Journal of Engineering Innovations & Technology (2024)

Journal of Engineering Innovations & Technology

journal homepage: www.journal-eit.online

Paper Type: Original Article

A Bayesian Networks-based Risk Prediction and Decision Support System for Financial Markets

Yufei He[1#], Haihao, Luo[2#], Zongze He[3#], Jiaxin Chen[4#], He Wang[2&], Tianheng Wang[2&]

[1]*King's College London, London, United Kingdom*

[2]*Shenzhen Aurora International Academy, Shenzhen, China*

[3]*Fairmont Schools, Anaheim, United States*

[4]*Shenzhen Senior High School, Shenzhen, China*

[#]Co-first author

[&]Co-second author

Abstract

With the rapid development and increased complexity of global financial markets, the effectiveness of risk management becomes more and more important. In this paper, a Bayesian network-based financial market risk prediction and decision support system (DSS) is proposed to enhance the accuracy and efficiency of financial decision making. The system architecture includes an input data module, an inference engine, and a user interaction interface, emphasizing data integration and real-time processing capabilities. The input module is responsible for collecting real-time and historical financial data, and the inference engine utilizes Bayesian inference algorithms to assess risks and generate decision recommendations. It is shown that the system is able to effectively capture the complex dependencies of variables in the financial market, provide accurate risk prediction, and thus help decision makers to make scientific judgments in a dynamic environment.

In addition, the user interaction interface is designed to make the system more user-friendly for quick understanding and application of risk assessment results. The integration and maintenance strategy of the system ensures the stability and scalability of the DSS and lays the foundation for future functionality enhancements. Despite the positive results of this study in financial market risk management, there are still some challenges, such as the dependence on the quality of input data and the increase in computational complexity. Future research could explore the integration with other machine learning techniques to further enhance the prediction capabilities.

Keywords: Bayesian Networks, Risk Prediction, Decision Support Systems, Financial Markets, Uncertainty Management

1 | Introduction

In today's financial markets, the importance of risk management is becoming more and more pronounced. With the continuous development of the global economy and the diversification of financial instruments, investors and financial institutions are faced with an ever-increasing variety and complexity of risks. These risks come not only from market volatility, but are also affected by multiple factors such as macroeconomic factors, policy changes and the social environment. Accurate identification and assessment of financial risks are therefore critical to achieving sustainable investment returns and maintaining financial stability.

In this context, Decision Support Systems (DSS), as an effective management tool, have gradually attracted extensive attention from researchers and practitioners. DSS help decision makers make more informed choices in the face of uncertainty by integrating data analysis, reasoning mechanisms and user interaction functions. In particular, DSS based on Bayesian networks has shown good potential for application due to its advantages in dealing with uncertainty and modeling complex dependencies.

However, despite the progress made by existing DSS in the field of risk management, some challenges remain. For example, how to efficiently integrate real-time and historical data, how to design a flexible and easy-to-use user interface, and how to ensure the maintainability and scalability of the system are all pressing issues. Therefore, this dissertation aims to deeply explore the method of building DSS based on Bayesian networks, to enhance the accuracy and practicality of financial market risk prediction through systematic architectural design and modular implementation, and to provide more effective support for financial decision-making.

2 | Literature Review

After the successive crises in the financial field, the economies of various countries have experienced different degrees of economic recession, and the people's living standards have declined drastically, resulting in economic turmoil in various regions, and the systemic financial risk has been continuously focused on the response. International research on the response to systemic financial risk is generally placed in the context of the financial crisis framework [1], and has not yet formed a theoretical system recognized by all scholars [2]. At present, the definition of systemic financial risk has not reached a unified conclusion, and there is no common understanding of the focus of financial risk. Federal Reserve Chairman Bernanke believes that systemic financial risk is large enough to pose a threat to the modern economy and financial system [3]. F. Kaufman [4] emphasizes the contagion property of systemic financial risk, i.e., a sudden event may lead to a cascade of losses in the relevant institutions and markets, which can be rapidly transmitted and lead to the spread of the losses. zigrand [5] believes that systemic financial risk is the risk of causing the financial market to suffer shocks and affecting the normal operation. Financial market shocks and affect the normal operation of the financial market, resulting in significant socio-economic losses. BIS and IMF [6] interpret systemic financial risk as the risk of severe damage to the financial services system, creating a widespread forced cessation of financial services, leading to a sharp downturn in the economic sector. Diamond and Dybvig [7] believe that the systemic financial risk is triggered by a localized event. The resulting risk can spill over into the financial system so that the crisis spreads and eventually leads to the deterioration and then collapse of the entire financial system, causing significant harm to the real economy and social wealth.

Initially, scholars believed that systemic financial risk was affected by changes in macroeconomic and financial variables, which were determined by analyzing historical data on financial crises and then summarized by statistical methods to derive risk assessment indicators. Alessi and Detken [8] used macro data such as GDP, inflation rate, and monetary aggregates to build an indicator system and predicted some national. Duca and Peltonen [9] used macro data from emerging markets as well as economic data from more developed countries to explore that credit size can be used as an indicator of the eve of risk outbreak.

Scholars have continued to explore financial system risk, observing that the macro and micro data indicator method does not have continuity in monitoring systemic financial risk, and lacks foresight due to the significant lag in most macro data, which are low-frequency data, and have gradually turned to the study of key nodes of financial institutions in the outbreak of crises. Freixas and Rochet [10] suggest that the network structure will accelerate the spread of inter-bank defaults. The network structure will accelerate the spread of inter-bank defaults. Over time many scholars believe that the higher the degree of network connectivity, the easier it is to trigger the spread of risk, and it is more contagious.Iori, Jafarey and Padilla [11] and other studies found that systemic financial risk is closely related to the degree of network connectivity, and in the era of high interconnectivity, high-intensity shocks could be devastating to the entire financial system. Ouyang Hongbing and Liu [14] used network analysis to simulate inter-bank transactions, identify nodes in the financial network, and explore the propagation mechanism of systemic financial risk.

Financial institutions settlement data often belong to the key data, the market is not public, and lack of dynamic disclosure, and its measurement results can only rely on the simulation of risk induction, based on assumptions rather than real shocks, which leads to the actual application of the network analysis method often has many deviations [15]. The continuous improvement of financial markets has dramatically increased the availability and accessibility of public market data, and many scholars have turned to public market data as the basis for measuring systemic financial risk.

3 | Overview of Bayesian Network Modeling of Financial Market Risk

3.1 Basic Components of Bayesian Networks

Bayesian networks are powerful graphical models that consist mainly of nodes and directed edges. Nodes represent random variables, while edges represent conditional dependencies between these variables. Each node is associated with a conditional probability table (CPT) that describes the probability distribution of that node given its parent. This structure allows Bayesian networks to clearly demonstrate the causal relationships of the variables in a system, making complex dependencies intuitively understandable.

With such a composition, Bayesian networks can effectively capture uncertainty and provide a basis for inference. When the states of certain variables are observed, Bayesian networks can utilize the known information to infer the possible states of other variables, thus supporting the decision-making process. This property makes Bayesian networks have a wide range of potential applications in areas such as risk prediction in financial markets, and can help decision makers make more accurate judgments in dynamic and complex environments.

3.2 Steps in Constructing a Bayesian Network

Constructing a Bayesian network is a systematic process aimed at building an effective model to deal with uncertainty by identifying variables and their relationships. The process is usually divided into several key steps, starting with the definition of random variables related to the topic of study, which will become the base nodes of the network. Next, the structure of the network needs to be constructed to specify the dependencies between the variables, which can be achieved through expert knowledge or data mining methods. Then, conditional probability tables (CPTs) are computed for each node through parameter estimation to ensure that the model accurately reflects the relationships between variables. Finally, validation and tuning are performed to ensure the validity and predictive power of the model. Through these steps, a Bayesian network that can effectively support risk prediction in financial markets can be constructed.

3.2.1 | Defining Variables

Financial risk is divided into two main blocks, namely systemic risk and non-systemic risk. Systemic risk refers to the threat of economic loss faced by financial institutions or investors engaged in financial activities due to the violent fluctuations of the economic system caused by the impact of external factors or changes in their own internal factors in the overall economic system. Systemic risk is mainly caused by socio-political, environmental, cultural and other macro elements, is engaged in financial activities of individuals can not control, and can not be eliminated through diversified investment, also known as uncontrollable risk. Non-systematic risk is the risk that affects the return of an industry or a stock, and does not produce a large impact on the entire securities market in the economy. Units engaged in financial activities can reduce or eliminate unsystematic risk by taking the right risk management measures, such as diversification can reduce or eliminate the possibility of asset loss. Based on the risk classification, we need to identify the random variables that are relevant to the topic of the study, which will act as nodes in the network.

Table 1 Risk Classification

Risk classification		Risk description
systemic risk	market risk	The loss of value of financial assets triggered by changes in market factors including interest rates, exchange rates, stock prices, etc. is called market risk.
	policy risk	Economic policies and regulatory decisions taken by the government can directly affect changes in the investment returns of financial practitioners.
	interest rate risk	Interest rate risk is the potential for losses on commercial bank assets due to unfavorable movements in market interest rates.
	purchasing power risk	Due to inflation, the purchasing power of the same amount of money is different at different moments, and this side is the purchasing power risk.
Non-systemic risk	credit risk	Credit risk is also called default risk. In the securities market, it refers to the possibility that the issuer of a security will not have sufficient funds to repay the investor's principal and interest upon maturity, resulting in a loss to the investor.
	financial risk	Financial risk can be observed through the capital structure of a company. The greater the ratio of the sum of the value of loans and bonds to the total assets of the company, the higher the financial risk, and vice versa.
	business risk	Uncertainty in the company's operating income caused by problems in the external operating environment (institutional, cultural, ecological, etc.) and the company's internal day-to-day management that are not under the company's control, thus causing uncertainty in the investor's return, is the business risk faced by the investor.
	liquidity risk	The uncertainty to investor returns when there are difficulties in realizing assets is generally referred to as liquidity risk.

3.2.2 | Constructing the Network Structure

First, we need to identify the dependencies between the variables. Set a set of random variables $V = \{V_1, V_2, ..., V_n\}$, and our goal is to construct a directed acyclic graph (DAG), where each node corresponds to a random variable and edges represent conditional dependencies. For each pair of random variables (V_i, V_j), we define the relation $R(V_i, V_j)$, which denotes V_i whether it directly affects V_j in the financial risk prediction. In financial risk forecasting, dependencies may include the effect of market volatility on stock prices, the effect of interest rates on trading volume, etc. Based on domain knowledge and experience, we can initially construct a network structure that contains these dependencies.

Next, we build the network structure. The core idea of the K2 algorithm is to determine the optimal network structure by maximizing a conditional probability distribution. Specifically, we wish to find a set of variables arranged such that given a parent node $Pa(V_i)$ given the parent node, the conditional probability $P(V|Pa(V))$ is maximized. This probability can be expressed as:

$$P(V|Pa(V)) = \prod_{i=1}^{n} P\left(V_i|Pa(V_i)\right)$$

Where $Pa(V_i)$ is the set of nodes V_i of the set of parent nodes. To find the optimal structure, we need to traverse all possible network structures. The K2 algorithm has low complexity and is suitable for dealing with relatively small sets of variables.

If sufficient historical data are available, we can apply statistical methods to determine the dependence between variables. Commonly used statistical methods include correlation analysis and independence tests. Taking mutual information as an example, we can use the following formula to evaluate the variables X and Y the dependence between them:

$$I(X;Y) = \sum_{x \in X} \sum_{y \in Y} P\left(X = x, Y = y\right) \log \frac{P(X = x, Y = y)}{P(X = x)P(Y = y)}$$

If $I(X;Y) > 0$, then it indicates that X and Y have some kind of dependency, and edges can be created between them in a Bayesian network. This process can be automated algorithmically, for example by using the PC algorithm (Peter-Clark algorithm) or the GES (Greedy Equivalence Search) algorithm to derive dependencies.

After establishing the dependencies, the next step is to specify the parent node for each node. In a Bayesian network, each node can have more than one parent node, and these parents directly affect the state of that node. For example, suppose we have three variables: market volatility M stock price P and interest rate R. If we consider that market volatility and interest rates jointly affect stock prices, then we can assign the M and R are set as the parent nodes of P . This can be expressed as:

$$P(P|M, R)$$

After identifying the dependencies, the Bayesian network structure can be visualized using graphical tools. This process not only helps to understand the relationships between the variables, but also provides an intuitive framework for subsequent parameter estimation and inference.

3.2.3 | Parameter Estimation

In Bayesian networks, parameter estimation is a key step in calculating the conditional probability table (CPT) for each node. This process ensures that the network structure effectively reflects the relationships between the variables.

First, historical data need to be collected on the selected variable. The data set should contain enough samples to ensure that the estimated probabilities accurately reflect the actual situation. Suppose we have a training set containing N samples $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$, where each pair (x_i, y_i) represents a set of observations of a variable.

For each node V_i, we need to define its conditional probability table $P(V_i|Pa(V_i))$, where $Pa(V_i)$ denotes the set of V_i the set of parent nodes of the node. The conditional probability table describes, given a parent node, the node'sV_i probability distribution. For example, suppose node V_i has k parent nodes $V_1, V_2, \dots, V_k$, then the conditional probability table can be expressed in the form:

$$P(V_i|V_1, V_2, \dots, V_k)$$

Maximum likelihood estimation methods are commonly used to estimate conditional probabilities in the absence of prior knowledge. For a given parent node configuration $Pa(V_i)$, we can calculate the nodeV_i frequency of occurrence under that configuration. LetN_{ijk} be a node in the parent node state$j$ and$k$ under which the nodeV_i takes the value of$x$ the observation frequency, then the conditional probability can be estimated using the following formula:

$$P(V_i = x|Pa(V_i) = j, k) = \frac{N_{ijk}}{N_{ij.}}$$

Where $N_{ij.}$ is in the parent node state$j$ and$k$ the total frequency under

In the presence of prior knowledge, Bayesian estimation can better handle the problem of sparse data. By introducing the prior distribution$P(\theta)$, we can compute the posterior distribution:

$$P(\theta|D) \propto P(D|\theta)P(\theta)$$

Where θ denotes the model parameters, and D is the observed data. We usually assume that the prior is a Dirichlet distribution because it combines well with multinomial distributions and is suitable for multiclass situations.

Let α be the parameters of the Dirichlet distribution, then for each$x$ conditional probability estimate can be expressed as:

$$P\big(V_i = x|Pa(V_i)\big) = \frac{N_{ijk} + \alpha_x}{N_{ij.} + \sum_{x'} \alpha_{x'}}$$

Here α_x is a parameter associated with the x associated a priori parameter.

Once the parameter estimation is complete, the conditional probability tables need to be validated. Cross-validation methods can be used to assess the performance of the model on unknown data and to check the accuracy and robustness of the predictions. If necessary, the conditional probabilities are adjusted based on the validation results to improve the model performance.

3.2.4 | Model Performance Evaluation

Cross-validation is a commonly used model evaluation method to effectively test the performance of Bayesian networks on unknown data. A common method is K-fold cross-validation, in which the dataset is divided into K subsets. One subset at a time is selected as the test set and the remaining $K - 1$ subsets are used as the training set, the model is trained on the training set and then evaluated on the test set. This process is repeated K times and finally the average performance metrics of the model such as accuracy, precision and recall are calculated. For example, for a binary classification problem, the accuracy of the model can be expressed as:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

of which TP is true positive, and TN is true negative, and FP is false positive, and FN is false negative.

In addition to accuracy, other evaluation metrics can be used, such as AUC (Area Under the Curve) and F1 score.AUC is the area under the ROC curve, which can reflect the classification performance of the model, and is given by the formula:

$$AUC = \int_0^1 TPR\,(x)\; d\big(FPR(x)\big)$$

where TPR (true rate) and FPR (false positive rate) are the sensitivity and specificity of the model.

3.2.4 | Model Restructuring

Restructuring primarily involves reassessing and optimizing the dependencies between variables in the network.

1. Assessing Dependencies

Through model validation and sensitivity analysis, identify which dependencies between variables are not being captured correctly. For example, if it is found that a change in some variable significantly affects the model output, but this relationship is not reflected in the network structure, then consideration should be given to adding the appropriate edges.

2. Add Edge

If the validation process recognizes the variables A and B there is a significant causal relationship between them, but the initial model does not incorporate it, adjustments can be made through the following steps:

Data support: analyzing variables using historical data A and B relationships to ensure that decisions to add edges are supported by data.

Update the network: new edges are added to the network and the conditional probability table is updated to reflect the added dependencies.

For example, suppose we find in our model that market volatility M directly affects the stock price P , and this dependency is missing in the original model, add a line from M to P the edge of the

3. Deletion of Edges

If certain edges fail to provide effective support for the predictive power of the model, or cause the model to overfit, consider removing these edges. This process involves:

Analytical contribution: assessing the contribution of specific edges to model performance can be determined by comparing the predictive power of the model.

Update the structure: remove these edges from the network and recompute the conditional probability table.

For example, if it is found that an edge between the stock price P and some external economic indicator does not show a significant relationship in the data and the model performance improves after deletion, the edge should be considered for deletion.

4 | Decision Support System (DSS) Construction

The Decision Support System (DSS) is designed based on a Bayesian network architecture, where each component plays an important role to ensure that the system is able to effectively deal with prediction and decision support for financial market risks.

4.1 Input Data Module

The Input Data Module is one of the fundamental components of a Decision Support System (DSS), whose main function is to collect, process and store real-time or historical financial data to ensure that the system is able to perform risk assessment and decision support based on accurate information. The diversity and complexity of data in the financial market requires the input module to have strong data acquisition and processing capabilities.

First, the input data module needs to obtain data from multiple sources, which may include real-time data streams, economic indicators, foreign exchange rates, company financial statements, and social media sentiment analysis. Real-time data streams usually come from financial exchanges and involve market dynamics such as stock prices and trading volumes, while economic indicators include national macroeconomic statistics such as Gross Domestic Product (GDP), unemployment rates, and consumer price indices, which reflect the health of the overall economy. Foreign exchange rates, on the other hand, are an important factor affecting the global financial markets, and monitoring exchange rate movements between major currencies can help policymakers to assess the associated risks.

Second, to ensure the quality of the input data, the module needs to perform data preprocessing. This process includes data cleansing, data normalization and data integration. Data cleansing aims to remove missing values and outliers to improve the accuracy and completeness of the data. In the financial sector, outliers can be caused by market fluctuations or data entry errors, so careful scrutiny of the data is critical. Data normalization is the process of converting data of different scales to a uniform scale to ensure that different types of data are comparable in subsequent analyses. In addition, data integration combines data from different sources into a consistent data format to facilitate subsequent processing and analysis.

The input data module also needs to have flexible interfaces to facilitate the receipt and updating of data in real time. This real-time capability enables DSS to respond quickly to market changes in a dynamic financial environment, providing timely and up-to-date risk assessment results and decision-making recommendations.

4.2 Inference Engine

The inference engine is a key component of a Decision Support System (DSS), whose main function is to perform risk analysis and decision making inference based on input data through Bayesian networks. The engine is designed to provide real-time financial market risk prediction to help decision makers effectively respond to complex market environments.

First, the inference engine relies on a previously constructed Bayesian network model. This model not only captures the conditional dependencies between variables in the financial markets, but also dynamically updates these relationships based on input data. By using a Bayesian inference algorithm, the inference engine is able to compute the posterior probability of each variable given a particular input condition. This process involves the use of Conditional Probability Tables (CPTs) to reflect the probability distribution of each node under specific parent conditions. This reasoning capability allows the system to effectively handle uncertainty and derive possible future states based on available information.

Second, the implementation of the inference engine includes a variety of inference algorithms, such as inference algorithms and sampling methods. Specifically, methods such as variational inference and Gibbs sampling are commonly used to deal with complex Bayesian networks, ensuring that the model can operate efficiently on large-scale datasets. With these algorithms, the inference engine is able to quickly compute relevant risk assessment results when faced with different risk scenarios. In addition, the engine can perform sensitivity analyses to assess the degree of influence of different variables on the final prediction results, which provides more in-depth insights for decision makers.

Another important function of the inference engine is to generate decision support information. After making a risk prediction, the inference engine translates the results into easy-to-understand risk assessment reports. These reports include not only quantitative analysis of risk levels, but also qualitative descriptions to help decision makers understand the nature of potential risks and their impact on business operations. In addition, the inference engine provides decision-making recommendations based on the risk assessment results, guiding the user on how to deal with the risks. These recommendations may include adjusting investment portfolios, implementing risk hedging strategies, or developing emergency response plans.

4.3 Output Module Module

The Output Module is a crucial part of the Decision Support System (DSS), responsible for translating the risk assessment results generated by the inference engine into easy-to-understand and implement decision recommendations. The module is designed to ensure that the user is able to quickly grasp the market risk situation and take decisions accordingly.

In the output module, the generated risk assessment results typically include a variety of key indicators, such as risk level, potential loss, and associated risk factors. These metrics are presented visually so that decision makers can intuitively understand the current risk profile. For example, the results of a risk assessment can be presented in the form of a risk matrix or dashboard, with risk levels color-coded (e.g., red, orange, and green), allowing users to quickly identify areas of concern.

In addition, the output module provides decision-making recommendations that are based on the results of the inference engine's analysis. For example, if the assessment results show that the credit risk of an asset has increased significantly, the output module may recommend that investors consider reducing their holdings of the asset or hedging their risk. The decision recommendations not only rely on the current risk assessment results, but also incorporate historical data and market trends to ensure that the recommendations are accurate and actionable.

To support decision making, the output module can also include simulation and forecasting capabilities. Users can adjust different input parameters (e.g. market volatility, interest rates, etc.) to simulate risk changes under different scenarios, thus assessing the potential impact of various decisions. This "what-if" function helps decision makers to be more thoughtful in the face of uncertainty.

In terms of implementation techniques, the output module can be enhanced by an automated report generation system. Such a system could generate risk assessment reports on a regular or real-time basis, pushing the necessary information to the relevant decision makers so that they can respond in a timely manner.

4.4 User Interaction Interface

The user interaction interface is an important part of a decision support system (DSS), which provides an intuitive and friendly operating environment for users to interact with the system effectively. Designing an efficient UI not only improves the usability of the system, but also enhances the user experience and ensures a smooth and efficient decision-making process.

First, the design of the user interaction interface should focus on the visualization of information. Through a graphical interface, users can intuitively view real-time financial data, risk assessment results and system-generated decision-making recommendations. For example, using visualization tools such as charts, curves and dashboards, market trends, risk levels and other key indicators can be clearly displayed. Such visualizations not only help users quickly understand complex data, but also effectively reduce the cognitive burden of information processing.

Second, the user interaction interface should support customization and personalization settings. Different users may have different needs and preferences, so it is important to provide flexible interface setting options. Users can choose the presentation of financial data, risk indicators, and decision-making suggestions they are concerned about according to their needs. In addition, the system can intelligently recommend relevant information based on the user's historical operation records and preferences to improve decision-making efficiency.

The interactivity of the interface is also a key factor. Users should be able to quickly access the information they need through simple actions such as clicking, dragging or sliding. For example, users can view the corresponding market data or risk assessment results by selecting a specific timeframe, or enter relevant parameters through a simple form and the system instantly updates and displays the results. This flexible interaction enhances the user's sense of engagement and strengthens their trust and reliance on the system.

Finally, the user interaction interface should also have a friendly feedback mechanism. After the user has performed an operation, the system should give timely feedback to ensure that the user understands the current state and the result of the operation. For example, when the user enters data or selects a parameter, the system should guide the user through prompt messages, warnings or confirmation dialog boxes. This feedback mechanism not only prevents users from making operational errors, but also enhances users' confidence in the system.

4.5 System Integration and Maintenance

System integration and maintenance is an important part of ensuring the long-term effective operation of a decision support system (DSS). An efficient DSS not only requires seamless integration among various modules, but also requires good maintenance and updating in daily use to adapt to the dynamic changes in the financial market and the changing needs of users.

First, the goal of system integration is to efficiently connect various components such as input data modules, inference engines and user interaction interfaces to form a coherent whole. To achieve this goal, standardized data formats and interface protocols must be adopted to ensure smooth data flow between different modules. For example, the input module should be able to receive information from external data sources (e.g., market data providers, economic indicator databases, etc.) in real time and seamlessly pass this information to the inference engine for analysis. After processing the data, the inference engine should feed the results back to the output module and the user interface in a timely manner, so that the user can obtain real-time risk assessment and decision-making recommendations.

Secondly, the maintenance of the system is equally critical. This process includes regular monitoring of the operational status of the system, assessing its performance and identifying and fixing potential problems in a timely manner. Maintenance should also include updating and upgrading the software to adapt to new market conditions and technological advances. For example, as new financial products emerge or market conditions change, inference algorithms may need to be adjusted or optimized to improve forecasting accuracy. In addition, user feedback is an important part of the maintenance process, and the system should regularly collect user experiences and suggestions for continuous improvement.

System security and stability are also aspects of the integration and maintenance process that cannot be ignored. Ensuring the security of data transmission and preventing data leakage and tampering are the basic requirements for protecting user information and system integrity. At the same time, the system should have good fault recovery capability and be able to quickly resume normal operation in the event of unforeseen circumstances. This can be achieved by implementing regular backup mechanisms and disaster recovery plans to ensure data security and system reliability.

Finally, as the financial market continues to evolve, the functional requirements of the DSS may change. Therefore, system integration and maintenance should also have a certain degree of flexibility and scalability to facilitate future expansion and optimization of functions according to user needs and market trends. By establishing an effective maintenance mechanism, DSS can continue to provide users with accurate decision-making support and enhance their ability to cope with the complex financial environment.

5 | Experimentation and Evaluation

5.1 Data Selection and Data Description

Under the efficient market hypothesis, all information of value is reflected in the financial market price movements in a timely, adequate and accurate manner. The stock market is regarded as the barometer of the economy, and the market price fluctuations of financial institutions can show the market risk in a timely and effective manner. In the empirical part of this paper, the stock data are selected as proxy variables to further explore the risk spillover effect.

In this paper, the selection of specific stock data, the current financial institutions listed companies are more, a separate study of which one or part of the listed company's data can not represent the overall industry, so the empirical research part of the data should be selected to reflect the specific circumstances of the industry to analyze in detail, this paper intends to select the industry sector index, this is because the industry sector index gathered the changes in the fluctuations of the stock price of the shares in the plate, can fully reflect the overall trend of the plate. According to the research object of this paper, the financial sub-industry is further subdivided into three core financial industries, namely banking, securities and insurance, which have irreplaceable roles in the national economy. The banking index, securities index and insurance index are selected as indicators of the business development of the financial sub-industry. These indices are all derived from the Shenwan Secondary Industry Index, which is based on the weighted average stock prices of listed companies in different industries according to the industry classification standards of the Securities and Futures Commission. At present, the weighted average industry stock price index has been adopted by many scholars in China, who recognize that this kind of index has representativeness and authority, and agree that it can show the real market operation situation of the industry. At the same time, CSI Financial Index is regarded as the representative of the financial industry system, and its construction rule is to be calculated by weighting the Shanghai and Shenzhen Stock Exchanges stocks that are closely related to the financial industry, such as financial information service, payment, investment and financing, etc., which have significant representativeness, so that it can more objectively present the overall operating condition of China's financial industry system.

In this paper, the closing price of the A-share market from January 5, 2016 to September 30, 2021, $(P_i t, P_{st}) = P_t$ as the sample data, and i denotes financial sub-sectors such as banking, securities and insurance. s denotes the CSI financial index, whose risk is used to measure the systematic financial sector risk. The logarithmic return is taken as the research object, and its calculation is announced as follows:

$$R_t = Ln\left(P_t / P_{t-1}\right)$$

R_t is the return of the industry index in period t. P_t , and P_{t-1} denote the period t and period t-1 of the three financial sub-industry indices and the financial market index, respectively. The closing prices of the industry indices in this paper are obtained from Flush iFinD.

5.2 Data Visualization

We plot the trend of financial sub-sectors such as banking, securities and insurance against the time series of systematic returns of the financial sector, using yh for banking, zq for securities, bx for insurance, and jr for financial sector, to observe the characteristics of their market returns.

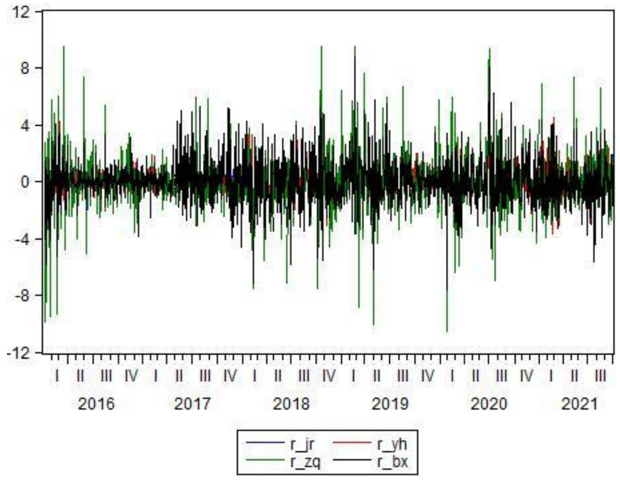

Figure 1 Time-series plot of banking, securities and insurance vs. financial sector system rates of return

The logarithmic returns of the financial system and the banking, securities and insurance sectors have been fluctuating up and down around the value of zero, with small fluctuations accompanied by small fluctuations and large fluctuations accompanied by large fluctuations.

The banking, securities and insurance sector has a more prominent volatility concentration. Due to the dramatic impact of the "meltdown mechanism" triggering the equity pledge crisis in early 2016, the escalating trade war between the US and China in 2018, and the sudden outbreak of the New Crown epidemic in the first quarter of 2020, the volatility of yields in the banking, securities, and insurance sectors and the financial system in 2016, 2018, and 2020 are all out of the ordinary.

There are very significant oscillations and fluctuations, showing obvious volatility clustering, in addition, the three major financial sub-sectors and the financial industry system have similar characteristics of the shape of the change in the rate of return, this synergistic fluctuation trend indicates that the financial sub-sectors on the financial system of the influence of the role of the relationship between the stronger and stronger, once the financial sub-sectors encountered risks are very likely to spread to the financial system, which may lead to a systemic financial risk.

Table 2 Description of the basic statistics of the yield series

	average value	maximum values	minimum value	(statistics) standard deviation	Skewness S	kurtosis K	JB Inspection
YH	0.006887	8.731280	-8.792649	1.177418	0.326711	7.528093	1220.079(0)
ZQ	-0.007104	9.531318	-10.52380	1.967549	0.145796	8.555581	1804.098(0)
BX	0.017407	8.762145	-7.647865	1.679438	0.216850	5.145799	279.3657(0)
JR	0.000959	8.542593	-8.289567	1.310419	-0.011932	7.994635	1454.199(0)

5.2 Risk Evaluation Process and Results

We bring the Bayesian network into the computation according to the risk classification.

In the first step, this paper first utilizes the Bayesian formula posterior probability to calculate, so as to derive the Bayesian probability distribution of each risk source, the specific Bayesian probability distribution is shown in the table:

Table3 Bayesian probability distribution for each risk source

		Risk impact value			weighted risk value			Total risk impact value
		State0	Statel	State2	State0	State1	State2	
market risk	State0	0.0652	0.0625	0.2273	0.03	0.02	0.05	0.2234
	Statel	0.1304	0.1563	0.5000	0.06	0.05	0.11	
policy risk	State0	0.1522	0.0938	0.1818	0.07	0.03	0.04	0.4156
	State1	0.1521	0.1874	0.045	0.07	0.06	0.01	
	State2	0.3043	0.6875	0.3636	0.14	0.22	0.08	
interest rate risk	State0	0.1087	0.3437	0.2727	0.05	0.11	0.06	0.2028
	Statel	0.2609	0.437	0.0909	0.12	0.14	0.02	
	State2	0.7608	0.7187	0.5454	0.35	0.23	0.12	
purchasing power exposures	State0	0.2174	0.5312	0.3182	0.1	0.17	0.07	0.2256
	Statel	0.0869	0.125	0.3451	0.04	0.04	0	
	State2	0.1739	0.9375	0.2345	0.08	0.3	0.22	
credit risk	State0	0.213	0.1563	0.2273	0.06	0.05	0.05	0.1785
	Statel	0.0217	0.2157	0.0454	0.01	0.03	0.01	
	State2	0.0652	0.1875	0.1364	0.03	0.06	0.03	
financial risk	State0	0.0652	0.2187	0.1818	0.03	0.07	0.04	0.1361
	Statel	0.0435	0.2187	0.2273	0.02	0.07	0.05	
	State2	0.1087	0.2187	0.5455	0.05	0.07	0.12	
business risk	State0	0.1956	0.0937	0.2727	0.09	0.03	0.06	0.2094
	Statel	0.0869	0.0313	0.1364	0.04	0.01	0.03	
	State2	0.0869	0.0625	0.0909	0.04	0.02	0.02	
shifting exposures	State0	0.1024	0.0937	0.1818	0.04	0.03	0.04	0.2618
	Statel	0.1304	0.0625	0.1818	0.06	0.02	0.04	
	State2	0.3478	0.375	0.1818	0.16	0.12	0.04	

We standardized the degree of impact of each risk source on the total risk using the standardization formula as shown in the table below.

Table 4 Impact Values for each Risk Source after Standardized Treatment

Risk attributes	probability of occurrence	Risk impact	Standardized process value
market risk	0.7014	0.2234	0.120548241

policy risk	0.3380	0.4156	0.224260738
interest rate risk	0.3878	0.2028	0.109432333
purchasing power risk	0.3388	0.2256	0.121735377
credit risk	0.2384	0.1785	0.096319879
financial risk	0.4304	0.1361	0.073440535
business risk	0.3104	0.2094	0.112993741
liquidity risk	0.3388	0.2618	0.141269156

5.2 Validity Analysis of Bayesian Network Models

According to the Bayesian network we constructed a Bayesian network model for project financing risk evaluation, so we need to progress to verify the effectiveness and rationality of the Bayesian network model. We convert the data collected beforehand into Bayesian probabilities and input them into the Bayesian network model, so as to verify the consistency of the impact of the occurrence of risk events on the overall risk of the item. Since the output of the model is a probability distribution, we choose the maximum probability of the overall risk when the probability of the risk of each different risk source is assumed to be 1. Therefore we first assume the maximum probability of occurrence of each risk source, and based on the results of the model thus derive the probability distribution when the overall risk is large.

Table 5 Bayesian network model validation results

Projected results	Low total risk probability	Of the total risk probability	High total risk probability
Low probability of each source of risk	0.969	0.019	0.010
Of the probability of each source of risk	0.050	0.909	0.039
High probability of each source of risk	0.092	0.004	0.901

When the probability of occurrence of each risk source is maximum, the overall risk is 97%; when the probability of occurrence of each risk source is medium, the overall risk is 91%; when the probability of occurrence of each risk source is small, the overall risk is 90%. Through the validation analysis, the model's prediction results are consistent, while the prediction results are relatively accurate. Therefore, the model is relatively good.

4 | Conclusion

This paper discusses the construction of a Bayesian network-based financial market risk prediction and decision support system, aiming to improve the accuracy and effectiveness of financial decision-making. Through the detailed design of each module of the system, including the input data module, inference engine, user interaction interface, as well as system integration and maintenance, we have built a comprehensive tool that can flexibly respond to the dynamic market environment.

The results show that Bayesian networks have significant advantages in dealing with uncertainty and modeling complex dependencies, enabling the system to provide accurate risk assessments based on real-time data. The design of the user interaction interface enhances the usability and user experience of the system, enabling decision makers to quickly understand and apply the risk assessment results. In addition, the system integration and maintenance strategy ensures the stability and scalability of the DSS for future enhancements and market adaptation.

Despite the results achieved in this study in the area of financial market risk management, there are still some limitations. For example, the performance of the model depends on the quality and completeness of the input data and may face computational complexity issues when dealing with large-scale data. Future research could further explore the incorporation of other machine learning techniques to enhance the system's predictive and real-time response capabilities.

Overall, the decision support system based on Bayesian network provides an effective solution for risk management in the financial market and has a broad application prospect. With the continuous progress of technology and changes in the market environment, it is expected that the system can provide more scientific and effective support for decision makers and promote the sustainable development of the financial field.

References

[1] Bank for International Settlements.78th BIS Annual Report, 2008; 79th BIS Annual Report, 2009.

[2] Zhang, Xiaopu. Systemic financial risk research: evolution, causes and regulation[J]. International Financial Studies,2010(7):58-67.

[3] Ben Bernanke.A Letter to Sen.Bob Corke.The Wall Street Journal,2009.

[4] F. Kaufman, G. Bank contagion: a review of the theory and evidence.Journal of Financial Services Research, 1994, 8(2):123-150.

[5] Zigrand J P. Systems and systemic risk in finance and economics [D]. London School of Economics and Political Science, LSE Library, 2014.

[6] BIS, FSB, IMF. Guidance to assess the systemic importance of financial institutions, markets and instruments: initial considerations [R]. BIS Meeting, 2009.

[7] Diamond D W,Dybvig P H.Bank runs,deposit insurance,and liquidity[J]. The journal of political economy,1983,91(3):401-419.

[8] Alessi L,Detken C.Quasi real time early warning indicators for costly asset price boom/bust cycles:A role for global liquidity[J]. European Journal of Political Economy, 2011,27(3): 520-533.

[9] Duca M L,Peltonen T A.Assessing systemic risks and predicting systemic events[J].Journal of Banking&Finance,2013,37(7):83-95.

[10] Freixas X,Rochet P J C. Part 2: what should central banks do?systemic risk interbank relations, and liquidity provision by the central bank[J]. Journal of Money Credit and Banking, 2000,32(3): 611-638.

[11] Iori G,Jafarey S,Padilla F G.Systemic risk on the interbank market[J]. Journal of Economic Behavior & Organization,2006,61(4):525-542.

[12] Ouyang Hongbing,Liu Xiaodong. Analysis of systemic importance and systemic risk contagion mechanism of financial institutions in China - A perspective based on complex networks[J]. China Management Science,2015,23(10):30-37.

[13] ZHU Yuanqian,MIAO Yufeng. A review of models on systemic risk measurement and early warning[J]. International Financial Studies, 2012(1): 79-88.