Exploiting Alternative Datasets to Identify Gender Bias in Doctor-Patient Communication

Sonam Gupta

PhD Candidate at Harrisburg University

Gender Bias Exists!

Types of bias

- Conscious
- Unconscious

Why is it important to identify gender bias?

- Lack of women subjects in clinical trials
- Different symptoms based on gender lead to misdiagnosis
- No detailed medical investigation for female patients
  - Ex.: drug prescriptions may not be suitable

PhD Research Question (Still evolving)

- Identify gender bias in doctor-patient communication
  - Is the bias more prominent towards one gender?
  - What do patients think of their physicians?
  - Can language of their communication help identify gender bias?

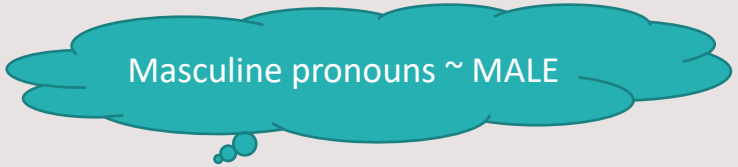
Patient's Perception of Gender Bias

- Online reviews data from ZocDoc.com
- Do patients' biases influence reviews for their physicians?
- Any identifiable patterns in the language to detect gender bias

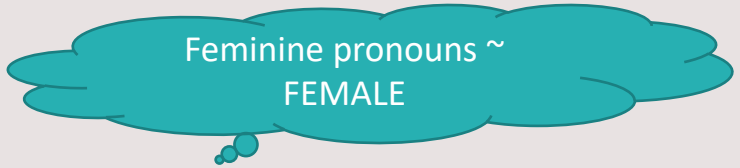
Reviews
A is amazing ! She made me feel totally comfortable and answered all my questions. I was in and out. No wait time. The medical assistant at the desk was also very sweet. I definitely will come back if I ever have to. :)
He is extremely professional and knowledgable. He doesn't waste time and diagnosed my situation very quickly. Dr. S explained the problem wrote a prescription (which he called in himself) and made a followup appointment - very impressive. I would recommend Dr. S without any hesitation.
Dr. A was great - she took the time to explain my diagnosis and her associate (admin/nurse) was very friendly and helpful as well.

Table 1: Sample patient reviews

Labeling Gender of Doctors



He is extremely professional and knowledgable. He doesn't waste time and diagnosed my situation very quickly. Dr. S explained the problem wrote a prescription (which he called in himself) and made a followup appointment - very impressive. I would recommend Dr. S without any hesitation.



Dr. A was great - she took the time to explain my diagnosis and her associate (admin/nurse) was very friendly and helpful as well.

More about the data

- Online reviews data
  - Data columns: Ratings (*OverallRating*, *BedsideMannerRating*, *WaitTimeRating)* and text reviews
  - 19,372 text reviews; 555 doctors; 2008 to 2015
  - 214 female doctors, 341 male doctors

Ratings (out of 5)	Mean	Standard Deviation
Overall	4.66	0.89
Bedside Manner	4.7	0.77
Wait Time	4.34	0.83

Table 2: Basic statistics on numerical ratings

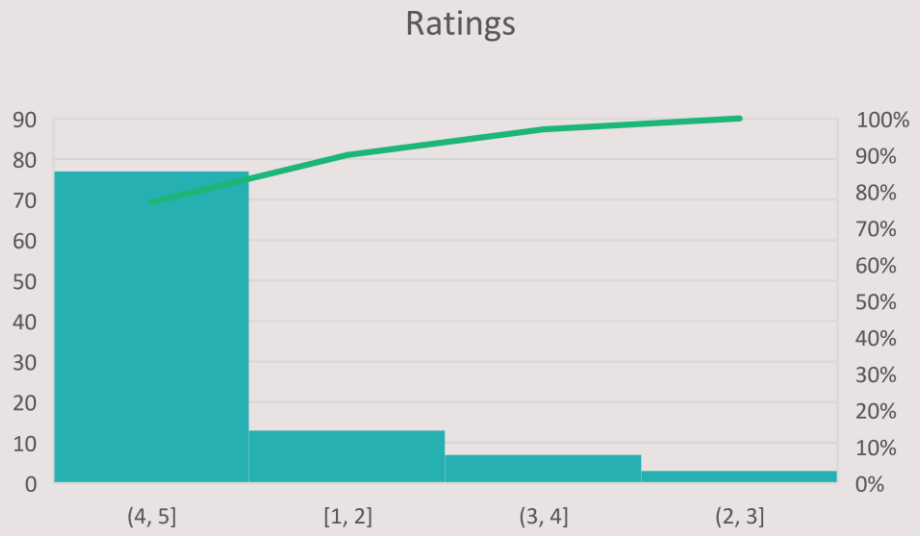


Fig 1: Popular ratings

Ratings vs Text Reviews

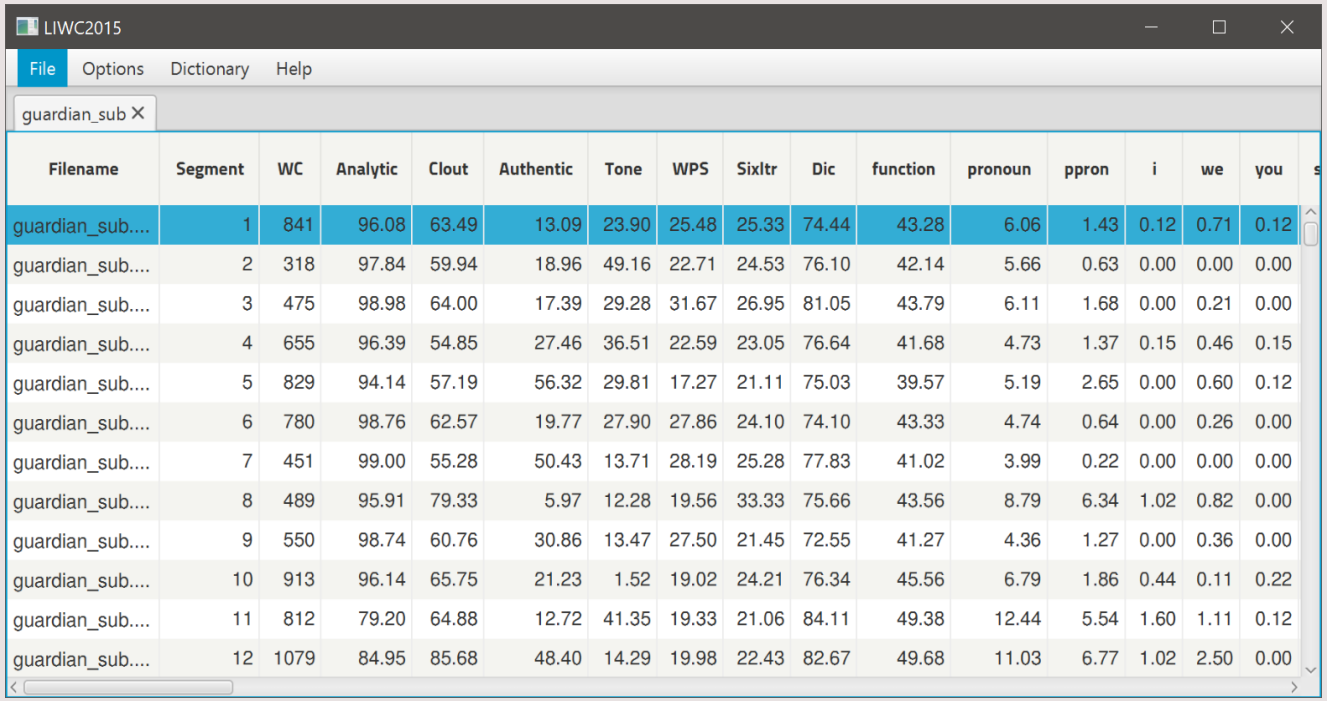
Insufficient information from numerical ratings about patient experience

Text reviews are more elaborated

Language analysis of text reviews

Text Reviews Analysis using LIWC

- LIWC (Linguistic Inquiry and Word Count) – A computerized text analysis tool



Filename	Segment	WC	Analytic	Clout	Authentic	Tone	WPS	Sixltr	Dic	function	pronoun	ppron	i	we	you
guardian_sub....	1	841	96.08	63.49	13.09	23.90	25.48	25.33	74.44	43.28	6.06	1.43	0.12	0.71	0.12
guardian_sub....	2	318	97.84	59.94	18.96	49.16	22.71	24.53	76.10	42.14	5.66	0.63	0.00	0.00	0.00
guardian_sub....	3	475	98.98	64.00	17.39	29.28	31.67	26.95	81.05	43.79	6.11	1.68	0.00	0.21	0.00
guardian_sub....	4	655	96.39	54.85	27.46	36.51	22.59	23.05	76.64	41.68	4.73	1.37	0.15	0.46	0.15
guardian_sub....	5	829	94.14	57.19	56.32	29.81	17.27	21.11	75.03	39.57	5.19	2.65	0.00	0.60	0.12
guardian_sub....	6	780	98.76	62.57	19.77	27.90	27.86	24.10	74.10	43.33	4.74	0.64	0.00	0.26	0.00
guardian_sub....	7	451	99.00	55.28	50.43	13.71	28.19	25.28	77.83	41.02	3.99	0.22	0.00	0.00	0.00
guardian_sub....	8	489	95.91	79.33	5.97	12.28	19.56	33.33	75.66	43.56	8.79	6.34	1.02	0.82	0.00
guardian_sub....	9	550	98.74	60.76	30.86	13.47	27.50	21.45	72.55	41.27	4.36	1.27	0.00	0.36	0.00
guardian_sub....	10	913	96.14	65.75	21.23	1.52	19.02	24.21	76.34	45.56	6.79	1.86	0.44	0.11	0.22
guardian_sub....	11	812	79.20	64.88	12.72	41.35	19.33	21.06	84.11	49.38	12.44	5.54	1.60	1.11	0.12
guardian_sub....	12	1079	84.95	85.68	48.40	14.29	19.98	22.43	82.67	49.68	11.03	6.77	1.02	2.50	0.00

Fig 2: LIWC

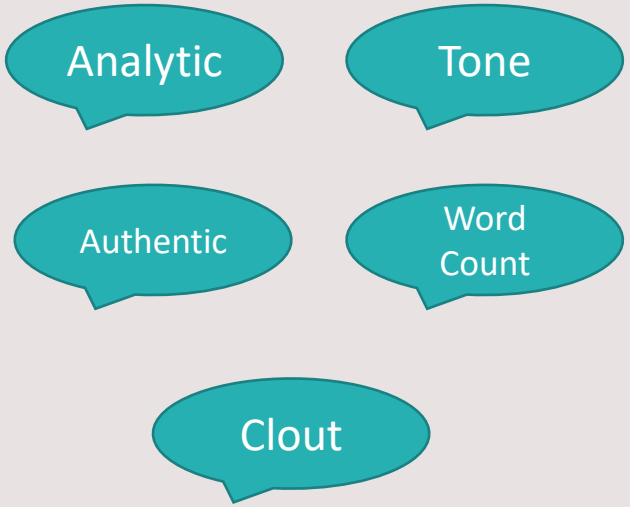


Fig 3: Examples of LIWC categories

Category 1: Informality

- Informality Metric:
  - Analytic thinking:-
    - High analytic score: more use of abstract language
    - Low analytic score: more use of personal language
  - Informal:-
    - High informality score: more use of pronouns, auxiliary verbs and adverbs
    - Low informality score: more use of prepositions, articles
- Expected higher analytic and lower informal for male doctors

Category 2 : Socio-Emotional Content

- Social Metric
  - Social – More use of social words like friend, family etc.
- Emotional Metric
  - Tone, positive & negative emotion - Emotional reaction or tone
- Expected higher in Tone and emotional content for female doctors

Category 3 : Gendered References

- Gendered References Metric:
  - Female: more use of words like ‘girl’, ‘lady’, ‘woman’ etc.
  - Male: more use of words like ‘boy’ , ‘man’ etc.
- Expected higher female references for female doctors

Statistical Analysis (LIWC ~ Doc Gender)

- Cohen's d calculations for gender differences
  - Standardized mean differences
  - Confidence intervals
- Hierarchical Linear Modeling Regression
  - DV ~ Ratings, LIWC scores
  - IV ~ Gender of the doctors

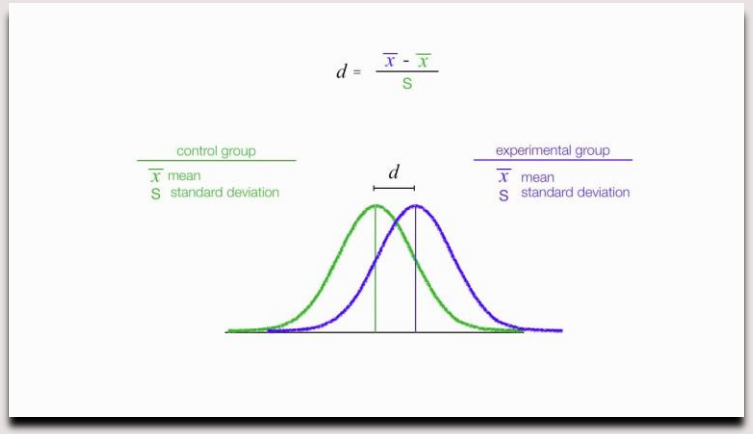


Fig 4: Cohen's d

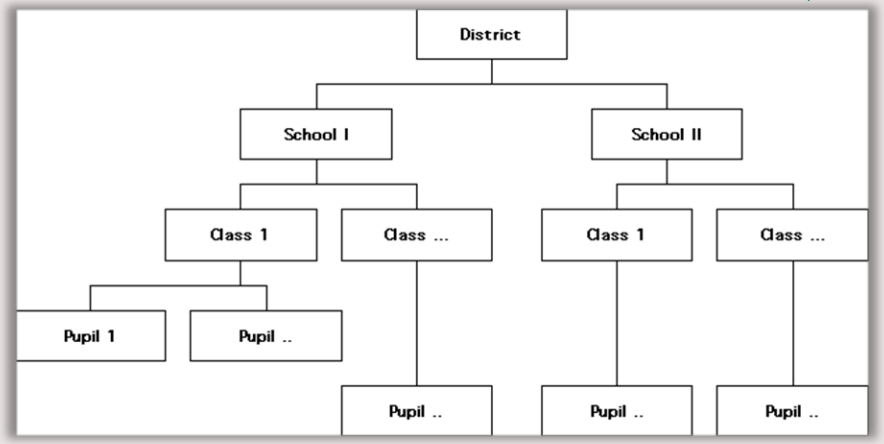
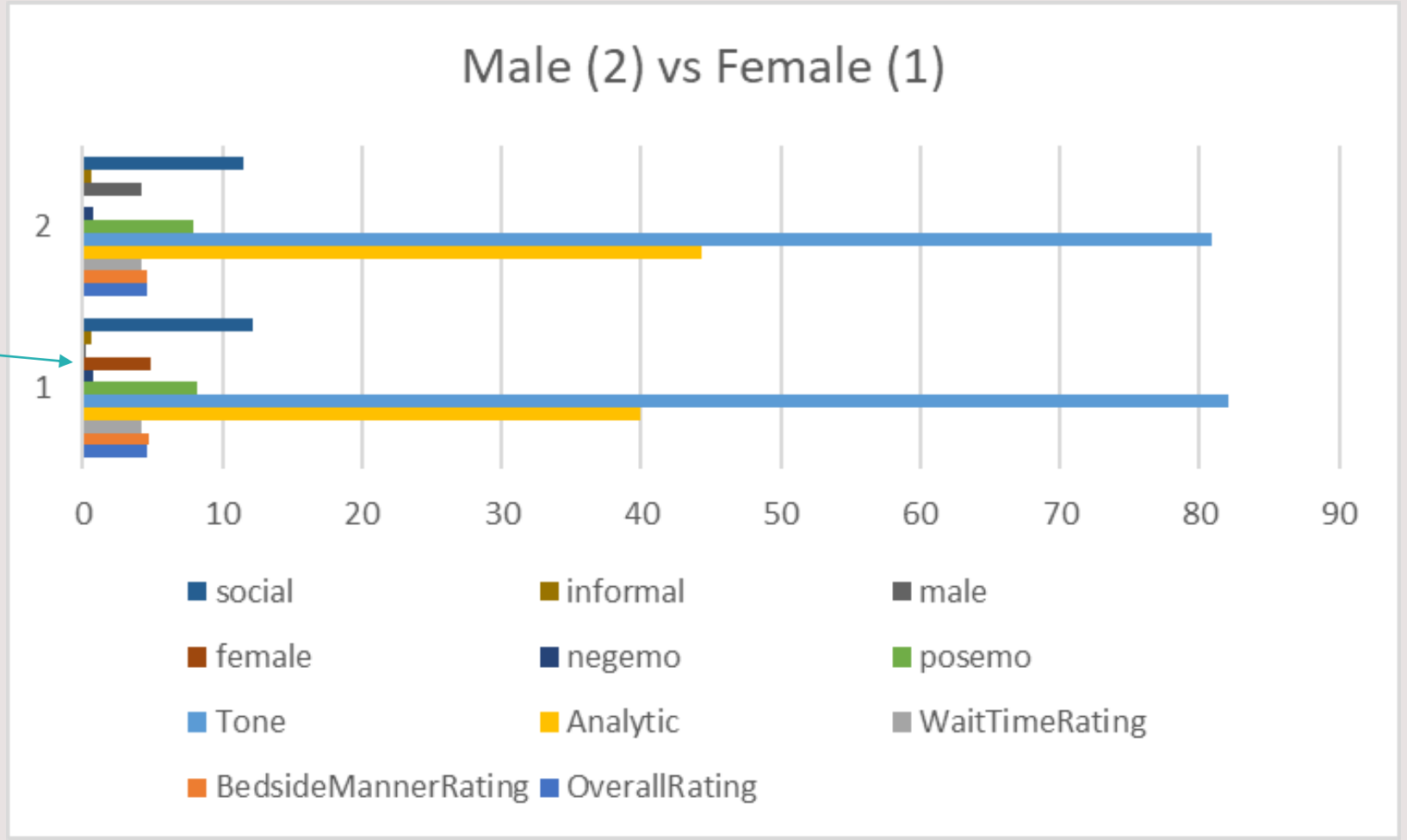


Fig 5: HLM model

Results

Female gendered reference is high



X-axis : scale of Means
Y-axis : 1 = female doc
2 = male doc

Fig 6: Means for male and female

Results (Contd.)

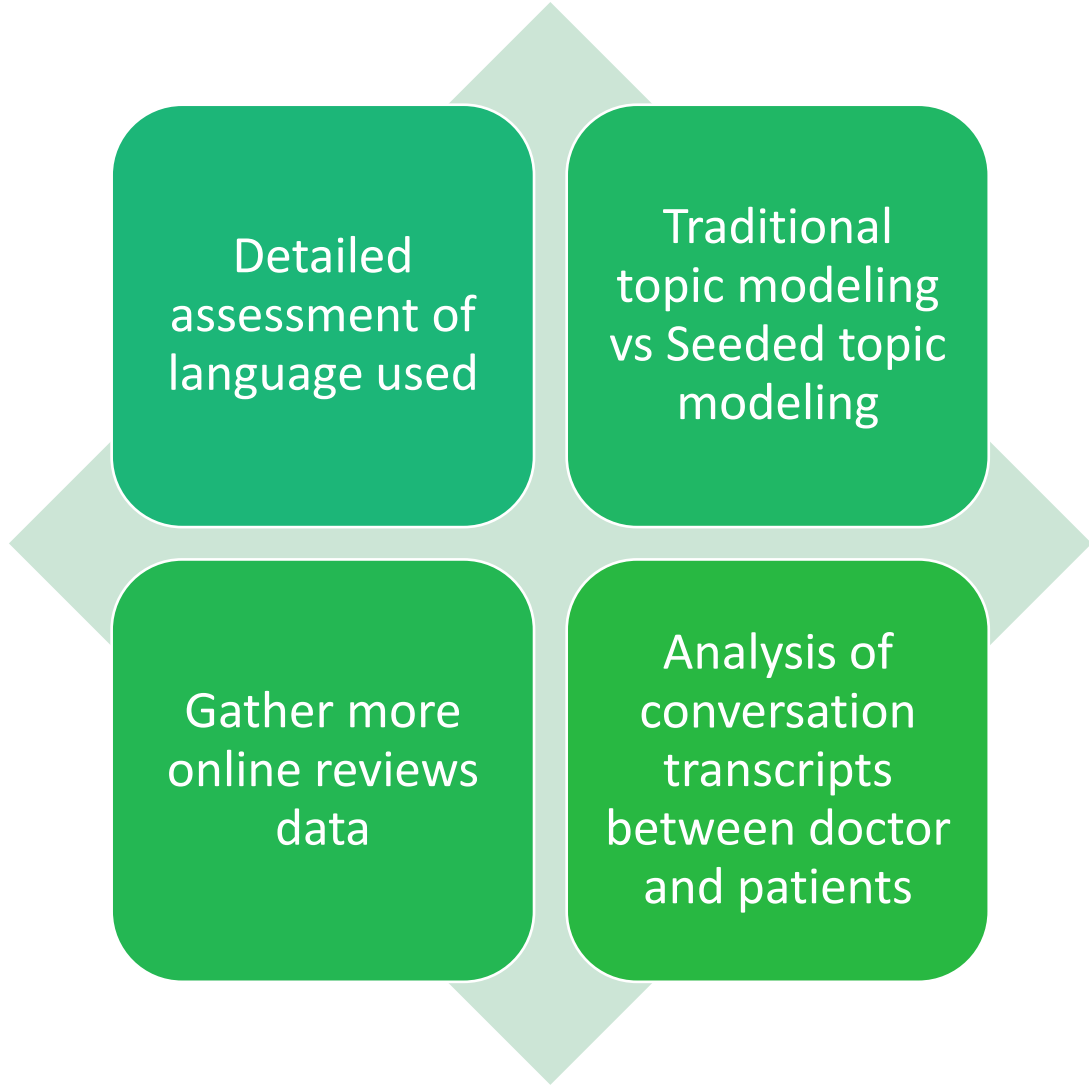
Dependent Var	Mean_female	StdDev_female	Mean_male	StdDev_male
OverallRating	4.59	2.46	4.58	2.62
BedsideMannerRating	4.66	2.15	4.63	2.36
WaitTimeRating	4.22	2.92	4.26	2.90
Analytic	39.96	49.32	44.35	48.31
Tone	82.04	63.65	80.84	61.18
posemo	8.19	11.46	7.87	11.08
negemo	0.70	3.12	0.71	3.22
female	4.85	10.56	0.13	2.62
male	0.28	5.56	4.24	8.49
informal	0.59	2.45	0.56	2.45
social	12.22	13.66	11.57	12.64

Table 3: Means & SDs for male/female doctors from Cohen's d

Conclusion

- Text analysis > numerical ratings
- Language usage → reveal gender bias
- Patients may have unconscious biases
- Analysis need more refinement

Future Work


Detailed assessment of language used
Traditional topic modeling vs Seeded topic modeling
Gather more online reviews data
Analysis of conversation transcripts between doctor and patients

Thank you!