SerendipiDater: Unexpected Connections in Dating (Group 12)

GIOVANNI FINCATO DE LOUREIRO, 4926854, Delft University of Technology, The Netherlands
SMRUTI KSHIRSAGAR, 5938643, Delft University of Technology, The Netherlands
ESHA DUTTA, 5915996, Delft University of Technology, The Netherlands
ZIXUAN HAN, 5916135, Delft University of Technology, The Netherlands

The landscape of online dating continues to evolve rapidly, with millions worldwide relying on dating platforms to connect with potential romantic partners. Users' actions in online dating often deviate from their explicitly stated preferences due to implicit preferences. In this paper, we have integrated serendipity into a collaborative filtering algorithm to enhance the diversity and utility of dating recommendations. Then, we evaluated the quality of recommendations against serendipity using various metrics on offline data. Our results for this data showed that adding serendipity to the recommendation algorithm leads to a decrease in the ranking and relevance quality of profile recommendations.

CCS Concepts: • **Information systems → Recommender systems**.

Additional Key Words and Phrases: Serendipity, Recommender Systems, Online Dating

ACM Reference Format:
Giovanni Fincato de Loureiro, Smruti Kshirsagar, Esha Dutta, and Zixuan Han. 2024. SerendipiDater: Unexpected Connections in Dating (Group 12). In *IN4325-Q3-24: Information Retrieval,TU Delft.* , 9 pages.

1 INTRODUCTION

According to a recent survey, 366 million people use online dating apps like Tinder and Bumble worldwide[18]. Moreover, one in ten partnered adults – meaning those who are married, living with a partner or in a committed romantic relationship – met their current significant other through a dating site or app[14]. Recommendation systems can match users based on explicit criteria, like the level of education or religious and political beliefs, that the user has provided. However, users' actions are often contrary to their stated preferences due to implicit preferences[15]. For example, "attractiveness" is an implicit preference which is difficult to quantify as it encompasses many subjective attributes that are not always directly stated in a user profile or preferences. Thus, recommending profiles based on users' explicit criteria may not be an effective way to provide good matches.

In this paper, we propose to integrate serendipity into a collaborative filtering recommendation algorithm to expand the recommendations generated to a more diverse yet useful set of profiles. The main research question of this paper is "What impact does adding serendipity to a collaborative filtering dating recommendation system have on the quality of the recommendations with respect to their relevance and ranking?"

The experiment has been performed on a dataset from the popular dating app OkCupid. We have created two models, one without serendipity and one with serendipity. We have employed the Root Mean Squared Error (RMSE), Precision, and Normalized Discounted Cumulative Gain (nDCG) as a measure of recommendation quality. Serendipity is also measured through a metric called SRDP that measures unexpectedness and usefulness[4]. A link to the GitHub repository for this project can be found here. The results of the experiment show that adding serendipity to the algorithm leads to a higher SRDP value but a decrease in Precision and nDCG.

The remainder of this paper is structured in the following way: the Related Work section provides a systematic review of literature pertaining to dating recommendation systems and serendipity. The Methodology section gives insights into the proposed solution and its implementation. The following Experiments section delves into results, discussion, limitations, and possible future work. Finally, the Conclusion summarises the results and implications of this experiment.

2 RELATED WORK

Online dating system recommendation systems have been extensively studied, with a broad range of methodologies. Earlier efforts focused on leveraging the information of user profiles and preference settings to make matches based on shared interests and important demographic factors[3]. These systems employed algorithms designed to maximise the compatibility of users based on explicit criteria, emphasising the importance of aligning the key personal and lifestyle preferences of the users. Despite the sophistication of the algorithms employed, the research highlighted their limitations as well. Namely, algorithms do not account for the most nuanced aspects of attraction and interpersonal chemistry; these aspects are not possible to explicitly state in user profiles[19]. This critique underscores the need for a dynamic recommendation approach that can better capture the complexity of human interaction.

2.1 Collaborative Filtering Recommendation Systems

A Content-Based reciprocal recommender was discussed in [16], where compatibility scores were calculated based on user preferences and activity, along with basic similarities such as gender, age, and body type. However, profile matching is subjective, and users may not always be looking for people who are similar to them. The authors of [9] addressed this by using Collaborative Filtering (CF) based on interactions between users. They concluded that CF worked well in generating suitable candidates and also provided a substantial improvement in success rate. The studies in [10] and [21] produced similar results, supporting the usage of CF for person-to-person recommendations. Different algorithms for finding similarities were discussed in [2]. User-user similarity was assessed by utilizing users with similar rating vectors (referred to as neighbors) to predict ratings for a particular user. To find the neighbors of a node, K-Nearest Neighbors (KNN) can be used [1]. The choice of K defines the locality of KNN. The similarity and distance measures also play a role in finding the appropriate set of neighbors. The rating patterns of users close to the target user should contribute more to the prediction than patterns that are further away. Thus, the concept of distance-weighted KNN [8] can be used to factor in this degree of similarity while predicting the ratings.

2.2 Serendipity in Recommendation Systems

The definition of serendipity is ambiguous and a major challenge which impacts evaluation methods. For instance, [17] considered serendipity as a measure of how surprising and successful the recommendations were. Another accepted definition of serendipity is the measure of the extent to which the recommended items are both attractive and surprising to the users[5]. One of the first comprehensive studies of serendipity usage in recommendation systems outlined a framework and potential benefits of incorporating recommendations that were unexpected, but pleasing to the users[7]. The work suggested that the algorithm introducing users to matches they would not normally consider could enhance user satisfaction.

There are a wide variety of evaluation methods for serendipity since there are different definitions of the word. Depending on the definition, various components contribute to serendipity. For example, if serendipity is defined as a measure of unexpectedness, a metric has been defined based on the idea that unexpectedness is low for easy-to-predict items and high for difficult-to-predict items[12].

Research concerning serendipity in online dating recommendation systems specifically is limited. However, studies from other domains provide valuable insight into the potential benefits. For example, there has been an investigation into the impact of serendipity in recommendation systems in the context of movie recommendations[13]. Another study extended this exploration into music recommendation systems and demonstrated that serendipity enhances user engagement and improves satisfaction by introducing users with a sense of variety into the recommendation process[23]. A study proposed a framework for improving recommendation systems[20]. The framework argued that metrics like serendipity should be used as key metrics for evaluating recommendation systems. This can be used to guide the development of dating recommender algorithms. These findings support the idea that users highly value unexpected recommendations if they are relevant. This principle can also be adapted to online dating. A 2017 study conducted reviews of user satisfaction in online dating, which identified factors that make the user experience more positive[22]. Despite not specifically mentioning serendipity, it showed the potential for pleasant surprises to increase engagement. This finding supports the idea that serendipity can play a significant role in satisfaction for users.

3 METHODOLOGY

3.1 Data and Compatibility Scores Calculation

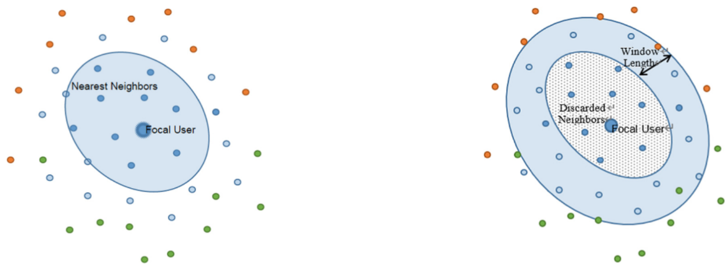
We have used OkCupid profile data from Kaggle[6] which comprises user profile data from approximately 59,946 users residing within a 25-mile radius of San Francisco during the 2010s. The data is anonymous and includes information such as age, gender, religious beliefs, certain habits and answers to questions posed by the website (essays). We scaled the data down to the first 5,000 entries of user profiles to reduce computing time.

Although the data is well-formatted and requires little pre-processing, it does not contain any column that can be explicitly used to predict potential compatibility between pairs of users. We took inspiration from a GitHub repository in which profile features are used to determine compatibility scores. Approaches include matching gender and orientation, and cosine scores of pairs of essays. The features are then weighted based on relevance using logistic regression, and a compatibility score is calculated for each user with every other user in the dataset. For our experiment, only the top 50 most compatible users are kept for each user. This provides a ground truth for matchmaking and evaluation. These ratings are bilateral and thus mutually common between two particular users.

3.2 Algorithm for Prediction of Ratings

A user-based collaborative filtering (CF) approach is employed to generate profile recommendations of users for other users. The K-Nearest Neighbour (KNN) algorithm is used to facilitate collaborative filtering by finding k users closest to the target user. The compatibility scores of these neighbors with their potential matches are used to calculate ratings for the target user with those same potential matches. These ratings are the neighbors' aggregated ratings weighted by the inverse of the distance of those neighbors from the target user. These ratings are sorted in descending order and the top ones are selected as potential matches for the target user.

To add serendipity to this algorithm, the K-Nearest Neighbor algorithm is replaced by the K-Surrounding Neighbors algorithm which is based on the theory of weak ties[11]. The proposed method discards some nearest neighbors based



(a) Traditional K-Nearest Neighbors (b) K-Surrounding Neighbors

Fig. 1

on their similarity, giving more weight to less-similar others. This can provide fresh information and new experiences. This concept is illustrated in Figure 1. The pseudocode for this algorithm can be found in Algorithm 1.

Input: user_item_matrix, target_user_index, X, n_neighbors, n_recommendations
Output: predicted_ratings
$distances, indices \leftarrow$ n_neighbors + X nearest neighbors and their distances for target_user_index
$relevant\_indices, relevant\_distances \leftarrow$ Exclude X nearest neighbours
for $item\_index$ ***in*** *all users* **do**
  $total\_rating \leftarrow 0$;
  $total\_weight \leftarrow 0$;
  for $neighbor\_index$ ***in*** $relevant\_indices$ **do**
    if $neighbor\_index$ *has rated* $item\_index$ **then**
      $weight \leftarrow 1/distance\_of\_neighbor\_from\_target\_user$
      $total\_rating \leftarrow total\_rating + neighbor\_rating * weight$
      $total\_weight \leftarrow total\_weight + weight$
    end
  end
  if $total\_weight > 0$ **then**
    Add $(item\_index : total\_rating/total\_weight)$ to $list\_of\_recommendations$
  end
end
$top\_recommendations \leftarrow$ top $n\_recommendations$ elements in $list\_of\_recommendations$ sorted by value in descending order
return $top\_recommendations$

Algorithm 1: predict_ratings

3.3 Metrics

The evaluation is done using the computed compatibility scores discussed in section 3.1 as the ground truth. The top 50 matches along with the compatibility scores are computed for each user ($N = 5000$). The top ten of these are used as user-item ratings for the collaborative filtering algorithm. The remaining forty matches and scores are considered to be relevant results and relevance scores respectively. The algorithm generates twenty recommendations ($r = 20$), i.e, predicted matches, along with the predicted user-item ratings, i.e., the predicted compatibility scores. In case a match recommended by the algorithm is not present in the top 50 matches for a particular user, it is considered to be an irrelevant result with a relevance score of 0.

For the evaluation of the predicted scores, Root Mean Squared Error(RMSE) is used. Mean is taken over compatibility scores (y) for each match of each user. It is calculated as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} \sum_{j=1}^{r} (y_{ij} - \hat{y}_{ij})^2}{N \times r}} \tag{1}$$

The recommendation algorithm is evaluated using Precision at K (P@K) and Normalised Discounted Cumulative Gain at K (nDCG@K). These metrics are chosen because Precision factors in the number of relevant matches while nDCG measures the quality of the rankings. These factors represent the quality of recommendations. Calculation of nDCG is done using ground truth relevance scores. P@k and nDCG@k are both calculated for k = 5, 10, 20. The following formulae are used:

$$\text{Average P@k} = \frac{1}{N} \times \sum_{i=1}^{N} \frac{|\text{relevant\_matches}_k \cap \text{recommended\_matches}_k|}{\text{k}} \tag{2}$$

$$\text{Average nDCG@k} = \frac{1}{N} \sum_{i=1}^{N} \frac{DCG_i@k}{IDCG_i@k} \tag{3}$$

The serendipity metric used to evaluate both models (with and without serendipity) is two-fold[4]. In the first step, the unexpected set of recommendations is calculated as follows:

$$UNEXP = RS/PM \tag{4}$$

Here, PM is a set of recommendations generated by a primitive prediction model and RS denote the recommendations generated by our recommendation system. This metric requires a baseline primitive model, the recommendations from which can be compared to both of our CF models to find UNEXP. A popularity-based recommendation system was implemented for this purpose. This model recommends users which have the highest compatibility score in the ground truth. Then, the serendipity is calculated as:

$$SRDP = \frac{\sum_{i=1}^{N} u(RS_i)}{N} \tag{5}$$

Here, RS_i is defined as an element in $UNEXP$. When $u(RS_i) = 1$, RS_i is considered to be useful and when $u(RS_i) = 0$, it means RS_i is useless to the user. N is the total number of elements in $UNEXP$. The usefulness of RS_i should ideally be judged by the user. However, we have set threshold values of 0.356 and 0.350 for the models with and without serendipity respectively. We consider that if the aggregate rating is above these values, the user finds that recommendation useful. These values have been set as the median values of all aggregate ratings. Figure 2 illustrates boxplots for the aggregate ratings for recommendations generated for all users. This metric was chosen because it takes into consideration the usefulness as well as the unexpectedness of the recommendations. Similar to the other metrics, SRDP@K was calculated as the SRDP value for the first K recommendations.

4 EXPERIMENTS

The metrics mentioned in section 3.3 were used to evaluate both models. The results can be found in Table 2 where Model 1 corresponds to the algorithm without serendipity and Model 2 is the one with serendipity. To optimise k for k-NN, RMSE was calculated for the predicted ratings for k values 5, 10, 20 and 25. The findings are shown in Table 1. The values show that RMSE increases with increasing number of neighbours. However, for k=5 and k=10, the algorithm is unable to generate a sufficient number of recommendations for certain users. For k=20, the algorithm generates 20 recommendations for 99% and 96% of users for models with and without serendipity respectively, and at least 10 recommendations for all users in both models. Thus, we use 20 neighbours for k-NN.

Model	**RMSE**			
	k=5	**k=10**	**k=20**	**k=25**
Model 1	0.221	0.253	0.265	0.266
Model 2	0.273	0.273	0.276	0.277

Table 1. RMSE values for different values of k

4.1 Results

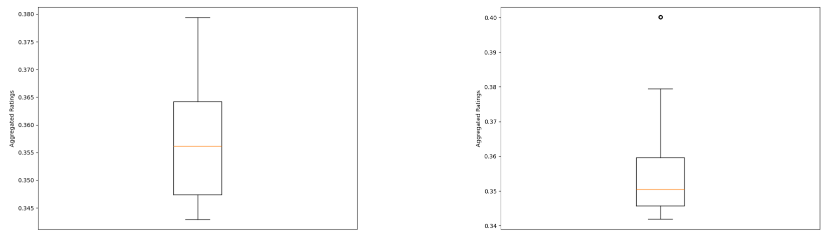
In the model with serendipity, the 5 nearest neighbors are excluded. This model only has a slightly higher SRDP than the one without serendipity. Also, the model without serendipity consistently performs better with respect to both precision and ranking. Thus, there is a clear trade-off between serendipity against precision and ranking.

Moreover, the precision and nDCG values are quite low in the model with serendipity. Also, the SRDP values in this model are only slightly higher compared to the model without serendipity. This low difference may be because the compatibility scores, which constitute the ground truth, are calculated based only on explicit preferences. They are not indicative of users' implicit preferences. Integrating real-world user-interaction data that sheds more light into users' actual match preferences can lead to establishment of a more nuanced ground truth. This will provide more representative data to quantify relevance of the recommendations, which can influence the serendipity model and potentially lead to better performance with respect to all metrics.

Despite decreased precision, the addition of serendipity may encourage users to explore a wider range of potential matches beyond their explicit preferences. This can lead to increased diversity in the profiles users interact with, fostering a more inclusive dating community which provides more beneficial matches to users. However, overly random recommendations may lead to frustrated users, while overly tailored recommendations may limit chance discoveries. Thus, there should be a balance between providing serendipitous recommendations and ensuring relevance to the user's preferences.

		Metrics@5			**Metrics@10**			**Metrics@20**		
Model	**RMSE**	**P**	**nDCG**	**SRDP**	**P**	**nDCG**	**SRDP**	**P**	**nDCG**	**SRDP**
Model 1	0.265	0.581	0.765	0.455	0.524	0.771	0.501	0.415	0.752	0.557
Model 2	0.276	0.503	0.694	0.471	0.454	0.704	0.525	0.370	0.692	0.581

Table 2. Metrics of both models



(a) Without serendipity

(b) With serendipity

Fig. 2. Boxplot for aggregated ratings of both model

4.2 Limitations

This paper does not take any user feedback into account. It only relies on the aforementioned metrics to gauge the quality of recommendations. However, user feedback can help find a correlation between user satisfaction and serendipity in dating apps. Moreover, the algorithm uses an existing static dataset and has not been executed in a live app scenario. Thus, it is difficult to ascertain the performance in the real world, where diverse user behaviours lead to outcomes that cannot be anticipated by simulations.

The methodology used for establishing ground truth static compatibility scores is based on explicit user attributes which oversimplifies the complex nature of relationships. This introduces difficulty in evaluating the algorithm. This may also be the reason behind reduced precision and serendipity values for the models. A more sophisticated method would account for more nuances, like messaging data and profile visits, displayed by users while searching for a partner.

The bidirectional interaction process of dating apps or mutual interest is not emulated in the recommendation system. This simplification compromises the relevance of the model in applying it to actual dating app scenarios, where the efficacy of the recommendation depends on reciprocal user engagement.

Furthermore, the serendipity metric used to evaluate the algorithm depends on users deciding whether a serendipitous recommendation is useful or not. Currently, this has been simulated with a threshold. In an ideal scenario, the opinions of users should be considered.

5 CONCLUSION

Adding serendipity to dating recommendation systems could give more surprising and unexpected matches to users. Our investigation aimed to compare two collaborative filtering dating recommendation algorithms, one with serendipity and one without it, to evaluate the effect of adding serendipity on the relevance and ranking of the recommendations. The findings reveal that including serendipity reduces the precision and ranking quality of the algorithm. However, it also suggests that the potential for increasing the diversity of profiles recommended to users can enrich the users' overall experience on dating platforms. Replicating this experiment on real-world data and receiving and incorporating user feedback would provide more valuable insights into the effect of adding serendipity on the quality of recommendations.

REFERENCES

[1] Taushif Anwar, V. Uma, Md. Imran Hussain, and Muralidhar Pantula. 2022. Collaborative filtering and kNN based recommendation to overcome cold start and sparsity issues: A comparative analysis. *Multimedia Tools and Applications* 81, 25 (March 2022), 35693–35711. https://doi.org/10.1007/s11042-

021-11883-z

[2] Lukas Brozovsky and Vaclav Petricek. 2007. Recommender System for Online Dating Service. arXiv:cs/0703042 [cs.IR]

[3] Fernando Diaz, Donald Metzler, and Sihem Amer-Yahia. 2010. Relevance and Ranking in Online Dating Systems. *Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval* (2010).

[4] Mouzhi Ge, Carla Delgado-Battenfeld, and Dietmar Jannach. 2010. Beyond accuracy: evaluating recommender systems by coverage and serendipity. In *Proceedings of the Fourth ACM Conference on Recommender Systems* (Barcelona, Spain) *(RecSys '10)*. Association for Computing Machinery, New York, NY, USA, 257–260. https://doi.org/10.1145/1864708.1864761

[5] Jonathan L. Herlocker, Joseph A. Konstan, Loren G. Terveen, and John T. Riedl. 2004. Evaluating collaborative filtering recommender systems. *ACM Trans. Inf. Syst.* 22, 1 (jan 2004), 5–53. https://doi.org/10.1145/963770.963772

[6] Albert Kim and Adriana Escobedo-Land. 2015. OkCupid Data for Introductory Statistics and Data Science Courses. *Journal of Statistics Education* 23 (07 2015). https://doi.org/10.1080/10691898.2015.11889737

[7] Denis Kotkov, Shuaiqiang Wang, and Jari Veijalainen. 2016. A survey of serendipity in recommender systems. In *Knowledge-Based Systems*, Vol. 111. 180–192.

[8] Oliver Kramer. 2013. *K-Nearest Neighbors.* Springer Berlin Heidelberg, Berlin, Heidelberg, 13–23. https://doi.org/10.1007/978-3-642-38652-7_2

[9] Alfred Krzywicki, Wayne Wobcke, Xiongcai Cai, Ashesh Mahidadia, Michael Bain, Paul Compton, and Yang Sok Kim. 2010. Interaction-Based Collaborative Filtering Methods for Recommendation in Online Dating. In *Web Information Systems Engineering – WISE 2010*, Lei Chen, Peter Triantafillou, and Torsten Suel (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 342–356.

[10] A. Krzywicki, W. Wobcke, Y.S. Kim, X. Cai, M. Bain, A. Mahidadia, and P. Compton. 2015. Collaborative Filtering for people-to-people recommendation in online dating: Data analysis and user trial. *International Journal of Human-Computer Studies* 76 (2015), 50–66. https://doi.org/10.1016/j.ijhcs.2014.12.003

[11] Huidi Lu. 2023. k-Surrounding Neighbors: Incorporating Serendipity in Collaborative Recommendations. In *Proceedings of the 2023 2nd International Conference on Artificial Intelligence, Internet and Digital Economy (ICAID 2023)*. Atlantis Press, 115–122. https://doi.org/10.2991/978-94-6463-222-4_11

[12] Tomoko Murakami, Koichiro Mori, and Ryohei Orihara. 2008. Metrics for evaluating the serendipity of recommendation lists. In *New Frontiers in Artificial Intelligence: JSAI 2007 Conference and Workshops, Miyazaki, Japan, June 18-22, 2007, Revised Selected Papers*. Springer, 40–46.

[13] Tien T Nguyen, F Maxwell Harper, Loren Terveen, and Joseph A Konstan. 2018. User Personality and User Satisfaction with Recommender Systems. *Information Systems Frontiers* 20 (2018), 1173–1189. Published: 03 September 2017, Accesses: 4852, Citations: 60, Altmetric: 1, Mention: 1.

[14] Pew Research Center. 2023. *Key Findings About Online Dating in the U.S.* https://www.pewresearch.org/short-reads/2023/02/02/key-findings-about-online-dating-in-the-u-s/

[15] Luiz Pizzato, Thomas Chung, Tomasz Rej, Irena Koprinska, Kalina Yacef, and Judy Kay. 2010. Learning user preferences in online dating. In *Proceedings of the Preference Learning (PL-10) Tutorial and Workshop, European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD)*. Citeseer.

[16] Luiz Pizzato, Tomek Rej, Thomas Chung, Irena Koprinska, and Judy Kay. 2010. RECON: a reciprocal recommender for online dating. In *Proceedings of the Fourth ACM Conference on Recommender Systems* (Barcelona, Spain) *(RecSys '10)*. Association for Computing Machinery, New York, NY, USA, 207–214. https://doi.org/10.1145/1864708.1864747

[17] Guy Shani and Asela Gunawardana. 2011. Evaluating recommendation systems. *Recommender systems handbook* (2011), 257–297.

[18] Statista. Year of Access. Online Dating - Statistics Facts. https://www.statista.com/topics/7443/online-dating/. Accessed: April 4, 2024.

[19] Stephanie T Tong, Jeffrey T Hancock, and Richard B Slatcher. 2016. Online dating system design and relational decision making: Choice, algorithms, and control. *Personal Relationships* 23, 4 (2016). https://doi.org/10.1111/pere.12158 Citations: 21, This work was supported by NSF Grant 1520723. Portions of this paper were presented at the 2016 Human Computer Interaction International Conference in Toronto, ON..

[20] Guido van Capelleveen, Chintan Amrit, Devrim Murat Yazan, and Henk Zijm. 2019. The recommender canvas: A model for developing and documenting recommender system design. *Expert Systems With Applications* 129 (2019), 97–117. https://doi.org/10.1016/j.eswa.2019.04.001 UvA-DARE (Digital Academic Repository), Final published version, License: Article 25fa Dutch Copyright Act.

[21] Peng Xia, Benyuan Liu, Yizhou Sun, and Cindy Chen. 2015. Reciprocal Recommendation System for Online Dating. In *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015* (Paris, France) *(ASONAM '15)*. Association for Computing Machinery, New York, NY, USA, 234–241. https://doi.org/10.1145/2808797.2809282

[22] Peng Xia, Bruno Ribeiro, Benyuan Liu, and Cindy Chen. 2013. A study of user behavior on an online dating site. In *Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*. IEEE/ACM. https://doi.org/10.1145/2492517.2492659

[23] Yuan Cao Zhang, Daniele Quercia, Diarmuid Ó Séaghdha, and Tamas Jambor. 2012. Auralist: Introducing serendipity into music recommendation. In *Proceedings of the Fifth International Conference on Web Search and Web Data Mining, WSDM 2012*. ACM, Seattle, WA, USA. https://doi.org/10.1145/2124295.2124300 Source: DBLP.

A SELF-ASSESSMENT OF CONTRIBUTIONS

- Esha: Project proposal, identifying method to add serendipity to algorithm, implementation of both models of prediction algorithm, finding a serendipity metric and implementing it, report writing

- Giovanni: Related work on serendipity in dating apps, evaluation of applicability, impact and limitations of research, data pre-processing, report writing
- Smruti: Usage of collaborative filtering for recommendation, analysis for selection of k in KNN, selection of metrics for evaluating recommendations and implementation of the same, report writing
- Zixuan: Data pre-processing, calculation of ground truth compatibility score and limitations of ground truth, report writing

Received 5 April 2024