Galaxy Zoo - The Galaxy Challenge

Classify the morphologies of distant galaxies in our Universe

$16,000 · 326 teams · 3 years ago

Training Data

6 files

- all_ones_benchmark.z...
- all_zeros_benchmark....
- central_pixel_benchm...
- images_test_rev1.zip
- images_training_rev1...
- training_solutions_r...

all_ones_benchmark.zip

Download File

File size 265.31 KB

Data Introduction

- **images_training:** JPG images of 61578 galaxies. Files are named according to their GalaxyId.
- **solutions_training:** Probability distributions for the classifications for each of the training images.
- **images_test:** JPG images of 79975 galaxies. Files are name according to their GalaxyId. You will provide probabilities for each of these images.
- **all_ones_benchmark:** Sample submission file corresponding to the All Ones Benchmark
- **all_zeros_benchmark:** Sample submission file corresponding to the All Zeros Benchmark
- **central_pixel_benchmark:** Simple benchmark that clusters training galaxies according to the color in the center of the image and then assigns the associated probability values to like-colored images in the test set.

The first column in each solution is labeled GalaxyID; this is a randomly-generated ID that only allows you to match the probability distributions with the images. The next 37 columns are all floating point numbers between 0 and 1 inclusive. These represent the morphology (or shape) of the galaxy in 37 different categories as identified by crowdsourced volunteer classifications as part of the Galaxy Zoo 2 project. These morphologies are related to probabilities for each category; a high number (close to 1) indicates that many users identified this morphology category for the galaxy with a high level of confidence. Low numbers for a category (close to 0) indicate the feature is likely not present.

Visit the Galaxy Zoo Decision Tree page for a detailed description of the data.