

FastParam

Strategie di Machine Learning per l'accelerazione dei modelli parametrici per la rilevazione delle onde gravitazionali nei segnali delle pulsar

Luca Naso, Koexai S.r.l.

Spoke 3 Progetti Bandi a Cascata, 09/01/2025

FastParam

Strategie di Machine Learning per l'accelerazione dei modelli parametrici per la rilevazione delle onde gravitazionali nei segnali delle pulsar

Area Tematica: 6

Metodologie avanzate di analisi dati

Sotto-tematica: d)

Porting di codici di analisi Bayesiana basati su diverse decine/centinaia di parametri liberi su architetture HPC

Project Overview: challenges and proposed solutions I

Context and challenges

Gravitational Waves (GWs) are ripples in spacetime caused by accelerating masses, such as merging black holes. They provide experimental confirmation of General Relativity and insights into cosmological models and alternative theories of gravity. GWs are difficult to detect due to their extremely small amplitude, requiring highly sensitive instruments. **Pulsars** are fast-rotating neutron stars that emit regular pulses of radiation, allowing for precise timing measurements. Fluctuations in the arrival times of these pulses can indicate the presence of propagating GWs. The **European Pulsar Timing Array (EPTA)** is an international collaboration aimed at measuring nanohertz GWs by analyzing pulsar signals, since 2006.

The **models** used for analyzing pulsar signals are complex due to the numerous parameters that must be considered to account for various astrophysical phenomena. This complexity leads to a vast parameter space, resulting in significant computational challenges and extended processing times. Consequently, researchers have to make simplifying assumptions about certain parameters to expedite the process, which can compromise the accuracy of the results

Project Overview: challenges and proposed solutions II

Proposed solutions

FastParam aims to enhance the detection of GWs by optimizing the analysis of pulsar signals through advanced **Machine Learning techniques**.

The objective is to **reduce the analysis time** of current scientific models.

This will be achieved by exploring two different strategies:

1. The first one focuses on intelligent **sampling practices of the parameter space**, inspired by various Machine Learning techniques used for similar problems, such as Hyper-Parameter Optimization.
2. The second strategy seeks solutions based on **dimensionality reduction** techniques.

Technical Objectives and Methodologies

Parameter Sampling: search for the best techniques for intelligent sampling of the parameter space. This includes the use of grid search, random search, and genetic algorithms, which allow for a thorough examination of potential parameter configurations. Additionally, gradient-based algorithms and Neural Architecture Search (NAS) algorithm For deep neural networks, are considered.

Dimensionality Reduction: search for the best dimensionality reduction techniques, including Principal Component Analysis (employed to select the most significant parameters or combinations of parameters, simplifying the model without losing critical information), Linear Discriminant Analysis (used to project data into lower-dimensional spaces, facilitating easier analysis and interpretation), Least Absolute Shrinkage and Selection Operator (LASSO, regression analysis performing both regularization and parameter selection), and Autoencoders (a type of deep learning model utilized to encode high-dimensional spaces into lower-dimensional representations).

Accelerated Pipeline: integrated pipeline that combines the most promising techniques from both hyper-parameter tuning and dimensionality reduction approaches.

Involved Staff and new recruitments

Spoke Scientific Advisor: Prof. Massimo Meneghetti, INAF Osservatorio di Astrofisica e Scienza dello Spazio (OASS) Bologna

Koexai Team size: 3,05 FTE = 5.240 hours

1. Luca Naso: Project Lead (PhD in Astrophysics)
2. Marco Cataldo: Project Administrative Manager
3. Romina Anfuso: Project Manger
4. Vincenzo Del Zoppo: Data Scientist Team Lead
5. Luigi D'Amico: Senior Data Scientist
6. Aldo Barca: Data Scientist
7. Francesco Alì: Data Scientist
8. Fatima Modica Bittordo: Data Scientist

Gantt

OR	AR	Titolo	T1			T2			T3			T4		
			1	2	3	4	5	6	7	8	9	10	11	12
1	<i>Studio e caratterizzazione del dominio astrofisico e dei dataset disponibili</i>		█	█	█									
	1.1	Studio sullo stato dell'arte di onde gravitazionali da segnali pulsar e relativi modelli scientifici	█	█	█									
	1.2	Studio e qualificazione dei dataset disponibili e collaborazioni internazionali		█										
2	<i>Ricerca e caratterizzazione di strategie di campionamento e di tecniche di riduzione della dimensionalità, per modelli complessi ad alta dimensionalità</i>					█	█	█	█	█	█	█	█	
	2.1	Studio e definizione di strategie di campionamento per modelli complessi ad alta dimensionalità				█	█	█	█	█	█			
	2.2	Ricerca e determinazione di tecniche di riduzione della dimensionalità per modelli complessi ad alta dimensionalità				█	█	█	█	█	█			
	2.3	Evoluzione e caratterizzazione delle strategie di campionamento e delle tecniche di riduzione della dimensionalità							█	█	█	█	█	
3	<i>Studio e caratterizzazione della pipeline di accelerazione e condivisione dei risultati</i>							█	█	█	█	█		
	3.1	Studio e definizione di un sistema integrato di accelerazione delle analisi						█	█	█	█	█		
	3.2	Evoluzione e caratterizzazione della pipeline di accelerazione e condivisione dei risultati									█	█		

Timescale, Milestones, SAL

Project start: Dec 2024

