

Intelligent Decision Support System for Software Project Management using Data Mining Techniques

Satish Kumar
Student, M-Tech (SE) Department,
U.I.E.T., M.D.U., India
nit9292@gmail.com

Abstract

This Paper aims proposed a decision support system model for a number of different learning schemes that can be used for practical data mining which is vital for the growth of any individual as well as society. Effective data mining strategies, if implemented in software engineering, can be a great help for the Software development Industry as well as gives a practical exposure to researchers. Data mining is a technique of finding useful and interesting patterns from a large amount of data in order to take accurate and precise business decisions., if implemented on software requirements, initial investigation and software design of any software project, the mining algorithms will generate different patterns that can be used for prediction of effort estimation, staffing, Size estimation, development time estimation etc. Furthermore, the research may prove to be crucial for education institutions and professional organizations of any kind.

Keywords: *Software Engineering, Project Management, Data Mining, Decision Support system, Software Project Management, Software development.*

1. Introduction

Software development activities are inherently complex and difficult to conceptualize. The intricacy, caused by various dependencies and rigid programming paradigms, slows the development process and maintenance activities, leads to faults and deficiencies and ultimately increases the cost of software. Many software development organizations develop some sort of processes for the management of various development activities. [1]

To improve software efficiency and quality, software engineers are increasingly applying data mining algorithms to various software engineering tasks. The development of any Software largely depends on the software engineering principles. Software

engineering can be best defined as ‘A discipline whose core objective is to produce the quality product that satisfies user requirement with in time’.

To satisfy the above definition a clear & correct understanding of different estimations (Size, Effort, Development, Average Staff Size, Productivity and Cost) is essential. This will help software professionals to decide on various factors affecting development process. This definitely makes the process faster than the manual one and provides the developers extra time to concentrate on other technical processes.

Data mining is a discipline which uses various analytical tools to extract patterns and information from huge datasets. In business now a days, we have a very large numbers of datasets. We are really efficient at storing data than extracting knowledge from it, mainly the precise and meaningful information needed to create good software. This paper focuses on how to bridge a gap between data mining and software engineering and what data mining techniques can be applied to software engineering

2. Literature Survey

SE is a backbone for any Business today, and study on its problems has proposed various schemes for improvement. Thayer al. [2] introduced planning problems, organizing problems, staffing problems, and controlling problems as major challenges in this area. Ramamoorthy et al. [3] stated that as more complex software applications are required, programmers will fall further behind the demand. This causes the development of poor quality software and higher maintenance costs. The problems stated

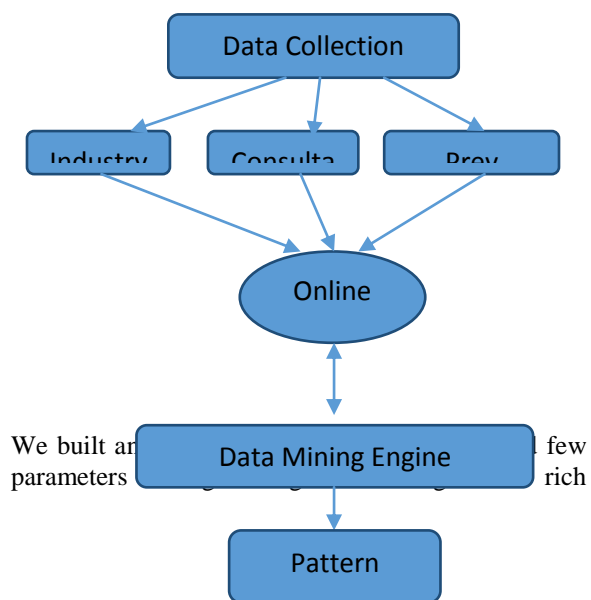
IJESPR

www.ijesonline.com

by Thayer are more closely tied to the processes of SE project management, while those introduced by Ramamoorthy are more closely tied to the limitations of human beings [4]. Later work by Clarke [5] identified challenges in SE associated with the complexity of the software development process. Similar to our study, he stated that the complexity of software development causes the software to become harder to maintain. Data perturbation techniques for preserving privacy in data mining were proposed by Islam and Brankovic [6]. About proposed the clustering technique to identify patterns in the underlying data [7]. Later work by Ma and Chan [8] suggested iterative mining for mining overlapping patterns in noisy data.

3. Methodology

Decision support system [9][10] is a class of computer software systems that helps manager to take routine and non routine decisions so that the organization can grow and business can be flourished. It uses the components of information technology to develop the system like user interface that can accept the data, processing unit for data processing and output unit that provides information to managers. For the use of data mining techniques in software engineering, we use a descriptive methodology.



descriptions and explanations for the study purpose. We collect data from various resources (approached companies, consultants, previously published papers and many other data repositories) to populate our online repository designed for the purpose of study. After identifying the current need of the business, we analyze which technique of data mining is best and then extract the interesting patterns from the data which will be useful for the present business scenario.

4. Proposed Model

According to the proposed model the initial requirements for the new projects would be taken through the graphical user interface and converted into an acceptable format by the connectors, then these inputs would be feeded into a data mining engine. The knowledgebase is used to initiate and restrict the patterns searching. Then these patterns with the help of data mining [11] will further be evaluated and converted into a format which is understandable by the user.

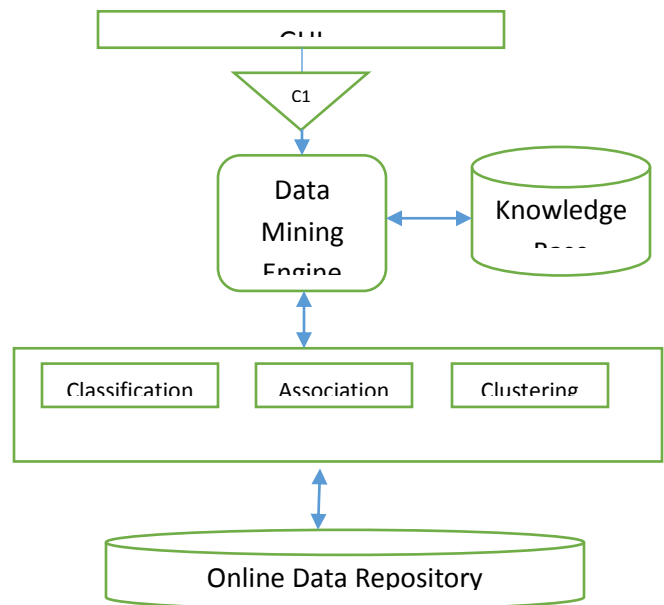


Figure 2: Model

Here the requirements of the new project will be feeded and the above process will provide solutions for the following questions like.

1. Which software models were used in the past for such projects?
2. What is the success rates for each models used in the past by different companies?
3. What would be the average staff size of the team responsible for development?
4. What other resources are required?
5. What cost/budget would be involved for such type of projects?
6. What are the different risks involved?
7. What should be the ideal testing strategy?
8. What implementation method should be adopted for successful execution of the project?

The model bridges the gap between software engineering process and data mining technology in order to take better decisions by providing the answers to the above stated questions and provide software engineers a new insight by using data mining approach

5. Conclusion

Predicting the above parameters would give the companies a deep insight of the overall environment required for developing such a software and would decrease the time and risk of the failure of the project and hence increase the profits and Data mining can effectively use to solve complex software engineering problems.

References

- [1] Quinn Taylor and Christophe Giraud-Carrier, "Applications of data mining in software engineering", *Int. J. Data Analysis Techniques and Strategies*, Vol. 2, No. 3, 2010
- [2] R. H. Thayer, A. Pyster, and R. C. Wood, "Validating solutions to major problems in software engineering project management," *IEEE Computer Society*, pp. 65-77, 1982.
- [3] C. V. Ramamoorthy, A. Prakash, W. T. Tsai, and Y. Usuda, "Software engineering: problems and perspectives," *IEEE Computer Society*, pp. 191-209, Oct 1984.
- [4] Husain et al., "Application of Data Mining Techniques for Improving Software Engineering", *ICIT 2011 The 5th International Conference on Information Technology*
- [5] J. Clarke et al., "Reformulating software engineer as a search problem," *IEEE Proceeding Software.*, Vol. 150, No. 3, pp. 161-175, June 2003.
- [6] M. Z. Islam and L. Brankovic, "Detective: a decision tree based categorical value clustering and perturbation technique for preserving privacy in data mining," *Third IEEE Conference on Industrial Informatics (INDIN)*, pp. 701-708, 2005.
- [7] M. Aouf, L. Lyanage, and S. Hansen, "Critical review of data mining techniques for gene expression analysis," *International Conference on Information and Automation for Sustainability (ICIAFS) 2008* , pp. 367- 371, 2008
- [8] P. C. H. Ma and K. C. C. Chan, "An iterative data mining approach for mining overlapping coexpression patterns in noisy gene expression data," *IEEE Trans. NanoBioscience*, Vol. 8 No. 3, pp. 252-258, Sept 2009.
- [9] Kobus Ehlers et al, "A Decision Support System for Institutional Research Management in Higher Education Data Mining to determine Research Focus, Intensity and Synergy " *IEEE International Conference on Computational Science and Engineering 2009*
- [10] Marina Segura et al, "Decision support systems for forest management: A comparative analysis and assessment " *Computers and Electronics in Agriculture 101 (2014) 55-67*, Elsevier.
- [11] Iparrukh Saleem, "Implementation of Data Mining Approach for Building Automated Decision Support Systems", *International Conference on Information Society (i-Society 2012)*