International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-8 Issue-10, August 2019

A Methodology for Discovering Upstream and Downstream Causal Relationships in Stock Market

Harchana Bhoopathi, B Rama

Abstract: ***Causal relationships between events pertaining to stock market have potential to influence the stakeholders of the companies associated with those events. Understanding causal relationships in stock markets help in making intelligent decisions. Traditional prediction approaches cannot estimate the upstream and downstream causal relationships. Therefore, it is inevitable to consider portfolios that exhibit causal relationships. Simple correlation between variables may not reflect causal relationships unless there is an event that is the result of occurrence of another event. Finding upstream and downstream causal relationships is challenging. In the literature it is found that inter-transactional details can help in finding causal relationships. Based on this idea, in this paper, we planneda methodology to mine upstream and downstream causal relationships. An algorithm by name Upstream Downstream - Causal Relationship Mining (UD-CRM) is proposed to achieve this. The framework and underlying algorithm produce specific rules that are used to conclude causal relations. Experiments are made with stock dataset using a prototype application built. The experimental results revealed that the proposed framework is useful and performance better than existing approach.***

Keywords: ***Data mining, causal relationships, upstream and downstream causal relationships***

I. INTRODUCTION

Stock market is the aggregation of many stakeholders like buyers and sellers of stocks besides securities and public stock exchanges. National Stock Exchange (NSE) and Bombay Stock Exchange (BSE) are two Indian stock markets where trading takes place. The effectiveness of stock analysis and prediction has it impact on decision making. Stock analysis is crucial for all the businesses and individuals who invest in stock market. The entities involve in stock analysis exploit the visible data and their analysis knowhow to predict trends in stocks. From the literature [1], [2], [4], [5] and [6] it is understood that causal relationship mining also can be used to predict stock business trends. However, it is still in its infancy and there is need for new ways of mining causal rules. One of the ways for its improvement is considering upstream and downstreamstock business incompanies and applies causal relationship mining. Before entering into the proposed upstream and downstream causal relationship mining, it is important to understand causal relationship in more formal way. Causal relationship is something like cause and effect relationship. It indicates that an event E1 causes another event E2 denoted as E1→E2. Here E1 is called cause and E2 is known as effect. However, when two variables like x and y are correlated and change in one variable causes change in another variable may not essentially mean causal relationship. Rather it is correlation as proved by Pearson's correlation effect.

$$r_{xy} = \frac{n\sum xy - \sum x \sum y}{\sqrt{(n\sum x^2 - (\sum x)^2)(\sum y^2 - (\sum y)^2)}} \quad (1)$$

As shown in Eq. (1), the correlation between variables x and y are measured. The point here is that the correlation between x and y does not imply causation. The result of Pearson's correlation can be between -1 to 1 where 1 indicates strong positive correlation while -1 indicates strong negative correlation. As the correlation is not equal to causal relationship, it is essential to understand the causal relationship in a right way. Nevertheless, positive correlation may have causal relationship in it sometimes. For instance, the number of hours study made by student caused improved grading. The bottom-line here is that all associations may not be causal relationships. The existing approaches pertaining to data mining include clustering, web mining, association rule mining and sequential pattern mining. If two stocks such as S1 and S2 have certain relationship they may show up and down together. This kind of knowhow is not suitable for future investments. The sequential patterns and association rules have their limitations in supporting decision making. This problem is solved in this paper by using data mining approach known as inter-transactions mining. We find causal upstream and downstream relationships to achieve better forecasting capabilities. Stock value of a company C1 is increased which will cause other company C2's stock value to be increased most likely with certain (70% probability for example). Our contributions in this paper are as follows.

- We proposed a novel framework for discovering upstream and downstream causal relationships and causal relationship chains so as to analyse trends in stock market.
- We proposed an algorithm known as Upstream Downstream - Causal Relationship Mining (UD-CRM) to obtain causal relationships among stocks of upstream and downstream companies.
- We built a prototype application to demonstrate proof of the concept. A moderated stock market dataset is used for experiments. The results revealed the effectiveness of the proposed framework.

Manuscript published on 30 August 2019.
*Correspondence Author(s)
Harchana Bhoopathi, Department of Computer Science, Kakatiya University, Warangal, Telengana State, India.
B Rama, Department of Computer Science, Kakatiya University, Warangal, Telengana State, India.

The remainder of the paper is structured as follows. Section 2 reviews literature on causal relationship mining. Section 3 provides preliminary details. Section 4 presents the proposed framework for mining upstream and downstream causal relationships. Section 5 presents experimental results. Section 6 presents evaluation of the proposed approach. Section 7 provides threats to validity of the methodology used while section 8 concludes the paper besides giving directions for future work.

II. RELATED WORK

This section reviews literature on causal relationships. Dritsaki and Melina Dritsaki [1] considered stock market and credit market and analysed economic development and the causal relationship affecting it. Causal relationship is found between economic growth and banking sector while no such relationship is found between baking sector and stock market. Ali et al. [2] on the other hand proposed a methodology to find causal relationship between stock exchange prices and macro-economic indicators. There was no causal relationship found between stock exchange and macroeconomic indicators. Garcia and Liu [3] found that there are macroeconomic determinants for the growth of stock markets. However, there is relationship between stock markets and macroeconomic variables like inflation and exchange rate [5]. The sensitivity of Indian stock markets of international markets is found affirmative [4]. Importance of data mining for various applications is explored in [6].

Alagidede et al. [7] studied exchange rates and stock prices to know causal relationships. Some relationship is found between them in the study. In another study made in [8] it is found that media pressure influences stock markets. Yet another study made in [9] claimed that there is relationship between inflation and stock prices. Yang et al. [10] found that there is possibility of causal and upstream and downstream relationships among related companies listed in stock markets. This study is close to our research in this paper.

The role of text mining is studied in [11] and [15] for acquiring business intelligence. It is used in educational data mining as discussed in [12], [23]. It comes under machine learning [27], [36] which is widely used in different real time applications [13]. Many algorithms related to machine learning are found in [28]. Often data mining and visualization go hand in hand for better performance [14]. Causal relationship is explored in questions and answers in [16]. Various techniques are available for acquiring business intelligence. They include soft computing [17], clustering and classification [18], statistical inference [19] and Bayesian networks [20].

Sentimental causal rules are discovered from social media. From Twitter data different aspects are studied and causal relationships are found [22]. Extraction of causal relationships on human proteins is studied in [24]. Causal mining rule discovery is studied in [25] for making intelligent applications. With respect to movement patterns, candidate causal relationships are extracted in [26] using data mining techniques. Mining of profit rules from the dataset containing inter-transaction item sets is explored in [29]. Important data mining techniques [30] and data science [31] are investigated. Causal inference methods [32], intelligent method for stock prediction [33] dependency driven data analytics [34] and usage of artificial intelligence in data mining [35] are other researchers found in the literature. Causal mining research is also found in [37] and [38]. From the literature it is understood that there is inadequate research on finding upstream and downstream causal relationships. Therefore, in this framework we proposed a framework for throwing light on this.

III. PROBLEM FORMULATION

Data mining techniques[37] such as association rule mining is widely used for discovering knowledge from databases. Associations in transactions can provide hidden information in the form of trends or patterns. A database D can have set of associations A that can be discovered by generating rules R. The problem of mining association rules can be divided into two parts. In the first part, all item sets with given support are obtained from data source. Then the frequent item sets are used to generate association rules. The problem with this kind of mining association rules is that any item set which does not have the support is considered an uninteresting item set[37] and [38]. However, we believe that associations can be of different kinds. They are known as direct, indirect and exception associations. With respect to stock market data, it is essential to have comprehensive business intelligence before making decisions. Therefore, it is inevitable to have a framework that can cater to the needs of such expert decision making. The rationale is described here. In stock market dataset item sets that are infrequent also can provide information needed towards converging decisions. Two items a and b have no direct relationship but they may have strong relationship through another item set Y. In this case the, b pair is said to have indirection relationship or association. Mining such rules is the focus of this paper. We defer the details of direct and exception relationships besides effect of events and government decisions are deferred to our next research paper.

IV. PROPOSED FRAMEWOK

In this paper, we proposed a framework that has provision for finding causal relationships in terms of direct, indirect and exception relationships. When all these relationships are investigated, it can lead to more intelligent decision making in stock market business. In addition to this certain events and unexpected decisions made by government can have impact on stock market. When all these are considered, the error rate in the prediction can be minimized. Though the framework has provision for discovering different relationships, in this paper we present finding indirect relationships of stock tickers.

Retrieval Number: J10740881019/19©BEIESP
DOI: 10.35940/ijitee.J1074.0881019
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

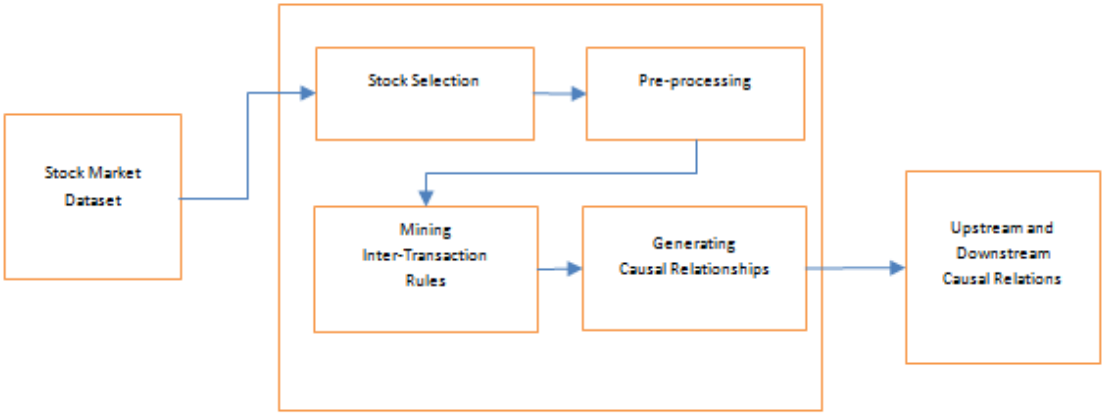

Figure 1: Proposed Framework for Comprehensive Stock Business Intelligence

As shown in Figure 1, it is evident that the basis for mining causal relationships is to have frequent item sets from given stock dataset. The ensuing section provides more details on stock dataset used for empirical study. The framework has provision to analyse one or more stock datasets. After extracting frequent item sets using any of the existing data mining algorithms, it focuses on mining direct, indirect and exception relationships from the frequent item sets. The resultant rules obtained are integrated to form comprehensive business intelligence. As far as implementation is concerned, in this paper, we present only the study on indirect relationships. In other words, our focus is on inter-transactions that have no direct relationship but through a mediator. This can show potential tickers that have influence on others.

Causal Relationship Discovery in Stock Market

Since stock market is in financial domain and influences the lives of all stakeholders and countries even, it is considered in the proposed research. Moreover, upstream and downstream causal relationships can be found in stock market and can exploit for expert decision making.

Algorithm:Upstream Downstream - Causal Relationship Mining (UD-CRM)
Input:Stock market database D, support *sup*, confidence *conf*, sliding window *w*, updown percentage *udp*
Output:Causal Relationships *R*
01**Initialize support vector *S***
02**Initialize frequent itemset vector *F***
03***F* = FIN(*D*, *w*)**
04**For each *itemset* in *F***
05**Find *support* value**
06 **Find *confidence* value**
07 **IF *support*<*sup*and confidence < conf THEN**
08 **Prune it from *F***
09 **END IF**
10**End For**
11**For each *itemset* in *F***
12**Mine inter-transaction rules using *udp***
13 **Add rules to *R***
14**End For**
15**Return *R***

Algorithm 1:Upstream Downstream - Causal Relationship Mining

As provided in Algorithm 1, the sliding window, support, confidence and stock dataset are considered as inputs. The algorithm used FIN algorithm and sliding window in order to generate frequent item sets associated with given sliding window. Then the item sets are pruned in order to meet the required support and confidence (interestingness of users) metrics. Afterwards, the inter-transaction rules are mined and finally the rules are returned. Here are more details of the algorithm. Lines 1 and 2: Arrays or vectors are created to hold a collection of support values and frequent item sets. Line 3: FIN is known as fast item set mining algorithm. It is used to mine item sets very faster with high efficiency. The reason behind this is that it makes use of data structure which makes it faster. The underlying data structure is known as POC-tree (Pre Order Coding tree). It takes two parameters such as Dataset D and sliding window w. Sliding window is something related to the model which considers data processing incrementally. FIN returns frequent item sets that are stored in the vector F. POC tree is constructedas underlying data structure which makes mining frequent item sets faster. The tree needs either pre-order or post-order traversal so that it is computationally effective and improves performance of item set mining. More technical details of the POC-tree and FIN algorithm can be found in [1]. Line 4: Iterative statement which is repeated based on the number of values present in F. Line 5: Computes support value for given item set. Line 6: Computes confidence value for the given item set. Both support and confidence are metrics that quantify the frequency of item sets. Support is the measure to know frequency of given pattern in the dataset. Considering A and B are the two item sets that are part of a pattern, the support indicates frequent of these items. In other words, support is the percentage of the transactions where A and B items are present. Support is mathematically represented as follows.

Support = Probability (A and B)
Support = (# of transactions involving A and B) / (total number of transactions).

Confidence on the other hand denotes the strength of implication of an association (frequent item set) rule. In other words, it is the percentage of transactions that contain B if they contain A. It is mathematically represented as follows.

Confidence = Probability (B if A) = P(B/A)
Confidence = (# of transactions involving A and B) / (total number of transactions that have A).

Line 7 checks fitness or usefulness of the item set based on the given support and confidence (inputs to the algorithm). Line 8 removes item set from F if the item set is not meeting requirements. Line 9 terminates the conditional statement. Line 10 terminates iterative statement. Line 11 starts an iterative statement which repeats number of times equal to the items present in the F vector. Line 12: Inter transaction rules are mined using up down percentage which is used for upstream and downstream causal relationship. Mined rules are added to vector R in Line 13 while Line 14 ends the iterative statement. Line 15 returns the rules that reflect causal relationships.

V. EXPERIMENTAL RESULTS

Experiments are made with select groups of stocks. A prototype application is built for carrying out experiments. The upstream and downstream computations are made with certain computations described here. The ups and downs of the stock prices are computed with a measure known as UDP (Up or Down Percentage). Stock is considered up when it satisfies the following condition.

UDP=(Closing Price – Opening Price) / Opening Price

Stock price is considered up when UDP is greater than 0.005 and considered down when the same is less than -0.005. In case of other results, the situation is known as a tie and it is not interested by the investors in general. In the results, therefore, the tie situations are removed intentionally.

After performing causal mining for discovering upstream and downstream relationships, rules are formed as the result of algorithm. For instance, rulesare formed as follows.

Rule 1: If stock *A* goes up *p*% of time, stock *B* will go up after *d* days.
Rule 2: If stock *A* goes up *q*% of time, stock *B* will go down after *e* days.
Rule 3: If stock *A* goes up *r*% of time, stock *B* will go up after *f* days.

This way causal upstream and downstream rules are generated. The results are presented in this section. The proposed algorithm is employed to obtain causal relationships. The selected groups and stocks are provided in Table 1.

Group	Stocks
Dow Jones Indus	DJI01, DJI02
FT-SE 100	FTSE01, FTSEB02
Nikkei 225	Nikkei01, Nikkei02
Hang Seng	HS01, HS02
Singapore Straits	SS01, SS02
IBM	I01,I02
Microsoft	M01,M02

Table1: Selected groups and their stock values

The selected groups and their stocks are provided. Now it is important to have predictions and then provide causal rules that are useful to investors. Up and Down percentages are used in order to identify a down or up trend in the stock market. The sliding window considered is 5 days. The first mining results are provided in Table 2.

Rule	Day-0	Day-1	Day-2	Day-3	Day-4	conf
1	FTSE01D()	FTSE02D()				0.52
2	HS01D()	HS02D()				0.53
3	SS01D()	HS02D()				0.55
4	Nikkei01D()	HS02D()				0.50
5	I01D()	HS02D()				0.57
6	Nikkei01D()	HS02D()				0.51
7	SS01D()	Nikkei02D()				0.51
8	SS02D()		HS02D()			0.51
9	FTSE02D()		HS02D()			0.54
10	I02D()			HS02D()		0.50
11	FTSE01D()			HS02D()		0.51
12	SS02D()			HS02D()		0.58
13	HS02D()			HS02D()		0.51
14	IO1D()			HS02D()		0.50
15	Nikkei01D()			HS02D()		0.57
16	FTSE02D()			HS02D()		0.50
17	Nikkei02D()			FTSE02D()		0.50
18	I02D()			FTSE02D()		0.54
19	SS02D()			HS02D()		0.55
20	I01D()					

Table 2: Results of first mining

Table 2 shows the results with different parameters considered. The sliding window used is 5. In the same fashion support is considered as 0.2 while confidence is 0.5. As shown in Table 2, in Day 0 there is FTSE1D (229) where stock id is FTSE1D. Here D refers to downward stock and 229 refer to its frequency in the source data. The Rule 1 indicates that when FTSE1D goes down it causes the FTSE2D to go up the next day with 52% probability (confidence). Similar kind of relation exists in Rule 2 to Rule 7 with 0.53%, 0.55%, 0.50%, 0.57%, 0.51% and 0.51% probability. Incase of Rule 8 and Rule 9, there is similar kind of relationship witnessed on Day-0 and Day-2 with 0.51% and 0.54% probability. From Rule 10 to Rule 19, there is this kind of relationship occurred with observations on Day-0 and Day-3 with different probabilities. From the observed rules one stock id goes up or down it causes the stock id goes up odown next day with high or low percentage of probability. Another experiment is made by lowering the support value to 10% and increase confidence to 60%. It has resulted in 10 rules as shown in Table 3.

Rule	Day-0	Day-1	Day-2	Day-3	Day-4	Conf
1	I02D,HS02D,M02D()			FTSE02D()		0.63

2	DJI01D,I02D()			FTSE02D()	0.60
3	SS01U(143)		FTSE02D()		0.62
4	HS02U(140)		FTSE02D()		0.61
5	HS02D,FTSE02D,M02D(91)			FTSE02D()	0.60
6	Nikkei01U, Nikkei02U,M02U,FTSE02D(53)				0.66
7	Nikkei01U, Nikkei02U,M02U(80)			FTSE02D()	0.64
8	FTSE02U(148)			FTSE02D()	0.64
9	M02U(117)		FTSE02D()		0.63
10	M01U()		FTSE02D()		0.63

Table 3: Results of second mining

As shown in Table 3 different rules showing causal relationship with specified confidence or probability are presented.

Confidence	Records
0.7	1
0.69	1
0.68	9
0.67	8
0.66	13
0.65	11
0.64	50
0.63	75
0.62	57
0.61	80
0.60	94

Table 4: Results showing confidence and instances

The confidence values used in experiments and the instances are presented in Table 4. It shows the impact of confidence on the results.

VI. PERFORMANCE EVALUATION

The causal mining framework proposed in this paper is evaluated. The metrics used for the same are precision, recall, memory consumption and execution time.

Size of Dataset (No. of Instances)	Execution Time (Sec)			
50000	0.220	2.828	6.409	14.084
100000	0.193	2.351	5.043	12.295
150000	0.196	2.508	5.354	13.685
200000	0.223	3.520	6.241	13.098

Table 5: Execution time

As shown in Table 5, the execution time of the algorithm is recorded. The results are observed against number of instances in dataset.

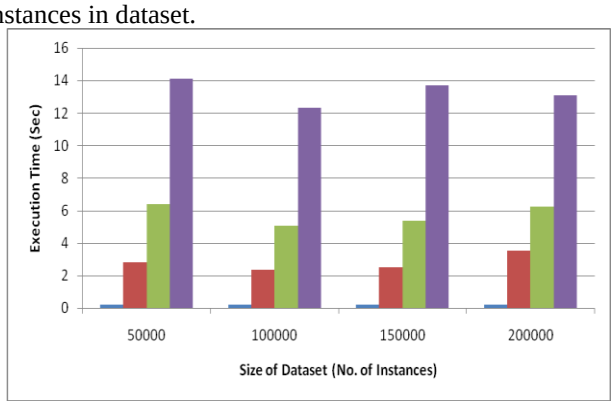

Figure 3: Impact of dataset size on execution time

As presented in Figure 3, the size of dataset is presented in horizontal axis and vertical axis is showing execution time. The results revealed that the execution time is different with different number of instances in the dataset.

Size of Dataset (No. of Instances)	Memory (MB)			
50000	2.947838	3.8514	5.6231	7.3677
100000	2.98597	3.2761	5.7986	7.0214
150000	2.928062	3.6817	6.1708	7.2894
200000	2.928093	4.6732	7.8204	8.5607

Table 6: Memory usage

As shown in Table 6, it is evident that the memory usage is recorded against different number of instances in dataset.

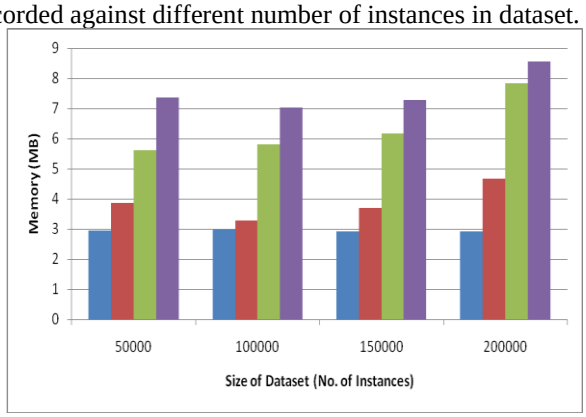

Figure 4: Impact of size of dataset on memory usage

As presented in Figure 4, there is impact of size of dataset on the memory usage of the algorithm proposed.

	Ground Truth (correct prediction)	Ground Truth (incorrect prediction)
Result of UD-CRM (correct prediction)	True Positive (TP)	False Positive (FP)
Result of UD-CRM (incorrect prediction)	False Negative (FN)	True Negative (TN)

Table 6:Shows confusion matrix used in the evaluation of results

Based on the confusion matrix shown in Table 6, three measures are used for evaluation. They are known as precision, recall and F-measure. Precision reveals fraction of relevant relationships among the retrieved relationships. The recall indicates the fraction of relevant relationships that have been found over all the relationships. Eq. 2 and Eq. 3 show the computation of precision and recall.

Precision = (TP/(TP+FP) *100 (2)

Recall = (TP/(TP+FP))*100 (3)

The precision and recall are used to have a weighted harmonic mean of the measures precision and recall. It is computed as in Eq. 4.

F-Measure = 2*((precision*recall)/(precision+recall)) (4)

The evaluation results of the proposed algorithm are compared with an existing approach. The evaluation results are presented in Table 7.

Technique	Precision	Recall	F-Measure
Base Line Approach (existing)	83.4	84.2	84.8
Proposed	89.8	90.7	90.24

Table 7: Evaluation results

As shown in Table 7, the precision and recall and F-measure are computed for both existing and proposed approaches and the results are presented.

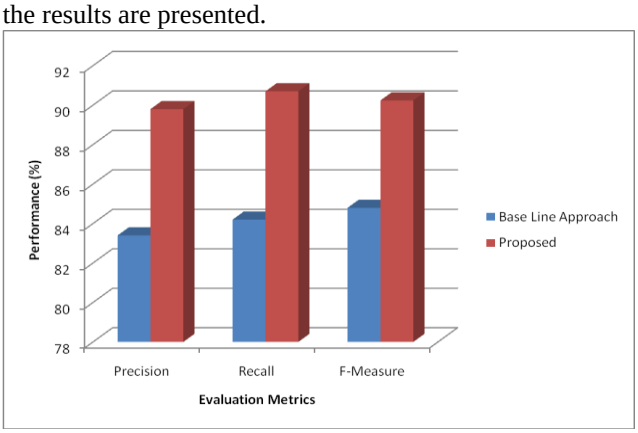

Figure 5: Evaluation results

As presented in Figure 5, the measures are provided in horizontal axis. They are precision, recall and F-measure. These metrics are computed against existing and proposed causal mining approaches. The results reveal that the proposed method outperforms the existing method in terms of all the metrics used for evaluation. This shows the efficiency of the Upstream Downstream - Causal Relationship Mining (UD-CRM) algorithm. The algorithm proposed in this paper is made more effective with consideration on the usage of FIN algorithm which is faster and provides accurate frequent item sets. Moreover, the algorithm considered different parameters like support, confidence, sliding window and stock dataset of different size.

VII. CONCLUSIONS AND FUTURE WORK

Investigation into cause-and-effect relationships has its utility in many real time applications. In the case of stock market which shows unexpected trends, it is essential to consider causal relationships. Moreover, upstream and downstream relationships associated with cause-and-effect can have more impact on the prediction of stock performance of selected portfolios. A group of companies may be related to a domain and all of them may get affected due to some cause-and-effect relationship. Mining such trends needs careful consideration of stock portfolios and mine useful rules that govern cause and effect. In this paper, we proposed a framework that provides required guidelines with an underlying algorithm to mine upstream and downstream causal relationships in stock markets. A prototype application is built to demonstrate proof of the concept. The experimental results are evaluated using performance metrics like execution time, memory usage, precision, recall and f-measure. The proposed framework showed better results than an existing state of the art method. In future we intend to consider causal-chain relationships that provide finer grained knowhow on the prediction of causal relationships and the portfolios in stock market.

REFERENCES

1. Chaido DritsakiI and Melina Dritsaki-Bargiota. (2015). The Causal Relationship between Stock, Credit Market and Economic Development: An Empirical Evidence for Greece. Economic Change and Restructuring, p1-16.
2. Imran Ali1, Kashif Ur Rehman, Ayse Kucuk Yilmaz, Muhammad Aslam Khan and Hasan Afzal. (2010). Causal relationship between macro-economic indicators and stock exchange prices in Pakistan. ISSN. 4 (.), p1-8.
3. Valeriano. Garica and Lin Liu. (1999).Macroeconomic Determinants of Stock Market Development. Journal of Applied Economics. 2, p1-31.
4. Dr. Naliniprava Tripathy. (2011). Causal Relationship between Macro-Economic Indicators and Stock Market in India. Asian Journal of Finance & Accounting. 3, p1-19.
5. Dharmendra Singh1. (2005). Causal Relationship Between Macro-Economic Variables and Stock Market: A Case Study for India. PJSS., p1-12.
6. Cristobal Romero and Sebastian Ventura. (2013). Data mining in education. DOI. 3, p1-16.
7. Paul Alagidede, Theodore Panagiotidis and Xu Zhang. (2010). Causal Relationship between Stock Prices and Exchange Rates. IEEE, p1-22.
8. Paul and Tetlock. (2007). Giving Content to Investor Sentiment: The Role of Media in the Stock Market. FINANCE. LXI, p1-30.
9. Basabi Bhattacharya and Jaydeep Mukherjee. (2010). The Nature of the Causal Relationship between Stock Market and Macroeconomic Aggregates in India: An Emparical Analysis.IEEE., p1-21.
10. Y.L. Hsieh, Don-Lin Yang, and Jungpin Wu. (2006). Using Data Mining to Study Upstream and Downstream Causal Relationship in Stock Market. IEEE., p1-5.
11. Sarvnaz Karimi and Chen Wang and Alejandro Metke-Jimenez and Raj Gaire and Cecile Paris, Csiro. (2015). Text and Data Mining Techniques in Adverse Drug Reaction Detection. ACM., p1-40.
12. Ryan S.J.D. Baker. (2009). The State of Educational Data Mining in 2009: A Review and Future Visions. Journal of Educational Data Mining, 1, p1-14.
13. H. Witten and Eibe Frank. (2009). Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations. SIGMOD. 31, p1-3.
14. Ben Shneiderman. (2002). Inventing discovery tools: combining information visualization with data mining. ACM, p1-18.
15. Marti A. Hearst. (2009). Untangling Text Data Mining. ACM, p1-8.
16. Roxana Girju. (2009). Automatic Detection of Causal Relations for Question Answering. ACM, p1-8.
17. Sushmita Mitra, Senior Member, Sankar K. Pal, Fellow, and Pabitra Mitra. (2002). Data Mining in Soft Computing Framework: A Survey. IEEE. 13 (.), p1-12.
18. E.W.T. Ngai, Li Xiu b, D.C.K. Chau a. (2009). Application of data mining techniques in customer relationship management: A literature review and classification. Expert Systems with Applications., p1-11.
19. Clark Glymour, David Madigan, Daniel Pregibon and Padhraic Smyth. (2009). Statistical Inference and Data Mining. cmu, p1-18.
20. David Heckerman. (1997). Bayesian Networks for Data Mining. IEEE., p1-41.
21. Rahim Dehkharghani, Hanefi Mercan, Arsalan Javeed and Ycel Saygn. (2014). Sentimental Causal Rule Discovery from Twitter. ISSN, p1-8.
22. Ryan S.J.d. Baker. (2010). Educational Data Mining and Learning Analytics. International Journal of Computer Applications., p1-19.
23. Kalpana Raja, Suresh Subramani and Jeyakumar Natarajan. (2013). PPInterFinder—a mining tool for extracting causal relations on human proteins from literature. ACM., p1-11.
24. Philippe Fournier-Viger, Usef Faghihi, Roger Nkambou, and Engelbert Mephu Nguifo2. (2012). CMRules: Mining Sequential Rules Common to Several Sequences. ACM., p1-18.
25. Antony Galtonb, Patrick Laubec and Jarod LyondSusanne Bleischa, Matt Duckhama, Antony Ga. (2012). Susanne Bleischa, Matt Duckhama, Antony Galtonb, Patrick Laubec, Jarod Lyond. Mining candidate causal relationships in movement patterns, p1-21.
26. Tony C. Smith and Eibe Frank. (2016). Introducing Machine Learning Concepts with WEKA. IEEE, p1-26.

27. Giulio Caravagnaa, Alex Graudenzia, DanieleRamazzottia, Rebeca Sanz-Pamplonad, Luca De Sanoa, Giancarlo Mauria, Victor Morenod, Marco Antoniottia and Bud Mishra. (2016). Algorithmic methods to infer the evolutionary trajectories in cancer progression. IEEE, p1-10
28. Yu-Lung Hsieh, Don-Lin Yang, Jungpin Wuand Yi-Chung Chen. (2016). Efficient Mining of Profit Rules from Closed Inter-Transaction Itemsets. Journalof Information Science and Engineering. 32, p575-595.
29. Keito Yoshimura, TakeshiOkanoue, HayaoEbise, Tsuyoshi Iwasaki, Masayuki Mizuno, ToshihideShima, Junji Ichihara and Kazuto Yamazaki. (2016). Identification of Novel Noninvasive Markers for Diagnosing Nonalcoholic Steatohepatitis and Related Fibrosis by Data Mining. Hepatology. 63 (2), p1-12.
30. Julia Stoyanovich, Bill Howe, Serge Abiteboul, Gerome Miklau, Arnaud Sahuguet, Gerhard Weikum. (2017). Fides: Towards a Platform for Responsible Data Science. IEEE, p1-14.
31. Thuc Duy Le. (2017). Causal inference methods and applications in Bioinformatics. IEEE, p1-16.
32. Yuvraj Wadghule, I.R. Shaikh. (2017). Neural Network Approach for Stock Prediction using Historical Data. International Journal of Engineering Technology, Management and Applied Sciences. 5 (7), p1-8.
33. Ruslan Mavlyutov, CarloCurino, BorisAsipov, Philippe Cudre-Mauroux. (2017). Dependency-Driven Analytics: A Compass for Uncharted Data Oceans. IEEE, p1-9.
34. Yaguang Li, Rose Yu, Cyrus Shahabi, Yan Liu. (2018). Diffusion Convolutional Recurrent Neural Network: Data-Driven Traffic Forecasting. IEEE, p1-16.
35. Yaguang Li, Cyrus Shahabi. (2018). A Brief Overview of Machine Learning Methods for Short-term Traffic Forecasting and Future Directions. IEEE, p1-7.
36. Harchana Bhoopathi and B. Rama. (2017). Causal rule mining for knowledge discovery from databases. International Conference on Intelligent Computing and Control Systems. p1-25.
37. Harchana Bhoopathi and B. Rama. (2017). A novel framework for stock trading analysis using causal relationship mining. Third International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics, p1-25.
38. Vurukonda, S., Ranadheer, G., Mounika, B., & Reddy, S. (2013). A survey on tree based association rules (TARs) from XML Documents. Proceedings of International Journal of Research and Computational Technology, 5.
39. Naveen Kumar, R., Vege, H.K. & Sreeram, G. 2019, "Patient treatment interval used in forecast algorithm and solicitations in hospital queuing management", International Journal of Innovative Technology and Exploring Engineering, ISSN:2278-3075, vol. 8, no. 7, pp. 3003-3007.
40. Reddy, L.H., Thamognudu, Y. & Sreeram, G. 2019, "Deployment of a secured web application using cryptanalysis in cloud environment", International Journal of Engineering and Advanced Technology, ISSN:2249-8958, vol. 8, no. 4, pp. 1841-1844.
41. Sreeram, G., Kanumuri, M.K. & Bodduluri, M. 2019, "Improving cloud data storage performance based on calculating score using data transfer rate between the internetwork drives", International Journal of Engineering and Advanced Technology, ISSN:2249-8958, vol. 8, no. 4, pp. 1830-1835.
42. Naveen Kumar, R., Vege, H.K. & Sreeram, G. 2019, " A Perspective of Probabilistic Misbehavior Detection Scheme in Vehicular Ad-hoc Networks", International Journal of Innovative Technology and Exploring Engineering,ISSN:2278-3075, vol. 8, no. 6, pp.1098-1102.

AUTHORS PROFILE

Harchana Bhoopathi Completed M.C.A. from Alluri Institute of Management and Sciences. Pursuing Ph.D in the Department of Computer Science, Kakatiya University. Area of subject interested in Causal Data Mining. Published papers in IEEE Conferences and International Journals.

Dr. B.RAMA received her Ph.D. Degree in Computer Science from Padmavati Mahila Visvavidyalayam, Thirupathi, India in the year of 2009. She is working as Assistant Professor in Computer Science at Department of Computer Science, University Campus College, Kakatiya Universitysince ten years. Her area of interest is Artificial Intelligence and Data Mining. She is the author or co-author of various scientific, technical papers mainly in IEEE, Springer International Conferences and International Journals.