Nirav Joshi
Vinod Kushvaha
Priyanka Madhushri *Editors*

Machine Learning for Advanced Functional Materials

Springer

Machine Learning for Advanced Functional Materials

Nirav Joshi · Vinod Kushvaha · Priyanka Madhushri
Editors

Machine Learning for Advanced Functional Materials

Springer

Editors
Nirav Joshi
University of Sao Paulo, Sao Carlos
Institution of Physics, Grupo de Polimeros
São Paulo, Brazil

Priyanka Madhushri
Proof of Concept and Innovation Group
Stanley Black and Decker (United States)
Atlanta, GA, USA

Vinod Kushvaha
Department of Civil Engineering Materials and Structures
Indian Institute of Technology Jammu
Jammu, India

ISBN 978-981-99-0392-4 ISBN 978-981-99-0393-1 (eBook)
https://doi.org/10.1007/978-981-99-0393-1

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

The evolving field of Machine Learning (ML) and Artificial Intelligence (AI) contributed tremendously to the advancement of various branches of science and technology. Machine learning has attracted great interest from the research community in material science, because of its ability to statistically analyze a large collection of data. Along with the computational task, time-efficient tools of machine learning have also been applied for the prediction of the design and properties of new materials. A noticeable shift from trial and error-based laboratory approach to the modeling and simulation-based software techniques in the preparation and characterization of functional materials manifests the emergence of big data in the field of material science. The efficient algorithms enable data collection, and storage with high security, fast processing, and interpretation of physically generated results. It has put the research of physical and chemical science at the forefront with the advancements in image processing, photonics, optoelectronics, and other emerging areas of material science. Although there have been a lot of research articles on this topic, no comprehensive scientific reference book on machine learning for advanced functional materials has been published so far. This book aims to provide an in-depth examination of recent achievements in material science by focusing on topical issues using machine learning methods.

In this book, there are a total 13 chapters. *Muhammad Abdul Basit* describes the ML about organic and inorganic solar cells, making a discussion about the use of machine learning, various classes of machine learning, common algorithms, and basic steps for ML. Detailed discussion about specific types of ML for solar cells, and application of ML for prediction of suitable materials, for optimization of device structure and fabrication processes, and reconstruction of measured data for solar cells. *Shirong Huang* demonstrates the application of machine learning toward gas identification focusing on two main strategies: the conventional electronic nose strategy relying on sensor arrays with different materials or functionalization providing different maximum response values with each, and the utilization of single sensors but utilizing multiple transient features of their response. Similarly, *Elsa M. Materón* discusses artificial intelligence tools utilized to assist electrochemical, optical, and gas sensors, including the advantages, advances, limitations, and

strategies of the most commonly used machine learning algorithms. *Nitesh Sureja* discusses the design, implementation, and effectiveness of the application of ML on data generated by healthcare devices. *Gisela Ibáñez-Redin* outlines a general overview of the application of machine learning algorithms to wearable technologies with special emphasis has been placed on the application of this approach to health monitoring, sports analytics, veterinary medicine, and agriculture. *Humaira Rashid Khan* demonstrates the need and necessity of developing new energy materials to contribute to global carbon neutrality with the most recent advancements in data-driven materials research and engineering reviewed, including alkaline ion battery materials, photovoltaic materials, catalytic materials, and carbon dioxide capture materials. *Purvi Bhatt* reviews applications of machine learning algorithms to study experimentally obtained results of physical systems and concludes the chapter with a discussion of future directions and challenges in the acceptability of this advanced technique to the existing vast area of material science. *Ramandeep Kaur* includes a brief discussion of perovskite materials and the developments made in the lead-free perovskite for photovoltaics. Further, this discussion turns to the collection and analysis of materials data and extends to the descriptors used to describe the performance and properties of lead-free perovskites. *Shulin Yang* discusses the models applied for machine learning and the enhanced mechanism of the novel sensors or sensor array to better understand the machine learning methods. *R. Vignesh* covers the challenges in advanced functional materials research and the role of machine learning in design, simulation, and evaluation. Also, significant pointers to successful machine learning applications are addressed, as well as the remaining hurdles in machine learning for next-generation functional materials. *Tulsi Satyavir Dabodiya* provides a brief introduction to the ML process that could benefit the photocatalysis field. Further, the chapter provides basic PC research knowledge that could potentially be useful for machine learning methods. Additionally, we also describe the pre-existing ML practices in PC for quick identification of novel photocatalysts. *G. Sudha Priyanga* explains the state of the art in materials informatics as it pertains to photocatalysts and indicates the combination of ML with photocatalysis expertise could pave the way for the creation of a comprehensive catalyst screening platform that would greatly help in making photocatalysis prominent and technologically relevant, at the same time. *V. Balasubramani* focuses on impedance-based sensors such as chemical and biosensors using graphene-based nanocomposite materials and the role of machine learning in addressing the challenges of these sensors in the detection of chemical and biomolecules, advantages, drawbacks, current improvement, and future direction are illustrated. This book will benefit the students, researchers, and working professionals working in the field of machine learning, pattern recognition, and artificial intelligence using advanced functional materials.

São Paulo, Brazil — Nirav Joshi
Jammu, India — Vinod Kushvaha
Atlanta, USA — Priyanka Madhushri

Contents

Solar Cells and Relevant Machine Learning

Muhammad Abdul Basit, Muhammad Aanish Ali, and Mamoona Yasmeen

Abstract Machine learning (ML) and data science is the most emerging computation tool which has been recently incorporated in emerging fields of materials science and engineering including but not limited to solar cells. It helps us to optimize materials and their photovoltaic performance for various types of solar cells through algorithms and models, which is easy, cost-efficient, and rapid compared to conventional programming methods. Although the family of solar cells has been classified into various types based on their generations, however, the basic two types (i.e., organic, and inorganic solar cells) are more specific owing to the contrast in their materials, fabrication techniques, and corresponding characterizations. A large number of materials can be used for developing photoanode/photocathode in solar cells; however, it is too difficult and complex to design the most proficient one practically. In this chapter, we will comprehensively review ML about organic and inorganic solar cells, making a discussion about the use of machine learning, various classes of machine learning, common algorithms, and basic steps for ML. A detailed discussion about specific types of ML for solar cells and the application of ML for the prediction of suitable materials, optimization of device structure and fabrication processes, and reconstruction of measured data for solar cells are given. In the end, we shall cover the current research status and future challenges, and expected progress of ML, and will propose suggestions that can enhance the usefulness of machine learning.

1 Introduction

Modern day world faces unprecedented challenges caused by global warming and energy crisis. Researchers and scientists continue to work for finding novel and smart solutions to address these challenges.

M. A. Basit (✉) · M. Aanish Ali · M. Yasmeen
Department of Materials Science and Engineering, Institute of Space Technology, Islamabad 44000, Pakistan
e-mail: ab_saim@hotmail.com; m.abdulbasit@mail.ist.edu.pk

N. Joshi et al. (eds.), *Machine Learning for Advanced Functional Materials*,
https://doi.org/10.1007/978-981-99-0393-1_1

There is a large gap between the supply of fossil fuels and the energy demands, and on the other hand, burning fossil fuels for energy production is a cause of global warming. Green, renewable solar energy can maintain a balance between the environment and the demands of energy [1]. Annually 23,000 TWy of energy can be harvested from solar radiation and the annual consumption of energy in the world is 1,600 TWy. Similarly, the current energy needs of the global can be met for eight years if per day of sunlight that reaches the earth could be harvested. The above facts and figures are good evidence that solar energy can meet the ever increasing energy demands and resolve environmental problems of the world [2].

1.1 Generations of Solar Cells

In 1839, Alexandre-Edmond Becquerel gave the concept of photovoltaics, and in 1940, the first silicon-based solar cell was developed in Bell lab [3]. The efficiency of silicon-based solar cells is around 47.1% to date [4]. Further discussion is about the categorization of solar cells, in three generation based on the photoanode junction's cost of manufacturing, performance, and limitations. Figure 1 presents a broad overview of major categories of solar cells and their classifications according to type [5].

1.1.1 First-Generation Solar Cells

The first generation belongs to wafer-based solar cells like mono-crystalline silicon, polycrystalline silicon, and multi-junctions. Mono-crystalline silicon-based solar cells are expensive but durable under ambient conditions, stable under high temperatures, and have a longer life span. Polycrystalline silicon solar cells have a short life span, less durability under ambient conditions, and stability compromises under elevated temperatures but are low cost. Multi-junction solar cells consist of multiple $p-n$ junctions. In these $p-n$ junctions, different semiconductors are used and can respond to a wider range of light, so more electron-hole generation occurs, and as a result, more current is produced, and the efficiency of multi-junction solar cells is increased [6].

1.1.2 Second-Generation Solar Cells

This type of solar cell is made of thin films and is cheaper than the first generation. The absorbing layer is mounted on a suitable substrate, and mostly three types of substrates are used in this generation; copper, cadmium, and silicon. Amorphous silicon solar cells are extremely low cost, but the problem is degradation under sunlight; as a result, efficiency is decreased [7]. Cadmium telluride solar cells (CdTe) are inexpensive and chemically stable and has a wide range of absorbances due to

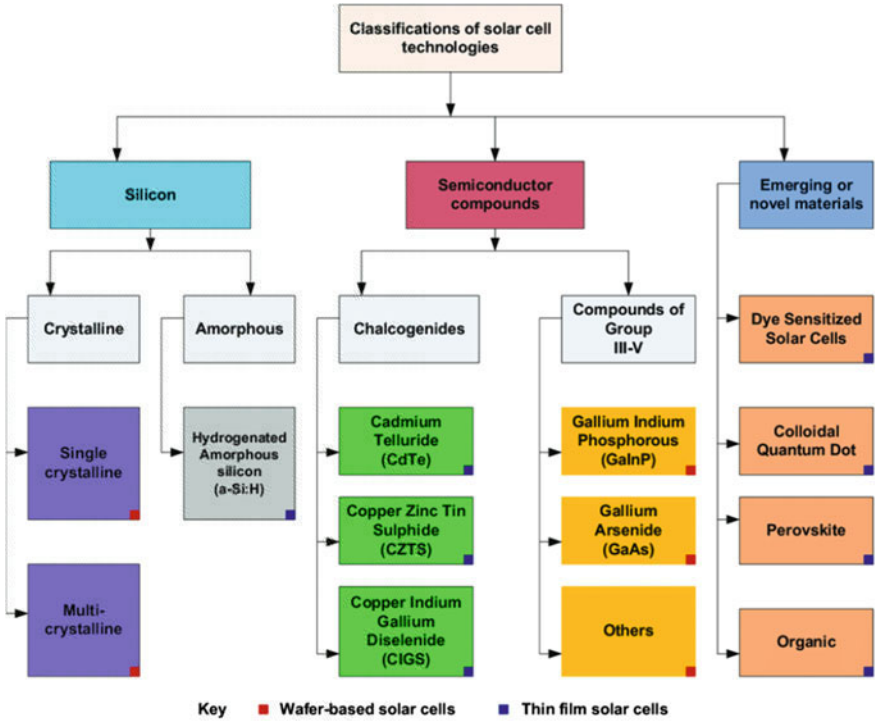

Fig. 1 Classification of solar cells. Reprinted with permission from Ref. [5]. Copyrights 2017 Elsevier

a low band gap, but cadmium is highly toxic. Copper Indium Gallium Selenide (CIGS) solar cells are less expensive, toxic, and efficient around 20%, but expensive in manufacturing [8].

1.1.3 Third-Generation Solar Cells

In 1839, calcium titanium oxide ($CaTiO_3$) perovskite was discovered by German researcher Govstave Rose. Organic cations are present in the corner, inorganic cations are present at the body-centered, and anions are positioned at the face center. In 2009, perovskites were reported for solar cells, with around 4% efficiency. 29% efficiency is reported for tandem perovskite solar cells [9]. Perovskites are cheaper and have a wide range of absorbance due to a low bang gap, but the problem with perovskite is a short life scan due to instability under ambient conditions, and high temperature. For the last few decades, Quantum Dots Sensitized Solar Cells (QDSSCs) have been a hot area of research for researchers due to their wide range of absorbance; highly stable under thermal, and moisture, as well as facile fabrication methods; and low

cost. Various photoanode sensitizers, electrolytes, and counter electrodes have been investigated to increase the efficiency of QDSSCs [10, 11].

1.2 Machine Learning

Machine learning is a subcategory of Artificial Intelligence (AI) that utilizes knowledge of computer science, mathematics, and statistics to build algorithms for a specific purpose. Now algorithms and tools of ML are providing new devise ways to investigate the chemical and physical properties of the material and their corresponding efficiencies. ML allows us to investigate hidden properties of the material. Generally, ML is categorized into three groups namely Supervised Learning (SL), Unsupervised Learning (UL), and Reinforcement Learning (RL) as shown in Fig. 2. The basic difference between them is input and output data. In SL, labeled data is used as input and output is known. In UL unlabeled data is used as input, and output is discovered after understanding the pattern. RL is semi-supervised learning in which learning data is not predefined and work on trial-and-error method. The performance of the material is significantly dependent on the selection of the ML algorithm. For the analysis of ML, the first step is data preparation. Data can be collected from computational simulations or experimental measurements with at least 50 data points necessary for a reasonable ML model. In the second step, data is perfectly cleared to avoid inconsistency and normalize for better interpretation. The third is the proper selection of an algorithm for analysis. Clustering, classification, regression, and probability estimates are widely used in material science. The last step is model evolution; for this, purpose root means squared error (RMSE) and coefficient of determination (R2) are used. Its values range is 0–1 and close to 1 indicates accuracy [12].

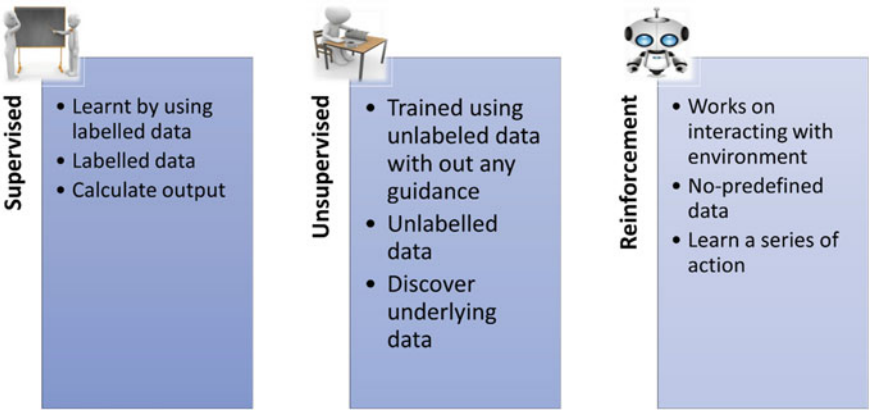

Fig. 2 Various types of machine learning relevant to solar cells

In the first section of this chapter, we have tried to reveal the importance of ML to predict the PCE of solar cells by exploiting the properties of materials and optimal device structure. In the second section of this chapter, the basic workflow of ML in material science and solar cells is elaborated. In the third section, we have briefly described the ML algorithms which can be used to predict the band gap of materials, and the role of acceptor and donor (D/A) for the performance of the solar cells. In the fourth section, we have explored the role of ML algorithms to predict highly efficient materials and have the potential to discover new efficient materials to improve the performance of solar cells. ML algorithms also helped to optimize the device structure of solar cells. This work provides practical guidance for the acceleration of the discovery of new efficient materials to improve the performance of solar cells.

2 Workflow of Machine Learning

The main purpose of ML is to explore hidden patterns of data and help predict new data patterns. For this purpose, ML adopts a systematic workflow like data collection and its preparation, model building, and its evaluation. Here the discussion is about a few steps of machine learning regarding solar cells.

2.1 Data Collection and Preparation

The first and most crucial step in the machine learning process is data collection. The type or amount of data collected depends on algorithms. For SL labeled input data is required but for UL labeled input data is not compulsory. The amount of data also varies from model to model, e.g., ANN required a large amount of data but in general, almost 50 points are necessary for the response of ML. The performance of the model also depends on the splitting of data methods, and the original dataset can be split into training and testing subsets and can be split into ratios. ML gives knowledge of the hidden patterns of given data and helps to predict the patterns of new data. For efficient performance of model data, cleaning and transformation are performed to avoid noise and redundant information. Properties of various descriptors fall on different scales, like structural which can be controlled by synthetic rules. If the descriptor features are higher than the observation features, then we used dimensional reduction tools like principle component analysis (PCA), and independent component analysis (ICA) [13].

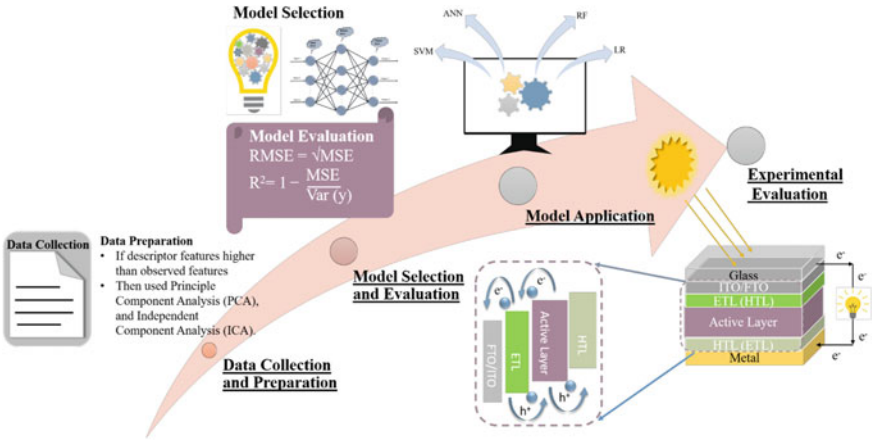

Fig. 3 The general workflow of ML for solar cells includes data collection and feature engineering, model selection and evaluation, model application, and experimental evaluation

2.2 Model Building and Evaluation

Every algorithm has its pros and cons. Clustering, regression, and classification are mostly used for material properties prediction. Probability estimation is used for the discovery of new material. Common conditions for a good model are to perform well for both training and testing data. To check the performance of the model, different metrics are used; metrics vary from algorithm to algorithm, and mostly root mean squared error (RMSE) and coefficient of determination (R^2) metrics are used. Value range between 0 and 1, near to 1 indicates efficiency of prediction of the model [13]. To verify the performance of the model, split the original data into test and training sets several times, and observe each set with different groups. Finally, sum up the performance of all training and test sets; this way observe the performance of the model. This method is called cross-validation [14]. Figure 3 presents a general workflow of ML for solar cells.

3 Machine Learning for Solar Cells

3.1 Naïve Bayes (NB)

Naïve Bayes classifier is a Supervised ML algorithm that works on the principle of Naïve Theorem. In the eighteenth century, Thomas Bayes and Clergy describe the concept of probability of events.

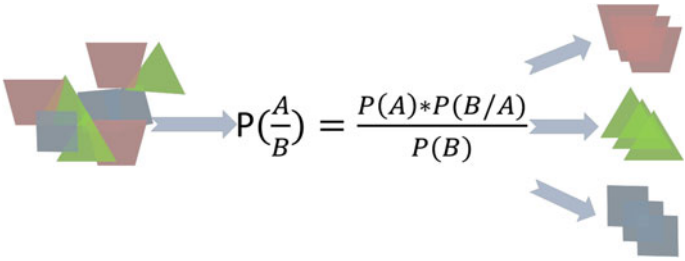

Fig. 4 Structure of Naïve Bayes

$$P\left(\frac{A}{B}\right) = \frac{P(A)*P(B/A)}{P(B)} \tag{1}$$

Here posterior probability $\boldsymbol{P}(\frac{\boldsymbol{A}}{\boldsymbol{B}})$ describes the conditional probability that calculates the hypothesis in the light of relevant observations, prior probability $\mathbf{P}(\mathbf{A})$ describes probability before collection of new data, likelihood probability $\mathbf{P}(\frac{\mathbf{B}}{\mathbf{A}})$ is the inverse of the posterior probability, and at the end, marginal likelihood probability $\mathbf{P}(\mathbf{B})$ is evidence. NB classifier selects the feature of class independently to the other classes as an input, applies the Naïve Bayes algorithm, and works on conditional probabilities by using prior knowledge after the calculation results were obtained based on prior records, as shown in Fig. 4 [15]. NB classifier is simple and computationally efficient and can operate categorical features directly without pre-processing, and can also handle missing, noisy data pretty well [16].

3.2 Artificial Neural Network (ANN)

In 1943, Warren McCullah and Water Pitts conceived the idea of ANN; the purpose of the ANN approach is to create a computational system that can solve real-life problems like the human brain. In the human brain, neutrons are interconnected similarly, in ANN neurons are interconnected in various layers, and these neurons form nodes [17]. Generally, ANN is consists of three layers as shown in Fig. 5. The input layer accepts inputs in various formats inserted by the user. Hidden layers are present between input and output layers, its main function is to calculate and find out hidden features and patterns of data. After the operation of the ANN algorithm data is transferred to output layers. Output layers present a series of transformations of hidden layers. The main advantages of the ANN algorithm are it can perform properly even if inadequate data is provided, programming data is stored in a network, not on a database, and can perform more than one simultaneous task.

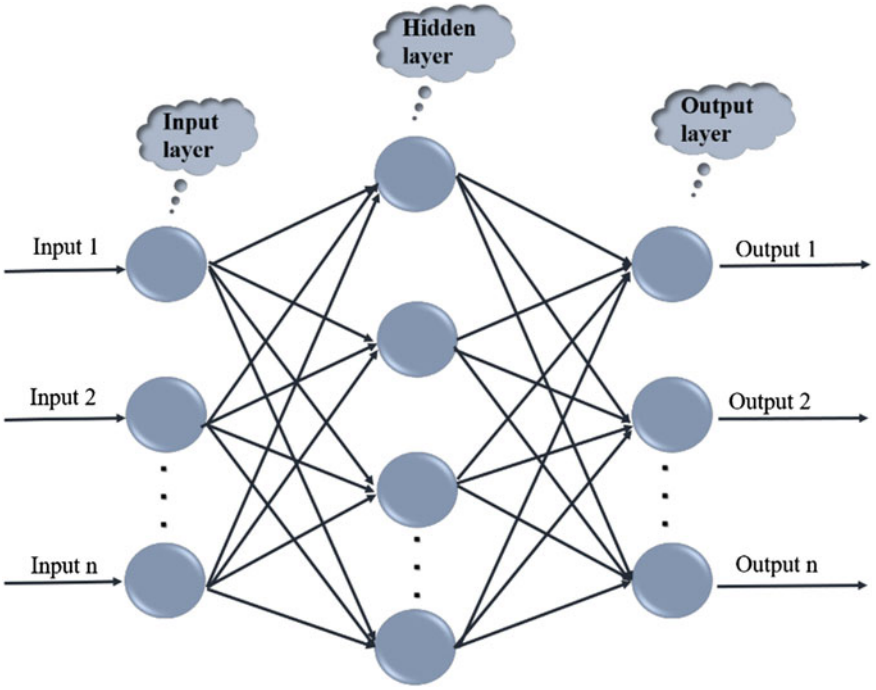

Fig. 5 Structure of an artificial neural network

3.3 Decision Trees (DT)

DT belongs to supervise ML and is very popular, user-friendly compared to other algorithms. Its tree-like structure represents the relationship between prediction and outcome. It works for both discrete (classification tree) and continuous (regression tree) [18]. It works on the principle of recursive partition, a single partition known as the root node. Nodes are conditions, branches represent the results of conditions, and leaf nodes are decisions based on conditions. Partition of the tree depends on decision-based conditions, splitting the original dataset into subsets until each subset is homogenous. Leaf nodes (tree terminates) represent predicted results depending on the conditions of roots and decision nodes. The advantage of DT is that it can handle missing and noisy data very well and can better work with discrete and continuous features. Figure 6 shows the structure of the decision tree.

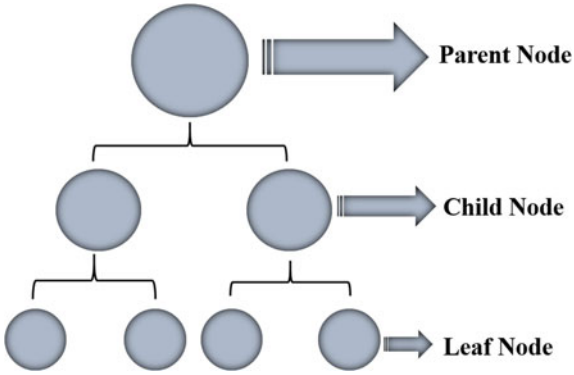

Fig. 6 Structure of decision trees

3.4 Other Machine Learning Techniques

ANN, NB, and DT SVM machine learning models are mostly used to investigate optimal material prediction, fabrication process, and device structure for solar cells. But not limited to these techniques, several other models can be used for this purpose like Random Forest (RF) [19], K-nearest neighbors (K-NN) [20], Boosted Regression Trees (BRT) [21], Kernel Ridge Regression (KRR) [22], Convolutional Neural Network (CNN) [23], Gradient Boosting Regression Tree (GRBT) [24], Artificial Neural Network (ANN) [25] Support Vector Machine (SVM) [26], and Genetic Algorithm (GA) [27].

4 Typical Applications of ML Tools for Solar Cells

Solar cells can meet human energy demands but finding suitable materials for high-performance solar cells is a crucial problem. Most often designing or predicting materials based on conventional methods (lab experiments) is expensive, time-consuming, and needs a lot of manpower. ML is a promising avenue for finding suitable and most efficient materials without a trial-and-error approach. Recently, researchers have published significant articles to predict properties of the material by ML like band gaps, solubility, density, refractive index, scattering, absorption, transmittance, reflectance, etc. ANN algorithm is most extensively used due to its superior properties and other algorithms are also used for this purpose as shown in Table 1 [28].

Table 1 List of ML models used for material prediction

Sr. No	Material	Property	Algorithm for ML	Precision accuracy	References
1.	Perovskite	Optical	ANN	66.5% prediction coefficient improved	[29]
2.	Y6 non-fullerene acceptor	PCE	DFT, RF	Means Absolute Error (MAE) ≈ 0.43, Correlation Efficient $r \approx 0.97$	[30]
3.	Organic solvent	Solubility of Solvent	Light GBM	log ± 0.59	[31]
4.	Dye-sensitizer	Charge transfer, optoelectronic properties, PCE of device	DFT, TD-DFT, GA-MLR	Average relative error ≈ 3.1%	[32]
5.	Pentacene	Charge carrier mobility	KRR	MAE of 3.0 ± 0.2, 5.3 ± 0.4, 9.7 ± 0.6	[33]
6.	ZnO	Morphology, PCE	DT, RF, ANN	Adj. R2 = 0.7232	[34]
7.	Perovskite	Stability and Band Gap	XGBoost, RF	Prediction accuracy of XGBoost R2 = 0.9935 and MAE = 0.0126, and for Random Forest R2 = 0.9410 and MAE = 0.1492	[35]
8.	Kesterite I2-II-IV-V4 Semiconductors	Band Gap	SVG with Radial-Bias Kernel	Root mean squared error = 283 meV	[36]
9.	Perovskite	PCE	SCAPS-1D	N/A	[37]
10.	Mono-crystalline module, CIGS Thin Film	I–V	ANN	Means Absolute Percentage Error (MAPE) = 0.874%	[38]
11.	Benzene	HOMO-LUMO	GA	N/A	[39]

4.1 Effect of Material Properties on PCE of Solar Cells

4.1.1 Effect of Band Gap

Li and his co-workers used real experimental data from the literature and predicted the values of band gaps and power conversion efficiency (PCE) using the ANN algorithm of ML, which further correlated with their experimental results of thin

film. In the first model, material composition is used as input and output is band gap values. In the second model, PCE is found by using band gap, ΔH, and ΔL as input [39]. The research is basically theoretical as DT algorithm is used to predict the PCE of perovskite solar cells, as shown in Fig. 7. For simulation parameters of the DT algorithm, solar cell capacitance simulator-one-dimensional (SCAPS-1D) software has been used. PCE of perovskite solar cells especially depends on band gap but many other input parameters are important like hole mobility, electron affinity, the concentration of donor and acceptor, etc. Furthermore, the device efficiency is improved from 13.29 to 16.68% by optimizing the band gap, thickness, number of defects, relative permeability, band gap energy, etc. [37].

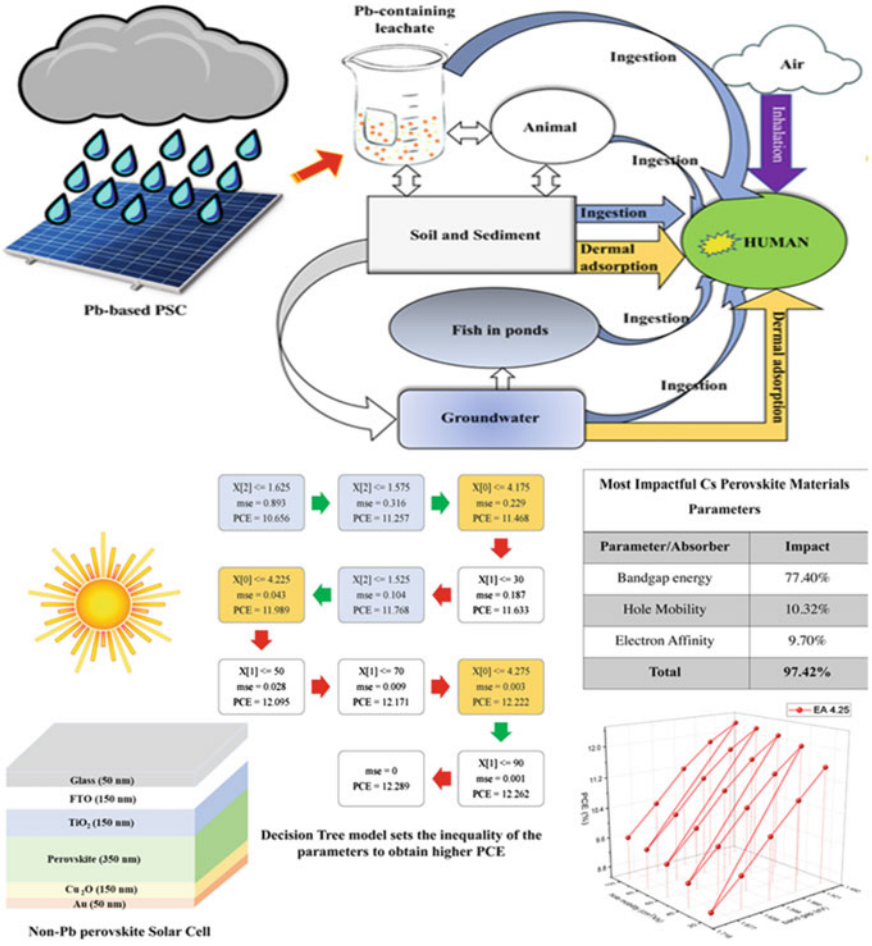

Fig. 7 Machine learning approach to delineate the impact of material properties on physics of solar cell device. Reprinted with permission from Ref. [37]. Copyrights 2022 American Chemical Society

4.1.2 Effect of Non-Fullerene Acceptor

Non-fullerene acceptors are a special type of acceptor used for organic solar cells. Recently, Zhao et al. aimed to predict the efficiency of novel non-fullerene acceptors by the leave-one-group out (LOGO) cross-validation method and give more accurate prediction than conventional ML methods [40]. Mahmood and his co-workers theoretically investigated the planar non-fullerene acceptors by DFT. Planar is introduced by restricting specific bond movement, these planars helped to tune band gap and improved absorbance of material, due to high planar conjugate system charge carrier and found improvement in mobility and photovoltaic performance [41]. Similarly, Xiao and his co-workers designed small molecule acceptors for P_3HT-based organic solar cells. Voc increased and device efficiency improved from 4.93 to 6.08% [42]. In this research article five algorithms RL, MLR, RF, ANN, and BRT have been constructed by Yoa and his co-workers; after predicting the ability evaluation of these five models, RL, MLR, and ANN algorithms show low correlation coefficient, and RF and BRT algorithms exhibit batter performance. To discover optimal D/A pairs for non-fullerene solar cells 32 million donor-acceptor (D/A) pairs were screened using RF and BRT algorithms, and six photovoltaic D/A pairs were selected and synthesized by a high screening method. Both models show good correlation efficient $r \approx 0.72$ of RF and $r \approx 0.70$ of BRT for experimental and predicted PCE [21]. Padula et al. evaluated 280 datasets of organic donors and acceptors and exploited the electronic and chemical properties of organic molecules for the prediction of the efficiency of solar cells. K-NN and R-NN are used for this purpose, and further, found the prediction capability of models to be up to $r \approx 0.7$. Such prediction capability of models may allow us to discover new materials [43].

4.1.3 Effects of Morphology

For PCE of OPVs Random Forest (RF), Gradient Boosting (GB), Artificial Neural Network (ANN), and k-Nearest Neighbor (k-NN) algorithms are used and considered 13 important microscopic properties of molecules as descriptors [35]. To predict the performance of OPVs, the thickness of film, morphology, topology, and band gap are studied but here we elaborate the importance of donors and acceptors in photocurrent generation. Firstly, Bayesian Machine Scientist (BMS) algorithm is used for intrinsic J_{SC}-vol% for material that is synthesized by a high-throughput screening method. Secondly, Random Decision Forest (RDF) is used for normalized J_{sc}-vol% of binary OPVs and the mean absolute error (RME) is below 0.20. The Decision Tree (DT) algorithm can be used to exploit the effects of morphology by hydrothermal experimental parameters while on the other hand, the effect of synthetic parameters on PEC of ZnO DSSCs can be investigated by ANN and Regression Tree (RT) [34]. In this article, the investigation is about the effect of annealing temperature on morphology, grain size, and PCE of organic solar cells (OSCs). The optimal grain size was found to be 9.5 nm at 110 °C for stimulatory discussion, in which DT, and Support Vector Machine (SVM) algorithm were used [45].

4.2 Prediction of Optimal Device Structure

To study the effect of mesoporous structure TiO_2 on the efficiency of DSSCs and PSCs, for the sake of analysis and optimization RF algorithm was employed. Simulation data was collected from SCAPS-1D software and more than 200 simulation samples were seeded. The thickness of mmp-TiO_2 increased from 0.0 to 0.5 μm, and comp-TiO_2 increased from 0.1 to 1.0 μm [46].

4.2.1 Anti-Reflective Coating (ARC)

Yan et al. investigated anti-reflection (AR) coating by a genetic algorithm [46]. There is a comparison of the double and four discrete layers and low refractive index AR coating of the SiO_2/TiO_2 and fabricated by RF magnetron sputtering method. In four discrete layers, low refractive indexes show omnidirectional AR characteristics because 34.4% J_{sc} improvement is observed and contribute to the enhancement of the photovoltaic efficiency of inverted metamorphic (IMM) solar cells [47]. They were able to successfully deposit 4-layer AR coatings on Si substrate and the results of SEM (at a scale of 200 nm) can be viewed in Fig. 8. Our careful observation of this micrograph clearly identifies the distinct layer boundaries as shown in Fig. 8a. This SEM image further indicates that such AR coating would have least nanoporosity of thin film and hence the lesser refractive index. Figure 8b shows the top view of the same substrate. It can be seen that the diameter of the nanorod is much smaller than the wavelength of visible light.

Ant Colony Algorithm (ACA) was used for three layers of refractive index for Si-based solar cells. 2.98% average reflectance is obtained, when the angle of incident of light varies from 0° to 80°, and by further increasing the angle of incident light from 0° to 90°, average reflectance also increased 6.56% [48]. GA algorithm was used to get the optimal combination of almost 1.4×10^9 potential candidates. The optimal combination of uniform and opal like structure of active layer show 98.1% quantum efficiency which is relatively higher than homogenous photoactive perovskite active materials [49]. AR coating for CuIn-GaSe (CIGS) solar cells is inquired from Simulated Annealing (SA) algorithm. 8.46% average reflectance is observed in MgF_2 layer, the range of wavelength is 350–1200 nm, and the angle of incident is from 0° to 80°. On the other hand, 11.30% average reflection is reduced, when a double layer of SiO_2/TiO_2 AR coating is investigated from 400 to 1100 nm wavelength and incident angle from 0° to 80° [50]. Figure 9 shows the schematic of a typical perovskite solar cell. As indicated by the arrow, the incident radiation comes from a medium with $n = 1.5$. As soon as the incident ray hits the surface, it interacts with fluorine-doped tin oxide, commonly referred to as FTO. The "t_{FTO}" refers to the thickness of this layer and is typically between 50 and 800 nm. Right above the perovskite layer is a uniform, homogeneous layer of TiO_2 of thickness 40 nm with the purpose of hole blocking. Since the optical band gap of TiO_2 is 385 nm, it is able to absorb some of the visible light. Below is the hybrid photoactive layer of methylammonium lead

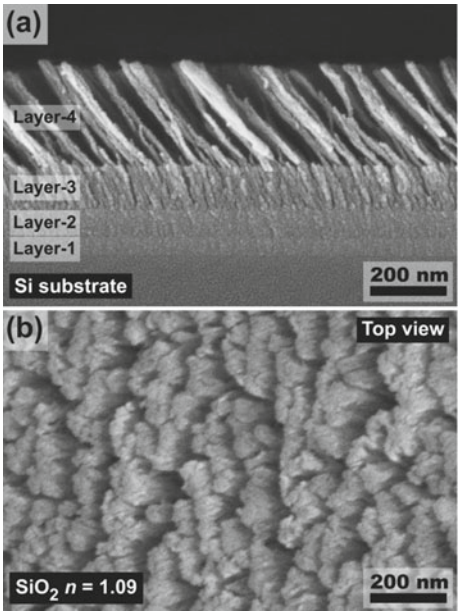

Fig. 8 **a** SEM micrograph of multiple layers of AR coatings on Si wafer. **b** Top view of the substrate. Reprinted with permission from Ref. [47]. Copyrights Material Views 2012

triiodide ($MAPbI_3$) perovskite. It is arranged in face centered cubic (FCC) structure with "ABC" stacking sequence. The radius of spheres and the thickness of this perovskite layer are important attributes that play a significant role in determining its photoactive nature. It is pertinent to mention here that due to the presence of hybrid perovskite layer beneath, hybrid character is induced in the photoactive layer. The "t_{uni}" as indicated in the diagram varies between 0 (absence of layer) to 100 nm. The hole transporting layer is placed below this uniform layer of perovskite and above the gold (Au) counter electrode. The "t_{spiro}" ranges from 50 to 800 nm. This optimum level of thickness allows the layer to maximize the reflection of non-absorbed light by Au counter electrode. Moreover, it also aids in constructive interference which in turn increases the light absorption inside the hybrid photoactive layer. Generally, the thickness of Au is limited to 100 nm so that it is able to completely block and reflect the light remaining. In conclusion, the iQE value increased to 2% more than the unstructured photoactive layer [49].

4.2.2 Effects of Light Scattering

Generally, absorbance is improved by introducing light-trapping structure, but on the other hand, electrical degradation occurs. Wang and his co-worker employed finite element method (FEM) software for addressing the issue of electrical degradation [50]. Optical properties were improved by introducing GaP (high refractive index > 3) as the first layer, and flat a-Si:H active layer has been used to avoid electrical degradation in hydrogenated amorphous silicon (a-Si:H) solar cells. Through this

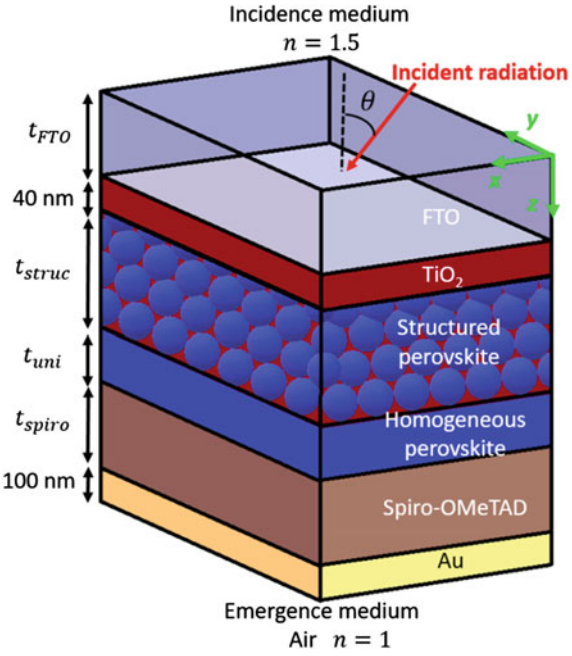

Fig. 9 Schematic of perovskite solar cell composed of hybrid photoactive structured and homogeneous layer designed to increase the light absorption. Reprinted with permission from Ref. [49]. Copyrights 2020 MDPI

strategy, the power conversion efficiency of solar cells was enhanced [51]. Alsaigh et al. [52] improved the power conversion efficiency of crystalline and amorphous solar cells by introducing them to the top layer of multi-layer optical designs. Two designs were heightened, inverted multi-element lenslet array (MELA) and non-uniform MELA. With the help of these layer reflection losses, they are reduced and promote light trapping from a wide range of incident angles. For these optical designs, optimal simulations are performed by COMSOL Multiphysics [53]. A bit more about the prediction of optimal solar cells (device optimization) through machine learning is summarized in Table 2.

Table 2 List of ML models used to predict optimal device structure

Sr. No	Material	Properties	ML (model)	References
1.	SiO_2 TiO_2	Anti-reflection coating	GA	[54, 55]
2.	TiO_2	Light scattering	RF	[56]
3.	–	Doping profile	GA	[57]
4.	Polymer-fullerene	Charge transfer, optoelectronic properties, PCE of device	RF, ANN	[58]
5.	GaN/AlGaN quantum wells	Structural parameter of GaN/AlGaN quantum well	GA	[12]

5 Conclusion and Future Recommendations

This book chapter aimed to introduce the importance of machine learning in material science and engineering. This chapter reviewed solar cells and relevant ML algorithms like ANN, DT, and NB. Review was classified into four categories, the aim of the first category was to briefly introduce the workflow of machine learning, e.g., raw data collection and its future engineering, selection and evaluation of models, and its applications. The second category briefly introduced ML algorithms that have a role in efficient material prediction and device optimization. The third category reviewed the impact of morphology, band gap, non-fullerene acceptor, solubility, and charge transfer on PCE of solar cells. Furthermore, device optimization was investigated, e.g., anti-reflective coating of low refractive index materials, and trapping of light from high refractive index materials. GA and ANN algorithms are extensively used for this purpose because they are user-friendly. This review highlights that ML technology has a lot of potential to accelerate the discovery of high-performance efficient material for solar cells.

Recently, research on ML for Organic Solar Cells (OSCs) has tremendously increased. The performance of OSCs specifically depends on solvents, crystallinity, molecular orientation of absorbing layer, and morphology of active and interfacial layers. The complex nature of organics is demanding more efficient and eco-economic, and eco-friendly ML models such as photovoltaic phenomena are related to microscopic properties and require high-accuracy quantum calculations. For high accuracy, large-scale virtual screening is required, but on the other hand, high computational cost made it difficult for large-scale virtual screening. To get trade-off situation between cost and accuracy, molecular fingerprints and descriptors are used for organic solar cells. Experimental validation is another important point to get a fruitful result of ML algorithms. Materials are screened through heuristic rules, but it is not sure whether they are synthesizable or not [60]. There are few papers in literature such as Saeki et al. in which they synthesized donor and fabricated device and its PCE was 0.53% and the prediction of RF model was 5.0–5.8% [58]. Much more work is needed in this aspect. In the present era, some degree of programming ability is required to use ML tools. It could be an easy task for those who have deep knowledge about data science and programming. There are many materials scientists who are working on OSCs and do not have such good command on ML tools. This may lead to incorrect interpretation of results. This problem can be solved by interfacing user-friendly graphics for material scientists and engineers. With the evolution of modern photo-energy sensitization in solar cells,i.e., quantum-dots (QDs) and metal clusters (MCs), the ML tools have gained even more importance as it has become critical to control the size of these sensitizers (i.e., QDs/MCs) to achieve well-optimized solar light harvesting, electron-hole transfer, and recombination control. All such characteristics are more often dependent on the size of the photosensitizers in emerging photovoltaic devices and relative energy band positions. The experimental development of such materials is quite expansive as compared to conventional absorbing

layers, which makes it vital to investigate and apply the ML concepts in solar cell technology for the cost-effective and time-saving development of modern solar cells.

References

1. Shaikh, M. R., Shaikh, S., Waghmare, S., Labade, S., & Tekale, A. (2017). A review paper on electricity generation from solar energy. *International Journal for Research in Applied Science and Engineering Technology, 887*. https://doi.org/10.22214/ijraset.2017.9272
2. This month in physics history. https://www.aps.org/publications/apsnews/200904/physicshistory.cfm
3. Fraas, L. M. (2014). History of solar cell development. In *Low-Cost Solar Electric Power* (p. 1).
4. Ibn-Mohammed, T., et al. (2017). Perovskite solar cells: An integrated hybrid lifecycle assessment and review in comparison with other photovoltaic technologies. *Renewable and Sustainable Energy Reviews, 80*, 1321–1344. https://doi.org/10.1016/j.rser.2017.05.095
5. (PDF) Systematic review elucidating the generations and classifications of solar cells contributing towards environmental sustainability integration | published in Reviews in Inorganic Chemistry. https://www.researchgate.net/publication/343261055_Systematic_review_elucidating_the_generations_and_classifications_of_solar_cells_contributing_towards_environmental_sustainability_integration
6. Ballaji, A., Mh, A., Swamy, K., Oommen, S., & Ankaiah, B. (2019). A detailed study on different generations of solar cell technologies with present scenario of solar PV efficiency and effect of cost on solar PV panel. *International Journal of Research in Advent Technology, 7*, 364–372. https://doi.org/10.32622/ijrat.74201963
7. (PDF) Review on life cycle assessment of solar photovoltaic panels. https://www.researchgate.net/publication/338384189_Review_on_Life_Cycle_Assessment_of_Solar_Photovoltaic_Panels
8. Pseudo-halide anion engineering for α-$FAPbI_3$ perovskite solar cells. https://www.researchgate.net/publication/350641338_Pseudohalide_anion_engineering_for_a-FAPbI3_perovskite_solar_cells
9. Monolithic perovskite/silicon tandem solar cell with >29% efficiency by enhanced hole extraction. Science. https://www.science.org/doi/10.1126/science.abd4016
10. Chebrolu, V. T., & Kim, H.-J. (2019). Recent progress in quantum dot sensitized solar cells: an inclusive review of photoanode, sensitizer, electrolyte, and the counter electrode. *Journal of Materials Chemistry C, 7*(17), 4911–4933. https://doi.org/10.1039/C8TC06476H
11. Choudhary, R., & Gianey, H. K. (2017). Comprehensive review on supervised machine learning algorithms. in *2017 International Conference on Machine Learning and Data Science (MLDS)* (pp. 37–43). https://doi.org/10.1109/MLDS.2017.11
12. Mahmood, A., & Wang, J.-L. (2021). Machine learning for high performance organic solar cells: Current scenario and future prospects. *Energy & Environmental Science, 14*(1), 90–105. https://doi.org/10.1039/D0EE02838J
13. Parikh, N., et al. (2022). Is machine learning redefining the perovskite solar cells? *Journal of Energy Chemistry, 66*, 74–90. https://doi.org/10.1016/j.jechem.2021.07.020
14. Practical Machine Learning in R | Wiley. Wiley.com. https://www.wiley.com/en-us/Practical+Machine+Learning+in+R-p-9781119591535
15. Abdualgalil, B., & Abraham, S. (2020). Applications of machine learning algorithms and performance comparison: A review. in *2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE)* (pp. 1–6). https://doi.org/10.1109/ic-ETITE47903.2020.490
16. A review on machine learning algorithms to predict daylighting inside buildings—ScienceDirect. https://www.sciencedirect.com/science/article/abs/pii/S0038092X20303509

17. Sustainability | Free Full-Text | Review on Machine Learning Techniques for Developing Pavement Performance Prediction Models. https://www.mdpi.com/2071-1050/13/9/5248
18. Sun, W., et al. (2019). Machine learning-assisted molecular design and efficiency prediction for high-performance organic photovoltaic materials. *Science Advances, 5*(11), eaay4275. https://doi.org/10.1126/sciadv.aay4275
19. Padula, D., & Troisi, A. (2019). Concurrent optimization of organic donor-acceptor pairs through machine learning. *Advances Energy Materials, 9*(40), 1902463. https://doi.org/10.1002/aenm.201902463
20. Machine learning for accelerating the discovery of high-performance donor/acceptor pairs in non-fullerene organic solar cells | *NPJ Computational Materials*. https://www.nature.com/articles/s41524-020-00388-2
21. Effect of increasing the descriptor set on machine learning prediction of small molecule-based organic solar cells | chemistry of materials. https://pubs.acs.org/doi/abs/10.1021/acs.chemmater.0c02325
22. Pokuri, B. S. S., Ghosal, S., Kokate, A., Sarkar, S., & Ganapathysubramanian, B. (2019). Interpretable deep learning for guided microstructure-property explorations in photovoltaics. *NPJ Computational Materials, 5*(1). https://doi.org/10.1038/s41524-019-0231-y
23. Sahu, H., & Ma, H. (2019). Unraveling correlations between molecular properties and device parameters of organic solar cells using machine learning. *The Journal of Physical Chemistry Letters, 10*(22), 7277–7284. https://doi.org/10.1021/acs.jpclett.9b02772
24. Majeed, N., Saladina, M., Krompiec, M., Greedy, S., Deibel, C., & MacKenzie, R. C. I. (2020). Using deep machine learning to understand the physical performance bottlenecks in novel thin-film solar cells. *Advanced Functional Materials, 30*(7), 1907259. https://doi.org/10.1002/adfm.201907259
25. Pilania, G., Balachandran, P. V., Kim, C., & Lookman, T. (2016). Finding new perovskite halides via machine learning. *Frontier in Materials, 3*. https://www.frontiersin.org/articles/10.3389/fmats.2016.00019
26. Boosting photoelectric performance of thin film GaAs solar cell based on multi-objective optimization for solar energy utilization. *Solar Energy, 230*, 1122–1132. https://doi.org/10.1016/j.solener.2021.11.031
27. A review on machine learning algorithms, tasks and applications. https://www.researchgate.net/publication/320609700_A_Review_on_Machine_Learning_Algorithms_Tasks_and_Applications
28. Kim, S. M., Naqvi, S. D. H., Kang, M. G., Song, H.-E., & Ahn, S. (2022). Optical characterization and prediction with neural network modeling of various stoichiometries of perovskite materials using a hyperregression method. *Nanomaterials Basel Switzerland, 12*(6), 932. https://doi.org/10.3390/nano12060932
29. Zhang, Q., et al. (2022). High-efficiency non-fullerene acceptors developed by machine learning and quantum chemistry. *Advanced Science, 9*(6), 2104742. https://doi.org/10.1002/advs.202104742
30. Ye, Z., & Ouyang, D. (2021). Prediction of small-molecule compound solubility in organic solvents by machine learning algorithms. *Journal of Cheminformatics, 13*(1), 98. https://doi.org/10.1186/s13321-021-00575-3
31. Accelerated discovery of high-efficient N-annulated perylene organic sensitizers for solar cells via machine learning and quantum chemistry—ScienceDirect. https://www.sciencedirect.com/science/article/abs/pii/S2352492820326155
32. Machine Learning—Based Charge Transport Computation for Pentacene—Lederer—2019—Advanced Theory and Simulations—Wiley Online Library. https://onlinelibrary.wiley.com/doi/abs/10.1002/adts.201800136
33. Analysis and prediction of hydrothermally synthesized ZnO-based dye-sensitized solar cell properties using statistical and machine-learning techniques | ACS Omega. https://pubs.acs.org/doi/10.1021/acsomega.1c04521
34. Machine learning stability and band gap of lead-free halide double perovskite materials for perovskite solar cells—ScienceDirect. https://www.sciencedirect.com/science/article/abs/pii/S0038092X21007878

35. Weston, L., & Stampfl, C. (2018). *Physical Review Materials, 2*(8), 085407. https://doi.org/10.1103/PhysRevMaterials.2.085407
36. Machine learning approach to delineate the impact of material properties on solar cell device physics | ACS Omega. https://pubs.acs.org/doi/10.1021/acsomega.2c01076
37. Applied Sciences | Free Full-Text | Prediction Model for the Performance of Different PV Modules Using Artificial Neural Networks | HTML. https://www.mdpi.com/2076-3417/12/7/3349/htm
38. Huwig, K., Fan, C., & Springborg, M. (2017). From properties to materials: An efficient and simple approach. *The Journal of Chemical Physics, 147*(23), 234105. https://doi.org/10.1063/1.5009548
39. Predictions and Strategies Learned from Machine Learning to Develop High-Performing Perovskite Solar Cells—Li—2019—Advanced Energy Materials—Wiley Online Library. https://onlinelibrary.wiley.com/doi/abs/10.1002/aenm.201901891
40. Zhao, Z.-W., del Cueto, M., & Troisi, A. (2022). Limitations of machine learning models when predicting compounds with completely new chemistries: possible improvements applied to the discovery of new non-fullerene acceptors. *Digital Discovery, 1*(3), 266–276. https://doi.org/10.1039/D2DD00004K
41. Mahmood, A., Tang, A., Wang, X., & Zhou, E. (2019). First-principles theoretical designing of planar non-fullerene small molecular acceptors for organic solar cells: manipulation of noncovalent interactions. *Physical Chemistry Chemical Physics, 21*(4), 2128–2139. https://doi.org/10.1039/C8CP05763J
42. Xiao, B., et al. (2017). Non-fullerene acceptors with A2 = A1 – D – A1 = A2 Skeleton containing Benzothiadiazole and Thiazolidine-2,4-Dione for high-performance P_3HT-based organic solar cells. *Solar RRL, 1*(11), 1700166. https://doi.org/10.1002/solr.201700166
43. Combining electronic and structural features in machine learning models to predict organic solar cells properties—Materials Horizons (RSC Publishing). https://pubs.rsc.org/en/content/articlelanding/2019/mh/c8mh01135d
44. Lan, F., Jiang, M., Wei, F., Tao, Q., & Li, G. (2016). Study of annealing induced nanoscale morphology change in organic solar cells with machine learning. in *2016 IEEE 16th International Conference on Nanotechnology (IEEE-NANO)* (pp. 329–332). https://doi.org/10.1109/NANO.2016.7751398
45. Al-Saban, O., & Abdellatif, S. O. (2021). Optoelectronic materials informatics: Utilizing random-forest machine learning in optimizing the harvesting capabilities of mesostructured-based solar cells. in *2021 International Telecommunications Conference (ITC-Egypt)* (pp. 1–4). https://doi.org/10.1109/ITC-Egypt52936.2021.9513898
46. Yan, X., et al. (2013). Enhanced omnidirectional photovoltaic performance of solar cells using multiple-discrete-layer tailored- and low-refractive index anti-reflection coatings. *Advanced Functional Materials, 23*(5), 583–590. https://doi.org/10.1002/adfm.201201032
47. Guo, X., et al. (2014). Design of broadband omnidirectional antireflection coatings using ant colony algorithm. *Optics Express, 22*(104), A1137–A1144. https://doi.org/10.1364/OE.22.0A1137
48. Lobet, M., et al. (2020). Opal-like photonic structuring of perovskite solar cells using a genetic algorithm approach. *Applied Sciences, 10*(5). https://doi.org/10.3390/app10051783
49. Broadband omnidirectional antireflection coatings for metal-backed solar cells optimized using simulated annealing algorithm incorporated with solar spectrum. https://opg.optica.org/oe/abstract.cfm?uri=oe-19-s4-a87
50. Wang, D., & Su, G. (2015). New strategy to promote conversion efficiency using high-index nanostructures in thin-film solar cells. *Scientific Reports, 4*(1), 7165. https://doi.org/10.1038/srep07165
51. Jäger, K., Fischer, M., van Swaaij, R. A. C. M. M., & Zeman, M. (2013). Designing optimized nano textures for thin-film silicon solar cells. *Optics Express, 21*(S4), A656. https://doi.org/10.1364/OE.21.00A656
52. Alsaigh, R. E., Alsaigh, R. E., Bauer, R., Lavery, M. P. J., & Lavery, M. P. J. (2020). Multi-layer light trapping structures for enhanced solar collection. *Optics Express, 28*(21), 31714–31728. https://doi.org/10.1364/OE.403990

53. Schubert, M. F., Mont, F. W., Chhajed, S., Poxson, D. J., Kim, J. K., & Schubert, E. F. (2008). Design of multilayer antireflection coatings made from co-sputtered and low-refractive-index materials by genetic algorithm. *Optics Express, 16*(8), 5290–5298. https://doi.org/10.1364/OE.16.005290
54. Zhang, Y.-J., Li, Y.-J., Lin, J., Fang, C.-L., & Liu, S.-Y. (2018). Application of millimeter-sized polymer cylindrical lens array concentrators in solar cells. *Chinese Physics B, 27*(5), 058801. https://doi.org/10.1088/1674-1056/27/5/058801
55. Al-Sabana, O., & Abdellatif, S. O. (2022). Optoelectronic devices informatics: optimizing DSSC performance using random-forest machine learning algorithm. *Optoelectronics Letters, 18*(3), 148–151. https://doi.org/10.1007/s11801-022-1115-9
56. Alì, G., Butera, F., & Rotundo, N. (2013). Geometrical and physical optimization of a photovoltaic cell by means of a genetic algorithm. *Journal of Computational Electronics, 13*(1), 323.
57. Nagasawa, S., Al-Naamani, E., & Saeki, A. (2018). Computer-aided screening of conjugated polymers for organic solar cell: Classification by random forest. *The Journal of Physical Chemistry Letters*. https://doi.org/10.1021/acs.jpclett.8b00635
58. Radosavljević, S., Radovanović, J., Milanović, V., & Tomić, S. (2014). Frequency up-conversion in nonpolar a-plane GaN/AlGaN based multiple quantum wells optimized for applications with silicon solar cells. *Journal of Applied Physics, 116*(3), 033703. https://doi.org/10.1063/1.4890029

Machine Learning-Driven Gas Identification in Gas Sensors

Shirong Huang, Alexander Croy, Bergoi Ibarlucea, and Gianaurelio Cuniberti

Abstract Gas identification plays a critical role in characterizing our (chemical) environment. It allows to warn of hazardous gases and may help to diagnose medical conditions. Miniaturized gas sensors, and especially those based on chemiresistive detection mechanisms, have gained rapid development and commercialization in the past decades due to their numerous advantageous characteristics, such as simple fabrication, easy operation, high sensitivity, ability to detect a wide range of gases, and compatibility with miniaturization as well as integration for portable applications. However, they suffer from a remarkable limitation, namely their low selectivity. Recently, machine learning-driven approaches to enhance the selectivity of gas sensors have attracted considerable interest in the community of gas sensors, increasing the analyte gas identification ability. In this chapter, firstly, we introduce the general approaches to enhance the selectivity of gas sensors implemented by machine learning techniques, which consists of the architecture scheme design of gas sensors (sensor array and single sensor architecture), the selection of gas sensing response features (steady-state feature and transient-state feature), and the utilization of gas sensing signal modulation techniques (sensing materials modulation, concentration modulation, and temperature modulation). Afterward, a specific application case using a machine learning-enabled smart gas sensor for the identification of industrial gases (PH_3 and NH_3) is presented, which is based on a single-channel device and utilizes multiple transient features of the response. We believe machine

S. Huang (✉) · B. Ibarlucea · G. Cuniberti (✉)
Institute for Materials Science and Max Bergmann Center for Biomaterials, TU Dresden, 01062 Dresden, Germany
e-mail: shirong.huang@tu-dresden.de

G. Cuniberti
e-mail: gianaurelio.cuniberti@tu-dresden.de

A. Croy
Institute for Physical Chemistry, Friedrich Schiller University Jena, Helmholtzweg 4, 07743 Jena, Germany

G. Cuniberti
Dresden Center for Computational Materials Science (DCMS), TU Dresden, 01062 Dresden, Germany

N. Joshi et al. (eds.), *Machine Learning for Advanced Functional Materials*,
https://doi.org/10.1007/978-981-99-0393-1_2

learning in combination with efficient sensing signal modulation techniques could be a feasible way to gain the gas identification capability of gas sensors.

Keywords Chemiresistive gas sensors · Selectivity · Machine learning · Electronic nose · Transient response features · Signal modulation

1 Introduction

Air pollution is the presence of substances in the atmosphere that are detrimental to the health of humans and other living beings, which is presently one of the major global social and environmental issues accompanied by terrible consequences. According to an estimate of the world health organization (WHO), there are seven million people killed by the air pollution around the world annually [1]. The European Union (EU) has identified seven main air pollutants (excluding greenhouse gases): nitrogen oxides (NO_x), carbon monoxide (CO), particulate matter (PM), sulfur dioxide (SO_2), ozone (O_3), ammonia (NH_3), and volatile organic compounds (VOCs) [2].

Take NH_3 as an example, it is an inorganic compound widely utilized in many industrial processes, such as manufacturing, food processing, refrigeration systems, and fertilizer production [3]. However, even exposure to ultra-low concentrations of NH_3 adversely affects human health [4]. Another example, phosphine (PH_3) is a colorless, flammable, toxic gas, which is widely used in the semiconductor industry (silicon semiconductors and photovoltaic process applications) [5]. Nevertheless, PH_3 is extremely toxic and exhibits an acute lethal effect on humans and animals by inhibiting aerobic respiration at even extremely low concentrations. Another large category of pollutants, volatile organic compounds (VOCs), are organic chemicals with a high vapor pressure at room temperature, among which some are prominent and representative indoor pollutants, such as benzene, toluene, ethylbenzene, xylene (BTEX), formaldehyde, and acetaldehyde [6] The United States Environmental Protection Agency (U.S. EPA) estimates that the VOCs level in indoor air is typically 2–5 times higher than that of outdoor air [7]. Commonplace items in our dwellings such as building materials, paints, furniture, cleaning agents, and cosmetics are potential sources of VOCs. However, considerable evidence suggests that a substantial number of these VOCs could cause adverse health effects, including sensory irritation, respiratory symptoms (asthma, allergy, etc.), and even cancer [8], as shown in Fig. 1. Indoor air quality of residential units and workplaces has been a serious concern since human beings spend more than 80% of their lifetime indoors, including domestic residences and working places [9].

Gas sensors play vital roles in monitoring and detecting hazardous gases, and ensuring public safety, air quality, or analyzing environments throughout many different fields [10]. They are useful not only in monitoring toxic gases in the atmosphere emitted from industry, but also in the control of indoor air quality [11], as well as safety in the vehicle. Their applications are virtually countless in the industry and span across a lot of industrial branches, including automotive, underground