Love in Time of Algorithms:

Ethics & Data in Digital Dating Landscape

April 23, 2024

Group 2

Brandon Lewis, Cassie Ren, Ko Choi, Teagan Milford

1. Introduction

1.1. Project Overview

Whether you are looking for the love of your life or casual dating, there is a good chance that you have come across a dating platform. Regardless of your objective, finding the right person often takes time and effort.

This project aims to analyze potential risks that come with using machine learning in dating platforms. While machine learning might try to match two ideal candidates, its algorithms could also result in biases, missed matches, and discrimination.

To investigate further, we examine a dataset from speed dating events from 2002 to 2004, an ancient time before the rise of mobile dating apps. In this dataset, participants rated their date after a four-minute date on attributes such as attractiveness, sincerity, intelligence, fun, ambition, and shared interests. They were also asked if they would like to see their date again. Additionally, the dataset contains participants' demographic information as well as their interests.

1.2. Characterization of Use

Using machine learning algorithms for matchmaking is a foundation of modern dating, owing to the success of major dating applications like Tinder, Bumble, and Hinge. While applications use different algorithms, the basic idea is the same: match users based on their preferences and perceptions of other users.

Some applications leverage pre-existing algorithms. In the case of Hinge, matchmaking occurs by applying the Gale-Shapley algorithm[1]. This algorithm is predicated on the concept of "stable pairing", which seeks to find mutually agreeable matches with input from both sides of the potential pair[2]. Consequently, users receive the matches in order of predicted compatibility, considering information from both ends of the match.

Other applications opt to use custom algorithms. OkCupid, for example, relies on extensive information from users through an introductory survey. Its algorithm considers three main factors for each question: what you prefer your potential match to answer, the importance of your potential match's answer aligning with your expectation, and a prediction of how much your potential match will like your answer[3].

Given that these algorithms are empowered by human labeling, they are vulnerable to the biases that accompany human subjectivity. The use of an application may exacerbate indirect biases, disproportionately impacting effectiveness for groups of users. For example, a matching algorithm may reward a potential match having similar traits to the user being matched. While this sounds like a logical process for matching, it may have adverse impacts when considering which racial, sexual-oriented, and income groups are represented in the data used for algorithmic construction.

The speed dating data used for our modeling has a similar foundation with the intention of gaining information about participant preferences and perceptions. Mutually agreeable matchings are possible to predict based on the data collected, with the data collection methodology as a predecessor to modern dating applications. The data emerges from voluntary surveys by

---

[1] https://getstream.io/blog/dating-app-algorithms/
[2] https://www.geeksforgeeks.org/stable-marriage-problem/
[3] https://getstream.io/blog/dating-app-algorithms/

participants and is supplemented by demographic data that add complexity to the dataset. In our case, the target variable is “match”, which signifies whether the participant indicated that they want to see their “date” again.

2. Ethical Considerations

2.1. Sampling Bias

Our speed dating dataset has a mean age of 24 and a median age of 25. 54% of the group is aged 26 or younger, and 89% of the group is 30 or under. This overrepresentation of younger groups suggests that the data does not accurately account for the entire population of people in online dating. For example, Match.com reports the average age of people registered in their application as 36 years[4]. Certainly, other matchmaking applications have a lower mean, like Hinge at 25 years old, though this may be a reaction to non-representative algorithms.

Additionally, participants in the speed dating dataset do not reflect the populations expected compared to true population measures. According to the self-reported race factor, a little over half of the group is Caucasian, at 56%, and almost a quarter of the group is Asian, at 24%. About 8% of participants identify as Latino and Hispanic American and about 5% identified as Black or African American. If you consider the 2003 US population by race[5], White/Caucasians, Black/African American, and Latino/Hispanic are underrepresented while Asian is overrepresented. This disparity raises concerns about the generalizability of the greater population.

This may be a result of geographic, social, or demographic factors related to the Columbia Business School, which could impact the type of people who participated in the study. Similarly, education level could impact the responses of participants, making this data difficult to generalize due to differences from the US population as a whole.

Several factors could contribute to this sampling bias. Speed dating might be more popular among younger individuals and specific racial communities (White/Asian). Cultural preferences may also influence participation rates. Additionally, the geographical location of the speed dating events could be skewed if they were held in areas with higher concentrations of certain races. Sampling bias can lead to underrepresented groups’ preferences in a partner not being recognized, leading to exclusion.

2.2. Differential Subgroup Validity

Within the dataset, participants were asked to rate the importance of certain features in a potential partner, such as attractiveness, intelligence, interests, etc. The predictive power of the features may vary vastly amongst the different subgroups. For example, younger people may place a higher emphasis on physical attraction, whereas older people may prioritize shared interests and experiences. It is also possible that people looking for casual dating could prioritize physical attraction and that people looking for long-term relationships could value the compatibility of interests and hobbies more.

---

[4] https://roast.dating/blog/match-com-statistics
[5] https://getstream.io/blog/dating-app-algorithms/

If the algorithm for matching partners relies on features that have different predictive power across subgroups, it can lead to bias for certain users. With a greater number of younger participants and fewer racial minorities, the predictive power of certain features may be skewed toward the preferences of overrepresented groups, like younger Asian people. This can lead to a feedback loop where the algorithm continues to optimize the majority group's preferences, which further marginalizes the underrepresented groups.

2.3. Algorithmic Bias

Given the skewed nature of the dataset and the previously mentioned differentiating feature importance within the subgroups, algorithmic bias can creep into the model when no actions are taken. For example, if the algorithm learns that the majority of successful matches occur between younger individuals or those of certain races, it may disproportionately recommend profiles of similar demographics to users.

The consequences of this outcome go beyond online dating, as it can shape social relationships and perpetuate systemic inequalities. If certain groups consistently receive less favorable recommendations it hurts their ability to form meaningful connections in the world of dating and further, within society. Additionally, from a legal standpoint, avoiding discriminatory practices in online dating platforms ensures fairness for all users.

In our case, we practiced reweighting to balance the different age groups in the training set. By assigning higher weights to underrepresented age groups, we aimed to ensure that the model learns to reduce the majority group’s preferences. We also performed an intersectional analysis across different combinations of attributes, like age and gender. Through this, we identified disparities in the model’s predictions and addressed them through fine-tuning of the logistic regression model parameters. This ensured more balanced and equitable outcomes for all users.

3. Code and Results

3.1. Algorithmic Bias Assessment

Built into this dataset is the assumption that the participant is dating a person of the opposite sex. Race is a demographic feature surveyed, along with questions about the participant’s perception of how important having a same-race partner is. Similarly, participants were asked how important having a partner of the same religion is.

When investigating the data using feature importance, we did not find that race or religion were of high importance overall. However, this does not mean that those features did not factor into other feature perceptions indirectly.

According to our investigation of feature importance, the top 2 important features are ‘decision_o’ and ‘decision’, representing the decision that each party makes on the night of the event. This makes sense, as it would correlate directly with the match outcome decided. Following these features, the top 3 features of importance are ‘like’, ‘attractive_partner’, and ‘attractive_o.’

The ‘like’ feature represents whether the participant reported liking their partner. Again, it is easy to see the connection between a match and whether the participant enjoyed meeting with their partner. ‘Attractive_partner’ represents the participant’s rating of the attractiveness of their partner on a scale of 1-10, with 10 being the most attractive.‘Attractive_o’ represents the partner’s rating of the attractiveness of the survey taker, on a scale of 1-10, with 10 being most attractive. These are features that could potentially act as proxies for race preference, as participants may have indirect biases against or for people belonging to certain racial groups.

3.2. Analysis of Potential Harms

Potential harms associated with this analysis could include omitting commensurate input from underrepresented groups in creating future matchmaking algorithms. This could undermine the experience of minority group members who use dating applications, minimizing the impact of their preferences proportional to people with other group memberships in the data. The societal impact of this could include limiting minority group users' matches, effectively excluding them from benefiting from the matchmaking service. Additionally, users with non-minority group membership may miss out on matching with those in minority groups, perpetuating a feedback loop in the preference data.

3.3. Meaningful Metrics

To investigate algorithmic bias, we looked at accuracy, false positive rates, precision, and recall for different groups in the dataset.

All of the model versions attempted had high accuracies, with most exceeding 95% accuracy. However, this does not necessarily mean that the models are desirable for understanding algorithmic bias; rather, it suggests that the models are correctly making predictions based on potentially flawed inputs.

The recall, or true positive rate, is variable, depending on the conditions considered. The highest recall rates occurred using logistic regression on gender, with a 90% recall for men and 94% recall for women. The lowest-performing recall occurred in the intersectional analysis, with an average recall of about 72%. This information suggests that the logistic regression model is the best performing out of the models attempted, supporting the efficacy of its use for understanding variable dynamics in the matchmaking context or with similar applications.

4. Conclusion

As we look to the future of digital dating, developers must consider fairness and discrimination in algorithmic designs. Our investigation into the ethical considerations of using machine learning algorithms for matchmaking reveals some challenges, including sampling bias, algorithmic bias, and differential subgroup validity. These challenges not only impact the accuracy and fairness of matchmaking outcomes but also reflect larger social issues regarding equity and representation in dating spaces.

For next steps, dating apps or platforms could try to collect data that is more representative of the population and include a wider range of age groups, racial and ethnic backgrounds, sexual orientations, and geographical locations. This would help mitigate some of the sampling bias.

In addition, collecting more information from participants on what they value could have more consistent predictive power across demographic groups. For example, instead of focusing solely on physical attractiveness, include features related to personality traits, values, and communication styles. Participants' intentions in using a platform could also possibly explain some of their responses. Finally, developing separate predictive models for different demographic groups might better capture the unique relationships between features and outcomes within each subgroup.