



**HAL**  
open science

## Use of metaheuristics for machine learning in the context of ALS

Thibault Anani, Jean-François Pradat-Peyre, François Delbot, Pierre-François  
Pradat

► **To cite this version:**

Thibault Anani, Jean-François Pradat-Peyre, François Delbot, Pierre-François Pradat. Use of metaheuristics for machine learning in the context of ALS. 8th ALS and MND research meeting, Oct 2022, Paris, France. hal-04246306

**HAL Id: hal-04246306**

**<https://hal.science/hal-04246306v1>**

Submitted on 17 Oct 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

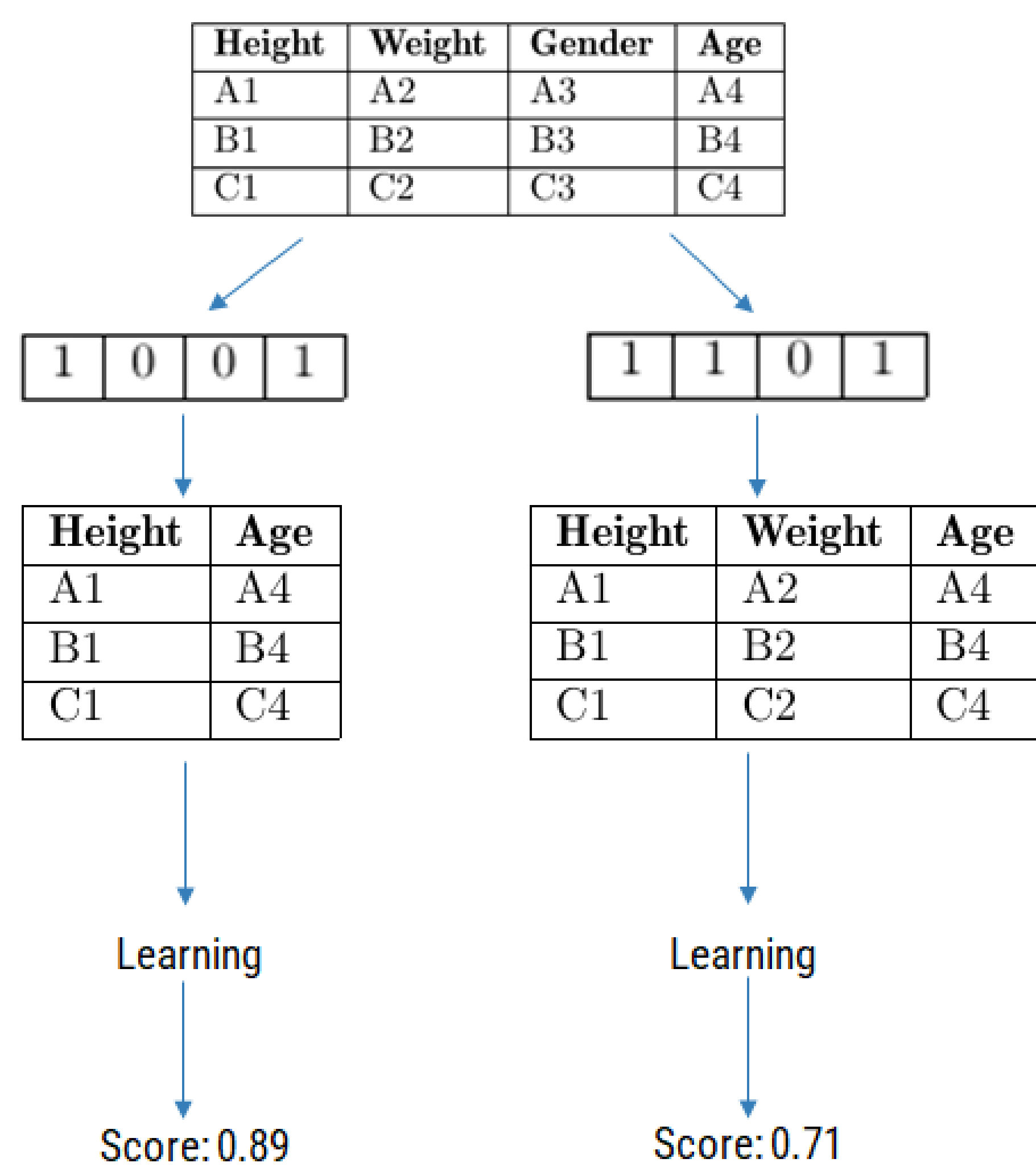
## Background

- Strong patient heterogeneity leads to unreliable prognosis.
- Machine learning methods based on large patients datasets have successfully identified underlying correlations in patient data as in the Pro-Act challenge [1].
- Limited performance of machine learning models because of too many irrelevant features leads to poorly usable results to predict disease progression.

## Method

- The data consist of 2983 patients from 2 different datasets: Clinical trials from Pro-Act [1] and Exhonor [2].
- The data available across the datasets are age, sex, site of onset, ALSFRS, weight, height, time since first symptom, etc. both at T0 (first visit) and at T3 (next visit 3 months later) with a total of 131 features to forecast the evolution of the disease over the next few months (patients' ALSFRS)
- An autoregressive model allows each forecast of ALSFRS in a patient to feed the value back into the model to forecast its new value for a patient 3 months later.
- Feature selection consists in choosing a part of the features considered as the most relevant in the construction of the model to improve the performance of the learning algorithm used.

Figure 1. Illustration of feature selection



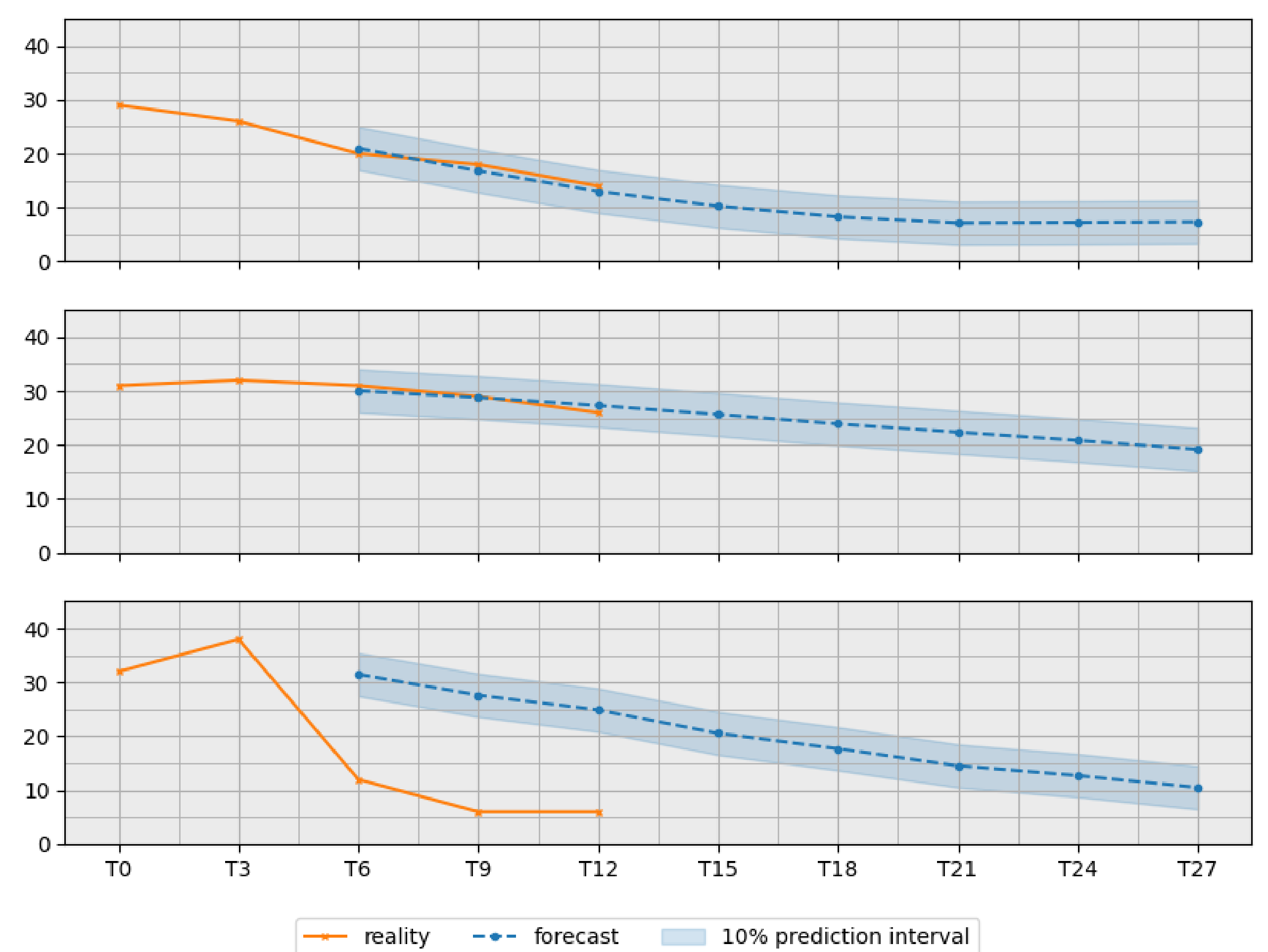
- The evaluation of all possible combinations leads to a combinatorial explosion due to the overly high number of features.
- Metaheuristics are computational methods for solving complex optimisation problems and finding a solution close to the optimal solution.
- An improved version of differential evolution [3] was developed to find the best possible feature subset.

## Results

Method	RMSE	R <sup>2</sup> -adjusted	Nb features
No selection	3.64	77.84	131
Metaheuristic	3.61	78.29	83

- Feature selection significantly reduces the number of features required to design the model while increasing its predictive efficiency.
- Overall our model is close to 3.61 on average compared to the actual value.
- The selected features explain 78% of the variation in ALSFRS among patients.

Figure 2. Comparison of forecast with reality in 3 patients



- An estimate of the evolution of ALSFRS in a patient is plotted against the patient's characteristics.
- The model works well when a patient has slow or moderate disease progression. However, the model fails to forecast a recovery or a very strong progression.

## Discussion

- Implementation: Clinicians enter patient stats to obtain a forecast of disease progression using a web interface.
- Additional data will provide a more relevant model.
- The model can be used regardless of where the patient is in the progression of the disease.

## References

- N et al Atassi. The pro-act database: design, initial analyses, and predictive features. *Neurology*, 83.19:1719-1725, 2014.
- V et al Meininger. Pentoxifylline in als: a double-blind, randomized, multicenter, placebo-controlled trial. *Neurology*, 66.1:88-92, 2006.
- T et al Anani. Experimental comparison of metaheuristics for feature selection in machine learning in the medical context. pages 194-205, 2022.
- Ali Furkan Kalay. Uq beamerposter template. *Overleaf Poster Templates*, 2021.