International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2019): 7.583

Advanced Testing Frameworks for Next - Generation Semiconductor Devices Using Machine Learning

Botlagunta Preethish Nandan

Senior Lead Business Intelligence
Email: *preethishnananbotlagunta[at]gmail.com*
ORCID ID: 0009 - 0008 - 3617 - 8149

Abstract: ***Semiconductor devices are the essential building blocks in today's information technology and society. Next - generation semiconductor devices such as FinFETs, GAA - FETs, and nanowires have been proposed to serve for high - performance and low - power applications in the deep - submicron technology nodes. However, with the rapid scaling of semiconductor technology nodes, traditional test generation methods face large runtime and memory footprint challenges to ensure efficient fault diagnosis and reliability screening. In this regard, machine learning (ML) has emerged as a new paradigm to address the above issues with great success. This paper presents recent advanced testing frameworks for next - generation semiconductor devices by enabling ML. Several successful demonstrations are covered, creating and exploiting novel ML models of various levels of complexities. At the device level, fast test generation methods using shallow ML models are presented, yielding significant improvement in runtime and memory efficiency compared to existing commercial tools. At the circuit level, the applicability of deep learning - based approaches for stuck - at fault identification and location is explored. Finally, a massive data generation and representation learning framework for deep neural network - based built - in self - test (BIST) generation is presented to improve the design robustness of the built - in test (BIT) architectures for large - scale applications. The basic concepts and implementations are first introduced, followed by successful demonstrations and industrial applications. This opens up questions on how advanced ML models can improve existing approaches, where they might fail, and how to mitigate the biases. Manual test development for semiconductor chips relies on designerś knowledge, experience & heuristics, requiring substantial time & effort. ML models representative of test - influencing factors and their likely trade - offs can be created directly from the available test data. Such auto - generated ML models can then be exploited for test generation or validation, or to estimate the test cost based on semiconductor models. A bottleneck in this approach is the time - consuming, complex, intensive operation of taking different process steps to generate quality physical or logical test data on the specific test chips. ML methods are exceptionally data - hungry, requiring a large volume to train and generalize successfully. However, there are many untapped sources of ample simulation data both from older chips and different fabrication processes. It is possible to create new problem representations, or simulation domains, that standardize/normalize all factors not relevant to a particular problem. Various progressive domain adaptation methods can then be used to adapt existing ML models from these pre - trained representations to the new problem domain. Once adapted, the production ML models can be used for fully automated test generation or engineering studies.***

Keywords: machine learning in testing, semiconductor device reliability, automated test generation, deep neural networks, domain adaptation strategies

1. Introduction

As semiconductor device technologies advance based on novel architectures and new materials, the need for complex and advanced testing is anticipated to increase. Consequently, it is vital to develop appropriate testing techniques and tools that can effectively and efficiently test next - generation semiconductor devices. To date, enormous efforts have been made in testing methodologies and frameworks. Traditional testing techniques are insufficient for capturing important parameters. Consequently, researchers are focusing on integrating machine - learning with testing methodologies and frameworks introduction to enhance test quality.

The context of this paper is on next - generation semiconductor devices, namely new types of FinFET Stress Test Structures, Gate - All - Around FET Logic Devices, and 2D Materials for Future Semiconductor Applications. These types of devices involve complex structures and novel mechanisms, making advanced testing methodologies necessary. However, conventional testing methodologies do not consider the new characteristics presented by the devices. Various parameters must be analysed to ensure good performance, which significantly increases the complexity of test vector generation and design. Machine - learning methods are able to significantly reduce the complexity and enhance the quality of the generation while analysing the new characteristics of the devices.

Machine - learning - based testing methodologies are still in the early stage of development. There has been a surge of interest in exploring the testing methodologies that leverage machine - learning techniques and frameworks. This paper provides a comprehensive survey of the major testing philosophies, procedures, and techniques, accompanied by illustrative selected applications, that bring machine learning into the physical assurance of semiconductor devices. To compare existing works based on different criteria, a taxonomy of machine - learning - based testing methodologies and frameworks is presented for the first time. Additionally, challenges and opportunities in testing next - generation semiconductor devices using machine - learning techniques and frameworks are further discussed to motivate future research. The growing interactions between testing and machine - learning domains hold enticing potential for contributing to the emergence of new methodologies and techniques, which will benefit the efficiency and reliability of semiconductor devices.

2. Overview of Semiconductor Devices

Semiconductors have become a part of everyday life in different forms: TVs, laptops, game consoles, and cars, to name a few. All of these products require semiconductor devices inside to perform the different tasks. The workhorse of the semiconductor industry is a field - effect transistor known as the metal - oxide - semiconductor field - effect transistor (MOSFET). This device is typically 1–2 μm in size, and it is built using the semiconductor material, silicon (Si). There are some inherent properties of Si that make it a popular choice over the other materials. Most importantly, it facilitates the construction of Metal - Oxide - Semiconductor (MOS) structures which can be used for device fabrication. This property can be exploited to fabricate thin oxide layers (~1 nm) on the Si dielectrics, with the gate oxide influencing the device performance characteristics. The compact model describes the circuit or electrical behavior of the device based on the underlying physics of the device; hence, it is very useful for fast simulations and typical design scenarios in an industry

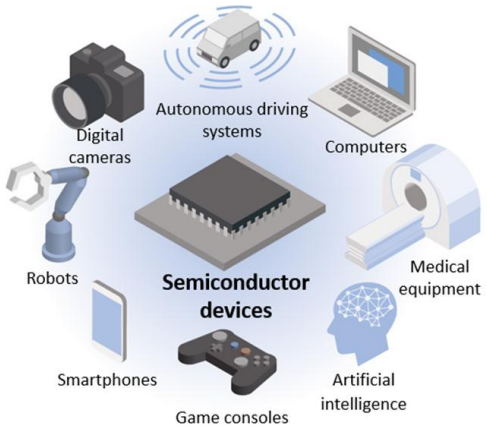

Figure 1: Semiconductor Lithography Equipment

In the recent past, there has been a significant amount of research in developing accurate compact models for the sub - 100 nm technology nodes. The advent of machine learning (ML) has given rise to new generation models which can predict the device characteristics very fast; hence they are very useful for design iterations. Use of ML for modeling device characteristics is a classical nonlinear regression problem; hence different techniques can be applied. With the rapid advancement of the semiconductor industry, the number of processes involved in making one device has drastically increased. Semiconductor fabrication is an interdisciplinary field with a cumbersome process flow. The difficulty lies in the fact that there are many chemical reactions and multidisciplinary knowledge is necessary.

With the rise of ML, it becomes more feasible to use ML algorithms to describe the complex relations between semiconductor processes: etching and deposition at the plasma level, oxidation and wafer cleaning at the fluid level, and lithography in 2D at the solid state. The efficacy of supervised learning is very high but comes with one major drawback: there is a requirement of a large amount of crucial labeled data. In the real manufacturing environment, data sparsity is a critical issue, as not all the sequences of processes are measured throughout the process. The best modeling technique for the semiconductor industry is to mix the idea of ideal physics, which gives the coherent insight into the physics of the problem, with the ML algorithms, which can learn the pattern and behavior from the data at the same time [3].

3. Importance of Testing in Semiconductor Manufacturing

Semiconductors are critical and indispensable key materials for the manufacture of modern electronic devices. Forward - looking technologies such as artificial intelligence (AI), quantum computing (QC), new energy vehicles (NEV), and HiGHTS —high performance, high energy density, high frequency, high temperature, high voltage, and high reliability —gas turbine engine, all face growing requirements for improved performance, efficiency, size, and costs. Continued progress in these advanced applications critically hinge upon the development of next - generation semiconductors, which require substantial reductions in size and improved performance, efficiency, and yield. Next - generation semiconductors include both new materials and new structures of the classical silicon devices that have been dominating the semiconductor industry for more than half a century since the invention of transistor. New materials are made of quantum dots or other two - dimensional materials, such as transition metal dichalcogenides (TMDs) MoS2, WSe2, etc. New structures of classical silicon devices include three - dimensional structures, such as gate - all - around field effect transistors (GAA FETs) and Fin field effect transistors (FinFETs). Such next - generation semiconductors pose multiple unique and unprecedented grand challenges in terms of practical realizability, fabrication, testing, and reliability, all of which must be addressed with innovations in materials and method development.

Conventionally, a testing framework is composed of a number of oftentimes independently designed and implemented test patterns, instruments, and techniques in hardware and software. While a conventional testing framework is effective, it often suffers from difficulties in scalability, performance, and time. Specifically, as more and more next - generation semiconductor devices under test (DUTs) are introduced into the testing framework, the issue of the in scalabilities of test patterns, instruments, and techniques becomes particularly pronounced. On the efficiency side, testing throughput demands assistance by modern algorithms including both classical and emerging AI techniques. Finally, fixed hardware testers and software testers can incur high engineering costs and tedious time to design test patterns for new devices [4]. In addition, much of the existing test patterns are often underuse when modern devices change quite a bit. Here, ML may also help identify transferable test patterns across various next - generation semiconductor devices.

4. Challenges in Traditional Testing Methods

Semiconductor testing involves a set of operations associated with performing electrical tests on semiconductor dies, in particular on those that have been fabricated in the fabrication process. Semiconductor testing consists of two parts including Wafer - Test (Pre - bond test) and Final - Test (Post

- bond test). The wafer test serves as an early reliability and yield screening step in the semiconductor manufacturing flow, and therefore could significantly reduce the test cost if it is effective. Because dies are tested at the wafer level, there is also a chance to test a large number of pins simultaneously. The die, as such, has enhanced accessibility and fewer physical limitations [4]. However, it is also harder to find the parameters with respect to measure the test effectiveness due to the non - linearities of the die.

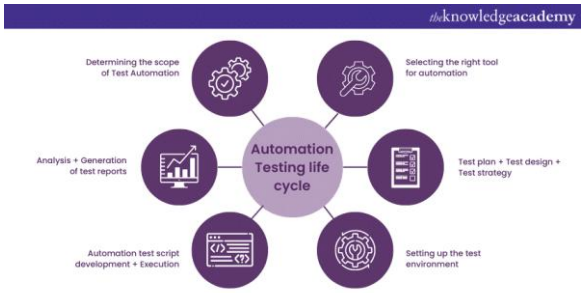

Figure 2: Top Challenges in Automation Testing

At the wafer level, a die that has got from the wafer probing test will have the pin assignments according to the layout design database since it flows to the Final - Test at the first time. A partitioning stage called translation is needed, which will determine the design pin assignments. The translation could affect the test effectiveness. Loss of signal propagation across the active silicon area can influence the test predictability and severity. A test - deck could be produced for the wafer probe test (WPT) using the translation data but there are infrastructure and coordination issues. The test could not be easily reusable and it wastes time to generate the test - deck for each wafer probing. This could also refer to the limited test replacement and difficulty in the generation of the site - specific insertion loss tests.

With respect to the scanning architecture, there have previously focused on test access and scan architecture visualization. Masked testing is needed for the complete and necessary output fault simulation to perform full - semiconductor fault coverage fault simulation. Deep - learning based approaches to classify the test - data and test passes [3]. Integrated pattern compaction reduces the amount of test - data stored on - chip to ease the process of downloading to the chip during test mode while allowing more test - data to persist in the compactor output to achieve higher fault coverage.

5. Introduction to Machine Learning

Machine learning (ML) focuses on developing and evaluating computer programs that improve performance at some tasks with experience. The simplest form of ML is based on supervised learning approaches. In supervised learning, data sets with input - output pairs are provided, and the algorithm attempts to predict the output from the input. Unsupervised learning evaluates training scripts without a target outcome. The algorithm attempts to cluster input data with no target or a limited understanding of the value. Reinforcement learning techniques are provided with limited supervision, consisting of guidance that incorporates the desired outcome value within restriction boundaries. Often a reward history is processed to develop an evaluation outcome. The ML process consists of three main stages: data preparation, training, and evaluation. The data preparation phase transforms inputs into usable data. The training phase consists of providing the ML models with the prepared data enabling networks to identify relevant features. The evaluation phase uses other existing datasets to evaluate how well the trained model performs on new data [1].

ML has been developing over recent decades but transitioned to faster development with emergent approaches and solutions. ML research is now presented with introductory principles and hands - on tutorials, resulting in academic courses and many fields where it can be applied. Many ML introduction packages are now in programming libraries, taking seconds to apply. With models and well - structured input data, surprising is often outputted. Generated ML algorithms' complexity prevents a clear understanding of how inputs will be processed. Furthermore, testing the validity of such algorithms becomes hard for created ML models. It is unsure which eventualities and observations were considered in the testing loop during evaluation in the simplest libraries, while knowledge on input data is lost in reprocessing in complex models.

6. Machine Learning Applications in Testing

Machine Learning is one of the popular Artificial Intelligence (AI) approaches for developing early fault detection applications in semiconductors. ML techniques can analyze system data and identify potential problem patterns. Also, ML can optimize a systematic approach involving manual or semi - manual processes and development cycles [1]. A systematic approach prevents space exploration or mission drift and systematic testing. Avoid over - testing and under - testing test buttons and cycles. In infrastructures involving multiple components, ML can precisely identify and aid the verification of newly updated components.

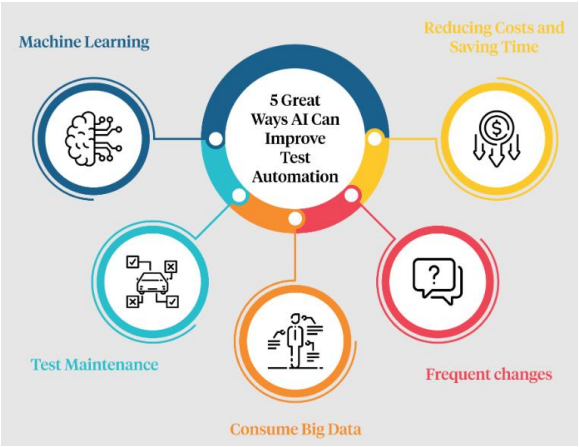

Figure 3: Machine Learning in Test Automation

Some survey efforts summarize various uses and classes of ML approaches in verification and validation, listing potential applications in testing and design verification and constraint satisfiability. Many automatic testing or verification approaches, including different verification levels in different groups, involve manually crafted or semi - manual process resources. Finding the optimized input to perform system

signing without a specific model or coverage target is a challenge. Computed Static Analysis inputs static analysis into ML to enhance the fault detection coverage of the verification process and reduce runtime. Firm Sample Recovery exploits the temporal relationship between γ - universal and context - free regular expressions to synthesize initial compact samples and generate new and diverse test cases. Clustering applies a uniform - vector similarity comparison using multi - style learned vectors at abstraction levels to narrow the search space. To be efficient, most of these approaches require a large amount of design information such as event logs, signal traces, and Intent.

7. Frameworks for Machine Learning in Semiconductor Testing

New advanced technologies have been introduced, including a 5G high - speed ADC that can be tested with builtin testing methods, a high - speed SAR ADC with testing frameworks based on SL, and a multi - channel pressure sensor with digital background calibration. Interconnects, as a primary component of Integrated Circuits (ICs) for over four decades, show an increasing delay and cost, along with diminishing returns in circuitry [1]. By the year 2020, approximately 90% of the area in a standard cell IC is allocated to interconnects, and efforts to reduce delay via technology are restricted by the trade - offs in alternative costs, routing congestion, and the emphasis on longer channels. Despite incorporating new materials, there is no confidence that these approaches will permit excessively large chip sizes and connectably heavy systems on ICs [3].

Machine learning is currently a major theme in scientific research, offering promise for a variety of applications, including failure detection in physical systems of ever - increasing complexity. It offers an opportunity for a complete reappraisal of test data metrics, as well as the development of new test metric generation and evaluation approaches at various testing stages. A framework is introduced that brings together various aspects of machine learning that may be of relevance in the context of semiconductor testing, including some practical applications. It outlines important requirements and offers many examples of the potential use of machine learning at testing stages, as well. However, it is clear that much work is needed to ensure that the approaches can lead to the desired outcomes in an economical manner.

7.1 Data Collection and Preprocessing

In semiconductor manufacturing, process and equipment monitoring are becoming increasingly complex due to generational technology advances from planar devices to fin field - effect transistors or gate - all - around field - effect transistors. Artificial Intelligence has evolved to become a powerful tool in this field for detecting and predicting wafer yield well before its fabrication stage. However, with an increasing number of wafer fabrication factories, process tool and recipe diversities, different concerns for yield degradation, and evolving sophisticated device structures, test information grows both more complex and massive, rendering value extraction increasingly elusive [3]. As the first step of any machine learning or deep learning methodologies, data collection and processing involve data selection, filtering, preparation, and transformation to render them suitable for input into models. In practice, this is realized in conjunction with data feature discovery and extraction, emphasizing efficiently capturing data distribution without losing valuable information. Recently, pre - trained data collection and transfer learning techniques have come into play. However, in this absence of pre - trained models, some issues have received less attention.

When devising a novel methodology, novel tests are often performed corresponding to requirements not previously observed, in which cases data collection is performed afresh. As test information is in the form of time signals or curves, it possesses the inherent dimension of time. Taking this into account, this section reviews signal feature extraction methods suited to the investigation of novel requirements. Detecting the considerable amount of process data 24/7 in a foundry is challenging. Since achieving complete fault coverage with a finite number of tests is infeasible, tests need to be prioritized to ensure maximum fault detection capabilities [5]. Due to the size of the search space, heuristic search algorithms are often applied for test prioritization. With equipment becoming increasingly complex and automated, mass logs are generated beyond human capability to analyze manually. Patterns representing equipment health status and abnormalities need to be automatically discovered from these logs. Similarity between test conditions is a prerequisite for using an existing test to qualify new devices. When transferring a Neural Network, the lower the similarity, the more retargeting is needed. Feature extraction methods are reviewed for data collection in various situations, and how a particular type of data collection could suffer from its own failings when evaluated a posteriori.

7.2 Model Selection and Training

The deep learning frameworks, including deep neural networks, convolutional neural networks, and recurrent neural networks, with various structures and hyperparameters can be used to ascertain machine learning architectures for target applications. It is essential to choose the right structure and hyperparameters that can enhance the efficiency, robustness, and accuracy of analysis. Hence, a sub - optimal model can lead to poor combinations that may waste the training time and deteriorate the test performance. The selection of the model structure can have exaggerated repercussions for complex problems. Therefore, the model architecture recommendation is necessary for each task and dataset.

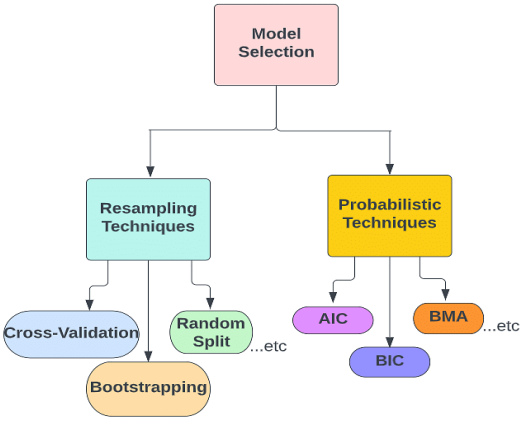

Figure 4: Machine Learning Model Selection and Parameter Tuning

An in - depth survey for the related works is performed that provide automated methods for architecture searches. Many approaches focus on heuristics that typically use fixed - length encoding. A poor search strategy is often employed for optimization algorithms. Only a handful of works leverage recent advances in optimization. The neural architecture search with reinforcement learning can easily obtain results, but their training times are too long. The genetic programming - based methods can be effective but require more customization of the prior knowledge and framework. Besides these, a framework that handles hyperparameter, architecture search, and training with a more user - friendly way is developed. For the end - user, it is proposed to automatically tune the hyperparameters, architecture search, and training simultaneously.

The preliminary results can find highly competitive architectures within hours of running on the end - user PC, and the plug - and - play experience is nice. Though techniques have never been introduced to advance deep learning architecture search, the required experts are quite needed to tune the hyperparameters which can take hours or days. Nevertheless, this framework can benefit this testing machine learning community. Surprisingly, it has been scarcely studied for computer vision tasks even though its performance well above the set standard. The existing frameworks do not include natural ways for experts to inject prior knowledge.

8. Case Studies of Machine Learning in Semiconductor Testing

Machine learning algorithms have been researched to enhance the functional verification process. These algorithms were applied to the coverage directed test generation, coverage driven verification, and runtime optimization of constrained random verification. The clustering approach was used for coverage - driven verification of the communication core. In addition, clustering based approaches were used to debug test failures in UVM - based tests at both the RTL and gate level. Coverage driven test generation is mostly model - based, thus more limited in application scope

Primarily, this paper describes the machine learning algorithms suitable for improving the coverage directed test generation process either by steering or altering the conditions of the test generation engines or by tuning the test generation proofs. Meanwhile, promising future directions and areas of research are outlined. It is presented that coverage - driven test generation is proactive in generating potentially valid tests, while test - selection from generated test cases is reactive. Therefore, machine learning algorithms can be applied to the test generation and test selection processes. As a follow - up, the challenges and some initial ideas of applying machine learning algorithms in automated coverage - driven test generation are discussed.

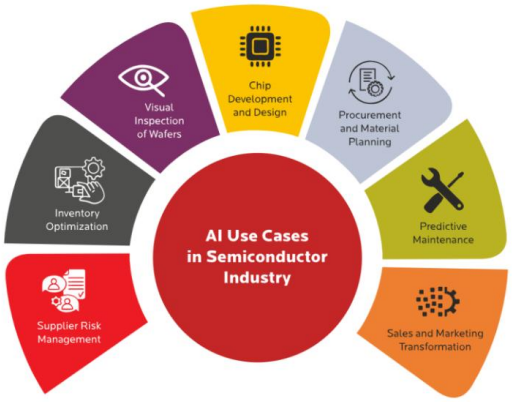

Figure 5: AI in Semiconductors Industry

Big data and machine learning rapidly change life styles, and they can potentially change semiconductor research and development. Despite substantial advancements, the industry still struggles with electro - migration and stress void, fin - cycle - time - variation, hardware security, and new device architecture. The industry faces challenges in analyzing coverage holes in verification patents, accelerating qualitative error understanding, and analyzing performance - drift defects in post - silicon physical characterization [3].

This paper showcases some of the successful achievements of applying big data and machine learning for semiconductor failure analysis. Several remaining challenges are also summarized for future research efforts. Semiconductor devices are fashionable applications of hardware - oriented machine learning. Modern semiconductor devices have nanoscale dimensions and increased complexity, which makes them difficult to manufacture and test. As device dimensions shrink, testing is increasingly reliant on electromagnetic simulation models and less test hardware. Simulation models need to be verified for precision, which is complex and time - consuming.

8.1 Case Study 1: Predictive Maintenance

This case study presents a recently developed approach for predictive maintenance based on measurement and prognosis of logical failure mechanisms. The approach consists of a Tangible Digital Twin which predicts the time - to - failure of a product, reasoning about its Condition Indicator and Failure Avoidance actions at the same time. It actively operates with real measurements and simulated failures in a changing environment, with learning capabilities to continuously refine its models with new data. The proposed method is grounded on qualitative change points, which project the high dimensionality of raw time series data onto unsupervised, lower size, interpretable, and actionable qualitative variables. In the last decades, predictive maintenance applications have become one of the most impactful Artificial Intelligence and Machine Learning areas, with successful industry implementations reducing costs and failures rates significantly [5].

A wide variety of data - driven methodologies for product condition monitoring and prognosis have been proposed, supporting on ‘offline’ exploratory analytics or/and ‘online’ decision making. They significantly differ in accuracy, interpretability, and tractability, but all these methods are

using a 'Digital Twin', large datasets with simulations or historical measurements of condition changes and an expensive trustable simulator to forecast the evolution of products and systems. The methods for real - time applications usually sacrifice the product simulator. An alternative enabling the breathing digital twin concept is currently being addressed by fundamentally different Boolean - state modelling and analysis methods which do not require simulation of the conditional behaviour, but directly operate by encoding measurements and business rules in Boolean states and rules. Such methods have numerous advantages over digital twins: a very small state space, fully interpretable model structures, expressive visualisations and analyses, and direct coupling with mathematical programming, simulation and statistical learning.

8.2 Case Study 2: Quality Assurance

In semiconductor manufacturing, despite a lot of efforts on manufacturing and design improvement of advanced technology nodes, technology is becoming more complicated and defect density is being reduced. However, increasing design complexity, device performance, and manufacturing process variations can still lead to larger variability and defects [3]. In an attempt to capture these changes, qualification testing approaches are challenged to efficiently and effectively plan characterization at - worst - case (AWC) corner settings. Although running a lot of dummy tests at AWC corners can ensure the correctness of a design, it increases time - to - market and testing costs. To address this, a novel data analytics approach to characterize the test cost and effectiveness of all test conditions is proposed.

In general, a large amount of collected data from past technology nodes may not be reused due to the obstinately increasing difficulty in predicting defects. In an attempt to solve this, a data - augmented optimization approach is proposed to select AWC corners. Note that even with current designs having a similar structure, design - flavors can be so diverse that narrowing down the corner settings can be ambiguous. Fortunately, the most widely exploited timing delay defects in static timing analysis can be predictable by constructing rejection region models. Instead of finding the exact test corner settings, the target test conditions can be narrowed down to rejection regions [5].

Based on that, selective data - augmented optimization targeting at the design - specific situation is carried out repeatedly and thus competitors can be analyzed, enabling the selection of corner settings. The experimental results, using different cases ranging from dual/multiple supply voltage designs, are also investigated. Numerous potential test settings involving a lot of process corner variations, different parameters are tested and assessed in terms of their cost and effectiveness on the timing path faults are provided.

9. Evaluation Metrics for Testing Frameworks

The proposed testing framework incorporates ML techniques, but its reliability depends on the availability of reliable and robust datasets. In addition, the following evaluation metrics are proposed for comprehensive DEDs testing framework evaluation:

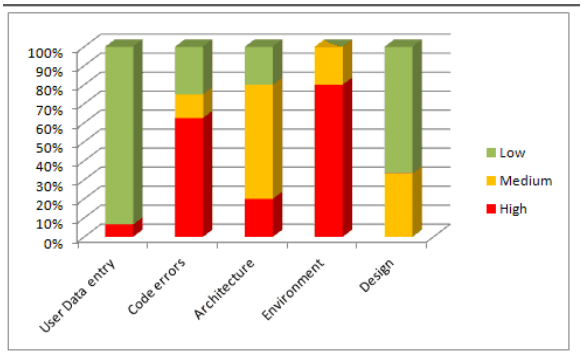

Figure 6: 4 Essential Testing Metrics for Measuring Quality Assurance Success

ML model performance evaluation metric: This metric includes several common performance metrics generally used to evaluate ML model performance, including confusion matrix representation, accuracy, precision, recall, F1 score, ROC, and AUC. The input is an ML model, DEDs dataset for testing, and performance metric option. The output is the desired performance metric values. The recommended metric usage is based on whether the dataset is imbalanced or balanced. The confusion matrix (true positive (TP), true negative (TN), false positive (FP), and false negative (FN)) are used if the dataset is imbalanced, while accuracy, precision, recall, and F1 are mainly utilized if there are several classes (K>2), otherwise accuracy, precision, and recall are sufficient (K=2). ROC and AUC are also recommended if there is class imbalance.

Testing framework robustness evaluation metric: This metric includes several robustness metrics to evaluate the testing framework's robustness by determining the robustness of individual classifiers. The input is a DEDs dataset and the generated performance metrics values (using the above testing framework robustness evaluation metric). The output is a score used to evaluate the robustness of the testing framework. This metric demonstrates the robustness of individual classifiers based on the robustness postulates, and if they are not strictly followed, the classifiers might be sensitive, especially to testing data variation. The estimated score can be utilized to evaluate individual classifier robustness. Therefore, the robustness analysis of the testing framework can be done based on the robustness evaluation scores of individual classifiers.

Testing framework consensus evaluation metric: This metric incorporates the proposed consensus function to evaluate the effectiveness of the consensus function. This metric is introduced since the final classification decision of the testing frameworks is determined using consensus functions. Additionally, scoring rules are also utilized in decision fusion than consensus functions, so equivalent metrics are also incorporated for scoring rules. The input is a DEDs testing dataset, individual classifiers, and final decision using a consensus function (both centering and scoring rule functions). The output is the evaluation metric value. The consensus metrics' value must not be zero [6]; else, the classifiers do not agree on the final decision. Therefore, these metrics are used to evaluate the effectiveness of the consensus function in a testing framework.

10. Integration of Machine Learning with Existing Testing Protocols

There are several prominent testing protocols used for next - generation semiconductor devices that are integrated into machine learning - based approaches starting from machine learning - based circuit design for a reliable post - fabrication testing mechanism to data processing modules for registration - based electrical test data specifications to hardware - in - the - loop methods for extraction and simplification of complex testing specifications, among many others presented in literature. These protocols and applications can be analyzed based on a machine - learning - based approach's data requirements for training reliable models based on both the data content and the utilized pre - processing steps for improved model performance.

Eqn 1: Conceptual Equation for ML Integration in Testing

$$T_{ML} = f(T_{legacy}, D, M)$$

Where:

- T_{ML}: The output of the test protocol enhanced with machine learning.
- T_{legacy}: Existing (legacy) testing results, protocols, or rules.
- D: Data collected from past tests, logs, defect reports, and usage patterns.
- M: Machine learning model(s) trained on data D (e.g., classifiers, regressors, anomaly detectors).
- f: A function representing the integration logic — combining ML outputs with traditional test mechanisms (e.g., rule-based decision logic, model predictions, heuristics).

Models trained on testing data generated with specific protocols can be reused towards the testing of other devices designed with a similar architecture. By only adjusting the sampling rate of the input signals such models can produce almost identical testing procedures for other devices of the same family, while more comprehensive test specifications remain needed to derive functional tests from scratch for completely different designs. Perfectly conditionally - independent electrothermal data sets can be trained with no cost, if at least one test chain circuitry is available in each chip design family, towards the generalization of test specifications among all devices from the same family [1]. There are no hardware requirements besides DFT logic; however, the runtime is longer than that of ML approaches. The former must have kernel approximation during the entire chip testing, while the latter is only needed during model training.

11. Future Trends in Semiconductor Testing

In the last four decades, due to shrinking geometries, the semiconductor industry has observed increasingly multifaceted integrated circuits (ICs). These ICs pose testing challenges that are far ahead of traditional methodologies. Following a time of total focus on problems arising due to Technology Scaling (TS), the focus has now shifted to Integrated Device Manufacturing (IDM) and System on Chip processes (SoC) [7]. Process variability and new manufacturing technologies bring in a new host of challenges, many of which overlap with performance variation challenges and incursions in test methodology. Along with these changes, a new phenomena Baker's Dozen has emerged. Due to riding the Learning Curve, Electro - migration, Frequency Drifting, and so forth, early silicon death is a real cause to worry and is extremely difficult to catch at the design stage. A rising pattern for SoCs or IPs is to protoboard. As Soon as Possible in silicon prototyping, either to cut first silicon timing, power and area, or to have some proper working die to validate algorithm vs hardware. Aspects like grouping embedded memory snapshots to sampling rates, start up signals, frequency changes, etc. have to be dealt with at design - time as they serve as major break - points to concurrent testing. Moreover, given the host of test challenges, and to prevent a cat and mouse game, designing dedicated hardware test circuitry to assist normal chip use seems a befitting solution.

With the advent of Artificial Intelligence (AI) (especially Machine Learning (ML)), a new arsenal is available to test engineers. Use of ML in testing includes screening of test patterns, fault detection and classification, and automating test - circuit synthesis and rehearsal. ML is at the heart of new - age automatic test equipment. Highlights AI/ML's effect on VLSI design & test are also presented by domain.

11.1 Automated Testing Solutions

The semiconductor industry is rewarding, but it is facing new opportunities and challenges. Technologies such as artificial intelligence, 5G, and the internet of things are taking center stage in computing and communications systems. Applications like autonomous vehicles and smart cities will bring new opportunities for high - quality chips with reduced size, power, and cost. Extensive verification of next - generation chips is necessary to keep up with aggressive scaling of technologies. Novel methods using machine learning are needed to keep up with the rapid growth of the design space.

Testing is a fundamental and essential task for every chip at every level of abstraction, design phase, and application. Production testing is crucial to find silicons that do not conform to the specification before deploying them to the customer. Each defect has its own physical manifestation on the chip and causes a corresponding unexpected response. There are generally two types of production tests: unit tests, which determine whether a device is good or fail, and enhanced tests, which aim to specify the failing defect. As the chips grow more complicated, many of them can no longer be tested exhaustively. As a result, an effective test generation algorithm that guarantees quality or even correctness and a robust and efficient infrastructure for its integration into the design environment and yielding into a usable testcase and testprogram are essential.

Eqn 2: Automated Testing Solutions

$$ATS = f(C, S, T, A, R)$$

Where:

- ATS: Effectiveness or output of the Automated Testing Solution.
- C: Codebase or components under test.
- S: Test scripts and scenarios (automated).
- T: Test execution tools/platforms (e.g., Selenium, JUnit, Cypress).
- A: Automation framework features (e.g., CI/CD integration, reporting, parallelization).
- R: Results analysis and feedback loop (e.g., dashboards, alerts, ML models).

Tests are generated in parallel, with a variety of constraints. Some biases of the initial pseudo - random/testbench sequences could be filtered out at the input levels, thus gaining high coverage. Fixing this issue and generating efficient tests in a probabilistic manner can be done using genetic algorithms, which manipulate the input states and the procedure program as chromosomes. An average of over 80% stuck - at or delay fault coverage was achieved. To ensure that no relevant testbenches were missed in the design environment, some contextual constraints could be circumvented; algorithms were developed to generate input programs from multi - granularity ones that still retained the behavior of interest.

11.2. Real - Time Data Analysis

With the technological advancement of semiconductor chips, the integration densities of devices in a single semiconductor chip is rapidly increasing. The ever - evolving structure and additional complex components in semiconductor devices greatly enhance their performance but make it increasingly difficult to detect defects or defects. Conventionally, using machine learning (ML) to perform analysis on semiconductor failures is to train a model in batch mode. Most of the time the data is flushed to other analytics hardware, and ML guided failure analysis is conducted after a set of chips are processed. However, when the chips are being manufactured in a semiautomatic set of processes, a new type of failure occurs. The ML model that is previously trained can only detect the types of failures it has seen. Recently some ML techniques have been proposed to retrain the model online, as additional data points come in the early failure detection becomes more urgent. But how to adapt models designed for large batch data to a single - pass online fashion on semiconductor failure analysis is more challenging and more meaningful [3].

This work discusses novel quantile online learning to detect semiconductor failures online in a single batch. Classical regression is designed for a fixed distribution and can't deal with unseen data distributions. The novelty of this method is to take advantage of quantile regression which builds quantile functions adaptively on the observations for each identically distributed random variable. To handle the ambiguity problem with data that is not labeled, as many commonly used performance metrics are based on quantile models, this paper combines quantile regression with online learning. By taking advantage of binary encoding fashions of quantile prediction functions proposed previously, it is shown this combination generalizes classical quantile regression learning from batch mode to online fashion.

The proposed method is on a family of semiconductor device - level defects. They are FinFET bridge defect, GAA - FET bridge defect, GAA - FET dislocation defect. The first three device - level defects are thermally triggered hotspots. Conventional test sockets break during wafer probing considering the excellent electrical performance of devices in scenarios matching corners. The SEM imaging of a micro - bump in Chip 3 was also shared for this family of device - level defects. Further abnormal gates are shown to distribute more in Chip 3. Evaluation of multi - channels and scales of control are present for a public database: SECOM. To test the robustness and applicability of the proposed quantile online learning method, this work is on a public metallization defects problem: SECOM, which is used widely to evaluate different ML techniques.

12. Ethical Considerations in Machine Learning Applications

There is an inherent tension in the process of how things get constructed in the world of automated computer systems where machine learning makes their own decisions to classify things and carry out actions in response to them. On the one hand, automated action and a large measure of autonomy is what makes machine learning - based systems useful and powerful. As more sophisticated machine learning techniques are adopted in various industries and domains, concern regarding the ethical implications of their deployment and usages intensifies. Based on a scoping literature review on the notion of ethics in the study of automated computer systems, offers a machine learning - inspired understanding of the ethics of machine learning. While using this crudely articulated notion of machine learning ethics as a basis, analysis notes that much work remains to be done to advance a more incisive understanding of machine learning ethics [8]. By singing the praises of machine learning as a concept tool that allows organizations to anticipate the future and act on predictions in a preemptive manner, a jubilant response to and reinforcement of the celebratory discourse has been offered. However, this overly positive appraisal begs the question of whom machine learning 'serves'. Such pushback provides a more complex account of who is accountable for precision medicine and related conceptually derivative technologies. In addition, a convincing case is made about how to think critically about society scanning citizens faces well informed by machine learning.

Eqn 3: Ethical Considerations in Machine Learning Applications

$$E_{risk} = w_1 \cdot B + w_2 \cdot L + w_3 \cdot P + w_4 \cdot A^{-1}$$

Where:

- E_{risk}: Overall ethical risk score of the ML system (higher means more ethically problematic).
- B: Bias in data or model (measured via fairness metrics like disparate impact).
- L: Lack of transparency (e.g., opacity of decision-making).
- P: Privacy intrusion (risk of re-identification, data leakage).
- A: Accountability level (e.g., traceability, auditability) — higher accountability reduces risk.
- w_1, w_2, w_3, w_4: Weights representing the relative importance of each factor (domain-specific).

Machine learning classifiers are central decision makers in large measure, if not entirely, automated computer systems. As such, understanding machine learning and the underlying mathematics, engineering principles, and their implications has been an academic pursuit for years. Regardless of the technology underlying them, decisions taken in the world of machine learning inherently have a lacking of human agency and ethical accountability.

13. Regulatory Implications for Semiconductor Testing

The advanced GDT system is compatible with other TSK - based predictive tests. For example, one can define a TSK model per DTM yield or product confidence metric, and both types of models could be managed by the same GDT system. One can also digest the results of TSK Pareto prediction on

process parameters per sort. Not leading to design changes, but either process improvements or alternative mitigation plans. Finally, engaging Fault Classification (FC) on those tests, so that different and real time responses can be triggered, whether the anomaly produces a fail or is processed for Possible Failing Products (PFP) analysis.

The TSK approach to modelling advanced IC test measurements and responses has shown appealing results. Given the complex nature of modern devices, no exhaustiveness can be justified on the number of tests or the characteristics of the metrics. Tests remain state - of - the - art when defined by a TSK technique at a given frequency, as one needs the MDA state robust under fabrication and packaging variations and to engage the fault classification on such cases [5]. These tests could evolve to some sort of deep learning methodology in the future.

Another candidate for future improvement is the GDT system, the main results of which show that in some situations good operating points of the DTM models exist when not controlling the problematic chip class percentages. Other predictive models could be defined per DTM metric: relevant test selections could be anticipated by engaging the DTM prediction methodology. Yet the positive results concerning the non - compromised safety net show the high capacity and level of automation of the GDT system. Extreme 2D metal triplets gating and added procedure on the damaged devices would lead to a different supervised learning modelling approach, such as the one used for using statistical attributes and regression trees at wafer - level data [3]. The major focus for immediate deployment in current manufacturing would be in advanced gates new operation points, with models tending less to chronic chip classification and (80%) while leading surgeries capable of changing or replacing many gate in run.

14. Industry Collaborations and Partnerships

In various industries, collaboration between semiconductor manufacturers, technology suppliers, and research institutions is crucial. The semiconductor industry in Taiwan boasts several leading semiconductor companies and an advanced supply chain for semiconductor equipment, materials, and design. Thus, the industry is an appropriate work area for case studies on AI applications. Given that the semiconductor industry is rapidly transforming through the introduction of AI solutions, there are many opportunities for further research. Most semiconductor manufacturers and related vendors are seeking AI - related personnel, showing that the market for this topic is hot. Research in this area is highly datacentric and can employ various data science methodologies. In addition, case studies can be pursued on multiple levels, including the entire industry, organizations, systems, and processes. All these aspects contribute to increasing both theoretical and practical relevance in this research topic [9].

The diversity of both the participating actors and data sources offers many interesting research opportunities. Interestingly, there is no coherent theoretical framework to understand the ecosystem of AI applications in the semiconductor manufacturing context. Most prior research has focused on only one dimension or industry actor. For example, some research has addressed different AI application areas in semiconductor processes and value chains. Others have elaborated on issues related to AI in one area of the industry, such as investment in AI or its value creation potential [2]. In sum, there has been little exploration of how semiconductor manufacturers interact with consulting firms, software providers, universities, and R&D centers regarding AI applications. Importantly, no empirical research exists to investigate if such collaboration leads to systemic success. How data is created, shared, and processed is significant in understanding this multi - actor ecosystem for AI applications.

15. Economic Impact of Advanced Testing Frameworks

Implicitly towards reducing the risks associated with failure of newly qualified devices, there is further examination of the entire testing process, including the information and cost - effectiveness of the approach adopted. The extent of the intelligent scheduling of preferred methodologies to use in new environments is substantial, but other areas for progress should be expected. The overall cost of testing would be reduced if it was possible to predict failure before a large amount of expenditure was incurred. The risk of an expensive failure is a target here. The approach is very general and not tied to a single personality. The techniques mentioned are used only to illustrate the point made. Teams of engineers become very adept at running through knock - down tests and following established procedures, taking no account of passing evidence or device history [5]. For any candidate test, an appropriate and sufficient amount of data passed, and failed distributions are examined and passed to the data conversion stage. Discussion of the distribution shapes assumed is presented. The conversion relies on estimates of common parameters for both pass and fail populations, and the score needs to be modified to accommodate this for tests showing different population shapes. Assuming a preliminary defined target cost per loss of persuading a rejection, decisions are made on what to suck to meet the cost. In FLEX, Planning is Complete, but sometimes uncertain [1]. Testing decisions generally sit within the Operations Environment. They target part costs, production release/holding periods, and reworking constraints. The simplest options are economic policies compliant with the analytical model, but greater savings must be gained by more computationally intensive mining of huge associated distribution data sets in real time. Such processes follow acceleratory coding of loss tolerance, including devising additional economic costs for suspected unit performance. In this dull scenario of daily costly computation, outcomes of pure analyst fiction must better predicted than hindsight, for earlier production returns, disposable reserves, and routine monitoring of performance trends. Progress, initially at least, must fully capitalize on the extensive expertise and enforced control elsewhere in Operations.

16. Conclusion

VLSI technology is a highly challenging field in which to develop and test advanced IC devices, due to the extreme reliability requirements of next - generation devices,

aggressive scaling, and the multi - faceted nature of device technologies. Current testing frameworks are dominated by tools that perform sampling, which are necessary to achieve the needed test quality, and by coupled physical/simulation tools that provide device - in - the - loop circuit test scenarios. A key challenge is the exhaustiveness of testing, which in general cannot be achieved. Additionally, first - order testing frameworks cascade multiple tests for stress and excursions. A need exists for a comprehensive, end - to - end testing framework that can integrate machine learning into device testing to overcome these existing limitations. Nevertheless, current machine learning methods could not be adapted directly to IC testing due to the difficulty of defining reward functions. Extensive research is needed to provide high - fidelity device models that are amenable to an adaptable, automated, physics - guided testing and design - for - manufacturing methodology.

The objectives of the testing frameworks include the following goals. (1) To use machine learning to cast the most difficult testing challenges into a framework, where a heuristic obtain - a - device for a given scenario is defined in terms of a simulation tool, while the simulator itself is fixed. The outputs of a simulation tool are fed to a simulation engine that selects which excitation to assess. This engine allows for adaptation by feeding a single value under conditions that differ from those used to compile sampled tests. (2) To apply machine learning to predict regions in the testing criteria space where sensor - testing events are possible. (3) To apply machine learning to extend the capability of conventional tests to enable control over greater portions of the manufactured devices.

Extensive computational and theoretical work details the machine learning approach for testing. Anticipating that the most difficult aspects of modeling the testing scenario to be addressed will be the selection of exciting devices, initial efforts are focused on integration with multi - level static analysis tools and accomplished testing methods. While preliminary tests reveal promise, extensive theoretical modeling is still needed before these can be applied to avionics IC devices. In conjunction with this is exploratory computation to implement a basic, monolithic simulation/intelligence framework for devices that can run instability tests.

References

[1] Karthik Chava, "Machine Learning in Modern Healthcare: Leveraging Big Data for Early Disease Detection and Patient Monitoring", International Journal of Science and Research (IJSR), Volume 9 Issue 12, December 2020, pp.1899 - 1910, https: //www.ijsr.net/getabstract. php?paperid=SR201212164722, DOI: https: //www.doi. org/10.21275/SR201212164722
[2] Data Engineering Architectures for Real - Time Quality Monitoring in Paint Production Lines. (2020). International Journal of Engineering and Computer Science, 9 (12), 25289 - 25303. https: //doi. org/10.18535/ijecs. v9i12.4587
[3] Vamsee Pamisetty. (2020). Optimizing Tax Compliance and Fraud Prevention through Intelligent Systems: The Role of Technology in Public Finance Innovation. International Journal on Recent and Innovation Trends in Computing and Communication, 8 (12), 111–127. Retrieved from https: //ijritcc. org/index. php/ijritcc/article/view/11582
[4] Mandala, V. (2019). Integrating AWS IoT and Kafka for Real - Time Engine Failure Prediction in Commercial Vehicles Using Machine Learning Techniques. International Journal of Science and Research (IJSR), 8 (12), 2046 - 2050.
[5] Ghahramani, M., Qiao, Y., Zhou, M., O'Hagan, A., & Sweeney, J. (2020). AI - based modeling and data - driven evaluation for smart manufacturing processes. IEEE/CAA Journal of Automatica Sinica, 7 (4), 1026–1037. https: //doi. org/10.1109/JAS.2020.1003114