Predicting and Optimizing the Power Performance of Dye-Sensitized Solar Cells Using Machine Learning Techniques

Deepanjali Chowdhury^{†‡}, Leen Al Homoud[†], Khandaker Akramul Haque[†], Katherine Davis[†]

[†]Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX, United States [‡]Department of Materials Science and Engineering, Cornell University, Ithaca, NY, United States dc997@cornell.edu, leen.alhomoud@tamu.edu, akramwired@tamu.edu, katedavis@tamu.edu

Abstract-Solar cells hold great potential for the future of energy production. Dye Sensitized Solar Cells (DSSCs) emerge as a promising contender among third-generation solar cells, due to their cost-effectiveness and flexibility. Ongoing experiments aim to further enhance their performance by exploring alternative materials, with a specific emphasis on refining the photoanode for increased efficiency. Within the photoanode, the semi-conducting oxides demonstrate excellent conductivity and porosity. However, despite their promise, testing various material variations in DSSCs is challenging due to time and resource constraints. This research builds on past studies by introducing machine learning to predict efficiencies for alternative DSSC configurations. specifically emphasizing the semiconducting oxide. This machine learning algorithm aims to be a cost-effective approach to identify low-cost, high-performance configurations for DSSCs and predict the system's performance across various parameters, including Power Conversion Efficiency (PCE) and Fill Factor (FF). Using machine learning, this study aims to simplify the evaluation process and speed up the identification of optimal materials and their weight percentages for the semi-conducting oxide, contributing to the advancement of sustainable and efficient DSSC technologies.

Index Terms—Dye-Sensitized Solar Cells, Materials Optimization, Machine Learning, Performance Prediction

I. INTRODUCTION

Solar energy is one of the most sustainable and abundant resources available, capable of meeting global energy demands. While silicon-based solar cells have achieved efficiencies of 15-21% with a lifespan of over 25 years [1], their bulkiness and manufacturing challenges have led to high costs and significant waste [2]. Second-generation thin-film solar cells, though more durable, suffer from lower efficiencies and scaling difficulties [1]. Third-generation solar cells, including quantum dots, perovskites, and DSSCs, are being explored for their non-toxic and organic properties, although they currently exhibit relatively low PCE [3].

DSSCs (Fig. 1) utilize non-toxic, sustainable materials for electricity production via photoelectrochemical processes [3]. This study emphasizes the role of semiconductor oxide materials in the photoanode, a critical component for light absorption and electron generation. While titanium dioxide (TiO_2) remains the top-performing material, alternative materials such as Au-TiO₂, ZnO, Ag₂O-ZnO, and others are also investigated [4, 5].



Fig. 1. Dye-Sensitized Solar Cell Schematic (Image reproduced under open access license) [6]

The ultimate goal of this research is to enhance the performance of DSSCs by optimizing the photoanode material composition, focusing on improving PCE by utilizing machine learning (ML) models. The expected outcomes include accurate predictions of PCE and bandgap for various semiconductor oxide materials in DSSC photoanodes. This study utilizes ML models trained with open-source material data to predict the performance of DSSCs, offering a cost-effective and efficient computational alternative to experimental methods.

A. DSSC Fabrication Process and Components in ML Model

DSSCs aim to utilize non-toxic, sustainable materials for cost-effective electricity production, leveraging photoelectrochemical principles to convert solar energy into electricity [3]. Their conductivity and porosity make them ideal for electrochemical energy production and storage [7].

1) DSSC Components: Key components of DSSCs include:

- **Transparent Conductive Oxide (TCO) Substrate:** Typically made of fluorine-doped tin oxide (FTO) or indium tin oxide (ITO) glass, allowing light into the cell while protecting it from the weather.
- **Photoanode:** Comprising of a porous semiconducting oxide (SCO) like titanium dioxide (TiO_2) and dye, it determines how much charge transport occurs.



Fig. 2. Fabrication Process of a DSSC

- **Counter Electrode:** Conductive glass coated with a catalyst, often platinum (Pt), completing the electrical circuit.
- Electrolyte: Typically containing a redox couple such as iodide/triiodide ions $(3I^-/I_3^-)$, enabling electron transfer between the photoanode and counter electrode to keep the cell stable.

The photoanode is crucial for light absorption and electron generation. While TiO_2 is the current top-performing material due to its stability and electron mobility, other materials such as $Au - TiO_2$, ZnO, $Ag_2O - ZnO$, Nb_2O_5 , SnO_2 , WO_3 , and Fe_2O_3 are being explored [4, 5].

2) DSSC Fabrication Process: The fabrication process follows below [8], as shown in Fig. 2:

- 1) Prepare conductive glass plates and coat them with a TiO_2 suspension.
- 2) Apply a thin layer of TiO_2 , spread it evenly over the glass plate, dry on a hot plate, and anneal in an oven at 150°C for 2 hours.
- 3) Soak the photoanode in dye in a dark room for 20 minutes to 24 hours.
- Coat the counter electrode with platinum and anneal in an oven at 150°C for 2 hours.
- 5) Apply the tri-iodide electrolyte to the counter electrode, followed by placing the soaked TiO_2 layer to complete the DSSC.

In this project, the ML model assumes a fabrication process similar to the above, with variations in the photoanode semiconductor oxide composition. Specifically, the composition of the photoanode semiconductor oxide varies, with a suggested ratio of 50% TiO_2 and 50% of the desired material selected by the user.

B. Feature Selection for ML Model PCE & I-V Curve Prediction

While ML models have previously predicted PCE and DSSC performance [9], this project aims to predict performance for DSSCs not created in a lab, using open-sourced



Fig. 3. Simple Flowchart of ML Model

material property data. The models will predict performance efficiencies of DSSCs with various semi-conducting oxide materials & compositions in the photoanode, enhancing solar energy conversion and storage [10–14].

Six features significantly impact DSSC performance: Absorbance & Transmittance, Bandgap, Electrical Conductivity & Impedance, and Recombination Rate [14].

- Absorbance & Transmittance: These quantify light absorbed by the SCO, essential for energy conversion efficiency in DSSCs. Higher absorbance increases photon capture, leading to more electron-hole pairs and improved photocurrent and PCE. Optimizing these factors is vital for maximizing DSSC performance and understanding light energy availability for current generation.
- **Bandgap:** Determines the energy needed for electron promotion between the conduction and valence bands, affecting device stability. A narrower bandgap allows lower-energy photons to create more electron-hole pairs, while a wider bandgap requires higher-energy photons, reducing light absorption efficiency. Extremely low bandgaps can lead to instability, while high bandgaps limit electron transitions.
- Electrical Conductivity & Impedance: These evaluate current flow within the SCO, impacting overall performance. High conductivity in materials like titanium dioxide (TiO_2) facilitates efficient electron transport, while impedance hinders charge movement, arising from material resistance and interfaces. High impedance reduces solar cell efficiency, making optimization crucial for DSSC stability.
- Recombination Rate: Optimal levels are essential to prevent the rapid degradation of solar cells. In DSSCs, recombination occurs when photo-generated electrons and holes combine before generating electrical output, reducing efficiency. Minimizing recombination is crucial for enhancing charge carrier separation and collection. Understanding and controlling the recombination rate in semiconductor materials is vital for optimizing DSSC performance. Strategies to reduce recombination include optimizing material morphology, removing surface defects, and precise doping. The user must input hole and electron mobility to calculate the recombination rate in this ML model.

II. METHODOLOGY

This investigation employed both experimental and computational methods to achieve a comprehensive understanding of the topic. The flow chart (Figure 4) outlines the various aspects investigated and what was aimed to be completed.



Fig. 4. High-Level Diagram of ML Model

A. Datasets

The data collection process involves two methods, as illustrated in Figure 5. The MaterialsProject.org API [15] was utilized to obtain the dielectric constant and bandgap of all compounds and elements accessible on the Materials Project platform. However, while Materials Project offers approximately 5,000 dielectric constants for various compounds and elements and 17,000 bandgaps for different compounds and elements, it lacks single-element predictions, especially for elements doped with TiO_2 . Thus, an ML model is employed to predict these values for any elements and compounds doped with TiO_2 as well as other metal oxides.

B. Machine Learning Models

This research employs four primary ML models to identify the most accurate ones. These models are Linear Regression, K-nearest neighbors (KNN), Partial Least Squares (PLS), and Support Vector Machine (SVM). Within each model type, a dual-model system is used: one model assesses the absorbance coefficient, while the other evaluates the bandgap.

a) Linear Regression: Linear Regression models the relationship between input features and the target variable by determining the best-fitting line that minimizes the differences between predicted and actual values [16]. In this study, it is used under the assumption that the relationship between predictors and the target in DSSC prediction is linear.

b) K-Nearest Neighbors: The K-Nearest Neighbors (KNN) algorithm classifies or predicts target values based on the characteristics of the nearest neighbors in the feature space [17]. In this research, KNN predicts the target by averaging

the Materials Project					
Naterials Explorer				References	Documentation
s	earch for materials inf	ormation by chemistry,	composition, or	property.	
Materials e.g. Li-Fe or Li,Fe or Li3Fe or mp-19017					Search
Filters	Reset	7,290 materials match your search Columns V			
Composition		× Iotal Dielectric Constant: 0.1 or more			
 Thermodynamics 		Material ID	Formula	Crystal System	Space Group Symbol
Structural Properties		mp-866056	AoN	Heragonal	P8ame
 Symmetry 		mp-36526	AdOF	Cubic	H43m
Calculated Properties	c	np-23558	Ag ₂ 8iO ₅	Orthorho	Prin2
 calculated Properties 	2	🛊 mp-4691	Ag ₂ CO ₅	Monodinic	P12/m1

Fig. 5. Materials Project Website (Not API)

or classifying based on the majority of k-nearest neighbors, where the value of k is chosen to balance sensitivity to local variations.

c) Partial Least Squares: Partial Least Squares (PLS) is employed to address datasets with many correlated predictors by combining elements of principal component analysis and multiple regression [18]. This method identifies latent variables that capture maximum variance in both predictors and the target variable, iteratively extracting components to make predictions.

d) Support Vector Machine: Support Vector Machine (SVM) is utilized to construct hyperplanes in high-dimensional

space for classification or regression tasks, optimizing the margin between classes or predicted values for enhanced accuracy [19]. The model leverages support vectors and kernel functions to manage non-linear decision boundaries in the data efficiently.

$$J_{sc} = \frac{q\Phi L\alpha}{1 - L^2 \alpha^2} \left[-L\alpha + \tanh\left(\frac{d}{L}\right) + \frac{L\alpha e^{-d/L}}{\cosh\left(\frac{d}{L}\right)} \right] \quad (1)$$

$$V_{oc} = \frac{kTm}{q} \ln \left[\frac{LJ_{sc}}{qD_n \tanh\left(\frac{d}{L}\right)} + 1 \right]$$
(2)
$$J = \Lambda + \left(3.25 + J_{sc} - \frac{qD_n e \cosh\left(\frac{d}{L}\right)}{L} \right)$$
$$\times \tanh\left(\frac{d}{L}\right) \left(e^{\frac{qV}{kTm}} - 1 \right)$$
(3)

1) Math Modeling: To find the I-V Curve of each DSSC configuration with the X-Doped metal oxide as the semiconducting oxide in the photoanode, math modeling, and assumptions were utilized [20], as seen in Equations 1–3. Table I summarizes the assumptions in the ML model equations. Values labeled as "ML Model" are predicted by the model itself. The result is an I-V curve that plots from 0 to 0.85 volts. Typically, the V_{oc} of the circuit is from 0.7 to 0.8, so the graph plots to 0.85. The only portion of the equation that changes when generating the I-V curve is the absorption coefficient and length or area of the cell depending on user input.

Parameters/Features	Value	Information				
k	1.381×10^{-23} J/K	Boltzmann Constant				
Q	$1.602 \times 10^{-19} \text{ C}$	Electron Charge				
L	$2.2361 \times 10^{-3} { m ~cm}$	Length of electron diffusion				
d	User Input	Length of TiO_2 (SCO)				
α	ML Model	Absorption Coefficient				
m	4.5	Ideal Factor				
D	$2.3 imes10^{-5}$ cm ² /s	Diffusion Coefficient				
n_0	10 ¹⁶ e ⁻ /cm ²	Electron Concentration				
t	0.01 seconds	Lifetime				
φ	$1 \times 10^{17} \text{ cm}^{-2}/\text{s}^{-1}$	Sunlight Intensity				
Т	300 K	Temperature				
FF	0.65	Fill Factor				
ε	MP_api	Dielectric Constant				
λ	600 nm	Wavelength of Light				
Λ	-	I-V Curve Multiplier				

TABLE I PARAMETERS AND FEATURES

III. RESULTS

A. Feature Results

1) Absorbance and Transmittance: In this ML model, key assumptions were derived from literature [20]. A dataset from the Materials Project API, comprising approximately 7,300 compounds, was used to determine the absorbance coefficient. Equation 4 converts the dielectric constant to absorbance:

$$\alpha = \frac{2\sqrt{2\pi}}{\lambda}\epsilon\tag{4}$$

Dielectric constants were found on the Materials Project API. With the absorbance coefficient established, Equations 1–3 were applied to calculate short circuit current density and open circuit voltage, both critical for assessing PCE.



Fig. 6. Bandgap Prediction Results

2) Bandgap: The bandgap is essential for electron transitions between the valence and conduction bands in semiconductor materials. Based on the data in Figure 6, the ML model's predicted bandgap values differ from literature values. For instance, the predicted bandgap for TiO_2 (Rutile) is higher than the actual value, while FeO and Fe_2O_3 show significantly lower predictions. Conversely, the model performs well for CuO and V_2O_3 , where predictions are closer to literature values. These results indicate areas for potential refinement in the model's predictive capabilities while demonstrating its ability to detect factors influencing bandgap fluctuations.

3) Electrical Conductivity and Impedance: Conductivity and impedance are critical for the performance of DSSCs, as they influence charge carrier flow [14]. The equations 5-6 were used to determine impedance and conductivity, with the Length (L) and Width (W) of the SCO set to a default of 1 cm.

$$\rho = \frac{R \cdot A}{L} = \frac{V_{oc} \cdot (L \cdot W)}{I_{sc} \cdot L} \tag{5}$$

$$\sigma = \frac{1}{\rho} \tag{6}$$

4) Recombination Rate: The recombination rate is based on the relationship between hole (p) and electron (n) mobilities and a constant, B [21]. After analyzing further literature [22], it was found that the B constant is the Stefan-Boltzmann Constant (k) and emissivity. Additionally, if the solar cell is at thermal equilibrium, its emissivity equals the absorption coefficient, which is accounted for in the ML model. Thus, the recombination rate in the ML model is calculated by multiplying the Boltzmann constant, absorption coefficient, electron mobility, and hole mobility, as seen in Equation 7.

$$R = Bnp = (k\epsilon)np = (k\alpha)np \tag{7}$$

B. Front End Results

As seen in Figures 7 and 8, various values are calculated and put on display in the front end. All Front-End results are



Fig. 7. ML Model User Input



Fig. 8. Front End Output Results

locally hosted through Streamlit and are available on Github¹. The user can input a variety of values seen in Figure 7, notably

¹https://github.com/deepachowdhury4/dssc_machine_learning.git

able to adjust the percent of the dopant in the SCO metal oxide.



Fig. 9. PCE Prediction Results

The PCE was calculated based on Equation 3, which is derived from the ideal diode equation [23]. In this ML model, a multiplier (Λ) ranging from 3.2 to 6.4, which is the same for the absorption coefficient and bandgap, is necessary to obtain a value within a reasonable PCE range. This multiplier is derived from an intrinsic material property that requires further investigation. In the future, exploring this property could enhance the model's predictive accuracy and overall performance. The FF is determined through Equations 8 and 9. V_{oc} is the x-intercept of the I-V curve, and J_{sc} is the yintercept. P_{max} is calculated as $J_{max} \cdot V_{max}$, which is found by identifying the maximum area under the curve up to V_{oc} .

$$FF = \frac{P_{\text{MAX}}}{J_{\text{sc}} \cdot V_{\text{oc}}} \tag{8}$$

$$P_{\rm MAX} = J_{\rm MP} \cdot V_{\rm MP} \tag{9}$$

IV. CONCLUSIONS AND FUTURE WORKS

This paper explores the potential of Dye-Sensitized Solar Cells as a sustainable energy solution by investigating various semiconducting oxide materials and utilizing machine learning (ML) to enhance material selection and predict performance cost-effectively. The model demonstrated varying performance, with bandgap prediction significantly outperforming PCE prediction due to a greater volume of available data. This highlights the critical role of data availability in influencing model accuracy. Insights gained from this research could drive advancements not only in DSSC technology but also in other solar technologies, such as perovskites, underscoring the need for efficient and sustainable energy solutions to combat climate change.

The developed ML model offers a user-friendly interface for predicting DSSC performance, integrating several models with KNN as the top performer. It provides predictions for Absorption Coefficient, PCE, and Bandgap. This research highlights the potential of ML models in advancing solar energy technologies by addressing the need for improved accuracy and reduced computational demands, focusing on key areas for improvement such as data availability, expanding ML model range, and increasing user input combinations.

ACKNOWLEDGMENTS

This research project was supported by the Department of Electrical and Computer Engineering at Texas A&M University under the Senior Design Capstone Project Funds.

References

- J. Jean, P. R. Brown, R. L. Jaffe, T. Buonassisi, and V. Bulović, "Pathways for solar photovoltaics," *Energy Environmental Science*, vol. 8, no. 4, pp. 1200–1219, 2015. [Online]. Available: http: //dx.doi.org/10.1039/C4EE04073B
- [2] T. M. Brown, F. De Rossi, F. Di Giacomo, G. Mincuzzi, V. Zardetto, A. Reale, and A. Di Carlo, "Progress in flexible dye solar cell materials, processes and devices," *Journal of Materials Chemistry* A, vol. 2, no. 28, pp. 10788–10817, 2014. [Online]. Available: http://dx.doi.org/10.1039/C4TA00902A
- [3] G. Nandan Arka, S. Bhushan Prasad, and S. Singh, "Comprehensive study on dye sensitized solar cell in subsystem level to excel performance potential: A review," *Solar Energy*, vol. 226, pp. 192–213, 2021. [Online]. Available: https://www.sciencedirect.com/science/articl e/pii/S0038092X21006927
- [4] S. K. Balasingam, M. Lee, M. G. Kang, and Y. Jun, "Improvement of dye-sensitized solar cells toward the broader light harvesting of the solar spectrum," *Chemical Communications*, vol. 49, no. 15, pp. 1471–1487, 2013. [Online]. Available: http://dx.doi.org/10.1039/C2CC37616D
- [5] D. Kishore Kumar, J. Kříž, N. Bennett, B. Chen, H. Upadhayaya, K. R. Reddy, and V. Sadhu, "Functionalized metal oxide nanoparticles for efficient dye-sensitized solar cells (dsscs): A review," *Materials Science for Energy Technologies*, vol. 3, pp. 472–481, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2589299120300185
- [6] T. H. Syed and W. Wei, "Technoeconomic analysis of dye sensitized solar cells (dsscs) with ws2/carbon composite as counter electrode material," *Inorganics*, vol. 10, no. 11, p. 191, 2022. [Online]. Available: https://www.mdpi.com/2304-6740/10/11/191
- [7] R. Singh and H.-W. Rhee, "The rise of bio-inspired energy devices," *Energy Storage Materials*, vol. 23, pp. 390–408, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2405829 719303265
- [8] M. I. Khan, A. Basir, S. K. Rahman, K. Rahat, and M. Huq, "Production and testing of dye sensitized solar cell," in 2015 International Conference on Electrical Engineering and Information Communication Technology (ICEEICT), Conference Proceedings, pp. 1–3.
- [9] A. Eibeck, D. Nurkowski, A. Menon, J. Bai, J. Wu, L. Zhou, S. Mosbach, J. Akroyd, and M. Kraft, "Predicting power conversion efficiency of organic photovoltaics: Models and data analysis," ACS Omega, vol. 6, no. 37, pp. 23764–23775, 2021, doi: 10.1021/acsomega.1c02156. [Online]. Available: https://doi.org/10.1021/acsomega.1c02156
- [10] Y. Cao, Y. Liu, S. M. Zakeeruddin, A. Hagfeldt, and M. Grätzel, "Direct contact of selective charge extraction layers enables high-efficiency molecular photovoltaics," *Joule*, vol. 2, no. 6, pp. 1108–1117, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/ S2542435118301326
- [11] S. Karthick, K. Hemalatha, S. K. Balasingam, F. Manik Clinton, S. Akshaya, and H.-J. Kim, *Dye-Sensitized Solar Cells: History, Components, Configuration, and Working Principle*, 2019, pp. 1–16. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/97 81119557401.ch1
- [12] Y. Ren, D. Zhang, J. Suo, Y. Cao, F. T. Eickemeyer, N. Vlachopoulos, S. M. Zakeeruddin, A. Hagfeldt, and M. Grätzel, "Hydroxamic acid preadsorption raises the efficiency of cosensitized solar cells," *Nature*, vol. 613, no. 7942, pp. 60–65, 2023.
- [13] M. Shakeel Ahmad, A. K. Pandey, and N. Abd Rahim, "Advancements in the development of tio2 photoanodes and its fabrication methods for dye sensitized solar cell (dssc) applications. a review," *Renewable* and Sustainable Energy Reviews, vol. 77, pp. 89–108, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364032 117304616
- [14] K. Sharma, V. Sharma, and S. S. Sharma, "Dye-sensitized solar cells: Fundamentals and current status," *Nanoscale Res Lett*, vol. 13, no. 1, p. 381, 2018, 1556-276x Sharma, Khushboo Sharma, Vinay Sharma, S S SR/FTP/PS-112/2012/Science and Engineering Research Board/ Journal Article Review United States 2018/11/30 Nanoscale Res Lett. 2018 Nov 28;13(1):381. doi: 10.1186/s11671-018-2760-6.

- [15] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, and K. A. Persson, "Commentary: The Materials Project: A materials genome approach to accelerating materials innovation," *APL Materials*, vol. 1, no. 1, p. 011002, 07 2013. [Online]. Available: https://doi.org/10.1063/1.4812323
- [16] D. Maulud and A. M. Abdulazeez, "A review on linear regression comprehensive in machine learning," *Journal of Applied Science and Technology Trends*, vol. 1, no. 2, pp. 140–147, 2020. [Online]. Available: https://jastt.org/index.php/jasttpath/article/view/57
- [17] P. Cunningham and S. J. Delany, "k-nearest neighbour classifiers a tutorial," ACM Comput. Surv., vol. 54, no. 6, p. Article 128, 2021. [Online]. Available: https://doi.org/10.1145/3459665
- [18] H. Kaneko, "Genetic algorithm-based partial least-squares with only the first component for model interpretation," ACS Omega, vol. 7, no. 10, pp. 8968–8979, 2022, 2470-1343 Kaneko, Hiromasa Orcid: 0000-0001-8367-6476 Journal Article United States 2022/03/22 ACS Omega. 2022 Mar 4;7(10):8968-8979. doi: 10.1021/acsomega.1c07379. eCollection 2022 Mar 15.
- [19] J. Gopala Krishna and K. Roy, "Qspr modeling of absorption maxima of dyes used in dye sensitized solar cells (dsscs)," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, vol. 265, p. 120387, 2022. [Online]. Available: https://www.sciencedirect.com/scie nce/article/pii/S1386142521009641
- [20] A. El Tayyan, "Dye sensitized solar cell: Parameters calculation and model integration," *Journal of Electron Devices*, vol. 11, pp. 616–624, 2011.
- [21] M. A. Green, Solar cells: operating principles, technology, and system applications. United States: Prentice-Hall, Inc.,Englewood Cliffs, NJ, 1982. [Online]. Available: https://www.osti.gov/biblio/6051511
- [22] F. Kreith and M. Bohn, *Principles of Heat Transfer*. PWS Publishing Company, 1997. [Online]. Available: https://books.google.com/books?i d=Cf4uAQAAIAAJ
- [23] K. McIntosh and C. Honsberg, "The influence of edge recombination on a solar cell's iv curve," 2000.