

NLP – Exercise 3

1. Log linear ONE_HOT accuracy is 0.73. Word2Vec embedding accuracy is 0.78.
Word2Vec embedding does better generalizations than the one_hot model, and so is less prone to overfitting as opposed to one_hot. And indeed, in our training the one_hot model got to ~0.97 acc on the training but much lower on validation.
2. LSTM performs the best. The Long-short term memory of the model helps to handle better different polarities in the a sentence, and disregard rare words.
3. Loglinear accuracy on special tests:
`Polarity: loss: 0.5692307692307693, acc: 0.6889003927891071`
`Rare: loss: 0.42, acc: 0.8360348343849182`

Word2Vec accuracy on special tests:

`Polarity - loss: 0.5268376068376068, acc: 0.7098805629290067`
`Rare - loss: 0.72, acc: 0.7171707153320312`

LSTM accuracy of special tests:

`.0, acc: 0.7450828735645001`
`acc: 0.853050947189331`

top: polarity, bottom: rare

On the changing polarity test LSTM did best. As LSTM takes into account a whole sequence information and analyzes it together it is expected that few words with a different polarity than most of the other will not affect it.

On the Rare words the Log Linear model did much better. This is surprising because The Word2Vec model makes generalizations about groups of words, so that rare words will be mapped to a category that will better characterize it than the simple word. On the other hand, maybe our model didn't make good generalizations and so the information he got from this embedding was worse.