

# ニューラルネットを用いることによる 株価予測への有効性について

提出日： 平成 20 年 2 月 18 日  
所属研究室： 計算神経科学研究室  
指導教官： 北野 勝則 准教授  
学籍番号： 2640040047-3  
氏名： 神原 悠祐

## 概要

先行研究の中でニューラルネットワークを用いての株価予測は数多く行われている。しかし、株価予測の先行研究のほとんどは X 日後の株価の騰落を予測しているものがほとんどである。また X 日後の株価の変動を求めようとしている研究は個別銘柄の変動に比べると変動の緩やかな日経平均株価や TOPIX のようなものを予測対象に選んでいる。そこで本研究では株式の個別銘柄の変動をニューラルネットワークによって予測しようと試みた。ニューラルネットワークの学習において重要なことは入力値の選択・処理方法、適切な学習データの選択である。そこで本研究ではニューラルネットワークの入力値と学習データに着目して個別銘柄の株価を予測してみた。結果、過去のデータ(学習期間データ)についてはほぼ完璧な予測精度であり、未来のデータ(学習期間外データ)についてもある程度の精度の予測ができた。

# 目次

概要	2
1. はじめに	4
2. 方法	5
2.1 ニューラルネットワークについて	5
2.2 階層型ニューラルネットワークについて	5
2.3 誤差逆伝搬学習法(Back Propagation)について	6
2.4 最急降下法について	7
2.5 関数近似能力について	8
2.6 ニューラルネットワークの入力値について	9
2.6.1 株価の時系列データの前処理について	9
2.6.2 正規化について	9
2.6.3 入力方法	9
2.6.4 入力方法	10
2.6.5 入力方法	11
3. 結果	12
3.1 学習データでの2つの入力方法の比較	13
3.1.1 入力方法による結果	13
3.1.2 入力方法による結果	16
3.2 入力方法を用いての学習期間外での予測結果	19
3.2.1 野村ホールディングスの学習期間外での予測結果	20
3.2.2 松下電器産業の学習期間外での予測結果	22
4. 考察	25
謝辞	26
参考文献	26
付録	27

# 1. はじめに

株式など時系列の予測は昔から様々な手法が考案されてきているが、様々な手法は大きく分けると「確率論」と「決定論」に分けることが出来る。

まず、確率論の立場の考え方は時系列の複雑な動きは確率的(ランダム)なため予測することは出来ないというもので、時系列から得られる情報は平均や分散といった統計量のみと考える。確率論での時系列を予測する方法は実際の時系列から得られた平均、分散を持つような数理モデルを再現することによって予測する。株価時系列予測において確率論的な手法による予測をしているのは金融工学、金融経済学などのアカデミックな世界である。

次に、決定論の立場の考え方は現実の市場は必ずしも理論上の仮定と一致しないとし、時系列には予測可能性が残っているというもので、実務の世界ではファンダメンタル分析やテクニカル分析などの決定論的な手法がよく用いられ、また近年では、ニューラルネット、遺伝的アルゴリズム、強化学習、フラクタル次元、サポートベクタマシンなどの様々な学習理論を用いての株価予測の研究はされてきている。

ここまでは今まで研究されてきた株価予測の種類を述べてきたが、これらの研究の結果により株価予測に対する一般的な認識は「金融市場はランダムな動きをする世界ではなく、マクロな経済動向を反映する市場として長期的には予測可能である」となっている。またもうひとつ一般的な認識として「将来の株価はピンポイントに確定できない」というものもある。この理由は現時点では予測できない要因(突発的な不祥事、政府の金融政策など)が存在し、その影響を受けるためである。しかしこれは株価予測のすべてを否定するものではない。実際先ほどの認識が生まれた理由から逆を言えば、現時点で予測できない要因が存在しない間は将来の株価を予測できるとも言える。また確かに株価に影響を与える要因がいつ起こるかは予測できないが、その要因が起こったことによって株価に大きな影響を与えた場合にもその後の影響については過去の類似性の高い要因の影響からある程度の精度で予測できる可能性がある。そこで本論文では予測できない要因の影響を受けやすい個別銘柄の株価を予測する。予測器として階層型ニューラルネットワークを用い、入力値と学習データについて着目しての予測を試みる。

## 2. 方法

### 2.1 ニューラルネットワークについて

ニューラルネットワークとは、図1に示すように人間の脳内のニューロンの結合からなる神経回路による情報処理機構を模倣した数理モデルであり、与えられたデータに対して学習し必要とされる情報処理を実現するものである。ニューラルネットワークには教師信号(正解)の入力を与えることにより問題に対して最適化していく教師あり学習と教師信号を必要としない教師なし学習に分けられる。教師あり学習の代表的なものとして階層型ニューラルネットワークが挙げられる。

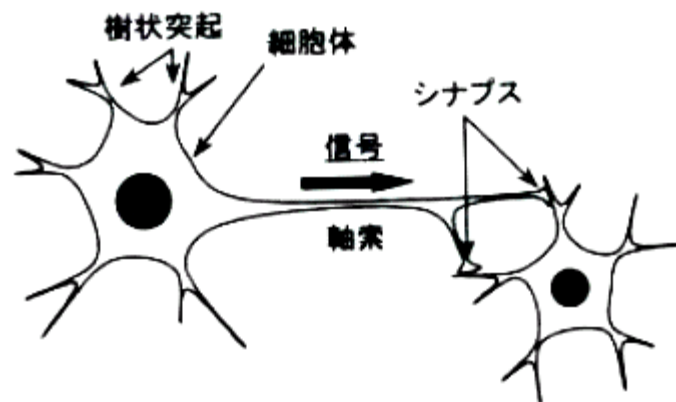


図1： ニューロンの結合(参考文献[3]より抜粋)

### 2.2 階層型ニューラルネットワークについて

ニューロン間での信号の流れは入力層へ信号が入れば、順次隣の層へ信号が伝搬していき最終的には出力層へ伝搬する。三層の階層型ネットワークの場合、入力層へ入力が入り、入力層の出力が中間層の入力となる。次に、中間層の出力が出力層への入力となる。そして、出力層の出力がネットワーク全体の出力となる。学習に使用されるアルゴリズムは最急降下法に基づく誤差逆伝搬学習法が有名である。

## 2.3 誤差逆伝搬学習法(Back Propagation)について

階層型ニューラルネットワークの学習技術であり、人間の脳内のニューロン結合からなる神経回路網を模擬した学習方法である。図2はネットワークの概略図である。

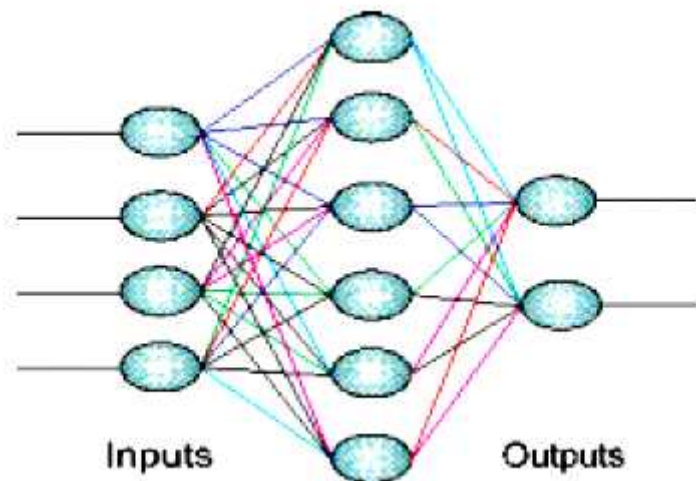


図2： 階層型ニューラルネットワーク(参考文献[3]より抜粋)

バックプロパゲーションとは入力層に入力されたデータから計算された出力層からの出力値を正しい値である教師信号と比較して、正しくない場合はそれぞれのユニット間の結合荷重を更新するといった処理を出力値と教師信号の誤差が一定以下になるまで繰り返し行う学習方法である。

今、 $m$ 層のネットワークと考えると、 $k$ 層の $i$ 番目のユニットの総入力を $x$ 、このユニットの出力を $y$ 、閾値を $\theta$ 、 $k-1$ 層の $i$ 番目のユニットから $k$ 層の $j$ 番目のユニットへの結合荷重を $\omega_{ij}$ と表記する。そのときの各ユニットの入力と出力は次のように定義されている。

$$\text{ユニットの出力： } y = f(x) = \frac{1}{1 + e^{-x}} \quad (1)$$

$$\text{ユニットへの総入力： } x = \sum_j \omega_{ij} y_j - \theta \quad (2)$$

上記の式(1)(2)によりネットワーク全体の出力を計算することができる。この出力された結果と教師信号を比較して学習終了を判断する。ニューラルネットワークの学習終了の判断は出力層からの出力値と教師信号  $t$  の二乗誤差を計算し-

定値以下であれば学習終了とする。二乗誤差の計算式は次の式で表される。

$$\text{二乗誤差: } E = \frac{1}{2} (y - t)^2 \quad (3)$$

二乗誤差が一定値以下でない場合は次式の最急降下法によって結合荷重を修正していく。

$$\Delta \omega_{ij} = -\varepsilon \frac{\partial E}{\partial \omega_{ij}} \quad (4)$$

$\varepsilon$ は学習率で学習の速度を決める定数である。ここでユニットの入出力関数がシグモイド関数であることから式(4)を変形すると

$$\Delta \omega_{ij} = -\varepsilon \delta^k y_j (1 - y_j) \quad (5)$$

となる。ここで、

$$\delta^k = \begin{cases} y - t, & \text{if } (k = m) \\ \sum \delta^k \omega_{ij}, & \text{else} \end{cases} \quad (6)$$

である。このように結合荷重を修正していく過程で必要な誤差の計算がネットワークの出力を求める際の計算とは逆の流れで出力層から入力層に向かって伝搬するようになっている。これが逆誤差伝搬法と呼ばれる所以である。

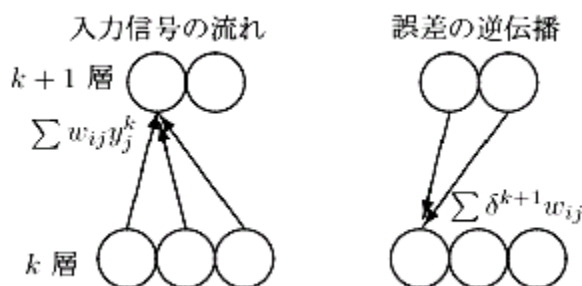


図3： 誤差逆伝播の模式図

## 2.4 最急降下法について

最急降下法は、ある適当な初期値からはじめて、その値を繰り返し修正することにより、最適なパラメータの値を求める方法(繰り返し最適化手法)の最も基本的で簡単な方法である。バックプロパゲーションで使用されている式(4)で簡単に説明するとネットワークの出力値と教師信号の誤差は結合荷重 $\omega$ によって定まる誤差関数とで表すことが出来る。誤差関数  $E$  は最小値と数多くの極小値が存在しているが、現在の結合荷重のときの誤差関数の一次微分(傾き)に学習率を乗じたものを減算することによって最小値にたどり着くかは初期値に依存するが、極小値(局所的最小解)にはたどり着くことは保障されている。

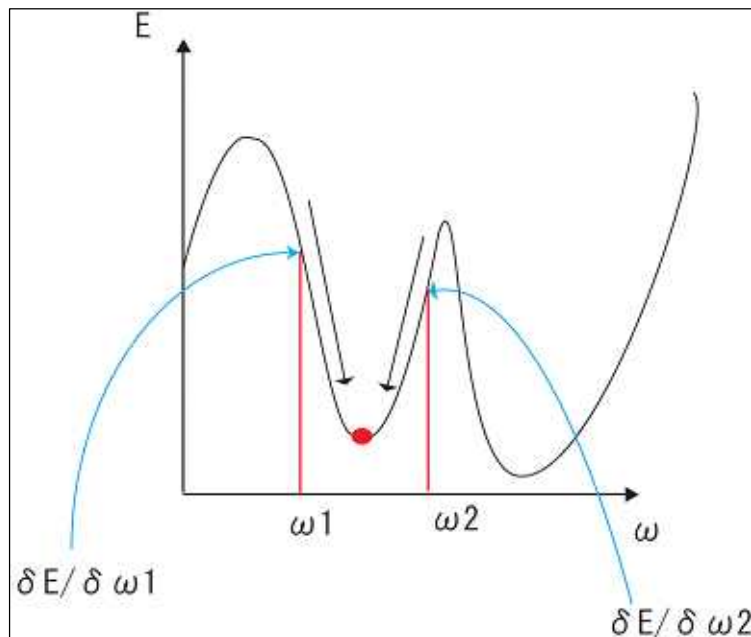


図 4 : 最急降下法の模式図

## 2.5 関数近似能力について

誤差逆伝播法による学習は、入力層に入力されたデータから計算された出力層ニューロンの出力値を教師信号と比較し、その誤差に応じて結合荷重を修正することによって出力値を教師信号に近似していくためのアルゴリズムである。ゆえにこの結果により構築されたネットワークはある入力ベクトル  $\mathbf{I}$  に対して適切な出力ベクトル  $\mathbf{O}$  を表現するための非線形関数  $F$  を近似していると言える。

関数近似を行う方法として他に有名なものにフーリエ級数による展開が挙げられるが、この2つの方法の一番の大きな違いは基底関数の性質である。フーリエ級数展開の基底関数は固定されている固定基底であるのに対して、誤差逆伝播法による学習は基底関数の形状を変化させることが出来る可変基底であり、固定基底の場合は入力次元が大きいために一定以上の近似精度を保障するためには必要な基底関数の数が膨大になる(次元の呪い)に対して、可変基底の場合は基底関数が増えるため次元の呪いを回避できる。こうした保障があるため、誤差逆伝播法による学習は応用上よく用いられている。



## 2.6 ニューラルネットワークの入力値について

### 2.6.1 株価の時系列データの前処理について

株価の時系列データには短期間に激しく変動させている短周期成分と長期間の間に緩やかに変動させている長周期成分が存在している。先行研究[1][2]の中では最初にローパスフィルタを用いることによって株価の時系列データから長周期成分を除去することによって短期の株価を正確に予測しようと試み、比較的正確な予測結果が得られている。しかし株価の時系列データから短周期成分と長周期成分を正確に分けることは難しく、また株価の短期予測の精度を求める際に長周期成分の影響を無視できるのか判断できないため、本論分ではローパスフィルタによる株価の時系列データの前処理は行わない。

### 2.6.2 正規化について

ニューラルネットワークへの入力には0~1の連続値であるため、株価の時系列を予測する際には株価の変動値を0~1の連続値の範囲になるように正規化して表現することである。以降に実験に使用した正規化方法について説明する。

以降の式に使用している変数

i番目の入力ユニットへの入力：  $input$

i番目の出力ユニットの出力：  $output$

i日目の株価：  $CP$

i-1日目からi日目への株価変動率：  $DCP = CP - CP$

予測に使用する期間内の最大株価変動率：  $maxDCP$

予測に使用する期間内の最小株価変動率：  $minDCP$

### 2.6.3 入力方法

入力値：  $input = (DCP - minDCP) / (maxDCP - minDCP)$

予測株価変動率：  $output \times (maxDCP - minDCP) + minDCP$

この入力方法は予測に使用する期間のDCPで最小株価変動率(minDCP)を0、最大株価変動率(maxDCP)を1になるように正規化した方法である。しかしこの方法は、入力の値の範囲を0~1にきれいに正規化することが出来るがネットワ

ークの出力も正規化された値なので元に戻す必要があり、ネットワークの出力から予測株価変動率を計算する際に予測に使用した日数内でのminDCP、maxDCPの影響がネットワークの出力に比べて非常に大きく、学習データに対してネットワークが正しく学習していたとしても、学習データの中のminDCP、maxDCPに大きく依存している特化したネットワークを作成している。ゆえに未学習データに使用するには有効ではないと考え、本研究ではこの入力方法は使用しない。

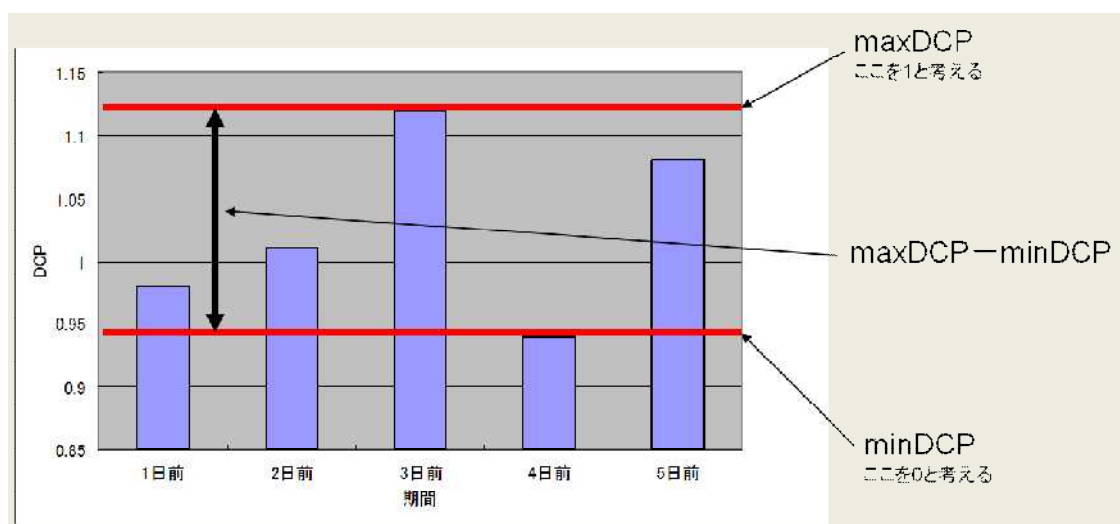


図 5： 入力方法 模式図

## 2.6.4 入力方法

入力値： $\text{input} = \text{DCP} - (\text{maxDCP} - 1.0)$

予測株価変動率： $\text{output} + (\text{maxDCP} - 1.0)$

この入力方法は上記の方法を改良したものでmaxDCPはこの方法でも使用するが、方法で予測株価変動率を計算する際に最も影響の大きいminDCPを使用しないことによってネットワークの出力の影響を大きくし、学習データへの特化を和らげ、汎化性のあるネットワークの構築が期待できる。

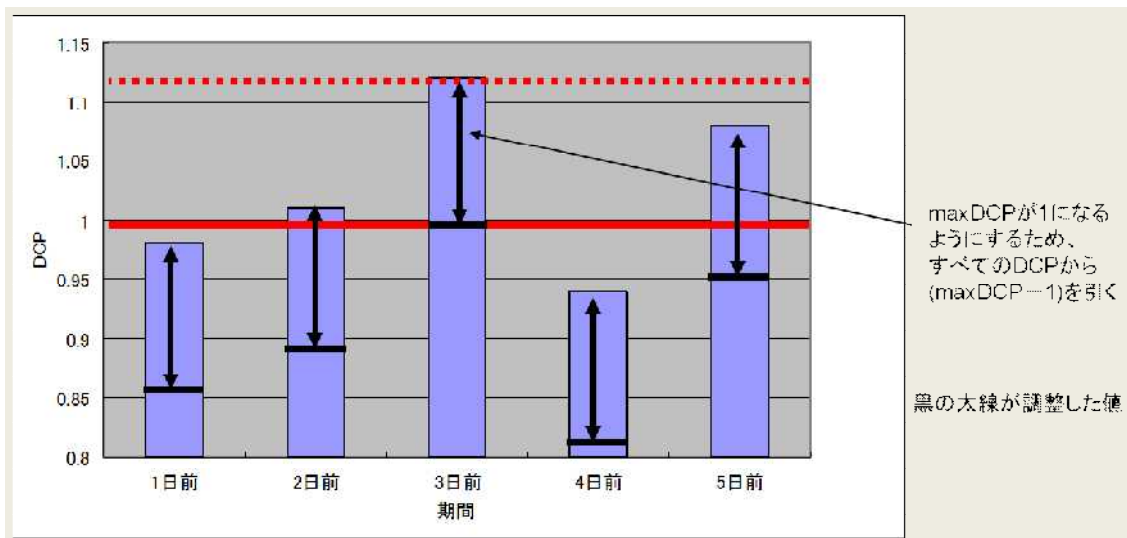


図 6： 入力方法 模式図

## 2.6.5 入力方法

$$\begin{aligned} \text{入力値：} \quad \text{input} &= \begin{cases} 1 : \text{if} ( \text{DCP} - 1.0 ) > 0 \\ 0 : \text{else} \end{cases} \\ \text{input} &= | \text{DCP} - 1.0 | \end{aligned}$$

$$\text{測株価変動率：} \quad S(\text{output}) \times \text{output}, \quad S(x) = \begin{cases} 1 : x > 0.5 \\ -1 : x < 0.5 \end{cases}$$

この入力方法は株価の変動を上昇しているか下降しているか、どれくらい変動しているかの2つの情報に分割して入力値として与える方法である。ゆえに入力ユニットは2ユニットで1セットである。入力値と同様に出力値も上昇下降の方向を表すユニットと変動量を表すユニット2つで構成し、2つの出力値のみ予測株価変動率を計算する。これにより予測株価変動率を計算する際にmaxDCP、minDCPといった学習データに特化した情報を使用する必要がなくなり、汎化性の高いネットワークが構築できると考えられる。しかしネットワークへの学習が株価の上昇下降の方向と変動量の2つになるため学習の難易度が上がると考えられる。

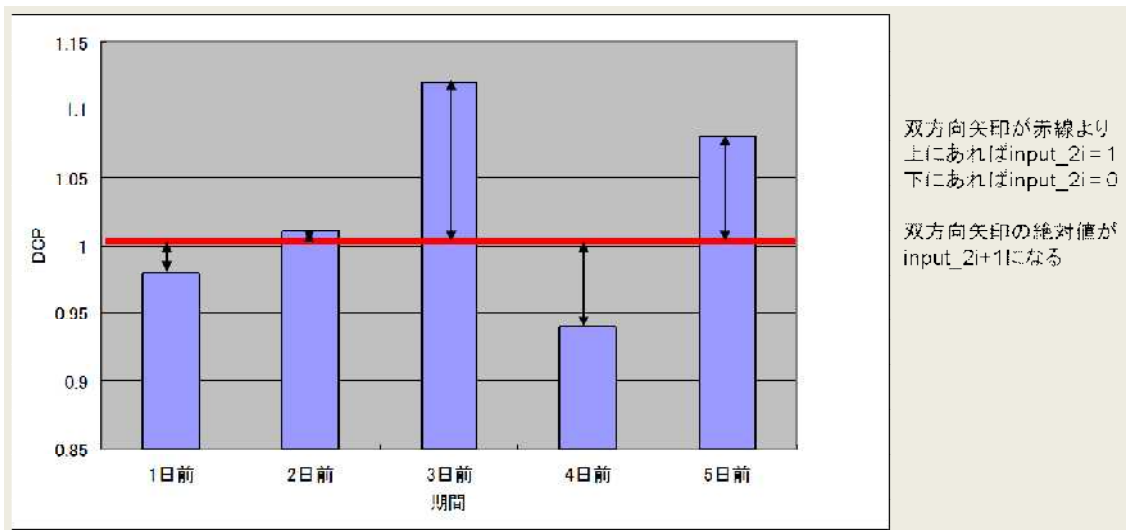


図 7： 入力方法 模式図

### 3. 結果

本研究における株価予測の精度上昇の着目点は入力方法と学習に使用するデータについてである。最初に入力方法、入力方法の入力方法で作成したニューラルネットワークを使用しての学習データに対する精度について確認することによって入力方法と入力方法の優劣を判断した。ニューラルネットワークを作成する際に使用したデータは東証1部の売買高・出来高の上位から選んだ50社であり、業種コードごとに分類してネットワークを構築した。

また学習期間外データに対する予測を行うニューラルネットワークの作成には東証1部の売買高・出来高の上位から選んだ50社の業種コードごとに分類し作成したネットワーク、東証1部の売買高・出来高の上位から選んだ300社の業種コードごとに分類し作成したネットワーク、予測する対象の企業のデータのみで作成したネットワーク、前述の300社すべてのデータから作成したネットワークを使用して予測精度の差を比較した。

なお業種コードは証券コード協議会の中分類33種で分類している。

ニューラルネット作成時のパラメータは付録に記載しておく。

以降のグラフの横軸は期間、縦軸は株価である。

### 3.1 学習データでの2つの入力方法の比較

入力方法の比較には方法、方法共に前述の50社の内の業種コード7100(証券・商品先物取引)、業種コード4050(電気・ガス業)に分類されている計5社の1日後の株価を予測することで確認した。

業種コード7100のネットワークの作成には野村ホールディングス、大和証券グループ本社の2社、業種コード4050のネットワークの作成には関西電力、東京電力、東京ガスの3社を選択し、業種コード7100、業種コード4050共に期間(2005/1/5～2006/12/29)の2年間で作成した。

#### 3.1.1 入力方法 入力値： $\text{input} = \text{DCP} - (\text{maxDCP} - 1.0)$ による結果

業種コード7100である2社の株価変動の波形はかなり類似性の高い。そのためだと考えられるが2社の株価を予測した結果、実際の株価と予測株価を比較すると平均誤差、最大誤差は近いものになっており、また波形を確認してみると大きな誤差が発生している時期もかなり一致している。

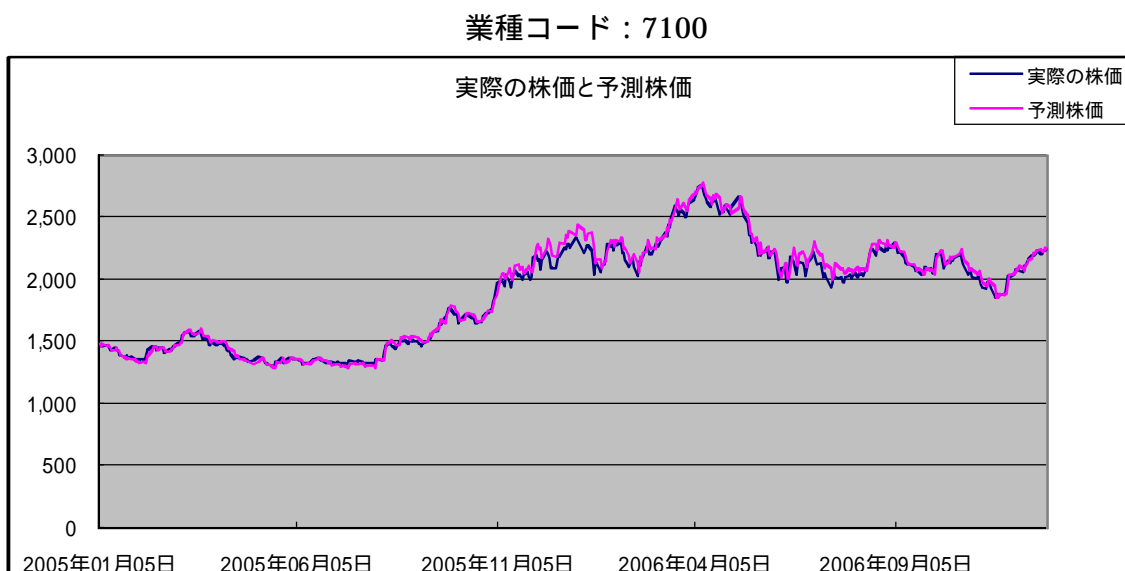


図8： 野村ホールディングス(平均誤差：0.018922 最大誤差：0.10096)

業種コード：7100

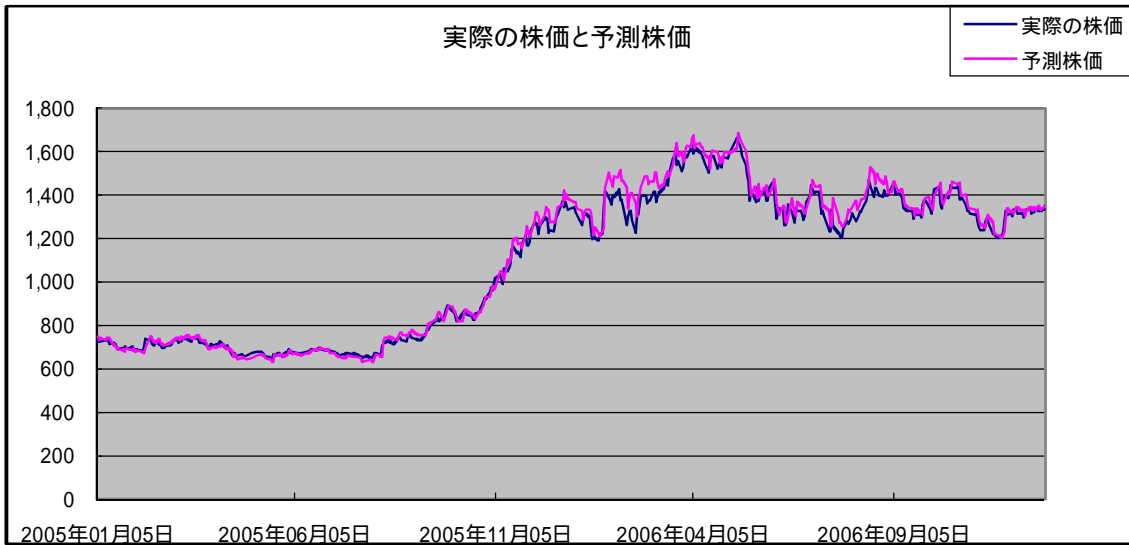


図 9： 大和証券グループ本社(平均誤差：0.025573 最大誤差：0.139617)

業種コード 4050 である 3 社の内、関西電力、東京電力の 2 社の波形の類似度はかなり高いが、東京ガスの波形は似てはいるが前述の 2 社に比べると変動が激しい。これは業種コードの分類では電気業とガス業は同じ業種に分類されているが、電気業とガス業との差が現れたためだと考えられる。しかし実際の株価と予測株価を比較したときの平均誤差、最大誤差は 3 社とも似た値になっている。また平均誤差、最大誤差が業種コード 7100 の 2 社と比較して小さい理由は株価の変動が業種コード 4050 の 3 社は少ないためだと考えられる。

業種コード：4050

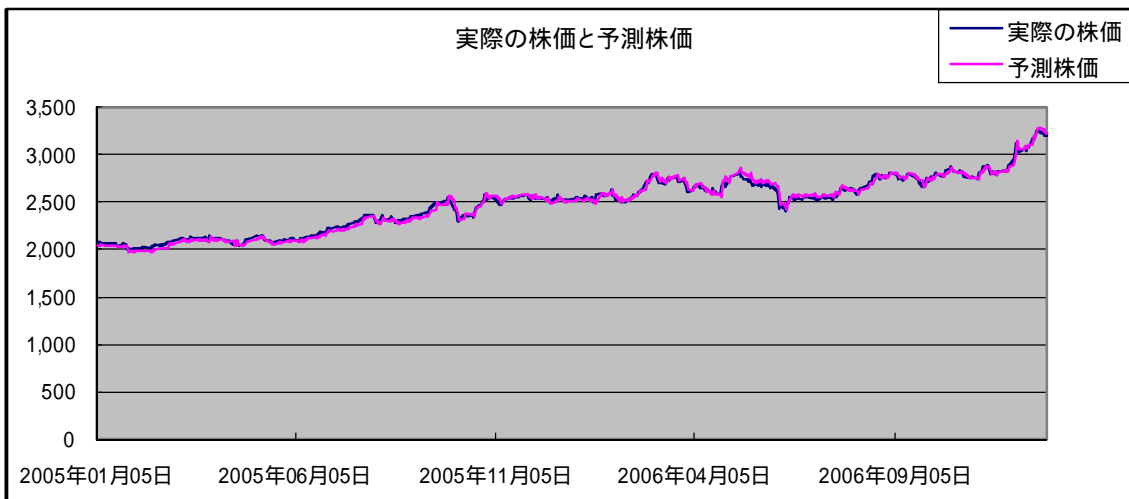


図 10： 関西電力(平均誤差：0.011529 最大誤差：0.071467)

業種コード：4050

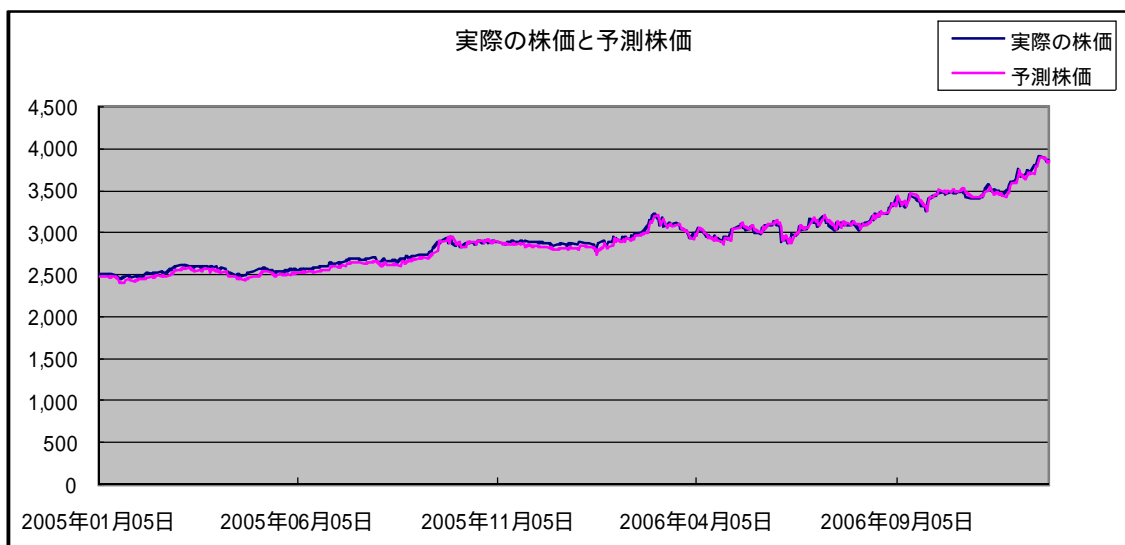


図 11： 東京電力(平均誤差：0.011601 最大誤差：0.066383)

業種コード：4050

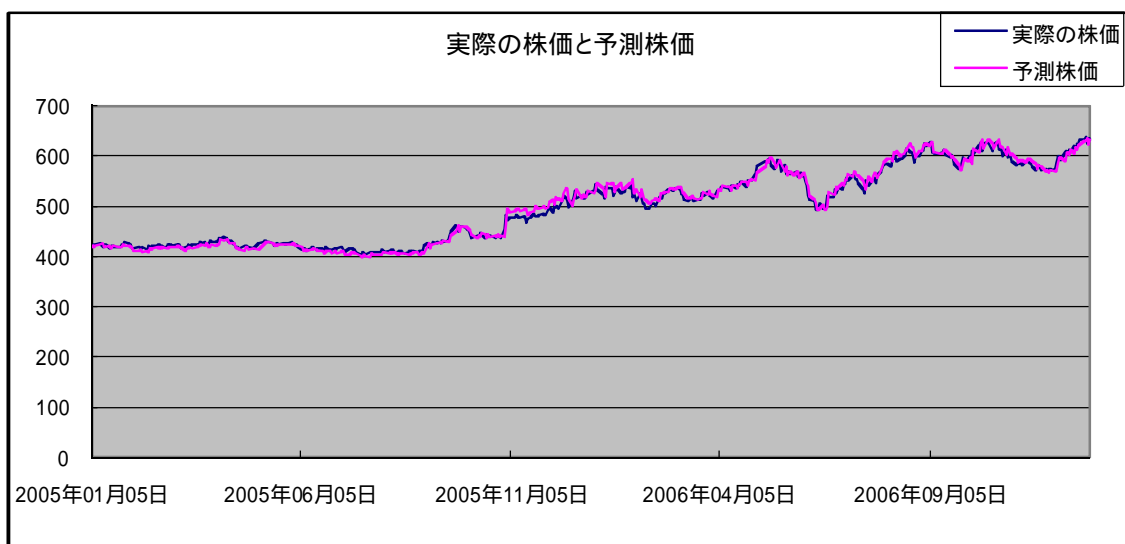


図 12： 東京ガス(平均誤差：0.013747 最大誤差：0.068009)

3.1.2 入力方法  $\text{input} = \begin{cases} 1 & \text{if } (DCP - 1.0) > 0 \\ 0 & \text{else} \end{cases}$ ,  $\text{input} = |DCP - 1.0|$

による結果

業種コード 7100 である 2 社の実際の株価と予測株価を比較すると株価が急変動している場所でも正確に予測できていることが見て取れる。また平均誤差、最大誤差で判断しても 2 社共にかなり小さい誤差となっており、平均誤差が 1% を下回っていることから予測精度の高さがわかる。

業種コード：7100

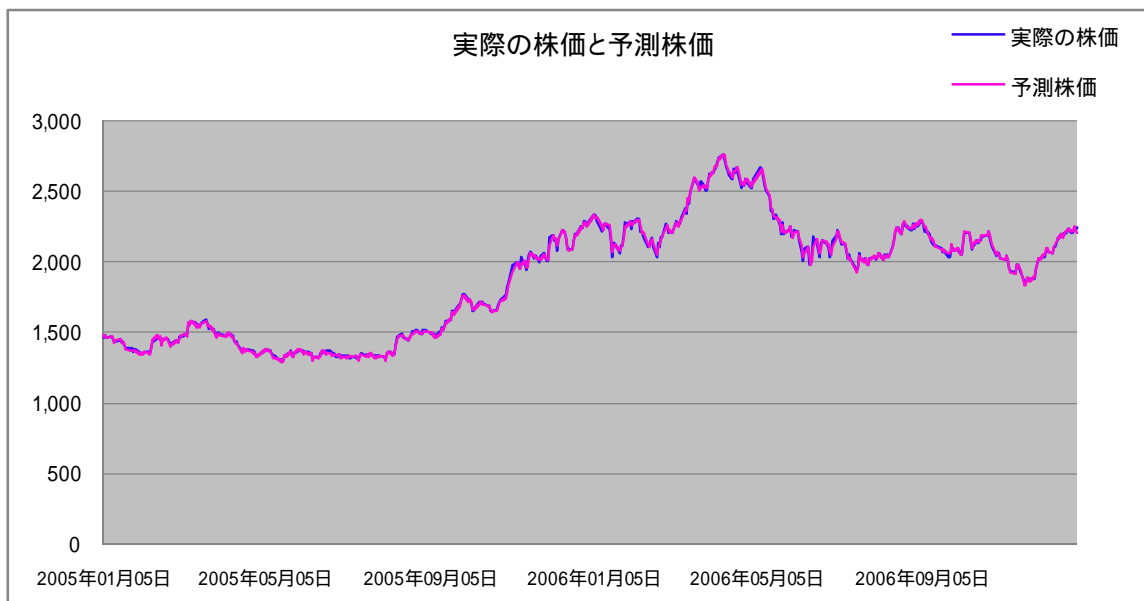


図 13：野村ホールディングス(平均誤差：0.005447 最大誤差：0.034324943)



業種コード：7100

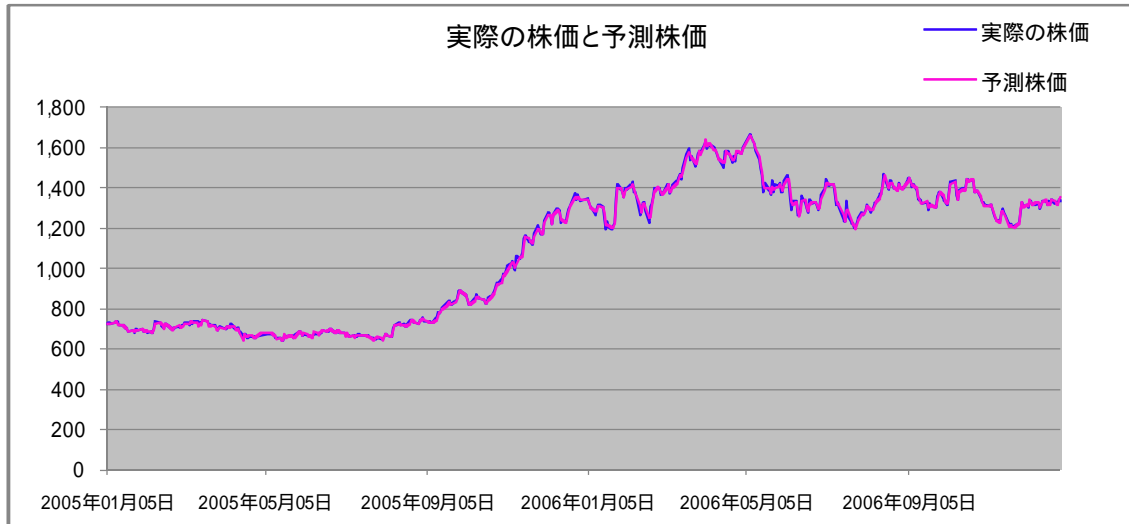


図 14：大和証券グループ本社(平均誤差：0.006933 最大誤差：0.040413534)

業種コード 4050 の 3 社についても業種コード 7100 の 2 社と同様に予測株価が実際の株価を正確に追従できていることが見て取れる。平均誤差、最大誤差についてもかなり小さいもので、3 社共に平均誤差が 1%を下回っていることから予測精度の高さがわかる。

業種コード：4050

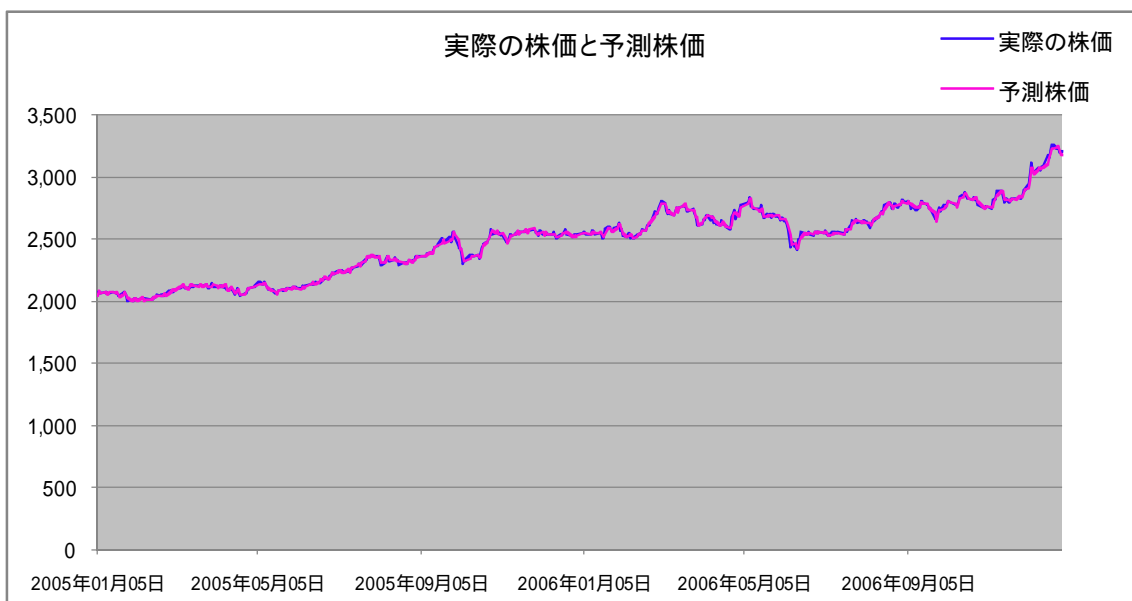


図 15：関西電力(平均誤差：0.004607 最大誤差：0.040983607)

業種コード：4050

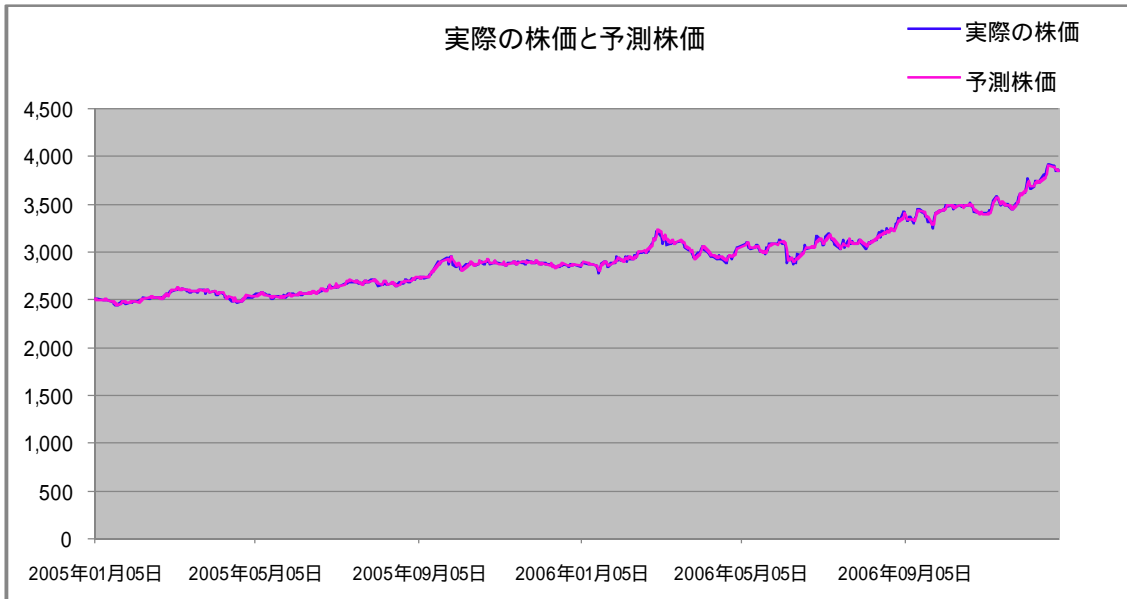


図 16：東京電力(平均誤差：0.004474 最大誤差：0.048442907)

業種コード：4050

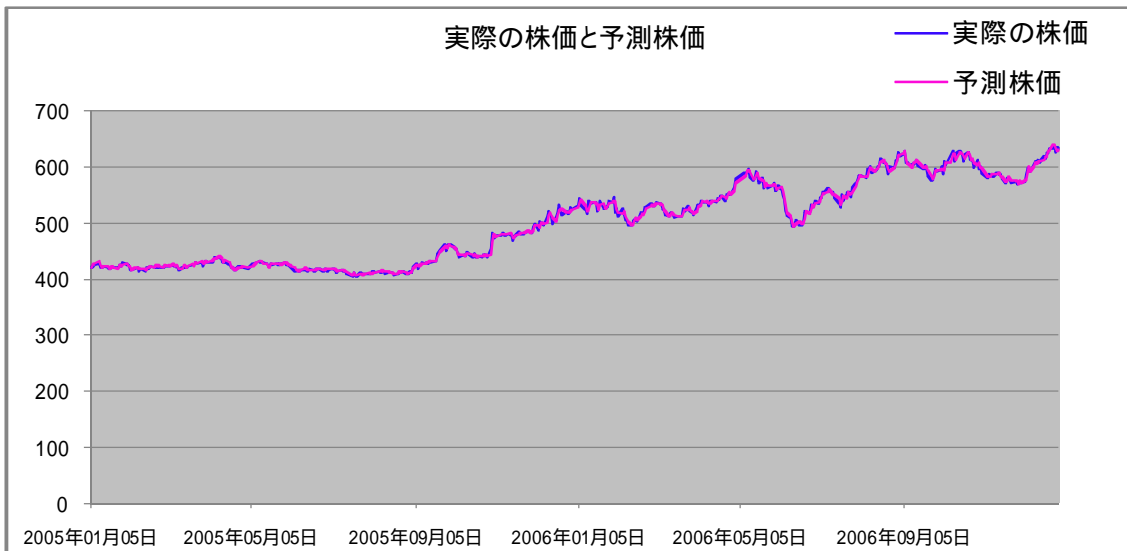


図 17：東京ガス(平均誤差：0.006023 最大誤差：0.033195021)

表1は入力方法、入力方法によって予測した計5社の平均誤差、最大誤差である。5社とも入力方法よりも入力方法で作成したニューラルネットワークでの予測のほうが平均誤差、最大誤差共に小さいのがわかる。この結果から入力方法のほうが入力方法よりも優れていると判断した。

入力方法	野村ホールディングス	大和証券グループ本社	関西電力	東京電力	東京ガス
平均誤差	0.018922	0.025573	0.011529	0.011601	0.013747
最大誤差	0.10096	0.139617	0.071467	0.066383	0.068009
入力方法	野村ホールディングス	大和証券グループ本社	関西電力	東京電力	東京ガス
平均誤差	0.005447	0.006933	0.004607	0.004474	0.006023
最大誤差	0.034324	0.040413	0.040983	0.048442	0.033195

表1：入力方法と入力方法との比較(入力方法でを使用した5社での比較)

### 3.2 入力方法を用いての学習期間外での予測結果

3.1の結果から入力方法のほうが高精度に予測できていることから、入力方法を使用しての学習期間外での1日後の株価の予測精度を確認した。精度確認のために作成したニューラルネットワークは前述の50社の内の各業種コードごとに作成したネットワーク、前述の300社の内の各業種コードごとに作成したネットワーク、各企業のデータのみで作成したネットワーク、前述の300社すべてのデータで作成したネットワークである。なおすべてのネットワークの学習期間は(2006/1/4～2006/12/29)の1年間である。

### 3.2.1 野村ホールディングスの学習期間外での予測結果

図 18 は前述の 300 社の内、野村ホールディングスと同じ業種コードの企業 4 社のデータで作成したニューラルネットワークで予測したものである。予測株価は実際の株価を正確に追従できている期間は少なく、また最大誤差を見ればわかるが、実際の株価から大きく外れた株価を予測することもあり正しく予測できているとは言えない。

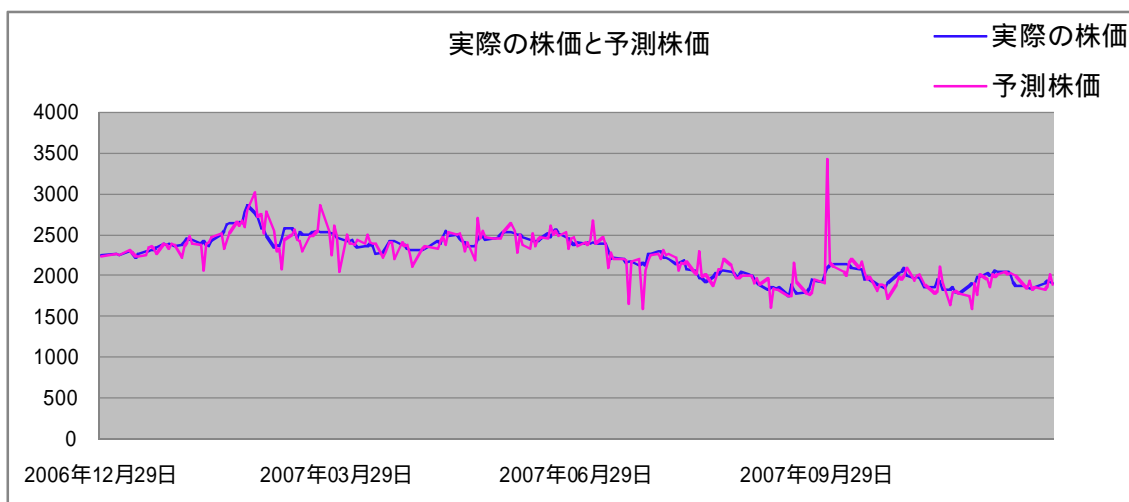


図 18：同業種 4 社(平均誤差：0.037021 最大誤差：0.636842105)

図 19 は前述の 50 社の内、野村ホールディングスと同じ業種コードの企業 2 社のデータで作成したニューラルネットワークで予測したものである。予測株価は実際の株価を比較的正確に追従できており、また最大誤差から見て実際の株価のから大きく外れた予測をしている場所は少ないことがわかる。

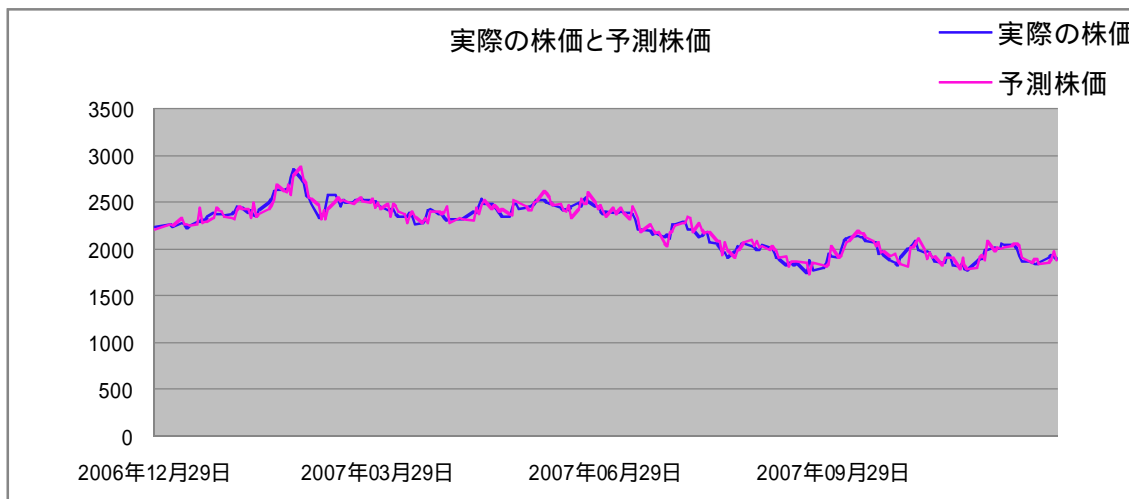


図 19：同業種 2 社(平均誤差：0.022412 最大誤差：0.0945)

図 20 は自社のデータのみで作成したニューラルネットワークで予測したものである。予測株価は図 19 のものほどではないが比較的正確に追従できており、また最大誤差から見ても図 19 よりは大きくなってはいるが、実際の株価から大きく外れた予測をしている場所は少ない。

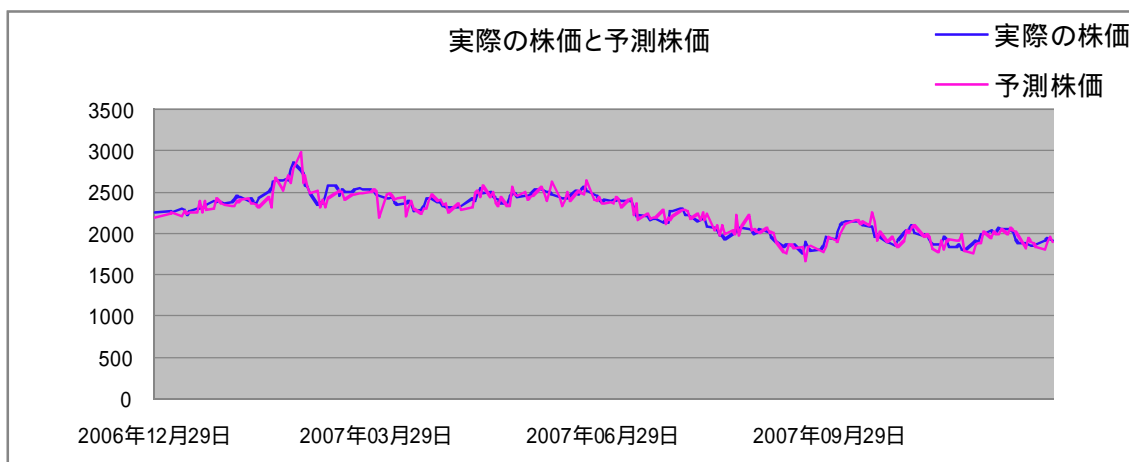


図 20 : 自社のみ(平均誤差 : 0.027325 最大誤差 : 0.118553961)

図 21 は前述の 300 社すべてのデータで作成したニューラルネットワークで予測したものである。実際の株価とはまったく違う株価を予測しており、予測しているとは言い難い結果となった。

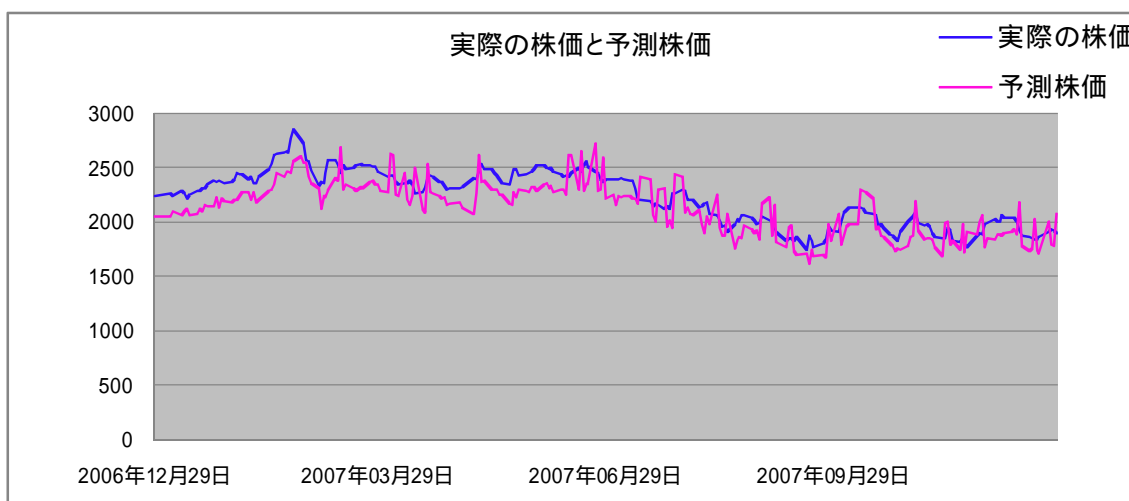


図 21 : 300 企業(平均誤差 : 0.069454 最大誤差 : 0.141945774)

### 3.2.2 松下電器産業の学習期間外での予測結果

図 22 は前述の 300 社の内、松下電器産業と同じ業種コードの企業 29 社のデータで作成したニューラルネットワークで予測したものである。最大誤差も小さく、予測株価が実際の株価を比較的正確に追従できている。

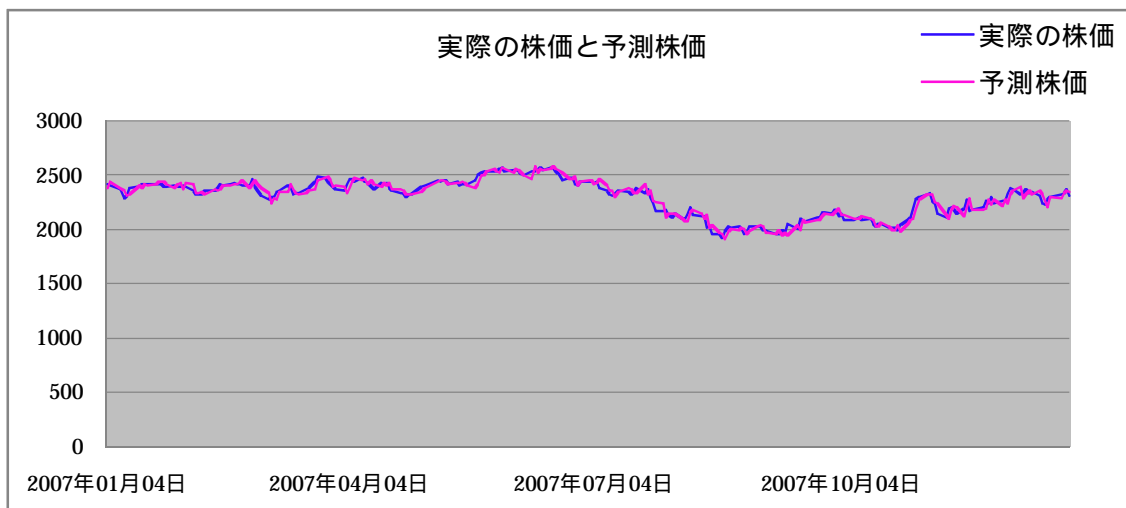


図 22 : 同業種 29 社(平均誤差 : 0.025939 最大誤差 : 0.116411)

図 23 は前述の 50 社の内、松下電器産業と同じ業種コードの企業 3 社のデータで作成したニューラルネットワークで予測したものである。平均誤差、最大誤差共に図 22 の半分になっており、図 22 よりもさらに予測株価が実際の株価を正確に追従できている。

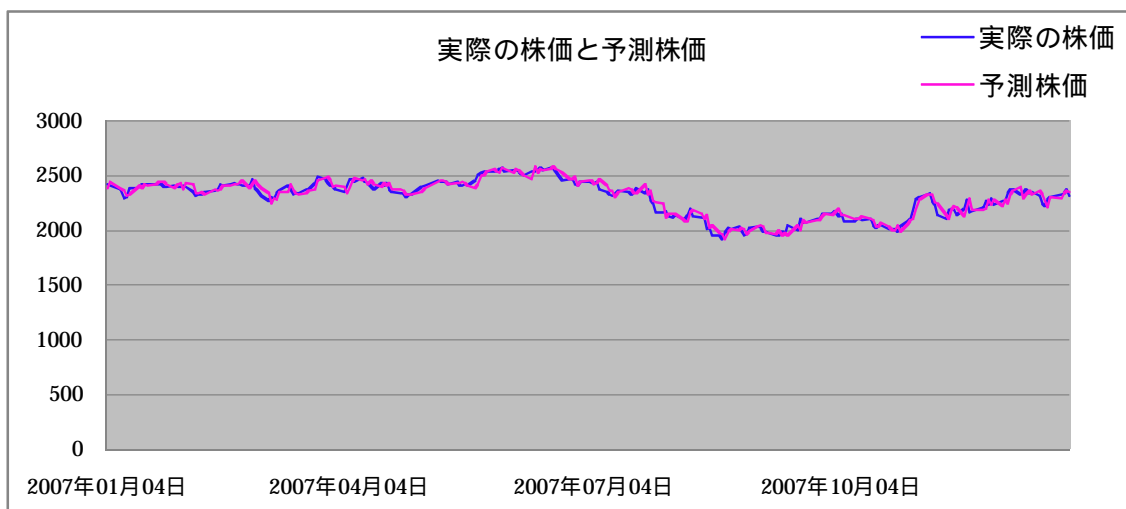


図 23 : 同業種 3 社(平均誤差 : 0.014503 最大誤差 : 0.058065)

図 24 は自社のデータのみで作成したニューラルネットワークで予測したものである。比較的正確に株価を予測できてはいるが、図 23 の予測ほど正確ではない。

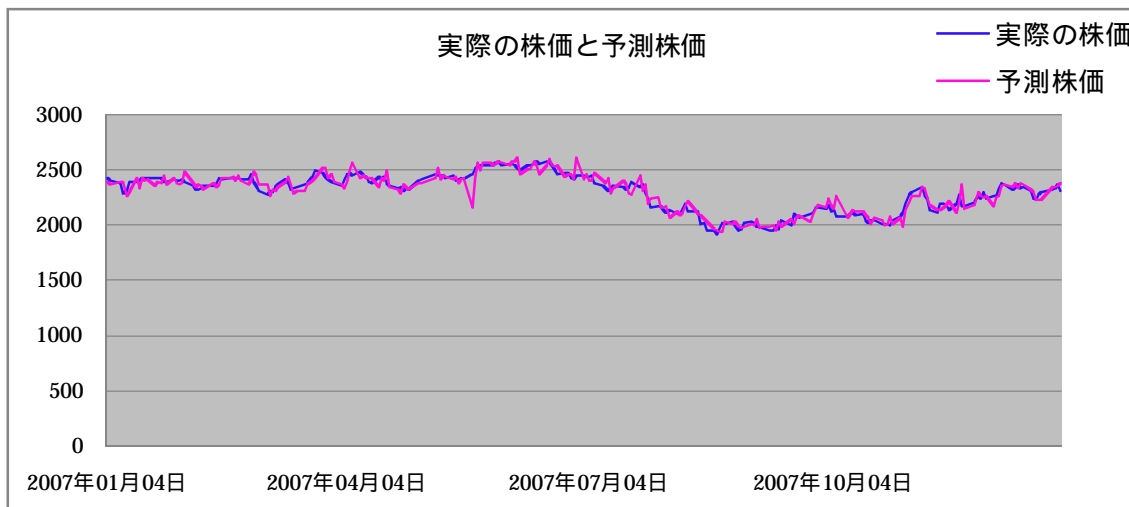


図 24 : 自社のみ(平均誤差 : 0.018501 最大誤差 : 0.12057)

図 25 は前述の 300 社すべてのデータで作成したニューラルネットワークで予測したものである。実際の株価とはまったく違う株価を予測しており、予測しているとは言い難い結果となった。

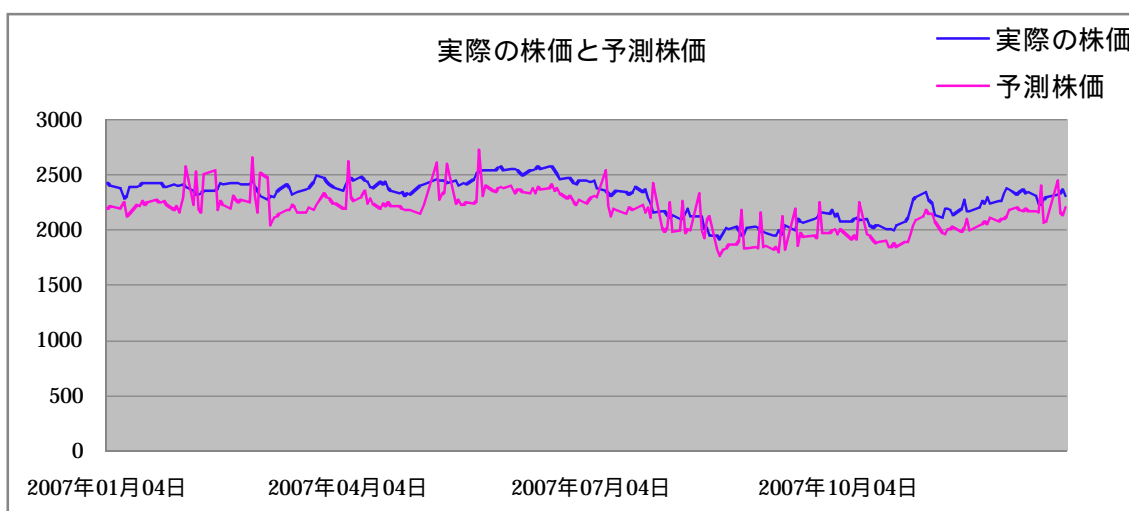


図 25 : 300 企業(平均誤差 : 0.072191 最大誤差 : 0.122402)

結果として野村ホールディングス、松下電器産業の2社共に学習期間外の株価を予測する際には同業種2,3社で作成したニューラルネットワークで予測したときに1番予測精度が高いという結果になった。また300社すべてのデータで学習させたネットワークでの予測は2社共に実際の株価から大きく外れた予測結果であった。そして自社のデータのみで作成したニューラルネットワークよりも同業種2,3社で作成したニューラルネットワークでの予測精度が高いことから、株価を予測させるニューラルネットワークの作成におけるデータは予測させる企業と同業種の企業のデータを適切数選択し、学習させることが最適だという結果になった。



## 4. 考察

株価予測の方法としてニューラルネットワークを使用する際の着目点として本研究では入力値と学習データについて検討した。

入力方法については結果 3.1 から入力方法 よりも、入力方法 のほうがよい結果が得られているのがわかる。これは入力方法 と比較して入力方法 のほうが元の株価データへの処理を少なくし、正確なデータを学習できたためだと考えられる。入力方法 による学習データに対する株価予測は 2 年間という決して長い期間ではないが、1 日後の株価をほぼ完全に再現できており、前述の誤差逆伝播法による学習の特徴の 1 つである関数近似能力という点から考えると、一定期間の過去の株価データを入力値として 1 日後の株価を出力する関数を再現できていると言え、これは次期の株価を予測することのできる関数が存在する可能性があると考えられる。

学習データの選択については結果 3.2 から同業種 2,3 社を学習データとして選択したときが最もよい予測結果となり、最も悪い結果は 300 社すべてを学習データとして選択したときである。300 社の学習データのものが 1 番悪い結果となったのは、予測する対象の企業にはその企業特有の株価変動特性があり、それに関連性のない企業のデータはノイズとなって学習を阻害したためだと考えられる。また興味深い結果として、学習外データを予測する際に自社のデータのみで学習させたニューラルネットワークより同業種 2,3 社のデータを学習させたニューラルネットワークのほうが高い予測精度であったことが挙げられ、これは先ほど述べていた対象企業特有の株価変動特性を否定させる結果である。この原因として「対象企業の現時点では予測できない要因(ノイズ)」が存在し、ノイズの影響を大きく受けたニューラルネットワークが作成されてしまっているのに対して、株価変動特性の似ている同業種のデータも使用することによって対象企業だけに影響のある要因(ノイズ)が発生してしたとしても、対象企業と同業種のデータはノイズのない変動をしているため、同業種のデータを含めることでノイズの影響が少ないニューラルネットワークが作成できているためだと考えられる。

しかし、入力方法 を使用しての株価予測は株価の上昇下降(方向一致率)はどの予測でも 50%程度になっていた。単純に株価の方向一致率のみを予測した場合の予測精度は 70%程度になっていることから考えてもかなり低いものになっている。これは入力値に株価の上昇下降の情報以外に変動量を使用しているが、これがノイズとなって予測精度を減少させていると考えられる。今後の課題として「株価の方向予測」と「株価の変動量」の 2 つの予測を両立させることが挙げられる。株価の方向一致率の予測精度が低い理由は入力値に株価の変動方向データと変動量データを一緒にして使用しているため変動量データがノイズになっていると考えられるため、変動量と変動方向の予測を行うニューラルネットワークをそれぞれ別に作成することによってノイズを減らし予測精度が上昇すると考えられる。

## 謝辞

本研究及び論文作成にあたって、ご指導と助言をいただいた北野勝則准教授に厚く御礼申し上げます。また本研究を進めていく上で様々な議論を賜った計算神経科学研究室の皆様にも心より感謝いたします。

最後になりましたが、プログラム作成及び研究方法にあたって有益な助言をいただいた佐藤裕望君にも心より御礼申し上げます。

## 参考文献

- [1] 齋藤進, 神田晋太郎, “フィルタリングによる日経平均株価の予測”, 日本オペレーションズ・リサーチ学会, pp.242-243, 2001
- [2] 高穂洋, 荒井隆行, 大竹敢, 田中衛, “ニューラルネットワークによる次期の株価予測 株価予測におけるフィルタリングによる特徴量抽出”, 信学技報, pp.13-16, November, 2002
- [3] 山口和孝, “ニューラルネットと遺伝的アルゴリズムを用いた株式売買支援システム”, 修士論文, 東京大学大学院情報理工学系研究科, 2003
- [4] 石本大貴, 中井暉久, “重回帰分析とAICを用いたニューラルネットワークによる株価予測”, 日本オペレーションズ・リサーチ学会, pp.204-205, 2006

## 付録

学習期間全体の学習回数	:	1000 回
各日にちの学習回数	:	100 回
学習率	:	1.0
入出力関数(シグモイド関数)の傾き	:	1.0
各層のユニット数(iの値は入力方法 は 1、入力方法 は 2)		
入力層	:	$30 \times i$
隠れ層	:	$60 \times i$
出力層	:	1

### 東証 1 部の売買高・出来高の上位から選んだ 50 社

企業コード	企業名	業種コード
1928	積水ハウス	2050
2503	麒麟ビール	3050
2914	J T	3050
3402	東レ	3100
4063	信越化学工業	3200
4502	武田薬品工業	3250
4568	第一三共	3250
4901	富士写真フイルム	3200
5001	新日本石油	3300
5108	ブリヂストン	3350
5201	旭硝子	3400
5401	新日本製鐵	3450
5411	JFE ホールディングス	3450
5802	住友電気工業	3500
6301	コマツ	3600
6460	セガサミーホールディングス	3600
6501	日立製作所	3650
6752	松下電器産業	3650

7011	三菱重工業	3600
7203	トヨタ自動車	3700
7267	ホンダ	3700
7741	HOYA	3750
7751	キヤノン	3650
7912	大日本印刷	3800
8031	三井物産	6050
8058	三菱商事	6050
8253	クレディ セゾン	7200
8267	イオン	6100
8306	三菱UFJフィナンシャル・グループ	7050
8316	三井住友フィナンシャルグループ	7050
8411	みずほフィナンシャルグループ	7050
8591	オリックス	7200
8601	大和証券グループ本社	7100
8604	野村ホールディングス	7100
8752	三井住友海上火災保険	7150
8766	ミレアホールディングス	7150
8801	三井不動産	8050
8802	三菱地所	8050
9020	東日本旅客鉄道	5050
9022	東海旅客鉄道	5050
9104	商船三井	5100
9432	NTT	5250
9433	KDDI	5250
9437	NTTドコモ	5250
9501	東京電力	4050
9503	関西電力	4050
9531	東京瓦斯	4050
9735	セコム	9050
9831	ヤマダ電機	6100
9984	ソフトバンク	6050

東証 1 部の売買高・出来高の上位から選んだ 300 社

企業コード	企業名	業種コード
1332	日本水産	50
1721	コムシスホールディングス	2050
1801	大成建設	2050
1802	大林組	2050
1803	清水建設	2050
1808	長谷工コーポレーション	2050
1812	鹿島建設	2050
1860	戸田建設	2050
1878	大東建託	2050
1925	大和ハウス工業	2050
1928	積水ハウス	2050
1942	関電工	2050
1951	協和エクシオ	2050
1963	日揮	2050
2002	日清製粉グループ本社	3050
2261	明治乳業	3050
2267	ヤクルト本社	3050
2282	日本ハム	3050
2331	総合警備保障	9050
2433	博報堂 D Y ホールディングス	9050
2501	サッポロホールディングス	3050
2502	アサヒビール	3050
2503	キリンビール	3050
2579	コカ・コーラウエストジャパン	3050
2593	伊藤園	3050
2651	ローソン	6100
2670	エービーシー・マート	6100
2730	エディオン	6100
2768	双日	6050
2779	三越	6100
2784	アルフレッサHD	6050

2801	キッコーマン	3050
2802	味の素	3050
2897	日清食品	3050
2914	J T	3050
3101	東洋紡績	3100
3105	日清紡績	3100
3332	J F E 商事ホールディングス	6050
3337	サークルKサンクス	6100
3382	7 & I H D	6100
3401	帝人	3100
3402	東レ	3100
3404	三菱レイヨン	3100
3405	クラレ	3100
3407	旭化成	3200
3591	ワコールホールディングス	3100
3861	王子製紙	3150
3893	日本製紙グループ本社	3150
4004	昭和電工	3200
4005	住友化学	3200
4043	トクヤマ	3200
4062	イビデン	3650
4063	信越化学工業	3200
4091	大陽日酸	3200
4118	カネカ	3200
4182	三菱ガス化学	3200
4183	三井化学	3200
4185	J S R	3200
4188	三菱ケミカルホールディングス	3200
4204	積水化学工業	3200
4208	宇部興産	3200
4217	日立化成工業	3200
4307	野村総合研究所	5250
4321	ケネディクス	9050
4324	電通	9050
4452	花王	3200
4502	武田薬品工業	3250

4503	アステラス製薬	3250
4507	塩野義製薬	3250
4519	中外製薬	3250
4523	エーザイ	3250
4535	大正製薬	3250
4543	テルモ	3750
4568	第一三共	3250
4631	大日本インキ化学工業	3200
4661	オリエンタルランド	9050
4676	フジテレビジョン	5250
4680	ラウンドワン	9050
4681	リゾートトラスト	9050
4704	トレンドマイクロ	5250
4716	日本オラクル	5250
4723	グッドウィル・グループ	9050
4732	ユー・エス・エス	9050
4756	CCC	9050
4768	大塚商会	5250
4901	富士写真フイルム	3200
4902	コニカミノルタ	3650
4911	資生堂	3200
5001	新日本石油	3300
5016	新日鉱ホールディングス	3300
5108	ブリヂストン	3350
5110	住友ゴム工業	3350
5201	旭硝子	3400
5214	日本電気硝子	3400
5233	太平洋セメント	3400
5333	日本碍子	3400
5334	日本特殊陶業	3400
5401	新日本製鐵	3450
5405	住友金属工業	3450
5406	神戸製鋼所	3450
5407	日新製鋼	3450
5411	JFEホールディングス	3450
5471	大同特殊鋼	3450

5486	日立金属	3450
5631	日本製鋼所	3600
5711	三菱マテリアル	3500
5713	住友金属鉱山	3500
5801	古河電気工業	3500
5802	住友電気工業	3500
5803	フジクラ	3500
5901	東洋製罐	3550
5929	三和シャッター工業	3550
5938	住生活グループ	3550
5943	ノーリツ	3550
5947	リンナイ	3550
5991	日本発条	3550
6113	アマダ	3600
6146	ディスコ	3600
6201	豊田自動織機	3700
6273	S M C	3600
6301	コマツ	3600
6302	住友重機械工業	3600
6305	日立建機	3600
6326	クボタ	3600
6361	荏原製作所	3600
6366	千代田化工建設	2050
6367	ダイキン工業	3600
6370	栗田工業	3600
6417	S A N K Y O	3600
6448	ブラザー工業	3600
6460	セガサミーホールディングス	3600
6471	日本精工	3600
6472	N T N	3600
6473	ジェイテクト	3600
6479	ミネベア	3650
6481	T H K	3600
6501	日立製作所	3650
6502	東芝	3650
6503	三菱電機	3650



6504	富士電機ホールディングス	3650
6581	日立工機	3600
6586	マキタ	3600
6701	NEC	3650
6702	富士通	3650
6723	NECエレクトロニクス	3650
6724	セイコーエプソン	3650
6752	松下電器産業	3650
6753	シャープ	3650
6758	ソニー	3650
6762	TDK	3650
6764	三洋電機	3650
6773	パイオニア	3650
6806	ヒロセ電機	3650
6841	横河電機	3650
6857	アドバンテスト	3650
6861	キーエンス	3650
6902	デンソー	3700
6923	スタンレー電気	3650
6952	カシオ計算機	3650
6954	ファナック	3650
6967	新光電気工業	3650
6971	京セラ	3650
6988	日東電工	3200
6991	松下電工	3650
7011	三菱重工業	3600
7012	川崎重工業	3700
7013	石川島播磨重工業	3600
7201	日産自動車	3700
7202	いすゞ自動車	3700
7203	トヨタ自動車	3700
7211	三菱自動車工業	3700
7259	アイシン精機	3700
7261	マツダ	3700
7267	ホンダ	3700
7269	スズキ	3700

7272	ヤマハ発動機	3700
7453	良品計画	6100
7459	メディセオ・パルタックホールディングス	6050
7532	ドン・キホーテ	6100
7731	ニコン	3750
7733	オリンパス	3750
7741	HOYA	3750
7751	キヤノン	3650
7752	リコー	3650
7832	バンナムHD	3800
7862	トッパンフォームズ	3800
7911	凸版印刷	3800
7912	大日本印刷	3800
7915	日本写真印刷	3800
7936	アシックス	3800
7951	ヤマハ	3800
7966	リンテック	3800
7984	コクヨ	3800
8001	伊藤忠商事	6050
8002	丸紅	6050
8012	長瀬産業	6050
8015	豊田通商	6050
8016	オンワード樫山	3100
8028	ファミリーマート	6100
8031	三井物産	6050
8035	東京エレクトロン	3650
8036	日立ハイテクノロジーズ	6050
8053	住友商事	6050
8056	日本ユニシス	5250
8058	三菱商事	6050
8060	キヤノンMJ	6050
8078	阪和興業	6050
8113	ユニ・チャーム	3200
8129	東邦薬品	6050
8136	サンリオ	6050
8184	島忠	6100

8218	コメリ	6100
8227	しまむら	6100
8233	高島屋	6100
8234	大丸	6100
8238	伊勢丹	6100
8242	阪急百貨店	6100
8252	丸井	6100
8253	クレディ セゾン	7200
8267	イオン	6100
8268	西友	6100
8270	ユニー	6100
8273	イズミ	6100
8281	ゼビオ	6100
8282	ギガスケーズデンキ	6100
8303	新生銀行	7050
8306	三菱UFJフィナンシャル・グループ	7050
8308	りそなホールディングス	7050
8309	三井トラスト・ホールディングス	7050
8316	三井住友フィナンシャルグループ	7050
8326	福岡銀行	7050
8328	札幌北洋ホールディングス	7050
8331	千葉銀行	7050
8332	横浜銀行	7050
8333	常陽銀行	7050
8355	静岡銀行	7050
8377	ほくほくフィナンシャルグループ	7050
8403	住友信託銀行	7050
8404	みずほ信託銀行	7050
8411	みずほフィナンシャルグループ	7050
8473	SBIホールディングス	7200
8515	アイフル	7200
8564	武富士	7200
8572	アコム	7200
8574	プロミス	7200
8583	UFJニコス	7200
8591	オリックス	7200

8601	大和証券グループ本社	7100
8604	野村ホールディングス	7100
8606	新光証券	7100
8615	三菱UFJ証券	7100
8752	三井住友海上火災保険	7150
8766	ミレアホールディングス	7150
8801	三井不動産	8050
8802	三菱地所	8050
8804	東京建物	8050
8815	東急不動産	8050
8830	住友不動産	8050
8840	大京	8050
8868	アーバンコーポレイション	8050
8870	住友不動産販売	8050
8874	ジョイント コーポレーション	8050
9005	東京急行電鉄	5050
9020	東日本旅客鉄道	5050
9021	西日本旅客鉄道	5050
9022	東海旅客鉄道	5050
9041	近畿日本鉄道	5050
9042	阪急ホールディングス	5050
9062	日本通運	5050
9064	ヤマトホールディングス	5050
9101	日本郵船	5100
9104	商船三井	5100
9202	全日本空輸	5150
9301	三菱倉庫	3100
9364	上組	3100
9401	TBS	5250
9404	日本テレビ放送網	5250
9432	NTT	5250
9433	KDDI	5250
9437	NTTドコモ	5250
9501	東京電力	4050
9503	関西電力	4050
9531	東京瓦斯	4050

9532	大阪瓦斯	4050
9602	東宝	5250
9613	NTTデータ	5250
9684	スクウェア・エニックス	5250
9735	セコム	9050
9737	CSKホールディングス	5250
9744	メイテック	9050
9747	アサツー ディーケイ	9050
9766	コナミ	5250
9831	ヤマダ電機	6100
9843	ニトリ	6100
9875	マツモトキヨシ	6100
9962	ミスミグループ	6050
9984	ソフトバンク	6050
9987	スズケン	6050
9989	サンドラッグ	6100