A NOVEL MACHINE LEARNING METHOD FOR SOFTWARE DEFECT ESTIMATION

MANINDRA **1** , VINAY **2** , RAGHURAM **3**

Mr.ARUN4  , DR. CHANDRAN5 ,DR.G.VICTO SUDHA GEORGE6

1,2 UG Student, Department of Computer Science and Engineering,

Dr. MGR Educational and Research Institute, Maduravoyal, Chennai 600095, TN, India

3,4,5 Professor, Department of Computer Science and Engineering,

Dr. MGR Educational and Research Institute, Maduravoyal, Chennai 600095, TN, India

[1manindragundeboina9@gmail.com](mailto:1manindragundeboina9@gmail.com)

[2vinaynaiducse@gmail.com](mailto:2vinaynaiducse@gmail.com)

[3raghuramreddygantla143@gmail.com](mailto:3raghuramreddygantla143@gmail.com)

[4 arun.cse@drmgrdu.ac.in@gmail.com](mailto:4%20arun.cse@drmgrdu.ac.in@gmail.com)

[5chandran.mech@drmgrdu.ac.in](mailto:5chandran.mech@drmgrdu.ac.in%20%20%20)

[6victosudhageorge@drmgrdu.ac.in](mailto:6victosudhageorge@drmgrdu.ac.in%20)

**ABSTRACT:**

Software failure prediction plays a crucial role in improving software quality and reducing testing time and costs. Machine learning has emerged as a powerful tool in this domain, offering adaptive capabilities that refine predictions based on past outcomes. By leveraging machine learning, previously undiscovered patterns and insights can be extracted from complex datasets, enhancing human decision-making and improving defect prediction accuracy. Effective software failure prediction requires robust data preprocessing techniques such as feature selection and noise removal to eliminate irrelevant and erroneous data. Metrics, including code complexity, defect density, and change frequency, are key indicators that guide predictive models. These metrics help identify defective modules, ensuring better software reliability.

This study evaluates existing software failure prediction methods and presents a novel machine learning approach that outperforms traditional models like Support Vector Machines (SVM), Naïve Bayes, and Decision Trees. The proposed method demonstrates improved prediction accuracy and reduced computational time, achieving 69.8% accuracy in just 3.24 minutes, showcasing its efficiency as a superior solution for defect estimation.

**Keywords**: Software defect prediction, Software Metrics, Machine learning techniques, SVM, Naive Bayes and Decision Tree, Data Cleaning, Predication Accuracy, Dataset Optimization