

Effective Utilization of Library Books among University Students Using Data Mining Tool

Babu Santhalingam, J. Devi, K. Revathi

Abstract: Data mining is the process of discovering previously unknown and potentially interesting patterns in large datasets. University Library is an excellent example of knowledge creation, application and utilization. "A library is a growing organism". This study examines the effective use of library books by the students of SCSVMV University, Kanchipuram. The analysis had been carried out by using a library books dataset and using the J48 decision tree algorithm implemented in Weka. WEKA is a collection of machine learning algorithms for data mining tasks. The objectives of this study are to provide an approximate estimation of existing computerized records of library books population and to evaluate how books are used. The results revealed that maximum use of the library books were made and which particular department books are highly borrowed from the library. A Sample of 460 volumes of 2010 year borrowed books is used for this study.

Keywords: Data Mining, University Library, Effective use of library books, J48 decision tree algorithm, WEKA, Utilization of library.

◆

1. INTRODUCTION

Education plays a vital role in the development of a country. It is one of the fundamental rights of every individual. It not only enhances knowledge but also develops overall personality of a person where library has become "a place entrusted with the acquisition, organization, preservation, storage, retrieval and dissemination of information in whatever format it might appear".

Dr S.R. Ranganathan defined the five laws of library science, focusing especially on the fifth law, which said, "A library is a growing organism". There couldn't have been a truer definition of a library, considering the pace at which libraries have grown as information storehouses, around the world, from being paper dependent to the present digital-oriented. The vision is to develop a 21st Century world-class knowledge resource centre that can play a pivotal role in spreading library awareness among masses through a focused approach.

The view of Ranganathan is well appreciated pointing out the relationship between books and the readers. Librarians therefore, should make sure that library collections are readily made available to readers and evaluated from time to time.

S.R. Ranganathan's five library science can be summed up as:

1. Books are for use
2. Every reader his book
3. Every book its read
4. Save the time of the reader
5. A library is a growing organization.

Globally, any university is faced with great challenges in developing an efficient university library. The importance of adequate library services for effective performance of researchers, lecturers and students are evident and must be facilitated by the concerned institution.

Students cannot acquire knowledge only through text books or classroom lectures. They are referred to other books also. The library is of great help in the fulfillment of their wishes, ambitions and inclinations, as it provides ample opportunities for acquiring knowledge. On the other hand, it is much more convenient for the students to collect required material from the library. It not only saves their time but also cost of books and their energy.

WEKA (Waikato Environment for Knowledge Analysis) was developed at the University of Waikato in New Zealand. WEKA is a collection of machine learning algorithms for data mining tasks. Weka's native storage method is ARFF (.arff) format. So a conversion has been performed to make the examination data available for analysis through Weka. It can be run on **Windows, Linux and Mac**. Extensive, automatically generated documentation of the WEKA source is available to guide the user through interactions with the system. Earlier, Unix-based versions of the WEKA system are further described in Holmes et al, 1994; Garner et al, 1995. The current version of WEKA is implemented as a suite of Java class libraries. It is freely available on the World-Wide Web www.cs.waikato.ac.nz/ml/weka. The software accompanies a new text on data mining (Witten and Frank, 1999) which documents and fully explains all the data mining algorithms incorporated in WEKA. Application programs written using the WEKA class libraries can be run on any computer with a WWW browser. WEKA uses the J48 algorithm, which is Weka's implementation of C4.5 Decision tree algorithm.

The **J48 algorithm** is actually a slight improved to the latest version of C4.5 algorithm, created by J. Ross Quinlan, using an inductive top down method for decision trees. They are built by testing every node of the tree starting with the first node, for every record. Each node represents the name of an attribute. The algorithm tries to introduce the record tested into an existing class, considering the similar aspects, evaluating the attribute that corresponds to

the current node. It was the last public version of this family of algorithms before the commercial implementation C5.0 was release. J48 algorithm is an implementation of the C4.5 decision tree learner. This implementation produces decision tree models. The algorithm uses the greedy technique to induce decision trees for classification. A decision tree model is built by analyzing training data and the model is used to classify unseen data. J48 generates decision trees, the nodes of which evaluate the existence or significance of individual features, e.g., heart rate. The decision trees are constructed in a top-down fashion by choosing the most appropriate attribute each time. This algorithm is chosen to compare the accuracy rate with other algorithms.

1.1 Role of University Library

The library is regarded as the 'nerve centre of knowledge', the centre of intellectual life and the heart and Soul of the academic institution. This means that discoveries are actually made in the library and subsequently tested in the laboratory. It occupies an important place in the modern education system and maintains the expensive educational resources of the academic institutions. It is the responsibility of the staff of university libraries to provide right information at the right time to right user to save the time of the user. The libraries are primarily responsible for the selection and collection of material appropriate for libraries, preservation and organization of the collection and dissemination of the material or the information, which it contains.

Libraries as centre of learning are playing an important role in sustaining and satisfying the information requirements of parent institutions. For the efficient, effective and scientific development of information resources and services, the libraries need to be designed and developed systematically based on the inputs from studies on existing resources and services, and the studies on users of university libraries.

From one step to another

Student's use of the university's library can be presented as a staircase that consists of five steps. Each step forms the basis to the next – a new step can be reached only if the previous ones are in good condition. The steps also emphasis the meaning of socializing to the academic community as well as the meaning of the basic information skills and how they are learned.

The available recourses can only be used if the student is encouraged and tutored in the use of those resources throughout their academic Studies. Philosophies based solely on the material recourses are thus superficial and too narrow. The foundation for the effective use of the library and its services is dependent on the student's Personal abilities – the courage to use the available material resources – as well as possession of good information skills. The next step that encourages and helps the students are other individuals – other students and the university

(library) staff. The final step consists of the library's material resources that can be used only if the previous steps are in good condition.

2. PURPOSE OF THIS STUDY

The study provide an approximate estimation of the existing computerized records of library books and to evaluate how books are used that is somehow related to the ways of students used the library. Focusing on this point, the current study attempts to understand effective use of university library books among students and clarify maximum use of library books and which particular department books are highly borrowed from the library. The analysis had been carried out by using a library books dataset and using the Data Mining's J48 decision tree algorithm implemented in Weka.

Student expectations and demands of the library are reflected in their library usage, and if these are related to certain educational outcomes, the results of this study may be useful to promote certain types of library use that are more efficient in attaining desired outcomes. Also, if it is possible to predict the level of outcomes affected by the identified types of library books use, this process can be examined further to develop into a method of outcomes assessment.

3. DATA MINING IN THE EXPERT SYSTEM

Data mining (sometimes called data or knowledge discovery) is the process of analyzing data from different perspectives and summarizing it into useful information - information that can be used to increase revenue, cuts costs, or both. Data mining software is one of a number of analytical tools for analyzing data.

It allows users to analyze data from many different dimensions or angles, categorize it, and summarize the relationships identified. Technically, data mining is the process of finding correlations or patterns among dozens of fields in large relational databases. In our paper, one of the data mining methods of J48 algorithm used for finding maximum usage of library books.

4. METHODOLOGY

The secondary sources of information were the computerized records of library books details from the year of 2010 sessions to determine maximum use of the library books were made and which particular department books are highly used by the students.

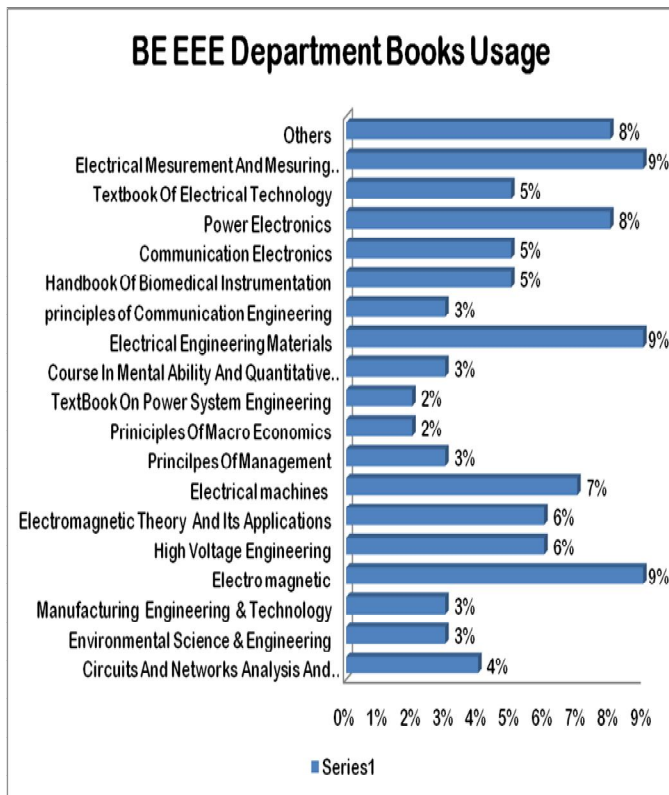
A total of 460 volumes of 2010 year borrowed books are used for this study, in this 150 data's are taken from BE (EEE Department), 110 data's from BE (ECE Department), 105 data's from BE (MECH Department) and finally 95 data's from BE (CSE) department. The data were analyzed using the J48 decision tree algorithm implemented in Weka.

5. RESULTS AND DISCUSSIONS

5.1 BE EEE Department Books Usage Detail

The effectiveness of a library as an instrument of learning is determined by the success with which it is able to provide the user with the information he or she seeks. The library can fulfill its function best by pursuing a policy of constant self-evaluation in order to be alert to the changing needs of its users. The utility of a library depends upon its proper organization which includes the distribution of books, their arrangement, the situation of the library, etc.

In this section, we used 150 volumes of 2010 year BE EEE department borrowed books and tried to find out maximum number of borrowed books among students using J48 Decision Tree algorithm in Data Mining WEKA Tool.



Here, we drawn a bar chart for BE EEE department books usage details and I found that the BE EEE department students are maximum borrowed the first highest level of these three books: Electromagnetic (9%), Electrical Engineering Materials (9%) & Electrical Measurement and Measuring Instrument (9%).

The second levels of maximum borrowed books are: 8% of power Electronics, 7% of Electrical machines, 6% of Electromagnetic theory and its Applications & 6% of High Voltage Engineering.

Then, the third level of maximum borrowed books are: 5% of Hand Book of Biomedical Instrumentation, 5% of Communication Electronics, Textbook of Electrical Technology, 4% of Circuits and Networks analysis and synthesis, 3% of Environmental Science & Engineering, 3% of Manufacturing Engineering & Technology, 3% of Principle of Management, 3% of Course In Mental Ability And Quantitative Aptitude For Competitive Examinations, 2% of Principles of Macro Economics & 2% Of Textbook On Power System Engineering. Finally, the BE EEE department students are rarely borrowed 8% of some other books.

J48 Pruned Tree For BE EEE Department

J48 is a module for generating a pruned or unpruned C4.5 decision tree. When we applied J48 onto refreshed data, we got the results shown as below.

```

=== Classifier model (full training set) ===

J48 pruned tree
-----
: NO (150.0/9.0)

Number of Leaves : 1

Size of the tree : 1

Time taken to build model: 0.03 seconds

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances    141    94%
Incorrectly Classified Instances    9    6%

Kappa statistic    0
Mean absolute error    0.0233
Root mean squared error    0.1081
Relative absolute error    68.3028 %
Root relative squared error    98.4705 %
Total Number of Instances    150

=== Detailed Accuracy by Class ===
TP Rate FP Rate Precision    Recall F-Measure ROC Area Class
1    1    0.94    1    0.969    0.447    No
0    0    0    0    0    0.047    Handbook Electronics
0    0    0    0    0    0.047    Electromagnetic
0    0    0    0    0    0.047    DS using C & C++
0    0    0    0    0    0.047    Ancient India
0    0    0    0    0    0.047    Training Emerging
0    0    0    0    0    0.047    World Power
0    0    0    0    0    0.047    Empirical Method
0    0    0    0    0    0.047    CA
0    0    0    0    0    0.047    Digit Signal Process

Weighted Avg.    0.94    0.94    0.884    0.94    0.911    0.423

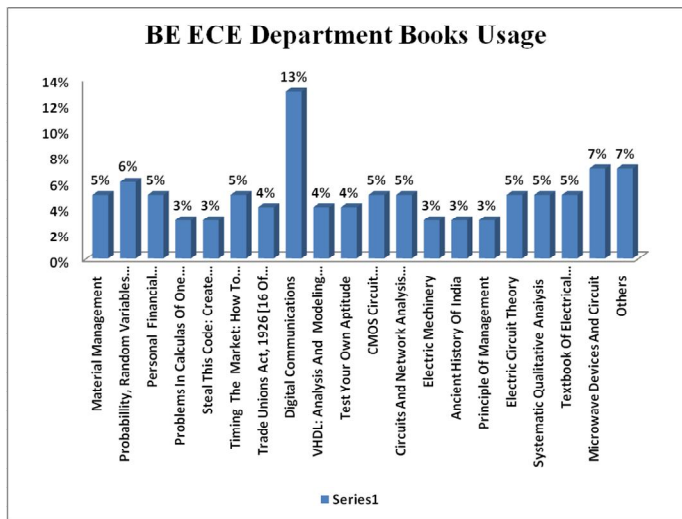
=== Confusion Matrix ===

 a b c d e f g h i j <-- classified as
141 0 0 0 0 0 0 0 0 0 0 | a = No
 1 0 0 0 0 0 0 0 0 0 0 | b = Handbook Electronics
 1 0 0 0 0 0 0 0 0 0 0 | c = Electromagnetic
 1 0 0 0 0 0 0 0 0 0 0 | d = DS using C & C++
 1 0 0 0 0 0 0 0 0 0 0 | e = Ancient India
 1 0 0 0 0 0 0 0 0 0 0 | f = Training Emerging
 1 0 0 0 0 0 0 0 0 0 0 | g = World Power
 1 0 0 0 0 0 0 0 0 0 0 | h = Empirical Method
 1 0 0 0 0 0 0 0 0 0 0 | i = CA
 1 0 0 0 0 0 0 0 0 0 0 | j = Digit Signal Process
    
```

5.2 BE ECE Department Books Usage Detail

The books in a library should be properly arranged and classified as they have as great an importance as the library itself. The importance does not consist in storing a large number of books as in having suitable books in a proper order. In this connection, attention should be directed towards the utility of the books. These books should be capacity and ability of the students so that they may read and understand themselves. When the student is able to understand a book himself he feels encouraged to read other books as well.

In this section, we used 110 volumes of 2010 year BE ECE department borrowed books and tried to find out the maximum number of borrowed books among students using J48 Decision Tree algorithm in Data Mining WEKA Tool.



Here, we drawn a column chart for BE ECE department books usage details and found that the first top levels of 13% of Digital Communications books are maximum borrowed by the BE ECE department students.

The second levels of maximum borrowed books are: 7% of Microwave Devices and Circuit, 6% Probability Random Variables and Random Signals Principles, 5% of Material Management, 5% of Personal Financial Management: Study Material-Vol.2, 5% of Timing the Market: How To Profit In Bull And Bear Markets With Technical Analysis, 5% of CMOS Circuit Design, Layout, and Simulation, 5% of Circuits and Network Analysis and Synthesis, 5% of Electric Circuit Theory, 5% of Systematic Qualitative Analysis & 5% of Textbook Of Electrical Technology - Vol.2: AC And DC Machines.

Then, the third levels of maximum borrowed books are: 4% of Trade Unions Act, 1926 [16 of 1926], 4% of VHDL: Analysis and Modeling of Digital Systems, 4% of Test Your Own Aptitude, 3% of Problems In Calculus Of One Variable With Elements Of Theory, 3% of Steal This Code: Create Reusable Software Components for Windows 95 And Windows NT, 3% of Electric Machinery,

3% of Ancient History of India & 3% of Principle of Management. Finally, the BE ECE department students are rarely borrowed 7% of some other books.

J48 Pruned Tree For BE ECE Department

J48 is a module for generating a pruned or unpruned C4.5 decision tree. When we applied J48 onto refreshed data, we got the results shown as below.

```

=== Classifier model (full training set) ===
J48 pruned tree
: NO (110.0/7.0)
Number of Leaves : 1
Size of the tree : 1
Time taken to build model: 0.02 seconds

=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances      103      93.6364 %
Incorrectly Classified Instances     7        6.3636 %
Kappa statistic                      0
Mean absolute error                  0.0309
Root mean squared error              0.1246
Relative absolute error              68.7136 %
Root relative squared error          98.3986 %
Total Number of Instances           110

=== Detailed Accuracy By Class ===
TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class
1        1        0.936     1        0.967     0.34      No
0        0        0         0         0         0.046     Flash memory
0        0        0         0         0         0.046     Control Sys
0        0        0         0         0         0.046     Data communication
0        0        0         0         0         0.046     Internet road rap
0        0        0         0         0         0.046     MSDOS
0        0        0         0         0         0.046     Psychological life
0        0        0         0         0         0.046     Realities

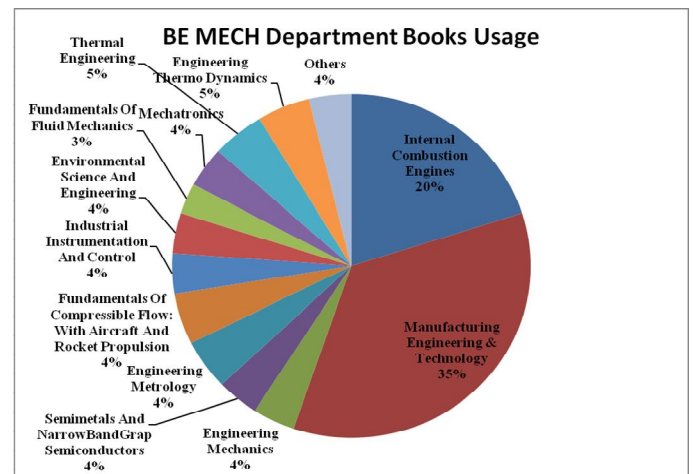
Weighted Avg. : 0.936  0.936  0.877  0.936  0.906  0.321

=== Confusion Matrix ===
 a  b  c  d  e  f  g  h  <-- classified as
1 0 0 0 0 0 0 0 1 a = No
1 0 0 0 0 0 0 0 1 b = Flash memory
1 0 0 0 0 0 0 0 1 c = Control Sys
1 0 0 0 0 0 0 0 1 d = Data communication
1 0 0 0 0 0 0 0 1 e = Internet road rap
1 0 0 0 0 0 0 0 1 f = MSDOS
1 0 0 0 0 0 0 0 1 g = Psychological Life
1 0 0 0 0 0 0 0 1 h = Realities
    
```

5.3 BE MECH Department Books Usage Detail

In this section, we used 105 volumes of 2010 year BE MECH department borrowed books and tried to find out the maximum number of borrowed books among students using J48 Decision Tree algorithm in Data Mining WEKA Tool.

Here, we had drawn a pie chart for BE MECH department books usage details.



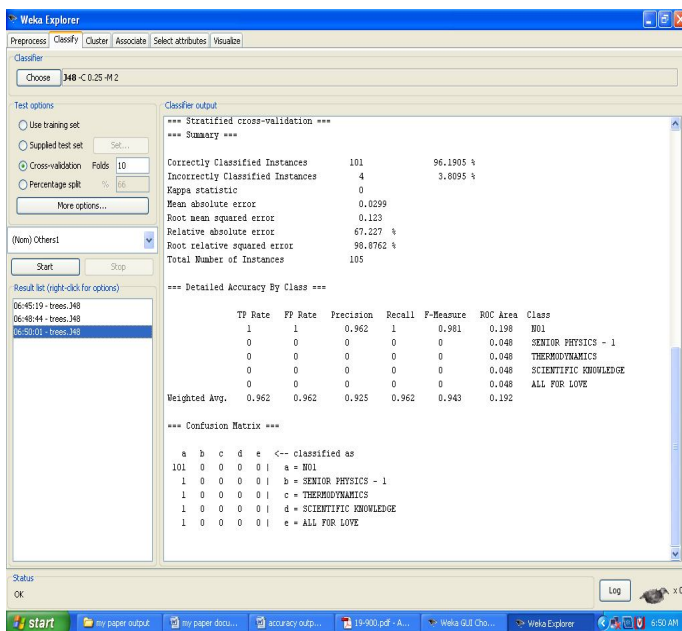
The most important point in this section is compared to all other above and below said department details the BE MECH department students are highly utilized the library books.

Because we found that the first highest levels of 35% of Manufacturing Engineering and Technology books are maximum borrowed by the BE MECH department students. So the highest amounts of library books are borrowed in this department.

Then, the second levels of maximum borrowed books are: The 20% of Internal Combustion Engines. The third level of maximum borrowed books are: 5% of Thermal Engineering, 5% of Engineering Thermo Dynamics, 4% of Engineering Mechanics, 4% of Semimetals and Narrow BandGap Semiconductors, 4% of Engineering Metrology, 4% of Fundamentals of Compressible Flow: With Aircraft and Rocket Propulsion, 4% of Industrial Instrumentation and Control, 4% of Environmental Science and Engineering, 4% of Mechatronics & 4% of Fundamentals of Fluid Mechanics. Finally, the BE MECH department students are rarely borrowed 4% of some other books.

J48 Pruned Tree For BE MECH Department

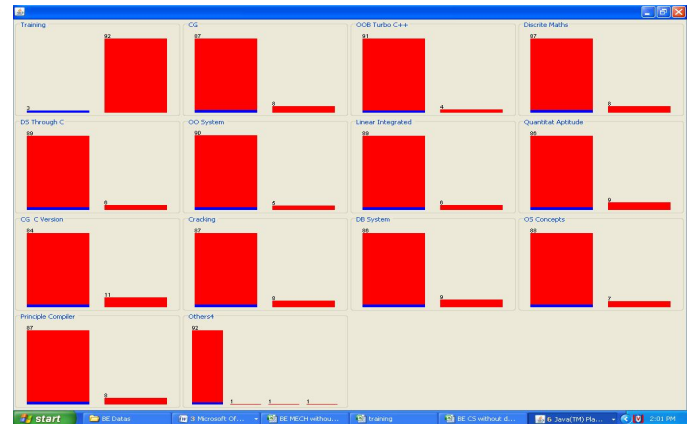
J48 is a module for generating a pruned or unpruned C4.5 decision tree. When we applied J48 onto refreshed data, we got the results shown as below.



5.4 BE CSE Department Books Usage Detail

It is become very clear that, "the bulk of the available information that might lead to knowledge acquisition is still paper based that is in book form Apeji (1991) stated further that, "the world without books would not only be a bore but would be completely devoid of any form of development". Thus, Okwilagwe (2001) has been prompted to declare that, "the book medium carries enormous information for development liberty and democracy".

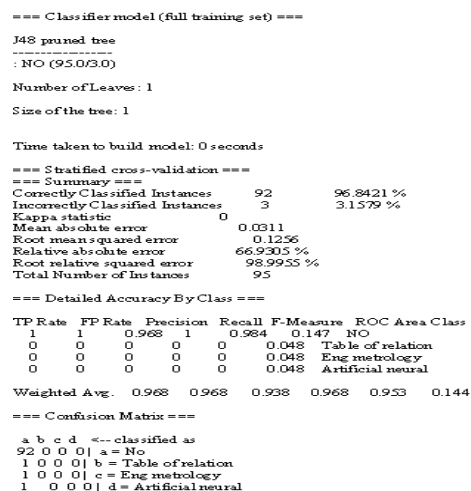
In this section, we used 95 volumes of 2010 year BE CSE department borrowed books and tried to find out the maximum borrowed books among the students using J48 Decision Tree algorithm in Data Mining WEKA Tool. Here, we attached a WEKA tool chart for BE CSE department books usage details and found that the first highest level of maximum borrowed books are: 11% of Computer Graphics C Version, 9% of Quantitative aptitude for competitive Examinations, 9% of Database System Concepts.



Then, the second level of maximum borrowed books are: 8% of computer Graphics, 8% of Mathematics, 8% of cracking the Gre, 8% of Principle of Compiler Design, 7% of Operating System Concepts, 6% of Data Structure through C & 6% of Linear Integrated Circuits. The third levels of maximum borrowed books are: 5% of Object Oriented System Development, 4% of Object Oriented Programming in Turbo C++ & 3% of Training and Development: Text, Research and cases. Finally, the BE CSE department students are rarely used 3% of some other books.

J48 Pruned Tree For BE CSE Department

J48 is a module for generating a pruned or unpruned C4.5 decision tree. When we applied J48 onto refreshed data, we got the results shown as below.



6. CONCLUSION

The results of this study give a remarkable revelation. It is to the young people of today that libraries are of the greatest benefit. The general conclusion to emerge from this study was that the students are maximally utilized the library. There was a progressive increase in the number of books borrowed and consulted from the year of 2010.

This study gives an insight into the University library Books usage among students using J48 Algorithm in Data Mining WEKA Tool and we found that 35% of Manufacturing Engineering and technology books are maximally borrowed by the BE MECH department students when compared with above said other engineering departments. The progressive increase in the utilization of library books is an indication that the library is meeting its primary role of supporting the objective of its parent institution. The library has been seen as a storehouse of knowledge. In future if it is possible to predict the level of outcomes affected by the identified types of library books use, this process can be examined further to develop into a method of outcomes assessment.

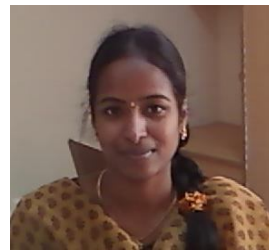
REFERENCES

- [1] Fowowe, S. O. "Students' Use of an Academic Library: A Survey at the University of Ilorin Libraries". Nigerian Library and Information Science Review, 7(1): 47-56.W.-K. Chen, Linear Networks and Systems. Belmont, Calif.: Wadsworth, pp. 123-135, 1993. (Book style)
- [2] Akin-Ojo, J. (1994). "Library orientation and effective use of the Library: a case study of the University of Lagos". Lagos Librarian, Vol.15: 11-17.
- [3] Whitmaire, E. (2002). "Academic library performance measures and undergraduates' library use and educational outcomes". Library & Information Science Research, 24, 107-28.
- [4] Sokoyz.A (2002). "Managing the library for user's convenience". Journal of Issues in Special Education 4(1) Pp 130 – 131.
- [5] Osagie, J.N. (2003). "User education in tertiary institution in Nigeria". Lagos Journal of Library and Information Science, No. 1:32-36.
- [6] Akande, Samson O. (2003). "Patterns of first-year students use of a university library": A Survey of University of Ibadan. Lagos Journal of Library and Information Science, 2(1); 22-26.
- [7] Ajayi N. A. & Adetayo J.O. "Utilization of Library Books to Enhance Academic Excellence in Nigeria Tertiary Institution: A Case Study of Hezekiah Oluwasanmi Library", O. A. U., Ile-ife, 10(2): 119-122 (2005).
- [8] Toda, A., Nagata, H., & Kytomaki, P. (2007). "Student's library use and their learning outcomes: A study on outcomes assessment in college and university library". Journal of the Japan Society of Library and Information Science, 53(1), 17-34.
- [9] Joan Giesecke. (2010). "Academic Library Futures: University of Nebraska-Lincoln".
- [10] <http://www.cs.waikato.ac.nz/ml/weka>, accessed 06/05/21.
- [11] http://grb.mnsu.edu/grbts/doc/manual/J48_Decision_Trees.html, accessed 06/06/12.

AUTHORS PROFILE



S. Babu MCA, Mphil working as an Assistant Professor in the Department of Computer Science and Applications in Sri Chandrasekharendra Saraswathi Viswa Mahavidyalaya, Enathur, Kanchipuram, Tamil Nadu. He has published more than 5 Papers in National, International journals and conferences. His research interest lies in the area of Data Mining, Distributed Computing and Database Management System. [mailto: babulingaa@gmail.com](mailto:babulingaa@gmail.com)



J. Devi Mphil, Research Scholar in the Department of Computer Science & Applications in SCSVMV University, Enathur, Kanchipuram. She received the degree in Master of Computer Applications from SCSVMV University in 2010. She has participated in many National and International Conferences and also published papers. Her interested area in research is Data Mining. [mailto: devi.gj88@gmail.com](mailto:devi.gj88@gmail.com)



K. Revathi Mphil, Research Scholar in the Department of Computer Science & Applications in SCSVMV University, Enathur, Kanchipuram. She received the degree in Master of Computer Applications from SCSVMV University in 2010. She has participated in many National and International Conferences and also published papers. Her interested area in research is Data Mining. [mailto: revathi.gk88@gmail.com](mailto:revathi.gk88@gmail.com)