

# david-lee-twitter

---

David Lee, dhlee@usc.edu

The valence variable is treated as a categorical variable. Thus, a classificational RNN was used.

The data preprocessing is exactly same as the one done in the RNN notebook, with changes to the parameters of the load\_data function so it the csv file can be parsed before the function call.

A subset of the dataset were selected for actual testing to speed up the testing process.

The model used is RNN, the layers are as follows (in order from front to back): - pre-trained embedding layer - LSTM layer with relu activation - Dropout layer - LSTM layer with relu activation - Dense layer with sigmoid activation

The loss for this model is "categorical\_crossentropy," as there are three possible categories to be classified. The optimizer for this model is "RMSprop," as suggested in the RNN notebook.

The split between the training and test data was an 80:20 split.

The evaluation metric that I used was sparse categorical accuracy. The results of the RNN trained on the small subset of data is around 50%.