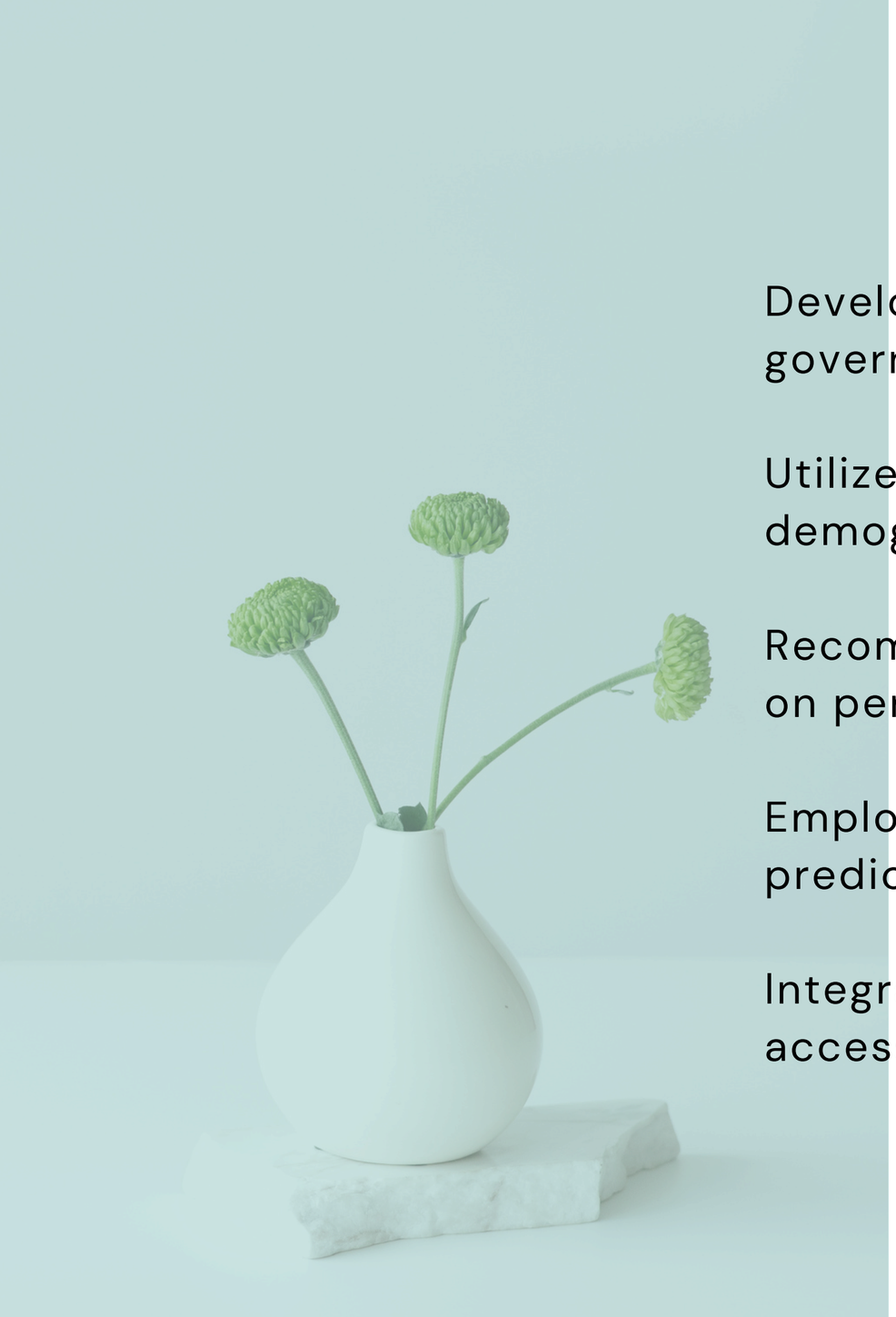
CROP
FARMER SCHEME ELIGIBILITY PREDICTION

SANJAY KS
PES1PG24CA307



Abstract

Develops a predictive system to help Indian farmers access eligible government schemes.

Utilizes machine learning to analyze farmer-specific data such as demographics, crop types, income, and landholding.

Recommends schemes like PMFBY, PM-Kisan, and FPO support based on personalized eligibility.

Employs Decision Forest-based algorithms for robust and accurate predictions.

Integrated into a Flask web application for easy and widespread accessibility.

Problem Statement

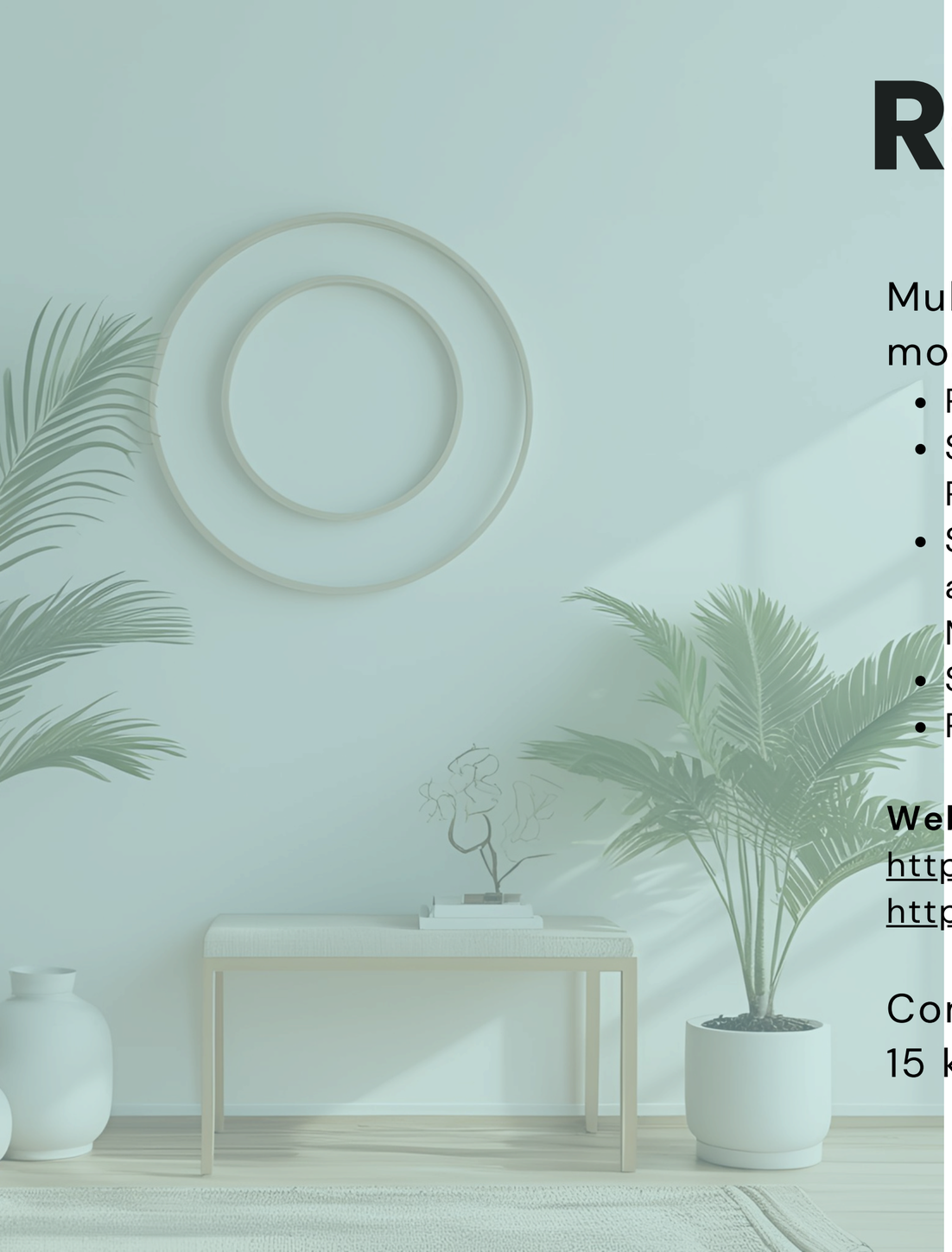
Despite numerous government schemes for the welfare of farmers, many remain unaware of their eligibility.
Manual identification is inefficient and lacks personalization.

This project addresses >>>

Automating the eligibility prediction process

Providing personalized scheme recommendations

Enhancing accessibility through a user-friendly web interface



Record of Dataset

Multiple datasets were used to simulate and train the model:

- PMFBY Enrollment and Claim Data (2018–2022)
- State UT wise Number of Farmer Producer Organisations (FPOs) Registered under 10,000 FPOs Scheme as on 22-07-2024
- StateUTs-wise Details of the Total Farmer Applications Enrolled and Farmer Applications Paid Claims under PMFBY (Pradhan Mantri Fasal Bima Yojana) from 2018-19 to 2021-22
- Scheme_Data
- PMFBY coverage

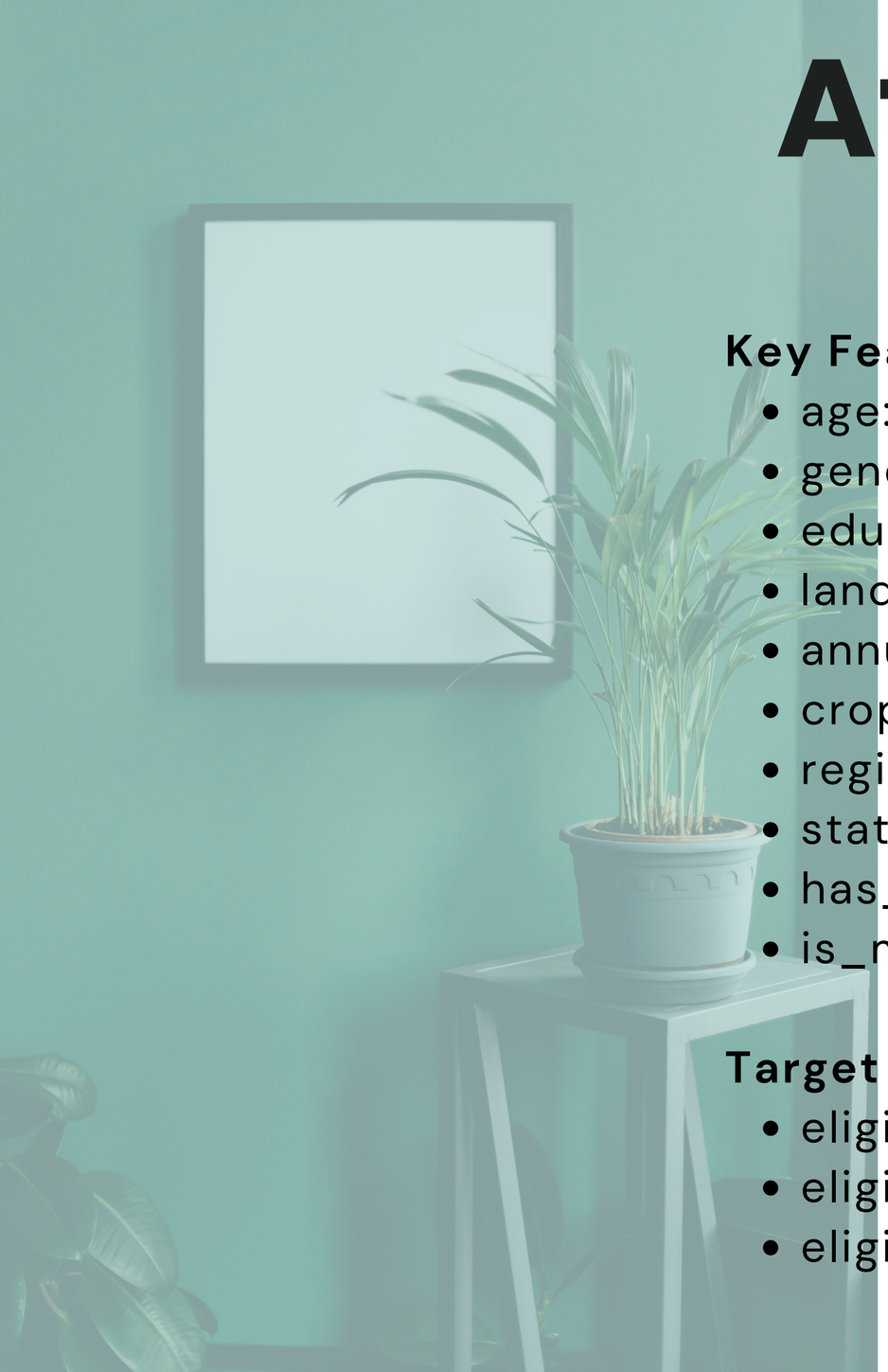
Website Links :
https://www.data.gov.in/
https://www.kaggle.com/

Combined to form a unified dataset with 10,000 rows and 15 key attributes.

Sample Dataset

1	farmer_id	age	gender	education	land_size_acres	annual_income	crop_type	state	district	region
2	1	56	Male	None	9.94	82896.6	Maize	Karnataka	Bengaluru	North
3	2	69	Female	Higher	5.6	428720.86	Maize	Bihar	Muzaffarpu	North
4	3	46	Female	Primary	7.25	313103.18	Wheat	Gujarat	Surat	East
5	4	32	Female	Secondary	1.03	308630.38	Wheat	Punjab	Jalandhar	North

is_member_fpo	has_taken_crop_insurance	eligible_pmfby	eligible_pmkisan	eligible_fpo_support
0	1	1	0	0
0	0	0	0	0
0	0	0	0	0



Attributes of the Dataset

Key Features Used in Prediction:

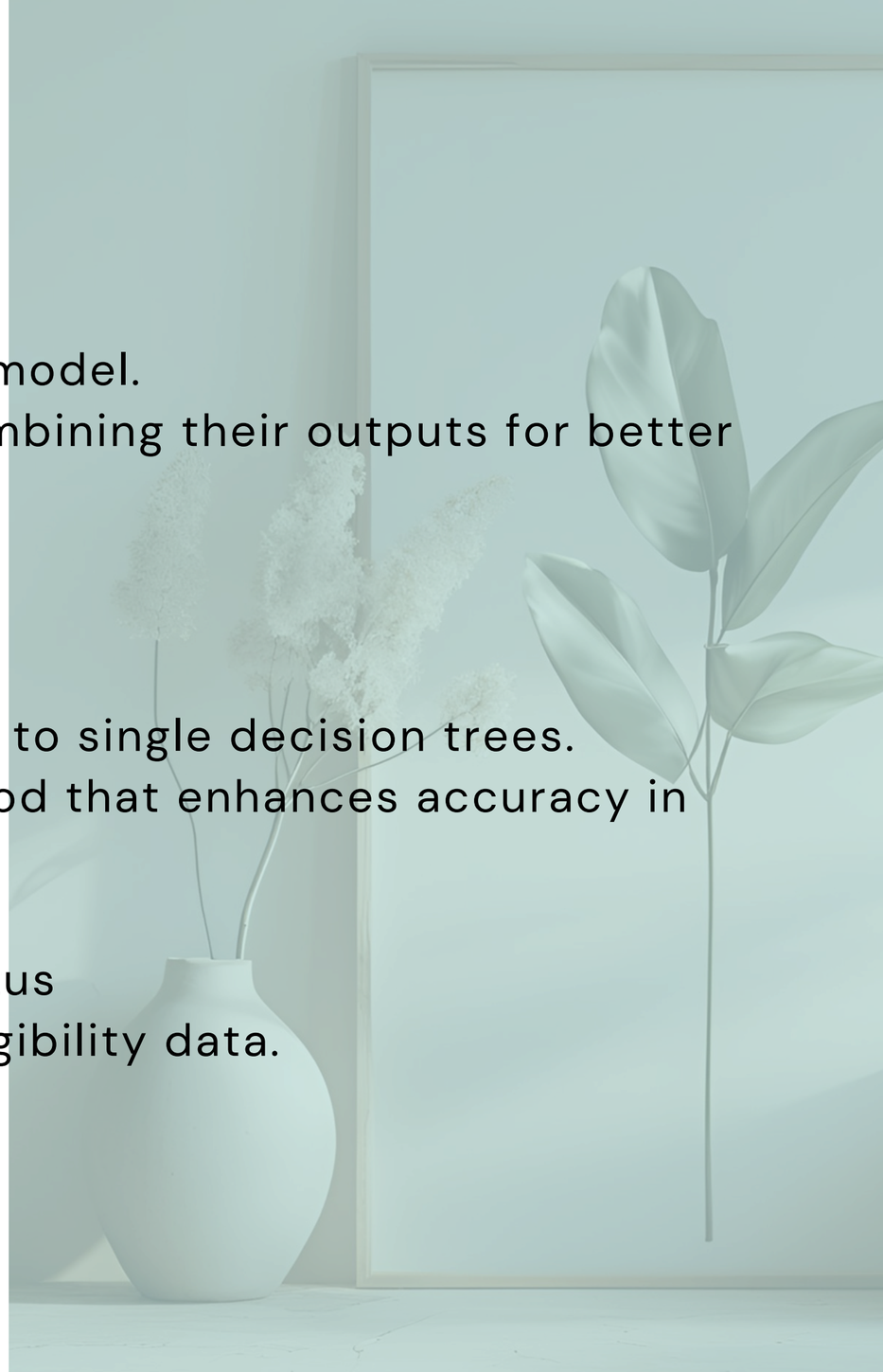
- age: Farmer's age
- gender: Male/Female/Other
- education: Literacy level
- land_size_acres: Size of landholding
- annual_income: Yearly income from agriculture
- crop_type: Primary crop grown
- region: Geographical region (North, South, etc.)
- state, district: Administrative region
- has_taken_crop_insurance: Boolean flag
- is_member_fpo: Farmer Producer Org membership

Target Variables (Labels):

- eligible_pmfby
- eligible_pmkisan
- eligible_fpo_support

Algorithm Used

- **Random Forest Classifier** used as the primary predictive model.
- Operates by constructing **multiple decision trees** and combining their outputs for better accuracy.
- Chosen for its ability to handle:
  - Both numerical and categorical data
  - Missing values and noisy data
- Provides **high performance and low overfitting** compared to single decision trees.
- Additionally evaluated **XGBoost**, a gradient boosting method that enhances accuracy in complex patterns.
- Models trained on features like:
  - Age, land size, crop type, income, region, insurance status
- Achieved accuracy over 99% on processed and refined eligibility data.



Thank you!