Analysis of Quantitative Investment Decision System in Financial Markets Based on Python

Ziju Hou

Abstract

The influence of computer technology in financial markets has escalated in recent years. Amidst the financially affluent markets brimming with data, decision analysis via quantitative investing leveraging computer technology enhances investment efficiency. Notably, it mitigates investment risks to some extent. From a business standpoint, the feasibility of quantitative investment projects can be observed with the assistance of computers, facilitating timely adjustments in business strategies to maximize corporate value. The resultant analysis serves as crucial reference material before listing.

This thesis will elaborate on the relevant theories of quantitative finance in economics, analyze cases of quantitative investing in financial markets, and establish an investment decision system with Python as the development platform supported by extensive data. It encompasses economic theory, data analysis, and computer technology, ultimately providing a practical platform of reference value for businesses and individuals, contributing to investors' final decision-making.

Keywords: Moving average strategy; Python; Financial engineering; Financial mathematics; Financial modeling.

1.1 Research Background and Significance

Amid the maturing landscape of technology, an increasing scope and range of fields within the financial industry are integrating with computers, and the significance of quantitative finance is on the rise within the financial domain. Concerning quantitative investment and traditional subjective investment, the two are complementary. The superiority of quantitative investment lies in its enhanced speed and breadth of data processing, with the ability to handle substantial weak-logic data within a given timeframe. However, traditional subjective investment has its advantage in investment depth. The research methodology proposed in this thesis aims to create a quantitative investment system using Python, allowing investors to combine financial tools with subjective investment, thus forming superior investment strategies.

Python, a widely used development environment, is suitable for developing and building various systems, including stock analysis and trading systems. As a programming language, Python possesses simplicity of operation, ease of learning, readability, and writability, along with extensive support for third-party databases. Additionally, utilizing Python can aid investors in more accurately analyzing the stock market and making investment decisions based on real data. Employing Python as a tool for financial quantitative analysis helps extract valuable information from vast data, providing investment decision support for businesses or individuals. This thesis aspires to innovate financial tools to construct a system that aids investors in quantitative investment analysis, thereby enhancing investment efficiency while reducing investors' risks.

1.2 Research and Applications of Quantitative Investment

Quantitative investment finds applications across funds, stocks, and bonds by surveying the global capital markets. Quantitative investment has undergone several decades of evolution in developed nations' capital markets and has now become a relatively mainstream investment approach [1]. Taking the fund industry as an example, the proportion of quantitative hedge funds is notably high. In some top-ranked companies, more than half are quantitative hedge funds, displaying a pronounced leading effect [1].

China's development of quantitative investment started later, transitioning from its inception to a rapid development phase in just over a decade. This results in a certain gap in scale compared to mainstream developed countries. Still, the vast market space in China, coupled with an abundance of supportive policies and resources, leaves no doubt about its future growth potential. According to statistics from the Asset Management Association of China and CITIC

Securities, the management scale of quantitative private equity funds was 400 billion yuan at the end of 2019, reaching 700 billion yuan by the end of 2020, 1.18 trillion yuan by the end of 2021, and 1.5 trillion yuan by the end of 2022, currently accounting for only about 25% of private securities funds [1]. Supported by these figures, the conclusion that China's development space in quantitative investment is immense becomes more evident.

The widespread application of quantitative investment can bring numerous conveniences to financial markets. For example, it can efficiently improve market liquidity, ensure more effective and reasonable pricing in the long term, enhance the capital market's ability to withstand risk and enable the preservation of investors' wealth.

1.2.1 Development of Quantitative Investment Strategies in China

China's quantitative investment strategy models primarily include quantitative stock selection strategies, quantitative long strategies, index-enhancement strategies, market-neutral strategies, and CTA strategies. This stems from the fact that China's initiation of quantitative investment occurred later, and the market development is still not fully mature.

A quantitative stock selection strategy is a stock-picking strategy that favors investors, is supported by a vast amount of market data and examples, and utilizes various algorithmic tools. The key and core of this strategy lie in data mining and analysis. In recent years, China has successively introduced policies related to digital currency, greatly promoting the development of quantitative stock selection strategies. In 2014, the People's Bank of China formed a digital currency research team to study the operability of legal digital currency [2]; in May 2017, the Institute of Digital Money of the People's Bank of China was officially launched [2]; in November 2021, the State Council proposed to accelerate the pilot of legal digital currency [2]; in January 2022, the State Council issued a plan to support the use of digital currency in scenarios such as retail transactions, daily payments, and government services [2]. Different quantitative stock selection strategies have diverse characteristics. For instance, stock return prediction models and investment portfolios emphasize the effectiveness of artificial intelligence algorithms in identifying heterogeneous factors and other complex models; the SCDF multi-factor stock selection model based on feature reordering has a good effect on predicting stock returns; the TWSVM algorithm has a certain advantage in reducing data noise. Overall, quantitative stock selection strategies place greater emphasis on the application of artificial intelligence and machine algorithms in the field of financial innovation. They have a high dependency on various financial market data, with advantages in processing efficiency and breadth, thereby providing convenience to investors.

Quantitative long strategies in China are mainly applied as stock long strategies, divided into fundamental and multi-factor strategies. The fundamental strategies belong to value investment strategies, selecting investment products through relative or absolute valuation; multi-factor strategies, centered around statistics, undergo multiple evolutions to mine decision factors and construct a factor library.

Index-enhancement strategy is a strategy that exhibits stronger performance based on anchoring the index, which can reflect the final strategy's correlation with the anchored index by controlling the deviation from the index [3]. The evolution of the index-enhancement strategy has progressed from a traditional focus on macroeconomics to a modern integration with computer science, artificial intelligence, and machine learning. This transformation has ultimately achieved the purpose of predicting data and determining stock prices.

The principle of the market-neutral strategy involves simultaneous operations in long and short positions, buying relatively undervalued stocks while selling related overvalued stocks. This approach, through hedging, reduces the portfolio's systematic risk and achieves excess returns [4]. Market-neutral strategy is divided into two types: statistical arbitrage strategy and stock market-neutral strategy. The stock market-neutral strategy leverages the principles of market neutrality, constructing a low-systemic-risk investment portfolio through stock hedging relationships, thus maximizing returns with minimal risk. Statistical arbitrage strategy relies on historical information in the financial market, observing stock performance through the organization of historical and released information, and once deviations occur relative to historical data, buying or selling the stocks for arbitrage.

CTA strategy, referring to Commodity Trading Advisor strategy, involves investors entrusting funds to professional advisors. Advisors, based on their judgments of price trends in the commodity and stock index markets, use financial derivatives such as options to trade, thus gaining investment returns. The usual trading targets for this strategy are commodity futures and stock index futures [5].

1.2.2 Application of Quantification in Financial Markets

In the financial markets, quantification relies heavily on a plethora of financial data, including publicly available data on aspects like price, volume, and fundamentals. The stock market offers diverse data types for quantification in finance, such as market indicators, valuation metrics, and profitability indicators, detailed as follows:

Table 1: Types and Names of Indicators

Types	Name
Market Indicators	Annual Turnover Rate
	Annual Fluctuation Range
Valuation Indicators	Price-to-Earnings Ratio
	Price-to-Book Ratio
	Price-to-Sales Ratio
	Price-to-Cash Ratio
Profitability Indicators	Return on Net Assets
	Total Asset Return Rate
	Sales Profit Margin

Annual Turnover Rate: This reflects the frequency of buying and selling shares of stock. A higher value for this indicator signifies more active trading of the stock.
Annual Fluctuation Range: This indicates the degree of stock activity to a certain extent.
Price-to-Earnings Ratio: A ratio of company market value to profits, it is one of the most commonly used indicators to measure the level of stock prices.
Price-to-book ratio, Price-to-Sales Ratio, and Price-to-Cash Ratio: These ratios, comparing company market value to net assets, sales revenue, and cash flow (all vital financial indicators for publicly traded companies), can be used to assess the level of a company's stock price. If these ratios are relatively high, it indicates that the stock price level is comparatively high, reflecting strong investor confidence in the company or a certain level of bubble in the stock price.
Return on Net Assets: This demonstrates the rate of return on shareholders' investment.
Total Asset Return Rate: Calculated by including total liabilities in the denominator, this measures the return on a company's total assets, reflecting the overall operational capability of the company.
Sales Profit Margin: This gauges the profit level per unit of sales. If this indicator is low, the profit from the company's products or services is low, warranting consideration as to whether the pricing is too low or cost control is poor.

Quantitative investment involves organizing and analyzing massive amounts of data and extracting valuable information to construct investment strategy models. It employs artificial intelligence and machine learning tools to analyze investment choices, thereby avoiding subjective emotional judgments. In the process of quantitative investment, quantitative researchers refine and analyze data related to stock prices, generating effective signals. These signals are then aggregated to form the final investment strategy. The next step involves live trading operations. This process begins by obtaining the latest market data. The strategic model then evolves the data to form new investment instructions. Finally, these instructions are executed through algorithmic trading[3].
As China's financial market continues to develop, the application examples of quantitative finance are becoming increasingly rich and diverse. This includes various machine-learning-based quantitative investment strategies such as Risk Metrics (RM), Artificial Neural Networks (ANN), Adaptive Boosting, and Gradient Boosting Decision Trees (GBDT). These approaches represent the integration of cutting-edge technology with traditional financial analysis, facilitating more precise, efficient, and objective investment decision-making.
(1) Random Forest (RF): Random Forest is an ensemble algorithm developed based on decision tree algorithms. The procedure entails conducting a bootstrap random sampling of the entire training sample, forming multiple training sets containing original samples. Decision trees are drawn on the training set data by randomly selecting nn features for each tree, followed by feature screening to select the best features for node generation. Each tree can grow without pruning, culminating in a simple voting system where the classification with the highest number of votes is output[6].
(2) Artificial Neural Networks (ANN): In comparison to traditional methods like linear regression models and ARIMA[7], neural network machine learning algorithms perform better in stock price prediction. Research on the Japanese stock market returns using neural network algorithms has demonstrated that ANN's predictive ability surpasses traditional BP algorithms[8].
(3) Adaptive Boosting (AdaBoost): The AdaBoost model shows promising performance in predicting the risks associated with derivative financial instruments[9]. Kim et al. (2014)[18] incorporated emotional factors into the research framework, using the AdaBoost model to predict stock prices. Their findings indicate that the predictive capability of the AdaBoost model, enriched with emotional factors, is enhanced.
(4) Gradient Boosting Decision Trees (GBDT): This

model offers higher precision compared to traditional linear regression models that follow time series[10]. Additionally, the model's higher predictive accuracy and true classification nature are prominent advantages[11-12].

In the specific application of strategies, China has several, including event-driven strategies, hedging strategies, relative value arbitrage strategies, global macro strategies, and CTA fund management strategies.

Event-Driven Strategies: These strategies demand a high ability from investors to grasp financial market information and trends, necessitating decisions based on existing investment indicators. However, as China's financial market continues to open and information remains opaque, the market is not entirely efficient. The occurrence and dissemination efficiency of investment events significantly influence the market's development, thus affecting the pricing of financial assets[13]. Hedging Strategies: Based on hedging principles, these strategies establish long or short positions to reduce systemic risk and attain excess returns. When implementing stock hedging strategies, the choice of quantitative models and the coverage of systemic risk directly relate to the profit model of stock hedging. This pattern performs best in bear markets[14]. Relative Value Arbitrage Strategies: These strategies emphasize asset value, realizing arbitrage by buying undervalued assets and selling overvalued related assets. Specific applications include correlation arbitrage, volatility arbitrage, and statistical arbitrage. Global Macro Strategies: Investors must comprehensively analyze and study information in the macroeconomic market, seeking profitable development patterns and engaging in multidirectional leveraged investment transactions[13]. During economic recovery, active investment in financial markets is required; in overheated situations, investment in commodities; in stagnant phases, holding substantial cash resources is needed; in recession, investment in debt is possible[15]. CTA Fund Management Strategies: Most applications in China start with trend tracking and strategies, supplemented by reversal trading strategies. In algorithmic trading, these are combined with historical information to reflect the strategy's advantage from a probabilistic perspective.

1.3. Research Methodology

Two primary methodologies are implemented in this study.

Literature Research: This approach involves locating, selecting, analyzing, and comparing relevant literature, thereby enabling an in-depth exploration of the research subject. During the research process, the researcher actively reviews literature in various ways and systematically studies relevant theories by combining online and offline resources and information.

Quantitative Analysis: This method emphasizes data dependency more than qualitative analysis, employing mathematical and statistical models as a technical means of measuring and analyzing the research object. Researchers can reach conclusions through specific numerical values. Machine learning algorithms, such as random forests, fall within the scope of quantitative analysis.

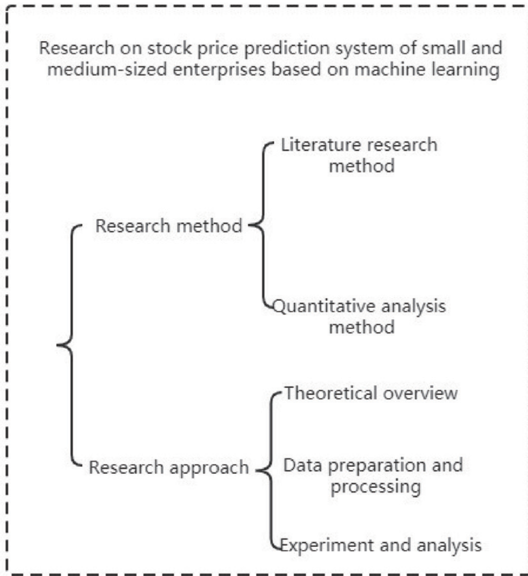

Fig. 1 Overall Flow Chart

2.1 Efficient Market Hypothesis

The Efficient Market Hypothesis (EMH) occupies a pivotal position in economics, encompassing monetary policy, investor behavior, financial development, and more. Arbitrage principles, the Capital Asset Pricing Model (CAPM), the Arbitrage Pricing Model (APM), and the Option Pricing Model form the theoretical bedrock of the EMH [16]. Adam Smith's assumption of rational individuals posits that profit maximization becomes their ultimate goal when investors enter the stock market. In this ideal state, market efficiency can be ensured. However, in real life, the stringent conditions of the "rational person" assumption are difficult to satisfy, affecting market efficiency. Hence, different schools of thought have provided varying interpretations and advancements in market efficiency over the past century.

In the early stages, economists adhered to the "Random Walk" theory, believing that investor behavior in financial markets was unordered and random and that stock returns followed a stochastic process. Under this assumption, common stock prices obeyed Brownian motion, rendering stock prices almost unpredictable. Yet, as the volume of data increased and technology advanced, researchers discovered that stock prices exhibit certain patterns in the long run. Building on this insight, economists proposed three forms of the Efficient Market Hypothesis: Weak-Form Efficiency, Semi-Strong Form Efficiency, and Strong-Form Efficiency. These classifications continue to influence contemporary financial strategy and investment approaches.

2.2 Moving Average Line Theory

Stock price analysis falls into two categories: qualitative analysis and quantitative analysis. Qualitative analysis focuses on the economic conditions and usually involves the international situation, commodity markets, political factors, and industry development. It requires researchers to grasp the macroeconomic environment comprehensively. The quantitative analysis relies on a vast amount of financial data. All techniques used to mine and analyze data are encompassed within quantitative analysis.

A particular method within quantitative analysis is the use of technical indicators, through which various stock market data are utilized to build models for predicting stock prices. These models analyze the stock's fundamental data, such as the opening price, highest price, and percentage change in stock price. A detailed representation of the data included in this analysis is illustrated in the figure below:

date	open	close	high	low	volume	code	close-open	high-low	pre-close	price_change	p_change
2015-01-05	8.168	8.688	9.068	7.998	6560835.0	000002	0.063663	0.133783	NaN	NaN	NaN
2015-01-06	8.378	8.138	8.768	7.828	3346346.0	000002	-0.028646	0.120082	8.688	-0.55	-6.330571
2015-01-07	8.038	8.008	8.278	7.778	5642051.0	000002	-0.003732	0.064284	8.138	-0.13	-1.597444
2015-01-08	8.098	7.368	8.148	7.238	2639394.0	000002	-0.090146	0.125725	8.008	-0.64	-7.992008
2015-01-09	7.318	7.228	7.998	7.068	3294584.0	000002	-0.012298	0.131579	7.368	-0.14	-1.900109

Figure 2: Basic Stock Data

The Moving Average (MA) indicator is a critical index for the trend of price movements, reflecting the strength and direction of prices over a certain period. The MA indicator is characterized by high stability and a small fluctuation amplitude. In actual trading, continuous limit up or limit down is not observed. The moving average line reacts when a price trend is formed, effectively identifying noise signals [17]. Technical indicator analysis generally presumes three prerequisites: the market is strongly efficient; future stock price trends change according to observed historical patterns; all factors in trading can be quantified [17]. The mathematical representation of the Moving Average is given as:

MA(N) is the moving average value over the past N trading days, and N is the number of trading days.

3.1 Calculation

The Moving Average (MA) is a statistical analysis method that averages security prices (indices) over a specific period. The average values at different times are then connected to form a moving average line used to observe the trend of security price changes. The MA theory is one of the most widely applied technical indicators today. It assists traders in confirming existing trends, discerning imminent trend cushions, and detecting overextended trends that are about to reverse. Its primary purpose is to help investors predict current and future market directions. As defined earlier, MA(N) represents the moving average value of the past N trading days, where N is the number of trading days.

When two moving averages (MA) of different lengths intersect, a situation may occur where the shorter moving average breaks upward through the longer one. Generally speaking, when the 5-day moving average (5MA) surpasses the 10-day moving average (10MA), this is referred to as a "golden cross." It signifies a potential short-term rise and significant price gains, suitable for entering long positions or exiting short ones. Conversely, a "death cross" happens when the 5MA falls below the 10MA, indicating a potential short-term decline and substantial price reduction, suitable for entering short positions or exiting long ones.

3.2 System Construction

The process of constructing the system, including some code snippets, is as follows:

Step 1 Obtaining Market Data from Tushare

```
import tushare as ts
import pandas as pd
import matplotlib
import matplotlib.pyplot as plt
import numpy as np
import talib
```

Figure 3: Database Retrieval

Step 2 Adjusting the Image, Reading the Data

```
matplotlib.rcParams['axes.unicode_minus'] = False
ts.set_token('token码')
pro = ts.pro_api()
df = pro.query('daily', ts_code='000001.SZ', start_date='20140801', end_date='20200810')
```

Figure 4: Data Reading

Step 3 Setting the Date Index, Closing Stock Price, and Percentage Change in Stock Price

```
df = df.sort_index()
df.index = pd.to_datetime(df.trade_date, format='%Y-%m-%d')

close = df.close

ret = df.change / df.close
```

Figure 5: Parameter Setting

Step 4 Calculating with an 8-Day Moving Average as the Target

```
df1 = talib.MA(np.array(close), timeperiod=8)
close8 = df1
```

Figure 6: Function Invocation for Calculation

Step 5 Processing Signals, Building a Time Series

```
SmaSignal = pd.Series(0, index=close.index)
s = 0
k = 0
for i in range(8, len(close)):
    if all([close[i] > close8[i], close[i - 1] < close10[i - 1]]):
        SmaSignal[i] = 1
        k += 1
    elif all([close[i] < close10[i], close[i - 1] > close10[i - 1]]):
SmaSignal[i] = 0
        k += 1
    else:
        SmaSignal[i] = SmaSignal[i - 1]
    s = s + SmaSignal[i]
    ds = len(close) - s
SmaTrade = SmaSignal.shift(1).dropna()
SmaRet = ret * SmaTrade.dropna()
```

Figure 7: Signal Processing

Step 6 Calculating the Cumulative Returns of Stocks and Strategy

```
cumStock = np.cumprod(1 + ret[SmaRet.index[0:]]) - 1
cumTrade = np.cumprod(1 + SmaRet) - 1
plt.plot(cumTrade, label="cumTrade", color='r', linestyle=':')
plt.plot(cumStock, label="cumStock", color='k')
plt.title("Cumulative stock return and 8-day average strategy return")
plt.legend()
print("Portfolio annualized return: {}".format(cumTrade[-2] * 250 / len(close)))
print("Operation {} times, short days {}, long days {}".format(k, s, ds))
```

Figure 8: Results

This thesis delves into the relevant theories of quantitative finance in economics, analyzing cases of quantitative investment in the financial market. It employs Python to build an investment decision-making system encompassing economic theory, data analysis, and computer technology. Ultimately, it offers a practical value platform to enterprises and individuals, providing references for investors' final decisions. In today's rapidly advancing computer science society, quantitative finance will play an increasingly significant role in investment. It is the hope that this thesis may bring assistance and insight to fellow researchers. The opportunity to contribute in this manner is indeed an honor.

References

[1] Zhao, Y. The quantitative market in China is still in the stage of rapid development [N]. Financial Times, 2023-07-28(007).

[2] Lin Sheng, Qi Ke, Wei Kaicong, et al. Research summary of machine learning in stock price forecasting [J]. China Economist, 2019, No.361(03):71-73+78.

[3] Liu, H., Huang Y., Yao Z. Research on quantitative stock selection strategy of digital currency stock market [J]. Marketing Management Review, 2023 (02): 34-36. DOI: 10.19932/J.CNKI.22-1256/F.2023.02.034.

[4] Nie, R. Design of exponential enhancement quantization strategy based on machine learning [D]. Chinese Academy of International Trade and Economic Cooperation, 2022. doi:

10.27054/d.cnki.ggjms.20022.200100001006
[5] Ye, B. Analysis of the Application of Market Neutral Strategy in China [J]. China Economic & Trade Herald, 2012(32):46-47. Guo, C. CTA quantitative trading strategy based on Boolean line research [D]. Huazhong University of Science and Technology, 2022. The DOI: 10.27157 /, dc nki. Ghzku. 2022.000794.
[6] Ma, M. Research on Credit Risk Assessment of Small and Medium-Sized Enterprises Based on Random Forest Method [J]. Special Zone Economy, 2023, No.408(01):141-144.
[7] Mostafa, M. M. Forecasting stock exchange movements using neural networks: tems with applications,2010,37(9):6302-6309
[8] Qiu, M. Song, Y., Akagi, F. Application of artificial neural network for predicting stock market returns The case of the Japanese stock market[J]. Chaos, Solitons & Fractals,2016,85:1-7.
[9] Zhang J., Sun Y. Derivative financial instrument risk prediction based on combination of AdaBoost algorithm [J]. Statistics & Decision, 2012, No. 355 (7): 41-44, DOI: 10.13546 / j.carol carroll nki tjyjc. 2012.07.006.
[10] Hassan, M. R., Nath, B., Kirley, M.A fusion model of HMM,ANN and GA for stock market forecacasting [J].Expert systems with Applications,2007,33(1):171-180.
[11] Zhang X., Wei Z., Yang T. The Application of GBDT Combination Model in Stock Forecasting [J]. Journal of Hainan Normal University (Natural Science), 2018,31(01):73-80.
[12] Liu P., Zeng Z. Financial credit risk assessment of pharmaceutical supply chain based on GBDT [J]. Friends of Accounting, 2021, No.664(16):24-31.
[13] Li L. Construction and application of multi-factor quantitative investment strategy [D]. Chengdu: University of Electronie Science and Technology of China, 2022.
[14] Li Z. Application of quantitative investment in China's investment market [J]. Capital, 2022(11):22-24.
[15] Zeng Q. Research on the risk of quantitative investment in capital market and risk control strategy [J]. Capital, 2022(6):4-6.
[16] Yu M. Market validity theory research review [J]. Liaoning Economy, 2020 (6) : 56-57, DOI: 10.14041 / j.carol carroll nki. 1003-4617.2020.06.023.
[17] Yu Y. Research on quantitative trading strategy of bonds based on moving average technical indicators [J]. Economic & Trade Update, 2022,19 (09): 40-45.doi: 10.19463/j.cnki. sdjm.2022.09.032.