Paris Session 2022

Restrictions on applying “black box” machine learning-based software

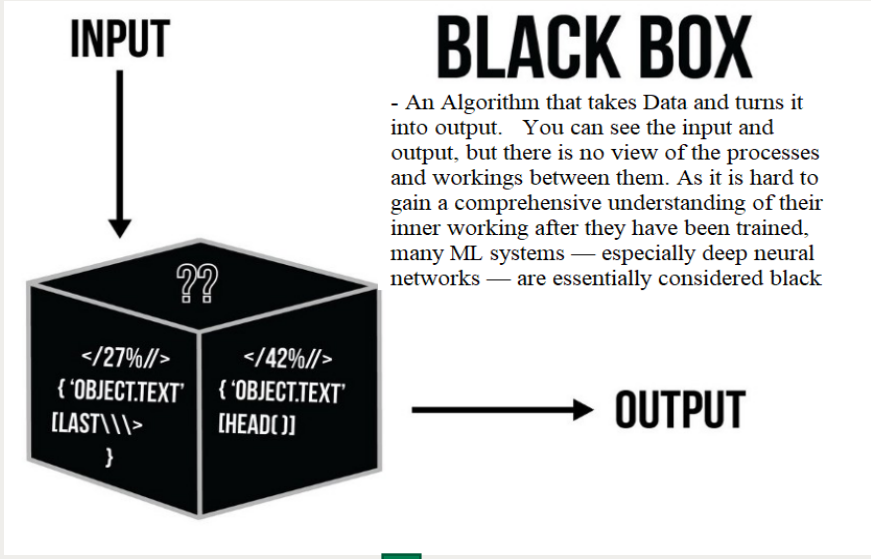
D2, PS 1 – Question 1.2

“Are there any restrictions on applying “black box” machine learning-based software for critical infrastructures like power industry? Is the application area of such software limited by decision supporting systems or some decision-taking systems do exist?”

Sarala M Naidu, Sweden

Restrictions on applying “black box” machine learning-based software

- Although the potential of ML in power industry has been recognized, obstacles to further development of ML models remains, of which the black-box nature of the most ML algorithms is an important factor. Thus, the restriction for applying black-box models are as follows:
  - The power grid is **highly regulated** and a **critical infrastructure area**, so machine learning models involving it must be **interpretable** in order to **increase user trust** and improve **system reliability**
  - Then, from and ML point of view :
    - with increase in the model complexity, the interpretability of ML algorithms has deteriorated
    - when ML is used in critical Decision -making process (like outages), professionals are hesitant to deploy such models, because the models error will induce a very large impact
  - Thus, the experts in the loop of decision making need to know where the model is failing or it needs to be tuned, and for doing that, they need to understand the black-box-based model in depth



from a mathematical point of view, they are very hard to explain & to be understood by experts

examples

- ANN-based models (CNN & GANs families)
- Hyperplanes, e.g. SVM
- probabilistic and combinatory logic, e.g. Markov networks, Bayesian networks
- based on instance or the lazy learnings e.g. kNN

Is the application area of such software limited by decision supporting systems or some decision-taking systems do exist?

The complexity may vary on a continuum:

- Simple situations are usually predictable and simple
- Complicated circumstances require more expertise in problem-solving or analysis to determine the cause and effect.
- Complex situations encompass multiple relationships and interdependence. An effective analysis may therefore require a holistic or systemic approach. Besides, simulations are necessary to understand how decision making in business intelligence affects widespread elements.

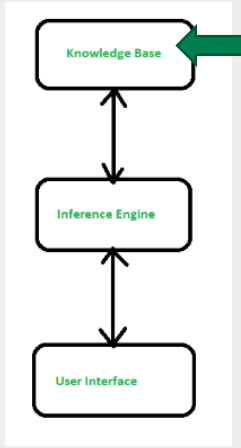
Algorithmic decision-making has **the potential to eliminate, introduce or amplify biases or discrimination**. However, this depends on how the software is deployed, and the quality and representativeness of the underlying data used by the algorithm.

Many advanced power quality monitoring systems are equipped with either off-line or on-line intelligent systems to evaluate disturbances and system conditions to make conclusions about the cause of the problem or even predict problems before they occur.

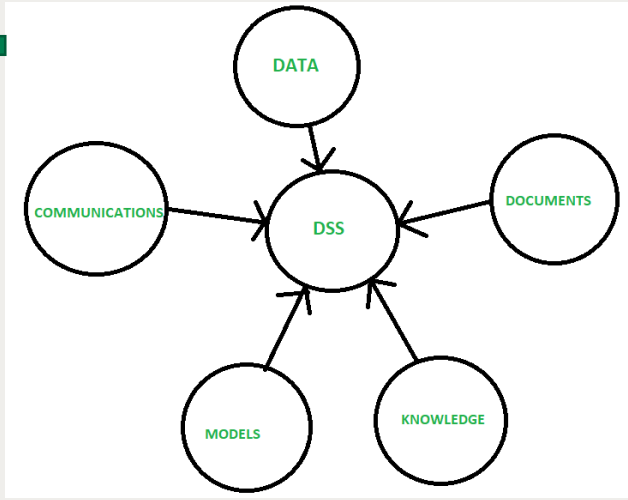
rule-based, frame-based, fuzzy systems, neural expert systems, and neuro-fuzzy systems

examples

Decision Making Expert system



Decision Support



Application area	per cent
operation	45
planning	18
diagnosis	37

The Power grid is a high regulated, complex interconnected and a critical infrastructure area and any model error will induce very large impacts. Thus, there is a need for an in-depth interpretability of the black-box-based models to increase the user trust for decision making process. All these restriction of application is entirely based on the complexity and the context of the interconnections & interdependencies of the subsystems.