**LAPORAN PROJECT MACHINE LEARNING**

A logo for a university

Description automatically generated

**EXPLORING XAU/USD TRENDS: LEVERAGING PREDICTION ANALYSIS USING MACHINE LEARNING APPROACH**

Oleh

Louis 2602147276 Computer Science

Clarice Arlin Wijaya 2602135143 Computer Science

Janssen Mitchellano Hamaziah 2602117525 Computer Science

**UNIVERSITAS BINA NUSANTARA JAKARTA**

**2024**

1. **LATAR BELAKANG**

Dalam medan yang dinamis dari pasar keuangan global, perhatian utama tertuju pada pergerakan nilai tukar mata uang. XAU/USD, pasangan yang membandingkan harga emas terhadap dolar Amerika Serikat, memiliki relevansi signifikan. Emas sebagai penanda stabilitas dan dolar AS sebagai mata uang global membentuk dinamika unik pada pasangan ini.

Dalam proyek ini, kehadiran teknologi kecerdasan buatan (AI) memainkan peran penting. Machine Learning akan diterapkan untuk mengurai dan memahami pola-pola kompleks dalam pergerakan nilai tukar XAU/USD. AI tidak hanya menjadi pendekatan analisis yang kuat, tetapi juga menjadi mitra strategis dalam merespons dinamika pasar yang cepat dan seringkali tidak terduga.

1. Pentingnya Pasangan XAU/USD dalam Konteks Ekonomi Global

Pasangan mata uang XAU/USD memainkan peran utama dalam keputusan bisnis dan investasi global, khususnya saat nilai emas terpapar ketidakpastian ekonomi.

1. Dampak pada Keputusan Bisnis dan Investasi

Prediksi pergerakan nilai tukar XAU/USD memberikan keunggulan dalam pengambilan keputusan, memungkinkan respons yang tepat terhadap perubahan dinamika pasar. Kombinasi antara keahlian manusia dan kecerdasan buatan menghadirkan solusi yang lebih cerdas.

1. Kompleksitas Pasar Forex

Pasar forex, terutama XAU/USD, dipengaruhi oleh faktor-faktor kompleks. Penerapan Machine Learning membuka pintu untuk pemahaman yang lebih mendalam terhadap kompleksitas ini.

1. Fleksibilitas Investasi

Informasi yang akurat tentang pergerakan XAU/USD memungkinkan para investor untuk mengoptimalkan portofolio, mengelola risiko, dan merespons perubahan pasar dengan lebih cepat.

1. **TUJUAN**

Proyek ini bertujuan untuk mengembangkan model machine learning yang dapat memprediksi prediksi pergerakan nilai tukar USD terhadap emas. Dengan prediksi yang akurat, proyek ini diharapkan dapat memberikan keunggulan dalam pengambilan keputusan bisnis dan investasi, memungkinkan para pelaku pasar merespons dengan lebih tepat terhadap perubahan pasar yang cepat dan membuka peluang untuk mengoptimalkan portofolio juga mengelola risiko.

1. **METODE**
   1. **Data Acquisition**

Dataset awal mencakup informasi penting seperti nilai pembukaan, penutupan, tertinggi, terendah, volume perdagangan, dan tanggal transaksi terkait dengan pasangan mata uang XAU/USD. Data ini memungkinkan pengembangan model Machine Learning untuk memprediksi potensi perubahan nilai tukar USD dan tren pasar dengan memanfaatkan pola-pola kompleks dalam perubahan harga.

A screenshot of a data

Description automatically generated

Dalam proyek ini, kami menggunakan dua jenis dataset:

* + 1. Dataset Statik

<https://www.investing.com/currencies/xau-usd-historical-data>

Dataset ini menyediakan informasi tentang harga historis dari pasangan mata uang XAU/USD (harga emas terhadap dolar AS) dalam rentang waktu 2010-2024. Data ini akan digunakan dalam tahap analisis historis dan pengembangan model prediksi.

* + 1. Dataset Realtime

<https://finazon.io/featured-datasets/us-stocks-market-data>

Dataset ini memberikan akses ke data pasar saham Amerika Serikat secara real-time, yang berguna untuk memperbarui model dan mengambil keputusan yang lebih tepat waktu berdasarkan informasi terbaru.

* 1. **Exploratory Data Analysis (EDA)**

Exploratory Data Analysis (EDA) merujuk pada proses penting untuk memahami karakteristik dan struktur dari dataset. Pada tahap ini, dilakukan penyelidikan awal terhadap data untuk menemukan pola, tren, dan anomali dengan bantuan summary statistics dan visualisasi grafis.

Informasi Dataset:

1. Rentang Waktu Data: 1 Januari 2010 hingga 1 Januari 2024.
2. Jumlah Baris: 1424
3. Jumlah Kolom: 7
4. Kolom dan Tipe Data:

* Date: Tanggal (datetime)
* Open: Harga pembukaan (float)
* High: Harga tertinggi (float)
* Low: Harga terendah (float)
* Close: Harga penutupan (float)
* Adj Close: Harga penutupan yang disesuaikan (float)
* Volume: Volume perdagangan (int)

Berikut adalah langkah-langkah EDA yang diterapkan:

* + 1. Analisis Statistik Deskriptif

Kami melakukan analisis statistik deskriptif untuk setiap atribut dalam dataset. Hal ini melibatkan peninjauan nilai-nilai utama seperti mean, median, dan deviasi standar. Ini membantu untuk mendapatkan gambaran awal tentang sebaran dan karakteristik dasar dari data.

* + 1. Visualisasi Distribusi Data

Berbagai jenis plot, seperti histogram atau box plot digunakan untuk mengeksplorasi distribusi data lebih lanjut. Visualisasi ini membantu dalam mengidentifikasi pola dan melihat sebaran nilai dalam dataset dengan lebih jelas.

* + 1. Identifikasi Outliers

Keberadaann outliers dalam data diidentifikasi menggunakan metode box plot. Outlier adalah nilai yang jauh berbeda dari sebagian besar data dan dapat memiliki pengaruh yang signifikan terhadap hasil analisis.

* + 1. Eksplorasi Korelasi

Dengan menganalisis korelasi antara atribut-atribut dalam dataset menggunakan heatmap, kami dapat memahami hubungan antar variabel. Ini membantu dalam menentukan atribut-atribut yang saling berkaitan dan relevan untuk analisis lebih lanjut.

* 1. **Pre-processing Data**

Sebelum data dapat digunakan untuk melatih model, pre-processing data pun dilakukan untuk membersihkan, menormalkan, dan mengatur data agar sesuai dengan kebutuhan analisis dan proses pemodelan lanjutan, termasuk dalam tahap pemilihan fitur, pelatihan model, dan evaluasi performa.

Berikut adalah langkah-langkah yang kami terapkan dalam pre-processing data untuk meningkatkan kemampuan model kami:

* + 1. Handling Null Values

Nilai yang hilang atau null dalam dataset dihapus untuk memastikan kebersihan data sebelum membangun model.

* + 1. Feature Scaling

Skala nilai setiap atribut disesuaikan agar tidak ada atribut yang mendominasi karena rentang nilai yang besar. Salah satu teknik yang umum digunakan adalah Min-Max scaling, di mana nilai setiap fitur dinormalisasi ke rentang antara 0 dan 1, menjaga konsistensi dalam pemodelan.

* + 1. Data Splitting

Pembagian data dilakukan dengan menghitung jumlah baris yang akan digunakan untuk data latih dan data uji, kemudian membagi dataset berdasarkan proporsi tersebut. Langkah ini memungkinkan kita untuk menguji kinerja model pada data yang belum pernah digunakan sebelumnya.

Link Colab: <https://colab.research.google.com/drive/1F7rlGtcbenNJtx0DQQnNHbLCZO1GLkuJ?usp=sharing>