**DLP Project Report  
Roll No: 21K-4870  
Section: BCS-DLP 8B**

**Project Title:**

**Deep Learning-Based Forecasting of Solar Energy Generation Using Weather Data**

**1. Objective**

The objective of this project is to build predictive deep learning models—**LSTM**, **GRU**, and **Transformer**—to forecast hourly solar power generation using historical solar output and corresponding weather data.

**2. Dataset Description**

Three years of solar energy generation data (2022–2024) were combined with averaged hourly weather data (temperature, humidity, wind speed, etc.).

* Data columns included: Year, Month, Hour, Size, weather features, and Solar\_Output\_kWh.
* Preprocessing involved:
  + Dropping missing and non-numeric columns (e.g., Installation)
  + Normalizing inputs using MinMaxScaler
  + Creating 24-hour time series sequences (sliding window method)
  + Train-validation split (80:20), **without shuffling** to preserve temporal sequence

**3. Model Architectures & Training Settings**

**A. LSTM Model**

* **Architecture:**
  + LSTM (128 units, return\_sequences=True)
  + Dropout (0.2)
  + LSTM (64 units)
  + Dropout (0.2)
  + Dense (32 units, ReLU)
  + Output: Dense(1)
* **Epochs:** 30
* **Batch Size:** 64
* **Learning Rate Scheduler:** ReduceLROnPlateau
* **Loss Function:** Mean Squared Error (MSE)
* **Early Stopping:** Not used

**Results:**

* **MAE:** 16,868.04
* **RMSE:** 29,633.49
* **R²:** 0.3672
* **MAPE:** 4.81%

*Analysis:*  
LSTM performed moderately well but suffered slightly from overfitting after epoch 10. Learning rate reduction helped, but the model didn't significantly improve after that.

**B. GRU Model (Bidirectional)**

* **Architecture:**
  + Bidirectional GRU (128 units, return\_sequences=True)
  + Dropout (0.3)
  + Bidirectional GRU (64 units)
  + Dropout (0.3)
  + Dense (64 units, ReLU)
  + Output: Dense(1)
* **Epochs:** 100 (early stopping triggered at epoch 48)
* **Batch Size:** 128
* **Callbacks:** ReduceLROnPlateau + EarlyStopping
* **Loss Function:** MSE

**Results:**

* **MAE:** 11,296.44
* **RMSE:** 19,406.13
* **R²:** 0.7286
* **MAPE:** 465.37%

*Analysis:*  
GRU achieved the best R² and RMSE, indicating strong correlation and low prediction error. However, an unusually high MAPE (>400%) suggests the presence of very small actual values in some records (causing division blow-up in percentage errors). GRU's bidirectional nature likely helped it capture both forward and backward temporal dependencies.

**C. Transformer Model**

* **Architecture:**
  + Transformer Encoder Block:
    - MultiHeadAttention (4 heads, 64 dim)
    - LayerNorm + Feed Forward (128 hidden units)
    - GlobalAveragePooling
  + Dense(64, ReLU), Dropout(0.1)
  + Output: Dense(1)
* **Epochs:** 50
* **Batch Size:** 64
* **Callbacks:** ReduceLROnPlateau + EarlyStopping
* **Loss Function:** MSE

**Results (Updated):**

* **MAE:** 20,191.61
* **RMSE:** 30,871.15
* **R²:** 0.3132
* **MAPE:** 1,515.82%

*Analysis:*  
Despite achieving good validation loss during training, the Transformer underperformed on final metrics, particularly due to outlier sensitivity (extreme MAPE). Overfitting and noisy predictions likely stem from data volume and sequence length limitations.

**4. Comparative Analysis**

| **Model** | **MAE** | **RMSE** | **R²** | **MAPE (%)** |
| --- | --- | --- | --- | --- |
| **LSTM** | 16,868 | 29,633 | 0.3672 | 4.81 |
| **GRU** | 11,296 | 19,406 | 0.7286 | 465.37 |
| **Transformer** | 20,191 | 30,871 | 0.3132 | 1,515.82 |

* **Best Overall:** GRU (based on RMSE and R²)
* **Worst Performance:** Transformer (due to volatile MAPE)
* **Balanced Option:** LSTM (stable MAPE and fair R²)

**5. Key Observations**

* All models used sliding windows with a sequence length of 24 hours.
* MAPE can be misleading when actual values are near zero—affecting Transformer and GRU scores.
* GRU handled time series best, with robust sequence modeling and generalization.
* Transformer models may need larger datasets or positional encoding refinement to outperform RNNs in such tasks.