CEUR-WS.org/Vol-2943/restmex_paper4.pdf

A Recommendation System for Tourism Based on Semantic Representations and Statistical Relational Learning

Eliseo Morales-González[0000−0002−3186−0308], David Torres-Moreno[0000−0001−7844−956X], Alexandra Ehrlich-López[0000−0001−6880−8426], Mauricio Toledo-Acosta[0000−0002−8047−0189], Bolívar Martínez-Zaldivar[0000−0002−2153−4031], and Jorge Hermosillo-Valadez[0000−−0001−9040−767X]

Centro de Investigación en Ciencias, Universidad Autónoma del Estado de Morelos, C.P. 62209, Cuernavaca, Morelos, México
eliseo.moralesgon@uaem.edu.mx,jhermosillo@uaem.mx

Abstract. In this paper, we describe the methods used to submit our results to the Rest-Mex Recommendation System for Mexican Tourism task of the Iberian Languages Evaluation Forum 2021. We propose a Bag of Part-Of-Speech representation for text and to use statistical relational learning in order to predict a relation between a tourist and a place in terms of a target label, or recommendation.

Keywords: Tourism Recommendation System · Semantic Bag of Part of Speech · Statistical Relational Learning.

1 Introduction

Tourism has become a crucial source of revenue worldwide. From a socioeconomic point of view, tourism has become one of the largest and fastest growing industries in the world, extending activity online in the most recent decade [2]. In Mexico, this phenomenon is no exception, accounting for 8.7% of the national GDP, generating around 4.5 million direct jobs. However, with the COVID-19 pandemic, which began in Mexico in mid-March 2020, tourism was one of the most affected sectors in this country [1].

In this context, the use of Artificial Intelligence (AI), and in particular, Natural Language Processing (NLP), could be of great help to identify problems based on the analysis of the semantic aspects of tourists' opinions. In the case of tourism, a significant number of users express their views and opinions regarding the experience of traveling to a certain place through social media. These opinions

IberLEF 2021, September 2021, Málaga, Spain.
Copyright © 2021 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

are subjective information that represents the user's feelings, and the user's assessment associated with that experience. Online customer reviews of hotels and restaurants for tourism play a key role in decision making. Text reviews on travel websites can potentially influence destination selection. Tourists use this information to satisfy their preferences. Similarly, managers of tourism services and public institutions dedicated to promoting tourism can use this information to improve customer service. In this way, tourism content shared through social networks has become a highly influential source of information that may impact tourism in many ways. Thus, mining the opinions of tourists in search of the polarity of this opinion could influence decision making throughout the value chain and support this industry.

In this paper, we describe the methods used to submit our results to the Rest-Mex Recommendation System task of the Iberian Languages Evaluation Forum 2021 [1]. For this competition edition, the recommendation system problem is defined as follows: "*Given a TripAdvisor tourist and a Mexican tourist place, the goal is to automatically obtain the degree of satisfaction (between 1 and 5) that the tourist will have when visiting that place.*" The motivation of this competition is that few recommendation systems for tourist sites are based on a user's profile's affinity, compared to each place's description. The data collections to train these types of systems are from users and places in English-speaking countries. Considering the importance of Ibero-American countries in tourism, it is vitally important to generate Spanish resources that allow the generation of systems that help develop intelligent systems in tourism [1].

Studies about recommendations, suggestions or content filtering for the tourism sector can be traced back to the 80's of the last century [5], [16]. Today, Recommendation has become one of the most important methods for marketing and selling products and services over the Internet. Recommendation Systems (RSs) are software tools and techniques that provide suggestions of items that are most likely of interest to a particular user [20]. Thus, RSs are becoming more present in many websites and applications, providing us with recommendations on where to travel (e.g. Expedia), what music to listen to (e.g. Spotify), what movies to watch (e.g. Netflix), what to eat (e.g. Ifood), who to date (e.g. Tinder), or even discover who shares your same lifestyle interests [12], [6], [7].

The problem that concerns us here is the following. The input consists of data records that include: personal information about a tourist T; a history of opinions about places T has visited, associated with ratings T has given to those places; the name of the current place of interest and the rating (label) that T has given to this place of interest. The problem consists of predicting the labels that each tourist would give to a new place, not included in tourist's history.

Traditional techniques employed in RSs can be categorized in three approaches: Content-Based (CB), Collaborative Filtering (CF), and Knowledge-Based (KB). Hybrid systems are based on a combination of these techniques trying to leverage the advantages and mitigate the disadvantages of one over another [20]. In this paper, we propose to approach the recommendation problem from a hybrid perspective.

We consider a CB model, based on linguistic features associated with each tourist and each place, and a CF model, which seeks to make a prediction of the label that a tourist would give to a new place, based on the prediction of a Tourist-Label-Place relationship. To this end, we propose a new Bag-of-POS (BoPOS) type representation for modeling linguistic features, and the use of the ComplEx [23] model, which computes a complex matrix factorization to predict relationships between entities. In this regard, we assume that the model can make a good prediction of the probability that a relationship (Tourist, Label, Place) exists.

The structure of the paper is as follows. Section 2 reviews related works. Section 3 introduces the proposed methods. Section 5 presents the experimental results discussing the main findings. We conclude the paper in Section 6.

2 Related work

A CB recommendation system (RS) learns to recommend items that are similar to those the user has liked in the past. The similarity of items is calculated based on the features associated to the compared items. This technique identifies the common characteristics of items that have received a favourable rating from a user, and then it recommends to him/her new items that share those characteristics [10], [17], [18]. For example, when a user rated (positively) a point of interest (POI), the system can recommend similar POIs by calculating how similar these two POIs are according to their features.

On the other hand, CF is the process of filtering or evaluating items using the opinions of other people [21]. These opinions can be obtained explicitly from users, or by using some implicit measures, such as records of previous purchasing. That is, CF is an algorithm for matching people with similar interests for the purpose of making recommendations [20]. In other words, CF algorithms extract patterns of similarity in previous opinions about products or services, so that profiles of people with similar interests can be matched for the purpose of making a recommendation. For instance, a system may recommend a customer who travelled to Paris and Barcelona, to travel to Rome, because other (similar) users that travelled to Paris and/or Barcelona, travelled to Rome as well. There are two types of CF algorithms: (1) memory-based, where user rating data is used to calculate the similarity between users or items, and (2) model-based, which use data mining and artificial intelligence tools to predict user ratings of unrated items [19], [4], [8].

The KB technique works by recommending items based on specific domain knowledge about how certain item features meet users' needs and preferences and, ultimately, how the item is useful for the user [20]. In other words, it generates recommendations to the user based on the knowledge about his needs towards a particular item. These recommendations are performed under measures of utility, derived from the knowledge of the relationship between a specific user and item. For instance, a KB tourism RS will generate recommendations not only based on the past travel experience of the user, but also based on what are the characteristics of the places/cities visited and the places available to recommend, that is, a KB RS exploits knowledge to map a user to the products he likes [14], [3].

Finally, Hybrid systems are based on a combination of the above mentioned techniques. Hybrid RSs have been proposed in the past. For instance, [13] used associative classification to predict context and improve the recommendation; [22] used user interaction and collective intelligence; [15] in their hybrid project use collaborative filtering, content-based recommendation and demographic profiling; [11] introduced a hybrid RSs combining the Markov Model and topic models in which a user is modelled as a mixture of topics, and a topic is modelled as a probabilistic distribution over landmarks. More recently, [9] used K-nearest neighbors (K-NN) for both collaborative filtering and content-based filtering, and a decision tree for the demographic filtering in order to enhance the recommendation accuracy.

In this paper, we propose a hybrid system based on a Statistical Relational Learning approach, namely the ComplEx model [23]. The model is motivated by the fact that Web-scale knowledge bases (KBs) provide a structured representation of world knowledge, but the incompleteness of these KBs has stimulated research on the prediction of missing entries, a task known as link prediction [23]. The core idea is to model the link prediction task as a 3D binary tensor completion problem, where each slice is the adjacency matrix of one relation type in the knowledge graph. In this case, a partially observed matrix or tensor is decomposed into a product of embedding matrices with much smaller rank, resulting in fixed-dimensional vector representations for each entity and relation in the database. Given a fact $r(s, o)$, in which subject s is linked to object o through relation r, the score can then be recovered as a multi-linear product between the embedding vectors of s, r and o [23].

3 Methods

3.1 Corpus characteristics and goal

Each instance in the training set consists of a $UserX$ who recommends a $placeY$ according to a satisfaction degree, or $Label \in \{1, \dots, 5\}$. The training set contains 1582 instances, and the test set, 681.

The available information about $UserX$ is the following.

- Tourist's gender.
- Tourist's place: a set of places that $UserX$ has visited.

– Tourist's history: a set of $UserX$'s opinions about the places in Tourist's place.

– Tourist's satisfaction: set of satisfaction degrees that correspond to each location and opinion in the columns Tourist's place and Tourist's history.

The available information about *placeY* is the following.

– Kind of place.

– Place's description.

Our goal is to find statistical patterns ($UserX, Label, PlaceY$), which represent the relationship between a $UserX$ and target $PlaceY$ given by $Label$. In this sense, we need to create triplets of this type with the information available in the corpus.

3.2 Data pre-processing

Each instance of the training set consists of a tourist $UserX$, which is accompanied by the gender information and the tourist's history. This history is a set of opinions of various places together with their corresponding satisfaction degrees. Additionally, each training example has a target *PlaceY* associated with a *Label*, rating that place. Each opinion in the tourist's history is a short text, which in some way describes the tourist's appreciation of each place visited in the past, correlating with the rating value (tourist's satisfaction degree) given to that place.

The problem here is twofold. On the one hand, for each training instance, there is one target *placeY* associated with one target *Label*, but there are many past places and opinions associated with their corresponding satisfaction degrees, which do not always match the value of the target *Label*. On the other hand, the target place may not be available in the tourist's history. Table 1 below illustrates the problem for the record *User622*, rating "1" the target place "Isla de Coral".

In order to generate the relationship graph, we divided each record, creating a new one for each satisfaction degree value available in the tourist's history (Tourist's satisfaction). In this way, each record contains one *UserX*, accompanied by the gender, together with those tourist's places and histories corresponding to one single *Label*, and including the target *placeY* and *Label*. Therefore, in relation to the example shown above, we will have two new records, one for the satisfaction rate of 5, and another for the rate of 4, with redundant information regarding the tourist (*User622*), his/her gender, and the target place ("Isla de Coral") along with its recommendation value, or target *Label* ("1").

Not the least, we used WordNet as a linguistic resource, so all texts were previously translated to English.

Tourist's place	**Tourist's history**	**Tourist's satisfaction**
Islas Marietas	If you love to go to ”Hidden Beach” then the place is Marieta Island in Mexico. The island was formed by volcano activities. This is hidden away in the remote ...	5.0
OYO 23635 Sher E Bengal	Food quality is good enough. The accommodation is also available. Surprisingly at each hotel room has a different name. Hooka Bar is also available (Different f...	5.0
Garh Jungle	It is a historical place. The place is known as ”Medhasram”. I have come several times. The place is 20 kilometers away from the city.	4.0
New Embassy Chinese Restaurant	Food quality is very good. Especially mixed chowmin and chicken with french mushroom. You will not find that anywhere else.	5.0

Table 1: Data in record *User622*. The target place is “Isla de Coral” with a recommendation $Label = 1$.

3.3 Bag-of-POS representation

In this paper, we propose a text representation based on Part-Of-Speech features and synonymy-antonymy relations.

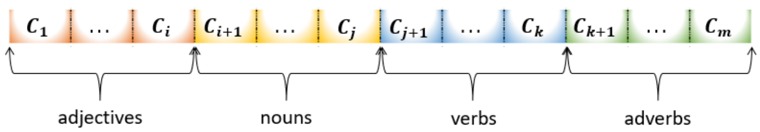
Part-of-Speech synonymy-antonymy relations. First, a collection of sets is created, chosen from 4 linguistic features, specifically the Part-Of-Speech, or the grammatical category of words, as follows.

We define V as the vocabulary of unique words from all the texts in the corpus. For each word $w \in V$ such that w is an adjective, noun, verb or adverb we create the set

$$C_i = \left\{ \begin{array}{c|c} \text{synonyms} & \text{antonyms} \\ \text{of} & \text{of} \\ w & w \end{array} \right\}.$$

However, if w already belongs to a previously created set, we discard its respective set C_i and move on to the next word. To obtain the synonyms (resp. antonyms) of w, we used the synonymy (resp. antonymy) relation provided by WordNet. We construct the sets in such a way that they only contain words with the corresponding grammatical form (i.e. adjectives, nouns, verbs or adverbs). At the end of the process, m sets are obtained.

Next, we order the sets according to their POS.



The first i sets correspond to adjectives, the next j-i sets are nouns, the following k-j-i are verbs, and the last m-k-j-i are adverbs.

Text vector representation. We now construct a text representation using the ordered sets. For each text t in the corpus, we assign an initial representation with the following structure

$R_t =$

0	0	0	…	0	0	0	…	0	0
1	2	3	…	$i-1$	i	$i+1$	…	$m-1$	m

We go through each word $w \in t$, if w belongs to a set C_i there are two options:

1. If w is on the left side of the set C_i, we add 1 in the i-th position of the representation of t. The result is as follows;

$R_t =$

0	0	0	…	0	1	0	…	0	0
1	2	3	…	$i-1$	i	$i+1$	…	$m-1$	m

2. If w is on the right side of the set C_i, we subtract 1 in the i-th position of the representation of t. The result is as follows.

$R_t =$

0	0	0	…	0	-1	0	…	0	0
1	2	3	…	$i-1$	i	$i+1$	…	$m-1$	m

After having processed each word w in the text t, the vector representation is the result of the addition and subtraction of 1's, which results in a vector of positive or negative integers which could look like this:

$$R_t = (0, -1, 3, 5, 5, 6, \ldots, 13, 4).$$

3.4 Relation representation

In order to model each triplet (user, label, place), we first represent the tourist using his/her gender, the tourist's history and the place's description, using the text representation method described in Section 3.3, and the name of the target place, concatenating each word by a hyphen. Thus, the Tourist representation is a vector consisting of three parts, concatenated by underscores as follows:

– Part 1: The name of *placeY*.

– Part 2: Gender of *UserX*.

– Part 3: The string consisting of the concatenation of the entries of R_t, the vector associated with the text in the Tourist's history.

To represent the place in the triplet, we use the Place's type and Place's description, that are associated with the triplet. Then, the representation of *placeY* is divided into the following parts:

– Part 1: The concatenation of each word in Place's type, removing stop-words.

– Part 2: The concatenation of the vector entries associated with the text in Place's description.

As an example, consider the case mentioned above, that is, the triplet (*User622*, *Type_1*, *Isla de Coral*). Using our BoPOS representation we obtain the new triplet $(R_{\text{User622}}, Type_1, R_{\text{Isla de Coral}})$, where

$$
\begin{aligned}
R_{\text{User622}} &= Part1\_Part2\_Part3 \\
&= \text{Isla-de-Coral\_Male\_0} - 13 - 5560 \cdots 134
\end{aligned}
$$

$$
\begin{aligned}
R_{\text{Isla de Coral}} &= Part1\_Part2 \\
&= \text{Isla-de-Coral\_0000} - 5560 \cdots 346 - 1.
\end{aligned}
$$

This final vectors is what we call a Bag-Of-POS, or BoPOS representation. Its interest is that it allows to compare tourists, or places, by encoding the semantic contents of the reviews provided. In the presence of a new tourist, with his own reviews, our BoPOS representation will allow to measure the similarity between tourists. A similar situation is expected in regard to the places. This is illustrated in Fig. 1.

To calculate the similarity between tourists the following applies. Consider two tourists t1 and t2, with their respective BoPOS m-dimensional representations R_{t2} and R_{t2}. The Jaccard coefficient is given by

$$
J(t1, t2) = \frac{|I_{t1} \cap I_{t2}|}{|I_{t1} \cup I_{t2}|}
$$

where, I_{t1} is the set of indices $i \in \{1, 2, 3, \ldots, m\}$ such that the i-th position in R_{t1} is nonzero. I_{t2} is similar to I_{t1}, but corresponding to R_{t2}. Then the similarity between the tourists t1 and t2 is calculated as follows,

$$
\texttt{sim}(t1, t2) = \begin{cases} J(t1, t2) + 1 & \text{if } t1\text{'s gender is equal to t2's gender} \\ J(t1, t2) & \text{otherwise.} \end{cases}
$$

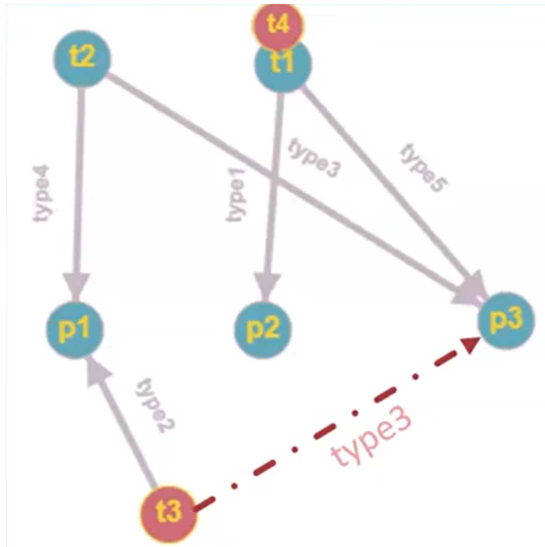


Fig. 1: Hypothetical relation graph illustrating the purpose of our BoPOS representation, aiming at measuring similarities between existing tourists (t1 and t2), and new tourists (t3 and t4).

3.5 The ComplEx model

In this subsection, we briefly review the ComplEx model, for further details see [23]. ComplEx models a link prediction task as a 3-dimensional binary tensor completion problem. Each slice of this tensor is the adjacency matrix of one relation type in the knowledge graph. Each relation type r consists of triplets (s, r, o), in which subject s is linked to object o through relation r.

ComplEx considers the tensor as the real part of a complex normal tensor, which is diagonalized, slice by slice, via the Spectral Theorem for normal matrices. Each slice X_r of the tensor is factored as

$$X_r = \mathrm{Re}\left(E W_r E^*\right),$$

where W_r is diagonal and E is unitary such that $EE^* = \mathrm{Id}$. It is worth noting that the matrix E is the same for every slice X_r. To predict whether a triplet exists, ComplEx uses the scoring function $\phi(s, r, o) = \mathrm{Re}\left(\langle w_r, e_s, e_o\rangle\right)$, where $\langle x, y, z\rangle$ is the component-wise multilinear dot product

$$\langle x, y, z\rangle = \sum_k x_k y_k z_k.$$

The vector w_r is the diagonal of the matrix W_r. The vector e_s (resp. e_o) is the real part of the row of E corresponding to the subject s (resp. the object o). The function ϕ returns the probability that the triplet exists.

The real and imaginary parts of the matrices E and W_r are initialized with vectors having a zero-mean normal distribution with unit variance. The model is trained using stochastic gradient descent optimizing the negative log-likelihood of the logistic model described with L2 regularization. After the training, the model learns the matrices E and W_r, and thus, the scoring function ϕ needed to predict the existence of missing links.

We use ComplEx in the following way. We consider the binary 3-dimensional tensor given by the 5 adjacency matrix defining each of the five relations type1, ..., type5 between representations of users and places described in Subsection 3.4. In our case, the subject s corresponds to $UserX$, and the object o to $placeY$. After training the ComplEx model, we use it to predict the probability of the existence of a triplet $(UserX, Type\_w, placeY)$.

4 Experiments

4.1 Experimental setup: Training Phase

The training instances for the ComplEx model (M) were constructed from the corpus of opinions, which was available for these purposes.

In this corpus, each record of a *UserX* recommending a *PlaceY*, with Label w, was modeled as described in Section 3. Then, we obtain a set of triples $(R_{UserX}, Type\_w, R_{placeY})$, denoted by $D1$, which is used to train model M.

To train M, we split the set $D1$ into two sets: $D1_{train}$ with the 80% of the triples in $D1$ and $D1_{test}$ with the remaining 20%. Thereafter, we performed validation tests to determine accuracy and recall.

4.2 Experimental setup: Test Phase

Once the models were trained, the unlabeled test data were received. We describe now the experimental setup used to produce the test results. The test corpus is made up of instances of a *UserX* recommending a *PlaceY*, for which we desire to predict the label.

The test corpus was modeled similarly as described in Section 3, but since, in this case, we do not have the labels, each instance of a *UserX* that recommends a *placeY* is modeled with the tuple $(Rep_{UserX}, Rep_{PlaceY})$. Thus, we obtain a set $D2$ of tuples from all the instances of the test set.

Since in $D2$ there exist representations of tourists that do not appear in $D1$, we used the similarity measure, described in Section 3, when performing the test task using $D2$. To obtain the recommendation Label $(Rep_{UserX}, Rep_PlaceY) \in D2$, in the range [1,5], all possible relationships, or degrees of satisfaction, between *UserX* and *placeY* are tested; that is, all possible triples

$$(Rep_{UserX}, Type\_w, Rep_{PlaceY})$$

are passed to ComplEx, with $i = 1, 2, ..., 5$. The predicted label is obtained by retaining the triple with the highest probability.

5 Results and discussion

Our results for the validation and test phases are shown in Table 2, including also the best participating test results.

Metric	**Validation** (training phase)	**Test phase**	**Baseline** (Majority Class)	**Best result**
MAE	–	1.65	0.73	0.31
Accuracy	39.85 %	20.91 %	53.81 %	77.28 %
Recall	39.70 %	19.54 %	10.76 %	52.85 %

Table 2: Prediction performance.

The Mean Absolute Error (MAE) was the primary metric used to determine the overall ranking of participants. With respect to this metric, our results were below the average. This is shown in Figure 2. This figure was obtained by taking the distribution of the results of all participants, in each metric reported by the competition organizers. The figure shows violin plots depicting the distribution of the results, where the average, the best result and our performance are depicted on this plot.

Looking at Table 2, we can see that, compared to the baseline (Majority Class), our model outperforms it 81.5% with respect to the Recall metric. However, with respect to Accuracy, our model is way below the baseline. This can be explained by the presence of the dominant class 5.

How can we explain these negative results? A possible reason is our representation method. Our aim was to extract the linguistic features in the text, specifically related to the lexical and semantic relations of synonymy-antonymy. In doing so, we expected to grasp similar degrees of satisfaction, as expressed by these semantic relations in each review. However, it is very likely that these similarities do not correlate well with the corresponding labels of each review. Therefore, it is necessary to model other aspects of language, and to delve into the linguistic features that could have a better incidence in the correlation between the tourist's opinion about a place and his satisfaction degree for that place. Another (correlated) cause is the similarity function applied to the tourists' representation. Indeed, the Jaccard coefficient measures representation similarities in terms of which POS terms are shared and which are distinct. The problem is that these similarities in POS contents, does not feature similarities in tourists' recommendations coincidences.

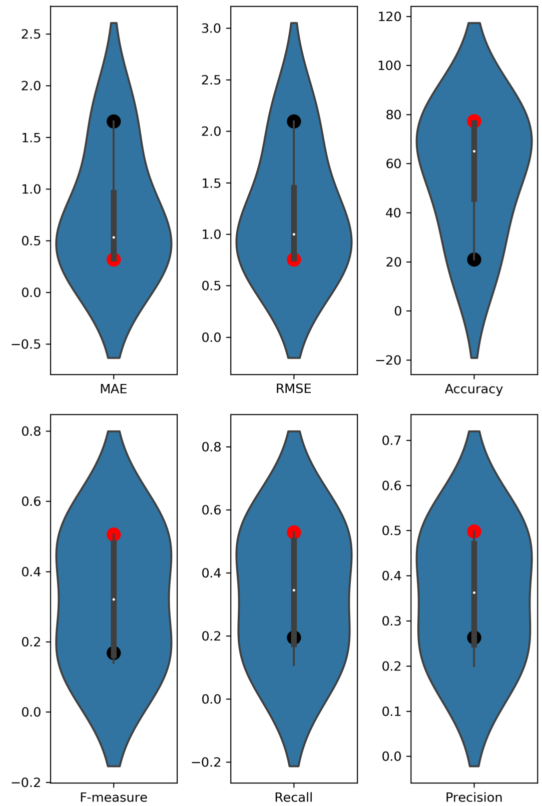


Fig. 2: Performances of all the participating runs in each of the metrics of the competition. The black dot in each metric is the performance of our method. The red point is the best performance in each metric. The white point is the average of the performances.

6 Conclusion

In this paper, we presented a method of text representation different from the methods of lexical co-occurrence in text. This method extracts the linguistic features in the text, specifically the lexical and semantic signals of synonymy-antonymy. We proposed to use the ComplEx model [23] for the recommendation task. The model was modified to perform the prediction of the target label, considering it as a relationship between a User and a Place. The results obtained are somewhat negative. However, better performance may be obtained by improving the text representation method, in order to improve similarity measures. Also, we hypothesize that performance may improve, if we focus on the linguistic features that could have a better incidence in the correlation between the tourist's opinion about a place, and his satisfaction degree for that place.

Funding

Research was partially funded by CONACYT Project A1-S-24213 of Basic Science and CONACYT grant 28268.

References

1. Álvarez-Carmona, M.Á., Aranda, R., Arce-Cárdenas, S., Fajardo-Delgado, D., Guerrero-Rodríguez, R., López-Monroy, A.P., Martínez-Miranda, J., Pérez-Espinosa, H., Rodríguez-González, A.: Overview of rest-mex at iberlef 2021: Recommendation system for text mexican tourism. Procesamiento del Lenguaje Natural **67** (2021)
2. Anis, S., Saad, S., Aref, M.: Sentiment analysis of hotel reviews using machine learning techniques. In: International Conference on Advanced Intelligent Systems and Informatics. pp. 227–234. Springer (2020)
3. Baraglia, R., Frattari, C., Muntean, C.I., Nardini, F.M., Silvestri, F.: A trajectory-based recommender system for tourism. In: Proceedings of the 8th International Conference on Active Media Technology. p. 196–205. AMT'12, Springer-Verlag, Berlin, Heidelberg (2012). https://doi.org/10.1007/978-3-642-35236-2_20, `https://doi.org/10.1007/978-3-642-35236-2\_20`
4. Chang, C.C., Chu, K.H.: A recommender system combining social networks for tourist attractions. In: Proceedings of the 2013 Fifth International Conference on Computational Intelligence, Communication Systems and Networks. p. 42–47. CICSYN '13, IEEE Computer Society, USA (2013). https://doi.org/10.1109/CICSYN.2013.52, `https://doi.org/10.1109/CICSYN.2013.52`
5. Chaudhari, K., Thakkar, A.: A comprehensive survey on travel recommender systems. Archives of Computational Methods in Engineering pp. 1–27 (2019)
6. Das, A.S., Datar, M., Garg, A., Rajaram, S.: Google news personalization: Scalable online collaborative filtering. In: Proceedings of the 16th International Conference on World Wide Web. p. 271–280. WWW '07, Association for Computing Machinery, New York, NY, USA (2007). https://doi.org/10.1145/1242572.1242610, `https://doi.org/10.1145/1242572.1242610`
7. Eksombatchai, C., Jindal, P., Liu, J.Z., Liu, Y., Sharma, R., Sugnet, C., Ulrich, M., Leskovec, J.: Pixie: A system for recommending 3+ billion items to 200+ million users in real-time. In: Proceedings of the 2018 World Wide Web Conference. p. 1775–1784. WWW '18, International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE (2018). https://doi.org/10.1145/3178876.3186183, `https://doi.org/10.1145/3178876.3186183`
8. Jiang, K., Wang, P., Yu, N.: Contextrank: Personalized tourism recommendation by exploiting context information of geotagged web photos. In: Proceedings of the 2011 Sixth International Conference on Image and Graphics. p. 931–937. ICIG '11, IEEE Computer Society, USA (2011). https://doi.org/10.1109/ICIG.2011.48, `https://doi.org/10.1109/ICIG.2011.48`
9. Kbaier, M.E.B.H., Masri, H., Krichen, S.: A personalized hybrid tourism recommender system. In: 2017 IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA). pp. 244–250 (2017). https://doi.org/10.1109/AICCSA.2017.12

10. Kofler, C., Caballero, L., Menendez, M., Occhialini, V., Larson, M.: Near2me: An authentic and personalized social media-based recommender for travel destinations. In: Proceedings of the 3rd ACM SIGMM International Workshop on Social Media. p. 47–52. WSM '11, Association for Computing Machinery, New York, NY, USA (2011). https://doi.org/10.1145/2072609.2072624, https://doi.org/10.1145/2072609.2072624
11. Kurashima, T., Iwata, T., Irie, G., Fujimura, K.: Travel route recommendation using geotags in photo sharing sites. In: Proceedings of the 19th ACM International Conference on Information and Knowledge Management. p. 579–588. CIKM '10, Association for Computing Machinery, New York, NY, USA (2010). https://doi.org/10.1145/1871437.1871513, https://doi.org/10.1145/1871437.1871513
12. Linden, G., Smith, B., York, J.: Amazon.com recommendations: Item-to-item collaborative filtering. IEEE Internet Computing **7**(1), 76–80 (Jan 2003). https://doi.org/10.1109/MIC.2003.1167344, https://doi.org/10.1109/MIC.2003.1167344
13. Lucas, J.P., Luz, N., Moreno, M.N., Anacleto, R., Almeida Figueiredo, A., Martins, C.: A hybrid recommendation approach for a tourism system. Expert Syst. Appl. **40**(9), 3532–3550 (Jul 2013). https://doi.org/10.1016/j.eswa.2012.12.061, https://doi.org/10.1016/j.eswa.2012.12.061
14. Lucchese, C., Perego, R., Silvestri, F., Vahabi, H., Venturini, R.: How random walks can help tourism. In: Proceedings of the 34th European Conference on Advances in Information Retrieval. p. 195–206. ECIR'12, Springer-Verlag, Berlin, Heidelberg (2012). https://doi.org/10.1007/978-3-642-28997-2_17, https://doi.org/10.1007/978-3-642-28997-2_17
15. Meehan, K., Lunney, T., Curran, K., McCaughey, A.: Context-aware intelligent recommendation system for tourism. In: 2013 IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops). pp. 328–331 (2013). https://doi.org/10.1109/PerComW.2013.6529508
16. Menk, A., Sebastia, L., Ferreira, R.: Recommendation systems for tourism based on social networks: A survey. arXiv preprint arXiv:1903.12099 (2019)
17. Pazzani, M., Billsus, D.: Learning and revising user profiles: The identification ofinteresting web sites. Mach. Learn. **27**(3), 313–331 (Jun 1997). https://doi.org/10.1023/A:1007369909943, https://doi.org/10.1023/A:1007369909943
18. Pazzani, M.J.: A framework for collaborative, content-based and demographic filtering. Artif. Intell. Rev. **13**(5–6), 393–408 (Dec 1999). https://doi.org/10.1023/A:1006544522159, https://doi.org/10.1023/A:1006544522159
19. Pu, Z., Du, H., Yu, S., Feng, D.: Improved tourism recommendation system. In: Proceedings of the 2020 12th International Conference on Machine Learning and Computing. p. 121–126. ICMLC 2020, Association for Computing Machinery, New York, NY, USA (2020). https://doi.org/10.1145/3383972.3384074, https://doi.org/10.1145/3383972.3384074
20. Ricci, F., Rokach, L., Shapira, B.: Recommender Systems Handbook. Springer Publishing Company, Incorporated, 2nd edn. (2015)
21. Schafer, J.B., Frankowski, D., Herlocker, J., Sen, S.: Collaborative Filtering Recommender Systems, p. 291–324. Springer-Verlag, Berlin, Heidelberg (2007)
22. Shen, J., Deng, C., Gao, X.: Attraction recommendation. Neurocomput. **173**(P3), 789–798 (Jan 2016). https://doi.org/10.1016/j.neucom.2015.08.030, https://doi.org/10.1016/j.neucom.2015.08.030

23. Trouillon, T., Welbl, J., Riedel, S., Gaussier, É., Bouchard, G.: Complex embeddings for simple link prediction. In: International Conference on Machine Learning. pp. 2071–2080. PMLR (2016)