

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pasar modal Indonesia telah menunjukkan pertumbuhan positif yang konsisten dalam beberapa tahun terakhir, yang dikaitkan dengan inisiatif pendidikan dan sosialisasi yang berkelanjutan oleh Bursa Efek Indonesia (BEI). Pada tahun 2024, jumlah investor pasar modal di Indonesia yang mencakup saham, obligasi, dan reksadana meningkat 22% menjadi 14,84 juta dari 12,17 juta pada tahun sebelumnya (IDX, 2024). Investasi pasar modal di Indonesia, khususnya pada instrumen saham, menawarkan berbagai kategori karakteristik risiko dan imbal hasil yang berbeda. Salah satu kategori saham yang banyak diminati adalah saham *bluechip*, yaitu saham yang memiliki fundamental yang kuat, stabil, dan memiliki pasar yang besar.

Perusahaan *bluechip* dikenal memiliki ketangguhan operasional yang tinggi, terbukti mampu melewati masa-masa sulit dan tetap mendapatkan keuntungan di tengah kondisi ekonomi yang tidak baik (Basyith dkk., 2015). Kemampuan ini menjadi dasar rekam jejak pertumbuhan yang konsisten dan dapat diandalkan dalam jangka panjang. Dengan kapitalisasi pasar yang sangat besar, saham *bluechip* sangat relevan untuk dijadikan investasi jangka panjang dan sumber pendapatan berkelanjutan (Lubis dkk., 2020). Karakteristik emiten besar, likuid, dan fundamental kuat ini tercermin dalam Indeks LQ45 yang diterbitkan oleh IDX. Daya tarik investasi dan stabilitas yang ditawarkan oleh saham *bluechip* mendorong peningkatan partisipasi investor, sehingga menciptakan kebutuhan akan alat prediksi yang andal bagi pengambil keputusan.

Pendekatan dalam peramalan harga saham atau indeks secara umum dapat dikategorikan ke dalam empat kelompok utama, yakni (i) pendekatan tradisional (fundamental dan teknikal), (ii) pendekatan ekonometrik seperti *Autoregressive Integrated Moving Average* (ARIMA) yang efektif untuk pola linear, (iii) pendekatan kecerdasan buatan (*Artificial Intelligence / AI*) seperti *Artificial Neural Network* (ANN), *Recurrent Neural Network* (RNN), dan *Long Short-Term Memory* (LSTM) yang unggul dalam menangkap hubungan non-linier, serta (iv) pendekatan hibrida yang menggabungkan dua atau lebih metode untuk memperoleh keunggulan komplementer. Berbagai penelitian menunjukkan bahwa model tunggal seperti ARIMA sering kali kurang mampu menangani pola non-linier yang kompleks sementara model AI murni cenderung rentan terhadap *overfitting* pada data finansial yang sangat bising (Zhang dkk., 1998). Oleh karena itu, pendekatan *hybrid* semakin banyak digunakan karena mampu menggabungkan keunggulan metodologis dari masing-masing model untuk meningkatkan akurasi peramalan (Bates & Granger, 1969).

Berdasarkan hal tersebut, peramalan dilakukan melalui pendekatan *hybrid* yang menggabungkan *Wavelet Transform* sebagai teknik pra-pemrosesan dan *Long Short-Term Memory* (LSTM) sebagai model prediksi utama. *Wavelet Transform* digunakan untuk memecah deret waktu ke dalam domain waktu frekuensi sehingga

memungkinkan pemisahan komponen frekuensi tinggi dan rendah guna melakukan *denoising* dan ekstraksi fitur multiskala. Sementara itu, LSTM, sebagai varian dari *Recurrent Neural Network* (RNN), memiliki kemampuan menangkap ketergantungan jangka panjang pada data sekuensial melalui mekanisme *gates* yang efektif mengatasi masalah *vanishing gradient* (Hochreiter & Schmidhuber, 1997). *Wavelet Transform* sangat sesuai untuk data runtun waktu karena mampu melakukan dekomposisi pada setiap level untuk mengekstraksi fitur yang relevan bagi prediksi (Alima dkk., 2024), sedangkan LSTM berfungsi mempelajari dinamika non-linier dari hasil dekomposisi tersebut. Dengan demikian, kombinasi keduanya menjadikan model *Wavelet Transform*–LSTM bersifat komplementer dan berpotensi memberikan hasil peramalan yang lebih akurat dibandingkan model tunggal.

Berbagai studi empiris telah menunjukkan efektivitas integrasi *Wavelet Transform* dengan model *deep learning* dalam meningkatkan akurasi peramalan harga saham. Misalnya, penelitian (Bao dkk., 2017) menemukan bahwa kombinasi *Wavelet Transform*–LSTM mampu meningkatkan performa prediksi harga saham secara signifikan, sedangkan (Jarrah & Salim, 2019) melaporkan bahwa model *Discrete Wavelet Transform*–RNN memiliki kinerja yang lebih baik dibandingkan ARIMA pada pasar finansial. Meskipun demikian, kesenjangan penelitian masih ditemukan pada konteks pasar modal Indonesia, khususnya terkait penerapan model *Wavelet Transform*–LSTM pada kumpulan saham *bluechip* lintas sektor. Sebagian besar penelitian terdahulu berfokus pada pasar luar negeri atau indeks spesifik, serta menggunakan horizon pengujian yang relatif sempit. Oleh karena itu, model *Wavelet Transform*–LSTM dipilih untuk menguji tingkat akurasi dalam melakukan peramalan harga saham *bluechip* di Indonesia sebagai representasi dari stabilitas dan dinamika pasar domestik.

Berdasarkan uraian latar belakang dan penelitian sebelumnya, peneliti tertarik menggali potensi penggunaan *Wavelet Transform* pada model *Long Short-Term Memory* (LSTM) dalam penelitian berjudul **“Penggunaan *Wavelet Transform* dalam Model *Long Short-Term Memory* (LSTM) pada Prediksi Harga Saham *Bluechip*.”** Penelitian ini diharapkan memperkaya literatur terkait penerapan *Wavelet Transform* pada LSTM untuk prediksi harga saham dan memberikan wawasan bagi investor dalam meningkatkan akurasi keputusan investasi saham *bluechip* di pasar modal.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka permasalahan pada penelitian ini adalah bagaimana akurasi *Wavelet Transform* pada Model *Long Short-Term Memory* (LSTM) dalam prediksi harga saham *bluechip* BBRI, ASII, TLKM, INDF, dan ANTM menggunakan data harga penutupan saham harian ?

1.3 Tujuan Penelitian

Penelitian ini bertujuan untuk menerapkan model *Hybrid Wavelet Transform* – LSTM, dan melakukan evaluasi hasil peramalan harga saham *bluechip* BBRI, ASII, TLKM, INDF, dan ANTM menggunakan data harga penutupan saham dengan menggunakan *Wavelet Transform* pada Model *Long Short-Term Memory* (LSTM).

1.4 Landasan Teori

1.4.1 Saham

Saham merupakan bukti kepemilikan nilai dalam suatu perusahaan. Investor yang memiliki saham pada perusahaan tertentu secara otomatis memiliki hak untuk mengontrol atas manajemennya (Hartono, 2022). Menurut Bursa Efek Indonesia (BEI), Saham merupakan salah satu instrumen keuangan yang paling banyak diminati. Bagi perusahaan, penerbitan saham merupakan salah satu langkah untuk pendanaan modal jangka panjang untuk perusahaan. Di sisi lain, saham dipilih oleh investor karena saham mampu memberikan tingkat keuntungan yang lebih menarik. Investor yang membeli atau memiliki saham berhak mendapatkan keuntungan dividen dan *capital gain*. *Capital gain* adalah keuntungan modal dari penjualan aset (saham). Melalui *capital gain*, investor banyak meraih keuntungan karena transaksi jual beli yang dilakukan di pasar modal. Secara umum, instrumen saham diklasifikasikan menjadi dua kategori, yaitu saham preferen dan saham biasa.

Saham preferen merupakan kombinasi antara saham dan hutang. Saham ini memiliki sifat sebagai saham karena mempunyai hak kepemilikan perusahaan dan tidak ada waktu jatuh tempo dan mempunyai sifat utang karena mendapatkan dividen tetap yang sudah pasti (Hartono, 2022). Maka dari itu, pemegang saham preferen ini memiliki keistimewaan, yaitu didahulukan dalam pemberian dividen. Namun, pemegang saham preferen tidak memiliki hak untuk mengikuti Rapat Umum Pemegang Saham (RUPS) (Suratna dkk., 2020). Saham Biasa merupakan saham yang menunjukkan kepemilikan perusahaan. Sebagai pemegang saham biasa, investor memiliki hak untuk memilih direksi dan hak untuk melakukan veto atas putusan – putusan perusahaan. Pemegang saham berhak mendapatkan dividen yang dibayarkan dari hasil laba dan bersedia menanggung risiko kerugian yang diderita perusahaan (Suratna dkk., 2020).

1.4.2 Saham Bluechip

Saham *bluechip* adalah saham perusahaan besar yang menonjol karena memiliki pertumbuhan laba yang konsisten serta ditopang oleh kemampuan finansial yang mapan untuk membagikan dividen secara berkelanjutan. Karakteristik saham *bluechip* muncul dari status perusahaan yang dikenal secara nasional, memiliki dasar finansial sehat, dan menghasilkan produk atau layanan berkualitas tinggi yang diterima luas oleh pasar (Basyith dkk., 2015). Berdasarkan pencapaian dan kinerja yang unggul, Bursa Efek Indonesia (IDX) mengklasifikasikan dan memberikan urutan tertinggi kepada perusahaan-perusahaan ini sebagai kategori saham *bluechip*, yang dianggap memiliki tata kelola yang baik dan fundamental bisnis yang baik (Dewa & Sundari, 2025; Widyarti dkk., 2021). Kelompok saham *bluechip* dengan likuiditas tertinggi di pasar kemudian diwakili oleh Indeks LQ45, yang terdiri dari 45 saham yang dipilih setiap enam bulan berdasarkan kriteria kapitalisasi pasar, likuiditas, dan kinerja keuangan, menjadikannya tolok ukur bagi perusahaan yang mematuhi praktik bisnis yang transparan (Basyith dkk., 2015; Dewa & Sundari, 2025).

Saham-saham *bluechip* tersebut umumnya tersebar di berbagai sektor strategis yang menopang perekonomian nasional. Sektor perbankan memegang peranan vital sebagai lembaga intermediasi keuangan yang menghimpun dana dari masyarakat dan menyalurkannya kembali dalam bentuk kredit serta layanan jasa keuangan lainnya guna mendorong aktivitas ekonomi (Garnia dkk., 2024). Selain itu, terdapat sektor konglomerasi atau perusahaan induk multiindustri yang menjalankan strategi diversifikasi dengan mengelola berbagai lini bisnis berbeda di bawah satu naungan grup perusahaan untuk meminimalkan risiko operasional (Putra dkk., 2021). Di sektor telekomunikasi berfokus pada penyediaan layanan jaringan, data, dan konektivitas digital yang telah menjadi kebutuhan pokok masyarakat modern (Purba dkk., 2021). Stabilitas pasar juga didukung oleh sektor barang konsumen (*consumer goods*), yaitu kelompok perusahaan yang memproduksi kebutuhan sehari-hari seperti makanan, minuman, dan produk perawatan diri dengan tingkat permintaan yang cenderung stabil dan tahan terhadap fluktuasi ekonomi. Sementara itu, sektor pertambangan mencakup perusahaan yang bergerak dalam eksplorasi dan ekstraksi sumber daya alam, seperti batu bara, mineral, dan logam yang kinerjanya sangat dipengaruhi oleh harga komoditas global namun sering kali memberikan kontribusi kapitalisasi pasar yang signifikan (Thorbecke, 2023).

1.4.3 Wavelet Transform

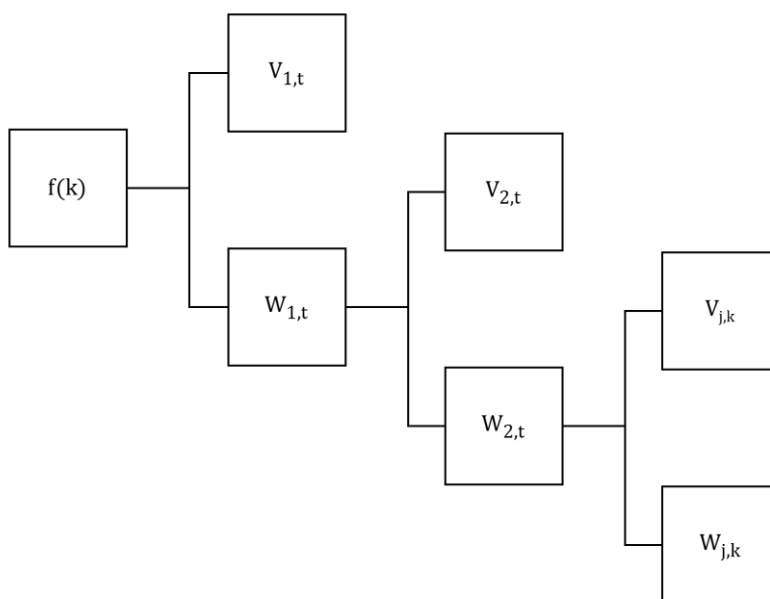
Banyak data deret waktu ekonomi dan keuangan memiliki sifat non-stasioner, penuh dengan *noise*, dan menunjukkan variasi temporal dalam frekuensi yang kompleks. Kondisi ini sering kali menyulitkan model prediksi konvensional maupun jaringan saraf tiruan untuk menangkap pola data yang sebenarnya. Dalam konteks ini, *Wavelet Transform* menawarkan solusi efektif sebagai metode pra-pemrosesan data yang mampu menangani sinyal non-stasioner dan non-linier secara simultan (Liang dkk., 2019).

Secara definisi, *wavelet* adalah gelombang kecil yang berosilasi dalam rentang waktu tertentu dan menunjukkan fluktuasi sepanjang durasi yang ditentukan (Aprilianti dkk., 2022). Tidak seperti analisis deret waktu tradisional yang sering memerlukan diferensiasi data untuk mencapai stasioneritas, yang berisiko menghilangkan memori jangka panjang. Analisis *wavelet* memungkinkan dekomposisi sinyal ke dalam berbagai skala resolusi tanpa mengubah integritas informasi aslinya. Pendekatan ini memungkinkan pemisahan antara struktur dasar data (tren) dan gangguan acak (*noise*), sehingga model prediksi seperti LSTM dapat belajar dari data yang lebih bersih dan terstruktur (Jarrah, 2024; Yang, 2022).

Wavelet Transform merupakan alternatif yang lebih unggul dibandingkan Transformasi Fourier (Ismail & Asfour, 1999). Jika *Fourier* hanya mampu menganalisis sinyal dalam domain frekuensi dan kehilangan informasi waktu, *Wavelet Transform* mampu melakukan analisis *Time-Frequency Localization*. Artinya, metode ini dapat mengungkap kapan dan di frekuensi mana suatu peristiwa atau lonjakan harga terjadi, sebuah fitur yang sangat krusial dalam analisis pasar saham yang dinamis (Tang dkk., 2021). Keuntungan ini memungkinkan analisis lokal

yang mendetail, mengungkap tren diskontinuitas, serta mengompresi dan menghilangkan *noise* pada sinyal tanpa kehilangan energi sinyal yang signifikan.

Secara umum, transformasi ini dikategorikan menjadi dua jenis, yaitu *Continuous Wavelet Transform* (CWT) dan *Discrete Wavelet Transform* (DWT). Dalam penerapannya pada data keuangan, DWT lebih sering digunakan karena efisiensinya dalam komputasi dan kemampuannya untuk mendekomposisi sinyal tanpa redundansi informasi (Jarrah, 2024). Oleh karena itu, penelitian ini menerapkan DWT untuk memecah data harga saham.



Gambar 1. Dekomposisi dari *Discrete Wavelet Transform*

Gambar 1 merupakan proses dekomposisi DWT bekerja dengan memecah sinyal asli $f(k)$ menjadi dua komponen utama, yakni koefisien aproksimasi ($W_{j,k}$) dan koefisien detail ($V_{j,k}$). Komponen aproksimasi berisi informasi frekuensi rendah yang merepresentasikan tren utama pergerakan harga, sedangkan komponen detail berisi informasi frekuensi tinggi yang menangkap fluktuasi cepat atau *noise* pasar. Sinyal aproksimasi kemudian dapat didekomposisi lagi ke level berikutnya. Dalam metode *Wavelet transform*, berbagai jenis keluarga *wavelet* bisa digunakan, seperti *Haar*, *Meyer*, *Daubechies*, *Morlet*, dan *Coiflets*. Penelitian ini menggunakan keluarga *wavelet Daubechies*, khususnya *Daubechies 4*). Pemilihan *Daubechies 4* didasarkan pada karakteristiknya yang memiliki kemiripan bentuk dengan sinyal harga saham yang tidak teratur namun memiliki pola tren, serta kemampuannya yang efektif dalam menghaluskan data tanpa merusak fitur penting dari sinyal asli (Alima dkk., 2024).

Secara matematis, konstruksi *wavelet Daubechies 4* memiliki dua filter, yaitu filter *wavelet* dan filter skala. Agar koefisien filter *wavelet* dapat dikonstruksi dengan benar,

koefisien filternya harus memenuhi dua syarat, yaitu sifat ortonormalitas dan sifat *vanishing moment*. Kedua kondisi tersebut dinyatakan dalam persamaan (1) dan (2).

$$\sum_{l=0}^3 h_l h_{l-2k} = \delta(k) \quad (1)$$

$$\sum_{l=0}^3 (-1)^l l^p h_l = 0 \quad (2)$$

Fungsi delta *Kronecker* yang dinotasikan sebagai $\delta(k)$, didefinisikan bernilai 1 apabila $k = 0$ dan bernilai 0 apabila $k \neq 0$. Definisi ini menjadi dasar penyelesaian persamaan (1), di mana dengan menyubstitusi nilai $k = 0$ dan $k = 1$, diperoleh hubungan dasar antar koefisien filter. Selanjutnya untuk melengkapi persyaratan matematis pembentukan filter *wavelet*, digunakan *vanishing moment* yang dinyatakan secara spesifik pada persamaan (2). Dengan menyubstitusi nilai $p = 0$ dan $p = 1$ ke dalam persamaan tersebut, maka akan didapatkan hasilnya sebagai berikut

$$h_0^2 + h_1^2 + h_2^2 + h_3^2 = 1 \quad (3.a)$$

$$h_0 h_2 + h_1 h_3 = 0 \quad (3.b)$$

$$h_0 - h_1 + h_2 - h_3 = 0 \quad (3.c)$$

$$-h_1 + 2h_2 - 3h_3 = 0 \quad (3.d)$$

Penyederhanaan proses aljabar dilakukan dengan memisalkan notasi koefisien menjadi $h_0 = a$, $h_1 = b$, $h_2 = c$, dan $h_3 = d$. Langkah pertama penyelesaian adalah mencari hubungan koefisien tersebut dengan mengkuadratkan persamaan (3.c). Mengingat bahwa $a^2 + b^2 + c^2 + d^2 = 1$ dan $ac + bd = 0$, maka perhitungannya sebagai berikut.

$$(a - b + c - d)^2 = a^2 + b^2 + c^2 + d^2 + 2(-ab + ac - ad - bc + bd - cd)$$

$$(a - b + c - d)^2 = 1 + 2(ab + bd) - 2(ab + ad + bc + cd)$$

$$0 = 1 + 0 - 2(ab + ad + bc + cd)$$

$$2(ab + ad + bc + cd) = 1$$

Langkah selanjutnya adalah menyubstitusikan persamaan $2(ab - ad - bc - cd) = 1$ ke dalam persamaan $(a + b + c + d)^2$.

$$(a + b + c + d)^2 = a^2 + b^2 + c^2 + d^2 + 2(ab + ac + ad + bc + bd + cd)$$

$$(a + b + c + d)^2 = a^2 + b^2 + c^2 + d^2 + 2(ab + ad + bc + cd)$$

$$(a + b + c + d)^2 = 1 + 1$$

$$(a + b + c + d)^2 = 2 \Rightarrow a + b + c + d = \pm\sqrt{2} \quad (4)$$

Persamaan (4) menjadi kunci untuk mengevaluasi variabel satu persatu. Dengan mengombinasikan persamaan (4) dan persamaan (3) Selanjutnya menambahkan persamaan (4) dengan persamaan (3.c) dan mendapatkan nilai variabel a .

$$a + c = \frac{\sqrt{2}}{2}$$

Setelah mendapatkan nilai dari persamaan $a + c$ dan mengelompokkan persamaan (3) menjadi $(a + c) - (b + d)$, maka kita mendapatkan hasil dari $b + d$ sebagai berikut.

$$b + d = \frac{\sqrt{2}}{2}$$

Hasil dari persamaan $b + d$ disubstitusi ke persamaan (3.d) dan mendapatkan nilai variabel c .

$$c - d = \frac{\sqrt{2}}{4}$$

$$d = c - \frac{\sqrt{2}}{4}$$

Nilai dari persamaan $b + d$ dan persamaan $c - d$ di gunakan untuk mencari persamaan $b + c$ sehingga dapat menghasilkan nilai variabel b .

$$b + c = \frac{3\sqrt{2}}{4}$$

$$b = \frac{3\sqrt{2}}{4} - c$$

Nilai – nilai dari variabel a, b , dan d , kemudian disubstitusikan ke dalam persamaan (3.a) untuk menghasilkan nilai dari variabel c yang kemudian akan menjadi koefisien filter wavelet.

$$a^2 + b^2 + c^2 + d^2 = 1$$

$$\left(\frac{\sqrt{2}}{2} - c\right)^2 + \left(\frac{3\sqrt{2}}{4} - c\right)^2 + (c)^2 + \left(c - \frac{\sqrt{2}}{4}\right)^2 = 1$$

$$4c^2 - 3\sqrt{2}c + \frac{7}{4} - 1 = 0$$

$$\left(c - \frac{3\sqrt{2}}{8}\right)^2 - \frac{18}{64} + \frac{12}{64} = 0$$

$$\begin{aligned}\left(c - \frac{3\sqrt{2}}{8}\right)^2 &= \frac{6}{64} \\ c &= \frac{3\sqrt{2}}{8} \pm \sqrt{\frac{6}{64}} \\ c &= \frac{3 \pm \sqrt{3}}{4\sqrt{2}}\end{aligned}$$

Nilai c merepresentasikan koefisien h_2 . Dengan menyubstitusikan nilai c ke variabel a, b , dan d , maka nilai untuk seluruh koefisien lainnya dapat ditentukan. Adapun hasil dari perhitungan koefisien h_0, h_1, h_2 dan h_3 adalah sebagai berikut.

$$h_0 = \frac{1 + \sqrt{3}}{4\sqrt{2}}, h_1 = \frac{3 + \sqrt{3}}{4\sqrt{2}}, h_2 = \frac{3 - \sqrt{3}}{4\sqrt{2}}, h_3 = \frac{1 - \sqrt{3}}{4\sqrt{2}} \quad (5)$$

Filter *wavelet* (h_l) dengan indeks $l = 0, 1, \dots, L - 1$ dibentuk dari suatu bilangan bulat positif yang didefinisikan bahwa L adalah panjang filter. Untuk *Daubechies* 4, panjang filter didefinisikan sebesar $L = 4$, sehingga terdapat empat koefisien filter *wavelet* yakni h_0, h_1, h_2 dan h_3 . Agar koefisien – koefisien ini dapat membentuk basis *wavelet* yang ortogonal dan memiliki dukungan kompak, maka harus dipenuhi syarat ortogonalitas dan normalisasi momen. Persyaratan matematis tersebut dinyatakan dalam persamaan berikut.

$$h_0 + h_1 + h_2 + h_3 = \frac{1+\sqrt{3}}{4\sqrt{2}} + \frac{3+\sqrt{3}}{4\sqrt{2}} + \frac{3-\sqrt{3}}{4\sqrt{2}} + \frac{1-\sqrt{3}}{4\sqrt{2}} = \sqrt{2} \quad (6.a)$$

$$h_0^2 + h_1^2 + h_2^2 + h_3^2 = \left(\frac{1+\sqrt{3}}{4\sqrt{2}}\right)^2 + \left(\frac{3+\sqrt{3}}{4\sqrt{2}}\right)^2 + \left(\frac{3-\sqrt{3}}{4\sqrt{2}}\right)^2 + \left(\frac{1-\sqrt{3}}{4\sqrt{2}}\right)^2 = 1 \quad (6.b)$$

$$h_0 h_2 + h_1 h_3 = \left(\frac{1+\sqrt{3}}{4\sqrt{2}} \times \frac{3-\sqrt{3}}{4\sqrt{2}}\right) + \left(\frac{3+\sqrt{3}}{4\sqrt{2}} \times \frac{1-\sqrt{3}}{4\sqrt{2}}\right) = 0 \quad (6.c)$$

Selanjutnya, filter skala merupakan deret bernilai cacah positif yang dinotasikan sebagai (g_l) dengan indeks ($l = 0, 1, 2, \dots, L - 1$) yang di mana L menyatakan lebar filter yang merupakan bilangan bulat positif. Filter skala dapat diperoleh menggunakan persamaan (7).

$$g_l = (-1)^l h_{L-1-l} \quad l = 0, 1, 2, 3. \quad (7)$$

Berdasarkan persamaan tersebut, maka diperoleh nilai koefisien filter skala untuk g_0, g_1, g_2 dan g_3 sebagai berikut.

$$\begin{aligned}g_0 &= h_3 = \frac{1 - \sqrt{3}}{4\sqrt{2}}, & g_1 &= -h_2 = -\frac{3 - \sqrt{3}}{4\sqrt{2}}, \\ g_2 &= h_1 = \frac{3 + \sqrt{3}}{4\sqrt{2}}, & g_3 &= -h_0 = -\frac{1 + \sqrt{3}}{4\sqrt{2}}\end{aligned} \quad (8)$$

Sama halnya dengan filter *wavelet*, koefisien filter skala juga harus memenuhi sifat ortogonalitas dan normalisasi energi agar rekonstruksi sinyal dapat dilakukan dengan sempurna. Sifat-sifat tersebut adalah:

$$g_0 + g_1 + g_2 + g_3 = \frac{1-\sqrt{3}}{4\sqrt{2}} + \left(-\frac{3-\sqrt{3}}{4\sqrt{2}}\right) + \frac{3+\sqrt{3}}{4\sqrt{2}} + \left(-\frac{1+\sqrt{3}}{4\sqrt{2}}\right) = 0 \quad (9.a)$$

$$g_0^2 + g_1^2 + g_2^2 + g_3^2 = \left(\frac{1-\sqrt{3}}{4\sqrt{2}}\right)^2 + \left(-\frac{3-\sqrt{3}}{4\sqrt{2}}\right)^2 + \left(\frac{3+\sqrt{3}}{4\sqrt{2}}\right)^2 + \left(-\frac{1+\sqrt{3}}{4\sqrt{2}}\right)^2 = 1 \quad (9.b)$$

$$g_0g_2 + g_1g_3 = \left(\frac{1-\sqrt{3}}{4\sqrt{2}} \times \frac{3+\sqrt{3}}{4\sqrt{2}}\right) + \left(-\frac{3-\sqrt{3}}{4\sqrt{2}} \times -\frac{1+\sqrt{3}}{4\sqrt{2}}\right) = 0 \quad (9.c)$$

Dalam implementasi *Discrete Wavelet Transform* (DWT), jumlah data input (N) idealnya merupakan perpangkatan dari dua ($N = 2^J$), dengan J adalah bilangan bulat. Proses dekomposisi dilakukan dengan mengonvolusikan sinyal *input* X terhadap filter wavelet dan filter skala, yang kemudian diikuti oleh proses *downsampling* sebesar $\frac{N}{2}$. Rumus perhitungan untuk mendapatkan koefisien wavelet ($W_{j,k}$) dan koefisien skala ($V_{j,k}$) pada level j dan translasi k sebagai berikut.

$$W_{j,k} = \sum_{l=0}^{L-1} h_l W_{j+1,2k+l \bmod N} \quad (10)$$

$$V_{j,k} = \sum_{l=0}^{L-1} g_l V_{j+1,2k+l \bmod N} \quad (11)$$

Dengan $k = 0, 1, 2, \dots, \frac{N}{2} - 1$. Di mana koefisien $W_{j,k}$ disebut koefisien *wavelet* pada level ke j dan pada waktu ke t . Koefisien *wavelet* merupakan komponen frekuensi rendah yang menangkap tren keseluruhan dalam deret waktu. Sedangkan, $V_{j,k}$ disebut koefisien skala pada level ke j dan pada waktu ke t . Koefisien detail merupakan komponen frekuensi tinggi yang menangkap variasi jangka pendek, lonjakan, dan tepi data. Setelah sinyal didekomposisi dan diproses, sinyal asli dapat dibentuk kembali ke domain waktu melalui proses rekonstruksi atau *Inverse Discrete Wavelet Transform* (IDWT). Fungsi $f(k)$ merupakan kumpulan dari koefisien $W_{j,k}$ dan $V_{j,k}$ untuk $j = 1, 2, \dots, J$. Berikut rumus rekonstruksi sinyal $f(k)$ dari koefisien tersebut yang ditunjukkan pada persamaan (12).

$$f(k) = \sum_{l=0}^{L-1} W_l \tilde{h}_{k-2l} + \sum_{l=0}^{L-1} V_l \tilde{g}_{k-2l} \quad (12)$$

Persamaan tersebut menunjukkan bahwa sinyal asli $f(k)$ merupakan superposisi dari komponen aproksimasi W_j pada level terakhir dan akumulasi komponen detail V_j dari seluruh level dekomposisi. Adapun $\tilde{h}$ dan $\tilde{g}$ merepresentasikan koefisien filter rekonstruksi (*synthesis filter*) atau filter invers,

yang digunakan untuk mengembalikan sinyal dari domain frekuensi waktu kembali ke domain waktu aslinya.

1.4.4 Normalisasi Data

Min-max Scaling adalah salah satu metode normalisasi yang paling terkenal. Dengan menerapkan metode tersebut akan mengubah data ke dalam rentang nilai [0, 1] (Nugraha et al., 2024), artinya nilai minimum dan maksimum suatu variabel akan menjadi 0 dan 1 masing-masing. Tujuannya adalah untuk membuat variabel diukur pada skala yang sama sehingga setiap variabel memiliki kontribusi yang sama dalam proses penyesuaian model. (Yang, 2022) menyatakan bahwa model yang dilatih pada data yang diskalakan cenderung memiliki kinerja yang lebih baik daripada yang dilatih pada data yang tidak diskalakan, dan ANN memerlukan normalitas input untuk mempercepat dan menstabilkan proses pembelajaran. Rumus matematika dari normalisasi *min - max* yang digunakan adalah:

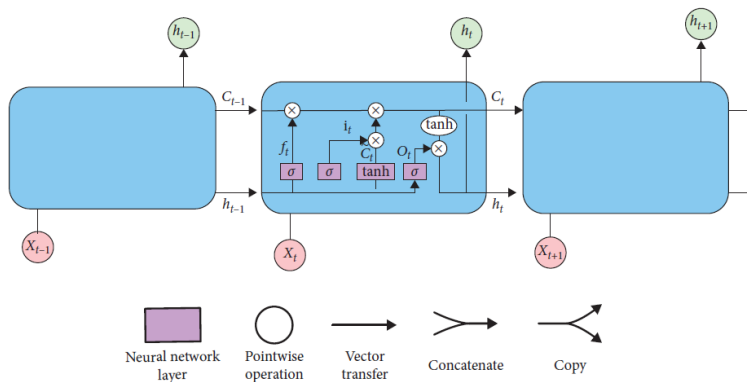
$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (13)$$

Nilai hasil normalisasi adalah x' , x adalah nilai aktual, x_{min} adalah nilai terkecil dari data aktual, dan x_{max} adalah nilai terbesar dari data aktual.

1.4.5 Long Short-Term Memory

Long Short-Term Memory (LSTM) adalah *Recurrent Neural Network* (RNN) atau jenis jaringan saraf rekursif yang dirancang untuk menangani masalah dalam pengolahan data urutan panjang. LSTM pertama kali diusulkan oleh Hochreiter dan Schmidhuber pada tahun 1997, sebagai pengembangan dari RNN untuk mengatasi keterbatasan dalam mengingat informasi dari waktu yang jauh pada urutan panjang (Nugraha et al., 2024).

LSTM terdiri dari tiga gerbang utama: *forget gate*, *input gate*, dan *output gate*. Setiap gerbang ini berfungsi untuk mengontrol aliran informasi dalam jaringan, dengan *forget gate* yang memutuskan informasi mana yang harus dilupakan dari memori, *input gate* yang mengatur informasi mana yang akan diterima dan disimpan dalam sel, serta *output gate* yang menentukan informasi mana yang akan diteruskan ke jaringan berikutnya (Tang et al., 2021). Struktur ini memungkinkan LSTM untuk mempertahankan informasi penting dalam waktu yang lama, sementara mengabaikan informasi yang tidak relevan.



Gambar 2. Arsitektur dari LSTM

Sumber: (Tang dkk., 2021)

LSTM sangat efektif dalam mengatasi tantangan pengolahan data yang melibatkan urutan panjang, karena kemampuannya untuk menyimpan dan memodifikasi informasi dalam jangka waktu yang lama. Hal ini menjadikannya pilihan utama dalam aplikasi-aplikasi seperti pengenalan suara otomatis, pemrosesan bahasa alami, dan analisis keuangan, di mana informasi dalam urutan sebelumnya memiliki pengaruh terhadap prediksi di masa depan (Liang dkk., 2019). Dengan adanya mekanisme gerbang yang canggih, LSTM mampu menangani data berurutan dan memodelkan ketergantungan jangka panjang yang ada pada data tersebut (Jarrah, 2024).

Selain itu, pemilihan fungsi aktivasi yang tepat sangat penting dalam pelatihan LSTM, seperti penggunaan fungsi aktivasi *sigmoid* dan *tanh* yang membantu jaringan untuk memproses faktor *non-linier* dalam data (Liang dkk., 2019). Keunggulan ini menjadikan LSTM lebih unggul dibandingkan dengan model RNN biasa, yang lebih terbatas dalam mengelola informasi jangka panjang. Fungsi *sigmoid* memiliki keluaran antara 0 dan 1, jika keluarannya adalah 0, fungsi aktivasi tidak aktif dan jika keluarannya adalah 1, fungsi aktivasi aktif. Fungsi *tanh* memiliki nilai antara -1 hingga 1. Persamaan fungsi *sigmoid* dan *tanh* dapat dilihat pada persamaan (14) dan (15).

$$\sigma(x) = \frac{1}{(1 + e^{-x})} \quad (14)$$

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (15)$$

Adapun rumus perhitungan untuk setiap gerbang adalah sebagai berikut. Untuk gerbang pertama, yaitu *forget gate*:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (16)$$

Di mana f_t adalah *forget gate*, x_t adalah nilai dari waktu t , σ mewakili fungsi aktivasi *sigmoid*, W_f adalah *weight* dari *forget gate*, h_{t-1} adalah nilai *output* sebelum waktu t , dan b_f adalah bias dari *forget gate*.

Setelah itu, gerbang *input* berfungsi untuk memilih informasi baru yang akan disimpan dalam *cell state*. Langkah ini akan terbagi menjadi dua bagian, yaitu bagian pertama adalah lapisan *sigmoid* yang disebut *input gate* dan bagian kedua adalah lapisan *tanh*. Lapisan *sigmoid* menentukan nilai yang akan diperbarui, sedangkan lapisan *tanh* menciptakan vektor baru yang akan ditambahkan ke *cell state*. Rumus *input gate* dapat dilihat pada persamaan (17) dan (18).

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (17)$$

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (18)$$

i_t adalah *input gate*, W_i merupakan *weight* dari *input gate*, b_i adalah bias dari *input gate*, $\tilde{C}_t$ nilai baru yang akan ditambahkan ke *cell state*, *tanh* adalah fungsi aktivasi *tahn*, W_c adalah *weight* untuk *cell state*, dan b_c merupakan nilai bias pada *cell state*. Selain itu, keluaran *forget gate* (f_t) dikalikan dengan nilai *cell state* yang lama dan dijumlahkan dengan hasil perkalian i_t dan $\tilde{C}_t$ seperti pada persamaan (19).

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \quad (19)$$

C_t adalah *cell state* order ke- t , dan C_{t-1} adalah *cell state* sebelum order ke- t . Gerbang terakhir, yaitu *output gate*, berfungsi untuk menentukan bagian dari keadaan sel yang akan dikeluarkan dengan menggunakan lapisan *sigmoid*. Setelah itu, keadaan sel diproses melalui *tanh*, yang menghasilkan nilai dalam rentang -1 hingga 1.

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (20)$$

$$h_t = o_t \cdot \tanh(C_t) \quad (21)$$

o_t adalah *output gate*, W_o adalah *weight* dari *output gate*, dan b_o adalah bias dari *output gate*.

1.4.6 Evaluasi Model

Root Mean Square Error (RMSE) dan *Mean Absolute Percentage Error* (MAPE) digunakan sebagai indikator akurasi untuk mengevaluasi kinerja model pada set uji dan memvalidasi kemampuan generalisasi. MAPE mengukur kesalahan relatif (persentase) sehingga memudahkan perbandingan lintas saham/periode harga yang berbeda, sedangkan RMSE memberi bobot lebih besar pada kesalahan besar

karena pengkuadratan, sehingga tepat ketika kesalahan ekstrem perlu ditekan. MAPE mudah sederhana untuk ditafsirkan, namun perlu kehati-hatian saat nilai aktualnya hampir nol (Hyndman & Koehler, 2006). Selain itu, RMSE dipilih karena tidak ditentukan oleh ambang ukuran sampel tertentu, melainkan oleh sifat penalti terhadap *error* besar dan asumsi galat yang mendekati *Gaussian* (Chai & Draxler, 2014).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (22)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100 \quad (23)$$

Dengan y_i adalah nilai aktual dan $\hat{y}_i$ adalah prediksi yang dihasilkan oleh model peramalan, serta n adalah jumlah prediksi. Adapun pedoman interpretasi MAPE yang sering dipakai sebagai berikut.

Tabel 1. Kategori MAPE

MAPE (X)	Kategori Peramalan
$X < 10\%$	Sangat Baik
$10\% \leq X < 20\%$	Baik
$20\% \leq X < 50\%$	Cukup
$X \geq 50\%$	Buruk

BAB II

METODOLOGI PENELITIAN

2.1 Pendekatan dan Jenis Penelitian

Penelitian ini menggunakan metodologi kuantitatif, yang berfokus pada analisis data numerik untuk menghasilkan temuan yang dapat diukur secara objektif. Metodologi penelitian yang digunakan adalah studi kasus kuantitatif yang menganalisis data harga saham historis dari lima perusahaan *bluechip* yang tercatat di Bursa Efek Indonesia (BEI) dan masuk ke dalam indeks LQ45. Objek penelitian ini meliputi lima emiten, yaitu BBRI, ASII, TLKM, INDF, dan ANTM yang dipilih melalui teknik *purposive sampling* dengan kriteria pemimpin pasar yang memiliki karakteristik unik dan dinamika pada sektornya masing – masing.

1. BBRI dipilih untuk merepresentasikan sektor perbankan karena BBRI merupakan bank dengan total aset terbesar dan memiliki jaringan pelayanan yang luas. Selain itu, likuiditas transaksi harian yang sangat tinggi sehingga menjadikan representasi utama pada sektor perbankan.
2. ASII dipilih untuk mewakili sektor konglomerasi karena ASII memiliki diversifikasi bisnis yang luas sehingga tidak bergantung pada satu segmen saja. Dominasi pada industri otomotif menjadikan saham ini indikator penting bagi kondisi ekonomi rill.
3. TLKM dipilih mewakili sektor telekomunikasi, Sebagai perusahaan yang masuk ke dalam BUMN dengan kapitalisasi pasar terbesar di sektornya, TLKM memiliki keunggulan dalam dukungan permodalan untuk ekspansi infrastruktur telekomunikasi dan transformasi bisnis yang agresif di bandingkan dengan kompetitornya.
4. INDF dipilih untuk mewakili sektor barang konsumen karena memiliki dominasi pasar yang signifikan pada segmen produk makanan kemasan, didukung oleh ekuitas merek yang sangat kuat, terutama pada produk mie instan melalui merek Indomie.
5. ANTM dipilih untuk mewakili sektor pertambangan, Pemilihan emiten ini didorong oleh lonjakan emas dunia akibat ketidakpastian geopolitik, serta nikel yang permintaannya meningkat akibat produksi baterai untuk kendaraan listrik, sehingga memperkuat posisi nikelnya. Faktor tersebut menjadikan pergerakan saham ANTM memiliki volatilitas tinggi.

Secara keseluruhan, penelitian ini bertujuan untuk memodelkan dan memprediksi harga saham perusahaan-perusahaan tersebut melalui analisis data keuangan yang relevan untuk mengidentifikasi tren dan faktor yang mempengaruhi fluktuasi harga saham.

2.2 Data Penelitian

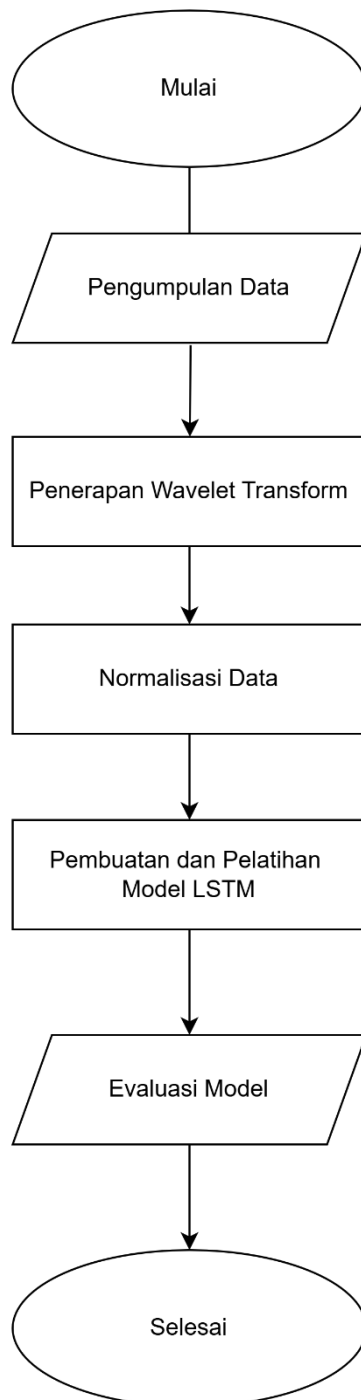
Penelitian ini menggunakan data sekunder berupa data harga penutupan harian saham yang diperoleh dari situs web <https://finance.yahoo.com>. Data tersebut mencakup rentang waktu pengamatan mulai tanggal 26 Oktober 2016 hingga 31 Januari 2025. Pemilihan periode ini bertujuan untuk merekam dinamika pasar yang beragam, mulai dari bencana pandemi COVID-19 dan pemulihan pasca pandemi COVID-19, ketidakpastian geopolitik global, hingga sentimen pasar domestik yang dipengaruhi oleh pergantian presiden. Data historis saham dari Yahoo Finance tidak hanya memuat variabel Harga Penutupan, tetapi juga informasi rinci seperti Harga Pembukaan, Harga Tertinggi, Harga Terendah, Harga Penutupan Yang Disesuaikan), dan Volume. Yahoo Finance dipilih sebagai sumber data karena keandalannya, kemudahan akses, dan kepraktisannya untuk keperluan penelitian.

2.3 Metode Analisis Data

Keseluruhan proses analisis data menggunakan bahasa pemrograman *Python* di Google Colab. Metodologi penelitian mencakup langkah-langkah sistematis sebagai berikut:

1. Mengumpulkan data saham dari situs web Yahoo Finance.
2. Menerapkan *Wavelet Transform* untuk memisahkan data menjadi komponen frekuensi yang berbeda.
3. Melakukan Normalisasi komponen *wavelet transform* dengan menggunakan *Min-Max Scaler*.
4. Membangun dan melatih model *Long Short Term-Memory* (LSTM) dengan data yang telah diproses
5. Melakukan evaluasi akurasi peramalan model menggunakan *Root Mean Square Error* (RMSE) dan *Mean Absolute Percentage Error* (MAPE)

2.4 Alur Kerja



Gambar 3. Alur Kerja *Hybrid* WT - LSTM