DESERT HEIGHTS
CHARTER SCHOOLS

DESERT HEIGHTS
CHARTER SCHOOLS

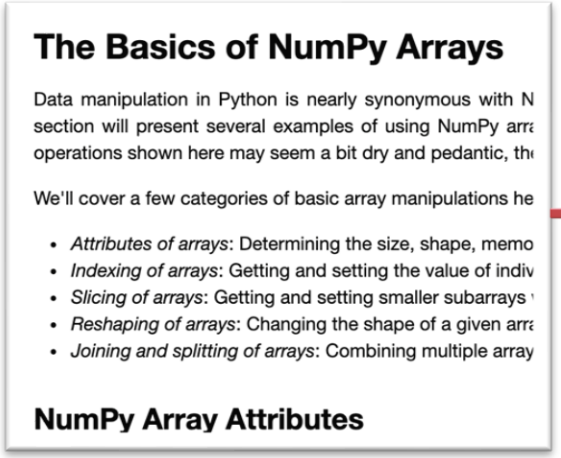
RET Project: Power Predictions in PV Panels

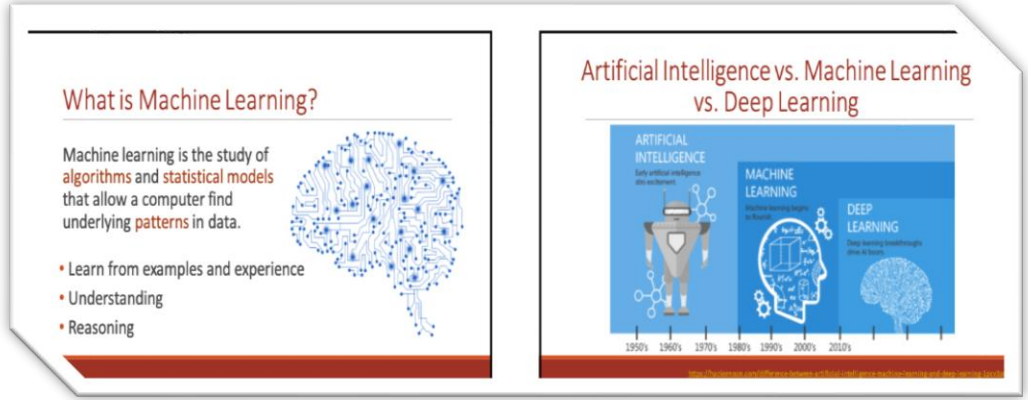
Deborah Sweeting

High School Algebra 2
Desert Heights Preparatory Academy
Mentors: Mentors G. Muniraju, Dr. A. Spanias

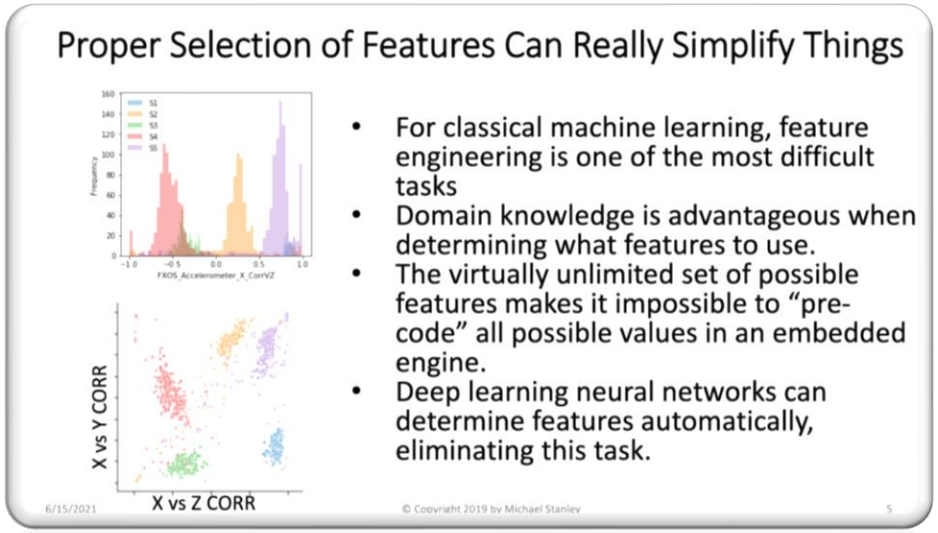
NSF Award 1953745

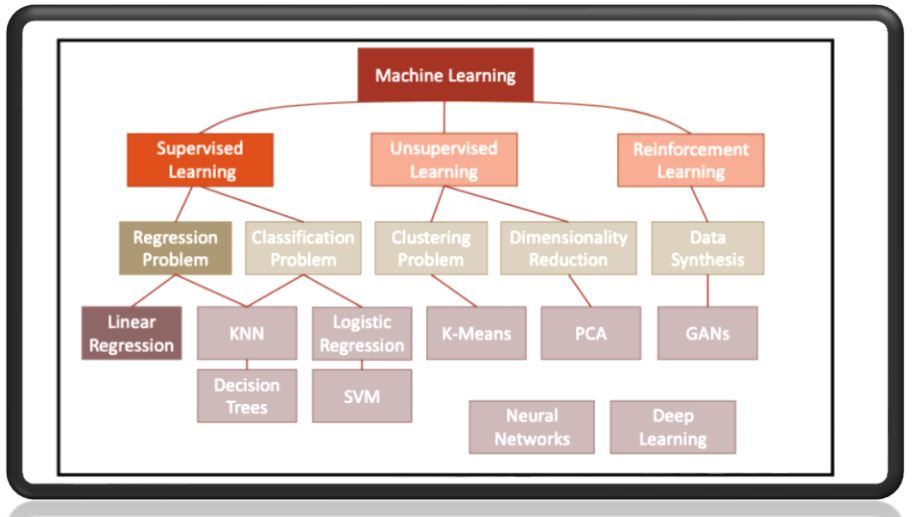
https://sensip.engineering.asu.edu/ret/





RET Research and Training





https://sensip.engineering.asu.edu/ret/

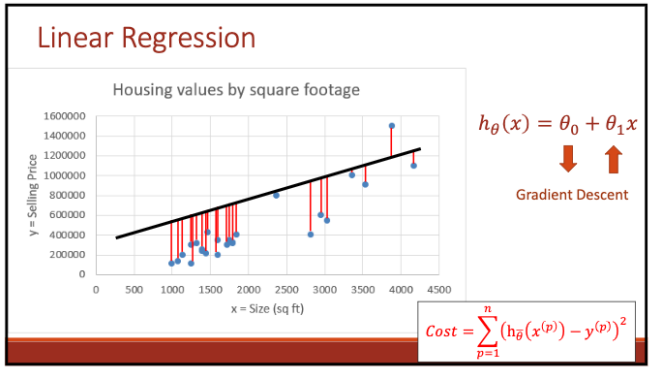
RET Schedule and Training

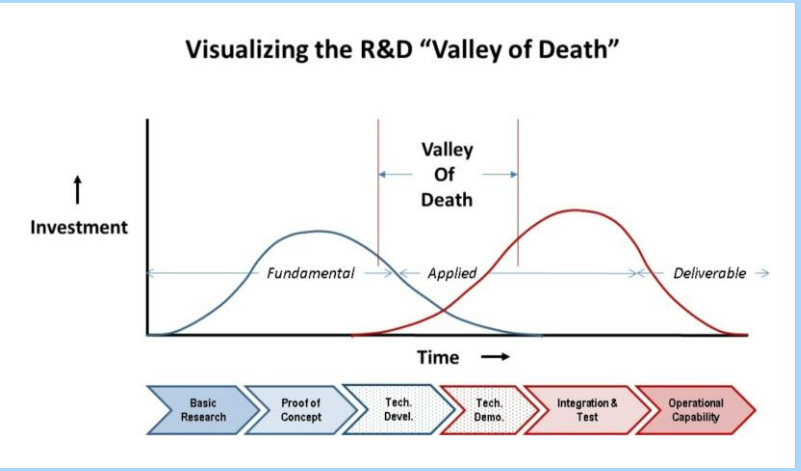
Figure 1.2: The SenSIP Solar Monitoring Facility at the ASU Research Park [8].

Hands On Technical Training

Topics covered:

- Python/Colab uses and functionality
- Types of Machine Learning (focus on clusters)
- Practice in Colab – Used SciKit library to run kmeans algorithm
- Regression
  - KNN
  - ML & Gradient Descent
  - Linear Regression - Outputs
- Classification
  - Logistic Regression – Probabilities vs. Outputs
- Embedded ML
- Gary Mastin's presentations on Research & Development
  - Cultural bias and management behavior usually determine if collaboration occurs.





https://sensip.engineering.asu.edu/ret/

RET Lab Experience Research Summary

Research Objectives

- Problem: Once researchers have a more efficient way of training sensors to find faults in solar panels, there needs to be an efficient analysis of the data to quickly predict power outputs.
- Hypothesis: There is little to know residual in the actual data's power output and the predicted power output.

Deborah.Sweeting@dhschools.org
gmuniraj@asu.edu
spanias@asu.edu

***Abstract*—The application of Smart Monitor Devices (SMD) to solar arrays has enabled the collection of data for 14 solar features on each panel of an array. Features include open circuit voltage, maximum voltage, maximum current, short circuit current, temperature of module, irradiance of module, fill factor, gamma, and power. This data can be used to predict DC Power output based on common solar array faults such as soiling, arc faults, degraded modules, and shading. After collecting feature data, principal component analysis can be used to determine the importance of features to an artificial neural network. Determining the importance of factors can allow for a simplification of the neural network model used for fault detection.**

***Index Terms*—Feature analysis, Nonlinear PCA, Neural networks, Machine Learning, PV modules.**

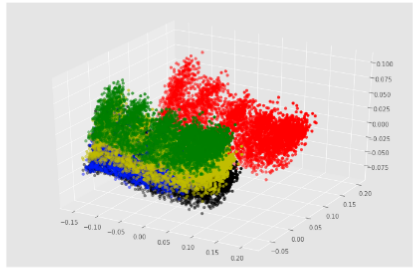


Fig. 1. 3D Component Space Fault Clusters Generated Using KPCA.

I. Introduction

Previous neural network implementations of the SMD solar array data used all nine features as inputs [1]. Initial attempts at redesigning the network to use fewer input features resulted in significant reduction of the networks accuracy in fault classification. It would be desirable to reduce the number of features the model is dependent on for computational efficiency, and to reduce the number of sensors needed for each SMD.

One method of dimensionality reduction is principal component analysis (PCA). Principal component analysis allows us to determine the importance of variables and the correlation between variables [2]. Ideally, in this application, principal component analysis will allow us to determine a set of four or fewer input features that dominate the networks ability to classify faults. Principle component analysis can be achieved by scaling the feature data such that all features are of similar magnitude and using the scaled data to create an association matrix and covariance matrix. The eigenvectors of these matrices. known as the load and score matrices. can principal component analysis (NLPCA) techniques. The advantage of NLPCA is revealing nonlinear correlation between features that can not be observed through PCA. One method for NLPCA is using Kernel Principle Component Analysis (KPCA). KPCA operates on the understanding that a nonlinear system can be projected into a higher dimension to make the data linearly separable. [3] This method requires the use of a nonlinear kernel function to enable the projection.

NLPCA can also be implemented using an autoencoder neural network configuration that attempts to recreate input data at the network output. [4]. The autoencoder neural network consists of a encoding, bottleneck, and decoding layer. The bottleneck layer is frequently used in noise reduction applications and similar principles apply for the dimensionality reduction. By funneli.

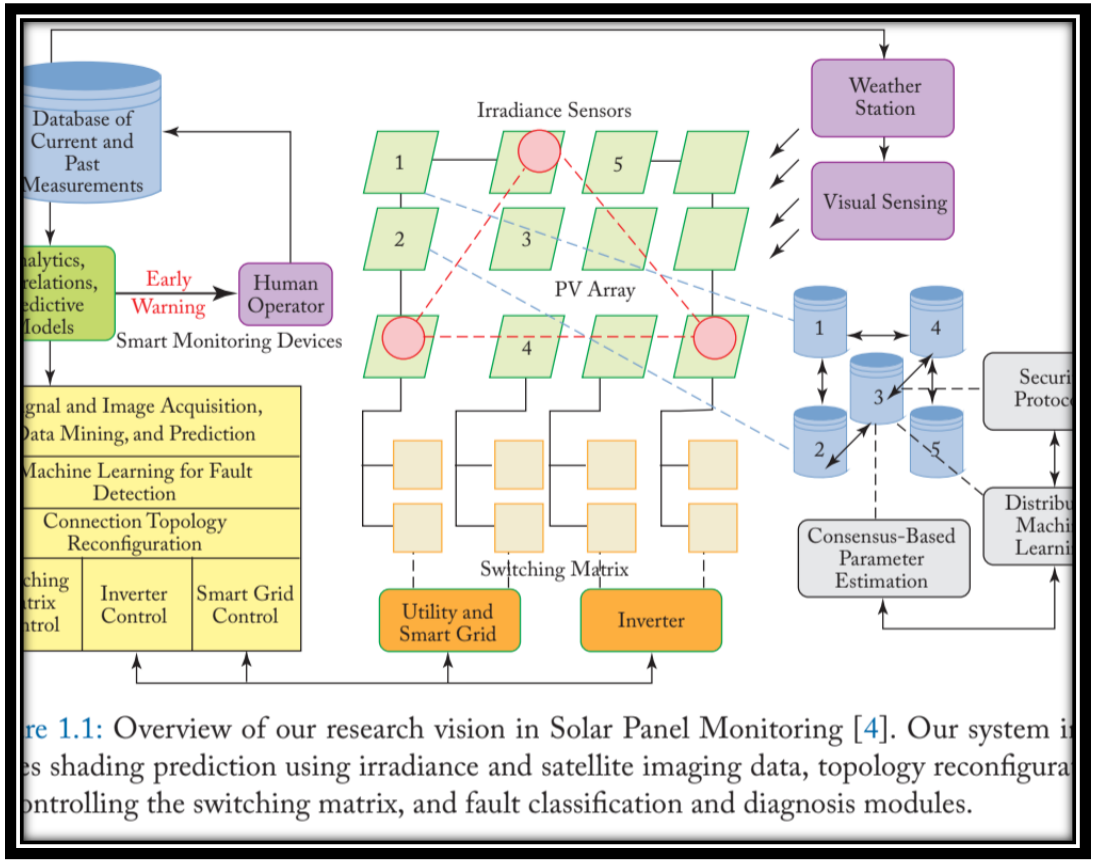
References

[1] S. Rao, S. Katoch, V. Narayanaswamy, G. Muniraju, C. Tepedelenlioglu,

RET Lab Experience Research Summary

Research Background

1. Power prediction in solar farms is well studied
2. In the presence of different faults and under the influence of changing weather conditions, accurate prediction of power is still an open problem.
3. In this project, after the data is collected and analyzed on PV faults, researchers need to know what will result to see if some type of regression will be useful.



re 1.1: Overview of our research vision in Solar Panel Monitoring [4]. Our system i
s shading prediction using irradiance and satellite imaging data, topology reconfigura
ontrolling the switching matrix, and fault classification and diagnosis modules.

RET Research Experiment Setup

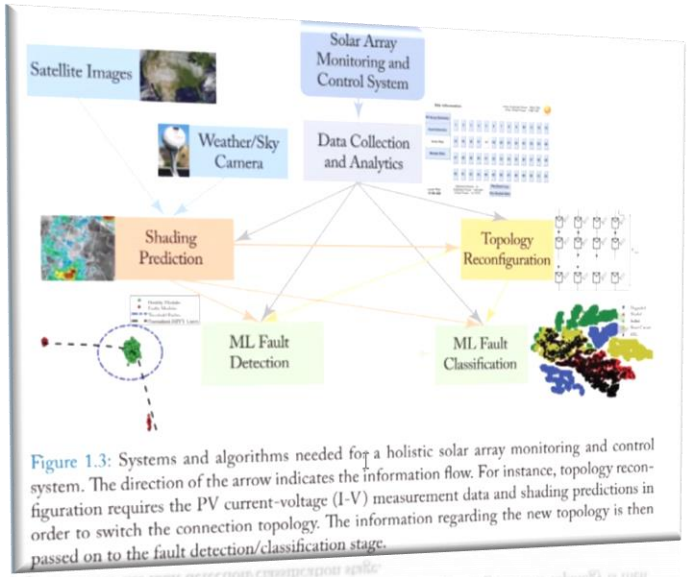
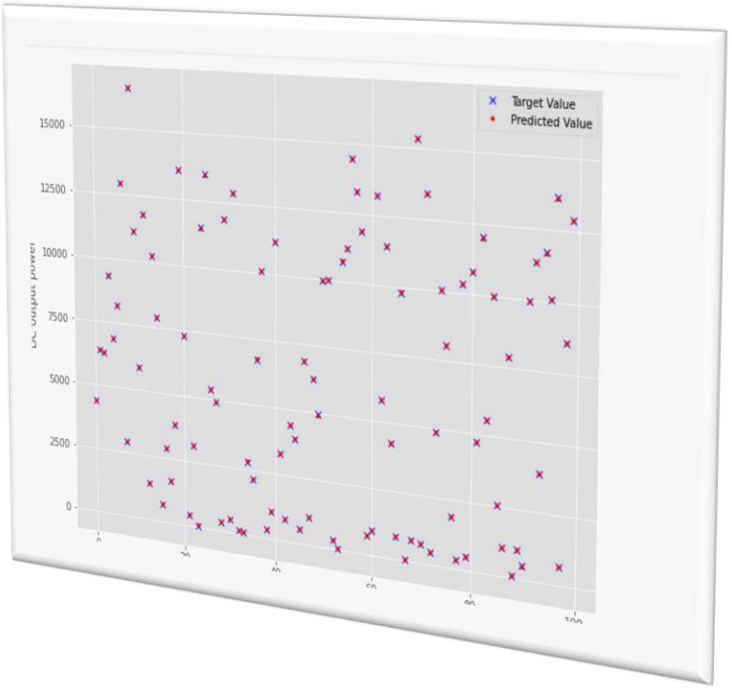


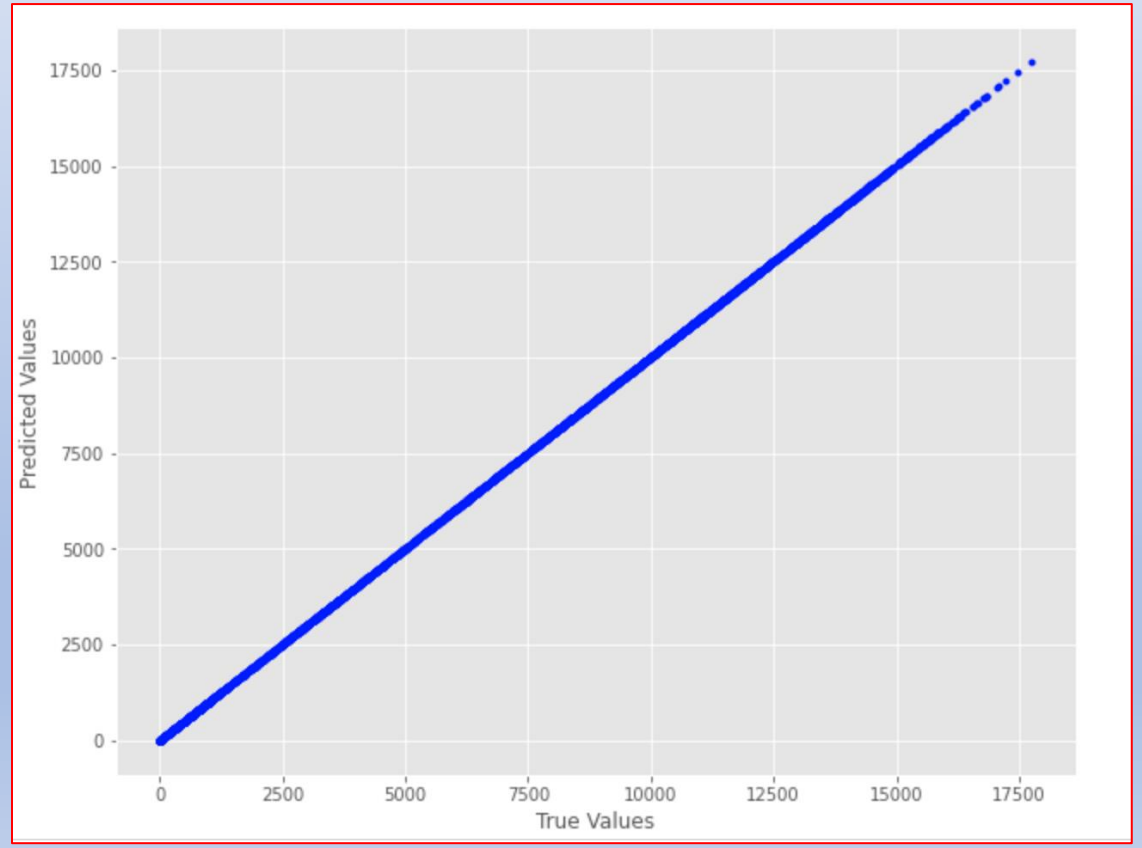
Figure 1.3: Systems and algorithms needed for a holistic solar array monitoring and control system. The direction of the arrow indicates the information flow. For instance, topology reconfiguration requires the PV current-voltage (I-V) measurement data and shading predictions in order to switch the connection topology. The information regarding the new topology is then passed on to the fault detection/classification stage.

- Run Data Simulation
- Linear Regression ML/Neural Network
- Training Models
- DC Power Predictions
- Analyzing Results



RET Research Experiment Results

True Values coincide with Predicted Values



The results are consistent with the hypothesis. After many iterations, the machine learning algorithm learned the optimal linear regression to fit the true values of power outputs.

RET Lab Experience Research Summary

Research Conclusions

- Based on the Start Tech Academy, the linear regression algorithm is a more basic machine learning model that can help beginners to understand more complex regression algorithms and beyond
- This maybe necessary for the purpose of predicting DC power outputs in PV panels.
- Could prove useful to researchers in analyzing PV fault data and definitely for other research in business and academia.

RET Instructional Lesson Implementation

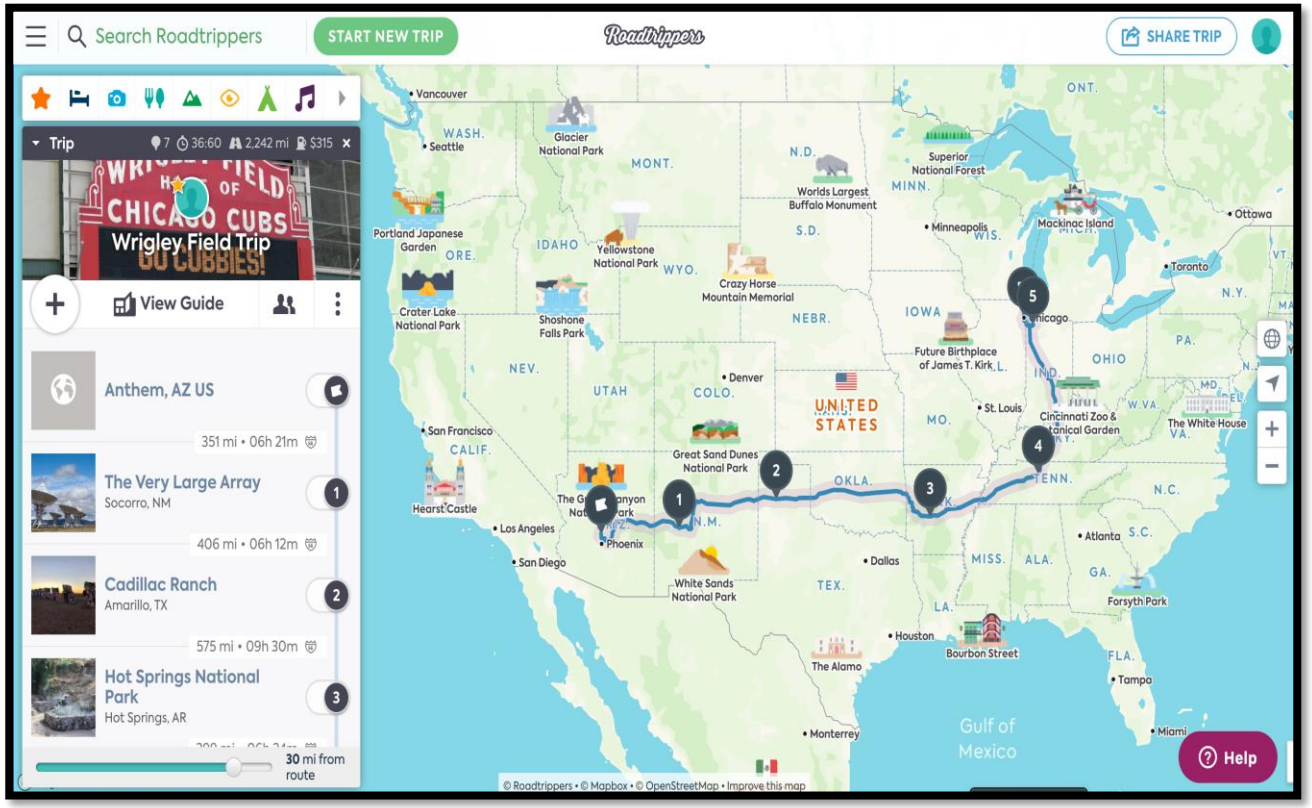
Lesson Objectives

How do you intend to connect the real-world problem you researched in the lab to a learning experience for students?

- Linear regression is applicable to many algebra based courses. In Algebra 2, students already have prerequisite knowledge of linear regression and will use it for statistical analysis in a Mileage versus Cost problem.

What specific learning objectives will be met through the lesson?

- Students will be able to develop data from a simulated road trip and plot in excel and Google Colab, to find the line of best fit and make cost predictions by extrapolating using the model.



RET Instructional Lesson Implementation

Lesson Description

i. Students navigate the fundamentals of machine learning through linear regression.

ii. Students first collect data for a long trip that is at least 1000 miles.

iii. They organize the data in an excel sheet to calculate simple totals for miles driven and total gas costs.

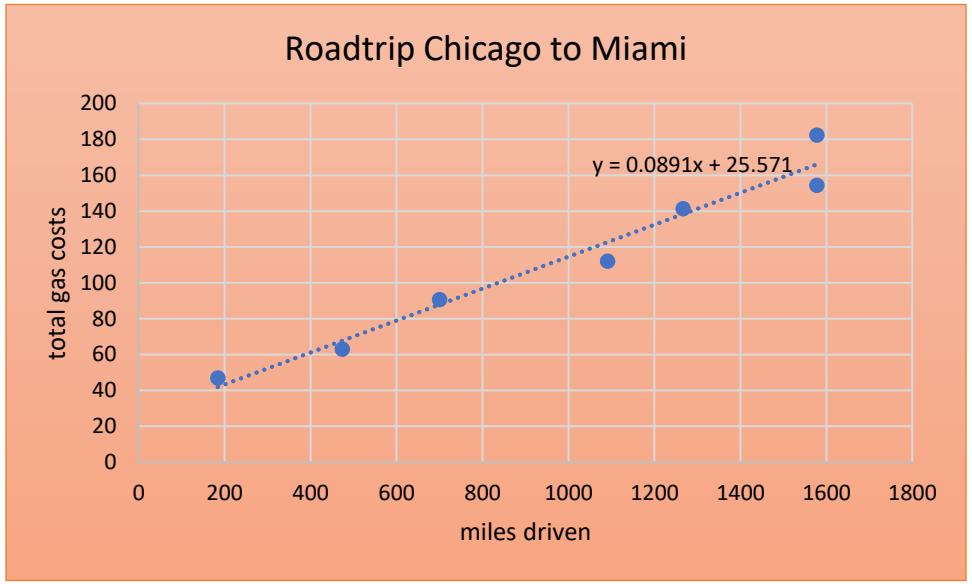
iv. Students create a scatter plot to ultimately find the line of best fit.

v. Students use the line to make predictions (interpolating and extrapolating).

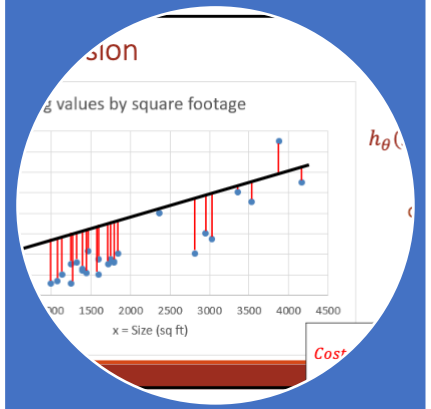
vi. Students use machine learning to find the line of best fit. They upload their data to Colabs, then run the program.

vii. Students compare the linear regressions from excel and Colabs.

viii. Assessment will be a formal presentation of findings and explanation of linear regression as it relates to their specific data set.



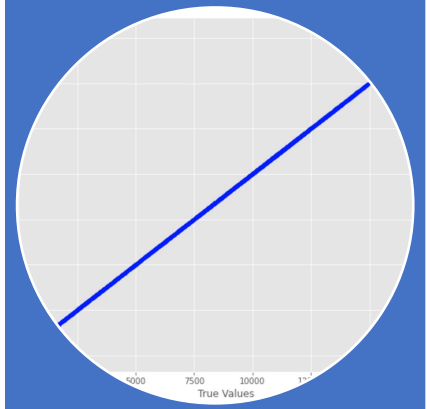
Reflection on RET Summer Experience



TRAINING

effective lectures

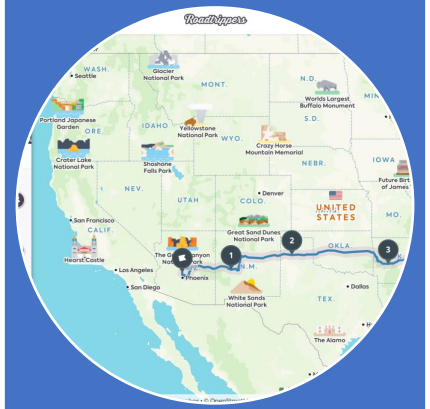
need more interactive videos



RESEARCH

excellent support

struggled with independent research



Lesson Plans

excellent support

great tools for PBLs

Networking

-online environment slowed the process

-mentors & participants provided consistent support

-Final Report and Poster

-Delivering Lecture in Class

-Evaluation

Questions & Feedback

Teacher Input:

- What are the essential questions students need to answer by the end of the lesson?
- What are some potential problems to consider if this is to be a PBL?
- What's one thing you liked about the lesson and one thing you think I should still be thinking about?

References

- [1] Braun, Henry, et al. Signal Processing for Solar Array Monitoring, Fault Detection, and Optimization. 1st ed., vol. 4, Morgan & Claypool Publishers, 2012, doi:10.2200/S00425ED1V01Y201206PEL004.
- [2] Academy, Start-Tech. Machine Learning for Beginners: Linear Regression Model in R. 1st edition, Packt Publishing, 2020.
- [3]Rao, Katoch, et al. Machine Learning for Solar Array Monitoring, Optimization, and Control., Morgan & Claypool Publishers, 2020, DOI 10.2200/S01027ED1V01Y202006PEL013