

深層学習を用いた株価動向推定

○ 片倉 賢治 , 高橋 大志 (慶應義塾大学)

Analyzing Stock markets through Deep Learning

* K. Katakura and H. Takahashi (Keio University)

Abstract— In recent years, a technique called deep learning is drawing attention in the field of machine learning. In this paper, we utilize the deep learning method to analyze financial markets. This study also makes attempts to investigate the relationship between sentiment indices and factor returns, which have been widely used in finance to elucidate the mechanism of price fluctuations.

Key Words: Deep Learning, Capital Asset Pricing, Sentiment Analysis

1 はじめに

社会システムにおける金融市場の役割は大きい。世界的な金融危機以降、先進諸国のみならず多くの国々において経済金融政策運営は大きな転換点を迎えており、個々の金融機関の健全性を確保するだけでは、金融システム全体としての安定を必ずしも実現できるわけではないとの見方が強まっている。このような中、例えば、金融安定のためのマクロプルーデンス政策、インフレ目標の設定、物価安定、ゼロ金利政策、量的緩和政策等、様々な政策がなされており金融システムへの関心は益々増している。

金融システムにおいては、さまざまなステークホルダーが存在するが、主要なステークホルダーの一つとして企業を挙げることができる。企業は様々な金融に関する意思決定を行っているが、その中で、企業価値評価は重要な要素の一つとなっている。企業価値の最大化は、株式会社 (Corporate) として組織された企業経営者の意思決定において拠り所とすべき基準であり¹⁾、投資や資金調達といったコーポレートファイナンスの領域における企業の財務意思決定も企業価値の最大化、株主価値の最大化が主要な目的である²⁾。

企業価値評価を行うためにはいくつかの要素が必要となるが、とりわけ資本コスト (Cost of Capital) は重要な役割を果たす。この資本コストは、(1) 投資家や株主にとっての期待リターン、(2) 資金調達コスト、(3) 投資判断基準収益率 (キャッシュフロー割引率)、(4) 業績評価基準 (ハードルレート)、といった意味を含み、企業価値評価にとって必要不可欠な要素である。日本において金融ビッグバンによる直接金融の規制緩和等、金融市場の変化によって、従来の企業統治 (コーポレートガバナンス) システムが変化し、企業価値や株主重視の声が高まってきたが、この資本コストを意識した経営指標である EVA®* (Economic Value Added) やキャッシュフローを導入する企業が相次いで増加した³⁾。その意味においても資本コストの重要性は大きい。

資本コストの推計においては、一般にマーケットデータをを用いた推計が行われるが、その際、資本資産価

格評価モデル (CAPM: Capital Asset Pricing Model) や、Fama-French の 3 ファクターモデルといったモデルが広く用いられている^{4,5)}。また、CAPM は、合理的な投資家等を仮定したモデルであるが、必ずしも合理的でない投資家を考慮した行動ファイナンス (Behavioral Finance)⁶⁾ に関する議論も近年関心を集めており、その議論に広がりを見せている^{7,8,9)}。更に、近年の情報処理技術の向上等を背景とし、市場参加者が利用可能な情報は飛躍的に増大しており、従来からの分析手法に加え、新たな手法によって分析する要望が高まっている。

このような中、新たな分析手法の一つとして、ニュースや Twitter 等マイクロブログの言語情報や機械学習手法を用いた分析が関心を集めつつあり、これまで数多くの研究報告が行われている。例えば、Twitter フィードから二種類の手法、(1) Opinion Finder による Positive, Negative の 2 値分類、(2) Google-Profile of Mood States (GPOMS) による Calm, Alert, Sure, Vital, Kind, Happy の 6 値分類、で個人の感情を分類し、それらの情報を用いて、投資意思決定に及ぼす影響について分析したもの¹⁰⁾ や、ニュース記事の言語情報を SVM (Support Vector Machines) により分類し、株価動向に関して行った分析¹¹⁾ といった報告などが挙げられる。このように、機械学習手法を取り入れ、金融市場を説明しようとする分析も数多く提案されている。

機械学習の分野において数多くの手法が提案されてきているが、とりわけ、近年、深層学習 (Deep Learning) と呼ばれる手法が関心を集めている。深層学習は、従来手法と比較し、圧倒的に高い性能を達成しており、その性能の高さから、数多くの海外および国内の研究者が、意欲的に研究に取り組んでいる¹²⁾。

これらを背景として、本研究では、言語情報処理に深層学習手法を用いた資産価格変動の分析を行う。日本の株式市場を分析対象とし、ファイナンス分野において広く用いられているファクターモデル及びマクロ指標を取り込んだ分析を行い、価格変動メカニズムの解明を試みる。本分析は、深層学習を通じて資産価格変動メカニズムの解明を試みるものであり、当手法の資産価格分野における有用性を示すものである。

* EVA は、米スターン・スチュワート社の登録商標である。

2 深層学習と自然言語処理

深層学習とは、多層に重ねたニューラルネットワークを用いて抽象度の高い複雑なデータの表現を学習する手法であり、表現学習 (Representation Learning) とも呼ばれている¹³⁾。多層ニューラルネットワークは過去にも提案されてきたが、学習の際に訓練データのみを学習し未知データに対しては適合できないといった過学習(Overfitting)が起りやすく、学習に膨大な計算が必要なため非常に長時間かかるといった、学習の困難さや計算処理能力が課題であった。しかしながら、昨今、(1) 大幅なチップ処理能力の増加、HWコストの顕著な低下、大規模データ獲得の容易さ等といった背景的な進展、(2) 膨大なデータを多層ニューラルネットワークに入力し、層ごとに事前学習 (pre-training) という教師なし学習を行うことによる過学習の回避、(3) 事前学習後、ネットワーク全体の微調整を行う教師あり学習 (fine-tuning) 、によってこれらの課題解決が提案された^{14,15,16)}。これらによって学習精度が高まっており、画像認識分野¹⁷⁾、音声認識分野¹⁸⁾、さらに化合物反応予測等において、従来手法と比較して圧倒的な性能を達成しており、様々な分野への応用が試みられている。

これは、自然言語処理分野 (NLP, Natural Language Processing) においても例外ではない。単語を固定長ベクトルで表現し、類似する単語が類似したベクトルを持つよう意味を埋め込み (word embedding) 表現する分散表現 (Distributed Representation)^{19,20)}、言語における単語の連りの条件付き確率を学習するNNLM (Neural Net Language Model) と呼ばれる統計的言語モデル²¹⁾、そして、品詞タグ付け、構文解析 (Parse) を行うチャンキング、固有名詞等を抽出する固有表現抽出、動作主や相手、対象等に単語を分類するための意味役割付与、構造解析といったNLPタスクを統一的に扱う畳み込みアーキテクチャー (Convolutional Neural Network) の適用²²⁾、といった様々なタスクにおいて、深層学習の圧倒的に優れた精度が報告されている。とりわけ、構成的意味論^{23,24)} 分野において、Socherらが提案したRNTN (Recursive Neural Tensor Network) の応用である句、文の評価極性分類手法²⁵⁾ は、Stanford Sentiment Analysisとも呼ばれており、従来のRNN (Recursive Neural Network)^{26,27)} やMV-RNN (Matrix-Vector Recursive Neural Network)²⁸⁾、ナイーブベイズ (Naive Bayes)、SVM等の分類性能を超え、長文に対する評価極性分類で最先端の精度80%~85.4%を達成している。

RNTNは、文章を単語群から成る二分木構造によって表現しており、入力として一様分布からランダムなサンプリングによって生成した単語ベクトルを使用している。そして、中間ノードでテンソル (Tensor) と呼ばれる極性の重み付け演算を行い、ボトムアップ方式で句や文の評価極性を合成するモデルであり、これによってパラメータの指数的増加を防ぐことに成功している。従来のRNNには、テンソルという概念はなく、否定文では高い精度を達成することが出来なかった。また、MV-RNNは、RNNに比べて高い性能を持つものの扱う文書の語彙数によってパラメータ数が膨大になるという課題があった。

RNTNの学習データは、きめ細かいセンチメントラ

ベルが付与された、215,154句、11,855文のレビューデータを用いており、Pang and Lee(2005)²⁹⁾ のデータを利用したものとなっている。デモ、訓練・テストコード、Stanford Sentiment Treebankのデータセットが、Webページ (<http://nlp.stanford.edu/sentiment/>) において公開されている。

本研究では、これら手法を用い株式市場を対象とした分析を試みるものである。

3 データ

本稿では、テキストデータおよび市場関連データを用い分析を行う。テキストデータに関しては、世界のマーケット動向に関するニュースとして、Thomson Reuters社提供のNews Feed Direct (NFD)を用いた。NFDは、News Scope Directとしても知られており、ニュースのヘッドラインや経済イベントを極小の遅延で配信し、発表時刻もミリ秒単位で保持している等、分析に適した特徴を有している。

サンプル期間は、2003年1月1日~2012年7月31日とした。NFDは、世界各国の市場を対象としたニュースが含まれており、ニュースの言語も英語、フランス語、ドイツ語、日本語などをはじめ多岐にわたる。このような膨大な量のニュースの中から本分析では、とりわけ、日本市場及び日本企業に関する英語ニュース記事411,531件を対象として分析を行った。日次のNFDのニュースの件数は、最もニュースが多い日で947件、最もニュースが少ない日で0件となり、平均すると117.6件であった (Fig. 1)。また、対象記事の内、日本企業に関連する記事は、363,970件、該当する企業数は308件であった。

また、市場関連データについては、日経NEEDSおよびThomson Reuters Datastreamより入手した。また、本分析ではファクターリターンを対象とした分析を行う。ファクターモデルに関するデータ (以下、FFファクター) は、久保田、竹原(2007)³⁰⁾ に従い、日本における東証1部、東証2部から構成される銘柄から算出した3ファクター (マーケットファクターである Rm-Rf、Small minus BigのSMBファクター、High minus LowのHMLファクター) データを用いた。

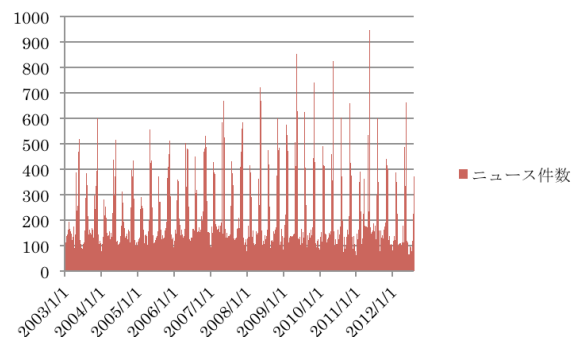


Fig. 1: 日次 NFD 件数.

4 分析方法

本研究では、金融市場において配信されているニュー

ースとマーケットとの関係性を調査する。具体的には日次NFDをStanford Sentiment Analysisモジュールによってセンチメント・インデックスを算出し、マーケットのファクターリターンとの関係性に関して分析を行う。(Fig. 2)

Methodology

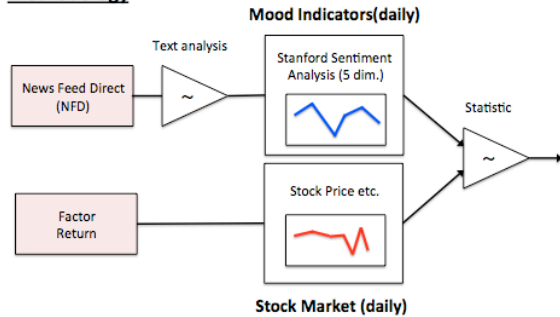


Fig. 2: システム概要図.

具体的な分析方法は次のとおりである。

- (1) NFDを日次記事群として1日分を1ファイルとして抽出する。ただし、1日分の記事群は、東証の取引時間(9:00~15:00)を考慮して、あるt日に発表された記事群は、日本標準時+0900(JST)で t-1日(前日)の15:00以降、t日(当日)の15:00迄の間に発表された記事を含めるものとした。
- (2) 全ての日次記事群に対して、URLやEメールアドレス、記事途中のピリオドの除外処理を行う。
- (3) 全ての日次記事群を1ファイル毎に学習済みのStanford Sentiment Analysisモジュールにインプットする。
- (4) モジュールがアウトプットした評価極性分類を結果ファイル(1日分1ファイル)として保存する。
- (5) 結果ファイル中の評価極性毎に設定したセンチメントスコア(評価極性のスコアは、Very Positive (+2), Positive (+1), Neutral (0), Negative (-1), Very Negative (-2)として設定)を合計し、日次センチメント・インデックスを算出する。
- (6) 算出した日次センチメント・インデックスとファクターリターンとの関係性を調べる。日次センチメント・インデックスは、他手法との比較可能性を考慮し、Bollen(2011)¹⁰⁾に従い、Z-scoreに規格化した時系列データを算出した。Z-scoreの定義式を以下に示す。

$$Z_{X_t} = \frac{X_t - \bar{x}(X_{t \pm k})}{\sigma(X_{t \pm k})}$$

ここで、 X_t はスコアの時系列データである。 $\bar{x}(X_{t \pm k})$ 、 $\sigma(X_{t \pm k})$ は、それぞれ、k日間平均、k日間標準偏差である。

5 分析結果

2003年~2012年の9年間のセンチメント・インデック

スを算出し、同期間におけるファクターリターンとの相関関係を統計的手法により分析する。Fig.3 および Fig.4は、2003年1月1日~2003年3月8日の日次NFD記事群を用いて、センチメント・インデックス(Z-score)を計算したものである。本分析では、それぞれ以下の条件のものについて分析を行っている。

- (1) 東証取引時間内のみの記事群を対象とした分析 (Fig. 3).
- (2) 東証取引時間内外全ての記事群を対象とした分析 (Fig. 4).

図の横軸は、記事の発表日を表し、縦軸はセンチメント・インデックスのZ-score値を表す。

Fig.3 取引時間内の記事群を対象とした日次センチメントZ-scoreは、東証取引日に発表された記事を対象として算出したZ-scoreであるが、全体的に-1.0~+1.0の間に収まっていることを確認できる。図より、Z-scoreの値がとりわけ大きい日があることを確認できるが、これらの日においては、何らかのイベントが生じている可能性がある。

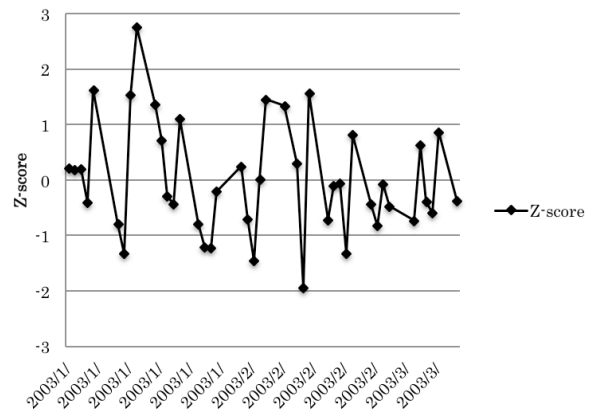


Fig. 3: 取引時間内の記事群を対象とした日次センチメント Z-score.

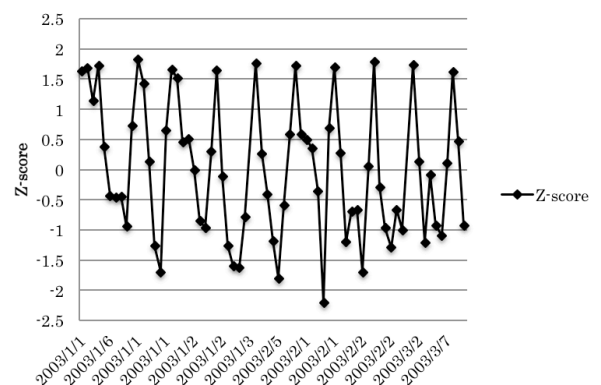


Fig. 4: 取引時間内外全ての記事群を対象とした日次センチメント Z-score.

一方、Fig.4 取引時間内外全ての記事群を対象とした日次センチメント Z-score においては、東証取引日以外の休業日(日曜日、国民の祝日、前日及び翌日が国民の祝日である日、土曜日、年始3日間及び12月31日)

を含む全ての記事群から Z-score を算出したものである。Fig.3 と比較してこれらの値は、一定周期で大きな値になっていることがわかる。これは、休業日に発表される記事数が取引時間内と比較して少ないことが要因の1つと考えられる。時系列分析手法等を用いた詳細な分析は、今後の課題である。

6 まとめおよび今後の予定

本分析においてセンチメント・インデックスの挙動について分析を行い、インデックスの値がとりわけ大きい日があることなどの結果を確認した。センチメント・インデックス株式リターンの関連性についての詳細な分析は、今後の課題である。また、従来の機械学習手法との比較等も今後の課題として挙げられる。

参考文献

- 1) 久保田敬一：決定版コーポレートファイナンス。東洋経済新報社, (2006).
- 2) Brealey, Richard A., Myers, S., and Allen, F.: Principles of corporate finance., Tata McGraw-Hill Education, (2012).
- 3) 加護野忠男, 砂川伸幸, 吉村典久. コーポレート・ガバナンスの経営学: 会社統治の新しいパラダイム. 有斐閣, (2010).
- 4) Jensen, Michael C.: Some anomalous evidence regarding market efficiency. , Journal of financial economics 6.2 , 95/101. , (1978).
- 5) Fama, Eugene F., and French, Kenneth R.: Common risk factors in the returns on stocks and bonds. , Journal of financial economics 33.1, 3/56. (1993).
- 6) Shleifer, Andrei. Inefficient markets : An introduction to behavioral finance. , Oxford university press, (2000).
- 7) 砂川伸幸, 山崎尚志.: マーケットの非効率性と企業の投資・財務戦略. , 国民経済雑誌 186.3 , 65/77, (2002).
- 8) Ikenberry, David, Lakonishok, Josef, and Vermaelen, Theo.: Market underreaction to open market share repurchases. , Journal of financial economics 39.2 , 181/208, (1995).
- 9) DeLong, J. Bradford.: Noise trader risk in financial markets. , (1989).
- 10) Bollen, Johan, Huina Mao, and Xiaojun Zeng.: Twitter mood predicts the stock market. , Journal of Computational Science 2.1 , 1/8, (2011).
- 11) Schumaker, Robert P., and Hsinchun Chen.: Textual analysis of stock market prediction using breaking financial news: The AZFin text system. , ACM Transactions on Information Systems (TOIS) 27.2, 12, (2009).
- 12) 吉原輝, 藤川和樹, 関和広.: 深層学習による経済指標動向推定. , 人工知能学会全国大会論文集 28 , 1/4, (2014).
- 13) Bengio, Yoshua, Aaron Courville, and Pascal Vincent.: Representation learning: A review and new perspectives. , Pattern Analysis and Machine Intelligence, IEEE Transactions on 35.8: 1798/1828. (2013).
- 14) G.E. Hinton, S. Osindero, and Y. Teh.: A fast learning algorithm for deep belief nets, Neural Computation, vol 18, (2006).
- 15) Bengio, Yoshua, et al.: Greedy layer-wise training of deep networks. , Advances in neural information processing systems 19: 153. (2007).
- 16) Poultney, Christopher, Sumit Chopra, and Yann L. Cun.: Efficient learning of sparse representations with an energy-based model. , Advances in neural information processing systems. (2006).
- 17) Le, Quoc V.: Building high-level features using large scale unsupervised learning. , Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on. IEEE, (2013).
- 18) Deng, Li, et al.: Recent advances in deep learning for speech research at Microsoft. , Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on. IEEE, (2013).
- 19) Hinton, Geoffrey E. "Learning distributed representations of concepts." Proceedings of the eighth annual conference of the cognitive science society. Vol. 1. (1986).
- 20) Mikolov, Tomas, et al.: Distributed representations of words and phrases and their compositionality. , Advances in Neural Information Processing Systems. (2013).
- 21) Bengio, Yoshua, et al.: Neural probabilistic language models. , Innovations in Machine Learning. Springer Berlin Heidelberg, 137/186, (2006).
- 22) Collobert, Ronan, and Jason Weston.: A unified architecture for natural language processing: Deep neural networks with multitask learning. , Proceedings of the 25th international conference on Machine learning. ACM, (2008).
- 23) Socher, Richard, et al.: Dynamic pooling and unfolding recursive autoencoders for paraphrase detection. , Advances in Neural Information Processing Systems. (2011).
- 24) Socher, Richard, et al.: Semi-supervised recursive autoencoders for predicting sentiment distributions. , Proceedings of the Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics, (2011).
- 25) Socher, Richard, et al.: Recursive deep models for semantic compositionality over a sentiment treebank. , Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP). (2013).
- 26) Goller, Christoph, and Andreas Kuchler.: Learning task-dependent distributed representations by backpropagation through structure. , Neural Networks, 1996., IEEE International Conference on. Vol. 1. IEEE, (1996).
- 27) Socher, Richard, et al.: Parsing natural scenes and natural language with recursive neural networks. , Proceedings of the 28th International Conference on Machine Learning (ICML-11). (2011).
- 28) Socher, Richard, et al.: Semantic compositionality through recursive matrix-vector spaces. , Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning. Association for Computational Linguistics, (2012).
- 29) Pang, Bo, and Lillian Lee.: Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales. , Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics. Association for Computational Linguistics, (2005).
- 30) 久保田敬一, 竹原均.: Fama-french ファクターモデルの有効性の再検証. , 現代ファイナンス 22, 3/23 , (2007).