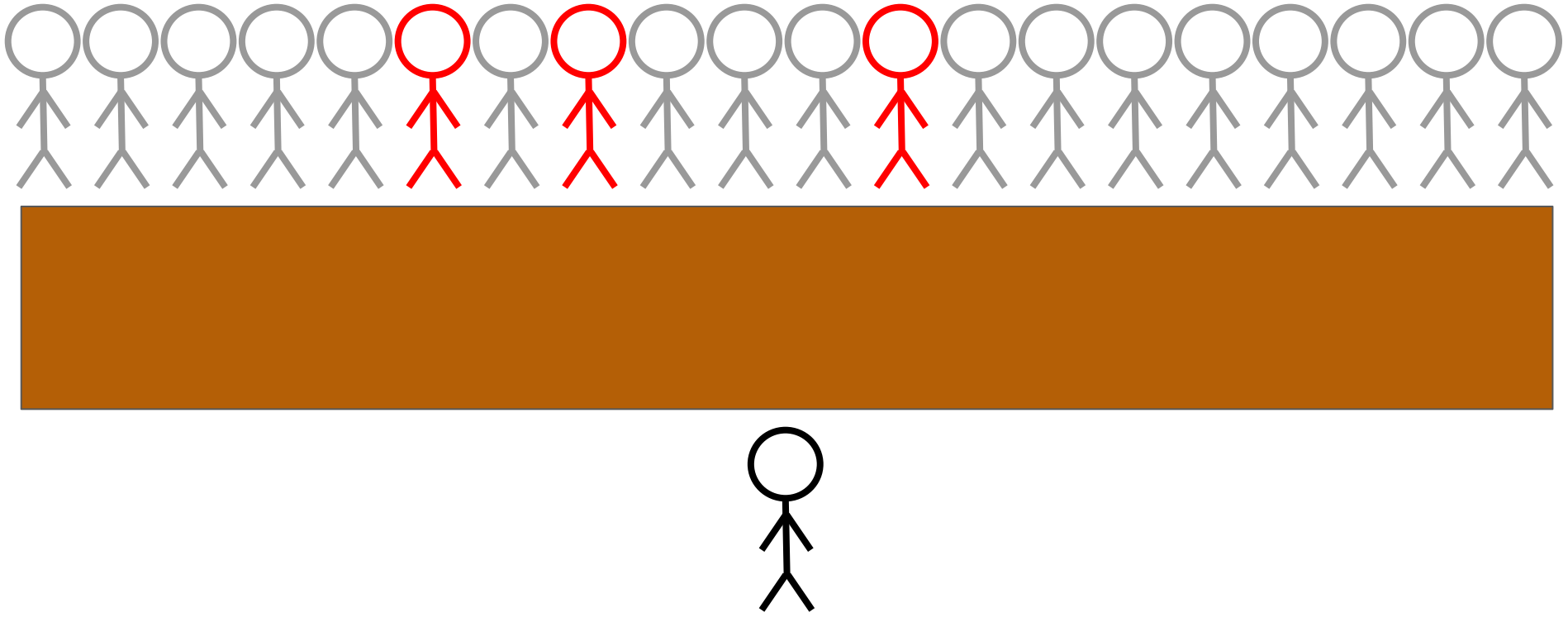
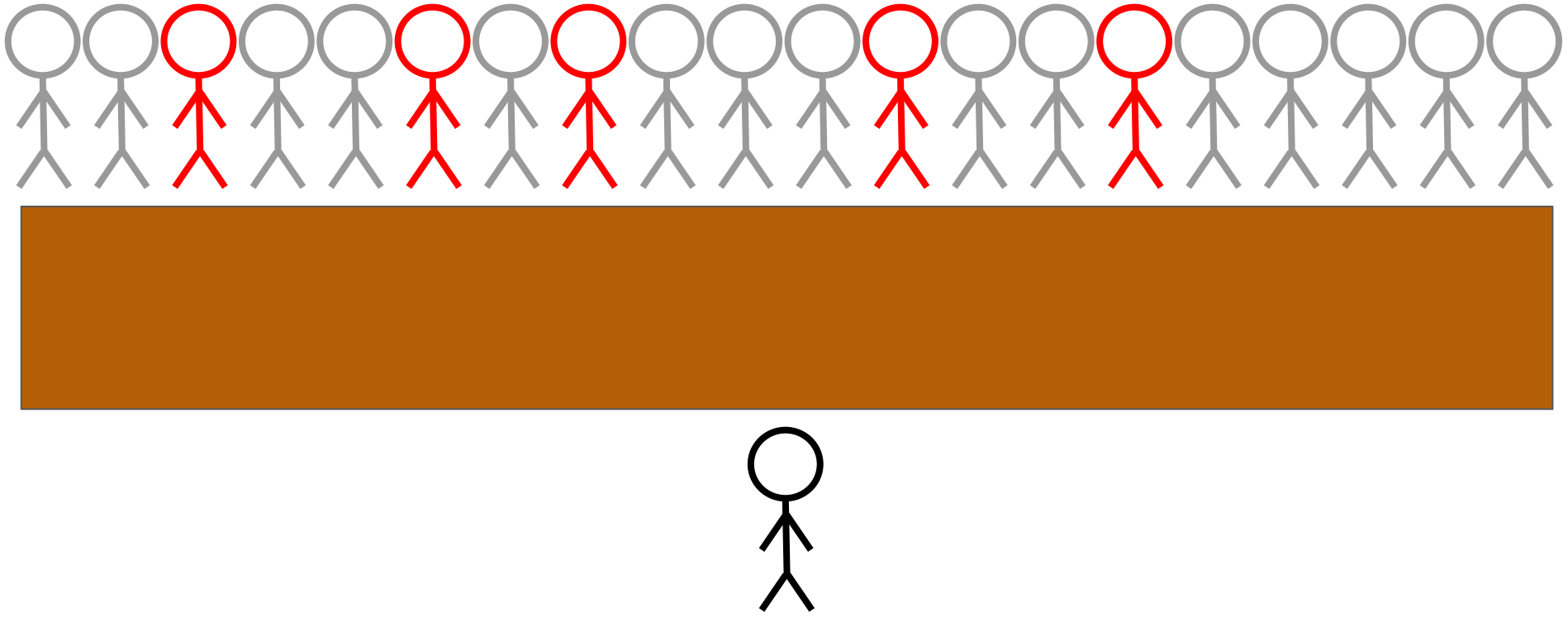
Improving Speed Dating Using Analytics

Anil Kumar
July 12, 2021

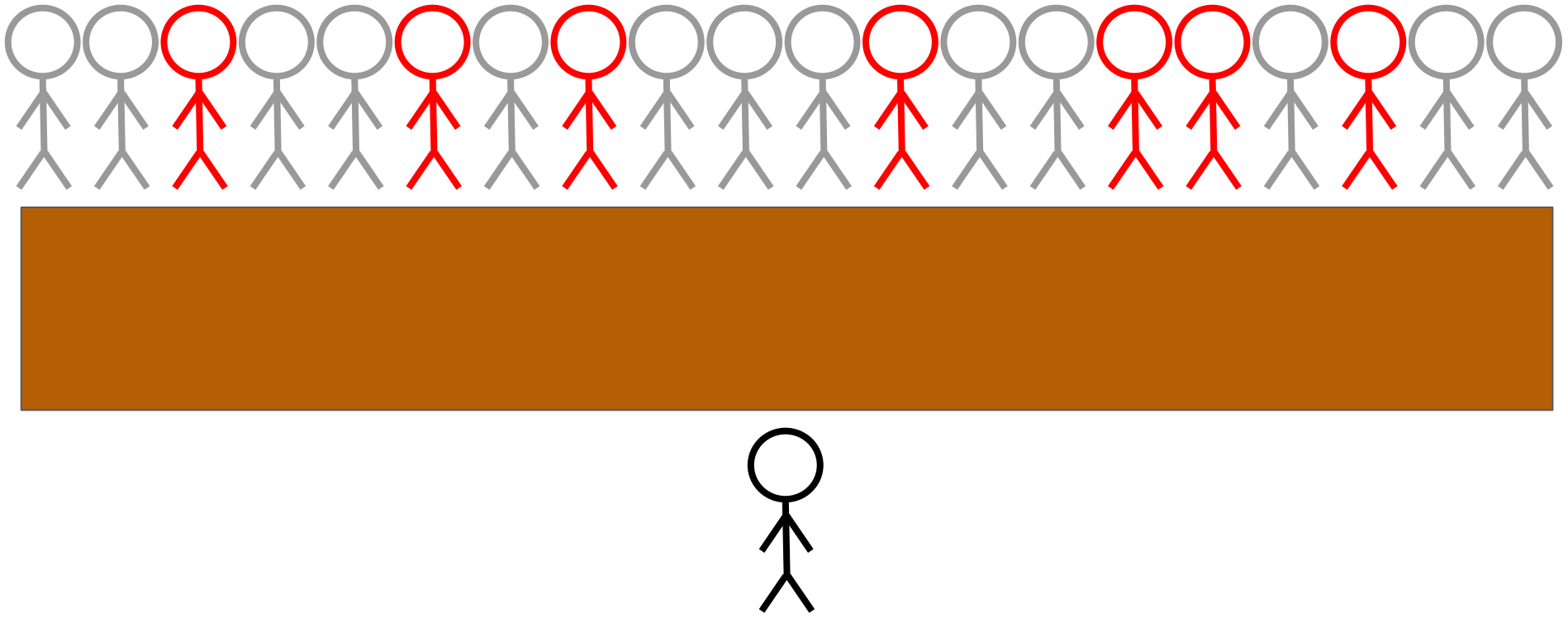
Current Situation



Goal



Even Better



Data and Diagnostic Model

Overview

- **8,378** speed dating encounters
  - **6,998 (83.5%)** non-matches
    - ~17 out of 20
  - **1,380 (16.5%)** matches
    - ~3 out of 20

Features

- **Attractiveness, Intelligence, Sincerity, Funny, Ambition, and Shared Interests**
  - Person's rating of partner
  - Partner's rating of person
  - Person's self-rating
  - How important is the trait to person?
  - How important is the trait is to partner?

- **Socio-demographic**
  - Gender
  - Age
  - Race
  - Are the person and partner the same race
  - Importance that partner is the same race
  - Importance that partner is the same religion

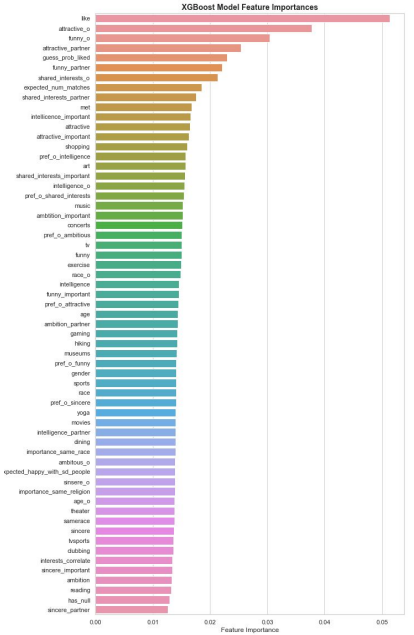
- **Interests**
  - Sports, TV Sports, Exercise, Dining, Museums, Art, Hiking, Gaming, Clubbing, TV, Reading, Theater, Movies, Concerts, Music, Shopping, Yoga

- **Other**
  - How happy does the person expect to be with the people that they meet?
  - How much does the person like their partner?
  - How likely does the person think it is that their partner likes them?
  - How many matches does the person expect to get?
  - Have the person and partner have met before?

Most Predictive Features

- **Created a machine learning model that analyzed the feature impacts on predicting a match**

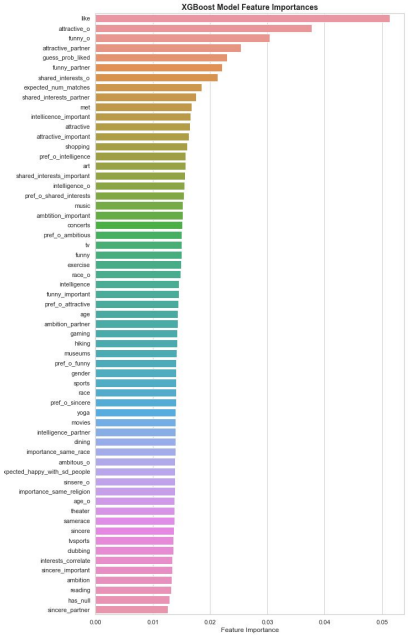
- **Features with the most impact**
  - Attractiveness
  - Funny
  - Liking each other
  - Shared Interests



Most Predictive Features

- **Created a machine learning model that analyzed the feature impacts on predicting a match**

- **Features with the most impact**
  - Attractiveness
  - Funny
  - Liking each other
  - **Shared Interests**



Match Rates Based on Shared Interest Ratings

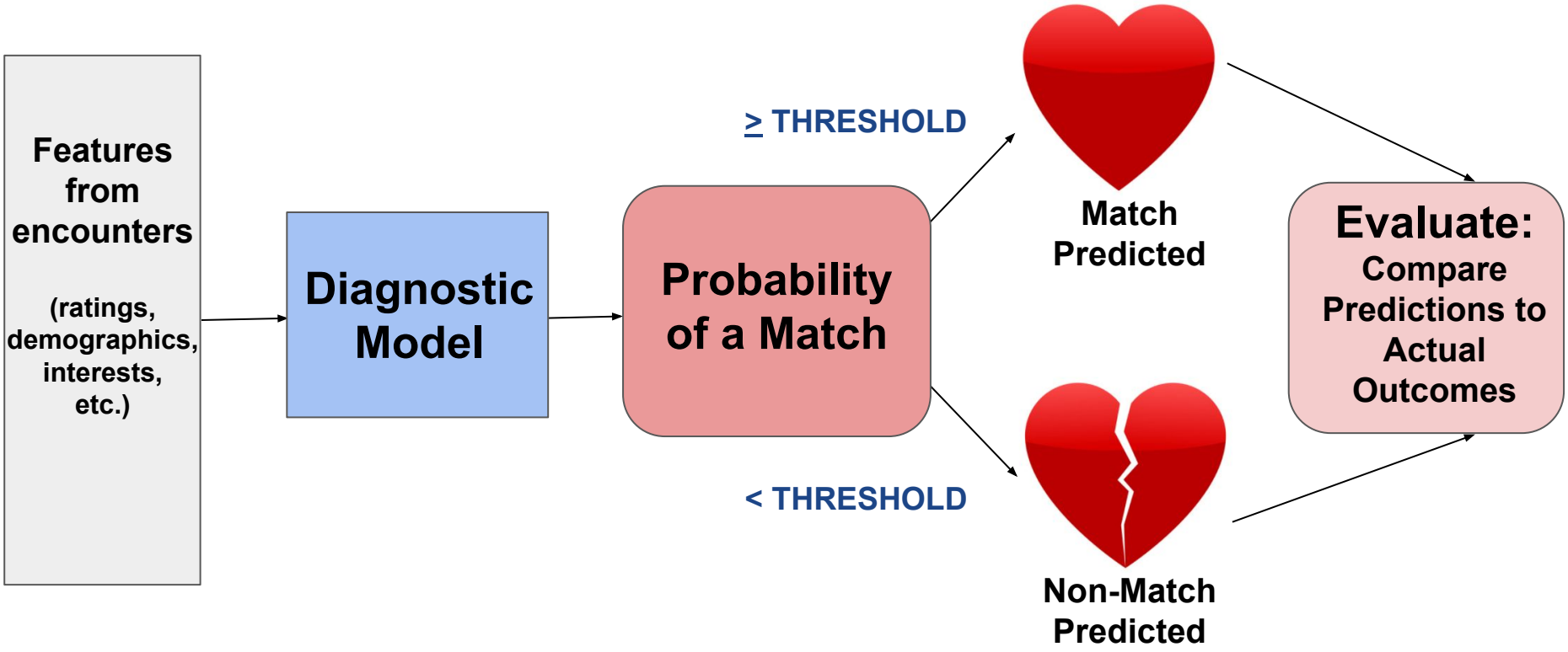
	Match Rate
Both Rated 0-5	5.3%
One Rated 0-5, Other Rated 6-8	14.2%
One Rated 0-5, Other Rated 9-10	18.7%
Both Rated 6-8	32.1%
One Rated 6-8, One Rated 9-10	44.9%
Both Rated 9-10	60.0%

Match Rates Based on Shared Interest Ratings

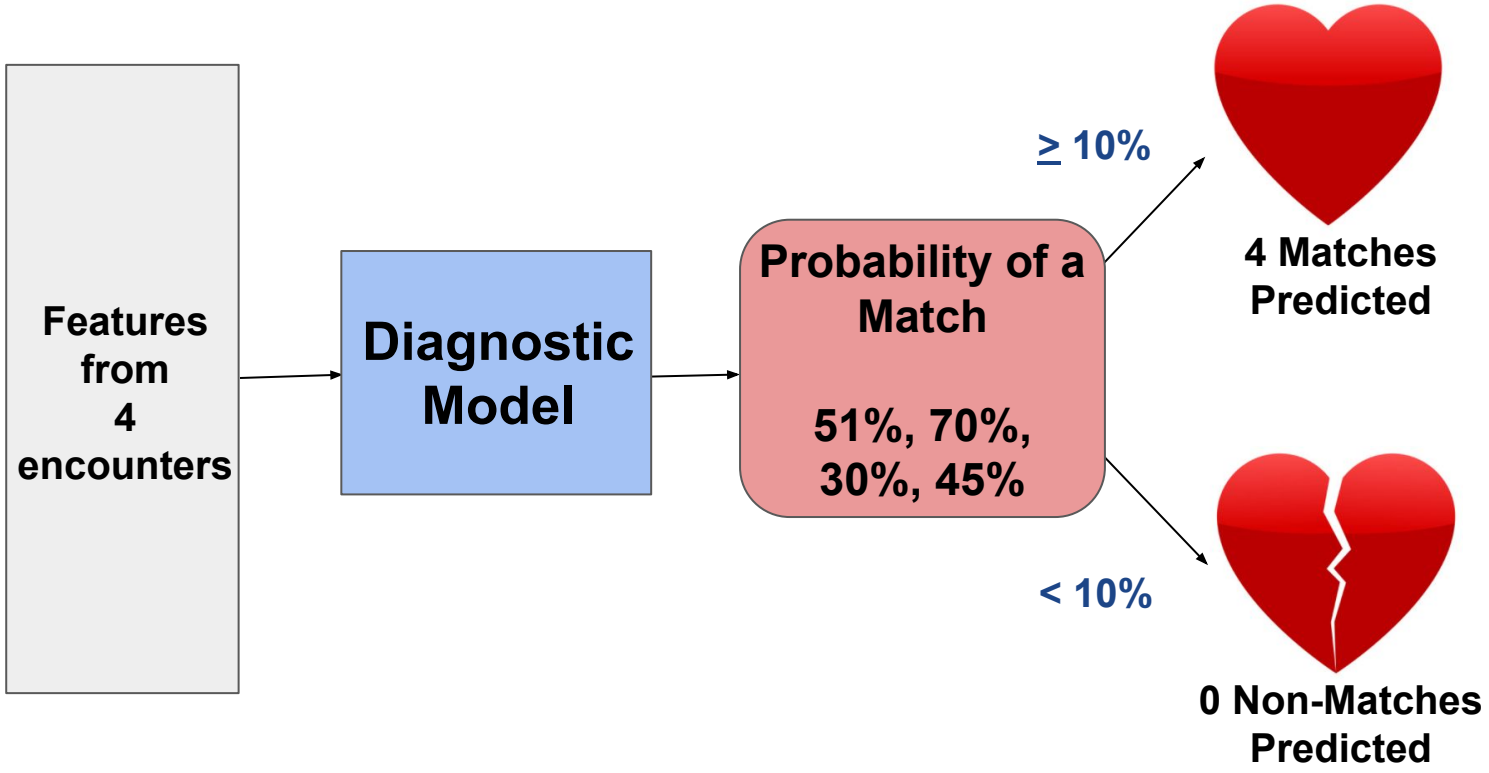
	Match Rate	% of Encounters
Both Rated 0-5	5.3%	35.5%
One Rated 0-5, Other Rated 6-8	14.2%	36.7%
One Rated 0-5, Other Rated 9-10	18.7%	4.3%
Both Rated 6-8	32.1%	16.8%
One Rated 6-8, One Rated 9-10	44.9%	5.6%
Both Rated 9-10	60.0%	1.2%

72.2%
of encounters

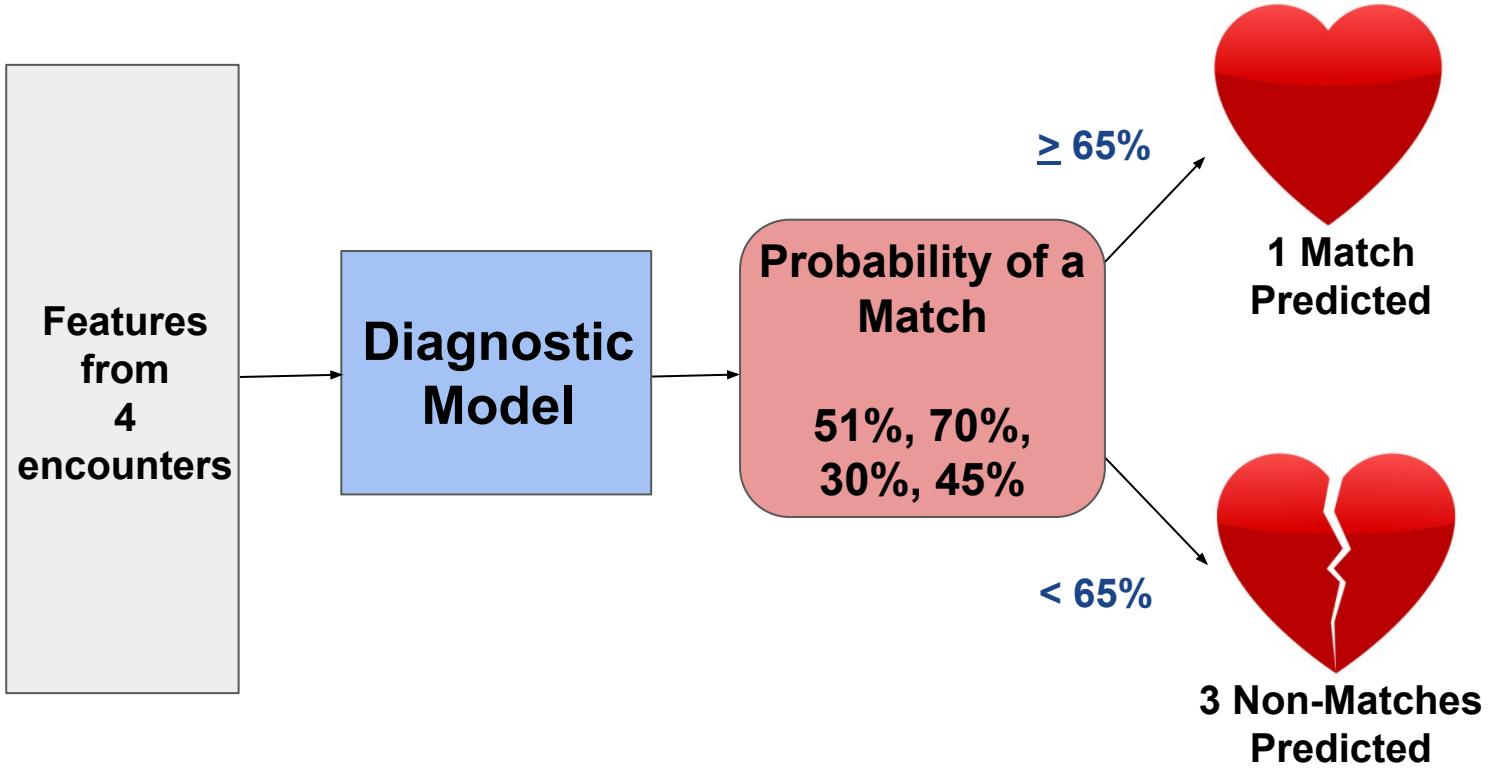
Diagnostic Model - Evaluation



Hypothetical Threshold 10%



Hypothetical Threshold 65%



Results from Diagnostic Model

100 encounters (expect 16-17 matches)	Threshold 10%	Threshold 50%	Threshold 65%
% Encounters predicted as matches	33.7% 33-34 matches predicted	11.2% 11-12 matches predicted	7.3% 7-8 matches predicted

Future Directions

Create Pre-Event Survey

- **Use diagnostic model results to create a pre-event survey that collects data on the most predictive features**
  - Shared Interests
  - Preferences for a partner
  - If repeat customer, use data from previous events
- **Appropriate length**
  - Reduce data quality issues (i.e. missing or invalid entries)

Create Prediction Model

- **Create model to predict matches based on pre-event data**
- **Arrange clients into events that give them the best chances at finding matches**
  - Enhanced experiences
  - Referrals
  - Reviews
  - Reputation
  - Repeat customers

Longitudinal Data Collection

- **Clients that matched**
  - Did matches last in long-term romantic relationships or friendships?
- **Clients that didn’t match**
  - What were the issues?
  - What can we do to help them find matches?

Customer Lifetime Value

CLTV = *Customer Value x Average Customer Lifespan*

CLTV = $30/event x 1 event = **$30**

Over the course of 1,000 customers: **$30,000**

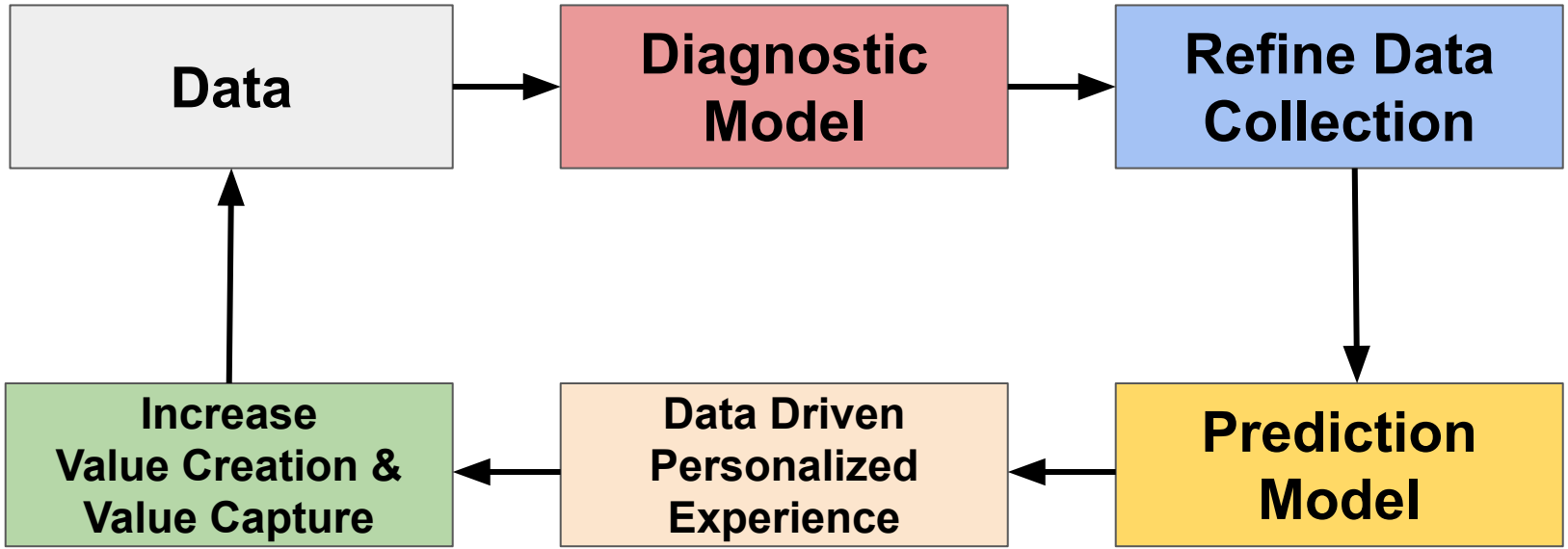
If 500 out of 1,000 customers elect to attend 2 events:

Average Customer Lifespan = 1.5 events

CLTV = $30/event x 1.5 events = **$45**

Over the course of 1,000 customers: **$45,000**

Summary


Data
Diagnostic Model
Refine Data Collection
Prediction Model
Data Driven Personalized Experience
Increase Value Creation & Value Capture

Questions?