# FRE 9733 Big Data in Finance Week 14 Homework

Mengheng Xue

May 7, 2019

## 1 Smallest Model for CNN

In this section, I try to find the CNN model with least parameters while keeping the prediction accuracy above 95%. I only adjust parameters in convolution layers and dense layer which are relevant to parameters counting. Based on my test, my final smallest CNN structure layers with parameters counting are

- **Convolution Layer 1 :** kernel size $[2,2]$, stride 1, filter number 3, so number of parameters in this layer is $(2 \times 2 + 1) \times 3 = 15$.

- **Convolution Layer 2 :** kernel size $[2,2]$, stride 2, filter number 3, so number of parameters in this layer is $(2 \times 2 \times 2 + 1) \times 3 = 27$.

- **Dense Layer :** inputs 65, outputs 10, so number of parameters in this layer is $(65 + 1) \times 10 = 660$

The total number of parameters is $15 + 27 + 660 = 702$, and the prediction accuracy I obtain is 0.9527.

## 2 Activation Function

Based on my previous model, I change different activation functions in convolution layers and dense layer, the prediction performances are shown as follows

| test | C1 | C2 | D | accuracy |
|------|----------|----------|----------|----------|
| 0 | IDENTITY | IDENTITY | RELU | 0.9527 |
| 1 | RELU | RELU | RELU | 0.9430 |
| 2 | SIGMOID | SIGMOID | RELU | 0.7995 |
| 3 | SIGMOID | SIGMOID | SIGMOID | 0.1032 |
| 4 | TANH | TANH | RELU | 0.9236 |
| 5 | TANH | TANH | TANH | 0.9272 |

Based on my test, when both convolution layers and dense layer use activation function RELU will achieve the best performance. RULE and TANH seem to be much better than SIGMOID as nonlinear activation functions in this case. However, we can find that if we don't add nonlinear activation functions in convolution layers will achieve best performance which is shown in case 0.

## 3 Pooling Type

Based on the previous best performance setting, I test two different pooling types, the performances are shown as follows

| test | C1 | C2 | D | PoolingType | accuracy |
|------|----------|----------|------|-------------|----------|
| 0 | IDENTITY | IDENTITY | RELU | MAX | 0.9527 |
| 1 | IDENTITY | IDENTITY | RELU | AVG | 0.9118 |

We can see that pooling type MAX will achieve better performance than AVG.