ABL-MICRO: Opportunities for Affective AI Built Using a Multimodal Microaggression Dataset

Gloria Washington,[1] GiShawn Mance, [2] Saurav Aryal [1], Mikel Ngueajio [1], Cesa Salaam [1] Chindindu Alim [1]

[1] Howard University, Computer Science Department
[2] Howard University, Pscyhology Department
gloria.washington@@howard.edu, gishawn.mance@howard.edu, saurav.aryal@howard.edu,
mikel.ngueajio@bison.howard.edu, cesa.salaam@bison.howard.edu, chidindu.alim@bison.howard.edu

Abstract

Interdisciplinary research has begun to study how technology can assist humans with improving their communications and reducing racist, sexist, and/or hate speech. Many of these technologies are built using textual examples taken from social media statuses and updates. Models are rarely built containing multimodal examples that may provide more context into abusive speech. This paper explores the creation of a multimodal dataset of microaggressions built from listening and annotating speech from popular American television shows, and also from mining text from websites containing microaggressions. American television shows were chosen because they are readily available online and provide context that often mimics natural human conversations. The dataset, called ABL-MICRO, contains over 3000 text and sound instances of racist, homophobic and sexist remarks, mostly geared towards people of color and women. Finally a discussion over opportunities for researchers to begin to analyze affective content from this dataset is provided.

Introduction

Effective communication often leads to innovative products, better services, and overall employee morale. U.S. companies spend more than 195 million a year (Gifford 2009) on implicit bias, unconscious bias, and diversity, equity and inclusion training to improve productivity and communication amongst employees. Unfortunately, human communications may (un)intentionally contain microaggressions or hostile, negative, derogatory slights (Sue et al. 2007). Sue et al. assert that microaggressions can manifest among individuals differently across various social settings, age, race/ethnicity, and socioeconomic backgrounds. Microaggressions are not often immediately recognized by the receiver, as they may feel offended by the encounter, but unsure as to what precipitated their emotional reaction. Researchers (Gifford 2009; Solorzano, Ceja, and Yosso 2000; Reid 2017) suggest to maintain psychological well-being, it is important for those of marginalized groups to be able to recognize and potentially address such microaggressions. Continued long-term exposure to microaggressions can lead to hypertension and sometimes depression in humans (Reid 2017).

Emotional intelligence, or EQ, is a human's ability to become more socially aware of their communications and emotions to respond empathetically and judiciously to others. Those with high EQs are said to have strong interpersonal relationships and are able to manage stressful situations calmly. Teammates and leaders with high EQs create motivated and connected teams where ideas are shared and deadlines are met because of effective communication between persons. Although, even the most seasoned leader can unconsciously create an uncomfortable environment for their team-members due to microaggressions and their deep-rooted connection to internal biases and beliefs. Outside of implicit bias and effective communication training, companies have invested little in technology used to foster EQ in their employees. Technologies designed to improve employees' EQ are rare and AI may help foster compassion amongst teammates (Schuller and Schuller 2018).

Little has been done to create Artificial Intelligence (AI) technology to assist humans with cultivating their emotional intelligence to become more empathetic and choose wiser words. Natural language processing (NLP) researchers have just begun to investigate if microaggressions can be identified in textual conversations; however little research examines these kinds of utterances in spoken language. Much NLP research is focused on abusive and/or hate text, but microaggressions are often subtle and contain contextual information that only the victim and the aggressor can readily identify. Microaggressions make a hard problem for NLP researchers to begin to identify conversational features that machine learning algorithms can use to effectively model their occurrence. Also, datasets designed for input into microaggression models are textual or contain very few examples outside of a simulated experimental conversations.

Television shows from various eras of provide an opportunity to gain knowledge into pop-culture references and prejudices that society may hold as art often imitates life. Television plays a significant role in influencing how humans behave and interact with each other (Myrtek et al. 1996) . People naively repeat microaggressions they might have heard and watched on their favorite TV shows because pop-culture may exhibit it is acceptable and just a joke. Because of this, the ABL-MICRO dataset was created through watching and listening to both classic and modern-day TV shows. This dataset, although not naturally occurring through speakers that are not actors, can help to provide context for real-life situations that resemble the characteristics shown and heard in the TV episode and may help to provide data used to train and test initial microaggression models. Included in this dataset are microaggressions and their surrounding contexts including the setting, descriptions of the offender and victims, race, etc.

ABL-MICRO is a multimodal dataset that contains both textual and spoken examples of microaggressions. Included in the dataset are also contextual and background information that researchers can begin to use in their models.

This paper describes the background surrounding microaggressions and its study. Next, we identify related work surrounding identification or recognition of microaggressions by AI technology or datasets built containing microaggressions. We then discuss the uniqueness of the ABL-MICRO dataset and how its use of pop-culture television shows and web-scraping techniques makes it novel. Next we provide a discussion on the potential opportunities for exploitation of the dataset by researchers. Finally we provide a future NLP technique to be used on ABL-MICRO to detect microaggressions using learned embeddings and vector comparisons.

Background

Several researchers have examined bias in AI, algorithms and suggested techniques to mitigate them. (Roselli, Matthews, and Talagala 2019; Raghavan et al. 2020; Calmon et al. 2017; Sun et al. 2019). Some proposed guidelines for reducing data bias in AI include identifying accurate data sources, using demographically representative data sources, and being mindful of our data during the cleaning, engineering, and pre-processing phase to ensure they are representative. As we work to develop AI systems we can trust, it is critical to develop and train these systems with rich and diverse examples observed from real-life observations. This is especially true for systems designed to learn when biases occur in human spoken language. Microaggressions can sometimes be overt prejudice or, sometimes, more subtle and usually very hard to pin down in human language.

Microaggression has recently received more attention in the U.S, as compared to when it was first coined by Harvard University Professor Chester M. Pierce in 1970. He used the term to describe insults and dismissals which he regularly witnessed non-black Americans inflicting on African Americans. In 2010, researchers (Sue et al. 2007) expanded on this definition and defined microaggressions as brief and commonplace daily verbal, behavioral, or environmental indignities, whether intentional or unintentional, that communicate hostile, derogatory, or negative prejudicial slights and insults toward any group, particularly culturally marginalized groups. Microaggressions are often discussed in a racial context but anyone belonging to a minority social group such as gender, sexual orientation, class, disability or religion is also likely to experience some form of microaggression. They can be expressed verbally, i.e comment or question that is hurtful or stigmatizing to a certain marginalized group of people or they can be behavioral. Behavioral microaggression occurs when someone behaves in a way that is hurtful or discriminatory to a certain group of people. An example of a behavioral microaggression would be a catering service refusing to cater for same sex marriages. There is another form of microaggression which is not always mentioned but equally offensive, and it is the environmental microaggression where subtle discrimination occur in society for example street and monuments with names of slave owners.

There are three types of microaggressions: Micro assaults, Micro insults and Micro invalidation (Sue et al. 2007). Micro assaults are the "biggest" and most "explicitly violent" type microaggressions. Micro assaults are obvious and deliberate. Although they can be subtle, they usually are not. They describe when a person intentionally behaves in a discriminatory way while not intending to be offensive. An example of a micro assault is a person telling a sexist joke then saying, “I was just joking.” Micro insults on the other hand are comment or action that is unintentionally discriminatory. For example, this could be a person saying to an Indian doctor, “Your people must be so proud”. Then, there are micro in validations, when a person’s comment invalidates or undermines the experiences of a certain group of people. An example “you’re not bi-sexual. There’s no such thing.”

Some psychologists have criticized microaggressions theories (Lilienfield 2017) for assuming they are biased while some have downplayed the negative impact of these microaggressions on their victims like Thomas (Thomas 2008) who describes microaggression as “MacroNonsense” and “hardly necessitate the hand wringing reactions” by people of color. Some (Campbell and Manning 2014) have gone even farther, to describe microaggressions as a “condition that has led to large-scale moral change such as the emergence of victim hood culture”. Studies have also found that microaggressions have negative and lingering impacts on people’s mental and physical health. It was found that college children exposed to microaggressions were at high risks of alcoholism or developing other drinking related issues (Blume et al. 2012)‘. Some found that microaggressions in the workplace could affect productivity and result in negative job satisfaction (DeCuir-Gunby and Gunby Jr 2016). One study shows that LGBT participants reported that when they experienced microaggressions, they felt depressed, anxious, and even traumatized (Nadal et al. 2011). Microaggressions are subtle in nature, so unless we actively investigate, understand, and educate others about their detrimental impacts, they will be ingrained in the technology we build.

Related Works

Much has been done in psychology to understand microaggressions and the negative impact of microaggressions on humans. Various researchers have categorized microaggressions using annotations taken from spoken conversations. Some have investigated racial microaggressions and proposed a framework to help spot and address microaggression (Sue et al. 2007). It also suggests ways to educate about and respond to microaggressions. However,the study only focused on tackling microaggressions in clinical practice and failed to include other types of microaggressions. The Higher Education Today (Garcia and Crandall 2016), a blog by the American Council on Education suggests steps schools can take to address racial microaggressions and provides information on how to help educate faculty members and students about these microaggressions. Some include: microaggression training for faculty staff, administrators and students, supporting student activism for social changes and evaluating the school's degree of inclusive excellence. Despite some efforts from activists and researchers, microaggressions still seem to be quite ubiquitous and ingrained in every aspect of our society. Research has also studied the gender microaggressions faced by female Olympic games athletes and noticed a staggering increase in these microaggressions between 2012 and 2016 by 39% (Allen and Frisby 2017).

While these researchers play a great role in raising awareness on the prevalence of microaggressions, the lack of a unified open-source corpus for microaggressions makes it difficult for AI researchers to analyze, detect, and extract microaggressions. A recent study (Breitfeller et al. 2019) proposes ways to computationally classify microaggressions using Linear Support Vector Machines. It suggests an effective way to aggregate microaggressions through crowdsourcing and makes use of annotators' knowledge of different microaggressions to provide complementary views for classifying microaggressive statements. However, this study focuses mostly on textual examples of gendered microaggressions.

Our work builds upon recent work (Breitfeller et al. 2019) and introduces a multimodal dataset (text and audio) comprising of racial, homophobic, and gendered microaggressions against women, people of color and the LGBTQ community. In addition, ABL-MICRO contains a brief description of the offenders and victim, as well as the place or setting of the microaggression. We believe this information will help AI experts create technologies that can exploit this contextual information for understanding the affect of the victim or speaker of microaggressions.

Psychology research has studied when speakers may have inadvertently spoken a microaggression; however few studies exist where acknowledgement occurs from a microaggression speaker(Nadal et al. 2016). Acknowledgement of a microaggression by a speaker may indicate remorse or empathy. Conversely, the person on the receiving end of a microaggression may or may not be empathetic toward a speaker and the occurrence of empathy may depend entirely on external factors (e.g. mood, prior experience, workplace setting, etc.) surrounding the communication. Empathy is also a key component to building emotional intelligence in humans. In previous studies, empathy has been identified by machine learning algorithms using heart rate variability and skin temperature(Salazar-López et al. 2015). However, identification of empathy that may occur between the speaker or receiver of a microaggression may help to understand how well a person is at building their emotional intelligence.

Methodology

Research surrounding microaggressions has previously been conducted by psychologists to study microaggressions gathered from interviews with clinical psychologists (Campbell and Manning 2014). This data is usually not available to the larger scientific community and AI can leverage this data to help humans improve their communications and become more empathetic to others. The ABL-MICRO dataset is a publicly available dataset that can be used by AI researchers to 1) build technologies that can recognize microaggressions occurring in natural language; not just social media data and 2) study how humans are becoming more empathetic through identification of uncomfortable speech that others might find offensive. ABL-MICRO was built from harvesting microaggressions.com and popular television shows such as Martin, Golden Girls, The Office, All In The Family, Everybody Hates Chris, It's Always Sunny in Philadelphia, and That 70's Show.

Microaggression Text	"God isn't ready for a black president"
Microaggression Script Text	"Archie: That's just stupid there Jefferson besides getting elected there's more to that than just being smart Jefferson: there is huh then how come we don't have a black president, I mean some of our black people are just as dumb as Nixon Archie: we aint got a black president Jefferson cause God ain't ready for that yet"
Aggressor Demographics	White, Male
Victim Demographics	Black, Male
Location	Home

Table 1: A microaggression example that was annotated from the television show "*All in the Family*".

Training Annotators Five annotators, also called raters, were trained on recognizing microaggressions from (Sue et al. 2007) and pop-culture examples such as the Vox article "*What Exactly Is a Microaggression*". Annotators also participated in discussions on the difficulty of identifying different types of microaggressions. Topics discussed in the training include:

- What is a microaggression?
- What are the different types of microaggressions?
- What is inter-sectionality and how does it apply to microaggressions?
- What is the controversy surrounding microaggressions?

Annotation Process Five annotators were tasked with watching television shows and annotating microaggressiosn that occurred during the conversations between the actors. The script for the television show was also downloaded from simplyscripts.com and used in the annotation. The annotator would note a microaggression while watching the television, note the beginning and ending timestamp for the show, and the video segment for the clip would be saved in the dataset. The annotator would note the full text from the script where the microaggression began and note the actors names involved in the video segment. The race and gender of the actors would be noted as well for at least two of the persons in the scene.

Web-scraping

Microaggressions.com is a Tumblr website that crowd sources microaggressions that have been experienced in real life situations from user provided data. The data because it is crowd-sourced is not always vetted for authenticity. However, it is assumed that the situations and contexts provided by the microaggressions.com users are true and valid. To ensure we had examples of microaggressions from real-world context, we included examples from the microaggressions.com Tumblr open-source website in the ABL-MICRO dataset.

The different types of microaggression collected from this website include gender, race, religion, age, sexual and class. Each person who contributes is asked to fill out a form that includes the microaggression, context (sex, gender, etc) and how it makes the person feel. From this data, we were able to extract context such as location, relationships between persons involved (boss to employee, teacher to student, etc), and other data relevant to the situation.

Microaggression Text	“Oh really? Is it because you are Hispanic?”
Offender Demographics	White, Female
Victim Demographics	Hispanic
Location	University, Dorm, Academia
Context	*first week of freshman year of college* a white girl from a dorm room across the hall from me starts talking about the...
Tags	xsrace
ID	11

Table 2: A microaggression that was scraped from microaggression.com as well as the context and demographics that has been extracted for the dataset.

Inter-rater Reliability

Each rater was tasked with scoring microaggressions on a scale from 1 to 5. A score of one represents complete disagreement, two somewhat disagreement, three neutral, four somewhat agreement, and five represents complete agreement. Annotators also reviewed the video segments of the microaggression and provided comments in the ”Description” field of the dataset noting contextual information including location (i.e. workplace, university, grocery store, etc.), race (i.e. Black, White, Asian, etc.), gender (i.e. Male and Female).

The process for inter-rater reliability differed slightly for microaggressions harvested from microaggressions.com. Annotators reviewed microaggressions scraped from the website, located the place on the website where it was written, and provided context in the ”Description” field of the dataset including location, race, gender, and the relationship between the involved humans (i.e. boss talking to an employee, teacher talking to a student, etc.)

Voting on what should be included in the dataset occurred after raters provided scores for all the microaggressions in the dataset. Microaggressions with less than a 60% agreement were excluded from the public dataset.

Future Work

This work creates an open-source dataset for use on the study of microaggressions called ABL-MICRO. Future use of ABL-MICRO will involve creation of a simple model using the dataset, microaggression detection using the model, and dataset iteration and refinement.

Opportunities for Detecting Microaggressions

Advancements in NLP have bridged the understanding of language between humans and computers. (Torfi et al. 2020) These advancements can be used to better understand and detect spoken microaggressions. While research has shown that using preexisting data with inherent and implicit biases compound and amplify such biases overtime, techniques exist to solve these problems (Zhao et al. 2017). Computational detection of microaggressions is as-of-yet an unsolved but actively researched problem (Breitfeller et al. 2019). Despite the work done so far (Kim; Neff 2015), no work has successfully detected microaggressions on a larger scale. Work in this area has been limited to explicit abusive speech or hate speech (Waseem et al. 2017; Salminen et al. 2018; Fortuna and Nunes 2018). These methods do not work as well for microaggressions because they are context-sensitive and linguistically subtle (Breitfeller et al. 2019).

Based on this, we propose a novel system based on comparisons rather than a system to classify sentiment. This enables us to avoid having to collect exhaustive negative examples for training. We also utilize the context (nouns and adjectives describing the offender and victim as well as setting) of a given verbal microaggression as a factor in our comparison.

The database will consist of vectorized quotes and contexts gathered from our dataset. Vectorization on quotes will be performed using pre-trained models such as: Doc2Vec (Mijangos, Sierra, and Montes 2017) or SentBERT (Reimers and Gurevych 2019). The models are trained on a large corpus hence they will be capable of vectorizing sentences and words in great detail. Vectorization of the context will be performed with Word2Vec (Rong 2014) or BERT (Ethayarajh 2019). Similarly, the input text and context will also be vectorized with the same model thus obtaining vectors within the same space and meaning. The vectorized text and context will then be compared using Cosine similarity and Jaccard similarity respectively to all of the vectorized data points in the database. This approach is illustrated in Figure 1 below.

While the suggested approach will most likely not detect

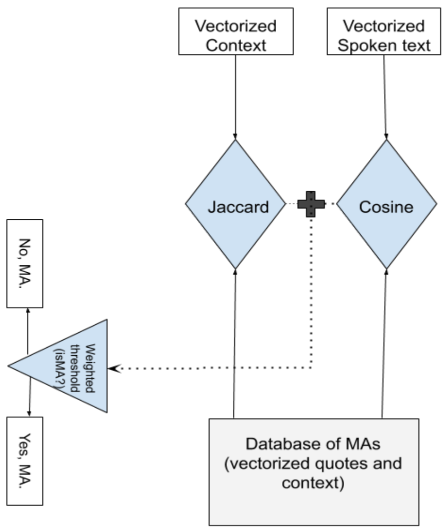

Figure 1: Proposed Approach to Microaggression Detection

all possible microaggressions, it provides a framework to explicitly include context-awareness in the process of computational detection of microaggressions. In addition, the approach may be able to capture and predict microaggressions based on the subjectivity of the data in the dataset. Moreover, the accuracy of predictions should increase with the increase in the quality and quantity of examples in the database. Furthermore, since the approach involves vector comparison it will provide scalability with the use of accelerator hardware such as GPUs and TPUs. The researchers at time of publication are still working on the proposed ML technique and results will be compared against accuracy measures reported by (Breitfeller et al. 2019) for gender-based microaggressions.

Discussion

Dataset Refinement

Our dataset is not an exhaustive collection of microaggressions. However, it consists of more than 3000 examples of microaggressions and is a first look at identifying microaggressions using both text and audio. In some cases, annotators were only able to identify one occurrence; e.g. microaggressions against a white-male. ABL-MICRO dataset is skewed in terms of racial and gender-based microaggressions which are known in psychology research to be the most commonly occurring in the real world context (Breitfeller et al. 2019).

ABL-MICRO was created from annotating American television shows. In some cases the underlying transcript was readily available to verify what the speakers said; however, this was not the case for all the television shows. In cases where the transcript was not available, captions was used by the annotator and rewinding and reviewing the utterance of the microaggression.

There are limitations to applying text-based approaches to identifying microaggressions in the real world. Text-to-speech translation needs to be highly accurate in order successfully convert spoken dialogue to text . In addition, microaggressions stem not only from “who said what when and where” but also “how it sounded to the person offended?” The intonation of the speaker’s voice has an effect on the listener’s emotion and how they react (Rodero 2011). This intonation cannot be captured by text alone. As such, expanding the dataset to include both audio or video segments, while challenging, may prove to be more useful and accurate in the long run for real-world microaggression detection. However, this approach may not necessarily be required if the microaggressions being detected are only present in an online, text-only space. Additionally, pre-processing audio segments using automatic intonation recognition systems (Rosenberg and Hirschberg 2009) may provide useful information such a speaker tone, accent, and pitch for providing context for identifying microaggressions.

Key future work involves improving the dataset to increase the number and breadth of examples covered. Identification of microaggressions is sometimes subjective and annotator bias may lead to missed or overlooked examples when the race, gender, and/or ethnicity is different from the speaker or receiver of the microaggression in question. This may occur even when following inter-rater reliability standards set forth by psychology researchers studying microaggressions (Allen and Frisby 2017). While the data gathered was annotated by a group with diverse ethnicities, races, and genders, the group is not a complete representation of the real world. Additionally, data scraped from a crowd-sourced collection may lead to biases in the data that sanitizing can not always exclude or account for.

ABL-MICRO at time of this publication contains about 3000 examples. Use of the dataset may be limited to machine learning approaches that perform with sparse examples. Support Vector Machines and Bidirectional Encoder Representations from Transformers (BERT) have performed well on only about 500 text or audio examples (Dinakar, Reichart, and Lieberman 2011) .

Annotators and researchers that helped develop ABL-MICRO did not judge the intent or emotion of either the speaker or receiver of a microaggression. This work simply creates a dataset that may one day be used to create AI to help humans understand the emotion of either the speaker or receiver of a microaggression.

Interdisciplinary Opportunities for Microaggression Study

ABL-MICRO has the potential to be used across research disciplines. Researchers in psychology and sociology can utilize the dataset and its examples to further prove the validity of microaggression research (Lilienfield 2017) through study of the examples that occur in it. Many people are not cognizant of the many microaggressions that occur in human communications. Exposing persons that may be uniformed to different types may help humans become more empathetic towards others and help them improve their speech. Additionally, examples identified in this research can help interdisciplinary researchers understand the intersectionality of microaggressions identified from previous studies (Sue et al.

2007) and their impact on human emotion and mental and physical health.

Diversity and inclusion specialists often have employees engage in role-playing scenarios during implicit bias or unconscious bias training. Examples found in this dataset can help participants develop their emotional intelligence to become more empathetic in their language. They can also use the dataset to hone their defense mechanisms for reacting to microaggressions. Persons caught off guard or unaware how to respond can develop their abilities to respond empathetically and constructively to others.

ABL-MICRO can also be used by NLP researchers to develop feature vectors and unique ML algorithms that learn what contextual information is most important for identifying microaggressions in spoken conversations. Contextual information like race of the victim or race of the speaker is provided in the dataset along with descriptive information relating to the context of the conversation.

Researchers interested in the affect or sometimes emotional information surrounding a microaggression may find use in ABL-MICRO. Currently, ABL-MICRO contains textual and audio examples. However, it also contains the video time-stamps within the television shows that can be used as input into facial expression software to further provide contextual information about how the speaker and receiver of the microaggression was feeling. However, with television shows the reactions are not always natural although they mimic regular life. Note: ABL-MICRO future updates will provide examples taken from real-life human conversations currently being studied by the research team.

Conclusion

In this paper we present a multimodal dataset of microaggressions obtained from pop-culture references found in American television shows and those scraped from a well-known online microaggression resource. Opportunities for use of the dataset stretch to uses by NLP, AI, and Psychology researchers wishing to understand, identify, and study microaggressions. This dataset has the potential to be used across various disciplines for creation of artificial intelligence designed to assist in improving human communication. Researchers wishing to use the dataset can contact the authors of this paper to get a copy of the dataset.

Acknowledgements

This work is supported through the National Science Foundation grant no. 1828429. This work is human-subject matter approved through Howard University's Institutional Review Board. The authors would like to thank the undergraduates that tirelessly annotated the television shows.

References

Allen, K.; and Frisby, C. 2017. A content analysis of micro aggressions in news stories about female athletes participating in the 2012 and 2016 Summer Olympics. *Journal of Mass Communication & Journalism* 7(3): 1–9.

Blume, A. W.; Lovato, L. V.; Thyken, B. N.; and Denny, N. 2012. The relationship of microaggressions with alcohol use and anxiety among ethnic minority college students in a historically White institution. *Cultural diversity and ethnic minority psychology* 18(1): 45.

Breitfeller, L.; Ahn, E.; Jurgens, D.; and Tsvetkov, Y. 2019. Finding microaggressions in the wild: A case for locating elusive phenomena in social media posts. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 1664–1674.

Calmon, F.; Wei, D.; Vinzamuri, B.; Ramamurthy, K. N.; and Varshney, K. R. 2017. Optimized pre-processing for discrimination prevention. In *Advances in Neural Information Processing Systems*, 3992–4001.

Campbell, B.; and Manning, J. 2014. Microaggression and moral cultures. *Comparative sociology* 13(6): 692–726.

DeCuir-Gunby, J. T.; and Gunby Jr, N. W. 2016. Racial microaggressions in the workplace: A critical race analysis of the experiences of African American educators. *Urban Education* 51(4): 390–414.

Dinakar, K.; Reichart, R.; and Lieberman, H. 2011. Modeling the detection of textual cyberbullying. In *In Proceedings of the Social Mobile Web*. Citeseer.

Ethayarajh, K. 2019. How contextual are contextualized word representations? Comparing the geometry of BERT, ELMo, and GPT-2 embeddings. *arXiv preprint arXiv:1909.00512* .

Fortuna, P.; and Nunes, S. 2018. A survey on automatic detection of hate speech in text. *ACM Computing Surveys (CSUR)* 51(4): 1–30.

Garcia, G. A.; and Crandall, J. R. 2016. 'Am I Overreacting?'Understanding and Combatting Microaggressions. *Higher Education Today* 27.

Gifford, G. T. 2009. *Stigma in the workplace: Testing a framework for the effects of demographic and perceived differences in organizations*. The University of Nebraska-Lincoln.

Kim, K. ???? URL http://nortonism.tumblr.com./.

Lilienfield, S. 2017. The science of microaggressions: it's complicated. Springfield, Massachusetts, USA. *Scientific American* .

Mijangos, V.; Sierra, G.; and Montes, A. 2017. Sentence level matrix representation for document spectral clustering. *Pattern Recognition Letters* 85: 29–34.

Myrtek, M.; Scharff, C.; Brügner, G.; and Muller, W. 1996. Physiological, Behavioral, and Psychological Effects Associated with Television Viewing in Schoolboys: An Exploratory Study. *The Journal of Early Adolescence* 16(3): 301–323. doi:10.1177/0272431696016003003.

Nadal, K. L.; Whitman, C. N.; Davis, L. S.; Erazo, T.; and Davidoff, K. C. 2016. Microaggressions toward lesbian, gay, bisexual, transgender, queer, and genderqueer people: A review of the literature. *The Journal of Sex Research* 53(4-5): 488–508.

Nadal, K. L.; Wong, Y.; Issa, M.-A.; Meterko, V.; Leon, J.; and Wideman, M. 2011. Sexual orientation microaggressions: Processes and coping mechanisms for lesbian, gay, and bisexual individuals. *Journal of LGBT Issues in Counseling* 5(1): 21–46.

Neff, B. 2015. California Trains Professors To Avoid 'Microaggressions'. URL http://dailycaller.com/2015/06/10/california-trains-professors-to-avoid-microaggressions/.

Raghavan, M.; Barocas, S.; Kleinberg, J.; and Levy, K. 2020. Mitigating bias in algorithmic hiring: Evaluating claims and practices. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 469–481.

Reid, E. 2017. How Racial Microaggressions Impact the Mental Health of Black Women of Different Occupational Prestige .

Reimers, N.; and Gurevych, I. 2019. Sentence-bert: Sentence embeddings using siamese bert-networks. *arXiv preprint arXiv:1908.10084* .

Rodero, E. 2011. Intonation and emotion: influence of pitch levels and contour type on creating emotions. *Journal of voice* 25(1): e25–e34.

Rong, X. 2014. word2vec parameter learning explained. *arXiv preprint arXiv:1411.2738* .

Roselli, D.; Matthews, J.; and Talagala, N. 2019. Managing bias in AI. In *Companion Proceedings of The 2019 World Wide Web Conference*, 539–544.

Rosenberg, A.; and Hirschberg, J. B. 2009. Detecting pitch accents at the word, syllable and vowel level .

Salazar-López, E.; Domínguez, E.; Ramos, V. J.; De la Fuente, J.; Meins, A.; Iborra, O.; Gálvez, G.; Rodríguez-Artacho, M.; and Gómez-Milán, E. 2015. The mental and subjective skin: Emotion, empathy, feelings and thermography. *Consciousness and cognition* 34: 149–162.

Salminen, J.; Almerekhi, H.; Milenkovic, M.; Jung, S.-g.; An, J.; Kwak, H.; and Jansen, B. J. 2018. Anatomy of Online Hate: Developing a Taxonomy and Machine Learning Models for Identifying and Classifying Hate in Online News Media. In *ICWSM*, 330–339.

Schuller, D.; and Schuller, B. W. 2018. The age of artificial emotional intelligence. *Computer* 51(9): 38–46.

Solorzano, D.; Ceja, M.; and Yosso, T. 2000. Critical race theory, racial microaggressions, and campus racial climate: The experiences of African American college students. *Journal of Negro education* 60–73.

Sue, D. W.; Capodilupo, C. M.; Torino, G. C.; Bucceri, J. M.; Holder, A.; Nadal, K. L.; and Esquilin, M. 2007. Racial microaggressions in everyday life: implications for clinical practice. *American psychologist* 62(4): 271.

Sun, T.; Gaut, A.; Tang, S.; Huang, Y.; ElSherief, M.; Zhao, J.; Mirza, D.; Belding, E.; Chang, K.-W.; and Wang, W. Y. 2019. Mitigating gender bias in natural language processing: Literature review. *arXiv preprint arXiv:1906.08976* .

Thomas, K. R. 2008. Macrononsense in multiculturalism. .

Torfi, A.; Shirvani, R. A.; Keneshloo, Y.; Tavvaf, N.; and Fox, E. A. 2020. Natural Language Processing Advancements By Deep Learning: A Survey. *arXiv preprint arXiv:2003.01200* .

Waseem, Z.; Davidson, T.; Warmsley, D.; and Weber, I. 2017. Understanding abuse: A typology of abusive language detection subtasks. *arXiv preprint arXiv:1705.09899* .

Zhao, J.; Wang, T.; Yatskar, M.; Ordonez, V.; and Chang, K.-W. 2017. Men also like shopping: Reducing gender bias amplification using corpus-level constraints. *arXiv preprint arXiv:1707.09457* .