A Review Approach on various form of Apriori with Association Rule Mining

Ms. Pooja Agrawal[1], Mr. Suresh kashyap[2], Mr.Vikas Chandra Pandey[3,] Mr. Suraj Prasad Keshri[4]

Research Scholar (Ph.D.), Dr.C.V.RamanUniversity, Kargi Road Kota,Bilaspur,India[1]
Research Scholar (M.Tech.), Dr.C.V.RamanUniversity, Kargi Road Kota,Bilaspur,India[2]
Research Scholar (Ph.D.), Dr.C.V.RamanUniversity, Kargi Road Kota,Bilaspur,India[3]
Research Scholar (M.Tech.), Dr.C.V.RamanUniversity, Kargi Road Kota,Bilaspur,India[4]

Abstract: - Data mining is a computerized technology that uses complicated algorithms to find relationships in large data bases Extensive growth of data gives the motivation to find meaningful patterns among the huge data. Sequential pattern provides us interesting relationships between different items in sequential database. Association Rules Mining (ARM) is a function of DM research domain and arise many researchers interest to design a high efficient algorithm to mine association rules from transaction database. Association Rule Mining plays a important role in the process of mining data for frequent pattern matching. It is a universal technique which uses to refine the mining techniques. In computer science and data mining, Apriori is a classic algorithm for learning association rules Apriori algorithm has been vital algorithm in association rule mining. . Apriori algorithm - a realization of frequent pattern matching based on support and confidence measures produced excellent results in various fields. Main idea of this algorithm is to find useful patterns between different set of data. It is a simple algorithm yet having many drawbacks. Many researches have been done for the improvement of this algorithm. This paper shows a complete survey on few good improved approaches of Apriori algorithm. This will be really very helpful for the upcoming researchers to find some new ideas from these approaches. The paper below summarizes the basic methodology of association rules along with the mining association algorithms. The algorithms include the most basic Apriori algorithm along with other algorithms such as AprioriTid, AprioriHybrid.

Keywords: Data Mining,Association Rule Mining, Apriori, Support, Confidence.

*****

I.INTRODUCTION

Due to growth of the data volume in the last decade, a set of different techniques for deletion of repetitive data and conversion of data to more usable forms has been proposed under the name of Data Mining. There exist a set of different techniques concerning the Data Mining, Some of which are decision trees, associative rules and Data Clustering. There are two categories of data mining namely, Descriptive data mining and Predictive data mining.

Descriptive data mining is used for carrying out summarizations or generalizations. Wherein, for finding out the inference or predictions, Predictive data mining is used.

Basic Data Mining Tasks

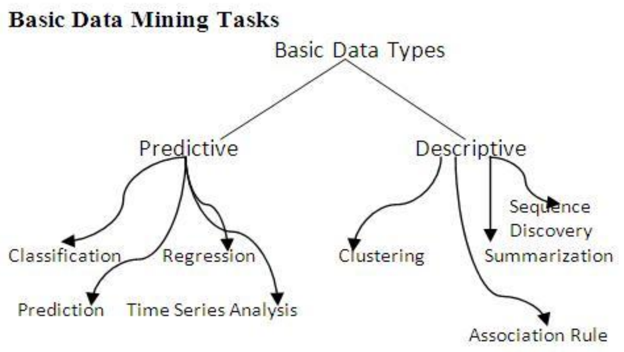

Table below show the above mentioned data mining techniques:

Techniques	Description	Example
Classification	It is often referred as "supervised learning". It has a predefined set of groups or models based on that values we predict.	Airport security maintains a set of metrics and tries to predict the terrorist.
Regression	The regression using known data formats like linear or logistic assumes the future data format will fall into the data structure. It then tries to predict the value by applying some athematical algorithms on the data set.	Investing on Pension fund. Calculating your annual income and trying to predict what you need after you retire. Then based on the present income and needed income makes investment decision. The Prediction is done by simple

		regression formula to revise every year.
Time series analysis	With time series analysis, every attribute value is determined by the different time intervals.	Buying a company stock. Take X, Y, Z companies' month by month performance and try to predict their next one year performance and based on the performance, you buy stocks.
Prediction	Prediction is related with time series but not time bound. It is used to predict values based on past data and current data.	Water flow of a river will be calculated by various monitors at different levels and different time intervals. It then uses that information to predict the water flow of future.
Clustering	It is widely called as unsupervised learning. It is similar to classification except it won't have any predefined groups. Instead the data itself defines the group.	Consider a super market has buying details like age, job and purchase amount we can group by age against percentage as well job against percentage to make meaningful business decisions to target the specific user group.
Summarization	Summarization is associating the sample subset with small description or snippet.	Display the data as a graph and calculate mean, median, mode etc.
Association rules	It is also called as linked analysis. It is all about under covering relationship among data.	Amazon "People bought this also bought this" model.
Sequence discovery	Sequence discovery is about finding sequence of an activity.	In a shop, people may oftenbuy toothpaste after toothbrush. It is all about what sequence user buying the product and based on the shop owner can arrange the items nearby each other.

Table 1: Data Mining Techniques

II. ASSOCIATION RULE MINING

Association rule mining, one of the most important and well researched techniques of data mining. It aims to extract interesting correlations, frequent patterns, associations or casual structures among sets of items in the transaction databases or other data repositories. Association rule mining is to find out association rules that satisfy the predefined minimum support and confidence from a given database. Association rules are widely used in various areas such as telecommunication networks, market and risk management, inventory control etc.

The problem is usually decomposed into two sub problems. One is to find those item sets whose occurrences exceed a predefined threshold in the database; those item sets are called frequent or large item sets. Then other rule are generated by deleting the last items in the antecedent and inserting it to the consequent, further the confidences of the new rules are checked to determine the interestingness of them. The first sub-problem can be further divided into two sub-problems: candidate large item sets generation process and frequent item sets generation process. We call those item sets whose support exceed the support threshold as large or frequent item- sets, those item sets that are expected or have the hope to be large or frequent are called candidate item sets.

In many cases, the algorithms generate an extremely large number of association rules, often in thousands or even millions. It is nearly impossible for the end users to comprehend or validate such large number of complex association rules, thereby limiting the usefulness of the data mining results. Several strategies have been proposed to reduce the number of association rules, such as generating only "interesting" rules, generating only "non redundant" rules, or generating only those rules satisfying certain other criteria such as coverage, leverage, lift or strength. An association rule is an *implication* or *if-then-rule* which is supported by data. The motivation given in for the development of association rules is *market basket analysis* which deals with the contents of point-of sale transactions of large retailers[5]. A typical association rule resulting from such a study could be 90 percent of all customers who buy bread and butter also buy milk". Insights into customer behavior may also be obtained through customer surveys, but the analysis of the transactional data has the advantage of being much cheaper and covering all current customers. Compared to customer surveys, the analysis of transactional data does have some severe limitations, however. For example, point-of-sale data typically does not contain any information about personal interests, age and occupation of customers. Nonetheless, market basket analysis can provide new insights into customer behavior and has led to higher profits through better customer relations, customer retention, better product placements, product development and fraud detection.

III. CONCEPT OF ASSOCIATION MINING

Item : It is a field of the transaction database.

Transaction : It is corresponding to a record of the database. Transaction usually is marked as small letter t to mark item i. ti={i1,i2,…,ip}. Each transaction has an only identifier called TID. The whole set of transaction ti constitutes a database D. D={t1,t2,…,tn}

Support: The support of association rule X-> Y in transaction database is a ratio. The ratio is between the count of item set which contains X and Y, and the count of all of item set. That marks support(X->Y). That is the percent of the item set containing X and Y at the same time in the transaction database.

Confidence: It is the ratio between the count of transaction containing X and Y and the count of transaction containing X. That is marked as confidence(X ->Y). Confidence is the percent of the transaction sets containing X and Y at the same time in the transaction database.

Frequent Item set: The item set, whose support is not lower than the minimum support (Min Sup).

Strong rule and Weak rule: If support(X -> Y) >=MinSupport and Confidence (X->Y)>=MinConf, then mark association rule X-> Y as strong rule, otherwise mark it as a weak rule.

IV. SEARCHING FREQUENT ITEMSET

Frequent patterns, such as frequent item sets, substructures, sequences term-sets, phrase-sets, and sub graphs, generally exist in real-world databases. Identifying frequent item sets is one of the most important issues faced by the knowledge discovery and data mining community. Frequent item set mining plays an important role in several data mining fields as association rules [6] warehousing, correlations, clustering of high-dimensional biological data, and classification. Given a data set d that contains k items, the number of item sets that could be generated is 2k - 1, excluding the empty set[6]. In order to searching the frequent item sets, the support of each item sets must be computed by scanning each transaction in the dataset. A brute force approach for doing this will be computationally expensive due to the exponential number of item sets whose support counts must be determined. There have been a lot of excellent algorithms developed for extracting frequent item sets in very large databases. The efficiency of algorithm is linked to the size of the database which is amenable to be treated. There are two typical strategies adopted by these algorithms: the first is an effective pruning strategy to reduce the combinational search space of candidate item sets (Apriori techniques). The second strategy is to use a compressed data representation to facilitate in-core processing of the item sets (FP-tree techniques)

V. APRIORI ALGORITHM

In 1994 Agrawal etc. put forward famous Apriori algorithm according to the property of association rule: the sub sets of the frequent item set is also frequent item set, the supersets of non-frequent item set is also non-frequent item set. The algorithm each time makes use of k-frequent item set carrying on conjunction to get k+1 candidate item set.

The key idea of Apriori algorithm is to make multiple passes over the database. It employs an iterative approach known as a breadth-first search (level-wise search) through the search space, where k-item sets are used to explore (k+1)-item sets. The working of Apriori algorithm is fairly depends upon the Apriori property which states that” All nonempty subsets of a frequent item sets must be frequent” [5]. It also described the anti monotonic property which says if the system cannot pass the minimum support test, all its supersets will fail to pass the test.Therefore if the one set is infrequent then all its supersets are also frequent and vice versa. This property is used to prune the infrequent candidate elements. In the beginning, the set of frequent 1-itemsets is found. The set of that contains one item, which satisfy the support threshold, is denoted by . In each subsequent pass, we begin with a seed set of item sets found to be large in the previous pass. This seed set is used for generating new potentially large item sets, called candidate item sets, and count the actual support for these candidate item sets during the pass over the data. At the end of the pass, we determine which of the candidate item sets are actually large (frequent), and they become the seed for the next pass. Therefore, is used to find , the set of frequent 2-itemsets, which is used to find , and so on, until no more frequent k-item sets can be found. The feature first invented by in Apriori algorithm is used by the many algorithms for frequent pattern generation.

The basic steps to mine the frequent elements are as follows:

1.Generate and test: In this first find the 1-itemset frequent elements by scanning the database and removing all those elements from which cannot satisfy the minimum support criteria.

2.Join step: To attain the next level elements join the previous frequent elements by self join i.e. known as Cartesian product of . i.e. This step generates new candidate k-item sets based on joining with itself which is found in the previous iteration. Let denote candidate k-item set and be the frequent k-item set.

3.Prune step: is the superset of so members of may or may not be frequent but all frequent item sets are included in thus prunes the to find frequent item sets with the help of Apriori property. i.s This step eliminates some of the candidate k-item sets using the Apriori property A scan of the database.

To illustrate this, suppose n frequent 1-itemsets and minimum support is 1 then according to Apriori will generate and so on. The total number of candidates generated is greater than Therefore suppose there are 1000 elements then 1499500 candidate are produced in 2 itemset frequent and 166167000 are produced in 3-itemset frequent .
It is no doubt that Apriori algorithm successfully finds the frequent elements from the database. But as the dimensionality of the database increase with the number of items then:

1) More search space is needed and I/O cost will increase.
2) Number of database scan is increased thus candidate generation will increase results in increase in computational cost. Therefore many variations have been takes place in the Apriori algorithm to minimize the above limitations arises due to increase in size of database.
These subsequently proposed algorithms adopt similar database scan level by level as in Apriori algorithm, while the methods of candidate generation and pruning, support counting and candidate representation may differ.

The algorithms improve the Apriori algorithms by:
1) Reduce passes of transaction database scans
2) Shrink number of candidates
3) Facilitate support counting of candidates

Pseudocode for Apriori algorithm

Join Step: C_k is generated by joining Lk-1with itself

Prune Step: Any (k-1)-itemset that is not frequent cannot be a subset of a frequent k-itemset

C_k: Candidate itemset of size k
L_k: frequent itemset of size k
L_1= {frequent items};
for(k= 1; L_k!=∅; k++) do begin
C_{k+1}= candidates generated from L_k;
for eachtransaction tin database do
increment the count of all candidates in C_{k+1}that are contained in t
L_{k+1}= candidates in $C_{k+1}$with min_support
end
return$\cup_k L_k$;

Apriori-tid Algorithm:

Similar to the Apriori algorithm, the AprioriTid algorithm also uses the apriori-gen function to determine the candidate itemsets but the difference is that the database is not used for counting support after the first pass. Instead, set of candidate itemsets is used for this purpose of k>1. In case a transaction does not have any candidate k-itemset then the set of candidate itemsets would not have any entry for that transaction which will eventually decrease the number of transaction in the set containing the candidate itemsets as compared to the database. As value of k increases each entry will be smaller than the corresponding transactions because the number of candidates in the transactions will decrease. Apriori performs better than AprioriTid in the initial passes but in the later passes AprioriTid has better performance than Apriori.

Apriori –Hybrid:

Based on a different concept, this algorithm supports the idea that it is not necessary to use the same algorithm in all passes over data. As mentioned in [7], Apriori demonstrates better performance in earlier passes, and Apiori-TID outperforms Apriori in later passes. Based on the experimental observations, the Apriori-Hybrid technique was developed which uses Apriori in the initial passes and switches to Apriori-TID when it expects that the set *Ck* at the end of the pass will fit in memory. Therefore, an estimation of *Ck* at the end of each pass is necessary. Also, certain cost is involved in switching from Apriori to Apriori-TID. The performance of this technique was also evaluated by conducting experiments for large datasets. It was observed that Apriori-Hybrid performs better than Apriori except in the case when the switching occurs at the very end of the passes [5].

VI. REVIEW ON VARIOUS IMPROVEMENTS OF APRIORI ALGORITHM

Several improved algorithms have been proposed to conquer drawbacks of Apriori algorithm in several ways. Here presents six different approaches that face the common drawback.

1 Intersection and Record filter approach

1.1 Enlightenment

To present proposed algorithm, Goswami D.N., Chaturvedi Anshu and Raghuvanshi C.S.[4] has given Record filter and Intersection approach. In Record filter approach, count the support of candidate set only in the transaction record whose length is greater than or equal to the length of candidate set, because candidate set of length k , can not exist in the transaction record of length k-1 , it may exist only in the transaction of length greater than or equal to k. In Intersection approach, to calculate the support , count the common transaction that contains in each element‟ s of candidate set. This approach requires very less time as compared to classical Apriori. In Proposed Algorithm, set theory concept of intersection is used with the record filter approach. In proposed algorithm, to calculate the support, count the common transaction that contains in each element‟ s of candidate set. In this approach, constraints are applied that will consider only those transaction that contain at least k items.

1.2 Disadvantage Memory optimization is done but still it needs much optimization.

2 Improvement based on set size frequency

2.1 Enlightenment

To eradicate non significant candidate keys the modified algorithm introduces issues such as set size and set size frequency .These issues can diminish candidate keys in a more efficient way. The improved algorithm for Apriori[5] takes for the set size which is the number of items per transaction and set size frequency which is the number of transactions that have at least set size items. Initially database is given with set size and second database is of set size frequency of the initial database. Remove items with frequency less than the minimum support value initially and determine initial set size to get the highest set size whose frequency is greater than or equal to minimum support of set size. Set size which are not greater than or equal to min set size support are eliminated.

2.2 Disadvantage

Ideal starting size of combination size for pruning candidate keys is not given.

3 Improvement by reducing candidate set and memory utilization

3.1 Enlightenment

This algorithm[6] introduces a more efficient way to achieve the pruning operation. The algorithm only needs to search Lk-1 one time to complete the deletion and the remaining of each element X in Ck. The idea of the algorithm is as follows. Ik is a k-dimensional itemset. If the number of (k-1)-dimensional subsets of all (k-1)-dimensional frequent itemset Lk-1, which contains Ik, is less than k, then Ik is not a k-dimensional frequent itemset..So the improved algorithm only needs to match up the count of each element of Lk-1with the count of each element (X) of Ck (each element X has a count). If the count of the element X equals to k, then keep X. Otherwise X must be deleted. I/O speed can be deduced by cutting down unnecessary transaction records.The item that not appears in Lk-1 will no longer appear in Lk. So we can revise these items to null in the transaction database. Then we can pay no attention to these data information in any search work to D. At the same time, delete the transaction records (T) of which the number of valid data is less than k so as to deduce the database.[4] Then the candidate set Ck will be generated by latest D. The deletion of D will greatly reduce the number of transaction records which will effectively increase the speed of the implementation of the algorithm. Ultimately this will increase efficiency and I/O speed of algorithm.

4 Algorithm based on Trade list

4.1 Enlightenment

This algorithm scans the database at the start only once and then makes the undirected item set graph.[7] From this graph by considering minimum support it finds the frequent item set and by considering the minimum confidence it generates the association rule. If database and minimum support is changed, the new algorithm finds the new frequent items by scanning undirected item set graph. That is why it‟ s executing efficiency is improved conspicuously compared to classical algorithm. It makes each item as a node(V) and at the same time it makes the supporting trade list for each node. Supporting trade list is a binary group T= {Tid, Itemset} (where Tid is transaction id and Itemset is trade item set). So the side between nodes can be accomplished by corresponding trade list operation. The algorithm does the intersection of two nodes with supporting trade list. For search strategy select a node Vi from node set V. If the number of times Vi appears in the database is not less than the minimum support minsup, then {Vi} will belong to the frequent 1-item set. If count of node Vi adjacent to node Vj‟ s side is not less than support S, then { Vi, Vj } will belong to the item in frequent 2-iterm set. When there are three nodes in undirected item set graph and count of each side of the node is not less than minimum support minsup, these three nodes { Vk,Vm,Vn} will belong to frequent 3-item set. When there more than three nodes in undirected item sets graph then count of each side of the node should not be less than minimum support minsup and all the subset of these n nodes should be frequent. Subsequently nodes are added to k-item set. Main advantage of this approach is scanning of database is done once and after that the graph to find frequent item set.

5 Algorithm based on frequency of items

5.1 Enlightenment

In this paper, Mamta dhanda suggests a ground-breaking and attentive approach for the mining of interesting association patterns from transaction database.[8] First, frequent patterns are discovered from the transactional database using the Apriori algorithm. From the frequent patterns mined, this approach extracts novel interesting association patterns with emphasis on significance, quantity,profit and confidence. To overcome the weakness of the traditional association rules mining,Weighted association rule mining[6] have been proposed. Weighted association rule mining considers both the frequency and significance of itemsets.It is helpful in identifying the most precious and high selling items which contribute more to the company‟ s profit . This approach proposes an efficient idea based on mainly weight factor and utility for mining of high utility patterns. Initially, the proposed approach makes use of the classical Apriori algorithm to generate a set of association rules from a database. Firstly it uses attributes to get frequent item set. These attributes are like profit ratio calculation using Q-factor. Q – Factor = P / ΣPi (1) Than it gives Transactional database where each item‟ s frequency is counted in each transaction. From that pruning is done with minsup and confedence. Finally calculation of Weighting –factor is done based on frequency of itemset and Q-factor. n PW = Σ frequency * Q – Factor (2) i=1 Finally efficient frequent pattern is selected based on min PW-factor.

5.2 Disadvantage Initially classical algorithm is used. To improve efficiency some improvement can be done on pruning for faster execution.

6 Utilization of Attributes

6.1. Enlightenment

In this approach[10] using Tanagra Tool frequent item set is found by applying Apriori algorithm on database. Main problem of finding all association rules that satisfy minimum support and confedence thresholds given by users. Work illustrates that Association rule mining has several problems that it only tells whether item is present in database or absent , it treats all present or absent items equally ,it does not consider importance of item to user/business perspective and it fails to associate output i.e. frequent items with user and business objectives. These disadvantages can be removed by using attributes like profit, quantity, frequency of items which will give important information to user and business.

6.2 Disadvantage Various attributes like frequency, weight can be associated with frequent item set which can provide more information for business and user point of view ,which is not done here.

VII CONCLUSION

It is very important to have a data mining algorithm with high efficiency because transaction database usually are very large. Various algorithms have been proposed for mining association rule but in every algorithm there founds a common drawback of various scans over the database.The aim of this paper is to introduce association rule with apriori approach in various form.After doing survey of above algorithms conclusion can be given by this paper is that mostly in improved Apriori algorithms ,aim is to generate less candidate sets and yet get all frequent items. In the approach of Intersection and Record filter, intersection is used with the record filter approach where to calculate the support, count the common transaction that contains in each element‟ s of candidate set. In this approach, only those transactions are considered that contain at least k items. In other approach set size and set size frequency are considered.

As we know that each and everything has its required advantages and disadvantages So the Apriori algorithms also has its advantages and its various uses are given by

1. Initial information: transactional database D and user-defined numeric minimum support threshold min_sup
2. Algorithm uses knowledge from previous iteration phase to produce frequent itemsets.
3. This is reflected in the Latin origin of the name that means ”from what comes before”.

The various limitations of the Apriori algorithm are given by

1. Needs several iterations of the data.
2. Uses a uniform minimum support threshold .
3. Difficulties to find rarely occurring events .
4. Alternative methods (other than Apriori) can address this by using a non-uniform minimum support threshold .
5. Some competing alternative approaches focus on partition and sampling .

VII. REFERENCES

[1] Ben franklin, Genealogical data Mining,2006 .
[2] Maragatham G ,Lakshmi M,”A Recent Review On Association Rule ining”,Computer Science & Engg Department, Sathyabama University,Chennai, Tamil Nadu, India.
[3] Ramakrishnan Srikant,”Fast Algorithms for Mining Association Rules and Sequential Patterns”, Ph.D Dissertation, 1996, University of Wisconsin, Madison.
[4] Qiankun Zhao,Sourav S. Bhowmick, “Association Rule Mining: A Survey”, Nanyang Technological University, Singapore.
[5] Wael A. AlZoubi, Azuraliza Abu Bakar, Khairuddin Omar, " Scalable and Efficient Method for Mining Association Rules ", Proc. Of IEEE, vol. 1,no.8, 2009, pp. 36-41. [6] Lijuan Zhou, Shuang Li, Mingsheng Xu “Research on Algorithm of Association Rules in Distributed Database System”, Proc. Of IEEE, vol. 3,no.9, 2010, pp. 216-219
[6] Agrawal, R., Imielinski, T., and Swami, A. N. 1993. Mining association rules between sets of items in large databases.
[7] Rakesh Agrawal, Tomasz Imielinski, and Arun N. Swami,”Mining Association Rules Between Sets of Items in Large Databases”, Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data, May 1993, pp. 207-216.
[8] Mamta Dhanda,” An Approach To Extract Efficient Frequent Patterns From Transactional Database”,In: *International Journal of Engineering Science and Technology (IJEST), Vol.3 No.7 July 2011, ISSN:0975-5462*
[9] Andrew Kusiak, Association Rules-The Apriori algorithm[Online],Available: http://www.engineering. uiowa. edu/~comp /Public Apriori.pdf
[10] Mamta Dhanda, Sonali Guglani , Gaurav Gupta, ”Mining Efficient Association Rules Through Apriori Algorithm Using Attributes”, In: *International Journal of Computer Science and Technology Vol 2,Issue 3,September 2011,ISSN:0976-8491*
[11] David Wai-Lok Cheung, Vincent T. Ng, Ada Wai-Chee Fu, and Yongjian Fu,”Efficient Mining of Association Rules in Distributed Databases”, IEEE Transactions on Knowledge and Data Engineering, Vol. 8, No.6, December 1996.
[12] Sotiris Kotsiantis, Dimitris Kanellopoulos,”Association Rules Mining: A Recent Overview”,Educational Software Development Laboratory Department of Mathematics, University of Patras, Greece.

XI .BIOGRAPHY

Pooja Agrawal received her Master Degree in Computer Science from Guru Ghasi Das University, Bilaspur(C.G.).She Completed her M.Phil In computer science from Dr. C. V. Raman University, Kargi Road Kota,Bilaspur. She is currently pursuing Ph.D. in Computer Science in Dr. C. V. Raman University, Kargi Road Kota, Bilaspur. She is having 10 years of teaching experience. Her Research interest area in Data mining.

Suresh Kashyap received his Master Degree in Master Diploma in Computer Science and application from Guru Ghasi Das University, Bilaspur(C.G.).He Completed his M.Phil In Information Technology from Dr. C. V. Raman University,Kargi Road Kota,Bilaspur.He is currently pursuing M.Tech. in Computer Science in Dr. C. V. Raman University,Kargi Road Kota,Bilaspur. He is having 5 years of teaching experience.His Research interest area in Software Engineering.

Vikas Chandra Pandey received his Master Degree in Information Technology from Dr. C. V. Raman

.

University,Kargi Road kota ,Bilaspur(C.G.).He Completed his M.Phil In computer science from Dr. C. V. Raman University,Kargi Road Kota,Bilaspur.He is currently pursuing Ph.D. in Computer Science in Dr. C. V. Raman University,Kargi Road Kota,Bilaspur.He is having 5 years of teaching experience.His Research interest area in Data mining.

Suraj Prasad Kesari received his Master Degree in Master Diploma in Computer Science and application from CSVTU Bhilai raipur(C.G.).He Completed his M.Phil In Information Technology from Dr. C. V. Raman University, Kargi Road Kota,Bilaspur. He is currently pursuing M.Tech. in Computer Science in Dr. C. V. Raman University, Kargi Road Kota,Bilaspur. He is having 5 years of teaching experience. His Research interest area in Software Engineering.