Tourism Integrated Recommender System: Setubal Peninsula Case Study

Mohamma Julashokri[1][a], Suzana Monteiro Leonardi[2][b] and Pedro Seabra[1][c]
[1]*Viatecla SA, Estr. Algazarra 72, 2810-013 Almada, Portugal*
[2]*Faculdade de Letras, Universidade de Lisboa, Lisboa, Portugal*

Keywords: Tourism, Recommender System, Collaborative Filtering, Content-based Filtering, Point of Interest.

Abstract: The diversity and a huge number of different places and attractions can make the decision for tourists difficult. The recommender systems are developed to facilitate people's decision-making process that can help in the area of tourism as well. In this paper, we proposed and implement a recommender system that works integrated with a Setubal peninsula portal to help tourists to choose experiences and points of interest to visit. The implementation was done using collaborative and content-based filtering to make recommendations based on user profiles and activities within the portal.

1 INTRODUCTION

The tourism market has grown enormously during the past 60 years. The number of international tourist trips in 1950 was 25 million, however, this number in 2000 became 674 million, and increased to 1.235 million in 2016. From a profit perspective international tourism in 1950 has a market of $ 2 billion and in 2000 became $ 495 billion but the market increased to $1.220 billion in 2016. The prediction for market growth in international travel is 3.3% per year until 2030 and will reach $1.8 billion. (Huang et al 2017).

Recommender systems can help tourists to select between a large number of destinations. In the tourism market, the recommender systems suggest different things like experiences, places to visit, hotels, etc. using different methods like content-based filtering, collaborative filtering, and knowledge-based filtering or their composition (Santos et al 2016).

In the content-based filtering method, characteristics of users' interests are extracted and other items that have the same or similar characteristics will be recommended (Park & Chang, 2009, Julashokri et al, 2011); While collaborative filtering considers similarities between user profiles. A user's profile in collaborative filtering consists of user-specified ratings to items and user demographics and the recommendations will build based on similarities between user profiles (Montaner et al., 2003; Park & Chang, 2009, Julashokri et al, 2011). Knowledge-based filtering is to use any kind of knowledge in the recommendation-making processes, like when the system has knowledge that how a specific item relates to a specific user (Alpekin & Buyukozkan, 2011).

The rest of this work is organized as follows. Section 2 outlines the background and reviews related works on customer profiles and recommendation methods. Section 3 illustrates the proposed methods. Section 4 describes the implementation of the proposed model. Finally, Section 5 draws conclusions and future works.

2 LITERATURE REVIEW

Many studies have dealt with recommender systems in the tourism area. Some of them proposed mobile recommender systems to help tourists to choose points of interest and others focused on the recommendation-making process.

Noguera et al (2012) present a novel mobile recommender system that brings together a hybrid

[a] https://orcid.org/0000-0003-0012-9456
[b] https://orcid.org/0000-0002-4434-5567
[c] https://orcid.org/0000-0002-5771-3648

Julashokri, M., Leonardi, S. and Seabra, P.
Tourism Integrated Recommender System: Setubal Peninsula Case Study.
DOI: 10.5220/0011102100003179
In *Proceedings of the 24th International Conference on Enterprise Information Systems (ICEIS 2022) - Volume 1*, pages 623-627
ISBN: 978-989-758-569-2; ISSN: 2184-4992

recommendation engine and a mobile 3D GIS architecture. Their system gives users a 3D map-based interface based on user location.

Nilashi et al (2017) proposed a recommendation method based on multi-criteria collaborative filtering to enhance the predictive accuracy of recommender systems in the tourism domain using clustering, dimensionality reduction, and prediction methods. They also used a cluster ensembles approach to improve the recommendation accuracy. Based on the TripAdvisor dataset, they evaluated their model to confirm that cluster ensembles can provide better predictive accuracy.

Smirnov et al (2014) proposed a context-aware recommender system that gathers information about the point of interest from internet sources like Wikipedia, wikivoyage, wikitravel, panoramio and flickrm. In their system, tourists can rate attractions that they like, or dislike and they used this information to find similar users using the collaborative filtering method. Smirnov et al (2014) presented a mobile app that recommends point of interest and the appropriate transportation to that POIs using users' context like location and weather as well as user profiles like trip length, interaction mode, etc.

Jorro-Aragoneses et al (2017) also proposed a context-aware recommender system to recommend leisure activities in Madrid, called Madrid Live. Their system asks for restrictions and preferences of the tourist and tries to satisfy restrictions to recommend a set of activities based on tourist preferences. The contextual information that the Madrid Live system uses to make recommendations are time, location, weather, users' budget, and if they wish to use public transport.

In Khallouki et al (2018), authors presented a new approach to make context-aware mobile recommender system that combines Internet of Things (IoT) technologies with semantic web services to predict the suitable recommends for user.

Alrasheed et al (2020) proposed a multi-level tourism recommender system framework that provides the user with a set of destinations liked by similar users to allow constructing a list of potential destinations. Then their system ranks the selected destinations depending on user preference and user constraint using data gathered from different Web portals to consider the dynamic context of each trip.

3 PROPOSED MODEL

This paper proposes the creation of an integrated recommendation system based on users' profiles and ratings to suggest both POIs and experiences based on the accuracy level of that content (POI, Experience, etc.). This algorithm shows the recommendation building process in the current work:

1: Choose the class of recommendations (can be more than a class)
- *1.1: Content-based: if the customer booked experience on that class or if he read an article about that class*
- *1.2: Collaborative-based: make an SVM model based on customer profile (based on user demographics, user bookings and user logs) and predict the class of experience that this user likes*

2: Get the location of person and make recommendation list based on (region, on make itinerary)

3: Filter the recommendations if they are open at the time, if the customer already visit that experience

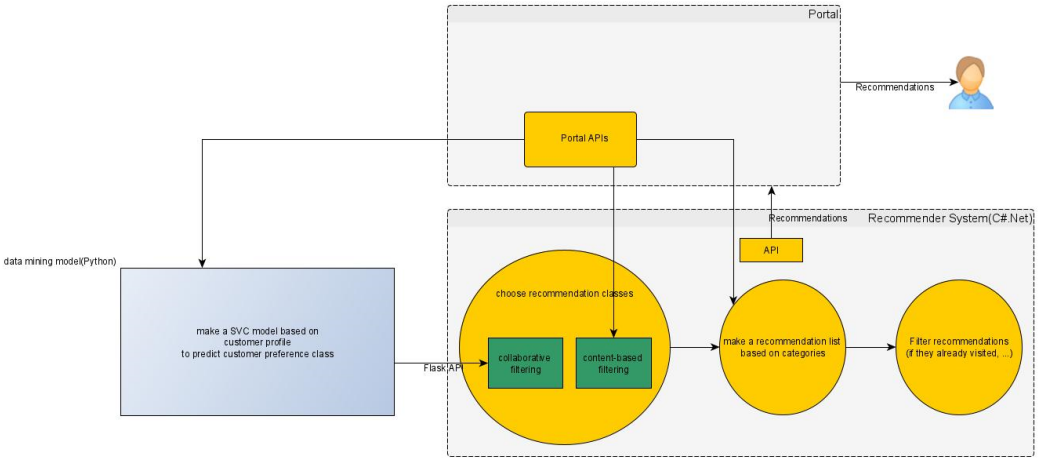

Figure 1: High-level architecture of proposed model.

In Figure 1, a high-level architecture is shown.

A. Portal

The "Peninsula de Setubal" is implemented to share the experience, articles, etc. to make people familiar with Setubal attractions.

For receiving data from this portal to build a data mining model in python or to different parts of the recommender system in C#.Net, a list of APIs was implemented.

The portal is responsible for calling the recommender system for each user and then sending the recommendations to the user using interaction channels like email.

B. Data Mining Model

To implement the data mining model for predicting customer preferences, the python application is implemented. The application receives a data file containing the user profiles labeled with the category of the content (experience, route, article, …). This historical data which is collected from user data in the Portal uses for building an SVM model using the scikit-learn library of python. Then an API that has been implemented with Flask library will listen to receive user data to call the model and prepare the recommendation categories(classes) based on customer profile.

C. Recommender System

The main part of the recommender system is an application written in C#.Net.

This application has an API that listens to receive, requests containing a user profile and will respond to the request with a list of recommendations that should send to the user.

In the first step, the application prepares the categories of recommendations from two sources: collaborative filtering and content-based filtering.

Collaborative filtering is implemented in python and described before as python block. The content-based filtering will give the list of categories that users booked or have logged on to them.

After having the categories from two resources a unique list with different weights for each list will be prepared.

In the second step, the list of recommendations in the categories will be extracted.

And in the third step, the recommendation list will be filtered and the extra recommendations like the ones that have already been booked by the user will be removed.

4 MODEL IMPLEMENTATION

The proposed model is implemented to integrate with the portal of "Enjoy Peninsula de Setubal" which is a complete tourist guide to promote attractions of Setubal peninsula in Portugal (https://visitsetubalpeninsula.scriptorserver.com/).

After preparing the data using Portal API and saving them in CSV file to start building the model, first, we compared the different types of data mining algorithms to see which one has the better answer on our data.

We used the python scikit-learn library to build our model using different algorithms:

- 'LR': Logistic Regression
- 'LDA': Linear Discriminant Analysis
- 'KNN': K Nearest Neighbors Classifier
- 'CART': Decision Tree Classifier
- 'NB': Gaussian NB
- 'SVM': Support Vector Machine

In the separated script we compare mentioned algorithms on our data to see which one can have better output results for us and we used k-Fold Cross-Validation.

Cross-validation is a resampling procedure to evaluate machine learning models based on a limited data sample. Cross-validation is primarily used in applied machine learning to estimate how the model is expected to perform in general when used to make predictions on data not used during the training of the model.

Figure 2 shows the comparison between different algorithms on sample data.

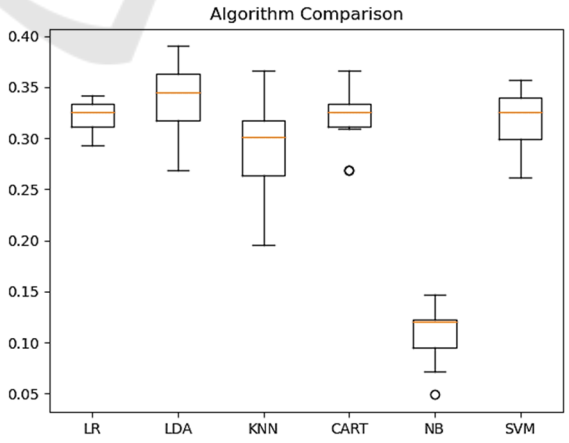

Figure 2: Comparison between different algorithms on sample data.

Based on figure 2 the best results appeared with LDA and SVM algorithms. As our data still test data

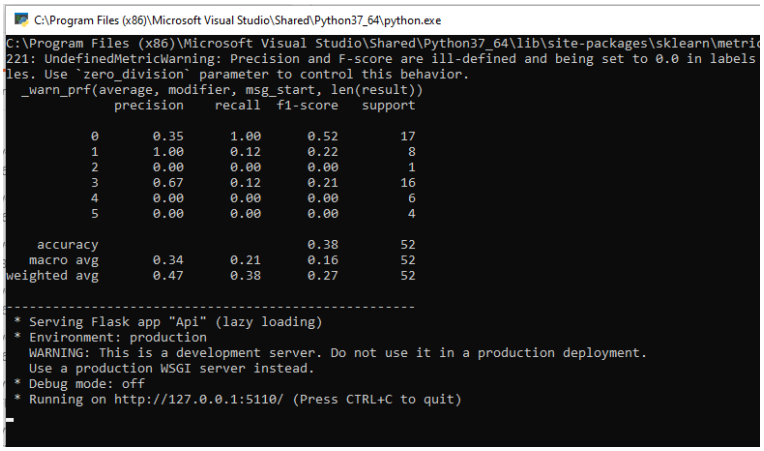

Figure 3: Python data mining model listen to receive new data sample for prediction.

and will be grown during the productional use of portal we decided to use the SVM algorithm for now, because the LDA algorithm supposes that the data points have the same covariance, and the probability density is normally distributed but SVM has no such assumption.

Finally, we made our data mining model based on SVM, and to have better performance on run time, we build the model one time then use an API to receive a new data sample and respond to the prediction for that data sample. As is shown in Figure 3, we developed the API using the Python Flask library.

5 CONCLUSION AND FUTURE WORKS

Nowadays, Recommender systems play a major role in any business that tries to sell products or services to customers and plays in a huge and diverse market. In the tourism area, the companies face with a huge number of peoples that have totally different wishes as well as loads of places and experiences which can be suggested to these customers.

There are a huge number of destinations that can make customers choose his-her next destination for holidays.

The aim of this study is to propose and implement a complete and integrated system that support all data and process from booking of experience, reading of the articles about POI, etc. to track the user feedback and recommend new POIs and experiences based on the user profile.

This system is making the recommendations using data mining algorithms for collaborative filtering and content-based filtering and works as part of a complete and integrated system to support all tourism needs to make the decision. The automated DevOps test showed that the recommendation system works correctly at the Unit and integration levels.

This integrated system will also have the location track and will improve its recommendations based on the user location and the time of the day in the future.

Also, after having more interactions with real customers the datamining algorithm parameters will be tuned; Because depending on data behavior the input fields' impact in the data mining model should change.

ACKNOWLEDGMENT

The research leading to these results has received funding from the Portugal 2020 program, project XPLORE_MARKET (ALT20-03-0247-FEDER-024296).

REFERENCES

Alptekin, G.I., Buyukozkan, G., 2011. An integrated case-based reasoning and MCDM system for Web based tourism destination planning. Expert Systems With Applications 38, 2125–2132.

Alrasheed, H., Alzeer, A., Alhowimel, A., Shameri, N., Althyabi, A. (2020). A Multi-Level Tourism Destination Recommender System, Procedia Computer Science, 170(333-340), https://doi.org/10.1016/j.procs.2020.03.047

Berger H., Denk M., Dittenbach M., Merkl D., Pesenhofer A. (2007) Quo Vadis Homo Turisticus? Towards a Picture-based Tourist Profiler. In: Sigala M., Mich L., Murphy J. (eds) Information and Communication Technologies in Tourism 2007. Springer, Vienna.

C.D. Huang, J. Goo, K. Nam, C.W. Yoo, Information & Management 54, 757 (2017), smart Tourism: Traveler, Business, and Organizational Perspectives.

Cufoglu, A. (2014). *User Profiling-A Short Review*. *108*(3), 1–9.

Cufoglu, A., Lohi, M., & Everiss, C. (2017). Feature weighted clustering for user profiling. *International Journal of Modeling, Simulation, and Scientific Computing*, *08*(04), 1750056. https://doi.org/10.1142/s1793962317500568

Cufoglu, A., Lohi, M., & Madani, K. (2009). A comparative study of selected classifiers with classification accuracy in user profiling. *2009 WRI World Congress on Computer Science and Information Engineering, CSIE 2009*, *3*, 708–712. https://doi.org/10.1109/CSIE.2009.954

Inoue, C. R. (2015). Tipos de revisão de literatura. *Tipos de Revisão de Literatura*, 9. Retrieved from http://www.fca.unesp.br/Home/Biblioteca/tipos-de-evisao-de-literatura.pdf

Jorro-Aragoneses, J.L., Díaz Agudo, M.B., Recio García, J.A. (2017). Madrid live: a context-aware recommender systems of leisure plans, 29th International Conference on Tools with Artificial Intelligence (ICTAI), Boston, DOI: 10.1109/ICTAI.2017.00125

Julashokri, M., Fathian, M., Gholamian, M.R., Mehrbod, A. (2011). Improving electronic customers' profile in recommender systems using data mining techniques, Management Science Letters 1, 449–456

Khallouki, H., Abatal, A., Bahaj, M. (2018). An Ontology-based Context awareness for Smart Tourism Recommendation System, Proceedings of the International Conference on Learning and Optimization Algorithms: Theory and Applications(LOPAL '18), 43(1-5), https://doi.org/10.1145/3230905.3230935

Kotsiantis, S. B. (2007) Supervised Machine Learning: A Review of Classification Techniques. Informatica **31,** 249-268

K.R.Ananthapadmanaban, A., & Srivatsa, S. K. (2011). Personalisation of User Profile: Creating User Profile Ontology for Tamilnadu Tourism. *International Journal of Computer Applications*, *23*(8), 42–47. https://doi.org/10.5120/2903-3808

Linden, Greg & Smith, Brent & York, Jeremy. (2003). Amazon. com recommendations: Item-to-item collaborative filtering. Internet Computing, IEEE. 7. 76-80.

Martín-Bautista, M. J., Kraft, D. H., Vila, M. A., Chen, J., & Cruz, J. (2003). User profiles and fuzzy logic for web retrieval issues. *Soft Computing - A Fusion of Foundations, Methodologies and Applications*, *6*(5), 365–372. https://doi.org/10.1007/s00500-002-0190-x

Nilashi, M., Bagherifard, K., Rahmani, M., Rafe, V. (2017). A recommender system for tourism industry using cluster ensemble and prediction machine learning techniques, Computers & Industrial Engineering, 109 (357-368), https://doi.org/10.1016/j.cie.2017.05.016

Montaner, M., Lopez, B., & De La Rosa, J. L. (2003). A taxonomy of recommender agents on the Internet. Artificial Intelligence Review, 19(4), 285-330.

Noguera, J.M., Barranco, M.J., Segura, R.J., Martínez, L. (2012). A mobile 3D-GIS hybrid recommender system for tourism, Information Sciences, 215(37-52), https://doi.org/10.1016/j.ins.2012.05.010

Norma Sandra Ferreira de Almeida. (2002). As pesquisas denominadas estado de arte. *Educação & Sociedade*, *79*(257–272), 257–272. Retrieved from http://www.scielo.br/pdf/es/v23n79/10857.pdf

Park, Y.-J., & Chang, K.-N. (2009). Individual and group behavior-based customer profile model for personalized product recommendation. Expert Systems with Applications, 36(2), 1932 -1939.

Ricci, Francesco. (2002). Travel recommender systems. IEEE Intelligent Systems. 17. 55-57.

Santos, F., Almeida, A., Martins, C., Oliveira, P. (2016). Gonçalves, R., Tourism Recommendation System based in user's profile and functionality levels, C3S2E '16: 9th International Conference on Computer Science & Software Engineering, 93-97, https://doi.org/10.1145/2948992.2948995

Schiaffino, S., Scientific, N., & Scientific, N. (2009). *Artificial Intelligence An International Perspective*. *5640*(January 2009). https://doi.org/10.1007/978-3-642-03226-4

Schiaffino S., Amandi A. (2009) Intelligent User Profiling. In: Bramer M. (eds) Artificial Intelligence An International Perspective. Lecture Notes in Computer Science, vol 5640. Springer, Berlin, Heidelberg

Smirnov, A., Kashevnik, A., Ponomarev, A., Teslya, N., Shchekotov, M., Balandin, S.I.(2014). Smart space-based tourist recommendation system: Application for mobile devices, Internet of things, smart spaces, and next generation networks and systems, Springer International Publishing, Cham, 40-51, DOI: 10.1007/978-3-319-10353-2_4