Will you predict this rose?:

Examining the sociobiological underpinnings of short-term attraction in a highly structured environment using a deep learning framework.

C. Caggiano, M. Jones, A. Kung

BMI 209
10 May 2018

Motivation: Darwinian evolution posits that an individual's fitness is determined by its ability to successfully reproduce

Figure 1: individuals with high fitness

Figure 2: individual with low fitness

Motivation: Humans often participate in non-random mating which exerts selective pressure.

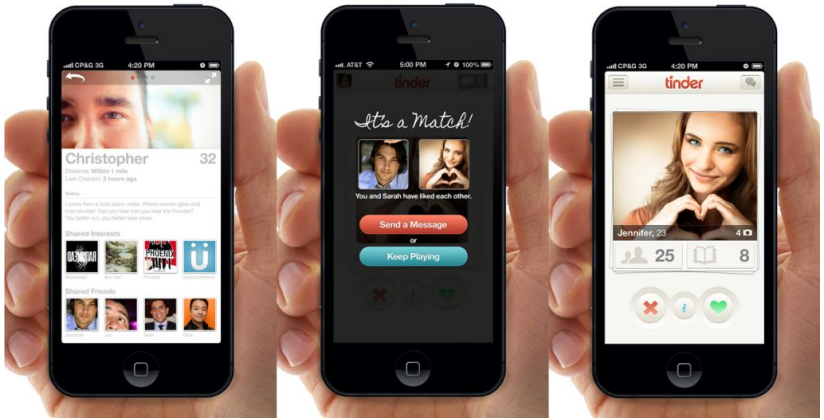


Figure 3: examples of non-random mating behavior in modern human populations

Question: What behavioral choices motivate forces of attraction and non-random mating in human populations?

Population genetics provide two models of non-random assortative mating.

1) Positive assortative mating

Figure 4: mating with positively assorted phenotypes

2) Negative assortative mating

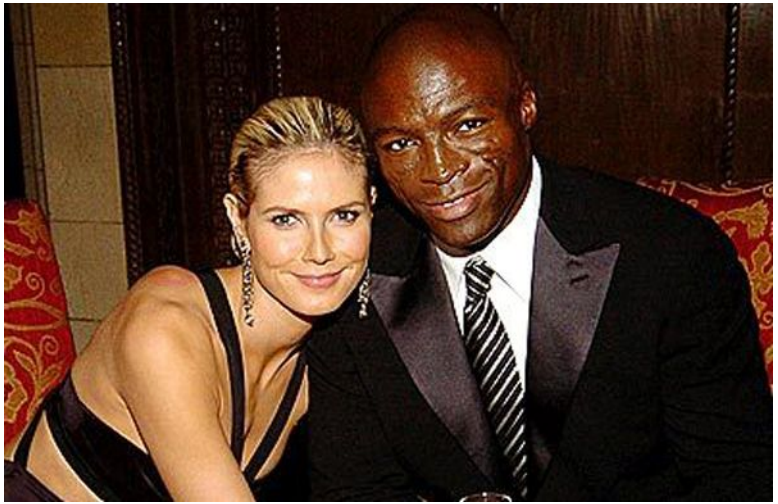


Figure 5: mating with negatively assorted phenotypes

The Bachelor provides an opportunity to generate rich behavioral data in a tightly controlled environment.

One male, 25-30 females

Female contestants iteratively chosen week-by-week until one "winner" remains

**premise reversed on The Bachelorette*

Overarching Goal: Unlock the Mysteries of the Heart

Given a set of contestants, can we predict the weekly selections of The Bachelor using deep learning?

Previous (i.e. Inferior) Efforts

FiveThirtyEight

Politics Sports Science & Health Economics **Culture** Politics Podcast: Primary Season Gets Into Full Swing

How To Spot A Front-Runner On The 'Bachelor' Or 'Bachelorette'

What we learned from analyzing all 33 seasons.

By Ella Koeze and Walt Hickey
Filed under TV
Get the data on GitHub
Published May 22, 2017

HumanGeo Blog

(Machine) Learning About Love: Who will leave The Bachelor next?

FEB 20, 2015 • JIMMY F OPEN-SOURCE PYTHON DATA-SCIENCE BIG-DATA

betches

THE SUP LIFESTYLE POP CULTURE VIDEO PODCASTS BOOKS SHOP

A Breakdown Of The Army Of Skanks Competing For Arie Luyendyk Jr.'S Heart On 'The Bachelor'

BY: SGT. OLIVIA BETCHSON / DECEMBER 11, 2017

Data Collection: Contestant Bios

Becca K.

Bachelorette

contestant biography

Age: 27

Occupation: Publicist

Hometown: Prior Lake, MN

Height: 5'7"

Tattoos: Yes, three. One on my right foot, one on my right wrist and hand.

What are your three best attributes? Loyal, honest, charismatic

What are your top three all-time favorite movies? *Sister Act 2, Gladiator, Pitch Perfect*

What's the most embarrassing thing you listen to? Country (when I'm feeling sad) or the *Sister Act 2* soundtrack, which I don't think is embarrassing at all, but my friends disagree.

What are five things you can't live without? Chapstick, facial lotion, bobby pins, popcorn, and stamps.

What is your biggest date fear? That we will run out of things to talk about and not mesh well together or having food stuck in my teeth.

What are the top three things on your bucket list? Hot-air ballooning, grape stomping, fall in love and own a house and dog with my partner.

<abc.go.com/shows/The-Bachelor/cast>

Data Collection: Feature Extraction

Age

Hometown: size, region

Occupation: broad category, education needed, full-time?

Physical features: hair color, hair length, waviness

Ethnicity

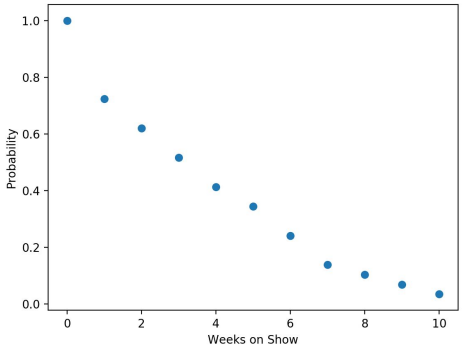
Introduction order

Other Suitors: Random Chance

Real life analogy: meeting a stranger at a bar

$$P(winning) = \frac{1}{N\ contestants} = \frac{1}{25\ to\ 30} \approx 3 - 4\%$$

$$P(surviving\ x\ weeks) = \frac{contestants\ remaining}{total\ contestants}$$



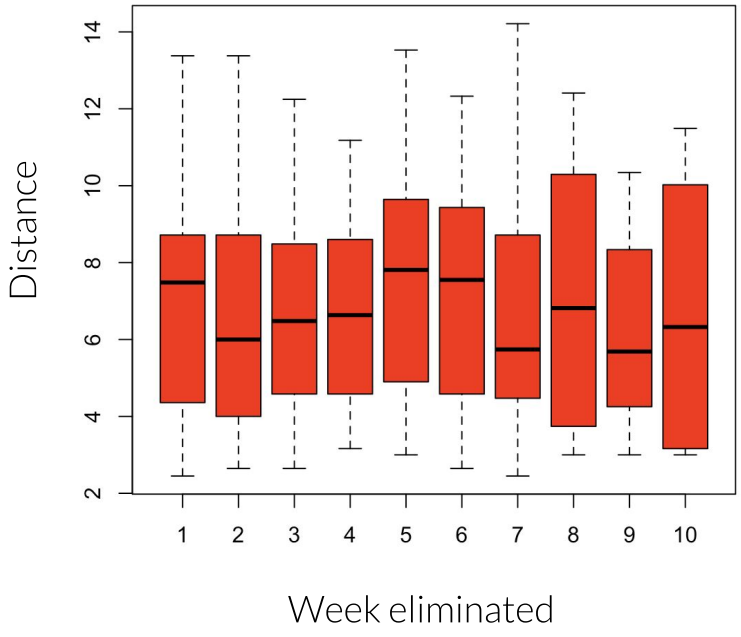
Other Suitors: Linear Regression

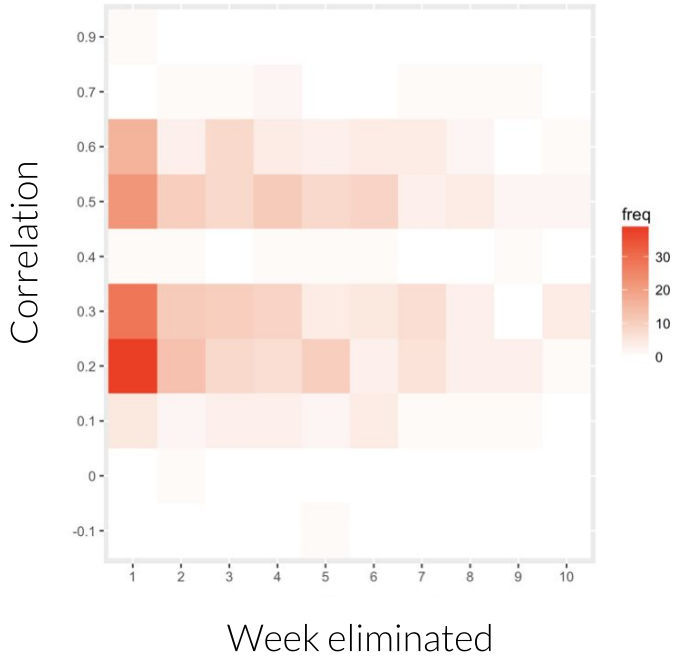
Real life analogy: introduced by a mutual friend who thinks you have a lot in common

Model: correlation between bachelor and contestant ~ success in competition

Linear Regression - results

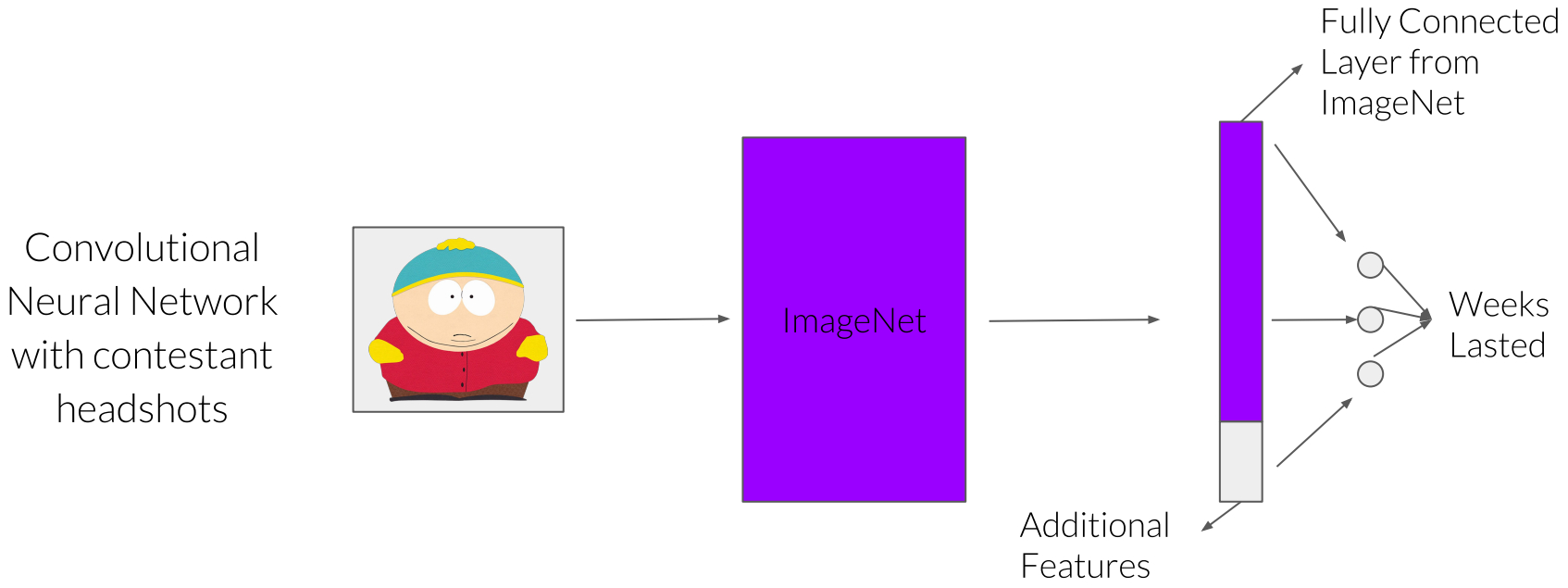
Real life analogy: much like my track record in setting up friends, linear regression does not model a significant interaction (P-value = 0.16)





Deep Learning Architecture

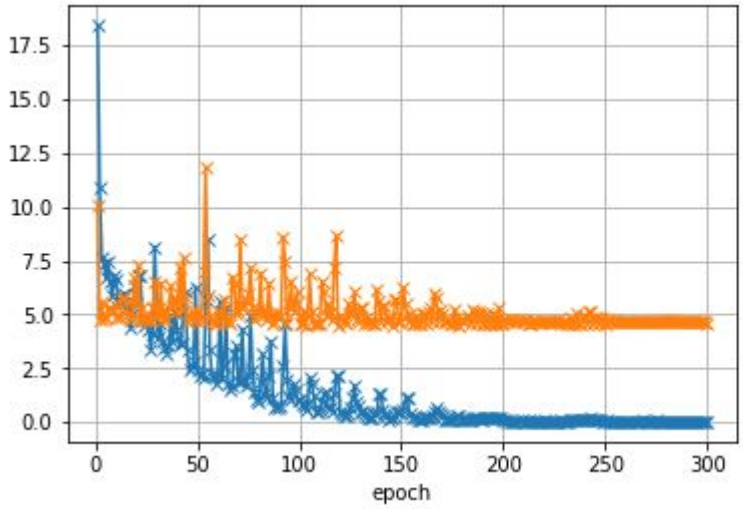
Real life analogy: blindly trusting the judgment of a black box process (either reality television or deep learning)



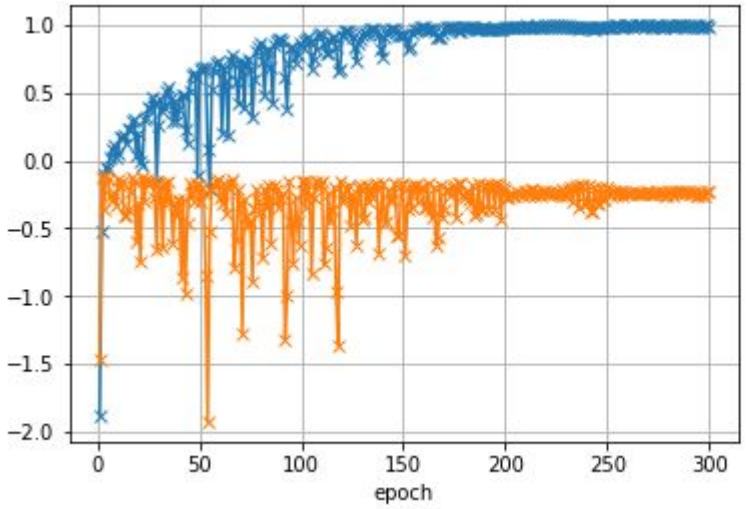
Deep Learning Results

Real life analogy: spending countless hours on a faithless model.
This actually is not an analogy.

Mean Squared Error



R2 Score



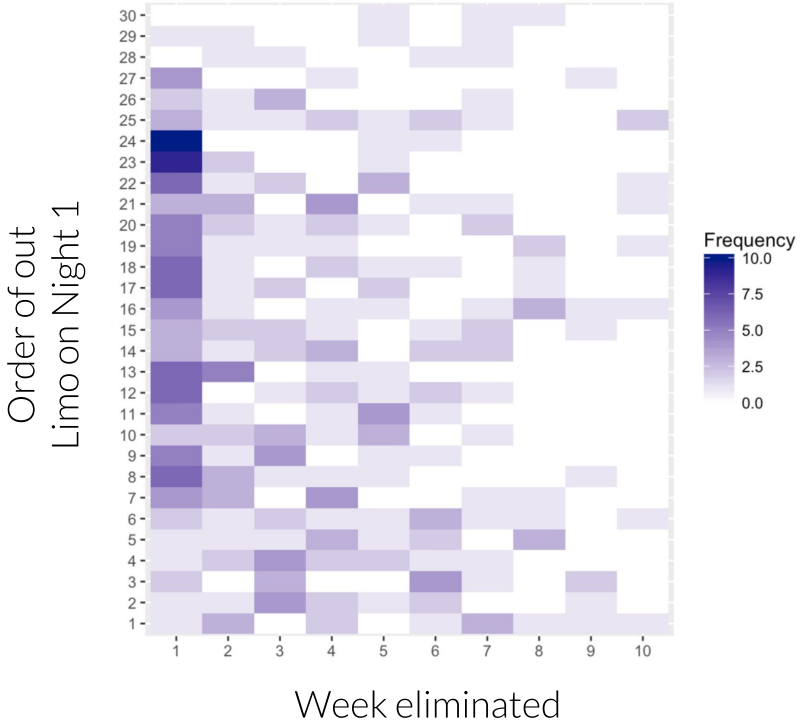
Poor Track Record: A Good Fit After All?

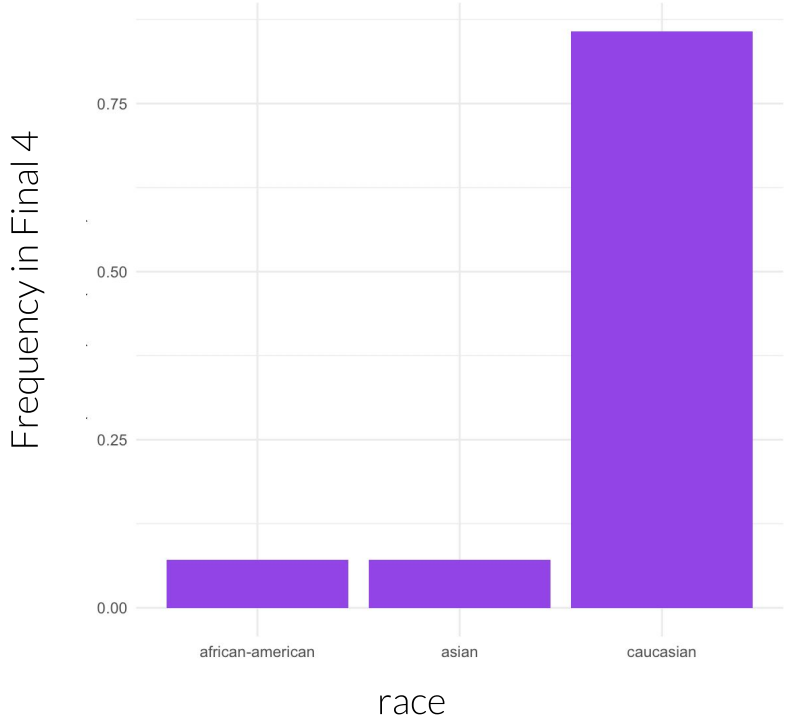
Seasons [edit]

#	Original run	Bachelor	Number of contestants	Winner	Runner(s)-up	Proposal	Still together
1	March 25–April 25, 2002	Alex Michel	25	Amanda Marsh	Trista Rehn	No	No
2	September 25–November 20, 2002	Aaron Buerge		Helene Eksterowicz	Brooke Smith	Yes	No
3	March 24–May 21, 2003	Andrew Firestone		Jen Schefft	Kirsten Buschbacher	Yes	No
4	September 24–November 20, 2003	Bob Guiney		Estella Gardinier	Kelly Jo Kuharski	No	No
5	April 7–May 26, 2004	Jesse Palmer		Jessica Bowlin	Tara Huckeby[21]	No	No
6	September 22–November 24, 2004	Byron Velvick	27	Mary Delgado	Tanya Michel	Yes	No
7	March 28–May 16, 2005	Charlie O'Connell	25	Sarah Bryce	Krisily Kennedy	No	No
8	January 9–February 27, 2006	Travis Lane Stork		Sarah Stone	Moana Dixon	No	No
9	October 2–November 27, 2006	Lorenzo Borghese	27	Jennifer Wilson	Sadie Murray	No	No
10	April 2–May 22, 2007	Andrew Baldwin	25	Tessa Horst	Bevin Nicole Powers	Yes	No
11	September 24–November 20, 2007	Brad Womack		None	DeAnna Pappas & Jenni Croft	No	No
12	March 17–May 12, 2008	Matt Grant		Shayne Lamas	Chelsea Wanstrath	Yes	No
13	January 5–March 3, 2009	Jason Mesnick		Melissa Rycroft	Molly Malaney	Yes	No[a]
14	January 4–March 1, 2010	Jake Pavelka		Vienna Girardi	Tenley Molzahn	Yes	No
15	January 3–March 14, 2011	Brad Womack[46]	30	Emily Maynard	Chantal O'Brien	Yes	No
16	January 2–March 12, 2012	Ben Flajnik[48]	25	Courtney Robertson	Lindzi Cox	Yes	No
17	January 7–March 11, 2013[50]	Sean Lowe[51]	26	Catherine Giudici	Lindsay Yenter	Yes	Yes
18	January 6–March 10, 2014	Juan Pablo Galavis[55]	27	Nikki Ferrell	Clare Crawley	No	No
19	January 5–March 9, 2015	Chris Soules[59]	30	Whitney Bischoff	Becca Tilley	Yes	No
20	January 4–March 14, 2016[62]	Ben Higgins[63]	28	Lauren Bushnell	Joelle "JoJo" Fletcher	Yes	No
21	January 2–March 13, 2017[65]	Nick Viall[66]	30	Vanessa Grimaldi	Raven Gates	Yes	No
22	January 1–March 6, 2018	Arie Luyendyk Jr.[68]	29	Rebecca "Becca" Kufrin	Lauren Burnham	Yes	No[b]

Wikipedia

But… is there a more cynical way to predict the rose?





Future Directions

1) Siamese network: two identical neural networks for one-shot image recognition
2) CelebA dataset: 200,000 images of celebrity faces
3) OKCupid (dating app) in San Francisco dataset
4) Features of interest that we were unable to currently extract
5) Collect more data (i.e. hope the show doesn't get cancelled)

Acknowledgements

Kangway Chuang

Mike Keiser/Chris Harrison (?)

Nobu Ota

BACHELOR NATION

The BACHELOR

THEBACHELOR hot