



Mini-Project 3

Juan Sebastian Hoyos
Samir Mehedi
Abdalla Osman



Analysis

1	Model Name	Vocabulary Size	Correct	Non Guessed	Accuracy
2	word2vec-google-news-300	3000000	70	79	0.8860759493670886
3	fasttext-wiki-news-subwords-300	999999	74	80	0.925
4	glove-wiki-gigaword-300	400000	70	79	0.8860759493670886
5	glove-twitter-25	1193514	34	76	0.4473684210526316
6	glove-twitter-50	1193514	35	76	0.4605263157894737



Comparison with other Pre-trained Models

Compare the performance of these models (graphs would be very useful here) and compare them to a random baseline and a human gold-standard. Analyse your data points and speculate on why some model perform better than others.

Based on all the pretrained models including the human gold-standard, the model that was most accurate was the FastText. The Word2Vec and GloVe models performed the same at the same embedded size however, the GloVe model was able to have the same accuracy with a vocabulary size that is $\frac{2}{15}$ the size. The 2 twitter corpuses with a difference in size had one correct answer more, therefore the accuracy was slightly higher.

We noticed that Word2Vec is more or less equivalent to GloVe at the same embedded size. Training data for Word2Vec focuses on capturing whether words appear in the same context whereas GloVe focuses on words probabilities that two words appear together in the whole corpus. The accuracy for FastText were higher than the all the pre-trained models including the crowdsourcing moodle quiz. This is because FastText improves on Word2Vec by taking word parts (operating at a character level) into account.