Optimizing the search for underground water on Mars with explainable artificial intelligence

Marco Lippi

`marco.lippi@unifi.it`

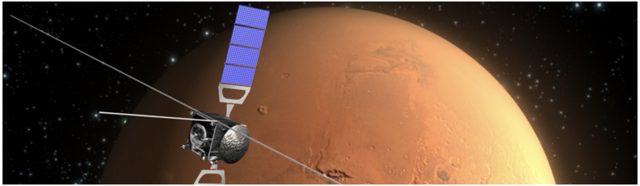
Department of Information Engineering

UNIVERSITÀ DEGLI STUDI FIRENZE

Table of Contents

Mars Express

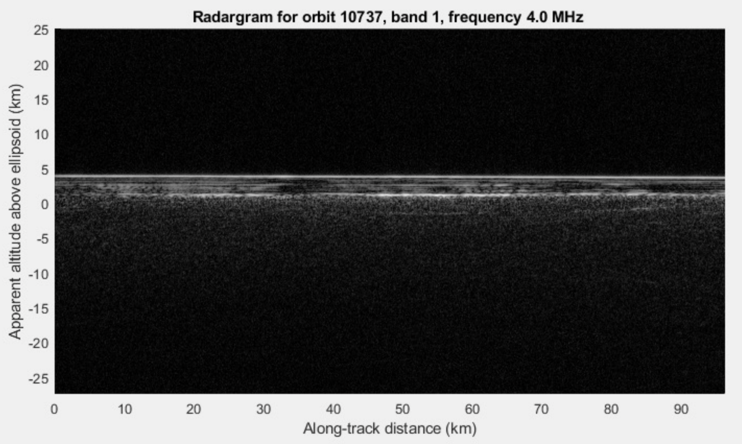
ESA Mission launched in 2003 to study ionosphere and analyze subsurface of Mars



R. Orosei et al., Radar evidence of subglacial liquid water on Mars, *Science*, 2018

Radargrams from MARSIS

1 Optimizing the search for underground water on Mars with XAI



Radargrams from MARSIS

How are radargrams collected from MARSIS?

- MARSIS activities are planned **manually** on a monthly basis
- The available resources are very scarce: only 134 seconds of observations per day
- Over 4 million samples collected on the South Pole since 2005

Can we improve over manual scheduling? How can machine learning help?

MARS Observation Scheduling Problem

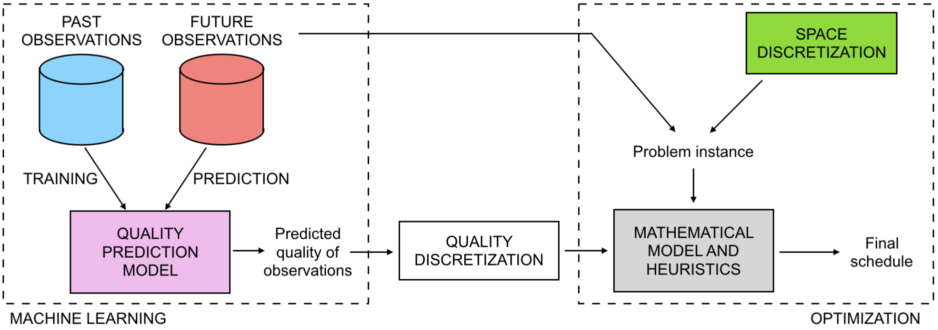
We are given a set O of observation opportunities, each belonging to an orbit on a specific day of a set D; each observation covers a specific area of the South Pole of Mars and lasts for L seconds; for each day, we can choose at most N observations

We aim at **maximizing the area covered by the observations** and, more specifically, the overall **quality** of the collected data, which is unknown a priori

We propose an Integer Linear Programming (ILP) formulation and two greedy heuristics and a fix-and-optimize matheuristic as comparative solutions

Predict-and-optimize

1 Optimizing the search for underground water on Mars with XAI



B. Ferrari, M. Delorme, M Iori, M. Lippi, R. Orosei, *submitted for publication*, 2025

Predict-and-optimize

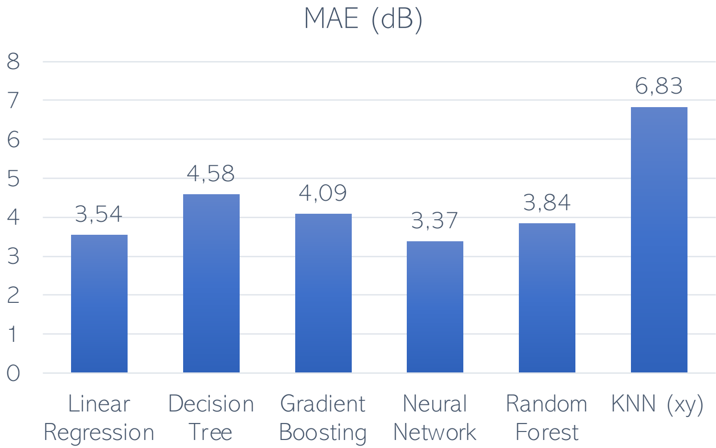
Machine learning algorithms should predict the **quality of the signal** as a function of some **variables** describing observation conditions, using a data set of 2 millions examples

The most important features include the following

- Sun elevation angle, Mars–Sun distance
- Latitude, longitude, roughness
- Magnetic field of Mars
- Altitude of radar, orientation of antenna

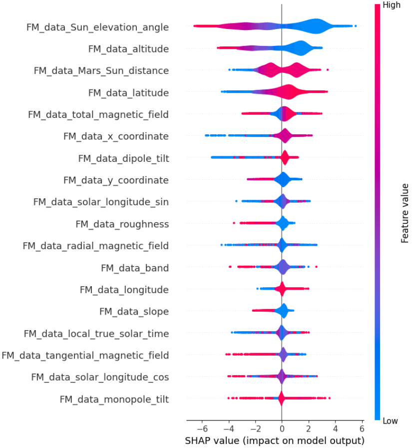
Experimental Results

1 Optimizing the search for underground water on Mars with XAI



Explainability: SHAP values

1 Optimizing the search for underground water on Mars with XAI



Interpretability: EBM

Astrophysicists are very much interested in understanding the impact of different variables on signal quality: there is need either to **explain the predictions** of black-box models, or to exploit **interpretable-by-design** approaches

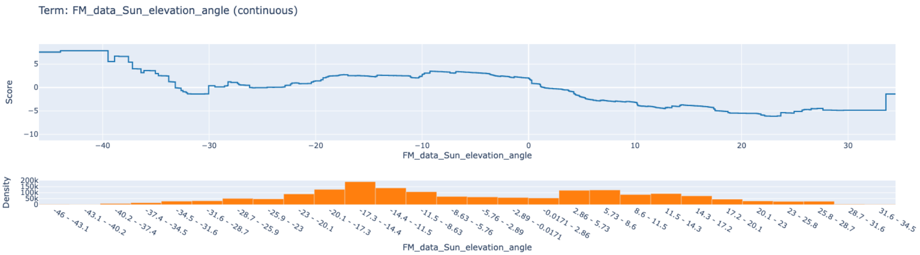
Now investigating with Explainable Boosting Machines [Lou et al., 2016]

$$y = \sum_{j=1}^{p} f_j(x_j) + \sum_{h=1}^{p} \sum_{k=h+1}^{p} f_{hk}(x_h, x_k)$$

being p the number of features (i.e., independent variables) and y the target of prediction

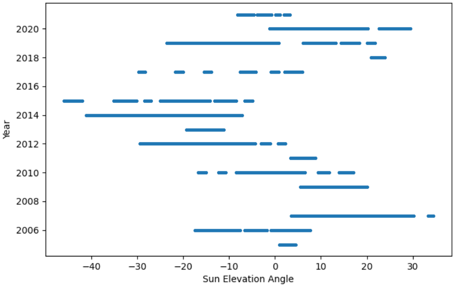
The impact of solar activity

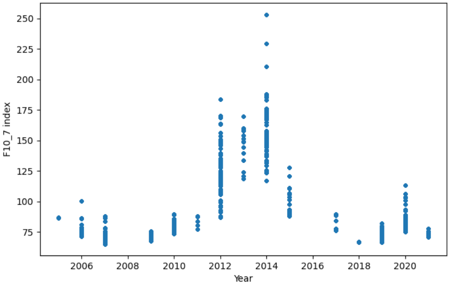
Solar activity has a strong impact on the quality of the signal, as it has some effect on the polarization of the ionosphere surrounding Mars



The impact of solar activity

Bias in data: observations with low Sun elevation angle (i.e., by night) were mostly performed during the peak of Solar cycle! Need to model differently ionosphere variables





Next steps

- Adding more variables to the problem
- Use symbolic regression (learn mathematical formulas)
- Directly analyze spectrograms (i.e., images) maybe with prototypes?

Table of Contents

2 Understanding conditions for harmful algal bloom

Copernicus data observations

Copernicus is the European Union’s Earth observation programme, looking at our planet and its environment to benefit all European citizens.

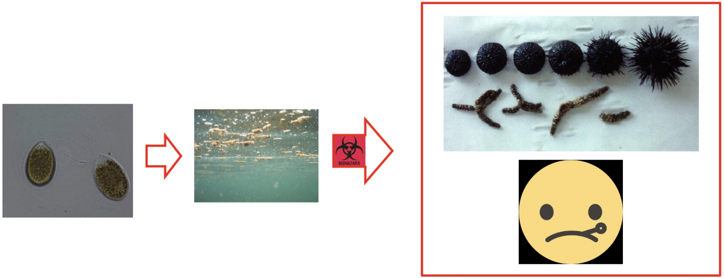
Material shared by Claudia Lapucci, CNR

Ostreopsis ovata

2 Understanding conditions for harmful algal bloom

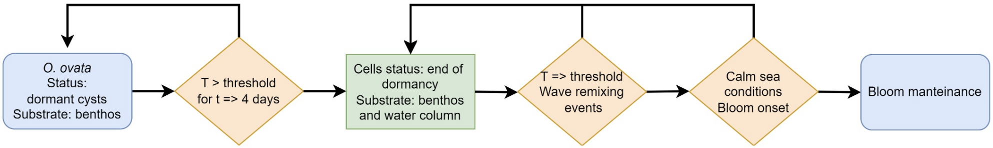
O. ovata harmful algal bloom has been detected in Italian coastal areas in the last years

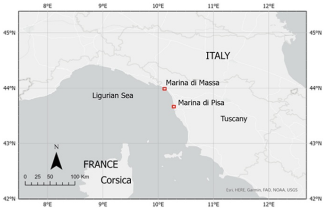
Blooms are sudden events and have toxic effects on marine organisms and humans



Ostreopsis ovata

2 Understanding conditions for harmful algal bloom





J. Mar. Sci. Eng. 2022, 10, 461

The role of (explainable) machine learning

2 Understanding conditions for harmful algal bloom

Is it possible to use XAI to **anticipate** the prediction of harmful algal bloom, so that policies can be adopted to mitigate the harm for the environment?

Random forests have been used in the past for similar tasks (i.e., compute the TRIX index for water quality and monitoring)

In this case, we aim to study the physical phenomenon, so it would be interesting to obtain an **interpretable** model: which variables are responsible?

Table of Contents

Safety-critical systems

The role of AI in Safety Engineering is that of **supporting safety-critical decisions**, in situations where **errors** can produce catastrophic and fatal consequences

Therefore, there is a growing need for AI systems that are...

- reliable
- trustworthy
- interpretable
- interactive with humans
- able to exploit domain knowledge

[Image from telematicswire.net]

Cyber-physical systems

Cyber-physical systems offer an **ideal scenario** for NeSy AI

- Difficulty of **designing a model** for the system (typically provided by a human)
- Traditional **verification** systems typically too complex for large-scale domains
- **Data-driven** approaches are becoming crucial for many subtasks

XAI for safety-critical systems

- **Model learning**
  Learn the structure of a model directly from data (e.g., system traces, logs)
- **Formal verification**
  Use neural networks to handle uncertainty in formal verification
- **Program synthesis**
  Synthesize programs from formal specifications (Large Language Models?)
- **Testing**
  Model flow graphs; maximize probability to execute selected test cases

XAI for safety-critical systems

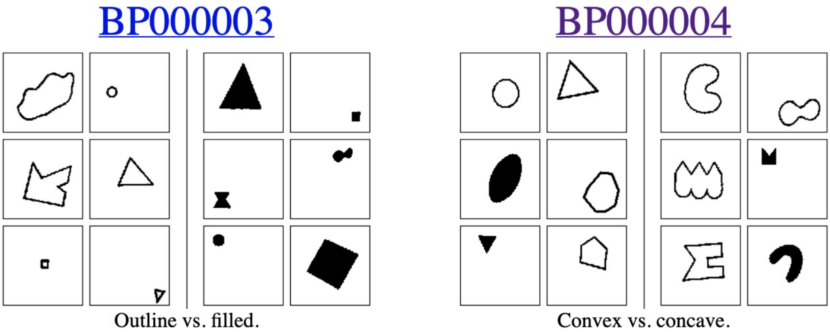
Example: log of a timed system (e.g., timed automata, Petri nets, etc.)

Traces: events that occur at certain timestamps: e.g, `event(a,100).`

Possible tasks: event recognition, fault detection, forecasting, diagnostics, …

Table of Contents

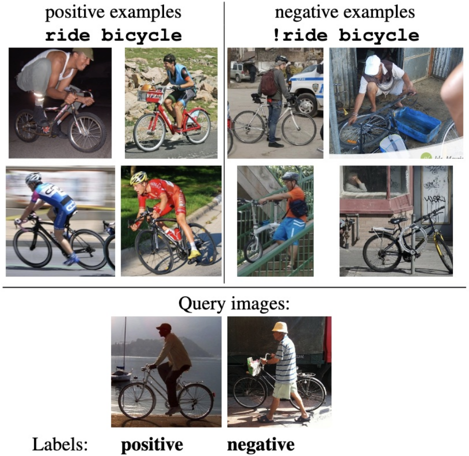
4 Visual reasoning tasks



Bongard problems

Difficult tasks for (deep) neural networks

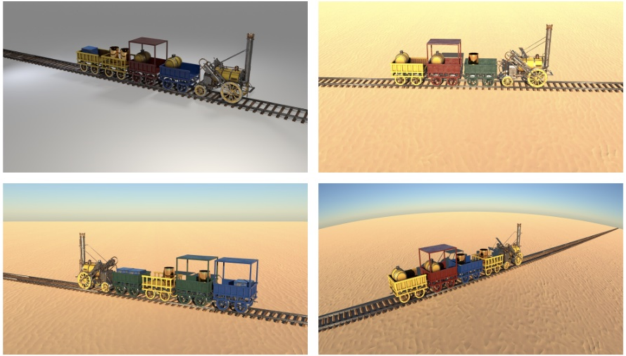
4 Visual reasoning tasks



Bongard-HOI (Human-Object Interaction)

Difficult tasks for (deep) neural networks

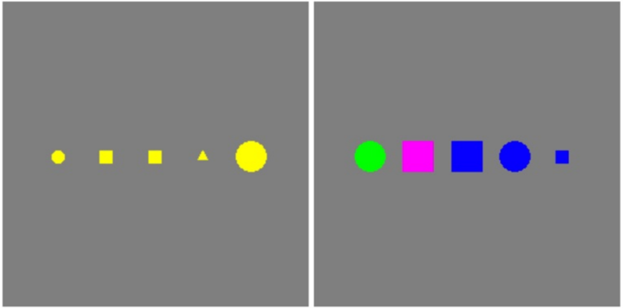
4 Visual reasoning tasks



V-LoI (Michalski trains...)

The KANDY benchmark

4 Visual reasoning tasks

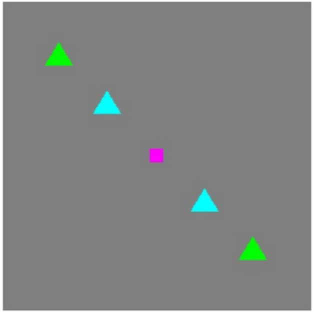


(a) Positive: Every object is yellow (b) Negative: There is no shared attribute

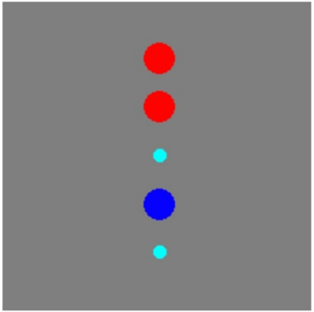
Lorello et al., 2025

The KANDY benchmark

4 Visual reasoning tasks



(a) Positive: A diagonal palindrome

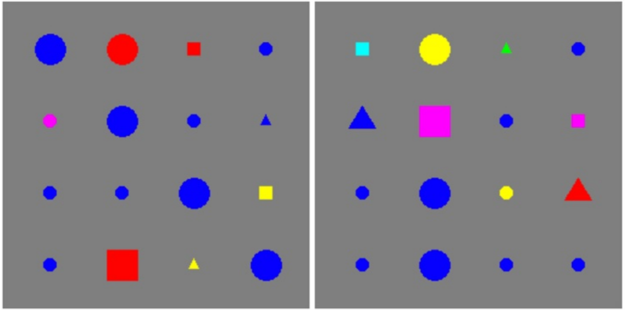


(b) Negative: A non palindromic vertical line

Lorello et al., 2025

The KANDY benchmark

4 Visual reasoning tasks



(a) Positive: The two diagonals contain only blue circles.

(b) Negative: Blue circles are not positioned at diagonals.

Lorello et al., 2025

Table of Contents

5 Explainable music classification

Music classification

There exist several tasks for music classification

- Genre classification
- Real vs. synthetic piece classification
- Authorship attribution
- ...

Most of the approaches are purely neural (and work reasonably well)
Could we use XAI? For example, ProtoPNets applied to spectrograms...

Music classification

Music can be generated by Large Language Models too

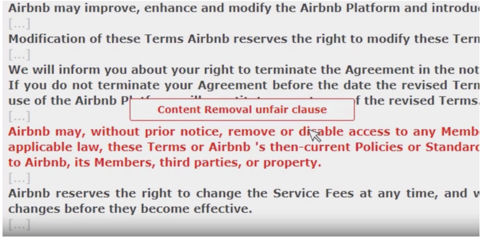
- Can we distinguish LLM-generated vs. human-generated music?
- Can we distinguish among music generated by different LLMs?

Most of the approaches are purely neural (and work reasonably well)

Table of Contents

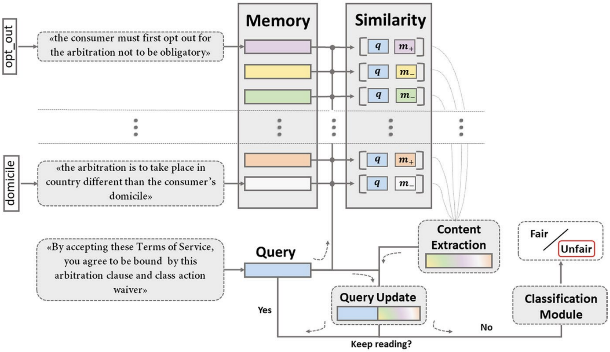
Unfair clause detection in contracts

Automatically detect **potentially unfair clauses** in online contracts (a.k.a. Terms of Service): build an automated CLAUse DETecTEr (CLAUDETTE)



M. Lippi, P. Palka, G. Contissa, F. Lagioia, H.-W. Micklitz, G. Sartor P. Torroni, *Artificial Intelligence and Law.*, 2019

Unfair clause detection in contracts



(a)

F. Ruggeri, F. Lagioia, M. Lippi, P. Torroni, *Artificial Intelligence & Law*, 2022

Unfair clause detection in contracts

Unfair clause detection in contracts

Potentially unfair clause #4

EXCEPT FOR CERTAIN TYPES OF DISPUTES MENTIONED IN THE ARBITRATION CLAUSE , YOU AND HEADSPACE AGREE THAT DISPUTES RELATING TO THESE TERMS OR YOUR USE OF THE PRODUCTS WILL BERESOLVED BY MANDATORY BINDING ARBITRATION , AND YOU WAIVE ANY RIGHT TO PARTICIPATE IN A CLASS-ACTION LAWSUIT OR CLASS-WIDE ARBITRATION .

Unfairness categories: **Arbitration**

Hide/show rationales

Potentially unfair clause #5

1.4 CHANGES TO TERMS Headspace reserves the right to change or update these Terms , or any other of our policies or practices , at any time , and will notify users by posting such changed or updated Terms on this page .

Unfairness categories: **Unilateral Change**

Hide/show rationales

The clause is potentially unfair for **Unilateral Change** since the provider has the right for unilateral change of the contract, services, goods, features for any reason at its full discretion, at any time (score = 0.834)

Potentially unfair clause #6

Your continued use of the Products constitutes your agreement to abide by the Terms as changed .

Unfairness categories: **Contract by Using**

Hide/show rationales

Unfair clause detection in contracts

Possible directions to explore...

- Local explanations (e.g., LIME, SHAP)
- Decision rules or trees from handcrafted features (e.g., synonyms, concepts)
- Measure XAI performance with legal rationales