Muhammad Rakibul Islam

Professor Lisa Rhody

DATA78000

22 May 2023

An Attempt to Create a Feminist Model for Computational Text Analysis

At the beginning of this course, we were asked the question “Can there be such a thing as feminist text analysis”. Throughout the course, we apply feminist inquiry and build a critique at each of the five stages of the text analysis process as identified in our first reading *How we do things with words: Analyzing text as social and cultural data* [1]. The five stages are “research questions, data selection, conceptualization, and operationalization, and analysis and the interpretation of results”.

In this paper, we will modify our question as “can there be a feminist model for text analysis”, apply several concepts to build a theoretical model for a particular use case and then critique that again through a feminist lens in an attempt to outline insights for further research.

But why should we even care about this in the first place? Why is a feminist text analysis needed? To understand that we draw insight from two sources: the book *Raw Data Is an Oxymoron*, edited by Lisa Gitelman [2] and the chapter *What Get Counted Counts* from *Data Feminism* by D’Ignazio and Klein [3].

What the book *Raw Data Is an Oxymoron* tells us is that essentially there is nothing as “raw” data which are rather always “cooked”. In simple terms, since data is simply a form of representation of the world, even in its fundamental form it already contains the biases, constructing and perpetuating existing inequalities and power dynamics present in the real

world. Also, data always needs human labor to be able create, collect, interpret, and analyze it which ends up translating more of the human bias in the process.

From this, we understand that data is not objective and is itself biased but why does that matter? The chapter from *Data Feminism* mentioned above answers that. Data is being used increasingly in the world to impact decisions whether that is in industry, education, public policy etc. Since data is biased and perpetuates the existing inequalities and power dynamics present in the world, the conclusions we draw from the analysis of such data and the consequential decision making that follows also inherits the same issues. As such “What Gets Counted” in the data literally “Counts” because it has real world implications.

What is the goal of a feminist model for text analysis in the first place? The book *The Resisting Reader* by Judith Fetterley [4] provides us with concepts that can be applied to our use case. She outlines that "the first act of the feminist critic must be to become a resisting reader rather than an assenting reader ... to begin the process of exorcising the male mind that has been implanted in us." From this we can derive that to build a feminist model, it must be able to detect the dominant, oppressive messaging and narratives that are embedded in a text which can allow us “resist” or challenge them – this defines the goal of our model.

To build on this further, I like the following quotes from the book *Intersectionality* by Collins and Bilge [5] “The events and conditions of social and political life and the self can seldom be understood as shaped by one factor” and they “are better understood as being shaped not by a single axis of social division … but by many axes that work together and influence each other” What this tells us is that in intersectional feminism, we have to recognize not just one social factor that influences and progresses discrimination but a multitude of them

(like gender, race, ethnicity etc.) as well as understand how they interplay, interact and build on each other. As such, any feminist model we create for CTA must incorporate these ideas.

Applying this intersectional approach, a feminist model for text analysis needs to be not a unidirectional & scalar one but a multi-directional & spatial one where the framework itself can mimic the complex, interconnected, multi-dimensional nature of intersectional feminism.

In our case, the word “data” can be replaced by “text” and all our previous discussions will still hold as in computation text analysis (CTA), text itself is the data being analyzed. But what guideline do we have in converting text to a form that computers can understand? Do we just think of random text or its unit of words as singular, unrelated entities? From Jerome J. McGann’s book *The Textual Condition* [6] and our following use of his ideas in our weekly notebook, we have identified that texts exist in social, historical, and cultural context and as such, the units of text or words themselves would also apply the same idea. From this, our model’s treatment of text must be contextual or in a term applicable to CTA “context aware”.

Applying these ideas discussed so far to our attempt at creating a feminist model for studying text, we arrive at the following guidelines that must be met. That our model needs to be context aware, able to detect not just one or few but a multitude of social factors (at varying degrees – say a probabilistic rather than a nominal representation) that exists within the text and be a multi-directional & multi-dimensional model (feminism as a spatial mathematical concept) with scope for intersectional relation recognition abilities.

Since building a completely new model is not feasible, we will try to find one that is already available that can fit these guidelines. Methods like LDA can decode topics but is not context aware. Meanwhile using TF-IDF with LinearSVC and other such models can consider

data in higher planes but not context aware. Earlier word embedding like Word2Vec are vectorized but also not context aware.

As such, context-aware word-embedding models such as BERT can be utilized. BERT can differentiate between the meanings in which a word is used. Using BERT, we could potentially identify the context in which different social factors of influence such as gender, race, ethnicity etc. are used with a probabilistic output (to recognize things aren't always a strict classification). And being multi-dimensional, we can perform intersectional analysis, understanding the relationships between these different cases. With its sentiment analysis and opinion mining capabilities, we can uncover those expressed in the text in relation to the different social factors. Given the model's customizability, it can be trained to the specific domain. So, BERT can fit with the initial guideline we set.

So, have we solved the problem? Did we figure out the holy grail of building a feminist model for CTA? Let's dig deeper. From both Richard Jean So's *All Models Are Wrong* [7] and Nan Z Da's *The Computational Case against Computational Literary Studies* [8] we have read how CTA models usually end up creating a rather simplified interpretation, neglecting the depth and nuances of language, context, meaning and the overall humanistic value humans perceive. Sure, BERT may be better at detecting that words are being used in a different context, but it cannot grasp the humanistic, social and cultural context that texts are situated in.

Furthermore, as we learn from Meredith Broussard's *Artificial Unintelligence: How Computers Misunderstand the World* [9], that AI models themselves are first trained on a provided dataset which itself can introduce further bias both inherent in the data and from the humans who prepare, clean and encode the data into the model used.

What if we could produce a training dataset and build the model in such a way that all context, representation etc. are utilized? While in a rather fictional, theoretical world that might be true, unfortunately, since we are humans ourselves with our own set of biases, using data and AI models that are also biased, the reality is that we cannot truly ever completely escape this vicious cycle. Also, even if that did somehow work, such a model would be so vast that it would require an unfathomable amount of computational and human resources, making it highly inaccessible and not environment friendly, which itself are feminist concerns.

As ironic as it may be, any attempts to create a feminist model for CTA will itself continually give rise to feminist concerns and consequently the need for a feminist critique in a cycle that can never be fully perfected. Deviously, it will always churn out new obstacles to be assessed and improved with every attempt we make.

But does that mean all hope is lost? No! Perfection is something that rarely, absolutely exists. It goes against the very rubric, the fundamental law of nature itself to have a system void of any chaos. But in attempts to make things better, we can achieve better models, more accurate representations and a continually improving world of CTA which is still in its infant stages to attempt to encapsulate the breadth of the human experience itself. Basically, the conversations need to be ongoing, critique must be made at every stage and in cycles with proper documentation about limitations and potential for bias in every step – so that others can use it and build upon it. After pursuing an answer to the question from multiple angles, I reach the conclusion that a feminist text analysis is not an absolute destination, but a perpetual process that reiterates and builds on itself through revisions, rectifications, and enrichment.

References

[1] Nguyen, Dong, Maria Liakata, Simon DeDeo, Jacob Eisenstein, David Mimno, Rebekah Tromble, and Jane Winters. “How we do things with words: Analyzing text as social and cultural data” arXiv:1907.01468v1 [cs.CL] 2 Jul 2019

[2] Gitelman, Lisa. “Raw Data” Is an Oxymoron. MIT Press, 2013

[3] D’Ignazio, Catherine and Lauren Klein. “What Gets Counted Counts.” Data Feminism. Cambridge, MA: MIT Press, 2020

[4] Fetterley, Judith. “Introduction & Chapter 1.” The Resisting Reader: A Feminist Approach to American Fiction, Indiana Univ. Press, Bloomington, 1989

[5] Collins, Hill Patricia, and Sirma Bilge. “What Is Intersectionality?” Intersectionality, Polity Press, Cambridge, UK, 2018

[6] Jerome McGann “Introduction: Texts and Textualities” “The Textual Condition” and “How to Read a Book” McGann, Jerome J. The Textual Condition. Princeton, N.J.: Princeton UP, 1991. Print. Princeton Studies in Culture/power/history

[7] So, Richard Jean. “All Models Are Wrong.” PMLA : Publications of the Modern Language Association of America, vol. 132, no. 3, 2017, pp. 668–73. cuny-gc.primo.exlibrisgroup.com, https://doi.org/10.1632/pmla.2017.132.3.668

[8] DA, Nan Z. “The Computational Case against Computational Literary Studies.” Critical Inquiry. v. 45, n.3, Mar 2019, pp 601-39. journals.uchicago.edu (Atypon), https://doi.org/10.1086/702594

[9] Broussard, Meredith. Artificial Unintelligence: How Computers Misunderstand the World. The MIT Press, 2018