

## Phase 1 – Problem Identification & Data Acquisition

### Problem Statement

Build an AI-based solution to predict stock prices using time series analysis, enabling short-term and long-term forecasting, trend detection, and model interpretability.

### Data Collection

#### - Sources Used:

- Yahoo Finance via `yfinance` (for live, real-time stock data)
- Kaggle datasets:
  - Nifty 50 Stock Market Data
  - Financial Analysis Stocks Dataset
  - Nifty 100 5-min Interval Dataset

### Data Overview

- Multiple tickers, daily and intraday data
- Data attributes include Open, High, Low, Close, Volume
- Combined data covered Indian and global stocks

### Challenges

- Varying formats and symbols
- Missing values and inconsistent column names
- Need to unify static and live formats