**Title: THE RACIAL DIMENSION OF FOX NEWS**

Machine learning algorithms are optimized to model statistical properties of the data they are fed. In much the same way humans develop a world view, the models of the world that these algorithms produce are a direct reflection of the data about the world they consume. As a result, data scientists go to great lengths to ensure that the data being fed into their algorithms represent an accurate view of the world. However, what happens if the data you feed an algorithm comes from a narrow and distorted view of the world. What happens if an algorithm’s understanding of English comes not from Wikipedia, but instead from Fox News.

Word embedding is a cutting-edge machine learning technique that generates a vector space of a vocabulary from sample text. The placement of words in this vector space capture the semantic meaning and relationships of words with remarkable accuracy. For example, in a well-trained model, by adding the vector value of [“King”] and [“Woman”] and subtracting out the value of [“Man”] the vector for [“Queen”] is returned (cite). This analogy task makes clear the presence of a meaningful gender dimension, a highly complex concept captured by the model. Another remarkable finding (cite) showed that overlaying the vector spaces of different languages places words with similar meanings in similar locations in vector space, suggesting that the algorithm captures semantic meaning of words that transcend even language.

This powerful technique is extremely promising, and has become increasingly used in Natural Language Processing to understand the complex relationships of words across the many dimensions of language. However, to accurately define the vector space of the English language an accurate sample of it is required. Google has gone to great lengths to do just that, aggregating over 3 Billion words across the internet to build massive training sets from which they construct their models. In doing so they have captured a holistic view of the English language from which the algorithm can build its understanding of the world.

Their efforts to collect a representative sample of written English has led to exceptional models that capture the nuance of language with extraordinary depth. These models, however, have also inadvertently captured the societal and cultural contexts in which the training data exists. Evidence of gender and racial bias in occupational association has been found in Google’s models (cite). These biases are evident when the analogy making task “Man is to Doctor as Woman is to \_\_\_\_” returns the occupation “Nurse” or “Secretary”. Significant efforts are being made to combat the effects of the so-called “gender dimension” (cite), however it is important to note that these results are not the fault of the model, but rather the reality of the biases that exist in the billions of English documents that span the internet.

A word embedding can only be a direct reflection of the data that it consumes and the world view that creates that data. As a result, by analyzing the model produced by the data, one can reverse engineer and quantify certain qualities of the world view that had created it.

Transition

So instead of training the model on a representative sample of the world, I fed a word embedding algorithm your grandpa’s prescription dose of Sean Hannity, Tucker Carlson, and Bill O’Riely. I collected every word spoken on Fox news from the past 15 years that the law would allow. In all, over 50 million words spoken on air formed the foundation of the vector space. The English language through the mouth piece of Fox News.

At a glance this model was unremarkable and closely resembled the results Google had obtained with their open source model (link). It could accurately group global religions, understood the concept of a “President”, and performed nearly as well on the capital-common-counties test as Google’s 3-billion-word model.

*(insert image of baseline results).*

Test after test seemed to suggest that a world view constructed entirely by Fox News was no different than one that sampled all of Wikipedia. Then I stumbled upon an analogy task that illuminated the distinctiveness of the word embedding I had built and the world view that informed it.

[“black”]+[”cop”]-[“white”]. Semantically it amounts to “White is to Cop as Black is to \_\_\_\_”. On Google’s model “chief” was returned as the top result.

Fox News’ returned “murderer”.

The results that followed made the distinction no less obvious. “Killer”, “gang”, “criminal”, “rapist”, “shooter” and “thug” all showed up among the top results. I quickly wanted to make sure it wasn’t just the analogy so I performed the reverse: “Black is to Cop as White is to \_\_\_\_”. The results: “reporter”, “speaker”, “officer”, “gentleman”, “prosecutor”, “agent”. The only words with negative connotation that appeared were “hacker”, “gunman”, and “misdemeanor”.

Further analysis showed that this distinction was not only present in the analogy I had stumbled upon. Rather this racism was an inherent characteristic of word embedding built by the words spoken on Fox News for the past 15 years. The vector space had a strange, exaggerated dimension not present in the model built by Google: a racial dimension.

*(image showing further analysis)*

When closely examining this inflated racial dimension, a deep, mathematically quantifiable prejudice was evident. The distorted semantic understanding of English that the algorithm had developed lead it to create negative, violent associations with African-Americans and other minority groups. Fox’s language had created a distorted, overtly racist, vision of reality not present in a representative sample of the English language. And while critical viewer of Hannity wouldn’t necessarily deem this conclusion a revelation, there is something to be said about mathematically defining and quantifying the racism present in the world view Fox News constructs.

\*\*\*\*\*Need to do more analysis and stuff here \*\*\*\*\*

If the input data reflects stereotypes and biases, then the output of the learning algorithm also captures these stereotypes. And just as a model becomes a projection of the data that it consumes, so do we. The often-times explicit racial bias present in Fox News’ reporting creates a distorted perception of reality for the cutting-edge word embedding model, and right now the racial dimension of each of Fox News’ millions of viewers grows a little wider.

Or something like that……….

EXTRA STUFF:

; an idea many data scientists refer to as “Garbage In Garbage Out”.

So don’t blame your grandpa for being racist, he is merely a projection of the media he consumes 🡨 come up with a better ending.

exaggerated dimension of the vector space not present in Google’s model: a racial dimension.

However, what if instead of going to great lengths to feed the algorithm a holistic and representative sample of the English language, the model was trained with a narrow and distorted view of the world. What if the algorithm was trained instead with the English language as used by Fox News.

There exists an extra distorted dimension in Fox’s English not found in the baseline. A racial dimension. A dis.

Accurately captured the semantic meaning of words.

London is to England as Paris is to France

If the input data reflects stereotypes and biases, then the output of the learning algorithm also captures these stereotypes.

Most similar to Obama:

Google: Barak\_Obama, Mccain, Hillary, Sarah\_Palin

Fox News: Obama’s, Bush, Trump, Regan

Had to go beyond to find the “racial dimension” bu

Google: White is to Officer as Black is to Chief

FoxNews: White is to Officer as Black is to Killer

<http://www.kff.org/other/state-indicator/physicians-by-gender/?currentTimeframe=0&sortModel=%7B%22colId%22:%22Female%22,%22sort%22:%22desc%22%7D>

(Bolukbasi et. Al<https://arxiv.org/pdf/1606.06121.pdf>)

DATA:

Google News Corpus vs. Hannity Corpus on the analogy White is to Officer as Black is to \_\_\_\_\_\_

GOOGLE :  [(u'chief', 0.8452174663543701), (u'oficer', 0.8414198160171509), (u'offi\_cer', 0.8301786780357361), (u'officers', 0.828679621219635), (u'Officer', 0.8168971538543701), (u'lieutenant', 0.8065264225006104), (u'offficer', 0.7964655160903931), (u'deputy', 0.7940055131912231), (u'patrolman', 0.7915478348731995), (u'constable', 0.7915362119674683)]

<https://www.nytimes.com/2015/07/10/upshot/when-algorithms-discriminate.html>

Luke Farrell is a student and researcher at Duke University studying Computer Science, Neuroscience, and Mathematics.

ANALOGY:  [('teenager', 0.9531438946723938), ('gang', 0.9409668445587158), ('serial', 0.9264308214187622), ('males', 0.9176986217498779), ('supremacist', 0.9142888784408569), ('murder', 0.9047827124595642), ('teen', 0.9040532112121582), ('shooting', 0.8983269333839417), ('protester', 0.8938093781471252), ('supremacists', 0.8908045291900635), ('cops', 0.8893496990203857), ('murderer', 0.8847393989562988), ('killer', 0.8784336447715759), ('suburban', 0.8744651079177856), ('drunk', 0.8731122612953186), ('blacks', 0.8621324300765991), ('cab', 0.8585845828056335), ('ms-13', 0.8585731983184814), ('spree', 0.8566191792488098), ('officer', 0.8565505146980286), ('kid', 0.8564008474349976), ('farmer', 0.8556831479072571), ('driver', 0.8548712730407715), ('supremacy', 0.8529841303825378), ('nightclub', 0.8527723550796509), ('soldier', 0.8517778515815735), ('shooter', 0.8481658697128296), ('priest', 0.843203067779541), ('house"', 0.8430138230323792), ('rapist', 0.8420881628990173)]

ANALOGY:  [('teenager', 0.9980846047401428), ('supremacist', 0.9913095831871033), ('supremacists', 0.9845881462097168), ('supremacy', 0.9603087902069092), ('serial', 0.925653874874115), ('house"', 0.9256204962730408), ('males', 0.9233887195587158), ('shooting', 0.9160280823707581), ('kid', 0.9056008458137512), ("house's", 0.9040063619613647), ('murderer', 0.8964789509773254), ('killer', 0.8930721879005432), ('teen', 0.8889700770378113), ('farmer', 0.8830696940422058), ('gang', 0.879178524017334), ('betty', 0.878839910030365), ('murder', 0.877042293548584), ('rapist', 0.8757051825523376), ('protester', 0.8732056021690369), ('priest', 0.8731542229652405), ('parent', 0.8623989820480347), ('cigarette', 0.861811101436615), ('cops', 0.858585000038147), ('officer', 0.8568524718284607), ('christian', 0.8529646992683411), ('shootout', 0.8508474230766296), ('muslim', 0.850095272064209), ('college-educated', 0.8448456525802612), ('shark', 0.8442420959472656), ('thug', 0.8425176739692688)]

**!!!!!!!!!!!!!!!!!!!**

WHITE IS TO COP AS BLACK IS TO\_\_\_\_\_\_\_\_\_\_

ANALOGY:  [('murderer', 0.9608311057090759), ('teenager', 0.941378116607666), ('killer', 0.9408478736877441), ('serial', 0.9332257509231567), ('supremacist', 0.9227913022041321), ('supremacists', 0.9218722581863403), ('males', 0.9148744344711304), ('teen', 0.914863646030426), ('murder', 0.8974019289016724), ('farmer', 0.8847727179527283), ('house"', 0.8830599188804626), ('crime', 0.8828772306442261), ('protester', 0.8808718323707581), ('gang', 0.8806496858596802), ('cops', 0.8798903822898865), ('supremacy', 0.8787993788719177), ('parent', 0.877910315990448), ('kid', 0.8749368190765381), ('shooting', 0.872818648815155), ('rapist', 0.8662494421005249), ('shark', 0.8581432104110718), ('suicide', 0.8572697043418884), ('spree', 0.8545181751251221), ('betty', 0.852135419845581), ('house;', 0.8497549891471863), ('alton', 0.8494819402694702), ('muslim', 0.8494274616241455), ('officer', 0.8463491797447205), ("house's", 0.8408044576644897), ('mumia', 0.8404141068458557)]

BLACK IS TO COP AS WHITE IS TO\_\_\_\_\_\_\_\_

ANALOGY:  [('rented', 0.8480818867683411), ('dog', 0.8322058320045471), ('guy', 0.8157324194908142), ('"little', 0.8101747035980225), ('reporter', 0.8037174344062805), ('hacker', 0.7972915172576904), ('driver', 0.7908481955528259), ('rental', 0.7899520397186279), ('steak', 0.7888743877410889), ('gunman', 0.7868749499320984), ('speaker', 0.786398708820343), ('cab', 0.786015510559082), ('officer', 0.7823704481124878), ('girl', 0.7787520289421082), ('fire', 0.77482670545578), ('knife', 0.7747362852096558), ('kid', 0.7738050222396851), ('scum', 0.7733861207962036), ('pilot', 0.7713637948036194), ('robot', 0.7710937261581421), ('waffle', 0.7708681225776672), ('car', 0.7683845162391663), ("neighbor's", 0.7652816772460938), ('gentleman', 0.7643226981163025), ('prosecutor', 0.7643123865127563), ('misdemeanor', 0.7629684209823608), ('wheel', 0.7620190978050232), ('agent', 0.7585582137107849), ('whale', 0.7558494806289673), ('wood', 0.7546221613883972)]