



**ADSP**

# CORN YIELD FORECAST

---

Parsa Taati

Hamed Goldoust

Afsoon Abbasi



**Politecnico  
di Torino**



PASSION FOR INNOVATION

# Project Value Proposition

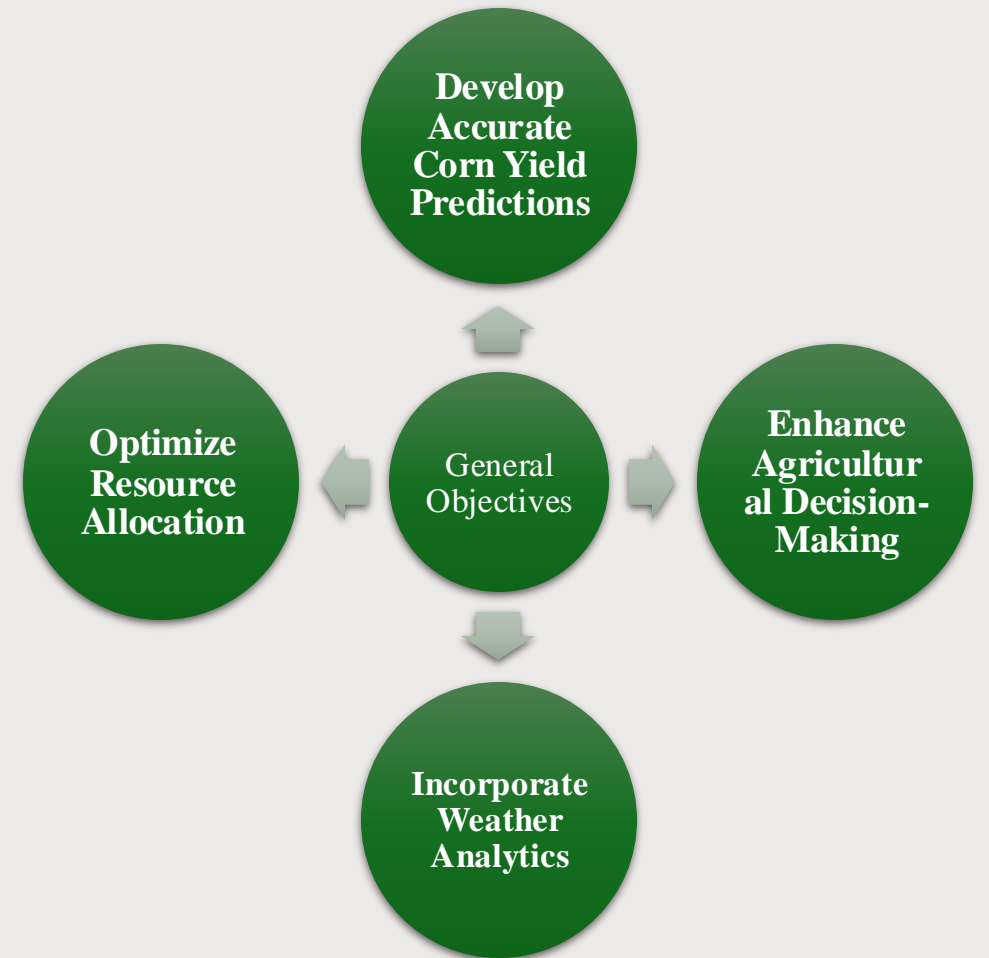
**For farmers,** internal team members, and academic advisors, **who** require actionable insights for crop management, reliable datasets for accurate model development, and well-documented progress for academic evaluation,

**our project, Corn Yield Forecast,** is a data-driven decision-support system leveraging predictive models and weather analytics,

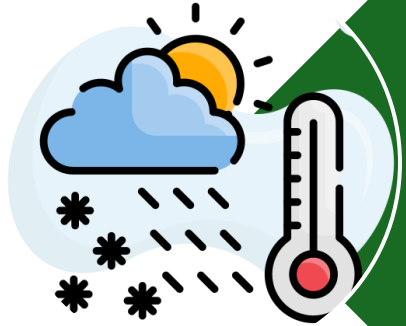
**that** provides optimal planting and harvesting recommendations, clean and accessible datasets, and comprehensive project documentation to enhance agricultural efficiency, streamline development processes, and ensure academic rigor.

Unlike existing solutions, our system integrates advanced predictive analytics with user-centric design, ensuring accurate yield forecasting, seamless data accessibility, and actionable insights tailored to the unique needs of farmers, researchers, and development teams, while promoting sustainability and academic excellence.

# General objectives



# Hypotheses

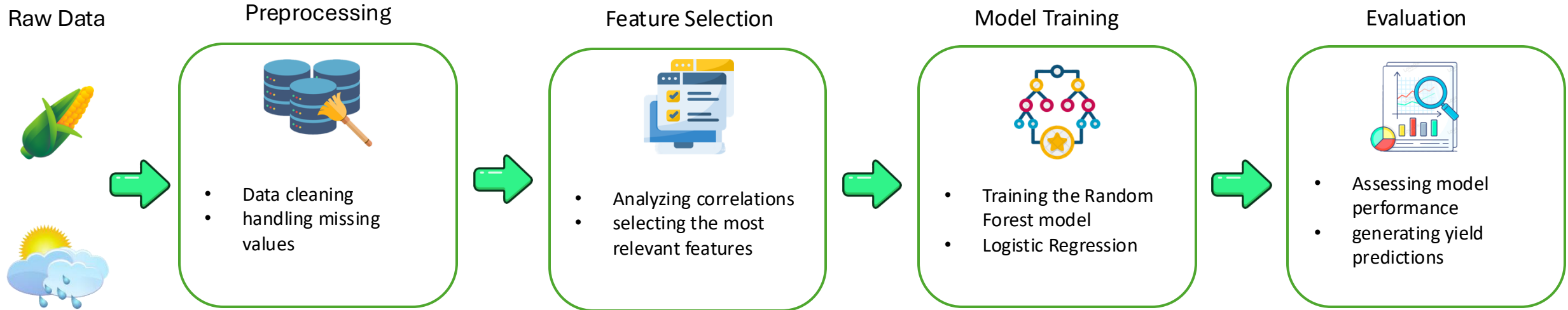


**Higher cumulative  
rainfall positively  
correlates with corn yield**



**Solar radiation  
significantly impacts  
crop growth**

# Random Forest Pipeline for Corn Yield Prediction:



## Step 01

- Data Merging



## Step 02

- EDA
- Outlier Detection
- Normalization
- Feature Selection



**We are doing**



## Step 03

- Split features and targets
- Model Selection



## Step 04

- Evaluation

# Merge Datasets:

## Features :

- ☐ YEAR
- ☐ FAO Cicle
- ☐ Month
- ☐ Frost Days
- ☐ Heavy Rain Events
- ☐ Average Temperature
- ☐ Maximum Temperature
- ☐ Minimum Temperature
- ☐ Growing Degree Days
- ☐ Heat Stress Days
- ☐ Cold Stress Days
- ☐ Temperature Variability
- ☐ Total Precipitation
- ☐ Rainy Days
- ☐ Drought Days
- ☐ Maximum Daily Rainfall

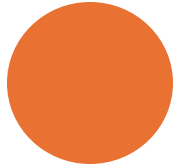
- ☐ Rainfall Variability
- ☐ Consecutive Dry Days
- ☐ Cumulative Solar Radiation
- ☐ Average Solar Radiation
- ☐ Potential Evapotranspiration
- ☐ Average Wind Speed
- ☐ Maximum Wind Speed
- ☐ Average Relative Humidity
- ☐ Humidity Variability

☐ Fresh Ear Amount

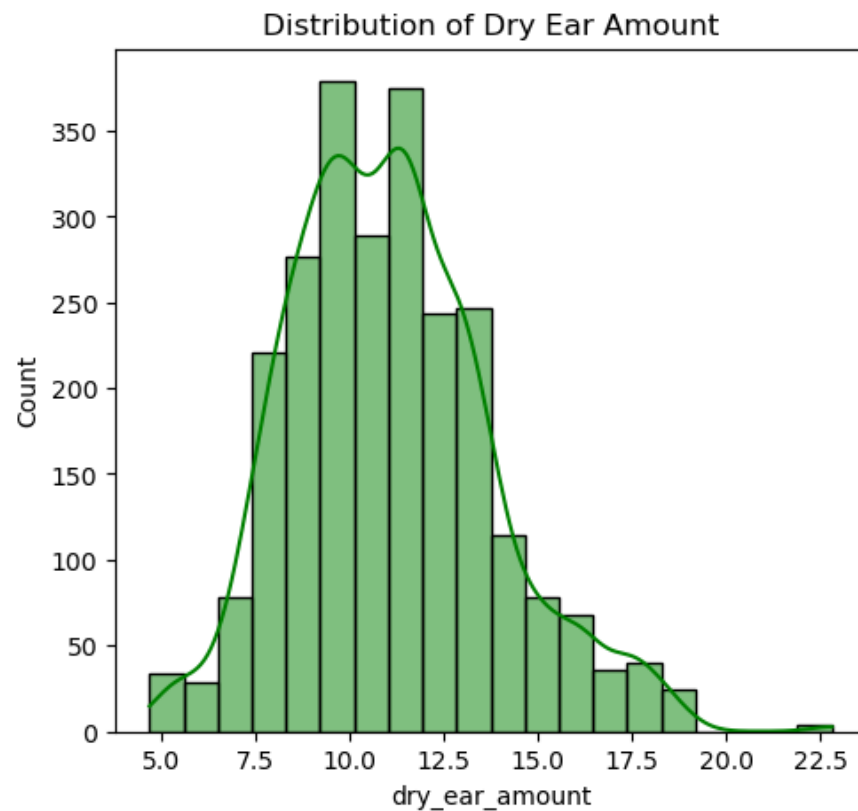
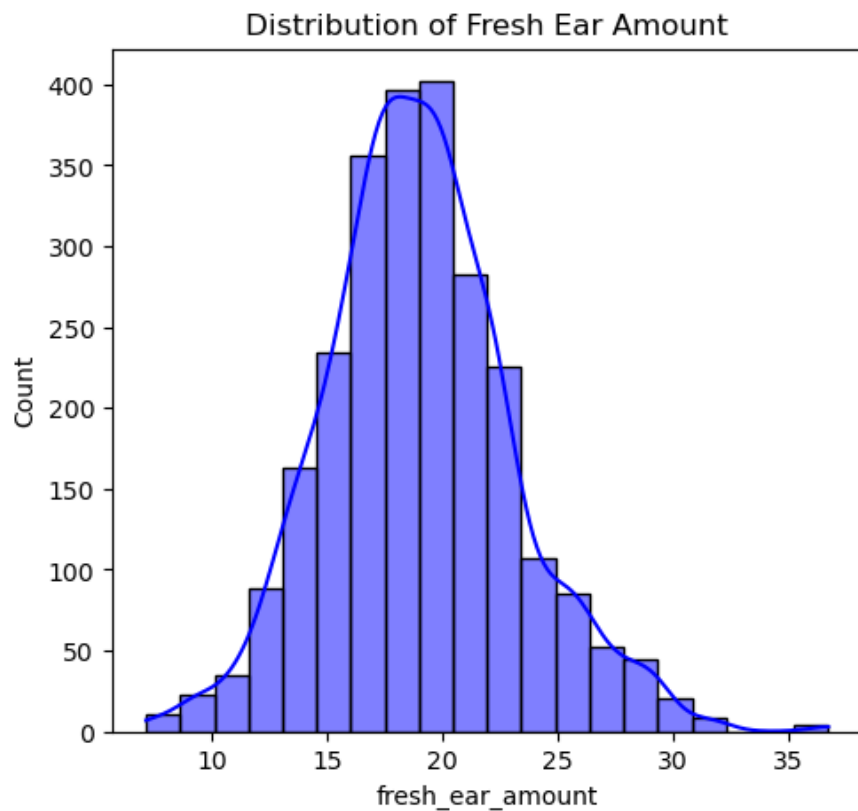
☐ Dry Ear Amount



Target Columns

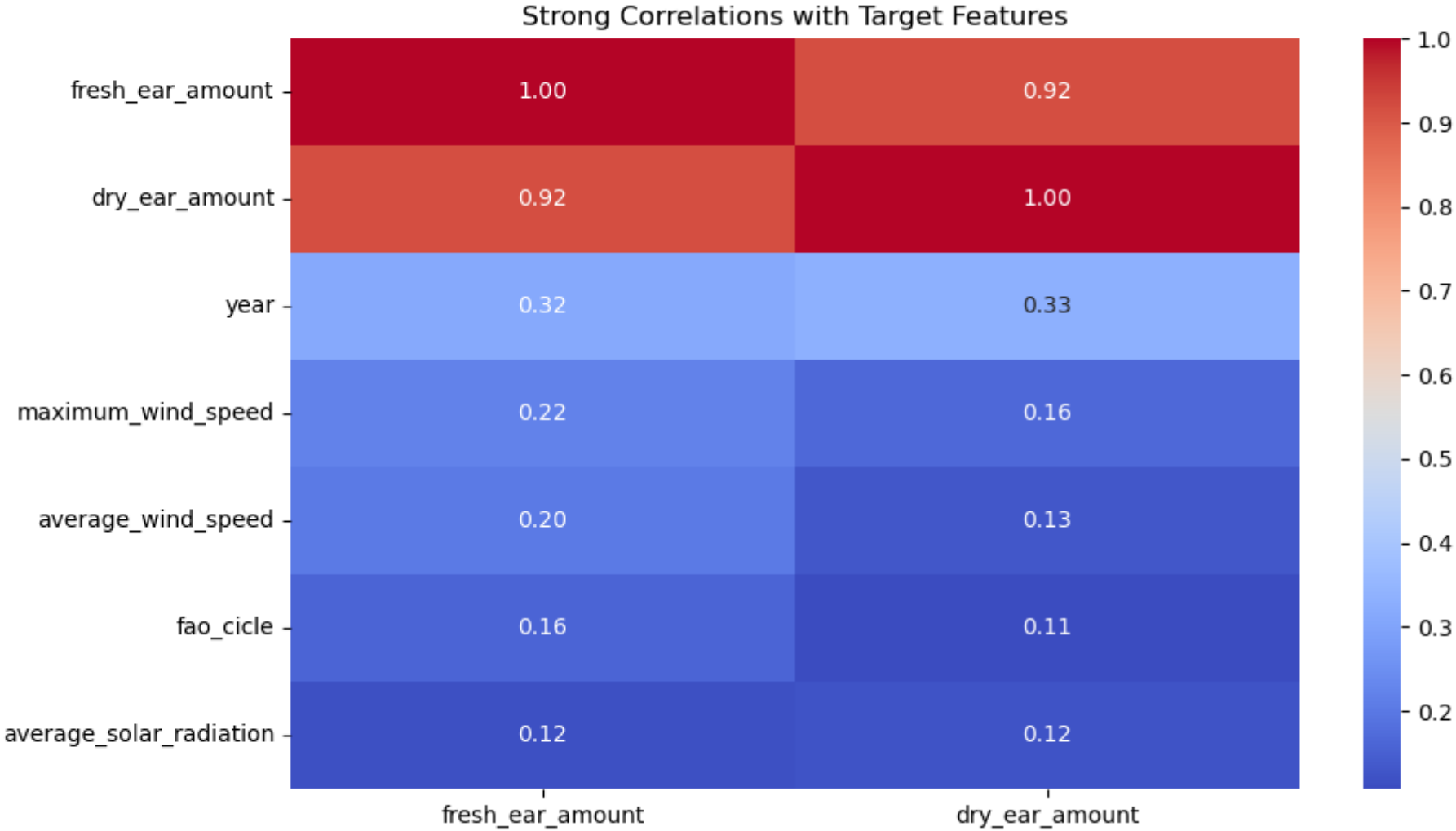


# Target Normal Distribution





# Strong Correlation with Target Features





# Dataset Split

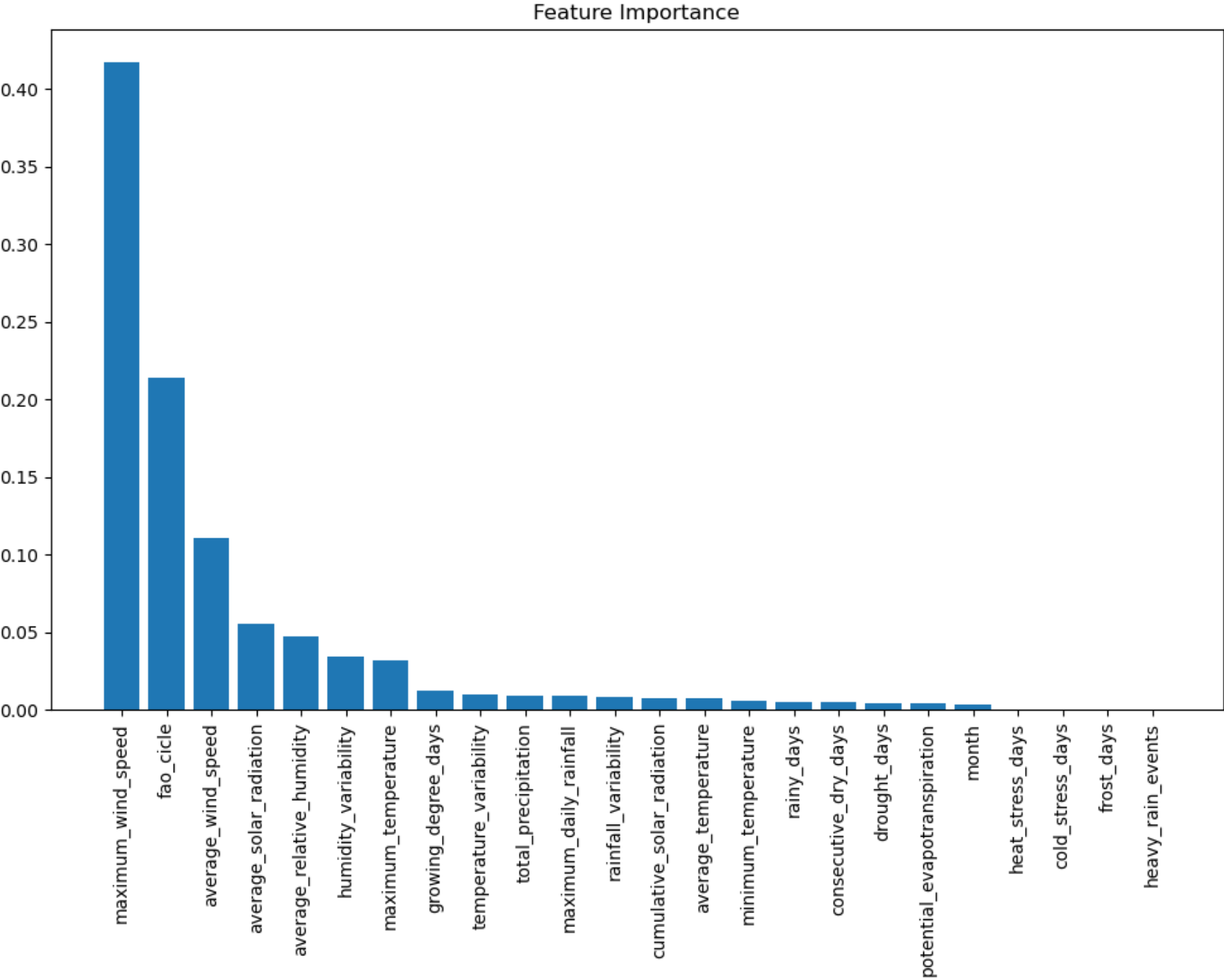
- Dataset Duration: 2011-2023
- Train\_Set: 2011-2020
- Validation\_Set: 2021-2023

# Method

- **Random Forest Regression:**
  - **Data Preparation:**
    - Features and targets split from training and validation datasets.
    - Train data further divided into train/test splits (80/20).
  - **Model Training:**
    - Random Forest Regressor initialized with 100 estimators.
    - Model trained on train split.
  - **Evaluation:**
    - Metrics: **Mean Absolute Error (MAE)** and **Root Mean Squared Error (RMSE)**.
  - **Hyperparameter Tuning:**
    - GridSearchCV with 3-fold cross-validation used to find optimal hyperparameters:
    - Parameters: *n\_estimators*, *max\_depth*, *min\_samples\_split*, *min\_samples\_leaf*.



# Feature Importance



# **Experiment & Conclusion**



# Result

- **Random Forest results:**

- **Training Result:**

- RMSE: 0.0885
    - MAE: 0.0680

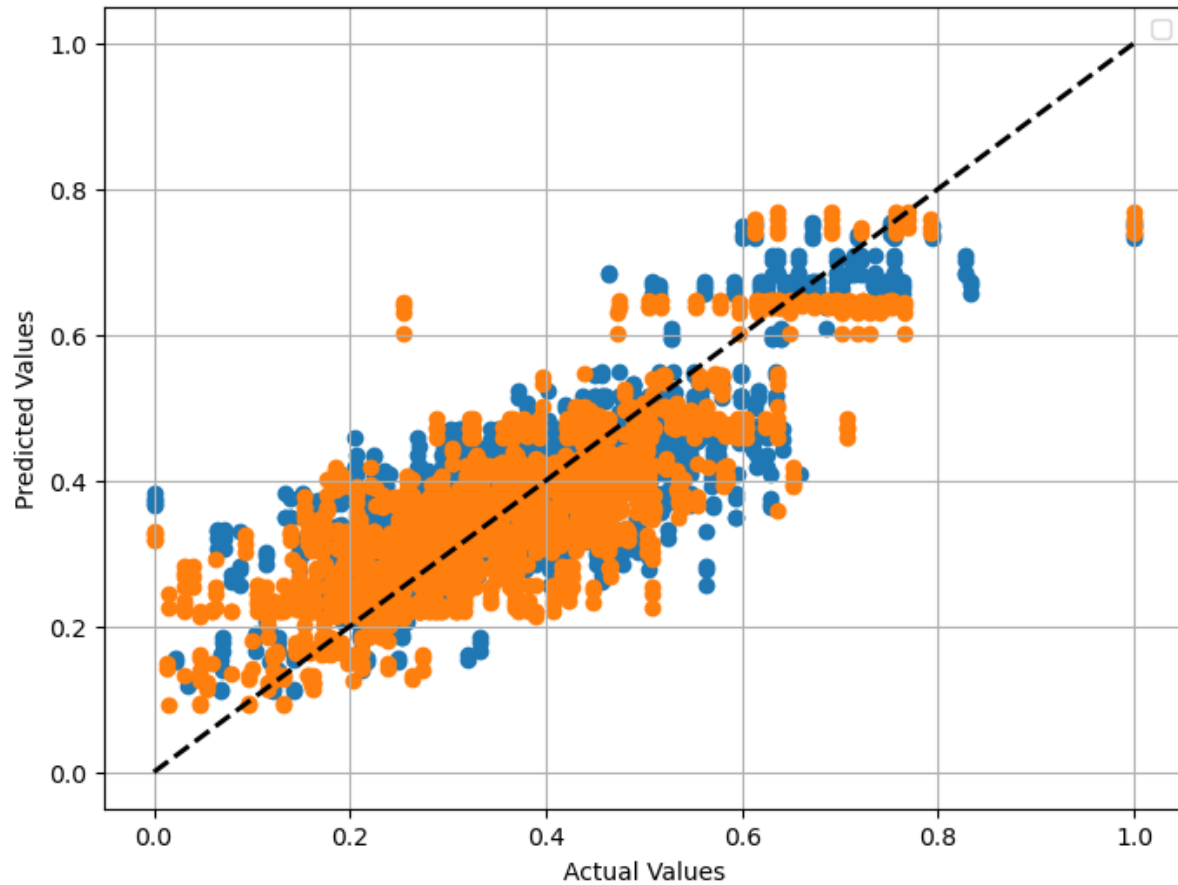
- **Validation Result:**

- RMSE: 0.1864
    - MAE: 0.1487

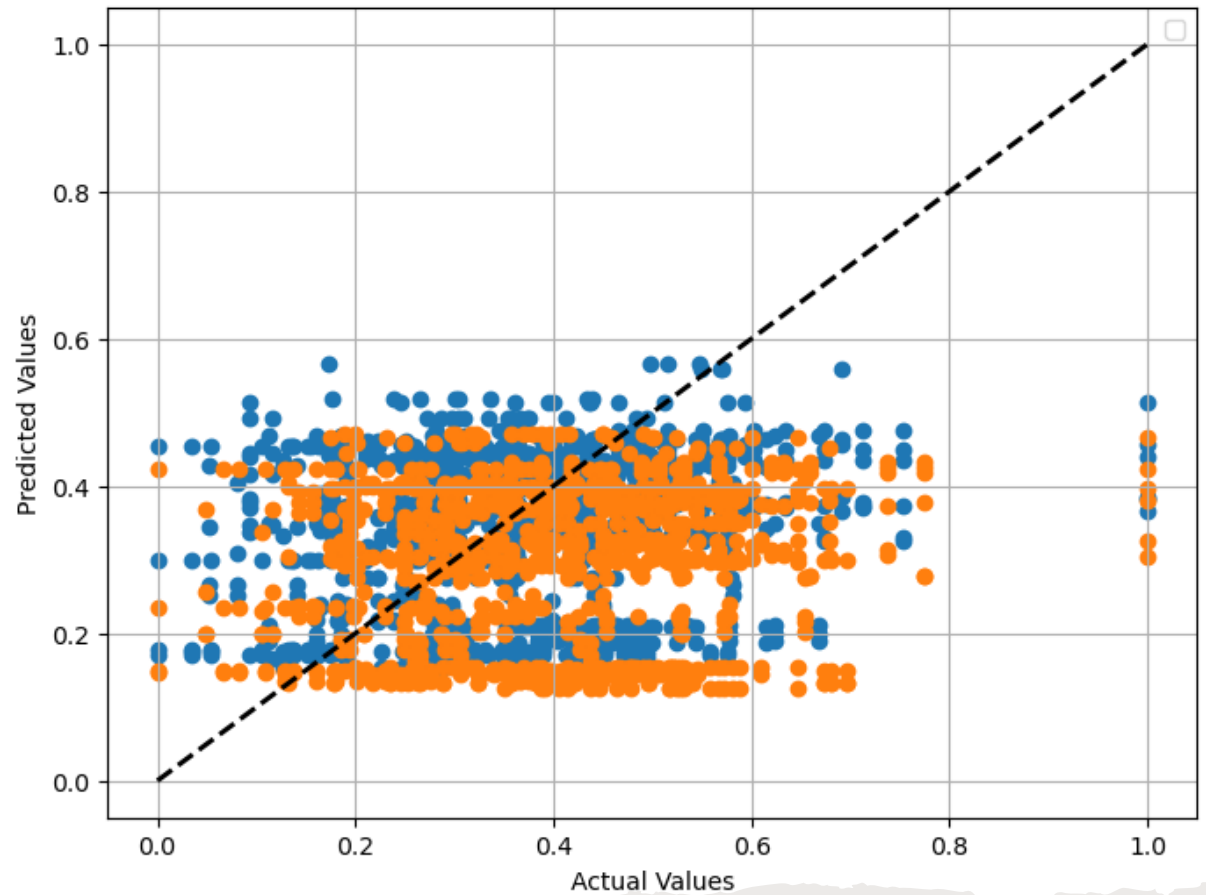


# Actual VS Predicted by Model

Actual vs Predicted (Training)



Actual vs Predicted (Validation)





**Thanks for your  
attention**

---