

GAX-Kepler: Identificador de morfología de galaxias mediante el uso de redes neuronales convolucionales

Brian Gonzalez, Dylan Jara

Universidad de Santiago de Chile

17 de diciembre de 2025



1 Introducción

2 Descripcion de datos de estudio

3 Hipótesis

4 Esquema Metodologico

5 Metodología

6 Ideas Futuras

7 Refs

8 Conclusiones

1 Introducción

- Introducción al problema
- Justificación
- Estado del arte

2 Descripción de datos de estudio

3 Hipótesis

4 Esquema Metodológico

5 Metodología

6 Ideas Futuras

1 Introducción

Introducción al problema

Justificación

Estado del arte

2 Descripción de datos de estudio

3 Hipótesis

4 Esquema Metodológico

5 Metodología

6 Ideas Futuras

Dataset Galaxy Zoo

- Una galaxia es un entorno en el cual las estrellas nacen y mueren, y las galaxias lejanas son faros luminosos que nos permiten explorar el universo lejano. (SB01, B01)
 - En este trabajo, usamos Galaxy Zoo (Lintott et al., 2008, 2011; Willett et al., 2013) para la clasificación y validación de nuestros modelos de clasificación.

Dataset Galaxy Zoo

- Una galaxia es un entorno en el cual las estrellas nacen y mueren, y las galaxias lejanas son faros luminosos que nos permiten explorar el universo lejano. (SB01, B01)
- En este trabajo, usamos Galaxy Zoo (Lintott et al., 2008, 2011; Willett et al., 2013) para la clasificación y validación de nuestros modelos de clasificación.
- El proyecto original Galaxy Zoo fue lanzado en julio de 2007 y obtuvo más de 100,000 voluntarios en sus primeros diez días.

Dataset Galaxy Zoo

- Una galaxia es un entorno en el cual las estrellas nacen y mueren, y las galaxias lejanas son faros luminosos que nos permiten explorar el universo lejano. (SB01, B01)
- En este trabajo, usamos Galaxy Zoo (Lintott et al., 2008, 2011; Willett et al., 2013) para la clasificación y validación de nuestros modelos de clasificación.
- El proyecto original Galaxy Zoo fue lanzado en julio de 2007 y obtuvo más de 100,000 voluntarios en sus primeros diez días.
- Los voluntarios clasificaron imágenes del Sloan Digital Sky Survey (SDSS) de galaxias en seis categorías: elíptica, espiral en sentido horario, espiral en sentido antihorario, edge-on , estrella/no se, o fusión de galaxias.

1 Introducción

Introducción al problema

Justificación

Estado del arte

2 Descripción de datos de estudio

3 Hipótesis

4 Esquema Metodológico

5 Metodología

6 Ideas Futuras

¿De qué nos sirve?

- This growing catalog of exoplanets has challenged the uniqueness of our solar system and informed theories of planet formation (Johnson, 2009).
- Basicamente facilita la exploracion espacial. Y esto ayuda a entender mejor el universo y el origen de la vida

1 Introducción

Introducción al problema

Justificación

Estado del arte

2 Descripción de datos de estudio

3 Hipótesis

4 Esquema Metodológico

5 Metodología

6 Ideas Futuras

1 Introducción

2 Descripcion de datos de estudio

Tipos de galaxias

Analisis preliminar y exploratorio de los datos

3 Hipótesis

4 Esquema Metodologico

5 Metodología

6 Ideas Futuras

7 Refs

1 Introducción

2 Descripcion de datos de estudio

Tipos de galaxias

Analisis preliminar y exploratorio de los datos

3 Hipótesis

4 Esquema Metodologico

5 Metodología

6 Ideas Futuras

7 Refs

Definiciones

Dominio

Debe ir categorización de galaxias con galaxy zoo

Esquema de Relación

Detección usando IA

Relación

Clasificación multicategoría pero con binary-CE usando redes neuronales

Ejemplo de una imagen

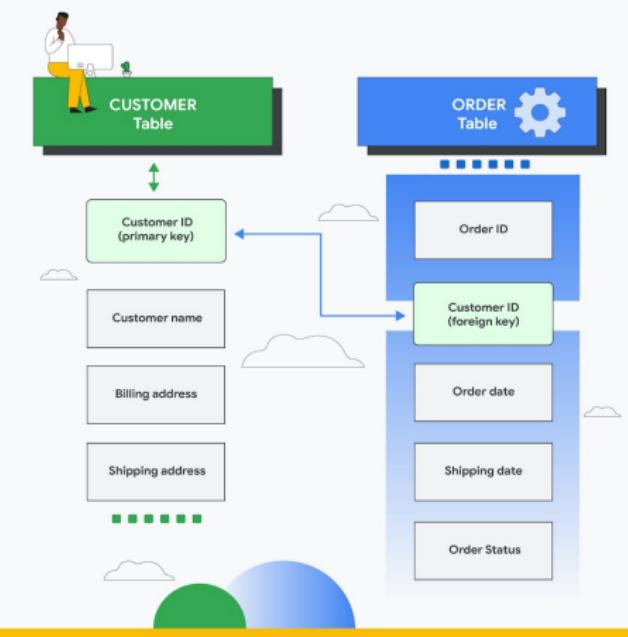


Figura 1: Tablas de Clientes y Pedidos (Google Cloud)

1 Introducción

2 Descripcion de datos de estudio

Tipos de galaxias

Analisis preliminar y exploratorio de los datos

3 Hipótesis

4 Esquema Metodologico

5 Metodología

6 Ideas Futuras

7 Refs

- Se uso data augmentation para rotar las imágenes

1 Introducción

2 Descripcion de datos de estudio

3 Hipótesis

¿Qué queremos lograr?

4 Esquema Metodologico

5 Metodología

6 Ideas Futuras

7 Refs

1 Introducción

2 Descripcion de datos de estudio

3 Hipótesis

¿Qué queremos lograr?

4 Esquema Metodologico

5 Metodología

6 Ideas Futuras

7 Refs

aaa

1 Introducción

2 Descripcion de datos de estudio

3 Hipótesis

4 Esquema Metodologico

Preprocesamiento de datos

5 Metodología

6 Ideas Futuras

7 Refs

1 Introducción

2 Descripcion de datos de estudio

3 Hipótesis

4 Esquema Metodologico

Preprocesamiento de datos

5 Metodología

6 Ideas Futuras

7 Refs

1 Introducción

2 Descripcion de datos de estudio

3 Hipótesis

4 Esquema Metodologico

5 Metodología

Arquitectura de la red

6 Ideas Futuras

7 Refs

1 Introducción

2 Descripcion de datos de estudio

3 Hipótesis

4 Esquema Metodologico

5 Metodología

Arquitectura de la red

6 Ideas Futuras

7 Refs

aaa

1 Introducción

2 Descripcion de datos de estudio

3 Hipótesis

4 Esquema Metodologico

5 Metodología

6 Ideas Futuras

7 Refs

- Usar GPU para entrenar mas rapido
- Divide y vencerás para distribuir
- Usar la CNN para trapasar a un Decision Tree o Random Forest
- Galaxy Zoo is a citizen science project which provides a visual morphological classification for nearly one million galaxies in its first phase (Galaxy Zoo 1) distinguishing elliptical from spiral galaxies.
-

1 Introducción

2 Descripcion de datos de estudio

3 Hipótesis

4 Esquema Metodologico

5 Metodología

6 Ideas Futuras

7 Refs

8 Conclusiones

- We use scikit-learn (Pedregosa et al., 2011) python library to perform the experiments and procedures reported in this Section
- For a further analysis on the problem with two classes, we use the Receiver Operating Characteristic (ROC) curve and the Area Under the ROC curve (AUC, Bradley, 1997)
- Reducir la cantidad de clases existentes para disminuir la no-linealidad de los datos
- Throughout this work, we use Galaxy Zoo (Lintott et al., 2008, 2011; Willett et al., 2013) classification as supervision and validation (ground truth) to our classification models.
- Several authors (Abraham et al., 1996; Conselice et al., 2000; Conselice, 2003; Lotz et al., 2004) studied and presented results about objective galaxy morphology measures with Concentration, Asymmetry, Smoothness, Gini, and M20 (CASGM system).
- The main purpose of this investigation is to answer the

1 Introducción

2 Descripcion de datos de estudio

3 Hipótesis

4 Esquema Metodologico

5 Metodología

6 Ideas Futuras

7 Refs

8 Conclusiones

- ① Las bases de datos son esenciales en nuestra vida
- ② No son perfectas y necesitan manejo especializado
- ③ Perdurarán hasta mucho tiempo más

References

Silk, J., y Bouwens, R. (2001). The formation of galaxies. *New Astronomy Reviews*, 45(4), 337-350. Descargado de <https://www.sciencedirect.com/science/article/pii/S1387647300001640> (Understanding the Universe at the close of the 20th century) doi: [https://doi.org/10.1016/S1387-6473\(00\)00164-0](https://doi.org/10.1016/S1387-6473(00)00164-0)