Journal of Information Technology Management

Print ISSN: 2008-5893

Online ISSN: 2423-5059

An Effective Model for Ontology Relations Efficacy on Stock prices: A Case Study of the Persian Stock Market

Mohammad Hossein Samani

PhD Candidate, Faculty of Industrial and Systems Engineering, Tarbiat Modares University, Tehran, Iran. E-mail: mhsamani@modares.ac.ir

Amir Albadvi *

*Corresponding author, Professor, Faculty of Industrial and Systems Engineering, Tarbiat Modares University, Tehran, Iran. E-mail: albadvi@modares.ac.ir

Abstract

The unpredictability of the stock market makes it a serious area of study and analysis. With the help of the accumulated information available in the current digital age and the power of high-performance computing machines, there is a great focus on using these capabilities to design algorithms that can learn stock market trends and successfully predict stock prices. The main goal is to create an intelligent system that provides these features for predicting short-term stock price trends to facilitate the investment decision process. To increase the accuracy and productivity of these systems and facilitate the routine of using common-sense knowledge in machine learning systems, developing or enriching knowledge bases and ontology for market modeling will be one of the effective measures in this field. In this research, an attempt has been made to strengthen and enrich the basic ontology created by the authors by using other global ontologies related to the subject of the stock market, and parts of the target space that were not addressed have been added to the ontology. By combining reference ontologies, a level of standardization is also created for the ontology and stability in the representation of concepts and relationships is ensured. In the next step, it has been tried to test the impact of the concepts and relations of the ontology in predicting stock price movements. For this purpose, news in the field of economy is considered as input and a model is created that first filters the textual inputs related to the desired stock symbol and then observes their effect on the price changes of the related stock. After improving the performance and comprehensiveness of the ontology, the study conducted in this report presented a model to measure and prove the effect of the relationships in this ontology on price changes. In practice, according to human limitations and the tools used, this effect was observed and confirmed with a proper level of certainty by checking the economic news.

Keywords: Stock forecasting, Stock exchange, Financial markets, Ontology.

Journal of Information Technology Management, 2024, Vol. 16, Issue 3, pp. 1-20
Published by University of Tehran, Faculty of Management
doi: https:// 10.22059/jitm.2024.373635.3640
Article Type: Research Paper

Received: February 06, 2024
Received in revised form: May 06, 2024
Accepted: May 18, 2024
Published online: August 26, 2024

Introduction

It is very difficult to predict stock and commodity price indexes due to many influential uncertainties. With the help of the accumulated information available in the current digital age and the power of high-performance computing machines, there is a lot of focus on designing algorithms that can learn stock market trends and successfully predict stock prices. Therefore, to increase the accuracy and productivity of these systems and facilitate the use of common knowledge in machine learning systems, developing or enriching knowledge bases and ontology for market modeling will be one of the most effective measures in this field.

If we want to have a basic definition, ontology as a word is a combination of the two words onto (being) and logy (study). Ontology has emerged initially in the knowledge of philosophy to study the existence of world entities through a comprehensive classification of them. Ontology is used in computer science, especially in semantics, to model and display the meaning of words and the relationship between them in a desired context. In general, it can be said that ontology is an understandable expression of the knowledge concepts of an arbitrary field, which is presented to create the possibility of sharing and reusing knowledge. Due to the importance and role of ontology in knowledge-based systems, various definitions of ontology have been given in different references. The definition we are looking for, which was presented by Gruber in 1993 (Gruber, 1993), is: "Ontology, is an explicit specification of a conceptualization".

Creating an ontology is a time-consuming and precise activity, and it requires the participation of experts in the two fields of knowledge engineering and ontology, as well as experts in the desired field to create an ontology. The best method is the semi-automatic development of ontology with little human intervention (learning ontology). In this regard, it is generally necessary to use knowledge discovery, natural language processing, information retrieval, machine learning, deduction, and database management.

To improve the accuracy of stock market forecasting using ontology, the authors of this research previously developed the Persian ontology of the stock market and financial markets. This ontology comprises 565 concepts, 496 hierarchical relationships, 137 non-hierarchical relationships, and 937 examples. It was produced as one of the outputs of a thesis at Tarbiat Modares University (Samani et al., 2023). In this research, at first, it was tried to strengthen and enrich the created ontology by using other global ontologies related to the subject of the stock market, and to add parts of the target space that were not addressed to the ontology. By combining reference ontologies, a level of standardization was also created for the ontology and stability in the representation of concepts and relationships is ensured.

In the next step, we tried to test the impact of the concepts in the ontology in predicting stock price movements. For this purpose, news in the field of stock exchange and financial markets as well as news trends created in social networks were considered as input, and a model was created that first filters the textual inputs related to the desired stock symbol and their impact on the changes in the mentioned stock prices can be seen.

The rest of this article is structured as follows: Section 2 reviews related work. Section 3 compares and summarizes the reference ontologies, and it creates a model to discover the impact of ontology relationships on stock prices and explains the steps of its production. In section 4, the details of the evaluations and the discussion of the paper are presented. Finally, in section 5, a summary is made and a conclusion is made.

Literature Review

In this section, we review related research in two main categories: 1) methods expressing the application of ontology in financial markets, and 2) existing knowledge sources related to financial markets.

Application of ontology in financial markets

In this field, according to current knowledge, relatively few research studies have been published, which are introduced below.

In the research conducted by Wang et al. (2008), the impact of various types of news on business activities in financial markets, such as trading price fluctuations, trading volume, and frequency of trading, was measured by presenting an ontology for news related to financial instruments. This ontology consists of two parts: The first part presents a hierarchical framework for domain knowledge, primarily including news classes, market participants classes, financial market classes, and basic relations between these classes. In the second part, a causal map is used to illustrate how classes of news are related to areas of these markets. The process of creating the ontology to demonstrate the concept was conducted manually and in a limited manner using news related to terrorist attacks and their impact on the capital markets. Finally, the study concluded that: First, understanding the knowledge of news in financial fields is a complex and comprehensive process. Second, building this knowledge base helps in constructing trading models based on news in financial instrument markets. Third, developed systems such as news-based stock price prediction systems and support systems for financial market participants searching for relevant news are facilitated and supported by ontology.

In a study conducted by Zhang et al. (2019), explicit knowledge was sought to be extracted from news sources. Based on the general mechanism of domain knowledge, knowledge reasoning and knowledge discovery, this paper develops a framework for discovering implicit knowledge from news and applying knowledge in stock forecasting. At first, according to the characteristics of the financial field and the conceptual cube, the conceptual structure of the industry-company-product is constructed and the ontological framework is proposed in a limited field (polyester industry). Second, by constructing the ontology of the financial field, the financial news knowledge management framework is proposed. In addition, using the ontology features and domain rules extracted from the news text, a financial news knowledge reasoning mechanism is built to achieve financial news knowledge discovery, and a rule-based model for stock price prediction is presented, which is claimed to have a proper performance in empirical analysis of the polyester industry.

In the study by Zaki et al. (2019), it is noted that there is a scarcity of research focusing on the development of ontologies, particularly in the areas of stock markets and fraud detection within financial securities and transactions. This paper attempts to address this gap by providing a systematic framework for developing and managing an ontology for stock markets. To achieve this goal, researchers use techniques to gather and integrate domain knowledge from existing sources alongside new knowledge obtained from unstructured source analysis based on stock market fraud cases reported by the Securities and Exchange Commission. The proposed framework and knowledge are built through six case studies that evaluate work in different types of stock market maintenance and monitoring programs.

In the study by Zdraveski et al. (2017), a data model for a stock market was developed. This model is based on a built ontology for the stock market that covers the entities and relationships in the stock market and its stakeholders. The article, which is written very briefly, states that the main purpose of using this ontology is to connect specific stock market and stock transaction data with general company information from DBPedia, and it is expected that merging these two datasets will lead to more accurate statistics or new indicators that help traders see a better big picture of companies and stocks being traded.

In the study by Sharma et al. (2023), it is stated that the news related to the stock market in the last decade has a vital role for brokers or users, and therefore the focus of this article is on predicting the sentiments of the stock market news based on their polarity and textual information using a Convolutional Neural Network (CNN), which is based on ontology knowledge. The details of this ontology are not mentioned in the article. However, the ontology is built based on users' information, which affects the polarity of opinions.

The paper by Lamparter et al. (2006) describes the design procedure of an ontology-based marketplace for trading web services. Based on the analysis of the requirements to create a web service market, a market is designed that has these advantages: The market uses a communication language based on ontology, which can represent the semantics of described requests, offers, and agreements. In addition, semantic information has been used to divide the entire market into several independent submarkets. This concept is shown to be more efficient than the existing mechanism, it can be said that the use of background knowledge has reduced the overall complexity.

In the study by Mishra et al. (2023), it is stated that the stock market depends on various factors such as price, volume, etc. that these factors can be easily identified by indicators and also in the literature, there are various indicators available that can be used to reduce risk in the stock market within investing or trading. However, none of the indicators alone can provide complete protection against capital risk. Therefore, it is necessary to either create a new index or propose a framework that consists of a set of several indices that can reduce the investor's risk compared to any of the existing indices. This paper aims to propose an ontology-based framework that simultaneously presents the characteristics of three indicators, i.e., Bollinger Bands, Fibonacci Bands, and Heikin-Ashi. In this paper, these indicators are mapped with an ontology in the form of classes, features, and examples.

According to the reviewed research and the confirmation of the importance of knowledge in identifying factors and characteristics influencing the price prediction process or the direction of stock movement in the research, as well as the lack of comprehensive knowledge bases, especially in the domestic stock market (Iran), the necessity of creating and using a comprehensive ontology for Persian financial markets seems to be confirmed. As mentioned in the introduction, to use ontology in more accurate forecasting of the domestic stock market, previous Persian ontology of the stock market and financial markets, which includes 565 concepts, 496 hierarchical relationships, 137 non-hierarchical relationships, and 937 examples, has been developed by the authors of this research (Samani et al., 2023). In the present research, in order to create comprehensiveness and standardization of the created ontology, an attempt has been made to strengthen and enrich the ontology using other global ontologies related to the subject of the stock market, and the necessary resources to do this have been reviewed in section 2.2.

Introduction of knowledge resources related to financial markets

In the scientific space of financial fields, many knowledge bases have been introduced and used. In general, there is no definitive or universally agreed-upon knowledge base that all stock market analyses refer to. The choice of knowledge base depends on various factors, including the specific research objectives, the expertise of the domain analysts, and the availability of relevant resources (Chen et al., 2021).

However, when performing stock market analysis, analysts often rely on a combination of different sources of knowledge, including:

1. Financial databases: These are comprehensive repositories of financial data, such as stock prices, historical performance, financial statements, and company profiles. Examples include Bloomberg, Thomson Reuters, and Yahoo Finance. While these databases are not fully knowledge bases, they provide valuable information for analysis (Zhao et al., 2023).

2. Economic Indicators: Analysts often consider macroeconomic indicators such as GDP growth, interest rates, inflation rates, and employment data. These indicators provide insights into the overall health of the economy and can influence stock market trends (Omar et al., 2022).

3. Financial Research Reports: Analysts rely on research reports from financial institutions, investment banks, and independent research firms. These reports provide in-depth analysis, forecasts, and recommendations on specific stocks or sectors (Chen et al., 2021b).

4. Academic Research and Publications: Researchers may refer to academic articles, journals, and publications that focus on stock market analysis, financial modelling, and related topics. These resources contribute to the existing body of knowledge and provide insights into various quantitative and qualitative analysis techniques (Gopinathan et al., 2019).

5. Regulatory information: Analysts often refer to regulatory filings such as annual reports, quarterly reports, and proxy statements. These filings contain important information about the financial performance, risks, and governance of publicly traded companies (Odilon Tilly Sardou, 2021).

6. Industry-Specific News and Reports: Analysts may follow industry-specific sources of news, market reports, and expert opinions to learn about developments and trends in specific sectors or industries (Mohti et al., 2019).

The important point is that the field of stock market analysis is dynamic and new sources and knowledge bases emerge over time. Analysts constantly review and evaluate various sources of information to enhance their understanding and make informed investment decisions. The mentioned items are the input of the created anthologies in this field, the most important of which are mentioned below.

1. Financial Industry Business Ontology (FIBO): It is a comprehensive ontology that covers a wide range of financial concepts including stocks, bonds, currencies, and commodities. This knowledge base provides a structured representation of financial data and can be used to facilitate the integration of financial information in different systems (Bennett, 2015).

2. Financial Instrument Global Identifier (FIGI) Ontology: The FIGI ontology is designed to provide a standard identifier for financial instruments such as stocks, bonds, and options. It can

be used to link different sources of financial data and provide a consistent method to identify financial instruments in different systems (Houstoun et al., 2014).

3. Stock Market Ontology (SMO): A domain-specific ontology that provides a structured representation of stock market data. It covers concepts such as stock price, trading volume, and company finances. It can be used to facilitate the integration of stock market data in different systems and provide a basis for stock price prediction models (Zdraveski et al., 2017).

In the following, these ontologies will be described to use them in enriching the basic ontology.

Financial Industry Business Ontology (FIBO)

This knowledge base is a standard and comprehensive ontology that aims to define concepts, relationships, and laws in the field of the financial industry. FIBO provides a common language and framework for the presentation and exchange of financial information across different systems and organizations and covers a wide range of financial areas, including securities, derivatives, banking, loans, and corporate actions. This ontology is designed to depict complex relationships between financial entities such as issuers, investors, intermediaries, and regulatory bodies, and defines concepts such as financial instruments, parties involved in financial transactions, market data, pricing, and risk management. and incorporates industry best practices, standards, and regulations to ensure consistency and accuracy in the presentation of financial data (Bennett, 2015).

FIBO can be used in a variety of applications in the financial industry, including regulatory compliance, risk management, data integration, and financial analytics. By using FIBO, financial institutions can increase data interoperability, reduce data integration costs, and improve overall data quality and consistency.

Overall, FIBO serves as a basic framework for modeling and presenting financial knowledge, enabling efficient data management and facilitating interoperability within the financial industry, and offers a rich set of relationships and rules (including ownership, classification, aggregation, dependency, law, and time) that will enable detailed visualization and analysis of financial data and support various use cases in the financial industry such as risk management, regulatory compliance, and financial reporting (Bennett, 2015). An example of the relationships in this ontology is given in Table 1.

Table 1. Example of FIBO ontology relationships

The name of the	Example
Ownership relation	Company A owns 50% of the shares of Company X.
Classification relation	Financial instrument Z is classified as a bond in the fixed income asset class.
Aggregation relation	Portfolio P1 aggregates the performance of individual financial accounts A1, A2
Dependency relation	The valuation of contract A depends on the market data M1 and M2.
Rule-based relation	Credit risk limit rule: The exposure to a counterparty must not exceed a specified
Temporal relation	The historical price of Stock A on a specific date was $100 per share.

Financial Instrument Global Identifier (FIGI) Ontology

The Financial Instrument Global Identifier (FIGI) Ontology is a standardized framework that provides a consistent and unique identification system for financial instruments. It is designed to facilitate the accurate and efficient identification and classification of financial instruments across various systems and platforms (Houstoun et al., 2014).

The FIGI Ontology defines a set of concepts, relationships, and attributes to represent the characteristics and metadata associated with financial instruments. It enables the integration and interoperability of financial instrument data by providing a common language for identifying and categorizing instruments.

The ontology includes concepts such as FIGI codes, security types, asset classes, exchange codes, and issuer information. It captures the relationships between these concepts, allowing for the classification, grouping, and querying of financial instruments based on their attributes.

The FIGI Ontology supports various use cases in the financial industry, including trading, portfolio management, risk analysis, and regulatory reporting. By utilizing FIGI codes and the associated ontology, market participants can achieve accurate and standardized identification of financial instruments, leading to improved data quality, reduced operational risk, and enhanced data integration capabilities.

Overall, the FIGI Ontology provides a robust and standardized framework for identifying, classifying, and organizing financial instruments. It serves as a foundation for efficient data management, enabling seamless integration and interoperability of financial instrument data across different systems and platforms. The relations and rules in the FIGI ontology (including Is-a, Has attribute, reference, is-denominated-in, Has exchange, and rule-based) enable the precise identification, classification, and organization of financial instruments. By using these relationships and rules, market participants can ensure stability as well as interoperability and improve data quality in identifying financial instruments in different systems and frameworks (Houstoun et al., 2014). An example of the relationships in this ontology is given in Table 2.

Table 2. Example of FIGI ontology relationships

The name of the relationship	Example
Is-a relation	Equity ABC is an equity instrument.
References relation	Equity ABC references Issuer Company X.
Is-denominated-in relation	Bond XYZ is denominated in EUR.
Has-exchange relation	Equity ABC is traded on Exchange A.
Rule-based relation	FIGI code rule: A FIGI code must follow the format AAAA99999999.

Stock Market Ontology (SMO)

The Stock Market Ontology is a knowledge representation framework that aims to capture the concepts, relationships, and rules related to stock markets and their associated entities. It provides a standardized structure for modeling and analyzing stock market data, enabling better understanding, integration, and utilization of information in the domain of stock trading and investment (Zdraveski et al., 2017).

The ontology encompasses various key components of stock markets, including stocks, exchanges, trading activities, market data, market participants, and regulatory aspects. It defines concepts such as stock symbols, stock prices, trading volumes, market indices, trading orders, and investor profiles.

The Stock Market Ontology captures the relationships between these concepts, allowing for the representation of stock market dynamics and interactions. It represents relationships such as stock ownership, trading relationships between buyers and sellers, market data dependencies, and regulatory compliance rules.

By utilizing the Stock Market Ontology, stakeholders in the stock market domain, including investors, traders, analysts, and regulators, can gain insights into market behavior, perform data analysis, develop trading strategies, and monitor compliance. It provides a standardized framework that promotes interoperability and data integration across different stock market systems and platforms.

Overall, the Stock Market Ontology serves as a foundation for representing and analyzing stock market-related information. It enables a better understanding of stock market dynamics, facilitates data integration, and supports various applications in stock trading, investment analysis, and regulatory compliance. The relationships and rules in SMO (including ownership, trading, market data, regulatory, classification, financial performance and time-based relationships) enable the display, analysis and integration of stock market data. Using these relationships and rules, stock market stakeholders can gain insights, perform data analysis, develop investment strategies, and ensure compliance with regulatory requirements. (Zdraveski et al., 2017) An example of the relationships in this ontology is given in Table 3.

Table 3. Example of SMO ontology relationships

The name of the	Example
Ownership relation	Institutional investor ABC holds a 10% stake in Company XYZ.
Trading relations	Institutional investor XYZ executes a trade on Exchange A.
Market data relations	Stock price of Company ABC influences the performance of Market Index X.
Regulatory relations	Trades above a certain threshold must be reported to the regulatory authority.
Classification relations	Stock XYZ is categorized under the Technology sector.
Financial performance	Stock XYZ's price-to-earnings ratio is 15.
Time-based relations	Moving averages provide insights into short-term and long-term stock price

Methodology

In this research, the main aim is to improve the ontology of the Persian stock market made in (Samani et al., 2023) and to create a model of the influence of relationships on stock prices. In the first part, a comparison of the reference ontologies is made and the details of the enriched ontology are stated. In the second part, how to use ontology of the stock market and financial markets to filter textual inputs and measure their impact on stock prices is described.

Comparison and summary of reference ontologies

According to the available sources that were introduced in section 2.2, in this section, a comparison of these three enrichment references is presented in Table 4.

Table 4. Comparison of reference ontologies

	FIBO	FIGI	SMO
Scope	Focuses on representing concepts, relationships, and rules within the financial industry, covering areas such as banking, insurance, securities, and regulations.	Specifically designed to represent and manage unique identifiers for financial instruments globally, providing a standardized approach to identify and track financial instruments.	Targets the representation of concepts, relationships, and rules within stock markets, including stocks, exchanges, trading activities, market data, and regulatory aspects.
Primary Focus	Covers a wide range of financial industry domains, including banking, insurance, securities, and regulations, with a focus on standardization and interoperability.	Primarily focuses on the identification and tracking of financial instruments through the use of unique identifiers.	Concentrates on the representation and analysis of stock market data, including stocks, exchanges, trading activities, market data, and market participants.
Key Concepts	Concepts include financial instruments, market participants, financial transactions, regulations, legal entities, and financial markets.	Concepts revolve around financial instruments, unique identifiers, instrument attributes, and standardization.	Concepts encompass stocks, exchanges, trading activities, market data, market indices, market participants, and regulatory aspects.

Relations and Rules	Represents relations and rules related to financial transactions, market regulations, legal entities, and data standardization within the financial industry.	Focuses on relations and rules associated with identifying and tracking financial instruments through unique identifiers and instrument attributes.	Captures relations and rules related to stock trading, market data dependencies, market participants, regulatory compliance, and financial performance.
Applications	Used for data integration, semantic interoperability, regulatory compliance, risk management, and financial analysis within the financial industry.	Utilized for unique identification and tracking of financial instruments, improving data quality and streamlining financial operations.	Applied in stock trading, investment analysis, market research, trading strategies, and regulatory compliance within the stock market domain.

According to the review and analysis of these anthology, some measures have been taken to enrich it. In the first stage, the desired domains that were neglected in the proposed basic ontology were checked and added. Then, concepts and their classification were updated and new concepts were added. In the next step, the existing relations will be expanded according to the existing relations in reference ontologies. For this purpose, new domain relationships are formed and new relationships are rebuilt among different domains. New rules or principles are added in ontology to apply constraints or infer new knowledge. These rules can redefine the logical relationships, the restrictions related to the values of the attributes or requirements based on concepts and relationships existed in the ontology. These actions have been done semi-automatically and with human supervision. Here we use all 3 mentioned ontologies to enrich our base ontology.

The basic ontology created by the authors included 565 concepts, 496 hierarchical relationships, 137 non-hierarchical relationships, and 937 examples (Samani et al., 2023, after merging it with reference ontologies and reconstructing the concepts and relationships based on the new domains and ranges, the new ontology includes 1142 concepts, 543 types of hierarchical relationships, 211 non-hierarchical relationships, and 1823 examples.

Creating a model to discover the impact of ontology relations on stock prices

In this section, the process of creating a model for identifying the impact of knowledge extracted from ontology on stock price changes will be discussed. To create this model, we will first need to check the text entries to detect the connection with the desired stock symbols. Therefore, for each symbol (for each company's stock), a one-week interval is considered, and we consider the set of economic news of that week as the primary data set. Now, for each news, we have to do two main pre-processing:

1- Extracting keywords and important words of each news text, to match the concepts and examples in the ontology.
2- Identifying the polarity of each news text using sentiment analysis methods

In the following, we describe these sections.

keywords extraction

To extract the important words and expressions of the text, we tokenize them, these tokens are filtered based on the concepts and examples extracted from the ontology, meaning that if a single or multi-word phrase in the text is found in one of the examples and concepts of ontology, that term is considered as a token. In the next step, we calculate the number of occurrences of each tokenized expression. In this step, the 5 most frequent terms (which repeat more than 2 times) of the text are extracted.

Determining the polarity of the news text

The purpose of sentiment analysis projects is to categorize a desired comment or text generally into three categories: positive and negative or neutral, but in this application, it is necessary to recognize the polarity of the input text in a 2way approach: positive way (effect on increasing the price) and negative way (effect on price reduction). therefore, a total of 2000 thousand news from the following three economic sites were collected and labeled:

- https://www.eghtesadnews.com/
- https://www.eghtesadonline.com/
- https://donya-e-eqtesad.com/

These websites were chosen because they represent good examples of common economic news sites. They maintain archives of daily news in each section, and all domain-related news has been archived.

To do this activity, first input texts were pre-processed (normalization, tokenization, removal of redundant and stop words and punctuation marks). Next, the process of extracting the features of the text is done. For this purpose, different word embedding methods are used in academic researches, but here the sentiment analysis module is implemented using FastText and Sklearn libraries in Python programming language. More precisely, a ft.supervised classifier was trained on labeled data and evaluated with 10 folds. The output accuracy of the desired system was measured as 84.8%.

Preparation of ontology relations

In the next step, after the preparation of each text, the polarity determination process is performed according to the concepts in all hierarchical and non-hierarchical relations of ontology. Here, the goal is to determine for existing relationships whether the influence between two concepts will work directly or inversely. For example, the growth of foreign exchange rates has direct relation to the increase in the stock prices of companies that have imported raw materials. Also, the decrease in the price of the products produced in an industry affects the decrease in the value of the stocks of the companies in that industry. On the other hand, the excessive growth of the share in the reopening of transactions will lead to the suspension of transactions and cancellation of the price. Therefore, the relations in the ontology should receive direct and reverse labels with human supervision. All hierarchical relationships and most non-hierarchical relationships in this activity will be labeled directly, and only relationships such as "stops", "does not affect", and "cancels" will be labeled inversely.

Filtering Texts Using Ontology

At this stage, after pre-processing, a set of filtered news is formed for each stock symbol and each working week. In this way, the entire collection of news texts of that week will be the input, and according to the sharing of the keywords of each news text and the examples and concepts in the ontology, a subset of the news related to the mentioned symbol will be pruned in that week.

Modelling impact of news on stock prices using extracted knowledge

After creating a subset of filtered news, the impact of the news is predicted for each news according to its positivity and negativity, as well as the positivity and negativity of the relationship that exists in the ontology with the example of the stock symbol or its parent concepts in the ontology.

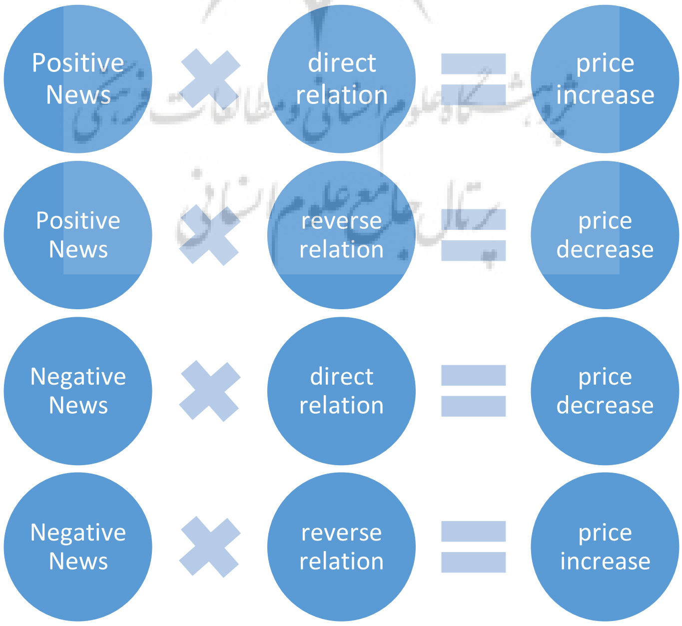

Figure 1. Determining the impact of the news on stock price

Figure 1 is a symbolic representation of how positive and negative impacts are calculated in this research. As presented in the figure, for each stock symbol, according to the news of the last week and its sentiment, the trend of price change was observed and it is determined whether the price of that week is increasing or decreasing. Finally, the accuracy of the prediction is evaluated based on whether the real price is increasing or decreasing.

Results

Enriched ontology evaluation

According to Raad et al. (2015), an ontology can generally be evaluated through different approaches, such as comparison with a reference standard, data-driven evaluation, user-based evaluation, and applied (task-oriented) evaluation. According to the knowledge of the authors, there is no other ontology for the domain of stock exchange and financial market of Iran, so the first option, that is comparing it with the reference standard, is not possible. Furthermore, data-driven evaluation is the process of comparing the ontology with existing data about the domain that the ontology models.

In this section, this evaluation is compared with the values obtained for the base system (Samani et al., 2023) with the new values obtained by integrating the reference ontologies. The last case is also practical, which will be used and evaluated after using this ontology in the automatic analysis system of the stock market, part of which will be examined in the fourth chapter and the other part in future works and not available at the time of writing this article.

In addition, to provide a better evaluation of the introduced ontology, we also considered the criteria introduced in several other studies:

Firstly, the measure introduced by Fernández et al. (2009) involves examining the depth and breadth of the ontology. According to studies, it is stated that among the different measures of depth and breadth, the most important are the variance of breadth and variance of depth, and the best ontologies are generally those that have higher values of variance of depth and breadth in their structure. The depth and breadth of the ontology for the basic model and merged model are given in Table 5.

The second measure, as stated in the reference by Batet et al. (2014), involves using the semantic distance between two classes. This criterion is based on the comparison of concepts and their semantic commonalities. Based on this, a feature-based metric is proposed that measures semantic distance based on the number of fathers in the non-shared hierarchy divided by the total number of fathers. The relationships used are given in Figure 2. In these relations, T(c) is the set of fathers of a concept including itself.

$$d(c_1,c_2) = \log_2\left(1+\frac{|T(c_1)\cup T(c_2)|-|T(c_1)\cap T(c_2)|}{|T(c_1)\cup T(c_2)|}\right)$$

$$Dispersion(O) = \sqrt{\frac{\sum_{c_i \in C} d(c_i, Root(O))^2}{|C|}}$$

Figure 2. Relationships used to calculate semantic sparsity

As concluded by the authors of the article, higher values of sparsity indicate a proper distribution of concepts in the ontology. The distribution of the base and pooled ontologies is given in the evaluation table.

Table 5. Values of ontology evaluation measures for basic and merged model

EVALUATION MEASURE	VALUE			
PRECISION	85.7% >> 91.4%			
THE DEPTH AND BREADTH OF THE ONTOLOGY BASED ON (FERNÁNDEZ ET AL., 2009)		Minimum	Maximum	Average
	Depth	2 >> 2	9 >> 11	4.2 >> 6.4
	Breadth	2 >> 2	72 >> 76	3.6 >> 5.7
SEMANTIC SPARSITY (BATET ET AL., 2014)	82.0% >> 86.7%			

Evaluation of the built model to discover the impact of ontological relationships on stock prices

According to what was mentioned in the previous section, the period of one year from March 2022 to March 2023 was selected for the evaluation of 6 stock symbols from the industrial group of basic metals in the Tehran Stock Exchange. This group was selected because the authors wanted to choose a narrow subdomain from original basic industries with good comprehensiveness and versatility. The 'basic metals' industry is the most valuable in the Tehran stock market, accounting for 17% of the capital market value. Among this group, the most important symbols have been selected considering comprehensiveness, aiming to include both small and large shares (6 symbols: Foolad, Zob, Fakhuz, Febahonar, Femelli, and Fasmin). No sampling method has been used; all historical data has been gathered.

Currently, the basic metals group is known as the second largest group in this market, having more than 17% of the capital market value. The total number of reviewed news from the three selected web sources mentioned in the previous section was equal to 121,622, and the filtering process was done for each stock symbol. For each symbol, the forecasting process was carried out in 49 working weeks, and the accuracy criterion was calculated by comparing it with the direction of the actual price movement.

The accuracy criterion has been used to predict each share according to the degree of connection. You can see how to calculate the accuracy for each symbol below (Metz, n.d.1987).

$$accuracy = \frac{number\ of\ true\ positive\ of\ forcasts\ +\ number\ of\ true\ negative\ forcasts}{number\ of\ all\ forcasts}$$

In this way, for each working week and each symbol, for example, several news have been filtered, some of which predict the direction of the stock price correct and some incorrect. You can see the prediction accuracy output for the stock symbols in the table.

Table 6. Evaluation of prediction accuracy according to ontological relationships

symbol	Number of filtered news	Average number of predictions (news) per week	Market value/thousand billion tomans	accuracy (percentage)
Foolad	16824	343	452.8	70.85%
Zob	14101	288	26.8	67.01%
Fakhuz	11852	242	71.7	64.04%
Febahonar	11741	240	13.1	60.41%
Femelli	13952	285	418.8	68.07%
Fasmin	10870	222	16.8	63.96%
Total	**79340**	**1620**	**1000**	**66.17%**

As can be seen in Table 6, the measured accuracy for these 6 symbols is set at around 65%. Considering that a method for polarity analysis with 85% accuracy has been used in the pre-processing stages of texts, it is evident that the overall accuracy of the results will increase if the accuracy of polarity detection improves. Additionally, employing human data entry for impact detection systems is likely to further enhance the accuracy of predictions.

According to these results, the most important and most probable conclusion that can be drawn is that the relations in the ontology have a direct and indirect effect on the forecasting process. In other words, the identification of influencing factors on stock price changes is facilitated by using knowledge resources such as ontology. Considering the proof of the effect of these relations on the direction of stock price movement, it is possible to embed this enriched knowledge base in machine learning models and deep learning methods in future studies. The general evaluation of the prediction can be seen in Figure 3.

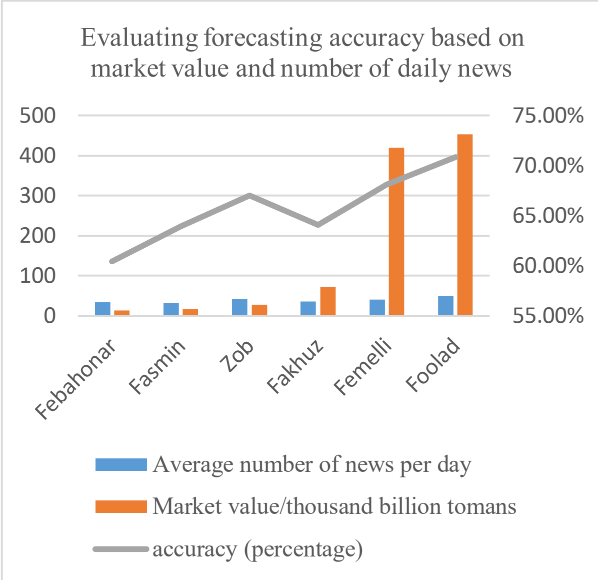

Figure 3. Evaluating forecasting accuracy based on market value and number of daily news

Another point that can be seen from these observations and predictions in Figure 3 is about the stocks of large and high-margin companies. In general, the bigger the market value of a symbol, the more the number of news related to them in the economic news, and therefore the impact of news and relationships on them is more visible, and as a result, better predictions are provided. For example, the two symbols Femelli and Foolad, which are considered index-making stocks, have had a higher number of news and, consequently, the number of predictions and accuracy compared to the symbols of Febahonar and Fasmin, which are considered smaller stocks. Of course, the type of supply chain and the metal used as raw material also play a role in these changes, and their impact must also be measured. The same observation is seen for Zob symbol, which has a relatively lower market value, because Zob Ahan Company is a newsmaker, it can be seen that the number of news related to the Zob symbol is high compared to other symbols with low market value.

Discussion and Conclusion

A stock market is a place where the shares of companies that are publicly listed are traded. The unpredictability of the stock market makes it a serious area of study and analysis. In these types of markets, an investor's success depends on the quality of information used to support decision-making and how quickly decisions can be made. Due to the practical importance of investment decisions, financial market developments in recent decades have been extensively studied in the fields of financial sciences, engineering, and mathematics (Cavalcante et al., 2016). Algorithm predictions based on traditional technical analysis in stock prices are generally founded on statistical analysis of data related to the price and volume of stock transactions. Stock markets, though unpredictable, are said to be influenced by certain factors such as trends in stocks over time, news trends, social media, and many other factors.

Therefore, by expanding the range of factors that affect stock prices, more accuracy can be obtained.

It is very difficult to predict stock and commodity price indexes due to many influential uncertainties. With the help of the accumulated information available in the current digital age and the power of high-performance computing machines, there is a lot of focus on designing algorithms that can learn stock market trends and successfully predict stock prices. Therefore, to increase the accuracy and productivity of these systems and facilitate the routine of using common-sense knowledge in machine learning systems, developing or enriching knowledge bases and ontology for market modeling will be one of the effective measures in this field.

In this research, in the first stage, an attempt has been made to enrich and strength the ontology output from the first study by comparing it with the international reference ontologies. As mentioned earlier, the basic ontology includes 565 concepts, 496 hierarchical relationships, 137 non-hierarchical relationships, and 937 examples. After merging with the reference ontologies, it included 1142 concepts, 543 types of hierarchical relationships, 211 non-hierarchical relationships, and 1823 samples.

In the next step, to create a model for discovering the impact of ontological relations on price changes, textual inputs were filtered using relationships, concepts, and examples in the ontology. A method was employed to determine the polarity of the impact of these news texts on stock price changes. Observations indicated that this effect on the change of the stock price is undeniable.

In this research, according to the proposed topic, an attempt was made to improve the ontology created in the area of the Tehran Stock Market, and then a model was presented to prove the effect of the relationships in the ontology on stock prices. According to the scope defined for the research, the upcoming challenges and uncertainties will be as follows:

• Time limits for manual advancement and improvement using expert knowledge

• Limitations of Persian tools for text processing, determining polarity, and analysing news and social networks

• Human limitations to more accurately assess the impact of relationships on price changes

After improving and enriching the ontology created in the area of the Tehran Stock Exchange and proving the effectiveness of the existing relationships in the ontology, as a future work, creating and implementing a machine learning prediction model using the improved ontology and other prediction methods The stock price will be discussed and we will try to provide a better result for automatically predicting stock price changes.

The authors could not find research similar to this work for comparison. Moreover, the output of this research serves solely as proof of concept regarding the effect of ontology relations on stock price movement. As a future endeavor, a prediction model will be constructed using the ontological knowledge developed herein, enabling comparison with other existing studies in this field.

Conflict of interest

The authors declare no potential conflict of interest regarding the publication of this work. In addition, the ethical issues including plagiarism, informed consent, misconduct, data fabrication and, or falsification, double publication and, or submission, and redundancy have been completely witnessed by the authors.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

Batet, M., & Sánchez, D. (2014). A semantic approach for ontology evaluation. *2014 IEEE 26th International Conference on Tools with Artificial Intelligence*, 138–145.

Bennett. (2015). *Financial Instrument Global Identifier®.*

Cavalcante, R. C., Brasileiro, R. C., Souza, V. L. F., Nobrega, J. P., & Oliveira, A. L. I. (2016). Computational Intelligence and Financial Markets: A Survey and Future Directions. *Expert Systems with Applications*, *55*, 194–211. doi: 10.1016/j.eswa.2016.02.006

Chen, J., & Yang, L. (2021a). A Bibliometric Review of Volatility Spillovers in Financial Markets: Knowledge Bases and Research Fronts. *Emerging Markets Finance and Trade*, *57*(5), 1358–1379. doi: 10.1080/1540496X.2019.1695119

Chen, J., & Yang, L. (2021b). A Bibliometric Review of Volatility Spillovers in Financial Markets: Knowledge Bases and Research Fronts. *Emerging Markets Finance and Trade*, *57*(5), 1358–1379. doi: 10.1080/1540496X.2019.1695119

Fernández, M., Overbeeke, C., Sabou, M., & Motta, E. (2009). What makes a good ontology? A case-study in fine-grained knowledge reuse. *The Semantic Web: Fourth Asian Conference, ASWC 2009, Shanghai, China, December 6-9, 2009. Proceedings 4*, 61–75.

Gopinathan, R., & Durai, S. R. S. (2019). Stock market and macroeconomic variables: new evidence from India. *Financial Innovation*, *5*(1). doi: 10.1186/s40854-019-0145-1

Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, *5*(2), 199–220. doi: 10.1006/knac.1993.1008

Houstoun, K. J., Milne, A., & Parboteeah, P. (2014). Preliminary Report on Standards in Global Financial Markets. *SSRN Electronic Journal*, *May*, 1–88. doi: 10.2139/ssrn.2531210

Lamparter, S., & Schnizler, B. (2006). Trading services in ontology-driven markets. *Proceedings of the 2006 ACM Symposium on Applied Computing*, 1679–1683.

Metz, C. E. (n.d.). *Basic Principles of ROC Analysis*.

Mishra, A. K., Anand, S., Debnath, N. C., & Patel, A. (2023). An Ontological Framework for Risk Mitigation in Stock Market. In Intelligent Systems and Applications: Select Proceedings of ICISA 2022 (pp. 517–527). Springer.

Mohti, W., Dionísio, A., Ferreira, P., & Vieira, I. (2019). Contagion of the subprime financial crisis on frontier stock markets: A copula analysis. *Economies*, *7*(1). doi: 10.3390/economies7010015

Odilon Tilly Sardou. (2021). *Regulatory sanctions and their effects on the stock market*.

Omar, A. Bin, Ali, A., Mouneer, S., Kouser, R., & Al-Faryan, M. A. S. (2022). Is stock market development sensitive to macroeconomic indicators? A fresh evidence using ARDL bounds testing approach. *PLoS ONE*, *17*(10 October). doi: 10.1371/journal.pone.0275708

Raad, J., & Cruz, C. (2015). A survey on ontology evaluation methods. *Proceedings of the International Conference on Knowledge Engineering and Ontology Development, Part of the 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*.

Samani, M., & Albadvi, A. (2023). Constructing the ontology of the Persian stock market. *Iranian Journal of Information Processing & Management*.

Sharma, N., Soni, M., Kumar, S., Kumar, R., Deb, N., & Shrivastava, A. (2023). Supervised Machine Learning Method for Ontology-based Financial Decisions in the Stock Market. *ACM Transactions on Asian and Low-Resource Language Information Processing*, *22*(5), 1–24.

Wang, S., Zhe, Z., Kang, Y., Wang, H., & Chen, X. (2008). An ontology for causal relationships between news and financial instruments. *Expert Systems with Applications*, *35*(3), 569–580. doi: 10.1016/j.eswa.2007.07.022

Zaki, M., Theodoulidis, B., & Diaz, D. (2019). Ontology-driven framework for stock market monitoring and surveillance. In HANDBOOK OF GLOBAL FINANCIAL MARKETS: Transformations, Dependence, and Risk Spillovers (pp. 75–103). World Scientific.

Zdraveski, V., Jovanoski, M., & Franke, U. (2017). Stock Market Ontology. *ICT Innovations*, 1–7.

Zhang, L., Zhao, M., & Feng, Z. (2019). Research on Knowledge Discovery and Stock Forecasting of Financial News Based on Domain Ontology. *International Journal of Information Technology and Decision Making*, *18*(3), 953–979. doi: 10.1142/S0219622019500160

Zhao, C., Hu, P., Liu, X., Lan, X., & Zhang, H. (2023). Stock Market Analysis Using Time Series Relational Models for Stock Price Prediction. *Mathematics*, *11*(5). doi: 10.3390/math11051130

Bibliographic information of this paper for citing:

Samani, Mohammad Hossein & Albadvi, Amir (2024). An Effective Model for Ontology Relations Efficacy on Stock prices: A Case Study of the Persian Stock Market. *Journal of Information Technology Management,* 16 (3), 1-20.
https://doi.org/10.22059/jitm.2024.373635.3640