# Neural Networks for Stepwise Classification of Multi-Attribute LIBS Spectra under Simulated Martian Conditions

*Alexander von Canal, German Aerospace Center, Institute of Optical Sensor Systems, Berlin Adlershof, alexander.voncanal@dlr.de*
*Kristin Rammelkamp, German Aerospace Center, Institute of Optical Sensor Systems, Berlin Adlershof*

## 1. Introduction and goals

The goal of this study is to build a stepwise classification scheme for multi-attribute laser-induced breakdown spectroscopy (LIBS) data using back propagation neural networks (BPNN). The classification pipeline is depicted in **Fig. 1**. LIBS uses a pulsed laser to induce a plasma on the surface of a sample (see **Fig. 2)**. Light emitted by the plasma contains characteristic emission depending on the chemical composition of the target. Since there is no need for direct contact to the sample, LIBS is highly relevant for in-situ exploration of extraterrestrial bodies **[1]**, and is currently deployed on several Mars missions **[2, 3, 4]**.

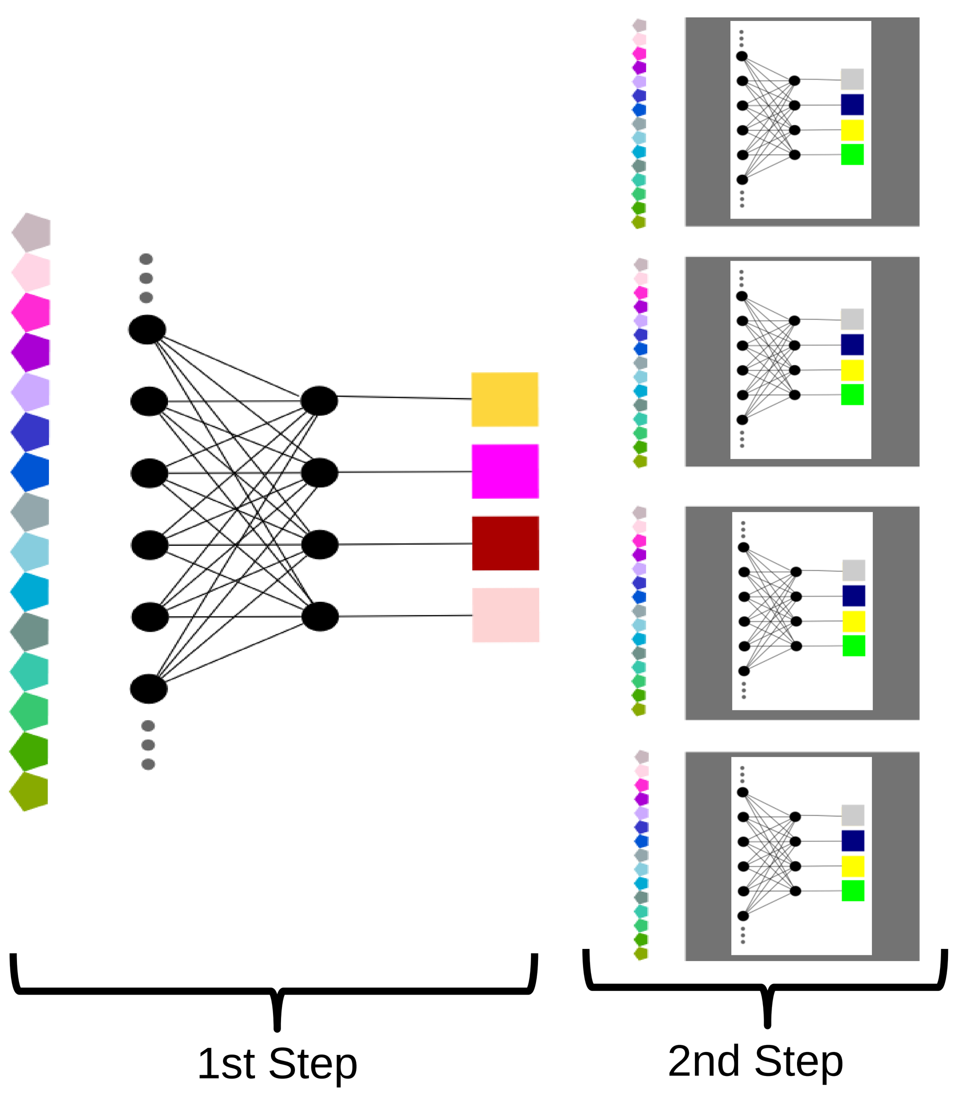



**Fig. 1: Classification pipeline for multi-attribute LIBS data.**

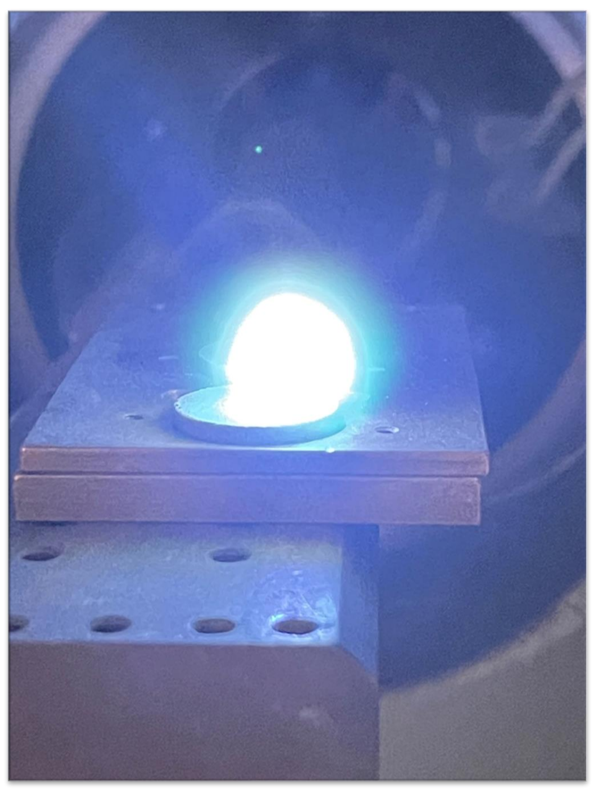


**Fig. 2: Picture of the plasma plume during a LIBS measurement under Martian atmospheric conditions**

## 2. Sample preparation and data set

We measured 2500 LIBS spectra under Martian atmospheric conditions from 100 prepared samples. **Fig. 3** depicts the sample preparation schematically and **Fig. 4** shows one example LIBS spectrum. Specifically, each sample consists of:

- One basaltic Mars Simulants (JEZ-1, MGS-1, MGS-1C, MGS-1S). **[5]**
- Mixed with one out of four salts (NaCl, MgCO3, CaSO4 2H20, MgSO4 H2O).
- Salt concentrations were varied between ~0.5 -15 wt% to account for realistic variance of water-deposited salts and cements in Martian sedimentary rocks.
- Four samples without added salt.

Each sample was measured five times with five different laser energies (~5-51mJ) to account for varying laser irradiances as it is the case for real measurements on Mars. Pulse duration and spot diameter of the laser are 6ns and 300µm respectively.

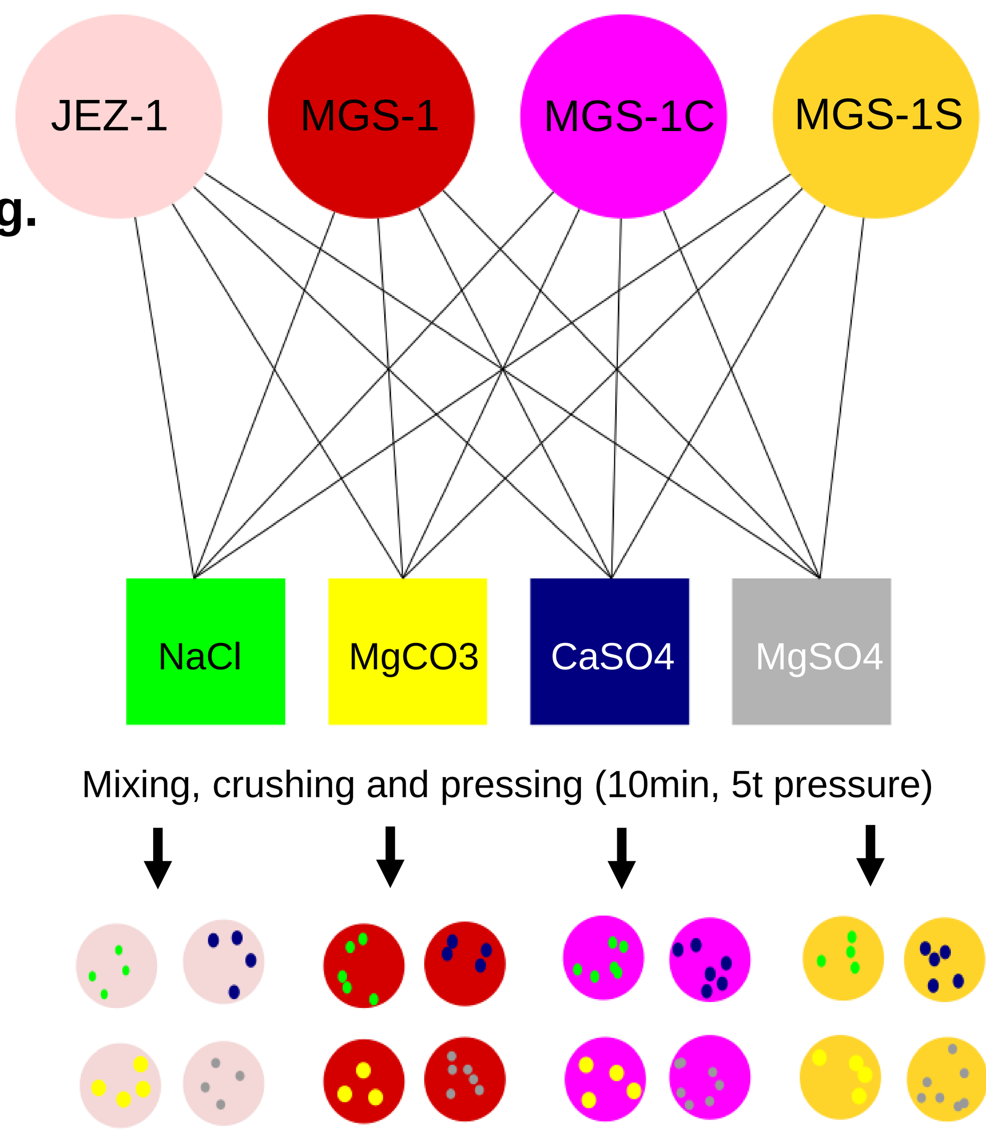
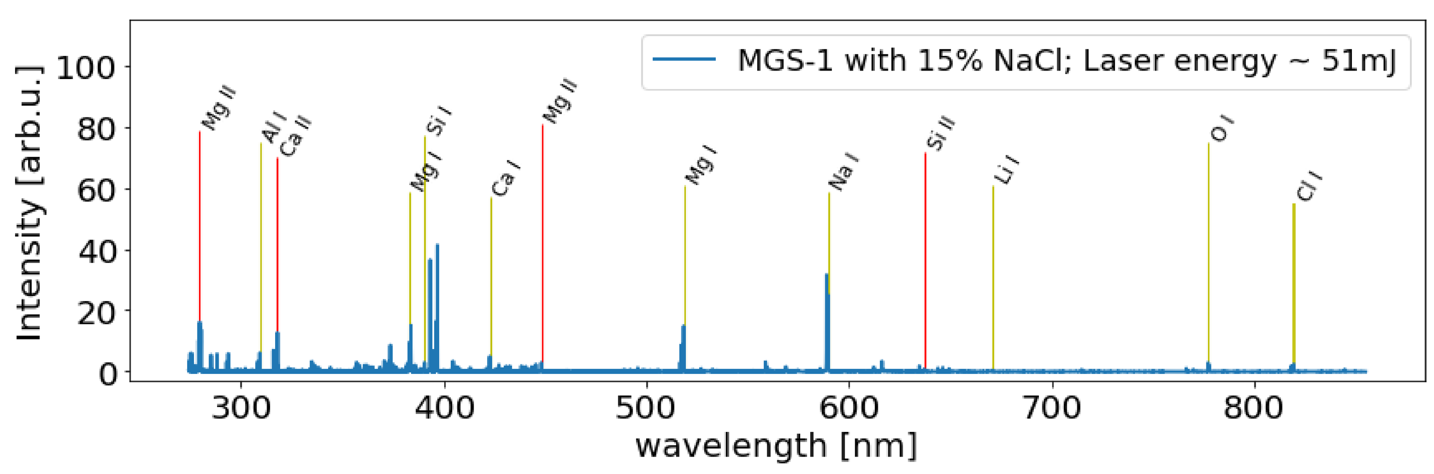



**Fig. 3: Sketch of sample preparation as described in the text.**

**Fig. 4: Example of a standardized LIBS spectrum with some annotated neutral (I) and ionized (II) atomic emission lines.**



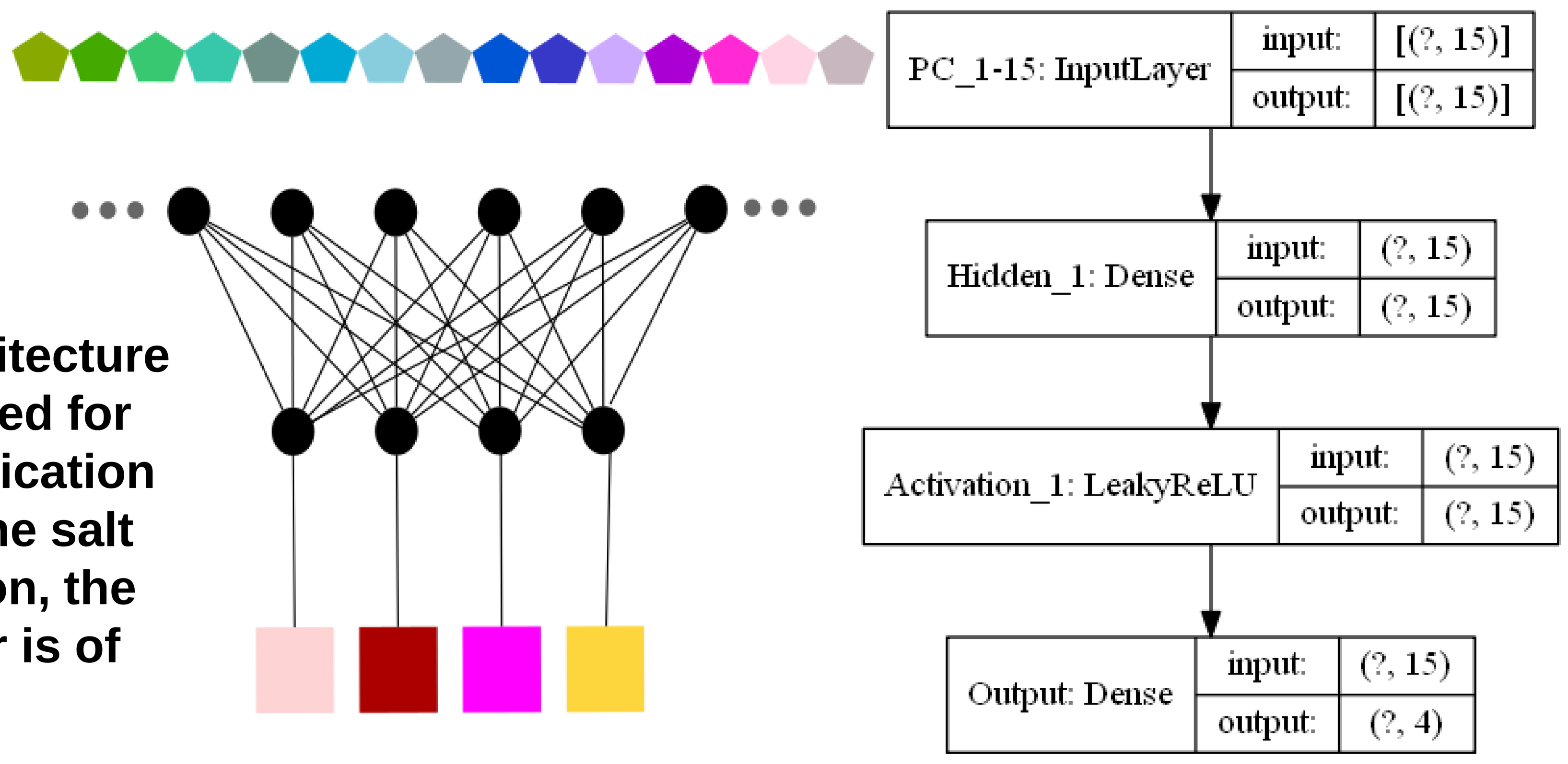




**Fig. 5: Architecture of BPNN used for both classification steps (for the salt classification, the output layer is of size five).**

## 3. Results of classification steps

- Before training, the dimensionality of the data set was reduced with principal component analysis (PCA).
- First 15 principal components (PC) served as inputs for the first BPNN classifier.
- The BPNN architecture for both classification steps is the same, namely one fully connected hidden layer of size 15 (as shown in **Fig. 5**).
- For the second classification step, a new PCA was down for each Mars simulant. Again the first 15 PC served as inputs for the BPNNs.
- Train-test splitting was always chosen to be 80%-20%
- All BPNNs were trained with a learning rate of 0.01 and Adam optimizer (implemented with PyTorch).
- **Fig. 6** shows the final classification accuracies over 50 different data splittings after 30 epochs of training, for (a) Mars simulant and (b) salt classification. Different values on the x-axis correspond to different splittings of the train-test data. All splits were chosen such that all five measurement form the same sample were not split up.

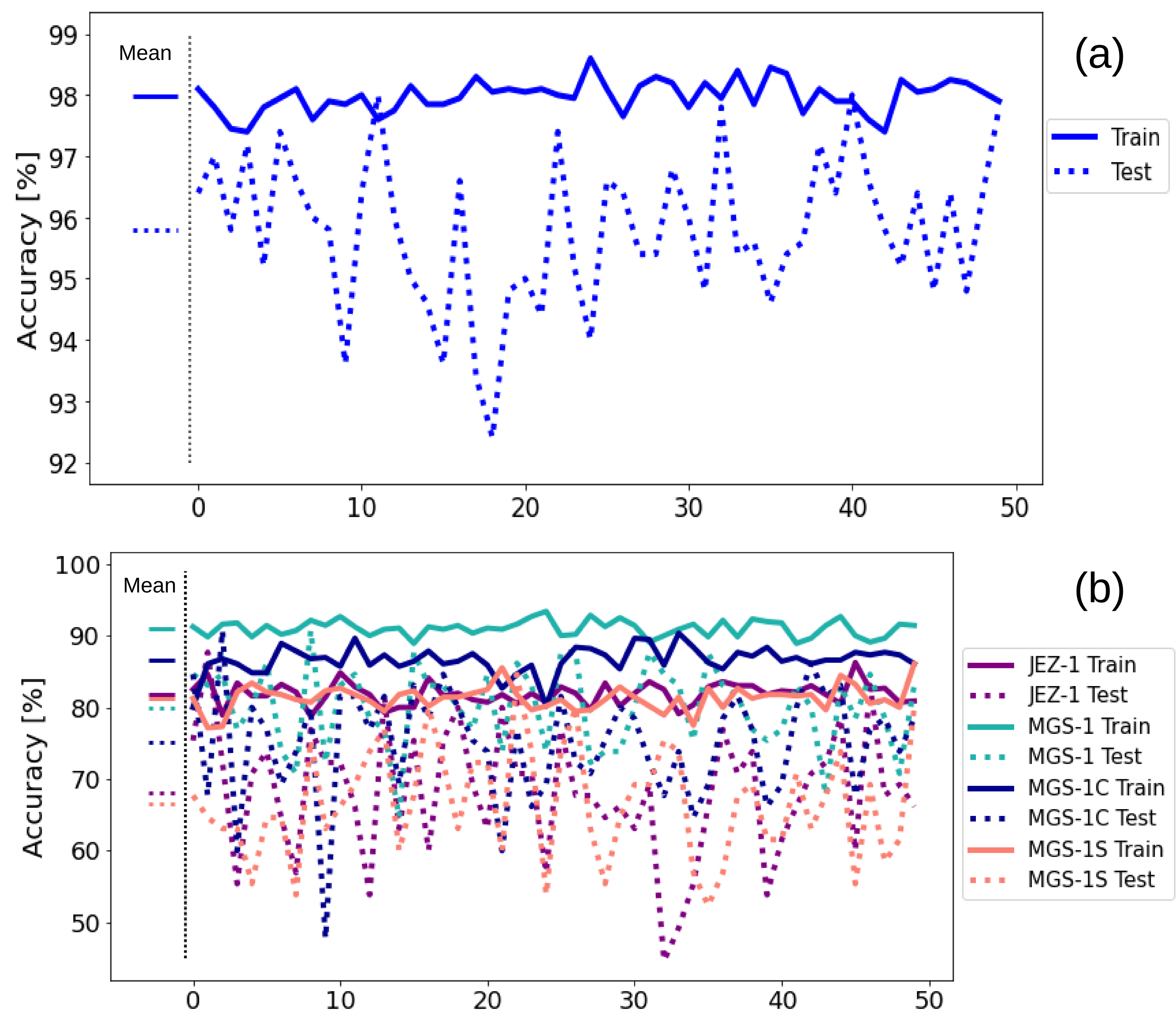



**Fig. 6: Train and test accuracies after 30 epochs for different train-test data splittings. a) First classification step; b) second classification step.The mean accuracies over all splitings are depicted next to the vertical black dotted lines (see text for details).**

## 3. Conclusion and outlook

- The Mars simulant classification works well, with an average train / test accuracy of ~ 98% / 96%.
- The salt classification is more difficult, since after the first classification step, training set sizes are only 0.25 of the initial full data set. Furthermore, some salt concentrations are very small, which could make classification inaccurate. Future work will focus on improving this classification step.
- Compared to other machine learning approaches, namely random forest classifier (RF) and partial least squares discriminant analysis (PLS-DA), we achieved slightly better accuracies with BPNNs.
- A general challenge is the small size of our data set. Since only few labeled LIBS data is available, we aim to work on generative models for synthetic data extension in the future. Especially variational autoencoders (VAEs) and generative adversarial networks (GANs) will be investigated.

**References:**

**[1]** Knight et al. (2000), Applied Spectroscopy, 54(3); **[2]** Maurice et al. (2016), Journal of Analytical Atomic Spectrometry (Vol. 31, Issue 4, pp. 863–889); **[3]** Xu et al. (2021), Space Science Reviews (Vol. 217, Issue 5); **[4]** Maurice et al. (2021), Space Science Reviews (Vol. 217, Issue 3); **[5]** Cannon et al. (2019), Icarus, 317, 470–478.

DLR Deutsches Zentrum für Luft- und Raumfahrt