A Web Application For Photometric Redshift Evaluation


1Dipartimento di Scienze Fisiche, Universita Federico II di Napoli
2Istituto Nazionale di Astrofisica – OACT
3California Institute of Technology

Abstract

In the era of massive astronomical databases, efficient identification of candidate quasars and the reconstruction of their three-dimensional position in the Universe is a key requirement for constraining some of the main issues regarding the formation and evolution of QSOs. A method for the accurate determination of the photometric redshifts of QSOs based on multivariated absolute photometry and a combination of data mining techniques will be discussed. This procedure, specifically suited for accommodating the candidate selection method discussed in (D'Abrusco et al. 2008), uses specific tools developed under the EuroVO and NVO frameworks for data gathering, pre-processing and mining, while relying on the scaling capabilities of the computing grid. This method allows us to obtain photometric redshifts with an increased accuracy (up to 3%) with respect to the literature.

1. A method to evaluate distances when spectroscopic estimates become impossible due to either poor signal-to-noise ratio or to instrumental systematics, or to the fact that the sources under study are beyond the spectroscopic limit;

2. An economical way to obtain, at a relatively low price in terms of observing and computing time, redshift estimates for large samples of sources.

Photometric Redshifts

Clustering is an unsupervised method for partitioning the parameter space (quantities, colors, space) into different regions according to a given definition of distance. In our approach clustering is used to enhance the performance of the photometric estimation. Since there is no preferable criterion to determine when clustering is better, we use a feedback approach, which is described below.

Exploiting the Virtual Observatory and GRID Computing Capabilities

The Artificial Neural Networks Training is a computing intensive task when the number of training examples is large as it is in astronomical Data Mining. In our approach, we have two different optimization problems: the ANN training process itself and the optimization of the input networks. The latter refers to the case when clusters are already defined.

The DAME/VONeural Framework makes extensive use of the GRID technology to store the data and perform calculations in a "Cloudy" fashion, by means of a "robot certificate" to allow the safe access to the GRID resources.

As it is shown in the picture above, the spectroscopic "features" shift the photometric bands. In order to avoid degeneracies and to increase accuracy one can federate different tables from different surveys. The Virtual Observatory provides us with all the tools needed to build multivariated absolute photometry on the fly. The final goal of our approach is to create a fully automated application that queries the VO resources for all the photometric information available for a catalogue of extragalactic sources, so to provide the best estimation of the photometric redshift available with all the information at our disposal in the VO infrastructure.

The Web Application

This is the web application of this project. It is a platform independent, service oriented framework for Data Mining. The user can store his data and save his experiments and sessions. The developer will download the code from JAR files. The grid independent: GRID is the main platform, but different "drivers" will be available. Extensible: Data Mining models, complex tasks, middleware drivers and even the client application can all be extended and augmented by means of a consistent, Object Oriented plug-in architecture.