Data Challenge ‘22

The submission evaluation process is outlined in the following slides

Scoring Guidelines

Breakdown of the score per task

1 – AOILabel Score

Overview

For the 1st part of the challenge, we use classification *F1-Score of the defect class* to score the capability to predict the correct AOILabel.

- **NOITCE**: We evaluate the label in terms of **presence only**
  - i.e., it is not required to predict the AOILabel reported in the AOI dataset, but just if a component will appear in the AOI dataset.
- We compute the classification accuracy using **the Test dataset only**.
- To compute it, we use the ground truth labels and the labels produced by the **classification_1() only**.
  - The ground truth reports if a component is present in the AOI dataset.

2 – OperatorLabel Score

Overview

For the 2nd part of the challenge, we use the classification *F1-Score of the operator label on finding the Bad class* to score the capability to predict the correct OperatorLabel.

- **NOTICE**: We compute the performance considering only components for which a ground truth label is available
  - i.e., components that have been flagged by the AOI.
  - F1-Score is computed as: $F1 - Score_{Bad}$
- We use **the Test dataset only**.
- To compute it, we use the ground truth labels and the labels produced by the **classification_2() only**.

3 – RepairLabel Score

Overview

For the 3rd part of the challenge, we use the classification *Macro average F1-Score of the Repair label* to score the capability to predict the correct RepairLabel.

- **NOTICE**: We compute the performance considering only components having FalseScrap or NotPossibleToRepair ground truth label
  - i.e., components that have been inspected by the Repair operator.
  - Macro Average F1-Score is computed as: $\frac{F1-Score_{FalseScrap}+F1-Score_{NotPossibleToRepair}}{2}$
- We use **the Test dataset only**.
- To compute it, we use the ground truth labels and the labels produced by the **classification_3() only**.

Final Score

Data challenge Score

For each team, the data challenge score is computed as the average of the scores.

$$\textit{Data challenge} = \frac{AOILabel\ Score\ + OperatorLabel\ Score + RepairLabel\ Score}{3}$$

The evaluation code will be available on GitHub.