Journal of Advanced Research in Applied Sciences and Engineering Technology

Journal homepage:
https://semarakilmu.com.my/journals/index.php/applied_sciences_eng_tech/index
ISSN: 2462-1943

A Machine Learning Approach for Efficient Mach-Zehnder Modulator Design

Hany Mahrous[1,*], Mostafa Fedawy[1,2], Mira Abboud[3], W. Fikry[4], Michael Gad[4]

[1] Electronics and Communications Department, Faculty of Engineering, Arab Academy for Science and Technology and Maritime Transport, Cairo 2033, Egypt
[2] Center of Excellence in Nanotechnology, Arab Academy for Science and Technology and Maritime Transport, Cairo 2033, Egypt
[3] Department of Computer Sciences, Faculty of Sciences, Lebanese University, Fanar 2611, Lebanon
[4] Engineering Physics and Mathematics Department, Faculty of Engineering, Ain Shams University, Cairo 11517, Egypt

ARTICLE INFO

Keywords:
Integrated optics; silicon photonics; silicon on insulator; machine learning; mach-zehnder modulator

ABSTRACT

The advancement of silicon photonics modulators is vital for achieving high-speed optical communications. However, designing optical modulators is a complex and resource-intensive task due to the large number of design parameters. Traditionally, the design of silicon photonics modulator depends on electrical and optical simulations. This work introduces employing machine learning models as a powerful design approach for MOS-like Mach-Zehnder Modulators to overcome the traditional design complexity. The proposed design and methodology build upon prior successful efforts in developing electro-optic modulators at the device level. The RandomForestRegressor model is developed to predict the effective refractive index (n_e) and a HistGradientBoostingRegressor for the absorption coefficient (k). These models show high prediction accuracy, with a mean absolute percentage error (MAPE) of 0.02% for the effective refractive index (n_e) and 0.80% for the absorption coefficient (k) for the test dataset. The developed models can predict the performance of MOS-like Mach-Zehnder Modulators within a few milliseconds, exhibiting a minimal margin of error. These results highlight the potential of integrating machine learning in photonic device design to simplify optimization, reduce the high computational and enable efficient exploration of new design spaces.

1. Introduction

The recent developments in the Internet of Things (IoT) and the fifth generation (5G) of telecom mobile networks require ultra-high-speed modulation Given the advanced technology available, it is necessary for several devices used in everyday life to be connected to the network, both wireless and wired. The backbone circuits of such networks necessitate a substantial amount of bandwidth. Ref. [1] states that the required speeds for different types of transport in 2020 are as follows: 1 $Pbps$ for core transport, one 1 $Tbps$ for metro transport, 100 Gbps for backhaul truck, and $1 - 10$ $Gbps$ for end-user devices.

* *Corresponding author.*
E-mail address: hanyhabib@student.aast.edu

https://doi.org/10.37934/araset.64.4.173186

Integrated photonics has attracted significant attention in this particular environment in recent years [2] Specifically, silicon photonics (SiP) has become a promising platform for creating affordable, high speed and low power optical modulators. This is because it can be easily integrated with complementary metal-oxide-semiconductor (CMOS) technology [3], allowing the integration with the electronic components. Additionally, silicon photonics benefits from existing fabrication expertise and a well-established manufacturing infrastructure. Unlike polymers, used in traditional photonics, such as $LiNbO_3$, Si possesses a centrosymmetric crystalline structure, resulting in limited parametric electro-optic effects such as the Pockels and Kerr effects [4]. However, the semiconducting properties of silicon (Si) enable the introduction and removal of unbound charge carriers, and this phenomenon is called the plasma dispersion effect (PDE). PDE enables the control of the distribution and the concentration of the free charge carriers with the applied voltage enabling the implementation of Silicon phase shifters. Silicon phase shifters can be utilized to manipulate the interferometric patterns in several interferometer setups, including micro-ring resonators (MRR), Michelson modulators, and Mach–Zehnder modulators (MZM) [5-7].

When comparing these various types of modulators, it can be observed that MRRs and Michelson interferometer modulators (MIM) possess characteristics such as small footprint, low power consumption, and good modulation efficiency. However, these modulators are constrained by limited bandwidth [6-9]. In contrast, Mach-Zehnder modulators (MZMs) offer a favorable balance between bandwidth, power consumption, and insertion loss, rendering them appropriate for high-speed systems. Mach-Zehnder modulators (MZMs) also demonstrate superior thermal resistance and minimize signal distortion known as chirp. High-performance Mach-Zehnder Modulators (MZMs) employ two-phase shifters to minimize the V_π values. These MZMs incorporate several innovations, such as resistance and capacitance (RC) equalizers, traveling-wave electrodes (TWEs) with 'T'-shaped extensions, and serial push-pull (SPP) driving topologies, to improve both modulation efficiency and bandwidth [10].

Recent developments in silicon photonic phase shifters have brought attention to three main mechanisms: carrier injection, carrier depletion, and carrier accumulation. Carrier injection, accomplished by applying a forward bias to PN junctions, can generate substantial phase shifts of up to π radians. However, this method is hindered by its high-power consumption, which can reach several milliwatts, and its slower operating speeds, which are below $16\,Gbps$ [11-14]. These limitations arise from the process of electron-hole recombination, as described by Soref and Bennett in 1987 [12]. Carrier depletion utilizes reverse-biased PN junctions to provide high-speed modulation ($> 40\ GHz$) while consuming reduced power (in the sub-milliwatt range). However, it only offers small phase shifts ($0.3 - 0.5\,\pi\ radians$) [5]. Carrier accumulation, utilizing MOS capacitor topologies, enables the attainment of high efficiency with speeds exceeding $100\ GHz$ and low signal degradation [15]. When comparing these options, it is important to note that carrier injection is restricted by both power and speed. On the other hand, carrier depletion strikes a compromise between speed and power, while maintaining reasonable efficiency. Lastly, MOS capacitor topologies have shown significant potential as carrier accumulation modulators, achieving high-speed operation with decreased power consumption. Consequently, it is necessary to conduct further research on the design and optimization of MOS-based modulators in order to optimize and improve their performance. This will make them more suitable for large-scale, high-speed applications [16-20].

The design process of an electro-optical modulator involves two main stages: simulating the electronic side to determine the change in charge concentration or current due to the modulating signal and simulating the optical side to observe how these electronic variations affect optical parameters, such as phase and losses, impacting the optical carrier signal. This process demands significant memory, processing power, and time, often requiring multiple programs. This complex

simulation process, alongside many design parameters (including dimensions, doping concentrations, and bias voltage) makes the optimization process of the electro-optical modulators challenging and computationally expensive; thus, finding alternatives to save computing resources is desirable [21].

Over the last decade, numerous researchers have carried out studies with different approaches for designing silicon photonics components in order to address the challenges posed by conventional design techniques [22]. Heuristic optimization algorithms, such as the genetic algorithm (GA), are extensively employed for optimizing the parameters of components [23-24]. Artificial intelligence (AI) is a potential option to overcome the complexity of the design and optimization of silicon photonic devices. AI models can be trained to achieve high accuracy in predicting the performance of an optical modulator [21]. The work of Gandhi Alagappan and Ching Eng Png [25] is among the first articles to utilize deep learning models for classifying optical waveguide geometries as either single-mode or multi-modal. Ma *et al.,* conducted a detailed evaluation of the achievements in deep learning-based photonic design in [21]. The authors explore the potential of deep learning as a very effective methodology for designing photonic structures. The researchers highlight that deep learning offers data-driven methodologies that complement traditional physics and rule-based techniques. The paper discusses the drawbacks of conventional numerical simulations, emphasizes different model structures for certain photonic tasks, and provides insights into the problems and prospects of this new research field.

In 2023, RA de Paula Jr *et al.,* [10] demonstrated the integration of artificial neural networks with the differential evolution optimization technique to enhance the performance of silicon-based Mach-Zehnder modulators. The authors implemented an Artificial Neural Network (ANN) model and used a numerically produced dataset from their previous research to train the ANN model. The model, which had undergone training, successfully predicts the performance of the examined MZM. Subsequently, the researchers integrated the model within the optimization procedure to optimize the modulator bandwidth, optical losses and operating voltage. In this work, machine learning models are developed to accurately predict the performance of MOS-Like Mach-Zehnder modulator. The modulator structure is based on our former research [17]. We picked the data for this modulator to train and evaluate the machine learning models. The implemented models are tested on data from the same research, where the testing data is unseen for the models during the training phase. This model can be used to study new design space.

The rest of the paper is organized as follows:

i. Section 2 reviews the selected modulator structure and its principle of work. In addition, data preparation and preprocessing are reviewed. Moreover, the selection and training of the machine learning model are explained.
ii. Section 3 discusses the results of the implemented models.
iii. Finally, Section 4 provides the conclusion, summarizing the findings.

2. Methodology

The main objective of this study is to improve the efficiency of MOS-Like Mach-Zehnder modulators (MZMs) by developing a machine-learning model(s). Initially, a dataset is gathered from our formed research [17]. This dataset contains the results of numerical simulation for such modulator. The dataset has been cleansed and preprocessed to ensure its suitability for the models. Next, a machine learning model has been trained and evaluated. The data preprocessing and the machine learning model are implemented using Python programming language [26] in the Google

Colab platform [27]. Following these steps, we will have a fully functional model ready for the MZM optimization.

2.1 MOS-Like Mach-Zehnder Modulators (MZMs)

In a previous work [17], a MOS-like Mach-Zehnder modulator that has low optical loss with the highest reported this-time operation speed to our knowledge has been presented. The modulator employs the plasma dispersion effect in changing the free charge carrier concentration in a silicon waveguide core, resulting in a phase shift in the optical signal. The modulator cross-section consists of a silicon waveguide core with a rectangular cross-section of height H and width W as shown in Figure 1. Two metallic electrodes, with dimensions equal to the core waveguide, are placed on the left and the right sides of the waveguide. The structure is encapsulated in silica (SiO_2). Four different rectangular waveguide sizes, as shown in Table 1, were studied, with dimensions chosen to realize single-mode conditions and ensure maximum mode confinement factor.

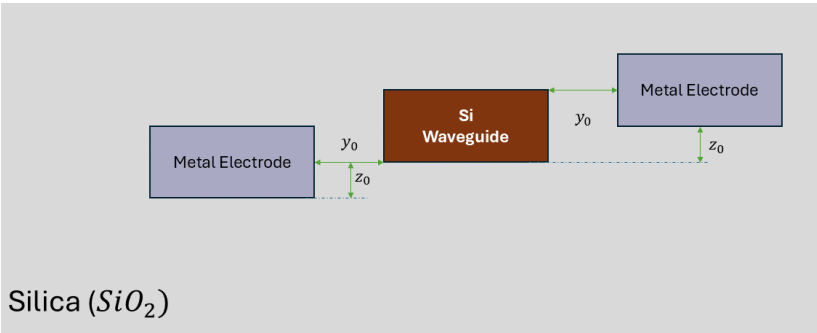


Fig. 1. A schematic cross section for the MOS-Like MZM

Table 1
Different waveguide configuration

Configuration	Height (nm)	Width (nm)
A	100	750
B	150	600
C	220	450
D	300	400

The electrodes are placed at distances y_0 and z_0 away from the core in the vertical and horizontal directions, respectively. $z_0 = 0$, means that the electrodes are aligned horizontally with the core. As z_0 increases, the right electrode moves upwards, and the left electrode moves downwards. Similarly, as y_0 increases, the right electrode moves to the right and the left electrode moves to the left. The electrodes were moved away the core by changing y_0 from 100 to $500\ nm$ and z_0 from 0 to $500\ nm$ in steps of $100\ nm$, creating 30 different combinations of electrode positions.

The right electrode is connected to the control signal and named Anode. The left electrode is grounded and is named Cathode. The same configuration applies to the other arm of the MZI, with the anode of one arm having a voltage equal to the negative of the other. In the 'ON' state, the anode voltage is zero, causing the optical signals to be in phase and the output to be high. In the 'OFF' state, different anode voltages induce a phase difference, minimizing the output signal. The design parameters of the modulator are shown in Table 2.

Table 2
MOS-Like Mach-Zehnder modulators design parameters

Parameter	Start value	End value	Step
Applied voltage	$-20\ V$	$20\ V$	$1\ V$
Concentration	$1x10^{16}\ cm^{-3}$	$1x10^{17}\ cm^{-3}$	$1x10^{16}\ cm^{-3}$
Electrode horizontal separation distance $\boldsymbol{y_0}$	$100\ nm$	$500\ nm$	$100\ nm$
Electrode vertical separation distance $\boldsymbol{z_0}$	$0\ nm$	$500\ nm$	$100\ nm$

Before delving into the detailed process of the design, it is important to review the performance metric of that symmetric MZM. One of the key metrics is the switching voltage V_π, which is the voltage required to switch the output optical power between its maximum and minimum values. The output power switching is dependent on the phase difference between the two arms: when the waves in the two arms are in phase, the output is maximum; when the waves are completely out of phase, the output is minimum. The switching voltage induces this phase difference by altering the charge distribution and consequently changing the effective refractive index (n_e) of the propagating modes inside the waveguides forming the MZM arms. The phase shifter length L_π, is a design metrics that is calculated by dividing the wavelength of the propagated light by $(2x\Delta n_e)$, where Δn_e is the change in effective refractive index (n_e), and is the length of the phase shifter that is required to give phase shift between the two phases by π. The switching voltage V_π and the phase shifter length L_π metrics are driven from the effective refractive index (n_e), which is the real part of the complex refractive index [16].

The efficiency of the modulator in terms of optical losses is assessed using insertion loss (IL) and extinction ratio (ER). Insertion loss is calculated from Eq. (1), where I_o is the input optical intensity and I_H is the maximum output optical intensity. The extinction ratio is calculated from Eq. (2), where I_L is the minimum output optical intensity [16]. These metrics are essential for evaluating the efficiency and performance of the MZM-based modulator. The insertion loss (IL) and extinction ratio (ER) of an optical modulator can be influenced by the absorption coefficient (k), which represents the material's absorption losses. Therefore, in this work, we will use the effective refractive index (n_e) and the absorption coefficient (k) as optimization parameters for the optical modulator.

$$IL = 10\ log\ (I_0/I_H) \tag{1}$$

$$ER = 10\ log\ (I_\mathrm{H}/I_L) \tag{2}$$

2.2 Data Preparation and Preprocessing

The dataset used in this study results from a numerical simulation of the reported MOS-like modulator in [17] Table 3 displays a representative sample (first 10 rows) extracted from the dataset used for training the models. The dataset is determined through the traditional design methodology which consists of two basic stages. Initially, the modulator wase simulated by a charge solver to calculate the charge density change inside the waveguide due to changing the applied voltage. Next, the waveguide section was analyzed by solving Maxwell's equations using a Finite-Difference Eigenmode (FDE) numerical simulation [28]. This allowed for the calculation of the effective refractive index (n_e) and the absorption coefficient (k) for the modulator.

Table 3
Sample data from the training dataset for MOS-Like Mach-Zehnder Modulators

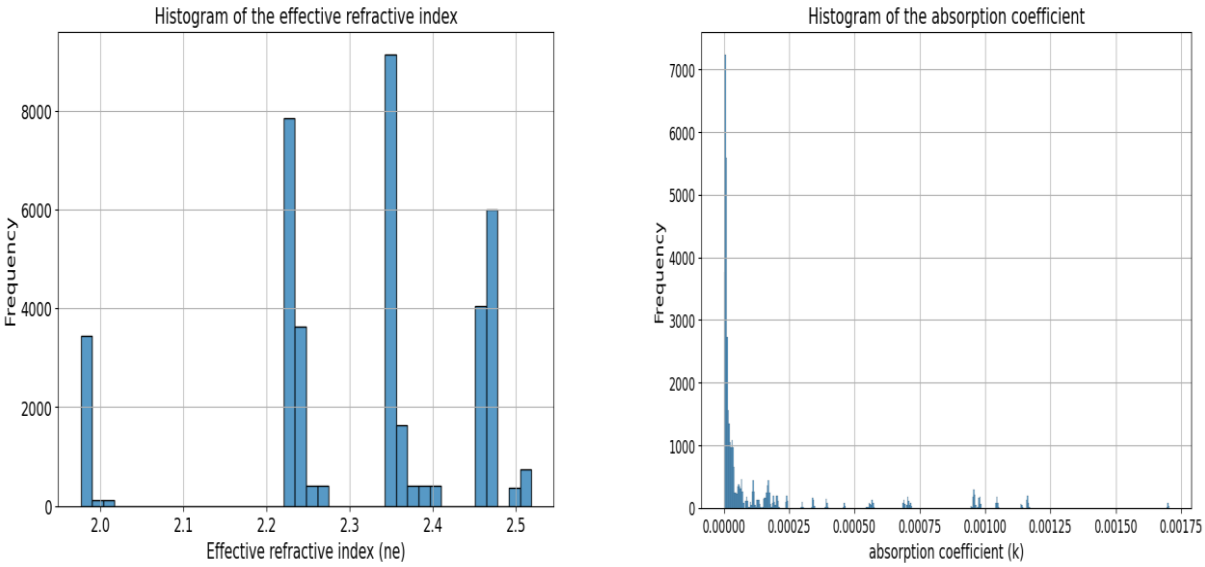
Waveguide height (m)	Doping concentration (cm^{-3})	z_0 (m)	y_0 (m)	Applied voltage (V)	Effective refractive index (n_e)	Absorption coefficient (k)
0.0000001	2E+22	0	0.0000001	-20	2.015784458	0.001701962943
0.0000001	2E+22	0	0.0000001	-19	2.015789768	0.001701740686
0.0000001	2E+22	0	0.0000001	-18	2.015795151	0.001701517474
0.0000001	2E+22	0	0.0000001	-17	2.015800611	0.001701293238
0.0000001	2E+22	0	0.0000001	-16	2.015806152	0.001701067904
0.0000001	2E+22	0	0.0000001	-15	2.015811782	0.001700841387
0.0000001	2E+22	0	0.0000001	-14	2.015817507	0.001700613594
0.0000001	2E+22	0	0.0000001	-13	2.015823335	0.001700384417
0.0000001	2E+22	0	0.0000001	-12	2.015829274	0.001700153734
0.0000001	2E+22	0	0.0000001	-11	2.015835334	0.001699921405

The data set includes the waveguide's height, describing the waveguide cross-section area, doping concentration, separation distance between the electrodes and the waveguide in both horizontal and vertical directions, applied voltage, the effective refractive index (n_e) and the absorption coefficient (k). Each row in the dataset represents a single run at a specific voltage step, with every 41 rows corresponding to one modulator configuration. The dataset contains 49200 rows, which equates to 1200 unique modulator configurations.

The descriptive statistics of the dataset, that is used for the training and evaluation of the model, provide essential information about the range and distribution of the parameters, as shown in Table 4. During the data preprocessing step, the input features are normalized to provide consistency across the different scales of the parameters. Normalization is required in order to prevent any individual feature from dominating the learning process due to its magnitude. The technique of Min-Max scaling was utilized to normalize all feature values within the range of [0, 1], enabling rapid and precise training and evaluation of the machine learning model. The histogram of the effective refractive index (n_e), Figure 2(a), shows that its values is uniformly distributed in groups along the range. This distribution shows that this output objective can be processed without more preprocessing. While the histogram of the absorption coefficient (k), Figure 2(b), shows a highly skewed distribution towards the lower end of the scale with more than 7000 occurrences at $k \approx 0$. To manage this discrepancy, the log transformation technique is used to improve the efficiency of the machine learning models for this output.

Table 4
Statistical data for the MOS-Like Mach-Zehnder Modulators training dataset (post-cleaning)

	Waveguide height (m)	Doping concentration (cm^{-3})	z_0 (m)	y_0 (m)	Applied voltage (V)	Effective refractive index (n_e)	Absorption coefficient (k)
count	39155	39155	39155	39155	39155	39155	39155
mean	2.09E-07	5.38E+22	2.49E-07	2.99E-07	0	2.3163	1.21E-04
std	6.79E-08	2.78E+22	1.70E-07	1.42E-07	11.832311	0.141446	2.66E-04
min	1.00E-07	1.00E+22	0.00E+00	1.00E-07	-20	1.976813	1.69E-07
25%	1.50E-07	3.00E+22	1.00E-07	2.00E-07	-10	2.233821	6.92E-06
50%	2.20E-07	5.00E+22	2.00E-07	3.00E-07	0	2.353751	1.54E-05
75%	3.00E-07	8.00E+22	4.00E-07	4.00E-07	10	2.464247	6.93E-05
max	3.00E-07	1.00E+23	5.00E-07	5.00E-07	20	2.518619	1.70E+00



(a) (b)

Fig. 2. Histogram for (a) Effective refractive index (n_e) (b) Absorption coefficient (k)

2.3 Machine Learning Model Development

2.3.1 Machine learning model selection

The selection of the appropriate machine learning model in this study depends on the characteristics of the selected dataset. As discussed in the previous section, the dataset consists of numerical features and variables related to the performance of the modulator. The correlation matrix, Figure 3, highlights the connections between these features, indicating potential non-linear interactions that need to be effectively captured by the models. The correlation matrix indicates a strong correlation between the waveguide height and the effective refractive index (n_e). This correlation can be explained physically by the fact that the height of the waveguide determines the cross-sectional area, which in turn affects the number of free charge carriers inside the waveguide. This, in turn, impacts the refractive index (n_e) of the waveguide. The other noticeable relation is a moderately negative correlation between the electrode's separation, and the absorption coefficient (k). The physical meaning of this is that the long distance that separates the electrodes from the waveguide leads to low absorption loss for the propagated light. Other features show weak correlations with the output objectives. One more critical relation between the two output objectives is that there is a very weak correlation with the output objectives. This relation enables us to use two separate models, one for each output.

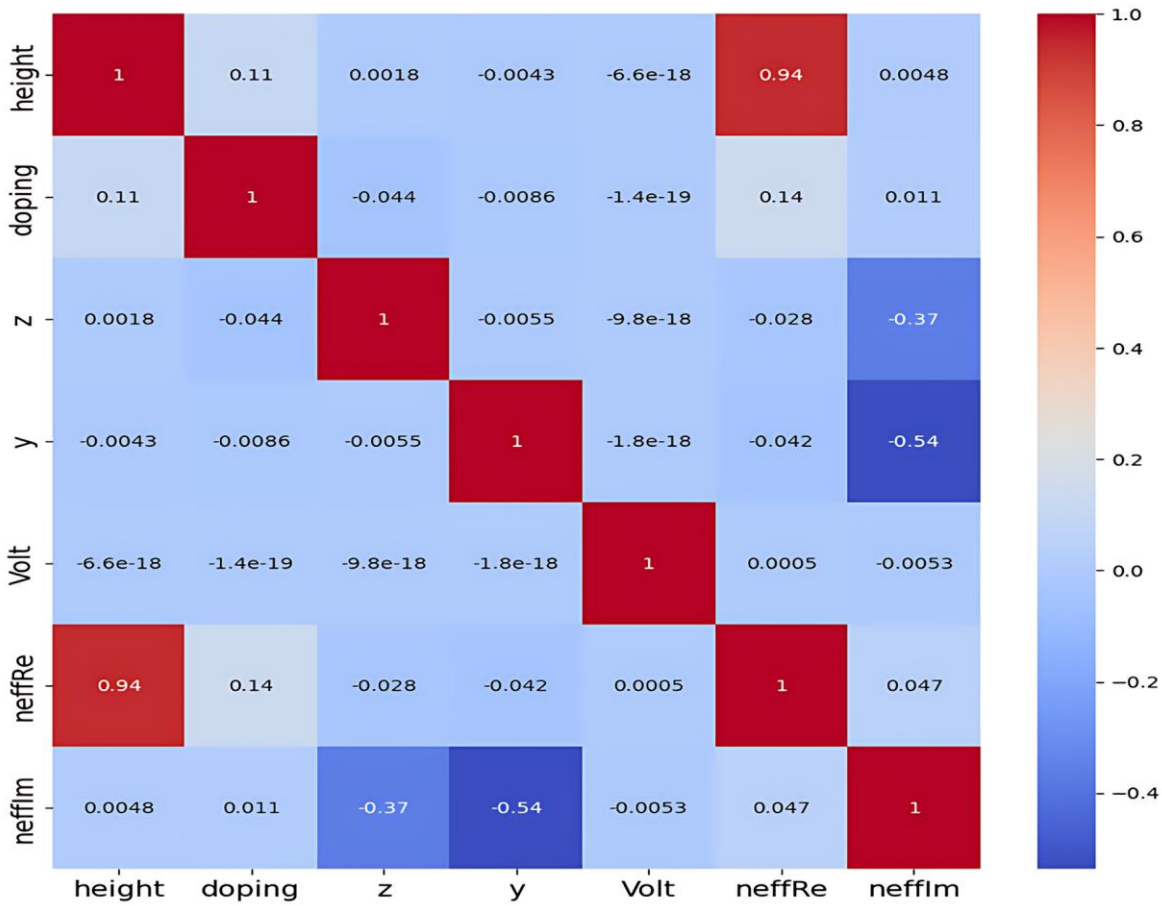


Fig. 3. The correlation matrix for the training dataset

Considering the attributes and correlations of the dataset, the RandomForestRegressor model [29] is used for predicting the effective refractive index (n_e) because of its ability to manage linear and nonlinear data and capability to manage a large and complex relationships. This ability makes it well-suited to the large correlation between the height of the waveguide and the effective refractive index (n_e). The ensemble technique of such a model generates multiple decision trees throughout the training process and combines their predictions, resulting in improved accuracy and reduces the risk of overfitting. The feature importance ranking capability of this model provides vital insights into the features that have the greatest impact on the effective refractive index (n_e).

The highly skewed histogram of the absorption coefficient (k), Figure 2(b), leads to increase the error when using the RandomForestRegressor model. The HistGradientBoostingRegressor model [29] belongs to the gradient boosting family, which sequentially builds an ensemble of trees to minimize the remaining errors from previous steps. This advantage enables the use of this model with absorption coefficient (k).

2.3.2 Machine learning model training and evaluation

The model training procedure involved splitting the dataset into training and testing sets. The dataset is divided into three different sets. First, a 10% of the primary dataset is pickad to test the developed model after training it. This dataset has been manually selected from the original dataset and deliberately excluded from both the training and validation steps of the machine learning models. As a result, the models have not been exposed to this dataset during either the training or validation processes. This testing dataset is employed following the training and validation of the

models. The remaining 90% of the primary dataset is split as 80% for training and 20% for validation during the training phase. Each dataset is split into input features and output objectives. The input features are the waveguide height, the doping concentration, y_0, z_0 and the applied voltage, and the output objectives are the effective refractive index (n_e) and the absorption coefficient (k).

The RandomForestRegressor model is utilized to predict the effective refractive index (n_e), while the HistGradientBoostingRegressor model was implemented to predict the absorption coefficient (k). Both models were incorporated into preprocessing pipelines to standardize the characteristics using Min-Max scaling. RandomForestRegressor was built with 100 trees, while HistGradientBoostingRegressor is initialized with a learning rate of 0.01 and $10{,}000$ iterations, as shown in Table 5. The model parameters are selected after many trials, considering the trade-off between accuracy and computational time. The performance of models is evaluated by predicting the output objectives and subsequently computing the Mean Square Error (MSE) and Mean Absolute Percentage Error (MAPE) to verify their accuracy. Table 6 comprehensively explains how the models are trained and evaluated.

Table 5
Parameters and values for RandomForestRegressor and HistGradientBoosting Regressor models

Model	Parameter	Value
RandomForestRegressor	n_estimators (number of trees)	100
	min_samples_split	2
	min_samples_leaf	1
	loss	'Squared error'
HistGradientBoostingRegressor	learning_rate	0.01
	max_iter	10000
	min_samples_split	2
	min_samples_leaf	1
	loss	'Squared error'

Table 6
Steps to train the machine learning models

Algorithm
Input:
X, y1, y2 #training features and target output objectives
1. y2_train←log10(y2_train) #Log transformation
2. Split data into training and validation sets
3. Create preprocessing pipeline
4. Initialize and fit models
5. Evaluate models (X_test, y1_test, y2_test)
6. Calculate and print metrics

The evaluation of the performance of the machine learning models in this study is conducted by employing Mean Squared Error (MSE) and Mean Absolute Percentage Error (MAPE). These metrics are chosen for their widespread use in regression analysis and their ability to provide a comprehensive view of the model's prediction accuracy and reliability. During the training phase the Mean Squared Error (MSE) was set as an error validation parameter for both models as shown in Table 5. This parameter is used to measure the average of the squared difference between the actual and predicted values. MSE is chosen due to its ability to apply higher penalty on large variances,

making it support variance reduction as a feature selection criterion and helps to reduce the least squares error [30].

Mean Absolute Percentage Error (MAPE) is used to evaluate the trained model with the validation dataset. MAPE expresses the error as percentage and is defined by Eq. (3). Where y_i is the actual value, $\hat{y}_i$ is the predicted value, and n is the number of observations. MAPE helped us evaluate our models' performance parameter prediction by explaining the prediction error in relative terms, regardless of data scale. The trained models are tested by using the testing dataset, which is considered a new dataset for the model since it is picked before the training phase.

$$MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \; x \; 100 \tag{3}$$

3. Results

Assessing the performance of machine learning models with new, unseen datasets is crucial for evaluating its generalization. While the models realize exceptionally low MAPE with the training and validation dataset. It still provides a very low MAPE with the testing dataset. Table 7 shows the MAPE for the models in the evaluation and the testing stages.

Table 7
Mean absolute percentage error for the evaluation and testing datasets

Model	MAPE for evaluation phase	MAPE for testing phase
RandomForestRegressor	0.00000016 %	0.02 %
HistGradientBoostingRegressor	0.13 %	0.80 %

The testing dataset, after removing the outliers, contains twenty different modulators configuration with 820 sample rows. The developed machine learning models, after the training and evaluation, are used to predict the effective refractive index (n_e) index and the absorption coefficient (k) for a testing dataset. The absolute error between the predicted and the numerically simulated values, as reported in [17], is calculated for both outputs. The results are displayed in Figure 4 and Figure 5 where the x-axis represents the index of the observations, while the y-axis represents the absolute error (e). For the effective refractive index (n_e) output, Figure 4 highlights that most of the data samples show an exceptionally low absolute error, close to 0 %, indicating high accuracy in the predictions for most of the observations, while all the data range experience absolute error below 0.01 %. For the absorption coefficient (k), Figure 5 displays that for a massive portion of the data points, the absolute error is again close to 0 %, indicating that the model performs well in predicting the absorption coefficient (k) with the whole range of the absolute error which is less than 0.00025 %. These error ranges are considered a great for most of the physics applications for the machine learning, and it is acceptable in silicon photonics field which mean that the developed models can be ralied upon and employed in an optimization algorithm to search the unvisited design space for the MOS-like MZM.

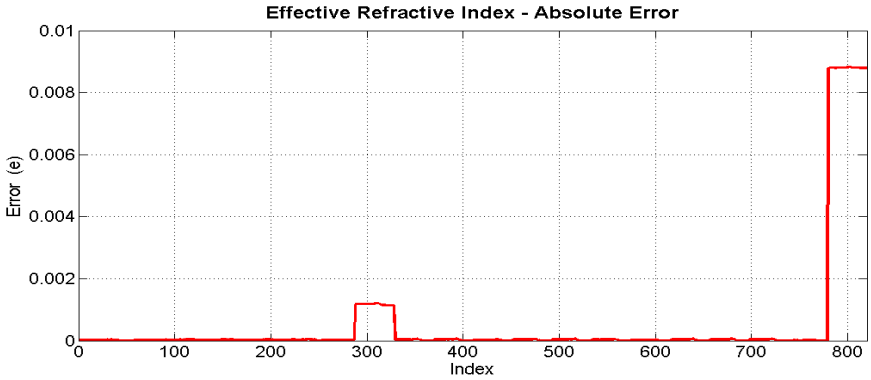


Fig. 4. Absolute error for the effective refractive index (n_e) for the testing dataset

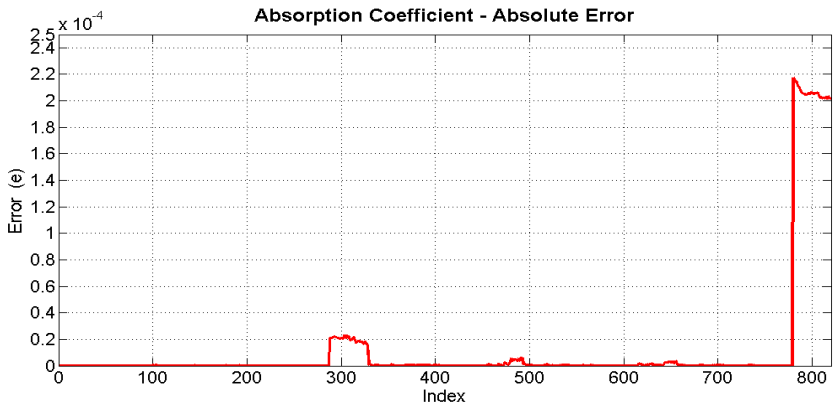
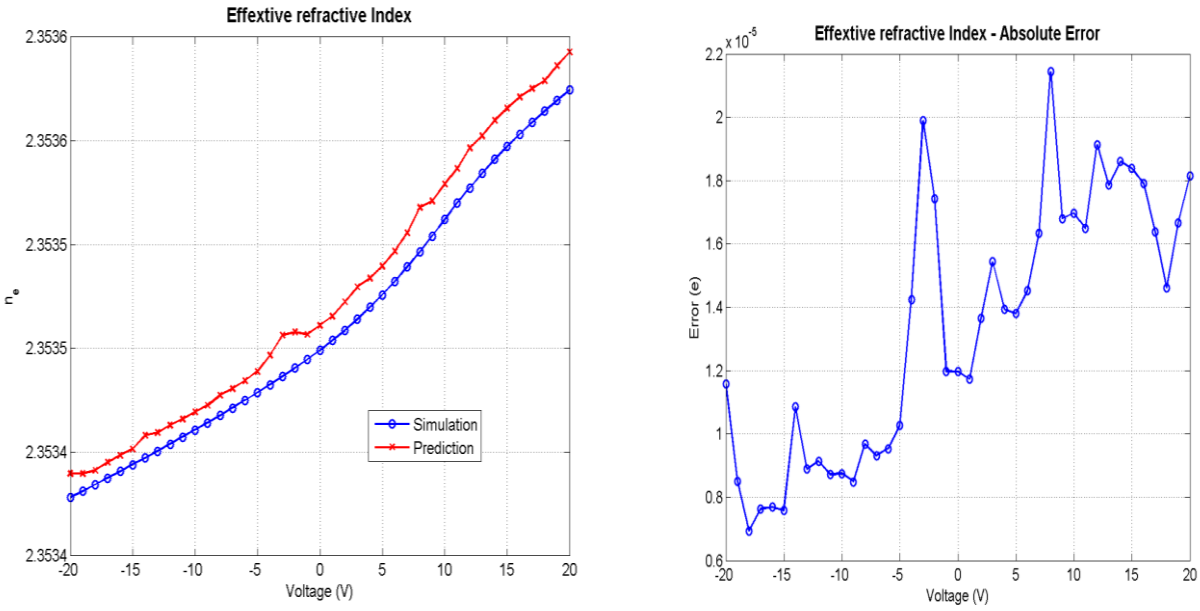


Fig. 5. Absolute error for the absorption coefficient (k) for the testing dataset

All the good results are for the whole dataset, but it needs to be examined from the silicon photonics point of view. So, we chose randomly one of the modulator configurations from the testing dataset and used the developed models to predict its performance. For this sample modulator configuration with parameters as listed in Table 8, a comparison between the simulated and predicted outputs is illustrated in Figure 6. The effective refractive index (n_e) as a function of applied voltage (V) is illustrated in Figure 6(a). The x-axis represents the voltage in volts (V), while the y-axis represents the effective refractive index (n_e). Both the predicted and simulated curves follow the same increasing behavior as the voltage increases. This indicates that the RandomForestRegressor model can capture the overall relationship between applied voltage and the effective refractive index (n_e). The deviation between predicted and the simulated curves is considering the model error and is shown in Figure 6(b) as an absolute error.

Table 8
Sample MOS-Like Mach-Zehnder modulator parameters

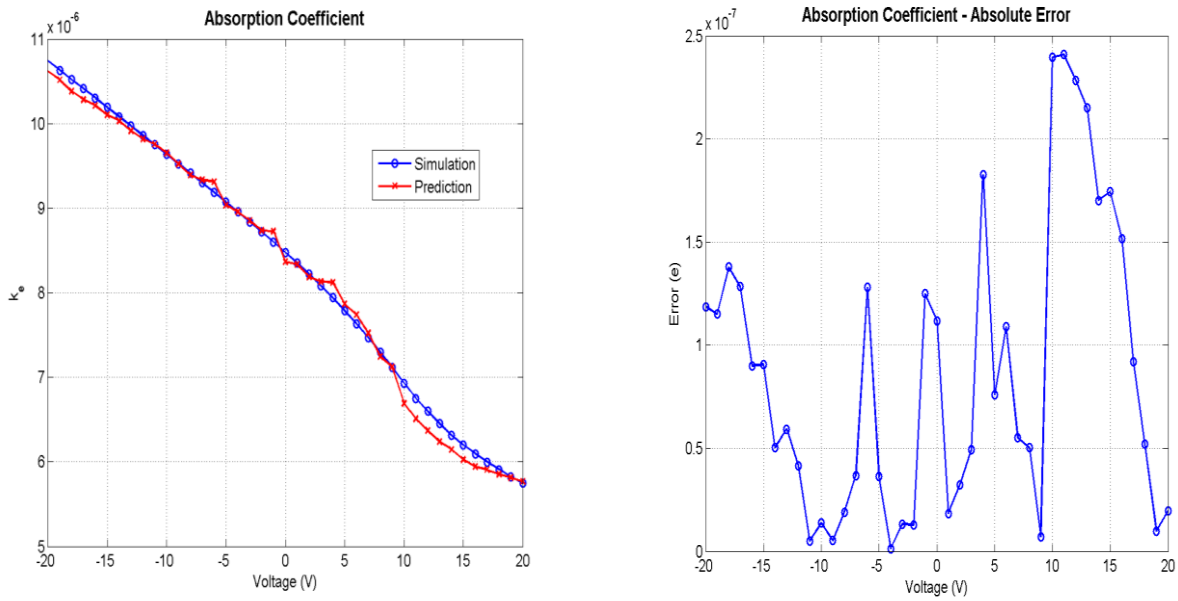
Parameter	Value
Waveguide height	220 nm
Waveguide width	450 nm
Doping concentration	$1x10^{23}$ cm^{-3}
Electrode horizontal separation distance $\boldsymbol{y_0}$	400 nm
Electrode vertical separation distance $\boldsymbol{z_0}$	500 nm



(a) (b)

Fig. 6. Effective refractive index (n_e) modelling results (a) Simulated vs predicted (b) Absolute error

Figure 7(a) displays the predicted and simulated curves for absorption coefficient (k) for the same modulator. The x-axis represents the voltage in volts (V), while the y-axis represents the absorption coefficient (k_e). The two curves have the same decreasing trend with the voltage increase, indicating that the HistGradientBoostingRegressor model can capture the relationship between applied voltage and the absorption coefficient (k). The absolute error between the predicted curve and the simulated curve is shown in Figure 7(b). The error curves show that most of the data points for the sample modulator utilize absolute error below $2x10^{-5}$ for effective refractive index (n_e) and below $2.5x10^{-7}$ for the absorption coefficient (k).



(a) (b)

Fig. 7. Absorption coefficient (k) modelling results (a) Simulated vs predicted (b) Absolute error

In comparison, the previously used numerical simulation for the Mach-Zehnder Modulator (MZM) design realizes high accuracy but is time-consuming and resource-intensive. Simulating a single modulator takes approximately 20 minutes. In contrast, While the training process for machine learning models can take nearly one hour, these models can predict device performance in almost a few milliseconds after training. ML models have good accuracy; however, more research is needed to reduce prediction errors. The ML technique is efficient, allowing rapid examinations and exploration of large design spaces. It supports optimization algorithms like genetic algorithms for fast configuration discovery, making it a valuable tool in silicon photonics developments.

4. Conclusions

In this study the application of machine learning models is demonstrated for predicting the performance characteristics of MOS-like Mach-Zehnder Modulators (MZMs) in silicon photonics. We built on our former work to find out a dataset for training and evaluation these models. Using the RandomForestRegressor model, we accurately predicted the effective refractive index (n_e). Additionally, we used the HistGradientBoostingRegressor model for predicting the absorption coefficient (k). Our predictions had a mean absolute percentage error (MAPE) of 0.02% for the effective refractive index (n_e) and 0.80% for the absorption coefficient (k). These findings demonstrate the ability of machine learning to greatly decrease the computational complexity and time needed for the development and improvement of photonic devices.

Acknowledgement
This paper is based upon work supported financially by The Arab Academy for Science, Technology and Maritime Transport under grant number (2056).

References
[1] Alavi, S. E., M. R. K. Soltanian, I. S. Amiri, M. Khalily, A. S. M. Supa'At, and Harith Ahmad. "Towards 5G: A photonic based millimeter wave signal generation for applying in 5G access fronthaul." *Scientific reports* 6, no. 1 (2016): 19891. https://doi.org/10.1038/srep19891

[2] Sharma, Tarun, Jiaqi Wang, Brajesh Kumar Kaushik, Zhenzhou Cheng, Roshan Kumar, Zhao Wei, and Xuejin Li. "Review of recent progress on silicon nitride-based photonic integrated circuits." *Ieee Access* 8 (2020): 195436-195446. https://doi.org/10.1109/ACCESS.2020.3032186

[3] Siew, Shawn Yohanes, Bo Li, Feng Gao, Hai Yang Zheng, Wenle Zhang, Pengfei Guo, Shawn Wu Xie et al. "Review of silicon photonics technology and platform development." *Journal of Lightwave Technology* 39, no. 13 (2021): 4374-4389. https://doi.org/10.1109/JLT.2021.3066203

[4] Mulcahy, Jack, Frank H. Peters, and Xing Dai. "Modulators in silicon photonics—heterogenous integration & and beyond." In *Photonics*, vol. 9, no. 1, p. 40. MDPI, 2022. https://doi.org/10.3390/photonics9010040

[5] Kim, Younghyun, Jae-Hoon Han, Daehwan Ahn, and Sanghyeon Kim. "Heterogeneously-integrated optical phase shifters for next-generation modulators and switches on a silicon photonics platform: A review." *Micromachines* 12, no. 6 (2021): 625. https://doi.org/10.3390/mi12060625

[6] Gad, Michael, David Yevick, and Paul E. Jessop. "Tunable polymer/silicon over insulator ring resonators." *Optical Engineering* 47, no. 12 (2009): 124601-124601. https://doi.org/10.1117/1.3050355

[7] Gad, Michael, David Yevick, and Paul E. Jessop. "High-speed polymer/silicon on insulator ring resonator switch." *Optical Engineering* 47, no. 9 (2008): 094601-094601. https://doi.org/10.1117/1.2978947

[8] Rahim, Abdul, Artur Hermans, Benjamin Wohlfeil, Despoina Petousi, Bart Kuyken, Dries Van Thourhout, and Roel Baets. "Taking silicon photonics modulators to a higher performance level: state-of-the-art and a review of new technologies." *Advanced Photonics* 3, no. 2 (2021): 024003-024003. https://doi.org/10.1117/1.AP.3.2.024003

[9] Xu, Mengyue, Wenjun Chen, Mingbo He, Xueqin Wen, Ziliang Ruan, Jian Xu, Lifeng Chen, Liu Liu, Siyuan Yu, and Xinlun Cai. "Michelson interferometer modulator based on hybrid silicon and lithium niobate platform." *Apl Photonics* 4, no. 10 (2019). https://doi.org/10.1063/1.5115136

[10] Aparecido de Paula Jr, Romulo, Ivan Aldaya, Tiago Sutili, Rafael C. Figueiredo, Julian L. Pita, and Yesica RR Bustamante. "Design of a silicon Mach–Zehnder modulator via deep learning and evolutionary algorithms." *Scientific reports* 13, no. 1 (2023): 14662. https://doi.org/10.1038/s41598-023-41558-8
[11] Lira, Hugo LR, Sasikanth Manipatruni, and Michal Lipson. "Broadband hitless silicon electro-optic switch for on-chip optical networks." *Optics Express* 17, no. 25 (2009): 22271-22280. https://doi.org/10.1364/OE.17.022271
[12] Soref, R. I. C. H. A. R. D. A., and B. R. I. A. N. R. Bennett. "Electrooptical effects in silicon." *IEEE journal of quantum electronics* 23, no. 1 (1987): 123-129. https://doi.org/10.1109/JQE.1987.1073206
[13] Watts, Michael R., William A. Zortman, Douglas C. Trotter, Ralph W. Young, and Anthony L. Lentine. "Low-voltage, compact, depletion-mode, silicon Mach–Zehnder modulator." *IEEE Journal of Selected Topics in Quantum Electronics* 16, no. 1 (2010): 159-164. https://doi.org/10.1109/JSTQE.2009.2035059
[14] Barrios, C. Angulo, V. R. Almeida, R. Panepucci, and M. A. L. M. Lipson. "Electrooptic modulation of silicon-on-insulator submicrometer-size waveguide devices." *Journal of Lightwave Technology* 21, no. 10 (2003): 2332-2339. https://doi.org/10.1109/JLT.2003.818167
[15] Ramesh Gabriel, Jesuwanth Sugesh, and Sivasubramanian Arunagiri. "Performance analysis of carrier depletion silicon PIN phase shifter." *Journal of Optical Communications* 44, no. s1 (2024): s131-s137. https://doi.org/10.1515/joc-2019-0259
[16] Mahrous, Hany, Michael Gad, Mona El Sabbagh, Mostafa Fedawy, and W. Fikry. "A high-speed electro-optic modulator with optimized electrode positions." In *2018 13th International Conference on Computer Engineering and Systems (ICCES)*, pp. 530-535. IEEE, 2018. https://doi.org/10.1109/ICCES.2018.8639436
[17] Mahrous, Hany, Mostafa Fedawy, Mona El Sabbagh, W. Fikry, and Michael Gad. "130 Gbps low-loss electro-optic modulator based on metal-oxide-semiconductor technology." *Optik* 217 (2020): 164928. https://doi.org/10.1016/j.ijleo.2020.164928
[18] Mahrous, Hany, Mostafa Fedawy, Mona El Sabbagh, W. Fikry, and Michael Gad. "A compact 120 GHz monolithic silicon-on-silica electro-optic modulator." *Optical and Quantum Electronics* 52, no. 2 (2020): 111. https://doi.org/10.1007/s11082-020-2239-4
[19] Mahrous, Hany, Mostafa Fedawy, Mona El Sabbagh, W. Fikry, and Michael Gad. "Design of a 90 GHz SOI fin electro-optic modulator for high-speed applications." *Applied Sciences* 9, no. 22 (2019): 4917. https://doi.org/10.3390/app9224917
[20] Mahrous, Hany, Mina Azmy, Ahmed Afifi, Abdallh AbouElainain, Amr Kotb, Mostafa Fedawy, W. Fikry, Michael Gad, and Dalia Selim. "Design of compact, high-speed and low-loss silicon-on-silica electro-optic modulators." *Semiconductor Science and Technology* 35, no. 9 (2020): 095017. https://doi.org/10.1088/1361-6641/ab9d09
[21] Ma, Wei, Zhaocheng Liu, Zhaxylyk A. Kudyshev, Alexandra Boltasseva, Wenshan Cai, and Yongmin Liu. "Deep learning for the design of photonic structures." *Nature photonics* 15, no. 2 (2021): 77-90. https://doi.org/10.1038/s41566-020-0685-y
[22] Li, Zean, Zhipeng Zhou, Cheng Qiu, Yongyi Chen, Bohan Liang, Yubing Wang, Lei Liang et al. "The intelligent design of silicon photonic devices." *Advanced Optical Materials* 12, no. 7 (2024): 2301337. https://doi.org/10.1002/adom.202301337
[23] Mahrous, Hany, Mostafa Fedawy, Mira Abboud, Ahmed Shaker, W. Fikry, and Michael Gad. "A multi-objective genetic algorithm approach for silicon photonics design." In *Photonics*, vol. 11, no. 1, p. 80. MDPI, 2024. https://doi.org/10.3390/photonics11010080
[24] Mahrous, Hany, Mostafa Fedawy, Mira Abboud, Dalia Selim, and Michael Gad. "Optimizing the Design of a Silicon-Photonics Interleaver Circuit." In *2022 39th National Radio Science Conference (NRSC)*, vol. 1, pp. 293-300. IEEE, 2022. https://doi.org/10.1109/NRSC57219.2022.9971168
[25] Alagappan, Gandhi, and Ching Eng Png. "Modal classification in optical waveguides using deep learning." *Journal of Modern Optics* 66, no. 5 (2019): 557-561. https://doi.org/10.1080/09500340.2018.1552331
[26] Python.org. "Welcome to Python.org," n.d. Welcome to Python.org.
[27] "Google Colab," n.d. Welcome To Colab - Colab (google.com).
[28] Ansys. "Photonics Simulation Software | Ansys Lumerical," n.d. Optical Simulation and Design Software | Ansys Optics.
[29] "Scikit-learn: Machine Learning in Python — Scikit-learn 1.5.1 Documentation," n.d. scikit-learn: machine learning in Python — scikit-learn 1.5.1 documentation.
[30] Chai, Tianfeng, and Roland R. Draxler. "Root mean square error (RMSE) or mean absolute error (MAE)?–Arguments against avoiding RMSE in the literature." *Geoscientific model development* 7, no. 3 (2014): 1247-1250. https://doi.org/10.5194/gmd-7-1247-2014