Agro-Advisory Services

A Project Report

Submitted by:

Vatsalkumar Patel (AU2040043)

in partial fulfillment for the award of the degree

of

BACHELOR OF TECHNOLOGY

in

(COMPUTER SCIENCE AND ENGINEERING)

at

Ahmedabad University

School of Engineering and Applied Science (SEAS)
Ahmedabad, Gujarat

May, 2024

DECLARATION

I hereby declare that the project entitled "**Agro-Advisory Services**" submitted for the B. Tech. (**Computer Science and Engineering**) degree is my original work and the project has not formed the basis for the award of any other degree, diploma, fellowship or any other similar titles.

Signature of Student

Date:

Place:

CERTIFICATE

This is to certify that the project titled "**Agro-Advisory Services**" is the bona fide work carried out by **Vatsalkumar Patel**, a student of B. Tech. (**Computer Science and Engineering**) of School of Engineering and Applied Science at Ahmedabad University during the academic year 2023-2024, in partial fulfillment of the requirements for the award of the degree of Bachelor of Technology in **Computer Science and Engineering** and that the project has not formed the basis for the award previously of any other degree, diploma, fellowship or any other similar title.

This project was done under the supervision of the faculty mentor **Professor Sanjay Chaudhary**.

Signature of Faculty Mentor

Date:

Place:

Abstract

This report is a documentation of the project created in the span of 4 months as part of BTEP. The initial plan was to develop a service to facilitate urban planning, but after thorough discussion with Prof. Sanjay sir, it was decided that the project will focus on helping farmers make informed decisions in the agricultural domain. This is because we realised technology is not being used up to par in the field of agriculture. Only facilities provided are some websites that are managed by the government. Hence I have worked on developing a platform that hosts GIS layers which contain valuable information for farmers. Further a machine learning model is integrated for personalised suggestions as well. I believe farmers have also become smart nowadays, and they are capable of using such facilities themselves. If not, the government can guide them in using these tools. Keeping this is mind Agro-Advisory Service was created.

Table of Contents

List of Figures

Gantt Chart

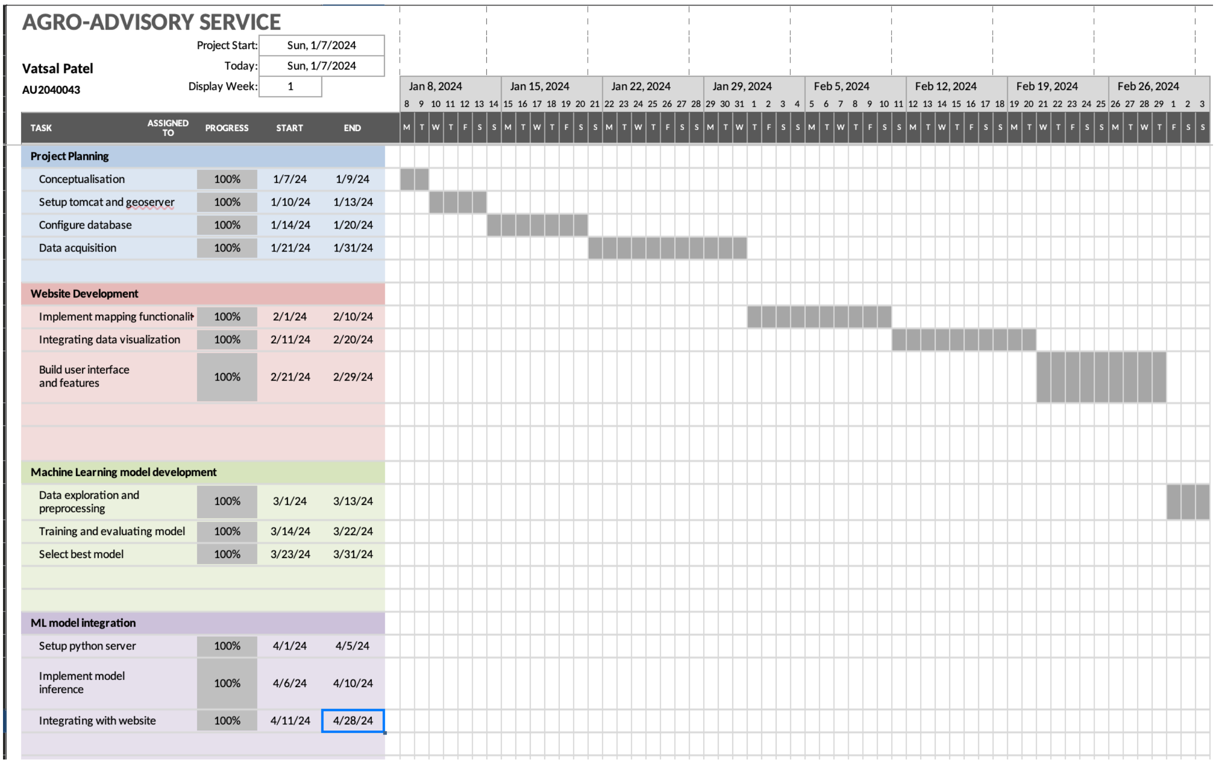

Figure 0.0.1: Gantt chart

Chapter 1

Introduction

While several government-managed websites were designed to serve agricultural needs, there was a deficiency of competent platforms created with farmers' requirements in mind. It was this blank of information that gave rise to the idea of Agro-Advisory Service, which was designed to address the problem of farmers obtaining valuable information and getting personalized advice by means of GIS layers and machine learning models integration.

The primary functionality is based on layers of GIS that host important data for farmers, which refers to agriculture related infrastructure, soil composition, etc. Furthermore, the integration of machine learning model in the system to provide personalised suggestions make the platform more attractive to use.

The project was built with the fact in mind that modern day farmers are smart enough to use such facilities on their own. If otherwise, the government can take initiative to help them out by providing guidance in using such facilities.

1.1 | Project Definition

The goal is to build a platform that can render spatial data related to agriculture. Additionally, features will be added to website to help farmers navigate and use the available data. Other than that, farmers can specify their particular use case to get personalised recommendation.

1.2 | Project Objectives

1. Create backend to host base map to feed agriculture related data.
2. Find data from trusted sources such as government websites to load on to the website
3. Add features to website to make data accessible.
4. Find and train the best suited ML model for prediction.
5. Integrate ML model into website to create recommendation system.

Chapter 2

Literature Survey

As computer science is evolving rapidly, there are major advances in terms of software availability is every domain. In agriculture field most of the facilities made available are provided and maintained by the Indian government.

2.1 | Related Work

Some websites made available by the government include the following:

1. **Bhuvan - Indian Geo Platform of ISRO:** Bhuvan provides maps of India along with geospatial data services. It can be used for urban planning, infrastructure development, disaster monitoring, etc.

2. **AGMarknet:** On AGMarknet website, one can check the prices of different commodities in different mandis. These prices are updated on a daily basis. It can help farmers make an informed decision for the price of their own crop.

3. There are other websites which only provided data collected by the government such as ICRISAT, agriexchange, etc. The data available on these websites can be put to good use if utilised effectively.

2.2 | Tools and Technologies

1. **Frontend:** Bootstrap and JQuery-UI for building the user-interface.

   [a] **Bootstrap** is a library of HTML, CSS and JS which assists in the development of website pages. It provides ready to use HTML elements which result in a uniform layout across the platform.

   [b] **JQuery-UI** is a library consisting of GUI widgets and visual effects and themes implemented using the jQuery javascript library.

2. **Backend:** Apache tomcat server for hosting the platform and geoserver for managing spatial data.

   [a] **Apache Tomcat Server** is a java servlet which is used to make HTTP request and response. It is a scalable servlet option.

   [b] **Geoserver** is used to manage spatial data layers efficiently.

3. **Database:** PostgreSQL is the database used to store shapefiles. PostGIS extension enables storage of spatial data.

    [a] **PostgreSQL** is a relational database management system which is open-sourced.

    [b] **PostGIS** is an external extension for postgres to provide support for geographic objects. It follows the specification of the Open Geospatial Consortium.

4. **APIs:** Openlayers API for managing maps on the platform.

    [a] **Openlayers** provides javascript libraries to render maps on websites.

5. **Languages:** HTML, Javascript and Python.

6. **Machine Learning:** Light Gradient Boost Machine model.

    [a] **LGBM** is a scalable and fast machine learning algorithm which works on the logic of decision trees, histogram bins and gradient based one side sampling for speed and scalability.

Chapter 3

Methodology

3.0.1 | Learn tech stack

The first phase of the project was to learn how to create website that can render maps online. I finalised the tech stack to use. It includes Apache Tomcat server, Geoserver, Openlayers mapping API, and PostgreSQL database with PostGIS extension.

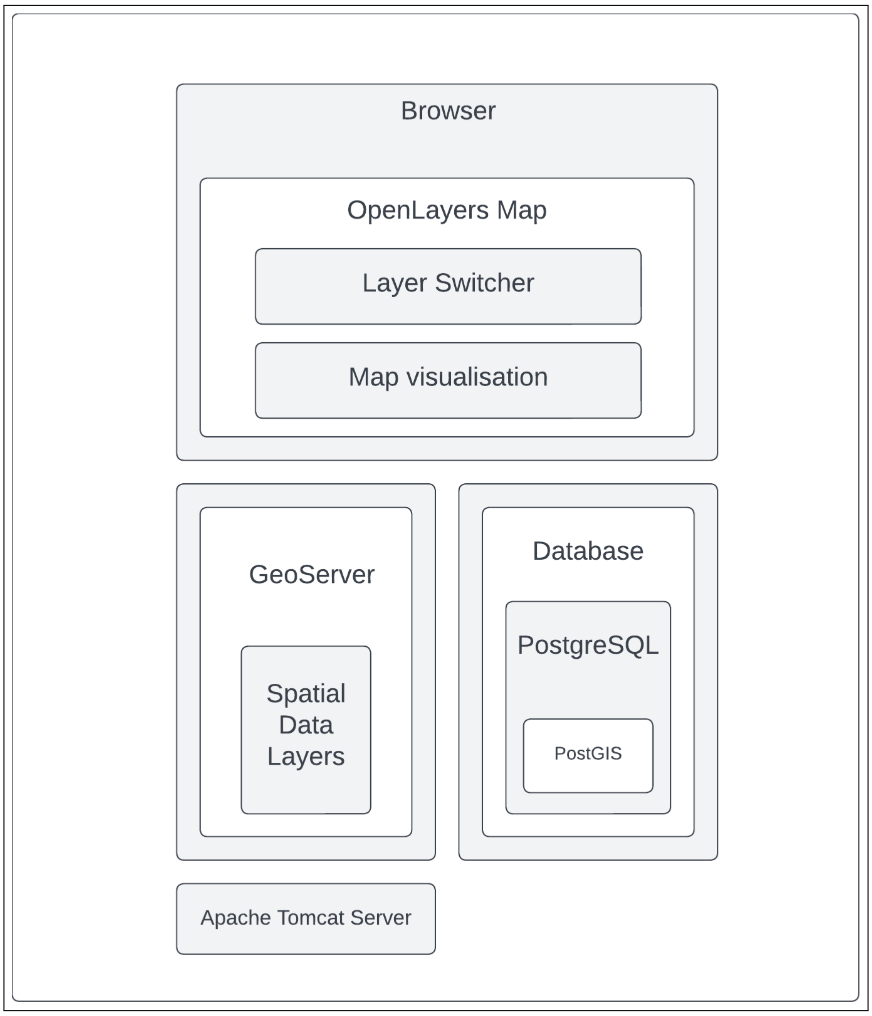

Figure 3.0.1: Website architecture

3.0.2 | Server Setup

Then I set up Apache Tomcat and Geoserver on my localhost. For that I installed the required files for both of them. I also installed other necessary dependencies and wrote the required code to get them up and running. I also installed other libraries such as jQuery, LayerSwitcher and Bootstrap for future use in implementing functionalities.

3.0.3 | Base Map setup

Then to implement the maps, I fetched base layer from Open Street Maps provided by OpenLayers. The coordinate reference used across the project for uniformity is EPSG:4326, also known as WGS84. It is the same used by other GPS systems like google maps. It uses latitude and longitude coordinates. To check basic functionality I added shapefiles for India state and India district to geoserver. They are one of the most common file formats for geospatial data. Geoserver is running on top of Apache tomcat server. Even though geoserver can run an a standalone application, it is deployed on top of tomcat for security, integration with other web apps and scalability. Geoserver uses WMS and WFS standards provided by OGC.

3.0.4 | Database

Next step is to setup database for storing shapefiles. So I setup the server for PostgreSQL. Now to make it capable of handling geospatial data, I installed the PostGIS extension. I connected the database to geoserver. Now it is possible to fetch any datafiles available on the database onto geoserver. The main use of database is only to store the data in form of shapefile. So no complex schema is required.

3.0.5 | Feature Addition

Now that database is ready, the task at hand is to find data to add features to the website. The important thing is that the data should be taken from a reliable source such as government provided datasets. I started looking for data related to agricultural infrastructure such as cold storage warehouses or APMC locations. Soon I realised this data is not easily available to be readily used. Even searching on google maps did not help much. Finally I found a website of goverment of gujarat which had pdf files containing address and contact details of both APMC and cold storage warehouses. Now from these pdf files I mapped the addresses to the latitude/longitude by running a python script. I created a csv for the same, opened it in QGIS and exported as shapefile. Then I used the shp2pgsql command to upload the shapefile to my database. Then on geoserver, I modified the design of these GIS layers by modifying their respective SLD files. Now these

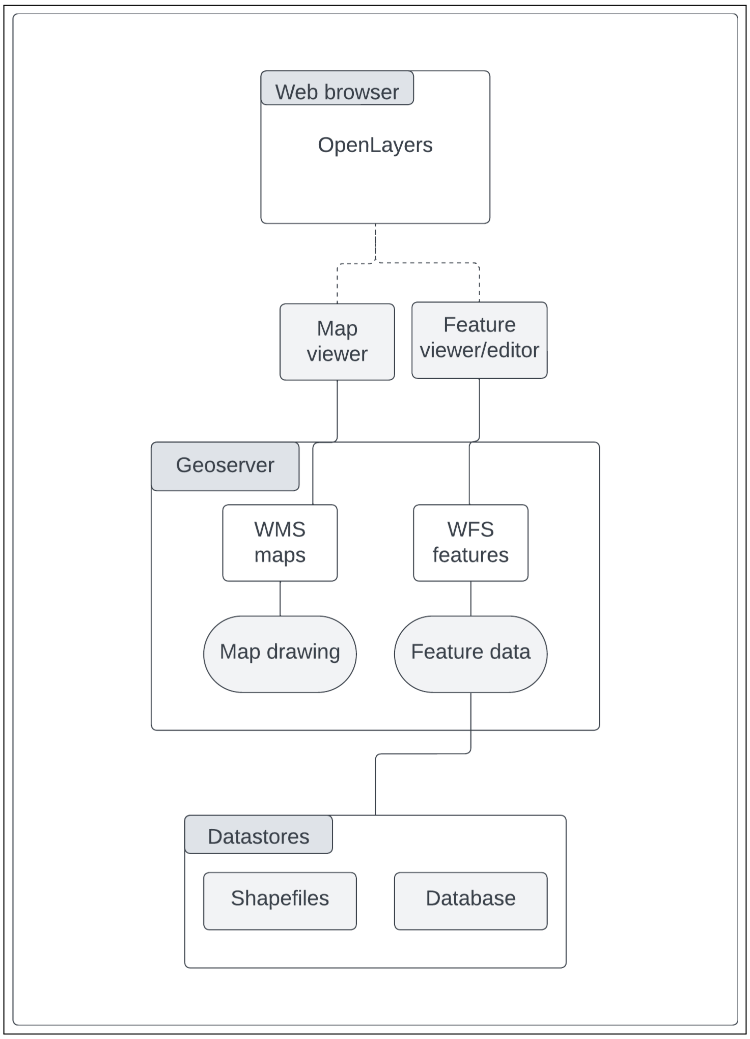

Figure 3.0.2: Geoserver architecture

data-points can be viewed on the frontend, in form of red points for cold storage and green points for APMC location. In the background can be seen the gujarat state provided by OSM base map. They both are aligned properly because of same datum WGS84.

Now just showing these points on the website is meaningless as they don't hold any value. The data that those points hold such as address, contact details, storage type in case of cold storages, etc are what are helpful for someone using the site. So I implemented the GetFeatureInfo functionality. When this is active, clicking on the data point will pop-up all information for that point.

3.0.6 | Recommendation System

Now is a major task, the recommendation system. The plan is to develop a crop recommendation system using machine learning on soil nutrients and weather data. Usually farmers have enough knowledge about which crop is best to grow, but because of worsening weather conditions, changes in soil because of chemicals, problems are being faced by farmers. Machine learning can solve this problem in the best manner. First step is data. The soil data I am using is provided by Indian Chamber of Food and Agriculture. It has soil data with features of nitrogen, phosphorous and potassium content (NPK fertiliser). Other than that it also has features of temperature, humidity, pH and rainfall. It has total 21 crops including both agriculture and horticulture crops. Because of this diverse geography consideration and various crops, this dataset is great for model training.

3.0.7 | Model training

First of all, I performed exploratory data analysis, data pre-processing and feature scaling. After that I tried out multiple different models namely, k nearest neighbours, decision tree, xgboost, SVM and random forest.

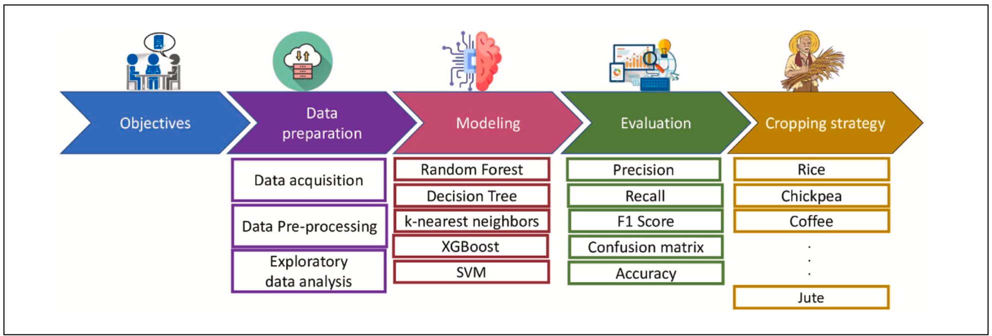

Figure 3.0.3: Procedure for creating recommendation model

As per the results, it is clear that gradient boosting is the best approach for the particular use case. LightGBM is based on decision trees that grow leaf wise based on the gain. It uses a histogram approach where data is divided into bins, and bins are iterated to calculate value of loss function and accordingly split the data. For sampling in lightGBM, GOSS (gradient based one side sampling) is used. Here the data points with higher

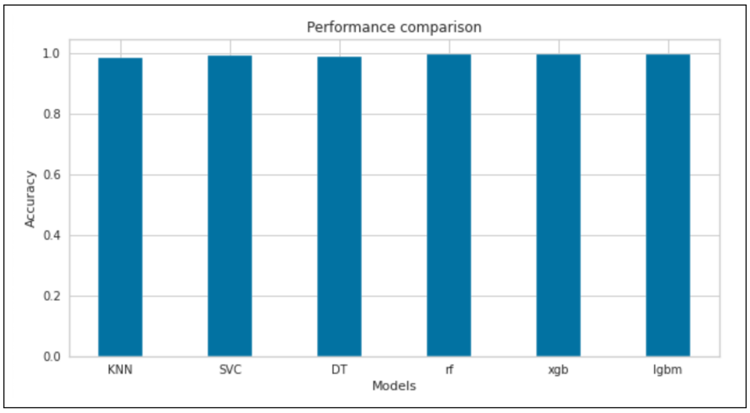

Figure 3.0.4: Accuracy comparison of models

gradient are given more weight while calculating gain. Hence I decided to move forward with using light gradient boosting model for my project.

3.0.8 | Model integration

Now the next checkpoint is to integrate this model into the website. For that I first set up a flask server to handle HTTP requests and responses. When you enter the NPK fertiliser data and weather data on the website and submit the details, the predict route is called. Inside the predict function, all the data is passed to the pickle file of the ML model. A pickle file is a pre-trained version of an ML model. The model gives output in terms of numbers which are re-mapped to crop names. The output is then sent to the website using render_template.
With this all main functionalities of the website are complete.

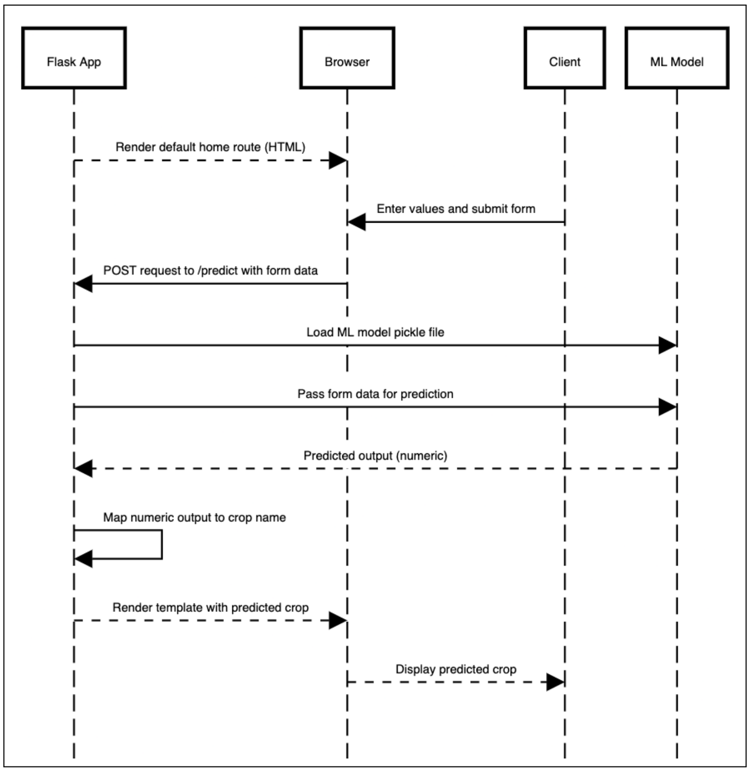

Figure 3.0.5: ML integration - Sequence Diagram of calls made

Chapter 4

Results

The goal of the project was to create a platform for people in the agriculture field to access vital knowledge and information to support their decision making. The result is a web application that brings together geospatial data, machine learning brain and user-friendly features to provide farmers with useful information.

4.1 | Project Outcomes

Figure 4.1.1 refers to the look of the landing page of the website. It opens with India at the center. Clickable data points for cold storage and apmc location can be seen in gujarat. On clicking the cold storage point following data is visible:

1. Name
2. Address
3. Capacity
4. Sector
5. Products

Following information is available for apmc:

1. Name
2. Address
3. Manager
4. Contact Number
5. Email

Polygon layer for soil data is available with NPK fertiliser information and weather condition data.
Figure 4.1.2 is how the information popup appears when clicked on a point. Information from all layers below the point that is clicked is visible in the popup.
Figure 4.1.3 is the form that opens when clicked on the Suggest button in the top right corner of the website. Fill the soil information and press submit to get crop recommendation in the yellow box.

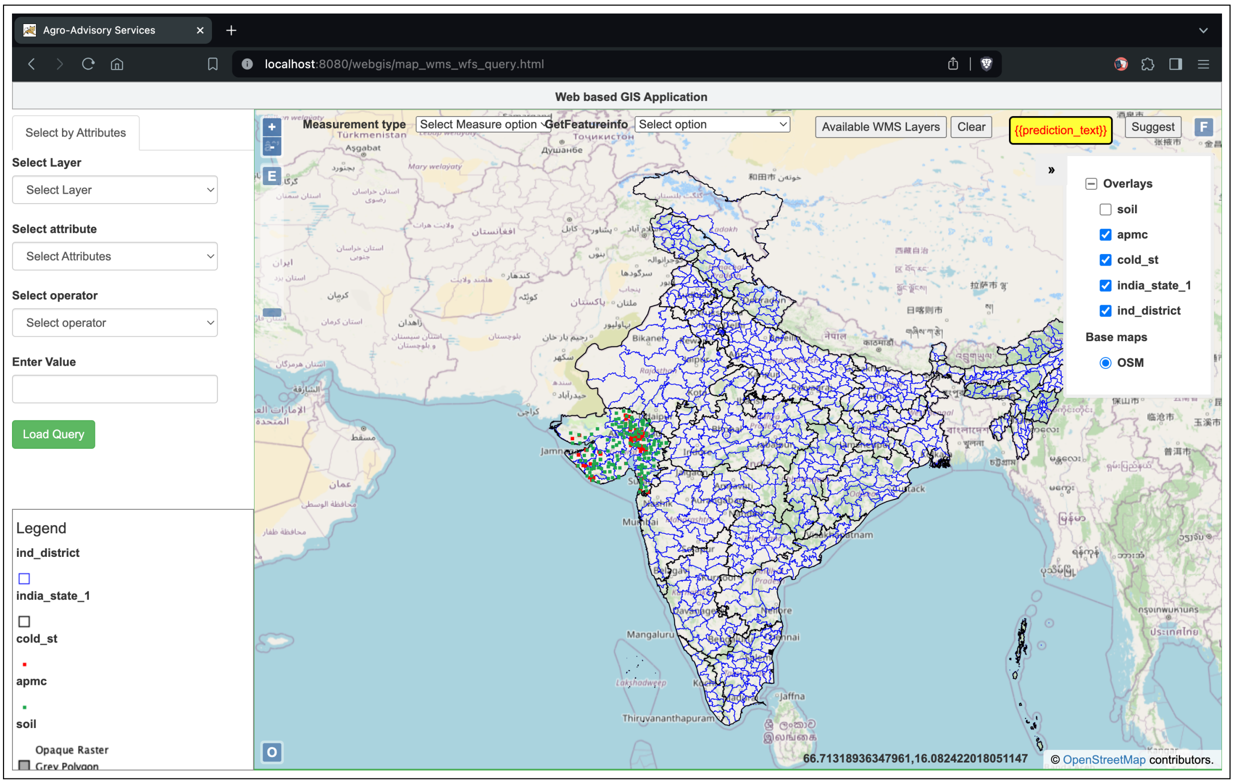

Figure 4.1.1: Landing page of website

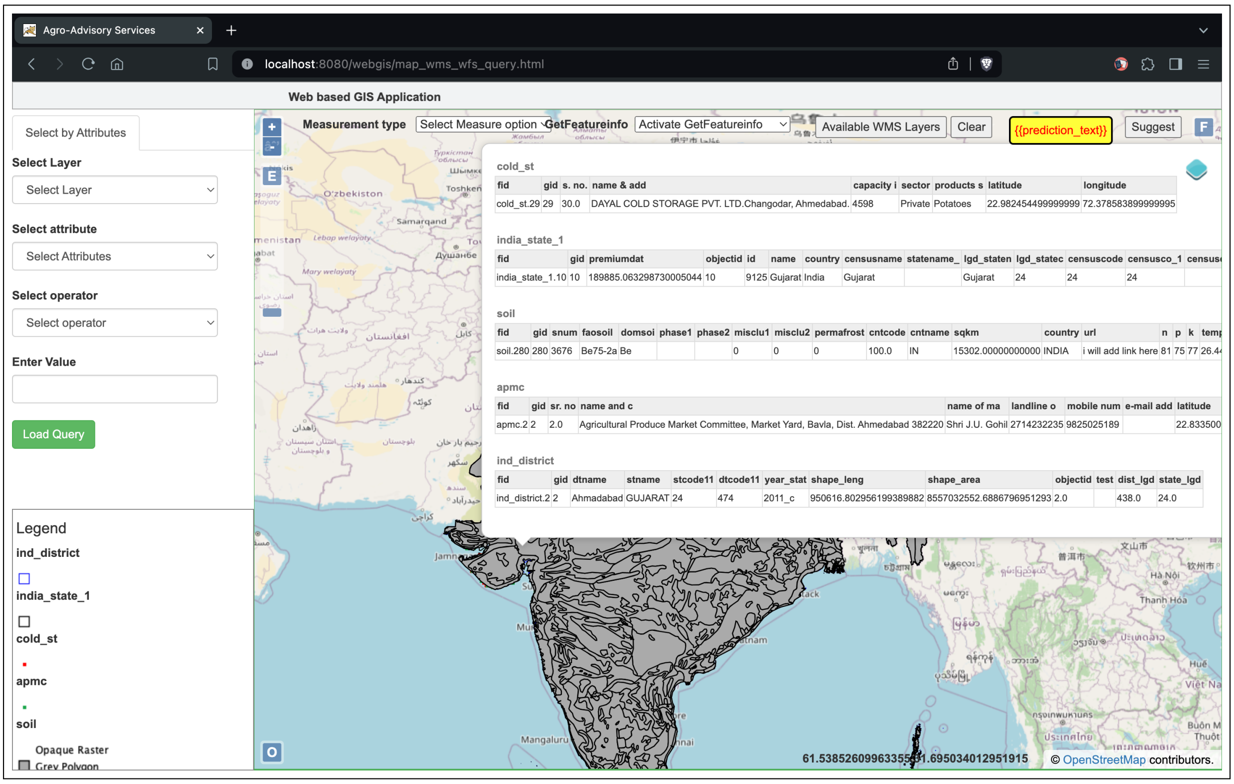

Figure 4.1.2: Information popup on click example

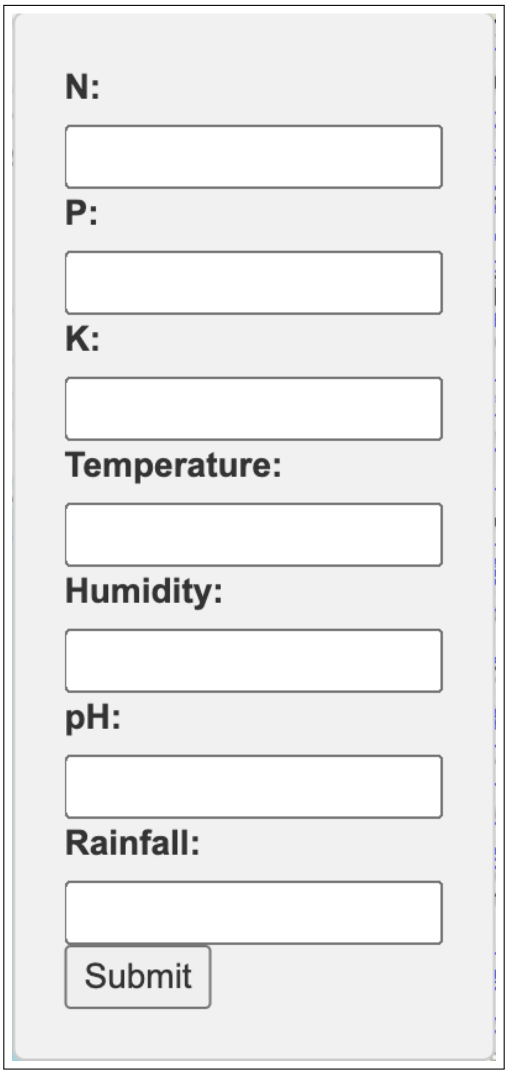

Figure 4.1.3: ML input parameters form

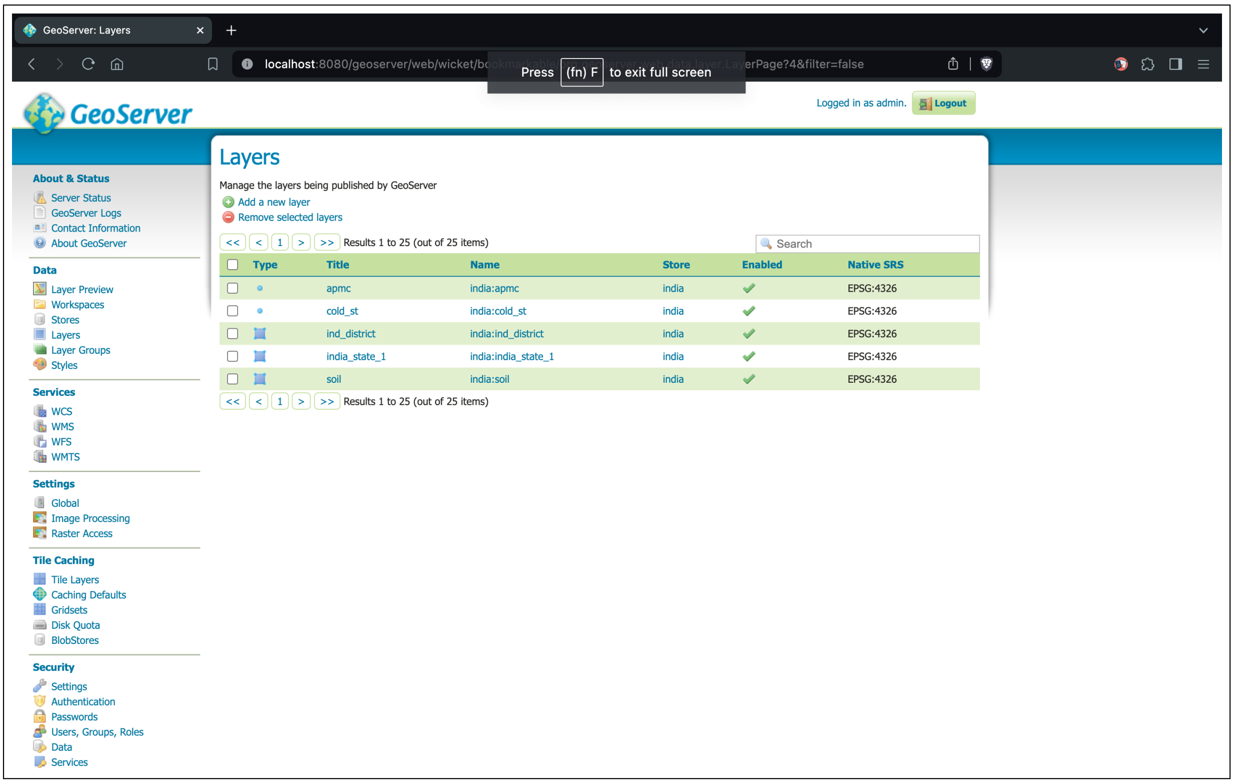

Figure 4.1.4: Shapefile layers in geoserver

Figure 4.1.4 is the retreival of data from database onto geoserver in form of layers. Here the bounds for the layers are computed so that they align according to same reference cordinates. Other than that styling of the layers can be done here as well.

4.2 | Learning Outcomes

In the process of working on this project, I have gained technical skills, problem-solving abilities and critical thinking abilities. The learning outcomes cover a wide scope including geospatial data handling, mapping technology, machine learning and web development.

I was able to use my existing knowledge from the machine learning course while creating the recommendation system. I used my knowledge about QGIS which I learnt from the data science course. I had learnt about software development in the software engineering course which was utilised while making the website.

1. **GIS:** Learnt about vector data formats, working with shapefiles. Learnt about data scraping while collecting relevant geospatial data from various sources. Learnt about OGC standards, WGS84 datum which are required for uniformity of data especially when data comes from various sources.
2. **Openlayers:** Learnt about the mapping library and how various map services that we use daily, work.
3. **QGIS:** Learnt using QGIS, in which I learnt about attribute tables, spatial data managment, designing of layers.

4. **Machine Learning:** Learnt about light gradient boosting model. Learnt how it works and is effiicient and scalable. Learnt how to integrate ML models into a website. Learnt about flask app in python.

4.3 | Real World Applications

This project has numerous real-world applications that can benefit the agricultural community and contribute to advancement of sustainable farming practices.

1. **Precision Agriculture:** The platform can be used to tailor farming practices to specific field conditions. By making use of geospatial data coupled with machine learning, farmers can take educated decisions about crop selection, nutrient management and irrigation strategies.

2. **Government Assistance to farmers:** Some farmers might be unable to use the website on their own. In this case government can assign tasks to people to help farmers by visiting each one of them and using the platform and giving them all the insights. In case someone wants to test soil nutrients separately and then use the ML model independently, it is also possible.

3. **Farm planning and management:** Based on the knowledge of agricultural infrastructure availability, farmers can identify suitable areas for their farms, analyze soil properties, weather conditions, proximity of markets or storage facilities, enabling them to make better decisions.

Overall the project has the potential to help agricultural community in various ways.

Chapter 5

Conclusion and Future Work

At the end of the project, we have got a feasible as well as interactive web-based platform where not only we can use the benefit of geospatial data, AI and user-friendly interfaces but can also assist farmers in making better decisions as well as improve their productivity.

The platform's integration of top-notch technologies, including GIS layers, OpenLayers mapping API, PostgreSQL with PostGIS, and Apache Tomcat and GeoServer establish a reliable and scalable foundation for the presentation and assessment of geometric data such as soil, weather, crop geometry, and market locations.

The notable feature is the Light Gradient Boosting machine learning model as it gives precise crop recommendations as well as crafts each recommendation to the very conditions the soil and weather of the place. This way the farmers will be empowered to improve their crop choices and resource utilization for the better of their crops.

With the use of simple interface, visual interactive maps, data visualization and customization to suit needs of farmers one can expect accessibility and convenience. The exhaustive testing and commensal optimization of the platform resulted in increased performance and quick response, which became a source of positive feedback from users.

This platform has the potential to be used for precision agriculture, under the soil analysis, farm planning, environmental monitoring, extension services, research, and education. This whole in turn leads to the promotion of the sustainable practices, increase in productivity, and the development of the sector.

Depending on expansions and correspondences with cutting-edge technologies, the project is likely to be merely a step towards leveraging technology to address the agricultural problems, generating new approaches and methodologies for decision, management and sustainability.

At the present stage, the existing source for informed decision-making of the farmers via the provision of geospatial data, machine learning, and usability factors is available. However, it is clear that there are several areas that need improvement or additional elements that need to be added to make it even more useful for farmers.

1. **Multilingual Support:** To include a wider audience in the platform we could build

in a native support for other languages such for instance Hindi. This will make the platform to cater for the individuals who are not expert in English to communicate with the platform with ease. This improves service delivery, and in turn, will lead to the wide use of the platform.

2. **Voice Commands:** The platform could provide commands in different languages to cater for farmers who have different levels of literacy. This part would enable farmers with reading and writing challenges to navigate through the platform and access all the features by using voice commands so that the platform becomes more inclusive and simple to use for everyone on it.

3. **Warehouse Integration:** Continuing the supply chain integration with warehouses and storage departments, farmers can be given (i.e) actual time data on storage capacity and pricing. Through the applications, the growers would have an access to making informed decisions on where to place their crops in a manner that would optimize logistics and efficiently use resources.

4. **APMC Mandi Price Integration:** One of the effective strategies can be integrating actual data of APMC mandi prices into the application for which time saved will be used by farmers in making the right decisions they need to take. Such knowledge would then help them in formulation of timely and right sales points for their products, which in turn might translate to increased sales revenue.

5. **Soil Deficiency Recommendation System:** Expand the platform by introducing a capacity to generate recommendations for soil deficiency will extend more value to it. The system can be facilitated by the analysis of soil data and identification of nutrient deficiencies which will further recommend appropriate remediation strategies like fertilizer applications or soil amendments meant to encourage microbial activities and boosting crop growth.

Continued development and integration of up and coming technology would be paramount in maintaining relevance of the platform against the ever changing needs of the agricultural community. Domain specialists, farmers, and research institutions' engagement would be crucial in terms of identifying other improvement areas as well.

Bibliography

1. Cold storages in Gujarat. (n.d.). https://agriexchange.apeda.gov.in/Ready%20Reckoner/Cold_Storage/WesternRegion/GUJARAT.aspx

2. Latest-APMC-Contact-List.pdf. (n.d.). gsamb.gujarat.gov.in/Images/gsamarketingboard/pdf/Latest-APMC-Contact-List.pdf

3. FAO Map Catalog. (n.d.). https://data.apps.fao.org/map/catalog/srv/eng/catalog.search#/metadata/446ed430-8383-11db-b9b2-000d939bc5d8

4. Dey, B., Ferdous, J., & Ahmed, R. (2024). Machine learning based recommendation of Agricultural and horticultural crop farming in India under the regime of NPK, soil ph and three climatic variables. *Heliyon*, *10*(3). https://doi.org/10.1016/j.heliyon.2024.e25112