

Classification of Stars, Galaxies and Quasars

Enio Linhares Junior

5/19/2019

Contents

0.1	1 Introduction	1
0.2	2 Downloading Installing and Starting R	1

0.1 1 Introduction

0.1.1 1.1 Content

The Sloan Digital Sky Survey offers public data of space observations and the task here is to build a model that is able to predict the different classes of objects (Stars, Galaxies and Quasars) based on the data acquired through the scientific equipment. The data consists of 10,000 observations of space taken by the SDSS. Every observation is described by 17 feature columns and 1 class column which identifies it to be either a star, galaxy or quasar.

Our model achieved an **accuracy** of 0.9915019

0.1.2 1.2 Feature Description

The table results from a query which joins two tables (actuacly views): “PhotoObj” which contains photometric data and “SpecObj” which contains spectral data.

During our data exploratory analysis we will be explaining the features as they appear.

The data released by the SDSS is under public domain. Its taken from the current data release RD14.

0.2 2 Downloading Installing and Starting R

0.2.1 2.1 Installing the DataSet

The url can be found here: <https://www.kaggle.com/lucidlenn/sloan-digital-sky-survey/version/1>