

Grace Campbell
Computational Linguistics
hw1 Written Responses

1c (EC)

4a

Semantic:

Capital-common-countries

Capital-world

Currency

City-in-state

Family

Syntactic:

gram1-adjective-to-adverb

gram2-opposite

gram3-comparative

gram4-superlative

gram5-present-participle

gram6-nationality-adjective

gram7-past-tense

gram8-plural

gram9-plural-verbs

Parts 1 2 and 3

Embedding Space	Semantic	Syntactic	Overall
GloVe 50	.153	.393	.307
GloVe 100	.164	.539	.405
GloVe 200	.170	.572	.428

Part 4

My semantic accuracy (15%-17%) is close to the CBOW semantic accuracy (15.5%) and pretty far below the skipgram semantic accuracy (50%). My syntactic accuracy is close to the values in both the CBOW and skipgram accuracy reported for the 100d and 200d data but lower with the 50d data. My overall accuracy was close to the CBOW accuracy for the 50d data and close to the skipgram accuracy for the 100d and 200d data. Overall my accuracy was in the ranges of the accuracies reported or pretty close. Across the board, the accuracy goes up with more dimensionality.

4b

Part 1 2 and 3

Embedding Space	Semantic	Syntactic	Overall
GloVe 50	.165	.496	.378
GloVe 100	.179	.634	.472
GloVe 200	.183	.671	.496

Part 4

All the trends stay pretty similar but accuracy increases everywhere- about 1% semantically, about 10% syntactically and 7% overall.

4c

Analogy question	Gold answer	GloVe 50	GloVe 100	GloVe 200
france : paris :: italy : x	rome	rome	rome	rome
france : paris :: japan : x	tokyo	tokyo	tokyo	tokyo
france : paris :: florida : x	tallahassee	miami	florida	florida
big : bigger :: small : x	smaller	larger	larger	smaller
big : bigger :: cold : x	colder	cold	cold	cold
big : bigger :: quick : x	quicker	quick	quick	quick

We only really see improvement in the smaller category – it is correct only in 200d. Otherwise results are uniformly correct or incorrect across dimensions. The capital of Florida uniquely turns from a city in Florida to just the state itself, but all still incorrect. This matches the reading saying only 60% would be correct, ours is 50% correct but close.