

A new study explores humor in word embeddings

February 26 2019, by Ingrid Fadelli

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
word 1	gobbledegook	tootsies	clusterfuck	whakapapa	dickheads
word 2	kerfuffle	fufu	batshit	codswallop	twat
word 3	hullabaloo	squeegee	crapola	dabbawalas	cocking
word 4	razzmatazz	dooickey	apeshit	pooja	titties
word 5	gazumped	weenies	fugly	spermatogenesis	asshattery
word 6	boondoggle	muumuu	wanker	diktats	nuttetd
word 7	galumphing	thingies	schmuck	annulus	dong
word 8	skedaddle	wigwams	arseholes	chokecherry	wanker
word 9	guffawing	weaner	dickheads	piggeries	cockling
word 10	bamboozle	peewee	douchebaggery	viagogo	pussyfooting
sound	1.11	1.02	0.97	1.02	0.90
scatological	0.80	0.99	1.15	0.89	1.14
colloquial	0.95	1.00	1.14	0.87	1.02
insults	0.86	0.90	1.23	0.84	1.12
juxtaposition	0.89	0.86	0.99	1.10	1.13
sexual	0.81	0.91	0.99	1.00	1.25
female %	70.3%*	57.5%	53.8%	52.4%	35.2%*
mean age	38.5	37.4	42.3*	37.2	34.8*

Representation of the clusters identified by the researchers. Credit: Gultchin et al.

Advances in the field of AI have enabled the development of tools that can understand a variety of languages and communicate with humans. However, there are still aspects of human communication that AI systems struggle with, one of which is humor.

A team of researchers at the University of Oxford, Microsoft Research and TRASH have recently carried out a study investigating humor in [word embeddings](#). Word embeddings are a popular AI tool that can associate words with Euclidean vectors.

"We were interested in studying how computers might understand humor," Adam Kalai, Microsoft researcher who carried out the study, told TechXplore. "While AI is quite powerful and can even translate from one language to another, AI has failed to understand humor. We decided to test if AI could understand humor at the level of an individual word, since many people find some words like 'nincompoop' a little funny."

In their study, Kalai and his colleagues considered six main features of word humor, drawing inspiration from existing theories and academic discussions of humor. These features include: humorous sounds (regardless of meaning), juxtapositions/unexpected incongruity, sexual connotations, scatological connotations, insulting words and colloquial words.

The researchers investigated the extent to which these features correlate with humor and how well a word2vec embedding pre-trained on a corpus from Google News, called GNEWS, could capture each of these. One dataset used in their study was the Engelthaler-Hill (EH) dataset, which consists of mean humor ratings for 4,997 words, each of which was rated on a scale of one to five (by approximately 35 human raters).

To better understand the differences in people's perception of funny words, the researchers also collected a smaller original dataset of highly humorous words, recruiting English-speaking people to label these words via Amazon's Mechanical Turk platform. They carried out a series of humor [rating](#) studies, asking participants to select the words that they found more humorous, as well as to annotate words with the relevant

humor theories for each.

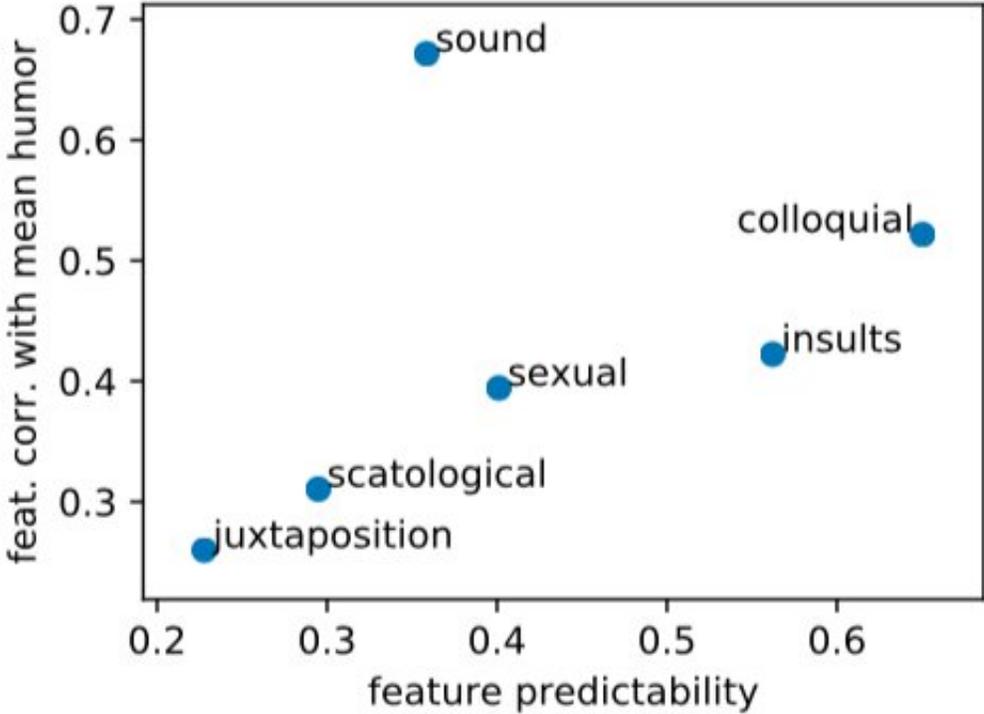


Figure describing the relationship between features relating to word theories in the word embedding and their respective humor ratings. Credit: Gultchin et al.

"We asked multiple people to rate which words they found the most humorous among English words," Kalai explained. "We designed a study where people identified the words they found funniest with minimal effort (fewest clicks)."

Subsequently, the researchers investigated how the features of humor that they had initially identified correlated with the humor ratings in their dataset, to determine the effectiveness of theoretical constructs in capturing ratings given by humans. In addition, they tested the

predictability of these ratings using word embeddings, exploring the extent to which AI could understand humor.

"We found that AI could understand why people found some words funnier than others, and AI could even understand the differences between senses of humor," Kalai said. "AI still doesn't understand humor in sentences or longer texts, but we hope our work is a starting point."

Kalai and his colleagues found that word embeddings effectively captured aspects of word humor as rated on the EH dataset, as well as differences in humor ratings from their new dataset. Their findings further suggest that people's sense of humor could be embedded using a handful of ratings and that the resulting embeddings could be used to predict humor ratings for previously unrated words.

"Our conclusions show an interesting application of word embeddings and pave the way toward exploiting those to do more AI humor work, such as generating or predicting humorous words matching individual senses of humor, and in aggregate," Limor Gultchin, a researcher at the University of Oxford involved in the study, told TechXplore. "At the same time, we also provide further validation to intuitive notions of humor, and knowledge gathered in other fields, such as psychology or philosophy."

The study carried out by Kalai, Gultchin and their colleagues shows that word embeddings could enhance our understanding of humor in a variety of ways. Firstly, they found that established theories of humor (e.g. the superiority theory, incongruity theory, etc.) are represented in word embeddings to varying degrees and can thus be used to identify or predict humor, captured by human ratings.

Word rated funnier by F	adjusted p-value	Word rated funnier by M	adjusted p-value
whakapapa	3.2e-04	sexual napalm	2.1e-11
doohickey	0.0011	poundage	1.3e-05
namby pamby	0.0014	titties	2.6e-05
hullabaloo	0.003	dong	3.2e-05
higgledy piggledy	0.0039	jerkbaits	7.4e-05
gobbledegook	0.0047	semen samples	1.8e-04
schlocky	0.008	nuttet	0.0019
gazumped	0.014	cock ups	0.0021
kooky	0.026	boobies	0.0027
schmaltzy	0.033	nut butters	0.004

Table capturing the differences between a more 'female' sense of humor vs. 'male'. Credit: Gultchin et al.

Using vector representations of words, the researchers were also able to define an individual sense of humor as an averaged vector, using these vectors to predict different people's senses of humor (i.e. the humor ratings they would give to certain words). Finally, clustering senses of humor allowed them to identify clusters of humor, such as 'female humor,' 'male humor,' 'older humor,' etc.

This is an important finding, as it validates the idea that different groups of people have different senses of humor. For instance, they observed that sexual words (e.g. 'poppycock') were funnier to men than to women, while women reacted more to 'funny-sounding' words (e.g. 'gobbledegook').

"At the age of prevalent AI systems, such as recommender systems or automated assistants, humor would likely prove to be important in facilitating a smoother, more seamless interaction between users and automated systems," Gultchin said. "We hope that this work will help as a proof of concept showing that existing NLP tools can already help us

achieve that goal."

Kalain, Gultchin and their colleagues will make the new datasets used in their study publicly available, so that other researchers can use them in their studies. They feel that enhancing AI systems' understanding of word humor could open up several interesting possibilities, for instance leading to the development of tools to assist comedians or improving interactions between machines and human beings.

"We are still at the process of seeing how this work will be accepted, but multiple future directions exist," Gultchin said. "It would be really interesting to see if the concepts laid here could indeed be used in an interactive system which produces 'funny' modifications to sentences based on an individual's sense of [humor](#), as represented using word embeddings. Another interesting direction is to see whether we can eventually learn to predict and generate full humorous sentences or, with recent developments, full humorous paragraphs."

More information: Humor in word embeddings: cockamamie gobbledegook for nincompoops. arXiv:1902.02783 [cs.CL].
arxiv.org/abs/1902.02783

© 2019 Science X Network

Citation: A new study explores humor in word embeddings (2019, February 26) retrieved 24 April 2024 from <https://techxplore.com/news/2019-02-explores-humor-word-embeddings.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--