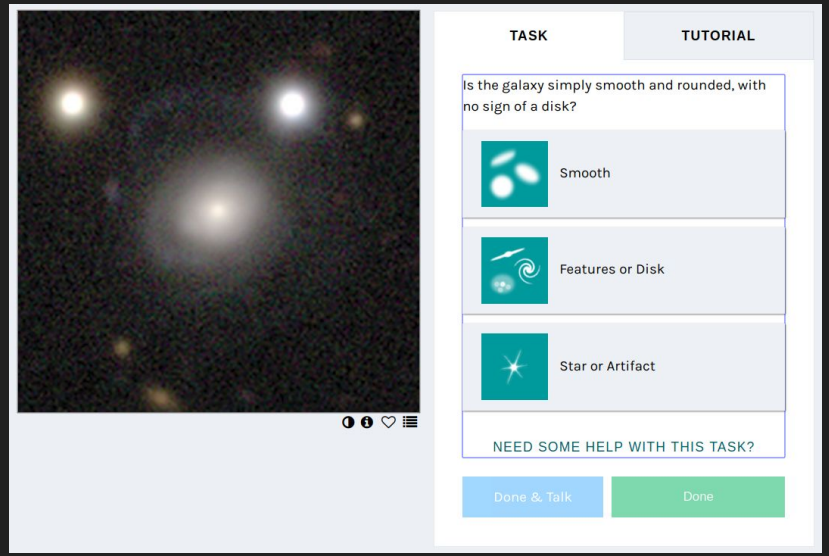

Astronomy Final
Presentation

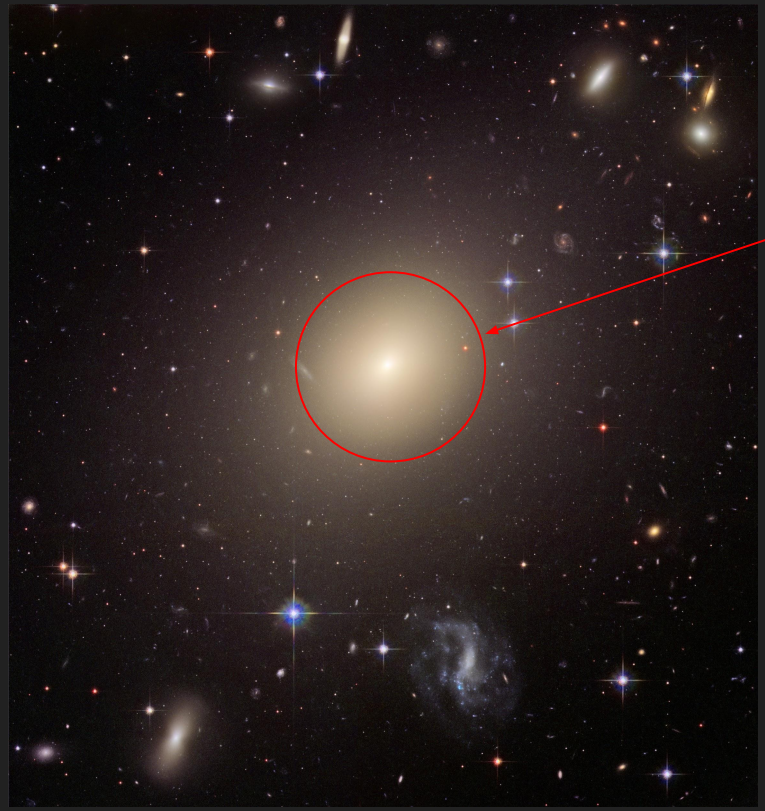
Hello!

Hello. Today we are going to talk about types of galaxies and their classifications.



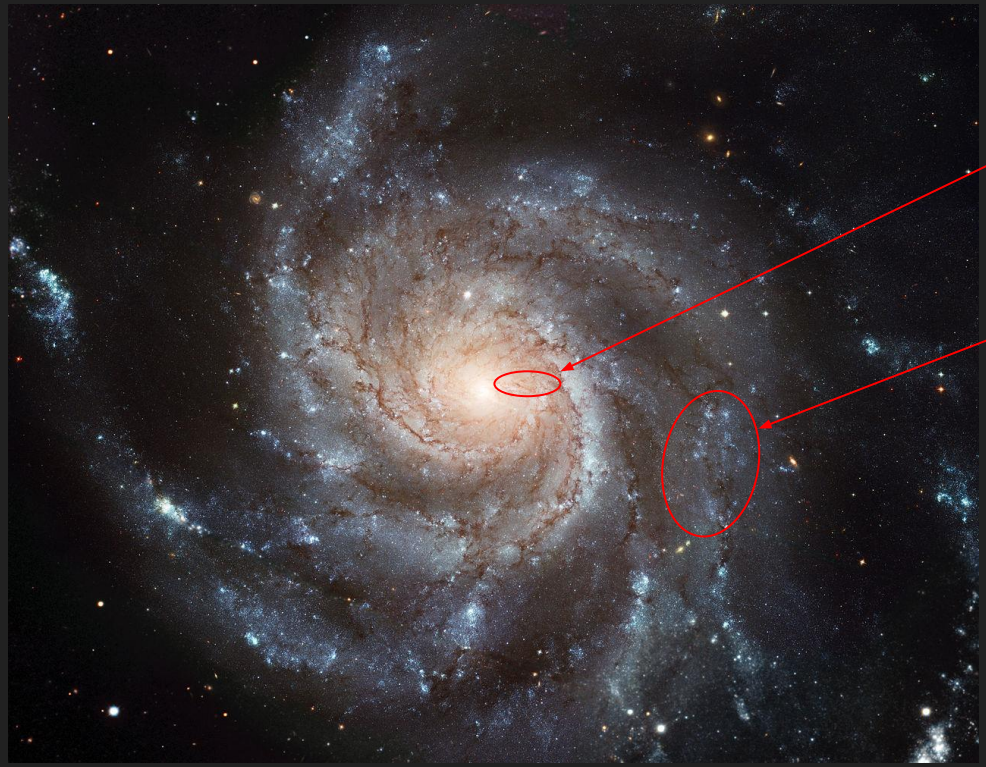
There are three types of galaxies out there: elliptical, spiral, and irregular.

Types of Galaxies



The elliptical galaxies have no disc, no spiral arms, and almost no gas and dust.

Types of Galaxies



The spiral galaxies are disc-shaped, usually have spiral arms and contain gas and dust.

Types of Galaxies

The irregular galaxies are generally shapeless and tend to be rich in gas and dust.

You may ask, why did not we have the third category in Galaxy Zoo questions?

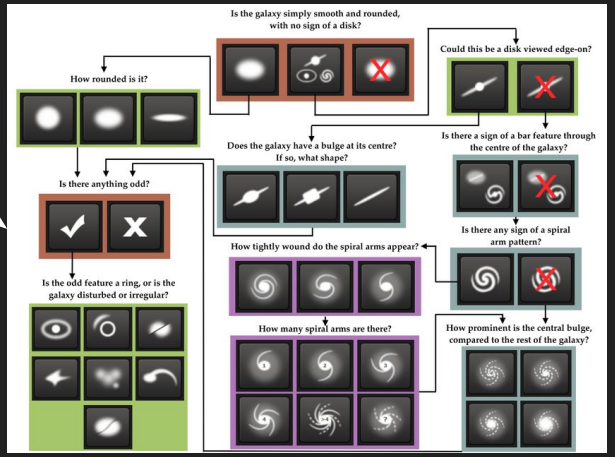
Now for some info about the Galaxy Zoo Project...

It was created more than a decade ago by astronomers. Its main purpose is to get volunteers to help scientists explore galaxies near and far, and there are at least 300 billions of them in the observable Universe.

Kaggle Competition

Almost five years ago there was a competition on the website called Kaggle.com. Its topic was classification of galaxies. Kaggle.com is the website for data scientists - people who analyze the data and create models to make predictions and gain valuable insights.

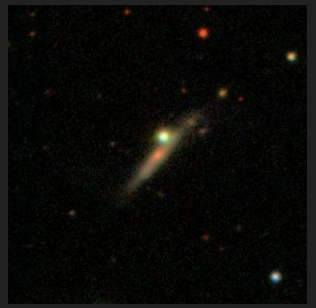
It would say whether it is elliptical or spiral, how many arms it has, is there a bar, etc.

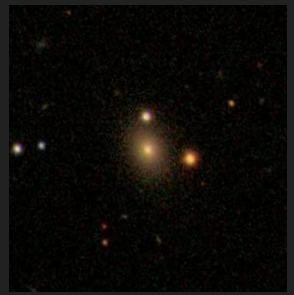


Data Pre-processing and augmentation

In this case we are working

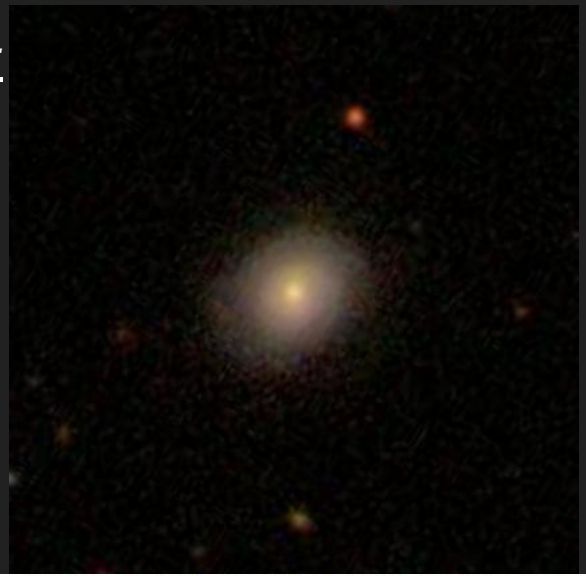
With galaxies...





Data Pre-processing and augmentation

One Example:



There are actually 38 parameters. But need be we only look at the first three.

Class 1.1 : The ratio of people who think it's smooth.
Class 1.2 : The ratio of people who think it's disc.
Class 1.3 : Is artifact, etc.

	GalaxyID	Class1.1	Class1.2	Class1.3	Class2.1	Class2.2	Class3.1	Class3.2	Class4.1	Class4.2	...
0	100008	0.383147	0.616853	0.0	0.0	0.616853	0.038452	0.578401	0.418398	0.198455	...

Data Pre-processing and augmentation

	GalaxyID	Smooth
0	100008	0

0 in “Smooth” column means people think it is disc galaxy.

If it’s 1 then people think it’s a smooth galaxy

Before we proceed...

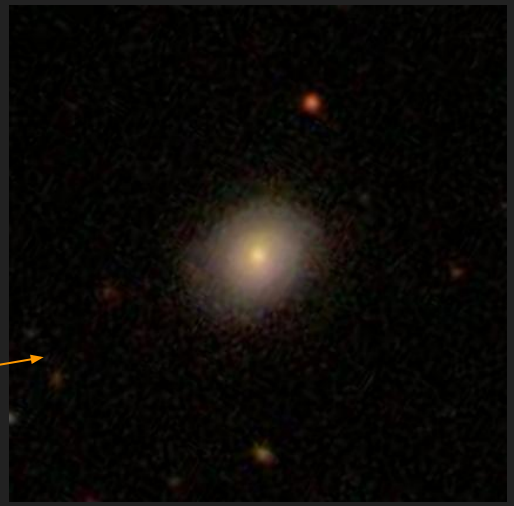
We had to get rid of images that were NOT depicting galaxies.

Like these!

Data Pre-processing and augmentation

We needed to make adjustments to the images we had.

Like this one!



===>

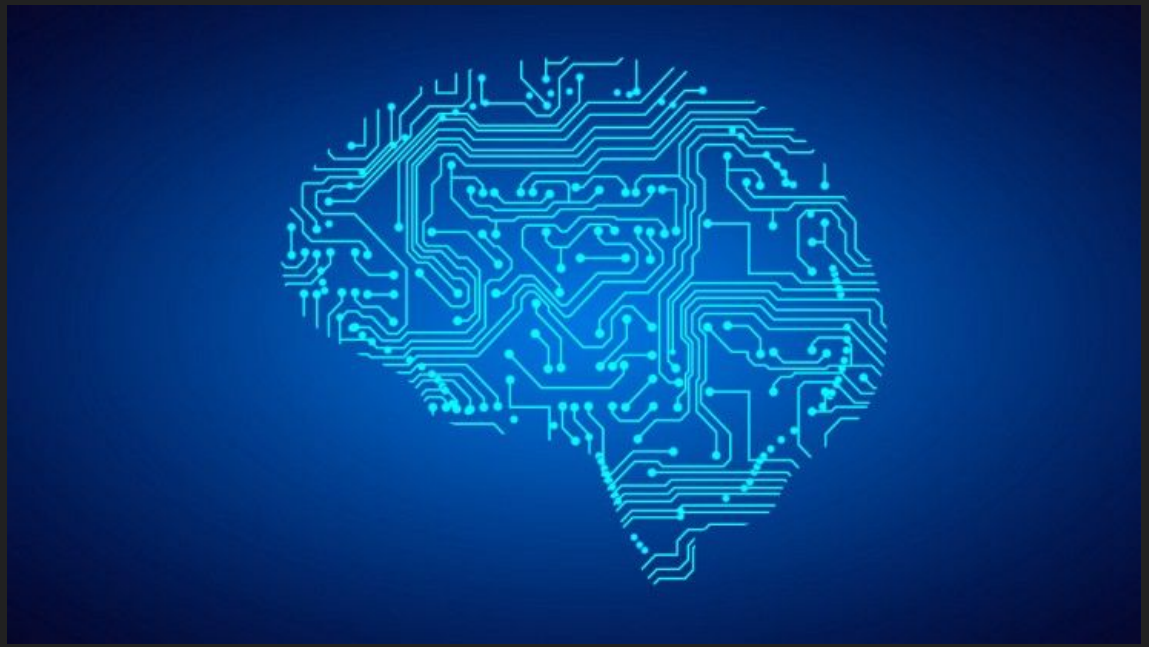
Cropped 200x200 from 424x424 and reduced resolution to 128x128

Research Method

Review Possible Methods

Machine Learning

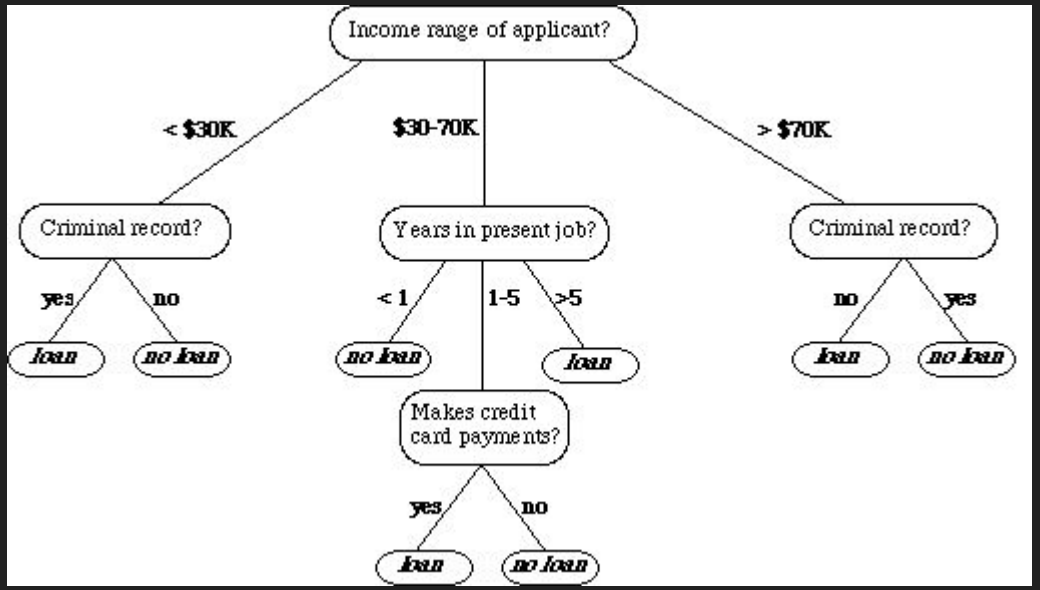
Gradient Boosting



Decision Trees

Simple Branching

How to Apply to Galaxies?

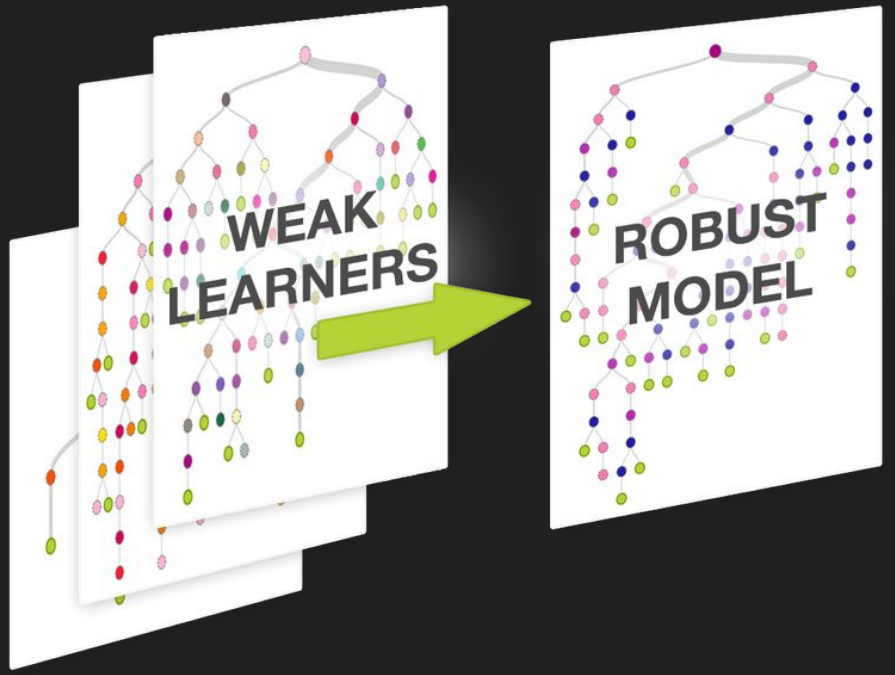


Gradient Boosting Algorithm

Comparing Results

Adjusting Model

Repeat!



Results

80% precision on 40000 images

Training took a night!

20000 images were used to train the model

Code is available github.com/Temurson/GalaxyClassification

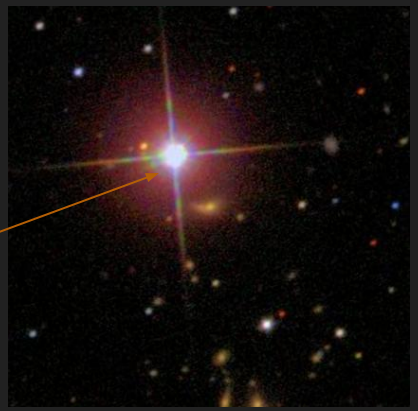
One thing to understand about computers...

Is that computers *cannot* identify strange things.

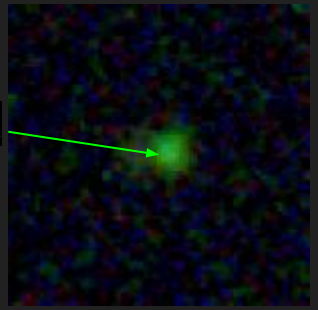
Examples:

The program says this is a smooth galaxy but it's really nothing there.

Program says its a smooth galaxy but it's really a star.

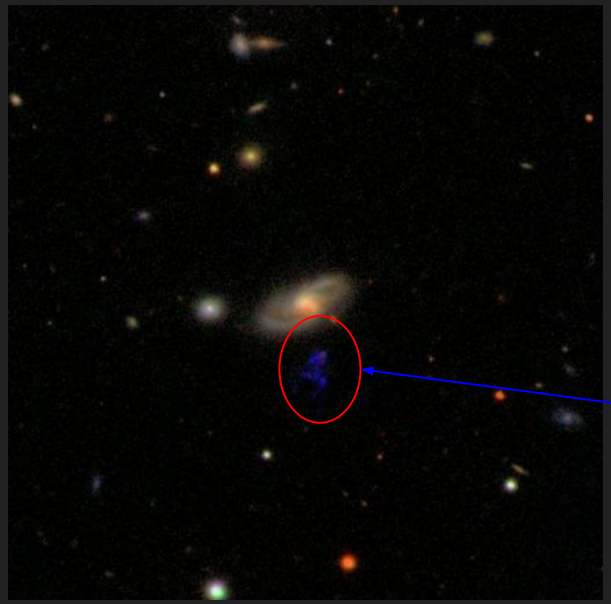


More *Examples :*

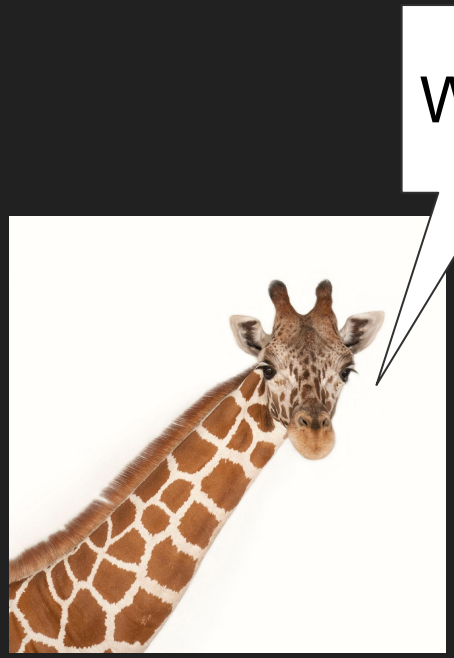


When discovered by the public they were thoroughly analyzed.

These samples are really valuable to the astronomers.

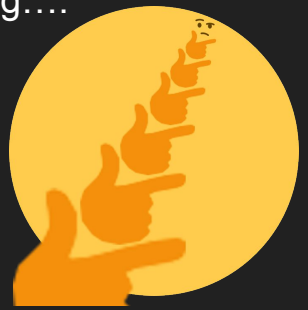


Now for the *weird* part...



The model actually says that this image is a ***smooth*** galaxy….

……the program may be onto something….



Bibliography

- Willett, K. W., Lintott, C. J., Bamford, S. P., Masters, K. L., Simmons, B. D., Casteels, K. R., . . . Thomas, D. (2013, 09). Galaxy Zoo 2: Detailed morphological classifications for 304 122 galaxies from the Sloan Digital Sky Survey. Monthly Notices of the Royal Astronomical Society, 435(4), 2835-2860. doi:10.1093/mnras/stt1458
- Smethurst, R. (2014, July 11). The SEVEN wonders of Galaxy Zoo. Retrieved from https://blog.galaxyzoo.org/2014/07/11/the-seven-wonders-of-galaxy-zoo/
- (2018). Screenshot showing classification process. [Screenshot]. Retrieved from https://www.zooniverse.org/projects/zookeeper/galaxy-zoo/
- Miller, G. (Illustrator). (2012). Official Zooniverse avatar for Galaxy Zoo. [Logotype]. Retrieved from https://commons.wikimedia.org/wiki/File:Galaxyzoo.jpg
- Galaxy Zoo Project. (2013). Galaxy Zoo - The Galaxy Challenge [images of galaxies and data of classifications]. Retrieved from https://www.kaggle.com/c/galaxy-zoo-the-galaxy-challenge
- NASA, ESA, The Hubble Heritage Team, Blakeslee, J. (2005 and 2006). Hubble Illuminates Cluster of Diverse Galaxies (Abell S0740). [Processed Telescope Photograph]. Retrieved from https://en.wikipedia.org/wiki/File:Abell_S740,_cropped_to_ESO_325-G004.jpg
- European Space Agency & NASA. (2006). Galaxy Messier 101. [Processed Telescope Photograph]. Retrieved from https://en.wikipedia.org/wiki/File:M101_hires_STScI-PRC2006-10a.jpg
- NASA, ESA, and The Hubble Heritage Team. (2003). The irregular galaxy NGC1427A. [Processed Telescope Photograph]. Retrieved from https://en.wikipedia.org/wiki/File:Irregular_galaxy_NGC_1427A_(captured_by_the_Hubble_Space_Telescope).jpg

Bibliography (continuation)

- Machine learning artist interpretation. [Image]. Retrieved from https://cdn-images-1.medium.com/max/1000/1*60gs-SFYyooZZBxatuoNJw.jpeg
- Plapinger, Thomas. (Artist). (2017). Decision tree for credit consideration. [Diagram]. Retrieved from https://towardsdatascience.com/what-is-a-decision-tree-22975f00f3e1
- Illustration of gradient boosting. [Image]. Retrieved from https://littleml.files.wordpress.com/2016/07/gradient_boosting-ok.png?w=768&h=647