1.1 **Project Title:**

<u>Solar Power Generation Prediction Using Machine Learning Techniques</u>

1.2 **Team Members:**

1. Kartikeya Sehgal (2022244)
2. Krishna Shukla (2022254)

1.3 **Problem:**

The project aims to predict AC power generation by employing machine learning techniques to analyze various factors related to solar insolation and system efficiency. The key motivation is to provide a data-driven approach for evaluating the benefits of solar power installations, considering factors such as weather patterns, and system characteristics.

1.4 **Method:**

1. Data Preprocessing: Initial preprocessing involves cleaning and timestamping solar insolation data and system performance metrics.
2. Feature Extraction: Relevant features such as solar irradiance levels, ambient temperature, module temperature, and DC are extracted.
3. Dimensionality Reduction: Data analysis is conducted carefully to reduce the number of features.
4. Regression Modeling: Regression models such as Linear Regression, Random Forest, and XGBoost are utilized to predict AC power output based on input features.
5. Performance Evaluation: Model performance is evaluated using metrics such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared (R2) score to assess prediction accuracy.
6. Ensemble Techniques: Ensemble techniques like bagging and boosting are explored to improve prediction accuracy and robustness.

1.5 **Dataset:**

The project utilises a dataset containing historical solar insolation data, system performance metrics, and environmental factors collected from various solar power installations across different geographical locations. The dataset includes details such as timestamp, solar irradiance levels, ambient temperature, module temperature, and AC power output.

1.6 **Result/ Evaluation Metric:**

1. Accuracy: Percentage of correctly predicted AC power output values.
2. Mean Squared Error (MSE): Measure of the average squared difference between predicted and actual AC power output values.
3. Root Mean Squared Error (RMSE): Square root of the MSE, providing a measure of the average magnitude of prediction errors.
4. R-squared (R2) Score: Proportion of the variance in the AC power output that is predictable from the input features.

1.7 **References:**

1. "Solar Power Prediction Using Machine Learning" by E. Subramanian et al.: This paper employs Support Vector Machines, Gradient Boosting, and Random Forest as

machine learning algorithms to predict generated solar power, using metrics such as Area Under Curve (AUC) and F1-score to evaluate the models.

2. "A Survey of Solar Power Forecasting Machine Learning Techniques" by Debojyoti Chakraborty et al.: This literature discusses various ensemble ML algorithms applied to meteorological data for solar power generation forecasting.
3. "Short-term Solar Power Forecasting Using XGBoost with Numerical Weather Prediction" by Quoc-Thang Phan et al.: This paper advocates for the use of Kernel Principal Component Analysis along with XGBoost as the main regression technique for accurate prediction of one-hour ahead solar power forecasts.