Predicting Multi-Component Mineral Compositions in Gale crater, Mars with Label Distribution Learning

The CheMin X-ray diffraction (XRD) instrument onboard the Mars Science Laboratory rover *Curiosity* analyzes drilled rock fines and scooped soils in Gale crater, Mars. The CheMin team estimates mineral abundances and unit-cell parameters of major crystalline phases found in each of the CheMin samples with Rietveld refinement of the XRD patterns. The main crystalline phases identified by CheMin include plagioclase, sanidine, pyroxenes, olivine, magnetite, and alunite-jarosite group minerals. In order to better understand the formational conditions and geologic history of the minerals found in Gale crater, the CheMin team developed a crystal-chemical method to predict limited chemical compositions of the minerals observed in the CheMin samples [1,2]. However, limitations in the statistical algorithms inhibited the prediction of chemical compositions beyond three elements. In this study, we adapt a machine learning technique, Label Distribution Learning (LDL) [3], to predict multicomponent chemical compositions of Gale crater mineral phases, thereby allowing for more detailed petrologic interpretation of the geologic history of the martian surface.

LDL is a novel framework for classification problems with small datasets and has been widely applied to facial recognition problems such as age estimation. In this study, we adapt the LDL algorithm such that it can predict chemical elements (labels) and their abundances (degrees) for each martian mineral sample, based on crystallographic parameters. We evaluate performance using distance and similarity between label distributions as well as mean square error and also compare the results to traditional machine learning methods.

[1] Morrison et al. (2017) Relationships between unit-cell parameters and composition for rock-forming minerals on Earth, Mars, and other extraterrestrial bodies, Am Min, 103(6): 848-856
[2] Morrison et al (2017) Crystal chemistry of martian minerals from Bradbury Landing through Naukluft Plateau, Gale crater, Mars, Am Min, 103(6): 857-871
[3] Geng (2016) Label distribution learning. IEEE Transactions on Knowledge and Data Engineering, 28(7), 1734-1748

## Authors

- **Shaunna M Morrison**
  - *Geophysical Laboratory of the Carnegie Institution for Science*
- Feifei Pan
  - *Rensselaer Polytechnic Institute*
- Olivier C Gagné
  - *Geophysical Laboratory of the Carnegie Institution for Science*
- Anirudh Prabhu
  - *Rensselaer Polytechnic Institute*
- Ahmed Eleish
  - *Rensselaer Polytechnic Institute*
- Peter Arthur Fox
  - *Rensselaer Polytechnic Institute*
- Robert T Downs
  - *University of Arizona*

- Thomas Bristow
  - *NASA Ames Research Center*
- Elizabeth B Rampe
  - *NASA Johnson Space Center*
- David Frederick Blake
  - *NASA Ames Research Center*
- David Vaniman
  - *Planetary Science Institute*
- Cherie Achilles
  - *NASA Goddard Space Flight Center*
- Douglas W Ming
  - *NASA Johnson Space Center*
- Albert Yen
  - *NASA Jet Propulsion Laboratory*
- Allan H Treiman
  - *Lunar and Planetary Institute*
- Richard V Morris
  - *NASA Johnson Space Center*
- Steve Chipera
  - *Chesapeake Energy*
- Patricia Craig
  - *Lunar and Planetary Institute*
- Valerie Tu
  - *Jacobs at NASA Johnson Space Center*
- Nicholas Castle
  - *Lunar and Planetary Institute*
- Philippe Sarrazin
  - *SETI Institute Mountain View*
- David J Des Marais
  - *NASA Ames Research Center*
- Robert Hazen
  - *Carnegie Institution for Science*