Impacts of climate change on food Security

Group No: 17

Group Members Names:

1. Mohamed Fadlelseed
2. Fulchany Arvelin Nanitelamio
3. Bacoro dit Elhadj lansar
4. Madalena Makiesse
5. Zeinab Mohmmed
6. Israa Abdulrahman Mohammed kheir

1 Concept Note

1.1 Project Overview

- **Project Overview**: This project aims to explore the impacts of climate change on food security, focusing on how variations in climate variables affect crop yields and overall food availability. By aligning with Sustainable Development Goals (SDGs) 13 (Climate Action) and 2 (Zero Hunger), the project seeks to provide actionable insights to enhance food security and inform climate adaptation strategies. The project's potential impact includes improving predictive models for agricultural productivity and identifying key factors that influence food security.

- **Problem Statement**: Climate change is increasingly recognized as a major threat to global food security. Rising temperatures, changing precipitation patterns, and more frequent extreme weather events are significantly affecting agricultural productivity, crop yields, and food availability. This, in turn, exacerbates food insecurity, particularly in vulnerable regions such as Africa and South Asia. Despite numerous studies documenting these impacts, there remains a critical need for more accurate and actionable insights into how climate change will affect food security across different regions and time scales.
- Potential Impact of the Solution:

Our project uses machine learning and deep learning to predict climate change impacts on food security. Key benefits include:

- Enhanced Predictive Accuracy: More precise forecasts of climate effects on agriculture for better planning and resource allocation.
- Regional Insights: Detailed understanding of regional food security challenges for tailored adaptation strategies.
- Informed Policy Making: Data-driven recommendations for effective climate change mitigation in agriculture.
- Global Food Security: Supports global goals of Zero Hunger and Climate Action by improving resilience and understanding of climate-related risks.

1.2 Objectives

Objective 1: Develop predictive models to estimate the impact of climate change on crop yields using historical climate and agricultural data.

Objective 2: Identify and analyze key climate variables that significantly affect food security.

Objective 3: Study the Global Hunger Index (GHI): Examine the relationship between key climate variables and the Global Hunger Index to understand their combined effect on global hunger levels..

Objective 4: Enhance understanding of regional vulnerabilities to climate change and suggest targeted interventions.

1.3 Background

- **Contextualizing the Problem:**

Climate change has increasingly become a critical threat to global food security. Rising temperatures, erratic precipitation patterns, and more frequent extreme weather events impact agricultural productivity and food availability. These changes pose significant risks to food supply chains and exacerbate hunger, particularly in vulnerable regions. Understanding these impacts is crucial for developing effective strategies to ensure food security and adapt to the evolving climate.

- **Existing Solutions and Initiatives:**

Several initiatives address climate change's impact on agriculture and food security:

- **Climate Models:** Traditional climate models project future climate scenarios but often lack specific predictions on their effects on food security.

- **Agricultural Research:** Institutions and organizations work on crop resilience and adaptive practices, but they may not fully integrate climate variability.

- **Early Warning Systems:** Systems exist for predicting weather patterns and natural disasters but may not link these predictions directly to food security outcomes.

- **Machine Learning Approach:**

Machine learning offers several advantages for addressing this issue:

1. **Data Integration:** ML models can process and integrate large datasets, including climate data, agricultural productivity records, and hunger indices, to identify complex patterns and relationships.
2. **Predictive Accuracy:** Advanced ML algorithms can provide more precise forecasts of climate impacts on food security, enabling proactive measures.
3. **Adaptability:** Machine learning models can adapt to new data, improving predictions as more information becomes available.
4. **Insight Generation:** ML can reveal insights from data that might be missed by traditional analytical methods, informing better policy and decision-making.

1.4 Methodology

Machine Learning Techniques:

1. **Regression Models:**
    - **Purpose:** Predict crop yields and food security indicators based on climate variables.
    - **Some Algorithms: Examples**
        - **Random Forest:** An ensemble learning method that builds multiple decision trees and combines their predictions to improve accuracy and robustness. It's particularly effective for handling complex and high-dimensional datasets.
        - **Gradient Boosting:** A boosting technique that builds models sequentially, each correcting the errors of the previous one. It helps in achieving higher predictive performance by combining the strengths of multiple weak models.
2. **Deep Learning Models:**
    - **Purpose:** Analyze time-series climate data to capture long-term dependencies and patterns that might affect food security.
    - **Algorithm:**
        - **Long Short-Term Memory (LSTM) Networks:** A type of Recurrent Neural Network (RNN) designed to learn and remember long-term dependencies in time-series data. LSTMs

are effective in capturing temporal relationships and trends in climate variables over extended periods.

Frameworks:

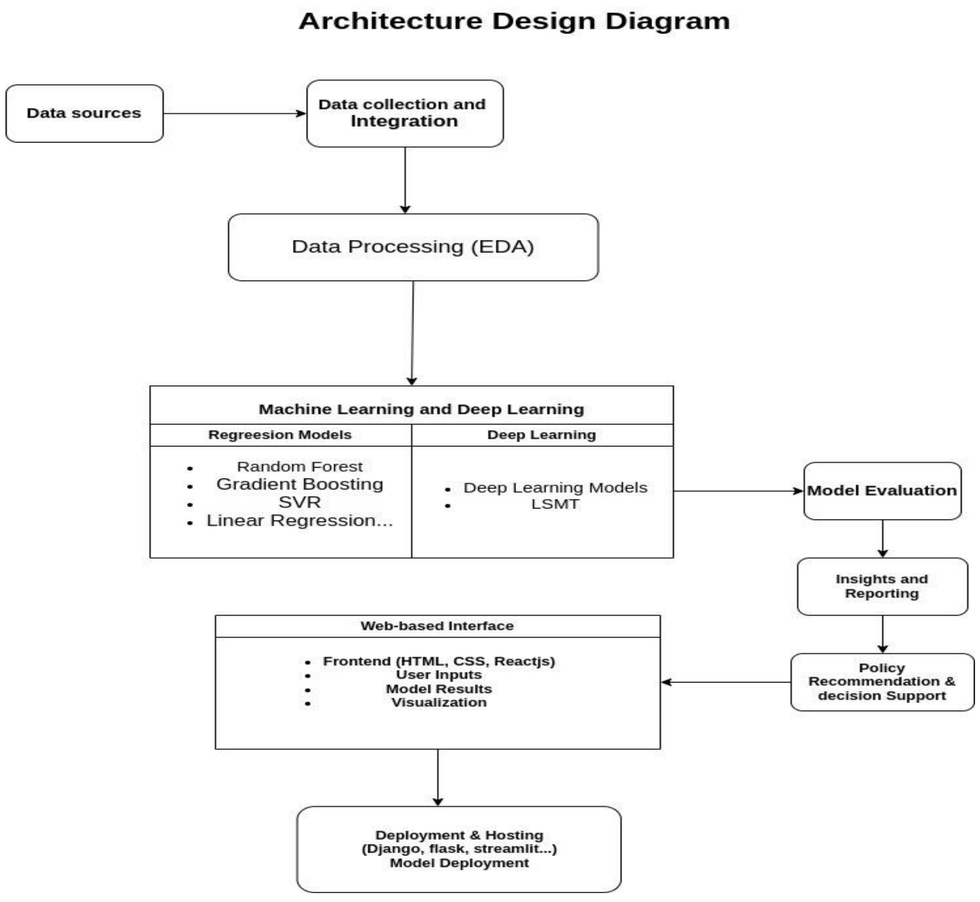
1. **Keras:**
   - **Purpose:** Simplify the implementation and training of deep learning models, specifically LSTMs.
   - **Features:** Keras provides an intuitive API for building and training neural networks, including LSTM networks, making it easier to experiment with and fine-tune complex models.
2. **scikit-learn:**
   - **Purpose:** Facilitate the development and evaluation of traditional machine learning models.
   - **Features:** scikit-learn offers a comprehensive suite of tools for implementing and evaluating regression models like Random Forest and Gradient Boosting. It provides functions for data preprocessing, model training, and performance metrics.

Implementation Overview:

1. **Data Collection and Preparation:**
   - Gather historical climate data, crop yield records, and food security indicators.
   - Preprocess data to handle missing values, normalize features, and split data into training and testing sets.
2. **Model Development:**
   - **For Regression Models:**
     - Implement Random Forest or Gradient Boosting models using scikit-learn.
     - Train the models on historical data to predict future crop yields and food security indicators.
   - **For Deep Learning Models:**
     - Build and train LSTM networks using Keras to analyze time-series climate data.
     - Fine-tune the LSTM models to capture long-term dependencies and improve prediction accuracy.
3. **Model Evaluation:**
   - Evaluate regression models using metrics like Mean Absolute Error (MAE) and R-squared.
   - Assess LSTM model performance using metrics such as Mean Squared Error (MSE) and accuracy on test data.
4. **Integration and Interpretation:**

- Integrate predictions from both regression and deep learning models to generate comprehensive insights.
- Use the results to inform strategies for improving crop yields and food security.

1.5 Architecture Design Diagram



Detailed Component Descriptions:

1. **Data Sources:**
    - **Role:** Collects raw data from various sources, including Climate impact features, agricultural databases, and food security reports.
    - **Functionality:** Provides the foundational data necessary for analysis.
2. **Data Collection and Integration:**
    - **Role:** Aggregates and integrates data from multiple sources.
    - **Functionality:** Merges and cleans data, ensuring consistency and readiness for preprocessing.

3. **Data Preprocessing:**
   - **Role:** Prepares data for analysis.
   - **Functionality:** Includes data cleaning (handling missing values), normalization, and splitting into training/testing sets.
4. **Machine Learning & Deep Learning Models:**
   - **Role:** Develops predictive models to analyze the impact of climate change.
   - **Functionality:**
     - **Regression Models:** Algorithms like Random Forest and Gradient Boosting predict crop yields and food security indicators.
     - **Deep Learning Models:** LSTM networks capture long-term dependencies in time-series climate data.
5. **Model Evaluation:**
   - **Role:** Assesses model performance.
   - **Functionality:** Uses metrics (e.g., MAE, MSE) to evaluate the accuracy and reliability of predictions.
6. **Insights and Reporting:**
   - **Role:** Provides insights based on model outputs.
   - **Functionality:** Generates visualizations and reports to communicate findings.
7. **Policy Recommendations & Decision Support:**
   - **Role:** Supports decision-making and policy development.
   - **Functionality:** Provides actionable recommendations for improving food security and adapting to climate change.
8. **Web-Based Interface:**
   - **Role:** Allows users to interact with the system.
   - **Functionality:** Users can input data, view model results, and access visualizations through a web interface.
     - **Frontend (HTML/CSS/JS):** Design and develop the user interface to facilitate interaction with the system.
9. **Deployment & Hosting:**
   - **Role:** Manages the deployment of the system.

1.6 Data Sources

1. **Climate Data**
    - **Source:** Search for a Dataset - Humanitarian Data Exchange (humdata.org).
    - **Type:** Precipitation, CO2 emissions, Renewable energy consumption, Droughts, floods, extreme temperatures, Agricultural land, Arable land etc
    - **Relevance:** Crucial for understanding climate impacts on agriculture
    - **Preprocessing:** Clean, normalize, aggregate
2. **Agricultural Data**
    - **Source:** World Bank
    - **Type:** Crop yields, planting dates, soil types
    - **Relevance:** Key for predicting crop productivity
    - **Preprocessing:** Clean, engineer features
3. **Food Security Data**
    - **Source:** FAOSTAT
    - **Type:** Food availability, malnutrition rates
    - **Relevance:** Directly related to food security outcomes
    - **Preprocessing:** Clean, normalize, aggregate

1.7 Literature Review (2nd Assignment)

Summary

Existing research underscores the critical need for integrating machine learning (ML) and deep learning (DL) techniques in addressing climate change's impacts on food security. Parry et al. (2004) reveal that climate change generally has a slight to moderate adverse effect on global crop yields, with varying regional impacts. Their work highlights the necessity of adaptation strategies and informs policymakers on mitigating these effects, establishing a foundation for incorporating both socio-economic and climatic factors in agricultural planning. Khan et al. (2024) further this understanding by applying ML and DL models to predict the impacts of temperature extremes on crop yields using CMIP6 data. Their study demonstrates the power of advanced analytics in forecasting future agricultural challenges and provides insights into regional variations in temperature effects. Xiong et al. (2024) utilize ML to evaluate food security across

169 economies, revealing disparities and suggesting policy interventions. Their approach, which improves data quality and completeness through ML, enhances the accuracy of food security assessments and supports effective policymaking. Collectively, these studies validate the application of ML and DL in predicting climate impacts on agriculture and food security, offering a basis for our project's methodology and extending the current research by integrating diverse data sources and advanced analytical techniques.

2 Implementation Plan

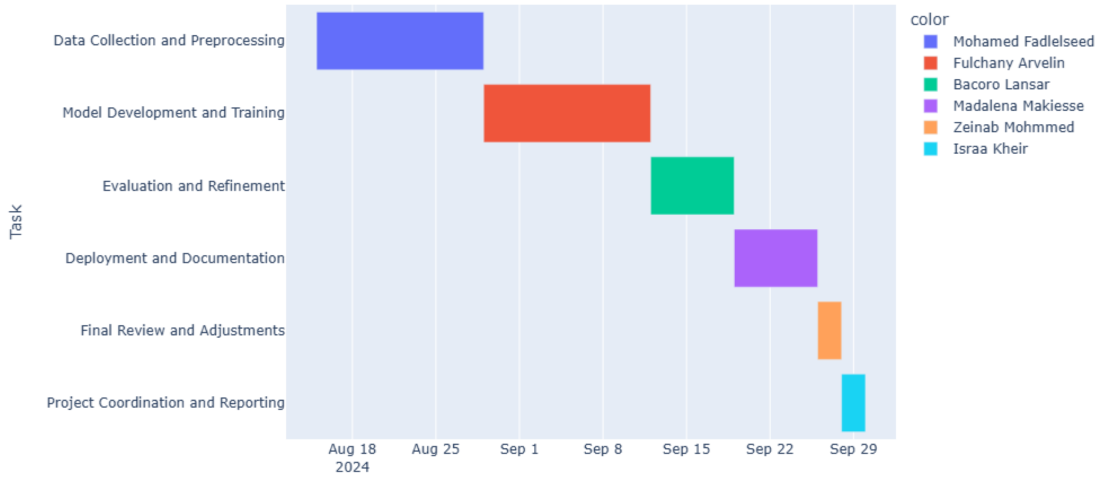
2.1 Technology Stack

- ❖ Programming Languages: Python
- ❖ Libraries: scikit-learn, Keras, TensorFlow, Pandas, NumPy
- ❖ Frameworks: Flask or Django for deployment
- ❖ Tools: Jupyter Notebook for development, GitHub for version control

2.2 Timeline

- **Data Collection and Preprocessing**: 2 week (Aug 15 - Aug 29)
- **Model Development**: 2 weeks (Aug 29 - September 12)
- **Training and Evaluation**: 1 weeks (September 12- September 19)
- **Deployment**: 1 week (September 19 - September 26)
- **Final Review and Adjustments**: 4 days (September 26 – 30 September)

Gantt Chart:



Task Distribution:

- **Mohamed Fadlelseed**: Data Collection and Preprocessing
- **Fulchany Arvelin**: Model Development and Training
- **Bacoro dit Elhafji Lansar**: Evaluation and Refinement
- **Madalena Makiesse**: Deployment and Documentation
- **Zeinab Mohmmed**: Final Review and Adjustments
- **Israa Abdulrahman Mohammed Kheir**: Project Coordination and Reporting

2.3 Milestones

- Data Collection Completion: August 29
- Model Development Completion: September 12
- Training and Evaluation Completion: September 19
- Deployment Completion: September 26
- Final Review and Submission: September 29

2.4 Challenges and Mitigations

- **Data Quality:**
  - **Challenge**: Incomplete, noisy, or inconsistent data can severely impact

the performance of machine learning models, leading to inaccurate predictions and unreliable insights.

- **Mitigation**: To address this, we will implement robust data preprocessing techniques, including data cleaning, normalization, and transformation. Missing values will be handled using imputation methods like mean, median, or mode imputation, or by employing more sophisticated techniques such as k-Nearest Neighbors (k-NN) imputation or iterative imputation. Additionally, outlier detection and removal processes will be conducted to ensure the data used for modeling is clean and representative of real-world scenarios.

- **Model Performance:**

  - **Challenge**: Achieving high model accuracy and generalization to unseen data can be difficult, especially in complex datasets with high variability.

  - **Mitigation**: To enhance model performance, we will utilize cross-validation techniques such as k-fold cross-validation. Additionally, we will perform grid search or randomized search for hyperparameter tuning. Grid search involves exhaustively searching through a specified hyperparameter space to find the optimal combination that maximizes model performance. Alternatively, randomized search samples a subset of hyperparameter combinations, which can be more efficient in large parameter spaces, and can still lead to near-optimal results. By integrating these techniques, we ensure that the model parameters are fine-tuned to achieve the best possible predictive accuracy.

- **Technical Constraints:**
  - **Challenge:** Ensuring the system is scalable, reliable, and compatible across various platforms can be challenging, particularly when dealing with large datasets and complex models.

  - **Mitigation:** To overcome technical constraints, we will regularly test the system components throughout the development process, using unit testing, integration testing, and system testing methods. This approach ensures that each part of the system functions correctly on its own and works seamlessly when integrated.

- **Deployment Constraints:**
  - **Challenge:** Deploying machine learning models effectively while managing resource limitations and infrastructure compatibility.

  - **Mitigation:** Deploy the application using Django or Flask frameworks to create a scalable and manageable web service for the machine learning model. Flask, being lightweight, allows for rapid deployment, while Django provides a robust framework for larger, more complex applications.

2.5 Ethical Considerations

1. **Data Privacy**:

Challenge: Ensuring that the data used in the project respects individual privacy and complies with legal regulations such as GDPR or CCPA. Mitigation: All data will be thoroughly anonymized before processing to remove any personally identifiable information (PII). Access to sensitive data will be restricted, and encryption techniques will be applied to protect data both in transit and at rest. Additionally, data usage will comply with relevant data protection laws and regulations, with consent obtained from data providers where necessary. Regular audits will be conducted to ensure ongoing compliance with privacy standards.

2. **Bias:**

Challenge: Machine learning models are susceptible to biases that can lead to unfair or inaccurate predictions, potentially exacerbating inequalities. Mitigation: Techniques such as fairness-aware algorithms will be employed to detect and mitigate biases in the data and models. The data will be scrutinized for any inherent biases, and efforts will be made to ensure a balanced and representative dataset. During model training, strategies such as re-sampling, re-weighting, and adversarial debiasing will be utilized to reduce bias. Post-modeling, fairness metrics will be evaluated to ensure the model's decisions do not disproportionately affect any particular group.

3. **Impact:**

Challenge: The model’s predictions and recommendations could have unintended socio-economic impacts, particularly on vulnerable populations. Mitigation: A thorough impact assessment will be conducted to evaluate the socio-economic implications of the model's outputs. This will involve consulting with stakeholders, including representatives from affected communities, to understand potential risks and benefits. The model will be designed with a focus on inclusivity, ensuring that its recommendations consider the needs of the most vulnerable populations. Additionally, a feedback mechanism will be implemented, allowing continuous monitoring and adjustment of the model's predictions to avoid negative consequences. Ethical guidelines will be established to govern how the model’s outputs are used in decision-making processes, ensuring they promote social equity and do not harm vulnerable groups.

2.6 References

1. Parry ML, Rosenzweig C, Iglesias A, Livermore M, Fischer G. Effects of climate change on global food production under SRES emissions and socio-economic scenarios. Global Environ Chang. 2004;14(1):53–67.

2. Khan F, Khan F, Liou YA, Khan F, Spöck G, Wang X, et al. Assessing the impacts of temperature extremes on agriculture yield and projecting future extremes using machine learning and deep learning approaches with CMIP6 data. Int J Appl Earth Obs Geoinf. 2024;132(300).

3. Xiong R, Peng H, Chen X, Shuai C. Machine learning-enhanced evaluation of food security across 169 economies. Environ Dev Sustain [Internet]. 2024;(0123456789). Available from: https://doi.org/10.1007/s10668-024-05212-1

4. Tubiello FN. System of Environmental-Economic Accounting for Agriculture, Forestry and Fisheries
- Food Security Indicators [Presentation]. 2016;(June):24–6. Available from: https://seea.un.org/sites/seea.un.org/files/food_security_indicators.pdf