Democrats are to Republicans as Batman is to Wolverine? Detecting and Removing Political Biases in Word Embeddings

Julia Buffinton, April Kim, Jennifer Podracky

August 10, 2018

Abstract

Text data used for machine learning and natural language processing contain the biases carried by the human authors of that text. Thus, blindly generating word embeddings from this biased data will result in applications that have bias built-in. Inspired by Bolukbasi et al.'s 2016 paper, "Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings," [1] in which they define metrics to quantify gender bias in embeddings and propose algorithms to remove that bias, we extend their approach to political bias. We apply their approach to word embeddings trained on two different corpora to replicate and bolster their results using both their dataset and a dataset of recent news articles, and then apply this approach to political party bias in both texts. Using human judgments and the gender bias results as benchmarks, we demonstrate that this approach can also identify and remove complex biases such as political bias. These debiased embeddings can then be used in a variety of traditional NLP applications without incorporating and amplifying the bias of the original embeddings.

1 Introduction

Word embeddings are a commonly used tool in natural language processing to turn text-based content into a set of features that can be fed to any number of powerful machine learning algorithms. The process of turning English words into vectors of numbers (referred to as "embeddings") requires a large body of text to be fed into a machine, which then determines how many unique words exist in its vocabulary. The machine then assigns each word a unique vector representation in such a way that words that appear in similar contexts are located near each other in vector space. Another feature, which we'll make frequent use of in this paper, is that the distance between these word embeddings represents relationships between words. This concept is exciting to AI and machine learning researchers, as this improves performance in such areas as machine translation.

Unfortunately, however, the word embeddings themselves are very dependent on the body of text that is used to generate those embeddings, and they are privy to any biases or prejudices that exist in those bodies of text. Previous research has addressed the existence of gender biases in these embeddings, as well as their negative impacts; for example, as noted in "Semantics derived automatically from language corpora contain human-like biases" [2], Google Translate translates Turkish sentences with gender-neutral pronouns ("O bir doktor. O bir hemşire.") to these English sentences: "He is a doctor. She is a nurse." What has not been addressed in detail is the existence and prevalence of other biases in word embeddings, in particular political biases, and how those built-in biases manifest themselves in embedding applications. In this paper, we present an approach to identify those biases and correct them.

2 Background

This paper was inspired by "Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings" [1], a paper published in 2016 by a group of researchers at Microsoft and Boston University. The focus of that work was to examine the machine-generated vector embeddings of words associated with gender, and to find word embeddings of words that should

Democrat	Republican
Working-class	Rich
Unionized	Non-Unionize
Younger	Older
Women	Men
Secular	Evangelical
African-American	White
LGBTQ+	Heterosexual
Northern	Southern

Table 1: Perceived characteristics of the Democratic and Republican parties, as written in [5].

not have included a gender influence but were geometrically represented in the gender subspace. The researchers then corrected and regenerated the word embeddings excluding those biases.

The above mentioned paper, along with a wealth of other research that addresses bias in machine learning and natural language processing [2, 3], focused solely on gender bias. However, there are numerous sources of bias that could be present, including, but not limited to: race, religion, sexuality, age, and political leanings (liberal/conservative). In this paper, we focus on political party bias, as it is relatively binary (democratic/republican) and less controversial than other types of biases.

2.1 Origins of Political Party Bias in the U.S.

Since the genesis of the Democratic and Republican parties as we know them today, the two opposing parties that form the foundation of the United States' political system have been drifting further and further apart; a 2017 Pew Research Center study found that the gap between Republican-leaning individuals and Democratic-leaning individuals grew from 15 percentage points in 1994 to 36 points in at the end of 2017 [4]. The controversy surrounding the 2016 U.S. presidential election has no doubt contributed; the same Pew study found that Donald Trump's job approval ratings are the most polarized of any first-year president dating back to Dwight D. Eisenhower in 1953 [4]. As the parties and their constituents become more and more polarized, it becomes more important to examine the stereotypes and caricatures being applied to members on both sides. This is especially important when, as a 2015 survey found, each side holds these stereotypes to be true for a much larger percentage than is truly applicable [5]. Some typical dichotomies between Democrats and Republicans generally seen in literature are listed in Table 1.

3 Methods

The following methodology was undertaken to understand the gender and political biases present in two examples of word embeddings, with the ultimate goal of using this knowledge to apply debiasing algorithms that remove the biases from each of the embeddings. We first pursue this for gender bias as a baseline, to replicate the finding in [1], and then extend it to political party bias.

3.1 Collecting Data

Two bodies of text were used to generate word embeddings for our research. The first corpus was from the Google News dataset, the same used in [1].

The second corpus is a series of news articles published across 337 different news sources from January to June 2018. This was collected by Boston company RANE (the employer of one of our researchers), using a web crawling application. The text data crawled and retrieved included the full HTML body, which included HTML tags, embeddings, and photo captions. A subset of sources crawled can be found in the Appendix.

3.2 Generating Word Embeddings

The Google News word embedding used in our baseline was sourced from the Google Code Archive[1]. This embedding was generated using word2vec and has 300 dimensions and a vocabulary of 26,423

[1] https://code.google.com/archive/p/word2vec/

unique words. This is the embedding used in [1].

The RANE News embeddings was also generated using word2vec; this algorithm was chosen to mitigate any potential differences that could be attributed to using a different embedding algorithm. The text was preprocessed to remove capitalization, punctuation, and special characters. The preprocessed text was then run through the word2vec algorithm for 3 iterations, outputting a 300-dimensional embedding and a vocabulary of 462,643 unique words. The embedding set was then filtered, to improve usability of the outputs and the performance of subsequent word relationship algorithms, to contain a vocabulary of only the top 50,000 English words[2].

3.3 Analyzing the Embeddings for Bias

To identify the presence of bias in our embeddings, we capitalize on one of the hallmarks of word embeddings: that the cosine similarity between two word vectors represents the semantic similarity of the corresponding words. This enables us to make inferences about the relationships between words based on where they are relative to each other and other words.

One application of understanding relationships between words is analogies. Using the embeddings of a set of three words, one can use simple vector math to find the most appropriate fourth word. For example, "*frog* is to *tadpole* as *cow* is to what?" could be represented as WE(result) = WE(*tadpole*) - WE(*frog*) + WE(*cow*), where WE represents the word embedding for the italicized word. Additionally, a small modification of this task enables us to start with two words and generate a second pair of words with a similar geometric relationship, and thus an inferred similar semantic relationship.

Figure 1: **Stereotypical** "Democrats-Republicans" analogies, automatically generated from the RANE News embedding.

anarchist-elitist	sisters-disciples	urbanism-modernist
bourgeoisie-despots	belinda-bobby	lotions-whitening
darwinism-determinism	midwife-surgeon	bourgeoisie-despots
synagogue-congregants	basketballer-newsman	unionism-reflation

These analogies were automatically generated using the RANE News word embedding. The first analogy in Figure 1 is interpreted as *democrats:republicans :: anarchist:elitist*. Each automatically generated analogy was evaluated by the researchers and determined to be stereotypical, definitional, or unrelated to the original word pair *democrats-republicans*. The most stereotypical analogies for both political parties are listed in Figure 1. In these analogies, different elements of the stereotypes listed in Table 1 (including race, gender, classism, and religion) are represented. We compare these stereotypical analogies with the definitional analogies, as seen in Figure 2.

Figure 2: **Definitional** "Democrats-Republicans" analogies, automatically generated from the RANE News embedding.

democratic-republican	democrat-republican	populist-protectionist

Another way to understand relationships between words is to find the group of classifiers, e.g. colors or professions, that are most related to a given word. By creating an axis between two words, or the two vectors that represent those words, one can see which classifiers are closer to one end of the axis than the other. This was done in [1] for the word pair *she-he* and professions; we repeated this examination with *democrats-republicans* and professions, using the RANE News embedding. Any occupations specifying an elected office (e.g., *president*) were excluded.

Interestingly, the biases discussed in the earlier section can also be found in the professions most closely related to either *Democrats* or *Republicans*. Under *Democrats*, we see two words related to working-class/blue-collar occupations: *worker* and *butler* (Figure 3). *Worker* is actually the most "poor" of the professions classified. We do see *minister* and *bishop*, two traditionally

[2] https://github.com/dwyl/english-words

Figure 3: The 15 Most Extreme "Democrats" Occupations, automatically generated from the RANE News embedding.

worker	ambassador	scientist	minister	artist
broadcaster	pollster	butler	bishop	councilor
psychologist	chancellor	sociologist	researcher	advocate

religious professions, even though this defies the stereotypes we've previously observed. We also see *advocate* under *Democrats*, which was the most "minority" profession. In contrast, under *Republicans*, we see several professions associated with upper class and high salaries: *salesman*, *banker*, *broker*, *surgeon*, and *physician* (Figure 4). We still see two religious professions, *monk* and *parishioner*. We see *salesman* and *patrolman*, two of the most "old" professions, as well as *negotiator*, one of the most "he" professions.

Figure 4: The 15 Most Extreme "Republicans" Occupations, automatically generated from the RANE News embedding.

salesman	bureaucrat	welder	assassin	physician
patrolman	parishioner	broker	gangster	fisherman
surveyor	banker	surgeon	negotiator	monk

By solely comparing the distances between political leanings and professions, we can see evidence of six different types of biases that are present in U.S. society. These professions align with the human-identified stereotypes (Table 1) for each political party.

3.4 Identifying Geometric Subspaces for Each Bias

To identify the subspace for a type of bias, we employed similar methods as in [1]. Using several pair difference vectors for each type of bias (*maleVector* - *femaleVector*, *democratVector* - *republicanVector*, etc.), we computed their principal components (PCs). The PC(s) of the difference vectors can be used to identify a bias direction $b \in \mathbb{R}^d$ that will be used to quantify bias in words and associations.

It is not necessary that this subspace is single-dimensional. Especially with a bias that is more complex, such as political party, it may be the case that multiple dimensions capture this subspace.

A decrease in the proportion of variance captured by each PC is expected; however, the PC(s) that capture the majority of variance in the pair difference vectors should be significantly larger than the rest of the components, indicating that they can be used to explain the bulk of the difference between word pairs, namely, the bias we are examining.

The following steps can be followed to apply these methods to any sort of bias:

1. Get 5-10 pairs of words that indicate definitional analogies from human-generated evaluations.
2. Generate a vector for difference between vectors in pair ('pair difference vector').
3. Compute PCs of the pair difference vectors.
4. Identify direction(s) that explain the majority of variance in these vectors.
5. Attribute this space as the bias subspace.

3.5 Debiasing

The subspace identified using the techniques above was used to "debias" the words that contain stereotypes. We elected to pursue the "hard" debiasing option proposed in [1], where they **neutralize and equalize**. First, this ensures that neutral or definitional words (e.g., "ballerina" and "father" and "conservative" and "progressive") remain zero in the subspace. Second, words outside of the subspace are equalized with respect to the word pairs used above. For example, if *democrats-republicans* was an equality set word pair, then after equalization, *surgeon* should be equally as close to *democrats* as it is to *republicans*. Depending on the amount of polysemy of the words in the bias attribute subspace, this may equalize words that should not equalized, but overall this approach seems to do more good than harm.

Once the subspace is identified, we used this space B, a list of words to neutralize ($N \subseteq W$), and a family of equality pairs ($\varepsilon = E_1, E_2, ..., E_m, E_i \subseteq W$) to complete the debiasing. Each word $w \in N$ is re-embedded to be:

$$\vec{w} := (\vec{w} - \vec{w}_B / \|\vec{w} - \vec{w}_B\|)$$

Additionally, for each set $E \in \varepsilon$:

$$\mu := \sum_{w \in E} w/|E|$$

$$v := \mu - \mu_B$$

$$\vec{w} := v + \sqrt{1 - \|v\|^2} \frac{\vec{w} - \mu_B}{\|\vec{w} - \mu_B\|}$$

3.6 Evaluation

After debiasing a set of word embeddings, the resulting embeddings need to be evaluated to confirm that the bias has been removed and ensure that the true, unbiased word meanings and relationships are still present.

To confirm that words have been correctly debiased, we tested that only stereotypical words have been debiased and definitional words remain unaffected. We must ensure that the bias attribute has not been removed from definitional words, and that definitional analogies have not changed, such as *democrats:liberalism::republicans::conservatism*. Then, we examined whether the bias component is removed from other stereotypical word embeddings.

Once the bias has been removed, we must then ensure that the debiased word embeddings maintain the same utility for applications that build upon the meanings of the word embeddings. This can be done as in [1], by examining how well the debiased word embeddings perform in applications that rely on their meaning, relative to their biased counterparts.

Two such applications with well-established benchmarks are similarity comparisons and analogy tasks. For each set of embeddings and bias type, scores were generated for performance on these tasks using standard benchmarks: Rubenstein and Goodenough (RG)[3] similarity pairs [7], Finkelstein et al. similarity pairs (WS)[4] [8], and Mikolov et al. analogies (M-Analogy)[5] [9]. Regardless of their absolute performance, the biased and debiased word embeddings should perform similarly on all of these metrics.

4 Results and Discussion

4.1 Identifying Bias Subspace

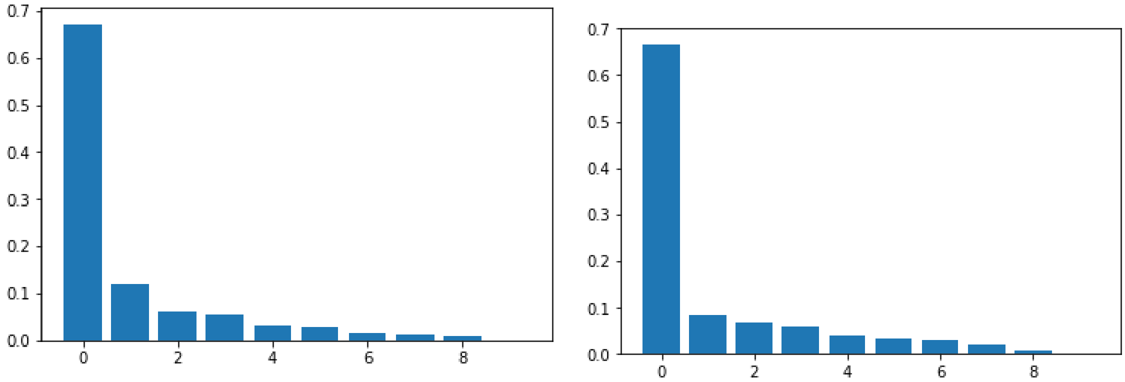
For both types of bias, PCA revealed one principal component that was used to identify the subspace of the bias type of interest. This was achieved by compiling sets of word pairs that occurred on each end of the extreme for the bias type (*she-he*, *democrats-republicans*, etc.), calculating a difference vector for each word pair, and identifying the PCs of the difference vectors.

It is clear that there is a single direction that explains the majority (over 60%) of the variance in the gender pair difference vectors for both word embeddings. This supports the hypothesis proposed by [1] that this top PC captures the gender subspace. See Figure 5 for the gender pair results.

Although we hypothesize that the variance in the political party pairs may be more complex, PCA reveals that one direction also explains the majority of variance in these difference vectors.

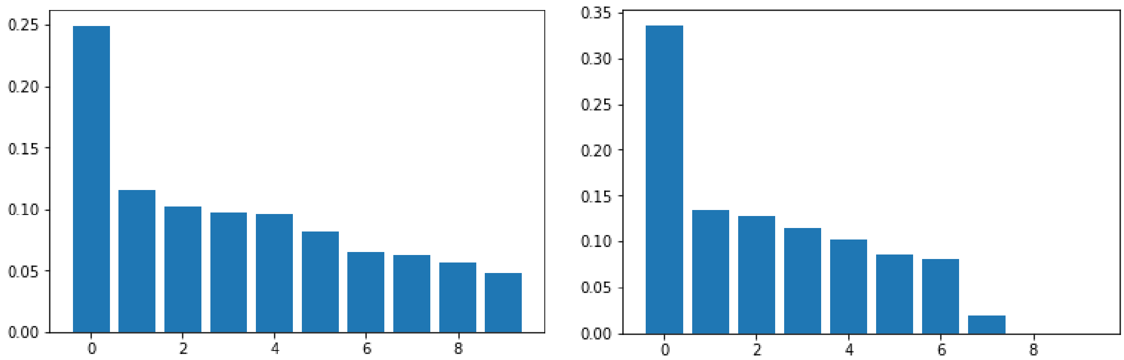
[3]https://github.com/mfaruqui/eval-word-vectors
[4]https://github.com/mfaruqui/eval-word-vectors
[5]http://www.fit.vutbr.cz/ imikolov/rnnlm/word-test.v1.txt

Figure 5: Percentage of variance explained by PCs of Gender pair difference vectors, using the Google News (left) and RANE (right) embeddings.



It is much less definitive than the PC for gender, with only approximately 25-35% of the variance explained by this vector. However, this is still significantly higher than other dimensions, much higher than if it was just a random set of vectors. See Figure 6 for political party pair results.

Figure 6: Percentage of variance explained by PCs of Political Party pair difference vectors, using the Google News (left) and RANE (right) embeddings.



4.2 Debiasing

Each word embedding was debiased separately against both biases of interest: gender and political party. A set of professions was plotted and scored along the bias axis, both on the original embeddings and on the debiased version of those same embeddings. The scores for each profession were compared between the original embedding and the debiased embedding. A summary of those score differences is show in Table 4.2. A random subset of professions is also plotted to the bias axes in Figure 7, and Figures 8, 9, and 10 in the Appendix.

These tables and figures show that the amount of bias, as interpreted by the distance from 0 on the bias axis, is reduced from the original word embedding to the debiased word embedding. Debiasing was more effective by both sum and average measures on the gender axis in both embeddings. Debiasing was least effective on the political party axis in the RANE News embedding (pictured in Figure 7.

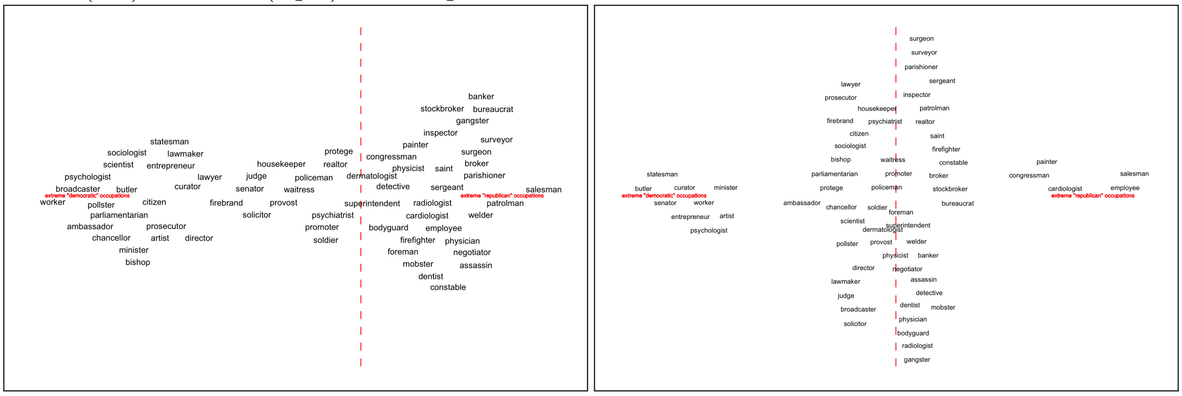
4.3 Evaluation

As shown in Table 4.3, performance of both the Google News and RANE embeddings on RG, WS, and M-Analogy tasks did not change significantly between the original embedding and the debiased versions of the same embedding. This suggests that the usability of these embeddings in other semantic tasks was not negatively affected by the embedding debiasing.

	Sum Bias Distance	Average Bias Distance
Google News, Original (gender)	25.65	0.09
Google News, Debiased (gender)	8.08	0.03
Google News, Original (political)	17.19	0.05
Google News, Debiased (political)	9.55	0.03
RANE News, Original (gender)	24.75	0.08
RANE News, Debiased (gender)	8.12	0.03
RANE News, Original (political)	14.24	0.05
RANE News, Debiased (political)	12.04	0.04

Table 2: Performance for both sets of embeddings on similarity and analogy tasks. Performance does not degrade significantly after either debiasing.

Figure 7: A distribution of professions on the Political Party axis using the RANE News embedding, before (left) and after (right) debiasing.



It is interesting to note that the overall performance on the analogy/word relationship tasks was lower for the RANE embedding than for the Google News embedding. We hypothesize that this is due to the quality/cleanliness of the RANE dataset, with numerous non-English and non-alphabetic words included when generating the embedding.

5 Conclusion and Next Steps

Through the research detailed above, we confirmed that it is possible to detect, identify, and remove a large amount of both gender and political party bias from word embeddings by building two embeddings off of separate text corpora and reducing the bias in those embeddings while retaining their usability for other NLP/ML applications.

However, we also observed that political party bias is much less cut-and-dry than gender, in terms of geometric subspace; while one dimension was enough to explain 65% of the gender variance, the first dimension of the political party pair difference only explained 25-35% of that

	RG (53)	WS (318)	M-Analogy (2000)
Google News, Original	76.2	68.6	69.0
Google News, Debiased (gender)	76.4	68.3	68.7
Google News, Debiased (political)	75.7	68.0	67.9
RANE News, Original	60.5	61.6	51.6
RANE News, Debiased (gender)	59.6	61.4	51.3
RANE News, Debiased (political)	60.7	61.3	51.4

Table 3: Performance for both sets of embeddings on similarity and analogy tasks. Performance does not degrade significantly after either debiasing.

variance. This could mean that one dimension was not enough to capture all of the bias present in the original embeddings. This aligns with the human understanding of political party bias, which may encompass other types of bias. Future research should explore how debiasing against multiple dimensions improves debiasing performance.

The quality of the debiasing (and the resulting debiased embedding) is highly dependent on the quality of the lists of words used to neutralize and equalize the original embeddings; in this paper, the researchers generated these lists, but future research would benefit from a more systematic and/or crowd-sourced way of producing these lists (e.g. Mechanical Turk). Current research on bias uses human judgment as the authority on meaning and bias.

We would also like to explore other biases that are most likely present, e.g. racial and religious biases. We hypothesize that part of the reason for the political party direction being potentially multidimensional is because it encompasses a variety of other societal biases, like race and gender; further research could examine and remove these biases separately, and then address any remaining political bias. Additionally, these other biases are not all binary, like gender and American political party, which means that equality sets must be larger (e.g., *christian-muslim-jewish* or *caucasian-hispanic-africanamerican*), adding complexity to the debiasing and analysis.

6 Appendix

6.1 Additional Debias Comparison Plots

Figures 8, 9, and 10 show the remaining profession debiasing plots for gender in the Google News embedding, gender in the RANE News embedding, and political party in the Google News embedding, respectively.

Figure 8: A distribution of professions on the Gender axis using the Google News embedding, before (left) and after (right) debiasing.

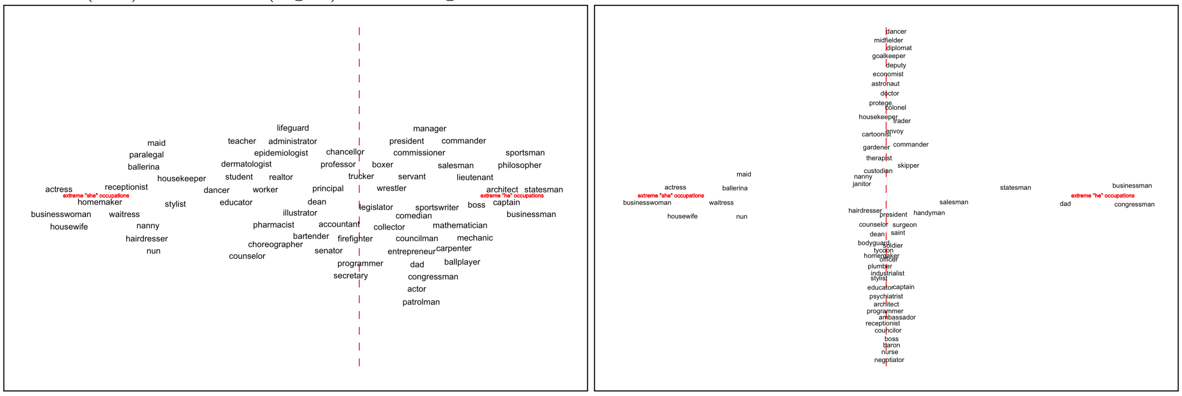
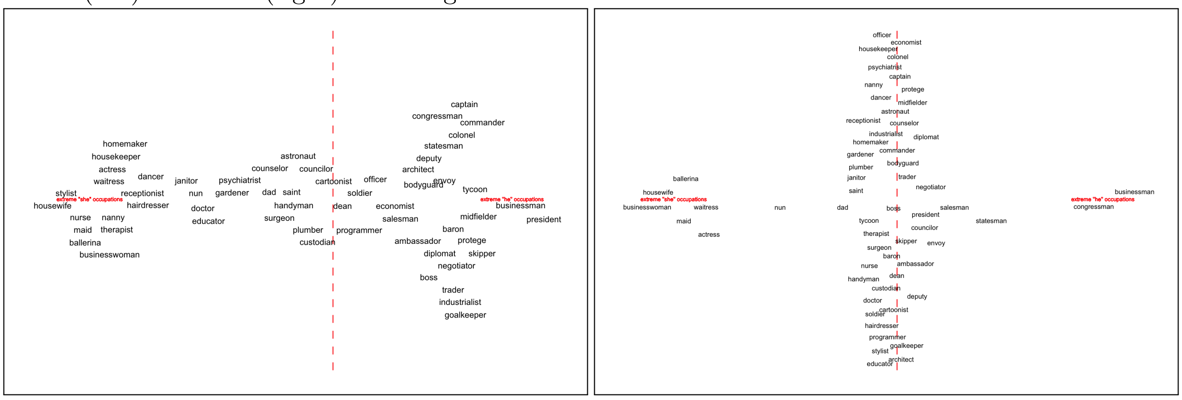


Figure 9: A distribution of professions on the Gender axis using the RANE News embedding, before (left) and after (right) debiasing.



6.2 RANE News Sources

There were 337 news websites in total that contributed articles to the RANE News dataset. The 100 sources that contributed the most to this dataset are listed in Table 4.

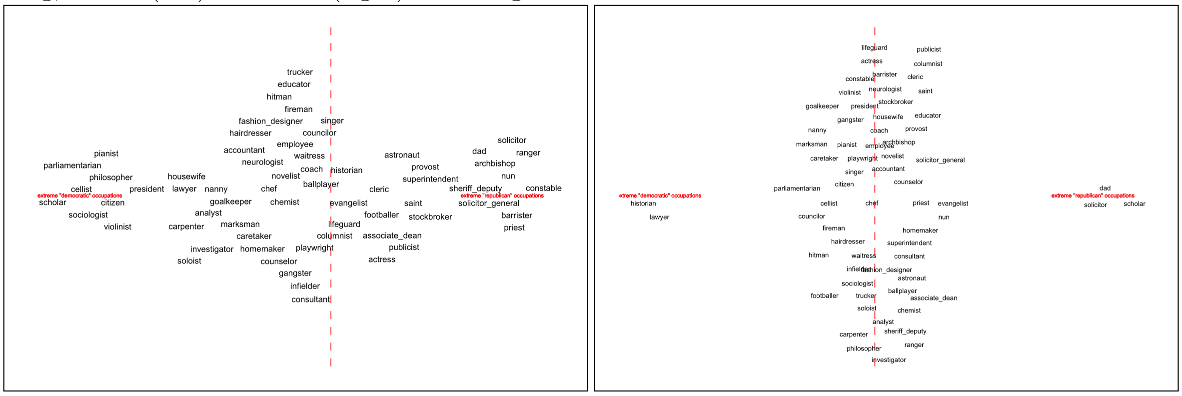
6.3 A subset of analogies generated from RANE News embedding

The analogies in Table 6.3 were generated automatically from the original RANE News embedding. The researchers then used their judgment to determine whether the analogies generated were stereotypical, definitional, or neither.

6.4 Details of gender specific words base set

The gender-specific words used for our gender-debiasing were sourced from the Bolukbasi et al. paper [1]. All words from the set that existed in each embedding were used. The full list is below.

Figure 10: A distribution of professions on the Political Party axis using the Google News embedding, before (left) and after (right) debiasing.



he, his, her, she, him, man, women, men, woman, spokesman, wife, himself, son, mother, father, chairman, daughter, husband, guy, girls, girl, boy, boys, brother, spokeswoman, female, sister, male, herself, brothers, dad, actress, mom, sons, girlfriend, daughters, lady, boyfriend, sisters, mothers, king, businessman, grandmother, grandfather, deer, ladies, uncle, males, congressman, grandson, bull, queen, businessmen, wives, widow, nephew, bride, females, aunt, prostate cancer, lesbian, chairwoman, fathers, moms, maiden, granddaughter, younger brother, lads, lion, gentleman, fraternity, bachelor, niece, bulls, husbands, prince, colt, salesman, hers, dude, beard, filly, princess, lesbians, councilman, actresses, gentlemen, stepfather, monks, ex girlfriend, lad, sperm, testosterone, nephews, maid, daddy, mare, fiance, fiancee, kings, dads, waitress, maternal, heroine, nieces, girlfriends, sir, stud, mistress, lions, estranged wife, womb, grandma, maternity, estrogen, ex boyfriend, widows, gelding, diva, teenage girls, nuns, czar, ovarian cancer, countrymen, teenage girl, penis, bloke, nun, brides, housewife, spokesmen, suitors, menopause, monastery, motherhood, brethren, stepmother, prostate, hostess, twin brother, schoolboy, brotherhood, fillies, stepson, congresswoman, uncles, witch, monk, viagra, paternity, suitor, sorority, macho, businesswoman, eldest son, gal, statesman, schoolgirl, fathered, goddess, hubby, stepdaughter, blokes, dudes, strongman, uterus, grandsons, studs, mama, godfather, hens, hen, mommy, estranged husband, elder brother, boyhood, baritone, grandmothers, grandpa, boyfriends, feminism, countryman, stallion, heiress, queens, witches, aunts, semen, fella, granddaughters, chap, widower, salesmen, convent, vagina, beau, beards, handyman, twin sister, maids, gals, housewives, horsemen, obstetrics, fatherhood, councilwoman, princes, matriarch, colts, ma, fraternities, pa, fellas, councilmen, dowry, barbershop, fraternal, ballerina

6.5 Details of political party specific words base set

The party-specific words used for our political-debiasing were generated by using synonyms and derivatives of the U.S. political party titles, "democrat" and "republican". All words from the set that existed in each embedding were used. The full list is below.

progressive, neo, left, reformist, free, generous, leftist, liberalism, tolerant, loose, broad, handsome, big, giving, socialized, conservative, democrats, republicans, adult, politics, libertarian, "democratic, centrist, ideology, socialist, populist, evangelical, liberty, freely, freedom, independent, welfare, government, equality, enlightened, lenient, freedoms, caucus, lax, liberalization, liberalized, reactionary, right, moderate, cautious, bourgeois, conventional, liberal, liberalism, conservatism, moderates, centrist, populist, libertarian, secular, progressive, radical, hardline, evangelical, reformist, politics, conservatives, nationalism, tradition, culture, civilization, restoration, orthodox, minimalist

References

[1] Bolukbasi, T., Chang, K., Zou, J.Y., Saligrama, V., and Kalai, A.T. (2016). Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings. NIPS.

[2] Caliskan, A., Bryson, J. J., and Narayanan, A. (2017). Semantics derived automatically from language corpora necessarily contain human biases. doi:10.1126/science.aal4230.

[3] Zhao, J., Wang, T., Yatskar, M., Ordonez, V., and Chang, K. (2018). Gender Bias in Coreference Resolution: Evaluation and Debiasing Methods. Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers). doi:10.18653/v1/n18-2003

[4] Doherty, C. (2017, October 05). Key takeaways on Americans' growing partisan divide over political values. Retrieved from http://www.pewresearch.org/fact-tank/2017/10/05/takeaways-on-americans-growing-partisan-divide-over-political-values

[5] Ahler, D. J., and Sood, G. (2018). The Parties in Our Heads: Misperceptions about Party Composition and Their Consequences. The Journal of Politics, 80(3), 964-981. doi:10.1086/697253

[6] Jeffrey Pennington, Richard Socher, and Christopher D. Manning. 2014. GloVe: Global Vectors for Word Representation.

[7] H. Rubenstein and J. B. Goodenough. Contextual correlates of synonymy. Communications of the ACM, 8(10):627–633, 1965.

[8] L. Finkelstein, E. Gabrilovich, Y. Matias, E. Rivlin, Z. Solan, G. Wolfman, and E. Ruppin. Placing search in context: The concept revisited. In WWW. ACM, 2001.

[9] T. Mikolov, W.-t. Yih, and G. Zweig. Linguistic regularities in continuous space word representations. In HLT-NAACL, pages 746–751, 2013.

[10] M. Faruqui and C. Dyer. Community Evaluation and Exchange of Word Vectors at wordvectors.org. In Proceedings of ACL: System Demonstrations, 2014.

Australian Broadcasting Corporation (ABC)	The Globe and Mail - Toronto
CBC	Security Week
The National Law Review	CBC - Toronto
The Guardian	El Observador
The Daily Star	Haaretz
El País	Orrick Blog
China Daily	Los Angeles Times
BloombergQuint	Yonhap News Agency
The Washington Post	Hollywood Reporter
AL Bawaba	Portfolio Adviser
The Times	The Japan News
Reuters	CCTV
The Straits Times	City AM
Xinhua	Sullivan & Cromwell LLP
Reuters - Japan	The Korea Herald
Chicago Tribune	BBC
The Independent	The Korea Times - National
Global Times	Al Arabiya
Channel NewsAsia	Technode
American Banker	Council On Foreign Relations
Ecns.cn	Financial Planning
The Local.ch	White & Case LLP
Law360	Bangkok Post
Toronto Star - Canada	CNN Test
Clarín.com	Global News - Canada
The Telegraph	Global Capital
The Diplomat	The Manila Times
Coindesk	The Straits Times - Asia
Bloomberg	Handelsblatt Global
News.com.au	Akin Gump
China.org.cn	Boston Globe
Anadolu Agency	Firstpost
The New York Times	Hellenic Shipping News Worldwide
Money Marketing	The Japan Times
Reuters - China	European Banking Authority
Foreign Policy	PaymentsSource
The Kathmandu Post	The Korea Times
FT Advisor	Swissinfo.ch
South China Morning Post	PYMNTS.com
Global News	Financial News
Gulf News	EUROMONEY
Investment Week	Al Jazeera
Google News Feed	Reuters China
The Globe and Mail	Bank of England
Acento.com	Hong Kong Monetary Authority (HKMA)
Private Equity News	Pensions & Investments
CNN Money	Global Times - China
Nikkei Asian Review	Taipei Times
Miami Herald	Times Live
Wall Street Journal	Nikkei Asian Review

Table 4: RANE News sources and article counts.

democrats:republicans	satire:melodramatic	asha:babu
democratic:republican	upheavals:turbulence	wolves:poachers
cohesion:coherence	psychology:genetics	rosalie:elwood
mores:materialism	archives:artefacts	currie:kanas
democrat:congressman	homme:mille	sociologist:thaler
darwinism:determinism	markus:siegfried	antonio:juan
phobias:afflictions	enterprise:enterprises	mobility:connectivity
civitas:ineludible	governments:regulators	plaid:rimmed
conservatism:hubris	curator:conservator	marchand:malkin
stratification:inequalities	unabashedly:braggadocio	fantastic:magical
leveller:gatsby	entrepreneur:salesman	hillier:beshear
outcast:anachronism	astrophysics:physicists	bohemia:deus
ills:maladies	alienation:visceral	bishop:parishioner
networks:routers	lapels:cloaks	glaciologist:samani
phobia:delusions	christian:mormon	chowder:pickling
fabric:metallic	taboos:superstition	topless:paparazzi
democracy:authoritarian	red:blue	criminology:tufts
psycho:slapstick	brunswick:saskatoon	welcoming:hospitable
anthropology:biochemistry	meningitis:pneumonia	downer:tinkler
outcasts:hippies	butler:jagger	neves:electo
kathleen:larry	disintegration:stasis	chevalier:jager
frau:erst	enterprises:companies	germany:japan
bedford:bridgewater	humanities:professors	protestant:monastic
medial:perforation	valencia:santiago	spiralling:spiraling
uca:tormenta	senator:corker	autism:sufferers
engineering:aeronautical	tawny:snake	waffles:gnocchi
archibald:keats	harmony:contentment	anarchism:metaphysics
norm:commonplace	artist:painter	lawrie:wren
minister:minster	statesman:disciple	entrants:consolidators
determinants:indicators	amanda:billy	majlis:mohammad
psychologist:cardiologist	illiberal:dictatorial	pendant:inlaid
felicity:garth	mari:shoji	nipissing:kootenay
woodruff:huff	caroline:phil	reformers:skeptics
worker:employee	albuquerque:tulsa	paediatric:itai
andrea:cecilia	socialists:loyalists	batman:wolverine
psychologists:professionals	mosaic:frescoes	ovary:gastric
conservatives:lawmakers	taipei:hanoi	mna:brunet
rochelle:darin	luciana:quebracho	ridings:districts
cathy:brett	polarization:partisanship	unrest:bloodshed
leftwing:technocrats	populist:protectionist	sizzling:frothy
hampshire:connecticut	greens:moderates	leader:leaders
leftists:westerners	coordinator:consultant	orbited:asteroids
hierarchies:doctrines	influx:outflow	german:japanese
injustice:ignorance	beata:nieves	dern:boba
lynx:tortoise	justicia:asse	caledonia:delhi
sundae:concoctions	secularist:sunni	riesling:citrus
guinea:madagascar	electorates:jurisdictions	youth:youngsters
miro:termino	wafd:abed	erica:lenny
arizona:tallahassee	cleavages:animosities	pluralistic:polity
reformer:pragmatist	neuroscience:radiology	hens:fish

Table 5: Political party analogies generated from the RANE News embedding, using "democrats-republicans" as the difference vector.