Literature Review on Solar Energy Production Forecasting Using Machine Learning and Meteorological Data

Focus on Solar Energy Forecasting

- Review academic papers, articles, and reports on solar energy forecasting.
- Identify key methodologies used in previous studies, including statistical and machine learning models.
- Examine different data sources, including meteorological data (solar irradiance, temperature, humidity) and historical solar power generation.
- Analyze challenges such as variability in weather conditions, data quality issues, and computational complexity.

Data Preprocessing Techniques

- Explore methods for handling missing data using techniques like mean/mode imputation, KNN imputation, and interpolation.
- Identify approaches for detecting and removing outliers in meteorological and energy generation data.
- Examine data transformation techniques such as normalization and standardization to improve model performance.
- Consider feature extraction and selection methods to enhance prediction accuracy.

Machine Learning Models

- Investigate different machine learning models used for solar energy production forecasting.
- Explore **time-series models** like ARIMA and Prophet for capturing temporal patterns.
- Examine **tree-based models** like Random Forest and XGBoost for feature importance and ensemble learning.
- Assess **deep learning models** like LSTMs and CNN-LSTM hybrids for capturing long-term dependencies and spatiotemporal relationships.
- Evaluate the strengths and weaknesses of each model in terms of accuracy, interpretability, and computational efficiency.

Optimization Strategies

- Research hyperparameter tuning techniques such as Grid Search and Bayesian Optimization to improve model performance.

- Explore feature selection methods to identify the most influential variables for forecasting.
- Investigate ensemble learning techniques to combine multiple models for better accuracy.
- Identify key evaluation metrics such as RMSE, MAE, and R^2 to assess model effectiveness.

Dataset Exploration

Data Loading and Inspection

- Load historical solar power generation and meteorological data into Python using Pandas.
- Inspect the dataset structure, including the number of rows, columns, and variable types.
- Identify relevant features such as solar irradiance, temperature, humidity, wind speed, and time-based attributes.

Data Cleaning

- Identify and handle missing values using imputation or removal techniques.
- Detect and manage inconsistencies or outliers that may affect model performance.
- Convert timestamps into suitable formats for time-series analysis.

Exploratory Data Analysis (EDA)

- Perform statistical analysis to understand data distribution and variability.
- Visualize relationships between meteorological parameters and solar power generation using histograms, scatter plots, and correlation heatmaps.
- Identify seasonal trends, daily variations, and other patterns in the data.

Feature Identification

- Select key features influencing solar power production, such as solar irradiance, temperature, cloud cover, and time-based variables.
- Consider additional external factors (e.g., holidays, location-specific weather conditions) that may impact solar energy output.

Data Transformation

- Apply feature scaling techniques such as Min-Max Scaling and Standardization for numerical variables.
- Encode categorical variables if present, using One-Hot Encoding or Label Encoding.
- Explore dimensionality reduction techniques if necessary for optimizing model training.

Deliverables

- **Literature Review Summary**: A comprehensive summary of past research, highlighting key findings, methodologies, and challenges in solar energy forecasting.
- **Dataset Exploration Report**: A detailed report documenting dataset inspection, data cleaning, EDA, feature identification, and data transformation steps.

References

 Chen, C., Duan, S., Cai, T. and Liu, B. (2011) 'Online 24-h solar power forecasting based on weather type classification using artificial neural network', *Solar Energy*, 85(11), pp. 2856–2870. Available at: https://doi.org/10.1016/j.solener.2011.08.027.

 Çolak, I. and Qahwaji, R. (2021) *Solar energy forecasting*. Amsterdam: Elsevier. Available at: https://doi.org/10.1016/C2019-0-02200-0.

 Hyndman, R.J. and Athanasopoulos, G. (2018) *Forecasting: principles and practice*. 2nd edn. Melbourne: OTexts. Available at: https://otexts.com/fpp2/.

 Inman, R.H., Pedro, H.T.C. and Coimbra, C.F.M. (2013) 'Solar forecasting methods for renewable energy integration', *Progress in Energy and Combustion Science*, 39(6), pp. 535–576. Available at: https://doi.org/10.1016/j.pecs.2013.06.002.

 Li, Z., Han, D., Wang, J. and Zhao, L. (2021) 'Solar power forecasting based on a hybrid CNN-LSTM model with spatiotemporal correlations', *IEEE Access*, 9, pp. 102928–102938. Available at: https://doi.org/10.1109/ACCESS.2021.3097983.

 Liu, Y., Zheng, Y., Li, Z., Wang, J. and Liu, Y. (2022) 'Short-term solar power forecasting: A review', *Journal of Cleaner Production*, 368, p. 133104. Available at: https://doi.org/10.1016/j.jclepro.2022.133104.

 Makridakis, S., Spiliotis, E. and Assimakopoulos, V. (2020) 'The M4 Competition: Results, findings, conclusion and way forward', *International Journal of Forecasting*, 36(1), pp. 54–74. Available at: https://doi.org/10.1016/j.ijforecast.2019.04.014.

 Pedro, H.T.C. and Coimbra, C.F.M. (2012) 'Assessment of forecasting techniques for solar power production with no exogenous information', *Solar Energy*, 86(7), pp. 2017–2028. Available at: https://doi.org/10.1016/j.solener.2012.04.004.

 Rahman, M.M. and Saha, T.K. (2022) 'Machine learning-based forecasting of renewable energy production: A review', *Renewable Energy Reports*, 8, pp. 229–248. Available at: https://doi.org/10.1016/j.rer.2021.12.004.

 Voyant, C., Notton, G., Kalogirou, S., Nivet, M.L., Paoli, C., Motte, F. and Fouilloy, A. (2017) 'Machine learning methods for solar radiation forecasting: A review', *Renewable Energy*, 105, pp. 569–582. Available at: https://doi.org/10.1016/j.renene.2016.12.095.

 Wang, F., Mi, Z., Su, S. and Zhao, H. (2019) 'Short-term solar irradiance forecasting model based on EEMD-LSTM', *Energy*, 174, pp. 1053–1061. Available at: https://doi.org/10.1016/j.energy.2019.03.179.