

Model Evaluation Platform ML/AI

Diogo Magalhães¹[0009-0008-4323-2692], Emanuel Marques¹[0009-0001-4017-8351],
Leonardo Almeida¹[0009-0006-5669-2378], Pedro Rodrigues¹[0009-0009-5916-4062],
Rafael Gonçalves¹[0009-0002-8628-8262], and Mário
Antunes¹[0000-0002-6504-9441]

DETI, University of Aveiro, Aveiro, Portugal
[d.magalhaes, emanuel.gmarques, leonardoalmeida7,
pedrofrodrigues4, rfg, mario.antunes]@ua.pt

Abstract. The need to handle large amounts of data and advances in processing technologies have led to the mass development of intelligent systems. Machine Learning algorithms are often applied to optimise various real-life scenarios, leading to cost savings and increased productivity. Given the urgency to promote a better education in this field of study, we developed an intuitive and collaborative platform, where students can improve their skills, by solving exercises created by teachers and uploading the results of their models for automatic evaluation.

Keywords: Education · Model Evaluation · Performance Metrics

1 Introduction

Contrary to traditional software education, validating the results from Machine Learning (ML) applications is considerably harder [2]. To tackle this issue there are some platforms that allow one to issue ML competitions where the participants receive feedback on their models. However, these platforms are not aligned for education. In fact, they have some aspects that limit their applicability in the university environment, namely they do not support the assignment of exercises to a restricted group of students, there is no clear distinction between the roles of student and teacher as actors in the systems. In addition, such platform should be modular to allow the addition of new evaluation metrics, support new types of learning problems and allows for the improve data and results visualisation.

2 Proposed Solution

We developed a multi-layered application, where teachers can create and manage exercises, import and manage their classes, track the students' progress, and add new evaluation metrics, while students can view or solve exercises, get quick feedback on their models and rank themselves among their peers. The platform is preconfigured with the most commonly used metrics for evaluating model performance. This selection was based on a comprehensive study of widely employed metrics in the field [1].

The proposed solution architecture can be found in Figure 1. It is important to mention that the platform provides the exercises and evaluates the results from the students. It does not execute the students code.

We should not assume that all users are trustworthy and allow them to run arbitrary code on our server. So, to run the performance metrics, we relied on a sandbox environment, where Python methods are restricted to those that are strictly necessary.

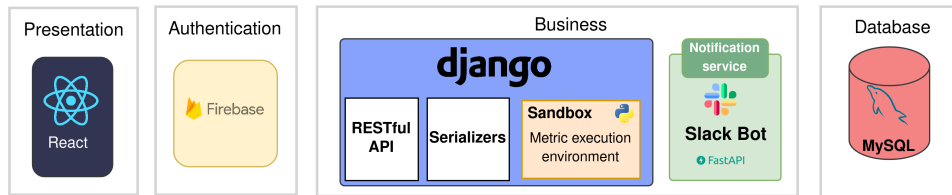


Fig. 1. Proposed solution architecture.

3 Evaluation

To assess how easily and intuitively users interact with the system, identify any potential usability issues, and collect valuable feedback for system enhancement, we conducted usability tests. These tests were carried out in a class of one of the Machine Learning courses offered as part of the Master's degree program in Informatics Engineering and Data Science at the University of Aveiro.

The feedback we received from the users was overwhelmingly positive, as they found the system easy to navigate and intuitive. However, users also provided us with valuable suggestions that could be incorporated into future iterations of the system, aiming to further improve user experience and overall usability.

4 Conclusions

The proposed solution was developed as a tool for helping the education of ML within the University of Aveiro. The initial tests show that the platform is intuitive and easy to use. It also provides a simpler way for the teaching personnel to guide the students.

References

1. Blagec, K., Dorffner, G., Moradi, M., Samwald, M.: A critical analysis of metrics used for measuring progress in artificial intelligence pp. 4–12 (08 2020)
2. Sanusi, I.T.: Teaching machine learning in k-12 education. In: Proceedings of the 17th ACM Conference on International Computing Education Research. ACM (aug 2021). <https://doi.org/10.1145/3446871.3469769>