

一目均衡表を用いた機械学習による日経平均 株価予測

畑仲, 優輝 / HATANAKA, YUKI

(出版者 / Publisher)

法政大学大学院理工学研究科

(雑誌名 / Journal or Publication Title)

法政大学大学院紀要. 理工学研究科編

(巻 / Volume)

65

(開始ページ / Start Page)

1

(終了ページ / End Page)

6

(発行年 / Year)

2024-03-24

(URL)

<https://doi.org/10.15002/00030755>

一目均衡表を用いた機械学習による 日経平均株価予測

NIKKI STOCK AVERAGE PREDICTION USING ICHIMOKU KINKO HYO

畑仲優輝

YUKI HATANAKA

指導教員 平原誠

法政大学大学院理工学研究科応用情報工学専攻修士課程

In this investigation, our objective is to predict the Nikkei Stock Average (Nikkei 225), which represents the trend of the Japanese stock market, with a high degree of accuracy. Technical analysis, which predicts future price movements based on past trends and patterns, is widely used to forecast stock prices. Ichimoku Kinko Hyo is one of the well-known methods in technical analysis. We predict the Nikkei Stock Average using Ichimoku Kinko Hyo as an input to a neural network. We demonstrate the effectiveness of the proposed method by observing the forecasting accuracy and operational performance of the Nikkei Stock Average.

Key Words : Machin learning, ,Neural network, Stock, Ichimoku

1. はじめに

近年, 株式市場では株価の値動きなどの時系列データを用いたテクニカル分析や, ニュースや決算説明書類などのテキストデータを用いたファンダメンタル分析により株価の予測が行われている。また機械学習を活用した株価予測の研究が活発に行われており, 様々なモデルが存在する[1-2]。精度の高い株価予測を行うことにより利益が期待できる。しかし, 景気や為替, 企業業績などの数多くの株価変動要因が株の需給に影響を及ぼし, 株価が複雑に変動するため精度の高い株価予測のモデルの開発は難しいとされている。

テクニカル分析手法に一目均衡表という手法が存在しており, 本研究では, 過去の日経平均株価指数の始値, 終値, 高値, 安値からなる四本値データと, 一目均衡表の情報から得たデータの入力から, NN(Neural Network)により翌営業日の日経平均株価指数を予測する。

2. 日経平均株価

本研究での予測対象となる日経平均株価は, 東京証券取引所に上場する銘柄のうち, 日本経済新聞社が選定した 255 銘柄の株価をもとに, 株式が取引されている時間帯に算出され配信されている。

3. 一目均衡表

一目均衡表とは, 細田悟一氏により考案されたテクニカル分析手法であり, 今日でも, 金融商品の取引の判断に国内外で活用されている[3]。

一目均衡表には時間によって相場の転換点を分析する時間論, 波動の形によって相場の方向性を分析する波動論, 及び相場の上昇幅や下落幅を計算し分析する水準論の 3 つの理論があり, これらを取り入れた三役好転(取引価格上昇の目安)と三役逆転(取引価格下落の目安)の条件が存在する。三役好転と三役下落は, 現在の金融商品の取引価格と時間ごとの金融商品の取引価格からなるチャートから算出した, 基準線, 転換線, 先行スパン 1, 先行スパン 2, 遅行線の比較により条件を全て満たすか否かが判断される。

三役好転の条件は, ①転換線が基準線より上に位置する状態, ②遅行線が取引価格より上に位置する状態, ③取引価格が先行スパン 1 及び先行スパン 2 より上に位置する状態である。

三役逆転の条件は, ①転換線が基準線より下に位置する状態, ②遅行線が取引価格より下に位置する状態, ③取引価格が先行スパン 1 及び先行スパン 2 より下に位置する状態である。

本研究ではこの三役好転と三役逆転の条件を NN への入力として取り入れる。

4. 実験

(1) 入力

日経平均株価の過去データから四本値と一目均衡表を編集し加工する。加工したデータを NN への入力として, 翌日の日経平均株価の終値の前日比を予測する。

NN への四本値の入力は, 連続 $d(= 3, 5, 10)$ 営業日分

の始値, 高値, 安値, 終値それぞれの前日比とする. 以降この入力を四本値入力と呼ぶ.

一目均衡表を加工した入力として, t 日の日経平均株価の終値 E_t を用い, 基準線 S_t , 転換線 C_t , 先行スパン 1 O_t , 先行スパン 2 T_t , 遅行線 L_t を求める.

$$S_t = \frac{\max(E_{t-25}, E_{t-24}, \dots, E_t) + \min(E_{t-25}, E_{t-24}, \dots, E_t)}{2} \quad (1)$$

$$C_t = \frac{\max(E_{t-8}, E_{t-7}, \dots, E_t) + \min(E_{t-8}, E_{t-7}, \dots, E_t)}{2} \quad (2)$$

$$O_t = \frac{C_{t-26} + S_{t-26}}{2} \quad (3)$$

$$T_t = \frac{\max(E_{t-77}, E_{t-76}, \dots, E_{t-26}) + \min(E_{t-77}, E_{t-76}, \dots, E_{t-26})}{2} \quad (4)$$

$$L_t = E_{t-26} \quad (5)$$

これら 5 つの指標から三役好転と三役逆転の条件を考慮した, 以下の 4 つの入力(一目入力)を提案する.

$$\frac{C_t - S_t}{E_t} \quad (6)$$

$$\frac{E_t - \max(O_t, T_t)}{E_t} \quad (7)$$

$$\frac{E_t - \min(O_t, T_t)}{E_t} \quad (8)$$

$$\frac{L_t - E_t}{E_t} \quad (9)$$

本研究では四本値入力の NN モデル(四本値モデル)と四本値入力に一目入力を合わせた NN モデル(提案モデル)の予測精度, 及び売買運用性能を比較する.

(2) データセット

日経平均株価データを「Fusion Media Ltd.」が運営するポータルサイト「Investing.com」から取得する[4]. モデルの学習データを 950 件, 学習を評価するための検証データを 50 件, テストデータを 1 件としてデータセットを作成する. データセットは連続した営業日のデータを作成しており, 古い営業日のデータから学習データ, 検証データ, テストデータの順序で構成し実験を行う. データセットの日付を 1 営業日分ずつずらし, 2015 年から 2019 年までの予測を行う.

(3) モデルの学習

連続 $d(=3, 5, 10)$ 営業日分の, 四本値入力をを用いる計 12, 20, 40 ($=3 \times 4, 5 \times 4, 10 \times 4$) 入力の四本値モデル 3 種類と, 四本値入力に一目入力を加えた計 24, 40, 80 ($=3 \times 8, 5 \times 8, 10 \times 8$) 入力の提案モデル 3 種類を構築する. 損失関数を平均二乗誤差とし, 中間素子数を 16, 32, 64, 128, 256 の 5 段階で最適化をする. バッチサイズは 128

で統一し, 学習回数は 5000 回とするが, 検証データの損失関数が 100 回連続で改善されない時点で学習を打ち切り, テストデータでの予測を行う. 平均的な予測精度を観測するため, 各モデルの学習, 及び予測について, NN の初期の重みをランダムに変更し 5 回繰り返す.

(4) 予測精度評価

四本値モデルと提案モデルのそれぞれ 3 種類について 2015 年から 2020 年の 1 年おき, 5 試行分の予測精度を評価する. なお, 評価指標として実際の終値前日比と予測値との平均的な乖離を示す平均絶対誤差を用いる. 5 試行の平均絶対誤差 ($\pm 1SD$) を図 1~6 に示す. なお比較のため, 翌営業日の終値が前の終値と同じとする「予測なしモデル」を導入し, 予測なしモデルの平均絶対誤差を点線で表示している. 図 1~6 の各々について, 予測なしモデルと各モデル(四本値モデル 3 種類, 提案モデル 3 種類)との間で有意水準 0.05 の t 検定を実施し, 予測なしモデルに対して予測精度が有意に上回ったモデルに青色の*を, 有意に下回ったモデルに赤色の*を付記する. 加えてモデル(四本値モデル 3 種類, 提案モデル 3 種類)間で Bonferroni 法を用いた多重比較の結果を表 1~6 に示す. 有意水準を 0.05, 実施する検定数 $N = 15$ として補正を行い, 有意差のある組に対して○を記す.

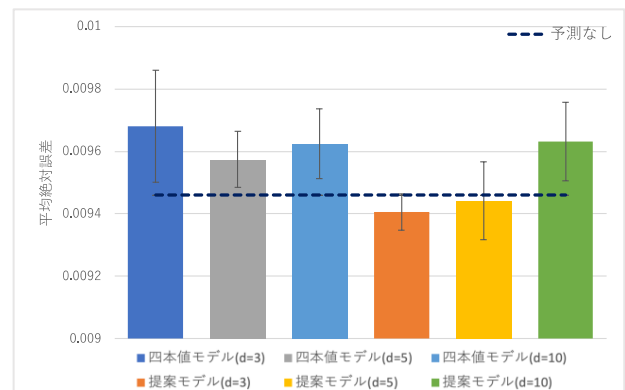


図 1 2015 年モデル別平均絶対誤差

表 1 2015 年モデル別平均絶対誤差の多重比較

2015	四本値モデル(d=3)	四本値モデル(d=5)	四本値モデル(d=10)	提案モデル(d=3)	提案モデル(d=5)	提案モデル(d=10)
四本値モデル(d=3)						
四本値モデル(d=5)						
四本値モデル(d=10)						
提案モデル(d=3)						
提案モデル(d=5)						
提案モデル(d=10)						

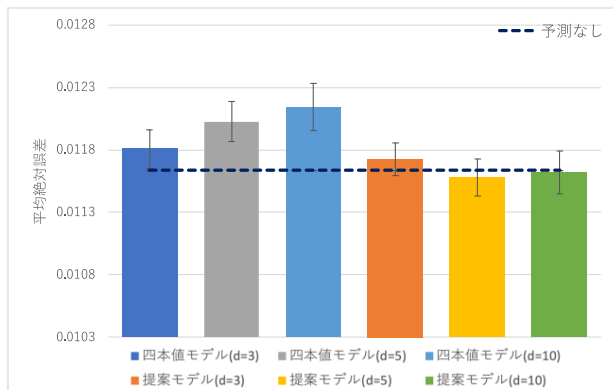


図 2 2016 年モデル別平均絶対誤差

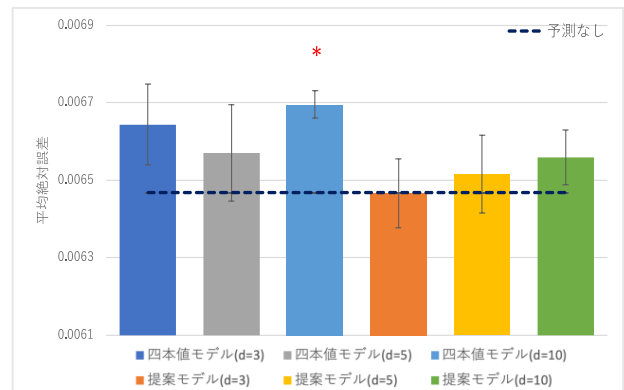


図 5 2019 年モデル別平均絶対誤差

表 2 2016 年モデル別平均絶対誤差の多重比較

2016	四本値モデル(d=3)	四本値モデル(d=5)	四本値モデル(d=10)	提案モデル(d=3)	提案モデル(d=5)	提案モデル(d=10)
四本値モデル(d=3)						
四本値モデル(d=5)						
四本値モデル(d=10)						
提案モデル(d=3)						
提案モデル(d=5)						
提案モデル(d=10)						

表 5 2019 年モデル別平均絶対誤差の多重比較

2019	四本値モデル(d=3)	四本値モデル(d=5)	四本値モデル(d=10)	提案モデル(d=3)	提案モデル(d=5)	提案モデル(d=10)
四本値モデル(d=3)						
四本値モデル(d=5)						
四本値モデル(d=10)						
提案モデル(d=3)						
提案モデル(d=5)						
提案モデル(d=10)						

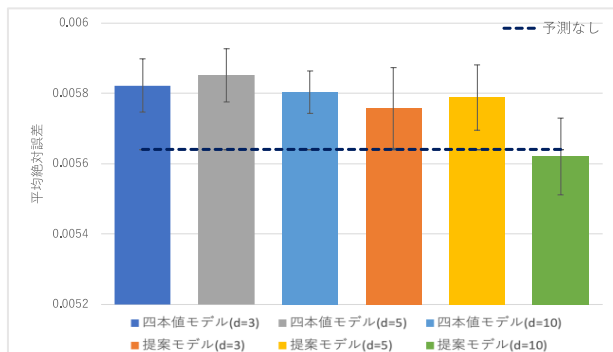


図 3 2017 年モデル別平均絶対誤差

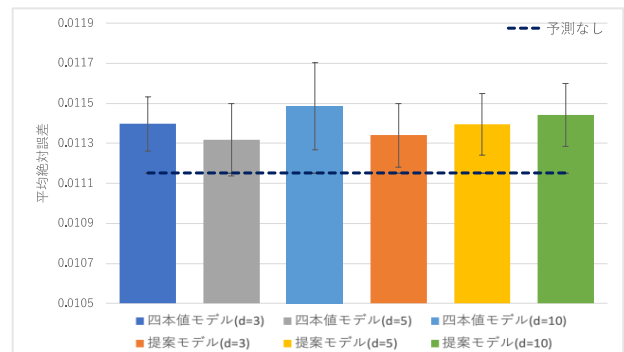


図 6 2020 年モデル別平均絶対誤差

表 3 2017 年モデル別平均絶対誤差の多重比較

2017	四本値モデル(d=3)	四本値モデル(d=5)	四本値モデル(d=10)	提案モデル(d=3)	提案モデル(d=5)	提案モデル(d=10)
四本値モデル(d=3)						
四本値モデル(d=5)						
四本値モデル(d=10)						
提案モデル(d=3)						
提案モデル(d=5)						
提案モデル(d=10)						

表 6 2020 年モデル別平均絶対誤差の多重比較

2020	四本値モデル(d=3)	四本値モデル(d=5)	四本値モデル(d=10)	提案モデル(d=3)	提案モデル(d=5)	提案モデル(d=10)
四本値モデル(d=3)						
四本値モデル(d=5)						
四本値モデル(d=10)						
提案モデル(d=3)						
提案モデル(d=5)						
提案モデル(d=10)						

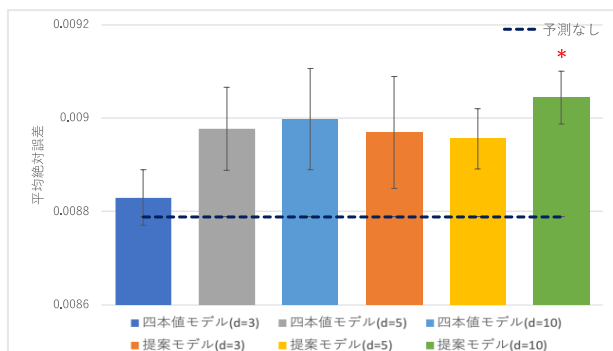


図 4 2018 年モデル別平均絶対誤差

表 4 2018 年モデル別平均絶対誤差の多重比較

2018	四本値モデル(d=3)	四本値モデル(d=5)	四本値モデル(d=10)	提案モデル(d=3)	提案モデル(d=5)	提案モデル(d=10)
四本値モデル(d=3)						
四本値モデル(d=5)						
四本値モデル(d=10)						
提案モデル(d=3)						
提案モデル(d=5)						
提案モデル(d=10)						

(5) 運用性能評価

四本値モデルと提案モデルのそれぞれ 3 種類, 5 試行分の予測結果から簡易的な運用シミュレーションを行う。

運用シミュレーションでは日経平均株価自体を仮想的な金融商品(日経株)と仮定する。2015 年から 2020 年の各年の年初に 100 ポイント保有し, 年初に保有している日経株はないものとする。各モデルの予測値により行動を決定し, 予測値が 1.0 を上回った場合(日経株の上昇が予測される場合)には「購入」の行動をとり, 予測値が 1.0 を下回った場合(日経株の下落が予測される場合)には「売却」の行動をとる。

例として予測値が 1.01 の場合, 「購入」の行動をとり, 実際の終値前日比が 1.02 ならばその後に日経株の株価を 1.02 倍にする。このように日経株の株価を日経平均株価に連動させる。また予測値が 0.99 の場合, 「売却」の行

動を取り、その後に日経株の株価を日経平均株価に連動させる。

「購入」の行動を選択した際、日経株を保有していない場合には、現在保有しているポイントを全て使用し日経株を購入する。事前に日経株を保有している場合には、日経株の保有をそのまま継続する。「売却」の行動を選択した際、事前に日経株を保有している場合には、現在の株価で日経株を全て売却しポイントを得る。日経株を保有していない場合には、何もせず日経株を保有しない。なお本シミュレーションでは売買による費用や管理費用などの手数料は考慮しないものとする。

2015年から2020年の四本値モデルと提案モデルそれぞれ3種類、5試行分の予測結果から各年の年初の日経平均株価を100ポイントとして、運用シミュレーションを行い、各年の年末に「売却」を行なった結果のポイントを図7~12に示す。なお、運用性能の比較のため、年初に日経株を購入し保持し続ける「予測なしモデル」を導入し、運用結果を点線で表示している。図7~12の各々について予測なしモデルと各モデル(四本値モデル3種類、提案モデル3種類)との間で有意水準0.05のt検定を実施し、予測なしモデルに対して予測精度が有意に上回ったモデルに青色の*を、有意に下回ったモデルに赤色の*を付記する。加えてモデル(四本値モデル3種類、提案モデル3種類)間でBonferroni法を用いた多重比較の結果を表7~12に示す。有意水準を0.05、実施する検定数 $N=15$ として補正を行い、有意差のある組に対して○を記す。

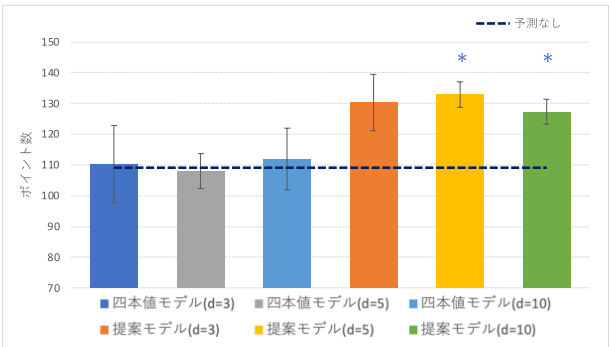


図7 2015年モデル別運用性能評価

表7 2015年モデル別運用性能の多重比較

2015年	四本値モデル(d=3)	四本値モデル(d=5)	四本値モデル(d=10)	提案モデル(d=3)	提案モデル(d=5)	提案モデル(d=10)
四本値モデル(d=3)						
四本値モデル(d=5)						
四本値モデル(d=10)						
提案モデル(d=3)						
提案モデル(d=5)						
提案モデル(d=10)						

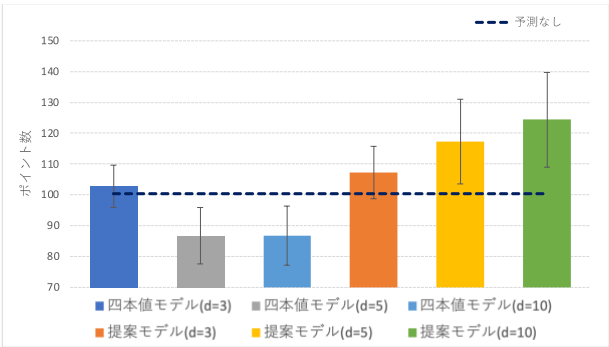


図8 2016年モデル別運用性能の多重比較

表8 2016年モデル別運用性能評価有意差検定

2016年	四本値モデル(d=3)	四本値モデル(d=5)	四本値モデル(d=10)	提案モデル(d=3)	提案モデル(d=5)	提案モデル(d=10)
四本値モデル(d=3)						
四本値モデル(d=5)						
四本値モデル(d=10)						
提案モデル(d=3)						
提案モデル(d=5)						
提案モデル(d=10)						

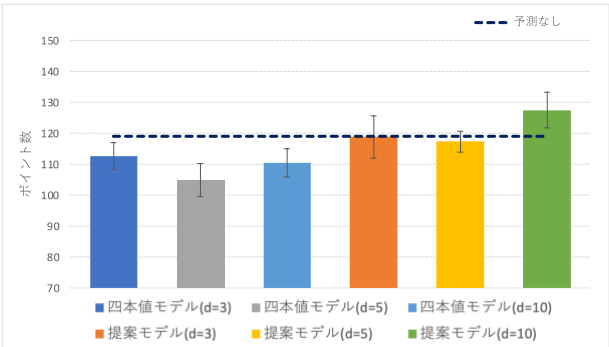


図9 2017年モデル別運用性能評価

表9 2017年モデル別運用性能の多重比較

2017年	四本値モデル(d=3)	四本値モデル(d=5)	四本値モデル(d=10)	提案モデル(d=3)	提案モデル(d=5)	提案モデル(d=10)
四本値モデル(d=3)						
四本値モデル(d=5)						
四本値モデル(d=10)						
提案モデル(d=3)						
提案モデル(d=5)						
提案モデル(d=10)						

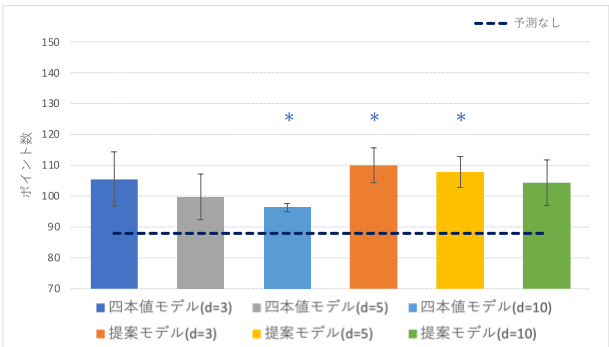


図10 2018年モデル別運用性能評価

表10 2018年モデル別運用性能の多重比較

2018年	四本値モデル(d=3)	四本値モデル(d=5)	四本値モデル(d=10)	提案モデル(d=3)	提案モデル(d=5)	提案モデル(d=10)
四本値モデル(d=3)						
四本値モデル(d=5)						
四本値モデル(d=10)						
提案モデル(d=3)						
提案モデル(d=5)						
提案モデル(d=10)						

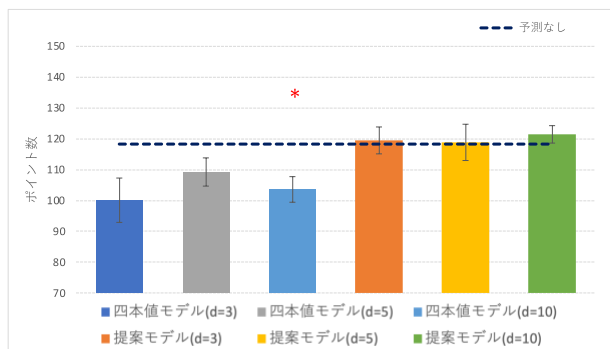


図 11 2019 年モデル別運用性能評価

表 11 2019 年モデル別運用性能の多重比較

2019年	四本値モデル(d=3)	四本値モデル(d=5)	四本値モデル(d=10)	提案モデル(d=3)	提案モデル(d=5)	提案モデル(d=10)
四本値モデル(d=3)				○	○	○
四本値モデル(d=5)						○
四本値モデル(d=10)				○	○	○
提案モデル(d=3)	○		○			
提案モデル(d=5)	○		○			
提案モデル(d=10)	○	○	○			

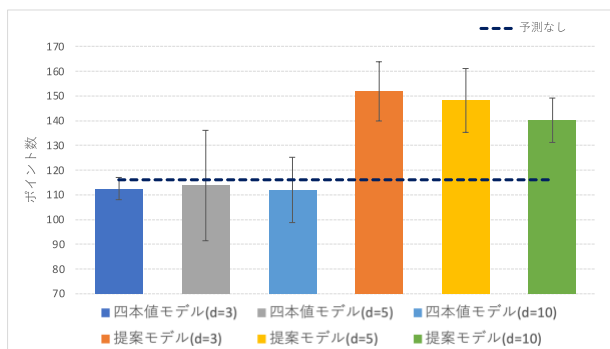


図 12 2020 年モデル別運用性能評価

表 12 2020 年モデル別運用性能の多重比較

2020年	四本値モデル(d=3)	四本値モデル(d=5)	四本値モデル(d=10)	提案モデル(d=3)	提案モデル(d=5)	提案モデル(d=10)
四本値モデル(d=3)				○	○	○
四本値モデル(d=5)						○
四本値モデル(d=10)				○	○	○
提案モデル(d=3)	○		○			
提案モデル(d=5)	○		○			
提案モデル(d=10)	○		○			

5. 考察

2015 年から 2020 年の四本値モデルと提案モデルそれぞれ 3 種類、5 試行の平均絶対誤差を示した図 1~6 に着目する。2018 年で四本値モデルが提案モデルより予測精度が有意に高い組み合わせが存在した。一方 2016 年と 2019 年においては提案モデルが四本値モデルより有意に高い組み合わせが存在した。2018 年は他の 5 年と異なり年初から年末にかけて日経平均株価が下落する相場であったことから、下落相場での前日比の数値予測には四本値モデルが優れており、四本値モデルは下落方向へのネガティブな予測をする傾向があると考えられる。2015 年、2016 年、2017 年、2019 年の提案モデルは、前日比を 1.0 とする予測なしモデルと比較して、有意差はなかったものの精度の平均が勝っているものが見られた。そのため試行回数を増やすことで上昇相場での提案モデルの有効性を示せると考える。

図 7~12 で示した売買シミュレーションにおける運用性能評価では、2015 年から 2020 年の全ての年に関して、提案モデルが四本値モデルよりも有意に勝る組み合わせが存在した。また提案モデルは運用性能で 2015 年と 2018 年に有意差をもって予測なしモデルに勝り、他の年においても平均が予測なしモデルより高い傾向が見られた。唯一予測なしモデルが 100 ポイントを切り、年間を通して下落局面を示した 2018 年において、四本値モデル(d=10)以外で平均が 100 ポイントを切ることはなかった。また四本値モデル(d=10)は予測なしモデルと比較して有意に運用性能が勝っており、四本値モデルと提案モデル共に下落を予測し売却行動が取れたと考える。

ここで予測精度評価では予測なしモデルと比較して四本値モデルと提案モデルそれぞれ 3 種類とも有意に勝った結果が得られなかった一方、運用性能評価では有意に勝った結果が得られた理由を考える。四本値モデルと提案モデルそれぞれ 3 種類、5 試行分の予測結果における株価上昇(P)と下落(N)の 2 値分類評価(正解率、適合率、陰性的中率)を図 13、14 に示す。図 13 では年末にかけて上昇相場であった 2015 年の評価を示し、図 14 では年末にかけて下落相場であった 2018 年の評価を示す。また 2015 年と 2018 年の年間における、上昇した営業日数、下落した営業日数及び全営業日数を表 13 に示す。

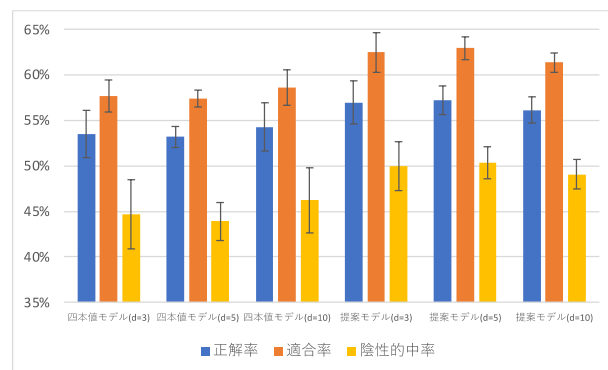


図 13 2015 年モデル別 2 値分類評価

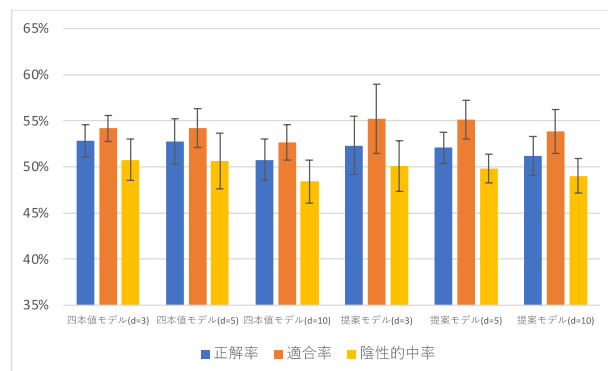


図 14 2018 年モデル別 2 値分類評価

表 13 年間営業日騰落数

	上昇営業日数(割合)	下落営業日数(割合)	全営業日数
2015年	139(57.0%)	105(43.0%)	244
2018年	128(52.2%)	117(47.8%)	245

図 13 の 2015 年では四本値モデルと比較して提案モデルが正解率、適合率、陰性的中率ともに平均的に優れていた。運用性能評価での予測なしモデルは、年初から年末にかけて「日経株」を保有し続けるため、常に株価が上がると予測するモデルであるとも考えられる。従って上昇営業日数の割合が予測なしモデルの正解率と適合率になり、2015 年においては 57.0%となる。予測なしモデルの正解率は 3 種類の提案モデルの正解率と同等の数値であった。一方提案モデルの適合率は 60%以上の結果を示していた。図 14 の 2018 年でも同様に適合率の観点で提案モデルは優れており、モデルが上昇すると予測した日に関しての的中率が高い結果となった。3 種類の提案モデルの陰性的中率においては 2015 年と 2018 年ともに 50%前後であった。下落営業日数の割合をみると、2015 年は 43.0%、2018 年は 47.8%であり、陰性的中率の方が平均的に高い。従って提案モデルが下落すると予測した日に関しての的中率も高い結果であったと考える。

以上のことから、四本値モデルよりも提案モデルの方が的確に上昇を予測し、「日経株」を保有する行動を取っていたため、年間を通して損失を回避しながら運用が行えていたと評価できる。

6. 今後の展望

運用性能評価で予測なしモデルを有意に上回った 2015 年提案モデル($d=10$)における、5 試行中の 1 試行分の出力(estimated y)と実際の終値前日比(actual y)の散布図を図 15 に示す。

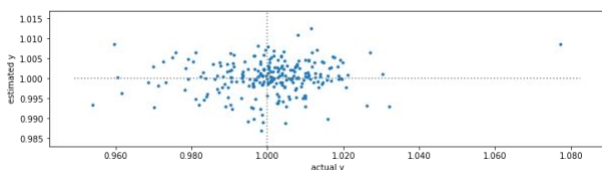


図 15 2015 年提案モデル($d=10$)の出力散布図

図 15 では、実際の終値前日比と比較してモデルの出力が幅狭い範囲に限られていた。本研究で作成した、他の四本値モデルや提案モデルの出力にも同じ傾向が観察された。これは NN の損失関数に平均二乗誤差を採用したことで、出力が大きく乖離しないように前日比 1.0 付近をとる消極的な学習が行われたと考えられる。従って損失関数を変更することにより、更なる予測精度の向上が期待できる。

提案モデルの入力に使用した基準線 S_t 、転換線 C_t 、先行スパン 1 O_t 、先行スパン 2 T_t 、遅行線 L_t は慣例的に用

いられる営業日数分過去のデータを参照した。この営業日数も一種のパラメータとして考えることができ、最適な営業日数を探索する余地がある。

謝辞：

本研究を進めるにあたり、多大なご指導を賜りました、修士論文指導教員の平原誠准教授に心からの感謝の意を表します。

参考文献

- 1) 左穀, 北栄輔: “ベイジアンネットワークを用いた株価予測について”, 情報処理学会論文誌トランザクション数理モデル化と応用, vol.3, no.3, pp.80-90, 2010.
- 2) Adebisi AA, Ayo KC, Adebisi MO, Otokiti SO “Stock price prediction using neural network with hybridized market indicators”, J Emerg Trends Comput Inf Sci, vol.3, no.1, pp.1-9, 2012.
- 3) 佐々木 英信, 一目均衡表の研究, (株)投資レーダー, 2003.
- 4) Fusion Media Ltd, 日経 225 過去のレート (N 225)-Investing.com, <https://www.investing.com/indices/japan-ni225-historical-data>, (参照 2023-9-1).