SCIENCE AUTONOMY USING MACHINE LEARNING AND DATA SCIENCE FOR PLANETARY MISSIONS: APPLICATION TO THE EXOMARS MISSION

V. Da Poian[1,2], E. I. Lyness[1,2], R. M. Danell[3], B. P. Theiling[1], W. B. Brinckerhoff[1], [1]NASA Goddard Space Flight Center (Greenbelt, MD 20771, da-poian.victoria@hotmail.com, [2]Microtel LLC, Greenbelt, MD, 20770, [3]Danell Consulting Inc., Winterville, NC, 28590)

Introduction: Most robotic space exploration missions return one primary product: science data. Regardless of the target – the Earth, planets in our solar system, the sun, or deep space, new science instruments will continue to generate more data, requiring more efficient processing approaches that enable advanced mission operations and speed scientific progress.

To maximize the value of each bit returned to Earth, instruments will increasingly need to select carefully the subset of total data to be prioritized for return, as in the future, compression and transmission of the full data volume will not be feasible. To perform this selection optimally, a fundamental goal is to enable the concept of *science autonomy*, where instruments collect measurement data, perform selected science data analyses onboard, and then autonomously and iteratively act upon those analyses through self-adjustment and tuning of instrument parameters, to yield an optimal and compact data set for return. Indeed, the selection of the next operation(s) to be run following preliminary measurements, without requiring ground-in-the-loop, increases mission efficiency and allows for successful shorter duration missions in hazardous planetary environments. This could also allow missions to prioritize the most compelling or time-critical data, yielding a more efficient and productive scientific investigation overall.

Ultimately instruments destined for outer solar system mission targets will need to baseline efficient analysis of their own data to make decisions about mission operations and prioritize data returned to Earth. As a first step in the area of separation mass spectrometry, we have begun to apply machine learning (ML) techniques to data from the flight-like engineering model of the Mars Organic Molecule Analyzer (MOMA). MOMA is a dual-source mass spectrometer-based investigation onboard the ExoMars rover Rosalind Franklin, that will search for past or present life on the Martian surface and subsurface through solid sample analysis. This effort, in preparation for operating on Mars, is aimed at 1) helping the ExoMars science and operations team quickly analyze new data and, 2) getting a better understanding of the challenges to enable science autonomy in future missions.

Methodology: We present our implementation using different ML techniques (i.e., clustering, classification) and the two significant challenges we faced, namely 1) sufficient data volume from these unique and highly optimized instruments to train neural networks (NN), and 2) sufficient trust in the system. Both challenges will be common to most, if not all, planetary science instruments. To deal with the data volume challenge, we discuss adopting transfer learning techniques to fine-tune NN trained on large amounts of commercial instrument data to operate on our limited MOMA dataset (fig. 1). For the 'trust' challenge, as it is not always clear what we are looking for in planetary science, we must consider agile ML applications proving to scientists that it will not filter out potentially groundbreaking data. We will discuss our concept of a Trust Readiness Level (TRL) for science autonomy akin to the NASA Technology Readiness Level.

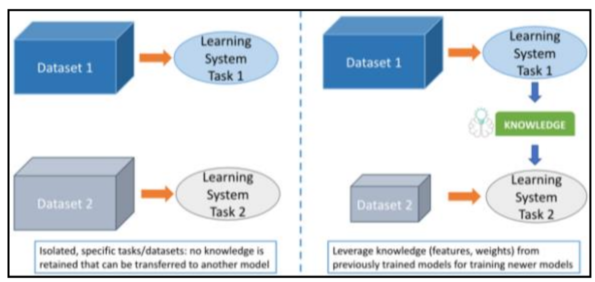

Figure 1: Illustration of ML techniques: basic learning (left) and transfer learning (right)

Preliminary results: We used MOMA Engineering Test Unit (ETU) data to develop initial ML algorithms and strategies as a proof of concept and to design software for supporting intelligent operations for autonomous systems. First results of this study show that 1) the preliminary categorization achieved using unsupervised algorithms as a filtering stage could permit autonomous operations such as prioritization of data to be sent to Earth, and 2) the prediction made using supervised learning algorithms could assist the scientists in their decision-making process during space operations. These initial results illustrate a path where on board processes can make decisions about re-tuning parameters specific for the studied sample, and therefore enable first generation science autonomy. Full science autonomy, which is our ultimate goal, will *minimize* ground-to-space interactions and similarly *maximize* the science return from future missions to challenging locations in our solar system and beyond.

Acknowledgments: This research is supported within the MOMA project with funding from NASA's Mars Exploration Program (MEP). It has been using the performance computing resources from the GSFC Science Data Processing branch. We thank the group of MOMA scientists who participated in the different workshops, Xiang Li, Melissa Trainer, Desmond Kaplan, Andrej Grubisic, Friso van Amerom, Stephanie Getty, and Marco Castillo, for their guidance and support. We also thank the software team development who supported the project, Joe Cirillo, Sam Larson, Nick Dobson and Joe Avolio.

References:

[1] Goesmann, F., Brinckerhoff, W.B., Raulin, F., Goetz, W., Danell, R.M., et al. (2017). The Mars Organic Molecule Analyzer (MOMA) Instrument: Characterization of Organic Material in Martian Sediments. International Journal of Astrobiology 17.

[2] X. Li, R. M. Danell, V. T. Pinnick et al., "Mars Organic Molecule Analyzer (MOMA) laser desorption/ionization source design and performace characterization," International Journal of Mass Spectrometry 422, 2017.

[3] NASA Procedural Requirements, Appendix J, 7120.8, " NASA Research and Technology Program and Project Management Requirements, 2018.