

# Peramalan Harga Saham Pertambangan Pada Bursa Efek Indonesia (BEI) Menggunakan *Long Short Term Memory* (LSTM)

**Roby Julian<sup>\*1</sup>, Muhammad Rizky Pribadi<sup>2</sup>**

<sup>1,2</sup>Universitas MDP; Jalan Rajawali No. 14, (0711) 376400

<sup>3</sup>Program Studi Teknik Informatika, Fakultas Ilmu Komputer dan Rekayasa Universitas MDP, Palembang

e-mail: <sup>\*1</sup>robyjulian212@gmail.com, <sup>2</sup>rizky@mdp.ac.id

## **Abstrak**

*Investasi saham merupakan salah satu pilihan yang tepat untuk mendapatkan keuntungan lebih. Akan tetapi dalam melakukan investasi saham diperlukan ilmu analisis terhadap data sebuah perusahaan yang dapat menentukan naik atau turunnya suatu harga saham pada perusahaan tersebut. Pergerakan yang sangat dinamis memerlukan pemodelan data untuk melakukan prediksi harga saham agar mendapatkan tingkat akurasi yang tinggi. Sebuah algoritma dikembangkan untuk mengatasi masalah data jangka panjang atau data historis yaitu Long Short Term Memory (LSTM). Dengan menggunakan Long Short Term (LSTM) penelitian ini menghasilkan nilai RMSE yang cukup baik dengan peningkatan nilai RMSE berdasarkan penambahan jumlah variasi epoch. Variasi epoch optimal didapatkan dengan jumlah epoch sebesar 200. Sedangkan nilai RMSE optimal yang dihasilkan metode Long Short Term Memory (LSTM) dihasilkan oleh emiten TINS dengan RMSE sebesar 31.71.*

**Kata kunci**— *Long Short Term Memory, Data Historis, Harga Saham, Pertambangan*

## **Abstract**

*Stock investment is one of the right choices to get more profit. However, in investing in stocks, it is necessary to analyze the data of a company that can determine the rise or fall of a stock price in the company. Very dynamic movements require data modeling to predict stock prices in order to get a high level of accuracy. An algorithm was developed to solve the problem of long-term data or historical data, namely Long Short Term Memory (LSTM). By using the Long Short Term (LSTM) this study produces a fairly good RMSE value with an increase in the RMSE value based on the addition of the number of epoch variations. The optimal epoch variation was obtained with the number of epochs of 200. Meanwhile, the optimal RMSE value produced by the Long Short Term Memory (LSTM) method was generated by TINS issuers with an RMSE of 31.71.*

**Keywords**— *Long Short Term Memory, Historical Data, Stock Price, Mining*

## 1. PENDAHULUAN

**I**nvestasi di pasar bursa dalam bentuk kepemilikan saham perusahaan atau bentuk komoditas lainnya sangat menarik dan menggoda karena dapat menjanjikan keuntungan yang cukup besar. Namun, investasi saham diklasifikasikan sebagai risiko tinggi, karena sensitifnya pergerakan harga saham yang dipengaruhi oleh faktor luar maupun dalam negeri. Faktor

tersebut bisa berupa perubahan finansial, peraturan dalam undang-undang pada industri, perusahaan yang mengeluarkan saham itu sendiri dan tentunya pengaruh dari ekonomi negara itu sendiri.

Harga saham sangat sulit ditebak pergerakan arah kelajuannya. Pergerakan harga saham tersebut biasanya dapat prediksi dengan melakukan tiga teknik faktor yang dapat dilihat, yaitu faktor teknikal, faktor fundamental dan faktor sentimen. Faktor teknikal adalah gerakan dengan mengamati harga di masa lalu, faktor fundamental adalah teknik analisis bisnis yang terjadi, faktor sentimen adalah pergerakan harga saham yang dipengaruhi oleh faktor bisnis, berita dan operasi bisnis.

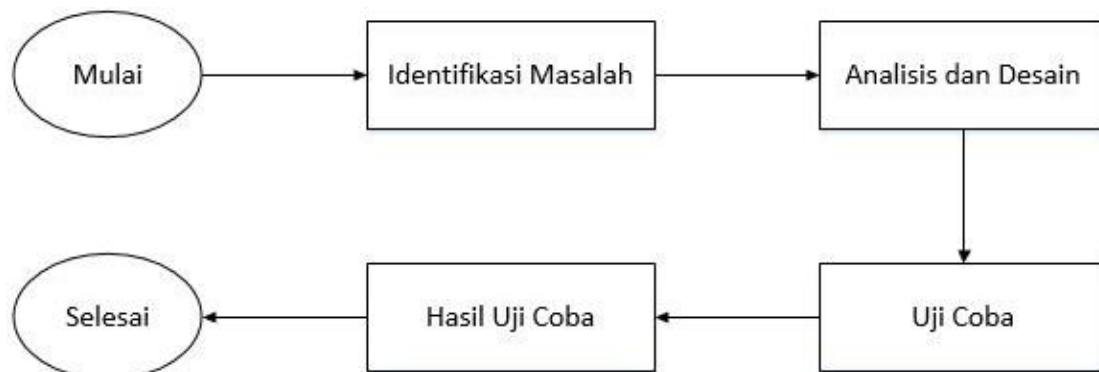
Memprediksi dalam hal jual beli saham terdapat banyak sekali metode ataupun caranya. Sebuah metode dari kecerdasan buatan lebih tepatnya sub dari artificial intelligence yaitu machine learning adalah metode yang dapat melakukan pendekatan dalam memprediksi harga saham. Machine Learning merupakan sub dari Artificial intelligence yang memiliki tujuan untuk meningkatkan pengetahuan atau performa [7]. Recurrent Neural Network (RNN) merupakan salah satu algoritma dari machine learning. RNN memiliki tingkat keakuratan yang tinggi untuk melakukan prediksi data dalam bentuk time series [1]. Sebuah algoritma dikembangkan berdasarkan algoritma RNN adalah algoritma Long Short Term Memory (LSTM), dimana algoritma tersebut dapat mengekstraksi informasi dari data jangka panjang, time series atau sequential. Long Short Term Memory (LSTM) mampu mengatasi vanishing gradient atau keadaan nilai gradien adalah 0 atau dekat dengan 0 dengan mekanisme gate.

Dalam penelitian sebelumnya oleh Suyudi et al [9], penelitian dilakukan dalam 3 tahap untuk memprediksi harga saham yaitu pra-proses, proses pelatihan data dan proses pengujian. Penelitian tersebut menunjukkan bahwa metode RNN tidak pandai memprediksi harga saham. Berbeda dengan penelitian yang dilakukan oleh Fauzi [3], di mana penelitiannya memperkirakan model menggunakan algoritma LSTM untuk variasi epoch. Didalam penelitian tersebut menggunakan 20 epoch menghasilkan nilai MSE sebesar 0.00019 dan RMSE sebesar 0.014. Dalam penelitian Arfan & Lussiana [2], menjelaskan rentang data yang semakin besar digunakan dalam SVM maka nilai MSE yang dihasilkan semakin besar. Sedangkan nilai MSE yang dihasilkan oleh LSTM memiliki nilai yang sama meskipun menggunakan rentang data yang berbeda. Pada penelitian yang dilakukan Riyantoko et al [8], membandingkan tiga optimasi pada algoritma LSTM yaitu SGD, Adam dan RMSprop. Menghasilkan bahwa optimasi Adam dengan variasi epoch sangat menunjukkan stabil turun nilai RMSE-nya dari rendah ke tinggi. Beda halnya dengan Adhitho Satyo Bayangkari Karno, [6] ia melakukan penelitian terhadap satu emiten telekomunikasi yaitu TLKM melakukan transformasi pada data dan menghasilkan prediksi saham TLKM dengan hasil prediksi mencapai tingkat akurasi 99% dengan nilai RMSE 1%, dimana penelitian tersebut data melalui proses transformasi loging kemudian Moving Average atau loging kemudian EWMA. Berdasarkan hasil prediksi harga saham menunjukkan nilai error atau kesalahan yang kecil, maka bisa disimpulkan bahwa LSTM mampu mengatasi ketergantungan data time series atau jangka panjang dan mampu menghasilkan prediksi harga saham dengan hasil akurat.

Penelitian ini akan menggunakan data dari emiten atau perusahaan yang terdaftar pada Bursa Efek Indonesia (BEI) pada sektor pertambangan dan mengalami pertumbuhan harga saham tercatat meningkat drastic serta terdapat fenomena yang kuat sehingga emiten saham tersebut layak untuk dilakukan penelitian. Mengacu pada informasi bahwa Long Short Term Memory (LSTM) dapat mengolah permasalahan data time series atau jangka panjang. Dalam melakukan penelitian ini beberapa tahapan akan dijalankan yaitu mengidentifikasi masalah, analisis dan desain, uji coba dan hasil uji coba. Sehingga tercapai tujuan penelitian yaitu untuk meramalkan masa depan dari harga saham pertambangan pada bursa efek Indonesia berdasarkan data historisnya menggunakan metode Long Short Term Memory (LSTM).

## 2. METODE PENELITIAN

Pada penelitian akan dilakukan beberapa tahapan mulai dari identifikasi masalah yang ada berdasarkan fenomena yang terjadi sekarang dan berdasarkan referensi penelitian terdahulu yang berkaitan, kemudian dengan melakukan studi literatur untuk mencari dan mengumpulkan beberapa referensi dari jurnal maupun buku yang berkaitan dengan penelitian. Tahap berikutnya adalah analisis dan desain, dimulai dengan pengumpulan data harga saham emiten pertambangan yang terdaftar pada Bursa Efek Indonesia (BEI) yaitu ANTM, TINS dan INCO. Data tersebut diambil di website Yahoo Finance dan data yang tercatat dari tanggal 1 Januari 2016 sampai 31 Desember 2020. Setelah pengumpulan data penelitian dilanjutkan dengan perancangan sistem dan akan diimplementasikan menjadi sebuah sistem untuk peramalan harga saham emiten/perusahaan pertambangan. Tahap berikutnya dilakukan uji coba pada sistem dengan skenario pengujian yang telah dirancang, kemudian pada tahap akhir akan dilakukan evaluasi terhadap hasil skenario uji coba yang telah dilakukan dengan penentuan nilai RMSE. Tahapan penelitian dapat dilihat pada Gambar 1 berikut.



Gambar 1. Tahapan Penelitian

### 2.1 Identifikasi Masalah

Penelitian dimulai dengan melakukan identifikasi masalah yang berkaitan dengan peramalan atau forecasting harga saham yang terdaftar pada Bursa Efek Indonesia (BEI) serta mengidentifikasi fenomena-fenomena yang terjadi pada pasar bursa saham. Tahapan ini merupakan pencarian dan mempelajari beberapa jurnal dan buku yang berkaitan dengan peramalan atau forecasting harga saham menggunakan metode yang bermacam-macam, jurnal yang berkaitan dengan metode Long Short Term Memory (LSTM), Jurnal dan buku yang berkaitan dengan ilmu saham sebagai pendukung teori pada penelitian.

### 2.2 Analisis dan Desain

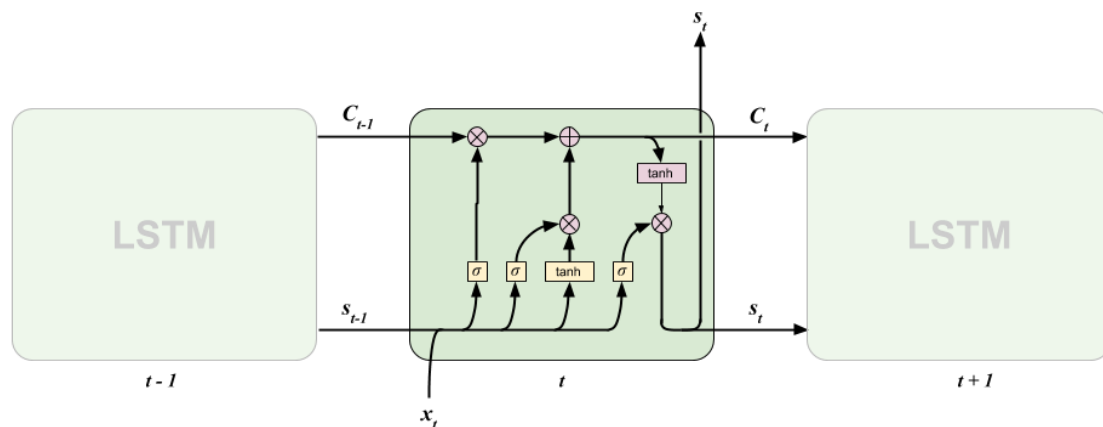
Pada tahapan ini dilakukan pengumpulan data untuk peramalan harga saham yang berupa data historis sebuah emiten selama 5 tahun kebelakang. Data yang digunakan dataset public yang berasal dari website Yahoo Finance. Pengambilan merupakan harga saham dari emiten pertambangan yang telah terdaftar atau listing pada Bursa Efek Indonesia (BEI). Masing-masing emiten yaitu ANTM, TINS dan INCO yang merupakan data historis tercatat dari tanggal 1 Januari 2016 sampai 31 Desember 2020 [5]. Data historis tersebut terdiri dari Date, Open, High, Low, Close, Adj Close, Volume. Date merupakan tanggal, bulan dan tahun perdagangan pasar saham. Open merupakan harga pembuka perdagangan pasar saham. High merupakan harga tertinggi yang dicapai perdagangan pasar saham. Low merupakan harga terendah yang dicapai perdagangan pasar saham. Close merupakan harga penutupan perdagangan pasar saham berakhir. Adj Close merupakan harga penutupan perdagangan pasar saham yang sudah

disesuaikan ketika terjadi aksi korporasi perusahaan. Volume merupakan jumlah saham yang diperdagangkan pada priode tertentu.

Date	Open	High	Low	Close	Adj Close	Volume
2016-01-04	315	315	296	296	289,2284	96322900
2016-01-05	297	305	297	302	295,0911	40252100
2016-01-06	303	307	298	300	293,1369	79060500
2016-01-07	298	301	294	294	287,2741	60458100
2016-01-08	292	302	292	299	292,1598	43796800
2016-01-11	299	308	296	303	296,0682	105459000
...						
2020-12-22	1950	1960	1830	1835	1835	678922000
2020-12-23	1830	1850	1720	1790	1790	659296400
2020-12-28	1825	2000	1780	2000	2000	992175000
2020-12-29	2050	2070	1935	1945	1945	739754900
2020-12-30	1945	1985	1905	1935	1935	535943500

Gambar 2. Dataset ANTM Penelitian

Penelitian menggunakan metode Long Short Term Memory (LSTM) yang difokuskan pada fitur harga penutupan close harian pada setiap emiten sebagai data input. Harga penutupan (close) sangat penting dikarenakan menjadi acuan untuk harga pembukaan di keesokan harinya [4]. Data tersebut akan dilakukan proses pre processing normalisasi pada dataset untuk meminimalkan error mengubah data aktual menjadi nilai dengan range [0,1] dengan teknik MinMaxScaler dari sklearn. Teknik ini dipakai untuk mengatasi masalah gap nilai yang cukup besar dari harga close per harinya. Selanjutnya data akan dialokasikan menjadi 2 dataset yaitu data uji dan data latih masing-masing sebesar 80% dan 20%.

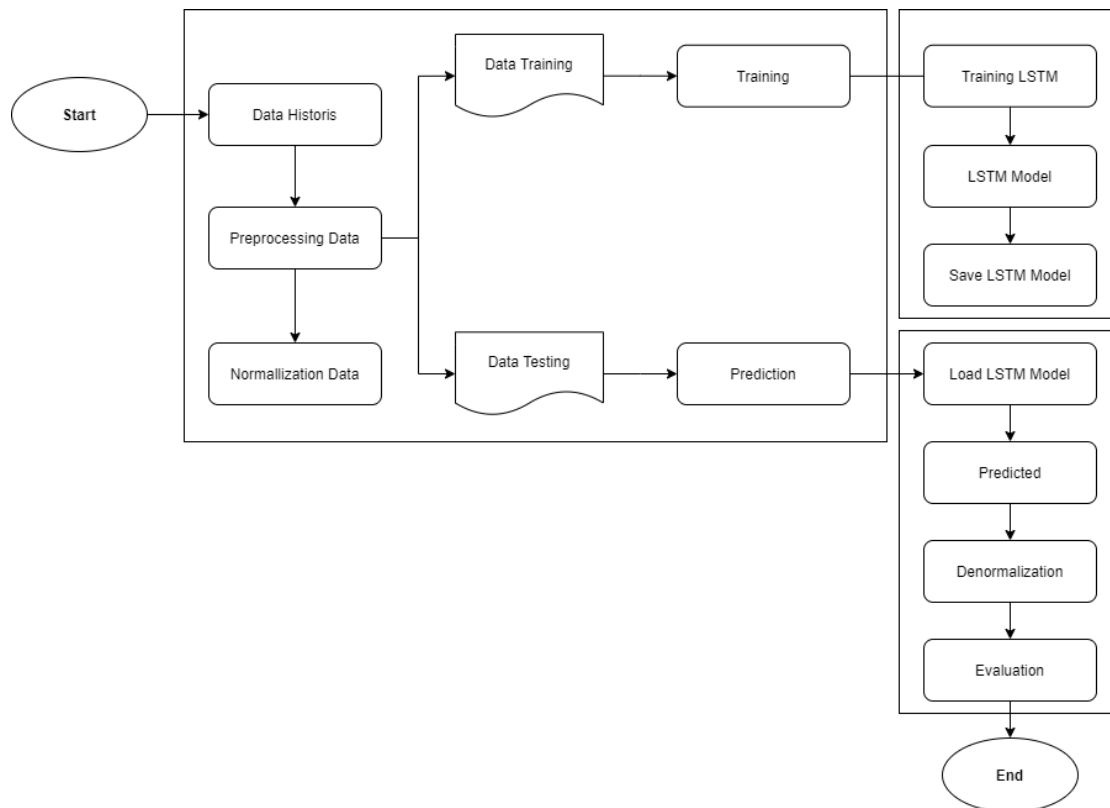


Gambar 3. Arsitektur LSTM

Desain perancangan model Long Short Term Memory (LSTM) dimulai dengan inialisasi parameter yang diproses Long Short Term Memory (LSTM), parameter yang digunakan adalah lapisan tersembunyi (hidden layer), memori sel (neurons), putaran (epoch) dan jumlah sampel data (batch size). Long Short Term Memory (LSTM) dalam perancangannya akan menggunakan optimasi ADAM. Hasil dari sistem ini akan berupa nilai Root Mean Square Error (RMSE) sebagai penentu seberapa dekat nilai prediksi yang dihasilkan dengan nilai aktual. Pada tahap ini diimplementasikan ke dalam Bahasa pemrograman Python 3.8.10 64-bit.

### 2.3 Uji Coba

Tahap ini merupakan tahapan pengujian terhadap algoritma atau model yang telah diimplementasikan pada tahap sebelumnya. Ditahapan ini pengujian akan dilakukan seperti gambar 4 berikut.



Gambar 4. Tahapan Pengujian

Pada tahap pengujian ini penelitian akan mengikuti skenario pengujian dengan variasi epoch sebagai berikut neurons sejumlah 50, hidдем layer sejumlah 2, batch size sejumlah 2 dan variasi epoch 50, 100, 150, 200.

#### 3.2.4 Hasil Uji Coba

Hasil dari uji coba sistem akan berupa prediksi harga saham dan nilai RMSE dari ketiga emiten saham pertambangan yaitu ANTM, INCO dan TINS. Nilai tersebut akan dilakukan perbandingan berdasarkan variasi epoch dan ditampilkan dalam bentuk grafik serta tabel yang nantinya dianalisis membentuk suatu kesimpulan serta rekomendasi penelitian.

## 3. HASIL DAN PEMBAHASAN

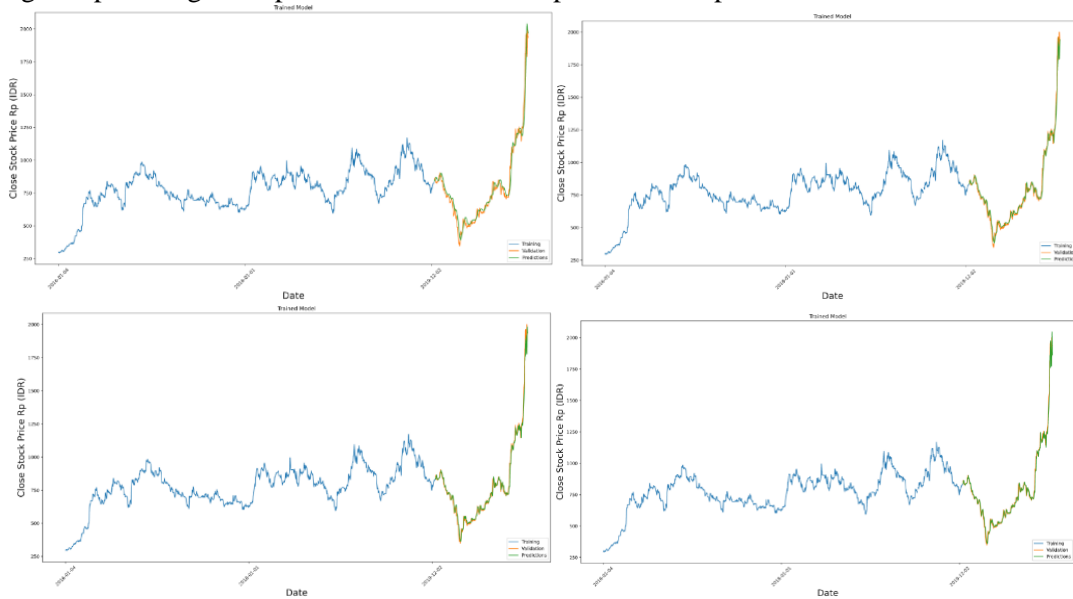
### 3.1 Hasil Prediksi Harga Saham ANTM

Hasil prediksi harga saham ANTM dibagi sesuai dengan jumlah variasi epoch yang dilakukan yaitu 50, 100, 150 dan 200. Pada tabel 1 berikut hasil keseluruhan dari prediksi harga saham ANTM.

Tabel 1. Hasil Prediksi Saham ANTM

Epoch	RMSE	Waktu Komputasi
50	54.84999664170528	20.13856816291809
100	42.3883300193954	40.100128412246704
150	40.64737513030087	54.13262486457825
200	40.14714151597957	66.41139078140259

Berdasarkan hasil pada tabel 1 menunjukkan nilai RMSE terbesar didapat pada variasi epoch 50 dan nilai RMSE terkecil didapat pada variasi epoch 200. Ini menunjukkan bahwa hasil yang didapat mengalami penurunan nilai RMSE pada variasi epoch.



Gambar 5 Grafik Hasil Prediksi Saham ANTM

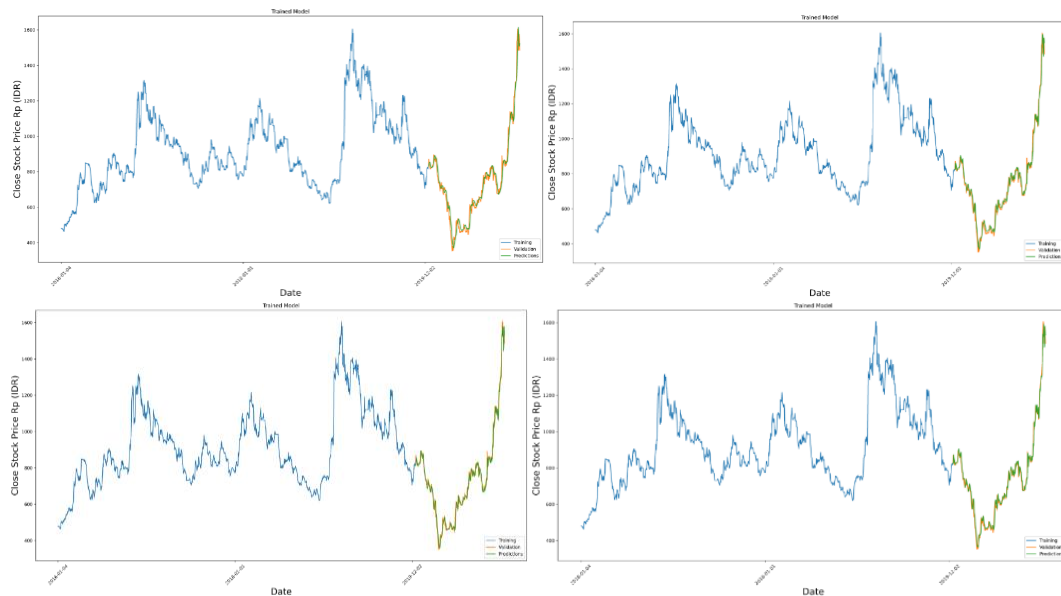
### 3.2 Hasil Prediksi Harga Saham TINS

Prediksi yang dilakukan pada emiten TINS menggunakan variasi epoch 50, 100, 150 dan 200 menghasilkan prediksi pada tabel 2 sebagai berikut.

Tabel 2. Tabel Hasil Prediksi Saham TINS

Epoch	RMSE	Waktu Komputasi
50	40.73122599735746	22.21646213531494
100	33.2015574850126	37.27807855606079
150	32.703005232318745	51.356966495513916
200	31.764224892121756	67.21452713012695

Berdasarkan hasil yang didapat pada tabel 2 menunjukkan bahwa nilai RMSE terbesar didapat pada variasi epoch 50 dan nilai RMSE terkecil didapat pada variasi epoch 200. Dan menunjukkan bahwa hasil yang didapat berdasarkan variasi epoch nilai RMSE semakin mengecil akan tetapi waktu komputasi yang dilakukan semakin lama. Pada gambar 6 menunjukkan bahwa ketepatan yang didapatkan semakin membaik disetiap penambahan jumlah epoch.



Gambar 6. Grafik Prediksi Harga Saham TINS Epoch 50, 100, 150 dan 200

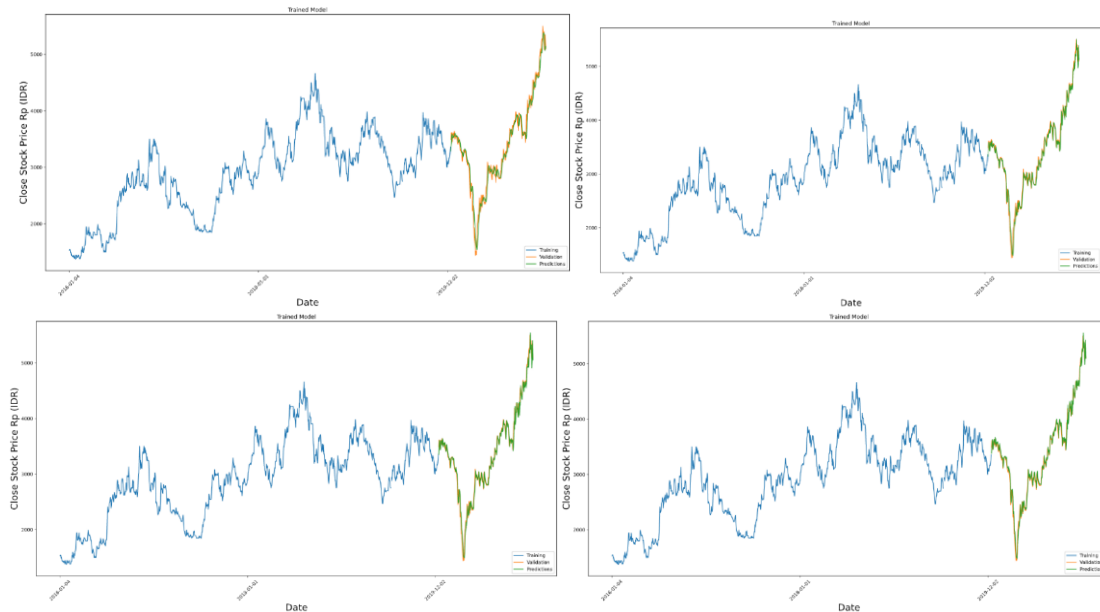
### 3.3 Hasil Prediksi Harga Saham INCO

Hasil prediksi harga saham pada emiten INCO menggunakan variasi epoch sejumlah 50, 100, 150 dan 200 dapat dilihat pada tabel 3 berikut.

Tabel 3 Hasil Prediksi INCO

Epoch	RMSE	Waktu Komputasi
50	139.67756358966105	22.42471408843994
100	117.0294986575377	40.3596875667572
150	114.3784314352723	59.13053059577942
200	111.47021370192789	73.68336915969849

Berdasarkan pada tabel 3 menunjukkan bahwa nilai RMSE terkecil didapat pada variasi epoch 50 dan nilai RMSE terbesar didapat pada variasi epoch 200. Menunjukkan bahwa penambahan jumlah epoch pada prediksi mempengaruhi penurunan nilai RMSE akan tetapi waktu komputasi yang dilakukan semakin besar. Pada gambar 7 dapat dilihat bahwa ketepatan prediksi terhadap data aslinya mengalami kerapatan yang semakin baik.



Gambar7. Grafik Prediksi Harga Saham INCO Epoch 50, 100, 150, 200

### 3.4 Perbandingan dari Hasil Prediksi Emiten Saham

Pada data pengujian yang dihasilkan ditabel 4 terlihat bahwa proses pelatihan dan pengujian prediksi dipengaruhi oleh jumlah epoch, waktu komputasi dan RMSE. Dari ketiga emiten saham yang diprediksi semakin tinggi nilai epoch semakin tinggi waktu komputasi yang dibutuhkan. Untuk waktu komputasi tercepat dengan nilai 20.13 detik di saham emiten ANTM saat nilai epoch sejumlah 50. Sedangkan untuk waktu komputasi terlama pada saat epoch sejumlah 200 di saham emiten INCO dengan waktu komputasi sebesar 73.68 detik.

Tetapi apabila dilihat dari data Root Mean Square Error (RMSE) memiliki nilai yang monoton turun walalupun terdapat penurunan yang sangat tipis. Dalam hal monton turun dapat diartikan bahwa data prediksi yang dihasilkan semakin akurat. Untuk nilai terendah Root Mean Square Error (RMSE) pada data harga saham emiten TINS dengan nilai sebesar 31,76 pada saat nilai epoch 200, sedangkan nilai Root Mean Square Error (RMSE) tertinggi terjadi pada harga saham emiten INCO dengan nilai 139,67 pada saat nilai epoch 50. Semakin tinggi waktu komputasi yang dibutuhkan, maka semakin rendah pula nilai Root Mean Square Error (RMSE) yang dihasilkan.

Tabel 4 Hasil Prediksi Waktu Komputasi dan RMSE

Epoch	ANTM		TINS		INCO	
	Waktu Komputasi	RMSE	Waktu Komputasi	RMSE	Waktu Komputasi	RMSE
50	20.13	54.84	22.21	40.73	22.42	139.67
100	40.10	42.38	37.27	33.20	40.35	117.02
150	54.13	40.64	51.35	32.70	59.13	114.37
200	66.41	40.14	67.21	31.76	73.68	111.47

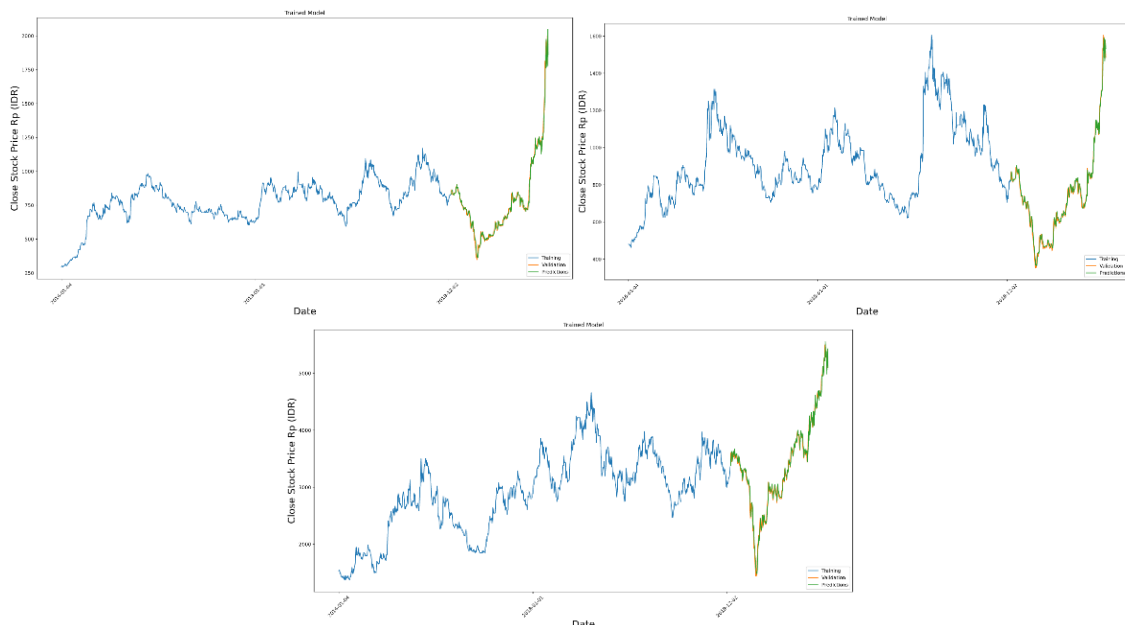
Berdasarkan data pada tabel 4 prediksi menggunakan metode Long Short Term Memory (LSTM) nilai epoch 200 adalah jumlah epoch paling optimal berdasarkan nilai RMSEnya. Dan pada gambar 8, dapat dilihat antara data pengujian dan data prediksi sangat tidak jauh berbeda.



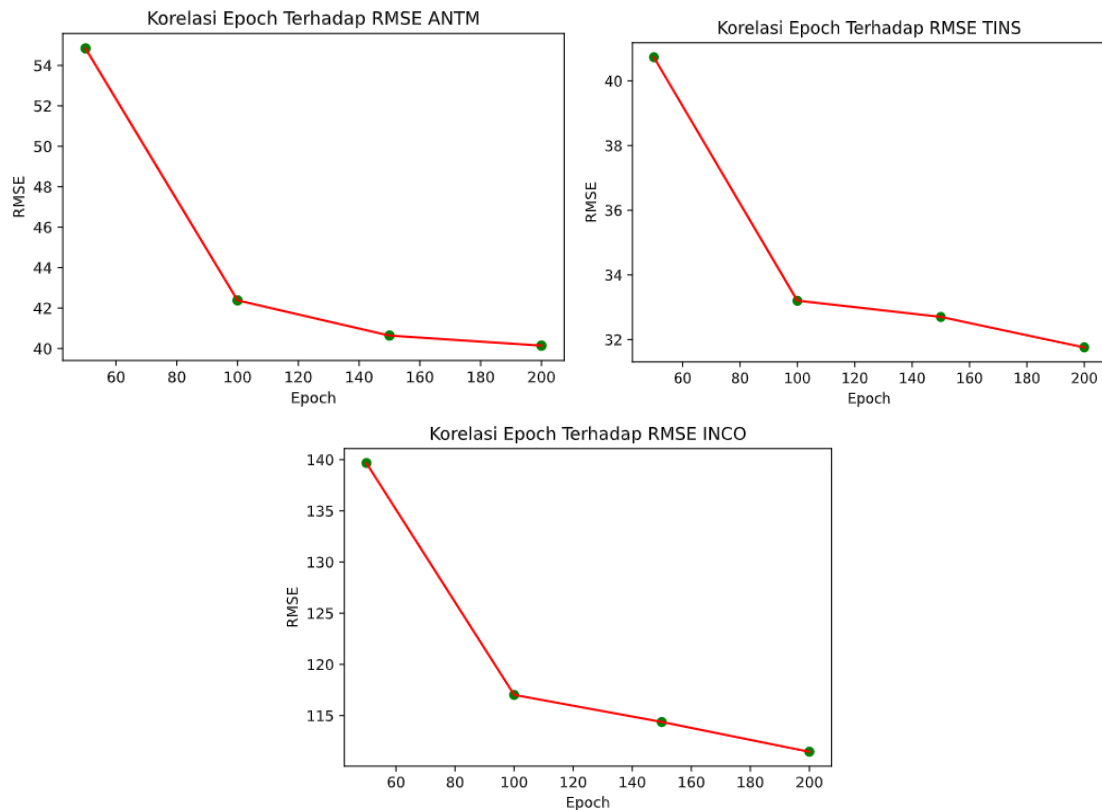
Dimana 20% dari data aktual berwarna oren yaitu data pengujian berjalan dari tahun 2019 akhir sampai 2020 akhir, garis grafik hampir berhimpitan tepat dengan data prediksi berwarna hijau. Sehingga dapat dikatakan bahwa semakin tinggi nilai epoch, semakin tinggi juga tingkat akurasi. Untuk lebih dapat dikatakan bahwa variasi *epoch* mempengaruhi nilai RMSE dilakukan uji korelasi terhadap *epoch* dengan RMSE menghasilkan grafik pada gambar 9. Dari hasil uji korelasi didapatkan bahwa nilai korelasinya mencapai nilai negatif seperti tabel 5 berikut.

Tabel 5 Nilai Korelasi Epoch dan RMSE

No	ANTM	TINS	INCO
Nilai Korelasi	-0,850292	-0,856458	-0,873683



Gambar 8. Data Prediksi Untuk Saham ANTM (Kiri Atas), TINS (Atas Kanan) dan INCO (Tengah Bawah) Dengan Epoch 200



Gambar 9. Grafik Korelasi Epoch dan RMSE

## 5. KESIMPULAN DAN SARAN

Penelitian ini memiliki kesimpulan bahwa nilai *Root Mean Square Error* (RMSE) mengalami penurunan pada setiap peningkatan jumlah epoch dalam proses prediksinya yang dimana dapat dilihat pada uji korelasi yang dihasilkan antara jumlah epoch dan nilai RMSE pada gambar 9 dan tabel 5. Akan tetapi penambahan jumlah epoch juga mempengaruhi waktu komputasi yang dibutuhkan yang dapat dilihat pada tabel 4. Untuk hasil nilai RMSE terkecil didapatkan oleh LSTM pada emiten TINS dimana nilai RMSE terkecil sebesar 31,76 dengan variasi epoch 200 sedangkan hasil nilai RMSE terbesar didapatkan pada emiten INCO dimana nilai RMSE yang didapatkan sebesar 139,67 dengan variasi epoch 50.

1. Penambahan jumlah data yang lebih besar untuk prediksi, dikarenakan pada waktu 5 tahun data yang terkumpul hanya 1261 baris data. Untuk mendapatkan hasil prediksi yang semakin baik.
2. Pengkajian ulang kembali terhadap penentuan jumlah parameter karena kemungkinan besar jumlah parameter berdampak pada hasil prediksi. Pada penelitian ini hanya melakukan pengkajian pada variasi epoch saja.

## DAFTAR PUSTAKA

- [1] Arfan, A. & ETP., Lussina 2019. *Prediksi Saham di Indonesia Menggunakan Algoritma Long Short Term Memory*, SeNTIK.
- [2] Arfan, A. & ETP., Lussina 2020. *Perbandingan Algoritma Long Short-Term Memory dengan SVR pada Prediksi Harga Saham di Indonesia*. LSTM, PETIR Vol.13, No. 1.
- [3] Fauzi, A. 2019. *Forecasting Saham Syariah Dengan Menggunakan LSTM*. Research Article: AlMasraf Jurnal Lembaga Keuangan dan Perbankan.
- [4] Gemilang, FA, 2017, *Prediksi Harga Penutupan Saham Menggunakan Fuzzy Time Series*, [https://repository.usd.ac.id/11870/2/135314089\\_full.pdf](https://repository.usd.ac.id/11870/2/135314089_full.pdf), Diakses 23 Maret 2021.
- [5] *Harga Saham Pertambangan pada Bursa Efek Indonesia*, <https://finance.yahoo.com/>, Diakses 7 Mei 2021.
- [6] Karno, ASB. 2020. *Analisis Data Time Series Menggunakan LSTM (Long Short Term Memory) dan ARIMA (Autocorrelation Integrated Moving Average) Dalam Bahasa Python*, ULTIMA InfoSys, Vol XI, No.1.
- [7] Provost, F. & Fawcett, T. 2013. *Data Science and It Relationship to Big Data and Data Driven Decision Making*. Big Data, 1(1): 51-59.
- [8] Riyantoko, PA., Fahrudin, TM., Maulida, K. & Safitri, HE. 2020. *Analisis Prediksi Harga Saham Sektor Perbankan Menggunakan Algoritma Long Short Term Memory*, SEMNASIF.
- [9] Suyudi, MAD., Djamal, EC., dan Asri Maspupah. 2019. *Prediksi Harga Saham Menggunakan Metode Recurrent Neural Network*. Makalah Disajikan Dalam Seminar Nasional Aplikasi Teknologi Informasi, Yogyakarta.