universidade de aveiro

Milestone 4: Transition

ML/AI model evaluation platform

Plataforma de avaliação de modelos ML/AI

Course: "Projeto em Informática"
Supervisor: Prof. Mário Luís Pinto Antunes

2023-06-06

Team:

Diogo Oliveira Magalhães	102470
Rafael Fernandes Gonçalves	102534
Leonardo Almeida	102536
Emanuel Gaspar Marques	102565
Pedro Henrique Figueiredo Rodrigues	102778

Our team

Mário Antunes

Supervisor

Leonardo Almeida

Team Manager
Front-end developer

Rafael Gonçalves

Product Owner
Back-end developer

Pedro Rodrigues

Architect
Front-end developer

Emanuel Marques

DevOps Master
Back-end developer

Diogo Magalhães

QA Tester
Developer (IdP)

Context

The need to handle **large amounts of data** and advances in processing technologies have led to the mass development of intelligent systems.

Machine Learning (ML) algorithms are often applied to optimise **multiple real-life scenarios**, leading to cost saving and increased productivity.

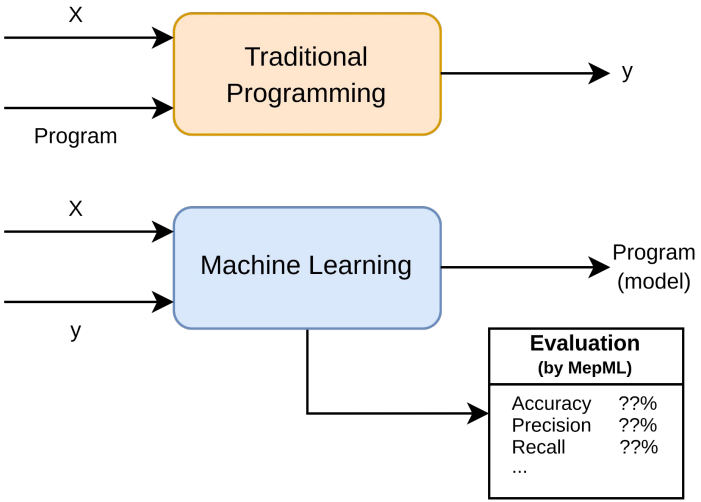
ML/AI examples

ChatGPT

GitHub Copilot

That's why promoting a better ML/AI education is so important!

Our product



MepML

Concept:
ML/AI model evaluation system, based on performance metrics provided by teachers.

The models are trained by users, who then **just upload the results**.

The server side performs a quick **analysis** of the models.

Related Work

	Kaggle	CodaLab	Weights and Biases	Machine Hack	MepML
Professor can create exercises and students can join them	✓	✓	✗	✗	✓
Define exercise visibility	✓	✓	✗	✗	✓
Exercise Leaderboard	✓	✓	✗	✓	✓
Define exercise deadline	✓	✓	✗	✓	✓
Define exercise maximum number of tries	✗	✓	✗	✗	✓
Add/remove/import students to a restricted group	✗	✗	✗	✗	✓
Add new and reuse Metrics	✗	✗	✗	✗	✓
Use UA IdP to authenticate students and professors	✗	✗	✗	✗	✓
Check assigned exercises	✓	✓	✓	✓	✓
Download exercise related content	✓	✓	✓	✓	✓
View Students Code and results	✓	✗	✗	✗	✓
Students and professor user types	✗	✗	✗	✗	✓

Teacher workflow

What will the Teacher upload?

Training dataset

X_train

Delay	Class	Wi-Fi	Temperature	Satisfied
3 min.	1st	Yes	20ºC	Yes
15 min.	2nd	Yes	17ºC	No
1 min.	2nd	Yes	34ºC	No
…	…	…	…	…

y_train

Test dataset

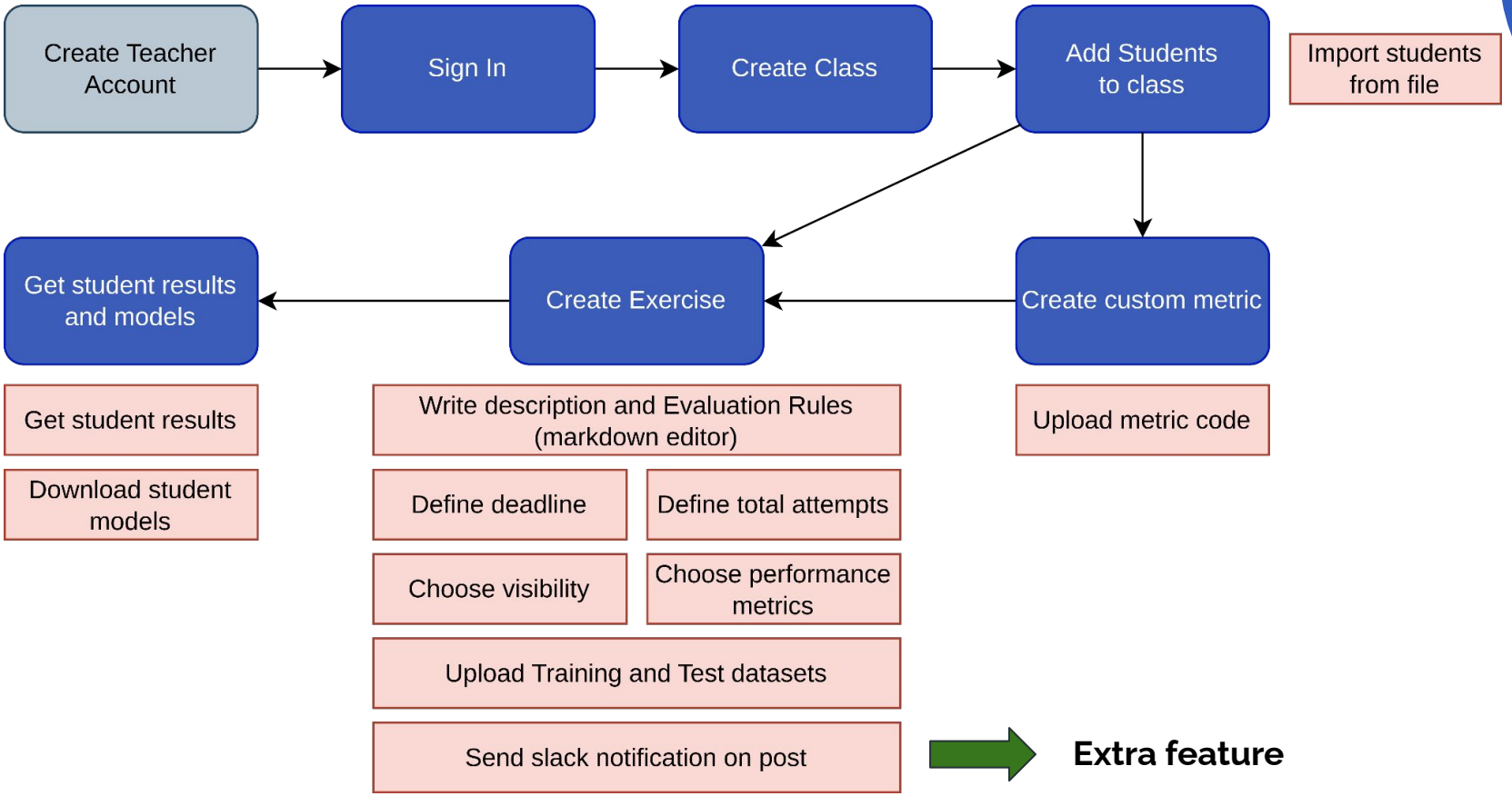
X_test

Delay	Class	Wi-Fi	Temperature	Satisfied
0 min.	2nd	Yes	22ºC	Yes
30 min.	1st	No	18ºC	No
…	…	…	…	…

y_test

Students can't see it

Teacher workflow



Student workflow

Student solves the exercise

Delay	Class	Wi-Fi	Temperature	Satisfied
0 min.	2nd	Yes	22°C	Yes
30 min.	1st	No	18°C	No
2 min.	1st	Yes	19°C	Yes
16 min.	2nd	Yes	17°C	No
3 min.	2nd	Yes	34°C	No

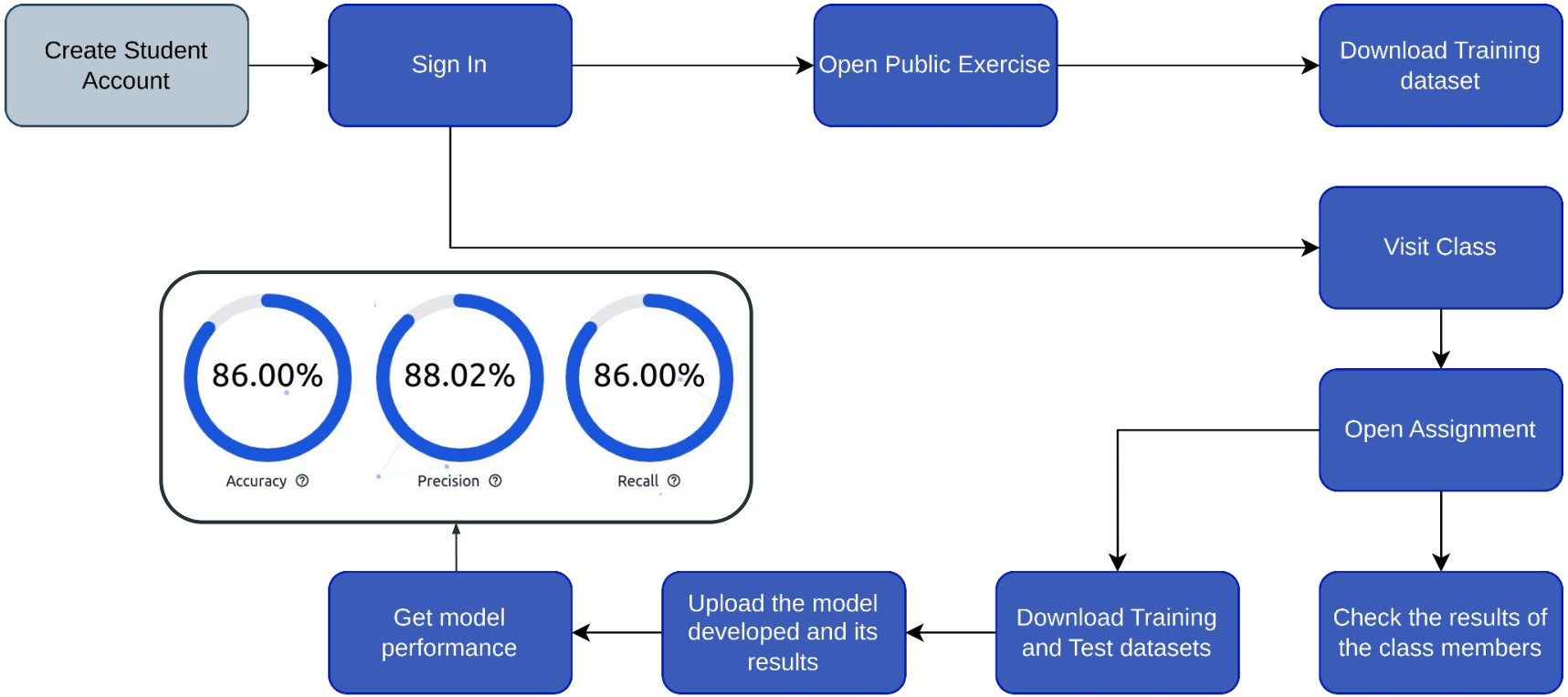
X_test y_test

Delay	Class	Wi-Fi	Temperature	Satisfied	Prediction
0	2	1	22	1	1
30	1	0	18	0	0
2	1	1	19	1	1
16	2	1	17	0	1
3	2	1	34	0	0

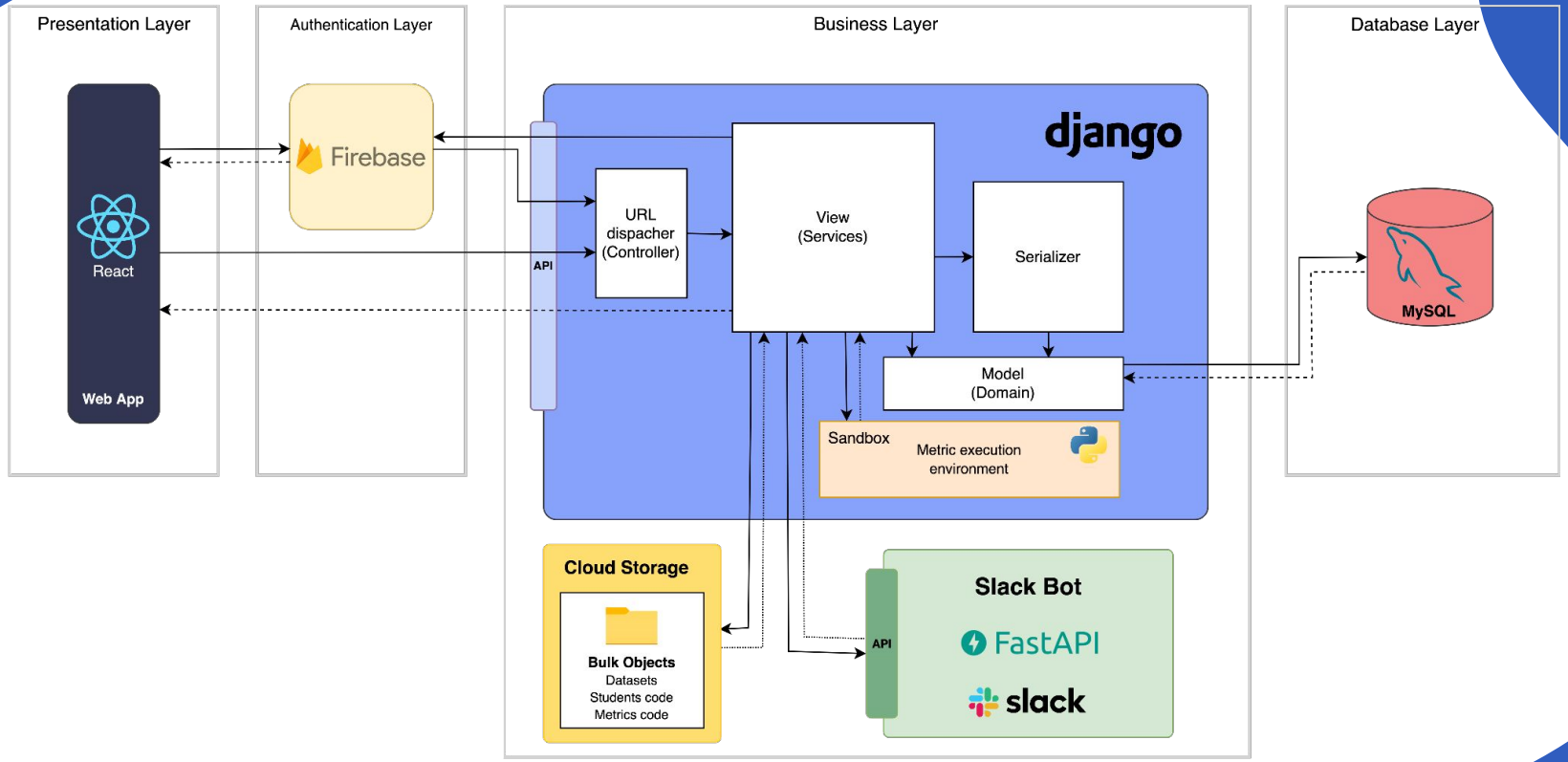
y_pred

Students only **know** X_test
and are trying to **predict** y_test.

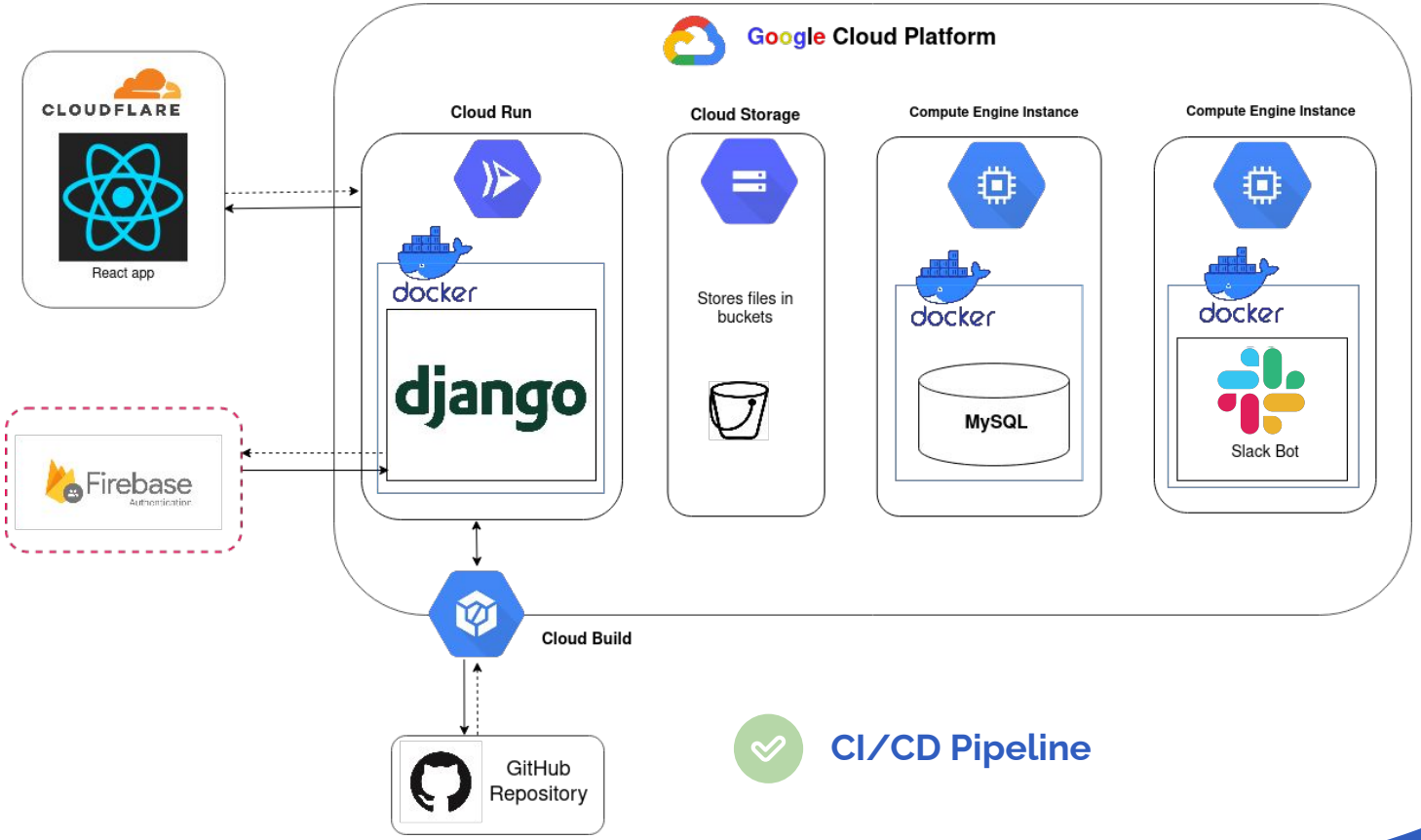
Student workflow



System Architecture



Deployment Architecture



CI/CD Pipeline

Quality Assurance

- Limited Development time
- Still learning Quality Assurance
- Low quantity of developers

Usability Tests

Usability Tests

Why?
Validate the final product with real users

Who?
Students and professor of CAA course

How?
Performing tasks & Usability questionnaire

Results?
Identification of major usability issues & interesting functionalities

What We Learned

- Understand each other differences
- Better communication with all the people involved
- Applying agile methodology
- Importance of planning and team organisation

Future Work

- UA IdP implementation
- Usability tests
- Feedback analysis
- Web analytics monitoring
- More types of ML problems
- More dataset formats
- Kubernetes-oriented deployment

Resources

- https://slidesgo.com/theme/retato-slideshow#position-38&results-1357
- https://storyset.com
- https://www.pexels.com/
- https://www.utm.mx/~caff/doc/OpenUPWeb/index.htm
- https://www.kaggle.com
- https://wandb.ai/site
- https://machinehack.com/
- https://codalab.lisn.upsaclay.fr
- https://online.visual-paradigm.com/pt/
- https://www.analyticsinsight.net/top-5-model-evaluation-metrics-for-machine-learning-projects/
- https://blog.idexlab.com/state-of-the-art-example
- https://towardsdatascience.com/4-data-science-competition-platforms-other-than-kaggle-6d1795ff46a

Relevant paper:

Alex Serban, Koen van der Blom, Holger Hoos, and Joost Visser. 2020. **Adoption and Effects of Software Engineering Best Practices in Machine Learning**. In *Proceedings of the 14th ACM / IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM)*, ACM . DOI:https://doi.org/10.1145/3382494.3410681

ML research has much to improve, hence the need for our platform.

Join us at
mepml.pages.dev

MepML

Thank you!