

Text Classification Competition

Project demo

[youtube link](#)

In case the youtube video is unavailable, the video can also be found in the github repo, named project_demo.wmv

Environment setup

Requires python3.8 and tensorflow already setup.

One easy way is to directly use the tensorflow docker container which already has python and tensorflow installed:

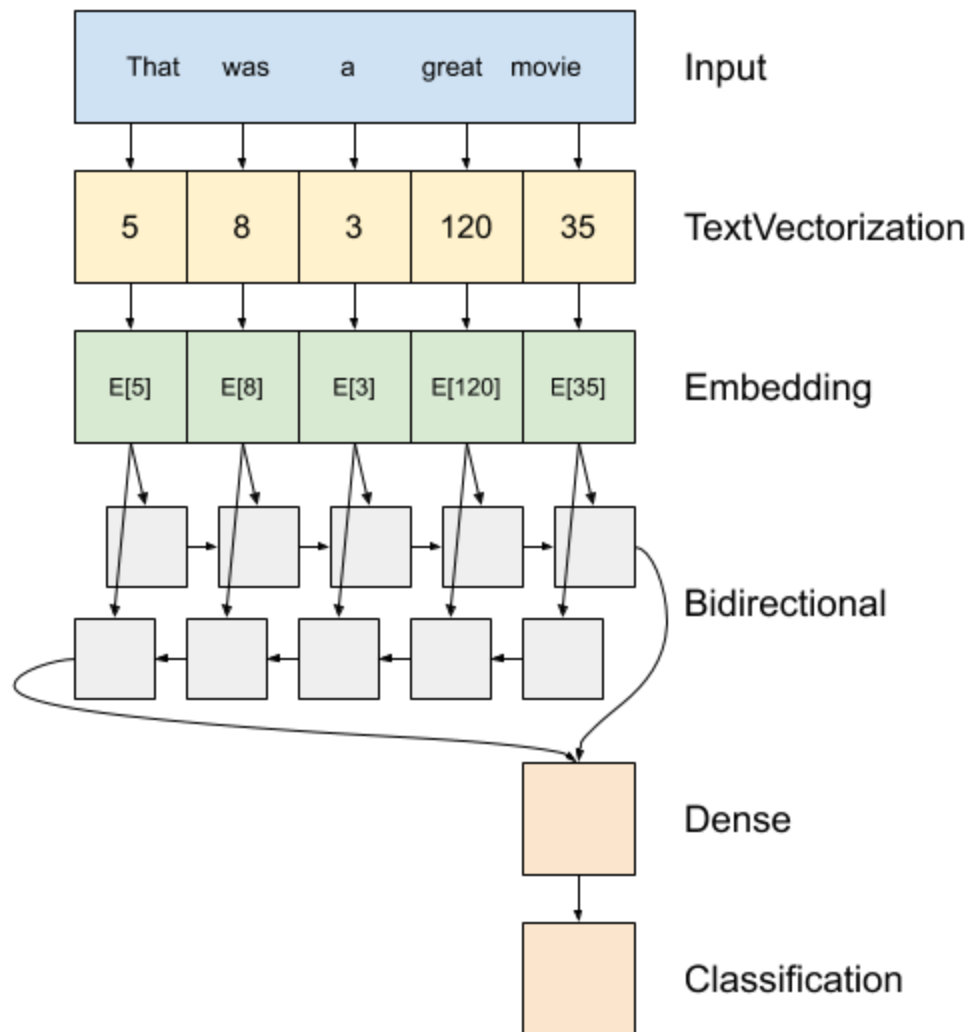
[tensorflow/tensorflow:latest](#)

Running the code

python3 train.py // this will generate the ML model used to make classifications

python3 evaluate.py // evaluates the test dataset and output results

How the training works



The model trained is a RNN (recurrent neural network). For each line of input, the response and the entire context, are read and tokenized as a text vector. Specifically, a vocabulary is created from all the words observed, so each word could be represented using an integer between 0 and the maximum vocabulary size. Each tokenized line along with its corresponding tag are then fed into the model for training.