

Assignment #2: RNN Practice

Youngseok Yoon

20172911

Computer Science and Engineering, POSTECH

yys8646@postech.ac.kr

1. Introduction

RNN practice with city name classification dataset.

2. Dataset

The city name dataset have nine classes; training data have 3000 instances in each class and validation data have 100 instances in each class.

3. Model

In this assignment, three models are tested: RNN, LSTM and GRU, each have one hidden layer and one output layer. The number of hidden layer's node is 128 dimension. From city name character length, the input feature have 46 dimension and the number of class is nine. Weights are randomly initialized with uniform distribution $\mathcal{U}[-\sqrt{k}, \sqrt{k}]$ with $k = \frac{1}{in_features}$

4. Experiment

Each model trained with 10000 iteration and Adam optimizer with learning rate 10^{-3} (see Fig. 1).

5. Discussion

Validation accuracy of three model is $GRU > LSTM > RNN$.

References

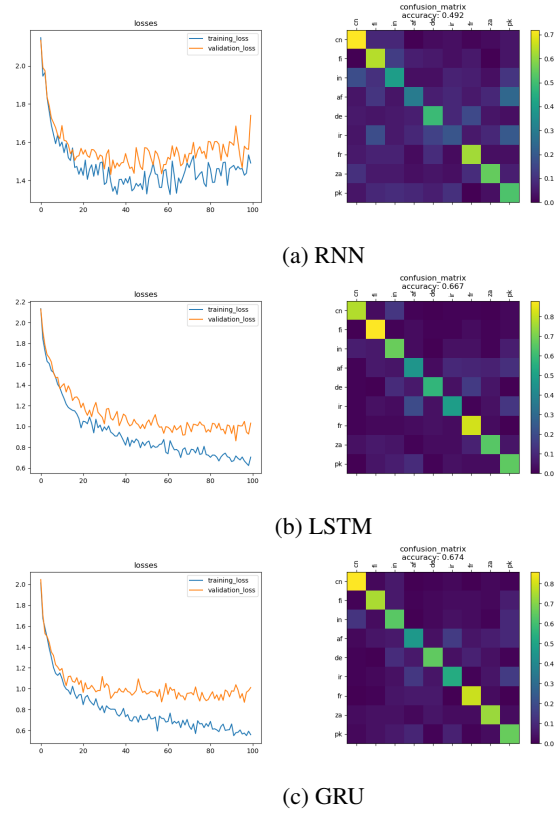


Figure 1: Loss function and Confusion matrix