Overview:

The goal of this homework assignment is for you to demonstrate your ability to create a quick classifier based on customer chat data, and as a stretch goal to stand it up for inference using TensorFlow Serving.

For this assignment you will be working with a small sample of our training data (customer_chat_sample.csv). We'd like you to write a ML pipeline using Python and TensorFlow (feel free to use Keras) that:

1. Reads the dataset into memory
2. Computes or uses pre-trained word embeddings. Some available ones are:
    a. BERT word embeddings using TensorFlow hub (or feel free to download your own): https://tfhub.dev/google/collections/bert/1
    b. GloVe word embedding: there are a number of TensorFlow tutorials online for this
3. Trains any classifier to predict the `label` class
4. Freezes the graph for serving

Optional: Expose the trained model in TensorFlow Serving

Expected Output

The final result should be a Python script that we can run locally to replicate your work. We would like to be able to understand your approach to training the model as well as the steps you take to achieve the best accuracy of the model.

Once you complete the assignment, please send it back to Chris Henk (chris.henk@carvana.com) with any instructions on how to run the script. In a possible next Zoom or on-site interview, we'll ask you to give a short presentation about your approach and other techniques you would have tried if you had more time.

Good luck!