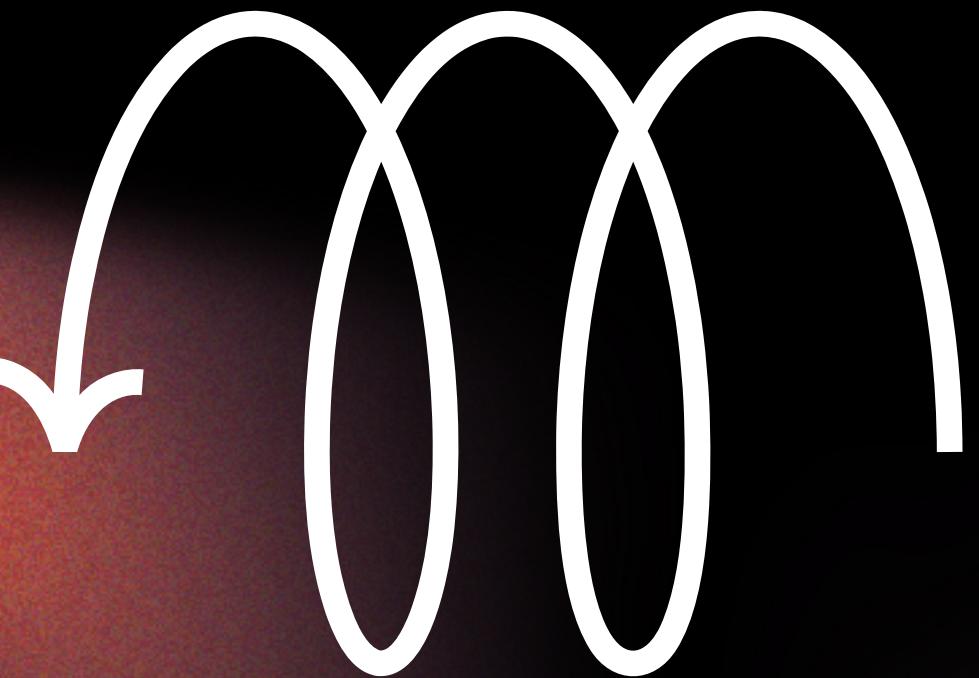


PREDICTIVE ANALYTICS AND TREND FORECASTING OF THE NIFTY 50 INDEX



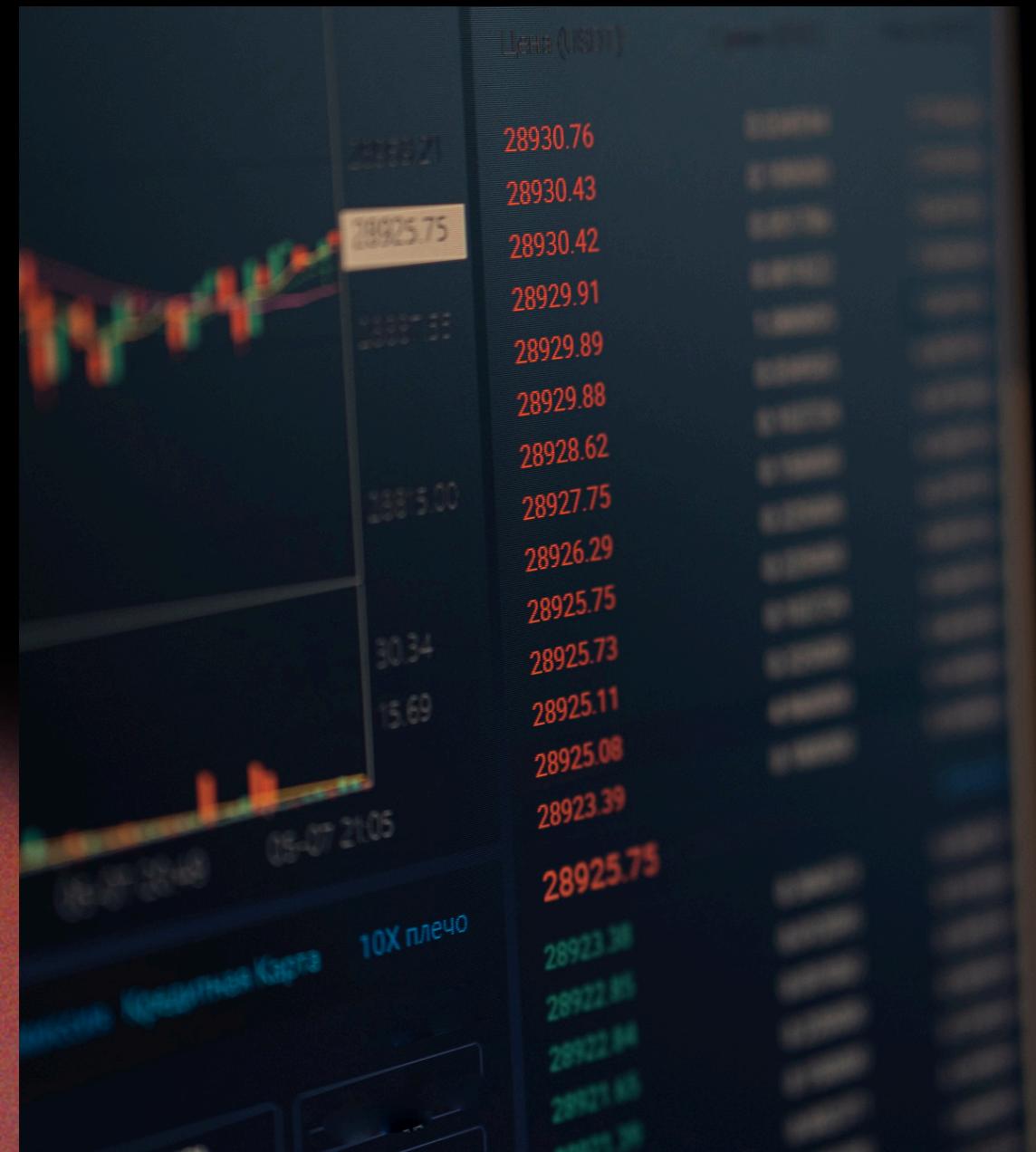
Quantitative Finance and Machine Learning

This project implements an automated pipeline to forecast the Nifty 50 index using 5 years of daily historical data. The system integrates statistical models with technical indicators to deliver risk aware financial insights. A final accuracy with a Mean Absolute Percentage Error of 7.15 percent was achieved across volatile market conditions.



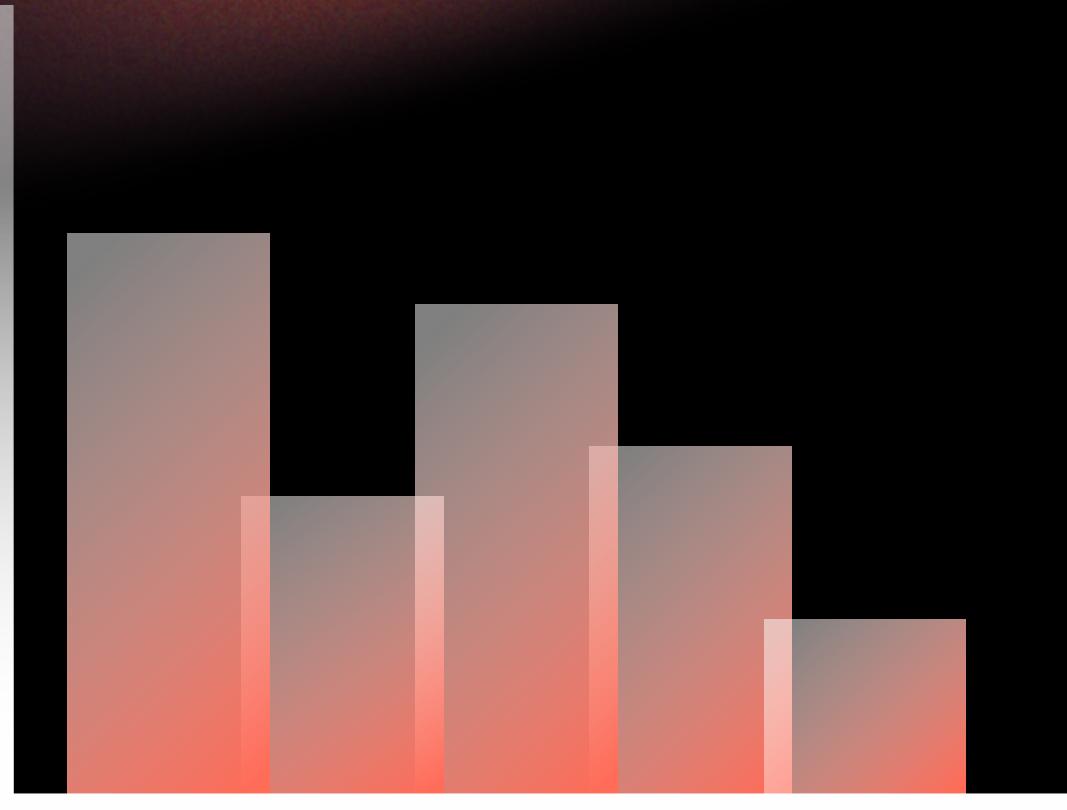
Data Engineering and Preprocessing

The dataset consists of 5,582 daily records spanning 22 years, with a focused analysis on the most recent 1,827 records. A forward fill method was utilized to handle missing data points during market weekends and holidays to ensure a continuous time series. The Augmented Dickey Fuller test was conducted to assess stationarity, showing an initial p value of 0.84 which indicated non stationary data. First order differencing was applied to stabilize the mean, resulting in a stationary series with a p value of 9.78e-20. Technical indicators including RSI, MACD, and SMAs were processed as exogenous variables to provide market momentum context to the models.



Time Series Decomposition and Analysis

The closing price series was decomposed into Trend, Seasonality, and Residual components using the statsmodels library. The Trend component identified a clear long term bullish trajectory despite short term fluctuations and global economic shocks. Seasonality analysis revealed recurring annual cycles, providing insights into periodic market buying and selling pressures. Residual analysis isolated random noise, specifically highlighting the extreme volatility during the early 2020 global crisis. This decomposition allowed the predictive models to focus on the underlying structural signals rather than transient market noise.

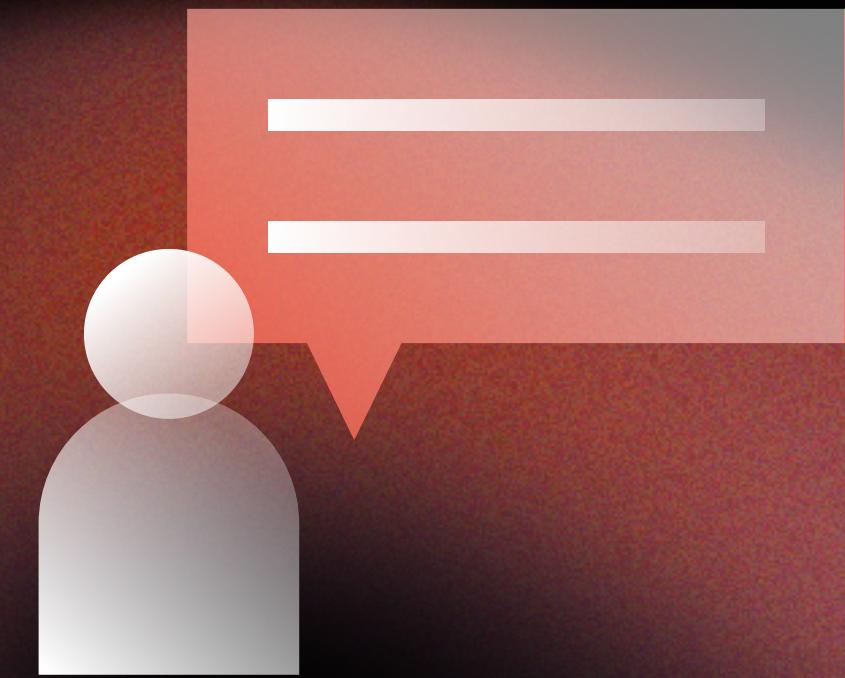


Modeling Framework and Evaluation

The ARIMA(2,1,2) model was implemented after optimizing parameters using AIC and BIC penalties to prevent overfitting.

The Facebook Prophet model was enhanced by incorporating technical indicators as extra regressors for contextual intelligence. ARIMA achieved a Root Mean Square Error of 1472.69, which is approximately 8.66 percent of the mean closing price. Prophet achieved a Mean Absolute Percentage Error of 7.15 percent, demonstrating superior performance in capturing external market influences. Validation was performed using 95 percent confidence intervals to provide a visual and mathematical representation of potential market risk.

Financial Insights and Conclusion



The forecasting pipeline successfully identifies a robust long term upward trend for the Nifty 50 index. Strategic recommendations suggest that investors should orient their decisions around the core trend line rather than reacting to daily residuals. Confidence intervals generated by the system serve as a quantitative guide for setting stop loss levels and managing portfolio risk. Residual distribution analysis confirmed the models were not overfitted, as errors were approximately normally distributed. Future improvements may involve integrating sentiment analysis from financial news or using deep learning architectures like LSTM for higher precision.