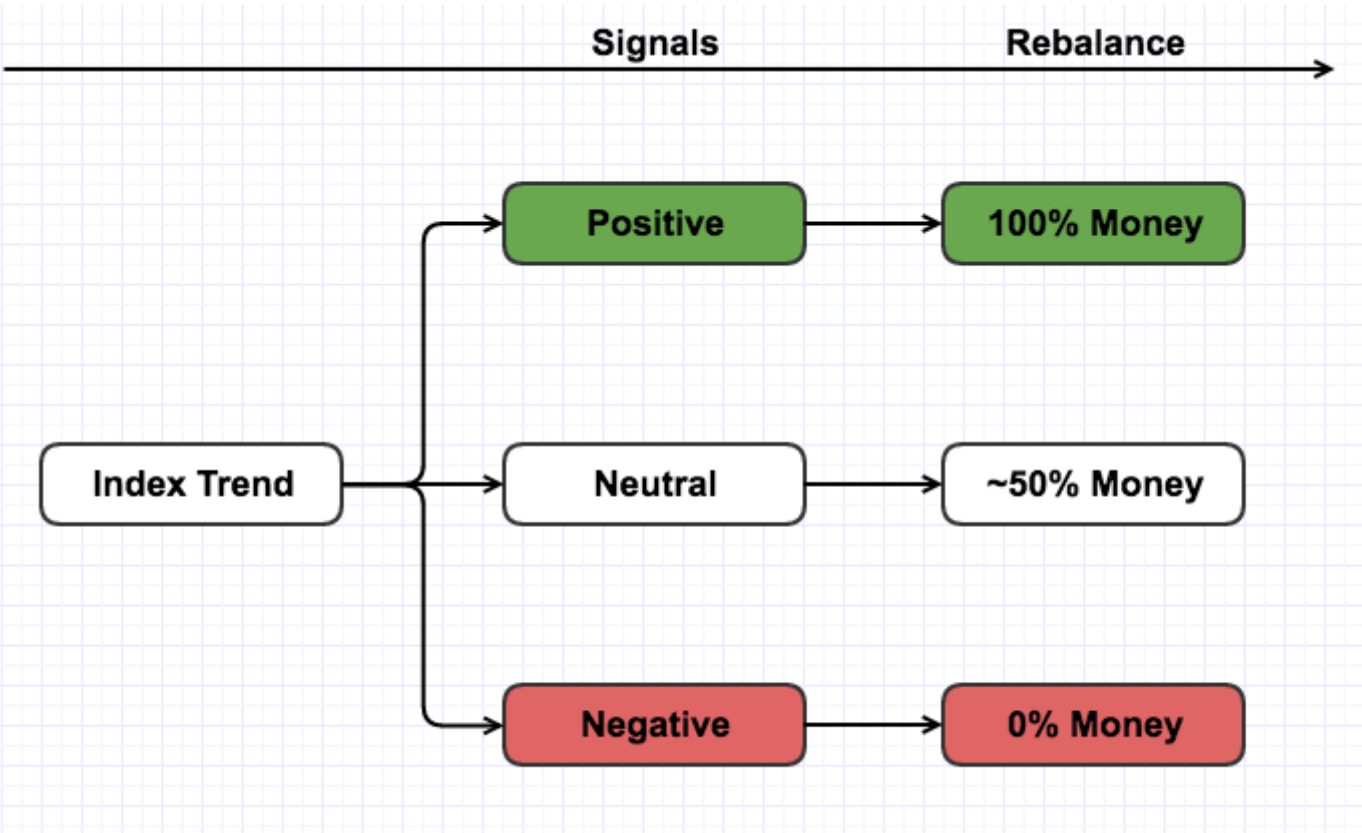


Spark-in-Finance-Quantitative-Investing

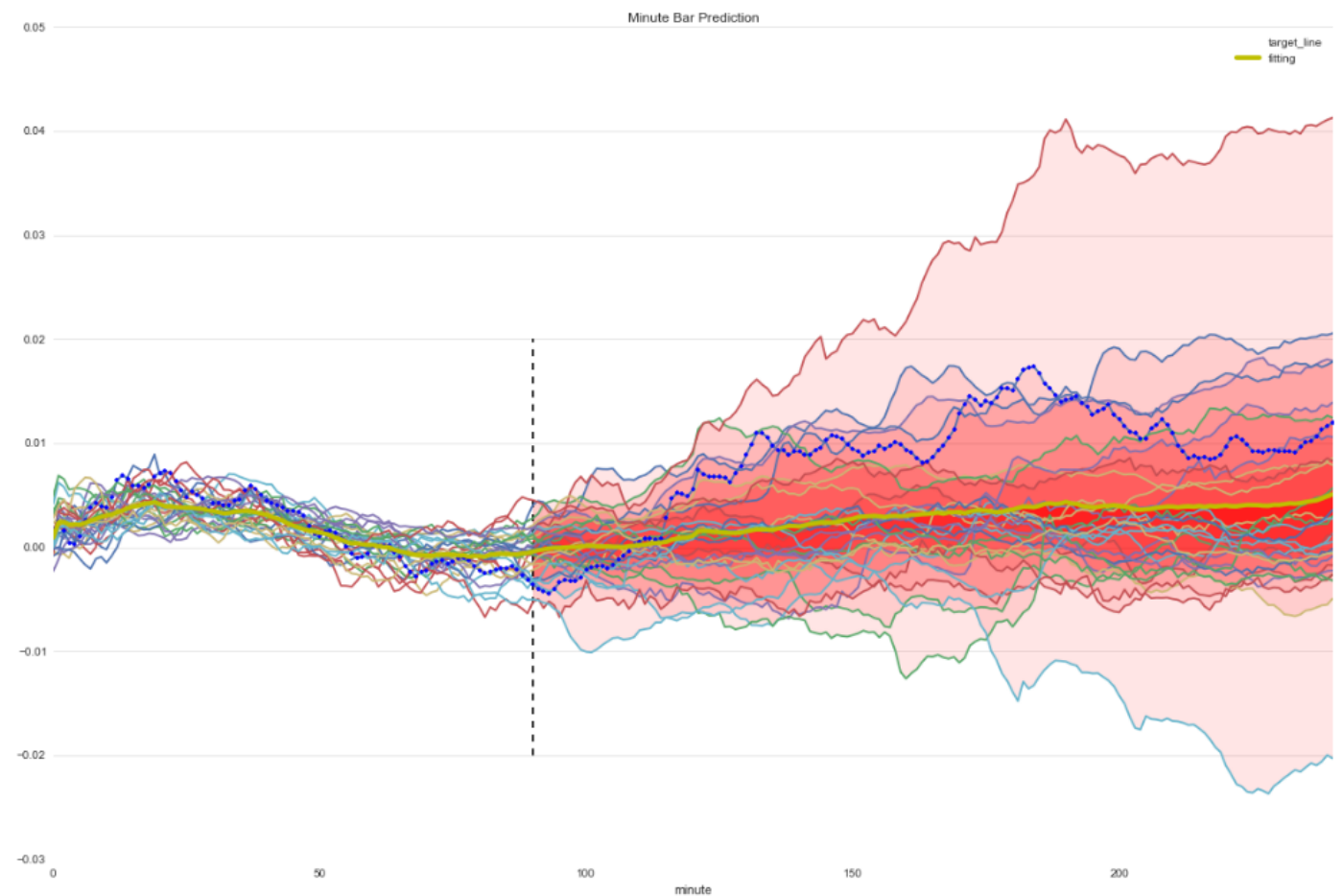
Product

This is a module in some financial investment strategies. Through analysing the trend of index, we can build some macro timing or portfolio adjustment signals.



Screenshot

similarity line and predict trend



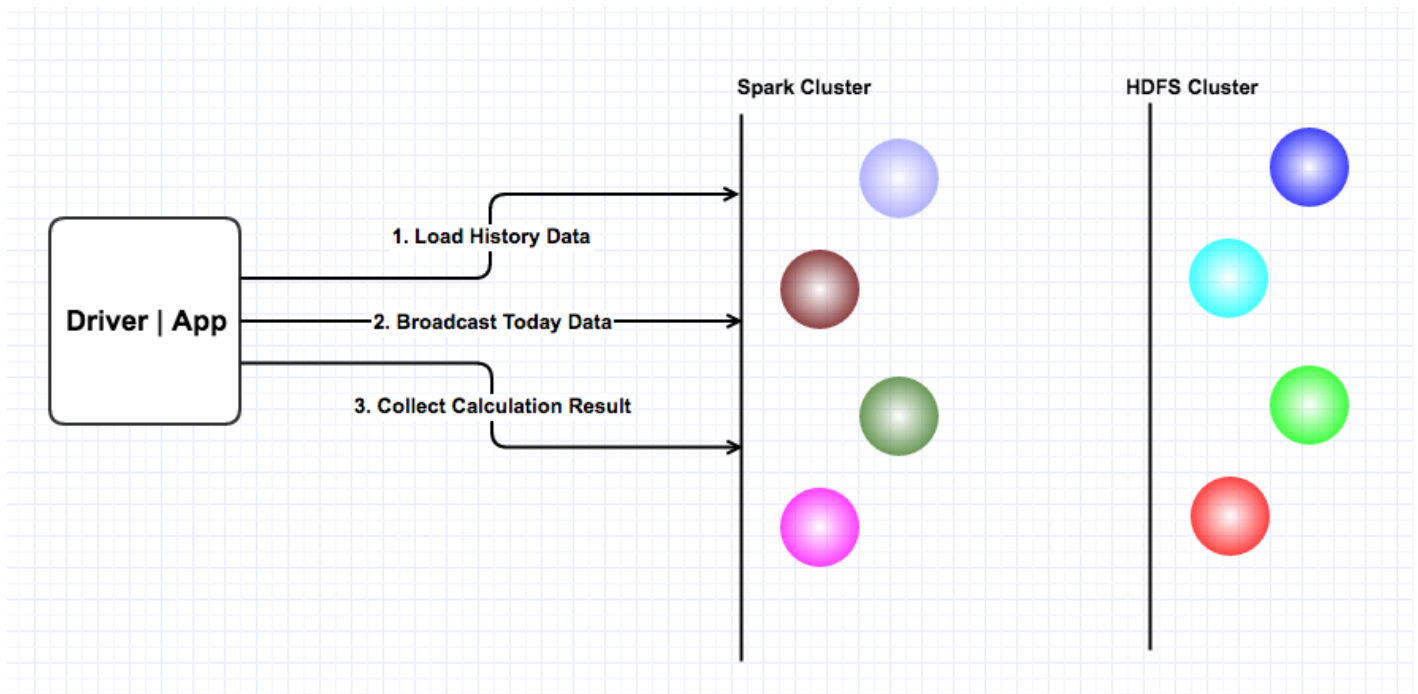
prediction close index

	fitting	target_line
236	0.003979	NaN
237	0.004014	NaN
238	0.004172	NaN
239	0.004300	NaN
240	0.004506	NaN

Architecture

Bellow is the brief indroduction of this app:

- During the transaction time:
 - Spark cluster loads all the history data;
 - Dirver loads today's data;
 - Driver broadcasts today's data to the cluster;
 - Spark cluster parallely calculated similarity data;
 - Dirver collects the calculation results;
 - Dirver parse the calculation results;



▸ Data Used

This application used several data below:

- Index minute bar in the previous several years;
- Index minute bar of today, refresh every one minute;

▸ Algorithms

Just for the demo, I used the basic similarity algorithm.

▸ Value of Product

This application is a module in some quantitative fund, it can be used in macro timing and portfolio rebalancing.

For the future, there are endless imaging and extension space. For example, as the data growing more and more, we can put more data in the algorithm, and design several different algorithms to do the prediction parallelly, leverage the power of big data and Spark, complete a calculation round within 1 second, build high-frequency signals in the market. Which will be a revolution in the financial market.

I currently use a more complex algorithm to calculate the similarity and build macro signals in my private investment account, it really works, and I believe it will do much better in the future.

▸ Improvement

- More data:
 - Using more history data
 - Using more kinds data
 - price
 - volume
 - money flow
- Signals
 - Macro timing
 - Portfolio rebalance
 - Ticker timing
 - Ticker pair trading strategy
- Larger cluster

Future

In recent years, as the big data announcing its power, more and more frameworks show their muscles to the world. Apache Spark is one of the most powerful frameworks in my eyes, and now there are many companies that start put Spark in their business systems. But, as known, most of today's Spark applications are dealing with logs and machine learning models, to be frankly, we do not leverage the power of big data and Spark until the Financial Market pays attention to big data and Spark.

In quantitative investing field, we use some mathematic methods to analyse market data, in order to build some signals in our strategies. As the data grows more and more, as the models become more and more complex, we need a powerful tool to do parallelly computing on mass market data.

This application, ***Spark-in-Finance-Quantitative-Investing***, is just a demo, but it's really powerful and useful. I believe through this application, the quant can find that Spark is the thing they are seeking to do the analysing, computing, modeling, etc job.

Github

[Spark-in-Finance-Quantitative-Investing](#)

Video & Documents

- [Spark-in-Finance-Quantitative-Investing.pdf](#)