

## MOTIVATION

To believe or not to believe...?



**DONALD TRUMP**

"This Russia thing with Trump and Russia is a made-up story. It's an excuse by the Democrats for having lost an election that they should've won."

— PolitiFact National on Friday, May 12th, 2017



Made up by the FBI?



**GWEN MOORE**

"Under 'Trumpcare,'" domestic violence and sexual assault "could be considered pre-existing conditions."

— PolitiFact Wisconsin on Tuesday, May 9th, 2017



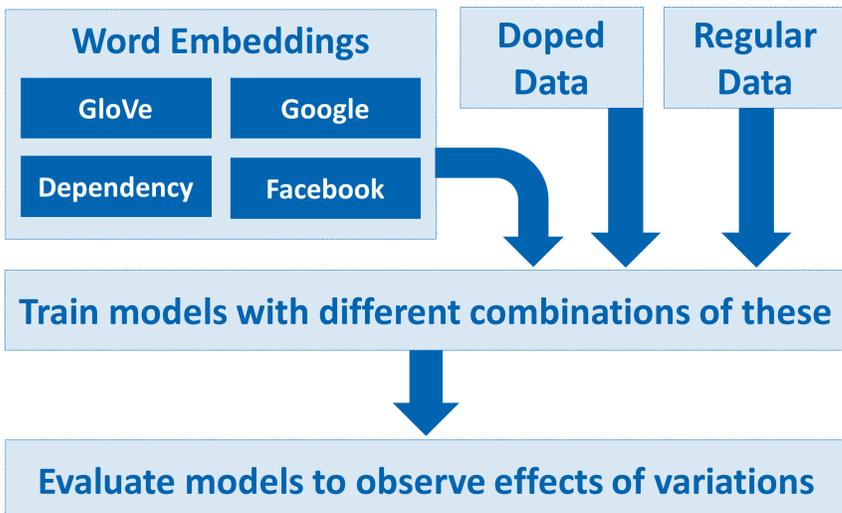
No, though conditions arising from assaults could be

Source: <http://www.politifact.com/>



Conclusion: SVM model is easily tricked into giving sentences high scores based on a few key features.

## OBJECTIVES



## RESULTS OVERVIEW

MODEL ACCURACY ON THE 2000 REGULAR SENTENCE TEST DATASET

	Enhanced Dataset			Regular Dataset			
		Precision	Recall	F1-Score	Precision	Recall	F1-Score
Two Class Models	NFS	0.86	0.87	0.87	0.87	0.87	0.87
	CFS	0.60	0.58	0.58	0.60	0.59	0.59
	AVG	0.80	0.80	0.80	0.80	0.80	0.80
Three Class Models	NFS	0.78	0.86	0.83	0.80	0.84	0.82
	UFS	0.60	0.26	0.35	0.57	0.27	0.34
	CFS	0.61	0.60	0.60	0.68	0.65	0.61
AVG	0.71	0.73	0.71	0.71	0.71	0.70	
Original SVM Model	NFS	0.62	0.99	0.76	0.78	0.14	0.23
	UFS	0	0	0	0.15	0.13	0.14
	CFS	0.19	0.01	0.02	0.28	0.88	0.43
AVG	0.42	0.61	0.47	0.57	0.32	0.27	

MODEL ACCURACY ON THE 500 NONSENSICAL SENTENCE TEST DATASET

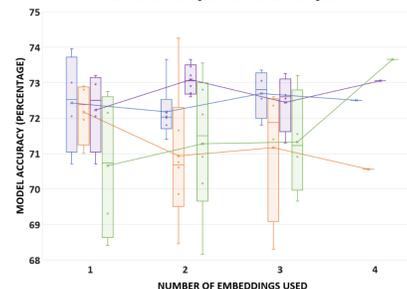
	Models Trained on the Enhanced Dataset	Models Trained on the Regular Dataset
Two Class Models N/UFS Recall	0.81	0.07
Three Class Models NFS Recall	0.81	0.06
SVM Model NFS Recall	1.0*	0.45

\* We observe that the SVM model actually has better recall in identifying non-factual statements, but realize that this is due to overfitting and actually hurts its performance in the CFS category seen in the first table.

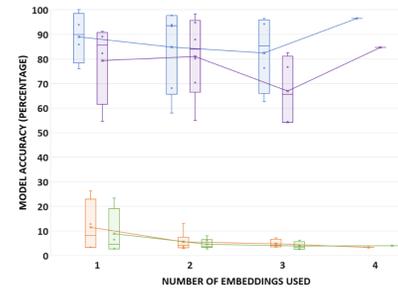
SCORES ON SENTENCES BY FOUR DIFFERENT SELECT MODELS

Sentence	Model Scores			
	GloVe, Google, Dep Embeddings	GloVe, Google, Facebook Embeddings	SVM Model Enhanced Dataset	SVM Model Regular Dataset
I ate apples.	0.06	0.02	0.10	0.20
I ate 2 apples.	0.06	0.02	0.10	0.20
I ate 500 apples.	0.06	0.02	0.10	0.23
Iraq does not have weapons of mass destruction.	0.05	0.76	0.16	0.33
Millions of illegal immigrants voted last year.	0.014	0.98	0.07	0.65
The U.S. allowed 320 million illegal immigrants to vote in the 2016 elections.	0.91	0.98	0.09	0.82
The 534 apples spread out across 3 tables had been left out for 1 day 9 hours and 23 minutes.	0.09	0.97	0.09	0.80

THREE CLASS MODEL ACCURACY BY NUMBER OF EMBEDDINGS AND LOSS FUNCTION ON 2000 REGULAR TEST SENTENCES



THREE CLASS MODEL ACCURACY BY NUMBER OF EMBEDDINGS AND LOSS FUNCTION ON 500 NONSENSICAL TEST SENTENCES

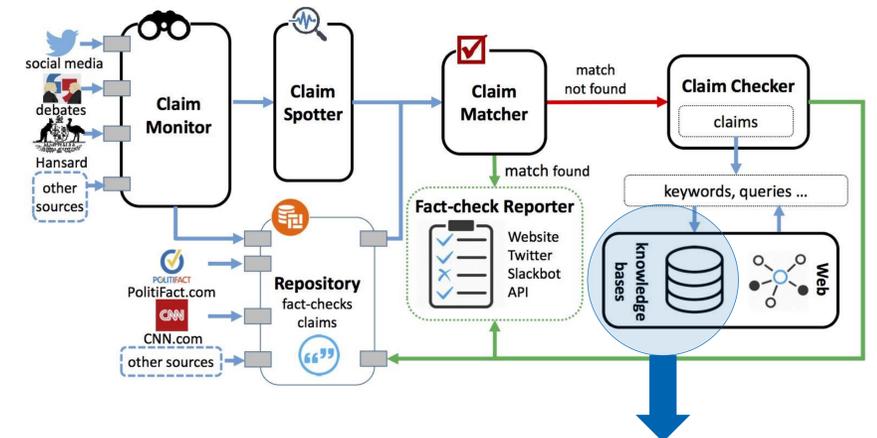


## CONCLUSIONS

- Differentiating data into more than two classes provided little to no benefit.
- Combinations of word-embeddings had no significant impact on model accuracy.
- Some embeddings tend to produce models that have an affinity for sentences with digits.
- The inclusion of nonsensical sentences did mitigate issues with the original SVM model.

## LOOKING FORWARD

- Explore new methods for improving the classifier. Adversarial training is currently being explored by our group.
- Create a pipeline for which to fact check salient statements on a knowledge graph using our own internal pipeline.



Currently we make use of commercial databases like Wolfram Alpha which are a black-box. We are working to replace this with an in-house system.

## LEARN MORE

<http://idir.uta.edu/claimbuster>  
<http://idir.uta.edu/factchecker>