VISVESVARAYA TECHNOLOGICAL UNIVERSITY

"Jnana Sangama", Belagavi, Karnataka– 590 018

A
Report
on

INNOVATION / ENTREPRENEURSHIP / SOCIETAL INTERNSHIP (21INT68)

Submitted for the partial fulfilment of academic requirements for VI Semester

BACHELOR OF ENGINEERING
IN
ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

Submitted by

DISHITHA K S 4CI21AI012

Carried out at

DLITHE, BENGALURU

Under the Internal Guidance of

Mr. MURALIDHAR B M
Assistant Professor
Department of AI&ML

DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

Kodava Education Society®

COORG INSTITUTE OF TECHNOLOGY

(Accredited by NAAC 'A' grade, Approved by the AICTE, New Delhi & Affiliated to VTU, Belagavi)

Ponnampet-571216

2023-24

Kodava Education Society®

COORG INSTITUTE OF TECHNOLOGY

(Accredited by NAAC 'A' grade, Approved by AICTE, New Delhi & Affiliated to VTU, Belagavi)

PONNAMPET - 571216, Kodagu, Karnataka

Department of Artificial Intelligence and Machine Learning

CERTIFICATE

This is to certify that the course entitled "**Innovation / Entrepreneurship / Societal Internship (21INT68)**" is carried out by **DISHITHA K S (4CI21AI012)** bonafide student of **Coorg Institute of Technology** in partial fulfilment of internship work for **Bachelor of Engineering** in **Artificial Intelligence and Machine Learning** under **Visvesvaraya Technological University**, **Belagavi** during the academic year **2023-2024**. The report has been verified to satisfy the academic requirements in respect of work prescribed by the university.

----------------------------	-------------------------------	--------------------------
Prof. Muralidhar B M	**Dr. Ramakrishna B B**	**Dr. M Basavaraj**
Internal Guide	Associate Professor & Head	Principal

INTERNSHIP CERTIFICATE

DLithe®
Be Agile..Be in Demand

Dlithe Consultancy Services Pvt. Ltd.
CIN: U72900KA2019PTC121035

27 November 2023

TO WHOMSOEVER IT MAY CONCERN

This is to certify **Dishitha K S**, bearing USN No: **4CI21AI012** from **Coorg Institute of Technology, Coorg** has successfully completed one-month internship starting from **27-10-2023** to **27-11-2023** under the mentorship of DLithe's development team. **Dishitha K S** has worked on Fundamentals of Python, Introduction to Data Science and Machine learning, Building the Supervised learning models for real-world scenario and deployment, Classification and Regression model implementation, exposer to Artificial Nueral network working, architecture and network parameters.

The domain & agile development process exposure was given along with usage of GitHub tool. During the internship, **Dishitha K S** demonstrated good coding skills with good design thoughts.

We wish all the best for future endeavours!

DLITHE CONSULTANCY SERVICES PVT. LTD. Bengaluru

For Dlithe Consultancy Services Pvt Ltd
Director

Certificate ID: NOV2023AIML7458217

ACKNOWLEDGEMENTS

I express my deep sense of gratitude and indebtedness to the management of "***Coorg Institute of Technology***", Ponnampet which provided me an opportunity to fulfil the requirement of completing the internship.

I express my sincere gratitude *to* ***Dr. M Basavaraj***, ***Principal***, Coorg Institute of Technology, Ponnampet for his kind-hearted cooperation in the completion of the internship.

I feel elated in expressing my profound sense of gratitude to ***Dr. Ramakrishna B B***, ***Associate Professor and Head***, Department of Artificial Intelligence & Machine Learning for his constant encouragement in carrying out the internship.

My profound thanks to ***Prof. Muralidhar B M***, ***Internal Guide*** and Asst. Professor, Department of Artificial Intelligence & Machine Learning for his constant encouragement and valuable guidance in completion of the internship.

I also thank the company ***DLithe, Bengaluru*** for providing me with this golden opportunity to do the internship which helped me in learning new things.

I would like to thank all the teaching and non-teaching staff members of Department of Artificial Intelligence & Machine Learning for their support and suggestions.

DISHITHA K S [4CI21AI012]

ABSTRACT

The crop recommendation system powered by machine learning offers a robust framework for aiding farmers in making informed decisions about crop selection. Harnessing the capabilities of advanced algorithms and data analytics, this system analyzes various agronomic factors including soil properties, climate conditions, historical yield data, and crop characteristics. By integrating these diverse datasets, it generates personalized recommendations tailored to the unique needs and constraints of individual farming contexts. Through predictive modeling techniques, the system forecasts the performance of different crops under specific environmental conditions, enabling farmers to optimize their choices for maximum yield and profitability.

Moreover, the crop recommendation system serves as a valuable tool for promoting sustainable agriculture practices. By considering factors such as soil type, soils pH, and temperature, it encourages the adoption of crops that are well-suited to local conditions, thereby minimizing resource wastage and environmental impact. Additionally, by taking into account market trends and demand forecasts, the system assists farmers in aligning their production strategies with market dynamics, enhancing their competitiveness in the agricultural sector. Overall, the crop recommendation system represents a significant advancement in agricultural decision support technology, offering farmers actionable insights to improve productivity, profitability, and sustainability in their farming operations. Its integration of machine learning algorithms holds immense potential for revolutionizing crop management practices and driving innovation in the agricultural industry.

CONTENTS

LIST OF FIGURES

CHAPTER 1

COMPANY PROFILE

DLithe, headquartered in Bangalore, is a pioneering firm specializing in data science, machine learning, and artificial intelligence solutions. With a steadfast commitment to innovation and excellence, DLithe serves clients across various sectors, aiding them in leveraging data-driven insights to foster business growth and drive innovation. Boasting a team of seasoned data scientists and engineers, DLithe excels in crafting advanced algorithms and predictive models that unearth actionable insights from intricate datasets. The company's service portfolio encompasses a wide array of offerings, including data analytics, machine learning solutions, AI consulting, and comprehensive training programs.

DLithe's approach is distinguished by its emphasis on tailoring solutions to suit the unique needs and challenges of each client, ensuring maximum impact and value creation. By combining domain expertise with cutting-edge technologies, DLithe empowers organizations to unlock the full potential of their data assets and stay ahead in today's dynamic business landscape. Committed to delivering excellence and fostering long-term partnerships, DLithe continues to be a trusted advisor and solution provider for companies seeking to capitalize on the transformative power of data and AI.

CHAPTER 2

TASKS PERFORMED

During a 30-day internship at DLithe focusing on machine learning, participants engage in a comprehensive learning journey designed to develop practical skills and knowledge in the field. The internship typically involves a mix of theoretical learning, hands-on projects, and mentorship from experienced data scientists. Interns may begin by gaining a solid understanding of fundamental machine learning concepts such as supervised and unsupervised learning, regression, classification, and clustering.

As the internship progresses, interns are likely to delve into more advanced topics and techniques, such as deep learning, neural networks, and model evaluation. They may work on real-world projects, applying machine learning algorithms to solve practical problems and gain practical experience in data preprocessing, feature engineering, and model tuning.

Throughout the internship, participants receive guidance and feedback from mentors, allowing them to refine their skills and build a strong foundation in machine learning. By the end of the program, interns are equipped with the knowledge and experience to tackle complex machine learning challenges and make meaningful contributions in the field.

2.1 PROBLEM STATEMENT

Implement crop recommendation system using Machine Learning by using data sets such as pH, temperature and soil type of a particular agricultural area.

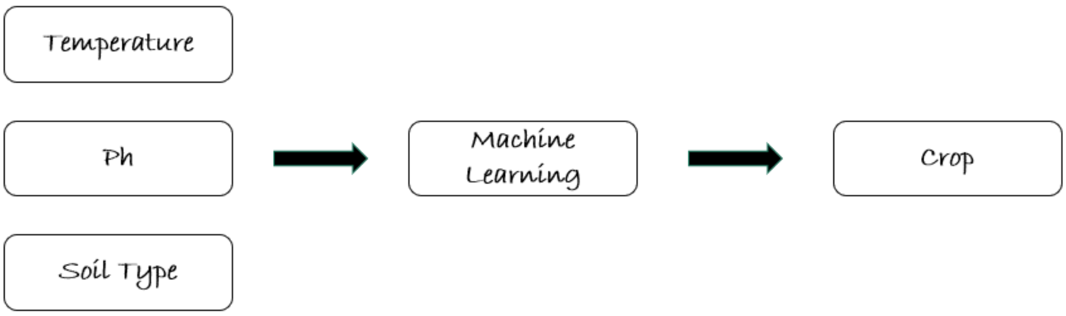

Figure 2.1 : Implementing the given problem statement

2.2 PROPOSED SYSTEM

The proposed crop recommendation system utilizing machine learning (ML) aims to revolutionize agricultural decision-making by providing personalized and accurate crop suggestions to farmers. Here's an outline of the proposed system:

1. Data Collection: Gather comprehensive datasets including soil characteristics, climate data, historical crop yields, crop growth patterns, and farmer preferences.

2. Preprocessing: Clean, preprocess, and integrate the collected data to ensure consistency and quality. This involves handling missing values, outlier detection, and normalization.

3. Feature Engineering: Extract relevant features from the data and engineer new features that capture important relationships between variables, such as soil fertility indices and climate suitability metrics.

4. Model Development: Experiment with a variety of ML techniques, including regression, classification, and ensemble methods, to develop predictive models. Consider algorithms such as decision trees, random forests, gradient boosting machines, and deep neural networks.

5. Personalization: Incorporate a personalized recommendation engine that accounts for individual farmer preferences, land characteristics, and socio-economic factors to tailor crop suggestions.

6. Validation and Testing: Evaluate the performance of the developed models using validation techniques such as cross-validation and holdout validation. Test the system on unseen data to assess its generalization capabilities.

7. Deployment: Integrate the developed crop recommendation system into user-friendly platforms accessible via web or mobile applications. Ensure scalability, reliability, and real-time updates to accommodate changing environmental conditions.

8. Continuous Improvement: Implement mechanisms for feedback collection from farmers and monitoring of model performance. Continuously update and refine the system based on new data and user feedback to enhance accuracy and relevance.

By leveraging the power of ML algorithms and comprehensive datasets, the proposed crop recommendation system aims to empower farmers with actionable insights, optimize crop selection, and improve agricultural productivity and sustainability.

2.3 SOFTWARE CONFIGURATION

- Coding Language: Python
- Tools:
    1. Pandas
    2. Numpy
    3. Tensorflow
    4. Sickitlearn

2.4 IMPLEMENTATION

2.4.1 DATA ANALYSIS

One of the first steps we perform during implementation is an analysis of the data. This was done by us in an attempt to find the presence of any relationships between the various attributes present in the dataset.

Acquisition of Training Dataset: The accuracy of any machine learning algorithm depends on the number of parameters and the correctness of the training dataset. In this project we analysed mutiple datasets given to us from our trainee and carefully selected the parameters that would give the best results. Many work done in this field have considered environmental parameters to predict crop sutainability some have used yield as major factor where as in some works only economic factors are taken into consideration. We have tried to combine both environmental parameters like temperature,ph in soil, soil type to provide accurate and reliable recommendation to the farmer on which crop will be most suitable for his land. The crops recommended in our dataset include rice, maize, soyabeans, beans, peas, groundnut, cowpeas, banana, mango, grapes, watermelon, apple, orange, cotton and coffee.

	A	B	C	D
1	temperature	Soil Type	Ph	Crop Recommendation
2	20.87974371	loam	6.502985292	rice
3	21.77046169	gravels	7.038096361	rice
4	23.00445915	clay	7.840207144	rice
5	26.49109635	alluvium	6.980400905	rice
6	20.13017482	red loamy	7.628472891	rice
7	23.05804872	red sandy	7.073453503	rice
8	22.70883798	black	5.70080568	rice
9	20.27774362	medium black	5.718627178	rice
10	24.51588066	shallow black	6.685346424	rice
11	23.22397386	loam	6.336253525	rice
12	26.52723513	gravels	5.386167788	rice
13	23.97898217	red sandy	7.50283396	rice
14	26.80079604	black	5.108681786	rice
15	24.01497622	red sandy	6.98435366	rice
16	25.66585205	shallow black	6.94801983	rice
17	24.28209415	loam	7.042299069	rice
18	21.58711777	gravels	6.249050656	rice
19	23.79391957	clay	6.970859754	rice
20	21.8652524	alluvium	5.953933276	rice
21	23.57943626	red loamy	5.85393208	rice
22	21.32504158	red sandy	6.442475375	rice
23	25.15745531	black	5.070175667	rice
24	21.94766735	medium black	6.012632591	rice
25	21.0525355	shallow black	6.254028451	rice
26	23.48381344	loam	7.375482851	rice
27	25.0756354	gravels	7.778915154	rice
28	26.35927159	red sandy	6.286500176	rice
29	24.52922681	black	7.070959995	rice
30	20.77576147	red sandy	6.244841491	rice
31	22.30157427	shallow black	6.043304899	rice
32	21.44653958	loam	5.824709117	rice
33	22.17931888	gravels	6.357389366	rice
34	24.52783742	red sandy	6.364134968	rice
35	20.26707606	black	5.01450727	rice

Figure 2.2 Data Set

CHAPTER 3

THEORETICAL BACKGROUND

3.1 OVERVIEW ON MACHINE LEARNING

Machine learning is an application of artificial intelligence (AI) that gives systems the ability to automatically learn and evolve from experience without being specially programmed by the programmer. The process of learning begins with observations or data, such as examples, direct experience, or instruction, in order to look for patterns in data and make better decisions in the future based on the examples that we provide. The main aim of machine learning is to allow computers to learn automatically and adjust their actions to improve the accuracy and usefulness of the program, without any human intervention or assistance. Traditional writing of programs for a computer can be defined as automating the procedures to be performed on input data in order to create output artifacts. Almost always, they are linear, procedural and logical. A traditional program is written in a programming language to some specification, and it has properties like:

- We know or can control the inputs to the program.
- We can specify how the program will achieve its goal
- We can map out what decisions the program will make and under what conditions it makes them.
- Since we know the inputs as well as the expected outputs, we can be con dent that the program will achieve its goal

3.1.1 SUPERVISED AND UNSUPERVISED LEARNING

Machine learning techniques can be broadly categorized into the following types:

Supervised learning takes a set of feature/label pairs, called the training set. From this training set the system creates a generalised model of the relationship between the set of descriptive features and the target features in the form of a program that contains a set of rules. The objective is to use the output program produced to predict the label for a previously unseen, unlabelled input set of features, i.e. to predict the outcome for some new data. Data with known labels, which have not been included in the training set, are classified by the generated model and the results are compared to the known labels. This dataset is called the test set. The accuracy of the predictive model can then be calculated as the proportion of the correct predictions the model labeled out of the total number of instances in the test set.

Unsupervised learning takes a dataset of descriptive features without labels as a train- ing set. In unsupervised learning, the algorithms are left to themselves to discover interesting structures in the data. The goal now is to create a model that finds some hidden structure in the dataset, such as natural clusters or associations. Unsupervised learning studies how systems can infer a function to describe a hidden structure from unlabeled data. The system does not figure out the right output, but it explores the data and can draw inferences from datasets to describe hidden structures from unla- beled data. Unsupervised learning can be used for clustering, which is used to discover any inherent grouping that are already present in the data. It can also be used for association problems, by creating rules based on the data and finding relationships or associations between them.

Semi-supervised machine learning falls somewhere in between supervised and unsupervised learning, since they use both labeled and unlabeled data for training typically a small amount of labeled data and a large amount of unlabeled data. The systems that use this method are able to considerably improve learning accuracy. Usually, semisupervised learning is chosen when the acquired labeled data requires skilled and relevant resources in order to train it / learn from it. Otherwise, acquiring labeled data generally does not require additional resources.

Reinforcement machine learning algorithms is a learning method that interacts with its environment by producing actions and discovers errors or rewards.Machine learning algorithms are tools to automatically make decisions from data in order to achieve some over-arching goal or requirement. The promise of machine learning is that it can solve complex problems automatically, faster and more accurately than a manually specified solution, and at a larger scale. Over the past few decades, many machine learning algorithms have been developed by researchers, and new ones continue to emerge and old ones modified.

3.2 SCIKIT LEARN

SciKit learn is an open source machine learning library built for python. Since its release in 2007, Scikit-learn has become one of the most popular open source machine learning libraries. Scikit-learn (also called sklearn) provides algorithms for many machine learning tasks including classification, regression, dimensionality reduction and clustering.

The documentation for scikit-learn is comprehensive, popular and well maintained. Sklearn is built on mature Python Libraries such as NumPy, SciPy, and matplotlib. While languages such as R and MATLAB are extremely popular and useful for machine learning, we decided to choose Python along with its SciKit-learn libraries as our programming language of choice.

3.3 ADVANTAGES

Crop recommendation systems offer several advantages that contribute to improved agricultural practices and enhanced productivity. Here are some key benefits:

1. Increased Yield: By analyzing various agronomic factors such as soil type, climate conditions, and historical data, crop recommendation systems can suggest the most suitable crops for a specific location. This optimization leads to increased crop yields as farmers are guided towards selecting crops that thrive in their environment.

2. Resource Efficiency: Crop recommendation systems help optimize resource utilization by suggesting crops that require minimal inputs such as water, fertilizers, and pesticides. By matching crops to local soil and climate conditions, farmers can reduce waste and environmental impact while maximizing resource efficiency.

3. Risk Mitigation: By diversifying crop selection and recommending resilient crop varieties, recommendation systems help farmers mitigate risks associated with factors like climate variability, pests, and diseases. This reduces the likelihood of crop failure and minimizes financial losses.

4. Improved Profitability: By maximizing yields and minimizing input costs, crop recommendation systems contribute to improved profitability for farmers. By selecting

crops with higher market demand and better price prospects, farmers can further enhance their economic returns.

5. Sustainable Agriculture: Crop recommendation systems promote sustainable agricultural practices by encouraging the use of environmentally friendly farming methods. By selecting crops that are well-suited to local conditions and require fewer chemical inputs, these systems contribute to soil health, biodiversity conservation, and ecosystem resilience.

6. Data-Driven Decision Making: Crop recommendation systems empower farmers to make informed decisions based on data-driven insights rather than relying solely on traditional knowledge or intuition. By leveraging advanced analytics and machine learning algorithms, these systems provide actionable recommendations that optimize agricultural outcomes.

Overall, crop recommendation systems play a crucial role in modernizing agriculture, improving farm profitability, and promoting sustainable farming practices. By harnessing the power of technology and data-driven insights, these systems help farmers navigate complex agricultural challenges and achieve better outcomes.

3.4 DISADVANTAGES

Certainly, here are four major disadvantages of crop recommendation systems:

1. Data Dependence and Quality: Crop recommendation systems heavily rely on the availability and quality of data. Inaccurate or incomplete data can lead to unreliable recommendations, impacting farmers' decisions and potentially reducing crop yields.

2. Lack of Local Context: Some crop recommendation systems may overlook localized farming practices, preferences, and socio-economic factors. Recommendations based solely on general environmental data may not consider specific regional nuances, leading to inappropriate crop choices for certain areas.

3. Overreliance on Technology: Farmers overly dependent on crop recommendation systems may diminish their own agricultural knowledge and decision-making skills. This

overreliance can limit adaptability in responding to unexpected challenges or changing conditions.

4. Implementation Challenges and Accessibility: Implementing and maintaining sophisticated crop recommendation systems can be complex and costly, especially for small-scale farmers with limited resources and technical expertise. Accessibility to these systems may also be limited in rural areas with poor internet connectivity or limited access to technology, exacerbating disparities in agricultural practices and outcomes.

CHAPTER 4

SNAPSHOTS

4.1 CODE

```
import pandas as pd
import numpy as np
from sklearn.ensemble import RandomForestClassifier
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score, precision_score,
recall_score, f1_score
from sklearn.preprocessing import LabelEncoder
import pandas as pd

data = pd.read_excel(r"C:\Users\HP\Downloads\crop recommendation data
set.xlsx")
pd.set_option('display.max_rows', None)
print(data)
data('Crop Recommendation').value_counts()
encoder=LabelEncoder()
for column in data.columns:
  if data[column].dtypes=='object':
    data[column]=encoder.fit_transform(data[column])
print(data)
data['Crop Recommendation'].value_counts()
data.head()
X=data.drop('Crop Recommendation', axis=1)
y=data['Crop Recommendation']
x.head()
y.head()
X_train, x_test, y_train, y_test = train_test_split(x,y,test_size=0.2)
model= RandomForestClassifier(n_estimators=100)
model.fit(X_train, y_train)
predicted_targets = model.predict(X_test)
accuracy-accuracy_score(y_test, predicted_targets)
precision = precision_score(y_test, predicted_targets, average='micro')


recall = recall_score(y_test, predicted_targets, average='micro')
f1 = f1_score(y_test, predicted_targets, average='micro')
print('Accuracy:', accuracy)
print('''
o for alluvial
1 for alluvium
2 for black
3 for black soil
4 for clay
5 for clay loam
6 for gravels
7 for Loam
8 for loamy sand
9 for loamy soil
10 for medium black
11 for mixed red and black
12 for red
13 for red loamy
14 for red sandy
15 for sandy clay
16 for sandy loam
17 for sandy-clay loam
18 for shallow black
        ''')
i=(input("Select the soil type and enter the corrosponding number of
the soil by refering the above soil chart\n:"))
j=(input("enter the ph of the soil\n:"))
k=(input("enter the temperature\n:"))
new_data = [[i,j,k]]
predicted_crop = model.predict(new_data)
```

Figure 4.1 Importing model and assigning values

```
if i==0:
  print("Soil Type: alluvial")
elif i==1:
  print(" Soil Type: alluvium")
elif i==2:
  print(" Soil Type: alluvium")
elif i==3:
  print(" Soil Type: alluvium")
elif i==4:
  print(" Soil Type: alluvium")
elif i==5:
  print(" Soil Type: alluvium")
elif i==6:
  print(" Soil Type: alluvium")
elif i==7:
  print(" Soil Type: alluvium")
elif i==8:
  print(" Soil Type: alluvium")
elif i==9:
  print(" Soil Type: alluvium")
elif i==10:
  print(" Soil Type: alluvium")
elif i==11:
  print(" Soil Type: alluvium")
elif i==12:
  print(" Soil Type: alluvium")
elif i==13:
  print(" Soil Type: alluvium")
elif i==14:
  print(" Soil Type: alluvium")
elif i==15:
```

Figure 4.2 Soil Type

```
  print(" Soil Type: alluvium")
elif i==16:
  print(" Soil Type: alluvium")
elif i==17:
  print(" Soil Type: alluvium")
elif i==18:
  print(" Soil Type:alluvium")
print("Soil Ph",j)
print("Temperature", k)
# print('Precision:', precision)
# print('Recall:', recall)
# print('F1 score:', f1)
# print(predicted_crop)
if predicted_crop==13:
  print(" the recommended crop is rice")
elif predicted_crop==1:
  print("the recommended crop is Apple")
elif predicted_crop==2:
  print("the recommended crop is Banana")
elif predicted_crop==3:
  print("the recommended crop is Beans")
elif predicted_crop==4:
  print("the recommended crop is coffee")
elif predicted_crop==5:
  print("the recommended crop is Cotton")
elif predicted_crop==6:
  print("the recommended crop is Cowpeas")
elif predicted_crop==7:
  print("the recommended crop is grapes")
elif predicted_crop==8:
  print("the recommended crop is groundnut")
```

Figure 4.3 Recommending crop

4.2 OUTPUT

```
0 for alluvial
1 for alluvium
2 for black
3 for black soil
4 for clay
5 for clay loam
6 for gravels
7 for  loam
8 for loamy sand
9 for loamy soil
10 for medium black
11 for mixed red and black
12 for red
13 for red loamy
14 for red sandy
15 for sandy clay
16 for sandy loam
17 for sandy-clay loam
18 for shallow black
```

Figure 4.4 Display of Soil Types

```
Select the soil type and enter the corrosponding number of the soil by refering the above soil chart
:8
enter the ph of the soil
:8
enter the temperature
:36
==================================================================================================
Selected Soil No: 8
Soil Ph: 8
Temperature: 36
predicted_crop: [11]
Recommended Crop:Orange
```

Figure 3.5 Crop Recommendation

CONCLUSION

In conclusion, the development and utilization of crop recommendation systems mark a significant advancement in agricultural practices. By harnessing the power of data analytics, machine learning, and agronomic expertise, these systems offer tailored guidance to farmers, enabling them to make informed decisions about crop selection, cultivation techniques, and resource management. The integration of such systems not only enhances agricultural productivity and profitability but also promotes sustainability by optimizing resource use and minimizing environmental impact. Moreover, crop recommendation systems contribute to food security by helping farmers adapt to changing climatic conditions and market demands. As technology continues to evolve, these systems hold promise for further revolutionizing agriculture, ensuring its resilience and viability in the face of global challenges.

BIBILIOGRAHY

[1] AgroConsultant: Intelligent Crop Recommendation System Using Machine Learning Algorithms.

[2] https://engrxiv.org/preprint/view/2959/

[3] https://jespublication.com/uploads/2023-V14I9095.pdf

[4] Tom M. Mitchell, Machine Learning, India Edition 2013, McGraw Hill Educa- tion.