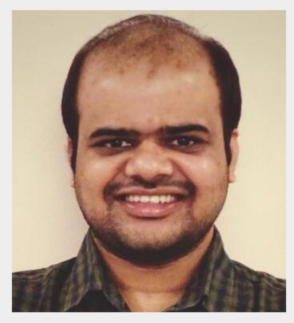
Estimating physical properties of galaxies using deep learning

Omkar Bait[1], Yogesh Wadadekar[2], Harsh Grover[3], Shraddha Surana[4]

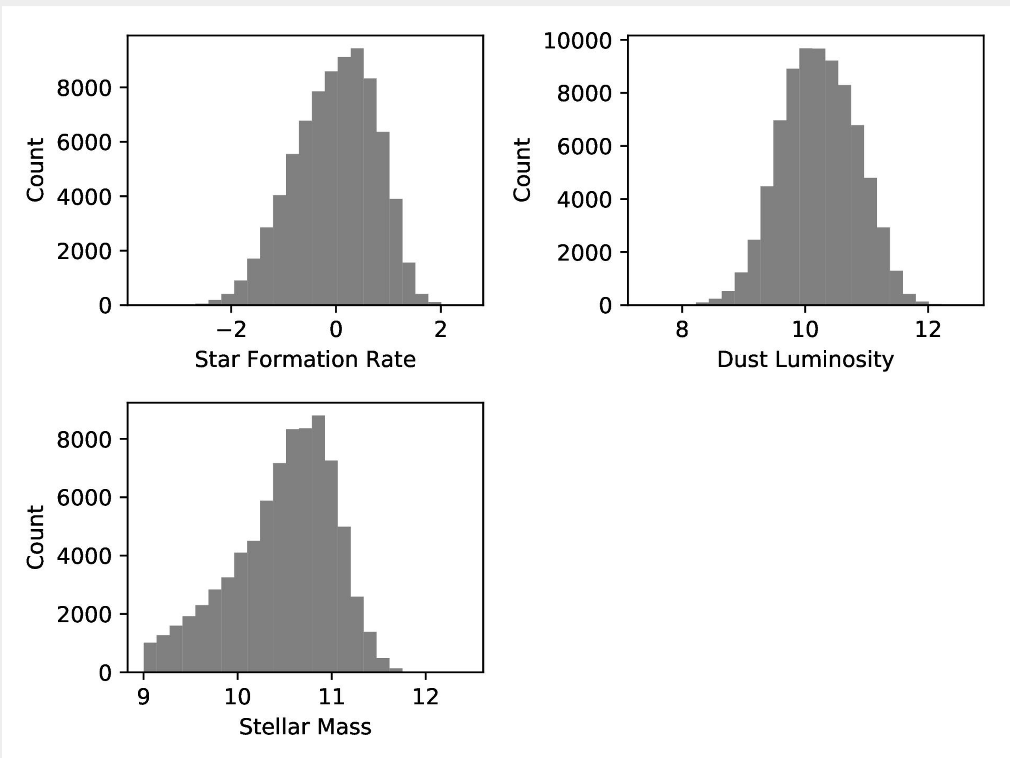
[1]Observatoire de Geneve, Universite de Geneve, Ch. Pegasi 51, 1290 Versoix, Switzerland, [2]National Centre for Radio Astrophysics, Pune-411007, India, [3]Birla Institute of Technology and Science, Pilani, Rajasthan, 333031, India, [4]ThoughtWorks Technologies, Yerwada, Pune, 411006, India



Motivation

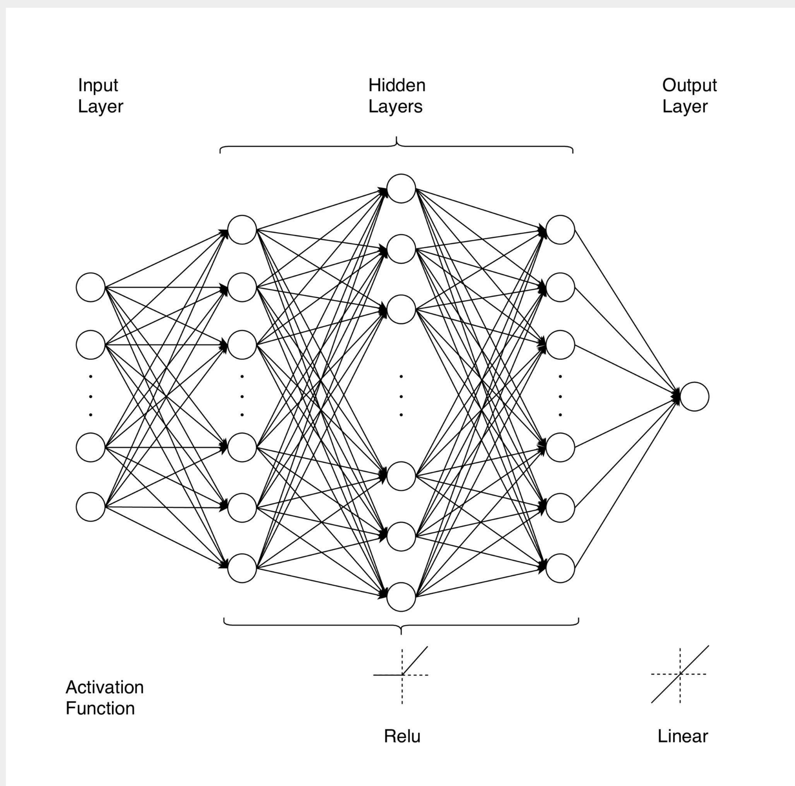
- Estimating the star formation and morphological properties of galaxies is an important exercise in galaxy evolution studies.
- Traditional approaches require a significant amount of human intervention and are computationally expensive.
- Deep learning models have the potential to overcome these limitations.
- We present two deep learning models to predict: 1) star-formation related properties and 2) r-band bulge-to-total luminosity ratio (B/T) of nearby galaxies.

Model 1: Estimation of star formation properties



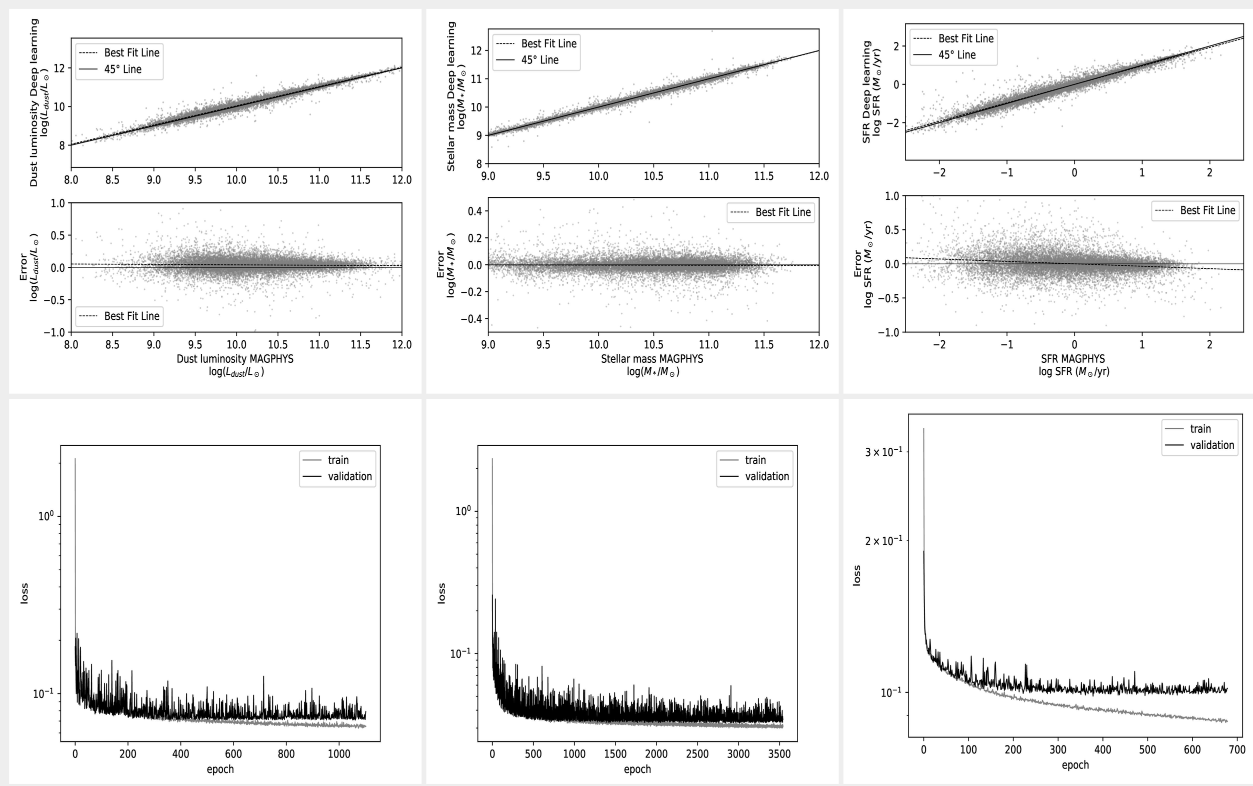
Training data

- DATA: GAMA (Galaxy And Mass Assembly; Driver et al. 2011) survey.
- GAMA is a spectroscopic survey of ~300,000 galaxies down to r < 19.8 mag over ~286 deg^2 , which was carried out using the AAOmega multi-object spectrograph on the Anglo-Australian Telescope (AAT).
- Fluxes in 21 bands from far-UV to far-IR, redshift and SED fits from MAGPHYS ~ 120,000 galaxies.
- We selected all galaxies with redshift (z) ≤ 0.5 and stellar mass ≥ 10^9 M$_\odot$.
- Galaxies are rejected if:
  - it occupies a sparsely populated region of the parameter space
  - it has flux measurements that are missing or unreliable (SNR < 3 in less than 6 bands)
- Input data for training: sample of ~80K galaxies with upto 21 band flux data and redshift information.
- Output: stellar mass, star formation rate and dust luminosity.



Model Architecture

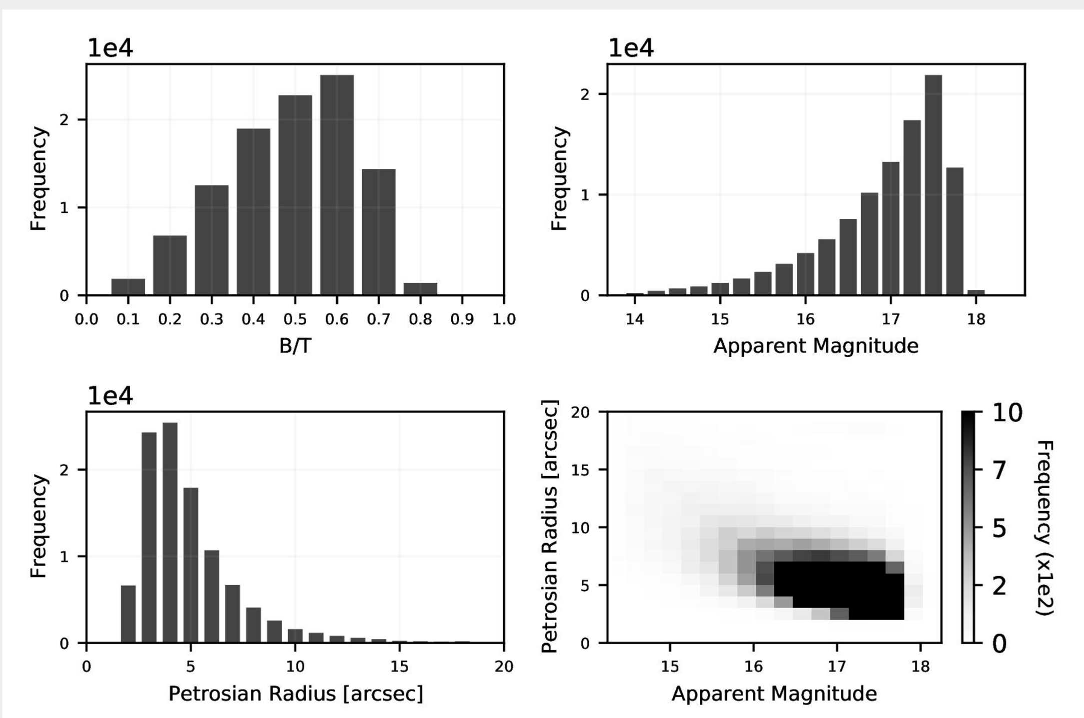
- We experimented with several machine learning models: random forests, support vector machines and deep neural networks.
- After some trial we found that deep neural networks with 3-5 hidden layers (with several hundreds of nodes in each layer) gave the best results.
- The nodes were activated using a Rectified Linear Unit (ReLU) function.
- We used mean absolute error (MAE) as a loss function which was optimized using an Adam optimizer.
- The model was trained for several thousands of epochs so as to minimize the loss function.
- We found best results when we trained three different models for each of the three output parameters: SFR, stellar mass and dust luminosity.



Results

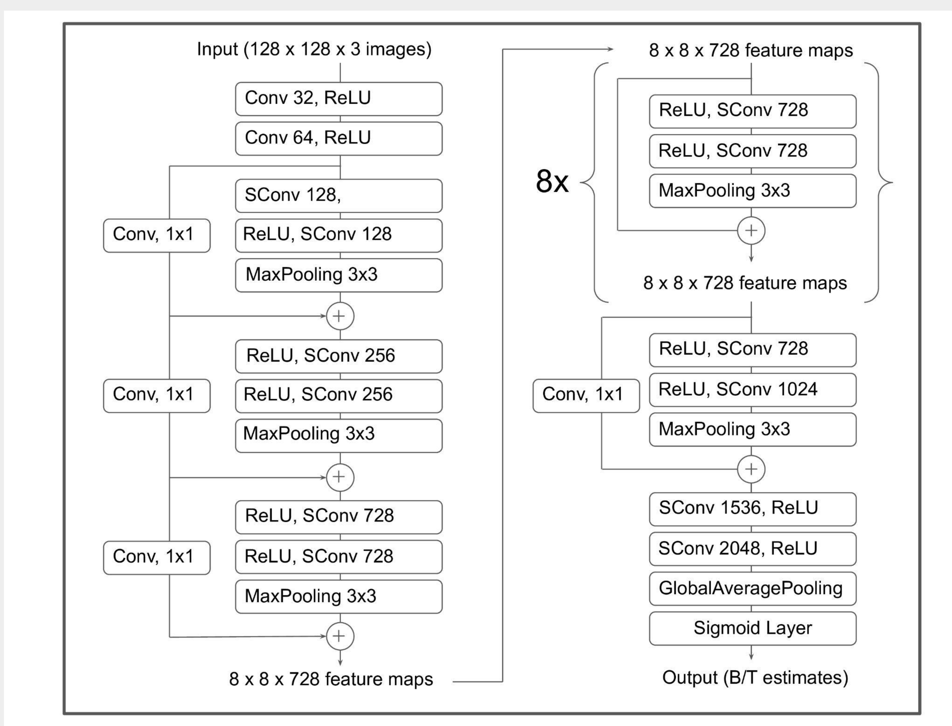
- The uppermost panels show the relation between the true (from MAGPHYS) and predicted value (from our model) for the three output parameters: stellar mass, SFR and dust luminosity respectively for our test sample.
- The middle panels show the difference between the predicted and model parameters.
- We overall get good estimation of stellar mass, dust luminosity but the SFR shows more scatter in the prediction.
- The SFR prediction also shows a small overprediction at low SFRs.
- The lowermost panels show the comparison between the loss function of training and validation data.
- The loss function for stellar mass and dust luminosity compare very well, however for SFR there seems to be a problem of overfitting.
- For SFR estimation we have thus used early stopping so as to avoid overfitting.

Model 2: Estimation of bulge-to-total ratio of galaxies



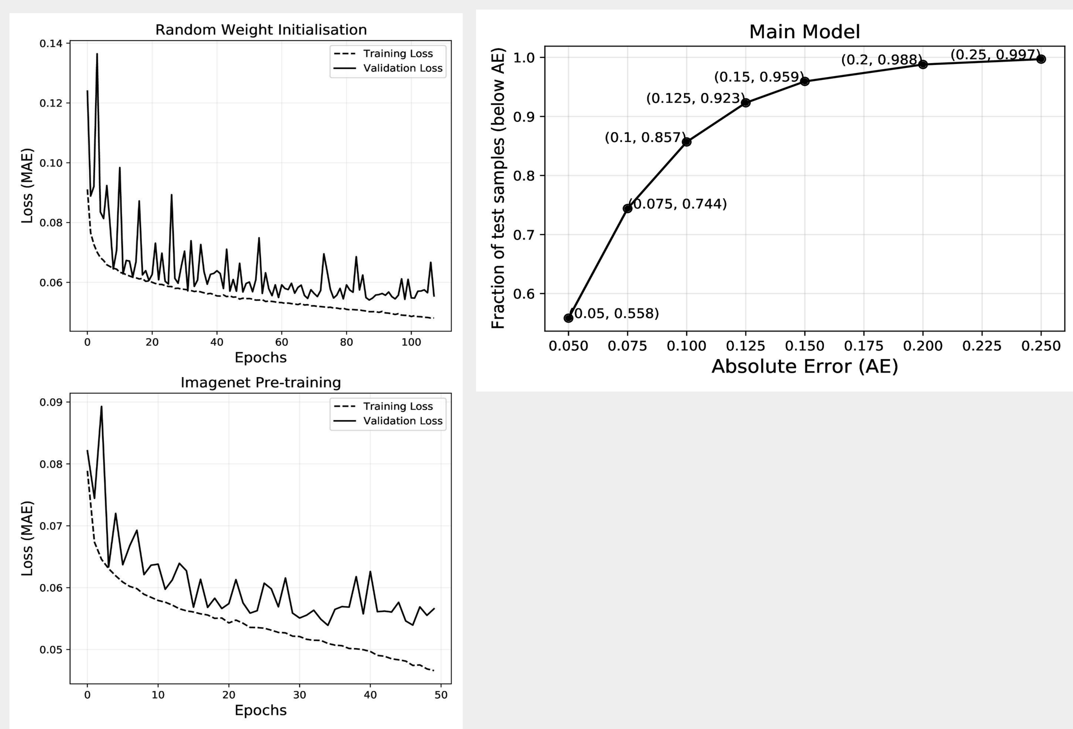
The distribution of B/T, apparent magnitude and petrosian radius

- We used only the SDSS g,r,i color composite jpeg cutout images to train our model.
- Our parent sample is from the Meert et al. (2015) catalogue which contains a 2D photometric decompositions of about ~670,000 spectroscopically selected low-z galaxies from the Sloan Digital Sky Survey (SDSS) Data Release.
- The fits in the Meert catalogue are performed on r-band SDSS images using the GALFIT code.
- We select galaxies having only "good fits" with the two-component model from the Meert catalogue.
- Our final sample contains a total of 103,757 galaxies which we split in 60-20-20 ratio for training, validation and testing.



Xception architecture (Chollet 2016) used in our work. SConv represents separable convolution. Number of filters are specified with every Conv and SConv operations.

- We trained our model using a convolutional neural network (CNNs) using the Xception architecture.
- We experimented several other architectures e.g., InceptionV3, Resnet50, but Xception gave the best results.
- Loss function: Mean absolute error (MAE)
- Optimizer: Adam optimizer.
- See Grover et al. 2021 for more details.



Left panel: Learning curves for the Xception model trained using random weight initialization and Imagenet pre-training. The dashed curves show changes in MAE loss for the training set with each epoch. The solid curve shows the same for validation set.

Right panel: Performance as a function of absolute error cut-off. For any point (x, y) on this graph, x is the absolute error and y represent fraction of test samples with absolute error less than x. On the entire test set of 20 000 images, we achieved accuracy of 85.7 percent with the main model. Accuracy is defined as the fraction of test samples with absolute error less than 0.1.

Conclusion

- We predict the various star formation related properties derived via traditional SED fitting using a deep neural network model.
- We predict the r-band B/T ratio of galaxies using an end-to-end deep learning model with no manual intervention and high accuracy for most of our sample.
- Both these models give a tremendous improvement (upto three orders of magnitude) in the computational speeds compared to traditional methods.
- Our methods have the potential to save tremendous amount of time, computational and financial resources, along with reduced human effort in predicting an important parameter that characterizes galaxies.
- This will be particularly important in the era of large scale era of next-generation sky surveys such as the Legacy Survey of Space and Time (LSST) and the Euclid sky survey.

Selected References:

- Grover, H., Bait, O., Wadadekar, Y., et al. 2021, MNRAS, 506, 3313
- Surana, S., Wadadekar, Y., Bait, O., et al. 2020, MNRAS, 493, 4808