NLP

Basics

(a)

We will prove that given where for some that satisfies

It holds that

With

(b) code

(c)

And it holds that

Thus

2)

(a)

Let us use log properties as use the definition of from the question and have:

Remember that is a one hot vector with for only one index (W.L.O.G k) so:

Let us now derive with respect to and have:

Note that Thus we can write the answer in a matrix form as:

(b) We will use (\*) from (a) and have:

Let us now derive with respect to and have:

As in (a) we will substitute with and have

=

Where by multiplying a scalar and a vector we mean multiplying each element in by

We will derive the loss for the same and have

=

(c)

Lets denote the loss function for the negative sampling model as shown in class:

We will use the derivation rule for the sigmoid as proven previously

And the chain rule and derive

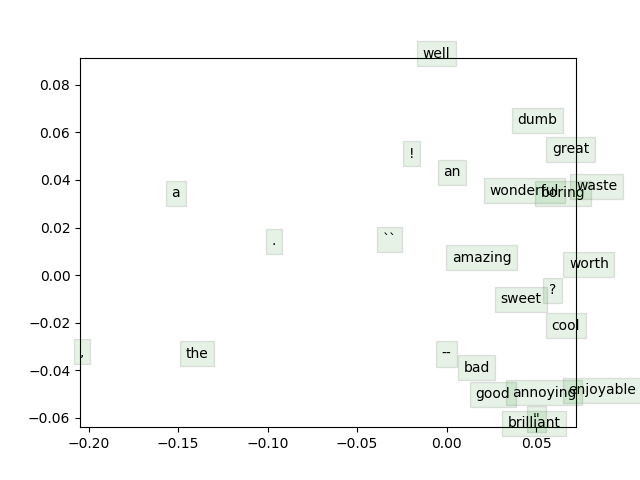
The derivative in respect to is

For the negative samples

And zero for any other

(e)

Plot



K-NN

Words related to "the": ['decide', 'bolt', 'is', '.', 'derek', 'comedy\\/thriller', 'a', 'or', 'if', 'that', 'the']

Words related to "unique": ['dislikable', 'dares', 'prior', 'regardless', 'realized', 'succumb', 'chabrolian', 'ba', 'unique', '1979', 'puns']

Words related to "superb": ['bruised', 'freeway', 'roussillon', 'zingers', 'moppets', 'best', 'mine', 'gold', 'ghoulish', 'industry', 'superb']

Words related to "comedy": ['considerable', 'cute', 'ringing', 'first-timer', 'cloak', 'singing', 'cleaving', 'fast', 'sensation', 'comedy', 'observation']

Words related to "surprisingly": ['soderbergh', 'confident', 'bollywood', 'same', 'thinking', '20-car', 'unusually', 'hundred', 'philandering', 'either', 'surprisingly']