INTERNATIONAL JOURNAL OF RESEARCH IN COMMERCE & MANAGEMENT

IJRCM

A Monthly Double-Blind Peer Reviewed (Refereed/Juried) Open Access International e-Journal - Included in the International Serial Directories

Indexed & Listed at:

Ulrich's Periodicals Directory ©, ProQuest, U.S.A., EBSCO Publishing, U.S.A., Cabell's Directories of Publishing Opportunities, U.S.A.,

Open J-Gage, India [link of the same is duly available at Inflibnet of University Grants Commission (U.G.C.)],

Index Copernicus Publishers Panel, Poland with IC Value of 5.09 & number of libraries all around the world.

Circulated all over the world & Google has verified that scholars of more than 3130 Cities in 166 countries/territories are visiting our journal on regular basis.

Ground Floor, Building No. 1041-C-1, Devi Bhawan Bazar, JAGADHRI – 135 003, Yamunanagar, Haryana, INDIA

http://ijrcm.org.in/

CONTENTS

CALL FOR MANUSCRIPTS

We invite unpublished novel, original, empirical and high quality research work pertaining to recent developments & practices in the areas of Computer Science & Applications; Commerce; Business; Finance; Marketing; Human Resource Management; General Management; Banking; Economics; Tourism Administration & Management; Education; Law; Library & Information Science; Defence & Strategic Studies; Electronic Science; Corporate Governance; Industrial Relations; and emerging paradigms in allied subjects like Accounting; Accounting Information Systems; Accounting Theory & Practice; Auditing; Behavioral Accounting; Behavioral Economics; Corporate Finance; Cost Accounting; Econometrics; Economic Development; Economic History; Financial Institutions & Markets; Financial Services; Fiscal Policy; Government & Non Profit Accounting; Industrial Organization; International Economics & Trade; International Finance; Macro Economics; Micro Economics; Rural Economics; Co-operation; Demography: Development Planning; Development Studies; Applied Economics; Development Economics; Business Economics; Monetary Policy; Public Policy Economics; Real Estate; Regional Economics; Political Science; Continuing Education; Labour Welfare; Philosophy; Psychology; Sociology; Tax Accounting; Advertising & Promotion Management; Management Information Systems (MIS); Business Law; Public Responsibility & Ethics; Communication; Direct Marketing; E-Commerce; Global Business; Health Care Administration; Labour Relations & Human Resource Management; Marketing Research; Marketing Theory & Applications; Non-Profit Organizations; Office Administration/Management; Operations Research/Statistics; Organizational Behavior & Theory; Organizational Development; Production/Operations; International Relations; Human Rights & Duties; Public Administration; Population Studies; Purchasing/Materials Management; Retailing; Sales/Selling; Services; Small Business Entrepreneurship; Strategic Management Policy; Technology/Innovation; Tourism & Hospitality; Transportation Distribution; Algorithms; Artificial Intelligence; Compilers & Translation; Computer Aided Design (CAD); Computer Aided Manufacturing; Computer Graphics; Computer Organization & Architecture; Database Structures & Systems; Discrete Structures; Internet; Management Information Systems; Modeling & Simulation; Neural Systems/Neural Networks; Numerical Analysis/Scientific Computing; Object Oriented Programming; Operating Systems; Programming Languages; Robotics; Symbolic & Formal Logic; Web Design and emerging paradigms in allied subjects.

Anybody can submit the **soft copy** of unpublished novel; original; empirical and high quality **research work/manuscript** ***anytime*** in ***M.S. Word format*** after preparing the same as per our **GUIDELINES FOR SUBMISSION**; at our email address i.e. infoijrcm@gmail.com or online by clicking the link **online submission** as given on our website (***FOR ONLINE SUBMISSION, CLICK HERE***).

GUIDELINES FOR SUBMISSION OF MANUSCRIPT

1. **COVERING LETTER FOR SUBMISSION:**

DATED: _____

THE EDITOR
IJRCM

Subject: **SUBMISSION OF MANUSCRIPT IN THE AREA OF .**

(e.g. Finance/Marketing/HRM/General Management/Economics/Psychology/Law/Computer/IT/Engineering/Mathematics/other, please specify)

DEAR SIR/MADAM

Please find my submission of manuscript entitled '_____' for possible publication in your journals.

I hereby affirm that the contents of this manuscript are original. Furthermore, it has neither been published elsewhere in any language fully or partly, nor is it under review for publication elsewhere.

I affirm that all the author (s) have seen and agreed to the submitted version of the manuscript and their inclusion of name (s) as co-author (s).

Also, if my/our manuscript is accepted, I/We agree to comply with the formalities as given on the website of the journal & you are free to publish our contribution in any of your journals.

NAME OF CORRESPONDING AUTHOR:
Designation:
Affiliation with full address, contact numbers & Pin Code:
Residential address with Pin Code:
Mobile Number (s):
Landline Number (s):
E-mail Address:
Alternate E-mail Address:

NOTES:

a) The whole manuscript is required to be in ***ONE MS WORD FILE*** only (pdf. version is liable to be rejected without any consideration), which will start from the covering letter, inside the manuscript.
b) The sender is required to mentionthe following in the **SUBJECT COLUMN** of the mail:
New Manuscript for Review in the area of (Finance/Marketing/HRM/General Management/Economics/Psychology/Law/Computer/IT/ Engineering/Mathematics/other, please specify)
c) There is no need to give any text in the body of mail, except the cases where the author wishes to give any specific message w.r.t. to the manuscript.
d) The total size of the file containing the manuscript is required to be below **500 KB**.
e) Abstract alone will not be considered for review, and the author is required to submit the complete manuscript in the first instance.
f) The journal gives acknowledgement w.r.t. the receipt of every email and in case of non-receipt of acknowledgment from the journal, w.r.t. the submission of manuscript, within two days of submission, the corresponding author is required to demand for the same by sending separate mail to the journal.

2. **MANUSCRIPT TITLE:** The title of the paper should be in a 12 point Calibri Font. It should be bold typed, centered and fully capitalised.

3. **AUTHOR NAME (S) & AFFILIATIONS:** The author (s) **full name**, **designation**, **affiliation** (s), **address**, **mobile/landline numbers**, and **email/alternate email address** should be in italic & 11-point Calibri Font. It must be centered underneath the title.

4. **ABSTRACT:** Abstract should be in fully italicized text, not exceeding 250 words. The abstract must be informative and explain the background, aims, methods, results & conclusion in a single para. Abbreviations must be mentioned in full.

5. **KEYWORDS:** Abstract must be followed by a list of keywords, subject to the maximum of five. These should be arranged in alphabetic order separated by commas and full stops at the end.

6. **MANUSCRIPT:** Manuscript must be in ***BRITISH ENGLISH*** prepared on a standard A4 size ***PORTRAIT SETTING PAPER***. It must be prepared on a single space and single column with 1" margin set for top, bottom, left and right. It should be typed in 8 point Calibri Font with page numbers at the bottom and centre of every page. It should be free from grammatical, spelling and punctuation errors and must be thoroughly edited.

7. **HEADINGS:** All the headings should be in a 10 point Calibri Font. These must be bold-faced, aligned left and fully capitalised. Leave a blank line before each heading.

8. **SUB-HEADINGS:** All the sub-headings should be in a 8 point Calibri Font. These must be bold-faced, aligned left and fully capitalised.

9. **MAIN TEXT:** The main text should follow the following sequence:

INTRODUCTION

REVIEW OF LITERATURE

NEED/IMPORTANCE OF THE STUDY

STATEMENT OF THE PROBLEM

OBJECTIVES

HYPOTHESES

RESEARCH METHODOLOGY

RESULTS & DISCUSSION

FINDINGS

RECOMMENDATIONS/SUGGESTIONS

CONCLUSIONS

SCOPE FOR FURTHER RESEARCH

ACKNOWLEDGMENTS

REFERENCES

APPENDIX/ANNEXURE

It should be in a 8 point Calibri Font, single spaced and justified. The manuscript should preferably not exceed ***5000 WORDS***.

10. **FIGURES &TABLES:** These should be simple, crystal clear, centered, separately numbered &self explained, and **titles must be above the table/figure**. **Sources of data should be mentioned below the table/figure**. It should be ensured that the tables/figures are referred to from the main text.

11. **EQUATIONS:**These should be consecutively numbered in parentheses, horizontally centered with equation number placed at the right.

12. **REFERENCES:** The list of all references should be alphabetically arranged. The author (s) should mention only the actually utilised references in the preparation of manuscript and they are supposed to follow **Harvard Style of Referencing**. The author (s) are supposed to follow the references as per the following:

- All works cited in the text (including sources for tables and figures) should be listed alphabetically.
- Use (**ed.**) for one editor, and (**ed.s)** for multiple editors.
- When listing two or more works by one author, use --- (20xx), such as after Kohl (1997), use --- (2001), etc, in chronologically ascending order.
- Indicate (opening and closing) page numbers for articles in journals and for chapters in books.
- The title of books and journals should be in italics. Double quotation marks are used for titles of journal articles, book chapters, dissertations, reports, working papers, unpublished material, etc.
- For titles in a language other than English, provide an English translation in parentheses.
- The location of endnotes within the text should be indicated by superscript numbers.

PLEASE USE THE FOLLOWING FOR STYLE AND PUNCTUATION IN REFERENCES:

BOOKS

- Bowersox, Donald J., Closs, David J., (1996), "Logistical Management." Tata McGraw, Hill, New Delhi.
- Hunker, H.L. and A.J. Wright (1963), "Factors of Industrial Location in Ohio" Ohio State University, Nigeria.

CONTRIBUTIONS TO BOOKS

- Sharma T., Kwatra, G. (2008) Effectiveness of Social Advertising: A Study of Selected Campaigns, Corporate Social Responsibility, Edited by David Crowther & Nicholas Capaldi, Ashgate Research Companion to Corporate Social Responsibility, Chapter 15, pp 287-303.

JOURNAL AND OTHER ARTICLES

- Schemenner, R.W., Huber, J.C. and Cook, R.L. (1987), "Geographic Differences and the Location of New Manufacturing Facilities," Journal of Urban Economics, Vol. 21, No. 1, pp. 83-104.

CONFERENCE PAPERS

- Garg, Sambhav (2011): "Business Ethics" Paper presented at the Annual International Conference for the All India Management Association, New Delhi, India, 19–22 June.

UNPUBLISHED DISSERTATIONS AND THESES

- Kumar S. (2011): "Customer Value: A Comparative Study of Rural and Urban Customers," Thesis, Kurukshetra University, Kurukshetra.

ONLINE RESOURCES

- Always indicate the date that the source was accessed, as online resources are frequently updated or removed.

WEBSITES

- Garg, Bhavet (2011): Towards a New Natural Gas Policy, Political Weekly, Viewed on January 01, 2012 http://epw.in/user/viewabstract.jsp

DATA MINING APPLICATIONS IN BANKING AND FINANCIAL SECTORS

M JAHNAVI
RESEARCH SCHOLAR
DEPARTMENT OF MBA
SRI PADMAVATI MAHILA VISVAVIDYALAYAM
TIRUPATI

DR. J KATYAYANI
PROFESSOR
DEPARTMENT OF MBA
SRI PADMAVATI MAHILA VISVAVIDYALAYAM
TIRUPATI

ABSTRACT

This article considers building of mathematical models with financial data by using data mining techniques. In general, data mining methods such as neural networks and decision trees can be a useful addition to the techniques available to the financial analyst. It describes data mining in finance by discussing financial task and techniques in this data mining area. Currently, huge electronic data repositories are being maintained by banks and other financial institutions. Valuable bits of information are embedded in these data repositories. The huge size of these data sources make it impossible for a human analyst to come up with interesting information (or patterns) that will help in the decision making process. A number of commercial enterprises have been quick to recognize the value of this concept, as a consequence of which the software market itself for data mining is expected to be in excess of 10 billion USD. This paper is intended for those who would like to get aware of the possible applications of data mining to enhance the performance of some of their core business processes. In this paper discussion is about the broad areas of application, like risk management, portfolio management, trading, customer profiling and customer care, where data mining techniques can be used in banks and other financial institutions to enhance their business performance.

KEYWORDS
Data mining, banking, financial sector.

INTRODUCATION

Data mining involves the use of sophisticated data analysis tools to discover previously unknown, valid patterns and relationships in large data sets. These tools can include statistical models, mathematical algorithms, and machine learning methods (algorithms that improve their performance automatically through experience, such as neural networks or decision trees). Consequently, data mining consists of more than collecting and managing data, it also includes analysis and prediction.

Data mining can be performed on data represented in quantitative, textual, or multimedia forms. Data mining applications can use a variety of parameters to examine the data. They include association (patterns where one event is connected to another event, such as purchasing a pen and purchasing paper), sequence or path analysis (patterns where one event leads to another event, such as the birth of a child and purchasing diapers), classification (identification of new patterns, such as coincidences between duct tape purchases and plastic sheeting purchases), clustering (finding and visually documenting groups of previously unknown facts, such as geographic location and brand preferences), and forecasting (discovering patterns from which one can make reasonable predictions regarding future activities, such as the prediction that people who join an athletic club may take exercise classes).

As an application, compared to other data analysis applications, such as structured queries (used in many commercial databases) or statistical analysis software, data mining represents a *difference of kind rather than degree*. Many simpler analytical tools utilize a verification-based approach, where the user develops a hypothesis and then tests the data to prove or disprove the hypothesis. For example, a user might hypothesize that a customer who buys a hammer, will also buy a box of nails. The effectiveness of this approach can be limited by the creativity of the user to develop various hypotheses, as well as the structure of the software being used. In contrast, data mining utilizes a discovery approach, in which algorithms can be used to examine several multidimensional data relationships simultaneously, identifying those that are unique or frequently represented.

DATA MINING COMMONLY INVOLVES FOUR CLASSES OF TASKS

- **Classification** - Arranges the data into predefined groups. For example, an email program might attempt to classify an email as legitimate or spam. Common algorithms include decision tree learning, nearest neighbor, naïve Bayesian classification and neural networks.
- **Clustering** - Is like classification but the groups are not predefined, so the algorithm will try to group similar items together.
- **Regression** - Attempts to find a function which models the data with the least error.
- **Association rule learning** - Searches for relationships between variables. For example a supermarket might gather data on customer purchasing habits. Using association rule learning, the supermarket can determine which products are frequently bought together and use this information for marketing purposes. This is sometimes referred to as market basket analysis.

HOW DATA MINING WORKS?

1. Business Understanding: The objectives and problems of business are determined and converted to data mining problem.A preliminary plan is prepared.'
2. Understanding the data: The data is collected initially.Information in relation to structure,quality and subset of data are found out.
3. Data Preparation: Final data set is constructed.After sorting and arranging the data and removing unwanted data,the modeling tools are directly applied on final data set.
4. Modeling:There are various modeling techniques like decision trees,rule induction,case base reasoning,visualization techniques, nearest neighbor technique,clustering algorithms etc.Best suited modeling technique is selected Models are combined with different parameters They are compared and ranked for validity and accuracy.
5. Evaluation: Models and steps in modeling are verified with business goals.
6. Deployment: Depending on the assessment and process review, a report is prepared or new data mining project is again set up.

THE SCOPE OF DATA MINING

Data mining derives its name from the similarities between searching for valuable business information in a large database — for example, finding linked products in gigabytes of store scanner data — and mining a mountain for a vein of valuable ore. Both processes require either sifting through an immense amount of material, or intelligently probing it to find exactly where the value resides. Given databases of sufficient size and quality, data mining technology can generate new business opportunities by providing these capabilities:

- **Automated prediction of trends and behaviors**.
- **Automated discovery of previously unknown patterns**.

APPLICATION OF DATA MINING IN BANKING AND FINANCIAL SECTORS

1. MARKETING

Data mining carry various analysis on collected data to determine the consumer behavior with reference to product, price and distribution channel. The reaction of the customers for the existing and new products can also be known based on which banks will try to promote the product, improve quality of products and service and gain competitive advantage. Bank analysts can also analyze the past trends, determine the present demand and forecast the customer behavior of various products and services in order to grab more business opportunities and anticipate behavior patterns. Data mining technique also helps to identify profitable customers from non-profitable ones. Another major area of development in banking is Cross selling i.e. banks makes an attractive offer to its customer by asking them to buy additional product or service. For example, Home loan with insurance facilities and so on. With the help of data mining technique, banks are able to analyze which products and service are availed by most of the customers in cross selling and which type of consumers prefer to purchase cross selling of products and so on.

2. RISK MANAGEMENT

Banks provide loan to its customers by verifying the various details relating to the loan such as amount of loan, lending rate, repayment period, type of property mortgaged, demography, income and credit history of the borrower. Customers with bank for longer periods, with high income groups are likely to get loans very easily. Even though, banks are cautious while providing loan, there are chances for loan defaults by customers. Data mining technique helps to distinguish borrowers who repay loans promptly from those who don't. It also helps to predict when the borrower is at default, whether providing loan to a particular customer will result inbadloansetc. Bank executives by using Data mining technique can also analyze the behavior and reliability of the customers while selling credit cards too. It also helps to analyze whether the customer will make prompt or delay payment if the credit cards are sold to them.

3. PORTFOLIO MANAGEMENT

Risk measurement approaches on an aggregated portfolio level quantify the risk of a set of instrument or customer including diversification effects. On the other hand, forecasting models give an induction of the expected return or price of a financial instrument. Both make it possible to manage firm wide portfolio actively in a risk/return efficient manner.

The application of modern risk theory is therefore within portfolio theory, an important part of portfolio management.With the data mining and optimization techniques investors are able to allocate capital across trading activities to maximise profit or minimise risk. This feature supports the ability to generate trade recommendations and portfolio structuring from user supplied profit and risk requirement. With data mining techniques it is possible to provide extensive scenario analysis capabilities concerning expected asset prices or returns and the risk involved.With this functionality, what if simulations of varying market conditions e.g. interest rate and exchange rate changes) can be run to assess impact on the value and/or risk associated with portfolio, business unit counterparty, or trading desk. Various scenario results can be regarded by considering actual market conditions. Profit and loss analyses allow users to access an asset class, region, counterparty, or custom sub portfolio can be benchmarked against common international benchmarks.

4. FRAUD DETECTION

Nowadays terrorism is one of the biggest troubles that almost every country faces. It mainly influences the economy and the well being of the citizens and this effect is relatively larger in the developed countries. Since the financial sources of terrorist groups can be regarded as black money, the solutions against the money laundering actions can be expected to identify the transactions of the terrorists. Then, blocking their accounts could slow down their actions if cannot stop. In many countries, the financial institutions are expected to inform compliance regulation bodies about any persons or transactions that they think suspicious. Sometimes the given demographics and transaction history of the customers are likely to defraud the bank.Data mining technique helps to analyze such patterns and transactions that lead to fraud.

5. CUSTOMER RETENTION

Today in this competitive environment, customers have wide range of products and services provided by different banks. Hence, banks have to cater the needs of the customer by providing such products and services which they prefer. This will result in customer loyalty and customer retention. Data mining techniques helps to analyze the customers who are loyal from those who shift to other banks for better services. If the customer is shifting from his bank to another reasons for such shifting and the last transaction performed before shifting can be known which will help the banks to perform better and retain its customers.

6. CURRENCY EXCHANGE RATES

The stability of currency plays a very important role in the economic progress of a nation. It is of paramount importance to the investors. It will also serve as a catalyst to economies and yield positive results to support market economy in a situation where the present tsunami is expected to last for a longer time. Therefore, there has to a line of action which may serve as a cushion for the economies of all countries and allow all governments sufficient scope to handle their economic situation. Even small efforts in this regard will count a lot. The Apriori-Algorithm approach determines the selection of most frequently occurring currencies values in a year which is represented by a universal (U) set. The selection from the whole set of various currencies is represented by (S). In this set the most frequently occurring currency with specific dates are noted down and separated from the universal set to yield the result in (X) sets. Specific entries in set (S) are again considered to select more frequently occurring entries among them. This process is continued until desired results are achieved. The simulations have been carried out along with results on the algorithm.

6.1. EXCHANGE RATE

An exchange rate is the rate at which one currency can be exchanged for another. In other words, it is the value of a country's currency compared to that of the other.

6.2. FIXED RATES

There are two ways the price of a currency can be determined against another. A fixed or pegged rate is the rate which the government (central bank) sets and maintains as the official exchange rate. A set price is determined against a major world currency (it is usually the U.S. dollar, but some other major currencies such as euro, yen, pound or a basket of currencies are also considered). In order to maintain the local exchange rate, the central bank buys and sells its own currency in the foreign exchange market in return for the currency to which it is pegged (8).

6.3. FLOATING RATES

Unlike the fixed rate, a floating exchange rate is determined by the private market through supply and demand. A floating rate is often termed "self-correcting", as any differences in supply and demand will automatically be corrected in the market. To be simpler, if demand for a currency is low, its value will decrease, thus making imported goods more expensive and thus stimulating demand for local goods and services. This in return will generate more jobs, and hence an auto-correction would occur in the market. A floating exchange rate is subject to constant change.

7. TRADING

Trading is based on the idea of predicting short term movements in the price/value of a product (currency/equity/interest rate etc.). With a reasonable guesstimate in place one may trade the product if he/she thinks it is going to be overvalued or undervalued in the coming future. If he/she thinks the product is not priced properly he/she may sell/buy it. This instinct is usually based on past experience and some analysis based on market conditions. The price of a financial asset is influenced by a variety of factors which can be broadly classified as economic, political and market factors. Participants in a market observe the

relation between these factors and the price of an asset, account for the current value of these factors and predict the future values to finally arrive at the future value of the asset and trade accordingly.

Data mining techniques are used to discover hidden knowledge, unknown patterns and new rules from large data sets, which may be useful for a variety of decision making activity. Advancements made in technology have enabled to create faster and better prediction systems. These systems are based on a combination of data mining techniques and artificial intelligence methods like Case Based Reasoning (CBR) and Neural Networks (NN). A combination of such a forecasting system together with a good trading strategy offers tremendous opportunities for massive returns. CBR methods can be used in real time which makes analysis really quick and helps in real time decision making resulting in immediate profits. Thus data mining and business intelligence (CBR and NN) techniques may be used in conjunction in financial markets to predict market behavior and obtain patterned behavior to influence decision making.

DATA MINING ALGORITHMS

The most commonly used techniques in data mining are:

Data mining uses a variety of techniques to find hidden patterns and relationships in large pools of data and infer rules from them that can be used to predict future behavior and guide decision making. The *data mining algorithm* is the mechanism that creates a data mining model. To create a model, an algorithm first analyzes a set of data and looks for specific patterns and trends. The algorithm uses the results of this analysis to define the parameters of the mining model. These parameters are then applied across the entire data set to extract actionable patterns and detailed statistics.

The mining model that an algorithm creates can take various forms, including:

- A set of rules that describe how products are grouped together in a transaction.
- A decision tree that predicts whether a particular customer will buy a product.
- A mathematical model that forecasts sales.
- A set of clusters that describe how the cases in a dataset are related.

Types of Data Mining Algorithms:

1. **CLASSIFICATION ALGORITHM -** predict one or more discrete variables, based on the other attributes in the dataset. An example of a classification algorithm is the Microsoft Decision Trees Algorithm.

1.1 MICROSOFT DECISION TREES ALGORITHM

The Microsoft Decision Trees algorithm is a classification and regression algorithm used in predictive modeling of both discrete and continuous attributes.

For discrete attributes, the algorithm makes predictions based on the relationships between input columns in a dataset. It uses the values, known as states, of those columns to predict the states of a column that you designate as predictable. Specifically, the algorithm identifies the input columns that are correlated with the predictable column. For example, in a scenario to predict which customers are likely to purchase a bicycle, if nine out of ten younger customers buy a bicycle, but only two out of ten older customers do so, the algorithm infers that age is a good predictor of bicycle purchase. The decision tree makes predictions based on this tendency toward a particular outcome.

For continuous attributes, the algorithm uses linear regression to determine where a decision tree splits. If more than one column is set to predictable, or if the input data contains a nested table that is set to predictable, the algorithm builds a separate decision tree for each predictable column.

1.2 MICROSOFT NAIVE BAYES ALGORITHM

The Microsoft Naive Bayes algorithm is a classification algorithm used in predictive modeling. The name Naive Bayes derives from the fact that the algorithm uses Bayes theorem but does not take into account dependencies that may exist, and therefore its assumptions are said to be naive.

This algorithm is less computationally intense than other Microsoft algorithms, and therefore is useful for quickly generating mining models to discover relationships between input columns and predictable columns. You can use this algorithm to do initial explorations of data, and then later you can apply the results to create additional mining models with other algorithms that are more computationally intense and more accurate.

1.3 MICROSOFT NEURAL NETWORK ALGORITHM

Neural Network algorithm combines each possible state of the input attribute with each possible state of the predictable attribute, and uses the training data to calculate probabilities. You can later use these probabilities for classification or regression, and to predict an outcome of the predicted attribute, based on the input attributes.

A mining model that is constructed with the Microsoft Neural Network algorithm can contain multiple networks, depending on the number of columns that are used for both input and prediction, or that are used only for prediction. The number of networks that a single mining model contains depends on the number of states that are contained by the input columns and predictable columns that the mining model uses.

2. **REGRESSION ALGORITHM -** predict one or more continuous variables, such as profit or loss, based on other attributes in the dataset. An example of a regression algorithm is the Microsoft Time Series Algorithm.

2.1 MICROSOFT TIME SERIES ALGORITHM

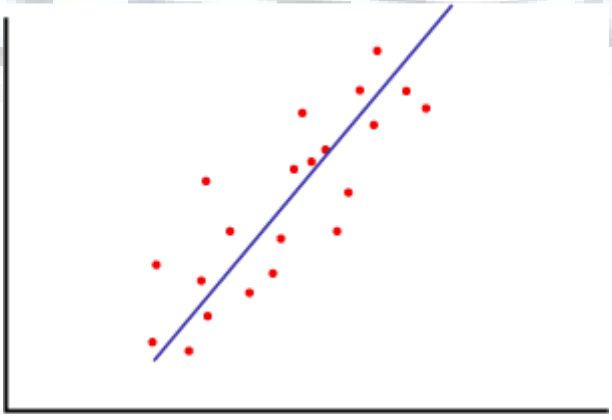
The Microsoft Time Series algorithm provides regression algorithms that are optimized for the forecasting of continuous values, such as product sales, over time. Whereas other Microsoft algorithms, such as decision trees, require additional columns of new information as input to predict a trend, a time series model does not. A time series model can predict trends based only on the original dataset that is used to create the model. You can also add new data to the model when you make a prediction and automatically incorporate the new data in the trend analysis.

An important feature is that it can perform cross prediction. If you train the algorithm with two separate, but related, series, you can use the resulting model to predict the outcome of one series based on the behavior of the other series. For example, the observed sales of one product can influence the forecasted sales of another product. Cross prediction is also useful for creating a general model that can be applied to multiple series.

2.2 MICROSOFT LINEAR REGRESSION ALGORITHM

The Microsoft Linear Regression algorithm is a variation of the Microsoft Decision Trees algorithm that helps you calculate a linear relationship between a dependent and independent variable, and then use that relationship for prediction.

The relationship takes the form of an equation for a line that best represents a series of data. For example, the line in the following diagram is the best possible linear representation of the data.



Each data point in the diagram has an error associated with its distance from the regression line. The coefficients *a* and *b* in the regression equation adjust the angle and location of the regression line. You can obtain the regression equation by adjusting *a* and *b* until the sum of the errors that are associated with all the points reaches its minimum. There are other kinds of regression that use multiple variables, and also nonlinear methods of regression. However, linear regression is a useful and well-known method for modeling a response to a change in some underlying factor.

2.3 MICRSOFT LOGISTIC REGRESSION ALGORITHM

The Microsoft Logistic Regression algorithm is a variation of the Microsoft Neural Network algorithm. Logistic regression is a well-known statistical technique that is used for modeling binary outcomes, such as a yes-No outcome.

Logistic regression is highly flexible, taking any kind of input, and supports several different analytical tasks:

- Use demographics to make predictions about outcomes, such as risk for a certain disease.
- Explore and weight the factors that contribute to a result. For example, find the factors that influence customers to make a repeat visit to a store.
- Classify documents, e-mail, or other objects that have many attributes.

3. **SEGMENTATION ALGORITHM** divide data into groups, or clusters, of items that have similar properties. An example of a segmentation algorithm is the Microsoft Clustering Algorithm.

3.1 **MICROSOFT CLUSTERING ALGORITHM**

The Microsoft Clustering algorithm is a segmentation algorithm. The algorithm uses iterative techniques to group cases in a dataset into clusters that contain similar characteristics. These groupings are useful for exploring data, identifying anomalies in the data, and creating predictions.

Clustering models identify relationships in a dataset that you might not logically derive through casual observation. The clustering algorithm differs from other data mining algorithms, such as the Microsoft Decision Trees algorithm, in that you do not have to designate a predictable column to be able to build a clustering model. The clustering algorithm trains the model strictly from the relationships that exist in the data and from the clusters that the algorithm identifies.

4. **ASSOCIATION ALGORITHM -** it find correlations between different attributes in a dataset. The most common application of this kind of algorithm is for creating association rules, which can be used in a market basket analysis. An example of an association algorithm is the Microsoft Association Algorithm.

4.1 MICROSOFT ASSOCIATION ALGORITHM

The Microsoft Association algorithm is an association algorithm that is useful for recommendation engines. A recommendation engine recommends products to customers based on items they have already bought, or in which they have indicated an interest. Association models are built on datasets that contain identifiers both for individual cases and for the items that the cases contain. A group of items in a case is called an *itemset*. An association model consists of a series of itemsets and the rules that describe how those items are grouped together within the cases.

5. **SEQUENCE CLUSTERING ALGORITHM -** summarize frequent sequences or episodes in data, such as a Web path flow. An example of a sequence analysis algorithm is the Microsoft Sequence Clustering Algorithm.

5.1 MICROSOFT SEQUENCE CLUSTERING ALGORITHM

The Microsoft Sequence Clustering algorithm is a sequence analysis algorithm. You can use this algorithm to explore data that contains events that can be linked by following paths, or *sequences*. The algorithm finds the most common sequences by grouping, or clustering, sequences that are identical. This algorithm is similar in many ways to the Microsoft Clustering algorithm. However, instead of finding clusters of cases that contain similar attributes, the Microsoft Sequence Clustering algorithm finds clusters of cases that contain similar paths in a sequence.

APPLYING THE ALGORITHMS

Task	Microsoft algorithms to use
Predicting a discrete attribute. For example, predict whether the recipient of a targeted mailing campaign will buy a product.	Microsoft Decision Trees Algorithm Microsoft Naive Bayes Algorithm Microsoft Clustering Algorithm Microsoft Neural Network Algorithm
Predicting a continuous attribute. For example, forecast next year's sales.	Microsoft Decision Trees Algorithm Microsoft Time Series Algorithm
Predicting a sequence. For example, perform a click stream analysis of a company's Web site.	Microsoft Sequence Clustering Algorithm
Finding groups of common items in transactions. For example, use market basket analysis to suggest additional products to a customer for purchase.	Microsoft Association Algorithm Microsoft Decision Trees Algorithm
Finding groups of similar items. For example, segment demographic data into groups to better understand the relationships between attributes.	Microsoft Clustering Algorithm Microsoft Sequence Clustering Algorithm

CONCLUSION

Comprehensive data warehouses that integrate operational data with customer, supplier, and market information have resulted in an explosion of information. Competition requires timely and sophisticated analysis on an integrated view of the data. However, there is a growing gap between more powerful storage and retrieval systems and the users' ability to effectively analyze and act on the information they contain. Both relational and OLAP technologies have tremendous capabilities for navigating massive data warehouses, but brute force navigation of data is not enough. A new technological leap is needed to structure and prioritize information for specific end-user problems. The data mining tools can make this leap. Quantifiable business benefits have been proven through the integration of data mining with current information systems, and new products are on the horizon that will bring this integration to an even wider audience of users. Data mining techniques can be of immense help to the banks and financial institutions in this arena for better targeting and acquiring new customers, fraud detections in real time, providing segment based products for better targeting the customers, detection of emerging trends to take proactive stance in a highly competitive market adding a lot more value to existing products and services. To be successful a data mining project should be driven by the application needs and results should be tested quickly. Financial applications provide a unique environment where efficiency of the methods can be tested instantly, not only by using traditional training and testing data but making real stock forecast and testing it the same day. This process can be repeated daily for several months collecting quality estimates.

REFERENCES

1. Artificial Intelligence: A Modern Approach (2nd Edition) . New Jersey: Pearson Education.
2. Berthold, M., & Hand, D. J. (2007). Intelligent Data Analysis. Springer.
3. D. Hand, H. Mannila, and P. Smyth, Principles of Data Mining, MIT Press, 2001.
4. Freitas, A. A. (2002). Data Mining and Knowledge Discovery with Evolutionary Algorithms. Springer.
5. Hastie, R. Tibshirani, J. Friedman, The Elements of Statistical Learning: Data Mining, Inference and Prediction (2nded), Springer Verlag, 2008

6. J. Han and M. Kamber, Data Mining: Concepts and Techniques, Morgan Kaufmann, 2000.
7. M. Berry and G. Linoff, Mastering Data Mining: The Art and Science of Customer Relationship Management, Wiley, 2000.
8. Richard J.Roiger, M. G. (2002). Data Mining: A Tutorial Based Primer. Addison Wesley. Russel, S., & Norvig, P. (2003).
9. Tan, P.-N., Steinbach, M., & Kumar, V. (2005). Introduction to Data Mining. Addison-Wesley.
10. Weigend, A. S. (1999, January 20). Data mining in finance - Course notes. Retrieved 2011, from http://wwwpsych.stanford.edu/~andreas /Teaching/ Data MiningFinance/S99/Notes/01Description.html

REQUEST FOR FEEDBACK

Dear Readers

At the very outset, International Journal of Research in Commerce & Management (IJRCM) acknowledges & appreciates your efforts in showing interest in our present issue under your kind perusal.

I would like to request you tosupply your critical comments and suggestions about the material published in this issue as well as on the journal as a whole, on our E-mail**infoijrcm@gmail.com** for further improvements in the interest of research.

If youhave any queries please feel free to contact us on our E-mail infoijrcm@gmail.com.

I am sure that your feedback and deliberations would make future issues better – a result of our joint effort.

Looking forward an appropriate consideration.

With sincere regards

Thanking you profoundly

Academically yours

Sd/-
Co-ordinator

DISCLAIMER

The information and opinions presented in the Journal reflect the views of the authors and not of the Journal or its Editorial Board or the Publishers/Editors. Publication does not constitute endorsement by the journal. Neither the Journal nor its publishers/Editors/Editorial Board nor anyone else involved in creating, producing or delivering the journal or the materials contained therein, assumes any liability or responsibility for the accuracy, completeness, or usefulness of any information provided in the journal, nor shall they be liable for any direct, indirect, incidental, special, consequential or punitive damages arising out of the use of information/material contained in the journal. The journal, nor its publishers/Editors/ Editorial Board, nor any other party involved in the preparation of material contained in the journal represents or warrants that the information contained herein is in every respect accurate or complete, and they are not responsible for any errors or omissions or for the results obtained from the use of such material. Readers are encouraged to confirm the information contained herein with other sources. The responsibility of the contents and the opinions expressed in this journal is exclusively of the author (s) concerned.

ABOUT THE JOURNAL

In this age of Commerce, Economics, Computer, I.T. & Management and cut throat competition, a group of intellectuals felt the need to have some platform, where young and budding managers and academicians could express their views and discuss the problems among their peers. This journal was conceived with this noble intention in view. This journal has been introduced to give an opportunity for expressing refined and innovative ideas in this field. It is our humble endeavour to provide a springboard to the upcoming specialists and give a chance to know about the latest in the sphere of research and knowledge. We have taken a small step and we hope that with the active co-operation of like-minded scholars, we shall be able to serve the society with our humble efforts.

Our Other Journals