International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-6, March 2020

Software Reusability of Object-Oriented Systems using Data Mining Techniques

Bharti Bisht, Parul Gandhi

Abstract: ***Due to fast advancement in software industry, there was a demand to cut down time and efforts during process of software development. While designing product and services it is very essential to assure quality of product in order to strengthen market value of the product. To accomplish both quality as well as productivity objectives, it is suggested to go for software reuse. Reusability is an essential measure that can be used to improve overall software quality with lesser cost and efforts. This paper gives insights into various literature studies related to software reusability of Object-oriented software using data mining techniques. In this paper even comparative analysis of various techniques related to prediction and enhancement of reusability of Object-Oriented software systems has been done. This would help to get better understanding of need of reusability enhancement of Object-Oriented systems using data mining techniques***

Keywords: ***Data Mining techniques, Object-Oriented, Reusability Process, Software Reusability.***

I. INTRODUCTION

This Slow growth in software productivity and raised demands means that most of the software industries have progressively found themselves in the midst of software crisis. This conceals their abilities to build manageable, good quality as well as profitable software. Software Reusability is an aspect that indicates software component's expected reuse potential. It also means using reusable components when software development takes place rather than generating code from scratch. Component can be treated as independent and replaceable element of the application that implements clear as well as distinct function [6]. Reusability is not only confined to the source code only but also covers reusability of design, document as well as test cases. Latest techniques are needed for reusing the actual knowledge from software factual data that includes project's progress and its status. Currently various data mining techniques are used to extract useful information from software factual data. Mapping of different software metrics for various data mining [10] techniques is being done that can be applied for reusability process. Object-Oriented based languages have made possible to build software even in component form and has also enables software reuse on large scale.

Manuscript received on February 10, 2020.
Revised Manuscript received on February 20, 2020.
Manuscript published on March 30, 2020.
* Correspondence Author
Bharti Bisht*, Department of Computer Applications MRIIRS, Faridabad, Haryana, India.
Dr. Parul Gandhi, Department of Computer Applications MRIIRS, Faridabad, Haryana, India.

There is urgency for Object-Oriented techniques that would implement an efficient approach to improve level of reusability of an organization in a reuse environment. There are some of the cases available in which existing reusable components are used to develop new systems. This makes software development process totally reuse driven process and even the modification of the requirements can consider the use of these components. The remaining paper is structured as:

Section II discusses about Software Reusability, Reusability Process and its need. Section III discusses about data mining and its techniques and Section IV discusses about various related work. Section V demonstrates the comparative analysis of various techniques used for reusability of Object-Oriented software. Section VI demonstrates the observations made after literature survey. Finally, Section VII includes conclusion and Future Work.

II. SOFTWARE REUSABILITY

Software Reusability means to amend software system using already existing components rather than constructing complete system from scratch. Reusability results into increased productivity as well as quality by overcoming the risk from construction of new projects. The various crisis related to reuse are [16]:

1) Inefficient Software [16]
2) Low quality Software [16]
3) Software not meeting the requirements
4) Unmanageable Project

Software Reusability can be a success only if there is a strong organizational foundation. There is a different way to look over software known as Product Line which contains products that are similar and addresses a particular domain. This process helps to create new system every time from a set of existing software assets instead of constructing a new system from scratch.

(A) Software Reusability Process:

Reusability process covers the following phases[3]:

i. Identification of Software Components:

In this phase, different components are searched in Software Repository such as Internal Memory and External. In Internal Memory components like keywords, natural languages etc (Ramamoorthy C. V 1988) are searched whereas in External, software libraries are looked over to search for any component.

ii. Context Understanding:

In this phase, internal as well as external representations of the software program are analysed using design patterns from existing documentation.

iii. Using Software Reuse Techniques:

During this phase, various software reuse techniques are applied such as Code Reusability, Design Reusability, and System Reusability etc.

iv. Integrating and Evaluating:

During this step, techniques such as Code cloning, Code Invocation and COTS (Ramamoorthy C. V 1988) are integrated and their accuracy is evaluated.

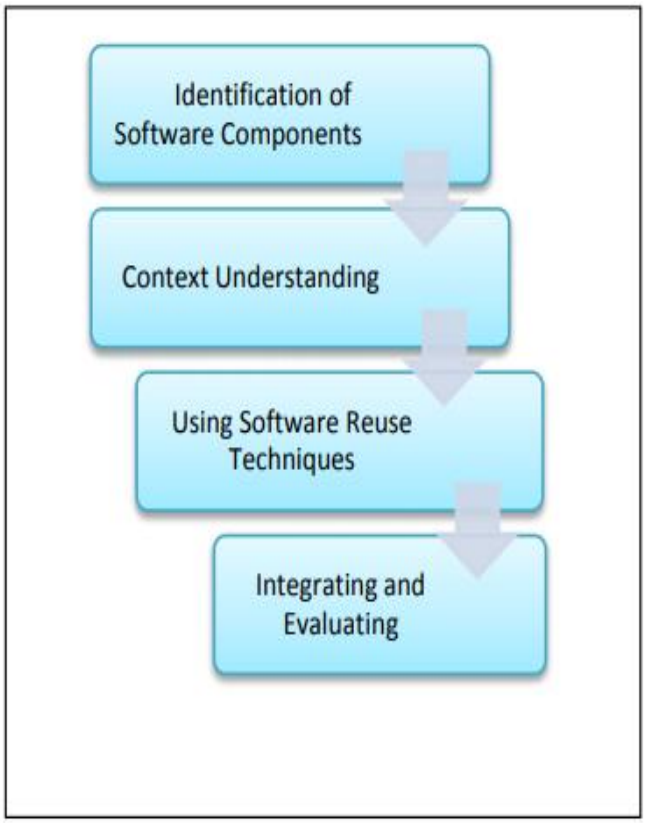

Fig.1. Reusability Process Phases [4]

Reusability provides some benefits [16]:

i. Reduces development cost
ii. Increased Reliability
iii. Reduced Market Time
iv. Reduction in maintenance cost[16]

(B) Software Reusability Techniques:

Software reusability of software component is performed using various reusability techniques. The various software reusability techniques are as follows:

i. **Domain Analysis:** It is very important part for a success of reuse program. This phase identifies elements and the operations required for processing of information of a specific domain. It would also help to identify various domain specific artifacts for finding out the best reusable component. This objective can be achieved by two basic approaches: Bottom-up approach and Top-down approach

ii. **Object-Oriented Based Technique:** In object-oriented language, objects represent real world entity. Class represents a list of all functions that can be implemented on the class elements. The functions which belong to a specific class are known as member functions of that class. Abstract class acts as a reusable component because it demonstrates the functionality which is used widely.

iii. **Standard Interfaces:** The use of these interfaces plays a crucial role in software reuse process. Information hiding among modules cannot be enforced without standards. If these standards are more general, then there is the possibility that software component can easily interface with other components leading to better chance of reuse of these components.

iv. Reuse Designing:

As we have discussed software reuse as a process which includes analyzing, organizing as well as evaluating different artifacts related to software. But a different approach can also be applied in which reuse can also be implemented with the help of building new systems. This would in turn increase the capability of reusing some of the components of these new systems.

v.Use of Metrics:

Metric usage is also an important part of reuse process. Metrics can also be used to check on the reuse programs status, quality of components assessment and even track the cost involved in the reuse process.

III. DATA MINING AND TECHNIQUES

Data Mining is the technique of extracting hidden patterns or information from large collection of raw data. Data Mining also known as KDD(Knowledge Discovery and Data Mining) that can be used by organizations in order to take decisions by analyzing as well as accessing data. Tools like Query tools, Analytical tools and DSS tools can be used for this purpose.

The various steps involved in KDD Process are depicted in Fig 2(Han & Kamber 2001).

(1) **Data Cleaning:** The data which has been collected from various sources are not clean and even contain missing values .So there is a need of applying various techniques to clean these anomalies. In this phase, noise as well as inconsistent data is eliminated.

(2) **Data Integration**: In this step, data from several sources i.e. heterogeneous are integrated and combined into a common source(Dogra &Tanuj 2015)

(3) **Data Selection**: This deals with recovering of appropriate data for the particular task from the given database

(4) **Data Transformation**: This phase is also known as Data Consolidation Phase. Data is consolidated into forms which are very suitable for mining process through summaries.

(5) **Data Mining:** This is the important process which helps in extracting data patterns from the available data

(6) **Pattern Evaluation**: This deals with recognizing of meaningful patterns or information on the basis of given measures.

(7) **Knowledge Representation**: This is the final step in KDD process. In this phase, extracted information is displayed using various visualization techniques.

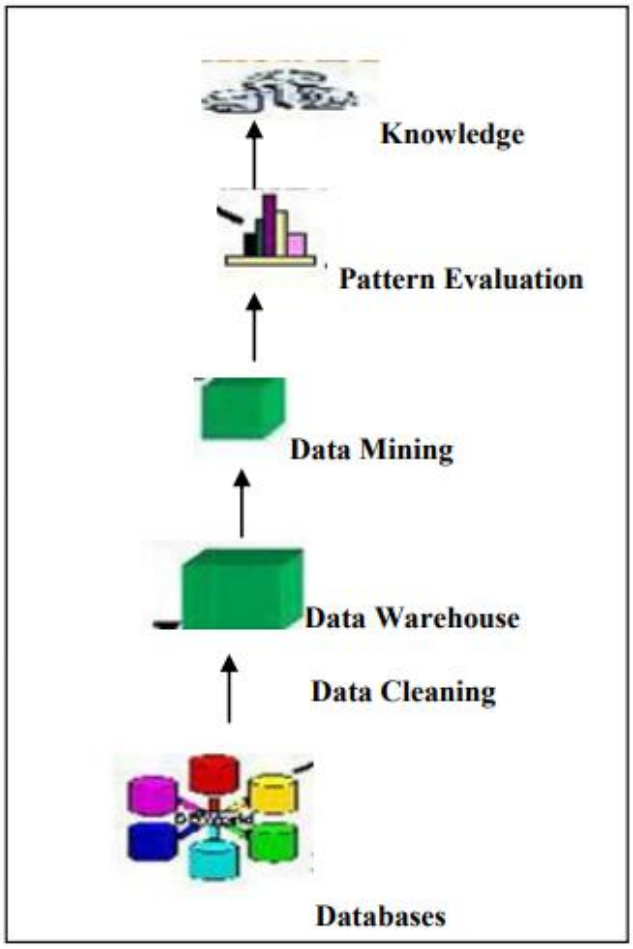

Fig.2. Steps in KDD Process [8]

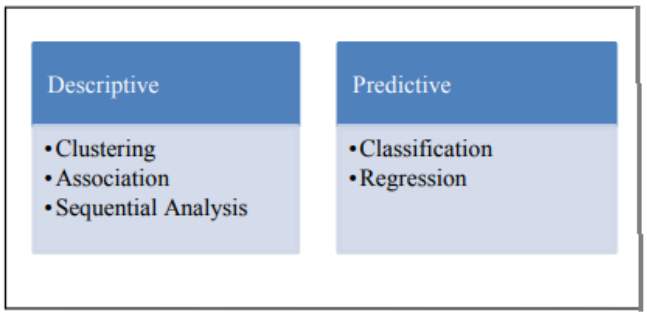

Fig.3. Data mining Techniques [8]

(A) Classification Technique

This technique is also known as Supervised learning technique. This technique involves two steps. In first and foremost step, model is developed by analyzing data rows from training data consisting set of all attributes.
Various Classification techniques are as follows:

i. Decision Tree :

These are induction techniques that are used to find classification rules for a selected attribute from a given data set by subdividing information in a systematic way. It contains Root node, Branch nodes as well as Leaf Nodes

ii. Bayesian Classification:

It is based on Bayes Theorem. Naive Bayes model is very easy to build which makes it useful for very large datasets.

iii. K-nearest Neighbour Technique:

This technique is one of the basic but essential technique in Classification Approach. This is the supervised approach that has vast application in pattern recognition process, intrusion detection process.

iv. Genetic Algorithm:

These algorithms are adopted from Darwin's principle (principle of fittest survival).It is used to find solutions for finding problems which are based on natural selection theory.

(B) Clustering Technique

It is used to find out the clusters in which objects of one cluster are similar to each other and different from objects in other clusters. It is one of the important unsupervised learning approaches.
Various Clustering techniques are:

i. Partitioning Method:

These methods are used to classify observations into multiple groups from a given similar sets. In this algorithm it is necessary to mention the number of clusters which are to be generated.

ii. Hierarchical Method:

It is a method in which data partition is done in series of steps rather than single step.

iii. Density Based Method:

It is used for identifying non-linear shaped structures on the basis of density. DBSCAN is one of the most widely used algorithm that uses density connectivity concept.

iv. Grid Based Method:

This technique develops hierarchical structure of the available data and even provides solutions to various queries effectively.

(C)Association Mining:

Association Rules depicts statements that can be utilized to find out the relationship among unrelated data present in the database. It is a two-step approach:

i. Generation Of Frequent Item set Generation:

It is generation of rules by using transactional dataset. In this technique, those items which are purchased frequently are kept together.

ii. Generation of Rules:

Once the frequent item set is generated then the rules are identified out of these item sets. These rules are generated with the help of binary partition.

(D) Regression

It is a technique in which numeric values range (Continuous) for a particular given dataset are predicted. The oldest technique is a linear regression which is used to find the relationship among two variables. Equation used in this technique is:
$Y=mx+b$
Various Regression techniques are as follows:

i. Standard Multiple regression:

It is used to find relationship between various independent variables and dependent variables.

ii. Stepwise Multiple regression:

It is used to find out the best predictors.

It evaluates the order of predictor variables and then find out the appropriate subset.

iii. Hierarchical Regression:

It is like stepwise method but firstly the predictor variables are mentioned in a pre-defined order. This algorithm do not contains any built-in equations.

(E) Sequential Analysis

It is a technique that extracts sequential patterns. These pattern's support value exceeds the given minimal support threshold value. It also mines the sequences that represent frequent behaviours. It is also used in organizations to analyse behaviour of customers.

Algorithms used in this technique are:

i. Apriori-like Algorithms
ii. BFS-Based Algorithms
iii. DFS-Based Algorithms
iv. Incremental-Based Algorithms

IV. LITERATURE REVIEW

The detailed literature review of software reusability models and methods related to Object-oriented system is discussed in this section.

Sandhu P., Singh H. (2006) designed a model for predicting reusability of various software components. In their work, they used C&K (Chidamber and Kemerer**)** metrics and designed Neuro-Fuzzy [17] engine to predict the reusability. Further an algorithm was proposed in which inputs was values from metrics and output achieved was in form of Reusability Index of various software components. The already trained Neuro-Fuzzy [17] system will extract values of metrics from previous phases and then will calculate the Reusability Index.

Gill N. (2006) discusses about need of characterization of software components for enhancing software reusability of component-based systems. This paper also points out further advantages of component characterization [6] and even mentions issues related to component-based systems.

Mierswa I., Wurst M., Klinkenberg R., Scholz M., Euler T. (2006) KDD is complicated and critical task. Rapid Prototyping is the technique that should support maximum reuse process. This paper defines Yale, free and open-source environment that can be used for KDD process. It also demonstrates the advantages of rapid prototyping for KDD process.

Sandhu P., Singh P., Singh H. (2007) designed Neuro-Fuzzy system using Taguchi Algorithm. This system is used to optimize reusability of object-oriented data by combining decision tree algorithm with results of Taguchi algorithm. Comparison of various other techniques with Neuro-Fuzzy algorithm showed Neuro-Fuzzy algorithm to be superior one.

Veras R., Meira S., Oliveira A., Melo B. (2007) presented comparison of clustering techniques which were applied to software repositories for software reuse process. Techniques used in this paper were SOM [30] and GHSOM [30]. Comparison showed that GHSOM [30] is much more suitable technique for software reuse.

Rao C., Niranjan P. (2008) Reuse helps in improving productivity as well as quality of software products. In order to reuse software components [20] it is very important to find out the component that can be further reused. In this paper, various classification techniques were compared and even new approach for classifying software component's information was also proposed.

Sharma A., Grover, P., Kumar R.(2009) Software reuse has emerged as a tool to cut down the cost of development as well as time required for a software. Currently, large numbers of software modules are being constructed by integrating various reusable software modules or components. In this paper, artificial neural network [19] approach is used to predict the reusability of software modules or components. This approach will help to find out the right module or component for reusability analysis.

Sandhu S., Kaur H., Singh A. (2009) designed a reusability prediction model for object-oriented based software system. In this paper, they designed their software metrics and worked on neural-network approach. In this approach value

of metrics was taken as input and value of reusability level of object-oriented software was calculated. The experiments showed that the calculated value of reusability level was same as the value being calculated manually. But the proposed system produces much more good quality software.

Drown D. J., Khoshgoftaar M., Seiya N. (2009) used Genetic algorithm using evolutionary model that maximize the performance. This approach is not applicable to small projects. The risk analysis of these projects requires highly skilled resources.

Fokaefs M., Tsantalis N., Chatzigeorgiou A., Sander J. (2009) In object-oriented systems, classes may become very large and less cohesive[].This papers deals with agglomerative clustering method. This helps to identify solutions according to their impact on design quality [9].

Shri A., Sandhu P., Gupta V. (2010) Identification of reusability metrics during design stage or coding stage can cut down the rework process resulting in improved quality of software. In this paper, C&K (Chidamber and Kemerer**)** metrics are used to analysis reusability level of object-oriented software systems. In this, K-means Clustering approach along with Decision –tree algorithm is also used. The proposed reusability prediction model produces the desired precision results.

Manhas S., Vashisht R., Sandhu P., Nirvair N. (2010) developed a model that evaluates the reusability level of software modules. They used five neural network based techniques considering values of proposed metrics as input. The calculated value of reusability level helps to identify the good reusable software component automatically.

Anju S. (2010) used K-means clustering technique and decision tree approach to distinguish between reusable components with "Nil" value and "Excellent" value. It also uses Rapid Miner tool. The proposed model produced the high precision results along with satisfactory low error values.

J. Bhagwan, Oberoi A. (2011) mentioned in their paper that evolution of software is increasing in size but the work force to manage this evolution is not increasing accordingly, which results in software failure [10]. So, to solve this crisis researcher suggested that software reuse could play a powerful role.

In this paper, researchers work on reusability with help of clustering technique.

Parashar A., Chhabra K. (2011) Reusability process is successful if there is appropriate identification of reusable software component. In case of object-oriented software, finding out appropriate class plays a important role in it's reusability. In this paper K-Mean clustering technique has been used.

Krishna T., Vasumathi D. (2011) In order to deliver optimal software, reuse of existing software components proves to be better than developing software products from scratch. In this paper, Netminer approach is used for this purpose. It takes source code as input and then various data mining techniques are applied on the input. Patterns are even identified in the input data. As a result optimal software is developed.

Alkhalid A., Alshayeb M., Mahmoud S. (2011) As software adapts new requirements [1],it becomes more and more complex as well as quality also decreases. In this paper, method to identify ill-structured software at the class level has been proposed. For this purpose, Adaptive K-nearest Neighbour algorithm has been applied.

Kumar A. (2012) Slowly and steadily software community is shifting towards software reusability concept. This reusability process helps to construct new software from the existing systems only. Code reusability is the new approach emerging in recent years. In this paper, software reusability is being performed by using SVM algorithm. This algorithm is run under MATLAB environment.

Kakkar P., Sharma M., Sandu P. (2012) constructed their own modules to predict the reusability level of different software components. They developed flexible model to perform pattern recognition by finding out features that helped them to measure reusability aspect of software components and even used Naïve Bayes Approach to study reusability level of software components.

Prakash A., Ashoka D., Aradhya M. (2012) used typical KDP (Knowledge Discovery Process). Clustering, classifications as well as visualization are the techniques that have been used to evaluate reusable components. C&K (Chidamber and Kemerer) tool has been utilized to construct software metrics for Object-Oriented programming language.

Shatnawi A., Abdelhak S. (2013) proposed an approach for mining reusable components from similar set of Object-Oriented software. The main aim behind this approach is to improve reusability of these reusable components. They have even validated this approach using two open source JAVA applications.

Gupta C., Rathi M. (2013) proposed an approach to identify reusable components from existing software data using Genetic algorithm, classification, K-means method and as well as Rabin Karp algorithm. This would help to find out those software parts that can be reused resulting in reduced efforts, cost as well as time.

Banu M., Thoufeeque N., Archana K. (2013) Software component is the part and parcel of software development process. Reusability is one of the major attribute of the software component. This is one of the best approach to overcome challenges related to existing systems. In this paper, various data techniques have been applied on software component.

Hajizadeh N., Keshtgari M., Ahmadzadeh M. (2014) Reusability is the technique of utilizing reusable component during software development process rather than working on new component. This not only cuts down development cost but also reduces the extra time. They used 42 classification algorithms on reuse data set(PROMISE data set) to predict success or failure of Software reusability as well as they have used 8 evaluation metrics and then comparing of these metrics showed that SGD and Multi class classifier updateable algorithms gave better results from other algorithms.

Prakash A., Ashoka D., Aradhya V. (2014) Existing software components are being included during software development process to cut down the time needed for software delivery and even produce good quality software. Reuse process requires existing information from repositories. So for this, data mining could be the best approach. In this paper, data mining technique has been applied on 167 input data from existing software systems.

Srinivas C., Radhakrishna V., Rao C. (2014) Software libraries are being developed continuously. So, selection of suitable software component from these libraries is a challenging process. Mining and then clustering of these components for software reuse process is becoming more popular among researchers. In this paper, main idea was to extract and cluster the best software component from software libraries for reuse process. Feature selection technique is also applied.

Keswani R., Joshi S., Jatain A. (2014) Software Reuse helps to assist the software development process. Constructing systems out of pretested components [17] can save design cost and time taken to write new code.This paper provides overview of software reuse and it's research contribution.

Vispute J. N., Sahu K. D., Rajput A. (2015) discussed data mining as a process of analyzing large amount of data which extracts hidden patterns from available data. Various data mining algorithms are used to test the accuracy of datasets. Classification technique predicts results on the basis of given input. Algorithms like Naïve Bayes, SGD and H-SOM are used to compare data and then the final results are simulated.

Srinivas C., Rao C. (2015) Clustering, an unsupervised technique can be used to create clusters from different software modules. This paper focuses on K-means algorithm along with feature vector approach. This algorithm uses distance measure as input and then the algorithm is applied on collected inputs.

Chopra N., Dhanda N. (2015) Object-oriented techniques helps to increase reusability of software components. Object-oriented software systems are much easier to reuse than component based software systems. These systems utilize the advantage being provided by their inheritance property. This paper analyzes the advantages of object-oriented systems and even suggests the ways to improve reusability of these systems.

Zafar M., Aslam R., Ilyas M. (2015) Reuse is a technique of using current available software modules to update systems. This paper develops a method by combining classification as well as clustering technique. In this each reusable modules are assigned some clusters and then classification approach is applied.

Vodithala S., Pabboju S. (2015) In order to attain software reusability, an adequate design of software archive is required. This paper discussed clustering technique for designing of software component archives. Clustering technique is applied on the basis of behaviour of different software components.This technique helped in to increase the rate of similarility measure between various software components.

Deeba K., Amutha B. (2016) proposed that data mining is a technique of extracting information from available data. Data mining uses several techniques to anticipate relationship between available data and the expected results. The paper discusses about software reusability for classifying data using data mining approach.

Hudaib A., Huneiti A. (2016) Reusability plays an important role in improving quality of software products. It cuts down

cost, time as well as efforts required during software development cycle. Reusability detection models are developed early in the process so that early assessment of reusability level is possible. This paper uses C&K (Chidamber and Kemerer**)** metrics in order to identify reusability level of Object-oriented system.SOM technique is also used to create clusters of values extracted from these metrics.

Araar I., Seridi H. (2016) From past many years, various organizations have started working on Reusability techniques in order to improve the productivity. In this, authors have analysed the overlapping clustering techniques to extract functional features of object-oriented systems from their source code.

Mateen A., Kausar S., Sattar A. (2017) Reusability of software products has increased in recent years. A good organised software reuse process can help to improve productivity as well as quality of produced software. The main idea behind this research was to find out how software reusability can leads to good quality software.

Sondhi T., Ghildyal S., Sabharwal S., Nagarkar A., Lavanya K. (2017) studied various techniques available for reuse from traditional technique like REBOOT to the later techniques like PULSE as well as object-oriented methodologies. These techniques helped to decrease cost and even time also.

Rathee A., Chhabra J. (2017) Object-Oriented programming in software development process is being widely adopted, so there is a need of well designed software systems [ ].Cohesion is one of the essential quality attribute in Object-Oriented software systems used to measure [] quality of these systems. In this paper software is analyzed to detect those classes with unrelated member functions [] and then code problem [] is detected. This helps to improve software quality. Hierarchical clustering algorithm is also applied on the detected classes.

Divanshi P. W., Singh A. (2018) presented a model for component reusability prediction along with right goals which was selected for reusability level. This model uses classification as well as decision tree approach to enhance quality of a software component.

Maheswari G., Chitra K. (2019) presented comparison of H-SOM and Naïve Bayes algorithms using R data mining tool. The comparison results proved that H-SOM [12] algorithm is more time efficient approach rather than Naïve Bayes [12] Algorithm.

V. RESULTS

This section depicts the comparative analysis of different techniques used for reusability of Object-oriented software system with their merits as well as demerits

Table- I: Analysis of Various Used Data Mining Techniques in Different Research Studies

Techniques/Approach	Dataset/Input	Merits	Demerits
Genetic algorithm, classification, k-means method and as well as Rabin Karp algorithm	36 software components collected from various websites	Provided better approach of software reuse	It cannot deal with complex datasets
NetMiner	Structured SE data	Provided optimal software	It cannot deal with unstructured data
Agglomerative Clustering	Object-Oriented classes	Identifies solutions according to their impact on design quality	Coincident clusters may result
Adaptive K-nearest Neighbour Algorithm	Object-Oriented classes	Reduces Software complexity and improves Quality	Only fixed number of classes were consider
Feature Extraction Approach	Object-oriented source code	Identifies the best software component that can be reused	Only fixed number of classes were consider

Self-Organizing Map(SOM)	CK metric along with Object-oriented classes	It can help to visualize the relationship between software metrics and its reusability level	Less time - efficient
Clustering, Classification and Visualization	36 software components collected from various websites	Provided better approach of software reuse	It cannot deal with complex datasets
Artificial Neural Network	NASA repository	It is accurate due to its own learning capability	It cannot deal with inaccurate information
Classification with Clustering	36 software components collected from various websites	Provided a better approach to extract reusable software component	In this proposed approach, clustering was done in specified format only
Decision Tree	Object-Oriented Metrics(WMC,DIT,NOC,CBO,LCOM)	Produced high results as compared to other techniques	Structuring of this approach is very complex
SVM	NASA Repository	Useful for reusability of software modules with different functionality	It can't be used for large collection of software metrics
K-means	NASA MDP	Worked on previous existing data	Required software metrics continuous values
Hierarchical Clustering	Object-Oriented source code classes	Improved Reusability of source code	It cannot deal with large datasets
Genetic algorithm using evolutionary model	Object-Oriented source code	Efficient model	This approach is not applicable to small projects.
Hybrid K-mean and Decision Tree	Object-Oriented software components	It helps to differentiate between 'Nil' and 'Excellent' Reusability value components	It cannot deal with large as well as complex datasets

VI. DISCUSSION

On analyzing various literatures, some observations can be summarized as follows:

i. Software Reusability is the best approach for improving quality of software and cutting down the cost as well as time required during software development process.
ii. Object-Oriented software components [20] are much more reusable than component based software systems.
iii. In Object-Oriented software, classes are common modules which can be easily reused.
iv. The Existing solutions do not fully answer the different reusability levels related to Object-Oriented based software components.
v. From [13],it was observed that H-SOM [13] is more time efficient [13] than other data mining technique
vi. H-SOM [13] is a good combination of pattern recognition [13] technique as well as classification technique.

VII. CONCLUSION AND FUTURE WORK

Despite the fact that software reusability helps to improve productivity as well as quality of software still it is considered as challenging approach. In this paper, we have presented various literature review of current research studies related to prediction of reusability of Object-oriented systems and even discussed the various data techniques or tools used for this.

We analyzed that the existing solutions do not fully answer the different reusability levels related to Object-Oriented based software components So after all analysis of various studies, working on different Reusability factors (design reuse, code reuse etc.) with the help of SOM technique could be a subject of future work.

ACKNOWLEDGMENT

I would like to acknowledge and thanks to Dr Parul Gandhi, Associate Professor and Ph.D. Coordinator in Faculty of Computer Applications, MRIIRS, and Faridabad for her support and guidance constantly during the preparation of this article.

REFERENCES

1. Alkhalid A., Alshayeb M., Mahmoud S.,2011, "Software Refactoring at the Class Level using Clustering Techniques" *Journal of Research and Practice in Information Technology*, Vol. 43, No. 4, pp. 285-306
2. Araar I., Seridi H., 2011," Software Features Extraction from Object-Oriented Source Code Using an Overlapping Clustering Approach" *Informatica* ,vol. 40 245–255.
3. Arifa B., Mohamed N., Archana K., 2013," Study of Software Reusability in Software Components", ,*IJET*, Vol. 5, pp. 2455-2460.
4. Basili J., B. V.R., Melo W., 1996, "How reuse influences productivity in object-oriented systems", *Communications of the ACM*, Oct. 1996, vol. 30, no.10, pp.104-114.
5. Bibi M., Mehboob R., Sirshar M., 2015, "Analytical Study of Data Mining Techniques for Software Quality Assurance", *International Journal of Computer and Communication System Engineering*, vol.2 (3), pp.377-386.
6. Deeba K., Amutha B., 2016, "Classification Algorithms of Data Mining*", Indian Journal of Science and Technology*, vol. 9, pp. 2-5.
7. Divanshi P. W., Singh A.,2018, "A Classification based Predictive Cost Model for Measuring Reusability Level of Open Source Software", *International Journal of Recent Research Aspects*, Vol. 5, pp. 19-23.
8. Dogra A., Wala T., 2015, "A Review Paper on Data Mining Techniques and Algorithms", *International Journal of Advanced Research in Computer Engineering & Technology*, Vol.4, pp.1976-1979.
9. Fokaefs M, Tsantalis N, Chatzigeorgiou A, Sander J ,2009, "Decomposing object-oriented class modules using an agglomerative clustering technique., *"In: Proceedings of the 25th international conference on software maintenance. Edmonton, Canada*, pp 93–101.
10. Frakes W B., Kang K., 2005, Software Reuse Research: Status and Future [J]. *IEEE Transaction on Software Engineering*, vol. 31(7) pp.529 - 536
11. Gill S. N., 2006, "Importance of Software Component Characterization for Better Software Reusability", *In ACM SIGSOFT Software Engineering Notes*, vol. 31, pp. 1-3.
12. Gupta C., Rathi M., 2013, "A Meta level data mining approach to predict reusability", *International Journal of Information Engineering and Electronics Business*, vol.6, pp. 33–39.
13. Hudaib A., Huneiti A., 2016, "Software Reusability classification and prediction using Self-Organising Map (SOM)*", in Communications and network* , vol. 8, pp. 179-192.
14. J. Bhagwan , A. Oberoi, 2011,, "Software Modules Clustering: An Effective Approach for Reusability", *Journal of Information Engineering and Applications*, Vol. 1, No.4.
15. Kakkar P., Sharma M., Sandu P.,2012, "Modeling of Reusability of Procedure based Software Components using Naïve Bayes Classifier Approach", *International Journal of Computer Applications*, Vol. 55 , pp. 12-17.
16. Karpagavadivu M., Dr. K., 2012, "A Survey of Different Software Fault Prediction Using Data Mining Techniques Methods", *International Journal of Advanced Research in Computer Engineering & Technology*, Vol. 1 pp. 1-3.
17. Keswani R., Joshi S., Jatain, 2014, "A Software Reuse in Practice" *2014 Fourth International Conference on Advanced Computing & Communication Technologies* vol. 57 pp. 159-162.
18. Liaghat Z., Rasekh A. H. , Tabebordbar A. R., 2012, "Enhance Software Quality Using Data Mining Algorithms", *S-CET Spring Congress*, pp. 1-5.
19. Maheswari G., Chitra K., 2019, "Enhancing Reusability and Measuring Performance Merits of Software Component using Data Mining", *International Journal of Innovative Technology and Exploring Engineering*, Vol.8, pp.1577-1583.
20. Manhas S., Vashisht R., Sandhu P., Neeru N.,2010, " Reusability Evaluation Model for Procedure Based Software Systems", *International Journal of Computer and Electrical Engineering*, Vol.2, pp. 1107-1110
21. Mateen A., Kausar S., Sattar A,2017,"A Software Reuse Approach and Its Effect On Software Quality, An Empirical Study for The Software Industry", *International Journal of Management, IT & Engineering* ,Vol. 7 Issue 2, pp. 266-279.
22. Mierswa I., M. wurst, R. Klinkenberg, M. Scholz and T.Euler, 2006,"Yale: Rapid prototyping for complex data mining task" International conference on knowledge discovery and data mining, ACM, pp. 935–940.
23. Morisio, Ezran M., Tully C., 2002, "Success and Failure Factors in Software Reuse", *IEEE Transactions on Software Engineering*, vol. 28, pp. 340–357.
24. Najadat H., Alsmadi I, 2012," Enhance Rule Based Detection for Software Fault Prone Modules", *International Journal of Software Engineering and Its Applications*, vol.6, pp. 75-86.
25. Prakash A., Ashoka D., Aradhya M.., 2012 ,"Application of Data Mining Techniques for Software Reuse process", *in Procedia Technology*, vol. 4, pp . 384 – 389.
26. Rao C., Niranjan P., 2008,"An integrated classification scheme for efficient retrieval of components", *Journal of computer science*, vol. 10, pp. 821-825
27. Rathee A., Chhabra J.,2017," Restructuring Of Object-Oriented Software Through Cohesion Improvement Using Frequent Usage Patterns", *ACM Sigsoft*, Vol. 42, pp. 1-8.
28. Sandhu P., Singh H., 2006, "A Reusability Evaluation Model for OO-Based Software Components", *International Journal of Computer Science*, vol. 1, no. 4, pp. 259-264.
29. Sankar K., Kannan S., Jennifer P., 2014 , "Prediction Of Code Fault Using Naïve Bayes And SVM Classifiers", *Middle-East Journal Of Scientific Research*, vol. 20, No. 1, pp. 108-113.
30. Sharma, A., Grover, P.S. and Kumar, R.,2009," Reusability Assessment for Software Components", *ACM SIGSOFT Software Engineering Notes*, vol. 34, pp. 1-6.
31. Shatnawi A., Abdelhak D. S., 2013, "Mining reusable software components from object-oriented source code of a set of similar software, Information Reuse and Integration", *2013 IEEE 14th International Conference on IEEE*, pp.2-18.
32. Shri A., Sandhu P. , Gupta V., Anand S., 2010,"Prediction of Reusability of Object Oriented Software Systems Using Clustering Approach*", International Journal of Computer and Information Engineering*, Vol. 4, pp. 853-858.
33. Sindre G., Conradi R. , Karlsson A., 1995 , "The REBOOT Approach to Software Reuse", *Journal of Systems and Software*, vol. 30, pp. 201-212.
34. Swathy V.,Niranjan R.,2012, "A Resolved Retrieval Technique for software Components", *IJARCET*, volume 1, issue 4
35. Thwin T., Quah S., 2005, "Application of neural networks for software quality prediction using object-oriented metrics", *Journal of systems and software*, vol. 76, no. 2, pp. 147-156.
36. Touw E., 2017, "Multi-faceted Reliability Assessment Techniques: An Industrial Case Study*", Software Architecture Workshops 2017 IEEE International Conference on IEEE*, pp.17-22.
37. Veras C., Meira, S.R., Oliveira A.L. , Melo, B.J. ,2007, "Comparative Study of Clustering Techniques for the Organization of Software Repositories", *19th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2007)*, vol. 4
38. Vodithala S., Pabboju S., 2015,"A clustering technique based on the specifications of software components," *2015 International Conference on Advanced Computing and Communication Systems*, Coimbatore, pp. 1-6.

AUTHORS PROFILE

Bharti Bisht, is doing her P.hD in Computer Science under the guidance of Dr Parul Gandhi in Manav Rachna International Institute of Research and Studies (MRIIRS), Faridabad. She has published paper in IEEE Digital Library and even presented her papers in International Conferences. She is active participant of various workshops and seminars conducted by educational institutions. Her research area is Software Reusability using Data Mining.

Dr Parul Gandhi, is working as Associate Professor, Manav Rachna International Institute of Research and Studies (MRIIRS), Faridabad and also as Associate Coordinator of Non Engineering Ph.D. Programs. A Doctorate in the subject of Computer Science with the study area in Software Engineering from Guru Jambheshwar University, Hisar. She is also a Gold Medalist in M.Sc. Computer Science. She is having a strong inclination towards academics and research. She has 13 yrs. of academic, research and administrative experiences. She has published many research papers in reputed International/ National journal and conferences. Her current research interests include software reusability, soft computing, and Software metrics and Component Based Software Development etc.