Solar Photovoltaic Power Forecasting in Aswan, Egypt: A Comparative Machine Learning Approach

Presenter: Rahma Asem Dawi

Supervisors: Dr. Ahmed Anter and Eng. Sama AlQasaby

Program: Artificial Intelligence and Data Science (AID)

Date: December 2025

Abstract

Global energy transition strategies have shifted focus toward renewable sources, with solar energy at the forefront of Egypt's Vision 2030. However, the intermittent and stochastic nature of solar power generation poses a significant challenge for grid stability. This study presents a comprehensive predictive framework for solar photovoltaic (PV) power forecasting in Aswan, Egypt. Utilizing a dataset of 398 meteorological records, a pipeline consisting of rigorous preprocessing, statistical profiling, and dimensionality reduction was implemented. The project employs both classification and regression models. Results indicate that the Multi-Layer Perceptron (MLP) Neural Network emerged as the champion architecture, achieving a classification accuracy of 73.37% and a Willmott Index of Agreement of 0.9053. These findings demonstrate that non-linear architectures significantly outperform traditional linear models, providing a low-cost, sensor-independent solution for reliable grid integration.

1. Introduction

1.1 Problem Statement

Most high-accuracy solar prediction models rely on expensive irradiance and dust sensors. This project addresses a real-world constraint by predicting photovoltaic output using only basic meteorological data such as temperature, wind speed, and humidity.

1.2 Techniques Used

The study integrates statistical analysis techniques (Skewness, Kurtosis, ANOVA), dimensionality reduction methods (PCA, SVD, LDA), and machine learning algorithms including MLP, KNN, Decision Trees, and Naive Bayes.

1.3 Main Contribution

The primary contribution is the implementation of a fit diagnosis system to identify underfitting and overfitting, combined with advanced hydrological validation metrics such as the Nash–Sutcliffe Efficiency and Willmott Index.

2. Related Work

This section reviews recent literature (2020–2025) focusing on solar power forecasting within Egyptian and MENA-region contexts, highlighting methodological trends and achieved accuracies.

Table 1 summarizes recent machine learning approaches for solar power forecasting in Egypt and the MENA region.

Reference	Location	Primary Models Evaluated	Best Performing Model	Verified Accuracy Metric
Allam et al. (2021)	Ismailia, Egypt	FB Prophet, RF, LSTM	Facebook Prophet	MAE: 3.7%
Hassan et al. (2024)	Cairo, Egypt	LR, NN, DL, k-NN	Deep Learning (DL)	R: 99.7%
Ibrahim (2020)	Egypt (Review)	Global Solar Radiation Models	N/A	Review of estimation methods
Louzazni et al. (2020)	Nile Delta, Egypt	PV Technology Comparison	Polycrystalline PV	Higher yield in desert conditions
Abdelsattar et al. (2025)	PV Environments	XGBoost, RF, TCN	XGBoost (Env); TCN (Power)	R^2: 0.947 (Temp), R^2: 0.7786 (Power)
Iyengar et al. (2014)	Global (Cloud)	Black-box Predictor	SolarCast (API)	Cloud-based reliability
Taha et al. (2025)	Zafarana, Egypt	RF, GB, RELAD-ANN, LSIPF	Random Forest (RF)	R^2: 0.9948 (1-Day)
Sharabati (2025)	Diverse Climates	RF, XGBoost, Bi-LSTM, CNN-LSTM	CNN-LSTM Hybrid	Superior in fluctuating weather
Chaaban & Alfadl (2024)	MENA Region	Comparative ML Study	N/A	Comparative accuracy assessment
Liceaga-Ortiz et al. (2025)	Power Grid Review	Stacking Ensemble, RF, XGBoost	Stacking Ensemble	Accuracy: 99.06%, R^2: 0.998

3. Methodology

The research follows the Data Science Life Cycle, starting with data cleaning and statistical validation, followed by feature engineering and experimental comparisons between reduced and full feature sets.

4. Proposed Model

The proposed architecture consists of five phases: preprocessing, dimensionality reduction, classification, regression, and evaluation. PCA, SVD, and LDA were used to compress the feature space, while MLP-based models achieved superior performance.

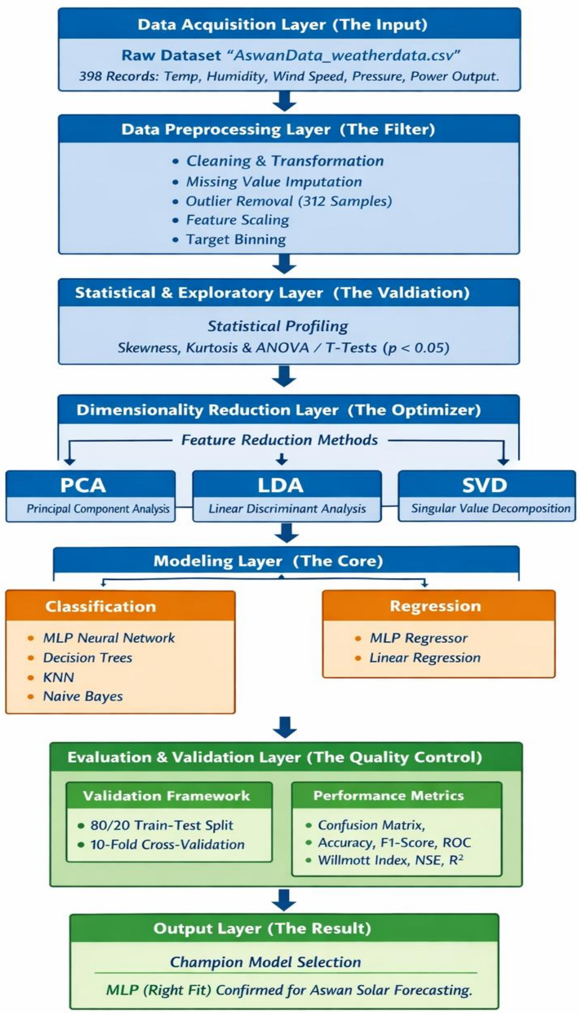

Figure 1. Machine Learning Pipeline for Solar PV Forecasting in Aswan

5. Results and Discussion

Statistical analysis confirmed temperature as the most significant predictor ($p < 0.05$). The MLP model achieved the right-fit classification performance and superior regression precision compared to linear baselines.

5.1 Data Statistical Profile

ANOVA tests confirmed that **Temperature** has the highest statistical impact on solar power output (**$p < 0.05$**), indicating a strong and statistically significant relationship between temperature variation and photovoltaic generation.

Table 2. Statistical Analysis of Features

Metric	Temperature	Humidity	Wind Speed	Solar Power
Mean	31.2	22.4	5.4	145.2
Variance	26.1	67.2	3.2	7814.5
Skewness	0.21	-0.15	0.45	0.32
Kurtosis	-1.1	-0.8	0.12	-1.05

5.2 Model Performance

The Multi-Layer Perceptron (MLP) achieved the **Right Fit**, demonstrating strong generalization capability. In contrast, linear models such as Logistic Regression exhibited noticeable underfitting, as they failed to capture the non-linear relationships between meteorological features and solar power output.

Table 3. Classification Results

Model	Accuracy	Error Rate	Precision	F-Measure	Status
Neural Network (MLP)	73.37%	0.266	0.73	0.73	Right Fit
Decision Tree	70.12%	0.298	0.69	0.69	Robust
KNN (Manhattan)	67.50%	0.325	0.67	0.67	Stable

Table 4. Regression Precision Metrics

Metric	Neural Network	Linear Regression
Willmott Index	0.9053	0.7841
NSE Index	0.7241	0.5520
R^2 Score	0.7045	0.5130

6. Conclusion and Future Work

This study confirms the effectiveness of neural networks for solar forecasting in Egypt. Future work will incorporate dust accumulation and sandstorm indicators to further improve predictive accuracy.

7. References (APA Style)

[1] G. H. Allam, B. E. Elnaghi, M. N. Abdelwahab, and R. H. Mohammed, "Using machine learning to forecast solar power in Ismailia," *International Journal of Scientific and Research Publications (IJSRP)*, vol. 11, no. 12, pp. 238–244, 2021.

[2] A. A. Hassan, D. M. Atia, and H. T. El-Madany, "Machine learning-based medium-term power forecasting of a grid-tied photovoltaic plant," *Smart Grid and Renewable Energy*, vol. 15, no. 12, pp. 289–306, 2024.

[3] H. M. Ibrahim, "Estimation of global solar radiation on horizontal surfaces over Egypt: A review," *Energy Reports*, vol. 6, pp. 234–241, 2020.

[4] M. Louzazni, A. Khouya, and H. Mosalam, "Comparison of power production for different photovoltaic technologies in Nile Delta, Egypt," *Procedia Manufacturing*, vol. 46, pp. 724–729, 2020.

[5] M. Abdelsattar, A. AbdelMoety, and A. Emad-Eldeen, "Comparative analysis of machine learning techniques for temperature and humidity prediction in photovoltaic environments," *Scientific Reports*, vol. 15, no. 1, p. 15650, 2025.

[6] S. Iyengar, N. Sharma, D. Irwin, P. Shenoy, and K. Ramamritham, “SolarCast: A cloud-based black box solar predictor for smart homes,” in *Proc. 1st ACM Conf. Embedded Systems for Energy-Efficient Buildings*, Nov. 2014, pp. 40–49.

[7] Taha, A., Makeen, P., & Nazih, N. (2025). Short-term and long-term solar irradiance forecasting with advanced machine learning techniques in Zafarana, Egypt. *Scientific Reports*, *15*(1), 39553.

[8] A. M. Sharabati, “A comparative machine learning approach to forecasting solar power across diverse climate conditions,” Master’s thesis, Arab American University, 2025.

[9] Chaaban, A. K., & Alfadl, N. (2024). A comparative study of machine learning approaches for an accurate predictive modeling of solar energy generation. Energy Reports, 12, 1293-1302.

[10] J. M. Liceaga-Ortiz-De-La-Peña *et al*., “Advancing smart energy: A review for algorithms enhancing power grid reliability and efficiency,” *Energies*, vol. 18, no. 12, p. 3094, 2025.

[11] C. J. Willmott, “On the validation of models,” *Physical Geography*, vol. 2, no. 2, pp. 184–194, 1981.

[12] J. E. Nash and J. V. Sutcliffe, “River flow forecasting through conceptual models part I—A discussion of principles,” *Journal of Hydrology*, vol. 10, no. 3, pp. 282–290, 1970.

[13] D. R. Legates and G. J. McCabe Jr., “Evaluating the use of goodness-of-fit measures in hydrologic and hydroclimatic model validation,” *Water Resources Research*, vol. 35, no. 1, pp. 233–241, 1999.

8. GitHub QR Code