

A Survey: Machine learning approach used in Industry for bug prediction

Nidhi Srivastava¹ Manisha Agarwal² Tripti Lamba³

¹ Research Scholar, Banasthali Vidyapith

²Associate Professor, Banasthali Vidyapith,

³Associate Professor, IITM GGSIPU,

Artificial Intelligence (AI) is one of the important subjects in today's scenario. AI is been widely used in our daily life in one or another way. ML is a part of AI or we can say that ML is one of the techniques of AI from which we can evaluate different training algorithms which always helps us for better results by taking the right decision with correct results. ML always focuses on having filtered information that can be extracted from the data of different dimensions. ML is also called as the dimension of AI as ML always try that the system starts learning and also improve its functionality with the real-time experiences without human intervention. Machine learning is a technical branch of AI and it also helps in finding the output by extracting the knowledge from the observations without any regressed programming skills required. The machine learning approach is used in the industry to get better results. The main objective of the paper was to study the machine learning approach adopted by the industry for software bug prediction [11]. The data set used by the software industries for maintenance and bug prediction was static because the activities were performed manually in the organizations. No doubt by changing the approach from static to the machine, will provide a faster and accurate software defect prediction but accepting this new change was a challenge for employees [1]. There were few industries like Volvo, Ericson, and ABB Robotics were initially working on the traditional method[2]. The Ericsson Connected Vehicle Cloud gives Volvo Cars the power to control and manage this growing ecosystem. Our industrialized Ericsson Connected Vehicle Cloud platform, provided as a managed service, enables Volvo Cars to swiftly deliver services at scale, to more cars and further More ABB Robotics is a Robotics industry which deals with ABB is a leading industry of robot software and equipment, industrial robots and complete application solutions and these all industries believe in working with their traditional system. The Ericsson and Volvo Car Group developed a framework for adopting machine learning. They adopted the approach of machine learning and analysis was done by taking different parameters and evaluated that the machine learning approach helps in finding the best outcomes to get software accuracy, efficiency, and reduction of the error rate and also the cost of the software. A Study on ABB Robotics declared that the current state of practice in software development seems to be unrealistic. The main hurdle of defect prediction is the failure to identify new patterns and low generalizability of the machine learning algorithm. The study concluded that it is desirable to adopt the machine learning approach for software bug prediction as it increases the accuracy and reduces the software cost further best machine learning model can be implemented by some feature selection techniques like Boruta and F Selector with this we can find out the important variable and then on important variables the implementation of the different algorithm is done to get the highest accuracy and minimum error in software bug prediction.

Keywords: Machine Learning, Bug Prediction, AI, Software Accuracy

I. Introduction

Machine learning is fast growing technology but still due to some factors, industries are not able to adapt machine learning technology. In the paper the case study on different industries are discussed and author tried to find out the major reasons which is affecting the approach of implementing machine learning in the industry and gave the factors need to be revisited[3]. At the initial level it was difficult to find the factors which are important for the adoption of machine learning algorithm[2], so interviews were conducted among selected staff of the industry like in **Volvo** (an automobile industry), team leaders were selected who were responsible for all the project related activities, managers were selected who were responsible for testing and managing the data of the industry same with the **Ericsson** which is a telecom domain[7], their manager and team leaders were also interviewed to get the data whereas in the case of **ABB Robotics**[1,2] which mainly focus on developing embedded software with electronic systems, team of designer/ developer and a manger was taken to get interviewed.

This paper has the following sections. The literature review is presented in Section 2. The related study is presented in Section 3. The adoption of machine learning in industry is discussed in Section 4. The paper is concluded in section 5 and section 6 discussed about the future work that can be done further in the paper and at the end references are included.

II. Literature Review

Machine learning is an approach that helps in improving accuracy and reduces the error rate of the software from the available data and facts. In industry adoption of machine learning approach is very difficult. MacDonald et al. [1] given a report from the study on ABB Robotics and the work was done on both the end that is at academia and company both. This report proved that by implementing machine driven approach will also increase in the productivity effectiveness and efficiency of the software. The whole study is done with the help of creating research questions which includes the requirement of an industry, factors affecting the industry if the machine learning is adopted by them and which type of machine learning model is most suitable to the industry.

[2] This paper discuss the case study done on Ericsson and AAB Robotic by creating two categories of people from both the companies which will help in taking the decision that weather machine driven approach should be implemented in the industry and then the main focus was made on modeling UML so that the productivity of the industry will also increase these groups consist of architect , manager and technical coordinator from ericsson as it is the company of telecommunication and they selected manger and developer/designer from ABB Robotic as it is a software development company.

In this paper the all [3] discussion was on bug detection in machine learning used in an industry they worked on Volvo and Ericson, they also took some hypothesis which discussed about the positive and negative effects of adopting the techniques of machine learning[12].

III. Related study / Survey

There are many studies done on the adoption of machine learning techniques in industry and these are done on the basis of many hypotheses, questionnaire, interviews etc [6]. Table 1 shows the discuss changes before and after adoption of machine learning approach. Table 1 shows that all factors i.e. accuracy, software cost, data comparison, Tools, Data acquisition and maintenance cost all give good results as compare to the present state of practice[8]. This is also shown in the graphical representation in Figure 1 where Red represents machine learning technology and Blue represents present state of practice.

Factors	Present state of practice	Machine Learning implementation
Accuracy	4	5
Software Cost	2	4
Data Comparison	0	1
Tools	2	4
Data acquisition	1	1
Maintenance Cost	3	5

Table 1 : Description about the comparison study among the system (0=NO,1=YES,2-5 less to high)

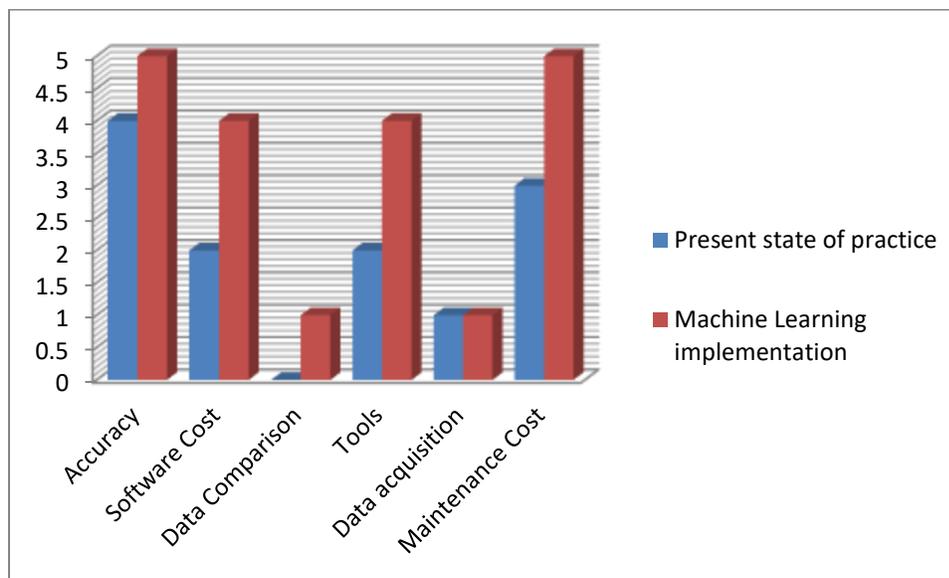


Figure 1 : Graph Representation of comparison study.

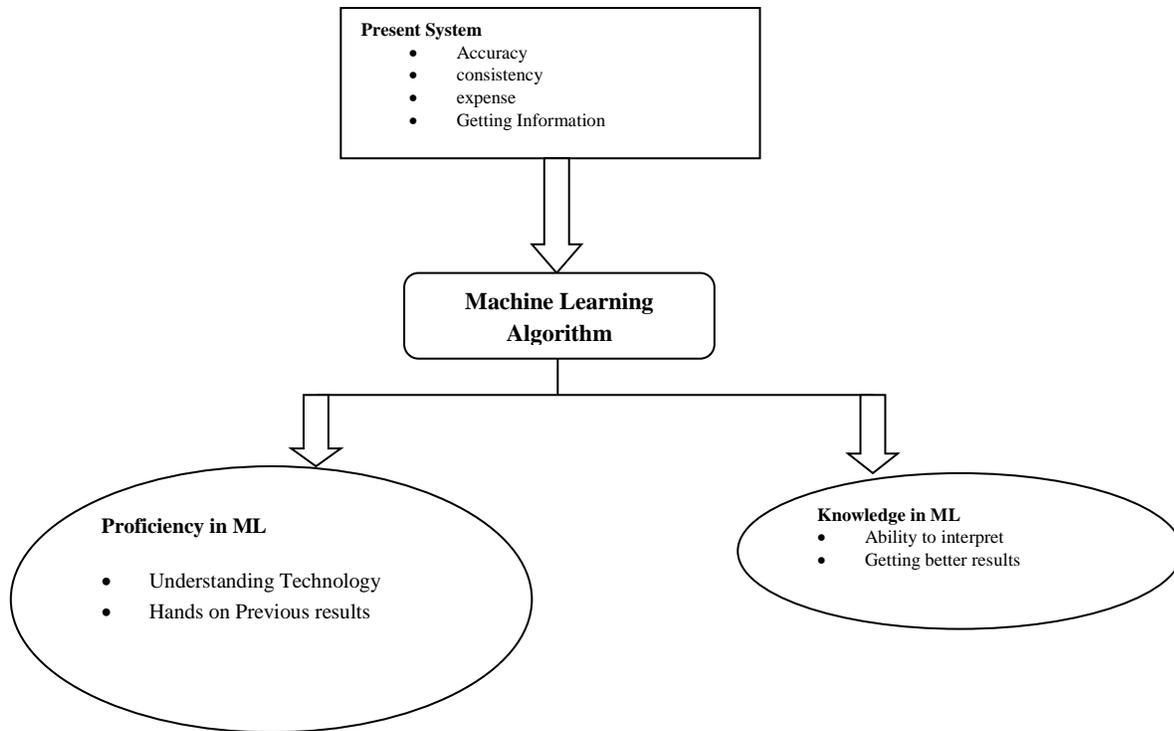


Figure 2: Overview of the attribute of ML & Present system

Industries in present state are having many attribute like accuracy, consistency, expense and getting information from many recourses [4].where as if we implement machine learning to the industry which is shown in Figure 2 they give two main attributes which are proficiency and knowledge and through this we get better results with better ability to interpret the data[9].

IV. Adoption of machine learning in Industry

Some studies were done to adopt the machine learning algorithm in companies [5]. Machine learning can be adopted by the industry only when it include all the dimensions, these dimensions can be **External Environment** which include all the competitors etc.[10] , second and most important dimension was to know **about the industry** itself where machine learning is to be implemented which include challenges face to accept the machine learning approach, what are the satisfaction level of employee while working with present system, how much the employees already familiar with the machine learning approach and lastly what are the **benefits of using the machine learning** approach to the industry[7]. The study indicated that there were some dimensions which were important factor that needs to be taken care and that can be many attributes but we cannot ignore the size of the industries on which we are implementing the machine learning[6]. Majorly the questions were divided into three categories they were discussed in the Table 2 given below:

S No.	Categories	Sub-Categories
A.	Industry requirement for adopting machine learning.	<ul style="list-style-type: none"> • Cost Estimation. • Quality improvement. • Increase Product Development Communication. • Keeping track of software artifacts

B.	Search for the factor for adopting machine learning.	<ul style="list-style-type: none"> • Software adopting Tools. • Software adopting and developing Cost. • Employee knowledge regarding the software • Maintaining the quality of software.
C.	What would be the best conditions for implementing machine learning in the industry?	<ul style="list-style-type: none"> • It was suggested to implement UML as per the requirement of data. • It is preferred to use notations which are easy to understand but at the initial stage then gradually go for advance notations.

Table 2 : Discussing about the major categories and sub category for adopting machine learning technique

V. Conclusion

This paper discussed case studies which were done on different industries, where some factors were found to adopt the machine learning model in the industry [9]. Many round of the interview was held among manager, team leaders and developer of the industry and then some new dimensions were found which includes how industry is presently working, what all changes comes when the industry start working with machine learning and what all factors can be implemented into the industry.

Though all the factors cannot be implemented simultaneously but few of the factors can be taken which will give accurate and correct result

VI. Future work

In the current paper we discussed about the industries where machine learning approach is being implemented. We have seen that these industries at some places accepted the approach of machine learning where as at some places they become reluctant to accept the approach of machine learning but somehow the cost of maintenance of software, accuracy of results produced maintenance of the software and other factors are having better results as compares to the existing system [11, 12,13]. For future work, it is being recommended to propose a model from the comparative study done in the paper

VII. References

1. MacDonald, A., Russell, D., & Atchison, B. (2005, March). Model-driven development within a legacy system: an industry experience report. In 2005 Australian Software Engineering Conference (pp. 14-22). IEEE.
2. Staron, M. (2006, October). Adopting model driven software development in industry—a case study at two companies. In International Conference on Model Driven Engineering Languages and Systems (pp. 57-72). Springer, Berlin, Heidelberg.
3. Rana, R., Staron, M., Hansson, J., Nilsson, M., & Meding, W. (2014, August). A framework for adoption of machine learning in industry for software defect prediction.

- In 2014 9th International Conference on Software Engineering and Applications (ICSOFT-EA) (pp. 383-392). IEEE.
4. Rahman, M. S., Rivera, E., Khomh, F., Guéhéneuc, Y. G., & Lehnert, B. (2019). Machine learning software engineering in practice: An industrial case study. arXiv preprint arXiv:1906.07154.
 5. Lamba, T., & Mishra, A. K. (2019). Optimal machine learning model for software defect prediction. *International Journal of Intelligent Systems and Applications*, 11(2), 36.
 6. Lamba, T., & Mishra, A. K. (2017). Optimal Metrics Selection for Software Defect Prediction. *International Journal of Data Mining And Emerging Technologies*, 7(2), 82-91.
 7. Khanna, A., Goyal, R., Verma, M., & Joshi, D. (2018, February). Intelligent Traffic Management System for Smart Cities. In *International Conference on Futuristic Trends in Network and Communication Technologies* (pp. 152-164). Springer, Singapore.
 8. Garg, D., Goel, P., Kandaswamy, G., Ganatra, A., & Kotecha, K. (2018, July). A Roadmap to Deep Learning: A State-of-the-Art Step Towards Machine Learning. In *International Conference on Advanced Informatics for Computing Research* (pp. 160-170). Springer, Singapore.
 9. Sharma, N., & Singh, A. (2018, July). Diabetes Detection and Prediction Using Machine Learning/IoT: A Survey. In *International Conference on Advanced Informatics for Computing Research* (pp. 471-479). Springer, Singapore.
 10. Kumar, Y., & Singh, P. K. (2018). Improved cat swarm optimization algorithm for solving global optimization problems and its application to clustering. *Applied Intelligence*, 48(9), 2681-2697.
 11. Suresh, Y., Pati, J., & Rath, S. K. (2012). Effectiveness of software metrics for object-oriented system. *Procedia technology*, 6, 420-427.
 12. M. B. Kursu, "Feature Selection with the Boruta Package," vol. 36, no. 11, 2010. P. Narayana, "Software defect prevention – in a nutshell," 2016. [Online]. Available:<https://www.isixsigma.com/industries/software-it/software-defect-prevention-nutshell/>. Accessed: Oct 22, 2016.
 13. Aggarwal, K. K., Singh, Y., Kaur, A., & Malhotra, R. (2006). Empirical Study of Object-Oriented Metrics. *Journal of Object Technology*, 5(8), 149-173.