POI-based Recommender System for the Support of Academics in a Smart Campus

Guilherme Xavier de Carvalho\textsuperscript{1}\textsuperscript{a}, Isabela Gasparini\textsuperscript{1}\textsuperscript{b}, Guilherme Medeiros Machado\textsuperscript{2}\textsuperscript{c}, Leandro Krug Wives\textsuperscript{2}\textsuperscript{d} and José Palazzo M. de Oliveira\textsuperscript{2}\textsuperscript{e}

\textsuperscript{1}Santa Catarina State University (UDESC), Joinville, SC, Brazil
\textsuperscript{2}Informatic Institute - PPGCC, Federal University of RS (UFRGS), Porto Alegre, Brazil

Keywords: Recommender Systems, Smart Campus, Points of Interest.

Abstract: Recommender Systems are software used to suggest user items in a personalized and automated way. When combined with Points of Interest (POI), they can set locations as referable items. This type of approach is useful when the amount of POI available for the user is large. In the context of Universities, students have different needs and have to look for different locations to experience the Universities’ resources. The goal of this paper is to present a POI-based Recommender System to improve student’s well-being and to support their academic journey in a Smart Campus. The recommender system was implemented by an application called AONDE, which was used by 110 students, where 63 answered a satisfaction questionnaire allowing the data collection needed for the system evaluation. An accuracy of 61\% in the recommendations of items to students was measured, as well as a high satisfaction rate, where 90.5\% of respondents said they were satisfied or very satisfied with the locations suggested by the app. The purpose of this experience paper is to demonstrate that the approach here described proved to be useful for students’ routine, impacting positively their academic journey.

1 INTRODUCTION

Recommender systems (RS) were first proposed as a solution to deal with the problem of user cognitive overload, where the amount of information to be analyzed exceeds the user capability (Machado et al., 2018). RS is a subclass of information filtering systems that seeks to predict the “rating” or “preference” a user would give to an item (Ricci et al., 2011). They are used in a variety of applications, such as suggestion of movies (e.g., Netflix), music (e.g., Spotify), or videos (e.g., YouTube). When combined with Points of Interest (POI), they can set locations as referable items. This type of approach is useful when the amount of POI available for the user is bigger than the user capacity to analyze the full set. Therefore, there is a need to customize recommendations to meet the interests and needs of the user. Such systems can be used in intelligent scenarios to improve the quality of life of individuals in this environment; since data filtering reduces the number of items associated with users’ domain, their cognitive and informational load are reduced.

The goal of this paper is to present a POI-based Recommender System in a Smart Campus to support the academic journey in the environment. For the system development, a literature review was performed in search of related work and the basic concepts were identified. The recommender system was created, aligned primarily with the interests and needs of its target audience. The algorithm implemented represents the users and the items to be recommended through a vector of tags, which are keywords that describe such an entity. Both users and items are structured in matrices and weighted by the metric of Term Frequency–Inverse Document Frequency (TF-IDF). Finally, the similarity between items and users is calculated and generates the recommendations that are presented through a web application, called AONDE, which allows interaction with the user.

AONDE was used by 110 students of the Santa Catarina State University - UDESC, and 63 answered...
a satisfaction questionnaire and thus allowed the collection of inputs used in the evaluation of the system. Through these, an accuracy of 61% in the recommendations of items to users was measured, as well as a high satisfaction rate, where 90.5% of respondents said they were satisfied or very satisfied with the locations suggested by the app.

2 DEVELOPED RECOMMENDER SYSTEM

In this section, we present the system that was designed and developed to provide POI recommendations inside a smart campus. The strategy used in the system is content-based since we did not have an annotated dataset (items previously evaluated). The system is developed to attend specific student stereotypes, i.e., newcomers and veterans. More specifically, it is designed to help newcomers, i.e., students that recently arrived at the campus and do not know the environment, the routines, neither the resources available to them. It is also able to help veteran students to get involved with current activities and news, since we found that they are usually more focused on finishing their courses and do not have time to keep up to date.

Based on these stereotypes, and also the user interaction, the system is able to identify the current user needs and interests, classifying and prioritizing POI accordingly to each user profile. Thus, not only the most interesting items are recommended, but also the ones most necessary to the academic life of each profile.

2.1 POI and User Modeling and Representation

Both items to be recommended by the Recommender System and user interests are represented within the application through tags, keywords that define an element in a short, straightforward and clear manner. Therefore, it was necessary to define the best tags for each POI. A questionnaire was designed to collect students’ perceptions of the locations found on the university campus, as well as how they would describe them using tags.

The questionnaire was organized into three sections, which sought to generate mechanisms so that we could understand, through user responses, some aspects such as: which locations on campus were interesting for students to know, with which tags certain locations could be described, what was the profile of freshman and senior student, what were the interests of students as members of a college campus, what was the relationship of the student’s course and the places of interest.

Before being applied to the general public, a pilot test was accomplished with students of the Human-Computer Interaction discipline. This pilot test was conducted to verify if the questionnaire was well structured and organized and the interaction to answer were analyzed. Such students contributed to the refinement of the research instrument and once validated, it was sent for access to all students on the university through institutional emails and shares in social networking groups of the institution. The questionnaire was available from August 6 to September 12, 2019; and 133 responses were collected. It is noteworthy that there was at least one student response from each of the 21 courses offered by the university, among undergraduate and graduate students.

Based on information from this questionnaire and three standard university documents the Freshman Manual, the Extension Project Guide, and the Pedagogical Curriculum Guidelines from UDESC, we were able to define the representation of POIs, courses, user interests and freshman stereotype using a folksonomy.

According to Pereira e Silva (Pereira and da Silva, 2008) the folksonomy technique represents an initiative to assist in the process of organization and attribution of meaning to the contents available on the web, as well as for the treatment of information overload. Thus, it is a content analysis and organization approach based on the view of the interviewed audience and the author. The process of categorization of information is based on the three pivots of folksonomy, the user - who performs the categorization; the object - which is categorized; and the tags - which categorize this object. It is worth mentioning that this is exactly the path to be followed through the questionnaire directed at the students.

Applying folksonomy can be understood as two very distinct moments: information categorization and information retrieval. In the first, users are given a freedom of work, which enables them to generate hypotheses and great information through a practical and inexpensive process of both time and cognitive effort. The second moment is related to the classification of terms chosen during categorization, in short, is to evaluate the quality of the answers and make decisions from them. At this time freedom of work is on the side of the experimenter, so it is possible to make decisions by looking at the inputs collected from users in the first phase of the approach and also from the convictions that the project scope already has (Pereira
The following sections explain how each of these elements was elaborated, considering the glued sampling and the applied folksonomy techniques, to be implemented in the web application.

A survey had been carried out to define primarily which POIs would be recommended by the system. After data collection, it was possible to understand which of those POIs were indeed relevant to the students’ context. Such definitions emerged by applying folksonomy techniques to the information base collected and thus reached the number of 86 recommendable POIs within the UDESC Smart Campus.

Once the POIs were chosen, attention turned to defining the tags that would describe a particular location. At this time, the students’ answers were also analyzed in order to understand their perceptions about the content that would serve as a description of a particular POI. It was also through them that it was possible to extract the keywords for each of the recommended places. However, it is important to note that not all tags that make up the RS came directly from the students’ verbalization or responses in the questionnaire.

Having the folksonomy presented by Pereira and Silva (2008) as the basis for this process of analysis and definition, decisions were made looking at data sampling based on the scope of the project. Thus, some tags were elaborated through the interpretation of the author of this work under the profile of the students who participated in the research, relating the academic journey of a student with the locations listed as POI. We defined 83 distinct keywords, which were scattered among the items to be recommended contextually, where each POI has at least 3 tags describing it.

Still on this process it is useful to emphasize that the task of establishing the keywords for the POI needs some care. Since if many POIs have a high tag similarity rate, it will be difficult for the RS to differentiate the specialties of each recommendation item. Therefore, such a system may recommend items that are not necessarily interesting to a user but which have a high similarity index between the suggested tags. On the other hand, if each POI has tags that are significantly specific to it, it will be difficult to correlate the items to be recommended with the user’s interests more generally, thus requiring explicit and individualized statements about the interest in each POI.

We have designed two scenarios to show the feasibility of the system, one for each stereotype previously defined.

The first one considers a student named Aurora, a freshman student at the university. She is informed by the administration that a application is available to help students get localized and to know the university. She installs it on her smartphone and fills in her profile and chooses some categories of interest: Culture, entertainment and research. Right after, the system starts calculating the recommendations and shows different POIs. Aurora starts her tour in the campus, very excited, and starts sharing her experience with colleagues, which, in turn, start using the application too. She finds out that some recommendations, even if her colleagues have chosen the same categories, are different (most of them are older than her in the university). Hers interests are more related to bureaucratic activities, needed in her initial stage. The second time she used the application, the profile was already filled and the recommendations were immediate, but she was able to change her interests, personalizing the recommendations.

The second scenario considers a student named Tobias, who is in the 3rd year. He checked the bulletin board and found the information about a new app. Surprised by such news, he first wondered “When did they create it?”. Even suspicious about what it was really about, he accessed the application. Tobias created his profile by entering his personal details and selecting the categories that seemed convenient. A few moments later the system presented him with places that would be of interest to him, respecting the newly created profile. Analyzing the suggestions that appeared on his list, Tobias realized that even though he was a veteran, there were still several locations on campus he didn’t know about. The student was interested in the functionality of the application, because it became more updated with the actions that take place at the university. Using other times, he noted that he can modify his profile and disregard elements that are associated with activities he already knows, such as the restaurant for example.

As our RS is content-based, the way in which the content of the items is represented should be made explicit. In this case, the items to be recommended through the system are the POI. Therefore, these are modeled through a structure where each POI is associated with a set of keywords that describe that respective POI, functioning as tags. An array of characteristics (items x tags), where each line represents a POI, so there will be a vector for each POI.

Considering these aspects and the feasibility of using tags to describe items, the second phase of the content representation process deals with the weighting of the items. According to Jannach et al. (Jannach et al., 2010), the purpose of a content-based recommender approach is not to maintain a list of items and their static meta-information, but to categorize their
relevant keywords in the form of weights. Therefore, items are typically described using the TF-IDF encoding format (Term Frequency – Inverse Document Frequency) (Cantador et al., 2010). Following the approach via tags, where tags are used instead of documents, the TF-IDF encoding was adjusted and according to the literature it can be applied in two ways: in the first, only the Term Frequency (TF) is calculated, in the second the full scope of the formula is taken into account. In this work, only the first version of the formula is applied. Which describes an item by means of tags, and precisely by using this artifice there is no need to check the inverse frequency of a term in the document (Cantador et al., 2010), after all there are only a set of tags describing each item to be recommended. Thus, the profile of an item is represented by the vector \( i_n \) expressed by the equation below:

\[
i_{n,t} = t f_{n}(t) \tag{1}
\]

Where \( t f_{n}(t) \) corresponds to the number of times \( (tf) \) an item \( i_n \) was highlighted with the tag \( t_j \) to a certain item \( i_{n,j} \).

At the end of the weighting process, the set of items will be represented as a vector of the weights computed by the TF.

The representation of the user occurs in a similar way to that of the item. In this scheme, user profile is represented through a matrix of interest structure (user x tags), where each line represents a user, so there is a vector for each user. Each cell in the matrix is linked to a specific tag and through this feature it is possible to know which tags a user is associated with. It is worth mentioning that these tags must be collected through the user’s past interactions or explicitly asking the user when the user creates a profile in the application.

Weights are also plotted in a vector space, adjusted noting only by interest, but also by the user’s needs. Therefore, the weighting of a user’s profile consists of the vector \( u_m \) expressed by:

\[
u_{m,l} = t f_{m}(t) \tag{2}
\]

Where \( t f_{m}(t) \) corresponds to the number of times \( (tf) \) that user \( u_m \) has been associated with \( t \) for a given user \( u_{m,t} \). After calculating the TF the user vector proceeds to another weighting step, since still in the user representation, it is necessary to consider the recommendations for the freshman user profile. For this purpose, a calibration in the value of the items is incorporated.

Knowing that the user is a freshman, it is possible to evaluate the interest vector of that user to give more or less weight to certain interests, through pre-processing. The evaluation in this stage considers the stereotype of a freshman user. This stereotype is a user vector \( F \), which followed the representation established for the user, but which already had predefined tags according to the items (POI) mapped as necessary for a freshman.

These POI considered necessary for freshmen were taken from the University Freshman Manual, which defines some main locations within the University campus. Therefore, this \( F \) stereotype is applied to all users who use RS and are freshmen, that is, they have been in the university for less than a year. It is, therefore, a calibration imposed on the vector already calculated for these users. Such calibration is performed by means of a simple average between the vector of the user already weighted and the vector \( F \), in order to increase the weight of tags related to the items defined as necessary for the freshmen. After the user modeling process is finished, it is possible to proceed to the recommendation algorithm, where the similarity between item and user is calculated.

### 2.2 Finding Relevant POI

After the item and user were processed, the third stage of processing deals with the similarity between the item and user profiles. This is when the recommendation is actually generated. For that, the cosine similarity method, established as state of the art in this context is applied (Ricci et al., 2015).

The proximity calculation can be performed using the following equation, which is applied to the item and users weighted vectors only by TF:

\[
sim_{Cos}F(u_m,i_t) = \frac{\sum_{l=1}^{N} t f_{m}(t_l) \cdot t f_{0}(t_l)}{\sqrt{\sum_{l=1}^{N} (t f_{m}(t_l))^2 \cdot \sum_{l=1}^{N} (t f_{0}(t_l))^2}} \tag{3}
\]

Where \( \sim_{Cos}F(u_m,i_t) \) corresponds to the similarity coefficient between the user vectors and items (value belonging to the set \([0,1] \)): \( t f_{m}(t_l) \) is the number of times that the user \( u_m \) has been associated with textit tag \( t_l \); and \( t f_{0}(t_l) \) deals with the number of times that the item \( i_t \) has been marked with the textit tag \( t_l \).

It is worth mentioning that for the calculation of similarity, the corresponding mathematical expression was implemented in an algorithm in the Python language, without the use of any specific library.

At the end of this step, the items to be recommended to the user have already been defined. Therefore, they are subject to suggestion and can be presented to the user.
2.3 Architecture of the System

With regard to architecture, the system is based on the client-servers concept. A structure where service providers, i.e. servers, are separated from those who request the service, e.g. customers (Lee et al., 2004). This relationship is conducted in such a way that customers use server resources, without having direct access to them. For that, it is on the basis of the exchange of requests that the application works.

The architecture is organized in three layers: presentation layer, business layer and data layer. This model is implemented to organize the main entities that make up the app. The presentation or interface layer is the layer that interacts directly with the user. Therefore, it is through it that requests for consultations are made and the results are displayed. The business layer, on the other hand, is the logical part, where the RS itself is, in order to relate to the other two layers for its full functioning. The third layer of this organization is the data layer. It is the entity composed of the information repository that receives requests from the business layer and executes these requests in the database.

The user receives recommendations through an application. When accessing it, the user is interacting in the presentation layer, without the need to know the processes behind the RS itself. It is worth mentioning that the application is on the client side, in the client-server relationship, and relates to the server side when sending and receiving data.

On the server side are the data and business layers. Respectively, the first layer represents the database with all the necessary information in the recommendation process that is manipulated. The second, on the other hand, stores the recommender’s business rules, which is in fact the logical and coded definitions of the RS. Once the architectural organization is established, it is possible to specify the technical aspects of the system in each of the layers. For the presentation layer, a Web App was developed, a responsive web application accessed from any browser installed on computers or smartphones. Such a system was developed using HTML5, CSS3, JavaScript and PHP technologies.

The Web App communicates with the other layers through RESTful API services developed in Python. It is worth mentioning that the application design process was guided by the Double Diamond model, based on the Design Thinking approach. The data layer of the system is structured by the PostgreSQL DBMS.

As for the business layer, this is where the logical part of the recommendation algorithm is located. RS was implemented in Python and using the tools of the language itself, a server was also built. Such business layer resources were hosted in a structure that allows users to access remotely. In this sense, the hosting was performed on a server within the University’s infrastructure. Figures 1 and 2 present the recommendation from AONDE application.

3 EVALUATION AND RESULTS

The AONDE application was available to the University students and the experiment was carried out between October 21 and 31, 2019. The release was communicated through an institutional email, which
would make it possible to reach a large part of the students on the university campus.

During this current period of the experiment, 110 students registered in the application. The inputs generated from the interactions where they were monitored periodically through the evaluation questionnaire and the database. After this period the access to the application has been restricted and the data manipulation and analysis phase has started.

For the process of evaluating the AONDE application and its recommender system, the metrics of accuracy of the Recommender System (RS) and user satisfaction regarding the use of the application were established. So that such metrics could be analyzed, two instruments were created. The first is related to the formalism of precision and recall used to calculate the accuracy of an RS, which is combined with a feedback system (like and dislike) inserted in each recommended POI. While the second is a questionnaire created to collect evidence related to the usability of the system and user satisfaction.

The analysis of the data from this experimental mechanism is based on the calculation of precision and recall. In order to precision and recall to be calculated in a Recommender System, the top- $k$ items of greatest relevance delivered by the system to the user are used. In the context of our system, it was defined that the 8 items with the highest recommended similarity rate would represent this set. Precisely because this is the number of POIs shown on the first page. Therefore, the accuracy and recall were calculated for each user who interacted with the system.

As an example, imagine that a student named Bob accessed the system and received his recommendations after creating his profile. Bob interacts with the system, navigating between pages, give “likes” and “dislikes” in some of the suggested locations. This interaction provides the necessary data to assess the accuracy of the recommendations generated for Bob. Suppose he liked 5 of the first 8 POIs presented, that is, those on the first page, and that in total he liked 13 of all 86 recommended items, including those on other pages. In this case, the precision would be $5/8 \approx 0.63$, that is, 63%, where 5 corresponds to the number of items defined as relevant selected by the user, and 8 the total number of items considered relevant. In this case the recall would result in the proportion $5/13 \approx 0.38$, that is, 38%, where 5 corresponds to the number of items defined as relevant selected by the user, and 13 the total number of items selected by the user.

Applying this logic and analyzing the data collected through the interactions of users with the items recommended by AONDE, we can measure an accuracy of 61% and a recall equal to 66%, this having limited to 8 the number of items considered relevant. An accuracy of 61% represents that on average, for every 10 items recommended, 6 met the expectations of users, were in line with the profile and aligned with the context. Therefore, it can be said that 61% of the suggestions through AONDE were in fact assertive. Thus, the probability that a selected item is the one considered relevant for a student who interacted with AONDE is 61%.

The RS recall was 66%, this represents the average number of times a relevant item is selected from the total number of items selected. Thus, the greater the recall, the greater the number of items considered relevant to have been selected from the total. In this context, with a 66% recall, it can be said that of all items selected by the student during use, 66% were among those considered relevant (top-8).

In short, the metrics collected through the data samples point to a system with good recommendation accuracy, considering the values of 61% for precision and 66% for recall. The numbers indicate that at least more than 60% of the recommendations generated by the RS were relevant.

Of the 110 students who used AONDE application and received recommendations, 63 students (i.e. 57%) answered the satisfaction questionnaire. Of these students, 52.4% are male and 49.2% are freshmen. Students from 16 different courses participated in the assessment (from 21 courses available in the campus).

Regarding some usability issues from the questionnaire, they were presented in a 5-point likert scale. Figure 3 presents the results. As Figure shows, most students strongly agreed with the usability issues of the app.

Regarding the question in which the student is invited to say how useful for the academic journey of a university student is to know the Points of Interest (POI) recommended by the system, using a scale from 0 to 10, where 0 represents a lot useless and 10 very useful, the Figure 4 presents the results. It can be seen that the vast majority of academics agree that knowing such places is of great use, when voting 41.3% for option 9 and 38.1% for option 10, totaling 50 of the 63 users who participated in the evaluation.

Students were also asked how satisfied they were with the recommendations they received. As can be seen in the Figure 5, the vast majority of those who used the system and participated in the evaluation said they were satisfied or very satisfied with the recommendations received. Together these two groups represent 57 students, 90.5% of those who responded to the evaluation. Finally, the students had the op-
Figure 3: Usability and User satisfaction with the application.

Figure 4: Utility of the system.

portunity to leave suggestions to AONDE improvements. Among the collected ideas are to present the geographical location of the POIs, as if in the form of a map or using a GPS system in real time. In addition, several students showed interest in learning more about POIs, so that each one had more details. Another aspect raised was the possibility of knowing the value of similarity for each recommended item, in the same way creating gamification strategies or integrating the application to social networks.

4 CONCLUSION

Recommendation systems (RS) are artifacts used to manipulate large data sets, in order to reduce their complexity by identifying and suggesting to users only items of most interest or need. These software tools seek to reduce the information burden while they want to extract elements relevant to the context of use. In the case of an environment integrated by an expressive set of information, a Smart Campus proves to be a suitable space for the application of an RS. Therefore, an RS was proposed in the domain of the University, whose objective is to support the students in their academic journey, relieving their cognitive and informational burden, with regard to the recommendation of Points of Interest (POI) within the university campus.

After deepening and understanding about RS, the proposal of this work was conceived, mainly aligned with the vision of its target audience, since this project is concerned with the use experience that it will offer. Personas were drawn, representing the user’s profile in the Smart Campus domain, resulting in a content-based POI Recommender System, linked to a web application called AONDE.

The proposed algorithm represents the users and the items to be recommended through tags, that are keywords that describe such an entity. In this way, both users and items are structured in matrices and weighted by the TF metrics.

An interesting point of the RS developed in this work is its concern in recommending items that are necessary for a student in his/her first steps at the university, the freshman. A bias is applied to the user’s profile, considering this stereotype, so that items fundamentally necessary for this type of student receive a greater weight.

Finally, the similarity between items and user is calculated, using the cosine similarity, and generate the outputs for recommendation. During the devel-
POI-based Recommender System for the Support of Academics in a Smart Campus

Figure 5: User satisfaction with the recommendations received.

The development phase, several tests were carried out in order to align users’ expectations. In addition, a pilot test was conducted before the system was released to the public. University students were able to interact with AONDE, personalize their profile and receive POI recommendations within the campus. In the period of the experiment, 110 users were created, which enabled the present project to collect sufficient data for analysis.

Having the accuracy of RS and user satisfaction as metrics for evaluation, the RS was evaluated. Through a feedback system present in each recommended item and a satisfaction questionnaire, evidence was collected that guarantees the benefits of the system. It can be concluded that the system corresponds to the desires of its target audience and actually contributes to the academic journey of a student on campus, since 90.5% of the students say they felt satisfied or very satisfied with the recommendations received. Still on this evaluative perspective, the system presented an accuracy in the recommendations with an accuracy of 61% and a recall of 66%.

As future work the Recommender System can be enhanced, since other recommender approaches could be included in the system. For example, as more students use the system on a regular basis, the collaborative approach can be combined with content filtering. This work offers the opportunity to connect several other elements that may contribute to its dissemination and use in the university community, such as a gamification system, whose goal would be to create mechanisms of engagement and contribution to the platform, disseminating knowledge and sharing experiences in a playful and natural way.

ACKNOWLEDGEMENT

The authors would like to thank the funding provided by FAPESC (public call FAPESC/CNPq No. 2017TR1755 - Ambientes Inteligentes Educacionais com Integração de Técnicas de Learning Analytics e de Gamificação) and UNIVERSAL MCTI/CNPq Nº 01/2016 - Recomendação adaptativa para cidades inteligentes. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) Finance Code 001.

REFERENCES


06/2016 support the infrastructure of CTI for young researchers, project T.O. No.: 2017TR1755 - Ambientes Inteligentes Educacionais com Integração de Técnicas de Learning Analytics e de Gamificação) and UNIVERSAL MCTI/CNPq Nº 01/2016 - Recomendação adaptativa para cidades inteligentes. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) Finance Code 001.