

Service ID S00259

Location Remote



Evaluation of AI model for agrifood applications

Provider service

GRADIANT

Link to content

<https://agrifoodtef.eu/catalogue-of-services/evaluation-ai-model-agrifood-applications>

Type of Sector

Arable farming, Food processing, Greenhouse, Horticulture, Livestock farming, Tree Crops, Viticulture, Aquaculture, Forestry

Accepted type of products

Data, Design / Documentation, Software or AI model

Type of service

Performance evaluation

Description

This service helps businesses and organisations evaluate how well their AI models perform in real-world conditions. We conduct thorough testing of your AI model using either your own dataset or our reference datasets, depending on your needs (see the related services). Our evaluation process goes beyond basic accuracy metrics to provide a comprehensive understanding of your model's strengths and limitations. We analyse various aspects of performance that matter for your specific use case, whether it's detecting diseases in crops, sorting produce, or monitoring livestock behaviour. We work with you to determine the most relevant performance metrics for your application, taking into consideration factors like accuracy, speed, reliability, and resource usage. This helps you understand if your AI solution is ready for deployment, needs improvement, or requires adjustments for specific conditions. The evaluation provides clear, actionable insights about your model's performance, helping you make informed decisions about its readiness for real-world agricultural applications.

How can the service help you

This service addresses a crucial challenge faced by companies developing AI solutions for agriculture: knowing whether their AI models are truly ready for real-world deployment. Before using our service, you may be uncertain about your model's actual performance capabilities and limitations.

You might have questions about how well it handles different scenarios, edge cases, or varying conditions typical in agricultural settings. After using our evaluation service, you'll have a clear, comprehensive understanding of your AI model's performance. You'll know exactly how well it performs across different metrics that matter for your specific use case. For example, if you have an AI model for crop disease detection, you'll learn not just its overall accuracy, but also its false positive rate, how well it performs under different lighting conditions, and its speed of detection. This information helps you make confident decisions about deployment, identify areas needing improvement, and understand any limitations that need to be addressed before putting the model into production.

The service transforms uncertainty about your AI model's capabilities into actionable insights, helping you save time and resources by identifying potential issues before real-world deployment. This knowledge is particularly valuable for meeting

How the service will be delivered

The evaluation service can be tailored in several ways to match your specific needs. We can customise the evaluation metrics based on your application—for instance, if response time is critical for your use case, we can emphasise performance speed and latency measurements.

If accuracy in specific conditions is paramount, we can focus on detailed analysis of edge cases and failure modes.

- The testing process can be customised through:
- Selection of reference datasets best matching your deployment environment
- Definition of custom performance thresholds and success criteria - Focus on specific aspects such as robustness, fairness, or resource efficiency.
- Custom evaluation scenarios that simulate your specific use conditions

Service customisation

The service execution begins with an initial consultation to understand your AI model and specific evaluation needs. You'll need to provide us with your trained AI model and, if applicable, your dataset. If you're using TEF available datasets, we'll help you select the most appropriate ones for your use case.

Finally, if you need a new dataset tailored to this specific case, we offer some other services to create datasets in different scenarios (see S00243 or S00254 in the related services section). The evaluation process typically takes 2-4 weeks, depending on the complexity of your model and the required scope of testing. The evaluation is performed remotely through our secure testing environment, meaning there's no requirement for your physical presence. Upon completion, you'll receive a comprehensive evaluation report that includes detailed performance metrics, visualisations of the results, and specific insights about your model's behaviour.

The report will highlight both strengths and areas for potential improvement, along with recommendations for optimisation if needed. We also provide a technical appendix with all raw performance data for your reference. To begin the service, you need to provide your trained AI model in a standard format (we support most common frameworks, like, for example, TensorFlow, PyTorch, XGBoost, or scikit-learn, among others), documentation describing your model's intended use and current implementation, and if you're using your own dataset, the properly formatted data along with any relevant annotations or labels.