The impact of different parameter sets on the classification of asteroid types

Hanna Klimczak, Wojciech Kotłowski, Dagmara Oszkiewicz, Francesca DeMeo, Agnieszka Kryszczyńska, Tomasz Kwiatkowski and Emil Wilawer

The aim of the project is the classification of asteroids according to the most commonly used asteroid taxonomy (Bus-Demeo et al. 2009) with the use of various machine learning methods like Logistic Regression, Naive Bayes, Support Vector Machines, Gradient Boosting and Multilayer Perceptrons. Different parameter sets are used for classification in order to compare the quality of prediction with limited amount of data, namely the difference in performance between using the 0.45mu to 2.45mu spectral range and multiple spectral features, as well as performing the Prinicpal Component Analysis to reduce the dimensions of the spectral data.

Contact: hanna.klimczak@gmail.com
University of Adam Mickiewicz, September 2, 2021

Methodology

The classification of taxonomic types according to the Bus-DeMeo taxonomy [DeMeo et al., 2009] is performed with the use of the following machine learning methods: **Naive Bayes, Logistic Regression, Support Vector Machines, Gradient Boosting and Multilayer Perceptron**. The objects used for classification are 370 from [DeMeo et al., 2009] and 195 objects from [Binzel et al., 2019].

Three feature sets are used for classification:

- 41 spectral points from visible and near-infrared ranges ($0.45\mu m$ to $2.45\mu m$) processed as in [DeMeo et al., 2009] and slope of the spectral curve.
- Two main Principal Components of the aforementioned features and slope of the spectral curve.
- 35 reflectance values computed at chosen wavelengths to map known spectral features described by various authors and to maintain a link to previous classifications.

Results

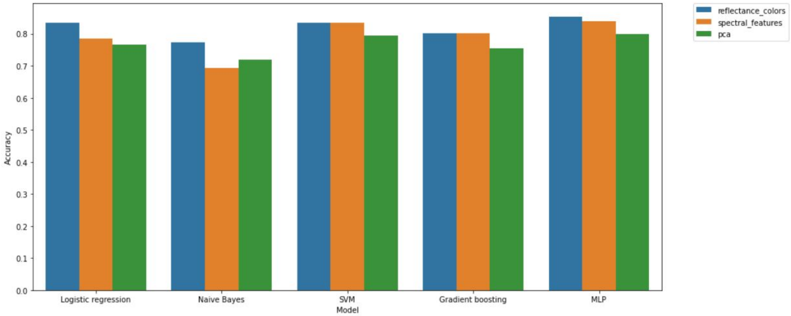

Figure: Average accuracy scores obtained by performing 10 runs of 5-fold cross validation

SVM and MLP achieve the best results across all feature sets, and Logistic Regression also performs well on the reflectance color set. The best scores per feature set are the following: **80%** for pca data (MLP), **84%** for spectral features (MLP), **86%** for reflectance colors (MLP).

Results

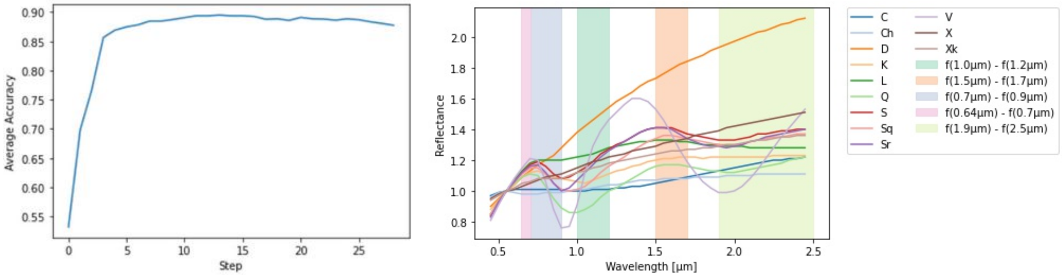

Figure: Average accuracy per step during Feature Selection (left) and most important reflectance colors against average spectra of each type (right).

The results of Feature Selected proved that no more than 5 reflectance colors were required to obtain **87%** accuracy (MLP). Most commonly selected features were: 0.9 μm feature, 0.9 μm absorption band, f(0.64μm) - f(0.7μm) absorption feature, the gentle concave up and 2.0 μm absorption band.

Conclusions

The main contributions of this work were the following.

- The classification of taxonomic types from [DeMeo et al., 2009] was performed on three different feature sets and a wide range of classification methods. Overall, the best results across all feature sets were obtained by the MLP with the following hyperparameters: two layers of 64 neurons, Stochastic Gradient Descent solver, batch size of 32 and adaptive learning rate with the initial value of 0.1. The results obtained with reflectance colors were not worse, or even slighlty better, than using the $0.45\mu m$ to $2.45\mu m$ spectral data.
- Sequential Feature Selection was performed on the reflectance colors set to quantify the importance of each color for classification. No more than 5 features were sufficient for 87% accuracy, which drastically reduces the amount of data needed to assign the taxonomic type.

Thank you!

This work has been supported by grant No. 2017/25/B/ST9/00740 from the National Science Centre, Poland.

References:

Binzel, R., DeMeo, F., Turtelboom, E., Bus, S., Tokunaga, A., Burbine, T., Lantz, C., Polishook, D., Carry, B., Morbidelli, A., Birlan, M., Vernazza, P., Burt, B., Moskovitz, N., Slivan, S., Thomas, C., Rivkin, A., Hicks, M., Dunn, T., Reddy, V., Sanchez, J., Granvik, M., and Kohout, T. (2019).
Compositional distributions and evolutionary processes for the near-earth object population: Results from the mit-hawaii near-earth object spectroscopic survey (mithneos).
Icarus, 324:41–76.

DeMeo, F. E., Binzel, R. P., Slivan, S. M., and Bus, S. J. (2009).
An extension of the bus asteroid taxonomy into the near-infrared.
Icarus, 202(1):160–180.