NAME : Antoine Marot
COUNTRY : France
REGISTRATION NUMBER :

GROUP REF. : C2
PREF. SUBJECT : 1
QUESTION N° : 8

Developing Framework & Benchmark & Organisation for AI

Recently, the field of artificial intelligence (AI) went through a deep learning revolution, which drastically enlarges its potential for real-world application. Deep-learning AI systems are able to digest large volumes of information, memorize historical datasets, and learn to quickly infer effective actions in context by taking time horizons and forecast uncertainties into account. This revival started through the ImageNet Large Dataset [1] and Benchmark push in 2010 leading to a striking result at the 2012 "Visual Recognition challenge" with a deep neural network approach outperforming all others. Since then, Datasets, Competitions and Benchmarks have been driving major advances in the AI community such as in NLP [2,3] or protein folding [4]. NeurIPS largest scientific conference for AI encourages new releases of such datasets, competitions and benchmarks every year through dedicated tracks. This is how AI power and developments are first harnessed.

To drive similar progress, formulation of power-system use cases and challenges in an AI-friendly manner would be a catalyst. Such an example is the "Learning to run a power network" [5,6,7] series of competition which have demonstrated a real potential for AI to assist for action recommendations, with the contribution of AI teams without initial power system domain knowledge. It relied on the release of a large synthetic dataset and the AI-friendly Grid2op Framework [8] which models and emulates system operation scenarios and leverage available powerflow simulator of one's choice. Given real-world data confidentiality or scarcity, synthetic data generation will probably become a key driver. Among other AI-oriented datasets and benchmarks, we can mention the comprehensive RE-Europe dataset [9] for operational, investment or evaluation studies with markets. More recently, we can list the LIPS benchmark (Learning Industrial Physical Simulation) [10] which evaluates learnt model of powerflow simulator, or "A multi-scale time-series dataset in decarbonized grids" [11] for dynamic disturbance events; robust hierarchical forecasting of load and renewable energy with uncertainties and extreme events; and realistic synthetic generation of measurement time series. Such open developments and common ground will enable harnessing the power of AI for power systems.

In addition to open-source frameworks (Tensorflow, Pytorch, OpenGym, HuggingFace) and libraries, open and reproducible papers with code, and besides compute resources, the rapid success of Deep Learning was also largely possible thanks to a new organisational mindset relying on short iteration and feedback cycles with mutualized ressources between R&D and Deployment such as at Google, Deepmind, OpenAI, Microsoft, Amazon, NVIDIA.

In order to enable and ultimately master the design, development, validation, training and deployment of AI solutions, each organisation indeed needs to develop sufficient capabilities in the data, platform and people pillars [12].

Beside open-datasets and benchmarks, the data pillar is about organising data in terms of data accessibility, governance, and quality. In the majority of cases, availability of both historical and real-time process data is of key enabling importance to swiftly start with an AI initiative. Within the TSO/DSO organisation, it is often worth considering setting up a dedicated database to store structured and non-structured data for long-term, and to consolidate and democratize data sources, improve data quality and security via central data governance, and simplify data access. In other words, an ambitious digitalization step is crucial within companies since quality and rich data is an essential raw material: it should now be considered as a company product/asset.

The people pillar represents human resources including skill, culture, and organisation required for the successful launching of AI initiatives, their realisation, as well as their long-term maintenance. In general, a broad range of different human expertise and skills are necessary to identify, design, build, validate, and maintain AI solutions. Also to educate others about AI possibilities and promote AI developments internally in an organisation and connect and co-develop with similar AI communities externally. Finally, best practices including ethic guidelines when developing, deploying and maintaining AI solutions should be defined.

The platform pillar is about tools and technology used to develop, test and implement AI solutions. For development, platforms should be developer-friendly and supported by set of data analytics, visualisation, and programming tools, database pipelines, scalable compute and storage resources, in order to quickly develop prototypes and demonstrate value. In addition, the platform should be flexible enough to enable seamless yet secure cooperation with external parties when co-developing AI solutions. For a production environment, scalability and reliability are key aspects. Crucial is the ability to support continuous integration and deployment of new improved versions of the AI solution for maintenance purpose without affecting the rest of the system. C3.AI [13] platform could be highlighted as a first example that offers a comprehensive technical stack towards such generic platform.

The above three pillars mostly focus on required enablers for AI solution realisation within a company. Nevertheless, for a seamless and rapid production cycle from design to implementation of AI solutions, the company needs to be organised to facilitate collaboration between Research, Development, and Deployment innovation phases. However, the process of technology innovation often takes place in a linear fashion today in which Data, Code, People, and Testing resources are not shared. The innovation process of largely independent phases prevents the virtuous cycle of rapid development/testing/deployment that other sectors have used to good effect, in particular in AI-driven companies. To gain the most in the shortest time when harnessing AI solutions, it is crucial firstly that each company integrates the phases of the innovation process, and secondly that the integration of system operation and the global AI R&D ecosystem is increased. Near real-world testing environments, providing realistic scenarios and validation criteria, need to be continuously defined and refined back and forth from research to deployment. The use of open data, models and code with open access papers should be encouraged, much like within the LF Energy Initiative (https://www.lfenergy.org/), to enlarge the ecosystem for AI tool development. First steps in this direction are currently taking place, as exemplified by the L2RPN competitions and the European AI on Demand Platform (https://www.ai4europe.eu/).