

INTERNATIONAL JOURNAL OF RESEARCH IN COMPUTER APPLICATION & MANAGEMENT

I
J
R
C
M



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-Gate, India [link of the same is duly available at Inlibnet 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 4600 Cities in 180 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

Sr. No.	TITLE & NAME OF THE AUTHOR (S)	Page No.
1.	CUSTOMER SATISFACTION IN THE COMMERCIAL BANKS WITH SPECIAL REFERENCE TO NEPAL <i>KESHAV RAJ BHATTA & DR. B. P. SINGHAL</i>	1
2.	COMPUTER VIRTUALIZATION: STRATEGIC ADVANTAGE TO BUSINESS <i>SOMANING TURWALE</i>	5
3.	ROLE OF MANAGEMENT INFORMATION SYSTEM FOR DECISION MAKING PROCESS IN THE ORGANIZATION <i>DR. SAMPADA GULAVANI</i>	7
4.	GAP OF LEADERSHIP SKILLS AMONGST MALE AND FEMALE STUDENTS <i>DR. D M KUMTHEKAR & DR. PAUL D MADHALE</i>	10
5.	CASE STUDY ON MINING BIG DATA <i>M.DHANAMALAR & B.AYSHWARYA</i>	12
6.	SECURE DATA TRANSMISSION IN MOBILE AD-HOC NETWORK USING RSDSV PROTOCOL <i>A. P. PANDE & A. U. PATIL</i>	14
7.	ENHANCEMENT OF TEMPORAL DATA CLUSTERING WITH CLIPPED GAUSSIAN DISTRIBUTION <i>SARALA.R</i>	18
8.	A COMPARATIVE STUDY ON IDENTIFYING USAGE OF MOBILE BANKING SERVICES IN SELECTED PUBLIC AND PRIVATE SECTOR BANKS <i>DR. S.T. BHOSALE</i>	23
9.	A REVIEW ON THE ROLE OF DATA MINING IN BANK <i>VARSHA MOHITE</i>	30
10.	AN EVALUATION OF WEB SERVICES USED BY CUSTOMERS OF SELECTED BUSINESS SECTORS IN ISLAMPUR CITY <i>MANE HIMMATRAO MAHADEO</i>	32
11.	A PERFORMANCE EVALUATION OF SOFTWARE INSTALLED IN RAJARAMBAPU MILK DAIRY, ISLAMPUR <i>D. R. PATIL</i>	36
12.	IMPACT OF 'MAKE IN INDIA' ON FOREIGN DIRECT INVESTMENT <i>SHRADDHA PRASAD KULKARNI</i>	40
13.	PATIENT'S SATISFACTION TOWARDS SERVICE QUALITY: A COMPARATIVE ANALYSIS OF GOVERNMENT AND MISSION HOSPITAL, MIRAJ <i>M. M. SAMUDRE</i>	43
14.	EFFECTIVE USE OF JOB SHARING IN WORK-LIFE BALANCE <i>M. M. PATIL</i>	46
15.	AN REVIEW OF INTERNET BANKING IN INDIA <i>R. T. THORAT</i>	48
16.	FINDING INVISIBLE AND NON-MEASURABLE IMPEDENCE VALUES INSIDE THE OBJECT <i>A. R. RASAL</i>	50
17.	A ROLE OF DATA MINING TECHNIQUES IN SHARE MARKET INVESTMENT <i>H. SANKAPAL</i>	54
18.	WICAT: CHAT AND FILE TRANSFER APPLICATION <i>PRADNYAWANT MALAGE</i>	58
19.	DIGITAL PRINTING IN INDIA: SWOT ANALYSIS <i>S.V. CHAVAN</i>	60
20.	A STUDY ON THE IMPORTANCE OF CORPORATE SOCIAL RESPONSIBILITY IN INDIAN COMPANIES <i>AMRUTHA SANTHOSH, LEKSHMY P.R. & NAYANA N NAIR</i>	62
	REQUEST FOR FEEDBACK & DISCLAIMER	65

CHIEF PATRON

PROF. K. K. AGGARWAL

Chairman, Malaviya National Institute of Technology, Jaipur
(An institute of National Importance & fully funded by Ministry of Human Resource Development, Government of India)
Chancellor, K. R. Mangalam University, Gurgaon
Chancellor, Lingaya's University, Faridabad
Founder Vice-Chancellor (1998-2008), Guru Gobind Singh Indraprastha University, Delhi
Ex. Pro Vice-Chancellor, Guru Jambheshwar University, Hisar

FOUNDER PATRON

LATE SH. RAM BHAJAN AGGARWAL

Former State Minister for Home & Tourism, Government of Haryana
Former Vice-President, Dadri Education Society, Charkhi Dadri
Former President, Chinar Syntex Ltd. (Textile Mills), Bhiwani

FORMER CO-ORDINATOR

DR. S. GARG

Faculty, Shree Ram Institute of Business & Management, Urjani

ADVISORS

PROF. M. S. SENAM RAJU

Director A. C. D., School of Management Studies, I.G.N.O.U., New Delhi

PROF. S. L. MAHANDRU

Principal (Retd.), Maharaja Agrasen College, Jagadhri

EDITOR

PROF. R. K. SHARMA

Professor, Bharti Vidyapeeth University Institute of Management & Research, New Delhi

EDITORIAL ADVISORY BOARD

DR. RAJESH MODI

Faculty, Yanbu Industrial College, Kingdom of Saudi Arabia

PROF. PARVEEN KUMAR

Director, M.C.A., Meerut Institute of Engineering & Technology, Meerut, U. P.

PROF. H. R. SHARMA

Director, Chhatrapati Shivaji Institute of Technology, Durg, C.G.

PROF. MANOHAR LAL

Director & Chairman, School of Information & Computer Sciences, I.G.N.O.U., New Delhi

PROF. ANIL K. SAINI

Chairperson (CRC), Guru Gobind Singh I. P. University, Delhi

PROF. R. K. CHOUDHARY

Director, Asia Pacific Institute of Information Technology, Panipat

DR. ASHWANI KUSH

Head, Computer Science, University College, Kurukshetra University, Kurukshetra

DR. BHARAT BHUSHAN

Head, Department of Computer Science & Applications, Guru Nanak Khalsa College, Yamunanagar

DR. VIJAYPAL SINGH DHAKA

Dean (Academics), Rajasthan Institute of Engineering & Technology, Jaipur

DR. SAMBHAVNA

Faculty, I.I.T.M., Delhi

DR. MOHINDER CHAND

Associate Professor, Kurukshetra University, Kurukshetra

DR. MOHENDER KUMAR GUPTA

Associate Professor, P. J. L. N. Government College, Faridabad

DR. SHIVAKUMAR DEENE

Asst. Professor, Dept. of Commerce, School of Business Studies, Central University of Karnataka, Gulbarga

DR. BHAVET

Faculty, Shree Ram Institute of Engineering & Technology, Urjani

ASSOCIATE EDITORS**PROF. ABHAY BANSAL**

Head, Department of Information Technology, Amity School of Engineering & Technology, Amity University, Noida

PROF. NAWAB ALI KHAN

Department of Commerce, Aligarh Muslim University, Aligarh, U.P.

ASHISH CHOPRA

Sr. Lecturer, Doon Valley Institute of Engineering & Technology, Karnal

FORMER TECHNICAL ADVISOR**AMITA**

Faculty, Government M. S., Mohali

FINANCIAL ADVISORS**DICKIN GOYAL**

Advocate & Tax Adviser, Panchkula

NEENA

Investment Consultant, Chambaghat, Solan, Himachal Pradesh

LEGAL ADVISORS**JITENDER S. CHAHAL**

Advocate, Punjab & Haryana High Court, Chandigarh U.T.

CHANDER BHUSHAN SHARMA

Advocate & Consultant, District Courts, Yamunanagar at Jagadhri

SUPERINTENDENT**SURENDER KUMAR POONIA**

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/Mkt./HRM/General Mgt./Engineering/Economics/Computer/IT/ Education/Psychology/Law/Math/other, please specify)

DEAR SIR/MADAM

Please find my submission of manuscript entitled ' _____ ' for possible publication in one of 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 co-authors of this manuscript have seen the submitted version of the manuscript and have agreed to their inclusion of names as co-authors.

Also, if my/our manuscript is accepted, I agree to comply with the formalities as given on the website of the journal. The Journal has discretion to publish our contribution in any of its journals.

NAME OF CORRESPONDING AUTHOR :

Designation :

Institution/College/University with full address & Pin Code :

Residential address with Pin Code :

Mobile Number (s) with country ISD code :

Is WhatsApp or Viber active on your above noted Mobile Number (Yes/No) :

Landline Number (s) with country ISD code :

E-mail Address :

Alternate E-mail Address :

Nationality :

NOTES:

- a) The whole manuscript has to be in **ONE MS WORD FILE** only, which will start from the covering letter, inside the manuscript. **pdf. version is liable to be rejected without any consideration.**
 - b) The sender is required to mention the following in the **SUBJECT COLUMN of the mail:**
New Manuscript for Review in the area of (e.g. Finance/Marketing/HRM/General Mgt./Engineering/Economics/Computer/IT/ Education/Psychology/Law/Math/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 expected to be below **1000 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 within twenty four hours** and in case of non-receipt of acknowledgement 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 a separate mail to the journal.
 - g) The author (s) name or details should not appear anywhere on the body of the manuscript, except the covering letter and the cover page of the manuscript, in the manner as mentioned in the guidelines.
2. **MANUSCRIPT TITLE:** The title of the paper should be **bold typed, centered and fully capitalised**.
 3. **AUTHOR NAME (S) & AFFILIATIONS:** Author (s) **name, designation, affiliation (s), address, mobile/landline number (s), and email/alternate email address** should be given underneath the title.
 4. **ACKNOWLEDGMENTS:** Acknowledgements can be given to reviewers, guides, funding institutions, etc., if any.
 5. **ABSTRACT:** Abstract should be in **fully italicized text**, ranging between **150 to 300 words**. The abstract must be informative and explain the background, aims, methods, results & conclusion in a **SINGLE PARA**. **Abbreviations must be mentioned in full.**
 6. **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 stop at the end. All words of the keywords, including the first one should be in small letters, except special words e.g. name of the Countries, abbreviations.
 7. **JEL CODE:** Provide the appropriate Journal of Economic Literature Classification System code (s). JEL codes are available at www.aeaweb.org/econlit/jelCodes.php, however, mentioning JEL Code is not mandatory.
 8. **MANUSCRIPT:** Manuscript must be in **BRITISH ENGLISH** prepared on a standard A4 size **PORTRAIT SETTING PAPER**. **It should be free from any errors i.e. grammatical, spelling or punctuation. It must be thoroughly edited at your end.**
 9. **HEADINGS:** All the headings must be bold-faced, aligned left and fully capitalised. Leave a blank line before each heading.
 10. **SUB-HEADINGS:** All the sub-headings must be bold-faced, aligned left and fully capitalised.
 11. **MAIN TEXT:**

THE MAIN TEXT SHOULD FOLLOW THE FOLLOWING SEQUENCE:**INTRODUCTION****REVIEW OF LITERATURE****NEED/IMPORTANCE OF THE STUDY****STATEMENT OF THE PROBLEM****OBJECTIVES****HYPOTHESIS (ES)****RESEARCH METHODOLOGY****RESULTS & DISCUSSION****FINDINGS****RECOMMENDATIONS/SUGGESTIONS****CONCLUSIONS****LIMITATIONS****SCOPE FOR FURTHER RESEARCH****REFERENCES****APPENDIX/ANNEXURE****The manuscript should preferably range from 2000 to 5000 WORDS.**

12. **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.*
13. **EQUATIONS/FORMULAE:** These should be consecutively numbered in parenthesis, horizontally centered with equation/formulae number placed at the right. The equation editor provided with standard versions of Microsoft Word should be utilised. If any other equation editor is utilised, author must confirm that these equations may be viewed and edited in versions of Microsoft Office that does not have the editor.
14. **ACRONYMS:** These should not be used in the abstract. The use of acronyms is elsewhere is acceptable. Acronyms should be defined on its first use in each section: Reserve Bank of India (RBI). Acronyms should be redefined on first use in subsequent sections.
15. **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. **Also check to make sure that everything that you are including in the reference section is duly cited in the paper.** 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 parenthesis.
 - **Headers, footers, endnotes and footnotes should not be used in the document.** However, **you can mention short notes to elucidate some specific point**, which may be placed in number orders after the references.

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, 19–23

UNPUBLISHED DISSERTATIONS

- 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 Gas Policy, Political Weekly, Viewed on January 01, 2012 <http://epw.in/user/viewabstract.jsp>

A ROLE OF DATA MINING TECHNIQUES IN SHARE MARKET INVESTMENT

H. SANKAPAL
ASST. PROFESSOR
MIRAJ MAHAVIDYALAYA
MIRAJ


ABSTRACT

Knowledge discovery in financial organization have been built and operated mainly to support decision making using knowledge as strategic factor. Data mining can be defined as the process of selecting, exploring and modeling large amounts of data to uncover previously unknown patterns. In the Share Market in equity sector, data mining can help firms gain business advantage Many attempts have been made to predict stock market data using statistical and traditional methods, but these methods are no longer adequate for analyzing this huge amount of data. In this paper, we investigate the use of various data mining techniques for knowledge discovery in investment in Share Market in equity sector. We introduce different exhibits for discovering knowledge in the form of association rules, clustering, classification suitable for data characteristics. Data mining is one of most important powerful information technology tool in today's competitive business world, it is able to uncover hidden patterns and predict future trends and behavior in stock market.

KEYWORDS

share market, equity, association rules, clustering, classification, data mining.

1. INTRODUCTION

 stock market is a place where buying and selling of shares takes place. When an investor buys stocks of a certain company he becomes a part owner of that company according to the number of shares held by him. To make maximum profit, right investment should be made at the right time. The main function of a stock market is the dealings of stocks between investors. Stocks are grouped into industry groups according to their primary business focus. A transaction is the willing of an investor to sell some stocks and the request of another to buy them. Normally purchasing of share depends upon occupation and income of investor. Following table shows current trend of investment based on occupation and income.

Data mining can be defined as the process of selecting, exploring and modeling large amounts of data to uncover previously unknown patterns. In the Share Market in equity sector, data mining can help firms gain business advantage. For example, by applying data mining techniques, companies can fully exploit data about customers' buying patterns and behavior – as well as gaining a greater understanding of their business to help reduce fraud, improve underwriting and enhance risk management. This paper discusses how help to customers can benefit by using modern data mining methodologies and thereby selection of shares in equity, increase profits, acquire new customers. Data mining methodology often can improve upon traditional statistical approaches to solving business solutions. Stock market produces huge datasets that deals enormously complex and dynamic problems with data mining tool. Potential significant benefits of solving these problems motivated extensive research for years . The research in data mining has gained a high attraction due to the importance of its applications and increasing generated information Data mining often can improve existing models by finding additional, important variables, identifying interaction terms and detecting nonlinear relationships. Models that predict relationships and behaviors more accurately lead to greater profits and reduced Risk taking ability. Data classification can be done by many different techniques; one of those methods is the classification by using Decision Tree. It is graphical demonstration of all possible paths and outcomes by which they can be reached.

Specifically, data mining can help stockbrokers in the business of bringing buyers and sellers of stocks and securities together ,increasing generated information. A stock market or equity market is a private or public market for the trading of company stock and derivatives of company stock at an agreed price; there are securities listed on a stock exchange as well as those only traded privately.

- ★ Acquiring new customers for Opening New Demat Account.
- ★ Retaining existing customers.
- ★ Performing sophisticated classification

2. ACQUIRING NEW CUSTOMERS FOR OPENING NEW DEMAT ACCOUNT

An important business problem is the acquisition of new customers. Although traditional approaches involve attempts to increase the customer base by simply expanding the efforts of the stock brokers efforts that are guided by more quantitative data mining approaches can lead to more focused and more successful results. A traditional approach is to increase the number of Demat account holders by simply targeting those who meet certain Shares constraints. A drawback to this approach is that much of the marketing effort may yield little return. Hence in this situation it is important to identify population segments among already demat account holders customers in trading company could be targeted. A statistical technique called "cluster analysis," sometimes used in the private sector to identify various market segments. Clustering is a technique of partitioning or segmenting the data into groups that might or might not be disjointed. The clustering usually accomplished by determining the similarity among the data on predefined attributes. The most similar data are grouped into clusters. Since clusters are not predefined, a domain expert is often required to interpret the meaning of the created clusters.

EXAMPLE 1.1

Trading companies can create special catalogs targeted to various demographic groups based on attributes such as income, occupation, and age as physical characteristic of potential customers. The company then can perform a clustering of potential customers based on determined attribute values to create new catalogs. The results of the clustering exercise can be then used by management to create special catalogs for different Shares and distribute them to the correct target population based on the cluster for that policy.

A Trading company can group its customers based on common features. Company management does not have any predefined for this label. Based on the outcome of the grouping they will target marketing and advertising campaigns to the different groups for a particular type of policy. Following is the general trend observed by experts in share market.

TABLE 1.1: TREND FOR INVESTMENT IN SHARE MARKET

Occupation	Income	Low level shares	Mid level shares	High level shares
Employee	3,000 to 10000	10%		
	10000 to 20000	10%	10%	
	20000 to 30000		20%	
	30000 to 40000	10%		
	40000 to 50000			40%
	>50000		20%	20%
Business man	3,000 to 10000	5%		
	30000 to 40000		40%	
	>50000			60%
Professional	20000 to 30000		40%	
	>50000			60%
Student	3,000 to 10000	10%		
	20000 to 30000		20%	
	>50000			40%
Others	20000 to 30000		10%	
	>50000			60%

TABLE 1.2: SAMPLE DATA BASED ON TRENDS

Occupation	Income	Investment	Type of Share
Employee	8500	1000	Low level
Employee	18000	4000	Low, Mid level
Employee	60000	25000	Low, Mid level
Business Man	38000	15000	Mid level
Professional	28000	6000	Mid level
Employee	125000	30000	Mid, High level
Business Man	70000	45000	High level
Employee	55000	20000	Mid, High level
Professional	120000	75000	High level

The sample data shows the investment done by the customer for particular type of shares (Low-level risk, Mid-level, High-level) depending upon their occupation, and income. Hence depending upon the occupation and income, we could target the customers having less, high income and occupation as employee, professional, students, others, Business man. Hence the group of people, is of younger Customers having knowledge of share market suitable for investment in low, mid, high level shares in equity.

DEFINITION 1.1. Given a database $D = \{t_1, t_2, \dots, t_n\}$ of tuples and an integer value k , the clustering problem is to define a mapping $f : D \rightarrow \{1, \dots, k\}$ where each t_i is assigned to one cluster k_j , $1 \leq j \leq k$. A cluster k_j , contains precisely those tuples mapped to it that is, $k_j = \{t_i \mid f(t_i) = k_j, 1 \leq i \leq n, \text{ and } t_i \in D\}$

Algorithm 1.1 k-means Clustering

K-means is an iterative clustering algorithm in which items are moved among sets of clusters until the desired set is reached

Input:

$D = \{t_1, t_2, t_3, \dots, t_n\}$ //Set of elements

k //Number of desired clusters

Output:

K //set of clusters

Algorithm:

assign initial values for means $m_1, m_2, \dots,$

m_k ; repeat

assign each item t_i to the cluster which has closest

mean; calculate new mean for each cluster;

until convergence criteria is met.

Note that the initial values for means are arbitrarily assigned and the algorithm could stop when no or very small number of tuples are assigned to different clusters. As per the algorithm, first we have to find mean of each cluster. Hence accordingly mean for first cluster is 28500, Employee, 15000, Occupation, Income. Similarly mean for second cluster is 60000, Employee, Professional, Business Men, 25500, while the same for third cluster is 80000, Business, Employee, Professional, 80000. Suppose a customer with income is 25000, occupation Employee, Investment is 14000 and will provide differences 3000, 0, 1000, 0. Once the customer is added to one of the clusters its new mean will be automatically calculated.

3. CLASSIFICATION: SEGMENT DATABASES

To improve predictive accuracy, databases can be segmented into more homogeneous groups. Then the data of each group can be explored, analyzed and modeled. Depending on the business question, segmentation can be done using variables associated with risk factors, profits or behaviors. Segments based on these types of variables often provide sharp contrasts, which can be interpreted more easily. Classification maps data into predefined groups or segments. Classification algorithms require that the classes be defined based on data attributes values. They often describe these classes by looking at the characteristics of data already known to belong to the classes. As a result, Trading companies can more accurately predict the likelihood of a shares in equity based on the upon income and occupation.

EXAMPLE 1.2

Trading company can find a segment based on the income, Occupations. Such patterns can be stored in database. So while to customer, broker can get the information of customer like income and occupation. This pattern can be compared to entries in a database and broker can suggest low level share, mid level share, high level shares in equity customer based on matched patterns.

TABLE 1.2 SAMPLE DATA FOR EXAMPLE

Occupation	Income	Investment	Type of Share
Employee	8500	1000	Low level
Employee	18000	4000	Low, Mid level
Employee	12000	2500	Low, Mid level
Employee	15000	3500	Low, Mid level
Employee	17500	4500	Low, Mid level
Employee	60000	25000	Low, Mid level
Employee	15000	15000	Mid, High level
Employee	125000	30000	Mid, High level
Business man	38000	15000	Mid level
Business man	70000	45000	High level
Business man	9000	500	Low level
Professional	28000	6000	Mid level
Professional	120000	75000	High level
Student	8500	1000	Low level
Student	25000	5000	Mid level
Student	60000	25000	High level
Others	27000	3000	Mid level
Others	60000	35000	High level

This example assumes that the problem is to classify customers in terms of different investments attributes such as occupation, income, mid point. The investments classification can simply be done using income as main criteria shown below:

TABLE 1.3

3,000	≤ Income ≤ 20,000	Low level
15,000	≤ Income ≤ 40,000	Mid level
Income ≥ 40,000		High level

The investments term require complicated set of divisions using both income and Occupation. Similarly Risk shares level require complicated set of divisions using both Income and Occupation while investments require much more complicated set of divisions using Age, Income and Occupation.

u N such that $sim(t,u) \leq sim(t,d)$, then begin

$N = N - \{u\}$;

$N = N \{d\}$;

DEFINITION 1.1. Given a database $D = \{t_1, t_2, \dots, t_n\}$ of tuples (items, records) and a set of classes $C = \{C_1, \dots, C_m\}$, the classification problem is to define a mapping $f : D \rightarrow C$ where each t_i is assigned to one class. A class C_j , contains precisely those tuples mapped to it that is, $C_j = \{t_i \mid f(t_i) = C_j, 1 \leq i \leq n, \text{ and } t_i \in D\}$

Algorithm 1.2 K Nearest Neighbors

When classification is to be made for new item using K Nearest Neighbors algorithm, its distance to each item in the training set must be determined. The new item is then placed in the class that contains the most items from the (K) closest set.

Input:

T//Training data

K //Number of neighbors

t //Input tuple to classify

Output:

c //class to which t is

assigned Algorithm:

$N = \emptyset$

//Find the set of neighbors, N, for

t For each d T do

If $|N| \leq K$, then

$N = N \{d\}$; \uparrow

else i

else if

end

//Find class for classification

C=class to which the most u N are classified;

For example, investment there can be two groups as first is for customer with Income 3000/- to 10000/- and Occupation Employee with low risk level share purchasing and 10% amount of their income. Similarly second one is for customer Income >50000/- and Occupation Employee with low risk level share purchasing and 60% amount.

4. CONCLUSION

With the increase of economic globalization and evolution of information technology, financial data are being generated and accumulated at an unprecedented pace. As a result, there has been a critical need for automated approaches to effective and efficient utilization of massive amount of financial data to support companies and individuals in strategic planning and investment decision making. Data mining techniques have been used to uncover hidden patterns and predict future trends and behaviors in financial markets. Clustering is a tool for data analysis, which solves classification problems. Its objective is to distribute cases (people, objects, events etc.) into groups, so that the degree of association can be strong between members of the same cluster and weak between members of different clusters. In clustering, there is no pre classified data and no distinction between independent and dependent variables. Instead, clustering algorithms search for groups of records. The algorithms discover these similarities. This way each cluster describes, in terms of data collected, the class to which its members belong. Clustering is a discovery tool. It may re-veal associations and structure in data which, though not previously evident, nevertheless are sensible and useful once found. As part of a stock market analysis and prediction system consisting of an expert system and clustering of stock prices, data is needed. The competitive advantages achieved by data mining include increased revenue, reduced cost, and much improved marketplace responsiveness and awareness.

REFERENCES

1. Alex Berson and Stephen J. Smith, "Data Warehousing, Data Mining, And OLAP", MC Grauw-Hill, 1997.
2. Bigus and Joseph P, "Data Mining With Neural Networks", MC Graw-Hill, New York 1996.
3. Dasrathy B. V., Ed, "Nearest Neighbor Norms: NN Pattern Classification Techniques", IEEE, Computer Society Press, Calif. 1990.

4. Mrs. Keerti. S. Mahajan & R. V. Kulkarni. Application Of Datamining Tools For Stock Market,vol(4)1,19-27,IJCTA ,JAN-FEB 2013
5. Graig Silverstein, Sergey Brin and Rajeev Montwani, "Beyond Market Baskets: Generalizing Association Rules to Dependence Rules", Data Mining and Knowledge Discovery, Vol. 2, No. 1, Jan 1998, Kluwer Academic Publishers.
6. Huan Liu, Farhad Hussain, Chew Lim Tan and Manoranjan Dash, "Discretization: An Enabling Technique", Data Mining and Knowledge Discovery", vol. 6 No. 4, October 2002.
7. A. B. Devale and Dr. R. V. Kulkarni (IJDKP)" Aplications Of Data Mining Techniques In Life Insurance" Vol.2, No.4, July 2012..
8. Jiawei Han, Laks V. S. Lakshmanan and Raymond T.NG, "Constraint-Based Multidimensional Data Mining", IEEE, August 1999.
9. H. A. Sankpal And R. V. Kulkarni,"A Review Of Expert System For Investment Portfolio With Reference Of Stock Market In BSE", year:2014 vol:3 iss:5
10. Krzysztof J. Cios, Witold Pedrycz and Roman W. Surniarski, " Data Mining Methods for Knowledge Discovery", Kluwer Academic Publishers 1998 Second Printing 2000.
11. Ming-Syan chen, Jiawei Han and Philip S. Yu, "Data Mining: An Overview From a Database Perspective", IEEE Transactions on Knowledge and Data Engineering Vol. 8, No. 6, Dec. 1996

REQUEST FOR FEEDBACK

Dear Readers

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

I would like to request you to supply 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 you have 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, neither 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 are 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

