SDSS MOC4 Asteroids' Color Classifications

Benjamin Montgomery
Department of Computer Science
University of Southern Maine
Portland, Maine 04104
Email: benjamin.montgomery@maine.edu

Daniel Shepard II
IDEXX
Westbrook, Maine 04092
Email: daniel.shepard@maine.edu

Emily Wall
Departments of Physics and Mathematics
University of Southern Maine
Portland, Maine 04104
Email: emily.wall@maine.edu

***Abstract*—The range of taxonomic categories of asteroids present in the Sloan Digital Sky Survey (SDSS) has varied markedly within the span of a decade, bringing to question the particular methods used by machine learning algorithms to analyze and classify this data. We propose to use the Agglomerative, DBSCAN, OPTICS, K-Means++, X-Means, and Hysteresis algorithms to attempt to independently verify the methods used by previous papers.**

I. Introduction

The Sloan Digital Sky Survey is a multi-spectral survey of celestial objects, most notably of asteroids. Multiple efforts [1], [2], [3], [4] have been presented to classify these asteroids based on their observed spectra and colors; the expected number of classes existent in the data ranges from two classes [4] to sixteen [3]. These analyses fail to capture the accepted twenty-six [5] classes found using other asteroid spectral datasets. The variance appears largely due to the kind of classification methods used; original attempts were made using unsupervised clustering algorithms [2], [3], whereas later efforts were supervised learning algorithms that relied on human-specified templates [6] of what each class of asteroid could look like. This brings to question if similar results could be produced by an unsupervised learning algorithm using the modern color correction methods found in recent research.

II. Methods

We propose to use a classification scheme similar to that presented in Carvano's [3] work, namely, classification based on the SDSS' recorded values of the u', g', r', i', and z' filters. Similarly, we adopt the color correction steps made for finding reflectance [3] colors, and for normalizing each asteroid observation by its reflectance color gradients[3]. After this data pre-processing, we expect to use Python's implementation of Tensorflow [7] and the library PyClustering [8] to independently apply Agglomerative, DBSCAN, OPTICS, K-Means++, X-Means, and Hysteresis unsupervised learning algorithms on our normalized color data to classify our data into discrete compositions. These independent attempts at classification will then be aggregated onto a single prediction; the classification associated with each asteroid will be compared to that of the original effort in Autoclass [4] and Carvano's more recent effort [3].

We expect this selection of algorithms to be particularly effective in our dataset, given that subsets of these algorithms have been demonstrated to work well for other astronomical objects recorded through the SDSS' color filters, notably with galaxy classification [9].

Acknowledgment

The authors would like to thank Professor Julie Ziffer of the University of Southern Maine's Department of Physics for her role as an acting consultant in the interpretation of astronomical data.

References

[1] C. R. Chapman, D. Morrison, and B. Zellner, "Surface properties of asteroids - A synthesis of polarimetry, radiometry, and spectrophotometry," , vol. 25, pp. 104–130, May 1975.

[2] V. Zappalà, P. Bendjoya, A. Cellino, P. Farinella, and C. Froeschlé, "Asteroid families: Search of a 12,487-asteroid sample using two different clustering techniques." , vol. 116, pp. 291–314, Aug. 1995.

[3] J. M. Carvano, P. H. Hasselmann, D. Lazzaro, and T. Moth-Diniz, "SDSS-based taxonomic classification and orbital distribution of main belt asteroids," *Astronomy and Physics*, 2009.

[4] Ž. Ivezić et al., "Solar system objects in the SDSS commissioning data," *The Astronomical Journal*, 2001.

[5] S. J. Bus and R. P. Binzel, "Phase II of the Small Main-Belt Asteroid Spectroscopic Survey. A Feature-Based Taxonomy," , vol. 158, pp. 146–177, Jul. 2002.

[6] R. P. B. et al., "Searching for v-type and q-type main-belt asteroids based on sdss colors," 2007. [Online]. Available: https://www.lpi.usra.edu/meetings/lpsc2007/pdf/1851.pdf

[7] M. A. et al., "TensorFlow: Large-scale machine learning on heterogeneous systems," 2015, software available from tensorflow.org. [Online]. Available: http://tensorflow.org/

[8] A. Novikov, "annoviko/pyclustering: pyclustering 0.8.1 release," May 2018. [Online]. Available: https://doi.org/10.5281/zenodo.1254845

[9] J. Sánchez Almeida, J. A. L. Aguerri, C. Muñoz-Tuñón, and A. de Vicente, "Automatic Unsupervised Classification of All Sloan Digital Sky Survey Data Release 7 Galaxy Spectra," , vol. 714, pp. 487–504, May 2010.