7th International Conference on
Artificial Intelligence and Speech Technology 2025
AIST 2025

IEEE Delhi Section
Technical Sponsor

# LEVERAGING AI AND DATA-DRIVEN TECHNIQUES FOR SOLAR POWER PREDICTION {PAPER ID - 332}

AFFILIATION OF PRESENTING AUTHORS
Indira Gandhi Delhi Technical University for Women (IGDTUW)
Department of Electronics and Communication Engineering (Specialization in Artificial Intelligence)
Kashmere Gate, Delhi – 110006, India

AUTHORS - Anjali , Himanshi Kataria, Livya Jajoria
Nirma Peter, Isha Malhotra, Nidhi Goel

# INDEX

# ABSTRACT

Accurate prediction of solar energy generation is a key requirement for the reliable integration of Photovoltaic (PV) systems into modern power grids. Because PV production depends heavily on changing environmental factors including sunlight intensity, ambient temperature, moisture levels, wind patterns, and cloud movements, short-term forecasting remains a difficult challenge.

**Research Objective:** Comparative analysis of Machine Learning (ML) and Deep Learning (DL) techniques for short-term PV forecasting using an open-access multivariate dataset.

**Models Evaluated:** Random Forest, XGBoost, ANN, LSTM, GRU, 1D-CNN, and Transformer architectures.

**Key Findings:** Random Forest performed best among ML methods, while GRU achieved strong results within DL. ML models are effective for fast, resource-efficient deployment, whereas DL methods are suitable when highest accuracy is priority.

# INTRODUCTION

### Global Context

- Rapid adoption of renewable energy
- Solar as sustainable power source
- Declining PV installation costs
- Supportive government policies

### Key Challenge

- Variability of solar output due to changing weather patterns remains a critical barrier to large-scale grid integration .

### Environmental Factors

- Solar Irradiance
- Temperature
- Humidity
- Wind Speed
- Cloud Movement
- Module Temperature

### Benefits

- Stable grid operations
- Efficient plant scheduling
- Optimized storage
- Improved energy trading

# LITERATURE REVIEW

| Category | Methods/Models | Characteristics | Limitations |
|---|---|---|---|
| **Classical Methods** | Persistence, ARIMA, Linear Regression | Simple, Low computation, Fast training | Linear assumptions, Poor with variability |
| **Machine Learning** | SVR, Random Forest, XGBoost | Handles non-linearity, Multi-variate inputs | Feature engineering, Manual tuning |
| **Deep Learning** | LSTM, GRU, CNN, Transformers | Auto features, Temporal patterns, High accuracy | High computation, Data intensive |
| **Hybrid Models** | CNN-LSTM, CNN-GRU, Ensembles | Combines strengths, Enhanced accuracy | Very complex, Long training time |

TABLE I :Forecasting Models: Classical, ML, Deep Learning, and Hybrid

# RESEARCH GAPS

- **Limited Comprehensive Comparisons**

  Most studies focus on specific model types without systematic comparison across ML and DL architectures

- **Accuracy vs Efficiency Trade-off**

  Insufficient analysis of computational efficiency alongside prediction accuracy for real-world deployment

- **Real-Time Integration Gap**

  Limited demonstration of models integrated with real-time weather APIs for operational deployment

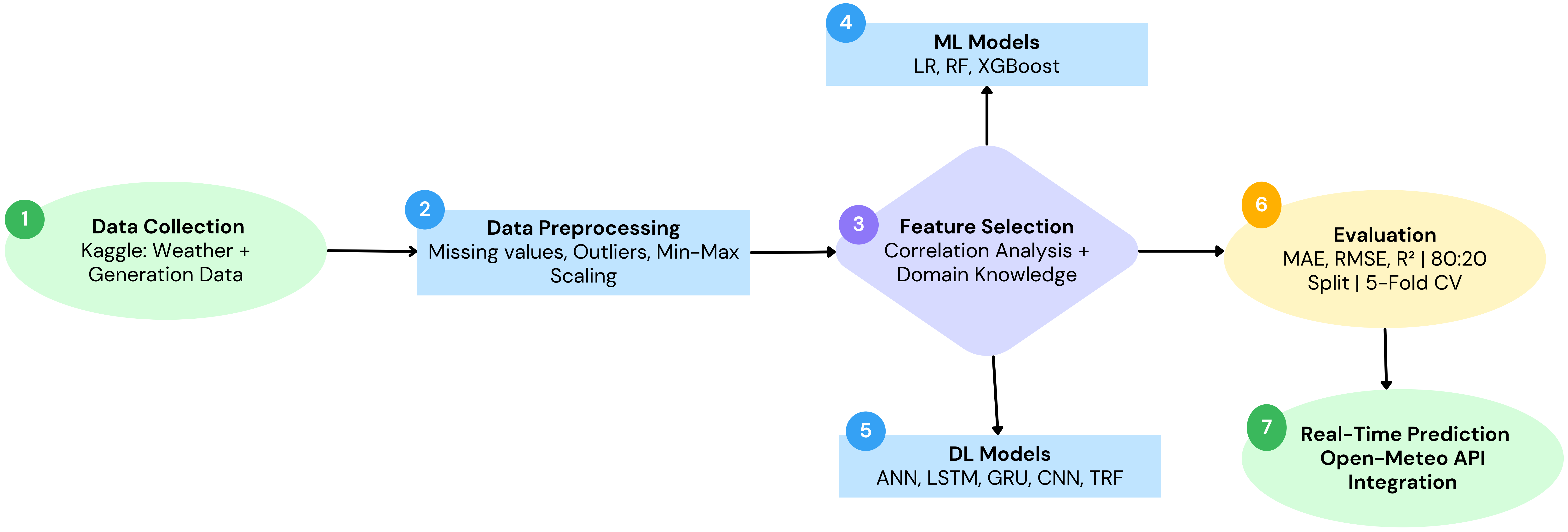
- **Resource-Constrained Deployment**

  Need for lightweight models suitable for edge devices and embedded systems

- **Transformer Model Exploration**

  Limited investigation of attention-based architectures for solar forecasting

# PROPOSED METHODOLOGY



# RESULTS & DISCUSSION

| Model | Category | MAE (kW) | RMSE (kW) | R2 Score | Time (s) |
|---|---|---|---|---|---|
| Linear Regression | ML | 268.50 | 567.03 | 0.98 | 0.5 |
| **Random Forest** | **ML** | **168.38** | **469.05** | **0.99** | **2.4** |
| XGBoost | ML | 187.36 | 472.26 | 0.99 | 3.2 |
| ANN | DL | 211.31 | 527.86 | 0.87 | 5.1 |
| LSTM | DL | 195.79 | 556.98 | 0.91 | 15.6 |
| **GRU** | **DL** | **195.27** | **557.08** | **0.92** | **13.8** |
| 1D-CNN | DL | 236.20 | 607.46 | 0.89 | 6.8 |
| Transformer | DL | 506.31 | 795.55 | 0.88 | 20.5 |

TABLE II :Model Performance Comparison: Accuracy and Training Time

# COMPARATIVE ANALYSIS

**Best ML Model**
- Random Forest
- MAE: 168.38 kW | $R^2$: 0.99
- Training: 2.4s

**Best DL Model**
- GRU
- MAE: 195.27 kW | $R^2$: 0.92
- Training: 13.8s

**Recommendation**
- Use ML (Random Forest) for production deployment: Fast, efficient, best accuracy
- Use DL (GRU) when temporal patterns are critical and resources available

**ML Advantages**
- Faster training time
- Lower resource requirements
- Better interpretability
- Best overall accuracy (RF)

**DL Advantages**
- Automatic feature learning
- Temporal pattern recognition
- Complex relationships
- Scalable to large datasets

# CONCLUSION & FUTURE WORK

| Aspect | Finding | Performance Metrics |
|---|---|---|
| **Best Overall Model** | Random Forest (ML) | MAE: 168.38 kW \| RMSE: 469.05 kW \| $ $R$: 0.99 \| Time: 2.4s |
| **Best DL Model** | GRU (Deep Learning) | MAE: 195.27 kW \| RMSE: 557.08 kW \| $ $R$: 0.92 \| Time: 13.8s |
| **ML Advantage** | Superior efficiency & accuracy | Fast training, Low resources, Best performance |
| **DL Advantage** | Temporal pattern learning | Auto features, Complex relationships, Scalable |
| **Real-Time Integration** | Open-Meteo API Success | Live weather data, Instant predictions, Production-ready |

TABLE III : Key Findings and Model Advantages in Forecasting

| Area | Proposed Work | Expected Impact |
|---|---|---|
| **Hybrid Models** | RF-LSTM, XGB-GRU ensembles | Combine ML efficiency with DL temporal learning |
| **Multi-Source Data** | Satellite imagery, radar integration | Enhanced short-term forecast accuracy |
| **Edge Deployment** | Model compression, quantization | IoT & resource-constrained devices |
| **Explainable AI** | SHAP, LIME techniques | Model transparency & interpretability |
| **Multi-Horizon** | 1-hr, 6-hr, 24-hr forecasts | Comprehensive grid planning support |
| **Grid Integration** | Smart grid optimization | Energy management system deployment |

TABLE IV : Future Work and Expected Impact for Enhanced Forecasting

# REFERENCES

**1. Solar Power Forecasting Reviews**
F. Shaker et al., "Forecasting Solar Power Generation: A Survey of Current Techniques," IEEE Access, 2019.
- Why: Overview of solar forecasting techniques

P. T. Nguyen et al., "Deep Learning in Solar Power Forecasting: A Comprehensive Review," Renewable and Sustainable Energy Reviews, 2021.
- Why: Deep learning approaches for solar prediction

M. Abuella and M. Chowdhury, "Solar Power Forecasting Using Machine Learning Techniques: A Review," Energies, 2020.
- Why: ML techniques review for solar forecasting

**2. Machine Learning Models**
T. Hastie et al., The Elements of Statistical Learning, Springer, 2009.
- Why: Foundation for Random Forest and ensemble methods

[T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," ACM SIGKDD, 2016.
- Why: XGBoost algorithm you used

M. Lari and A. Rezaei, "Solar Energy Forecasting Using Hybrid Machine Learning Models," Energy Reports, 2020.
- Why: ML models for solar forecasting

**3. Deep Learning Architectures**
I. Goodfellow et al., Deep Learning, MIT Press, 2016.
- Why: Foundation for all DL concepts (ANN, CNN)

S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Computation, 1997.
- Why: LSTM architecture you implemented

K. Cho et al., "Learning Phrase Representations Using RNN Encoder–Decoder," arXiv:1406.1078, 2014.
- Why: GRU architecture you implemented

N. Sharma et al., "Deep Learning-Based Framework for Short-Term Solar Power Prediction," Applied Energy, 2020.
- Why: DL framework for solar prediction

# REFERENCES

**4. Comparative Studies**

S. Yeong et al., "Comparative Analysis of Deep Learning Models for Solar Forecasting," Energy AI, 2021.

- Why: Comparative analysis like yours

A. Elbreki et al., "Artificial Neural Networks for Solar PV Forecasting," Renewable Energy Journal, 2021.

- Why: ANN for PV forecasting

**5. Implementation Tools**

M. Abadi et al., "TensorFlow: Large-Scale Machine Learning," arXiv:1603.04467, 2016.

- Why: Framework for DL implementation

F. Chollet, "Keras: The Python Deep Learning Library," 2018.

- Why: Keras for model building

**6. Dataset**

Anikó Komal, "Solar Power Generation Data," Kaggle, 2020.

- Why: Your actual dataset source

# QUESTION & ANSWER SESSION

**Leveraging AI and Data-Driven Techniques for Solar Power Prediction**

**Paper ID: 332 | Department of ECE (AI), IGDTUW**

## We're Ready to Answer Your Questions!

Time Allocated: 2 minutes | Team Members: Anjali, Himanshi, Livya