**Lite Survey**  
  
The first survey discusses the use of a data mining technique called naive Bayes classification to classify different types of agricultural land soils. The study aimed to extract knowledge from a large dataset of soil profile data and compare the effectiveness of different data mining techniques. The findings of the study could have potential benefits for agriculture, soil management, and environmental management.

The second survey explores factors that affect the yield and quality of potato crops in Canterbury, New Zealand. The study found that soil-borne diseases, soil compaction, and irrigation management were consistent factors in reduced yields. The study also examined the impact of different treatments, such as soil fumigants and pesticides, on soil-borne pathogens and crop yield.

The third survey provides a retrospective and literature survey of a statistical design and numerical optimization technique called response surface methodology (RSM). The study discusses the history and widespread application of RSM, especially in the chemical and process industries. The study also identifies current areas of research and potential areas for future research in the field of RSM.  
  
**Implementation:**  
These are the parameter we used;  
Nitrogen

Phorsporous

Potassium

Temperature

Humidity

Ph(of soil)

Rainfall(in mm)

**Methodology:**  
  
  
  
  
1..First. Data Preparation and Model Building:

The dataset is imported, processed, and cleaned.

Data visualization is performed using seaborn and matplotlib.

The dataset is split into training and testing sets. Various machine learning models are built, including feature selection methods, KNN, Naive Bayes, Bagging Classifier, Random Forest, Decision Tree, SVM, Gradient Boosting, and Voting Classifier. These models are trained on the dataset for prediction.

2…Flask Framework Integration:

Flask Framework is integrated with SQLite for user registration and login.

The user gives input as feature values which are preprocessed for prediction.

3…Prediction and Output Display:

The trained model is used for prediction and the predicted result is displayed.

The final outcome is displayed through the frontend.  
  
**Results:**  
After visualization, the data is cleaned and split into the train and test modules. Feature selection is applied to several algorithms, and the predicted and tested data are compared to calculate accuracy using the score() method. The results show that the voting classifier has the highest accuracy and can be used for agricultural analytics.  
  
**Conclusion:**  
  
so to conclude, This paper employed various feature selection and classification techniques to predict the yield size of plant cultivations, which is a challenging task in agriculture. The results show that an ensemble technique provides higher prediction accuracy than existing classification techniques, enabling the planning of sowing structures for cereals, potatoes, and other energy crops at the farm and country scale, leading to measurable financial benefits through the use of modern forecasting techniques.  
  
**future scope:**Increasing the range of crops and adding more crop yield data will enhance the project's usefulness. Additionally, deploying the website prototype with cloud support can make it accessible to a wider audience in the digital world. Furthermore, the project can also include recommendations for fertilizers and pesticides for each crop.