

## Enhanced Software Defect Prediction By Support Vector Machine With Radian Basis Function Kernel

Jasvir Kaur  
Student  
Rayat Bahra University  
Punjab, India  
[Kaurjasvir214@gmail.com](mailto:Kaurjasvir214@gmail.com)

Ekta Gupta  
Assistant Professor  
Rayat Bahra University  
Punjab, India  
[83ekta@gmail.com](mailto:83ekta@gmail.com)

**Abstract-** It is valuable to predict the software that is defect-prone. There have been many studies and learning approaches that are used to measure the performance of software. A meta-analysis of all relevant, high quality primary studies of defect prediction are used to determine that what factors influence predictive performance. The quality of the software can be measured with the different features such as cyclomatic complexity, design complexity, effort, time estimator, length of the program, operands, operators, line count etc. we can use to SVM-RBF kernel technique to improve accuracy complexity and function recalling.

**Keywords-** About four key words or phrases in alphabetical order, separated by commas. (Mention 4-5 keywords)

### I. INTRODUCTION

#### 1.1 Software Engineering: Introduction

Software Engineering is defined as the systematic and well defined approach to the development, operation, maintenance and retirement of the software. By the word 'systematic' means that the methodologies used for the development of the software are repeatable. The goal of software

engineering is to take software development closer to science and engineering that solves the problems of the clients and away from those approaches for development whose outcomes are not predictable.

### SATISFIES

#### 1.1.1 Software Quality Attributes

Software Quality attributes can be defined as follows:

- Functionality
- Reliability
- Usability
- Efficiency
- Maintainability
- Portability

1. **Functionality:** It is the capability that provides functions which meets the defined and implied needs of the software when it is used.

2. **Reliability:** It is the capability that maintains the specified level of performance.

3. **Usability:** The capability to be understood, learned and used.

4. **Efficiency:** It is the capability to measure the performance relative to the amount of resources used.

5. **Maintainability:** It is the capability to be updated and modified for purposes of making corrections, improvements or adaptation.

6. **Portability:** It is the capability to be adapted for different environments without applying actions.

## 1.2 Defect Prediction in Software module

### a) Data cleansing process

- Initial Preprocessing of the data
- Removal of Constant attributes
- Removal of Repeated attributes
- Replacement of Missing Values
- Enforce Integrity with Domain Specific Expertise

- Removal of Repeated and Inconsistent Instances

### b) Prediction Performance measures

- Precision
- Recall
- Accuracy

### c) Data Extraction

- Classifier family
- Data set family
- Metric family
- Researcher Group

## 1.3 Machine learning

Machine learning is a science that explores the building and study of algorithms that can learn from the data. Machine learning process is the union of statistics and artificial intelligence and is closely related to computational statistics. Machine learning takes decisions based on the qualities of the studied data using statistics and adding more advanced artificial intelligence heuristics and algorithms to achieve its goals.

Machine learning tasks are classified into three broad categories:

- Supervised learning.

- Unsupervised learning
- Reinforcement learning

### 1.4 Data mining

Data mining is related with the discovery of new and interesting patterns from large data sets for analysis and executive decision making. Data mining is described as the union of historical and recent developments in statistics, artificial intelligence and machine learning. Data mining and machine learning are used together to study data and find previously-hidden trends or patterns within.

#### 1.3.1 Scope of Data Mining

- Automation in prediction of behavior and trends
- Automated discovery of previously unknown patterns

#### 1.3.2 Data Mining Process

Data mining consists of five major elements:

- Extraction and transformation of data onto the data warehouse system.
- Run data on multidimensional database system in a managed way

- Providing data access to business analysts and other professionals
  - Data analyzing
  - Presentation of data in useful and required formats such as tables and graphs.
- Knowledge

Interpretation Evaluation

### 3.3 Classification Algorithms

- 1 Statistical Algorithms
- 2 Neural Networks
- 3 Genetic algorithm
- 4 Decision trees
- 5 Nearest neighbor method
- 6 Rule induction

---

#### Future Scope

The quality of the software can be measured with the different features such as cyclomatic complexity, design complexity, effort, time estimator, length of the program, operands, operators, line count etc. we can use to SVM-RBF kernel technique to improve accuracy complexity and function recalling. We can reduce the complexity, improve efficiency, improve function recalling.

## REFERENCES

- [1] Gray, David, et al. "The misuse of the NASA metrics data program data sets for automated software defect prediction." *Evaluation & Assessment in Software Engineering (EASE 2011), 15th Annual Conference on*. IET, 2011.
- [2] Menzies, Tim, Butcher, A., Marcus, A., Zimmermann, T., & Cok, D. "Local vs. global models for effort estimation and defect prediction." *Proceedings of the 2011 26th IEEE/ACM International Conference on Automated Software Engineering*. IEEE Computer Society, 2011
- [3] Song, QinbaoJia, Z., Sheppard, M., Ying, S., & Liu, J. "A general software defect-proneness prediction framework." *Software Engineering, IEEE Transactions on* 37.3 (2011): 356-370.
- [4] Li, M., Zhang, H., Wu, R., & Zhou, Z. H. "Sample-based software defect prediction with active and semi-supervised learning." *Automated Software Engineering* 19.2 (2012): 201-230.
- [5] Okutan, Ahmet and Olcay Taner Yıldız. "Software defect prediction using Bayesian networks." *Empirical Software Engineering* 19.1 (2014): 154-181.
- [6] Sheppard, Martin, David Bowes, and Tracy Hall. "Researcher bias: The use of machine learning in software defect prediction." *Software Engineering, IEEE Transactions on* 40.6 (2014): 603-616.
- [7] Czibula, Gabriela, Susanna Marian, and Is van Greely Czibula. "Software defect prediction using relational association rule mining." *Information Sciences* 264 (2014): 260-278.
- [8]<http://rspublication.com/ijst/ijst%20pdf%20feb%2012/18.pdf>
- [9] Data mining  
[http://en.wikipedia.org/wiki/Data\\_mining](http://en.wikipedia.org/wiki/Data_mining)
- [10] Data mining Process,  
<http://www.google.co.in/images>.
- [11]  
<http://www.unc.edu/~xluan/258/datamining.html>
- [12]  
[http://www.tutorialspoint.com/data\\_mining/dm\\_dti.htm](http://www.tutorialspoint.com/data_mining/dm_dti.htm).