TRACING GALAXY EVOLUTION WITH AUTOENCODERS

Leon Butterworth

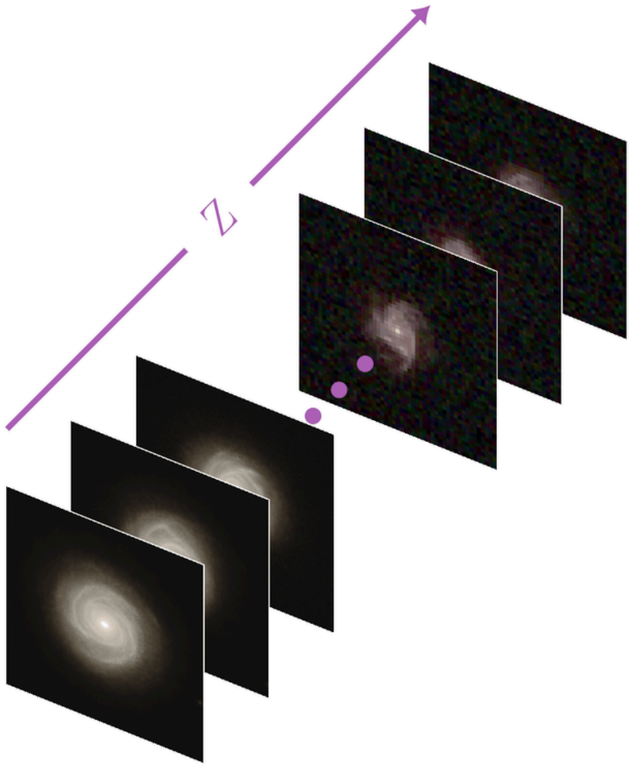
2nd September 2026

University of Hertfordshire UH

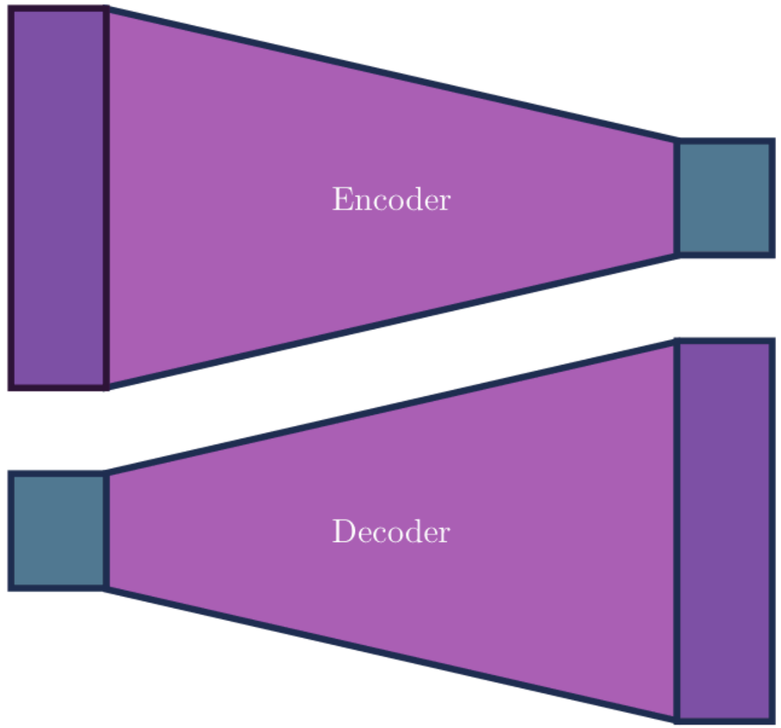
MACHINE LEARNING FOR ASTROPHYSICS

3RD EDITION VALLETTA, 31 AUG - 4 SEP, 2026

Training an autoencoder

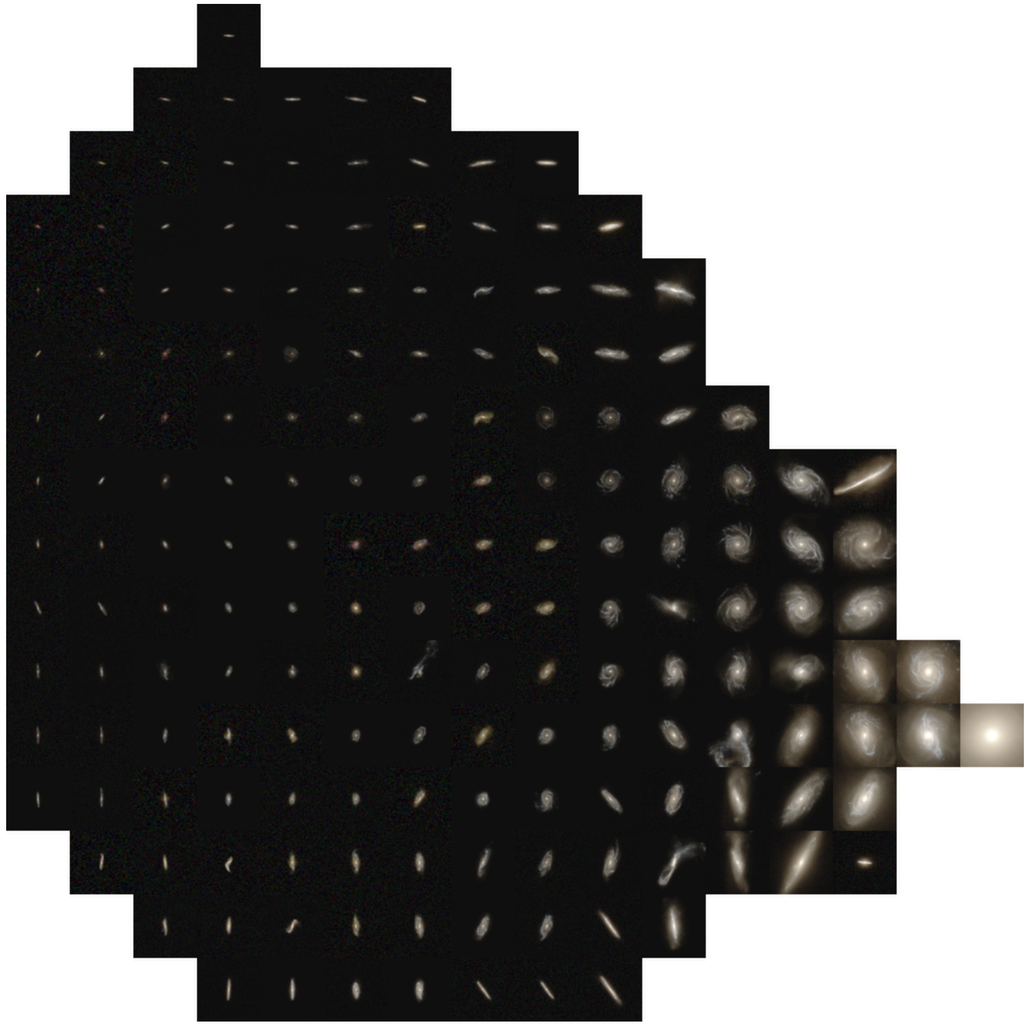


- Images of galaxies between redshift of 0 < z < 1 are created from the IllustrisTNG-50 simulation
- Images are created from 50 simulation snapshots, creating a galaxy chain



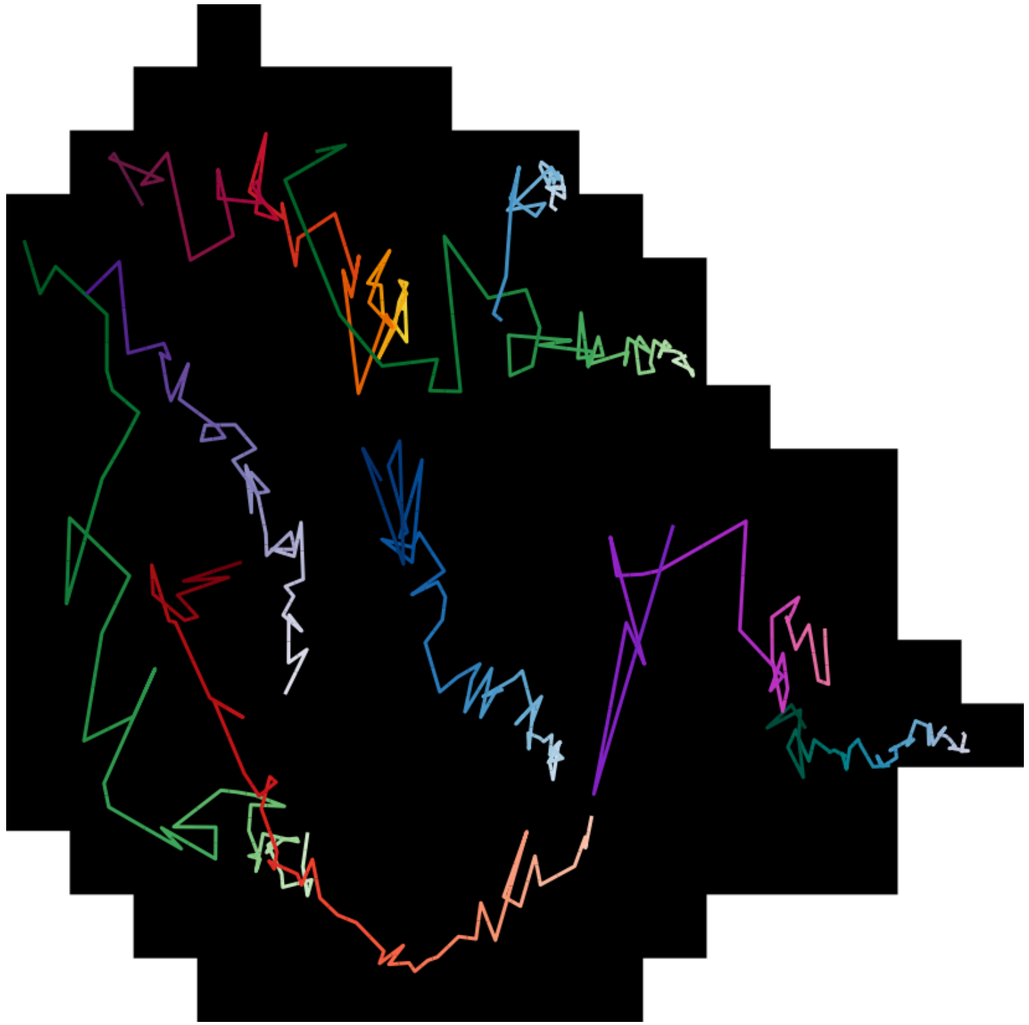
- We use these images to train an autoencoder
- The encoder creates a representation space that encodes the entire set of galaxy images

The representation space



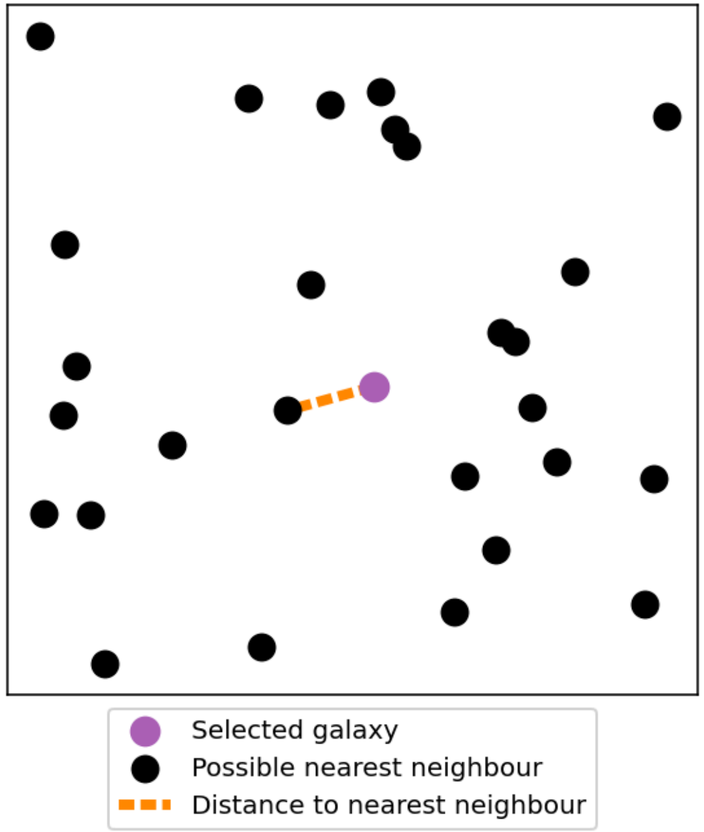
- This is a 2D projection of the encoder's representation space
- This is how the model categorises the galaxies
- The galaxies are separated by morphology

The representation space



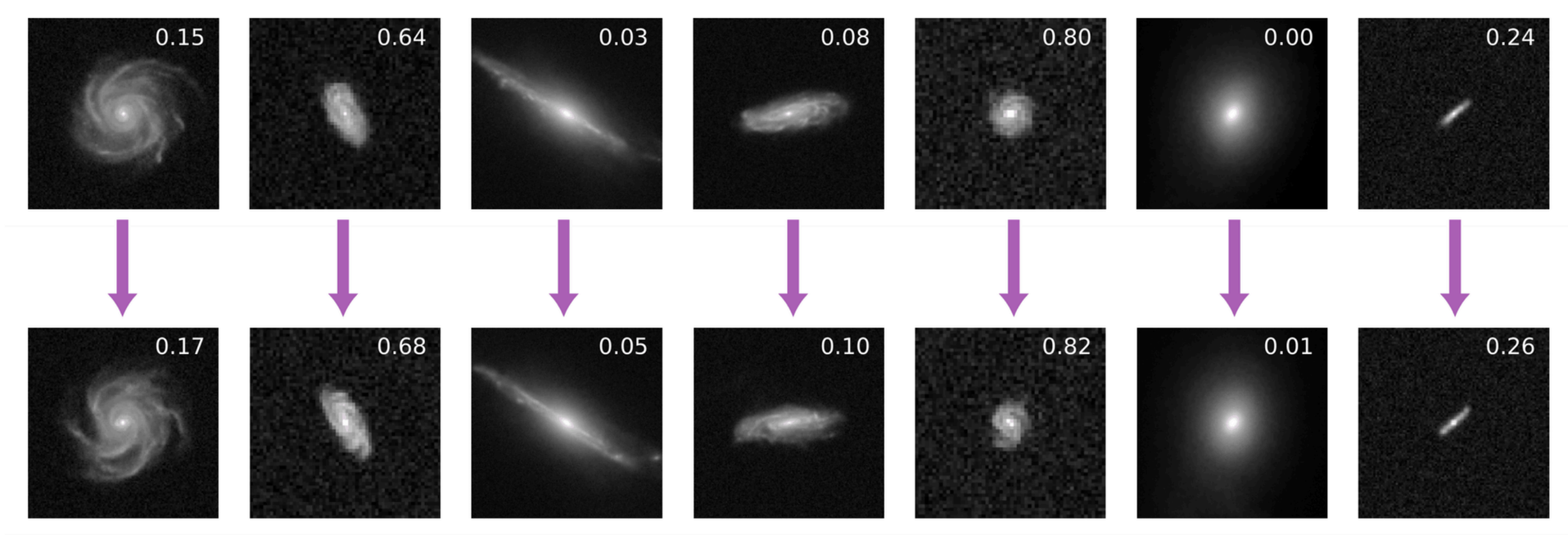
- Each coloured line is a galaxy chain from the simulation
- As the galaxies evolve with redshift, they also move through the representation space
- We aim to recreate these chains using our autoencoder

Nearest neighbour searches



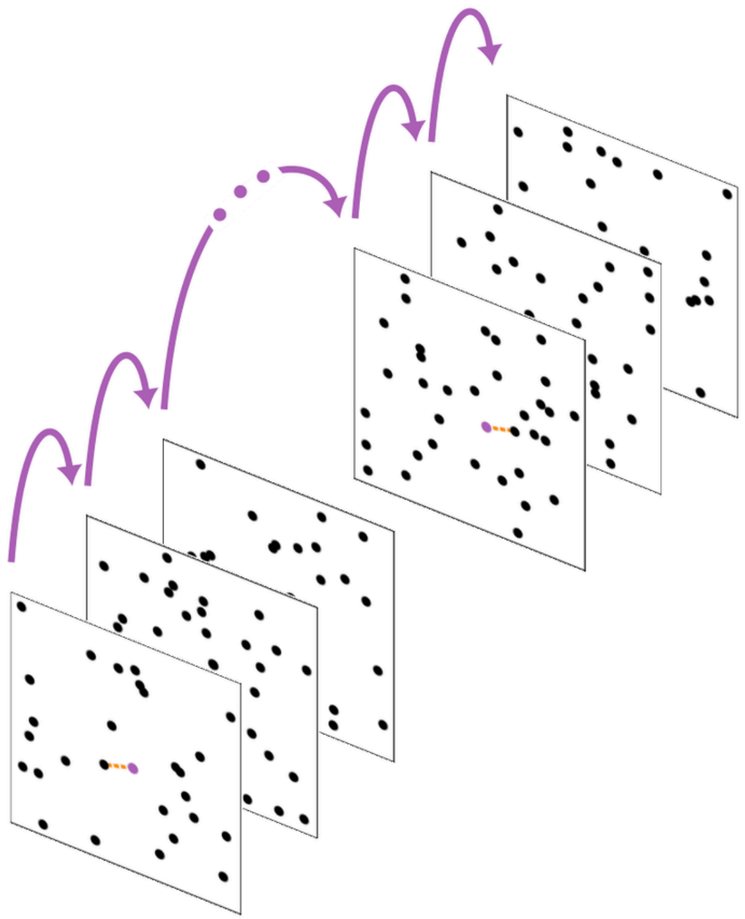
- We create our chains by performing successive nearest neighbour searches
- An example search is shown
- A nearest neighbour search finds the closest galaxy in the next snapshot to a selected target

Galaxy morphology



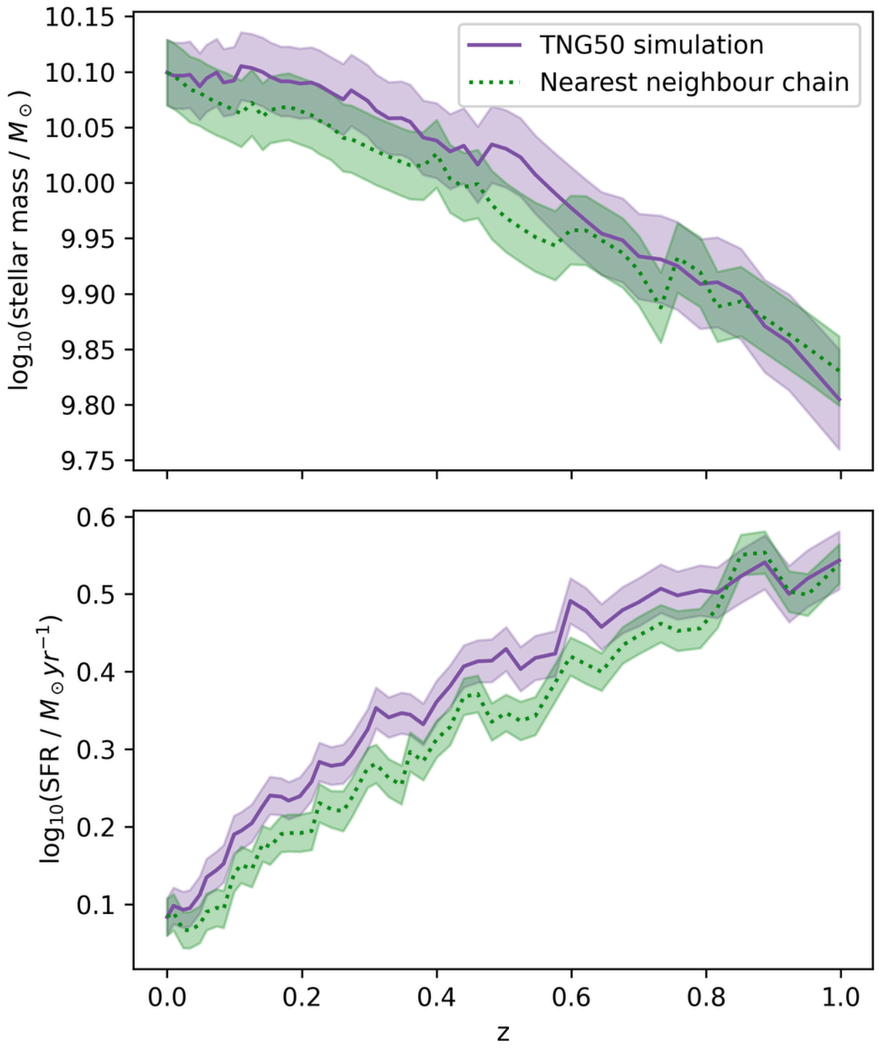
- Here are the results of some searches, representing one step of the galaxy chains
- The searches find galaxies that share similar morphologies

Creating galaxy chains



- We combine individual nearest neighbour searches through each snapshot, creating a new galaxy chain
- We then compare this new chain to the actual galaxy chain from the TNG50 simulation

Galaxy properties



- We have the evolution of stellar mass and star formation rate
- The galaxy chains we create through our searches produce evolutionary histories that are similar to the actual evolution of the galaxy chains from the TNG50 simulation
- We plan to use this method on real observational data

Conclusions – Thank you for listening

- We train an autoencoder on galaxy images from IllustrisTNG-50
- We use the representations space of the autoencoder to perform nearest neighbour searches on each image
- These individual searches are combined to create a new galaxy chain
- Our tests show that these new chains are similar to the real evolution of the simulated galaxies
- For future work, we will use this method on real observational data in an attempt to uncover how galaxies evolve

email: l.h.butterworth@herts.ac.uk