Predicting pathological risks and dynamics: innovative approaches based on mechanistic modelling augmented by artificial intelligence

MICHELA SPERTI, XXXVI Cycle Department of Mechanical and Aerospace Engineering, Polito[BIO]Med Lab, Polytechnic University of Turin, Turin, Italy

TRAINING ACTIVITIES

HARD SKILLS (120 hours)

Computational hard skills

Courses on **I) Data mining concepts and algorithms** (20 hours); **II)** deep learning focusing on **Adversarial training of neural networks** (15 hours); **III)** statistical sampling approaches such as **The Monte Carlo method** (30 hours); **IV)** GPU and HPC programming **provided by NVIDIA** on **Compute Technical Curriculum** (8 hours) and **Virtualization Technical Curriculum** (4 hours).

Biomedical hard skills

Course on **V) Pianificazione, gestione e analisi di ricerca clinica e di laboratorio** (15 hours) focusing on experimental activities; webinar on **(VI) Nanoscience in cancer immunotherapy** (10 hours) with emphasis on emerging technologies in cancer immunology; Training event MSCA-ITN project PARENT on **(VII) Neurodevelopment after preterm birth: diagnosis, ethics and social impact** (18 hours).

SOFT SKILLS (42 hours)

I) Communication (5 hours); **II) Project management** (5 hours); **III) Public speaking** (5 hours); **IV) Thinking out of the box** (1 hour); **V) Time management** (2 hours); **VI) Writing Scientific Papers in English** (15 hours), **VII) Navigating the hiring process: CV, tests, interview** (2 hours); **VIII) Responsible research and innovation, the impact on social challenges** (5 hours); **IX) IEEE Italy Authorship Symposium** (2 hours).

PERIODS ABROAD

03/05/2022-26/08/2022: **University of Alberta (Edmonton, Canada)**
Research topic: AI toward novel drug cardiotoxicity prediction tools.

UNIVERSITY OF ALBERTA

RESEARCH ACTIVITIES

Research context

- **Enormous quantity of data about medical conditions** available.
- **AI-based algorithms** proven successful in analyzing big data sources and unravel complex hidden relationships.

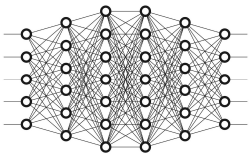
Clinical decision support systems can help physicians in their complex decision-making processes.

Challenge

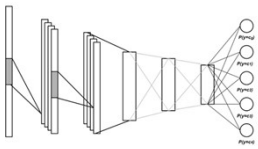
Early risk prediction in the clinical domain, with specific focus on cardiovascular diseases.

- **Improving patient stratification** according to risk.
- Identifying principal pathology variables characterizing the patient (**personalized therapeutics**).
- **Analyzing the risk evolution over time**.

Computational methodologies and tools



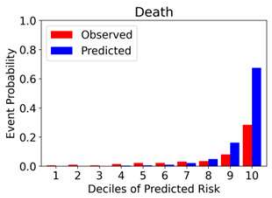
python · GitHub · jupyter · colab · learn · NumPy · pandas · matplotlib · TensorFlow · Keras · SHAP



Supervised Machine Learning (e.g., Random Forest, Support Vector Machine, Naïve Bayes, etc.) **and Deep Learning** (Feed-Forward Neural Networks, Convolutional Neural Networks, Autoencoders, etc.) **for classification and regression tasks**. **Traditional statistics** (e.g., Cox-proportional hazard regression) methods as benchmark for comparison in the medical domain.

Research Topics

***(I)** Cardiovascular **risk prediction and stratification** in the **general** population…*

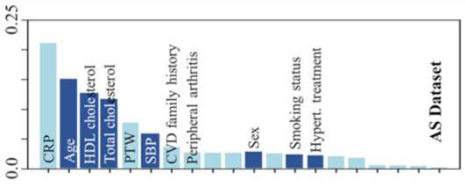


- **ML-based risk stratification model** (named the **PRAISE model**) to predict all-cause death, recurrent acute myocardial infarction, and major bleeding after ACS.
- **AUC = 0.92 (0.90-0.93)** in the external validation cohort for 1-year all-cause death.
- **Tool available for physicians' use**.

SCAN ME

DOI: 10.1016/S0140-6736(20)32519-8

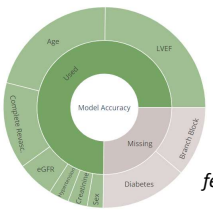
*…and in patients with **comorbidities (II)***



- ML-based cardiovascular risk prediction model for patients affected by **inflammatory arthritis**.
- Study of the **main variables (global perspective) affecting the predictions.**

DOIs: 10.1093/rheumatology/kez677, 10.6084/m9.figshare.12887285

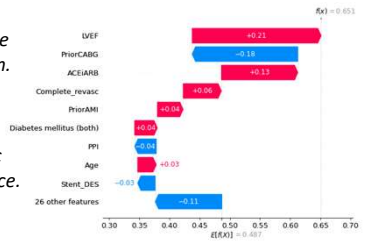
***(III-A) Explainability** of the cardiovascular risk prediction*



Reliability of the given prediction.

+

Patient-specific feature importance.

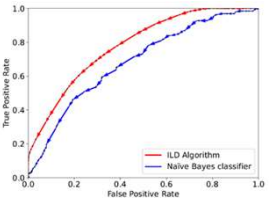


- **Explainable AI technique (SHAP)**.
- For a clinician, this is far more useful than simply having a risk prediction in a **black-box** way.
- This has been implemented to augment the usability of the **PRAISE model** in the clinical domain.

(III-B)** Determination of the **intrinsic limit of a dataset

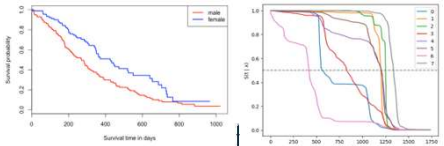
- **ILD Algorithm**, model-agnostic.
- Novel technique to determine the best possible performance (**AUC and accuracy**), achievable from a specific dataset.
- Applications in binary classification problems with categorical features.

DOIs: 10.3390/a14110301



(IV)** Individual survival analysis to study the **risk evolution over time

- Predicting the **time to an event**.
- **Kaplan-Meier** vs **Individual Survival Curves**.



Publications

9 articles, 3 conference papers (h-index 5)
SCOPUS ID: 57217331713

[1] D'Ascenzo et al., The Lancet (2021) **[2]** Halasz et al., Journal of Medical Internet Research (2021) **[3]** Michelucci et al., Algorithms (2021) **[4]** Sperti et al., Minerva Cardiology and Angiology (2022) **[5]** Sperti et al., Optical Sensing and Detection VII SPIE (2022) **…**

Tutors: Prof. Umberto Morbiducci, Prof. Marco A. Deriu, Dr. Umberto Michelucci

PhD Day 2022
25th October 2022, Dental School, Torino