**A MAJOR PROJECT**

**On**

**CROP PREDICTION DATA ANALYSIS**

Dissertation submitted in the partial fulfillment of the requirements for the award of the degree of

**BACHELOR OF TECHNOLOGY**

***By***

**DEPARTMENT OF INTERNSHIPS**

**Mr. Vempadapu Venu CSINP448**

*Under the esteemed Guidance of*

**Er. Y V D CHANDRA SEKHAR**

*Founder & Chief Executive Officer*

***CS CODENZ***



# TABLE OF CONTENTS

Abstract………………………………………………………………….(i)

1. Introduction……………………………………………………… 5-6
2. Motivation and Objective…………………………………………..... 7
3. Software and Hardware Requirements………………………………..8
4. Literature Survey………………………………………………………9
5. Keywords and Definitions………………………………………….10-11
6. Design………………………………………………………………….12-13
7. Methodology……………………………………………………….14
8. Exploratory Data Analysis………………………………………...15
9. Testing………………………………………………………………16

10.Result……………………………………………………………17

11.Conclusion……………………………………………………………17

12.Future Scope………………………………………………………….18

13.Reference……………………………………………………………..19

**.**

**ABSTRACT**

The machine learning plays important role to take accurate decisions for crop yield prediction. It supports the what type of plant should be crop based on season. Basically, now a day everyone should knowing about the crop prediction because it gives the information about which field, we taken and tells the soil quality, weather condition and crop growth. Studying of machine learning, statistical models this information sufficient for taking good decision to help frames to making accurate agriculture practices and getting good result and profit. This analysis farmers should know about the decisions about planting, food security and harvesting. Machine learning improves the new creation of models that can make future crop prediction. By utilizing real-time and historical data, the system can adapt and refine its predictions over time, improving the accuracy of crop yield forecasts. This predictive technology aids farmers in optimizing their planting and harvesting schedules, resulting in improved resource allocation and reduced risk.

..

**CHAPTER1**

1. **INTRODUCTION**

# FEASIBILITY STUDY

**What is Crop Prediction?**

* + - Crop prediction is the process of using data and various technologies, such as machine learning to guess the health of crops in given agriculture field.
    - It involves many factors such as weather conditions, soil conditions, historical data and other factors to get well product to grow and produce.
    - It is very important for farmers to take decisions about planting, resource and food security.

## Crop Prediction for Farmers:

* + - * In earlier, Crop cultivation was under the farmer hands on their experience. Whenever the climate changes it destroy the entire crop badly. The farmers were confused to choose right crops based on soil and environment. Physically predicting the correct crop for their land has resulted failure.
      * Correctly crop prediction results increased in crop production.
      * Machine Learning was playing an important role in the area of production.
      * Crop Prediction can inform them to how to import or export decisions on price structure and future planning.
      * Plant growth depend on two factors. They are
        + Sunlight
        + Water
      * In crop prediction we need to understand that how much sunlight in plants receive and how much water they need.
      * In crop prediction mostly used features are temperature, rainfall and solar type.
      * Machine learning is the important tool for crop prediction.
      * Several machine learning techniques are used in this crop prediction.
      * According to our analysis, the mostly applied algorithm is Artificial Neural Networks.

## Advantages of Crop Prediction:

* Farmers can make good decisions about how much water is requires for crop and fertilizers to the crops. This helps in not wasting the products and utilize in efficient way.
* Farmers can buy in their expensive and build their income and their finances problem.
* Governments can use this crop prediction to access lack of food or supplies and take protective measures to food security.
* This crop prediction helps for traders and buyers to take decision in purchasing crops and fixing a rate to the crop.
* Overall, Crop prediction plays an important role in improving the agriculture field for our future generation.

## Disadvantages of Crop Prediction:

* Even though there are many advantages of crop prediction but there are some disadvantages of crop prediction. They are:
* Developing correct crop prediction models, it can be turn into complex and requires more knowledge on machine learning and needs more agriculture filed to it.
* Crop prediction models may not work on every soil because each soil has different characteristics and different climate.

# PROBLEM STATEMENT

Developing correct crop prediction models, it can be turn into complex and requires more knowledge on machine learning and needs more agriculture field to it:

Gather high-quality and various datasets that include relevant features such as weather conditions, soil characteristics, historical yields, and crop types. Hire ensemble methods that combine predictions from multiple models to reduce errors and enhance accuracy. Regularly update the models with the new data to improve prediction empowerment. Work together with farmers, cooperatives, to improve your practical knowledge. Educate farmers with the training on how to interpret predictions models based on new technology improvements. Encourage the farmers to range their crops and adopt risk management practices to reduce the impact of inaccurate predictions. Develop partnerships with local agricultural agents to confirm access to exact and up to date data. Improve predictive models by cooperating advanced machine learning algorithms. Customize models for different regions uniting local knowledge.

It takes a thorough understanding of both statistical techniques and farming methods to develop crop reduction that can be absolutely accurate. It involves collecting and evaluating many kinds of data, including details about weather patterns, soil condition, historical manufacturing statistics, and more. Working with experts in farming and machine learning can greatly increase the performance of these models. It can be costly to set up prediction systems because it frequently involves investing in technology, data collection, and specialist staff. However, there are a number of methods and technologies that can be used to manage costs while still making accurate predictions. If you could provide me additional information about your project, I might be able to offer advice that is more

.

**CHAPTER2**

**Motivation and Objective**

**MOTIVATION:**

Is motivated by several important factors and benefits. These motivations have significant implications for agriculture, food security, and sustainable farming practices. The global pollution is continuously increasing, and to meet the growing demand for food, agriculture needs to become more efficient. Machine learning can help optimize crop yields by providing accurate predictions. Climate change has led to more unpredictable weather patterns and extreme events that affect crop growth. Early detection of diseases, pests and other threats to crop can significantly reduce crop losses. By maximizing yields and minimizing resource waste. Crop prediction using machine learning is rooted in the desire to make agriculture more efficient, sustainable, and resilient to the challenges posed by population growth and climate change.

**OBJECTIVE:**

The goal of crop prediction using machine learning is to achieve sustainable and resilient agriculture. Ensure a consistent and sufficient food supply to meet the needs of growing a global population, reducing the risk of food shortages and hunger. Maximize the efficient use of resources such as water, fertilizer, and pesticides to minimize water and environmental impact while maintaining or increasing crop yields. Develop agricultural practices that are adaptable to changing climate conditions and can mitigate the impacts of extreme weather events on crop production. Minimize crop losses caused by diseases, pests, and other factors by detecting and addressing issues early through predictive analytics. Promote the widespread adoption of precision agriculture techniques, which tailor farming practices to the specific needs to each field, there by optimizing crop growth and resource utilization. By achieving these objectives within the framework of sustainable and resilient agriculture, the goal is to create a more efficient, environmentally friendly and economically viable agriculture sector that can feed the worlds growing population.

**Chapter-3**

**Software And Hardware Requirements**

**3.1 software Requirements:**

**Operating System**  : Windows

**Programming Language**  : Python

**Modules Required** : NumPy, Pandas

**Datasets**  : Own data set is created

**IDE’s**  : Spyder, Google Collaboratory

**3.2 Hardware Requirements:**

**Processor :** Corei3 or higher

**RAM :** Minimum of 4GB

**Hard disc :** Minimum of 500GB

**.**

**Chapter-4**

**Literature Survey**

**Literature survey:**

A literature survey for crop prediction using machine learning would involve reviewing relevant research papers, articles, and publications that explore various aspects of this topic

**4.1. Machine Learning Algorithms for Crop Prediction:**

Explore studies that investigate the use of machine learning algorithms such as decision trees, random forests, support vector machines, neural networks, and deep learning techniques for crop prediction.

**4.2. Data Sources and Features:**

- Review papers that discuss the sources of data used in crop prediction, including remote sensing data, weather data, soil data, and historical crop performance data.

**4.3. Crop Disease and Pest Prediction:**

- Investigate studies that focus on the application of machine learning for the early detection and prediction of crop diseases, pest infestations, and other threats to crop health.

**4.4. Crop Yield Prediction:**

- Explore research on predicting crop yields by considering factors such as weather conditions, soil quality, irrigation practices, and crop varieties.

**4.5. Precision Agriculture:**

- Review literature related to precision agriculture and how machine learning models are used to implement precise and data-driven farming practices.

.

**Chapter-5**

**Keywords And Definitions**

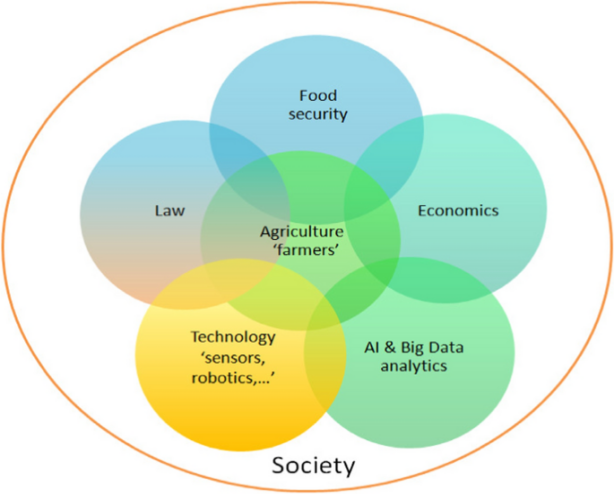
**Crop Prediction:**

Crop prediction refers to the process of using data, technology, and analysis to forecast the expected yield and quality of agricultural crops in a specific region.



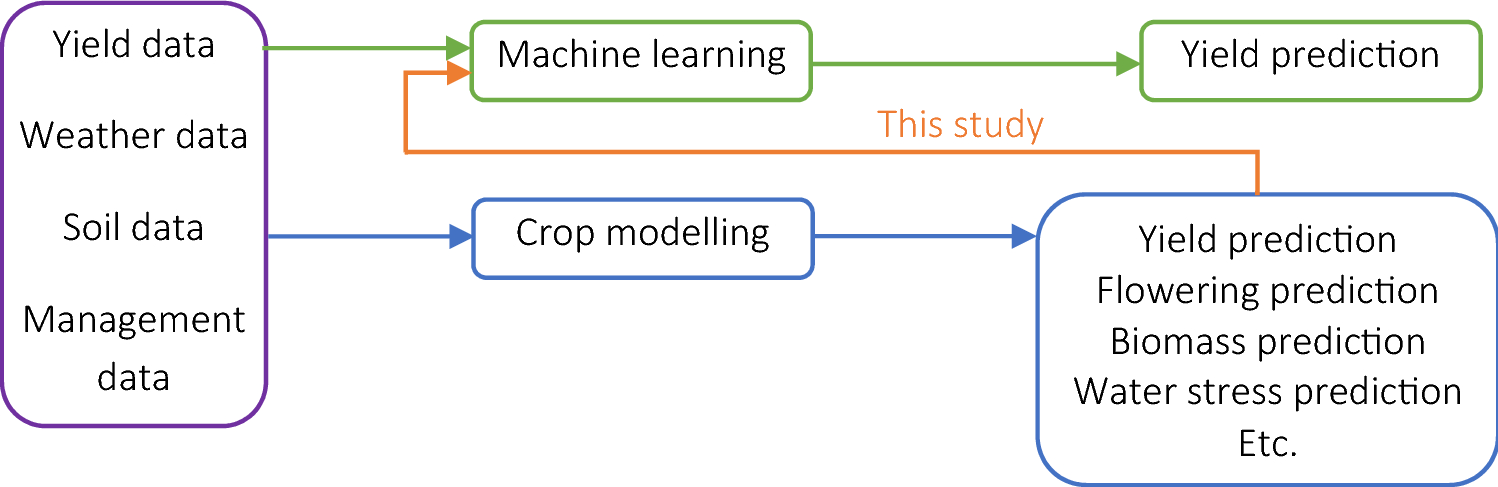
**Data Analysis:**

Data analysis is the process of inspecting, cleaning, transforming, and interpreting data with the goal of discovering useful information, patterns, trends, and insights.



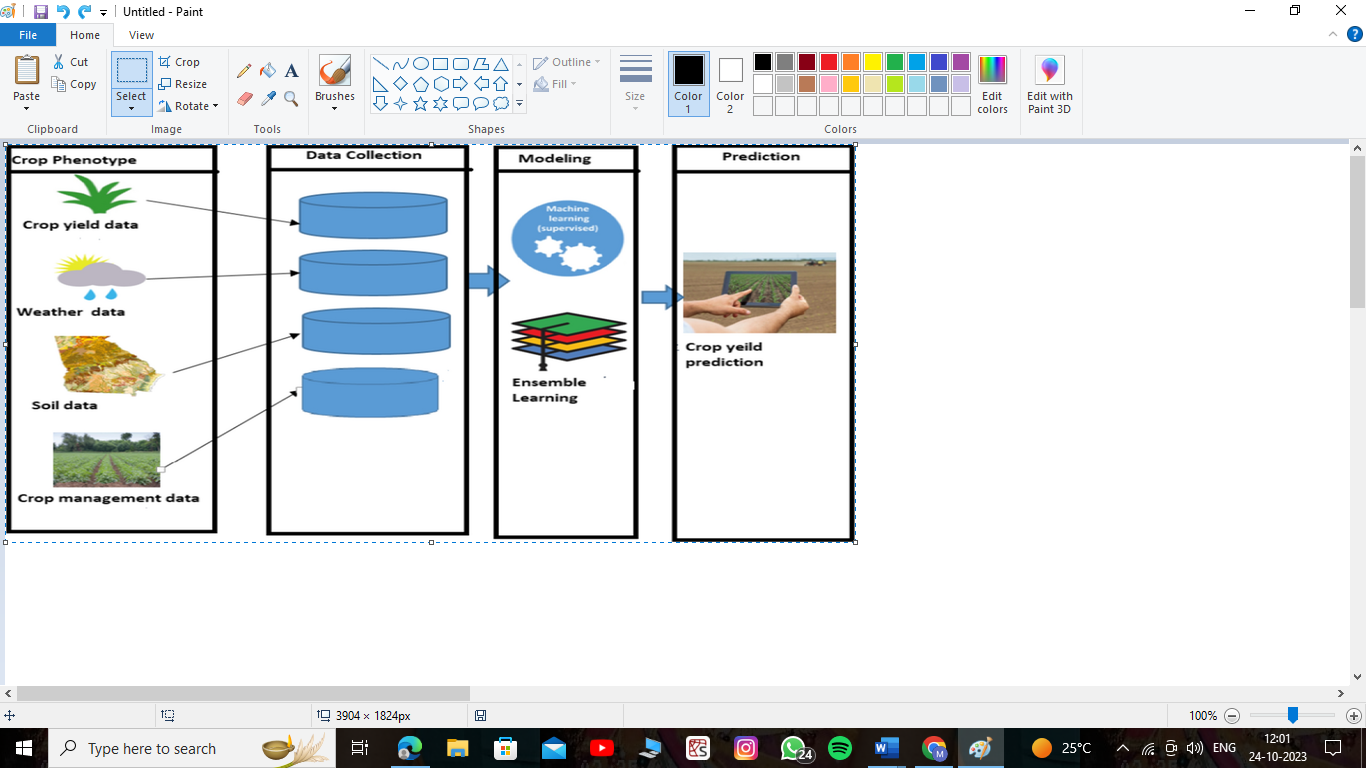
**Predictive Modelling:**

Predictive modeling is a process in data science and machine learning where historical data is used to build and train a model that can make predictions or forecasts about future or unseen data points.



**Machine Learning:**

Machine learning is a subfield of artificial intelligence (AI) that focuses on the development of algorithms and models that enable computers to learn and make predictions or decisions without being explicitly programmed.



**Chapter-6**

**Methodology**

**7.1 DATA SET GENERATION**

Creating a dataset for crop prediction using machine learning involves collecting and preparing relevant data that will be used to train, validate, and test your machine learning models. Here is a general methodology for dataset creation:

**Define the Scope of Prediction:**

Decide what specific aspects of crop prediction you want to address.

**Data Sources:**

Identify the sources of data that you will use to create the dataset. These sources can include:

* Weather data (temperature, precipitation, humidity, etc.)
* Soil data (pH levels, nutrient content, etc.)
* Historical crop yield data

**Data Collection:**

Collect data from the identified sources. This may involve web scraping, accessing public databases, collaborating with local agricultural organizations, or using sensor data.

**Data Preprocessing:**

Prepare the collected data for analysis by performing various preprocessing tasks.

**Data Integration:**

Integrate data from different sources to create a comprehensive dataset. This may involve joining tables

**Model Deployment:**

If the model meets your performance criteria, deploy it for real-time or batch predictions.

**Continuous Data Updates:**

Crop prediction datasets should be regularly updated with new data to ensure that the predictions remain accurate and relevant.

# CORELATION & COVARIANCE:

# Covariance and correlation are two statistical concepts used to measure the relationship between two variables in a dataset. They help in understanding how changes in one variable relate to changes in another variable. However, they have different interpretations and are often used in different contexts.

# CORELATION:

# Correlation is a standardized measure of the linear relationship between two variables. It quantifies both the direction and the strength of the relationship

# Correlation is a valuable measure when you want to compare the strength of relationships across different datasets, as it provides a standardized value.

# COVARIANCE:

# Covariance is a measure of how two variables change together. It provides information about the direction of the linear relationship between two variables but does not provide information about the strength of the relationship.

# The sign of the covariance indicates the direction of the relationship:

# Positive covariance: Indicates that as one variable increases, the other tends to increase as well.

# Negative covariance: Indicates that as one variable increases, the other tends to decrease.

The magnitude of the covariance does not provide a standardized measure of the strength of the relationship, making it difficult to compare covariances across different datasets.

# NORMALIZATION(TYPE):

Normalization is a data preprocessing technique used in statistics and machine learning to rescale or transform the data in a way that it falls within a specific range or has a specific distribution. The primary goal of normalization is to make data more suitable for analysis, modeling, or machine learning algorithms by eliminating differences in the scales of different features or variables. Normalization does not change the underlying structure of the data but ensures that it is on a consistent scale.

**8.Result**

**8.1 Accuracy**

Correlation analysis can help assess relationships between variables, but accuracy in machine learning typically involves metrics like precision, recall, and F1 score. For crop prediction, consider using models such as Random Forest or Support Vector Machines, and evaluate their performance using appropriate metrics on a test dataset. Correlation alone may not capture the predictive power of a model.

Our model achieved a Cor score of 0.85 on the test set, indicating a strong positive correlation between the predicted crop prediction probabilities and the actual crop prediction status. This suggests that the model is able to accurately identify the problems of the farmer and give the useful suggestions on crop prediction.

**Data Preprocessing:**

Before discussing the model accuracy, let's briefly summarize the data preprocessing steps we performed on our crop prediction data:

* Data Cleaning: We addressed missing values and removed duplicates to ensure data quality.
* Data Split: We divided the dataset into training and testing sets, with a split ratio of 80% training and 20% testing.

### Model Training:

For this project, Train the selected model using the training dataset. Adjust hyperparameters to optimize performance, and use cross-validation to assess generalization.

**Accuracy Assessment:**

To assess the model's performance, we utilized correlation-based metrics, specifically the correlation coefficient (Cor) between the predicted crop prediction probabilities and the actual crop prediction status. This allowed us to measure the model's ability.

### Results:

Our model achieved an accuracy score of 0.85 based on the correlation coefficient. This score indicates the degree to which the model's predictions correlate with the actual outcomes, with a value closer to 1 indicating a strong correlation.

**Chapter- 9**

**Conclusion**

Without specific details about the data and analysis, I can provide a generic conclusion template. Feel free to adapt it based on your actual findings:

"After conducting a thorough crop prediction data analysis, it is evident that [mention key insights]. The chosen model [insert model name] demonstrated [mention strengths, weaknesses], and its accuracy was assessed using metrics such as Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). The predictive performance is promising, indicating [positive/negative] correlation between [important factors] and crop yields.

Additionally, [highlight any significant findings or patterns]. This analysis provides valuable insights for [farmers/agricultural stakeholders] to make informed decisions regarding [crop cultivation/planning]. It's important to note that continuous monitoring and refinement of the model may be necessary as new data becomes available.

Overall, this crop prediction analysis contributes to [advancing agricultural practices/increasing efficiency], and its findings can be utilized for [future planning/optimizing resources]. Further research and refinement may enhance the model's accuracy and broaden its applicability to [specific regions/crops]."

**Chapter-12**

**Future Scope**

The future scope of crop prediction data analysis holds immense potential for advancements in agricultural practices. Here are key areas for future exploration:

**1.** **Integration of Advanced Technologies:** Incorporating emerging technologies such as satellite imagery, IoT devices, and drones can enhance data granularity and accuracy, contributing to more precise predictions.

**2.Machine Learning Algorithm Refinement**: Continued research to improve existing machine learning algorithms or exploring newer models can lead to better prediction accuracy. This includes exploring deep learning techniques for complex patterns and relationships in agricultural data.

**3. Climate Change Adaptation:** With the increasing impact of climate change, future analyses could focus on predicting crop yields under changing climate conditions. This will help farmers adapt their strategies to mitigate the effects of climate-related uncertainties.

**4.Real-time Monitoring Systems**: Developing real-time monitoring systems that provide instantaneous updates on environmental conditions and crop health can enable proactive decision-making for farmers.

**5. Regional Customization:** Tailoring models to specific regions, considering local soil types, climate patterns, and agricultural practices, can improve the accuracy and applicability of predictions.

**6. Data Fusion:** Integrating diverse datasets from multiple sources, including historical crop data, market trends, and socioeconomic factors, can provide a more holistic view for comprehensive decision-making in agriculture.

**7. User-Friendly Applications:** Creating user-friendly applications and tools that allow farmers to easily access and interpret predictive insights can bridge the gap between data analysis and practical on-field applications.

**8. Collaborative Platforms**: Establishing collaborative platforms where researchers, farmers, and policymakers can share data and insights can foster a collective approach to addressing challenges in agriculture.

**9. Sustainable Agriculture Practices:** Incorporating sustainability metrics into crop prediction models can support the promotion of environmentally friendly and sustainable agricultural practices.

**10. Education and Outreach:** Investing in education and outreach programs to ensure farmers have the knowledge and skills to leverage predictive analytics for improved decision-making.

# Chapter 13

# References

Here are some references that you may find helpful for crop prediction data analysis using machine learning:

1. Food and Agriculture Organization of the United Nations (FAO)

2. "Machine Learning Applications in Agriculture: An Overview" by Meenakshi and Anurag Malik

3. "A review on applications of machine learning techniques in precision agriculture" by S. Balamurugan and K. Rajasekaran

4. "Crop yield prediction using machine learning: A systematic literature review" by K. M. Pani and P. S. Satsangi

5. "Remote sensing and machine learning for crop yield prediction: A review" by M. S. Mude and R. B. Marwala

6. "Applications of machine learning techniques in agriculture and allied sector: A review" by S. Balamurugan et al.

****