

# CS378 assignment4

Qinxin Wang

April 15, 2020

## 1 RNNs for Classification

**Q1** (1) I implemented a RNN model to predict whether a string is followed by consonants or vowels. In my code, I created a class RNN to build up my classification model. It consists of an embedding layer, a single-layer GRU and a linear layer for classification. In my forward method, I take the the last hidden layer as output and pass it through a linear layer, then predict the probability over two classes.

In order to load training data efficiently, I inherited the class Dataset to load provided cons and vowel examples, and used a Dataloader to load them in batch.

(2) I set learning rate to  $1e-3$  and used Adam as optimizer. After training for 10 epoches in 42 seconds, the accuracy is 77.1%.

## 2 Implementing a Language Model

**Q2** (1) I implemented the language model by training every input character to predict its next character. I chunked the input data with size=50, which means that the input data is a string starting with a space as start symbol and followed by 49 characters. The target output is the 50 character string. I took the output layer of RNN for every input character, and classify them on 27 characters.

I implemented `get_log_prob_sequence` by feeding the previous output hidden state as the input hidden state to the next prediction, which can save much time. For every classification output, I calculated the log probability by `LogSoftmax` on the target output and sum up over `next_chars`.

(2) I trained the RNN Language Model for 15 epochs with learning rate  $1e-3$ . The final perplexity and average log probability are 4.37 and -1.47 respectively. The model is trained in 41.7 seconds.