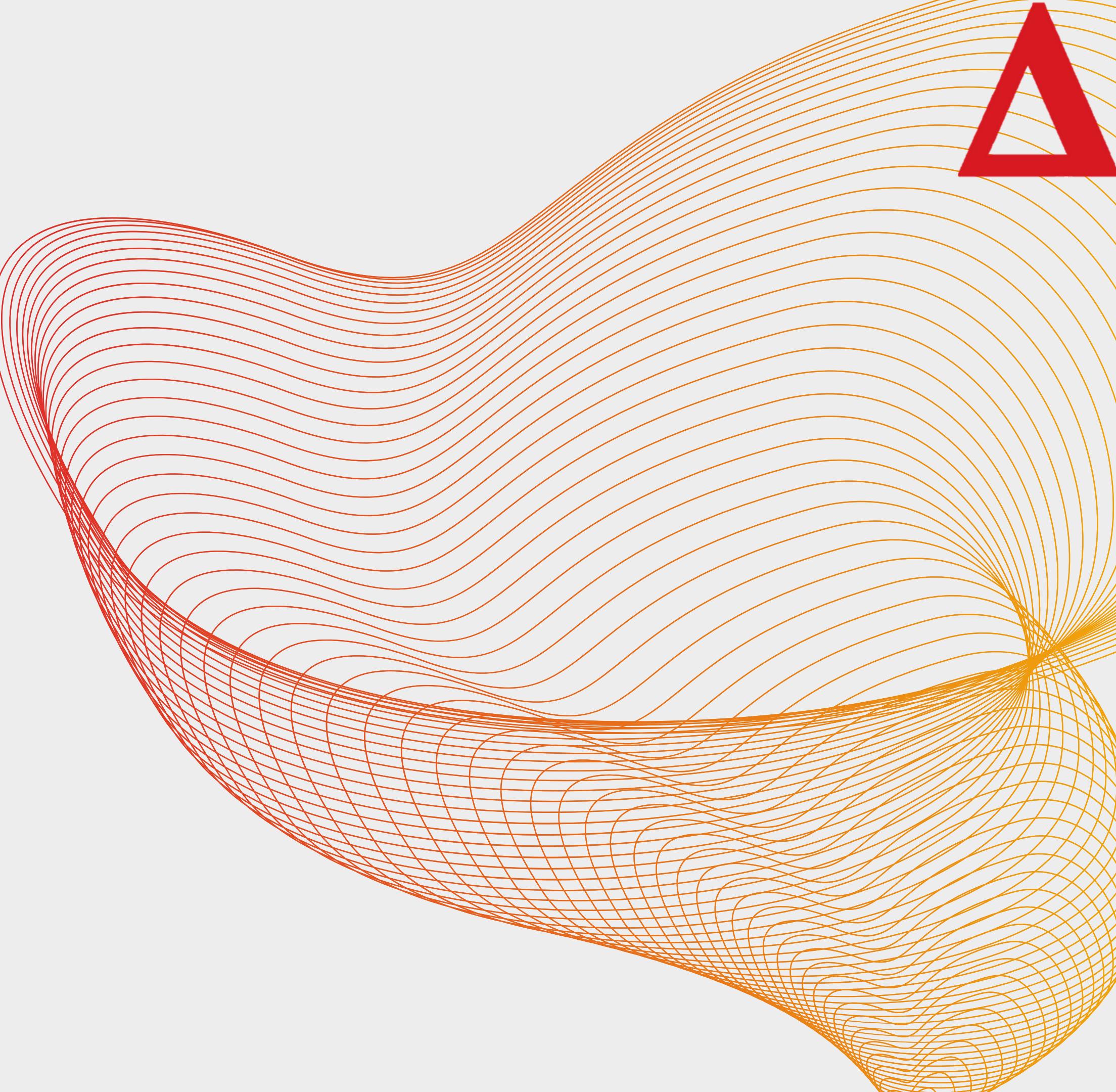
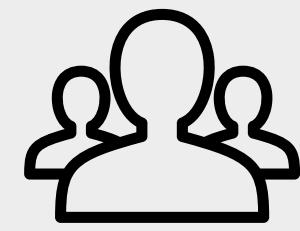


Optiver volatility Project

Created by
Optiver Team 3





Our Team



Kellen Liew

Final Year Software
Engineering

- Trading Intern
- Data Science and Engineering



Richard
Zhong

Final Year Software
Engineering

- Neural networks
- Data engineering



Rutvik

MS in IT

- 3.5 Years Work Ex. at Siemens.
- Advanced Algorithms



Arya

Computer Science

- Data, Insights, and Decisions
- Data analysis





Contents



- **Introduction**
 - Our Team
 - Project Background
 - Aims and Scope
 - Current Market Models
- **Investigation**
 - Model Development Storyline
 - Final Model
 - Global vs Local
- **Findings**
 - Evaluation metrics
 - Results and Performance
 - Limitations
 - Recommendations
- **Value Proposition**
- **Summary**



Project Background



Why calculate Volatility

- Volatility is a variable in calculating option prices
- Being confident in volatility allows Optiver to submit more competitive orders

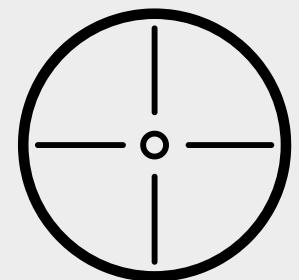
Established methods at predicting volatility

- GARCH
- Exponentially Weighted Moving Average (EWMA)
- Machine learning

Problems with current methods

- Zero Shot/Few Shot Capability
- Interpretability
- Model Risk





Aims and Scope



The Idea

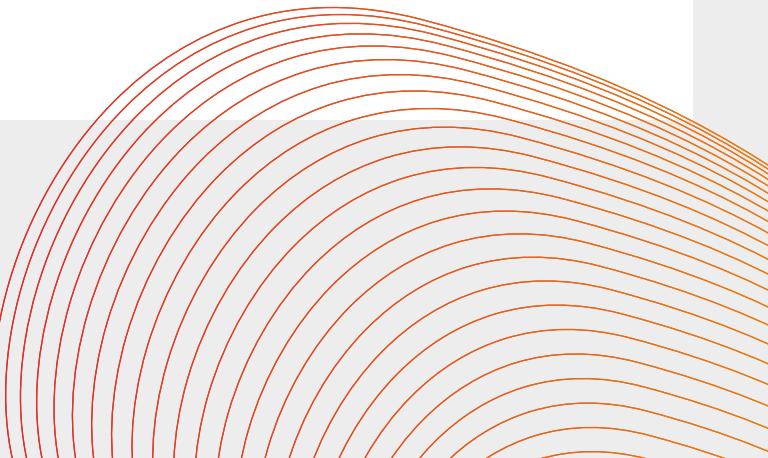
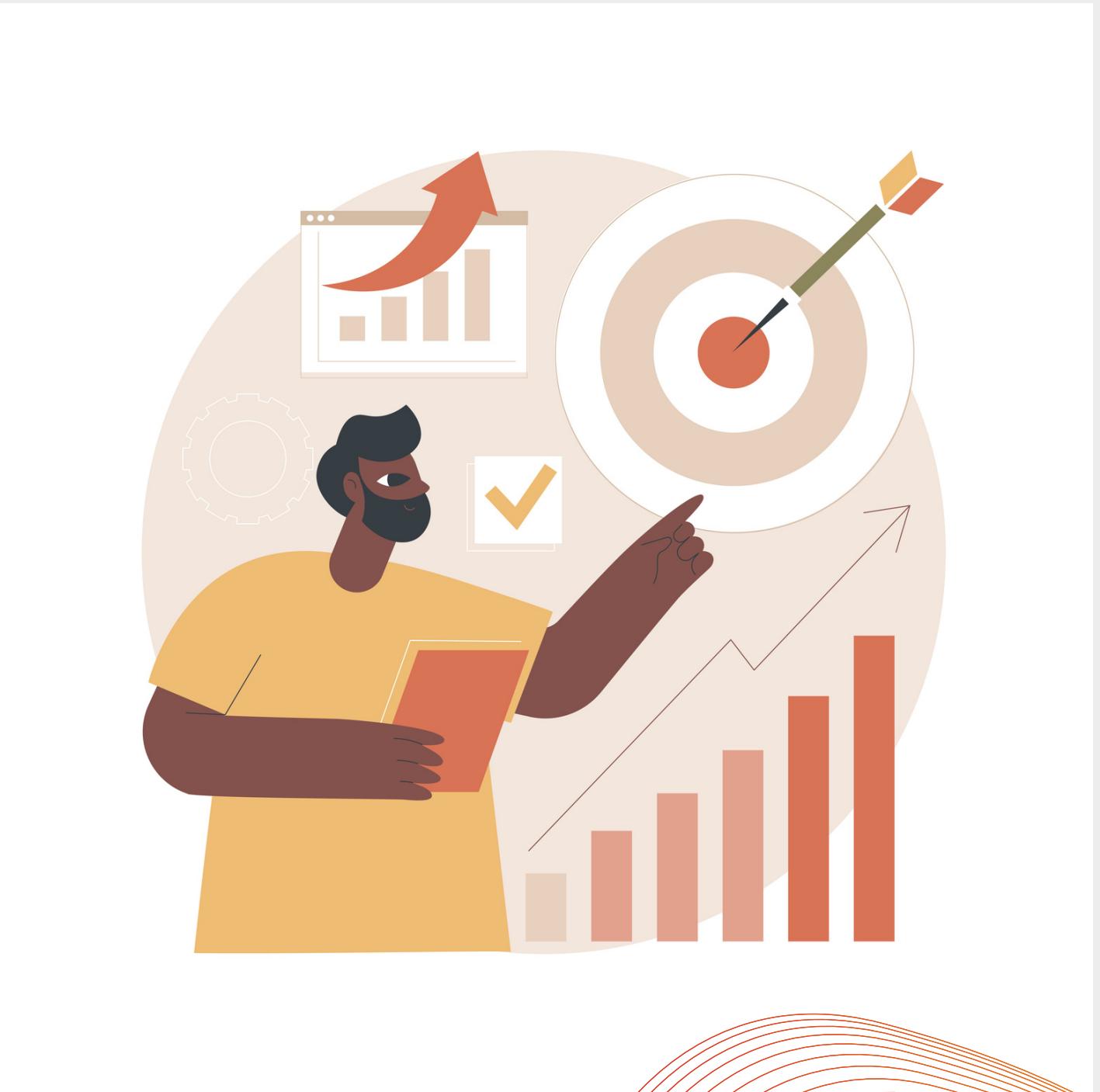
Correlation (Beta) can be used to reduce the error in volatility estimation.

Intended users

- Researchers
- Traders

Project Goals

- Iterate upon previous attempts with modern methods
- Produce a viable sparse data solution to calculating volatility
- Determine the viability of correlation as an error improving factor in estimating volatility.





Current Market Comparison



Comparing Against other Prediction Approaches

	Improves upon base method	Sparse Data	Explainability	Considers wider market variable
GARCH	✓	✓	✓	
Machine Learning	✓			✓
Proposed Beta Method	✓	✓	✓	✓

Strengths of the proposed model

- Low Cost and Faster as its raw computations
- Understanding Correlation Relationships help manage risk in a portfolio.
- Model equation is easy to understand and can be improved upon.
- Can be used on variety of financial instruments.



MODEL DEVELOPMENT STORYLINE

A brief on development



WK 3



Created baseline prediction model based on WAP movement

WK 4



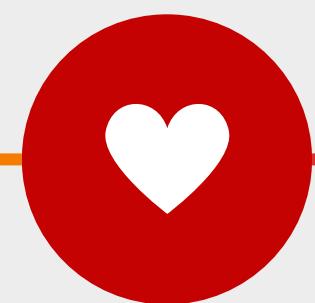
Conceptualised Beta Correlation Component

WK 5



Log returns for estimating base volatility

WK 7



Best Beta calculation

WK 8

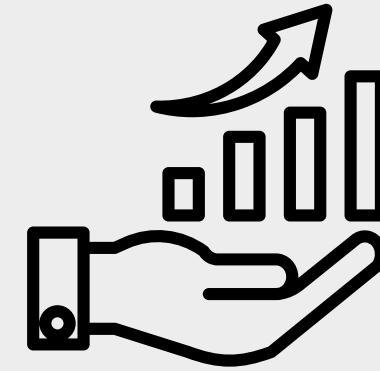


EWMA for volatility estimation

PRESENT



Final model



Equation

Volatility estimation is done using Exponentially Weighted Moving Average (EWMA)

Two ways to calculate Beta:

- Local Beta
- Global Beta

$$\sigma_{target} = \sigma_{estimated} + \phi$$

ϕ = error

σ = standard deviation/volatility

$$\phi = \beta * abs(\sigma_{SPY,curr} - \sigma_{SPY,prev})$$

β = correlation b/w stock and S&P500 index

$$\beta = \frac{\text{Covariance(stock, S\&P500)}}{\sigma_{stock} * \sigma_{SPY}}$$



Comparison



Global Beta:

- Considers entire time series data for beta calculation

Pros:

- Considers entire history of the stock performance
- Shows overall correlation of the stock to index
- Good for longer time period volatility estimation

Cons:

- Events that might not be relevant now, are considered
- Data heavy approach

Local Beta:

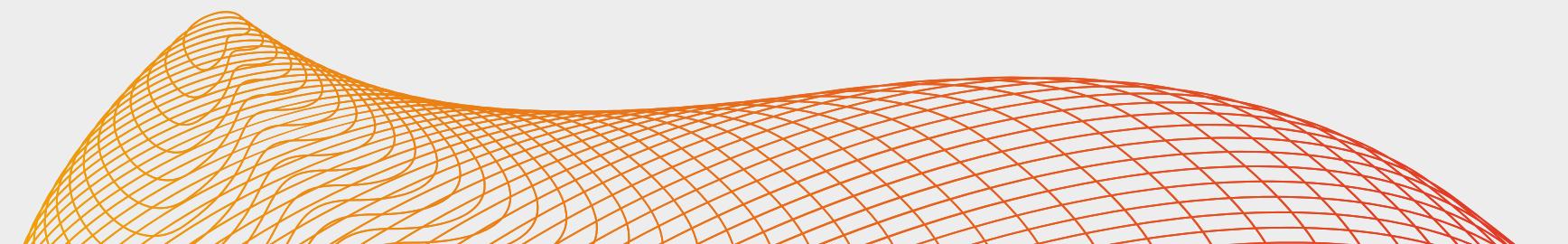
- Considers time series data for the required time period

Pros:

- Shows relevant correlation for that time period
- Not data heavy
- Good for shorter time period

Cons:

- Beta is calculated every time period which as per definition should be a more constant value





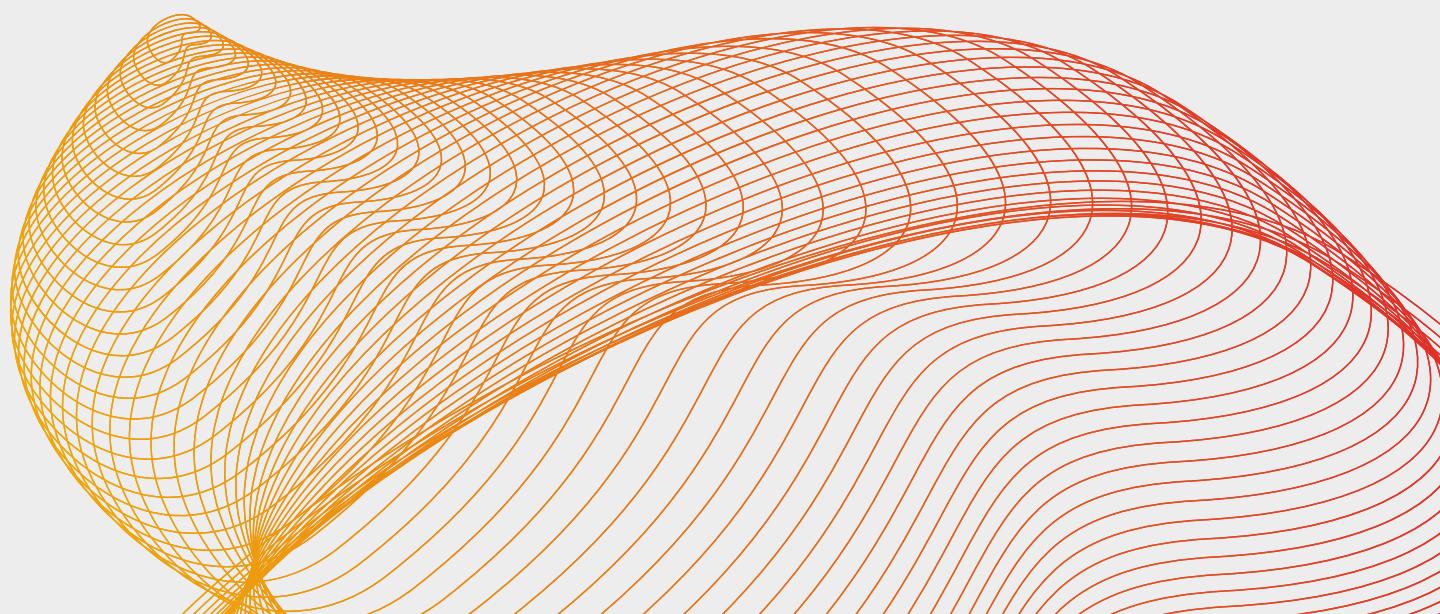
Evaluation Metrics

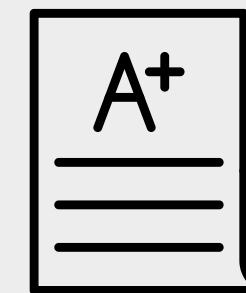
- Performance analysis of the overall model can be done by seeing the overall R2 and RMSPE scores.
- Individual stock performances (gain and loss)
- Beta Error – frequency of beta leading us in the wrong direction i.e. instead of reducing error it increased it.
 - Happens if the estimated volatility is actually lesser than target volatility.

Note:

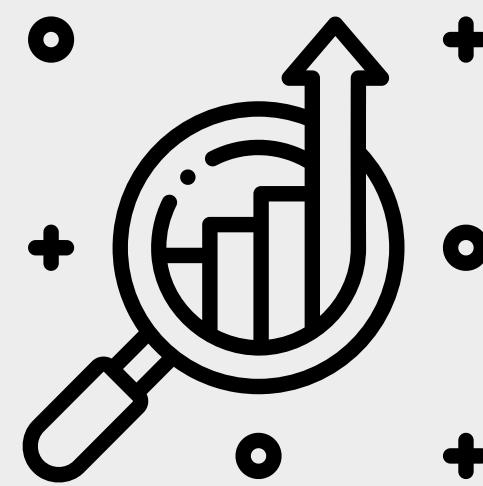
R2 score: Higher the better

RMSPE: Lower the better

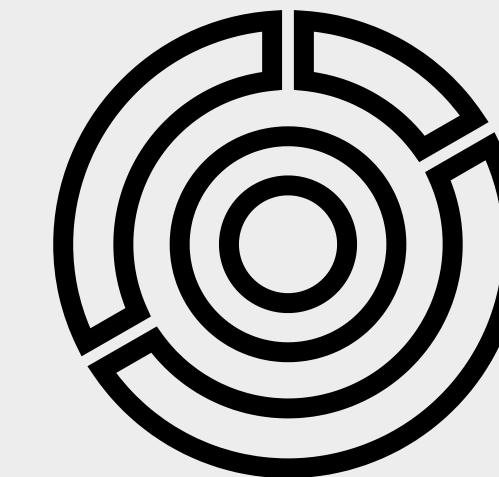




Results



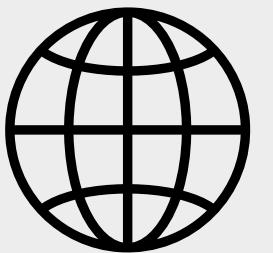
Naive Model:
R2: **0.845**
RMSPE: **0.21**



Global Beta Model:
R2: **0.865**
RMSPE: **0.198**



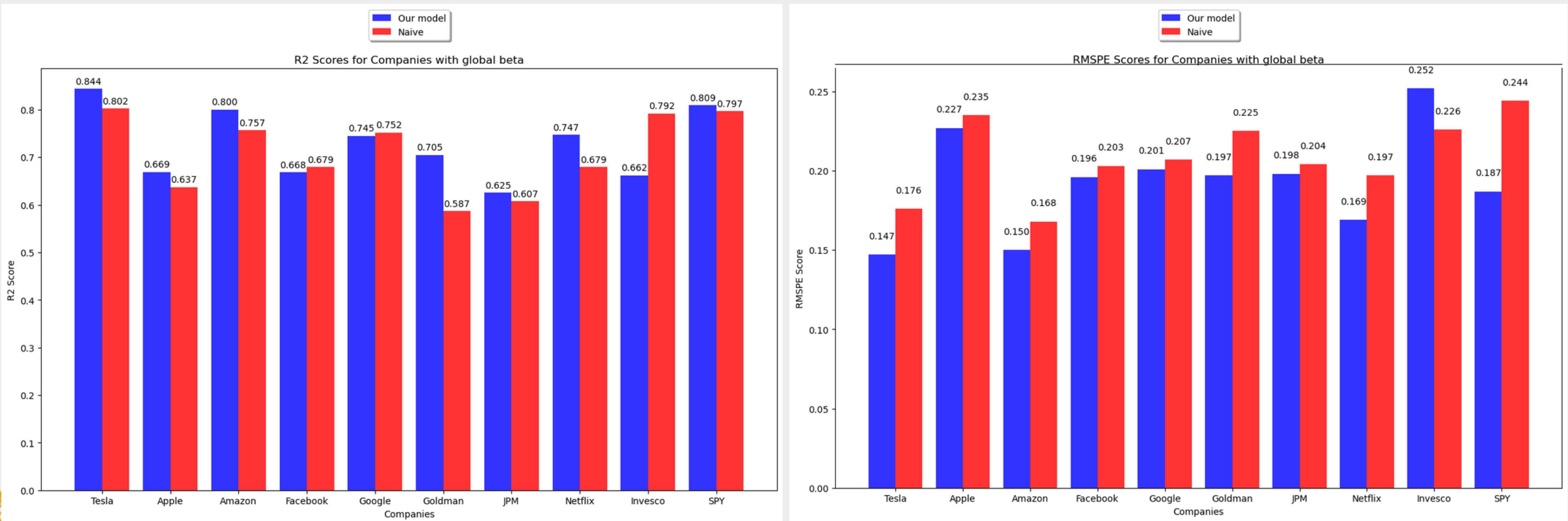
Local Beta Model:
R2: **0.864**
RMSPE: **0.198**

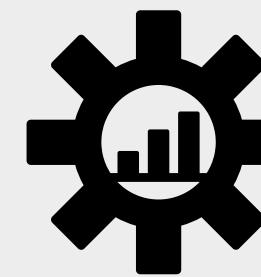


Global Beta performance



Deep dive into per stock evaluation scores:

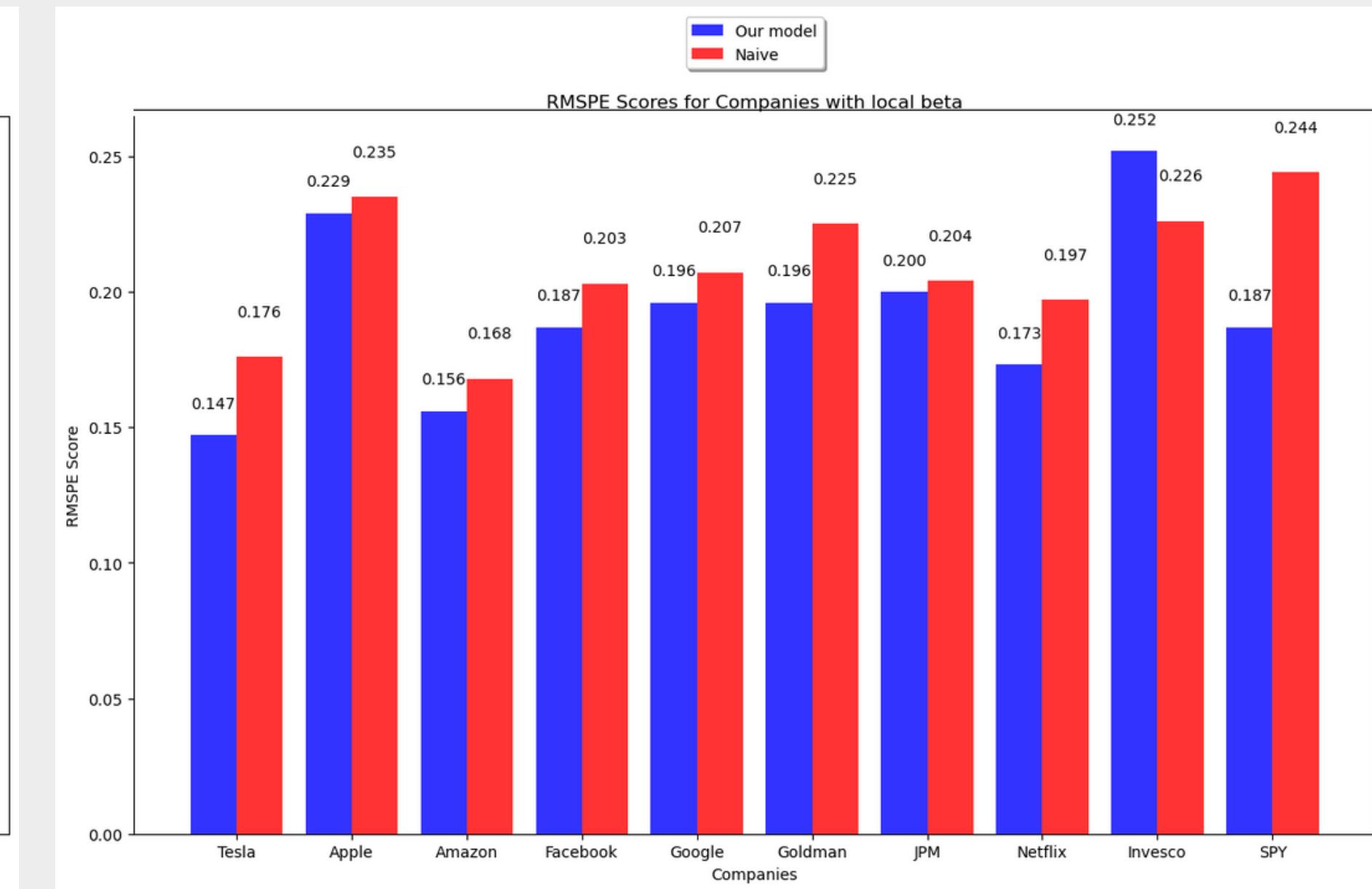
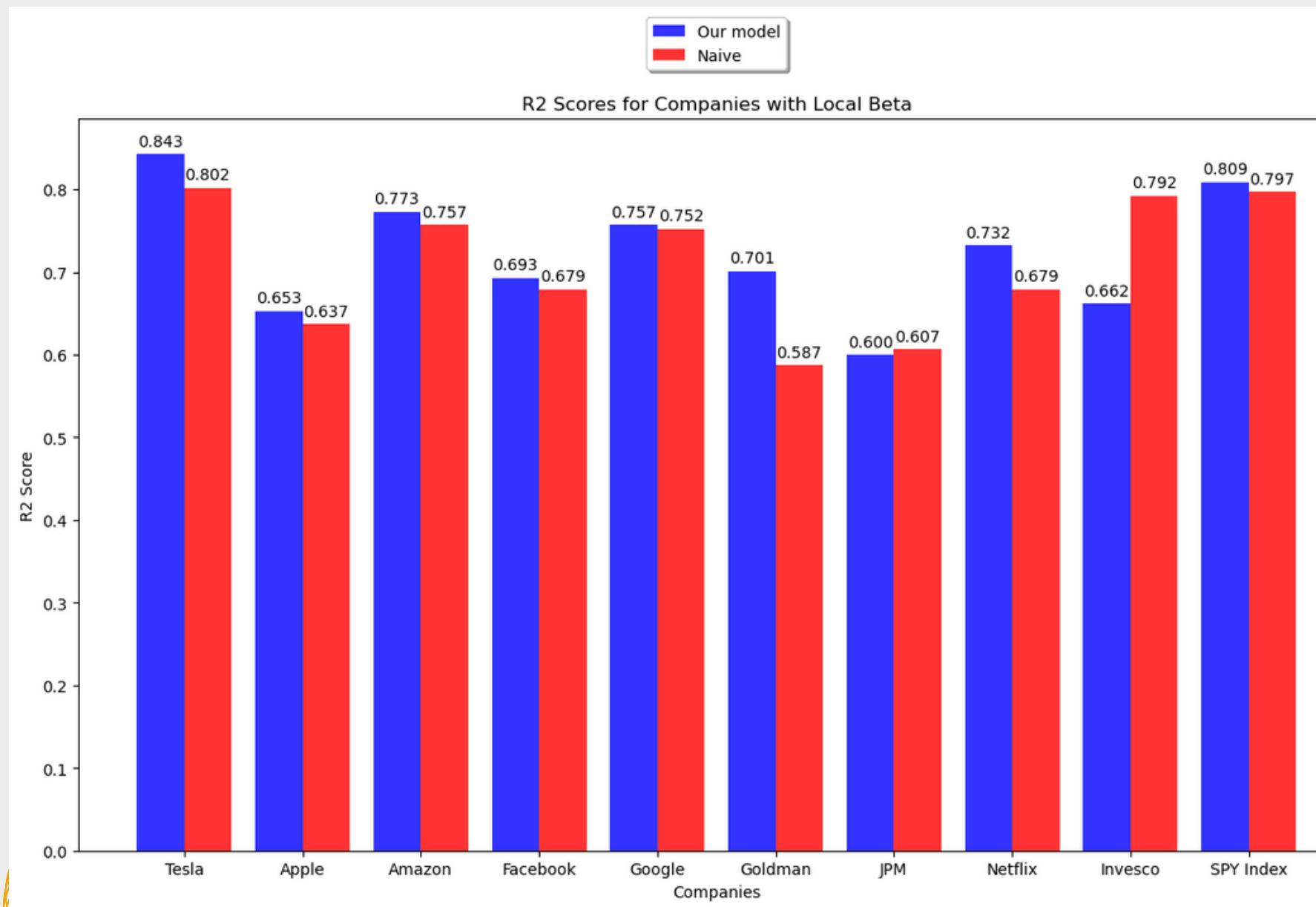


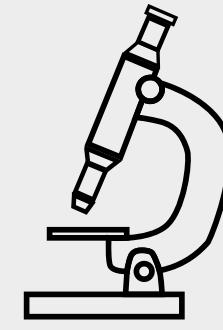


Local Beta performance



Deep dive into per stock evaluation scores:





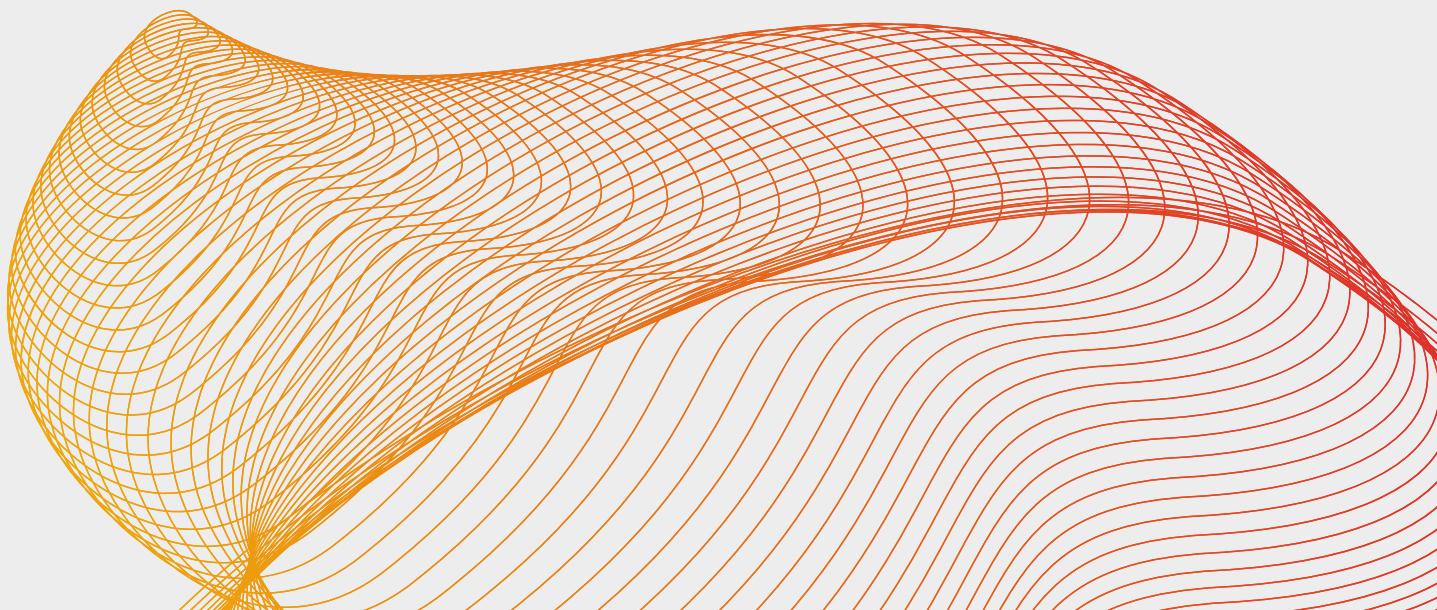
Highlights

Evaluation of per stock performance:

- Tesla (-ve correlation)
 - 4.5% increase in R2
- Goldman Sachs
 - 13% increase in R2
- Netflix
 - 5% increase in R2
- Invesco
 - 13% decrease in R2

! General:

The magnitude of gains on certain stocks is higher in global beta model attributed to a more data centric beta value used for correction leading to a higher overall R2.

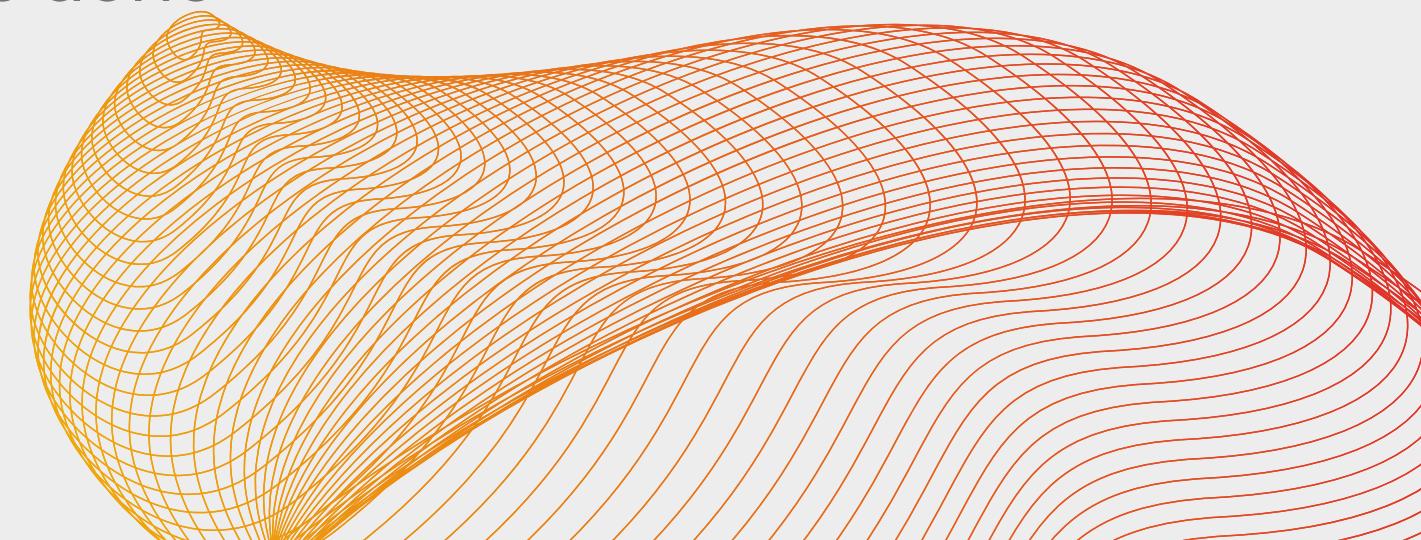




S&P500 limitation

S&P500 Limits:

- S&P500 cannot be considered in individual stock performance and beta error as it is used as a baseline for other stocks and is an input to the model.
- Correlation of S&P500 with itself is 1 so essentially we would not be estimating the volatility for the time period from the equation.
- For overall results, volatility estimation of S&P500 is done using EWMA only.



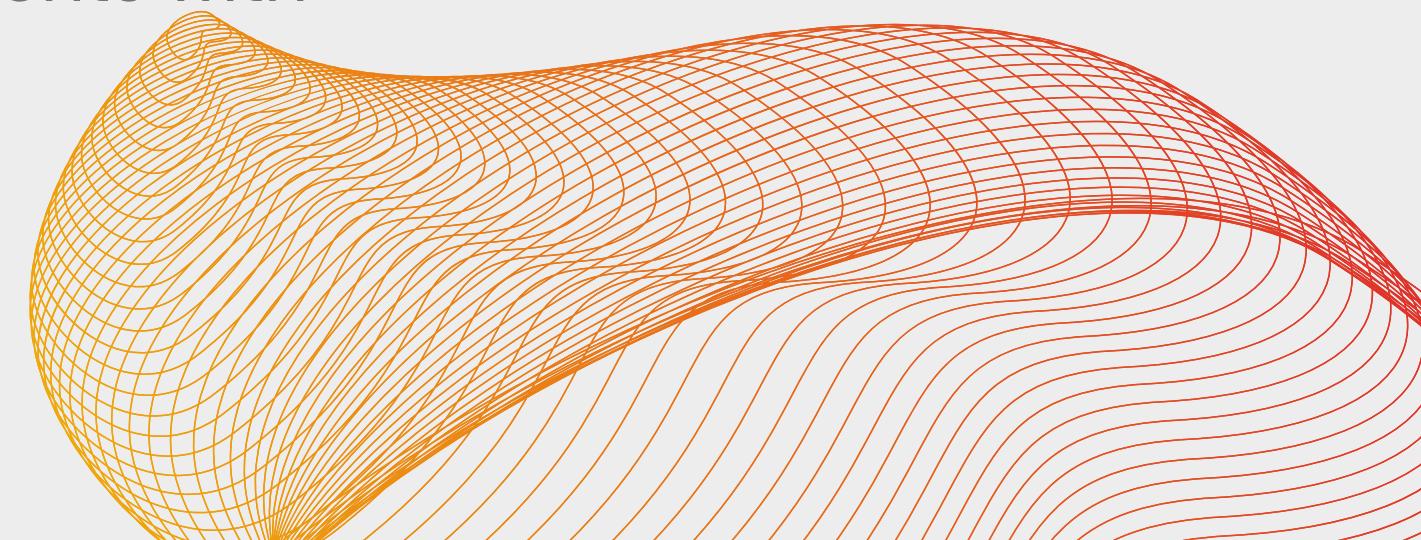


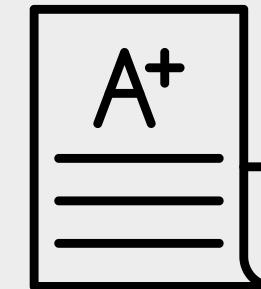
Invesco limitation



Invesco Failure:

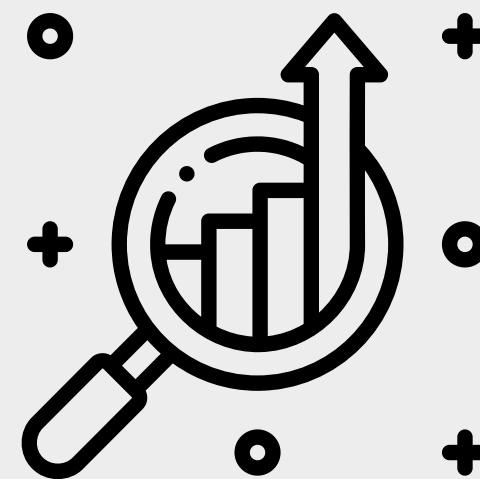
- Invesco is not a part of S&P500, rather tracks NASDAQ.
- Correlation with S&P500 does not guarantee accurate results, should be rather correlated to NASDAQ.
- Invesco is also a more volatile stock with sharp turns and gains.
- EWMA volatility estimator adapts poorly for sharp increases and decreases as it does not weight the recent events with equal weights like Naive model.



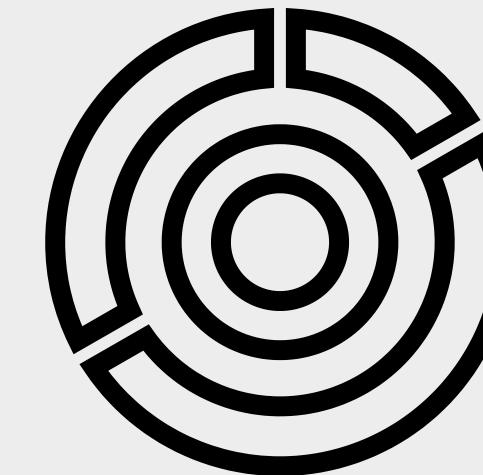


Adjusted Results

Growth in model metrics without limitations i.e. without S&P500 and Invesco



Naive Model:
R2: **0.799**
RMSE: **0.203**



Global Beta Model:
R2: **0.828**
RMSE: **0.187**



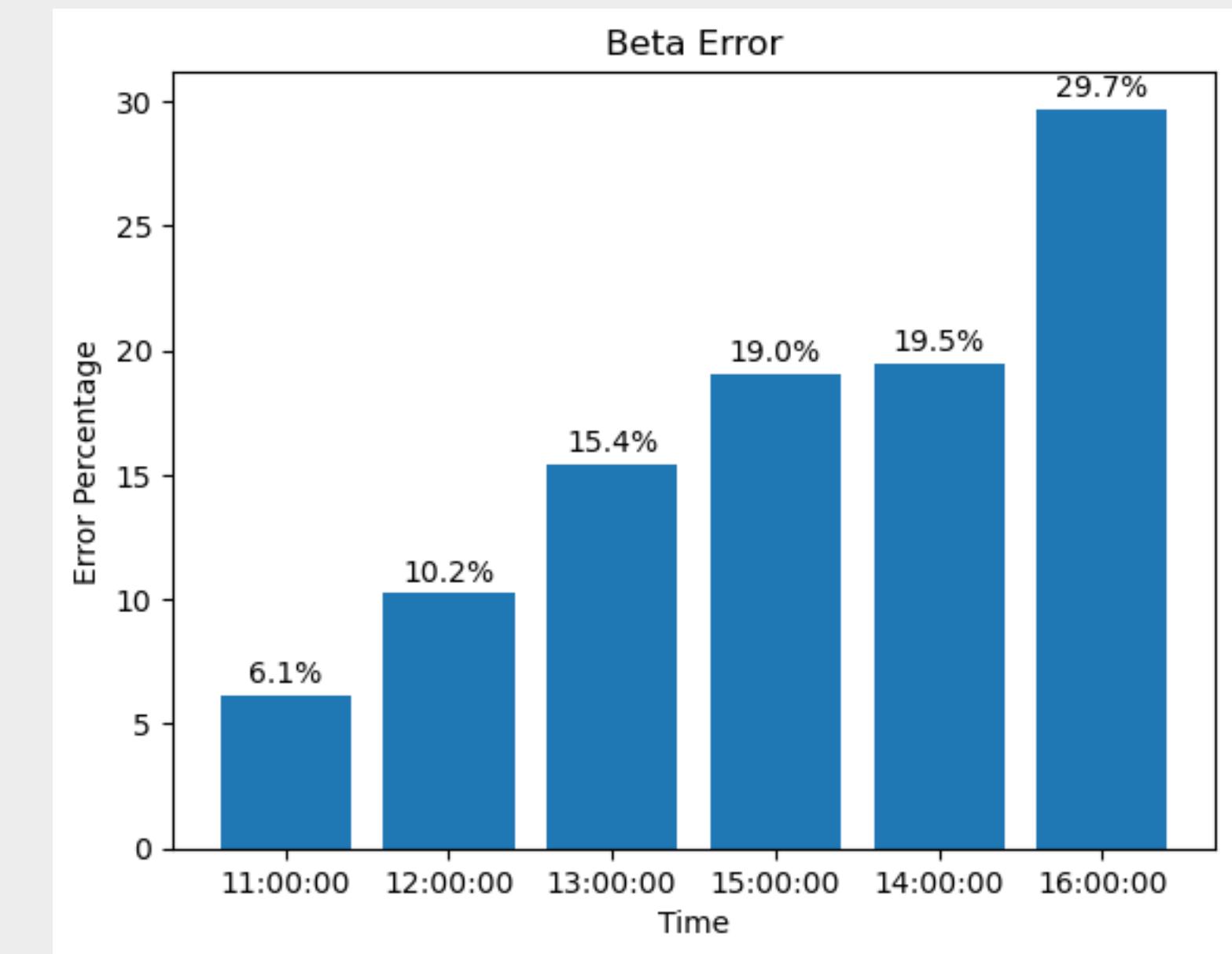
Local Beta Model:
R2: **0.828**
RMSE: **0.187**

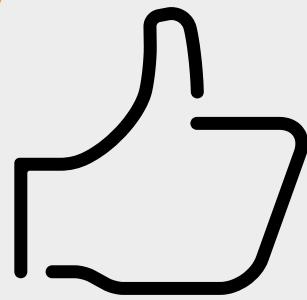


Beta Error

Analyzing Beta Error:

- Perfect performance for opening hours 11am and 12pm (6% error)
- Inaccuracies with direction are very high for 4pm (30% error)





Recommendations

- Trying out new methods
- Neural networks and machine learning
- Ensemble methods
- Limitation on the index Stock is listed on. (Invesco - QQQ not part of SPY)
- Add a flag for beta for 16:00 hr of the day to minimize the error.
- Weighted Correlation





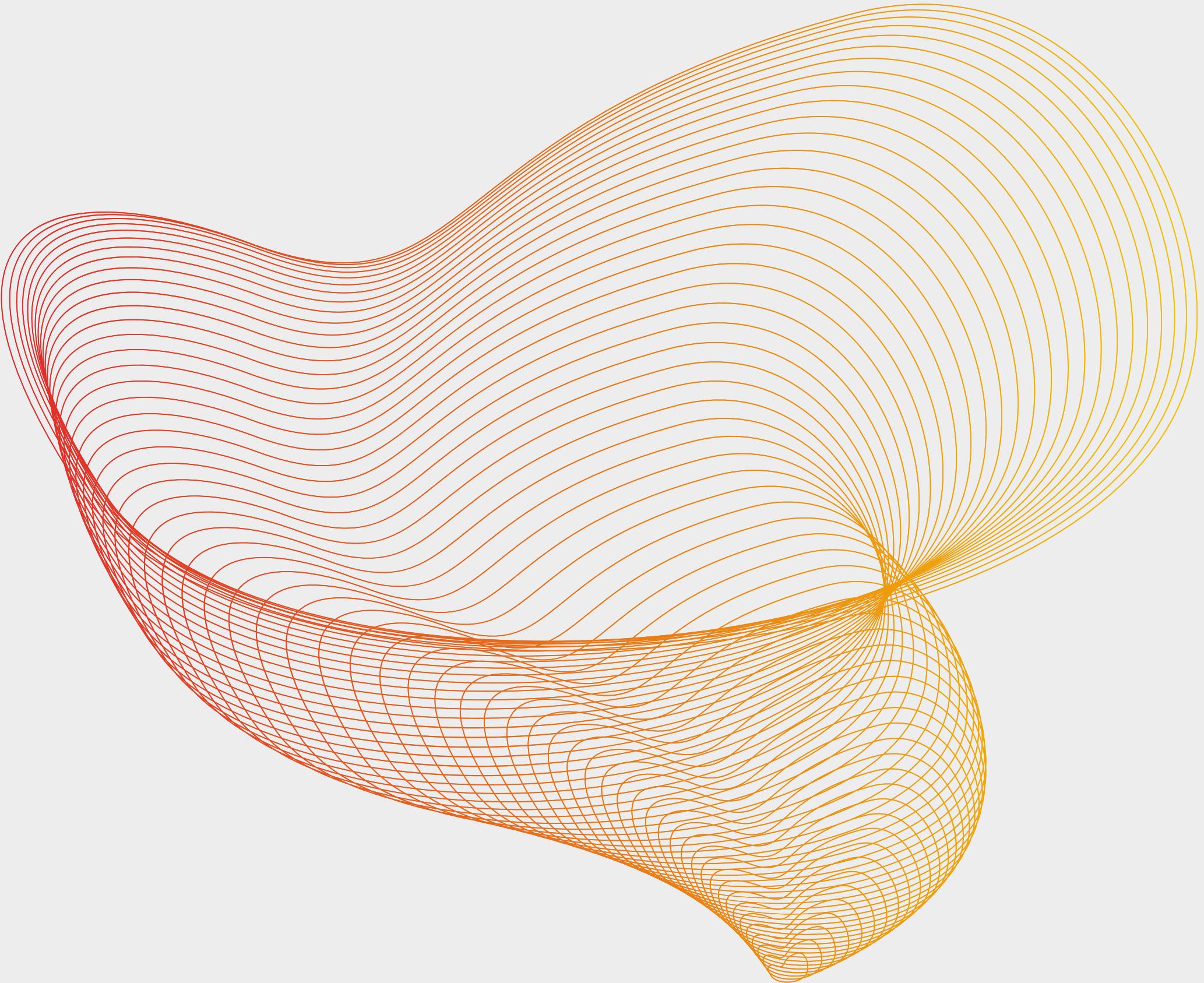
Value proposition

- Proved the hypothesis that correlation can be used to reduce error in volatility estimation.
- Works best for opening hours and negatively correlated stocks.
- Good runtime. (1 sec)
- No model training, its plug and play.



Summary

- Model Idea
- Model creation process
- Overall accuracy improved
- Future improvement suggestions
- Value proposition





Thank You!