INTERNATIONAL JOURNAL OF RESEARCH IN COMPUTER APPLICATION & MANAGEMENT

A Monthly Double-Blind Peer Reviewed Refereed 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.
as well as inOpen J-Gage, India [link of the same is duly available at Inflibnet of University Grants Commission (U.G.C.)]
Registered & Listed at: Index Copernicus Publishers Panel, Poland
Circulated all over the world & Google has verified that scholars of more than 1500 Cities in 141 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

CONTENTS

CALL FOR MANUSCRIPTS

Weinvite unpublished novel, original, empirical and high quality research work pertaining to recent developments & practices in the area of Computer, Business, Finance, Marketing, Human Resource Management, General Management, Banking, Insurance, Corporate Governance and emerging paradigms in allied subjects like Accounting Education; 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; Monetary Policy; Portfolio & Security Analysis; Public Policy Economics; Real Estate; Regional Economics; Tax Accounting; Advertising & Promotion Management; Business Education; Management Information Systems (MIS); Business Law, Public Responsibility & Ethics; Communication; Direct Marketing; E-Commerce; Global Business; Health Care Administration; Labor 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; Public Administration; Purchasing/Materials Management; Retailing; Sales/Selling; Services; Small Business Entrepreneurship; Strategic Management Policy; Technology/Innovation; Tourism, Hospitality & Leisure; Transportation/Physical Distribution; Algorithms; Artificial Intelligence; Compilers & Translation; Computer Aided Design (CAD); Computer Aided Manufacturing; Computer Graphics; Computer Organization & Architecture; Database Structures & Systems; Digital Logic; Discrete Structures; Internet; Management Information Systems; Modeling & Simulation; Multimedia; Neural Systems/Neural Networks; Numerical Analysis/Scientific Computing; Object Oriented Programming; Operating Systems; Programming Languages; Robotics; Symbolic & Formal Logic and Web Design. The above mentioned tracks are only indicative, and not exhaustive.

Anybody can submit the soft copy of his/her manuscript **anytime** in M.S. Word format after preparing the same as per our submission guidelines duly available on our website under the heading guidelines for submission, at the email address: infoijrcm@gmail.com.

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 mention the 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, KurukshetraUniversity, 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

AN EMPIRICAL STUDY ON BREAST CANCER USING DATA MINING TECHNIQUES

GOMATHI.K
ASST. PROFESSOR
KARPAGAM UNIVERSITY
COIMBATORE

ABSTRACT

Data mining is taking out of hidden patterns from huge database. It is commonly used in marketing, surveillance, fraurd detection and scientific discovery. In data mining, machine learning is mainly focused as research which is automatically learnt to recognize complex patterns and make intelligent decisions based on data. Nowadays, Breast cancer occurs when a malignant (cancerous) tumor originates in the breast. As breast cancer tumors mature, they may metastasize (spread) to other parts of the body. This deals with the some of classification models to predict the causes of breast cancer using Naive bayes, Ada BoostM1 Meta Classifier, PART Rule Classifier, J48 Decision Tree Classifier and Random Forest Classifier.

KEYWORDS
Breast cancer, Decision Tree, Naive Bayes, PART Rule Classifier, Random Forest Tree Classifier.

INTRODUCTION

Cancer is a group of diseases that cause cells in the body to change and grow out of control. Most types of cancer cells eventually form a lump or mass called a tumor, and are named after the part of the body where the tumor originates.

Breast cancer begins in breast tissue, which is made up of glands for milk production, called lobules, and the ducts that connect the lobules to the nipple. The remainder of the breast is made up of fatty, connective, and lymphatic tissue.

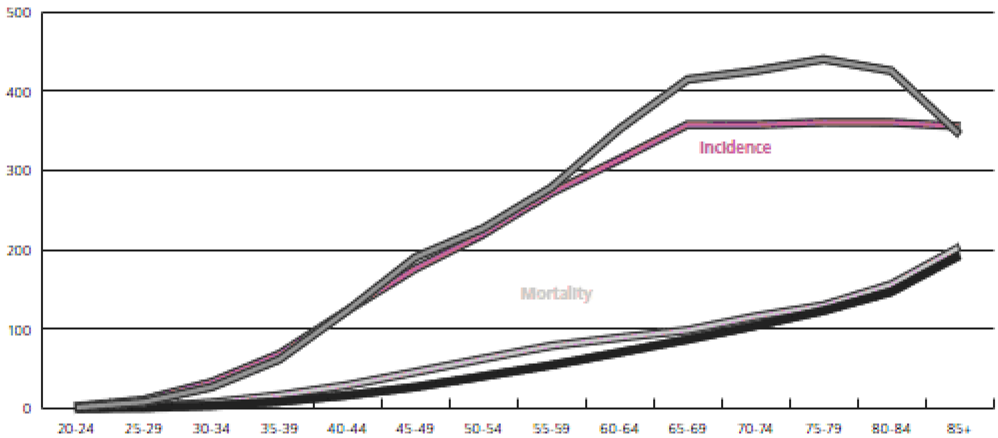
- Most masses are benign; that is, they are not cancerous, do not grow uncontrollably or spread, and are not life-threatening.
- Some breast cancers are called in situ because they are confined within the ducts (ductal carcinoma in situ or DCIS) or lobules (lobular carcinoma in situ or LCIS) where they originated. Many oncologists believe that LCIS (also known as lobular neoplasia) is not a true cancer, but an indicator of increased risk for developing invasive cancer in either breast.

The majority of in situ breast cancers are DCIS, which accounted for about 83% of in situ cases diagnosed during 2004-2008.

LCIS is much less common than DCIS, accounting for about 11% of female in situ breast cancers diagnosed during 2004-2008. · Other in situ breast cancers have characteristics of both ductal and lobular carcinomas or have unspecified origins. Most breast cancers are invasive, or infiltrating. These cancers started in the lobules or ducts of the breast but have broken through the duct or glandular walls to invade the surrounding tissue of the breast.

BREAST CANCER STATISTICS

FIGURE 1: FEMALE BRESAT CANCER INCIDENCE (2004-2008) AND MORTALITY (2003-2007) RATES



Breast cancer cases have doubled in India in the last two decades. The number of women estimated to be dying of breast cancer every year has also been steadily rising.

As against an estimated 48,170 women who died of breast cancer in 2007, the number breached the 50,000 mark in 2010. The figure for the year was put at 50,821.

Uttar Pradesh recorded the highest number of breast cancer deaths among states in 2010 - 8,882 followed by Maharashtra (5,064), Bihar (4,518), West Bengal (4,095), Andhra Pradesh (3,863), Madhya Pradesh (3,179) and Rajasthan (3,097).

Gujarat recorded 2,632 deaths, Kerala 1,618, Haryana 1,118 and Orissa 1,885. Delhi recorded an estimated 810 deaths due to breast cancer in 2010 compared to 779 in 2009 and 749 in 2008.

When it comes to states recording low breast cancer mortality rate, Lakshwadeep recorded the lowest with three deaths followed by Andaman and Nicobar Islands with 19 deaths. The north-eastern states also showed low levels of breast cancer deaths. Sikkim recorded 30 deaths, Mizoram 49 and Arunachal an estimated 63 deaths.

MoS health S Gandhiselvan said, "According to Indian Council of Medical Research, there is a significant increase in the incidence of breast cancers in various urban population based cancer registries between 1982 and 2005."

Globally, breast cancer is the most common female cancer accounting for an estimated 1.4 million cases each year, with more than half of the 400,000 breast cancer deaths occurring in low and middle income countries.

A landmark analysis of cancer cases in Delhi, Mumbai, Chennai and Bangalore between 1982 and 2005 (24 years) by ICMR had found that while cervical cancer cases -- earlier the most common -- dipped, in some cases by almost 50%, the incidence of breast cancer doubled.

While Bangalore saw breast cancer cases more than double since 1982 -- 15.8 in a population of one lakh in 1982 to 32.2 in 2005 -- Chennai recorded 33.5 new cases of breast cancer per one lakh women in 2005 against 18.4 in 1982.

Delhi recorded 24.8 new cases of breast cancer a year per 100,000 women which rose to 32.2 in 2005. Mumbai recorded 20.8 new cases of breast cancer per 100,000 population in 1982 which increased by almost 10% in 2005.

TABLE 1: AFTER THE 10 YEARS AGO

Age	InSitu Cases	Invasive Cases	Deaths
Under 40	1780	11330	1160
Under 50	14240	50430	5240
50-64	23360	81970	11620
65+	20050	98080	22660
Allages	**57650**	**230480**	**39520**

FIGURE2: ACCORDING TO THE 2011

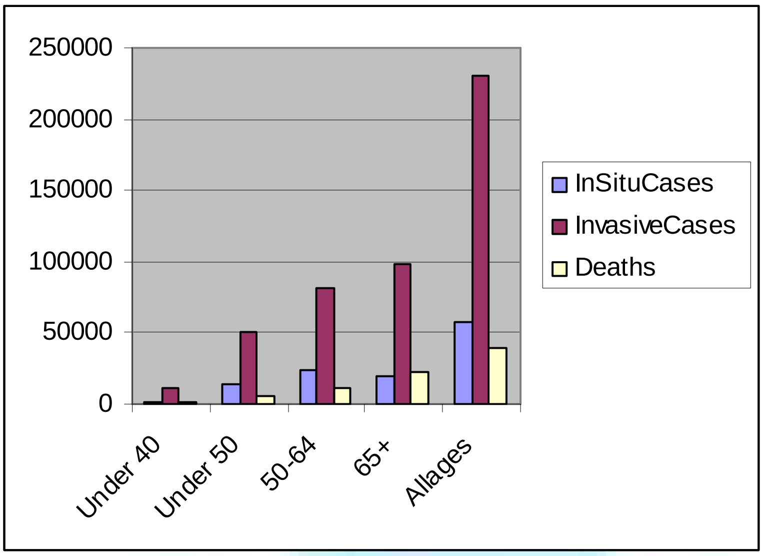


TABLE 2: ACCORDING TO FEMALE AGE

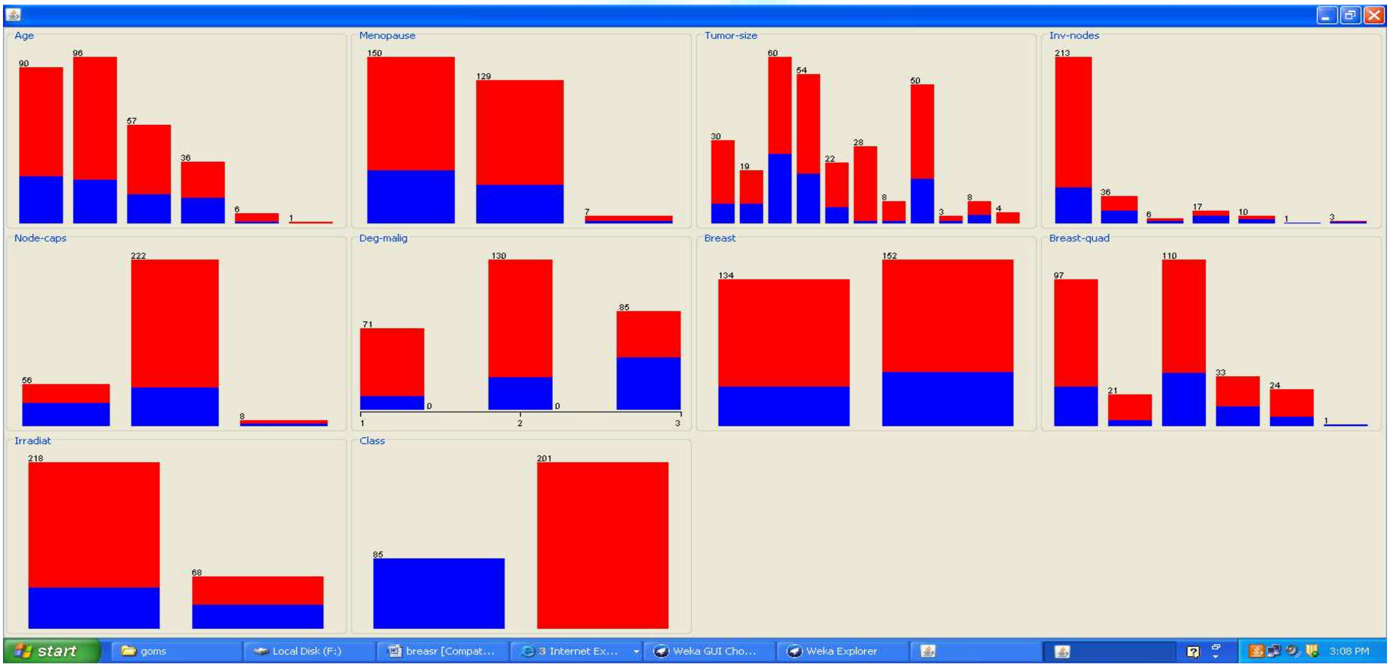
If Current Age is	Developing breastcancer in the next 10 years is:	Or 1 in
20	0.06%	1681
30	0.43%	232
40	1.45%	69
50	2.38%	42
60	3.45%	29
70	3.74%	27

Lifetime risk 12.15% 8

EXPERIMENTAL RESULTS

A major focus of machine learning [3, 8] research is to automatically learn to recognize complex patterns and make intelligent decisions based on data. Hence, machine learning is closely related to fields such as artificial intelligence, adaptive control, statistics, data mining, pattern recognition, probability theory and theoretical computer.

FIGURE 3: CLUSTERING OF BREST CANCER DATA



NAIVE BAYESIAN CLASSIFIER

A Naive Bayesian classifier [21] is a simple probabilistic classifier based on applying Bayesian theorem (from Bayesian statistics) with strong (naive) independence assumptions. By the use of Bayesian theorem we can write

$$p(C \mid F_1....F_n) = \frac{p(C)p(F_1.....F_n \mid C)}{p(F_1.....F_n)}$$

ADVANTAGES

- It is fast, highly scalable model building and scoring
- Scales linearly with the number of predictors and rows
- Build process for Naive Bayes is parallelized
- Induced classifiers are easy to interpret and robust to irrelevant attributes
- Uses evidence from many attributes, the Naive Bayes can be used for both binary and multiclass classification problems

TABLE 3: EVALUATION ON TRAINING SET

Correctly Classified Instances	73.7762 %
Incorrectly Classified Instances	26.2238 %
Kappa statistic	0.3338
Mean absolute error	0.3077
Root mean squared error	0.4315
Relative absolute error	73.5638 %
Root relative squared error	94.4037 %
Total Number of Instances	286

ACCURACY BY CLASS

TP Rate	FP Rate	Precision	Recall	F-Measure	ROCArea	Class
0.459	**0.144**	**0.574**	**0.459**	**0.51**	**0.759**	**recurrence-events**
0.856	**0.541**	**0.789**	**0.856**	**0.821**	**0.759**	**no-recurrence-events**
Weighted Avg. 0.738	**0.423**	**0.725**	**0.738**	**0.729**	**0.759**	

=== Confusion Matrix ===
a b <-- classified as
39 46 | a = recurrence-events
29 172 | b = no-recurrence-events

J48 DECISION TREE CLASSIFIER

J48 is a simple C4.5 decision tree, it creates a binary tree. C4.5 builds decision trees from a set of training data which is like an ID3, using the concept of information entropy [20].

ALGORITHM

- Check for base cases
- For each attribute „a" find the normalized information gain from splitting on „a"
- Let a_best be the attribute with the highest normalized information gain
- Create a decision node that splits on a_best
- Recourse on the sub lists obtained by splitting on a_best, and add those nodes as children of node

ADVANTAGES

- Gains a balance of flexibility and accuracy
- Limits the number of possible decision points
- It had a higher accuracy

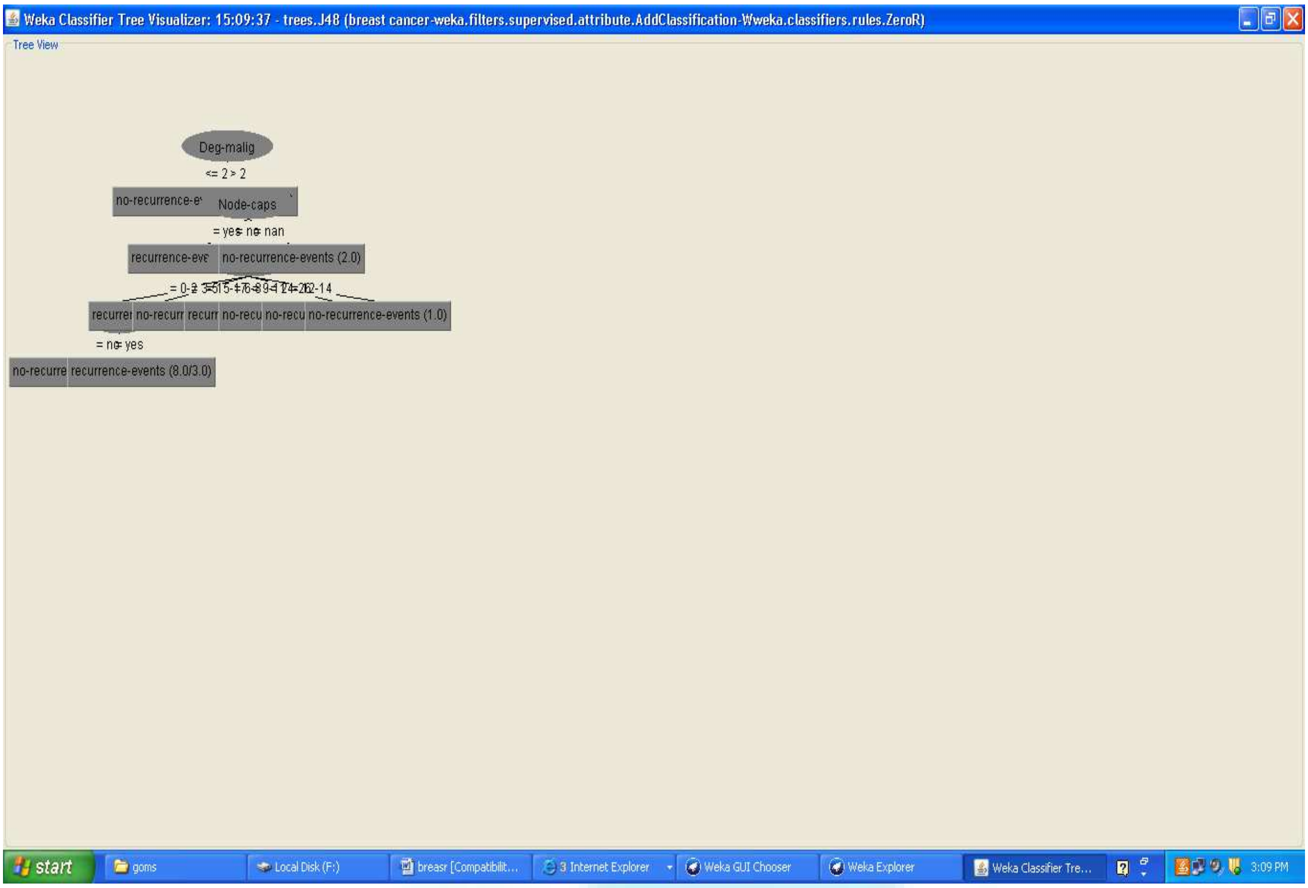
TABLE 4: EVALUATION ON TRAINING SET

Correctly Classified Instances	75.5245 %
Incorrectly Classified Instances	24.4755 %
Kappa statistic	0.3057
Mean absolute error	0.3528
Root mean squared error	0.4295
Relative absolute error	84.3212 %
Root relative squared error	93.9673 %
Total Number of Instances	286

ACCURACY BY CLASS

TP Rate	FP Rate	Precision	Recall	F-Measure	ROCArea	Class
0.318	0.06	0.692	0.318	0.435	0.64	recurrence-events
0.94	0.682	0.765	0.94	0.844	0.64	no-recurrence-events
Weighted Avg. 0.755	0.497	0.744	0.755	0.722	0.64	

FIGURE 4: J48 TREE CLASSIFIER



=== Confusion Matrix ===

a b <-- classified as

27 58 | a = recurrence-events

12 189 | b = no-recurrence-events

AdaBoostM1 Classifier

Adaptive Boosting [13] is a meta-algorithm in the sense that it improves or boosts an existing weak classifier. Given a weak classifier (error close to 0.5), AdaBoostM1 algorithm improves the performance of the classifier so that there are fewer classification errors.

ALGORITHM

- All instances are equally weighted
- A learning algorithm is applied
- The weight of incorrectly classified example is increased and correctly decreased
- The algorithm concentrates on incorrectly classified "hard" instances
- Some "had" instances become "harder" some "softer"
- A series of diverse experts are generated based on the reweighed data.

ADVANTAGES

- Simple and trained on whole (weighted) training data
- Over-fitting (small subsets of training data) protection
- Claim that boosting "never over-fits" could not be maintained.
- Complex resulting classifier can be determined reliably from limited amount of data

TABLE 5: EVALUATION ON TRAINING SET

Correctly Classified Instances	75.5245 %
Incorrectly Classified Instances	24.4755 %
Kappa statistic	0.3574
Mean absolute error	0.341
Root mean squared error	0.4188
Relative absolute error	81.5216 %
Root relative squared error	91.6302 %
Total Number of Instances	286

ACCURACY BY CLASS

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.435	0.109	0.627	0.435	0.514	0.751	recurrence-events
0.891	0.565	0.789	0.891	0.836	0.751	no-recurrence-events
Weighted Avg. 0.755	0.429	0.741	0.755	0.741	0.751	

=== Confusion Matrix ===

a b <-- classified as

37 48 | a = recurrence-events

22 179 | b = no-recurrence-events

PART (PARTIAL DECISION TREES)

Classifier PART is a rule based algorithm [12] and produces a set of if-then rules that can be used to classify data. It is a modification of C4.5 and RIPPER algorithms and draws strategies from both. PART adopts the divide-and-conquer strategy of RIPPER and combines it with the decision tree approach of C4.5.

PART generates a set of rules according to the divide-and conquer strategy, removes all instances from the training collection that are covered by this rule and proceeds recursively until no instance remains [5].

To generate a single rule, PART builds a partial decision tree for the current set of instances and chooses the leaf with the largest coverage as the new rule. It is different from C4.5 because the trees built for each rules are partial, based on the remaining set of examples and not complete as in case of C4.5.

ADVANTAGES

It is simpler and has been found to give sufficiently strong rules.

TABLE 6: EVALUATION ON TRAINING SET

Correctly Classified Instances	80.7692 %
Incorrectly Classified Instances	19.2308 %
Kappa statistic	0.4694
Mean absolute error	0.2967
Root mean squared error	0.3851
Relative absolute error	70.919 %
Root relative squared error	84.2706 %
Total Number of Instances	286

ACCURACY BY CLASS

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.447	0.04	0.826	0.447	0.58	0.77	recurrence-events
0.96	0.553	0.804	0.96	0.875	0.77	no-recurrence-events
Weighted Avg. 0.808	0.4	0.811	0.808	0.788	0.77	

=== Confusion Matrix ===

a b <-- classified as

38 47 | a = recurrence-events

8 193 | b = no-recurrence-events

RANDOM FOREST TREE CLASSIFIER

A random forest [14] consisting of a collection of tree structured classifiers (h(x,_k), k = 1, . .) where the _k are independent identically distributed random vectors and each tree casts a unit vote for the most popular class at input x.

ALGORITHM

- Choose T number of trees to grow
- Choose m number of variables used to split each node. m<<M, where M is the number of input variables, m is hold constant while growing the forest
- Grow T trees. When growing each tree do
- Construct a bootstrap sample of size n sampled from Sn with the replacement and grow a tree from this bootstrap sample
- When growing a tree at each node select m variables at random and use them to find the best split
- Grow the tree to a maximal extent and there is no pruning
- To classify point X collect votes from every tree in the forest and then use majority voting to decide on the class label

ADVANTAGES

- It is unexcelled in accuracy among current algorithms and it runs well on large data bases.
- It can handle thousands of input variables without variable deletion and also the learning is so fast.
- It has an effective method for estimating missing data and maintains accuracy.
- The new generated forests can be saved for future use on other data.
- It computes proximities between pairs of cases that can be used in clustering, locating outliers or give interesting views of the data.

TABLE 7: EVALUATION ON TRAINING SET

Correctly Classified Instances	**97.9021 %**
Incorrectly Classified Instances	**2.0979 %**
Kappa statistic	**0.9508**
Mean absolute error	**0.0221**
Root mean squared error	**0.1052**
Relative absolute error	**5.2937 %**
Root relative squared error	**23.0237 %**
Total Number of Instances	**286**

ACCURACY BY CLASS

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
1	0.03	0.934	1	0.966	0.999	recurrence-events
0.97	0	1	0.97	0.985	0.999	no-recurrence-events
Weighted Avg. 0.979	0.009	0.98	0.979	0.979	0.999	

OUTLIER PREDICTION

Time taken to build model: 0.06 seconds

=== Predictions on test data ===

inst#,actual,predicted,error,prediction

1,1:recurrence-events,2:no-recurrence-events,+,0.808

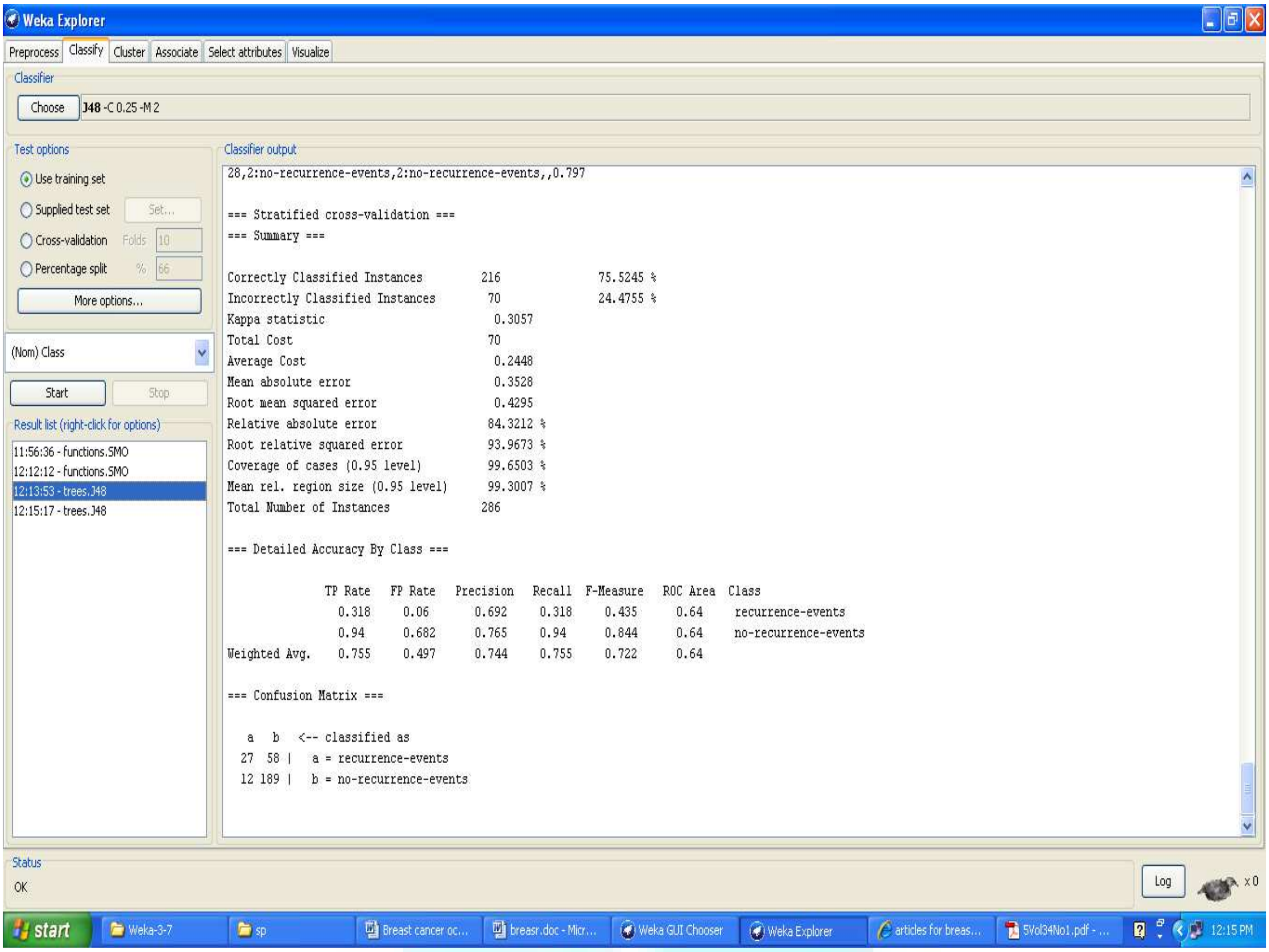
2,1:recurrence-events,1:recurrence-events,,0.759

3,1:recurrence-events,2:no-recurrence-events,+,0.808

4,1:recurrence-events,2:no-recurrence-events,+,0.808

5,1:recurrence-events,2:no-recurrence-events,+,0.727

FIGURE 5: EVALUATION ON TEST SPLIT



CONCLUSION

This paper provides causes of breast cancer diagnosis and statistical analysis of breast cancer. Problems and explores that data mining techniques using breast cancer data set. For this purpose the Experimenter in Weka-3.6.6 will be used.

REFERENCES

1. American Cancer Society. Breast Cancer Facts & Figures 2005-2006. Atlanta: American Cancer Society, Inc. (http://www.cancer.org/).
2. Benjamin F. Hankey, et. al. The Surveillance, Epidemiology, and End Results Program: A National Resource. Cancer Epidemiology Biomarkers & Prevention 1999; 8:1117-1121.
3. Cios KJ, Moore GW. Uniqueness of medical data mining. Artificial Intelligence in Medicine 2002; 26:1-24.
4. Cox DR. Analysis of survival data. London: Chapman & Hall; 1984.
5. Delen D, Walker G, Kadam A. Predicting breast cancer survivability: a comparison of three data mining methods. Artificial Intelligence in Medicine. 2005 Jun; 34(2):113-27.
6. Houston, Andrea L. and Chen, et. al.. Medical Data Mining on the Internet: Research on a Cancer Information System. Artificial Intelligence Review 1999; 13:437-466.
7. Ian H. Witten and Eibe Frank. Data Mining: Practical machine learning tools and techniques, 2nd Edition. San Fransisco:Morgan Kaufmann; 2005.
8. Lundin M, Lundin J, Burke HB, Toikkanen S, Pylkkanen L, Joensuu H. Artificial neural networks applied to survival prediction in breast cancer. Oncology 1999; 57:281-6.
9. Surveillance, Epidemiology, and End Results (SEER) Program (www.seer.cancer.gov) Public-Use Data (1973-2002), National Cancer Institute, DCCPS, Surveillance Research Program, Cancer Statistics Branch, released April 2005, based on the November 2004 submission.
10. Weka: Data Mining Software in Java, http://www.cs.waikato.ac.nz/ml/weka/
11. Zhou ZH, Jiang Y. Medical diagnosis with C4.5 Rule preceded by artificial neural network ensemble. IEEE Trans Inf Technol Biomed. 2003 Mar; 7(1):37-42.

REQUEST FOR FEEDBACK

Dear Readers

At the very outset, International Journal of Research in Computer Application and 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

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