

Classification Technique for Scrum Analysis in Data Mining

Amanpreet Kaur¹, Amanpreet Kaur Bath²

¹Research Scholar of M.Tech, Department of Computer Science and Engineering, Global Institute of Management & Emerging Technologies, Amritsar, Punjab, India

²Assistant Professor, Department of Computer Science and Engineering, Global Institute of Management & Emerging Technologies, Amritsar, Punjab, India

Abstract— The data mining is the technique which is applied to extract the useful information from the rough data. The scrum analysis is the technique of data mining which is applied to predict the consequences according to the data. In this work, the teacher performance is predicted from the students results. The k-mean clustering is applied which will cluster similar and dissimilar type of data. The algorithm of SVM is applied for the scrum analysis. In the proposed work, the SVM is replaced with the knn classifier for the scrum analysis. The performance of SVM and Knn classifier is compared in terms of accuracy, execution time.

Keywords— K-means, scrum, SVM, KNN, Prediction.

I. INTRODUCTION

Data mining (now and again called data or knowledge discovery) is the way toward analyzing data from alternate points of view and summarizing it into helpful information that can be utilized to build revenue, cuts costs, or both. Data mining software is one of various analytical tools for analyzing data [1]. Cluster examination or clustering is the task of grouping a set of objects in a manner that objects in the same group (called a cluster) are more comparable (in some sense or another) to each other than to those in different groups (clusters). It is a primary task of exploratory data mining, and a typical technique for statistical data investigation, utilized as a part of many fields. Classification comprises of predicting a specific result based on a given input. Keeping in mind the end goal to predict the result, the algorithm forms a preparation set containing a set of attributes and the respective result, ordinarily called goal or prediction attribute [2]. The algorithm tries to find relationships between the attributes that would make it conceivable to predict the result. Next the algorithm is given a data set not seen some time recently, called prediction set, which contains the same set of attributes, except for the prediction attribute, not yet known. The algorithm investigations the input and produces a prediction. The prediction accuracy defines how "good" the algorithm is [3]. The supervised learning is a method that is based on the classified data which is generated from the set of data which has the exact classes known before. There are various scenarios that involve the classification mechanism. There are various decision problems arising in day to day lives with the growth in technology [4]. The tree main categories involved within this technology are statistical, machine learning and neural network. There are various goals as per the technologies.

SCRUM is a management methodology, created with a specific end goal to improve and maintain an existing system or a production model [5]. This methodology assumes the existence of a project and some source code sequences, which quite often exist in the object-oriented software development because of class libraries. SCRUM is not addressing to development efforts for totally new or legacy systems. In the SCRUM terminology, a SPRINT is an arrangement of development exercises which are embraced during a pre-decided timeframe, more often than not from one to four weeks. The interval depends on the complexity of the product, on the risk assessments and on the required degree of skills and expertise [6]. The speed and the intensity of a SPRINT are dictated by its agreed duration. The risk is assessed persistently and permanently, and adequate measures are gone for each risk event. The Scrum framework is build upon three pillars: roles, events and artifacts. Utilizing those three pillars will enable one to understand how Scrum works. This section explains the different roles available in Scrum, to be specific the team, the ScrumMaster, the product owner, the customer, the management and the users [7]. One can differ between roles that belong to the project team (the team itself obviously and the ScrumMaster), roles that influence the project team (the product owner) and more external roles (the customer, the management and the user).

II. LITERATURE REVIEW

Raoul Vallon, et al., (2013) proposed in this paper [8] a case study that investigates an agile approach from a real world project involving two unaffiliated IT organizations that collaborate in a distributed development environment. The objective of this case study is to find root causes of failure of integration of two different organizations in distributed development environment. An external observer is hired to participate in scrum meetings, observe the problems and the last phase of presentation and discussion of problems with team members. The result of this case study suggests formation of single-site self-organizing teams instead of multi-sites.

Sebastian Hanschke, et al., (2010) focused in this paper [9], to answer two main questions whether and how agile methods such as scrum can be used to create architecture deliverables. Both the questions in this paper are answered using two different procedure models, in which one to be used on TOGAF's enterprise strategic and segment architecture

level and other on capability architecture level. TOGAF ADM is split up into four different scrum projects, each of them consist of various components. Overall, the interviewers considered it promising to apply Scrum to areas to see, whether the integration, especially the collaboration between EA and ASD is going to work in practice.

Gangjun Yang, et al., (2010) proposed in this paper [10], the method of agile industrial design, and builds the agility model with three major factors of element group, lightweight process, and flexible methods. The method is adapted to the conceptual design process filled with uncertainty, particularly in creativity and demand. Because of agile method emphasizes the communication of implicit learning and application of flexible method, the controllability and stability of teamwork design appears poor. The rapid design reconstruction and vigor of agile industrial design is the further study task.

Julio Ariel, et al., (2010) developed in this paper [11], AVISPA, a tool for restricting error patterns in software process models determined with EPF. In this paper the Scrum process model is broken down with the AVISPA tool and it is found that the specification that is broadly utilized by the community has been not entirely defined. This may explain, in any event to some extent, the gap between the expected and the reported performance of Scrum. As a major aspect of this work the capacity to prove its helpfulness for process engineers as methods for supporting their work mostly when the process evolves is achieved.

Hycinta Andrat, et al., (2015) proposed in this paper [12], that risk management is a vital part in project development which helps in eliminating the risks prevailing in software project. Despite the fact that risk management is associated with agile approach, still there prevails the scope for improvement. Along these lines the motto of this paper is to investigate agile methodology, Scrum and highlight the degree to which risk is overseen in it and propose a model to overcome its limitations in risk analysis stage. To overcome the limitation, agile risk arranges model is proposed as an alternative approach. It aids better comprehension of the impact of each identified risk.

Zulkarnain Azham, et al., (2011) proposed in this paper [13], keeping in mind the end goal to deliver the product speedier, a transformation to their customary software development lifecycle to agile development method which can empower them towards quick delivery of software adapting to the requirements-change phenomenon. With a specific end goal to address the aforementioned issue, this paper proposes the integration of security principles in development phases utilizing scrum. The result of the proposed arrangement will be presented soon after enough data has been collected from different studies, meetings and experiments that are in progress and planned.

III. RESEARCH METHODOLOGY

In this work, data regarding placement of the students based on some criteria on which student’s selection is done is being collected. That involves data based on the selection of students without giving any inputs for placements i.e. raw data of final year students who are eligible for placements based on

certain criteria. Sectioning of students based on this is done by using k-mean clustering for generating clusters of similar and dissimilar type of data. Neural network classifier is used that consists of units arranged in layers, which convert an input vector into some output. In this each unit takes an input, applies a function to it and then passes the output to the next layer. Neural network is used as it is relevant for applying scrum practices i.e. the result one iteration at the end of sprint is used as input in the next iteration in next sprint for generation of final outcome at the end of sprint. The method is proposed without affecting the current work and comparing the result with and without using scrum practices.

In this work, the existing scrum analysis system will be improved using the knn classifier. In the existing system the following steps are followed:-

1. The dataset will be taken as input for the classification.
2. In the second step, the technique of k-mean clustering is applied which will cluster the similar and dissimilar type of data.
3. In the final step, the technique of SVM classifier will be applied which will classify the similar and dissimilar data into two classed.

In the proposed system the improvement in the existing system will be proposed in which following steps are applied.

1. The dataset will be taken as input for the classification.
2. In the first step, the technique of k-mean clustering will be applied which will cluster the similar and dissimilar type of data.
3. In the last step, the technique of knn classifier will be applied which will classify the similar and dissimilar type of data.

IV. EXPERIMENTAL RESULTS

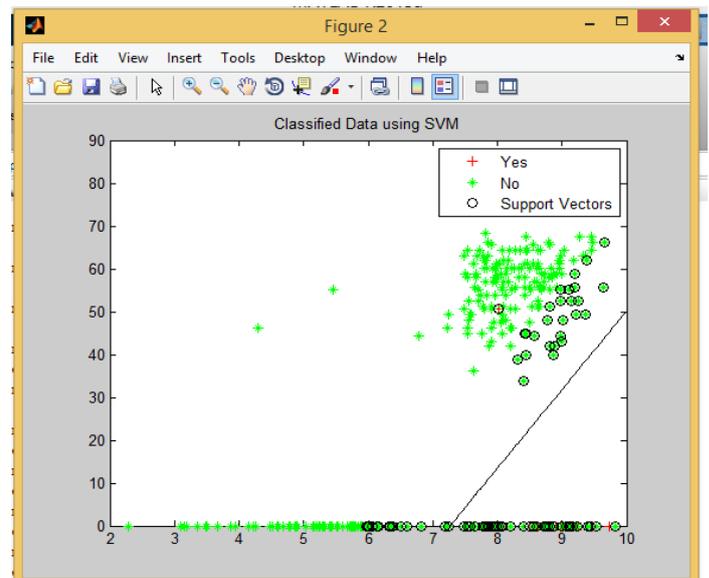


Fig. 1. SVM classification.

As shown in the figure 1, the k-mean clustering is applied which will cluster the similar and dissimilar type of data. The technique SVM classification is applied which will classify the similar and dissimilar type of data.

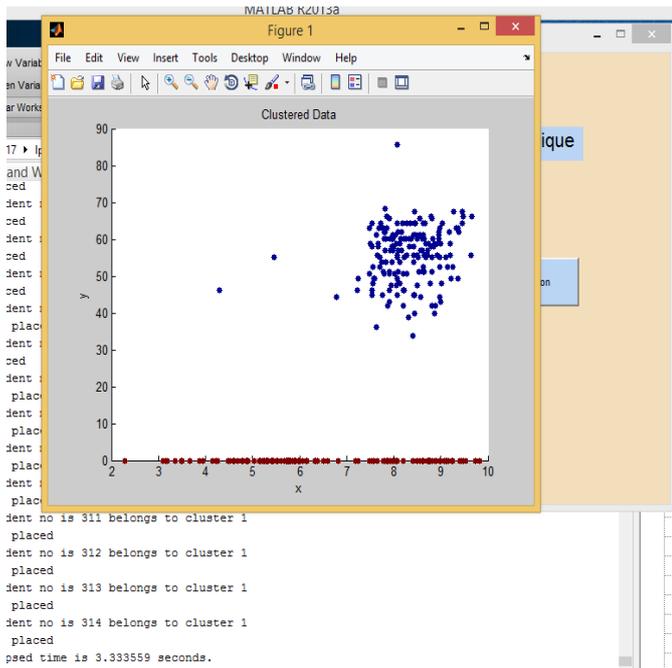


Fig. 2. Clustering of data.

As shown in the figure 2, the technique of k-mean is applied which will cluster similar and dissimilar type of data. The two type of data is shown with two different colors in the figure.

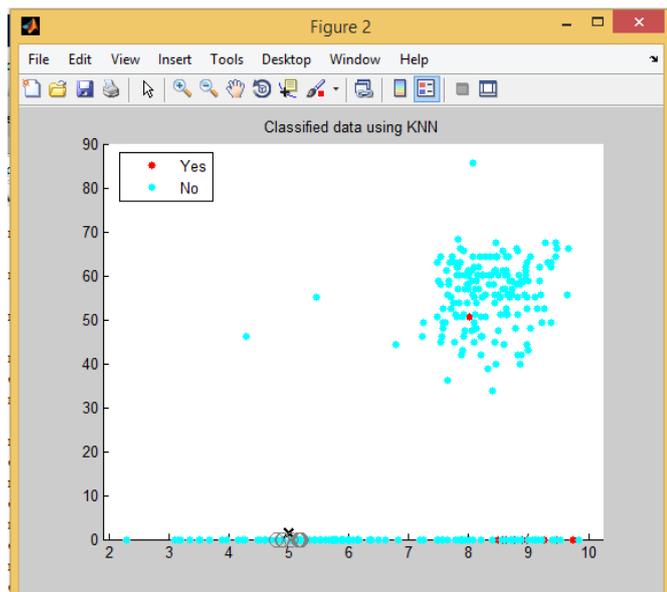


Fig. 3. Knn classification.

As shown in figure 3, the clustered data is taken as input for this step and technique knn is applied which will cluster similar and dissimilar type of data.

V. CONCLUSION

In this work, it is been concluded that scrum analysis is the technique of prediction analysis which is applied to predict teacher performance on the basis of student marks. The k-mean algorithm is applied which will cluster similar and dissimilar type of data. The SVM classifier is applied on the cluster data for the scrum analysis. In the proposed work, the SVM classifier is replaced with KNN classifier which leads to increase in accuracy of scrum analysis. In future the proposed algorithm can be compared with other scrum analysis techniques.

REFERENCES

- [1] Sundar.C, M. Chitradevi, and Dr. G. Geetharamani, "Classification of Cardiotocogram data using neural network based machine learning technique," *International Journal of Computer Applications*, vol. 47, no. 14, pp. 19-25, 2012.
- [2] Smitha .T, and V. Sundaram, "Comparative study of data mining algorithms for high dimensional data analysis," *International Journal of Advances in Engineering & Technology (IJAET)*, vol. 4, issue 2, pp. 173-178 2012.
- [3] K Priya, R. Geetha Ramani, and S. Gracia Jacob, "Data mining techniques for automatic recognition of Carnatic Raga Swaram notes," *International Journal of Computer Applications*, vol. 52, no. 10, pp. 4-10, 2012.
- [4] M. Elallaoui, K. Nafil, and R. Touahni, "Automatic generation of UML sequence diagrams from user stories in Scrum process," *IEEE 10th International Conference on Intelligent Systems: Theories and Applications (SITA)*, 2015.
- [5] J. Sutherland, A. Viktorov, J. Blount, and N. Puntikov, "Distributed scrum: Agile project management with outsourced development teams," *40th Hawaii International Conference on System Sciences*, 2007.
- [6] M. Mahalakshmi, and Dr. M. Sundararajan, "Tracking the student's performance in web-based education using scrum methodology," *IEEE International Conference on Computing and Communications Technologies (ICCCCT)*, 2015.
- [7] G. K. Hanssen, B. Haugset, T. Stålhane, T. Myklebust, and I. Kulbrandstad, "Quality assurance in scrum applied to safety critical software," *LNBIP* 251, pp. 92-103, 2016.
- [8] R. Vallon, S. Strobl, M. Bernhart, and T. Grechenig, "Inter-organizational co-development with Scrum: Experiences and lessons learned from a distributed corporate development environment," *Agile Processes in Software Engineering and Extreme Programming*, pp. 150-164, 2013.
- [9] S. Hanschke, J. Ernsting, and H. Kuchen, "Integrating agile software development and enterprise architecture management," *Proceedings of the Annual Hawaii International Conference on System Sciences*, pp. 4099-4108, 2015.
- [10] G. Yang, S. Yu, G. Chen, and I. Chu, "Agile industrial design management based on Serum," *IEEE 11th International Conference on Computer-Aided Industrial Design & Conceptual Design (CAIDCD)*, 2010.
- [11] J. Ariel H. Alegria, M. Cecilia Bastarrica, and A. Bergel, "Analyzing the Scrum process model with AVISPA," *IEEE International Conference of the Chilean Computer Science Society*, 2010.
- [12] H. Andrat, and S. Jaswal, "An alternative approach for risk assessment in Scrum," *Intl. Conference on Computing and Network Communications (CoCoNet)*, 2015.
- [13] Z. Azham, I. Ghani, and N. Ithnin, "Security backlog in Scrum security practices," *5th Malaysian Conference in Software Engineering (MySEC)*, 2011.