Interactive machine learning

[Insert Project Name Here]

DUBOS Cyril   MORVANY Trévis   POINTAL Brice   Université Paris-Saclay

Introduction

Scenario 1: Help a novice in ML build a robust model with limited data

Help the user explore his previous results in order to provide better data

Type of data: small dataset of images (webcam pictures)

Machine learning task: classification (pre-processing model + classifier)

Related works

Data-iteration :

- Evolving real-life data
- Data unfit for the goal

Model visualization :

- Provides insight
- Helps uncover issues

Fred Hohman, Kanit Wongsuphasawat, Mary Beth Kery, and Kayur Patel. Understanding and visualizing data iteration in machine learning, 2020.

James Wexler, Mahima Pushkarna, Tolga Bolukbasi, Martin Wattenberg, Fernanda Viegas, and Jimbo Wilson. The what-if tool: Interactive probing of machine learning models, 2019.

Proposed design

Giving the user full control over the data

Create and update datasets, flexibility

Comparing model effectiveness

Influence on characteristics

Comparing model prediction

Influence on predictions

Implementation

The tool is based on three main sections

Dataset structure: snapshot information stored within an object when training

Dataset visualization: (re)implemented components for snapshot (using genericChart...)
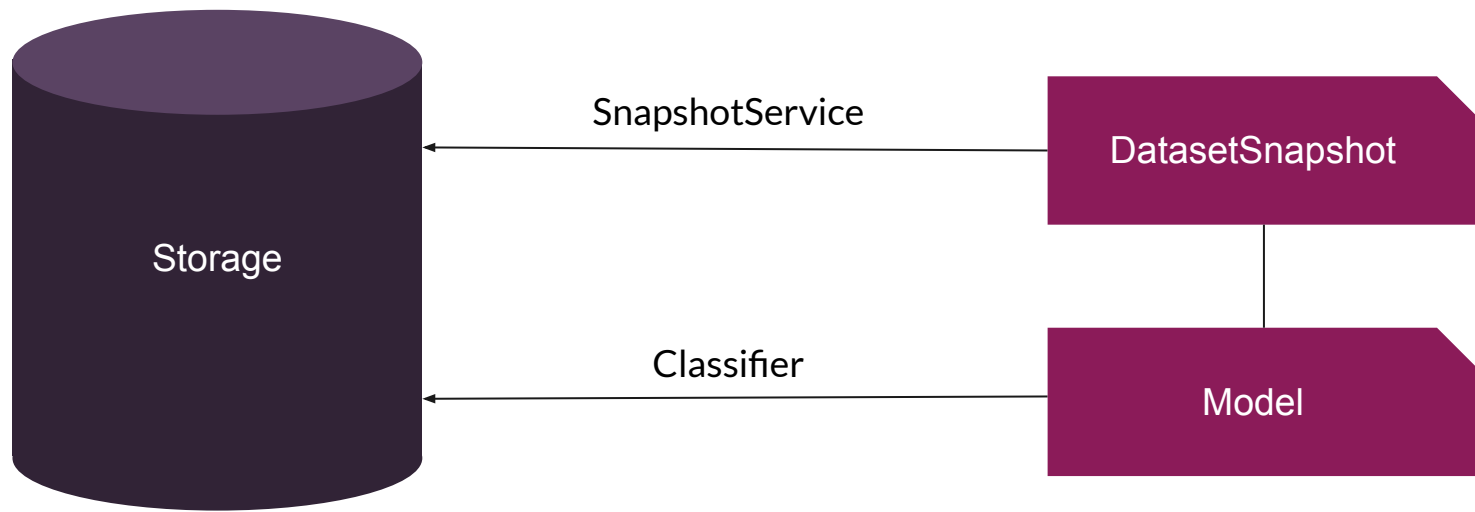
Prediction with Marcelle Multi-layer Perceptron (MPL) classifier

Types

```
type StoredObject = {

    id: string,

    createdAt: string,

    updatedAt: string

}
```

```
type DatasetSnapshot = StoredObject & {

    name: string;

    instances: string;

    model_id: string;

    training_metrics: any;

}
```
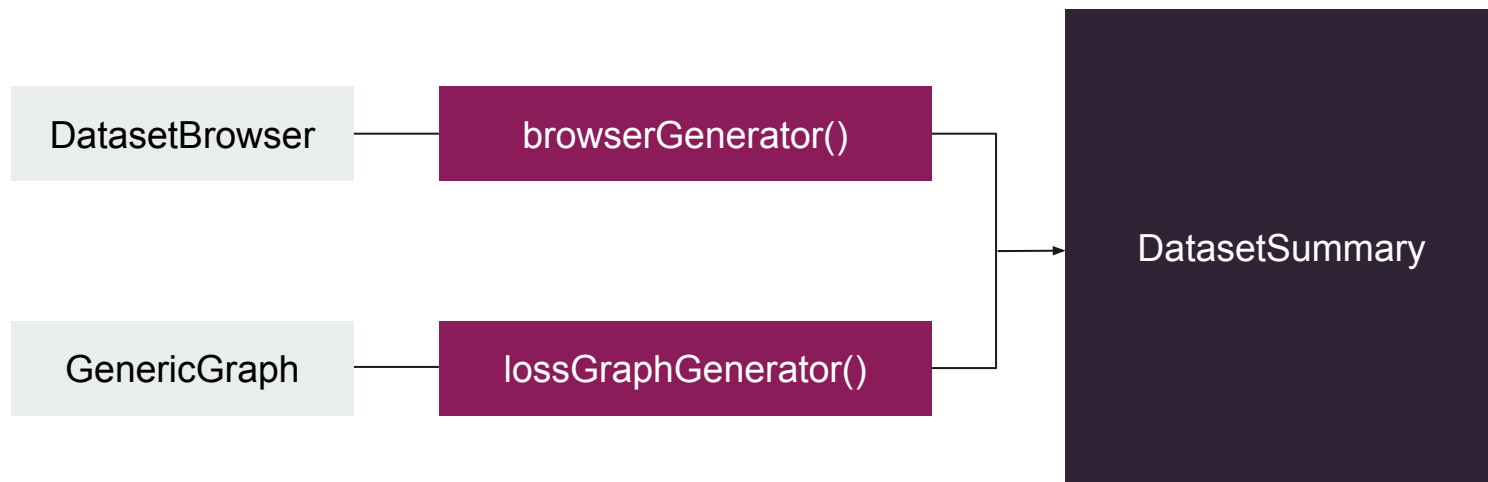
Structure and storage

Visualization generators

```
function browserGenerator(snapshot: DatasetSnapshot): DatasetBrowser {

  let dataset = marcelle.dataset(marcelle.dataStore());

  dataset.upload(new File(snapshot.instances));

  return marcelle.datasetBrowser(dataset);

}
```

Dataset summary

DatasetBrowser

browserGenerator()

GenericGraph

lossGraphGenerator()

DatasetSummary

Demonstration
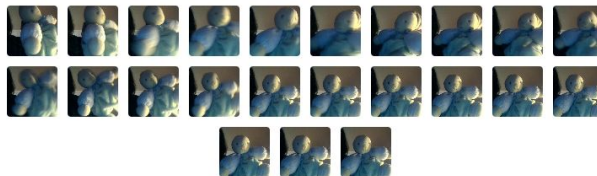
Data entries management
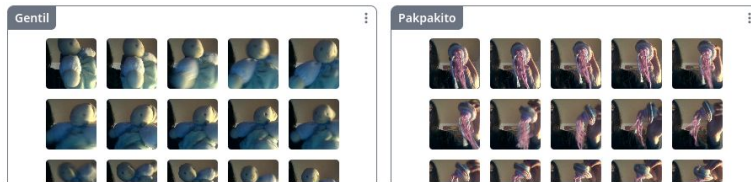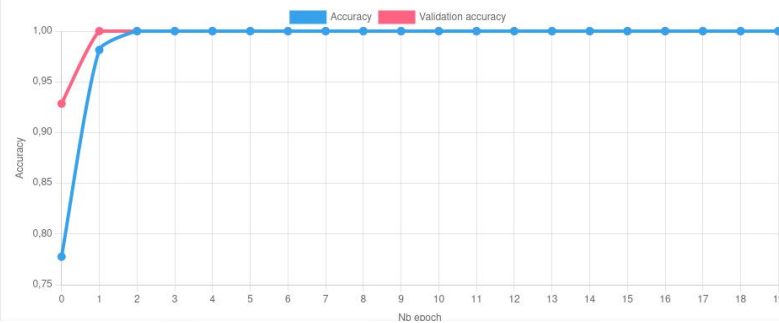
Model characteristics
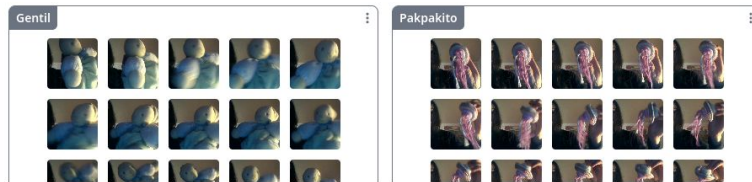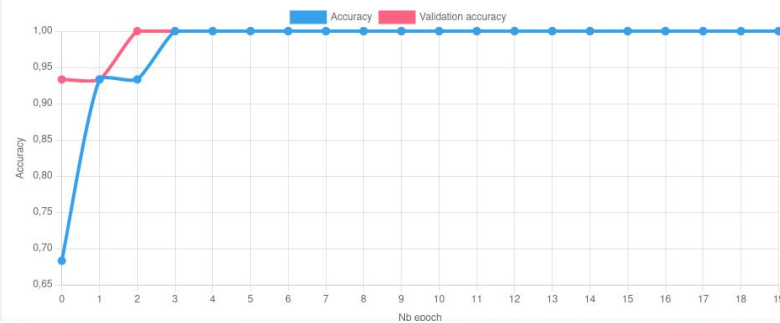
Project demo!   Welcome   Dataset comparison   Dataset predictions   ⚙

webcam

🔵 activate video

Dataset summary

choose a dataset

first ▾

confidence plot

Predicted Label: bourse

■ Confidences

1,0
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0

Confidence

bourse   eau   monster
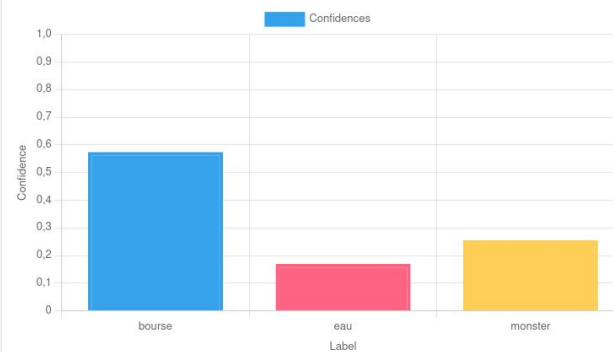
Label

Dataset summary

choose a dataset

second ▾

confidence plot

Predicted Label: bourse

■ Confidences

1,0
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0

Confidence

bourse   eau   monster

Label

Prediction comparison

Welcome   Dataset Comparison   Predictions   **Dataset Management**

Dataset List

| Search... | Sort by name | Sort by date |

☐	**Name**	**Created at**	
☐	a	24/02/2023 13:17:24	Load instances
☐	b	24/02/2023 13:17:42	Load instances
☐	c	24/02/2023 13:17:48	Load instances
☐	aa	24/02/2023 13:22:52	Load instances
☐	ab	24/02/2023 13:28:17	Load instances
☐	a	24/02/2023 13:40:06	Load instances
☐	truc	24/02/2023 13:40:43	Load instances

Delete

Delete selected datasets

> Model a_model was loaded from data store at location localStorage

Datasets management

Conclusion

Helps to understand data - prediction relationship

Facilitate corrections

Intuitive / Flexible

More natural workflow

Thank you.