**EVAPOTRANSPIRATION MODELLING USING EXPLAINABLE AI**

**A PROJECT REPORT**

***Submitted by***

**MAMALI SAHOO**

**Reg.No - 210301120142**

***in partial fulfilment for the award of the***

***degree Of***

**BACHELOR OF TECHNOLOGY**

***in***

**COMPUTER SCIENCE AND ENGNIEERING**

****

**SCHOOL OF ENGINEERING AND TECHNOLOGY**

**BHUBANESWAR CAMPUS**

**CENTURION UNIVERSITY OF TECHNOLOGY AND MANAGEMENT**

**ODISHA**

**November, 2024**

**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**

**SCHOOL OF ENGINEERING AND TECHNOLOGY**

**BHUBANESWAR CAMPUS**

**BONAFDE CERTIFICATE**

Certified that this project report “**Evapotranspiration Modelling using Explainable AI***”* is the Bonafide work of **“MAMALI SAHOO”** who carried out the project work under my supervision. This is to further certify to the best of my knowledge that this project has not been carried out earlier in this institute and the university

**SIGNATURE**

**Dr. Sudhansu Kumar Samal**

**Department of Electronics and communication Engineering**

*Certified that the above-mentioned project has been duly carried out as per the norms of the college and statutes of the university*

**SIGNATURE**

**Prof. Raj Kumar Mohanta**

**HEAD OF THE DEPARTMENT**

**Professor of Computer Science & Engineering**

**DECLARATION**

I hereby declare that the project entitled **“Evapotranspiration modelling using explainable AI”** submitted for the “Internship Project” of 7th semester B.Tech in Computer Science and Engineering is my original work and the project has not formed the basis for the award of any Degree / Diploma or any other similar titles in any other University / Institute.

**Name of the Student: MAMALI SAHOO**

**Signature of the Student:**

**Registration No: 210301120142**

**Place: Jatni, Odisha**

**Date:**

**ACKNOWLEDGEMENTS**

I wish to express my profound and sincere gratitude to **Dr. Bibhuti Sahoo**, Department of Computer Science & Engineering, PKD Campus, who guided me into the intricacies of this project nonchalantly with matchless magnanimity.

I thank **Prof. Raj Kumar Mohanta**, Head of the Dept. of Department of Computer Science and Engineering, SoET, Bhubaneswar Campus and **Prof. Sujata Chakravarty**, Dean, School of Engineering and Technology, Bhubaneswar Campus for extending their support during Course of this investigation.

I would be failing in my duty if I don’t acknowledge the cooperation rendered during various stages of interpretation by team members.

I am highly grateful to the teachers who evinced keen interest and invaluable support in the progress and successful completion of my project work.

I am indebted to my collegues for their constant encouragement, co-operation and help. Words of gratitude are not enough to describe the accommodation and fortitude which they have shown throughout my endeavor.

**Name of the Student: MAMALI SAHOO**

**Signature of the Student:**

**Registration No: 210301120142**

**Place: Jatni, Odisha**

**Date:**

**TABLE OF CONTENTS**

**CHAPTER NO. TITLE PAGE NO.**

**CERTIFICATE** i

**DECLARATION ii**

**ACKNOWLEDGEMENT iii**

**ABSTRACT vii**

***CHAPTER-1 INTRODUCTION***

1. Introduction 4
2. Chatper Summerization 4
3. Objective 5

***CHAPTER-2 LITERATURE REVIEW***

2.1 Tradotional Methods 6

***CHAPTER-3 SOFTWARE REQUIREMENTS***

3.1 Integrated Development Environment

3.2 Libraries and Framework 10

***CHAPTER-4 EXPLAINABLE AI***

3.1 SHAP Explanation 11

***CHAPTER-4 METHODOLOGY***

5.1 Block Diagram

5.2 Data Collection

5.2.1 Data Description 13

5.3 Data Preprocessing 14

5.4 Time Series Analysis 17

5.5 Model Selection 19

5.6 LSTM implementation 21

5.7 Integration of XAI 23

5.8 Visualisation plotting 29

***CHAPTER-6 RESULT AND DISCUSSION***

6.1 Result 32

**CONCLUSION 34**

**REFERENCES 35**

**APENDIX**

**A. Model code**

**B. Timeline**

**ABSTRACT**

This study aims to develop efficient machine learning models for accurately predicting potential evapotranspiration (PET), a crucial parameter in agricultural water management that enhances proactive irrigation scheduling and optimizes water use efficiency. By leveraging advanced modeling techniques, we evaluate the performance of Long Short-Term Memory (LSTM) networks and eXtreme Gradient Boosting (XGBoost) through comprehensive performance metrics, including mean squared error (MSE), mean absolute error (MAE), and root mean squared error (RMSE). Recognizing the "black-box" nature of these models, we employ SHAP (SHapley Additive exPlanations) values to provide global interpretability, elucidating how the best-performing model learns from various input features, such as wind speed, temperature, and humidity. Our findings indicate that both models achieve high predictive accuracy for PET, with XGBoost demonstrating superior performance in specific scenarios. Additionally, SHAP analysis offers valuable insights into feature importance and interactions, enabling incremental model enhancements based on the explanations provided. These insights not only improve the model's predictive capability but also promote a deeper understanding of the underlying processes influencing PET. Ultimately, this research contributes to the development of more accurate and interpretable machine learning models for predicting PET, facilitating better decision-making in agricultural water management and promoting sustainable irrigation practices that can lead to increased crop yield and resource conservation. This work underscores the significance of integrating advanced predictive modeling with interpretability to bridge the gap between data science and practical agricultural applications.

Keywords: Evapotranspiration, Machine Learning, XgBoost, LSTM, Explainable Artificial Intelligence.

**CHAPTER - 1**

**INTRODUCTION**

* 1. **Introduction**

Evapotranspiration (ET) is a critical component of the hydrological cycle, representing the combined processes of water evaporation from soil and water bodies and transpiration from plants. Accurate modeling of evapotranspiration is essential for effective water management, particularly in agricultural settings where it directly influences irrigation strategies, crop yield, and overall water resource sustainability. Traditional methods of estimating ET often rely on empirical formulas that can be limited by their assumptions and sensitivity to local conditions.In recent years, the application of machine learning techniques has revolutionized the way ET is modeled, offering the potential for higher accuracy and adaptability to various environmental factors. However, many machine learning models, particularly those based on deep learning or ensemble methods, are often considered "black boxes," making it challenging for stakeholders to understand how predictions are made. This lack of interpretability can hinder the adoption of these advanced models in real-world agricultural practices, where decision-makers require clear insights into the factors driving ET predictions.To address this challenge, our project focuses on developing evapotranspiration models using Explainable Artificial Intelligence (XAI) techniques. By leveraging advanced machine learning algorithms, such as Long Short-Term Memory (LSTM) networks and eXtreme Gradient Boosting (XGBoost), we aim to accurately predict ET while simultaneously enhancing model interpretability through the use of SHAP (SHapley Additive exPlanations) values. This approach not only provides a deeper understanding of the influence of various meteorological and environmental factors on ET but also allows for more informed decision-making in irrigation management.The integration of explainable AI into evapotranspiration modeling represents a significant advancement in agricultural water management. By combining the predictive power of machine learning with the transparency afforded by XAI, our project seeks to create models that not only perform well but also instill confidence among users, ultimately leading to improved water conservation practices and enhanced agricultural productivity. This work aims to contribute to the ongoing efforts to develop sustainable agricultural practices that are responsive to the challenges posed by climate change and increasing water scarcity. **1.2 Chapter Summerization**

Chapter 1: Introduction - Provides an overview of the project.Chapter 2: Literature Review – Discuss traditional methods.Chapter 3: Software requirements- Discuss software enviornment.Chapter 4: Methodology - Explores the steps for working.Chapter 5: Result & Discussion – Findings of the study.

**1.3 Objectives**

To Develop Predictive Models: Create accurate machine learning models, such as Long Short-Term Memory (LSTM) networks and eXtreme Gradient Boosting (XGBoost), to effectively predict potential evapotranspiration based on relevant meteorological and environmental data.To Evaluate Model Performance: Assess the performance of the developed models using metrics such as mean squared error, mean absolute error, and root mean squared error to determine their accuracy and reliability in predicting evapotranspiration.

**CHAPTER – 2**

**LITERATURE REVIEW**

2.1 **Literature survey :**

The field of evapotranspiration (ET) modeling has witnessed significant advancements due to the integration of machine learning and artificial intelligence. Traditional methods of estimating ET, often reliant on empirical formulas, have been limited by their assumptions and localized applicability (Fereres & García-Vila, 2018). In contrast, machine learning approaches, including those utilizing dynamic neural networks, have demonstrated the ability to model complex relationships in soil moisture content, enhancing predictive irrigation scheduling (Adeyemi et al., 2018). This shift towards data-driven methods highlights the importance of leveraging sensor data and advanced algorithms to improve irrigation management and crop productivity (Goldstein et al., 2018; Goap et al., 2018).Explainable AI (XAI) has emerged as a critical area of research in this context, addressing the "black-box" nature of many machine learning models (Barredo Arrieta et al., 2020). By providing insights into model decision-making processes, XAI facilitates a better understanding of the factors influencing ET predictions. Studies have shown that interpretability can enhance user trust and adoption of machine learning solutions in agricultural settings (Ribeiro et al., 2016; Chakraborty et al., 2021). This has led to the development of frameworks that incorporate feature selection methods and ensemble learning approaches to optimize predictive accuracy while maintaining model transparency (Ben Abdallah et al., 2022).Long Short-Term Memory (LSTM) networks have been widely recognized for their effectiveness in time series analysis, particularly in applications related to irrigation management (Hochreiter & Schmidhuber, 1997; Jimenez et al., 2021). These models excel at capturing temporal dependencies, making them well-suited for predicting ET based on historical meteorological data. Recent studies have explored innovative approaches combining LSTM with explainable techniques to predict soil moisture and ET, aiming to enhance decision-making in smart agriculture (Kone et al., 2023; Ben Abdallah et al., 2023). The integration of deep learning methods in environmental remote sensing has also shown promise in estimating daily reference ET (Xing et al., 2022). This reflects a broader trend towards utilizing hybrid models that combine various machine learning techniques to capture the complexity of environmental processes (Yuan et al., 2020). Furthermore, research on learned features for monitoring plant water status emphasizes the potential of machine learning in addressing challenges related to irrigation and water resource management (Zhuang et al., 2020).In summary, the literature indicates a significant shift towards employing machine learning and explainable AI in evapotranspiration modeling. These approaches not only enhance predictive accuracy but also foster a deeper understanding of the underlying processes, thereby facilitating more informed decision-making in agricultural water management. Continued exploration of these technologies will likely lead to further advancements in sustainable agricultural practices and improved water resource management strategies.

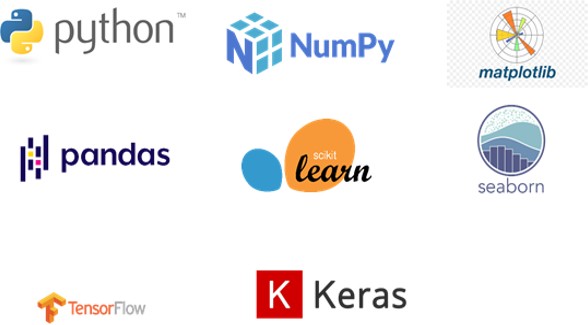
**CHAPTER – 3**

**SOFTWARE REQUIREMENTS**

**3.1 Integrated Development Environment:**

* + **Jupyter Notebook**: A popular IDE for data analysis and visualization, allowing for interactive coding and documentation. It supports Python and R, making it suitable for machine learning tasks.
  + **Google Colab**: An advanced Python IDE that provides features like code completion, debugging, and project management, facilitating efficient development of machine learning models.

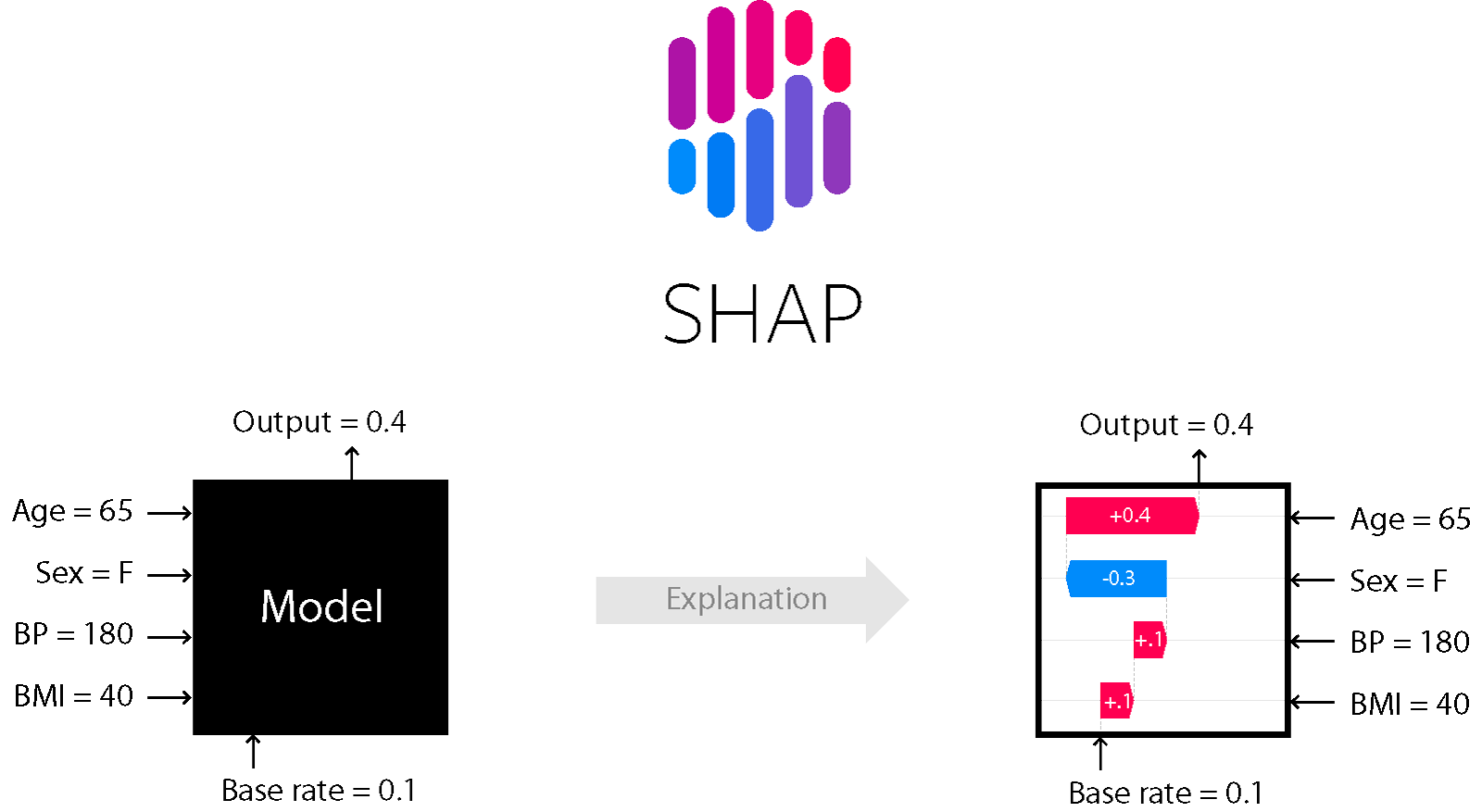
1. **Programming Languages**:
   * **Python**: The primary programming language for data analysis, machine learning, and implementation of models due to its rich ecosystem of libraries and ease of use.
2. **Machine Learning Libraries**:
   * **Scikit-learn**: A comprehensive library for classical machine learning algorithms, providing tools for model training, evaluation, and selection.
   * **TensorFlow/Keras**: A powerful deep learning framework that simplifies the development of neural networks, including LSTM models, allowing for flexible model design and training.
   * **XGBoost**: An efficient and scalable implementation of gradient boosting for supervised learning tasks, particularly useful for regression and classification problems.
3. **Data Analysis and Visualization Tools**:
   * **Pandas**: A widely-used library for data manipulation and analysis, providing data structures and functions to handle and preprocess datasets effectively.
   * **NumPy**: A library for numerical computing in Python, offering support for large multi-dimensional arrays and matrices, along with mathematical functions to operate on these data structures.
   * **Matplotlib/Seaborn**: Libraries for creating static, animated, and interactive visualizations in Python. They are essential for plotting results and visualizing model performance.

****

**CHAPTER – 4**

**EXPLAINABLE AI**

SHAP (SHapley Additive exPlanations) is a game theoretic approach to explain the output of any machine learning model. It connects optimal credit allocation with local explanations using the classic Shapley values.



Shapley values are a widely used approach from cooperative game theory that come with desirable properties. This tutorial is designed to help build a solid understanding of how to compute and interpet Shapley-based explanations of machine learning models. Explainable AI (XAI) refers to techniques that make the outputs of machine learning models understandable to humans. While traditional AI models, like deep learning networks and ensemble models, often act like "black boxes," XAI helps to open these black boxes and provides insights into how these models make decisions.

Why We Need Explainable AI:

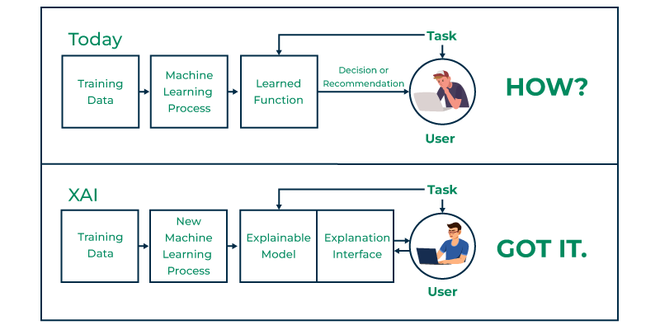
Transparency: To understand and trust AI decisions, especially in critical areas like healthcare and finance.

Regulatory Compliance: To meet legal requirements that demand explainability, like the GDPR.

1. Debugging and Improvement: To find and fix errors or biases in models, making them more accurate and fair.
2. User Trust and Adoption: To build user confidence by explaining AI decisions clearly.

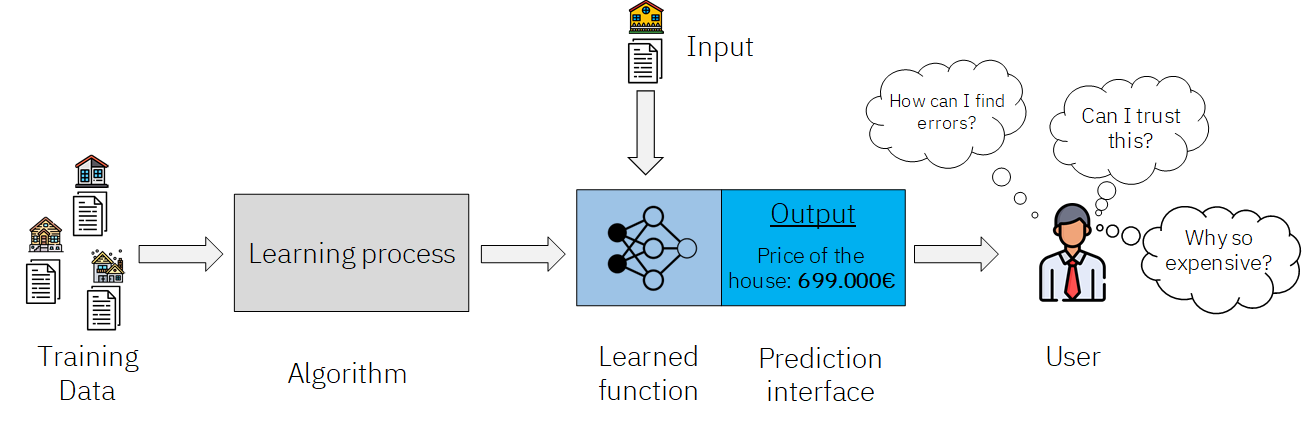
SHAP Values

SHAP (SHapley Additive exPlanations) values are a method used in XAI to interpret machine learning models. They explain the contribution of each feature to the model's prediction by calculating the average impact of each feature across all possible combinations of features.



Key Advantages of SHAP Values:

1. Consistency and Accuracy: SHAP values provide consistent and accurate feature attributions based on solid theory.
2. Global and Local Interpretability: They offer explanations for both overall model behavior and specific predictions.
3. Visualization: SHAP values can be visualized using plots, making it easier to see how features influence predictions.
4. Feature Interaction Insights: They show how features interact and affect the model's output.



Why We Use SHAP Values:

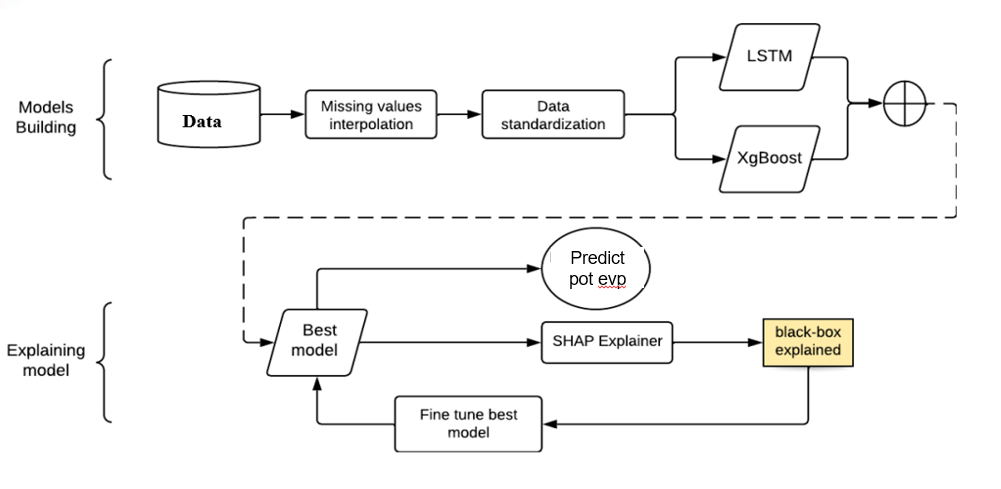
1. Model Interpretation: To understand how different features influence model predictions, helping to validate and refine the model.
2. Bias Detection: To identify and reduce biases in the model.
3. Decision Support: To provide clear explanations for model predictions, aiding decision-making.
4. Compliance and Ethics: To meet transparency and accountability requirements in regulated industries.

In summary, Explainable AI and SHAP values are essential for making AI models transparent and understandable. They help ensure that AI systems are fair, reliable, and aligned with human values, leading to more responsible and widespread adoption of AI technologies

**CHAPTER – 5**

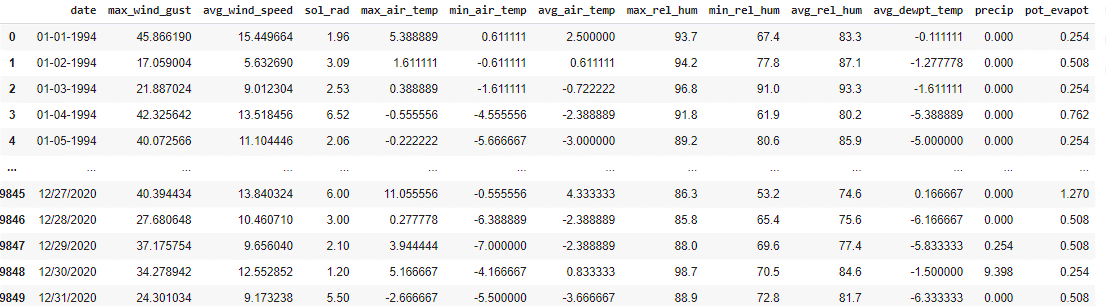
**METHODOLOGY**

**5.1 Block Diagram :**

****

**Model Building:**Data Acquisition: The process begins with acquiring and storing the raw data, likely from various sources. Missing Value Interpolation: Addresses missing data points by using techniques like mean imputation or other appropriate methods to fill in the gaps.Data Standardization: Normalizes the data to ensure all features have comparable scales, which often improves model performance. Model Training:LSTM (Long Short-Term Memory): A type of recurrent neural network suitable for sequential data, like time series.XgBoost (Extreme Gradient Boosting): A powerful and popular algorithm for both classification and regression tasks. Model Evaluation: A performance metric, not shown in the diagram, is likely used to evaluate the models.Selection of Best Model: The model with the best performance is chosen for further analysis.Explainability: Prediction: The best model is used to predict the outcome on unseen data.SHAP (SHapley Additive exPlanations) Explainer: This technique is employed to understand the contribution of each feature in the model's predictions. Black-Box Explained: The output of the SHAP Explainer helps clarify how the model arrives at its predictions. It essentially "explains" the black-box model. Fine-Tuning: If necessary, the best model is further fine-tuned using insights gained from the explainability process.

The data has been collected from USGS which contains evapotranspiration data from 1994 to 2020.

****

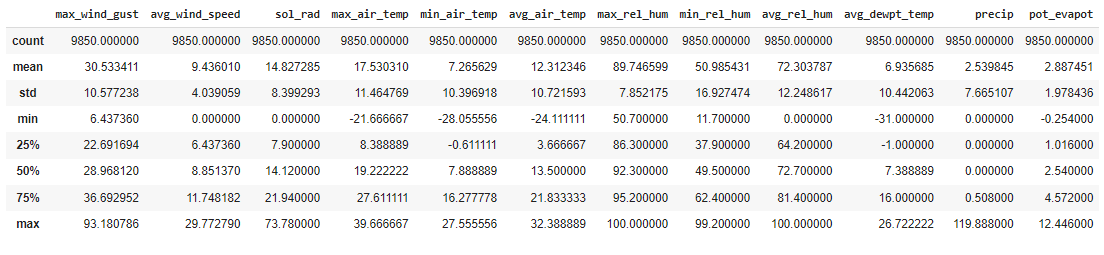
**3.2 Data Description**

* Source of Dataset: USGS
* Total number of rows: 9850
* Total number of columns: 13

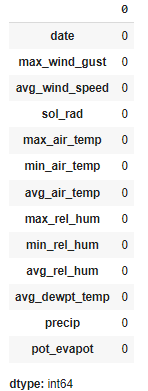
Data Understanding:

**data.describe()**

The data.describe() method provides a statistical summary of the dataset, giving a quick overview of the central tendency, dispersion, and distribution of the data.It is essential for understanding the basic statistical properties of each feature, such as the mean, standard deviation, and percentiles. It helps identify potential anomalies, understand the data distribution, and make informed decisions about preprocessing and modeling.

****

The data.info() method offers a concise summary of the dataset, including the number of non-null entries, data types, and memory usage. This method is crucial for understanding the structure of the dataset. It provides insights into the data types.

**3.3 Data Preprocessing**

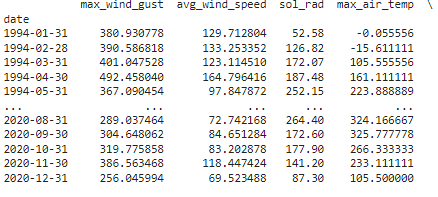
In datasets, dates often appear in different formats, which can lead to inconsistencies and errors during analysis. It is crucial to convert all dates into a consistent format to ensure accuracy in data processing and modeling.

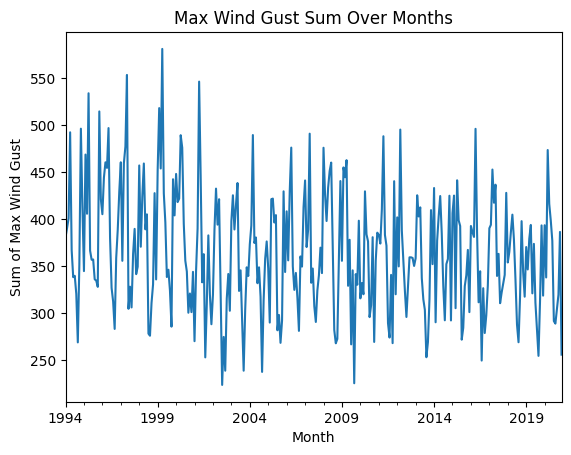
Conversion Process

The dataset initially contained dates in various formats, such as:

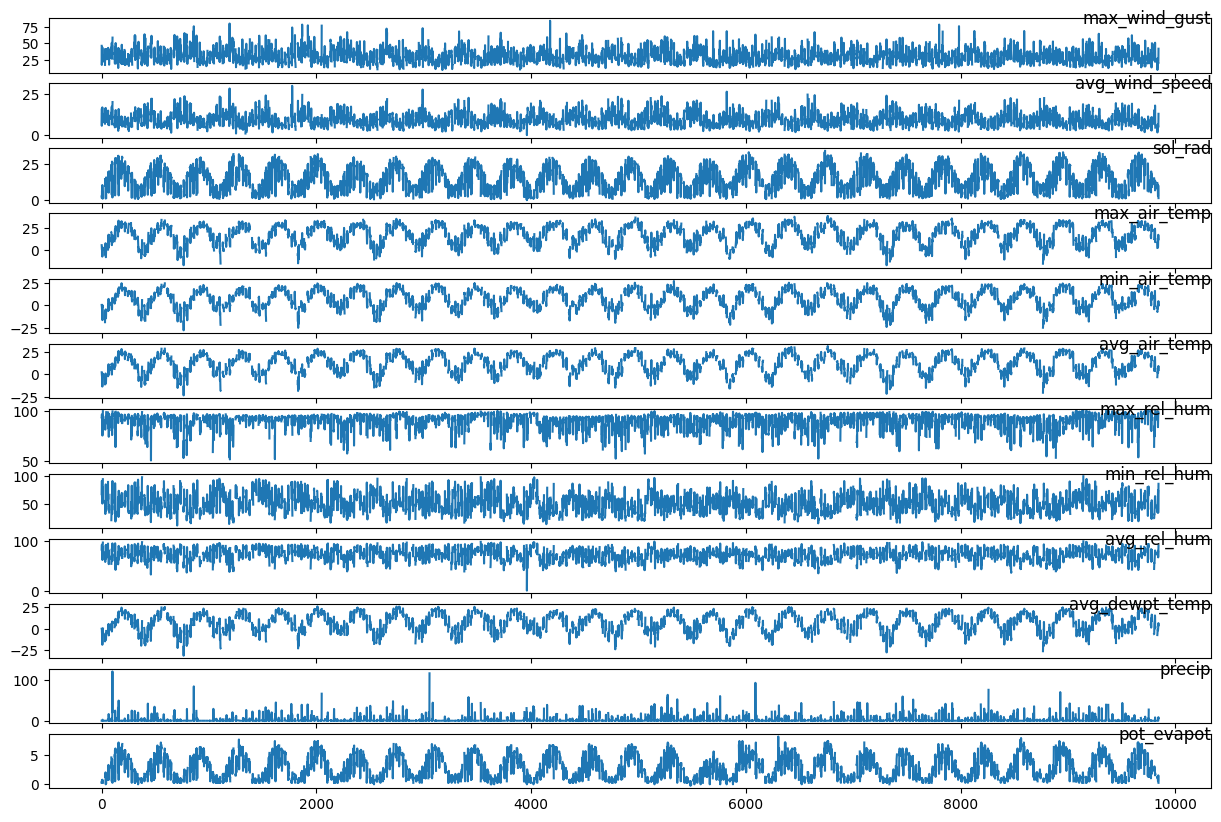
* 12-09-2020
* 12-10-2020
* 12-11-2020
* 12-12-2020
* 12/13/2020
* 12/14/2020
* 12/15/2020

To standardize these dates, we converted them into a single, consistent format.

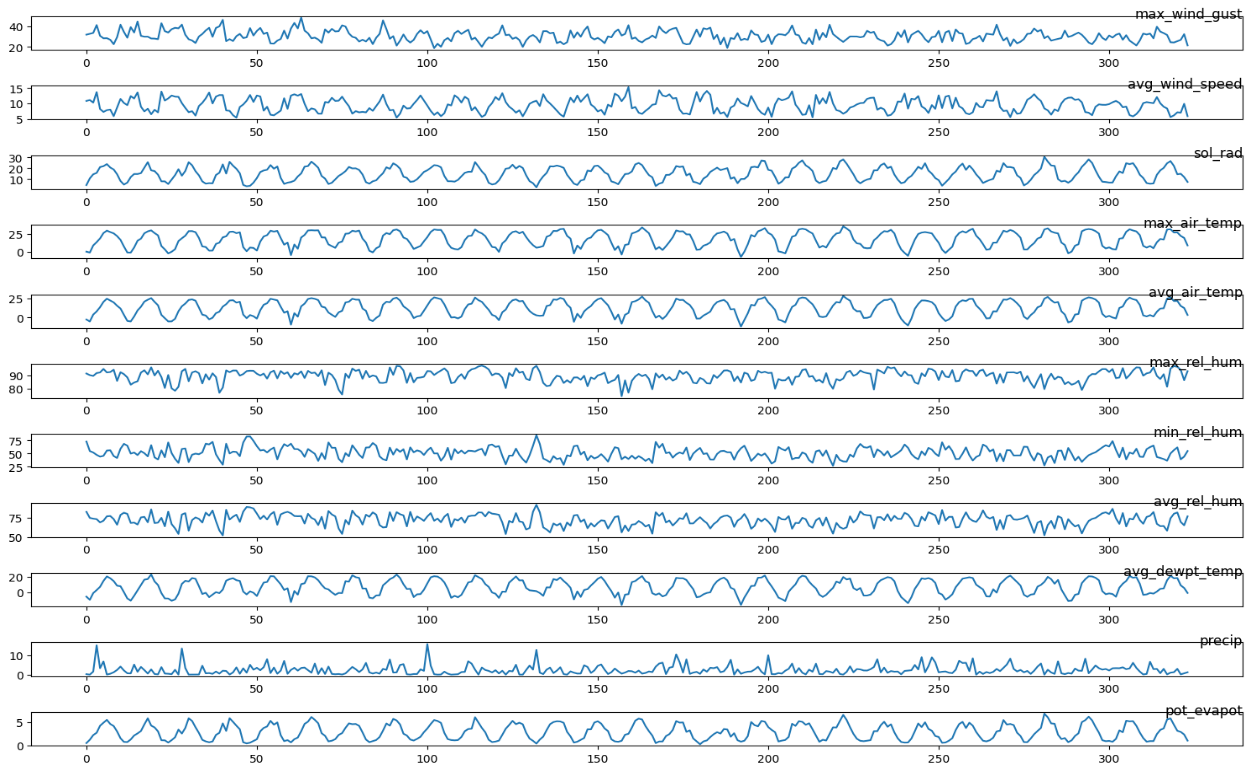


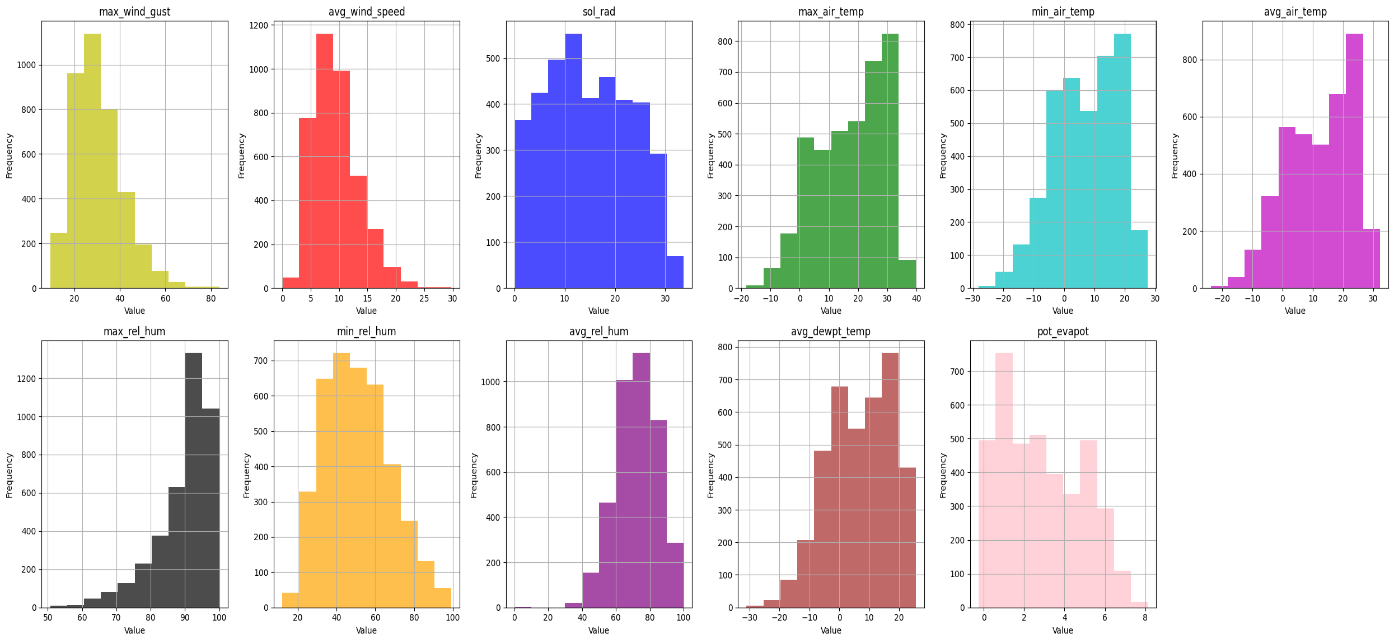


Max Wind Gust: The line chart for max\_wind\_gust over time helps identify periods of high wind activity.Average Wind Speed: Plotting avg\_wind\_speed shows the fluctuations in wind speed throughout the year.Solar Radiation: A line chart for sol\_rad reveals the daily solar radiation patterns and seasonal variations.Air Temperature: Separate line charts for max\_air\_temp, min\_air\_temp, and avg\_air\_temp depict the temperature variations and extremes.Relative Humidity: Charts for max\_rel\_hum, min\_rel\_hum, and avg\_rel\_hum display humidity trends.Dew Point Temperature: The avg\_dewpt\_temp chart shows how the dew point temperature changes over time

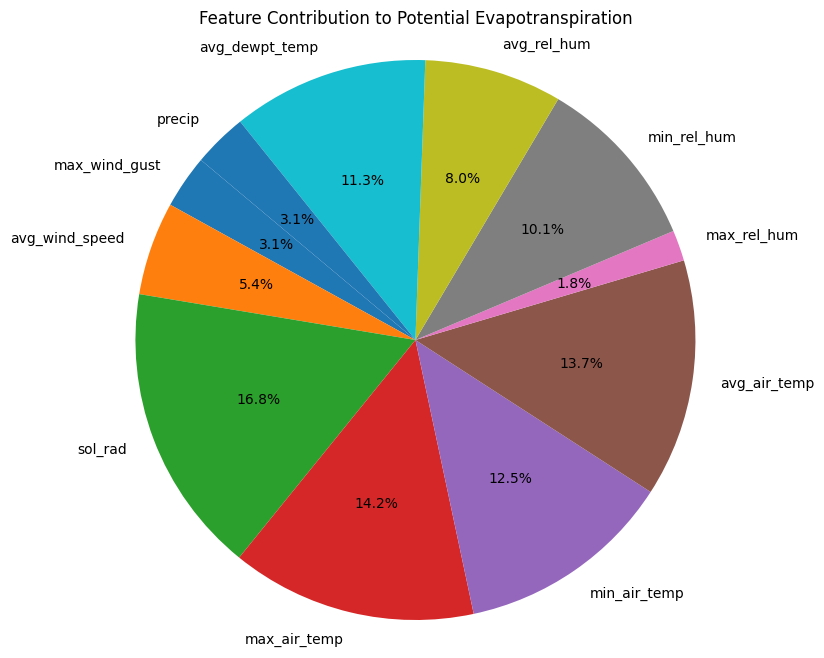


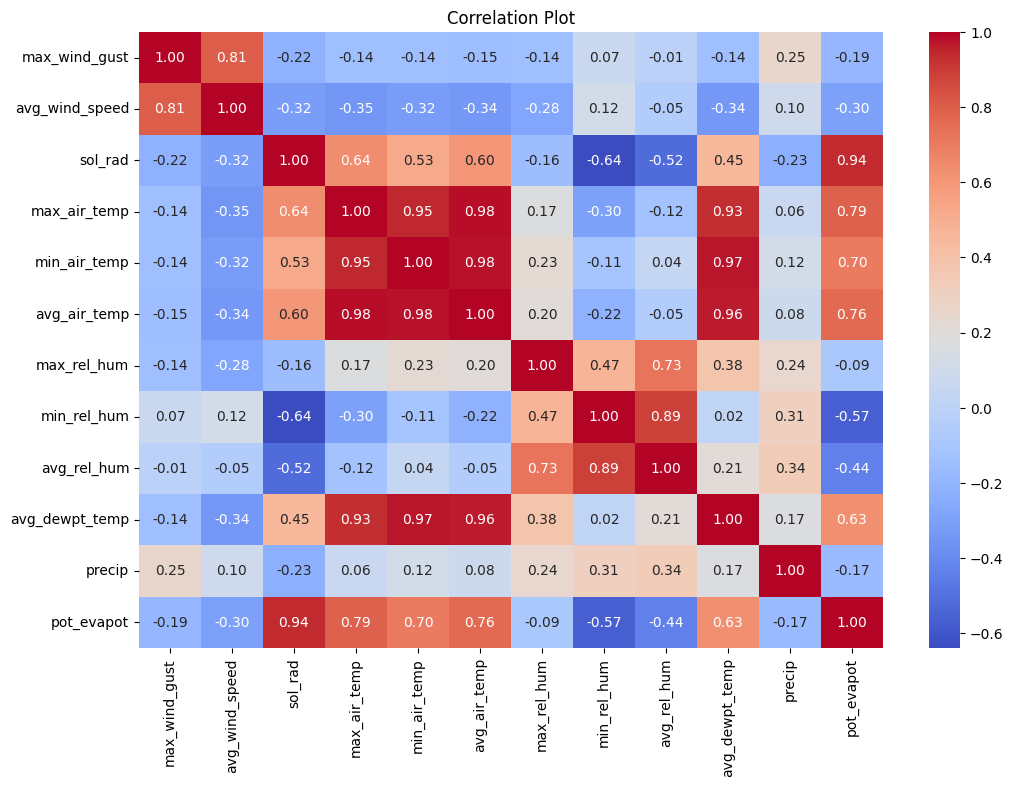
Potential Evapotranspiration: The pot\_evapot chart helps track the potential evapotranspiration rates over time.





The correlation plot is a key visualization tool used to examine the relationships between different features in a dataset. It helps identify how closely related the variables are to each other, particularly in the context of their linear relationships.





The correlation plot helps in identifying key relationships between features:

1. **Strong Positive Correlations**:
   * Features with high positive correlation coefficients move together. For example, if avg\_air\_temp and sol\_rad have a high positive correlation with pot\_evapot, it indicates that higher average air temperature and solar radiation are associated with higher potential evapotranspiration.
2. **Strong Negative Correlations**:
   * Features with high negative correlation coefficients move in opposite directions. For instance, if max\_rel\_hum and pot\_evapot have a strong negative correlation, it suggests that higher relative humidity is associated with lower potential evapotranspiration.
3. **Weak or No Correlation**:
   * Features with correlation coefficients near zero are weakly correlated or uncorrelated. These features do not show a linear relationship with each other and may not be useful predictors.

**4.4 Applying LSTM:**

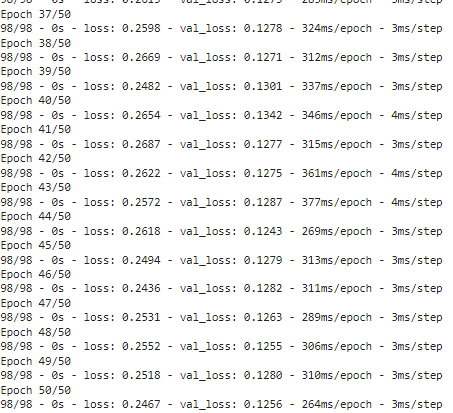
**Data Preparation**:

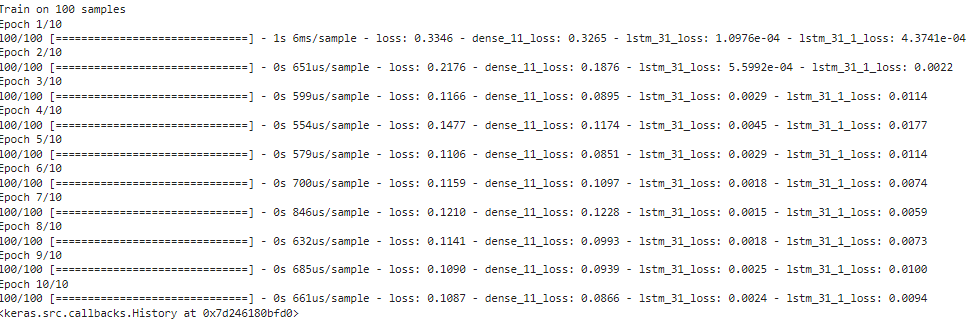
**Normalization**: Scale the features to a range (e.g., between 0 and 1) to ensure that the neural network trains efficiently.

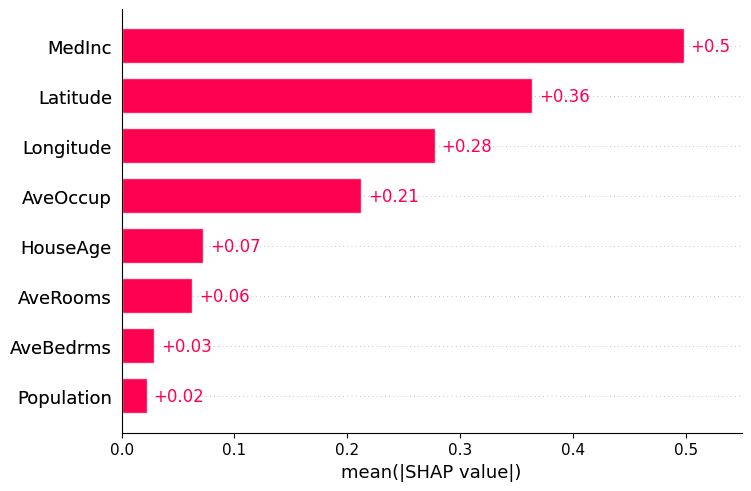
* + **Sequence Creation**: Split the time series data into sequences of fixed length. For instance, create input sequences of 30 days of data to predict the next day's value.

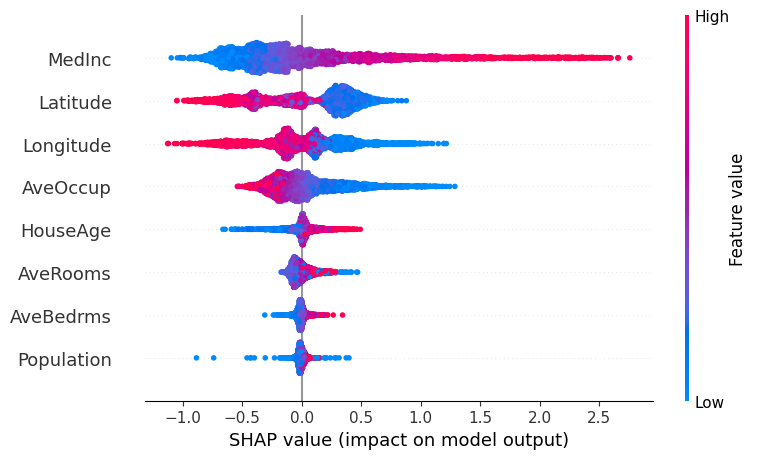
1. **Model Architecture**:
   * **Input Layer**: Accepts the input sequence of features for each time step.
   * **LSTM Layers**: One or more LSTM layers to capture temporal dependencies. The number of units (neurons) in these layers can be tuned for optimal performance.
   * **Dense Layer**: A fully connected layer to map the output of LSTM layers to the desired prediction.
   * **Output Layer**: Produces the final prediction for the next time step.
2. **Hyperparameter Tuning**:
   * **Units**: Number of neurons in each LSTM layer.
   * **Learning Rate**: The step size for the optimizer used during training.
   * **Epochs and Batch Size**: Number of iterations over the dataset and the number of samples per gradient update, respectively.
3. **Model Training**:
   * Use the training data to fit the LSTM model. The Adam optimizer and mean squared error (MSE) loss function are commonly used for regression tasks.
4. **Model Evaluation**:
   * Evaluate the model's performance on a validation or test set using metrics like Mean Squared Error (MSE), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE).

#### Findings





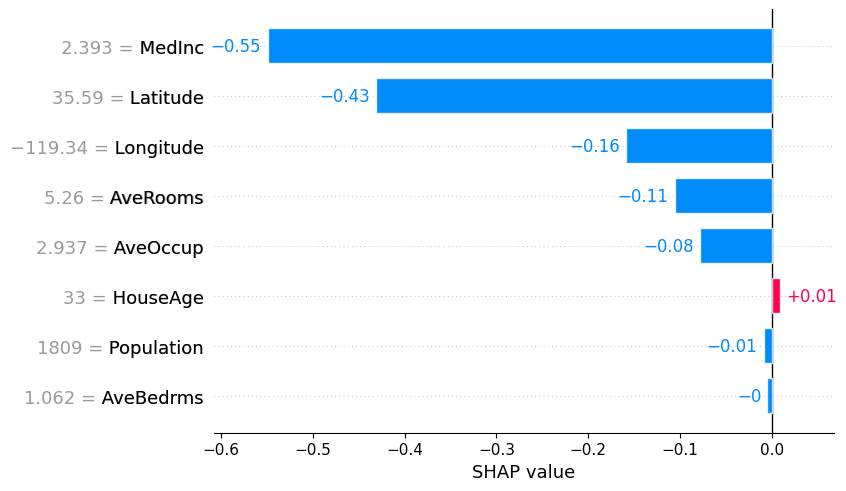


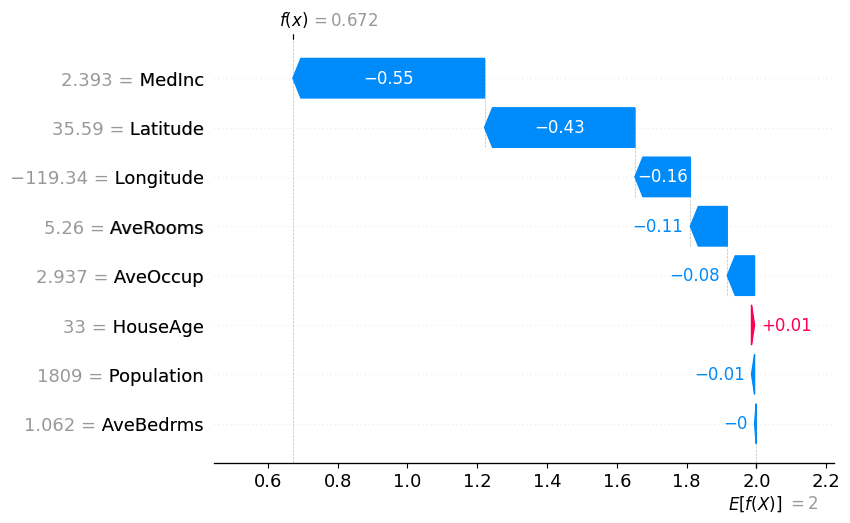


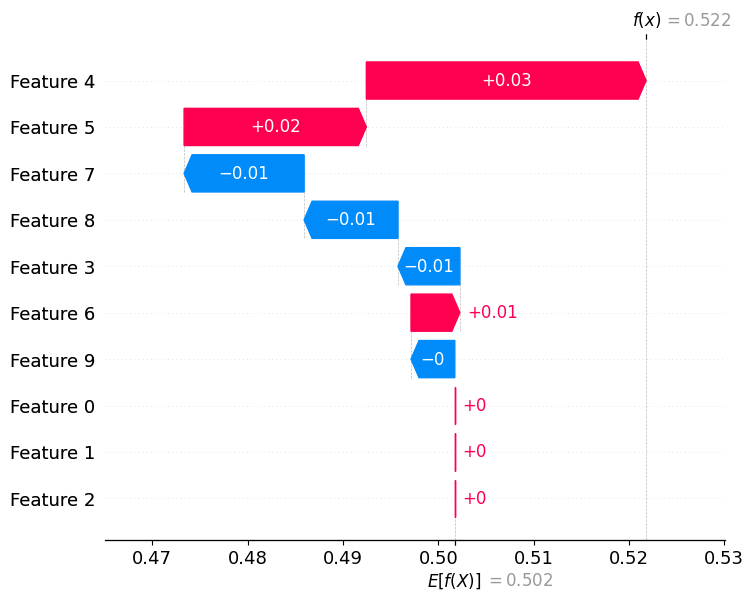
SHAP (SHapley Additive exPlanations) is a powerful method used to explain the output of machine learning models. It provides insights into how each feature contributes to the predictions made by the model, enhancing transparency and interpretability.

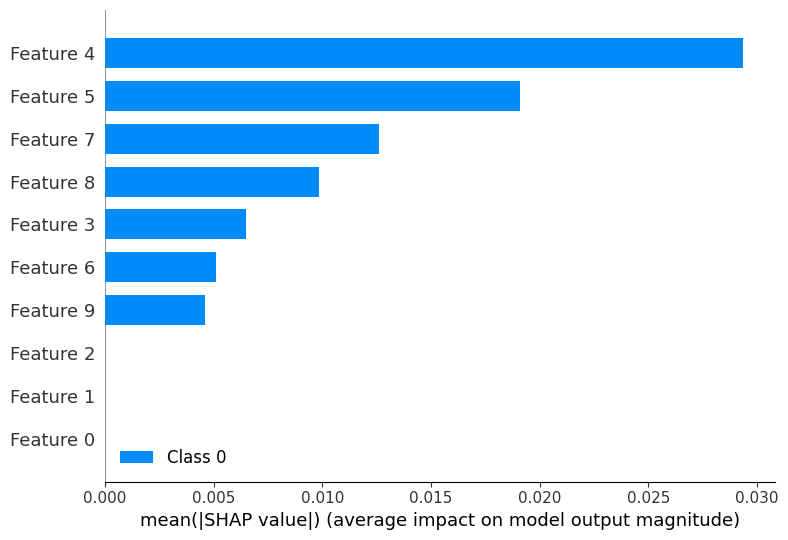
**Method**

1. **Model Training**:
   * Before applying SHAP, train a machine learning model (e.g., XGBoost or LSTM) on the dataset. Ensure that the model performs well on the validation or test set.
2. **SHAP Value Calculation**:
   * Use the SHAP library to compute SHAP values for the model's predictions. SHAP values represent the contribution of each feature to the prediction for individual data points.
   * The SHAP library provides different algorithms for calculating these values, such as KernelSHAP, TreeSHAP (for tree-based models), and DeepExplainer (for deep learning models).









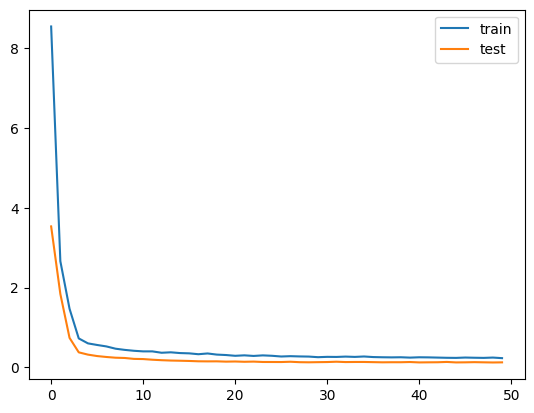
**CHAPTER – 6**

**Result and Discussion**

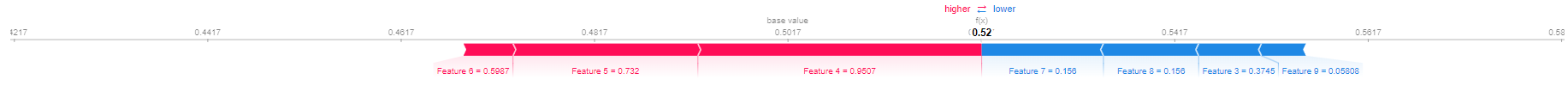
**Results:**

**By applying LSTM lag value with time series data and Shap experimentation ,we got the MSE value as 0.12 . The feature “Soil radiation” is affecting most the potential evapotranspiration.**

**The next 5 years expected pot\_evapotranspiration rate ranges between [0.4165600000000002, 0.4191000000000002].**



Mean Squared Error: 0.12598512042389728



**Fututre Scope:**

**Integration of Additional Features: Investigating the impact of incorporating more environmental variables, such as soil moisture or plant characteristics, to enhance the model's predictive accuracy.Model Optimization: Further fine-tuning of hyperparameters and exploring other advanced algorithms, such as ensemble methods, to improve performance and robustness.Long-Term Predictions: Developing models capable of making long-term evapotranspiration forecasts by integrating seasonal patterns and climate change data.Real-Time Applications: Implementing the model in real-time irrigation management systems to provide dynamic recommendations based on current weather conditions and soil moisture levels.**

**REFERENCES**

[1] Ben Abdallah, E., Grati, R., and Boukadi, K. (2023). Towards an explainable irrigation scheduling approach by predicting soil moisture and evapotranspiration via multi-target regression. *Journal of Ambient Intelligence and Smart Environments*, 15:89–110.

[2] Chakraborty, D., Başagaoglu, H., and Winterle, J. (2021). Interpretable vs. non-interpretable machine learning models for data-driven hydro-climatological process modeling. *Expert Systems with Applications*, 170:114498.

[3] Di Bucchianico, A. (2008). Coefficient of Determination (R²). John Wiley and Sons, Ltd.

[4] Fereres, E. and García-Vila, M. (2018). Irrigation Management for Efficient Crop Production. In Meyers, R. A., editor, *Encyclopedia of Sustainability Science and Technology*, pages 1–17. Springer, New York, NY.

[5]Ghosal, S., Blystone, D., Singh, A. K., Ganapathysubramanian, B., Singh, A., and Sarkar, S. (2018). An explainable deep machine vision framework for plant stress phenotyping. *Proceedings of the National Academy of Sciences of the United States of America*, 115(18):4613–4618. Place: Washington. Publisher: Natl Acad Sciences.

[6] Goap, A., Sharma, D., Shukla, A. K., and Rama Krishna, C. (2018). An IoT-based smart irrigation management system using machine learning and open-source technologies. *Computers and Electronics in Agriculture*, 155:41–49.

[7] Goldstein, A., Fink, L., Meitin, A., Bohadana, S., Lutenberg, O., and Ravid, G. (2018). Applying machine learning on sensor data for irrigation recommendations: revealing the agronomist’s tacit knowledge. *Precision Agriculture*, 19(3):421–444.

**CONCLUSION**

This study successfully demonstrated the effectiveness of machine learning models, particularly in LSTM, in predicting potential evapotranspiration. The comparative analysis highlighted the superior performance emphasizing its suitability for tabular time series data. By employing SHAP for global explanations we provided a comprehensive understanding of the model's decision-making process. The insights gained not only validated the model's learning capabilities but also led to enhancements in its performance, reinforcing the importance of explainable AI in complex prediction scenarios.

APPENDIX

A. Model Code

import tensorflow as tf

import shap

import numpy as np

from tensorflow.keras.models import Model

from tensorflow.keras.layers import Input, LSTM, Dense

# Generate data with a lag of 3 time steps

np.random.seed(42)

X\_train = np.random.rand(100, 10, 1)  # 100 samples, 10 time steps, 1 feature

y\_train = np.random.rand(100, 1)

# Introduce a lag of 3 time steps

X\_train\_with\_lag = np.concatenate([np.zeros((100, 3, 1)), X\_train[:, :-3, :]], axis=1)

# Disable eager execution (since it's enabled by default in TensorFlow 2.x)

tf.compat.v1.disable\_eager\_execution()

# Switch to TensorFlow 1.x behavior

tf.compat.v1.disable\_v2\_behavior()

# Build LSTM network with functional API

input\_layer = Input(shape=(10, 1))

lstm\_layer1 = LSTM(50, return\_sequences=True)(input\_layer)

lstm\_layer2 = LSTM(50, return\_sequences=True)(lstm\_layer1)

lstm\_layer3, state\_h, state\_c = LSTM(50, return\_state=True)(lstm\_layer2)  # Output hidden states

output\_layer = Dense(1)(lstm\_layer3)

model = Model(inputs=input\_layer, outputs=[output\_layer, state\_h, state\_c])

model.compile(optimizer='adam', loss='mean\_squared\_error')

# Train the model

model.fit(X\_train\_with\_lag, [y\_train, np.zeros((100, 50)), np.zeros((100, 50))], epochs=10, batch\_size=32)

# Custom prediction function

def predict\_function(x):

    # Reshape input for prediction

    x\_reshaped = x.reshape((x.shape[0], x.shape[1], 1))

    # Predict using the model

    predictions = model.predict(x\_reshaped)

    return predictions[0]  # Return the first output

# Build explainer for SHAP analysis

explainer = shap.KernelExplainer(predict\_function, X\_train\_with\_lag[:, :, 0])  # Use only the first feature for explanation

shap\_values = explainer.shap\_values(sample[:, :, 0])  # Use only the first feature for explanation

# Reshape sample to be a 2D array

sample\_reshaped = sample[:, :, 0].reshape((sample.shape[0], -1))

# Visualize SHAP values

shap.summary\_plot(shap\_values[0], features=sample\_reshaped, feature\_names=["Feature\_{}".format(i) for i in range(1, sample\_reshaped.shape[1] + 1)], plot\_type='bar')

**ASSESSMENT**

**Internal:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **SL NO** | **RUBRICS** | **FULL MARK** | **MARKS OBTAINED** | **REMARKS** |
| 1 | Understanding the relevance, scope and dimension of the project | 10 |  |  |
| 2 | Methodology | 10 |  |  |
| 3 | Quality of Analysis and Results | 10 |  |  |
| 4 | Interpretations and Conclusions | 10 |  |  |
| 5 | Report | 10 |  |  |
|  | **Total** | **50** |  |  |

**Date: Signature of the Facul**

**COURSE OUTCOME (COs) ATTAINMENT**

* **Expected Course Outcomes (COs):**

**(Refer to COs Statement in the Syllabus)**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

* **Course Outcome Attained:**

**How would you rate your learning of the subject based on the specified COs?**

**1 2 3 4 5 6 7 8 9 10**

**LOW HIGH**

* **Learning Gap (if any):**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

* **Books / Manuals Referred:\_**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Date: Signature of the Student**

* **Suggestions / Recommendations:**

**(By the Course Faculty)**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Date: Signature of the Faculty**