Capstone Project Concept Note and Implementation Plan

Project Title: Forecasting Future Energy Consumption for Households and Businesses

Team Members

1. Ousmane OUEDRAOGO
2. Lucien KALMOGO
3. Yann Boris OUEDRAOGO
4. Edene Vasco BENDANE

1 Concept Note

1.1 Project Overview

This project focuses on developing a real-time energy consumption forecasting system leveraging machine learning techniques. It directly contributes to **Sustainable Development Goal 7 (Affordable and Clean Energy)** by optimizing energy resource utilization and reducing waste.
The project addresses the challenges of fluctuating energy demand, inefficient resource management, and environmental sustainability. The proposed solution aims to provide accurate forecasts that enable energy providers to improve supply-demand management and reduce carbon emissions.

1.2 Objectives

- Develop a robust machine learning model for forecasting short-term and long-term energy consumption.
- Integrate a real-time prediction system using Django for deployment.
- Create interactive dashboards in Power BI to visualize forecasts and empower decision-making.
- Support energy providers in optimizing resources, minimizing costs, and enhancing sustainability practices.

1.3 Background

Energy consumption forecasting is a cornerstone of efficient energy management. Traditional approaches like ARIMA models struggle with nonlinear patterns, while advanced machine learning methods (e.g., XGBoost and LSTM) have shown superior performance in handling large and complex datasets. This project builds on these advancements to address the limitations of existing methods by incorporating real-time data and visualization tools.

1.4 Methodology

- **Machine Learning Models**: XGBoost for short-term forecasts and LSTM for long-term projections.
- **Frameworks**: Python for data processing and modeling; Django for web deployment; Power BI for visualization.
- **Steps**:

    1. Data Collections
    2. Exploratory Data Analysis (EDA) to understand patterns.
    3. Feature engineering to incorporate weather, seasonality, and socioeconomic factors.
    4. Data Splitting
    5. Model Selection
    6. Hyperparameter Tuning
    7. Model Validation
    8. Integration into Application
    9. Interactive Dashboards
    10. Deployment
    11. Monitoring and Maintenance

1.5 Architecture Design Diagram

The system architecture includes:

- **Data Ingestion**: Collect real-time data from energy meters and historical records.
- **Preprocessing**: Data cleaning and transformation pipelines.
- **Modeling**: Training and deploying the forecasting model using machine learning frameworks.
- **Visualization**: Dashboards in Power BI for interactive insights.

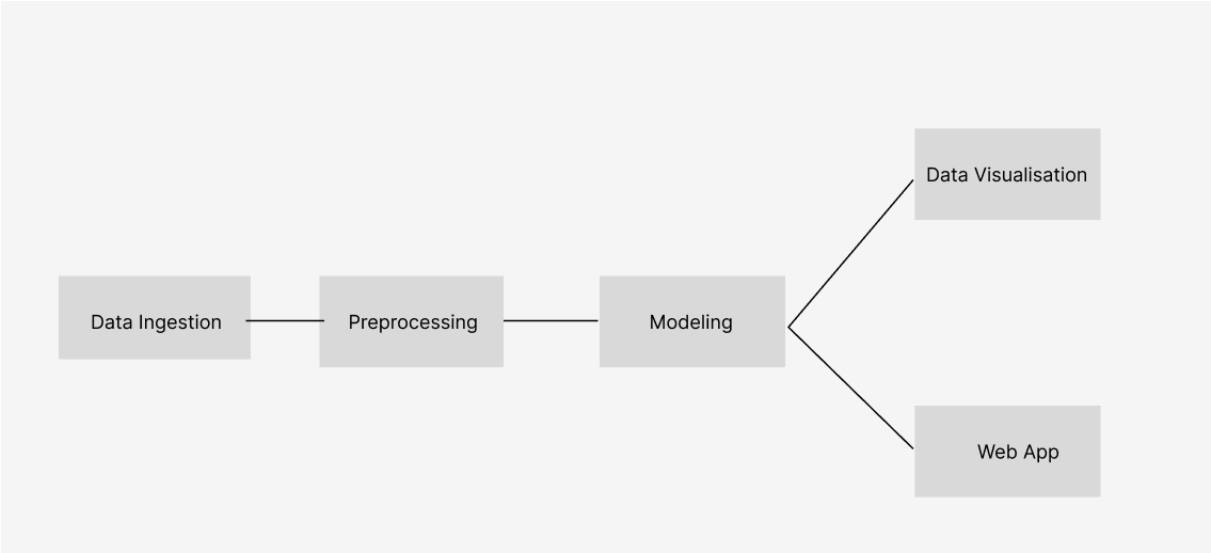

Figure 1 : Architecture Design Diagram

1.6 Data Sources

- **Primary Data Source:** An existing dataset from Kaggle has been chosen for the project.
- **Reasoning:** This decision was made to address challenges related to acquiring IoT tools and the training required for real-time data collection.
- **Data Details :**

- Includes historical energy consumption records.
- Features weather data and other relevant external factors.
- **Preprocessing Steps:** Handling missing values, normalization, and feature extraction to ensure data readiness for modeling.

1.7 Literature Review

Existing studies highlight the efficiency of XGBoost in capturing nonlinear patterns and the applicability of LSTM for sequential data like energy consumption. Hybrid approaches combining statistical and machine learning models improve forecast accuracy. This project builds upon these insights to integrate multiple techniques and real-time capabilities.

2 Implementation Plan

2.1 Technology Stack

- **Programming Language**: Python
- **Libraries**: Pandas, NumPy, Scikit-learn, XGBoost, TensorFlow/Keras
- **Frameworks**: Django for deployment

- **Visualization Tool** : Power BI
- **Other Tools**: Git for version control, Docker for containerization

2.2 Timeline

Task	Assigned Member	Duration	Start Date	End Date
1. Data Collection	Ousmane OUEDRAOGO	2 days	Nov 30, 2024	Dec 1, 2024
2. Exploratory Data Analysis	Lucien KALMOGO	2 days	Dec 2, 2024	Dec 3, 2024
3. Feature Engineering	Yann Boris OUEDRAOGO	2 days	Dec 4, 2024	Dec 5, 2024
4. Data Splitting	Edene Vasco BENDANE	1 day	Dec 6, 2024	Dec 6, 2024
5. Model Selection	Ousmane OUEDRAOGO	2 days	Dec 7, 2024	Dec 8, 2024
6. Hyperparameter Tuning	Lucien KALMOGO	2 days	Dec 9, 2024	Dec 10, 2024
7. Model Validation	Yann Boris OUEDRAOGO	1 day	Dec 11, 2024	Dec 11, 2024
8. Integration into Application	Edene Vasco BENDANE	2 days	Dec 12, 2024	Dec 13, 2024
9. Interactive Dashboards	Ousmane OUEDRAOGO	2 days	Dec 14, 2024	Dec 15, 2024
10. Deployment	Lucien KALMOGO	1 day	Dec 16, 2024	Dec 16, 2024

2.3 Milestones

- Completion of EDA and feature engineering.
- Selection of the best forecasting model based on evaluation metrics.

- Creation of fully functional dashboards in Power BI.

2.4 Challenges and Mitigations

- **Data Quality**: Use robust preprocessing techniques to handle missing or noisy data.
- **Model Performance**: Employ hyperparameter tuning and ensemble methods for better accuracy.
- **Technical Constraints**: Use scalable cloud resources if local computational power is insufficient.

2.5 Ethical Considerations

- Ensure data privacy by anonymizing sensitive information.
- Mitigate bias in models by balancing datasets and testing for fairness.
- Assess the potential societal impact of inaccurate predictions and emphasize transparency in reporting.