Choice Models in Tourism Recommender Systems

Almomani, A., Saavedra, P., Barreiro, P., Durán, R. and Sánchez E.
CITIUS, University of Santiago de Compostela
Santiago de Compostela, Spain

Crujeiras R.
School of Mathematics, University of Santiago de Compostela
Santiago de Compostela, Spain

Loureiro M
School of Business, University of Santiago de Compostela
Santiago de Compostela, Spain

***Abstract*—Choice models (CM) are proposed in the field of tourism recommender systems (TRS) with the aim of providing algorithms with both a theoretical understanding of tourist's motivations and a certain degree of transparency. The goal of this work is to overcome some of the limitations of current state-of-art algorithms used in TRSs by providing: (1) accurate preferences, which are learnt from user choices rather than from ratings, and (2) interpretable coefficients, which are achieved by means of the set of estimated parameters of CM. The study was carried out with a gastronomic data set generated in an ecological experiment in the tourism domain. The performance of CM has been compared with a set of baseline algorithms (rating-based and ensembles) by using two evaluation metrics: precision and DCG. The CM outperformed the baseline algorithms when the size of the choice set was limited. The findings suggest that CM may provide an optimal trade-off between theoretical soundness, interpretability and performance in the field of TRS.**

***Keywords*— Knowledge engineering, Artificial intelligence, Recommender Systems, Choice models, Ensembles and Tourism.**

I. Introduction

The main drawbacks of the techniques applied so far in the development of TRSs are: the lack of a theoretical background to understand the underlying motivational factors conditioning the tourist decision-making, and their interpretability to explain the recommendations. Our work focuses on providing a theoretical background to the algorithms behind tourism recommender systems to alleviate these problems. The contribution to the field of TRSs can be summarized as follows: 1. the application of choice models (CM) as a tool to learn tourist's preferences with a sound methodology, and 2. the exploration of the potential of CM by comparing them with algorithms used in TRSs: (1) advanced rating-based algorithms and (2) ensemble strategies.

II. Methods

The methods were chosen to compare the performance of CM against rating-based models and popular ensemble strategies. The analysis was carried out with a gastronomic data set generated in an ecological experiment in the tourism domain. The details about the design, dataset, CM and baseline algorithms can be found in [1].

A. Dataset

The data collected in the experiment were used to build two datasets: the choice and the rating dataset. The choice dataset consisted on choice observations, where each observation included the vectors containing the attribute values of the chosen *tapa* as well as the tapas of the choice set. To describe a *tapa*, the following attributes were considered: "type" and "character". Traditional *tapas* were created following well-known, popular recipes, while daring *tapas* were new and creative. In terms of data preparation, "type" and "character" attributes were transformed into eight dichotomous or binary variables associated with each value. The choice dataset codified this way was used to fit the choice-based models.

B. Choice models

The standard logit model as well as the mixed logit model, assuming Gaussian distribution on the coefficients, were chosen as basic representatives of the family of random utility choice-based models. The CM estimate the probability P_{ni} as the ratio between the relevancy of the item a_i for user c_n, estimated by the $e^{V_{ni}}$ term, and the aggregated relevancy of all items a_j in the choice set. This set is the collection of items that the user considers/analyzes at the time of choice. The values of the probability P_{ni} depend on the representative utilities, V_{ni}. As V_{ni} increases, reflecting a higher match between the observed attributes of the alternative and the preferences of the decision-maker, P_{ni} approaches the value one. The representative utility is usually specified as linear in the set of alternative attributes: $V_{nj} = \beta_{nj} \cdot x_j$, where x_j is a vector including the observed attribute's values of the alternative a_j, and β_{nj} denotes the model coefficients vector describing the preferences of decision-maker c_n for the attributes of the alternatives a_j. The preferences β_{nj} (model coefficients) are estimated by fitting the CM to a data set of choices.

C. Baselines

The following types of baseline models were chosen: (1) basic rating-based collaborative filtering algorithms, (2) advanced rating-based collaborative algorithm, (3) single decision-trees, and (4) tree-based ensemble strategies. The choice of tree-based from among other types of ensembles is explained by the fact that trees produce meaningful predictions, and they thus become a natural alternative to choice-based models to overcome the interpretability problem.

Partially supported by Spanish Thematic Network on Recommender Systems (Action RED2022-134302-T funded by MCIN/AEI/10.13039/501100011033)

D. Performance evaluation

Two metrics were applied in order to evaluate the performance of choice-based and rating-based algorithms: “Precision” and “Discounted Cumulative Gain” (DCG). For each *tapa's* item included in the choice set, either its rating or its choice probability was predicted. Thereafter, the *tapas* were ranked and only the item with highest value was considered the predicted choice and therefore recommended (Top-1 scheme).

III. RESULTS

A. Data description

The choice dataset characterized includes the choices of 5517 individuals regarding a set of 113 *tapas* available during the Santiago(e)Tapas contest. There were three choice sets corresponding to the three locations of restaurants in the city. In the new area of the city 2030 users consumed 3888 *tapas* that were chosen from 37 alternatives: 18 of traditional character, and 19 of daring character. As for the old area, 3953 participants tasted 8948 *tapas* chosen from the set of 62 available tapas: 32 of traditional character, and 30 of daring character. Finally, in the outlying area of the city, the least popular area, 436 users consumed 743 *tapas* from 14 available choices: 3 of traditional nature and 11 of daring nature. Further details about the data description can be found in [1].

B. Fitting of choice models

The CM, both the standard and mixed logit models, were fitted to the data for the three choice problems described before. For the mixed logit model, a Gaussian distribution of the coefficients was assumed, and the number of draws, R, was set to 100. Details about the fitting can be found in [1].

Most of the coefficients proved significant. In terms of preferences, the sign of coefficients represent the positive or negative preference of users for the *tapas'* attribute. For instance, participants revealed a positive preference for egg, meat and shellfish *tapas* in the old area, but a negative one for egg and traditional *tapas* in the new area.

C. Performance evaluation

In terms of “Precision”, the results show that CM perform slightly better than both CF and Ensemble algorithms, but quite similarly to the Single-Tree approach. However, in most cases, Precision is zero or close to zero, indicating that the predicted *tapas* item does not usually correspond to the one actually chosen. DCG, in turn, is more informative for analyzing and comparing the performance of the different models. CM showed a superior performance in the Outlying area of the city, but Single-Tree and Ensemble algorithms provided better results in the other two areas.

The ranking of the methods in terms of DCG show that Choice-models were ranked first in the outlying area (with only 14 alternative *tapas*) but the comparative performance was lower in the other two areas, in which the choice sets are larger. Ensembles show a somewhat opposite behaviour: they performed comparatively better when more choices, users and observations are available. Surprisingly, the Single-Tree method, including only one learner, provided competitive results in the outlying and new areas. As expected, however, for larger data sets (old part of the city) the performance was not as good as that of ensembles and CM. On the other hand, CF approaches are far from being competitive and occupied the lowest position, except for the outlying area of the city. Another interesting finding is that the performance of all models in terms of DCG was lower as long as the choice set, i.e. the number of available *tapas*, increased from the outlying area to the old area. This may suggest that the prediction problem is probably more complex when the choice set increases and the choice becomes more difficult for the decision-maker.

Further details about the obtained results can be found in [1].

IV. DISCUSSION

CM offer a theoretical background to generate sound recommendations. The robustness of the approach is achieved by means of estimating preferences from a reliable source. In CM, preferences are learnt by fitting choice-based models with choice data. The underlying assumption is that user's choices are the result of the direct matching between the user's preferences and the item's attributes. So, by observing the choices and gathering the attribute's values of the alternatives in the choice sets, the unknown preferences can be discovered. On the other hand, rating-based models rely on user's ratings, which represent a post-experience satisfaction. This outcome mainly depends on the comparison between the real and the expected satisfaction. Preferences and tastes are thus out of this equation, meaning that ratings and preferences have no clear relationship.

Transparency and interpretability are another contribution of choice models. The estimated preferences, i.e. model coefficients, provide a means of easily explaining why some items are more likely to be recommended than others. Moreover, this benefit does not mean a decrease in performance. The results suggest that choice models may be inferior to ensembles only in situations where larger data sets (old area) are available. The fact that ensembles show a superior performance in the case of larger datasets was expected as the N predictors of the ensemble could take advantage of sampling the data in a more efficient way.

REFERENCES

[1] Almomani, A., Saavedra, P., Barreiro, P., Durán, R., Crujeiras, R., Loureiro, M. and Sánchez, E., 2023. Application of choice models in tourism recommender systems. Expert Systems, 40(3), p.e13177.