International journal of basic and applied research

www.pragatipublication.com

ISSN 2249-3352 (P) 2278-0505 (E)

Cosmos Impact Factor-5.86

A Novel Machine Learning Method for Software Defect Estimation

Ms. A. Prasmitha[1], Gopasani Sai Mahitha[2], Syed Jabida[3],

Billu Akshitha[4], Nijam Raghu[5]

#1Associate Professor in Department of CSE,in PBR VITS,KAVALI.

#2#3#4#5 B.Tech with Specialization of Computer Science and Engineering in PBR Visvodaya Institute of Technology & Science , Kavali.

Abstract_ Recent improvements in software defect prediction (SDP) involve the combination of different classification algorithms to form an ensemble or hybrid approach. This technique was developed to increase prediction performance by overcoming the constraints of individual classification techniques. This study conducts a systematic literature review on the application of the ensemble learning approach to software defect prediction. The review is performed after thoroughly reviewing research papers released since 2012 in four well-known online libraries: ACM, IEEE, SpringerLink, and ScienceDirect. This paper addresses five research problems related to the use of ensemble learning for software defect prediction. To extract the answers to the identified queries, 46 most relevant papers are picked after a rigorous systematic research procedure.In this research, we use ensemble machine learning algorithms including Random Forest, Logistic Regression, and Linear egression to forecast software problems. Software flaws constantly present the issue of extending development time and investing additional money.

1.INTRODUCTION

With the continual expansion of software scale and the enhancement of function, the complexityofsoftware is increasing, and the possibility of software defects is increasing, which is more likelytolead to software failures [1]. The most common

International journal of basic and applied research

www.pragatipublication.com

ISSN 2249-3352 (P) 2278-0505 (E)

Cosmos Impact Factor-5.86

software quality assurance activity is softwaretesting, which can effectively check software errors, but also the most time-consumingandresource-consuming stage of the software development life cycle. Therefore, the concept of softwaredefect prediction arises at the historic moment. Its purpose is to predict software defects as soonaspossible according to the basic characteristics of software, allocate test resources reasonably accordingto the prediction results, give test priority, shorten development cycle, and reduce software cost[2]. At present, machine learning has been widely used in the field of software defect prediction[3]. From the existing research results, the defect prediction effect of single machine learning model is not

ideal[4]. A single model either over-fits or lacks generalization ability. To solve this problem, wepropose an ensemble learning model. Ensemble learning is a machine learning paradigmthat isapplicable to both supervised learning and unsupervised learning. We use k-nearest neighbour, logisticregression, random forest, and artificial neural network to build our ensemble leaning model andmakean experiment.r.

2.LITERATURE SURVEY

2.1 S. Parnerkar, A. V. Jain, and C. Birchha, ‘‘An approach to efficient software bug prediction using regression analysis and neural networks,’’ Int. J. Innov. Res. Comput. Commun. Eng., vol. 3, no. 10, Oct. 2015.

Machine Learning approaches are good in solving problems that have less information. In most cases, the software domain problems characterize as a process of learning that depend on the various circumstances and changes accordingly. A predictive model is constructed by using machine learning approaches and classified them into defective and non-defective modules. Machine learning techniques help developers to retrieve useful information after the classification and enable them to analyse data from different perspectives. Machine learning techniques are proven to be useful in terms of software bug prediction. This study

International journal of basic and applied research

www.pragatipublication.com

ISSN 2249-3352 (P) 2278-0505 (E)

Cosmos Impact Factor-5.86

used public available data sets of software modules and provides comparative performance analysis of different machine learning techniques for software bug prediction. Results showed most of the machine learning methods performed well on software bug datasets. The advancement in software technology causes an increase in the number of software products, and their maintenance has become a challenging task. More than half of the life cycle cost for a software system includes maintenance activities.

2.2 B. Liu, H. Qin, Y. Gong, W. Ge, M. Xia, and L. Shi, ‘‘EERA-ASR: An energy-efficient reconfigurable architecture for automatic speech recognition with hybrid DNN and approximate computing,’’ IEEE Access, vol. 6, pp. 52227–52237, 2018.

This paper proposes a hybrid deep neural network (DNN) for automatic speech recognition and an energy-efficient reconfigurable architecture with approximate computing for accelerating the DNN. To accelerate the hybrid DNN and reduce the energy consumption, we propose a digital–analog mixed reconfigurable architecture with approximate computing units, including a binary weight network accelerator with analog multi-chain delay-addition units for bit-wise approximate computing and a recurrent neural network accelerator with approximate multiplication units for different calculation accuracy requirements. Implemented under TSMC 28nm HPC+ process technology, the proposed architecture can achieve the energy efficiency of 163.8TOPS/W for 20 keywords recognition and 3.3TOPS/W for common speech recognition. Deep Neural Networks (DNNs) that have many hidden layers have been proven to outperform traditional models (i.e., Markov models, Gaussian mixture models) on a variety of speech recognition benchmarks by a large margin [1], [2].

2.3 N. Cummins, S. Amiriparian, G. Hagerer, A. Batliner, S. Steidl, and B. W. Schuller, ‘‘An image-based deep spectrum feature representation for the recognition of emotional speech,’’ in Proc. 25th ACM Multimedia Conf. (MM), 2017, pp. 478–484.

International journal of basic and applied research

www.pragatipublication.com

ISSN 2249-3352 (P) 2278-0505 (E)

Cosmos Impact Factor-5.86

The outputs of the higher layers of deep pre-trained convolutional neural networks (CNNs) have consistently been shown to provide a rich representation of an image for use in recognition tasks. This study explores the suitability of such an approach for speech-based emotion recognition tasks. First, we detail a new acoustic feature representation, denoted as deep spectrum features, derived from feeding spectrograms through a very deep image classification CNN and forming a feature vector from the activations of the last fully connected layer. We then compare the performance of our novel features with standardised brute-force and bag-of-audio-words (BoAW) acoustic feature representations for 2- and 5-class speech-based emotion recognition in clean, noisy and denoised conditions. The presented results show that image-based approaches are a promising avenue of research for speech-based recognition tasks. Key results indicate that deep-spectrum features are comparable in performance with the other tested acoustic feature representations in matched for noise type train-test conditions; however, the BoAW paradigm is better suited to cross-noise-type train-test conditions. Convolutional neural networks (CNNs) have become increasingly popular in machine learning research.

4.PROPOSED SYSTEM

In the proposed approach, ensemble learning is used to combine the strengths of different classifiers to improve defect detection in the dataset. Over the past decade, numerous studies have found that ensemble approaches outperform individual classifiers in terms of accuracy..

3.1 IMPLEMENTATION

Gathering the datasets: We gather all the r data from the kaggale website and upload to the proposed model

Generate Train & Test Model: We have to preprocess the gathered data and then we have to split the data into two parts training data with 80% and test data with 20%

Run Algorithms: For prediction apply the machine learning models on the dataset by splitting the datasets in to 70 to 80 % of training with these models and 30 t0 20 % of testing for predicting

Obtain the accuracy: In this module we will get accuracies

Predict output: in this module, we will get output based input data

International journal of basic and applied research

www.pragatipublication.com

ISSN 2249-3352 (P) 2278-0505 (E)

Cosmos Impact Factor-5.86

4.RESULTS AND DISCUSSION

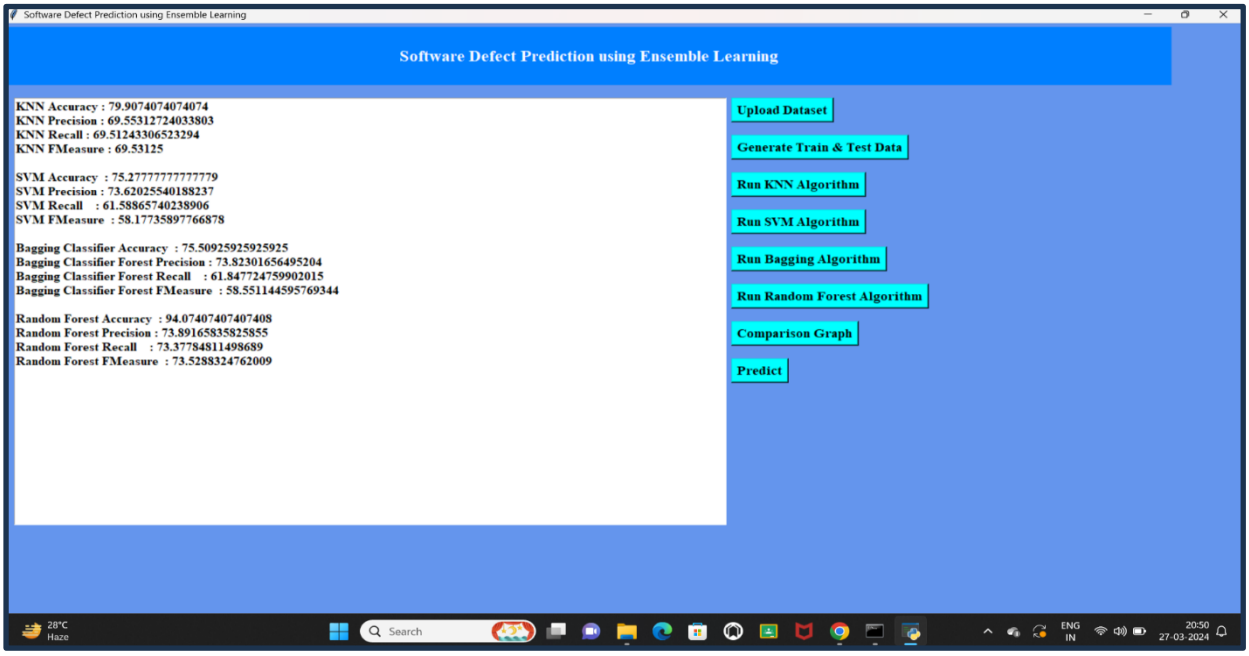

Fig 1: Performing Algorithms on the data:

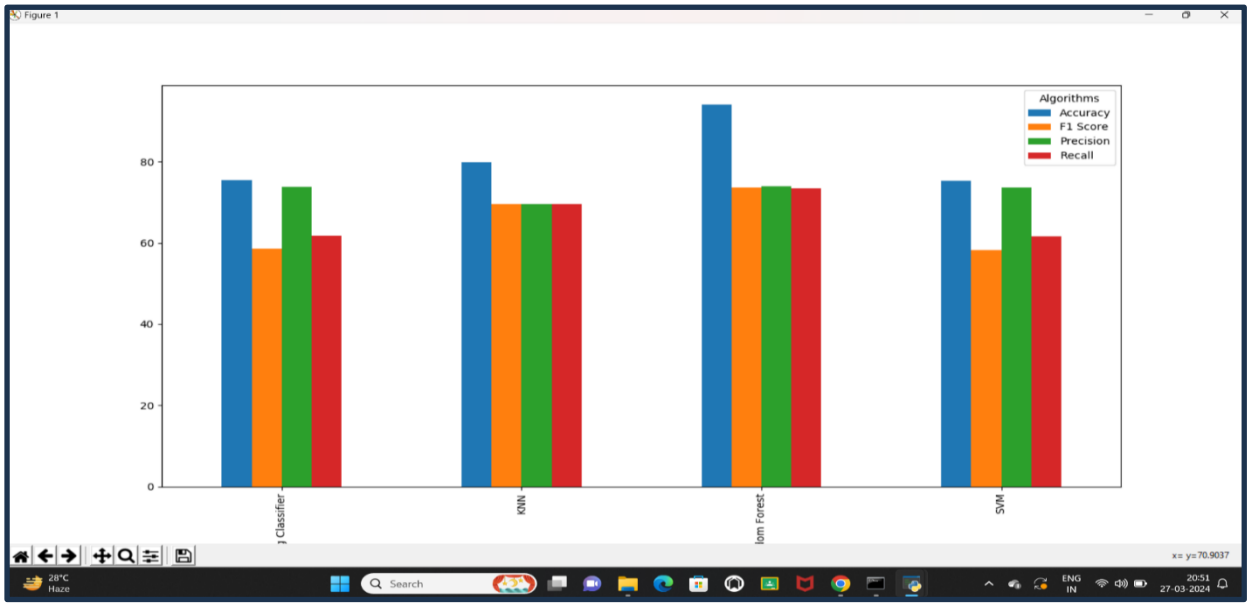

Fig 2: Graphical representation of the prediction:

International journal of basic and applied research

www.pragatipublication.com

ISSN 2249-3352 (P) 2278-0505 (E)

Cosmos Impact Factor-5.86

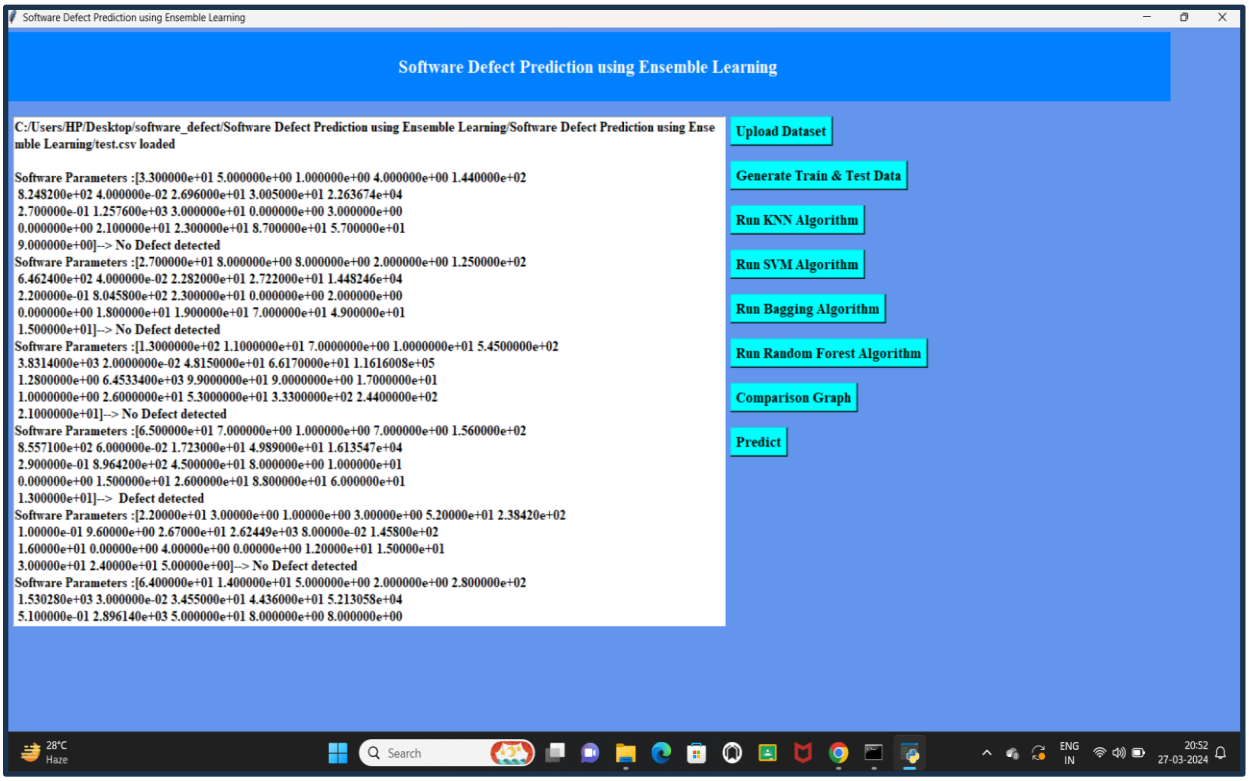

Fig 3: Prediction Screenshot

5.CONCLUSION

This study employs an SLR to track the most recent research breakthroughs in ensemble learning approaches for software defect prediction. This review is carried out by systematically evaluating the most significant research papers published in three well-known online libraries: ACM, IEEE, Springer Link, and Science Direct. This paper outlines and discusses five research issues concerning the various aspects of research advancement in the use of ensemble learning techniques for software defect prediction. It is concluded that ensemble learning procedures outperform individual classifiers. In the future, the effects of feature selection algorithms on ensemble learning should be investigated.

REFERENCES

[1] Assim M, Obeidat Q and Hammad M 2020 Software Defects Prediction using MachineLearning Algorithms International

International journal of basic and applied research

www.pragatipublication.com

ISSN 2249-3352 (P) 2278-0505 (E)

Cosmos Impact Factor-5.86

Conference on Data Analytics for Business and Industry:Way Towards a Sustainable Economy (ICDABI) (IEEE) pp 1-6

[2] Thota M K, Shajin F H and Rajesh P 2020 Survey on software defect prediction techniquesJ. International Journal of Applied Science and Engineering 17(4)

[3] Geanderson E, Eduardo F, Adriano V, Markos V and Nivio Z 2020 Understanding machinelearning software defect predictions J. Automated Software Engineering 27(3-4)

[4] Matloob F, Aftab S and Iqbal A 2019 A Framework for Software Defect PredictionUsingFeature Selection and Ensemble Learning Techniques J. International Journal of ModernEducation and Computer Science (IJMECS) 11(12)

[5] Yao W 2020 Research on Static Software Defect Prediction Algorithm Based on BigDataTechnology J. International Conference on Virtual Reality and Intelligent Systems (ICVRIS)(IEEE) pp 610-613

[6] Jiang P C 2021 Research on software defect prediction technology based on deep learning2ndInternational Conference on Computing and Data Science (CDS) pp 104-107

[7] Kussul N, Lavreniuk M, Skakun S and Shelestov A 2017 Deep Learning Classification of LandCover and Crop Types Using Remote Sensing Data J. IEEE Geoscience and Remote SensingLetters 14(5) pp 778-782

[8] Grazia C, Donato I, Vincenzo D, et al. 2021 Human Gait Analysis in NeurodegenerativeDiseases: a Review J. IEEE journal of biomedical and health informatics p 1.

Author's Profiles

Ms. A. Prasmitha working as Associate Professor in Department of CSE, PBR Visvodaya Institute of Technology and Science KAVALI.

International journal of basic and applied research

www.pragatipublication.com

ISSN 2249-3352 (P) 2278-0505 (E)

Cosmos Impact Factor-5.86

Team Members

Gopasani Sai Mahitha B.Tech with Specialization of Computer Science and Engineering in Visvodaya Engineering College, Kavali.

Syed Jabida B.Tech with Specialization of Computer Science and Engineering in Visvodaya Engineering College, Kavali.

Billu Akshitha B.Tech with Specialization of Computer Science and Engineering in Visvodaya Engineering College, Kavali.

Nijam Raghu B.Tech with Specialization of Computer Science and Engineering in Visvodaya Engineering College, Kavali.