

Dataset Magician

Objective : Build a dataset from scratch for a classification task !

Assignment guidelines :

- Take the animal of your first project.
- Using an Image Generation Model, build a dataset from scratch images which will allow you to re-train your model of project 1.
- Use the Diffusers library.
- Compare the results of your project 1 model with the one you train in this project.

Dataset :

- NA

Ressources:

- [Diffusers](#)

Deliverables :

- **A notebook** (html) :
 - Data Generation
 - Classifier training
 - Performance evaluation
- **BONUS: Experiment tracker:**
 - Implement an experiment tracker such as MLFlow.
 - Submit a public link toward the dashboard summarizing your experiment or a screenshot of the dashboard.

Evaluation criterias (110 / 100 pts) :

Skill	Description	Points
Documentation (markdown)	<ul style="list-style-type: none">• Your strategy is explained.• Your code is commented when needed.• The model selection and hyperparameters selection is explained.• The performances are commented on.• Bibliographical references are present.	30
Code (python)	<ul style="list-style-type: none">• All blocks necessary to implement your strategy are present.• Specialized libraries have been used.• All notebook cells have been executed successfully sequentially.	30
Performances	<ul style="list-style-type: none">• A baseline is defined.• More than one model is tested.• All necessary comparisons are done.• Figures are readable and legends are present.• A proper evaluation metric was selected.	40
Bonus	<ul style="list-style-type: none">• The dashboard is submitted.	10