Introduction

One of the most important decisions a person must make is who they will marry. In modern Western civilizations, this choice typically comes after a protracted learning period in which individuals participate in more casual, frequently polygamous relationships, or dating, which is the subject of this essay. We specifically examine gender disparities in dating preferences. In the dating markets, it is challenging to infer dating preferences from equilibrium results since there is frequently a correlation between partner traits and different preference structures. By employing an experimental Speed Dating market to research dating behavior, we are able to solve this issue. In our experimental paradigm, participants spend four minutes with each possible partner. The next day, if both parties agree to another meeting, they each get the other's email address. The target variable of this study will be whether the partner wants to see the individual again and vice versa after each "date" or interview. This will be a classification problem where we will try to predict a participant's final decision based on various features. Although there are many shared sexual preferences between the genders, psychologist have long noted significant difference between genders when it comes to premarriage determinants.

There are 190 features for 8,378 rows. The data is collected from restaurants and bars where volunteers participate in speeding dating rounds. There are multiple waves which are basically events that happen on different days. There were a total of 21 speed dating events. At each wave or event, there were different numbers of participants. There are general features that track the race, age, major, test scores, income and city of origin. In addition, the participants were also asked for their religious, racial, and social preferences. Social preference questions include things like how often the individual goes out on dates. Individuals were asked about their hobbies and future career aspirations. They were also asked about what expectations they had about the experiment and what they had hoped to gain. The core of the data would be based on 6 attributes: attractiveness, sincerity, intelligence, fun, ambition, and shared interest. Participants had to rate their partners after each date. They would also rate themselves. The participants would share what they thought other people looked for in potential partners. They would split each of the six values and the total has to add to 100. The participants are also told to rate what they thought others perceived of them. Data is also collected after the speed dating event. However, I decided to omit those features in this study.

**In this study we only focus on heterosexual couples. We will only be using he/she pronouns. We are aware that there are other genders and sexuality, but due to the limitation of data, we are going to solely address heterosexual couples. Furthermore, the complexity would rise substantially if we were to start evaluating the various non-binary identitfying individuals.*

Existing Literature

The data set I decided to use is collected with the intention to conduct research. Raymond Fishman, Sheena Iyengar, Emir Kamenica, and Itamar Simonson published their paper on gender difference in mate selection and used the same data set. In their study they found that women's selectivity depends on the number of potential partners. In smaller sessions, women and men have nearly identical selectiveness. As the number of participants of the opposite sex increased, women became more selective while men's selectivity did not change with respect to group size. The results show a more rapid diminishing return curve when it comes to dating for females. If women simply had a higher cost per date, we would not expect to see any gender difference in the relationship between selectivity and the number of potential partners. The observed differences point to the possibility that females have a more concave benefit function and or a more convex cost function with respect to the number of dates relative to their male counterparts. There are several theories that attempt to explain this discrepancy. Women may have to invest more to prepare for each date. There could also be negative social stigma associated with a woman who goes on too many dates.

Exploratory Data Analysis

The target variable I am trying to predict is the whether or not two people match at the end of a date. The data set is balanced, it has 83.5% of the values as no match and 16.5% of the values as yes match. I removed all the features that were collected after the result of matches were revealed. Despite removing so many features, I still had over 103 columns. I found it very fascinating that during the speed dates the male and females were about the same level selectivity (figure. 1). In my experience, women tend to be much more selective than males.

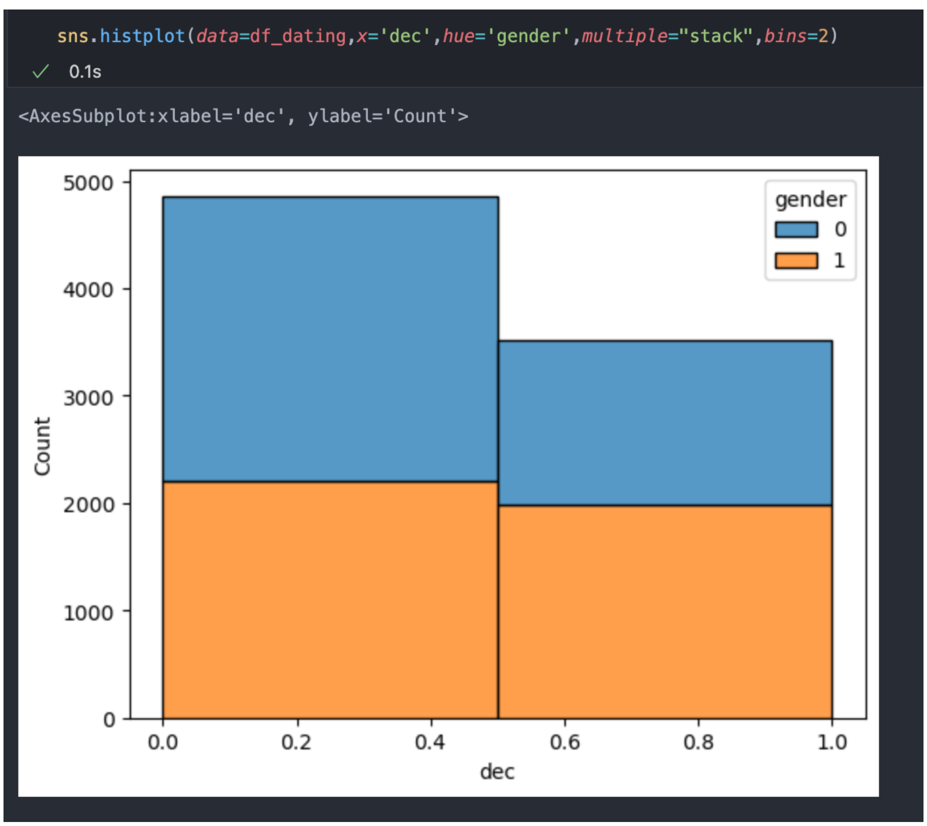


Figure 1: Decision from males and females.

In the research that I am somewhat trying to replicate, the researchers said female selectivity is essentially the same as male selectivity in small groups. When the group sizes becomes larger, women become significantly more selective. I tried to visualize this data, but it was not the most obvious.

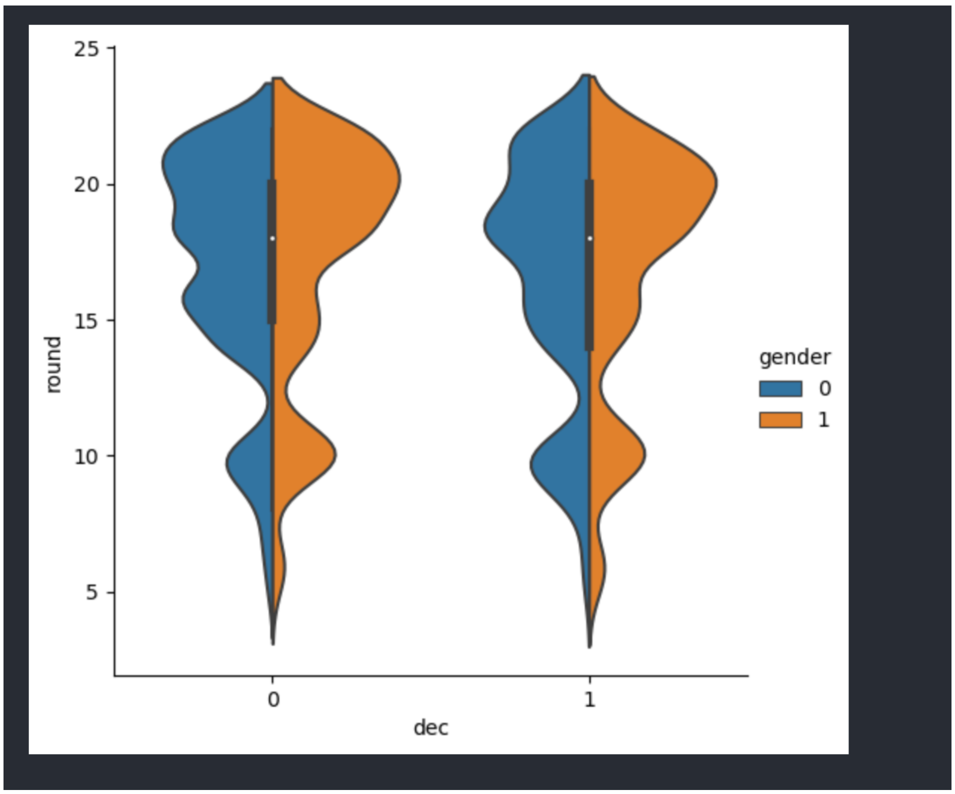


Figure 2: violin chart displaying selectivity based on gender and group size.

There were lots of interesting trends I discovered in the dataset. Out of the 6 core attributes the candidates were ranked on, only three varied significantly between the genders. Males were likely to be seen as attractive when they were rated high in intelligence and ambition. Although female ambition and intelligence mattered, it wasn't as big of an indicator for likeability. This was true in my exploration. However, I found discrepancies when the research claimed that women's intelligence and ambition were negligible when they exceeded those of the male partner. The graph doesn't show a clear distinction between women who are smarter and more ambitious than their male counterparts versus females who are not. What it does show is that in general, men perceive themselves as more intelligent and ambitious than women (figure 3).

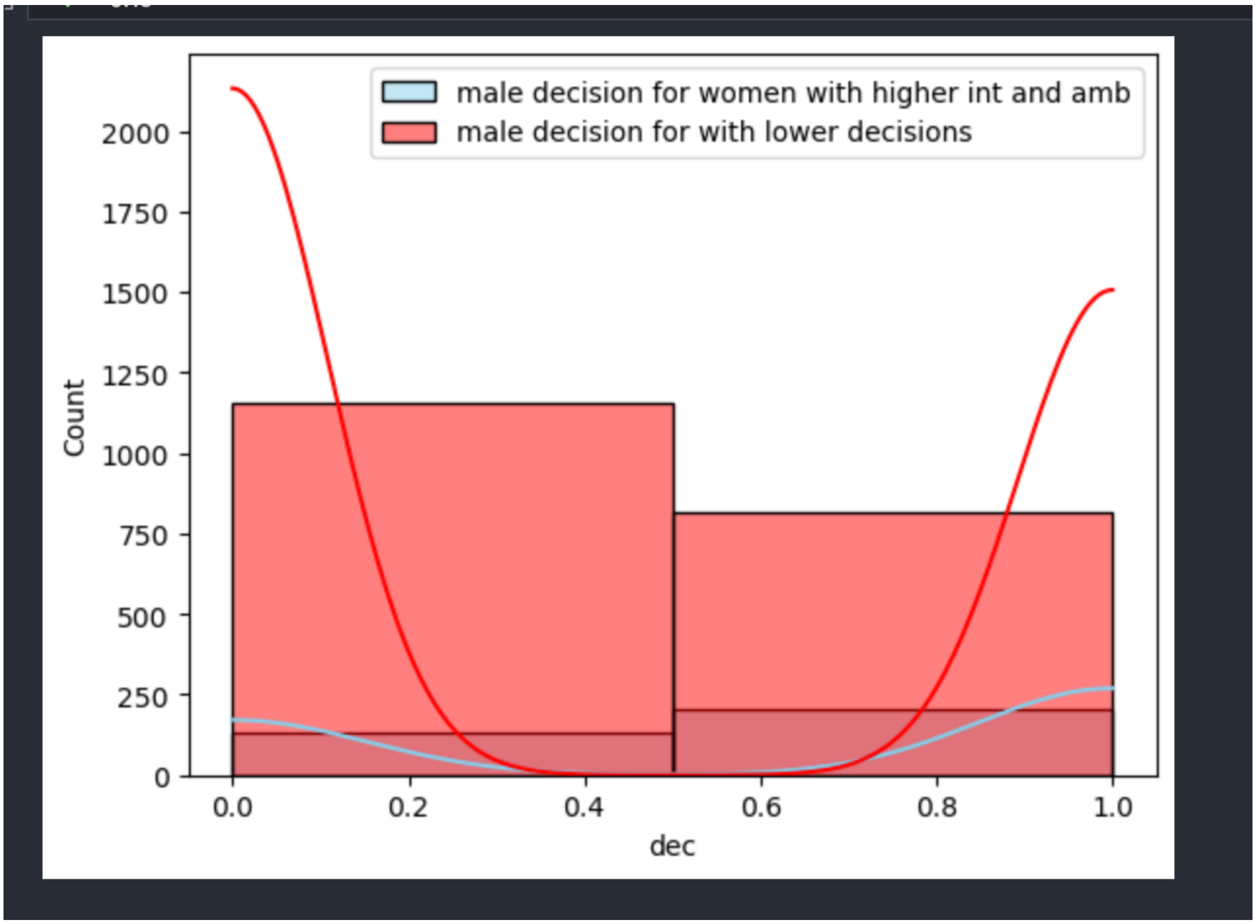


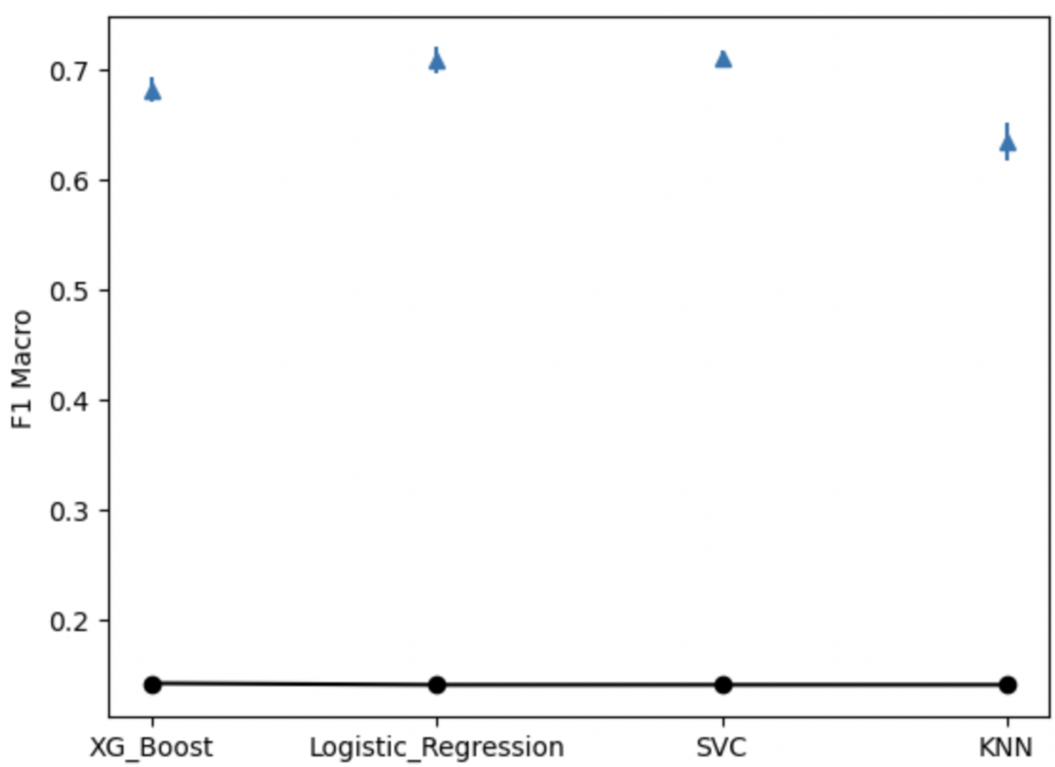
Figure 3: The overlapping histograms do not show a significant difference in preference.

Methods: Data Preprocessing, Cross Validation and Modeling

I split the data using group shuffle split. My data was collected over 21 events. I wanted to keep the events together so I used "wave" as the unique identifier variable. My data set is IID because whatever other people decide to rate you has no impact on what you rate them because you do not have the access to the data. The data parameter and distribution is the same across the board. My data does have group structure. The event variable keeps track of the date when the speed dates occur. Although the events happen on different days, I do not consider the data time series. The day when the event happens is not a key variable. I split the model 60, 20, and 20 because that seems reasonable. First, I used the basic split to remove the test set. Then I used group shuffle split to cross validate my model. I have a medium data set so I was able to be flexible with my splitting. I was lucky in that my data set was specifically prepared for research.

Most of the columns were already encoded and had dummy variables. I used the standard scaler because I had over 100 features so I wanted to be time efficient. The preprocessed data had 103 features. It would have been very unnecessary and time consuming for me to do different scalings. Finally, I used feature engineering to create an additional column based on the zipcode and the partner zipcode column to determine if people had matching zipcodes. I believed that people who came from similar locations were more likely to match.

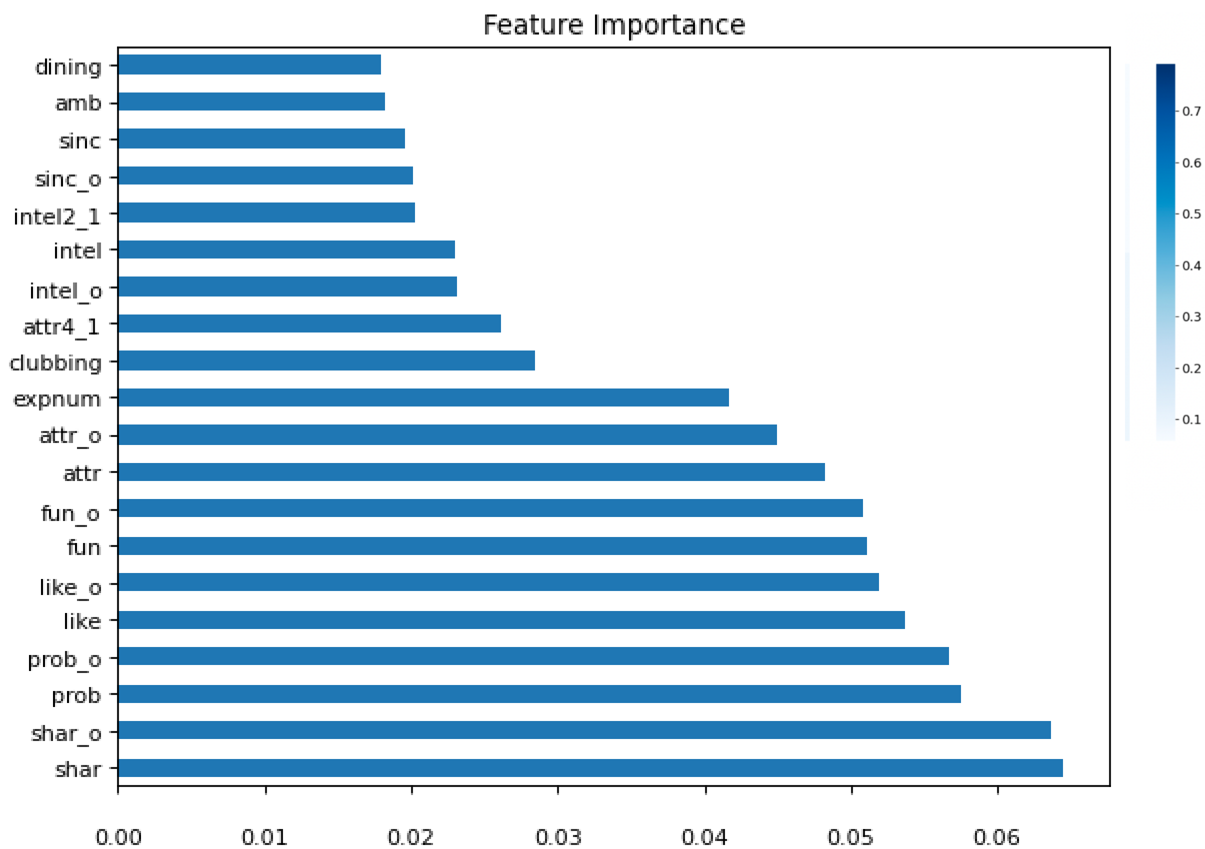
For the first model, I used XGBoost. I had a for loop that would change the random states 3 different times. I kept this number low because it was very time consuming to run the model. Under the for loop, I calculated the baseline score by assuming the model predicted positive for every single row. We predicted positive every time even though it's not the majority class because the goal of our model is to maximize the number of times people match. The stakes are higher for false negatives than false positives. In applications, models like this one would be used to match potential participants. It might waste a person's time if he/she had to look at the profile and reject it. However, if the app did not recommend a person that he/she might have matched with, the person could be missing out on his/her soul mate. Originally, I wanted to use recall as a metric because of our objective. However, I think recall is a bad metric in all scenarios because a model can predict all positives and have a perfect score. I eventually went with F1 Macro to emphasize the importance of the minority class. I tuned three hyper hyperparameters: lambda, gamma, and alpha. I tuned lambda to prevent overfitting. Overfitting could be a problem because of how many features I had. I tuned Alpha to make sure my run time wasn't unreasonably high. Again, I had a lot of features, and I want to reduce my overall run time. I tuned gamma to adjust the threshold difference in the loss function that was required to make the split. The models and the scores we calculate with XG boost are appended to two different lists. I used the model with the highest score and saved the parameter settings. After doing the XG boost model, I created a function that could take in various ML algorithms. The first model I



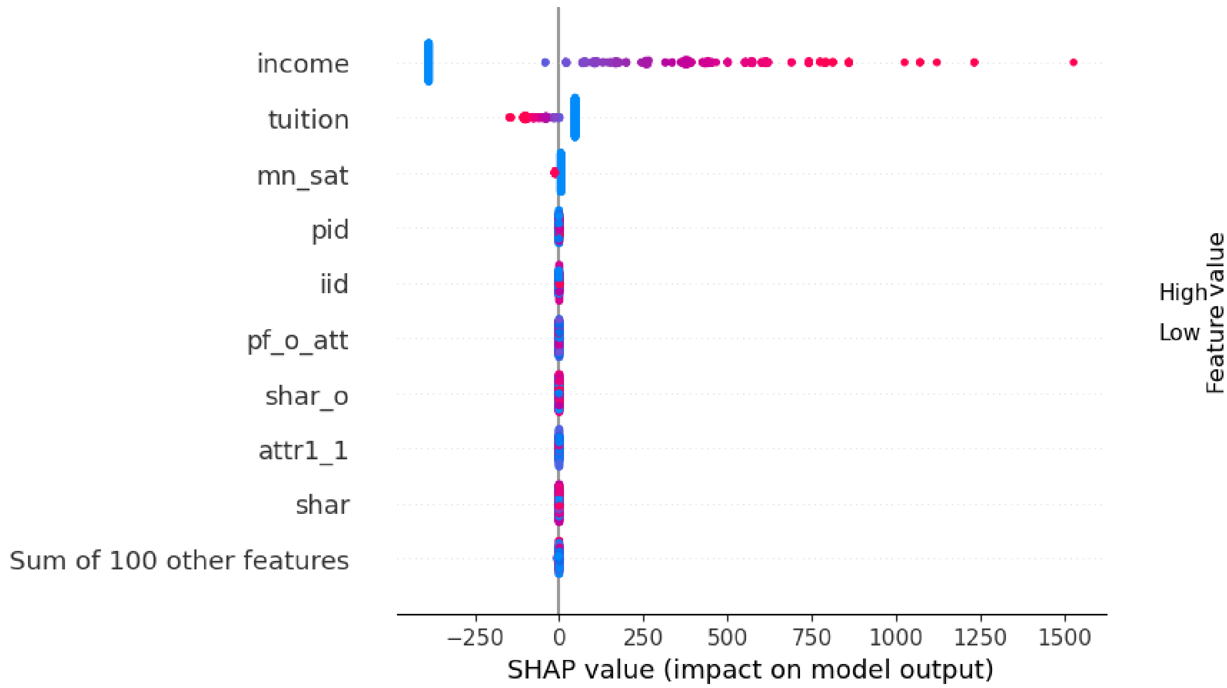
used was logistic regression. In this model, I only used recommended hyper parameter tuning increments. I used gridsearch cv to optimize the process. I used the same splitting strategy as the previous model. I also created a generic function to run logistic regression, support vector machine classification, and K-nearest neighbor.

Results

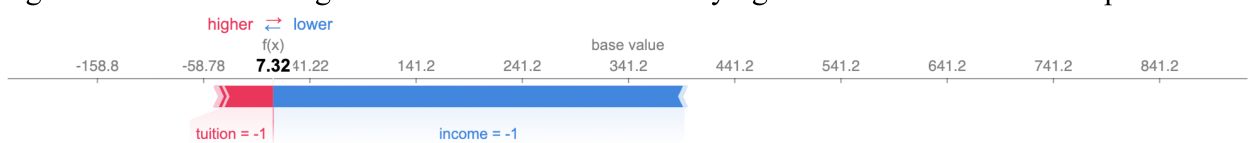
I used the same splitting strategy. I only tuned the recommended hyper parameters from recommendations by data science experts. My two best models were logistic regression and support vector machine classifier. I created confusion for each of the algorithms.



I chose the logistic regression model because it had a higher true positive rate. The logistic regression model is on average 20.35 times better than the baseline model. I show all the largest coefficients for logistic regression. People with shared interests are more likely to match. People who are likable and fun are more likely to match. Physical attractiveness was also an important factor. Expected number of people who will be interested in dating you was also important. I think this is interesting for two reasons. Social experiments in the past shows that individuals are good at predicting their own sexual market value based on how people treat them. In addition, I believe ‘expnum’ is also correlated with confidence. Confidence is very important to attracting potential mates. The graph below shows the shap values for global feature importance.



After finding all the feature importance, I looked at how the features influenced my false negative values. False negatives are the worst. We are trying to reduce this as much as possible.



Outlook

The biggest limitations in this project were time and computational power. If I had more time and computational power, I would have tuned more parameters after understanding what they do. I would also create more columns. The one column I engineered 'zip match' really helped my model. I would also create an ensemble model to see if that would perform better than the individual models. I would also run more tests to see if there are gaps in my conclusion. I would make sure that my findings were rigorous enough to be published.

References

Fisman , R., IYENGAR, S. H. E. E. N. A. S., Kamenica, E., & Simonson, I. (n.d.).

*GENDER DIFFERENCES IN MATE SELECTION: EVIDENCE FROM A SPEED DATING EXPERIMENT**. Retrieved October 19, 2022, from http://www.stat.columbia.edu/~gelman/stuff_for_blog/sheena.pdf

Kristi Gustafson Barlette on February 22, 2016 at 7:00 A. M. (2016, February 22).

About 80 percent of men unattractive, women say. Kristi Gustafson Barlette. Retrieved October 19, 2022, from https://blog.timesunion.com/kristi/2016/02/22/about-80-percent-of-men-unattractive-women-say/#:~:text=But%2C%20apparently%2C%20visual%20satisfaction%20is,This%20chart%20breaks%20it%20down

Stephens-Davidowitz, S. (2022, May 10). *People are dating all wrong, according to Data*

Science. Wired. Retrieved October 23, 2022, from https://www.wired.com/story/data-marriage-behavior-love-psychology-romance/

.

Github link:

https://github.com/hjiangbrown/Final-Project