For queries:

- Dr. P. Shunmuga Perumal (9790914537)
- Mr. S. Maheshkumar (6379288324)

Round 1 Instructions (25.07.2025)

1) Each team is expected to annotate the following dataset using CVAT (offline).
   i) Crop disease dataset (50 images): https://drive.google.com/drive/folders/1-mgBbJnEOSewrKsgmn4aDWbHKaUAeGrD?usp=sharing
   ii) Crop insect dataset (50 images): https://drive.google.com/drive/folders/1szQoqBQdBH4xm8Yo_FxQ34xVGpJOnTgl?usp=sharing

2) Each team is expected to implement a multimodal:
   → ***Refer Table 1:Round-1 Implementation details***

3) Each team is expected to submit the round-1 document to ***shunmugaperumal.p@vit.ac.in*** with the following contents on or before **02-AUG-2025 (Saturday)**:

   i) CVAT (offline) setup procedure.

   ii) Annotation techniques used.

   iii) Annotated dataset (*Attach Google Drive link in the round-1 document*).

   iv) Architecture of the multimodal with
      → short & clear description.
      → Pictorial representation (Clear Hand-drawn pictures also fine)

   v) Implementation codes of the multimodal in an organized manner.

   vi) Procedure to run the entire application (end to end) in the terminal.
      → GUI is not required.

   vii) Output screenshots
      → Project/tabulate all possible highlights (model performance parameters) of your multimodal.

Note:

→ ***Model accuracy is NOT the round-1 selection criteria***.

→ The main objective of this round-1 process is to select suitable teams for the Round-2 of the Hackathon by assessing the following skills:

→ Linking/integrating different models (YOLOv8s-seg, YOLOv8s, TabNET) together to form a multimodal.

→ Documentation skills.

Table 1: Round-1 Implementation details

1	Annotation Tool to be used	CVAT (offline)	• Teams are expected to explore the full potential of the CVAT (offline)
2	Annotation types	• Bounding box (crop Insect images) • Segmentation (crop disease images)	
3	Augmentation (after annotation):	• Normalization • Rotation • Flipping • Color Jittering • Contrast enhancement	• Participants are expected to use other possible augmentation techniques.

4	Multimodal implementation	Input 1: Image disease Input: ● Crop disease Image will be the input. ● Crop disease must be detected (**output-1**).	● YOLOv8s-seg model trained on segmented dataset must be used to detect crop disease.
		Input 2: Image insect Input: ● Crop insect Image will be the input. ● Crop insect must be detected (**output-2)**.	● YOLOv8s model trained on segmented dataset must be used to detect crop disease.
		Input 3: Text input from CSV file: ● A list of questions (symptoms of crop disease) along with the corresponding crop disease images (from the sample collection) will be displayed to the farmers. ● Farmers will look into the crop disease images displayed by the system and answer the questions (Yes/No type inputs). ○ Eg: Question by the system: Do you see a brown color infection on the leaf? ○ Answer by the Farmer: Yes	● TabNet model

		• Farmer's input (text type) will be taken as input by the TabNet model to predict the presence of the crop disease (**output-3**). • Refer the csv file for Farmer's text input (**crop_disease_characteristics.csv**) to TabNET.	

		Input 4: Text input from CSV file:
		• A list of questions (symptoms of crop insect) along with the corresponding crop insect images (from the sample collection) will be displayed to the farmers.
		
		• Farmers will look into the crop insect images displayed by the system and answer the questions (Yes/No type inputs).
		  ○ Eg: Question by the system: Do you see a green color larva on the crop leaf?
		  ○ Answer by the Farmer:	

		Yes • Farmer's input (text type) will be taken as input by the TabNet model to predict the presence of the crop insect (**output-4**). • Refer the csv file for Farmer's text input (**crop_insect_characterists.csv**) to TabNET.	

		• Fusing the outputs of all models together to get the final conclusion:	• The multimodal pipeline must fuse all outputs (***output-1, output-2, output-3, output-4***) to get the final conclusion (final_output).
			
			• **Final_output:**
			  ■ *Crop disease present or not present*
			  ■ *Crop insect present or not present*