

COMPARISON OF STOCK PRICE PREDICTION ACCURACY WITH VARIATION NUMBER OF HIDDEN LAYER CELL IN BACKPROPAGATION ALGORITHM

Doohan Kristiawan¹, Rosita Herawati²

^{1,2}Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Katolik Soegijapranata

¹doohanhahaha@gmail.com, ²rosita@unika.ac.id

Abstract

Nowadays there are various ways to invest money. The capital market becomes a tantalizing choice for potential investors. The profit from this investment is obtained from the difference between the selling and buying price of the shares. Created program is expected to provide an overview comparison of prediction accuracy from different hidden layer cells amount. Using historical data from yahoo finance to get stock prices in last 12 months. The data is processed using backpropagation algorithm and the process is done by the number of different hidden layer. Eventually the program will give the stock price predictions from four different number of hidden layer. Of all the proceeds will be matched to determine which one has the most precise accuracy.

Keywords: *backpropagation, neural network, artificial intelligence*

Pendahuluan

Banyak cara yang bisa dilakukan untuk berinvestasi. Kemajuan teknologi membuatnya menjadi lebih mudah. Salah satu investasi yang paling menarik adalah saham. Investasi saham tergantung pada pergerakan harga. Investor akan membeli saham tersebut saat harga rendah dan menjualnya ketika harganya tinggi. Namun ada banyak faktor yang mempengaruhi pergerakan harga saham seperti rumor dan sentimen pasar, kebijakan pemerintah, fluktuasi nilai tukar mata uang asing, dll. Pergerakan harga saham yang cenderung tidak stabil membuat banyak orang mencari solusi untuk menganalisa pergerakan. Padahal data harga saham yang bisa dianalisis itu tersedia di finance.yahoo.com. Dengan mengolah data di sana, siapapun bisa melakukan analisa.

Ada dua metode untuk menganalisa prediksi saham yaitu fundamental dan teknis. Analisis fundamental meliputi kinerja perusahaan, analisis industri, analisis persaingan usaha, analisis ekonomi dan makro mikro. Analisis fundamental adalah cara yang sulit karena butuh banyak informasi yang tidak bisa didapatkan oleh masyarakat. Karena analisa fundamental memang sulit, maka analisa teknikal menjadi solusi terbaik bagi investor. Analisis teknis meliputi moving average convergence divergence, indeks kekuatan relatif, stochastic oscillator dan lain-lain. Karena banyaknya data dan parameter yang akan diuji, analisa teknikal menjadi sangat sulit dilakukan secara manual. Salah satu cara untuk mengatasi masalah ini adalah dengan menggunakan jaringan syaraf tiruan.

Jaringan syaraf tiruan memiliki algoritma yang disebut backpropagation. Algoritma ini bisa memberikan prediksi berdasarkan data di masa lalu. Sangat cocok untuk prediksi

saham karena ada data historis yang bisa diolah. Algoritma backpropagation dapat belajar dari kesalahan dan memperbaiki nilai untuk memberikan hasil yang paling optimal.

Algoritma backpropagation membutuhkan data saham pada masa lampau untuk diproses terlebih dahulu. Data masa lampau tersebut tersedia di finance.yahoo.com bisa diambil dengan teknik webscraping. Webscraping adalah cara untuk mengambil data yang terdapat di situs web. Data yang diambil adalah harga saham selama 6 bulan terakhir. Kemudian akan dianalisis dengan 4 cara yang berbeda, dengan jumlah sel lapisan tersembunyi yang berbeda mulai dari 2, 4, 6 dan 8 sel. Lapisan sel tersembunyi adalah suatu tempat yang memproses bobot dari input sehingga menghasilkan suatu prediksi. Setiap susunan lapisan tersembunyi akan memberikan hasil prediksi dan hasilnya akan dibandingkan dengan harga realita. Hasil analisis ini dapat digunakan untuk memprediksi harga saham.

Landasan Teori

Investor saham memiliki masalah dalam menganalisa pergerakan harga saham dan memprediksi harga. Ada dua cara untuk menganalisa pergerakan harga saham dengan analisa fundamental dan analisa teknikal. Analisis fundamental membutuhkan banyak data yang tidak bisa diakses oleh masyarakat, sedangkan analisa teknikal sangat sulit untuk dihitung secara manual. Oleh karena itu, penulis akan membuat sebuah program yang bisa memprediksi harga saham pada masa lampau. Program ini akan mudah digunakan dan jauh lebih cepat daripada harus menghitung secara manual.

Program ini akan dibuat dengan menggunakan algoritma backpropagation yang merupakan sub bab dari jaringan syaraf tiruan. Algoritma backpropagation dapat memberikan prediksi berdasarkan data history. Data historis tersedia di finance.yahoo.com dan dapat diambil dengan menggunakan webscraping.

Langkah – langkah yang akan dilakukan:

1. Mengambil data historis dari finance.yahoo.com

Pada jurnal berjudul “Penerapan Metode Jaringan Syaraf Tiruan Backpropagation Untuk Meramalkan Harga Saham (IHSG)” karangan Andri Triyono dkk, menjelaskan bahwa data saham dapat diambil dari finance.yahoo.com [1]. Data saham yang akan diambil adalah TLKM, HMSB, BBCA, BBNI dan ASII. Data akan diambil mulai Januari 2016 sampai November 2017. Data dari Januari sampai Desember 2016 akan digunakan sebagai data pelatihan dan Januari - November 2017 akan digunakan sebagai data testing. Data akan disimpan dalam file teks.

2. Normalisasi data

Data yang sudah didapat akan dinormalisasi. Pada jurnal berjudul “Prediksi Harga Saham Menggunakan Jaringan Syaraf Tiruan Backpropagation” karangan Siti Amiroch, dijelaskan tujuan normalisasi adalah menyamarakan data yang berbeda besarnya agar menjadi setara dengan cara mengubah semua bilangan menjadi kisaran 0.1 sampai 0.8 [2].

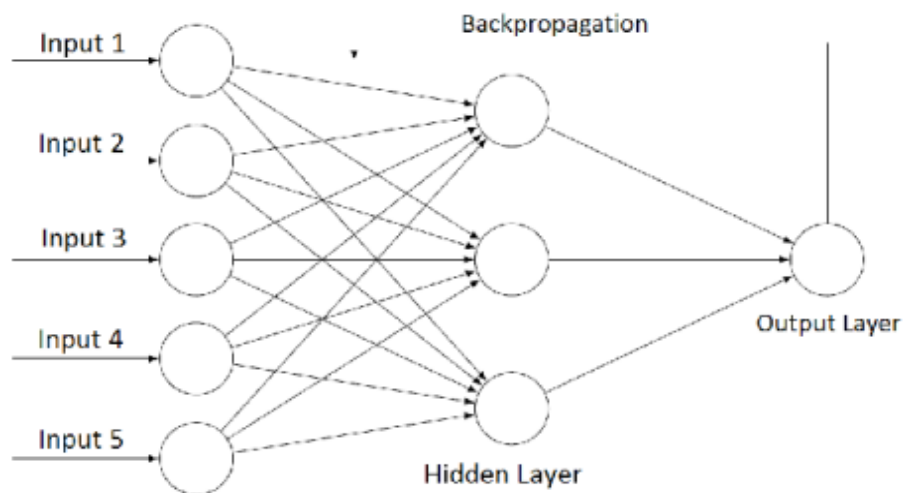
$$x' = 0.8 x \frac{(x - \text{MinValue})}{\text{MaxValue} - \text{MinValue}} - 0.1$$

x : data asli

MinValue : nilai terendah dari kumpulan data.

MaxValue : nilai tertinggi dari kumpulan data.

3. Proses Backpropagation



Gambar 1 : Proses Backpropagation

Pada buku yang berjudul “Kecerdasan Buatan” karangan T Sutojo dkk, dijabarkan langkah – langkah algoritma backpropagation untuk memperoleh suatu hasil prediksi [3].

Langkah – langkah tersebut yaitu :

Langkah 0: Inisialisasi semua bobot dengan bilangan acak kecil.

Bobot bias masukan (v_{0j}) = bilangan acak dari $-\beta$ dan β .

Masukan berat (V_{ij}) = bilangan acak $-0,5$ dan $0,5$.

Bobot bias hidden layer (W_{0k}) = bilangan acak dari -1 dan 1

Berat hidden layer (W_{jk}) = bilangan acak dari -1 dan 1 .

Langkah 2: Untuk setiap data training lakukan langkah 3-8.

Langkah 3 : Hitung semua outputs di hidden unit ($Z_j, j = 1, \dots, p$)

$$z_{in_j} = v_{0j} + \sum_{i=1}^n x_i v_{ij}$$

$$z_j = f(z_{in_j})$$

Langkah 4 : Hitung fungsi aktivasi pada hidden layer

$$Z_1 = \frac{1}{1 + e^{-z_{jn}}}$$

Langkah 5 : Operasi pada output layer.

$$y_{in_k} = w_{0k} + \sum_{j=1}^p z_j w_{jk}$$

$$y_k = f(y_{in_k})$$

Langkah 6 : Hitung fungsi aktivasi pada output layer

$$Y = \frac{1}{1 + e^{-y_{in}}}$$

Langkah 7 : Hitung δ pada output unit berdasarkan error pada tiap unit of output

$$\delta_k = (t_k - y_k) f'(y_{in_k})$$

Hitung selisih bobot hidden and bobot hidden menggunakan learning rate α

$$\Delta w_{jk} = \alpha \delta_k z_j$$

$$\Delta w_{0k} = \alpha \delta_k$$

Langkah 8 : Hitung δ pada hidden units berdasarkan error pada tiap hidden unit Z_j

$$\delta_{in_j} = \sum_{k=1}^m \delta_k w_{jk}$$

$$\delta_j = \delta_{in_j} f'(z_{in_j})$$

Hitung perbedaan bobot nilai input dan bobot input bias dengan learning rate α

$$\Delta v_{ij} = \alpha \delta_j x_i$$

$$\Delta v_{0j} = \alpha \delta_j$$

Langkah 9 : Hitung semua perubahan bobot

$$w_{jk}(\text{baru}) = w_{jk}(\text{lama}) + \Delta w_{jk}$$

$$v_{ij}(\text{baru}) = v_{ij}(\text{lama}) + \Delta v_{ij}$$

Langkah 10 : Tes kondisi perhentian berdasarkan square error

Langkah 11 : Ulangi langkah 3-10 sampai squared error < 0.01

Langkah 12 : Hitung dengan jumlah cell lain

4. Proses Pengujian

Hasil backpropagation akan dibandingkan dengan data historikal pada bulan Januari - November 2017. Pada jurnal berjudul "Prediksi Pergerakan Harga Saham Menggunakan Metode Backpropagation Neural Network" karangan Rio Bayu Afrianto, dijelaskan bahwa untuk mengecek keakuratan setiap prediksi maka hasil akan dibandingkan dengan hasil yang lain untuk menentukan jumlah sel lapisan tersembunyi mana yang bisa memberikan akurasi paling akurat [4].

Metodologi Penelitian

Langkah-langkahnya dijalankan satu per satu dan pada akhir penelitian akan didapatkan data perbandingan. Dari data perbandingan dapat disimpulkan dimana lapisan tersembunyi paling akurat untuk prediksi saham.

Langkah yang dilakukan:

1. Mempelajari jurnal dan buku

Jurnal dan buku dipelajari terlebih dahulu untuk mendukung penelitian. Ada 3 jurnal dan 1 buku yang diteliti. Dari berbagai jurnal peneliti akan belajar dan mencoba menerapkannya pada penelitian ini sambil memberikan sudut pandang yang berbeda sehingga bisa memberikan gambaran yang berbeda mengenai proses pada jaringan syaraf tiruan.

2. Menganalisa masalah

Analisis semua masalah yang akan dipecahkan oleh program. Hasil analisis akan digunakan untuk membuat desain proyek, sehingga implementasinya akan sesuai dengan tujuan utama dan tetap efisien untuk memenuhi tenggat waktu yang disepakati.

3. Pengambilan data dan proses coding

Dalam penelitian ini harga saham yang akan digunakan adalah PT Telekomunikasi Indonesia Tbk (TLKM), PT HM Sampoerna Tbk (HMSP), PT Bank Central Asia Tbk (BBCA), PT Bank Negara Indonesia Tbk (BBNI) dan PT Astra International Tbk (BBNI) ASII). Saham tersebut dipilih untuk masuk dalam kategori lima saham utama di Indonesia. Pasar saham buka mulai Senin sampai Jumat sehingga dalam 12 bulan mulai Januari - Desember 2016 akan diperoleh sebanyak 151 data dan total 5 perusahaan adalah 755 data. Dari 12 bulan data yang diperoleh akan digunakan untuk data training dan testing. Proses pengambilan data akan menggunakan metode webscraping menggunakan Java HTML parser (JSOUP). Setelah diperoleh data akan disimpan dalam file teks. Proses pengkodean akan menggunakan bahasa pemrograman java.

4. Laporan penelitian

Langkah ini melaporkan proses implementasi yang dapat memberikan prediksi harga saham. Prediksi harga akan dikumpulkan dan dibandingkan. Hal ini juga dilaporkan bahwa jumlah sel lapisan tersembunyi dapat memberikan akurasi yang paling akurat.

Hasil dan Pembahasan

Aplikasi ini dibuat dengan menggunakan bahasa pemrograman Java. Linked list sebagai struktur data digunakan untuk menyimpan data saham. Linked list dipilih karena stok data selalu tumbuh. Algoritma backpropagation digunakan oleh variasi jumlah sel lapisan tersembunyi yang berbeda. Variasi jumlah sel lapisan tersembunyi adalah 2, 4, 6 dan 8. Semua perbedaan lapisan tersembunyi akan menghasilkan prediksi harga yang berbeda. Beberapa percobaan akan dilakukan dengan menggunakan aplikasi ini. Hasil percobaan ini akan memberikan prediksi yang paling optimal dan akurat.

Aplikasi ini akan memproses data history stock yang diambil dari finance.yahoo.com. Data stok yang akan diambil adalah TLKM, HMSP, BBKA, BBNI dan ASII. Data akan diambil mulai Januari 2016 sampai November 2017. Data dari Januari sampai Desember 2016 akan digunakan sebagai data pelatihan dan Januari - November 2017 akan digunakan sebagai data testing. Data stok diambil dan disimpan ke dalam file teks.

Data yang telah disimpan akan dinormalisasi sebelum digunakan dalam proses perhitungan. Data harus dinormalisasi karena data volume puluhan juta namun harga terbuka, tinggi, rendah dan menyesuaikan data hanya dalam ribuan. Jadi normalisasi digunakan untuk menyamakan nilai data yang berbeda. Setelah data dinormalisasi maka tahap "feed forward" akan terus berlanjut. Tahap umpan maju akan menghasilkan prediksi sementara. Prediksi sementara dibandingkan dengan harga asli dan selisihnya akan dihitung. Perbedaan tersebut digunakan untuk proses backpropagation sehingga bobotnya akan disesuaikan. Kemudian dengan bobot yang disesuaikan, proses diulang sampai mencapai target zaman.

1. Mengambil Data Saham

Tahap pertama dari aplikasi ini adalah mengambil data saham dari finance.yahoo.com. Teknik pengambilan data dari internet disebut web scraping. Dalam bahasa pemrograman java ada sebuah perpustakaan bernama JSOUP yang bisa digunakan untuk web scraping. Sebelum mengambil data, aplikasi akan memeriksa data saham yang telah tersimpan dalam file teks dan melihat data terakhir yang dimilikinya. Kemudian lakukan web scrap untuk mengambil data baru yang belum tersimpan dalam file teks. Data yang lebih baru ditambahkan ke dalam file teks.

2. Normalisasi

Data dalam linked list akan masuk ke dalam proses normalisasi data. Normalisasi digunakan agar data yang berbeda menjadi lebih menyamakan kedudukan dan memiliki rentang 0.1 – 0.8. Normalisasi membutuhkan nilai minimum dan maksimum suatu data

yang harus dicari sebelum masuk dalam proses formula. Hasil dari proses ini adalah akan disimpan ke dalam linked list.

3. Feed Forward

Daftar linked yang dinormalisasi masuk ke dalam proses umpan maju. Pada awalnya ada beberapa bobot yang harus diinisiasi seperti masukan bobot ke lapisan tersembunyi, bobot bias ke lapisan tersembunyi, bobot lapisan tersembunyi ke lapisan keluaran dan bobot bias ke lapisan keluaran. Data berat secara acak diisi dengan kisaran -0,5 sampai 0,5. Bobot ini hanya sebagai angka awal saja, saat memasuki tahap Backpropagation maka bobotnya akan berubah sesuai perhitungan.

Tahap feed forward melakukan perhitungan operasi pada lapisan tersembunyi, fungsi aktivasi pada lapisan tersembunyi, operasi pada lapisan output dan fungsi aktivasi lapisan keluaran. Hasil prediksi akan muncul pada fungsi aktivasi lapisan output.

Proses pada tahap Feed Forward ini adalah menghitung input dengan bobot awal untuk menghasilkan prediksi harga sementara. Disebut prediksi sementara karena pada tahap selanjutnya prediksi ini akan diperbaiki lagi dengan hasil prediksi yang lebih baik. Kelima input berupa harga terbuka, tinggi, rendah, harga penyesuaian dan volume akan menghasilkan output yang harganya mendekati.

4. Menghitung Persentase Error

Setelah hasil prediksi muncul, hitunglah error rate dengan menggunakan rumus RMSE. Dari hasil RMSE yang bisa diketahui apakah datanya optimal atau tidak. Rumus untuk menghitung RMSE:

$$RMSE = \sqrt{\sum \frac{(y_{pred} - y_{ref})^2}{N}}$$

Ypred : Harga prediksi

Yref : Harga asli

N : Jumlah data

Setiap kali proses jaringan syaraf berjalan maka hasil yang diprediksi akan bertambah akurat antara harga asli dan kuadrat. Hasilnya akan terus ditambahkan sampai semua proses jaringan syaraf selesai dan hasilnya dibagi dengan jumlah data yang digunakan.

Perhitungan rumus RMSE adalah untuk mengetahui keakuratan prediksi. Semakin kecil nilai RMSE hasil prediksi yang lebih akurat dengan harga sebenarnya. Nilai RMSE dianggap kecil jika nilainya di bawah 0,01.

5. Proses Backpropagation

Tahap Backpropagation adalah memperbaiki nilai bobot agar lebih optimal dengan jaringan. Harga prediksi sementara dari tahap Feed Forward dibandingkan dengan harga

semula. Perbedaan harga untuk mengubah semua nilai bobot. Nilai bobot akan diperbarui sehingga bobot menjadi lebih akurat. Proses menyesuaikan bobot membuat algoritma bisa belajar dari kesalahan. Bila diprediksi harga selisih dengan harga asli masih terlalu besar, algoritma akan mengganti bobot sampai mencapai nilai yang akurat.

Setelah semua bobot diperbarui, proses feed forward dan backpropagasi diulang. Fase yang berulang disebut epoch. Oleh karena itu algoritma membutuhkan banyak waktu untuk mengulangi prosesnya. Setiap epoch berjalan, maka bobotnya akan lebih disesuaikan sehingga pada akhir epoch target bisa memberikan hasil prediksi yang akurat. Dalam penelitian ini digunakan target epoch sebanyak 10.000 kali.

6. Hitung Prediksi

Hitung prediksi yang dilakukan setelah proses feed forward dan proses backpropagation berjalan sebanyak 10.000 epoch. Untuk menghitung prediksi diperlukan input harga terbuka, tinggi, rendah, menyesuaikan harga dan volume tanggal yang harus diprediksi. Data masukan dihitung dengan menggunakan tahap Feed Forward dan menggunakan bobot terakhir yang diperbarui. Hasilnya adalah prediksi harga yang mendekati.

Hasil Pengujian

Penelitian ini menggunakan data saham dari PT Telekomunikasi Indonesia (TLKM), PT HM Sampoerna (HMSP), PT Bank Central Asia (BBCA), PT Bank Rakyat Indonesia (BBRI) dan PT Astra International (ASII). Data pelatihan yang digunakan selama 1 tahun mulai Januari 2016 sampai desember 2016 (berjumlah sekitar 252 data). Pengujian data yang digunakan selama bulan januari 2017 - november 2017 (sekitar 241 data). Menggunakan epoch 10.000 dan learning rate 1.

Tabel 1. Hasil Pengujian

Stock	Number of Hidden Layer Cell	Minimum	Maximum	Average
PT Telekomunikasi Indonesia	2	96.60%	100%	99.75%
	4	97.21%	100%	99.75%
	6	97.21%	100%	99.76%
	8	95.84%	100%	99.74%
PT HM Sampoerna	2	96.38%	100%	99.57%
	4	96.99%	100%	99.65%
	6	97.17%	100%	99.66%
	8	97.17%	100%	99.67%
PT Bank Central Asia	2	98.31%	100%	99.72%
	4	95.31%	100%	99.71%
	6	94.26%	100%	99.71%
	8	96.72%	100%	99.71%
PT Bank Negara Indonesia	2	95.45%	100%	99.46%
	4	95.57%	100%	99.53%
	6	95.62%	100%	99.54%
	8	95.65%	100%	99.54%
PT Astra International	2	97.65%	100%	99.80%
	4	98.23%	100%	99.80%
	6	98.23%	100%	99.79%
	8	98.22%	100%	99.78%

Kesimpulan

Penelitian ini bertujuan untuk mengetahui jumlah sel lapisan tersembunyi yang memberikan prediksi paling akurat. Komposisi jumlah sel lapisan tersembunyi adalah 2,

4, 6 dan 8 sel. Contoh kasus yang digunakan adalah memprediksi harga saham. Prediksi saham dilakukan dengan menggunakan data masa lalu

Hasil penelitian ini menemukan bahwa prediksi hasil semua jumlah sel memiliki akurasi di atas 99%. Setiap nomor sel pada lapisan tersembunyi memberikan hasil yang baik dalam menghitung prediksi harga saham.

Dalam penelitian ini disimpulkan bahwa variasi jumlah sel pada lapisan tersembunyi tidak mempengaruhi hasil prediksi. Perbedaan akurasi antara semua variasi jumlah sel pada hidden layer kurang dari 1% sehingga dianggap tidak berpengaruh.

Daftar Pustaka

- [1] A. Triyono, A. J. Santoso, and Pranowo -, “Penerapan Metode Jaringan Syaraf Tiruan Backpropagation Untuk Meramalkan Harga Saham (IHSG),” *Jurnal Sistem dan Informatika (JSI)*, vol. 11, no. 1, Art. no. 1, 2016.
- [2] S. Amiroch, “PREDIKSI HARGA SAHAM MENGGUNAKAN JARINGAN SYARAF TIRUAN BACKPROPAGATION,” *Unisda Journal of Mathematics and Computer Science (UJMC)*, vol. 1, no. 01, Art. no. 01, Jun. 2015.
- [3] T. Sutojo, *Kecerdasan buatan*. Andi Offset, 2011.
- [4] R. B. Afrianto, H. Tjandrasa, and I. Arieshanti, “PREDIKSI PERGERAKAN HARGA SAHAM MENGGUNAKAN METODE BACK PROPAGATION NEURAL NETWORK,” vol. 3, no. 3, p. 10, 2013.