# RET Project: Solar Fault Detection and Classification using Machine Learning

## Research Experience for Teachers (RET) Summer 2020

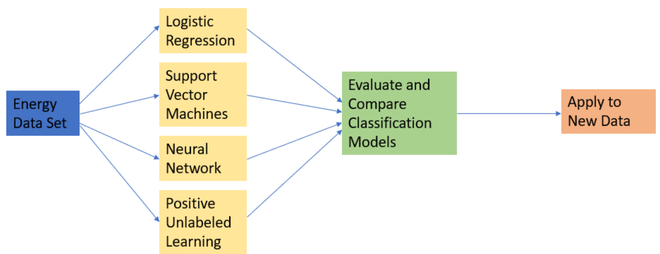
Milton Johnson, Bioscience High School, Phoenix, Az
*Kristen Jaskie, Dr. Andreas Spanias; SenSIP Center, School of ECEE, Arizona State University.*

## RESEARCH BACKGROUND/DESCRIPTION

- Similar work has been done already with applying classification algorithms to fault detection in solar systems.
- Our goal is to also apply the Positive Unlabeled algorithm to the data set as well.

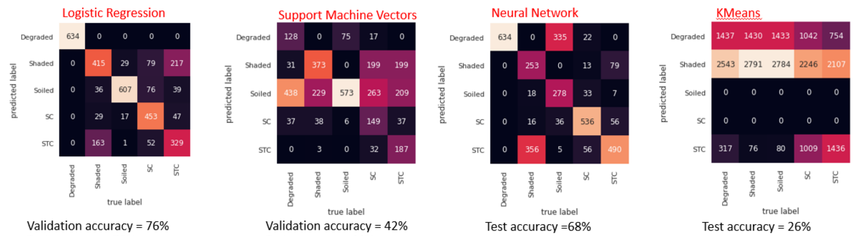

## RESEARCH OBJECTIVES/PLAN

- Create classification models using labeled data for solar arrays that are affected by soiling, weather, ground leakage and short circuits.
- Evaluate the models and use our assessments to create new methods that can be used for unlabeled solar array data sets.




## RESEARCH RESULTS/REMARKS

- The logistic regression algorithm worked best for the labedl data. Neural network showed good results but was unstable.
- The modified logistic regression (PU) model performed very well with unlabeled data.




Validation accuracy = 76%
Test accuracy = 76%

Validation accuracy = 42%
Test accuracy = 44%

Test accuracy =68%

Test accuracy = 26%

Figure . Accuracy results of classification testing. A Kmeans test is also included

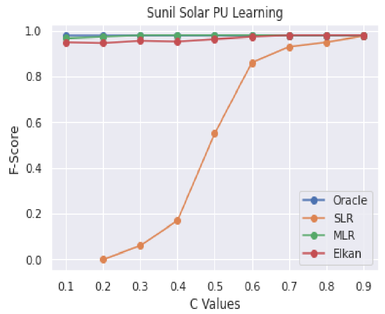



Figure . PU classificaton vs standard logistic regression and others

**Solar Array Data Features**
DC Array Output
Voltage-max power
Current-max power
Cell Temperature-C
Plane of Array Irradiance-W/m^2
Fill Factor
Gamma
Power-max power
Voltage-open current
Current-short circuit

Figure . Labels used for classification tests

## REFERENCES

[1] Rao, Sunil, et al. "An 18 kW solar array research facility for fault detection experiments." 2016 18th Mediterranean Electrotechnical Conference (MELECON). IEEE, 2016.
[2] Maghami, Mohammad Reza, et al. "Power loss due to soiling on solar panel: A review." Renewable and Sustainable Energy Reviews 59 (2016): 1307-1316.
[3] Fault Classification in Photovoltaic Arrays via Graph Signal Processing Jie Fany, Sunil Raoy, Gowtham Munirajuy, Cihan Tepedelenlioglu and Andreas Spanias SenSIP Center, School of ECEE, Arizona State University

## LESSON PLAN OBJECTIVES

Students will:

- manually develop the concept of the Kmeans algorithm
- demonstrate the use of Kmeans Python algorithm in Google Colab
- apply Kmeans algorithm to choosing an optimal location for a business and provide a report

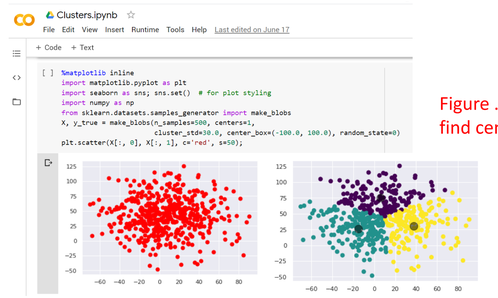
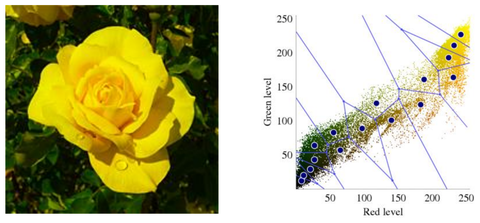



Figure . Using Kmeans clustering to find centroids of data sets.

## LESSON IMPLEMENTATION/OUTCOMES

- Explain the math behind the Kmeans algorithm
- Demonstrate use of python code to run Kmeans algorithm
- Apply algorithm to solving real-world questions
- Communicate results in technical report




Wikipedia images showing Kmeans applied to image compression

Ira A. Fulton Schools of Engineering
ARIZONA STATE UNIVERSITY
Sensor Signal and Information Processing Center
https://sensip.asu.edu
This work is funded in part by SenSIP center and NSF Award 1953745
NATIONAL SCIENCE FOUNDATION