参赛学生姓名：王 傲

中学：上海外国语大学附属外国语学校

省份：上海

国家/地区：中华人民共和国

指导老师姓名：潘丽坤（英才计划导师）

指导老师单位：华东师范大学

论 文 题 目：Machine Learning-Driven Development of Semiconductor Photothermal Materials by Tuning Bandgap for Enhancing Solar Water Evaporation

2026 S.-T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

Machine Learning-Driven Development of Semiconductor Photothermal Materials by Tuning Bandgap for Enhancing Solar Water Evaporation

Ao Wang

Shanghai Foreign Language School Affiliated to SISU, Shanghai, 200083, China

2026 STYou High school Science Award
仅用于2026年友邦中学科学奖论文公示

Abstract

Solar water evaporation outstands as a promising method to solve the global freshwater shortage due to its lower cost, green and sustainable. Semiconductors photothermal materials (SPMs), with modifiable bandgap, and high performance, exhibit excellent prospects for practical solar evaporation. However, the precise influence of SPMs properties on evaporation performance remains unclear, necessitating long time and large costs to develop by trial-and-error tests.

Machine learning (ML) using data-driven approach to achieve feature classification and performance prediction. This work employed four ML models were employed to analyze the database collected from the latest research articles of SPMs, using data-driven approach to achieve feature classification and performance prediction. Shapley Additive Explanation analysis revealed the impact features for evaporation performance of SPMs. Moreover, experimental validation for the ML model's reliability was achieved doping sodium borohydride ($NaBH_4$) to titanium dioxide (TiO_2) for narrowing bandgap by calcining under 300 °C, 500 °C and 700 °C for one hour, respectively. Various characterizations were analyzed through SEM, XPS and UV-Vis confirmed Ti^{4+} converted into Ti^{3+}, resulting in forming oxygen vacancy to narrow bandgap of four samples from 3.36 eV (P25), 3.33 eV (300 °C), 3.12 eV (500 °C 5) to 2.09 eV (700 °C). The experimental results indicate that the evaporation rate of TiO_2 -700 °C improved 17.5% than primary literature, with evaporation rates of 1.12, 1.62, 1.89 and 2.28 kg m^{-2} h^{-1} under 1sun, 2sun, 3sun and 4sun, respectively. This study highlights that ML strategies can enhance the efficiency of exploring high-performance SPMs for solar-driven water evaporation.

Keywords: Solar evaporation performance, Machine learning, Semiconductor photothermal materials, Bandgap, Absorbance

Contents

1. Introduction.....1

2. Methods….....3

2.1 Dataset construction and feature screening.....3

2.2 Training ML models and accuracy assessment.....3

3. Experimental system….....4

3.1 Materials and instruments.....4

3.2 Samples preparation.....5

3.3 Membranes preparation.....5

3.4 Solar water evaporation experimental tests.....6

4. Experimental results and discussion.....7

4.1 Doping mechanism of TiO_2.....7

4.2 Pearson correlation coefficient study.....8

4.3 ER prediction.....9

4.4 EE prediction.....10

4.5 Feature importance and SHAP values analysis.....11

4.6 Experimental validation.....12

5. Conclusions.....21

References.....22

Acknowledgements.....26

Appendix A: Supplementary Information (data collection and testing).....28

Appendix B: Validating machine learning model stability.....48

Appendix C: Experimental Process.....56

Appendix D: 半导体光热材料特性分析的机器学习算法源代码.....57

2026 S.T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

1. Introduction

Currently the problem of global freshwater shortage is becoming more and more serious due to rapid population growth and accelerating industrialization. Seawater desalination has emerged as an effective approach to address this issue [1,2]. Solar-driven photothermal evaporation technology is an eco-friendly and sustainable solution to the freshwater crisis [3,4]. It primarily operates via three modes: bottom heating, bulk heating, and interfacial heating [5-7]. Interfacial solar evaporation (ISE) captures solar energy and converts into heat at the air-photothermal material-water interface. This improves thermal conversion efficiency, minimizes heat loss, and boosts freshwater production [8,9]. The interfacial heating process involves two distinct stages: evaporation and diffusion. Evaporation occurs when disruption breaks interfacial hydrogen bonds, and evaporation enthalpy is lower than bonded free water, resulting in thermal diffusion across the interface [10]. Photothermal materials float on the water surface and directly convert absorbed solar energy into heat, driving water evaporation at the water-air interface. The performance of ISE is governed by material microstructures, insulation layer performance, and solar-thermal conversion efficiency [11,12].

Recently, a variety of solar photothermal materials have been reported, including carbon, metal, semiconductor materials [13]. Particularly, semiconductor photothermal materials (SPMs) have gained extensive attention owing to excellent light-to-heat conversion capability, good chemical stability and broad applicability across multiple industries [6,14-16]. For ISE, the performance of SPMs is mainly influenced by their narrow bandgap and solar absorbance. When photon energy exceeds the semiconductor bandgap, absorbed photons generate electron–hole pairs under sunlight [17]. Photo-generated carriers relax toward band extrema, converting surplus energy into heat via thermalization. A narrower bandgap facilitates light absorption over a wider wavelength range. High light absorbance boosts heat generation by forming charge carriers, which then recombine to release thermal energy. This process also involves the absorption of electromagnetic waves, such as infrared radiation [13,15]. Moreover, solar irradiation, environmental temperature and humidity exert significant impacts on evaporation efficiency (EE) in the desalination process [18]. However, the interplay between these factors and solar water evaporation remains complex and unclear. Current SPMs perform well in desalination, but novel SPMs are still needed to achieve high evaporation rate (ER) in practice. Conventional design approaches involve complex, multi-step processes including material synthesis, property analysis, and evaporation evaluation, develop by trial-and-error tests. Therefore, optimal and resource-conserving methods for designing high-performance SPMs are imperative.

Machine learning (ML) has become a transformative tool in materials science, offering

2026 S.T. Yau High School Science Award
仅用于2026年丘成桐中学科学奖论文公示

a powerful means to automatically identify patterns and learning from complex datasets, and generate data-driven predictions [19-25]. ML technology integrating with theory and experimental enables predicting material properties, optimize parameters and structures, designing novel materials, and uncovering new mechanisms beyond intuition [26,27]. The operational mechanism of photothermal material can be accurately predicted by training various ML models with large datasets curated from relevant literatures. According to Gradient Boosting Decision Tree (GBDT) ensure each new tree correct the residuals of the previous trees in iterative process, Wang et al. applied GBDT regressor to predict the desalination capacity of faradic materials [28]. Ali et al. explored the extreme gradient boosting (XGBT) improves speed and prediction accuracy by using a new tree learning algorithm, leveraging parallel and distributed computing [29]. Elsheikh et al. applied support vector machines (SVM) for improving the performance of solar stills and water yield prediction [30]. In our previous study, voting regressor (VR) was employed to improve regression and explore carbon-based materials for solar evaporation. These works investigated solar desalination capabilities of carbon-derived photothermal absorbers, so ML-assisted exploration of SPMs remains largely underexplored.

In this study, four ML models (GBDT, XGBT, SVM and VR) were employed to predict the ER and EE of SPMs, and to establish the evaporation performance relationships with structure properties and experiment conditions. Particularly, two key material features of bandgap (BG) and absorbance (A); a property parameter of evaporation system of thermal conductivity (TC), and three operational conditions illumination conditions (sun), room temperature (T), and relative humidity (H); which were considered as inputs. The schematic process of predicting the evaporation performance of SPMs through ML methods is demonstrated in Fig. 1.

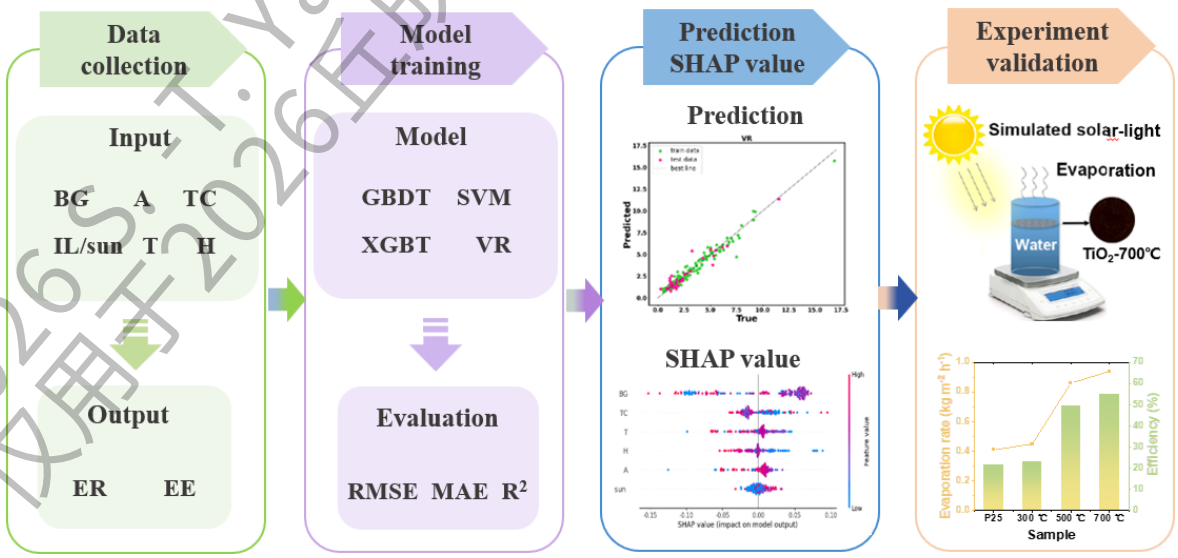


Fig. 1. Diagram for predicting ER and EE of SPMs using ML.

2026 S.T. Yau High School Science Award
仅用于2026丘成桐中学科学奖公示

The key research objectives of this study are:

- Collect dataset (six input and two output features) from published literatures researched on semiconductor photothermal materials for solar water evaporation.
- Train machine learning models and evaluate performance by root mean square error (RMSE), mean absolute error (MAE) and R^2 values.
- Predictive ER and EE by four ML models, and analyze feature importance of six input features in the model was evaluated using Shapley Additive Explanation (SHAP).
- Prepare black TiO_2 samples by calcining mix of P25 and $NaBH_4$ under temperature 300 ℃, 500 ℃, 700 ℃, respectively.
- Evaporation performance will test under different solar irradiation (1 sun, 2 sun, 3 sun, 4 sun) to confirm the reliability and applicability of ML method.

By achieving these objectives, this study highlights that ML strategies can accelerate development of high-performance SPMs for solar water evaporation, improving the efficiency of seawater desalination.

2. Methods

2.1. Dataset construction and feature screening

The data collected for this research comprises 196 samples, all sourced from the published literature. Comprehensive details of the dataset are presented in Table S1, as shown in Supplementary Information document. Considering that data of features are missing exceeding 15% in Table S1, Multiple Imputation was applied to handle missing data by synthesizing all available data to derive optimal imputed values [20]. In this dataset, ER and EE were used as the prediction target to evaluate solar water evaporation performance. The key properties of semiconductor materials BG, A are regarded to effectively influence the evaporation performance. Additionally, operational condition sun determines the light energy providing heat for evaporation process. Ambient environmental factors T and H also affect evaporation performance. To evaluate and identify the most accurate ML model for predicting the solar water evaporation performance, six commonly used features (BG, A, TC, sun, T, and H) served as input variables for the ML algorithms for predicting ER and EE.

2.2. Training ML models and accuracy assessment

In this work, VR used for ensemble method to combining the predictions from GBDT, XGBT, and SVM models, which leads to more robust predictions compared to individual model. The performance of the four ML models was evaluated through random allocation of the dataset with 80% for training sets and 20% for testing sets. Cross validation with five-

2026 S.T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

fold was performed to train each ML model, each algorithm underwent five training and evaluation cycles. In addition, Bayesian optimization was applied to optimize the hyperparameters of each model, and Table S2 and Table S3 summarize the optimized hyperparameters for each model. Using the optimal hyperparameters, each model's performance was evaluated on the training sets and testing sets using RMSE, MAE, and coefficient of determination (R^2). These evaluation indicators are defined as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (1)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (2)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \quad (3)$$

where, n denotes the sample size, y_i and $\hat{y}_i$ are actual value and predicted value, respectively, and $\bar{y}_i$ is the mean of y_i. Notably, for both values of RMSE and MAE are small signify higher predictive accuracy, smaller RMSE, and MAE indicate prediction, while the larger R^2 means better performance of the model ($0 \le R^2 \le 1$).

In the study, the weights for each base model in the VR were determined based on their performance on the validation set (specifically, the R^2).

3. Experimental

3.1. Materials and instruments

A commercial fine titanium dioxide (TiO_2) powder (P25) was used as starting material. The sodium borohydride ($NaBH_4$) was used as reduction agent. The P25 and $NaBH_4$ were acquired from Sinopharm Chemical Reagent Co. Ltd., and all chemicals were used without any further treatment. All experiments used deionized water.

The main instruments used in the sample preparation of this project are as follows:

(1) High-precision electronic analytical balance, model FA1204B, Jucheng Group.

(2) Tubular furnace, model OTF-1200X, Hefei Kexing Materials Technology Company.

(3) Ultrasonic cleaning machine, model UC-7180L, Shenzhen Langjie Ultrasonic Electrical Appliance Company.

(4) High-speed centrifuge, model TGL-15B, Shanghai Anting Scientific Instrument Factory.

(5) Vacuum drying oven, model DHG-9055A, Shanghai Yiheng Scientific Instrument Company.

2026 S.T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

3.2. Samples preparation

Samples preparation processes are demonstrated in Fig. 2, for the preparation of each sample, 0.4g TiO2 and 0.15g $NaBH_4$ weighed and placed in a ceramic crucible. The two materials should be ground to ensure complete mixing. Then, they should be separately placed in ceramic crucible boxes and placed in a tubular heating furnace for baking. Four types of samples were prepared according to this mass ratio: the original sample, heated at 300°C for 1 hour, heated at 500°C for 1 hour, and heated at 700°C for 1 hour.



Fig. 2. Samples preparation process.

Place the crucible in a tubular heating furnace, protected by nitrogen gas (flow rate of 0.2 L/min), and raise the temperature at a rate of 5°C per minute to the predetermined temperature, then hold it for 1 hour. Reach the predetermined temperatures (300°C, 500°C, and 700°C) respectively. After cooling to room temperature, remove the obtained titanium dioxide and grind it. Mix it for 1 hour by ultrasonic centrifugation, then wash and filter with water. Put the filtered sample into a vacuum box at 70°C to dry for 2 hours to obtain the prepared sample, as shown in Fig. 2. To facilitate distinction and comparison, we uniformly name the prepared samples. The original sample is named TiO_2-P25, while the samples after 1-hour heating at 300°C, 500°C and 700°C are respectively named TiO_2-300°C, TiO_2-500°C, and TiO_2-700°C.

3.3. Membranes preparation

Membranes preparation process as demonstrated in Fig. 3, we weighed 10 mg of

different pure TiO2 samples, respectively, at 300°C for 1 hour, 500°C for 1 hour, and 700°C for 1 hour. Then, we took 10 ml of deionized water for each sample and mixed them together. We placed the mixture in an ultrasonic machine for 1 hour to dissolve. After that, we removed the mixture and used a standard cellulose paper with a diameter of 3.8 cm. We performed filtration to produce 4 different water-evaporated TiO_2 membranes.

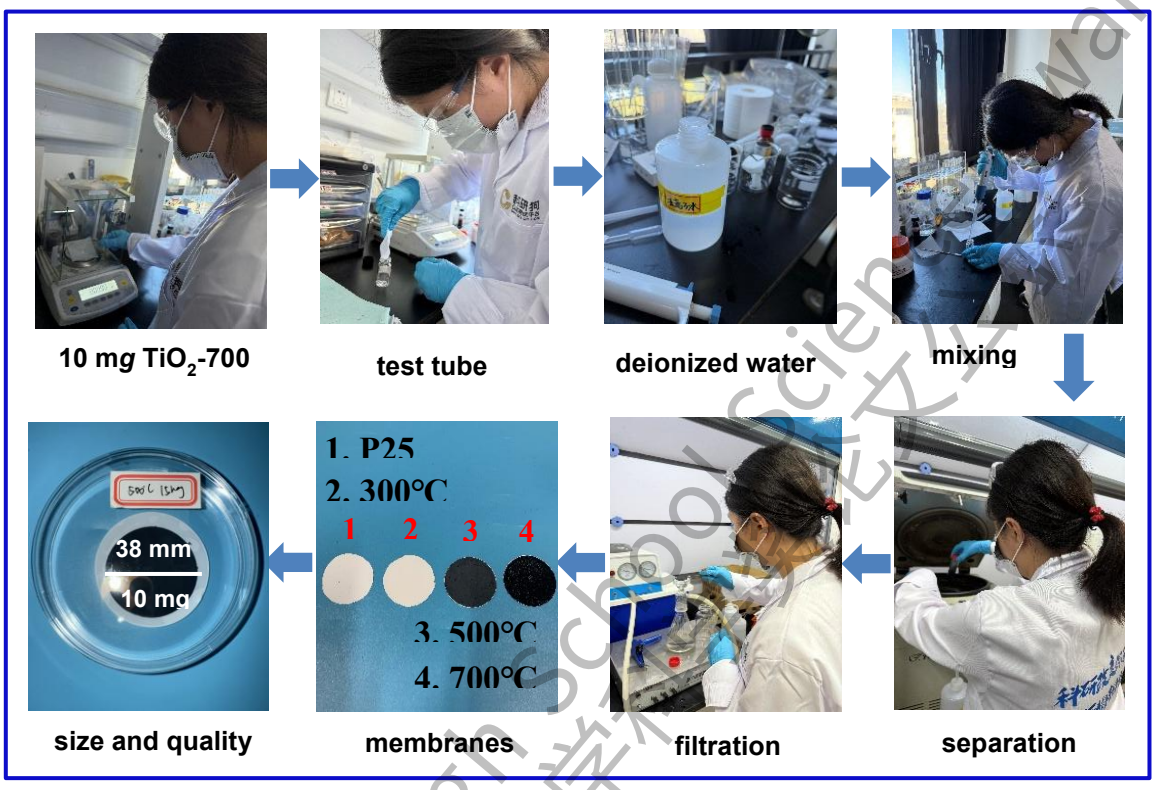


Fig. 3. Membrane preparation process.

3.4. Solar water evaporation experimental tests

The black TiO_2 powder was prepared by a $NaBH_4$-assisted method. In brief, 2 g P25 powder were mixed with 0.75 g $NaBH_4$ powder, and then the mixed power was homogenized by grinding. Moreover, the powder mixture was transferred into tube furnace and purged with argon gas for 30 min. After that, the mixtures were calcined at 300 ℃, 500 ℃, and 700 ℃ under an argon gas for 1 h, respectively, and obtained black TiO_2 samples were named 300 ℃, 500 ℃ and 700 ℃. Finally, the black TiO_2 samples were washed with deionized water two times, respectively.

The efficiency of solar-driven water evaporation was quantified through the experimental setup presented in Fig. 4(a). Light source is a xenon lamp as a chamber optical system. Under the sun illumination, the mass reduction was continuously measured using a high precision analytical balance. As shown in Fig. 4(b), 10 mg TiO_2 samples (P25, 300 ℃, 500 ℃ and 700 ℃) were dissolved in 10 mL deionized water with ultrasound dispersion for 1 h at room temperature, and then deposited onto mixed cellulose ester (MCE) membranes

2026 S. J. Yau High School Science Award 仅用于2026丘成桐中学科学奖

via vacuum filtration process to fabricate into evaporator, respectively. These membranes were cut into 38 mm diameter circle. Which were floated on water using a polyvinylidene foam insulating support and overlaid with an airlaid paper. Thermal conductivity of MCE membranes is 0.073 W m^{-1} K^{-1} as shown in Table S4. The membrane evaporator converts absorbed light into heat, leading to localized evaporation and steam generation. The variations in water surface temperature and the corresponding thermal distribution were captured using an infrared camera, as presented in Fig. 4(c). Absorbed light energy is rapidly converted into heat by the membrane evaporator, thereby inducing a rapid increase in its surface temperature and stabilize at approximately 51.3 °C within just 10 min under three sun irradiations.

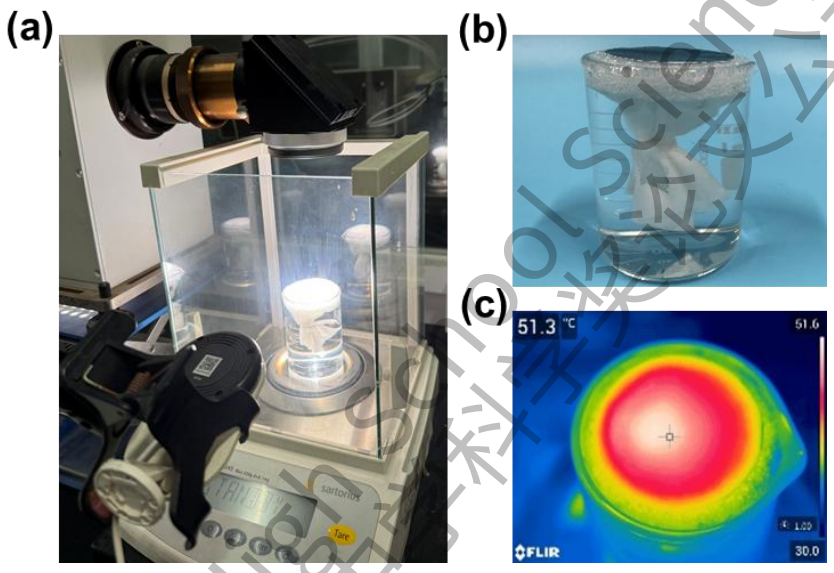


Fig. 4. Solar water evaporation experimental tests. (a) Testing system setup, (b) Modified TiO_2 membrane evaporators, (c) IR image of 700 °C membrane surface**.**

All tests were performed under controlled ambient conditions (26 °C, 65% relative humidity). Following each test cycle, the samples were thoroughly dried at 50 °C in a vacuum oven for subsequent reuse. Comprehensive details regarding the materials, characterization methods, and Supplementary Information provided evaporation testing procedure.

4. Experimental results and discussion

4.1. Doping mechanism of TiO_2

To reduce bandgap of original TiO_2 sample for increasing the visible light absorption, $NaBH_4$ was employed as reduction agent, which can provide active hydrogen to convert Ti^{4+} into Ti^{3+} and form oxygen vacancies, as the primary origin of the Ti 3d bandgap states [31]. The thermal decomposition of $NaBH_4$ follows the below reaction:

2026 S.T. Yau High School Science Award
仅用于2026丘成桐中学科学奖公示

$$NaBH_4 \rightarrow \sum_i solid_i\,(Na, B, H) + \sum_j gas_j\,(B, H) \quad (4)$$

where the solid residue and unreacted reactant can be easily removed by water.

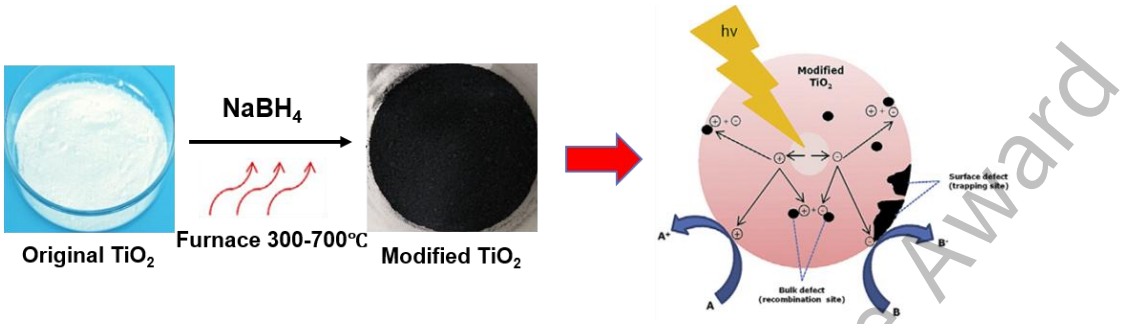


Fig. 5. Modified schematic diagram of TiO_2 [31].

As described in Fig. 5, by applying $NaBH_4$ modified TiO_2 in the tube furnace under 300-700 °C, defects confirmed to be oxygen vacancies mainly formed bulk defect which trigger recombination of electrons and holes. These introduced defects (oxygen vacancies) mean that the excess electrons generated from the active hydrogen reaction remain at the positions of oxygen vacancies, and thus act as F center that has a tendency to absorb visible light. Theoretically, the concentration of the defects should be high enough to create a new band that leads to the visible light absorption of the modified TiO_2.

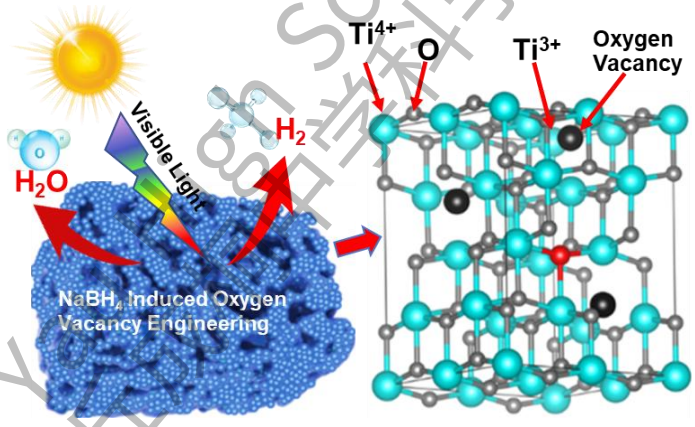


Fig. 6. Supercell model and oxygen vacancy [32].

To study the doping mechanisms, P. J. Ying et. al. applied DFT + U calculations to analyze atomic model of the 2 × 2 ×1 supercell, as described in Fig. 6 [32]. It contains 48 atoms (16 Ti and 32 O atoms). An oxygen vacancy in TiO_2 will induce two excess electrons, which will cause a reduction of two Ti^{4+} cations and the formation of two Ti^{3+} ions. In addition, because a H substitutes an O, that is, H^- replaces O^{2-}, it should reduce one Ti^{4+} to Ti^{3+}. The defects (oxygen vacancy) are uniformly distributed in TiO2 to reduce the interaction between them.

4.2. Pearson correlation coefficient study

In this work, the interrelationships among features (BG, A, TC, sun, T, H) and targets

(ER, EE) were evaluated using Pearson correlation coefficients, as presented in Fig. 7. Each matrix element represents the correlation coefficient between its respective row and column variables. This value falls in between -1 and 1, providing information on how the two variables correlate to each other. As shown in Fig. 7(a), the results show that a very high positive correlation (correlation coefficient of 0.92) between sun and ER, indicating that higher irradiance directly enhances solar energy capture. Mechanistically, increased photon flux elevates light-induced carrier generation efficiency in semiconductors, thus enhancing energy conversion. In contrast, no significant linear correlations were observed between the other features and the predicted variables of ER and EE, a key consideration in multivariate analysis.

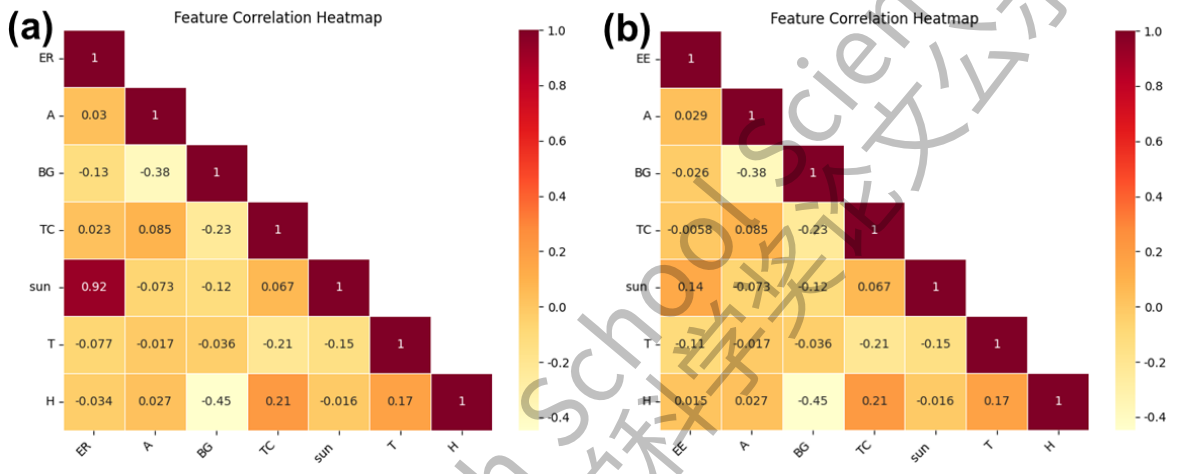


Fig. 7. Pearson correlation coefficient matrix between the six input features（A, BG, TC, sun, T, H) and (a) target variable (ER), (b) target variable (EE).

4.3. ER prediction

ER was employed as the primary performance metric to evaluate the evaporation performance of SPMs. Model efficacy was quantified using RMSE and MAE on separately training sets and test sets, as shown in Table S5 and Table S6, with lower values indicating better performance. Comparative analysis indicates that the VR model as optimal, given its better error statistics.

Additionally, scatter plots illustrate the error by comparing the predicted versus actual ER for the four ML models, as shown in Fig. 8. Blue and red points denote the training and test data for ER, respectively. Obviously, these blue points and red points in Fig. 8(a) closely cluster around the best-fit line in the VR model. Furthermore, the prediction error (RMSE, MAE, R^2) of VR model are lower than other models as shown in Table S5, which indicates VR model has high accuracy. This advantage of VR model is due to integrate three distinct sets of ER prediction results of GBDT, SVM and XGBT models which are approximate and without overfitting, resulting in overperforms the individual models present as its estimators.

2026 S. Tung High School Science Award
仅用于2026江成桐中学科学奖公示

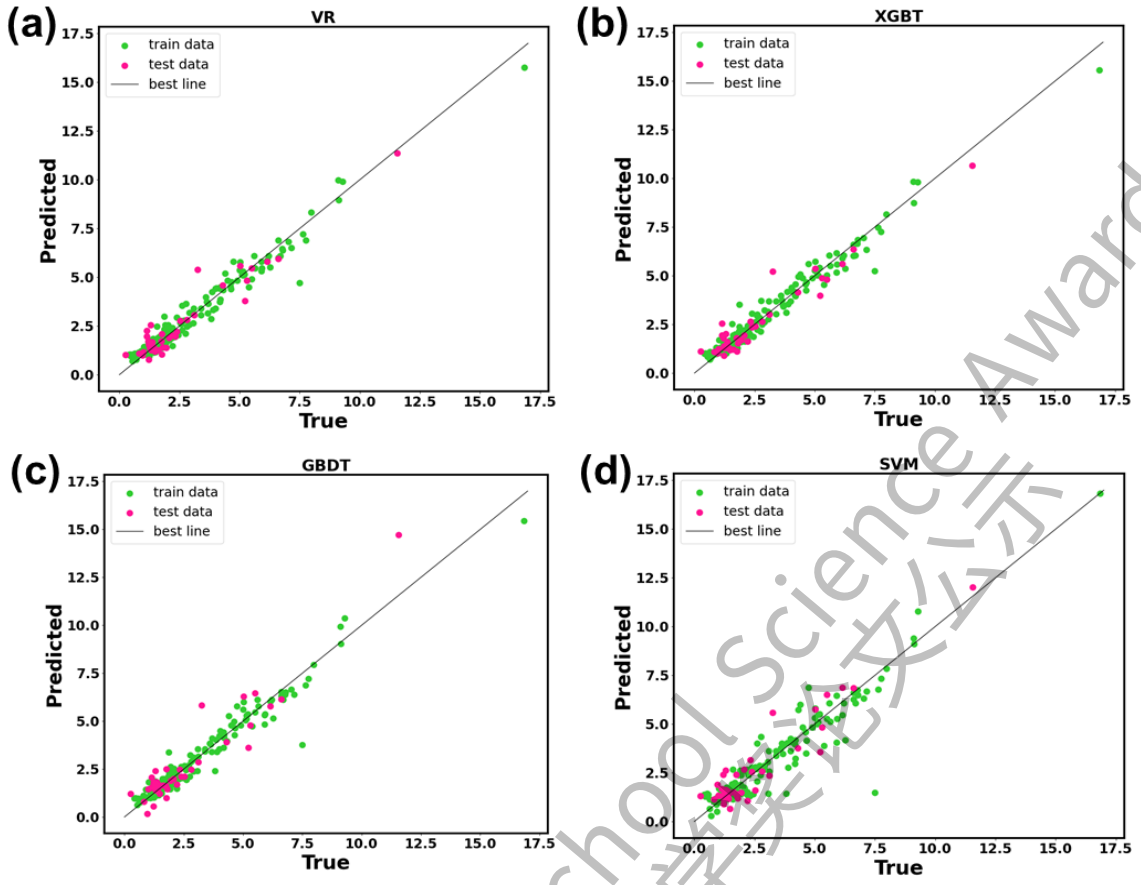


Fig. 8. ER prediction scatter plots for the four ML models: (a) VR, (b) XGBT, (c) GBDT, (d) SVM.

4.4. EE prediction

The EE is a critical factor in evaluating performance of SPMs on converting full-spectrum solar energy into recoverable heat and depends on their specific properties, including bandgap, and light absorption capability. The EE prediction error of four ML models on the training sets and test sets are present in Tables S7 and Table S8. Fig. 9. displays scatter plots comparing the predicted versus actual values for the four ML models. In these plots, training data for EE are shown in blue, while test data are denoted in red. Comparison of the predicted against actual values are illustrated via scatter plots in Fig. 9, where blue and red points denote the training and test data for EE, respectively. Fig. 9(a) indicates the VR model performs better due to it combining the predictions of multiple diverse estimators to smooth out errors and captures different aspects of the data. Therefore, VR model can achieve superior accuracy than single model on EE training data.

2026 S. T. You High School Science Award
仅用于2026丁佑同中学科学奖论文公示

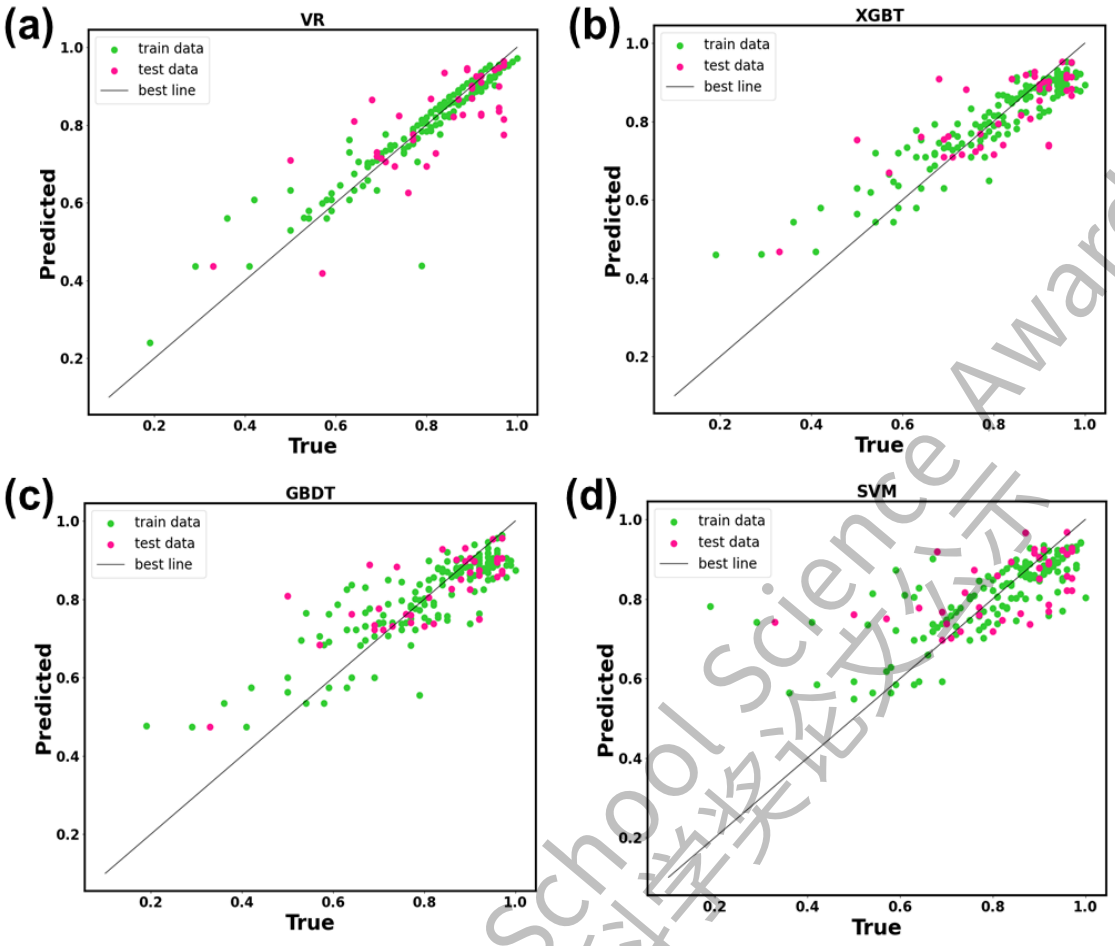


Fig. 9. EE prediction scatter plots for the four ML models: (a) VR, (b) XGBT, (c) GBDT, (d) SVM.

4.5. Feature importance and SHAP values analysis

Enhancing model interpretability is crucial for developing ML. SHAP values quantify the contribution of each feature to a specific prediction, and thus it can be used to analyze which features are most important for the model and how they affect the outcome. The SHAP plots shown in Fig. 10(a) highlight the impact of six features (BG, A, TC, sun, T, and H) on the model's output of ER. As shown in Fig. 10(a), among the six features, sun demonstrates a marked positive association with ER, indicating significant influence on ER prediction. When increasing sun light intensity, SPMs can achieve more solar energy absorption, resulting in improving water mass loss to increase ER. Similarly, TC is a system-level property dominated by the membrane, resulting in primary influencing ER. It has been reported that the interfacial heating is concentrated on the upper surface of the membrane evaporation with low thermal conductivity materials to prevent heart diffusion to the container, thereby improving ER [33]. Absorbance is essential for efficient photothermal conversion in SPMs, with higher absorbance leading to greater light harvesting and heat generation to improve ER [34]. BG reduction can decrease the intrinsic energy barrier for electron excitation from valence band to conduction band in SPMs, and prolong the activity

2026 S. T. Yau High School Science Award 仅用于2026丘成桐中学科学奖论文公开

under visible light irradiation. Ambient temperature and relative humidity which also directly impact ER because they affect energy available and vapor pressure, respectively.

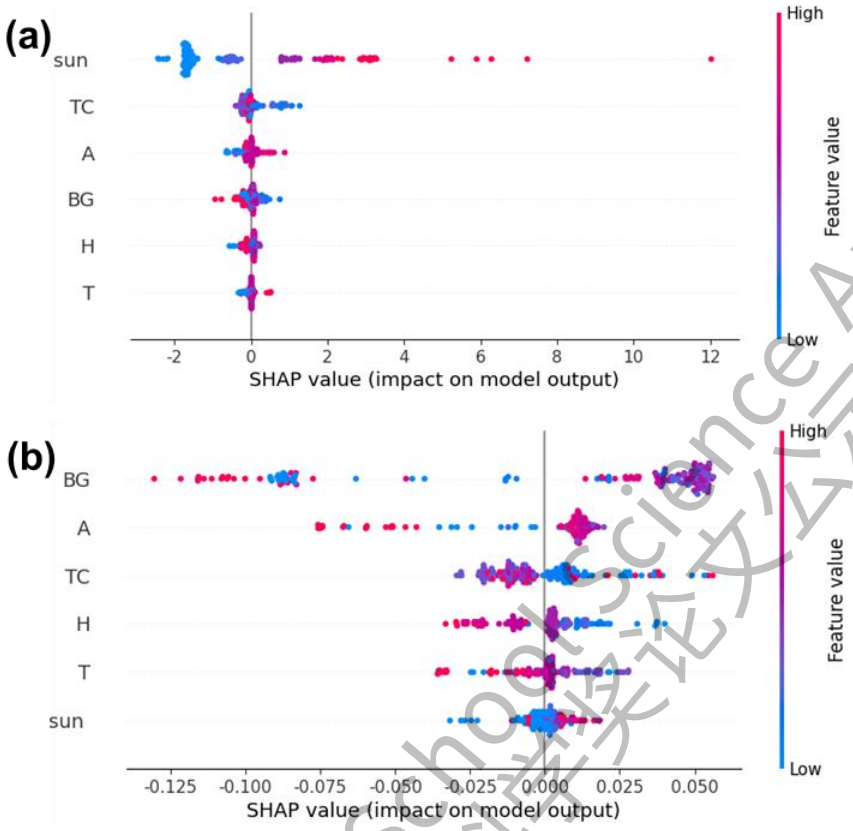


Fig. 10. Analyzing feature contribution with SHAP: (a) predicted ER using the VR model, (b) predicted EE using the VR model**.**

As shown in Fig. 10(b), the SHAP value for EE prediction indicates that no single feature exhibits a dominant influence, which is consistent with the model's difficulty in capturing the underlying physics of EE as shown in Fig. 4. Relatively speaking, BG and A have the most significant impacts on the SHAP value among all features, providing reliable physical interpretability. TC shows smaller influences on SHAP value of EE. This lack of clear, dominant feature contributions is consistent with the lower R^2 for EE prediction, suggesting that EE is governed by more complex, multi-factor interactions that are not yet fully captured by the current dataset. Larger and more diverse dataset may improve both the accuracy and the interpretability of EE prediction. To further validate the ML prediction, in the following section the solar water evaporation tests of SPMs was measured.

4.6. Experimental validation

To validate the predictive efficacy of the ML models, we employed black TiO_2 as a photothermal material for performance evaluation in interfacial evaporation. By using P25 as the starting material and $NaBH_4$ as the reducing agent, we thoroughly mixed them and then placed the mixture in a tube furnace for reaction. As depicted in Fig. S1, with the

2026 Si Yuan High School Science Award
仅用于2026年思源科学奖论文公示

increase in reaction temperature, the color of the samples gradually changed from white to black.

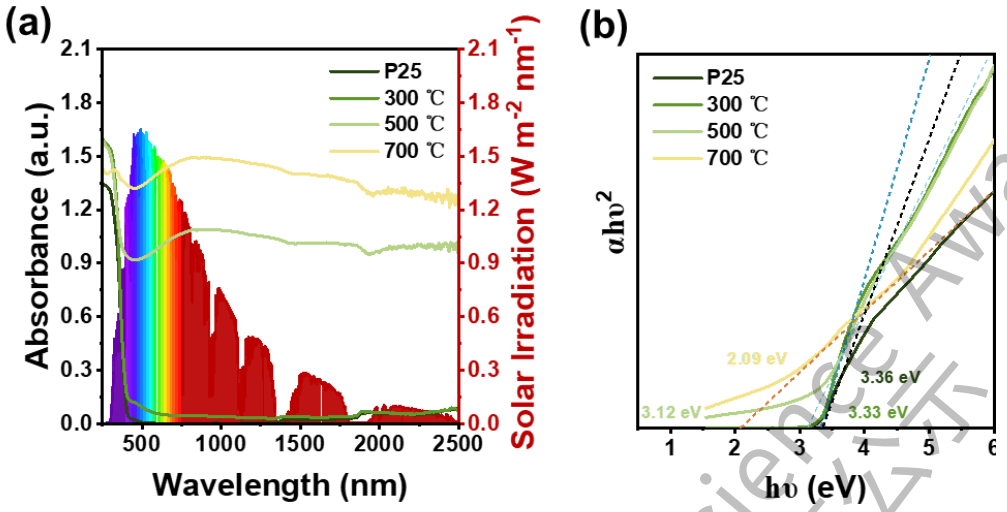


Fig. 11. The four different samples properties (a) absorbance, (b) bandgap.

As shown in Fig. 11(a), the UV-Vis-NIR absorption spectra of the gray-black and black samples clearly demonstrate their excellent light absorption capabilities, whereas the white raw material and the sample calcined at 300°C exhibit extremely poor light absorption. As the reaction temperature rises, the intensity of light absorption gradually increases, which is consistent with the color change of the samples. The bandgap energy of the samples is calculated based on their absorption spectra. As shown in Fig. 11(b), the bandgap of P25 is 3.36 eV. With the increase in reaction temperature, the bandgap gradually decreases, which is consistent with the phenomenon of gradually enhanced light absorption intensity in the absorption spectra. The gradual decrease in bandgap resulted from the creation of oxygen vacancies induced by reduction, which causes the transformation of Ti^{4+} to Ti^{3+} [35, 36].

To delve into the influence mechanism of valence states and oxygen vacancies on the bandgap, four samples were characterized by X-ray photoelectron spectroscopy (XPS) [37]. As shown in Fig. 12a, the XPS survey indicates that the sample is composed solely of Ti, O and C. As shown in Fig. 13(a), Ti 2p spectral analysis was carried out on P25 and TiO_2 calcined at different reaction temperatures. The analysis results reveal that within the energy range of 455 to 470 eV, XPS peaks for Ti $2p_{3/2}$ and Ti $2p_{1/2}$ can be observed at binding energies of 458.88 eV and 464.68 eV, respectively. These peaks are typical characteristic peaks of Ti^{4+}-O bonds. For the calcined TiO_2, the XPS peak position of Ti $2p_{3/2}$ shifts from the initial 458.88 eV to 458.68 eV, 458.58 eV, and 458.58 eV successively. At the same time, the XPS peak position of Ti $2p_{1/2}$ also shifts from 464.68 eV to 464.48 eV, 464.18 eV, and 464.18 eV. In combination with Figs. 12(b)-(e), such shifts are observable in peak positions indicate the formation of Ti^{3+} [38]. Specifically, the peak at 458.88 eV can be decomposed into two Gaussian peaks, which are attributed to $Ti^{3+}$and Ti^{4+}, respectively. Through peak-

2026 S.T. Yau High School Science Award 仅用于2026年丘成桐中学科学奖论文公示

fitting calculation, the Ti^{4+}/Ti^{3+} peak area ratios for the samples reduced at different temperatures were determined to be 1.86, 1.56, 1.25, and 1.18, respectively. This ratio shows a gradual decreasing trend with increasing calcination temperature, indicating that the reduction degree of Ti^{4+} to Ti^{3+} is progressively enhanced.

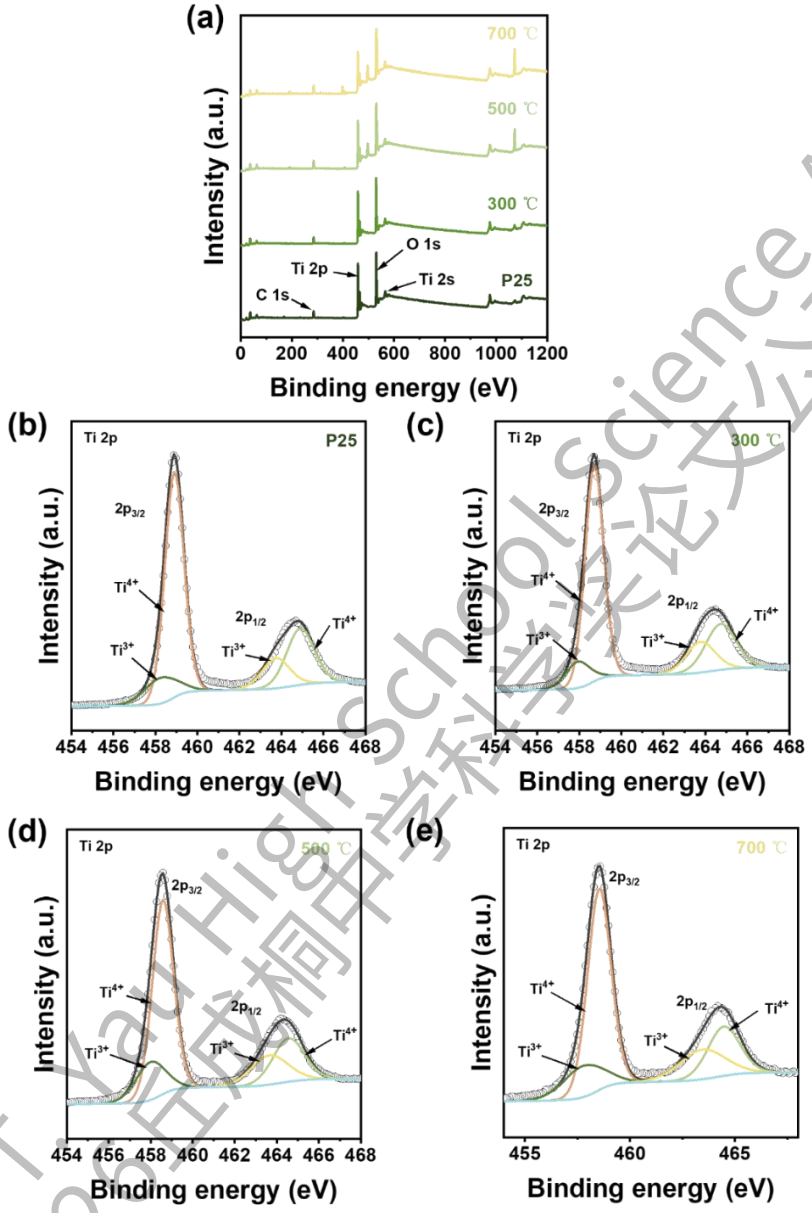


Fig. 12. (a) XPS survey. High-resolution Ti 2p spectra of (b) P25, (c) 300 ℃, (d) 500 ℃ and (e)700 ℃.

Fig. 13(b) presents the O 1s XPS spectrum. From the spectrum, it can be observed that the XPS peak of Ti-O shows a certain degree of shift. In addition, there is a broad O 1s peak at 532.18 eV, which is attributed to the Ti-OH bond. This result confirms the presence of hydroxyl groups on the surface of black TiO_2 after the reduction reaction. To confirm the presence of hydroxyl groups, further analysis was conducted on the O 1s XPS spectra of each sample, with the results shown in Figs. 13(c)-(f). The analysis revealed the emergence of two peaks at 530.18 eV and 531.78 eV. The peak at 530.18 eV corresponds to lattice

2026 S.-T. Yau High School Science Award 仅用于2026年丘成桐中学科学奖论文公示

oxygen (Ti^{4+}-O), while the other peak at 531.78 eV corresponds to oxygen vacancies (Ti^{3+}-O). Additionally, a new peak appeared at 532.68 eV in the calcined TiO_2, which corresponds to adsorbed oxygen or surface hydroxyl groups (Ti-OH). These results of XPS analysis indicate abundant oxygen vacancies and Ti^{3+} into surface of TiO_2 sample with $NaBH_4$ reduction under high temperature.

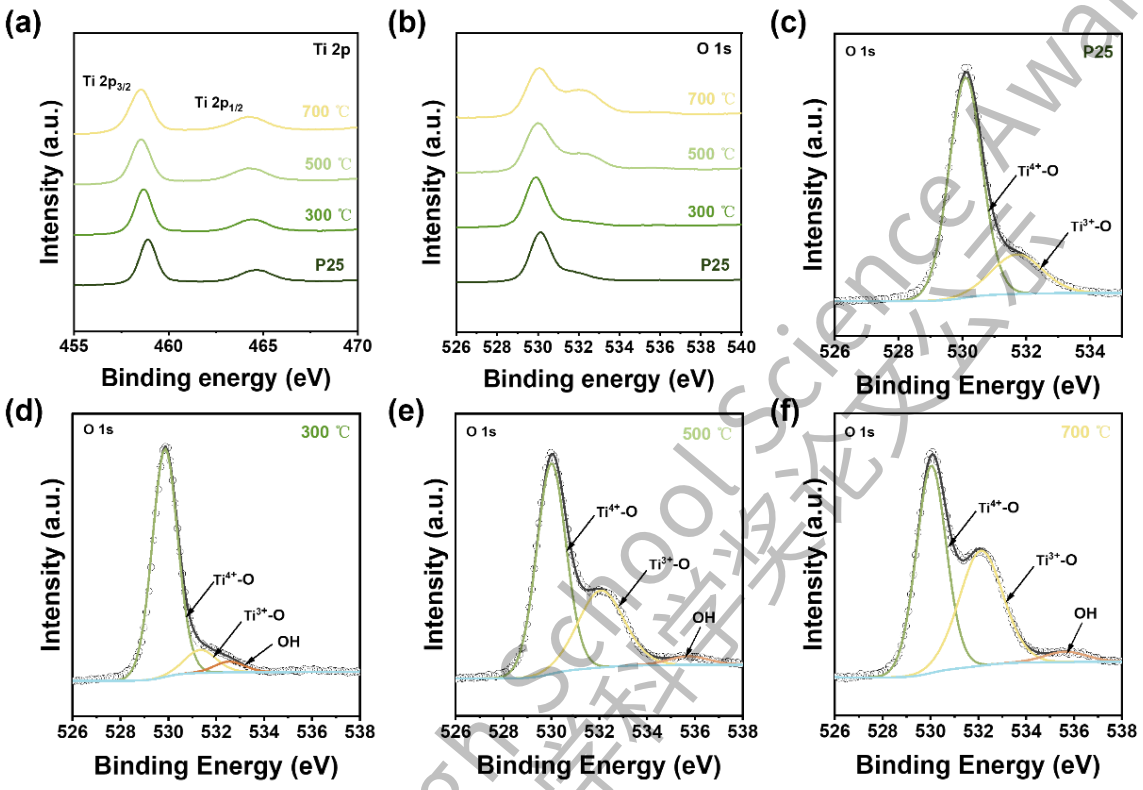


Fig. 13. (a) Ti 2p XPS spectra and (b) O 1s XPS spectra of the four TiO_2 samples. High-resolution O 1s spectra of (c) P25, (d) 300 °C, (e) 500 °C and (f) 700 °C.

To further investigate the effect of calcination temperature on the material structure, XRD analysis was performed on the samples reduced at different temperatures, and the results are shown in Fig. 14. For the pristine P25, the diffraction peaks at 25.3°, 38.0°, 48.1°, 53.9°, 54.8°, and 62.7° can be assigned to the (101), (004), (200), (105), (211), and (204) crystal planes of anatase-phase TiO_2 (JCPDS No. 21-1272), respectively. When the calcination temperature was 300 °C, the diffraction peak positions of the sample were essentially consistent with those of P25, indicating that the moderate-temperature reduction did not induce obvious phase changes. As the temperature increased to 500 °C, new diffraction peaks appeared at 27.3°, 36.2°, and 54.0° in addition to the anatase peaks. These peaks correspond to the (110), (101), and (211) crystal planes of rutile-phase TiO_2 (JCPDS No. 21-1276), suggesting that the anatase-to-rutile transformation had begun at this temperature. When the calcination temperature was further raised to 700 °C, the XRD pattern changed dramatically: the characteristic peaks of anatase almost completely disappeared and were replaced by numerous diffraction peaks attributable to Ti_3O_5 (JCPDS

2026 S. T. Yau High School Science Award
仅用于2026丘成桐中学科学奖公开

No. 76-1066) along with a small amount of TiO (JCPDS No. 72-0020) [39]. This indicates that TiO_2 was deeply reduced to low-valence titanium oxides, and the anatase phase was completely converted to the rutile phase.

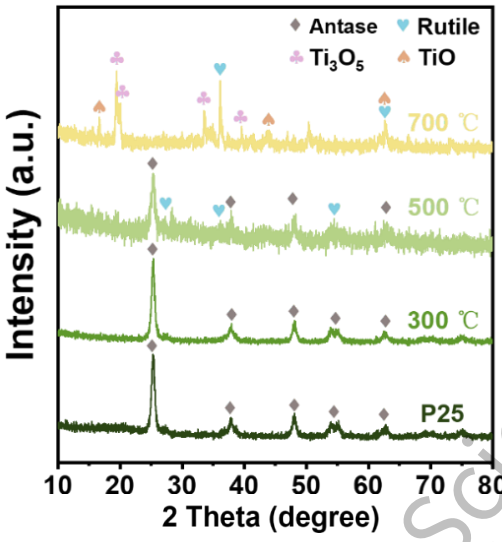


Fig. 14. XRD pattern.

SEM observations were performed on the samples treated at different temperatures, and the results are shown in Fig. 15. As the calcination temperature increased, the pristine P25 nanoparticles gradually agglomerated and grew. When the temperature reached 700 °C, obvious rod-like crystals appeared in the sample. This morphological evolution further confirms that when the calcination temperature exceeds 500 °C, a phase transformation occurs within the material, with the crystal structure gradually converting from the anatase phase to the rutile phase. Moreover, under the high-temperature reduction conditions at 700 °C, low-valence titanium oxides, namely Ti_3O_5 and a small amount of TiO, were further generated.

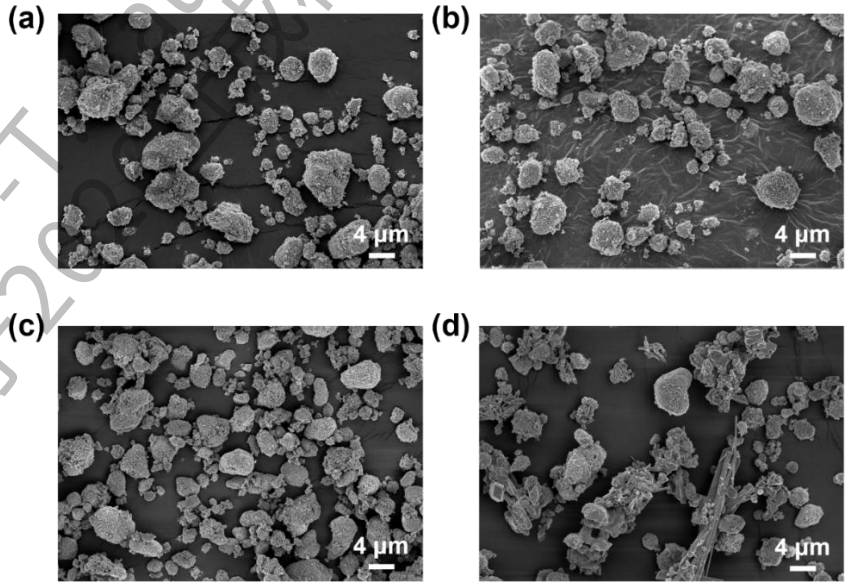


Fig. 15. SEM images of (a) P25, (b) 300 ℃, (c) 500 ℃ and (d) 700 ℃.

To further verify the phase transformation behavior occurring during the high-

2026 S. - Tsinghua High School Science Award
仅用于2026清华附中科学奖论文公示

temperature reduction process, TEM characterization was performed on samples treated at different temperatures, with the results shown in Fig. 16. As the calcination temperature increased, the pristine spherical particles gradually agglomerated. When the temperature reached 500 °C, amorphous flocculent edges appeared in some regions, and the particle boundaries became blurred, indicating the onset of lattice distortion, which is consistent with the initial stage of the anatase-to-rutile phase transition observed in the XRD results. At this stage, the lattice fringe spacing measured from High-resolution (HR-TEM) images is about 0.36 nm, which corresponds to the (101) plane of the anatase phase. When the temperature was further increased to 700 °C, the sample morphology changed significantly, with the spherical particles completely transforming into rod-like crystals. HR-TEM images revealed lattice fringes with a spacing of approximately 0.46 nm in the 700 °C sample, which can be indexed to the (110) crystal plane of the low-valence titanium oxide Ti_3O_5, confirming that the phase had been completely reconstructed into a new crystalline phase with good crystallinity. In addition, lattice fringes corresponding to the rutile phase were also observed in the 700 °C sample, with a spacing of approximately 0.34 nm, assigned to the (110) crystal plane of rutile (Fig.S2). Notably, this spacing is larger than the theoretical value (0.3247 nm) for the standard rutile (110) plane, which can be attributed to lattice distortion and expansion caused by a high concentration of oxygen vacancies. The TEM observations are in good agreement with the XRD analysis, confirming that upon calcination at 700 °C, a phase transition indeed occurred and low-valence titanium oxides were generated.

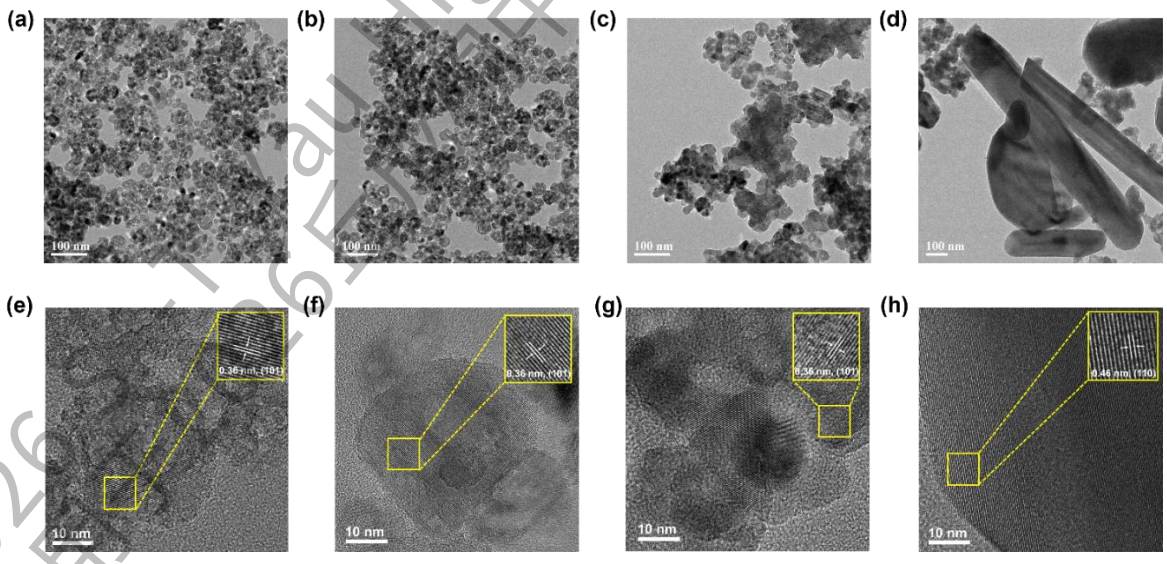


Fig. 16. TEM images of (a) P25, (b) 300 °C, (c) 500 °C and (d) 700 °C. HR-TEM images of (e) P25, (f) 300 °C, (g) 500 °C and (h) 700 °C.

Fig. 17(a) shows the Raman spectra of the four samples. The spectra of P25 and the sample reduced at 300 °C are consistent, exhibiting characteristic peaks of anatase TiO_2 at

147 cm^{-1}, 401 cm^{-1}, 516 cm^{-1}, and 641 cm^{-1}, indicating that low-temperature reduction did not induce phase transformation. The 500 °C sample still retains the anatase peak shape, but the strongest peak (144 cm^{-1}) blue-shifts to 154 cm^{-1} with a slight increase in peak width, and the intensities of the other peaks decrease, suggesting that the sample is in the initial stage of the anatase-to-rutile phase transition, accompanied by lattice distortion caused by a small number of oxygen vacancies. For the 700 °C sample, the characteristic anatase peaks completely disappear and are replaced by a set of broadened new peak groups, indicating that at this temperature, TiO_2 is deeply reduced and the phase reconstructs into low-valence titanium oxides such as Ti_3O_5, while the high concentration of oxygen vacancies, Ti^{3+} and lattice disorder lead to significant broadening of the Raman peaks.

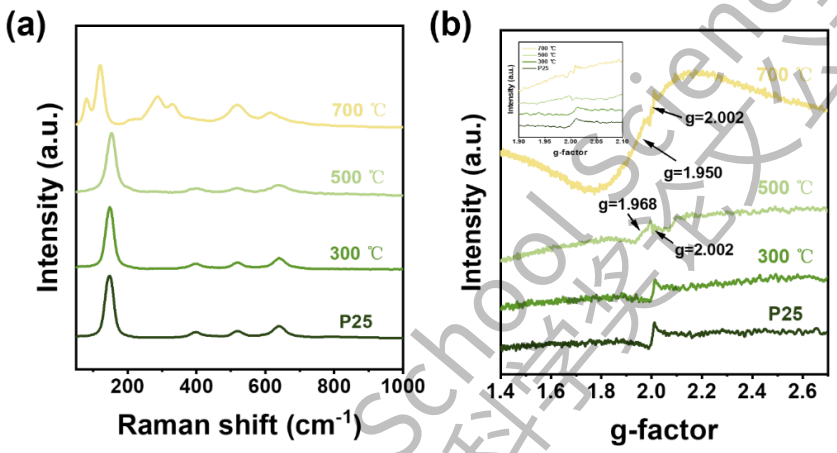


Fig. 17. (a) Raman spectroscopy of P25, 300 °C, 500 °C and 700 °C. (b) EPR spectra of P25, 300 °C, 500 °C and 700 °C, the inset shows a magnified view of the region within g =1.90 to 2.10.

Fig. 17(b) shows the EPR spectra of pristine P25 and $NaBH_4$ reduced TiO_2 samples at different temperatures. No obvious paramagnetic signal is detected for pristine P25 or the sample reduced at 300 °C. As shown in the magnified view in Figure 11b, the weak sharp peaks observed in both samples are attributable to instrument baseline noise or paramagnetic impurities, rather than characteristic signals of oxygen vacancies or Ti^{3+}. When the reduction temperature is increased to 500 °C, the EPR signal intensifies significantly, with the peak at g≈2.002 assigned to oxygen vacancies and the peak at g≈1.968 corresponding to Ti^{3+} species. Upon further increasing the temperature to 700 °C, the signal becomes substantially stronger, accompanied by pronounced broadening and an asymmetric line shape. These features are closely related to the deep reduction of the sample, the generation of abundant Ti^{3+} and oxygen vacancies, and the strong magnetic interactions arising from structural reconstruction into low-valence titanium oxides (Ti_3O_5), which is fully consistent with the Raman and XRD results. The significant increase in EPR linewidth further indicates a markedly higher local density of unpaired electrons and enhanced spin-spin coupling, confirming that treatment at 700 °C introduces high concentrations of oxygen

2026 S.-T. Yau High School Science Award
仅用于2026丘成桐中学科学奖公示

vacancies and Ti^{3+} defects into TiO_2.

The surface temperature of TiO_2-700 membrane under 1 sun as shown in Fig. 18 under 1 sun for 0 min, 60 min, 70 min, 80 min and 90 min, respectively.

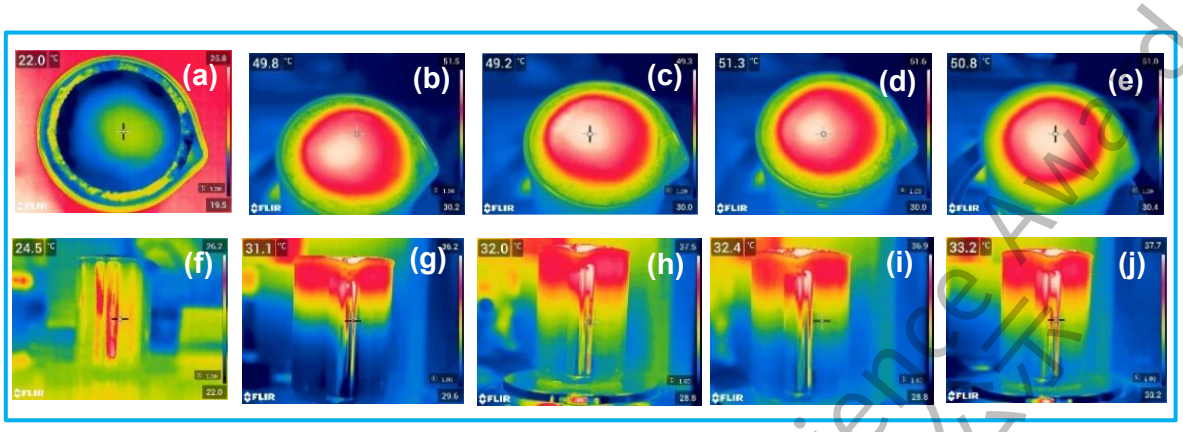


Fig. 18. Infrared radiation images of TiO_2-700 membrane surface under 1 sun (a)-(f) 0 min, (b)-(g) 60 min, (c)-(h) 70 min, (d)-(i) 80 min, (e)-(j) 90 min.

The evaporation performance of four different samples was evaluated to validate the ML prediction. To quantify solar water evaporation of the four different samples, we recorded the mass change versus time under different light densities of one sun, two sun, three sun, and four sun, respectively. Fig. 19(a)-(b) show the variations in mass and surface temperature over time for pure water and four different TiO_2 membranes under one sun.

Fig. 19(c) compares the measured ER and EE data for four different samples under one sun. The ER values of pure water, P25, 300 ℃, 500 ℃ and 700 ℃ are 0.22, 0.41, 0.45, 0.86 and 1.12 kg m^{-2} h^{-1}, respectively, with corresponding EE values of 10.9%, 21.8%, 23.2%, 49.7% and 67.8%. Comparing the 500 ℃ and 700 ℃ with 300 ℃, ER and EE achieves a significantly improvement, primarily attributed to the bandgap of samples decreases with increasing of calcined temperature, and it can absorb light across the full spectrum, with a stronger light absorption capacity. Overall, for the four different samples, the consistency between the experimental correlation of these factors and the ML results confirm BG plays a pivotal role in determining the EE.

To further verify the influence of sun on evaporation performance, we tested ER and EE of the membrane with 700 ℃ under different irradiance levels. Fig. 19(d) displays water mass loss of the 700 ℃ sample exhibits a continuous increase with rising sun. Fig. 19(e) shows the surface temperature fast increases at initial time, eventually reaching saturation and keep stability with prolonging the irradiation time. Fig. 19(f) presents the ER values of 700 ℃ evaporator are 1.12, 1.62, 1.89 and 2.28 kg m^{-2} h^{-1} under one sun, two sun, three sun, and four sun, respectively, with corresponding EE values of 67.8%, 51.2%, 40.7% and 37.7%. Which manifests the four sun have significant enhancement of ER for 700 ℃

2026 Siu Yau High School Science Award 仅用于2026年度科学奖论文公示

evaporator. Obviously, EE exhibits a gradual decline with increasing sun. This phenomenon is primarily ascribed to the inherently limited light absorption of the material calcined at 700 °C, which restricts the maximum surface temperature increment. Therefore, the evaporation rate fails to grow linearly with increasing light intensity, demonstrating a typical nonlinear behavior. In addition, elevated heat loss via thermal radiation and convection occurs under high illumination, accompanied by severe interfacial overheating. These combined effects deteriorate the photothermal conversion performance and gradually reduce the evaporation efficiency [40-42].

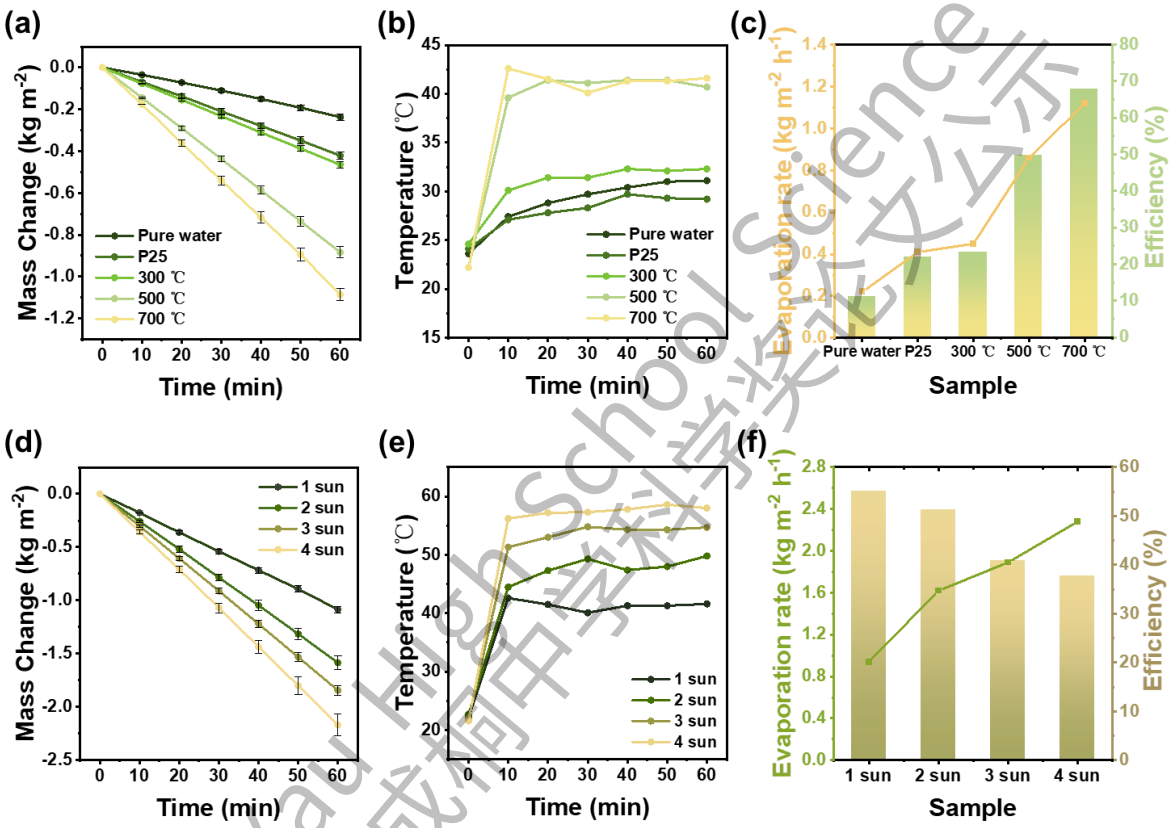


Fig. 19. (a) Water mass change for Pure water and four samples (1 sun). (b) Surface temperature profiles for Pure water and four membranes (1 sun). (c) Comparative analysis of ER and EE for Pure water and four membranes (1 sun). (d) Water mass change with 700 ℃ sample under various irradiance levels. (e) Changes in surface temperature of 700 ℃ sample under various irradiance levels. (f) The ER and EE values of 700 ℃ sample under different simulated intensities of sunlight.

Additionally, as described in Table 1, the 700 ℃ demonstrated a higher water evaporation performance superior to those of previously reported in the literature. This indicates that narrower BG promotes the adsorption of sunlight. Collectively, research results verify the distinct yet significant influences of material feature A on the ER and BG on the EE of SPMs. It confirming the SHAP analysis and establishing the validity of material performance forecasts by the ML methodologies employed.

2026 S.T. Yau High School Science Award
仅用于2026丘成桐中学科学奖公示

Table 1. Comparison of evaporation performance of different semiconductor materials

Semiconductor materials	Solar intensity (kw m^{-2})	Evaporation Efficiency (%)	Evaporation rate (kg m^{-2} h^{-1})	Refs.
Black TiOx	1	50.3	0.8	[6]
Fe_3O_4 MNPs	1	63	0.9	[8]
Ti_3C_2	1	53.7	0.45	[10]
Fe_3O_4/C	1	67	1.07	[26]
WO_3	1	57.3	1	[38]
Cu_2-xS NWs	1	67	0.95	[49]
TiN	1	64.1	0.94	[59]
TiO_2 NWAs	1	59	1.25	[68]
TiO_2-700 ℃	1	67.8	1.12	This work
Fe3O4 MNPs	2	65	1.85	[8]
MoS_2	2	61	1.95	[21]
Cu-Fe_3O_4	2	36	1.2	[83]
Cu-Fe_3O_4	2	42	1.3	[83]
TiO_2-700 ℃	2	51.2	1.62	This work

5. Conclusions

In this study, ML models VR, XGBT, GBDT and SVM were employed to accurately forecast the evaporation rate and evaporation efficiency of SPMs. As the limited sample size in the current dataset (n=196), we applied Bayesian optimization and treated cross-validation as the objective function, minimizing the number of cross-validation evaluations, reducing overfitting to a single validation split. Moreover, the correlations between critical features and evaporation performance were investigated. The VR model achieves the highest R^2 for ER on test sets (R^2=0.9337). However, the R^2 of EE prediction is only moderate (R^2=0.724), suggests that the current dataset may not fully capture the complex factors governing evaporation, expanding the dataset samples to enhance model performance in the future work. SHAP value analysis identified sun and TC as the most significant factors for ER, while BG was the most influential for EE. The results highlight BG as critical parameters deserving focused attention in the future study and design of novel SPMs. Four modified TiO_2 samples prepared by utilizing hydrogen-induced reduction reaction to reduce the band-gap of TiO_2 under different high temperature. To evaluate the prediction performance of the ML models, solar evaporation tests of four different samples were tested under different light density of one sun, two sun, three sun, and four sun, respectively. The characterization and evaporation experimental results show that when the

2026 S. T. You High School Science Award
仅用于2026丘成桐中学科学奖论文公示

calcination temperature reaches 700 °C, a phase transition occurs within the P25 sample, generating low-valence titanium oxides, with Ti^{4+} reduced to Ti^{3+} accompanied by the formation of oxygen vacancies. These changes lead to bandgap narrowing and enhanced light absorption, thereby effectively improving the water evaporation capability. This result further validates the reliability and applicability of the machine learning (ML) method. This study highlights the promise of ML approaches in innovating design strategies for maximizing energy conversion, and optimizing material features and operational conditions in photothermal evaporation technologies. Overall, this work presented a ML approach to support the development of SPMs, establishing a predictive framework between material characteristics and evaporation performance for material design.

References

[1] B. Yang, Z.M. Zhang, P.T. Liu, et al., Flatband λ-Ti_3O_5 towards extraordinary solar steam generation, Nature 622 (2023) 499-+.

[2] H. Wang, X. T. Xu, X. Y. Gao, Y. Q. Li, T. Lu, L. K. Pan, Design of three-dimensional faradic electrode materials for high-performance capacitive deionization, Coord. Chem. Rev. 510 (2024) 215835.

[3] W. Zhang, Y.Z. Chen, Q.H. Ji, Y.Y. Fan, G. Zhang, X. Lu, C.Z. Hu, H.J. Liu, J.H. Qu, Assessing global drinking water potential from electricity-free solar water evaporation device, Nat. Commun. 15 (2024) 6784.

[4] X. Wu, Y. Lu, X.H Ren, P. Wu, D.W Chu, X.F. Yang, H.L. Xu, Interfacial solar evaporation: from fundamental research to applications, Adv. Mater. 36 (2024) 2313090.

[5] A. Selvam, G. Jain, R.G. Chaudhuri, M.K. Mandal, S. Chakrabarti, Avant garde solar-thermal nanostructures: nascent strategy into effective photothermal desalination, Sol. RRL 6 (2022) 2200321.

[6] H. Abdullah, H. Shuwanto, N.A. Astarini, J. Lie, R.T. Ginting, M.L. Tsai, S.J. Shih, M. Sillanpää, A brief review of emerging strategies in designing interfacial solar steam generation for desalination, water purification, power generation, and sea farming, ACS Appl. Energy Mater. 8 (2025) 2663-2704.

[7] Srishti, K. Khandelwal, A. Kumar, A. Sinhamahapatra, Progress on TiO_2-based materials for solar water interfacial evaporation, Front. in Chem. Eng. 4 (2022) 1046019.

[8] Y.Q. Sun, X.Y. Tan, X. Yuan, J. Li, Solar-driven interfacial evaporation: Research advances in structural design, Chem. Eng. J. 495 (2024) 153316.

[9] M. Yu, X. Li, S.T. Jia, H.Y. Wang, C.C. Li, D.Y. Bao, W.Y. Tian, P. Wang, J. Guo, L. Huang, R. Li, Z. Zhang, Solar-driven interfacial evaporators from traditional microstructures to smart microstructures, Sep. Purif. Technol. 369 (2025) 133099.

2026 Siyou High School Science Award 仅用于2026年成都市学科学论文公示

[10] J.W. Liang, D.Q. Fan, C.K. Cai, Z.Y. Su, Y. Lu, Synergistic structural alignment and interfacial switching enable fast and low-energy water release in solar-driven atmospheric water harvesting, Adv. Funct. Mater. 36 (2026) e18960.

[11] Y. Lu, Z.Z Li, D.Q. Fan, X.L. Xu, H.L. Xu, X.F. Yang, Scalable solar-tracking ion-selective evaporation system for concurrent desalination and radionuclide removal, Chem. Eng. J. 528 (2026) 172221.

[12] S.Y. Xia, X.A. Wang, D.Q. Fan, Y.Y. Li,, Y. Lu, Electrospun bilayer nanofibrous membranes for solar-driven water purification with CO2 mitigation, J. Environ. Chem. Eng. 13 (2025) 120366.

[13] I. Ibrahim, D.H. Seo, A.M. McDonagh, H.K. Shon, L. Tijing, Semiconductor photothermal materials enabling efficient solar steam generation toward desalination and wastewater treatment, Desalination, 500 (2021) 114853.

[14] Y. Xia, Q. Hou, H. Jubaer, Y. Li, Y. Kang, S. Yuan, H. Liu, M.W. Woo, L. Zhang, L. Gao, Spatially isolating salt crystallization from water evaporation for continuous solar steam generation and salt harvesting, Energy Environ. Sci. 12 (2019) 1840-1847.

[15] H.W. Liu, C.J. Chen, H. Wen, R.X. Guo, N.A. Williams, B.D. Wang, F.J. Chen, L.B. Hu, Narrow bandgap semiconductor decorated wood membrane for high efficiency solar-assisted water purification, J. Mater. Chem. A 6 (2018) 18839.

[16] H.L. Liu, F. Wu, X.Y. Liu, J.Y. Yu, Y.T. Liu, B. Ding, Multiscale synergetic bandgap/structure engineering in semiconductor nanofibrous aerogels for enhanced solar evaporation, Nano Letters, 23 (2023) 11907-11915.

[17] J. Wang, Y.Y. Li, L. Deng, N.N. Wei, Y.K. Weng, S. Dong, D.P. Qi, J. Qiu, X.D. Chen, T. Wu, High-performance photothermal conversion of narrow-bandgap Ti_2O_3 nanoparticles, Adv. Mater. 29 (2017) 1603730.

[18] X. Lin, P. Wang, R. Hong, X. Zhu, Y. Liu, X. Pan, X. Qiu, Y. Qin, Fully lignocellulosic biomass-based double-layered porous hydrogel for efficient solar steam generation, Adv. Funct. Mater. 32 (2022) 2209262.

[19] Merchant, A., Batzner, S., Schoenholz, S.S. et al. Scaling deep learning for materials discovery. Nature, 624 (2023) 80-85.

[20] G. Bai, Z.H. Wang, H. Wang, C.L. Wang, XJ. Liu, H.B. Li, G. Yang, M. Xu, L.K. Pan, Machine learning-guided exploration of carbon-based photothermal materials for solar evaporation, Sep. Purif. Technol. 373 (2025) 133627.

[21] G.S. Xu, Y. Li, J.F. Li, J.L. Li, X.J. Liu, C.L. Wang, W.J. Mai, G. Yang, L.K. Pan, Toward stable zinc anode: an AI-assisted high-throughput screening of electrolyte additives for aqueous zinc-ion battery, Angew. Chem. Int. Ed. 64 (2025) e202511389.

[22] C. Gao, X. Min, M. Fang, T. Tao, X. Zheng, Y. Liu, X. Wu, Z. Huang, Innovative materials science via machine learning, Adv. Funct. Mater. 32 (2022) 2108044.

[23] H. Wang, Y. Q. Li, X. Y. Xuan, K. Wang, Y. F. Yao, L. K. Pan, Machine learning accelerated discovery of covalent organic frameworks for environmental and energy applications, Environ.

2026 S-T Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

Sci. Technol. 59 (2025) 6361-6378.

[24] G. Xu, M. Jiang, J. Li, X. Xuan, J. Li, T. Lu, L. Pan, Machine learning-accelerated discovery and design of electrode materials and electrolytes for lithium ion batteries, Energy Storage Mater. 72 (2024) 103710.

[25] H. Wang, Y. Li, Y. Liu, X. Xu, T. Lu, L. Pan, Advancement of capacitive deionization propelled by machine learning approach, Sep. Purif. Technol. 354 (2025) 129423.

[26] L.A. Yang, H. Wang, D.Y. Leng, S.P. Fang, Y.N. Yang, Y.R. Du, Machine learning applications in nanomaterials: Recent advances and future perspectives, Chem. Eng. J. 500 (2024) 156687.

[27] K. Guo, Z.Z. Yang, C.H. Yu, M.J. Buehler, Artificial intelligence and machine learning in design of mechanical materials, Mater. Horiz. 8 (2021) 1153-1172.

[28] H. Wang, Y. Zhu, K. Han, C.L. Wang, J.F. Li, Y.Q. Li, Y. Liu, G. Yang, L.K. Pan, A data-driven machine learning approach for interpretable prediction of desalination stability of carbon materials for capacitive deionization, J. Mater. Chem. A. 376 (2025) 134181.

[29] Z.A. Ali, Z.H. Abduljabbar, H.A. Tahir, A.B. Sallow, S.M. Almufti, Exploring the power of eXtreme gradient boosting algorithm in machine learning: a Review, Academic Journal of Nawroz University, 12 (2023) 320-334.

[30] A.H. Elsheikh, S. Shanmugan, R. Sathyamurthy, et al. Low-cost bilayered structure for improving the performance of solar stills: Performance/cost analysis and water yield prediction using machine learning, Sustain. Energy Technol. Assess. 49 (2022) 101783.

[31] D. Ariyanti, L. Mills, J.Z. Dong, Y. Yao, W. Gao, $NaBH_4$ modified TiO_2: Defect site enhancement related to its photocatalytic activity, Mater. Chem. Phys. 199 (2017) 571-576.

[32] P.J. Ying, M. Li, F.L. Yu, Y. Geng, L.Y. Zhang, J.J. He, Y.J. Zheng, R. Chen, Band Gap Engineering in an Efficient Solar-Driven Interfacial Evaporation System, ACS Appl. Mater. Interfaces 12 (2020) 32880-32887.

[33] L. Zhou, X.Q. Li, G.W. Ni, S. Zhu, J. Zhu, The revival of thermal utilization from the Sun: interfacial solar vapor generation, Nati. Sci. Rev. 6 (2019) 562-578.

[34] X. Wu, G.Y. Chen, G. Owens, D.W. Chu, H.L. Xu, Photothermal materials: A key platform enabling highly efficient water evaporation driven by solar energy, Mater. Today energy 12 (2019) 277-296.

[35] J. Li, E.H. Wu, Z. Xu, J. Hou, W.J. Peng, H. Li, X. Li, Superior photothermal conversion performance of black titanium-based materials, iScience 28 (2025) 112188.

[36] S.J. Guan, Y.L. Cheng, L. Hao, H. Yoshida, C. Tarashima, T.Z. Zhan, T. Itoi, T.B. Qiu, Y. Lu, Oxygen vacancies induced band gap narrowing for efficient visible-light response in carbon-doped TiO_2, Sci. Rep. 13 (2023) 14105.

[37] L. Phuthu, R.S. Dima, N.E. Maluta, J.K. Kirui, R.R. Maphanga, DFT study of TiO_2 brookite (210) surface doped with silver and molybdenum, Mater. Res. Express 9 (2022) 095901.

[38] M.F. Rahman, Z.R. Melody, M.H. Ali, A. Ghosh, P. Barman, M.R. Islam, M.K. Hossain, First-principles analysis of how Cobalt doping affects the structural, electronic, and optical properties

2026 S.T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

of α-MoO_3, Indian J. Phys. 98 (2024) 2695-2704.

[39] R. Ricka, A. Wanag, E. Kusiak-Nejman, M. Reli, M.F. Edelmannova, M. Lapinski, G. Slowik, A.W. Morawski, K. Koci. Effect of $NaBH_4$ loading and reduction temperature on defect-driven CO_2 photoreduction over TiO_2, Appl. Surf. Sci. Adv. 31 (2026) 100925.

[40] Y. Su, L. Liu, X.C. Gao, W. Yu, Y. Hong, C. Liu, A high-efficient and salt-rejecting 2D film for photothermal evaporation, iScience. 26 (2023) 107347.

[41] S.D. Zhang, H.X. Ma, D. Guo, P. Guo, J.J. Wang, M.Q. Liu, S. Wu, C.L. Bao, Multiscale preparation of graphene oxide/carbon nanotube-based membrane evaporators by a spray method for efficient solar steam generation, ACS Appl. Nano Mater. 5 (2022) 7198-7207.

[42] C. Xue, C. Qi, W.Z. Li, J. Lian, A novel rGO-Fe_3O_4-mixed cellulose ester membrane for solar evaporation and seawater desalination, Sol. Energ. Mat. Sol. C. 290 (2025) 113708.

2026 S.-T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

Acknowledgements

This research project was completed under the guidance of my Yingcai Program mentor, Professor Likun Pan, and granted the honor of "Outstanding Scholar" (Physics). I would like to sincerely thank to Professor Pan for guiding in research routes, and applying machine learning to materials development, significantly reducing the time and cost associated with trial-and-error experimentation in new material testing. For the research topic, Professor Pan not only fully respected my interests but also provided crucial and timely guidance and suggestions on specific research orientations. I would also like to thank Dr. Zihui Wang for her dedicated guidance on experimental directions, and assistance in instrument operation.

I am sincerely grateful to Shanghai government for supporting this Yingcai Program, which has given me the opportunity to engage in in-depth scientific research under the guidance of scientists. I am also thank my school for support me to research activities. I sincerely appreciate my parent for their strong support for my scientific exploration.

I declare: This research report and thesis were not produced using AI.

My project research under mentor guidance

- **2025-05: Literature review**
  - Review the literatures on photothermal materials for solar water evaporation.
  - Searching literatures on machine learning in material development.
- **2025-06～08: Collect dataset and machine learning model**
  - Collected dataset (six input and two output features) from published literatures researched on semiconductor photothermal materials for solar water evaporation.
  - Written the code for four machine learning model (GBDT, SVM, XGBT, VR) program.
  - Train ML models and evaluate performance by RMSE, MAE and R^2 values.
- **2025-09～10: Prepare samples and experimental test**
  - Prepare black TiO_2 samples by calcining mix of P25 and $NaBH_4$ under temperature 300°C, 500°C, 700°C, respectively.
  - Samples material characteristic parameter testing.
  - Evaporation performance test under different solar irradiation.
- **2025-11: Prepared report**
  - Analyzed the experimental results.
  - Written the project report.
- **2026-01～06: Improved the project**
  - Expanded the database, and optimized the machine learning model.
  - Repeated the experiment and plot the error bars.

2026 S.T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

A. Author CV

Ao Wang, Grade 12

- GPA: 1/165, awarded Merit Student honor every year since 2020.
- Participated in Research Project at MIT, 2026.06-2026.08.
- 2026 ISEF (USA), Second Prize Award, Non-Trivial Ventures Special Award, 2026.05.
- Participated in Stanford University Mathematics Camp(Residential Program), 2025.06-07.
- Offer letter of 2026 Stanford University Mathematics Camp (P2, Residential Program).
- Participated in 2025 China Science Talent Program (Yingcai, Physics), Awarded Outstanding Scholar.
- Ao Wang, Peter Lu*, Oreoluwa Alao. Paper accepted by IEEE-ICDM 2026.
- Ao Wang, Zihui Wang*, Likun Pan*. Sep. Purif. Technol., 2026, 400: 138440.
- Ao Wang, Likun Pan. Computer Software Copyright Registration Certificate, 17299658.
- Participated in 2024 Ross Mathematics Program (Indiana), 2024.06-08.
- BPhO Top Gold 2025 (Top 0.2%), Physics Bowl Top Gold 2025 (Global Top 63 equal).
- Awarded First prize, Research practice for Shanghai high school students, 2024.12.
- 2024 USAMO Qualifier, 2023 AMC12 (Top 1%).
- 2022 USAJMO Qualifier, 2021 AMC10 (Top 1%).

B. Mentor CV

Likun Pan, Professor, Doctoral advisor, School of Physics, East China Normal University.

He graduated from Fudan University with a bachelor's degree in Applied Chemistry, from the Acoustics Research Institute of the Chinese Academy of Sciences with a master's degree, and from Nanyang Technological University in Singapore with a PhD in Microelectronics, in 1997, 2001 and 2004, respectively. He is a globally highly cited scientist by Clarivate Analytics and an outstanding Chinese researcher by Elsevier. His current research focus is on the design and application of functional materials enabled by artificial intelligence. He has published over 500 SCI papers, with 40,000 citations and an H-index of 115. Currently, he serves as the deputy editor-in-chief of Korean Journal of Chem. Eng., and as the editorial board of Journal of Colloid and Interface Science, etc.

2026 S.T. Yau High School Science Award
仅用于2026丘成桐中学科学奖公示

Appendix A: Supplementary Information (data collection and testing)

Table S1. Collected dataset from previous works.

EE %	ER kg/m^2 h	BG eV	A %	TC W/mK	Sun kw/m^2	T ℃	H %	Ref
81	1.4	1.3	0.959	0.02	1	23	90	[1]
87	2.8	1.3	0.959	0.02	2	23	90	
96	4.3	1.3	0.959	0.02	3	23	90	
96	3.68	1.3	0.838	0.02	3	23	90	
96	3.05	1.3	0.765	0.02	3	23	90	
80.5	1.3	1.2	0.93	0.06	1	26	75	[2]
83.5	2.8	1.2	0.93	0.06	2	26	75	
91	4.18	1.2	0.93	0.06	3	26	75	
91	5.65	1.2	0.93	0.06	4	26	75	
91	7.03	1.2	0.93	0.06	5	26	75	
67.7	1.15	0.45	0.99	0.525	1	23	55	[3]
70	2.2	0.45	0.99	0.525	2	23	55	
76.5	3.6	0.45	0.99	0.525	3	23	55	
82	5.1	0.45	0.99	0.525	4	23	55	
86.2	6.6	0.45	0.99	0.525	5	23	55	
84.5	1.46	/	0.9225	0.02	1	25	50	[4]
88.78	6.8	/	0.9225	0.02	5	25	50	
88	1.32	/	0.95	0.05	1	20	58	[5]
50.3	0.8	/	0.913	/	1	21	55	[6]
76.7	1.229	1.82	0.97	/	0.5	25	50	[7]
85.6	1.255	1.82	0.97	/	1	25	50	
86.5	2.52	1.82	0.97	/	2	25	50	
63	0.9	0.3	0.95	0.48	1	23	55	[8]
65	1.85	0.3	0.95	0.48	2	23	55	
67	3.1	0.3	0.95	0.48	3	23	55	
70	4.5	0.3	0.95	0.48	4	23	55	
71	5.6	0.3	0.95	0.48	5	23	55	
68.6	1.1	0.3	0.95	0.48	1	23	55	
71	2.25	0.3	0.95	0.48	2	23	55	
73	3.4	0.3	0.95	0.48	3	23	55	
75	4.9	0.3	0.95	0.48	4	23	55	
76.5	6.2	0.3	0.95	0.48	5	23	55	
75	1.35	0.3	0.95	0.48	1	23	55	
78	2.75	0.3	0.95	0.48	2	23	55	
80	4	0.3	0.95	0.48	3	23	55	
82	5.3	0.3	0.95	0.48	4	23	55	
83	6.7	0.3	0.95	0.48	5	23	55	
70.8	0.63	/	0.94	0.32	0.6	25	60	[9]
76.9	0.92	/	0.94	0.32	0.8	25	60	
81.4	1.22	/	0.94	0.32	1	25	60	

2026 S. T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

86.2	1.56	/	0.94	0.32	1.2	25	60	
53.7	0.45	/	0.868	/	1	25	50	[10]
85.8	1.44	/	0.94	/	1	25	50	
90	2.56	/	0.94	/	2	25	50	
90	3.85	/	0.94	/	3	25	50	
88	5	/	0.94	/	4	25	50	
92	1.48	3.2	0.944	/	1	21	50	[11]
92	1.48	3.2	0.944	/	1	21	50	[12]
90.4	1.43	2	0.9	/	1	25	50	[13]
91.7	2.91	2	0.9	/	2	25	50	
93,3	5.93	2	0.9	/	4	25	50	
68.6	1.09	2	0.85	/	1	25	50	[14]
75.1	2.38	2	0.85	/	2	25	50	
81.2	5.15	2	0.85	/	4	25	50	
63.9	1.01	2	0.93	0.067	1	26	/	[15]
94.9	1.63	/	0.95	0.04	1	24	25	[16]
88.4	4.04	/	0.95	0.04	3	24	25	
97.5	7.16	/	0.95	0.04	5	24	25	
86.6	1.657	/	0.97	0.06	1	/	/	[17]
75.7	0.81	1	0.96	/	0.76	25	/	[18]
59.3	0.64	1	0.96	/	0.76	25	/	
81.4	6.15	1	0.96	/	5.35	25	/	
62.6	4.73	1	0.96	/	5.35	25	/	
94	1.95	/	0.98	0.03	1	/	/	[19]
91.5	6.6	/	0.95	0.06	5	25	/	[20]
80	1.27	/	0.9615	0.037	1	20	60	[21]
61	1.95	/	0.9615	0.037	2	20	60	
67	5.9	/	0.9615	0.037	5.5	20	60	
92.1	1.32	0.1	0.925	0.02	1	20	55	[22]
95.3	5.03	0.1	0.925	0.02	5	20	55	
77.8	1.24	/	0.837	/	1	25	40	[23]
84.45	1.23	/	0.9627	/	1	25	55	[24]
94	1.45	/	0.9	0.243	1	26	50	[25]
67	1.07	/	0.903	/	1	25	45	[26]
86.9	1.28	2.5	0.906	0.02	1	/	/	[27]
92.4	1.81	/	/	0.035	1	20	50	[28]
/	4.63	/	/	0.035	3	20	50	
91.5	1.337	1.7	/	0.03	1	27	50	[29]
93.5	3.99	1.7	/	0.03	3	27	50	
94.1	6.78	1.7	/	0.03	5	27	50	
77.1	1.2	2.7	0.99	/	1	22	40	[30]
78.5	1.25	2.43	0.879	/	1	22	40	[31]
/	0.69	3.11	0.634	/	1	22	40	

2026 S. -T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

93.8	1.22	1.7	/	0.03	1	/	/	[32]
89.5	1.86	2.76	0.948	0.145	1	28	30	[33]
72.5	4.61	2.76	0.948	0.145	3	28	30	
69.4	7.75	2.76	0.948	0.145	5	28	30	
87.5	3.25	2.4	0.657	/	5	/	/	[34]
93.8	4.4	1.1	0.834	/	5	/	/	
77.5	1.24	/	0.902	0.3	1	23	55	[35]
73.7	0.53	/	0.895	/	1	/	/	[36]
69.1	/	/	0.9	0.04	1	25	55	[37]
57.3	1	2.68	0.95	/	1	25	30	[38]
74.5	1.3	2.35	0.95	/	1	25	30	
83.1	1.45	2.35	0.99	/	1	25	30	
79.9	1.39	2.35	0.99	/	1	25	30	
81.6	1.86	0.7	/	0.08	1	24	55	[39]
97.7	7.5	0.7	/	0.08	1	24	55	
80.4	/	1.79	0.95	0.069	0.5	32	45	[40]
94.5	3.12	1.79	0.95	0.069	1	32	45	
89.2	4.2	1.79	0.95	0.069	1.5	32	45	
88.3	5.38	1.79	0.95	0.069	2	32	45	
95.9	2.2	/	0.98	0.77	1	25	60	[41]
95.6	2.09	/	/	/	1	25	25	[42]
87	2.2	/	0.93	0.39	1	/	/	[43]
91.2	1.57	/	0.97	0.55	1	24	25	[44]
94.5	1.5	/	0.98	/	1	25	60	[45]
91	1.09	2	0.95	/	1	25	40	[46]
91	2.38	2	0.95	/	2	25	40	
91	5.15	2	0.95	/	4	25	40	
98.5	1.79	/	0.983	0.71	1	25	60	[47]
85.1	1.355	/	0.94	0.06	1	25	31	[48]
91.4	2.911	/	0.94	0.06	2	25	31	
97.5	4.659	/	0.94	0.06	3	25	31	
96.6	6.152	/	0.94	0.06	4	25	31	
95.9	7.643	/	0.94	0.06	5	25	31	
67	0.95	1.27	/	/	1	26	23	[49]
79	4.31	1.27	/	/	4	26	23	
89.9	9.29	1.27	/	/	8	26	23	
97.1	1.546	0.85	/	/	1	/	/	[50]
90	/	0.85	/	/	7.6	/	/	
88	/	0.85	/	/	9.84	/	/	
90	/	0.85	/	/	14.5	/	/	
99	1.43	1.54	/	0.043	1	/	/	[51]
77.3	1.71	/	0.65	0.04	1.5	27	/	[52]
66	2.33	/	0.65	0.04	2.7	27	/	

2026 S. T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

58	2.61	/	0.65	0.04	3.4	27	/	
91.5	1.92	/	/	0.04	1	27	/	[53]
53	2.45	2.65	/	/	1	25	50	[54]
67	3.07	2.65	/	/	1	25	50	
81.4	3.81	2.65	/	/	1	25	50	
18.5	0.259	2.6	0.9	/	1	25	/	[55]
28.7	0.41	2.6	0.99	/	1	25	/	
33	0.48	2.6	0.99	/	1	25	/	
41.3	0.578	2.6	0.99	/	1	25	/	
78.7	1.101	2.6	0.99	/	1	25	/	
82.5	1.4	2.35	0.94	0.4	1	15	12	[56]
87	2.59	2.35	0.94	0.4	2	15	12	
94.5	5.51	2.35	0.94	0.4	4	15	12	
92	7.98	2.35	0.94	0.4	6	15	12	
89.5	9.12	2.35	0.94	0.4	7	15	12	
57.4	2.77	3.2	/	0.067	5	14	/	[57]
/	1.56	3.2	0.92	0.6	1	24	50	[58]
93.4	1.68	3.2	0.92	0.6	1.5	24	50	
64.1	0.94	3.2	/	0.3	1	20	/	[59]
89.4	1.77	/	0.968	0.1	1	/	/	[60]
91.3	1.96	/	0.99	0.1	1	/	/	
87	1.89	/	0.955	0.246	1	26	50	[61]
93	3.9	/	0.955	0.246	2	26	50	
98	5.9	/	0.955	0.246	3	26	50	
94.4	2.08	1.3	0.962	/	1	/	50	[62]
94.4	2.2	1.3	0.978	/	1	/	50	
100	1.61	/	0.94	0.432	1	25	50	[63]
85.3	1.36	1.7	0.97	0.35	1	27.3	50	[64]
84	2.7	1.7	0.97	0.35	2	27.3	50	
87	4.2	1.7	0.97	0.35	3	27.3	50	
83	5.2	1.7	0.97	0.35	4	27.3	50	
84	6.6	1.7	0.97	0.35	5	27.3	50	
95.5	2.53	/	/	/	1	23	40	[65]
88.7	2.07	1.89	0.92	/	1	25	50	[66]
93.5	1.01	1	0.969	0.02	0.5	25	40	[67]
95.3	2.06	1	0.969	0.02	1	25	40	
95.9	3.11	1	0.969	0.02	1.5	25	40	
96.5	4.17	1	0.969	0.02	2	25	40	
96.6	5.22	1	0.969	0.02	2.5	25	40	
96.7	6.27	1	0.969	0.02	3	25	40	
59	1.25	2.8	0.93	/	1	25	40	[68]
72.2	1.42	2.67	0.965	/	1	25	/	
67.5	1.31	1.19	0.96	/	1	25	/	
91	1.76	/	/	/	1	25	15	[69]

2026 Si Yau High School Science Award
仅用于2026丝绸中学科学奖论文公示

/	1.4	3	0.95	0.04	1	/	/	[70]
84.8	1.37	3	0.9	0.04	1	/	/	
91.5	3.7	3	0.9	0.04	3	/	/	
92	5.5	3	0.9	0.04	5	/	/	
72	1.14	/	/	/	1	30	72	[71]
94	1.525	/	0.95	/	1	/	/	[72]
75.6	3.63	2.4	/	/	3	25	50	[73]
93.5	1.16	2.1	/	/	1	/	/	[74]
95	2.3	0.47	/	/	1	25	55	[75]
96.6	1.54	1.7	/	/	1	/	/	[76]
85	1.21	1,7	0.86	0.17	1	25	50	[77]
92	1.917	/	/	/	1	/	/	[78]
93.8	1.62	/	0.97	/	1	22	50	[79]
90.4	1.64	/	0.95	0.4	1	28	Nan	[80]
78	3.4	/	0.72	/	3	/	/	[81]
82	2.6	/	0.967	/	1	/	30	[82]
36	1.2	0.3	0.99	0.2	2	28	60	[83]
54	1.74	0.3	0.99	0.2	2	28	60	
58	1.85	0.3	0.99	0.2	2	28	60	
42	1.3	0.3	0.99	0.08	2	28	60	
59	1.91	0.3	0.99	0.08	2	28	60	
63	2.02	0.3	0.99	0.08	2	28	60	
50	1.6	0.3	0.99	0.06	2	28	60	
64	2.06	0.3	0.99	0.06	2	28	60	
69	2.19	0.3	0.99	0.06	2	28	60	
81.3	2.26	2	/	/	2	22	50	[84]
91.3	2.59	2		/	2.5	22	50	
90	3.6	0.1	/	0.06	1	26	30	[85]
65	2.5	0.1	/	0.06	1	26	30	
78	3	0.1	/	0.06	1	26	30	
74	2.6	0.1	/	0.06	1	26	30	
67	2.3	0.1	/	0.06	1	26	30	
92	3	1.6	0.92	/	1	18	60	[86]
28	0.36	1.3	0.97	/	0.4	26	60	[87]
59	0.75	1.3	0.97	/	0.6	26	60	
79	1.19	1.3	0.97	/	0.8	26	60	
93	1.66	1.3	0.97	/	1	26	60	
81	1.18	/	0.87	0.05	1	22	60	[88]
85	2.48	/	0.87	0.05	2	22	60	
91	5.27	/	0.87	0.05	4	22	60	
62	4.95	2.07	/	/	1	22	60	[89]
0.64	1.53	2.88	/	0.05	1	22	60	[90]
80	1.17	1.5	0.91	0.05	1	22	60	[91]

2026 S.-T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

87	2.53	1.5	0.91	0.05	2	22	60	
92	5.28	1.5	0.91	0.05	4	22	60	
80	1.12	1.5	0.94	0.56	1	22	50	[92]
67	0.46	/	0.9	0.03	0.5	26	60	[93]
76	1.62	/	0.9	0.03	1.5	26	60	
78	2.18	/	0.9	0.03	2	26	60	
90	1.52	0.9	0.95	/	1	26	60	[94]
31	0.5	1.6	0.27	/	1	25	60	[95]
36	0.58	1.6	0.56	/	1	25	60	
48	0.76	1.6	0.87	/	1	25	60	
55	0.88	1.6	0.91	/	1	25	60	
62	0.99	1.6	0.95	/	1	25	60	
92	1.32	1.12	0.86	/	1	28	40	[96]
90	3.84	1.12	0.87	/	3	28	40	
92	6.26	1.12	0.87	/	5	28	40	
97	1.55	1.2	0.84	/	1	26	40	[97]
83	1.44	1.2	0.96	/	1	26	40	
87	1.4	1.1	0.95	/	1	25	60	[98]
73	1.68	1.12	0.99	/	1	22	30	[99]
84	1.27	0.8	0.9	/	1	22	40	[100]
84	3.87	0.8	0.9	/	3	22	40	
84	6.61	0.8	0.9	/	5	22	40	
75	1.88	1.38	0.99	0.03	1	22	30	[101]
79	3.95	1.38	0.99	0.03	1	22	30	
79	5.65	1.38	0.99	0.03	1	22	30	
92	2.46	1.7	0.95	0.06	1	26	52	[102]
94	1.57	1.5	0.9	0.04	1	25	50	[103]
82	1.39	1.5	0.87	0.04	1	25	50	
84	1.42	1.5	0.85	0.04	1	25	50	
71	1.23	1.5	0.82	0.04	1	25	50	
68	1.79	1.22	0.95	/	1	23	30	[104]
95	2.35	1.22	0.97	/	1	23	30	
70	1.75	1.22	0.93	/	1	23	30	
64	1.58	1.22	0.89	/	1	23	30	
4	0.84	1.22	0.8	/	1	23	30	
85	1.78	1.8	0.97	0.47	1	25	40	[105]
81	1.23	1.8	0.91	0.47	1	25	40	
97	2.17	1.35	0.78	0.05	1	25	60	[106]
66	0.98	/	0.71	/	1	22	60	[107]
90	1.5	1.2	0.91	0.6	1	21	50	[108]
42	0.77	1.2	0.62	0.6	1	21	50	
73	1.25	1.2	0.73	0.6	1	21	50	
81	1.37	1.2	0.78	0.6	1	21	50	
85	1.43	1.2	0.87	0.6	1	21	50	

2026 S. - T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

84	1.86	2.26	0.95	0.12	1	25	65	[109]
78	2.18	2.26	0.95	0.12	1	25	65	
93	2.25	2.26	0.95	0.12	1	25	65	
80	2.05	2.26	0.95	0.12	1	25	65	
76	1.13	2.26	0.95	0.12	0.5	25	65	
94	3.1	2.26	0.95	0.12	1.5	25	65	
96	4.19	2.26	0.95	0.12	2	25	65	
96	5.16	2.26	0.95	0.12	3	25	65	
97	6.71	2.26	0.95	0.12	4	25	65	
97	8.27	2.26	0.95	0.12	5	25	65	
93	1.78	1.92	0.95	0.22	1	22	49	[110]
95	2.5	1.92	0.95	0.22	1.5	22	49	
97	3.36	1.92	0.95	0.22	2	22	49	
89	1.36	1.3	0.89	0.05	1	22	60	[111]
88	4	1.3	0.89	0.05	3	22	60	
89	6.8	1.3	0.89	0.05	5	22	60	
76	0.84	1.7	0.98	/	0.5	23	65	[112]
96	1.58	1.7	0.98	/	1	23	65	
86	2.11	1.7	0.98	/	1.5	23	65	
89	2.7	1.7	0.98	/	2	23	65	
86	1.36	1.7	0.98	/	1	23	65	
60	0.96	1.7	0.98	/	1	23	65	
90	1.34	3.11	0.95	0.06	1	25	60	[113]
90	2.34	3.11	0.95	0.06	1	25	60	
99	2.82	2.1	0.95	0.05	1	35	6	[114]
97	1.67	0.45	0.99	/	2	25	50	[115]

Table S2. Optimized hyperparameters of four models for predicting ER.

GBDT	n_estimators = 200 learning_rate = 0.25, max_depth = 3, random_state=42
SVM	C = 2, gama = 0.1, kernel = 'linear'
XGBT	n_estimators = 120, learning_rate = 0.2, max_depth = 5, random_state=35
VR	n_estimators = 100, weight_gb = 2, weight_svm = 1,

2026 S.-T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

	weight_xgb = 4 random_state=35

Table S3. Optimized hyperparameters of four models for predicting EE.

GBDT	n_estimators = 180, learning_rate = 0.2, max_depth = 5, random_state=42
SVM	C = 6, gama = 1, kernel = 'rbf'
XGBT	n_estimators = 200, learning_rate = 0.2, max_depth = 5, random_state=35
VR	n_estimators = 120, weight_gb = 1, weight_svm = 6, weight_xgb = 2 random_state=35

Table S4. Thermal conductivity of MCE membrane

Sample	Ambient temperature ℃	Thermal conductivity W/(m K)
MCE membrane	25	0.07326

Table S5. RMSE, MAE, R^2 of predicted ER for each model on training sets.

ML Models	**RMSE**	**MAE**	**R^2**
VR	0.1818 ± 0.0343	0.3012 ± 0.012	0.9670 ± 0.0061
XGBT	0.2096 ± 0.0541	0.3302 ± 0.0936	0.9620 ± 0.01
GBDT	0.2732 ± 0.0085	0.3404 ± 0.0062	0.9504 ± 0.0078
SVM	0.2180 ± 0.0316	0.4939 ± 0.0358	0.8697 ± 0.0327

2026 S.T. Yau High School Science Award
仅用于2026丘成桐中学科学奖公示

Table S6. RMSE, MAE, R^2 of predicted ER for each model on test sets.

ML Models	RMSE	MAE	R^2
VR	0.3101 ± 0.0153	0.3756 ± 0.0173	0.9337 ± 0.0033
XGBT	0.3260 ± 0.0376	0.3927 ± 0.0166	0.9225 ± 0.0191
GBDT	0.5530 ± 0.0318	0.4952 ± 0.0325	0.8960 ± 0.0107
SVM	0.6188 ± 0.0126	0.6155 ± 0.0072	0.8708 ± 0.0031

Table S7. RMSE, MAE, R^2 of predicted EE for each model on training sets.

ML Models	RMSE	MAE	R^2
VR	0.0021 ± 0.0001	0.0283 ± 0.001	0.9082 ± 0.006
XGBT	0.0039 ± 0.0003	0.043 ± 0.0052	0.8302 ± 0.0113
GBDT	0.0028 ± 0.0016	0.0302 ± 0.0015	0.8641 ± 0.0097
SVM	0.0018 ± 0.0008	0.0196 ± 0.0027	0.9087 ± 0.0011

Table S8. RMSE, MAE, R^2 of predicted EE for each model on test sets.

ML Models	RMSE	MAE	R^2
VR	0.0096 ± 0.0001	0.0735 ± 0.0003	0.5984 ± 0.0011
XGBT	0.0106 ± 0.0004	0.0765 ± 0.0031	0.5598 ± 0.0221
GBDT	0.0108 ± 0.0002	0.0781 ± 0.0053	0.5511 ± 0.0161
SVM	0.0076 ± 0.0001	0.0655 ± 0.0003	0.6240 ± 0.0038

1.1 Programming languages of ML and the core software

All ML models were implemented in Python, installing Python 3 and packages (Numpy, Pandas, Scipy, Scikit-learn, Shap, Matplotlib, and Interpreter).

Open source software libraries lists:

(1) Voting Regression (VR) model

https://scikit-learn.org/stable/modules/ensemble.html

https://github.com/scikit-learn/scikit-learn

https://scikit-learn.org/stable/modules/ensemble.html

(2) Extreme gradient boosting (XGBT) model

2026 S.-T. Yau High School Science Award 仅用于2026丘成桐中学科学奖论文公示

https://github.com/dmlc/xgboost
https://xgboost.readthedocs.io/
(3) Gradient Boosting Decision Tree (GBDT) model
https://github.com/liudragonfly/GBDT
https://github.com/yarny/gbdt
https://scikit-learn.org/stable/modules/ensemble.html
(4) Support vector machines (SVM) model
https://github.com/xinlianghu/svm
https://github.com/AFAgarap/cnn-svm

1.2 Characterizations

The crystal structure was measured by X-ray diffraction spectroscopy (XRD, Holland Panalytical PRO PW3040/60). The elemental composition and atomic ratios were characterized by X-ray photoelectron spectroscopy (XPS, Thermo Scientific K-Alpha). The surface temperature of membranes was taken by an infrared thermal imager (FLTR E95). The band gap was measured using an ultraviolet-visible-near-infrared spectrophotometer (Shimadzu UV-3600i Plus). Raman spectral data were measured using a Raman spectrometer (Horiba LabRAM HR Evolution, Japan). Scanning electron microscopy (SEM) images were acquired using a Zeiss GeminiSEM 500 field-emission scanning electron microscope (FESEM). EPR measurements were performed on a Bruker ELEXSYS E580 X-band spectrometer with an ER4108 TMHS cavity. TEM images were acquired on a FEI Talos F200X transmission electron microscope.

1.3 Solar evaporation experiments

In solar steam generation experiment, a beaker containing 180 mL of distilled water was used as the container. The membrane evaporator was installed, supported by foam wrapped with dust-free paper. The entire setup was placed on an electronic balance (Sartorius BSA224S) to continuously monitor and record the mass change of the distilled water due to evaporation. Subsequently, the beaker was exposed to a 500 W xenon lamp equipped with a standard AM 1.5 G filter to conduct the evaporation experiment. The simulation of solar radiation intensity was measured using an optical power meter (CEL-NP2000-2). Every 10 minutes, surface temperature and mass changes of the membranes were recorded with an infrared camera (FLIR E95) and an electronic balance.

The solar evaporation efficiency (η, %) was calculated using the following equations:

$$\eta = \frac{\dot{m}\Delta H}{C_{opt}P} \tag{S1}$$

2026 S.T. Yau High School Science Award
仅用于2026日成同学科学奖论文公示

$$\Delta H = H_V + Q \tag{S2}$$

$$Q = c\left(T - T_1\right) \tag{S3}$$

$$H_V = 1.91846 \times 10^6 \times \left[\frac{T}{T - 33.91}\right]^2 \tag{S4}$$

where $\dot{m}$ is the evaporation rate (kg m^{-2} h^{-1}), $\dot{m} = \dot{m}_{light} - \dot{m}_{dark}$, $\dot{m}_{light}$ and $\dot{m}_{dark}$ are the solar water evaporation rate under the sunlight and dark conditions, respectively. C_{opt} is the optical concentration. P is the incident solar power (1 kW m^{-2}, one sun). ΔH is the liquid-vapor phase change enthalpy, H_V is the phase-change enthalpy (J kg^{-1}), and Q is the sensible heat of water. T is the surface temperature of solar evaporators, T_1 is the initial temperature of the water, and c is the specific heat of water (4.2 J g^{-1} K^{-1}). In this work, only the phase-change enthalpy is considered when calculating the evaporation efficiency.



Fig. S1. Four different prepared samples.

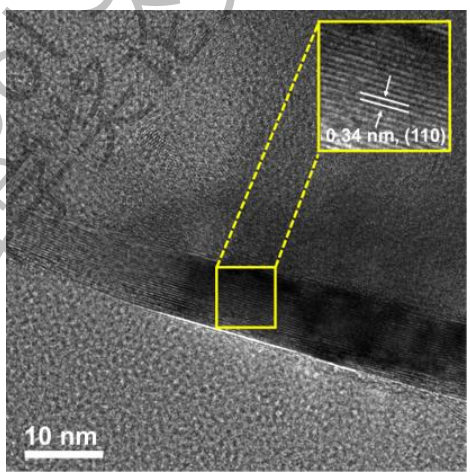


Fig. S2. HR-TEM image of 700 ℃.

References

[1] C.J. Chen, H.Y. Liu, H.T. Wang, Y.W. Zhao, M. Li. A scalable broadband plasmonic cuprous telluride nanowire-based hybrid photothermal membrane for efficient solar vapor generation, Nano Energy 84 (2021) 105868.

[2] Z.Z. Guo, G. Wang, X. Ming, T. Mei, J.Y. Wang, J.H. Li, J.W. Qian, X.B. Wang, PEGylated Self-Growth MoS_2 on a Cotton Cloth Substrate for High-Efficiency Solar Energy Utilization, ACS Appl. Mater. Interfaces10 (2018) 24583-24589.

[3] H.W. Liu, C.J. Chen, H. Wen, R.X. Guo, N.A. Williams, B.D. Wang, F.J. Chen, L.B. Hu, Narrow bandgap semiconductor decorated wood membrane for high-efficiency solar-assisted water purification, J. Mater. Chem. A, 6 (2018) 18839.

[4] C.H. Mu, Y.Q. Song, K. Deng, S.A. Lin, Y.T. Bi, F. Scarpa, D. Crouse. High Solar Desalination

2026 S.-T. Yau High School Science Award
仅用于2026丘成桐中学科学奖公示

Efficiency Achieved with 3D Cu_2ZnSnS_4 Nanosheet-Assembled Membranes, Adv. Sustainable Syst. 1 (2017) 1700064.

[5] Y. Shi, R.Y. Li, L. Shi, E. Ahmed, Y. Jin, P. Wang, A Robust $CuCr_2O_4$ SiO_2 Composite Photothermal Material with Underwater Black Black Property and Extremely High Thermal Stability for Solar-Driven Water Evaporation, Adv. Sustainable Syst. 2 (2018) 1700145.

[6] M.M. Ye, J. Jia, Z.J. Wu, C.X. Qian, R. Chen, P.G. O'Brien, W. Sun, Y.C. Dong, G.A. Ozin, Synthesis of Black TiOx Nanoparticles by Mg Reduction of TiO_2 Nanocrystals and their Application for Solar Water Evaporation, Adv. Energy Mater. 7 (2017) 1601811.

[7] Q.C. Lu, Y. Yang, J.R. Feng, X. Wang, Oxygen-Defected Molybdenum Oxides Hierarchical Nanostructure Constructed by Atomic-Level Thickness Nanosheets as an Efficient Absorber for Solar Steam Generation, Sol. RRL. 3 (2019) 1800277.

[8] Y.C. Liu, X.Q. Wang, H. Wu, High-performance wastewater treatment based on reusable functional photo-absorbers, Chem. Eng. J. 309 (2017) 787-794.

[9] D.S. Li, D.T. Han, C.W. Guo, C.L. Huang, Facile Preparation of MnO_2-Deposited Wood for High-Efficiency Solar Steam Generation, ACS Appl. Energy Mater. 4 (2021) 1752-1762.

[10] X.J. Zha, X. Zhao, J.H. Pu, L.S. Tang, K. Ke, R.Y. Bao, L. Bai, Z.Y. Liu, M.B. Yang, W. Yang, Flexible Anti-Biofouling MXene Cellulose Fibrous Membrane for Sustainable Solar-Driven Water Purification, ACS Appl. Mater. Interfaces 11 (2019) 36589-36597.

[11] M. Kaur, S. Ishii, S.L. Shinde, T. Nogao, All-Ceramic Solar-Driven Water Purifier Based on Anodized Aluminum Oxide and Plasmonic Titanium Nitride, Adv. Sustainable Syst. 3 (2019) 1800112.

[12] Y. Xu, J.X. Ma, Y. Han, J.J. Zhang, F.Y. Cui, Y. Zhao, X. Li, W. Wang, Multifunctional CuO Nanowire Mesh for Highly Efficient Solar Evaporation and Water Purification, ACS Sustainable Chem. Eng. 7 (2019) 5476-5485.

[13] F.J. Tao, Y.L. Zhang, K. Yin, S.J. Cao, X.T. Chang, Y.H. Lei, D.S. Wang, R.H. Fan, L.H. Dong, Y.S. Yin, X.B. Chen, Copper Sulfide-Based Plasmonic Photothermal Membrane for High-Efficiency Solar Vapor Generation, ACS Appl. Mater. Interfaces 10 (2018) 35154-35163.

[14] F.J. Tao, Y.L. Zhang, S.J. Cao, K. Yin, X.T. Chang, Y.H. Lei, R.H. Fan, L.H. Dong, Y.S. Yin, X.B. Chen, CuS nanoflowers/semipermeable collodion membrane composite for high-efficiency solar vapor generation, Mater. Today Energy 9 (2018) 285-294.

[15] M.Y. Shang, N. Li, S.D. Zhang, T.T. Zhao, C. Zhang, C. Liu, H.F. Li, Z.Y. Wang, Full-spectrum solar-to-heat conversion membrane with interfacial plasmonic heating ability for high-efficiency desalination of seawater, ACS Appl. Energy Mater. 1 (2018) 56-61.

[16] X. Wu, M.E. Robson, J.L. Phelps, J.S. Tan, B. Shao, G. Owens, H.L. Xu, A flexible photothermal cotton-CuS nanocage-agarose aerogel towards portable solar steam generation, Nano Energy 56 (2019) 708-715.

[17] Y.W. Yang, H.Y. Zhao, Z.Y. Ying, J.Q. Zhao, X.T. Ying, N. Li, D.D. Ying, Y.N. Li, B. Lei, Y.P. Du,

2026 S.T. Yau High School Science Award
仅用于2026已发利用学校公示

W.X. Que, A general salt-resistant hydrophilic/hydrophobic nanoporous double layer design for efficient and stable solar water evaporation distillation, Mater. Horiz. 5 (2018) 1143-1150.

[18] D. Ghim, Q.S. Jiang, S.S. Cao, S. Singamaneni, Y.S. Jun, Mechanically interlocked 1T/2H phases of MoS_2 nanosheets for solar thermal water purification, Nano Energy 53 (2018) 949-957.

[19] W.G. Li, M.C. Tekell, Y. Huang, K. Bertelsmann, M. Lau, D.L. Fan, Synergistic High-Rate Solar Steaming and Mercury Removal with MoS_2/C @Polyurethane Composite Sponges, Adv. Energy Mater. 15 (2018) 1802108.

[20] X.D. Yang, Y.B. Yang, L.N. Fu, M.C. Zou, Z.H. Li, A.Y. Cao, Q. Yuan, An Ultrathin Flexible 2D Membrane Based on Single-Walled Nanotube–MoS_2 Hybrid Film for High-Performance Solar Steam Generation, Adv. Funct. Mater. 28 (2018) 1704505.

[21] R. Chen, X. Wang, Q.M. Gan, T.Q. Zhang, K.H. Zhu, M.M. Ye, A bifunctional MoS_2-based solar evaporator for both efficient water evaporation and clean freshwater collection, J. Mater. Chem. A 7 (2019) 11177-11185.

[22] J.Wang, Y.Y. Li, L. Deng, N.N. Wei, Y.K. Weng, S. Dong, D.P. Qi, J. Qiu, X.D. Chen, T. Wu, High-Performance Photothermal Conversion of Narrow-Bandgap Ti_2O_3 Nanoparticles, Adv. Mater. 29 (2017) 1603730.

[23] J. Huang, Y.R. He, M.J. Chen, X.Z. Wang, Separating photo-thermal conversion and steam generation process for evaporation enhancement using a solar absorbe, Appl. Energy 236 (2019) 244-252.

[24] P.F. Wang, Y.F. Gu, L. Miao, J.H. Zhou, H. Su, A.Y. Wei, X.J. Mu, Y.Z. Tian, J.Q. Shi, H.F. Cai, Co_3O_4 nanoforest/Ni foam as the interface heating sheet for the efficient solar-driven water evaporation under one sun, Sustainable Mater. and Technol. 17 (2019) e00106.

[25] D.D. Wu, D. Qu, W.S. Jiang, G. Chen, L. An, C.Q. Zhuang, Z.C. Sun. Self-floating nanostructured Ni–NiOx/Ni foam for solar thermal water evaporation, J. Mater. Chem. A 7 (2019) 8485-8490.

[26] R. Chen, K.H. Zhu, Q.M. Gan, Y.Q. Yu, T.Q. Zhang, X.W. Liu, M.M. Ye, Y.D. Yin. Interfacial solar heating by self-assembled Fe_3O_4@C film for steam generation, Mater. Chem. Front. 1 (2017) 2620-2626.

[27] L. Sun, Z. Li, R. Su, Y.L. Wang, Z.L. Li, B.S. Du, Y. Sun, P.F. Guan, F. Besenbacher, M. Yu, Phase-Transition Induced Conversion into a Photothermal Material: Quasi-Metallic $WO_{2.9}$ Nanorods for Solar Water Evaporation and Anticancer Photothermal Therapy, Angew. Chem. Int. Ed. 130 (2018) 10826-10831.

[28] W.X. Li, W.Q. Feng, S.X. Wu, W. Wang, D. Yu, Synergy of photothermal effect in integrated 0D Ti_2O_3 nanoparticles 1D carboxylated carbon nanotubes for multifunctional water purification, Sep. Purif. Technol. 292 (2022) 120989.

[29] L.F. Su, Y.Q. Hu, Z.Q. Ma, L. Miao, J.H. Zhou, Y.Y. Ning, Z.P. Chang, B. Wu, M. Cao, R. Xia, J.S. Qian, Synthesis of hollow copper sulfide nanocubes with low emissivity for highly efficient solar steam generation, Sol. Energy Mater. Sol. Cells 210 (2020) 110484.

2026 S. T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

[30] P.J. Ying, M. Li, F.L. Yu, Y. Geng, L.Y. Zhang, J.J. He, Y.J. Zheng, R. Chen, Band Gap Engineering in an Efficient Solar-Driven Interfacial Evaporation System, ACS Appl. Mater. Interfaces 12 (2020) 32880-32887.

[31] X.D. Chen, C. Meng, Y. Wang, Q.Q. Zhao, Y.J. Li, X.M. Chen, D.W. Yang, Y.X. Li, Y. Zhou, Laser-Synthesized Rutile TiO_2 with Abundant Oxygen Vacancies for Enhanced Solar Water Evaporation, ACS Sustainable Chem. Eng. 8 (2020) 1095-1101.

[32] N.S. Fuzil, N.H. Alias, N.H. Othman, M.S.M. Shayuti, N.A. Rahman, Z.A. Rashid, A.F. Ismail, A. UI-Hamid, I. Ichinose, Mechanistic insights of facile synthesized-solar driven TiO_2 intercalated 2D molybdenum disulfide evaporator for one-pot clean water production, Desalination 613 (2025) 119108.

[33] R. Meng, Y. Lu, L.M. Zou, H.D. Du, X.L. Li, B. Zhu, D.K. Macharia, M. Amidpour, Z.G. Chen, L.S. Zhang, Growth of TiO_2/Ti- MOF nanorod array with enhanced photoabsorption and photocatalytic properties on carbon cloth for efficient auto-cleaning solar desalination, Desalination 578 (2024) 117455.

[34] J. Li, E.H. Wu, Z. Xu, J. Hou, W.J. Peng, H. Li, X. Li, Superior photothermal conversion performance of black titanium-based materials, iScience 28 (2025) 112188.

[35] L.C. Yi, S.Q. Ci, S.L. Luo, P. Shao, Y. Hou, Z.H. Wen, Scalable and low-cost synthesis of black amorphous Al-Ti-O nanostructure for high-efficient photothermal desalination, Nano Energy 41 (2017) 600-608.

[36] C.Y. Ma, X.Y. Li, H. Zhang, J.Q. Li, J.L. Zhao, G. He, J.T. Li, T. Qi, Preparation and Photothermal Conversion Performance of Submicron Ti_4O_7, Mater. Rev. 32 (2018) 4079-4083. (Chinese paper)

[37] X.H. Liu, B.F. Hou, G. Wang, Z.Q. Cui, X. Zhu, X.B. Wang, Black titaniagraphene oxide nanocomposite films with excellent photothermal property for solar steam generation, J. Mater. Res. 33 (2018) 674-684.

[38] H.F. Yang, Y.S. Liu, Z.Q. Qiao, X.D. Li, Y.C. Wang, H.B. Bao, G.C. Yang, X.M. Li, One-step ultrafast deflagration synthesis of N-doped $WO_{2.9}$ nanorods for solar water evaporation, Appl. Surf. Sci. 555 (2021) 149697.

[39] X.P. Qiu, H.R. Kong, Y.T. Li, Q.H. Wang, Y. Wang, Interface engineering of a Ti_4O_7 nanofibrous membrane for efficient solar-driven evaporation, ACS Appl. Mater. Interfaces 14 (2022) 54855-54866.

[40] R.X. Zheng, T. Lin, W.L. Zhao, R.T. Yin, H.Q. Li, Z.Q. Deng, W. Chen, Y.F. Song, Hierarchical CoMn-LDH based photothermal membrane with low evaporation enthalpy and narrow bandgap toward highly efficient Solar-Driven evaporation, Chem. Eng. J. 470 (2023) 144103.

[41] D. Sun, L.L. Fan, W. Gong, D.Z. Ye, Z.K. Bai, J. Xu, A facile and versatile preparation method of sodium alginate-copper sulfide photothermal coating for efficient solar evaporation, Int. J. of Biol. Macromol. 279 (2024) 135164.

[42] C.R. Xue, Q. Zhang, Y. Shen, S.L. Hu, Q. Chang, H.Q. Wang, N. Li, J.L. Yang, Laser thermal

2026 S.-T. Yau High School Science Award
仅用于2026丘成桐中学科学奖公示

synthesis of reduced graphene oxide/CuS nanocomposites for efficient solar-driven water purification, Sep. Purif. Technol. 294 (2022) 121168.

[43] Z.G. Wang, X.F. Zhang, L. Shu, J.F. Yao, Copper sulfide integrated functional cellulose hydrogel for efficient solar water purification, Carbohydr. Polym. 319 (2023) 121161.

[44] S.J. Zhao, H.F. Wei, X. Zhang, F.Y. Wang, Z.Q. Su, Clay-based aerogel combined with CuS for solar-driven interfacial steam generation and desalination, J. of Colloid and Interface Sci. 653 (2024) 1504-1513.

[45] J. Xiong, Z.H. Yang, B. Wu, M. Li, X. Min, CuS-enhanced light-absorbing washable solar evaporator based on polydopamine-cotton fabric for efficient water purification, Int. J. Energy Res. 46 (2022) 16979-16990.

[46] F.J. Tao, Y.L. Zhang, S.J. Cao, K. Yin, X.T. Chang, Y.H. Lei, R.H. Fan, L.H. Dong, Y.S. Yin, X.B. Chen, CuS nanoflower/semipermeable collodion membrane composite for high-efficiency solar vapor generation, Mater. Today Energy 9 (2018) 285-294.

[47] D. Zhang, Y.X. Cai, Q.Q. Liang, Z.T. Wu, N. Sheng, M.H. Zhang, B.X. Wang, S.Y. Chen, Scalable, Flexible, Durable, and Salt-Tolerant CuS/Bacterial Cellulose Gel Membranes for Efficient Interfacial Solar Evaporation, ACS Sustainable Chem. Eng. 8 (2020) 9017-9026.

[48] M.Y. Shang, S.H. Xu, J.L. Li, H.J. Sun, J. Peng, S. Wang, M. Zhang, CuS Hollow Nanospheres/Cellulose Composite Film as a Recyclable Interfacial Photothermal Evaporator for Solar Steam Generation, Energy Technol. 10 (2022) 2100805.

[49] N. Li, D.D. Yin, L.L. Xu, H.Y. Zhao, Z.Q. Liu, Y.P. Du, High-quality ultralong copper sulphide nanowires for promising applications in high efficiency solar water evaporation, Mater. Chem. Front. 3 (2019) 394.

[50] C.J. Song, L. Wang, X. Li, L. Guo, Y.M. Leng, X.L. Jin, L.Q. Ye, A Chitin/CuS composite film for efficient solar seawater desalination, Inorg. Chem. Commun. 133 (2021) 108886.

[51] F.A. Bezza, S.A. Iwarere, S.M. Tichapondwa, E.M.N. Chirwa, Fabrication and Application of Ag, Black TiO_2 and Nitrogen-Doped 3D Reduced Graphene Oxide (3D Black TiO_2/Ag/N@rGO) Evaporator for Efficient Steam Generation, Catalysts 13 (2023) 514.

[52] W.Z. Wu, X.T. Han, X. Gong, Stable superhydrophobic interfacial solar evaporator based on CuS CDs hybrid photothermal composites towards ultrahigh-efficiency seawater evaporation, Appl. Energy 375 (2024) 124099.

[53] Y.Y. Liu, Y.J. Lao, Y.S. Wang, Q. Zhang, F.P. Du, Y.F. Zhang, Synergistic photothermal conversion in CuS-MoS_2/poly (vinyl alcohol) sponge for efficient solar water evaporation, Sep. Purif. Technol. 359 (2025) 130860.

[54] T.F. Chala, C.M. Wu, M.H. Chou, Z.L. Guo, Melt Electrospun Reduced Tungsten Oxide/Polylactic Acid Fiber Membranes as a Photothermal Material for Light-Driven Interfacial Water Evaporation, ACS Appl. Mater. Interfaces 10 (2018) 28955-28962.

[55] X. Ming, A.K. Guo, G. Wang, X.B. Wang, Two-dimensional defective tungsten oxide nanosheets as

2026 S.T. Yau High School Science Award
仅用于2026丘成桐中学科学奖公示

high performance photo-absorbers for efficient solar steam generation, Sol. Energy Mater. Sol. Cells 185 (2018) 333-341.

[56] Z.K. Li, M. Zheng, N. Wei, Y. Lin, W.L. Chu, R.Q. Xu, H.H. Wang, J. Tian, H.Z. Cui, Broadband-absorbing WO_3-x nanorod-decorated wood evaporator for highly efficient solar-driven interfacial steam generation, Sol. Energy Mater. Sol. Cells 205 (2020) 110254.

[57] Y. Zhang, L. Liu, K.L. Li, D.Y. Hou, J. Wang, Enhancement of energy utilization using nanofluid in solar powered membrane distillation, Chemosphere 212 (2018) 554-562.

[58] Y.P. Gao, X.J. Zhou, N. Fu, S.H. Su, B.P. Ma, Q.F. Ruan, D. Wu, N. Zhang, Z.W. Deng, R.B. Jiang, Superior solar water purification achieved by the excellent photothermal conversion and thermal management of porous plasmonic TiN nanoparticles, Chem. Eng. J. 475 (2023) 146078.

[59] Y. Zhang, K.L. Li, L. Liu, L.P. Wang, J. Xiang, D.Y. Hou, J. Wang, Titanium nitride nanoparticle embedded membrane for photothermal membrane distillation, Chemosphere 256 (2020) 127053.

[60] D.D. Pi, T.T. Xiong, W.Q. Gong, H.T. Zou, C. Gao, X. Liu, Study on preparation and performance of TiN modified vertically aligned carbon fiber evaporator, J. Alloys and Compd. 1003 (2024) 175702.

[61] J. Jiang, X.J. Lai, Y. Fang, R.R. Yang, X.Y. Zheng, S. Tian, F.F. Yang, A.K. Liu, L.H. Ai, “One stone two birds” strategy for sustainable water purification Bifunctional photothermic-catalytic manganese oxide/carbon composites enabling solar evaporation, seawater desalination and periodate activation, Desalination 597 (2025) 118402.

[62] X.X. Song, X.L. Li, B. Zhu, S.M. Sun, Z.G. Chen, L.S. Zhang, MnO_2/Poly-L-lysine Co-decorated Carbon Fiber Cloth with Decreased Evaporation Enthalpy and Enhanced Photoabsorption/ Antibacterial Performance for Solar-Enabled Anti-fouling Seawater Desalination, Adv. Fiber Mater. 6 (2024) 1569-1582.

[63] M. Shafaee, Z. Niazi, M. Asarnia, E.K. Goharshadi, R. Dehghani, Modified pine cone with MnO_2 nanoparticles as a photoabsorber for highly efficient seawater desalination and wastewater treatment, Appl. Phys. A-Mater 129 (2023) 656.

[64] A.M. Saleque, S. Ma, A.K. Thakur, R. Saidur, T.K. Han, M.I. Hossain, W. Qarony, Y. Ma, R. Sathyamurthy, Y.H. Tsang, MXene/MnO_2 nanocomposite coated superior salt-rejecting biodegradable luffa sponge for efficient solar steam generation, Desalination 554 (2023) 116488.

[65] S. Li, T. Liu, Y. Yao, M.Z. Yang, F.F. Li, Z.B. Zhao, S. Zhang, G.H. He, X.C. Li, MoS_2 supported hydrophilic porous membrane for solar water evaporation, Sep. Purif. Technol. 353 (2025) 128566.

[66] L.F. Cui, H.N. Che, B. Liu, Y.H. Ao, Multifunctional MoS_2 membrane for integrated solar-driven water evaporation and water purification, Commun. Mater. 5 (2024) 94.

[67] H.H. Shen, Z.H. Zheng, H. Liu, X.D. Wang, A solar-powered interfacial evaporation system based on MoS_2-decorated magnetic phase-change microcapsules for sustainable seawater desalination, J. Mater. Chem. A 10 (2022) 25509-25526.

[68] B.H. Yuan, L.S. Meng, C.F. Zhang, L.X. Yang, L.J. Bai, H.W. Yang, D.L. Wei, F. Wang, Q.Y. Wang,

2026 S.T. Yau High School Science Award 仅用于2026丘成桐中学科学奖论文公示

W.X. Wang, H. Chen, Enhancement of pollutant degradation and solar-driven water evaporation by architecting hierarchical 1D/2D TiO_2@MoS_2 core-shell networks, Appl. Surf. Sci. 570 (2021) 151143.

[69] H.J. Jiang, H.N. Fang, D. Wang, J.B. Sun, Coated Commercial PTFE Membrane from MoS_2/LaF_3/PDMS Ink as Solar Absorber for Efficient Solar Steam Generation, Sol. RRL. 4 (2020) 2000126.

[70] Q. Zhu, K. Ye, W. Zhu, W.J. Xu, C.W. Zou, L. Song, E. Sharman, L.J. Wang, S.Y. Jin, G.Z. Zhang, Y. Luo, J. Jiang, A Hydrogenated Metal Oxide with Full Solar Spectrum Absorption for Highly Efficient Photothermal Water Evaporation, J. Phys. Chem. Lett. 11 (2020) 2502-2509.

[71] S.Y. Zuo, D.S. Xia, Z.Y. Guan, F. Yang, S. Cheng, H.M. Xu, R.Z. Wan, D.Y. Li, M. Liu, Dual-functional CuO/CN for highly efficient solar evaporation and water purification, Sep. Purif. Technol. 254 (2021) 117611.

[72] P. Ren, J. Li, X. Zhang, X. Yang, Highly efficient solar water evaporation of TiO_2@TiN hyperbranched nanowires-carbonized wood hierarchical photothermal conversion material, Mater. Today Energy 18 (2020) 100546.

[73] M. Aziznezhad, E. Goharshadi, M. Namayandeh-Jorabchi, Surfactant-mediated prepared VO_2 (M) nanoparticles for efficient solar steam generation, Sol. Energy Mater. Sol. Cells 211 (2020) 110515.

[74] M. Aziznezhad, E.K. Goharshadi, R. Mehrkhaha, M. M. Ghafurian, Alkaline earth metals doped VO_2 nanoparticles for enhanced interfacial solar steam generation, Mater. Res. Bull. 149 (2022) 111705.

[75] T. A. Wani, P. Garg, S. Bera, S. Bhattacharya, S. Dutta, H. Kumar, A. Bera, Narrow-Bandgap $LaMO_3$ (M=Ni, Co) nanomaterials for efficient interfacial solar steam generation, J. Colloid Interface Sci. 612 (2022) 203-212.

[76] C.R. Xue, P.F. Li, Q. Zhang, D. Li, Y.S. Li, Q. Chang, J.L. Yang, Y. Li, J.F. Zhang, S.L. Hu, Fabrication of high-performance graphene oxide/CuO/Cu_2O film-coated copper foam for interfacial solar-driven water evaporation, J Mater Sci. 57 (2022) 3322-3336.

[77] J.H. Zhao, Z. Liu, S.C. Low, Z.Z. Xu, S.H. Tan, Biomimetic hydrophilic CuO-nanocluster/nanofibrous thin-film skeleton/hydrophilic CuO-nanocluster architecture membrane for efficient and stable solar water evaporation, J. Alloys Compd. 969 (2023) 172377.

[78] H.W. Wang, Y.M. Li, Y.Y. Shi, S.M. Gao, CMF/Cu@CuS as photothermal conversion material for solar-driven water evaporation in wet sand systems, Mater. Lett. 337 (2023) 133947.

[79] S. Chhetri, A.T. Nguyen, S. Song, D.H. Park, T.W. Ma, N. Gaillard, S.H. Yoon, W. Lee, Enhanced Photothermal Effect Assisted by Resonance Energy Transfer in Carbon/Covellite Core-Shell Nanoparticles toward a High-Performance Interfacial Water Evaporation Process, CS Appl. Mater. Interfaces 15 (2023) 54773-54785.

[80] X.X. Shi, D. Sun, D.M. Yue, B.B. Li, J.R. Zhu, L.Y. Zu, F. Liu, X.J. Li, A novel hierarchical chromium black membrane prepared by electroplating for desalination, Desalination 583 (2024)

2026 S.-T. Yau High School Science Award 仅用于2026丘成桐中学科学奖论文公示

117704.

[81] H.S. Sajjadizadeh, M. Karimi-Nazarabad1, E.K. Goharshadi,·A. Ebrahimi, F. Asadi, Bi-doped VO_2 coated onto wood as a highly efficient photoabsorber in interfacial solar steam generation, J. Iran. Chem. Soc. 21 (2024) 503-509.

[82] L.Z. Xing, X.K. Li, R.P. Wang, Y. Ha, D. Li, B. Chen, Z.M. Li, Fe_3O_4/Au@SiO_2 nanocomposites with recyclable and wide spectral photo-thermal conversion for a direct absorption solar collector, Renewable Energy 235 (2024) 121269.

[83] M.Q. Tang, L.F. Yue, C. Qi, L. Liang, Solar interface evaporation performance of Cu-Fe_3O_4 nanoparticle film coupled with thermal insulation materials, Sol. Energy 273 (2024) 112543

[84] C.B. Zhang, C. Yan, Z.J. Xue, et al. Shape-controlled synthesis of high-quality Cu_7S_4 nanocrystals for efficient light-Induced water evaporation, small 38 (2021) 5320-5328.

[85] Y.H. Guo, X.Y. Zhou, F. Zhao, et al. Synergistic energy nanoconfinement and water activation in hydrogels for efficient solar water desalination, ACS Nano 13 (2019) 7913-7919.

[86] X.C. Zhang, X.N. Wang, W.D. Wu, et al. S Self-floating monodisperse microparticles with a nano-engineered surface composition and structure for highly efficient solar-driven water evaporation, Journal of Materials Chemistry A 7 (2019) 6963.

[87] H.Y. Cheng, X.H. Liu, L.X. Zhang, et al. Self-floating Bi2S3/poly (vinylidene fluoride) composites on polyurethane sponges for efficient solar water purification, Solar Energy Materials and Solar Cells 203 (2019) 110127.

[88] Y.L. Zhang, F.J. Tao, S.J. Cao, et al. Hierarchical K2Mn4O8 nanoflowers: A novel photothermal conversion material for efficient solar vapor generation, Solar Energy Materials and Solar Cells 200 (2019) 110043.

[89] D.D. Ding, W.C. Huang, C.Q. Song, et al. Non-stoichiometric MoO3−x quantum dots as a light-harvesting material for interfacial water evaporation, Chemical Communications 53 (2017) 6744.

[90] Z.T. Hua, B. Li, L.L. Yin, et al. Designing a novel photothermal material of hierarchical microstructured copper phosphate for solar evaporation enhancement, The Journal of Physical Chemistry C 121 (2017) 60-69.

[91] F.J. Tao, Y.L. Zhang, K. Yin, et al. A plasmonic interfacial evaporator for high-efficiency solar vapor generation, Sustainable Energy & Fuels 2 (2018) 2762.

[92] Z.Z. Guo, X. Ming, G. Wang, et al. Super-hydrophilic copper sulfide films as light absorbers for efficient solar steam generation under one sun illumination, Semiconductor Science and Technology 32 (2018) 1088.

[93] W. Cai, X.W. Mu, Y. Pan, et al. Black phosphorous nanosheets: a novel solar vapor generator, Solar RRL 4 (2020) 1900537.

[94] Z.Z. Guo, Z.H. Chen, Z.X. Shi, et al. Stable metallic 1T phase engineering of molybdenum disulfide for enhanced solar vapor generation, Solar Energy Materials and Solar Cells 204 (2020) 110227.

[95] X.D. Li, D.Y. Wang, Y. Zhang, et al. Surface-ligand protected reduction on plasmonic tuning of one-

2026 S. T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

dimensional MoO3-x nanobelts for solar steam generation, Nano Research 13 (2018) 3025-3032.

[96] Z.G. Wang, K. Yu, S.J. Gong, et al. Cu3BiS3/MXenes with excellent solar–thermal conversion for continuous and efficient seawater desalination, Nano Research 13 (2021) 16246-16258.

[97] Y.R. Gu, X. Li, X.J. Li, et al. Facile preparation of Cu2S/Cu mesh for high-performance solar water evaporation, ChemistrySelect 6 (2021) 7901-7905.

[98] J. Xiong, Z.N. Zhang, J. Yi, et al. Hierarchical MnO2 nanosheets grown on cotton fabric as a flexible and washable solar evaporator for seawater desalination, ACS Applied Nano Materials 4 (2021) 13724-13733.

[99] P.F. Cheng, H.L. Wang, H.G. Wang, et al. High-efficiency photothermal water evaporation using broadband solar energy harvesting by ultrablack silicon structures, Advanced Energy and Sustainability Research 2 (2021) 2000083.

[100] Z.G. Wang, W.Q. Xu, K. Yu, et al. NiS2 nanocubes coated Ti3C2 nanosheets with enhanced light-to-heat conversion for fast and efficient solar seawater steam generation, Solar RRL (2021) 2100183.

[101] L.L. Hou, N. Wang, L.J. Yu, et al. High-performance Janus solar evaporator for water purification with broad spectrum absorption and ultralow heat loss, ACS Energy Letters 8 (2023) 553-564.

[102] Y.M. Ge, Z.W. Su, L. Wang, et al. Self-rotating spherical evaporator based on hydrogel and black titanium oxide for continuous desalination of seawater, ACS Materials Letters 5 (2023) 2576-2583.

[103] Y.L. Zheng, H. Zhang, P. Han, et al. Micro-flower like 1T-2H-MoS2@ZIF8@C composites for efficient interfacial solar vapor generation, New Journal of Chemistry 47 (2023) 16531.

[104] W.T. Wang, K. Zhang, Y.H. Wang, et al. A CuO/ZnO heterostructure interpenetrated with hydrogel network for efficient solar eater evaporation, Industrial & Engineering Chemistry Research 63 (2024) 15890-15900.

[105] R. Song, N.S. Zhang, P. Wang, et al. Three-dimensional NiO/Ni self-floating porous composite materials for efficient solar interfacial evaporation, Colloids and Surfaces A: Physicochemical and Engineering Aspects 695 (2024) 134261.

[106] J.H. Hu, X. Yao, K.Y. Han, et al. Boosted near-infrared photothermal conversion in rare earth ions-doped 2D SnSe nanosheets for solar-powered water evaporation systems, Small 20 (2024) 2405742.

[107] W.X. Mo, Q.Q. Hu, J. Guan, et al. Enhanced dispersion of prussian blue via intercalation into layered double hydroxides for efficient solar seawater evaporation, Dalton Transactions 53 (2024) 10285.

[108] C. Xue, C. Qi, W.Z. Li, et al. A novel rGO-Fe3O4-mixed cellulose ester membrane for solar evaporation and seawater desalination, Solar Energy Materials and Solar Cells 290 (2025) 113708.

[109] X. Jiang, C.M. Tang, X.L. Li, et al. Atomically precise Ni nanoclusters supported on CNT confined in hydrogel for efficient solar water evaporation: Understanding synergistic effect of Ni Nanoclusters and CNT on photothermal conversion, Chemical Engineering Journal 521 (2025) 166473.

[110] P. Xiang, C.M. Tang, K. Ma, et al. Highly efficient and robust oxygen vacancy-rich molybdenum

2026 S. T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

trioxide aerogel evaporator for Photothermal conversion and clean water generation, Desalination 594 (2025) 118281.

[111] Z.P. Chen, Z.R. Qian, Y.F. Wu, et al. Micro-interface engineering of CuO-coated copper mesh for enhanced solar absorption and localized heating in continuous seawater desalination, Solar Energy Materials and Solar Cells 282 (2024) 113373.

[112] C.P. Li, Y.Q. Zhang, Z.Y. Wang, et al. Tannin-iron complex-modified CoO/NiCo2O4 nanorod arrays enable anti-oil-fouling and durable solar driven desalination, Journal of Environmental Chemical Engineering 13 (2025) 118480.

[113] T. Matsuoka, C. Numako, H. Itasaka, et al. Morphology- and valence-tailored plasmonic MoO3-x nanoarchitectures achieving high-performance photothermal evaporation for water desalination, Surfaces and Interfaces 98 (2026) 110356.

[114] H. Tahzibi, S. Azizian, M. Hajivaliei, Deposition of CuxS thin layer on melamine foam and utilization as high-performance interfacial solar vapor generator, Surfaces and Interfaces 84 (2025) 108437.

[115] M.F. Maqsood, K.C. Fong, F. Ghafoor, et al. Solar vapor generation: Advances in materials engineering and structural design for efficient water evaporation, Materials Today Physics 59 (2025) 101938.

2026 S.-T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

Appendix B: Validating machine learning model stability

To verify the ML model's stability, I have repeated the experimental model multiple times (recommended 10 times) and present the data appropriately and quantitatively analysed the performance metrics (RMSE, MAE, R^2) as shown in the following Tables R1-R16.

Table R1. RMSE, MAE, R^2 of predicted ER for VR model on training sets.

Number	RMSE	MAE	R^2
1	0.2087	0.3219	0.9558
2	0.1951	0.3129	0.9583
3	0.1625	0.2994	0.9642
4	0.1706	0.3012	0.9627
5	0.191	0.3075	0.959
6	0.1416	0.2923	0.968
7	0.2272	0.3206	0.9524
8	0.2283	0.2908	0.9704
9	0.1457	0.2932	0.9672
10	0.1373	0.2915	0.9688
Mean	0.1818	0.3012	0.9670
Standard deviations	0.0343123	0.0119856	0.0060904

Table R2. RMSE, MAE, R^2 of predicted ER for VR model on test sets.

Number	RMSE	MAE	R^2
1	0.3345	0.4015	0.9284
2	0.3305	0.3984	0.9293
3	0.3174	0.3795	0.9321
4	0.2978	0.3603	0.9364
5	0.3044	0.3718	0.9349
6	0.2956	0.3599	0.9369
7	0.3244	0.3965	0.9306
8	0.305	0.3673	0.9348
9	0.2944	0.3598	0.9371
10	0.2974	0.3607	0.9365
Mean	0.31014	0.37557	0.9337
Standard deviations	0.0152647	0.0172505	0.00332

Table R3. RMSE, MAE, R^2 of predicted EE for VR model on training sets.

Number	RMSE	MAE	R^2

2026 S.-T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

1	0.002	0.0279	0.9128
2	0.002	0.0275	0.9136
3	0.0022	0.029	0.9039
4	0.0021	0.0286	0.9057
5	0.0022	0.0289	0.9038
6	0.0022	0.0292	0.9021
7	0.0022	0.0295	0.9005
8	0.0021	0.0282	0.9084
9	0.002	0.0276	0.9117
10	0.0018	0.0263	0.9191
Mean	0.00208	0.02827	0.90816
Standard deviations	0.0001317	0.000971	0.0059815

Table R4. RMSE, MAE, R^2 of predicted EE for VR model on test sets.

Number	RMSE	MAE	R^2
1	0.0097	0.0742	0.592
2	0.0097	0.0743	0.593
3	0.0096	0.0735	0.5938
4	0.0096	0.073	05942
5	0.0096	0.0738	0.5943
6	0.0096	0.0736	0.5944
7	0.0096	0.0735	0.5944
8	0.0096	0.0739	0.5947
9	0.0096	0.0739	0.5949
10	0.0096	0.0736	05961
Mean	0.0096	0.0735	0.5984
Standard deviations	0.0000421637	0.000283	0.0010993

Table R5. RMSE, MAE, R^2 of predicted ER for XGBT model on training sets.

Number	RMSE	MAE	R^2
1	0.3018	0.4412	0.947
2	0.133	01985	0.9776
3	0.1338	0.2007	0.9775
4	0.2236	0.3949	0.9594
5	0.2208	0.3797	0.9635
6	0.2267	0.3821	0.9625
7	0.2463	0.3994	0.9549
8	0.228	0.3829	0.9553
9	0.2381	0.3427	0.9598
10	0.151	0.2058	0.9716
Mean	0.2096	0.3302	0.9620
Standard deviations	0.0541	0.0936	0.01

2026 S.T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

Table R6. RMSE, MAE, R^2 of predicted ER for XGBT model on test sets.

Number	RMSE	MAE	R^2
1	0.3327	0.402	0.9275
2	0.2772	0.3689	0.9396
3	0.2704	0.4142	0.9411
4	0.3677	0.4095	0.9199
5	0.3477	0.3906	0.9206
6	0.3526	0.3865	0.8796
7	0.3567	0.3921	0.9036
8	0.3452	0.3863	0.9295
9	0.3372	0.4105	0.9227
10	0.2727	0.3667	0.9412
Mean	0.326	0.3927	0.9225
Standard deviations	0.0376	0.0166	0.0191

Table R7. RMSE, MAE, R^2 of predicted EE for XGBT model on training sets.

Number	RMSE	MAE	R^2
1	0.0037	0.0442	0.8404
2	0.0044	0.0486	0.8419
3	0.0038	0.0411	0.8585
4	0.0037	0.0442	0.8693
5	0.0035	0.0306	0.8761
6	0.0036	0.0417	0.8752
7	0.0043	0.0471	0.8571
8	0.0038	0.0405	0.8612
9	0.0043	0.0478	0.8505
10	0.0039	0.0437	0.8526
Mean	0.0039	0.043	0.8302
Standard deviations	0.0003	0.0052	0.0113

Table R8. RMSE, MAE, R^2 of predicted EE for XGBT model on test sets.

Number	RMSE	MAE	R^2
1	0.0101	0.0746	0.5797
2	0.0109	0.0785	0.5406
3	0.0104	0.0715	0.5242
4	0.0101	0.0746	0.5793
5	0.0098	0.0725	0.5957
6	0.0108	0.0796	0.5525
7	0.011	0.0767	0.5626
8	0.0111	0.0811	0.5409

2026 Si-Ti Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

9	0.0108	0.0779	0.5511
10	0.0105	0.0742	0.543
Mean	0.0106	0.0765	0.5598
Standard deviations	0.0004	0.0031	0.0221

Table R9. RMSE, MAE, R^2 of predicted ER for GBDT model on training sets.

Number	RMSE	MAE	R^2
1	0.2807	0.354	0.9486
2	0.2528	0.3296	0.9537
3	0.2398	0.3251	0.9542
4	0.295	0.3549	0.9461
5	0.3136	0.3905	0.9537
6	0.2353	0.3336	0.9514
7	0.2766	0.3325	0.9476
8	0.2655	0.336	0.9514
9	0.2671	0.3402	0.9505
10	0.2589	0.3191	0.9661
Mean	0.2785	0.3416	0.9503
Standard deviations	0.0332	0.0206	0.008

Table R10. RMSE, MAE, R^2 of predicted ER for GBDT model on test sets.

Number	RMSE	MAE	R^2
1	0.5068	0.4464	0.91
2	0.5541	0.5137	0.8898
3	0.5732	0.5103	0.8856
4	0.5178	0.4801	0.8976
5	0.6047	0.5344	0.8788
6	0.5509	0.5045	0.8904
7	0.6301	0.5581	0.8733
8	0.6106	0.462	0.8992
9	0.5657	0.4849	0.8951
10	0.5485	0.4968	0.8975
Mean	0.5562	0.4991	0.8907
Standard deviations	0.0398	0.033	0.0091

Table R11. RMSE, MAE, R^2 of predicted EE for GBDT model on training sets.

Number	RMSE	MAE	R^2
1	0.0034	0.0339	0.8545
2	0.0028	0.03	0.8708

2026 S.T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

3	0.0027	0.0302	0.8653
4	0.0028	0.0302	0.8766
5	0.003	0.0313	0.8703
6	0.0027	0.0297	0.8615
7	0.0029	0.0308	0.8727
8	0.0027	0.0309	0.8752
9	0.0027	0.0302	0.8531
10	0.0027	0.0306	0.8593
Mean	0.0028	0.0308	0.8642
Standard deviations	0.0002	0.0012	0.0085

Table R12. RMSE, MAE, R^2 of predicted EE for GBDT model on test sets.

Number	RMSE	MAE	R^2
1	0.0106	0.073	0.538
2	0.0104	0.0745	0.5462
3	0.0109	0.0685	0.5685
4	0.0103	0.0727	0.5501
5	0.0105	0.0742	0.5426
6	0.0109	0.0705	0.5631
7	0.0102	0.0712	0.5579
8	0.0101	0.071	0.5616
9	0.0099	0.0678	0.5723
10	0.010	0.0689	0.5697
Mean	0.0108	0.0781	0.5511
Standard deviations	0.0003	0.0024	0.0121

Table R13. RMSE, MAE, R^2 of predicted ER for SVM model on training sets.

Number	RMSE	MAE	R^2
1	0.216	0.4909	0.8928
2	0.2179	0.4855	0.8925
3	0.217	0.4906	0.8927
4	0.2135	0.4943	0.8933
5	0.2133	0.4961	0.8933
6	0.2153	0.4968	0.8916
7	0.2121	0.4929	0.8961
8	0.2146	0.4975	0.8912
9	0.2198	0.4876	0.8935
10	0.2173	0.4912	0.8937
Mean	0.2180	0.4939	0.8697
Standard deviations	0.0024	0.004	0.0013

2026 S.T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

Table R14. RMSE, MAE, R^2 of predicted ER for SVM model on test sets.

Number	RMSE	MAE	R^2
1	0.6058	0.6088	0.8705
2	0.6134	0.6048	0.8689
3	0.6013	0.607	0.8715
4	0.6282	0.617	0.8657
5	0.6357	0.6195	0.864
6	0.6128	0.6067	0.8689
7	0.6305	0.6158	0.8672
8	0.6185	0.6029	0.8651
9	0.6243	0.6102	0.8696
10	0.6131	0.6079	0.8703
Mean	0.6184	0.61	0.8682
Standard deviations	0.0112	0.0055	0.0025

Table R15. RMSE, MAE, R^2 of predicted EE for SVM model on training sets.

Number	RMSE	MAE	R^2
1	0.0019	0.0221	0.91
2	0.0018	0.0175	0.9122
3	0.0019	0.022	0.9101
4	0.0018	0.0175	0.9122
5	0.0018	0.0175	0.9112
6	0.0018	0.0175	0.9122
7	0.0019	0.022	0.9106
8	0.0018	0.0175	0.9115
9	0.0019	0.022	0.9105
10	0.0018	0.0175	0.9122
Mean	0.0018	0.0193	0.9113
Standard deviations	0.0001	0.0023	0.0009

Table R16. RMSE, MAE, R^2 of predicted EE for SVM model on test sets.

Number	RMSE	MAE	R^2
1	0.0076	0.0715	0.6272
2	0.0077	0.0721	0.6217
3	0.0076	0.0715	0.6271
4	0.0078	0.0657	0.6171
5	0.0078	0.0656	0.6181
6	0.0077	0.0649	0.6248
7	0.0076	0.0653	0.6237
8	0.0076	0.0647	0.6265

2026 S. T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

9	0.0077	0.0653	0.6189
10	0.0078	0.0656	0.6186
Mean	0.0077	0.0651	0.6240
Standard deviations	0.0001	0.0005	0.004

To clarify integration weights, I would like to provide quantitative data to clarify the weight determination method for the voting regressor, as shown in Tables R17-R20. The weights for each base model in the voting regressor were determined based on their performance on the validation set (specifically, the R^2). The voting regressor was configured as a soft voting ensemble using the voting regressor class from scikit-learn, with weights set proportional to each model's cross-validated R^2 score. The tables R17-R20 show the cross-validation R^2 values and assigned weights for each base model.

Table R17. RMSE, MAE, R^2 of predicted ER for VR model on training sets based on different weights of each model.

Weights of each model (SVM:XGBT:GBDT)	**RMSE**	**MAE**	**R^2**
5:8:5	0.2187	0.2919	0.9558
1:2:2	0.2051	0.2829	0.9583
1:5:5	0.1725	0.2694	0.9642
1:5:7	0.1806	0.2712	0.9627
1:5:10	0.201	0.2775	0.959
1:10:5	0.1516	0.2623	0.968
1:2:10	0.2372	0.2906	0.9524
1:10:2	0.2383	0.2608	0.9704
1:10:6	0.1557	0.2632	0.9672
1:10:4	0.1473	0.2615	0.9688
Mean	0.1908	0.27313	0.96268
Standard deviations	0.0343123	0.0119856	0.0060904

Table R18. RMSE, MAE, R^2 of predicted ER for VR model on test sets based on different weights of each model.

Weights of each model (SVM:XGBT:GBDT)	**RMSE**	**MAE**	**R^2**
5:8:5	0.3345	0.4015	0.9284
1:2:2	0.3305	0.3984	0.9293
1:5:5	0.3174	0.3795	0.9321
1:5:7	0.2978	0.3603	0.9364
1:5:10	0.3044	0.3718	0.9349
1:10:5	0.2956	0.3599	0.9369
1:2:10	0.3244	0.3965	0.9306

2026 S.-T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

1:10:2	0.305	0.3673	0.9348
1:10:6	0.2944	0.3598	0.9371
1:10:4	0.2974	0.3607	0.9365
Mean	0.31014	0.37557	0.9337
Standard deviations	0.0152647	0.0172505	0.00332

Table R19. RMSE, MAE, R^2 of predicted EE for VR model on training sets based on different weights of each model.

Weights of each model (SVM:XGBT:GBDT)	**RMSE**	**MAE**	**R^2**
4:2:15	0.002	0.0279	0.9128
2:1:2	0.002	0.0275	0.9136
3:1:6	0.0022	0.029	0.9039
3:1:9	0.0021	0.0286	0.9057
3:1:13	0.0022	0.0289	0.9038
4:1:13	0.0022	0.0292	0.9021
4:1:14	0.0022	0.0295	0.9005
4:1:15	0.0021	0.0282	0.9084
4:1:16	0.002	0.0276	0.9117
5:1:16	0.0018	0.0263	0.9191
Mean	0.00208	0.02827	0.90816
Standard deviations	0.0001317	0.000971	0.0059815

Table R20. RMSE, MAE, R^2 of predicted EE for VR model on test sets based on different weights of each model.

Weights of each model (SVM:XGBT:GBDT)	**RMSE**	**MAE**	**R^2**
4:2:15	0.0076	0.0632	0.722
2:1:2	0.0076	0.0633	0.723
3:1:6	0.0075	0.0625	0.7238
3:1:9	0.0075	0.063	0.7242
3:1:13	0.0075	0.0628	0.7243
4:1:13	0.0075	0.0626	0.7244
4:1:14	0.0075	0.0625	0.7244
4:1:15	0.0075	0.0629	0.7247
4:1:16	0.0075	0.0629	0.7249
5:1:16	0.0075	0.0626	0.7261
Mean	0.00752	0.06283	0.72418
Standard deviations	0.0000421637	0.000283	0.0010993

2026 S.T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

Appendix C: Experimental Process.



Fig. B1. Experimental tests of solar water evaporation.

2026 S.-T. Yau High School Award
仅用于2026丘成桐中学科学奖

Appendix D: 半导体光热材料性能研究的机器学习算法源代码（用 Python 语言编写，程序太多只附上了部分）

1、半导体多重插补处理缺失值 2.py

#处理缺失值程序

```
import numpy as np      #安装 numpy 库用于科学计算
import pandas as pd     #安装 pandas 库加载数据，并使用 numpy 进行数值操作
from sklearn.experimental import enable_iterative_imputer
from sklearn.impute import IterativeImputer
from sklearn.linear_model import BayesianRidge

pd.set_option('display.max_rows', None)
pd.set_option('display.max_columns', None)

filename = '半导体光热脱盐材料特征数据收集(终）.csv'
data = pd.read_csv(filename,encoding='gbk')        #加载数据
print(data.head())        #查看数据
print(data.describe())
print(data.info())
print(data.shape)

# 将 DataFrame 转换为 NumPy 数组
#data1 = data.drop(['吸光材料','文献'],axis=1)  #假设['吸光材料','文献']是目标变量
data1 = data.replace({pd.NaT:np.nan})
data2 = data1.to_numpy()

# 初始化 IterativeImputer
imputer = IterativeImputer(BayesianRidge(), max_iter=10)

# 拟合数据并生成插补的数据
imputed_data = imputer.fit_transform(data2)

# 将插补的数据转换回 Pandas DataFrame
imputed_df = pd.DataFrame(imputed_data, columns=data1.columns)

imputed_df.to_csv('半导体插补数据.csv',index=False,encoding='gbk')
```

2、GBDT_model_training_best1.py

GBDT-ER 模型训练的程序—半导体光热材料的蒸发速率数据的机器学习

```
import numpy as np
import pandas as pd
```


2026 Si-Yu High School Science Award 仅用于2026年成桐中学科学奖论文公示

```
from sklearn.model_selection import train_test_split,cross_val_score
from sklearn.metrics import mean_squared_error,mean_absolute_error,r2_score
from sklearn.ensemble import GradientBoostingRegressor
from sklearn.model_selection import KFold
from sklearn.model_selection import GridSearchCV
from skopt import gp_minimize
from skopt.space import Real, Integer
from sklearn.preprocessing import StandardScaler
from matplotlib import pyplot as plt

plt.rcParams['font.family'] = ['Microsoft YaHei']
plt.rcParams['axes.unicode_minus'] = False    # 正确显示负号

#导入数据
filename = '半导体插补数据.csv'
df = pd.read_csv(filename,encoding='gbk')
df2 = df.iloc[:,range(7)]
data = df.iloc[:,[2,3,4,5,6,7]] #所需要的材料特征
target = df.iloc[:,[0,1]] #前面两列
#target1 = df.iloc[:,0]
#target2 = df.iloc[:,1]

X = data.values
#Y_1 = target1.values
#Y_2 = target2.values
Y = target.values

#标准化数据，归一化特征
X = StandardScaler().fit_transform(X)

X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=0.2, random_state=10)
Y_train_1 = Y_train[:,0]
Y_train_2 = Y_train[:,1]
Y_test_1 = Y_test[:,0]
Y_test_2 = Y_test[:,1]

# 创建梯度提升回归模型
gb_reg = GradientBoostingRegressor()

# 定义超参数搜索空间
param_space = {
    'n_estimators': Integer(50, 300),
    'learning_rate': Real(0.01, 0.3),
    'max_depth': Integer(2, 5),
    'alpha': Real(0.5, 0.9),
    'min_samples_split': Integer(2, 10),
```

2026 S.T.Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

```
    'min_samples_leaf': Integer(2, 10),
}

# 转换超参数空间格式
dimensions = [
    param_space['n_estimators'],
    param_space['learning_rate'],
    param_space['max_depth'],
    param_space['alpha'],
    param_space['min_samples_split'],
    param_space['min_samples_leaf'],
]

# 定义贝叶斯优化目标函数
def objective(params):
    n_estimators = params[0]
    learning_rate = params[1]
    max_depth = params[2]
    alpha = params[3]
    min_samples_split = params[4]
    min_samples_leaf = params[5]

    model = GradientBoostingRegressor(
        n_estimators=n_estimators,
        learning_rate=learning_rate,
        max_depth=max_depth,
        alpha=alpha,
        min_samples_split=min_samples_split,
        min_samples_leaf=min_samples_leaf,
        random_state=42
    )

    #训练模型

# 使用交叉验证评估模型，对于回归问题使用负均方误差作为评分标准
    scores = cross_val_score(model, X_train, Y_train_2, cv=5,
scoring='neg_mean_squared_error')
    # 返回负平均分数，因为 gp_minimize 尝试最小化目标函数
    return -scores.mean()

# 执行贝叶斯优化
res = gp_minimize(
    objective,
    dimensions,
    n_calls=300,
    random_state=42,
```

2026 Siyuan High School Science Award
仅用于2026年成都中学科学奖论文公示

```
        n_jobs=-1
)

best_params = {
        'n_estimators': res.x[0],
        'learning_rate': res.x[1],
        'max_depth': res.x[2],
        'alpha': res.x[3],
        'min_samples_split': res.x[4],
        'min_samples_leaf': res.x[5],
}

print("最佳超参数：", best_params)

# 使用最佳超参数训练模型
best_model = GradientBoostingRegressor(
        n_estimators=best_params['n_estimators'],
        learning_rate=best_params['learning_rate'],
        max_depth=best_params['max_depth'],
        alpha=best_params['alpha'],
        min_samples_split=best_params['min_samples_split'],
        min_samples_leaf=best_params['min_samples_leaf'],
        random_state=42
)
best_model.fit(X_train, Y_train_2)

# 使用最佳模型进行预测

# 训练数据的预测
y_train_pred2 = best_model.predict(X_train)

# 测试数据的预测
predictions_2 = best_model.predict(X_test)
#result = pd.DataFrame([predictions_2, Y_test_2])
#result.to_csv('GBDT_PT_2.csv',index=False,encoding='gbk')

#计算训练数据的 MSE 和 MAE，评估模型性能

train_mse = mean_squared_error(Y_train_2,y_train_pred2)
train_mae = mean_absolute_error(Y_train_2,y_train_pred2)
train_r2 = r2_score(Y_train_2,y_train_pred2)

# 计算测试数据的 MSE 和 MAE
test_mse = mean_squared_error(Y_test_2, predictions_2)
test_mae = mean_absolute_error(Y_test_2, predictions_2)
test_r2 = r2_score(Y_test_2,predictions_2)
```

2026 S. T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

```
print("训练数据的 MSE:", train_mse)
print("训练数据的 MAE:", train_mae)
print("训练数据的 R2:", train_r2)
print("测试数据的 MSE:", test_mse)
print("测试数据的 MAE:", test_mae)
print("测试数据的 R2:", test_r2)

#绘制模型性能散点图
plt.figure(figsize=(12,10))
plt.plot([0,17], [0,17], color='black',linewidth=1,label='best line')
plt.scatter(Y_train_2, y_train_pred2,label='train data',color='lightblue',s=80)
plt.scatter(Y_test_2, predictions_2,label='test data',color='lightcoral',s=80)
plt.xlabel('True',fontsize=30,fontweight='bold')
plt.ylabel('Predicted',fontsize=30,fontweight='bold')
plt.title('GBDT',fontsize=24,fontweight='bold')
plt.legend(fontsize=20)
plt.xticks(fontsize=20,fontweight='bold')
plt.yticks(fontsize=20,fontweight='bold')
ax = plt.gca()
ax.spines['bottom'].set_linewidth(3)
ax.spines['left'].set_linewidth(3)
ax.spines['top'].set_linewidth(3)
ax.spines['right'].set_linewidth(3)
plt.show()
```

3、SVM_model_training_best1.py

 # SVM-ER 模型的训练程序—半导体光热材料的蒸发速率数据的机器学习

```
import numpy as np
import pandas as pd
from sklearn.model_selection import train_test_split,cross_val_score
from sklearn.metrics import mean_squared_error,mean_absolute_error,r2_score
from sklearn.svm import SVR
from sklearn.model_selection import KFold
from sklearn.model_selection import GridSearchCV
from hyperopt import hp, fmin, tpe, Trials
from sklearn.preprocessing import StandardScaler
from matplotlib import pyplot as plt

plt.rcParams['font.family'] = ['Microsoft YaHei']
plt.rcParams['axes.unicode_minus'] = False    # 正确显示负号

#导入数据
filename = '半导体插补数据.csv'
```

2026 S_T_Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

```
df = pd.read_csv(filename,encoding='gbk')
df2 = df.iloc[:,range(7)]
data = df.iloc[:,[2,3,4,5,6,7]]
target = df.iloc[:,[0,1]]
#target1 = df.iloc[:,0]
#target2 = df.iloc[:,1]

X = data.values
#Y_1 = target1.values
#Y_2 = target2.values
Y = target.values#标准化数据
X = StandardScaler().fit_transform(X)

X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=0.2, random_state=10)
Y_train_1 = Y_train[:,0]
Y_train_2 = Y_train[:,1]
Y_test_1 = Y_test[:,0]
Y_test_2 = Y_test[:,1]

# 创建五折交叉验证
#kf = KFold(n_splits=5, shuffle=True, random_state=100)

# 创建 SVM 模型
svm = SVR()

# 定义超参数搜索空间
space = {
    'C': hp.loguniform('C', np.log(1e-3), np.log(1e3)),  # 正则化参数
    'epsilon': hp.loguniform('epsilon', np.log(1e-3), np.log(1e3)),  # ε - 不敏感管的宽度
    'gamma': hp.loguniform('gamma', np.log(1e-3), np.log(1e3)),  # 核函数的系数
    'kernel': hp.choice('kernel', ['linear', 'poly', 'rbf', 'sigmoid'])  # 核函数类型
}

# 定义目标函数
def objective(params):
    params = {
        'C': params['C'],
        'epsilon': params['epsilon'],
        'gamma': params['gamma'],
        'kernel': params['kernel']
    }

    # 使用交叉验证来评估模型性能
    model = SVR(**params)
    scores = cross_val_score(model, X_train, Y_train_2, cv=5, scoring='r2')  # 使用 R 平方作为评估指标
```

2026 S.T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

```
        loss = -np.mean(scores)    # 因为 hyperopt 是用于最小化目标函数，所以取负值

        return loss

# 运行贝叶斯优化
trials = Trials()
best = fmin(
        fn=objective,
        space=space,
        algo=tpe.suggest,
        max_evals=50,    # 最大迭代次数
        trials=trials
)

# 输出最佳超参数
best_kernel = ['linear', 'poly', 'rbf', 'sigmoid'][best['kernel']]
best_params = {
        'C': best['C'],
        'epsilon': best['epsilon'],
        'gamma': best['gamma'],
        'kernel': best_kernel
}
print("最佳超参数：", best_params)

# 使用最佳超参数训练最终模型
best_model = SVR(**best_params)

# 使用最佳模型进行预测
best_model.fit(X_train, Y_train_2)

# 训练数据的预测
y_train_pred2 = best_model.predict(X_train)

# 测试数据的预测
predictions_2 = best_model.predict(X_test)
#result = pd.DataFrame([predictions_2, Y_test_2])
#result.to_csv('GBDT_PT_2.csv',index=False,encoding='gbk')

#计算训练数据的 MSE 和 MAE
train_mse = mean_squared_error(Y_train_2,y_train_pred2)
train_mae = mean_absolute_error(Y_train_2,y_train_pred2)
train_r2 = r2_score(Y_train_2,y_train_pred2)

# 计算测试数据的 MSE 和 MAE
test_mse = mean_squared_error(Y_test_2, predictions_2)
test_mae = mean_absolute_error(Y_test_2, predictions_2)
```

2026 S-T Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

```
test_r2 = r2_score(Y_test_2,predictions_2)

print("训练数据的 MSE:", train_mse)
print("训练数据的 MAE:", train_mae)
print("训练数据的 R2:", train_r2)
print("测试数据的 MSE:", test_mse)
print("测试数据的 MAE:", test_mae)
print("测试数据的 R2:", test_r2)

#绘制模型性能散点图
plt.figure(figsize=(12,10))
plt.plot([0,17], [0,17], color='black',linewidth=1,label='best line')
plt.scatter(Y_train_2, y_train_pred2,label='train data',color='lightblue',s=30)
plt.scatter(Y_test_2, predictions_2,label='test data',color='lightcoral',s=30)
plt.xlabel('True',fontsize=30,fontweight='bold')
plt.ylabel('Predicted',fontsize=30,fontweight='bold')
plt.title('SVM',fontsize=24,fontweight='bold')
plt.legend(fontsize=20)
plt.xticks(fontsize=20,fontweight='bold')
plt.yticks(fontsize=20,fontweight='bold')
ax = plt.gca()
ax.spines['bottom'].set_linewidth(3)
ax.spines['left'].set_linewidth(3)
ax.spines['top'].set_linewidth(3)
ax.spines['right'].set_linewidth(3)
plt.show()
```

4、XGboosting_model_training_best1.py

```
#    XGBT-ER 模型的训练程序—半导体光热材料的蒸发速率数据的机器学习

import numpy as np
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.metrics import mean_squared_error,mean_absolute_error,r2_score
from sklearn.model_selection import KFold,cross_val_score
from sklearn.model_selection import GridSearchCV
from sklearn.preprocessing import StandardScaler
from skopt import gp_minimize
from skopt.space import Real, Integer
from xgboost import XGBRegressor
from matplotlib import pyplot as plt

pd.set_option('display.max_rows', None)
pd.set_option('display.max_columns', None)
```

2026 S.T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

```
plt.rcParams['font.family'] = ['Microsoft YaHei']
plt.rcParams['axes.unicode_minus'] = False    # 正确显示负号

#导入数据
filename = '半导体插补数据.csv'
df = pd.read_csv(filename,encoding='gbk')
df2 = df.iloc[:,range(2,7)]
data = df.iloc[:,[2,3,4,5,6,7]]
target = df.iloc[:,[0,1]]
#target1 = df.iloc[:,0]
#target2 = df.iloc[:,1]

X = data.values
#Y_1 = target1.values
#Y_2 = target2.values
Y = target.values

#标准化数据
X = StandardScaler().fit_transform(X)
X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=0.2, random_state=10)
Y_train_1 = Y_train[:,0]
Y_train_2 = Y_train[:,1]
Y_test_1 = Y_test[:,0]
Y_test_2 = Y_test[:,1]

# 初始化 XGBoost 回归模型
xgb_model = XGBRegressor()

# 定义超参数搜索空间
param_space = {
    'learning_rate': Real(0.01, 0.8),
    'max_depth': Integer(3, 6),
    'n_estimators': Integer(50, 300),
    'gamma': Real(0, 0.5),
    'subsample': Real(0.5, 0.8),
    'reg_alpha': Real(0, 1)
}

dimensions = [
    param_space['learning_rate'],
    param_space['max_depth'],
    param_space['n_estimators'],
    param_space['gamma'],
    param_space['subsample'],
    param_space['reg_alpha']
]
```

2026 Si Yau High School Science Award
仅用于2026年成桐中学科学奖论文公示

```
# 定义贝叶斯优化目标函数
def objective(params):
    learning_rate = params[0]
    max_depth = params[1]
    n_estimators = params[2]
    gamma = params[3]
    subsample = params[4]
    reg_alpha = params[5]

    model = XGBRegressor(
        learning_rate=learning_rate,
        max_depth=max_depth,
        n_estimators=n_estimators,
        gamma=gamma,
        subsample=subsample,
        reg_alpha=reg_alpha,
        random_state=42
    )

    scores = cross_val_score(model, X_train, Y_train_2, cv=5,
scoring='neg_mean_squared_error')
    return -scores.mean()

# 执行贝叶斯优化
res = gp_minimize(
    objective,
    dimensions,
    n_calls=30,
    random_state=42,
    n_jobs=-1
)

best_params = {
    'learning_rate': res.x[0],
    'max_depth': res.x[1],
    'n_estimators': res.x[2],
    'gamma': res.x[3],
    'subsample': res.x[4],
    'reg_alpha': res.x[5]
}

print("最佳超参数：", best_params)

# 使用最佳超参数训练模型
best_model = XGBRegressor(
    learning_rate=best_params['learning_rate'],
```

2026 S.-T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

```
        max_depth=best_params['max_depth'],
        n_estimators=best_params['n_estimators'],
        gamma=best_params['gamma'],
        subsample=best_params['subsample'],
        reg_alpha=best_params['reg_alpha'],
        random_state=42
)
best_model.fit(X_train, Y_train_2)
# 使用最佳模型进行预测

# 训练数据的预测
y_train_pred2 = best_model.predict(X_train)

# 测试数据的预测
predictions_2 = best_model.predict(X_test)
#result = pd.DataFrame([predictions_2, Y_test_2])
#result.to_csv('GBDT_PT_2.csv',index=False,encoding='gbk')

#计算训练数据的 MSE 和 MAE
train_mse = mean_squared_error(Y_train_2,y_train_pred2)
train_mae = mean_absolute_error(Y_train_2,y_train_pred2)
train_r2 = r2_score(Y_train_2,y_train_pred2)

# 计算测试数据的 MSE 和 MAE
test_mse = mean_squared_error(Y_test_2, predictions_2)
test_mae = mean_absolute_error(Y_test_2, predictions_2)
test_r2 = r2_score(Y_test_2,predictions_2)

print("训练数据的 MSE:", train_mse)
print("训练数据的 MAE:", train_mae)
print("训练数据的 R2:", train_r2)
print("测试数据的 MSE:", test_mse)
print("测试数据的 MAE:", test_mae)
print("测试数据的 R2:", test_r2)

#绘制模型性能散点图
plt.figure(figsize=(12,10))
plt.plot([0,17], [0,17], color='black',linewidth=1,label='best line')
plt.scatter(Y_train_2, y_train_pred2,label='train data',color='lightblue',s=80)
plt.scatter(Y_test_2, predictions_2,label='test data',color='lightcoral',s=80)
plt.xlabel('True',fontsize=30,fontweight='bold')
plt.ylabel('Predicted',fontsize=30,fontweight='bold')
plt.title('XGBT',fontsize=24,fontweight='bold')
plt.legend(fontsize=20)
plt.xticks(fontsize=20,fontweight='bold')
plt.yticks(fontsize=20,fontweight='bold')
```

2026 Siyu High School Science Award
仅用于2026年思源中学科学奖论文公示

```
ax = plt.gca()
ax.spines['bottom'].set_linewidth(3)
ax.spines['left'].set_linewidth(3)
ax.spines['top'].set_linewidth(3)
ax.spines['right'].set_linewidth(3)
plt.show()
```

5、Voting_model_training_best1.py

```
#    VR-ER 模型的训练程序—半导体光热材料的蒸发速率数据的机器学习

from sklearn.model_selection import train_test_split
from sklearn.ensemble import VotingRegressor
from sklearn.metrics import mean_squared_error,mean_absolute_error,r2_score
from sklearn.preprocessing import StandardScaler
from sklearn.ensemble import GradientBoostingRegressor
from sklearn.svm import SVR
from xgboost import XGBRegressor
from matplotlib import pyplot as plt
import numpy as np
import pandas as pd

plt.rcParams['font.family'] = ['Microsoft YaHei']
plt.rcParams['axes.unicode_minus'] = False   # 正确显示负号

#导入数据
filename = '半导体插补数据.csv'
df = pd.read_csv(filename,encoding='gbk')
df2 = df.iloc[:,range(7)]
data = df.iloc[:,[2,3,4,5,6,7]]
target = df.iloc[:,[0,1]]
#target1 = df.iloc[:,0]
#target2 = df.iloc[:,1]

X = data.values
#Y_1 = target1.values
#Y_2 = target2.values
Y = target.values

#标准化数据
X = StandardScaler().fit_transform(X)

X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=0.2, random_state=10)
Y_train_1 = Y_train[:,0]
Y_train_2 = Y_train[:,1]
```

2026 S. T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

```
Y_test_1 = Y_test[:,0]
Y_test_2 = Y_test[:,1]

# 创建模型
svr = SVR(C=42.448789, kernel='rbf', epsilon=0.005106,gamma=0.004153)
xgb = XGBRegressor(gamma=0.491157,learning_rate=0.227627,n_estimators=289,max_depth=4,
                   reg_alpha=1.0,subsample=0.613525)
gb = GradientBoostingRegressor(alpha=0.9, learning_rate=0.3, n_estimators=50, max_depth=2,
                               min_samples_leaf=2, min_samples_split=2)

# 创建投票集成模型
voting_regressor = VotingRegressor([
      ('svr', svr),
      ('xgb', xgb),
       ('gb', gb)
],weights=[80,800,20])

voting_regressor.fit(X_train,Y_train_2)
# 训练数据的预测
y_train_pred2 = voting_regressor.predict(X_train)

# 测试数据的预测
predictions_2 = voting_regressor.predict(X_test)

#计算训练数据的 MSE 和 MAE
train_mse = mean_squared_error(Y_train_2,y_train_pred2)
train_mae = mean_absolute_error(Y_train_2,y_train_pred2)
train_r2 = r2_score(Y_train_2,y_train_pred2)

# 计算测试数据的 MSE 和 MAE
test_mse = mean_squared_error(Y_test_2, predictions_2)
test_mae = mean_absolute_error(Y_test_2, predictions_2)
test_r2 = r2_score(Y_test_2,predictions_2)

print("训练数据的 MSE:", train_mse)
print("训练数据的 MAE:", train_mae)
print("训练数据的 R2:", train_r2)
print("测试数据的 MSE:", test_mse)
print("测试数据的 MAE:", test_mae)
print("测试数据的 R2:", test_r2)

#绘制模型性能散点图
plt.figure(figsize=(12,10))
```

2026 Yau High School Science Award
仅用于2026亚成杯中学科学奖论文公示

```
plt.plot([0,17], [0,17], color='black',linewidth=1,label='best line')
plt.scatter(Y_train_2, y_train_pred2,label='train data',color='lightblue',s=80)
plt.scatter(Y_test_2, predictions_2,label='test data',color='lightcoral',s=80)
plt.xlabel('True',fontsize=30,fontweight='bold')
plt.ylabel('Predicted',fontsize=30,fontweight='bold')
plt.title('VR',fontsize=24,fontweight='bold')
plt.legend(fontsize=20)
plt.xticks(fontsize=20,fontweight='bold')
plt.yticks(fontsize=20,fontweight='bold')
ax = plt.gca()
ax.spines['bottom'].set_linewidth(3)
ax.spines['left'].set_linewidth(3)
ax.spines['top'].set_linewidth(3)
ax.spines['right'].set_linewidth(3)
plt.show()
```

10、Shap_value_VOTE_vra.py

```
#     Shap-ER 模型的训练程序（用 VR 预测）-半导体光热材料的蒸发效率数据的机器学习
import numpy as np
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from sklearn.ensemble import GradientBoostingRegressor
from sklearn.svm import SVR
from xgboost import XGBRegressor
from sklearn.ensemble import VotingRegressor
from matplotlib import pyplot as plt
import shap

plt.rcParams['font.family'] = ['Microsoft YaHei']
plt.rcParams['axes.unicode_minus'] = False    # 正确显示负号

#导入数据
filename = '半导体插补数据.csv'
df = pd.read_csv(filename,encoding='gbk')
df2 = df.iloc[:,range(7)]
data = df.iloc[:,[2,3,4,5,6,7]]
target = df.iloc[:,[0,1]]
#target1 = df.iloc[:,0]
#target2 = df.iloc[:,1]

X = data.values
#Y_1 = target1.values
#Y_2 = target2.values
Y = target.values
#标准化数据
```

2026 SJT · Yau Hing School Science Award
仅用于2026丘成桐中学科学奖论文公示

```
X = StandardScaler().fit_transform(X)

X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=0.2, random_state=10)
Y_train_1 = Y_train[:,0]
Y_train_2 = Y_train[:,1]
Y_test_1 = Y_test[:,0]
Y_test_2 = Y_test[:,1]

# 创建模型
svr = SVR(C=42.448789, kernel='rbf', epsilon=0.005106,gamma=0.004153)
xgb = XGBRegressor(gamma=0.491157,learning_rate=0.227627,n_estimators=289,max_depth=4,
                   reg_alpha=1.0,subsample=0.613525)
gb = GradientBoostingRegressor(alpha=0.9, learning_rate=0.3, n_estimators=50, max_depth=2,
                               min_samples_leaf=2, min_samples_split=2)

# 创建投票集成模型
voting_regressor = VotingRegressor([
    ('svr', svr),
    ('xgb', xgb),
     ('gb', gb)
],weights=[80,800,20])

voting_regressor.fit(X_train,Y_train_2)

def vr_predict(X):
    return voting_regressor.predict(X)

# 使用 Shap 库计算 Shap 值
explainer = shap.PermutationExplainer(vr_predict,X_train)
shap_values = explainer.shap_values(X_train)
#print(shap_values)
# 显示 Shap 值的密度散点图
shap.summary_plot(shap_values,X_train,feature_names=data.columns)
```

11、Shap_value_VOTE_cve.py

 # Shap-EE 模型的训练程序（用 VR 预测）--半导体光热材料的蒸发效率数据的机器学习

```
import numpy as np
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from sklearn.ensemble import GradientBoostingRegressor
from sklearn.svm import SVR
from xgboost import XGBRegressor
```

2026 Siyau High School Science Award
仅用于2026年成桐中学科学奖论文公示

```
from sklearn.ensemble import VotingRegressor
from matplotlib import pyplot as plt
import shap

plt.rcParams['font.family'] = ['Microsoft YaHei']
plt.rcParams['axes.unicode_minus'] = False    # 正确显示负号

#导入数据
filename = '半导体插补数据.csv'
df = pd.read_csv(filename,encoding='gbk')
df2 = df.iloc[:,range(7)]
data = df.iloc[:,[2,3,4,5,6,7]]
target = df.iloc[:,[0,1]]
#target1 = df.iloc[:,0]
#target2 = df.iloc[:,1]

X = data.values
#Y_1 = target1.values
#Y_2 = target2.values
Y = target.values

#标准化数据
X = StandardScaler().fit_transform(X)

X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=0.2, random_state=1)
Y_train_1 = Y_train[:,0]
Y_train_2 = Y_train[:,1]
Y_test_1 = Y_test[:,0]
Y_test_2 = Y_test[:,1]

# 创建模型
svr = SVR(C=4.170161, kernel='rbf', epsilon=0.019841,gamma=0.044303)
xgb                                                                                    =
XGBRegressor(gamma=0.265641,learning_rate=0.8,n_estimators=237,max_depth=6,
                                reg_alpha=0.0,subsample=0.657072)
gb  =  GradientBoostingRegressor(alpha=0.9,  learning_rate=0.080013,  n_estimators=76,
max_depth=2,
                                                    min_samples_leaf=4, min_samples_split=2)

# 创建投票集成模型
voting_regressor = VotingRegressor([
      ('svr', svr),
      ('xgb', xgb),
       ('gb', gb)
],weights=[30,5,5])
```

2026 S.T.Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示

```
voting_regressor.fit(X_train,Y_train_1)

def vr_predict(X):
    return voting_regressor.predict(X)

# 使用 Shap 库计算 Shap 值
explainer = shap.PermutationExplainer(vr_predict,X_train)
shap_values = explainer.shap_values(X_train)
#print(shap_values)
# 显示 Shap 值的密度散点图
shap.summary_plot(shap_values,X_train,feature_names=data.columns)
```

12、数据相关性绘图 1.py

```
# Feature Correlation Heatmap 的 ER 相关性热图程序
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

plt.rcParams['font.family'] = ['Microsoft YaHei']
plt.rcParams['axes.unicode_minus'] = False  # 正确显示负号

#导入数据
filename = '半导体插补数据.csv'
df = pd.read_csv(filename,encoding='gbk')
df1 = df.iloc[:,[1,2,3,4,5,6,7]]
df2 = df.iloc[:,[0,2,3,4,5,6,7]]
data = df.iloc[:,range(2,7)]
target = df.iloc[:,[0,1]]
target1 = df.iloc[:,0]
target2 = df.iloc[:,1]

X = data.values
Y_1 = target1.values
Y_2 = target2.values

#相关性分析
# 计算特征之间的相关性矩阵
correlation_matrix = df1.corr()
# 使用热力图可视化相关性矩阵
plt.figure(figsize=(8, 6))
#for (i, j), z in np.ndenumerate(correlation_matrix):
    #plt.text(j, i, f'{z:.2f}', ha='center', va='center', fontsize=12)
```

略，程序太多只附上部分代码。

2026 S.T. Yau High School Science Award
仅用于2026丘成桐中学科学奖论文公示