



Machine learning-based classification for predicting the power flow of surface plasmon polaritons in nanoplasmonic coupler

Zahraa S. Khaleel¹ · A. Mudhafer²

Received: 25 January 2023 / Accepted: 29 May 2023 / Published online: 16 June 2023
© The Optical Society of Japan 2023

Abstract

Using a combination of the finite element method (FEM) applied in COMSOL Multiphysics and the machine learning (ML)-based classification models, a computational tool has been developed to predict the appropriate amount of power flow in a plasmonic structure. As a plasmonic coupler, a proposed structure formed of an annular configuration with teeth-shaped internal corrugations and a center nanowire is presented. The following representative data mining techniques: standalone J48 decision tree, support vector machine (SVM), Hoeffding tree, and Naïve Bayes are systematically used. First, a FEM is used to obtain power flow data by taking into consideration a geometrical dimensions, involving a nanowire radius, tooth profile, and nanoslit width. Then, we use them as inputs to learn about machine how to predicate the appropriate power flow without needing FEM of COMSOL, this will reduce financial consumption, time and effort. Therefore, we will determine the optimum approach for predicting the power flow of the proposed structure in this work based on the confusion matrix. It is envisaged that these predictions' results will be important for future optoelectronic devices for extraordinary optical transmission (EOT).

Keywords Surface plasmon polaritons · Plasmonic coupler · Power flow · Machine learning · FEM

1 Introduction

The optical excitation of surface plasmon polaritons (SPPs) in metallic nanostructures is of fundamental interest and is essential for various applications such as optical data storage, optical lithography, and hyperlens [1–3]. This is mainly attributable to the low-dimensional nature of SPPs which represent the coupled modes of an electromagnetic waves and free charges on a surface metallic nanostructures. In addition, their capability to maintain the field confined tightly at short wavelength and not constrained by optical diffraction limit [4–9]. In particular, the emphasis of research area is which the comparatively limited light–SPP coupling is improved, which leads to more beneficial applications. Since then, numerous interesting topics have arisen that are related to plasmonic structures,

involving circular aperture arrays [10, 11], array of nano-holes [12, 13], annular aperture arrays [14, 15], and single aperture [16, 17], have been studied to investigate the SPPs excitation. In addition, increasing the effectiveness of this phenomenon necessitates the use of various geometric structures for intense and highly localized SPPs around the nanoslit [18, 19]. The theoretical direction on SPPs excitation is critical for studies into the practical application of plasmonic nanostructures. The analytical solution of SPPs excitation phenomena on plasmonic nanostructures of simple geometries can be obtained readily. However, the lack of an accurate solution for configurations with irregular geometries and anisotropic plasmonic characteristics restricts the analytical method's widespread application. To overcome this issue, a variety of numerical analysis techniques such as finite element method (FEM), finite-difference time-domain (FDTD), and discrete dipole approximation (DDA) have been developed to determine the SPPs excitation phenomenon. These methods allow for accurate optical behavior harvesting from plasmonic nanostructures of any configuration, but usually only after extensive calculation operations. The obvious conflict between numerical simulation's accuracy and rapidity

✉ A. Mudhafer
a.mudhafer@stu.edu.iq

¹ Open Educational College, Basra, Iraq

² Chemical and Petrochemical Techniques Engineering Department, Southern Technical University, Basra, Iraq

restricts its application in disciplines requiring assistance from numerous theoretical models. With the advent of free database creation, ML, as a subset of artificial intelligent methods has been extensively employed in the progress of novel materials. The common idea of ML methods is finding patterns in an input data set by applying computational algorithms. Currently, considered approaches which exploit the ML in investigating the properties and phenomena related with SPPs. Joshua Baxter et al. [20] apply ML to predict of new colors using the simulation and experimental data. They predict colors from laser parameters by leveraging the data set linking laser parameters to real, physical colors. also predict colors from nanoparticle geometries using the data set linking the simulation geometries to computed colors. Jing He et al. [21] utilized the machine learning method, to establish a mapping between the far-field spectra/near-field distribution and dimensional parameters of three types of plasmonic nanoparticles including nanospheres, nanorods, and dimmers. Arzola-Flores and Gonzalez [22] propose an investigation using ML tools, for predicting the wavelength of the dipole Surface plasmon resonance (SPR) of Concave Gold Nanocubes of different sizes and depths of the concavities dispersed in water. Emilio Corcione et al [23] study the applicability of ML methods of regression for sensor calibration and explore the limitations of surface-enhanced infrared absorption glucose sensing from the measured reflectance spectra for plasmonic nanoantenna glucose sensing. Among the previously mentioned plasmonic structures, the modeling with the corrugation on a surface is of high importance, as it allows light-surface coupling, which provides a light collection system that has a high level of efficiency and light steering and has given them a high importance for certain applications, for instance, high-power applications [24], THz optical sensing [25], nonlinear optical processes [26], photochemistry [27], highly integrated plasmonic circuits [28], and biophotonics [29]. Even though plasmonic features of nanostructures have garnered a great deal of attention in the literature, to the best of our knowledge, limited studies have been carried out on the crucial topics of power flow which represent the signal features of SPPs in real applications. In this paper, we present a study using ML-based classification methods and FEM of COMSOL Multiphysics, to predict the time-averaged power flow in a nanoplasmonic coupler. This proposed construction is an annular design with teeth-shaped interior corrugations and a nanowire at its center. The structure is distinctive because it combines the features of annular and corrugated designs. Firstly, by modifying the tooth profile, such as the teeth number, teeth height, teeth obliquity angle, and the teeth rotation angle, the interactions with the structure's interior walls enable

it to capture extra energy. Second, nanowire's illuminating mechanism precisely converts light to SPPs.

2 Simulation structure and data analysis methods

A schematic illustration of the simulated configuration is depicted in Fig. 1 [30]. The design consists of annular construction with nanoslit positioned at its top and a nanowire positioned in structure's center. The teeth produce grooves along the structure's internal sidewalls by cutting the inner wall.

A TM polarization plane wave illumination with an incidence wavelength of $\lambda = 500$ nm around a nanowire. Under this illumination, SPP modes can excite and propagate surround the nanowire. Corrugations primarily backscatter excited SPPs, which are subsequently confined and focused in nanogrooves and nanoslit. SPPs that propagated within the structure would interact with the accumulated SPPs. The following parameters indicate the geometrical dimensions: outer, inner, and root radii of the structure are R_0 , R_i , and R_t , respectively, and nanowire radius is denoted by R_w . Here, $R_0 = 6.21 \mu\text{m}$, $R_i = 5.07 \mu\text{m}$, and $R_t = 5.85 \mu\text{m}$ are remain unchanged. The nanoslit width is denoted by W_s , where $W_s < \lambda$, this leads to the nanoslit behaving as a diffractive element, and its near-field effects become significant. The tooth profile is described as outlined below: n , H_t , and W_t , respectively, stand for the number, height, and width of teeth. α denotes the oblique angle of teeth, while θ denotes the rotation angle of teeth. The metallic material in the design is silver (Ag), It's a popular material for the design of plasmonic devices, it has a relatively low absorption coefficient in the visible spectrum, which means that it can efficiently scatter and confine light in the near-field without significantly attenuating it. The interior dielectric material is simulated as a homogeneous SiO_2 with a refractive index of $n = 1.464$. The wavelength-dependent permittivity and the material parameters are specified by experimental data of Johnson Christy [31], presented in COMSOL. The data set was obtained through a two-dimensional numerical implementation of the COMSOL Multiphysics, based on the FEM [30], where the designer try with a set of different data at each time for a set of parameters that have a basic effect on obtaining the appropriate the time-averaged power flow (P_{avg}). The data set includes 150 record samples, each of them has eight attributes and one class with two possible yes or no. The data set description, attributes, statistical analysis, and their values are shown in Table 1 and the visualization of all attributes is shown in Fig. 2.

The following representative data mining techniques: standalone J48 decision tree, support vector machine (SVM), Hoeffding Tree, and Naïve Bayes are systematically used

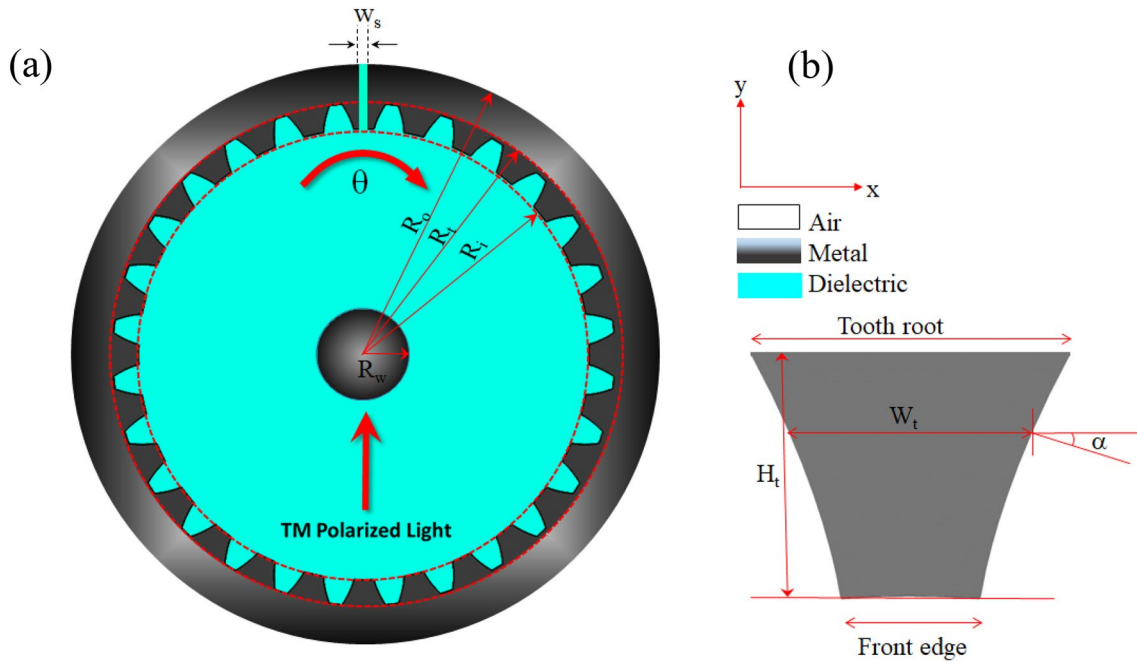


Fig. 1 a Schematic configuration of design, b The tooth cross-section [30]

Table 1 The visualization of data set

S.NO	Attributes	Attributes type	Range of attribute	Mean	Standard deviation of attribute
1	R_w	Numeric	50–500 (nm)	275	144.095
2	N_t	Numeric	20–38	29.333	5.754
3	H_t	Numeric	213–405(nm)	287.267	59.094
4	α	Numeric	0.35 (rad)	0.349	0
5	θ	Numeric	90–180 (deg.)	120.6	42.777
6	W_s	Numeric	70 (nm)	70	0
7	P_{avg}	Numeric	0.25–4732.44 (W/m ²)	513.2	875.445
8	Class	Class label	Yes, no		

to perform the prediction of the P_{avg} in the plasmonic coupler. The C4.5 algorithm is a classification model, which results a decision tree founded on the information entropy theory [32]. It is an improvement of Ross Quinlan’s earlier ID3 algorithm also described in Waikato Environment for Knowledge Analysis (Weka), as J48, J standing for Java. C4.5 is referred to as a statistical classifier since it produces decision trees that are used for classification. The J48 for C4.5 algorithm implementation includes additional features such as treating missing values, pruning the decision tree, continuous attribute value ranges, and derivation of rules. This algorithm builds a decision tree by depending a training data set such as the ID3 algorithm, utilizing the information entropy concept. The information entropy can be calculated as,

$$Entropy = - \sum_{i=0}^{c-1} p(i|t) \log_2 p(i|t) \tag{1}$$

where $p(i|t)$, denotes the records fraction that belongs to class i at a specified node t , while c denoted the classes number. When entropy is a measure of the randomness or unpredictability of a system the variance between the entropy of the data set before and after splitting on a particular feature is known as information gain which calculated as,

$$Gini(t) = 1 - \sum_{i=0}^{c-1} [p(i|t)]^2 \tag{2}$$

SVM is a well-known algorithm for supervised learning. [33], which is employed for regression and

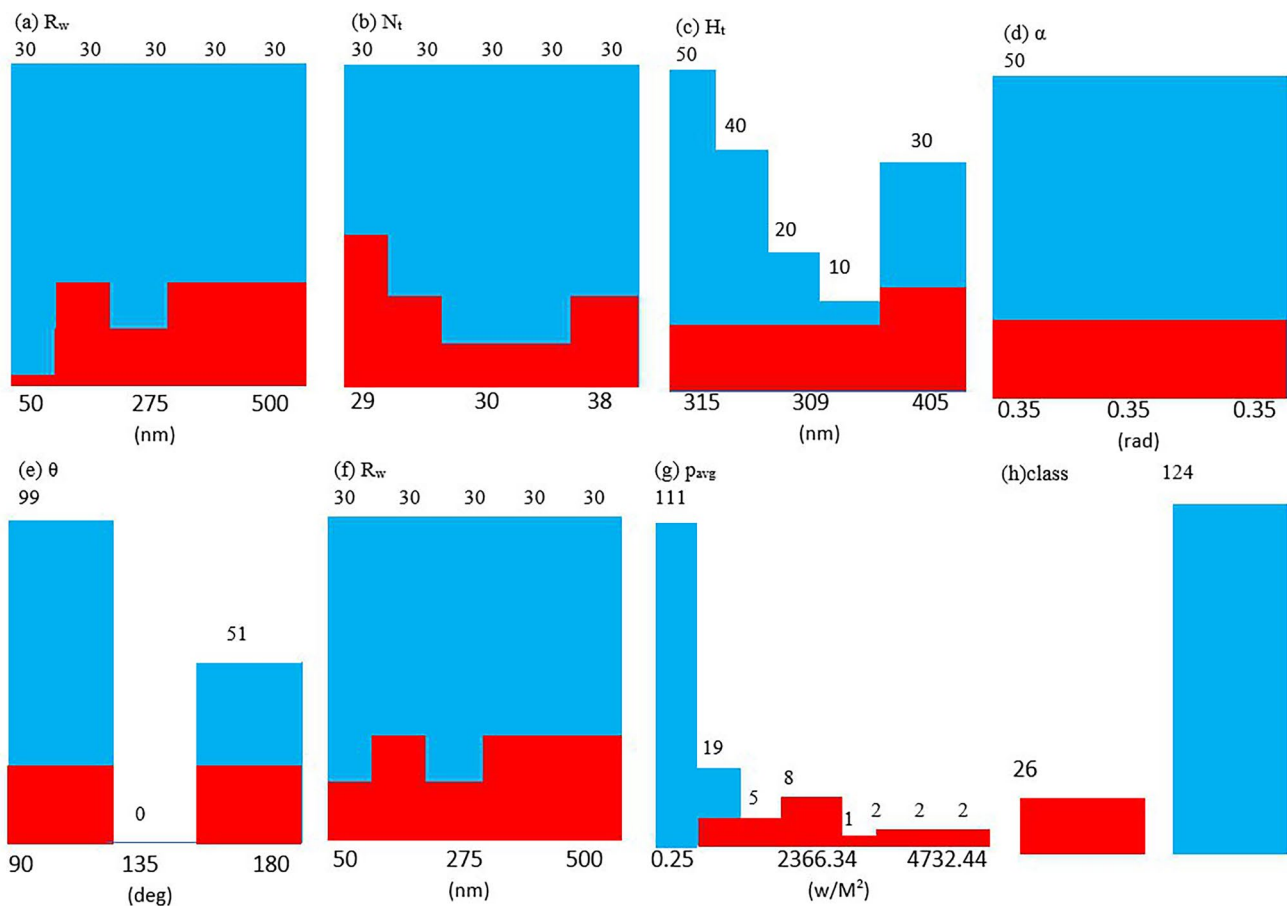


Fig. 2 The visualization of all attributes

classification problems. However, it is applied to classification problems in ML. The objective of the SVM algorithm is finding the hyperplane that best separates different classes in a given data set. In a binary classification task, the hyperplane is the decision boundary that separates the two classes, maximizing the margin between the two closest data points from different classes. We can consider This accurate decision as a hyperplane. SVM selects the maximum points and vectors contribute to the hyperplane formation. These cases are known as support vectors, and the corresponding algorithm is known as SVM. Considering Fig. 3, which depicts the classification of two distinct groups using a decision hyperplane or boundary.

In a data stream environment where it is impractical to keep all data, the most challenging aspect of constructing a decision tree is reusing instances to estimate the appropriate splitting attributes. Instead of reusing instances, Domingos and Hulten [34] presented the Hoeffding Tree, a decision tree algorithm for data streams that waits for more instances to emerge. The most intriguing aspect of

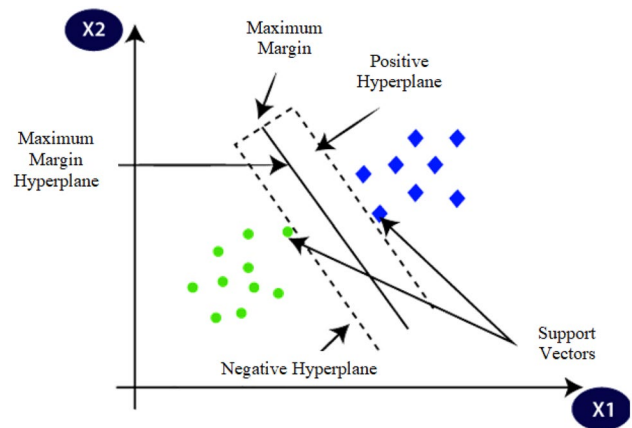


Fig. 3 Decision boundary and margin of SVM [34]

the Hoeffding Tree is that it constructs a tree that evidently converges to the tree constructed by a batch learner with sufficient data. The pseudo code of the Hoeffding

$$\epsilon = \sqrt{\frac{R^2 \ln 1/\delta}{2n}} \quad (3)$$

where R represents the range of random variable, δ is refers to the probability of the estimation *not* being within ϵ of its an expected value, and n is the number collection of examples at the node. In information gain, the entropy range is $[0, \dots, \log n_c]$ for n_c refers to class values. Even though use of Hoeffding's bound in this case is erroneous, as previously shown, most applications continue to use it. Reasonable outcomes may be attributable to an overestimation of the genuine probability of error in the majority of circumstances. A Hoeffding Tree, from massive data streams a decision tree induction algorithm is able for learning in any time, considering that the distribution generating examples does not change over time. A theoretically enticing aspect of Hoeffding trees that is not provided by other incremental decision tree learners is its solid performance guaranteed. By Using Hoeffding bound, it is possible to demonstrate that its output is approximately identical to that of a non-incremental learner given an infinite number of examples. Given the class label y and the assumption that the attributes are conditionally independent, a naive Bayes classifier [35] determines the class-conditional probability. The following is a formal description of the conditional independence assumption:

$$P(X|Y = y) = \prod_{i=1}^d P(X_i|Y = y) \quad (4)$$

where each attribute set $X = \{X_1, X_2, \dots, X_d\}$ consists of d attributes. With the conditional independence assumption, we just need to determine the conditional probability of each X_i , given Y , rather than estimating the class-conditional probability for each combination of X . The latter method is more useful since it can produce a reliable probability estimate using a smaller training set. The naive Bayes classifier determines the posterior probability for each class Y in order to classify a test record:

$$P(X|Y) = \frac{P(Y) \prod_{i=1}^d P(X_i|Y = y)}{P(X)} \quad (5)$$

Since $P(X)$ is fixed for every Y , it is sufficient to choose the class that maximizes the numerator term, $P(Y) \prod_{i=1}^d P(X_i|Y)$. In the next two subsections, we describe several approaches for estimating the conditional probabilities $P(X_i|Y)$ for categorical and continuous attributes.

3 Results and discussion

The evaluation of ML model is important as building it, because we work on new or previously unseen data, so the evaluation should be thorough and versatile to create a reliable and tough model. Here, we are depending on some concepts (Accuracy, sensitivity, Specificity, ROC Area, and Error Rate) to evaluate. After we obtained a data from COMSOL, load it to Weka, applied cross-validation with tenfold data and used four classification techniques to predicate if the amount of parameters (attributes) which the designer chose is leading to the appropriate power flow or not. Accuracy is one metric, accuracy and the fraction that predict the model got right as formally, accuracy can be define as:

$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}} \quad (6)$$

Binary classification in terms of positives and negatives can be calculated as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (7)$$

where, TP is true positive that refers to observations total number to the positive class which have been correctly predicted, FP is number of false positives represent the observations total number which have been predicted to belong to the positive class, but , actually instead, belong to the negative class, TN is true negative, number of observations that belong to the negative class and have been predicted correctly. and FN is false negative, is observations total number of that have been predicted as negative class but it is belongs to the positive class. Sensitivity is the metric that, evaluates ability of classification model to predict true positives of each available category and can be obtained as :

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (8)$$

Specificity is the evaluation of models capability to predict a true negatives of all available categories. These metrics could be applied on any categorical model

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (9)$$

The area under the receiver operator characteristic curve (ROC) is metric used to evaluate the problems of binary classification, it is a probability curve plots TPR opposed to FPR at several threshold values. The area under the curve (AUC) is the measurement of the capability of a classifier to identify among the classes and it is used as summary of the ROC curve. Error rate is inversely proportional to accuracy (1-accuracy).

$$ErrorRate = \frac{FP + FN}{P + N} \tag{10}$$

Figure 4 shows the comparison between the classification methods which used in this predication model, for J4.8 has the highest accuracy with value 99.3333% that indicate the models is making large number of correct predication, while has the lowest error rate equals 0.6667% indicates that the model is making a small number of incorrect predication, for Hoeffding and Naïve Bayes have the high accuracy 96.6667% and 97.3333% and low error rate 2.6667% and 3.3333%, For SVM it considered a good accuracies (92.6667%) and acceptable error rate (7.3333%) but it is not the best to make the predication for this model.

In Fig. 5, if the sensitivity equals one, it means that a change in the input variable will result in an equivalent change in the output variable. In other words, the output variable is highly responsive to changes in the input variable, Here the sensitivity of J4.8 equals one that, means all required power flow level is correctly identify as high while, Naïve Bayes and Hoeffding get nearly > 0.96 thus, they considered as efficiency method and indicates a strong and semi direct relationship between the input and output

variables, SVM has 0.577 sensitivity that means just 50% of power flow identify as high which, indicates a moderate relationship between the input and output variables, with the output variable responding to changes in the input variable, but not as strongly as J4.8, Naïve Bayes and Hoeffding. All about the specificity is a measure of how well a test can correctly identify negative, SVM and J4.8 the value of specificity equals one where succeeded to find all true negative in data set and did not produce any false positives. In other words, the test has perfect discrimination between positive and negative cases. However, it is important to note that a test with high specificity does not necessarily mean it is a good test overall, as other factors such as sensitivity and accuracy also need to be considered, Hoeffding and Naïve Bayes > 0.965 so their performance is satisfactory to identify the low power in data set.

It is clear from Fig. 6 J4.8, Naïve Bayes and Hoeffding produced the highest AUC equal 0.996 hence, the performance of these classification models at distinguishing between the positive and negative classes is outstanding which means these classifiers have perfect discrimination ability. AUC of SVM > 0.5 the ability of this model to characterize the positive and negative classes is very good but lower than other classifiers.

Fig. 4 Comparison classification models depending on accuracy and error rate

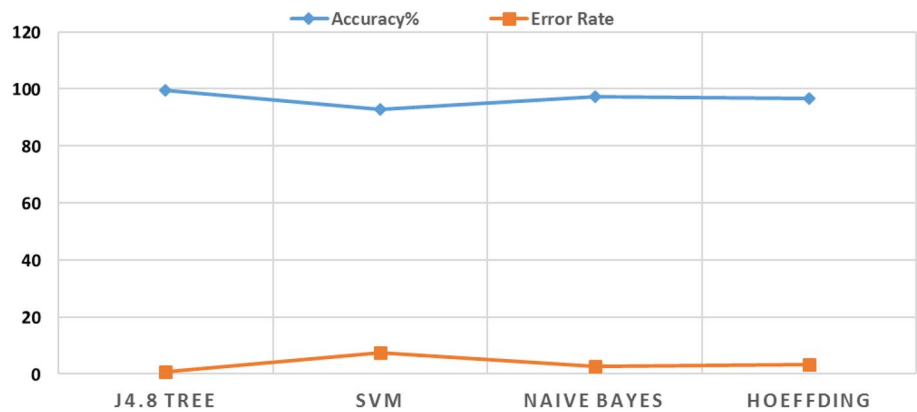
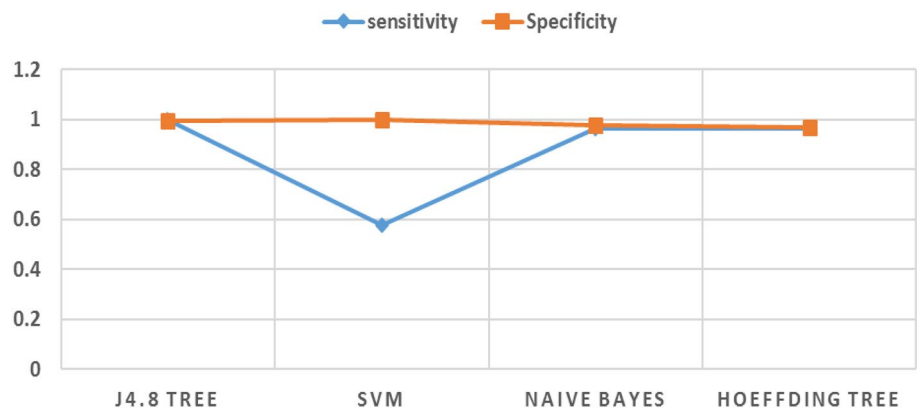


Fig. 5 Comparison classification models depending on sensitivity and specificity



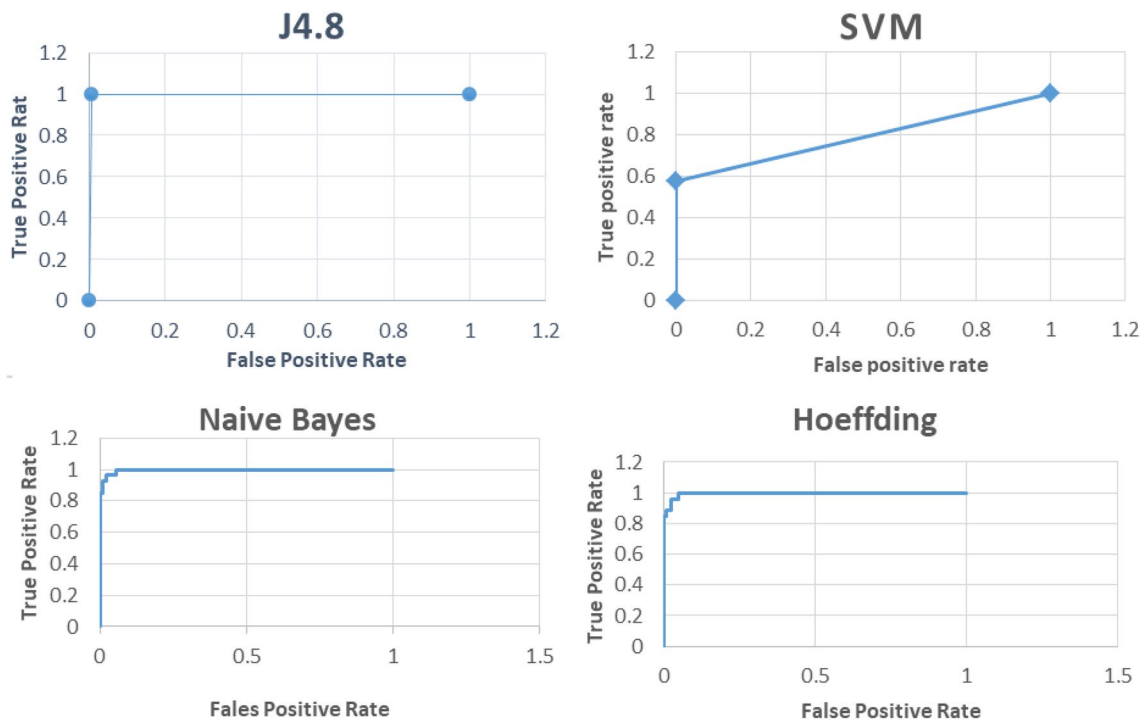


Fig. 6 AUC of j48=0.996, SVM=0.788, Naïve Bayes=0.996 and Hoeffding=0.996 obtained with ROC

4 Conclusions

Numerical methods are commonly performed for finding the value of P_{avg} that consume enormous computational resources, time and tedious procedures. We have determined a goal of predicting the P_{avg} at nanoslit in a nanoplasmonic coupler. To achieve this the FEM of COMSOL is used to acquire the data set of P_{avg} considering the geometrical parameters of the plasmonic coupler. Thus, the same set of variables for cross validation (tenfold) was used with the following four different ML algorithms: standalone J48 decision tree, SVM, Hoeffding Tree, and Naïve Bayes. Although none of these algorithms have the same interior architecture, we found after doing comparison between them, perfect accuracy in the prediction models is possible as the coefficient of standalone J48 decision tree is 0.9933, the possibility of error is almost non-existent and the capability to find all high P_{avg} is perfect because sensitivity equal one. Consequently, we have a good instance of the prediction based classification models with j48 algorithm. In fact, ML could be applied to a wide range of issues in materials science and engineering, including the design and optimization of new materials, the prediction of material properties, and the analysis of complex systems such as metamaterials [36–38]. ML algorithms can be used to analyze large data sets of material properties and performance metrics to identify patterns and relationships that can be used to guide the design and optimization of these materials. For example,

ML based classification algorithms can be used to classify different types of hyperbolic metamaterials based on their optical properties or critical coupling behavior. By training a machine learning model on a data set of material properties and performance metrics, the model can learn to accurately predict the behavior of new materials and guide the design of more efficient and effective hyperbolic metamaterials. Overall, ML has the potential to revolutionize the field of materials science and engineering, enabling researchers to discover new materials and optimize their performance with unprecedented speed and accuracy.

References

1. Song, M., Wang, D., Kudyshev, Z.A., Xuan, Y., Wang, Z., Boltasheva, A., Shalaev, V.M., Kildishev, A.V.: *Laser Photonics Rev.* **15**, 3 (2021)
2. Dhawan, A.R., Nasilowski, M., Wang, Z., Dubertret, B., Maitre, A.: *Adv. Mater.* **34**, 11 (2022)
3. Su, D., Zhang, X.Y., Ma, Y.L., Shan, F., Wu, J.Y., Fu, X.C., Zhang, L.J., Yuan, K.Q., Zhang, T.: *IEEE Photonics J.* **11**, 10 (2017)
4. Maier, S.A.: *Plasmonics fundamentals and applications*. In: Maier, S.A. (ed.) Springer Science Business Media. Springer, US, New York (2007)
5. Barnes, W.L., Dereux, A., Ebbesen, T.W.: *Nature* **424**, 6950 (2003)
6. Zayats, A.V., Smolyaninov, I.I., Maradudin, A.A.: *Phys. Rep.* **408**, 3–4 (2005)
7. Novotny, L., Van Hulst, N.: *Nat. Photonics* **5**, 2 (2011)

8. Ding, L., Wang, L.: *Appl. Phys. A* **119**, 3 (2015)
9. Aparna, U., Mruthyunjaya, H.S., Sathish, K.M.: *J. Opt.* **49**, 1 (2020)
10. Ebbesen, T.W., Lezec, H.J., Ghaemi, H.F., Thio, T., Wolff, P.A.: *Nature* **391**, 6668 (1998)
11. Beruete, M., Sorolla, M., Campillo, I.: *Opt. Express* **14**, 12 (2006)
12. Sturman, B., Podivilov, E.: *Gorkunov. phys rev B* **77**(7), 075106 (2008)
13. Wang, Q., Wang, L.: *Lab-on-fiber Nanoscale*. **12**, 14 (2020)
14. Duan, Y., Yin, S., Gao, H., Zhou, C., Wei, Q., Du, C.: *Opt. Laser Technol.* **44**, 1 (2012)
15. Liu, Y.J., Si, G.Y., Leong, E.S., Xiang, N., Danner, A.J., Teng, J.H.: *Adv. Mater.* **19**, 23 (2012)
16. Daniel, S., Saastamoinen, K., Saastamoinen, T., Rahomäki, J., Friberg, A.T., Visser, T.D.: *Opt. Express* **24**, 17 (2015)
17. Poujet, Y., Salvi, J.: and fadi issam baida. *Opt. Lett.* **32**, 20 (2007)
18. Wang, C., Bai, M., Jin, M., Zhang, Y.: *Optik* **123**, 20 (2021)
19. Xie, Y., Zakharian, A.R., Moloney, J.V., Mansuripur, M.: *Opt. Express* **13**, 25 (2004)
20. Corcione, E., Pfezer, D., Hentschel, M., Giessen, H., Tarín, C.: *Sensors* **21**, 1 (2021)
21. Valdivia-Valero, F.J., Nieto-Vesperinas, M.: *J. Nanophotonics* **5**, 1 (2011)
22. Valdivia-Valero, F.J.: and manuel nieto-vesperinas. *Opt. Express* **18**, 7 (2010)
23. Baxter, J., Calà Lesina, A., Guay, J.M., Weck, A., Berini, P., Ramunno, L.: *Sci. Rep.* **9**, 1 (2019)
24. He, J., He, C., Zheng, C., Wang, Q., Ye, J.: *Nanoscale* **11**, 37 (2019)
25. Arzola-Flores, J.A., González, A.L.: *J Phys Chem C* **124**, 46 (2020)
26. Genevet, P., Tétienne, J.P., Gatzogiannis, E., Blanchard, R., Kats, M.A., Scully, M.O., Capasso, F.: *Nano Lett.* **8**, 12 (2010)
27. Alam, M.Z., Yang, Z., Sheik-Bahae, M., Aitchison, J.S., Mojahedi, M.: *Sci. Rep.* **11**, 1 (2021)
28. Tu, Q., Liu, J., Ke, S., Wang, B., Lu, P.: *Plasmonics* **15**, 3 (2019)
29. Mishra, M., Sharma, M., Gupta, P.: *Physica E* **1**, 130 (2021)
30. Ioannidis, T., Gric, T., Rafailov, E.: *Opt. Quant. Electron.* **52**, 1 (2020)
31. Khoshdel, V., Shokooh-Saremi, M.: *JOSA B* **38**, 5 (2021)
32. Kehn, M.N., Mou., and Jen Yung Li. *IEEE Trans. Microw. Theory Tech.* **66**, 4 (2018)
33. Mudhafer, A., Zahraa, S.: Khaleel, and Ra'ed Malallah. *J. Comput. Electron.* **20**, 5 (2021)
34. Singh, L., Maccaferri, N., Garoli, D., Gorodetski, Y.: *Nanomaterials* **11**, 5 (2021)
35. Tan, Pang-Ning, Michael Steinbach, and Vipin Kumar. (2016) *Introduction to data mining India: Pearson Education* (ed.) 231
36. Xiang, Y., Dai, X., Guo, J., Zhang, H., Wen, S., Tang, D.: *Sci. Rep.* **4**, 1 (2016)
37. Lv, T.T., et al.: *Sci. Rep.* **6**, 1 (2016)
38. Jinhui, S., et al.: *J Mater Chem* **6**, 6 (2018)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.