The intriguing life of massive galaxies
Proceedings IAU Symposium No. 295, 2012
D. Thomas, A. Pasquali & I. Ferreras, eds.

doi:10.1017/S1743921313004675

Discriminating Quasars from Stars Based on SDSS and UKIDSS Databases

He Ma[1,2], Yanxia Zhang[1], Yongheng Zhao[1] and Bo Zhang[2]

[1]Key Laboratory of Optical Astronomy, National Astronomical Observatories, Chinese Academy of Sciences, 100012 Beijing, China
[2]Hebei Normal University, 050016 ShiJiaZhuang, Hebei, China

Abstract. In this work, two different algorithms: Linear Discriminant Analysis (LDA) and Support Vector Machines (SVMs) are combined for the classification of unresolved sources from SDSS DR8 and UKIDSS DR8. The experimental result shows that this joint approach is effective for our case.

Keywords. Classification, Astronomical databases: miscellaneous, Catalogs, Methods: Data Analysis, Methods: Statistical

1. Sample

The data are retrieved from SDSS DR8 and UKIDSS DR8. The training sample and test sample for stars comprises 334474 and 61338, respectively. The number of quasars in the two samples is 12061 and 12061, separately.

2. Algorithm and Result

The main idea of LDA is to find a projection in a low dimensional space for different classes to be well separated. The main principle of SVMs is to transfer input vectors into a high-dimensional feature space and find the optimal separating hyperplane in this space. Seven kinds of colors ($u-g$,$g-r$,$u-r$,$g-z$,$u-z$,$g-z$,$Y-K$) are taken as the input of SVMs because of their better color-color diagram separation. Anothr six kinds of LDA output for $u-r$ vs. $g-J$, $u-r$ vs. $r-Y$, $u-z$ vs. $r-z$, $u-Y$ vs. $r-Y$, $u-J$ vs. $u-r$, $g-Y$ vs. $i-K$ are also used as the input of SVMs simultaneously. In addition, r magnitude and i magnitude are used as the input of SVMs. As a result, the precision and completeness of stars is 99.60% and 99.80%, respectively, while that of quasars is 98.93% and 97.99%, respectively.

3. Conclusion

Multi-band data provide a large amount of information about objects. We use multi-band data to classify objects effectively. On the other hand, we combine traditional color-color, magnitudes and outputs of LDA as the input pattern of SVMs to create a new classifier. The classification result is rather satisfactory.

This paper is funded by National Natural Science Foundation of China under grant No.10778724, 11178021 and No.11033001, the Natural Science Foundation of Education Department of Hebei Province under grant No. ZD2010127 and by the Young Researcher Grant of National Astronomical Observatories, Chinese Academy of Sciences.