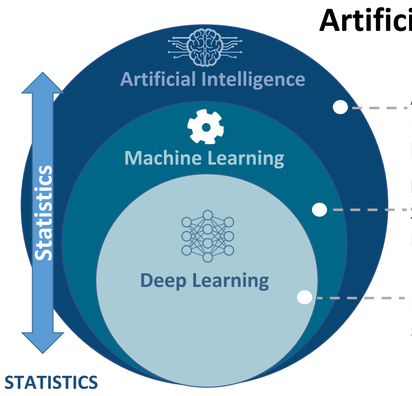
Artificial intelligence and machine learning: just a hype or a new opportunity for pharmacometrics?

R. Bartolucci, S. Grandoni, N. Melillo, G. Nicora, E. Sauta, E. M. Tosca, P. Magni

Department of Electrical, Computer and Biomedical Engineering, University of Pavia, via Ferrata 5, Pavia, I-27100, Italy

BACKGROUND:

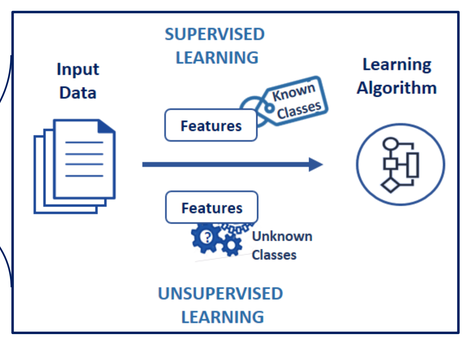
Artificial Intelligence, Machine Learning & Statistics definitions



ARTIFICIAL INTELLIGENCE
Intelligent system ("agent") able to perform human-like tasks or solve complex problems

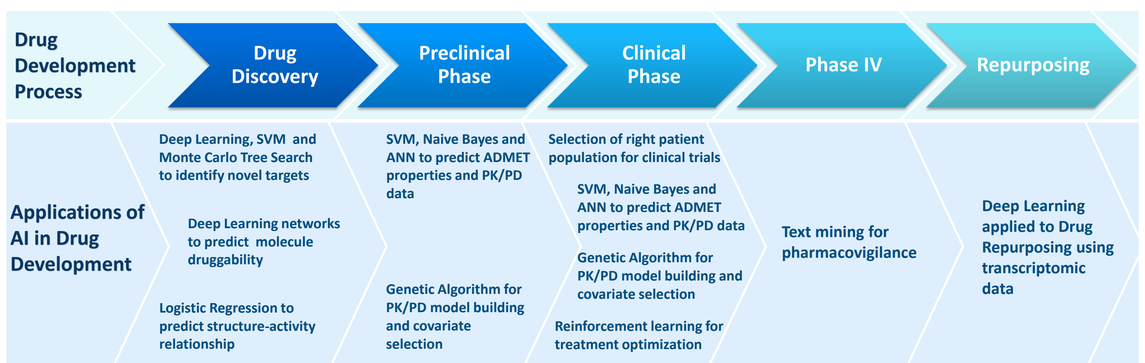
MACHINE LEARNING
Subfield of AI which allows machines to learn and make predictions based on their experience (data)

DEEP LEARNING
ML method which exploits neural architecture for self-learning patterns from a huge data amount



STATISTICS
Statistical data modeling can be exploited and integrated by ML techniques

AI/ML in drug discovery and development

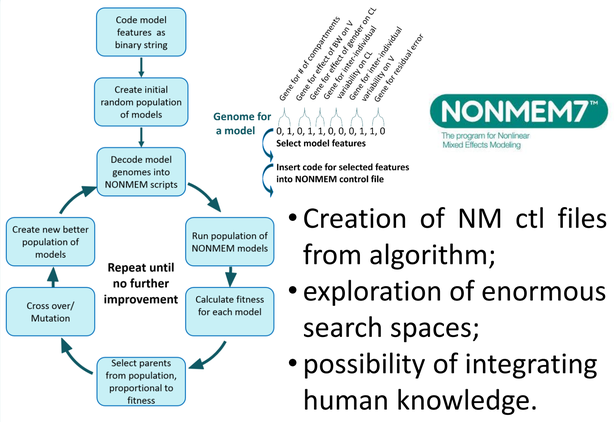


OBJECTIVE: Providing an overview of the AI/ML applications in pharmacometrics, with the aim of understanding how AI/ML can support, substitute or be integrated with model-based approaches and trying to clarify their effective role in this field.

METHODOLOGY: From a literature review, four main pharmacometric tasks, that have been approached by AI/ML, were identified:

Model building and covariate selection

Genetic Algorithm (GA) approaches:



- Creation of NM ctl files from algorithm;
- exploration of enormous search spaces;
- possibility of integrating human knowledge.

Stochastic Approximation for Model Building Algorithm (SAMBA):

Rsmlx R speaks Monolix

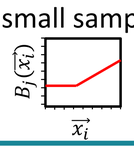
Monolix

Automatic iterative procedure identifying at each step how best to improve some of the model features (residual error, RE, covariates, correlations) with the aim of finding the optimal statistical model which minimizes some information criterion.

Multivariate Adaptive Regression Splines (MARS):

Non-parametric regression technique allowing to identify agnostic PK covariate and their nonlinear interactions using small sample sizes.

$$y_i = \sum_{j=1}^{n} c_j B_j(\vec{x_i})$$



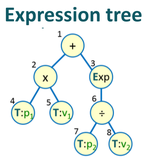
Gene Expression Programming GA:

GA with a special gene-encoding scheme that allows to model non-polynomial covariate relationships.

Mathematical equation

$p_1 \times v_1 + \exp\{p_2 \div v_2\}$

Expression tree



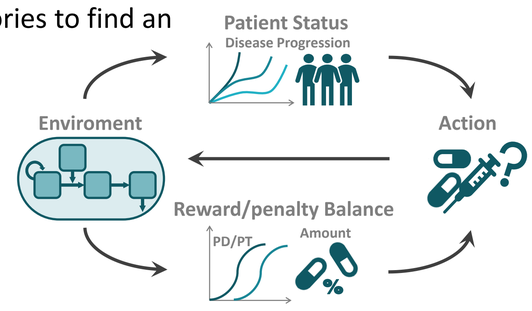
Gene code

1	2	3	4	5	6	7	8
+	x	E	T	T	÷	T	T
p_1	v_1	p_2	v_2	p	p	p	p

- ✓ Automation and optimization only of the actual search part
- ✗ No inference on causality or biological mechanisms
- ✓/✗ Need of subjective evaluation
- ✓/✗ Lower human efforts, higher computation load
- ✓ Generation of understanding if non-prespecified solutions are further explored as primary hypothesis

Therapy Optimization

An iterative process (e.g. Reinforcement Learning techniques, Monte Carlo Tree Search algorithm) based on optimal control and dynamic adaptation theories to find an optimal treatment strategy over a period of time.

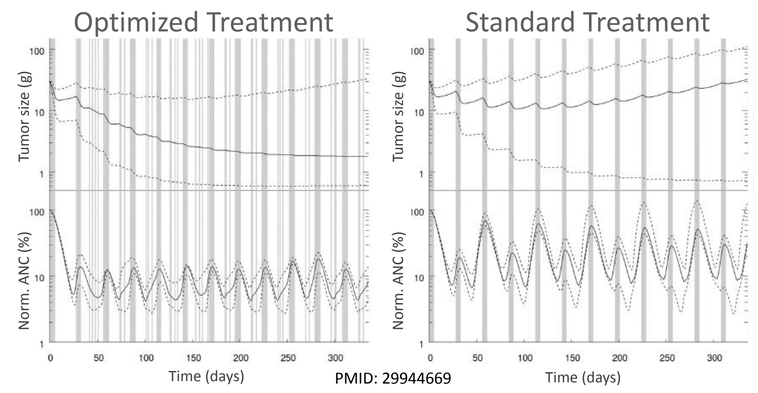


Algorithm elements:

- Set of patient status (e.g. tumor size)
- Set of possible actions (e.g. dose percentage)
- An enviroment (e.g. model simuations)
- A reward/penalty function (e.g. to balance efficacy and toxicity)

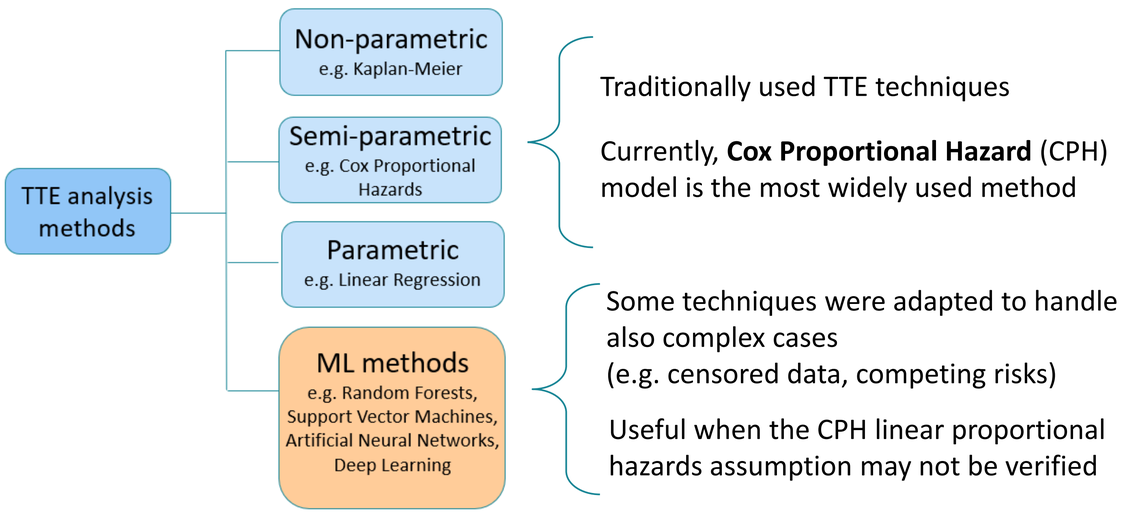
Process:

- At each time step, the patient status is evaluated to choose the right dose;
- the dose is chosen from a list of possibilities;
- a model simulation is run over the entire time frame to obtain the patient status;
- other factors can be taken into account for the dose selection.



- ✓ The algorithm learns the relationship between patient status and action
- ✓ Both drug efficacy and toxicity can be taken into account
- ✓ Treatment can be optimized for a group or for each patient
- ✗ Addition of other parameters that require tuning (e.g. penalty/reward balance)
- ✓ Possible applications in oncology and immunotherapy case studies
- ✓/✗ A new strategy, currently not yet applied and evaluated in the clinical practice

Time-to-Event (TTE) analysis



Traditionally used TTE techniques

Currently, **Cox Proportional Hazard** (CPH) model is the most widely used method

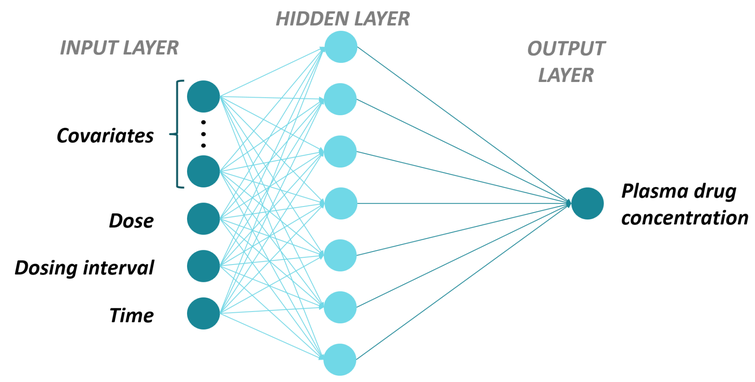
Some techniques were adapted to handle also complex cases (e.g. censored data, competing risks)

Useful when the CPH linear proportional hazards assumption may not be verified

- ✓ Suitable to discover non-linear relationships between covariates and risk difficult to establish without prior deep knowledge
- ✓ Identification of influential variables on high-dimensional data
- ✓ Handling of high dimensional data (even when variables are more than observations);
- ✓/✗ Black box-like approaches;
- ✗ High complexity of the underlying model architecture;
- ✗ Success of implementation dependent on the fine tuning of parameters.

PK/PD

Artificial neural networks (ANN) were used to describe the PK/PD of drugs. To perform this task, ANN can be trained on dataset containing, for example, *covariates*, *dose*, *dosing interval* and *time of blood drawn* as input variables and *plasma drug concentration* as output variable. Performances comparable to those obtained by using classical NLME models with NONMEM are reported.



- ✓ Reportedly more easy to use than NLME models
- ✗ Not a dynamical system: its behaviour (e.g. stability) is not easily predictable and thus, the extrapolation would be difficult
- ✗ Difficulty to understand which are the most important covariates in determining the model output (plasma drug concentration)

CONCLUSIONS: AI and ML methodologies are not novel approaches and are well established in other fields different than pharmacometrics. Recently, their application to the drug development process has raised a relevant interest showing their usefulness for the improvement of drug discovery phase and patient selection.

From our analysis, it emerges that attempts to adopt AI/ML approaches in the model building and covariate selection, prediction of PK/PD endpoints, identification of optimal treatment strategy and TTE analysis are widely documented in literature. Results suggest that these techniques could be an efficient and valid alternative to the standard forward-addition/ backward-elimination methods in the automation of the model building process, even if a user evaluation of the selection steps remains essential.

They are, also, extremely promising strategies for therapy optimization, especially in the perspective of personalized medicine. Conversely, in other circumstances, they cannot substitute a model-based strategy and, due to their data-driven approach, they seem to be in contrast to the current pharmacometrician efforts towards the building of more mechanistic models.

In summary, AI/ML methods could be successfully exploited in pharmacometrics and their capabilities should be assessed for further tasks, such as model-based meta-analysis, but only after a careful and critical evaluation of the investigated problem characteristics and an assessment of their real applicability.

REFERENCE: see the on line abstract.