

BAB V

HASIL DAN PEMBAHASAN

Pada bab tiga ini akan menjelaskan hasil dari perancangan serta pembahasan mengenai hasil yang telah dicapai selama penelitian ini.

5.1 Hasil

Pada subbab ini, dipaparkan hasil dari penelitian yang dimulai dari pra-pemrosesan data seperti seleksi fitur, normalisasi, pembuatan *dataset*. Setelah pra-pemrosesan data telah dicapai, hasil analisa dengan menerapkan pemodelan dari metode *LSTM* akan dipaparkan.

5.1.1 Pra-Pemrosesan Data

Pada subbab ini akan memaparkan hasil dari seleksi fitur dengan melibatkan statistik deskriptif, visualisasi, uji korelasi hubungan antar variabel atau fitur terhadap target, kesimpulan dari seleksi fitur, normalisasi data, dan pembuatan *dataset*. Tabel 5. 1 Tabel statistik deskriptif emas

5.1.1.1. Statistik Deskriptif

5.1.1.1.1. Statistik deskriptif harga emas

Tabel 5. 1 Tabel statistik deskriptif emas

	Harga Pembuka	Harga Terendah	Harga Tertinggi	Harga Close	Volume Harga
Valid	2051,000	2051,000	2051,000	2051,000	2051,000
Missing	0,000	0,000	0,000	0,000	0,000
Median	1766,2	1759,400	1776,000	1767,400	187,000
Rerata	1724,617	1716,623	1732,328	1724,359	5270,296
Std. Dev	403,046	400,091	405,624	402,747	29550,255
Skewness	0.693	0.695	0.686	0.692	7447,000
Kurtosis	0.139	0.136	0.129	0.138	61066,000
Minimum	1160,400	1146,500	1160,800	1155,200	0,000
Maksimum	2947,900	2933,100	2957,900	2949,100	386334,000

Berdasarkan tabel statistik deskriptif harga emas, terlihat bahwa jumlah data valid untuk setiap variabel mencapai 2.051 dan tidak terdapat data yang hilang. Nilai median dari harga pembuka, harga terendah, harga tertinggi, dan harga penutupan berada pada kisaran 1.759 hingga 1.776, yang mencerminkan harga tengah dari data dalam rentang tersebut. *Mean* dari keempat harga tersebut juga berada pada kisaran yang relatif dekat, yaitu sekitar 1.716 hingga 1.732. Jika diperhatikan lebih lanjut, nilai *mean* sedikit lebih rendah dibandingkan median, yang secara umum dapat mengindikasikan distribusi miring ke kiri (*negatively skewed*). Namun, nilai *skewness* yang tercatat justru berada pada kisaran positif (antara 0,686 hingga 0,695), yang menandakan adanya kemiringan ke kanan. Ketidaksesuaian ini dapat terjadi karena bentuk distribusi yang tidak sepenuhnya simetris atau karena pengaruh nilai-nilai ekstrem atau outlier di sisi atas distribusi.

Nilai standar deviasi yang berada di kisaran 400 untuk seluruh harga mengindikasikan adanya variasi harga yang cukup signifikan dari rata-ratanya. Sementara itu, nilai kurtosis yang sangat rendah (sekitar 0,13) menunjukkan bahwa distribusi data cenderung lebih mendatar, yang berarti kemungkinan terjadinya harga ekstrem atau outlier relatif rendah.

Untuk volume harga, terjadi perbedaan yang sangat mencolok antara nilai median (187.000) dan rata-rata (5.270.296), serta nilai *skewness* dan kurtosis yang sangat tinggi masing-masing sebesar 7.447 dan 61.066. Hal ini menunjukkan bahwa distribusi volume harga sangat tidak simetris dan memiliki ekor panjang di sebelah kanan (*positively skewed*), mengindikasikan adanya beberapa hari perdagangan dengan volume transaksi yang sangat besar dibandingkan hari-hari lainnya. Nilai minimum volume adalah 0, yang mengindikasikan adanya hari tanpa transaksi, sementara maksimum mencapai 386.334.000, menunjukkan lonjakan volume yang sangat besar pada titik tertentu.

5.1.1.1.2. Statistik Deskriptif Harga Minyak Mentah

Tabel 5. 2 Tabel statistik deskriptif minyak mentah

	Harga Pembuka	Harga Terendah	Harga Tertinggi	Harga Close	Volume Harga
Valid	2051,000	2051,000	2051,000	2051,000	2051,000

<i>Missing</i>	0,000	0,000	0,000	0,000	0,000
Median	67,540	66,410	68,520	67,500	405596,000
Rerata	66,109	64,956	67,248	66,146	457327,722
<i>Std. Dev</i>	17,702	17,419	17,931	17,661	225152,781
<i>Skewness</i>	0.073	-0.060	0.254	0.135	0.943
Kurtosis	0.964	1,159	0.536	0.638	2611,000
Minimum	-37,630	-40,320	13,690	-14,000	0,000
Maksimum	123,700	120,790	130,500	124,660	2288,000

Berdasarkan tabel statistik deskriptif harga minyak mentah, diketahui bahwa seluruh variabel memiliki 2.051 data valid tanpa adanya nilai yang hilang. *Median* harga pembuka, harga terendah, harga tertinggi, dan harga penutupan berkisar antara 66,41 hingga 68,52, sedangkan nilai rata-rata dari keempat harga tersebut berada sedikit di bawah median, yaitu antara 64,96 hingga 67,25. Hal ini menunjukkan bahwa distribusi data secara umum cenderung sedikit miring ke kiri (*negatively skewed*), meskipun nilai *skewness* yang tercatat sangat kecil (berkisar antara -0,060 hingga 0,254) menunjukkan bahwa distribusi harga relatif simetris atau hanya memiliki kemiringan yang sangat ringan.

Nilai standar deviasi sekitar 17 untuk seluruh harga menunjukkan adanya variasi yang signifikan, meskipun tidak sebesar variabilitas pada harga emas. Seluruh nilai kurtosis berada di bawah 3, termasuk harga terendah yang tercatat sebesar 1,159, yang menunjukkan distribusi datar dengan kemungkinan outlier ekstrem yang rendah. Namun, perhatian khusus perlu diberikan pada nilai minimum negatif, seperti -40,32 untuk harga terendah dan -37,63 untuk harga pembuka, yang secara logis tidak mungkin terjadi pada harga minyak mentah.

Fenomena harga minyak mentah yang turun hingga ke nilai negatif ini tidak selalu mencerminkan harga fisik negatif, melainkan lebih berkaitan dengan mekanisme pasar berjangka dan kondisi ekstrem mendasar. Berdasarkan artikel *Energy Information Administration (EIA)* dari AS yang ditulis oleh (French, 2021), pada April 2020 terjadi kondisi pasar yang luar biasa: permintaan minyak turun drastis akibat pembatasan mobilitas selama pandemi COVID-19, sementara

pasokan tetap tinggi dan kapasitas penyimpanan sudah penuh. Akibatnya, kontrak berjangka minyak seperti *West Texas Intermediate (WTI)* sempat diperdagangkan di bawah nol bahkan sempat mencapai $-37,63$ USD per barel. Hal ini terjadi karena para pemegang kontrak yang hampir jatuh tempo tidak ingin mengambil pengiriman fisik jika tidak ada tempat untuk menyimpannya, sehingga mereka bersedia membayar agar kontrak tersebut diambil alih orang lain. Pada titik itu, biaya penyimpanan dan potensi biaya *shutdown* produksi menjadi lebih mahal dibandingkan menjual kontrak berjangkanya dengan harga negatif.

Dengan demikian, penurunan harga hingga negatif dalam data deskriptif kemungkinan bersumber dari perilaku pasar kontrak berjangka yang terjadi dalam periode ekstrem tersebut, bukan perwakilan harga riil minyak fisik. Oleh karena itu, data ini harus ditandai sebagai anomali namun valid. Untuk volume harga, median tercatat sebesar 405.596, sementara rata-ratanya lebih tinggi yaitu 457.327, mengindikasikan sedikit kemiringan ke kanan (*positively skewed*), yang juga didukung oleh nilai *skewness* sebesar 0,943. Nilai kurtosis volume yang sangat tinggi (2611) menunjukkan adanya outlier signifikan, yaitu hari-hari dengan volume transaksi yang jauh lebih besar dibandingkan hari-hari lainnya. Nilai minimum volume adalah 0, dan maksimum mencapai 2.288.000, menunjukkan rentang yang sangat lebar.

5.1.1.1.3. Statistik Deskriptif Indeks Harga S&P 500

Tabel 5. 3 Tabel deskriptif S&P 500

	Harga Pembuka	Harga Terendah	Harga Tertinggi	Harga Close	Volume Harga
Valid	2051,000	2051,000	2051,000	2051,000	2051,000
Missing	0,000	0,000	0,000	0,000	0,000
Median	3727,750	3686,250	3774,750	3725,000	1475,000
Rerata	3727,765	3698,823	3751,914	3725,848	1553,000
Std. Dev	1028,804	1023,365	1033,926	1029,119	559545,450
Skewness	0.513	0.521	0.503	0.512	1195,000
Kurtosis	-0.678	-0.671	-0.689	-0.678	4037,000
Minimum	2220,500	2174,000	2259,500	2220,250	0,000

Maksimum	6163,000	6129,250	6166,500	6153,750	5715,000
----------	----------	----------	----------	----------	----------

Berdasarkan tabel statistik deskriptif indeks S&P 500, terlihat bahwa jumlah data valid untuk setiap variabel mencapai 2.051 dan tidak terdapat data yang hilang (missing). Nilai median dari harga pembuka, harga terendah, harga tertinggi, dan harga penutupan berada pada kisaran 3.686 hingga 3.775, yang mencerminkan posisi tengah dari distribusi data. Rata-rata (*mean*) dari keempat harga tersebut juga berada dalam kisaran yang sangat dekat, yaitu sekitar 3.698 hingga 3.751. Jika diperhatikan lebih lanjut, nilai *mean* pada semua kategori harga cenderung sedikit lebih tinggi dibandingkan median. Hal ini secara umum mengindikasikan adanya distribusi miring ke kanan (*positively skewed*).

Kondisi ini juga dikonfirmasi oleh nilai *skewness* yang seluruhnya berada di atas nol (sekitar 0,503 hingga 0,521), yang menunjukkan adanya kemiringan ke kanan, artinya terdapat sebagian kecil nilai-nilai harga yang lebih tinggi dari rata-rata, sehingga menarik ekor distribusi ke arah kanan. Nilai standar deviasi yang berada di kisaran 1.023 hingga 1.033 menunjukkan adanya variabilitas harga yang cukup besar, mencerminkan fluktuasi harga indeks S&P 500 dalam periode pengamatan yang signifikan. Meskipun besar, variasi ini dapat dianggap wajar mengingat tingkat harga yang memang berada dalam ribuan poin. Nilai kurtosis untuk seluruh kategori harga adalah negatif (sekitar -0,67 hingga -0,68), yang menunjukkan bahwa distribusi data cenderung lebih mendatar dibandingkan distribusi normal. Artinya, data harga tidak memiliki puncak yang tajam dan probabilitas terjadinya nilai ekstrem atau outlier cenderung rendah.

Untuk volume harga, median tercatat sebesar 1.475 dan *mean* sedikit lebih tinggi sebesar 1.553, yang menunjukkan kemiringan positif yang ringan, sesuai dengan nilai *skewness* sebesar 1.195. Sementara itu, kurtosis volume sangat tinggi, yakni mencapai 4.037, yang mengindikasikan adanya outlier atau lonjakan volume ekstrem dalam beberapa hari perdagangan. Hal ini diperkuat dengan perbedaan besar antara nilai minimum (0) dan maksimum (5.715.000), menunjukkan adanya hari-hari tanpa transaksi sama sekali, serta hari-hari dengan volume transaksi yang sangat tinggi.

5.1.1.1.4. Statistik Deskriptif Harga Perak

Tabel 5. 4 Tabel statistik deskriptif harga perak

	Harga Pembuka	Harga Terendah	Harga Tertinggi	Harga Close	Volume Harga
Valid	2051,000	2051,000	2051,000	2051,000	2051,000
Missing	0,000	0,000	0,000	0,000	0,000
Median	21,533	21,355	21,675	21,510	45,000
Rerata	21,220	21,087	21,364	21,231	1821,461
Std. Dev	5,020	4,961	5,078	5,021	10554,935
Skewness	0.417	0.418	0.418	0.421	7596,000
Kurtosis	-0.844	-0.834	-0.849	-0.838	63080,000
Minimum	11,735	11,735	12,205	12,070	0,000
Maksimum	34,831	34,425	34,835	34,555	131415,000

Berdasarkan tabel statistik deskriptif harga perak, seluruh variabel memiliki jumlah data valid sebanyak 2.051 dan tidak terdapat data yang hilang. Median dari harga pembuka, harga terendah, harga tertinggi, dan harga penutupan berada pada kisaran 21,355 hingga 21,675, yang mencerminkan nilai tengah dari distribusi harga. Sementara itu, nilai rata-rata (*mean*) sedikit lebih rendah dibandingkan median, yaitu antara 21,087 hingga 21,364. Pola ini biasanya mengindikasikan adanya kemiringan ke kiri (*negatively skewed*), namun dalam tabel terlihat bahwa nilai skewness untuk keempat harga berada di kisaran positif (sekitar 0,417 hingga 0,421), yang berarti secara statistik, distribusi memiliki kemiringan ke kanan (*positively skewed*). Ini menunjukkan bahwa meskipun secara visual tampak seimbang antara mean dan median, masih terdapat beberapa nilai tinggi yang mendorong ekor distribusi ke kanan.

Nilai standar deviasi berada di kisaran 4,961 hingga 5,078, yang menunjukkan adanya tingkat variasi harga yang cukup moderat jika dibandingkan dengan nilai rata-ratanya. Ini berarti harga perak relatif stabil dalam kisaran yang

tidak terlalu ekstrem, setidaknya dalam periode pengamatan. Nilai kurtosis seluruhnya bernilai negatif (sekitar $-0,83$ hingga $-0,84$), menunjukkan bahwa distribusi harga cenderung lebih mendatar dibandingkan distribusi normal. Artinya, kemungkinan terjadinya nilai ekstrem atau outlier relatif rendah dan distribusinya memiliki puncak yang lebar.

Sementara itu, untuk volume harga, terdapat perbedaan yang sangat mencolok antara median (45.000) dan rata-rata (1.821.461), serta nilai *skewness* dan kurtosis yang sangat tinggi masing-masing sebesar 7.596 dan 63.080. Hal ini menunjukkan bahwa distribusi volume perdagangan sangat tidak simetris dan memiliki ekor panjang ke kanan (*positively skewed*), yang menunjukkan adanya beberapa hari dengan volume transaksi yang jauh lebih besar dibandingkan hari-hari lainnya. Nilai minimum volume adalah 0, menunjukkan kemungkinan hari tanpa transaksi, sedangkan nilai maksimum mencapai 131.415.000, yang memperkuat indikasi adanya lonjakan besar dalam volume di waktu tertentu.

5.1.1.1.5. Statistik deskriptif indeks dolar AS

Tabel 5. 5 Tabel statistik deskriptif indeks harga dolar AS

	Harga Pembuka	Harga Terendah	Harga Tertinggi	Harga Close	Volume Harga
Valid	2051,000	2051,000	2051,000	2051,000	2051,000
Missing	0,000	0,000	0,000	0,000	0,000
Median	97,469	97,230	97,720	97,490	19691,000
Rerata	98,509	98,207	98,804	98,503	21945,578
Std. Dev	5,405	5,338	5,457	5,400	10879,227
Skewness	0.308	0.286	0.330	0.308	1850,000
Kurtosis	-0.723	-0.753	-0.683	-0.716	7384,000
Minimum	88,505	88,150	88,890	88,490	0,000
Maksimum	114,047	113,260	114,745	114,105	113489,000

Berdasarkan tabel statistik deskriptif indeks dolar AS (USDIX), seluruh variabel memiliki jumlah data valid sebanyak 2.051 dan tidak terdapat data yang hilang, menunjukkan kelengkapan data yang baik. Nilai median untuk harga pembuka, harga terendah, harga tertinggi, dan harga penutupan berada pada kisaran

97,230 hingga 97,720, sementara rata-rata (mean) dari keempat harga tersebut berada pada kisaran 98,207 hingga 98,804. Karena rata-rata lebih tinggi daripada median di seluruh variabel harga, hal ini menunjukkan bahwa distribusi data cenderung miring ke kanan (*positively skewed*). Informasi ini diperkuat oleh nilai *skewness* positif di semua harga (sekitar 0,286 hingga 0,330), yang menandakan adanya ekor distribusi yang lebih panjang di sisi kanan. Ini berarti ada sejumlah nilai tinggi yang mungkin menjadi outlier atau hari dengan penguatan nilai indeks secara signifikan.

Nilai standar deviasi berada di kisaran 5,338 hingga 5,457, yang menunjukkan tingkat variasi harga yang relatif moderat dalam konteks nilai indeks dolar. Artinya, meskipun terdapat fluktuasi, indeks ini tergolong cukup stabil dalam periode yang dianalisis. Nilai kurtosis berada pada kisaran negatif (sekitar -0,683 hingga -0,753), menunjukkan bahwa distribusi data lebih mendatar dari distribusi normal. Ini mengindikasikan bahwa data harga memiliki kemungkinan lebih kecil terhadap outlier dan lebih tersebar merata di sekitar rata-ratanya.

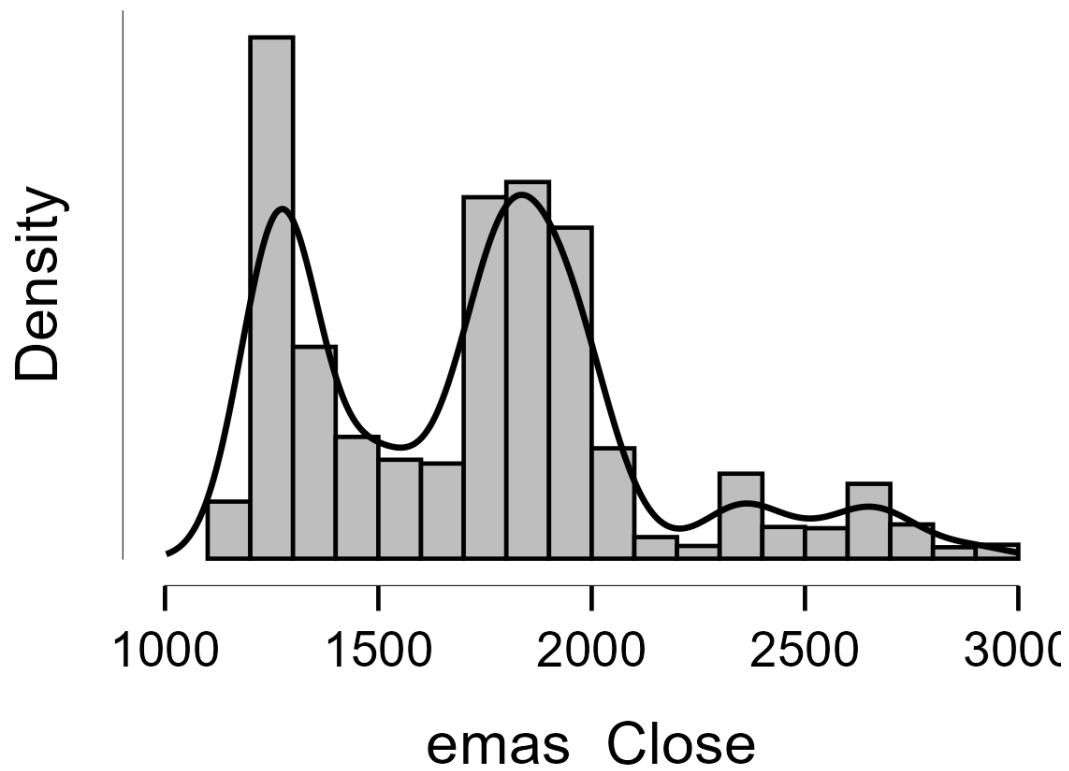
Untuk volume harga, terdapat perbedaan mencolok antara median (19.691) dan rata-rata (21.945,578), disertai nilai *skewness* dan kurtosis yang sangat tinggi masing-masing sebesar 1.850 dan 7.384. Hal ini menunjukkan bahwa distribusi volume perdagangan sangat tidak simetris dan memiliki ekor panjang di kanan (*positively skewed*), mengindikasikan adanya hari-hari tertentu dengan volume transaksi yang jauh lebih tinggi dibandingkan rata-rata. Nilai minimum volume adalah 0, yang bisa menunjukkan tidak adanya transaksi pada hari tertentu, sementara nilai maksimumnya mencapai 113.489, mengindikasikan lonjakan volume yang sangat tinggi.

5.1.1.2. Visualisasi *Plotting*

Untuk memahami karakteristik distribusi data secara visual, dilakukan eksplorasi menggunakan tiga jenis metode visualisasi, yaitu *distribution plot*, *boxplot*, dan *scatter plot*. Ketiga metode ini dipilih karena mampu memberikan gambaran menyeluruh mengenai pola penyebaran, potensi outlier, dan hubungan antar variabel secara intuitif.

Dalam visualisasi ini, digunakan harga penutupan sebagai representasi utama dari masing-masing komoditas dan indeks, yaitu harga emas, harga perak, harga minyak mentah, indeks S&P 500, dan indeks dolar AS (USDIX).

5.1.1.2.1. *Distribution Plot atau Grafik Distribusi*

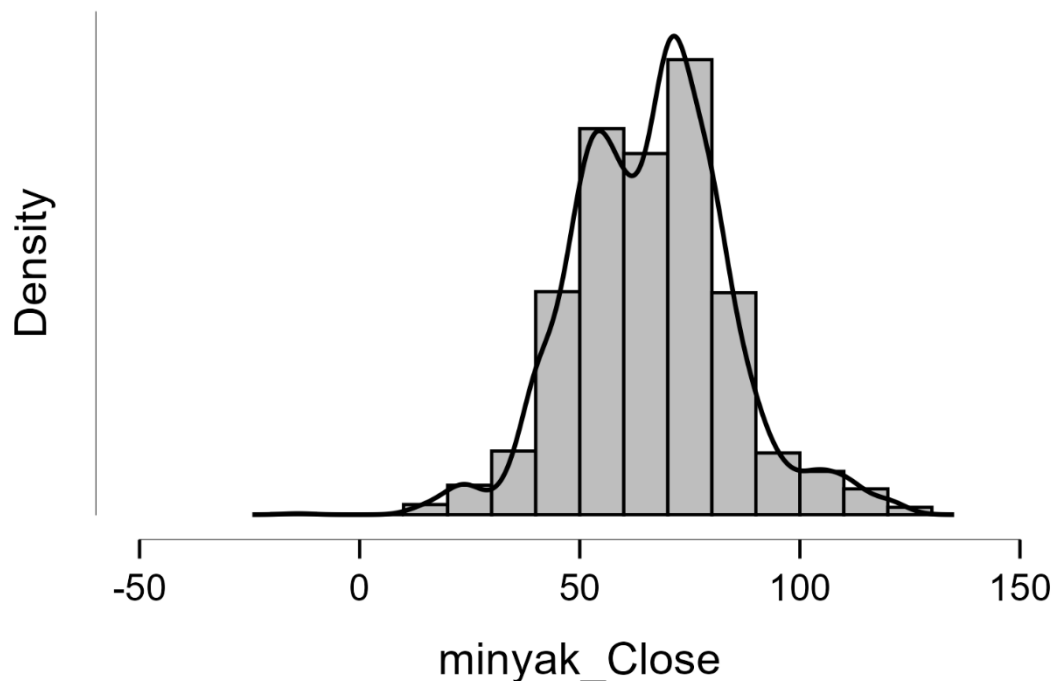


Gambar 5. 1 Gambar visualisasi *distribution plot* harga penutupan emas

Berdasarkan grafik distribusi harga emas_Close menunjukkan pola distribusi yang multimodal, dengan beberapa puncak yang terlihat pada kisaran harga sekitar 1250, 1850, dan indikasi lainnya di atas 2500. Pola ini selaras dengan hasil statistik deskriptif, yang menunjukkan bahwa nilai median harga emas berada di kisaran 1767,400, sementara *mean* sedikit lebih rendah, yaitu sekitar 1724,359. Perbedaan antara *mean* dan median ini umumnya mengarah pada interpretasi distribusi miring ke kiri (*negatively skewed*), namun nilai *skewness* positif yang tercatat 0.692 justru mengonfirmasi bahwa distribusi miring ke kanan (*positively skewed*). Hal ini juga tampak jelas dalam plot distribusi, di mana terdapat ekor

panjang di sisi kanan (harga di atas 2000), yang mengindikasikan adanya nilai-nilai ekstrem atau outlier atau fase lonjakan harga yang signifikan.

Rentang harga yang tercakup dalam distribusi, yakni dari 1155,20 dan 2949,10, mencerminkan tingkat volatilitas yang tinggi, yang juga diperkuat oleh nilai standar deviasi yang cukup besar 402,75 dalam analisis deskriptif. Selain itu, kepadatan data tertinggi yang muncul di kisaran 1250 sampai 1350 dan 1800 sampai 1900 mencerminkan dua fase harga emas yang paling sering terjadi, yang kemungkinan besar berkaitan dengan dua periode historis dengan karakteristik pasar berbeda. Sementara itu, nilai kurtosis yang rendah sekitar 0,13, menandakan bahwa distribusi data cenderung lebih mendatar dibanding distribusi normal, sehingga kemungkinan terjadinya lonjakan ekstrem sebenarnya relatif rendah, meskipun tetap terlihat pada sebagian kecil data.



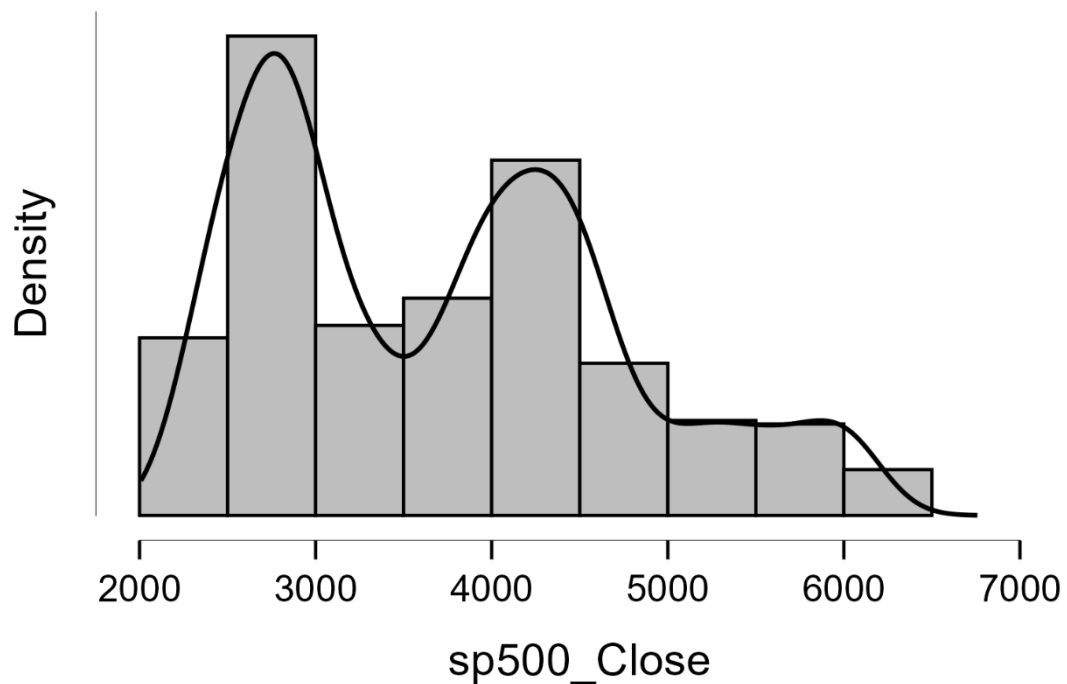
Gambar 5. 2 Gambar visualisasi *distribution plot* harga penutupan minyak mentah

Berdasarkan grafik distribusi harga penutupan minyak mentah atau minyak_close yang mencakup histogram dengan tumpang-tindih kepadatan (*density plot*), terlihat bahwa distribusi memiliki pola unimodal dengan satu puncak utama di sekitar harga 75,00. Kepadatan harga tertinggi terjadi pada kisaran ini, yang menunjukkan bahwa harga minyak paling sering berada dalam rentang

tersebut selama periode analisis. Ini selaras dengan statistik deskriptif, di mana median harga tercatat sebesar 67,50, sementara rata-rata sedikit lebih rendah di 66,15, menandakan bahwa sebagian besar nilai harga berpusat pada rentang tersebut dengan beberapa ekstrem yang menarik distribusi.

Kepadatan distribusi juga memperkuat temuan mengenai kemiringan data yang ringan ke kiri (*negatively skewed*), meskipun nilai *skewness* yang sangat kecil (0,135) menunjukkan bahwa distribusi relatif simetris. Dalam *density plot*, terdapat ekor panjang di sisi kiri, yang menunjukkan adanya sejumlah harga ekstrem di bawah 0, yang juga tercermin dalam nilai minimum yang tercatat -14,00. Keberadaan ekor ini menunjukkan bahwa walaupun sebagian besar data terkonsentrasi di sekitar harga 75, beberapa nilai rendah tetap berkontribusi pada persebaran distribusi.

Volatilitas harga minyak juga terlihat dalam kepadatan distribusi, di mana rentang harga -14 hingga 124,6 memperlihatkan variasi yang cukup signifikan, diperkuat oleh nilai standar deviasi sebesar 17,66. Grafik distribusi menunjukkan bentuk distribusi yang lebih mendatar dibandingkan distribusi normal, sesuai dengan nilai kurtosis rendah di nilai 1,16, yang mengindikasikan bahwa probabilitas munculnya lonjakan harga ekstrem tetap ada, tetapi tidak dominan. Secara keseluruhan, kepadatan distribusi memberikan gambaran tambahan bahwa harga minyak memiliki pola persebaran yang cukup stabil tetapi tetap menunjukkan beberapa titik ekstrem yang memengaruhi pergerakan harga secara keseluruhan.



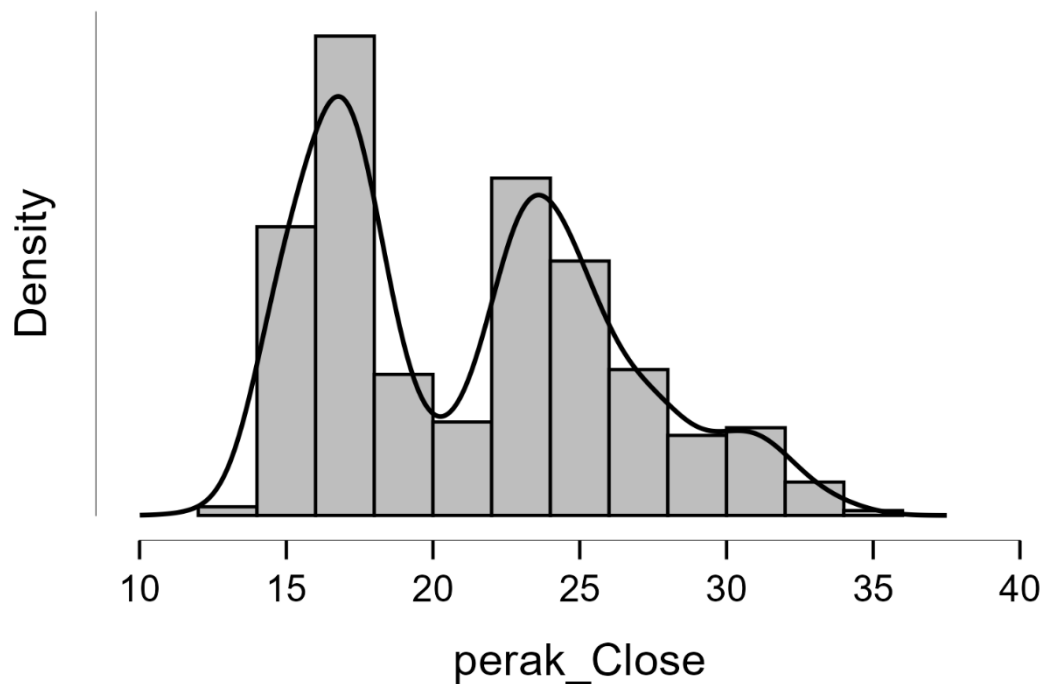
Gambar 5. 3 Gambar visualisasi *distribution plot* harga penutupan indeks S&P 500

Berdasarkan grafik distribusi harga penutupan indeks S&P 500, distribusi menunjukkan pola unimodal dengan satu puncak utama yang berada di kisaran harga sekitar 2500 sampai 4250 , di mana kepadatan tertinggi terjadi. Kepadatan yang tumpang-tindih pada grafik distribusi memperlihatkan bahwa mayoritas nilai harga terkonsentrasi di sekitar puncak ini, sedangkan ekor panjang di sisi kanan menunjukkan adanya sejumlah harga tinggi yang lebih jarang terjadi namun tetap memengaruhi persebaran distribusi.

Kepadatan distribusi memperkuat interpretasi bahwa indeks S&P 500 memiliki kemiringan ke kanan (*positively skewed*), dengan nilai *skewness* positif antara 0.512 yang menunjukkan bahwa sebagian kecil harga yang lebih tinggi dari rata-rata menarik ekor distribusi ke arah kanan. Nilai median harga penutupan 3725,00, yang sedikit lebih rendah dibandingkan rata-rata 3725,85, juga menjadi indikasi bahwa distribusi harga lebih berat di sisi kiri dibandingkan kanan.

Selain itu, nilai standar deviasi 1029,12 mengonfirmasi bahwa indeks ini mengalami variabilitas harga yang cukup besar. Namun, tingkat variabilitas ini tetap dalam batas wajar untuk indeks pasar yang memang bernilai tinggi. Kepadatan distribusi juga menunjukkan bahwa probabilitas terjadinya lonjakan harga ekstrem

relatif rendah, sejalan dengan nilai kurtosis negatif $-0,678$, yang mengindikasikan bahwa bentuk distribusi lebih mendatar dibandingkan distribusi normal. Secara keseluruhan, kepadatan distribusi mempertegas pola pergerakan indeks S&P 500 yang stabil namun tetap mengalami beberapa harga tinggi yang memengaruhi kemiringan distribusi ke kanan.

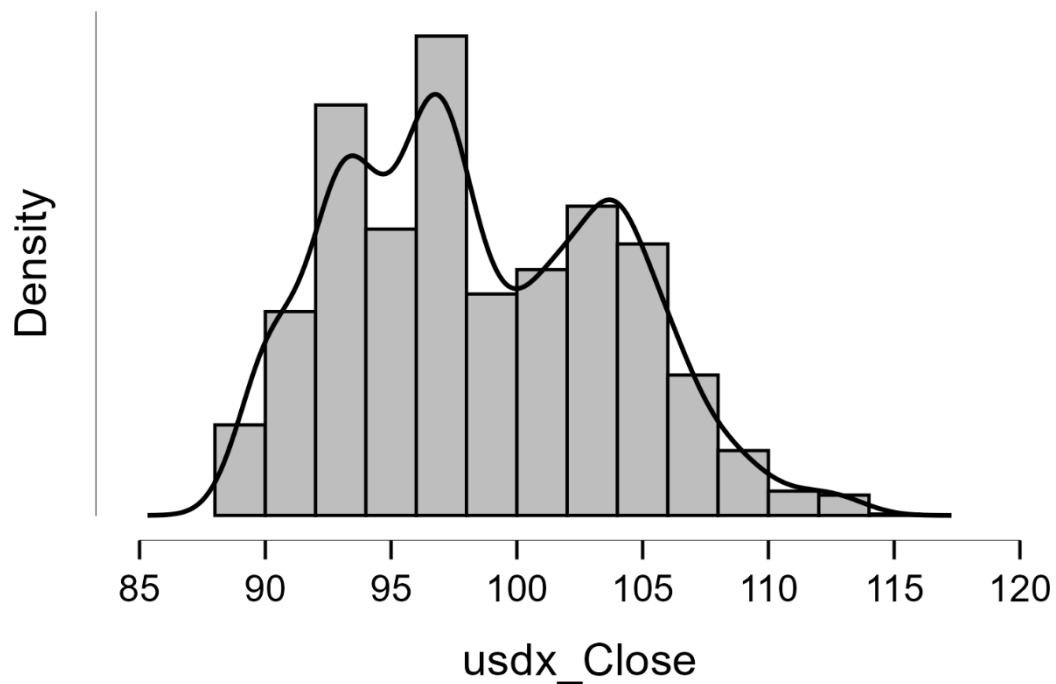


Gambar 5. 4 Gambar visualisasi *distribution plot* harga perak

Berdasarkan grafik distribusi harga penutupan perak, pola distribusi menunjukkan karakteristik bimodal, dengan dua puncak utama yang terlihat di kisaran 17 dan 24, yang mencerminkan adanya dua kelompok harga dominan dalam periode pengamatan. Kepadatan tertinggi terjadi di sekitar kedua puncak ini, menandakan bahwa harga perak paling sering berada dalam dua rentang yang berbeda, yang mungkin dipengaruhi oleh kondisi pasar tertentu dalam waktu yang berbeda. Distribusi ini selaras dengan hasil statistik deskriptif, di mana median harga penutupan tercatat di 21,51, sementara nilai rata-rata sedikit lebih rendah, yaitu sekitar 21,23. Perbedaan kecil antara *mean* dan median ini biasanya mengarah pada interpretasi distribusi miring ke kiri (*negatively skewed*), tetapi nilai *skewness* positif (sekitar 0,417 hingga 0,421) mengonfirmasi bahwa distribusi sebenarnya miring ke kanan (*positively skewed*).

Kepadatan yang tumpang-tindih pada grafik distribusi memperjelas bahwa meskipun mayoritas harga terkonsentrasi di sekitar dua puncak utama, terdapat ekor panjang di sisi kanan, yang menunjukkan beberapa harga ekstrem yang lebih tinggi dari rata-rata. Ini mengindikasikan adanya peristiwa tertentu yang menyebabkan lonjakan harga di luar rentang normalnya. Selain itu, nilai standar deviasi sekitar 5,02 menunjukkan bahwa variasi harga cukup moderat, mengisyaratkan stabilitas relatif dalam pergerakan harga perak, meskipun tetap menunjukkan fluktuasi dalam kisaran yang signifikan. Nilai kurtosis negatif -0,838 mempertegas bahwa distribusi data relatif mendatar dibandingkan distribusi normal, sehingga probabilitas terjadinya lonjakan harga ekstrem lebih rendah dan harga perak cenderung memiliki distribusi yang lebih merata tanpa puncak tajam.

Secara keseluruhan, kepadatan distribusi semakin memperjelas pola bimodal dengan kecenderungan miring ke kanan, serta menunjukkan dua kelompok harga yang paling sering muncul. Fluktuasi harga cukup stabil tetapi tetap menunjukkan adanya periode tertentu di mana harga perak mengalami pergerakan signifikan, terutama di sisi harga yang lebih tinggi.



Gambar 5. 5 Gambar visualisasi distribution plot harga indeks dolar AS

Berdasarkan grafik distribusi harga penutupan indeks dolar AS, pola distribusi menunjukkan kemiringan ke kanan (*positively skewed*), dengan ekor panjang di sisi kanan yang mengindikasikan adanya beberapa nilai tinggi yang lebih jarang terjadi tetapi signifikan. Grafik kepadatan memperlihatkan pola distribusi yang lebih halus dibandingkan histogram, dengan puncak utama yang menunjukkan frekuensi tertinggi dari pergerakan harga indeks. Kepadatan tertinggi tampak pada kisaran 95–100.

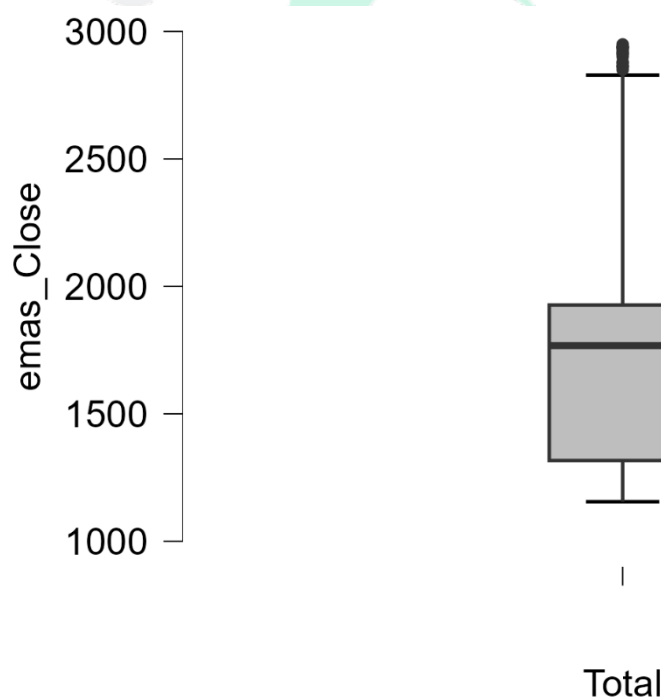
Distribusi ini sejalan dengan hasil statistik deskriptif, di mana median harga penutupan berada di 97,49, sedangkan rata-rata lebih tinggi, yakni 98,50, yang mengonfirmasi adanya kemiringan ke kanan. Grafik kepadatan juga memperjelas keberadaan ekor panjang di sisi kanan, yang menunjukkan beberapa nilai harga ekstrem yang lebih tinggi dari rata-rata dan dapat dianggap sebagai outlier, yang mendorong distribusi ke arah kanan.

Nilai *skewness* positif sebesar 0,308 semakin mempertegas bahwa distribusi lebih berat di sisi kiri, dengan sejumlah harga tinggi menarik ekor distribusi ke kanan. Dalam grafik kepadatan, bentuk yang tidak sepenuhnya simetris semakin menguatkan karakteristik ini. Sementara itu, nilai standar deviasi 5,40 menunjukkan

bahwa variabilitas harga cukup moderat dibandingkan dengan tingkat harga indeks, menandakan bahwa fluktuasi harga masih dalam rentang yang wajar. Selain itu, nilai kurtosis negatif (-0,716) mengindikasikan bahwa distribusi data lebih mendatar dibandingkan distribusi normal, sehingga probabilitas terjadinya lonjakan harga yang sangat ekstrem tetap rendah.

Secara keseluruhan, grafik kepadatan semakin menegaskan pola distribusi indeks dolar AS yang menunjukkan kemiringan ke kanan dengan stabilitas relatif, di mana beberapa nilai harga tinggi menjadi faktor utama dalam pergerakan indeks. Bentuk distribusi yang terstruktur ini mencerminkan fluktuasi yang terjadi dalam periode pengamatan tanpa indikasi perubahan yang sangat tajam.

5.1.1.2.2. Boxplot Atau Grafik Kotak-Garis



Gambar 5. 6 Gambar visualisasi *boxplot* pada harga penutupan emas

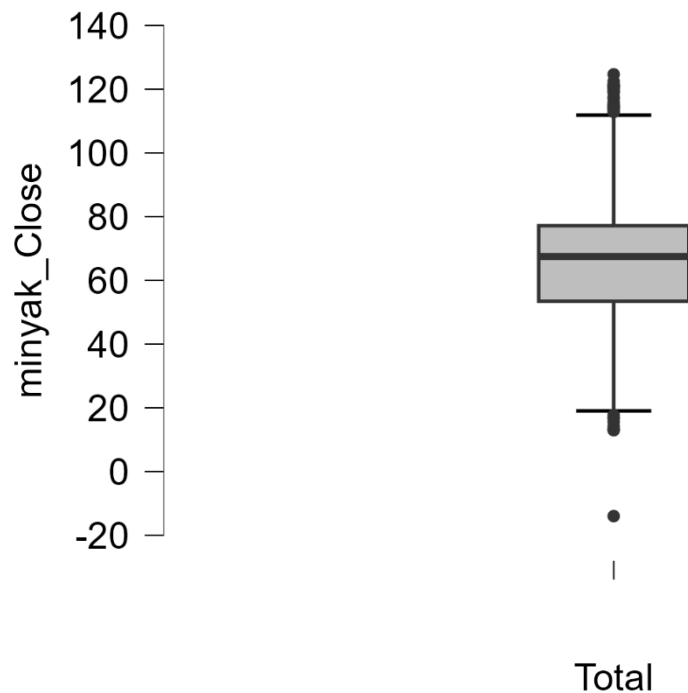
Boxplot harga penutupan emas memberikan wawasan visual mengenai distribusi harga yang telah dianalisis sebelumnya. Median harga emas tercatat berada di kisaran 1767,4, dengan rata-rata sedikit lebih rendah sekitar 1724,359, yang biasanya menunjukkan distribusi miring ke kiri. Namun, dalam analisis statistik, nilai *skewness* positif di 0.692 justru mengonfirmasi bahwa distribusi ini

sebenarnya miring ke kanan (positively skewed), dengan ekor panjang yang terlihat di sisi kanan.

Keberadaan outlier di atas 2000 dalam *boxplot* semakin memperkuat karakteristik *skewness* positif yang telah diidentifikasi dalam distribusi harga. Outlier ini mencerminkan fase lonjakan harga yang signifikan, di mana beberapa periode mengalami kenaikan harga yang lebih tinggi dibandingkan mayoritas data lainnya. Selain itu, *whisker* yang cukup panjang terutama ke arah atas menegaskan adanya fluktuasi harga yang cukup tinggi. Dengan rentang harga yang luas, yaitu antara 1155,20 dan 2949,10, volatilitas harga emas tampak jelas dan diperkuat dengan nilai standar deviasi yang cukup besar sekitar 402,75. Ini menunjukkan bahwa pergerakan harga dalam periode pengamatan cukup dinamis dan mengalami perubahan signifikan dari waktu ke waktu.

Kepadatan data tertinggi yang ditemukan di kisaran 1250 sampai 1350 dan 1800 sampai 1900 dalam distribusi menandakan bahwa terdapat dua fase harga yang paling sering terjadi, yang kemungkinan besar berhubungan dengan peristiwa historis tertentu yang mempengaruhi pasar emas. Sementara itu, nilai kurtosis yang rendah (sekitar 0,13) menunjukkan bahwa distribusi data lebih mendatar dibandingkan distribusi normal, sehingga probabilitas lonjakan harga yang sangat ekstrem sebenarnya relatif rendah, meskipun tetap terlihat dalam beberapa outlier yang berada di sisi kanan distribusi.

Secara keseluruhan, *boxplot* ini mengonfirmasi hasil analisis statistik sebelumnya dengan menunjukkan pola distribusi yang luas, volatilitas tinggi, dan kecenderungan harga ekstrem yang berada di sisi kanan distribusi.



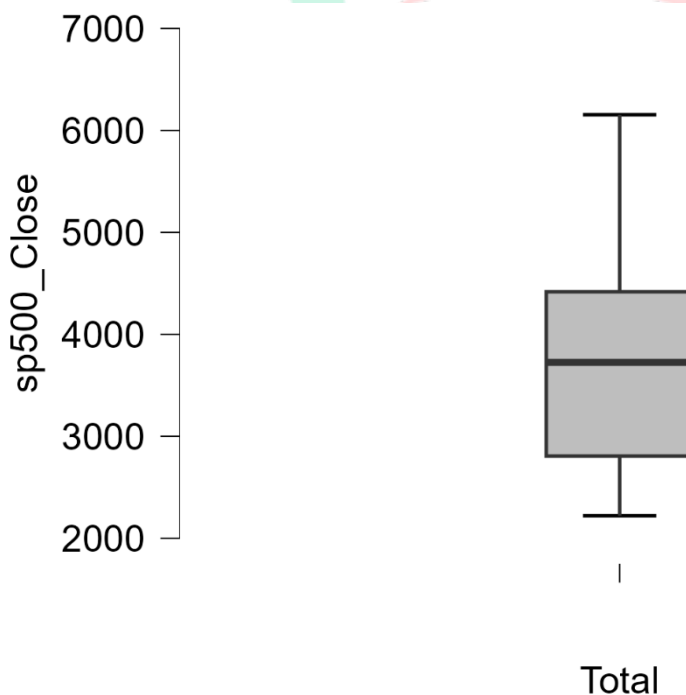
Gambar 5. 7 Gambar visualisasi *boxplot* pada harga penutupan minyak mentah

Boxplot harga penutupan minyak mentah memberikan gambaran visual yang memperkuat analisis distribusi sebelumnya, dengan menunjukkan persebaran harga yang relatif luas serta beberapa nilai ekstrem yang signifikan. Median harga berada di sekitar 70, yang selaras dengan temuan sebelumnya bahwa harga minyak paling sering berada dalam kisaran sekitar 67,50–75,00. Posisi median dalam *boxplot* terlihat lebih dekat ke bagian bawah kotak, yang menandakan bahwa sebagian besar harga cenderung berada di rentang bawah, meskipun ada beberapa nilai tinggi yang memperpanjang distribusi ke sisi kanan.

Rentang *IQR* dalam *boxplot* menunjukkan bahwa mayoritas harga minyak terkonsentrasi antara 60 hingga 80, mencerminkan variabilitas yang cukup stabil dalam periode analisis. Namun, garis *whisker* yang panjang ke arah bawah serta keberadaan outlier di bawah 0 semakin mempertegas bahwa terdapat beberapa harga ekstrem yang mendorong distribusi ke sisi kiri. Hal ini konsisten dengan distribusi sebelumnya, di mana ekor panjang ke kiri terlihat dalam plot kepadatan, menunjukkan bahwa walaupun sebagian besar harga berada di kisaran stabil, terdapat beberapa titik harga yang berada dalam rentang sangat rendah.

Di sisi lain, outlier di atas 120 memperkuat identifikasi bahwa meskipun distribusi secara umum cukup stabil, terdapat sejumlah lonjakan harga ekstrem yang terjadi pada periode tertentu. Ini selaras dengan nilai standar deviasi sebesar 17,66, yang menunjukkan bahwa harga minyak mengalami fluktuasi cukup besar, tetapi tetap dalam batas yang wajar. Selain itu, *boxplot* mengonfirmasi bahwa distribusi harga lebih mendatar dibandingkan distribusi normal, sebagaimana tercermin dalam nilai kurtosis yang rendah (sekitar 1,16), menandakan bahwa probabilitas munculnya lonjakan harga ekstrem tetap ada, tetapi tidak dominan.

Secara keseluruhan, *boxplot* harga minyak mentah menegaskan hasil analisis distribusi sebelumnya, dengan menunjukkan *skewness* ringan ke kiri, volatilitas harga yang cukup tinggi, dan keberadaan beberapa outlier ekstrem. Bentuk distribusi ini mencerminkan pola harga minyak yang cenderung stabil dalam rentang tertentu, tetapi tetap menunjukkan beberapa fluktuasi signifikan yang dapat mempengaruhi tren pasar.



Gambar 5. 8 Gambar visualisasi *boxplot* pada harga penutupan indeks *S&P 500*

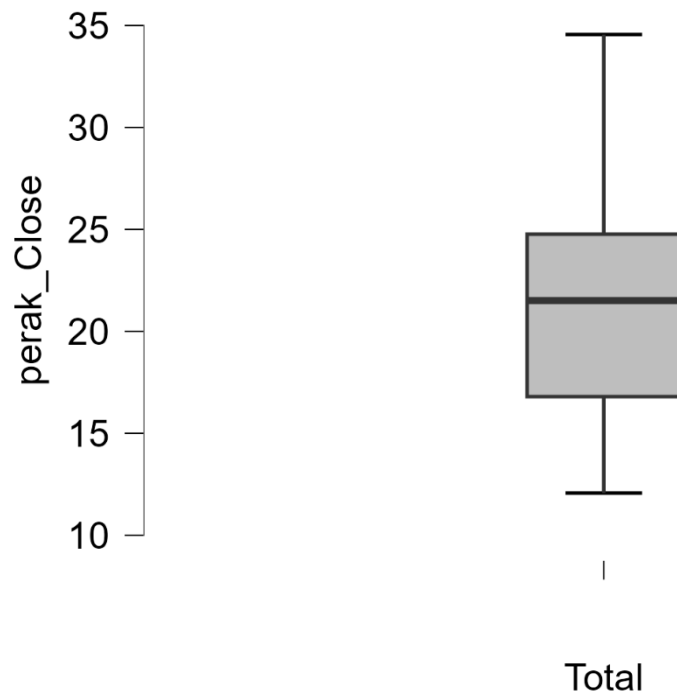
Berdasarkan *boxplot* harga penutupan indeks *S&P 500*, semakin memperkuat analisis distribusi sebelumnya dengan memberikan visualisasi yang jelas mengenai persebaran harga. Median harga berada di sekitar 4000, yang sesuai dengan pola distribusi sebelumnya yang menunjukkan bahwa mayoritas harga

indeks berkisar antara 2500 hingga 4250, di mana kepadatan tertinggi terjadi. Kotak dalam *boxplot* menggambarkan rentang *IQR*, yakni antara 3000 dan 5000, menunjukkan bahwa sebagian besar harga indeks terkonsentrasi di dalam rentang ini, meskipun ada beberapa nilai ekstrem yang memperpanjang distribusi ke arah luar.

Keberadaan outlier di bawah 2000 semakin menegaskan adanya harga ekstrem yang lebih rendah dari batas normal. Outlier ini sesuai dengan ekor panjang yang terlihat dalam distribusi, yang mencerminkan periode volatilitas yang lebih tinggi dalam harga indeks S&P 500. Namun, di sisi lain, *boxplot* juga menunjukkan bahwa harga tertinggi yang bukan outlier mencapai sekitar 6000, yang mengonfirmasi bahwa meskipun ada fluktuasi harga yang signifikan, sebagian besar pergerakan tetap dalam kisaran yang terstruktur.

Struktur *boxplot* juga mendukung temuan mengenai kemiringan ke kanan (positively skewed), seperti yang telah dianalisis dalam distribusi sebelumnya, dengan nilai *skewness* positif sebesar 0.512, yang menunjukkan adanya sebagian kecil harga yang lebih tinggi dari rata-rata yang menarik ekor distribusi ke arah kanan. Standar deviasi sebesar 1029,12 semakin mempertegas adanya variabilitas harga yang cukup besar, namun tetap dalam batas wajar untuk indeks saham bernilai tinggi seperti S&P 500. Selain itu, kurtosis negatif sebesar $-0,678$ mengindikasikan bahwa distribusi harga lebih mendatar dibandingkan distribusi normal, menandakan bahwa probabilitas terjadinya lonjakan harga ekstrem lebih kecil.

Secara keseluruhan, *boxplot* memberikan perspektif tambahan terhadap distribusi harga penutupan indeks S&P 500, dengan mengonfirmasi bahwa meskipun terdapat beberapa outlier yang mendorong distribusi, indeks ini tetap menunjukkan pola yang stabil dengan persebaran harga yang cukup terstruktur dalam *IQR* utama. Hal ini semakin mempertegas bahwa pergerakan harga indeks S&P 500 memiliki karakteristik volatilitas yang tinggi tetapi tetap dalam pola distribusi yang dapat diprediksi.

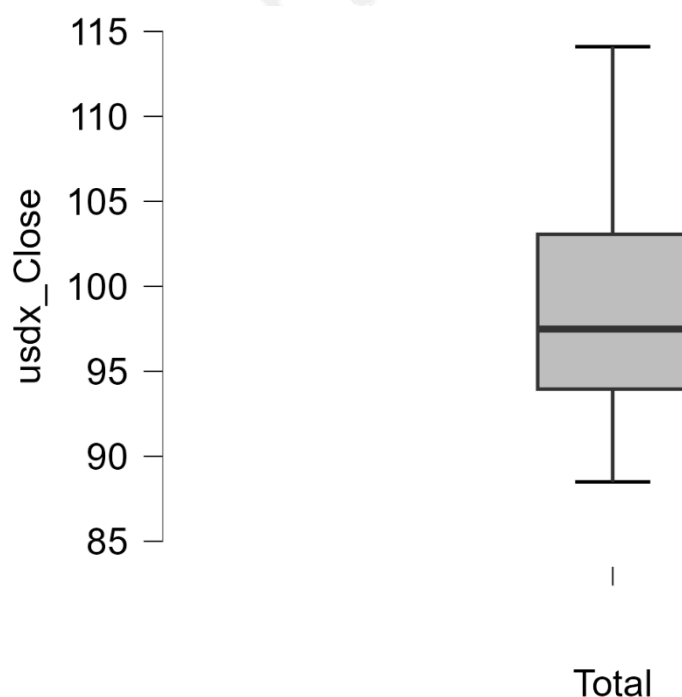


Gambar 5. 9 Gambar visualisasi *boxplot* pada harga penutupan perak

Berdasarkan *boxplot*, harga penutupan perak memberikan gambaran visual yang memperkuat hasil analisis distribusi sebelumnya. Hal ini menegaskan pola bimodal dengan kecenderungan miring ke kanan (*positively skewed*) serta keberadaan beberapa harga ekstrem. Median harga berada di sekitar 21,51, yang sesuai dengan kepadatan distribusi sebelumnya, di mana sebagian besar harga terkonsentrasi di dua puncak utama di kisaran 17 dan 24. Posisi median yang lebih dekat ke bagian bawah kotak dalam *boxplot* menunjukkan bahwa sebagian besar harga berada dalam rentang yang lebih rendah, meskipun terdapat sejumlah harga tinggi yang memperpanjang distribusi ke sisi kanan.

Dari *Boxplot* diperlihatkan *IQR* yang cukup lebar, mengindikasikan bahwa mayoritas harga perak berada dalam kisaran 20 hingga 25, namun dengan beberapa harga ekstrem yang muncul di luar batas ini. Keberadaan *whisker* yang panjang di sisi atas mengonfirmasi adanya harga tinggi yang terjadi dalam beberapa periode tertentu, yang juga tercermin dalam distribusi kepadatan yang menunjukkan ekor panjang di sisi kanan. Di sisi lain, terdapat outlier di bawah 10, yang memperkuat bukti bahwa terdapat nilai-nilai harga yang jauh lebih rendah dari sebaran utama, meskipun kejadian ini tergolong jarang.

Selain itu, nilai *skewness* positif 0.421 yang terlihat dalam analisis deskriptif juga tercermin dalam *boxplot*, dengan persebaran harga yang lebih berat di sisi bawah namun tetap memiliki beberapa harga tinggi yang menarik distribusi ke kanan. Standar deviasi sekitar 5,02 menunjukkan variasi harga yang cukup moderat tetapi tetap memberikan fluktuasi signifikan. Kurtosis negatif -0,838 semakin mempertegas bahwa distribusi lebih mendatar dibandingkan distribusi normal, sehingga probabilitas terjadinya lonjakan harga yang sangat ekstrem tetap rendah.



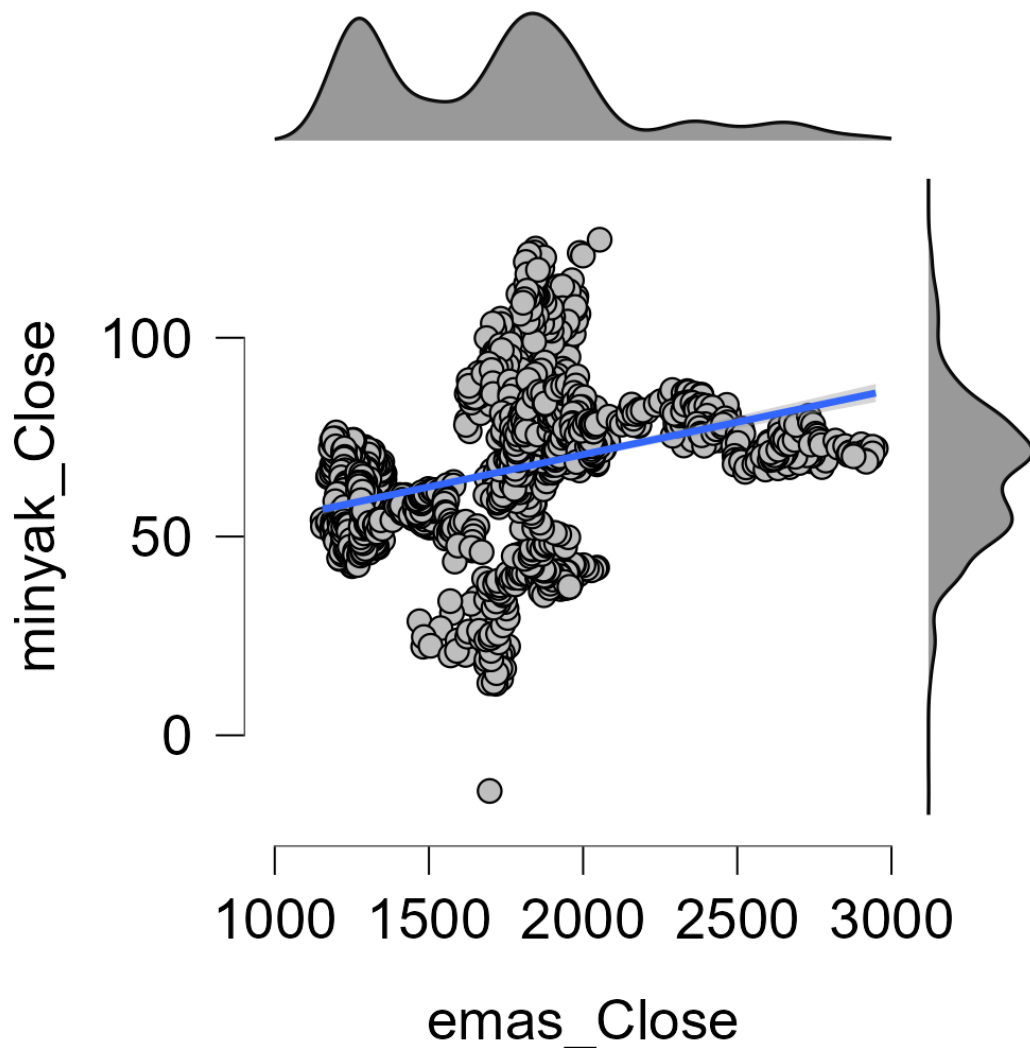
Gambar 5. 10 Gambar visualisasi *boxplot* pada harga penutupan indeks dolar AS

Berdasarkan *boxplot*, harga penutupan indeks dolar AS memberikan gambaran visual yang mempertegas hasil analisis distribusi sebelumnya, menunjukkan kemiringan ke kanan (*positively skewed*) dengan persebaran data yang cukup stabil tetapi tetap memperlihatkan beberapa outlier di sisi kanan distribusi. Median harga berada di sekitar 100, yang sesuai dengan hasil distribusi sebelumnya yang menunjukkan kepadatan tertinggi pada kisaran 95–100, mengindikasikan bahwa harga indeks dolar paling sering muncul dalam rentang ini.

Rentang *IQR* dalam *boxplot* menunjukkan bahwa mayoritas harga indeks terkonsentrasi di antara 95 dan 105, mencerminkan variabilitas yang moderat dalam pergerakan harga. Sementara itu, *whisker* yang lebih panjang di sisi atas menunjukkan adanya harga ekstrem yang lebih tinggi dari batas normal distribusi, yang sesuai dengan karakteristik ekor panjang di sisi kanan yang terlihat dalam grafik kepadatan. Keberadaan outlier di bawah 90 semakin memperjelas bahwa meskipun sebagian besar harga terkonsentrasi di sekitar median, terdapat beberapa nilai rendah yang relatif jarang terjadi tetapi tetap berpengaruh pada persebaran data.

Selain itu, nilai *skewness* positif sebesar 0,308 yang terlihat dalam analisis distribusi juga tercermin dalam *boxplot*, dengan persebaran harga yang lebih berat di sisi bawah tetapi tetap memiliki beberapa harga tinggi yang menarik distribusi ke kanan. Standar deviasi sekitar 5,40 semakin mempertegas adanya variabilitas harga yang cukup moderat, mencerminkan pergerakan indeks dolar yang cukup stabil tetapi tetap mengalami beberapa lonjakan harga dalam periode tertentu. Kurtosis negatif (-0,716) menunjukkan bahwa distribusi lebih mendatar dibandingkan distribusi normal, yang mengindikasikan bahwa probabilitas munculnya lonjakan harga yang sangat ekstrem tetap rendah.

5.1.1.2.3. Scatter Plot Atau Grafik Sebar

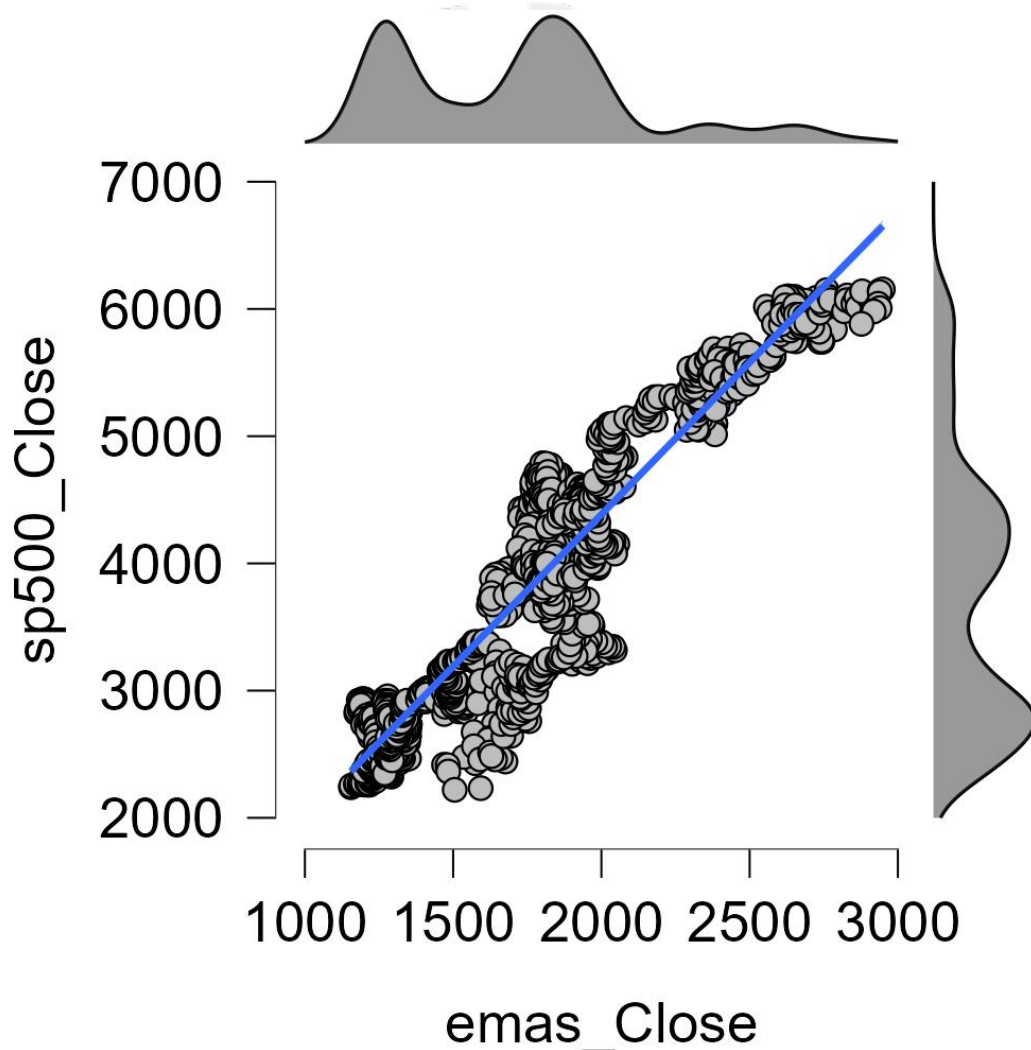


Gambar 5. 11 Gambar visualisasi *scatter plot* persebaran harga emas terhadap harga minyak mentah

Scatter plot yang ditampilkan menunjukkan hubungan antara harga penutupan minyak mentah dengan harga penutupan emas. Pada grafik ini, harga penutupan emas berada pada sumbu X, sementara harga penutupan minyak ditempatkan pada sumbu Y. Titik-titik data yang tersebar mengindikasikan bagaimana pergerakan harga kedua komoditas saling berkaitan. Dari distribusi titik-titik, terlihat bahwa terdapat kecenderungan positif, di mana peningkatan harga emas sering kali diikuti oleh kenaikan harga minyak. Meskipun pola ini tidak sepenuhnya linier, keberadaan garis regresi linier berwarna biru memberikan indikasi bahwa ada hubungan proporsional antara kedua variabel. Selain itu, grafik juga menyertakan *density plot* di bagian atas dan kanan, yang menggambarkan penyebaran data masing-masing variabel. Ini memperjelas pola distribusi harga,

yang mungkin menunjukkan adanya konsentrasi data di rentang tertentu. Hal ini terbukti dari analisis distribusi dan analisis dengan *boxplot* yang memiliki kepadatan pada rentang-rentang masing-masing harga. Dengan demikian, *scatter plot* ini berfungsi sebagai alat eksplorasi awal untuk memahami korelasi harga emas dan minyak, serta memberikan wawasan awal sebelum dilakukan uji korelasi.

emas_Close - sp500_Close

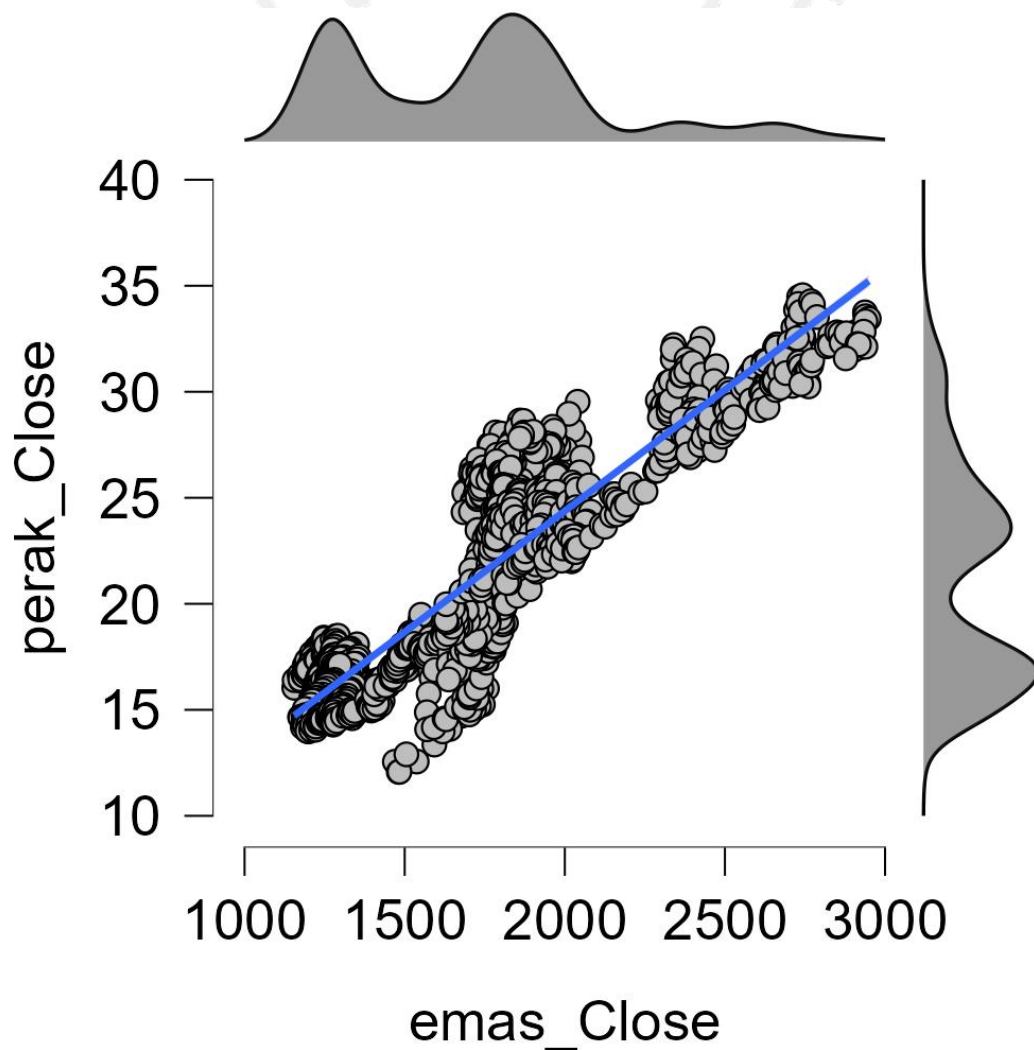


Gambar 5. 12 Gambar visualisasi *scatter plot* persebaran harga emas terhadap indeks *S&P 500*

Scatter plot yang ditampilkan menggambarkan hubungan antara indeks penutupan S&P 500 dengan harga penutupan emas. Harga emas ditampilkan pada sumbu X, sedangkan indeks S&P 500 berada pada sumbu Y. Distribusi titik-titik data menunjukkan pola yang cenderung membentuk hubungan positif, di mana

kenaikan harga emas sering kali sejalan dengan peningkatan indeks S&P 500. Garis regresi linier berwarna biru memberikan indikasi bahwa kedua variabel memiliki tren yang bergerak bersama dalam rentang tertentu. Selain itu, *density plot* di bagian atas dan kanan *scatter plot* menampilkan distribusi data masing-masing variabel, membantu dalam memahami pola penyebaran harga dan indeks. Pola ini dapat menjadi dasar analisis lebih lanjut untuk memahami bagaimana pergerakan pasar ekuitas dan harga emas berinteraksi dalam konteks ekonomi dan investasi.

emas_Close - perak_Close

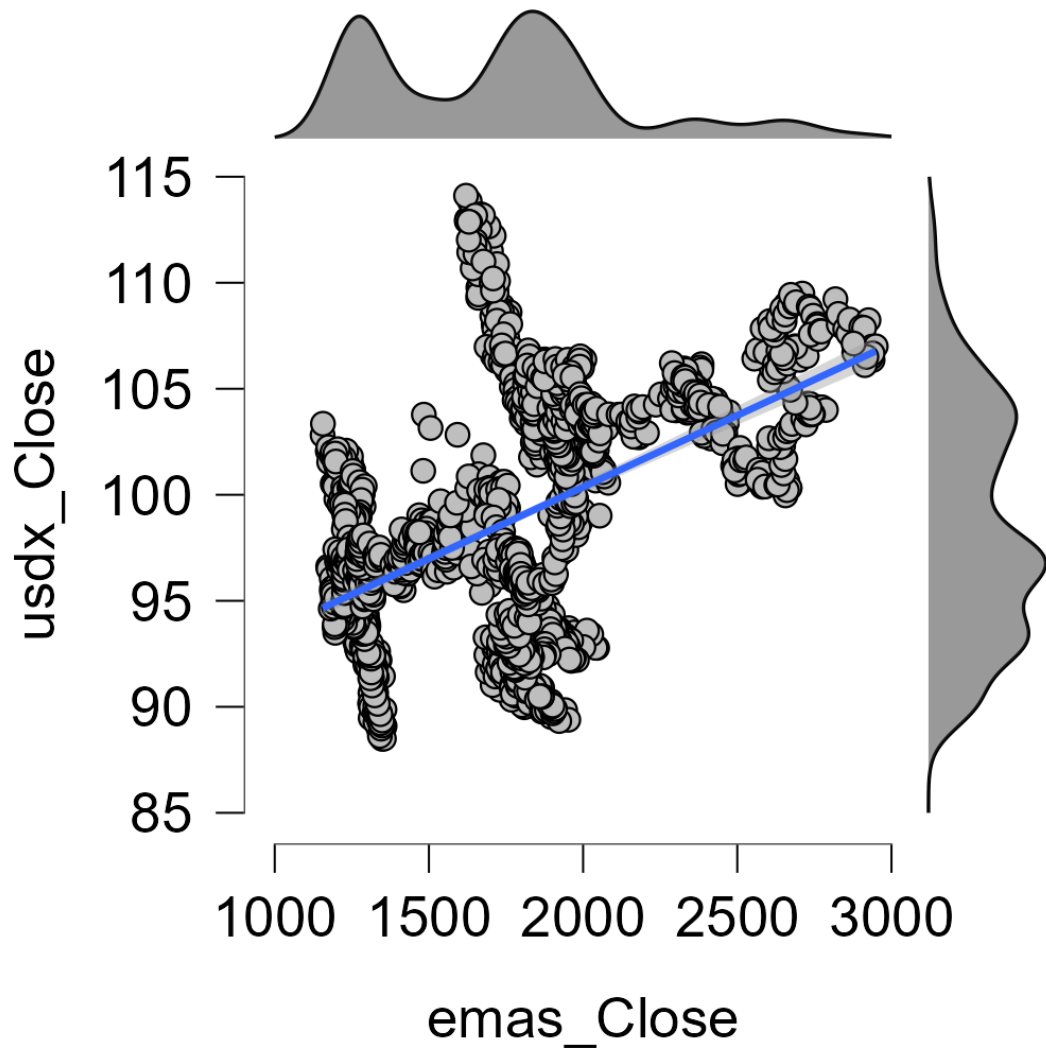


Gambar 5. 13 Gambar visualisasi *scatter plot* persebaran harga emas terhadap harga perak

Scatter plot yang ditampilkan menunjukkan hubungan antara harga penutupan perak dengan harga penutupan emas. Harga emas berada pada sumbu X, sementara harga perak ditempatkan pada sumbu Y, dengan titik-titik data yang

mewakili pasangan nilai dari kedua variabel pada waktu yang sama. Dari pola distribusi titik-titik tersebut, terlihat adanya hubungan positif, di mana kenaikan harga emas cenderung diikuti oleh kenaikan harga perak.

Keberadaan garis regresi linier berwarna biru dalam *scatter plot* memberikan indikasi bahwa hubungan antara harga emas dan perak memiliki tren yang relatif proporsional dalam rentang tertentu. Meskipun titik-titik data menunjukkan sedikit penyebaran yang mungkin mengindikasikan fluktuasi harga, secara umum pola ini mencerminkan hubungan yang erat antara kedua komoditas. Selain itu, histogram yang berada di sisi atas dan kanan *scatter plot* memberikan gambaran distribusi harga masing-masing variabel, memperlihatkan bagaimana harga perak dan emas tersebar dalam rentang tertentu. Visualisasi ini dapat digunakan sebagai alat eksplorasi awal dalam memahami keterkaitan pergerakan harga perak terhadap emas, sebelum dilakukan analisis statistik formal seperti uji korelasi.



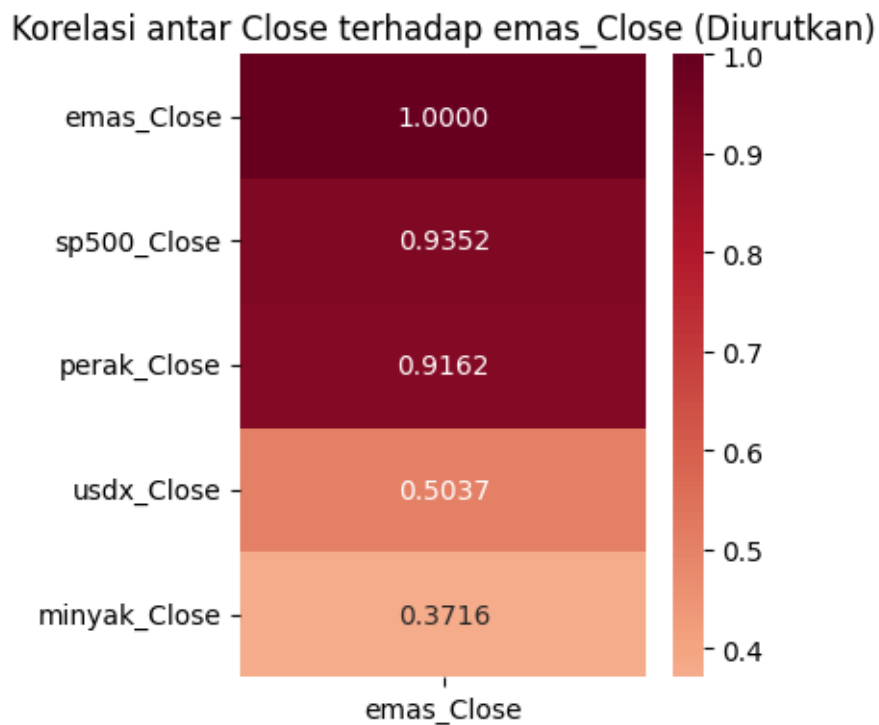
Gambar 5. 14 Gambar visualisasi *scatter plot* persebaran harga emas terhadap indeks dolar AS

Scatter plot yang ditampilkan menggambarkan hubungan antara indeks penutupan dolar AS dengan harga penutupan emas. Pada sumbu X, variabel yang ditampilkan adalah harga penutupan emas, sedangkan pada sumbu Y adalah indeks penutupan dolar AS. Titik-titik data pada grafik merepresentasikan pasangan nilai dari kedua variabel pada waktu yang sama, memungkinkan analisis terhadap pola hubungan antara harga emas dan pergerakan indeks dolar AS. Dari distribusi titik-titik pada *scatter plot*, terlihat bahwa hubungan antara kedua variabel ini cenderung negatif, di mana peningkatan harga emas sering kali disertai dengan penurunan nilai indeks dolar AS. Garis regresi linier yang ditambahkan dalam grafik memberikan indikasi kecenderungan hubungan positif yang mengisyaratkan adanya pergerakan kenaikan yang bersamaan antara harga emas dan indeks dolar AS. Selain itu, *density*

plot yang berada di sisi atas dan kanan *scatter plot* menunjukkan distribusi masing-masing variabel, membantu dalam memahami pola penyebaran nilai indeks dan harga emas dalam rentang tertentu. Visualisasi ini berfungsi sebagai langkah eksplorasi awal untuk memahami bagaimana pergerakan indeks dolar AS berkorelasi dengan harga emas sebelum dilakukan analisis statistik lebih mendalam, seperti uji korelasi.

5.1.1.3. Uji Korelasi

Uji korelasi dilakukan untuk mengevaluasi kekuatan dan arah hubungan linear antar harga penutupan dari lima instrumen global, yaitu emas, minyak mentah, perak, indeks S&P 500, dan indeks dolar AS (*USD*X). Dengan menggunakan metode *Pearson Correlation Coefficient* pada tingkat signifikansi 5% ($\alpha = 0,05$), analisis ini bertujuan mengidentifikasi apakah terdapat keterkaitan signifikan secara statistik di antara pasangan variabel tersebut. Nilai koefisien korelasi yang diperoleh berkisar antara -1 hingga 1, yang merepresentasikan arah dan kekuatan hubungan, mulai dari negatif kuat, tidak berkorelasi, hingga positif kuat. Untuk mempermudah interpretasi, hasil uji ini divisualisasikan dalam bentuk *heatmap*, yang menggambarkan nilai korelasi antar variabel secara visual menggunakan gradasi warna. Visualisasi ini membantu dalam mengidentifikasi pola keterkaitan antar aset yang dapat digunakan sebagai dasar dalam strategi analisis hubungan antar pasar maupun untuk seleksi fitur dalam analisis prediktif sehingga dapat membuang fitur yang tidak memiliki korelasi.



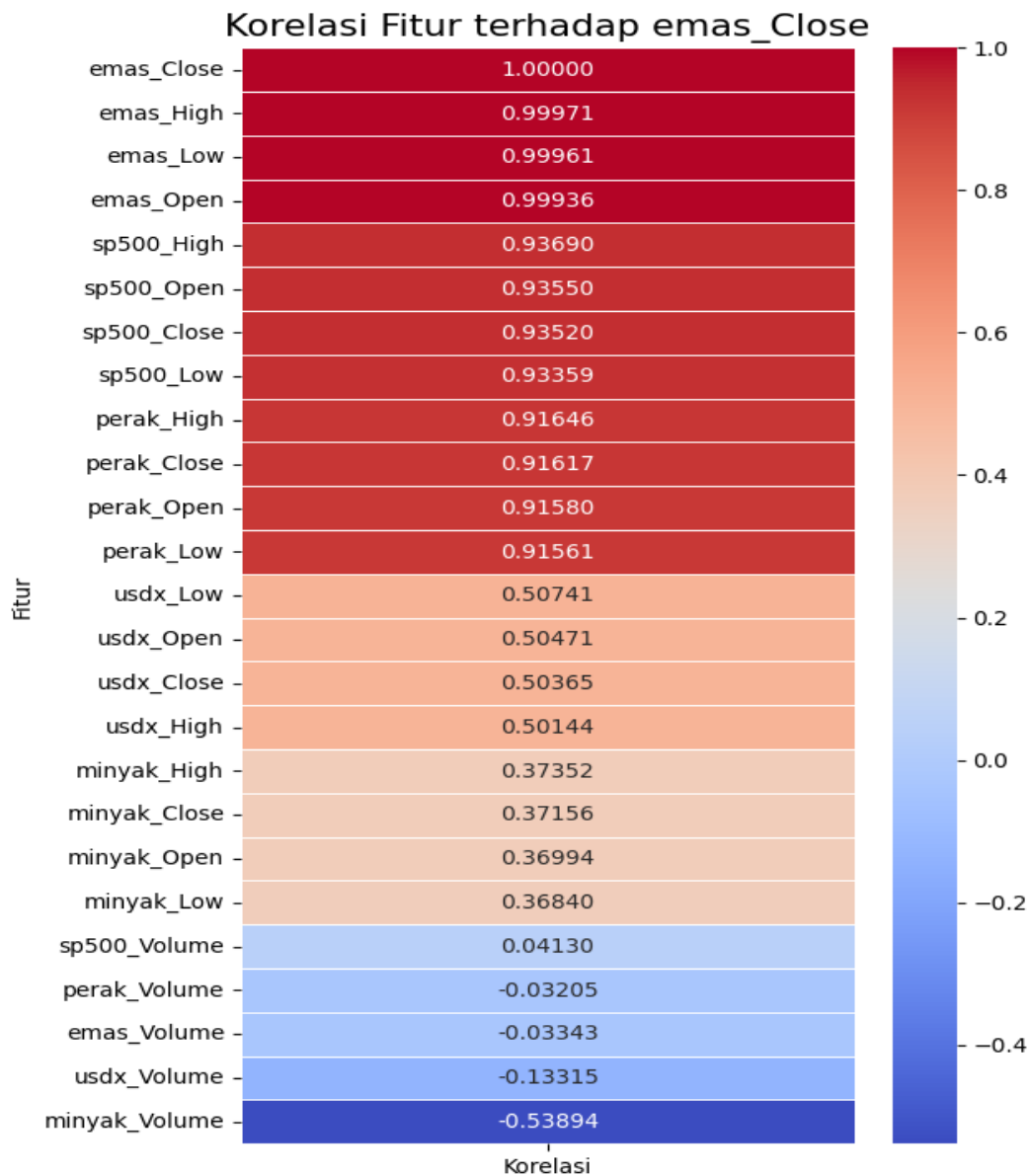
Gambar 5. 15 Gambar *heatmap* korelasi *Pearson* antar variabel harga penutupan masing-masing harga

Berdasarkan *heatmap* hasil uji korelasi *Pearson* menunjukkan hubungan linier antara harga penutupan emas (emas_Close) dengan berbagai variabel lainnya, termasuk minyak_Close, perak_Close, S&P 500 (sp500_Close), dan indeks dolar AS (usdx_Close). Visualisasi ini memberikan wawasan tentang keterkaitan antar aset global, di mana intensitas hubungan direpresentasikan melalui gradasi warna yang mencerminkan kekuatan korelasi.

Dari analisis ini, terlihat bahwa emas_Close memiliki korelasi yang sangat kuat dengan dirinya sendiri (1.0000), seperti yang diharapkan. Korelasi yang tinggi juga ditemukan antara emas_Close dan S&P 500 (0.9352) serta perak_Close (0.9162), menunjukkan bahwa pergerakan harga emas memiliki hubungan erat dengan pasar saham AS serta harga perak, yang sering dianggap sebagai aset komplementer dalam dunia investasi. Sebaliknya, korelasi dengan indeks dolar AS (0.5037) menunjukkan hubungan yang lebih moderat, mengindikasikan bahwa walaupun emas sering digunakan sebagai lindung nilai terhadap dolar, pergerakannya tidak selalu berbanding terbalik dalam pola yang konsisten. Korelasi terlemah dalam *heatmap* ini terlihat antara harga penutupan emas dan harga

penutupan minyak mentah (0.3716). Walaupun terdapat hubungan positif, nilainya relatif rendah dibandingkan dengan aset lainnya, menunjukkan bahwa harga minyak tidak selalu bergerak sejalan dengan harga emas.

Sebagai tambahan untuk uji korelasi terhadap semua variabel dari semua instrumen harga yang didapatkan, berikut adalah visualisasi dari *heatmap* uji korelasi Pearson untuk semua variabel terhadap harga penutupan emas.



Gambar 5. 16 Gambar *heatmap* korelasi *Pearson* semua variabel harga

Berdasarkan dari *heatmap* semua variabel dari semua instrumen, dapat disimpulkan bahwa volume perdagangan dari berbagai aset menunjukkan hubungan

yang lemah atau bahkan negatif terhadap harga penutupan emas. Korelasi negatif ini mengindikasikan bahwa peningkatan volume transaksi tidak selalu sejalan dengan perubahan harga emas, dan dalam beberapa kasus, bahkan bergerak berlawanan.

Secara keseluruhan, hubungan negatif atau lemah antara volume perdagangan dan harga penutupan emas menunjukkan bahwa aktivitas transaksi tidak memiliki keterkaitan linear yang kuat dengan pergerakan harga. Volume sering kali mengalami fluktuasi yang tidak sejalan dengan perubahan harga emas, mencerminkan bahwa faktor-faktor harga lainnya memiliki pengaruh yang lebih signifikan dalam menentukan tren harga. Korelasi negatif pada beberapa aset, seperti volume minyak terhadap harga emas, menandakan bahwa peningkatan jumlah transaksi tidak selalu beriringan dengan kenaikan harga emas, melainkan dapat berfluktuasi secara independen.

Selain itu, korelasi yang sangat kecil antara volume perdagangan emas sendiri terhadap harga penutupan emas mengindikasikan bahwa jumlah transaksi dalam satu periode tidak dapat digunakan sebagai fitur prediksi untuk pergerakan harga emas. Meskipun emas dan perak sering bergerak dalam pola yang serupa, volume perak juga menunjukkan korelasi yang rendah terhadap harga emas, yang menegaskan bahwa meskipun harga kedua logam mulia dapat saling terkait, aktivitas perdagangan tidak selalu merefleksikan hubungan harga yang kuat.

Dengan demikian, volume perdagangan dalam berbagai aset lebih bersifat sebagai indikator aktivitas pasar daripada penentu langsung pergerakan harga emas, sehingga dalam analisis harga, faktor-faktor lain yang lebih mencerminkan pergerakan harga perlu mendapatkan prioritas lebih tinggi dibandingkan volume transaksi.

5.1.1.4. Kesimpulan Seleksi Fitur

Kesimpulan seleksi fitur dalam penelitian ini didasarkan pada analisis statistik deskriptif, analisa berdasarkan visualisasi, dan uji korelasi antar variabel terhadap harga penutupan emas (emas_Close). Dari hasil eksplorasi data, beberapa

aspek utama yang diperhatikan meliputi distribusi data, hubungan antar fitur, serta signifikansi dalam representasi karakteristik aset.

Berdasarkan statistik deskriptif, seluruh variabel memiliki jumlah data valid yang lengkap tanpa adanya nilai yang hilang. Distribusi harga emas menunjukkan pola *positively skewed*, dengan median yang sedikit lebih tinggi dibandingkan rata-rata, serta standar deviasi yang mencerminkan tingkat volatilitas yang cukup signifikan. Harga perak dan indeks S&P 500 menunjukkan pola distribusi yang relatif stabil, meskipun tetap menunjukkan variasi harga yang cukup besar. Sebaliknya, harga minyak mentah memiliki beberapa nilai ekstrem, termasuk harga negatif yang terjadi akibat fenomena pasar berjangka saat krisis pandemi. Indeks dolar AS menunjukkan pola distribusi yang cenderung simetris dengan sedikit kemiringan ke kanan.

Uji korelasi Pearson mengonfirmasi hubungan yang kuat antara harga emas dengan harga perak (0.9162) dan indeks S&P 500 (0.9352), menandakan keterkaitan erat antara logam mulia dan pasar ekuitas. Korelasi dengan indeks dolar AS (0.5037) bersifat moderat, menunjukkan bahwa meskipun emas sering digunakan sebagai lindung nilai terhadap dolar, pergerakannya tidak selalu berkebalikan dalam pola yang konsisten. Sementara itu, harga minyak mentah memiliki korelasi paling rendah terhadap harga emas (0.3716), mengindikasikan bahwa fluktuasi harga minyak tidak selalu bergerak selaras dengan pergerakan emas.

Dalam penelitian ini, volume perdagangan dari berbagai instrumen menunjukkan korelasi yang lemah atau negatif terhadap harga penutupan emas, sehingga tidak digunakan sebagai fitur dalam analisis *time-series* dengan pendekatan *DL* metode *LSTM*. Korelasi negatif antara volume minyak dan harga emas, misalnya, menunjukkan bahwa peningkatan jumlah transaksi tidak selalu beriringan dengan kenaikan harga emas, melainkan dapat berfluktuasi secara independen. Selain itu, volume perdagangan emas sendiri memiliki korelasi yang sangat kecil terhadap harga penutupannya, sehingga tidak dapat dijadikan fitur prediksi dalam pemodelan harga berbasis *time-series*.

Dengan mempertimbangkan berbagai analisis deskriptif seperti statistik deskriptif, distribusi data, hasil visualisasi, khususnya uji korelasi Pearson, penelitian ini memutuskan untuk mengecualikan variabel volume dalam pemodelan prediktif menggunakan *LSTM*. Fokus analisis diarahkan pada variabel harga yang memiliki keterkaitan lebih erat dengan harga emas, sehingga model dapat menangkap pola pergerakan harga dengan lebih akurat tanpa dipengaruhi oleh fluktuasi volume yang tidak memiliki hubungan linear yang konsisten terhadap harga emas. Keputusan ini didasarkan pada evaluasi statistik yang menunjukkan bahwa volume lebih bersifat sebagai indikator aktivitas pasar daripada sebagai penentu langsung pergerakan harga emas. Dengan demikian, pemodelan prediktif dengan metode *LSTM* dapat lebih optimal dalam mengidentifikasi tren harga berdasarkan variabel yang memiliki hubungan lebih kuat dan signifikan.

5.1.1.5. Normalisasi Data

Setelah menerapkan seleksi fitur, semua fitur yang digunakan dalam analisis telah dipilih berdasarkan relevansi terhadap harga penutupan emas (*emas_Close*). Untuk memastikan bahwa data berada dalam skala yang seragam dan mengurangi pengaruh perbedaan rentang nilai antar fitur, dilakukan proses normalisasi menggunakan teknik *Min-Max Scaling*.

Dengan pendekatan ini, setiap fitur akan memiliki nilai yang dipetakan dalam rentang 0 hingga 1, sehingga skala variabel menjadi lebih seragam tanpa mengubah pola distribusi data secara signifikan. Berdasarkan hasil statistik deskriptif, fitur harga yang digunakan dalam penelitian ini, seperti harga penutupan emas, perak, minyak, indeks S&P 500, dan indeks dolar AS, memiliki rentang nilai yang cukup beragam. Oleh karena itu, normalisasi Min-Max dilakukan untuk setiap fitur harga guna memastikan bahwa model pembelajaran tidak dipengaruhi oleh perbedaan skala antar variabel.

Implementasi normalisasi ini bertujuan untuk meningkatkan stabilitas dan konvergensi model saat proses pelatihan dengan metode *LSTM*, sehingga model dapat lebih akurat dalam menangkap pola pergerakan harga tanpa terpengaruh oleh skala awal dari fitur yang berbeda-beda. Setelah proses normalisasi selesai, *dataset* yang telah diubah ke skala $[0,1]$ siap digunakan dalam tahap pemodelan berikutnya.

Berikut tabel lima baris terakhir dari semua fitur yang telah dinormalisasi dengan *Min-Max Scaling* yang dimulai dengan harga emas, minyak mentah, indeks *S&P 500*, harga perak, dan indeks dolar AS.

Tabel 5. 6 Tabel hasil normalisasi 5 baris terakhir pada harga emas

Date	emas_Open	emas_High	emas_Low	emas_Close
2025-02-24	1	1	1	0,991081
2025-02-25	0,97572	0,99182	0,975876	0,994202
2025-02-26	0,982601	0,977241	0,98875	0,980155
2025-02-27	0,963804	0,980469	0,964905	0,982942
2025-02-28	0,937846	0,955039	0,944587	0,959864

Tabel 5. 7 Tabel hasil normalisasi 5 baris terakhir pada harga minyak mentah

Date	minyak_Open	minyak_High	minyak_Low	minyak_Close
2025-02-24	0,671481	0,490112	0,683508	0,604356
2025-02-25	0,66051	0,492852	0,676556	0,612433
2025-02-26	0,658588	0,475901	0,67457	0,59938
2025-02-27	0,669311	0,486688	0,676122	0,597288
2025-02-28	0,665654	0,484548	0,679412	0,607024

Tabel 5. 8 Tabel hasil normalisasi 5 baris terakhir pada indeks *S&P 500*

Date	sp500_Open	sp500_High	sp500_Low	sp500_Close
2025-02-24	0,958846	0,974661	0,965931	0,971272
2025-02-25	0,951046	0,961479	0,948107	0,962565
2025-02-26	0,951237	0,963463	0,953543	0,956337
2025-02-27	0,927267	0,961095	0,935213	0,955828
2025-02-28	0,949334	0,949962	0,928892	0,931168

Tabel 5. 9 Tabel hasil normalisasi 5 baris terakhir pada harga perak

Date	perak_Open	perak_High	perak_Low	perak_Close
2025-02-24	0,902321	0,912506	0,918466	0,923505
2025-02-25	0,868808	0,894167	0,884354	0,891928
2025-02-26	0,888379	0,885904	0,904275	0,897621
2025-02-27	0,868808	0,887539	0,869766	0,89682
2025-02-28	0,843609	0,859037	0,852799	0,867467

Tabel 5. 10 Tabel hasil normalisasi 5 baris terakhir pada indeks dolar AS

Date	usdx_Open	usdx_High	usdx_Low	usdx_Close
2025-02-24	0,704878	0,687488	0,712266	0,701542
2025-02-25	0,693877	0,691549	0,714855	0,70974
2025-02-26	0,697831	0,684007	0,71366	0,688073
2025-02-27	0,730992	0,709534	0,724811	0,703299
2025-02-28	0,745948	0,723458	0,755476	0,727894

5.1.1.6. Pembuatan *Dataset*

Setelah seluruh variabel dinormalisasi dengan teknik *Min-Max Scaling*, *dataset* diorganisasikan menggunakan pendekatan *sliding window* dengan ukuran jendela waktu 30 hari. Artinya, dalam proses prediksi harga emas pada hari t , data historis dari 30 hari sebelumnya digunakan sebagai *input*. Penerapan teknik ini memungkinkan model untuk menangkap pola tren harga berdasarkan informasi dari periode sebelumnya, yang berperan penting dalam analisis *time-series*.

Selain itu, *dataset* atau kumpulan data dipartisi menjadi tiga ukuran berdasarkan rasio data pelatihan dan data pengujian, yaitu 70:30, 80:20, dan 90:10. Pembagian ini dilakukan untuk mengevaluasi dampak jumlah data pelatihan terhadap performa prediksi model. Berikut adalah perbandingan jumlah baris data pelatihan dan data pengujian pada masing-masing ukuran *dataset* yang dijabarkan dalam bentuk tabel.

Tabel 5. 11 Tabel perbandingan jumlah baris data setiap ukuran *dataset*

Proporsi Ukuran <i>Dataset</i>	Jumlah Baris Data Pelatihan	Jumlah Baris Data Pengujian
Ukuran <i>Dataset</i> 70:30	1414	606
Ukuran <i>Dataset</i> 80:20	1616	404
Ukuran <i>Dataset</i> 90:10	1818	202

Dengan struktur *dataset* ini, model *LSTM* dapat memanfaatkan data berderet waktu untuk membangun prediksi berdasarkan pola historis yang relevan. Variasi dalam pembagian *dataset* memungkinkan evaluasi efektivitas model dengan berbagai tingkat pelatihan dan pengujian, sehingga memberikan gambaran optimal tentang bagaimana pendekatan time series dapat digunakan dalam analisis harga emas.

5.1.2 Analisa Data Dengan Metode *LSTM*

Pada subsubbab ini akan memaparkan hasil pelatihan dan pengujian model *LSTM* dengan berbagai kumpulan data atau *dataset* yang telah diusulkan dalam perancangan. Hasil dari prediksi untuk model dalam memprediksi data tahunan akan dipaparkan dalam sub subbab ini. Hasil pelatihan ini mencakup nilai evaluasi yaitu *MSE*, *MAE*, *RMSE*, dan *MAPE*.

5.1.2.1. Hasil Pelatihan dan Pengujian Model *LSTM*

A. Hasil pelatihan model dengan kumpulan data ukuran 70:30

Tabel 5. 12 Tabel hasil pelatihan dan pengujian dengan ukuran kumpulan data 70:30

Mode l	<i>MSE</i> pelatiha n	<i>RMSE</i> pelatiha n	<i>MAE</i> pelatiha n	<i>MAPE</i> pelatiha n	<i>MSE</i> pengujia n	<i>RMSE</i> pengujia n	<i>MAE</i> pengujia n	<i>MAPE</i> pengujia n
Mode 11	431,49	20,77	17,18	1,10%	5510,31	74,23	50,87	2,12%
Mode 12	97,31	9,86	7,34	0,47%	4981,73	70,58	46,92	1,89%
Mode 13	316,07	17,78	15,94	1,05%	2233,96	47,26	32,43	1,35%

Mode 14	233,47	15,28	12,52	0,82%	2018,39	44,93	31,23	1,30%
Mode 15	104,08	10,20	7,52	0,48%	2212,72	47,04	34,00	1,41%

Tabel 5. 13 Tabel kenaikan dari pelatihan ke pengujian ukuran *dataset* 70:30

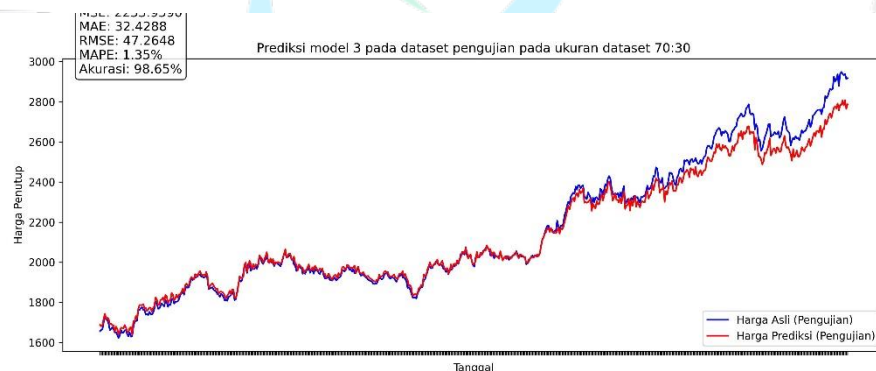
Model	MSE pelatihan→ MSE pengujian	$RMSE$ pelatihan→ $RMSE$ pengujian	MAE pelatihan→ MAE pengujian	$MAPE$ pelatihan→ $MAPE$ pengujian
Model 1	1177,04%	257,39%	196,10%	92,73%
Model 2	5019,44%	615,82%	539,24%	302,13%
Model 3	606,79%	165,80%	103,45%	28,57%
Model 4	764,52%	194,04%	149,44%	58,54%
Model 5	2025,98%	361,18%	352,13%	193,75%

Model 1 menunjukkan adanya *overfitting* yang jelas, di mana kinerjanya pada data pelatihan jauh lebih baik dibandingkan dengan data pengujian. Hal ini terlihat dari kenaikan yang signifikan pada metrik evaluasi. Untuk MSE , terjadi kenaikan sebesar 1.177,04% antara MSE pelatihan dan MSE pengujian, yang menunjukkan bahwa model mampu mempelajari data pelatihan dengan sangat baik, tetapi kesulitan saat diterapkan pada data pengujian. Begitu pula dengan MAE , yang mengalami kenaikan sebesar 196,10%, mengindikasikan bahwa model lebih akurat pada data pelatihan namun prediksinya jauh lebih buruk pada data pengujian. Kenaikan juga terjadi pada $RMSE$, dengan peningkatan sebesar 257,39%, dan $MAPE$ yang meningkat sebesar 92,73%, semakin memperlihatkan bahwa model terlalu "menghafal" data pelatihan dan tidak dapat menggeneralisasi dengan baik pada data pengujian.

Model 2 menunjukkan adanya *overfitting* yang lebih parah dibandingkan Model 1, di mana kinerjanya pada data pelatihan jauh lebih baik dibandingkan dengan data pengujian. Hal ini terlihat dari kenaikan yang sangat signifikan pada metrik evaluasi. Untuk MSE , terjadi kenaikan sebesar 5.019,44% antara MSE pelatihan dan MSE pengujian, yang menunjukkan bahwa model mampu mempelajari data pelatihan dengan sangat baik, tetapi kesulitan ketika diterapkan pada data pengujian. Begitu pula dengan MAE , yang mengalami kenaikan sebesar

539,24%, mengindikasikan bahwa model lebih akurat pada data pelatihan namun prediksinya jauh lebih buruk pada data pengujian. Kenaikan juga terjadi pada $RMSE$, dengan peningkatan sebesar 615,82%, dan $MAPE$ yang meningkat sebesar 302,13%, semakin memperlihatkan bahwa model terlalu "menghafal" data pelatihan dan tidak dapat menggeneralisasi dengan baik pada data pengujian.

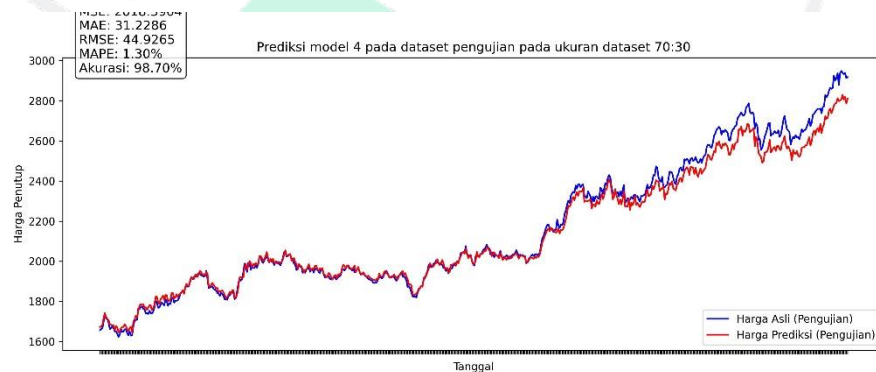
Model 3 menunjukkan adanya *overfitting* yang cukup jelas, di mana kinerjanya pada data pelatihan jauh lebih baik dibandingkan dengan data pengujian. Hal ini terlihat dari kenaikan yang signifikan pada metrik evaluasi. Untuk MSE , terjadi kenaikan sebesar 606,79% antara MSE pelatihan dan MSE pengujian, yang menunjukkan bahwa model mampu mempelajari data pelatihan dengan baik, namun kesulitan saat diterapkan pada data pengujian. Begitu pula dengan MAE , yang mengalami kenaikan sebesar 103,45%, mengindikasikan bahwa model lebih akurat pada data pelatihan namun prediksinya kurang memadai pada data pengujian. Kenaikan juga terjadi pada $RMSE$, dengan peningkatan sebesar 165,80%, dan $MAPE$ yang meningkat sebesar 28,57%, semakin memperlihatkan bahwa model terlalu "menghafal" data pelatihan dan tidak dapat menggeneralisasi dengan baik pada data pengujian. Meskipun demikian, model 3 merupakan model yang paling stabil dibanding kelima model lainnya karena kenaikan yang tidak setajam model lainnya untuk ukuran *dataset* 70:30.



Gambar 5. 17 Prediksi model 3 pada *dataset* pengujian dengan ukuran 70:30

Model 4 menunjukkan hasil yang sangat baik dan merupakan model yang paling optimal dibandingkan dengan model-model sebelumnya. Meskipun ada sedikit indikasi *overfitting*, kenaikan pada metrik evaluasi tetap terjaga dalam batas yang wajar. Untuk MSE , terjadi kenaikan sebesar 764,52% antara MSE pelatihan

dan MSE pengujian, yang menunjukkan bahwa model ini berhasil mempelajari data pelatihan dengan sangat baik dan tetap mempertahankan kinerja yang cukup baik pada data pengujian. MAE mengalami kenaikan sebesar 149,44%, yang menunjukkan bahwa prediksi pada data pelatihan jauh lebih akurat, namun tetap menunjukkan performa yang relatif baik pada data pengujian. Kenaikan pada $RMSE$ sebesar 194,04% dan $MAPE$ yang meningkat sebesar 58,54% semakin memperlihatkan bahwa meskipun terjadi peningkatan kesalahan saat diuji pada data pengujian, model ini tetap lebih baik daripada model-model sebelumnya dalam menjaga keseimbangan antara akurasi pelatihan dan generalisasi pada data pengujian. Dengan hasil ini, Model 4 dapat dianggap sebagai model yang paling efektif dalam menghasilkan prediksi yang stabil dan lebih dapat diandalkan.



Gambar 5. 18 Prediksi model 4 pada *dataset* pengujian dengan ukuran 70:30

Model 5 menunjukkan adanya *overfitting* yang cukup signifikan, meskipun kinerjanya pada data pelatihan jauh lebih baik dibandingkan dengan data pengujian. Kenaikan yang terjadi pada metrik evaluasi sangat tajam. Untuk MSE , terjadi kenaikan sebesar 2.025,98% antara MSE pelatihan dan MSE pengujian, yang menunjukkan bahwa model mempelajari data pelatihan dengan sangat baik, namun kesulitan dalam menerapkan pengetahuan tersebut pada data pengujian. Begitu pula dengan MAE , yang mengalami kenaikan sebesar 352,13%, menandakan bahwa model jauh lebih akurat pada data pelatihan, tetapi prediksinya sangat buruk pada data pengujian. $RMSE$ meningkat sebesar 361,18%, dan $MAPE$ yang meningkat sebesar 193,75%, semakin memperlihatkan bahwa model ini terlalu "menghafal" data pelatihan, namun tidak dapat menggeneralisasi dengan baik pada data pengujian. Meskipun demikian, model 5 memiliki performa yang lebih buruk dibandingkan model-model sebelumnya dalam hal generalisasi, menandakan

bahwa meskipun kinerja pada pelatihan sangat baik, model ini terlalu sesuai dan kurang efektif untuk prediksi pada data yang belum dilihat sebelumnya.

B. Hasil pelatihan model dengan kumpulan data ukuran 80:20

Tabel 5. 14 Tabel hasil pelatihan dan pengujian dengan kumpulan data 80:20

Model	MSE pelatiha n	$RMSE$ pelatiha n	MAE pelatiha n	$MAPE$ pelatiha n	MSE pengujia n	$RMSE$ pengujia n	MAE pengujia n	$MAPE$ pengujia n
Model 1	240,87	15,52	13,42	0,87%	3864,10	62,16	45,67	1,79%
Model 2	134,51	11,60	8,99	0,55%	3666,76	60,55	49,62	2,00%
Model 3	615,41	24,81	21,39	1,31%	770,34	27,76	24,46	1,11%
Model 4	98,40	9,92	7,28	0,45%	1543,11	39,28	28,95	1,15%
Model 5	167,81	12,95	10,32	0,66%	1264,04	35,55	26,85	1,07%

Tabel 5. 15 Tabel kenaikan dari pelatihan ke pengujian ukuran *dataset* 80:20

Model	MSE pelatihan→ MSE pengujian	$RMSE$ pelatihan→ $RMSE$ pengujian	MAE pelatihan→ MAE pengujian	$MAPE$ pelatihan→ $MAPE$ pengujian
Model 1	1504,23%	300,52%	240,31%	105,75%
Model 2	2626,01%	421,98%	451,95%	263,64%
Model 3	25,18%	11,89%	14,35%	-15,27%
Model 4	1468,20%	295,97%	297,66%	155,56%
Model 5	653,26%	174,52%	160,17%	62,12%

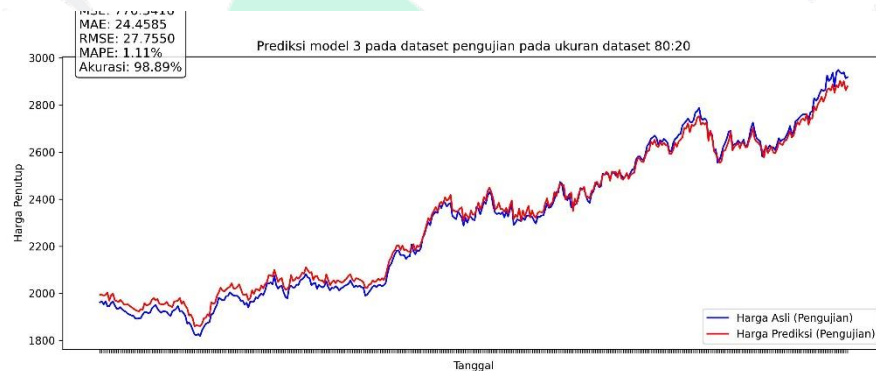
Model 1 dengan *dataset* 80:20 menunjukkan adanya *overfitting* yang cukup jelas, di mana kinerjanya pada data pelatihan jauh lebih baik dibandingkan dengan data pengujian. Hal ini terlihat dari kenaikan yang signifikan pada metrik evaluasi. Untuk MSE , terjadi kenaikan sebesar 1.504,23% antara MSE pelatihan dan MSE pengujian, yang menunjukkan bahwa model mampu mempelajari data pelatihan dengan sangat baik, tetapi kesulitan saat diterapkan pada data pengujian. Begitu

pula dengan *MAE*, yang mengalami kenaikan sebesar 240,31%, mengindikasikan bahwa model lebih akurat pada data pelatihan, namun prediksinya jauh lebih buruk pada data pengujian. Kenaikan juga terjadi pada *RMSE*, dengan peningkatan sebesar 300,52%, dan *MAPE* yang meningkat sebesar 105,75%, semakin memperlihatkan bahwa model terlalu "menghafal" data pelatihan dan tidak dapat menggeneralisasi dengan baik pada data pengujian. Kenaikan yang tajam pada metrik ini menandakan bahwa meskipun model sangat baik dalam mempelajari data pelatihan, performanya menurun secara signifikan ketika diuji pada data yang tidak terlihat sebelumnya.

Model 2 menunjukkan adanya *overfitting* yang lebih parah dibandingkan dengan Model 1, di mana kinerjanya pada data pelatihan jauh lebih baik dibandingkan dengan data pengujian. Hal ini terlihat dari kenaikan yang sangat signifikan pada metrik evaluasi. Untuk *MSE*, terjadi kenaikan sebesar 2.626,01% antara *MSE* pelatihan dan *MSE* pengujian, yang menunjukkan bahwa model mampu mempelajari data pelatihan dengan sangat baik, namun kesulitan saat diterapkan pada data pengujian. Begitu pula dengan *MAE*, yang mengalami kenaikan sebesar 451,95%, mengindikasikan bahwa model lebih akurat pada data pelatihan, namun prediksinya sangat buruk pada data pengujian. Kenaikan juga terjadi pada *RMSE*, dengan peningkatan sebesar 421,98%, dan *MAPE* yang meningkat sebesar 263,64%, semakin memperlihatkan bahwa model terlalu "menghafal" data pelatihan dan tidak dapat menggeneralisasi dengan baik pada data pengujian. Kenaikan yang sangat tajam pada metrik ini menandakan bahwa meskipun model mempelajari data pelatihan dengan sangat baik, kinerjanya menurun sangat drastis ketika diterapkan pada data yang tidak terlihat sebelumnya.

Model 3 menunjukkan hasil yang lebih optimal dan stabil dibandingkan dengan model-model lainnya, meskipun masih terdapat indikasi *overfitting*. Kenaikan pada metrik evaluasi antara data pelatihan dan data pengujian terjaga dalam batas yang wajar. Untuk *MSE*, terjadi kenaikan sebesar 25,18%, yang relatif kecil dibandingkan dengan model-model lainnya, menunjukkan bahwa meskipun model mampu mempelajari data pelatihan dengan baik, kinerjanya pada data pengujian tetap cukup stabil. *MAE* mengalami kenaikan sebesar 14,35%, yang juga menunjukkan stabilitas prediksi pada data pengujian. *RMSE* mengalami

peningkatan sebesar 11,89%, sementara *MAPE* justru menurun sebesar -15,27%, yang menunjukkan bahwa meskipun ada sedikit peningkatan kesalahan pada data pengujian, model ini masih dapat memperbaiki akurasi dan generalisasinya, dengan *MAPE* yang lebih rendah pada data pengujian dibandingkan data pelatihan. Penurunan *MAPE* ini mengindikasikan bahwa meskipun model menghadapi beberapa peningkatan kesalahan, ia tetap lebih mampu dalam menjaga tingkat ketepatan yang lebih baik dan konsisten pada data pengujian. Dengan kenaikan yang relatif moderat pada metrik evaluasi dan penurunan *MAPE*, Model 3 menjadi model yang paling stabil dan efektif dalam menjaga keseimbangan antara akurasi pada data pelatihan dan kemampuannya untuk menggeneralisasi pada data pengujian, menjadikannya pilihan terbaik di antara model lainnya.

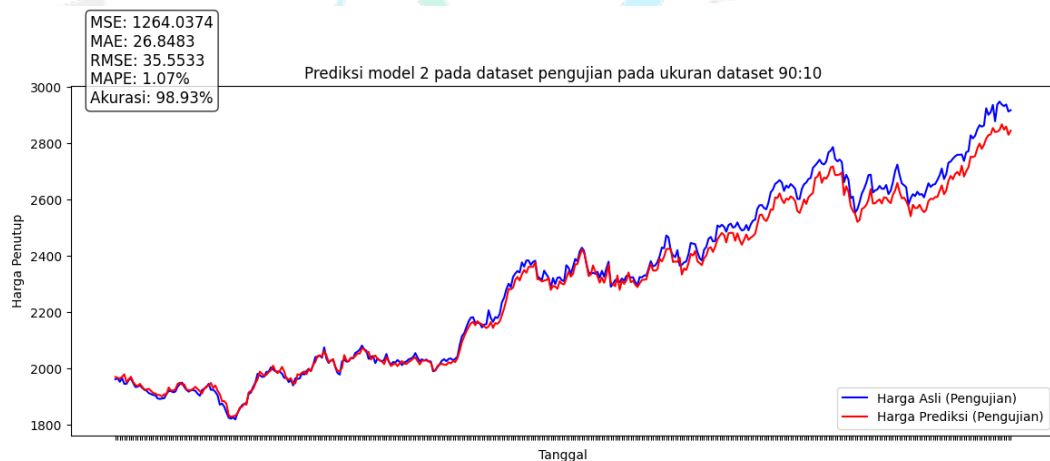


Gambar 5. 19 Prediksi model 3 pada *dataset* pengujian dengan ukuran 80:20

Model 4 menunjukkan adanya *overfitting* yang cukup jelas, meskipun masih lebih stabil dibandingkan dengan Model 1 dan Model 2. Kinerja model pada data pelatihan jauh lebih baik dibandingkan dengan data pengujian, yang terlihat dari kenaikan yang signifikan pada metrik evaluasi. Untuk *MSE*, terjadi kenaikan sebesar 1.468,20%, menunjukkan bahwa meskipun model mampu mempelajari data pelatihan dengan baik, performanya menurun saat diterapkan pada data pengujian. Begitu pula dengan *MAE*, yang mengalami kenaikan sebesar 297,66%, mengindikasikan bahwa prediksi model pada data pengujian lebih buruk dibandingkan dengan data pelatihan. Kenaikan juga terjadi pada *RMSE*, dengan peningkatan sebesar 295,97%, dan *MAPE* yang meningkat sebesar 155,56%, semakin memperlihatkan bahwa meskipun model belajar dengan baik pada data pelatihan, ia kesulitan dalam menggeneralisasi pada data pengujian. Meskipun

demikian, Model 4 menunjukkan peningkatan yang lebih moderat pada metrik ini dibandingkan dengan Model 1 dan Model 2, menjadikannya sedikit lebih stabil meskipun masih ada tanda-tanda *overfitting* yang cukup jelas.

Model 5, meskipun masih menunjukkan indikasi *overfitting*, menunjukkan performa yang lebih baik dibandingkan dengan Model 1, Model 2, dan Model 4. Kenaikan pada metrik evaluasi antara data pelatihan dan data pengujian cukup signifikan, tetapi tidak sebesar pada model-model sebelumnya. Untuk *MSE*, terjadi kenaikan sebesar 653,26%, yang menunjukkan bahwa meskipun model mampu mempelajari data pelatihan dengan baik, performanya menurun pada data pengujian. *MAE* mengalami kenaikan sebesar 160,17%, yang mengindikasikan bahwa prediksi model pada data pengujian lebih buruk dibandingkan dengan data pelatihan. *RMSE* mengalami peningkatan sebesar 174,52%, dan *MAPE* meningkat sebesar 62,12%, yang menunjukkan bahwa model masih "menghafal" data pelatihan namun kesulitan saat diuji pada data pengujian. Namun, kenaikan pada metrik ini lebih moderat dibandingkan model-model lainnya selain Model 3, dan dengan *MAPE* yang relatif rendah pada data pengujian, Model 5 menunjukkan kinerja yang cukup baik dalam menjaga akurasi pada data pengujian meskipun tetap ada sedikit penurunan dibandingkan data pelatihan.



Gambar 5. 20 Prediksi Model 5 pada *dataset* pengujian dengan ukuran 80:20

C. Hasil pelatihan model dengan kumpulan data ukuran 90:10

Tabel 5. 16 Tabel hasil pelatihan dan pengujian dengan kumpulan data 90:10

Mode l	<i>MSE</i> pelatiha n	<i>RMSE</i> pelatiha n	<i>MAE</i> pelatiha n	<i>MAPE</i> pelatiha n	<i>MSE</i> pengujia n	<i>RMSE</i> pengujia n	<i>MAE</i> pengujia n	<i>MAPE</i> pengujia n
Mode 11	359,31	18,96	16,45	1,03%	389,71	15,30	19,74	0,60%
Mode 12	74,45	8,63	6,03	0,36%	165,77	9,36	12,88	0,37%
Mode 13	204,09	14,29	11,67	0,70%	265,81	12,85	16,30	0,51%
Mode 14	108,81	10,43	7,51	0,45%	205,09	10,65	14,32	0,42%
Mode 15	79,24	8,90	6,41	0,39%	232,15	15,24	11,95	0,46%

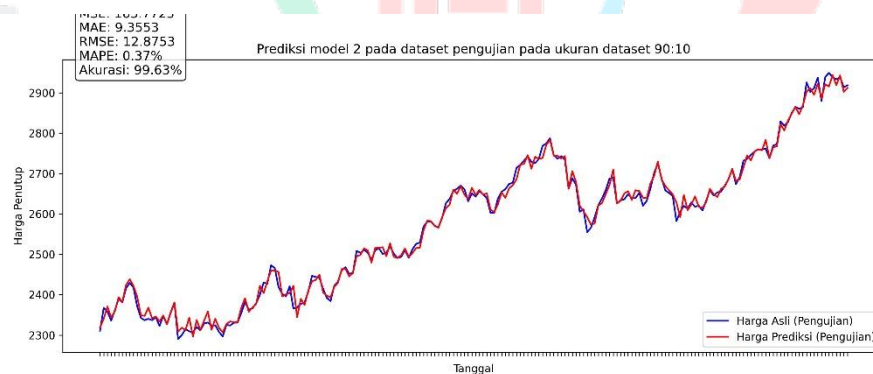
Tabel 5. 17 Tabel kenaikan dari pelatihan ke pengujian ukuran *dataset* 90:10

Model	<i>MSE</i> pelatihan→ <i>MSE</i> pengujian	<i>RMSE</i> pelatihan→ <i>RMSE</i> pengujian	<i>MAE</i> pelatihan→ <i>MAE</i> pengujian	<i>MAPE</i> pelatihan→ <i>MAPE</i> pengujian
Model 1	8,46%	-19,30%	20,00%	-41,75%
Model 2	122,66%	8,46%	113,60%	2,78%
Model 3	30,24%	-10,08%	39,67%	-27,14%
Model 4	88,48%	2,11%	90,68%	-6,67%
Model 5	192,97%	71,24%	86,43%	17,95%

Model 1 dengan *dataset* 90:10 menunjukkan adanya peningkatan yang signifikan pada metrik evaluasi, meskipun tidak separah pada *dataset* 80:20. Kinerja model pada data pelatihan lebih baik dibandingkan dengan data pengujian, dengan kenaikan *MSE* sebesar 8,46%, yang menunjukkan bahwa model dapat mempelajari data pelatihan dengan baik tetapi kesulitan saat diterapkan pada data pengujian. Begitu pula dengan *MAE* yang mengalami kenaikan sebesar 20,00%, yang menandakan bahwa model lebih akurat pada data pelatihan, namun prediksinya agak buruk pada data pengujian. *RMSE* mengalami penurunan sebesar -19,30%, dan *MAPE* justru menurun sebesar -41,75%, menunjukkan bahwa meskipun ada sedikit penurunan kesalahan, model ini lebih mampu menggeneralisasi dengan lebih baik pada data pengujian dibandingkan dengan data pelatihan. Dengan penurunan pada *RMSE* dan *MAPE*, Model 1 menunjukkan

perbaikan dalam kemampuan generalisasi meskipun masih ada tanda-tanda *overfitting* yang harus diperhatikan.

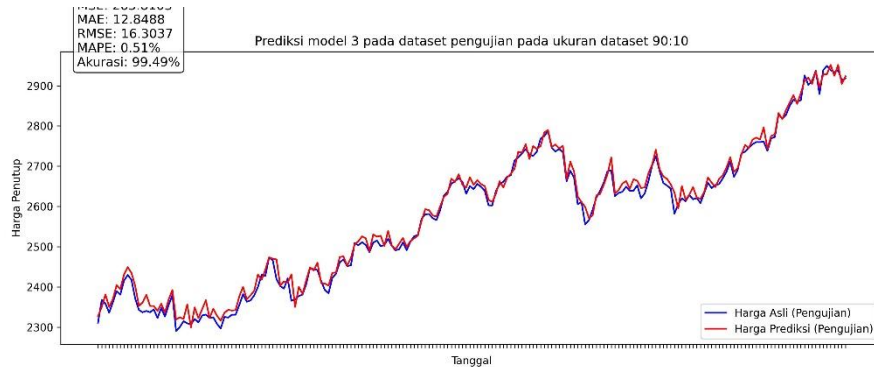
Model 2 menunjukkan adanya *overfitting* yang lebih moderat dibandingkan dengan Model 1, meskipun masih terdapat perbedaan signifikan antara kinerja pada data pelatihan dan data pengujian. Kenaikan *MSE* sebesar 122,66% menunjukkan bahwa meskipun model belajar dengan sangat baik pada data pelatihan, performanya menurun pada data pengujian. Begitu pula dengan *MAE* yang mengalami kenaikan sebesar 113,60%, yang mengindikasikan bahwa model lebih akurat pada data pelatihan, namun hasil prediksinya pada data pengujian lebih buruk. *RMSE* mengalami kenaikan sebesar 8,46%, dan *MAPE* yang sedikit meningkat sebesar 2,78%. Hal ini menunjukkan bahwa meskipun ada sedikit peningkatan pada kesalahan pada data pengujian, Model 2 mampu mempertahankan akurasi yang baik, dengan perbedaan antara data pelatihan dan data pengujian yang tidak terlalu tajam, menjadikannya model yang cukup stabil meskipun masih menunjukkan sedikit *overfitting*. Meskipun demikian, Model 2 merupakan model yang menghasilkan nilai terkecil dibanding model lainnya.



Gambar 5. 21 Prediksi Model 2 pada *dataset* pengujian dengan ukuran 90:10

Model 3 menunjukkan hasil yang lebih stabil dibandingkan dengan Model 1 dan Model 2, meskipun masih ada indikasi *overfitting*. Untuk *MSE*, terjadi kenaikan sebesar 30,24%, yang masih relatif rendah dibandingkan dengan model lainnya, menunjukkan bahwa model dapat mempertahankan kinerjanya pada data pengujian meskipun tidak sebaik pada data pelatihan. *MAE* mengalami kenaikan sebesar 39,67%, yang mengindikasikan sedikit penurunan akurasi pada data pengujian dibandingkan data pelatihan. *RMSE* mengalami penurunan sebesar -

10,08%, dan *MAPE* juga menurun sebesar -27,14%, yang menunjukkan bahwa meskipun terjadi sedikit peningkatan kesalahan pada data pengujian, model ini lebih baik dalam menggeneralisasi dan mempertahankan akurasi yang lebih konsisten. Penurunan *MAPE* ini mengindikasikan bahwa Model 3 mampu menjaga akurasi yang lebih baik pada data pengujian, menjadikannya pilihan yang lebih stabil di antara model lainnya.



Gambar 5. 22 Prediksi Model 3 pada *dataset* pengujian dengan ukuran 90:10

Model 4 menunjukkan adanya peningkatan yang moderat pada metrik evaluasi antara data pelatihan dan data pengujian, dengan *MSE* meningkat sebesar 88,48%, yang menunjukkan penurunan kinerja model pada data pengujian meskipun masih cukup stabil pada data pelatihan. *MAE* mengalami kenaikan sebesar 90,68%, yang menandakan bahwa prediksi model lebih buruk pada data pengujian, meskipun ada peningkatan yang lebih moderat dibandingkan dengan model lainnya. *RMSE* mengalami kenaikan sebesar 2,11%, dan *MAPE* mengalami penurunan sebesar -6,67%, yang menunjukkan bahwa meskipun ada sedikit peningkatan kesalahan pada data pengujian, model ini cukup baik dalam mempertahankan akurasi pada data pengujian. Dengan kenaikan yang lebih terkendali pada metrik evaluasi dan penurunan *MAPE*, Model 4 menunjukkan performa yang cukup stabil meskipun masih ada sedikit *overfitting*.

Model 5 menunjukkan adanya *overfitting* yang lebih jelas dibandingkan Model 4, meskipun tetap cukup stabil dalam hal kinerja. *MSE* mengalami kenaikan yang signifikan sebesar 192,97%, menunjukkan penurunan performa pada data pengujian. *MAE* mengalami kenaikan sebesar 86,43%, yang menunjukkan bahwa model lebih akurat pada data pelatihan, namun hasil prediksinya lebih buruk pada

data pengujian. *RMSE* mengalami kenaikan sebesar 71,24%, dan *MAPE* mengalami kenaikan sebesar 17,95%, yang mengindikasikan bahwa model mengalami kesulitan saat menggeneralisasi pada data pengujian meskipun tetap dapat mempertahankan akurasi yang cukup baik. Dengan kenaikan pada metrik evaluasi yang lebih besar dibandingkan Model 4, Model 5 masih menunjukkan adanya peningkatan pada prediksi data pengujian, tetapi kurang stabil dibandingkan dengan model lainnya.

5.1.2.2. Hasil Prediksi Model *LSTM*

Pada sub subbab ini akan memaparkan hasil prediksi model *LSTM* dengan data tahunan berfrekuensi harian selama lima tahun terakhir, dengan kata lain terdapat lima kali pengujian. Lima kali pengujian ini terdiri dari awal tahun 2020 hingga akhir tahun 2020 sampai awal tahun 2024 hingga akhir tahun 2024. Model-model yang digunakan merupakan model terbaik dari ketiga ukuran *dataset*, sehingga jumlah model yang diprediksi berjumlah tiga buah. . Perwakilan model terbaik dari ukuran *dataset* 70:30 adalah Model 4, dalam pengujian analisis *time-series* ini Model 4 dinamakan Model 1. Model 2 dalam pengujian ini merupakan perwakilan model terbaik dari ukuran *dataset* 80:20, yaitu Model 3. Model 3 atau model ketiga dalam pengujian ini merupakan perwakilan model terbaik dari ukuran *dataset* 90:10, yaitu Model 2. Proses analisis yang dilakukan selama pengujian analisis *time-series* ini dengan membandingkan kenaikan tingkat kesalahan atau selisih serta akurasi. Hasil prediksi model ini mencakup nilai evaluasi yaitu *MSE*, *RMSE*, *MAE*, dan *MAPE*, serta persentase kenaikan dari setiap nilai evaluasi per pengujian dengan tahun yang berbeda.

Tabel 5. 18 Tabel hasil prediksi dengan model 1 pada pengujian tahunan

Metrik/ Tahun	awal 2020 - akhir 2020	awal 2021 - akhir 2021	awal 2022 - akhir 2022	awal 2023 - akhir 2023	awal 2024 - akhir 2024
<i>MSE</i>	446,91	155,75	346,46	154,99	3193,08
<i>MAE</i>	17,33	9,84	16,07	9,73	49,07
<i>RMSE</i>	21,14	12,48	18,61	12,45	56,51
<i>MAPE</i>	0,99%	0,55%	0,90%	0,50%	1,95%

Tabel 5. 19 Tabel persentase kenaikan relatif pada prediksi Model 1

Metrik/ Tahun	2020→2021	2021→2022	2022→2023	2023→2024
<i>MSE</i>	-34,76%	33,10%	-51,77%	1552,68%
<i>MAE</i>	-15,35%	14,50%	-36,21%	316,37%
<i>MAE</i>	-19,23%	15,37%	-30,55%	306,53%
<i>MAPE</i>	-16,04%	14,61%	-41,18%	215,00%

Secara keseluruhan, Model 1 menunjukkan performa yang cukup baik pada tahun 2021 dan 2023, namun mengalami penurunan yang signifikan pada tahun 2024. Pada tahun 2021, sebagian besar metrik, termasuk *MSE* yang turun sebesar 34,76%, *MAE* yang turun -15,35%, *RMSE* yang turun -19,23%, dan *MAPE* yang turun -16,04%, menunjukkan peningkatan akurasi Model 1. Namun, pada 2022, terjadi penurunan performa, yang tercermin dari peningkatan *MSE* sebesar 33,10%, *MAE* yang meningkat 14,50%, *RMSE* yang meningkat 15,37%, dan *MAPE* yang meningkat 14,61%, meskipun perbaikan kembali terlihat pada 2023 dengan penurunan *MSE* sebesar -51,77%, *MAE* -36,21%, *RMSE* -30,55%, dan *MAPE* -41,18%. Lonjakan besar pada 2024, dengan kenaikan yang sangat drastis di semua metrik, seperti *MSE* yang meningkat sebesar 1552,68%, *MAE* yang meningkat 316,37%, *RMSE* yang meningkat 306,53%, dan *MAPE* yang meningkat 215,00%, menunjukkan adanya masalah yang sangat serius pada Model 1.

Tabel 5. 20 Tabel hasil prediksi dengan Model 2

Metrik/ Tahun	awal 2020 - akhir 2020	awal 2021 - akhir 2021	awal 2022 - akhir 2022	awal 2023 - akhir 2023	awal 2024 - akhir 2024
<i>MSE</i>	840,49	377,38	1475,15	1236,87	465,19
<i>MAE</i>	26,19	17,73	36,85	33,82	18,45
<i>RMSE</i>	28,99	19,43	38,41	35,17	21,57
<i>MAPE</i>	1,48%	0,99%	2,07%	1,74%	0,77%

Tabel 5. 21 Tabel persentase kenaikan relatif pada prediksi Model 2

Metrik/ Tahun	2020→2021	2021→2022	2022→2023	2023→2024
<i>MSE</i>	-55,10%	290,89%	-16,15%	-62,39%
<i>MAE</i>	-32,30%	107,80%	-8,22%	-45,45%

<i>MAE</i>	-32,99%	97,71%	-8,43%	-38,67%
<i>MAPE</i>	-33,11%	109,09%	-15,94%	-55,75%

Model 2 menunjukkan fluktuasi performa yang cukup signifikan sepanjang periode yang diamati, dengan beberapa penurunan dan peningkatan yang cukup mencolok pada masing-masing metrik. Pada tahun 2021, model menunjukkan perbaikan yang jelas dibandingkan tahun 2020, terlihat dari penurunan *MSE* sebesar -55,10%, *MAE* yang turun -32,30%, *RMSE* yang menurun -32,99%, dan *MAPE* yang turun -33,11%. Namun, pada tahun 2022, terjadi lonjakan besar pada *MSE* yang meningkat 290,89%, serta peningkatan *MAE* sebesar 107,80%, *RMSE* yang meningkat 97,71%, dan *MAPE* yang meningkat 109,09%, menunjukkan penurunan performa yang sangat signifikan. Tahun 2023 memberikan sedikit perbaikan, dengan *MSE* yang turun -16,15%, *MAE* yang turun -8,22%, *RMSE* yang turun -8,43%, dan *MAPE* yang turun -15,94%, menunjukkan bahwa model mulai kembali menunjukkan hasil yang lebih baik. Namun, pada tahun 2024, model mengalami penurunan yang signifikan lagi, dengan *MSE* turun -62,39%, *MAE* turun -45,45%, *RMSE* turun -38,67%, dan *MAPE* turun -55,75%, meskipun secara keseluruhan tetap lebih baik dibandingkan dengan 2022.

Tabel 5. 22 Tabel hasil prediksi dengan Model 3

Metrik/ Tahun	awal 2020 - akhir 2020	awal 2021 - akhir 2021	awal 2022 - akhir 2022	awal 2023 - akhir 2023	awal 2024 - akhir 2024
<i>MSE</i>	149,24	64,43	105,75	86,66	160,98
<i>MAE</i>	8,79	6,03	7,71	6,71	9,07
<i>RMSE</i>	12,22	8,03	10,28	9,31	12,69
<i>MAPE</i>	0,50%	0,34%	0,43%	0,34%	0,37%

Tabel 5. 23 Tabel persentase kenaikan relatif pada prediksi Model 3

Metrik/ Tahun	2020→2021	2021→2022	2022→2023	2023→2024
<i>MSE</i>	-56,83%	64,12%	-18,05%	85,76%
<i>MAE</i>	-31,43%	27,95%	-12,98%	35,08%
<i>MAE</i>	-34,29%	28,11%	-9,47%	36,29%
<i>MAPE</i>	-32,00%	26,47%	-20,93%	8,82%

Model 3 menunjukkan kinerja yang relatif stabil dengan beberapa fluktuasi yang terlihat jelas pada metrik yang digunakan, meskipun ada perbaikan yang cukup baik di beberapa tahun. Pada tahun 2021, model menunjukkan penurunan yang signifikan dalam *MSE* sebesar -56,83%, *MAE* turun -31,43%, *RMSE* turun -34,29%, dan *MAPE* turun -32,00%, yang menandakan perbaikan akurasi model dibandingkan dengan tahun 2020. Namun, pada tahun 2022, model mengalami peningkatan *MSE* sebesar 64,12%, serta kenaikan pada *MAE* 27,95%, *RMSE* 28,11%, dan *MAPE* 26,47%, yang mengindikasikan penurunan performa yang signifikan dibandingkan tahun sebelumnya. Setelah lonjakan di 2022, model kembali menunjukkan perbaikan pada 2023, dengan penurunan *MSE* sebesar -18,05%, *MAE* turun -12,98%, *RMSE* turun -9,47%, dan *MAPE* turun -20,93%, yang menunjukkan pemulihan kinerja. Namun, pada tahun 2024, meskipun metrik-metrik tersebut kembali menunjukkan penurunan, ada lonjakan besar pada *MSE* yang meningkat sebesar 85,76%, *MAE* meningkat 35,08%, *RMSE* meningkat 36,29%, dan *MAPE* meningkat 8,82%. Hal ini menunjukkan adanya penurunan performa model yang signifikan, yang perlu dianalisis lebih lanjut untuk mengetahui penyebabnya, apakah terkait dengan perubahan data atau potensi kesalahan lainnya.

Secara keseluruhan, Model 3 menunjukkan fluktuasi yang cukup besar dalam performanya, dengan perbaikan yang terlihat pada 2021 dan 2023, namun penurunan signifikan terjadi pada 2022 dan 2024. Fluktuasi ini mengindikasikan bahwa meskipun model mampu beradaptasi dengan baik di beberapa titik, ada kemungkinan bahwa faktor eksternal seperti perubahan dalam data atau pengaturan model mempengaruhi kinerjanya secara tidak konsisten.

Tabel 5. 24 Tabel hasil prediksi tahunan pada tiga buah model

Metrik / Tahun	Model 1 (2020-2024)	Model 2 (2020-2024)	Model 3 (2020-2024)
2020			
<i>MSE</i>	446,91	840,49	149,24
<i>MAE</i>	17,33	26,19	8,79
<i>RMSE</i>	21,14	28,99	12,22
<i>MAPE</i>	0,99%	1,48%	0,50%
2020→2021			

<i>MSE</i>	-65,15%	-55,10%	-56,83%
<i>MAE</i>	-43,26%	-32,30%	-31,43%
<i>RMSE</i>	-40,97%	-32,99%	-34,29%
<i>MAPE</i>	-44,44%	-33,11%	-32,00%
2021→2022			
<i>MSE</i>	122,44%	290,89%	64,12%
<i>MAE</i>	63,39%	107,80%	27,95%
<i>RMSE</i>	49,14%	97,71%	28,11%
<i>MAPE</i>	63,64%	109,09%	26,47%
2022→2023			
<i>MSE</i>	-55,26%	-16,15%	-18,05%
<i>MAE</i>	-39,48%	-8,22%	-12,98%
<i>RMSE</i>	-33,12%	-8,43%	-9,47%
<i>MAPE</i>	-44,44%	-15,94%	-20,93%
2023→2024			
<i>MSE</i>	1960,21%	-62,39%	85,76%
<i>MAE</i>	404,55%	-45,45%	35,08%
<i>RMSE</i>	353,90%	-38,67%	36,29%
<i>MAPE</i>	290,00%	-55,75%	8,82%

Secara keseluruhan, Model 3 menunjukkan performa yang paling konsisten dan stabil dibandingkan dengan Model 1 dan Model 2, meskipun ada beberapa fluktuasi dalam hasilnya. Pada tahun 2020, Model 3 memiliki nilai *MSE*, *MAE*, *RMSE*, dan *MAPE* yang jauh lebih rendah daripada kedua model lainnya, menunjukkan akurasi yang lebih baik di awal periode. Model 1 dan Model 2 memiliki nilai *MSE* yang jauh lebih tinggi pada tahun 2020, dengan Model 2 bahkan mencapai *MSE* sebesar 840,49, jauh lebih besar dibandingkan Model 3 yang hanya 149,24. Model 1 juga menunjukkan performa yang lebih baik dibandingkan Model 2 pada 2020, tetapi masih jauh di belakang Model 3 dalam hal akurasi.

Pada tahun 2021, ketiga model mengalami perbaikan signifikan, dengan Model 1 menunjukkan penurunan yang paling tajam pada *MSE* -65,15% dan *MAE* -43,26%. Meskipun demikian, Model 2 dan Model 3 juga mengalami perbaikan, meskipun tidak sebesar Model 1. Namun, ketika memasuki tahun 2022, terjadi lonjakan besar dalam *MSE*, *MAE*, *RMSE*, dan *MAPE* untuk semua model. Model 2 mengalami lonjakan terbesar, dengan *MSE* meningkat sebesar 290,89%, sementara

Model 1 juga mengalami peningkatan *MSE* sebesar 122,44%. Model 3 mengalami lonjakan lebih kecil dengan *MSE* meningkat 64,12%, tetapi tetap menunjukkan peningkatan yang lebih moderat dibandingkan dua model lainnya.

Pada tahun 2023, ketiga model menunjukkan pemulihan yang lebih baik dibandingkan tahun 2022. Model 1 mengalami penurunan *MSE* yang signifikan -55,26%, *MAE* -39,48%, dan *RMSE* -33,12%, yang menunjukkan perbaikan performa yang cukup solid. Model 2 dan Model 3 juga menunjukkan penurunan, meskipun lebih kecil daripada Model 1. Namun, pada tahun 2024, ada lonjakan besar yang sangat signifikan pada Model 1, dengan *MSE* meningkat sebesar 1960,21%, *MAE* meningkat 404,55%, *RMSE* meningkat 353,90%, dan *MAPE* meningkat 290,00%. Hal ini menunjukkan penurunan performa yang sangat besar dan menunjukkan bahwa Model 1 mengalami masalah serius di tahun 2024.

Sementara itu, Model 2 dan Model 3 menunjukkan penurunan yang lebih moderat pada tahun 2024. Model 2 mengalami penurunan yang sangat besar pada *MSE* -62,39%, *MAE* -45,45%, dan *MAPE* -55,75%, yang menunjukkan pemulihan yang baik setelah lonjakan besar di tahun 2022. Model 3, meskipun ada lonjakan kecil pada *MSE* 85,76% dan *MAE* 35,08%, tetap menunjukkan performa yang lebih stabil dibandingkan dengan Model 1, dengan *MAPE* yang tetap rendah 8,82%.

5.2. Pembahasan

Dalam penelitian analisis *time-series* untuk prediksi harga emas menggunakan pendekatan *DL* metode *LSTM*, dilakukan berbagai proses untuk dapat mengoptimalkan analisis prediktif dengan pemodelan *LSTM*. Proses-proses seperti pra-pemrosesan data yang meliputi seleksi fitur, normalisasi, pembuatan *dataset* hingga pelatihan model *LSTM* serta analisa prediksi.

Dimulai dengan seleksi fitur, seleksi fitur merupakan salah satu tahap yang sangat penting dalam penelitian ini karena menentukan fitur-fitur yang paling relevan dengan pergerakan harga emas, sehingga model prediktif dapat bekerja dengan lebih akurat. Proses seleksi fitur diawali dengan analisis statistik deskriptif, visualisasi distribusi data, serta uji korelasi *Pearson* untuk mengevaluasi hubungan

linear antar variabel dengan harga penutupan emas. Hasil analisis menunjukkan bahwa variabel volume perdagangan dari berbagai aset memiliki variabilitas yang sangat tinggi dan distribusi yang tidak konsisten. Selain itu, uji korelasi *Pearson* mengungkapkan bahwa volume memiliki hubungan yang sangat lemah atau bahkan negatif terhadap harga penutupan emas, yang menandakan bahwa peningkatan jumlah transaksi tidak selalu beriringan dengan perubahan harga emas. Oleh karena itu, penelitian ini memutuskan untuk mengecualikan variabel volume dari analisis prediktif guna memastikan bahwa model hanya menggunakan fitur yang memiliki keterkaitan erat dengan pergerakan harga.

Setelah seleksi fitur, dilakukan normalisasi data dengan teknik *Min-Max Scaling* untuk memastikan bahwa seluruh fitur berada dalam rentang yang seragam [0,1]. Normalisasi ini berfungsi untuk menghilangkan perbedaan skala antar variabel agar model *LSTM* dapat lebih efektif dalam menangkap pola pergerakan harga tanpa dipengaruhi oleh perbedaan rentang nilai yang terlalu besar. Selanjutnya, pembuatan *dataset* dilakukan dengan pendekatan *sliding window* berukuran 30 hari, yang berarti setiap observasi untuk prediksi harga emas menggunakan informasi dari 30 hari sebelumnya. Hal ini bertujuan agar model dapat memahami pola historis pergerakan harga dan memanfaatkannya dalam analisis *time-series*. Selain itu, *dataset* dibagi menjadi tiga skenario berdasarkan proporsi data pelatihan dan pengujian, yaitu 70:30, 80:20, dan 90:10.

Dengan kombinasi seleksi fitur yang tepat, normalisasi data, dan strategi pembuatan *dataset* yang mempertimbangkan ketergantungan temporal dalam data *time-series*, penelitian ini memastikan bahwa model *LSTM* memiliki fondasi yang kuat untuk menangkap pola pergerakan harga emas dengan lebih optimal. Langkah-langkah ini bertujuan untuk meningkatkan akurasi prediksi dan memberikan pemahaman yang lebih mendalam terhadap faktor-faktor yang memengaruhi pergerakan harga emas dalam periode waktu tertentu.

Pembahasan selanjutnya membahas mengenai hasil pelatihan serta pengujian dari model *LSTM* dengan ukuran *dataset* 70:30. Berdasarkan dari lima buah model dengan spesifikasi berbeda, Model 4 dan Model 3 merupakan model yang terbaik untuk ukuran *dataset* 70:30. Meskipun kedua model ini merupakan

model terbaik, hasil *MSE* menunjukkan bahwa kedua model ini mengalami *overfitting* dengan nilai *MSE* pada model 4 sebesar 2018,39. Ini menandakan bahwa model terlalu menghafal data pelatihan sehingga kinerjanya menurun pada data pengujian. Namun, berdasarkan Gambar 5. 17 dan Gambar 5. 18, dapat disimpulkan bahwa data pengujian ini memiliki outlier pada harga emas di atas yang terbukti oleh Gambar 5. 6 bahwa adanya outlier pada harga tinggi emas. Ini memperkuat bahwa baik Model 3 dan Model 4 kurang mempelajari dinamika harga emas yang terdapat outlier sehingga performanya kurang baik.

Selanjutnya, berdasarkan dari lima buah model dengan spesifikasi berbeda dan ukuran *dataset* 80:20, Model 3 dan Model 5 merupakan model terbaik. Meskipun kedua model ini merupakan model terbaik, hasil *MSE* menunjukkan bahwa model masih mengalami *overfitting* atau keadaan model terlalu menghafal data pelatihan sehingga kurang baik performanya dalam memprediksi data pada data pengujian. Namun, Model 3 merupakan model terbaik yang memiliki *MSE* pengujian sebesar 770,34 dan model yang terstabil dibandingkan model lainnya dalam ukuran *dataset* 80:20. Jika dilihat dari hasil prediksi data pengujian pada Gambar 5. 19 dan Gambar 5. 20, model masih kurang baik dalam menganalisis harga emas yang mengalami outlier. Hal ini dapat disimpulkan bahwa dalam data pelatihan ini kurang menggambarkan dinamika harga emas yang cukup untuk menangkap pergerakan harga emas yang fluktuatif.

Berdasarkan hasil pelatihan lima buah model dengan spesifikasi yang berbeda pada ukuran *dataset* 90:10, Model 2 dan Model 4 merupakan model yang terbaik bahkan untuk semua model pada ukuran *dataset* yang berbeda. Model 2 merupakan model terbaik dengan *MSE* pengujian sebesar 165,77 dan estimasi akurasi 99,63%. Jika dilihat hasil prediksi kedua model ini pada data pengujian yaitu Gambar 5. 21 dan Gambar 5. 22, dapat disimpulkan bahwa kedua model dapat menangkap pola-pola pergerakan harga emas yang fluktuatif. Dapat disimpulkan bahwa pada data pelatihan ini memiliki referensi yang cukup untuk menggeneralisasi serta menangkap pola-pola pergerakan harga emas yang fluktuatif.