International Journal of Scientific Engineering and Research (IJSER)
ISSN (Online): 2347-3878
SJIF (2025): 8.036

ML-Based Bayesian Optimization of Indoor Perovskite Solar Cells for IoT Integration

Mitashi Jain[1], Lakshya[2]

[1]Department of Energy Science and Engineering, Indian Institute of Technology Delhi, Hauz Khas, New Delhi – 110016, India
Email: *es1230177[at]iitd.ac.in*

[2]Department of Biochemical Engineering and Biotechnology, Indian Institute of Technology Delhi, Hauz Khas, New Delhi – 110016, India
Email: *bb1230593[at]iitd.ac.in*

Abstract: ***The rapid growth of Internet of Things (IoT) devices has created a need for maintenance-free indoor power sources that can replace primary batteries. Indoor photovoltaics (IPVs) based on metal-halide perovskites are well suited to this role because their bandgap can be tuned to the narrow emission spectra of LED and fluorescent lamps. This paper reports an integrated machine-learning (ML) and experimental workflow for indoor perovskite solar cells (IPSCs). A curated dataset drawn from the Perovskite Database Project, published literature and in-house laboratory records was used to train a gradient-boosted forward model that predicts optical bandgap from A-, B- and X-site composition, achieving R^2 = 0.90–0.92 with a mean absolute error below 0.05 eV. The trained model was then embedded in an Optuna-driven Bayesian inverse search of 10,000 trials to locate compositions with a bandgap near the indoor-optimal value of 1.8 eV. The recommended composition $MA_{0.37}FA_{0.31}Cs_{0.31}PbI_{1.33}Br_{0.44}Cl_{1.23}$ ($E_g \approx 1.812$ eV) was cross-validated against independently retrained models. Triple-cation n–i–p devices fabricated on FTO/TiO_2 substrates delivered power conversion efficiencies of 14.08–24.7 % under AM 1.5 G illumination and successfully drove an LED load, demonstrating feasibility for self-powered IoT nodes.***

Keywords: Indoor photovoltaics, perovskite solar cells, Bayesian optimization, machine learning, IoT energy harvesting

1. Introduction

Wireless sensor networks and wearable electronics are now being deployed at a rate that primary batteries cannot sustainably support: cells have finite lifetimes, require periodic replacement and generate disposal waste. Indoor photovoltaics (IPVs), which harvest ambient light from LED and fluorescent lamps, offer a maintenance-free alternative for the milliwatt-scale loads typical of IoT hardware. Metal-halide perovskites are particularly attractive in this niche because of their tunable bandgap, high absorption coefficient and strong low-light performance, combined with solution processability and low-temperature fabrication [6], [7]. Crystalline silicon, with an indirect bandgap of 1.12 eV and high-temperature processing, is poorly matched to indoor spectra, which are concentrated between roughly 400 and 700 nm (Figure 1); for such sources the optimum absorber bandgap shifts upward to about 1.8 eV. This work therefore targets indoor perovskite solar cells (IPSCs) through a three-stage workflow: ML-driven Bayesian prediction of composition, fabrication and characterisation of the predicted material, and assessment of its suitability for powering low-energy IoT devices.

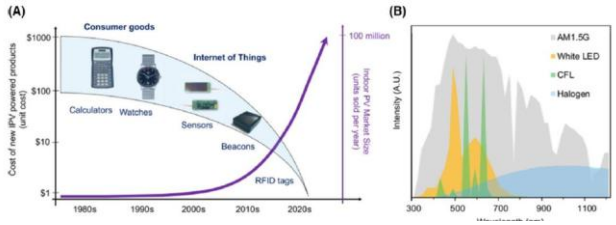

Figure 1: (a) Falling cost and rising volume of IPV-powered consumer and IoT products; (b) emission spectra of common indoor light sources compared with the AM 1.5 G solar spectrum.

2. Objective 1: AI–ML-Driven Prediction Models

The first objective was to build Bayesian-optimisation models capable of predicting the composition and processing parameters best matched to indoor illumination, so that experimental effort could be concentrated on a small number of promising candidates.

2.1 Dataset construction and preprocessing

The training set combined three sources: experimental records from the group's laboratory archive, verified entries from the Perovskite Database Project [1], and bandgap–composition pairs extracted from internal and external publications. All records were standardised into a common schema of A-site (MA, FA, Cs), B-site (Pb, Sn) and X-site (I, Br, Cl) fractions. Rows with missing bandgap values were dropped rather than imputed, which raised test accuracy from roughly 50 % to 87 %. A site-sum validation step retained only chemically valid stoichiometries ($\Sigma A = 1$, $\Sigma B = 1$, $\Sigma X = 3$ within a tolerance of 1×10^{-3}); near-duplicate rows within the same tolerance were merged to remove redundancy, and a correlation matrix was used to prune highly correlated features while preserving compositional diversity. The cleaned, normalised dataset was exported to CSV for training [2].

2.2 Forward bandgap prediction model

The dataset was split 80/20 into training and test subsets. Logistic regression underfitted the strongly nonlinear coupling between ionic composition and bandgap, while support-vector and k-nearest-neighbour regressors proved sensitive to feature scaling and generalised poorly at high computational cost. Bagging and boosting ensembles- random forest, extra trees, XGBoost, LightGBM and CatBoost- were therefore benchmarked instead. Hyperparameters were tuned

with Optuna's Bayesian sampler rather than grid or random search, balancing exploration against exploitation; each of 100 trials was scored by five-fold cross-validation and minimised mean absolute error while tracking R^2. The best model (XGBoost, with LightGBM close behind) reached $R^2 = 0.90–0.92$ on held-out data with MAE < 0.05 eV [3].

SHAP (SHapley Additive exPlanations) analysis was used to interpret the trained model (Figure 2). Iodine content is by a clear margin the most influential descriptor, followed by formamidinium and bromine, with caesium providing moderate fine-tuning; lead and tin contribute in opposite directions, as expected from their differing ionic radii and orbital energies. These rankings agree with established composition–bandgap trends and with the laboratory's own measurements, supporting the physical plausibility of the surrogate model rather than leaving it as a black box.

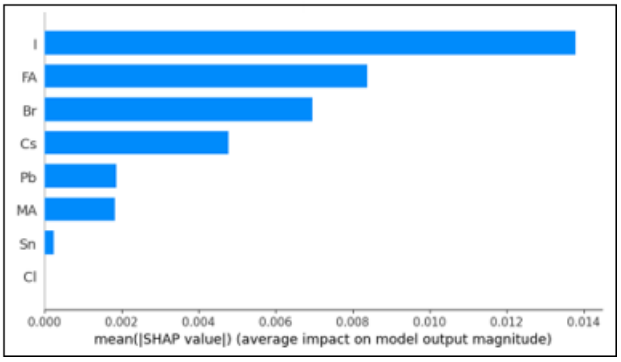

Figure 2: SHAP feature-importance ranking showing the mean absolute contribution of each ionic component to the predicted bandgap

2.3 Inverse design by Bayesian optimisation

The trained forward model was then used as an oracle inside an inverse search for compositions with a target bandgap of 1.8 eV, the value that maximises the Shockley–Queisser limit under typical indoor illumination. Optuna minimised the objective

$$\text{Error} = | E_{predicted} - 1.8 | \quad (1)$$

Each trial generated a candidate composition, normalised it to $\Sigma A = 1$, $\Sigma B = 1$ and $\Sigma X = 3$ to preserve physical stoichiometry, and evaluated its predicted bandgap; chemically invalid candidates were penalised with a fixed error of 10.0. Tin was excluded from the search space because the available laboratory setup permits only Pb-based synthesis in air. After 10,000 trials the 100 candidates closest to the target were retained for validation.

2.4 Validation and final composition

To guard against overfitting to any single data source, the forward model was retrained on experimental data from independent publications and on internal laboratory data, and only models achieving $R^2 > 0.90$ were kept. The bandgaps of the 100 shortlisted candidates were re-predicted with each retrained model, absolute errors against 1.8 eV were computed and sorted, and the composition recurring most frequently among the top-ranked predictions across datasets was adopted as the final recommendation [4].

As summarised in Table 1, the resulting composition is $MA_{0.37}FA_{0.31}Cs_{0.31}PbI_{1.33}Br_{0.44}Cl_{1.23}$, whose predicted bandgap of 1.8123 eV corresponds to an error of only 0.0123 eV against the target, and which lies within the composition space reachable with the precursors available in the laboratory.

Table 1: ML-optimised indoor perovskite composition

Ion / metric	Value	Ion / metric	Value
MA	0.37	I	1.33
FA	0.31	Br	0.44
Cs	0.31	Cl	1.23
Pb	1	Predicted E_g	1.8123 eV
Sn	0	Absolute error	0.0123 eV

3. Objective 2: PSC Fabrication and Characterization

3.1 Device fabrication

Cells were built in the n–i–p architecture on FTO-coated glass (≈15 Ω/sq, 2.5 × 2.5 cm), later masked into six pixels of ≈0.09 cm² each. Substrates were scrubbed with laboratory detergent, rinsed in DI water, sonicated sequentially in Hellmanex, DI water, acetone and isopropanol, dried under N_2 and treated with UV-ozone for 15 min; the decreasing-polarity solvent sequence removes organic and ionic residues without leaving a film, while UV-ozone oxidises residual hydrocarbons and hydroxylates the FTO, improving wetting and interfacial adhesion.

A compact TiO_2 electron transport layer was deposited from titanium diisopropoxide bis(acetylacetonate) in isopropanol, spin-coated at 3000 rpm for 30 s in a 0.15 M / 0.30 M / 0.15 M sequence with 130 °C, 5 min interlayer anneals, then crystallised to anatase at 500 °C for 1 h in air. A subsequent $TiCl_4$ treatment hydrolyses to an ultrathin conformal oxide inside residual pores, giving a smoother and more continuous ETL.

The triple-cation absorber $Cs_{0.05}(FA_{0.83}MA_{0.17})_{0.95}Pb(I_{0.83}Br_{0.17})_3$ was prepared in DMF:DMSO (4:1), in which DMSO coordinates Pb^{2+} to form a PbI_2·DMSO adduct that slows crystallisation and enlarges grains. A two-step spin programme (1200 rpm for 15 s, then 5000 rpm for 30 s) with ≈200 µL of chlorobenzene antisolvent dripped at the fifteenth second of the second step produced smooth, pinhole-free films. These were dried at 100 °C for 10 min and annealed at 150 °C for 30 min to convert the solvated intermediate into the crystalline α-phase:

$$FAI + PbI_2 + MABr + CsI \rightarrow Cs_{0.05}(FA_{0.83}MA_{0.17})_{0.95}Pb(I_{0.83}Br_{0.17})_3 \quad (2)$$

Spiro-OMeTAD doped with Li-TFSI and 4-tert-butylpyridine (72 mg per mL of chlorobenzene, filtered at 0.45 µm immediately before use) was spin-coated at 3000–4000 rpm for 30 s to a target thickness of 150–200 nm. Approximately 80 nm of gold was then thermally evaporated through a six-pixel shadow mask at a base pressure ≤ 5×10^{-6} mbar, beginning with a slow 0.1–0.2 Å s^{-1} pre-wetting ramp to avoid metal spiking through microscopic pinholes in the HTL. The complete process flow is shown in Figure 3.

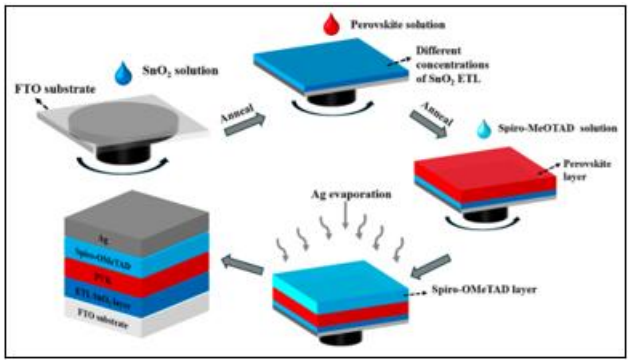

Figure 3: Sequential deposition steps for the n–i–p perovskite solar cell, from FTO cleaning through ETL and absorber coating to metal evaporation

3.2 Characterization

Current–voltage characteristics were recorded on an Ossila solar cell I–V measurement system under 1000 W m^{-2} (AM 1.5 G) illumination, sweeping from −0.5 V to +1.18 V in 5 mV steps with a 0.05 s settling delay in reverse-scan mode; pixels were switched manually across the 2 × 3 array. Open-circuit voltage, short-circuit current density, fill factor, maximum power point, series and shunt resistance and power conversion efficiency were extracted for each pixel.

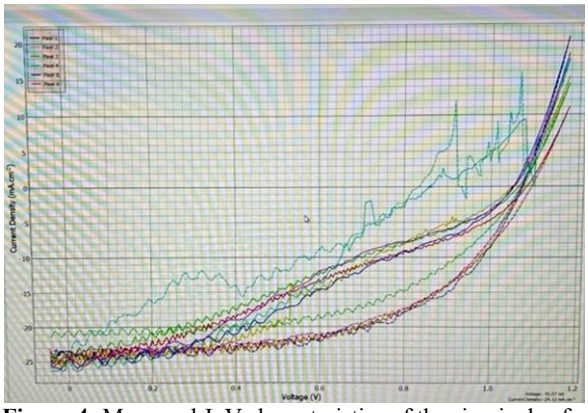

Figure 4: Measured J–V characteristics of the six pixels of a fabricated device under AM 1.5 G illumination (reverse scan).

Representative curves for all six pixels are shown in Figure 4. Power conversion efficiency ranged from 14.08 % to 24.7 %, with maximum-power-point output of roughly 1–3.5 mW per 0.09 cm^2 pixel, open-circuit voltages above 1.0 V and short-circuit current densities near 24 mA cm^{-2}. The spread between pixels is attributed to local film-uniformity differences and contact-resistance variation introduced during spin-coating rather than to any compositional difference, since all six pixels share a single absorber film. These figures set the load budget for the integration study described next.

4. Objective 3: IoT Integration

The final stage assessed whether the fabricated cells could act as a self-charging supply for small, low-power electronics such as wearable health monitors, smart watches and environmental sensors.

4.1 Matching cells to IoT loads

Compact IoT nodes typically draw 30–200 mW while active and up to 1–2 W during charging pulses, whereas a single 0.09 cm^2 pixel supplies a few milliwatts. Direct continuous operation is therefore not realistic at this device area; instead the cells suit duty-cycled sensing, in which pixels are interconnected into a larger module and harvest continuously into an energy buffer that is discharged in short bursts. Heart-rate and SpO_2 sensors, fitness trackers, and wireless temperature and humidity nodes all fit this intermittent-charging profile, and it is exactly the regime in which the low-light efficiency and solution processability of perovskites are most valuable.

4.2 Experimental demonstration

To verify energy transfer out of the device, a pre-encapsulated cell of the type shown in Figure 5 was wired on a breadboard with one terminal on the FTO front contact and the other on the gold back contact. Under 1-Sun illumination (≈100 mW cm^{-2}) the cell drove several LEDs simultaneously, with brightness tracking the incident intensity [5]. This confirms photo-induced charge generation, extraction and delivery into an external load, and provides a functional bridge toward a self-sustaining IoT node.

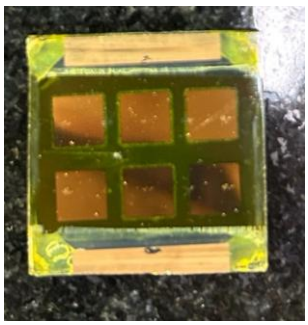

Figure 5: Encapsulated six-pixel perovskite module used for the LED demonstration.

4.3 Practical Constraints

Two constraints limited the integration stage. Perovskite deposition requires an ambient relative humidity below about 40 %; persistent moisture during the fabrication window repeatedly forced deposition and annealing runs to be paused or restarted. Second, encapsulation- essential to protect the absorber from oxygen, moisture and prolonged illumination- could not be completed within the project timeline, so integration with IoT hardware was deferred rather than risk rapid degradation of unprotected films.

4.4 Future Scope

- **Power management:** buck–boost DC–DC conversion with a supercapacitor or rechargeable buffer, to turn the variable instantaneous DC output into a regulated supply that supports simultaneous harvesting, storage and load operation.
- **Encapsulation and miniaturisation:** UV-curable epoxy or polymeric barrier films, followed by lamination of scaled-down modules onto flexible substrates for smartwatch and wearable-patch form factors.

- **Self-charging prototype:** mini modules based on the validated composition ($E_g \approx 1.81$ eV) driving heart-rate, temperature and motion sensors for continuous, battery-free operation.
- **Indoor characterisation:** repeat measurement under white LED and fluorescent sources at realistic indoor illuminance to obtain a true indoor PCE and long-term stability data.

5. Results

1) A Bayesian-optimised gradient-boosting forward model predicts perovskite bandgap from composition with R^2 = 0.90–0.92 and MAE < 0.05 eV; SHAP attributes most of the model output to iodine and formamidinium content, consistent with experiment and literature.
2) A 10,000-trial inverse search identified $MA_{0.37}FA_{0.31}Cs_{0.31}PbI_{1.33}Br_{0.44}Cl_{1.23}$ (E_g = 1.8123 eV, error 0.0123 eV) as the best Sn-free, Pb-based composition for indoor operation, reproducible across independently retrained models.
3) Triple-cation n–i–p cells fabricated with the optimised process gave power conversion efficiencies of 14.08–24.7 % across six pixels under AM 1.5 G, with stable V_{oc} and J_{sc}, confirming good film quality and process reproducibility.
4) Encapsulated cells delivered ≈1–3.5 mW per pixel and drove an LED load on a breadboard, demonstrating that the devices can power duty-cycled IoT sensing once pixels are interconnected and buffered.

6. Discussion

Coupling a forward surrogate model to a Bayesian inverse search converts material selection from an exhaustive experimental sweep into a targeted one: 10,000 candidate compositions were screened computationally in minutes, and only the most frequently recurring candidate was carried to the bench. The agreement between the predicted bandgap of 1.81 eV and the fabricated triple-cation film is the principal validation of the approach, and the SHAP ranking supplies an interpretable, physically consistent account of why the model behaves as it does.

Two caveats temper the result. First, the reported efficiencies were measured under AM 1.5 G rather than indoor illumination, so they characterise material and process quality but do not yet constitute an indoor figure of merit; performance under a low-lux LED source must be measured before indoor PCE can be claimed. Second, the spread from 14.08 % to 24.7 % across nominally identical pixels indicates that process reproducibility, not composition, is currently the limiting factor- which is itself an argument for spending experimental effort on deposition control now that the composition search is largely automated.

7. Challenges and Learnings

The dominant practical constraint throughout was humidity. Perovskite crystallisation is strongly moisture-sensitive, and even under dehumidified conditions fabrication had to be paused whenever external humidity exceeded the system's capacity, producing incomplete films or visible degradation. Minor defects- dust particles, fingerprints, uneven coatings - frequently rendered a device non-functional and required a full restart from substrate cleaning, which makes purely experimental optimisation both slow and fragile.

This is precisely the regime in which data-driven design repays its cost: by narrowing the composition space before any material is consumed, the ML stage reduced the number of fabrication cycles required to reach a working device. The broader lesson is that computational prediction and experimental validation are complementary — the model constrains the search and the bench decides — and that for indoor perovskite photovoltaics the remaining bottlenecks are environmental control, encapsulation and power conditioning rather than material discovery.

References

[1] The Perovskite Database Project. [Online]. Available: https://www.perovskitedatabase.com/. [Accessed: Oct. 20, 2025].
[2] M. Jain and Lakshya, "Data collection and preprocessing notebook," Google Colab, 2025. [Online]. Available: https://colab.research.google.com/drive/1y00UGccQNcSHFwrg757_5FJEoFcWUdQi
[3] M. Jain and Lakshya, "Bandgap prediction and inverse-design notebook," Google Colab, 2025. [Online]. Available: https://colab.research.google.com/drive/12AcLAFjkMIOCUTeGGFRRFA9wjIFL6_SY
[4] M. Jain and Lakshya, "Cross-validation and final composition derivation notebook," Google Colab, 2025. [Online]. Available: https://colab.research.google.com/drive/1Bur-K5-1b-UeLLo3guBM6zXm0yYIzLU4
[5] M. Jain and Lakshya, "LED demonstration of a fabricated perovskite solar cell," video recording, 2025. [Online]. Available: https://drive.google.com/file/d/1Fn4MJTmxv3_SxW3J6eHW76QfXWUoqXP8/view
[6] A. K. Jena, A. Kulkarni and T. Miyasaka, "Halide perovskite photovoltaics: Background, status, and future prospects," Chemical Reviews, vol. 119, no. 5, pp. 3036–3103, 2019.
[7] S. Bhattarai et al., "A detailed review of perovskite solar cells: Introduction, working principle, modelling, fabrication techniques, future challenges," Micro and Nanostructures, vol. 172, art. no. 207450, 2022.

Author Profile

Mitashi Jain is a fourth-year undergraduate student in the Department of Energy Science and Engineering at the Indian Institute of Technology Delhi (Entry No. 2023ES10177). Her research interests include machine learning for materials discovery, Bayesian optimization and inverse design, and explainable AI applied to structure–property prediction. She can be reached at es1230177[at]iitd.ac.in

Lakshya is a fourth-year undergraduate student in the Department of Biochemical Engineering and Biotechnology at the Indian Institute of Technology Delhi (Entry No. 2023BB10593). His research interests include solution-processed perovskite thin films, device fabrication and interface engineering, and indoor photovoltaics for self-powered IoT systems. He can be reached at bb1230593[at]iitd.ac.in