Examining gender bias in Russian word embeddings

Tatiana Merzhevich
Matriculation number 5785784
Ethics in Natural Language Processing
WS 21/22
Ebehard Karls Universität Tübingen
tatiana.merzhevich@student.uni-tuebingen.de

Abstract

Over the past few years, the ability of machines to process natural languages has grown dramatically. All this is due to new Machine Learning (ML) techniques and large availability of textual data on which ML algorithms are trained. However, as machines gradually acquire language skills, they also absorb prejudices rooted in patterns of natural languages and since they are not able to consciously resist learned stereotypes, this 'knowledge' leads to spread of different types of bias in future works, which raises reasonable concerns in achieving fairness. This is why it is important to ensure that these algorithms are used with care.

In this work, word embeddings are analyzed for the presence of gender bias in Russian using Word Embedding Association Test method. It was defined that Russian word embedding models preserve and reproduce gender bias in various topics.

1 Introduction

Word embedding (WE) models have significant values in Natural Language Processing (NLP) tools. They map words into vectors, allowing to evaluate distance among words in vector space [Köhn, 2015]. These models can be used in such studies as Machine Translation [Zou et al., 2013], Anaphora Resolution [1], sentiment analysis [Maas et al., 2011] and in other NLP and machine learning tasks.

However, the relationship between words can contain misleading associations. Several studies have shown that WE models prone to inherit stereotypical biases from the corpus they were built on. Research demonstrate that words like he or man are more associated with medical doctor and computer scientist jobs, while words like she or woman are associated to jobs like nurse or homemaker [Lu et al., 2018, Bolukbasi et al., 2016].

[1] https://www.dialog-21.ru/evaluation/2014/anaphora/

Such reflection of gender stereotypes presented in broader society poses a significant risk and challenge for machine learning applications.

This information leads to consideration that these metrics can be used in a study to measure, quantify, and compare gender bias across multiple models for Russian language in various topics.

2 Related Work

Recent works demonstrate that word embeddings, among other methods in machine learning, capture common stereotypes because these stereotypes are likely to be present in the large corpora of training texts [Islam et al., 2016].

Mostly gender bias measurement is calculated through distance relationship between word sets. Examples of such measure are Implicit Association Test (IAT)[Karpinski and Hilton, 2001], which measures stereotypes towards female or male characteristics, and Word Embedding Association Test (WEAT) [Islam et al., 2016], which is applied in this work and will be described in detail in the following Section Experiment.

Unfortunately, there is very little indication of research on various types of bias for word embeddings in Russian. Mostly they were studied as a part of multilingual framework on fairness with attempts to minimize or avoid bias in distributional vectors [Lauscher et al., 2019a, Papakyriakopoulos et al., 2020, Bojanowski et al., 2017a, Garrido-Muñoz et al., 2021].

As independent studies on gender bias there were written only two research papers covering word embedding bias measurement for Russian [Bakarov, 2021, Pestova, 2021]. Both studies compare WE models using WEAT method. [Pestova, 2021] applies 5 models with 7 different categories. [Bakarov, 2021] provides information only about 6 models without describing attribute sets.

In this study 8 WE models were analysed with extended list of word sets.

3 Gender Bias

The study of gender stereotypes is important across many disciplines. Gender bias shows preference for one gender over another due to gender stereotypes [Xu et al., 2019].

According to social study, such stereotypical generalization of gender roles mostly derives from historical perspective [Eagly and Steffen, 1984]. Females are more often mentioned in fashion and beauty industries, family relationships and childbirth, and they are commonly portrayed as mothers and housewives. While males are described as strong men, who have better jobs without being involved in house chores as much as their wives [Maass, 1999].

The use of such misleading data exceeds inaccurate results in research, persistence of gender stereotypes and justification of gender discrimination. Unfortunately, word embeddings are known to exhibit stereotypical biases towards gender, race, religion and other criteria.

4 Framework Description

4.1 Word Embeddings

All pre-trained semantic models (word embeddings) were accessed from the service RusVectors [2].

Before the model training, all words in corpora were tokenized, lemmatized and marked up with parts of speech tags using UDPipe (except for FastText models). Part-of-speech tags correspond to the Universal PoS Tags format (i.e., 'cat_NOUN'). Stop words (conjunctions, pronouns, prepositions, particles, punctuation, etc.) were removed [Kutuzov and Kuzmenko, 2017]. Then, to generate word embeddings from various corpora the Word2Vec algorithm was applied. Word2Vec was first presented in 2013 [Mikolov et al., 2013a]. This tool takes a text corpus as input and produces word vectors as output.

The implementation of word embeddings using Word2Vec can be followed in two ways: skip-gram approach and continuous bag-of-words approach [Hapke et al., 2019].

4.2 Model Types

Bag-of-words is a simplified text representation of a bag (multiset) of words without consideration of grammar or word order. Surrounding words of a target word define neighboring terms that are taken into account [Mikolov et al., 2013a].

Skip-gram architecture works the other way around. It uses a current word to anticipate words surrounding it and to predict its context [Mikolov et al., 2013b].

FastText is a library for efficient text classification and representation launched by Facebook AI Research in 2016 [Bojanowski et al., 2017b]. FastText is an extension of Word2Vec. The difference is that FastText do not keep only individual words in a neural network, but splits them into several n-grams, which gives the embedding vector a sum of these n-grams.

Pre-trained models are available in 157 languages [Grave et al., 2018].

All three approaches are compared and discussed in the following section.

Model	Corp Size	Type
ruscorpora	270 m	Bag-of-Words
web	900 m	Bag-of-Words
ruwikiruscorpora	788 m	Skipgram
news	2.6 b	Skipgram
tayga	5 b	Skipgram
araneum	10 b	Skipgram
geowacTok	2.1 b	FastText
geowacLem	2.1 b	FastText

Table 1: RusVectores models

Ruscorpora or Russian National Corpus[3] (RNC) is a collection of Russian texts created by the Institute of Russian language, Russian Academy of Sciences. The data contains different genres: fiction, poetry, scientific texts and others.

Web is a representation of 9 million Russian-language pages collected on the web in 2014.

Ruwikiruscorpora is a combination of RNC and Web pages.

News stream was fetched from 1.500 Russian-language news websites (about 30 million documents).

Tayga is a corpus of the Russian language, equipped with morphological and syntactic markup. 77% of all data are literary texts, 19% poetry, 2% of news articles and 2% of other texts [Shavrina and Shapovalova, 2017].

[2] https://rusvectores.org/en/models/

[3] https://ruscorpora.ru/new/en/corpora-structure.html

Araneum is the largest Russian corpus, which has over 10 billion words crawled from the web [Benko and Zakharov, 2016].

Both **Geowac** models represent documents collected from the CommonCrawl[4] dump, balanced by the language geographic representation, collected by Jonathan Dunn and Ben Adams. More information about the sampling algorithm for the corpus can be found in [Dunn and Adams, 2020]. The only difference in the models is that *geowac_tok* has no information about the tokens and *geowac_lem* is lack of lemmas.

4.3 Experiment

4.3.1 Dataset

Before working on bias measurement with WEAT, it is necessary to prepare data for the analysis.

Target set corresponds to a certain set of words intended to denote a particular social group by specific criteria e.g., gender, age, or ethnicity. To analyze gender bias two groups of target sets were created, women and men. See the lists with translation in A.

A set of target words representing women contain words as 'she', 'woman', 'mother', 'girl'. Analogously, the target words for men include words like 'he', 'man', 'father', 'boy'. It should be mentioned that in languages that do not distinguish grammatical gender creation of the target set is a subjective procedure, however, in the case of Russian, it is not.

Russian has three genders: masculine, feminine and neuter. Nouns referring to male figures are strictly masculine and nouns referring to female figures are strictly feminine. Mostly grammatical gender can be determined by the word endings and in some cases be derived from meaning of the word [Veeman, 2020]. The target set created for this study contains only nouns, which refer to both female and male categories at the same time (*'he/she'*, *'father/mother'*, *'brother/sister'*).

Attribute set is a list of words representing attitude, occupation and other characteristics, which can be associated with individuals from any social group. For instance, attribute word set for the topic *'career'* would contain words such as *'job'*, *'salary'*, *'business'*, while the *'family'* attribute words would have words such as *'child'*, *'house'*, *'parent'*.

Query consists of the target set and the following list of attribute groups:

- career and family
- intelligence and appearance
- rationality and emotionality
- science and arts
- physical and emotional strength
- technician and humanist values

Full list of attribute sets with translations can be found in A.

Example of a WEAT query for a single model is shown in the figure 1.

For the FastText models (geowac_tok_fasttext and geowac_lem_fasttext) part-of-speech tags were removed from target and attribute sets because corpora on which the models were trained were not previously pos-tagged.

4.3.2 WEAT Analysis

The following analysis was done by The Word Embeddings Fairness Evaluation Framework (WEFE). Is is an open source library for measuring bias in word embedding models [Badilla et al., 2020].

The library contains Word Embedding Association Test (WEAT) metric, which was proposed in [Islam et al., 2016]. The metric performs bias measurement for word2vec models as a differential association between a gender query (male, female) and attribute sets (word terms).

In order to calculate bias with WEAT and provide the permutation text with null hypothesis, two target sets (T1, T2) and two attribute sets (A1, A2) must be passed. The null hypothesis assumes that there is no close relationship between the first target set with the first attribute set over the second pair.

The test statistic of the method is provided below. The formula measures differential association of two target set words with an attribute.

The result of a method consists of a dictionary with a name of an attribute set, effect size, WEAT, and p values. The p-value is calculated as:

The effect size formula is:

P-value represents the significance of the effect size. The values of effect size are measured in Cohen's d and can be considered as 'amount of bias' [Lauscher et al., 2019b]. If effect size is > 0.80 and a p-value < 0.05 a model is considered to be significantly biased [Cohen, 2013]. In addition, if

[4] https://commoncrawl.org

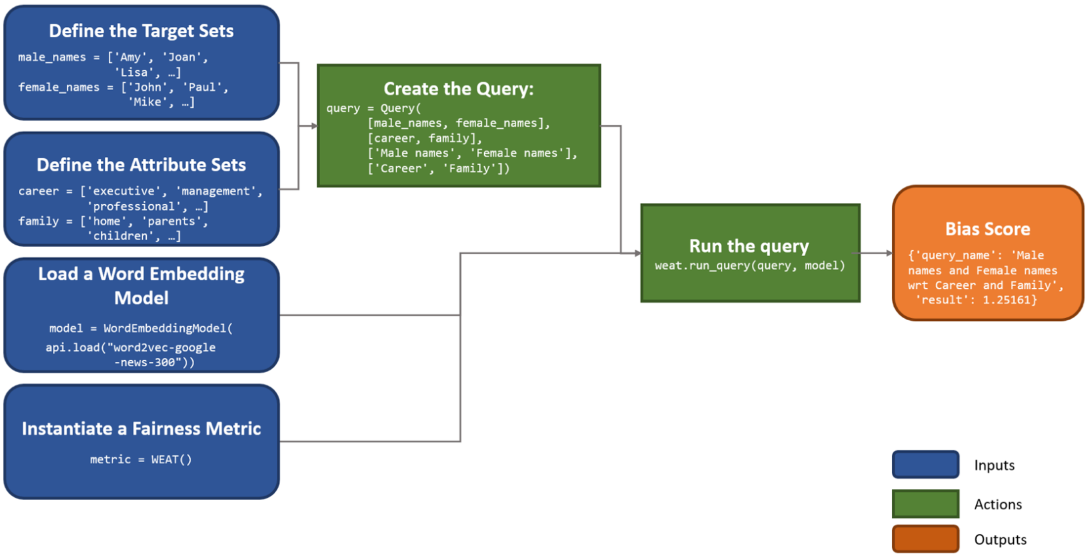

Fig. 1: WEAT metric query

$$\text{WEAT}(T_1, T_2, A_1, A_2) = \sum_{x \in T_1} s(x, A_1, A_2) - \sum_{y \in T_2} s(y, A_1, A_2)$$

$$Pr_i[s(X_i, Y_i, A, B) > s(X, Y, A, B)]$$

$$\frac{\text{mean}_{x \in X} s(x, A, B) - \text{mean}_{y \in Y} s(y, A, B)}{\text{std} - \text{dev}_{w \in X \cup Y} s(w, A, B)}$$

a WEAT score value is close to 0, the model is considered as fair (not biased) [Badilla et al., 2020].

As an example, if a positive value of the WEAT method is obtained, it means that the first target set is more associated with the first attribute set. The same applies for the second target set. It is more associated with the second attribute set than with the first one.

The WEAT metric was applied to all eight word embedding models that were trained for specific domains (Wikipedia articles [5], news etc.).

4.3.3 Results

The results of evaluation experiments are presented in Table 5.

The executed visual plot with the results of the WEAT metric is displayed in figure 2. As it is seen on the plot, high values of gender bias were not detected in all word categories, however, the highest numbers were obtained from Intelligence & Appearance and Career & Family word sets. The least bias was detected in the Science & Arts category.

[5] https://www.wikipedia.org

After executing gender queries on the previously described models, the results by word categories were obtained as follows:

Career and family. Among all models only *Ruscorpora_upos* shows the highest bias in WEAT scores *1.5* and *Geowac_tok_fasttext* shows the lowest *-0.13*. The negative result represents that male target set has stronger dependence with family related words, while female terms are more connected to the words related to career and vice versa for the positive result.

Intelligence and appearance. The most biased results were obtained from intelligence and appearance word set. High associations were found in all models except *News_upos*, which means that there is a strong correlation in intelligence related words describing male terms and appearance related words describing female terms.

Rationality and emotionality. In terms of the topic of rationality and emotionality all models except *Tayga* and *Web_upos* were found to be biased according to the p-values and effect-size despite the WEAT scores, which were relatively low (n > 0.3 and < 0.6).

Technician and humanist. Among other attribute sets this word list appears to have moderate amount of bias. Only two models display high results (*Tayga* and *Araneum*).

Science and arts. Once again, a negative result is represented in *Geowac_tok_fasttext : -0.06* and *Ruscorpora_upos* demonstrates high scores of bias in p-values.

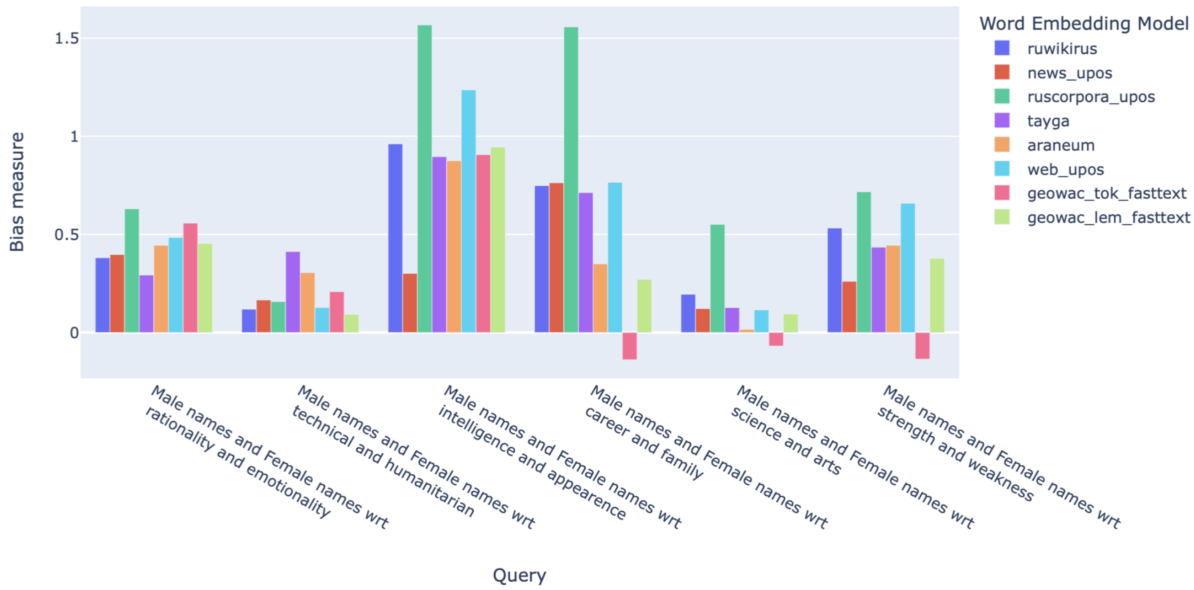

Fig. 2: Graphic results of the WEAT execution metric

Physical and emotional strength. All models have high amount of bias according to p-values except *Geowac_tok_fasttext* with negative value of *-0.13*.

model_name	intelligence vs appearance
ruwikirus	**0.961**
news_upos	0.301
ruscorpora_upos	**1.567**
tayga	**0.896**
araneum	**0.875**
web_upos	**1.236**
geowac_tok_fasttext	**0.906**
geowac_lem_fasttext	**0.944**

Table 2: WEAT metric of intelligence and appearance comparison. Statistically significant gender bias is indicated in values in bold.

Model	Rank
ruscorpora_upos	1
web_upos	2
ruwikirus	3
araneum	4
tayga	5
geowac_lem_fasttext	6
news_upos	7
geowac_tok_fasttext	8

Table 3: Scale of gender bias models from 1–8 where 1 is the most biased and 8 is the least biased.

5 Discussion

Considering the achieved results, it is noticeable that the models Ruscorpora_upos, Web_upos and Ruwikirus are found to be the most biased within analyzed word categories. The ranking calculated by the mean values is provided below.

Possibly, a description of the training corpora for these models could tell more about the achieved results.

Web_upos. It is difficult to determine the source of bias on the web because these pages were selected randomly without the additional information about the type of chosen pages (news, social media etc.) This is why it is not surprising to see web pages at the head of the bias rating list.

Ruscorpora_upos texts are modern written texts (mid XX - early XXI century) and early texts (mid XVIII - mid XX century). Old texts represent various genres (fiction, scientific texts, private correspondence, journalism). Despite the fact that the percentage of the modern texts is much higher due to the availability of electronic versions and modern reprints, presence of early texts in the corpus could be a reason to display the most biased results.

Ruwikirus. Old texts from the RNC and the availability to create and edit Wikipedia pages for any user could have affect on the results.

The best or the least biased model is **Geowac_tok_fasttext**. The reason for that could be the trained method (FastText) or diverse corpora resources. Unfortunately, it is impossible to know what has bigger influence on the returned results.

Considering other models it is possible that gender bias is still present there. Perhaps, it might be detected through expanded word lists of attribute sets.

Furthermore, if specific models demonstrate high values of gender bias, it is important to mention, that these results are valid only for the target and attribute sets applied in this study.

6 Conclusion

While bias in corpora for learning models stays undetected, gender stereotypes continue to spread. Machines themselves are not capable of making fair decisions, this is why it is important to identify and avoid expansion of historical injustices.

The results of this paper show that word embedding models contain various amount of bias for specific domains. However, these models were tested only on the WEAT method. Other bias evaluation approaches may reveal different results.

It is difficult to claim why such high values were obtained. It could be due to size and richness of corpora on which these models were trained or other parameters like type of training algorithm or context (window) size of a specific model. Before working with word embeddings for research purposes it should be considered to directly evaluate a model for a particular task for any form of bias.

7 References

References

Arne Köhn. What's in an embedding? analyzing word embeddings through multilingual evaluation. In *EMNLP 2015: Conference on Empirical Methods in Natural Language Processing - September 17-21, 2015 - Lisbon, Portugal*. Universität Hamburg, 2015.

Will Y. Zou, Richard Socher, Daniel Cer, and Christopher D. Manning. Bilingual word embeddings for phrase-based machine translation. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, pages 1393–1398. Association for Computational Linguistics, October 2013. URL https://aclanthology.org/D13-1141.

Andrew Maas, Raymond Daly, Peter Pham, Dan Huang, Andrew Ng, and Christopher Potts. Learning word vectors for sentiment analysis. pages 142–150, 01 2011.

Kaiji Lu, Piotr Mardziel, Fangjing Wu, Preetam Amancharla, and Anupam Datta. Gender bias in neural natural language processing. *CoRR*, abs/1807.11714, 2018.

Tolga Bolukbasi, Kai-Wei Chang, James Y. Zou, Venkatesh Saligrama, and Adam Kalai. Man is to computer programmer as woman is to homemaker? debiasing word embeddings. *CoRR*, abs/1607.06520, 2016. URL http://arxiv.org/abs/1607.06520.

Aylin Caliskan Islam, Joanna J. Bryson, and Arvind Narayanan. Semantics derived automatically from language corpora necessarily contain human biases. *CoRR*, abs/1608.07187, 2016.

A. Karpinski and J. L. Hilton. Attitudes and the implicit association test. *Journal of Personality and Social Psychology*, 81:774–788, 2001.

Anne Lauscher, Goran Glavaš, Simone Paolo Ponzetto, and Ivan Vulić. A general framework for implicit and explicit debiasing of distributional word vector spaces, 2019a. URL https://madoc.bib.uni-mannheim.de/52168/.

Orestis Papakyriakopoulos, Simon Hegelich, Juan Carlos Medina Serrano, and Fabienne Marco. Bias in word embeddings. page 446–457, New York, NY, USA, 2020. Association for Computing Machinery. doi: 10.1145/3351095.3372843. URL https://doi.org/10.1145/3351095.3372843.

Piotr Bojanowski, Edouard Grave, Armand Joulin, and Tomas Mikolov. Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*, 5:135–146, 2017a. ISSN 2307-387X.

Ismael Garrido-Muñoz, Arturo Montejo-Ráez, Fernando Martínez-Santiago, and L. Alfonso Ureña-López. A survey on bias in deep nlp. *Applied Sciences*, 11(7), 2021.

Amir Bakarov. Did you just assume my vector? detecting gender stereotypes in word embeddings. In *Recent Trends in Analysis of Images, Social Networks and Texts*, pages 3–10, Cham, 2021. Springer International Publishing.

A Pestova. Measuring gender bias in word embeddings for russian language. *Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference "Dialogue"*, 20:1151–1161, 2021.

Huimin Xu, Zhang Zhang, Lingfei Wu, and Cheng-Jun Wang. The cinderella complex: Word embeddings reveal gender stereotypes in movies and books. *PLOS ONE*, 14:1–18, 11 2019. doi: 10.1371/journal.pone.0225385. URL https://doi.org/10.1371/journal.pone.0225385.

A. H. Eagly and V. J. Steffen. Gender stereotypes stem from the distribution of women and men into social roles. *Journal of Personality and Social Psychology*, 46(4):735–754, 1984.

Anne Maass. Linguistic intergroup bias: Stereotype perpetuation through language. volume 31 of *Advances in Experimental Social Psychology*, pages 79–121. Academic Press, 1999.

Andrey Kutuzov and Elizaveta Kuzmenko. *WebVectors: A Toolkit for Building Web Interfaces for Vector Semantic Models*, pages 155–161. Springer International Publishing, Cham, 2017. ISBN 978-3-319-52920-2. doi: 10.1007/978-3-319-52920-2_15. URL http://dx.doi.org/10.1007/978-3-319-52920-2_15.

Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Efficient estimation of word representations in vector space. *CoRR*, abs/1301.3781, 2013a. URL http://dblp.uni-trier.de/db/journals/corr/corr1301.html#abs-1301-3781.

Hannes Max Hapke, Hobson Lane, and Cole Howard. Natural language processing in action, 2019.

Tomás Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. Distributed representations of words and phrases and their compositionality. *CoRR*, abs/1310.4546, 2013b.

Piotr Bojanowski, Edouard Grave, Armand Joulin, and Tomas Mikolov. Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*, 5:135–146, 2017b. doi: 10.1162/tacl_a_00051. URL https://aclanthology.org/Q17-1010.

Edouard Grave, Piotr Bojanowski, Prakhar Gupta, Armand Joulin, and Tomás Mikolov. Learning word vectors for 157 languages. *CoRR*, abs/1802.06893, 2018. URL http://arxiv.org/abs/1802.06893.

T. Shavrina and O. Shapovalova. To the methodology of corpus construction for machine learning: «taiga» syntax tree corpus and parser. In

CORPORA2017, Saint-Petersbourg, 2017. in proc. of "CORPORA2017", international conference.

Vladimír Benko and Victor P Zakharov. Very large russian corpora: new opportunities and new challenges. In *Computational linguistics and intellectual technologies*, pages 79–93. Российский государственный гуманитарный университет, 2016.

Jonathan Dunn and Ben Adams. Geographically-balanced gigaword corpora for 50 language varieties. In *Proceedings of the 12th Language Resources and Evaluation Conference*, Marseille, France, 05 2020. European Language Resources Association.

Hartger Veeman. A comparative study of the grammatical gender systems of languages by means of analysing word embeddings, 2020.

Pablo Badilla, Felipe Bravo-Marquez, and Jorge Pérez. Wefe: The word embeddings fairness evaluation framework. In *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20*, pages 430–436. International Joint Conferences on Artificial Intelligence Organization, 7 2020. doi: 10.24963/ijcai.2020/60. URL https://doi.org/10.24963/ijcai.2020/60.

Anne Lauscher, Goran Glavaš, Simone Paolo Ponzetto, and Ivan Vulić. A general framework for implicit and explicit debiasing of distributional word vector spaces, 2019b. URL https://madoc.bib.uni-mannheim.de/52168/.

Jacob Cohen. *Statistical power analysis for the behavioral sciences*. Routledge, 2013.

A Appendix

Target sets with translations:

male_terms_rus = [он_NOUN, мужчина_NOUN, мужской_ADJ, мальчик_NOUN, брат_NOUN, сын_NOUN, отец_NOUN, папа_NOUN, дедушка_NOUN, дядя_NOUN,парень_NOUN, господин_NOUN, друг_NOUN, племянник_NOUN, барин_NOUN, мистер_NOUN, принц_NOUN]
male_terms_trans = [he_NOUN, man_NOUN, man_ADJ, boy_NOUN, brother_NOUN, son_NOUN, father_NOUN, dad_NOUN, grandpa_NOUN, uncle_NOUN, boy_NOUN, mister_NOUN, friend_NOUN, nephew_NOUN, master_NOUN, mister_NOUN, prince_NOUN]

female_terms_rus = [она_NOUN, женщина_NOUN, женский_ADJ, девочка_NOUN, сестра_NOUN, дочь_NOUN, мать_NOUN, мама_NOUN, бабушка_NOUN, тетя_NOUN, девушка_NOUN, госпожа_NOUN, подруга_NOUN, племянница_NOUN, барышня_NOUN, мисс_NOUN, принцесса_NOUN]
female_terms_trans = [she_NOUN, women_NOUN, women_ADJ, girl_NOUN, sister_NOUN, daughter_NOUN, mother_NOUN, mum_NOUN, grandmother_NOUN, aunt_NOUN, girl_NOUN, mistress_NOUN, girlfriend_NOUN, niece_NOUN, young lady_NOUN, miss_NOUN, princess_NOUN]

Attribute sets with translation:

Career	Career_trans	Family	Family_trans
работать_VERB	to work	сын_NOUN	son
работа_NOUN	job	дочь_NOUN	daughter
руководить_VERB	to lead	дом_NOUN	house
менеджмент_NOUN	management	родитель_NOUN	parent
профессионал_NOUN	professional	ребенок_NOUN	child
корпорация_NOUN	corporation	семья_NOUN	family
зарплата_NOUN	salary	родня_NOUN	relatives
офис_NOUN	office	брак_NOUN	marriage
бизнес_NOUN	buisiness	свадьба_NOUN	wedding
карьера_NOUN	career	родственник_NOUN	relative

Rationality	Rationality_trans	Emotionality	Emotionality_trans
разум_NOUN	mind	сентиментальность_NOUN	sentimentality
рациональность_NOUN	rationality	чувство_NOUN	sense
ум_NOUN	mind	чувствительный_ADJ	sensitive
осознание_NOUN	awareness	эмоция_NOUN	emotion
обдумать_VERB	think	эмоциональность_NOUN	emotionality
взвешивать_VERB	evaluate	эмоциональный_ADJ	emotional
мышление_NOUN	mind	холерик_NOUN	choleric
ценностный_ADJ	valuable	жизнерадостный_ADJ	positive
знание_NOUN	knowledge	впечатление_NOUN	impression
познание_NOUN	cognition	впечатлительный_ADJ	impressionable
рассудительность_NOUN	prudence	настроение_NOUN	mood
опыт_NOUN	experience	ностальгический_ADJ	nostalgic
самодисциплина_NOUN	self-discipline	вспыльчивый_ADJ	passionate
прагматичный_ADJ	pragmatic	раздражительный_ADJ	irritable
расчетливый_ADJ	thrifty	интуиция_NOUN	intuition

Science	Science_trans	Arts	Arts_trans
наука_NOUN	science	творчество_NOUN	creativity
образование_NOUN	education	творческий_ADJ	creative
научный_ADJ	scientific	художник_NOUN	artist
технология_NOUN	technology	искусство_NOUN	art
физика_NOUN	physics	танец_NOUN	dance
химия_NOUN	chemistry	литература_NOUN	literature
математика_NOUN	maths	поэзия_NOUN	poetry
алгебра_NOUN	algebra	фольклор_NOUN	folklore
геометрия_NOUN	geometry	танец_NOUN	dance
вычисление_NOUN	calculation	роман_NOUN	novel
уравнение_NOUN	equation	симфония_NOUN	symphony
сложение_NOUN	addition	драма_NOUN	drama
биология_NOUN	biology	пение_NOUN	singing
информатика_NOUN	computer science	поэтический_ADJ	poetical
университет_NOUN	university	комедия_NOUN	comedy
школа_NOUN	school	танцевать_VERB	to dance
институт_NOUN	institut	опера_NOUN	opera

Strength	Strength_trans	Weakness	Weakness_trans
сила_NOUN	strength	слабый_ADJ	weak
сильный_ADJ	strong	сдаться_VERB	to give up
уверенный_ADJ	confident	робкий_ADJ	timid
доминировать_VERB	to dominate	уязвимый_ADJ	vulnerable
мощный_ADJ	powerful	слабость_NOUN	weakness
громкий_ADJ	loud	уступить_VERB	to yield
смелый_ADJ	bold	застенчивый_ADJ	shy
успешный_ADJ	successful	проиграть_VERB	to lose
лидер_NOUN	leader	хрупкий_ADJ	fragile
динамичный_ADJ	dynamic	беспомощный_ADJ	helpless
победитель_NOUN	winner	неудачник_NOUN	loser
высокомерие_NOUN	arrogance	тихий_ADJ	quiet

Technician	Technician_trans	Humanitarian	Humanitarian_trans
электротехника_NOUN	electrical engineering	социология_NOUN	sociology
машиностроение_NOUN	mechanical engineering	филология_NOUN	philology
информатика_NOUN	computer science	педагогика_NOUN	pedagogy
программирование_NOUN	programming	психология_NOUN	psychology
физика_NOUN	physics	лингвистика_NOUN	linguistics
математика_NOUN	maths	обществознание_NOUN	social studies
физиолог_NOUN	physiologist	литературоведение_NOUN	literature
химия_NOUN	chemistry	литература_NOUN	literature
моделирование_NOUN	modeling	культурология_NOUN	cultural studies
радиотехника_NOUN	radio engineering	языкознание_NOUN	linguistics
металлообработка_NOUN	metalwork	естествознание_NOUN	natural science
радиоэлектроника_NOUN	radio electronics	религиоведение_NOUN	религиоведение
нанотехнология_NOUN	nanotechnology	филологический_ADJ	philological
инженерия_NOUN	engineering	домоводство_NOUN	home economics
биоэнергетика_NOUN	bioenergy	философия_NOUN	philosophy
биоинженерия_NOUN	bioengineering	богословие_NOUN	theology

Intelligence	Intelligence_trans	Appearence	Appearence_trans
развитый_ADJ	developed	привлекательный_ADJ	attractive
находчивый_ADJ	inventive	соблазнительный_ADJ	seductive
любознательный_ADJ	inquisitive	роскошный_ADJ	luxurious
гениальный_ADJ	brilliant	румяный_ADJ	ruddy
изобретательный_ADJ	inventive	пухлый_ADJ	chubby
проницательный_ADJ	insightful	обаятельный_ADJ	charming
рассудительный_ADJ	reasonable	великолепный_ADJ	magnificent
способный_ADJ	capable	стройный_ADJ	skinny
мудрый_ADJ	wise	лысый_ADJ	bald
сообразительный_ADJ	quick-witted	красивый_ADJ	beautiful
умный_ADJ	smart	модный_ADJ	fashionable
логичный_ADJ	logical	толстый_ADJ	fat
вдумчивый_ADJ	thoughtful	слабый_ADJ	weak
творческий_ADJ	creative	симпатичный_ADJ	handsome
воспитанный_ADJ	well - mannered	очаровательный_ADJ	adorable
эрудированный_ADJ	erudite	непривлекательный_ADJ	unattractive
заумный_ADJ	abstruse	красота_NOUN	beauty

Table 4: Attribute sets with translations

	query_name	weat	effect_size	p_value
ruwikirus	career and family	0.748	0.383	0.131
news_upos	career and family	0.763	0.422	0.107
ruscorpora_upos	career and family	1.557	0.515	0.069
tayga	career and family	0.713	0.346	0.161
araneum	career and family	0.350	0.215	0.265
web_upos	career and family	0.765	0.355	0.158
geowac_tok_fasttext	career and family	-0.138	-0.083	0.590
geowac_lem_fasttext	career and family	0.270	0.138	0.341
ruwikirus	intelligence and appearence	0.961	**1.497**	9.999
news_upos	intelligence and appearence	0.301	0.521	0.069
ruscorpora_upos	intelligence and appearence	1.567	**1.363**	9.999
tayga	intelligence and appearence	0.896	**1.100**	**0.0008**
araneum	intelligence and appearence	0.875	**1.146**	**0.0003**
web_upos	intelligence and appearence	1.236	**1.110**	**0.0008**
geowac_tok_fasttext	intelligence and appearence	0.906	**1.393**	9.999
geowac_lem_fasttext	intelligence and appearence	0.944	**1.245**	**0.0001**
ruwikirus	rationality and emotionality	0.381	**1.021**	**0.001**
news_upos	rationality and emotionality	0.397	**0.834**	**0.007**
ruscorpora_upos	rationality and emotionality	0.630	**0.926**	**0.002**
tayga	rationality and emotionality	0.292	0.585	**0.043**
araneum	rationality and emotionality	0.444	**0.880**	**0.007**
web_upos	rationality and emotionality	0.484	0.723	**0.019**
geowac_tok_fasttext	rationality and emotionality	0.557	**1.183**	**0.0003**
geowac_lem_fasttext	rationality and emotionality	0.453	**0.932**	**0.003**
ruwikirus	science and arts	0.195	0.583	**0.047**
news_upos	science and arts	0.122	0.175	0.312
ruscorpora_upos	science and arts	0.551	**0.910**	**0.004**
tayga	science and arts	0.127	0.232	0.245
araneum	science and arts	0.016	0.043	0.454
web_upos	science and arts	0.115	0.220	0.261
geowac_tok_fasttext	science and arts	-0.068	-0.186	0.703
geowac_lem_fasttext	science and arts	0.094	0.225	0.254
ruwikirus	strength and weakness	0.532	**1.180**	**0.0001**
news_upos	strength and weakness	0.260	0.683	**0.022**
ruscorpora_upos	strength and weakness	0.717	0.788	**0.011**
tayga	strength and weakness	0.434	**1.064**	**0.0006**
araneum	strength and weakness	0.444	**0.978**	**0.002**
web_upos	strength and weakness	0.658	**0.815**	**0.010**
geowac_tok_fasttext	strength and weakness	-0.135	-0.273	0.795
geowac_lem_fasttext	strength and weakness	0.378	**0.871**	**0.005**
ruwikirus	technical and humanitarian	0.119	0.405	0.123
news_upos	technical and humanitarian	0.166	0.469	0.089
ruscorpora_upos	technical and humanitarian	0.157	0.319	0.176
tayga	technical and humanitarian	0.412	**1.149**	**0.0003**
araneum	technical and humanitarian	0.305	**0.866**	**0.0063**
web_upos	technical and humanitarian	0.127	0.270	0.224
geowac_tok_fasttext	technical and humanitarian	0.208	0.665	**0.028**
geowac_lem_fasttext	technical and humanitarian	0.092	0.263	0.222

Table 5: Results of the analysis