Financial Time Series

Mahdi Behroozikhah
Guideline

1. Problem & Motivation
2. Common Approaches
3. Datasets
4. High-Frequency Trading Strategy Based on Deep Neural Networks
   A. Inputs
   B. NN Architecture
   C. Result
5. A deep learning framework for financial time series using stacked autoencoders and long-short term memory
   1. Wavelet Transforms
   2. Stacked Autoencoders
   3. NN Architecture
   4. Result
6. Conclusion & Discussion
1. **Problem & Motivation**

2. Common Approaches

3. Datasets

4. High-Frequency Trading Strategy Based on Deep Neural Networks
   A. Inputs
   B. NN Architecture
   C. Result

5. A deep learning framework for financial time series using stacked autoencoders and long- short term memory
   1. Wavelet Transforms
   2. Stacked Autoencoders
   3. NN Architecture
   4. Result

6. Conclusion & Discussion
What is The Financial Data?
Candle-Sticks

- Time
- Volume
- Index Name
- High Price
- Low Price
- Close Price
- Open Price
Goal?

Making money by Predicting the future!
Well... What is the problem then?

Chaotic Market, Dynamic Updates, Many Variables
Still have some hopes...
Two Sigma, Wall street, D.E. Shaw, ...
Let’s get started!
Guideline

1. **Problem & Motivation**

2. **Common Approaches**

3. Datasets

4. High-Frequency Trading Strategy Based on Deep Neural Networks
   A. Inputs
   B. NN Architecture
   C. Result

5. A deep learning framework for financial time series using stacked autoencoders and long-short term memory
   1. Wavelet Transforms
   2. Stacked Autoencoders
   3. NN Architecture
   4. Result

6. Conclusion & Discussion
Approaches

• Fundamental Analysis

  ➡ News data as source (e.g. Bloomberg)

  • Jan 13, 2014 - Google Acquires Smart Thermostat Maker Nest For for $3.2 billion.

  ➡ NLP and CNN As a solution

  • (Actor = Google, Action = acquires, Object = Nest, Time = Jan 13, 2014)
Approaches

- Technical Analysis
  - Technical Indicators
    - Moving Average, Relative Strength Index, …
  - Price Action
    - Trend lines, Japanese candlesticks
- Elliott Wave
Common Methods

- Support Vector Machine (SVM)
- Neural Network
- Random Forest
1. **Problem & Motivation**

2. **Common Approaches**

3. **Datasets**
   4. High-Frequency Trading Strategy Based on Deep Neural Networks
      A. Inputs
      B. NN Architecture
      C. Result
   5. A deep learning framework for financial time series using stacked autoencoders and long-short term memory
      1. Wavelet Transforms
      2. Stacked Autoencoders
      3. NN Architecture
      4. Result

6. **Conclusion & Discussion**
Datasets

• Forex
  • USD/JPY, EUR/GBP

• Stock Market
  • S&P 500, Apple

• Cryptocurrency
  • Bitcoin, Ethereum, …
Putting the World’s Money into Perspective

Total value of gold is 200x the total value of Bitcoin

Bill Gates’ net worth is 2x the entire market capitalization of Bitcoin

All Cryptocurrencies
Bill Gates
Larry Page
Bitcoin

USD in Circulation
Gold Market Cap
Physical Money
Stock Markets
All Money

$83.6T
$66.8T
$31T
$8.2T
$1.5T
$730B
$402B
$100B
$86B
$41B
$41B

Sources:
https://howmuch.net/articles/worlds-money-in-perspective
https://coinmarketcap.com
https://www.forbes.com
https://www.federalreserve.gov
https://www.cia.gov

1 All Money = money in any form including bank or other deposits as well as notes and coins.
2 Physical Money = money in forms that can be used as a medium of exchange, generally notes, coins, and certain balances held by banks.
Evaluation

- Return value & profits
- Portfolio management
- Accuracy & Mean Square Error
1. Problem & Motivation

2. Common Approaches

3. Datasets

4. High-Frequency Trading Strategy Based on Deep Neural Networks
   A. Inputs
   B. NN Architecture
   C. Result

5. A deep learning framework for financial time series using stacked autoencoders and long-short term memory
   1. Wavelet Transforms
   2. Stacked Autoencoders
   3. NN Architecture
   4. Result

6. Conclusion & Discussion
High-Frequency Trading Strategy Based on Deep Neural Networks

Andrés Arévalo¹(✉), Jaime Niño¹, German Hernández¹, and Javier Sandoval²

¹ Universidad Nacional de Colombia, Bogotá, Colombia
{ararevalom,jhninop,gjhernandezp}@unal.edu.co
² Algocodex Research Institute, Universidad Externado, Bogotá, Colombia
javier.sandoval@uexternado.edu.co

Abstract. This paper presents a high-frequency strategy based on Deep Neural Networks (DNNs). The DNN was trained on current time (hour and minute), and \( n \)-lagged one-minute pseudo-returns, price standard deviations and trend indicators in order to forecast the next one-minute average price. The DNN predictions are used to build a high-frequency trading strategy that buys (sells) when the next predicted average price is above (below) the last closing price. The data used for training and testing are the AAPL tick-by-tick transactions from September to November of 2008. The best-found DNN has a 66 % of directional accuracy. This strategy yields an 81 % successful trades during testing period.
Input Data

- Apple Stock One Min from Sep. to Nov. 2008

- Features:
  - Current time
  - Last $n$ pseudo-log-returns
  - Last $n$ standard deviation of prices
  - Last $n$ trend indicators
Method & Training

- NN: I, 1, 4I/5, 3I/5, 2I/5, I/5
- Tanh on each layer
For each trading minute

At the beginning of the minute, to forecast the next one-minute average price

Current closing price < predicted price
To buy the stock at the current price

Current price >= predicted price OR current minute ended
No
Yes
To sell the stock at the current price

Current closing price > predicted price
To sell the stock at the current price

Current price <= predicted price OR current minute ended
No
Yes
To buy the stock at the current price

Fig. 6. Strategy flowchart.
## Results

<table>
<thead>
<tr>
<th>Window Size</th>
<th>DNN Architecture</th>
<th>Maximum</th>
<th>Minimum</th>
<th>Mean</th>
<th>$\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>MSE</td>
<td>DA (%)</td>
<td>MSE</td>
<td>DA (%)</td>
</tr>
<tr>
<td>2</td>
<td>8 8:6:4:3:1 1</td>
<td>0.07832</td>
<td>65.71328</td>
<td>0.06768</td>
<td>61.63236</td>
</tr>
<tr>
<td>3</td>
<td>11 11:8:6:4:2 1</td>
<td>0.07678</td>
<td>66.15492</td>
<td>0.06823</td>
<td>63.67759</td>
</tr>
<tr>
<td>4</td>
<td>14 14:11:8:5:2 1</td>
<td>0.09158</td>
<td>65.71328</td>
<td>0.07197</td>
<td>63.30659</td>
</tr>
<tr>
<td>5</td>
<td>17 17:13:10:6:3 1</td>
<td>0.10132</td>
<td>66.05024</td>
<td>0.07569</td>
<td>64.30565</td>
</tr>
<tr>
<td>6</td>
<td>20 20:16:12:7:3 1</td>
<td>0.10512</td>
<td>65.74816</td>
<td>0.07574</td>
<td>62.99267</td>
</tr>
<tr>
<td>7</td>
<td>23 23:18:13:9:4 1</td>
<td>0.10383</td>
<td>65.63154</td>
<td>0.08251</td>
<td>63.22400</td>
</tr>
<tr>
<td>8</td>
<td>26 26:20:15:10:5 1</td>
<td>0.09873</td>
<td>65.60865</td>
<td>0.07813</td>
<td>63.41123</td>
</tr>
<tr>
<td>9</td>
<td>29 29:23:17:11:5 1</td>
<td>0.09020</td>
<td>65.49197</td>
<td>0.07628</td>
<td>63.78227</td>
</tr>
<tr>
<td>10</td>
<td>32 32:25:19:12:6 1</td>
<td>0.10250</td>
<td>65.50401</td>
<td>0.07400</td>
<td>63.09731</td>
</tr>
<tr>
<td>11</td>
<td>35 35:28:21:13:6 1</td>
<td>0.10565</td>
<td>65.24773</td>
<td>0.07702</td>
<td>62.90997</td>
</tr>
<tr>
<td>12</td>
<td>38 38:30:22:15:7 1</td>
<td>0.09440</td>
<td>65.32961</td>
<td>0.07746</td>
<td>63.69026</td>
</tr>
<tr>
<td>13</td>
<td>41 41:32:24:16:8 1</td>
<td>0.09442</td>
<td>64.61967</td>
<td>0.07437</td>
<td>61.89811</td>
</tr>
<tr>
<td>14</td>
<td>44 44:35:26:17:8 1</td>
<td>0.09833</td>
<td>65.32961</td>
<td>0.07849</td>
<td>62.39972</td>
</tr>
<tr>
<td>15</td>
<td>47 47:37:28:18:9 1</td>
<td>0.09871</td>
<td>64.86392</td>
<td>0.08201</td>
<td>61.86322</td>
</tr>
</tbody>
</table>

**Fig. 9.** Cumulated profit of the trading strategy
1. Problem & Motivation

2. Common Approaches

3. Datasets

4. High-Frequency Trading Strategy Based on Deep Neural Networks
   A. Inputs
   B. NN Architecture
   C. Result

5. A deep learning framework for financial time series using stacked autoencoders and long-short term memory
   1. Wavelet Transforms
   2. Stacked Autoencoders
   3. NN Architecture
   4. Result

6. Conclusion & Discussion
A deep learning framework for financial time series using stacked autoencoders and long-short term memory

Wei Bao¹, Jun Yue²*, Yulei Rao¹

¹ Business School, Central South University, Changsha, China, ² Institute of Remote Sensing and Geographic Information System, Peking University, Beijing, China

* jyue@pku.edu.cn
Input financial time series

S(J)  D(J)  ...  D(j)  ...  D(2)  D(1)

Denoised financial time series

Multi-resolution discrete wavelet transformation

Stacked autoencoder

long-short term memory
Wavelet Transform

- Noise reduction
- Consider frequency & time simultaneously
- Haar function as the wavelet basis
- $O(n)$ runtime complexity
- Applied twice
Stacked autoencoders
Inputs

• From 1st Jul. 2008 to 30th Sep. 2016

• Developed Market
  • S&P500 and DJIA index in New York

• Between Developed and Developing
  • Hang Seng index in Hong Kong and & Nikkei 225 index in Tokyo

• Developing
  • CSI 300 and Nifty 50 index In India
<table>
<thead>
<tr>
<th>Name</th>
<th>Definition/Implication</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A. Daily Trading Data</strong></td>
<td></td>
</tr>
<tr>
<td>Open/Close Price</td>
<td>nominal daily open/close price</td>
</tr>
<tr>
<td>High/Low Price</td>
<td>nominal daily highest/lowest price</td>
</tr>
<tr>
<td>Trading volume</td>
<td>Daily trading volume</td>
</tr>
<tr>
<td><strong>Panel B. Technical Indicator</strong></td>
<td></td>
</tr>
<tr>
<td>MACD</td>
<td>Moving average convergence divergence: displays trend following characteristics and momentum characteristics.</td>
</tr>
<tr>
<td>CCI</td>
<td>Commodity channel index: helps to find the start and the end of a trend.</td>
</tr>
<tr>
<td>ATR</td>
<td>Average true range: measures the volatility of price.</td>
</tr>
<tr>
<td>BOLL</td>
<td>Bollinger Band: provides a relative definition of high and low, which aids in rigorous pattern recognition</td>
</tr>
<tr>
<td>EMA20</td>
<td>20 day Exponential Moving Average</td>
</tr>
<tr>
<td>MA5/MA10</td>
<td>5/10 day Moving Average</td>
</tr>
<tr>
<td>MTM6/MTM12</td>
<td>6/12 month Momentum: helps pinpoint the end of a decline or advance</td>
</tr>
<tr>
<td>ROC</td>
<td>Price rate of change: shows the speed at which a stock’s price is changing</td>
</tr>
<tr>
<td>SMI</td>
<td>Stochastic Momentum Index: shows where the close price is relative to the midpoint of the same range.</td>
</tr>
<tr>
<td>WVAD</td>
<td>Williams’s Variable Accumulation/Distribution: measures the buying and selling pressure.</td>
</tr>
<tr>
<td><strong>Panel C. Macroeconomic Variable</strong></td>
<td></td>
</tr>
<tr>
<td>Exchange rate</td>
<td>US dollar Index</td>
</tr>
<tr>
<td>Interest rate</td>
<td>Interbank Offered Rate</td>
</tr>
</tbody>
</table>
Data Splitting
The diagrams depict the price trends of various indices over time. Each chart shows multiple lines representing different models or data series, labeled as "Actual Data," "WSAE-LSTM," "WLS-LSTM," "LSTM," and "RNN." The x-axis represents the trading day, and the y-axis shows the price scale for each index.

1. CIX 300 Index
2. Nifty 50 Index
3. Hang Seng Index
4. Nikkei 225 Index
5. S&P/PS Index
6. DJIA Index
<table>
<thead>
<tr>
<th></th>
<th>CSI 300 Index</th>
<th>Nifty 50 Index</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Year</strong></td>
<td><strong>Year 1</strong></td>
<td><strong>Year 2</strong></td>
</tr>
<tr>
<td><strong>WSAEs-LSTM</strong></td>
<td>46.428</td>
<td>59.580</td>
</tr>
</tbody>
</table>

**Panel B. Relatively developed market**

<table>
<thead>
<tr>
<th></th>
<th>Hang Seng Index</th>
<th>Nikkei 225 Index</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Year</strong></td>
<td><strong>Year 1</strong></td>
<td><strong>Year 2</strong></td>
</tr>
<tr>
<td><strong>WSAEs-LSTM</strong></td>
<td>75.844</td>
<td>81.890</td>
</tr>
<tr>
<td><strong>WLSTM</strong></td>
<td>25.697</td>
<td>19.221</td>
</tr>
</tbody>
</table>

**Panel C. Developed market**

<table>
<thead>
<tr>
<th></th>
<th>S&amp;P 500 Index</th>
<th>DJIA Index</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Year</strong></td>
<td><strong>Year 1</strong></td>
<td><strong>Year 2</strong></td>
</tr>
<tr>
<td><strong>WSAEs-LSTM</strong></td>
<td>71.316</td>
<td>48.351</td>
</tr>
</tbody>
</table>
Discussion

• Adding more technical indicators such as moving average in the input could be helpful?

• Trend direction probability as a classification VS Exact value as a regression?
Thank you for your attention