



Finanziato
dall'Unione europea
NextGenerationEU



Ministero
dell'Università
e della Ricerca



Italiadomani
PIANO NAZIONALE
DI RIPRESA E RESILIENZA



Artificial Intelligence for Distances in Astronomy

Prof. Emanuele Lattanzi, University of Urbino Carlo Bo

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

Involved Staff and new recruitments

Responsabile Scientifico



Emanuele Lattanzi / Professore Associato

IINF-05/A 09/IINF-05 Sistemi di elaborazione delle informazioni

Struttura	Sede	Telefono	Email
Dipartimento di Scienze Pure e Applicate (DiSPeA)	Urbino Piazza della Repubblica, 13	0722 304411	emanuele.lattanzi@uniurb.it



Valerio Freschi / Professore Associato

IINF-05/A 09/IINF-05 Sistemi di elaborazione delle informazioni

Struttura	Sede	Telefono	Email
Dipartimento di Scienze Pure e Applicate (DiSPeA)	Urbino Piazza della Repubblica, 13	0722 304418	valerio.freschi@uniurb.it



Alessia Elisabetta Kogoj / Professoressa Associata

MATH-03/A 01/MATH-03 Analisi matematica

Struttura	Sede	Telefono	Email
Dipartimento di Scienze Pure e Applicate (DiSPeA)	Urbino Piazza della Repubblica, 13	0722 304421	alessia.kogoj@uniurb.it



Borsista di ricerca dal 1° Febbraio

IINF-05/A 09/IINF-05 Sistemi di elaborazione delle informazioni

Struttura	Sede	Telefono	Email
Dipartimento di Scienze Pure e Applicate (DiSPeA)	Urbino Piazza della Repubblica, 13		

Scientific collaboration

Sezione di Fisica del Dipartimento di Scienze Pure e Applicate dell'Università degli Studi di Urbino Carlo Bo

La ricerca in Fisica Astroparticellare vede la partecipazione del personale della Sezione a grandi esperimenti internazionali.

- LIGO-VIRGO
- coinvolgimento nella progettazione della sonda interplanetaria eLISA per la rivelazione di onde gravitazionali a bassissima frequenza, con la partecipazione all'esperimento LISA Pathfinder
- La Sezione è inoltre impegnata nello sviluppo di nuovi detector per onde gravitazionali
- Personale:
 - Gianluca Guidi (PO)
 - Andrea Viceré (PO)
 - Michele Veltri (PA)
 - Catia Grimani (PA)
 - Filippo Martelli (PA)

Project Overview

- **Tematica 5** - Intelligenza Artificiale applicata a Missioni Spaziali.
- **Sotto-tematica C**: sviluppo di algoritmi e modelli con tecniche di Machine Learning/Deep Learning per stimare le distanze astronomiche per l'analisi dei dati provenienti dal satellite spaziale Gaia
- **Lo scopo** del progetto è progettare **modelli di Machine Learning (ML) e Deep Learning (DL)** che, a partire dai dati del catalogo di Gaia DR3 permettano di stimare in modo indipendente e **accurato i valori di parallasse**. Utilizzo dei modelli per acquisire conoscenze più dettagliate **sulla scala delle distanze astronomiche** misurate con i diversi indicatori della distanza, con lo scopo di **contribuire alla determinazione della costante di Hubble**.
- **Il risultato finale sarà costituito dalla costruzione di una pipeline per l'analisi dei dati astrometrici e fotometrici.**

Technical Objectives, Methodologies and Solutions

-Analisi preliminare dei dati

- **Input principali:** posizione, parallasse, moto proprio, fotometria (G, BP, RP), e parametri astrofisici (ad esempio temperatura, gravità superficiale e metallicità).
- **Selezione dei dati:** escludere le sorgenti con qualità bassa ed eventualmente integrare con altri cataloghi?
- **Gestire i bias noti:** Il catalogo Gaia ha noti bias nelle parallassi, soprattutto alle alte magnitudini.

- Definizione del problema

- **Regressione:** Il modello deve predire il valore di parallasse basandosi sugli altri parametri come input.
- **Ground truth:** usare stelle con **parallassi di alta qualità** come set di training (ad esempio, stelle vicine con errori di parallasse bassi, **oppure pesare la funzione di loss** in base alla qualità del dato?).

Technical Objectives, Methodologies and Solutions

-Progettazione del modello

- **Machine Learning (ML): Random Forests o Gradient Boosted Trees** possono catturare relazioni non lineari tra input e parallasse.
 - *Pro:* **velocità di addestramento.**
 - *Contro:* potrebbero non essere abbastanza flessibili per relazioni molto complesse.
- **Deep Learning (DL): Reti Convoluzionali (CNN), Reti Ricorrenti (RNN), Transformer:**
 - *Pro:* Capacità di modellare relazioni complesse e apprendere direttamente dai dati grezzi.
 - *Contro:* Richiedono grandi quantità di dati e risorse computazionali.

Technical Objectives, Methodologies and Solutions

- Sfide e considerazioni

- **Bias di parallasse:** I modelli rischiano di apprendere i bias già presenti nei dati di Gaia. È importante considerare metodi di debiasing o integrare cataloghi esterni?
- **Incertezze nei dati:** incorporare l'incertezza dei dati come pesi della funzione di loss usata in addestramento?
- **Numerosità dei dati:** L'addestramento potrebbe richiedere risorse significative per gestire l'enorme mole di dati.

Timescale, Milestones, SAL

AIDA - Project Gantt					ANNO 1											
					dic-24	gen-25	feb-25	mar-25	apr-25	mag-25	giu-25	lug-25	ago-25	set-25	ott-25	nov-25
WPs	Descrizione	Mese Inizio	Mese Fine	Durata [in mesi]	1	2	3	4	5	6	7	8	9	10	11	12
ATTIVITA' A - Creazione e processing del dataset		1	3	3												
Att A.1	Creazione del dataset di riferimento	1	3	3			MS1									
ATTIVITA' B - Sviluppo Pipeline		4	9	6												
Att B.1	Sviluppo degli algoritmi basati sulle reti neurali	4	6	3						MS2						
Att B.2	Combinazione degli algoritmi sviluppati in un unico pipeline	7	9	3									MS3			
ATTIVITA' C - Testing e Misurazioni		10	12	3												
Att C.1	Applicazione della pipeline su dati da diversi survey astronomici	10	12	3												MS4

Milestones

- **MS1**: Creazione del dataset rappresentativo (**M3**). Modalità di verifica: rilascio del deliverable D.A.1
- **MS2**: Creazione dei modelli di machine learning (**M6**). Modalità di verifica: rilascio dei modelli (deliverable D.B.1)
- **MS3**: Creazione della pipeline (**M9**). Modalità di verifica: rilascio dei modelli organizzati in pipeline (deliverable D.B.2)
- **MS4**: Finalizzazione del progetto. Rilascio report finale deliverable D.C.1 (**M12**)