

**PENGEMBANGAN SISTEM PREDIKSI BI-RATE
BERBASIS EXPLAINABLE ARTIFICIAL
INTELLIGENCE (XAI) DAN LARGE LANGUAGE
MODEL (LLM) MENGGUNAKAN INDIKATOR
MAKROEKONOMI**

Proposal Tugas Akhir

Oleh

**Erwan Poltak Halomoan
18222028**



**PROGRAM STUDI SISTEM DAN TEKNOLOGI INFORMASI
SEKOLAH TEKNIK ELEKTRO DAN INFORMATIKA
INSTITUT TEKNOLOGI BANDUNG
Desember 2025**

LEMBAR PENGESAHAN

PENGEMBANGAN SISTEM PREDIKSI BI-RATE BERBASIS EXPLAINABLE ARTIFICIAL INTELLIGENCE (XAI) DAN LARGE LANGUAGE MODEL (LLM) MENGGUNAKAN INDIKATOR MAKROEKONOMI

Proposal Tugas Akhir

Oleh

**Erwan Poltak Halomoan
18222028**

Program Studi Sistem dan Teknologi Informasi
Sekolah Teknik Elektro dan Informatika
Institut Teknologi Bandung

Proposal Tugas Akhir ini telah disetujui dan disahkan
di Bandung, pada tanggal 5 Desember 2025

Pembimbing

Ir. Windy Gambetta M.B.A.
NIP. 196902121997021001

DAFTAR ISI

| | |
|---|-----------|
| DAFTAR GAMBAR | v |
| DAFTAR TABEL | vi |
| I PENDAHULUAN | 1 |
| I.1 Latar Belakang | 1 |
| I.2 Rumusan Masalah | 3 |
| I.3 Tujuan | 4 |
| I.4 Batasan Masalah | 4 |
| I.5 Metodologi | 4 |
| II STUDI LITERATUR | 7 |
| II.1 Kebijakan Moneter dan Variabel Makroekonomi | 7 |
| II.1.1 Suku Bunga Acuan Bank Indonesia (BI-Rate) | 7 |
| II.1.2 Teori Penetapan Suku Bunga (Taylor Rule) | 8 |
| II.1.3 Indikator Makroekonomi Penentu Suku Bunga | 8 |
| II.1.3.1 Indikator Domestik | 8 |
| II.1.3.2 Indikator Global | 10 |
| II.1.4 Hubungan Antarvariabel Ekonomi | 11 |
| II.2 Model Prediksi Ekonomi Berbasis <i>Time Series</i> | 11 |
| II.2.1 Model Tradisional (Ekonometrika) | 12 |
| II.2.1.1 Autoregressive Integrated Moving Average (ARIMA) | 12 |
| II.2.1.2 Vector Autoregression (VAR) | 13 |
| II.2.1.3 Vector Error Correction Model (VECM) | 13 |
| II.2.2 Model <i>Machine Learning</i> | 13 |
| II.2.2.1 Random Forest Regressor | 14 |
| II.2.2.2 Extreme Gradient Boosting (XGB) | 14 |
| II.2.3 Model <i>Deep Learning</i> | 15 |
| II.2.3.1 Long Short-Term Memory (LSTM) | 15 |
| II.2.3.2 Bidirectional Long Short-Term Memory (Bi-LSTM) | 16 |
| II.2.4 Metode Evaluasi Model | 17 |
| II.2.4.1 Root Mean Squared Error (RMSE) | 17 |
| II.2.4.2 Mean Absolute Percentage Error (MAPE) | 17 |
| II.2.4.3 R-squared (R^2) | 17 |
| II.3 Explainable Artificial Intelligence (XAI) | 18 |

| | | |
|--------------------------------|---|----|
| II.3.1 | Taksonomi Metode XAI | 19 |
| II.3.1.1 | SHapley Additive exPlanations (SHAP) | 20 |
| II.3.1.2 | Local Interpretable Model-Agnostic Explanations (LIME) | 21 |
| II.3.1.3 | Permutation Feature Importance (PFI) | 22 |
| II.3.2 | Penerapan XAI dalam Prediksi Ekonomi | 23 |
| II.4 | Large Language Model (LLM) | 23 |
| II.4.1 | OpenAI GPT-4o | 24 |
| II.4.2 | Meta Llama | 24 |
| II.5 | Penelitian Terkait | 25 |
| II.5.1 | Machine Learning Applications in Central Banking (Chakraborty dan Joseph 2017) | 25 |
| II.5.2 | Explainable Machine Learning in Credit Risk Management (Bussmann dkk. 2021) | 26 |
| II.5.3 | Temporal Data Meets LLM – Explainable Financial Time Series Forecasting (Yu, Chen, dan Lu 2023) | 26 |
| II.5.4 | Explainable-AI powered stock price prediction using time series transformers: A Case Study on BIST100 (Calik dkk. 2025) | 26 |
| III ANALISIS MASALAH | 28 | |
| III.1 | Analisis Kondisi Saat Ini | 28 |
| III.1.1 | Business Understanding | 28 |
| III.1.2 | Gambaran Sistem Saat Ini | 29 |
| III.2 | Analisis Kebutuhan | 30 |
| III.2.1 | Identifikasi Masalah Pengguna | 30 |
| III.2.2 | Kebutuhan Fungsional | 30 |
| III.2.3 | Kebutuhan Non-Fungsional | 31 |
| III.3 | Analisis Pemilihan Solusi | 32 |
| III.3.1 | Seleksi Model Prediksi | 32 |
| III.3.2 | Seleksi Metode XAI | 34 |
| III.3.3 | Seleksi Model LLM | 35 |
| III.3.4 | Penentuan Solusi | 35 |
| III.4 | Data Understanding | 36 |
| III.4.1 | Data Overview | 36 |
| IV DESAIN KONSEP SOLUSI | 38 | |
| IV.1 | Desain Solusi | 38 |
| V RENCANA SELANJUTNYA | 41 | |

| | | |
|-------|--|----|
| V.1 | Rencana Implementasi | 41 |
| V.1.1 | Perkakas yang Dibutuhkan | 41 |
| V.1.2 | Lini masa Pengerjaan Tugas Akhir | 42 |
| V.2 | Rencana Anggaran Biaya | 43 |
| V.3 | Desain Pengujian dan Evaluasi | 43 |
| V.3.1 | Verifikasi Kinerja Model | 43 |
| V.3.2 | Validasi Ekonomi | 44 |
| V.4 | Analisis Risiko dan Mitigasi | 44 |

DAFTAR GAMBAR

| | | |
|-------|--|----|
| I.1 | Metodologi CRISP-DM (Martínez-Plumed dkk. 2021) | 5 |
| II.1 | Ilustrasi <i>Black Box</i> dan <i>Interpretable</i> Model (Hui dkk. January 2022) | 18 |
| II.2 | Ilustrasi Taksonomi Metode XAI (Adadi dan Berrada 2018) | 19 |
| III.1 | Proses Penetapan BI-Rate Saat Ini | 29 |
| III.2 | Gambaran Sistem Prediksi BI-Rate Terdahulu (Putri, Adikara, dan Adinugroho May 2019) | 30 |
| IV.1 | Desain Solusi Pengerjaan Tugas Akhir | 39 |

DAFTAR TABEL

| | | |
|-------|--|----|
| III.1 | Kebutuhan Fungsional | 31 |
| III.2 | Kebutuhan Non-Fungsional | 31 |
| III.3 | Matriks Seleksi Model Prediksi | 32 |
| III.4 | Matriks Seleksi Metode XAI | 34 |
| III.5 | Matriks Seleksi Model Generatif (LLM) | 35 |
| III.6 | Deskripsi Kolom Dataset | 36 |
| IV.1 | Perbandingan Sistem Saat Ini dan Ajuan Sistem Solusi | 40 |
| V.1 | Perkakas yang Dibutuhkan | 41 |
| V.2 | Liniwasa Pengerjaan Tugas Akhir | 42 |
| V.3 | Rencana Anggaran Biaya | 43 |
| V.4 | Skenario Verifikasi Kinerja Model | 44 |
| V.5 | Skenario Validasi Interpretasi Ekonomi | 44 |
| V.6 | Analisis Risiko dan Mitigasi | 45 |

BAB I

PENDAHULUAN

Bab ini membahas secara umum permasalahan yang menjadi dasar penyusunan tugas akhir. Bab ini bertujuan untuk memberikan gambaran awal dan pemahaman mendasar mengenai topik yang diangkat. Pada subbab latar belakang dijelaskan faktor-faktor yang menimbulkan permasalahan. Selanjutnya, subbab rumusan masalah, tujuan, serta batasan masalah digunakan untuk memperjelas dan membatasi ruang lingkup permasalahan dengan konkret. Terakhir, subbab metodologi membahas tentang proses perumusan solusi.

I.1 Latar Belakang

Suku bunga acuan Bank Indonesia, yang saat ini dikenal dengan nama BI-Rate, merupakan instrumen kebijakan moneter yang digunakan sebagai sinyal kebijakan untuk mengendalikan ekspektasi inflasi dan menjaga stabilitas nilai rupiah. Dalam perkembangannya, Bank Indonesia melakukan penguatan kerangka operasi moneter pada 19 Agustus 2016 dengan mengimplementasikan BI-7 Day Reverse Repo Rate (BI7DRR). Instrumen ini dipilih karena bersifat transaksional di pasar uang, sehingga memiliki hubungan yang lebih kuat dalam memengaruhi perbankan dan sektor riil. Selanjutnya, mulai 21 Desember 2023, Bank Indonesia kembali menggunakan nama BI-Rate menggantikan BI7DRR. Perubahan nama ini bertujuan untuk memperkuat komunikasi kebijakan moneter, tanpa mengubah makna, tujuan, maupun operasionalisasinya yang tetap mengacu pada transaksi *reverse repo* Bank Indonesia tenor 7 hari (Bank Indonesia 2023).

Penetapan BI-Rate dilakukan melalui Rapat Dewan Gubernur (RDG) Bank Indonesia dengan mempertimbangkan berbagai indikator makroekonomi yang kompleks, baik indikator domestik maupun global. Secara teoritis, keputusan ini sering berlandaskan pada konsep *Taylor Rule*, yang menyarankan bank sentral harus menye-

suaikan suku bunga nominal sebagai respon deviasi inflasi aktual dari targetnya dan kesenjangan *output* ekonomi (Taylor 1993). Namun, seiring berkembangnya dinamika ekonomi global, variabel yang memengaruhi keputusan kebijakan moneter semakin bertambah dan rumit, jauh melampaui Taylor Rule awal yang cukup sederhana. Mekanisme transmisi kebijakan moneter sangatlah kompleks, sehingga melandaskan keputusan hanya pada fungsi sederhana dari inflasi dan *output gap* menjadi tidak relevan dalam praktik bank sentral modern (L. E. O. Svensson 2003).

Secara historis, pemodelan ekonomi di lingkungan bank sentral sangat bergantung pada pendekatan ekonometrika tradisional untuk memproyeksikan variabel makroekonomi. Model statistik konvensional seperti *autoregressive* (AR) dan *vector autoregressive* (VAR) telah lama menjadi standar dalam analisis deret waktu untuk menangkap hubungan antar variabel (Chakraborty dan Joseph 2017). Namun, metode konvensional ini memiliki keterbatasan fundamental karena umumnya berdasumsi pada hubungan linear, sehingga sering kali kesulitan menangkap pola non-linier yang kompleks yang melekat pada data keuangan modern (Sezer, Gudelek, dan Ozbayoglu 2020). Ketidakmampuan model tradisional dalam menangkap dependensi non-linier ini dapat menyebabkan stagnasi atau penurunan akurasi prediksi, terutama ketika dihadapkan pada dataset dengan volatilitas tinggi.

Keterbatasan model tradisional ini mendorong pergeseran ke arah metode yang lebih modern. Pendekatan *machine learning* (ML) dan *deep learning* (DL) terbukti lebih andal dalam menangkap pola non-linier (Chen dan Guestrin 2016). Algoritma ML mampu menangkap struktur data keuangan yang kompleks dengan lebih baik. Model *deep learning* seperti *Long Short-Term Memory* (LSTM) mampu mengurangi tingkat kesalahan prediksi secara signifikan dibandingkan model ARIMA (Sezer, Gudelek, dan Ozbayoglu 2020). Kemampuan LSTM untuk mempelajari ketergantungan temporal jangka panjang menjadikannya relevan untuk menangani data ekonomi yang fluktuatif (Hochreiter dan Schmidhuber 1997).

Walau model ML dan DL menunjukkan akurasi prediksi yang tinggi, tantangan utama yang dihadapi adalah interpretabilitas. Model ML dan DL beroperasi sebagai *black box* yang menyulitkan pemahaman terhadap faktor-faktor yang mendasari hasil prediksi (Rudin 2019). Dalam konteks kebijakan ekonomi, akurasi semata tidak cukup, pembuat kebijakan membutuhkan pemahaman kausal mengenai faktor yang memengaruhi prediksi demi akuntabilitas publik (Bussmann dkk. 2021). Bank of England menyatakan bahwa penggunaan ML di bank sentral wajib disertai transparansi agar keputusan yang diambil memiliki legitimasi (Chakraborty dan Joseph

2017). Sesuai pernyataan tersebut, sektor keuangan memerlukan model yang dapat dijelaskan (*explainable*) untuk memenuhi standar regulasi dan manajemen risiko (Bussmann dkk. 2021).

Untuk menjawab kebutuhan tersebut, pendekatan *explainable artificial intelligence* (XAI) seperti *SHapley Additive exPlanations* (SHAP) berkembang pesat sebagai solusi untuk melihat kontribusi variabel dalam model kompleks (Adadi dan Berrada 2018). Penerapan ML di bank sentral global telah meluas, mulai dari analisis makroekonomi hingga stabilitas finansial (Chakraborty dan Joseph 2017). Studi terbaru menunjukkan model berbasis *Gradient Boosting* yang diinterpretasikan dengan SHAP dapat menangkap faktor risiko finansial secara akurat dan tetap transparan (Moscatelli dkk. 2020).

Namun, di Indonesia, penelitian prediksi BI-Rate yang mengintegrasikan ML/DL dengan XAI masih sangat terbatas. Di sisi lain, meskipun solusi XAI tersedia, hasil outputnya masih berupa grafik teknis atau matriks angka. Oleh karena itu, penelitian ini hadir untuk mengisi gap tersebut dengan mengintegrasikan *Large Language Models* (LLM) sebagai bantuan kognitif. Sebagaimana disarankan dalam riset prediksi finansial terkini, LLM tidak hanya berfungsi sebagai penerjemah teks, tetapi mampu menarasikan logika kausalitas dari interpretasi teknis XAI menjadi wawasan kebijakan yang intuitif dan komunikatif, sehingga hasil prediksi dapat langsung dipahami oleh pemangku kepentingan non-teknis dalam perumusan kebijakan moneter (Yu, Chen, dan Lu 2023).

I.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dibahas, terdapat kebutuhan akan model prediksi BI-Rate yang akurat sekaligus transparan, untuk mendukung terciptanya kebijakan ekonomi yang tepat. Hal tersebut melahirkan beberapa permasalahan sebagai berikut:

1. Bagaimana membangun model prediksi BI-Rate yang mampu menangkap hubungan linear maupun non-linier antar indikator makroekonomi menggunakan pendekatan ML dan DL?
2. Bagaimana perbandingan performa akurasi antara model ekonometrika tradisional, ML, dan DL dalam memprediksi BI-Rate?
3. Bagaimana penerapan teknik XAI untuk memperlihatkan kontribusi indikator makroekonomi terhadap hasil prediksi model secara transparan?
4. Bagaimana memanfaatkan LLM untuk menerjemahkan hasil interpretasi tek-

nis dari XAI menjadi penjelasan naratif yang dapat dipahami dalam konteks kebijakan ekonomi?

I.3 Tujuan

Berdasarkan rumusan masalah yang sudah dijelaskan, didapatkan tujuan dari tugas akhir sebagai berikut. Tujuan yang dibuat menjawab pertanyaan yang muncul dalam rumusan masalah.

1. Mengembangkan model prediksi BI-Rate berbasis ML dan DL dengan memanfaatkan data indikator makroekonomi domestik dan global.
2. Melakukan evaluasi komparatif terhadap kinerja model yang dikembangkan dibandingkan dengan model tradisional berdasarkan metrik *error* prediksi.
3. Mengimplementasikan teknik XAI untuk memberikan interpretasi global dan lokal dari hasil prediksi yang dihasilkan oleh model prediktif.
4. Menghasilkan mekanisme pelaporan otomatis berbasis LLM yang mampu menarasikan alasan di balik prediksi BI-Rate berdasarkan analisis fitur XAI.

I.4 Batasan Masalah

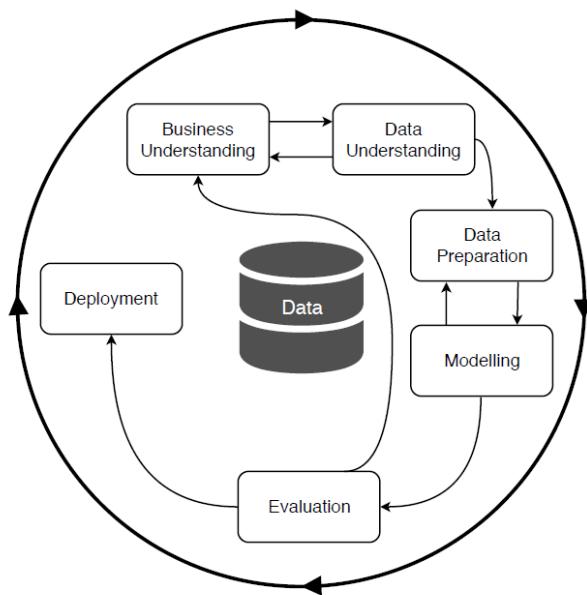
Dalam penggeraan tugas akhir ini dibutuhkan batasan masalah agar lingkup dari tugas akhir yang dikerjakan tidak terlalu luas. Berdasarkan tujuan yang sudah dijelaskan, dapat dibuat batasan masalah sebagai berikut.

1. Dataset yang digunakan bersumber dari data resmi Bank Indonesia, Badan Pusat Statistik (BPS), Bank for International Settlements (BIS) Data Portal, Federal Reserve Economic Data (FRED), dan sumber kredibel lainnya. Data mencakup periode waktu bulanan dari Juli 2005 hingga September 2024.
2. Penggunaan LLM dalam penelitian ini difokuskan hanya sebagai pembuat narasi untuk menjelaskan hasil interpretasi XAI. LLM tidak digunakan untuk melakukan kalkulasi prediksi angka suku bunga secara langsung.

I.5 Metodologi

Dalam penggeraan tugas akhir ini, dibutuhkan alur penggeraan yang terstruktur supaya tujuan yang telah dirumuskan dapat tercapai. Metodologi tugas akhir yang digunakan adalah CRISP-DM (*Cross-Industry Standard Process for Data Mining*) yang merupakan kerangka kerja standar yang bersifat independen terhadap industri dan telah menjadi acuan utama dalam proyek *data mining* (Schröer, Kruse, dan Gómez 2021). Pemilihan CRISP-DM didasarkan pada kemampuannya dalam mengelola

siklus proyek berbasis data dan telah terbukti banyak digunakan di berbagai sektor (Martínez-Plumed dkk. 2021). Penulis tidak melakukan tahap *deployment* pada tugas akhir ini karena fokusnya adalah pada pengembangan dan evaluasi model prediksi, bukan penerapan sistem secara operasional.



Gambar I.1 Metodologi CRISP-DM (Martínez-Plumed dkk. 2021)

Tahapan dalam CRISP-DM dijelaskan sebagai berikut:

1. *Business Understanding*

Pada tahap ini dilakukan identifikasi kebutuhan dan tujuan tugas akhir dalam konteks kebijakan moneter Indonesia. Tahap ini juga meliputi eksplorasi literatur untuk membantu perumusan masalah, objektif, dan parameter keberhasilan tugas akhir ini.

2. *Data Understanding*

Pada tahap *data understanding* dilakukan pengumpulan dan eksplorasi awal data makroekonomi dari sumber terpercaya, serta melakukan analisis struktur data, kelengkapan, pola dasar, identifikasi potensi masalah data, dan mencari *insights* yang mungkin belum didapatkan sebelumnya.

3. *Data Preparation*

Pada tahap *data preparation*, data dipersiapkan agar layak digunakan untuk proses pemodelan. Tahap ini meliputi pemrosesan data, transformasi format, dan konstruksi fitur yang diperlukan untuk meningkatkan kualitas *input* model. *Dataset* akhir merupakan data yang sudah layak dan siap untuk *training* dan pengujian model.

4. *Modelling*

Pada tahap *modelling*, dilakukan pengembangan model prediksi, konfigurasi parameter, pelatihan model, dan validasi model prediksi yang sesuai terhadap data terkait.

5. *Evaluation*

Pada tahap *evaluation*, hasil model dan aspek interpretabilitas dievaluasi berdasarkan metrik yang telah ditentukan untuk menilai apakah model yang dikembangkan telah memenuhi tujuan awal tugas akhir dan layak dinyatakan selesai. Jika hasil belum memuaskan, dapat dilakukan iterasi kembali ke tahap sebelumnya untuk perbaikan, baik data maupun model.

BAB II

STUDI LITERATUR

Pada bab ini dibahas berbagai literatur dan Tugas Akhir terdahulu yang dapat menjadi acuan dalam pengerjaan tugas akhir. Bab ini bertujuan untuk memberikan pemahaman terkait konsep dan metode yang digunakan pada tugas akhir ini. Bab ini dibagi menjadi lima subbab utama, yaitu pembahasan mengenai indikator makroekonomi penentu suku bunga acuan pada subbab II.1, model prediksi ekonomi berbasis *machine learning* dan *time series* pada subbab II.2, konsep *Explainable AI* (XAI) pada subbab II.3, pemanfaatan *Large Language Model* (LLM) pada subbab II.4, serta tinjauan penelitian terkait pada subbab II.5.

II.1 Kebijakan Moneter dan Variabel Makroekonomi

II.1.1 Suku Bunga Acuan Bank Indonesia (BI-Rate)

BI-Rate adalah suku bunga kebijakan yang mencerminkan *stance* kebijakan moneter Bank Indonesia. Instrumen BI-Rate memiliki hubungan yang kuat dengan suku bunga pasar uang, bersifat transaksional atau diperdagangkan di pasar, dan mendorong pendalaman pasar keuangan, khususnya penggunaan instrumen *repo*. Mulai 21 Desember 2023, Bank Indonesia kembali menggunakan nama BI-Rate sebagai suku bunga kebijakan menggantikan BI 7-Day Reverse Repo Rate (BI7DRR) untuk memperkuat komunikasi kebijakan moneter, tanpa mengubah makna dan tujuannya sebagai *stance* kebijakan moneter Bank Indonesia, dan operasionalisasinya tetap mengacu pada transaksi *reverse repo* Bank Indonesia tenor tujuh hari (Bank Indonesia 2023).

Pemilihan tenor tujuh hari didasarkan pada pertimbangan bahwa instrumen dengan tenor ini lebih aktif ditransaksikan di pasar uang dibandingkan tenor yang lebih panjang, sehingga lebih cepat memengaruhi kondisi pasar. Mekanisme penetapan BI-Rate dilakukan melalui Rapat Dewan Gubernur (RDG) Bank Indonesia yang dise-

lenggarakan setiap bulan dengan mempertimbangkan asesmen menyeluruh terhadap kondisi makroekonomi. Keputusan perubahan BI-Rate memengaruhi perekonomian riil melalui mekanisme transmisi kebijakan moneter yang bekerja melalui lima jalur utama, yaitu jalur suku bunga, jalur kredit, jalur nilai tukar, jalur harga aset, dan jalur ekspektasi (Warjiyo dan Juho 2019).

II.1.2 Teori Penetapan Suku Bunga (Taylor Rule)

Landasan teoritis yang paling banyak digunakan untuk menjelaskan penetapan suku bunga acuan oleh bank sentral adalah Taylor Rule, yang diperkenalkan oleh John B. Taylor pada tahun 1993 (Taylor 1993).

Secara matematis, Taylor Rule diformulasikan sebagai berikut dalam Persamaan II.1:

$$r = p + 0.5y + 0.5(p - 2) + 2 \quad (\text{II.1})$$

Pada Persamaan II.1, r adalah tingkat suku bunga *federal funds* (dalam persen), p adalah tingkat inflasi tahunan yang diukur menggunakan deflator PDB selama empat kuartal sebelumnya, sedangkan y menyatakan persentase deviasi PDB riil dari tren atau tingkat *output* potensialnya (*output gap*). Angka 2 yang muncul dalam bagian $(p - 2)$ adalah asumsi target inflasi jangka panjang sebesar 2 persen, sementara konstanta +2 pada bagian akhir persamaan adalah asumsi Taylor mengenai tingkat suku bunga riil ekuilibrium jangka panjang yang juga sebesar 2 persen. Dengan demikian, koefisien 0,5 pada y dan $(p - 2)$ menunjukkan bahwa bank sentral merespon deviasi inflasi dari target dan deviasi output dari potensinya dengan bobot yang seimbang dalam penentuan suku bunga kebijakan. Dinamika ekonomi modern menuntut bank sentral untuk mempertimbangkan variabel lain di luar inflasi dan output, seperti stabilitas nilai tukar dan kondisi likuiditas global (L. E. O. Svensson 2003).

II.1.3 Indikator Makroekonomi Penentu Suku Bunga

Penetapan suku bunga acuan didasarkan pada analisis komprehensif terhadap berbagai indikator makroekonomi, baik domestik ataupun global.

II.1.3.1 Indikator Domestik

Beberapa indikator makroekonomi domestik yang dibahas antara lain:

1. Inflasi

Bank Indonesia menerapkan kerangka Inflation Targeting Framework (ITF), yaitu target inflasi menjadi dasar utama kebijakan moneter. Tingkat inflasi yang tinggi umumnya akan mendorong bank sentral untuk menaikkan suku bunga untuk meredam permintaan agregat. Inflasi yang digunakan adalah inflasi *year-on-year* yang mencerminkan perubahan harga dalam satu tahun terakhir (Bank Indonesia 2023).

2. Produk Domestik Bruto (PDB)

Produk Domestik Bruto (PDB) mencerminkan kondisi aktivitas ekonomi suatu negara. Pertumbuhan PDB yang terlalu cepat dapat memicu tekanan inflasi, sementara pertumbuhan yang terlalu lambat mengindikasikan *slack* dalam perekonomian. *Output gap*, yaitu selisih antara PDB aktual dan PDB potensial, menjadi indikator penting dalam Taylor Rule untuk menentukan *stance* kebijakan moneter yang tepat (Taylor 1993).

3. Nilai Tukar (USD/IDR)

Nilai tukar mata uang, khususnya Rupiah terhadap Dolar Amerika Serikat (USD/IDR), menjadi indikator yang penting dalam ekonomi terbuka seperti Indonesia. Depresiasi nilai tukar dapat meningkatkan inflasi melalui jalur *imported inflation* (Mishkin 2019).

4. Jumlah Uang Beredar (M2)

Jumlah peredaran uang dalam perekonomian, yang diukur melalui agregat moneter M2, mempengaruhi tekanan inflasi melalui *demand-pull inflation*. Pertumbuhan M2 yang terlalu cepat dapat mengindikasikan kondisi moneter yang terlalu longgar dan memerlukan penyesuaian suku bunga untuk mengendalikan likuiditas (Mishkin 2019).

5. Kredit Perbankan

Kredit perbankan mencerminkan kondisi intermediasi keuangan dan aktivitas ekonomi riil. Pertumbuhan kredit yang tinggi dapat mengindikasikan ekspansi ekonomi yang kuat, namun juga berpotensi menciptakan *bubble* dan risiko stabilitas finansial. Bank Indonesia memantau pertumbuhan kredit sebagai salah satu indikator dalam perumusan kebijakan moneter dan makroprudential (Warjiyo dan Juhro 2019).

6. Indeks Harga Saham Gabungan (IHSG)

IHSG mencerminkan sentimen pasar terhadap kondisi ekonomi dan ekspektasi terhadap prospek perekonomian. Pergerakan IHSG dapat mempengaruhi *wealth effect* dan konsumsi masyarakat melalui jalur harga aset dalam mekanisme transmisi kebijakan moneter (Mishkin 2019).

7. Cadangan Devisa

Cadangan devisa merupakan indikator ketahanan eksternal dan kapasitas bank sentral dalam menstabilkan nilai tukar. Cadangan devisa yang memadai dapat memberikan kredibilitas pada kebijakan moneter dan meningkatkan kepercayaan investor terhadap stabilitas ekonomi (Obstfeld, Shambaugh, dan Taylor 2010).

II.1.3.2 Indikator Global

Beberapa indikator makroekonomi global yang digunakan dalam pengajaran tugas akhir ini, antara lain:

1. Federal Funds Rate (FFR)

Federal Funds Rate adalah suku bunga acuan yang ditetapkan oleh Federal Reserve (Fed) Amerika Serikat. FFR memiliki pengaruh signifikan terhadap kebijakan moneter negara *emerging markets*, termasuk Indonesia, melalui aliran modal dan ekspektasi pasar global (Rey 2015). Kenaikan FFR cenderung mendorong *capital outflow* dari *emerging markets* dan memberikan tekanan depresiasi pada mata uang domestik, yang selanjutnya mempengaruhi keputusan suku bunga Bank Indonesia (Obstfeld, Ostry, dan Qureshi 2019).

2. Harga Minyak

Harga komoditas internasional, khususnya minyak mentah, memengaruhi inflasi domestik melalui jalur biaya produksi dan ekspektasi inflasi. Kenaikan harga minyak meningkatkan biaya energi dan transportasi, yang kemudian ditransmisikan ke harga barang dan jasa secara luas (Hamilton 2009).

3. Harga Emas

Emas sering dianggap sebagai aset *safe haven* dan lindung nilai terhadap inflasi. Ditemukan bahwa kenaikan harga emas sering kali bertepatan dengan ketidakpastian ekonomi global atau ekspektasi inflasi tinggi (Baur dan McDermott 2010).

4. Volatilitas Pasar Global (VIX)

VIX atau *fear index* mengukur ekspektasi volatilitas pasar saham AS dan sering digunakan sebagai barometer sentimen risiko global. Peningkatan VIX mengindikasikan *risk-off sentiment* yang mendorong investor untuk melepas aset berisiko di *emerging markets* dan mencari *safe haven* di obligasi pemerintah negara maju (Bekaert, Engstrom, dan Xu 2019).

II.1.4 Hubungan Antarvariabel Ekonomi

Hubungan antara indikator makroekonomi dan suku bunga acuan bersifat dinamis, kompleks, dan tidak selalu mengikuti pola linear sederhana. Pertama, terdapat fenomena efek ambang batas atau *threshold effects*, ketika pengaruh suatu variabel terhadap suku bunga dapat berubah secara drastis tergantung pada level variabel tersebut. Hansen menunjukkan bahwa respons bank sentral terhadap inflasi mungkin lemah saat tingkat inflasi rendah, namun menjadi sangat agresif ketika inflasi melewati batas psikologis tertentu (Hansen 2011). Artinya, model linear biasa mungkin gagal menangkap urgensi respons kebijakan pada kondisi ekstrem.

Kebijakan moneter bekerja dengan jeda waktu (*time lags*) yang signifikan dan bervariasi. Keputusan perubahan suku bunga hari ini tidak langsung memengaruhi inflasi atau pertumbuhan ekonomi saat itu juga, melainkan memerlukan waktu transmisi sekitar 12 hingga 18 bulan untuk berdampak penuh pada tingkat harga (L. E. Svensson 1997). Terdapat mekanisme umpan balik (*feedback loops*) yang saling memengaruhi, yaitu keputusan suku bunga akan memengaruhi variabel makroekonomi yang pada gilirannya akan menjadi input bagi keputusan suku bunga periode berikutnya. Selain itu, hubungan antarvariabel ini rentan mengalami perubahan struktural (*structural breaks*) akibat guncangan eksternal besar seperti krisis finansial global atau pandemi, yang dapat mengubah pola korelasi antarvariabel secara permanen (Stock dan Watson 2012). Kompleksitas interaksi non-linear, jeda waktu, dan potensi perubahan struktural inilah yang menjadi dasar urgensi penggunaan metode *Machine Learning* dan *Deep Learning* dalam penelitian ini.

II.2 Model Prediksi Ekonomi Berbasis *Time Series*

Prediksi data *deret waktu* (*time series forecasting*) bertujuan untuk memodelkan perilaku dinamis suatu variabel berdasarkan urutan nilai historisnya untuk memproyeksikan nilai masa depan (Box dan Jenkins 1970). Dalam konteks ekonomi makro, tantangan utama prediksi ini terletak pada karakteristik data yang sering kali mengandung autokorelasi tinggi, volatilitas yang berubah-ubah (*heteroskedasticity*), dan interaksi non-linear antarvariabel yang kompleks. Secara historis, pendekatan ini didominasi oleh model ekonometrika linear. Namun, seiring dengan ledakan volume data (*Big Data*) dan kebutuhan akurasi yang lebih tinggi, terjadi pergeseran paradigma menuju penggunaan model *Machine Learning* dan *Deep Learning*. Variasi menegaskan bahwa teknik *Machine Learning* menawarkan alat baru bagi para ekonom untuk menangkap pola non-linear yang sering kali terlewatkan oleh model

regresi standar, menjadikannya komplemen yang penting bagi metode tradisional.

II.2.1 Model Tradisional (Ekonometrika)

Pada subbab ini dibahas model ekonometrika tradisional yang menjadi landasan awal dalam prediksi ekonomi karena kemampuannya dalam menjelaskan hubungan yang struktural antarvariabel secara transparan, meskipun kinerjanya yang terbatas pada asumsi linearitas yang kuat (ini tololong kasih sitasi).

II.2.1.1 Autoregressive Integrated Moving Average (ARIMA)

ARIMA adalah model statistik univariat yang banyak digunakan untuk *time series forecasting*. Model ini dikembangkan oleh Box dan Jenkins pada tahun 1970 dan menjadi standar dalam prediksi ekonomi (Box dan Jenkins 1970). ARIMA terdiri dari tiga komponen utama:

1. Autoregressive (AR)

Komponen ini memodelkan ketergantungan nilai saat ini terhadap nilai-nilai masa lalu. Parameter p menunjukkan jumlah *lag* yang digunakan.

2. Integrated (I)

Komponen ini menangani *non-stationarity* seperti ada tren tertentu, melalui proses *differencing*. Parameter d menunjukkan *order differencing* yang diperlukan untuk mencapai *stationarity*.

3. Moving Average (MA)

Komponen ini memodelkan ketergantungan terhadap kesalahan prediksi atau *error term* masa lalu. Parameter q menunjukkan jumlah *lag error* yang digunakan untuk mengoreksi prediksi.

Secara matematis, model ARIMA(p,d,q) dapat dituliskan seperti pada Persamaan II.2:

$$\phi(B)(1 - B)^d y_t = \theta(B)\epsilon_t \quad (\text{II.2})$$

Dengan $\phi(B)$ adalah polinomial AR, B adalah *backshift operator*, $\theta(B)$ adalah operator polinomial MA, dan ϵ_t adalah *white noise error* (Box dkk. 2015). Keunggulan utama ARIMA adalah efisiensinya pada data linear jangka pendek, tetapi model ini sering gagal menangkap volatilitas tinggi atau perubahan struktural mendadak (*structural breaks*) yang umum terjadi dalam data ekonomi modern.

II.2.1.2 Vector Autoregression (VAR)

VAR adalah perluasan multivariat dari model *autoregressive* yang dikembangkan oleh Sims (Sims 1980). VAR memodelkan sistem variabel endogen, yaitu setiap variabel dijelaskan oleh nilai *lag*-nya sendiri dan nilai *lag* variabel lain dalam sistem. Model VAR(p) untuk k variabel dapat dituliskan sebagai:

$$Y_t = c + A_1 Y_{t-1} + \cdots + A_p Y_{t-p} + u_t \quad (\text{II.3})$$

Dengan Y_t adalah vektor ($K \times 1$) variabel endogen, c adalah vektor konstanta, A_i adalah matriks koefisien, dan u_t adalah vektor *white noise* (Lütkepohl 2011).

II.2.1.3 Vector Error Correction Model (VECM)

VECM adalah model VAR yang telah dimodifikasi untuk menangkap hubungan kointegrasi antar variabel *non-stationary*. Model ini dikembangkan dari konsep *co-integration* oleh Engle dan Granger dan merupakan *special case* dari VAR dengan *restrictions* untuk hubungan jangka panjang (Engle dan Granger 1987). Model VECM dapat dituliskan sebagai:

$$\Delta Y_t = \Pi Y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta Y_{t-i} + \epsilon_t \quad (\text{II.4})$$

Dengan $\Pi = \alpha\beta'$ merepresentasikan hubungan kointegrasi jangka panjang (Johansen 1995).

II.2.2 Model *Machine Learning*

Keterbatasan model ekonometrika tradisional dalam menangkap pola non-linear dapat diatasi oleh algoritma ML. Dalam penelitian ini, model yang digunakan fokus pada pendekatan *Ensemble Learning*. *Ensemble Learning* adalah paradigma ML yang menggabungkan prediksi dari beberapa model dasar (*base learners*) untuk menghasilkan model prediktif yang lebih optimal (Hastie, Tibshirani, dan Friedman 2009). Dua teknik utama dalam *Ensemble Learning* yang digunakan adalah *Bootstrap Aggregating* (Bagging) dan *Boosting*. Bagging bertujuan mengurangi varians dengan melatih model secara paralel pada subset data acak, sedangkan *Boosting* bertujuan mengurangi bias dengan melatih model secara sekuensial untuk memperbaiki kesalahan model sebelumnya.

II.2.2.1 Random Forest Regressor

Random Forest, yang diperkenalkan oleh Breiman, adalah implementasi algoritma Bagging yang menggunakan pohon keputusan (*decision trees*) sebagai model dasar (**breiman2001random**). Algoritma ini bekerja dengan langkah-langkah berikut:

1. Bootstrap Sampling

Model akan mengambil N sampel acak dari data latih dengan pengembalian (*replacement*).

2. Konstruksi Pohon

Model membangun pohon keputusan untuk setiap sampel. Pada setiap pemecahan simpul (*node split*), algoritma hanya akan mempertimbangkan subset acak dari fitur, bukan seluruhnya. Hal ini akan memaksa pohon-pohon yang terbentuk menjadi tidak berkorelasi.

3. Agregasi

Prediksi akhir pada model ini diperoleh dengan menghitung nilai rata-rata output dari seluruh pohon individu.

Mekanisme ini membuat *Random Forest* sangat tahan terhadap *noise* dan *outlier* dalam data keuangan karena kesalahan acak pada satu pohon akan saling menghilangkan saat dirata-ratakan dengan pohon lain. Hal ini meminimalisir risiko *overfitting* pada model dengan signifikan.

II.2.2.2 Extreme Gradient Boosting (XGB)

XGB adalah implementasi algoritma *Gradient Boosting* yang sangat efisien dan terukur, dikembangkan oleh Chen dan Guestrin (Chen dan Guestrin 2016). Berbeda dengan *Random Forest* yang paralel, XGB menggunakan prinsip *Boosting*, yaitu pohon dibangun secara sekuensial. Setiap pohon baru dilatih secara spesifik untuk memprediksi dan memperbaiki sisa kesalahan (*residual*) yang dihasilkan oleh gabungan pohon-pohon sebelumnya. Proses kerja XGB secara rinci meliputi:

1. **Inisialisasi:** Model dimulai dengan prediksi awal konstan.
2. **Menghitung Gradien:** Model akan menghitung gradien dari *loss function* untuk menentukan arah perbaikan kesalahan sebelumnya.
3. **Fit Pohon Baru:** Model akan membangun pohon keputusan baru untuk memprediksi gradien kesalahan.
4. **Pembaruan model:** Model akan menambahkan hasil prediksi pohon baru dengan bobot tertentu (*learning rate*) untuk memperbarui prediksi akhir.
5. **Regularisasi:** XGB menerapkan penalti L1 (*lasso*) dan L2 (*ridge*) pada bobot daun pohon untuk mencegah kompleksitas model yang berlebihan, dan

berisiko terjadinya *overfitting*.

Fungsi objektif XGB dapat dituliskan sebagai:

$$L(\phi) = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k) \quad (\text{II.5})$$

Dengan $\Omega(f_k) = \gamma T + \frac{1}{2} \lambda ||w||^2$ mencegah *overfitting* dengan membatasi kompleksitas pohon. XGB dipilih dalam penelitian ini karena fitur *sparsity-aware split finding* yang dimilikinya dan mampu menangani *missing values* secara otomatis, kemampuannya untuk menangkap interaksi non-linear yang kompleks antarvariabel ekonomi, regularisasi bawaan yang membuatnya lebih aman untuk jumlah sampel terbatas, dan kecepatan komputasi yang tinggi dibanding algoritma *boosting* lainnya (Chen dan Guestrin 2016).

II.2.3 Model *Deep Learning*

Model *Deep Learning* berbasis jaringan saraf tiruan (*Neural Networks*) digunakan untuk menangkap struktur dependensi jangka panjang (*long-term dependencies*) yang sering kali luput dari model *Machine Learning* konvensional. Sezer dkk. mencatat bahwa dalam domain keuangan, model berbasis LSTM menjadi arsitektur dominan karena keunggulannya dalam memproses data sekuensial (Sezer, Gudelek, dan Ozbayoglu 2020).

II.2.3.1 Long Short-Term Memory (LSTM)

LSTM adalah varian khusus dari *Recurrent Neural Networks* (RNN) yang dirancang oleh Hochreiter dan Schmidhuber untuk mengatasi masalah *vanishing gradient*, yaitu kondisi ketika RNN gagal mempelajari pola dari data masa lalu yang jauh karena gradien kesalahan menyusut mendekati nol selama proses pelatihan (*backpropagation*) (Hochreiter dan Schmidhuber 1997). Solusi LSTM terletak pada struktur sel memori (*memory cell*) yang dilindungi oleh tiga gerbang logika (*gates*) untuk mengatur aliran informasi:

1. Forget Gate

Model memutuskan informasi mana dari memori masa lalu yang sudah tidak relevan dan harus dibuang.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (\text{II.6})$$

2. Input Gate

Model akan menentukan informasi baru apa dari input saat ini yang penting untuk disimpan ke dalam memori.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (\text{II.7})$$

3. Output Gate

Model akan mengatur informasi apa dari memori internal yang akan diteruskan sebagai output ke langkah waktu berikutnya.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (\text{II.8})$$

Mekanisme ini memungkinkan LSTM untuk mempertahankan informasi penting (seperti dampak kebijakan moneter tahun lalu) dalam jangka waktu panjang dan secara selektif melupakan gangguan jangka pendek (*noise*), menjadikannya sangat efektif untuk data ekonomi yang memiliki *lag effect*.

II.2.3.2 Bidirectional Long Short-Term Memory (Bi-LSTM)

Bi-LSTM merupakan pengembangan arsitektur LSTM yang diperkenalkan oleh Schuster dan Paliwal (Schuster dan Paliwal 1997). Jika LSTM standar hanya memproses informasi secara maju (*forward*) dari masa lalu ke masa kini ($t - 1 \rightarrow t$), Bi-LSTM memproses data dalam dua arah sekaligus menggunakan dua lapisan LSTM terpisah:

1. **Forward Layer:** Model memproses urutan data dari awal ke akhir untuk menangkap konteks masa lalu.
2. **Backward Layer:** Model akan memproses urutan data dari akhir ke awal untuk menangkap konteks masa depan dalam data latih.

Output akhir pada setiap langkah waktu (y_t) merupakan gabungan dari kedua lapisan tersebut yang biasanya melalui konkatenasi atau penjumlahan:

$$y_t = W_{\vec{h}} \vec{h}_t + W_{\overleftarrow{h}} \overleftarrow{h}_t + b_y \quad (\text{II.9})$$

Output dari kedua lapisan ini kemudian digabungkan untuk menghasilkan representasi fitur yang lebih baik. Dalam konteks prediksi *deret waktu*, pendekatan ini memungkinkan model untuk memahami struktur data secara lebih holistik, mendetecti pola kausalitas dan korelasi yang mungkin terlewat jika hanya dilihat dari satu arah sekuensial (Schuster dan Paliwal 1997).

II.2.4 Metode Evaluasi Model

Pada subbab ini dijelaskan berbagai metrik yang digunakan untuk menilai kinerja model prediksi secara objektif. Akan digunakan beberapa metrik evaluasi standar yang mengukur akurasi dan kecocokan model.

II.2.4.1 Root Mean Squared Error (RMSE)

RMSE mengukur standar deviasi dari kesalahan prediksi, yaitu perbedaan antara nilai yang diprediksi oleh model dan nilai sebenarnya (Hodson 2022). Semakin kecil nilai RMSE menunjukkan kesalahan prediksi yang lebih kecil, yang berarti model memiliki akurasi yang lebih tinggi. Rumus dari RMSE adalah sebagai berikut:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (\text{II.10})$$

Dengan n adalah banyaknya observasi, y_i adalah nilai sebenarnya, dan \hat{y}_i adalah nilai prediksi model. RMSE sensitif terhadap *outliers* dan dalam satuan yang sama dengan variabel target, memudahkan interpretasi (Hodson 2022).

II.2.4.2 Mean Absolute Percentage Error (MAPE)

MAPE mengukur *error* dalam persentase, memudahkan perbandingan antar *dataset* dengan skala berbeda (Hyndman dan Koehler 2006). Rumus dari MAPE adalah sebagai berikut:

$$MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (\text{II.11})$$

Dengan y_i adalah nilai aktual pada periode ke- i , \hat{y}_i adalah nilai prediksi pada periode ke- i , n adalah jumlah total observasi, dan nilai absolut $|\dots|$ memastikan bahwa setiap *error* selalu bernilai positif. Kelemahan utama MAPE adalah persentase kesalahan akan menjadi tak terhingga atau tidak terdefinisi jika terdapat nilai aktual $y_i = 0$ dalam data observasi (Hyndman dan Koehler 2006).

II.2.4.3 R-squared (R^2)

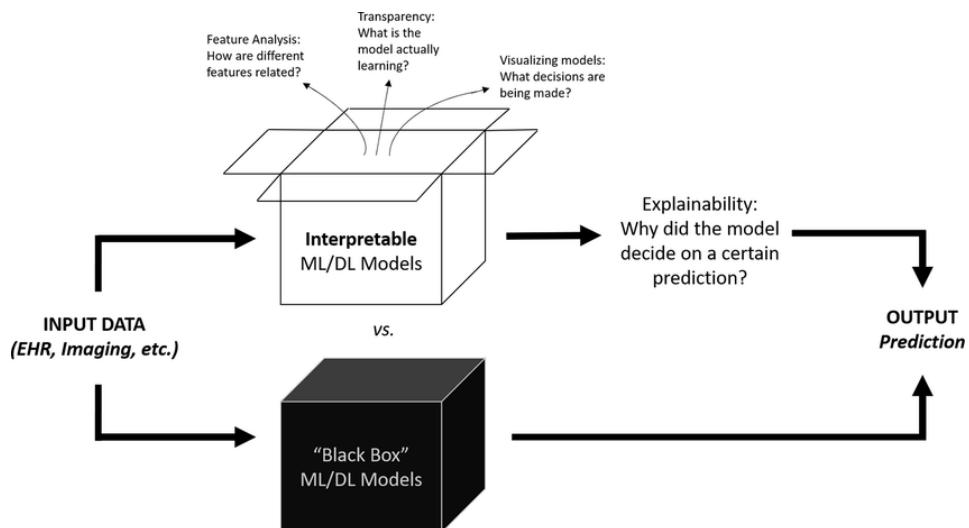
Koefisien determinasi atau R^2 mengukur seberapa baik variasi dalam variabel dependen dapat dijelaskan oleh model regresi. R^2 merepresentasikan proporsi varians

total yang diperlihatkan oleh garis regresi (Nagelkerke 1991). Rumus dari R^2 adalah sebagai berikut:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (\text{II.12})$$

Dengan y_i adalah nilai aktual observasi ke- i , \hat{y}_i adalah nilai prediksi observasi ke- i , dan \bar{y} adalah nilai rata-rata dari seluruh nilai aktual. Pembilang $\sum(y_i - \hat{y}_i)^2$ dikenal sebagai *Residual Sum of Squares* (RSS), yaitu jumlah kuadrat kesalahan prediksi, sementara penyebut $\sum(y_i - \bar{y})^2$ adalah *Total Sum of Squares* (TSS), yang merepresentasikan total variabilitas dalam data aktual. Nilai R^2 berkisar antara 0 hingga 1, di mana nilai yang mendekati 1 mengindikasikan bahwa model mampu menjelaskan sebagian besar variabilitas data, sedangkan nilai mendekati 0 menunjukkan kinerja prediksi yang buruk.

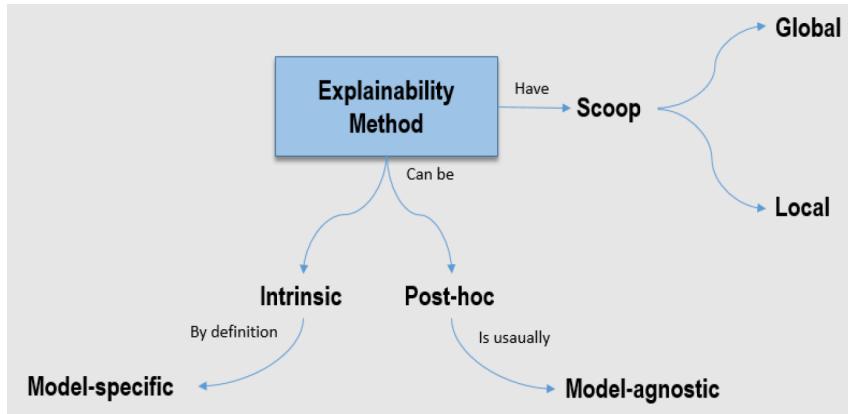
II.3 Explainable Artificial Intelligence (XAI)



Gambar II.1 Ilustrasi *Black Box* dan *Interpretable* Model (Hui dkk. January 2022)

Kompleksitas model ML dan DL seringkali mengorbankan interpretabilitas model dibanding akurasi prediksi. Rudin menyoroti bahwa untuk keputusan berisiko tinggi seperti kebijakan publik dan penetapan suku bunga, penggunaan model *black box* tanpa penjelasan yang memadai menjadi problematik (Rudin 2019). Oleh karena itu, ada XAI untuk membuat hasil prediksi A dapat dipahami dan dipercaya oleh para pengambil keputusan (Adadi dan Berrada 2018).

II.3.1 Taksonomi Metode XAI



Gambar II.2 Ilustrasi Taksonomi Metode XAI (Adadi dan Berrada 2018)

Adadi dan Berrada mengklasifikasikan metode XAI ke dalam tiga dimensi utama (Adadi dan Berrada 2018):

1. **Kompleksitas Model (Complexity Related Methods)**
 - a. *Intrinsic*: Metode ini menggunakan model yang mudah diinterpretasi sejak desain awal, seperti regresi linear. Model ini memang transparan, tetapi sering mengorbankan akurasi prediksi.
 - b. *Post-hoc*: Metode ini diterapkan setelah model dilatih untuk menjelaskan prediksi dari model kompleks yang tidak transparan. Pendekatan ini memungkinkan penggunaan model dengan akurasi tinggi sambil memberikan penjelasan, tanpa mengubah struktur internal model asli. SHAP dan LIME adalah contoh metode post-hoc yang populer.
2. **Cakupan Penjelasan (Scope Related Methods)**
 - a. *Global Interpretability*: Cakupan ini bertujuan untuk memahami logika model secara keseluruhan pada seluruh *dataset*. Metode ini menjawab pertanyaan agregat seperti "Fitur apa yang rata-rata memengaruhi prediksi Bi-Rate?" dan berguna untuk keputusan dalam kebijakan publik.
 - b. *Local Interpretability*: Cakupan ini fokusnya pada penjelasan untuk satu instans prediksi tunggal untuk menjawab pertanyaan spesifik seperti "Kenapa model memprediksi kenaikan BI-Rate pada bulan Januari 2024?". LIME adalah contoh representatif untuk penjelasan lokal yang dapat mengidentifikasi pengaruh fitur terhadap prediksi individual.
3. **Ketergantungan Model (Model Related Methods)**
 - a. *Model-Specific*: Metode ini dirancang untuk arsitektur model tertentu karena memanfaatkan struktur internalnya secara eksplisit. Contohnya

- visualisasi filter pada *Convolutional Neural Network*, *Attention Mechanism* pada model *Transformer*, atau *gradient-based methods* pada *neural networks*. Kelemahan metode ini adalah pembatasan pilihan model karena harus memilih model yang menyediakan interpretabilitas tertentu.
- b. *Model-Agnostic*: Metode ini dapat diterapkan pada semua jenis model ML, memperlakukan model sebagai *black-box* yang hanya mengamati hubungan *input-output*. Pendekatan ini lebih fleksibel dan memungkinkan perbandingan penjelasan antarmodel yang berbeda secara langsung (seperti XGB dengan LSTM).

II.3.1.1 SHapley Additive exPlanations (SHAP)

SHAP adalah metode interpretasi yang diperkenalkan oleh Lundberg dan Lee yang menyatukan berbagai teknik interpretasi sebelumnya di bawah kerangka teori permainan kooperatif (*cooperative game theory*) (Lundberg dan Lee 2017). Konsep dasarnya diadopsi dari *Shapley Value* yang dikemukakan oleh Lloyd Shapley pada tahun 1953 untuk mendistribusikan keuntungan secara adil di antara pemain dalam sebuah koalisi. Dalam konteks ML, "permainan" ini maksudnya adalah tugas prediksi, "keuntungan" adalah selisih antara prediksi aktual dengan rata-rata prediksi, dan "pemain" yang dimaksud adalah fitur-fitur yang memengaruhi model. SHAP menghitung kontribusi marginal rata-rata dari setiap fitur terhadap prediksi model di semua kemungkinan kombinasi fitur.

Nilai SHAP (ϕ_i) untuk fitur i didefinisikan secara matematis sebagai berikut:

$$\phi_i = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!(|F| - |S| - 1)!}{|F|!} [f_S(x_S \cup \{i\}) - f_S(x_S)] \quad (\text{II.13})$$

Dengan F adalah himpunan seluruh fitur dalam model, S adalah subset fitur yang tidak menyertakan fitur i atau $S \subseteq F \setminus \{i\}$. Notasi $|S|$ dan $|F|$ menyatakan jumlah elemen dalam subset S dan jumlah total fitur. Fungsi $f_S(x_S)$ adalah prediksi model ketika hanya fitur-fitur di dalam subset S yang diaktifkan, sedangkan $f_S(x_S \cup \{i\})$ adalah prediksi ketika fitur i ditambahkan ke subset yang sama. Selisih keduanya, yaitu $f_S(x_S \cup \{i\}) - f_S(x_S)$, menggambarkan kontribusi marginal fitur i dalam konteks subset S . Kontribusi tersebut lalu dirata-ratakan dengan bobot $\frac{|S|!(|F| - |S| - 1)!}{|F|!}$ untuk memastikan bahwa semua urutan munculnya fitur diperlakukan secara adil.

Lundberg dan Lee menunjukkan bahwa dalam kelas metode *additive feature attribution*, SHAP adalah satu-satunya solusi yang memenuhi tiga sifat penting sekaligus,

yaitu *local accuracy* (penjumlahan seluruh nilai SHAP ditambah nilai dasar sama dengan prediksi model), *missingness* (fitur yang absen tidak boleh berkontribusi), dan *consistency* (jika kontribusi suatu fitur di model meningkat, maka nilai SHAP-nya tidak boleh menurun) (Lundberg dan Lee 2017). Sifat-sifat ini menjadi dasar teoretis yang kuat sehingga SHAP banyak digunakan sebagai standar untuk menjelaskan model kompleks, baik secara global maupun lokal.

II.3.1.2 Local Interpretable Model-Agnostic Explanations (LIME)

LIME adalah metode interpretasi yang diperkenalkan oleh Ribeirio, Singh, dan Guestrin pada tahun 2016. Fokusnya adalah penjelasan lokal yang cepat dan intuitif. Perbedaannya adalah SHAP menghitung kontribusi eksak berdasarkan *game theory*, sedangkan LIME bekerja dengan asumsi bahwa model sekomples apa pun akan bersifat linear secara lokal di sekitar instans prediksi tertentu (Ribeiro, Singh, dan Guestrin 2016).

Mekanisme kerja LIME sebagai berikut:

1. LIME membangkitkan data sintetis baru di sekitar titik data yang ingin dijelaskan dengan menambahkan *noise* atau pengacakan (perturbasi data).
2. Data sintetis kemudian dimasukkan ke model asli (misalnya LSTM) untuk mendapatkan label prediksinya.
3. Data sintetis diberi bobot berdasarkan kedekatannya (jarak *euclidean*) dengan titik data asli. Data yang lebih dekat akan memiliki bobot yang lebih besar.
4. Sebuah model linear sederhana seperti regresi linear dilatih pada data sintetis berbobot tersebut. Koefisien model linear ini yang kemudian akan diambil sebagai penjelasan dari model (*explanation*).

Fungsi objektif LIME didefinisikan sebagai:

$$\xi(x) = \underset{g \in G}{\operatorname{argmin}} L(f, g, \pi_x) + \Omega(g) \quad (\text{II.14})$$

Dengan f adalah model asli yang ingin dijelaskan, g adalah model penjelasan dari kelas model interpretable G , $L(f, g, \pi_x)$ mengukur seberapa tidak setia (*unfaithful*) model penjelasan terhadap model asli pada lingkungan lokal yang ditentukan oleh π_x , dan $\Omega(g)$ adalah istilah yang mengontrol kompleksitas model penjelasan agar tetap sederhana dan mudah dipahami. Hasil akhirnya adalah kumpulan koefisien pada fitur-fitur *interpretable* yang menunjukkan fitur mana yang mendukung atau menentang prediksi model untuk instans tersebut. Kelebihan utama LIME adalah

sifatnya yang benar-benar *model-agnostic* dan fleksibel, sehingga dapat diterapkan pada berbagai jenis model dan data, sekaligus memberikan penjelasan lokal yang intuitif dan mudah dikomunikasikan kepada pengguna non-teknis (Ribeiro, Singh, dan Guestrin 2016).

II.3.1.3 Permutation Feature Importance (PFI)

Permutation Feature Importance (PFI) adalah metode interpretasi model yang diperkenalkan oleh Breiman (2001) untuk mengukur pentingnya fitur dengan mengevaluasi penurunan kinerja model setelah fitur tersebut di-permutasi atau diacak (**breiman2001random**). Ide PFI adalah bahwa sebuah fitur dianggap penting jika mengacak nilainya meningkatkan kesalahan prediksi model karena hal ini menunjukkan bahwa model bergantung pada fitur tersebut untuk membuat prediksi. Sebaliknya, jika mengacak fitur tidak mengubah kesalahan prediksi model yang signifikan, maka fitur tersebut dianggap tidak penting (**molnar2022interpretable**).

Algoritma PFI secara umum bekerja dengan langkah-langkah berikut:

1. Hitung kesalahan referensi model pada dataset asli.
2. Untuk setiap fitur j , buat data baru dengan mengacak (permutasi) nilai-nilai fitur j , sehingga hubungan antara fitur dan target hancur.
3. Hitung kesalahan model pada data yang telah diacak tersebut.
4. Terakhir, ukur pentingnya fitur sebagai selisih atau rasio antara kesalahan referensi dan kesalahan setelah permutasi.

Secara matematis, PFI dapat dituliskan sebagai:

$$\text{PFI}_j = e_{\text{perm},j} - e_{\text{orig}} \quad (\text{II.15})$$

Dengan e_{orig} adalah kesalahan referensi model pada data asli dan $e_{\text{perm},j}$ adalah kesalahan model setelah fitur j dipermutasi.

Keuntungan PFI adalah bersifat *model-agnostic*, dapat diterapkan pada berbagai jenis model, dan tidak memerlukan pelatihan ulang model setelah melakukan permutasi. Namun, PFI memiliki keterbatasan penting ketika fitur-fitur dalam dataset saling berkorelasi (multikolinieritas). Ketika fitur-fitur berkorelasi, permutasi satu fitur masih menyisakan informasi prediktif dari fitur-fitur yang berkorelasi dengannya, sehingga penurunan kinerja model menjadi kurang terlihat dan interpretasi menjadi bias (**molnar2022interpretable**).

II.3.2 Penerapan XAI dalam Prediksi Ekonomi

Penerapan XAI dalam sektor ekonomi ini bertujuan untuk menjadi penghubung antara kinerja prediktif dan akuntabilitas regulasi. Bussmann dkk. menunjukkan bahwa penggunaan SHAP dalam manajemen risiko kredit memungkinkan institusi keuangan untuk tidak hanya memprediksi gagal bayar (*default*), tetapi juga mengidentifikasi faktor pendorong spesifik (misalnya rasio likuiditas rendah) bagi setiap nasabah (Bussmann dkk. 2021). Dalam konteks bank sentral, teknik XAI memungkinkan pembuat kebijakan membedah model DL untuk memastikan bahwa prediksi suku bunga didasarkan pada sinyal ekonomi yang valid seperti *output gap*, bukan korelasi yang salah dalam data (Chakraborty dan Joseph 2017).

II.4 Large Language Model (LLM)

Large Language Model (LLM) adalah model probabilistik yang dibuat untuk memproses dan menghasilkan teks dengan tingkat koherensi tinggi untuk menyelesaikan berbagai tugas generalisasi bahasa (Zhao dkk. 2023).

Perkembangan teknologi pemrosesan bahasa alami telah melalui beberapa fase evolusi yang signifikan sebelum mencapai kapabilitas LLM modern. Perkembangan teknologi ini diklasifikasikan ke dalam empat fase utama (Zhao dkk. 2023):

- 1. Statistical Language Models (SLM)**

Generasi awal *language models* yang dikembangkan berdasarkan prinsip statistik, khususnya *markov assumption*, yang memprediksi kata berikutnya berdasarkan probabilitas kemunculan kata sebelumnya. SLM banyak digunakan dalam sistem pengambilan informasi (*information retrieval*), tetapi memiliki kelemahan mendasar dalam menangani data berdimensi besar karena ledakan kombinasi kata (*curse of dimensionality*).

- 2. Neural Language Models (NLM)**

NLM memperkenalkan penggunaan jaringan saraf tiruan (*neural networks*) seperti RNN dan LSTM untuk menghitung probabilitas rangkaian kata. Pendekatan ini memungkinkan model untuk mempelajari representasi fitur kata secara terdistribusi (*word embeddings*), sehingga mampu menangkap semantik kata dengan lebih baik dibandingkan model statistik murni. Namun, arsitektur ini masih memproses data secara sekuensial, membatasi kemampuannya dalam menangkap konteks jangka panjang.

- 3. Pre-trained Language Models (PLM)**

PLM dimulai dengan dikenalnya arsitektur *Transformer* yang memungkinkan

pemrosesan paralel melalui mekanisme *Self-Attention*. Model seperti BERT (*Bidirectional Encoder Representations from Transformers*) memperkenalkan paradigma *pre-training* pada korpus masif tanpa label, diikuti dengan *fine-tuning* untuk tugas yang spesifik.

4. *Large Language Models* (LLM)

LLM adalah pengembangan ekstrem dari PLM, baik dari segi jumlah parameter maupun ukuran data latih. Pengembangan ini memunculkan kemampuan baru (*emergent abilities*) yang tidak dimiliki model kecil, seperti penalaran logis (*reasoning*) dan pemahaman instruksi kompleks (*few-shot learning*). Contoh LLM adalah GPT (OpenAI) dan Llama (Meta) yang mampu menyelesaikan tugas tanpa pelatihan spesifik sebelumnya (Brown dkk. 2020).

II.4.1 OpenAI GPT-4o

GPT-4o (*Generative Pre-trained Transformer 4 - Omni*) adalah model LLM terbaru yang dikembangkan oleh OpenAI. Sebagai evolusi dari seri GPT-4, model ini dirancang untuk menerima input berupa teks, audio, dan gambar, serta menghasilkan respons teks dengan tingkat penalaran yang superior (OpenAI 2024).

Beberapa keunggulan GPT-4o (OpenAI 2024):

1. GPT-4o memiliki kinerja yang baik dalam tugas-tugas yang membutuhkan pemahaman kausalitas kompleks.
2. Kapasitas token yang besar (128K *context window*) memungkinkan model untuk memproses riwayat data yang panjang dan instruksi *prompt* dengan detail tanpa kehilangan konteks analisis.
3. Kecepatan *inference* yang signifikan, sehingga hemat biaya untuk eksperimen iteratif dalam generasi narasi.
4. Kemampuan multimodal native dalam satu model untuk memudahkan integrasi output numerik (CSV/JSON dari model prediksi) ke dalam format naratif.

II.4.2 Meta Llama

Llama (*Large Language Model Meta AI*) adalah keluarga model LLM yang dikembangkan dan dirilis oleh Meta AI dimulai dari Februari 2023 dengan tujuan mendemokratisasi akses terhadap model bahasa berskala besar untuk komunitas peneliti (AI February 2023). Berbeda dengan GPT-4o yang merupakan model proprietary, Llama dirilis dengan bobot model yang tersedia untuk publik, sehingga lebih mudah diakses oleh peneliti dan developer yang memiliki infrastruktur terbatas. Llama

menggunakan arsitektur *decoder-only transformer* yang mirip dengan GPT-3, tetapi dengan beberapa perbaikan, seperti penggunaan fungsi aktivasi SwiGLU, *rotary positional embeddings* (RoPE), dan RMSNorm *normalization* untuk keseimbangan kapabilitas dan efisiensi komputasi.

Beberapa keunggulan Llama (AI February 2023):

1. Fleksibilitas ukuran model dari 7B hingga 70B parameter yang bisa adaptasi dengan infrastruktur penelitian yang terbatas.
2. Komputasi yang cepat pada perangkat lokal tanpa memerlukan koneksi internet atau API, sehingga lebih hemat biaya untuk eksperimen skala besar.
3. Transparansi penuh atas arsitektur dan data pelatihan mampu memfasilitasi pemahaman mendalam tentang perilaku model untuk tugas pembuatan narasi dari output XAI.
4. Fleksibilitas *licensing* yang memungkinkan integrasi ke dalam sistem produksi dengan kontrol penuh atas *deployment* dan *governance* data.

II.5 Penelitian Terkait

Subbab ini membahas kondisi penelitian terkini terkait topik yang dibahas. Pemahaman terhadap kondisi aktual ini menjadi landasan dalam mengidentifikasi peluang pengembangan lebih lanjut. Selain itu, penelitian-penelitian terdahulu juga dapat menjadi acuan dalam pengerjaan tugas akhir ini.

II.5.1 Machine Learning Applications in Central Banking (Chakraborty dan Joseph 2017)

Penelitian ini diterbitkan dalam *Bank of England Staff Working Paper* yang bertujuan mengeksplorasi potensi algoritma *Machine Learning* (ML) untuk menggantikan model ekonometrika standar dalam tugas bank sentral. Menggunakan data indikator makroekonomi Inggris, dibandingkan kinerja model *Random Forest*, *Support Vector Machines*, dan *Neural Networks*, lalu dibandingkan dengan model tradisional. Hasil eksperimen menunjukkan bahwa model ML secara konsisten lebih unggul dibandingkan model tradisional dalam akurasi prediksi di luar sampel (*out-of-sample accuracy*), terutama pada rentang waktu yang lebih panjang. Namun, kesimpulan utama dari *paper* ini adalah meskipun ML menawarkan akurasi superior, penggunaannya dalam kebijakan terhambat oleh masalah interpretabilitas (*black box*).

II.5.2 Explainable Machine Learning in Credit Risk Management (Bussmann dkk. 2021)

Bussmann dkk. mempublikasikan penelitian di jurnal *Computational Economics* yang mengembangkan kerangka kerja penilaian risiko kredit menggunakan algoritma XGB yang dijelaskan dengan metode *Shapley Values* (SHAP). Metodologi ini diuji secara empiris pada dataset pinjaman *peer-to-peer* (P2P) untuk memprediksi gagal bayar (*default*). Hasil studi mendemonstrasikan bahwa integrasi SHAP mampu memberikan penjelasan lokal yang konsisten untuk setiap keputusan kredit, sekaligus mengidentifikasi faktor risiko global yang sejalan dengan intuisi ekonomi. *Paper* ini memberikan *blueprint* metodologis tentang bagaimana menggabungkan model *Gradient Boosting* dengan SHAP untuk menghasilkan sistem prediksi keuangan yang transparan dan sesuai dengan regulasi.

II.5.3 Temporal Data Meets LLM – Explainable Financial Time Series Forecasting (Yu, Chen, dan Lu 2023)

Penelitian ini jadi inisiator dalam ranah integrasi *Large Language Models* untuk data keuangan. Yu dkk. mengusulkan kerangka kerja multimodal, yaitu model prediksi numerik (LSTM) digabungkan dengan LLM yang memproses data teks berita keuangan untuk memprediksi pergerakan harga saham NASDAQ-100. Metode yang digunakan melibatkan teknik *prompt engineering* untuk memandu LLM melakukan penalaran terhadap sinyal pasar. Hasil studi menunjukkan bahwa integrasi ini tidak hanya meningkatkan akurasi prediksi dibandingkan model *baseline*, tetapi juga menghasilkan penjelasan tekstual yang logis mengenai penyebab pergerakan pasar. *Paper* ini membuktikan bahwa LLM dapat berperan efektif sebagai mesin interpretasi naratif untuk data *deret waktu*.

II.5.4 Explainable-AI powered stock price prediction using time series transformers: A Case Study on BIST100 (Calik dkk. 2025)

Penelitian terbaru oleh Calik dkk. yang dipublikasikan di *arXiv* ini secara spesifik meneliti integrasi metode XAI dengan model *Transformer-based* untuk prediksi harga saham sektor perbankan. Studi ini menggunakan model *Time Series Transformer* yang diperkaya dengan indikator teknikal, kemudian menerapkan SHAP dan LIME untuk memberikan transparansi pada *output* model. Hasil penelitian menunjukkan bahwa model berbasis *Transformer* mengungguli model tradisional, dan penggunaan kombinasi SHAP serta LIME terbukti efektif dalam memberikan wawasan interpretatif kepada investor mengenai fitur mana (misalnya volume transaksi

atau tren global) yang paling memengaruhi harga.

BAB III

ANALISIS MASALAH

Pada bab ini dibahas terkait permasalahan penetapan suku bunga acuan Bank Indonesia, analisis kondisi sistem saat ini, identifikasi kebutuhan pengguna, serta analisis pemilihan solusi yang tepat untuk mengatasi permasalahan tersebut. Analisis kondisi ini dilakukan dengan mengacu pada tahap awal CRISP-DM yang telah dibahas pada Bab I, yaitu *business understanding* dan *data understanding*.

III.1 Analisis Kondisi Saat Ini

Pada subbab ini dibahas terkait model konseptual sistem penetapan suku bunga acuan Bank Indonesia saat ini, serta masalah yang ada pada sistem tersebut. Analisis kondisi ini dilakukan dengan mengacu pada tahap awal CRISP-DM yang telah dibahas pada Bab I, yaitu *business understanding*. Tahap ini memberikan pemahaman terkait konteks masalah yang akan diselesaikan dalam tugas akhir ini.

III.1.1 Business Understanding

Penetapan BI-Rate merupakan salah satu bentuk kebijakan moneter utama Bank Indonesia untuk mencapai target inflasi dan menjaga stabilitas nilai tukar. Berdasarkan analisis kondisi saat ini, kebutuhan bisnis yang mendesak adalah transparansi dan edukasi publik.

Meskipun data ekonomi tersedia secara luas, proses pengambilan keputusan di balik penetapan BI-Rate sering kali dipersepsikan sebagai *black box*. Publik mengetahui data masukan (inflasi, PDB, kurs) dan hasil keluaran (BI-Rate), tetapi tidak memahami mekanisme internal atau bobot kontribusi setiap variabel terhadap keputusan tersebut.

Oleh karena itu, diperlukan sebuah sistem yang mampu menjembatani kesenjang-

an antara data kuantitatif yang kompleks dengan pemahaman masyarakat. Sistem ini harus mampu memprediksi arah kebijakan BI-Rate dengan akurasi yang optimal sekaligus membuka *black box* tersebut melalui penjelasan yang transparan. Tujuan akhirnya adalah meningkatkan akuntabilitas kebijakan, mengurangi asimetri informasi, dan membantu membentuk ekspektasi pasar yang lebih rasional melalui narasi berbasis data yang mudah dipahami.

III.1.2 Gambaran Sistem Saat Ini

Proses penetapan BI-Rate yang berjalan saat ini melibatkan analisis mendalam dan prosedur bertingkat di internal Bank Indonesia.



Gambar III.1 Proses Penetapan BI-Rate Saat Ini

Sebagaimana diilustrasikan pada Gambar III.1, proses dimulai dari pengumpulan data ekonomi makro domestik dan global. Departemen terkait, khususnya Departemen Kebijakan Ekonomi dan Moneter (DKEM), melakukan analisis dan simulasi model ekonometrika untuk menyusun rekomendasi kebijakan. Hasil analisis ini dibahas dalam Rapat Dewan Gubernur (RDG) yang berlangsung selama dua hari: hari pertama untuk evaluasi kondisi ekonomi dan hari kedua untuk penetapan kebijakan (Bank Indonesia 2023).

Setelah keputusan ditetapkan, Bank Indonesia mengumumkan hasil RDG melalui siaran pers resmi. Namun, terdapat keterbatasan dalam sistem komunikasi saat ini:

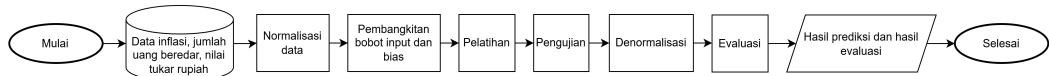
1. Penjelasan Kualitatif

Narasi dalam siaran pers umumnya bersifat makro dan kualitatif. Tidak ada rincian kuantitatif eksplisit yang menyatakan seberapa besar bobot kontribusi spesifik dari setiap indikator (misalnya berapa persen kenaikan inflasi memengaruhi keputusan kenaikan suku bunga bulan ini).

2. Kesenjangan Informasi

Pengguna awam dapat kesulitan memvalidasi rasionalitas keputusan penetapan BI-Rate secara mandiri karena ketiadaan alat bantu analisis yang transparan.

Selain sistem internal BI, terdapat juga berbagai model prediksi yang dikembangkan oleh penelitian terdahulu, seperti yang ditunjukkan pada Gambar III.2.



Gambar III.2 Gambaran Sistem Prediksi BI-Rate Terdahulu (Putri, Adikara, dan Adinugroho May 2019)

Sistem prediksi terdahulu hanya fokus pada akurasi prediksi. Alurnya bersifat linear, mulai dari *input* data, normalisasi, pemodelan, hingga *output* hasil prediksi. Sistem ini belum memiliki fitur interpretabilitas (*explainability*). Pengguna hanya menerima angka hasil prediksi tanpa mengetahui kenapa angka tersebut muncul, sehingga sistem ini kurang efektif untuk tujuan edukasi dan analisis kebijakan.

III.2 Analisis Kebutuhan

Pada subbab ini dibahas kebutuhan spesifik yang harus dipenuhi oleh sistem untuk menjawab permasalahan pengguna yang telah diidentifikasi. Analisis ini mencakup identifikasi masalah dari sudut pandang pengguna, serta spesifikasi kebutuhan fungsional dan non-fungsional sistem yang akan dikembangkan.

III.2.1 Identifikasi Masalah Pengguna

Berdasarkan observasi, masalah utama yang dihadapi pengguna adalah:

1. Pengguna awam sulit memahami hubungan sebab-akibat yang kompleks antara fluktuasi indikator ekonomi dengan perubahan BI-Rate tanpa bantuan alat analisis kuantitatif.
2. Pengguna tidak memiliki akses ke alat bantu yang dapat memberikan analisis naratif mengenai kondisi ekonomi terkini, sehingga sering kali bergantung pada opini pengamat yang subjektif atau tertunda (*lagging*).

III.2.2 Kebutuhan Fungsional

Kebutuhan fungsional mendefinisikan fitur dan kemampuan spesifik yang harus dimiliki oleh sistem yang akan dikembangkan. Rincian kebutuhan fungsional dapat dilihat di Tabel III.1.

Tabel III.1 Kebutuhan Fungsional

| Kode | Kebutuhan Fungsional | Deskripsi |
|------|-----------------------------|--|
| F-01 | <i>Input Data</i> | Sistem mampu menerima <i>input</i> data indikator makroekonomi yang akan dilatih melalui <i>dataset</i> . |
| F-02 | Prediksi BI-Rate | Sistem mampu memproses data input dan menghasilkan prediksi nilai BI-Rate. |
| F-02 | Interpretasi Fitur | Sistem mampu menghitung dan menampilkan skor kontribusi dari setiap variabel input terhadap hasil prediksi. |
| F-03 | Visualisasi Data Interaktif | Sistem mampu menampilkan grafik tren historis dan plot interpretasi XAI untuk memudahkan analisis visual pengguna. |
| F-04 | Generasi Narasi Otomatis | Sistem mampu mengonversi hasil prediksi dan interpretasi menjadi teks narasi kebijakan menggunakan LLM. |

III.2.3 Kebutuhan Non-Fungsional

Kebutuhan non-fungsional mendefinisikan batasan kualitas sistem. Rincian kebutuhan non-fungsional dapat dilihat di Tabel III.2.

Tabel III.2 Kebutuhan Non-Fungsional

| Kode | Kebutuhan Non-Fungsional | Deskripsi |
|-------|-------------------------------|---|
| NF-01 | <i>Understandability</i> | Output narasi dan visualisasi harus menggunakan bahasa dan format yang mudah dimengerti oleh pengguna. |
| NF-02 | <i>Performance Efficiency</i> | Sistem (terutama modul generasi narasi) harus mampu menghasilkan respons dalam waktu yang wajar (< 1 menit) setelah data diinput. |
| NF-03 | <i>Reproducibility</i> | Eksperimen model dan kode program harus dapat dijalankan ulang dengan hasil yang konsisten untuk memastikan validitas ilmiah. |

Berlanjut ke halaman berikutnya...

Tabel III.2 – lanjutan dari halaman sebelumnya

| Kode | Kebutuhan Non-Fungsional | Deskripsi |
|-------|--------------------------|--|
| NF-04 | <i>Maintainability</i> | Kode program harus terstruktur dengan baik, terdokumentasi dengan jelas, dan mudah dimodifikasi untuk perbaikan atau pengembangan. |
| NF-05 | <i>Scalability</i> | Sistem harus dapat dikembangkan menjadi lebih besar lagi di kemudian hari dengan mudah. |

III.3 Analisis Pemilihan Solusi

Pada subbab ini dilakukan analisis terhadap alternatif pendekatan solusi untuk memenuhi kebutuhan sistem, serta pemilihan solusi terbaik. Analisis ini menggunakan pendekatan *comparative selection*, yaitu berbagai kandidat metode dievaluasi berdasarkan karakteristik data, serta kebutuhan pengguna akan transparansi.

Berikut adalah analisis pemilihan metode untuk tiga komponen utama sistem, yaitu model prediksi, XAI, dan LLM.

III.3.1 Seleksi Model Prediksi

Tantangan utama dalam data deret waktu ekonomi adalah menangkap pola linier. Bagian ini akan menganalisis kandidat model dari ekonometrika tradisional, *machine learning*, dan *deep learning*.

Tabel III.3 Matriks Seleksi Model Prediksi

| Kategori | Algoritma | Analisis |
|--------------|-----------|--|
| Ekonometrika | ARIMA | Metode ini adalah standar dasar untuk pemodelan deret waktu univariat. Sangat relevan dijadikan <i>baseline</i> untuk mengukur seberapa baik model kompleks dapat menghasilkan prediksi yang lebih baik dibandingkan model linear sederhana. |

Berlanjut ke halaman berikutnya...

Tabel III.3 – lanjutan

| Kategori | Algoritma | Analisis |
|------------------|---------------|--|
| | VAR | Metode ini mampu menangkap hubungan linear timbal balik (interdependensi) antar variabel makroekonomi, sehingga bisa jadi pembanding yang penting untuk hasil prediksi. |
| Machine Learning | Random Forest | Menggunakan pendekatan <i>bagging</i> yang efektif mengurangi varians pada model. Sangat stabil dalam menangani data dengan <i>noise</i> tinggi dan tahan terhadap <i>outlier</i> tanpa asumsi distribusi data yang ketat. |
| | XGBoost | Menggunakan pendekatan <i>gradient boosting</i> yang secara iteratif memperbaiki kesalahan prediksi model sebelumnya. Memiliki performa tinggi pada data tabular dan mampu menangkap pola non-linear yang kompleks secara efisien. |
| Deep Learning | LSTM | Arsitektur ini mengatasi masalah <i>vanishing gradient</i> pada RNN standar. Sangat krusial untuk menangkap memori jangka panjang, mengingat dampak kebijakan moneter sering kali memiliki efek tunda (<i>lag effect</i>). |
| | Bi-LSTM | Pengembangan dari LSTM yang memproses urutan data dari dua arah (maju dan mundur). Hal ini memungkinkan model menangkap konteks pola temporal yang lebih baik dibandingkan LSTM standar. |

Berdasarkan analisis pada Tabel III.3, seluruh kategori model memiliki keunggulan masing-masing. Oleh karena itu, strategi yang diambil adalah melakukan komparasi performa antara model ekonometrika (ARIMA, VAR), ML (Random Forest, XGBoost), dan DL (LSTM, Bi-LSTM). Pendekatan ini memastikan bahwa eksperimen nanti dapat mengidentifikasi model terbaik yang paling sesuai dengan karakteristik data makroekonomi yang kompleks.

III.3.2 Seleksi Metode XAI

Komponen ini bertujuan untuk menginterpretasikan model *black box* agar transparan dan dapat diketahui faktor apa saja yang memengaruhi hasil prediksinya. Tantangan terbesar dalam data ekonomi adalah multikolinieritas yang ada antar variabel. Metode XAI yang dipilih harus dapat menangani kondisi ini.

Tabel III.4 Matriks Seleksi Metode XAI

| Metode XAI | Pendekatan | Analisis |
|---|------------------------|---|
| SHAP (<i>Shapley Additive Explanations</i>) | <i>Game Theory</i> | Satu-satunya metode yang menjamin konsistensi matematis melalui tiga sifat desirable: <i>local accuracy</i> , <i>missingness</i> , dan <i>consistency</i> . Sangat efektif menangani multikolinieritas fitur dengan memberikan kontribusi yang adil untuk setiap kombinasi subset fitur, aspek krusial untuk data makroekonomi dengan korelasi antar-variabel tinggi. |
| LIME (<i>Local Interpretable Model-agnostic Explanations</i>) | <i>Local Surrogate</i> | Menggunakan metode linear lokal untuk menjelaskan prediksi spesifik per-instans. Lebih ringan secara komputasi dan efektif untuk analisis mendalam pada periode waktu tertentu, serta dapat melengkapi penjelasan global dari SHAP. |
| PFI (<i>Permutation Feature Importance</i>) | <i>Error-based</i> | Bekerja dengan asumsi fitur-fitur independen dan mengukur pentingnya fitur berdasarkan peningkatan kesalahan prediksi. Kurang sesuai untuk data makroekonomi karena permutasi menghasilkan data instances yang tidak realistik pada konteks multikolinieritas, meningkatkan bias interpretasi. |

Berdasarkan analisis pada Tabel III.4, metode XAI yang digunakan pada tugas akhir ini adalah SHAP dan LIME untuk meminimalisir bias interpretasi pada data yang mengalami multikolinearitas. SHAP memberikan penjelasan yang konsisten dan matematis dengan latar belakang *game theory* dengan tiga sifatnya (*local accuracy*, *missingness*, dan *consistency*). LIME memberikan penjelasan lokal yang mudah dipahami melalui aproksimasi linear lokal. Keduanya lebih *robust* dibanding PFI dalam menangani fitur-fitur yang berkorelasi tinggi karena SHAP memperhitungkan semua kombinasi subset fitur secara eksplisit dan LIME bekerja pada representasi lokal yang dapat disesuaikan dengan struktur data.

III.3.3 Seleksi Model LLM

Komponen terakhir ini (LLML) akan membuat narasi kebijakan yang mudah dipahami awam berdasarkan hasil prediksi dan XAI. Kriteria utamanya adalah kemampuan penalaran (*reasoning*) model dan minimnya halusinasi.

Tabel III.5 Matriks Seleksi Model Generatif (LLM)

| Model | Tipe | Analisis |
|--------|---------------------------------------|---|
| GPT-4o | <i>Proprietary</i> / <i>Cloud API</i> | Model flagship OpenAI dengan skor <i>reasoning benchmark</i> yang tinggi untuk tugas penalaran kompleks. Kemampuan menyusun argumen kausalitas yang baik dengan 128K context window sangat dibutuhkan untuk meminimalkan risiko kesalahan narasi kebijakan moneter. Inferensi cepat dengan biaya kompetitif juga dapat mendukung eksperimen iteratif. |
| Llama | <i>Open Source</i> / <i>Local</i> | Fleksibilitas pengembangan model di lokal dan transparansi penuh atas model, mencegah <i>dependency</i> pada API eksternal. Namun, kemampuan reasoning domain-spesifik masih di bawah model <i>proprietary</i> skala besar, sehingga meningkatkan risiko halusinasi dan kesalahan kontekstual dalam narasi kebijakan. |

Berdasarkan analisis pada Tabel III.5, model LLM yang digunakan pada tugas akhir ini adalah OpenAI untuk memprioritaskan akurasi penjelasan di atas efisiensi biaya, mengingat kesalahan interpretasi dalam konteks keuangan dapat berdampak fatal terhadap kredibilitas sistem.

III.3.4 Penentuan Solusi

Berdasarkan analisis seleksi komponen yang telah dilakukan pada sub-bab sebelumnya, desain solusi akhir sistem prediksi BI-Rate dalam Tugas Akhir ini ditetapkan menggunakan pendekatan arsitektur *explainable artificial intelligence with generative narrative*.

Keputusan ini diambil dengan pertimbangan utama untuk menyeimbangkan akurasi prediksi dengan kebutuhan interpretasi yang menjadi masalah yang ingin diselesaikan. Secara ringkas, arsitektur solusi terpilih terdiri dari:

1. Modul Prediksi

Tidak mengandalkan satu model tunggal, melainkan menggunakan strategi komparasi performa antara model tradisional, ML, dan DL. Pendekatan ini dapat memastikan bahwa eksperimen nanti dapat mengidentifikasi model terbaik yang paling sesuai dengan karakteristik data makroekonomi yang kompleks.

2. Modul Interpretasi

Mengintegrasikan metode SHAP sebagai penjelas utama untuk menjamin konsistensi atribusi fitur secara global, lalu didukung oleh LIME untuk validasi penjelasan lokal pada kasus prediksi yang spesifik. Keduanya dipilih untuk memitigasi risiko misinterpretasi akibat multikolinearitas data makroekonomi yang ada.

3. Modul Narasi

Menggunakan OpenAI sebagai LLM untuk mengubah *output* hasil prediksi dan XAI menjadi teks narasi kebijakan yang mudah dipahami dan menjembatani kesenjangan teknis antara sistem AI dan pengguna awam.

Integrasi ketiga komponen ini diharapkan mampu menghasilkan sistem yang optimal dalam memprediksi BI-Rate, serta transparan dan komunikatif, sesuai dengan tujuan utama tugas akhir ini.

III.4 Data Understanding

Pada tahap ini, dilakukan eksplorasi dan analisis awal terhadap data makroekonomi yang akan digunakan untuk membuat model prediksi BI-Rate. Dataset yang digunakan dalam tugas akhir ini terdiri dari 11 variabel dengan frekuensi bulanan, mulai dari Juli 2005 hingga September 2025. Total data yang diobservasi adalah sebanyak 243 data point.

III.4.1 Data Overview

Dataset terdiri atas 15 kolom pada Tabel III.6.

Tabel III.6 Deskripsi Kolom Dataset

| No | Nama Kolom | Tipe | Deskripsi |
|----|------------|---------|---|
| 1 | ID | Object | Identitas unik untuk setiap observasi bulanan |
| 2 | Bulan | Integer | Periode bulan observasi (1–12) |

Berlanjut ke halaman berikutnya...

Tabel III.6 – lanjutan dari halaman sebelumnya

| No | Nama Kolom | Tipe | Deskripsi |
|----|-----------------------------|---------|--|
| 3 | Tahun | Integer | Periode tahun observasi |
| 4 | Inflation_YoY_Pct | Float | Tingkat inflasi tahunan (year-on-year) dalam persen (%) |
| 5 | Inflation_Gap | Float | Selisih antara target inflasi Bank Indonesia dengan nilai inflasi aktual dalam persen (%) |
| 6 | GDP_Growth_YoY_Pct | Float | Pertumbuhan Produk Domestik Bruto tahunan (year-on-year) dalam persen (%), mencerminkan output gap |
| 7 | USD_IDR_Monthly_Avg | Float | Nilai tukar rata-rata bulanan Rupiah terhadap Dolar Amerika Serikat (IDR/USD) |
| 8 | M2_Triliun_Rp | Float | Jumlah uang beredar (agregat M2) dalam triliun Rupiah |
| 9 | Credit_Growth_Triliun_Rp | Float | Total volume kredit perbankan yang beredar dalam triliun Rupiah |
| 10 | IHSG_End_of_Month | Float | Indeks Harga Saham Gabungan (IHSG) pada penutupan perdagangan akhir bulan |
| 11 | Foreign_Reserves_Miliar_USD | Float | Cadangan devisa negara dalam miliar Dolar Amerika Serikat (USD) |
| 12 | Federal_Funds_Rate_Pct | Float | Suku bunga acuan Federal Reserve (Fed) Amerika Serikat dalam persen (%) |
| 13 | Oil_Price_Brent_USD_per_Bbl | Float | Harga minyak mentah <i>Brent</i> dalam Dolar Amerika Serikat per barel (USD/bbl) |
| 14 | Gold_Price_US_D_per_Oz | Float | Harga emas dunia dalam Dolar Amerika Serikat per troy ounce (USD/oz) |
| 15 | VIX_Volatility_Index | Float | Indeks volatilitas pasar saham AS yang mengukur sentimen risiko global |
| 16 | BI_Rate_Pct | Float | Suku bunga acuan Bank Indonesia (target rate kebijakan moneter) dalam persen (%) |

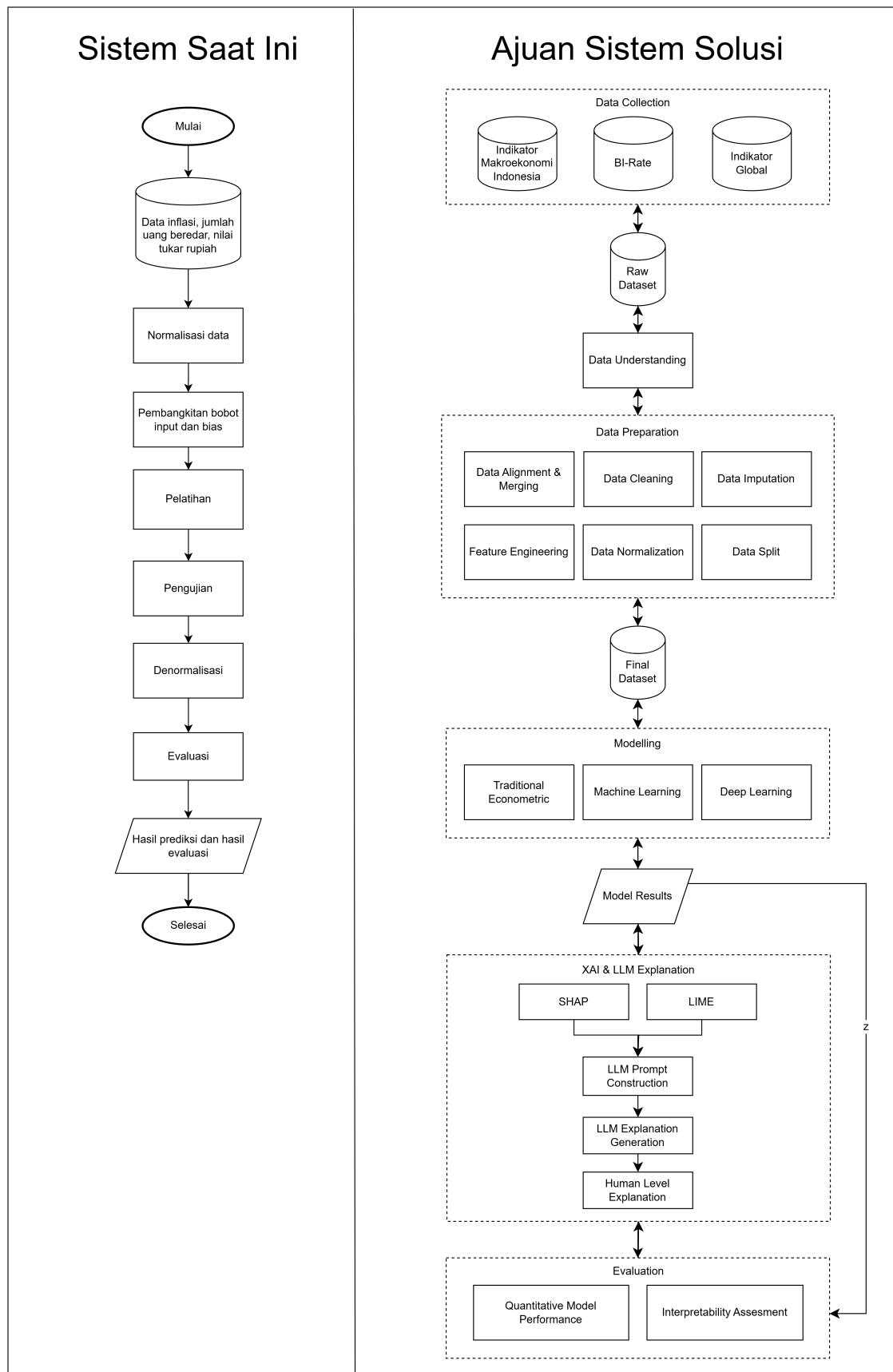
BAB IV

DESAIN KONSEP SOLUSI

Pada bab ini dijelaskan desain konseptual solusi yang diusulkan sebagai solusi atas permasalahan yang ada dalam sistem saat ini. Setiap tahapan dalam desain ini dirancang sejalan dengan metodologi CRISP-DM yang digunakan dalam tugas akhir ini.

IV.1 Desain Solusi

Tugas akhir ini akan dikerjakan sesuai dengan ajuan desain konseptual yang dapat dilihat pada Gambar IV.1. Diagram ini menggambarkan alur proses penggerjaan tugas akhir dengan lengkap, mulai dari tahap *data collection*, *data understanding*, *data preparation*, *modelling*, hingga tahap penjelasan hasil prediksi menggunakan integrasi XAI dan LLM, serta evaluasi kuantitatif dan interpretabilitas.



Gambar IV.1 Desain Solusi Pengerjaan Tugas Akhir

Perbedaan mendasar antara sistem yang ada saat ini dengan desain solusi yang diajukan dapat dilihat secara rinci pada Tabel IV.1.

Tabel IV.1 Perbandingan Sistem Saat Ini dan Ajuan Sistem Solusi

| Aspek | Sistem Saat Ini | Ajuan Sistem Solusi |
|-------------------------|--|---|
| Cakupan Data | Inflasi, jumlah uang beredar, dan nilai tukar rupiah. | Melibatkan indikator makroekonomi domestik dan global. |
| Pra-pemrosesan Data | Hanya berfokus pada normalisasi data sebelum pelatihan. | Lebih terstruktur, meliputi <i>data alignment, merging, cleaning, imputation, feature engineering</i> , hingga normalisasi dan <i>splitting</i> . |
| Pendekatan Model | <i>Single model approach.</i> | <i>Multi-model comparison</i> dengan membandingkan ekonometrika tradisional, ML, dan DL untuk memilih performa terbaik. |
| Interpretabilitas (XAI) | <i>Black-box</i> : hanya menghasilkan angka prediksi tanpa penjelasan konteks. | <i>Explainable</i> : menggunakan SHAP dan LIME untuk interpretasi fitur, serta LLM untuk menyusun narasi penjelasan yang mudah dipahami. |
| Evaluasi | Evaluasi tunggal berbasis metrik <i>error</i> kuantitatif. | Evaluasi performa kuantitatif dan asesmen interpretabilitas. |

Sistem yang diajukan dalam tugas akhir ini menjadi solusi atas kelemahan yang ada pada sistem prediksi saat ini. Sistem yang diajukan memiliki aspek transparansi yang lebih baik melalui integrasi metode XAI dan LLM untuk menjelaskan hasil prediksi, serta cakupan data yang lebih luas pada indikator makroekonomi lokal dan ditambahkan dengan indikator global.

Perbedaan utama terletak pada proses pasca pemodelan (*post-modelling*). Sistem ajuan tidak berhenti hanya pada *output* prediksi, tetapi juga mengintegrasikan XAI yang kemudian dinarasikan oleh LLM untuk memudahkan pemahaman terhadap faktor yang memengaruhi hasil prediksi bagi orang awam. Mekanisme ini memungkinkan sistem untuk tidak hanya memberikan hasil prediksi suku bunga ke depan, tetapi juga menjelaskan alasan dibalik hasil prediksi tersebut berdasarkan dinamika variabel ekonomi yang memengaruhinya. Selain itu, proses pra-pemrosesan data pada solusi yang diusulkan dirancang jauh lebih terstruktur dengan tahapan *cleaning, imputation, dan feature engineering* untuk menangani *raw dataset* yang kompleks. Hal ini bertujuan untuk memastikan kualitas data yang lebih baik sebelum masuk ke tahap pemodelan.

BAB V

RENCANA SELANJUTNYA

Pada bab ini dijelaskan terkait rencana pengerjaan tugas akhir, termasuk perkakas dan infrastruktur yang diperlukan, jadwal kerja, alokasi anggaran, desain pengujian dan evaluasi sistem, serta identifikasi dan mitigasi risiko yang berpotensi terjadi selama proses pengerjaan tugas akhir ini. Perencanaan ini dilakukan untuk memastikan kelancaran eksekusi tugas akhir dalam mencapai tujuan yang telah ditetapkan.

V.1 Rencana Implementasi

Pada subbab ini dijelaskan rencana teknis pelaksanaan penelitian, mencakup perkakas yang digunakan, serta linimasa pengerjaan tugas akhir.

V.1.1 Perkakas yang Dibutuhkan

Rincian perkakas yang dibutuhkan untuk mengerjakan tugas akhir ini dapat dilihat pada Tabel V.1.

Tabel V.1 Perkakas yang Dibutuhkan

| No | Perkakas | Deskripsi |
|----|-------------------|---|
| 1 | <i>Python</i> | Bahasa pemrograman utama untuk pengembangan model, analisis data, dan integrasi sistem. |
| 2 | <i>Pandas</i> | <i>Library</i> untuk manipulasi data tabular dan persiapan dataset. |
| 3 | <i>NumPy</i> | <i>Library</i> untuk operasi numerik. |
| 4 | <i>Matplotlib</i> | <i>Library</i> dasar untuk pembuatan visualisasi grafik. |

Berlanjut ke halaman berikutnya...

Tabel V.1 – lanjutan dari halaman sebelumnya

| No | Perkakas | Deskripsi |
|----|------------------------------|--|
| 5 | <i>Seaborn</i> | <i>Library</i> visualisasi data statistik tingkat tinggi yang dibangun di atas <i>Matplotlib</i> . |
| 6 | <i>Statsmodels</i> | <i>Library</i> untuk analisis statistik dan pemodelan ekonometrika (VAR) serta <i>time-series</i> klasik (ARIMA). |
| 7 | <i>Scikit-learn</i> | <i>Library</i> untuk <i>preprocessing</i> , metrik evaluasi, dan implementasi model <i>machine learning</i> (Random Forest). |
| 8 | <i>XGBoost</i> | <i>Library</i> khusus untuk implementasi algoritma XGB. |
| 9 | <i>TensorFlow</i> | <i>Framework</i> untuk pengembangan model <i>deep learning</i> sekuensial (LSTM dan Bi-LSTM). |
| 10 | <i>SHAP</i> | <i>Library</i> untuk interpretasi model secara global dan lokal berbasis <i>Game Theory</i> . |
| 11 | <i>LIME</i> | <i>Library</i> untuk interpretasi model secara lokal menggunakan pendekatan perturbasi. |
| 12 | <i>OpenAI API</i> | Layanan API LLM untuk mengubah hasil interpretasi numerik (XAI) menjadi narasi teks otomatis. |
| 13 | <i>Git</i> dan <i>GitHub</i> | Sistem kontrol versi untuk manajemen kode sumber dan kolaborasi. |

V.1.2 Linimasa Pengerjaan Tugas Akhir

Linimasa pengerjaan tugas akhir secara umum dapat dilihat pada Tabel V.2.

Tabel V.2 Linimasa Pengerjaan Tugas Akhir

| Bulan | Kegiatan |
|------------------------|--|
| September–Oktober 2025 | Identifikasi permasalahan, eksplorasi literatur pendukung, serta menentukan pendekatan dan metode Tugas Akhir yang akan digunakan. |
| November–Desember 2025 | Pengumpulan dan persiapan dataset, perangkat, dan teknologi yang diperlukan, serta penyusunan proposal tugas akhir. |

Berlanjut ke halaman berikutnya...

Tabel V.2 – lanjutan dari halaman sebelumnya

| Bulan | Kegiatan |
|-----------------------|---|
| Januari–Februari 2026 | Pengembangan sistem utama dan integrasi teknologi sesuai metode yang telah dirancang. |
| Maret–April 2026 | Pengujian sistem dan analisis hasil eksperimen. |
| Mei 2026 | Penyempurnaan sistem dan laporan berdasarkan temuan pengujian serta arahan dari pembimbing. |
| Juni 2026 | Penyelesaian laporan tugas akhir. |

V.2 Rencana Anggaran Biaya

Rincian estimasi biaya anggaran untuk mengerjakan tugas akhir ini dapat dilihat pada Tabel V.3.

Tabel V.3 Rencana Anggaran Biaya

| No | Komponen | Deskripsi | Biaya (IDR) |
|-------|----------------|--|-------------|
| 1 | OpenAI API Key | Biaya deposit kredit API untuk penggunaan model dalam menghasilkan narasi interpretasi prediksi. | 500.000 |
| Total | | | 500.000 |

V.3 Desain Pengujian dan Evaluasi

Pengujian dan evaluasi sistem dilakukan melalui dua aspek utama, yaitu verifikasi kinerja model secara teknis dan validasi dari perspektif ekonomi.

V.3.1 Verifikasi Kinerja Model

Seluruh skenario verifikasi kinerja model dapat dilihat pada Tabel V.4.

Tabel V.4 Skenario Verifikasi Kinerja Model

| No | Skenario | Tujuan Evaluasi | Metrik |
|----|----------------------------------|---|-------------------|
| S1 | <i>Baseline Comparison</i> | Membandingkan performa semua model (VAR, ARIMA, Random Forest, XGBoost, LSTM, Bi-LSTM) pada data uji. | RMSE, MAPE, R^2 |
| S2 | <i>Out-of-Sample Performance</i> | Menguji generalisasi model pada data yang tidak terlihat selama pelatihan. | RMSE, MAPE, R^2 |

V.3.2 Validasi Ekonomi

Seluruh skenario validasi ekonomi ini akan dilakukan kepada ahli di bidang terkait (ekonomi) dan rincian skenarionya dapat dilihat pada Tabel V.5.

Tabel V.5 Skenario Validasi Interpretasi Ekonomi

| No | Aspek Validasi | Deskripsi |
|----|--------------------------------|---|
| S3 | Kesesuaian Arah Pengaruh Fitur | Validasi apakah arah pengaruh fitur (positif/negatif) pada hasil SHAP dan LIME sesuai dengan teori ekonomi moneter. |
| S4 | Relevansi Peringkat Fitur | Validasi apakah urutan <i>feature importance</i> yang dihasilkan model sesuai dengan prioritas variabel dalam pengambilan keputusan BI-Rate menurut ahli. |
| S5 | Kejelasan Narasi Interpretasi | Validasi apakah narasi yang dihasilkan LLM dapat dipahami, akurat secara ekonomi, dan berguna bagi pengambil keputusan. |

V.4 Analisis Risiko dan Mitigasi

Terdapat beberapa potensi risiko dalam tugas akhir ini yang dapat memengaruhi hasil akhir, beserta upaya mitigasi yang telah dirancang untuk mengantisipasinya. Rincian lengkap analisis risiko dan mitigasi dapat dilihat pada Tabel V.6.

Tabel V.6 Analisis Risiko dan Mitigasi

| No | Risiko | Dampak | Mitigasi |
|----|---|--|---|
| 1 | Data target prediksi terbatas | Akurasi model rendah dan <i>overfitting</i> pada model <i>deep learning</i> . | <ol style="list-style-type: none"> 1. Menggunakan <i>time-series augmentation (sliding window)</i> untuk meningkatkan sampel pelatihan. 2. Implementasi <i>regularization (dropout, L1/L2)</i> pada model Bi-LSTM. 3. Melakukan <i>cross-validation</i> untuk validasi hasil prediksi. |
| 2 | Akurasi model <i>deep learning</i> rendah | Validasi model <i>deep learning</i> mungkin tidak optimal dibanding <i>machine learning</i> atau model ekonometrika tradisional. | <ol style="list-style-type: none"> 1. Melakukan <i>tuning hyperparameter</i> untuk mencari kombinasi parameter optimal yang menghasilkan kinerja model terbaik. 2. Melakukan proses <i>feature engineering</i> yang lebih luas lagi. |

Penerapan langkah-langkah di atas diharapkan dapat meminimalisir berbagai risiko yang terjadi, sehingga proses Tugas Akhir dapat berlangsung secara optimal sesuai dengan perencanaan.

DAFTAR PUSTAKA

- Adadi, Amina, dan Mohammed Berrada. 2018. “Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)”. *IEEE Access* 6:52138–52160. <https://doi.org/10.1109/ACCESS.2018.2870052>.
- AI, Meta. February 2023. *Introducing LLaMA: A foundational, 65-billion-parameter language model*. <https://ai.meta.com/blog/large-language-model-llama-meta-ai/>.
- Bank Indonesia. 2023. “Kebijakan Moneter”. Diakses pada 8 Oktober 2025. <https://www.bi.go.id/id/fungsi-utama/moneter/default.aspx>.
- Baur, Dirk G., dan Thomas K. McDermott. 2010. “Is Gold a Safe Haven? International Evidence”. *Journal of Banking & Finance* 34 (8): 1886–1898. <https://doi.org/10.1016/j.jbankfin.2009.12.008>.
- Bekaert, Geert, Eric C. Engstrom, dan Nancy R. Xu. 2019. *The World Fear Index*. Working Paper 26211. National Bureau of Economic Research. <https://doi.org/10.3386/w26211>.
- Box, George E. P., dan Gwilym M. Jenkins. 1970. *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day.
- Box, George E. P., Gwilym M. Jenkins, Gregory C. Reinsel, dan Greta M. Ljung. 2015. *Time Series Analysis: Forecasting and Control*. 5th edisi. Hoboken: John Wiley & Sons. ISBN: 978-1118675021. <https://doi.org/10.1002/9781118619193>.
- Brown, Tom, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D. Kaplan, Prafulla Dhariwal, Arvind Neelakantan, dkk. 2020. “Language Models are Few-Shot Learners”. Dalam *Advances in Neural Information Processing Systems*, 33:1877–1901. <https://proceedings.neurips.cc/paper/2020/hash/1457c0d6bfcb4967418bfb8ac142f64a-Abstract.html>.

- Bussmann, N., P. Giudici, D. Marinelli, dan J. Papenbrock. 2021. “Explainable Machine Learning in Credit Risk Management”. *Computational Economics* 57:203–216. <https://doi.org/10.1007/s10614-020-10042-0>.
- Calik, S. S., A. Akyuz, Z. H. Kilimci, dan K. Colak. 2025. “Explainable-AI powered stock price prediction using time series transformers: A Case Study on BIST100”. *arXiv preprint arXiv:2506.06345*, <https://doi.org/10.48550/arXiv.2506.06345>.
- Chakraborty, Chiranjit, dan Andreas Joseph. 2017. *Machine Learning at Central Banks*. Staff Working Paper 674. London: Bank of England. <https://doi.org/10.2139/ssrn.3031796>.
- Chen, Tianqi, dan Carlos Guestrin. 2016. “XGBoost: A Scalable Tree Boosting System”. Dalam *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. New York: ACM. <https://doi.org/10.1145/2939672.2939785>.
- Engle, Robert F., dan C. W. J. Granger. 1987. “Co-integration and Error Correction: Representation, Estimation, and Testing”. *Econometrica* 55 (2): 251–276. <https://doi.org/10.2307/1913236>.
- Hamilton, James D. 2009. “Causes and Consequences of the Oil Shock of 2007–08”. *Brookings Papers on Economic Activity* 2009 (1): 215–261. <https://doi.org/10.1353/eca.0.0047>.
- Hansen, Bruce E. 2011. “Threshold Autoregression in Economics”. *Statistics and Its Interface* 4 (2): 123–127. <https://doi.org/10.4310/SII.2011.v4.n2.a4>.
- Hastie, Trevor, Robert Tibshirani, dan Jerome Friedman. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd edisi. New York: Springer. <https://doi.org/10.1007/978-0-387-84858-7>. <https://hastie.su.domains/ElemStatLearn/>.
- Hochreiter, Sepp, dan Jürgen Schmidhuber. 1997. “Long Short-Term Memory”. *Neural Computation* 9 (8): 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>.
- Hodson, T. O. 2022. “Root Mean Square Error (RMSE) or Mean Absolute Error (MAE): When to Use Them or Not”. *Geoscientific Model Development* 15 (14): 5481–5487. <https://doi.org/10.5194/gmd-15-5481-2022>.

- Hui, Aaron, Shawn Ahn, Carolyn Lye, dan Jun Deng. January 2022. “Ethical Challenges of Artificial Intelligence in Health Care: A Narrative Review”. *Ethics in Biology, Engineering and Medicine: An International Journal* 12 (). <https://doi.org/10.1615/EthicsBiologyEngMed.2022041580>.
- Hyndman, Rob J., dan Anne B. Koehler. 2006. “Another Look at Measures of Forecast Accuracy”. *International Journal of Forecasting* 22 (4): 679–688. <https://doi.org/10.1016/j.ijforecast.2006.03.001>.
- Johansen, Søren. 1995. *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford: Oxford University Press. ISBN: 978-0198774501. <https://doi.org/10.1093/0198774508.001.0001>.
- Lundberg, Scott M., dan Su-In Lee. 2017. “A Unified Approach to Interpreting Model Predictions”. Dalam *Advances in Neural Information Processing Systems*, 30:4765–4774. <https://doi.org/10.48550/arXiv.1705.07874>.
- Lütkepohl, Helmut. 2011. “Vector Autoregressive Models”. Dalam *International Encyclopedia of Statistical Science*, disunting oleh Miodrag Lovric, 1645–1647. Berlin: Springer. https://doi.org/10.1007/978-3-642-04898-2_605.
- Martínez-Plumed, Fernando, Lidia Contreras-Ochando, Cèsar Ferri, José Hernández-Orallo, Meelis Kull, Nicolas Lachiche, María José Ramírez-Quintana, dan Peter A. Flach. 2021. “CRISP-DM Twenty Years Later: From Data Mining Processes to Data Science Trajectories”. *IEEE Transactions on Knowledge and Data Engineering* 33 (8): 3048–3061. <https://doi.org/10.1109/TKDE.2019.2962680>.
- Mishkin, Frederic S. 2019. *The Economics of Money, Banking, and Financial Markets*. 12th edisi. New York: Pearson. ISBN: 978-0134733821.
- Moscatelli, M., F. Parlapiano, S. Narizzano, dan G. Viggiano. 2020. “Corporate Default Forecasting with Machine Learning”. *Expert Systems with Applications* 161:113567. <https://doi.org/10.1016/j.eswa.2020.113567>.
- Nagelkerke, N. J. D. 1991. “A note on a general definition of the coefficient of determination”. *Biometrika* 78 (3): 691–692. <https://doi.org/10.1093/biomet/78.3.691>.

- Obstfeld, Maurice, Jonathan D. Ostry, dan Mahvash S. Qureshi. 2019. “A Tie That Binds: Revisiting the Trilemma in Emerging Market Economies”. *Review of Economics and Statistics* 101 (2): 279–293. https://doi.org/10.1162/rest_a_00740.
- Obstfeld, Maurice, Jay C. Shambaugh, dan Alan M. Taylor. 2010. “Financial Stability, the Trilemma, and International Reserves”. *American Economic Journal: Macroeconomics* 2 (2): 57–94. <https://doi.org/10.1257/mac.2.2.57>.
- OpenAI. 2024. “GPT-4o System Card”. <https://openai.com/index/gpt-4o-system-card/>.
- Putri, Yohana Yunita, Putra Pandu Adikara, dan Sigit Adinugroho. May 2019. “Prediksi Suku Bunga Acuan (BI 7-Day Repo Rate) Menggunakan Metode Extreme Learning Machine (ELM)”. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer* 3, no. 5 (): 4251–4258. ISSN: 2548-964X. <http://j-ptiik.ub.ac.id>.
- Rey, Hélène. 2015. *Dilemma Not Trilemma: The Global Financial Cycle and Monetary Policy Independence*. Working Paper 21162. National Bureau of Economic Research. <https://doi.org/10.3386/w21162>.
- Ribeiro, Marco Tulio, Sameer Singh, dan Carlos Guestrin. 2016. “Why Should I Trust You?: Explaining the Predictions of Any Classifier”. Dalam *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135–1144. New York: ACM. <https://doi.org/10.1145/2939672.2939778>.
- Rudin, Cynthia. 2019. “Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead”. *Nature Machine Intelligence* 1 (5): 206–215. <https://doi.org/10.1038/s42256-019-0048-x>.
- Schröer, Christoph, Felix Kruse, dan Jorge Marx Gómez. 2021. “A Systematic Literature Review on Applying CRISP-DM Process Model”. *Procedia Computer Science* 181:526–534. <https://doi.org/10.1016/j.procs.2021.01.199>.
- Schuster, Mike, dan Kuldip K. Paliwal. 1997. “Bidirectional Recurrent Neural Networks”. *IEEE Transactions on Signal Processing* 45 (11): 2673–2681. <https://doi.org/10.1109/78.650093>.

- Sezer, Omer Berat, Mehmet Ugur Gudelek, dan Ahmet Murat Ozbayoglu. 2020. “Financial Time Series Forecasting with Deep Learning: A Systematic Literature Review: 2005–2019”. *Applied Soft Computing* 90:106181. <https://doi.org/10.1016/j.asoc.2020.106181>.
- Sims, Christopher A. 1980. “Macroeconomics and Reality”. *Econometrica* 48 (1): 1–48. <https://doi.org/10.2307/1912017>.
- Stock, James H., dan Mark W. Watson. 2012. “Disentangling the Channels of the 2007–09 Recession”. *Brookings Papers on Economic Activity* 2012 (1): 81–135. <https://doi.org/10.1353/eca.2012.0005>.
- Svensson, Lars E. O. 2003. “What Is Wrong with Taylor Rules? Using Judgment in Monetary Policy through Targeting Rules”. *Journal of Economic Literature* 41 (2): 426–477. <https://doi.org/10.1257/002205103765762734>.
- . 1997. “Inflation Forecast Targeting: Implementing and Monitoring Inflation Targets”. *European Economic Review* 41 (6): 1111–1146. [https://doi.org/10.1016/S0014-2921\(96\)00055-4](https://doi.org/10.1016/S0014-2921(96)00055-4).
- Taylor, John B. 1993. “Discretion versus Policy Rules in Practice”. *Carnegie-Rochester Conference Series on Public Policy* 39:195–214. [https://doi.org/10.1016/0167-2231\(93\)90009-L](https://doi.org/10.1016/0167-2231(93)90009-L).
- Warjiyo, Perry, dan Solikin M. Juhro. 2019. *Central Bank Policy: Theory and Practice*. Book Excerpt accessed from Bank Indonesia Institute. Bingley: Emerald Publishing Limited. ISBN: 978-1-78973-752-3.
- Yu, Xin, Zilong Chen, dan Yuan Lu. 2023. “Temporal Data Meets LLM – Explainable Financial Time Series Forecasting”. *arXiv preprint arXiv:2306.11025*, <https://doi.org/10.48550/arXiv.2306.11025>.
- Zhao, Wayne Xin, Kun Zhou, Junyi Li, Tianyi Tang, Xiaolei Wang, Yupeng Hou, Yingqian Min, dkk. 2023. “A Survey of Large Language Models”. *arXiv preprint arXiv:2303.18223*, <https://doi.org/10.48550/arXiv.2303.18223>.