**EMPIR Call 2020 – Fundamental, Industry and Normative**

Selected Research Topic number: **SRT-f05**
Version: 1.0

# Title: Trustworthy machine learning


## Abstract

Machine learning (ML) is being adopted in a wide variety of metrology applications in which the underlying physical model is not well understood, including medical imaging, advanced manufacturing, autonomous transport and communications. However, the adoption of ML with all its benefits is hindered by the perceived untrustworthiness of its outputs. There is a pressing need to understand how to incorporate ML into a metrology framework. The proposal will address solution for evaluation the uncertainty of the output of ML



## Keywords

Machine learning, deep neural networks, validation, uncertainty quantification, robustness, interpretability, data-driven models, good practice guides


## Background to the Metrological Challenges

There have been huge advances in recent years in the capability of machine learning (and especially deep neural networks) to build accurate data-driven predictive models, thanks in large part to the availability of increasing volumes of data and advances in computational processing power. However, most of the academic research into machine learning has been focused upon the goal of optimising predictive accuracy, and relatively less attention has been directed towards providing confidence in the output of machine learning algorithms. Provided assumptions on model misfit can be made, uncertainty evaluation has long been established for algorithms based on classical statistical modelling such as linear and kernel regression. However, uncertainty evaluation is much more challenging for machine learning algorithms, such as deep neural networks. It is vital then that the predictions made by machine learning algorithms can be trusted, the various aspects of which can be distilled into three categories: uncertainty evaluation, robustness and interpretability.

A wide range of techniques for assessing and improving the robustness of deep neural networks have been proposed in the literature, most notably the use of adversarial training and data augmentation. However, there is a need to understand how these methods compare, which factors affect the choice of method, and how to apply them in the context of the rapidly developing area of transfer learning.

Machine learning is currently being employed within a range of European metrology projects, including various types of medical imaging (magnetic resonance, X-Ray and positron emission tomography (PET)), analysis of electrocardiogram (ECG) signals, digital pathology, free form surface reconstruction, mass spectrometry, nanoparticle image segmentation and reconstruction, and energy systems modelling. While trustworthy machine learning is being investigated in specific instances in the metrology community, the fundamental questions are not currently being addressed in a systematic way. Example applications that showcase the techniques developed and good practice guides are needed to increase trustworthiness.

## Objectives

Proposers should address the objectives stated below, which are based on the PRT submissions. Proposers may identify amendments to the objectives or choose to address a subset of them in order to maximise the overall impact, or address budgetary or scientific / technical constraints, but the reasons for this should be clearly stated in the protocol.

The JRP shall focus on metrology research necessary to support standardisation in machine learning and data-driven models.


**EURAMET EMPIR MSU**
Phone: +44 20 8943 6666 Email: EMPIR.msu@euramet.org
msu.euramet.org



The EMPIR initiative is co-funded by the European Union's Horizon 2020 research and innovation programme and the EMPIR Participating States

The specific objectives are

1. To develop fundamental methods, based on, for example, Monte Carlo and Gaussian dropout, variational inference or deep ensembles for the evaluation of uncertainties, in both training and test data, associated with the results of machine learning models such as deep neural networks.
2. To develop fundamental methods to evaluate the sensitivity of machine learning outputs to both random and systematic effects, along with approaches for improving robustness such as data augmentation and adversarial training. Assessing and optimising generalisability to new data sets should also be addressed in the context of transfer learning.
3. To explore the use of inherently transparent methods for machine learning in a metrology context, such as kernel linear regression, and investigate methods for extracting interpretable information from non-interpretable models such as neural networks.
4. To develop good practice guides on uncertainty quantification, robustness and interpretability of machine learning models and lay the foundations for the extension of the GUM framework. Case studies showcasing possible applications of machine learning, such as analysis of electrocardiogram (ECG) signals, free form surface reconstruction, nanoparticle characterisation, smart grids, digital pathology and mass spectrometry should be included.
5. To facilitate the take up of the technology and measurement infrastructure developed in the project by the measurement supply chain, standards developing organisations (ISO/IEC 20546, ISO/IEC 20547, ISO/TC 69/WG 12) and e.g. end users in medical imaging, advanced manufacturing, autonomous transport and communications

Proposers shall give priority to work that aims at excellent science exploring new techniques or methods for metrology and novel primary measurement standards, and brings together the best scientists in Europe and beyond, whilst exploiting the unique capabilities of the National Metrology Institutes and Designated Institutes.

Proposers should establish the current state of the art, and explain how their proposed project goes beyond this.

EURAMET expects the average EU Contribution for the selected JRPs in this TP to be 1.5 M€, and has defined an upper limit of 1.8 M€ for this project.

EURAMET also expects the EU Contribution to the external funded partners to not exceed 40 % of the total EU Contribution across all selected projects in this TP.

## Potential Impact

Proposals must demonstrate adequate and appropriate participation/links to the “end user” community, describing how the project partners will engage with relevant communities during the project to facilitate knowledge transfer and accelerate the uptake of project outputs. Evidence of support from the “end user” community (e.g. letters of support) is also encouraged.

You should detail how your JRP results are going to:

- Address the SRT objectives and deliver solutions to the documented needs,
- Feed into the development of urgent documentary standards through appropriate standards bodies,
- Transfer knowledge to the technology sector.

You should detail other impacts of your proposed JRP as specified in the document “Guide 4: Writing Joint Research Projects (JRPs)”

You should also detail how your approach to realising the objectives will further the aim of EMPIR to develop a coherent approach at the European level in the field of metrology and include the best available contributions from across the metrology community. Specifically, the opportunities for:

- improvement of the efficiency of use of available resources to better meet metrological needs and to assure the traceability of national standards
- the metrology capacity of EURAMET Member States whose metrology programmes are at an early stage of development to be increased
- organisations other than NMIs and DIs to be involved in the work.

## Time-scale

The project should be of up to 3 years duration.