# Stellar Mass Prediction Using IllustrisTNG Simulation

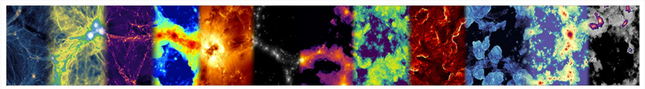
Arjina Islam
Advisors: Dr. Rachel Mandelbaum, Yesukhei Jagvaral & Dr. Sukhdeep Singh
McWilliams Center for Cosmology
Department of Physics
Carnegie Mellon University, Pittsburgh, PA

This internship is funded by the NSF AI Planning Institute for Data-Driven Discovery in Physics, Summer Research Program 2021

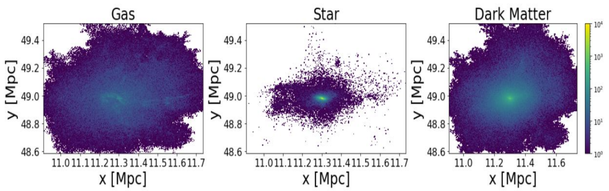
## Introduction

In this project we have trained artificial neural networks to learn the relationship between the galaxy stellar mass and the host dark matter halo mass in the IllustrisTNG simulation. The galaxy-halo connection is of great interest because dark matter makes up about 85% of matter in the universe; however, we can only observe it indirectly via its gravitational interaction with visible (stellar ) matter.

The IllustrisTNG project is a suite of cosmological hydro-simulations. The goal of the IllustrisTNG project is to understand galaxy formation and evolution, and to make predictions for current and future observational programs to broaden and deepen our understanding of galaxy physics.



*A snapshot of the Illustris simulation showing the distribution of various components of the universe at redshift zero, the present day.*




*In this plot we present the distribution of gas, star and dark matter particles in a single galaxy from the simulation. We can observe that the star particles are more concentrated than gas and dark matter particles.*

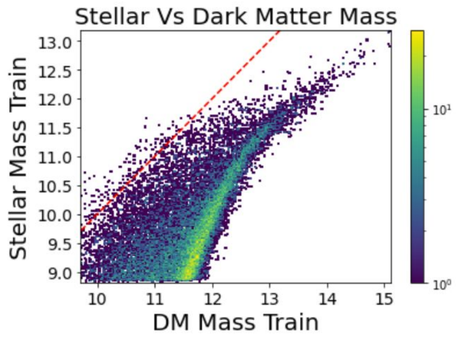
Multi-layered Perceptron (MLP) : A class of feedforward artificial neural network.

Neural network: An artificial neural network is an interconnected group of nodes, which learns from data (or are trained) by processing examples, each of which contains a known "input" and "result," learning probability-weighted associations between the two.

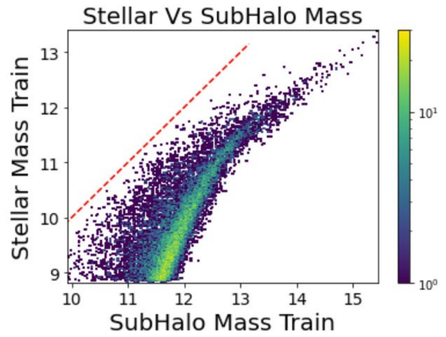
## Methodology

We used the galaxy and subhalo catalog from the IllustrisTNG100 simulation, and included galaxies that have more than 100 dark matter and star particles. The dark matter mass and the stellar mass of each galaxy were obtained from the simulation dataset. The dataset is then split into the training set and the testing set. We trained the MLPRegressor to predict the stellar mass when dark matter mass within the subhalo or the total subhalo mass (total mass of all matter components) is given.

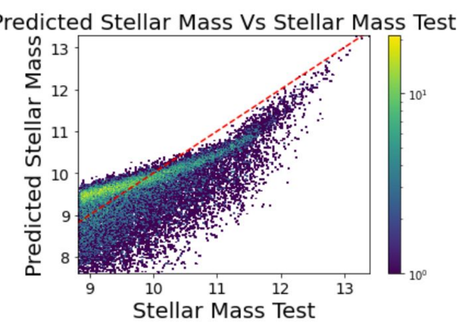
## Results



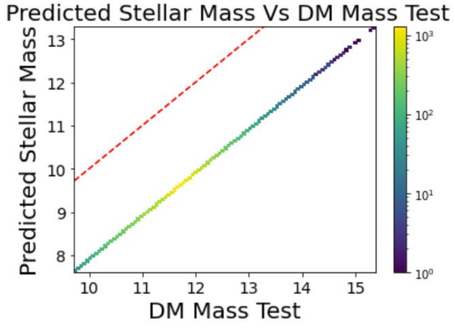

*Relationship between dark matter and stellar mass in the training set.*




*Relationship between total subhalo mass and stellar mass in the training set.*




*Relationship between the actual stellar mass for the test set and the predicted one from the model.*




*Relationship between dark matter mass from the test and predicted stellar mass.*

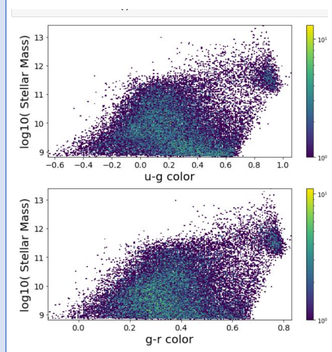
***SideNote***: All the masses are in $\log_{10}(h^{-1}\, M/M_{\odot})$. The red dashed line in all the plots represents x=y line.

## Results

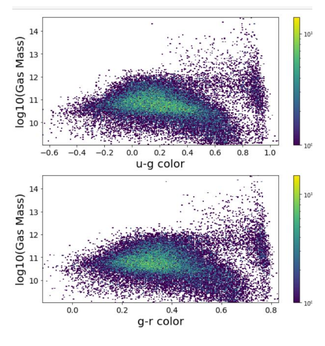
We believe that the model is underfitting the stellar mass-halo mass relation, given that we only provided one input feature. In reality this relation depends on a complex set of features that affect the galaxy physics. Current implementation of MLPRegressor is also not working well. For instance, the histogram that shows relationship between dark matter mass from the test and predicted stellar mass seems linear. However, theoretically, it is incorrect and the non-linear nature of the relation can be observed in the training sample.

## Future Endeavors

- We aim to explore other machine learning methods to improve the predictions.
- We will add more features to our network such as gas mass and galaxy color.




*Relationship between stellar mass and galaxy color*




*Relationship between gas mass and galaxy color*

## References

1. Risa H. Wechsler and Jeremy L. Tinker, “The Connection between Galaxies and their Dark Matter Halos.” Annu. Rev. Astron. Astrophys. 2018. AA:1–56.
2. Dylan Nelson , Annalisa Pillepich , Volker Springel, Rainer Weinberger , Lars Hernquist , Rudiger Pakmor , Shy Genel , Paul Torrey , Mark Vogelsberger , Guinevere Kauffmann , Federico Marinacci , Jill Naiman, “First results from the IllustrisTNG simulations: the galaxy color bimodality.” 31 Jan 2018.