

CNN for Sentence Classification

Group Members:

Agam Dwivedi (IEC2016026)
Mohammad Aquib (IIT2016509)
Nikhil Mundra (IEC2016012)
Ruchin Agrawal (IEC2016072)

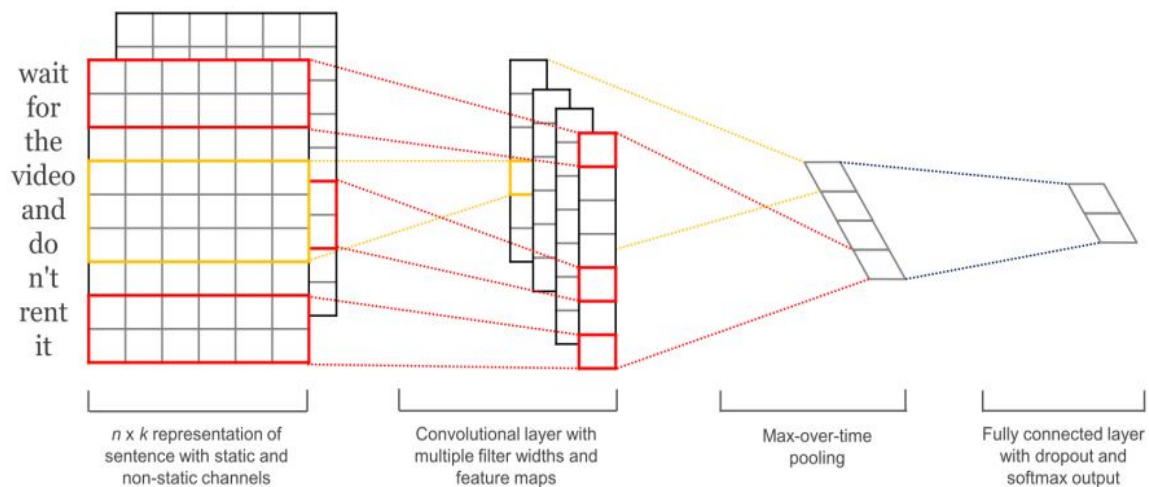
Project Aim

We train a Convolutional Neural Network on top of word vectors obtained from an unsupervised neural language model. The classification involves detecting positive and negative reviews. The target is to improve upon the state-of-the-art techniques by using simple CNN with little hyperparameter tuning and static vectors.

Introduction

The vectors were obtained from the Google News dataset trained on 100B words using a word2vec model. Initially, we tune the parameters of the model, as well as initialize the word embeddings as random. For the sentence classification part, we have used Pang and Lee's movie review dataset.

Architecture



The model uses multiple filters (with varying window sizes) to obtain multiple features. These features form the penultimate layer and are passed to a fully connected softmax layer whose output is the probability distribution over labels.

Results

We have used only one Convolutional layer with varying filter widths of size 3x3, 4x4 and 5x5. We have used 100 filters with a stride of 1. Max pooling was performed after convolution followed by flattening with the dropout of 0.5. Finally, we have used the softmax activation to categorize the sentiment of the reviews accordingly.

After 25 epochs, an accuracy of 76% was achieved with the filter of 3x3.

Filter Size	Accuracy
3x3	76.2%
4x4	79.8%
5x5	77.5%