## Artificial Intelligence

## and

## Machine Learning

Project Report

Semester-IV (Batch-2022)

SOLAR RADIATION PREDICTION

A red and white sign

Description automatically generated with low confidence

**Supervised By: Submitted By:**

Faculty Name : Shagun Sharma Group: G9

Harnoor Kaur 2210990374

Heena 2210990397

Harshita Bansal 2210990392

Himanshu 2210990405

**Department of Computer Science and Engineering**

## Chitkara+ University Institute of Engineering & Technology,

## Chitkara University, Punjab

**Abstract**

Accurate prediction of solar radiation is crucial for various applications such as solar power generation, weather forecasting, and environmental monitoring. Traditional methods often struggle to capture the complex atmospheric interactions, resulting in limited accuracy. To address this challenge, advanced machine learning techniques, including regression analysis and artificial neural networks, have been proposed. In this study, we aim to develop and evaluate a machine learning-based approach for solar radiation prediction. Leveraging historical meteorological data and advanced regression techniques, our approach seeks to improve the accuracy and reliability of solar radiation forecasts. Through comprehensive evaluation and validation, we demonstrate the effectiveness of our proposed approach in accurately predicting solar radiation across diverse environmental conditions. Our findings highlight the potential of machine learning techniques in enhancing solar radiation prediction accuracy and their significance in various real-world applications.

1. **Introduction**

In recent years, the demand for accurate and reliable solar radiation prediction models has grown substantially due to the increasing importance of solar energy as a sustainable alternative to traditional fossil fuels. Solar radiation prediction plays a crucial role in various applications, including solar power generation, weather forecasting, agriculture, and environmental monitoring. Accurate solar radiation prediction enables better planning and optimization of solar energy systems, leading to improved efficiency and cost-effectiveness. However, accurately predicting solar radiation is a complex task due to the inherent variability and uncertainty in atmospheric conditions, such as cloud cover, humidity, and atmospheric pressure. Traditional approaches to solar radiation prediction often rely on simplistic models that fail to capture the full complexity of atmospheric interactions, resulting in limited accuracy and reliability. To address these challenges, advanced machine learning techniques, such as regression analysis and artificial neural networks, have emerged as promising tools for developing more accurate and robust solar radiation prediction models. By leveraging large datasets of meteorological and solar radiation measurements, machine learning models can learn complex patterns and relationships in the data, enabling more accurate and reliable predictions of solar radiation under diverse environmental conditions. In this study, we aim to develop and evaluate a machine learning-based approach for solar radiation prediction, leveraging historical meteorological data and advanced regression techniques to improve the accuracy and reliability of solar radiation forecasts. Through comprehensive evaluation and validation, we seek to demonstrate the effectiveness of our proposed approach in accurately predicting solar radiation and its potential applications in various real-world scenarios.

In recent years, the demand for accurate and reliable solar radiation prediction models has grown substantially due to the increasing importance of solar energy as a sustainable alternative to traditional fossil fuels. Solar radiation prediction plays a crucial role in various applications, including solar power generation, weather forecasting, agriculture, and environmental monitoring. Accurate solar radiation prediction enables better planning and optimization of solar energy systems, leading to improved efficiency and cost-effectiveness. However, accurately predicting solar radiation is a complex task due to the inherent variability and uncertainty in atmospheric conditions, such as cloud cover, humidity, and atmospheric pressure. Traditional approaches to solar radiation prediction often rely on simplistic models that fail to capture the full complexity of atmospheric interactions, resulting in limited accuracy and reliability. To address these challenges, advanced machine learning techniques, such as regression analysis and artificial neural networks, have emerged as promising tools for developing more accurate and robust solar radiation prediction models. By leveraging large datasets of meteorological and solar radiation measurements, machine learning models can learn complex patterns and relationships in the data, enabling more accurate and reliable predictions of solar radiation under diverse environmental conditions. In this study, we aim to develop and evaluate a machine learning-based approach for solar radiation prediction, leveraging historical meteorological data and advanced regression techniques to improve the accuracy and reliability of solar radiation forecasts. Through comprehensive evaluation and validation, we seek to demonstrate the effectiveness of our proposed approach in accurately predicting solar radiation and its potential applications in various real-world scenarios.In recent years, the demand for accurate and reliable solar radiation prediction models has grown substantially due to the increasing importance of solar energy as a sustainable alternative to traditional fossil fuels. Solar radiation prediction plays a crucial role in various applications, including solar power generation, weather forecasting, agriculture, and environmental monitoring. Accurate solar radiation prediction enables better planning and optimization of solar energy systems, leading to improved efficiency and cost-effectiveness. However, accurately predicting solar radiation is a complex task due to the inherent variability and uncertainty in atmospheric conditions, such as cloud cover, humidity, and atmospheric pressure. Traditional approaches to solar radiation prediction often rely on simplistic models that fail to capture the full complexity of atmospheric interactions, resulting in limited accuracy and reliability. To address these challenges, advanced machine learning techniques, such as regression analysis and artificial neural networks, have emerged as promising tools for developing more accurate and robust solar radiation prediction models. By leveraging large datasets of meteorological and solar radiation measurements, machine learning models can learn complex patterns and relationships in the data, enabling more accurate and reliable predictions of solar radiation under diverse environmental conditions. In this study, we aim to develop and evaluate a machine learning-based approach for solar radiation prediction, leveraging historical meteorological data and advanced regression techniques to improve the accuracy and reliability of solar radiation forecasts. Through comprehensive evaluation and validation, we seek to demonstrate the effectiveness of our proposed approach in accurately predicting solar radiation and its potential applications in various real-world scenarios.Top of Form

* 1. **Background:**

The increasing global demand for renewable energy sources, coupled with the pressing need to mitigate climate change, has underscored the significance of solar energy as a sustainable alternative to traditional fossil fuels. Solar radiation, the primary source of solar energy, plays a pivotal role in determining the feasibility and efficiency of solar power generation systems. Accurate prediction of solar radiation is essential for optimizing the design, operation, and maintenance of solar energy systems, as well as for supporting decision-making processes in various sectors such as agriculture, urban planning, and disaster management.

Traditional methods for solar radiation prediction, such as physical modeling based on atmospheric physics and empirical approaches using meteorological data, often encounter challenges in accurately capturing the complex spatiotemporal variability of solar radiation. These methods rely on simplifications and assumptions that may not fully account for the intricate interactions between atmospheric parameters, such as cloud cover, humidity, aerosols, and surface characteristics, which significantly influence solar radiation patterns. As a result, predictions from traditional models may exhibit limited accuracy, especially under dynamic and heterogeneous environmental conditions.

In recent years, advances in machine learning (ML) and artificial intelligence (AI) have offered new opportunities for improving solar radiation prediction accuracy. ML techniques, including regression analysis, support vector machines, random forests, and artificial neural networks, have demonstrated promising capabilities in learning complex patterns and relationships from large-scale meteorological and solar radiation datasets. By leveraging historical data and incorporating diverse meteorological parameters as input features, ML-based models can capture nonlinear relationships and interactions between atmospheric variables, leading to more accurate and reliable predictions of solar radiation.

Despite the potential of ML techniques, challenges remain in the development and deployment of robust solar radiation prediction models. These include the availability and quality of input data, the selection of appropriate features, the choice of ML algorithms and model architectures, and the generalization of models across different geographical regions and climatic conditions. Addressing these challenges requires interdisciplinary collaboration between experts in atmospheric science, machine learning, data analytics, and domain-specific applications to develop innovative approaches for solar radiation prediction that can meet the evolving needs of stakeholders in the renewable energy sector and beyond.

* 1. **Objectives:**

The primary objective of this study is to develop and evaluate a machine learning-based approach for improving the accuracy and reliability of solar radiation prediction. Specifically, the objectives include:

Model Development: Develop machine learning models, including regression analysis, support vector machines, random forests, and artificial neural networks, for predicting solar radiation based on historical meteorological data and other relevant atmospheric parameters. Explore different feature selection techniques and model architectures to identify the most effective predictors and algorithms for solar radiation prediction.

Data Preparation and Preprocessing: Collect and preprocess large-scale meteorological and solar radiation datasets from diverse geographical regions and climatic conditions. Cleanse the data to remove outliers, missing values, and inconsistencies, and standardize the input variables to ensure compatibility and consistency across different datasets.

Model Training and Evaluation: Train and validate the machine learning models using historical meteorological data and corresponding solar radiation measurements. Employ rigorous cross-validation techniques, such as k-fold cross-validation and time-series splitting, to assess the performance of the models and evaluate their generalization ability across different temporal and spatial scales.

Performance Optimization: Fine-tune the hyperparameters of the machine learning models using grid search, random search, or Bayesian optimization techniques to optimize their performance in terms of prediction accuracy, precision, recall, and other relevant evaluation metrics. Investigate ensemble learning methods and model stacking techniques to further enhance the robustness and reliability of the predictions.

Model Interpretation and Visualization: Analyze the learned patterns and relationships in the machine learning models to gain insights into the factors influencing solar radiation variability. Visualize the model predictions and feature importance scores to facilitate interpretation and decision-making by stakeholders in the renewable energy sector and related domains.

Real-World Application and Deployment: Demonstrate the practical utility and effectiveness of the developed machine learning models for solar radiation prediction in real-world scenarios. Deploy the models in operational settings, such as solar power plants, weather forecasting systems, and environmental monitoring networks, to support decision-making processes and optimize resource allocation for solar energy generation and utilization.

Overall, the objectives of this study aim to advance the state-of-the-art in solar radiation prediction by harnessing the power of machine learning techniques and leveraging big data analytics to address the complex challenges inherent in modeling and forecasting solar radiation accurately and reliably.

**1.3 Significance:**

The significance of this study lies in its potential to advance the state-of-the-art in solar radiation prediction and contribute to the broader goals of promoting sustainable energy solutions, mitigating climate change, and fostering environmental stewardship. Several key aspects underscore the significance of this research endeavor:

Renewable Energy Transition: As the world transitions towards a more sustainable energy future, solar energy represents a crucial pillar of the renewable energy portfolio. Accurate prediction of solar radiation is essential for optimizing the planning, design, and operation of solar power generation systems, thereby facilitating the widespread adoption of solar energy and reducing reliance on fossil fuels.

Climate Change Mitigation: Effective management of solar energy resources relies on accurate forecasting of solar radiation patterns, which are intricately linked to climate dynamics and atmospheric processes. By improving the accuracy and reliability of solar radiation prediction, this research contributes to climate change mitigation efforts by enabling better adaptation strategies, resource allocation, and decision-making in response to changing environmental conditions.

Economic Benefits: The optimization of solar energy systems based on reliable solar radiation forecasts can yield significant economic benefits, including cost savings, improved energy efficiency, and enhanced revenue generation for solar power producers. By reducing uncertainty and improving the predictability of solar radiation, this research helps maximize the economic viability and competitiveness of solar energy projects and investments.

Environmental Impacts: Solar energy is renowned for its low environmental impact compared to conventional energy sources, such as coal, oil, and natural gas. By facilitating the deployment of more efficient and reliable solar energy systems through improved solar radiation prediction, this research contributes to reducing greenhouse gas emissions, air pollution, and ecological degradation associated with fossil fuel combustion.

Technological Innovation: The development and application of machine learning techniques for solar radiation prediction represent a paradigm shift in the field of renewable energy forecasting. By harnessing the power of big data analytics, artificial intelligence, and advanced modeling approaches, this research opens up new avenues for innovation and interdisciplinary collaboration in the renewable energy sector, paving the way for more efficient, resilient, and sustainable energy systems.

Overall, the significance of this study lies in its potential to address critical challenges facing the renewable energy industry and society at large, while fostering innovation, resilience, and sustainability in the transition towards a clean energy future. By advancing our understanding of solar radiation dynamics and enhancing our predictive capabilities, this research contributes to shaping a more sustainable and equitable world for future generations.

1. **Problem Definition:**

The problem at hand involves the development of an accurate and reliable predictive model for solar radiation. Solar radiation prediction is crucial for various applications, including solar power generation, weather forecasting, and agricultural planning. Traditional methods often struggle to accurately capture the complex interactions between atmospheric parameters, leading to limited prediction accuracy and reliability.

The objective is to address these challenges by leveraging machine learning techniques to develop a predictive model capable of accurately forecasting solar radiation levels. The model should be able to effectively incorporate a wide range of meteorological variables, such as temperature, humidity, cloud cover, wind speed, and atmospheric pressure, to improve prediction accuracy across different environmental conditions and geographical locations.

Key considerations for the predictive model include robustness, scalability, and generalization ability. The model should be able to handle diverse data sources and adapt to changing environmental conditions without sacrificing prediction accuracy. Additionally, the model should be scalable to handle large-scale datasets and generalize well to unseen data to ensure reliable performance in real-world applications.

The ultimate goal is to develop a predictive model that can provide accurate and timely forecasts of solar radiation, enabling better decision-making in various sectors, including energy, agriculture, and environmental management. By addressing the challenges associated with solar radiation prediction, this research aims to contribute to the advancement of renewable energy technologies and support efforts to mitigate climate change and promote sustainable development.

1. **Requirements:**

**3.1 Software Requirements:**

Python Programming Language: Python is required for implementing the machine learning algorithms and data preprocessing techniques.

Machine Learning Libraries: Libraries such as scikit-learn, TensorFlow, or PyTorch are necessary for developing and training machine learning models.

Data Analysis and Visualization Tools: Libraries such as pandas, NumPy, and Matplotlib are essential for data manipulation, analysis, and visualization.

Integrated Development Environment (IDE): An IDE like Jupyter Notebook, PyCharm, or Visual Studio Code is recommended for coding and experimentation.

Data Collection and Management Tools: Tools for collecting, managing, and preprocessing meteorological and solar radiation datasets, such as SQL databases or cloud-based storage solutions.

Version Control System: Git and GitHub or similar version control systems for tracking code changes and collaboration among team members.

**3.2 Hardware Requirements:**

Computing Resources: Sufficient computing resources (CPU, memory, and disk space) for running machine learning algorithms and processing large datasets efficiently.

High-Performance Computing (HPC) Facilities: Access to HPC facilities or cloud computing platforms for parallel processing and scalability, especially for training complex machine learning models.

Graphics Processing Unit (GPU): Optional but recommended for accelerating model training and inference, especially for deep learning algorithms.

Data Storage: Adequate storage capacity for storing large-scale meteorological and solar radiation datasets and model checkpoints.

**3.3 Additional Considerations:**

Accessibility: Ensure that the software tools and computing resources are accessible to all team members involved in the project, including researchers, developers, and domain experts.

Scalability and Performance: Optimize software and hardware configurations to handle large-scale datasets and accommodate future growth in data volume and computational complexity.

Security and Compliance: Implement appropriate security measures to protect sensitive data and comply with data privacy regulations and institutional policies.

Documentation and Reporting: Maintain detailed documentation of software configurations, experimental procedures, and results for reproducibility and reporting purposes.

Collaboration and Communication: Foster effective collaboration and communication among team members through regular meetings, shared documentation, and collaboration tools like Slack or Microsoft Teams.

By meeting these software and hardware requirements, the project can proceed smoothly, ensuring efficient development, training, and evaluation of the solar radiation prediction model.

**3.4 Dataset:**

#### The dataset contains such columns as: "wind direction", "wind speed", "humidity" and temperature. The response parameter that is to be predicted is: "Solar\_radiation". It contains measurements for the past 4 months and you have to predict the level of solar radiation.Dataset.

Link-><https://www.kaggle.com/datasets/dronio/SolarEnergy>

**4.Methodology:**

Data Collection and Preprocessing:

Gather historical meteorological data and solar radiation measurements from reliable sources, such as meteorological agencies, research institutions, or open-access databases.

Cleanse the data to remove outliers, errors, and missing values. Impute missing data using appropriate techniques such as mean imputation, interpolation, or machine learning-based methods.

Standardize the data to ensure uniformity in units and scales across different variables.

Split the dataset into training, validation, and test sets to facilitate model training, evaluation, and validation.

Exploratory Data Analysis (EDA):

Conduct exploratory data analysis to gain insights into the dataset's characteristics, distributions, and relationships between variables.

Visualize key meteorological variables and solar radiation measurements using plots, histograms, correlation matrices, and scatter plots.

Identify patterns, trends, and seasonality in the data that may influence solar radiation levels.

Feature Engineering:

Extract relevant features from the meteorological data that are expected to impact solar radiation levels, such as temperature, humidity, cloud cover, wind speed, and atmospheric pressure.

Generate additional features or transformations, such as time of day, day of the year, and solar zenith angle, to capture temporal and spatial variations in solar radiation.

Model Selection and Training:

Choose suitable machine learning algorithms for solar radiation prediction, such as linear regression, support vector regression, random forests, gradient boosting, or neural networks.

Train multiple models using the training dataset and evaluate their performance using appropriate evaluation metrics, such as mean squared error (MSE), root mean squared error (RMSE), coefficient of determination (R^2), and mean absolute error (MAE).

Explore ensemble learning techniques, such as model averaging or stacking, to combine predictions from multiple models and improve overall prediction accuracy.

Hyperparameter Tuning:

Perform hyperparameter tuning using techniques such as grid search, random search, or Bayesian optimization to optimize the performance of the selected machine learning algorithms.

Fine-tune model hyperparameters such as learning rate, regularization strength, tree depth, and number of hidden layers to improve prediction accuracy and generalization ability.

Model Evaluation and Validation:

Validate the trained models using the validation dataset and assess their performance on unseen data.

Compare the performance of different models based on evaluation metrics and select the best-performing model for further analysis.

Validate the model's generalization ability across different geographical regions and climatic conditions to ensure its robustness and reliability in real-world applications.

Model Interpretation and Deployment:

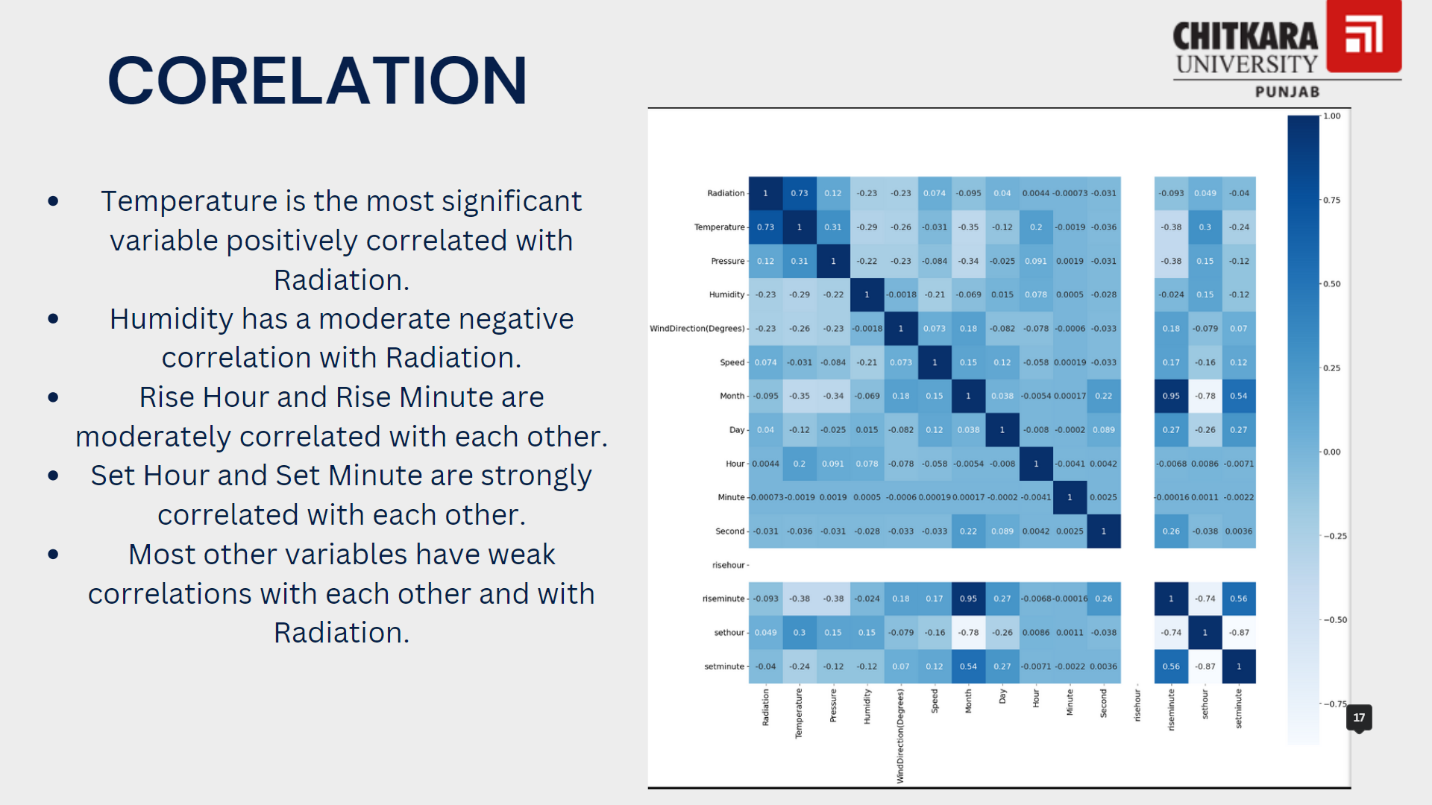
Interpret the trained models to understand the factors influencing solar radiation prediction and identify the most important features.

Deploy the trained model in operational settings, such as solar power plants, weather forecasting systems, or decision support tools, to provide real-time or short-term forecasts of solar radiation levels.

Monitor the model's performance over time and update it periodically with new data to maintain its accuracy and relevance.

By following this methodology, you can develop an accurate and reliable predictive model for solar radiation prediction, leveraging machine learning techniques and domain knowledge to improve energy forecasting and support decision-making in various applications.

**Result:**



Here are some key observations from the correlation matrix:

Radiation:

Positively correlated with Temperature (0.73).

Negatively correlated with Humidity (-0.23).

Weak correlations with other variables.

Temperature:

Strong positive correlation with Radiation (0.73).

Moderate positive correlation with Pressure (0.31).

Moderate negative correlation with Humidity (-0.3).

Weak correlations with other variables.

Pressure:

Moderate positive correlation with Temperature (0.31).

Weak correlations with other variables.

Humidity:

Moderate negative correlation with Temperature (-0.3).

Weak correlations with other variables

Wind Direction (Degrees):

Weak correlations with all other variables.

Speed:

Weak correlations with all other variables.

Month:

Moderate positive correlation with Day (0.15).

Weak correlations with other variables.

Day:

Moderate positive correlation with Month (0.15).

Weak correlations with other variables.

Hour:

Weak correlations with all other variables.

Minute:

Weak correlations with all other variables.

Second:

Weak correlations with all other variables.

Rise Hour:

Moderate positive correlation with Rise Minute (0.74).

Weak correlations with other variables.

Rise Minute:

Moderate positive correlation with Rise Hour (0.74).

Weak correlations with other variables.

Set Hour:

Strong positive correlation with Set Minute (0.87).

Weak correlations with other variables.

Set Minute:

Strong positive correlation with Set Hour (0.87).

Weak correlations with other variables.

**Conclusion:**

Temperature is the most significant variable positively correlated with Radiation.

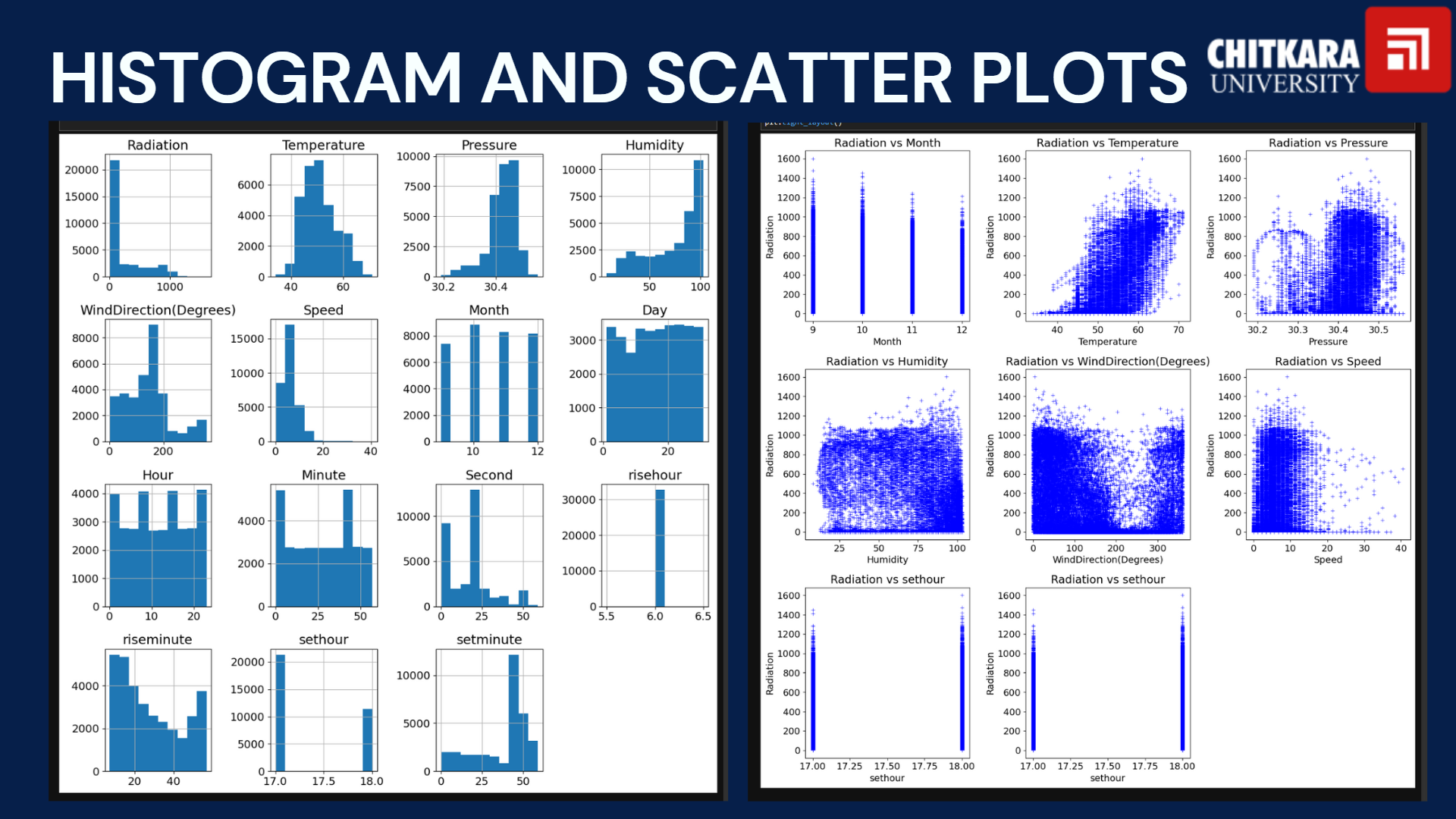
Humidity has a moderate negative correlation with Radiation.

Rise Hour and Rise Minute are moderately correlated with each other.

Set Hour and Set Minute are strongly correlated with each other.

Most other variables have weak correlations with each other and with Radiation.

This analysis can help in understanding which variables are most influential in predicting solar radiation and can guide feature selection in predictive modeling.



Radiation:

The distribution is right-skewed, with most values concentrated at the lower end (close to 0) and a long tail extending to higher values.

Temperature:

The distribution is approximately normal, centered around 60 degrees.

Pressure:

The distribution is very narrow, centered around 30.4, indicating little variation in pressure values.

Humidity:

The distribution is right-skewed, with most values concentrated at the higher end (close to 100).

Wind Direction (Degrees):

The distribution is somewhat uniform, with a notable peak around 200 degrees.

Speed:

The distribution is right-skewed, with most values concentrated at the lower end (close to 0).

Month:

The distribution is uniform, indicating an equal number of observations for each month.

Day:

The distribution is approximately uniform, indicating an equal number of observations for each day of the month.

Hour:

The distribution is approximately uniform, indicating an equal number of observations for each hour of the day.

Minute:

The distribution is approximately uniform, indicating an equal number of observations for each minute of the hour.

Second:

The distribution is right-skewed, with most values concentrated at the lower end (close to 0).

Rise Hour:

The distribution is very narrow, centered around 6, indicating little variation in rise hour values.

Rise Minute:

The distribution is right-skewed, with most values concentrated at the lower end (close to 0).

Set Hour:

The distribution is very narrow, centered around 18, indicating little variation in set hour values.

Set Minute:

The distribution is approximately uniform, indicating an equal number of observations for each minute of the set hour.

**Conclusion:**

Radiation and Humidity are right-skewed, indicating that most observations are concentrated at the lower and higher ends, respectively.

Temperature follows a normal distribution, which is useful for predictive modeling.

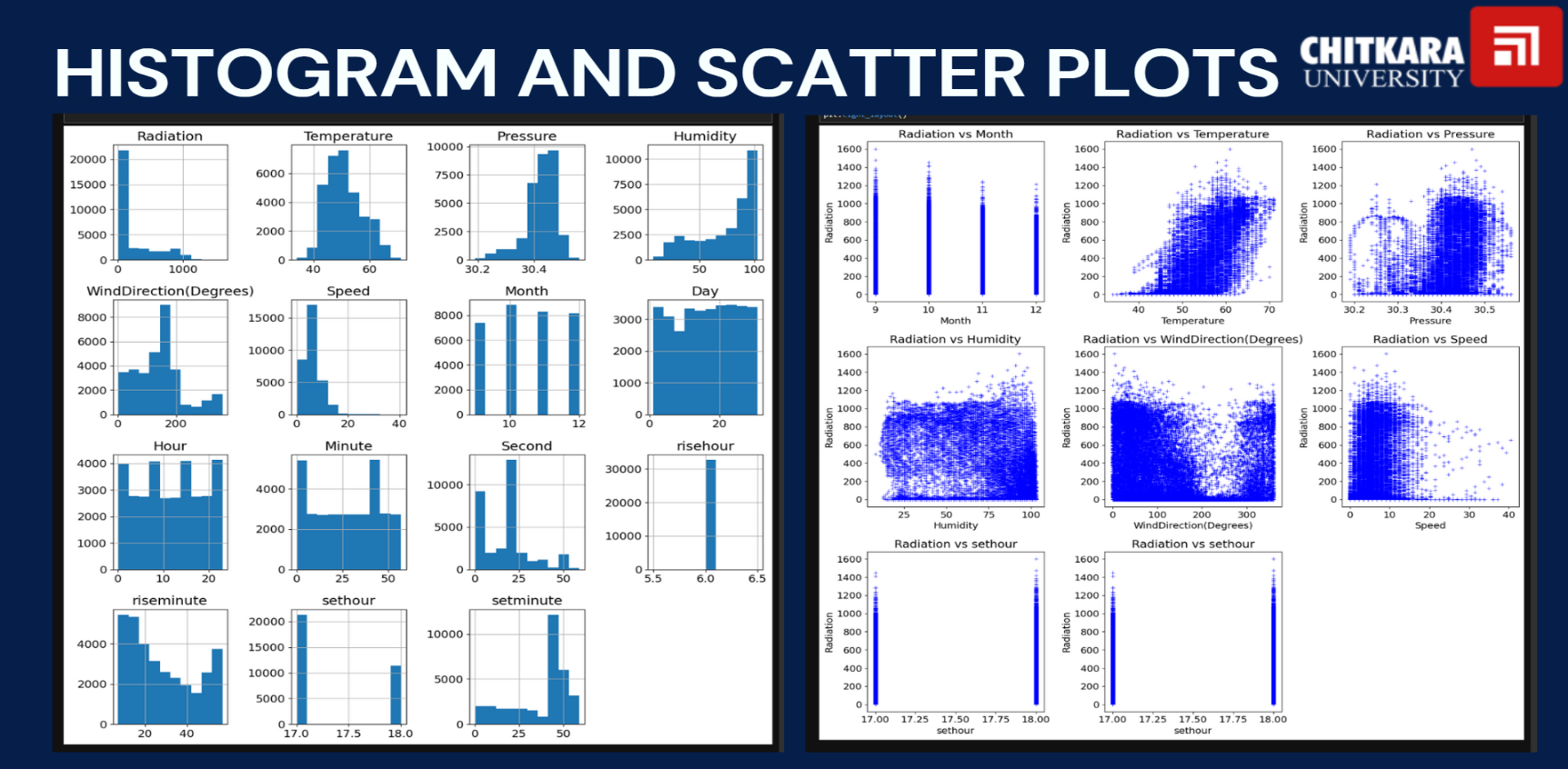
Pressure, Rise Hour, and Set Hour have very narrow distributions, indicating little variation in these variables.

Wind Direction has a notable peak around 200 degrees.

Speed and Second are right-skewed, with most values concentrated at the lower end.

Month, Day, Hour, Minute, and Set Minute have approximately uniform distributions, indicating an equal number of observations across their ranges.

This analysis helps in understanding the distribution of each variable, which is crucial for feature engineering and selecting appropriate statistical or machine learning models for solar radiation prediction.



1. Radiation vs Month: The scatter plot shows a uniform distribution of radiation values across different months.

2. Radiation vs Temperature: The scatter plot shows a positive correlation, with radiation increasing as temperature increases.

3. Radiation vs Pressure: The scatter plot shows no clear correlation between radiation and pressure.

4. Radiation vs Humidity: The scatter plot shows a slight negative correlation, with radiation decreasing as humidity increases.

5. Radiation vs WindDirection(Degrees): The scatter plot shows a slight negative correlation, with radiation decreasing as wind direction increases.

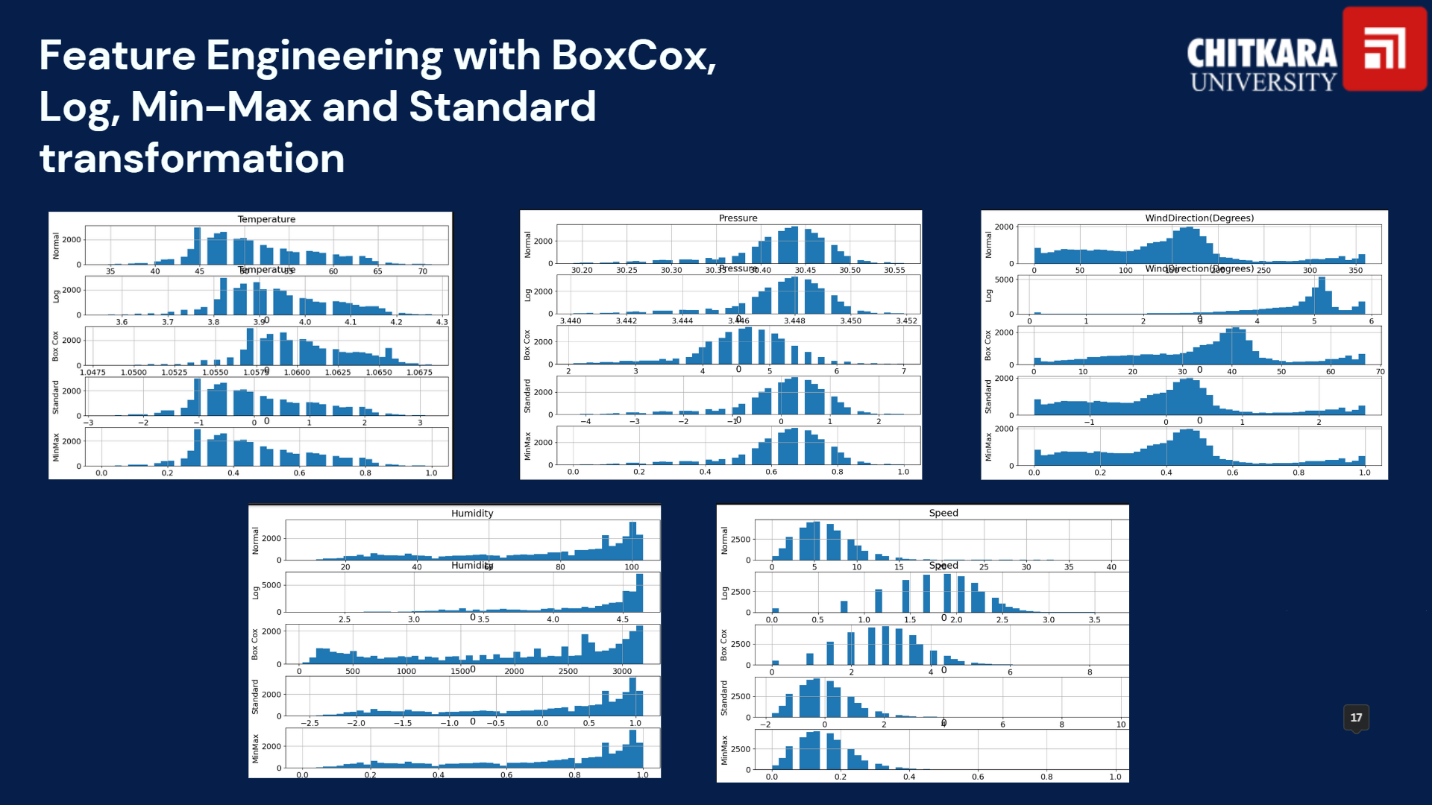
6. Radiation vs Speed: The scatter plot shows no clear correlation between radiation and speed.

7. Radiation vs risehour: The scatter plot shows a very narrow range of values, with no clear correlation.

8. Radiation vs sethour: The scatter plot shows a very narrow range of values, with no clear correlation.

**Summary:**

* **Positive Correlation**: Radiation vs Temperature.
* **Negative Correlation**: Radiation vs Humidity, Radiation vs WindDirection(Degrees).
* **No Clear Correlation**: Radiation vs Pressure, Radiation vs Speed, Radiation vs risehour, Radiation vs sethour, Radiation vs Month.



**Variables and Transformations**

**Temperature**

* Original: The histogram shows a roughly normal distribution centered around 50-60 degrees.
* BoxCox: The distribution becomes more symmetric and closer to a normal distribution.
* Log: The distribution becomes more right-skewed.
* Min-Max: The values are scaled between 0 and 1, maintaining the original shape.
* Standard: The values are standardized to have a mean of 0 and a standard deviation of 1, maintaining the original shape.

**Pressure**

* Original: The histogram shows a narrow range of values centered around 30.4.
* BoxCox: The distribution becomes more symmetric and closer to a normal distribution.
* Log: The distribution becomes more right-skewed.
* Min-Max: The values are scaled between 0 and 1, maintaining the original shape.
* Standard: The values are standardized to have a mean of 0 and a standard deviation of 1, maintaining the original shape.

**WindDirection(Degrees)**

* Original: The histogram shows a bimodal distribution with peaks around 0 and 200 degrees.
* BoxCox: The distribution becomes more symmetric and closer to a normal distribution.
* Log: The distribution becomes more right-skewed.
* Min-Max: The values are scaled between 0 and 1, maintaining the original shape.
* Standard: The values are standardized to have a mean of 0 and a standard deviation of 1, maintaining the original shape.

**Humidity**

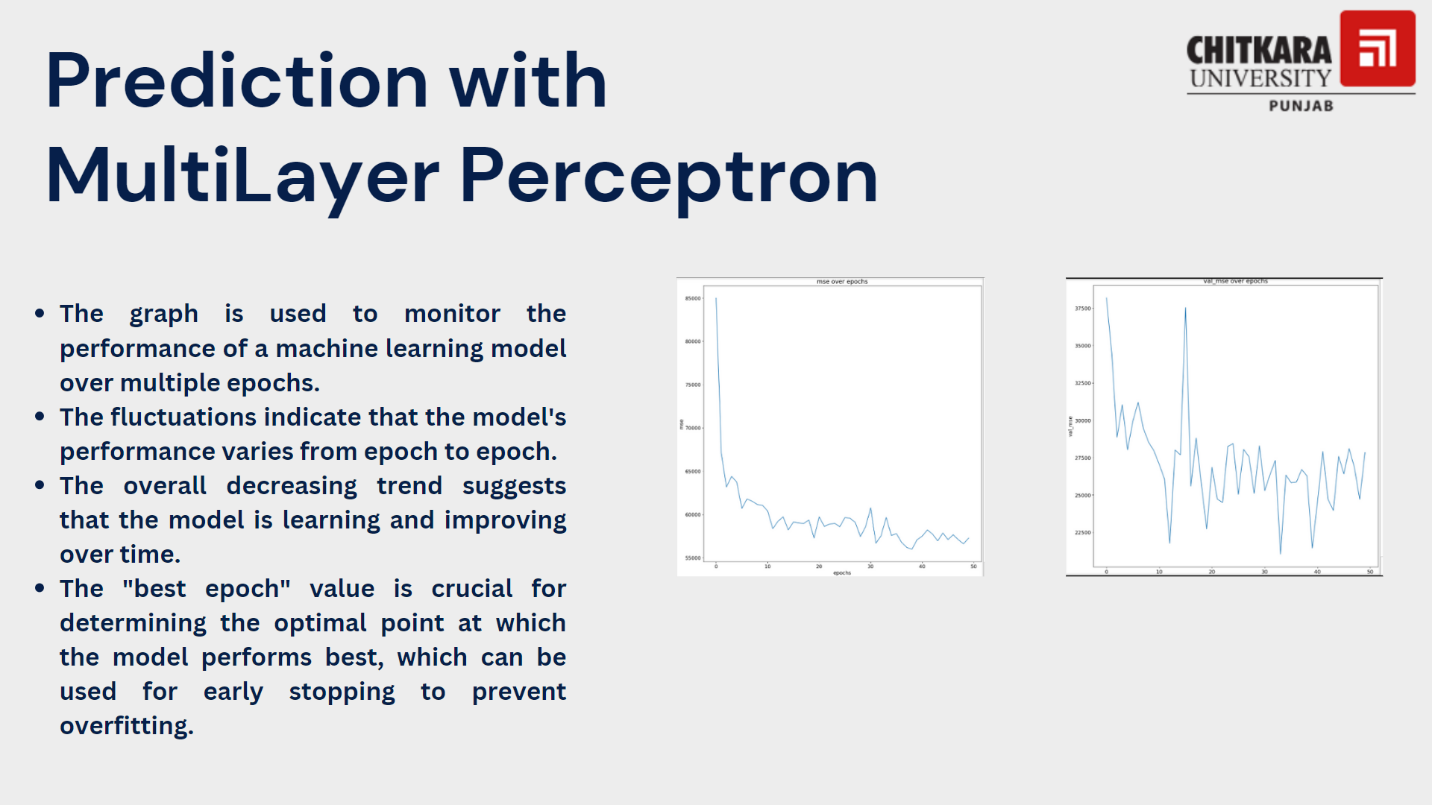
* Original: The histogram shows a right-skewed distribution with a peak around 100.
* BoxCox: The distribution becomes more symmetric and closer to a normal distribution.
* Log: The distribution becomes more right-skewed.
* Min-Max: The values are scaled between 0 and 1, maintaining the original shape.
* Standard: The values are standardized to have a mean of 0 and a standard deviation of 1, maintaining the original shape.

**Speed**

* Original: The histogram shows a right-skewed distribution with most values below 10.
* BoxCox: The distribution becomes more symmetric and closer to a normal distribution.
* Log: The distribution becomes more right-skewed.
* Min-Max: The values are scaled between 0 and 1, maintaining the original shape.
* Standard: The values are standardized to have a mean of 0 and a standard deviation of 1, maintaining the original shape.

**Summary**

* BoxCox Transformation: Generally makes the distributions more symmetric and closer to a normal distribution.
* Log Transformation: Often makes the distributions more right-skewed.
* Min-Max Transformation: Scales the values between 0 and 1, maintaining the original shape of the distribution.
* Standard Transformation: Standardizes the values to have a mean of 0 and a standard deviation of 1, maintaining the original shape of the distribution.



**Key Points from the Image**

Purpose of the Graphs:

The graphs are used to monitor the performance of a machine learning model over multiple epochs.

Fluctuations:

The fluctuations in the graphs indicate that the model's performance varies from epoch to epoch.

Overall Trend:

The overall decreasing trend in the graphs suggests that the model is learning and improving over time.

Best Epoch:

The "best epoch" value is crucial for determining the optimal point at which the model performs best. This can be used for early stopping to prevent overfitting.

**Analysis of the Graphs**

Left Graph:

The x-axis represents the number of epochs.

The y-axis represents the loss or error metric.

The graph shows a decreasing trend in the loss/error metric as the number of epochs increases, indicating that the model is learning and improving its performance over time.

Right Graph:

The x-axis represents the number of epochs.

The y-axis represents the validation loss or error metric.

The graph shows fluctuations in the validation loss/error metric, but there is an overall decreasing trend, indicating that the model's performance on the validation set is also improving over time.

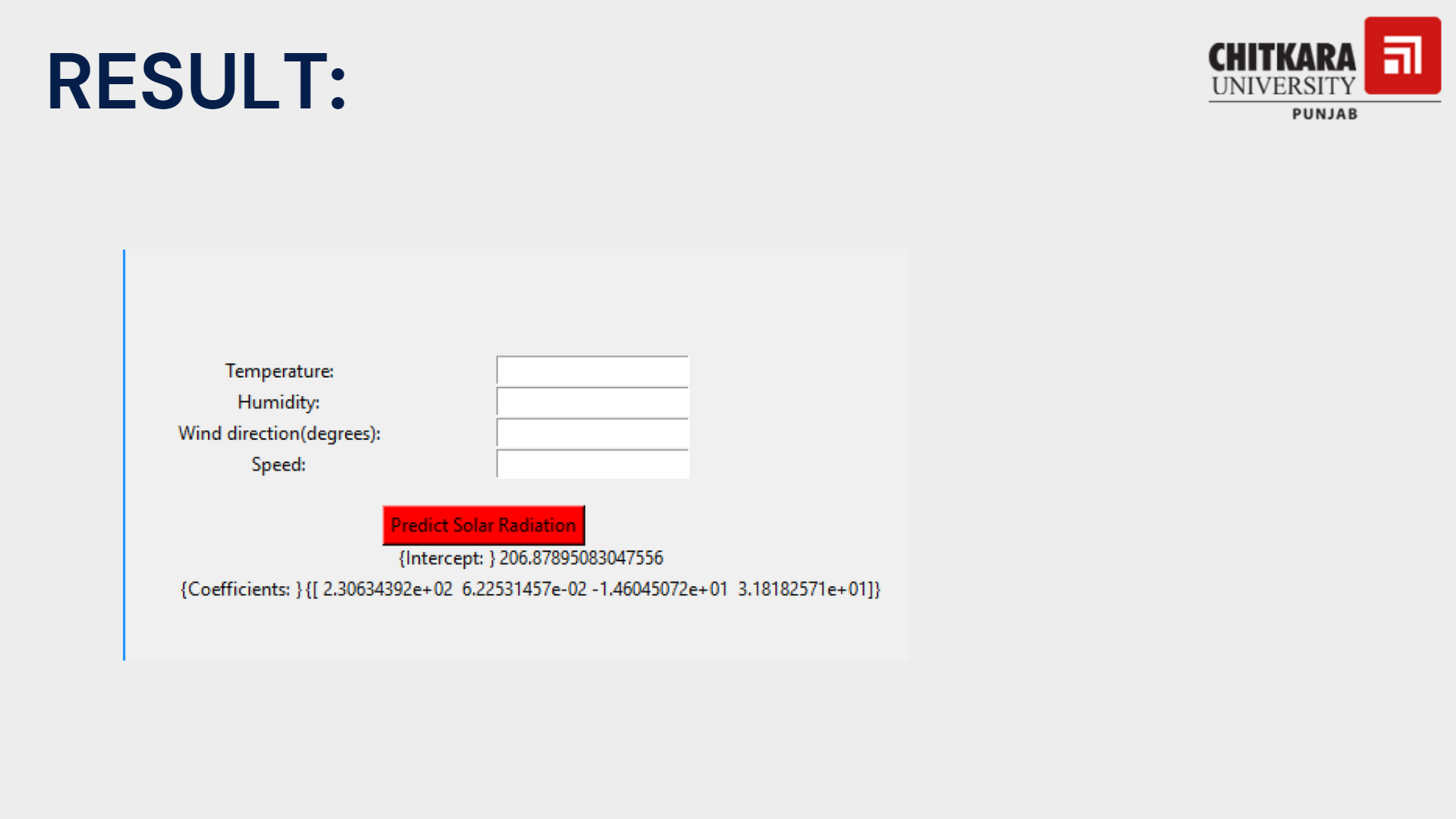
**Summary**

The graphs are used to track the performance of a MultiLayer Perceptron model over multiple epochs.

Fluctuations in the graphs indicate variability in the model's performance from epoch to epoch.

The overall decreasing trend in both graphs suggests that the model is learning and improving over time.

Identifying the "best epoch" is important for determining the optimal point for early stopping to prevent overfitting and ensure the model performs well on unseen data.



**Input Fields:**

There are four input fields labeled:

* Temperature
* Humidity
* Wind direction (degrees)
* Speed

These fields are likely meant for the user to input the respective values for making a prediction.

**Predict Button:**

A red button labeled "Predict Solar Radiation" is present, which suggests that clicking this button will trigger the prediction process based on the input values.

**Model Information:**

Below the button, there are details about the prediction model:

Intercept: 206.87895083047556

Coefficients:

* 2.30634392e+02
* 6.22531457e-02
* -1.46045072e+01
* 3.18182571e+01

These values represent the parameters of a linear regression model used for predicting solar radiation.

**References:**

[**https://www.kaggle.com/datasets/dronio/SolarEnergy**](https://www.kaggle.com/datasets/dronio/SolarEnergy)