240.

[31] János Podani. 1999. Extending Gower's general coefficient of similarity to ordinal characters. *Taxon* 48, 2 (1999), 331–340.

[32] Pearl Pu, Li Chen, and Rong Hu. 2011. A user-centric evaluation framework for recommender systems. In *Proceedings of the fifth ACM conference on Recommender systems*. ACM, 157–164.

[33] Pearl Pu, Li Chen, and Rong Hu. 2012. Evaluating recommender systems from the user's perspective: survey of the state of the art. *User Modeling and User-Adapted Interaction* 22, 4-5 (2012), 317–355.

[34] Amit Sharma and Dan Cosley. 2013. Do social explanations work?: studying and modeling the effects of social explanations in recommender systems. In *Proceedings of the 22nd international conference on World Wide Web*. ACM, 1133–1144.

[35] Rashmi Sinha and Kirsten Swearingen. 2002. The role of transparency in recommender systems. In *CHI'02 extended abstracts on Human factors in computing systems*. ACM, 830–831.

[36] Kirsten Swearingen and Rashmi Sinha. 2001. Beyond algorithms: An HCI perspective on recommender systems. In *ACM SIGIR 2001 Workshop on Recommender Systems*, Vol. 13. Citeseer, 1–11.

[37] Panagiotis Symeonidis, Alexandros Nanopoulos, and Yannis Manolopoulos. 2008. Providing justifications in recommender systems. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans* 38, 6 (2008), 1262–1272.

[38] Panagiotis Symeonidis, Alexandros Nanopoulos, and Yannis Manolopoulos. 2009. MoviExplain: a recommender system with explanations. *RecSys* 9 (2009), 317–320.

[39] Nava Tintarev and Judith Masthoff. 2007. Effective explanations of recommendations: user-centered design. In *Proceedings of the 2007 ACM conference on Recommender systems*. ACM, 153–156.

[40] Nava Tintarev and Judith Masthoff. 2012. Evaluating the effectiveness of explanations for recommender systems. *User Modeling and User-Adapted Interaction* 22, 4-5 (2012), 399–439.

[41] Nava Tintarev and Judith Masthoff. 2015. Explaining recommendations: Design and evaluation. In *Recommender systems handbook*. Springer, 353–382.

[42] Chun-Hua Tsai and Peter Brusilovsky. 2017. Providing Control and Transparency in a Social Recommender System for Academic Conferences. In *Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization*. ACM, 313–317.

[43] Chun-Hua Tsai and Peter Brusilovsky. 2018. Beyond the ranked list: User-driven exploration and diversification of social recommendation. In *23rd International Conference on Intelligent User Interfaces*. ACM, 239–250.

[44] Katrien Verbert, Denis Parra, and Peter Brusilovsky. 2014. The effect of different set-based visualizations on user exploration of recommendations. In *CEUR Workshop Proceedings*, Vol. 1253. University of Pittsburgh, 37–44.

[45] Katrien Verbert, Denis Parra, Peter Brusilovsky, and Erik Duval. 2013. Visualizing recommendations to support exploration, transparency and controllability. In *Proceedings of the 2013 international conference on Intelligent user interfaces*. ACM, 351–362.

[46] Jesse Vig, Shilad Sen, and John Riedl. 2009. Tagsplanations: explaining recommendations using tags. In *Proceedings of the 14th international conference on Intelligent user interfaces*. ACM, 47–56.

[47] Bo Xiao and Izak Benbasat. 2007. E-commerce product recommendation agents: use, characteristics, and impact. *MIS quarterly* 31, 1 (2007), 137–209.

[48] Kai Zeng, Kun She, and Xinzheng Niu. 2014. Feature selection with neighborhood entropy-based cooperative game theory. *Computational intelligence and neuroscience* 2014 (2014), 11.