

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

Supporting people with autism spectrum disorders in the exploration of Pals: An Inclusive Recommender System

This is the author's manuscript

Original Citation:

Availability:

This version is available <http://hdl.handle.net/2318/1835391> since 2022-06-22T11:31:15Z

Published version:

DOI:10.1145/3505267

Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

Supporting People with Autism Spectrum Disorders in the Exploration of PIs - An Inclusive Recommender System

Noemi Mauro
University of Torino
Corso Svizzera 185
Torino, Italy
noemi.mauro@unito.it

Liliana Ardissono
University of Torino
Corso Svizzera 185
Torino, Italy
liliana.ardissono@unito.it

Federica Cena
University of Torino
Corso Svizzera 185
Torino, Italy
federica.cena@unito.it

ABSTRACT

The suggestion of Points of Interest (PoIs) to people with autism spectrum disorders challenges the research about recommender systems by introducing an explicit need to consider both user preferences and aversions in item evaluation. The reason is that autistic users' perception of places is influenced by sensory aversions which can cause stress and anxiety when they visit the suggested PoIs. Therefore, the management of individual preferences is not enough to provide these people with suitable recommendations.

To address this issue, we propose a Top-N recommendation model that combines information about the user's idiosyncratic aversions with her/his preferences in a personalized way. The goal is that of suggesting the places that (s)he can like and smoothly experience at the same time. We are interested in finding a user-specific balance of compatibility and interest within a recommendation model that integrates heterogeneous evaluation criteria to appropriately take these aspects into account.

We tested our model on 148 adults, 20 of which were people with autism spectrum disorders. The evaluation results show that, on both groups, our model achieves superior accuracy and ranking results than the recommender systems based on item compatibility, on user preferences, or which integrate these aspects using a uniform evaluation model. These findings encourage us to use our model as a basis for the development of inclusive recommender systems.

1. INTRODUCTION

The personalized suggestion of Points of Interest (PoIs) to fragile users challenges the development of recommender systems [19] by broadening the factors to be taken into account in the identification of the most suitable items for the individual user. For instance, people with autism spectrum

A previous version of this paper, entitled "Personalized Recommendation of PoIs to People with Autism" was published in the Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization (Genoa, Italy, 2020), 163–172.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Copyright 2008 ACM 0001-0782/08/0X00 ...\$5.00.

disorders, who are the main target of this work, have idiosyncratic sensory aversions to noise, brightness, and other sensory features, which influence the way they perceive items, especially places [20]. Thus, a recommender system that overlooks these aversions could suggest PoIs that cause a high level of stress and anxiety on the user [7]. In order to address this issue, the preference data traditionally used to personalize item recommendation should be combined with information about people's aversions to estimate the likelihood that, rather than only being interested in exploring the suggested places, they can serenely experience them.

Starting from Multi-Criteria Decision Analysis [25], which provides techniques for the evaluation of multiple dimensions of items, and on match-making models based on user-to-item similarity [12], most recommender systems assume that the attributes of an item contribute to its utility to the user in an additive way. However, depending on individual idiosyncrasies and their strength, problematic features might make an item unsuitable, even though it meets the user's preferences. Moreover, the impact of compatibility on decision-making varies individually and it cannot be separately managed. For instance, some people with autism are determined to visit noisy and crowded places if they like them very much. Therefore, inclusive recommendation models must reflect personal evaluation criteria by balancing feature compatibility and preference satisfaction at the individual level. In the present work, we investigate the role of these two types of information in the personalized suggestion of PoIs to users with, or without autism spectrum disorders (neurotypical users). We propose a novel Top-N recommender system that applies heterogeneous evaluation criteria to take user preferences and compatibility requirements into account, by exploiting feature-based user profiles for the specification of individual needs.

Our work has two key aspects. Firstly, we acquire data about people's aversion to sensory features of places in terms of disturbance caused by low or high feature values; e.g., darkness or strong light. In this task, we try to limit the amount of information elicited from people as much as possible. For this purpose, we employ a questionnaire derived from [23], which provides data about a user's aversion to a subset of the values that each feature can take. Then, we interpolate her/his aversion to the whole range of values and we derive the compatibility of the feature as the complement of aversion. Secondly, for the estimation of item

ratings, we distinguish user preferences for broad categories of places from idiosyncratic sensory aversions. Moreover, as users might balance differently these aspects in item evaluation, we combine preferences and compatibility by applying user-specific weights, which we acquire by analyzing users' ratings, in conjunction with their declared preferences and idiosyncrasies.

An important challenge in the development of this type of system is that it must work under data scarcity because few users can be studied to learn their interests. Research studies indicate that autism spectrum disorders affect around 1 in 100 people in Europe [5]. Moreover, these people can be hardly contacted because they have interaction problems and a tendency to avoid new experiences. Finally, their attention problems cause difficulties in providing detailed feedback about items [13].

We tested our model on 148 adults: 20 of them were people with autism, while we did not have any information about the others. However, we can reasonably expect that the second sample respected the proportion of the entire population, including at most 1 or 2 autistic people. On both groups of participants, the accuracy and ranking capability achieved by our model was higher than that of a set of baseline recommender systems that singularly take item compatibility, or user preferences, into account. Moreover, our system outperformed baseline models that uniformly manage compatibility and preference information, without differentiating their contributions.

The approach presented in this paper is part of the Personalized Interactive Urban Maps for Autism project (PIUMA), which aims at developing novel digital solutions to help people with autism in their everyday movements [18]. PIUMA involves a collaboration among the Computer Science and Psychology Departments of the University of Torino, and the Adult Autism Center of the city of Torino, Italy. The result of this project is a mobile app that manages dynamic geographical maps specifically conceived for users with autism spectrum disorders, but which target neurotypical people, as well [3].

The remainder of this paper is organized as follows: first, we discuss the spatial needs of autistic people (Section 2), and we position our work in the related one (Section 3). Section 4 outlines how we gather data about users and PoIs, and Section 5 presents our model. Section 6 describes the validation methodology we applied to test our model. Sections 7 and 8 present and discuss the evaluation results. At last, Section 9 concludes the paper.

2. SPATIAL NEEDS OF AUTISTIC PEOPLE

Symptoms of autism span from severe language and intellectual disabilities to the absence of disabilities, and an Intelligence Quotient above the average. Autism entails an atypical sensory perception in over 90% of individuals [23], who can be overwhelmed by environmental factors that are easily managed by neurotypical subjects. At least in part because of these characteristics, people with autism spectrum disorders actively avoid places that may negatively impact their senses [22]. Sight, smell, and hearing are relevant to mobility in urban environments, and high sensory stimulation negatively influences individuals in their movements. Further relevant environmental dimensions that could impact the sense of safeness are the temperature, openness, and crowding of a place. These idiosyncratic sensory aversions

may result in anxiety, fatigue, disgust, sense of oppression or distraction [18].

In order to address this issue, there is a strong need for technological support able to satisfy the spatial needs of people with autism, focusing on aversions derived from their high sensitivity to sensory stimulation. Moreover, as these aversions seem to be highly idiosyncratic, there are no features of places that may reassure the entire autistic population, and the peculiarities of each person have to be considered [16]. Therefore, the provision of personalized solutions that adapt to the individual user is extremely important.

3. RELATED WORK

As people with autism spectrum disorders commonly exhibit an affinity with technology, Information and Communication Technologies are largely used to support them in the management of daily activities [16, 17]. However, the research about autism tends to pay more attention to children, and it overlooks adults' needs. This might be a consequence of the "medical model", which promotes the intervention toward a school-aged target. Moreover, the Human-Computer Interaction community seems to prefer addressing social interaction problems [16, 8], instead of dealing with spatial difficulties.

Most applications investigate the adoption of personalization strategies targeted to autism in the educational domain. For example, Judy et al. [11] present a personalized e-learning system that provides learning paths having different difficulty levels, based on the user's past performance. The authors define ontologies to describe learning materials, annotation schemas, and services, and they use a genetic algorithm as an optimization technique, by representing a set of learning objects as chromosomes.

García et al. [6] propose an adaptive web-based application that helps students with autism spectrum disorders overcome the challenges they might have to face when they attend university. The system adapts the presentation of the information site to autistic and neurotypical students, but the information is the same for everyone. The adaptive functionality is based on learning styles (visual vs. verbal, global vs. analytical, active vs. reflective) and on the user's history. For example, if the user is more visual than verbal, the video version of content is shown at the top of the learning object. Otherwise, it is moved to the bottom of the object. Hong et al. [10] propose to provide users with suggestions within a social network aimed at supporting young adults' independence. However, they focus on the organization of the social network, by relying on peer suggestions, rather than automatically generating recommendations.

Differently, Costa et al. [4] develop a task recommendation system that uses a machine learning technique to supplement the child's regular therapy. The system suggests the daily activities to be performed (related to eating, keeping clean, getting dressed, and so forth) based on age, gender, and time of day. It does not consider the child's preferences, and the difficulty level of the activities is manually set by the therapist. Moreover, in [15], Ng and Pera propose a hybrid game recommender for adult people with autism spectrum disorders, based on collaborative and graph-based recommendation techniques.

Our work differs from the previously listed ones in several aspects. Firstly, we focus on a different domain, i.e., spatial support. Secondly, we evaluated our model with autistic

people. This has rarely, if ever, been done in the related research. Thirdly, our approach employs personal preferences for item categories, and aversions to sensory features, to steer recommendation in a context where a limited amount of feedback about items can practically be collected.

Our work also differs from general content-based [12], feature-based [9], collaborative and multi-criteria [1] recommender systems, because we treat sensory features as sources of discomfort for users, rather than liking or disliking factors. In other words, we separately model the influence of idiosyncratic sensory aversions, which determine the compatibility of items with the user, from her/his preferences for different types of items. Notice that this separation also distinguishes our model from recommender systems that deal with negative preferences, such as [14], because we support the management of heterogeneous criteria to deal with user preferences and sensory idiosyncrasies. Previously, the INTRIGUE [2] tourist guide introduced the notion of compatibility requirements in PoI recommendation. However, it did not investigate their different meaning and impact on the evaluation of items, with respect to user preferences.

It is worth mentioning that, while constraint-based recommender systems are too knowledge-intensive for our purposes (we are not suggesting item bundles with constraint satisfaction requirements), the optimization of soft constraints for path-finding under suitability criteria is relevant to extend PoI suggestion with instructions for reaching the target places. This type of technique has been explored in recommender systems for routing, such as [24].

4. PRELIMINARY STUDY SETUP

This section describes how we gathered data about users and places to validate our model. Moreover, it describes the samples of users we involved.

4.1 Data

4.1.1 Knowledge about users

The acquisition of individual user profiles is a key step to personalize recommendations because it makes it possible to explicitly represent the user preferences and requirements to be considered in item evaluation. User profiles can be explicitly elicited from users, or they can be unobtrusively learned by tracking and analyzing user behavior [19]. In this work, we adopted the former technique, which makes it possible to initialize the user profile before starting to use the mobile guide, and thus supports the identification of unsuitable PoIs since the beginning of the interaction with the user. This approach does not preclude the adoption of dynamic user modeling techniques to update the user profiles while people use the mobile guide, and we have recently extended our work in this direction.

Our questionnaire, shown in Table 1, includes two sections. In the first one (left column of the table), it elicits user preferences about categories of PoIs such as restaurants, parks, etc., in order to learn which ones users like or dislike. In the second section (right column), questions concern users' aversions to sensory features of places.

The information about sensory aversions is hard to obtain: usually, very long and complex surveys have to be completed for this purpose [20]. Moreover, asking people with autism for such data is challenging because they have difficulties in social interactions and they tend to avoid new experiences

Table 1: Short questionnaire to elicit information about preferences and sensory idiosyncrasies (translated from Italian).

| <i>From 1 (not at all) to 5 (very much), how much do you like doing the following activities?</i> | <i>In a place, how much does it bother you:</i> |
|---|---|
| 1. be in nature, go to parks, gardens, green areas, ... | 1. too much light |
| 2. visit museums, exhibitions, cultural events | 2. very low light |
| 3. go to the cinema, theater, concerts | 3. a lot of people |
| 4. go to comic shops | 4. a lot of noise |
| 5. go to clothing stores | 5. strong smells |
| 6. go to malls and markets | 6. cramped places (narrow, small) |
| 7. go to the library | 7. large places |
| 8. go to the bookshop | |
| 9. play sport | |
| 10. go to pubs, cafe | |
| 11. go to the restaurant | |
| 12. go to the ice cream shop | |
| 13. stay in squares | |
| 14. go to the railway stations | |

[21]. Given our users' attention problems [13], and considering the application context of our project, which is not a clinical setting, we decided to avoid long and detailed surveys. Therefore, we carefully prepared with psychologists a short list of questions to capture such information.

We defined the questions about aversions by adapting a subset of the Sensory Perception Quotient (SPQ) test [23] on the basis of the findings reported in [18]. SPQ is a standard sensory questionnaire for adults that assesses basic sensory hyper- and hyposensitivity. We would have liked to directly use it since it is part of the battery of assessment tests proposed to the patients of the Autistic Adult center in Torino. However, it includes 92 items, too many to be proposed when bootstrapping a mobile guide. As shown in Table 1, our questionnaire is aimed at acquiring aversion information more quickly. Specifically, for some features (brightness and space), the user is asked to evaluate two extreme conditions, i.e., low or high levels, assuming that the middle ones are less problematic. In other cases (crowding, noise, and smell), the user is asked about her/his annoyance concerning the highest level, because low levels of these features are neutral.

In our experiment, users filled in the survey of Table 1, possibly in the presence of an operator (when needed), and they answered questions using the [1, 5] Likert scale. Then, we asked users to evaluate 50 specific PoIs located in Torino city center (e.g., *How much do you like Castle Square?*) in order to collect a dataset of user ratings to test our model. We used the same [1, 5] Likert scale as above, but we included the "I don't know this place" choice to support opting out.

4.1.2 Knowledge about PoIs

We used the Maps4All crowdsourcing platform (<https://maps4all.firstlife.org/\#>) as a source of information about places. Specifically, the 50 PoIs mentioned in Sec-

tion 4.1.1 are representative of all the categories of places defined in that platform. We selected those PoIs with the requirement that they had previously been mapped with the contribution of at least three different crowdsourcers each.

The reason for exploiting an ad-hoc platform as a source of information about PoIs, instead of relying on a public Open Data source, is the fact that Maps4All was explicitly designed to support the crowdsourcing of sensory features of places. In contrast, Open Data sources such as OpenStreetMap (<https://www.openstreetmap.org>) fail to provide the sensory information we needed for our experiment. In particular, for each place, Maps4all enables the user to rate in the $[1, 5]$ scale the level of i) brightness, ii) crowding, iii) noise, iv) smell, v) openness, and vi) temperature. These sensory features have been defined based on the findings of the user study presented in [18], and of state-of-art research [20]. Notice that, by interacting with Maps4All, the user can also provide a global rating of the place.

We populated the Maps4All platform through two experimental crowdsourcing sessions, during two lessons at the Master’s degree in Social Innovation and ICT at the University of Torino, in May and December 2019. About 120 students participated in the crowdsourcing tasks. In order to guarantee the collection of a reasonable amount of data about places, we asked each of them to provide evaluations for at least three PoIs in Torino city center. In total, during the two sessions, we collected the evaluations of 282 items.

4.2 Sample

For our study, we involved two groups of users:

- 20 adults with autism spectrum disorders (from 22 to 40 years old, mean age: 26.3, median 28; 11 men, 9 women), who are patients of the Autistic Adult Center of Torino, medium- and high-functioning.
- 128 neurotypical subjects (from 19 to 71 years old, mean age: 28.1, median 23; 63 men, 65 women), who are university students or contacts of the authors of this paper.

All participants signed a privacy consensus according to General Data Protection Regulation. Moreover, we obtained approval for the study from the research ethics committee of the University of Torino.

As far as the 50 PoIs we selected, the mean number of evaluations we obtained is 31 for autistic participants and 39 for neurotypical ones.

5. RECOMMENDATION MODEL

As previously discussed, we assume that both user preferences and item compatibility should be taken into account to identify the most relevant items that a user can smoothly experience and like, at the same time. However, evaluation criteria might be personal. Moreover, these aspects can be weighted differently in decision-making processes. For instance, in contrast to the tendency of people with autism spectrum disorders to visit places in which they feel comfortable, during our participatory design interview sessions we encountered a few subjects who face the challenges of noisy and crowded environments in order to be able to carry out the activities they like very much. We thus propose a recommendation model that, based on the observed item

evaluations, can weigh the contribution of compatibility and preferences in rating prediction, on a user-specific basis.

For clarity purposes, we split the presentation of our model as follows. In Section 5.1, we describe the input data for recommendation. In Section 5.2, we specify how we estimate the compatibility of the individual features of an item with the user. Then, we present the estimation of the overall compatibility of the item with the user (Section 5.3), and the preference-based item evaluation (Section 5.4). In Section 5.5, we describe how we combine compatibility and preference-based evaluation to predict item ratings.

Before describing our model, we introduce the notation we use:

- U is the set of users and I the set of items of the domain.
- C is the set of item categories, such as shops and cinemas.
- L is a Likert scale in $[1, v_{max}]$. In this work, $v_{max} = 5$.
- $F = F^\uparrow \cup F^V$ is the set of sensory features defined in our domain. We assume that each feature $f \in F$ takes values in L .

Specifically, F^\uparrow is the set of features f such that, the higher the value of f , the stronger its negative impact on the user. For instance, noise belongs to this class. Differently, F^V denotes features whose extreme values make users uncomfortable, while the middle ones are less problematic; e.g., brightness.

In our domain, there are no features such that people are expected to feel comfortable with high values and uncomfortable with low ones. Thus, we omit this class.

For each user $u \in U$, and item $i \in I$, we estimate u ’s evaluation of i (denoted as \hat{r}_{ui}) as a decimal number in the $[1, v_{max}]$ interval, by taking u ’s previous ratings, preferences for item categories, and idiosyncrasies into account.

5.1 Input Data

Our model takes the user and item profiles as input. The profile of $u \in U$, extracted from the questionnaire data, specifies:

- The ratings r_j in L that (s)he provided for a set of items $j \in I$.
- Her/his declared preferences for the categories $c \in C$, each one expressed in the L scale.
- Her/his declared sensory aversion to specific values of item features, expressed in L . We denote u ’s aversion to a value v of a feature $f \in F$ as a_{ufv} . For example, $a_{uf5} = 4$ means that u is fairly disturbed by items having $f = 5$.

For each feature $f \in F^\uparrow$, we assume by default that $a_{uf1} = 1$. Therefore, the user profile stores a single value, $a_{ufv_{max}}$, which specifies u ’s aversion to the maximum value of f . We denote the maximum value of f as v_{max} .

For each feature f in F^V , the user profile stores two values which express u ’s aversion to the minimum and maximum values of f , respectively: e.g., $\{a_{uf1} = 3, a_{ufv_{max}} = 4\}$.

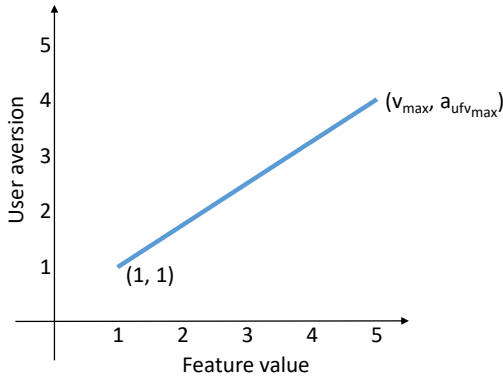


Figure 1: Interpolation of a user’s aversion to a feature of type F^\uparrow .

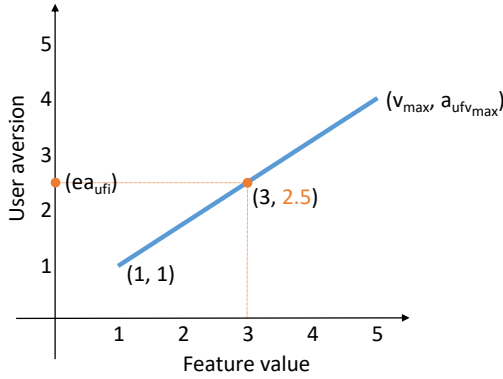


Figure 2: Example of the interpolation of a user’s aversion to a feature of type F^\uparrow .

Differently, the profile of an item i specifies the category $c \in C$ of the item, and a vector $\vec{\mathbf{i}}_f$ storing, for each feature $f \in F$, the value of f in item i retrieved by querying the Maps4All platform. For each feature f , Maps4All returns the mean evaluation $\vec{\mathbf{i}}_f$ it collected from crowdsourcers; $\vec{\mathbf{i}}_f$ takes values in the $[1, v_{max}]$ interval.

5.2 Compatibility of Individual Features with the User

We can define compatibility as the opposite of aversion to feature values. However, user profiles only include one or two aversion values declared by users for each feature. Thus, the missing ones have to be interpolated. In the following, we describe the patterns we apply to approximate a user’s aversion to item features, starting from the values stored in her/his profile.

For each $f \in F^\uparrow$, we approximate aversion as a linearly increasing function. Let us represent feature values in the X axis, and user aversion in the Y axis of a plane. Then, we can define this function as a line that connects point $(1, 1)$ to point $(v_{max}, a_{ufv_{max}})$, as in Figure 1:

$$line^\uparrow(x) = 1 + \frac{(a_{ufv_{max}} - 1)(x - 1)}{v_{max} - 1} \quad (1)$$

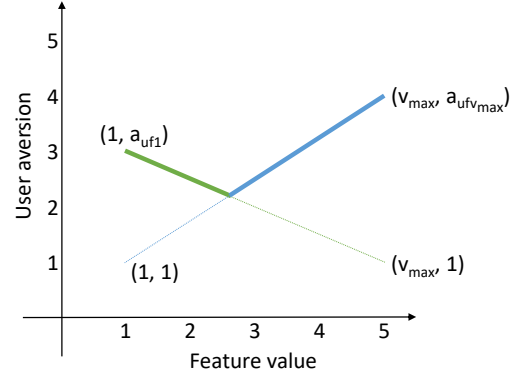


Figure 3: Interpolation of a user’s aversion to a feature of type F^V .

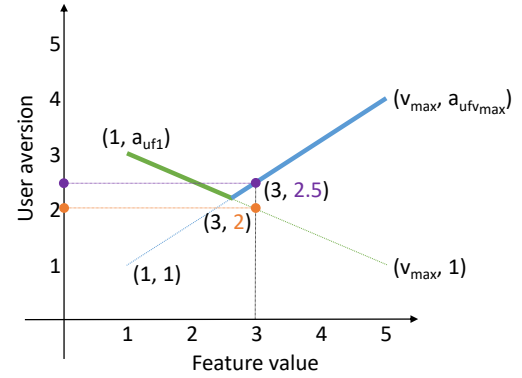


Figure 4: Example of the interpolation of a user’s aversion to a feature of type F^V .

We thus estimate u ’s aversion to f in i (ea_{ufi}) as follows:

$$ea_{ufi} = line^\uparrow(\vec{\mathbf{i}}_f) \quad (2)$$

For instance, the line in Figure 2 shows the interpolation of a user’s aversion to a feature f . Given a user u with $a_{ufv_{max}} = 4$, and a PoI i such that $\vec{\mathbf{i}}_f = 3$, $ea_{ufi} = line^\uparrow(3)$. Thus, u ’s aversion to f in i is approximated to 2.5.

Differently, for each $f \in F^V$, and given $\{a_{uf1}, a_{ufv_{max}}\}$ in u ’s profile, we interpolate aversion by means of a concave function on the range of f . The aversion function has a “V” shape, which we approximate by drawing two lines, as in Figure 3:

- $line^\uparrow$ connects points $(1, 1)$ and $(v_{max}, a_{ufv_{max}})$ to represent the increment of aversion towards the maximum value of f .
- $line_\downarrow$ connects points $(1, a_{uf1})$ and $(v_{max}, 1)$ to represent the decrease in aversion while f takes higher values than its minimum:

$$line_\downarrow(x) = 1 + \frac{(x - v_{max})(1 - a_{uf1})}{v_{max} - 1} \quad (3)$$

We estimate u ’s aversion to f in i by selecting the maximum values of the two lines:

$$ea_{ufi} = \max(line^\uparrow(\vec{\mathbf{i}}_f), line_\downarrow(\vec{\mathbf{i}}_f)) \quad (4)$$

Let's look at the example in Figure 4. Given a PoI i such that $\vec{i}_f = 3$, $ea_{ufi} = \max(\text{line}^\uparrow(3), \text{line}^\downarrow(3))$. Thus, u 's aversion to f in i is estimated as $\max(2.5, 2) = 2.5$.

Notice that ea_{ufi} takes values in the $[1, v_{max}]$ interval. Moreover, higher values of this measure mean that the feature generates more discomfort to u .

Given ea_{ufi} , the compatibility of f with u in i , denoted as $comp_{ufi}$, can thus be defined as:

$$comp_{ufi} = v_{max} + 1 - ea_{ufi} \quad (5)$$

For example, if $ea_{ufi} = 2.5$ and $v_{max} = 5$, $comp_{ufi} = 3.5$.

5.3 Overall Item Compatibility: Aggregation Measures

We propose alternative aggregation measures to compute the overall compatibility of an item i with a user u ($comp_{ui}$) by modeling different types of influence of individual features. In Section 7, we evaluate their performance, in combination with diverse recommendation algorithms.

- **Min.** This measure defines $comp_{ui}$ as the minimum compatibility of i 's features with u :

$$comp_{ui} = \min_{f \in F} comp_{ufi} \quad (6)$$

Min is conjunctive and it evaluates i as incompatible with u if the item has at least one incompatible feature.

- **Ave.** In this case, $comp_{ui}$ is the mean compatibility of the features of i :

$$comp_{ui} = \frac{\sum_{f \in F} comp_{ufi}}{|F|} \quad (7)$$

where $|\cdot|$ denotes set cardinality. This measure is additive (disjunctive) and equally balances the influence of the features on compatibility.

We also define two aggregation measures that estimate the overall compatibility of an item i with a user u in function of the distance between the features of i (stored in the \vec{i} vector) and those of an *ideal* item that best matches u 's idiosyncrasies. We denote this item as \vec{ideal}_u . For each feature $f \in F$, \vec{ideal}_{uf} is the most compatible value of f , based on u 's estimated aversion to sensory features. Specifically, for each $f \in F^\uparrow$, $\vec{ideal}_{uf} = 1$ (see the red point in Figure 5). Moreover, for each $f \in F^V$, \vec{ideal}_{uf} is represented by the value of f associated to the minimum aversion. For instance, $\vec{ideal}_{uf} = 2$ (violet point) in Figure 6.

The two vector-based aggregation measures for the computation of the overall compatibility of i with u are:

- **Cos.** In this measure, $comp_{ui}$ is the Cosine similarity between \vec{i} and \vec{ideal}_u :

$$comp_{ui} = 1 + (v_{max} - 1) * \frac{\vec{i} \cdot \vec{ideal}_u}{\|\vec{i}\|_F * \|\vec{ideal}_u\|_F} \quad (8)$$

where \cdot is the scalar vector product, $\|\cdot\|_F$ is the Frobenius Norm, and $*$ is the decimal product. A small the angle between \vec{i} and \vec{ideal}_u means that i is highly compatible with u , and *vice versa*.

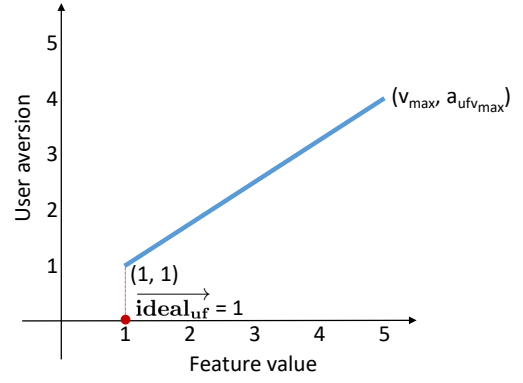


Figure 5: Identification of \vec{ideal}_{uf} for a feature of type F^\uparrow .

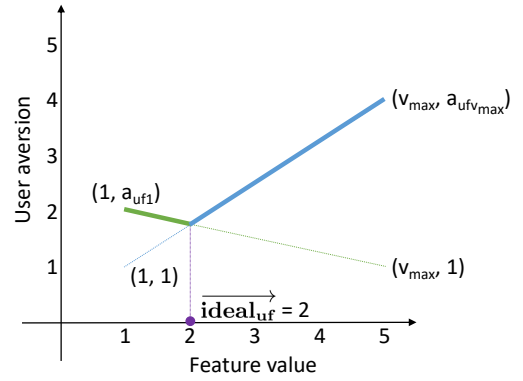


Figure 6: Identification of \vec{ideal}_{uf} for a feature of type F^V .

- **RMSD.** In this case, $comp_{ui}$ is the complement of the Root Mean Square Deviation between \vec{i} and \vec{ideal}_u :

$$comp_{ui} = v_{max} + 1 - \sqrt{\frac{1}{|F|} * \sum_{f \in F} (\vec{i}_f - \vec{ideal}_{uf})^2} \quad (9)$$

The smaller is the distance between \vec{i} and \vec{ideal}_u , the more compatible is i with u .

5.4 Preference-based Item Evaluation

While compatibility indicates whether the user can smoothly experience an item, it does not mean that (s)he will like it. User preferences have to be taken into account for this purpose. In our domain, the only preference that we consider is the interest in the category of the item to be evaluated. Thus, the preference value of a user u for an item of category $c \in C$ corresponds to the value of u 's preference for c stored in u 's profile. We denote this value as p_{uc} .

It is worth mentioning that, if more preferences had to be modeled, a Multi-Criteria Decision Analysis approach might be applied to compute an overall preference estimation as a weighted function of preferences for individual attributes. However, this is out of the scope of the present work.

5.5 Rating Prediction

In order to balance compatibility and preferences in a personalized way, we propose to identify user-dependent evaluation criteria by exploiting the user's idiosyncrasies and

preferences, in combination with the ratings of items (s)he provides. Specifically, we estimate the rating that a user u will give to an item i as a weighted mean of overall compatibility and user preferences:

$$\hat{r}_{ui} = \alpha * comp_{ui} + (1 - \alpha) * p_{uci} \quad (10)$$

where α takes values in the $[0, 1]$ interval, and $p_{uci} \in L$ is the preference-based evaluation of i , given u 's profile. This model, henceforth referred as **Ind** (that is, Individual), identifies a specific α value for each user to optimize item recommendation to her/him. We identify the value of α for each $u \in U$ as the one that minimizes the distance between estimated ratings and ground-truth ones.

6. VALIDATION METHODOLOGY

We aim at assessing whether a recommendation model that takes both item compatibility and user preferences into account is more effective than an approach based on a single type of information. Moreover, we aim at evaluating the usefulness of a personalized balance of these aspects, as specified by the α parameter of Equation 10. For these purposes, we compare our model to a set of recommender systems that (i) uniformly manage compatibility and user preferences, ignoring their possibly different impact on decision-making, or (ii) focus either on compatibility or on preferences. We consider the following baselines:

- **Multi-Criteria (MC)**. This recommender system estimates item ratings by uniformly treating idiosyncratic aversions and preferences on the basis of the aggregation measures described in Section 5.3. Given an item i , it computes \hat{r}_{ui} by fusing u 's preference for the category of i (p_{uci}) with the compatibility of individual features with u ($comp_{ufi}$) by means of a single aggregation function. For example, this function could be the mean of all these values, as in Equation 7.
- **C-only**. This is a configuration of our recommendation model in which $\alpha = 1$. In this case, items are evaluated exclusively on the basis of their compatibility with the user.
- **Pref-only**. In this configuration of our model, we set $\alpha = 0$ to evaluate items on the exclusive basis of the user's preferences.

We did not select as baselines any collaborative or feature-based recommenders such as those proposed in [9, 1], because the data about users is too small to train those algorithms.

We separately compare our model to the above baselines on the dataset of the users with autism spectrum disorders (henceforth denoted as AUT), and on the one regarding neurotypical users (NEU). For the comparison, we configure all the algorithms on each aggregation measure of Section 5.3. The resulting configurations are named by appending the name of the selected measure to that of the applied algorithm. For instance, Ind_{Cos} represents the application of the Cos aggregation measure to model Ind.

To evaluate recommendation performance, we consider ranking capability (MRR and MAP), accuracy (Precision, Recall, and F1), error minimization (MAE and RMSE), and user coverage. However, consistently with recent trends in

Table 2: Results on AUT dataset for N=5. The lines of the table are ordered by MAP. The best values of each measure across all algorithms are printed in bold. The best value obtained by the other category of algorithms is underlined. Stars denote statistical significance: **: $p < 0.01$; *: $p < 0.05$.

| Algorithm | Prec. | Recall | F1 | MAP | MRR | MAE | RMSE |
|------------------|---------------|---------------|---------------|-----------------|---------------|---------------|----------------|
| Ind_{Cos} | 0.6290 | <u>0.6207</u> | 0.6046 | **0.5384 | 0.8095 | 0.9927 | 1.4541 |
| Ind_{Min} | 0.6328 | 0.5832 | 0.5910 | 0.5125 | 0.7825 | 0.8691 | *1.3020 |
| Pref-only | 0.6220 | 0.5912 | 0.5860 | <u>0.5114</u> | 0.7858 | <u>0.9346</u> | <u>1.4276</u> |
| Ind_{Ave} | 0.6118 | 0.5710 | 0.5736 | 0.4960 | 0.7667 | 0.9168 | 1.3659 |
| C-only $_{Cos}$ | 0.6263 | 0.6224 | <u>0.6001</u> | 0.4877 | 0.7583 | 1.3675 | 1.6948 |
| Ind_{RMSD} | 0.5978 | 0.5545 | 0.5577 | 0.4799 | 0.7537 | 0.9965 | 1.4533 |
| MC $_{Ave}$ | 0.6255 | 0.5383 | 0.5575 | 0.4489 | 0.7792 | 1.1902 | 1.4861 |
| MC $_{RMSD}$ | 0.6080 | 0.5396 | 0.5463 | 0.4429 | 0.7775 | 1.2172 | 1.5426 |
| MC $_{Min}$ | <u>0.6305</u> | 0.5057 | 0.5344 | 0.4352 | <u>0.7950</u> | 1.4512 | 1.7943 |
| MC $_{Cos}$ | 0.5917 | 0.5558 | 0.5459 | 0.4336 | 0.7217 | 1.3534 | 1.6236 |
| C-only $_{Min}$ | 0.6065 | 0.4999 | 0.5230 | 0.4166 | 0.7583 | 1.3675 | 1.6816 |
| C-only $_{Ave}$ | 0.5912 | 0.5154 | 0.5270 | 0.4142 | 0.7192 | 1.3045 | 1.6060 |
| C-only $_{RMSD}$ | 0.5825 | 0.5009 | 0.5145 | 0.4036 | 0.7142 | 1.3702 | 1.7168 |

the evaluation of recommender systems, we pay special attention to ranking metrics because they help understand whether the items that the user likes are placed in the first positions of the suggestion list, or not.

We perform 5-fold cross-validation in which, for every fold, we use 80% as training set and 20% as test set. As the Ind models have to optimize the α parameter, we train each of them to find the best user-specific setting by optimizing its results with respect to MAP. Moreover, to be sure that the baselines are consistently evaluated, we run the other algorithms (MC, C-only, and Pref-only, which do not need any training) on the same test sets used for Ind.

7. EVALUATION RESULTS

Tables 2 and 3 show the Top-N evaluation results with $N=5$. That is, the list of suggested items has length=5. The tables omit the results concerning user coverage because it is 100% in all the cases.

We consider two categories of algorithms, i.e., the configurations of our model on the various aggregation measures, and the corresponding ones of the baselines. In the tables, we show the best value across all algorithms in bold. Moreover, the best value obtained by the other category of algorithms is underlined (when our model obtains the best value, we underline the best value achieved by the baselines, and *vice versa*). Stars indicate significant differences according to a Student T-Test between the best performing algorithm from each category; **: $p < 0.01$; *: $p < 0.05$.

The evaluation results suggest that Ind_{Cos} is the best recommender system because it achieves good accuracy and ranking capability. On both datasets, it outperforms all the other algorithms (baselines and own category) in F1 and MAP. Moreover, it has the best Recall of its own category. As a matter of fact, Ind_{Min} achieves better error minimization than Ind_{Cos} on both datasets. Specifically, it obtains the best MAE of all models, and it achieves the best RMSE in AUT. Furthermore, in NEU, it obtains better results than the other algorithms of its own category. However, as previously discussed, our primary evaluation criterion is ranking capability.

Interestingly, Ind_{RMSD} is the worst configuration of our

Table 3: Results on NEU dataset for N=5. We use the same notation as in Table 2.

| Algorithm | Prec. | Recall | F1 | MAP | MRR | MAE | RMSE |
|------------------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| Ind _{Cos} | 0.5790 | <u>0.5406</u> | 0.5349 | 0.4139 | 0.7475 | 1.1792 | 1.5232 |
| Ind _{Min} | 0.5791 | 0.5225 | 0.5250 | 0.4120 | 0.7688 | 1.0950 | <u>1.4024</u> |
| Ind _{Ave} | 0.5740 | 0.5261 | 0.5262 | 0.4108 | 0.7555 | 1.1085 | 1.4343 |
| Ind _{RMSD} | 0.5816 | 0.5286 | 0.5297 | 0.4108 | 0.7521 | 1.1427 | 1.4758 |
| Pref-only | <u>0.5795</u> | 0.5408 | <u>0.5347</u> | <u>0.4076</u> | 0.7304 | 1.1416 | 1.5270 |
| C-only _{Cos} | 0.5503 | 0.5414 | 0.5255 | 0.4000 | 0.7189 | 1.4374 | 1.7456 |
| MC _{Ave} | 0.5752 | 0.5154 | 0.5213 | 0.3995 | 0.7564 | <u>1.1238</u> | 1.3564 |
| MC _{Min} | 0.5664 | 0.4956 | 0.5053 | 0.3890 | <u>0.7583</u> | 1.1249 | 1.4052 |
| MC _{RMSD} | 0.5568 | 0.4840 | 0.4963 | 0.3767 | 0.7433 | 1.3320 | 1.6255 |
| C-only _{Ave} | 0.5476 | 0.4936 | 0.4979 | 0.3701 | 0.7168 | 1.2122 | 1.4668 |
| C-only _{Min} | 0.5507 | 0.4769 | 0.4899 | 0.3673 | 0.7359 | 1.1704 | 1.4213 |
| C-only _{RMSD} | 0.5460 | 0.4870 | 0.4936 | 0.3651 | 0.7223 | 1.4157 | 1.7281 |
| MC _{Cos} | 0.5274 | 0.5053 | 0.4974 | 0.3535 | 0.6591 | 1.2775 | 1.5795 |

model. On the AUT dataset, it obtains the lowest results of its own category on all evaluation metrics. However, it achieves better results than several baselines in MAP and other metrics, supporting the superiority of our model. It is also worth noting that Pref-only is the best baseline regarding MAP. Moreover, C-only_{Cos} has a lower ranking capability than Pref-only, but it has fairly good accuracy. It is the best or second-best baseline on the various measures.

Unfortunately, the low size of the AUT and NEU datasets does not support the statistical significance of results for several evaluation metrics. However, the results concerning MAP and RMSE on the AUT dataset are significant. This is important because our recommendation model is especially targeted to users with autism spectrum disorders. Thus, we can rely on the ranking capability of our model when recommending items to them. At the same time, the results are encouraging for neurotypical users. Therefore, it is worth investigating performance within a larger experiment that will possibly provide more statistically relevant results on both groups of people.

8. DISCUSSION

From the evaluation results, we draw two conclusions. The first one is that a customized model of item evaluation, which balances feature compatibility and preference satisfaction in a personalized way, achieves better performance than the recommender systems that manage only one of these aspects. As far as F1 and ranking capability are concerned, the configurations of the Ind model that take both preferences and compatibility into account (and, specifically, Ind_{Cos}) obtain higher results than Pref-only, which only employs user preferences in item suggestion. Moreover, they achieve better results than the C-only algorithms, which only use compatibility data. The performance of these algorithms is poorer than that of Pref-only, as well. This means that, not surprisingly, compatibility information alone is not enough to generate relevant recommendations for the user.

The second conclusion we draw is that a customized model of item evaluation, which balances feature compatibility and preference satisfaction in a personalized way, outperforms the recommender systems which uniformly manage both aspects. Specifically, the Ind configurations outperform the MC ones, regardless of the applied aggregation measure, in most evaluation metrics, and especially in ranking and F1 measures.

To summarize, preference information is useful to suggest relevant PoIs in Top-N recommendation. However, better results can be achieved by combining it with a compatibility evaluation aimed at assessing whether the user can smoothly experience the recommended items. Interestingly, a uniform management of compatibility and preference information, which does not distinguish the possibly heterogeneous evaluation criteria concerning them, does not bring good results. Conversely, the acquisition of user-specific weights to balance the impact of compatibility and interests improves item suggestion.

9. CONCLUSIONS

Users with autism spectrum disorders are a challenging target of PoI recommender systems because of their spatial needs. In order to suggest suitable solutions, which the user can like and serenely experience, her/his preferences for PoI categories, traditionally analyzed by researchers, and her/his aversions to sensory features, have to be jointly considered. The reason is that aversions can seriously affect the visit experience, causing negative feelings on the user.

In this paper, we presented a novel Top-N recommender of Points of Interest especially targeted to these people. The peculiarity of our model is that it takes the individual user’s idiosyncratic aversions to sensory features into account to generate suggestions that (s)he is expected to like and smoothly experience at the same time. We tested our model on autistic and neurotypical people. The evaluation results show that, on both user groups, our model achieves higher accuracy and ranking capability than baseline recommenders which (i) evaluate items on the sole basis of how closely they meet the user’s preferences, or how compatible they are with her/his idiosyncratic aversions to sensory features, and (ii) uniformly manage compatibility and preference information without distinguishing the different contributions of these aspects to item evaluation. We thus conclude that the integration of heterogeneous evaluation criteria about user interests and aversions is a promising approach to make recommender systems more inclusive.

10. ACKNOWLEDGMENTS

This work is supported by the Compagnia di San Paolo Foundation. We thank Stefano Cocomazzi, Stefania Brighenti, and Claudio Mattutino for their contributions to the work. We are also grateful to the Adult Autism Center of the city of Torino for their participation in the PIUMA project.

11. REFERENCES

- [1] G. Adomavicius, N. Manouselis, and Y. Kwon. *Multi-Criteria Recommender Systems*, pages 769–803. Springer US, Boston, MA, 2011.
- [2] L. Ardissono, A. Goy, G. Petrone, M. Segnan, and P. Torasso. INTRIGUE: personalized recommendation of tourist attractions for desktop and handset devices. *Applied Artificial Intelligence, Special Issue on Artificial Intelligence for Cultural Heritage and Digital Libraries*, 17(8-9):687–714, 2003.
- [3] F. Cena, N. Mauro, L. Ardissono, C. Mattutino, A. Rapp, S. Cocomazzi, S. Brighenti, and R. Keller. Personalized tourist guide for people with autism. In *Adjunct Publication of the 28th ACM Conference on User Modeling, Adaptation and Personalization*,

- UMAP '20 Adjunct, page 347–351, New York, NY, USA, 2020. Association for Computing Machinery.
- [4] M. Costa, A. Costa, V. Julián, and P. Novais. A task recommendation system for children and youth with autism spectrum disorder. In J. F. De Paz, V. Julián, G. Villarrubia, G. Marreiros, and P. Novais, editors, *Ambient Intelligence– Software and Applications – 8th International Symposium on Ambient Intelligence (ISAmI 2017)*, pages 87–94, Cham, 2017. Springer International Publishing.
 - [5] M. Elsabbagh, G. Divan, Y.-J. Koh, Y. S. Kim, S. Kauchali, C. Marcín, C. Montiel-Nava, V. Patel, C. S. Paula, C. Wang, et al. Global prevalence of autism and other pervasive developmental disorders. *Autism research*, 5(3):160–179, 2012.
 - [6] A. M. García, N. Stash, M. Fabri, P. D. Bra, G. H. L. Fletcher, and M. Pechenizkiy. Adaptive web-based educational application for autistic students. In K. Koidl and B. Steichen, editors, *Late-breaking Results, Demos, Doctoral Consortium, Workshops Proceedings and Creative Track of the 27th ACM Conference on Hypertext and Social Media (HT 2016), Halifax, Canada, July 13-16, 2016*, volume 1628 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2016.
 - [7] A. Gillott and P. Standen. Levels of anxiety and sources of stress in adults with autism. *Journal of Intellectual Disabilities*, 11(4):359–370, 2007.
 - [8] O. Grynszpan, P. L. Weiss, F. Perez-Diaz, and E. Gal. Innovative technology-based interventions for autism spectrum disorders: a meta-analysis. *Autism*, 18(4):346–361, 2014.
 - [9] E.-H. S. Han and G. Karypis. Feature-based recommendation system. In *Proceedings of the 14th ACM International Conference on Information and Knowledge Management, CIKM '05*, pages 446–452, New York, NY, USA, 2005. ACM.
 - [10] H. Hong, J. G. Kim, G. D. Abowd, and R. I. Arriaga. Designing a social network to support the independence of young adults with autism. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work, CSCW '12*, page 627–636, New York, NY, USA, 2012. Association for Computing Machinery.
 - [11] M. Judy, U. Krishnakumar, and H. A. Narayanan. Constructing a personalized e-learning system for students with autism based on soft semantic web technologies. In *2012 IEEE International Conference on Technology Enhanced Education (ICTEE)*, pages 1–5. IEEE, 2012.
 - [12] P. Lops, M. de Gemmis, and G. Semeraro. *Content-based recommender systems: state of the art and trends*, pages 73–105. Springer US, Boston, MA, 2011.
 - [13] D. Murray, M. Lesser, and W. Lawson. Attention, monotropism and the diagnostic criteria for autism. *Autism*, 9(2):139–156, 2005.
 - [14] C. Musto, G. Semeraro, P. Lops, and M. de Gemmis. Random indexing and negative user preferences for enhancing content-based recommender systems. In C. Huemer and T. Setzer, editors, *E-Commerce and Web Technologies*, pages 270–281, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg.
 - [15] Y. Ng and M. Pera. Recommending social-interactive games for adults with autism spectrum disorders (ASD). In *Proceedings of the 12th ACM Conference on Recommender Systems, RecSys '18*, pages 209–213, New York, NY, USA, 2018. ACM.
 - [16] C. Putnam, C. Hanschke, J. Todd, J. Gemmell, and M. Kollia. Interactive technologies designed for children with autism: Reports of use and desires from parents, teachers, and therapists. *ACM Trans. Access. Comput.*, 12(3), Sept. 2019.
 - [17] S. Ramdoss, W. Machalicek, M. Rispoli, A. Mulloy, R. Lang, and M. O'Reilly. Computer-based interventions to improve social and emotional skills in individuals with autism spectrum disorders: a systematic review. *Developmental Neurorehabilitation*, 15(2):119–135, 2012.
 - [18] A. Rapp, F. Cena, R. Castaldo, R. Keller, and M. Tirassa. Designing technology for spatial needs: Routines, control and social competences of people with autism. *International Journal of Human-Computer Studies*, 120:49 – 65, 2018.
 - [19] F. Ricci, L. Rokach, and B. Shapira. *Introduction to Recommender Systems Handbook*, pages 1–35. Springer US, Boston, MA, 2011.
 - [20] C. E. Robertson and S. Baron-Cohen. Sensory perception in autism. *Nature Reviews Neuroscience*, 18(11):671, 2017.
 - [21] E. Schopler and G. B. Mesibov. *Social Behavior in Autism*. Springer Science & Business Media, 1986.
 - [22] A. D. Smith. Spatial navigation in autism spectrum disorders: a critical review. *Frontiers in Psychology*, 6:31, 2015.
 - [23] T. Tavassoli, R. A. Hoekstra, and S. Baron-Cohen. The sensory perception quotient (SPQ): Development and validation of a new sensory questionnaire for adults with and without autism. *Molecular Autism*, 5:29, 2014.
 - [24] R. Verma, S. Ghosh, M. Saketh, N. Ganguly, B. Mitra, and S. Chakraborty. Comfride: a smartphone based system for comfortable public transport recommendation. In *Proceedings of the 12th ACM Conference on Recommender Systems, RecSys '18*, pages 181–189, New York, NY, USA, 2018. ACM.
 - [25] D. von Winterfeldt and W. Edwards. *Decision analysis and behavioral research*. Cambridge University Press, Cambridge, UK, 1986.