Service ID S00185

Location Austria, Remote

Performance evaluation of an AI model

Provider service

Josephinum Research (JR)

Link to content

https://agrifoodtef.eu/catalogue-of-services/performance-evaluation-ai-model

Type of Sector

Arable farming, Horticulture, Livestock farming, Tree Crops, Viticulture

Accepted type of products

Software or AI model

Type of service

Performance evaluation

Description

Our service provides a thorough evaluation of AI models by utilising a wide range of benchmark datasets—from images and sensor data to remote sensing and multimodal data. This comprehensive testing assesses the model's performance across standard metrics, specifically chosen based on the model's intended tasks. The results offer insights into the model's effectiveness and accuracy, guiding developers in refining their AI solutions to achieve optimal performance in real-world applications.

How can the service help you

The service helps customers by evaluating the performance of their AI models with diverse benchmark datasets. Before the service, customers may lack clarity on how their AI models perform in specific real-world applications. After the service, they gain detailed insights into the effectiveness, accuracy, and potential improvement areas of their AI models, enabling them to optimise performance for targeted use cases.

How the service will be delivered

Customisation options include selecting specific datasets, performance metrics, and testing scenarios aligned with the AI model's applications. Customers may request a focus on certain data types or performance benchmarks relevant to their use case. Limitations include compatibility with the provided datasets. If no benchmark dataset is available, the data can be collected in an additional service.

Service customisation

The service will be delivered remotely, leveraging a suite of benchmark datasets tailored to the AI model's intended use cases. The evaluation process includes multiple iterations to ensure comprehensive coverage of performance metrics. The duration of the service depends on the model's complexity and the volume of testing required, generally taking a few days to weeks. Customers receive a detailed performance report, which includes accuracy metrics, insights into strengths and weaknesses, and recommendations for improvement. Customers must provide their AI models and weights in a standard file format (e.g., .pb, .pkl, .yaml, .onnx, .h5) and any related documentation, including intended use case details.