Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
ICSC Centro Nazionale di Ricerca in HPC, Big Data and Quantum Computing

Deep learning of 87A-like supernovae progenitor characteristics: training the Inception model on synthetic data

M. Grassia, G. Mangioni, S. Cosentino*, M. L. Pumo

Spoke 3 General Meeting, Elba 5-9 / 05, 2024

* External collaborator

Scientific Rationale

- In the next few years, the **number of available supernovae observations will increase exponentially** thanks to the LSST Survey

- **Inferring the supernovae's progenitor characteristics is** currently performed using several **computationally expensive** methods
  - An example is **Bayesian inference**, that is computationally challenging as a Monte Carlo search is involved to test the parameters

- Given the increasing number of observations, **finding the characteristics of all the supernovae' progenitors will be** practically **unfeasible** with current methods
  - Supernovae classification (e.g., 87A, type 1, etc.) , on the other hand, is an easier task

Technical Objectives

- In this work, **we aim at training a machine to infer the characteristics of the** supernovae's **progenitor**: **energy, mass, radius, and nickel content**
  - **World-first** approach, never done before

- **Machine Learning is well-suited** for the task:
  - **Noise and error-tolerance**
  - **Fast computation**
  - Easily trainable on new data
  - Can provide insights on the problem itself trough interpretability tools

- **This work is also meant as a test-bench for the new HPC capabilities** that are under deployment

Methodologies and Solutions

- **We employ** the **state-of-the-art** of trainable algorithms, the ***InceptionTime* model**, **for time-series data to infer** (through regression) the characteristics
- Each **observation** includes **two timeseries** the **bolometric luminosity** and **photo-spheric velocity**
- Unfortunately, **real-world observations are scarce**
- **We train the model on generated** time-series **data from synthetic models**, that provide **bolometric luminosity** and **photo-spheric velocity**

Methodologies and Solutions

- **Synthetic models available**:
  - **Semi-analitical**: very **fast**, produces **rough approximations** that do not exactly match the real-world observations
  - **Hydrodynamical** simulations: **computationally expensive** and **slow**, accounts for many phenomena happening in the star
- **Dataset generated**:
  - **205k semi-analitical observations of 87A-like**
  - **59 hydrodynamical observations of 87A-like**

Accomplished Work, Results

- **First iteration:**
  - **Training and testing on** the **semi-analitical** data
  - Excellent results: **Mean Absolute Percentage Error** (MAPE) lower than 1% on all the four characteristics (radius, mass, energy, nickel content)

Accomplished Work, Results

- **Second iteration:**
  - **Training on** the **semi-analitical** data
  - **Validation and testing on** the **hydrodynamical** observations

  - Very **poor results (MAPE)**

Radius	Mass	Energy	Nickel
98%	48%	193%	186%

  - Removing the velocity curve and trimming the first ~30 days (that are way off in the semi-analytical model) and extending the tail (that only depends on the nickel content and that can be easily inferred) didn't help

Accomplished Work, Results

- **Third iteration: Curriculum Learning**
  - **Training on** the **semi-analitical** data
  - **Validation on** the **hydrodynamical** observations
  - The **best model is re-trained** on the **hydrodynamical validation set**
  - **Testing** is then performed on a different split of **hydrodynamical**

  - **Major improvement in results (MAPE)**

	Radius	Mass	Energy	Nickel
baseline	98%	48%	193%	186%
C. learning	**61%**	**20%**	**41%**	**39%**

  - Grid search was also implemented, but needs computational power to be run

Next Steps and Expected Results

- **Generalize** from synthetic data **to real-world observations**

- **Dataset improvement**
  - New and more accurate synthetic model to improve the training data
  - More hydrodynamical simulations

- **Grid search** over the hyper-parameters of the model

- **Model interpretability/explanation tools to understand** what are **the most important phases (in days) for the task**
  - This is meant **to help researchers spare observational resources** that can be employed elsewhere

Timescale, Milestones and KPIs

- **M7:**
  - **Model refinement**
  - **Code development and improvements**
  - **KPI: code / data**
- **M8**
  - **Grid search** of hyper-parameters
  - **KPI: numerical results, code / data**
- **M9**
  - Design and development of **timeseries key points extraction**
  - **Generalization to real-world examples**
  - **KPI: numerical results, code / data**
- **M10**
  - **KPI: paper preparation + results**

Thank you!

InceptionTime model

- Proposed for **Time Series Classification (TSC) tasks**
- **Stacks multiple Inception-v4 layers** initialized with various weights
  - **Convolutional Neural Networks** (CNNs) are **at the core**
  - Combines **pooling, batch normalization** and
  - optionally includes **residual connections**