Astrophysics Project:
Test Deep learning architectures to compress
the information of Stellar Profiles

Abstract
Stars are a basic component of structure in the Universe. The study of how stars born, evolve, die and interact with the interstellar medium is fundamental to understand many aspect of our Universe. Today, we have an established knowledge of the structure of Star and how it evolves. The structure of a Star is described by a set of partial differential equations while its evolution is accounted for by a network of nuclear interactions that produce the energy required to sustain the stellar luminosity and to maintain the Star in hydrostatic equilibrium. Numerical procedures are required to solve the set of differential equations and estimate the stellar profiles, i.e. how the stellar properties change as function of the stellar radius (or mass). The numerical complexity of shuch procedure limits its usage for stellar population studies (e.g. study of the properties of the black holes produced by different stellar populations). In such cases population synthesis codes are used. Such codes account for the stellar evolution in a simplified way using fitting equations of interpolation pre-computer stellar tracks. However, only the “macroscopic” properties of the star (such as the surface luminosity and temperature, and mass) are taken into account, the information about the detailed stellar profile is lost. The possibility to access this information also in population synthesis code could be useful to improve them and open the possibility to use more detailed formalism for both stellar and binary processes.

Project goal
The main aim of this project is to investigate whether it is possible to compress the information of the internal structure of Stars in a lower dimensional space by using Autoencoder Neural Network architectures.
The development of an efficient compression tool can make it possible to distribute and store in a more efficient way the stellar evolution models computer with hydrostatic stellar evolution code.
In addition the possibility to use interpolation in such lower dimensional space to obtain the stellar structure without running a full hydrostatic simulation has the potential to have a breakthrough impact on stellar

population studies. In particular, it can be included in population synthesis codes based on track interpolations such as SEVN. Those codes are at the moment limited to the interpolation of macroscopic properties such as the stellar radius, luminosity, etc.. the possibility to retrieve on-the-fly the detailed stellar structure could allow the use of new detailed formalism to take into account stellar and binary processes.

Organisation

The organization of the project is very flexible. We will not have lectures, but rather I will give you a number of resources that you will use at your own pace and preference. Then, we will schedule a number of zoom or in person meetings (depending on my and your availability). In the meeting we will discuss all the aspects of the project and its development.

Project milestones

These project milestones represent a tentative development line for the project. As usual in research projects, they can be adapted, updated or even totally changed depending on the results you obtained, your idea and personal preference.

- *Understand the basics of stellar structure and stellar evolution.*
  - In order to have a basic understanding of the physics involved in this project you have to look at the recorder lectures in Moodle under the Astrophysics Projects section (https://stem.elearning.unipd.it/course/view.php?id=8879#coursecontentcollapse6), in particular:
    - Summary of stellar evolution: https://stem.elearning.unipd.it/mod/kalvidres/view.php?id=499227 (neglect everything before minute 14, it was specific of the past year projects)

- *Write a code to load and handle the stellar evolution profile from MESA simulations.*
  - You have to familiarize with the dataset, e.g. stellar profiles obtained with the stellar evolution code MESA (see Data: stellar profile from MESA simulations). These datasets are the input for the AutoEncoder network, so you have to write codes to read the information you need from the file and save it in a format that can be used for the Autoencoder

training. You can find an example script here http://user.astro.wisc.edu/~townsend/resource/tools/mesa-web/mesa_web.py

- *Implement a basic Autoencoder and apply it to stellar profiles of selected properties*
  - You have to implement your own Autoencoder version using a framework such as Tensorflow or PyTorch. In this first test consider the properties density and mass as function of the radius. Develop and test autoencoders testing different dimension for the latent variables. It is important to notice that each profile can have a different number of points and the mass and radius of each shell can be very different. Therefore you have to think a way to standardize your data so that your input contains always the same number of points in the same range.

- *Now let's do a step forward, implement a more sophisticated convolutional autoencoder*

  Implement the method described in this paper: https://arxiv.org/pdf/2306.06938.pdf.

  The author of the papers tests something very similar to what we are trying with this project. The difference is that they are focusing on the stellar atmosphere (the outermost region of the stars) rather than we are considering the full stellar structure. In the paper they are considering the properties *mass*, *temperature, pressure, electron density*, you will instead consider *mass, temperature, density, energy* (see Tab. 2).

  If you have time and you want to expand further this analysis, add to the properties also the chemical abundance (idrogen), Y (helium) . The X is simply the property h1 (see Tab. 2), while Y is the sum of the property h3 and he4 (see Tab. 2).

Data: stellar profiles from MESA simulations

The Data used in this work are stellar structure profiles (e.g. how stellar parameters vary as function of the radius or the mass in the stellar interior) obtained with the detailed stellar evolution code MESA (https://docs.mesastar.org/en/release-r24.03.1). The evolution of a star is solely defined by its initial mass and its metallicity (i.e. the mass abundance of elements that are heavier than helium). During stellar life the internal profiles change due to the change in chemical composition and energy output. Therefore for each couple of initial mass and metallicity there are multiple profiles depending on age.

Pregenerated-dataset

A set of stellar profiles that you can start to explore can be found at this link: https://www.dropbox.com/scl/fo/9jcvhhgwpicddfzb5xuao/AKXwh39qalaIGXopaMkZwpk?rlkey=xly7k60fcr9uy68mz8uhhzt8v&dl=0.

Each folder contains the evolution of a star at given initial mass and metallicity. The folder name is: *MESA-WEB_MXX_ZKYYY* where XX is the initial mass of the star in solar masses and KYYY is the star metallicity in which the K represent the unit, so MESA-WEB_M20_Z002 refers to the evolution of a star with initial mass 20 Msun and metallicity Z=0.02. Inside the folder the stellar structures are stored in the file *profileJ.data* where J is a sequential number referring to different evolutionary times at which the profiles have been saved.

The file contains the following information:

- First three rows: general information about the star:
  - first row: index of the column
  - second row: name of the variable
  - third row: value

  There are a total of 52 variables, but the only that could be important for your project and the relative description is reported in Tab.1.
- Rest of the rows: stellar profile
  - 4th row: index of the column
  - 5th row: variable name
  - 6th-..: profiles for all the variables

  The stellar evolution codes divide the interior of the stars in a series of spherical shells, each shell is completely identified by the radius or by the amount of mass within the radius of

that shell (first two variables). So for example the density profile of a star can be either the density-radius relation or the density-mass relation.
The files contains 56 variables defined in each shell, the one that we are interested in for this project are reported in Tab. 2

Additional information on the files can be found at:
http://user.astro.wisc.edu/~townsend/static.php?ref=mesa-web-output

Expand the dataset
In order to expand the dataset of stellar profiles you can use the MESA web query at this link:
http://user.astro.wisc.edu/~townsend/static.php?ref=mesa-web-submit
Define the initial mass and the metallicity of the star in the Initial Properties section, then leave all the other properties to their default values. In General you can select how many profiles you want to produce (each profile is at a different time during the evolution). Finally insert your mail and then click on Submit. You will receive an email with the link to Download the data. A simulation can require up to 2 hours, but you can submit many simulations.

Index	Variable	Description
3	initial_mass	mass of the star at the beginning of the evolution in Msun
4	initial_z	stellar metallicity
5	star_age	age of the star in yeras
7	Teff	Surface stellar temperature in K
8	photosphere_L	Luminosity at the stellar surface in Lsun
9	photosphere_r	radius of the stellar surface in Rsun

18	star_mass	current stellar mass in Msun

Tab. 1: general properties columns, for additional information see http://user.astro.wisc.edu/~townsend/static.php?ref=mesa-web-output

Index	**Variable**	**Description**
1	mass	mass contained within the shell in Msun
2	radius	radius of the shell in Rsun
5	logRho	logarithm (log10) of the density of the shell
6	logT	logarithm (log10) of the Temperature of the shell
7	energy	Gas internal energy in erg/gr
31	h1	mass abundance of hydrgeon in the shell
32	he3	mass abundance of helium3 (isotope with two proton and one netron)
33	he4	mass abundance of helium4

Tab. 2: stellar structure variables, for additional information see http://user.astro.wisc.edu/~townsend/static.php?ref=mesa-web-output