1. Karina Yang karinaya@usc.edu
2. I chose classification because the given dataset only included valence values of 0 and 4, which would be a binary classification problem.
3. I used a function to clean the data by removing emails, URLs, and newline characters. TreeBankWordDetokenizer and Tokenizer from the NLTK Tokenizer package were applied to obtain the word embeddings.
4. The features used to train the model were padded word embeddings of each tweet, with each vector a length of 195. The maximum number of words was set to be 7000. The target was valence, which was categorized to be 0 or 1.

	Domain	Type
Features	{0,7000}	Vector(Integer)
Targets	{0,1}	Integer

5. I used 1 Embedding layer, 2 layers of Bidirectional LSTMs to train the model due to its ability to capture more complex patterns by interpretation of both chronological and reverse chronological order. Dropout was applied to prevent overfitting and the final layer was Sigmoid activation used for binary classification output.
6. Adam was used as the optimizer and binary crossentropy was chosen as the loss function.
7. Data was split into 60% train and 40% test using train_test_split.
8. The model was evaluated using binary accuracy metric for both train and test. The final results were ~82% accuracy for both.`