

## BAB 2

### LANDASAN TEORI

Pada bab ini berisi definisi dari konsep atau teori yang menjadi dasar atau pondasi didalam penelitian ini. Landasan Teori menjelaskan mengenai *Time-Series Forecasting* yang menjadi permasalahan utama dalam penelitian ini, *Deep-Learning* sebagai alternatif untuk mengatasi permasalahan, *Informer* yang termasuk *deep-learning* sebagai model utama, metode persiapan data, RMSE dan MAE sebagai metrik evaluasi standar dalam mengukur performa model, serta *Walk-Forward Cross Validation* salah satu metode *Cross-Validation* khusus pada *time-series forecasting*. Serta studi literatur mengenai penelitian terdahulu.

#### 2.1 Landasan Teori

##### 2.1.1 *Time-series Forecasting*

Inti dari *Time-series forecasting* (TSF) pada dasarnya mengacu pada pemanfaatan data historis dari variabel untuk memprediksi nilai masa depan dari variabel target [15]. Data *time-series* memiliki karakteristik yang unik seperti

- *Trend* merupakan arah pergerakan data secara keseluruhan dalam jangka panjang, dimana arah dapat cenderung naik, turun, atau stationer.
- *Seasonality* merupakan pola fluktuasi yang berulang secara teratur dalam periode waktu tertentu, seperti mingguan, bulanan, atau tahunan.
- *Cyclical Patterns* merupakan pola fluktuasi yang terjadi secara berulang, namun dalam periode waktu berbeda, biasa dipengaruhi oleh siklus bisnis atau ekonomi.
- *Noise* merupakan variasi acak atau gangguan yang tidak memiliki pola tertentu dan tidak dapat diprediksi.

Karakteristik data *time-series* yang unik menimbulkan tantangan tersendiri dalam pengembangan TSF. Model machine learning konvensional seperti ARIMA [9] banyak diterapkan sebagai model TSF, namun memiliki keterbatasan pada pola non-linear dan prediksi jangka panjang. Keterbatasan model konvensional dalam mengatasi masalah TSF mendorong pengembangan *machine-learning* secara lebih

mandalam, hingga *deep-learning* diperkenalkan. Penerapan DL dalam TSF menjadi solusi aktif yang memberikan hasil yang signifikan. Pendekatan DL berbasis RNNs seperti LSTM, GRU ([16], [17]) telah terbukti menunjukkan performa yang baik dalam menangkap pola jangka panjang. Namun, mengalami penurunan performa yang signifikan ketika menghadapi data dengan skala yang besar. Hal ini menjadi peluang model berbasis Transformer, melalui mekanisme *attention* yang mampu menangkap pola jangka panjang yang lebih efisien dan efektif.

### 2.1.2 *Deep-Learning*

Keterbatasan model konvensional dalam menangani pola data yang kompleks dan dinamis menimbulkan ruang bagi *deep-learning* untuk memberikan solusi yang lebih akurat dan adaptif. Perkembangan *Deep-Learning* memberikan kontribusi besar dalam meningkatkan kinerja solusi masalah TSF. Sebagai salah satu cabang *machine-learning*, DL berfokus dalam memahami dan merekayasa fungsi otak manusia melalui perancangan dan pelatihan jaringan saraf berlapis. Jaringan saraf ini bekerja dengan cara menggambarkan hubungan data yang kompleks melalui pengaturan bobot dan koneksi antar lapisan, yang dipelajari menggunakan data dalam skala besar. Berkat kemampuannya dalam mengidentifikasi pola yang penting pada data skala besar, membuat proses prediksi dan klasifikasi menjadi lebih akurat melalui ekstraksi fitur secara otomatis [18].

Berbagai penelitian telah banyak dilakukan terkait penerapan DL terutama pada bidang TSF. Kemampuan model DL dalam menangkap pola non-linear menunjukkan performa yang baik, namun mengalami kesulitan ketika data dalam skala yang sangat besar. Model berbasis Transformer menjadi sebuah pilihan yang unggul, karena memiliki mekanisme *attention* yang mampu menangkap pola jangka panjang dengan efektif dan komputasi yang efisien. Pada beberapa survey ([15], [19]) menyatakan bahwa model berbasis Transformer dan variasinya terutama informer memiliki performa akurasi yang tinggi dan efisiensi komputasi, karena arsitektur dan mekanisme khusus lain yang ada pada informer.

### 2.1.3 *Informer*

Model informer merupakan peningkatan dari model transformer dan memiliki struktur yang serupa dengan transformer, yaitu struktur berlapis yang dibuat dengan menumpuk beberapa blok informer [19]. Informer terdiri dari

dua komponen utama Encoder dan Decoder yang bertukar informasi melalui mekanisme *self-attention* dan modul Encoder-Decoder *attention* [20].

## A Encoder

Encoder bertugas untuk mengekstrak informasi temporal yang penting dari barisan input. Struktur Encoder sebagai berikut.

- *Embedding layer* digunakan untuk mengubah setiap token yang berisi input fitur pada setiap time step, menjadi vektor dense berdimensi tetap. Proses *embedding* dibagi menjadi 3, yaitu *value embedding* untuk proyeksi nilai data, *local timestamp embedding* untuk menangkap informasi posisi, dan *global timestamp embedding* untuk menangkap informasi waktu tambahan.
- *ProbSparse self-attention* bertujuan untuk mengidentifikasi dan memfokuskan hanya pada informasi yang paling penting dan berpengaruh, sehingga dapat mengurangi kompleksitas komputasi. Setiap titik waktu dibandingkan untuk menentukan kontribusinya terhadap prediksi, lalu informer menghitung perbedaan atau jarak yang dihasilkan dan hanya perhatian yang paling tinggi yang dianggap berpengaruh.
- *Feedforward sub-layer* adalah lapisan pemrosesan data setelah data melalui mekanisme *attention*. Pada lapisan ini data diproses untuk mempelajari pola yang lebih kompleks dan mendalam.
- *Self-attention distilling* sebagai mekanisme penyaringan yang digunakan untuk penggabungan atau pengurangan data agar hanya data yang paling penting yang dipertahankan. Sehingga model akan terfokus hanya pada fitur yang dominan dan mengurangi *overfitting*.

## B Decoder

Decoder bertugas untuk menghasilkan prediksi dengan memanfaatkan output dari encoder. Decoder menganut pendekatan *Generative-Style Decoding* untuk menghasilkan semua prediksi secara bersamaan dalam satu langkah [20]. Struktur Decoder sebagai berikut.

- *Masked ProbSparse Self-Attention* adalah mekanisme yang memilih informasi yang paling penting dan menerapkan masking casual agar model hanya memperhatikan informasi masa lalu dan bukan informasi masa depan.

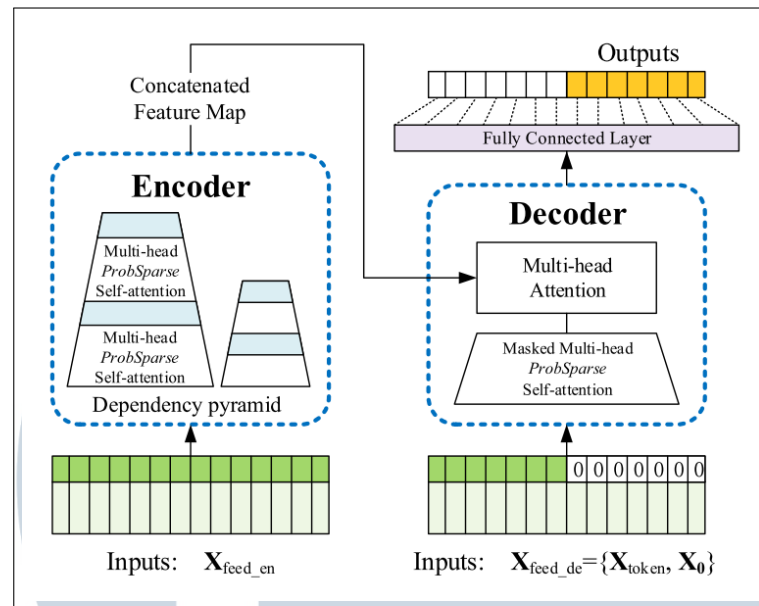
- *Encoder-decoder attention* adalah mekanisme *cross-attention* yang memungkinkan decoder memanfaatkan pola yang telah dipelajari encoder untuk memprediksi nilai masa depan. Decoder menghasilkan prediksi yang bersifat abstrak yang akan diubah menjadi angka melalui Fully Connected Layer.

Berdasarkan penelitian [14], diagram alur proses dari model informer dapat dilihat seperti pada gambar 2.1. Proses diawali dengan mengubah setiap titik waktu menjadi vektor berdimensi melalui *embedding layer*. Kemudian informer menerapkan mekanisme didalam encoder meliputi *ProbSparse self-attention*, *feedforward sub-layer*, dan *self-attention distilling* untuk menangkap pola penting dan mengurangi beban komputasi. Hasil dari encoder berupa *Concatenated Features Map* yang berisi informasi pola temporal utama dari data historis. Selanjutnya hasil encoder dikirimkan ke decoder.

Decoder secara paralel membentuk input yang berupa vektor hasil dari *embedding layer* dan digabungkan dengan *placeholder* yang digunakan sebagai wadah kosong penanda seberapa banyak target nilai yang akan diprediksi. Kemudian input tersebut akan diidentifikasi melalui mekanisme *Masked ProbSparse Self-Attention* yang hasilnya berupa representasi temporal. Selanjutnya hasil tersebut digabungkan dengan output dari encoder melalui *encoder-decoder attention*. Kemudian decoder melakukan prediksi dalam satu langkah sekaligus. Hasil prediksi dalam bentuk vektor abstrak akan diubah kedalam bentuk numerik melalui *Fully Concatenated Layer*.

#### 2.1.4 Penyiapan Data

Tingkat kualitas data sangat mempengaruhi hasil kinerja sebuah model [21]. Data mentah yang telah dikumpulkan seringkali diproses terlebih dahulu untuk memastikan data sesuai dan dapat digunakan dengan maksimal. Serangkaian proses yang dilakukan meliputi pembersihan data, pembagian data, dan normalisasi data [22]. Proses-proses tersebut bertujuan meminimalkan noise, mengatasi inkonsistensi data, dan meningkatkan kualitas data agar model mampu bekerja lebih efektif dan efisien.



Gambar 2.1. Gambar Model Informer

## A Data Cleansing

Data Cleansing adalah proses menyaring, memodifikasi atau menghapus inkonsisten data dari dataset untuk memastikan kualitas data sebelum pemodelan. Proses ini sangat penting karena kualitas data yang bersih akan menghasilkan model yang andal dan analisis yang akurat. Data cleansing secara umum dilakukan dengan cara menghilangkan data duplikat, data kosong, atau data yang tidak relevan serta memodifikasi format atau struktur data.

## B Data Splitting

Data Splitting adalah proses membagi dataset menjadi beberapa bagian seperti data pelatihan (training data), dan data pengujian (testing data) yang digunakan untuk melatih dan menguji model. Pembagian data bertujuan untuk menghindari kebocoran data dan mencegah model menjadi overfitting sehingga menghasilkan model dengan kemampuan generalisasi yang baik [23]. Rasio pembagian data yang umum dan digunakan pada penelitian ini adalah 80% untuk data latih dan 20% untuk data uji. Pembagian ini bertujuan untuk memastikan model informer mampu mempelajari pola jangka panjang dengan lebih maksimal dan data uji digunakan untuk memastikan generalisasi model dengan pengujian menggunakan data yang tidak terlihat selama pelatihan [11].



## C Data Normalization

Data Normalization adalah proses mengubah fitur kedalam skala yang sama tanpa mempengaruhi perbedaan rentang nilai atau informasi penting didalamnya. Proses ini memastikan tidak ada data yang dominan karena memiliki perbedaan nilai yang besar [6]. Salah satu metode umum data normalization adalah Min-Max Scaler, yang mengubah skala data kedalam rentang 0 dan 1. Metode ini sangat cocok digunakan dalam membangun model informer karena skala data yang terbatas [0,1] menyebabkan tidak ada data yang dominan dan proses training menjadi stabil [11], serta tetap mempertahankan pola temporan data. Persamaan MinMaxScaler menggunakan persamaan 2.1.

$$x_{norm} = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (2.1)$$

Keterangan:

$x_{norm}$  = nilai setelah normalisasi

$x$  = nilai sebelum normalisasi

$\max(x)$  = nilai tertinggi pada dataset

$\min(x)$  = nilai terendah pada dataset

### 2.1.5 Metrik Evaluasi

Metrik evaluasi berfungsi untuk mengukur tingkat akurasi suatu model dalam memprediksi data. Penelitian ini menggunakan dua metrik evaluasi yaitu RMSE dan MAPE. RMSE yang lebih menekankan nilai pada kesalahan yang besar dan tetap memberikan hasil skala yang sama dengan nilai asli, MAPE menyajikan kesalahan dalam bentuk persentase yang mudah dipahami.

#### A Root Mean Square Error (RMSE)

RMSE adalah metrik evaluasi yang diperoleh dengan cara akar kuadrat dari *Mean Square Root* (MSE). RMSE tetap mempertahankan karakteristik asal MSE pengkuadratan kesalahan, yaitu memberikan pinalti yang lebih besar pada kesalahan prediksi yang lebih jauh [11]. Namun, dengan menggunakan akar kuadrat membuat skala nilai error kembali sama dengan data asli. Sehingga hasil perhitungan menjadi lebih mudah dipahami. Hal ini membuat RMSE menjadi

metriks yang sering digunakan dalam mengukur performa hasil prediksi. Semakin kecil hasil RMSE, maka semakin baik tingkat akurasi prediksi [24]. Persamaan matematis dari RMSE menggunakan persamaan 2.2.

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (Actual - Predict)^2} \quad (2.2)$$

## B Mean Absolute Percentage Error (MAPE)

MAPE adalah metrik yang menghitung rata-rata persentase selisih antara nilai prediksi dan nilai aktual dibagi nilai aktual [25]. MAPE menghasilkan perhitungan berupa persentase sehingga tidak dalam satuan atau *dimensionless*. Hal ini sangat berguna ketika menggunakan dataset yang berbeda untuk membandingkan performa suatu model. Persamaan matematis dari MAPE menggunakan persamaan 2.3.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{Actual - Predict}{Actual} \right| \times 100\% \quad (2.3)$$

Penilaian tingkat akurasi model dapat dilihat melalui nilai MAPE. Nilai MAPE yang semakin kecil atau mendekati nol maka semakin baik akurasi model. Nilai MAPE yang semakin besar menandakan akurasi model yang buruk. Tabel 2.1 menyajikan katogori tingkat akurasi berdasarkan nilai MAPE [26].

Tabel 2.1. Tabel tingkat akurasi MAPE

MAPE	Tingkat Akurasi
< 5%	Sangat Baik
10% - 25%	Baik
> 25%	Buruk

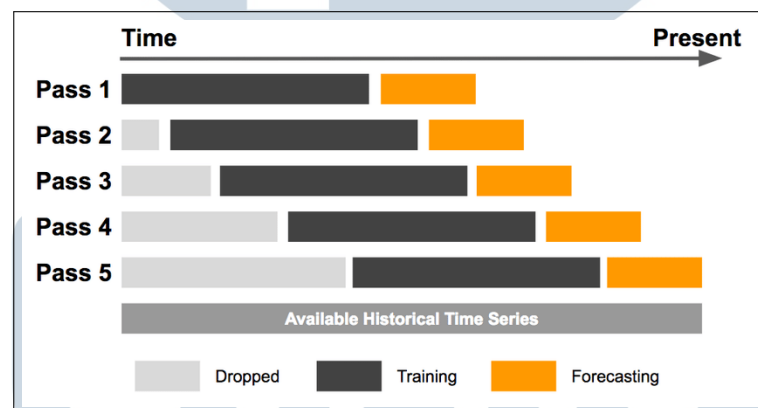
### 2.1.6 Validasi Model

Validasi Model memastikan model melakukan tugas sesuai tujuan, akurat, dan dapat digunakan pada data baru (generalisasi). Terdapat perbedaan antara validasi model secara umum dan validasi model time-series, yaitu validasi model time-series tetap mempertahankan urutan temporal data. Validasi model time-series seperti Walk-Forward Validation merupakan adaptasi dari K-Fold Cross Validation

yang dirancang khusus untuk time-series. Work-Forward Validation tetap menjaga urutan waktu dan memungkinkan simulasi secara realistis.

### A Walk-Forward Validation

Walk-forward validation adalah metode evaluasi khusus *time-series* yang mempertahankan urutan temporal data selama proses pelatihan dan pengujian data. Metode ini membagi data dan melatih data historis awal lalu diuji pada periode berikutnya, selanjutnya jendela pelatihan digeser (*sliding-window*) satu langkah atau satu timestamp dan proses dilakukan berulang. Ukuran jumlah data pelatihan, data pengujian dan pergeseran window dijaga tetap sama sehingga menghasilkan beberapa pasangan data pelatihan dan pengujian [27]. Berdasarkan penelitian [28], ilustrasi skema Walk-Forward Validation digambarkan seperti pada gambar 2.2. Metode ini memastikan data masa depan tidak digunakan untuk melatih data masa lalu, sehingga memungkinkan evaluasi secara realistis dan bebas dari kebocoran data (*data leakage*) [26].



Gambar 2.2. Gambar ilustrasi Walk-Forward Validation

Sebagai contoh jika menggunakan 1000 data, kemudian dipilih 100 data pertama sebagai window awal dengan pembagian data 80:20 sehingga menghasilkan 80 data latih dan 20 data uji. Kemudian proses pelatihan dan pengujian dilakukan dengan menggunakan masing-masing data latih dan data uji. Setelah proses pelatihan dan pengujian selesai, selanjutnya dihitung nilai evaluasi RMSE dan MAPE dan disimpan. Selanjutnya window bergeser sebanyak 100 data dan dilakukan kembali proses pembagian data, pelatihan, dan pengujian. Proses ini berlangsung hingga window mencapai data historis terakhir.



## 2.2 Penelitian Terdahulu

Beberapa tahun belakangan, Informer menjadi model yang banyak diterapkan dan dikembangkan khususnya dalam memprediksi data *time-series*. Pengaplikasian Informer sebagai model prediksi dikarenakan kemampuannya menangkap pola hubungan jangka panjang dan komputasi yang lebih efisien hingga menghasilkan prediksi dengan lebih akurat dibanding model konvensional. Beberapa penelitian terdahulu yang dirangkum pada tabel 2.2 menjadi dasar konseptual dalam penelitian ini dan sebagai bukti empiris bahwa Informer memiliki potensi signifikan dalam prediksi data *time-series* yang lebih akurat.

Tabel 2.2. Tabel penelitian terdahulu

No	Judul	Latar Balakang dan Hasil	Celah Penelitian
1	The Time Series Informer Model for Stock Market Prediction [11]	Membandingkan model Informer dan LSTM dalam memprediksi harga saham Google. Informer memiliki prediksi yang lebih akurat yang menghasilkan RMSE 1.75 lebih kecil dibandingkan LSTM 2.35.	Meskipun Informer menunjukkan performas yang lebih unggul, namun Informer memiliki waktu training yang lebih lama yaitu 60s dibandingkan waktu training LSTM 45s.

2	Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting [14]	Penelitian ini berfokus dalam meningkatkan model Transfomer dalam hal komputasi, memori dan efisiensi arsitektur. Pengujian menggunakan univariat dan multivariat dari 4 dataset berbeda (2 dataset Temperatur Transformator Listrik, Bebam Komsumsi Listrik, dan Cuaca). Peningkatan pada model memberikan hasil lebih akurat, yang mana ditunjukkan dari nilai MSE terendah dari model pembandingan lainnya.	Studi ini difokuskan untuk meningkatkan efisiensi dan mengurangi komputasi, namun tidak membahas reaksi model ketika menangani data dengan noise. Model diuji menggunakan data yang relatif halus dan beraturan, dan belum dilakukan pengujian terhadap data dengan tingkat volatilitas tinggi seperti data harga emas.
3	Time Series Forecasting of Motor Bearing Vibration Based on Informer [29]	Penelitian ini memprediksi getaran pada bearing motor dengan menggunakan model Informer yang telah dioptimasi menggunakan <i>random search</i> dan membandingkan dengan model lain seperti CNN, DeepRNNs, LSTM, dan Transformer. Hasil menunjukan Informer mampu memprediksi pola tren dan nilai yang ekstrim dengan lebih baik dari model pembandingan lain.	Studi ini menunjukkan keunggulan Informer dibandingkan model lain, namun belum menunjukkan perbedaan yang signifikan. Hal ini dikarenakan Informer masih terpengaruh oleh data dengan fluktuasi cepat dan tajam. Pengujian lebih lanjut perlu dilakukan menggunakan data dengan fluktuasi kuat.

4	A Multi-Source Time Series Forecasting Framework Based on Informer for Climate Finance and Industrial Economy [30]	Perkembangan <i>climate finance</i> dan ekonomi industri sangat dipengaruhi oleh banyak variabel. Penggunaan model Informer diharapkan mampu menangkap pola hubungan jangka panjang dan mengatasi data multivariat berskala besar. Model Informer mampu menunjukkan hasil dan performa yang lebih unggul dalam tingkat akurasi prediksi dibandingkan model lain (Transformer, LSTM, dan LightGBM).	Studi ini memanfaatkan data multi sumber yang cenderung berubah secara bertahap. Perbedaan karakteristik data terhadap harga emas yang berfluktuasi tinggi menjadi celah penelitian, karena belum dilakukan evaluasi mendalam mengenai efektivitas Informer dalam menangani data volatilitas ekstrem dan cepat seperti harga emas.
---	--	--	--



5	Applying Informer for Option Pricing: A Transformer-Based Approach [20]	Penentuan harga opsi sangat penting untuk trading dan manajemen resiko pada pasar keuangan, namun menjadi tantangan karena volatilitas pasar dan keterbatasan model konvensional. Penelitian ini menggunakan Informer untuk memprediksi harga opsi. Pada penelitian ini Informer mendapat nilai MAE 2.71 dan RMSE 3.68 menjadikannya terkecil dari semua model yang diuji (Black-Scholes, Heston, LSTM). Hal ini menunjukkan Informer menjadi model yang lebih unggul dalam prediksi harga opsi.	Studi ini berfokus pada pemetaan hubungan dan prediksi nilai pada jangka waktu pendek. Belum dilakukan evaluasi mengenai kinerja Informer terhadap horizon prediksi yang panjang dan berkelanjutan.
---	---	--	---

Penelitian terdahulu yang tertera pada tabel 2.2 menggunakan model berbasis *deep-learning* dalam memprediksi data *time-series*. Penelitian-penelitian tersebut menggunakan metrik evaluasi MAE, MSE, RMSE, MAPE untuk menguji tingkat akurasi model. Pengaplikasian Informer sebagian besar dilakukan pada domain yang berbeda, mulai dari harga saham, harga opsi, cuaca, penggunaan listrik, sampai getaran bearing motor, ini menunjukkan bahwa Informer sangat fleksibel digunakan pada data sekuensial jangka panjang.

Kemampuan Informer dalam mengatasi pola sekuensial yang panjang secara efisien dan mempelajari pola jangka panjang melalui mekanisme *self-attention* yang unik dan optimasi arsitektur menjadi alasan utama pemilihan model ini. Terbukti pada penelitian [14], melalui pengujian terhadap 4 dataset dengan domain dan rentang waktu yang berbeda, menghasilkan nilai MSE dan MAE terendah dibandingkan dengan model lain.

Informer telah banyak dikembangkan dan diterapkan dalam bidang keuangan. Seperti pada penelitian [11], yang berfokus pada harga pasar saham

(saham Google) dan penelitian [20] yang berfokus pada derivatif (harga opsi), yang mengungguli semua model pembandingnya dan memberikan hasil yang akurat. Penelitian lain [30], memanfaatkan data multi sumber untuk menemukan dampak variabel iklim terhadap operasi ekonomi industri dan pola prediksi. Penelitian-penelitian tersebut memiliki keterbatasan seperti waktu komputasi yang lebih lama, penggunaan horizon prediksi yang relatif pendek, dan perbedaan karakteristik data. Karakteristik data harga emas yang memiliki volatilitas ekstrim dan horizon prediksi yang panjang dan berkelanjutan menjadikan penerapan Informer pada prediksi harga emas sebagai celah penelitian yang perlu dieksplorasi lebih jauh.

