M K S S S' S Cummins
College of Engineering
for Women, Pune

STOCK SPHERE

Vaishnavi Machale: UCE2023534
Damini Pavare: UCE2023553
Vaishnavi Salunke: UCE2023559

Problem Statement:

The stock market is highly volatile and influenced by many unpredictable factors, making it difficult for investors to forecast future stock prices accurately. Traditional statistical methods often fail to capture complex, long-term patterns in stock price movements. There is a need for an intelligent system that can analyze historical data and provide reliable predictions of future stock trends.

project overview:

Goals:

- Predict stock prices using AI.
- Visualize trends and moving averages.
- Provide 30-day forecasts for better decisions.

Challenges:

- Handling noisy, volatile stock data.
- Ensuring prediction accuracy.
- Processing large historical datasets efficiently.

Abstract: Stock Market Prediction Using ML

Data Preprocessing :

- Collected historical stock data from Yahoo Finance or demo datasets.
- Cleaned data, handled missing values, and normalized prices.
- Computed technical indicators like moving averages (MA100, MA200).

Model Training & Evaluation :

- LSTM-based model trained on historical closing prices.
- Model evaluated using prediction accuracy and error metrics.
- Split data into training and testing sets for reliable evaluation.

Visualization & Insights :

- Interactive charts displaying stock trends and moving averages.
- Heatmaps for daily percentage changes over the last 6 months.
- 30-day price forecast with expected price change metrics.

Deployment :

- Streamlit-based web application for real-time predictions.
- User-friendly interface to explore historical data, model predictions, and forecasts.

Technology Stack

- Programming Languages & Libraries :

Python – Core language for data processing and model development.
NumPy & Pandas – Data manipulation and analysis.
Matplotlib & Seaborn – Visualization of stock trends and heatmaps.
Keras – LSTM model for time-series prediction.
scikit-learn – Data scaling, preprocessing, and evaluation metrics.

- Data Source :

Yahoo Finance API – Historical stock data acquisition.

- Web & App Development :

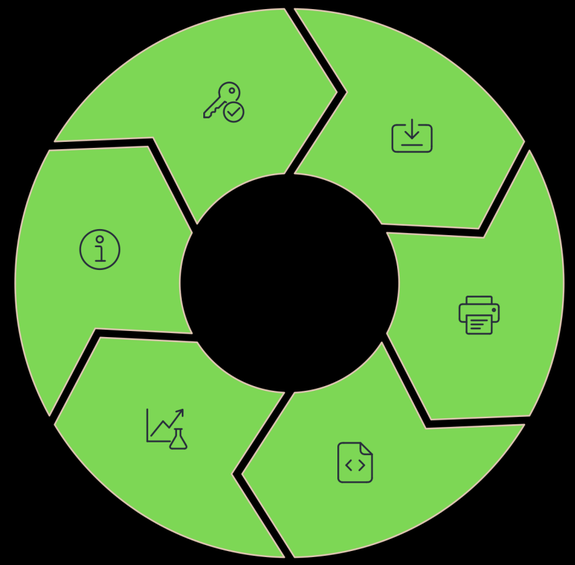
Streamlit – Frontend for user interaction and visualization.

- Tools & Environment :

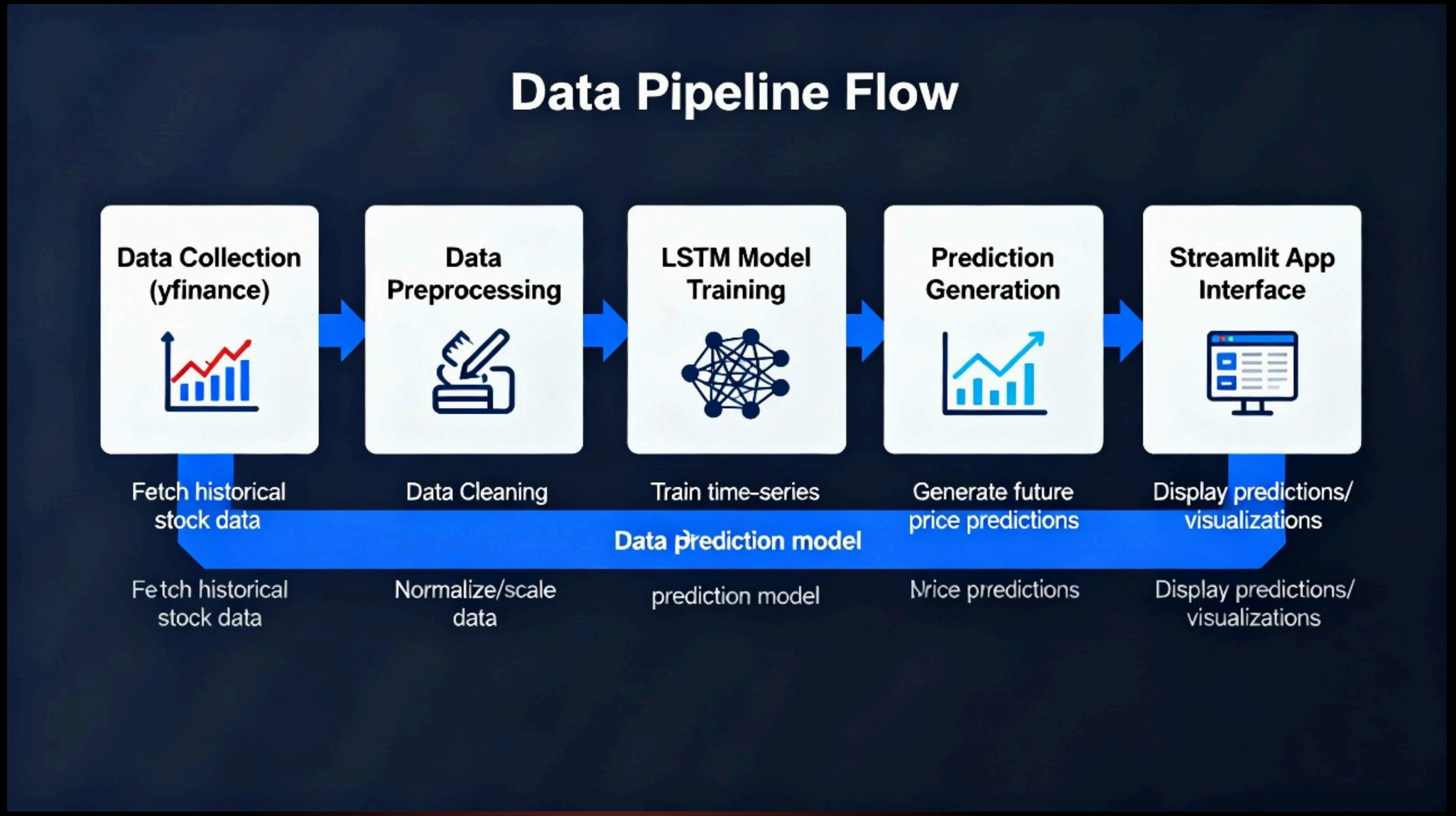
Jupyter / VS Code – Development and debugging.

- Model Storage

.keras file – Persist trained LSTM model for reuse.



System Architecture



Algorithms Used for Prediction

1. Data Collection

- Download 20 years of historical stock data using Yahoo Finance API (yfinance).
- Store Open, High, Low, Close, Volume for analysis.

2. Data Preprocessing

- Scale closing prices with MinMaxScaler.
- Create sequences of 100 past days to predict the next day.
- Split data into training (70%) and testing (30%) sets.

3. LSTM Model Training

- Build a multi-layer LSTM neural network.
- Train the model on training data to learn sequential patterns.
- Save the trained model for prediction.

4. Prediction on Test Data

- Prepare test sequences of 100 days.
- Predict closing prices and invert scaling to original values.
- Compare predicted vs actual prices using charts.

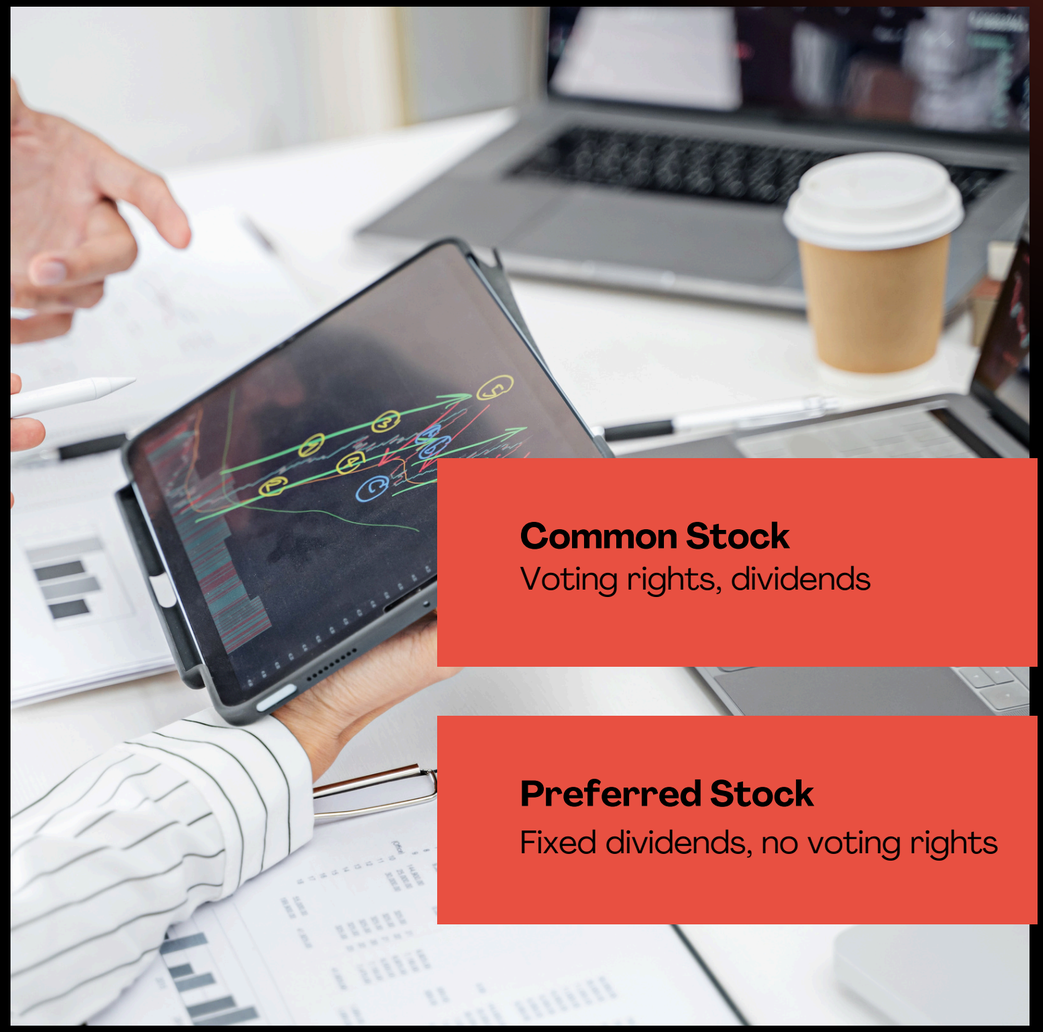
Dataset Overview

Source: Yahoo Finance via Python library yfinance
Purpose: To collect historical stock price data for analysis, visualization, and AI-driven predictions

Key Features:
Time Period: Last 15 years (from 2008–2023)

Columns:
Open – Opening price of the stock
High – Highest price of the day
Low – Lowest price of the day
Close – Closing price (used for predictions)
Volume – Number of shares traded

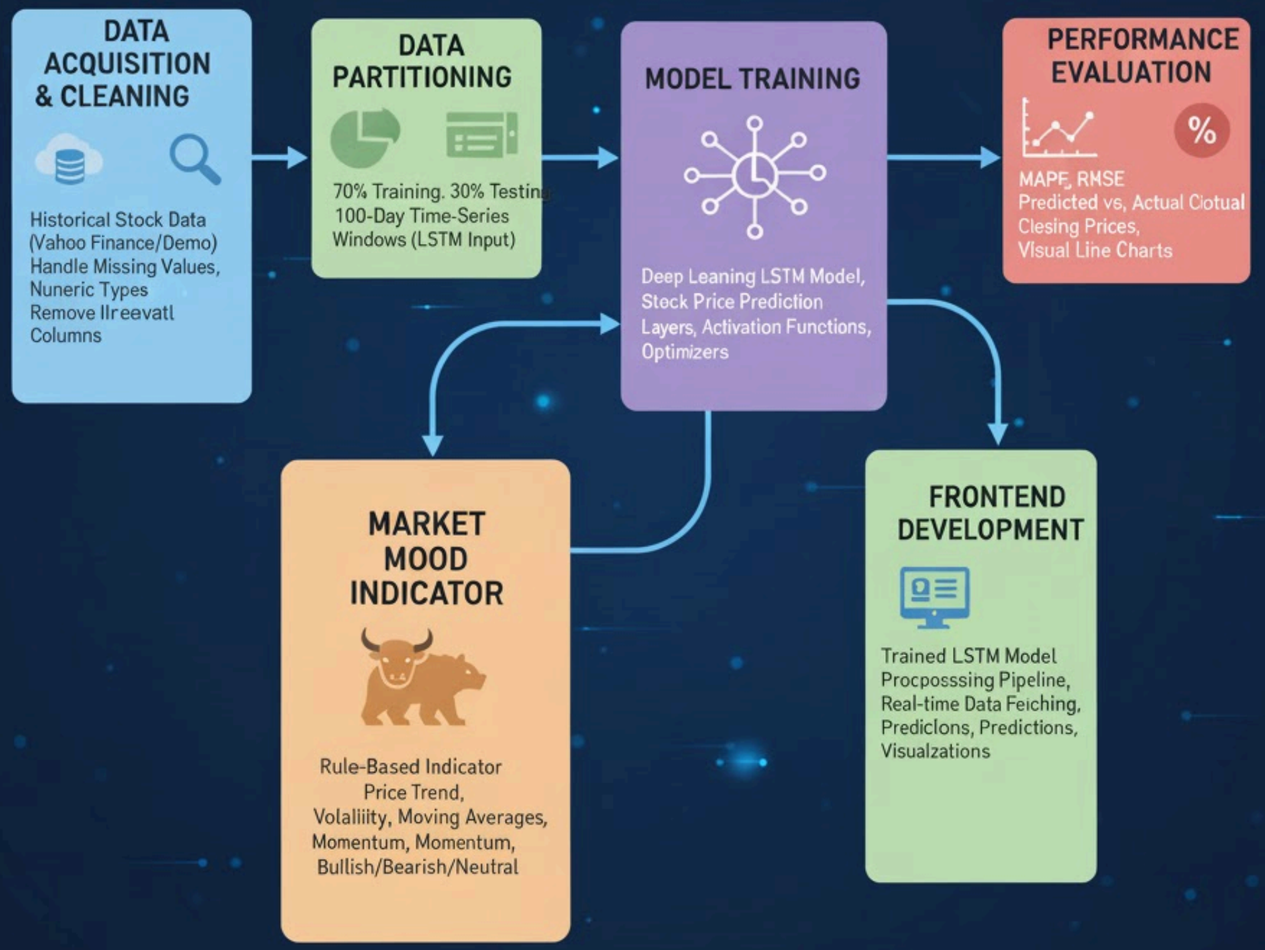
Frequency: Business days only (excludes weekends and holidays)



Common Stock
Voting rights, dividends

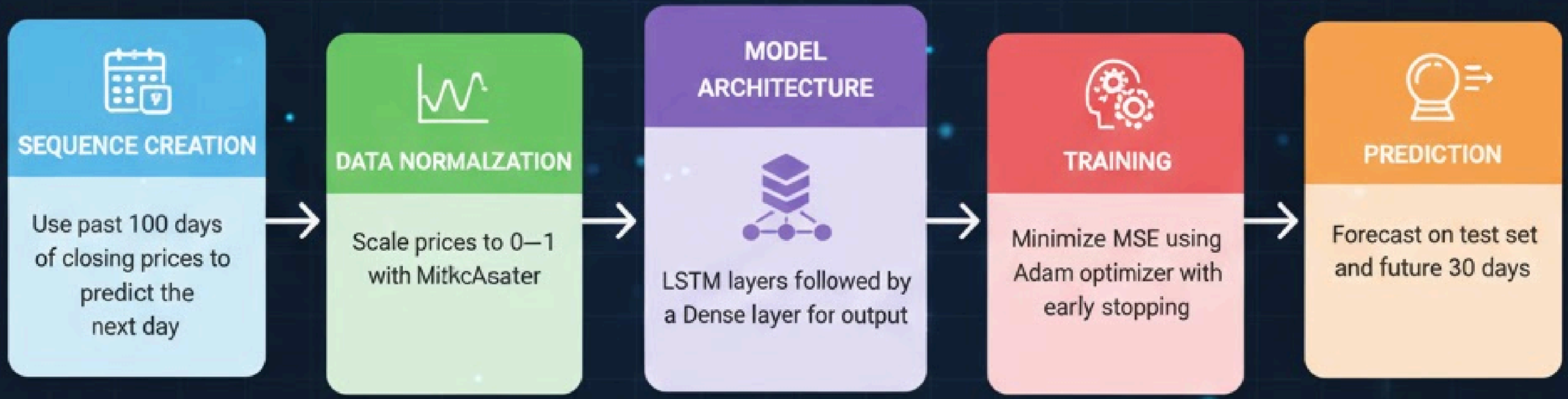
Preferred Stock
Fixed dividends, no voting rights

STOCK PREDICTION METHODOLOGY



ALGORITHM USED

Long Short-Term Memory (LSTM) Neural Network for stock price prediction.
prediction. Captures temporal depedencies.



Evaluation Metrics: Quantifying Performance

- Mean Absolute Error (MAE)

Measures the average magnitude of prediction errors:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

Lower MAE indicates better prediction accuracy.

- Mean Squared Error (MSE)

Penalizes larger errors more than MAE:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

- Root Mean Squared Error (RMSE)

Provides error in the same units as stock prices:

$$RMSE = \sqrt{MSE}$$

- Mean Absolute Percentage Error (MAPE)

Measures prediction error relative to actual values:

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


THANK
YOU.
www.reallygreatsite.com