Goldman Sachs

GSIH 2024

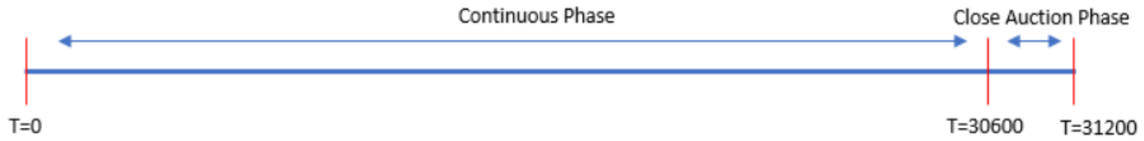
QUANT SECTION

Priyanshu Singh
IIT Kharagpur

PREDICTING CLOSE AUCTION VOLUME

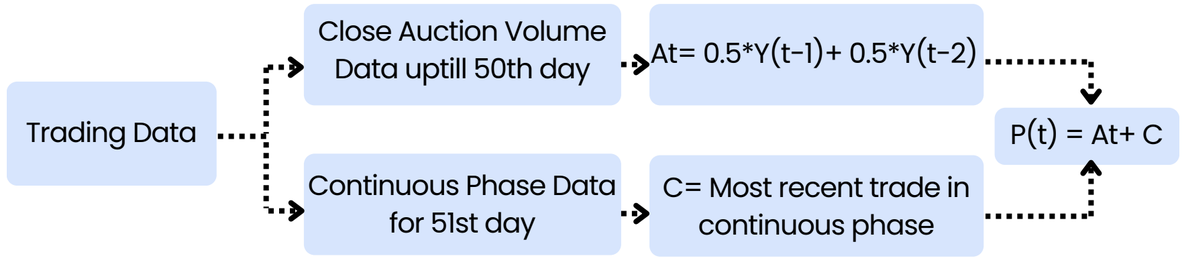
PROBLEM DESCRIPTION

- The task was to develop a model to predict the Close Auction Volume for stocks using historical trading data.



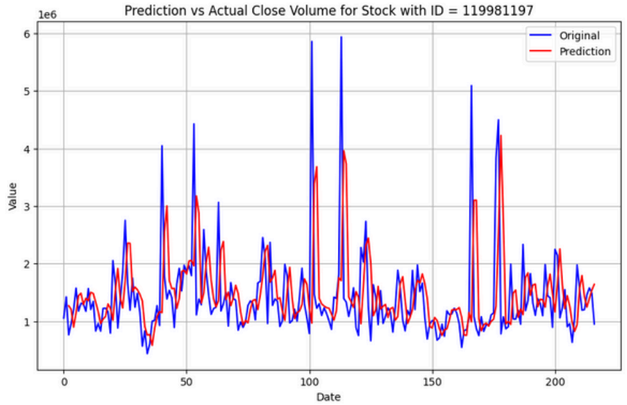
- The model was required to predict the 51st day's Close Auction Volume using the past 50 days' trading data and partial data from the 51st day.

PROPOSED SOLUTION



- Y(t-1), Y(t-2)= Close Auction Volume for (t-1)th day and (t-2)th day respectively and P(t)= Predicted Close Auction volume for (t)th day.

EXAMPLE TEST CASE



Shown is a plot comparing the predicted Close Auction Phase volume with the actual Close Auction Phase volume.

LIMITATIONS & IMPROVEMENTS

- The proposed solution underestimated the peaks and overestimated the valleys.
- To address the limitation in the proposed solution, a correction factor was introduced to improve accuracy.
- The correction factor was formulated using continuous phase trading data.

ALTERNATIVE SOLUTIONS

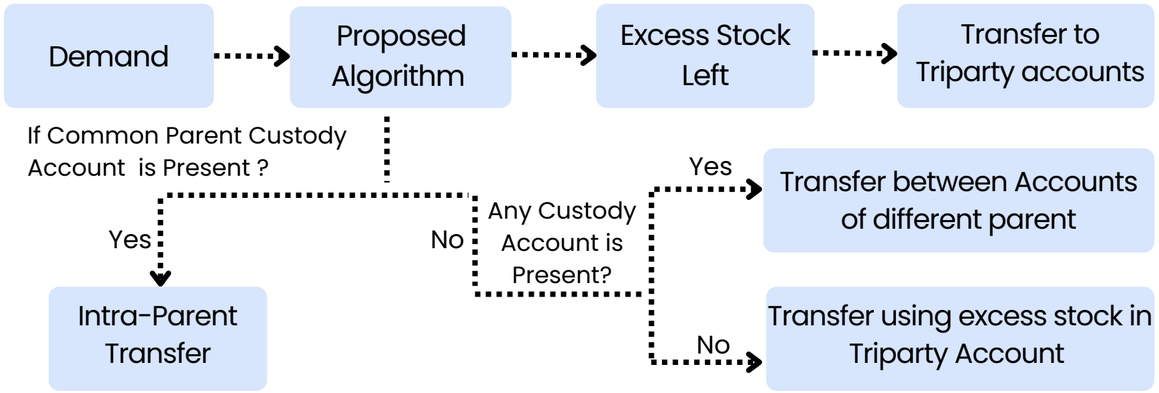
- Employing Time Series Forecasting models, including ARIMA, Boosting Regressor Trees, and LSTM .
- Leveraging Continuous Phase trading data more effectively in conjunction with Close Auction Phase data to create a robust predictive model.

MANAGING STOCK INVENTORY

PROBLEM DESCRIPTION

- The task was to develop a model to manage stock transfers, prioritizing the client demands, and minimizing the transaction costs within and across accounts.
- The algorithm needed to account for parent account hierarchies, where stock transfers within the same parent account incurred lower costs, and excess stocks could be pledged to triparty accounts to generate cash.

PROPOSED SOLUTION



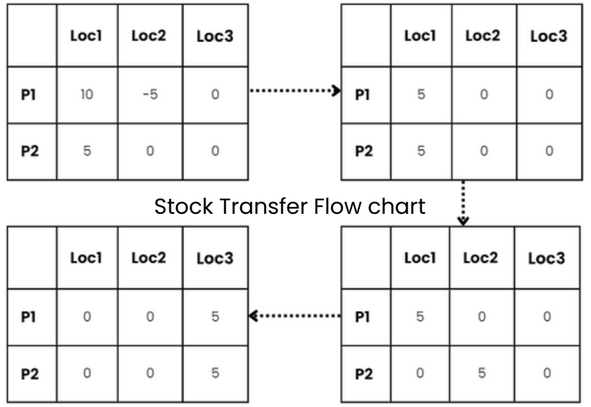
EXAMPLE TEST CASE

- Given Loc1 and Loc2 is Custody Account and Loc3 is Triparty Account.
- Loc2 and Loc3 have same parent id i.e. 2.
- Loc1 has parent with id=1.
- The matrix shown below represent eligible flow across the account.

P1	Loc1	Loc2	Loc3
Loc1	0	1	1
Loc2	1	0	1
Loc3	0	0	0

P2	Loc1	Loc2	Loc3
Loc1	0	1	0
Loc2	0	0	1
Loc3	1	0	0

- Aij =1 implies transfer is allowed from i to j



LIMITATIONS & IMPROVEMENTS

- The proposed solution exhaustively checks all possible flows, making it computationally expensive and inefficient.
- The solution employed a matrix representation to enhance efficiency and lower computational costs.

ALTERNATIVE SOLUTIONS

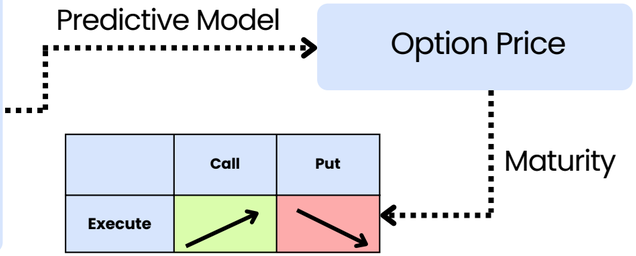
- A graph-based approach can be employed, utilizing optimized search techniques to efficiently determine stock movements.
- Utilizing linear programming to minimize transaction costs while satisfying demand and managing account constraints.

OPTION PRICING MODEL

PROBLEM DESCRIPTION

- The task was to develop a model to estimate the Option Price for derivatives, specifically Call and Put Options, based on historical data.

- Spot Value
- Strike Price
- Risk Free Interest Rates
- Time to expiry
- Type of Option
- Market Fear Index
- Buy-Sell Ratio

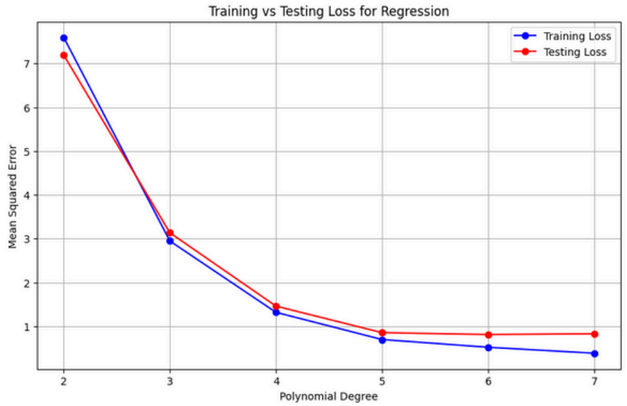


PROPOSED SOLUTION

- A polynomial regression model with L2 penalty and penalty coefficient 0.01, was utilized to predict Option prices based on historical data.
- Features were recursively projected into a higher-dimensional space to identify the optimal dimensionality, which was determined to be 5.

Input Vector X → Feature Transformation → Dimensionality Reduction → min || Xw-y ||2 +0.01 || w ||2

RMSE VS POLYNOMIAL DEGREE PLOT



The RMSE continuously decreased with the increase in degree, and the optimal point was achieved at degree = 5, as observed in the plot.

LIMITATIONS & IMPROVEMENTS

- The proposed solution utilizes feature projection into higher dimensions, resulting in a complex model that reduces the explainability.
- To overcome the limitations, dimensionality reduction using SVD and PCA was applied, resulting in feature reduction and improved explainability.

ALTERNATIVE SOLUTIONS

- Using models like Black-Scholes and Binomial Tree, which are statistically robust and offer clear explainability.
- For a ML-based solution, boosting techniques such as CatBoost, XGBoost, and GBM can be utilized.