SIMATS ENGINEERING
Engineer to Excel
TECH STAR SUMMIT 2024
Name: Mr. Yalaka Nikhil Reddy
Register Number: 192211188
Guided by E.K.Subramanian

Machine Learning based Software Fault Prediction in Software Efficiency using Artificial Neural Network Comparing with Support Vector Machines for Improved Accuracy

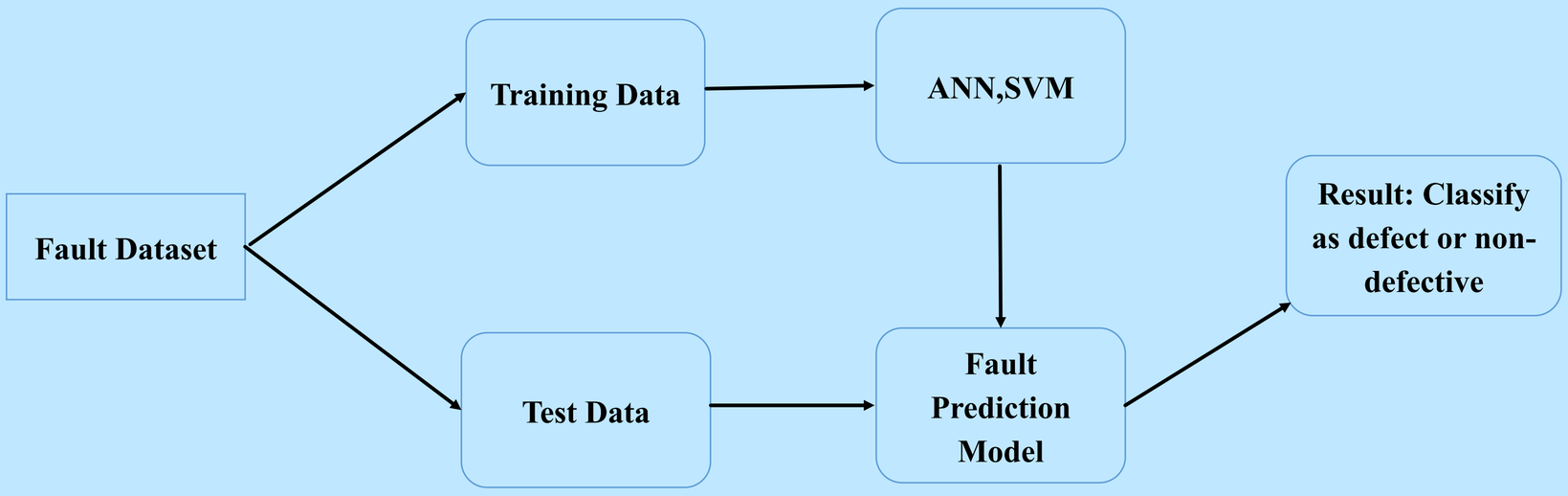
INTRODUCTION

- The aim of this study is to investigate the efficacy of machine learning algorithms, specifically Artificial Neural Networks (ANNs) and Support Vector Machines (SVMs), in predicting software faults to enhance software efficiency.
- Predicting software faults is critical in software development as it helps in identifying potential issues early in the development lifecycle, thereby reducing maintenance costs, enhancing system reliability, and ensuring user satisfaction.
- The application of machine learning techniques, particularly ANNs and SVMs, in software fault prediction offers a data-driven approach to analyze and predict potential software failures. By leveraging historical data, these algorithms can learn patterns and characteristics of faulty code, enabling developers to proactively address issues.
- ANNs are known for their ability to learn complex patterns in data through layers of interconnected neurons, while SVMs are effective in classifying data by finding the hyperplane that maximizes the margin between different classes. By comparing the performance of these algorithms, we aim to determine which approach yields higher accuracy in software fault prediction.
- The study utilizes a comprehensive dataset comprising various software metrics and fault occurrences collected from real-world software projects. Software fault prediction dataset in kaggle includes metrics like code lines, complexity, defect density, enabling robust model training.

Software fault prediction

MATERIALS AND METHODS

- The programming lab of the Saveetha Institute of Medical and Technical Sciences was the site of this investigation. There are two groups. There are two versions of software fault prediction; one is ANN and the other is SVM. The sample size was determined using Clinal. Test statistic with a 0.05 threshold, 80% power, and 95% confidence interval the optimal sample size.
- Intel Core i7 processor, 512 GB of RAM, Google Collab, Jupiter Notebooks, and Windows 11th generation. In order to determine accuracy, we conducted statistical analyses using IBM SPSS.



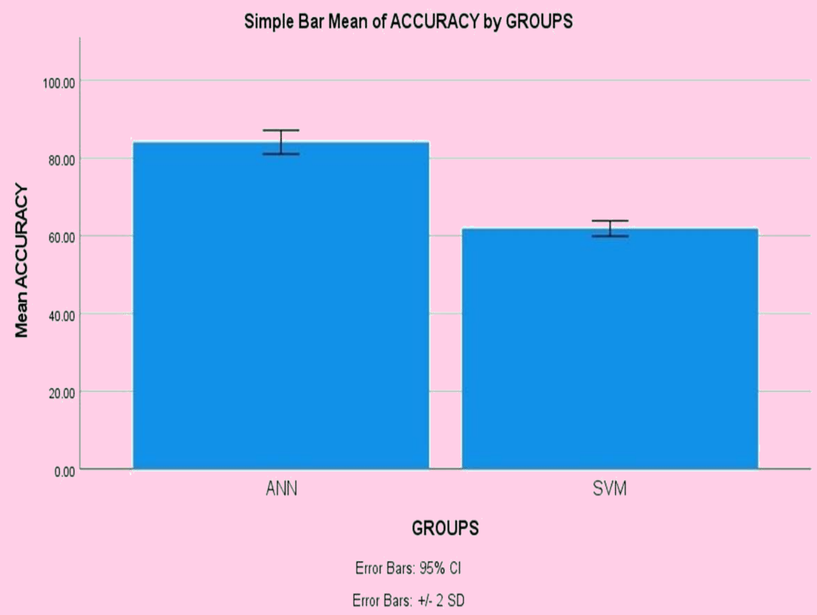
Software Fault Prediction

RESULTS

- Algorithm for software fault prediction using Artificial Neural Networks The T-tests for independent samples The following table shows the difference in significance, mean, and standard error between the ANN and SVM approaches, as determined by SPSS analysis ($p = 0.001$) ($p<0.05$).
- The results of the ANN and SVM algorithms were input into SPSS in order to find the average, dispersion, and error for the two networks. ANN attained 84.10% accuracy, with SVM and 62.50% also reaching their goals.
- Using a 95% confidence interval and a standard deviation of +/- 2 as parameters, SPSS calculated an accuracy of 84.10% for the suggested method ANN and 61.90% for the SVM.

Statistical Analysis of the difference in significance, mean, and standard error between the Artificial Neural Network and Support Vector Machines approaches

	Algorithm	N	Mean	Std. Deviation	Std. Error Mean
Accuracy	ANN	10	84.10	1.52388	0.48189
	SVM	10	61.90	0.99443	0.31447



Shows the comparison between Artificial Neural Network and Support Vector Machine with accuracy of 84.10% and 61.90%.

DISCUSSION AND CONCLUSION

- The T-test statistical analysis revealed a significant difference ($p=0.001$, independent sample T-test $p<0.05$) between the performance of Artificial Neural Network (ANN) and Support Vector Machines (SVM) in software fault prediction.
- This statistical validation strengthens the reliability of the findings and emphasizes the significance of the comparative analysis. The overall accuracy of Artificial Neural Networks Algorithm was 84.10%, whereas the Support Vector Machine (SVM) achieved an accuracy of 61.90%.This notable difference in accuracy highlights the efficacy of the Artificial Neural Network Algorithm in predicting software faults compared to SVM.
- The study concludes that the Artificial Neural Network algorithm demonstrates higher accuracy in software fault prediction when compared with Support Vector Machines.
- Both ANN and SVM contribute to improved accuracy, enabling stakeholders to anticipate and mitigate potential software faults, thus enhancing software efficiency and reliability . Future research could explore fine-tuning the algorithms for specific software development contexts or .
- Limitations such as dataset quality, feature selection, and algorithmic complexity should be considered in future studies to ensure the robustness of the predictive models.
- Future research efforts should focus on refining algorithms and addressing limitations to advance predictive accuracy in real-world software development scenarios.

BIBLIOGRAPHY

- Ali, Liaqat, Awais Niamat, Javed Ali Khan, Noorbakhsh Amiri Golilarz, Xiong Xingzhong, Adeeb Noor, Redhwan Nour, and Syed Ahmad Chan Bukhari. n.d. “An Optimized Stacked Support Vector Machines Based Expert System for the Effective Prediction of Heart Failure.” Accessed March 18, 2024. http://dx.doi.org/10.1109/ACCESS.2019.2909969.
- “Machine Learning Based Methods for Software Fault Prediction: A Survey.” 2021. Expert Systems with Applications 172 (June): 114595.https://doi.org/10.1016/j.eswa.2021.114595.
- Chakravarty, Krishna, and Jagannath Singh. n.d. “Optimizing Defect Removal Efficiency by Defect Prediction Using Machine Learning.” Accessed March 18, 2024. http://dx.doi.org/10.1109/OCIT56763.2022.00047
- Mahmood, Yasir, Nazri Kama, Azri Azmi, Ahmad Salman Khan, and Mazlan Ali. 2022. “Software Effort Estimation Accuracy Prediction of Machine Learning Techniques: A Systematic Performance Evaluation.” Software: Practice & Experience 52 (1): 39–65. https://doi.org/10.1002/spe.3009