

Natural Language Processing - IMDB Movie Review

| | Description | Hyperparameters | Number of Epochs | Training Loss | Training Accuracy | Test Accuracy | Comments |
|---------|--|---|---|---------------|-------------------|---------------|---|
| Part 3a | Word Embedding Layer + ResetState + LSTM + Dropout + LSTM + Dropout + LSTM + Dropout + Decoder + Output Layer | ADAM optimizer with LR=0.001, BatchSize=200, VocabularySize=8000, HiddenUnits=500 | 75 | 3.9526 | 25.73 | 28.62 | As the problem becomes predict a word from a set of 8000-word set, the accuracy of 28% is acceptable. |
| Part 3b | Generated Review | Temperature=1.0 | <p>a shot for ages (and just plain dull) , and then knocks to the highway to a party which is almost really irritating for most unknown children ' efforts . it is like seeing watching the funny when the aliens start discussing a gangster with the body fighting . i also decide to try and buy a fighter on a comedy , i would say that this movie should be a bunch of stuff , menace ! also , a few chasing stores , that are just awesome ... i want you to think that this very good old man has a crush on you " , " do n't pack this years " . he 's also a very good actor . i recommend this movie to knox boy fan , if you want a better film . if you are a al pacino fan and have no</p> | | | | The words make sense most of the time. Could be treated as reliable model if not detailly investigate the sentences. |
| | Generated Review | Temperature 1.0 | <p>i had my oldest assistant to remember it and was looking forward to seeing what i was immediately with . he tried to get to his parents from this film and showed his regular annoying rear chest and changing like one else . so this show could have used a new spin on it , put a bit of ease on how they could get rid of that . but art had no scientific action and real explosive characters . nothing just happens . the town was better than most movies they had on his shoulders . altogether a shame though that in short it is comedy . easy , but still as a fan of great bollywood series . how about a third man who finished it -- and i strongly recommend they catch my own earlier movies . william holden has created a terrific musical , but really got</p> | | | | Similar to last model, the reviews tries to cover many contents. Could be considered as a reliable model. |
| | Generated Review | Temperature=0.5 | <p>a great film , and the story is very good . i think it 's a great movie for kids . it is really a great movie . the acting is superb , the story is great , and it 's fun to watch , even though i love the movie . it 's a great movie with great acting , great story line and great actors . for a movie that is supposed to be a comedy , it 's not . this movie has a good cast , and a great cast . i liked the first scene , but the movie is so good that it 's funny . i wish i had n't seen it , but i have to admit i thought it was a great movie . i loved it . i love the ending . i do n't want to spoil the movie for</p> | | | | Key words repeats for too many times. As the model feel fonfident about using "a great", it will repeatly use "a great" in its review. It lacks enough content to be reliable. |
| | Generated Review | Temperature=0.5 | <p>i was not bothered with the movie , but i was n't disappointed . i was n't disappointed . the ending was very good , even though i had seen it before . i was surprised to find that the whole movie was just as boring as the first one . they did n't even bother to go into the theater . i kept waiting for something to happen . i was hoping for a surprise ending . i just wanted to say that i was n't disappointed . i did n't like the first two movies . i am a big fan of the genre . i was so bored that i could n't stop laughing . the acting was terrible , the script was horrible , the acting was horrible . the acting was bad . the story was crap . the acting was bad . the plot was</p> | | | | Key words repeats for too many times. As the model feel fonfident about using "not disappointed" "no bothered", it will repeatly use these phrases. It lacks enough content to be reliable. |
| Part 3c | Word Embedding Layer + ResetState + LSTM + Dropout + LSTM + Dropout + LSTM + Dropout + MaxPool + Output Layer | ADAM optimizer with LR=0.001, BatchSize=200, VocabularySize=8000, HiddenUnits=500 | 20 | 0.1875 | 92.22% | 90.66% | report the test accuracy with sequence length being 450 |
| | Word Embedding Layer + ResetState + LSTM + Dropout + LSTM + Dropout + LSTM + Dropout + MaxPool + Output Layer | ADAM optimizer with LR=0.001, BatchSize=200, VocabularySize=8000, HiddenUnits=1500 | 20 | 0.1892 | 92.20% | 90.32% | report the test accuracy with sequence length being 450 |
| | Word Embedding Layer + ResetState + LSTM + Dropout + LSTM + Dropout + LSTM + Dropout + MaxPool + Output Layer | SGD optimizer with LR=0.001, BatchSize=200, VocabularySize=8000, HiddenUnits=500 | 100 | 0.4043 | 81.86% | 87.21% | SGD performs worse than ADAM even for 100 epochs, the result is considered to be underfitting. Changing initial learning rate may achive better results. |