DEVELOPING NATURAL LANGUAGE PROCESSING AND SUPERVISED MACHINE LEARNING TECHNIQUES TO CLASSIFY MARS TASKS

Adam Kurth[1], Hunter Rehm[2], and Mona Matar[3]

[1]Arizona State University, Tempe AZ, [2]HX5, LLC., Cleveland, OH,
[3]NASA Glenn Research Center, Cleveland OH

As NASA's Human Research Program (HRP) prepares for long-duration Mars missions, understanding astronaut tasks is crucial. This study, conducted at NASA Glenn Research Center (GRC), employed Natural Language Processing (NLP) and machine learning techniques to analyze and classify Mars tasks.

A list of 1,058 Mars tasks was provided by HRP experts including binary labeling of 18 Human System Task Categories (HSTCs) [1]. We developed an NLP model using Google's BERT language model [2] to capture the semantic and syntactic nuances of these tasks. Supervised training was initially applied to a subset of the NLP-analyzed tasks to assess the model's effectiveness in classifying the remaining tasks. Incorporating HSTC descriptions significantly enhanced the classification accuracy for 9 out of the 18 HSTCs and reduced training time. To address the issue of severe class imbalance in the HSTC data, we introduced innovative weighting and sampling techniques for data augmentation [3].

We then fine-tune BERT to implement a pairwise relatedness scoring method, allowing us to cluster tasks based on their relatedness and similarity, getting a step closer to labeling the tasks without supervision.

In this presentation we guide you through data preprocessing, deciphering key syntax components using BERT, and performing supervised classification of the Mars tasks. This work showcases the potential use of advanced NLP techniques to analyze Mars missions to be incorporated into various crew health and performance analyses.

REFERENCES
[1] Fernandez W. et al (2023) *Journal of Space Safety Engineering* 10, no. 2: 231-238.
[2] Devlin J. et al (2019). arXiv preprint arXiv:1810.04805.
[3] Wei J. and Zou K. (2019). arXiv preprint arXiv:1901.11196 (2019).