Looking for Interpretable Artificial Intelligence Methods in Health Research

Encontro Ciência 2019

July 9, 2019

Carlos Geraldes

carlos.geraldes@nms.unl.pt

Ana Papoila

ana.papoila@nms.unl.pt

Centro de Estatística e Aplicações
Universidade de Lisboa

Outline

Centro de Estatística e Aplicações
Universidade de Lisboa

Introduction

Predictive Models

- Neural Networks
- Statistical models

GANN Models

- Model definition
- Partial Function
- Real case
- Partial Function Plot
- Odds Ratio Function Plot

FCT Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

Centro de Estatística e Aplicações
Universidade de Lisboa

Health Intelligence

- A major driver of initiatives such as precision medicine that uses artificial intelligence, data science tools, and other methods aiming to improve productivity in health services and clinical care.

FCT
Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

Centro de Estatística e Aplicações
Universidade de Lisboa

Health Intelligence

- A major driver of initiatives such as precision medicine that uses artificial intelligence, data science tools, and other methods aiming to improve productivity in health services and clinical care.
  - Disease management.

FCT Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

Introduction

Centro de Estatística e Aplicações
Universidade de Lisboa

Health Intelligence

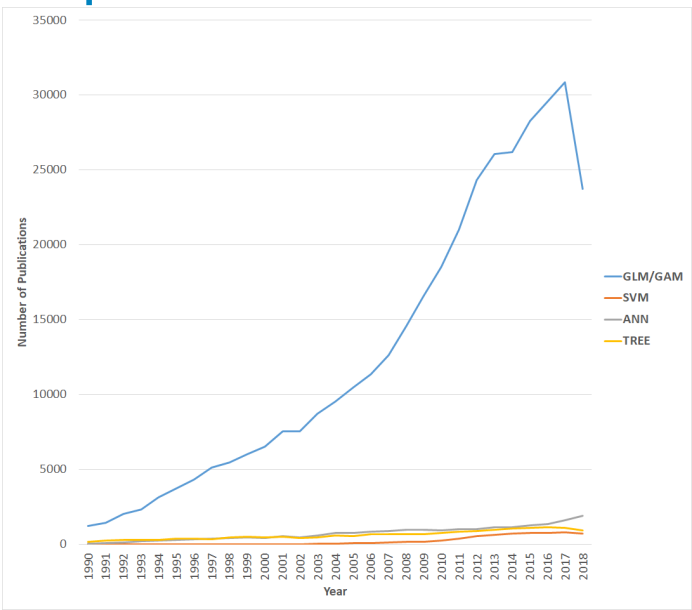
- A major driver of initiatives such as precision medicine that uses artificial intelligence, data science tools, and other methods aiming to improve productivity in health services and clinical care.
  - Disease management.
  - Personalized therapies or interventions.

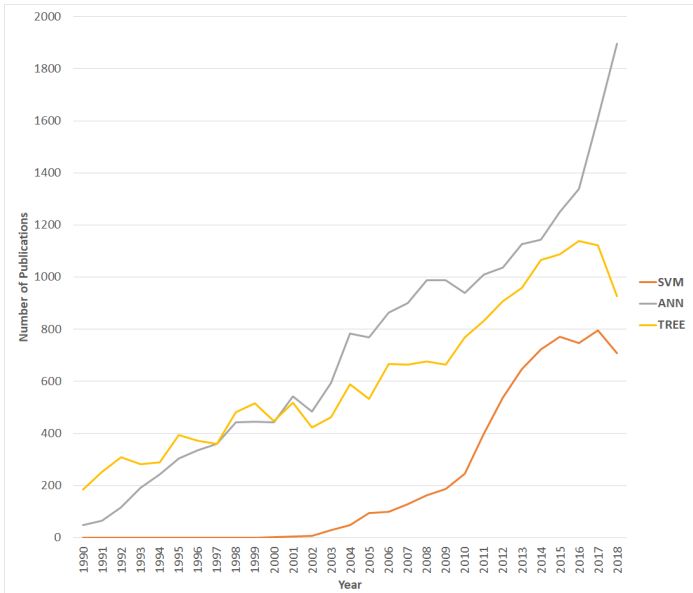
FCT Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

Predictive Models

Centro de Estatística e Aplicações
Universidade de Lisboa

Supervised Models





Number of publications in PubMed per year

FCT
Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

Centro de Estatística e Aplicações
Universidade de Lisboa

Accountability

- Dependence on AI-based dynamic systems demands clearer accountability to ensure trust and transparency in decisions (eXplainable Artificial Intelligence).

FCT Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

Predictive Models

Centro de Estatística e Aplicações
Universidade de Lisboa

Accountability

- Dependence on AI-based dynamic systems demands clearer accountability to ensure trust and transparency in decisions (eXplainable Artificial Intelligence).
- In Biomedical applications, model interpretability is mandatory.

FCT Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

Predictive Models

Centro de Estatística e Aplicações
Universidade de Lisboa

Interpretability

- Interpretability is the ability to explain or to provide the meaning in understandable terms to a human.

FCT
Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

Centro de Estatística e Aplicações
Universidade de Lisboa

Interpretability

- Interpretability is the ability to explain or to provide the meaning in understandable terms to a human.
  - Global Interpretability - A model allows humans to understand the whole logic of a model and follow the entire reasoning leading to all the different possible outcomes.

FCT Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

Centro de Estatística e Aplicações
Universidade de Lisboa

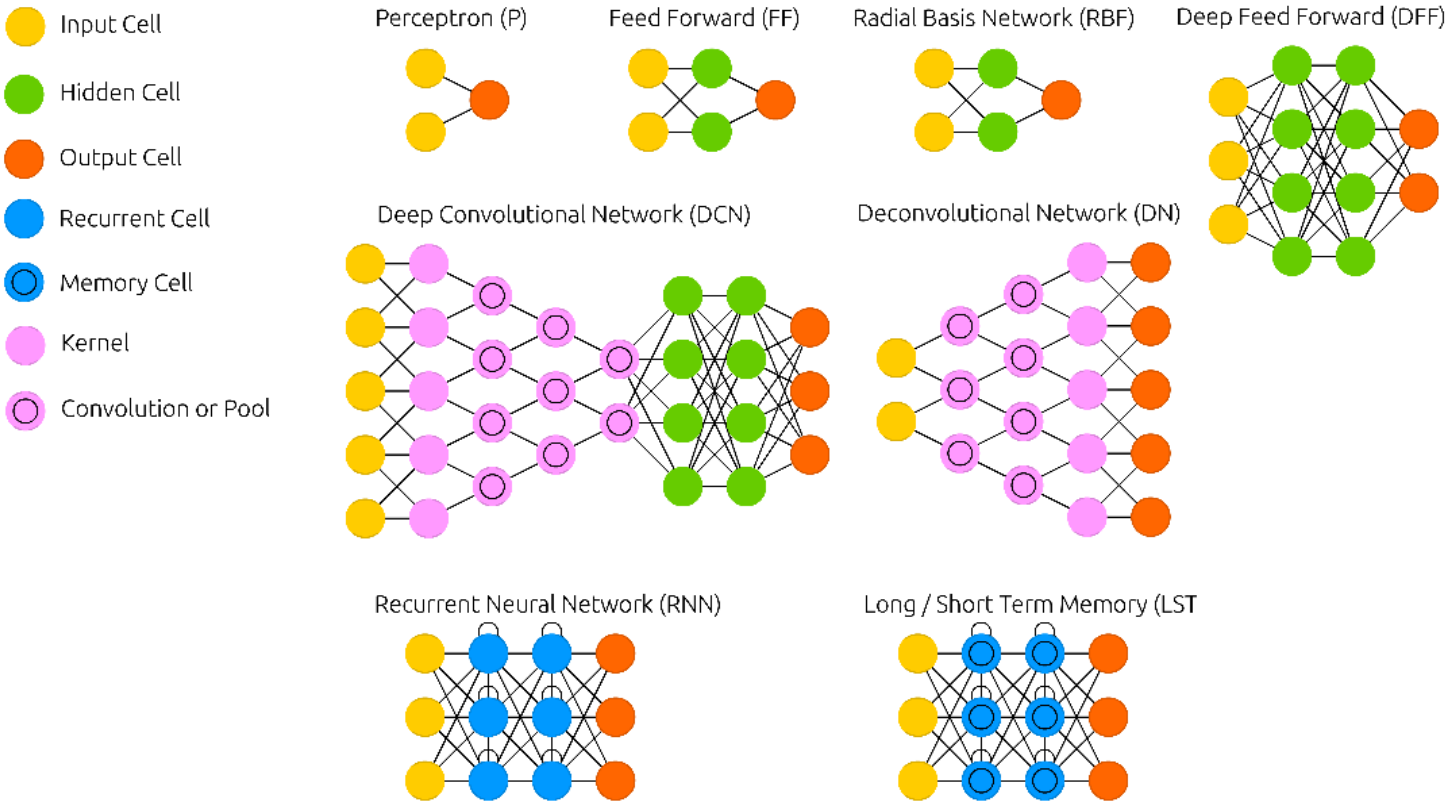
Interpretability

- Interpretability is the ability to explain or to provide the meaning in understandable terms to a human.
  - Global Interpretability - A model allows humans to understand the whole logic of a model and follow the entire reasoning leading to all the different possible outcomes.
  - Local Interpretability - Situation in which it is possible to understand only the reasons for a specific decision (only the single prediction/decision is interpretable).

FCT Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

Centro de Estatística e Aplicações
Universidade de Lisboa

Neural Networks



GLM/GAM

Centro de Estatística e Aplicações
Universidade de Lisboa

GLM

Linear function

$$E(Y|X) = h(\beta_0 + \beta_1(X_1) + \beta_2(X_2) + \ldots + \beta_p(X_p))$$

Link function

GAM

$$E(Y|X) = h(\beta_0 + f_1(X_1) + f_2(X_2) + \ldots + f_p(X_p))$$

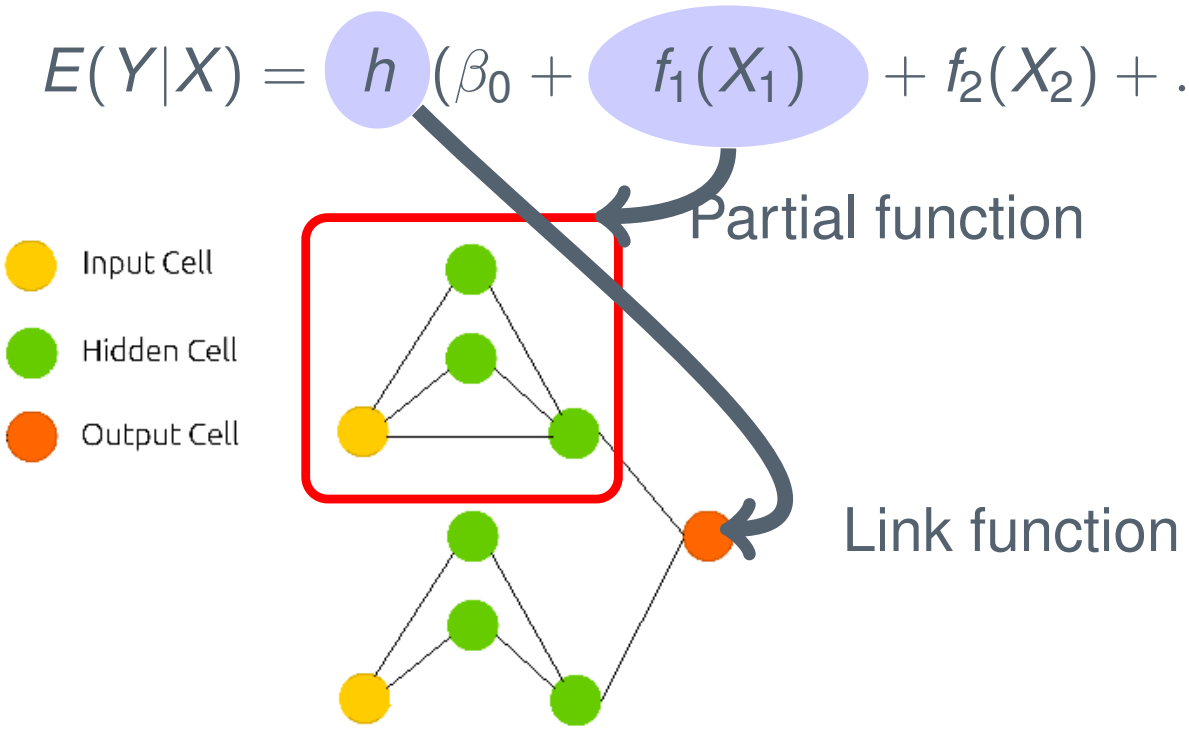
Partial (smooth) function

FCT Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

Centro de Estatística e Aplicações
Universidade de Lisboa

Model definition

$$E(Y|X) = h(\beta_0 + f_1(X_1) + f_2(X_2) + \ldots + f_p(X_p))$$



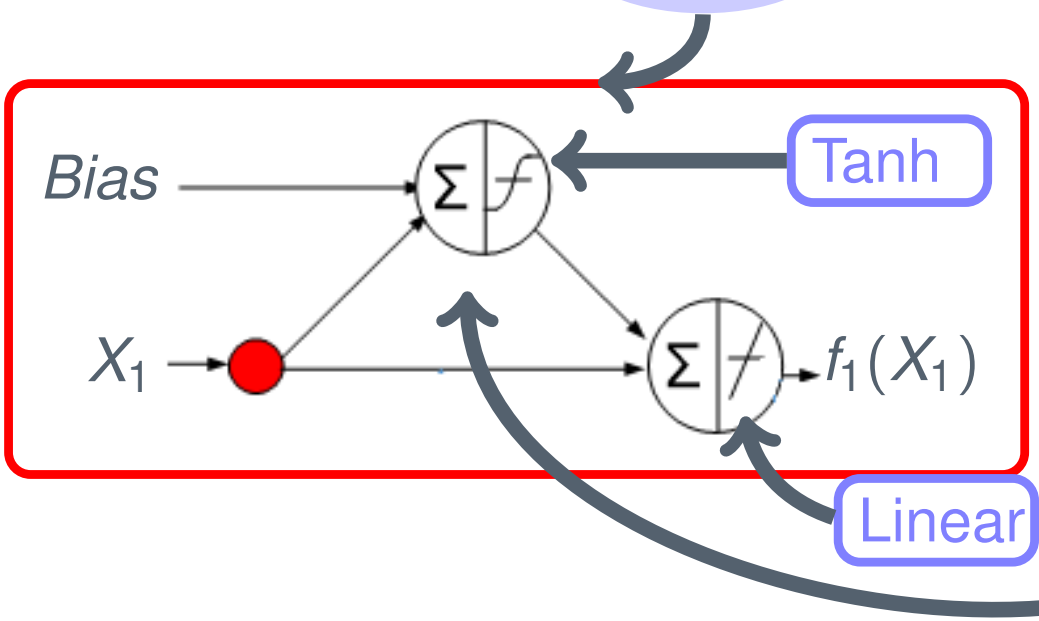
FCT Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

Generalized Additive Neural Networks

CEAUL
Centro de Estatística e Aplicações
Universidade de Lisboa

Partial Function

$$E(Y|X) = h(\beta_0 + f_1(X_1) + f_2(X_2) + ... + f_p(X_p))$$



Function flexibility depends on # hidden nodes

$$f_1(X_1) = \omega_{ac} X_1 + \omega_{bc} tanh(Bias + \omega_{ab} X_1)$$

FCT
Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

Centro de Estatística e Aplicações
Universidade de Lisboa

Real case

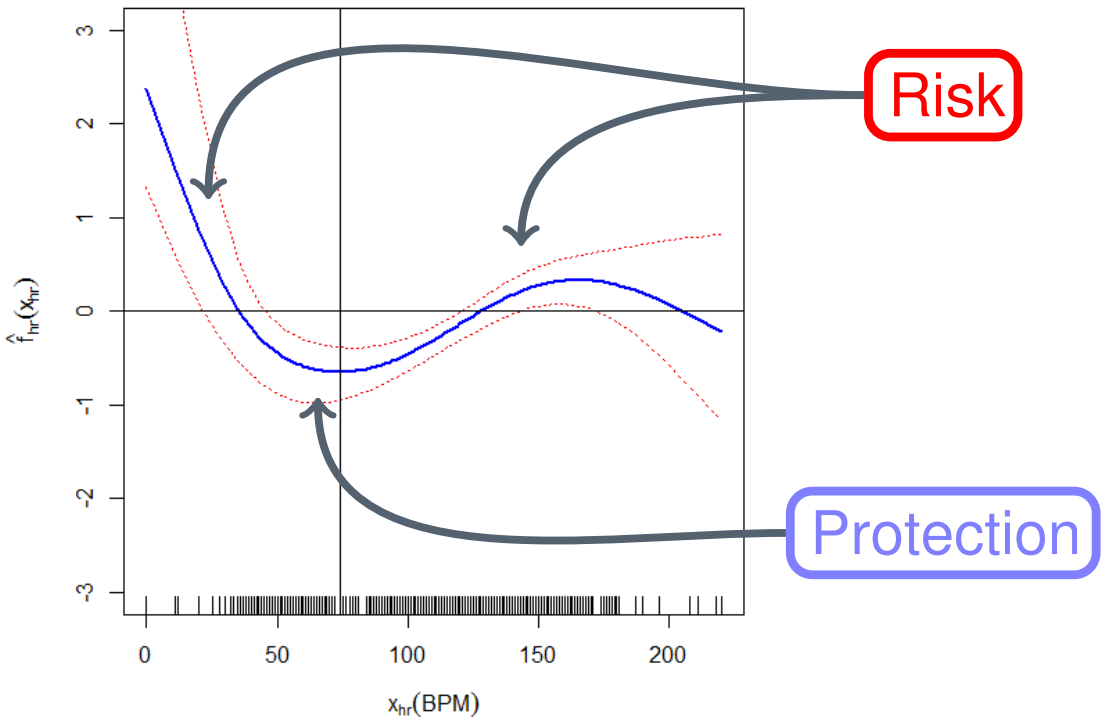
- Sample collected from an Intensive Care Unit (ICU) regarding 996 patients.
- The response variable Y refers to the death during hospitalization in the ICU. Of the 996 patients, 359 died.
- A multivariable model was built in which one of the studied variables was Heart Rate (BPM).

FCT
Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

Generalized Additive Neural Networks

Centro de Estatística e Aplicações
Universidade de Lisboa

Partial Function Plot



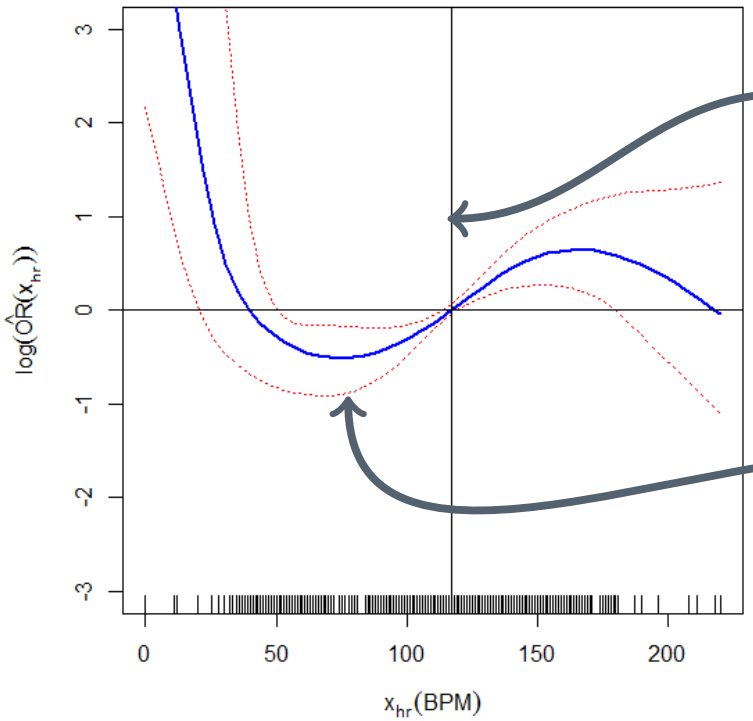
Brás-Geraldes, C., Papoila, A. & Xufre, P. (2019). Generalized additive neural network with flexible parametric link function: model estimation using simulated and real clinical data. Neural Computing & Applications. DOI 10.1007/s00521-017-3105-6

FCT Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

Generalized Additive Neural Networks

Centro de Estatística e Aplicações
Universidade de Lisboa

Odds Ratio Function Plot



Reference

Considering the reference value (117 BPM), the odds of death at $x_{hr} = 80\ BPM$ is $0.61[0.45, 0.81]$

Brás-Geraldes, C., Papoila, A. & Xufre, P. (2019). Odds ratio function estimation using a generalized additive neural network. Neural Computing & Applications. DOI 10.1007/s00521-019-04189-7

FCT Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

Thank You!