Q1-B] Neural Language Models (30 points): The main goal of the problem 1.B is to enable you to learn a neural language model on Google cloud. You should visit:
https://github.com/r-mal/utd-nlp/blob/master/neural_language_modeling_glove.ipynb

To open the notebook on Google's cloud, the students can just click the blue 'Open in Colab' button at the top of the webpage. Where we have prepared for you a framework for a simple feed-forward neural language model. You are provided with the Reuters newswire corpus that contains the text of 11,228 newswires from Reuters. These are split into 8,982 newswires for training and 2246 newswires for testing. You are instructed how to prepare your data, download the embeddings, and build the neural model. You are asked to train and test the neural model as a feedforward network with two intermediate or "hidden" layers, between the input and output (10 points), which is provided, as well as with one hidden layer (10 points) and three hidden layers (10 points). This will enable you to use the sparse categorical cross entropy loss function, which is provided. To obtain full credit for each model, you are requested to (1) generate a validation set, (2) train and evaluate the model; (3) create a graph indicating the change of accuracy and loss of the model over time and (4) provide the perplexity values of the model.

Solution:

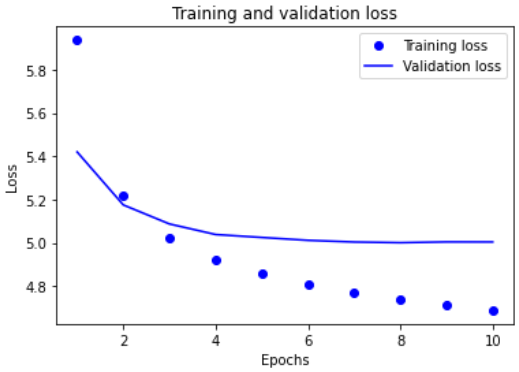
1. For 2 layers: (that were provided)

```
dense_19 (Dense)            (None, 16)                4016

dense_20 (Dense)            (None, 10000)             170000
```

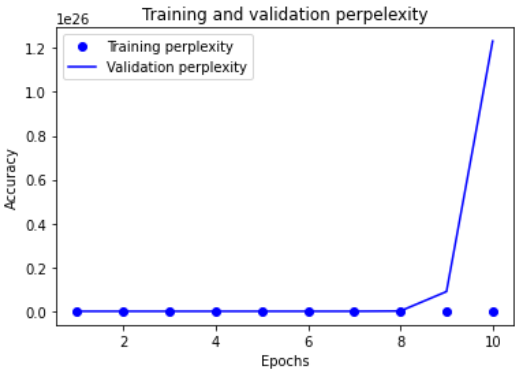
Evaluation:

```
loss: 4.9957
perplexity: 211298356093095625359 36.0000
acc@5: 0.3926
```

Loss over time:



Accuracy over time:



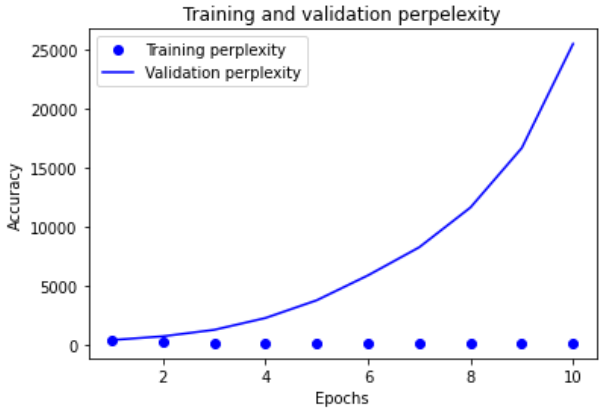
2. For 1 layer: (By removing the 16 node layer)

```
dense_18 (Dense)             (None, 10000)             2510000
```

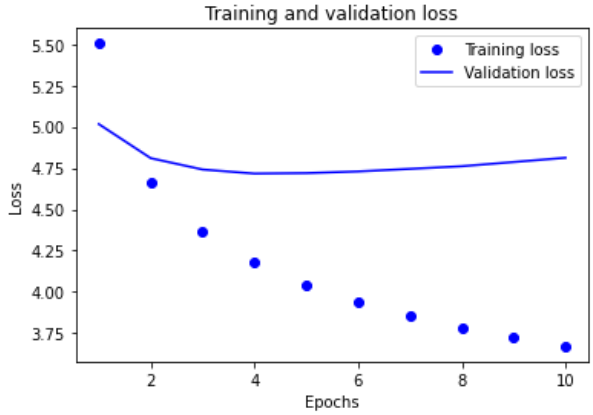
Evaluation:

```
loss: 4.7894
perplexity: 49192.6133
acc@5: 0.4415
```

Accuracy over time:



Loss over time:



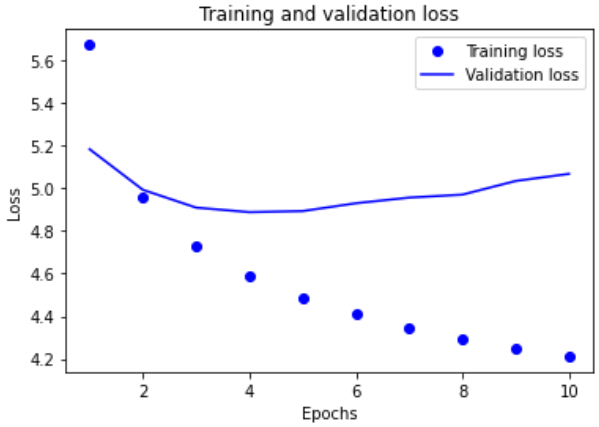
1. For 3 layers:

```
dense_15 (Dense)             (None, 16)                4016

dense_16 (Dense)             (None, 128)               2176

dense_17 (Dense)             (None, 10000)             1290000
```

Evaluation:

```
loss: 5.0839
perplexity: 97030163403224971066804503838 72.0000
acc@5: 0.4187
```

Accuracy over time:



Loss over time: