Low Level Design (LLD)

Climate Change Analysis and Forecasting

Document Version Control

Date	Version	Description	Author
24/12/2021	1	Initial HLD	Sachin Sarkar
26/12/2021	2	First Update	Sachin Sarkar
27/12/2021	3	Final Update	Sachin Sarkar

Contents

Table of Contents

Abstract 4
Introduction 5
What is Low Level Design Document? 5
Scope 5
General Description 6
Product Perspective : 6
Problem Statement : 6
Proposed Solution : 6
Dataset Used 7
Software and Tools 7
Architecture 9
Data Collection 9
Exploratory Data Analysis : 9
Data Modelling 10
Data Validation 10
Web Application Development 11
Deployment 12
Conclusion 13

Abstract

Climate change is undoubtedly one of the biggest problems in the 21st century. Artificial Intelligence methods have recently contributed in the advancement of accurate prediction tools for the estimation and assessment of extreme environmental events and investigation of the climate change time series. The recent advancement in Artificial Intelligence including the novel machine learning and deep learning algorithms as well as soft computing applications have greatly empowered prediction methods. Through this project, we have explore, analyze the global Climatic trend and pattern on temperature component and forecast the future temperature trends using a state of art time series deep learning model. After the research, exploration and analysis on the historical data and modelling, we build and deploy an end to end web solution on the frontend to view and explore historical data as well as future forecasts generated through the deep learning model.

Introduction

What is Low Level Design Document?

The goal of the Low-level design document (LLDD) is to give the internal logic design of the actual program code for the Heart Disease Diagnostic Analysis dashboard. LLDD describes the class diagrams with the methods and relations between classes and programs specs. It describes the modules so that the programmer can directly code the program from the document.

Scope

Low-level design (LLD) is a component-level design process that follows a step-by-step refinement process. The process can be used for designing data structures, required software architecture, source code and ultimately, performance algorithms. Overall, the data organization may be defined during requirement analysis and then refined during data design work.

General Description

Product Perspective :

This Climate Change Analysis and Forecasting Application is a Web Application powered by Streamlit, Plotly and a Deep Learning based Time Series Forecasting Model to analyze, visualize and forecast average temperatures per month.

Problem Statement :

Will analyze the change in temperatures across globe from the 17th century till now and build a multivariate deep learning-based time series model to forecast the U.S. Average temperature. Predictive models attempt at forecasting future value based on historical data. The main objective here is -

1. Analyse the changes in climate across the globe.
2. Alert if any unusual climate change happen.
3. Maintain a database to store each and every data.

Proposed Solution :

The solution here proposed is a Multi Paged Web Application through which user can input or set Country, State, Time or Year Range and based on that data (may historical or future) of average temperatures per month, statistical description like mean, median, max, min, count etc. This data also downloadable on a button click. An interactive Trend Plot to view trend based on moving average , an interactive seasonal bar plot to view monthly average temperatures and an interactive autocorrelation plot to show correlation upto 100 previous data. Except this all user also can send a feedback or message to development team.

Dataset Used :

	A	B	C	D	E	F	G	H	I
1	dt	LandAvera	LandAvera	LandMaxT	LandMaxT	LandMinTe	LandMinTe	LandAndO	LandAndO
2	1850-01-0	0.749	1.105	8.242	1.738	-3.206	2.822	12.833	0.367
3	1850-02-0	3.071	1.275	9.97	3.007	-2.291	1.623	13.588	0.414
4	1850-03-0	4.954	0.955	10.347	2.401	-1.905	1.41	14.043	0.341
5	1850-04-0	7.217	0.665	12.934	1.004	1.018	1.329	14.667	0.267
6	1850-05-0	10.004	0.617	15.655	2.406	3.811	1.347	15.507	0.249
7	1850-06-0	13.15	0.614	18.946	2.817	7.106	0.857	16.353	0.245
8	1850-07-0	14.492	0.614	19.233	2.84	8.014	0.786	16.783	0.238
9	1850-08-0	14.039	0.802	18.477	2.079	7.406	1.086	16.718	0.28
10	1850-09-0	11.505	0.675	15.846	2.692	4.533	1.798	15.886	0.254

For this project we used 3 csv file and 1 json. This are –

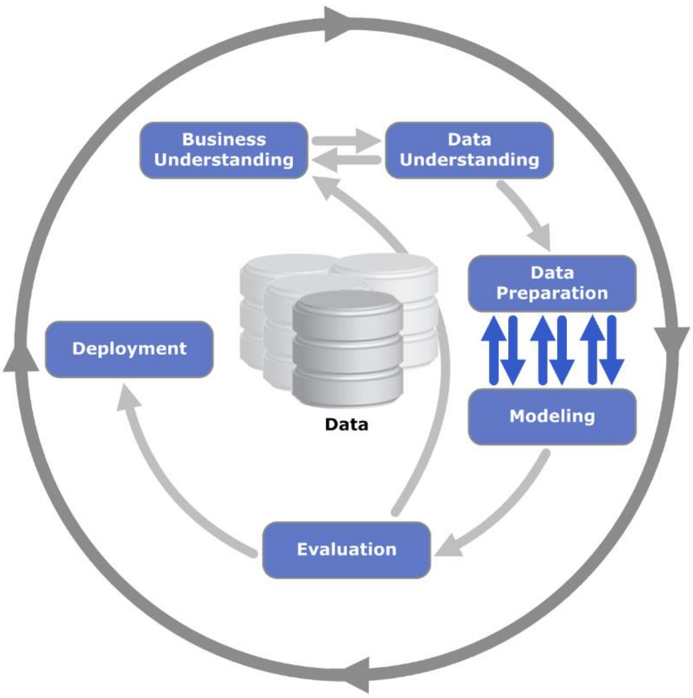
1) **GlobalTemperatures.csv :** For global average land temperatures.
2) **GlobalLandTemperaturesByCountry.csv :** For average temperatures group by countries.
3) **GlobalLandTemperaturesByState.csv :** For average temperatures group by states.
4) **Countries.geo.json :** Fo geographical locations of countries.

Software and Tools Used :

Here is the list of software and tools used in this project implementation.

- Python : used as the primary programming and scripting language.
- Jupyter Notebook : For python scripting, data analysis and research.
- Spyder : For application backend Programming.
- Pandas : used for data frame manipulation.
- NumPy : For array manipulation.
- Plotly Express : for interactive plots.
- Neural Prophet : For Deep Learning modelling and forecasting.
- Streamlit : For frontend development.
- Html / CSS : For frontend improvement.
- Git : For project version control.

Architecture



Data Collection :

Data used in this project already discussed before. This data was collected from http://berkeleyearth.org/data/

Exploratory Data Analysis :

Analyze the change in temperatures across globe from the 17th century till now.

This Exploratory Data Analysis contains,

1) Load and Show Dataset
2) Missing Values Imputation
3) Lag Plot and Analysis
4) ACF or Autocorrelation plots and Analysis
5) Trend Plot and Analysis
6) Seasonal Plot and Analysis
7) Chloropleth Map of Average Temperature by Countries and Analysis

Data Modelling :

Select, train and Validate a best Time Series Model to forecast Average Temperatures per month for next years.

This part contains,

1) Import All Required Libraries
2) Load and Show Dataset
3) Missing Values Imputation
4) Data Stationarity Check
5) Time Series Modeling : Neural Prophet
7) Preparing dataset for NeuralProphet Training
8) Spliting Dataset for training and validation
9) Model Training
10) Model Training History and Visualization
11) Model performence analysis on Validation Data
12) Demo forecasting for next 2 years

Data Stationarity Check

Strategy : Augmented Dickey Fuller test (ADF Test)

Time Series Modeling : Neural Prophet

What is Neural Prophet ?

NeuralProphet is an upgraded version of Prophet that is built using PyTorch and uses deep learning models such as AR-Net for time-series forecasting. The main benefit of using NeuralProphet is that it features a simple API inspired by Prophet, but gives you access to more sophisticated deep learning models for time-series forecasting.

Data Validation :

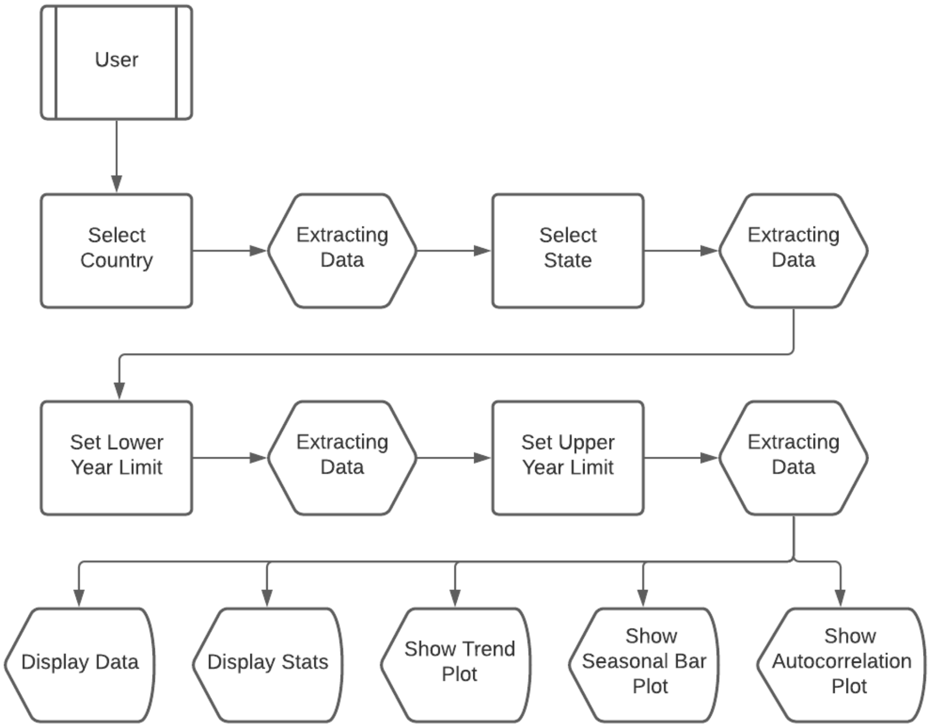
Testing model performance on validation data. Validation metrices used – RMSE and R2 Score.

Web Application Development :

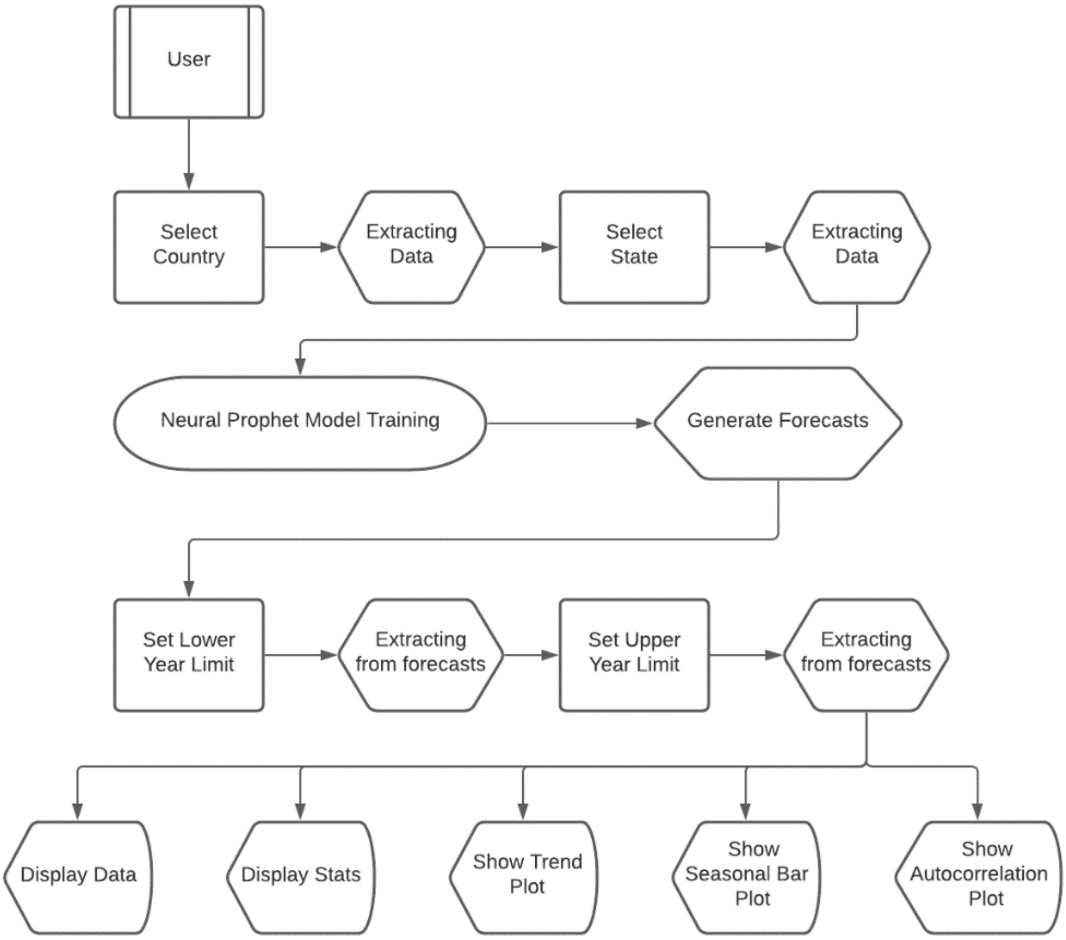
As per the proposed solution is a Multi Paged Web Application. To develop this frontend application, Streamlit is used powered by HTML / CSS. The whole coding for this frontend development is done on Spyder IDE.

Here 5 different web pages developed inside the application. This are –

- Documentation
- Historical Data and Plotting
- Future Data and Plotting
- Feedback Us
- About Us



1Architecture for Historical Data and Plotting

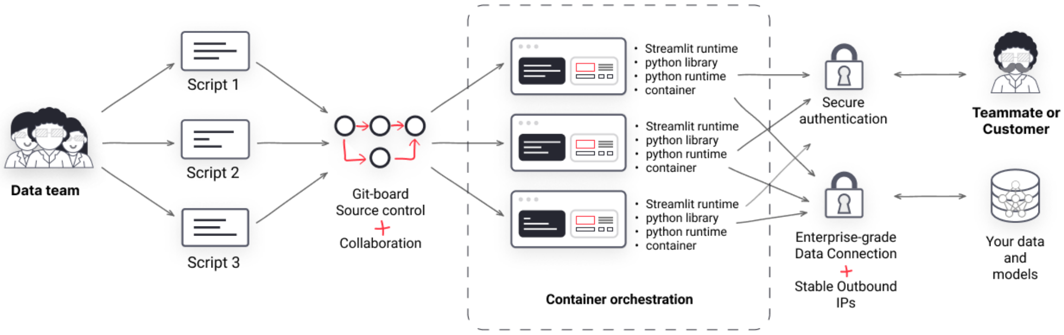


2 Architecture for Historical Data and Plotting

Deployment :

Streamlit Cloud used for deployment this application.

When you work on an app in Streamlit Cloud—be it a new model, data analysis, or idea—you're just a few clicks away from securely sharing it and collaborating on it with your team.



Conclusion

Undoubtedly, this application will help users to analyze past and future climatic trends and changes. As well as users can get to know the monthly seasonality and autocorrelation. Users can able to filter this insights as per there choice of country, state and time range.