

集団化する人工知能 鈴木教授による解説シリーズ ⑧ ～深層学習による株価予測（前編）～

お伝えしたいポイント

2018年8月2日

- ・ 2タイプの深層学習
- ・ 画像型の深層学習（目が見えるAI）
- ・ AIによる株価チャート分析（画像認識）
- ・ いじわる手書き文字認識実験
- ・ 非画像型の深層学習（回帰分析）

<2タイプの深層学習>

近年はAIブームの最中にありますが、飛躍的に向上した技術は深層学習（Deep Learning）です。画像からの物体認識や、AIスピーカーのように人間とのコミュニケーションの精度が高まりました。前者の画像認識には主に畳込みニューラルネットワーク（以下、画像型と呼ぶ）が用いられ、

画像以外にはリカレントニューラルネットワークやオートエンコーダ（以下、非画像型と呼ぶ）が用いられる傾向にあります。それぞれ用途や特徴が異なるため、株価予測の観点からどちらが有用なのか考察します。



出所：大和証券投資信託委託

当資料のお取り扱いにおけるご注意

■当資料は、ファンドの状況や関連する情報等をお知らせするために大和投資信託により作成されたものであり、勧誘を目的としたものではありません。■当資料は、各種の信頼できると考えられる情報源から作成していますが、その正確性・完全性が保証されているものではありません。■当資料の中で記載されている内容、数値、図表、意見等は当資料作成時点のものであり、将来の成果を示唆・保証するものではなく、また今後予告なく変更されることがあります。■当資料中における運用実績等は、過去の実績および結果を示したものであり、将来の成果を示唆・保証するものではありません。■当資料の中で個別企業名が記載されている場合、それらはあくまでも参考のために掲載したものであり、各企業の推奨を目的とするものではありません。また、ファンドに今後組み入れることを、示唆・保証するものではありません。

販売会社等についてのお問い合わせ⇒大和投資信託 フリーダイヤル 0120-106212(営業日の9:00～17:00) HP <http://www.daiwa-am.co.jp/>

<画像型の深層学習 (目が見えるAI)>

まず画像型の深層学習を考えてみます。仕組みの詳細には立ち入りませんが、畳込み層による特徴抽出とプーリング層によるデータ圧縮を繰り返すことで、認識対象に共通する特徴を絞っていく学習方法です。画像認識が得意なため、AIに視覚を与えることができます。すると人間の代わりにAIができる仕事の幅が広がります。

たとえば医療画像診断において、AIによる自動化が注目されています。CT検査では大量のX線断層画像を診断する必要があるため、人間では

疲労による見落としリスクが問題視されています。しかしAIならば疲れません。安定かつ高速で24時間フル稼働が可能です。ロジックの複製も容易ですから、医師不足も解消できます。

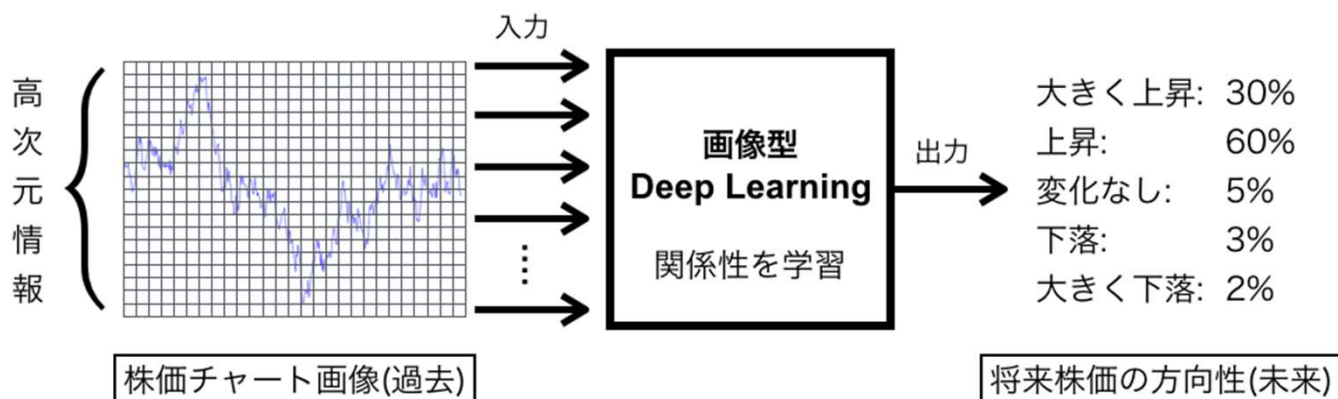
このようにAI技術の本質は、道具として人間を支援することにあります。速く安く安定的に大量処理を自動化できるツールとして、すでにAIは人間を超えています。シンギュラリティー※のようなSF話を持ち出すまでもありません。

※ 2045年頃にAIが人間以上の知性を獲得するという仮説 (レイ・カーツワイル)

<AIによる株価チャート分析 (画像認識)>

すると「株価予測もAIにやらせよう」という発想は極めて自然です。株価チャートを画像として扱えば、その後の値動きとの関係性を深層学習できるかもしれません。

しかしたとえ株価変動に法則があるとしても、悪性腫瘍のように、人間の目視の延長線上にある技術とは考え方が異なるはずで、つまり速さや安定性のみならず、人間以上の認識能力までもAIに要求することになります。これは可能なのでしょうか？



出所：大和証券投資信託委託

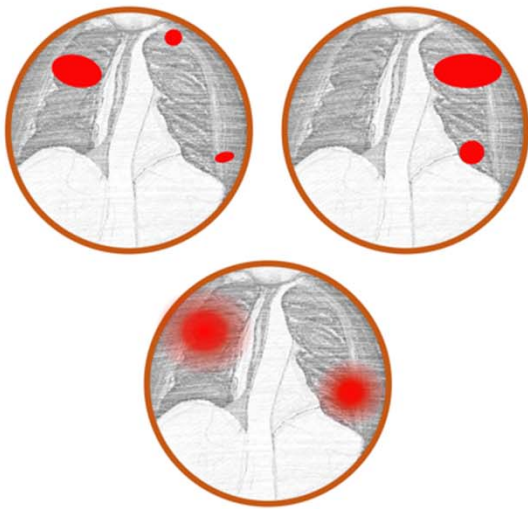
※1ページ目の「当資料のお取り扱いにおけるご注意」をよくお読みください。

結論を先に述べると「かなり難しい」でしょう。その根拠は画像だからです。非常に有望視されている医療画像診断においても、人間が教師画像を作成してAIに与えます。どこに腫瘍があるか人間が印を付けてあげるのです。人間が答えを教えている時点で、AIは認識能力において人間を超えることはできません。あくまで目的は自動化です※。株価チャート予測の場合、教師信号は実際の株価変動であるため人間が与えるものではありませんが、なおさら問題として難易度が高いと言えます。

※ 一部報道では、人間を上回る認識率を得た等の表現がなされますが、その理由は主に

- ・認識テストに時間制約があり、人間には見落としが起こる (高速で安定なAIが有利)
- ・認識対象の種類が膨大であり、人間には覚えきれない (大量処理できるAIが有利)

などであり、本来の認識能力 (形状の複雑性) とは趣旨が異なります。



【学習】

画像データを深層学習に入力し
どこに腫瘍があるか
答え (教師信号) を人間が教える。

【診断】

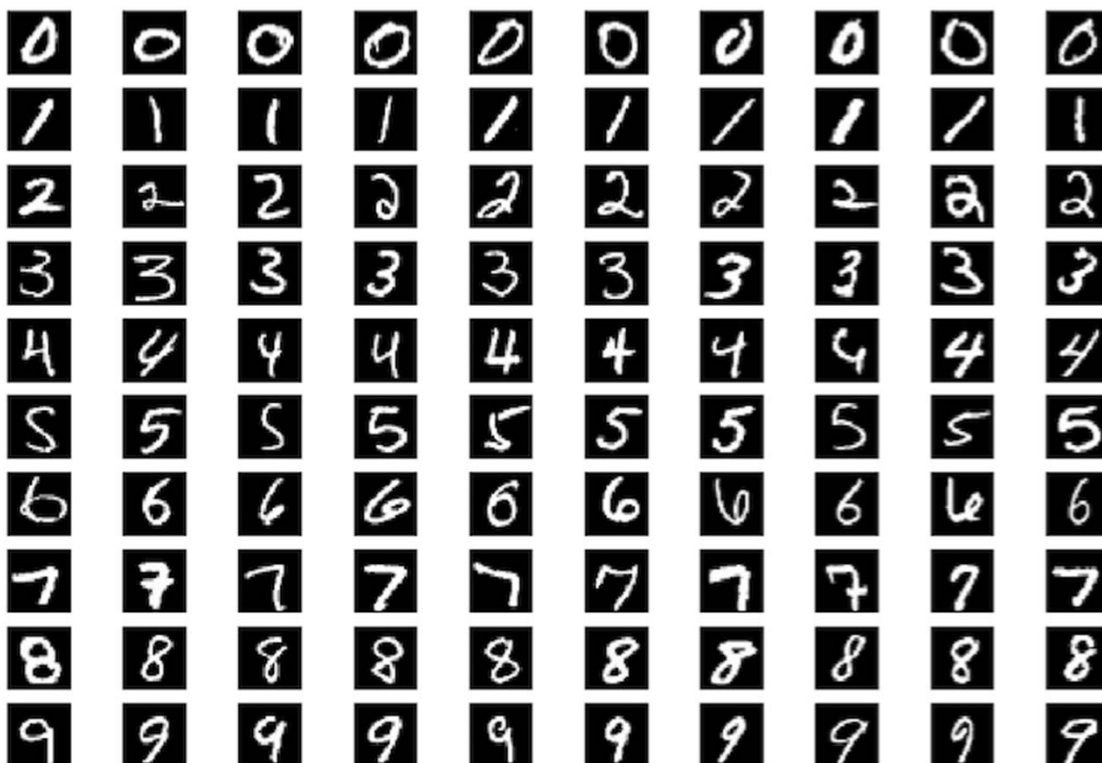
新規画像を学習済み深層学習に入力し
各ピクセルに対して
腫瘍がある確率を出力する。

出所: 茨城大学

※1ページ目の「当資料のお取り扱いにおけるご注意」をよくお読みください。

<いじわる手書き文字認識実験>

これを確かめるべく、簡単な実験をしてみます。画像認識の例題として最も有名なMNIST（手書き文字認識）を用います。ご覧の通り、それぞれの数字を認識することは難しくありません。画像型深層学習においても99%以上の精度で識別できます。

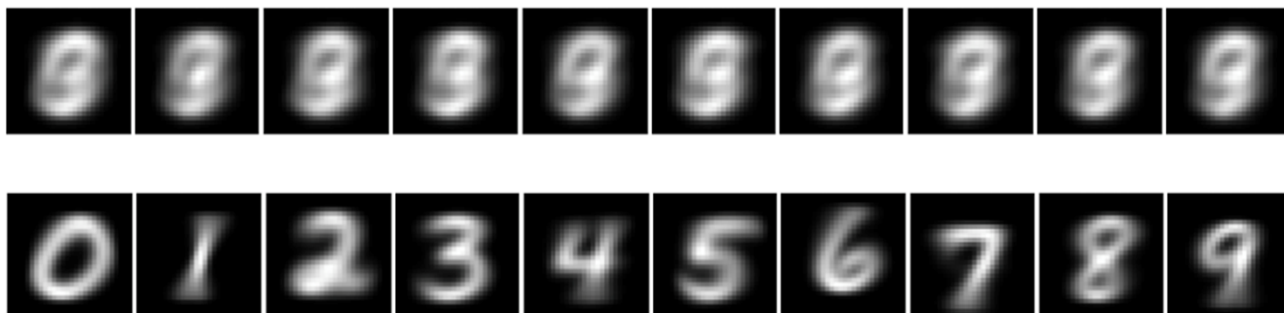


出所：茨城大学

当然ながら、株価チャートの予測はもっと難易度が高いはずですが。これを模擬すべく、手書き文字認識の正解ラベルの90%をランダムにシャッフルします。例えば8に見えても、別の数字が正解になるように意地悪をします。これにより金融市場のランダム性を表現します。しかし10%の法則性は維持されるため※、画像型深層学習によってこの法則性を検出できるかテストします。

※ 実際の金融市場に10%もの法則性があると主張するものではありません。金融市場の効率性の観点より、実際の法則性はまだ弱いと思われます。つまりこの実験は、実際の金融市場より易しい問題と言えます。

結果は全く識別できませんでした。画像型深層学習では同じ正解ラベルが付いた画像を集めて、共通の特徴を学ぼうとします。しかし90%の正解ラベルはでたらめですから、AIには次のように見えるのです。



同じ正解ラベルを持つ画像を平均化した様子: (上図)ランダム化あり、(下図)ランダム化なし。なお、上図と下図の正解ラベルは同一です。

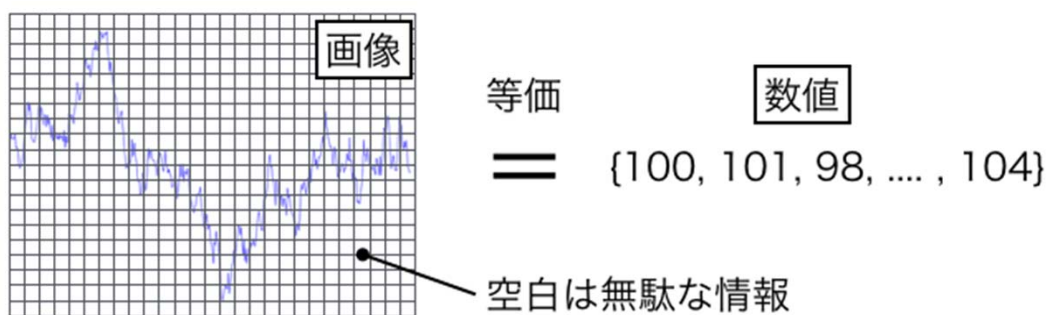
出所：茨城大学

10%の法則性は残存しているとは言え、上図は人間にも識別できません※。ゆえに深層学習でも手上げです。株価チャートを画像として識別するのは非常に難しい問題なのです。チャートの形状も数字に比べてはるかに複雑です。

※ 法則性を20%残存させると、人間にも識別できるようになります。同時に、画像型深層学習においても同等に識別できるようになります。

<非画像型の深層学習 (回帰分析)>

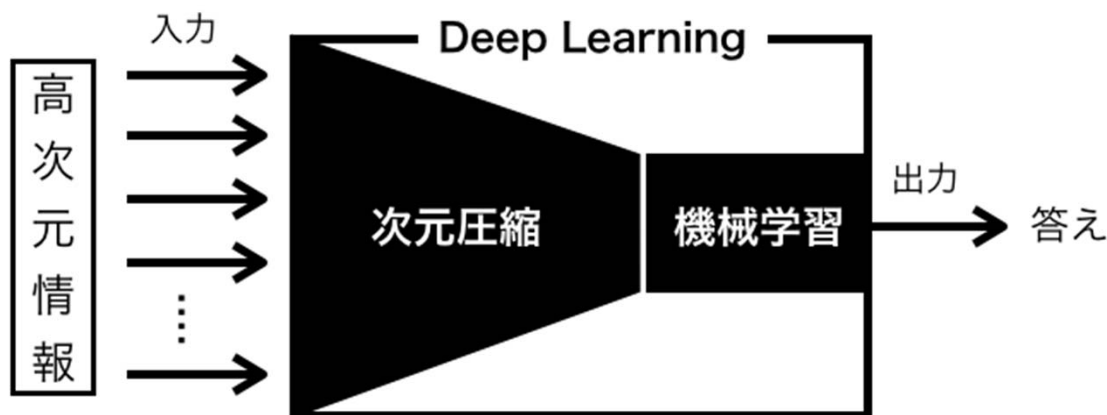
株価変動はそもそも数字の羅列 (時系列データ) なので、人間が見やすいように可視化 (画像化) するのは理にかなっていません。しかしAIにとっては情報として等価であれば、むしろ数値情報の方が演算処理しやすく無駄もありません。そこで画像以外の深層学習を考えてみましょう。



出所：大和証券投資信託委託

そもそも深層にする理由は、画像のような高次元情報※を各層において段階的に次元圧縮することで学習しやすくするためです。画像型深層学習ではプーリング層がその機能を担い、非画像型深層学習では主にオートエンコーダが用いられます。次元圧縮によりその後の機械学習を単純化できるため、過学習やデータ不足の問題を緩和できます。逆に高次元情報を用いないならば、わざわざ深層学習のような大鉈 (なた) を振る必要はありません。無用に多層化するほど、学習の失敗やブラックボックス化、さらには計算コストの増大につながります。

※ 画像の場合、各画素の輝度値が入力情報になるため、情報の種類は大規模になります。



出所：大和証券投資信託委託

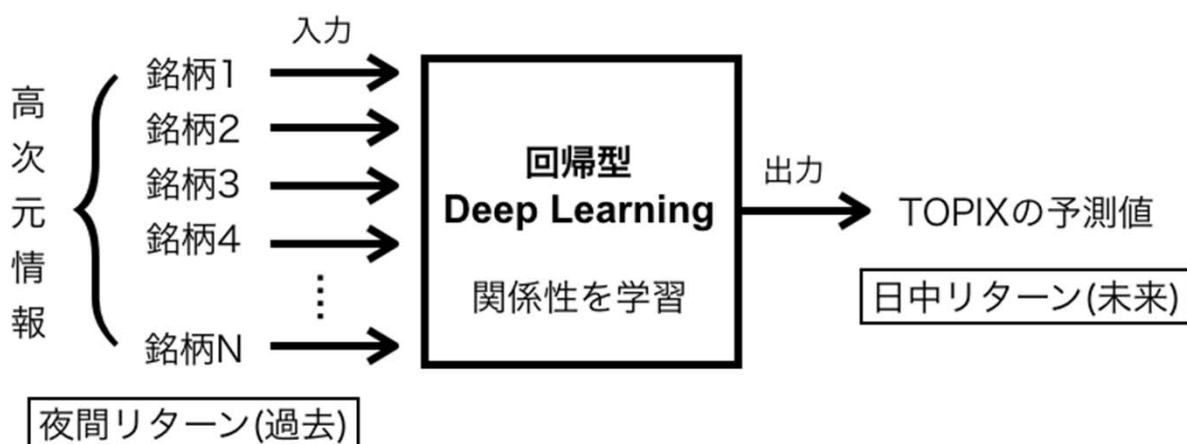
※1ページ目の「当資料のお取り扱いにおけるご注意」をよくお読みください。

このような高次元情報の利用が有効だと考えられる例題として「TOPIXの予測」が挙げられます※。

TOPIXは東証一部上場銘柄の全てから構成されるため、それぞれの個別銘柄の動きは次のTOPIXの動きに影響を与える可能性があります。そこで全銘柄の夜間リターンを説明変数に、TOPIXの日中リターンを目的変数とした回帰型の予測モデルを構築します。

リターンは変化率（数値）ですが、全銘柄を同時に入力するため画像のように高次元情報になります。そこでオートエンコーダを用いて次元圧縮を行います。その詳細については次回のMarket Letter（後編）にてご紹介いたします。

※ Tomoya Onizawa, Takehiro Suzuki, Tomoya Suzuki: "Predictability of Financial Market Indexes by Deep Neural Network," Proc. of Nonlinear Theory and its Applications (NOLTA), 4 pages, 2017.



出所：大和証券投資信託委託

【鈴木教授 プロフィール】

鈴木智也(すずきともや)

新潟県新潟市生まれ。IFTA国際検定テクニカルアナリスト(MFTA)。平成17年東京理科大学大学院理学研究科物理学専攻博士課程修了。理学博士。同年東京電機大学工学部電子工学科助手、平成18年より同志社大学工学部情報システムデザイン学科専任講師、平成21年より茨城大学工学部知能システム工学科准教授を経て、平成28年より同大学教授、さらに平成29年より大和証券投資信託委託(株)クウォンツ運用部特任主席研究員を兼務。平成30年より茨城大学大学院理工学研究科機械システム工学専攻長および領域長。研究分野は、時系列解析、テキスト解析、機械学習、人工知能、金融工学など実践的なデータサイエンスに従事。電子情報通信学会、情報処理学会、人工知能学会、日本テクニカルアナリスト協会、日本証券アナリスト協会各会員。

【Market Letter 鈴木教授による解説シリーズ バックナンバー】

第1回 AI運用に挑む	http://www.daiwa-am.co.jp/market/html_ml/ML20171207_1.html
第2回 集団化する人工知能	http://www.daiwa-am.co.jp/market/html_ml/ML20180125_1.html
第3回「2年目のジंकス」を集合知AIで緩和	http://www.daiwa-am.co.jp/market/html_ml/ML20180301_1.html
第4回 時系列データの見えない法則をつかむ	http://www.daiwa-am.co.jp/market/html_ml/ML20180409_1.html
第5回 愚かな人間心理・カモにするAI	http://www.daiwa-am.co.jp/market/html_ml/ML20180501_2.html
第6回 ナイトメア★アノマリーを狙え	http://www.daiwa-am.co.jp/market/html_ml/ML20180601_1.html
第7回 ブルーオーシャンAI戦略	http://www.daiwa-am.co.jp/market/html_ml/ML20180703_1.html

※1ページ目の「当資料のお取り扱いにおけるご注意」をよくお読みください。