**Project Report**

**Enhancing Climate Predictions with Advanced AI Algorithms for Optimizing Environmental Policies and Responses**

**Presented By : TEAM DELTA Phase #2**

**ITSOLERA DL PROJECT**

**Introduction**

The effects of climate change are increasingly impacting the planet, demanding more accurate and timely predictions. Traditional climate models often struggle with predicting future scenarios due to their inability to integrate diverse data sources and capture the intricate interactions within the climate system. This project aims to enhance climate simulations using advanced AI algorithms to improve predictive accuracy and support better environmental policy decisions.

**Problem Statement**

Current climate models are insufficient for providing precise, actionable predictions. They often lack the capability to handle large, varied data from sources like satellite observations, historical records, and real-time metrics. This deficiency limits the effectiveness of policies and strategies aimed at mitigating climate change. The objective of this project is to develop AI-based models that can perform complex climate simulations, ensuring real-time predictions with a high degree of accuracy.

**Data Collection**

The data for this project was sourced from Kaggle’s

URL: " <https://www.kaggle.com/datasets/lishuyu/gsod-shanghaichina-5station-collections>" which contains features such as `'DEWP', 'SLP', 'NAME','STP', 'VISIB', 'WDSP', 'MAX', 'MIN', 'PRCP', 'MONTH', 'DAY', 'WEEKDAY','TEMP' etc The dataset includes over 749096 rows x 19 columns observations, offering a rich foundation for climate predictions.

**Data Preprocessing**

To prepare the data for AI modeling, the following preprocessing steps were performed:

**Missing Values:** Checked and removed any missing or irrelevant data points.

**Feature Extraction:** Time-based features were extracted from `Date\_Time` (Month, Day, WeakDay).

**Categorical Encoding:** The `NAME` feature was encoded using label encoding.

**Scaling:** Continuous variables were scaled using `StandardScaler`.

**Data Aggregation:** Aggregated high-frequency data to balance the dataset and improve model efficiency.

The dataset was split into training, validation, and test sets to ensure robust model performance and generalization.

**Model Architecture**

The climate prediction model was built using a GRU (Gated Recurrent Unit) neural network to handle time-series data effectively. The model architecture is designed as follows:

**Input Layer:** Takes preprocessed climate data.

**GRU Layer 1:** First GRU layer with 128 units and return sequences enabled, followed by dropout to prevent overfitting.

**GRU Layer 2:** A second GRU layer with 64 units, followed by another dropout layer.

**Dense Output Layer:** Outputs a single continuous value, which represents the predicted temperature.

The model was trained using the `adam` optimizer and `mean\_squared\_error` as the loss function, ideal for regression tasks.

**Training Process**

The training process was designed to optimize model accuracy and efficiency. Key configurations included:

**Batch Size:** Set to 64 to optimize the training speed while maintaining model performance.

**Epochs:** Trained for 5 epochs with early stopping based on validation loss.

**Validation Split:** 20% of the training data was used for validation to monitor overfitting and performance.

**Patience:** Early stopping was applied with a patience of 10 epochs to avoid unnecessary training time when performance plateaued.

**Performance Evaluation**

The performance of the GRU model was evaluated using several metrics, including:

**Mean Absolute Error (MAE):** To measure the average difference between the predicted and actual values.

**Root Mean Squared Error (RMSE):** To highlight larger errors, which is crucial for identifying significant deviations in climate predictions.

**R² Score:** To evaluate how well the model captures the variability of the data.

The final model was able to make precise climate predictions with low error rates, enabling high-quality decision-making for environmental policies.

**Insights**

**Feature Importance:** SHAP values were used to identify the importance of features such as `Sea level pressure `, `Wind\_Speed\_kmh`, and `Precipitation\_mm`. These features significantly impacted the model’s predictions, providing insights into climate dynamics.

**Temporal Trends:** The model identified temporal patterns, such as seasonal variations in temperature, which can aid policymakers in preparing for extreme weather conditions.

**Policy Recommendations:** With accurate climate predictions, the model can support long-term planning for climate adaptation strategies, optimize the timing of resource allocations, and help governments implement more effective environmental policies.

**Conclusion**

This project demonstrates the successful application of advanced AI algorithms in enhancing climate predictions. By leveraging a GRU model and encoded categorical features, the model achieved high predictive accuracy, enabling faster, more reliable climate simulations. These predictions can significantly improve the quality of environmental policies and responses, contributing to more effective climate change mitigation and adaptation strategies. Further optimization and exploration of additional datasets may further enhance the model's accuracy and generalizability.