Christian Jutten | Editor-in-Chief | Christian.Jutten@grenoble-inp.fr

Trusting in the Sciences Requires Explainability

The July issue of *IEEE Signal Processing Magazine* (*SPM*) is a special issue focused on "Explainability in Data Science: Interpretability, Reproducibility, and Replicability." With increased enthusiasm for machine learning, it is a very timely topic, and I invite every IEEE Signal Processing Society (SPS) member to read these very instructive papers.

The "classical" scientific approach is based on models, designed with realistic assumptions on the problem to solve. The main advantage of such an approach is to provide direct explainability of the results. The limitation is the model's complexity, which is limited: it is usual to claim following G. Box: " a model is always wrong, but it can be useful" [1].

With the data deluge we have experienced in recent years, many more data-driven methods have been developed with great success. Their main advantage is that they do not first require the designing of a model. This is also their drawback as many of them are black boxes: explaining their results is tricky and requires a lot of effort and additional experiments. As data science, including machine learning and deep learning, are currently ubiquitous in all domains, explainability in data science is essential, especially for critical application domains like medicine and health, control of autonomous vehicles, and face recognition, to name a few.

In addition to intellectual satisfaction and trust for the user provided by explainability, in the European Union, as it is recalled in the article "Robust Explainability" in this month's issue (page 73), the General Data Protection Regulation adopted in May 2018 states that individuals have the right to an explanation of a decision based on automated processing [2].

This special issue, comprised of nine articles, illustrates very well the meaning and relevance of the key concepts of interpretability, replicability, and reproducibility. It provides a panorama of methods for achieving explainability for 1) different architectures, from two- or N-way arrays and graph representation or neural networks, for 2) different tasks like regression, classification of images or classification of time series, extraction of information from documents or images, and for 3) different applications in health care, or even more surprisingly, in art for understanding how generative/creative algorithms produce new ideas and artifacts [see the "Explaining Artificial Intelligence Generation and Creativity" article in this issue (page 85)].

I close with a few observations. In data-driven approaches, data are the core of the methods and of the quality of the results. Practically, the quality of the data (e.g., noiseless or noisy) and its representativity are essential to achieve relevant results. Poor data can be responsible for wrong or biased results. This was clearly shown in the articles of the SPM special issues on affective computing [3]: basic facial expressions (happiness, surprise, fear, disgust, anger, and sadness) are strongly variable and interdependent of the culture, gender, and so on. This variability can lead to bias in emotion classification. The recent examples in health care [4] report how algorithms trained with (gender or race) imbalanced data can lead to very bad performance for the underrepresented population. In fact, it is not surprising, as in data science, the black-box model is directly inherited from the data x. The output y = f(x), estimated by learning, will have poor generalization for inputs x, which are outside the subspace containing the training data x!

> **Practically, the quality of the data (e.g., noiseless or noisy) and its representativity are essential to achieve relevant results.**

2023: SPS 75th Anniversary

In 2023, we will celebrate the 75th anniversary of the SPS. On this occasion, a special issue of *SPM* will be published in June 2023. For its 50th anniversary, in 1998, the SPS worked with the IEEE History Center to prepare a monograph outlining the history of the Society from 1948 to 1998 [5]. You can also find important dates online at https://ethw.org/IEEE_Signal_Processing_Society_History.

In the 75th Anniversary Special Issue, we focus on what has happened over the last 25 years, the perspectives in SPS, and more generally, in the domain of signal and image processing. A Call for Papers has recently been published, and I invite authors to contribute to this special issue by submitting papers tracing the evolution during the last 25 years and perspectives on the concepts and high impact progress that has taken place in the areas of signal and image processing. If you have an idea that you want to discuss before submission, don't hesitate to contact me or the SPM area editors.

References

[1] G. E. P. Box, "Science and statistics," *J. Amer. Statist. Assoc.*, vol. 71, no. 356, pp. 791–799, 1976, doi: 10.1080/01621459.1976.10480949.

[2] B. Goodman and S. Flaxman, "European Union regulations on algorithmic decision-making and a 'right to explanation'," *AI Mag.*, vol. 38, no. 3, pp. 50–57, Oct. 2017, doi: 10.1609/aimag.v38i3.2741.

[3] B. W. Schuller, R. Picard, E. André, J. Gratch, and J. Tao, "Intelligent signal processing for affective computing [From the Guest Editors]," *IEEE Signal Process. Mag.*, vol. 38, no. 6, pp. 9–11, Nov. 2021, doi: 10.1109/MSP.2021.3096415.

[4] A. Kaushal, R. Altman, and C. Langlotz. "Health care AI systems are biased." Scientific American, Nov. 17, 2020. https://www.scientificamerican.com/article/health-care-ai-systems-are-biased/

[5] F. Nebeker, "The IEEE signal processing society, fifty years of service – From 1948 to 1998," IEEE History Center, New Brunswick, NJ, USA, 1998. [Online]. Available: https://ethw.org/w/images/9/9b/Signal_processing_society_history.pdf
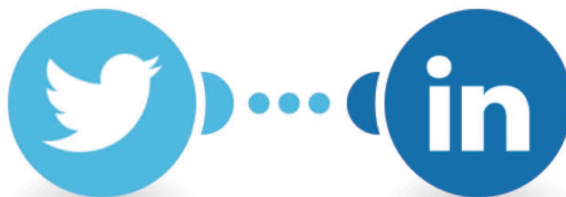
Twitter

LinkedIn

Interested in learning about upcoming SPM issues or open calls for papers?
Want to know what the SPM community is up to?
Follow us on twitter (@IEEEspm) and/or join our LinkedIn group
(www.linkedin.com/groups/8277416/) to stay in the know and share your ideas!