

BERLIN BUZZWORDS 2019

# Bias in NLP 101

By Sakshi Shukla

# Problems, Where?

Everywhere

Like?

- Movies
- Books
- Language Translation



# Best Company?



That means No Error, Right?

Let's Find Out

---

# Google Proves to be biased against Women!

## Women less likely to be shown ads for high-paid jobs on Google, study shows

Automated testing and analysis of company's advertising system reveals male job seekers are shown far more adverts for high-paying executive jobs



▲ One experiment showed that Google displayed adverts for a career coaching service for executive jobs 1,852 times to the male group and only 318 times to the female group. Photograph: Alamy

- Google
- Amazon
- Facebook

Are in the Same League

# Facebook and Amazon

## Campaigners begin action against male-targeted job ads on Facebook

Workers claim that Facebook provides job ads which allows employers to choose who they want based on their age and sex



## Amazon scraps secret AI recruiting tool that showed bias against women

Jeffrey Dastin

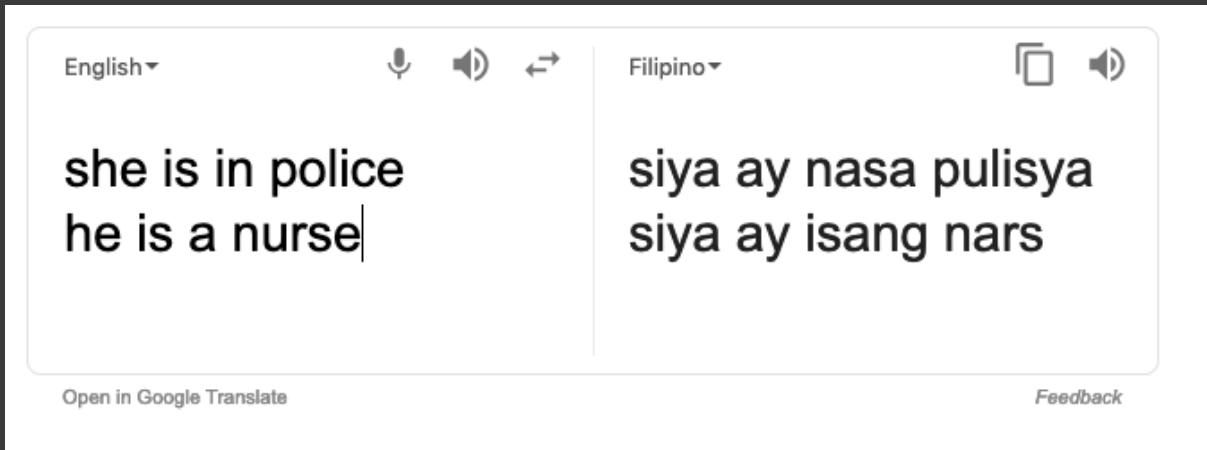
8 MIN READ








SAN FRANCISCO (Reuters) - Amazon.com Inc's (AMZN.O) machine-learning specialists uncovered a big problem: their new recruiting engine did not like women.

Renowned Companies didn't  
have the most efficient  
Algorithm!

# Only with Jobs?



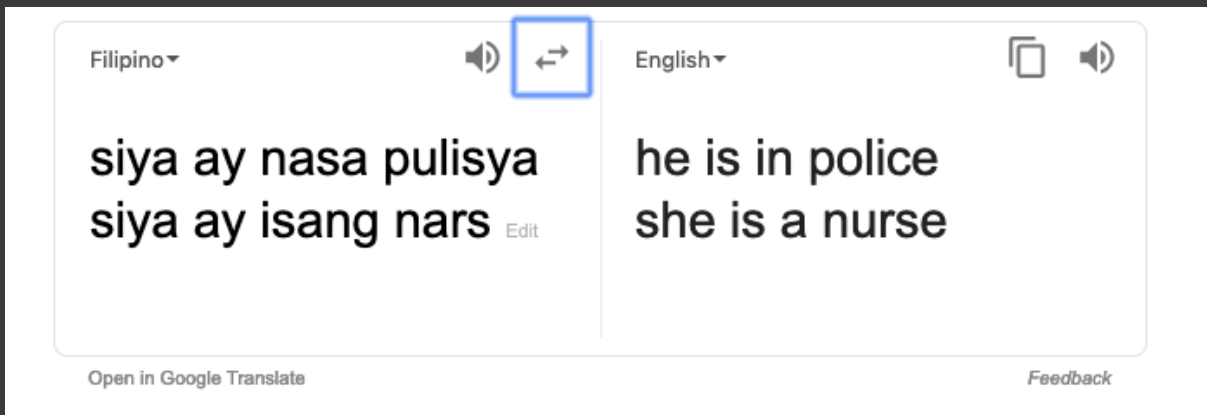
English ▾    Filipino ▾  





she is in police  
he is a nurse |

siya ay nasa pulisya  
siya ay isang nars

[Open in Google Translate](#) [Feedback](#)

Occupations are  
decided by genders



Filipino ▾   English ▾  

siya ay nasa pulisya  
siya ay isang nars Edit

he is in police  
she is a nurse

[Open in Google Translate](#) [Feedback](#)

---

Google Translation fails!

# Bias in Life: Books



# Bias in Life: Movies

**Our Movie**

It changed the plot of women waiting for a man to rescue them to men rescuing women.

Our audience is the same as before, young girls.

This shifted the bias from women to men. Instead of women waiting around for men, to a woman saving a man.

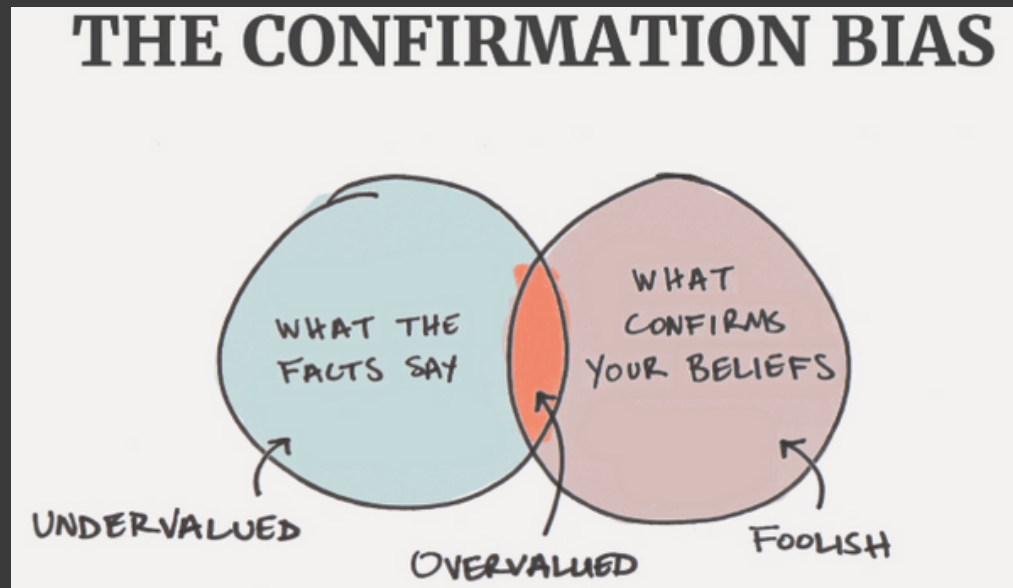
**THE CITY OF DREADFUL NIGHT**

## Gender Bias in Disney Princess Movies



# Bollywood Movie Dialogues

Rohit is an aspiring singer who works as a salesman in a car showroom, run by Malik. One day he meets Sonia Saxena, daughter of Mr. Saxena (Anupam Kher), when he goes to deliver a car to her home as her birthday present.



# Detected Bias?



How to...

- Analyse the Stereotypes
- Dataset Present for de-bias
- Develop De-bias Model



# Analyse the Stereotype: Words!

angry aspiring assistant brave **corrupt** criminal cruel decent  
disabled drunk dutiful eccentric fake **famous** foster furious greedy grown  
handsome happy homeless **honest** horrified illegitimate  
innocent lead lost loving middle mysterious nice  
**poor** powerful proud **real** respected  
**rich** rival ruthless shy **simple** single small spirited strict  
strong successful terrorist unemployed  
**wealthy**

(a) Adjectives used with males

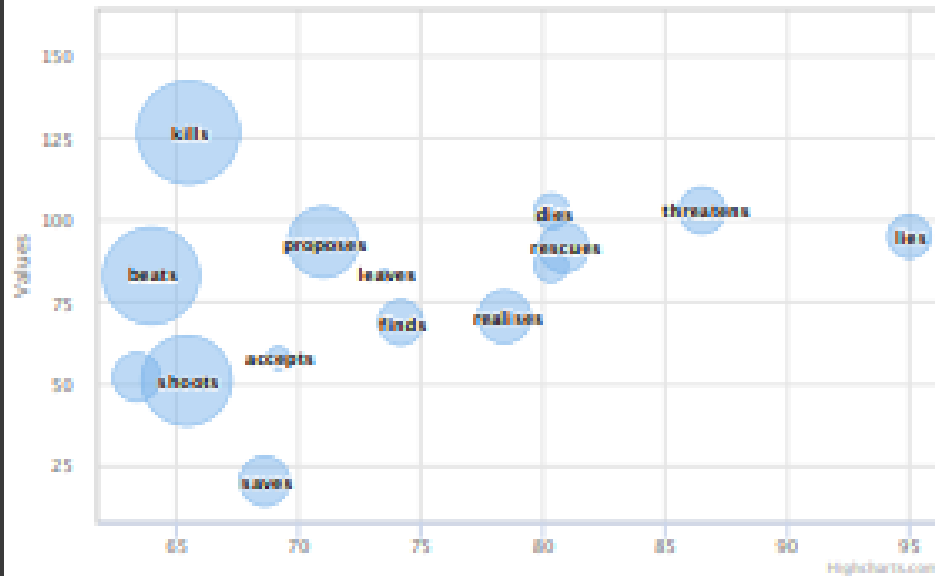
adopted ailing arrogant aspiring  
**attractive** **beautiful** beloved  
conservative **dead** drunk dumb elder happy illegitimate impoverished  
impulsive independent innocent insane intelligent lead local  
**lovely** married modern modest mysterious  
nagging perfect **poor** popular **pregnant**  
**pretty** reclusive reluctant **rich** secret sexy single  
spirited spoiled sweet talkative traditional unaware undercover  
unmarried wealthy **widowed** wrong

(b) Adjectives used with females

Highlighted words are more probable to the respective gender!

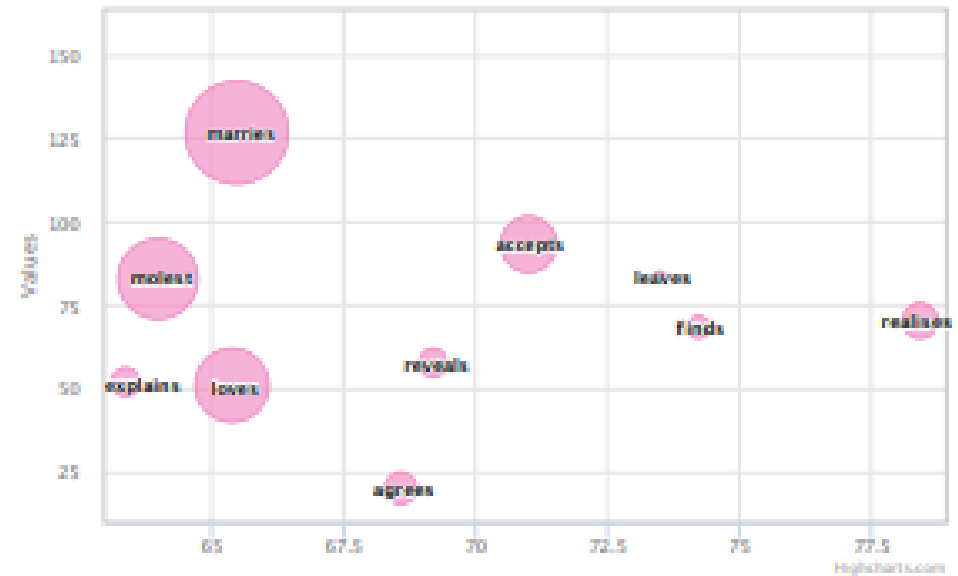
# Analyse the Stereotype: Actions!

Verbs used for MALES



(c) Verbs used with males

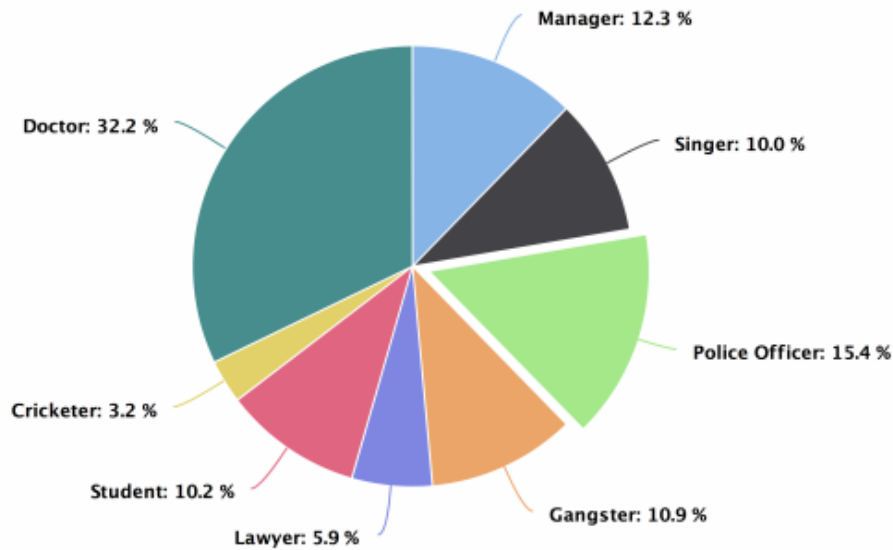
Verbs used for FEMALES



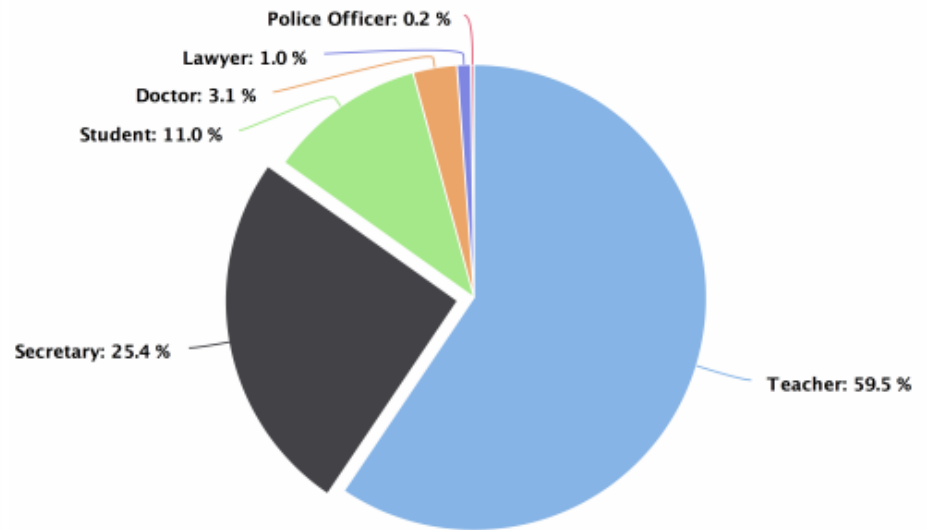
(d) Verbs used with females

Females are only considered to perform polite actions!

# Analyse the Stereotype: Occupations!



(e) Occupations used with males



(f) Occupations used with females

Females are only considered to perform polite actions!

# Dataset!

4000+ movies spanning the 1970-2017 time period.

Wikipedia plots with 5058 Female cast members.

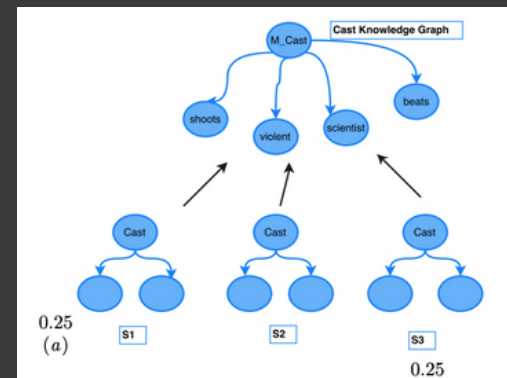
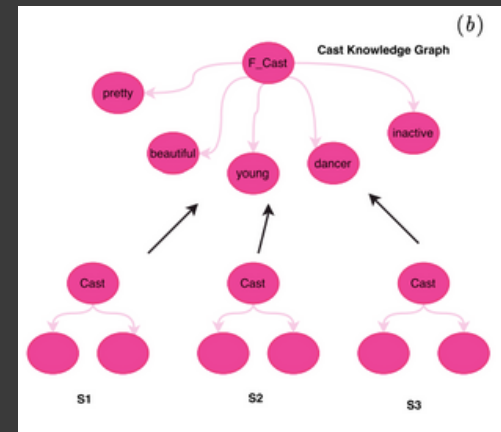
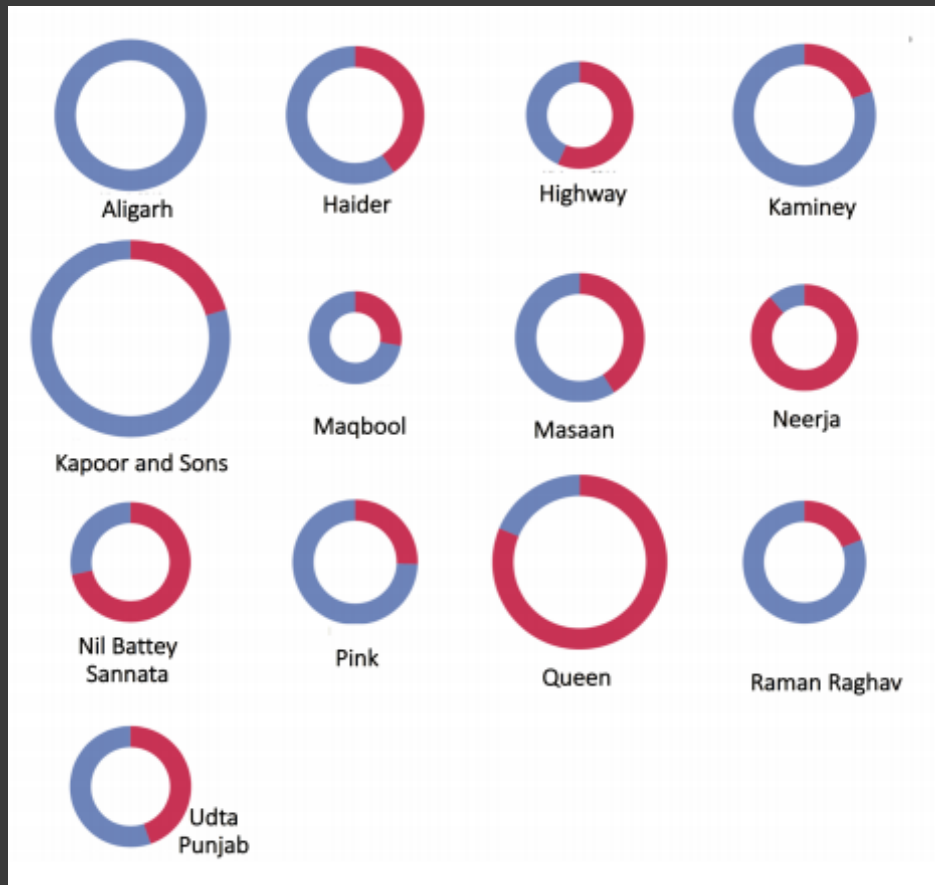
9380 Male cast members. 13 movie scripts. 880 movie trailers released for movies between 2008-2017.

Data set can be found at -

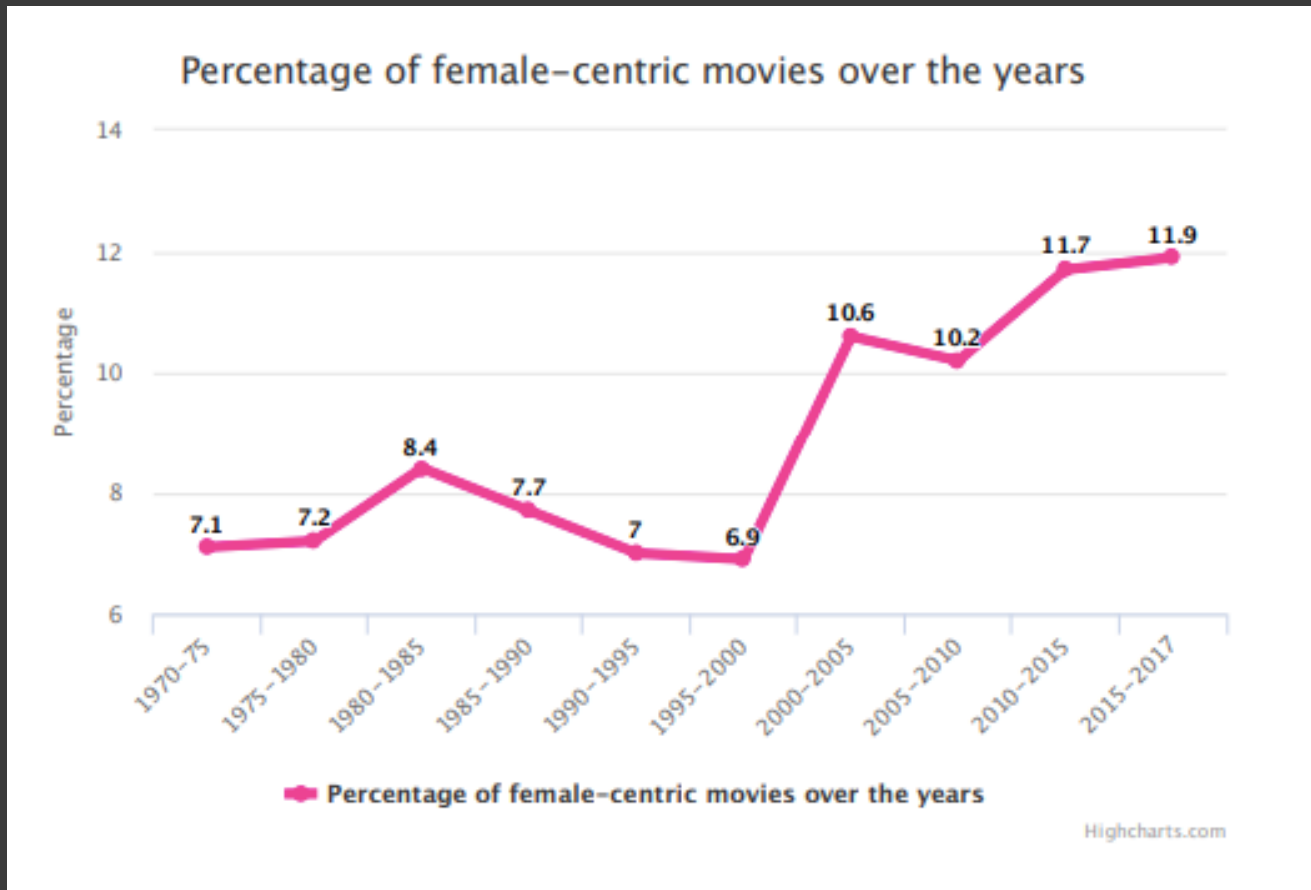
<https://github.com/BollywoodData/Bollywood-Data>

---

# Biassing Witnessed In Bollywood Movies!



# Biasing Witnessed In Bollywood Movies!



The growth is less than 5 percent!



# De-CogTeller: Bias Removal System !

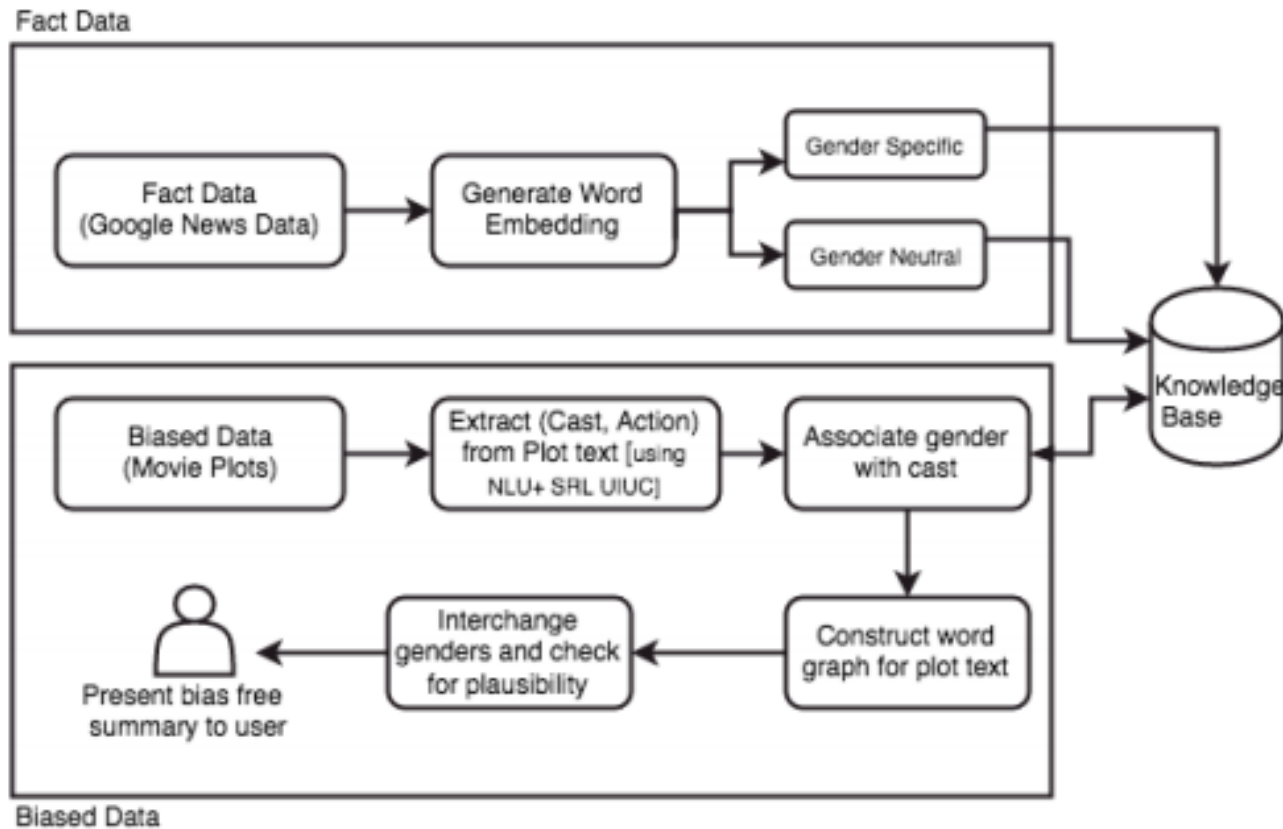
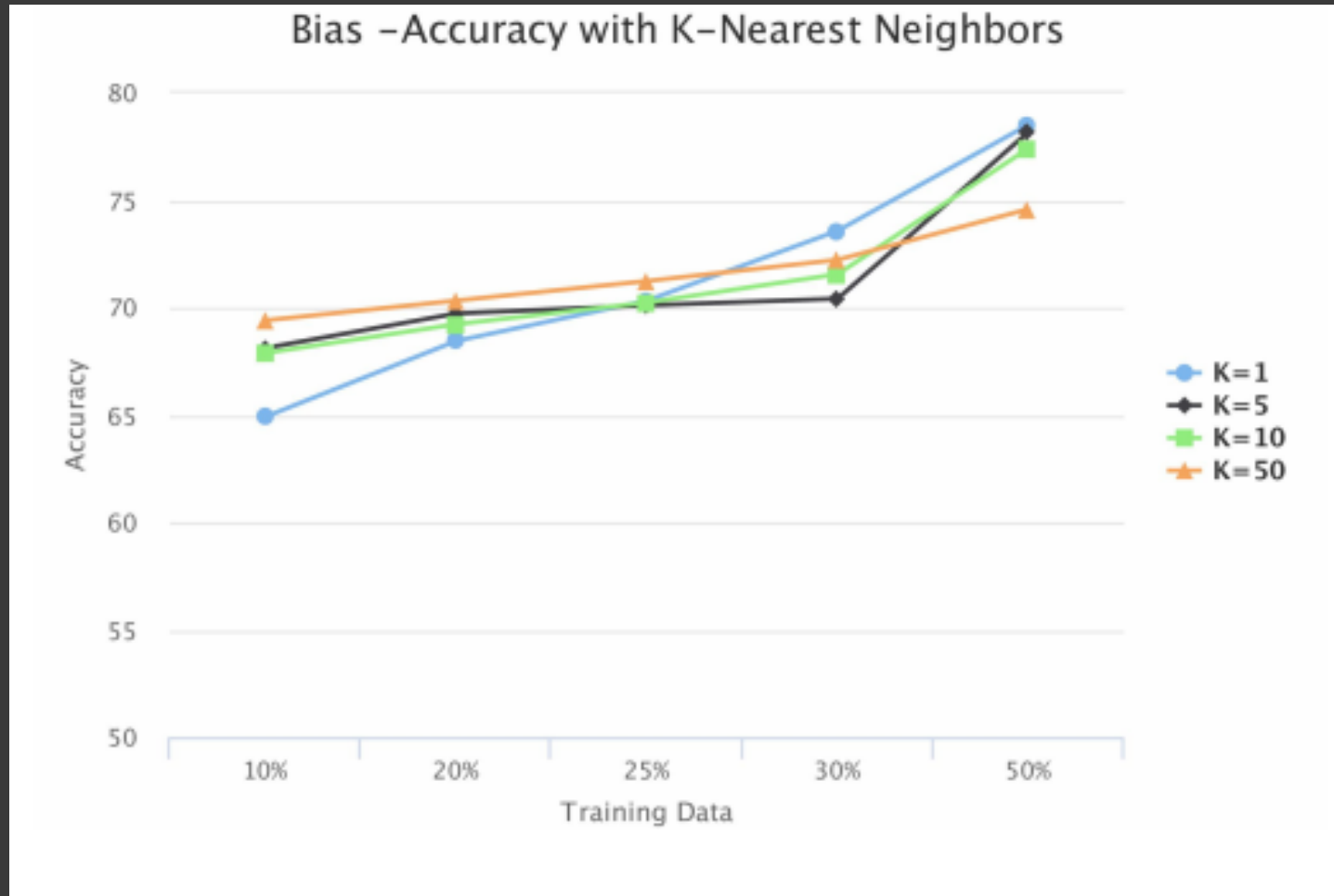


Figure 13: DeCogTeller- Bias Removal System

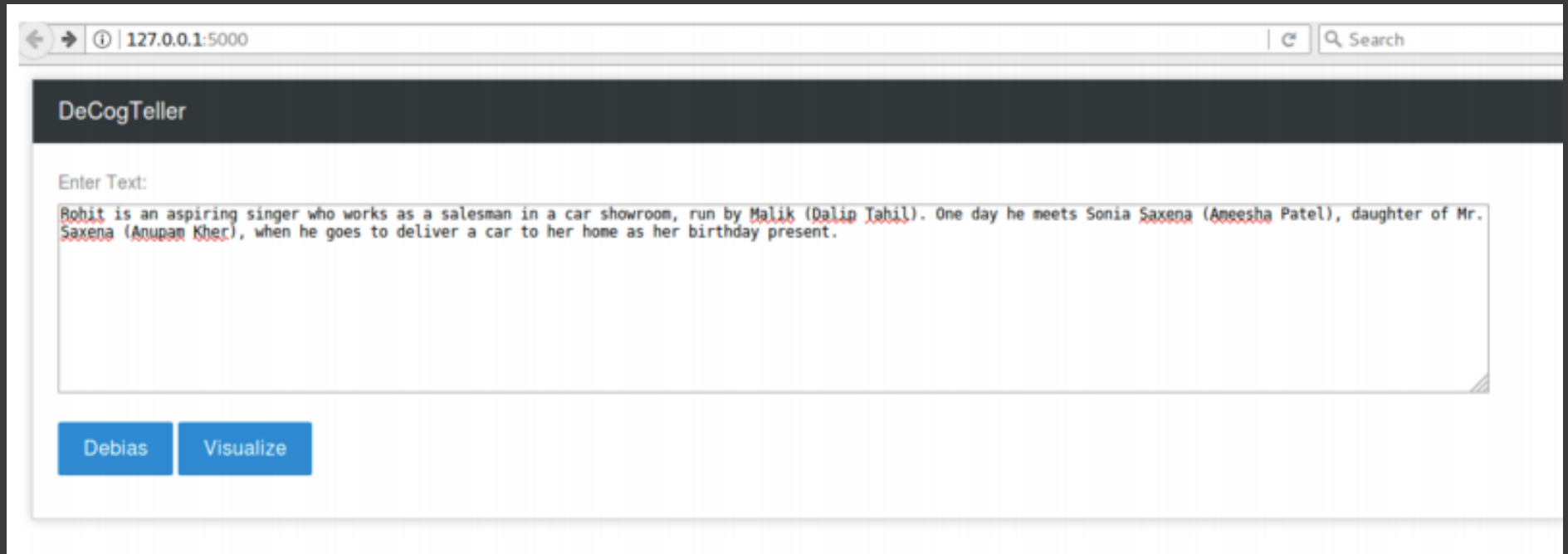
# De-CogTeller: Algorithm!

- Data Pre-processing
  - Generating word vectors
  - Extractions of analogical pairs
  - Classifying word pairs
  - Action Extraction from Biased Movie Data
  - Bias detection using Actions
  - Bias Removal
-

# Gender Prediction using Word Embeddings



# De-biasing System Testing



The screenshot shows a web browser window with the address bar displaying "127.0.0.1:5000". The page title is "DeCogTeller". Below the title, there is a text input field labeled "Enter Text:". The input field contains the following text: "Rohit is an aspiring singer who works as a salesman in a car showroom, run by Malik (Dalip Tahl). One day he meets Sonia Saxena (Aneesha Patel), daughter of Mr. Saxena (Anupam Kher), when he goes to deliver a car to her home as her birthday present." Below the input field, there are two blue buttons: "Debias" and "Visualize".

# De-biasing System Result

127.0.0.1:5000/debiased\_text/

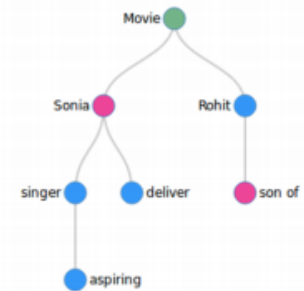
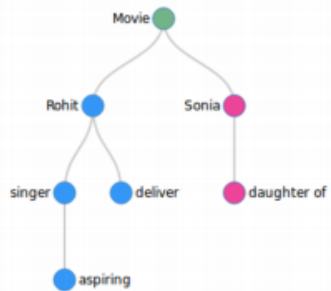
DeCogTeller

Debiased Text

<Rohit.he> is an aspiring singer who works as a salesman in a car showroom, run by Malik (Dalip Tahil). One day he meets <Sonia.she> Saxena (Ameesha Patel), daughter of Mr. Saxena (Anupam Kher), when he goes to deliver a car to her home as her birthday present.

<Rohit.she> is an aspiring singer who works as a saleswoman in a car showroom, run by Malik (Dalip Tahil). One day she meets <Sonia.he> Saxena (Ameesha Patel), son of Mr. Saxena (Anupam Kher), when she goes to deliver a car to his home as his birthday present. .

Visualize



# References

- Analyze, Detect and Remove Gender Stereotyping from Bollywood Movies Authors - Nishtha Madaan, Sameep Mehta, Tanea Agrawaal, Vrinda Malhotra, Aditi Aggarwal, Yatin Gupta, Mayank Saxena (Accepted in FAT\* and presented in Feb 2018 at New York University, USA)
  - <https://papers.nips.cc/paper/6228-man-is-to-computer-programmer-as-woman-is-to-homemade-r-debiasing-word-embeddings.pdf>
  - <http://www.cs.uml.edu/~arum/publications/Potash2017tracking-biased-news.pdf>
-

# Q and A



- Twitter: Sakshi\_ai
  - LinkedIn: sakshi-ai
  - Gmail: sakshishukla1996
  - Youtube: readtoreap
  - Mail: sakshi@wimlds.org
  - Facebook: readtoreap
  - Medium: readtoreap
-

Thank You

BERLIN BUZZWORDS 2019

---