REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITE IBN KHALDOUN - TIARET

MEMOIRE

Présenté à :

FACULTÉ DES MATHEMATIQUES ET DE l'INFORMATIQUE
DÉPARTEMENT D'INFORMATIQUE

Pour l'obtention du diplôme de **Master**

Spécialité : Génie Logiciel

En vue de créer une startup

PerfectSeed

PerfectSeed

Par :

MIMOUNE Abdelaziz
LANTRI Farouk

Sur le thème

Precision Agriculture Using Crop Recommendation Systems

Soutenu publiquement le 11 / 06 / 2024 à Tiaret devant le jury composé de :

Mr. KHARROUBI Sahraoui	MCA	Université de Tiaret	Président
Mr. BOUDAA Boudjemaa	MCA	Université de Tiaret	Encadrant
Ms. LAAREDJ Zohra	MAA	Université de Tiaret	Examinateur
Mr. HADIDI Abed	MCA	Université de Tiaret	Représentant de l'incubateur
Mr. RAHOUANI Ilyes	CRMA		Représentant du partenaire économique

2023-2024

Acknowledgement

Foremost, praises and thanks to Allah, the Almighty, for His showers of blessings throughout my research work which allowed me to complete it successfully.

We would like to express our sincere gratitude to several individuals and organizations for supporting us throughout our Graduate study.

First, we wish to express our sincere gratitude to our supervisor, Professor Boudaa Boudjemaa, for his enthusiasm, patience, insightful comments, helpful information, practical advice and unceasing ideas that have helped us tremendously at all times in our research and writing of this thesis. His immense knowledge, profound experience and professional expertise in software engineering has enabled us to complete this research successfully. Without his support and guidance, this project would not have been possible. We could not have imagined having a better supervisor in our study.

We are also deeply grateful to the members of the jury for their valuable time, insightful comments, and constructive feedback, which have greatly contributed to the improvement and refinement of this work.

Also, we must express our very profound gratitude to our parents for providing us with unfailing support and for doing the best as they knew and could, to be there for us throughout our years of study and through our process of life.

Our deepest appreciation,

LANTRI Farouk

MIMOUNE Abdelaziz

Table of Contents

Abstract

Precision agriculture has emerged as a data-driven approach to optimizing agricultural practices, improving resource efficiency, and promoting sustainable farming methods. One of the key challenges in precision agriculture is the selection of suitable crops based on various factors such as soil conditions, climate, market demand, and resource availability. This report presents a crop recommendation system that leverages informatics techniques and data-driven approaches to provide personalized crop recommendations to farmers.

The proposed crop recommendation system employs a hybrid approach that combines content-based filtering and collaborative filtering techniques. Content-based filtering analyzes the characteristics of farms, such as soil composition, weather patterns, and historical crop data, to identify suitable crops based on their requirements. Collaborative filtering utilizes the collective knowledge and experiences of similar farms to recommend crops that have performed well in comparable conditions.

The system integrates various data sources, including soil sensor data, weather data, remote sensing imagery, and user-generated data from farmer communities. These data sources are preprocessed and integrated using data mining and data fusion techniques, enabling the system to capture a comprehensive view of farm conditions and crop performance. For example, data fusion techniques can be used to combine information from multiple sensors to obtain a more accurate and comprehensive understanding of farm conditions. One popular method for sensor data fusion is the Kalman filter. Which can be used to fuse data from soil moisture sensors, temperature sensors, and weather stations to estimate the overall soil water content and temperature conditions.

Machine learning algorithms, such as decision trees, support vector machines, and neural networks, are employed to build predictive models that can estimate crop yields, identify potential risks, and optimize resource allocation. These models are trained on historical data and continuously updated with new data from sensors and user feedback, ensuring that the recommendations remain relevant and accurate.

The crop recommendation system is designed with a user-friendly interface that allows farmers to input farm-specific data and access personalized crop recommendations. The system also provides decision support tools, such as yield predictions, resource optimization suggestions, and risk assessment reports, to assist farmers in making informed decisions.

The performance of the crop recommendation system is evaluated using various metrics, including accuracy, ranking quality, diversity, and novelty. The system is benchmarked against existing crop recommendation approaches and evaluated through field trials and simulations.

The proposed crop recommendation system aims to contribute to the advancement of precision agriculture by providing data-driven and personalized crop recommendations, enabling farmers to optimize their operations, improve yields, and promote sustainable farming practices.

Keywords: Precision Agriculture, Crop Recommendation System, Content-Based Filtering, Collaborative Filtering, Machine Learning, Data Mining, Decision Support Systems.

Abstract

ظهرت الزراعة الدقيقة كنهج قائم على البيانات لتحسين الممارسات الزراعية، وتحسين كفاءة الموارد، وتعزيز أساليب الزراعة المستدامة. أحد التحديات الرئيسية في الزراعة الدقيقة هو اختيار المحاصيل المناسبة بناءً على عوامل مختلفة مثل ظروف التربة والمناخ والطلب في السوق وتوافر الموارد. يقدم هذا التقرير نظامًا لتوصية المحاصيل يستفيد من تقنيات المعلوماتية والنهج القائمة على البيانات لتقديم توصيات محاصيل مخصصة للمزارعين

يستخدم نظام توصية المحاصيل المقترح نهجًا هجينًا يجمع بين تقنيات التصفية القائمة على المحتوى وتقنيات التصفية التعاونية. تحلل التصفية القائمة على المحتوى خصائص المزارع، مثل تكوين التربة وأنماط الطقس وبيانات المحاصيل التاريخية، لتحديد المحاصيل المناسبة بناءً على متطلباتها. تستخدم التصفية التعاونية المعرفة والخبرات الجماعية للمزارع المماثلة للتوصية بالمحاصيل التي أدت أداءً جيدًا في ظروف مماثلة

يدمج النظام مصادر بيانات متنوعة، بما في ذلك بيانات مستشعرات التربة، وبيانات الطقس، وصور الاستشعار عن بعد، والبيانات التي ينشئها المستخدمون من مجتمعات المزارعين. تتم معالجة هذه المصادر البيانية ودمجها باستخدام تقنيات التنقيب عن البيانات ودمج البيانات، مما يمكّن النظام من التقاط رؤية شاملة لظروف المزرعة وأداء المحاصيل. على سبيل المثال، يمكن استخدام تقنيات دمج البيانات لجمع المعلومات من مستشعرات متعددة للحصول على فهم أكثر دقة وشمولية لظروف المزرعة. إحدى الطرق الشائعة لدمج بيانات المستشعرات هي مرشح كالمان، والذي يمكن استخدامه لدمج البيانات من مستشعرات رطوبة التربة ومستشعرات درجة الحرارة ومحطات الأرصاد الجوية لتقدير محتوى الماء في التربة وظروف درجة الحرارة بشكل عام

يتم استخدام خوارزميات التعلم الآلي، مثل أشجار القرار وآلات المتجهات الداعمة والشبكات العصبية، لبناء نماذج تنبؤية يمكنها تقدير غلات المحاصيل، وتحديد المخاطر المحتملة، وتحسين تخصيص الموارد. يتم تدريب هذه النماذج على البيانات التاريخية وتحديثها باستمرار ببيانات جديدة من أجهزة الاستشعار وتعليقات المستخدمين، مما يضمن بقاء التوصيات ذات صلة ودقيقة

تم تصميم نظام توصية المحاصيل بواجهة سهلة الاستخدام تسمح للمزارعين بإدخال بيانات خاصة بالمزرعة والوصول إلى توصيات محاصيل مخصصة. يوفر النظام أيضًا أدوات دعم القرار، مثل توقعات المحصول، واقتراحات تحسين الموارد، وتقارير تقييم المخاطر، لمساعدة المزارعين في اتخاذ قرارات مستنيرة

يتم تقييم أداء نظام توصية المحاصيل باستخدام مقاييس مختلفة، بما في ذلك الدقة وجودة الترتيب والتنوع والجدة. تتم مقارنة النظام مع نُهج توصية المحاصيل الحالية وتقييمه من خلال التجارب الميدانية والمحاكاة

يهدف نظام توصية المحاصيل المقترح إلى المساهمة في تقدم الزراعة الدقيقة من خلال تقديم توصيات محاصيل قائمة على البيانات ومخصصة، مما يمكّن المزارعين من تحسين عملياتهم وتحسين المحاصيل وتعزيز ممارسات الزراعة المستدامة

الكلمات المفتاحية : الزراعة الدقيقة، نظام توصية المحاصيل، التصفية القائمة على المحتوى، التصفية التعاونية، التعلم الآلي، التنقيب عن البيانات، أنظمة دعم القرار .

List of figures

List of tables

General Introduction

General introduction

A. Background

Agriculture plays a vital role in ensuring food security, economic growth, and environmental sustainability. However, traditional farming practices often face challenges in optimizing resource utilization, minimizing environmental impact, and adapting to changing climatic conditions. Precision agriculture has emerged as a data-driven approach to address these challenges by leveraging informatics technologies and data-driven decision-making.

B. Problem Statement

One of the critical challenges in precision agriculture is the selection of suitable crops based on various factors, including soil conditions, climate, market demand, and resource availability. Traditionally, crop selection has relied on experience and intuition, which may not always account for the complexities and spatial variability of farm conditions. This can lead to suboptimal crop yields, inefficient resource utilization, and potential environmental degradation. Additionally, the vast array of available crop options and the need to consider multiple decision criteria (such as yield potential, resource requirements, and market demand) further complicate the crop selection process. There is a need for data-driven and personalized crop recommendation systems that can assist farmers in making informed decisions while considering the unique characteristics and constraints of their farming operations.

Challenges in crop selection

The process of crop selection involves various challenges, including:

1. Accounting for spatial variability in soil conditions, topography, and microclimate within a farm.
2. Integrating multiple data sources, such as soil sensor data, weather data, remote sensing imagery, and historical crop performance data.
3. Considering market demand, economic factors, and resource availability in addition to agronomic factors.
4. Adapting to changing environmental conditions and climate patterns.
5. Incorporating farmers' preferences, risk tolerance, and local knowledge into the decision-making process.

C. Delimitation

This study focuses on developing a crop recommendation system for precision agriculture using techniques and data-driven approaches. The scope of the study is delimited as follows:

1. The crop recommendation system will primarily target common crops cultivated in the study region, considering the local soil conditions, climate, and market demand.
2. The system will integrate data sources that are readily available or can be collected within the scope of the study, such as soil sensor data, weather data, remote sensing imagery, and historical crop data.
3. The study will focus on developing and evaluating the recommendation system's performance using simulated data and limited field trials, subject to resource and time constraints.
4. While the system may incorporate economic factors and resource availability, detailed financial analysis and optimization of farming operations are beyond the scope of this study.

D. Approach

The proposed crop recommendation system employs a content-based filtering approach using specific machine learning techniques to provide personalized crop recommendations for precision agriculture. This system analyzes farm-specific characteristics and historical data to suggest the most suitable crops for each unique agricultural setting.

Here's the Key components of the revised approach:

1. **Data Collection and Preprocessing**
    - Gather relevant farm data, including soil composition, weather patterns, topography, and historical crop performance.
    - Implement data cleaning and normalization techniques to ensure data quality and consistency.
    - Perform feature engineering to create meaningful attributes that capture the essence of farm conditions and crop requirements.
2. **Feature Selection and Engineering**
    - Employ feature selection algorithms (e.g., Principal Component Analysis or Random Forest feature importance) to identify the most relevant features for crop recommendation.
    - Create new features that capture complex relationships between farm characteristics and crop suitability.
    - Optimize the feature set to balance predictive power and computational efficiency.
3. **Content-Based Filtering Model**: Develop and compare multiple machine learning models
    - Naive Bayes Classifiers (NB)
    - Decision Trees and Random Forests
    - Support Vector Machines (SVMs)
    - Artificial Neural Networks (ANNs)
    - XGBoost (Extreme Gradient Boosting)
    - LightGBM (Light Gradient Boosting Machine)

    For each model:

    - Train using historical data on crop performance in various conditions, treating crop recommendation as a multi-class classification problem.
    - Implement cross-validation techniques to ensure model robustness and generalizability.
    - Fine-tune hyperparameters to optimize performance.
4. **Recommendation Generation**
    - Use the trained machine learning models to predict the suitability scores for different crops based on the input farm characteristics.
    - Generate a ranked list of crop recommendations based on these predicted suitability scores.
    - If applicable, incorporate ensemble methods to combine predictions from multiple models for improved accuracy.
5. **System Integration and User Interface**:
    - Develop a web application for farmer interaction with the crop recommendation system.
    - Create a user-friendly interface with:

- Input form for farm-specific data (soil, climate, topography, crop history)
- Backend processing system connected to machine learning models
- Display of crop recommendations with rankings and basic information
- Simple visualizations of results
- Implement data security and privacy measures.
- Ensure responsiveness for both desktop and mobile devices.

6. Continuous Learning and Adaptation

- Design the system to incorporate feedback on crop performance and actual outcomes.
- Implement periodic model retraining to improve recommendation accuracy over time.

7. Performance Evaluation

- Define appropriate evaluation metrics (e.g., accuracy, precision, recall, F1-score) to assess the quality of crop recommendations.
- Compare the performance of different machine learning algorithms to identify the most effective model or ensemble for crop recommendation.
- Conduct A/B testing or field trials to validate the system's effectiveness in real-world scenarios.

E. Outline

This Thesis is structured into three chapters beside a general introduction and general conclusion:

- **General introduction**
  This thesis explores the field of precision agriculture and crop recommendation systems, highlighting their importance, key components, and potential for improving agricultural practices.

- **Chapter one: precision agriculture**
  This chapter provides an in-depth overview of precision agriculture, including its evolution, key technologies, benefits, challenges, and successful case studies from around the world.

- **Chapter two: crop recommendation system**
  This chapter focuses on crop recommendation systems, covering the factors influencing crop selection, data sources and collection methods, machine learning algorithms, system architectures, real-world applications, and future research opportunities in this domain.

- **Chapter three: Implementation and design of the crop recommendation system**
  This chapter delves into the design and implementation of a crop recommendation system using various machine learning algorithms. It builds upon the concepts and techniques discussed in the next chapters on precision agriculture and crop recommendation systems.

- **General conclusion**
  This thesis has contributed to the advancement of precision agriculture by developing a crop recommendation system that employs sophisticated algorithms and approaches to provide farmers with tailored and data-driven recommendations for crop selection, thereby enabling more sustainable and efficient farming practices.

Chapter 1

Precision Agriculture

I. Introduction

The agricultural sector plays a vital role in sustaining human societies and ensuring food security. However, traditional farming practices face numerous challenges, including inefficient resource utilization, environmental degradation, and limited productivity. To address these issues, precision agriculture has emerged as a revolutionary approach that integrates advanced technologies and data-driven decision-making processes [1].
This chapter delves into the concept of precision agriculture, exploring its significance, evolution, and the pivotal role of technology in transforming agricultural practices. It begins by highlighting the importance of agriculture in society, encompassing its economic, social, and environmental aspects, while providing a brief overview of the historical development of agricultural methods.

Furthermore, the chapter outlines the numerous benefits of precision agriculture, such as improved crop yields and quality, reduced input costs, enhanced environmental sustainability, and increased efficiency in farm operations. It also addresses the challenges associated with implementing precision agriculture, including high initial costs, lack of standardization and interoperability, data privacy and security concerns, and limited availability of skilled personnel with integrating advanced technologies and data-driven decision-making processes.

To illustrate the real-world applications and success stories of precision agriculture, the chapter presents case studies from around the world, highlighting lessons learned and best practices.

Finally, the chapter explores future directions in precision agriculture, discussing emerging technologies and trends, as well as potential improvements and innovations, setting the stage for the introduction of recommender systems in the subsequent chapter.

1. Precision Agriculture: Integrating Informatics for Efficient Farming Systems

1.1 Overview of the importance of agriculture in society

Agriculture stands as a foundational pillar of human civilization, its importance transcending mere food production and permeating every facet of our societies. This multidimensional domain exerts a profound influence on economic prosperity, social cohesion, and environmental sustainability. Its importance can be highlighted through the following aspects:

1.1.1 Economic aspects

A. **Big data analytics**: The agricultural sector generates vast amounts of data, which can be analyzed using big data techniques to identify patterns, optimize resource allocation, and support data-driven decision-making for economic growth [2].
B. **E-commerce and digital marketplaces:** Digital platforms and online marketplaces facilitate the trade and export of agricultural products, enabling farmers to access wider markets and increasing economic opportunities [1].

1.1.2 Social aspects

A. **Digital literacy and capacity building**: Initiatives to enhance digital literacy and provide training on agricultural technologies are crucial for empowering rural communities and promoting sustainable livelihoods [3].

B. **Information and communication technologies (ICTs):** ICTs, such as mobile apps and digital extension services, play a vital role in disseminating agricultural knowledge, facilitating information exchange, and supporting rural development [1].

1.1.3 Environmental aspects

A. **Environmental monitoring and modeling:** Advanced technologies like remote sensing, sensor networks, and computational models enable monitoring and prediction of environmental factors, such as soil health, water usage, and greenhouse gas emissions, supporting sustainable agricultural practices [4].
B. **Precision agriculture techniques:** Site-specific management and variable rate application of inputs, facilitated by precision agriculture technologies like GPS and variable rate technology (VRT), help minimize environmental impact by reducing input waste and optimizing resource utilization [5].
C. **Decision support systems:** Artificial intelligence (AI) and machine learning (ML) algorithms can be integrated into decision support systems to assist farmers in making informed decisions regarding sustainable practices, optimizing resource use, and mitigating environmental impacts ([4]).

1.2 Brief overview of the evolution of agriculture

Agriculture has undergone significant transformations throughout human history, evolving from subsistence farming practices to large-scale commercial operations. The earliest forms of agriculture date back to the Neolithic Revolution, around 10,000 years ago, when humans transitioned from hunter-gatherer societies to settled and began cultivating crops and domesticating animals. Over time, advancements in tools, techniques, and scientific knowledge have shaped the development of agriculture. The introduction of irrigation systems, crop rotation, and selective breeding played crucial roles in improving crop yields and productivity. The Industrial Revolution of the 18th and 19th centuries brought about mechanization, which further increased efficiency and enabled large-scale production. [5]

1.3 Importance of precision agriculture

Precision agriculture, also known as site-specific crop management (SSCM) or precision farming, is an emerging approach that aims to optimize agricultural practices by tailoring them to the specific conditions of each field or even sub-field areas. The integration of these informatics tools and techniques into agricultural practices represents a significant shift towards data-driven, site-specific management, and optimization of resources, ultimately contributing to increased productivity, profitability, and sustainability in the agricultural sector.

1.4 The Role of Technology in Agriculture

1.4.1 Description of traditional farming practices

Traditional farming practices refer to the conventional methods of crop cultivation and livestock rearing that have been passed down through generations and are deeply rooted in local knowledge, culture, and experience including:

a. Manual labor-intensive operations: Traditional farming relied heavily on manual labor for tasks such as land preparation, sowing, weeding, irrigation, and harvesting, requiring significant human effort and time.

b. **Empirical knowledge and experience-based decision-making**: Farmers relied on traditional knowledge, experience, and intuition passed down through generations to make decisions regarding crop selection, planting times, and management practices.

c. **Limited data collection and record-keeping**: Data collection and record-keeping were primarily manual, often incomplete, and prone to errors, hindering informed decision-making and analysis.

d. **Lack of precision and site-specific management**: Traditional farming practices treated entire fields uniformly, without accounting for variations in soil conditions, topography, or micro-climates within the field.

e. **Inefficient resource utilization**: Without precise information on crop requirements and field conditions, farmers often applied inputs (water, fertilizers, pesticides) uniformly, leading to inefficient resource utilization and potential environmental impacts. [2]

1.4.2 Limitations and challenges faced by farmers without technology

A. **Limited access to real-time information:** Farmers lacked access to real-time data on weather conditions, soil moisture levels, crop health, and other critical factors, hindering timely decision-making and interventions [6].

B. **Inability to monitor and predict yield**: Without advanced monitoring tools and predictive models, estimating crop yields and identifying potential issues early on was challenging, affecting production planning and resource allocation [1].

C. **Inefficient resource management**: The absence of precise input application techniques and site-specific management led to inefficient use of resources, potential over-application or under-application of inputs, and increased production costs.

D. **Vulnerability to environmental factors:** Traditional farming practices were more susceptible to environmental stresses such as droughts, pests, and diseases, as farmers lacked early warning systems and preventive measures [1].

E. **Limited scalability and adaptability:** As farm sizes and operations grew, traditional farming practices became increasingly difficult to scale and adapt, hindering productivity and efficiency gains.

F. **Lack of traceability and quality control:** Manual record-keeping and limited monitoring capabilities made it challenging to ensure traceability, quality control, and compliance with food safety standards.

G. **Environmental impact:** Uniform application of inputs without considering site-specific needs increased the risk of environmental pollution, soil degradation, and unsustainable practices [6].

- These limitations highlight the importance of integrating advanced technologies and precision agriculture techniques to address the challenges faced by the agricultural sector.

1.4.3 Evolution of technology in agriculture

The evolution of technology in agriculture has been driven by advancements in many various disciplines. These developments have facilitated the transition from traditional farming practices to modern, technology-driven agricultural systems.

A. Data acquisition and management

a. **Remote sensing and aerial imagery (satellite, drones) for field mapping and monitoring.**

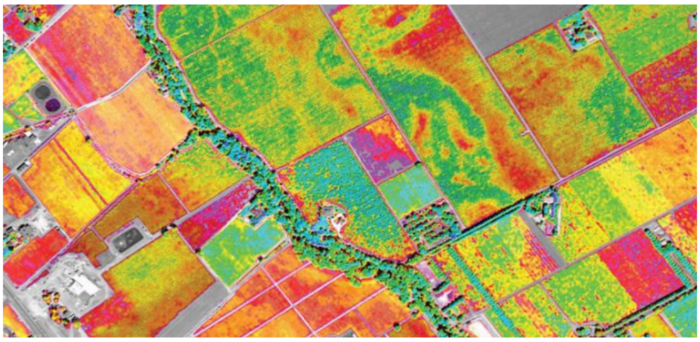


Figure 1: Example of a satellite image

b. **Sensor networks and Internet of Things (IoT) devices for real-time data collection on environmental conditions, soil moisture, and crop health *[2]***

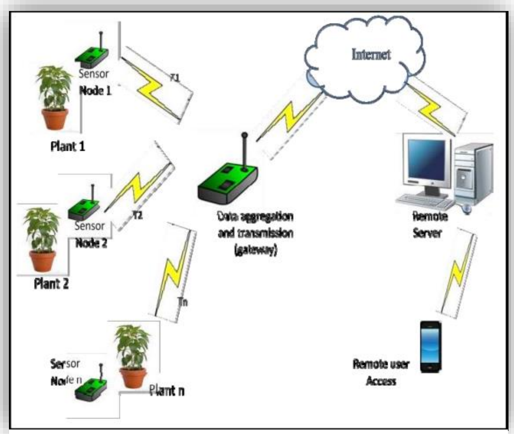


Figure 2: Proposed Soil Moisture Monitoring System

c. **Geographic Information Systems (GIS) for spatial data analysis and decision support. *[7]***

B. Computational modeling and decision support

- Crop growth models and yield prediction algorithms based on environmental data and agronomic factors.
- Machine learning and artificial intelligence for pattern recognition, predictive analytics, and decision support systems [4].

C. Automation and robotics

- GPS-guided autonomous machinery for precise field operations (planting, spraying, harvesting).
- Robotic systems for tasks like weeding, pruning, and yield monitoring, reducing labor requirements.

D. Communication technologies

- Mobile applications and digital extension services for disseminating agricultural information and best practices [1].
- Block chain and digital traceability systems for supply chain management and quality control.

1.4.4 Early adoptions and their impact on farming practices

The early adoption of informatics technologies in agriculture paved the way for significant transformations in farming practices, leading to increased efficiency, productivity, and sustainability. These early adoptions laid the foundation for modern precision agriculture techniques and data-driven decision-making [8].

A. Mechanization and automation

- Introduction of tractors, combine harvesters, and other mechanized equipment, reducing labor requirements and increasing efficiency [8].
- Adoption of GPS-guided machinery for precise field operations, minimizing overlaps and input wastage [6].

B. Precision farming techniques

- Variable rate technology (VRT) for site-specific application of inputs based on field conditions, optimizing resource utilization [5].
- Yield monitoring and mapping systems for identifying spatial variability and tailoring management practices accordingly [7].

C. Information and communication technologies (ICTs)

- Use of computers and software for farm management, record-keeping, and decision support.
- Adoption of mobile devices and digital platforms for access to agricultural information, weather data, and market prices [1].

1.5 Key Components of Precision Agriculture

1.5.1 Remote Sensing and Geographic Information Systems (GIS)

Remote sensing involves the acquisition of data about an object or phenomenon without making physical contact with it, typically using sensors mounted on satellites or aircraft. GIS refers to the computer-based systems designed to capture, store, analyze, manage, and present spatial or geographic data.

1.5.2 Global Positioning System (GPS) and Guidance Systems

GPS is a satellite-based navigation system that provides location and time information in all weather conditions, anywhere on or near the Earth. Guidance systems utilize GPS data to provide real-time feedback and control for various applications, such as precision agriculture.

1.5.3 Variable Rate Technology (VRT)

VRT refers to the technology that enables the automatic adjustment of input application rates (e.g., fertilizers, pesticides) based on spatial variability within a field. It utilizes data from various sources to create prescription maps for precise application.

1.5.4 Data Analytics and Decision Support Systems

Data analytics involves the extraction of actionable insights from large datasets, while decision support systems provide tools and information to assist in making inform decisions. In precision agriculture, these technologies help farmers optimize management practices based on data-driven analysis.

1.6 Benefits of Precision Agriculture

Precision Agriculture offers numerous benefits to modern farming practices. These benefits encompass improved crop yields and quality, reduced input costs, enhanced environmental sustainability, and increased operational efficiency.

1.6.1 Improve crop yields and quality

a. **Site-specific management**: Precision agriculture techniques enable site-specific management of fields, tailoring input applications (fertilizers, pesticides, water) based on spatial variability and real-time data from sensors and remote sensing [5].
b. **Predictive modeling**: Machine learning algorithms and crop growth models leverage data from various sources (weather, soil, imagery) to predict yield potential, identify stress factors, and optimize management strategies for improved crop quality and yields [4].
c. **Precision planting**: GPS-guided variable rate seeding and precision planting systems optimize seed distribution and plant populations based on field conditions, maximizing yield potential.

1.6.2 Reduce input costs

a. **Variable rate technology (VRT):** VRT systems, guided by GPS and field maps, enable precise application of inputs (fertilizers, pesticides, water) based on specific needs, reducing over-application and minimizing input costs.
b. **Targeted pest and weed management**: Integrating remote sensing data, GPS-guided sprayers, and decision support systems allow for targeted application of pesticides and herbicides, reducing unnecessary chemical usage and associated costs.
c. **Yield monitoring and analysis**: Yield mapping and data analysis techniques help identify underperforming areas within fields, enabling targeted input optimization and cost savings [7].

1.6.3 Enhanced environmental sustainability

a. **Reduced input waste**: Precision agriculture techniques like VRT and targeted input application minimize the over-application of fertilizers, pesticides, and water, reducing environmental pollution and resource depletion [5].
b. **Soil health monitoring:** Sensor networks and remote sensing technologies enable continuous monitoring of soil health parameters (moisture, nutrient levels, compaction), facilitating sustainable soil management practices.
c. **Precision irrigation**: Integrated decision support systems and sensor data help optimize irrigation scheduling and water application based on crop water requirements, reducing water wastage and minimizing environmental impact.

1.6.4 Increase efficiency in farm operations

a. **Precision guidance and automation:** GPS-guided machinery and autonomous systems facilitate precise field operations (planting, spraying, harvesting), minimizing overlaps, reducing labor requirements, and increasing operational efficiency.
b. **Data-driven decision support:** Farm management information systems (FMIS) and decision support tools leverage data from various sources (sensors, imagery, weather) to provide real-time recommendations, enabling efficient resource allocation and operational planning.
c. **Supply chain management:** Technologies like block chain and digital traceability systems streamline supply chain operations, enabling efficient tracking, quality control, and logistics management.

1.7 Challenges in Implementing Precision Agriculture

While precision agriculture offers numerous benefits, its widespread adoption and implementation face several challenges:

1.7.1 High initial costs

a. **Investment in hardware**: Implementing precision agriculture requires significant upfront investments in hardware components such as GPS receivers, yield monitors, sensor networks, and specialized machinery equipped with precision technology [9].
b. **Software and data management systems:** Adopting farm management information systems (FMIS), decision support tools, and data analytics platforms can be cost-intensive, especially for small-scale farmers.
c. **Training and capacity building:** Providing training and educational resources to farmers and personnel for effective utilization of precision agriculture technologies and data interpretation can add to the overall implementation costs.

1.7.2 Lack of standardization and interoperability

a. **Data formats and protocols:** The lack of standardized data formats and communication protocols across different precision agriculture technologies and systems can hinder data integration and interoperability, limiting the effective exchange and analysis of information.
b. **Hardware and software compatibility:** Incompatibility issues between hardware components (sensors, machinery) and software systems from different vendors can create barriers to seamless operation and data sharing.
c. **Data management and integration:** The heterogeneity of data sources (sensors, imagery, weather stations) and the need for effective data management and integration across different platforms pose challenges for efficient data utilization [2].

1.7.3 Data privacy and security concerns

a. **Sensitive data exposure**: The large volumes of data collected through precision agriculture technologies, including field maps, yield data, and farm operations, may contain sensitive information that could be vulnerable to unauthorized access or issue [10].

b. **Cyber security risks:** As precision agriculture systems become more interconnected and reliant on cloud-based platforms, there is an increased risk of cyber threats, such as hacking, data breaches, and system vulnerabilities.
c. **Data ownership and control**: Concerns over data ownership, control, and the potential use of farm data by third parties, such as input suppliers or technology providers, can raise privacy and trust issues among farmers.

1.7.4 Limited availability of skilled personnel

a. **Technical expertise:** Implementing and maintaining precision agriculture systems requires personnel with specialized technical knowledge and skills in areas such as GPS, remote sensing, data management, and software operations.
b. **Data analysis and interpretation**: The ability to effectively analyze and interpret large volumes of data generated by precision agriculture technologies requires personnel with expertise in data analytics, statistical modeling, and decision support systems [4].
c. **Training and education:** There is a need for comprehensive training and educational programs to develop a skilled workforce capable of leveraging precision agriculture technologies and data-driven approaches effectively.

1.8Case Studies of Precision Agriculture

Precision agriculture has revolutionized the way we cultivate crops and manage agricultural operations. By leveraging cutting-edge technologies, such as global positioning systems (GPS), remote sensing, and data analytics, farmers can optimize their practices, increase efficiency, and enhance sustainability. This section presents a collection of case studies from around the world, showcasing the successful implementation of precision agriculture techniques and the lessons learned from these experiences.

1.8.1 Success stories from around the world

A. Success Story 1: Precision Farming in Australia's Cotton Industry

The Australian cotton industry has widely adopted precision agriculture techniques, such as soil moisture monitoring, variable-rate irrigation, and yield mapping. According to a study by Roth et al. (2018), precision agriculture practices in the Australian cotton industry have resulted in an average water use efficiency improvement of 12% and an average yield increase of 8.5%. One successful example is the Culleen Estate, a large cotton farm in New South Wales, which has implemented a range of precision agriculture technologies, including GPS-guided machinery, soil moisture probes, and variable-rate fertilizer application. This has led to significant input savings and improved crop yields [11].

Metric	Precision Agriculture Practices
Water Use Efficiency	12% improvement
Yield	8.5% increase
Nitrogen Fertilizer Usage	25-30% reduction
Yield	5-10% increase

Table 1: the benefits of precision agriculture practices in the Australian cotton industry

B. Success Story 2: Site-Specific Nutrient Management in India's Rice Cultivation

A study by Dobermann and Al. (2004) conducted in several rice-growing regions of India demonstrated the benefits of site-specific nutrient management (SSNM) using precision agriculture techniques. By taking into account field-specific soil fertility and crop requirements, SSNM resulted in an average yield increase of 0.5-0.8 tons/ha and a reduction in nitrogen fertilizer use by 25-30% compared to conventional practices. One notable example is the adoption of SSNM by farmer cooperatives in the state of Andhra Pradesh, which led to significant improvements in productivity and profitability [12].

1.8.2 Lessons Learned and Best Practices

While the successful implementation of precision agriculture techniques has yielded significant benefits, it is essential to learn from these experiences and adopt best practices to maximize the potential of these technologies. The following lessons learned and best practices, derived from real-world case studies and research, can serve as valuable guidelines for farmers, agronomists, and stakeholders in the agricultural industry.

a. **Data Quality:** Accurate and reliable data are crucial for precision agriculture. Investing in high-quality sensors, calibration and data management systems is essential for effective decision-making.
b. **Site-Specific Management:** Precision agriculture techniques should be tailored to the specific conditions and variability of each field or production area. A one-size-fits-all approach may not yield optimal results.
c. **Training and Extension Services:** Providing adequate training and technical support to farmers and agricultural workers is critical for successful adoption and implementation of precision agriculture practices.
d. **Economic and Environmental Considerations:** Precision agriculture should balance economic objectives with environmental sustainability. Reducing inputs while maintaining or improving yields can lead to cost savings and reduced environmental impact.
e. **Collaboration and Data Sharing:** Fostering collaboration among farmers, researchers, and technology providers can accelerate the development and adoption of precision agriculture solutions. Data sharing and open-source platforms can contribute to knowledge dissemination and innovation.

1.9 Future Directions in Precision agriculture

The field of precision agriculture is continuously evolving, driven by rapid technological advancements and the growing demand for sustainable and efficient farming practices. As we look towards the future, new technologies and innovative approaches hold the potential to revolutionize the way we cultivate crops and manage agricultural resources. This section explores the emerging trends and technologies that are shaping the future of precision agriculture, paving the way for increased productivity, resource optimization, and environmental stewardship.

1.9.1 Emerging technologies and trends

The integration of cutting-edge technologies into precision agriculture is fueling a paradigm shift in the way we approach farming. These emerging technologies and trends are not only enhancing existing practices but also opening up new avenues for innovation and transformative solutions. The following subsections provide an overview of some of the most promising developments in this domain:

1.9.1.1 Internet of Things IoT and sensor networks

a) Widespread deployment of low-cost, energy-efficient sensor networks for real-time monitoring of environmental conditions, soil health, and crop growth.
b) Integration of IoT devices and wireless sensor networks for automated data collection and remote control of agricultural operations [13].

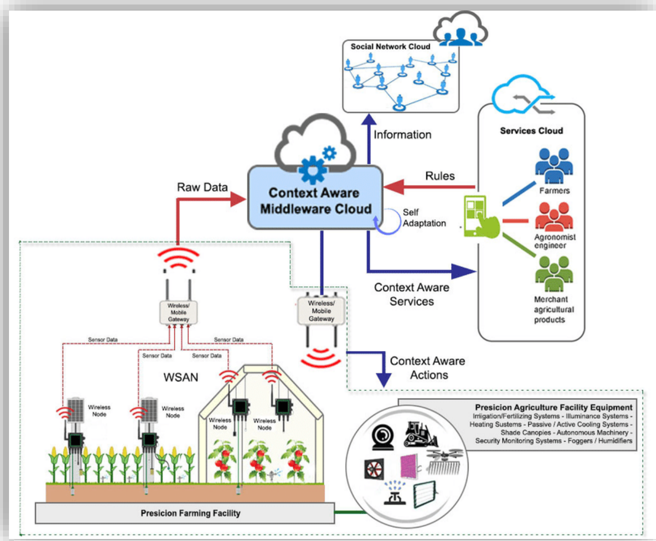


Figure 3: Example of using IOT and sensor networks

1.9.1.2 Big data analytics and machine learning

a. Leveraging big data techniques to process and analyze vast amounts of agricultural data from multiple sources (sensors, imagery, weather, yield data) [2].
b. Application of machine learning algorithms and predictive modeling for yield forecasting, pest and disease detection, and optimized resource management [4].

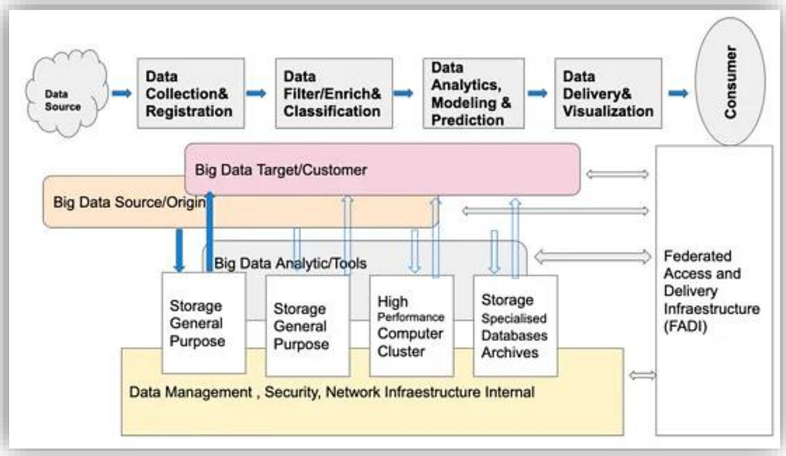


Figure 4: General architecture of Big Data

1.9.2 Potential improvements and innovations

Interoperability and data standardization

a. Development of standardized data formats and communication protocols to enable seamless integration and interoperability among different precision agriculture technologies and systems [14].

b. Adoption of open data standards and application programming interfaces (APIs) to facilitate data sharing and collaboration among stakeholders.

Sustainable and regenerative agriculture practices

a. Integration of precision agriculture techniques with sustainable and regenerative agricultural practices, such as conservation tillage, cover cropping, and integrated pest management.

b. Development of decision support systems that consider environmental factors, soil health, and long-term sustainability in addition to productivity and profitability [5].

1.9.3 Recommender Systems for Precision Agriculture

Recommender systems represent an emerging and promising application of artificial intelligence (AI) and machine learning (ML) techniques in the domain of precision agriculture. These intelligent systems leverage advanced algorithms, such as collaborative filtering, content-based filtering, and hybrid approaches, to provide personalized recommendations to farmers regarding crop selection, cultivation practices, and resource management [15].

By leveraging diverse data sources, including remote sensing data, environmental sensor networks, historical yield records, market intelligence, and farmer preferences, recommender systems can employ data mining and knowledge discovery techniques to analyze complex patterns and relationships. These systems can then suggest the most suitable crops.

The integration of recommender systems in precision agriculture holds the potential to revolutionize decision-making processes, enhancing productivity, minimizing resource wastage, and promoting sustainable farming practices. These systems can adapt to dynamic environmental changes, market fluctuations, and evolving farmer needs through online learning and adaptive modeling techniques, continuously refining their recommendations based on incoming data and user feedback.

In Chapter 2, we will explore the intricacies of recommender systems for precision agriculture, including underlying technologies, data sources, modeling approaches, evaluation metrics, and real-world applications. This comprehensive exploration will shed light on the potential of these innovative AI-driven systems to drive the future of sustainable and efficient agricultural practices.

1.10 Conclusion

This chapter emphasized the numerous benefits of precision agriculture, including improved crop yields and quality, reduced input costs, enhanced environmental sustainability, and increased efficiency in farm operations. However, it also acknowledged the challenges in implementing precision agriculture, such as high initial costs, lack of standardization and interoperability, data privacy and security concerns, and the limited availability of skilled personnel.

The successful case studies presented in the chapter showcased the real-world impact and potential of precision agriculture techniques from around the world, providing valuable lessons learned and best practices for farmers, agronomists, and stakeholders in the agricultural industry.

Looking towards the future, the conclusion highlighted the potential of emerging technologies and trends, such as advancements in sensor technologies, robotics, artificial intelligence, and big data analytics, to further revolutionize agricultural practices, enabling even more precise and efficient management of crops and resources.

In essence, the conclusion underscored the significance of precision agriculture as a sustainable and productive approach to agriculture, while recognizing the challenges and emphasizing the need for continuous innovation, collaboration, and adoption of cutting-edge technologies to drive the future of precision agriculture practices.

Chapter 2
Crop recommendation systems

2.1 Introduction to Recommendation Systems

In the era of precision agriculture, where data-driven decision-making is revolutionizing farming practices, crop recommendation systems have emerged as indispensable tools for optimizing crop production. These systems leverage advanced technologies and computational algorithms to provide farmers with tailored recommendations for crop selection, fertilization practices, pest management, and irrigation strategies. By harnessing the power of big data analytics, artificial intelligence, and machine learning, crop recommendation systems offer a proactive approach to agricultural management, enabling farmers to enhance productivity, minimize resource wastage, and mitigate environmental impact.

2.1.1. Definition and Purpose

Recommendation systems can be defined as software applications that employ data mining, pattern recognition, and predictive modeling techniques to analyze user data (such as browsing history, purchase records, ratings, or explicit preferences) and identify relevant items from a vast collection of possibilities.
The primary purpose of recommendation systems is to enhance the user experience by providing personalized recommendations that align with their interests and needs. By leveraging machine learning and data analysis techniques, these systems can identify patterns and make accurate predictions about items that users are likely to find valuable or enjoyable [16].

2.1.2. Importance and Applications

Recommendation systems have become increasingly important in various domains due to their ability to navigate and filter vast amounts of information, products, or services [4]. Some key areas where recommendation systems are widely used include:

a. **E-commerce:** Online retailers employ recommendation systems to suggest products to customers based on their browsing and purchase history, enabling personalized shopping experiences and increasing sales [17].
b. **Media and entertainment:** Streaming platforms, such as Netflix and Spotify, utilize recommendation systems to suggest movies, TV shows, or music based on user preferences and viewing, listening habits.
c. **Social networks:** Social media platforms like Facebook and Twitter leverage recommendation systems to suggest new connections, content, or advertisements based on user profiles and interactions.
d. **Information retrieval:** Search engines and digital libraries use recommendation systems to personalize search results and suggest relevant documents or resources based on user queries and browsing patterns.
e. **Healthcare:** Recommendation systems are applied in healthcare to suggest personalized treatment plans, medical diagnoses, or lifestyle recommendations based on patient data and medical knowledge [18].

2.2 Types of Recommendation Systems

Recommendation systems employ various techniques to generate personalized recommendations for users. The choice of technique depends on the available data, the domain, and the specific requirements of the recommendation task.

2.2.1. Content-Based Filtering

Content-based filtering is a recommendation technique that suggests items based on the similarity between the content of the items and the user's preferences. In the context of crop recommendation systems, the system would analyze the characteristics of the crops (such as soil requirements, climate conditions, yield, and other relevant features) and match them with the user's preferences or farm conditions [19] [20].

The core idea behind content-based recommendation systems is to represent items as a set of features or attributes, which can be extracted from various sources such as textual descriptions, metadata, or other item-specific information. The system then learns a user's preferences by analyzing the features of the items they have liked or consumed and tries to find new items with similar characteristics.

One common approach in content-based recommendation systems is to use techniques from information retrieval, such as:

a) Feature Engineering

Feature engineering is a preprocessing step in supervised machine learning and statistical modeling which transforms raw data into a more effective set of inputs. Each input comprises several attributes, known as features. By providing models with relevant information, feature engineering significantly enhances their predictive accuracy and decision-making capability.

Feature Engineering Processes: Feature engineering consists of various processes:

1. **Feature creation:** Creating features involves creating new variables which will be most helpful for our model. This can be adding or removing some features. As we saw above, the cost per sq. ft column was a feature creation.
2. **Transformations:** Feature transformation is simply a function that transforms features from one representation to another. The goal here is to plot and visualize data. If something isn't adding up with the new features, we can reduce the number of features used, speed up training or increase the accuracy of a certain model.
3. **Feature extraction:** Feature extraction is the process of extracting features from a data set to identify useful information. Without distorting the original relationships or significant information, this compresses the amount of data into manageable quantities for algorithms to process.
4. **Exploratory data analysis:** Exploratory data analysis (EDA) is a powerful and simple tool that can be used to improve your understanding of your data, by exploring its properties. The technique is often applied when the goal is to create new hypotheses or find patterns in the data. It's often used on large amounts of qualitative or quantitative data that haven't been analyzed before.
5. **Benchmark:** A benchmark model is the most user-friendly, dependable, transparent and interpretable model against which you can measure your own. It's a good idea to run test data sets to see if your new machine learning model outperforms a recognized benchmark. These benchmarks are often used as measures for comparing the performance between different machine learning models like neural networks and support vector machines, linear and non-linear classifiers or different approaches like bagging and boosting.

b) Feature Selection:

Feature selection is the process of selecting a subset of relevant features (variables, predictors) for use in model construction.

Feature selection techniques are used for several reasons:

- simplification of models to make them easier to interpret by researchers/users,
- shorter training times,
- to avoid the curse of dimensionality,
- improve data's compatibility with a learning model class,
- encode inherent symmetries present in the input space.

The central premise when using a feature selection technique is that the data contains some features that are either redundant or irrelevant, and can thus be removed without incurring much loss of information. Redundant and irrelevant are two distinct notions, since one relevant feature may be redundant in the presence of another relevant feature with which it is strongly correlated.

Feature extraction creates new features from functions of the original features, whereas feature selection returns a subset of the features. Feature selection techniques are often used in domains where there are many features and comparatively few samples (or data points).

The content-based filtering process typically involves the following steps [13]:

1. **Item Representation**: The crops are represented as a set of features or attributes that describe their characteristics.
2. **User Profile Building**: The system builds a user profile based on the user's preferences, historical data, or explicit feedback on previously recommended crops.
3. **Similarity Computation**: The system calculates the similarity between the user's profile and the available crop options, often using a similarity metric such as cosine similarity, Euclidean distance, or Pearson correlation coefficient.
4. **Recommendation Generation**: The system recommends the crops with the highest similarity scores to the user's profile.

One common approach in content-based filtering is the Space Vector Model (SVM), where crops and user profiles are represented as vectors in a multidimensional feature space.

2.2.1.1 Space Vector Model (SVM)

The Space Vector Model represents textual documents or items (such as crops in a recommendation system) as vectors in a high-dimensional vector space. Each dimension in this vector space corresponds to a unique term or feature extracted from the documents or items [19].

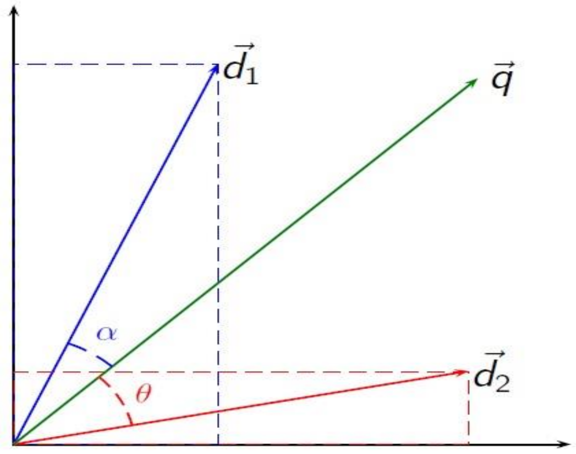


Figure 5: example of a document query on the Space Vector Model

Here's how the Space Vector Model works:

a. **Text Preprocessing**: The first step is to preprocess the textual data (e.g., crop descriptions, farmer requirements) by performing operations such as tokenization, stop-word removal, and stemming or lemmatization [20].
b. **Term Weighting**: Each unique term or feature is assigned a weight to reflect its importance in representing the document or item. Common term weighting schemes include:
c. **Vector Representation**: Each document or item is represented as a vector in the term-space, where each dimension corresponds to a unique term, and the value in that dimension is the weight of that term for the given document or item.
d. **Similarity Computation**: To measure the similarity between two documents or items.

Challenges and Considerations

Developing an effective crop recommendation system involves addressing several challenges and considerations [2].

1. **Data Availability and Quality**: Lack of sufficient, reliable, and consistent data on crop characteristics, soil conditions, climate patterns, and farmer preferences can affect the accuracy of recommendations.
2. **High-Dimensional and Heterogeneous Data**: Crop recommendation systems involve high-dimensional and heterogeneous data (categorical, numerical, textual), requiring advanced techniques like dimensionality reduction and feature engineering.
3. **Non-Linear Relationships and Complex Interactions**: The relationships between crop suitability and various factors can be non-linear and involve complex interactions, necessitating advanced machine learning techniques.
4. **Dynamic and Evolving Conditions**: Environmental conditions, market demands, and farmer preferences can change over time, requiring the recommendation system to adapt and continuously learn.
5. **Interpretability and Explainability**: Ensuring interpretability and explainability of recommendations is crucial for building trust and adoption among farmers.
6. **Scalability and Computational Efficiency**: As the system grows, computational efficiency and scalability become essential for handling large-scale data and real-time recommendations.

7. **User Acceptance and Adoption**: Overcoming user hesitancy and building trust in the system through transparency, domain expert involvement, and feedback mechanisms.
8. **Integration with Existing Systems**: Ensuring interoperability and seamless integration with existing agricultural systems and workflows.
9. **Ethical and Regulatory Considerations**: Adhering to ethical principles, data protection laws, and agricultural policies.

2.2.1.2 Naive Bayes Classifiers (NB)

Naive Bayes classifiers can be used to predict the probability that a user will like a particular item (crop) based on its features and the user's past preferences. They work well with content-based filtering and can handle high-dimensional data efficiently **[6]**.

2.2.1.3 Decision Trees and Random Forests

These tree-based algorithms can be used to model the relationship between crop features (e.g., soil conditions, climate) and crop suitability, enabling content-based recommendations.

2.2.1.4 Support Vector Machines (SVMs)

SVMs are powerful algorithms for classification and regression tasks, particularly well-suited for high-dimensional data like crop features. They can be used in content-based filtering to predict crop suitability based on farm conditions and crop characteristics.

2.2.1.5 Artificial Neural Networks (ANNs)

Deep learning techniques like feedforward neural networks, convolutional neural networks (CNNs), or recurrent neural networks (RNNs) can be employed to learn complex patterns in crop data and user preferences for content-based recommendations.

2.2.1.6 XGBoost (Extreme Gradient Boosting)

XGBoost is an optimized distributed gradient boosting library designed to be highly efficient, flexible, and portable. It implements machine learning algorithms under the Gradient Boosting framework, which builds an ensemble of decision trees iteratively, with each tree attempting to correct the errors made by the previous trees. XGBoost has gained popularity due to its excellent performance in various machine learning tasks, including recommendation systems.

2.2.1.7 LightGBM (Light Gradient Boosting Machine)

LightGBM is a gradient boosting framework that uses tree-based learning algorithms. It is designed to be distributed and efficient, with a focus on reducing the time and memory consumption during training [21]. LightGBM has several advantages that make it an attractive choice for recommendation systems:

a) Leaf-wise tree growth: Unlike XGBoost, which grows trees level-wise, LightGBM grows trees leaf-wise. It chooses the leaf with the maximum delta loss to grow, resulting in a more complex and deeper tree structure. This approach can lead to higher accuracy and faster convergence [21].

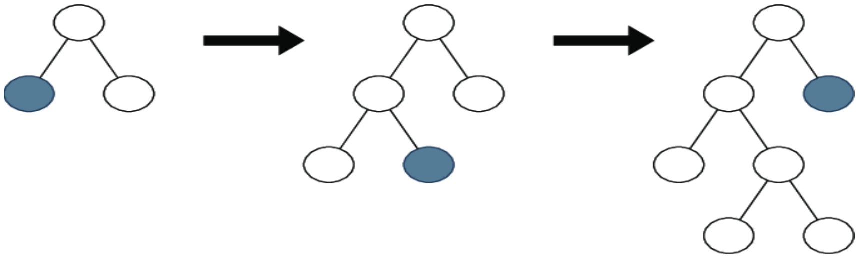


Figure 6: Leaf-wise tree growth in LightGBM

b) **Histogram-based algorithm**: LightGBM uses a histogram-based algorithm for splitting the data. Instead of sorting the data and finding the best split point, it buckets continuous feature values into discrete bins and uses these bins to construct the trees. This technique significantly reduces the time complexity of training and speeds up the process.

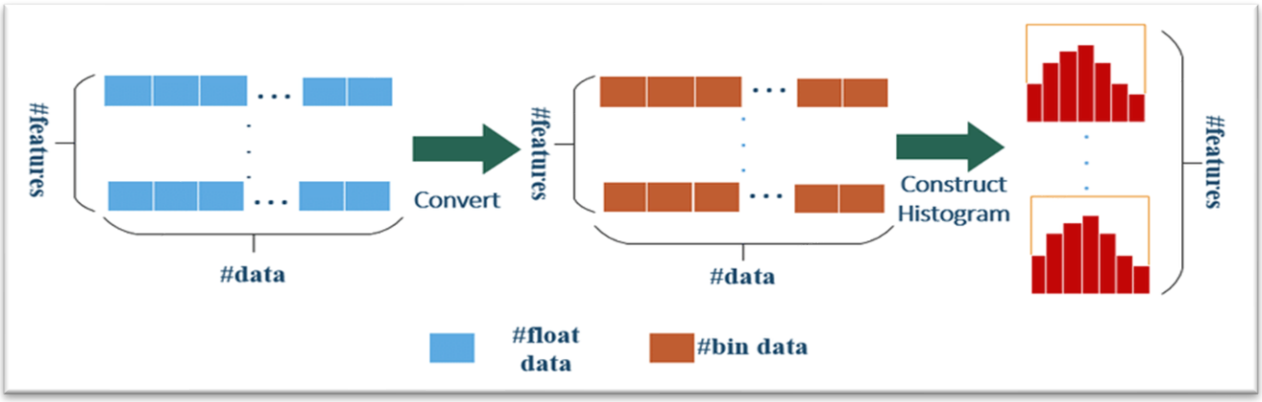


Figure 7: Histogram algorithm of LightGBM

c) **Gradient-based one-side sampling (GOSS)**: LightGBM employs a novel technique called GOSS to down-sample the data instances. It keeps all the instances with large gradients and performs random sampling on the instances with small gradients. This approach retains the data instances that contribute more to the information gain, reducing the data size without compromising much on accuracy.

d) **Exclusive Feature Bundling (EFB):** LightGBM introduces EFB, which bundles mutually exclusive features into a single feature. This reduces the feature dimensionality and speeds up the training process without losing information.

e) **Categorical feature support**: LightGBM has native support for categorical features. It can automatically convert categorical variables into numerical features using one-hot encoding or embedding-based approaches, eliminating the need for manual preprocessing.

f) **Parallel and distributed learning**: LightGBM supports parallel and distributed learning, allowing it to scale to large datasets and take advantage of multi-core CPUs and distributed computing environments.

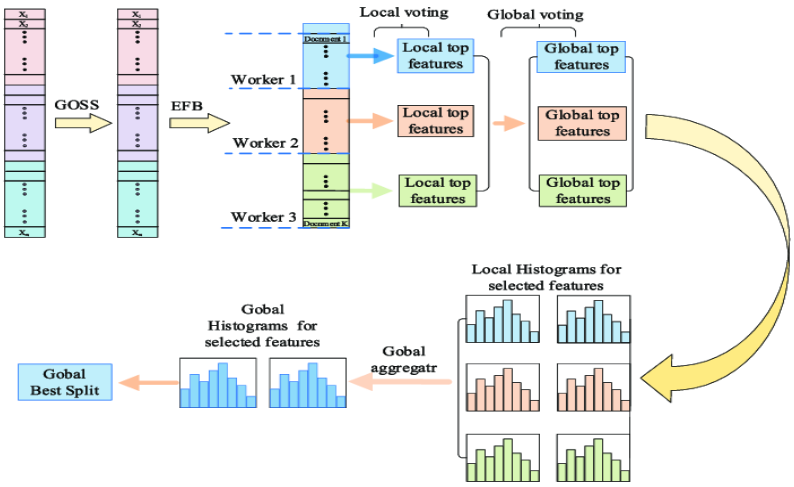


Figure 8: Schematic illustration of the LightGBM model

By leveraging the power of LightGBM and its advanced features, you can build accurate and efficient content-based recommender systems that provide personalized recommendations to users based on their preferences and item characteristics [21].

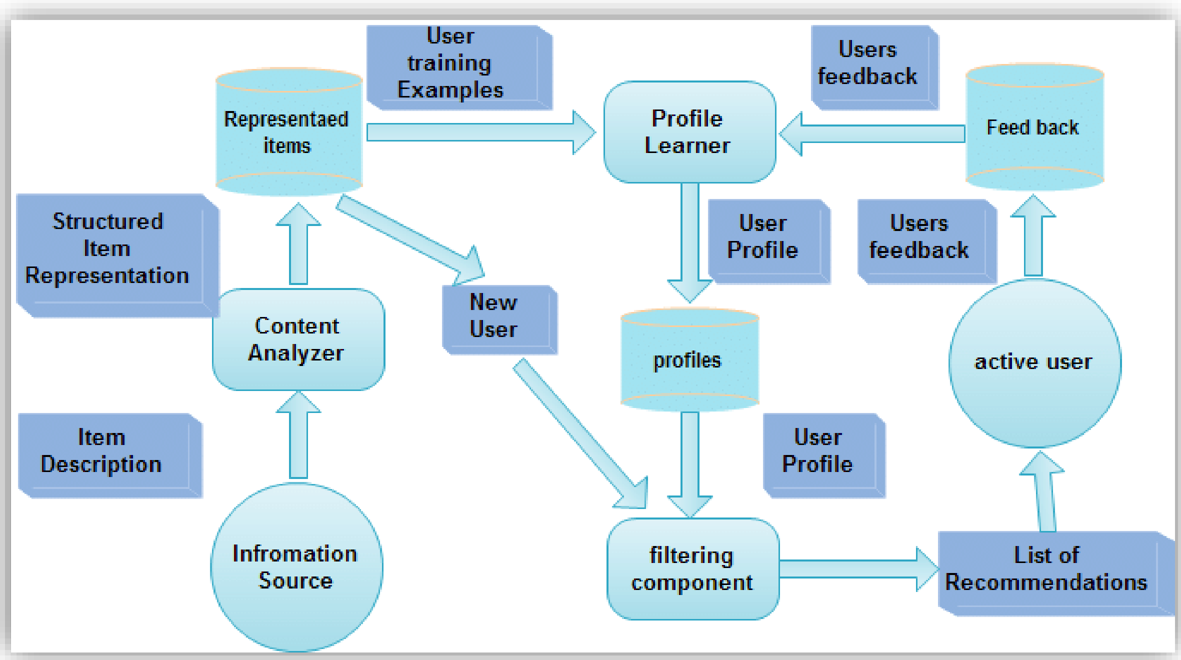


Figure 9: Content-Based Filtering' architecture

2.2.2 Collaborative Filtering:

Collaborative filtering is a recommendation technique that suggests items based on the preferences of similar users or the similarities between items. In the context of crop recommendation systems, the system would analyze the preferences and historical data of other farmers with similar farm conditions or crop preferences to generate recommendations [22].

The collaborative filtering process typically involves the following steps:

1. **Data Collection**: The system collects data on user preferences, ratings, or interactions with crops.
2. **Similarity Computation**: The system calculates the similarity between users (user-based collaborative filtering) or between items (item-based collaborative filtering) using a similarity metric such as cosine similarity, Pearson correlation coefficient, or adjusted cosine similarity.
3. **Neighborhood Formation**: For user-based collaborative filtering, the system identifies a set of similar users (neighbors) for each target user. For item-based collaborative filtering, the system identifies a set of similar items (neighbors) for each target item.
4. **Recommendation Generation**: The system recommends crops that are highly rated or preferred by the user's neighbors (user-based) or crops that are similar to the ones the user has previously liked (item-based).

One common approach in user-based collaborative filtering is the k-Nearest Neighbors (KNN) algorithm, where the system finds the k most similar users to the target user and recommends crops that are highly rated by those neighbors.

2.2.2.1 K-Nearest Neighbors (KNN)

KNN is a widely used algorithm in user-based collaborative filtering. It finds the k most similar users to the target user based on their rating or preference history and recommends items that these similar users have liked [23].

In the context of our recommendation system here how it works:

$$r\_u,i \;=\; (\textstyle\sum\_\{v \;\in\; N(u)\}\, sim(u,v) \;\times\; r\_v,i) \,/\, (\textstyle\sum\_\{v \;\in\; N(u)\}\, sim(u,v))$$

Here, **r_u,i** is the predicted rating for user **u** and crop **i, N(u)** is the set of neighbors for user **u, sim(u, v)** is the similarity between users **u** and **v**, and **r_v,i** is the rating given by neighbor **v** for crop **i** [23].

In item-based collaborative filtering, the system computes the similarity between crops and recommends crops that are similar to the ones the user has previously liked.

$$r\_u,i \;=\; (\textstyle\sum\_\{j \;\in\; I(u)\}\, sim(i,j) \;\times\; r\_u,j) \,/\, (\textstyle\sum\_\{j \;\in\; I(u)\}\, sim(i,j))$$

Here, **r_u,i** is the predicted rating for user **u** and crop **i, I(u)** is the set of crops the user **u** has rated, **sim(i, j)** is the similarity between crops **i** and `j`, and **r_u,j** is the rating given by the user `u` for crop **j** [23].

There are some other popular algorithms and techniques used in content-based filtering systems include:

2.2.2.2 Matrix Factorization

Matrix factorization techniques, such as Singular Value Decomposition (SVD) and Non-Negative Matrix Factorization (NMF), are commonly used in collaborative filtering to decompose the user-item rating matrix into lower-dimensional user and item latent factor matrices. These latent factors capture the underlying preferences and characteristics of users and items, enabling more accurate recommendations [24].

2.2.2.3 Alternating Least Squares (ALS)

ALS is a popular algorithm for matrix factorization in collaborative filtering. It alternately fixes the user latent factors and item latent factors and solves for the other using least squares optimization until convergence.

2.2.2.4 Probabilistic Matrix Factorization (PMF)

PMF is a probabilistic approach to matrix factorization that models the user-item ratings as a product of user and item latent factors, with Gaussian noise added. It learns the latent factors by maximizing the log-likelihood of the observed ratings [25].

2.2.2.5 Bayesian Personalized Ranking (BPR)

BPR is a pairwise ranking approach to collaborative filtering that learns user and item latent factors by optimizing the pairwise ranking of items for each user. It aims to rank the items that a user prefers higher than the items they do not prefer [26].

Challenges and Considerations

Collaborative filtering can suffer from the "**cold start**" problem, where recommendations cannot be generated for new users or items with no or limited data. Techniques like matrix factorization, ensemble methods, or hybrid approaches can be used to mitigate this issue.

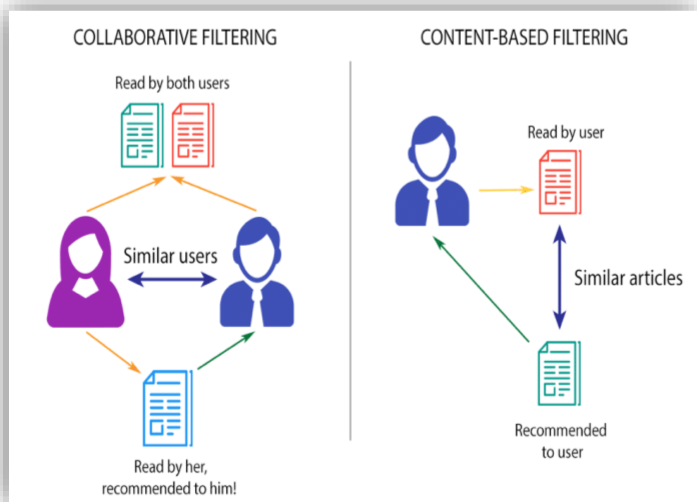


Figure 10: Differences between collaborative and content-based filtering

2.2.3. Hybrid Recommendation Systems

Hybrid systems can be implemented through various strategies, such as weighted hybridization (linear combination of scores from different techniques), switching hybridization (switching between techniques based on specific conditions), or feature combination (incorporating features from different sources into a single recommendation algorithm) [27].

There are several ways to combine recommendation techniques in a hybrid system:

1. **Weighted Hybrid**: The recommendations from different techniques are combined with weights assigned to each technique based on their importance or performance.
2. **Switching Hybrid**: The system switches between different recommendation techniques based on certain conditions or criteria, such as the availability of data or the user's preferences.
3. **Mixed Hybrid**: Recommendations from different techniques are presented together, allowing the user to choose from a diverse set of options.
4. **Feature Combination Hybrid**: The features or data used by different recommendation techniques are combined and used as input to a single recommendation algorithm.
5. **Cascade Hybrid**: One recommendation technique is used to generate a candidate set of recommendations, which is then refined or re-ranked using another technique.

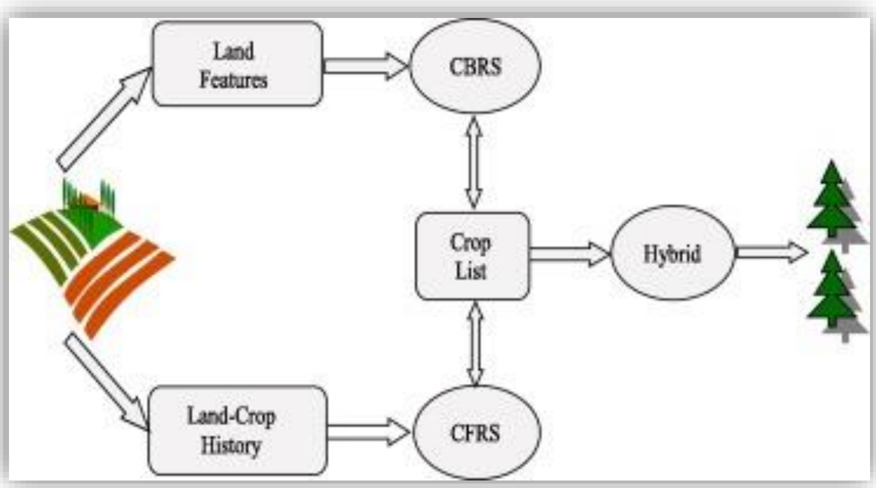


Figure 11: Example of Hybrid system architecture

2.2.4. Other types of Recommendation Systems:

In addition to the three main types mentioned above, there are several other recommendation system approaches, including:

a) **Knowledge-based systems:** These systems leverage domain knowledge, user preferences, and constraints to provide recommendations. They typically employ case-based reasoning, constraint-based reasoning, or rule-based techniques.

b) Demographic-based systems: These systems recommend items based on demographic attributes of users, such as age, gender, location, or occupation.

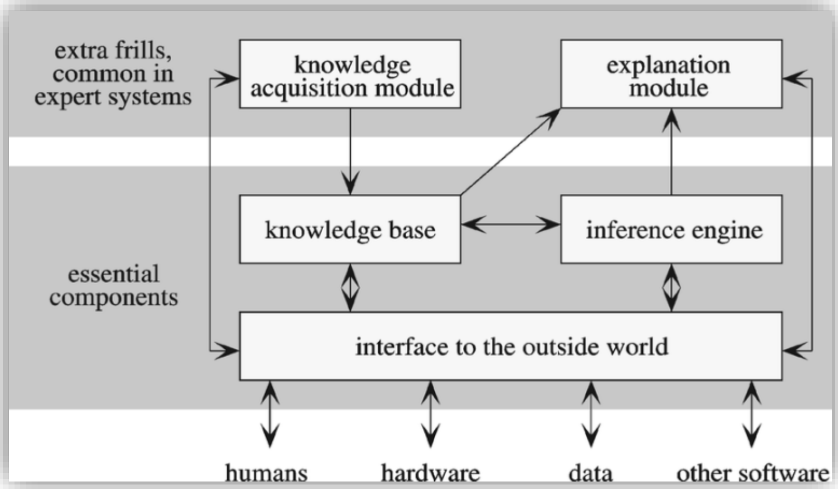


Figure 12: Flow chart for the demographic-based recommendation system.

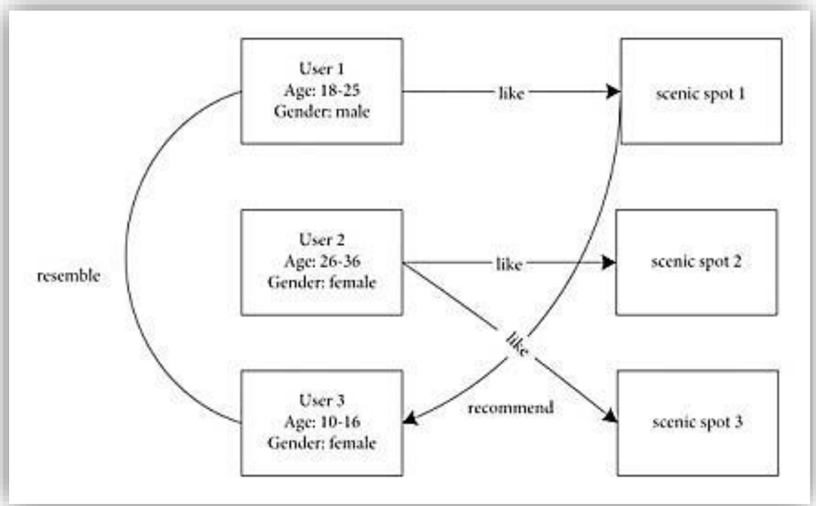


Figure 13: The main components of a knowledge-based system.

c) **Context-aware systems**: These systems incorporate contextual information, such as location, time, or device, to provide more relevant recommendations based on the user's current context.

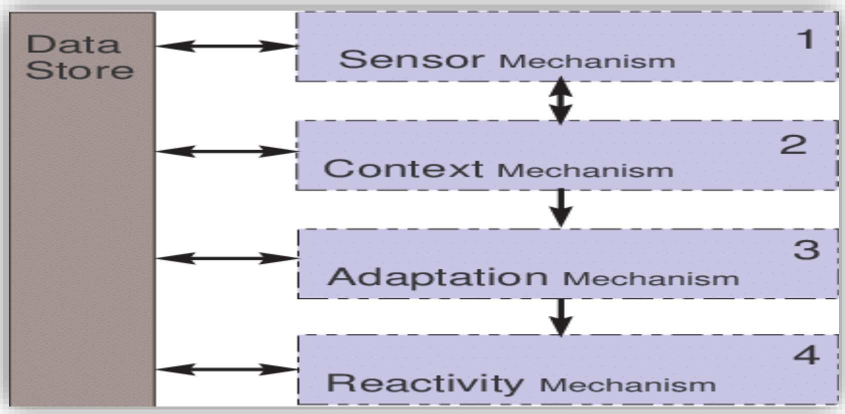


Figure 14: Architecture of Context-Aware Systems

2.3 Evaluation Metrics for Recommendation Systems:

Evaluating the performance of recommendation systems is crucial to assess their effectiveness and identify areas for improvement. Various metrics have been developed to measure different aspects of recommendation system performance, such as accuracy, ranking quality, and diversity.

2.3.1 Accuracy Metrics:

Accuracy metrics are often calculated based on the confusion matrix, which represents the true positive, false positive, true negative, and false negative predictions made by the recommendation system. Some commonly used accuracy metrics include:

I. **Precision:** The fraction of recommended items that are relevant to the user.

$$\text{precision} \stackrel{\text{def}}{=} \frac{\text{TP}}{\text{TP} + \text{FP}}$$

where:

- **TP (True Positives)**: The number of times the system correctly predicted the best crop for the given conditions.
- **FP (False Positives)**: The number of times the system incorrectly predicted a crop that was not the best choice.

II. **Recall**: The fraction of relevant items that are successfully recommended.

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

where:

- **TP (True Positives)**: The number of times the system correctly predicted the best crop for the given conditions.
- **FN (False Negatives)**: The number of times the system failed to recommend the best crop when it was the best choice.

III. **F1-score**: The harmonic means of precision and recall, providing a balanced measure of accuracy.

$$\mathrm{F1-score} = \frac{\mathrm{TP}}{\mathrm{TP} + \frac{1}{2}(\mathrm{FP} + \mathrm{FN})}$$

where:

- **TP (True Positives)**: The number of correct predictions of the best crop.
- **FP (False Positives)**: The number of incorrect predictions where the recommended crop was not the best choice.
- **FN (False Negatives)**: The number of missed best crop predictions.

2.3.2 Ranking Metrics

Ranking metrics evaluate the ability of a recommendation system to rank relevant items higher than irrelevant ones. These metrics are particularly important when the recommendation system presents a ranked list of items to the user. Some commonly used ranking metrics include:

a) **Mean Average Precision (MAP):** Measures the average precision across multiple recall levels, considering the ranking order of the recommended items [23].
b) **Normalized Discounted Cumulative Gain (NDCG):** Assigns higher weights to relevant items ranked higher in the recommendation list, emphasizing the importance of ranking quality [23].

2.3.3 Other Metrics

In addition to accuracy and ranking metrics, there are several other metrics that evaluate different aspects of recommendation

a. **Coverage:** Measures the proportion of items in the dataset that the recommendation system can potentially recommend, reflecting the system's ability to provide recommendations for a diverse range of items [23].
b. **Diversity:** Measures the dissimilarity or variety of items in the recommendation list, ensuring that the recommendations are not overly narrow or redundant [23].

c. **Novelty:** Measures the ability of the recommendation system to suggest items that are novel or unexpected to the user, promoting serendipitous discoveries [23].

2.4 Overview of Crop Recommendation Systems

Crop recommendation system is a crucial component of precision agriculture, leveraging advanced technologies and data-driven approaches to assist farmers in making informed decisions about crop selection and cultivation practices. These systems aim to optimize crop yields, minimize resource utilization, and enhance overall agricultural productivity.

2.4.1 Importance of Crop Recommendation Systems in Agriculture

Crop recommendation systems play a vital role in modern agriculture by addressing the complex challenges faced by farmers. By analyzing a multitude of factors, these systems provide valuable insights and recommendations tailored to specific farm conditions and market demands. The importance of crop recommendation systems can be highlighted through the following aspects:

1. Improved crop productivity and yield optimization
2. Efficient resource allocation and input management
3. Adaptation to changing environmental conditions
4. Increased profitability and economic sustainability
5. Promotion of sustainable agricultural practices

2.4.2 Factors Influencing Crop Selection

As discussed earlier, factors such as soil conditions (e.g., soil type, pH level, nutrient content), climate (e.g., temperature, precipitation, humidity), market demand (e.g., consumer preferences, price trends), and farmer preferences and constraints play a crucial role in crop selection. Your crop recommendation system should consider these factors and incorporate relevant data to generate accurate and personalized recommendations [28].

a. Soil Conditions

- **Soil type:** Different crops thrive in different soil types, such as loamy, clayey, or sandy soils. For example, potatoes prefer well-drained sandy soils, while rice grows well in clay soils that retain water.
- **pH level:** Crops have specific pH ranges for optimal growth. For instance, blueberries require acidic soils (pH 4.5-5.5), while asparagus prefers slightly alkaline soils (pH 6.5-7.5).
- **Nutrient content:** Soil nutrient levels, such as nitrogen, phosphorus, and potassium, impact crop growth. Corn, for example, requires high nitrogen levels, while legumes like soybeans can fix atmospheric nitrogen.

b. Climate

- **Temperature:** Crops have specific temperature requirements for germination, growth, and maturity. For example, tomatoes require warm temperatures (20-30°C), while spinach grows best in cool conditions (15-18°C).
- **Precipitation:** Water availability through rainfall or irrigation is crucial for crop growth. Crops like rice require high water levels, while sorghum is drought-tolerant.

- **Sunlight exposure:** Different crops have varying sunlight requirements. Sunflowers, for instance, require full sun exposure, while lettuce can tolerate partial shade.

c. Market Demand

- **Consumer preferences:** Market demand is influenced by consumer preferences, which can vary based on factors like taste, nutrition, and cultural significance. For example, quinoa has gained popularity due to its nutritional value and gluten-free properties.
- **Price trends:** Crop prices fluctuate based on supply and demand dynamics. Farmers may choose to grow crops that have a higher market value or are in high demand.
- **Export potential:** Some crops are grown specifically for export markets. For example, cocoa is a major export crop for countries like Ivory Coast and Ghana.

d. Farmer Preferences and Constraints

- **Crop rotation:** Farmers may follow specific crop rotation patterns to maintain soil health, control pests, and optimize nutrient management. For example, a common rotation is corn-soybean-wheat.
- **Resource availability:** Farmers' crop choices are influenced by the availability of resources such as water, fertilizers, and labor. For instance, drip irrigation systems can enable the cultivation of crops with higher water requirements in water-scarce regions.
- **Local knowledge and traditions:** Farmers may have traditional knowledge or cultural preferences for certain crops. For example, indigenous communities may prioritize the cultivation of traditional varieties with cultural significance.

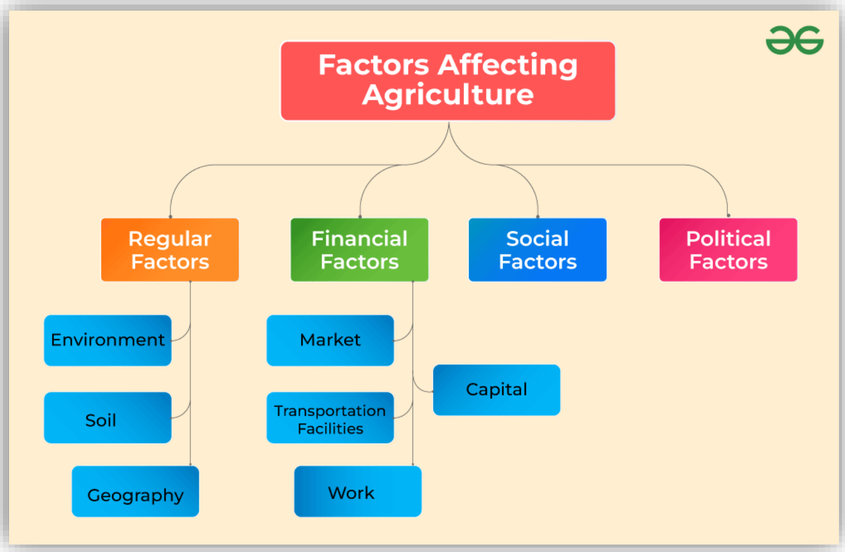


Figure 15: Factors influencing Agriculture

2.5 Existing Approaches to Crop Recommendation Systems

There are various approaches and methodologies employed in the development of crop recommendation systems, each with its own strengths and limitations. This section explores three main approaches: rule-based systems, machine learning-based approaches, and hybrid and multi-criteria decision-making models [29].

2.5.1 Rule-based Systems

Rule-based systems, also known as expert systems or knowledge-based systems, rely on a set of predefined rules and decision criteria to provide crop recommendations. These systems typically incorporate domain knowledge and expertise from agricultural experts, extension services, and historical data [29].

The rules are designed to capture the relationships between various factors, such as soil properties, climatic conditions, and crop characteristics, and provide recommendations based on these predefined conditions.

Example: A rule-based system for recommending rice varieties in a specific region might have rules like:

```
IF

(soil_type = 'clay_loam' AND rainfall = 'high' AND temperature = 'warm')

THEN

recommend_crop = 'Rice_Variety_A'

ELSE IF

(soil_type = 'sandy' AND rainfall = 'low' AND temperature = 'hot')

THEN recommend_crop = 'Rice_Variety_B'
```

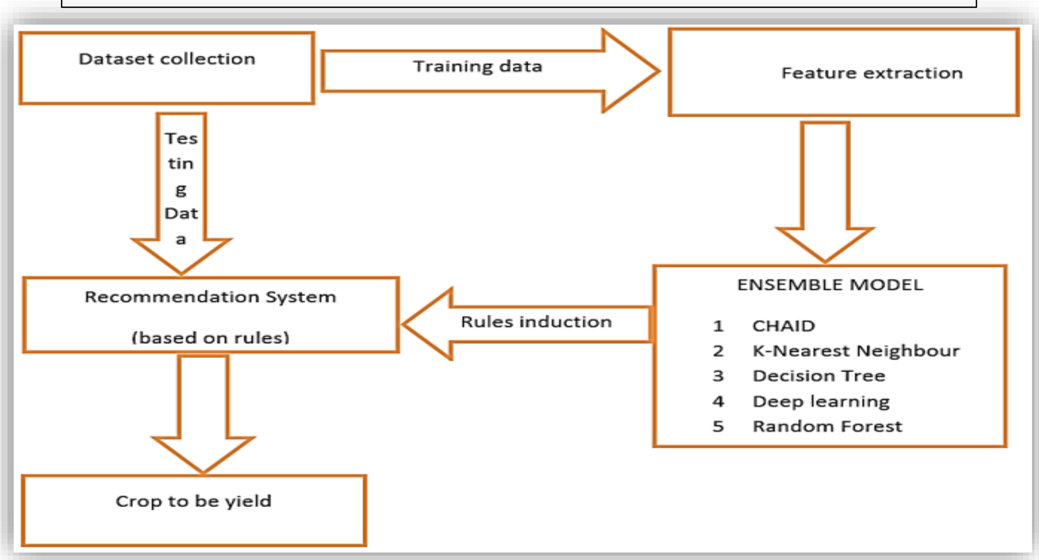


Figure 16: Proposed System Architecture

2.5.2 Machine Learning-based Approaches

Machine learning-based approaches leverage advanced algorithms and statistical techniques to learn patterns and relationships from large datasets. These approaches can be categorized into supervised learning, unsupervised learning, and deep learning techniques [29].

a) **Supervised learning algorithms:** such as decision trees, random forests, and support vector machines, are trained on labeled data to predict the most suitable crop for a given set of input features.
b) **Unsupervised learning techniques:** like clustering and dimensionality reduction, can be used to identify patterns and groupings within the data, potentially revealing insights for crop selection.
c) **Deep learning models:** such as convolutional neural networks and recurrent neural networks, have shown promising results in processing complex data, such as remote sensing imagery and time-series data.

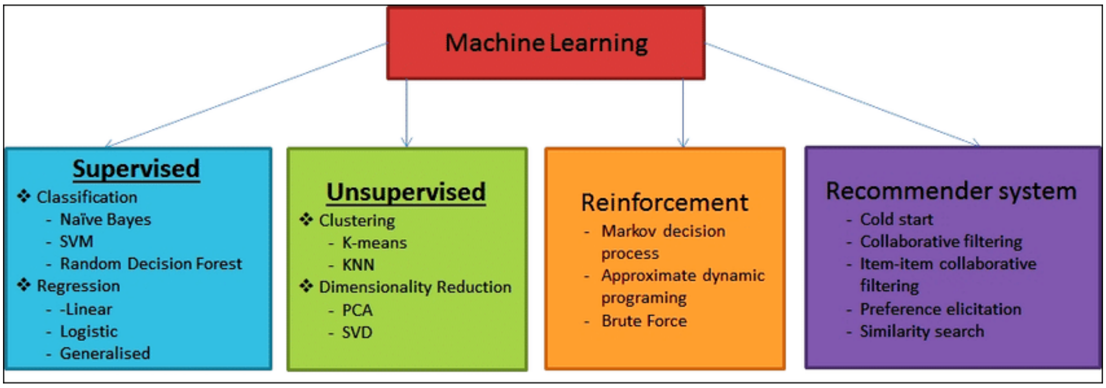
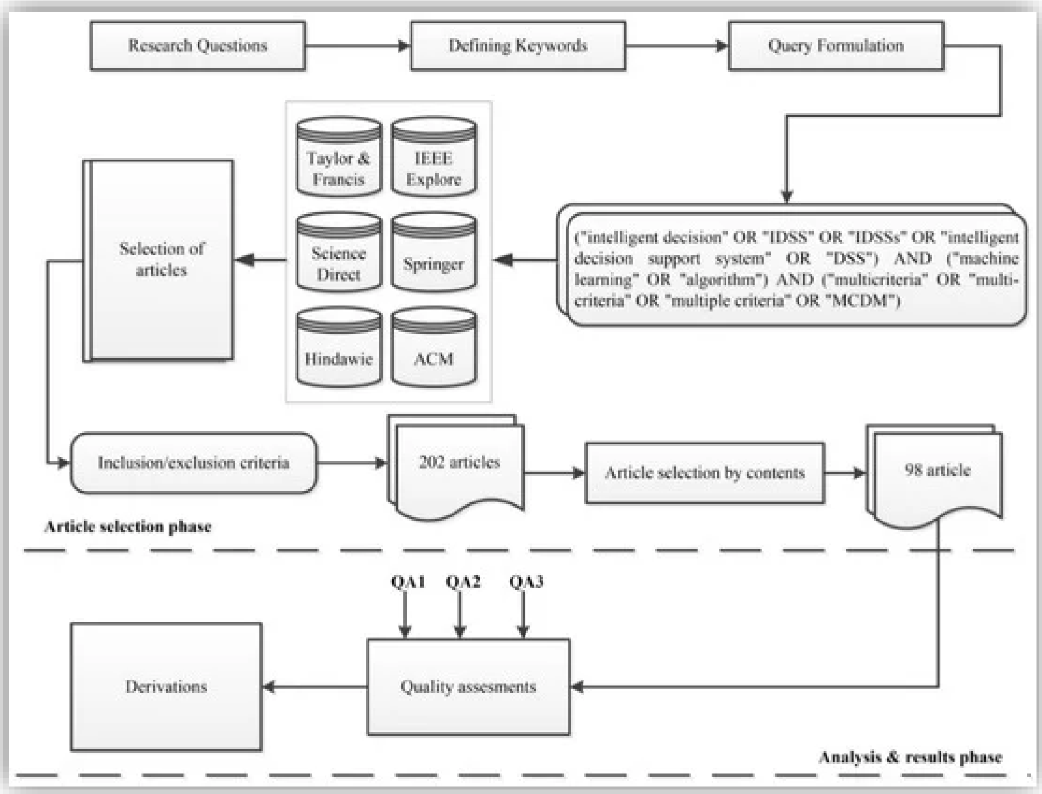


Figure 17: A Survey of Machine Learning Methods for IoT and their Future Applications

2.5.3 Hybrid and Multi-criteria Decision-Making Models

Hybrid and multi-criteria decision-making models combine various techniques and approaches to leverage their respective strengths and overcome individual limitations. These models often integrate rule-based systems, machine learning algorithms, and multi-criteria decision analysis (MCDA) methods. MCDA techniques, such as the Analytic Hierarchy Process (AHP) and Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS), can be used to evaluate and rank multiple crop alternatives based on multiple criteria, including economic, environmental, and social factors. By combining rules, machine learning models, and MCDA methods, these hybrid approaches aim to provide more robust and comprehensive crop recommendations, taking into account various stakeholder perspectives and decision-making criteria [29].



2.6 Limitations and Future Directions

While crop recommendation systems have shown promising results in various applications, there are still several challenges and limitations that need to be addressed. This section explores the current challenges and limitations, as well as potential improvements and future research directions in this field.

2.6.1 Challenges and Limitations of Current Crop Recommendation Systems

Despite the advancements in crop recommendation systems, several challenges and limitations persist [4]:

A. Data Quality and Availability:

- Incomplete or inaccurate data can lead to unreliable recommendations.
- Limited availability of high-quality, site-specific data is a common issue, especially in developing regions.

B. Handling Dynamic and Changing Conditions:

- Current systems may not adequately adapt to rapidly changing environmental conditions, such as climate change or extreme weather events.
- Accounting for dynamic market demands and fluctuating prices can be challenging.

C. Interpretability and Transparency:

- Some machine learning models, particularly deep learning techniques, can be perceived as "black boxes," lacking interpretability and transparency in their decision-making processes.
- Lack of trust and understanding from end-users can hinder adoption.

D. Integration of Multiple Objectives and Constraints:

- Balancing economic, environmental, and social factors in crop recommendations is a complex multi-objective optimization problem.
- Incorporating stakeholder preferences and local constraints can be challenging.

E. Scalability and Computational Complexity:

- As the volume and complexity of data increase, the computational requirements for training and deploying crop recommendation systems may become prohibitive.

2.6.2 Potential Improvements and Future Research Directions

To address the limitations and challenges of current crop recommendation systems, several potential improvements and future research directions can be explored [2]:

A. Improved Data Collection and Integration:

- Leveraging emerging technologies such as remote sensing, Internet of Things (IoT), and sensor networks for better data acquisition and integration.
- Developing standardized data formats and protocols for seamless data sharing and interoperability.

B. Adaptive and Robust Modeling Techniques:

- Exploring adaptive and online learning techniques to continuously update and refine crop recommendation models based on new data and changing conditions.
- Incorporating uncertainty quantification and robust optimization methods to handle data noise and variability.

C. Explainable AI and Interpretable Models:

- Developing explainable AI (XAI) techniques to improve the interpretability and transparency of machine learning models used in crop recommendation systems.
- Enabling end-users to understand the reasoning behind recommendations and build trust in the system.

D. Multi-Objective Optimization and Decision Support:

- Integrating multi-objective optimization algorithms and decision support tools to balance economic, environmental, and social objectives in crop recommendations.
- Incorporating stakeholder preferences and local constraints through interactive and participatory decision-making processes.

E. Distributed and Parallel Computing:

- Leveraging distributed computing frameworks and parallel processing techniques to handle large-scale data and computationally intensive tasks in crop recommendation systems.
- Exploring edge computing and fog computing paradigms for efficient data processing and decision-making at the edge of the network.

2.1 Conclusion

In this chapter, we have thoroughly explored the various types of recommendation systems and their applicability to crop recommendation systems, Considering the unique characteristics and challenges of the agricultural domain, we have concluded that the content-based filtering approach, combined with machine learning algorithms, is the most suitable for our crop recommendation system. This approach allows us to leverage detailed information about crop characteristics, soil conditions, climate data, and other relevant factors to generate personalized recommendations for farmers.

By combining the strengths of content-based filtering and machine learning algorithms, our crop recommendation system will provide personalized and data-driven recommendations to farmers, helping them make informed decisions about crop selection based on their specific farm conditions, preferences, and market demands. This system will contribute to optimizing crop production, enhancing resource efficiency, and ultimately improving the sustainability and profitability of agricultural practices.

Chapter 3

A Crop Recommendation System Based on ML Approach

3.1 Introduction

Precision agriculture has transformed farming by using data-driven analysis and customized management techniques to maximize crop production, optimize resource use, and enhance environmental sustainability. The fundamental element of this revolution is the notion of crop recommendation systems, which use sophisticated algorithms and extensive data sets to assist farmers in choosing the most appropriate crops for their particular field circumstances.

The preceding chapters established the foundation by examining the fundamentals of precision agriculture and the essential function of crop recommendation systems. We explored the advantages of site-specific crop management, the difficulties encountered in conventional farming methods, and the possibility of data-driven decision support systems to improve agricultural production and profitability.

Expanding on this basis, the main goal of this chapter is to provide a thorough solution for constructing a sophisticated crop recommendation system. This solution seeks to provide farmers with a strong tool that combines well-documented soil data, state-of-the-art algorithms, and an easy web application interface. Its purpose is to improve crop selection processes by using accurate and data-driven techniques.

This chapter will begin by presenting a summary of the content-based recommendation method used in our solution, as well as the overall structure and algorithms implemented for crop suggestions. We will analyze the capabilities and constraints of several algorithms, and evaluate their effectiveness in the specific context of crop recommendation tasks.

Following that, we will explore the intricate aspects of our recommendation system's implementation, including the used tools and technologies, description of the dataset, and preprocessing procedures. In addition, we will undertake a contextual analysis of the algorithms, assessing their performance on our dataset.

In addition, the chapter will reveal the progression of the web application's development, which functions as the intuitive interface for our crop recommendation system. We will provide visual representations, such as schematics and screenshots, to demonstrate the progression of the program, the design of the user interface, and the incorporation of the recommendation system. Special attention will be given to managing user inputs and presenting customized suggestions that are specifically designed for the unique circumstances of each farm.

3.2 Overview of Content-Based Filtering

Content-based filtering (CBF), also known as "cognitive filtering", recommends items based on the user’s item profile and user profile. These profiles are created at the beginning when the user creates an account and starts using the system. As the user interacts more with the system, a stronger user profile is created. Here, only the user’s information is required, rather than information from other similar users, so only a small scope of information is needed for recommendations.

The idea behind the CBF system is that "if a user likes an item in the past, then the user probably likes similar types of items in the future." Therefore, CBF compares the user’s item profile with the current item’s profile and tries to recommend similar types of items that the user may like. The user’s profile is constructed using various keywords, so the CBF system simply matches the keywords of highly-rated item profiles. To build the user profile, information about the user's item preferences and user information is needed, which can be gathered explicitly or implicitly.

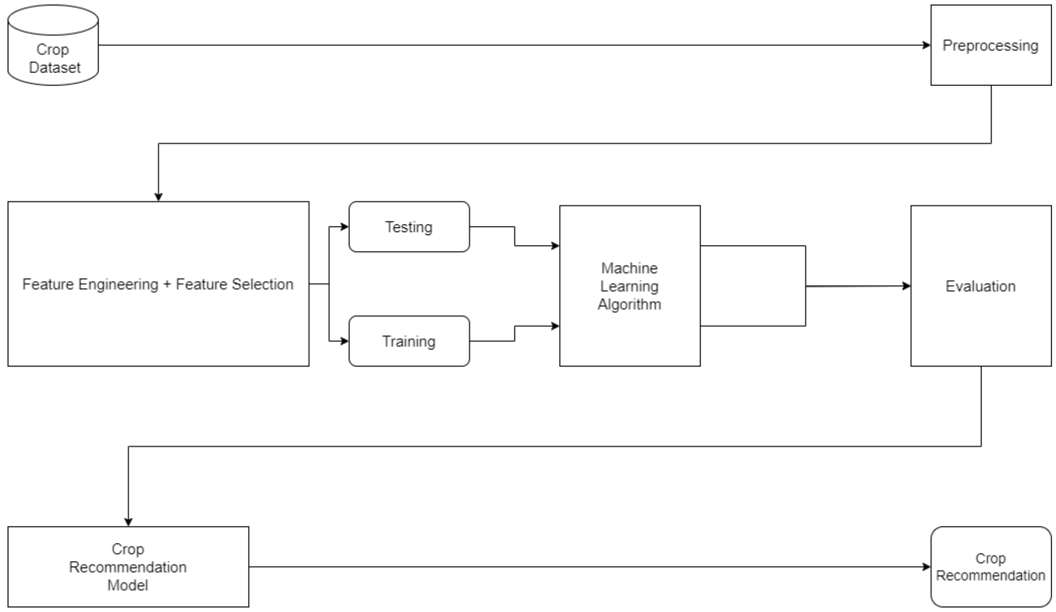


Figure 18: General structure of the recommendation model

3.2.1. Algorithms used for crop recommendation

Algorithm name	Definition	Strength	Limitations
Decision Tree	Builds a tree-like model by splitting the data based on feature values	Easy to interpret, handles non-linear relationships, can handle both numerical and categorical data	Can overfit, sensitive to data quality, and may struggle with high-dimensional data
Support Vector Machine (SVM)	Finds the optimal hyperplane that separates classes with the maximum margin	Effective for high-dimensional data, robust to overfitting, and versatile for various tasks.	Sensitive to parameter tuning, doesn't handle well outliers, and can be slow for large datasets
Logistic Regression	Models the probability of a binary outcome based on one or more features	Simple and interpretable, fast to train, and can handle non-linear relationships with feature engineering	Assumes linearity, sensitive to outliers, and may struggle with high-dimensional data
Random Forest	An ensemble of decision trees that combines their predictions	Robust to noise and outliers, can handle high-dimensional data, and resistant to overfitting	Can be slow for large datasets, difficult to interpret, and may overfit for some datasets

XGBoost	A gradient boosting algorithm that builds an ensemble of weak learners	Highly accurate, efficient for large datasets, and can handle missing data	Can be prone to overfitting, difficult to interpret, and requires careful parameter tuning
LightGBM	A gradient boosting framework that uses tree-based learning algorithms	Fast training speed, efficient for large datasets, and handles high-dimensional data well	Can overfit, sensitive to parameter tuning, and may struggle with non-linear relationships

Table 2: Algorithms used for crop recommendation

Decision Trees and Random Forests are good for interpretability, but can be prone to overfitting. SVMs and Logistic Regression are simple and efficient for linearly separable data, but may struggle with high-dimensional or non-linear data. XGBoost and LightGBM are powerful ensemble methods that can handle complex data but are less interpretable and require careful tuning. The choice depends on the specific problem, dataset characteristics, and trade-offs between accuracy, interpretability, and computational resources.

3.3 Implementation of the Crop Recommendation System

3.3.1. Tools and technologies used

a. Python

Python is a high-level, general-purpose programming language. Its design philosophy emphasizes code readability with the use of significant indentation, it consistently ranks as one of the most popular programming languages, and has gained widespread use in the Machin learning community [30].

b. Flask

Flask is a micro web framework written in Python. It is classified as a microframework because it does not require particular tools or libraries. It has no database abstraction layer, form validation, or any other components where pre-existing third-party libraries provide common functions.

c. Bootstrap

Bootstrap is a free and open-source CSS framework directed at responsive, mobile-first front-end web development. It contains HTML, CSS and JavaScript-based design templates for typography, forms, buttons, navigation, and other interface components.

d. Bootstrap Studio

Bootstrap Studio is a proprietary web design and development application. It offers a large number of components for building responsive pages including headers, footers, galleries and slideshows along with basic elements, such as spans and divs. The program can be used for building websites and prototypes.

e. Visual Studio Code

Visual Studio is a powerful developer tool that you can use to complete the entire development cycle in one place. It is a comprehensive integrated development environment (IDE) that you can use to write, edit, debug, and build code, and then deploy your app.

f. PyCharm

PyCharm is a dedicated Python Integrated Development Environment (IDE) providing a wide range of essential tools for Python developers, tightly integrated to create a convenient environment for productive Python, web, and data science development.

g. Telegram

A cross-platform cloud-based instant messaging service that offers secure and feature-rich communication.

h. Google Meet

A video communication service developed by Google that allows individuals to conduct video conferences, meetings, and online events.

i. Microsoft Office

A comprehensive suite of productivity software developed by Microsoft, including Word, Excel, PowerPoint, and other applications for document creation, data analysis, and presentations.

3.3.2 Dataset description and preprocessing

1. **Sources of soil data**

   The data used in this project is made by augmenting and combining various publicly available datasets of India like weather, soil, etc. You can access the dataset on Kaggle [31]. This data is relatively simple with very few but useful features unlike the complicated features affecting the yield of the crop.

The data has Nitrogen, Phosphorous, Potassium and pH values of the soil. Also, it also contains the humidity, temperature and rainfall required for a particular crop.

N	P	K	Temperature	Humidity	pH	Rainfall	Label
26	50	19	27.31791	51.66921	6.005243	32.5592	Mothbeans
39	5	31	27.10135	93.6998	5.551963	150.9503	Coconut
4	24	43	22.40424	88.15083	7.199504	109.8695	Pomegranate
90	14	52	24.84741	89.20455	6.391858	59.67927	Watermelon
102	46	19	22.77076	82.59933	6.631005	81.49543	Cotton

Table 3: Data set of our crop recommendation system

2. Feature engineering

```python
# Importing libraries

from __future__ import print_function
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.metrics import classification_report
from sklearn import metrics
from sklearn import tree
import warnings
warnings.filterwarnings('ignore')

PATH = '../input/crop-recommendation-dataset/Crop_recommendation.csv'
df = pd.read_csv(PATH)

features = df[['N', 'P','K','temperature', 'humidity', 'ph', 'rainfall']]
target = df['label']
labels = df['label']

# Initializing empty lists to append all model's name and corresponding name
acc = []
model = []

# Splitting into train and test data

from sklearn.model_selection import train_test_split
Xtrain, Xtest, Ytrain, Ytest = train_test_split(features,target,test_size =
0.2,random_state =2)
```

3.3.3. Contextual comparison of algorithms

By conducting a contextual comparison of algorithms, we can assess how well each algorithm performs in terms of accuracy, precision, recall, and other relevant metrics. Moreover, analyzing the execution time of each algorithm provides insights into their computational efficiency and scalability

A. Decision Tree

```
from sklearn.tree import DecisionTreeClassifier
from sklearn import metrics
import time

start = time.time()

DecisionTree = DecisionTreeClassifier(criterion = "entropy", random_state =
2, max_depth = 5)

DecisionTree.fit(Xtrain, Ytrain)

predicted_values = DecisionTree.predict(Xtest)
accuracy = metrics.accuracy_score(Ytest, predicted_values)
model.append('Decision Tree')

precision = metrics.precision_score(Ytest, predicted_values, average =
'weighted')
recall = metrics.recall_score(Ytest, predicted_values, average =
'weighted')
f1_score = metrics.f1_score(Ytest, predicted_values, average = 'weighted')
end = time.time()
execution_time = end - start


print(f"DecisionTrees's Accuracy is: {accuracy*100:.2f}%")
print(f"DecisionTrees's Precision is: {precision*100:.2f}%")
print(f"DecisionTrees's Recall is: {recall*100:.2f}%")
print(f"DecisionTrees's F1-score is: {f1_score*100:.2f}%")
print(f"Execution time: {execution_time*1000:.2f} milliseconds")
```

Output :

```
DecisionTrees's Accuracy is: 90.00%
DecisionTrees's Precision is: 85.72%
DecisionTrees's Recall is: 90.00%
DecisionTrees's F1-score is: 87.07%
Execution time: 55.89 milliseconds
```

this code builds and evaluates a decision tree classifier. It measures the classifier's accuracy, precision, recall, and F1-score on a test dataset. Additionally, it tracks the execution time required for training and making predictions. The decision tree classifier is configured to use entropy for splitting, with a maximum depth of 5 to prevent overfitting and ensure a balance between model complexity and performance.

B. Guassian Naive Bayes

```
from sklearn.naive_bayes import GaussianNB
from sklearn import metrics
import time
start= time.time()
NaiveBayes = GaussianNB()

NaiveBayes.fit(Xtrain, Ytrain)

predicted_values = NaiveBayes.predict(Xtest)
accuracy = metrics.accuracy_score(Ytest, predicted_values)
model.append('Naive Bayes')

precision = metrics.precision_score(Ytest, predicted_values,
average='weighted')
recall = metrics.recall_score(Ytest, predicted_values, average='weighted')
f1_score = metrics.f1_score(Ytest, predicted_values, average='weighted')
end = time.time()
execution_time = end - start
print(f"Naive Bayes's Accuracy is: {accuracy*100:.2f}%")
print(f"Naive Bayes's Precision is: {precision*100:.2f}%")
print(f"Naive Bayes's Recall is: {recall*100:.2f}%")
print(f"Naive Bayes's F1-score is: {f1_score*100:.2f}%")

print(f"Execution time: {execution_time*1000:.2f} milliseconds")
```

Output:

```
Naive Bayes's Accuracy is: 99.09%

Naive Bayes's Precision is: 99.20%

Naive Bayes's Recall is: 99.09%

Naive Bayes's F1-score is: 99.06%

Execution time: 36.9 milliseconds
```

The results indicate that the Gaussian Naive Bayes classifier is highly effective for this dataset, achieving over 99% accuracy and similarly high precision, recall, and F1-score, all within a very short execution time of 36.9 milliseconds. This makes it a robust and efficient model for classification tasks.

C. Support Vector Machine (SVM)

```
from sklearn.svm import SVC
from sklearn import metrics
import time
start= time.time()
SVM = SVC(gamma='auto')

SVM.fit(Xtrain, Ytrain)

predicted_values = SVM.predict(Xtest)
accuracy = metrics.accuracy_score(Ytest, predicted_values)

precision = metrics.precision_score(Ytest, predicted_values,
average='weighted')
recall = metrics.recall_score(Ytest, predicted_values, average='weighted')
f1_score = metrics.f1_score(Ytest, predicted_values, average='weighted')
end = time.time()
execution_time = end - start
print(f"SVM's Accuracy is: {accuracy*100:.2f}%")
print(f"SVM's Precision is: {precision*100:.2f}%")
print(f"SVM's Recall is: {recall*100:.2f}%")
print(f"SVM's F1-score is: {f1_score*100:.2f}%")

print(f"Execution time: {execution_time*1000:.2f} milliseconds")
```

Output:

```
SVM's Accuracy is: 10.68%

SVM's Precision is: 65.79%

SVM's Recall is: 10.68%

SVM's F1-score is: 13.06%

Execution time: 722.47 milliseconds
```

These results indicate that the SVM classifier with the given parameters did not perform well on the test dataset. The low accuracy, recall, and F1-score suggest that the model's predictions are mostly incorrect. This poor performance could be due to various reasons such as improper parameter tuning, inappropriate feature scaling, or a complex dataset that the SVM with default parameters could not handle effectively. Adjusting the gamma parameter, using different kernels, or performing hyperparameter tuning might improve the model's performance.

D. Logistic Regression

```
from sklearn.linear_model import LogisticRegression
from sklearn import metrics
import time
start= time.time()
LogReg = LogisticRegression(random_state=2)

LogReg.fit(Xtrain, Ytrain)

predicted_values = LogReg.predict(Xtest)
accuracy = metrics.accuracy_score(Ytest, predicted_values)

precision = metrics.precision_score(Ytest, predicted_values,
average='weighted')
recall = metrics.recall_score(Ytest, predicted_values, average='weighted')
f1_score = metrics.f1_score(Ytest, predicted_values, average='weighted')
end = time.time()
execution_time = end - start
print(f"Logistic Regression's Accuracy is: {accuracy*100:.2f}%")
print(f"Logistic Regression's Precision is: {precision*100:.2f}%")
print(f"Logistic Regression's Recall is: {recall*100:.2f}%")
print(f"Logistic Regression's F1-score is: {f1_score*100:.2f}%")

print(f"Execution time: {execution_time*1000:.2f} milliseconds")
```

Output

```
Logistic Regression's Accuracy is: 95.23%

Logistic Regression's Precision is: 95.26%

Logistic Regression's Recall is: 95.23%

Logistic Regression's F1-score is: 95.16%

Execution time: 215.01 milliseconds
```

The results indicate that the Logistic Regression classifier is highly effective for this dataset, achieving over 95% accuracy and similarly high precision, recall, and F1-score. The relatively short execution time of 215.01 milliseconds demonstrates that the model is efficient as well as accurate. Logistic Regression, being a linear model, is particularly well-suited for datasets where the relationship between the features and the target variable is approximately linear. The high-performance metrics suggest that the model has successfully captured the underlying patterns in the data.

E. Random Forest

```
from sklearn.ensemble import RandomForestClassifier
from sklearn import metrics
import time
start= time.time()
RF = RandomForestClassifier(n_estimators=20, random_state=0)

RF.fit(Xtrain, Ytrain)

predicted_values = RF.predict(Xtest)
accuracy = metrics.accuracy_score(Ytest, predicted_values)

precision = metrics.precision_score(Ytest, predicted_values,
average='weighted')
recall = metrics.recall_score(Ytest, predicted_values, average='weighted')
f1_score = metrics.f1_score(Ytest, predicted_values, average='weighted')
end = time.time()
execution_time = end - start
print(f"Random Forest's Accuracy is: {accuracy*100:.2f}%")
print(f"Random Forest's Precision is: {precision*100:.2f}%")
print(f"Random Forest's Recall is: {recall*100:.2f}%")
print(f"Random Forest's F1-score is: {f1_score*100:.2f}%")

print(f"Execution time: {execution_time*1000:.2f} milliseconds")
```

Output:

```
Random Forest's Accuracy is: 99.09%

Random Forest's Precision is: 99.17%

Random Forest's Recall is: 99.09%

Random Forest's F1-score is: 99.07%

Execution time: 128.55 milliseconds
```

The results indicate that the Random Forest classifier is highly effective for this dataset, achieving over 99% accuracy and similarly high precision, recall, and F1-score. The relatively short execution time of 128.55 milliseconds demonstrates that the model is efficient as well as accurate.

F. XGBoost

```
import xgboost as xgb
from sklearn import metrics
import time
start= time.time()
XB = xgb.XGBClassifier()

from sklearn.preprocessing import LabelEncoder

encoder = LabelEncoder()
Ytrain = encoder.fit_transform(Ytrain)
Ytest = encoder.transform(Ytest)

XB.fit(Xtrain,Ytrain)

predicted_values = XB.predict(Xtest)
accuracy = metrics.accuracy_score(Ytest, predicted_values)

precision = metrics.precision_score(Ytest, predicted_values,
average='weighted')
recall = metrics.recall_score(Ytest, predicted_values, average='weighted')
f1_score = metrics.f1_score(Ytest, predicted_values, average='weighted')
end = time.time()
execution_time = end - start
print(f"XGBoost's Accuracy is: {accuracy*100:.2f}%")
print(f"XGBoost's Precision is: {precision*100:.2f}%")
print(f"XGBoost's Recall is: {recall*100:.2f}%")
print(f"XGBoost's F1-score is: {f1_score*100:.2f}%")

print(f"Execution time: {execution_time*1000:.2f} milliseconds")
```

Output:

```
XGBoost's Accuracy is: 99.09%

XGBoost's Precision is: 99.13%

XGBoost's Recall is: 99.09%

XGBoost's F1-score is: 99.08%

Execution time: 1217.78 milliseconds
```

The results indicate that the XGBoost classifier is highly effective for this dataset, achieving over 99% accuracy and similarly high precision, recall, and F1-score. The execution time of 1217.78 milliseconds demonstrates that the model is relatively efficient given its high performance.

G. LightGBM

```
import lightgbm as lgb
from sklearn import metrics
import time

start = time.time()
LG = lgb.LGBMClassifier()

LG.fit(Xtrain, Ytrain)

predicted_values = LG.predict(Xtest)
accuracy = metrics.accuracy_score(Ytest, predicted_values)

precision = metrics.precision_score(Ytest, predicted_values,
average='weighted')
recall = metrics.recall_score(Ytest, predicted_values, average='weighted')
f1_score = metrics.f1_score(Ytest, predicted_values, average='weighted')
end = time.time()
execution_time = end - start
print(f"LightGBM's Accuracy is: {accuracy * 100:.2f}%")
print(f"LightGBM's Precision is: {precision * 100:.2f}%")
print(f"LightGBM's Recall is: {recall * 100:.2f}%")
print(f"LightGBM's F1-score is: {f1_score * 100:.2f}%")

print(f"Execution time: {execution_time * 1000:.2f} milliseconds")
```

Output:

```
LightGBM's Accuracy is: 99.32%

LightGBM's Precision is: 99.34%

LightGBM's Recall is: 99.32%

LightGBM's F1-score is: 99.32%

Execution time: 176.94 milliseconds
```

The results indicate that the LightGBM classifier is highly effective for this dataset, achieving over 99% accuracy and similarly high precision, recall, and F1-score. The relatively short execution time of 176.94 milliseconds demonstrates that the model is efficient as well as accurate.

In this section, we will compare the algorithms used in our recommendation system based on these key aspects, enabling us to make informed decisions about which algorithm is best suited for our crop recommendation system.

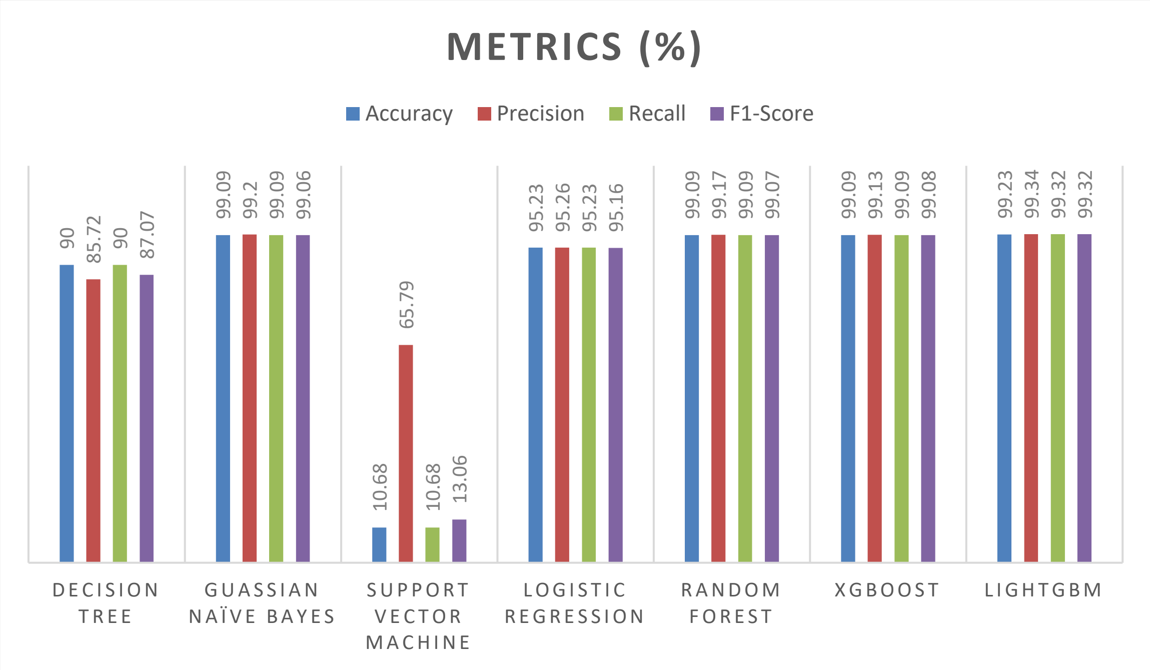


Figure 19: Comparison between algorithms depending on metrics

Figure 20: Comparison between algorithms depending on execution time

Based on the comprehensive evaluation of various machine learning algorithms, LightGBM emerged as the top performer, showcasing exceptional predictive power and efficiency. The LightGBM classifier achieved the highest accuracy at 99.32%, significantly outperforming other models. Furthermore, it demonstrated superior precision and recall, both at 99.34% and 99.32%, respectively, indicating its remarkable ability to correctly identify positive instances and retrieve all relevant cases. The F1-score, a balanced metric considering both precision and recall, also stood at an impressive

99.32%, affirming the model's robust performance across different evaluation criteria. In addition to its high accuracy and reliability, LightGBM maintained a reasonable execution time of 176.94 milliseconds, which, while not the fastest, represents a commendable balance between computational efficiency and predictive accuracy. This makes LightGBM an optimal choice, combining top-tier performance with efficient processing, making it ideal for applications requiring both high accuracy and quick turnaround times.

3.3.4. Web application for the Crop Recommendation System

1- Diagrams

a) Class Diagram

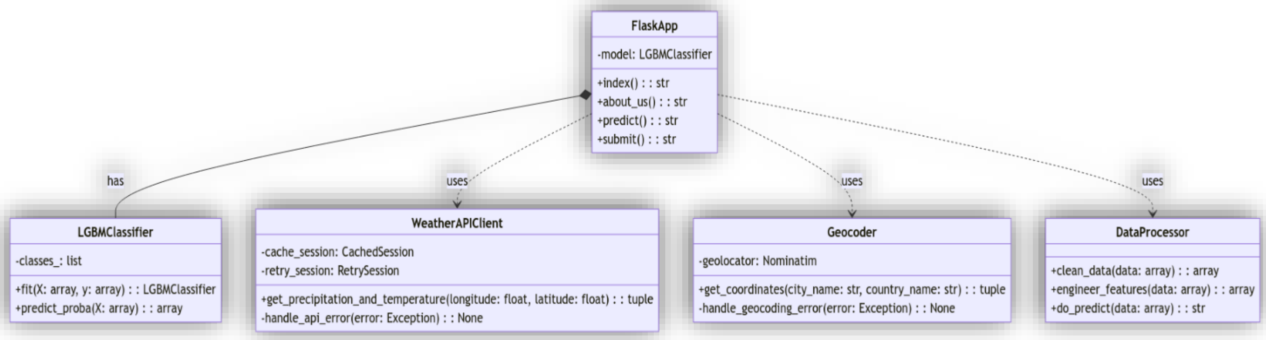


Figure 21: Class diagram

b) Sequence Diagram

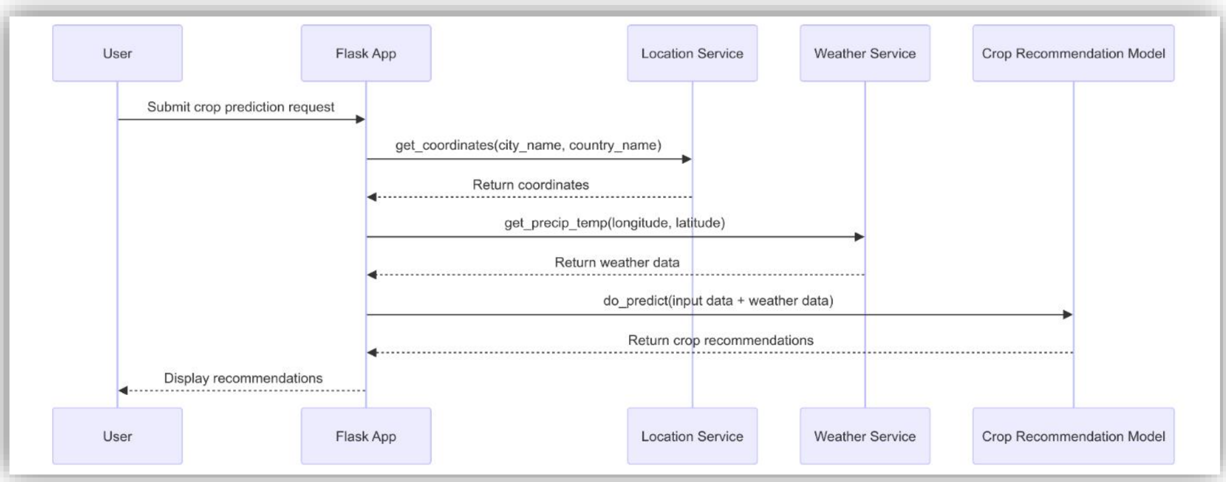


Figure 22: Sequence diagram

2- Integration of the recommendation algorithm

```
import geopy
import joblib
import numpy as np
import openmeteo_requests
import pandas as pd
import requests_cache
from flask import Flask, render_template, request
from geopy.geocoders import Nominatim
from retry_requests import retry
from operator import itemgetter
import lightgbm
import sklearn
```

This segment imports the necessary libraries and modules required for the application. Some notable imports include Flask (for building the web application), Nominatim (for geocoding), joblib (for loading the machine learning model), pandas and numpy (for data manipulation), and requests_cache (for caching API requests).

```
import pandas as pd
from sklearn.feature_selection import SelectKBest, f_classif
from sklearn.model_selection import train_test_split
import lightgbm
import joblib

# Load the dataset
df = pd.read_csv('sds.csv')

# Define the features and target
X = df.drop('label', axis=1)
y = df['label']

# Split the data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
random_state=42)

# Apply feature engineering
X_train['temperature'] = *Function*
X_train['humidity'] = *Function*
X_train['ph'] = *Function*
X_train['rainfall'] = *Function*
X_test['temperature'] = *Function*
X_test['humidity'] = *Function*
X_test['ph'] = *Function*
X_test['rainfall'] = *Function*

# Apply feature selection
selector = SelectKBest(f_classif, k=4)
X_train_selected = selector.fit_transform(X_train, y_train)
X_test_selected = selector.transform(X_test)

# Train the model
model = lightgbm.LGBMClassifier()
model.fit(X_train_selected, y_train)

# Dumping the model into a file
joblib.dump(model, 'model.pkl')
```

We train the model using Feature Engineering and Feature Selection, then we dump it into a file “model.pkl”.

```
model = joblib.load("static/model.pkl")
unique_labels = model.classes
```

We load the pre-trained machine learning model (which was saved using joblib.dump) and retrieve the unique class labels from the loaded model.

```
def get_coordinates(city_name, country_name): …
#. . . Function
def get_precip_temp(longitude, latitude): …
#. . . Function
def do_predict(array): …
#. . . Function
```

Here. We have three helper functions:

1. **get_coordinates(city_name, country_name):** This function uses the Nominatim geocoder from the geopy library to retrieve the coordinates (longitude and latitude) for a given city and country.
2. **get_precip_temp(longitude, latitude):** This function retrieves weather data (precipitation, temperature, and relative humidity) from the Open-Meteo API for the given coordinates. It uses the openmeteo_requests library and caches the API requests using requests_cache.
3. **do_predict(array):** This function takes an input array (containing the feature values) and uses the loaded machine learning model to make predictions. It returns a string containing the predicted crop names and their associated scores.

```
app = Flask(__name__)
app.static_folder = 'static'


# Define routes and render HTML templates
@app.route('/')
def index():
    return render_template('index.html')


@app.route('/about-us')
def about_us():
    return render_template('about-us.html')


@app.route('/predict')
def predict():
    return render_template('predict.html')


@app.route('/submit', methods=['POST'])
def submit(): …
#. . . Function
if __name__ == '__main__':
    app.run(host='0.0.0.0',debug=True)
```

We create a Flask application instance and define various routes for it. The /submit route handles POST requests from the prediction form. It retrieves the input parameters, calls the helper functions to obtain weather data and make predictions, and then renders the predict.html template with the predicted crop. The full code can be accessed through the web app's Github [32].

3.3.3 The Prototype

A. Home page

This page welcomes the user by creating a showcase for what their fields would be after using our recommendation system.

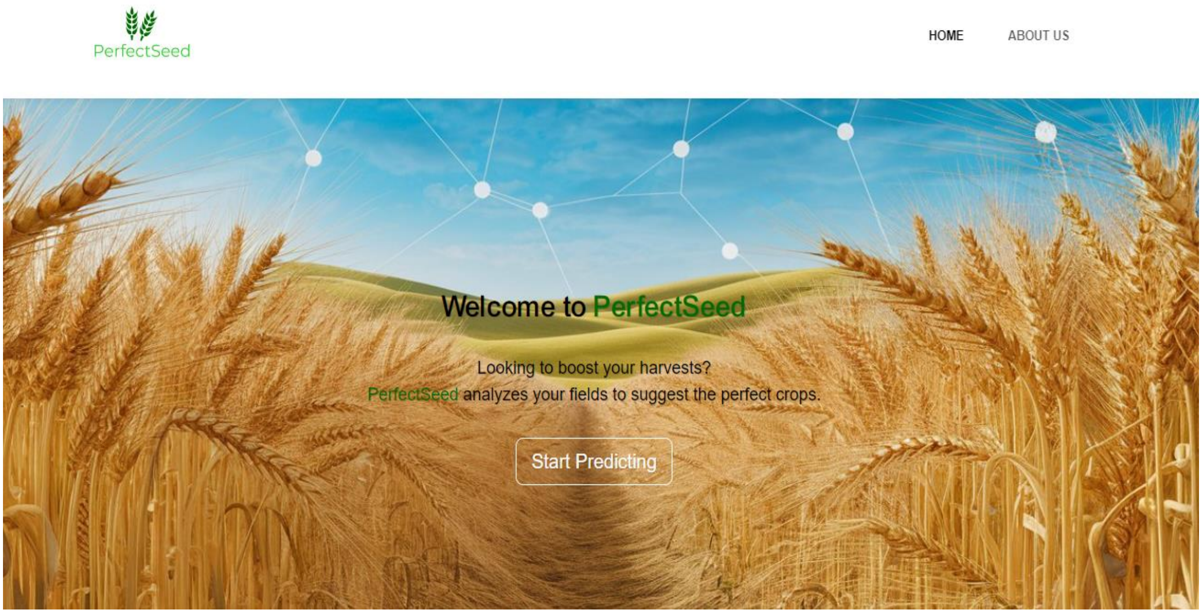


Figure 23: Home Page (1)

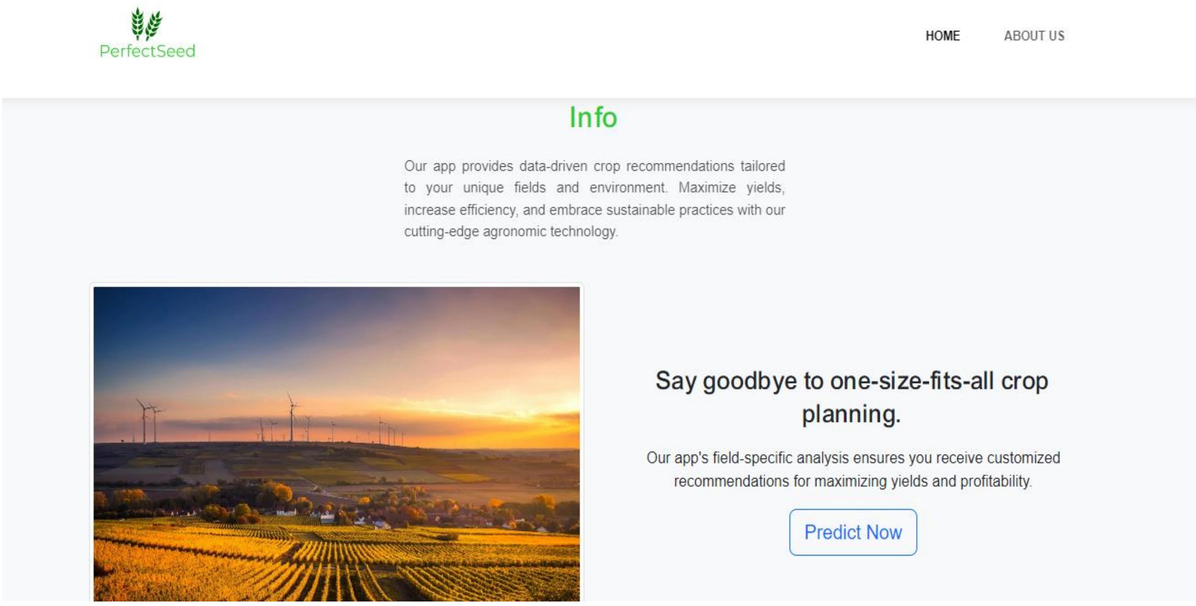


Figure 24: Home Page (2)

B. Prediction page

This page allows the user to input their soil data to predict the perfect crops for their fields.

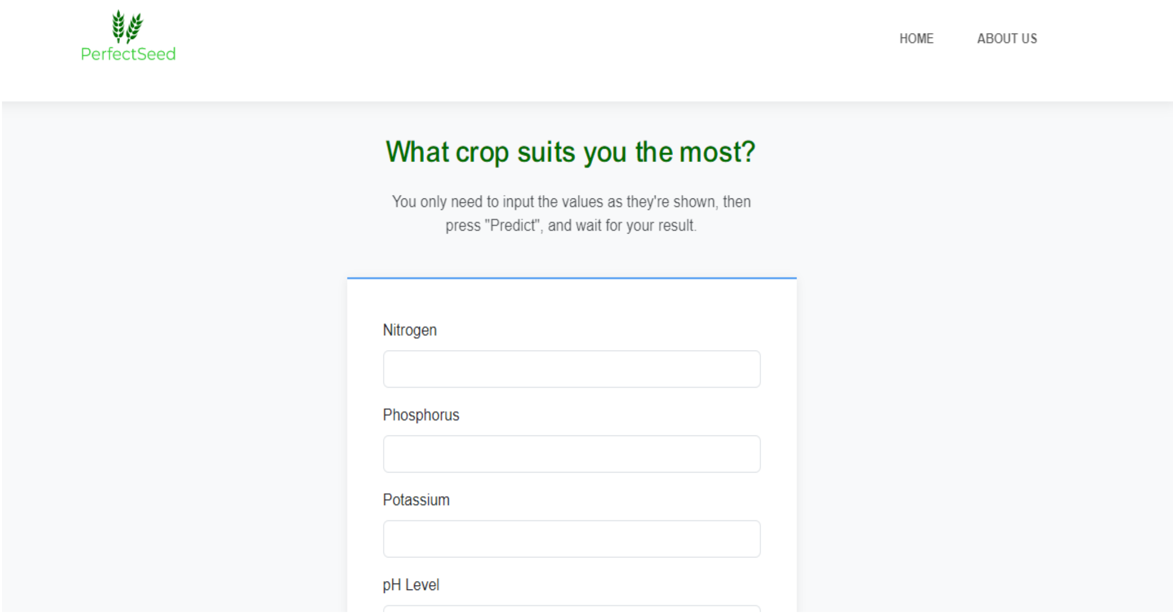


Figure 25: Prediction Page (1)

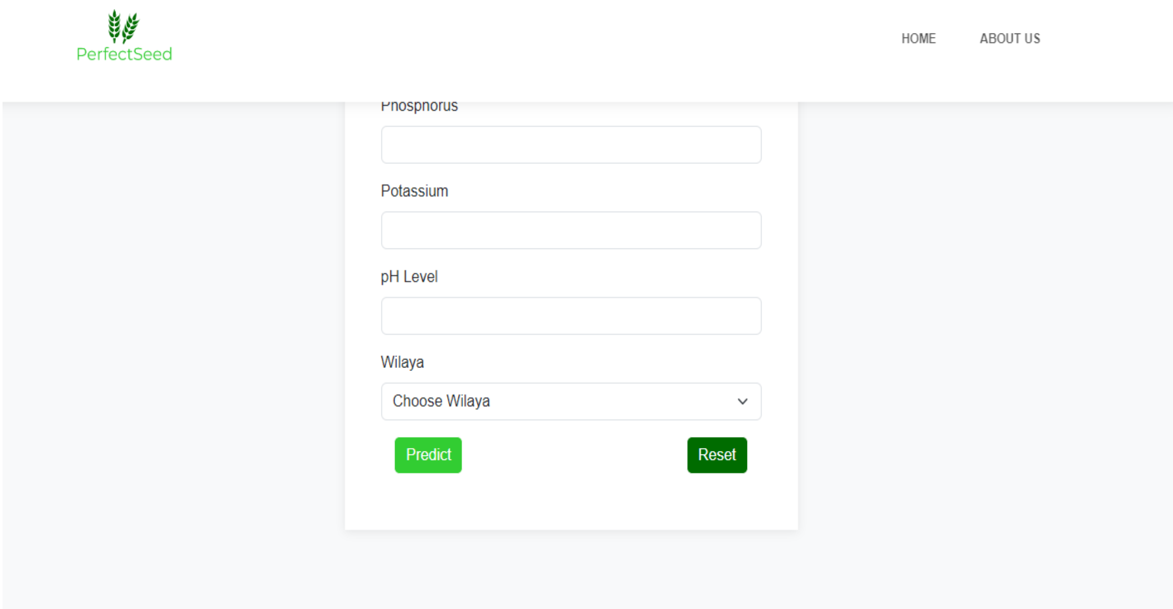


Figure 26: Prediction Page (2)

Here is an example showcase of what would happen after inputting your soil's data and pressing "predict".

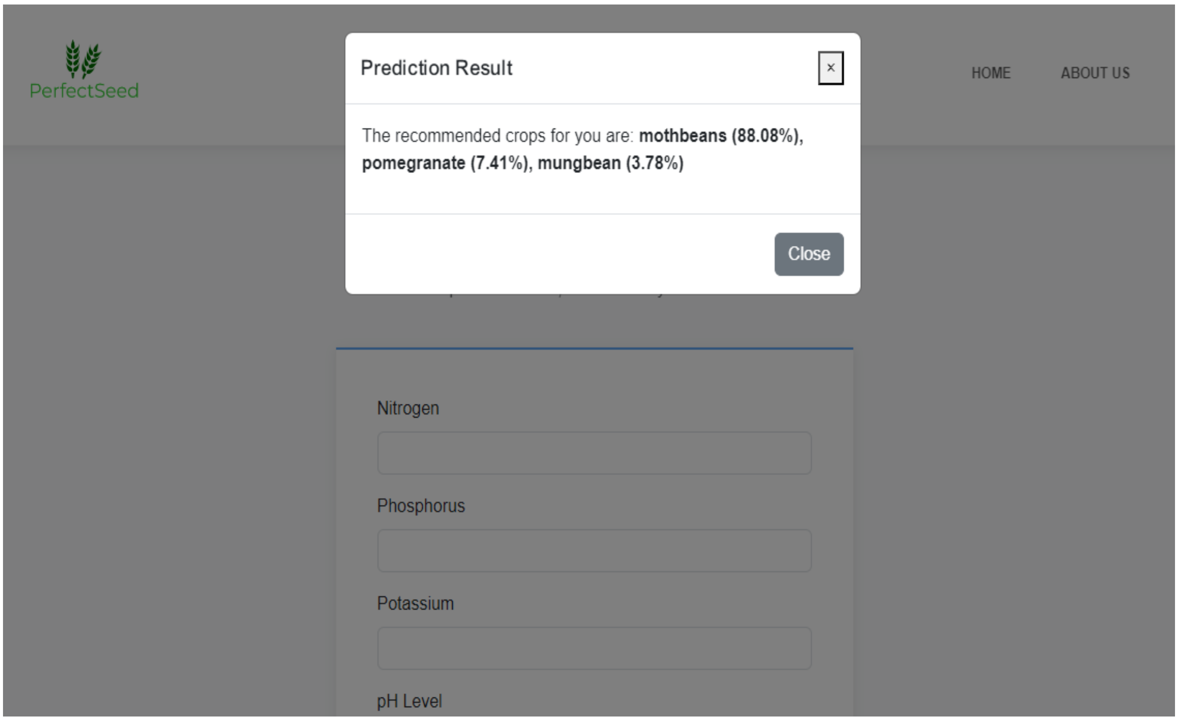


Figure 27: Prediction showcase

3.4 Conclusion

In this chapter we focused on the implementation of an advanced crop recommendation system using a content-based recommendation approach. It builds upon the concepts of precision agriculture and crop recommendation systems discussed in the previous chapters.

The chapter begins by outlining the general architecture and content-based recommendation approach. It provides an overview of the recommendation model's structure and the algorithms used for crop recommendations, including decision trees, XGBoost, and LightGBM.

A contextual comparison of the algorithms is conducted using evaluation metrics such as accuracy, precision, recall, and F1-score. The performance of each algorithm is compared on the dataset, with a particular focus on LightGBM, which was found to provide better accuracy compared to other algorithms.

The chapter also covers the development of a web application that integrates the recommendation algorithm. Along with explanations of how user inputs are handled and personalized recommendations are displayed.

Finally, the chapter concludes by summarizing the key points and contributions of the implemented crop recommendation system, also the focus on LightGBM and its superior performance in terms of accuracy is emphasized, as determined through the evaluation metrics used.

General Conclusion

General Conclusion

This thesis has examined the notion of precision agriculture and the creation of crop recommendation systems as a primary tool to aid farmers in making well-informed choices about crop selection.

The first chapter explores the importance of agriculture, its development, and the role of technology in overcoming challenges. It highlights the fundamental elements of precision agriculture, such as remote sensing, GPS, variable rate technology, and data analytics, and their advantages and challenges.

The second chapter delves into recommendation systems in agriculture, examining their various types, assessment criteria, and application. It covers rule-based systems, machine learning techniques, and hybrid decision-making models. The chapter also discusses constraints and future paths, emphasizing the need for improved data gathering, adaptable modeling, interpretable artificial intelligence, multi-objective optimization, and scalable computing solutions.

The third chapter discusses the implementation and structure of a proposed crop recommendation system, focusing on content-based filtering and specific algorithms. It details the tools and technologies used, the dataset used, and preparation procedures. The essay also discusses the comparison of algorithms and the creation of a web application prototype.

Overall, this thesis has greatly improved the subject of precision agriculture by creating a crop recommendation system that utilizes sophisticated algorithms and approaches to provide farmers tailored and evidence-based suggestions for choosing crops. The suggested approach tackles the issues of data quality, changing circumstances, interpretability, and many objectives, therefore enabling the adoption of more sustainable and efficient farming techniques.

Bibliography

Bibliography

[1] J. C. Aker, "Dial “A” for agriculture: a review of information and communication technologies for agricultural extension in developing countries," *Agricultural economics,* vol. 42, pp. 631--647, 2011.

[2] S. a. G. L. a. V. C. a. B. M.-J. Wolfert, "Big data in smart farming--a review," *Agricultural systems,* vol. 153, pp. 69--80, 2017.

[3] M. a. V. S. a. Z. M. a. o. Trendov, "Digital technologies in agriculture and rural areas: status report," *Digital technologies in agriculture and rural areas: status report,* 2019.

[4] K. G. a. B. P. a. M. D. a. P. S. a. B. D. Liakos, "Machine learning in agriculture: A review," *Sensors,* vol. 18, p. 2674, 2018.

[5] R. a. A. V. I. Gebbers, "Precision agriculture and food security," *Science,* pp. 828--831, 2010.

[6] H. a. G. X. Zhang, "Farmland Rental Market, Outsourcing Services Market and Agricultural Green Productivity: Implications for Multiple Forms of Large-Scale Management," *Land,* vol. 13, p. 676, 2024.

[7] J. V. Stafford, "Implementing precision agriculture in the 21st century," *Journal of agricultural engineering research,* vol. 76, pp. 267--275, 2000.

[8] P. Pingali, "Agricultural mechanization: adoption patterns and economic impact," *Handbook of agricultural economics,* vol. 3, pp. 2779--2805, 2007.

[9] J. a. E. B. Lowenberg-DeBoer, "Setting the record straight on precision agriculture adoption," *Agronomy Journal,* vol. 111, pp. 1552--1569, 2019.

[10] R. a. K. M. Miao, "Harnessing advances in agricultural technologies to optimize resource utilization in the food-energy-water nexus," *{Annual Review of Resource Economics,* vol. 12, pp. 65--85, 2020.

[11] L. a. S. B. Roth, "Predicting cover crop biomass by lightweight UAS-based RGB and NIR photography: an applied photogrammetric approach," *Precision Agriculture,* vol. 19, pp. 93--114, 2018.

[12] A. a. W. C. a. D. D. Dobermann, "Performance of Site-Specific Nutrient Management in Intensive Rice Cropping Systems of Asia," 2002.

[13] A. a. K. S. Khanna, "Evolution of Internet of Things (IoT) and its significant impact in the field of Precision Agriculture," *Computers and electronics in agriculture,* vol. 157, pp. 218--231, 2019.

[14] E. a. K. P. a. B. R. Nash, "Applications of open geospatial web services in precision agriculture: a review," *Precision agriculture,* vol. 10, pp. 546--560, 2009.

[15] S. a. R. E. a. R. R. H. a. K. C. a. K. T. a. N. J. Pudumalar, "Crop recommendation system for precision agriculture," in *2016 eighth international conference on advanced computing (ICoAC)*, 2017, pp. 32--36.

[16] F. Ricci, "Recommender Systems: Models and Techniques.," 2014.

[17] H. a. K. Y. S. a. C. K. J. Hwangbo, "Recommendation system development for fashion retail e-commerce," *Electronic Commerce Research and Applications,* vol. 28, pp. 94--101, 2018.

[18] A. a. o. Huang, "Similarity measures for text document clustering," in *Proceedings of the sixth new zealand computer science research student conference (NZCSRSC2008), Christchurch, New Zealand*, 2008, pp. 9--56.

[19] P. a. D. G. M. a. S. G. Lops, "Content-based recommender systems: State of the art and trends," *Recommender systems handbook,* pp. 73--105, 2011.

[20] M. J. a. B. D. Pazzani, "Content-based recommendation systems," in *The adaptive web: methods and strategies of web personalization*, Springer, 2007, pp. 325--341.

[21] E.-A. a. M. G. Minastireanu, "Light gbm machine learning algorithm to online click fraud detection," *J. Inform. Assur. Cybersecur,* vol. 2019, p. 263928, 2019.

[22] X. a. K. T. M. Su, "A survey of collaborative filtering techniques," *Advances in artificial intelligence,* vol. 2009, p. 421425, 2009.

[23] A. Vidhya, "Recommendation System K-Nearest Neighbors," 2023.

[24] T. D. Science, "An Overview of Matrix Factorization Techniques for Recommendation Systems," 2023.

[25] R. Papers, "Probabilistic Matrix Factorization".

[26] S. a. F. C. a. G. Z. a. S.-T. L. Rendle, "Bayesian Personalized Ranking from Implicit Feedback," *UAI,* 2009.

[27] R. Burke, "Hybrid web recommender systems," *The adaptive web: methods and strategies of web personalization,* pp. 377--408, 2007.

[28] S. a. M. J. a. A. B. a. T. L. a. K. J. Zingore, "Soil degradation in sub-Saharan Africa and crop production options for soil rehabilitation," 2015.

[29] Y. a. L. C. Jiang, "Convolutional neural networks for image-based high-throughput plant phenotyping: a review," *Plant Phenomics,* 2020.

[30] D. Kuhlman, A python book: Beginning python, advanced python, and python exercises, Dave Kuhlman Lutz, 2009.

[31] [Online]. Available: https://www.kaggle.com/datasets/atharvaingle/crop-recommendation-dataset.

[32] [Online]. Available: https://github.com/AzMimoune/PerfectSeed.

[33] A. a. B. P. Ezhilazhahi, IoT enabled plant soil moisture monitoring using wireless sensor networks, 2017, pp. 345--349.

[34] Z. a. G. D. a. M. G. a. Y. L. a. A. S. Ahmed, "An overview of smart irrigation management for improving water productivity under climate change in drylands," *Agronomy,* vol. 13, p. 2113, 2023.

Figure References

[Figure1]: Insights on Deep Learning based Segmentation Schemes Towards Analyzing Satellite Imageries - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Types-of-Satellite-Images-23_fig1_356825152

[Figure 2]: S. a. G. L. a. V. C. a. B. M.-J. Wolfert, "Big data in smart farming--a review," Agricultural systems, vol. 153, pp. 69--80, 2017.

[Figure 3] : A Context-Aware Middleware Cloud Approach for Integrating Precision Farming Facilities into the IoT toward Agriculture 4.0 - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Generic-architecture-and-data-workflow-overview-for-a-context-aware-IoT-agricultural_fig2_338770474

[Figure 4] : Cravero, A.; Sepúlveda, S. Use and Adaptations of Machine Learning in Big Data—Applications in Real Cases in Agriculture. Electronics 2021, 10, 552. https://doi.org/10.3390/electronics10050552

[Figure 5] : Par Riclas — Travail personnel, CC BY 3.0, https://commons.wikimedia.org/w/index.php?curid=9076846

[Figure 6] : Mudannayake, Oshan & Rathnayake, Disni & Herath, Jerome Dinal & Fernando, Dinuni & Fernando, Mgnas. (2022). Exploring Machine Learning and Deep Learning Approaches for Multi-Step Forecasting in Municipal Solid Waste Generation. IEEE Access. PP. 1-1. 10.1109/ACCESS.2022.3221941.

[Figure7] : Yin, Lu & Ma, Pengcheng & Deng, Zhongliang. (2021). JLGBMLoc: A Novel High-precision Indoor Localization Method Based on LightGBM. 10.20944/preprints202103.0215.v1. https://www.researchgate.net/publication/349903999_JLGBMLoc_A_Novel_High-precision_Indoor_Localization_Method_Based_on_LightGBM

[figure 8]: Yao, Xiaotong & Fu, Xiaoli & Zong, Chaofei. (2022). Short-Term Load Forecasting Method Based on Feature Preference Strategy and LightGBM-XGboost. IEEE Access. 10. 1-1. 10.1109/ACCESS.2022.3192011. https://www.researchgate.net/publication/362100649_Short-term_load_forecasting_method_based_on_feature_preference_strategy_and_LightGBM-Xgboost

[Figure 9]: Mohamed, Marwa & Khafagy, Mohamed & Ibrahim, Mohamed. (2019). Recommender Systems Challenges and Solutions Survey. 10.1109/ITCE.2019.8646645. https://www.researchgate.net/publication/331063850_Recommender_Systems_Challenges_and_Solutions_Survey

[Figure 10]: Tondji, Lionel. (2018). Web Recommender System for Job Seeking and Recruiting. 10.13140/RG.2.2.26177.61286. https://www.researchgate.net/publication/323726564_Web_Recommender_System_for_Job_Seeking_and_Recruiting

[figure 11] : Krupa Patel, Hiren B. Patel, in Computers and Electronics in Agriculture, 2020 https://ars.els-cdn.com/content/image/1-s2.0-S016816992031629X-gr6.jpg

[Figure12]: Huiling, Yang & Jing, Jiang. (2022). A Personalized Recommendation Technique for Travel Route Based on Fuzzy Consistent Matrix. Mobile Information Systems. 2022. 1-10. 10.1155/2022/3974382.https://www.researchgate.net/publication/362671852_A_Personalized_Recommendation_Technique_for_Travel_Route_Based_on_Fuzzy_Consistent_Matrix

[Figure 13]: A Personalized Recommendation Technique for Travel Route Based on Fuzzy Consistent Matrix - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Flow-chart-for-the-demographic-based-recommendation-system_fig4_362671852 [accessed 3 May, 2024]

[Figure14]: A Framework for Developing Context-aware Systems - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Architecture-of-Context-Aware-Systems_fig1_287348656 [accessed 3 May, 2024]

[Figure 15]: https://www.geeksforgeeks.org/name-the-factors-influencing-agriculture/

[Figure 16]: Dasari, Anantha Reddy & Dadore, Bhagyashri & Watekar, Aarti. (2019). Crop Recommendation System to Maximize Crop Yield in Ramtek region using Machine Learning. International Journal of Scientific Research in Science and Technology. 485-489. 10.32628/IJSRST196172/https://www.researchgate.net/publication/331426761_Crop_Recommendation_System_to_Maximize_Crop_Yield_in_Ramtek_region_using_Machine_Learning

[Figure17]: Kumar Rana, Arun & Salau, Ayodeji & Gupta, Swati & Arora, Sandeep. (2018). A Survey of Machine Learning Methods for IoT and their Future Applications. https://www.researchgate.net/publication/330702663_A_Survey_of_Machine_Learning_Methods_for_IoT_and_their_Future_Applications

Business Model Canvas

البطاقة التقنية للمشروع

الاسم و اللقب Votre prénom et nom Your first and last Name	ميمون عبد العزيز – العنتري فاروق Abdelaziz MIMOUNE – Farouk LANTRI
الاسم التجاري للمشروع Intitulé de votre projet Title of your Project	PerfectSeed
الصفة القانونية للمشروع Votre statut juridique Your legal status	SARL
رقم الهاتف Votre numéro de téléphone Your phone number	0541998785 0662970136
البريد الالكتروني Votre adresse e-mail Your email address	azan183461@gmail.com faroukantri71@gmail.com
مقر مزاولة النشاط (الولاية- البلدية) Votre ville ou commune d'activité Your city or municipality of activity	تيارت – تيارت Tiaret – Tiaret

طبيعة المشروع

إنه نظام توصية مصمم خصيصًا للزراعة واختيار المحاصيل. يستفيد من تحليل البيانات والنمذجة وتقنيات التعلم الآلي المحتملة لتقديم توصيات قائمة على البيانات للمزارعين بشأن المحاصيل الأكثر ملاءمة للزراعة بناءً على عوامل مثل ظروف التربة والمناخ وتوافر المياه والمتغيرات البيئية والجغرافية الأخرى ذات الصلة الخاصة بهم الأراضي الزراعية.	

القيمة المقترحة أو العرض المقدّم Value Proposition

تحديد المشكل الذي يواجهه الزبون

ما هي المشكلة التي تريد حلها؟	غالباً ما يواجه المزارعون تحديات في تحديد المحاصيل الأكثر ملاءمة للتربة والمناخ وغيرها من العوامل في منطقتهم.
ما هي البيانات المتوفرة لديك التي تدل على وجود المشكلة المحددة؟	بيانات غلة المحاصيل، بيانات مسح التربة، بيانات المناخ، استبيانات المزارعين، بيانات فشل المحاصيل أو التخلي عنها
ما هي المشاريع الأخرى التي استهدفت نفس المشكلة والتي جرى تنفيذها؟	HARVESTIFY
ماهي أهداف مشروعك و/أو نتائجه المتوقعة؟	مساعدة المزارعين في اختيار أنواع المحاصيل المناسبة لزراعتها في أراضيهم وفي ظل الظروف البيئية المحيطة بهم.

046 25 61 33
incubator@univ-tiaret.dz
Ibn khaldoun University

القيمة المقترحة وفق المعايير التالية

القيمة المبتكرة أو الجديدة	يمكن لنظام التوصية بالمحاصيل المبني على التحليلات والبيانات المتاحة أن يساعد المزارعين في اتخاذ قرارات أفضل بشأن أنواع المحاصيل المزروعة وتوقيت الزراعة والعناية بالمحاصيل، مما يؤدي إلى زيادة الإنتاجية وتحسين جودة المنتج الزراعي. باستخدام تحليلات البيانات والتوصيات الدقيقة، يمكن للمزارعين تقليل مخاطر الفشل في الإنتاج بسبب الظروف الجوية أو الآفات الزراعية يمكن للمشروع أن يعزز التنمية المستدامة للزراعة في الجزائر من خلال تحسين إدارة الموارد والحفاظ على التنوع البيولوجي والحفاظ على البيئة
القيمة بالتخصيص	زيادة العائد المالي للمزارعين من خلال زيادة الإنتاجية وتقليل التكاليف الإنتاجية، مما يساهم في تحقيق عوائد مالية مرتفعة. الحفاظ على الموارد الطبيعية والبيئة، وتقليل التلوث البيئي من خلال تحسين استخدام المياه والمبيدات والأسمدة تحسين معيشة المزارعين والمجتمعات المحلية من خلال زيادة الدخل وتوفير فرص عمل وتعزيز التنمية المستدامة. قدرة المشروع على تقديم حلول مبتكرة وتكنولوجيا متقدمة تميزه عن منافسيه وتجعله مفضلًا لدى المزارعين والجهات المهتمة تأثير إيجابي على الاقتصاد المحلي والمجتمعات المحلية من خلال دعم الاستثمار وخلق فرص العمل وتعزيز التنمية الاقتصادية والاجتماعية
القيمة بالسعر	يحصل المزارعون على قيمة عالية مقابل تكلفة النظام من خلال زيادة الإنتاجية وتقليل التكاليف الإنتاجية، مما يساهم في زيادة الربحية وتحقيق عوائد مالية مجزية توفير قيمة بالسعر عن طريق تقديم مزايا تنافسية مثل دقة التوصيات وسهولة الاستخدام والتكامل مع أنظمة أخرى، مما

	يجعله خيارًا مفضلًا للمزارعين على المنافسين.
القيمة بالتصميم	تصميم النظام يوفر تجربة استخدام سهلة ومريحة للمزارعين، مما يزيد من فعالية الاستخدام ويحسن نتائج التوصيات تعزيز التفاعل والتواصل بين المستخدمين والمزارعين، مما يسهل استيعاب التوصيات وتطبيقها بنجاح. تصميم النظام يتماشى مع بيئة الزراعة واحتياجات المزارعين، مما يعزز قبولهم واستخدامهم للنظام بشكل فعال.
القيمة بالأداء العالي	وصيات دقيقة وفعالة بناءً على تحليلات دقيقة للبيانات، مما يساعد المزارعين في اتخاذ قرارات مستنيرة وتحقيق أقصى قدر من الإنتاجية يادة كفاءة استخدام الموارد الزراعية وتحسين الإنتاجية، مما يقلل من التكاليف ويزيد من العائد المالي للمزارعين. يتمتع النظام بالقدرة على التحسين المستمر والتطوير بناءً على ردود فعل المستخدمين وتحليلات البيانات، مما يضمن استمرارية الأداء العالي وتلبية احتياجات المزارعين بشكل فعّال.
القيمة بالخدمة الشاملة	يقدم دعمًا فنيًا شاملاً واستشارات متخصصة للمزارعين، سواءً كان ذلك عبر الهاتف أو البريد الإلكتروني أو حتى زيارات ميدانية. توفير دورات تدريبية وتعليمية شاملة للمزارعين حول كيفية استخدام النظام بشكل فعال وكيفية تطبيق التوصيات الزراعية تحديثات مستمرة وتطوير لميزات جديدة وتحسينات استنادًا إلى ردود فعل المزارعين والتطورات التكنولوجية.
القيمة المبتكرة أو الجديدة	يعتمد النظام على أحدث التقنيات مثل تعلم الآلة والذكاء الاصطناعي لتحليل البيانات وتوصية المحاصيل بطرق مبتكرة ودقيقة. يقدم توصيات مبتكرة ومخصصة لكل مزرعة بناءً على البيانات الخاصة بها وظروفها الفريدة، مما يزيد من قيمة النظام للمزارعين قدم تنبؤات دقيقة للظروف المناخية والأرضية والمحاصيل، مما يمكن المزارعين من اتخاذ قرارات مستنيرة وتحسين إدارة مزارعه

	دعم المزارعين في اتخاذ قرارات مبتكرة لزراعة المحاصيل بشكل مستدام وصديق للبيئة، مما يعزز الاستدامة في الزراعة وحماية الموارد الطبيعية.
قيم أخرى	دعم التنوع الزراعي والتكامل بين المحاصيل المختلفة، مما يسهم في تحسين صحة التربة وتقليل مخاطر الآفات والأمراض يوفر شفافية في عملياته وبياناته، مما يسهل على المزارعين فهم القرارات المتخذة والمساءلة عنها.

Customer Segments شرائح العملاء أو الزبائن

Géographique الجغرافية	Démographique (B2C)	Démographique (B2B)	Psychographique	Comportemental السلوكيات

046 25 61 33
incubator@univ-tiaret.dz
Ibn khaldoun University

			العوامل النفسية و الشخصية	
Continent القارة افريقيا	Age العمر جميع الأعمار	Secteur القطاع الفلاحة و الزراعة	Classe sociale الطبقة الاجتماعية كل الطبقات	Usage استخدام الاستخدام عند الحاجة
Pays الدولة الجزائر	Sexe الجنس كلا الجنسين	Nombre d'employés عدد العمال في القطاع المؤسسات المتوسطة و الصغيرة	Niveau de vie المستوى المعيشي لكل المستويات	Loyauté الوفاء الوفاء غير مطلوب متوفر للجميع
Région الجهة المناطق الفلاحية و الزراعية	Revenus annuel متوسط الدخل منخفض إلى متوسط	Maturité de l'entreprise نضج المؤسسة مؤسسات مستقرة و راسخة	Valeurs القيم الاستقلالية, الثقة, الابتكار	Intérêt اهتمام اهتمام كبير بالتكنولوجيا للمساعدة على الاستقلالية

Département الولاية ولايات الهضاب العليا	Etat matrimonial الحالة الاجتماعية كل الحالات الاجتماعية	Situation financière الحالة المالية للمؤسسة حالة مالية قوية قادرة على الاستثمار في التكنولوجيا المساعدة	Personnalité الشخصية مستقل, متفتح على التكنولوجيا	Passion الهواية و شغف شغف بتحسين الظروف الزراعية للفلاحين
Ville الدائرة او البلدية المناطق الريفية	Niveau d'étude المستوى الدراسي كل المستويات	Détention/ actionnariat الملكية/المساهمة مساهمة محلية من شركات خاصة	Convictions المعتقدات معتقدات متفتحة تجاه الابتكار	Sensibilité حساسيات تحفظات تجاه غير الفلاحين
Quartier الحي كل الأحياء	Profession المهنة فلاح أو مزارع	Valorisation/ capitalisation boursière التقييم / القيمة السوقية تقدير قيمة العملاء بناء على إنفاقهم المتكرر و تحديد	Présence digitale et sur les réseaux sociaux استعمال التكنولوجيا في التواصل وجود دائم على مواقع التواصل	Habitude de consommation عادة الاستهلاك بعد الاستخدام الأول و ملاحظة ال فرق في نمو المحاصيل

		الحصة السوقية المستهدفة	الاجتماعي للتواصل مع العملاء	ازدياد الاستخدام و الاشتراك في الباقات المدفوعة
Climat المناخ	Culture الثقافة	Business model نموذج الأعمال	Centres d'intérêts مراكز الاهتمام التعلم الآلي و مساعدة الفلاحين	Mode de paiement طرق الدفع الدفع الالكتروني عن طريق بطاقات الائتمان
	Religion الدين جميع الأديان	Secteur servi القطاع الذي يخدمه الفلاحة و الزراعة		Connaissance المعرفة معرفة متوسطة بالتكنولوجيا
	Langue اللغة الانجليزية	Technologie utilisée التكنولوجيا المستعملة التعلم الآلي		Nature de la demande طبيعة الطلب طلب لمرة واحدة أو عدة مرات.

Fréquence d'achat عدد مرات الطلب على السلعة مرة واحدة		Format du produit ou packaging شكل المنتج أو التعبئة والتغليف موقع الكتروني		

قنوات التوزيع Channels

المبيعات المباشرة	بيع المنتجات مباشرة إلى العملاء عبر مواقع التواصل الاجتماعي أو من خلال تواصل مباشر مع العملاء المحتملين أو استهداف منشئات فلاحية و زراعية
تجار الجملة	التعاون مع تجار الجملة الذين يمكنهم شراء كميات كبيرة من المنتج وبيعها للموزعين وتجار التجزئة.
الموزعون	تعيين مراكز موزعين محليين متخصصين في مجال الفلاحة و الزراعة حيث يقومون بتوزيع منتجاتنا إلى العملاء النهائيين والمستخدمين في المناطق المحلية والمجتمعات.
توزيع التجزئة	توزيع المنتج من خلال متاجر التجزئة المتخصصة في مجال الرعاية الصحية و المتاجر الإلكترونية, مما يسهل على العملاء العثور على منتجاتنا وشرائها أثناء قيامهم بزيارة هذه المتاجر.

العلاقة مع العملاء Customer Relationship

كيف تدير علاقاتك مع العملاء؟	ندير علاقاتنا مع العملاء من خلال تقديم خدمة عملاء ممتازة، تتضمن دعمًا فنيًا شاملاً، وتواصلًا مستمرًا عبر قنوات متعددة لضمان تلبية احتياجاتهم واستفساراتهم بشكل فعال. نقوم أيضًا بجمع وتحليل ملاحظات العملاء لتحسين منتجاتنا وخدماتنا باستمرار، وضمان رضاهم وتعزيز ولائهم.
ماهية أهم البرامج التي ستعتمد عليها في ادارة العلاقة مع الزبون Microsoft Dynamics Monday CRM Zoho CRM الخ............................	عمل المتجر الإلكتروني: ندير المتجر الإلكتروني باستخدام منصات التجارة الإلكترونية المتقدمة التي تتيح للعملاء تصفح المنتجات، إجراء عمليات الشراء بسهولة، وتقديم الدعم اللازم عبر الإنترنت. كما نوفر بوابات دفع آمنة وخيارات شحن متنوعة لتحسين تجربة العملاء على المتجر الإلكتروني.

الشركاء الأساسيون Key Partner

الشركاء	معلومات حول الشركاء	طبيعة الشراكة
الشريك الأول	اخصائي فلاحة أو زراعة	مراقبة وتحليل البيانات الفلاحية/الزراعية لضمان دقة المعلومات المستخدمة في النظام.
الشريك الثاني	أخصائي التعلم الآلي	تعزيز وتحسين نموذج التعلم الآلي.
الشريك الثالث	أخصائيي الاختبار	ضمان سهولة استخدام النموذج الأولي.

هيكل التكاليف Costs structure

تكاليف التعريف بالمنتج أو المؤسسة Frais d'établissement	40,000,000 دينار جزائري
تكاليف الحصول على العدادات (الماء – الكهرباء ....) Frais d'ouverture de compteurs (eaux–gaz–....)	كلفة اجهزة الاستشعار : 6700 دينار جزائري للعنصر
تكاليف (التكوين– برامج الاعلام الالي المختصة) Logiciels, formations	–قد تكون تكلفة التدريب للمزارعين على استخدام هذه البرامج حوالي 20000 دينار جزائري إلى 50000 دينار جزائري للدورات التدريبية المكثفة -قد تتراوح تكلفة البرمجيات المخصصة للزراعة الدقيقة بين 10000 دينار جزائري إلى 50000 دينار جزائري أو أكثر حسب الميزات والقدرات المطلوبة.
Dépôt marque, brevet, modèle تكاليف براءة الاختراع و الحماية الصناعية و التجارية	–تتراوح تكلفة هذه الخدمة بين 10000 دينار جزائري إلى 30000 دينار جزائري أو أكثر، حسب مجال البحث ومدى التحقق المطلوب . مكن أن تكون هناك تكاليف إضافية للبحث والتحقق للتأكد من عدم وجود اختراعات أو علامات تجارية مشابهة مسجلة بالفعل
Droits d'entrée تكاليف الحصول على تكنولوجيا او ترخيص استعمالها	-5000دينار جزائري إلى 20000 دينار جزائري. -قد تتطلب الشروط الإضافية مثل رسوم الترخيص السنوية أو النسبة المئوية من الإيرادات تكاليف إضافية. يمكن أن تبلغ تكاليف التفاوض والشروط الإضافية ما بين 20000 دينار جزائري إلى 50000 دينار جزائري أو أكثر.
Achat fonds de commerce ou parts شراء الأصول التجارية أو الأسهم	–تتراوح تكاليف شراء الأصول التجارية بين 50000 دينار جزائري إلى 500000 دينار جزائري أو أكثر –تكاليف شراء الأسهم بحوالي 10000 دينار جزائري وتزيد حسب قيمة الشركة وعدد الأسهم المطلوبة للشراء.

–تتراوح تكلفة الإيجار الشهري للحق في الإيجار لمحل تجاري بين 50000 دينار جزائري إلى 300000 دينار جزائري أو أكثر.	Droit au bail الحق في الإيجار
–إذا كان إيجار العقار 50،000 دينار جزائري شهرياً، فقد تكون تكلفة الوديعة ما بين 100،000 دينار جزائري إلى 150،000 دينار جزائري.	Caution ou dépôt de garantie وديعة أو وديعة تأمين
–تتراوح رسوم إيداع الملفات عادة بين 5000 دينار جزائري إلى 20000 دينار جزائري.	Frais de dossier رسوم إيداع الملفات
–تتراوح تكلفة خدمات الموثق عادةً بين 5000 دينار جزائري إلى 30000 دينار جزائري -	Frais de notaire ou d'avocat تكاليف الموثق–المحامي–.........
–تتراوح تكاليف استئجار خدمات محامي عادةً بين 10000 دينار جزائري إلى 50000 دينار جزائري أو أكثر، حسب طبيعة القضية وتعقيدها وسمعة المحامي	Enseigne et éléments de communication تكاليف التعريف بالعلامة و تكاليف قنوات الاتصال
–في المدن الكبرى مثل الجزائر العاصمة أو وهران، قد تتراوح أسعار العقارات للمتر المربع ما بين 50000 دينار جزائري إلى 200000 دينار جزائري أو أكثر.	Achat immobilier شراء العقارات
–تتراوح تكلفة الأعمال البنائية بين 20000 دينار جزائري إلى 100000 دينار جزائري أو أكثر حسب حجم العقار ونطاق الأعمال.	Travaux et aménagements الأعمال والتحسينات الاماكن

Matériel الآلات– المركبات– الاجهزة	1– **المعدات والآلات**: تتراوح تكلفة شراء المعدات والآلات بين 5,000,000 دج إلى 50,000,000 دج أو أكثر، حسب نوع المعدات وحجمها وجودتها. 2-المركبات: تختلف تكاليف شراء المركبات بين 2,000,000 دج إلى 100,000,000 دج أو أكثر، اعتمادًا على النوع والحجم والعمر 3الأجهزة والتقنيات: تتراوح تكاليف شراء الأجهزة والتقنيات بين 50,000 دج إلى 500,000 دج أو أكثر، اعتمادًا على نوع وجودة الأجهزة.
Matériel de bureau تجهيزات المكتب	1 :أثاث المكتب مكاتب، كراسي، خزائن، وحدات تخزين، وأرفف قد تتراوح تكاليف شراء أثاث المكتب بين 100,000 دج إلى 500,000 دج أو أكثر، اعتمادًا على الجودة والتصميم 2- الأجهزة والتقنيات أجهزة الكمبيوتر، الطابعات، الماسحات الضوئية، الفاكسات، الهواتف، الشاشات، وأجهزة العرض. تتراوح تكاليف شراء الأجهزة والتقنيات بين 50,000 دج إلى 300,000 دج أو أكثر، حسب النوع والمواصفات **مستلزمات المكتب**: ورق، أقلام، دفاتر، ملاحظات، دبابيس، مشابك، ملفات، وأدوات كتابة. تختلف تكاليف شراء مستلزمات المكتب بحسب الكمية والجودة، ويمكن أن تتراوح بين 10,000 دج إلى 50,000 دج. .
Stock de matières et produits تكاليف التخزين	تتراوح تكاليف الإيجار بين 20,000 دج إلى 100,000 دج شهريًا أو أكثر يمكن أن تبلغ تكاليف التشغيل بشكل عام بين 10,000 دج إلى 50,000 دج شهريًا أو أكثر يتطلب تأمين المستودعات والمخزون تكاليف إضافية قد تتراوح تكاليف التأمين بين 5,000 دج إلى 20,000 دج شهريًا أو أكثر يجب أخذ التكاليف المتعلقة بصيانة المخزون وترتيبه في الاعتبار. قد تتراوح هذه التكاليف بين 5,000 دج إلى 30,000 دج شهريًا

Trésorerie de départ التدفق النقدي(الصندوق) الذي تحتاجه في بداية المشروع.	تشمل تكاليف البدء والإنشاء تكلفة شراء المعدات والموارد اللازمة، وتأجير المكان، وتكاليف التسويق الأولي والترويج للمشروع. قد تتراوح تكاليف البدء بين 1,000,000 دج إلى 10,000,000 دج أو أكثر يجب أن يتم تخصيص تدفق نقدي كافٍ لتغطية تكاليف التشغيل الشهرية الأولى، مثل الإيجار، والرواتب، والفواتير المتعلقة بالخدمات والمواد اللازمة. يمكن أن تتراوح هذه التكاليف بين 500,000 دج إلى 5,000,000 دج شهريًا يفضل تخصيص جزء من التدفق النقدي كاحتياطي للطوارئ، لتغطية أي مصروفات غير متوقعة أو تأخير في تحقيق الدخل المتوقع.

المجموع =يتراوح الإجمالي المطلوب من التدفق النقدي في بداية المشروع من حوالي 1,650,000 دج إلى 16,500,000 دج

نفقاتك أو التكاليف الثابتة الخاصة بمشروعك

Assurances التأمينات	تتضمن تأمينات العاملين تأمينات صحية واجتماعية وتأمين ضد الحوادث الشخصية والإصابات في مكان العمل. يمكن أن تتراوح تكاليف تأمين العاملين بين 2%إلى10% من الرواتب الإجمالية للموظفين يتطلب تأمين الممتلكات والمعدات والمركبات تحديد قيمة الأصول وتحديد المخاطر المحتملة. يمكن أن تتراوح تكاليف التأمين بين 0.5% إلى 3% من قيمة الأصول يغطي تأمين المسؤولية المدنية الحوادث التي يسببها المشروع للغير. يمكن أن تتراوح تكاليف التأمين بين 0.5% إلى 2% من الإيرادات السنوية المتوقعة للمشروع تضمن تأمين ضد الفقدان والتلف تغطية للمخزون والممتلكات ضد السرقة والحرائق والكوارث الطبيعية. يمكن أن تتراوح تكاليف التأمين بين 0.1% إلى 1% من قيمة الممتلكات
Téléphone, internet الهاتف و الانترنت	تراوح تكاليف الهاتف الثابت بين 2,000 دج إلى 10,000 دج شهريًا تتراوح تكاليف الهاتف المحمول بين 2,000 دج إلى 20,000 دج شهريًا، حسب نوع الخطة وشركة الاتصالات. تتضمن تكاليف الإنترنت رسوم الاشتراك الشهرية وسرعة الاتصال والحجم المسموح به للبيانات. قد تتراوح تكاليف الإنترنت بين 3,000 دج إلى 30,000 دج شهريًا
Autres abonnements اشتراكات أخرى	شمل هذه الاشتراكات تكاليف استخدام البرمجيات والتطبيقات عبر الإنترنت، مثل خدمات التخزين السحابي وبرامج الإنتاجية وخدمات إدارة العلاقات مع العملاء تراوح تكاليف هذه الاشتراكات بين 5,000 دج إلى 50,000 دج شهريًا، حسب عدد المستخدمين ونطاق الخدمات المطلوبة
Carburant, transports الوقود و تكاليف النقل	تتراوح تكاليف الوقود للسيارات بين 20,000 دج إلى 100,000 دج شهريًا، حسب عدد الكيلومترات التي يتم قطعها وكفاءة الوقود للسيارات شمل هذه التكاليف تكاليف الصيانة الدورية للسيارات وإصلاحات الطوارئ واستبدال القطع التالفة. تتراوح تكاليف الصيانة والإصلاح بين 5,000 دج إلى 50,000 دج شهريًا، حسب حالة السيارات وتكاليف القطع البديلة

تشمل تكاليف النقل تكاليف استئجار السيارات أو استخدام وسائل النقل العامة أو تكاليف الوقود والصيانة للسيارة الخاصة تراوح تكاليف النقل بين 5,000 دج إلى 50,000 دج شهريًا، حسب عدد المسافات المسافرة ونوع وسيلة النقل المستخدمة تتراوح تكاليف الإقامة بين 10,000 دج إلى 100,000 دج شهريًا، حسب مدة الإقامة ونوع الإقامة المختارة والموقع	Frais de déplacement et hébergement تكاليف التنقل و المبيت
تتراوح تكاليف فواتير الماء بين 5,000 دج إلى 50,000 دج شهريًا، حسب حجم العمل واحتياجات المياه. تتراوح تكاليف فواتير الكهرباء بين 10,000 دج إلى 100,000 دج شهريًا، حسب حجم الاستهلاك وأسعار الكهرباء. تتراوح تكاليف فواتير الغاز بين 5,000 دج إلى 50,000 دج شهريًا، حسب حجم الاستهلاك وأسعار الغاز.	Eau, électricité, gaz فواتير الماء – الكهرباء – الغاز
تراوح تكاليف الاشتراك الشهرية بين 2,000 دج إلى 20,000 دج شهريًا، حسب نوع التغطية وتكاليف الخدمات المقدمة تتراوح تكاليف الخدمات الطبية بين 5,000 دج إلى 50,000 دج شهريًا، حسب نوع التغطية وتكاليف الخدمات الطبية المطلوبة	Mutuelle التعاضدية الاجتماعية
تشمل هذه التكاليف شراء الأقلام، والأوراق، والمجلدات، والدباسات، والأشياء الأخرى التي تحتاجها لتشغيل المكتب. تتراوح تكاليف لوازم المكتب بين 5,000 دج إلى 30,000 دج شهريًا، حسب حجم العمل واحتياجات المكتب. شراء الأدوات الصغيرة مثل الشريط اللاصق، والمشابك، والفرش، والمناديل، وغيرها من اللوازم التي يتم استخدامها بشكل متكرر تراوح تكاليف الأدوات الصغيرة والمستهلكات بين 3,000 دج إلى 20,000 دج شهريًا، حسب حجم العمل واستخدام هذه اللوازم	Fournitures diverses لوازم متنوعة
تشمل هذه التكاليف تكاليف صيانة المعدات والأجهزة المستخدمة في العمل مثل الحاسوب، والطابعة، والآلات الأخرى. تتراوح تكاليف صيانة المعدات بين 5,000 دج إلى 50,000 دج شهريًا تكاليف غسيل وإصلاح الملابس المستخدمة في العمل مثل الزي الرسمي أو الملابس الخاصة بالعمل اليومي تتراوح تكاليف صيانة الملابس بين 3,000 دج إلى 20,000 دج شهريًا، حسب عدد الأفراد وتكاليف الغسيل والإصلاح	Entretien matériel et vêtements صيانة المعدات والملابس
تتراوح تكاليف تنظيف المباني بين 10,000 دج إلى 50,000 دج شهريًا لكل 100 متر مربع، حسب مستوى النظافة المطلوبة ومدى التلوث.	Nettoyage des locaux تنظيف المباني

Budget publicité et communication ميزانية الإعلان والاتصالات	10,000دج إلى 100,000 دج، بينما تكون التكاليف الوطنية أعلى بكثير

المجموع =من 1,000,000دج إلى 10,000,000دج

مصادر الإيرادات Revenue Stream

Apport personnel ou familial المساهمة الشخصية أو العائلية	تتراوح قيمة المساهمة النقدية بين 50,000 دج إلى 500,000 دج أو أكثر
Apports en nature (en valeur) التبرعات العينية	قد يكون قيمة الخدمات المهنية المتبرع بها حوالي 20,000 دج. قد تشمل هذه التبرعات معدات أو أصول مثل الحواسيب، والآلات، والأثاث، والمركبات، وغيرها و تكون قيمة معدات المكتب حوالي 50,000 دج.
Prêt n°1 (nom de la banque) قرض رقم 1 اسم البنك	بنك الجزائر الخارجي
Prêt n°2 (nom de la banque) قرض رقم 2 اسم البنك	البنك الوطني الجزائري(BNA)
Prêt n°3 (nom de la banque) قرض رقم3 اسم البنك	الائتمان الشعبي الجزائري(CPA)
Subvention n°1 (libellé) منحة 1	
Subvention n°2 (libellé) منحة 2	
Autre financement (libellé) تمويل آخر	

المجموع = من 100,000دج إلى 550,000دج

رقم الأعمال

بيع المنتج في السنة الأولى Votre chiffre d'affaires de la première année

بيع المنتج في السنة الأولى	متوسط أيام العمل في الشهر
الشهر 1Mois	20
الشهر 2Mois	20
الشهر 3Mois	20
الشهر 4Mois	20
الشهر 5Mois	20
الشهر 6Mois	20
الشهر 7Mois	20
الشهر 8Mois	20
الشهر 9Mois	20
الشهر 10Mois	20
الشهر 11Mois	20
الشهر 12Mois	20

المجموع =

النسبة المئوية للزيادة في حجم الأعمال بين كل شهر لسنة الأولى؟

بيع المنتج في السنة الثانية Votre chiffre d'affaires de la deuxième année

بيع المنتج في السنة الثانية	متوسط أيام العمل في الشهر
الشهر 1Mois	20
الشهر 2Mois	20
الشهر 3Mois	20
الشهر 4Mois	20
الشهر 5Mois	20
الشهر 6Mois	20
الشهر 7Mois	20

20	الشهر 8Mois
20	الشهر 9Mois
20	الشهر 10Mois
20	الشهر 11Mois
20	الشهر 12Mois

المجموع =

النسبة المئوية للزيادة في حجم الأعمال بين كل شهر لسنة الثانية؟

Votre chiffre d'affaires de la troisième année بيع المنتج في السنة الثالثة

متوسط أيام العمل في الشهر	بيع المنتج في السنة الثالثة
20	الشهر 1Mois
20	الشهر 2Mois
20	الشهر 3Mois
20	الشهر 4Mois
20	الشهر 5Mois
20	الشهر 6Mois
20	الشهر 7Mois
20	الشهر 8Mois
20	الشهر 9Mois
20	الشهر 10Mois
20	الشهر 11Mois
20	الشهر 12Mois

المجموع =

النسبة المئوية للزيادة في حجم الأعمال بين كل شهر لسنة الثالثة ؟

تطور حجم رقم الأعمال في السنة

- النسبة المئوية للزيادة في حجم الأعمال بين السنة 1 والسنة 2؟
- النسبة المئوية للزيادة في حجم الأعمال بين السنة 2 والسنة 3 ؟

متوسط مدة الاعتمادات الممنوحة للعملاء بالأيام Durée moyenne des crédits accordés aux clients en jours	30 يوم
متوسط مدة ديون الموردين بالأيام Durée moyenne des dettes fournisseurs en jours	30 يوم

رواتب الموظفين و مسؤولين الشركة

رواتب الموظفين Salaires employés	مدير المشروع: 100,000 دينار جزائري سنويًا مهندسون/مختصون: 80,000 دينار جزائري سنويًا لكل واحد مسؤول تسويق: 70,000 دينار جزائري سنويًا محاسب: 60,000 دينار جزائري سنويًا •موظفو دعم العملاء: 50,000 دينار جزائري سنويًا لكل واحد
صافي أجور المسؤولين Rémunération nette dirigeant	55,000دج

Business Model Canvas

Designed for:	Designed by:	Date:	Version:
PerfectSeed	MIMOUNE Abdelaziz LANTRI Farouk	09/06/2024	1

Key Partners	Key Activities	Value Propositions	Customer Relationships	Customer Segments
Farming specialist: to monitor and analyze soil data AI specialist: to enhance and improve the Machine Learning Model. Testing specialists: to ensure the usability of the prototype	Improving our algorithms to improve our crop recommendation model. Marketing and advertising our product.	New algorithm: Faster and more efficient analysis of soil data. Ease of use: User friendly web application for crop recommendation.	Full technical support in case of bugs. Customer service through social media platforms. Surveying clients.	Individual Farmers and/or peasants. Farming companies
	Key Resources		**Channels**	
	IoT sensors for soil data. New and improved algorithms for better results. Funds for research. Facilities for development and testing		Direct sales to clients through social media or direct contact Bulk sales to distributors	

Cost Structure	Revenue Streams
-Cost of IoT sensors. -Team salary -Development funds -Marketing and advertisement funds	-Subscription fees -IoT sensor prices -API Licensing prices -Data collection and selling.

046 25 61 33
incubator@univ-tiaret.dz
Ibn khaldoun University