

LSTM-RNN を用いたイベント考慮後の株価時系列予測

Predicting Equity Price With Corporate Action Events Using LSTM-RNN

南 正太郎 *1

Shotaro Minami

*1 あすかアセットマネジメント株式会社

Asuka Asset Management Co., Ltd

The forecasting the stock price of a particular has been a difficult task for many of analysts and researchers. In fact, investors are highly interested in the research area of stock price prediction. However, to improve the accuracy of forecasting a single stock price is a really challenging task, therefore in this paper, I propose a sequential learning model for prediction of a single stock price with corporate action event information and Macro-Economic indices using LSTM-RNN method. The results show the proposed model is expected to be a promising method in the stock price prediction of a single stock with variables like corporate action and corporate publishings.

1. はじめに

株式資産運用の実務においては、株価を予測することは重要な仕事の一つである。最終的には株価は需給によって決定するものの、企業の将来の業績を調べることで売買を判断する投資家もいれば、株価の動きだけを見て判断する投資家もいる。また、PER や PBR といった株式指標を参考に割安割高を判断する投資家もいるだろう。好きな企業だから買うという投資家もいるかもしれない。このようにさまざまな見通しや考え方をを持った市場参加者によって株価は決まるため、将来の株価を予測するのは簡単ではない。

企業の将来の業績から株価を予測する手法には、例えば将来の株式配当を現在価値に割り引くことで求める配当割引モデルや企業が事業を行った結果生み出される将来キャッシュフローの現在価値を求める割引キャッシュフローモデルなどがある。また、株価は EPS (一株当たり純利益) \times PER (株価収益率) によって構成されていると考える倍率法では、 PER は変わらないと想定した場合、将来の業績が好調のため EPS が倍になるのであれば、株価も単純に倍になると考えられる。少なくともその株価水準に長期的には収斂されていくだろうと期待される。

個別企業を分析する人が、多くの企業で四半期ごとに公表される企業の決算発表や日々のプレスリリースに着目するのは、業績の進捗を知ること、これまでの見通しが正しかったかどうかを判断したり、今後の企業の成長がどの程度期待できるのかを見極める意味がある。将来にわたって企業価値を高めると投資家の多くが考える情報があれば株価は上昇するからである。

このように、将来の株価を予測する上では、企業の決算発表やプレスリリースといった情報はとても重要な意味を持つ。また、過去の株価はこのような情報に基づいて判断してきた投資家の思惑が含まれているといえるだろう。

本稿で提案する LSTM-RNN (Long Short Term Memory Recurrent Neural Network) は、再帰的ニューラルネットワーク (Recurrent Neural Network) の一種であり、通常の RNN において長期の系列データを学習する際に発生する勾配問題を回避することから、長期の時間依存も短期の時間依存も学習

できる特性がある。従来の RNN では考慮できない、過去の情報をネットワーク内に保持し、必要なタイミングでのみ取得・置換したいとの課題から登場してきた背景がある。

発表された情報がどの時点の株価に影響しているのかは知ることは難しいものの、そのような状況を織り込んだ株価予測を行う上では、LSTM-RNN は有効な手段になるものと思われる。

2. 関連研究と本稿について

LSTM-RNN を用いた株価予測を行った関連研究について説明する。

株価だけに応用した研究としては、Murtaza らは、インドの取引所の NIFTY50 銘柄の株価について、Open、Close、High、Low それぞれのデータを組み合わせたとときの LSTM-RNN による株価の予測精度について比較している。結果としては4つのデータを用いて予測を行った場合が最も望ましいことを示した。また、Qun らは、個別銘柄の寄り付きの株価について、上海総合指数とフォーラム内の情報から取得した感情データを変数にして予測することを行っている。結論としては従来型の RNN を適用するよりも良好な結果が得られたようだ。

LSTM-RNN は時系列データ分析において適用する場合、効果が高いことが知られているが、株価を予測するという点においてもモデルの設定の仕方によっては先の論文のように有効な予測モデルの手段になりえる可能性がある。

そこで本稿では、LSTM-RNN を用いることで企業のイベントを織り込んだ株式予測の手法を提案する。具体的には、特定の企業を取り上げ、決算発表日、コーポレートアクションのプレスリリース、企業が投資家向けに公表している受注残高を変数に用いて、将来株価を予測する。

3. LSTM-RNN を用いた株価時系列予測

3.1 モデル

x_t は t 時点の入力値 (ベクトル)、 h_t は出力値とするとき、LSTM-RNN は以下に定式化される。

$$f_t = \sigma_g(W_f x_t + U_f h_{t-1} + b_f) \quad (1)$$

$$i_t = \sigma_g(W_i x_t + U_i h_{t-1} + b_i) \quad (2)$$

$$o_t = \sigma_g(W_o x_t + U_o h_{t-1} + b_o) \quad (3)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \sigma_c(W_c x_t + U_c h_{t-1} + b_c) \quad (4)$$

$$h_t = o_t \odot \sigma_h(c_t) \quad (5)$$

c_t はメモリセルの状態ベクトル、 W 、 U 、 b はパラメータ行列とベクトルである。また、 f_t は古い情報を記憶する重みベクトルである忘却ゲート、 i_t は新しい情報を得るための重みベクトルである入力ゲート、 o_t は出力候補のベクトルを出す出力ゲートを表す。活性化関数のうち、 σ_g はシグモイド関数、 $\sigma_{c,h}$ は双曲線正接関数である。

3.2 データについて

分析対象として、株式会社ツガミ（証券コード：6101）の株価を取り上げ分析を行った。工作機械メーカーであるツガミは、自動旋盤機を作っている世界的ニッチ企業である。生産と売り上げの大半は中国であることから、企業業績は国内の受注だけでなく海外の受注に大きく依存している。また、国内市場が縮小する中、海外へ事業の基盤を移す政策を進めており、また同時に明確な資本政策も発表していることから、生産工場の移管や自己株式の取得、ストックオプションなど積極的なコーポレートアクションを公表している。シクリカル銘柄にも分類されることがあるツガミは、中小型株であることに加えて、マクロ指標（工作機械受注高）や企業が月中に発表する前月の国内と海外の受注高推移、コーポレートアクションのプレスリリースに株価が反応しやすい特徴がみられる。

今回分析に用いる変数は、株価に加えて、企業がHP上に公表するイベント情報（プレスリリース）と企業の国内および海外の受注残高、内閣府が提供している工作機械受注高である。イベント情報は、決算発表日などの法定開示に関するものと、自己株式取得、ストックオプションの発行などコーポレートアクションに関するものを分けて変数にしている。

なお、日次で分析するために月次および四半期のマクロ指標については、多項式スプライン補完を行った。使用したデータの期間はデータ間で共通に取得できた2013年4月15日から9月11日までである。イベント情報については、データクリーニングの過程でIDとして数値にエンコーディングしている。データが公表された翌日から予測に使えるようにラグを考慮している。全期間1082日のうち700日をトレーニング期間、3割にあたる残りの期間をテスト期間とした。

3.3 モデルの検証

実運用を想定するならば実績値と乖離がないほうが望ましい。そこで、RMSE(Root Mean Squared Error)を使ってモデルの評価をすることにした。

$$RMSE = \sqrt{\frac{1}{N} \sum (\hat{y}_t - y_t)^2} \quad (6)$$

\hat{y}_t は予測値、 y_t は実績値。

3.4 実装

1999年にHochreiterらによって提唱されたLSTM-RNNに基づいている。計算するにあたって、学習時には、Epoch数を50、100、500の3つのパターンで計算。隠れ層のユニット数は50、LSTM層は1とした。学習のフレームワークとしては、フロントエンドにKeras、バックエンドにはTensorflowを使用して実装している。確率的勾配降下法の最適化を行う上で、Adam(Adaptive Moment Estimation)により学習率の設

定を行っている。また、GPUはGeForce GTX 1060 3GBを使用した。

4. 検証結果

表1は、説明する変数に株価のみを用いた場合、イベントのみを用いた場合、企業が公表する受注残高とマクロ指標を用いた場合、イベント及び受注残高、マクロ指標のすべてを用いた場合について、それぞれのRMSEにつきEpoch数別にまとめたものである。イベントを考慮した場合は、株価のみを用いた予測よりも総じて予測誤差が小さい。マクロ指標を用いた場合は、Epoch数が500の時のみ予測誤差が小さくなっていく。すべてを変数にした場合は、むしろ精度が低くなってしまった結果となった。運用実務においては、実績値の上昇・下降をどのくらい予測できているかが重要になることから、一致率も表2に掲載している。どの分析においても50%に満たないことからあまり実用的ではないものの、株価のみを使用した結果よりも一致率が高くなるパターンもあるようだ。ただし、使用したデータが少ないことから、改善する可能性はある。

RMSE	株価のみ	イベント	マクロ指標	すべて
Epoch = 50	19.5	18.9	19.9	20.4
Epoch = 100	17.1	16.9	17.6	18.1
Epoch = 500	15.0	14.9	14.8	15.3

表1: RMSE

一致率	株価のみ	イベント	マクロ指標	すべて
Epoch = 50	0.4842	0.4868	0.4711	0.4737
Epoch = 100	0.4895	0.4789	0.4921	0.4947
Epoch = 500	0.4579	0.4526	0.4553	0.4711

表2: 一致率

5. 結論

本研究は、LSTM-RNNを用いて特定の企業が出すイベント情報や受注残高を織り込んだ上での将来の株価を予測するというを試みた。説明する変数に株価のみを用いた場合、イベントのみを用いた場合、企業が公表する受注残高とマクロ指標である工作機械受注高を用いた場合、イベント及び受注残高、マクロ指標のすべてを用いた場合を比較して株価の予測値と実績値の差がどの程度異なるのかについて比較検討した。結果としてはイベントのみを用いた予測が最も実績値との誤差が小さいことが分かった。Epoch数を変えて計算してもその傾向は変わらないことから、当該企業についてはイベント情報が将来株価を予測する上で有効かもしれない。ただし、本稿は、このようなイベントを考慮する時系列予測の提案が目的であり、分析の精度を高めるためにはより長期にわたる分析を行うことが望ましいだろう。

本研究で提案したモデルは金融の世界では幅広い応用の可能性を持っている。株価に影響を与える変数は企業によって異なるであろうから、個別企業ごとに分析してみるのも興味深い。また、例えば、国内に限らず上場企業は法定開示資料や決算報告など定期的な情報の開示が必須であることから海外にも

同様の分析を行うことができる。今回の研究においては単にイベントがあったという情報だけであるが、その情報にポジティブ・ネガティブの属性を持たせたものでもよい。むしろそのほうが株価を説明する材料としては適切なのではないだろうか。株価を予測するだけでなく、企業が開示する受注状況からマクロ指標を予測することにも応用できるであろう。最適な自己株式の取得のタイミングを評価することに応用できれば企業にとっても役に立つ可能性もある。今後の課題である。

参考文献

- [南 17] 南正太郎, 光定洋介: AI の資産運用への応用の可能性と限界, 証券アナリストジャーナル, Vol.55, 8月号 (2017).
- [Jason 17] Jason Brownlee: Long Short-Term Memory Networks With Python Develop Sequence Prediction Models With Deep Learning, MACHINE LEARNING MASTERY (2017).
- [Argimiro 14] Argimiro Arratia: Computational Finance An Intoductory Course with R, ATLANTIS PRESS (2014).
- [Akhter 14] Akhter Mohiuddin Rather: A Hybrid Intelligent Method of Predicting Stock Returns, Hindawi Publishing Corporation, Advances in Artificial Neural Systems Volume 2014, Article ID 246484, 7 pages (2014).
- [Murtaza 2015] Murtaza Roondiwala, Harshal Patel, Shraddha Varma: Predicting Prices Using LSTM, International Journal of Science and Research, 2319-7064 (2015).
- [Qun 2017] Qun Zhuge, Lingyu Xu and Gaowei Zhang: LSTM Neural Network with Emotional Analysis for Prediction of Stock Price, Engineering Letters, 25:2, EL_25_2_09 (2017).