



Article

PreOBP_ML: Machine Learning Algorithms for Prediction of Optical Biosensor Parameters

Kawsar Ahmed ^{1,*} , Francis M. Bui ¹ and Fang-Xiang Wu ²

¹ Department of Electrical and Computer Engineering, University of Saskatchewan, 57 Campus Drive, Saskatoon, SK S7N 5A9, Canada; francis.bui@usask.ca

² Division of Biomedical Engineering, Department of Computer Science and Department of Mechanical Engineering, University of Saskatchewan, Saskatoon, SK S7N 5A9, Canada; faw341@mail.usask.ca

* Correspondence: k.ahmed@usask.ca or kawsar.ict@mbstu.ac.bd

Abstract: To develop standard optical biosensors, the simulation procedure takes a lot of time. For reducing that enormous amount of time and effort, machine learning might be a better solution. Effective indices, core power, total power, and effective area are the most crucial parameters for evaluating optical sensors. In this study, several machine learning (ML) approaches have been applied to predict those parameters while considering the core radius, cladding radius, pitch, analyte, and wavelength as the input vectors. We have utilized least squares (LS), LASSO, Elastic-Net (ENet), and Bayesian ridge regression (BRR) to make a comparative discussion using a balanced dataset obtained with the COMSOL Multiphysics simulation tool. Furthermore, a more extensive analysis of sensitivity, power fraction, and confinement loss is also demonstrated using the predicted and simulated data. The suggested models were also examined in terms of R^2 -score, mean average error (MAE), and mean squared error (MSE), with all of the models having an R^2 -score of more than 0.99, and it was also shown that optical biosensors had a design error rate of less than 3%. This research might pave the way for machine learning-based optimization approaches to be used to improve optical biosensors.

Keywords: machine learning; optical sensor; parameter estimations; prediction; performance



Citation: Ahmed, K.; Bui, F.M.; Wu, F.-X. PreOBP_ML: Machine Learning Algorithms for Prediction of Optical Biosensor Parameters. *Micromachines* **2023**, *14*, 1174. <https://doi.org/10.3390/mi14061174>

Academic Editor: Sabina Merlo

Received: 15 May 2023

Revised: 28 May 2023

Accepted: 29 May 2023

Published: 31 May 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Knight [1] first proposed the notion of photonic crystal fiber (PCF), which consists of a solid or hollow core with a microstructure arrangement extending throughout the fiber's length. PCFs are becoming more and more widespread because of their design freedom, low cost, resilience, rapid detection, compact size, high sensitivity, and versatility. As a result of these characteristics, it is utilized in optical sensors [2–4], optical lasers [5,6], spectroscopy [7,8], Raman scattering [9], and a variety of other applications. The finite element method (FEM) [10,11], block-iterative frequency domain methods [12,13], and plane wave expansion methods [14,15], are all numerical approaches for improving the properties of PCFs. Such procedures, however, take a long time to complete. Furthermore, determining a PCF's properties typically needs a considerable quantity of data, and all other parameters must be recalculated when one parameter is altered. As a result, parameter updates as well as data collection take a while to execute. Machine learning (ML) is now the most effective solution for these challenges, and it has begun to address them.

Machine learning is a technique for developing a standard function for a hypothesis or for learning specific indicators to distinguish different inputs as separate samples. A well-balanced dataset may be used to effectively train machine learning algorithms and help them in making highly accurate predictions. Over the last several years, machine learning techniques have been employed in a range of sectors, including medical research, transportation, agricultural science, plasmonics, biosensing, traffic categorization, network security, and many more. In the realm of PCF, ML has been utilized to optimize and

forecast the performance of photonic crystal nanocavities [16–18]. To our knowledge, only a few researchers have attempted to predict optical characteristics using machine-learning approaches. An artificial neural network (ANN) model for calculating the optical characteristics of a PCF was reported by Chugh et al. [19]. They raised the number of epochs to produce a more accurate prediction as well as when the wavelength exceeds 1.5 μm , the confinement loss is dispersed. They relied only on the ANN model, with no comparisons to other machine learning approaches. Ferreira and Malheiros-Silveira proposed an ANN model based on multilayer perceptron and Extreme Learning Machine for calculating the optical characteristics of PCF in 2018 [20]. They were only concerned with PCF geometrical models. Khan et al. addressed several machine learning techniques for performance monitoring in optical communications and networking applications at the start of 2019 [21]. They opined that the TensorFlow package was useful for building machine learning algorithms. At the end of 2019 [22], Chugh et al. investigated an ML regression technique for estimating the effective index, coupling length, and power confinement in various waveguides. According to the researchers, the PyTorch and MLPRegressor models exhibit absolute percentage errors of 7–10% and 1–4%, respectively. In December 2019, the same team presented an ML technique for calculating the optical properties of a PCF [19]. They only looked at the effective refractive index (ERI) and confinement loss for up to 5000 epochs (CL). They did not display the extra design features or optical attributes that they exhibited in the previous research [19]. The usage of an ANN and a Convolutional Neural Network (CNN) in the mode classification of PCF-based surface plasmon resonance (SPR) sensor design was proposed by Khare and Goswami in 2021 [19]. They proposed a variety of methods for simply categorizing modes. However, a number of design aspects were overlooked.

When evaluating PCF design factors and forecasting optical characteristics, different ML algorithms may play an essential role in enhancing PCF quality in various aspects in a short period and with less exertion. As a consequence, machine learning methods in the field of optical sensor design and improvement show a lot of opportunities. As a result, we utilized four distinct machine learning methods that, to our information, have never been used previously in this field. Other machine-learning approaches may have been utilized in the research, but we opted for those since they are relatively new in this field. The following are the primary objectives of this study:

1. To investigate the attributes of PCF sensors using various ML algorithms.
2. To make some changes to improve the model's accuracy as well as to minimize the error rate.
3. To estimate the output faster than direct numerical simulation strategies.

2. Parameter Estimation Methods

2.1. Machine Learning and Optimization

To make predictions or judgments without needing to be explicitly programmed to do so, ML algorithms build a model based on training data [23]. In many fields, such as health, computer vision, speech recognition, and email filtering, where constructing conventional algorithms to perform the necessary tasks is challenging or even impractical, ML methods are used [24,25]. The minimization of a loss function on a training set of instances is a common formulation for learning problems, which connects machine learning to optimization. Loss functions describe the discrepancy between the predictions of the model and the actual occurrences of the problem. These models are trained to properly predict the labels that have been assigned to a group of samples [26]. The four approaches used are briefly described below.

2.2. Least Squares Regression (LSR) Method

The LSR approach has traditionally been used to minimize the sum of squares of error terms with a homogeneous variance and normal distribution and, therefore, to improve the model [27]. In terms of optimization, it solves problems of the form [28]:

$$\text{Min}_{\beta} ||Y - X\beta||_2^2 \quad (1)$$

where X stands for the independent input variable, Y stands for the model output, and the symbol β stands for the model parameter. The LSR approach relies on a number of assumptions in order to produce trustworthy results. The regression model should be represented in a linear form, according to the starting assumption. This implies that the average error predicted by the regression model should be zero. The residual variance must be constant, and the residuals must not be correlated with one another (no autocorrelation) [29]. Certain assumptions must be made while using the LSR technique, such as the absence of a linear relationship between independent variables. The model reliability is contingent on the LSR assumptions being met. When all of these conditions are met, the LSR approach can produce the best results. Multicollinearity is defined as a linear relationship between the independent variables. While multicollinearity improves the covariance and variance of regression coefficients, none or only a few of the explanatory variables will be statistically significant if the model's R^2 value is changed. As a result, model misunderstanding is caused by multicollinearity [30].

The LSR method has a significant benefit over all other estimating methods, such as traverse adjustments, in which it is theoretically and statistically defensible and hence a fully rigorous procedure. By minimizing the sum of the squares of the mistakes to a minimal value, the LSR principle states that it is possible to determine the most likely values of a system of unidentified variables on which measurements have been made.

2.3. LASSO Method

For linear modeling, LASSO stands for Least Absolute Shrinkage and Selection Operator. In linear models, the LASSO regression method is a regularization technique for estimating unknown parameters [31]. It modifies model parameters based on a loss function while taking model complexity into account. When there are significant correlations between predictors, the LASSO will randomly select one and ignore the others, and when all predictor variables are comparable, it will collapse [32]. The LASSO penalty forces the majority of coefficients to be zero, with only a small percentage likely to be nonzero. It usually consists of a linear model with an additional regularization parameter. The following is the objective function [33] that must be minimized

$$\text{Min}_{\beta} \frac{1}{2n_{\text{samples}}} ||Y - X\beta||_2^2 + \lambda ||\beta||_1 \quad (2)$$

where $\lambda \geq 0$ signifies the regularization constant and $||\beta||_1$ denotes the coefficient vector's ℓ_1 -norm penalty. For correctly selected λ , the ℓ_1 penalty allows the LASSO to regularize the least squares fit while simultaneously decreasing certain components of β to zero. The cyclical coordinate descent [32] method is faster than the LRS method [28] at evaluating all LASSO solution routes for λ . The LASSO is a common feature selection approach due to these factors. Nonetheless, the LASSO has three major flaws: it lacks the oracle feature, is incompatible with high-dimensional data, and cannot select more features than the sample size without saturating [34]. On the other hand, because the coefficients have been decreased and deleted, it can deliver exact predictions without a large increase in biases.

2.4. Elastic-Net (ENet) Method

The ENet approach is a LASSO extension that can handle features with strong correlations [32]. The ENet was suggested to analyze high-dimensional data in order to avoid the instability of LASSO when predictors were significantly linked [34]. The ENet ap-

proach combines ℓ_1 (Lasso) and ℓ_2 (Ridge Regression) penalties, and it can be represented mathematically as [28]:

$$\text{Min}_{\beta} \frac{1}{2n_{\text{samples}}} \|Y - X\beta\|_2^2 + \lambda \rho \|\beta\|_1 + \frac{\lambda(1-\rho)}{2} \|\beta\|_2^2 \quad (3)$$

The regularization constant ρ is ℓ_1 ratio, and the ℓ_1 and ℓ_2 norm penalties of regularization coefficients are $\|\beta\|_1$ and $\|\beta\|_2^2$. When numerous properties are linked together, the ENet is effective. The ENet's ℓ_1 section automatically chooses variables, whereas the ℓ_2 segment allows for grouped selection and regulates solution paths to improve prediction. The ENet may discover and choose groups of correlated variables even when the groupings are unknown by using a grouping effect throughout the variable selection, such that a group of very correlated features appears to have coefficients of the same scale. When $p \gg n$, where n is the sample size and p is the predictor numbers, the ENet chooses more than n variables, unlike the LASSO. The ENet, on the other hand, lacks the oracle property.

2.5. Bayesian Ridge Regression (BRR) Method

BRR is a subset of Bayesian regression that belongs to the ridge regression category, and it possesses all of the properties of both Bayesian and ridge regression [35]. The Gaussian distribution is used in Bayesian regression, which is a probabilistic method [36]. In order to optimize posterior prediction, the ℓ_1 regularization must be used. The weighted coefficient W is determined from a Gaussian distribution, which is the only difference between Bayesian and BRR. The Gaussian prior considerably reduces the magnitude of all effects in BRR. A spherical Gaussian provides the prior W for the coefficient:

$$P(W|\lambda) = (W|0, \lambda^{-1}I_p) \quad (4)$$

The priors over α and λ are gamma distributions, which are the conjugate prior for Gaussian precision. By default, the value of λ_1 , λ_2 , α_1 , and α_2 is 10^{-6} . BRR is the name given to the resulting model, which is comparable to the classic Ridge. I_p represents the identity matrix. Although BRR takes a long time to run, it excels at adjusting to small data parameters and is simple to use in regularization problems and change hyper-parameters. Furthermore, the BRR focuses on feature selection in order to limit the number of inputs and then rank them according to their predictive value to the estimation method. In reality, BRR produces similar results to the maximum a posterior (MAP) technique.

3. Optical Sensor Numerical Models

3.1. Optical Sensor

Due to their great precision and sensitivity, optical sensors are becoming increasingly popular. According to their design and measurement procedures, the most common optical sensors can be grouped into three categories: (1) spectroscopy-based sensors, (2) SPR-based sensors, and (3) negative curvature-based sensors. Fiber optical sensors have been developed for various applications such as environmental monitoring, salinity monitoring, temperature monitoring, magnetic fluid monitoring, agricultural monitoring, industrial monitoring, food quality monitoring, health monitoring, and military operations over the last two decades in medicine, air force, and industry. Optical biosensors, among many uses of optical sensors, usher in a new age in the field of biological monitoring and detection [3,37].

Optical biosensors are analytical smart devices that are used to monitor a variety of biological components in conjunction with the appearance of chemicals or analytes. Nowadays, PCF-based biosensors are utilized to detect physiological components such as RBCs, hemoglobin, plasma cells, WBCs, cancer cells, tuberculosis, pregnancy, urine, medications, virus, SARS-CoV-2, and so on. The large variety of real-time applications of optical biosensors, as well as design freedom, fast detection time, compact size, simplicity, resilience, and low cost, have piqued researchers' interest. As a result, in recent years, a

number of researchers have looked at using plasmonic materials rather than gold and silver to improve the quality of optical biosensors. ML techniques are also used to speed up the detection process and improve accuracy. This quick biochemical diagnosis will help to reduce the danger of contamination and exposure [38–41].

Figure 1 shows the cross-sectional illustration of an optical biosensor with specifications. Core radius, cladding radius, pitch, analyte, and wavelength are five optical design factors for the proposed biosensor. The radius of the core and cladding in this study ranged from 2.8 μm to 3.00 μm , with a pitch of between 7.00 μm and 8.00 μm . Furthermore, during the operating range of wavelength 1.4 μm to 2.0 μm , the examined analyte range was 1.33 to 1.35. Some of the most important optical guiding properties are described as follows.

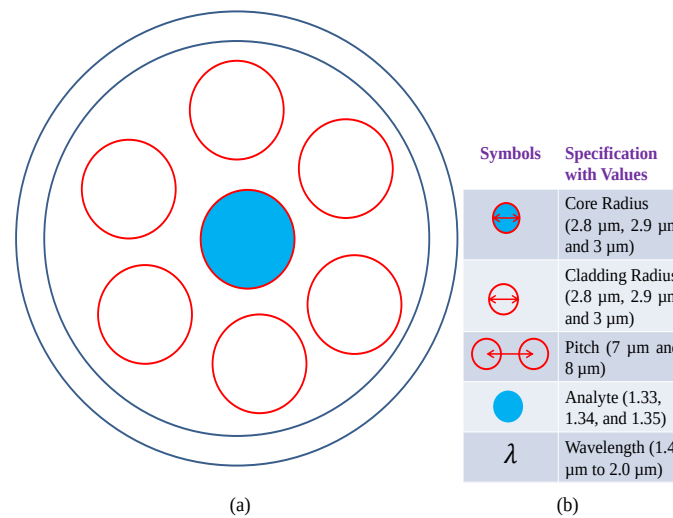


Figure 1. (a) Cross-sectional image of the proposed optical biosensor; (b) Design-dependent parameters or specifications.

3.2. Effective Refractive Index (ERI)

ERI is a number that compares the phase delay in a waveguide per unit length to the phase delay in a vacuum. ERI is symbolized as n_{eff} , with RIU as the unit. ERI can be expressed mathematically as follows:

$$n_{\text{eff}} = A \pm iB \quad (5)$$

where A stands for the real value of ERI and B stands for an imaginary value of ERI.

3.3. Optical Power Profiles (OPP)

Optical sensors are used to monitor buildings that originate, supply, distribute, and transform electrical power in the energy field. Core power, cladding power, and total power are the three categories of optical sensor energies. An optical sensor's performance is ensured by its high core power.

3.4. Optical Power Fraction (OPF)

The OPF is a unitless quantity that specifies how much light enters the optical sensor. Core and cladding optical power fractions are two different forms of optical power fractions. Core power fractions (CPF) with high values suggest the highest sensitivity and lowest cladding power fraction (CLPF). Mathematically, the total, core, and cladding powers are computed as follows [3,37–39]:

$$\text{Total Power} = \int_{\text{total}} \text{Re}(E_x H_y - E_y H_x) dx dy \quad (6)$$

$$\text{Core Power} = \int_{\text{Core}} \text{Re}(E_x H_y - E_y H_x) dx dy \quad (7)$$

$$\text{Cladding Power} = \int_{\text{Cladding}} \text{Re}(E_x H_y - E_y H_x) dx dy \quad (8)$$

where the transverse electric and magnetic fields are denoted by the symbols E_x , E_y , H_x , and H_y , respectively. Mathematically, the core and cladding power fractions are computed as follows [3,37–39]:

$$\text{CPF} (\%) = \frac{\text{Core Power}}{\text{Total Power}} \times 100 \quad (9)$$

$$\text{CLPF} (\%) = \frac{\text{Cladding Power}}{\text{Total Power}} \times 100 \quad (10)$$

3.5. Optical Effective Area (OEA)

The area that a waveguide or fiber mode successfully covers in the transverse dimensions is measured in this mode of operation. Fiber or other waveguide modes have smooth transverse profiles, making defining a mode area difficult, especially for sophisticated mode shapes where a $\frac{1}{e^2}$ intensity requirement, such as for Gaussian beams, is not appropriate. OEA is symbolized as A_{eff} , with m^2 as the unit. The OEA can be defined as follows [3,37–39]:

$$A_{\text{eff}}(\text{m}^2) = \frac{(\int \int |E|^2 dx dy)^2}{(\int \int |E|^4 dx dy)} \quad (11)$$

where E is the amplitude of the electric field.

3.6. Optical Loss Profiles (OLP)

To fulfill the boundary condition, a circular perfectly matched layer (PML) is utilized, which prevents possible reflection at the border. The imaginary component of the ERI can be used to calculate confinement or leakage loss. OLP is symbolized as L_c , with dB/cm as the unit. The confinement or leakage loss is calculated as follows [3,37–39]:

$$L_c(\text{dB/cm}) = 8.686 \times \frac{2\pi}{\lambda} \times \text{Im}(n_{\text{eff}}) \times 10^4 \quad (12)$$

In this case, the wave number in free space is $\frac{2\pi}{\lambda}$. L_c is the result of scattering light escaping from the core to the outer materials, which is calculated using the imaginary portion of n_{eff} at the specific wavelength λ .

3.7. Optical Sensing Profile (OSP)

The relative sensitivity coefficient measures the interaction between light and the analyte, which can be calculated as follows [3,37–39].

$$\text{Sensitivity}(\%) = \frac{n_r}{n_{\text{eff}}} \times \text{CPF} \quad (13)$$

where n_r denotes the refractive index of the detected liquid within the core and n_{eff} denotes the model effective index.

3.8. Limitations and Proposed Solutions

Due to restrictions, any changes in optical sensor design parameters will result in changes in sensor performance, which is time-consuming to optimize. Different ML

parameter estimate models that are faster and more accurate can be used to solve the optical sensor design optimization challenge.

4. Methodology

For the proposed research, quantitative data is necessary. The data was gathered through simulation. After some preparation, such as handling noisy data and measuring the correlation between variables using Pearson's correlation testing method, a valid dataset was formed. In order to determine the best one, regression techniques are applied. Figure 2 illustrates the full procedure of the research.

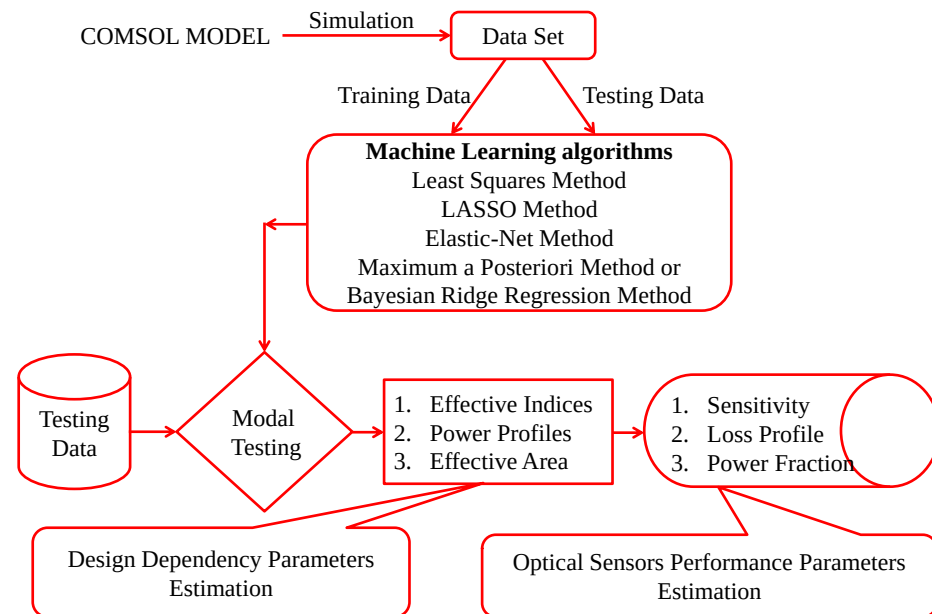


Figure 2. Flowchart of the research's procedure. First, the design dependency parameters are estimated, and then the optical Sensors' Performance Parameters are calculated.

4.1. Design and Dataset Collection

Data collection is a methodical process for gathering and analyzing particular information in order to respond to pertinent questions and evaluate results. It focuses on discovering all there is to know about a certain subject. Data is obtained in an effort to test hypotheses in an effort to comprehend a phenomenon. Here was created a rudimentary PCF model, which is shown in Figure 1.

4.2. Dataset Distribution

A technique or list that shows all potential values (or intervals) is a distribution of the data. Additionally, it displays the frequency of each value, which is important. Figures make it simple to evaluate both the numbers and the frequency in which they exist. A project's data is typically organized from the smallest to the largest. The dataset was split into training and testing portions by 70% and 30%, respectively, in this research.

4.3. Training, Testing and Evaluation

The dataset was fed into the ML algorithms for training, and the model was evaluated using a new tuple. In this research, 690 data points from 1402 were utilized to train the models, and 13 data points were used to visualize the model performance.

5. Result Analysis and Discussion

The trained models have been validated in this section. Unknown input parameters have been used to test the model. With wavelength changes ranging from 1.4 μm to 2.0 μm ,

the suggested model has been trained to predict effective indices, core power, and total power. After that, sensitivity, confinement loss, core power percentage, and effective mode area are calculated. This section includes various graphical visualizations and comparison tables in terms of the R^2 -score, MSE, and MAE.

5.1. ERI for X-Axis and Y-Axis

The predicted and simulated outcome values are represented by a straight line and different symbols in Figure 3a,c for real values of ERI in the X-axis and Y-axis, respectively. Methods are represented by different color dots, and the solid line represents the simulated values. The ERI values drop as the wavelength rises. The predicted values in Figure 3a,c are substantially equal to the simulated values. The greatest value of R^2 for both the X-axis and the Y-axis is 0.9994 using the LS and BRR methods. However, when compared to other applied techniques, the LS method has the lowest mean square error (MSE) of 3.9729×10^{-8} and 3.9921×10^{-8} for both the X-axis and Y-axis, respectively. Furthermore, all applied approaches have an MAE of 0.0001 in common. Table 1 shows the similarities in greater detail. All of the algorithms predict results that are nearly indistinguishable in the imaginary section of the ERI. At first, the predicted values are lower than the simulated values, but after $1.5 \mu\text{m}$, the predicted values are higher than the simulated values in Figure 4a,c. The LS method has the highest R^2 of 0.9436 and 0.9435, the lowest MSE of 5.7626×10^{-9} and 5.7656×10^{-9} , and the lowest MAE of 6.0969×10^{-5} and 6.1013×10^{-5} for the X-axis and Y-axis, respectively, when compared to all other strategies. Table 2 shows the results of the detailed comparisons and also exhibits that the ENet method has the weakest performance when compared to the other strategies. However, the ENet approach had an absolute error of 0.00015 at $1.85 \mu\text{m}$ wavelength, which is fantastic news. As a result, we may state that practically all used approaches have a 1–5 percent error rate. Furthermore, the results of the proposed sensors in both axes indicate similar results, which is why the fundamental mode is chosen for further investigation.

Table 1. Comparison of different applied methods for estimating the real values of effective refractive index (RIU) when analyte 1.35, pitch = $8 \mu\text{m}$, core and cladding radius = $3 \mu\text{m}$.

Applied Methods	R^2	X-Axis		R^2	Y-Axis	
		$MSE \times 10^{-8}$	MAE		$MSE \times 10^{-8}$	MAE
Least Squares	0.9994	3.9729	0.0001	0.9994	3.9721	0.0001
LASSO	0.9993	4.5486	0.0001	0.9993	4.5581	0.0001
Elastic-Net	0.9991	6.1648	0.0001	0.9992	5.8462	0.0001
B. Ridge Regression	0.9994	4.0301	0.0001	0.9994	4.5095	0.0001

Table 2. Comparisons of different applied methods for estimating the imaginary values of effective refractive index (RIU) when analyte 1.35, pitch = $8 \mu\text{m}$, core and cladding radius = $3 \mu\text{m}$.

Applied Methods	R^2	X-Axis		R^2	Y-Axis	
		$MSE \times 10^{-9}$	$MAE \times 10^{-5}$		$MSE \times 10^{-8}$	$MAE \times 10^{-5}$
Least Squares	0.9436	5.7626	6.0969	0.9435	5.7626	6.1013
LASSO	0.9421	5.9223	6.1785	0.9420	5.9274	6.1804
Elastic-Net	0.9318	6.0952	6.2524	0.9318	6.0964	6.2516
B. Ridge Regression	0.9319	8.3652	7.0367	0.9319	8.3645	7.0382

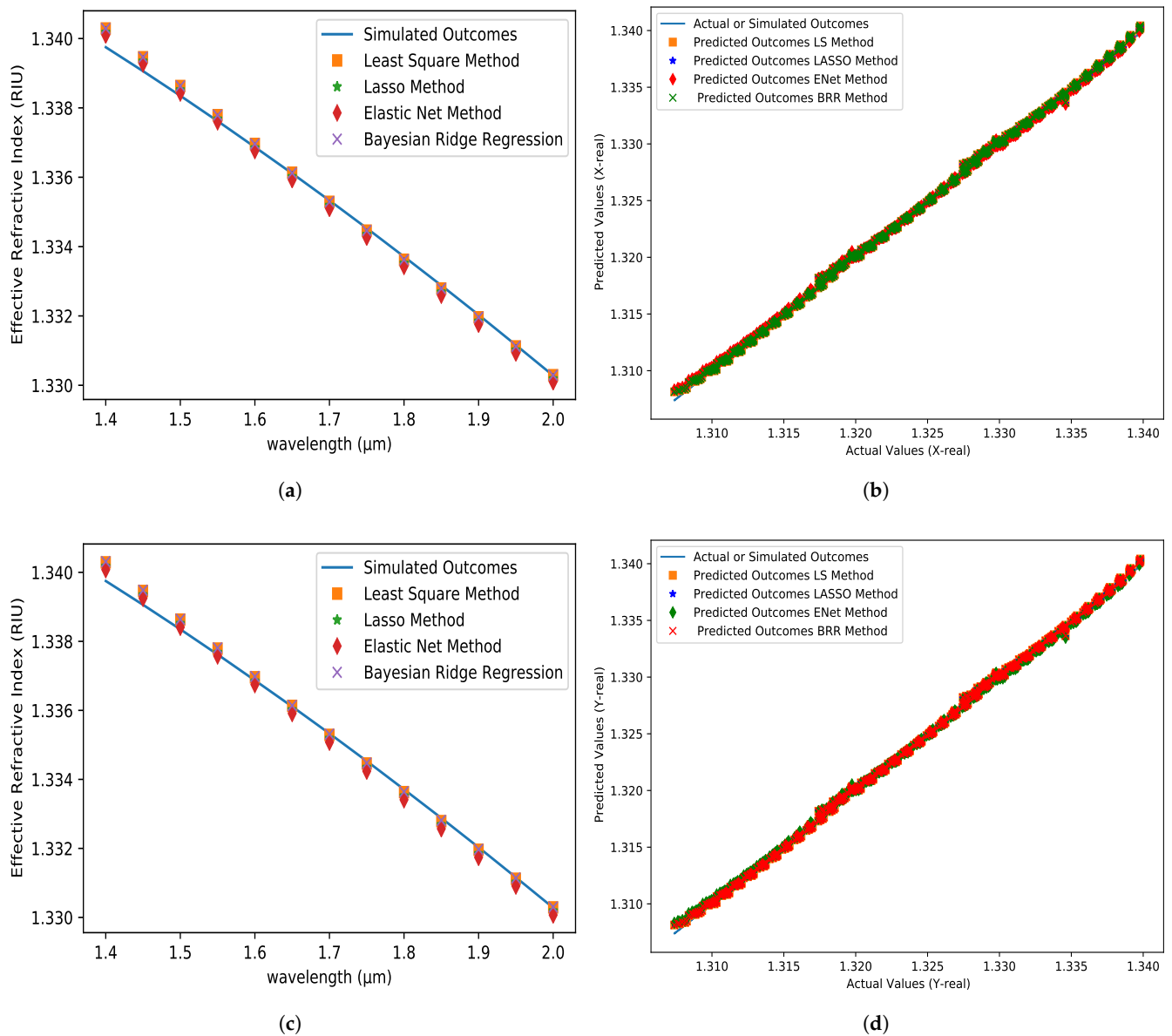


Figure 3. The real value of the effective refractive index (RIU) variations with respect to the wavelength at (a) X-axis and (c) Y-axis for analyte 1.35, pitch = 8 μm , core and cladding radius = 3 μm . The comparison between the predicted dataset (tuning) for different algorithms and the actual dataset obtained from FEM simulation at (b) X-axis and (d) Y-axis.

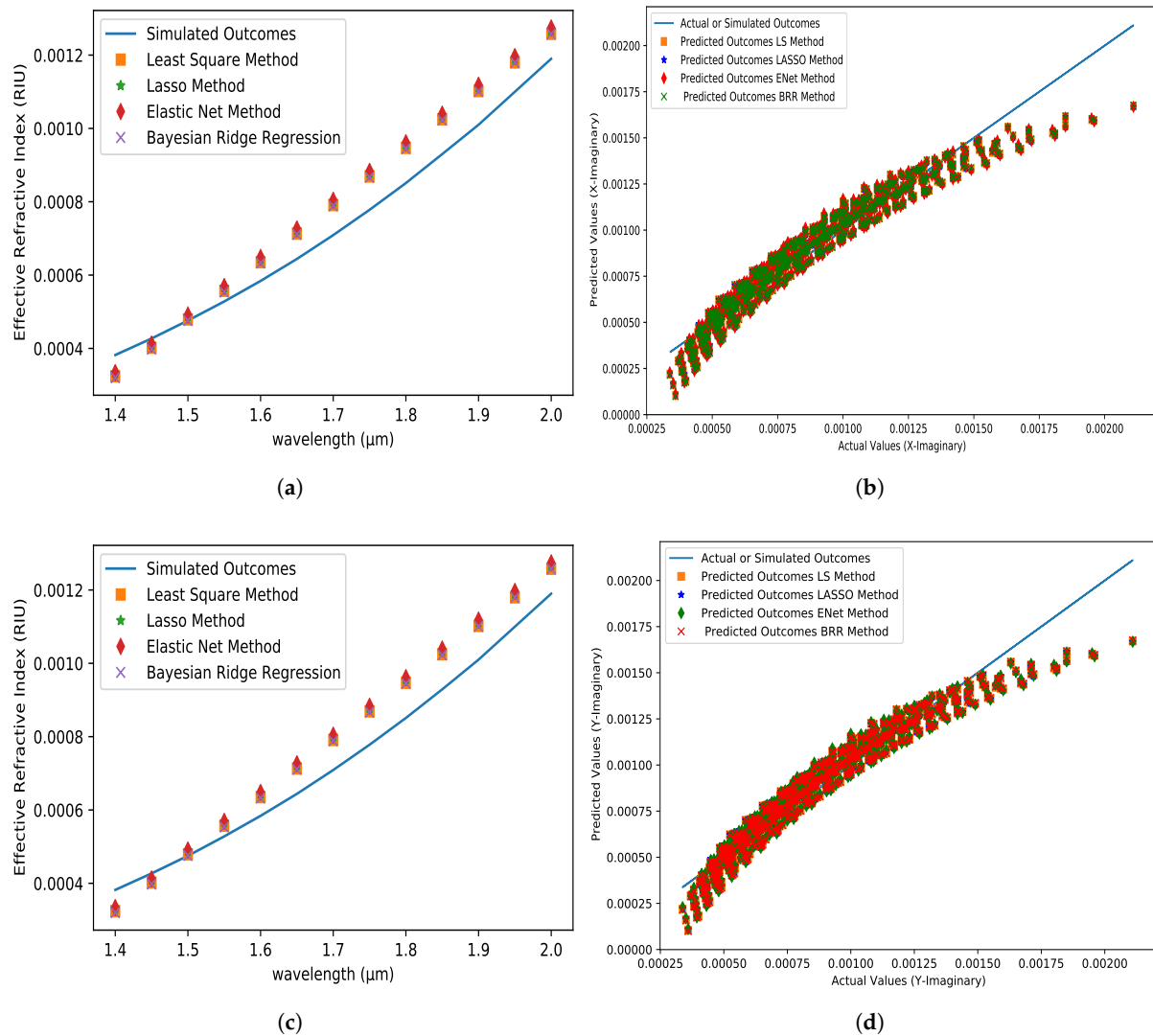


Figure 4. The imaginary value of ERI variations with respect to the wavelength at (a) X-axis and (c) Y-axis for analyte 1.35, pitch = 8 μm, core and cladding radius = 3 μm. The Comparison between the predicted dataset (tuning) for different algorithms and the actual dataset obtained from FEM simulation at (b) X-axis and (d) Y-axis.

5.2. Effective Mode Area (EMA)

Figure 5 depicts the EMA as a function of wavelength variations over the working range of 1.4 to 2.0 μm, which clearly shows that there is no noticeable difference between the estimated model results and the simulated outcomes. In comparison to the four applied models, it is also obvious that LS approaches perform the best and ENet performs the poorest. The absolute inaccuracy for the ENet approach is 3.1×10^{-13} at 1.8 μm wavelength, as seen in Figure 5. It demonstrates that the differences between predicted and simulated results are nearly identical. Furthermore, the MSE of least square, LASSO methods is of the order of 10^{-25} , whereas ENet, BRR methods are of the order of 10^{-24} . Furthermore, the MAE of all applicable methods is of the order of 10^{-13} .

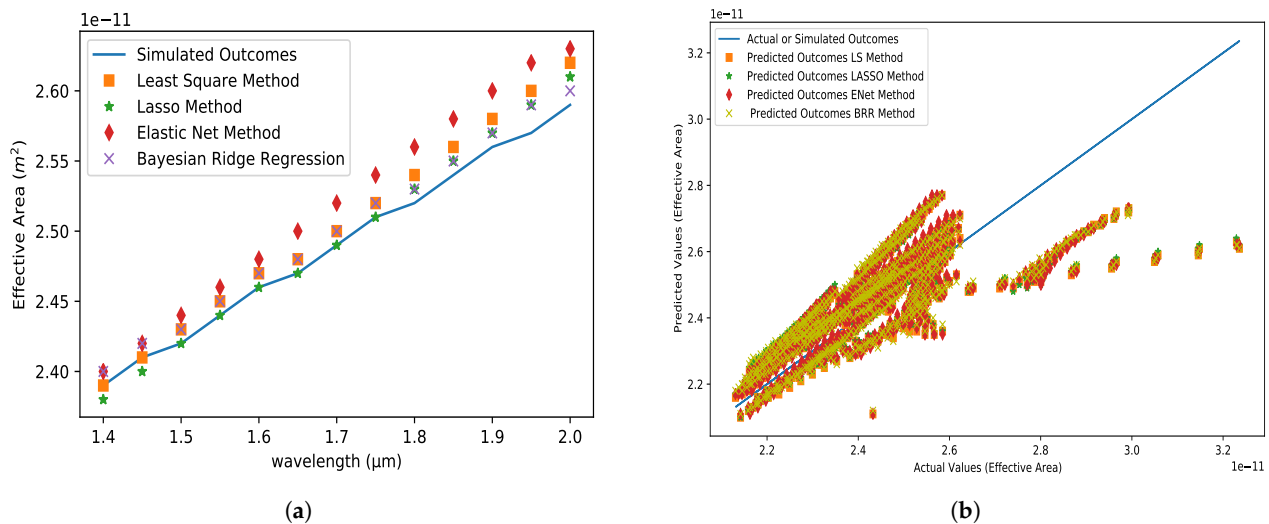


Figure 5. (a) Effective area variations with respect to wavelength for analyte 1.35, pitch = 8 μm , core and cladding radius = 3 μm and (b) the comparison between the predicted dataset for different algorithms and actual dataset obtained from FEM simulation.

5.3. Total Power and Core Power

Core power and total power fluctuate due to light intensity in the core and structural changes in the fiber. The total power values of the PCF model grow as the wavelength increases in Figure 6a. As the wavelength gets longer, the total power values for the least squares and maximum posterior approaches get closer to the simulated one. In contrast, the LASSO and ENet approaches represent the reversed scenario. For lower wavelength values, the expected values are virtually comparable. Figure 6b shows that the ENet and LASSO methods produce lower values than the simulated ones, but the LS and BRR methods produce higher values for wavelengths between 1.4 μm and 1.8 μm , and then the values approach the simulated values.

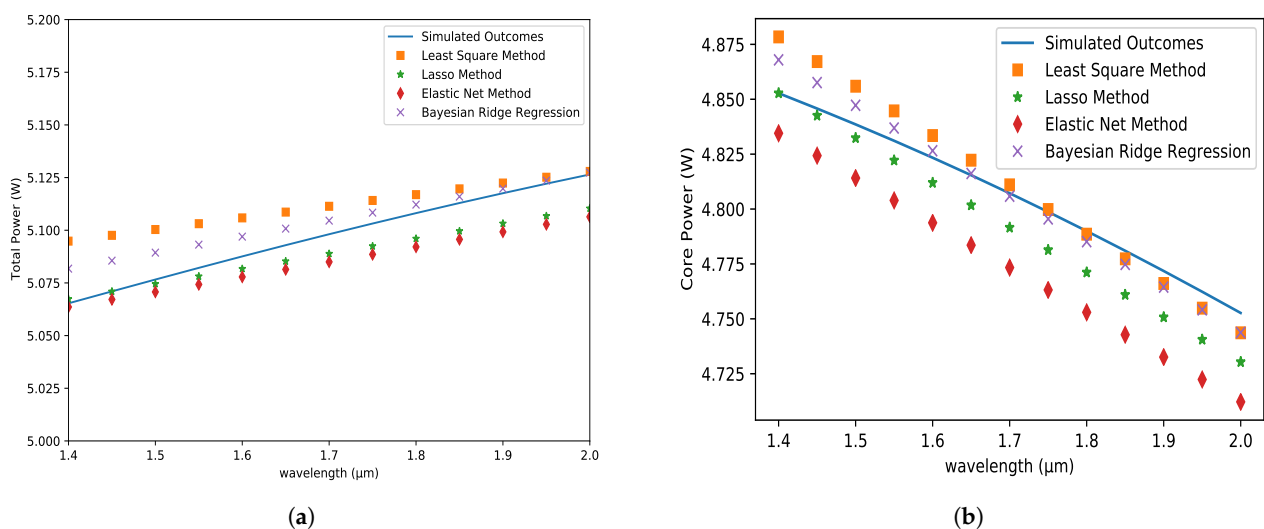


Figure 6. (a) Total Power and (b) Core Power variations with respect to the wavelength at the X-axis for analyte 1.35, pitch = 8 μm , core and cladding radius = 3 μm .

5.4. Core Power Fraction (CPF)

A power fraction is needed to calculate sensitivity. It refers to the ratio of total power to core power. The suggested PCF model has a lower power percentage for longer wave-

lengths. Figure 7a shows that all of the approaches had lower values than the simulated ones. Furthermore, the LS, LASSO, and BRR techniques predict closer outcomes at lower wavelengths and show absolute errors on the order of 10^{-5} at the lower operating wavelength ($1.45 \mu\text{m}$), respectively, while ENet displays 10^{-2} . At the upper working point of wavelength ($1.90 \mu\text{m}$), the LS, LASSO, and BRR approaches show an absolute error of the order of 10^{-3} , while the ENet method displays 10^{-2} .

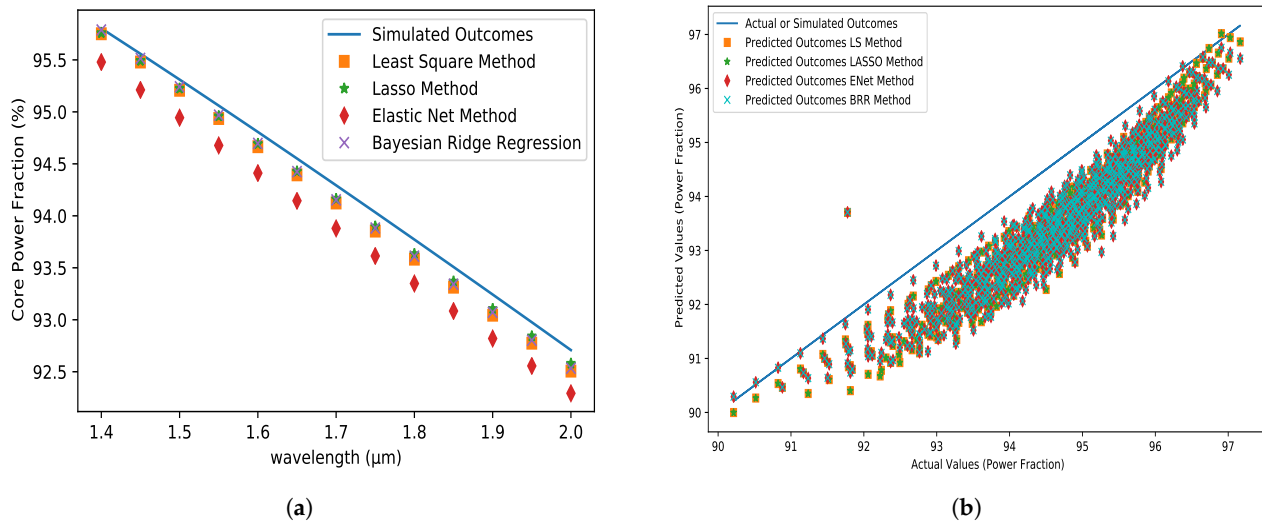


Figure 7. (a) CPF (%) variations with respect to the wavelength at the X-axis for analyte 1.35, pitch = $8 \mu\text{m}$, core and cladding radius = $3 \mu\text{m}$ and (b) the comparison between the predicted dataset for different algorithms and actual dataset obtained from FEM simulation.

5.5. Confinement Loss Profile (CLP)

CL is a critical feature of PCF, and the mode's leaky nature causes it. The wavelength and the imaginary component of the effective index of the core-guided mode determine it. The CL is increasing in response to the increase in wavelength. The CL of the proposed model is shown to be very small in this research. The CL values for all of the techniques in Figure 8a are the same, with predicted values smaller than simulated values in the range of $1.4 \mu\text{m}$ to $1.5 \mu\text{m}$. All of the other values are greater than the simulated values. When all approaches are used, the absolute inaccuracy is on the order of 10^{-5} .

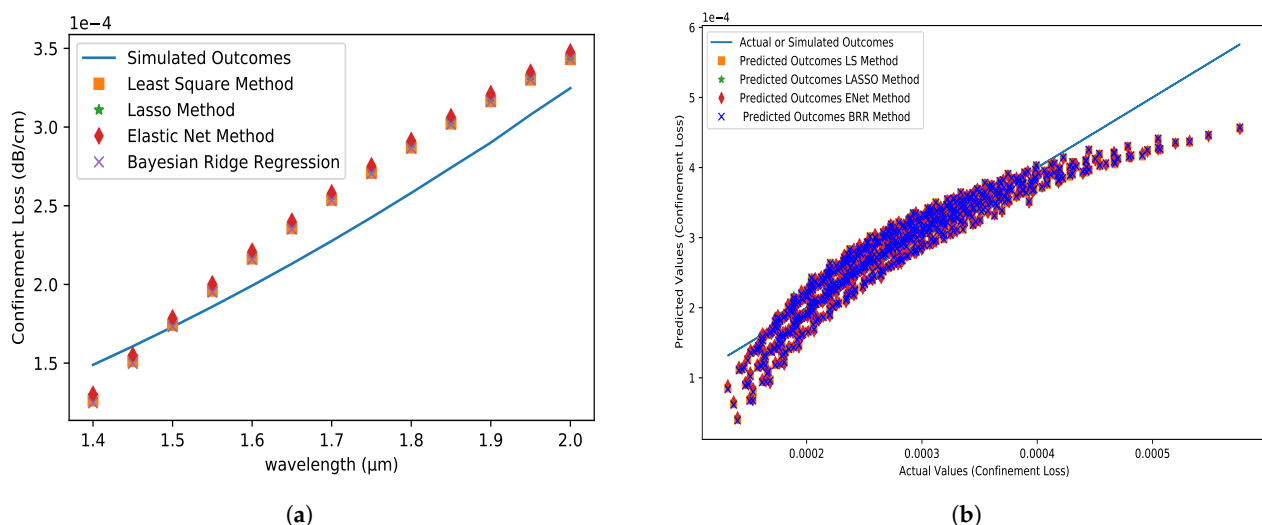


Figure 8. (a) CL (dB/cm) profile variations with respect to the wavelength at the X-axis for analyte 1.35, pitch = $8 \mu\text{m}$, core and cladding radius = $3 \mu\text{m}$ and (b) the comparison between the predicted dataset for different algorithms and actual dataset obtained from FEM simulation.

5.6. Optical Sensitivity Profile (OSP)

The power fraction, analyte, and ERI all influence PCF sensitivity. The predicted values of various ML models, as well as simulated values, were used. The suggested PCF model has a higher sensitivity for shorter wavelengths. As demonstrated in Figure 9a, all of the predicted values are identical to the simulated values. For assessing the sensitivity profile of the proposed optical sensor, almost all approaches function well, with absolute errors in the order of 10^{-8} . Table 3 demonstrates that the LS method outperforms other strategies, with an R^2 of 0.9994, an MSE of 3.9×10^{-8} , and an MAE of 1.5×10^{-4} . The ENet approach, on the other hand, has the lowest performance, with an R^2 of 0.9990, an MSE of 6.2×10^{-8} , and an MAE of 1.9×10^{-4} .

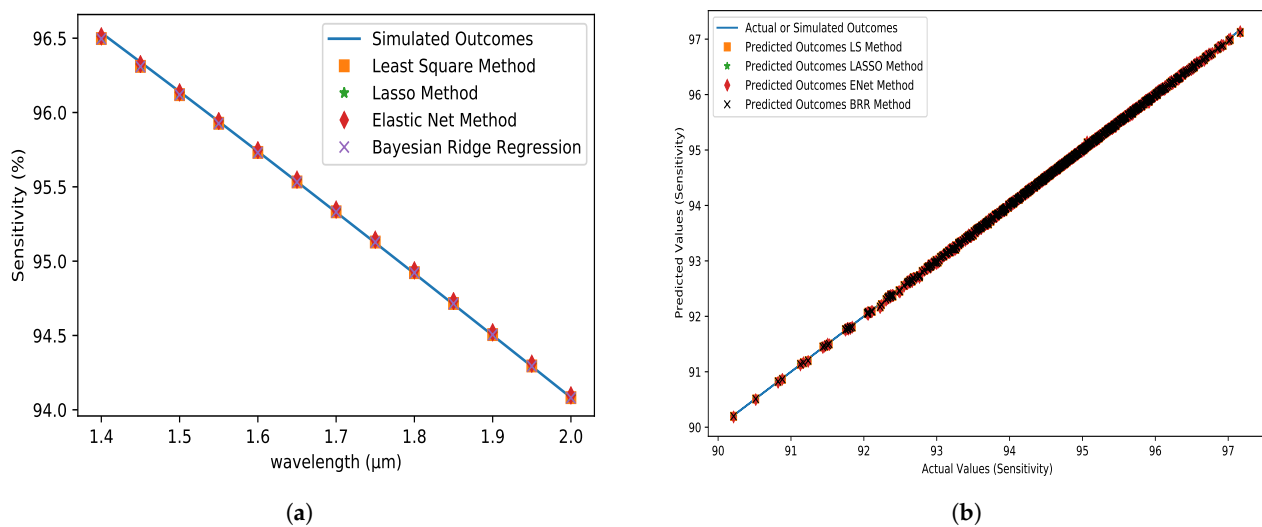


Figure 9. (a) Sensitivity (%) variations with respect to the wavelength at the X-axis for analyte 1.35, pitch = 8 μm, core and cladding radius = 3 μm. (b) The Comparison between the predicted dataset for different algorithms and the actual dataset obtained from FEM simulation.

Table 3. Comparison of performance evaluation of different Parameter Estimation Methods for sensitivity profile.

Applied Methods	R^2	$MSE \times 10^{-8}$	$MAE \times 10^{-4}$
Least Squares Method	0.9994	3.90	1.50
LASSO Method	0.9993	4.50	1.60
Elastic-Net Method	0.9990	6.20	1.90
Bayesian Ridge Regression Method	0.9994	4.00	1.90

5.7. OSP Evaluation for Different Volume of Datasets

The performance of different methods for the proposed sensor and various numbers of input datasets is shown in Figure 10. It is obvious from the illustration that the performance of applied algorithms improves as the amount of trained data grows. In comparison to dataset-1, datasets 2 and 3 include 1.5 and 2 times more data. It is now clearly demonstrated that when the volume of the dataset grows, the performance of the applied methods for the presented sensors improves as well. The changes in the approaches used are, however, relatively minor. The various performance indicator values of the employed methods are shown in Table 4. Finally, we can state that the applied ML algorithms exhibit remarkably consistent performance across a wide range of trained datasets.

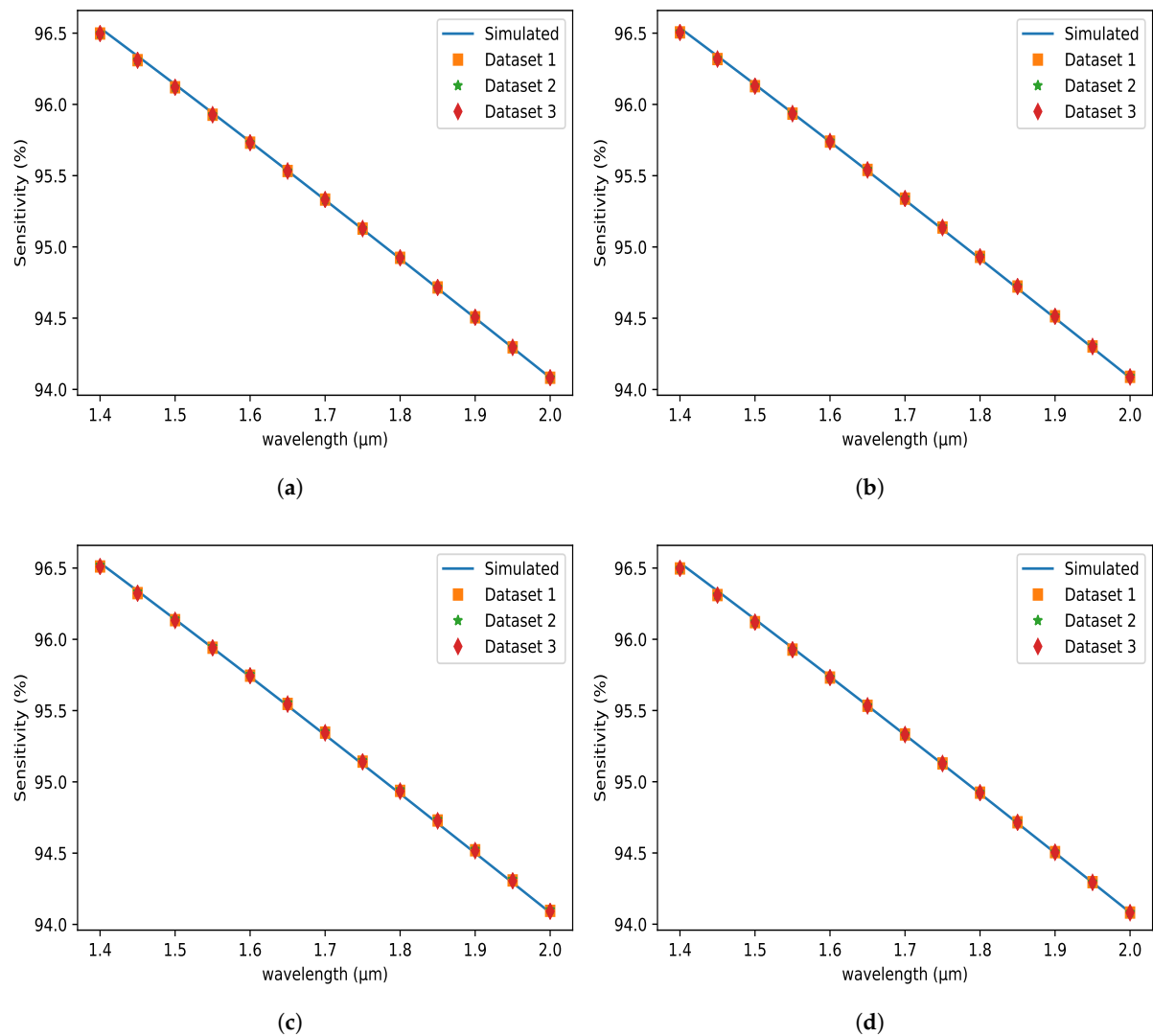


Figure 10. Sensitivity (%) variations with respect to the wavelength at the X-axis for different datasets using (a) LS (b) LASSO (c) ENet (d) BRR Methods.

Table 4. Comparison of performance evaluation of different Parameters Estimation Methods for different volumes of the dataset.

Applied Methods	Dataset	R^2	$MSE \times 10^{-8}$	$MAE \times 10^{-4}$
Least Squares Method	Dataset-1	0.9994	3.97	1.51
"	Dataset-2	0.9993	3.86	1.50
"	Dataset-3	0.9995	3.66	1.56
LASSO Method	Dataset-1	0.9993	4.55	1.58
"	Dataset-2	0.9994	4.41	1.66
"	Dataset-3	0.9994	4.63	1.61
Elastic-Net Method	Dataset-1	0.9991	6.16	1.91
"	Dataset-2	0.9991	6.30	2.01
"	Dataset-3	0.9993	5.82	1.82
Bayesian Ridge Regression Method	Dataset-1	0.9994	4.03	1.52
"	Dataset-2	0.9994	3.95	1.56
"	Dataset-3	0.9994	4.57	1.60

5.8. OSP Evaluation for Different Volume of Outliers

The sensitivity performance of several approaches for various volumes of outliers is depicted in Figure 11. It is obvious from the preceding study that the applied approaches produce stable results for the input data when precursors are absent. To test the performance of the applied methods, 5% and 10% outliers are introduced to the input data, and the methods are trained repeatedly. In comparison to Dataset-3, Datasets 4 and 5 include 5% and 10% outliers, respectively. The provided approaches then demonstrated stable performance for the proposed sensor with outlier data as well. Table 5 depicts the changes in technique performance by displaying several performance indicator values. The performance of different applied procedures, however, varies slightly. Even so, the effectiveness of the LS method is consistent across diverse groups of outliers.

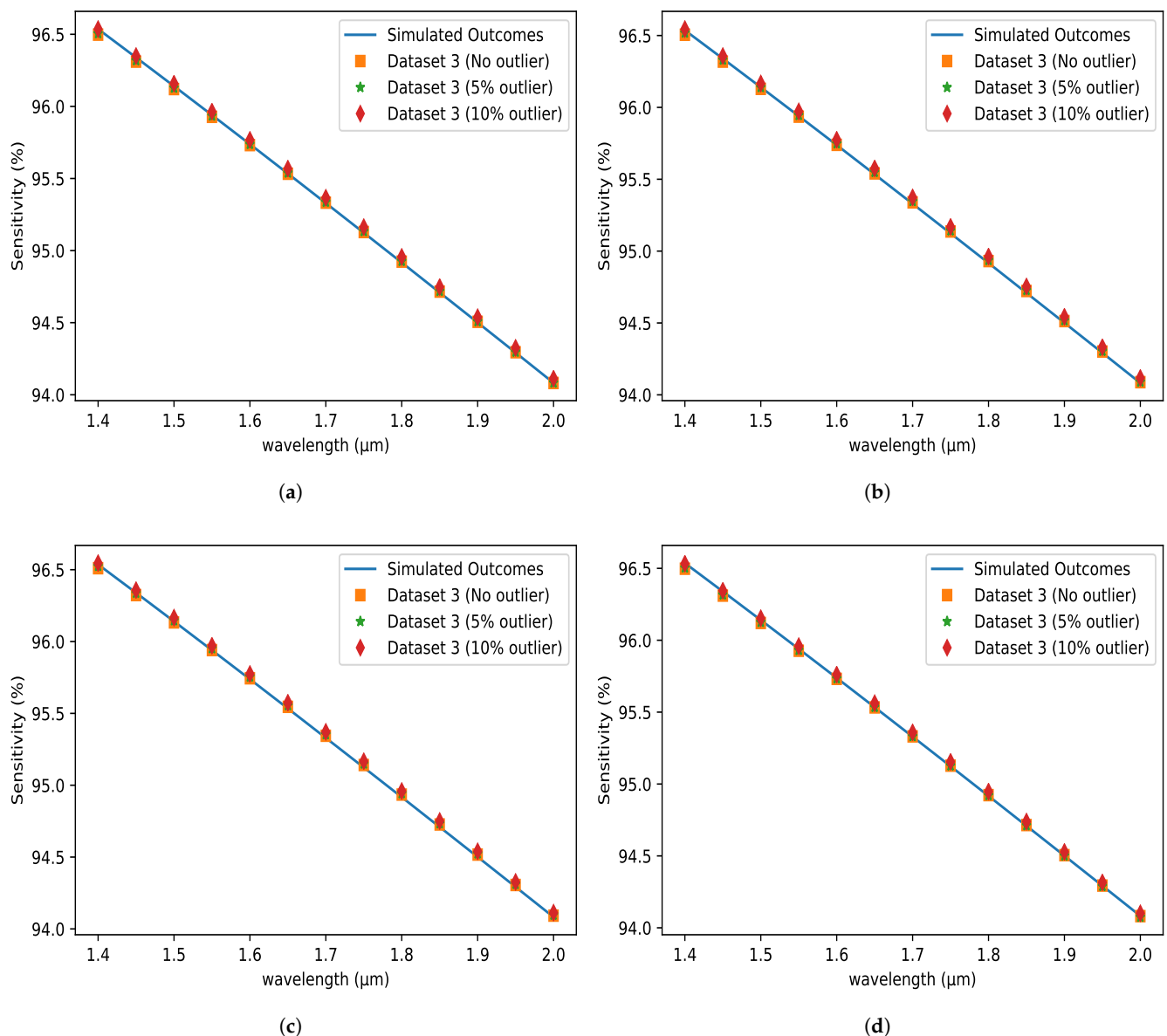


Figure 11. Sensitivity (%) variations with respect to the wavelength at the X-axis for different outliers using (a) LS (b) LASSO (c) ENet (d) BRR Methods.

Table 5. Comparison of performance evaluation of different Parameter Estimation Methods for different volumes of outliers.

Applied Methods	Dataset	R^2	MSE	MAE
Least Squares Method	Dataset-3	0.9995	3.66×10^{-8}	1.56×10^{-4}
"	Dataset-4	0.9897	8.29×10^{-7}	3.05×10^{-4}
"	Dataset-5	0.9657	2.73×10^{-6}	5.84×10^{-4}
LASSO Method	Dataset-3	0.9994	4.63×10^{-8}	1.61×10^{-4}
"	Dataset-4	0.9893	8.55×10^{-7}	3.04×10^{-4}
"	Dataset-5	0.9655	2.76×10^{-6}	6.00×10^{-4}
Elastic-Net Method	Dataset-3	0.9993	5.82×10^{-8}	1.82×10^{-4}
"	Dataset-4	0.9890	8.81×10^{-7}	3.20×10^{-4}
"	Dataset-5	0.9393	5.02×10^{-6}	6.01×10^{-4}
Bayesian Ridge Regression Method	Dataset-3	0.9994	4.57×10^{-8}	1.60×10^{-4}
"	Dataset-4	0.9854	1.20×10^{-6}	3.47×10^{-4}
"	Dataset-5	0.9395	5.01×10^{-6}	5.74×10^{-4}

5.9. Overall Performance Evaluation

As a consequence, the LS method clearly outperforms the other four techniques when it comes to predicting outcomes. Because the trained data was properly distributed with no outliers, the LS offers the best results. So, we also put the performance of the ML techniques to the test by merging outliers ranging from 5% to 10% of the total data. We noticed that LS R^2 -score performance was decreasing, from 99.94 percent to 96.52 percent, while other techniques did not degrade as significantly. After training the model, fast and precise prediction of PCF parameters is possible in a few milliseconds, which is approximately 99 percent faster than standard numerical simulation approaches. Finally, for optimizing an optical biosensor design, the proposed LS technique is the most effective when high-quality training data is available.

6. Conclusions

We proposed four alternative machine learning algorithms in this study to predict optical sensor design-dependent characteristics as well as optical sensor attributes. By adjusting the core and cladding radius (2.8 μm to 3 μm), pitches (7 μm and 8 μm), analytes (1.33 to 1.35) and wavelengths (1.4 μm to 2.00 μm), the mode effective indices, effective area, core, and total power confinement of suggested optical biosensors are obtained. A total of 690 data points are utilized for training the proposed ML model, with another 13 data points used to visualize the trained model performance. Almost all machine learning algorithms effectively anticipate optical design-dependent parameters and optical sensor qualities (about 99.94%), which are close to simulated values. The best LS technique achieves the highest R^2 -score of 0.9994, the lowest MAE of 3.9×10^{-8} , and the MSE of 1.5×10^{-4} for sensitivity profile prediction. On the other hand, the worst Elastic-Net approach has an R^2 -score of 0.9990, an MAE of 6.2×10^{-8} , and an MSE of 1.9×10^{-4} . As a result, it is evident that the least squares method outperforms the other three techniques in terms of prediction outcomes. We also test the performance of the ML approaches by combining outliers ranging from 5% to 10% of the total data. We found that the R^2 -score performance of LS fell from 99.94 percent to 96.52 percent. Finally, the proposed LS approach is efficient for optimizing an optical biosensor design, assuming that the model parameters are consistent and that training data conform to the true distribution without outliers.

Author Contributions: Conceptualization, K.A. and F.M.B.; methodology, F.-X.W., K.A. and F.M.B.; software, K.A.; validation, F.-X.W. and F.M.B.; formal analysis, K.A.; investigation, K.A.; resources, F.M.B.; data curation, K.A.; writing—original draft preparation, K.A.; writing—review and editing, F.-X.W. and F.M.B.; visualization, K.A.; supervision, F.-X.W. and F.M.B.; project administration, F.-X.W. and F.M.B.; funding acquisition, F.M.B. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by funding from the Natural Sciences and Engineering Research Council of Canada (NSERC).

Data Availability Statement: The dataset is available upon request to the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Knight, J.; Birks, T.; Russell, P.S.J.; Atkin, D. All-silica single-mode optical fiber with photonic crystal cladding. *Opt. Lett.* **1996**, *21*, 1547–1549. [\[CrossRef\]](#)
2. Paul, B.K.; Rajesh, E.; Asaduzzaman, S.; Islam, M.S.; Ahmed, K.; Amiri, I.S.; Zakaria, R. Design and analysis of slotted core photonic crystal fiber for gas sensing application. *Results Phys.* **2018**, *11*, 643–650. [\[CrossRef\]](#)
3. Ahmed, K.; Paul, B.K.; Vasudevan, B.; Rashed, A.N.Z.; Maheswar, R.; Amiri, I.; Yupapin, P. Design of D-shaped elliptical core photonic crystal fiber for blood plasma cell sensing application. *Results Phys.* **2019**, *12*, 2021–2025. [\[CrossRef\]](#)
4. Arif, M.F.H.; Hossain, M.M.; Islam, N.; Khaled, S.M. A nonlinear photonic crystal fiber for liquid sensing application with high birefringence and low confinement loss. *Sens. Bio-Sens. Res.* **2019**, *22*, 100252. [\[CrossRef\]](#)
5. Shi, J.; Chai, L.; Zhao, X.; Li, J.; Liu, B.; Hu, M.; Li, Y.; Wang, C. Femtosecond pulse coupling dynamics between a dispersion-managed soliton oscillator and a nonlinear amplifier in an all-PCF-based laser system. *Optik* **2017**, *145*, 569–575. [\[CrossRef\]](#)
6. Cheo, P.; Liu, A.; King, G. A high-brightness laser beam from a phase-locked multicore Yb-doped fiber laser array. *IEEE Photonics Technol. Lett.* **2001**, *13*, 439–441. [\[CrossRef\]](#)
7. Holzwarth, R.; Udem, T.; Hänsch, T.W.; Knight, J.; Wadsworth, W.; Russell, P.S.J. Optical frequency synthesizer for precision spectroscopy. *Phys. Rev. Lett.* **2000**, *85*, 2264. [\[CrossRef\]](#)
8. Markin, A.V.; Markina, N.E.; Goryacheva, I.Y. Raman spectroscopy based analysis inside photonic-crystal fibers. *TrAC Trends Anal. Chem.* **2017**, *88*, 185–197. [\[CrossRef\]](#)
9. Couny, F.; Carraz, O.; Benabid, F. Control of transient regime of stimulated Raman scattering using hollow-core PCF. *JOSA B* **2009**, *26*, 1209–1215. [\[CrossRef\]](#)
10. Bréchet, F.; Marcou, J.; Pagnoux, D.; Roy, P. Complete analysis of the characteristics of propagation into photonic crystal fibers, by the finite element method. *Opt. Fiber Technol.* **2000**, *6*, 181–191. [\[CrossRef\]](#)
11. Cucinotta, A.; Selleri, S.; Vincetti, L.; Zoboli, M. Hole fiber analysis through the finite-element method. *IEEE Photonics Technol. Lett.* **2002**, *14*, 1530–1532. [\[CrossRef\]](#)
12. Joannopoulos, J.D.; Villeneuve, P.R.; Fan, S. Photonic crystals: Putting a new twist on light. *Nature* **1997**, *386*, 143–149. [\[CrossRef\]](#)
13. Fanglei, L.; Gaoxin, Q.; Yongping, L. Analyzing point defect two-dimensional photonic crystals with transfer matrix and block-iterative frequency-domain method. *Chin. J. Quantum Electron.* **2003**, *20*, 35–41.
14. Shi, S.; Chen, C.; Prather, D.W. Plane-wave expansion method for calculating band structure of photonic crystal slabs with perfectly matched layers. *JOSA A* **2004**, *21*, 1769–1775. [\[CrossRef\]](#)
15. Hsue, Y.C.; Freeman, A.J.; Gu, B.Y. Extended plane-wave expansion method in three-dimensional anisotropic photonic crystals. *Phys. Rev. B* **2005**, *72*, 195118. [\[CrossRef\]](#)
16. Abe, R.; Takeda, T.; Shiratori, R.; Shirakawa, S.; Saito, S.; Baba, T. Optimization of an H0 photonic crystal nanocavity using machine learning. *Opt. Lett.* **2020**, *45*, 319–322. [\[CrossRef\]](#)
17. Christensen, T.; Loh, C.; Picek, S.; Jakobović, D.; Jing, L.; Fisher, S.; Ceperic, V.; Joannopoulos, J.D.; Soljačić, M. Predictive and generative machine learning models for photonic crystals. *Nanophotonics* **2020**, *9*, 4183–4192. [\[CrossRef\]](#)
18. Ghasemi, F.; Aliasghary, M.; Razi, S. Magneto-sensitive photonic crystal optical filter with tunable response in 12–19 GHz; cross over from design to prediction of performance using machine learning. *Phys. Lett. A* **2021**, *401*, 127328. [\[CrossRef\]](#)
19. Chugh, S.; Gulistan, A.; Ghosh, S.; Rahman, B. Machine learning approach for computing optical properties of a photonic crystal fiber. *Opt. Express* **2019**, *27*, 36414–36425. [\[CrossRef\]](#)
20. da Silva Ferreira, A.; Malheiros-Silveira, G.N.; Hernández-Figueroa, H.E. Computing optical properties of photonic crystals by using multilayer perceptron and extreme learning machine. *J. Light. Technol.* **2018**, *36*, 4066–4073. [\[CrossRef\]](#)
21. Khan, F.N.; Fan, Q.; Lu, C.; Lau, A.P.T. An optical communication's perspective on machine learning and its applications. *J. Light. Technol.* **2019**, *37*, 493–516. [\[CrossRef\]](#)
22. Chugh, S.; Ghosh, S.; Gulistan, A.; Rahman, B. Machine learning regression approach to the nanophotonic waveguide analyses. *J. Light. Technol.* **2019**, *37*, 6080–6089. [\[CrossRef\]](#)
23. Koza, J.R.; Bennett, F.H.; Andre, D.; Keane, M.A. Automated design of both the topology and sizing of analog electrical circuits using genetic programming. In *Artificial Intelligence in Design '96*; Springer: Berlin/Heidelberg, Germany, 1996; pp. 151–170.

24. Hu, J.; Niu, H.; Carrasco, J.; Lennox, B.; Arvin, F. Voronoi-based multi-robot autonomous exploration in unknown environments via deep reinforcement learning. *IEEE Trans. Veh. Technol.* **2020**, *69*, 14413–14423. [[CrossRef](#)]
25. Bishop, C.M. Pattern recognition. *Mach. Learn.* **2006**, *128*.
26. Le Roux, N.; Bengio, Y.; Fitzgibbon, A. 15—Improving first and second-order methods by modeling uncertainty. In *Optimization for Machine Learning*; MIT Press: Cambridge, MA, USA, 2011; p. 403.
27. Kutner, M.H.; Nachtsheim, C.J.; Neter, J.; Li, W. *Applied Linear Statistical Models*; McGraw-Hill: New York, NY, USA, 2005.
28. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
29. Mooi, E.; Sarstedt, M.; Mooi-Reci, I. Regression analysis. In *Market Research*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 215–263.
30. Çankaya, S.; Eker, S.; Abacı, S.H. Comparison of Least Squares, Ridge Regression and Principal Component Approaches in the Presence of Multicollinearity in Regression Analysis. *Turk. J. Agric.-Food Sci. Technol.* **2019**, *7*, 1166–1172. [[CrossRef](#)]
31. Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B* **1996**, *58*, 267–288. [[CrossRef](#)]
32. Friedman, J.; Hastie, T.; Tibshirani, R. Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* **2010**, *33*, 1. [[CrossRef](#)]
33. Efron, B.; Hastie, T.; Johnstone, I.; Tibshirani, R. Least angle regression. *Ann. Stat.* **2004**, *32*, 407–499. [[CrossRef](#)]
34. Zou, H. The Adaptive Lasso and Its Oracle Properties. *J. Am. Stat. Assoc.* **2006**, *101*, 1418–1429. [[CrossRef](#)]
35. Yang, Y.; Yang, Y. Hybrid prediction method for wind speed combining ensemble empirical mode decomposition and bayesian ridge regression. *IEEE Access* **2020**, *8*, 71206–71218. [[CrossRef](#)]
36. MacKay, D.J. Bayesian interpolation. *Neural Comput.* **1992**, *4*, 415–447. [[CrossRef](#)]
37. Islam, M.S.; Sultana, J.; Ahmed, K.; Islam, M.R.; Dinovitser, A.; Ng, B.W.H.; Abbott, D. A novel approach for spectroscopic chemical identification using photonic crystal fiber in the terahertz regime. *IEEE Sens. J.* **2017**, *18*, 575–582. [[CrossRef](#)]
38. Ahmed, K.; Ahmed, F.; Roy, S.; Paul, B.K.; Aktar, M.N.; Vigneswaran, D.; Islam, M.S. Refractive index-based blood components sensing in terahertz spectrum. *IEEE Sens. J.* **2019**, *19*, 3368–3375. [[CrossRef](#)]
39. Jabin, M.A.; Ahmed, K.; Rana, M.J.; Paul, B.K.; Islam, M.; Vigneswaran, D.; Uddin, M.S. Surface plasmon resonance based titanium coated biosensor for cancer cell detection. *IEEE Photonics J.* **2019**, *11*, 1–10. [[CrossRef](#)]
40. Mitu, S.A.; Ahmed, K.; Al Zahrani, F.A.; Grover, A.; Rajan, M.S.M.; Moni, M.A. Development and analysis of surface plasmon resonance based refractive index sensor for pregnancy testing. *Opt. Lasers Eng.* **2021**, *140*, 106551. [[CrossRef](#)]
41. Ahmed, K.; AlZain, M.A.; Abdullah, H.; Luo, Y.; Vigneswaran, D.; Faragallah, O.S.; Eid, M.; Rashed, A.N.Z. Highly sensitive twin resonance coupling refractive index sensor based on gold-and MgF2-coated nano metal films. *Biosensors* **2021**, *11*, 104. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.