Text Generation with LSTM Units

By Me

June 4, 2017

- I wanted to replicate the results shown in the blog post "The Unreasonable Effectiveness of Recurrent Neural Networks".

Introduction

- I wanted to replicate the results shown in the blog post "The Unreasonable Effectiveness of Recurrent Neural Networks".
- I also wanted to have fun.

How Neural Networks Work

- In general:

$$f_w : (-1, 1)^n \longmapsto (-1, 1)^m$$

How Neural Networks Work

- In general:

$$f_w : (-1, 1)^n \longmapsto (-1, 1)^m$$

- Or:

$$f_w : (0, 1)^n \longmapsto (0, 1)^m$$

How Neural Networks Work

- A neural network consists of layers.
- The first layer receives the input, transforms it, and passes it on to the next layer.
- Each subsequent layer receives an array from the previous layer, transforms it, and passes it on.
- The output of the last layer is the output of the neural network.

How Neural Networks Work

- A neural network consists of layers.
- The first layer receives the input, transforms it, and passes it on to the next layer.
- Each subsequent layer receives an array from the previous layer, transforms it, and passes it on.
- The output of the last layer is the output of the neural network.
- Transformation: Multiply the input with a matrix and apply a function to each element of the result.
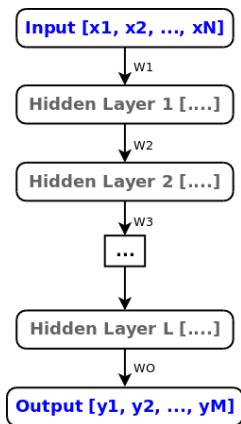
How Recurrent Neural Networks Work
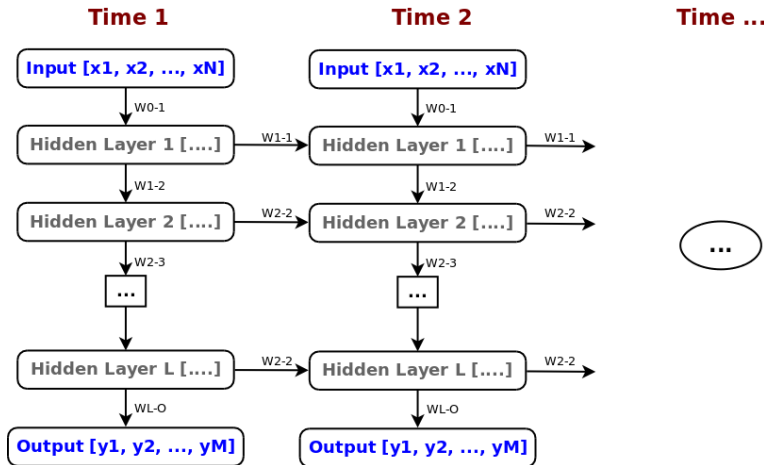
What Kinds of Layers are Needed

- I needed three kinds of layers:
  - Tanh Layers.
  - LSTM Layers.
  - Softmax Layers.

- It is a simple mapping: $f_w : (-1, 1)^n \longmapsto (-1, 1)^m$.

- It is also a mapping: $f_{w,s} : (-1, 1)^n \longmapsto (-1, 1)^m$.
- But it has an internal state, meaning that previous runs of the neural network may influence the output.

How Does a Softmax Layer Work

- It is a mapping: $f_w : (-1, 1)^n \longmapsto (0, 1)^m$.
- Where the output values add up to one, so that the output can be interpreted as a probability distribution.

The end