European Partnership on Metrology
Call 2023 –Fundamental Metrology, Industry, Normative and Research Potential

Selected Research Topic number: **SRT-f14**
Version: 1.0

Title: Machine learning for metrology

Abstract

Machine learning (ML) has the potential to bring transformation to a wide variety of metrology applications. However, ML methods in metrology still require significant work to establish trustworthiness and traceability. A generic metrological framework for uncertainty evaluation, explainability and terminology of ML models needs to be created. In addition, novel methods addressing the metrological challenges of uncertainty evaluation, uncertainty budgeting and sensitivity analysis need to be developed. This will create new fundamental methods for addressing the challenges of trustworthy ML, benefit Artificial Intelligence and data-driven modelling and result in the identification of new applications of metrology.

Keywords

Machine learning, uncertainty evaluation, sensitivity analysis, explainability, validation, neural networks

Background to the Metrological Challenges

ML is currently being employed within a range of European metrology projects, including various types of medical imaging (magnetic resonance, X-Ray and PET), analysis of ECG and PPG signals, digital pathology, free-form surface reconstruction, mass spectrometry, scatterometry, nanoparticles image segmentation and reconstruction and energy systems modelling. Machine learning has made huge advances in recent years in building accurate data-driven predictive models. This is mainly due to the availability of large volumes of data and advances in computational processing power. However, the widespread adoption of ML and the AI systems which they enable is hindered by the perceived unreliability of their outputs. It is vital then that the predictions made by ML algorithms can be trusted, especially when calculating uncertainties and when interpreting predictions.

Two fundamental issues need to be solved in order to give reliability to ML output data:

Uncertainty evaluation: It is fundamental that ML predictions are quality assured by being accompanied by a reliable quantitative assessment of the associated uncertainty. Many of the most effective ML approaches, such as deep neural networks, are challenging to analyse mathematically. This fact, together with their large-scale nature, make classical approaches for uncertainty evaluation, like those proposed by the GUM framework, not directly applicable. There has been significant recent research devoted to uncertainty evaluation of ML models and various methods have been proposed. However, there has been no systematic investigation into the requirements of uncertainty evaluation from a metrological point of view, and existing methods have deficiencies that need to be addressed before they meet the needs of metrological applications.

Explainability: Confidence in an ML model cannot be established unless its predictions can be explained in their physical context. Explainability of ML models is the understanding of which features in the data contributed to the model’s prediction, why they did it and to which extent (for example, by identifying which input variables or intermediate nodes of a network are most relevant in a classification task). An important aspect related to explainability is the interpretability of ML models, that is, the ease with which a human can understand their inner workings. ML models, in this respect, can vary from black-box models, inherently not interpretable, to opaque ones, up to white-box models, which are able to provide results clearly understandable to experts in the domain. Various methods for making deep neural networks interpretable have been proposed in the literature, though the focus is largely on image classification and a systematic investigation into the requirements of interpretability from a metrology point of view is still needed.

Research into trustworthy ML has yet to make its way into good practice guides and standards. Some standards dedicated to the topic of big data analytics are available or under development, such as ISO/IEC 20546 “Information technology - Big data - Overview and vocabulary” and ISO/IEC 20547 “Information technology - Big data reference architecture”. In connection to those, the Big Data Analytics Working Group of

ISO TC 69 “Applications of statistical methods” (ISO/TC 69/WG 12) is working on standards concerning Vocabulary and Symbols, Data Science Life Cycle and Model Validation.

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 metrology research necessary to support standardisation in machine learning (ML).

The specific objectives are:

1. To establish a framework for embedding ML in metrology by:
    a) Establishing a vocabulary that embeds ML terminology into the metrological context addressing terms related to different types of uncertainty.
    b) Defining benchmarking methods for assessing uncertainty evaluation and explainability methods for ML in metrological applications and compare them with classical ones.
    c) Validate state-of-the-art methods for uncertainty evaluation and explainability of ML models using the benchmarking methods used in 1b.
2. To develop reliable and scalable methods for evaluating the uncertainty of ML outputs. Methods should take account of model uncertainty, data uncertainty in input variables, dependencies between variables, errors in labels of training data and also uncertainty due to missing data.
3. To develop explainability and sensitivity analysis methods in order to establish for ML models a standardised counterpart to classical uncertainty budgets and classical sensitivity analysis. The developed methods would be applicable to typical metrological data, including time series data, and take into account all known data uncertainties.
4. To implement the developed methods and apply the framework developed in obj. 1 in the context of at least three metrological case studies to support present and future applications in metrology, in close collaboration with stakeholders from industry, the utility sector and/or government agencies, depending on the selected use cases.
5. To lay the foundation for an extension of the VIM and the GUM and to facilitate the take up of the technology and measurement infrastructure developed in the project by standards developing organisations (JCGM-WG1 & WG2, ISO/TC69, CEN/CENELEC JTC21), by the European Metrology Network for Mathematics and Statistics (EMN Mathmet) and end users (medical imaging autonomous transport and smart factories).

These objectives will require large-scale approaches that are beyond the capabilities of single National Metrology Institutes and Designated Institutes. 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, including other European Partnerships, 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. In particular, proposers should outline the achievements of the EMPIR 17IND12 Met4FoF, EMPIR 17IND02 SmartCom, and how their proposal will build on those.

EURAMET expects the average EU Contribution for the selected JRPs in this TP to be 2.0 M€ and has defined an upper limit of 2.5 M€ for this project.

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

Any industrial beneficiaries that will receive significant benefit from the results of the proposed project are expected to be beneficiaries without receiving funding or associated partners.

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,
- Facilitate improved industrial capability or improved quality of life for European citizens in terms of personal health, protection of the environment and the climate, or energy security,
- Transfer knowledge to the Artificial Intelligence 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 the Partnership 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.