

Convolutional Neural Network (CNN) Untuk Memprediksi Harga Saham

Mazi Prima Reza (10116038) dan Urfi NS Yogabama (10116105)

Institut Teknologi Bandung

Jl. Ganesha No. 10, Bandung, Jawa Barat

Email: maziprima@students.itb.ac.id, urfi@students.itb.ac.id

Abstrak- CNN dikenal sebagai pra-proses klasifikasi gambar dengan tujuan untuk memudahkan proses perhitungan di neural networks. Namun kali ini akan dicoba memprediksi harga saham menggunakan CNN dengan cara mengubah dimensi data sahamnya terlebih dahulu. Hasilnya didapat nilai prediksi yang cukup baik menggambarkan pergerakan harga saham.

Kata kunci: *pembelajaran mendalam, convolutional neural networks, saham, prediksi*

I. Pendahuluan

Untuk mengembangkan model dari data keuangan, dibutuhkan dua pendekatan yaitu analisis fundamental dan analisis teknis. Analisis fundamental merupakan suatu teknik analisis yang memperhitungkan kinerja perusahaan untuk mengetahui apakah perusahaan tersebut dalam kondisi baik atau tidak. Data analisis fundamental berupa aset, pendapatan, dan lain-lain yang biasanya disediakan oleh perusahaan sekuritas. Untuk analisis teknis, terdapat banyak sekali metode. Pada tugas kali ini, akan digunakan *Technical Indicators* dan *Convolutional Neural Networks* untuk melakukan analisis teknis prediksi pergerakan harga saham MFST.

a. Technical Indicators

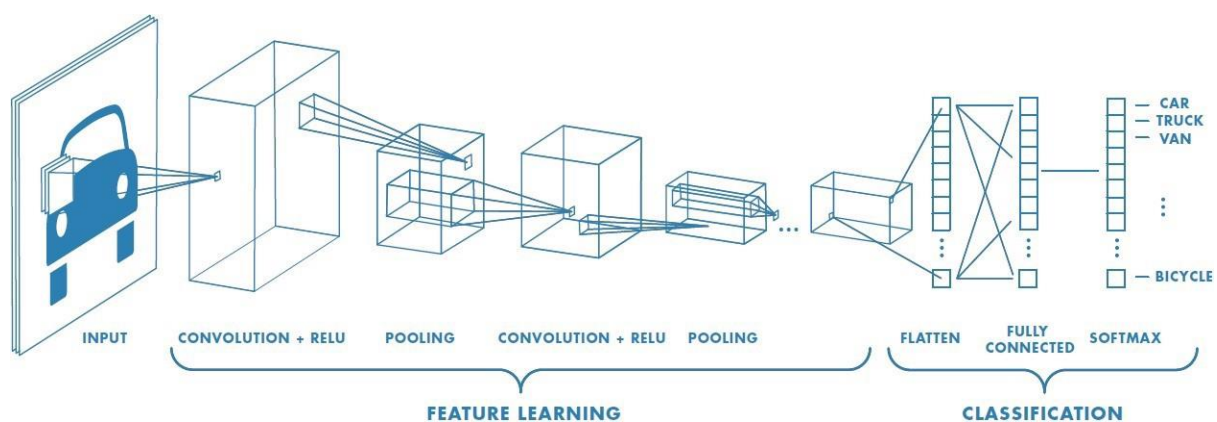
Technical indicators adalah rumus matematis yang digunakan dalam membantu mengambil keputusan dalam pembelian saham. Indikator biasanya memberikan informasi mengenai tren, yaitu pergerakan harga saham pada periode tertentu. Pada tugas ini, digunakan 3 jenis indikator, yaitu *simple moving average*, *relative strength index*, dan *stochastic oscillator*.

b. Convolutional Neural Networks (CNN)

Saat ini teknologi terus berkembang cepat begitu juga dengan perkembangan piksel gambar dan kualitas kamera. Jika pada tahun 2010 ukuran foto yang bisa ditangkap oleh ponsel pada umumnya masih dalam kualitas VGA atau memiliki ukuran piksel 640×480 , maka saat ini ponsel pada umumnya sudah dapat menangkap gambar dengan ukuran lebih dari 16 MP. Bayangkan jika harus melakukan klasifikasi gambar dengan ukuran matriks sebesar itu, maka

akan butuh komputer spesifikasi tinggi dan waktu *training* yang entah sampai kapan akan selesai. Oleh karena itu CNN dikembangkan sebagai alat yang dapat memudahkan klasifikasi gambar dengan ukuran piksel yang tinggi.

Convolutional Neural Networks atau CNN adalah sebuah algoritma pembelajaran mendalam, digunakan pada tahap pra-proses klasifikasi gambar. Secara umum CNN memiliki 2 layer: *convolutional layers* dan *sampling layers*, pada kasus ini *Max-pooling*. Di *convolutinal layer*, operasi konvolusi diterapkan, hasilnya dibawa ke layer berikutnya. Di subsampling layer banyaknya parameter dan ukuran spasial dikurangi. Di *subsampling* layer terakhir, data menjadi vektor satu dimensi yang siap diolah *neural networks*.



Gambar 1. Ilustrasi Proses CNN

Sumber gambar: <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>

II. Metodologi

Pada tugas ini digunakan bahasa pemrograman *python* dan diproses menggunakan *Google Colaboratory* sebuah fitur dari *google* sehingga pengguna dapat ‘meminjam’ GPU milik Google dengan harapan program dapat berjalan lebih cepat. Program ini juga dapat dijalankan di *jupyter notebook*.

1. Technical Indicators

a. Simple Moving Average (SMA)

SMA menggambarkan pergerakan sederhana saham. SMA biasa digunakan untuk melihat tren dan kemungkinan perubahan tren.

Nilai SMA didapat dengan menggunakan persamaan berikut.

$$SMA = \frac{A_1 + A_2 + \dots + A_n}{n}$$

Dengan A_n adalah harga saham pada periode ke- n dan n adalah banyaknya periode.

b. *Relative Strength Indeks (RSI)*

RSI menggambarkan apakah suatu saham jenuh jual (*over sold*) atau jenuh beli (*over bought*). RSI juga menggambarkan tren saham. Jika grafik berada di atas 50, maka saham kemungkinan sedang *uptrend* dan sebaliknya jika di bawah 50 maka kemungkinan saham sedang *downtrend*.

Nilai RSI dihasilkan dari persamaan

$$RSI = 100 - \frac{100}{1 + RS}$$

Dengan RS adalah rata-rata kenaikan harga dalam satu periode dikurangi periode saat harga naik dalam interval waktu tertentu.

c. *Stochastic Oscillator*

Stochastic Oscillator adalah indikator momentum yang membandingkan kisaran selama dengan harga penutupan selama periode waktu tertentu.

Persamaan indikator stokastik adalah

$$\%K = \left(\frac{C - L14}{H14 - L14} \right) \times 100$$

Dengan C adalah harga saham saat tutup, L14 adalah harga terendah 14 hari sebelumnya, H14 harga tertinggi 14 hari sebelumnya, dan %K adalah nilai stokastik.

Sedangkan %D adalah nilai stokastik dalam periode 3 hari.

2. *Persiapan Data*

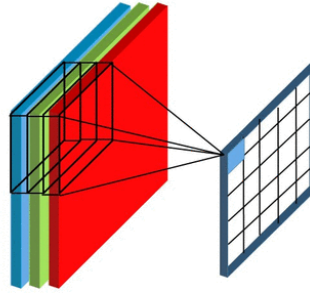
Inti dari tahap ini adalah mengubah data harga saham ke dalam matriks *time step* dengan nilai step 30 untuk tiap variabelnya dan mengubahnya ke dalam bentuk matriks 4 dimensi. Pada tahap ini dimensi data menjadi $\mathcal{M}^{(Nobservasi-step-1 \times step \times 1 \times 5)}$. Lalu bentuk data diubah menjadi data yang lebih dikenali CNN, yaitu $\mathcal{M}^{(Nobservasi \times RGB \times panjang \times lebar)}$. Jadi dimensi data menjadi $\mathcal{M}^{(4265 \times 1 \times 30 \times 5)}$. Data dianggap seolah-olah gambar hitam putih dengan panjang 30 piksel dan lebar 5 piksel sehingga ukuran data menjadi.

3. *Modelling*

Pada tahap CNN kami menggunakan satu layer *convolutional* dan satu layer *max-pooling*, jika lebih, dari yang sudah kami coba, maka model akan *overfitting*.

i. *Convolutional*

Convolutional merupakan layer dari CNN. Matriks filter yang digunakan berukuran 1×4 akan bergerak pada setiap baris di matriks hasil *reshape* ukuran 30×5 . Agar semua nilai matriks hasil *reshape* berhasil dilewati oleh matriks filter, maka setiap baris dilewati sebanyak dua kali. Setiap pergerakan filter memunculkan sebuah *cell* pada matriks baru.



Gambar 2. Ilustrasi convolution

ii. *Max pooling*

Ini adalah tahap kedua dari CNN. *Max pooling* bertujuan mengambil data terbesar dari sekelompok data yang berdekatan untuk memperkecil jumlah matriks data. Digunakan ukuran matriks 1×2 untuk membagi data untuk diambil nilai terbesarnya. Artinya, diambil data terbesar dari 30 baris pada masing-masing kolom.

iii. Flatten

Pada tahap ini matriks yang telah didapat pada layer CNN diubah bentuknya menjadi vektor.

iv. Dense

Dense merupakan hidden layer biasa seperti di neural networks. Layer ini memiliki matriks ‘weights’, bias, dan fungsi aktivasi. Dense layer di sini bertugas untuk melakukan klasifikasi gambar. Secara matematis dense menerapkan rotasi, translasi, dan scaling pada data.

Dense pada kasus ini dilakukan 2 kali. Yang pertama dengan 15 hidden layers dan yang kedua dengan 1 hidden layers. Hal ini bertujuan untuk mengubah dimensi data supaya data bisa diolah di neural networks.

4. Training dan prediksi data

Tahap terakhir adalah melakukan training dengan 85% data dan prediksi dengan 15% data. Pada tahap ini digunakan mean-squared error sebagai loss function dan ADAM optimizer sebagai metode optimasi.

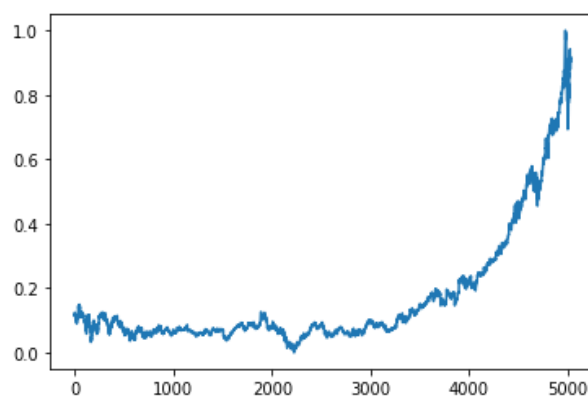
III. Hasil dan Diskusi

Pada program ini digunakan data sebanyak 5033 yang merupakan data harian selama 20 tahun terakhir yang didapat dari situs Yahoo Finance. 5033 observasi tadi dibagi sebanyak 85% untuk *training* dan 15% untuk *testing*.

Tabel 1. Sebagian Data Harga Saham Microsoft (MSFT)

Tanggal	Open	High	Low	Close	Volume
28/4/ 2020	35.375000	35.500000	34.125000	34.875000	78082600
1/5/2020	36.437500	37.000000	35.843750	36.718750	107811000
.
.
.
24/4/2000	172.059998	174.559998	170.710007	174.550003	34305300
27/4/2020	176.589996	176.899994	173.300003	174.050003	33158800

Jika data pada tabel 1 diplot, maka akan membentuk grafik pergerakan harga saham. Grafik di bawah ini adalah pergerakan harga saham saat penutupan.



Gambar 3. Pergerakan harga saham pada saat penutupan

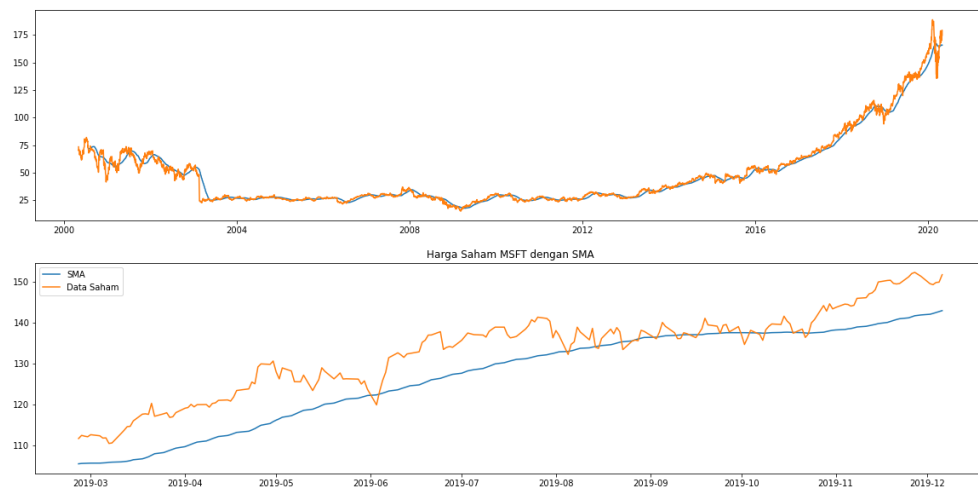
Dengan statistika deskriptifnya

Tabel 2. Statistika Deskriptif dari Data Harga Saham MSFT Harian dalam 20 Tahun

	Open	High	Low	Close	Volume
Count	5030				
Mean	44.114223	44.562321	43.660491	44.122416	54766470
Std	32.003050	32.296502	31.667534	32.010150	31045670
Min	15.200000	15.620000	14.870000	15.150000	7425600
25%	26.549999	26.841250	26.250000	26.526250	32861720
Median	29.709999	30.000000	29.42	29.724999	50197950
75%	46.330002	46.822501	45.985002	46.439999	68283480
Max	190.649994	190.699997	186.470001	188.699997	591052200

a. *Technical Indicator*

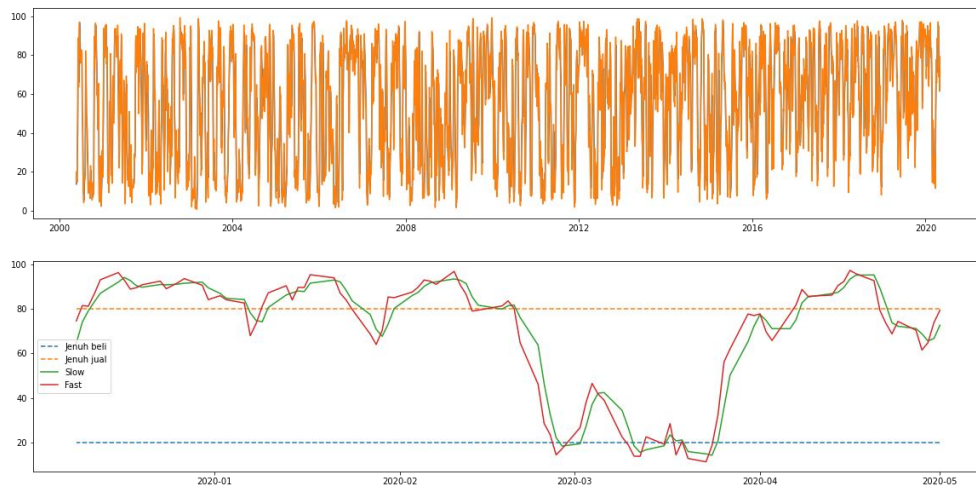
i. *Simple Moving Average(SMA)*



Gambar 4. Grafik Indikator SMA dengan time series dari MSFT

Pada gambar di atas, terdapat dua grafik. Grafik yang pertama adalah pergerakan SMA dengan harga saham MSFT dalam rentang waktu 20 tahun. Sedangkan grafik yang kedua dalam rentang waktu 200 hari supaya pergerakan dapat terlihat lebih jelas. SMA hanya menggambarkan pergerakan sederhana dari harga saham dan juga memprediksi kemungkinan terjadinya tren. Dari grafik kedua, dapat dikatakan bahwa saham MSFT pada rentang tersebut sedang *uptrend*.

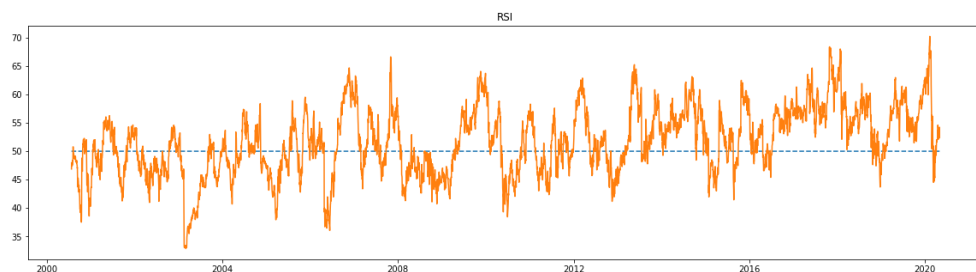
ii. *Stochastic Oscillation*



Gambar 5. Indikator *Stochastic Oscillation* dalam 20 tahun (atas) dan dalam 100 hari terakhir (bawah)

Pada grafik di bawah dapat dilihat bahwa grafik yang berada di atas nilai 80 menandakan saham sedang mengalami jenuh beli pada waktu tersebut dan grafik yang berada di bawah nilai 20 mengalami jenuh jual. Perhatikan juga bahwa pertemuan kedua grafik menandakan adanya kemungkinan perubahan tren.

iii. *Relative Strength Index (RSI)*



Gambar 6. Grafik Indikator RSI

Pada grafik menunjukkan bahwa pergerakan di bawah garis biru putus-putus menandakan bahwa saham sedang *downtrend* dan sebaliknya jika grafik berada di atas, maka saham sedang *uptrend*.

b. Pemodelan

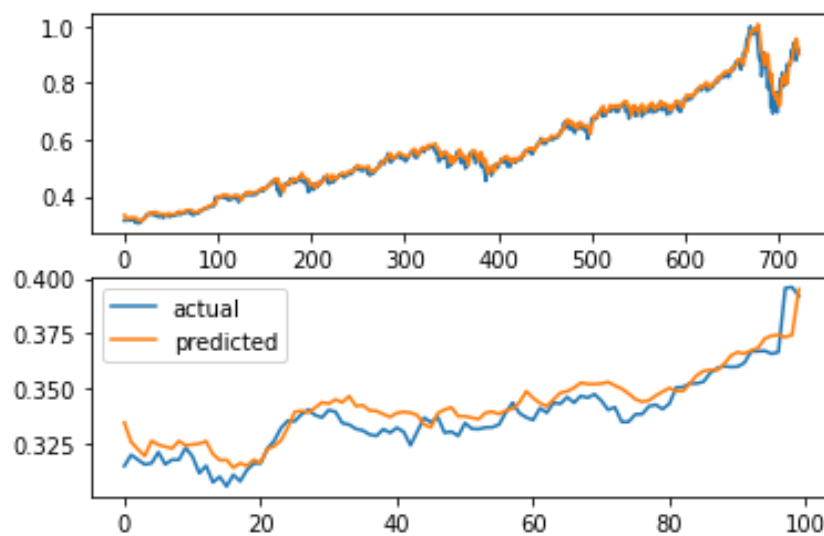
Tabel 3. Urutan Pemodelan dengan ukuran *outputnya*.

Tipe Layer	Ukuran Output
Convolution	(None, 1, 27, 10)

Max Pooling	(None, 1, 13, 10)
Flatten	(None,130)
Dense	(None,10)
Dense	(None,1)

c. *Training* dan Prediksi

Sekilas dilihat dari grafik, prediksi menunjukkan kemiripan dengan kondisi aktual. Hal tersebut didukung dengan nilai RMSE yaitu 0.018357185988200443, artinya selisih antara kedua grafik cenderung menuju nol.



Gambar 7. Grafik harga saham sebenarnya dengan prediksi

IV. Kesimpulan

Didapat kesimpulan sebagai berikut.

CNN biasa digunakan untuk mengolah data gambar dengan melakukan *input* data piksel pada gambar. Dalam pembelajaran ini, digunakan CNN untuk melakukan prediksi harga saham dengan mengubah data saham menjadi data matriks seolah-olah data tersebut adalah data gambar dalam piksel.

V. Kontribusi

Mazi Prima Reza (10116038)	Berdiskusi mengenai CNN, merumuskan technical indicator, membangun model, membuat laporan, mengedit video,
----------------------------	--

	membuat konten yang ada di video melalui power point.
Urfi NS Yogabama (10116105)	Berdiskusi mengenai CNN, membangun model, membuat laporan, menulis <i>script</i> video, mengedit video, membuat konten yang ada di video melalui power point.

VI. Referensi

- [1] Sezer, Omer Berat. *Financial Trading Model with Stock Bar Chart Image Time Series with Deep Convolutional Neural Networks*. arXiv.org. <https://arxiv.org/abs/1903.04610>
- [2] Saha, Sumit. 2018. *A Comprehensive Guide to Convolutional Neural Networks. Towards Data Science*. Diakses melalui <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53> tanggal 1 Mei 2020.
- [3] Yahoo Finance. Diakses melalui <https://finance.yahoo.com/quote/MSFT/history?p=MSFT> diakses tanggal 28 April 2020.