**Question 1: Elevator Pitch**

This project aims to develop an intelligent system for identifying medicinal plants using image processing and machine learning. By leveraging high-quality plant images and advanced feature extraction techniques like Histogram of Oriented Gradients (HOG) and Scale-Invariant Feature Transform (SIFT), the system enhances accuracy. Convolutional Neural Networks (CNNs), including ResNet and MobileNet, will be trained to classify plant species efficiently. This solution benefits researchers, herbalists, and healthcare professionals by enabling quick and accurate plant identification, supporting biodiversity conservation, traditional medicine, and agriculture.

**Question 2: Dataset Details**

**1) Collector(s):**  
Dataset contributors : **Mendeley** publicly available plant image datasets (e.g., PlantVillage, Medicinal Plant Image Dataset).

**2) Year:**  
2023

**3) Title of Dataset:**  
sources : Indian Medicinal Leaves Image Datasets

**4) Version Number (if any):**  
Version 3

**5) Publisher:**  
Organizations Published: 5 May 2023|Version 3|DOI:10.17632/748f8jkphb.3

Contributors:

Pushpa B R, Shobha Rani

**6) DOI or URL:**  
https://data.mendeley.com/datasets/748f8jkphb/3

**7) Study/Paper/Reason:**  
The dataset is used for training machine learning models to classify medicinal plant species based on their images, supporting applications in botany, medicine, and agriculture.

**Question 3: Language and Libraries**

**Language:**

* Python 3.13.1

**Libraries:**

1. **Image Processing:** OpenCV, Pillow
2. **Feature Extraction:** Scikit-image, SIFT, HOG
3. **Deep Learning:** TensorFlow, Keras, PyTorch
4. **Data Handling:** NumPy, Pandas
5. **Evaluation & Visualization:** Matplotlib, Seaborn

**Question 4: Custom Code**

1. **Dataset Preprocessing:** Image resizing, noise removal, and segmentation.
2. **Feature Extraction:** Implementing HOG, SIFT, or CNN-based feature extraction.
3. **Model Training:** Training CNNs (ResNet, MobileNet) on the dataset.
4. **Hyperparameter Optimization:** Tuning learning rates, batch sizes, and dropout rates.
5. **Model Evaluation:** Writing scripts for accuracy, precision, recall, and F1-score calculations.

**Question 5: Best Choice of Model**

**Model Choice:**

* Convolutional Neural Networks (CNNs) with architectures like ResNet or MobileNet.

**Why:**

* CNNs excel at image classification tasks due to their ability to capture spatial hierarchies in images.
* MobileNet offers efficiency, making it ideal for real-time applications.
* ResNet helps tackle vanishing gradient issues, improving accuracy for deep models.

**Question 6: Hyperparameters and Optimization**

**Key Hyperparameters:**

1. Learning Rate
2. Batch Size
3. Number of CNN Layers
4. Kernel Size
5. Dropout Rate

**Optimization Strategy:**

* Use **Grid Search** or **Random Search** for hyperparameter tuning.
* Implement **learning rate decay** for better convergence.
* Use **data augmentation** to improve model generalization.

**Question 7: Performance Evaluation**

**Metrics:**

1. **Accuracy:** Measures overall classification performance.
2. **Precision:** Ensures correct identification of medicinal plants.
3. **Recall:** Measures the model’s ability to detect all relevant plant images.
4. **F1-Score:** Balances precision and recall for better assessment.

**Techniques:**

* Use **confusion matrices** to analyze misclassifications.
* Apply **Grad-CAM** for model interpretability.
* Compare CNN-based approaches with traditional feature extraction methods (HOG/SIFT).