NIRMA UNIVERSITY JOURNAL OF ENGINEERING AND TECHNOLOGY VOL. 1 ISSUE 1

# Machine Learning Algorithms for Fault-Prediction

Sima Gonsai[A], Usha Mehta[B], *Senior Member, IEEE*

***Abstract*—With the advancement in VLSI technology, the number of transistors on a device increases along with the reduction in the size of transistor. The likelihood of a manufacturing failure is rises as feature size continue to contract. The overall testing cost and testing efforts are increases exponentially with each new technology node. Therefore, it is necessary to explore the techniques which guarantee the circuit functioning with less efforts and cost. As with each new technology node, not only the possible number of faults in circuit increases but also new types of faults are being introduced. In this scenario, this paper aims to explore the various existing Machine Learning (ML) methods for the prediction of number of faults in circuit. This paper also aims to categorize the fault prediction and prediction of test vector set. The paper includes the comparison analysis of different ML algorithms in fault prediction. With the use of ML algorithm, the Automatic Test Pattern generator (ATPG) shortens the time needed to generate test set required for manufacturing testing.**

***Index Terms*— Fault model, Fault equivalence, Fault Prediction, Stuck-at-Faults, Machine Learning, Test Data, Test Pattern Generation, Automatic Test Pattern Generator, Automatic Test Equipment.**

## I. INTRODUCTION

With shrinking technology, the number of transistors in a given integrated circuit is increasing exponentially. The fabrication at such small geometry is becoming very much complicated. So it is very natural that the probability of fabrication defects is increasing immensely for any fabricated IC. Such defects are generally modeled as fault. Further the advancement of transistor technology is coming up with new types of faults and hence new fault models. Hence, despite of drastic increase in number of transistors per chip, semiconductor industry has shown the reduction in the cost of manufacturing per transistor. But over the years, the cost of testing per transistor is not much reduced comparatively [13]. Among many parameters affecting test cost, key parameter is test data generation.

First A. Author is with the Institute of Technology, Nirma University, Ahmedabad, India. (e-mail: 23ptphde213@nirmauni.ac.in).

Second B. Author is with the Institute of Technology, Nirma University, Ahmedabad, India. (e-mail: usha.mehta@nirmauni.ac.in).

Every fabricated IC is being tested by Automatic Test Equipment (ATE). Speed of chip testing directly proportional to the number of test patterns generated by ATE. The test data set is being generated by Automatic Test Pattern Generator (ATPG). When the circuit is made of billions of transistors, the ATPG requires very complex algorithms to predict the possible fault types in circuit, to predict all possible faults in the circuit and to decide the test data set for this fault list. Hence, it is necessary to adopt the new methodologies to support the automation of fault prediction and test data generation. Further, with increase in number of transistors per chip within a small area, peripheral pin accessibility is also becoming a major concern of design process. To efficiently implement the pin-allocation, for pin-access prediction and optimization, speedy and effective algorithms are required [1].

With improved technology and with development of various machine learning algorithms, Machine Learning (ML) applications are becoming more and more powerful. Different learning algorithms are used to program machine learning models. These algorithms can learn from a given set of data. This data can be in any form such as integers, strings, images, videos, audio, etc. For example, voice acknowledgment frameworks, Siri and Cortana relies on machine learning and profound neural systems like Deep Neural Network (DNN) to mimic human communication. In general, to increase ML model accuracy, datasets should be as large as possible. To support massive data processing, high computational and processing power is needed. These massive computations are now more affordable because the amount of processing power available per dollar has likely increased by a factor of ten over the course of the last 25 years at regular intervals. Colab is a colaboratory suit from Google allows user to write, edit and execute the machine learning based python code with GPU access. It is web-based IDE released in 2017. As a result, various applications can use ML techniques with various algorithms. As transistors scales down to smaller size, the reliability of chips is extremely important [2].

In this paper, we have explored various familiar ML algorithms for the fault prediction of combinational circuits. We have used Synopsis and Cadence EDA tools for fault coverage and test set generation. These test pattern generation tools assure the fault coverage of the circuit or a chip. For the proposed ML approach these tools were used to generate the data set.

The section paper is arranged as follow. Section II describes proposed methodology. Section III and IV cover result

analysis and challenges respectively. Future scope is discussed in section V and section VI includes conclusion.

## II. Proposed Methodology

### A. Problem Statement

The main objective of the work is to do comparative analysis of fault prediction and find the best possible algorithm for predicting faults.

### B. Proposed Algorithms for prediction

To predict the faults using the dataset for given combinational circuits, various ML algorithm models are implemented. The ML models are categorized into two types:

1) Supervised Learning
2) Unsupervised Learning

They are used in various methods that are important for solving many real time problems. We have opted supervised learning for the prediction of faults, as training of the ML model with known example circuits is necessary.

In supervised machine learning, the inputs and outputs are given to the machine learning model and the model will learn by mapping the activation function. Supervised learning algorithms have been implemented in this research. Identifying the type of problem is the first step of ML implementation. In supervised learning typically; a problem can be of classification or to predict a target numeric value (Regression problem).

### C. Dataset

For any combinational circuit, the types of fault and number of fault depends upon the technology node, number of inputs, number of outputs, total gates needed in the circuit, total number of interconnections. Here for first proof of concept, we have considered the combinational circuit at gate level netlist. We have also considered that the circuit is containing the primitive gates only like AND, NAND, OR, NOR and NOT gates. Each gate is having 2 fan-in only. For this experimental work, the most popular fault models stuck-at-one and stuck-at-zero are considered.

For ML algorithms, it was required to have large number of training data set. For this purpose, more than 250 different combinational circuits with different complexity were considered. The data set has 60 tuples which include the number of gates, the number of inputs, number of outputs and number of faults after applying fault equivalence. The dataset consists of different features which have varied purpose and applicability. The parameters like Number of Gates, Number of Inputs, and the number of faults for analysis have been used.

### D. Fault Prediction

The advancement in scaling techniques of a device on silicon chip in highly dense systems and complex circuits can now be designed by Very Large Scale Integration. [3]. The need for fault prediction in today's world is necessary with the unprecedented growth scene in VLSI domain. The major concerns in chip testing are detection of the fault and identifying possible fault location. In large circuits, and fault prediction [4]. The likelihood of a manufacturing failure rises as feature size continues to reduce. Therefore, it is crucial to guarantee a circuit behaves correctly. As a result, the circuit has to be tested for faults. The suggested ML method offers a quicker method for estimating the number of possible faults in given circuit. The accurate prediction of possible faults in circuit can be helpful in future prediction of test pattern generation with a short execution time of ATPG algorithm. As a result, it shortens the time needed to confirm a circuit's proper operation. The prediction of the number of faults is done by giving minimal information about the circuit. With the basic information like the total gates required in the circuit and the total externally applied inputs, we are training the ML models to predict the number of faults in a circuit [5].

### E. Methodology

Seven machine learning algorithms have been implemented for the prediction of the faults in different combinational circuits.

1) *K-Nearest Neighbor:* The K-Nearest-Neighbors classification model is easy to implement yet efficient technique. [6]. KNN is a famous versatile supervised machine learning model which is especially applied to the domain where prior knowledge about circuits and their fault is unavailable. This algorithm is chosen because it performs better on large datasets in which values may change frequently. It has been observed that accuracy in the KNN model does not increase after grouping rather we see that F1-score in KNN model increasesby0.016.

2) *Decision tree Classifier:* In this algorithm data is classified like tree structure, where each node represents classification based on particular feature based on feature's value or range of vales, each branch from the parent node indicates a result of the test on a feature value and leaf node shows particular class label. A decision tree algorithm is a kind of supervised technique that functions using a hierarchical approach. The decision tree functions as a structure like the top down or bottom up. It applies branching to create a branch for every potential class. One branch classifies tuples in to children nodes and so on until the tuples in a particular node belongs to the same class. We have observed that neither the accuracy nor the f1-score improves after grouping. The rules for selection of particular node as decision node specifically depend on attribute's Information gain or gain ratio or Gini index value. Therefore, it is very essential to select a proper attribute wisely as it affects the performance of the model. [7].

3) *Support Vector Classifier:* SVM is a supervised machine learning method. It is an effective technique for classifying and analyzing data. Learning rate of SVM classifier is faster in large dataset compare to other classifier [8]. This algorithm separates different classes by forming line, plane or hyper plane depending on number of features selected for classification. These hyper planes in

the identify space to have a good separation between different classes. The significant margins between all classes are desirable to reduce loss in prediction. It is a kind of machine learning model which is used for both regression and classification type problems. The objective of the SVM method is to choose a hyper plane in an N-dimensional space that categorizes the data points without over fitting. The number of input features defines the type of hyper plane. If two input features are selected for the model training, then the hyper plane will be basically a line and if three input features are selected for the model then hyper plane becomes a 2-D plane. If more than three significant features are selected for the model, plane becomes multi dimensional and visualizing it becomes difficult. F1-score for the SVC models increases by 0.027 after grouping.

4) *Linear Regression:* In a linear regression model, output variable is predicted by plotting a line based on the value of independent variable. The dependent variable is the target variable predicted using one or more independent variable (input features) [9]. This algorithm works well for regression problem where continuous variable values to be determine, but due to its inability to differentiate among classes, it performs poorly for classification problem. So it is not recommended to use it for classification problems. The accuracy of the Linear Regression increases greatly after grouping, it increases by 45 percent, the F1-score also increases with smaller value.

5) *Naive Bayes Classifiers:* Naive Bayes algorithms are the statistical classification algorithms based on the Bayes' theorem help in identifying the probability of events, having knowledge of the other event occurrence. It is depends on the probabilities of occurrence of each event [10].The Bayes Theorem is used for deriving a naïve Bayes algorithm. The algorithm assumes that all the (input) independent features are mutually exclusive. In this classifier, it believes that if one feature is present in a class is unaffected by rest features presence. They are mutually exclusive with each other. The dataset is then classified depending on probabilities of all the class, which is known as posterior probability. The class with the highest posterior probability will be considered as the class label prediction. For applications with smaller available data set , this model predicts results faster and effectively than other Machine learning algorithm such as logistic regression.

6) *Logistic Regression:* This algorithm is a type of supervised machine learning model used for classification problems. It is used to identify the best possible class of the particular tuple. The nature of the dependant variable (output variable) is dichotomous. So, there would be two possible classes only. The accuracy of logistic regression model also increases by 16 percent (compare to which result) but we see that the F1-score decreases after grouping.

7) *Random Forest Classifier:* Random Forest comprises of big set of small decision trees, which act as an ensemble. Every single decision tree classifies the tuple and the class with the maximum votes is considered as the final result. Model is trained first then testing is performed. Like NN (neural network) , the model can be used for both regression and classification examples [11]. It also reduces the risk of over-fitting. Given that random forest can accurately perform both regression and classification tasks. The random forest classifier has advantage that for missing values in the dataset, accuracy is preserved by bagging. So, this is the machine learning model that is helpful for handling missing values in the dataset. Random Forest is also a slower and complex model and it needs large number of samples in dataset. Random forest makes it easy to evaluate variable importance. The Random Forest model, when applied for fault prediction, it gives the highest accuracy before grouping and also after grouping it is increase by around 30percent. It has been proved that the F1-score of random forest also nearly doubles after grouping.

Table 1: Accuracy before Grouping

| Model | Accuracy | F1-score |
|---|---|---|
| Linear Regression | 9.0909 | 0.0526 |
| Logistic Regression | 13.636 | 0.1083 |
| Decision Tree Classifier | 11.764 | 0.057 |
| Random Forest Classifier | 36.363 | 0.308 |
| Support Vector Classifier | 17.647 | 0.066 |
| Bernoulli Naives Bayes | 23.529 | 0.0761 |
| K-Nearest Neighbor | 5.882 | 0.0333 |

Table 2: Accuracy after Grouping

| Model | Accuracy | F1-score |
|---|---|---|
| Linear Regression | 54.545 | 0.541 |
| Logistic Regression | 29.411 | 0.156 |
| Decision Tree Classifier | 11.764 | 0.052 |
| Random Forest Classifier | 63.636 | 0.6335 |
| Support Vector Classifier | 17.647 | 0.093 |
| Bernoulli Naïve Bayes | 0 | 0 |
| K-Nearest Neighbor | 5.882 | 0.049 |

### F. *Reasons for Grouping of Data*

1) To overcome the problem of low accuracy and F1-score.
2) Splitting the range of data into small blocks of assigned bits, so as to reduce the complexity of result declaring in the ML models.
3) Class Prediction leads to fault prediction in a certain range of bits; though it reduces the resolution of prediction, it makes the implementation easier and accuracy higher.

## III. EXPERIMENTAL DATA

F1 Score and Accuracy computed by confusion matrices are very widely used performance parameters in classification problems [12]. In any machine learning application, the best way to best predictions of the faults rely on the use of the proper algorithm and on the quality of the dataset available. We have used two metrics for calculating

the algorithm's results are F1-score and the most critical accuracy. Accuracy is the one of the basic parameter can be verified that checks total correct predictions out of total tuples applied for testing the model.

$$Accuracy = \frac{no.of\ correct\ predictions}{Total\ no.of\ predictions}$$

F1-score is the harmonic mean of precision value and recall. It is one of the simplest ways to combine any performance parameters of the model. Also both parameters required to be high to achieve high F1 score. So, F1-score is more affected to low value in one of the parameter. Accuracy is more statistic parameter which is more intuitive and F1score is more preferred for imbalance dataset. We preferred both the performance metrics for our model's evaluation and also to select the best model for applications. It has been observed that the result has a better F1 score in Random Forest Classifier. Each model has varying accuracy due to quality of the available dataset. If dataset is more accurate, balanced as well as standardized can give us accurate results. This shows that variety of enrich datasets is required for the classifiers that can performs better after they are applied.

## IV. CHALLENGES

The biggest challenge for machine learning applications for VLSI test field is readily available training data set. For training purpose and result analysis as well as bench marking, the research needs a large pool of standard verified data set. When it comes about VLSI testing field, there are bench mark circuits and ready test data available like ISCAS-85 and ISCAS-89 combinational and synchronous circuits (International Symposium of Circuit and System) benchmark circuits, ITC-99 benchmark circuits, TAU circuits for timing analysis. These circuits and available test data set are considerable enough for optimization algorithms and their result analysis. In these recent trends of AI-ML applications for VLSI test, these circuits are found not enough for training and testing of AI-ML efficiency for this complex field of VLSI.

With the better and larger dataset available, many more features of VLSI design verification and testing can be better analyzed for a detailed further analysis and the same is applicable for the case of fault prediction. Dataset has basic details of the circuit's parameters like fault coverage and a few test patterns. However, it needs a enrich dataset with many tuples in the dataset. If such dataset is used for classification problem, it will be a robust algorithm that results better from many features available.

As the regression algorithms are never performs well for classification problem, so, they are not preferred for classification with higher values. Machine learning cannot predict accurately with false or missing data (garbage in, garbage out). Therefore, the challenge is to handle the missing data. This problem is also known as corruption of data. If '0' values are given instead of missing values ('NaN'), the output will consist of false predictions. NaN data can be a cause of large errors in machine learning models. The smaller size dataset is a big hurdle that needs to be tackled by proper documentation and the use of better software for accurate and verified documentation of these data points. Standard datasets are not easily accessible; So, the set applied to the ML model here needs to be preprocessed and verified for very low noise. We rely on locally generated circuits and test vectors to prepare datasets for training and testing of the model. They might have lower efficiency and they need to be checked before use.

## V. CONCLUSION

In this paper an evaluation method for identifying faults using various ML algorithms, from a perspective of F1-score measurement and accuracy of predictions. The final F1-score of this approach is an average of the result of every algorithm's specific sequence. The accuracy tells us about the ability of the particular model to predict faults. Various types of machine learning models were used throughout this research work. The implementation of this machine learning model has been done on the Google Colab platform using Python programming language. These techniques have been used to predict the number of faults for respective digital circuits.

The prediction of the number of faults was done more accurately by the Random Forest model with an accuracy of 63% after grouping the number of faults. The result analysis of better performance of Random Forest model compared to other model is out of the scope of this paper and will be considered in future scope.

## VI. FUTURE SCOPE

This paper has explored the fault prediction for a limited type of cases only. It means it covers only combinational circuits but the same experimental work can be done on sequential circuits and synchronous circuits. Further, the gate types considered here are only primitive gates like AND2, NAND2, OR2, NOR2 and INV. The most required work will be to apply these algorithms on fab specific gates being used in ASICs with different functions and different fan-ins. The future scope is also in examining the other types of fault models like delay faults, transistor level faults and bridge faults.

## REFERENCES

[1] Suren Abazyan, Vazgen Melikyan, "Enhanced pin access prediction and design optimization with machine learning integration", *Microelectronics Journal*,vol.116,pp.105198,2021.

[2] L. Lu, J. Chen, M. Ulbricht and M. Krstic, "A Methodology for Identifying Critical Sequential Circuits with Graph Convolutional Networks,"*IEEE Computer Society Annual Symposium on VLSI (ISVLSI)*, 2022, pp.20-25, doi: 10.1109/ ISVLSI54635. 2022.00 017.

[3] K. Khalil, O. Eldash, A. Kumar and M. Bayoumi, "Machine Learning Based Approach for Hardware Faults Prediction," in *IEEE Transaction son Circuits and Systems I: Regular Papers*, vol. 67, no. 11, pp. 3880-3892,Nov.2020,doi:10.1109/TCSI.2020.3010743.

[4] Jiang,B.L.,Wey,C.L.Fan,L.J. "Fault prediction for analog circuits. *Circuits Systems and Signal Process* ,95–109 (1988). https://doi.org/10.1007/BF01600009

[5] B. Yuce, N. F. Ghalaty, and P. R. Schaumont, "Microprocessor fault detection and response system," U.S. Patent 10452493, Oct.22, 2019.

[6] Guo, G., Wang, H., Bell, D., Bi, Y., Greer, K. (2003). KNN Model-Based Approach in Classification. In: *Meersman, R., Tari, Z.,*

*Schmidt, D.C. (eds) On The Move to Meaningful Internet Systems 2003: Coop IS, DOA, and ODBASE. OTM 2003*. Lecture Notes in Computer Science, vol 2888. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-39964-3_62

[7] W. Yanxia, "Student Information Management Decision System Based on Decision Tree Classification Algorithm," *2022 IEEE 5th International Conference on Information Systems and Computer Aided Education (ICISCAE)*, Dalian, China, 2022, pp. 827-831, doi: 10.1109/ICISCAE55891.2022.9927597.

[8] Parveen, A. Singh, ”Detection of brain tumor in *MRI images, using combination of fuzzy c-means and SVM,” 2nd International Conference on Signal Processing and Integrated Networks (SPIN)*, 2015, pp.98-102, doi:10.1109/SPIN.2015.7095308.

[9] Kavitha S, Varuna S and Ramya R, ”A comparative analysis on linear regression and support vector regression*,” Online International Conference on Green Engineering and Technologies (IC-GET)*, 2016, pp.1-5, doi:10.1109/GET.2016.7916627.

[10] G. Singh, B. Kumar, L. Gaur, A. Tyagi, ”Comparison between Multinomial and Bernoulli Na¨ıve Bayes for Text Classification,” *International Conference on Automation, Computational and Technology Management (ICACTM)*, 2019, pp.593-596, doi:10.1109/ICACTM.2019.8776800.

[11] S. Mishra, R. K. Mallick, D. A. Gadanayak, ”Islanding Detection of Micro grid using EMD and Random Forest Classifier,” *International Conference on Computational Intelligence for Smart Power System and Sustainable Energy (CISPSSE)*, 2020, pp. 1-5, doi:10.1109/CISPSSE49931.2020.9212279.

[12] Chicco, D., Jurman,G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics 21*, 2020, doi:10.1186/s12864-019-6413-7

**Sima Gonsai** is currently working as an assistant professor in the Department of Electronics and Communication at L. D. College of Engineering, Ahmedabad. She has received B.E. in EC Engineering from U. V. Patel college of Engineering in 2004 and M.E. in EC with a specialization in Communication Systems Engineering from the L. D. College of Engineering, in 2007. Currently she is research scholar at Institute of Technology, Nirma University, Ahmedabad.

**Prof. (Dr.) Usha Mehta**, currently a Professor and Head at EC department, Institute of Technology, Nirma University, Ahmedabad. She received her bachelor degree in Electronics & Communication Engineering in 1994 and master in 2005 and Ph. D. degree in VLSI Design in 2011.
She has more than 23 years of experience for teaching at Under Graduate and Post Graduate level. She has guided more than six Ph.D. students and M. Tech by Research students. She has one Patent on her credit. She has been the Principle Investigator for research projects of ISRO and GUJCOST. RHBD Chip developed by her and her teammates for ISRO has been fabricated at SCL, Chandigrah and in use by Space Application Centre. She has published more than 55 research paper in area of VLSI Design and Testing. She has delivered large number of technical talks at various conferences and workshops. She has contributed as the conference chair of international conferences like VDAT2015, NUiCONE2012, NUiCONE-2011 and so. She has volunteered IEEE Gujarat Section for various roles like WIE Chair, Section Secretary and Chapter Coordinator, Membership development chair etc. She has chaired the IEEE Gujarat section Women-In-engineering affinity group for more than six years and is actively involved in encouragement of Women Professionals. She is a senior member of IEEE, Associate member of CSI, Senior Member of IETE and Life member of ISTE. Her passion is teaching and research but beyond this, she is involved in academic administration. She is part of various committees involved with strategic planning and execution like Internal Quality Assurance Cell, University level Research Committee, Academic Council, etc.