The harvard organic photovoltaic dataset

Provided by the Springer Nature SharedIt content-sharing initiative Organic photovoltaic (OPV) materials are promising candidates for cheap, printable solar cells. However, there are a very large number of potential donors and acceptors, making selection of the best materials difficult.

The Harvard Clean Energy Project (CEP) is a distributed computing effort for screening OPV candidates carried out by volunteers connected to the IBM World Community Grid. 39 The first release of CEP screening data for a high-throughput virtual screening (HTVS) featured 1.3 million donor materials from an initial 26 common fragments in the organic ...

Organic photovoltaic (OPV) materials are of great interest because of their potential to generate cheap, printable semiconductor devices that convert light into electrical energy. They promise sustainable sources of clean energy if their efficiencies and stabilities can be improved.

Harvard Organic Photovoltaic Dataset. The second major dataset used in this work is the HOPV15 dataset, which consists of 350 different experimentally characterized organic solar cell donor structures that have been collated from various studies in the literature by Lopez et al. 22 The HOPV15 data is accessible as a single file from the ...

The Harvard organic photovoltaic dataset - experimental and quantum-mechanical PV data phonondb - database containing phonon band structures, DOS, and thermal properties for hundreds of materials PubChemQC PM6 dataset - database containing 221 million molecules with optimized molecular geometries and electronic properties

DFT-computed dataset- Harvard CEP to build more robust predictive models for relatively smaller HOPV datasets. Harvard ... Among current solar cell design paradigms, organic pho-tovoltaic cell technology is a promising technology for the inexpensive and ...

This Perspective introduces the Harvard Clean Energy Project (CEP), a theory-driven search for the next generation of organic solar cell materials, and gives a broad overview of its setup and infrastructure, present first results, and outline upcoming developments. This Perspective introduces the Harvard Clean Energy Project (CEP), a theory-driven search for ...

Performance of the Models Trained on CEPDB Data in Predicting the PCE Values of Organic Photovoltaics The results in Table 1 suggest that all models can achieve a reasonably good fit for the CEPDB data, with the largest test set MSE being 0.569 for random forests.

The Harvard Organic Photovoltaic Dataset: Conflicts of interest There are no conflicts to declare. Acknowledgements The authors acknowledge the support from the National Natural Science Foundation of China (No. 21672023, 21971014, and 21950410533) and the National Key Research and Development

Program of China (2018YFA0901800). Jin-Liang ...

The Harvard organic photovoltaic dataset. Sci. Data 3, 160086 (2016). Article Google Scholar Scharber, M. C. et al. Design rules for donors in bulk-heterojunction solar cells--towards 10% ...

However, it has been noted that computational predictions of PCE of OSCs often do not agree well with experimental measurements, and that machine learning approaches like Gaussian process regression are necessary to improve ...

The Harvard organic photovoltaic dataset. Sci. Data 3, 160086 (2016). ... Paul, A. et al. Transfer learning usingensemble neural networks for organic solar cell screening.

Here, we used the Harvard Photovoltaic Dataset (HOPV15) dataset 22 that includes data from quantum chemical calculations and that calculated by the Scharber model plus experimental

The HOPV datasets consist of the "Harvard Organic Photovoltaic Dataset. This dataset includes 350 small molecules and polymers that were utilized as p-type materials in OPVs. Experimental properties include: HOMO [a.u.], LUMO [a.u.], Electrochemical gap [a.u.], Optical gap [a.u.], Power conversion efficiency [%], Open circuit potential [V ...

The deep learning model was then fine-tuned (phase II, Figure 2) on a subset (194 molecules) of the Harvard organic photovoltaic dataset (HOPV15) with the weights being carried over from the first Phase I. The HOPV15 dataset consisted of around 350 molecules whose HOMO and LUMO levels were been i) experimentally measured (extracted from ...

The Harvard Organic Photovoltaic Dataset (HOPV15) presented in this work is a collation of experimental photovoltaic data from the literature, and corresponding quantum-chemical calculations performed over a range of conformers, each with quantum chemical results using a variety of density functionals and basis sets.

Here, we used the Harvard Photovoltaic Dataset (HOPV15) dataset 22 that includes data from quantum chemical calculations and that calculated by the Scharber model plus ...

The Harvard organic photovoltaic dataset. Scientific data 3, 1 (2016), 1--7. Google Scholar [32] Steven A Lopez, Benjamin Sanchez-Lengeling, Julio de Goes Soares, and Alán Aspuru-Guzik. 2017. Design principles and top non-fullerene acceptor candidates for organic photovoltaics. Joule 1, 4 (2017), 857--870. Crossref.

A wide variety of machine learning algorithms have been applied to predict the performance of organic photovoltaics using different target datasets. The Harvard clean energy project database (CEPDB), (8) is one such target dataset for ML models that contains computationally determined PCE values for 2.3 million organic photovoltaic candidates.

The harvard organic photovoltaic dataset

The Harvard Organic Photovoltaic Dataset (HOPV15) presented in this work is a collation of experimental photovoltaic data from the literature, and corresponding quantum-chemical calculations performed over a range of geometries, each with quantum chemical results using a variety of density functionals and basis sets. It is anticipated that this ...

However, it has been noted that computational predictions of PCE of OSCs often do not agree well with experimental measurements, and that machine learning approaches like Gaussian process regression are necessary to improve agreement.8,20,21 As a consequence, experimental OPV datasets, such as the Harvard organic photovoltaic dataset (HOPV15 ...

In this paper, the ability of five machine learning models and HDMR to predict the PCE of organic photovoltaics based on molecular structure information is assessed, including the impact and ...

Harvard organic photovoltaic 15 dataset (HOPV15). It was found that the neural-based. models generally performed better on the computational dataset with the attentive FP.

The Harvard Organic Photovoltaic Dataset (HOPV15) presented in this work is a collation of experimental photovoltaic data from the literature, and corresponding quantum-chemical calculations performed over a range of conformers, each with quantum chemical results using a variety of density functionals and basis sets. It is anticipated that this ...

The power conversion efficiencies of organic photovoltaics (OPVs) have grown tremendously over the last 20 years and represent a low-cost and sustainable solution for harnessing solar energy to power our residences, workplaces, and devices. Fullerene-containing OPVs are relatively expensive and have limited overlap absorbance with the solar spectrum. We used density ...

This Perspective introduces the Harvard Clean Energy Project (CEP), a theory-driven search for the next generation of organic solar cell materials. We give a broad overview of its setup and infrastructure, present first results, and outline upcoming developments. CEP has established an automated, high-throughput, in silico framework to study potential candidate ...

The Harvard Organic Photovoltaic Dataset (HOPV15) presented in this work is a collation of experimental photovoltaic data from the literature, and corresponding quantum-chemical calculations ...

A detailed python assisted statistical and visualization analysis of Harvard organic photovoltaic dataset is reported. Many new information are found. Structural similarity is performed to find the divergence between compounds in dataset and reference small molecule acceptors. Famous experimentally reported non-fullerene small acceptors are ...

An ensemble deep neural network architecture, called SINet, is presented, which harnesses both the SMILES

and InChI molecular representations to predict HOMO values and leverage the potential of transfer learning from a sizeable DFT-computed dataset- Harvard CEP to build more robust predictive models for relatively smaller HOPV datasets.

State-of-the-art results are achieved on the NREL OPV and Harvard CEP datasets. ... 2006) were developed for screening organic monomers used for photovoltaic applications and predicting organic solar cell efficiency (Lee, 2020, Paul et al., 2019c).

The data for machine learning analysis is retrieved from Harvard organic photovoltaic database [17]. This data consists of calculated HOMO and LUMO values of 51,000 molecules. The performance of machine learning models is strongly affected by quantity and quality of dataset [33]. This dataset is much large.

Performance of Models Trained on HOPV15 Data in Predicting the PCE Values of Organic Photovoltaics a Provided model accuracy metrics are given as a mean across all m -folds and the error bars are given as s, which for a normal distribution would correspond to a confidence level of 68%.