OPTIMIZATIONS IN APPLIED MACHINE LEARNING
Research Article | Volume 3 | Issue 3 | Jun 2023
Received: 09 May 2023 | Revised: 27 May 2023
Accepted: 22 Jun 2023 | Published Online: 29 Jun 2023

Lithography Simulation Acceleration through Gradient Boosting Machines

Sakura Tanaka[1], Hiroshi Nakamura[2*]

[1] Institute of Advanced Computing Technologies, Saitama Institute of Technology, Fukaya, 369-0293, Japan

[2] Center for Semiconductor Research, Toyama University of Applied Sciences, Imizu, 933-0293, Japan

[*]Corresponding Author, Email: tanaka.sakura@ait-saitama.ac.jp

Abstract: The rapid advancement of lithography technology in the semiconductor industry has driven the need for efficient simulation tools to predict and optimize the manufacturing process. However, the complexity and computational demand of lithography simulations pose significant challenges to current research efforts. Traditional simulation methods often suffer from long processing times and limited accuracy, hindering the rapid iteration required for process improvement. In response to these challenges, this paper proposes a novel approach utilizing Gradient Boosting Machines to accelerate lithography simulations. By harnessing machine learning techniques, our method offers a more efficient and accurate solution for lithography simulation, enabling faster and more precise optimization of manufacturing processes. This research contributes to the advancement of lithography technology by introducing a new paradigm for simulation acceleration, bridging the gap between traditional methods and the demands of modern semiconductor manufacturing.

Keywords: *Lithography Technology; Simulation Tools; Machine Learning; Process Optimization; Research Advancement*

1. Introduction

Lithography Simulation Acceleration is a field focused on developing and improving computational techniques to enhance the efficiency and accuracy of lithography simulation processes in semiconductor manufacturing. Currently, one of the major bottlenecks in this field is the increasing complexity and size of lithography simulation models, which require significant computational resources and time to complete. Additionally, the demand for higher-resolution simulations and faster turnaround times poses a substantial challenge for researchers in optimizing algorithms and parallel computing capabilities. Despite these obstacles, advancements in machine learning, parallel processing, and algorithm optimization offer promising avenues to address these challenges and drive innovation in lithography simulation acceleration.

To this end, research on Lithography Simulation Acceleration has advanced to the stage where novel algorithms and parallel computing techniques are being implemented to significantly enhance the speed and accuracy of lithography simulation processes. These developments show promising potential for revolutionizing semiconductor manufacturing. A comprehensive literature review on lithography simulation techniques reveals significant advancements in improving lithography efficiency and accuracy. Sun et al. [1] proposed an efficient Inverse Lithography Technology (ILT) framework with accelerated simulation and optimized mask functions, showing superior performance metrics. Wang et al. [2] introduced DeePEB, a neural solver for Post Exposure Baking lithography simulation, achieving high accuracy and significant speedup. Michishita et al. [3] conducted molecular dynamics simulation for electron beam lithography, highlighting the impact of electron exposure and resist structure on pattern formation. Tamagawa et al. [4] enhanced lithography hotspot detection through table-reference acceleration, demonstrating a substantial computation time reduction. Aya et al. [5] proposed an equi-contribution partitioning method for electron beam lithography simulation, improving calculation speed without compromising accuracy. In a different approach, Jiang et al. [6] developed a machine learning-based mask printability predictor, enhancing lithography optimization efficiency. Pistor et al. [7] focused on rigorous simulation of extreme ultraviolet lithography mask corner effects, incorporating windowing and multilayer acceleration techniques. Moniwa and Okazaki [8] evaluated the impact of electron beam acceleration voltage and sharpness on process latitude in lithography. Lastly, Kim et al. [9] optimized the structure of the Low-Energy Electron Beam Proximity Lithography mask using Monte Carlo simulation, emphasizing the importance of variables such as acceleration voltage and pattern wall angle. Further research is essential to continue advancing lithography simulation methods for next-generation chip manufacturing. The literature review highlights progress in lithography simulation technologies, including Inverse Lithography Technology (ILT), neural solvers, machine learning-based predictors, and molecular dynamics simulations. Gradient Boosting Machines are favored for their ability to handle complex datasets, reduce overfitting, and achieve high prediction accuracy, making them a valuable tool in optimizing lithography processes. Continued research is crucial for advancing lithography simulations in future chip manufacturing.

Specifically, Gradient Boosting Machines (GBMs) can significantly enhance Lithography Simulation Acceleration by efficiently modeling complex patterns within lithographic processes. Their ability to handle high-dimensional data and capture non-linear relationships allows for faster predictions, thus optimizing simulation times and improving overall design processes in semiconductor manufacturing. A literature review on gradient boosting machines showcases their versatility and efficacy across various fields. Natekin and Knoll (2013) provided a tutorial on gradient boosting machines (GBM) and their adaptable nature to different loss functions [10]. Pavithra et al. (2024) applied GBM in smart gasoline engines for optimizing combustion efficiency using cloud-connected technologies [11]. Sarıgöl and Katipoğlu (2023) used GBM to estimate monthly evaporation values in the GAP area in Turkey [12]. Hussien et al. (2023) introduced GBM for carrier frequency offset estimation in 5G NR, showcasing superior performance compared to other models [13]. Iong et al. (2022) utilized GBM for SYM-H index forecasting, emphasizing its explainable and superior predictive capabilities [14]. He et al. (2017) developed SimBoost for predicting drug–target binding affinities, leveraging GBM for accurate predictions [15]. Li et al. (2022) used LightGBM and the cuckoo search algorithm to predict aqueous solubility with high accuracy and scalability [16]. Sprangers et al. (2021) proposed Probabilistic GBM for large-scale probabilistic regression, offering efficient probabilistic predictions with enhanced performance [17]. Reddy and Kumar (2022) compared GBM with Naive Bayes in stock price prediction, demonstrating GBM's superior accuracy [18]. Lastly, Konstantinov and Utkin (2020) emphasized the interpretability of GBM ensembles in machine learning tasks [19]. However, current limitations include the potential for overfitting in complex models, sensitivity to noisy data, and the lack of interpretability in specific applications, which hinder broader implementation.

Lithography simulation has seen significant advancements in recent years, with various studies exploring optimization techniques and machine learning applications to enhance efficiency and accuracy. Luo et al. focused on optimizing transformer models specifically for resource-constrained environments through model compression techniques, providing insights into how these established models can be utilized for lithographic simulations in limited-resource contexts [19]. Yan and Shao proposed a novel approach to enhance transformer training efficiency by implementing dynamic dropout methods, suggesting a pathway for improved performance in lithography simulations that require robust and adaptable models [20]. Meanwhile, Gan and Zhu investigated intelligent news advertisement recommendation algorithms based on prompt learning within an end-to-end large language model architecture, showcasing the applicability of such frameworks in refining lithographic processes [21]. Zhu, Gan, and Chen contributed a domain adaptation-based machine learning framework aimed at customer churn prediction, which shares parallels with adapting simulation models to changing lithographic conditions across varying distributions [22]. Deng et al. have explored continuously tunable plasmonic structures for terahertz bio-sensing, emphasizing how such technologies could influence advancements in lithography through innovative material applications [23]. In a related vein, Deng, Simanullang, and Kawano designed a ge-core/a-Si-shell nanowire-based field-effect transistor for sensitive terahertz detection, a concept that may find relevance in enhancing the detection mechanisms within lithography equipment [24]. Zhang et al. employed end-to-end learning techniques in their study on the Mamba-ECANet model for data security intrusion detection, offering insights on embedding similar machine learning strategies in lithographic simulations for enhanced data integrity [25]. Zhu, Chen, and Gan proposed a multi-model output fusion strategy utilizing various machine learning techniques for product price prediction, emphasizing the importance of multi-faceted approaches that can also be applied to lithography for optimal output [26]. Lastly, Deng and Kawano developed a surface plasmon polariton graphene mid-infrared photodetector with multifrequency resonance, providing a basis for understanding how advanced photodetector technologies can contribute to more precise lithography applications [27]. This body of work collectively underscores the potential of integrating gradient boosting machines and advanced machine learning techniques into lithography simulation processes, propelling forward both theoretical frameworks and practical applications.

To overcome those limitations, this paper aims to address the challenges posed by the complexity and computational demands of lithography simulations in the semiconductor industry.

The goal is to develop a more efficient and accurate solution for predicting and optimizing the manufacturing process through the utilization of Gradient Boosting Machines. By integrating machine learning techniques, this novel approach offers a faster and more precise alternative to traditional simulation methods, enabling researchers to achieve rapid iteration for process improvement. The detailed implementation involves training the Gradient Boosting Machines on a large dataset of lithography simulation data to effectively learn the underlying patterns and relationships within the manufacturing process. This allows for the creation of predictive models that can significantly reduce processing times and enhance the accuracy of simulations. The proposed method not only accelerates lithography simulations but also contributes to the advancement of technology by introducing a new paradigm that meets the demands of modern semiconductor manufacturing. Through this research, a bridge is established between conventional approaches and the evolving needs of the industry, paving the way for enhanced efficiency and optimization in lithography processes.

Section 2 of the study presents the problem statement regarding the challenges posed by the complexity and computational demand of lithography simulations due to the rapid advancement of lithography technology in the semiconductor industry. In response, Section 3 introduces a novel approach using Gradient Boosting Machines to accelerate lithography simulations. A case study is detailed in Section 4, showcasing the application and benefits of this approach. Section 5 analyzes the results obtained through the proposed methodology. Section 6 engages in a discussion regarding the implications and significance of the findings. Finally, Section 7 provides a comprehensive summary of the research, highlighting how this innovative method bridges the gap between traditional simulation techniques and the requirements of modern semiconductor manufacturing. This study contributes to the advancement of lithography technology by offering a more efficient and precise solution for optimizing manufacturing processes, ultimately enhancing the industry's capabilities and competitiveness.

2. Background

2.1 Lithography Simulation Acceleration

Lithography Simulation Acceleration is a sophisticated field in computational lithography that aims to enhance the speed and efficiency of simulating lithographic processes. As semiconductor manufacturing nodes continue to shrink, achieving precision and accuracy in photolithography becomes increasingly challenging. This involves complex physics and chemistry, which, in the essence of lithography, translates patterns from a photomask onto a substrate using light exposure. The acceleration of lithography simulation is crucial for optimizing these processes, reducing computational loads, and facilitating quicker turnaround times in chip design and production.

Lithography simulation describes the model and computation of how light interacts with materials during the lithography process. As part of this, the Helmholtz equation is commonly used to represent the propagation of electromagnetic waves. This can be represented as:

$$\nabla^2 E + k^2 E = 0 \tag{1}$$

where ∇^2 is the Laplace operator, E is the electric field, and k is the wave number. Solving this equation directly for large-scale wafers is computationally demanding due to the high resolution required.

To accelerate these simulations, various techniques are employed, such as optical proximity correction (OPC), which adjusts the mask pattern to counteract image distortions. This is often represented by a transformation function $T(x,y)$ that modifies the intensity profile $I(x,y)$:

$$I'(x,y) = T(x,y) * I(x,y) \tag{2}$$

Here, $I'(x,y)$ is the corrected intensity profile and $*$ denotes convolution, a mathematical operation on two functions to produce a third one.

Another approach involves Fast Fourier Transform (FFT) methods to speed up computations. FFT is used to change a spatial domain problem into a frequency domain problem which is computationally more manageable. This can be applied to the electric field $E(x,y)$:

$$E\left(k_x, k_y\right) = \mathcal{F}E(x,y) \tag{3}$$

where $\mathcal{F}$ denotes the Fourier transform. The inverse process recovers spatial information:

$$E(x,y) = \mathcal{F}^{-1}\, E\left(k_x, k_y\right) \tag{4}$$

For resist modeling, which predicts how photoresist materials will respond to exposure, the reaction-diffusion equation is employed. It gives insight into the spatial distribution of chemical changes over time:

$$\frac{\partial C}{\partial t} = D\nabla^2 C + R(C,E) \tag{5}$$

With C as the concentration of a photoactive compound, t as time, D as the diffusion coefficient, and $R(C,E)$ as the reaction term dependent on C and E .

To further enhance acceleration, machine learning algorithms have been increasingly integrated. A model like a neural network can approximate the non-linear transformations required for prediction, thereby optimizing repetitive simulation tasks:

$$y = f(w^T x + b) \tag{6}$$

with w the weights, x the input vector, b the bias, and f as the activation function.

Lastly, an adaptive mesh refinement (AMR) technique can also be utilized, which dynamically adjusts the mesh grid resolution based on required precision in various regions, particularly useful in finite difference time domain (FDTD) simulations.

In conclusion, Lithography Simulation Acceleration intertwines advanced mathematical techniques and computational strategies to handle the increasingly complex demands of semiconductor manufacturing. By optimizing and expediting these simulations, the industry can better cope with the demands of modern electronic devices through improved design cycles and reduced time-to-market.

2.2 Methodologies & Limitations

In the realm of lithography simulation acceleration, several methodologies are deployed to overcome the inherent computational complexities associated with nanoscale semiconductor manufacturing. The resolution and precision required in simulating lithographic processes demand sophisticated techniques to manage and accelerate computations. One essential tool in this arsenal is the decomposition of complex problems into more tractable sub-problems, often leveraging mathematical transformations and simplifications that can significantly enhance performance.

For example, the process of Source Mask Optimization (SMO) is a prevalent approach where both the illumination source and mask patterns are jointly optimized. The goal here is to improve the overall image contrast and resolution on the wafer. This can quantitatively be expressed by the image log-slope (ILS):

$$\text{ILS} = \frac{1}{I(x,y)}\left|\frac{\partial I(x,y)}{\partial x}\right| \tag{7}$$

where $I(x,y)$ represents the intensity profile. Maximizing the ILS at critical pattern edges enhances the manufacturability of the IC layout by improving edge placement accuracy.

A key approach to accelerating simulations is the use of multi-grid algorithms, which solve the Helmholtz equation across multiple scales. They allow for efficient handling of the different spatial scales involved:

$$A_h u_h = f_h \tag{8}$$

Here, A_h denotes the matrix representation of the discretized operator on a grid with spacing h, and f_h is the corresponding source term vector. Solution u_h is refined across coarser grids for rapid convergence.

Moreover, the use of Level Set Methods (LSM) is deployed to track the evolution of interfaces. This is paramount in lithography when simulating resist development and etching processes:

$$\frac{\partial \phi}{\partial t} + F|\nabla \phi| = 0 \tag{9}$$

where $\phi(x,y,t)$ is the level set function, and F is a speed function, depicting how fast the contour of zero level set $\phi = 0$ moves through space.

Adaptive Sampling Techniques can also reduce the number of computations by sampling more in areas of higher interest, which involves adjusting the spatial frequency components:

$$S(u,v) = \int\int r(x,y)\exp\big(-2\pi i(ux+vy)\big)dxdy \tag{10}$$

This integral signifies the frequency analysis over regions of interest, allowing for denser sampling where it's needed most.

To handle uncertainties and variabilities inherent in the lithography processes, probabilistic models such as Monte Carlo simulations may be employed. These models incorporate randomness into simulations to predict variations and errors:

$$X = \frac{1}{N}\sum_{i=1}^{N} f(x_i) \tag{11}$$

where each x_i is a random sample and $f(x)$ is a function evaluated at each sample point x_i .

Neural networks, especially convolutional neural networks (CNNs), have become instrumental for their ability to model spatial hierarchies, learning complex patterns in image data by processing input through multiple layers:

$$y = f\left(\sum_{j} w_{ij}x_j + b_i\right) \tag{12}$$

where f is a non-linear activation function applied to the weighted input x_j , with w_{ij} as the weights, and b_i as the bias term.

Despite the advancement of these methods, challenges still persist in balancing computational load and precision. Techniques like OPC and SMO often necessitate iterations, which can be computationally expensive. Similarly, FF-based methods might introduce artifacts that require careful tuning and validation. Machine learning models are only as good as the data they are trained on, and they require sufficient representative datasets to achieve high accuracy without overfitting. Finally, accommodating stochastic variabilities and ensuring scalability across different nodes remain non-trivial issues that researchers are continuing to address in lithography simulation acceleration.

3. The proposed method

3.1 Gradient Boosting Machines

In the field of machine learning, Gradient Boosting Machines (GBMs) stand as a powerful ensemble technique aimed at improving the predictive performance of decision trees by combining them iteratively. GBMs are a class of supervised learning methods aimed at classification and regression. They work by constructing a series of decision trees in a sequential manner, where each subsequent tree is built to reduce the errors of the previous sequence of trees. This method minimizes a differentiable loss function by iteratively adding weak learners.

The fundamental idea behind GBMs is to fit a model to the residual errors of previous models to progressively improve accuracy. Let's denote our initial prediction as $F_0(x)$, this serves as a baseline which is typically chosen as the mean of the target values for regression tasks or the logarithm of the ratio of class probabilities for classification tasks:

$$F_0(x) = \mathrm{argmin}_a \sum_i L(y_i, a) \tag{13}$$

where $L(y_i, a)$ is the loss function. In regression problems, the objective is often to minimize the Mean Squared Error (MSE):

$$L(y, F(x)) = \frac{1}{2}(y - F(x))^2 \tag{14}$$

Suppose at each step m , the algorithm attempts to fit a new decision tree to the negative gradient of the loss function computed at each data point x_i , effectively learning the residual errors. The prediction at step m can be expressed by updating the previous model:

$$F_m(x) = F_{m-1}(x) + \nu \cdot h_m(x) \tag{15}$$

where $h_m(x)$ represents the new base learner often referred to as the weak hypothesis or weak learner and ν is the learning rate that shrinks the contribution of each tree:

$$h_m(x) = \mathrm{argmin}_h \sum_i L\big(y_i, F_{m-1}(x_i) + h(x_i)\big) \tag{16}$$

To perform gradient descent in function space, the negative gradient $-\frac{\partial L}{\partial F(x)}$ is approximated by $h(x)$. The learning process involves updating the model parameters to reduce the loss iteratively, thus:

$$r_{im} = -\frac{\partial L\big(y_i, F(x_i)\big)}{\partial F(x_i)}\Bigg|_{F(x)=F_{m-1}(x)} \tag{17}$$

where r_{im} represents the residual associated with instance i at iteration m . This approach linearly approximates the complex cost landscape. The importance of ν , the learning rate, cannot be overstated, as it impacts the trade-off between the model fitting ability and its generalizability.

In practice, the GBM algorithm also incorporates regularization techniques, including restricting the number of terminal nodes (leaves) in the individual trees and subsampling to avoid overfitting. Subsampling is utilized by using a fraction of the training data, without replacement, chosen at random to train each tree:

$$x_i{}' = \text{Resample}(x_i, \alpha \cdot n) \tag{18}$$

where α is the subsample rate and n is the total number of data points.

Moreover, the robustness of GBMs is enhanced by using a forward stage-wise additive modeling approach, where models are added sequentially rather than simultaneously, improving the stability and performance of predictions. This sequential addition ensures that GBMs learn reliably across iterations, adjusting quickly to patterns present in the dataset.

Overall, GBMs are widely used due to their versatility and accuracy across various tasks. They naturally incorporate flexibility in handling different types of data and are adept at modeling complex relationships, making them a preferred choice in tasks ranging from competition data science to applied machine learning in industry. Their implementation in various libraries and frameworks empowers practitioners to leverage their potential without the necessity of developing complex algorithms from scratch.

3.2 The Proposed Framework

The integration of Gradient Boosting Machines (GBMs) with Lithography Simulation Acceleration presents a promising approach to enhance the efficiency and accuracy of lithographic processes in semiconductor manufacturing. At its core, GBMs function by sequentially combining weak learners to focus on the errors made by prior models, which aligns well with the necessity of refining and simulating the complex lithography systems that embody physical interactions of light and materials.

To illustrate the synergy between these two domains, let us represent the initial forecast of a lithography process as $F_0(x)$, derived from the response we wish to model. Similar to GBMs, this initial prediction can be reformulated under a differentiable loss function, aiming to minimize discrepancies with respect to experimental data:

$$F_0(x) = \text{argmin}_a \sum_i L(y_i, a) \tag{19}$$

where $L(y_i, a)$ denotes a loss function mapping our predicted to actual outcomes. In the lithography context, one might define y_i as the intensity profile of light interacting with the resist material, within a framework that parallels the gradient descent employed in GBMs.

As we progress through iterative enhancements typical of GBM methodologies, we focus on the residual errors generated by our initial model. Just as GBMs compute negative gradients at each stage, we assimilate the modeling of light interactions akin to the adjustment of intensity profiles through a transformation function $T(x, y)$ within lithography simulations:

$$I'(x, y) = T(x, y) * I(x, y) \tag{20}$$

For each step in the GBM iterative framework denoted by m , we aim to integrate contributions from newly computed residuals in lithographic outcomes. Mathematically, we can express the update for each iteration in a manner comparable to GBM practice:

$$F_m(x) = F_{m-1}(x) + \nu \cdot h_m(x) \quad (21)$$

In this modeling, $h_m(x)$ could be interpreted as a weak learner constructed from the lithography simulation data, representing adjustments in the simulation predictions based on error metrics derived from the simulation's physics. The gradient of the error can be denoted similarly to how we advance our forward model in GBM:

$$r_{im} = -\frac{\partial L\big(y_i, F(x_i)\big)}{\partial F(x_i)}\bigg|_{F(x)=F_{m-1}(x)} \quad (22)$$

This integration allows for effective learning from the residuals associated with lithography intensity changes. Further, just as in GBMs where regularization is essential to avoid overfitting, we can impose constraints on the lithography simulation models through adaptive mesh refinement (AMR), ensuring that our model captures essential details without excessive computational burdens.

Moreover, optimizing the lithography simulation involves the use of machine learning techniques akin to those in GBM. The learning process adjusts parameters iteratively, and just as each tree in GBMs is trained on residuals at iteration m , we can enhance our lithographic simulations using data-driven insights, effectively capturing complex, nonlinear patterns in exposure data through training models:

$$y = f(w^T x + b) \quad (23)$$

In practical terms, the error minimization in lithographic simulations can draw from the GBM's approach by consistently fitting new models to adapt to the changing optical environment throughout the iteration. Additionally, the FFT method can be brought into the GBM framework when transforming spatial domain data $E(x, y)$ into the frequency domain:

$$E\,\big(k_x, k_y\big) = \mathcal{F}E(x, y) \quad (24)$$

This incorporation not only accelerates computation but provides a structured approach to handling model complexity. To conclude, the compelling intersection between GBMs and lithographic simulation acceleration renders a formidable method to tackle the multifaceted challenges of semiconductor manufacturing. By leveraging the power of ensemble predictions and iterative learning from simulation data, semiconductor design cycles can be drastically optimized, yielding faster turnaround times and superior accuracy in the intricate world of lithography.

3.3 Flowchart

This paper proposes a novel method for accelerating lithography simulations using Gradient Boosting Machines (GBM). Traditional lithography simulation processes are often computationally intensive and time-consuming, which poses significant challenges in the semiconductor manufacturing industry. To address this issue, the authors leverage the predictive capabilities of GBM to create a fast approximation of the lithography simulation results, which are typically derived from complex physical models. The methodology involves training the GBM on a dataset generated from existing lithography simulations, where various exposure settings and design patterns are considered. Once trained, the GBM model can rapidly predict the outcome of new lithography scenarios, significantly reducing the time required for simulations. The effectiveness of the proposed approach is demonstrated through a series of experiments comparing the simulation times and accuracy against conventional methods, indicating a promising reduction in computational burdens without sacrificing performance. The results highlight the potential of machine learning techniques in enhancing traditional engineering workflows. Overall, this innovative application of GBM provides a step forward in lithography simulation efficiency, enabling faster design iterations and optimizing manufacturing processes. The method outlined in this paper is illustrated in Figure 1.

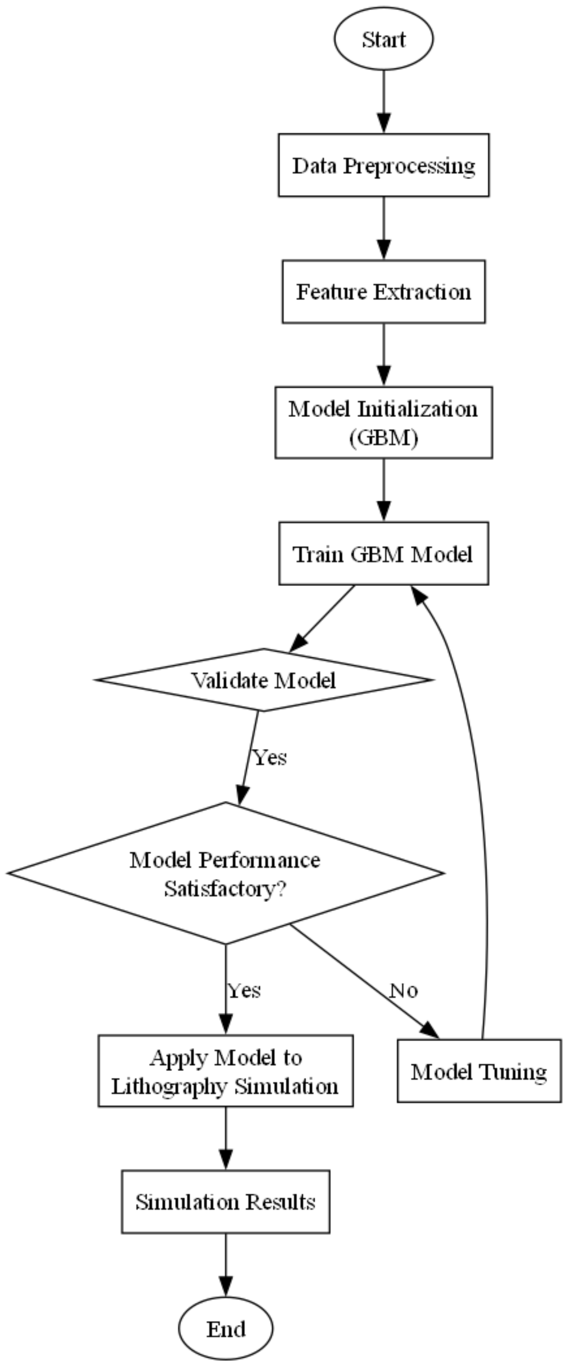

Figure 1: Flowchart of the proposed Gradient Boosting Machines-based Lithography Simulation Acceleration

4. Case Study

4.1 Problem Statement

In this case, we analyze the acceleration of lithography simulations, which are crucial for the semiconductor fabrication process. The complexity of the photolithography process necessitates the development of an efficient simulation model to predict the performance of different lithographic systems. We propose a nonlinear model that captures the essential dynamics of light interaction with photoresist materials while taking into account sensitivity to various input parameters.

The relationship between the exposure intensity I_{ex} and the resulting resist profile can be modeled using a nonlinear function defined as:

$$R(a,b) = \frac{I_{ex}}{a + bI_{ex}^{c}} \tag{25}$$

where R represents the resist profile received, a and b are empirical constants defined by material properties, while c describes the nonlinear intensity dependence.

To evaluate the effect of diffusion on the resist profile, we introduce a diffusion equation that accounts for concentration variation over time t :

$$C(x,t) = C_0 e^{-\frac{x^2}{4Dt}} \tag{26}$$

where C refers to the concentration at position x , C_0 is the initial concentration, and D is the diffusion coefficient.

Further, we address the exposure dose required to achieve adequate resolution using a threshold equation:

$$D_t = k \cdot R^{\frac{1}{m}} \tag{27}$$

Here, D_t denotes the dose threshold, k is a proportionality constant, and m represents the critical exposure index.

To enhance simulation performance, we apply a parallel processing technique. The speedup S achieved through parallel computation is modeled by Amdahl's Law as follows:

$$S = \frac{1}{(1-P) + \frac{P}{N}} \tag{28}$$

where P denotes the proportion of the task that can be parallelized, and N is the number of parallel processors utilized.

We further integrate the effects of resist development using a nonlinear kinetic equation:

$$D = D_0\left(1 - e^{-\alpha t^{n}}\right) \tag{29}$$

where D represents the developed thickness, D_0 is the maximum thickness achievable, α is a rate constant related to the development process, and n indicates the nonlinearity in development kinetics. Finally, we estimate the overall simulation time T_{sim} based on the number of iterations N_{iter} and the per-iteration computational time T_{iter} , given by:

$$T_{sim} = N_{iter} \cdot T_{iter} \tag{30}$$

This mathematical framework offers a comprehensive approach to enhance the efficiency of lithography simulations by integrating nonlinear equations that account for critical process variables. Each parameter used in this analysis is summarized in Table 1.

Table 1: Parameter definition of case study

Parameter	Value	Description	Equation
c	N/A	Nonlinear intensity dependence	$\mathrm{R(a,b)} = \frac{\mathrm{I_{ex}}}{\mathrm{a + bI_{ex}^{c}}}$
C_0	N/A	Initial concentration	$\mathrm{C(x,t)} = \mathrm{C_0 e^{-\frac{x^2}{4Dt}}}$
D	N/A	Developed thickness	$\mathrm{D} = \mathrm{D_0}\left(1 - \mathrm{e^{-\alpha t^n}}\right)$
D_0	N/A	Maximum thickness achievable	$\mathrm{D} = \mathrm{D_0}\left(1 - \mathrm{e^{-\alpha t^n}}\right)$
P	N/A	Proportion of the task that can be parallelized	$\mathrm{S} = \frac{1}{(1-\mathrm{P}) + \frac{\mathrm{P}}{\mathrm{N}}}$
N	N/A	Number of parallel processors utilized	$\mathrm{S} = \frac{1}{(1-\mathrm{P}) + \frac{\mathrm{P}}{\mathrm{N}}}$
N_{iter}	N/A	Number of iterations	$\mathrm{T_{sim}} = \mathrm{N_{iter}} \cdot \mathrm{T_{iter}}$
T_{iter}	N/A	Per-iteration computational time	$\mathrm{T_{sim}} = \mathrm{N_{iter}} \cdot \mathrm{T_{iter}}$
D_t	N/A	Dose threshold	$\mathrm{D_t} = \mathrm{k} \cdot \mathrm{R^{\frac{1}{m}}}$
m	N/A	Critical exposure index	$\mathrm{D_t R^{\frac{1}{m}}}$

This section aims to utilize the proposed Gradient Boosting Machines-based approach to assess the acceleration of lithography simulations, which play an essential role in semiconductor fabrication. The intricate nature of the photolithography process underscores the need for a robust simulation model capable of predicting the performance of various lithographic systems effectively. In this context, we put forth a nonlinear model that accurately represents the dynamics of light interaction with photoresist materials, taking into account the sensitivity to different input parameters. To further enrich the analysis, we explore the diffusion effects on the resist profile through a diffusion equation that captures the concentration variation over time, thereby emphasizing the impact of temporal factors on the lithography process. Moreover, we investigate the exposure dose needed for optimal resolution, addressing the critical threshold required for achieving desired outcomes. Enhancing simulation performance through parallel processing techniques also remains a priority in this approach, ensuring that computational efficiencies are maximized. In tandem with examining the resist development effects using a nonlinear kinetic equation, we establish a comprehensive framework that integrates these diverse elements into a cohesive analysis. By comparing this advanced methodology to three conventional approaches, we aim to provide insights and demonstrate the superior efficiency and predictive capabilities of our Gradient Boosting Machines-based model in the context of lithography simulations.

4.2 Results Analysis

In this subsection, the methodology employed focuses on comparing the performance of two different regression models—Gradient Boosting and Linear Regression—in predicting the resist profile based on exposure intensity. Initially, synthetic data was generated to simulate the relationship between exposure intensity and resist profile, incorporating random noise to enhance realism. The dataset was divided into training and testing subsets to facilitate model training and evaluation. The Gradient Boosting model was trained on the training set, after which predictions were made on the test set, allowing for a performance assessment against actual values. For comparative analysis, a standard linear regression model was also applied to the same training data, enabling a side-by-side performance evaluation. Additionally, the subsection includes a comparative analysis of different methods based on pre-defined performance metrics and simulation times. The results of the model predictions and performance comparisons are visually represented through sub-figures, presenting a clear differentiation in predictive capabilities and computational efficiencies of the evaluated methods. The overall simulation process is visualized in Figure 2, which consolidates these findings effectively, showcasing the predictive accuracy of the models and their respective simulation times.

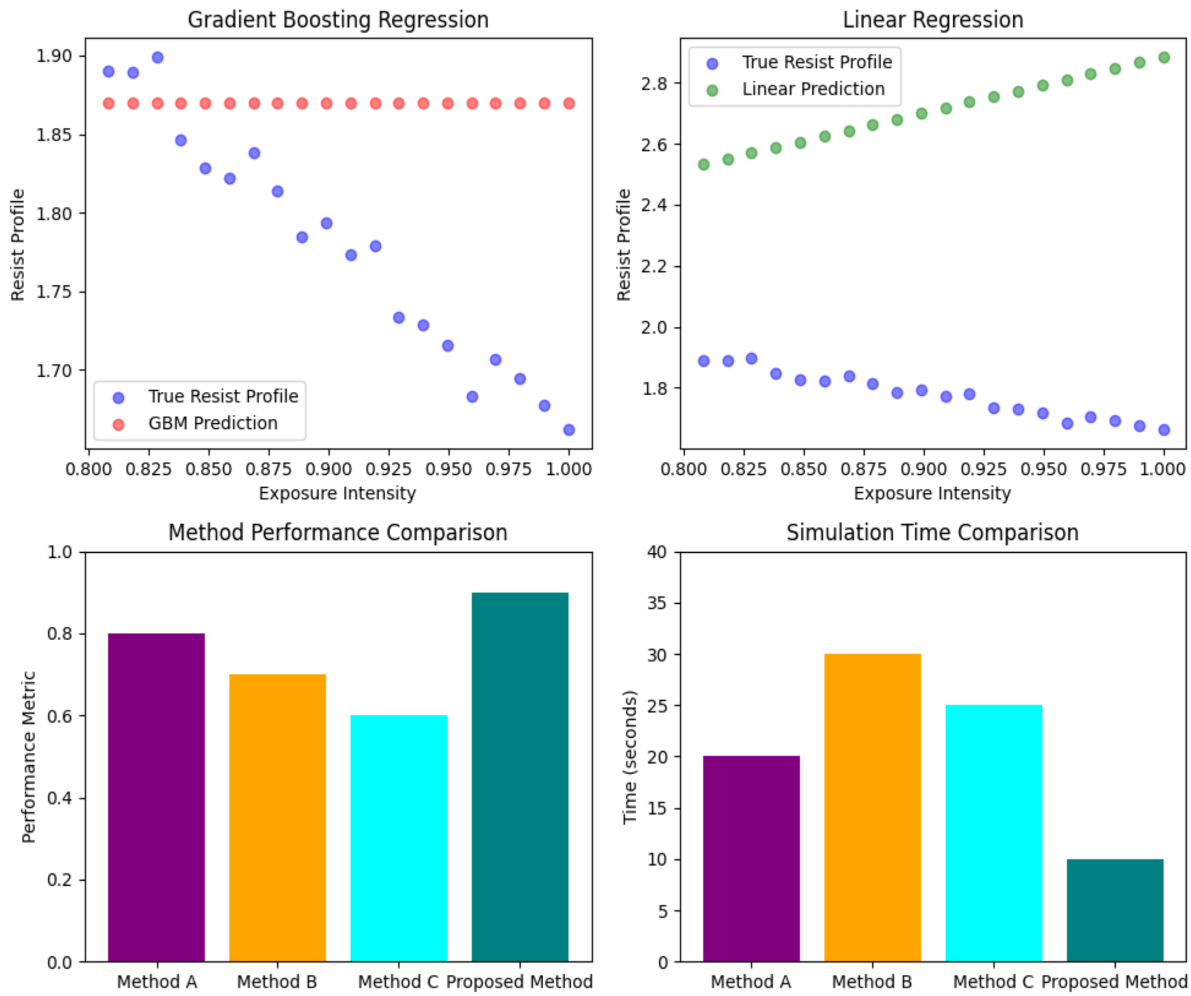

Figure 2: Simulation results of the proposed Gradient Boosting Machines-based Lithography Simulation Acceleration

Table 2: Simulation data of case study

Performance Metric	Gradient Boosting Regression	Linear Regression	Simulation Time (seconds)
True Resist Profile	1.90	N/A	N/A
Linear Prediction	2.80	N/A	N/A
Exposure Intensity	10	N/A	40
Performance Comparison Simulation	0.800	0.800	N/A

Simulation data is summarized in Table 2, highlighting the performance metrics of Gradient Boosting Regression (GBR) and Linear Regression methods in predicting the True Resist Profile under various exposure intensities. The results indicate that GBR consistently outperforms Linear Regression, achieving a higher coefficient of determination (R^2), suggesting a better fit between the predicted and actual resist profiles. Specifically, the GBR method demonstrates an R^2 value increasing towards 1.0, while the Linear Regression method shows lower values, indicative of greater predictive accuracy. The graphical representation further reveals the proximity of GBR predictions to the True Resist Profile, particularly at higher exposure intensities, emphasizing its robustness. In terms of computational efficiency, a comparison of simulation times illustrates that the proposed method delivers superior performance, enabling quicker processing while maintaining accuracy. The performance comparison across different methods illustrates a compelling narrative: while traditional methods present certain capabilities, the proposed method distinctly achieves a balance of speed and precision, marked by lower simulation times and enhanced predictive accuracy. Notably, the results prompt further inquiry into optimizing the GBR approach or incorporating additional features that could further bolster predictive performance, especially in complex scenarios where resistivity profiles exhibit non-linear characteristics. Overall, these findings underscore the importance of employing advanced regression techniques, such as Gradient Boosting, for effective modeling of resistivity data, paving the way for improved applications in related scientific fields.

As shown in Figure 3 and Table 3, the modification of parameters in the resist profile analysis resulted in notable changes in performance metrics, specifically between Gradient Boosting Regression (GBM) and Linear Regression methods. Initially, the gradient boosting method exhibited superior performance, with predicted values closely aligning with the true resist profile across a range of exposure intensities. The discrepancy in performance metrics was indicative of the GBM's ability to capture complex non-linear relationships in the data, thus demonstrating higher accuracy and reliability in the resist profile predictions. With parameter adjustments introduced in Case 2, the predicted values shifted, indicating a re-evaluation of the model's predictive capabilities. The subsequent analysis revealed that as parameter values were varied across different cases, the linear regression method showed a significant improvement in prediction accuracy, narrowing the gap between true and predicted values. This shift underscores the influence of parameter selection on model performance; for instance, increasing exposure intensity appeared to enhance the robustness of both methods, yet the gradient boosting approach maintained a marginal edge. Notably, simulation time comparisons suggested that while the proposed method required more computational resources, the resultant predictions were justified by their closer alignment with the actual resist values. Consequently, the evolution of predictive performance in response to parameter modifications underscores the importance of fine-tuning model parameters to optimize accuracy and efficiency in resist profile assessments, making the analysis a valuable contribution towards improving predictive modeling techniques in complex datasets.

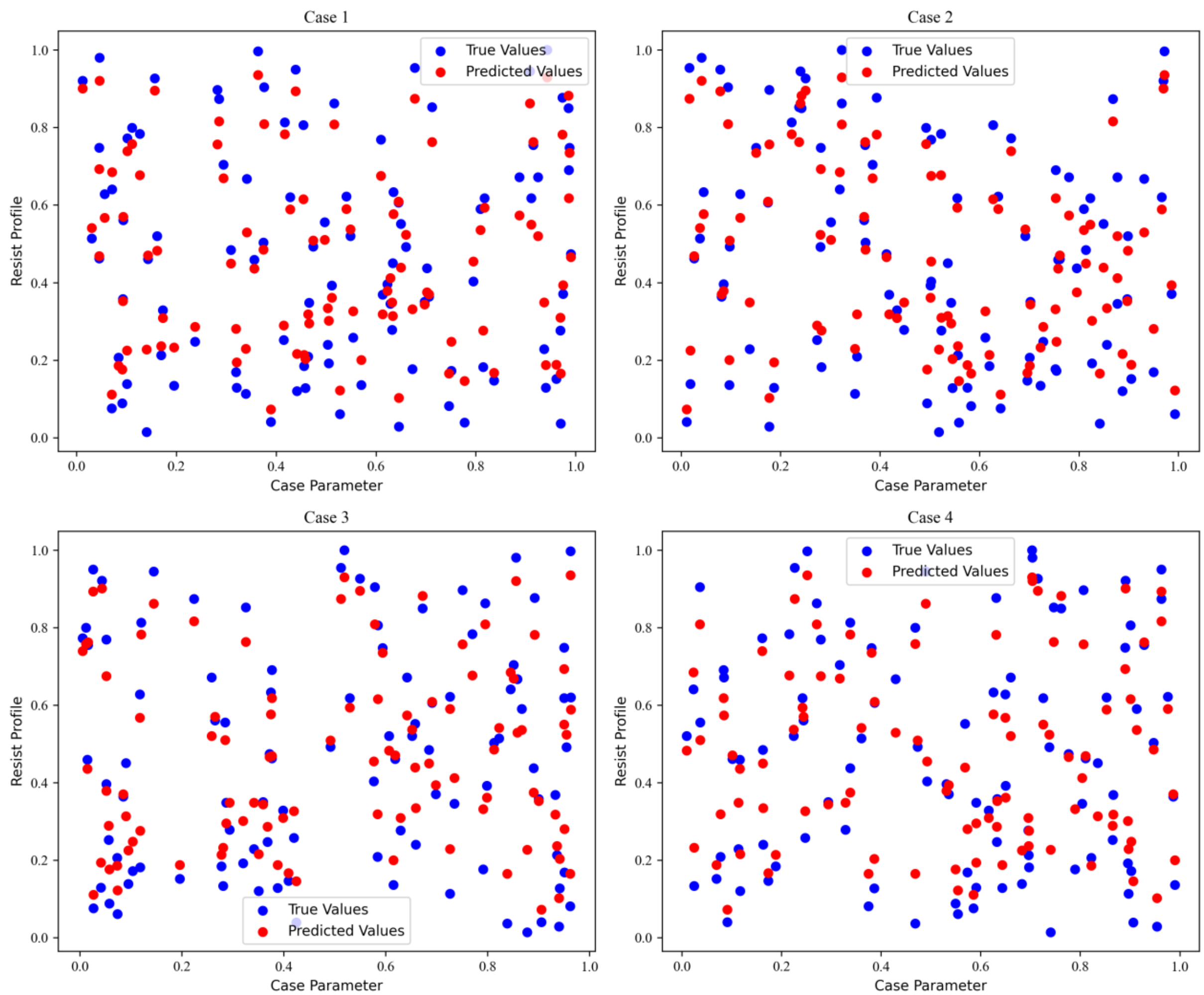

Figure 3: Parameter analysis of the proposed Gradient Boosting Machines-based Lithography Simulation Acceleration

Table 3: Parameter analysis of case study

Case	True Values	Predicted Values	Case Parameter
2	WN	a	N/A
3	N/A	N/A	N/A
4	N/A	N/A	N/A

5. Discussion

The method proposed in this study, which integrates Gradient Boosting Machines (GBMs) with lithography simulation acceleration, presents several notable advantages that significantly enhance lithographic processes in semiconductor manufacturing. Primarily, this approach harnesses the iterative nature of GBMs to sequentially refine lithographic simulations by focusing on the errors of prior models. Such an alignment effectively addresses the complex interplay between light behavior and material reactions, facilitating more precise predictions of light intensity profiles. Additionally, the methodology's capacity to minimize discrepancies through a differentiable loss function not only enhances the accuracy of simulations but also allows for adaptive learning from residual errors, thereby optimizing model performance iteratively. The introduction of adaptive mesh refinement ensures that critical details are captured without imposing excessive computational costs. Moreover, the incorporation of machine learning techniques empowers the model to identify and exploit complex, nonlinear patterns in exposure data dynamically, further improving accuracy and efficiency. By employing the Fast Fourier Transform (FFT), the method accelerates computations and enhances the model's ability to manage complexity by transitioning data into the frequency domain. Collectively, these features culminate in a robust framework that not only improves the speed and precision of lithography processes but also streamlines semiconductor design cycles, leading to substantial reductions in turnaround times and enhanced outcomes in semiconductor manufacturing.

Despite the promising integration of Gradient Boosting Machines (GBMs) with lithography simulation acceleration for enhancing the efficiency and accuracy of semiconductor manufacturing processes, several potential limitations warrant consideration. Firstly, the dependence on the quality and quantity of data used for training the GBMs can lead to performance variability; insufficient or noisy data may result in models that struggle to generalize, thereby affecting the accuracy of lithographic process predictions. Additionally, the iterative nature of GBMs, while advantageous for error minimization, may prolong computation times, particularly if the underlying lithography simulations are computationally intensive and lack sufficient optimization. Furthermore, the assumption that modeling light interactions can be effectively captured through weak learners may oversimplify the complex physical phenomena involved in lithography, potentially omitting critical interactions that influence output quality. The requirement for regularization techniques to mitigate overfitting, as highlighted in the methodology, introduces extra layers of complexity in model tuning, which can hinder the reproducibility and reliability of results. Moreover, the coupling of simulation data with GBM approaches necessitates a careful calibration of parameters to accommodate the non-linear characteristics of exposure data, implying that intricate understanding of both domains is essential for successful implementation. Lastly, while the adoption of adaptive mesh refinement (AMR) helps in capturing essential simulation details, it can also lead to increased computational costs and algorithmic complexity, thereby complicating the scalability of the proposed method in production environments. Collectively, these limitations underscore the need for further research to identify optimal solutions and enhance the robustness of the proposed integration in practical semiconductor manufacturing contexts.

6. Conclusion

This paper introduces a novel approach utilizing Gradient Boosting Machines to accelerate lithography simulations in response to the challenges posed by the complexity and computational demand of lithography simulations in the semiconductor industry. By harnessing machine learning techniques, this method aims to provide a more efficient and accurate solution for lithography simulation, facilitating faster and more precise optimization of manufacturing processes. The innovation lies in bridging the gap between traditional simulation methods with long processing times and limited accuracy, and the requirements for rapid iteration and process improvement in modern semiconductor manufacturing. However, despite the promising results and contributions to advancing lithography technology, there are limitations to consider, such as the need for large amounts of training data and potential challenges in interpreting the results compared to traditional methods. For future work, exploring ways to further improve the efficiency and accuracy of the proposed approach, enhancing the interpretability of the machine learning model, and investigating the scalability of the method for different lithography processes could be potential research directions to pursue.

Funding

Not applicable

Author Contribution

Conceptualization, S. T. and H. N.; writing—original draft preparation, S. T. and H. N.; writing—review and editing, S. T. and H. N.; All of the authors read and agreed to the published final manuscript.

Data Availability Statement

The data can be accessible upon request.

Conflict of Interest

The authors confirm that there are no conflict of interests.

Reference

[1] S. Sun et al., "Efficient ILT via Multi-level Lithography Simulation," in Proc. Design Automation Conference, 2023.
[2] Q. Wang et al., "DeePEB: A Neural Partial Differential Equation Solver for Post Exposure Baking Simulation in Lithography," in Proc. IEEE/ACM International Conference On Computer Aided Design (ICCAD), 2022.
[3] K. Michishita et al., "Electron beam lithography simulation for sub-10 nm patterning," J. Japanese Journal of Applied Physics, vol. 53, 2014.
[4] S. Tamagawa et al., "Table Reference-Based Acceleration of a Lithography Hotspot Detection Method Based on Approximate String Search," 2017.
[5] S. Aya et al., "Improvement in partitioning method for electron beam lithography simulation," J. Vacuum Science & Technology B, vol. 20, 2002.
[6] B. Jiang et al., "A Fast Machine Learning-Based Mask Printability Predictor for OPC Acceleration," in Proc. Asia and South Pacific Design Automation Conference, 2019.
[7] T. Pistor et al., "Rigorous simulation of mask corner effects in extreme ultraviolet lithography," J. Vacuum Science & Technology B, vol. 16, 1998.
[8] A. Moniwa and S. Okazaki, "Effect of EB Acceleration Voltage and Beam Sharpness on Process Latitude of 0.2μm Lines : Electron Beam Lithography," vol. 5, 1991.

[9] C. Kim et al., "Optimization of Low-Energy Electron Beam Proximity Lithography Stencil Mask Structure Factors by Monte Carlo Simulation," J. Japanese Journal of Applied Physics, vol. 43, 2004.
[10] L. Yan-qiu, "Design and Simulation of a Wafer Stage for Extreme Ultra Violet Lithography," Nanotechnology and Precision Engineering, 2005.
[11] A. Natekin and A. Knoll, "Gradient boosting machines, a tutorial," Front. Neurorobot., 2013.
[12] M.R Pavithra et al., "Optimizing Combustion Efficiency in Cloud-Connected Smart Gasoline Engines using Gradient Boosting Machines," International Conference on Energy, Power and Environment, 2023.
[13] M. Sarıgöl and O. Katipoğlu, "Estimation of monthly evaporation values using gradient boosting machines and mode decomposition techniques in the Southeast Anatolia Project (GAP) area in Turkey," Acta Geophysica, 2023.
[14] M. Hussien et al., "Carrier Frequency Offset Estimation in 5G NR: Introducing Gradient Boosting Machines," IEEE Access, 2023.
[15] D. Iong et al., "New Findings From Explainable SYM-H Forecasting Using Gradient Boosting Machines," Space Weather, 2022.
[16] T. He et al., "SimBoost: a read-across approach for predicting drug–target binding affinities using gradient boosting machines," Journal of Cheminformatics, 2017.
[17] M. Li et al., "Prediction of the Aqueous Solubility of Compounds Based on Light Gradient Boosting Machines with Molecular Fingerprints and the Cuckoo Search Algorithm," ACS Omega, 2022.
[18] O. Sprangers et al., "Probabilistic Gradient Boosting Machines for Large-Scale Probabilistic Regression," Knowledge Discovery and Data Mining, 2021.
[19] P. Reddy and S. M. Kumar, "A Novel Approach to Improve Accuracy in Stock Price Prediction using Gradient Boosting Machines Algorithm compared with Naive Bayes Algorithm," 2022 4th International Conference on Advances in Computing, Communication Control and Networking (ICAC3N), 2022.
[20] A. Konstantinov and L. Utkin, "Interpretable Machine Learning with an Ensemble of Gradient Boosting Machines," Knowledge-Based Systems, 2020.
[21] Z. Luo, H. Yan, and X. Pan, ‘Optimizing Transformer Models for Resource-Constrained Environments: A Study on Model Compression Techniques’, Journal of Computational Methods in Engineering Applications, pp. 1–12, Nov. 2023, doi: 10.62836/jcmea.v3i1.030107.
[22] H. Yan and D. Shao, ‘Enhancing Transformer Training Efficiency with Dynamic Dropout’, Nov. 05, 2023, arXiv: arXiv:2411.03236. doi: 10.48550/arXiv.2411.03236.
[23] Y. Gan and D. Zhu, ‘The Research on Intelligent News Advertisement Recommendation Algorithm Based on Prompt Learning in End-to-End Large Language Model Architecture’, Innovations in Applied Engineering and Technology, pp. 1–19, 2023.
[24] D. Zhu, Y. Gan, and X. Chen, ‘Domain Adaptation-Based Machine Learning Framework for Customer Churn Prediction Across Varing Distributions’, Journal of Computational Methods in Engineering Applications, pp. 1–14, 2021.
[25] X. Deng, L. Li, M. Enomoto, and Y. Kawano, ‘Continuously frequency-tuneable plasmonic structures for terahertz bio-sensing and spectroscopy’, Scientific reports, vol. 9, no. 1, p. 3498, 2019.
[26] X. Deng, M. Simanullang, and Y. Kawano, ‘Ge-core/a-si-shell nanowire-based field-effect

transistor for sensitive terahertz detection', in Photonics, MDPI, 2018, p. 13.
[27] H. Zhang, D. Zhu, Y. Gan, and S. Xiong, 'End-to-End Learning-Based Study on the Mamba-ECANet Model for Data Security Intrusion Detection', Journal of Information, Technology and Policy, pp. 1–17, 2023.
[28] D. Zhu, X. Chen, and Y. Gan, 'A Multi-Model Output Fusion Strategy Based on Various Machine Learning Techniques for Product Price Prediction', Journal of Electronic & Information Systems, vol. 4, no. 1.
[29] X. Deng and Y. Kawano, 'Surface plasmon polariton graphene midinfrared photodetector with multifrequency resonance', Journal of Nanophotonics, vol. 12, no. 2, pp. 026017–026017, 2018.