

Reporte 03

Jairo Andres Saavedra Alfonso

Universidad de los Andes

1 Objetivos

- Observar el balance de clases en los espectros identificados.
- Preprocesar los datos reconocidos.

2 Desbalance de clases

Para observar el desbalance de las clases *STARS*, *GALAXY*, *QSO* y *QSO_BAL* se creo un nuevo *database* con los 537.677 objetos identificados con su respectivo espectro. Se observ, independiente de las demas caractersticas, un desbalance entre la cantidad de objetos reconocidos como estrellas y galaxias de un orden de magnitud como se observa en la siguiente tabla.

Objeto	Cantidad
Estrellas	207.905
Galaxias	20.699
QSO	270.534
QSO_BAL	29.652
NC	8.887

Table 1. Cantidad de objetos en el nuevo *dataset*

Se obseva tambien un desbalance entre la clase QSO y QSO BAL. Por otro lado, se identific 8.887 objetos sin clasificacin original.

3 Preprocesamiento de los datos

Objetos

Dentro del proceso de preprocesamiento de los datos se contemplo eliminar los datos que no aportan a la clasificacin de nuestros espectros. Para nuestro nuevo *database* se eliminaron 8.887 objetos sin clasificacin para evitar clasificacin erronea. De igual forma se elimin, observando la columna de confianza en redshift, todos los objetos con una confianza de cero.

Finalmente se obtuvo 528.696 objetos entre los cuales están estrellas, galaxias, QSO y QSO BAL.

Espectro

Para la nueva *database* de espectros con objetos reconocidos, se ajustaron los flujos logaritmicamente en el mismo rango de longitud de onda de 443 nm a 886 nm. Este *database* contiene 528.696 espectros renormalizados.

4 Tareas pendientes

- Aprender a usar Pytorch para implementar una red neuronal básica con los espectros reconocidos.

5 Repositorio Github

En este repositorio se pueden ver todos los reportes semanales y el notebook.

<https://github.com/MrX1997/Reportes-Proyecto-de-Monografia>

References

1. Busca G., Balland C., (2018).: QuasarNET: Human-level spectral classification and redshifting with Deep Neuronal Networks. Retrieve from: <https://arxiv.org/abs/1808.09955>
2. Paris, I., et al., (2017), Astron. Astrophys...: The Sloan Digital Sky Survey Quasar Catalog: Twelfth data release. , 597, A79