

CL-as-a-Service: a novel tool to Continual and Real-time machine learning



powered by



Rudy Semola



<https://github.com/rudysemola>

r.semola@phd.unipi.it

rudysemola@gmail.com

About Me

Rudy Semola
PhD Student (ML System, Continual Learning) @ Unipi | Machine Learning Engineer Researcher @ TEACHING_H2020 | member @ ContinualAI community | Innovation enthusiast
Lucca, Tuscany, Italy · [Contact info](#)
500+ connections

TEACHING_H2020
Università di Pisa



Università di Pisa

Master's degree, Computer Science, curriculum Artificial Intelligence
2018 - 2021



Università degli Studi di Firenze

Bachelor's degree, Computer Engineering
2016 - 2018

Machine Learning Engineer Researcher
with startups, companies and R&D labs



Outline

Intro & Motivation

- Companies shortcut (open problems and challenges)
- Research Solution (Focus on Continual Learning, overall picture)

Tools Solution

- CL-as-a-Services system overview
- CL-as-a-Services system (Live) DEMO

Summary

Introduction and motivation

Analysis: scenario of companies in different industries

- Innovative trend is towards more and more **Real-time machine learning inferences** and **Continual Learning** update model

To move in these two direction companies found different challenges depend on their state

- Inferences ML system
 - Batch Prediction
 - Online Inferences with batch features
- Update model
 - Manual, Stateless Retraining
 - Automated Retraining
 - Automated, Stateful Training
 - Continual Learning

Introduction and motivation

Trend Companies direction: GrubHub* use case

- Recommendation System
- Scaling rapidly and they wanted to improve the user experience

In short, they did AB test, the results:

1. Increase 20% more in Purchase Through Rate
2. Model update costs reduced by x45

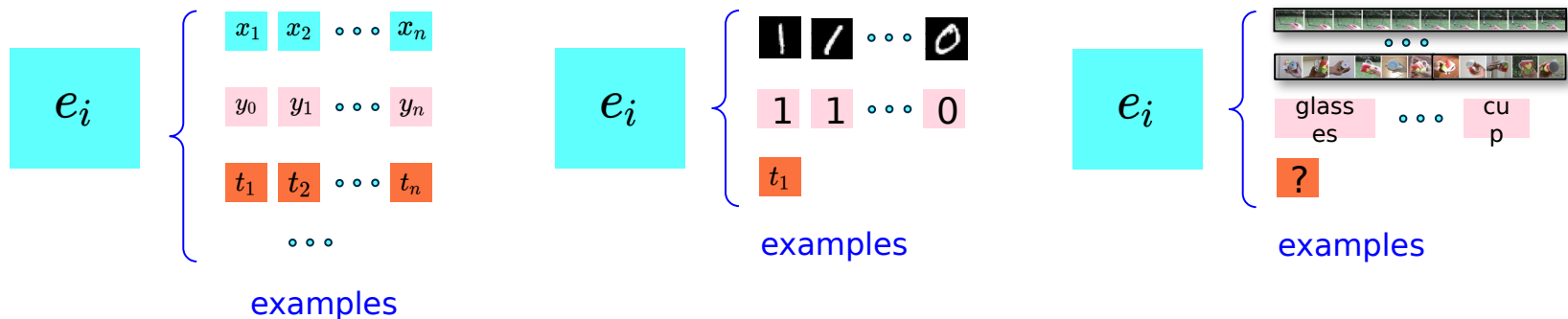
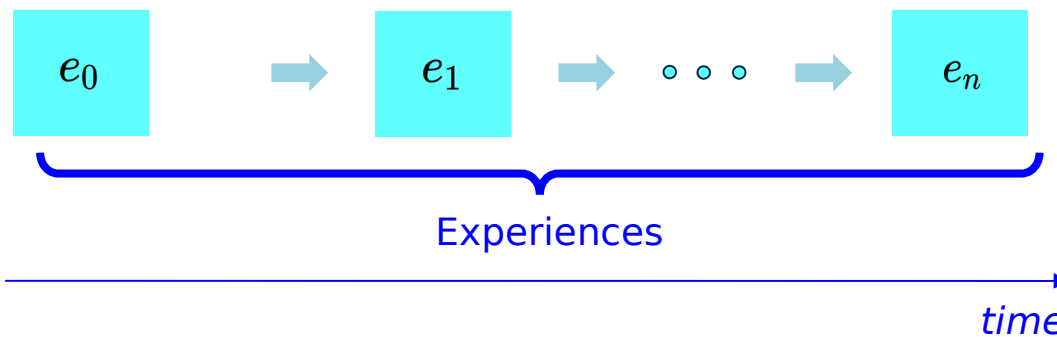
These results are based on an improvement by the company in their level of prediction, from batch to online, and retraining modalities, from stateless to stateful.

* Grubhub is an American online and mobile prepared food ordering and delivery platform owned by Just Eat Takeaway that connects diners with local restaurants



Research Solution: Continual learning

- ❑ **Deep learning** holds state-of-the-art performances in many tasks
- ❑ **Continual Learning** in a nutshell : Learn from a continual stream of experiences
 - ❑ Non stationary assumption (**Concept Drift**)
 - ❑ Learn new skill vs remember past skill (**Catastrophic Forgetting**)
 - ❑ limited computational and memory resources (**Efficiency**)



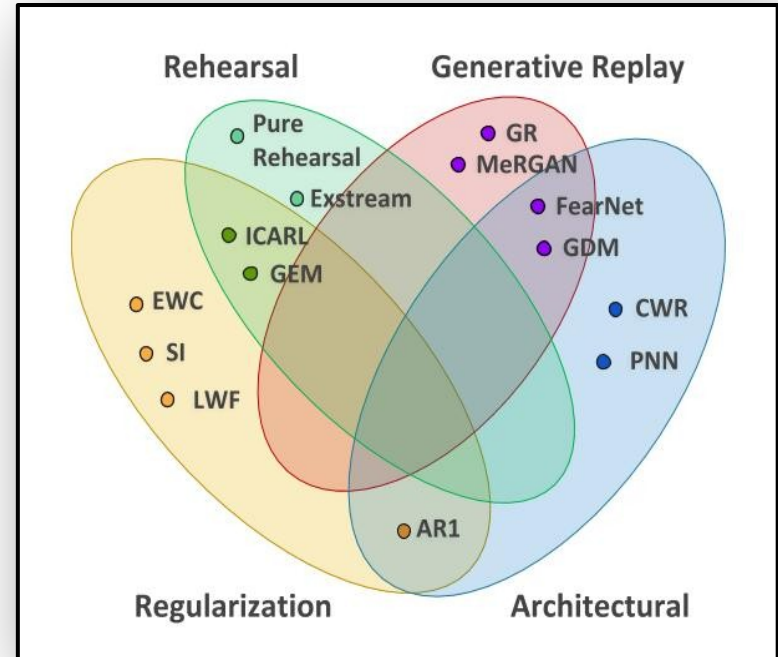
Research Solution: Continual learning

Methodologies

No formal definition

Common Baselines / Control Algorithms

- Naive / Finetuning** (just continuing backprop)
- JointTraining / Offline** (pure Multi-task learning): The best you can do with all the data starting from scratch
- Ensemble**: one model for each experience
- Cumulative**: for every experience, accumulate all data and re-train from scratch.




Research Solution: Continual learning

CL Research Tools

- Avalanche:** an End-to-End Library for Continual Learning
- Continuum:** Simple management of complex continual learning scenarios
- Sequoia:** A Software Framework to Unify Continual Learning Research
- CL-Gym:** Full-Featured PyTorch Library for Continual Learning


Avalanche: an End-to-End Library for Continual Learning

Powered by ContinualAI



Avalanche

powered by



ContinualAI

Avalanche is an *End-to-End Continual Learning Library* based on [PyTorch](#), born within [ContinualAI](#) with the unique goal of providing a **shared** and **collaborative** open-source (MIT licensed) **codebase** for *fast prototyping, training and reproducible evaluation* of continual learning algorithms.

Avalanche can help *Continual Learning* researchers and practitioners in several ways:

Research Solution: Continual learning

From research to production tool: [ModelCI-e](#)

- lightweight MLOps plugin
- Embraces CL and ML deployment techniques
- End-to-end supports for model updating and validation

Main component

- Model factory
- CL Server (drift detector, Replay-based)
- (nice) GUI interface

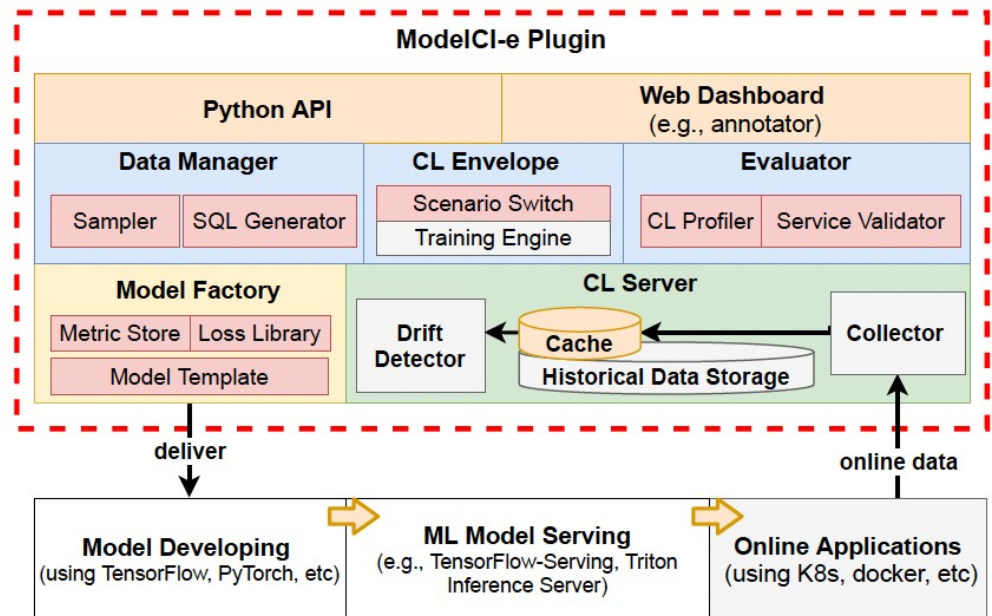


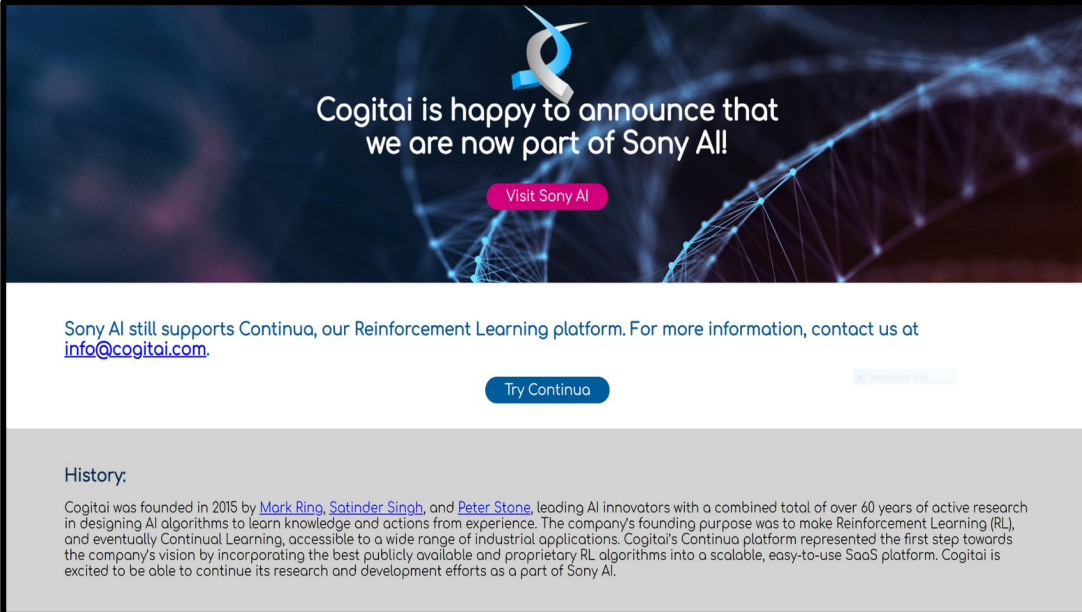
Figure 1: ModelCI-e architecture.

Research Solution: Continual learning

CL startup: Cogitai

By Mark Ring

- Based on Ring Continual Learning specific formulation (different from Lifelong Learning)
- Founded in 2015, sold in 2016 to Sony AI
- CLaaS platform, mostly focused on Reinforcement Learning algorithms




The image shows a screenshot of a webpage announcement. At the top, there is a dark blue header with a glowing network graphic. The text reads: "Cogitai is happy to announce that we are now part of Sony AI!". Below this is a pink button that says "Visit Sony AI". The main body of the page is white and contains the text: "Sony AI still supports Continua, our Reinforcement Learning platform. For more information, contact us at info@cocigitai.com". Below this is a blue button that says "Try Continua". At the bottom, there is a grey section titled "History:" followed by a paragraph of text: "Cogitai was founded in 2015 by [Mark Ring](#), [Satinder Singh](#), and [Peter Stone](#), leading AI innovators with a combined total of over 60 years of active research in designing AI algorithms to learn knowledge and actions from experience. The company's founding purpose was to make Reinforcement Learning (RL), and eventually Continual Learning, accessible to a wide range of industrial applications. Cogitai's Continua platform represented the first step towards the company's vision by incorporating the best publicly available and proprietary RL algorithms into a scalable, easy-to-use SaaS platform. Cogitai is excited to be able to continue its research and development efforts as a part of Sony AI."

Research Solution: Continual learning

CL startup: Gantry

Sort of OpenAI spin-off company

- Focusing on MLOps infrastructure for evolving data
- Still in the early days, but interesting set of investors and key human resources
- It suggests large opening in this fast emerging market



Data evolves. Build ML systems that adapt.

Gantry gives you full visibility into the state of your machine learning system. Decide when to retrain, what data to retrain on, and which models are performing best.

[Learn more about continual learning systems →](#)

[Sign up for our mailing list](#)

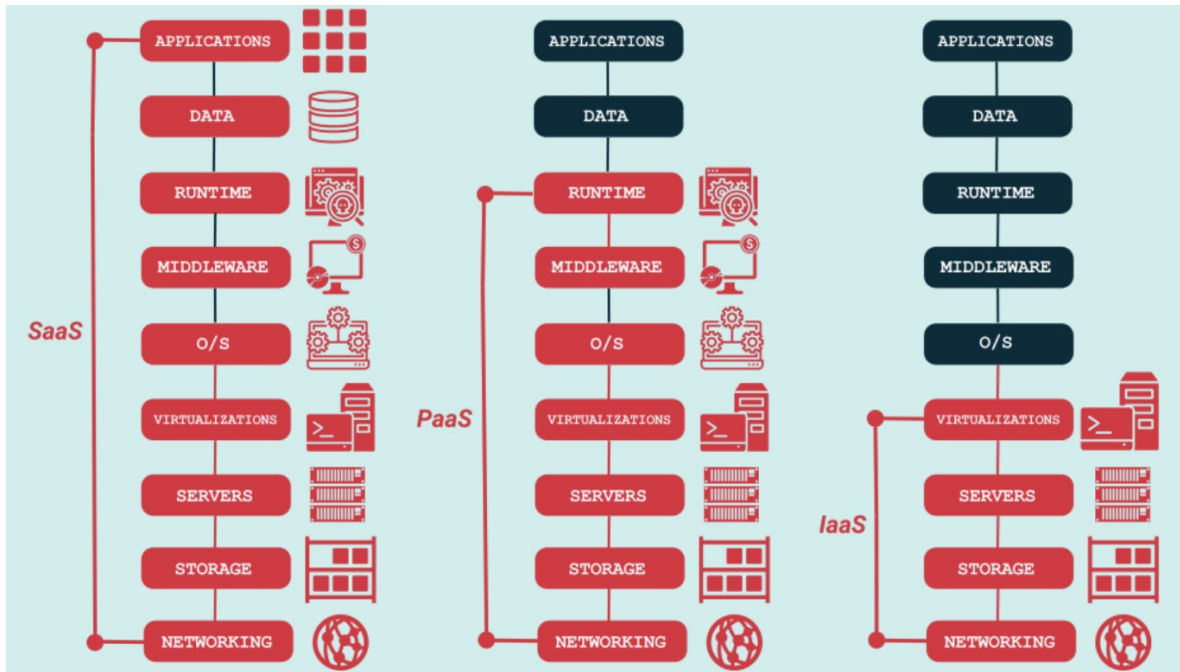
CL-as-a-Service system overview

History “as-a-service”

Different variant

☐ IaaS, PaaS, SaaS

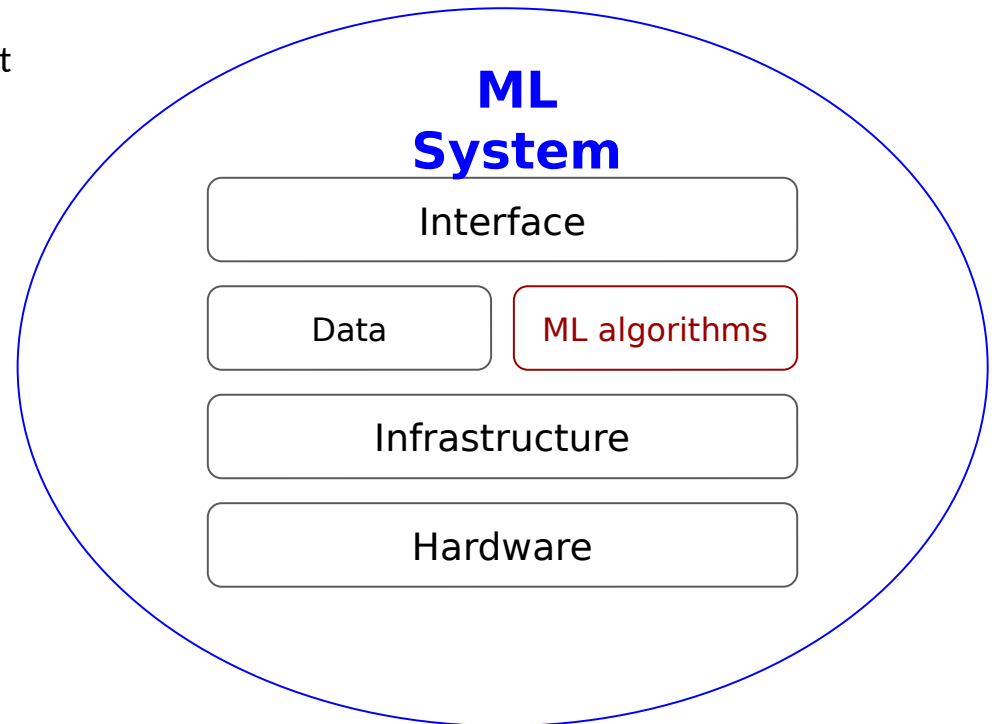
☑ Web-app: Backend-as-a-Services



CL-as-a-Service system overview

AI & ML world:

- AlaaS vs MLaaS vs CLaaS
- Helps organizations incorporate AI functionality with no expertise
- AI service, framework, and workflows built upon these infrastructures are offered to final customers for various use cases.



CL-as-a-Service system overview

AlaaS - types

Divided into distinct types

- Cognitive computing APIs
- Bots and digital assistance
- Fully-managed machine learning services
- Machine learning frameworks

AlaaS - advantages

Multitude benefits

- Reduced cost
- Ease of use
- Scalability and flexibility
 - Pay for what they use
 - Serverless computing
- Ecosystem growth and integration

CL-as-a-Service system overview

User and Core Features

Solution for

- Companies innovation toward real-time and continual machine learning system
- Startup and Companies at the early stage of the machine learning innovation
- MLOps solutions (as a platform and toolkit)
- Data Scientists and Developers to prepare, build, train, and deploy high-quality machine learning models in a fast and unique environment

All-in in a faster cheaper efficient fashion without loss in performance

This AlaaS infrastructure is mainly focused on **continual learning algorithm and tools** to optimize resource efficiency and therefore reduce the cost of the update ML model even with streaming data sequences

CLaaS is enable of fast, efficient, stateful and automated model updating and data drift detection.

CL-as-a-Service system overview

Cloud-based (indirect competitors)



GCP ML
(Vertex AI for MLOps)



**Amazon
SageMaker**

AWS
(SageMaker)



Azure ML
(mainly MLOps)

CL-as-a-Service (Toy) DEMO

Base use cases

As a	<role> (who)	I want to	<requirement> (what)	so that	<value> (why)
As a	User	I want to	setup the initial/next experiment with my options	so that	can then exec the experiment
As a	User	I want to	execute the experiment (setting done)	so that	can train the model continually
As a	User	I want to	retrieve the final results and model (training end)	so that	monitor the final results and save the model



URI	Request Type	Description
/base-scenario/setup/	POST	setup exp with options
/base-scenario/username /exec/exp_name	POST	train the model with CL for and return status training
/base-scenario/username /results/exp_name	GET	retrieve the final results (dict/json) and model (save-path)

REST API

CL-as-a-Service (Toy) DEMO

Tools



docker
Compose



Flask

PYTORCH

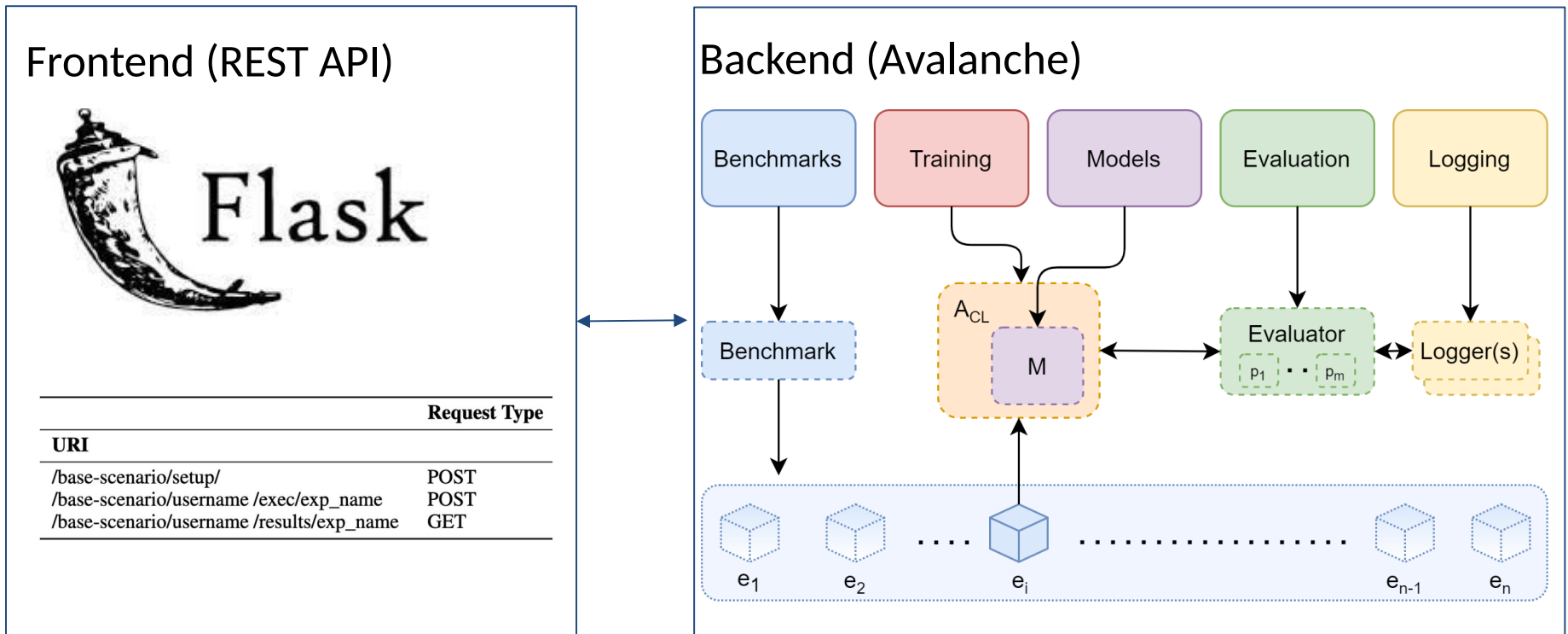


POSTMAN

CL-as-a-Service (Toy) DEMO

LIVE DEMO

CL-as-a-Service (Toy) DEMO



Key Takeaways

Summary

- Analysis
- CL research and tools solutions
- CLaaS system

Possible future works and research direction

- Set of PoC (Proof of concept)
- Novel CL hybrid strategies, intersersection with Auto-ML, Meta-Learning
- ML system platform as a Proto (MVP) for early adopter companies

Do you have any questions?

THANKS



powered by



Continual **AI**



UNIVERSITÀ DI PISA



Rudy Semola



<https://github.com/rudysemola>

r.semola@phd.unipi.it

rudysemola@gmail.com