Project Design Phase-I Solution Architecture

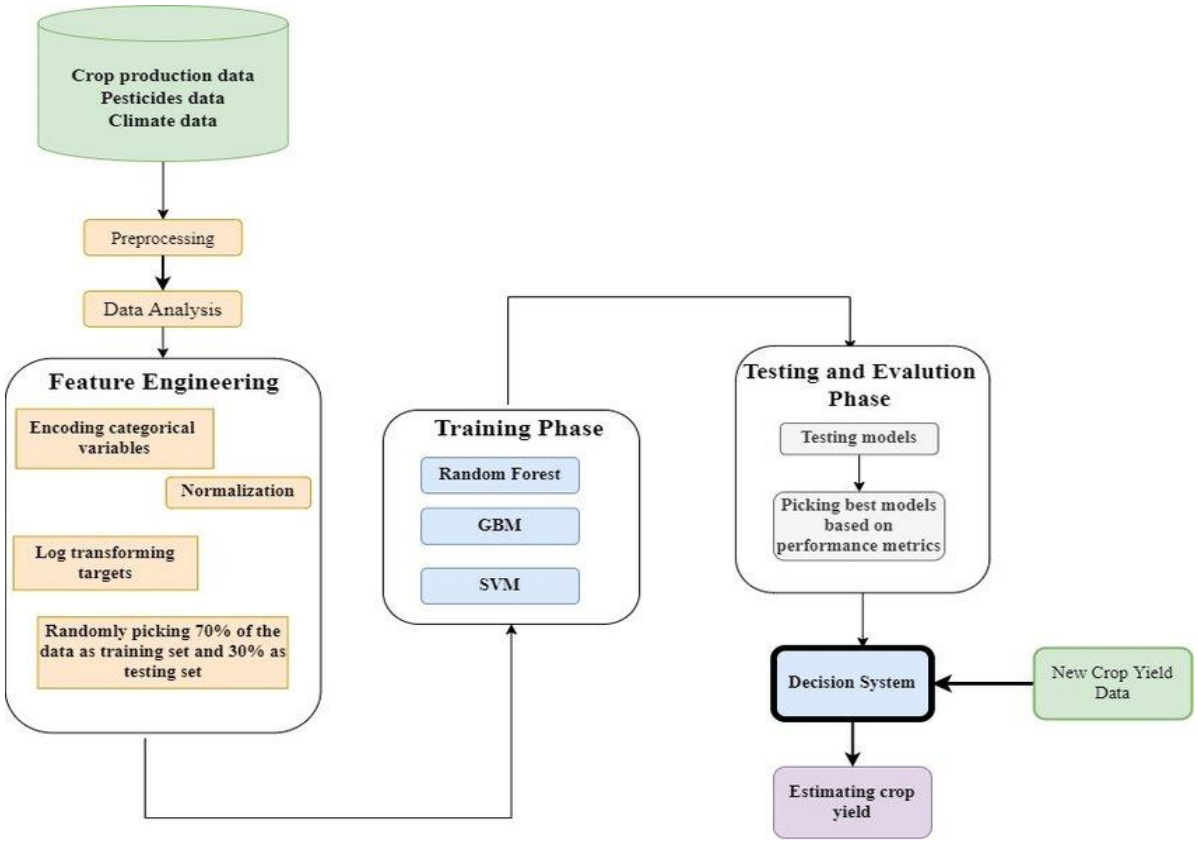
Date	19 September 2022
Team ID	PNT2022TMID27968
Project Name	Estimation of crop yield using data analytics.
Maximum Marks	4 Marks

Solution Architecture:

Solution architecture is a complex process – with many sub-processes – that bridges the gap between business problems and technology solutions. Its goals are to:

- Find the best tech solution to solve existing business problems.
- Describe the structure, characteristics, behaviour, and other aspects of the software to project stakeholders.
- Define features, development phases, and solution requirements.
- Provide specifications according to which the solution is defined, managed, and delivered.

Solution Architecture Diagram:



Data mining : It is the process of sorting through large data sets to identify patterns and relationships that can help solve business problems through data analysis. Data mining techniques and tools enable enterprises to predict future trends and make more-informed business decisions. It can be used in a variety of ways, such as database marketing, credit risk management, fraud detection, spam Email filtering, or even to discern the sentiment or opinion of users.

Data preprocessing : It is an important step to prepare the data to form a QSPR model. There are many important steps such as data cleaning, data transformation, and feature selection .It aims to facilitate the training/testing process by appropriately transforming and scaling the entire dataset. Preprocessing is necessary before training the machine learning models.

Data normalization **:** It is the process of reorganizing data within a database so that users can utilize it for further queries and analysis. Simply put, it is the process of developing clean data. This includes eliminating redundant and unstructured data and making the data appear similar across all records and fields. This includes creating tables and establishing relationships between those tables according to rules designed both to protect the data and to make the database more flexible by eliminating redundancy and inconsistent dependency.

Data visualization: It is the practice of translating information into a visual context, such as a map or graph, to make data easier for the human brain to understand and pull insights from. The main goal of data visualization is to make it easier to identify patterns, trends and outliers in large data sets. Data visualization is the representation of data through use of common graphics, such as charts, plots, and even animations. These visual displays of information communicate complex data relationships and data-driven insights in a way that is easy to understand.

Data Analysis: It is described as a particularized form of analytics. The purpose of Data Analysis is to extract useful information from data and taking the decision based upon the data analysis. A simple example of Data analysis is whenever we take any decision in our day-to-day life is by thinking about what happened last time or what will happen by choosing that particular decision. A Descriptive analysis can be performed on this. It analyzes the data by focusing on insights into business data.