What is the social impact of development practices on dating?

Jessica Pidoux

Digital Humanities Institute, EPFL

March, 2021

CC BY ND

A qualitative analysis of Tinder's algorithm patents

Reference: Jessica Pidoux « Toi et moi, une distance calculée. Les pratiques de quantification algorithmiques sur Tinder. » in Yann Calbérac, Olivier Lazzarotti, Jacques Lévy & Michel Lussault (dir.), Carte d'identités. L'espace au singulier, Paris, Hermann, 2019.
https://infoscience.epfl.ch/record/283981?&ln=en

Jessica Pidoux

Digital Humanities Institute, EPFL

March, 2021

CC BY ND

Objective of this section

To review development choices of Tinder's matching system and its social impact on gender, and more broadly on couple formation, through a qualitative analysis.

This type of analysis should help data scientists and engineers to reflect on their daily practices and learn to justify their choices when defining algorithms.

Outcomes

1. Situate the impact of inequalities produced by development choices through a practical case study; the dating app Tinder
2. Reflect on the relevance of critical thinking and the justification of choices for better development practices
3. Stay aware of the social and human implications of the data science and engineering profession
4. Understand the socioeconomic stakes within development work organization
5. Discuss collectively how a human rights-based approach could help implement better algorithms

Tinder outreach

Most downloaded app after Netflix:
around 56M users worldwide

The average subscribers worldwide were 6M in Q1 2020 as reported by the Match Group

Present in 190 countries

Mainly adopted by hetero young educated people (18-35 years old)

20% of Swiss in a 5-year relationship met a partner in a dating platform (2018)

CC BY ND

TINDER’S PATENTS

DATA COLLECTED

from Facebook and Tinder are of different **type**:

Demographics

Includes data such as the first name, last name, pseudonym, place of birth, date of birth, university, parental profession, income, ethnicity, level of education, sex, height, weight.

Linguistics

Analysis of keywords, interests and hobbies declared, total number of words, average number of words per sentence, total number of words with more than three syllables.

Behavioural

Takes into account the demonstration of a preference for a user (swipes), sending a message to a user, the number of times the user has been part of the list of results of other users

Declarative variables

Messaging

Online activity

CC BY ND

SOCIAL IMPACT

Data collection: profiling the user from different sources. Big Data fallacy: the more we collect the better, but what type of information matters to users for finding a date?

Feature selection: direct impact on recommendation results. Feature bias provides a distorted view of a type of user that is ‘good’ to recommend while discriminating other profiles that do not fit in the features selected.

Weighting features: not every feature is equally important which can be an advantage for some users (with a high level of education, good writing skills) and a disadvantage for others.

METRICS

Each user gets a **value** or score according to different **metrics**

Physical attractiveness

I.Q.

Readability

Nervousness

Similarity

CC BY ND

USER ORDERING

The system presents in a specific **order** someone **equivalent** according to:

- Geolocation
- Age difference
- Physical Attractiveness ratio
- Education level and occupation

Conditions might be applied if:

- If the score is high between users, preferences declared by the user on the application are ignored
- If the score is high between users, a closer geographical proximity is shown

CC BY ND

SOCIAL IMPACT

Establishment of metrics: matching users based on similarity reduces diversity.

Ordering: it establishes a hierarchical organization of users by popularity and attractivity (the “most popular” ones with more likes that are supposed to be “more beautiful” are presented in the top of recommendation results) which discriminates those that are not fitting the standards.

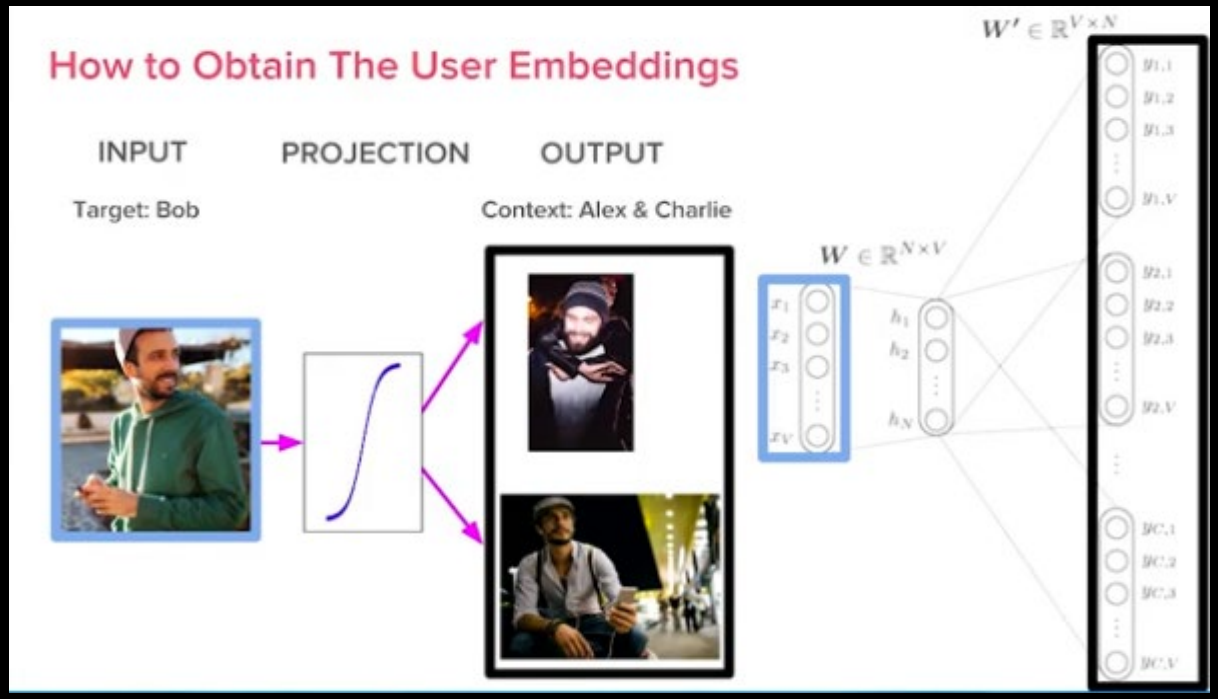
Reinforcement of stereotypes by favouring a “patriarchal model”: heterosexual users are matched based on the man’s superior socioeconomic position in respect to the woman with an inferior level. Same sex users are not covered in the patents.

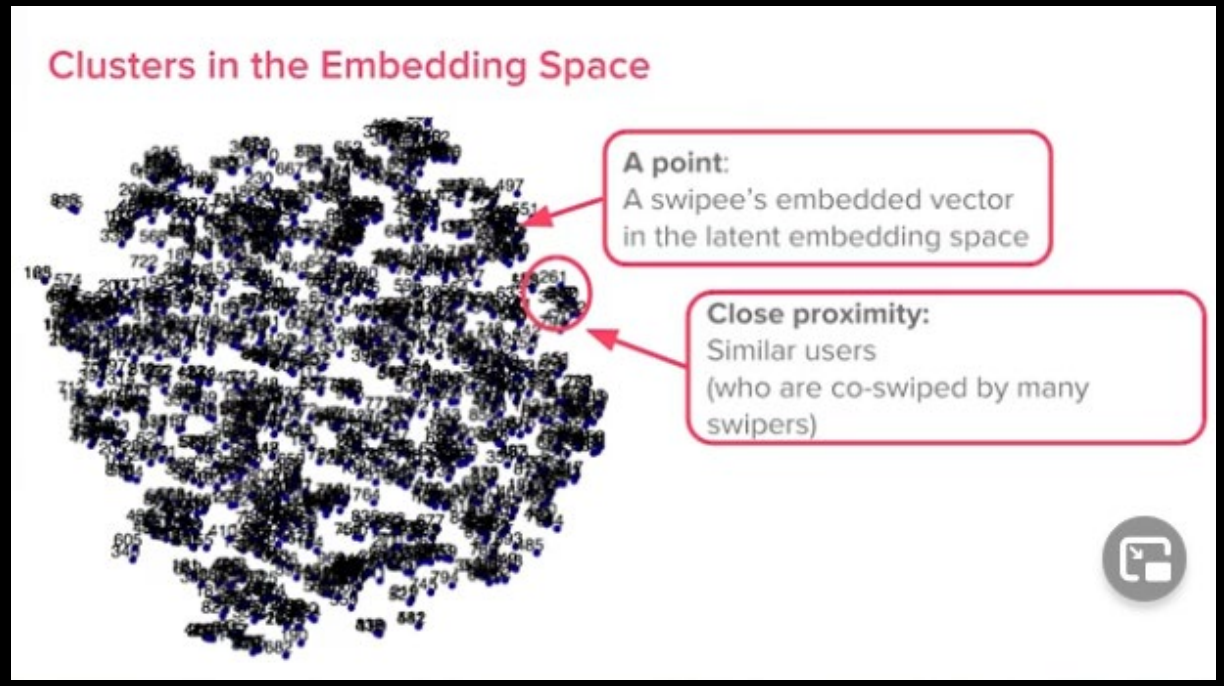
Older men are introduced to younger women. It discriminates as the system does not provide the same opportunities to everybody.

An important age difference between partners reinforces the patriarchal model as one of the partner will be in an advanced stage of the professional life that will be more difficult to the other person to catch up.

AN UPDATE ON TINDER'S ALGORITHMS PRESENTED IN A MACHINE LEARNING CONFERENCE

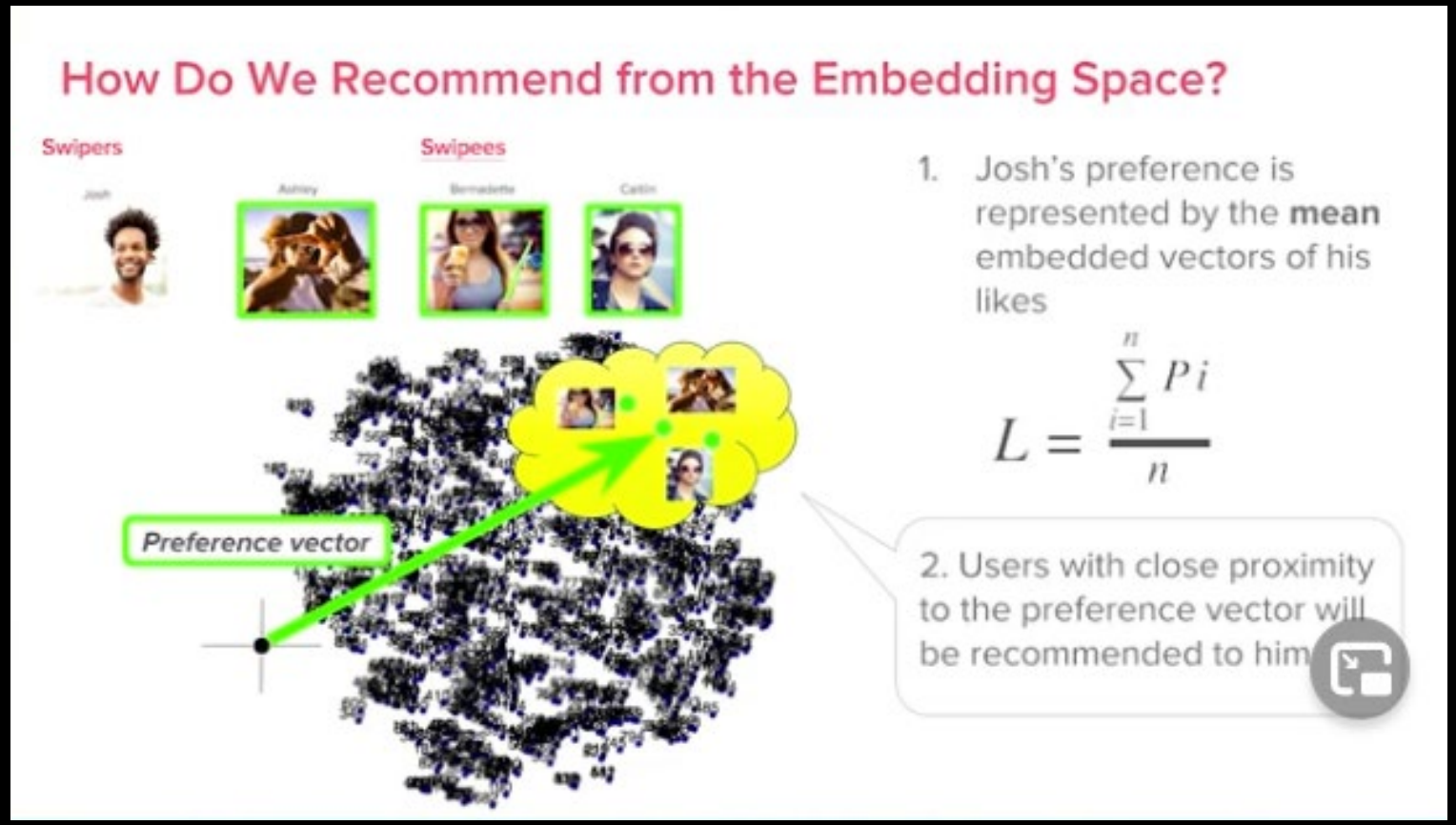
TINDER'S NEURAL NETWORKS





Probably thousands of dimensions, meaning that there are *thousands of individual values associated with each user*.

TINDER'S NEURAL NETWORKS



Recommendation: average all the users you've liked previously, then find other users that are similar (meaning their embedding representation has a low distance to the center).

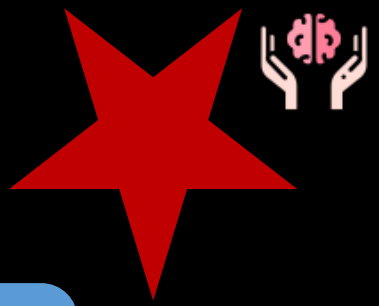
SOCIAL IMPACT

When implementing neural networks the interpretability of the results obtained is extremely difficult. Each value in the clusters describing users probably does not have a meaningful and legible interpretation for a human. How to explain and justify why some users are more or less favoured in the app?

Recommendations by similarity of preferences based in past experiences can trap users in a filter bubble. It reduces diversity in the future experiences.

A STUDY ABOUT DATING APP FOUNDERS' AND DEVELOPERS' PRACTICES

Developers' dynamics



Personal experience
(see Madeleine Akrich)
No user modeling

Lead Dev

PM

Data engineer

Test QA

Devs

User technical support

Tension between personalization and standardization (automatic traces and mainly direct feedback was useful)

Continuous experimentation (see Gerald Schermann) mobile diversity and user behavior impossible to capture, shared experiences, adapting general standards to specific problems

Reference: PIDOUX J., KUNTZ P., GATICA-PEREZ D., 2021, « Declarative Variables in Online Dating: A Mixed-Method Analysis of a Mimetic-Distinctive Mechanism », 5, CSCW1, p. 100-132: https://infoscience.epfl.ch/record/283917?&ln=fr

CONCLUSION

- The match is a mathematical distance defining a social distance that impacts couple formation at large scale!
- Development choices have real consequences in society: reinforcing a type of couple, a type of preferences by similarity while reducing the probability of finding different users.
- Work organization and methodologies of development are influenced by business logic. Development choices seek first efficiency and profit.
- It also affects from the end-user interface how users evaluate themselves and others (fast speed physical elimination) as they learn and adopt how the system works. But not everyone get matches or dates!