WXML Final Report: Optimizing neural networks

Aleksandr Aravkin
Peng Zheng

Rasika Bhalerao
Yijia Liu
Yijia Xue
N'Vida Yotcho

Spring 2016

1 Introduction

We can think of Artificial neural networks (**Figure 1**) to be a set of brains, in which each brain is driven by a unique function for a different purpose. Since their introduction in 1943, neural networks and deep learning became the hot topic of machine learning. They used in almost all large-

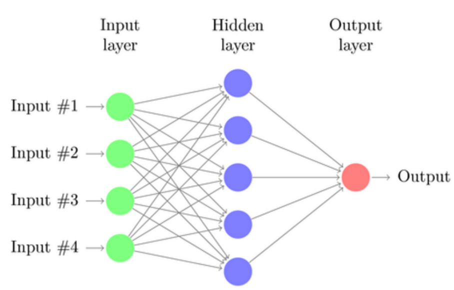

Figure 1: Simple Neural Network

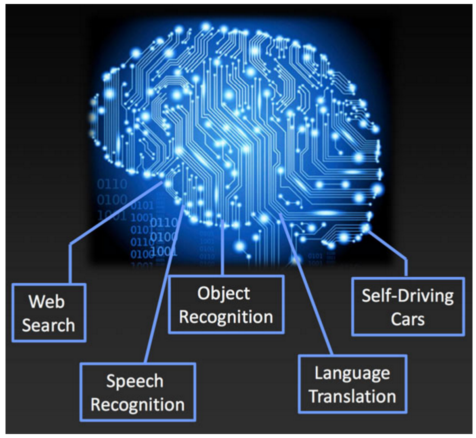

Figure 2: Brain Neural Network

scale inference problems (**Figure 2**). For example, they are used in image segmentation, speech recognition, regression, classification, and compression.

1.1 The initial problem

We decided to study and optimize the artificial neural network as a stock price predictor and as a speech recognizer.

1.2 Directions

Speech and stock prices are time series so that the recurrent neural network was the suitable model to work with. Indeed, the recurrent neural network creates an internal state of network that enables dynamic temporal behavior.

1.3 Theoretical

1. Stock Prediction Let's explain more about the structure of Recurrent Neural Nets. Our recurrent neural network mode are two hidden layers to capture the volatility of the stocks better.

    - **Input:** time series of historical stocks' returns in daily basis and other data, USD index and oil price.

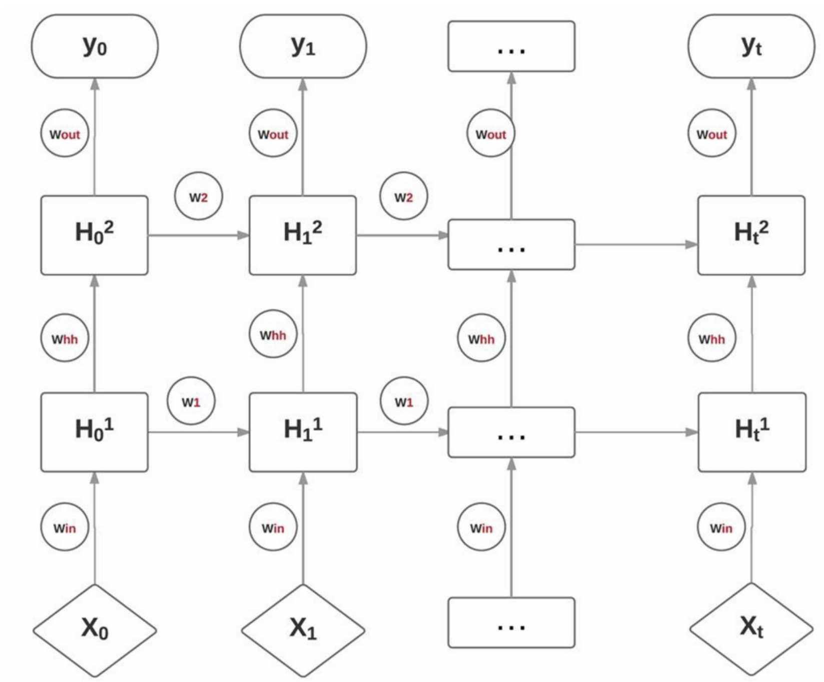

Figure 3: Recurrent Neural Network

- **Training weights:** solve the optimization problem to minimize discrepancy output and the real data over all weights(Circles, see in Figure 3).
- **Output:** time series of the relative change of stock prices(returns) in the near future.

2. Speech Recognition

- **Input:** audio of speech
- **Phoneme segmentation (Figure 4) and filtering**
  The audio is fragmented in phonemes so that each piece of audio contains one phoneme.
- **Phoneme recognition**
  Each audio is then translated in a sequence of phonemes (text)
- **Grammar model**
  The sequence of phonemes (text) is fully translated in word (text)

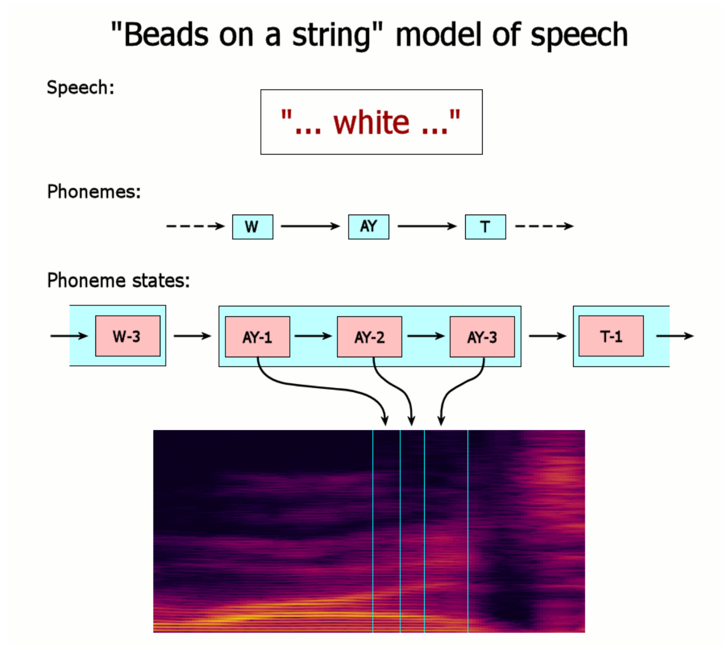

Figure 4: Example of phoneme segmentation

1.4 Computational

1. Stock Prediction

   - the inputs and the outputs are all in big matrices.
   - the tanh active function is used as the recurrent neural nets function
   - regulate the optimal function, the error between the predicted output and the real data, in order to avoid over-fitting.

2. Speech Recognition
   We worked on elaborating a more robust automatic phoneme segmentation algorithm that could minimize phonemes overlap

2 Progress

1. Stock Prediction

- We apply two hidden layers in Recurrent Neural Nets
- for the stock's trend of each training data that has been processed by Recurrent Neural Nets is close to the trend of actual data.
- for the stock's trend of each predicted data that has been processed by Recurrent Neural Nets is only as accurate as in the first three days.

2. Speech Recognition
   - for four out five speakers, the SA2 audio was successfully cut in sub-audios corresponding to the the 10 words in the SA2 sentence.
   - for 2 out five speakers, the SA2 audio was successfully cut in sub-audios corresponding to the phonemes and or pauses of the SA2 speech.

3 Future directions

1. Stock Prediction
   - find out more stock's data
   - categorize the data
   - improve the algorithm to predict the stock's price for long-term period(more general and accurate)
2. Speech Recognition
   - generalize the phoneme segmentation algorithm to work for the whole TIMIT corpus.
   - study the whole set of English phonemes to elaborate a more precise algorithm
   - elaborate an algorithm to differentiate unvoiced phoneme to voiced phoneme
   - study speech formation to explore or elaborate other tools for phoneme segmentation
   - build an automatic acoustic model using neural network