**MASTER THESIS PROPOSAL**

Supervision by Prof. Angelo Pio Rossi and Dr. Carlos Henrique Brandt

M.A student: Feven Legesse Mammo (Data Engineering Masters)

**RESEARCH TOPIC**

Machine Learning Model that Detects Minerals from the Surface of Mars.

**INTRODUCTION**

This study is intended to focus on mineral detection on the surface of mars based on the data collected from an instrument called CRISM spectrometer. Minerals are inorganic solid materials which we use them in different format for various purposes in our daily life routines. For example, a mineral called “gypsum” for building a house and a mineral “diamond” for jewelry and so much more. Mainly, minerals are the building block of rocks (Butler, 2018). According to International Mineralogical Association in 2018 there are 5,400 minerals are found on earth (America, 2020). Mars the second planet behind Earth, which is known as the “red planet”, it’s surface is rocky and mostly covered by a red dust which is the reason of thick layer of oxidized iron (Wild, 2017). For the reason that the surface of mars is somewhere among the basalt or andesite rocks on earth, the formation of minerals on mars has similarity with what is found on earth (wikipedia, 2019). CRISM (Compact Reconnaissance Imaging Spectrometer for Mars) is an instrument that is an imaging spectrometer with a scannable field of view that can cover wavelengths from 0.362 to 3.92 microns at 6.55 nanometers/channels so that can identify a broad range of minerals on the Martian surface (Webmaster, n.d.).

Spectral imaging is sampling of data in many wavelength bands and it yields a three-dimensional data that is called data cube. The third dimension of the data cube represents the spectrum of each pixel. Basically, spectral imaging can be divided in to two, one is multispectral that samples data from different and discrete but disconnected wavelengths the other is hyperspectral imaging which samples enough data to reconstruct a continuous or a connected spectrum over a given spectral range. The image data (data cube) I am going to work with is gathered through CRISM instrument within 9 different bands which is a multispectral image (ALABAMA, 2020). From this research project I am hoping to do a valuable contribution towards Asteroid minning and Planetary geology field.

**BACKGROUND (Should be included in Introduction part in thesis)**

**Surface of Mars and Mineral Extraction**

In general, since minerals has huge relevance to create something useful out of it for humankinds or for a whole living thing in the world, it is essential to discover more about it. On the point of mineralization on mars there is not so much known, yet which is an indication of a lot to discover, according to the exploration made by a machine (Robot) spacecraft the knowledge about it is less than about mars geology. Below Syrtis Major (“an extended plateau on the planet Mars which is the most dark and noticeable formation. (wikipedia, Syrtis major, 2017)”) pyroxene and olivine are stored up material in solidified system. Again, there is a suggestion that Ni–Cu and PGE ores can be found there also. Mars pathfinder figured out that iron contains one-fifth of the wait of the soil which is an indication of importance of iron-rich material on Mars. In the low land of Mars Infrared Mineralogical Mapping Spectrometer detected the sulfate mineral kieserite (Mg-sulfate). But there are some minerals deposits which are not found in Mars. However, recently nickel-iron meteorites (in number “three”) discovered by the Mars exploration rovers. Which those minerals are principal for steel manufacture. (West, 2018).

**Machine Learning predictive Algorithms**

Choosing an appropriate methodology or algorithm is necessary to obtain an accurate mineral potential map. Accuracy depends only on the capacity of the algorithm to learn multiform relationships among the input evidential features and mineral deposit occurrence. In addition, this, two things must be considered: transparency and Interpretability as seen in a journal under Ore Geology Reviews titled “Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines”. (V.Rodriguez-Galiano, 2015). There must be attentive selection out of machine learnning classification algorithms to do modeling, most importantly depending on the data entitled to trained by the machine learning algorithm. Based on this article explanation about finding the best tool for the classification model out of Random forest, Decition Tree, Support vector machine and Artificial Neural Network (ANN) algorithms the SVM and RF perform better. But still most studies did not seek to understand the machine learning algorithms performance during (at the time of) luck of trainning data.

**Multispectral Image classification Algorithm**

In Feb.2016 Institute of Electrical and Electronics Engineer (IEEE) published paper expressed that Neural network ensemble (NNE) classifier is the most robust and efficient classification algorithm for Multispectral image classification. As it is described on the paper NEE is a best model comparison with Bayes Maximum-Likelihood Classifier and K-NN (K-nearest Neighbors) Classifier based on the accuracy of each models on the multi-spectral image data took from SPOT image of Calcutta in India. (Fu, 2016).

**OBJECTIVE**

The objective of this research is listed below:

1. To help people who is working in the area of Planetary Geology by giving insight about which mineral is there in the chosen surface of mars.
2. In scientific aspect: to have better understanding about the evolution and content of the universe.
3. For extraction of minerals in the planet such as iron.
4. In resource point of view to upgrade the knowledge about what are the resources in planet mars.

**RESEARCH QUESTION**

1. Which one is the best machine learnning model that is more robust and efficient for feature detection and give more accurate prediction of multispectral image?
2. Which mineral are the most accumulated in the surface of mars and in which area?
3. Is it possible to infer and state a concrete information about the surface of mars related to minerology?

**METHODOLOGY**

I plan to answer the research questions and fulfill my objective, that I have mentioned above by taking the following steps: -

1. Pre-processing of the data and perform Exploratory data analysis (EDA).
   * Describe the dataset: what is the/are the information provided by the data; what is the data format
2. Implement machine learnning algorithms. The chosen algorithms are listed below.

* SVM (Support vector machine)
* Random forest
* Neural networks
* What are our features?

1. Best classifier model selection. And fine tunning.

**WORK PLAN (2020)**

|  |  |  |  |
| --- | --- | --- | --- |
| **Work** | **Start date** | **Submission date** | **HR/day** |
| 1. Dataset description 2. Data preprocess 3. EDA (Exploratory Data Analysis) | March-09 | March-28 | 8 hours |
| * SVM (Support Vector Machine) | March-30 | April-11 | 8 hours |
| * (RF)Random Forest | April-13 | April-25 | 8 hours |
| * Neural Networks | April-27 | May-16 |  |
| Thesis report writing | May-18 | June-05 | 8 hours |

**REFERENCES**

ALABAMA, U. O. (2020 ). *Spectral Imaging.* Retrieved from USA University of South Alabama: https://www.southalabama.edu/centers/bioimaging/spectral\_imaging.html

America, M. S. (2020). *Collector's Corner.* Retrieved from FREQUENTLY ASKED QUESTIONS MINERALOGY - GENERAL: http://www.minsocam.org/msa/collectors\_corner/faq/faqmingen.htm

Butler, C. (2018). *Minerals: Building Blocks of Rocks.*

Fu, X. (2016). Multispectral image classification based on neural network ensembles. *Eighth International Conference on Advanced Computational Intelligence (ICACI).* Chiang Mai: IEEE.

V.Rodriguez-Galiano, M.-C.-O.-R. (2015). Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. *ScienceDirect*, https://www.sciencedirect.com/science/article/pii/S0169136815000037.

Webmaster, J. (n.d.). *Instrument*. Retrieved from CRISM: http://crism.jhuapl.edu/instrument/design/overview.php

West, M. D. (2018, December 27). Planetary and Space science . *Potential martian mineral resources: Mechanisms and terrestrial analogues*, pp. 58(2010) 574-582.

wikipedia. (2017, December 21). *Syrtis major*. Retrieved from wikipedia: https://de.wikipedia.org/wiki/Syrtis\_Major

wikipedia. (2019, October 27). *Mineralogy of Mars*. Retrieved from wikipedia.org: https://en.wikipedia.org/wiki/Mineralogy\_of\_Mars

Wild, F. (2017, August 07). *What Is Mars?* Retrieved from NASA TV: https://www.nasa.gov/audience/forstudents/5-8/features/nasa-knows/what-is-mars-58.html