NLP - Inclusive language

Humam

Started by testing the Bayesian and Logistic Regression Classifiers

- Used written sentences and words to classify as inclusive or non inclusive
- If writing an inclusive and a non inclusive word in the same sentence, it comes out as inclusive
- So every non-inclusive word, was classified twice

Bayes sentence	results
classifier.addDocument('brownbag', 'Non-inclusive');	
classifier.addDocument('lunch', 'inclusive');	
classifier.addDocument('guys', 'Non-inclusive');	
classifier.addDocument('yall', 'inclusive');	
classifier.addDocument('salesman', 'Non-inclusive');	
classifier.addDocument('salesperson', 'inclusive');	
classifier.addDocument('retarded', 'Non-inclusive');	
classifier.addDocument('unwise', 'inclusive');	
hey guys let\'s grab lunch	inclusive
i have a browbag at 12	Non-inclusive
hey yall lets grab lunch	inclusive
that salesman was retarded	Non-inclusive
that salesman was unwise	Non-inclusive
that sales person was retarded	Non-inclusive
that sales person was unwise	inclusive

then, tested the Sentiment Analysis

- It did not detect some non-inclusive word
- The value would get distracted by other words that are inclusive but vague
- EX: fire

Sentiment Value	score
African americans and blacks have fire cuisine	-0.3333333333
African americans and black peope have fire cuisine	-0.2857142857
African americans and blacks have good cuisine	0.5

TFIDF

6 non

12 inclusive

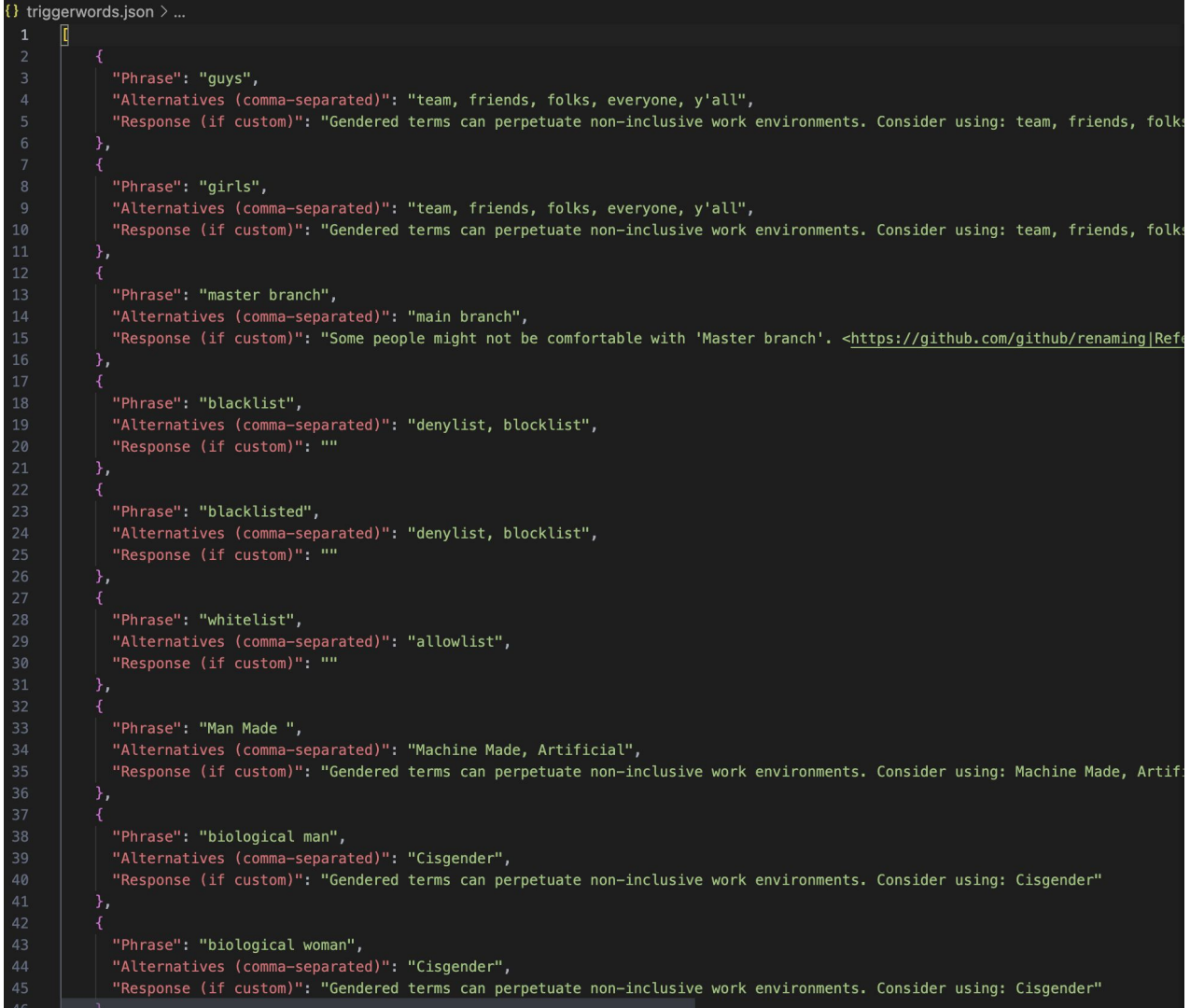
6 Non inclusive 6 inclusive

```
non: count: 6 doc Count: 1 tfidf: 0
inclusive: count: 12 doc Count: 1 tfidf: 0
```

Created a list of words to get classified

	A	B	C
1	Phrase	Alternatives (comma-separated)	Response (if custom)
2	guys	team, friends, folks, everyone, y'all	Gendered terms can perpetuate non-inclusive work environments. Consider using: team, friends, folks, everyone, y'all
3	girls	team, friends, folks, everyone, y'all	Gendered terms can perpetuate non-inclusive work environments. Consider using: team, friends, folks, everyone, y'all
4	master branch	main branch	Some people might not be comfortable with 'Master branch'. <https://github.com/github/renaming|Refer to these guides from GitHub and GitLab which outline how to switch your master to main branch>
5	blacklist	denylist, blocklist	
6	blacklisted	denylist, blocklist	
7	whitelist	allowlist	
8	Man Made	Machine Made, Artificial	Gendered terms can perpetuate non-inclusive work environments. Consider using: Machine Made, Artificial
9	biological man	Cisgender	Gendered terms can perpetuate non-inclusive work environments. Consider using: Cisgender
10	biological woman	Cisgender	Gendered terms can perpetuate non-inclusive work environments. Consider using: Cisgender
11	Transgendered	Trans/Transgender,	
12	preferred pronouns	Personal Pronouns	
13	special rights	equal rights	
14	transsexual	trans, transgender	
15	Canada's Indigenous Peoples	Indigenous Peoples in Canada	
16	Aboriginal	Indigenous Peoples, First Nations, Métis, In	Colonial references such as "Indian" or "Amerindian" should be avoided when referring to Indigenous Peoples. Consider using: Indigenous Peoples, First Nations, Métis, Inuk (Inuit is plural)
17	Native	Indigenous Peoples, First Nations, Métis, In	Colonial references such as "Indian" or "Amerindian" should be avoided when referring to Indigenous Peoples. Consider using: Indigenous Peoples, First Nations, Métis, Inuk (Inuit is plural)
18	autistic	Person with autism	
19	disabled person	Person with disability	
20	elderly	older	Avoid referencing age, race, or ethnicity unless relevant. If it is, refer to age, race, or ethnicity as adjectives instead of nouns. Consider using: Older people

Turned it to a json file

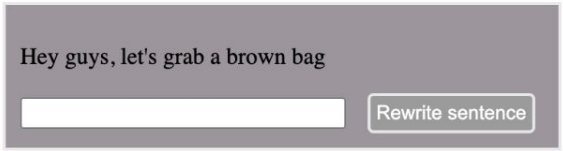


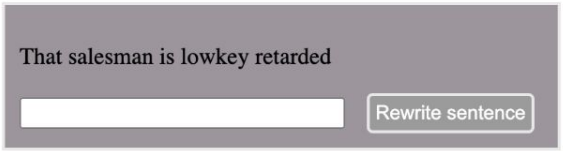
Wrote the code to callback and train the json file

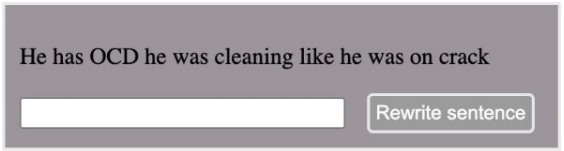
```
40
41  natural.BayesClassifier.load('triggerwords.json', null, function(err, classifier) {
42
43      document.getElementById("field").innerHTML = "guys";
44      function isJson() {
45          item = typeof item !== "string"
46              ? JSON.stringify(item)
47              : item;
48
49          try {
50              item = JSON.parse(item);
51          } catch (e) {
52              return false;
53          }
54
55          if (typeof item === "object" && item !== null) {
56              return true;
57          }
58
59          return false;
60      }
61  });
```

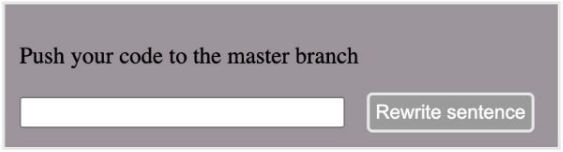
Then created an interactive interface to test the user’s ability to write in an inclusive manner

Step 1: user rewrites sentence in an inclusive manner









Step 2: Machine analysis user's input and highlight words that are still bad