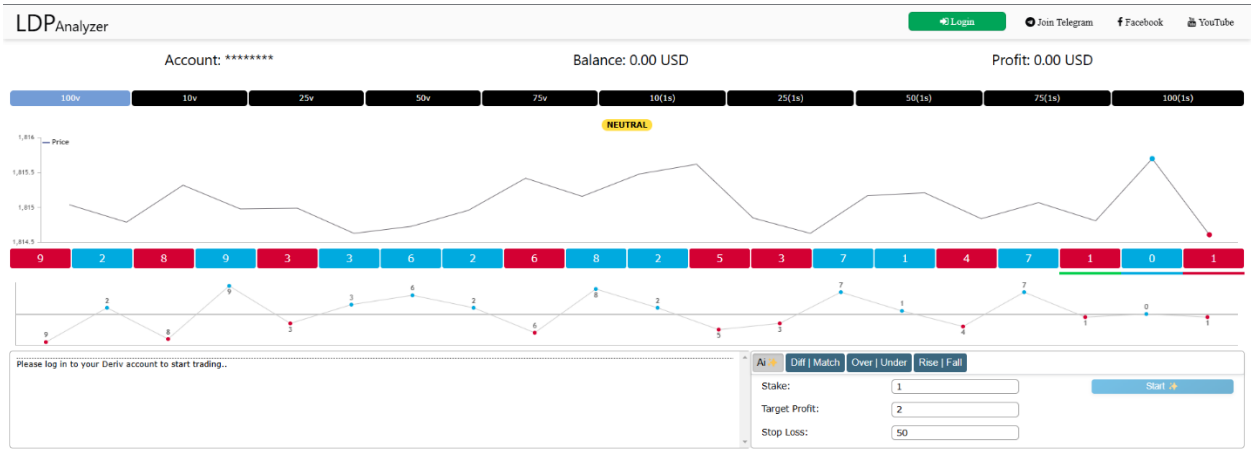
Machine Learning Engineer Task: Predicting the Next Number in the Byner Analyzer Sequence

Objective:

Develop a machine learning pipeline to predict the next number in the "Byner Analyzer" sequence (as shown in the image). The pipeline should also include saving the input data and predictions into a MongoDB database.



Website link: https://binarybot.live/ldp/

Requirements:

1. **Input Features**:
    - Historical sequence of numbers (e.g., numbers in blue and red tiles).
    - Corresponding trends or graph values over time.
    - Additional features, if applicable (e.g., time intervals, market indicators).
2. **Output**:
    - Predict the next number in the sequence.
        - Ex
            1. When 7 comes, 7 comes back after 12 digits. Or,

2. After 12 second, seven appears again. (show timer)

3. **Steps**:
    - **Data Extraction**: Create a method to extract the historical sequence and trends from a source (e.g., provided data files or simulated data).
    - **Data Storage**: Store the historical data, model predictions, and any other relevant information in a MongoDB database.
    - **Preprocessing**: Transform data into a format suitable for training (e.g., normalization or feature engineering).
    - **Model Development**: Use a suitable machine learning model for sequence prediction (e.g., RNN, LSTM, or transformers).
    - **Prediction**: Make predictions based on the most recent historical data.
    - **Visualization**: Compare actual vs. predicted values on a graph for validation.
4. **Database Requirements**:
    - Use MongoDB for storing:
        - Historical input data.
        - Model predictions.
        - Evaluation metrics for tracking model performance over time.
    - Ensure the database schema is designed efficiently for querying past predictions and trends.
5. **Evaluation Metrics**:
    - Mean Squared Error (MSE), Mean Absolute Error (MAE), or other metrics to evaluate the prediction quality.
6. **Deliverables**:
    - A Python script or Jupyter notebook containing:
        - Data preprocessing, model training, and testing logic.
        - MongoDB integration for saving input data, predictions, and metrics.
        - Provide requirement.txt
    - A visualization comparing actual vs. predicted values.

- Documentation explaining:
  - The MongoDB schema.
  - Steps to set up the MongoDB database.
  - Model architecture and reasoning.

Once the project is finished, share it to Github and submit it to dulanjana@cypsolabs.com and cc to office.cypsolabs@gmail.com.