

**PERAMALAN HARGA SAHAM INDONESIA MENGGUNAKAN
METODE *LONG SHORT TERM MEMORY* (LSTM)**

SKRIPSI



**UIN SUNAN AMPEL
S U R A B A Y A**

DISUSUN OLEH :

INDAH KUSUMAWATI

H06219003

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL
SURABAYA**

2023

PERNYATAAN KEASLIAN

Saya yang bertanda tangan di bawah ini,

Nama : Indah Kusumawati

NIM : H06219003

Program Studi : Sistem Informasi

Angkatan : 2019

Menyatakan bahwa saya tidak melakukan plagiat dalam penulisan skripsi saya yang berjudul: “PERAMALAN HARGA SAHAM INDONESIA MENGGUNAKAN METODE LONG SHORT TERM MEMORY (LSTM)”. Apabila suatu saat nanti terbukti saya melakukan tindakan plagiat, maka saya bersedia menerima sanksi yang telah ditetapkan.

Demikian pernyataan keaslian ini saya buat dengan sebenar-benarnya.

Surabaya, 27 Juni 2023

Yang menyatakan,




Indah Kusumawati

NIM H06219003

LEMBAR PERSETUJUAN PEMBIMBING

Skripsi oleh

NAMA : INDAH KUSUMAWATI
NIM : H06219003
JUDUL : PERAMALAN HARGA SAHAM INDONESIA
MENGGUNAKAN METODE LONG SHORT TERM
MEMORY (LSTM)

Ini telah diperiksa dan disetujui untuk diujikan.

Surabaya, 27 Juni 2023

Dosen Pembimbing 1



Dwi Rolliawati, M.T

NIP. 197909272014032001

Dosen Pembimbing 2



Khafid, M.Kom


NIP. 197906092014031002

PENGESAHAN TIM PENGUJI SKRIPSI

Skripsi Indah Kusumawati ini telah dipertahankan
di depan tim penguji skripsi
Surabaya, 13 Juli 2023.

Mengesahkan,
Dewan Penguji

Dosen Penguji 1



Ahmad Yusuf, M.Kom

NIP. 199001202014031003

Dosen Penguji 2



Bayu Adhi Nugroho, Ph.D

NIP. 197905182014031001

Dosen Penguji 3



Dwi Rolliawati, M.T

NIP. 197909272014032001

Dosen Penguji 4



Khalid, M.Kom

NIP. 197906092014031002

Mengetahui,

Dekan Fakultas Sains dan Teknologi
UIN Sunan Ampel Surabaya




Saepul Hamdani, M.Pd.

NIP. 196507312000031002



KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL SURABAYA
PERPUSTAKAAN

Jl. Jend. A. Yani 117 Surabaya 60237 Telp. 031-8431972 Fax.031-8413300
E-Mail: perpustakaan@uinsby.ac.id

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademika UIN Sunan Ampel Surabaya, yang bertanda tangan di bawah ini, saya:

Nama : Indah Kusumawati
NIM : H06219003
Fakultas/Jurusan : Sains dan Teknologi / Sistem Informasi
E-mail address : h06219003@uinsby.ac.id

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Perpustakaan UIN Sunan Ampel Surabaya, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah :

Sekripsi Tesis Desertasi Lain-lain (.....)

yang berjudul : PERAMALAN HARGA SAHAM INDONESIA MENGGUNAKAN

METODE LONG SHORT TERM MEMORY (LSTM)

beserta perangkat yang diperlukan (bila ada). Dengan Hak Bebas Royalti Non-Eksklusif ini Perpustakaan UIN Sunan Ampel Surabaya berhak menyimpan, mengalih-media/format-kan, mengelolanya dalam bentuk pangkalan data (database), mendistribusikannya, dan menampilkan/mempublikasikannya di Internet atau media lain secara *fulltext* untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan UIN Sunan Ampel Surabaya, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini yang saya buat dengan sebenarnya.

Surabaya, 17 Juli 2023

Penulis

(INDAH KUSUMAWATI)

nama terang dan tanda tangan

ABSTRAK

Peramalan harga saham indonesia menggunakan metode *long short term memory (lstm)*

Oleh :

Indah Kusumawati

Minat masyarakat terhadap saham mengalami pertumbuhan eksponensial karena return of investment yang cepat dan keuntungan yang besar. Adanya peramalan saham, dapat membantu investor saham untuk bertransaksi di waktu yang tepat. Peramalan saham menjadi menantang dikarenakan karakteristik data yang berjalan acak, tidak linier, dan *volatile*. Terdapat berbagai metode untuk melakukan peramalan saham. Penelitian ini, menggunakan metode LSTM untuk meramalkan harga saham di hari berikutnya. Metode LSTM dipilih karena dapat melakukan pembelajaran afiliasi jangka panjang pada data *sequence* dengan kompleksitas tinggi. Teknik dropout ditambahkan untuk mencegah adanya *overfitting* dan meningkatkan performa LSTM. Data yang digunakan adalah data saham dari AALI(Astra Agro Lestari Tbk.). Hasilnya adalah metode LSTM dengan *dropout* dapat memberikan akurasi peramalan hingga 98%. Namun, diperlukan tuning parameter yang tepat agar penggunaan dropout dapat menghasilkan akurasi yang cukup tinggi.

Kata kunci : Peramalan, Saham, LSTM, *Dropout*, *Tuning* Parameter

ABSTRACT

FORECASTING STOCK PRICE INDONESIA USING *LONG SHORT TERM MEMORY (LSTM)* METHOD

By :
Indah Kusumawati

Public interest in stocks has exponential growth because of the fast return on investment and large profits. The existence of stock forecasting, can help stock investors to transact at the right time. Stock forecasting becomes challenging due to the random, non-linear, and volatile of characteristic the data. There are various methods for forecasting stocks. This study uses the LSTM method to forecast stock prices in next day. The LSTM method was chosen because it can carry out long-term affiliate learning on sequence data with high complexity. Added dropout technique to prevent overfitting and improve LSTM performance. The data used is stock data from AALI (Astra Agro Lestari Tbk.).The result is the LSTM method with dropouts can provide forecasting accuracy of up to 98%. However, proper parameter tuning is required so that the use of dropouts can produce high accuracy.

Keywords : Forecasting, Stock, LSTM, Dropout, Parameter Tuning

DAFTAR ISI

PERNYATAAN KEASLIAN.....	i
LEMBAR PERSETUJUAN PEMBIMBING	ii
PENGESAHAN TIM PENGUJI SKRIPSI.....	iii
PERNYATAAN PERSETUJUAN PUBLIKASI	iv
ABSTRAK	v
ABSTRACT	vi
DAFTAR ISI	vii
DAFTAR GAMBAR	ix
DAFTAR TABEL.....	x
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah.....	3
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	4
BAB II KAJIAN PUSTAKA	5
2.1 Tinjauan Penelitian Terdahulu.....	5
2.2 Dasar Teori.....	5
2.2.1 Saham.....	5
2.2.2 <i>Time series Forecasting</i>	7
2.2.2 <i>Windowing</i>	8
2.2.4 <i>Deep learning</i>	9
2.2.4 <i>Long short term memory (LSTM)</i>	13
2.2.5 Arsitektur LSTM.....	23
2.2.6 Performance Measurement	25
2.3 Integrasi Keilmuan	26
BAB III METODE PENELITIAN.....	28
3.1 Desain Penelitian	28
3.1.1 Studi pustaka.....	29
3.1.2 Pengambilan Data	29
3.1.3 Windowing.....	29

DAFTAR GAMBAR

Gambar 2. 1 Proses sliding window.....	9
Gambar 2. 2 Arsitektur Deep learning	10
Gambar 2. 3 Contoh alur kerja deep learning	11
Gambar 2. 4 Taksonomi teknik deep learning.....	13
Gambar 2. 5 Arsitektur LSTM vanilla	14
Gambar 2. 6 LSTM tanpa forget gate	17
Gambar 2. 7 LSTM dengan forget gate	18
Gambar 2. 8 LSTM dengan peephole connection.....	19
Gambar 2. 9 Overfitting dan underfitting	23
Gambar 2. 10 Arsitektur LSTM	24
Gambar 3. 1 Alur metode penelitian	28
Gambar 3. 2 Alur proses build model	30
Gambar 4. 1 Tampilan grafik dataset saham harian dari AALI	32
Gambar 4. 2 Gambar pola dari 2 bulan terakhir data.....	34
Gambar 4. 3 Graphic loss activation function tanh dan sigmoid	38
Gambar 4. 4 Graphic loss dengan dropout rate 0.01	39
Gambar 4. 5 Graphic loss dengan dropout rate 0.1	39
Gambar 4. 6 Graphic loss dengan dropout rate 0.3.....	40
Gambar 4. 7 Hasil prediksi dari parameter terbaik model LSTM tanpa dropout.....	45
Gambar 4. 8 Hasil prediksi dari parameter terbaik model LSTM tanpa dropout.....	47
Gambar 4. 9 Hasil prediksi 4 hari berikutnya	51
Gambar 4. 10 (a) LSTM tanpa dropout (b) LSTM dengan dropout	51
Gambar 4. 11 (a) LSTM tanpa dropout (b) LSTM dengan dropout.....	52

DAFTAR TABEL

Tabel 2. 2 Parameter relevan.....	24
Tabel 4. 1 Tampilan dataset saham harian dari AALI	32
Tabel 4. 2 Penampilan data setelah dilakukan windowing	33
Tabel 4. 3 Tabel perbandingan hasil berdasarkan ukuran window size	35
Tabel 4. 4 Shape data setelah dilakukan windowing.....	35
Tabel 4. 5 Jumlah data setelah di split	36
Tabel 4. 6 Shape data setelah di split	36
Tabel 4. 7 Bentuk masing-masing data setelah di reshape.....	36
Tabel 4. 8 Performa beberapa epoch.....	41
Tabel 4. 9 Hasil R^2 model LSTM tanpa dropout.....	43
Tabel 4. 10 Hasil RMSE model LSTM tanpa dropout.....	43
Tabel 4. 11 Gap loss dari parameter terkecil.....	44
Tabel 4. 12 Hasil R^2 model LSTM dengan dropout	45
Tabel 4. 13 Hasil RMSE model LSTM dengan dropout.....	46
Tabel 4. 14 Perbandingan performa parameter terbaik antara model LSTM dengan dropout dan tanpa dropout	47
Tabel 4. 15 Perbandingan antar model dengan learning rate 0.1	48
Tabel 4. 16 Perbandingan antar model dengan learning rate 0.01	48
Tabel 4. 17 Perbandingan antar model dengan learning rate 0.001	49
Tabel 4. 18 Rata-rata performa antara model LSTM dengan dropout dan tanpa dropout	50

BAB I

PENDAHULUAN

1.1 Latar Belakang

Terdapat berbagai cara untuk menambah penghasilan, salah satunya adalah dengan investasi saham. Selama beberapa dekade terakhir, minat masyarakat terhadap saham mengalami pertumbuhan eksponensial (Kumbure dkk. 2022). Investasi saham yang memiliki *return of investment* cepat (Alkhatib dkk. 2022) dan keuntungan besar menjadikan saham sebagai aset bernilai miliaran dolar yang diperdagangkan di bursa saham setiap hari (Hoseinzadeh dan Haratizadeh 2019). Namun, investasi saham juga memiliki resiko yang tinggi. Dengan adanya peramalan saham, dapat membantu investor saham untuk menjual atau membeli saham di waktu yang tepat. Sehingga akan didapatkan keuntungan yang maksimal dan meminimalisir kerugian.

Peramalan saham menjadi salah satu prediksi yang menantang (Khan dkk. 2020) dikarenakan data saham memiliki karakteristik data yang berjalan acak (Lu dkk. 2020), tidak linier, *volatile* (Wang dkk. 2021) dan sulit diprediksi (Lee 2021). Almasarweh dan Irawan melakukan penelitian peramalan saham menggunakan metode tradisional ARIMA (Almasarweh dan Wadi 2018; Irawan, t.t.). Hasilnya adalah metode ARIMA dapat meramalkan harga saham dengan baik untuk peramalan jangka pendek, tetapi akurasi metode tradisional tidak cukup baik untuk peramalan jangka panjang (Lu dkk. 2020) sehingga diperlukan metode lain untuk dapat menemukan konsistensi akurasi dalam peramalan saham seperti *machine learning*, *deep learning*, dan sebagainya.

Riset terkait teknik *machine learning* (ML) yang digunakan untuk meramalkan harga saham, seperti Khan yang menggunakan 11 teknik ML (*gaussian Naive Bayes (GNB)*, *multinomial Naive Bayes (MNB)*, *Support Vector Machine (SVM)*, *Linear Regression*, *K-Nearest Network (KNN)*, *classification and regression tree (CART)*, *linear discriminant analysis (LDA)*, *AB*, *Gradient Boost Machine*, *Random Forest*, dan *ET*) (Khan dkk. 2022). Dengan akurasi tertinggi sebesar 75.16%. Abraham menggunakan *genetic algorithm (GA)* sebagai *feature selection* dan *random forest (RF)* sebagai klasifikasi tren saham (Abraham dkk. 2022).

Hasilnya adalah metode yang dipakai memberikan akurasi tertinggi sebesar 80% ketika diujikan di beberapa dataset. Beberapa metode ML yang digunakan memiliki performa yang baik, namun metode ML masih kurang cukup baik untuk memprediksi data dengan pola nonlinier yang kompleks (Hu, Zhao, dan Khushi 2021) sehingga teknik *deep learning* diusulkan untuk mengatasi masalah tersebut.

Teknik *deep learning* (DL) memberikan hasil yang sangat baik di berbagai bidang, sehingga memiliki potensi untuk diterapkan pada peramalan saham dikarenakan kemampuan dalam mendeteksi dinamika pergerakan data (Alkhatib dkk. 2022). Performa dari teknik DL lebih tinggi dibandingkan dengan teknik ML, seperti penelitian oleh (Nikou, Mansourfar, dan Bagherzadeh 2019) yang membandingkan metode LSTM dengan 3 teknik ML (SVR, NN, dan RF) untuk meramalkan harga saham di hari berikutnya. Hasilnya adalah metode LSTM memiliki nilai *error* terkecil dibandingkan 3 teknik ML yang diuji. Metode LSTM dipilih karena memiliki performa lebih baik untuk data *sequence* dengan kompleksitas tinggi dibandingkan GRU (*Gated recurrent unit*) (Cahuantzi, Chen, dan Güttel 2023). LSTM juga memungkinkan pembelajaran afiliasi jangka panjang yang tidak mungkin dilakukan oleh RNN (Nikou, Mansourfar, dan Bagherzadeh 2019).

Penelitian lain (Bhandari dkk. 2022) menggunakan LSTM untuk memprediksi harga saham S&P 500 dan menghasilkan nilai akurasi sebesar 0.9976. Penelitian lain oleh (Ding dan Qin 2020) juga menggunakan LSTM untuk memprediksi harga saham namun dengan data yang berbeda yakni menggunakan data dari Shanghai composite index dan PetroChina. Hasilnya adalah model memberikan nilai akurasi sebesar 0.7879.

Berdasarkan penelitian sebelumnya, metode yang sama diterapkan pada dataset yang berbeda akan menghasilkan nilai akurasi yang berbeda pula. Setiap dataset akan memberikan informasi berbeda yang akan mempengaruhi kemampuan performa model LSTM sehingga diperlukan *tuning hyperparameter*. Meskipun *tuning parameter* sudah dilakukan, terdapat masalah lain yang dapat timbul salah satunya *overfitting*. Beberapa teknik digunakan untuk mencegah adanya *overfitting* seperti *dropout* dan *batch normalization* (Alzubaidi dkk. 2021). Efek dari *batch*

normalization tergantung pada ukuran mini-batch dan tidak terlalu jelas bagaimana penerapannya pada RNN (Ba, Kiros, dan Hinton 2016) sehingga tidak disarankan untuk digunakan dengan LSTM. Maka dari itu penelitian ini akan menggunakan teknik *dropout* sebagai pencegahan *overfitting*. Penggunaan *dropout* dapat meningkatkan performa model LSTM seperti penelitian oleh (Ding dan Qin 2020) yang menambahkan *dropout* pada model LSTM dan berhasil meningkatkan nilai akurasi dari 0.7879 menjadi 0.9737.

Berdasarkan permasalahan di atas, penelitian mengajukan penggunaan metode *Long short term memory* (LSTM) dengan teknik *dropout* untuk mencegah *overfitting* dan untuk menghasilkan akurasi yang cukup tinggi. Penggunaan model ini diharapkan dapat meningkatkan performa metode LSTM ketika diterapkan dengan dataset yang berbeda

1.2 Perumusan Masalah

Berdasarkan uraian dari latar belakang yang telah dijelaskan diatas maka rumusan masalah dari penelitian ini antara lain:

1. Bagaimana implementasi model LSTM dengan *dropout* dalam meramalkan saham Indonesia?
2. Bagaimana performa model LSTM dengan *dropout* dalam meramalkan saham Indonesia?

1.3 Batasan Masalah

Pembahasan permasalahan dalam penelitian ini terbatas pada :

1. Penelitian ini berfokus pada peramalan harga saham
2. Data yang digunakan adalah data saham AALI (Astra Agro Lestari Tbk.) yang diperoleh dari situs kaggle melalui link <https://www.kaggle.com/datasets/muamkh/ihsgstockdata>
3. Variabel saham yang akan digunakan dalam peramalan hanya variabel close
4. Jenis LSTM yang digunakan terbatas pada *vanilla* LSTM (*Single layer* LSTM)
5. Penggunaan teknik pencegahan *overfitting* terbatas pada teknik *dropout*
6. Tuning parameter terbatas pada *hidden node*, *batch size*, dan *learning rate*

1.4 Tujuan Penelitian

Adapun tujuan dari penelitian ini antara lain:

1. Untuk menjelaskan implementasi metode LSTM dengan *dropout* dalam meramalkan harga saham
2. Untuk mengetahui performa metode LSTM dengan *dropout* dalam meramalkan harga saham

1.5 Manfaat Penelitian

a. Akademis

- Dapat memberikan informasi mengenai performa metode LSTM dengan *dropout* dalam meramalkan saham
- Dapat menjadi pertimbangan untuk menggunakan metode LSTM dengan *dropout* dalam bidang lain
- Dapat dijadikan referensi ilmiah bagi peneliti lain yang akan mengembangkan metode LSTM dengan *dropout*

b. Praktis

Dapat menjadi pertimbangan masyarakat dalam mengambil keputusan terhadap saham.

UIN SUNAN AMPEL
S U R A B A Y A

menerangkan bahwa pemilik kertas merupakan bagian dari kepemilikan perusahaan (Umar dan Savitri 2020). Terdapat dua jenis saham yakni

A. Saham biasa (*common stock*)

Saham yang menjelaskan nominal (*dollar*, rupiah, *euro*, dan sebagainya) yang pemegangnya memiliki hak untuk mengikuti RUPS (Rapat Umum Pemegang Saham) dan RUPSLB (Rapat Umum Pemegang Saham Luar Biasa), serta dapat menentukan keputusan untuk membeli *right issue* (penjualan saham terbatas) atau tidak. Pemegang saham akan mendapat keuntungan dalam bentuk dividen di akhir tahun.

B. Saham istimewa (*preferred stock*)

Saham yang menjelaskan nominal (*dollar*, rupiah, *euro*, dan sebagainya) yang pemegangnya akan memperoleh pendapatan tetap setiap kuartal (tiga bulan) dalam bentuk dividen.

Pada dasarnya, terdapat dua keuntungan yang diperoleh investor dari kepemilikan saham ("Saham" t.t.) yaitu

A. Dividen

Adalah pembagian keuntungan yang dihasilkan perusahaan. Dividen akan diberikan setelah pemegang memberikan persetujuan dalam RUPS. Investor akan mendapat dividen, apabila memegang kepemilikan saham dalam waktu relatif lama hingga periode pengakuan sebagai pemegang saham yang berhak mendapat dividen.

B. *Capital Gain*

Adalah selisih antara harga beli dan harga jual. *Capital gain* terbentuk dari aktivitas perdagangan saham di pasar sekunder. Investor akan mendapat *capital gain* apabila harga jual saham lebih tinggi dari harga beli.

Adapun risiko dari kepemilikan saham antara lain

A. *Capital Loss*

Kebalikan dari *capital loss*, yaitu kondisi ketika investor menjual saham dengan harga lebih rendah dari harga beli saham. Sehingga investor akan mendapat kerugian dari selisih perdagangan saham.

B. Risiko Likuidasi

Ketika perusahaan yang sahamnya dimiliki dinyatakan bangkrut, pemegang saham mendapat prioritas terakhir dari hasil penjualan kekayaan perusahaan setelah seluruh kewajiban perusahaan dilunasi. Apabila terdapat sisa dari hasil penjualan, maka sisa tersebut dibagi secara proporsional kepada seluruh pemegang saham. Namun apabila tidak ada sisa, maka pemegang saham tidak akan memperoleh hasil dari likuidasi tersebut. Kondisi ini adalah risiko terberat dari investor saham.

Pasar saham adalah kumpulan sekuritas pembeli dan penjual yang terbagi menjadi bursa saham swasta, bursa saham terbuka, dan bursa saham kepemilikan campuran. Bursa saham swasta melibatkan pertukaran saham perusahaan swasta, sedangkan bursa saham terbuka mencakup saham perusahaan yang terdaftar di pasar saham publik. Saham kepemilikan campuran ada di perusahaan yang sahamnya hanya dapat ditukar sebagian di pasar saham publik (Alkhatib dkk. 2022).

Terdapat beberapa saham yang diperdagangkan di Indonesia seperti AALI milik Astra Agro Lestari Tbk, BBKA milik Bank Central Asia Tbk, CAMP milik Campina Ice Cream Industry Tbk, dan lain sebagainya. Untuk dapat memaksimalkan keuntungan dan meminimalisir risiko atas kepemilikan saham, penggunaan *forecast* atau peramalan dapat membantu.

2.2.2 *Time series Forecasting*

Time series adalah kumpulan pengamatan yang dilakukan secara berurutan sepanjang waktu. Contohnya meliputi suhu udara yang diukur setiap jam, harga saham diukur setiap hari, dan tingkat karbon monoksida diukur setiap tahun (Chatfield 2005). *Diskrit time series* dikumpulkan menjadi tiga jenis yaitu

1. Menjadi sampel dari data kontinu (contoh, temperatur dengan interval tiap jam)

2. Di agregasi selama periode waktu tertentu (contoh, total penjualan dalam sebulan)
3. *Inherently discrete series* (contoh, dividen yang harus dibayarkan oleh perusahaan di tahun berikutnya)

Forecasting adalah sebuah teknik yang menggunakan data historis (data di masa lalu) untuk memberikan informasi di masa depan yang bersifat prediktif (Maricar, t.t.). *Time series forecasting* adalah memberikan informasi yang bersifat prediktif dengan menggunakan data historis berbentuk *time series*. Metode *forecast* secara luas diklasifikasikan menjadi tiga jenis (Chatfield 2000) yaitu

1. *Judgemental forecast* : Data yang digunakan berdasarkan penilaian subjek, intuisi, pengetahuan umum, dan informasi relevan lainnya.
2. *Univariate method* : Data yang digunakan rangkaian tunggal dari masa lalu dan saat ini. Kemungkinan juga ditambah dengan fungsi waktu seperti tren linier
3. *Multivariate method* : Data yang digunakan lebih dari satu variabel yang disebut sebagai variabel prediktor atau penjelas.

Berbagai macam metode diterapkan dalam *time series forecasting*, seperti metode statistik ARIMA yang dilakukan oleh (Almasarweh dan Wadi 2018) dan (Irawan, t.t.). Metode *machine learning* kemudian muncul untuk mengatasi kekurangan metode tradisional yang cukup sulit untuk membuat peramalan non linier seperti *support vector machine* (Sedighi dkk. 2019; Zhang dkk. 2020), *random forest*, *support vector regression* (Nikou, Mansourfar, dan Bagherzadeh 2019), *decision tree*, *adaboost*, dan *xgboost* (Nabipour dkk. 2020). Penggunaan *deep learning* mulai diterapkan karena keunggulannya yang dapat mengekstraksi fitur secara langsung dari data seperti metode *long short term memory* (Nikou, Mansourfar, dan Bagherzadeh 2019), *convolutional neural network* (Hoseinzade dan Haratizadeh 2019), dan lain sebagainya.

2.2.2 *Windowing*

Windowing adalah metode *sub-sampling* (Martínez-Galicia dkk. 2020) yang membuat hasil perkiraan sementara dari data aktual *time series* (Hota, Handa, dan Shrivastava, t.t.). Salah satu jenis *windowing* adalah *sliding window*. *Sliding window* mengumpulkan historis data *time series* untuk membuat prediksi nilai di hari

GPU akan mengoptimalkan operasi *deep learning* secara efisien. Oleh karena itu, diperlukan hardware GPU dan mesin dengan performa tinggi agar pelatihan *deep learning* dapat bekerja dengan baik.

3. *Feature Engineering Process* : Proses ekstraksi *fitur* (karakteristik, properti, dan atribut) dari data menggunakan pengetahuan domain. Perbedaan antara DL dengan ML adalah dapat mengekstraksi karakteristik tingkat tinggi langsung dari data. Dengan demikian, DL dapat mengurangi waktu dan upaya yang diperlukan untuk membuat ekstraktor *fitur*.
4. *Model Training and Execution time* : *Training* pseudocode DL membutuhkan waktu lama karena banyaknya parameter dalam pseudocode DL, sedangkan dengan pseudocode ML membutuhkan waktu yang relatif singkat. Namun, selama pengujian, pseudocode DL membutuhkan waktu yang sangat singkat dibandingkan dengan metode ML tertentu.

Deep neural network terdiri dari banyak *hidden layer* termasuk *input* dan *output layer*. Secara umum, struktur *deep neural network* memiliki *hidden layer* lebih dari 2 (*hidden layer* = N and $N \geq 2$). Teknik DL memiliki beberapa pendekatan seperti

1. Supervised : Penugasan menggunakan data latih berlabel
2. Unsupervised : Penugasan menggunakan kumpulan data tidak berlabel
3. Semi-supervised : Gabungan metode supervised dan unsupervised
4. Reinforcement : Pendekatan berbasis environment

(Sarker 2021) membuat taksonomi DL dengan tiga kategori, yakni metode *deep learning* untuk supervised atau *discriminative learning*, metode *deep learning* untuk unsupervised atau *generative learning*, dan metode *deep learning* gabungan. Taksonomi DL ditunjukkan oleh Gambar 2.4

2. Input Gate

Pada langkah ini, *input gate* di *update* dengan mengkombinasikan *input* terbaru $x^{(t)}$, *output* LSTM unit $y^{(t-1)}$ dan nilai *cell* $c^{(t-1)}$ di iterasi terakhir. Berikut persamaannya

$$i^{(t)} = \sigma(W_i x^{(t)} + R_i y^{(t-1)} + p_i \cdot c^{(t-1)} + b_i) \quad (14)$$

Dimana \cdot menunjukkan perkalian pointwise dua vektor, W_i, R_i dan p_i adalah *weight* yang terkait dengan $x^{(t)}, y^{(t-1)}$, dan $c^{(t-1)}$, sedangkan b_i adalah vektor bias. Pada langkah sebelumnya, *layer* LSTM menentukan informasi mana yang harus disimpan dalam jaringan *cell states* $c^{(t)}$. Termasuk juga pemilihan nilai $z^{(t)}$ yang berpotensi ditambahkan ke *cell state* dan nilai aktivasi $i^{(t)}$ dari *input gate*.

3. Forget Gate

Pada langkah ini, unit LSTM menentukan informasi mana yang harus dihapus dari *cell state* sebelumnya $c^{(t-1)}$. Oleh karena itu, nilai aktivasi $f^{(t)}$ dari *forget gate* pada time step t dihitung berdasarkan *input* saat ini $x^{(t)}$, *output* $y^{(t-1)}$ dan *state* $c^{(t-1)}$ dari memory *cell* time step sebelumnya ($t-1$), *peephole connections*, dan bias b_f dari *forget gate*. Berikut persamaannya

$$f^{(t)} = \sigma(W_f x^{(t)} + R_f y^{(t-1)} + p_f \cdot c^{(t-1)} + b_f) \quad (15)$$

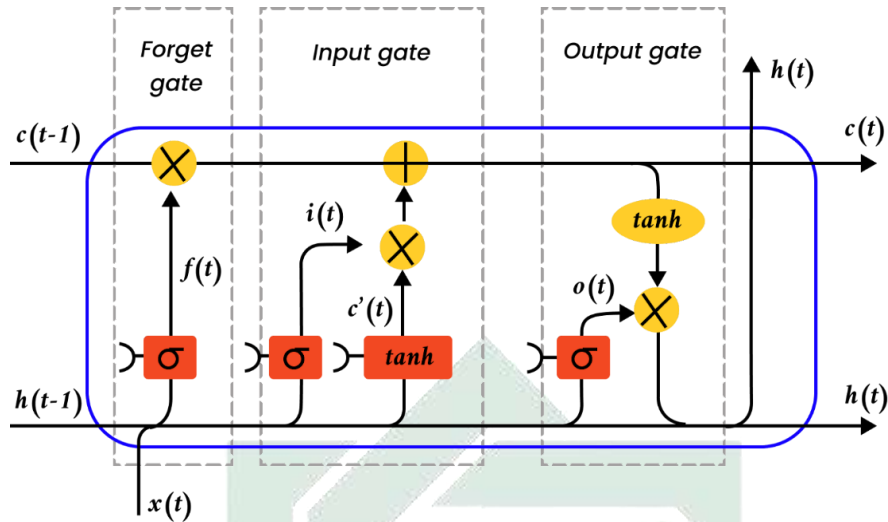
Dimana W_f, R_f , dan p_f adalah *weight* yang terkait dengan $x^{(t)}, y^{(t-1)}$, dan $c^{(t-1)}$, sedangkan b_f menunjukkan vektor *weight* bias.

4. Cell

Langkah ini menghitung nilai *cell*, yang menggabungkan nilai *block input* $z^{(t)}$, *input gate* $i^{(t)}$, dan *forget gate* $f^{(t)}$ dengan nilai *cell* sebelumnya. Persamaannya dapat dilihat berikut ini

$$c^{(t)} = z^{(t)} \cdot i^{(t)} + c^{(t-1)} \cdot f^{(t)} \quad (16)$$

5. Output Gate



Gambar 2. 7 LSTM dengan forget gate (Yu dkk. 2019)

LSTM dengan *forget gate* dapat dinyatakan secara matematis sebagai berikut

$$f_t = \sigma(W_{fh}h_{t-1} + W_{fx}x_t + b_f)$$

$$i_t = \sigma(W_{ih}h_{t-1} + W_{ix}x_t + b_i)$$

$$\dot{c}_t = \tanh(W_{ch}h_{t-1} + W_{cx}x_t + b_c)$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \dot{c}_t$$

$$o_t = \sigma(W_{oh}h_{t-1} + W_{ox}x_t + b_o)$$

$$h_t = o_t \cdot \tanh(c_t) \quad (20)$$

Forget gate memutuskan informasi apa yang akan dibuang dari *cell state*. Ketika nilai dari *forget gate* f_t adalah 1 informasi akan disimpan, apabila bernilai 0 informasi akan dihilangkan.

3. LSTM dengan *peephole connection*

Karena *gates* LSTM *cell* sebelumnya tidak memiliki koneksi langsung dari *cell state*, terdapat kekurangan informasi penting yang mengganggu kinerja jaringan. Pada tahun 2000 *cell* LSTM diperluas dengan memperkenalkan *peephole connections* untuk mengatasi masalah tersebut seperti ditunjukkan Gambar 2.13

- Euclidean *Loss* Function

Fungsi ini banyak digunakan dalam masalah regresi dan disebut juga mean square *error*. Representasi matematis dari euclidean *loss* dapat dilihat di persamaan 11

$$H(p, y) = \frac{1}{2N} \sum_{i=1}^N (p_i - y_i)^2 \quad (11)$$

- Hinge *Loss* Function

Fungsi ini umumnya digunakan dalam masalah yang berkaitan dengan klasifikasi biner. Biasanya digunakan untuk SVM yang berkaitan dengan klasifikasi berdasarkan margin maksimum, dimana *optimizer* berupaya memaksimalkan margin di sekitar kelas tujuan ganda. Rumus matematika dari hinge *loss* dapat dilihat di persamaan 12

$$H(p, y) = \sum_{i=1}^N \max(0, m - (2y_i - 1)p_i) \quad (12)$$

Margin m biasanya diatur di angka 1, p_i adalah *output* yang diprediksi, dan y_i adalah *output* yang diinginkan.

Learning rate didefinisikan sebagai ukuran langkah yang ditentukan untuk pembaruan parameter. Masa pelatihan mewakili pengulangan lengkap dari pembaruan parameter yang melibatkan kumpulan data pelatihan lengkap pada satu waktu. Pemilihan *learning rate* yang tepat diperlukan agar tidak memengaruhi proses pembelajaran secara tidak sempurna. Pengaturan *learning rate* dapat dilakukan ketika mengatur *optimizer* seperti pada Adaptive Moment Estimation (Adam). Berikut penjelasan lebih rinci mengenai Adam.

Adam merupakan tren terbaru *optimizer* dalam *deep learning*. Ini diwakili oleh Hessian matrix yang menggunakan turunan orde kedua. Adam adalah *strategi learning* yang dirancang khusus untuk melatih *deep learning*. Dua keunggulan Adam adalah penggunaan daya komputasi yang lebih sedikit dan penggunaan memori yang lebih hemat. Mekanisme Adam adalah menghitung adaptive LR untuk setiap parameter dalam model. Adam mengintegrasikan metode antara Momentum dan RMSprop. Yakni menggunakan gradien kuadrat untuk menskalakan *learning*

$$R^2 = 1 - \frac{\sum_{i=1}^n (u_i - \hat{u}_i)^2}{\sum_{i=1}^n (u_i - \bar{u})^2} \quad (23)$$

Dimana u_i adalah nilai aktual, \hat{u}_i adalah nilai prediksi, \bar{u} adalah rata-rata dari u_i , dan n adalah jumlah data.

2.3 Integrasi Keilmuan

Untuk mendukung penjelasan mengenai peramalan saham dengan integrasi keilmuan, maka dilakukan wawancara bersama Ibu Wiwin Luqna Hunaida selaku dosen studi al-quran di UIN Sunan Ampel Surabaya. Wawancara dilaksanakan pada tanggal 8 maret 2023 secara online. Beliau memaparkan bahwa peramalan sebuah saham diperbolehkan dalam islam berdasarkan surat yusuf ayat 46-49 yang berbunyi

يُوسُفُ أَيُّهَا الصِّدِّيقُ أَفْتِنَا فِي سَبْعِ بَقَرَاتٍ سِمَانٍ يَأْكُلُهُنَّ سَبْعٌ عِجَافٌ وَسَبْعِ سُنبُلَاتٍ خُضْرٍ وَأُخَرَ
يَبْسُتٍ لِّعَلِّي أَرْجِعَ إِلَى النَّاسِ لَعَلَّهُمْ يَعْلَمُونَ ٤٦

قَالَ تَزْرَعُونَ سَبْعَ سِنِينَ دَابًّا فَمَا حَصَدْتُمْ فَذَرُوهُ فِي سُنْبُلِهِ إِلَّا قَلِيلًا مِّمَّا تَأْكُلُونَ ٤٧

ثُمَّ يَأْتِي مِنْ بَعْدِ ذَلِكَ سَبْعٌ شِدَادٌ يَأْكُلْنَ مَا قَدَّمْتُمْ لَهُنَّ إِلَّا قَلِيلًا مِّمَّا تُحْصِنُونَ ٤٨

ثُمَّ يَأْتِي مِنْ بَعْدِ ذَلِكَ عَامٌ فِيهِ يُغَاثُ النَّاسُ وَفِيهِ يَعْرِضُونَ ٤٩

Terjemah :

[46](Dia berkata,) “Wahai Yusuf, orang yang sangat dipercaya, jelaskanlah kepada kami (takwil mimpiku) tentang tujuh ekor sapi gemuk yang dimakan oleh tujuh (ekor sapi) kurus dan tujuh tangkai (gandum) hijau yang (meliputi tujuh tangkai) lainnya yang kering agar aku kembali kepada orang-orang itu supaya mereka mengetahuinya.” [47](Yusuf) berkata, “Bercocok Tanamlah kamu tujuh tahun berturut-turut! Kemudian apa yang kamu tuai, biarkanlah di tangkainya, kecuali sedikit untuk kamu makan. [48] Kemudian, sesudah itu akan datang tujuh (tahun) yang sangat sulit (paceklik) yang menghabiskan apa yang kamu simpan untuk menghadapinya, kecuali sedikit dari apa (bibit gandum) yang kamu simpan. [49]Setelah itu akan datang tahun, ketika manusia diberi hujan (dengan cukup) dan pada masa itu mereka memeras (anggur).”

Ayat tersebut menceritakan tentang kisah Nabi Yusuf yang menakwilkan mimpi raja mesir atas petunjuk dari Allah SWT dan memberikan saran pada raja mesir untuk bercocok tanam dan tidak menghabiskan hasil panen dalam waktu yang singkat dikarenakan akan ada musim paceklik selama tujuh tahun kemudian setelah tujuh tahun bercocok tanam. Ayat tersebut menjelaskan tentang perlunya pengetahuan, persiapan, dan membuat perencanaan untuk menghadapi sesuatu di masa depan. Diperlukan pengetahuan dan persiapan untuk dapat meramalkan saham di masa depan.

Terdapat firman Allah yang lain dalam surat ar-Ra'd ayat 11 yang berbunyi

لَهُ مَعْقِبَاتٌ مِّنْ بَيْنِ يَدَيْهِ وَمِنْ خَلْفِهِ يَحْفَظُونَهُ مِنْ أَمْرِ اللَّهِ إِنَّ اللَّهَ لَا يُعَيِّرُ مَا بِقَوْمٍ حَتَّىٰ يُعَيِّرُوهُ مَا بِأَنْفُسِهِمْ وَإِذَا أَرَادَ اللَّهُ بِقَوْمٍ سُوءًا فَلَا مَرَدَّ لَهُ وَمَا لَهُمْ مِّنْ دُونِهِ مِنْ وَالٍ

Terjemah :

“Baginya (manusia) ada (malaikat-malaikat) yang menyertainya secara bergiliran dari depan dan belakangnya yang menjaganya atas perintah Allah. Sesungguhnya Allah tidak mengubah keadaan suatu kaum hingga mereka mengubah apa yang ada pada diri mereka. Apabila Allah menghendaki keburukan terhadap suatu kaum, tidak ada yang dapat menolaknya, dan sekali-kali tidak ada pelindung bagi mereka selain Dia”

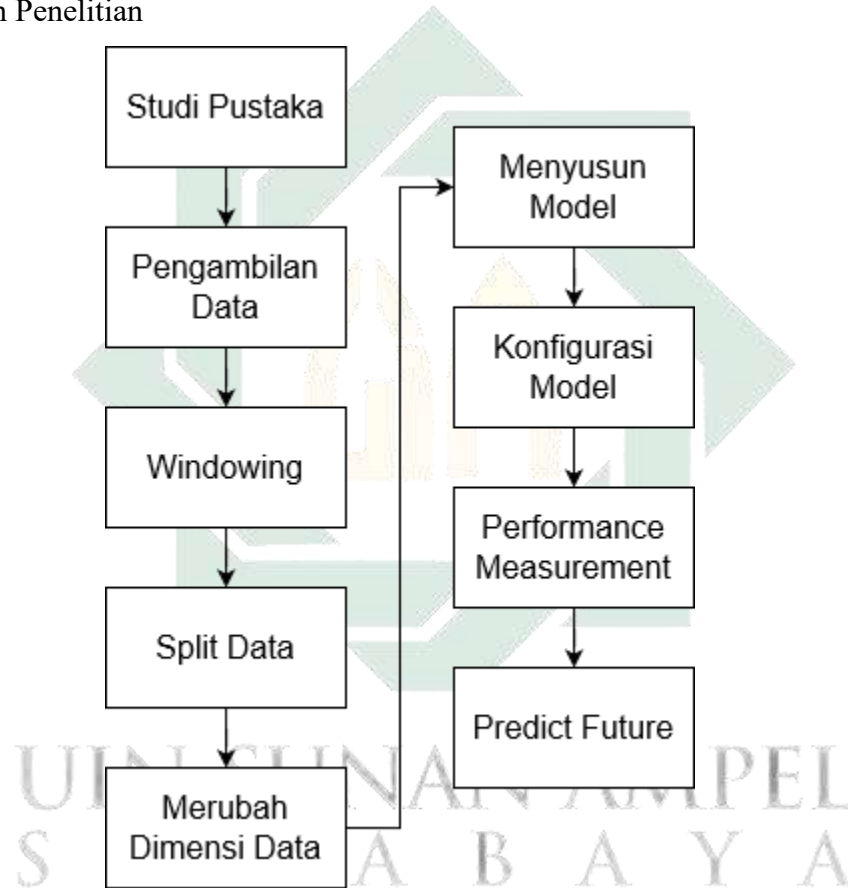
Diperlukan usaha agar dapat menjadi lebih baik, begitu juga dengan ilmu. Sebuah kegiatan akan menjadi lebih baik jika dilakukan oleh orang yang memiliki ilmu tentang hal atau kegiatan tersebut. Sebelum terjun di dunia saham juga diperlukan ilmu atau pengetahuan tentang bisnis tersebut untuk meminimalisir adanya kerugian.

Terdapat satu hadits yang berhubungan dengan peramalan saham juga yakni H.R Bukhari nomor 2537 yang bunyinya “...Sesungguhnya jika kamu meninggalkan ahli warismu dalam keadaan kaya itu lebih baik daripada kamu meninggalkan mereka dalam keadaan miskin lalu mengemis kepada manusia dengan menengadahkan tangan mereka...”. Hadits tersebut menjelaskan bahwa kita sebagai manusia harus berilmu agar tidak menjadi peminta-minta di masa depan. Memberikan pengetahuan dan ilmu yang cukup agar anak bisa hidup layak di masa

depan. Sehingga bisa disimpulkan bahwa dalam melakukan atau menyiapkan segala sesuatu di masa depan haruslah memiliki ilmu seperti dalam kegiatan meramalkan saham.

BAB III METODE PENELITIAN

3.1 Desain Penelitian



Gambar 3. 1 Alur metode penelitian

Alur penelitian ini ditunjukkan oleh Gambar 3.1. Penelitian dimulai dari studi literatur dengan mengumpulkan referensi dan penelitian terdahulu yang relevan dengan peramalan saham dan metode LSTM, kemudian mengambil dataset saham melalui website kaggle. Lalu data akan di split menjadi 3 bagian dan data akan di *reshape* sebelum dimasukkan ke model LSTM. Metode LSTM akan disusun di *build* model dan dilanjut dengan *konfigurasi model*. Lalu model akan diukur performanya dengan *performance measurement*. Dan setelah model memiliki performa yang baik maka model digunakan untuk *predict*.

3.1.1 Studi pustaka

Studi literatur adalah tahapan utama dari penelitian ini dengan mencari referensi dari penelitian terdahulu yang relevan dengan penelitian ini, khususnya pada peramalan saham dan metode LSTM. Pencarian referensi dilakukan dengan bantuan alat google scholar dan Garuda (Garba rujukan digital). Referensi yang didapatkan berasal dari beberapa sumber seperti sciencedirect, iee, springer, mdpi, dan lain sebagainya. Untuk pengecekan kualitas jurnal dilakukan dengan bantuan website SJR (scimago journal & country rank) dan SINTA (science and technology index).

3.1.2 Pengambilan Data

Data yang digunakan adalah salah satu data saham Indonesia yakni AALI milik Astra Agro Lestari Tbk yang diperoleh dari situs kaggle melalui link <https://www.kaggle.com/datasets/muamkh/ihsstockdata>. Data yang diperoleh dari periode tanggal 16 april 2001 hingga 6 januari 2023 dengan jumlah 5779 data. Data yang diperoleh memiliki enam variabel yakni date, open, close, low, high, dan volume. Namun, variabel yang akan digunakan hanya variabel close. Variabel close adalah variabel yang mewakili harga penutupan saham atau cerminan harga saham di hari tertentu.

3.1.3 Windowing

Windowing adalah metode subsampling (Martínez-Galicia dkk. 2020) atau mentransformasi data dari univariate menjadi multiple sample. Jenis windowing yang digunakan adalah sliding window yang akan menghasilkan dua variabel yakni x dan y . Variabel x berisi data historis dengan panjang masing-masing data sebesar ukuran window. Variabel y berisi satu data historis setelah variabel x yang digunakan sebagai label. Terdapat beberapa *window size* yang akan diuji yang disesuaikan dengan rentang yang membentuk pola pada data. *Window size* yang memiliki nilai RMSE yang akan digunakan sebagai proses selanjutnya.

3.1.4 Split Data

Data akan di split menjadi 80% *training set*, 10% *validation set*, dan 10% *testing set*. Split data digunakan untuk mencegah *overfitting* model dan membantu mengevaluasi model secara akurat (Alkhatib dkk. 2022). Dataset di split dengan

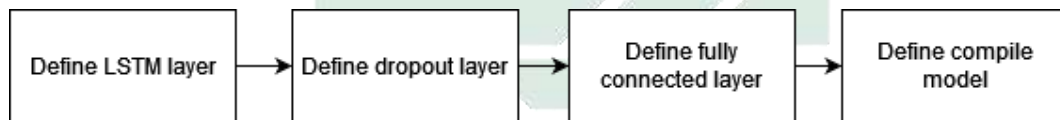
rasio 8:1:1 karena berdasarkan penelitian oleh (Zhao dkk. 2015) rasio tersebut memberikan performa terbaik ketika dibandingkan dengan rasio 6:2:2 dan 9:0.5:0.5.

3.1.5 Reshape Data

Model LSTM membutuhkan *input* data dengan shape 3 dimensi. Untuk memudahkan data diolah oleh LSTM, sebelum data dimasukkan ke model LSTM, data akan diubah menjadi bentuk *array* 3 dimensi. Data akan diubah menjadi (*len_data*, *n_step*, *feature*). *Len_data* merupakan panjang atau jumlah baris yang dimiliki data. Nilai *n_step* sama dengan nilai *window size*. *Feature* merupakan jumlah *fitur* atau variabel yang digunakan sebagai *input*.

3.1.6 Build Model

Model akan dibangun sesuai dengan arsitektur LSTM yang ditunjukkan oleh Gambar 2.14. *Build* model terdiri dari beberapa proses yang ditunjukkan oleh Gambar 3.2



Gambar 3. 2 Alur proses build model

- Define LSTM layer : Mendefinisikan LSTM layer seperti mengatur parameter *hidden node*, dan *activation function*. *Activation function* yang digunakan adalah ReLU (Rectified linear unit) karena fungsi ini dapat mengatasi gradien sangat kecil, memungkinkan model untuk belajar lebih cepat dan bekerja lebih baik (Ghimire dkk. 2022). Pendefinisian *input shape* juga dilakukan di dalam LSTM layer. *Input shape* untuk LSTM didefinisikan dengan (*batch size*, *timestep*, *feature*). *Timestep* adalah jumlah langkah (diisi dengan jumlah *window size*) dan *feature* adalah jumlah *fitur* atau variabel yang digunakan sebagai *input*.
- Define dropout layer : Mendefinisikan nilai *dropout rate*. *Dropout rate* ditulis antara angka 0 sampai 1.

BAB IV HASIL DAN PEMBAHASAN

4.1 Hasil

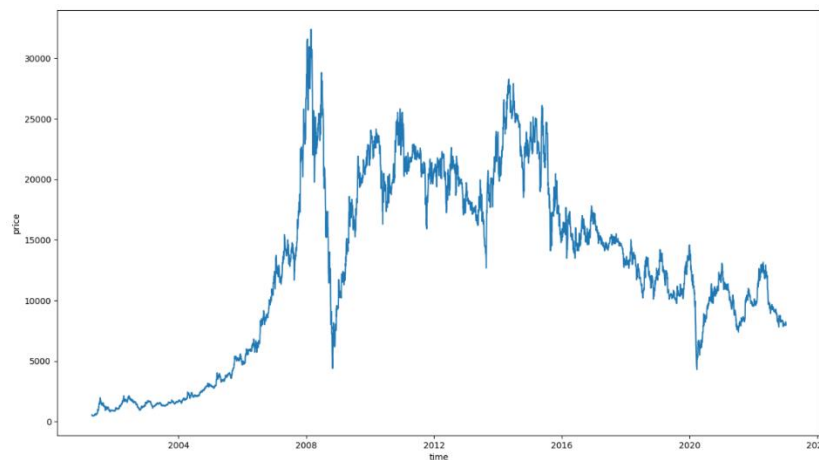
4.1.1 Pengambilan Data

Data yang digunakan adalah salah satu data saham Indonesia yakni AALI milik Astra Agro Lestari Tbk yang diperoleh dari situs kaggle melalui link <https://www.kaggle.com/datasets/muamkh/ihsstockdata>. Data yang diperoleh dari periode tanggal 16 april 2001 hingga 27 Juni 2023 dengan jumlah 5779 data. Data yang diperoleh memiliki enam variabel yakni date, open, close, low, high, dan volume. Namun, variabel yang akan digunakan hanya variabel close. Dataset berbentuk file .csv dan berikut adalah tampilan datanya pada Tabel 4.1

Tabel 4. 1 Tampilan dataset saham harian dari AALI

Date	Close
4/16/2001	547
4/17/2001	524
4/18/2001	524
4/19/2001	500
4/20/2001	452
...	...
6/23/2023	7550
6/26/2023	7475
6/27/2023	7475

Berikut tampilan harga close saham AALI dalam bentuk graphic



Gambar 4. 1 Tampilan grafik dataset saham harian dari AALI

4.1.2 Windowing

Sebelum memasukkan data ke dalam model LSTM, data akan ditentukan terlebih dahulu labelnya. Penentuan label untuk data univariat menggunakan teknik windowing. Teknik ini juga digunakan untuk merubah data univariat menjadi multiple sample agar pelatihan model terhadap data univariat menjadi lebih mudah. Berikut adalah konfigurasi kode untuk proses windowing

Pseudocode 4.1 Kode inisiasi fungsi windowing

```
Function sliding_window(data, window size)
```

```
Data, label = list(), list()
```

```
    for i in range(len(data)):
```

```
        end = i + window_size
```

```
        if end > len(data)-1:
```

```
            break
```

```
        x, y = data[i:end], data[end]
```

```
    append x to data
```

```
    append y to label
```

```
return array data, label
```

Pseudocode 4.1 digunakan untuk membuat fungsi dari windowing dengan nama “sliding_window”. Fungsi “sliding_window” membutuhkan dua masukan yakni data berupa dataset yang ingin dilakukan windowing dan *window size*. Pseudocode 4.1 menghasilkan 2 variabel yakni data dan label. data berisi data histori sebesar *window size* sedangkan label diambil dari satu langkah setelah data.

Untuk mengetahui perubahan dataset setelah dilakukan windowing, berikut hasil windowing dengan contoh *window size* 4 yang disajikan pada Tabel 4.2

Tabel 4. 2 Penampilan data setelah dilakukan windowing

Data 1	Data 2	Data 3	Data 4	label
547	524	524	500	452
524	524	500	452	452
524	500	452	452	471
500	452	452	471	471
452	452	471	471	471

4.1.5 Build Model

Model dibangun sesuai dengan arsitektur LSTM yang ditunjukkan oleh Gambar 2.14. Library python yang digunakan untuk membangun model LSTM adalah keras. Selama *build* model juga dilakukan tuning parameter. Berikut adalah konfigurasi kode untuk *build* model LSTM dengan *dropout* yang ditunjukkan oleh pseudocode 4.4

Pseudocode 4.4

```
Define sequence data()
Define LSTM layer(hidden_node, activation, input_shape)
Define dropout layer(dropout_rate)
Define fully connected layer(hidden_node)
Define compile model(optimizer, loss, metrics)
```

Define sequence data() digunakan sebagai wadah untuk data ketika diinputkan ke model. Pemaparan lebih lanjut mengenai pseudocode ... dijelaskan melalui beberapa sub bab di bawah

Berikut parameter yang akan di tuning di setiap alur dalam *build* model.

- Define LSTM layer

Parameter *hidden node* akan digunakan 5 nilai yang berbeda yakni 10, 30, 50, 100, dan 200. *Activation function* yang digunakan adalah ReLU. Karena penggunaan *activation function* tanh dan sigmoid membuat model LSTM mengalami *underfit* yang dapat dilihat melalui *graphic loss* pada Gambar 4.3

Pseudocode 4.5

```
Model = model.fit(  
input data train and label train,  
Define batch_size value,  
Define epochs value ,  
Input data validation and label validation)
```

Selama pendefinisian *konfigurasi model*, data *training*, nilai *batch size*, jumlah *epoch*, dan data validasi dimasukkan. Model akan diuji dengan 3 parameter *batch size* yang berbeda yakni 5, 10, dan 20. Untuk nilai *epoch* akan diukur terlebih dahulu performanya sebelum melakukan tuning parameter lainnya. Nilai *epoch* yang menghasilkan nilai RMSE terkecil dan R terbesar akan digunakan untuk proses *forecast* selanjutnya. Apabila terdapat *epoch* yang memiliki besaran nilai RMSE yang sama, proses seleksi akan dipilih berdasarkan nilai *gap loss* terkecil. Berikut konfigurasi untuk menghitung *gap loss* yang ditunjukkan pada pseudocode 4.6

Pseudocode 4.6

```
_, error_train = model.evaluate(data_train, label_train)  
_, error_validation = model.evaluate(data_val, label_val)  
print(error_train - error_validation)
```

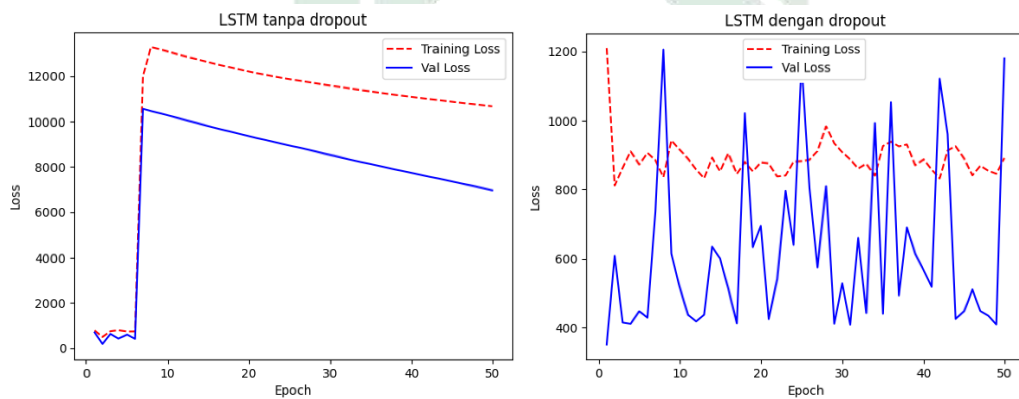
Variabel *error_train* digunakan untuk menghitung nilai rata-rata *loss* pada data *training*. Sedangkan variable *error_validation* digunakan untuk menghitung nilai rata-rata *loss* pada data validasi. Lalu *loss* data *training* akan dikurangi dengan *loss* data validasi dan menghasilkan *gap loss*. Berikut hasil performa dari masing-masing *epoch* yang ditunjukkan oleh Tabel 4.8

Tabel 4. 8 Performa beberapa epoch

<i>Epoch</i>	RMSE	R	<i>Gap loss</i>
10	197	0.98	156
20	182	0.98	137
30	181	0.98	136
40	182	0.98	134

penggunaan *dropout* tidak dapat digunakan untuk memperkecil nilai RMSE atau memperbesar nilai R^2 .

Peningkatan performa model LSTM tergantung pada kombinasi parameter yang digunakan. Seperti kombinasi parameter hidden node 100, batch size 5, dan learning rate 0.1 yang dapat meningkatkan performa model LSTM tanpa dropout yang semula terjadi underfit menjadi lebih baik. Model LSTM tanpa dropout menghasilkan nilai RMSE 5827 sedangkan setelah ditambahkan dropout nilai RMSE berkurang menjadi 1111. Berikut disajikan gap loss yang tertera pada Gambar 4.11



Gambar 4. 11 (a) LSTM tanpa dropout (b) LSTM dengan dropout

Penggunaan nilai parameter yang terlalu kecil juga membuat performa model LSTM dengan dropout menurun. Seperti penggunaan parameter hidden node 10 dengan batch size 5. Yang menghasilkan performa paling tinggi hanya sebesar R^2 0.56 dan RMSE 974 (Tabel 4.12 dan 4.13)

Hasil ini tidak selaras dengan penelitian yang dilakukan oleh Ding bahwasanya penggunaan *dropout* dapat meningkatkan performa model LSTM dalam meramalkan saham dengan masukan multivariat. (Ding dan Qin 2020).

Berdasarkan hasil penelitian ini, penambahan teknik dropout tidak selalu meningkatkan performa LSTM. Teknik dropout dapat membuat performa model LSTM menjadi meningkat ataupun menurun. Tergantung pada kombinasi parameter yang digunakan saat melakukan tuning parameter. Sehingga diperlukan kombinasi parameter yang tepat agar penambahan dropout dapat meningkatkan performa model LSTM.

BAB V PENUTUP

5.1 Kesimpulan

Implementasi model LSTM dengan *dropout* yang memiliki nilai RMSE terkecil dalam penelitian ini adalah menggunakan kombinasi parameter Adam *optimizer*, *learning rate* 0.001, *epoch* 50, *batch size* 20, *hidden layer* 1, *hidden node* LSTM 200, dan ReLU *Activation function*. Performa dari model LSTM dengan *dropout* menggunakan parameter tersebut menghasilkan nilai RMSE 184 dan R^2 0.98.

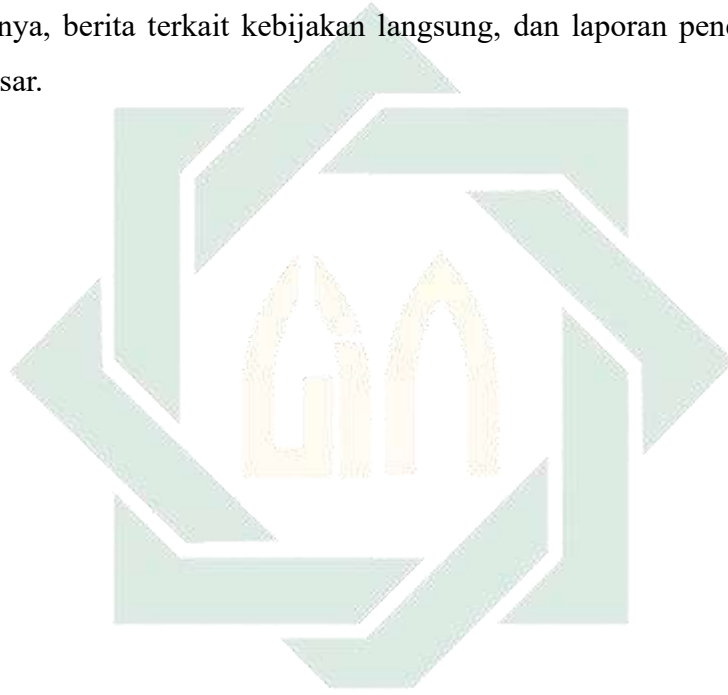
Performa model LSTM dengan *dropout* dalam meramalkan saham Indonesia pada penelitian ini sudah cukup baik. Terdapat model LSTM dengan *dropout* yang memiliki performa lebih tinggi dari model LSTM tanpa *dropout*. Namun, banyak kombinasi parameter dari model LSTM yang menghasilkan performa cenderung lebih rendah ketika ditambahkan *dropout*. Sehingga dapat disimpulkan bahwa tuning parameter dapat memengaruhi performa model LSTM dengan *dropout*.

5.2 Saran

Saran pengembangan untuk penelitian selanjutnya adalah

1. Hasil performa akan berbeda setiap melakukan running model meskipun menggunakan masukan data dan nilai parameter yang sama. Hal ini dipengaruhi oleh perubahan nilai *weight* yang dihasilkan secara random oleh mesin setiap running model. Penggunaan metode *weight* initialization atau save *weight* dari library model checkpoint Keras dapat digunakan untuk menghasilkan *weight* dengan range sama. Tapi perlu diperhatikan efek bias ketika menggunakan *weight* yang sama saat running model. Sehingga perbandingan antar model akan lebih mudah didefinisikan penyebabnya.
2. Penelitian selanjutnya dapat menggunakan nilai parameter yang lebih besar. Tuning parameter juga dapat dilakukan tidak terbatas pada *hidden node*, *batch size*, dan *learning rate*. Dapat ditambahkan parameter lain seperti *epoch* dan *optimizer* saat tuning parameter.

3. Penggunaan model LSTM lain seperti stacked LSTM (*multilayer* LSTM), bidirectional LSTM, convLSTM, atau mengkombinasikannya dengan metode lain dapat dilakukan untuk perbandingan performa dan menghasilkan performa yang lebih tinggi lagi.
4. Dapat mengkombinasikan data histori dengan data textual seperti seperti sentimen investor dari media sosial, laporan pendapatan dari perusahaan yang mendasarinya, berita terkait kebijakan langsung, dan laporan penelitian dari analisis pasar.



UIN SUNAN AMPEL
S U R A B A Y A

DAFTAR PUSTAKA

- Abraham, Rebecca, Mahmoud El Samad, Amer M. Bakhach, Hani El-Chaarani, Ahmad Sardouk, Sam El Nemar, dan Dalia Jaber. 2022. "Forecasting a Stock Trend Using Genetic Algorithm and Random Forest." *Journal of Risk and Financial Management* 15 (5): 188.
- Alhussein, Musaed, Khursheed Aurangzeb, dan Syed Irtaza Haider. 2020. "Hybrid CNN-LSTM Model for Short-Term Individual Household Load Forecasting." *IEEE Access* 8: 180544–57.
- Alkhatib, Khalid, Huthaifa Khazaleh, Hamzah Ali Alkhazaleh, Anas Ratib Alsoud, dan Laith Abualigah. 2022. "A New Stock Price Forecasting Method Using Active Deep Learning Approach." *Journal of Open Innovation: Technology, Market, and Complexity* 8 (2): 96.
- Almasarweh, Mohammad, dan S. Al Wadi. 2018. "ARIMA Model in Predicting Banking Stock Market Data." *Modern Applied Science* 12 (11): 309.
- Alzubaidi, Laith, Jinglan Zhang, Amjad J. Humaidi, Ayad Al-Dujaili, Ye Duan, Omran Al-Shamma, J. Santamaría, Mohammed A. Fadhel, Muthana Al-Amidie, dan Laith Farhan. 2021. "Review of Deep Learning: Concepts, CNN Architectures, Challenges, Applications, Future Directions." *Journal of Big Data* 8 (1): 53.
- Ba, Jimmy Lei, Jamie Ryan Kiros, dan Geoffrey E. Hinton. 2016. "Layer Normalization." arXiv.
- Bhandari, Hum Nath, Binod Rimal, Nawa Raj Pokhrel, Ramchandra Rimal, Keshab R. Dahal, dan Rajendra K.C. Khatri. 2022. "Predicting Stock Market Index Using LSTM." *Machine Learning with Applications* 9 (September): 100320.
- Cahuantzi, Roberto, Xinye Chen, dan Stefan Güttel. 2023. "A Comparison of LSTM and GRU Networks for Learning Symbolic Sequences." arXiv.
- Chatfield, C. 2000. *Time-Series Forecasting*. CRC Press.
<https://books.google.co.id/books?id=PFHMBQAAQBAJ>.
- Chatfield, Chris. 2005. "Time-series forecasting." *Significance* 2 (3): 131–33.
- Chung, Junyoung, Caglar Gulcehre, KyungHyun Cho, dan Yoshua Bengio. 2014. "Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling." arXiv.
- Dao, Dong, Hai-Bang Ly, Son Trinh, Tien-Thinh Le, dan Binh Pham. 2019. "Artificial Intelligence Approaches for Prediction of Compressive Strength of Geopolymer Concrete." *Materials* 12 (6): 983. <https://doi.org/10.3390/ma12060983>.
- "Dataset Saham Indonesia / Indonesia Stock Dataset." t.t. Diakses 7 Maret 2023.
<https://www.kaggle.com/datasets/muamkh/ihsghostockdata>.
- Desai, Prathamesh S., Victoria Granja, dan C. Fred Higgs. 2021. "Lifetime Prediction Using a Tribology-Aware, Deep Learning-Based Digital Twin of Ball Bearing-Like Tribosystems in Oil and Gas." *Processes* 9 (6): 922.
- Ding, Guangyu, dan Liangxi Qin. 2020. "Study on the Prediction of Stock Price Based on the Associated Network Model of LSTM." *International Journal of Machine Learning and Cybernetics* 11 (6): 1307–17.

- García, Salvador, Sergio Ramírez-Gallego, Julián Luengo, José Manuel Benítez, dan Francisco Herrera. 2016. "Big Data Preprocessing: Methods and Prospects." *Big Data Analytics* 1 (1): 9
- Ghimire, Sujan, Binayak Bhandari, David Casillas-Pérez, Ravinesh C. Deo, dan Sancho Salcedo-Sanz. 2022. "Hybrid Deep CNN-SVR Algorithm for Solar Radiation Prediction Problems in Queensland, Australia." *Engineering Applications of Artificial Intelligence* 112 (Juni): 104860.
- Haq, Anwar Ul, Adnan Zeb, Zhenfeng Lei, dan Defu Zhang. 2021. "Forecasting Daily Stock Trend Using Multi-Filter Feature Selection and Deep Learning." *Expert Systems with Applications* 168 (April): 114444.
- Hoseinzade, Ehsan, dan Saman Haratizadeh. 2019. "CNNpred: CNN-Based Stock Market Prediction Using a Diverse Set of Variables." *Expert Systems with Applications* 129 (September): 273–85.
- Hota, H S, Richa Handa, dan A K Shrivastava. t.t. "Time Series Data Prediction Using Sliding Window Based RBF Neural Network."
- Hu, Zexin, Yiqi Zhao, dan Matloob Khushi. 2021. "A Survey of Forex and Stock Price Prediction Using Deep Learning." *Applied System Innovation* 4 (1): 9.
- Huang, Chiou-Jye, dan Ping-Huan Kuo. 2018. "A Deep CNN-LSTM Model for Particulate Matter (PM2.5) Forecasting in Smart Cities." *Sensors* 18 (7): 2220.
- Irawan, Wici. t.t. "PERAMALAN HARGA SAHAM PT. UNIVELER TBK. DENGAN MENGGUNAKAN MODEL ARIMA."
- Jing, Nan, Zhao Wu, dan Hefei Wang. 2021. "A Hybrid Model Integrating Deep Learning with Investor Sentiment Analysis for Stock Price Prediction." *Expert Systems with Applications* 178 (September): 115019.
- Khan, Wasiat, Mustansar Ali Ghazanfar, Muhammad Awais Azam, Amin Karami, Khaled H. Alyoubi, dan Ahmed S. Alfakeeh. 2022. "Stock Market Prediction Using Machine Learning Classifiers and Social Media, News." *Journal of Ambient Intelligence and Humanized Computing* 13 (7): 3433–56.
- Kiranyaz, Serkan, Onur Avci, Osama Abdeljaber, Turker Ince, Moncef Gabbouj, dan Daniel J. Inman. 2021. "1D Convolutional Neural Networks and Applications: A Survey." *Mechanical Systems and Signal Processing* 151 (April): 107398.
- Kumbure, Mahinda Mailagaha, Christoph Lohrmann, Pasi Luukka, dan Jari Porras. 2022. "Machine learning techniques and data for stock market forecasting: A literature review." *Expert Systems with Applications* 197 (Juli): 116659.
- Kurniawan, Eliv, Hari Wibawanto, dan Djoko Adi Widodo. 2019. "Implementasi Metode Backpropogation dengan Inisialisasi Bobot Nguyen Widrow untuk Peramalan Harga Saham." *Jurnal Teknologi Informasi dan Ilmu Komputer* 6 (1): 49.
- Lee, Si Woon, dan Ha Young Kim. 2020. "Stock Market Forecasting with Super-High Dimensional Time-Series Data Using ConvLSTM, Trend Sampling, and Specialized Data Augmentation." *Expert Systems with Applications* 161 (Desember): 113704.
- Liu, Shuai, Hong Ji, dan Morgan C. Wang. 2020. "Nonpooling Convolutional Neural Network Forecasting for Seasonal Time Series With Trends." *IEEE Transactions on Neural Networks and Learning Systems* 31 (8): 2879–88.

- Lu, Wenjie, Jiazheng Li, Yifan Li, Aijun Sun, dan Jingyang Wang. 2020. "A CNN-LSTM-Based Model to Forecast Stock Prices." Disunting oleh Abd E. I.-Baset Hassanien. *Complexity* 2020 (November): 1–10.
- Maricar, M Azman. t.t. "Analisa Perbandingan Nilai Akurasi Moving Average dan Exponential Smoothing untuk Sistem Peramalan Pendapatan pada Perusahaan XYZ" 13 (2).
- Martínez-Galicia, David, Alejandro Guerra-Hernández, Nicandro Cruz-Ramírez, Xavier Limón, dan Francisco Grimaldo. 2020. "Windowing as a Sub-Sampling Method for Distributed Data Mining." *Mathematical and Computational Applications* 25 (3): 39.
- Moghaddasi, Hamid, Bahareh Ahmadzadeh, Reza Rabiei, dan Mohammad Farahbakhsh. 2017. "Study on the Efficiency of a Multi-Layer Perceptron Neural Network Based on the Number of Hidden Layers and Nodes for Diagnosing Coronary- Artery Disease." *Jentashapir Journal of Health Research* 8 (3).
- Nabipour, Mojtaba, Pooyan Nayyeri, Hamed Jabani, Shahab S., dan Amir Mosavi. 2020. "Predicting Stock Market Trends Using Machine Learning and Deep Learning Algorithms Via Continuous and Binary Data; a Comparative Analysis." *IEEE Access* 8: 150199–212.
- Nikou, Mahla, Gholamreza Mansourfar, dan Jamshid Bagherzadeh. 2019. "Stock Price Prediction Using DEEP Learning Algorithm and Its Comparison with Machine Learning Algorithms." *Intelligent Systems in Accounting, Finance and Management* 26 (4): 164–74.
- Rajagukguk, Rial A., Raden A. A. Ramadhan, dan Hyun-Jin Lee. 2020. "A Review on Deep Learning Models for Forecasting Time Series Data of Solar Irradiance and Photovoltaic Power." *Energies* 13 (24): 6623.
- "Saham." t.t. suit-baze. Diakses 12 April 2023. <https://www.idx.co.id/>.
- Sarker, Iqbal H. 2021. "Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions." *SN Computer Science* 2 (6): 420.
- Sedighi, Mojtaba, Hossein Jahangirnia, Mohsen Gharakhani, dan Saeed Farahani Fard. 2019. "A Novel Hybrid Model for Stock Price Forecasting Based on Metaheuristics and Support Vector Machine." *Data* 4 (2): 75.
- Sezer, Omer Berat, Mehmet Ugur Gudelek, dan Ahmet Murat Ozbayoglu. 2020. "Financial Time Series Forecasting with Deep Learning : A Systematic Literature Review: 2005–2019." *Applied Soft Computing* 90 (Mei): 106181.
- Shi, Lei, Zhiyang Teng, Le Wang, Yue Zhang, dan Alexander Binder. 2019. "DeepClue: Visual Interpretation of Text-Based Deep Stock Prediction." *IEEE Transactions on Knowledge and Data Engineering* 31 (6): 1094–1108.
- Torres, José F., Dalil Hadjout, Abderrazak Sebaa, Francisco Martínez-Álvarez, dan Alicia Troncoso. 2021. "Deep Learning for Time Series Forecasting: A Survey." *Big Data* 9 (1): 3–21.
- Umar, AhmadAhmad Ulil Albab Al, dan Anava Salsa Nur Savitri. 2020. "ANALISIS PENGARUH ROA, ROE, EPS TERHADAP HARGA SAHAM," 2, 4 (September): 92–98.
- Van Houdt, Greg, Carlos Mosquera, dan Gonzalo Nápoles. 2020. "A Review on the Long Short-Term Memory Model." *Artificial Intelligence Review* 53 (8): 5929–55.

- Vidal, Andrés, dan Werner Kristjanpoller. 2020. "Gold Volatility Prediction Using a CNN-LSTM Approach." *Expert Systems with Applications* 157 (November): 113481.
- Wang, Wei-Jia, Yong Tang, Jason Xiong, dan Yi-Cheng Zhang. 2021. "Stock Market Index Prediction Based on Reservoir Computing Models." *Expert Systems with Applications* 178 (September): 115022.
- Wibawa, Aji Prasetya, Agung Bella Putra Utama, Hakkun Elmunsyah, Utomo Pujianto, Felix Andika Dwiyanto, dan Leonel Hernandez. 2022. "Time-Series Analysis with Smoothed Convolutional Neural Network." *Journal of Big Data* 9 (1): 44.
- Wu, Jimmy Ming-Tai, Zhongcui Li, Norbert Herencsar, Bay Vo, dan Jerry Chun-Wei Lin. 2021. "A Graph-Based CNN-LSTM Stock Price Prediction Algorithm with Leading Indicators." *Multimedia Systems*, Februari.
- Yang, Jung-Jin, Gyeong Gang, dan Tae Kim. 2018. "Development of EOG-Based Human Computer Interface (HCI) System Using Piecewise Linear Approximation (PLA) and Support Vector Regression (SVR)." *Electronics* 7 (3): 38.
- Yu, Yong, Xiaosheng Si, Changhua Hu, dan Jianxun Zhang. 2019. "A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures." *Neural Computation* 31 (7): 1235–70.
- Zeng, Yongkang, Jingjing Chen, Ning Jin, Xiaoping Jin, dan Yang Du. 2022. "Air Quality Forecasting with Hybrid LSTM and Extended Stationary Wavelet Transform" 213 (April).
- Zhang, Jun, Yuan-Hai Shao, Ling-Wei Huang, Jia-Ying Teng, Yu-Ting Zhao, Zhu-Kai Yang, dan Xin-Yang Li. 2020. "Can the Exchange Rate Be Used to Predict the Shanghai Composite Index?" *IEEE Access* 8: 2188–99.
- Zhang, Jun, Yu-Fan Teng, dan Wei Chen. 2019. "Support Vector Regression with Modified Firefly Algorithm for Stock Price Forecasting." *Applied Intelligence* 49 (5): 1658–74.
- Zhao, Zongyuan, Shuxiang Xu, Byeong Ho Kang, Mir Md Jahangir Kabir, Yunling Liu, dan Rainer Wasinger. 2015. "Investigation and Improvement of Multi-Layer Perceptron Neural Networks for Credit Scoring." *Expert Systems with Applications* 42 (7): 3508–16.
- Zhou, Luyu, Chun Zhao, Ning Liu, Xingduo Yao, dan Zewei Cheng. 2023. "Improved LSTM-Based Deep Learning Model for COVID-19 Prediction Using Optimized Approach." *Engineering Applications of Artificial Intelligence* 122 (Juni): 106157.