AutoML Pipeline for Vision Data

Artin Sermaxhaj, Enes Hasani, Erlis Lushtaku EAEE

Modality 2

Introduction

The aim of this project is to develop a robust and efficient automated machine learning system that, given a vision dataset, outputs a trained and tuned model that achieves high results on image classification tasks. We utilized multifidelity HPO techniques to fine-tune a pre-trained model.

Our Approach

External Sources: ChatGPT API/Expert

User Prior

HPO:

PriorBand

Final model training: 100 epochs with best

hyper-parameter

configuration

Final model

evaluation

Week 1

Week 2

Week 3

Week 4

Week 5

Week 6

Week 7

Week 8

Week 9

Week 10

Bonus

Literature

Priors

- Manual expert priors through console input
- Prompting LLM gpt4o-mini Input:
 - Task context
 - Configuration space
 - Dataset metadata

Output:

- Mean and confidence of prior distribution

0.72

0.00 0.02 0.04 0.01 0.00 0.10 0.00 0.01 0.00

Emotions - Confusion Matrix

Resources Used

For development:

- 1 RTX3060 6GB GPU
- AMD Ryzen 5 CPU
- Apple M1 Pro CPU
- Total compute estimate: 250 CPU, GPU-h

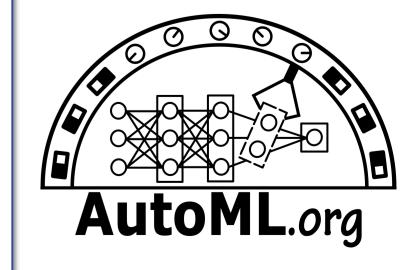
For AutoML:

- 1 RTX3060 6GB GPU
- Kaggle
- 20h

Workforce:

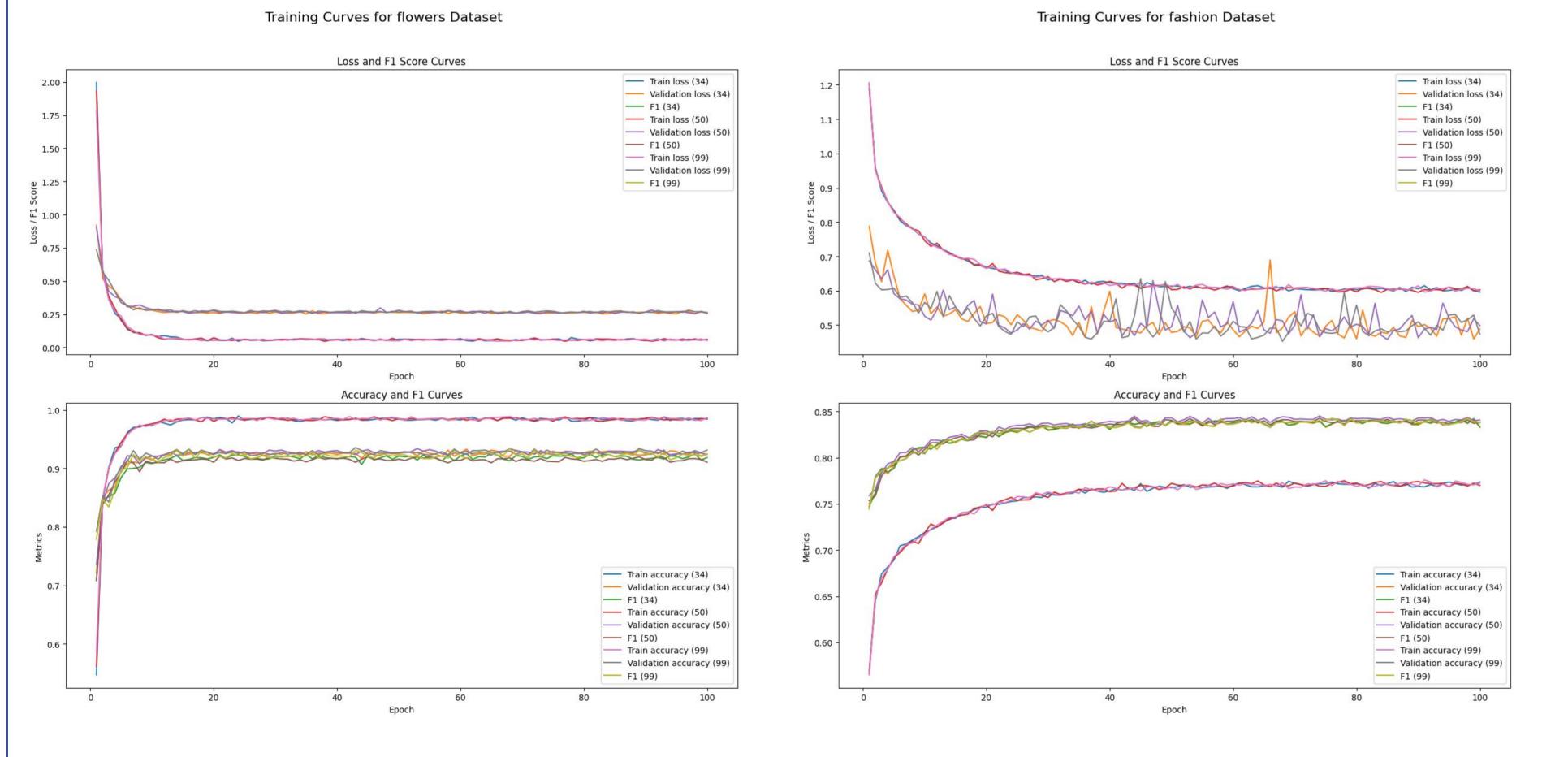
- 1 full week on average

Number of queries for test score generation: 1





Empirical Results



Fidelity 1 to 27

epochs

Pipeline

Data Augmentation

AutoAugment

Architechture Space:

Resnet18 and Resnet50

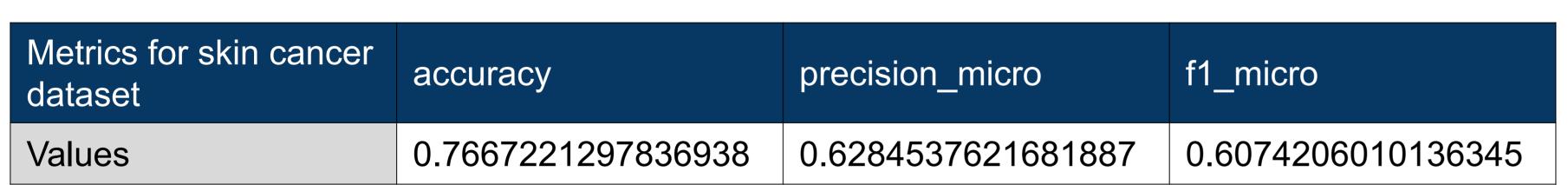
unfreezed layers)

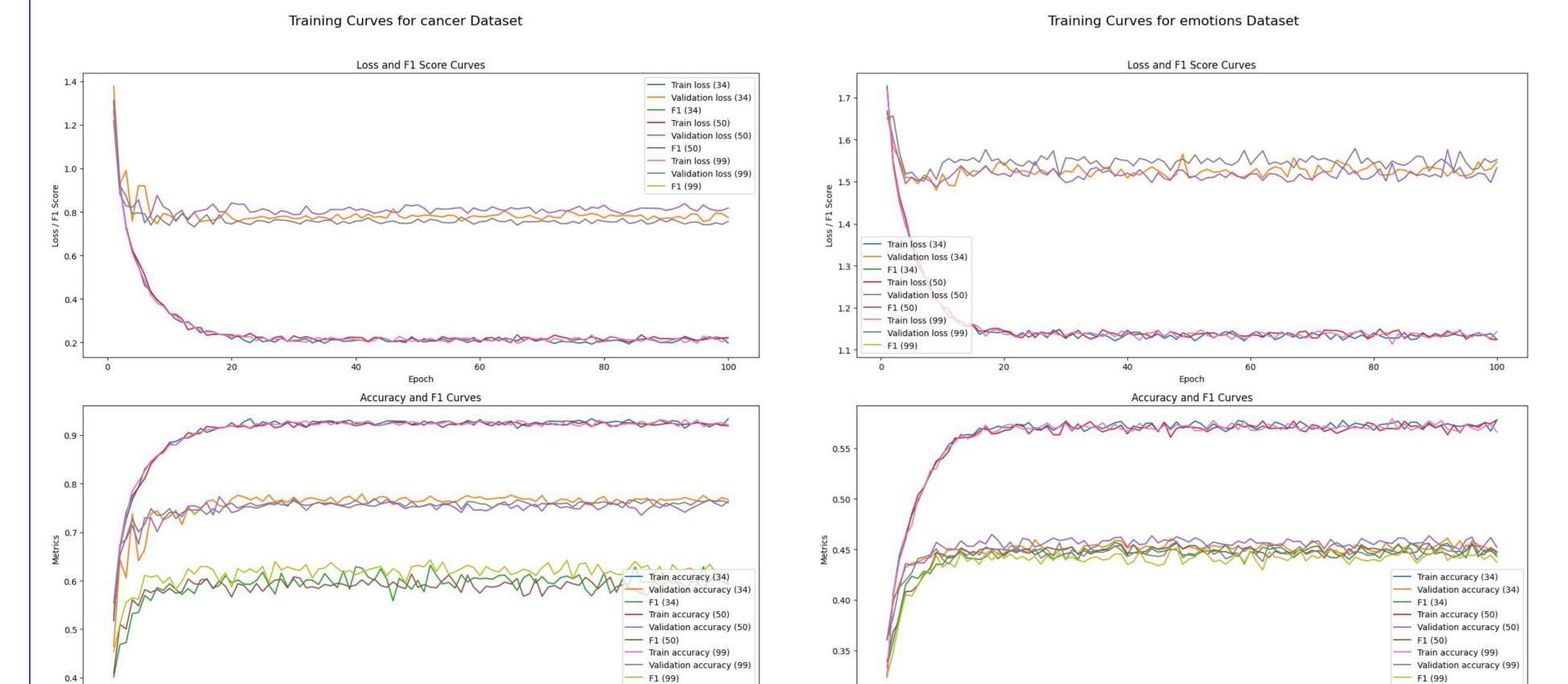
Hyper-parameter Space:

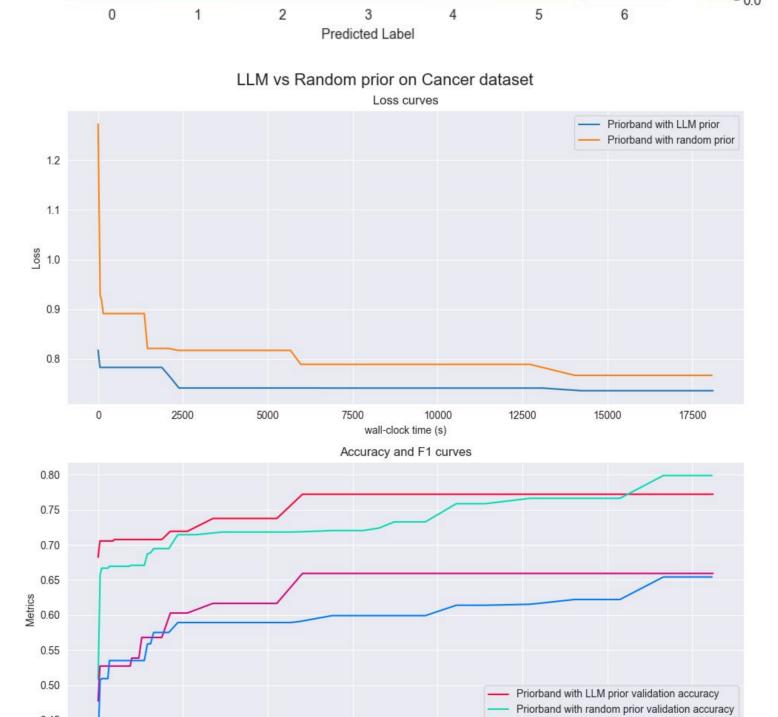
learning rate, weight

decay, scheduler gamma scheduler step size

Data splits







wall-clock time (s)

— Priorband with LLM prior F1 score

Priorband with random prior F1 score

0.04