Plant Leaves Disease Detection and Pesticide Recommendation

Sakshi Anil Meher

Date: 24th October, 2023

1. **Problem statement**:

The agricultural land mass is more than just a source of food in today's world. The Indian economy is highly dependent on agricultural productivity. Therefore, in the field of agriculture, the detection of disease in plants plays an important role. To detect a plant disease in its early stages, the use of automatic disease detection techniques is beneficial. The existing method for plant disease detection is simply naked-eye observation by experts, through which identification and detection of plant diseases are done. For doing so, a large team of experts as well as continuous monitoring of plants are required, which costs very much when we work on large farms.

2. Market need assessment:

Farmers do not have proper facilities or even the idea that they can contact experts. As a result, consulting experts is both expensive and time consuming. In such conditions, the suggested technique proves to be beneficial in monitoring large fields of crops. Automatic detection of diseases by just seeing the symptoms on the plant leaves makes it easier as well as cheaper. Plant disease identification by visual means is a more laborious task and, at the same time, less accurate and can be done only in limited areas. Whereas if the automatic detection technique is used, it will take less effort, less time, and become more accurate. In plants, some general diseases seen are brown and yellow spots, early and late scorch, and others are fungal, viral, and bacterial diseases.

3. Target Specifications and Characterization:

A: To assist the small-scale enterprises in product recommendation
B: To eliminate the customers (Farmers) plant disease related issues
C: To yield the better results in terms of productivity from crops and increase the overall customer satisfaction

4. External Searches:

Machine Learning (ML) techniques can be effectively utilized in the field of agriculture and specifically in the context of pesticide management and optimization. Here are some ways in which ML can be applied to the domain of pesticides:

1. Pest Identification and Monitoring: ML can help identify and monitor pests in crops through image recognition techniques. By training models with large datasets of images of various pests, it's possible to develop systems that can accurately identify and classify pests, thereby enabling early detection and timely intervention.

2. Optimized Pesticide Application: ML algorithms can be used to optimize the application of pesticides. By analysing data such as weather conditions, crop health, and pest population, models can recommend the precise amount and timing of pesticide application, reducing unnecessary usage and minimizing the environmental impact.

3. Prediction of Pest Outbreaks: ML can be employed to predict the likelihood of pest outbreaks based on historical data, environmental factors, and other relevant parameters. These predictions can help farmers take proactive measures to prevent or minimize the impact of pest infestations, leading to better crop management and reduced dependence on pesticides.

4. Recommendation Systems for Pesticides: ML-based recommendation systems can suggest suitable pesticides based on the specific crop, region, pest type, and other relevant factors. Such systems can take into account various parameters to recommend the most effective and environmentally friendly pesticide options for specific agricultural contexts.

5. Environmental Impact Assessment: ML can aid in assessing the potential environmental impact of pesticide use. By analysing data on soil quality, water systems, and biodiversity, ML models can help identify potential risks associated with pesticide use and suggest ways to minimize negative effects on the environment.

Implementing ML in the domain of pesticides can lead to more sustainable and efficient agricultural practices, reducing the overall environmental footprint while ensuring crop health and yield. However, it is crucial to ensure that the data used for training these models is representative and diverse, and that the models are regularly updated to incorporate new information and advancements in the field.

5. Benchmarking alternate products:

While there might not be specific off-the-shelf products tailored exclusively for machine learning in the field of pesticides, there are several software tools and platforms that can be used for implementing machine learning techniques in agriculture and pesticide management. These tools can aid in data analysis, modelling, and decision-making processes. Some widely used products and platforms that can be applied in the context of machine learning for pesticides include:

TensorFlow and Keras: These are popular open-source machine learning libraries that can be used for building and training custom deep learning models for tasks such as image recognition of pests, crop disease detection, and more.

Scikit-learn: A powerful machine learning library in Python that provides simple and efficient tools for data mining and data analysis. It offers various algorithms for classification, regression, clustering, and more, which can be applied to pesticide-related datasets.

Microsoft Azure Machine Learning: This cloud-based service provides a range of tools and services for building, training, and deploying machine learning models. It can be utilized for various agricultural applications, including the optimization of pesticide usage and crop management.

Google Cloud AI Platform: Google's AI Platform offers a suite of tools and services for machine learning tasks. It provides capabilities for data preprocessing, model training, and deployment, which can be used for developing customized solutions for pesticide-related challenges.

Amazon SageMaker: Amazon SageMaker is a fully managed service that enables developers and data scientists to build, train, and deploy machine learning models. It can be utilized for creating and deploying custom models for pest identification, prediction, and other pesticide-related tasks.

IBM Watson Studio: IBM Watson Studio is an integrated environment designed to help teams build, train, and deploy machine learning and deep learning models. It provides various tools and services for data scientists, which can be leveraged for developing ML solutions in the agricultural domain.

When implementing machine learning solutions in the context of pesticides, it is crucial to ensure that the data used for training the models is accurate, diverse, and representative of real-world scenarios. Additionally, considering the specific requirements of the agricultural context and the need for interpretability and transparency in decision-making, it's essential to design models that are not only accurate but also easily understandable by domain experts and stakeholders.

6. Applicable Regulations:

There were no specific regulations in India explicitly pertaining to the use of machine learning in the context of pesticides. However, it is essential to consult the latest legal and regulatory documents from relevant Indian government agencies such as the Ministry of Agriculture and Farmers' Welfare, the Central Insecticides Board and Registration Committee (CIBRC), and other related bodies for the most recent information.

That said, in the broader context of data usage and technology implementation in agriculture, India has been actively working on policies and initiatives to promote the adoption of digital technologies in farming practices. The Indian government has been emphasizing the integration of modern technologies, including machine learning and artificial intelligence, in the agricultural sector to improve productivity, sustainability, and the overall welfare of farmers.

7. Applicable Constraints:

Expertise:

7.1: Required Expert Advice for the supervised data creation for the target variable.
7.2: Need to validate the results of the prediction by consulting Pesticide Experts.

8. Business Model:

Small Scale Enterprises solely rely on the intuition and experience to predict and serve the right pesticide. From a monetization perspective, we can have a monthly subscription plan for the distributors with the broad classification of Regular (200 queries per month), Premium (500 queries), and VIP (for unlimited queries). Also we can charge omission from various pesticide companies for recommending their products to the farmers via small vendors.

9. Concept Generation:

The fundamental idea of this product is to serve the farmers better and introduce small scale farming distributors. As farming is a very crucial part in the rural area, this concept of assisting the distributors by incorporating Machine Learning techniques for better customer satisfaction can be very significant.

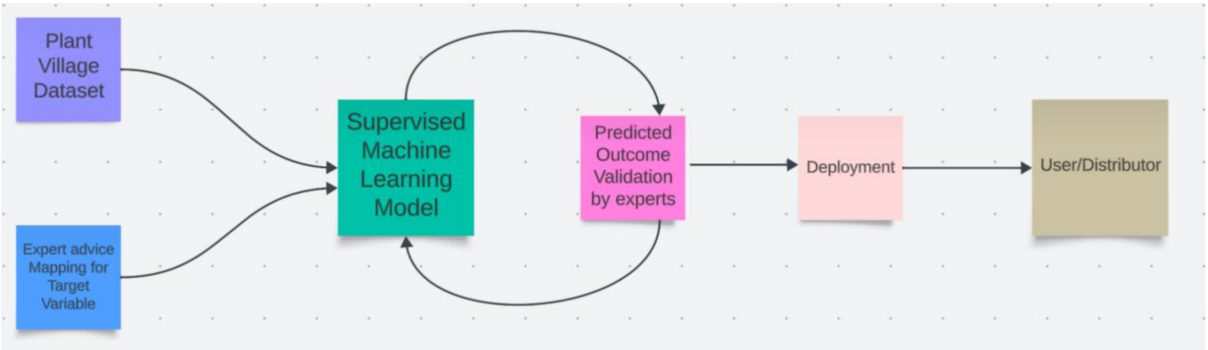
10. Concept Development:

The approach for developing the Pesticide Prediction Machine Learning product involves a systematic process that integrates advanced analytics, agriculture knowledge and experts' guidance.
The first and the major challenge is the collection of data, i.e Data Acquisition step, with the help of Kaggle this problem can be solved. Next thing is about Data Pre-Processing to limit the features as per our requirement and do the mapping for the supervised learning algorithms with the help of Soil and Distributors Experts.

The last challenge would be to pick up the correct ML model for Image Classification and the segmentation, in our case we will use SVMs or Random Forest Classifiers.

A trained model will take input as a plant image and categorise it into one of the diseases, and upon the result, the system will recommend the pesticides and quantity.

11. Final Product Prototype (abstract) with Schematic Diagram:



12. Product details

12.1 Data Sources:

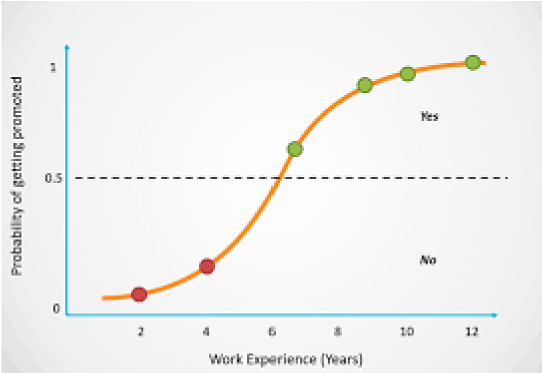
Plant disease classification from Kaggle has the PlantVillage dataset consists of 54303 healthy and unhealthy leaf images divided into 38 categories by species and disease.

A self-made dataset of the pesticides can be prepared with the help of Domain Experts in Pesticides, which will help our supervised ML algorithms for training purposes.

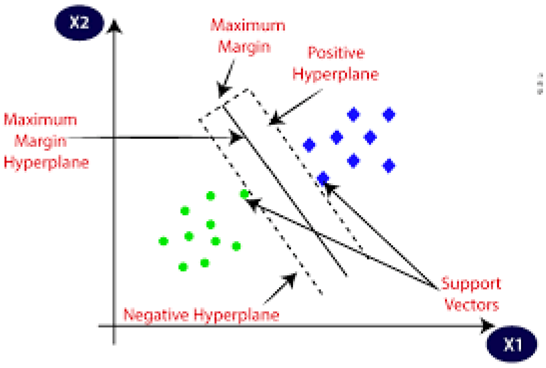
12.2 Algorithms:

Supervised Classification Model in our case predicts which pesticide to use based on the training data set, commonly used Classification models are

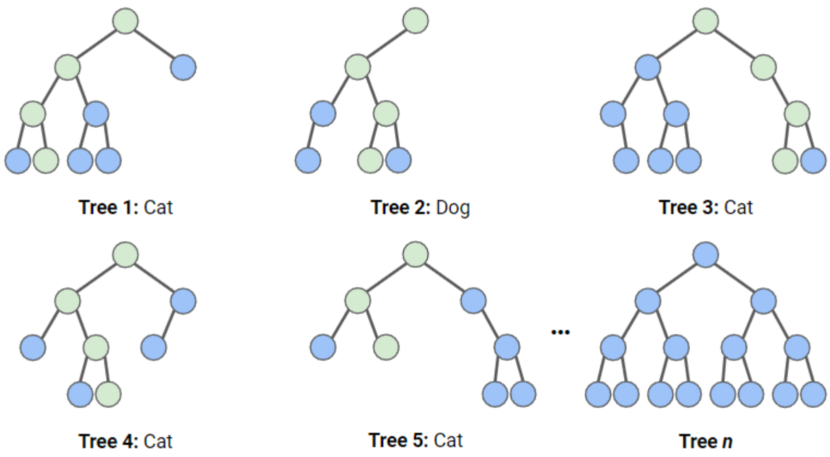
a. **Logistic Regression**: Logistic regression is a statistical method that predicts a binary outcome based on prior observations of a data set. It models the relationship between a dependent binary variable (like yes/no or 1/0) and one or more independent variables. The outcome is a probability, so the dependent variable is bounded between 0 and 1



b. **Support Vector Machine**: Support Vector Machine (SVM) is a powerful machine learning algorithm used for linear or nonlinear classification, regression, and even outlier detection tasks. SVMs can be used for a variety of tasks, such as text classification, image classification, spam detection, handwriting identification, gene expression analysis, face detection, and anomaly detection. SVMs are adaptable and efficient in a variety of applications because they can manage high-dimensional data and nonlinear relationships.



c. **Random Forest Classifier**: In a random forest classification, multiple decision trees are created using different random subsets of the data and features. Each decision tree is like an expert, providing its opinion on how to classify the data. Predictions are made by calculating the prediction for each decision tree, then taking the most popular result. (For regression, predictions use an averaging technique instead.)



12.3 Python Libraries for Pesticide Classification:

Pandas: Pandas is a popular Python library primarily used for data manipulation and analysis. While it is not specifically designed for machine learning, it plays a crucial role in the initial stages of the machine learning workflow, especially in data preprocessing and exploratory data analysis. Here's how Pandas can be used in the context of machine learning:

Syntax for Creating a Dataframe using Image Data Set

img_per_class = pd.DataFrame(nums.values(), index=nums.keys(), columns=["no. of images"])

img_per_class

Scikit Learn: SKLearn is a widely used open-source machine learning library for the Python programming language. Scikit-learn, often abbreviated as Sklearn, provides various tools for data mining and data analysis and is built on top of other scientific Python libraries, such as NumPy, SciPy, and matplotlib. It is designed to work with a range of machine learning tasks, including classification, regression, clustering, dimensionality reduction, model selection, and preprocessing.

Syntax for SVM: From sklearn.svm import SVC Syntax for Random Forest: From sklearn.ensemble import RandomForestClassifier

12.4 Teams required to Develop-

- Business Analyst
- Software Engineer
- Data Scientist expertise in Machine Learning
- Soil Expert/ Distributor
- Machine Learning Engineer

13. Conclusion

The incorporation of machine learning in pesticide management and overall agriculture has immense potential to change the way farming practices are being carried out right now. Despite the progress, there is a lot of room for improvement in terms of research, datasets, and the ethical use of these technologies to ensure the long-term sustainability of the agriculture ecosystem and the protection of human health. As we move forward, the ongoing collaboration between agricultural experts, data scientists, and policymakers will be critical in harnessing the full potential of machine learning for pesticide management while upholding the principles of environmental stewardship and sustainable agriculture.

14. References

[1] Detection of plant leaf diseases using image segmentation and soft computing techniques Vijai Singh , A.K. Misra

[2] Knn based Crop and Fertilizer Prediction Prakash U.M, Sristika Bora, Abhishek Gautam, K.R. Gokul Anand

[3] Multi-criteria Agriculture Recommendation System using Machine Learning for Crop and Fertilizesrs Prediction. KRUPA PATEL* and HIREN B. PATEL

[4] https://paperswithcode.com/dataset/plantvillage

[5] Plant leaf disease detection using computer vision and machine learning algorithmsAuthor links open overlay panelSunil S. Harakannanavar , Jayashri M. Rudagi , Veena I Puranikmath , Ayesha Siddiqua a, R Pramodhini

[6] https://www.mdpi.com/2077-0472/12/9/1350