CRYPTOLLA

Providing event-driven recommendations, intelligent reporting and personalized portfolio management based on artificial swarm intelligence and blockchain for individual and institutional cryptocurrency investors.

May 6, 2018

draft v2.0

Abstract

The CRYPTOLLA provides a intelligent set of tools and services including event-driven recommendations, intelligent reporting and personalized portfolio recommendation to the cryptocurrency investment activity, taking full advantage of artificial intelligence (AI) and knowledge graph (KG) technology in multi-source heterogeneous data processing and complex machine learning model contracting, while also designing a value generation and incentive mechanism through blockchain technology to enhance the application performance of structured data platforms and intelligent models. The CRYPTOLLA is a blockchain-based intelligent investment platform. It is powered by a swarm investment intelligence community to generate a more secure and reliable intelligent service for individuals and institutional investors of cryptocurrency assets, which effectively solves the problem of confusion in property rights of big data assets. Therefore, this design greatly stimulates intelligent innovation in the wealth management industry.

The complexity, variability, and long-term nature of cryptocurrency asset risk determine that only an swarm investment intelligent, event-driven and life-cycle risk monitoring tool can meet the needs of this type of asset investors. This is also the goal of our research and development of intelligent systems. In the uncertain environment, based on multi-source heterogeneous big data, we mainly adopt machine (deep) learning algorithms to train decision models that meet various needs of investors and potential investors, and ultimately develop a system for intelligent investment decisions in which knowledge graph technology and blockchain technology are also embedded. Specifically, from an application perspective, the project can be attributed to the development of four functional modules: asset risk life cycle monitoring, personalized information and portfolio recommendation, human-computer interaction intelligent investment advisor and blockchain-based swarm investment intelligent community.

The first three parts will be gradually developed in stages. The fourth part will run through the entire development process. Finally, a complete intelligent investment ecology based on the blockchain will be developed. Blockchain-based swarm investment intelligence community is mainly based on IPFS (InterPlanetary File System) and Ethereum smart contracts, which provides a decentralized investment intelligence platform for value exchange.

CONTENTS

I. EVENT-DRIVEN RECOMMENDATIONS

A. Cryptocurrency Related Hot Event Detecting & Profiling

Social media has become an important source of information for investment and material information can reach investors faster than ever before. Therefore, detecting hot events in real time becomes a vital function to satisfy the information needs of investors. In addition, comprehensive analysis and profiling for hotspot events have become an indispensable extension of the service quality of this function. In recent years, research on the detection of hot events based on real-time data flow in social media has become a hot topic in the field of social network research [1]. Some leading research results [2], [3] provide systematic solutions for realizing this function.

Cryptocurrency related hot events discovery and profiling functions are a comprehensive module. The first step is to build a real-time data stream of cryptocurrencies through collecting posts and microblogs which are published by the selected seed account pool in forums or social media platforms. The second step is performing data preprocessing and word vector representation, and then adopting a social media hot event detection algorithm to detect real-time hot events. In addition, while hot events detection results are being presented, the project will adopt relevant technologies for information visualization to design a three-dimensional full-angle interpretation of events for users. The specific aspects of the information visualization process are as follows:

- **Event tracking**
  After discovering hot events, the module uses some overlap stacked area maps to clearly track the history and latest developments.
- **Event real-time keyword cloud**
  While tracking hot events, this module applies word cloud technology to present the hot event keyword set in real time.
- **Network interactive**
  The event participant's account category information and related interactive information are displayed in this module.
- **Event-driven cryptocurrency value scoring**
  Event-driven cryptocurrency value scoring is designed to score the value of each cryptocurrecny in the asset pool generated by the event-driven cryptocurrency conceptual portfolio model.

According to the attributes of financial assets, hot events can bring investment opportunities or lead to asset risks. Therefore, the next modules will datamine the event-driven investment concepts and provide users with the automatically generated investment portfolio (the investment opportunities arising from the event). It will also extract risk factors and design monitoring and warning indicators (for asset impairment crisis triggered by events).

B. Event-driven Cryptocurrency Conceptual Portfolio

Hot events, especially those related to investment concepts, in the near or far future, will trigger price fluctuations of related assets. In the investment process, investors who understand the hot event associated portfolio earlier will be given the opportunity to gain positive returns. The models in this module involve three parts, hot events, investment concepts and portfolios. Correspondingly, these models include two layers of mapping, namely the mapping of hot events to investment concepts, and the mapping of investment concepts to portfolios.

The investment concepts is contained in a domain knowledge base build from cryptocurrency investment knowledge. Further, if the mapping relationship between the investment concept and the portfolio is determined meaning each investment concept has a portfolio corresponding to it, the matching of hotspot events and investment concepts can translate into the problem of questions and answers (QA) [4]. The problem of QA can be further abstracted into deep text matching problems [5], [6]. Therefore, relevant algorithms for deep text matching can be adopted to train the matching model of hot events and investment concepts.

Directly speaking, event-driven investment opportunity discovery is an intelligent system that integrates various models with different data objects such as social media data streams, hot events, investment concepts and portfolios. These models reflect the mapping between different data objects as an end-to-end machine (deep) learning algorithm framework that flows from raw data to investment strategy without the requirement for human intervention.

In the module, this project proposes two types of solutions. One is to use basic machine learning algorithms to directly map the relationship of events, invest concepts and cryptocurrency portfolios, and the other is to use deep learning algorithms to match the most relevant investment concepts. The two methods have similar concepts, but utilize different learning algorithms. Each solution of building a model is explained as follows:

1) Knowledge bases and data preprocessing.
    a) Word vector representation of hot event keywords (word2vec) [7].
    b) Word vector representation of related event keyword dictionary.
    c) Word vector representation of related concepts' description word set.
    d) Word vector representation of related cryptocurrencies' description.
2) Machine learning model training.
    a) Mapping learning system.
    b) Deep matching learning system.

Note that the model is a dynamic and scalable system, that is to say, as the amount of data increases, the prediction ability of the model will become stronger through an online learning mechanism.

C. Personalized Information Recommendation System

The goals of the blockchain design and application in this module are to apply the swarm intelligence for personalized information recommendation model training. Based on user preferences and information profiling features, recommendation system algorithms [8] are applied to design and train a personalized information recommendation model similar to the smart information recommendation software "Toutiao", to provide a personalized information service. The difference is that this module will focus on the area of cryptocurrencies and will design referral services based on in-depth and specialized domain knowledge. A more simple description can be the information profiling and user profiling process utilized in training the machine learning intelligent matching model to provide personalized recommendations. The two major modules included in the technical framework of the personalized information recommendation system are as follows:

1) User and item information.
    a) User information preferences.
        i) User browsing history.
        ii) User search log.
        iii) User account and attention information.
        iv) Customized cryptocurrency pool.
        v) Recommendation feedback log.
    b) Cryptocurrency information profiling.
        i) Topic distribution.
        ii) Entity relation extraction.
        iii) Summarization.
        iv) Keyword cloud.
        v) Sentiment tendency.
2) Recommendation engines.
    a) Collaborative filtering recommendation algorithm [9].
    b) Content-based recommendation algorithm [10].
    c) Social recommendation algorithm [11].
    d) Hybrid recommendation algorithm [12].

Rich data features and various basic recommendation algorithms will constitute an open source platform for training personalized information recommendation models. On this platform, users can design personalized recommendation systems with independent intellectual property rights.

D. Data Annotation & Data Object Profiling

Data annotation [13] has always been the bottleneck for big data related problems. Especially labeling tasks that require professional domain knowledge. Therefore, the use of blockchain technology provides a platform to utilize the swarm intelligence of experts who are able to accept labeling tasks. Similarly, data object profiling is also a complex task that requires swarm intelligence from domain experts to share solutions. Data object profiling is a concept that is similar to a user profiling [14] but more general. Data objects include users, information, cryptocurrencies, ico, and hot events. The data object profiling technologies developed in this project, namely user profiling, cryptocurrency profiling, information profiling and hot event detection and profiling, can effectively avoid the data storage confusion of the original data platform and greatly improve the utilization efficiency of the multi-source heterogeneous data.

a) Data Annotation: In general, when data preprocessing and algorithms have made great progress in building a machine learning model, if there are enough high-quality experimental samples, a highly accurate prediction model can be trained. High-quality training samples depend on high-quality data annotation. For example, advanced emotion analysis algorithms based on deep learning emerge in an endless stream [15]. When constructing a sentiment orientation analysis model for cryptocurrency information texts, those who have experience in investment annotation and are able to accurately annotate the emotional category of the information texts become the determinants of whether or not they can fully realize the practical application value of those algorithms. In addition, blockchain technology can effectively monitor labeling quality, create labeling motivation, and measure the labeling value of complex intellectual contributions.

b) Cryptocurrency & ICO Profiling: Apart from user profiling visualization, the whole profiling approach can be divided into three processes:

- **Multi-source Heterogeneous Data Acquisition**
  Three kinds of data sources are collected separately such as:
  - Official Website & Open Source
  - Web (Social) Media
  - Exchange Trading Data

- **Knowledge Extraction**
  Two stages of data preprocessing and predictive model training:
  - Data Pre-processing
  - Models
    * Text Mining
    * Deep Text Classification [16]
    * Natural Language Processing (NLP)

- **Profiling**
  The profiling stage includes three types analysis methods of the price or return impact factors for financial assets as follows:
  - Fundamental Analysis: The factors affecting the intrinsic value for cryptocurrencies or ICO.
  - Behavioral Analysis: The behavioral factors of investors who influence asset price movements based on behavioral economics.
  - Technical Analysis: The factors that are extracted from the momentum and price-volume of the underlying asset derived from exchange data.

Note that ICO profiling faces the dilemma of incomplete data sources and needs to find new sources of relevant data.

c) Cryptocurrency Information Profiling: The central task of information profiling is to label the categories based on topic and keyword models. This specific process is similar to that of a cryptocurrency.

d) User Profiling: User profiling focuses on analyzing investors' or potential investors' information preferences and risk-return preferences.

- **FinaMetrica questionnaire analysis**
  Analysing the investors' risk-return preference through the FinaMetrica questionnaire data [17].
- **Web log mining**
  Mining investors' information preference through cryptocurrencies and browsing history in online logs.

II. INTELLIGENT REPORTING

A. Custom Cryptocurrency Pool Management

Apart from information service capabilities, another function of the automation layer will be cryptocurrency asset portfolio management. In particular, the services provided include user portfolio return volatility assessment and the user investment risk preference assessment. Further, when it is found that the real-time ROI (return on investment) of the user is lower than their risk tolerance interval obtained from the questionnaire evaluation, real-time push notifications will be provided in advance to provide an automated service for managing user assets. It should be noted that this function must first be authorized by the user of the cryptocurrency asset management platform.

The management of a custom cryptocurrency pool is a function of simulatd asset trading [18], [19]. It can effectively improve the user's investment experience and capture the user's risk-return preferences. While enhancing user loyalty, it has accumulated a large amount of raw data that investors are concerned about. This provides raw data for the implementation of other module functions such as user profiling. The investment risk warning will be pushed to the user in real time, and the portfolio (custom cryptocurrency pool) analysis report will be regularly sent to the subscriber.

B. Cryptocurrency & ICO Credit Rating

The features of cryptocurrency & ICO profiling extraction is the data foundation for the development of this module's functions. These features directly or indirectly reflect the credit level of asset from different dimensions. Credit rating methods include traditional statistical models and machine learning algorithms.

The traditional statistical methods [20] are main linear models, which place more emphasis on finding the actual impact of each factor in computing. Through the factors (features) selected and the weights set by experts, the system automatically weight feature values of each profiling dimension to obtain a comprehensive score for the credit rating. In addition to the traditional statistical method, in this project, based on the accumulation of available training data, we also design machine learning models for forecasting credit levels [21]. Experts of different technical backgrounds, through the blockchain swarm intelligence community, invoke other people's or contribute their own credit rating models.

C. Asset Risk Holographic Detection and Reporting

This module design a risk detection and reporting function for the entire life cycle of the assets. To be specific, before and after purchasing asset, the risk can be systematically evaluated and dynamically monitored. This project uses text mining techniques and graph mining techniques [22], [23], such as correlation analysis [24] and outlier detection [25], to statically assess asset risk levels. Correspondingly, this module can also dynamically report the level of asset risk. Based on the event-driven cryptocurrency conceptual portfolio, a value (risk) scoring model for related assets is constructed.

Based on multi-source and heterogeneous big data, we apply data object profiling technology and knowledge graph technology to store data objects (information, users, cryptocurrencies, ICOs, events etc.) and their relationships in a graph database. This design idea of this project is open up part of the functions in the knowledge graph [26] data platform. Users of the blockchain community can call it to contribute their asset risk disposal experience for risk holographic detection and reporting.

D. Cryptocurrency Market Analysis Reports

Professional analysts write reports on cryptocurrencies, cryptocurrency portfolios, or entire virtual digital currency markets. In the blockchain community, the generation and quality of these analysis reports will be effectively supervised and the dissemination process will be clearly documented and ultimately provided to investors in the form of information.

E. Foundation for Knowledge Sharing

The premise of pooling swarm intelligence through blockchain technonlogy is to provide a foundation for knowledge sharing for contributors of investment intelligence. The shared knowledge includes raw data, training data sets, data object profiling features, and recommendation system algorithm libraries. This module will provide the initial shared knowledge for the swarm investment intelligent community.

III. PERSONALIZED PORTFOLIO

A. Personalized Asset Portfolio Recommendation System

In recent years, the application of artificial intelligence (AI) related technologies in the selection and recommendation of investment portfolios has become a research hotspot [27], [28]. The intellectual asset investment mentioned in this module mainly refers to the recommendation of a personalized asset allocation plan. Adopting the model design idea of ranking learning system and matching learning system (related investors' return-risk perference features) will realize the task of personalized recommendation. This module describes a complete process as follows, including data preprocessing, feature extraction, AI ranking prediction models and cryptocurrency personalized asset allocation. Note that many tasks in the data preprocessing model overlap with those in the data platform mentioned earlier.

1) **Data preprocessing**
   Clean the collected raw data and categorize it to the different requirements of the extraction feature.
2) **Feature engineering**
   Extracting the prediction target around the impact model using a deep understanding of domain knowledge.
3) **Customized AI recommendation model**
   Machine learning (deep learning) algorithms are used to train the ranking model to achieve a high-level model performance based on the relative prediction target.
4) **Cryptocurrency personalized asset allocation**
   Based on the output of the previous model, combined with the investment knowledge in the field of portfolio strategy, a personalized asset allocation plan is automatically generated.

In this section, the algorithm we use has two major advantages. First, setting the relative index as a objective function of the prediction model such as rankings of asset return rates will effectively controls the adverse effect of low signal-noise ratio and non-uniform sample distribution in financial data [29]. Second, a deep learning algorithm adopted will enable full captures of the time series features of financial data and improve the prediction accuracy of the model [30].

In summary, the personalized asset allocation system provides advanced intelligent services and builds a matching model based on users, assets, returns and risk assessment tripartite. The construction of a ranking forecasting model will function as the foundation for the overall model. This module service is different from the model of a simple forecast for absolute return in the past. The ranking prediction model will maintain the stability of the risk level while meeting the needs of users.

It should be pointed out that this system will cover the intelligent chat function of human-computer interaction by realistically replacing the artificial investment manager to meet the individual needs of users while reducing the cost of low-end manual investment to zero. This can also effectively improve user service quality.

B. Open Source Development

This module provides users with an open source development platform, including pre-processed data, custom algorithm libraries, and programming environments. Based on these resources, users can train models for specific tasks. Further, based on the blockchain design logic (guarantee the quality of the model), the user writes the trained model into the block, converting investment experience and wisdom into value blocks for other users in the community. Simultaneously, users can directly call other trained models through the blockchain-based investment intelligent community.

IV. SWARM INVESTMENT INTELLIGENCE COMMUNITY

In this section, we will focus on describing the design logic, application details, and technical framework of the blockchain in this project.

A. Swarm Investment Intelligence

A.1) Generative Process of Swarm Investment Intelligence: A core goal of this project is to use AI technology to provide intelligent services for cryptocurrency investors. Moreover, in order to better play the role of AI, blockchain technology is embedded to build a swarm investment intelligent community. It has some advantages that meeting the diversified needs of investors and improving the application of AI through the swarm investment intelligence. The generation process of the swarm investment intelligence is shown in the figure below Fig 1.
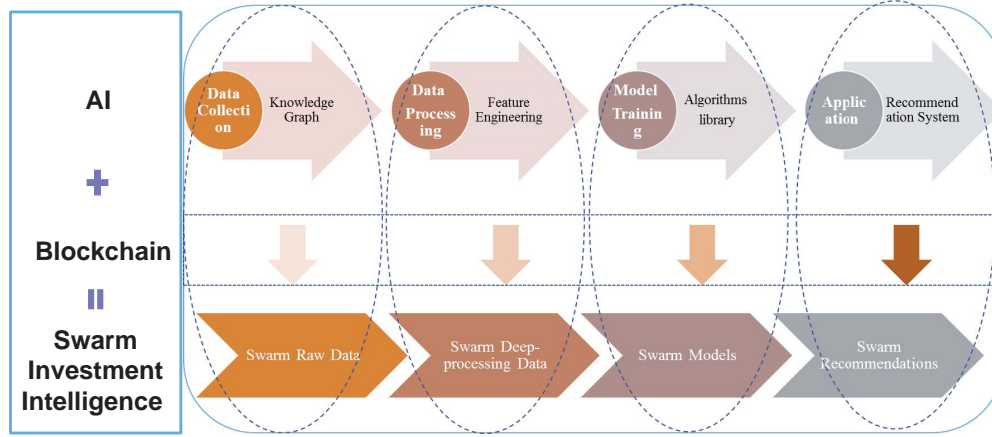

Fig. 1.   The generative process of swarm investment intelligence.

A.2) Standardized Items of Swarm Investment Intelligence: In the blockchain-based cryptocurrency intelligent investment platform framework, blockchain technology is embedded in the intelligent layer and data layer to introduce swarm intelligence, which can enhance the breadth and depth of the platform's intelligent investment services, and also satisfy the needs of investors in the largest range.

In the process of building an intelligent community, this platform extracts key tasks from the intelligent layer and data layer and standardizes these tasks for introduction to the cryptocurrency investment community. These tasks are pointed in the Fig 2 as follows:

- **Feature extraction**

  Based on modern asset pricing theory in the financial domain, participants extract factors that affect the return on investment and risk from pre-processed data sets, particularly (social) media data.

- **Data annotation**

  The labeling problem of training sets has always been the bottleneck of using machine learning algorithms to train prediction models in the financial field, especially those requiring professional domain knowledge
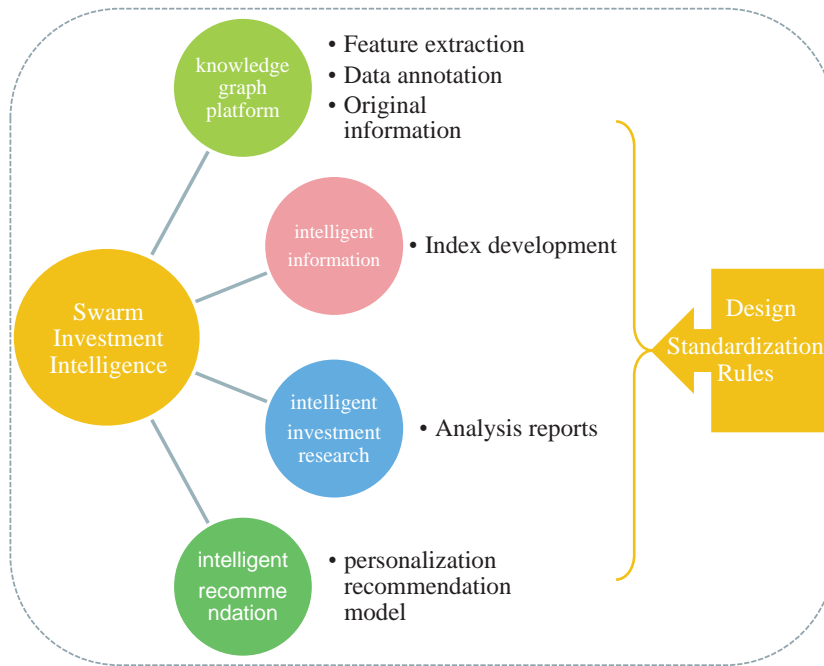
Fig. 2.   The standardized items of swarm investment intelligence in cryptolla designing framework.

to carry out the training set labeling. Therefore, data annotation will be an important manifestation of swarm intelligence.

- **Original information**
  The copyright issue of original information has always been the focus of debate in the press and publishing industry. The application of blockchain technology will make this problem easily solved, which ensure the author's intellectual property rights, and the resulting benefits. It will also promote the participants' creation enthusiasm.

- **Index development**
  Based on professional knowledge and practical experience, the contributors of community investment intelligence will select factors that affect returns or factors that reflect risk. Further, these experts use scoring methods and other methods to design indexes that accurately reflects the state of asset risk and returns.

- **Analysis reports**
  The analysis report refers to the analysis of the returns and risk status of assets in the past period of time and the forecast of price trends in the future. The methods of analysis and prediction will include text mining techniques and machine learning models which are applied to mine knowledge from multi-source heterogeneous big data.

- **Personalied recommendation models**
  Personalized asset allocation strategies and personalized information are what the platform will recommend. Noted that personalized information is the focus of the following technical framework, and the personalized asset allocation strategy will be the focus of the second development phase of the platform.

Based on these tasks, a swarm investment intelligence community is designed to stimulate the innovation abilities of professionalism investors.

B. Technology and Architecture of Blockchain-based Swarm Investment Intelligence Community

This module is used to upgrade the overall functionality of the system, which makes full use of blockchain technology to build a swarm investment intelligence community for professional investors. Through smart contracts, the investment intelligence of competent users will be transformed into value safely and conveniently. It also provides a decentralized trading platform for users to obtain swarm investment intelligence. Fig 3 shows the blockchain-based swarm investment Intelligence (CII) community, which contains three layers, including protocol layer, extension layer and application layer.
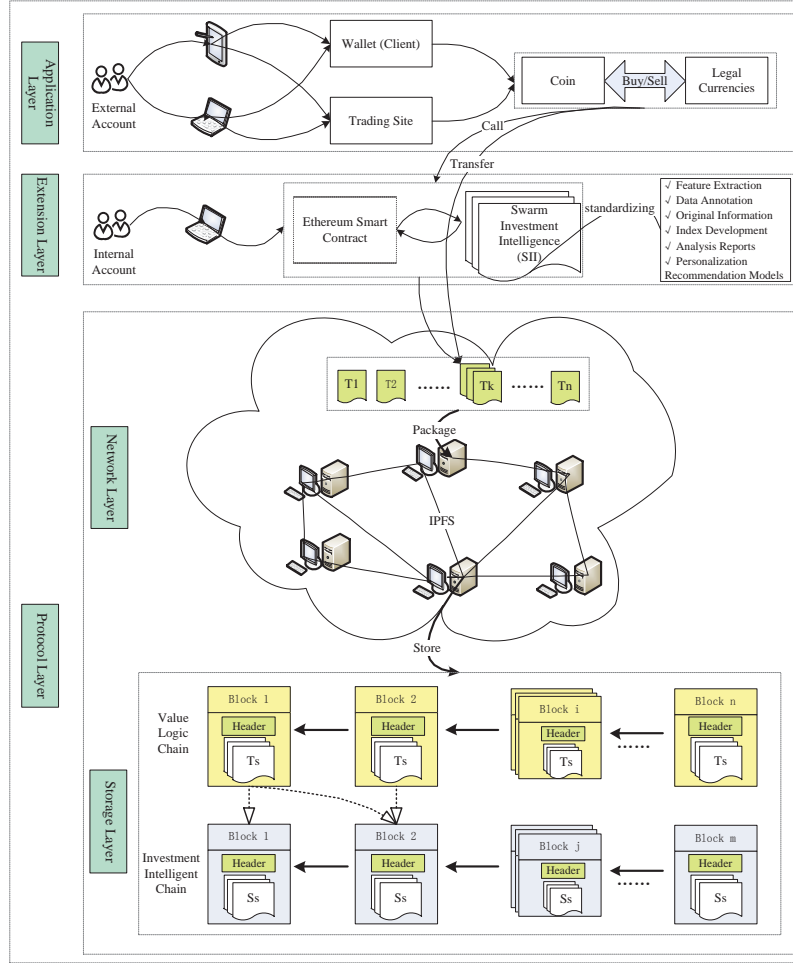

Fig. 3. The blockchain architecture of swarm investment intelligence community.

The blockchain design of this module is mainly implemented using Ethereum blockchains (smart contracts) and IPFS [31], [32]. Ethereum has a higher degree of abstraction and advantages in general data structures, inheritance, and encapsulation. IPFS is a point-to-point distribution file system, which connects all computing devices under the same file system. In other words, IPFS provides a high-throughput content-addressable block storage model, and content-addressed hyperlinks, which forms a generalized Merkle DAG (directed acyclic graph). In this data structure, we can build version control systems, blockchains, and even a permanent World Wide Web. IPFS combines distributed hash tables, incentive block switching, and a self-certified namespace. These three layers combine as whole to complete the transaction logic.

B.1) Protocol Layer: The protocol layer includes a network layer and a storage layer to realize different functions respectively and simultaneously. The network layer is a P2P network with distributed algorithm, encryption signatures and consensus algorithms to handle the generation of the blocks and verification.

The storage layer mainly contains a value logic chain and a investment intelligent chain according to their specific functions respectively. The cross-chain design effectively balances the relationship between decentralization and operational efficiency. In this layer, we will store and transfer the specific investment intelligence with the IPFS (InterPlanetary File System).

- **Value Logic Chain**
  The value logic chain gives liquidity to the value of the swarm investment intelligence (SII) and makes it possible to circulate, exchange, and trade the value. The value logic of SII includes the frequency of SII calls and the appreciation space of associated cryptocurrencies. Specifically, the higher the frequency with which data is called, the greater the value of the SII. The larger the appreciation space, the greater the value of the SII. As is shown in the Fig 4.
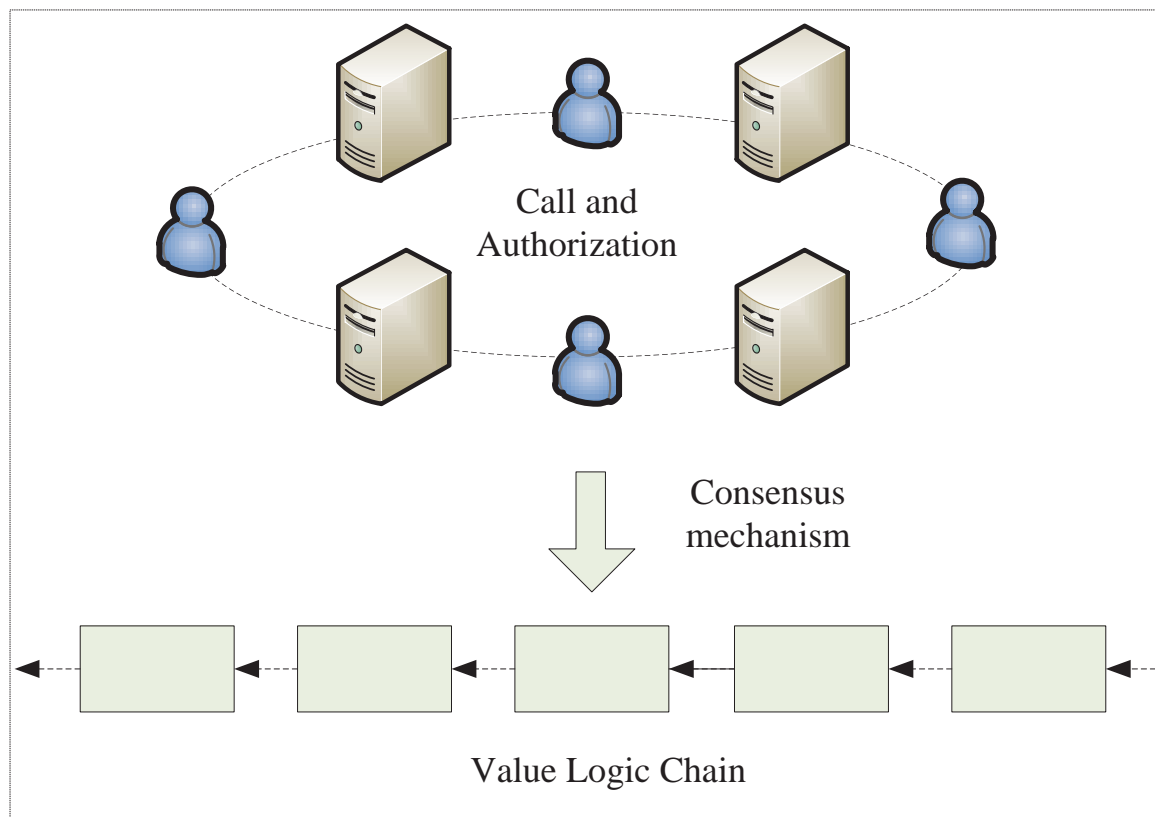

Fig. 4.   A flow chart about the value logic chain generation.

In the value logic chain, uses will play two roles: SII providers or SII consumers. Their value model is established in the value logic chain and stored in the investment intelligent chain. Tt also saves all transaction records of users.

The user value is relative stability, and the batch processing model doesn't require the value logic chain to be able to respond in real time. Therefore, we will adopt the traditional Proof of Work (PoW) as consensus mechanism to guarantee the security and fairness of the network.

- **Investment Intelligent Chain**

  The main task of the investment intelligent chain is to provide a complete and accurate SIIs. Through weak decentralized hierarchical management, data links format SII and store it on the chain.

  The investment intelligent chain manages users'SIIs through a two-tier structure, namely ordinary nodes and trusted nodes. The ordinary node is responsible for storing the user's formatted SIIs. A trust node is a higher-level node, that is, a trust node includes several ordinary nodes, has higher computational power and is responsible for node check. The ordinary nodes will store specific data according to the trust node's delegation. As is shown in the Fig 5.
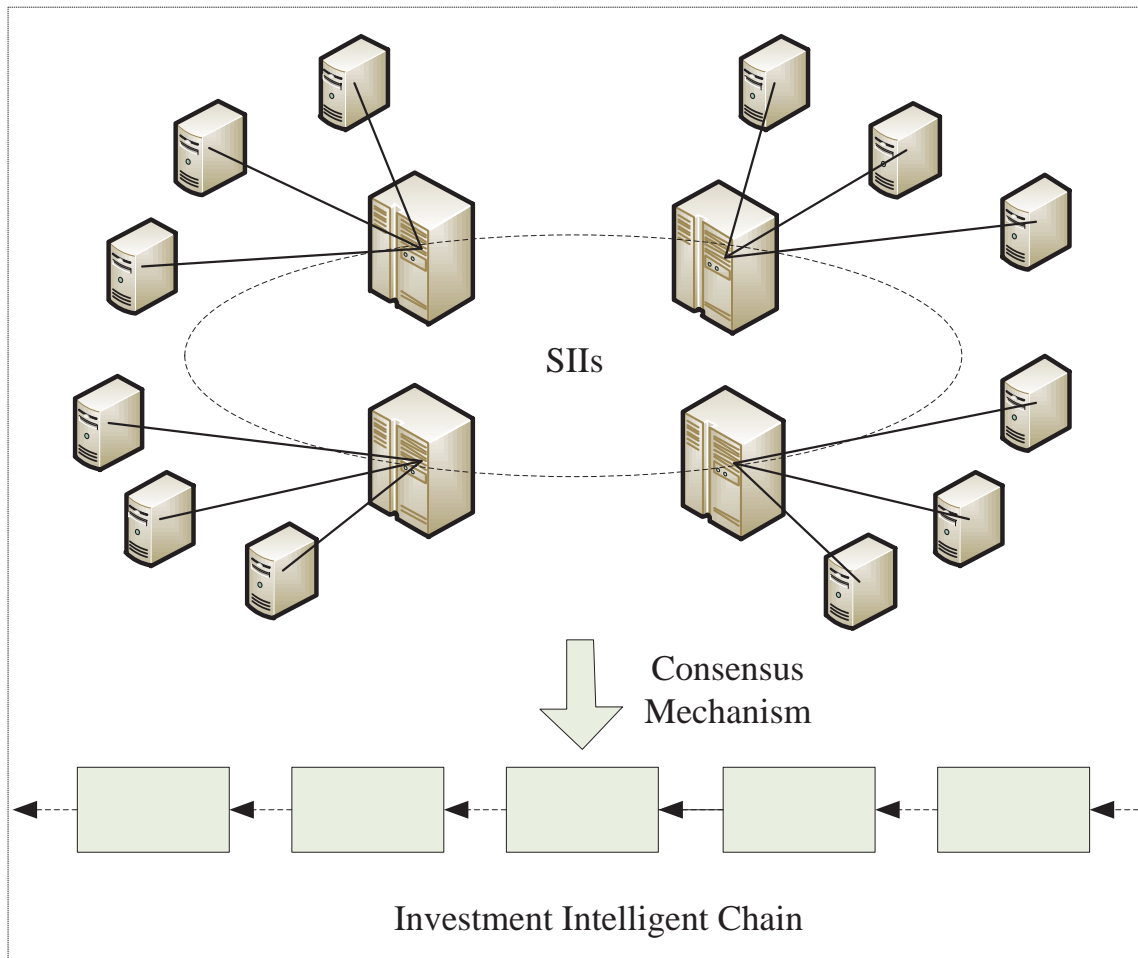

Fig. 5.   A flow chart about the investment intelligent chain generation.

Compared with the value logic chain, the investment intelligence chain needs faster transaction processing speed, shorter response time and more secure privacy protection mechanism, so we will introduce the Delegated Proof of Stake (DPOS). DPoS, as a variant of PoS, reduces the pressure on the network by lessoning the number of election nodes. It is a typical divide-and-conquer strategy. It divides all nodes into leaders and followers, and only after reaching a consensus among leaders can the followers be informed. This mechanism can effectively reduce network pressure without increasing computing resources, and it will have strong application value in excellent software implementation.

- **Data structures of accounts and blocks**

  In this section, we show the data structure details of the account and the block in the blockchain architecture. The internal or external account data structure contains some attributes to reflect itself. A series of transaction records about the proof of value are stored in the value logic chain. Each record includes a series of standardized attributes. The SIIs of the user are formatted and stored in the investment intelligence chain.
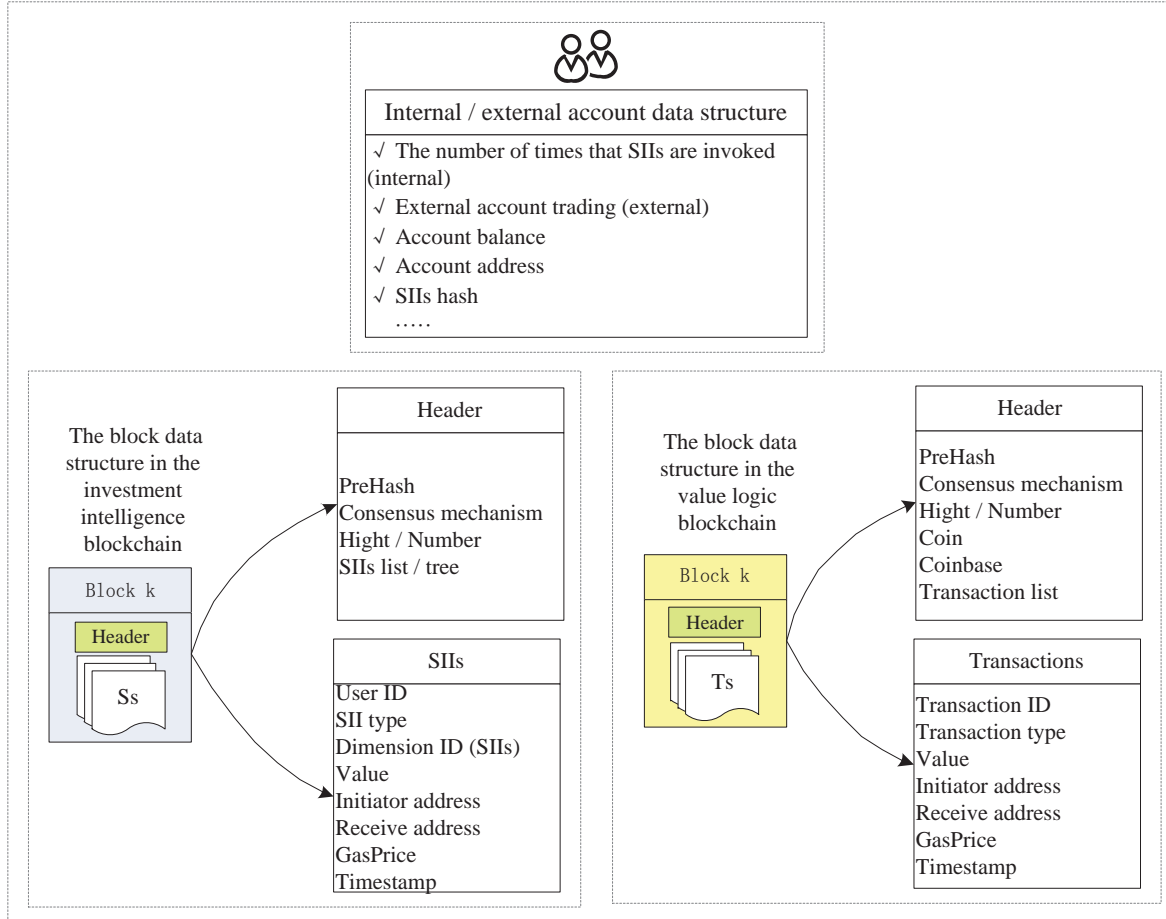

Fig. 6.   The data structures of accounts and blocks in blockchain architecture.

Each block in the blockchain is composed of two parts, the header and the main body. Fig 6 shows the block data structure information in different block chains and the information in the two account structures.

  B.2) Extension Layer: The extension layer is used to provide a series of smart contract APIs for users to generate all kinds of CIIs related to their own transaction logic. This flexible designing can enhance the robustness of the system, and satisfy the different requirements of the users' transactions.

Each smart contract is an irrevocable program that runs under the condition that the characteristic conditions are met. These smart contracts can be programmed with a pre-determined series of instructions, or they can be uploaded by an individual contract to an immutable and fully transparent virtual machine. Smart contracts can be quickly and cheaply determined and completely transparent. Similarly, logic can also be encoded into smart contracts in the form of Turing complete computer code.

- **Smart Contracts** Smart contracts create a possible for confirming and translating value. In the blockchain community, the contributors create investment intelligence which will be audited by smart contract. The consumers select some SIIs that can meet their individual needs and complete the entire transaction through smart contracts. As is shown in the Fig 7.

Fig. 7.   An application case about smart contracts.

The SIIs management protocol provides a auditing function which includes five steps, such as AI audition, expert reviewing and developer evaluation. This function can generate a comprehensive score for every SII. Smart contract will be executed based on thresholds and composite scores.

The key of AI technology is training the models to map the complex problems of the real world. The process of training model contains four stages. Therefore, in the blockchain community, the swarm intelligence of contributors can be attributed to these four aspects, including swarm raw data, swarm deep-processing data, swarm models and swarm recommendations.

B.3) Application Layer: The application layer mainly provides wallets and trading sites for users to realize the transfer of tokens or the calling of smart contracts by generating a series of transactions. The wallets include light clients, mainly for common users, and full clients, mainly for token miners.

It's worth mentioning that the system contains two types of accounts, including external account and internal

account, and only the external account can generate a transaction. The internal account is mainly used to store the smart contracts and the hash digest of CIIs, and the external account can call the contracts to use the CIIs by generating transactions.

• **CRYPTOLLA Value Ecosystem**

In the CRYPTOLLA value ecosystem (CVE), the value of SIIs are recognized and dynamically updated. The value score representative the current value of the SII. The audit comprehensive score is a initial envaluation for a SII. The call frequency score and the user value score (contributor) are evaluation index of dynamic updating. The infringement act judgement is adopted to guarantee intellectual property of SIIs. As is shown in the Fig 8.
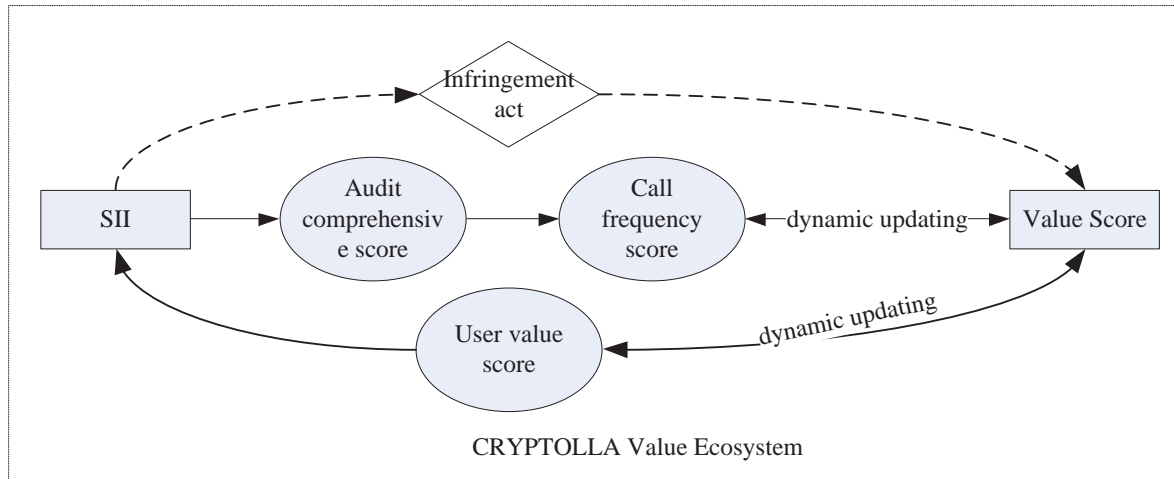
Fig. 8. The process of SII's value generation in the CRYPTOLLA value ecosystem.

In the figure, each node represents a scoring process, and if infringement, the score is zero. In addition, the value score of user can be automatically calculated based on the value of each SII and the total number of SIIs.

V. Appendices

A. Cryptolla System Architecture

The technical framework Fig 9 of the project is designed from the bottom up, starting with the underlying data platform and ending with the product application layer. Between them lies the big data layer, the automation layer, intelligence module layer and the integration functional module layer. The second and third layers are collectively referred to as the technical support layer. The main difference between the automation layer and the intelligence layer is the depth of the application of AI technology in their system design. This distinction also reflects the overall functional designing of the project which is based on a simple to complex layered concept. The functional module layer embodies the integration of similar services provided by the project and adopts a modular model for development. At the same time, it should be noted that the combination of AI and blockchain is always the primary focus throughout this framework.
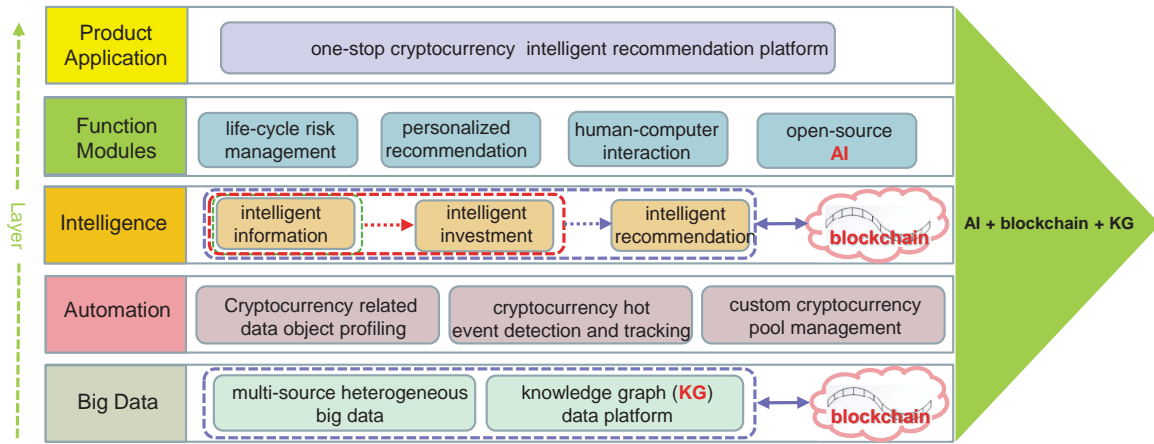
Fig. 9. The layered architecture design of the cryptolla.

According to the actual development needs, the functional design of the framework follows from simple to complex, from the basic data platform to the technical support layer, and to the application layer. The intelligent development level as shown below:

- Intelligence:
  - High-level: Intelligent recommendation
    * Middle-level: Intelligent investment research
      · Primary: Intelligent information

B. Cryptolla System Function Modules

Fig 10 shows the main function modules of cryptolla system. From the user's point of view, the main functions provided by the system include:

- **Deep knowledge search**
  Through the platform's knowledge retrieval engine, users can easily obtain relevant in-depth mining information of cryptocurrency & ICO profiling and credit scoring. Specially, these multi-dimensional information will be presented to the user in the form of information visualization.
- **Risk-return preference assessment**

When users access the system, if they participate in an online questionnaire, the system will automatically generate the user's risk-return preferences based on relevant models (FinaMetrica questionnaire analysis method).

- **Portfolio risk-return assessment**
The system provides users with a self-selected cryptocurrency service. If the user conducts simulation trading based on an optional cryptocurrency pool, the platform will automatically analyze the performance of the user's optional cryptocurrency portfolio.

- **Custom cryptocurrency pool management**
The system provides users with a push service for cryptocurrency portfolio risk early warning. Specifically, the user's risk tolerance and the user's cryptocurrency portfolio are matched in real time. When the value of the portfolio is below the risk safety line, an alert message is pushed to the user.

- **User profiling report**
Based on analysis of user log information and survey questionnaires, the system will provide users with a profiling reports, including risk-return preferences, information preferences, cryptocurrency & ICO preferences, and investment capability level assessment.

- **User information management**
User information management means that the system provides users with a series of personalized information services which include real-time hot event push and personalized information recommendation.

- **Event-driven cryptocurrency conceptual portfolio**
After detecting a hot-spot event, the system automatically invokes related models to provide users with related investment concepts and generate crypto-currency portfolios based on these investment concepts.

- **Event-driven cryptocurrency & portfolio analysis report**
After the system generates a cryptocurrency portfolio based on hot-spot events, the user may want to know how each cryptocurrency value changes in the portfolio as the event evolves. The system provides users with cryptocurrency profiling and event-related risk-return analysis reports. For example, there is a comprehensive index of dynamic reaction risk changes in the event-driven cryptocurrency & portfolio analysis report.

- **Investment intelligence contribution and acquisition**
Users can design relevant models and services according to their own needs through the open source artificial intelligence (AI) platform and knowledge graph (KG) data platform in the system, and can share the investment wisdom of others and contribute their own investment intelligence through the blockchain community of the system. This series of functions, is capable of satisfying the needs of users based on swarm investment intelligence and they can safely and conveniently convert their investment experience into intangible value (protected intellectual property).
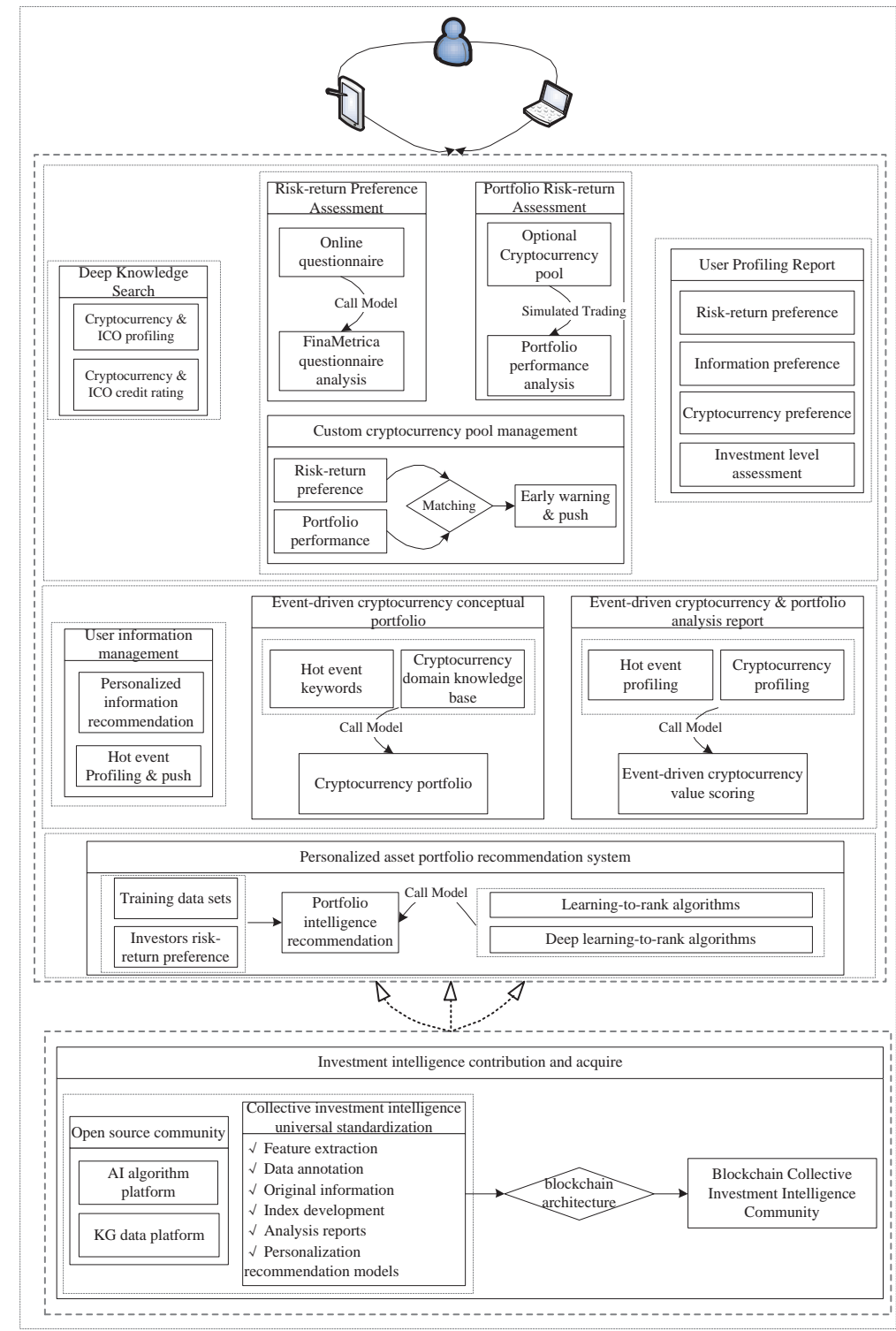
Fig. 10. The main function modules of cryptolla system.

a) S1 Fig.: **The flow chart of personalized asset portfolio recommendation system.** Fig 11 shows the flow description of personalized asset allocation recommendation system.
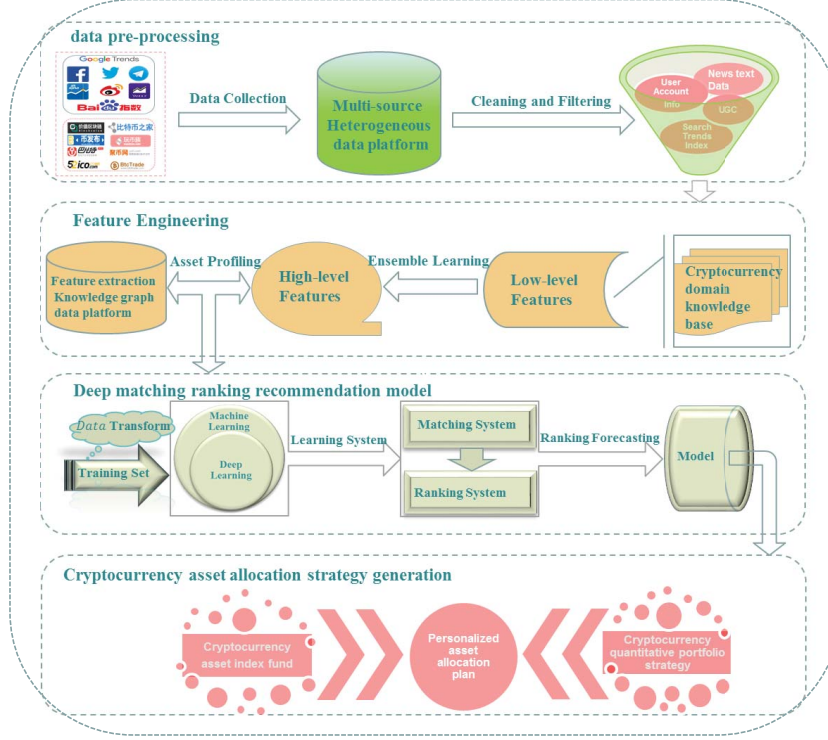

Fig. 11.   The flow chart of personalized asset allocation recommendation system.

b) S2 Fig.: **The development modules of cryptolla system.** Fig 12 shows the module groups description during system development.

c) S3 Fig.: **The model framework and data flow of cryptocurrency & ICO profiling.** Fig 13 shows a bottom-up implementation process for cryptocurrency or ICO profiling.

d) S4 Fig.: **The model framework and data flow of cryptocurrency information profiling.** Fig 14 shows the flow of information profiling.

e) S5 Fig.: **The model framework and data flow of user profiling.** Fig 15 shows the flow detail description in the problem of user profiling.

f) S6 Fig.: **The model framework and data flow of cryptocurrency related hot event detecting & profiling.** Fig 16 shows the specific description.

g) S7 Fig.: **The model framework and data flow of cryptocurrency & ICO credit rating.** Fig 17 shows the flow description in the problem solution.

h) S8 Fig.: **The model framework and data flow of personalized information recommendation system.** Fig 18 argue the technology architecture of personalized information recommendation system.

i) S9 Fig.: **The model framework and data flow of custom cryptocurrency pool management.** Fig 19 reflects the model characters of the custom cryptocurrency pool management.

j) S10 Fig.: **The model framework and data flow of event-driven cryptocurrency conceptual portfolio.** Fig 20 shows the model architecture of event-driven crypytocurrency conceptual portfolio.

k) S11 Fig.: **The model framework and data flow of event-driven cryptocurrency value scoring.** The factors depended by scoring and the methods applied to score are shown in the following Fig 21.
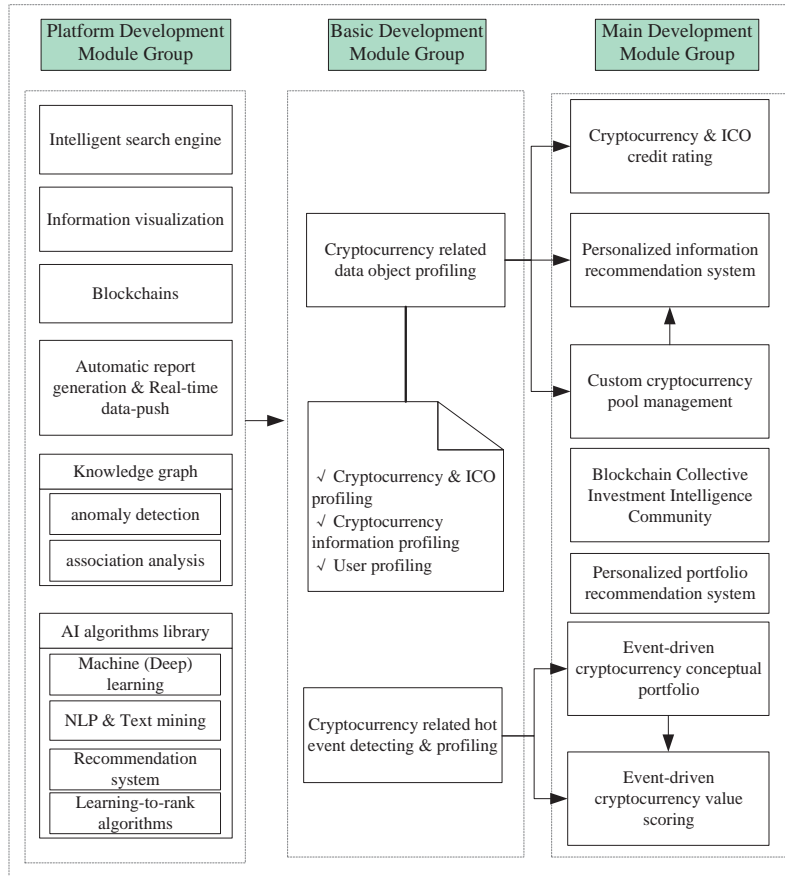
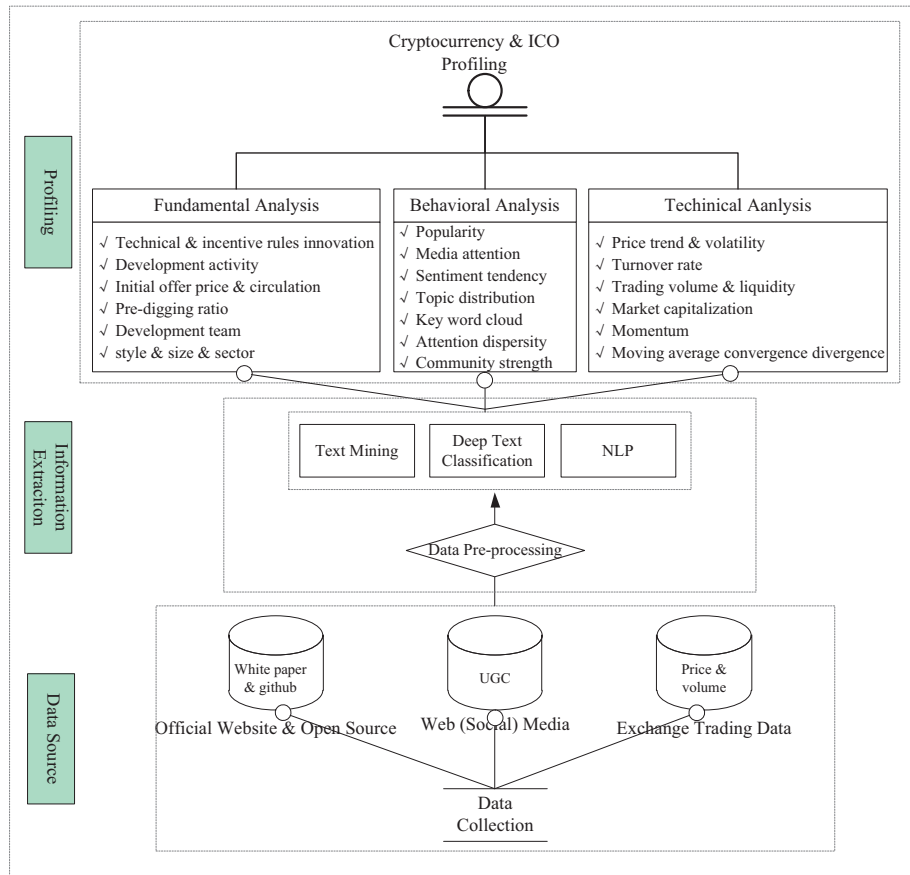Fig. 12. The development module groups in the first research and development phase.

Fig. 13.   The flow of the cryptocurrency or ICO profiling technology.
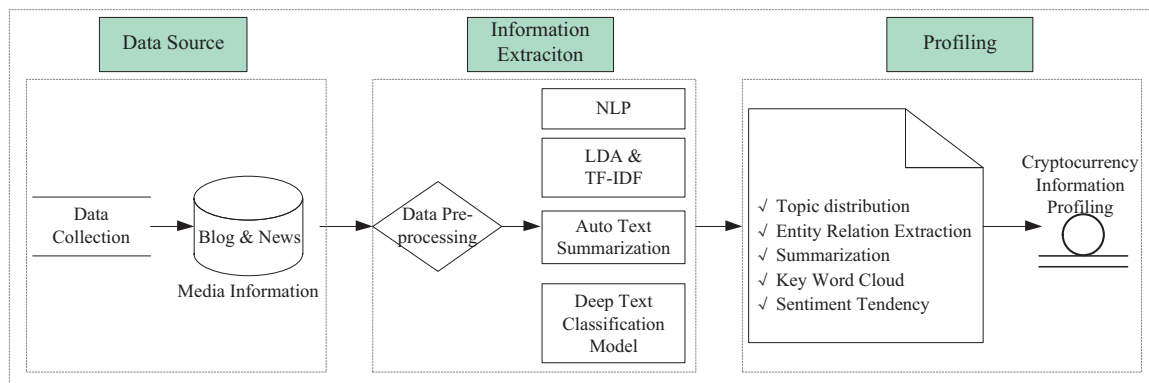

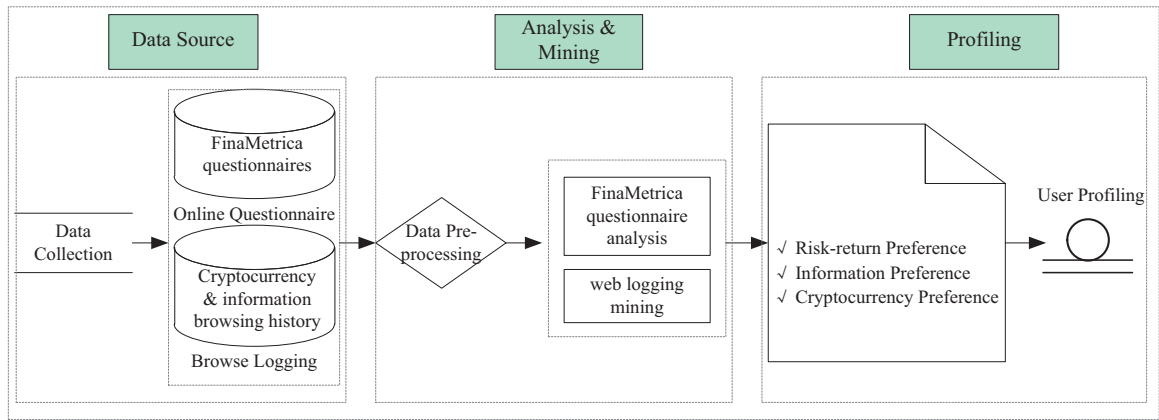Fig. 14.   The flow of the cryptocurrency information profiling technology.

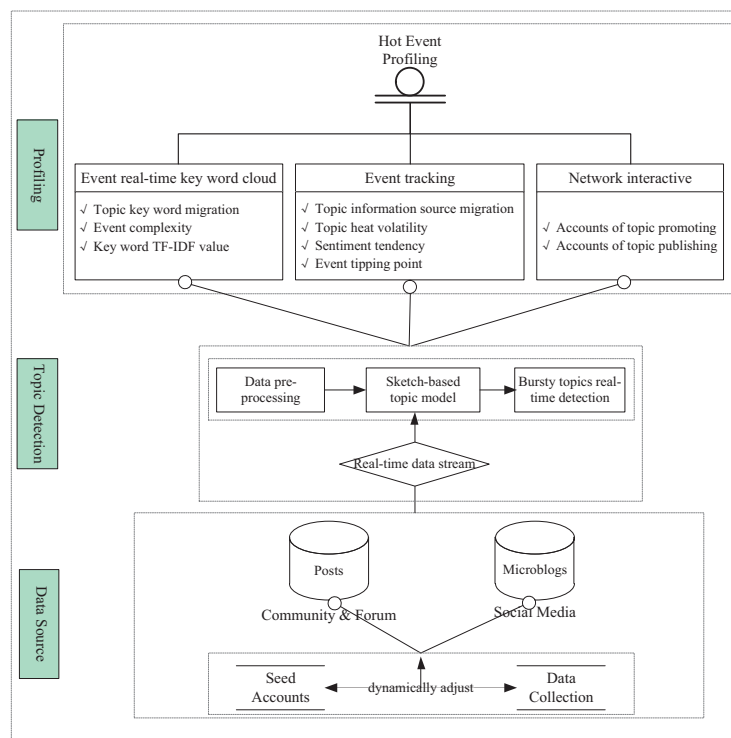Fig. 15.   The flow of user profiling technology.

Fig. 16.   The flow chart of cryptocurrency related hot event detecting and profiling.
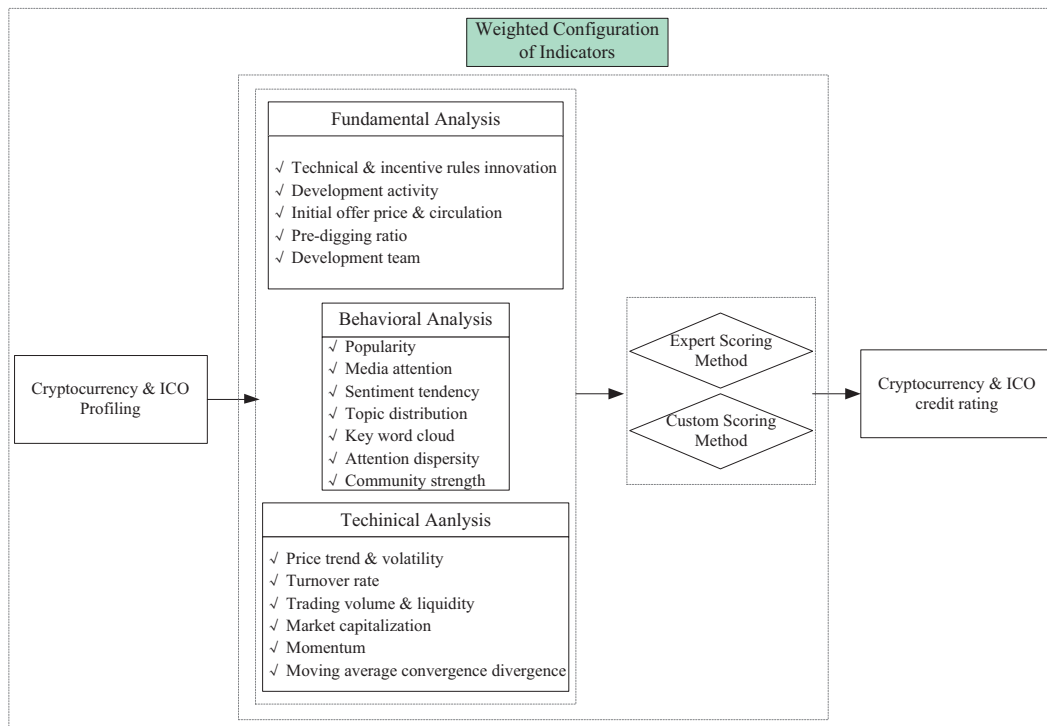
Fig. 17.   The approach of cryptocurrency or ICO credit rating.
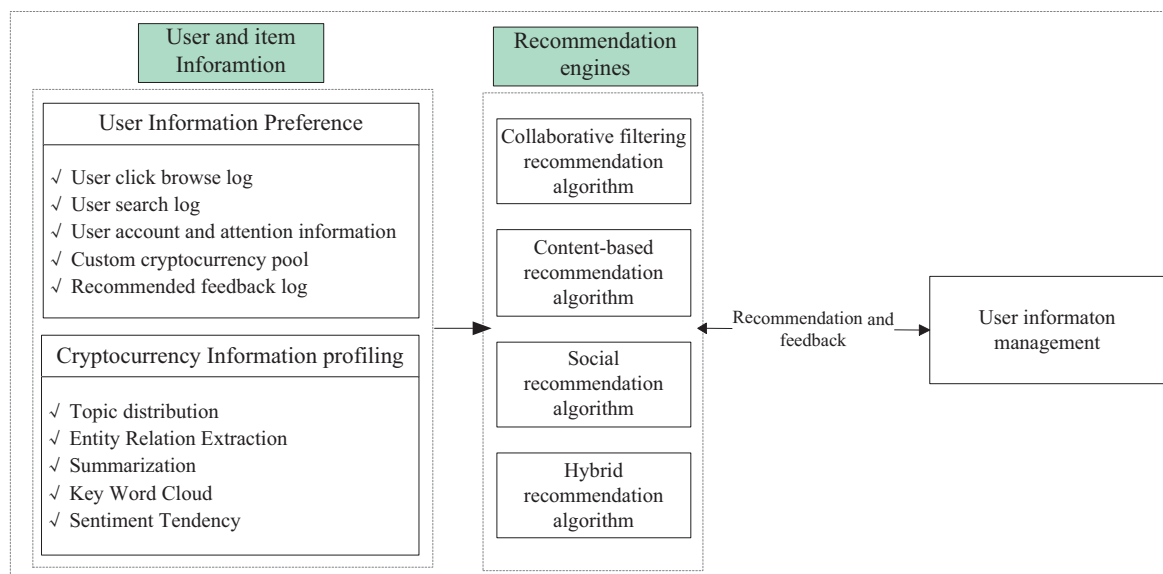

Fig. 18.   The technology architecture of personalized information recommendation system.
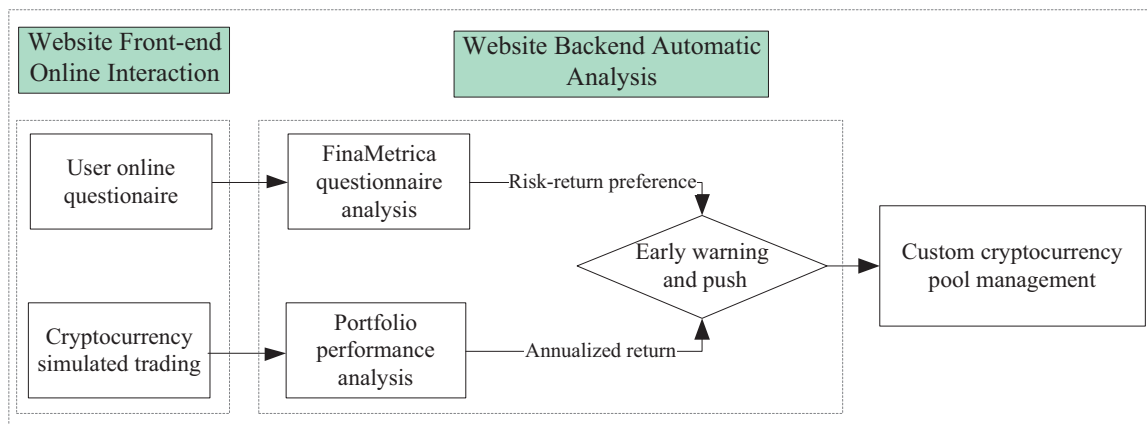
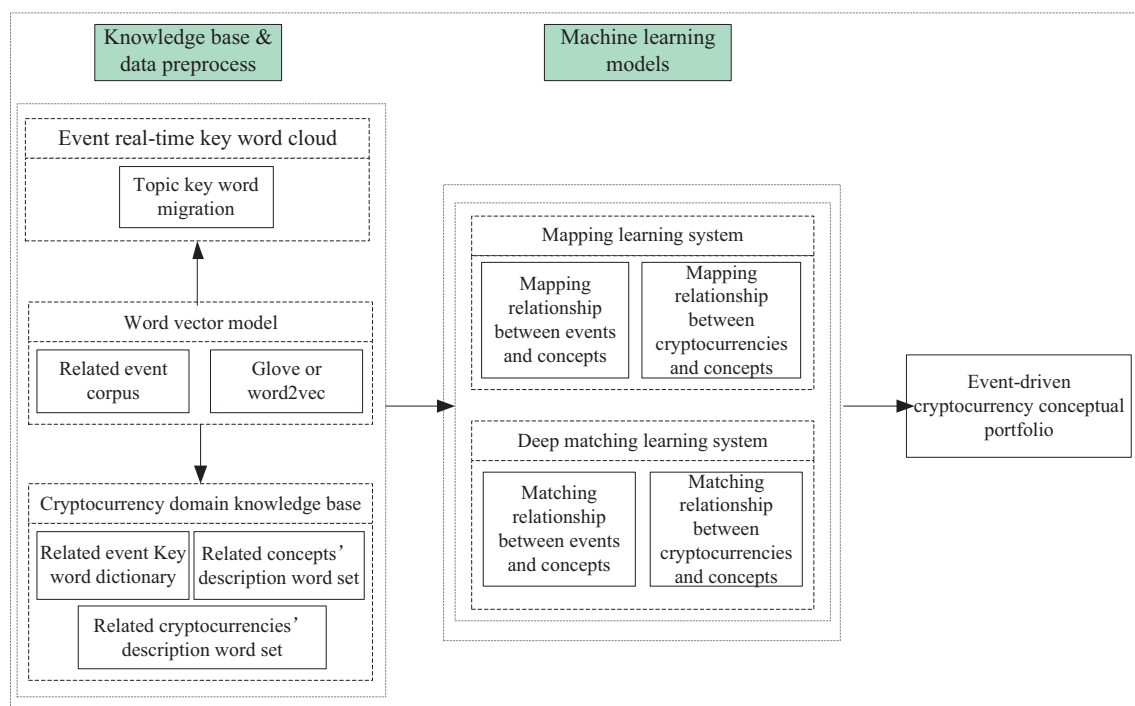Fig. 19. The model of custom cryptocurrency pool management.

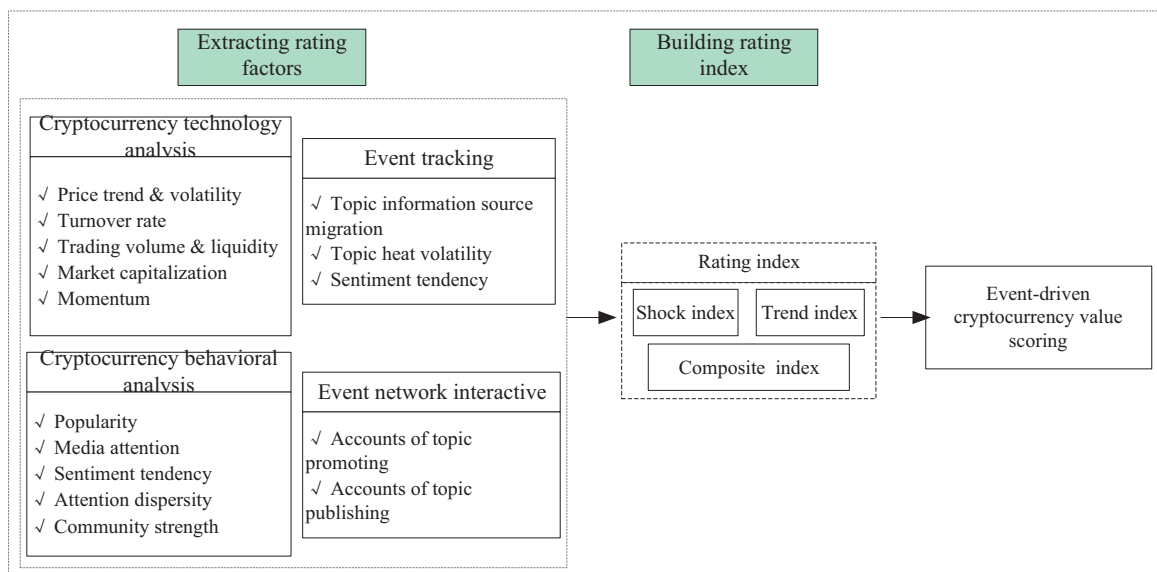Fig. 20. The model architecture of event-driven cryptocurrency conceptual portfolio.

Fig. 21.   The model architecture of event-driven cryptocurrency value scoring.

REFERENCES

[1] Atefeh F, Khreich W. A survey of techniques for event detection in twitter[J]. Computational Intelligence, 2015, 31(1): 132-164.

[2] Xie R, Zhu F, Ma H, et al. CLEar: a real-time online observatory for bursty and viral events[J]. Proceedings of the VLDB Endowment, 2014, 7(13): 1637-1640.

[3] Xie W, Zhu F, Jiang J, et al. Topicsketch: Real-time bursty topic detection from twitter[J]. IEEE Transactions on Knowledge and Data Engineering, 2016, 28(8): 2216-2229.

[4] Shen Y, Rong W, Sun Z, et al. Question/Answer Matching for CQA System via Combining Lexical and Sequential Information[C]//AAAI. 2015: 275-281.

[5] Severyn A, Moschitti A. Learning to rank short text pairs with convolutional deep neural networks[C]//Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM, 2015: 373-382.

[6] Lu Z, Li H. A deep architecture for matching short texts[C]//Advances in Neural Information Processing Systems. 2013: 1367-1375.

[7] Goldberg Y, Levy O. word2vec explained: Deriving mikolov et al.'s negative-sampling word-embedding method[J]. arXiv preprint arXiv:1402.3722, 2014.

[8] Sarwar B, Karypis G, Konstan J, et al. Item-based collaborative filtering recommendation algorithms[C]//Proceedings of the 10th international conference on World Wide Web. ACM, 2001: 285-295.

[9] Gong S. A collaborative filtering recommendation algorithm based on user clustering and item clustering[J]. JSW, 2010, 5(7): 745-752.

[10] Pazzani M J, Billsus D. Content-based recommendation systems[M]//The adaptive web. Springer, Berlin, Heidelberg, 2007: 325-341.

[11] Ma H, Yang H, Lyu M R, et al. Sorec: social recommendation using probabilistic matrix factorization[C]//Proceedings of the 17th ACM conference on Information and knowledge management. ACM, 2008: 931-940.

[12] Chen W, Niu Z, Zhao X, et al. A hybrid recommendation algorithm adapted in e-learning environments[J]. World Wide Web, 2014, 17(2): 271-284.

[13] Sorokin A, Forsyth D. Utility data annotation with amazon mechanical turk[C]//Computer Vision and Pattern Recognition Workshops, 2008. CVPRW'08. IEEE Computer Society Conference on. IEEE, 2008: 1-8.

[14] Middleton S E, Shadbolt N R, De Roure D C. Ontological user profiling in recommender systems[J]. ACM Transactions on Information Systems (TOIS), 2004, 22(1): 54-88.

[15] dos Santos C, Gatti M. Deep convolutional neural networks for sentiment analysis of short texts[C]//Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers. 2014: 69-78.

[16] Lai S, Xu L, Liu K, et al. Recurrent Convolutional Neural Networks for Text Classification[C]//AAAI. 2015, 333: 2267-2273.

[17] Van de Venter G, Michayluk D, Davey G. A longitudinal study of financial risk tolerance[J]. Journal of Economic Psychology, 2012, 33(4): 794-800.

[18] Ma F, Ng K, Lu N, et al. System for simulating trading of financial assets: U.S. Patent Application 10/047,134[P]. 2002-5-16.

[19] Mintz S P, West R, Kondilis C. System and method for simulating an electronic trading environment: U.S. Patent 7,672,895[P]. 2010-3-2.

[20] Tom M D. System and method for performing risk and credit analysis of financial service applications: U.S. Patent 5,696,907[P]. 1997-12-9.

[21] Huang Z, Chen H, Hsu C J, et al. Credit rating analysis with support vector machines and neural networks: a market comparative study[J]. Decision support systems, 2004, 37(4): 543-558.

[22] Zhu F, Qu Q, Lo D, et al. Mining top-k large structural patterns in a massive network[J]. Proceedings of the VLDB Endowment, 2011, 4(11): 807-818.

[23] Zhu F, Yan X, Han J, et al. gPrune: a constraint pushing framework for graph pattern mining[C]//Pacific-Asia Conference on Knowledge Discovery and Data Mining. Springer, Berlin, Heidelberg, 2007: 388-400.

[24] Mining W I D. Data Mining: Concepts and Techniques[J]. Morgan Kaufinann, 2006.

[25] Dai H, Zhu F, Lim E P, et al. Mining coherent anomaly collections on web data[C]//Proceedings of the 21st ACM international conference on Information and knowledge management. ACM, 2012: 1557-1561.

[26] Paulheim H. Knowledge graph refinement: A survey of approaches and evaluation methods[J]. Semantic web, 2017, 8(3): 489-508.

[27] Lai K K, Yu L, Wang S, et al. A double-stage genetic optimization algorithm for portfolio selection[C]//International Conference on Neural Information Processing. Springer, Berlin, Heidelberg, 2006: 928-937.

[28] Pareek M K, Thakkar P. Surveying stock market portfolio optimization techniques[C]//Engineering (NUiCONE), 2015 5th Nirma University International Conference on. IEEE, 2015: 1-5.

[29] Song Q, Liu A, Yang S Y. Stock portfolio selection using learning-to-rank algorithms with news sentiment[J]. Neurocomputing, 2017, 264: 20-28.

[30] Bao W, Yue J, Rao Y. A deep learning framework for financial time series using stacked autoencoders and long-short term memory[J]. PloS one, 2017, 12(7): e0180944.

[31] Christidis K, Devetsikiotis M. Blockchains and smart contracts for the internet of things[J]. IEEE Access, 2016, 4: 2292-2303.

[32] Benet J. IPFS-content addressed, versioned, P2P file system[J]. arXiv preprint arXiv:1407.3561, 2014.