Is Technical Indicator Really a Thing for Deep Learning?

107302002金融四張致中

ABSTRACT

自19世紀後基於道氏理論，激發出了後現代技術分析，並以此為基礎進行分析和利用，在發展了幾個世紀後，就成為了現在常人所使用的現代技術指標。技術分析相比起新聞、事件等外部訊息，偏向分析市場價格所反映出的相關資訊。「歷史不會重複，但是會押韻」這句話反映出了技術分析的價值所在:投資者會集體做出某種特定行為，使得資產價格移動行為因此擁有某種重複的模式。但是在現在日常交易裡，單純運用技術指標來做交易並不是一個能夠穩定獲利的行為，其非常依賴個人判斷技術以及經驗法則。技術指標有許多的參數，依據每個人對於資產標的理解的不同，會使用上不同的技術指標和參數。因此參數調教這件事情對於量化交易使用者來說變得相當重要，這不僅僅是程式撰寫的問題，也牽涉到他們過往的交易經驗。

本次實驗想要探討的是技術指標對於量化交易的表現究竟是有正面的影響還是負面的影響，實驗呈現方法會以深度學習模型呈現，預計使用至少四種模型作為量化交易之依據:一、simple baseline:使用fully-connected network，以及OHLC資料作為基準。二、baseline with technical indicator:使用和(一)相同的模型，但是在資料裡加入不同的技術指標，以獲得更多的feature幫助判斷。三、strong baseline:運用convolutional layer當作feature extractor，先對OHLC資料進行feature extraction，並再將資料通過fully-connected network作為預測。四、strong baseline with technical indicator:使用和(三)相同的模型，但是在資料裡加入不同的技術指標。其他模型:如果有時間，可能會參考其他論文的模型。

目標資料室採取幣安交易所的ETH/BTC、BNB/BTC、LTC/BTC、RVN/BTC以及BTC/USDT交易對資料為主。選用加密貨幣作為研究標的之理由如下:資料取得容易。由於股市資料在沒有特定券商帳號的情況下，難以取得交易週期小於一天的資料，使得資料量嚴重不足。至於使用比特幣交易對做為大部分的資料是因為以比特幣為訂價價格較穩定且合理。比特幣是最早的加密貨幣，被視為幣圈中的通用貨幣，也因為如此大部分的加密貨幣交易所都會放上比特幣的交易對。由於加密貨幣屬於新興市場，在美元計價的情況下容易有大漲大跌的情況發生，但是比特幣的交易對表現就相當穩定，故採用此交易對。

本次研究預期之結果，根據[1]之研究，應是模型(三)>模型(四)>模型(二)>模型(一)，但是實際結果還是得因應模型建構之完成度來做判斷。

DATA

本次實驗用的資料大部分採取了以BTC為底的五分鐘以及一小時交易對，以避免加密貨幣對法幣波動度過大影響模型的判斷。每份資料包含了Timestamp, Open, High, Low, Close以及Volume等六個欄位。在資料的處理上，每個交易對依照時間順序被分成了Training Data、Validating Data以及 Testing Data。這些資料的長度則是參考了論文中的作法:Training Data總共包含約兩年的資料，Validating Data包含了約兩個月的資料，而Testing Data則是兩星期的資料。除了最原本的資料以外，也用了不同的技術指標所生成的資料，包含了最基本的RSI、EMA，也包含了布林通道、Chaikin、True Range等其他包括不同時間週期，總共34種技術指標。除了最基本的資料，也做了其他的preprocessing以及labeling。

Data Preprocessing:

我對資料做了兩種處理，第一個是MinMax Scaler，第二種則是Rolling Window。會使用MinMax Scaler是因為他常常被用在處理這種在一定值以內變動的數值，如外匯價格。有鑑於資料都是以BTC為基底的交易對，可視為一種類似外匯的存在，且技術指標通常不會超過一個值，故採用之。

Rolling Window則是另一種常見處理時間序列資料的方式。這裡使用每24個data為一Rolling Window，並用此資料去預測兩個小時後價格的漲跌。

Data Labeling:

有關股票趨勢預測的Label方法有很多種，大部分採用的方法是下一期的價格是否高於現在的價格，但是由於加密貨幣高波動度的特性，很容易會導致摩擦成本的提高，所以在這裡加入了波動度做為評鑑價格移動趨勢的指標之一。所以下一期的價格必須得高於某個門檻後才會被標為「上漲」，若價格低於另外一個門檻則是會被標為「下跌」，其餘則是「不變」。在這裡波動度

是直接使用了一個Rolling Window的波動度。

MODEL

本次實驗使用了兩種不同的模型，分別是使用Fully-connected layer 建成的MLP model，以及[1]裡所使用的跨類別一維卷積模型(Cross datatype 1D CNN model，以下簡稱CDT-1D CNN)。

在MLP model中，共有六層的hidden layer。每層分別有300, 200, 100, 50, 25, 10個neurons，同時也使用了0.00001的L2 regularization，weight initialization則是都使用了random，activation function都是使用relu。除此之外，我也在每一層的後面加上Dropout layer，dropout rate設為0.7。Output layer則是有3個neurons，並使用了softmax function。梯度下降法使用了Adam，並使用其預設參數。其他模型只要有使用到類似的配置，都會使用一樣的超參數。

至於CDT-1D CNN模型，他是由三層卷積和池化層組成，其中較為特殊的是相對於傳統作法的1D CNN會一口氣將所有資料卷積在一起，導致原本應該不相關的資料混合在一起，降低判斷力，該模型會將每種不同的資料逐行掃過，除了能夠保持每種資料種類的獨立性以外，還能夠有parameter sharing的效果。這三層的卷積層分別使用了32, 64, 128個filter，filter size則是4, 3, 2，padding使用了same以避免input size改變。池化層的pool size和stride使用了4, 3, 2。所以input data的維度會在經過處理後由原本的24\*n\*1變為1\*n\*128，隨後通過flatten layer後丟入兩層Fully-connected layers，並預測其漲跌。

MLP model在時間序列資料裡的問題就是每個時間點的資料對於模型來說都是獨立的，模型因此不會考慮時間前後的相關性。至於CDT-1D模型則是因為將前後的資料卷積在一起而使得連續時間點下的前後資料會被模型考慮到。

EXPERIMENTS

本次實驗使用了兩種不同時間周期的資料，分別是5分鐘和1小時的資料。5分鐘的資料參考了[1]的時間週期，使用了一樣多的資料做為training set、validation set，以及testing set。至於1小時的資料則是使用了兩年的資料做為training set，兩個月的資料做為validation set，兩星期資料做為testing set。在這裡和[1]不同，不是以每兩個小時為周期才測一次，而是每個Timestep就拿前面24筆資料做測試，所以能夠做testing 的資料就相對較多。訓練模型的Batch size只用了12，是為了能夠有更加stochastic的梯度下降過程。

CONCLUSION

最終的結果如下面四張表所示，其中BTCUSDT的交易對的training、validating以及testing的正確率全部都是1。這是因為在做data labeling時是和其他項目使用一樣的threshold，但是其他項目都是以BTC為底的交易對，因此BTCUSDT的波動會明顯大於其他項目。所以在labeling的階段因此大部分的資料很可能都被大量標註了”Up”或”Down”，導致沒辦法用模型對資料做很好的解釋。

在技術指標使用與否的部分可以看出技術指標的使用其實是會增加模型預測的準確度，這點和[1]的結果是相反的。目前初步預估是因為加密貨幣市場相較於期貨市場缺乏一定的穩定性，時常因為**事件**的發生而暴漲暴跌，導致在價格上並沒有特定規律，**增加資料的噪音**，進而影響沒有技術指標資料模型的判斷。這點可以從實驗結果裡看出，在沒有技術指標資料下表現其實都是遜於有技術指標的資料的。

最後至於兩模型的比較，可以看出CDT-1D模型整體表現在這裡是稍微優於MLP模型，但是並沒有相差太多。很有可能是因為資料的雜訊真的太多導致模型沒辦法起到很好的效果。

|  |  |  |
| --- | --- | --- |
| MLP | Testing Accuracy | |
| 5m data | With TI | Without TI |
| BNBBTC | 0.6986 | 0.6301 |
| BTCUSDT | 1 | 1 |
| ETHBTC | 0.9178 | 0.0822 |
| LTCBTC | 0.3699 | 0.3699 |
| RVNBTC | 0.2329 | 0.2329 |

Table1: MLP model with 5 min data

|  |  |  |
| --- | --- | --- |
| CDT-1 | Testing Accuracy | |
| 5m data | With TI | Without TI |
| BNBBTC | 0.6438 | 0.6301 |
| BTCUSDT | 1 | 1 |
| ETHBTC | 0.6849 | 0.0822 |
| LTCBTC | 0.3699 | 0.3699 |
| RVNBTC | 0.6164 | 0.4384 |

Table2: CDT-1D model with 5 min data

|  |  |  |
| --- | --- | --- |
| MLP | Testing Accuracy | |
| 1h data | With TI | Without TI |
| BNBBTC | 0.524 | 0.524 |
| BTCUSDT | 1 | 1 |
| ETHBTC | 0.6741 | 0.6741 |
| LTCBTC | 0.5623 | 0.5623 |
| RVNBTC | 0.3035 | 0.3035 |

Table3: MLP model with 1 hour data

|  |  |  |
| --- | --- | --- |
| CDT-1D | Testing Accuracy | |
| 1h data | With TI | Without TI |
| BNBBTC | 0.524 | 0.2109 |
| BTCUSDT | 1 | 1 |
| ETHBTC | 0.6741 | 0.6741 |
| LTCBTC | 0.5815 | 0.4281 |
| RVNBTC | 0.5240 | 0.2428 |

Table4: CDT-1D model with 1 hour data

FUTURE

未來我應該還會去嘗試不同的方法來處理加密貨幣的資料，包括:

**資料處理方法改變:**

這裡我用的scaler是minmax scaler，之後也許可以改用每個timesteps的價格變化量做為scaler來處理價格資料，說不定會能夠比較好的解釋資料，尤其是那些以USDT為底的交易對。

**資料的噪音處理:**

由於加密貨幣還處在很早期的階段，和其他傳統金融市場相比，價格變化的噪音是相當大的。所以若真的非得做加密貨幣市場的處理，也許可以加上Kalman filter之類的濾波器來處理雜音，增加資料的可被解釋性。

**模型的改進:**

這裡我參考了[1]並成功的複製出了他們的模型。但是這個模型的缺點其實很明顯，由於只使用了CNN，所以無法讓相隔時間過長的資料同時被模型考慮。因此也許可以考慮在該模型的基礎上，加入LSTM、GRU，甚至是Self-attention來加強模型的長期預測能力。

REFERENCES

[1] Jia Wang, Tong Sun, Benyuan Liu, Yu Cao, Degang Wang: Financial Markets Prediction with Deep Learning；2021 April

[2] Qingsong, Wen, Liang Sun, Fan Yang, Xiaomin Song, Jingkun Gao, Xue Wang, Huan Xu: Time Series Data Augmentation for Deep Learning: A Survey；2021 Sep