Scholar Publishing Group
International Journal of Big Data Intelligent Technology
https://doi.org/10.38007/IJBDIT.2026.070114
ISSN 2790-0932 Vol. 7, Issue 1: 114-122

A High-Level Feature User Behavior Prediction Intelligent System Integrating Computer Technology with Financial Scenarios

Yongqiang Ma

School of Computer and Big Data, Jining Normal University, Ulanqab 012000, Inner Mongolia, China

Keywords: High order feature quantification, event driven information extraction, causal analysis, user behavior prediction, intelligent financial system

Abstract: This study focuses on the evolution of AI in the financial field, from AlphaGo to ChatGPT, as technology drives AI to penetrate into complex group game scenarios. Financial securities investment has become a core scenario for AI implementation due to its group game nature, but the existing evaluation system has shortcomings such as surface data listing, lack of time series and individual stock quantification, and lack of diagnosis of irrational behavior. To this end, a high-order feature user behavior prediction intelligent system integrating computer technology and financial scenarios has been constructed, which integrates four modules: structured data modeling, event information extraction, causal analysis, and feedback generation. The system quantifies operational behavior through SEA grid splitting and designs a attribution model for trend/stock selection/game ability; Adopting MRC framework for event extraction without trigger words and joint training scheme to extract event information; Quantify event impact using event analysis method and generate template comments based on user preferences. Research has found that the system has achieved a breakthrough in the quantitative system from the overall to individual stocks, filling the gap in causal attribution, promoting investors to transform towards a "data model dual drive" approach, demonstrating high-precision predictive capabilities (mAP verification) in financial digital management, strengthening decentralized architecture and transaction traceability functions, providing technical support for the digital upgrading of the financial industry, becoming a typical practice of technology integration in financial scenarios, laying the key technical foundation for future intelligent financial systems, ultimately optimizing decision-making efficiency, and promoting long-term stable development of the financial market.

1. Introduction

The evolution trajectory of artificial intelligence in the financial field has extended from AlphaGo breaking through the boundary of complete information games in 2016 to ChatGPT triggering a wave of generative AI in 2023. Technological breakthroughs continue to push AI to expand from simple rule scenarios to group games[1] with incomplete information and complex rules. Financial securities investment [2] has become a strategic high ground for the landing of AI technology due to its group game attributes intertwined with massive users, multi variety adjustment rules, and irrational behavior. However, there are significant shortcomings in the existing investment evaluation system: most attribution analyses[3] only focus on listing surface level data such as returns and risk indices, lacking a temporal decomposition of the investment process and quantification of the impact at the individual stock level; Investors often experience decision-making biases due to irrational behaviors such as herd behavior and loss aversion, and there is an urgent need for a closed-loop diagnostic tool that goes from "why profit and loss" to "how to improve". The aim of this study is to construct a high-order feature user behavior prediction intelligent system that integrates computer technology with financial scenarios. The core motivation comes from three points: breaking through the limitations of traditional investment behavior analysis models (such as the high complexity of Markowitz portfolio theory, strict CAPM assumptions, and Fama French multi factor model ignoring investor game ability), and constructing a quantitative system from overall evaluation to individual stock evaluation; Resolve the analysis gap of the impact of financial events on asset prices by using event extraction techniques to extract key events from news texts and quantify their effects; Realize personalized presentation of evaluation results, and use natural language comments to help investors understand investment styles, reasons for returns, and improvement directions. The research path of the system is divided into three progressive dimensions: high-order feature quantification modeling, constructing an evaluation system based on structured data that covers overall indicators such as profitability, risk control, asset liquidity, as well as individual securities return contributions, position adjustment capabilities, and other stock dimensions. Innovatively using SEA grid splitting method to quantify operational behavior, designing performance attribution models such as trend judgment ability, stock selection ability, and game ability to support position allocation decisions; Event driven information extraction, developing a trigger free word event extraction method based on MRC framework, and combining viewpoint based machine reading understanding to construct a joint training scheme of "text classification+machine reading understanding" and "named entity recognition+viewpoint based machine reading understanding", achieving accurate extraction of financial event types and event subjects; Causal analysis and personalized feedback are used to quantify the impact of specific events on securities prices through event analysis. Combined with user preferences, rule-based template reviews are generated to form a complete closed loop of "evaluation attribution feedback", ultimately outputting a natural language form of investment behavior diagnosis report. The system adopts a modular design, integrating four major modules: structured data modeling, event information extraction, causal analysis, and feedback generation. It not only fills the gap in in-depth analysis and causal attribution at the individual stock level in the existing evaluation system, but also promotes investors' transformation from an "experience driven" to a "data model dual driven" rational investment model through high-order feature extraction and intelligent feedback mechanisms, ultimately assisting in the long-term stable development of the financial market.

2. Correlation theory

2.1. User behavior prediction and intelligent decision-making architecture in financial service systems

In the data-driven upgrade process of the financial industry, financial service systems built on computer technology are centered around user behavior prediction, which drives intelligent decision-making and optimizes service experience through precise prediction. The system architecture consists of four modules: infrastructure, behavior monitoring, task management, and intelligent services, among which user financial behavior prediction is the key technical support.

The infrastructure module is data centric, relying on databases such as MySQL [4] to achieve efficient storage and transmission, and combined with cloud services such as Docker to complete report data management. It supports behavior prediction, intelligent decision-making, and outbound calls from the bottom layer, fully tapping into the value of data to drive service optimization. Securities investment data needs to be structured into a unified format, covering user position history and financial market data. User investment data includes time series operation records of stocks, bonds, reverse repos, and other varieties, such as fund transfers in and out, securities buying and selling, and dividend repayment. It has been obtained through simulation operations and stored after desensitization. As of April 2023, more than 380000 records have been accumulated; In the data processing stage, daily position and asset market value data should be generated based on transaction records, such as total market value, cash balance, and proportion of securities market value. At the same time, specific details such as the number of positions, sellable quantity, closing price, and market value of securities should be extracted. Financial market data is obtained from public financial data service providers through web crawling technology, standardized and stored, including key indicators such as daily securities opening prices, closing prices, and trading volumes. After structured definition of all data, it provides a foundation for quantitative evaluation of investment behavior, construction of performance attribution models, and visual display, ultimately achieving a complete analysis chain from the decomposition of investment history time series to the quantification of individual stock level impact, supporting intelligent decision-making and service optimization.

2.2. Functional analysis of behavior monitoring and task management module

The behavior monitoring module[5]focuses on heterogeneous data fusion processing, generates user behavior attributes by integrating multi-source financial data, and stores massive behavior information to build a lightweight buried point foundation, providing data support for intelligent decision-making; Simultaneously supporting the execution of diverse financial data processing tasks under specific program constraints. The task management module revolves around data capture, supervision and control, and customer dialing: real-time data capture obtains user financial behavior information from the behavior monitoring module, generates tasks, and intelligently assigns them to financial business personnel; Task monitoring provides diversified functions such as redirection and deletion; One click customer dialing enables the import and outbound of agent data, improving outbound calling efficiency while recording key information such as customer needs, follow-up time, and conversation skills. This effectively increases the success rate and avoids the risk of account suspension and complaints caused by repeated outbound calls, forming a complete closed loop from data collection to business implementation.

3. Research method

3.1. The core dimensions and implementation methods of quantitative evaluation of user investment

Quantitative evaluation of user investment requires comprehensive analysis of multidimensional indicators: yield evaluation covers absolute return, annualized return, and holding period return, etc. By comparing market benchmarks (such as indices) or other investor performance, the relative level of investment performance can be clarified; Risk assessment focuses on indicators such as volatility, maximum drawdown, and Sharpe ratio to measure the balance between returns and risks; Liquidity assessment reflects asset adjustment ability and trading efficiency through indicators such as turnover rate and bid ask spread; Diversify the evaluation and analysis of the proportion of assets such as stocks, bonds, and gold, and optimize the risk diversification effect; Performance attribution delves into the contribution of each asset, industry, or individual stock to earnings. The overall evaluation integrates the three elements of return, risk, and liquidity, and combines models such as time refined percentage method (TRR) to map quantitative values to 0-100 points, supporting comparative analysis between investors and the market, and promoting rational investment decisions. Individual evaluation focuses on a single held security and is conducted from two dimensions: security attributes (such as price increase and risk ranking) and investor perspectives (such as position adjustment ability and simple operation return rate). Using methods such as time cost weighted return (WWR) based on SEA grid, the adjustment return and operation win rate are accurately calculated to assist investors in optimizing their holding strategies and trading timing, ultimately forming a complete closed loop of "evaluation attribution feedback" to promote continuous improvement of investment ability.

3.2. Attribution model empowers high-order behavior prediction system

The core goal of building a user investment performance attribution model is to comprehensively analyze investors' abilities and investment status through quantitative evaluation, provide accurate decision support for investors, investment advisors, and institutions, and enhance the transparency and fairness of the securities market. This model focuses on quantitative analysis of three major investment strategies: trend investing measures the macro allocation ability of stock, bond, and gold positions through the ability to judge trends, and calculates the overall performance during the period based on the weighted increase of various index varieties; Value investing evaluates the ability to select high-quality securities and industries through stock selection ability (comparing stock holding returns with indices) and market selection ability (matching industry position ratios with index group gains); Game based investment quantifies the timing and control level of position adjustment through game ability, and compares the gap between expected and actual returns. In specific implementation, an overall evaluation framework is constructed using indicators such as time weighted return rate, maximum drawdown, and liquidity ranking. The SEA grid splitting method is combined to achieve refined analysis of individual securities' return rate, position adjustment ability, and simple operational behaviors (such as win rate and most profitable/least profitable operations). Ultimately, through visual means, the changes in user market value, quantitative indicator scores, performance attribution dimensions, and operational behavior decomposition are presented, forming a complete closed loop of "evaluation attribution feedback", supporting the optimization of investment strategies and the improvement of rational decision-making capabilities in high-order feature user behavior prediction intelligent systems.

3.3. Application verification and summary of long-term financial time series prediction models

In financial scenarios, event extraction technology [6]provides core data support for investor decision-making by identifying key event types and subject information of securities related companies. Traditional methods rely on trigger labeled datasets, which suffer from high annotation costs and low efficiency in trigger free scenarios. This article proposes a trigger free financial event extraction framework that integrates machine reading comprehension and ensemble learning, achieving joint optimization of event type recognition and subject extraction. The framework first defines eight types of financial events (equity changes, personnel changes, mergers and acquisitions, business difficulties, asset risks, executive accidents, violations and defaults, and illegal crimes), and based on the BERT classification model, captures the global semantics of the text through a pre trained language model, combines multi head self attention mechanism to learn long-distance dependency relationships, and finally uses a fully connected layer and sigmoid activation function to output the probability distribution of event types, achieving sentence level event type recognition; Simultaneously introducing FGM and PGD adversarial training to enhance the model's generalization ability to noise through small perturbations. On the basis of event type recognition, the machine reading comprehension[7] (MRC) framework is used to extract specific event subjects: the event type subclasses are merged into problem descriptions (such as "equity changes" corresponding to "stock issuance, equity transfer, stock repurchase..."), concatenated with the original text and input into BERT. The problem and text are separated by [SEP] markers, and the output vector is used to locate the starting and ending positions of the event subject words through the Span binary classification task. The final subject words are filtered and output based on thresholds, supporting nested entity recognition and focusing on the subject information associated with event types. The system adopts a pipeline architecture to integrate event classification and subject extraction tasks, achieves information exchange through sharing pre trained model parameters, and applies ensemble learning thinking for joint training, retaining the advantage of independent optimization of subtasks and capturing inter task dependencies through shared encoders. Adversarial training further enhances the stability of the model in financial text noise scenarios, forming a highly generalized event extraction intelligent system that fits the human reading logic of "word to sentence" in financial scenarios, providing a structured event data foundation for high-order user behavior prediction, and promoting the deep integration of computer technology in financial decision support.

4. Results and discussion

4.1. Financial event extraction dataset and experimental analysis

The financial text corpus constructed in this article includes two parts: annotated and unlabeled. The annotated data is sourced from the 2021 and 2020 financial event extraction tasks of the International Conference on Knowledge Graph and Semantic Computing [8](CCKS). The format is unified into four columns: event number, news text, event type, and event subject. Examples include "business difficulties", "mergers and acquisitions", and other types. For example, the news with the number 2045824 involves a "business difficulties" event, and the subject is "Huijin Shares". A total of 428739 samples of unlabeled data were collected from stock exchanges and trading platforms, covering the period from September 1, 2020 to September 1, 2022, with only news headlines retained. Examples include "December PMI Data Review: Industrial Product Prices and Employment Downward Dragging PMI".The experiment uses Ubuntu 18.04 system, based on Python 3.10.0 and PyTorch 2.0.0 framework. The data is divided into a training set and a validation set at a ratio of 9:1, with a total of 54756 samples covering 8 types of events such as stock changes, personnel changes, mergers and acquisitions. The specific distribution is shown in Table 1 below:

Table1 Event Type Distribution

Event Type	Sample Count	Event Type	Sample Count
Share Change	4,752	Asset Risk	4,381
Personnel Change	5,068	Executive Incident	4,176
M&A Restructuring	6,170	Breach/Violation	9,580
Business Difficulty	5,356	Criminal Offenses	10,851
Other Types	4,422	-	-

The pre trained model [9] uses FinBERT, with parameter configurations of 12 network layers, 768 hidden layer vector lengths, and 12 heads of multi head attention mechanism. The hyperparameter settings for the BERT-MRC-EE model based on the MRC framework include: 20 training iterations, 16 batch samples, and a maximum input sample length of 200. The evaluation indicators adopt weighted F1 value and micro average F1 value, and the calculation formula involves a weighted combination of precision, recall, and F1 value. The event subject evaluation also includes independent precision, recall, and F1 value. As shown in Table 2

Table 2 Comparison Results of Financial Event Extraction Experiments

Model	P_micro	R_micro	F1_micro
BERTBiLSTMCRF	85.97	87.22	86.59
BERTMultiPointer	90.88	92.07	91.47
BERTMRCEE+PGD	94.14	95.67	94.90
BERTOQMRCEE+PGD	94.28	95.39	94.83

The experimental results showed that after adding PGD adversarial training, the micro average F1 value of the BERT-MRC-EE model increased to 94.90%, which was better than mainstream models such as BERT BiLSTM CRF (86.59%) and BERT Multi Pointer (91.47%). After model fusion, the joint training model achieved the best micro average F1 value of 94.97%. Based on a multi task joint extraction model, a total of 68396 financial event data were extracted, covering text information sources, event types, and event subject word sets, providing core data support for building a user behavior prediction intelligent system that integrates high-order features.

4.2. Model experiment

Event analysis is a core method in the financial field for studying the impact of specific events on asset prices, with the core assumption that information is quickly and completely reflected in prices. This method assists investment decision-making, risk management, and market efficiency evaluation by quantifying event shocks. The specific implementation steps include: defining the type of research event (such as corporate mergers and acquisitions, policy changes, etc.), determining the event window (centered on event day t, considering news lag and selecting [t-5, t+10] for a total of 15 trading days), and estimating the window (located before the event window, selecting [t-200, t-6] for a total of 195 trading days to estimate normal returns); Calculate the normal rate of return using the Capital Asset Pricing Model (CAPM), with the formula$R_{i,t} = a_i + \beta_i R_{m,t} + \varepsilon_{i,t}$, where $R_{i,t}$is the increase of security i on day t, Rm, t is the increase of the market index, and a_i and β_i are estimated using the least squares method; The abnormal return rate (AR) is defined as the difference between the actual return rate and the expected return rate, i.e. $AR_{i,t} = R_{i,t} - \widehat{R_{i,t}}$ The cumulative abnormal return rate (CAAR) is the accumulation of the average abnormal return rate within the event window, i.e. $CAAR(t_1, t_2) = \sum_{t=t_1}^{t_2} AAR_t$ where $AAR_t = \frac{1}{n}\sum_{i=1}^{n} AR_{i,t}$(n is the number of securities in the sample). Finally, significance was verified through hypothesis testing. AR_t follows a normal distribution, null hypothesis H_0: μ=0 (event has no effect), alternative hypothesis H : μ ≠ 0, and the test statistic follows a t-distribution. At a confidence level of α=0.05, it was determined whether to reject the null hypothesis.

4.3. Effect analysis

Computer technology conducted in-depth analysis on 68396 financial event data from September 2020 to September 2022 through a multi task joint extraction model, and constructed a user behavior prediction intelligent system that integrates high-order features. The system divides the data into 8 subsets according to event types, each corresponding to a sample space of abnormal return rate (AAR) and cumulative abnormal return rate (CAAR), and reveals the impact of different events on asset prices through quantitative features: in mergers and acquisitions, AAR first rises and then falls within the event window, reaching a peak of 0.485 during t=-1 to t=1, and CAAR returns to market expectations after reaching 1.628 at t=1, reflecting that initial cash out mergers and acquisitions drive stock prices up but strategic risks lead to a pullback; In the executive accident, AAR dropped sharply to -1.932 at t=0, CAAR reached -4.527 at t=5, and asset prices continued to decline for 10 days, reflecting investors' delayed but persistent response to negative information; In the equity change event, the AAR was 0.106 at t=-1, rising to 0.594 at t=0, and the CAAR reached 1.221 at t=1. Within two days of the event, the CAAR increased sharply and returned to a normal distribution; In the personnel change event, AAR increased from 0.462 to 1.380 during t=-2 to t=0, CAAR reached 2.985 at t=0 and remained at 3.453 at t=10, indicating that investors were paying attention in advance and continued to be affected; In the illegal and criminal incident, the AAR was -1.549 at t=0, and the CAAR reached -3.293 at t=3. The stock price fell for three consecutive days and remained at -2.817 at t=10, reflecting the serious negative impact of public opinion risk; In the violation and breach events, the AAR increased from 0.175 to 0.400 during the period from t=-5 to t=-1, and the CAAR reached 1.229 at t=-1. After t=0, the AAR sharply dropped to -0.080, and the CAAR decreased to 0.637 at t=1, reflecting the early dividends of the event window and the bearish sentiment of investors after the news release; In business difficulties, the confidence levels of AAR and CAAR are generally not high, the null hypothesis is valid, and there is no significant impact due to investor differences; In the asset risk event, the AAR was -0.302 at t=-3 and the CAAR reached -1.897 at t=0. Investors perceived the risk in advance, leading to a decrease in prices and sustained decline before the news release. The system combines event analysis method with user position data to reverse search for the impact of impact events on profits and losses, such as significant profits caused by equity changes and significant losses caused by violations and defaults. Ultimately, personalized investment behavior analysis results are generated to achieve performance attribution and investment preference evaluation, fully demonstrating the core ability of computer technology to integrate high-order features and predict user behavior in financial scenarios.

5. Conclusion

The system achieves accurate prediction of user financial behavior [10] through multi-channel cross feature analysis algorithm, evaluates performance using mAP index, demonstrates high-precision prediction ability in financial digital management, and effectively promotes enterprise digital transformation. The continuous development of computer technology will promote the strengthening of decentralized architecture, tamper proof mechanisms, and transaction traceability functions in financial service systems, enhance security and operational efficiency within compliance frameworks, and provide technical support for the digital upgrading of the financial industry. The system deeply integrates computer technology and financial scenarios, achieves full chain intelligence through high-order feature extraction and cross validation mechanisms, optimizes decision-making efficiency while ensuring data integrity, and becomes a typical practice of technology integration in financial scenarios, laying a key technical foundation for the future development of intelligent financial systems.

References

[1] Jin Li. Performance Analysis of Efficient Microservice Architecture in the Financial Industry. Machine Learning Theory and Practice (2026), Vol. 6, Issue 1: 1-9.

[2] *Yixian Jiang. Performance Optimization and Improvement of Advertising Machine Learning Platform Based on Distributed Systems. International Journal of Big Data Intelligent Technology (2026), Vol. 7, Issue 1: 9-17*

[3] *Bukun Ren. Multimodal Learning Method for Cross-Modal Data Alignment and Retrieval. International Journal of Multimedia Computing (2026), Vol. 7, Issue 1: 1-8.*

[4] *Zhengle Wei. Research on Innovative Design of Financial Derivatives and Market Risk Management Strategies. International Journal of Social Sciences and Economic Management (2026), Vol. 7, Issue 1: 19-27*

[5] *Yuhan Zhou. Green Bonds and Sustainable Financing Models in Energy Finance. International Journal of Social Sciences and Economic Management (2026), Vol. 7, Issue 1: 28-35*

[6] *Yilin Fu. Research on the Application of Innovative Financial Technologies in Capital Market Risk Management. Socio-Economic Statistics Research (2026), Vol. 7, Issue 1: 1-9*

[7] *Linwei Wu. Data Visualization and Decision Support Analysis Based on Tableau. Socio-Economic Statistics Research (2026), Vol. 7, Issue 1: 10-18*

[8] *Xinran Tu. Resource Allocation Optimization and Cost Saving Analysis Based on Data Mining. International Journal of Business Management and Economics and Trade (2026), Vol. 7, Issue 1: 1-9*

[9] *Wang, C. (2026). Research on the Control of Uncertainty Risks in Investment Decision-making by Financial Modeling.*

[10] *Shuang Yuan. Research on Abnormal Detection and Transaction Risk Management Based on Machine Learning. International Journal of Social Sciences and Economic Management (2026), Vol. 7, Issue 1: 10-18*

[11] *Wang, Y. (2026). Research on Optimization of Neuromuscular Rehabilitation Program Based on Physiological Assessment. European Journal of AI, Computing & Informatics, 2(1), 21-30.*

[12] *Cai, Y. (2026). Design and Implementation of System Extensibility under High Concurrency Environment. International Journal of Engineering Advances, 3(1), 31-37.*

[13] *Liu, Y. (2026). The Application of Data-Driven Financial Risk Management in Multinational Enterprises. Economics and Management Innovation, 3(1), 20-26.*

[14] *Huang, J. (2026). Practice of Public Space Optimization and Functional Enhancement in Cultural Architecture. European Journal of Engineering and Technologies, 2(1), 9-21.*

[15] *Xu, D. (2026). AI-Driven Video Content Optimization Strategies for Immersive Media. European Journal of Engineering and Technologies, 2(1), 1-8.*

[16] *Qi, Y. (2026). AI Driven Payment System Security Improvement and User Privacy Protection Mechanism. Journal of Computer, Signal, and System Research, 3(1), 35-41.*

[17] *Sun, J. (2025). Research on Financial Systemic Risk Measurement Based on Investor Sentiment and Network Text Mining. Socio-Economic Statistics Research (2025), 6(2), 185-193.*

[18] *Lu, Z. (2025). Design and Practice of AI Intelligent Mentor System for DevOps Education. European Journal of Education Science, 1(3), 25-31.*

[19] *Zhang, X. (2025). Optimization and Implementation of Time Series Dimensionality Reduction Anti-fraud Model Integrating PCA and LSTM under the Federated Learning Framework. Procedia Computer Science, 262, 992-1001.*

[20] *Zhang, X. (2025, May). Automobile Finance Credit Fraud Risk Early Warning System based on Louvain Algorithm and XGBoost Model. In 2025 3rd International Conference on Data Science and Information System (ICDSIS) (pp. 1-7). IEEE.*

[21] *Dingyuan Liu. Measuring the Sensitivity of Local Skill Structures to AI Substitution Risks Based on Occupational Task Decomposition. Socio-Economic Statistics Research (2025), Vol. 6, Issue 2: 177-184*

[22] Chen, X. (2024, November). Cloud Storage User Behavior Analysis and Dynamic Replica Strategy Optimization Based on Improved RFM and Fuzzy Clustering. In International Conference on Cognitive based Information Processing and Applications (pp. 425-434). Singapore: Springer Nature Singapore.