RET Project: A Hybrid Regression Algorithm for Solar Power Forecasting
July 2020

Kathy Turner, EdD

11th Grade AP Computer Science
Sahuarita High School
Mentors: Graduate Student Kristen Jaskie, Jean Larson, Dr. Andres Spanias

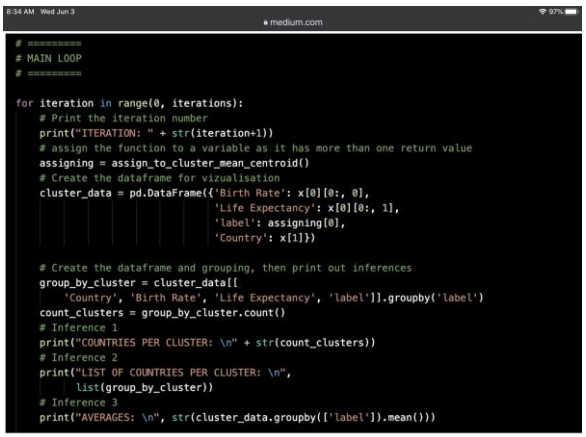
NSF Award 1953745

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

RET Research and Training

Introduction to Python

1. A variable can be used for any type
2. Arrays from Python lists
3. Creating arrays from scratch
4. Compare and contrast: xlrd, panda,list and numpy to determine which type of array may be best to use, and what function should be used to import data file.

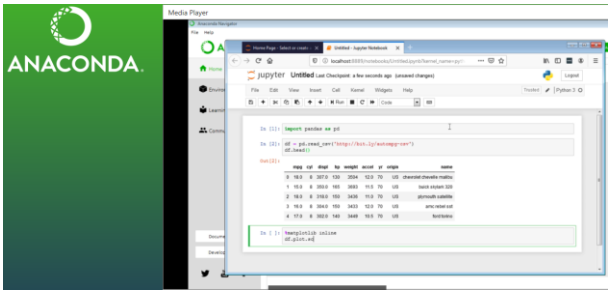


https://sensip.engineering.asu.edu/ret/

RET Schedule and Training

Hands On Technical Training

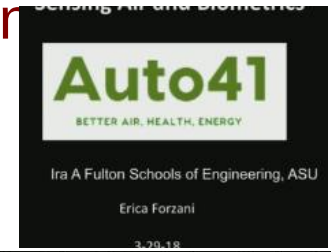
- Machine learning is new and challenging for me
- I was introduced to online platform
- During the daily afternoon meetings I learned about code notebooks. This was the most exciting for me because I had not used them before.

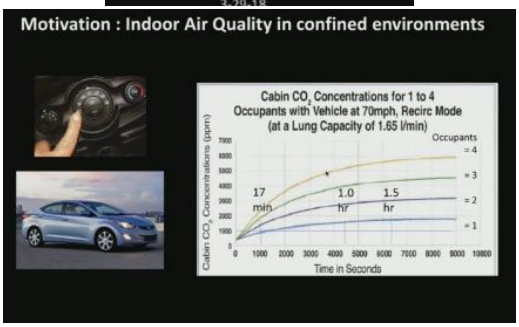


RET Schedule and Trainin

Technical Exposition

- There were 4 to 5 video presentations per week to watch.
- Guest speakers each week gave insight into what ML looks like in the real world practices.
- Topics on audio, motion, and video sensing were informing. The Arduino presentation was very enlightening.

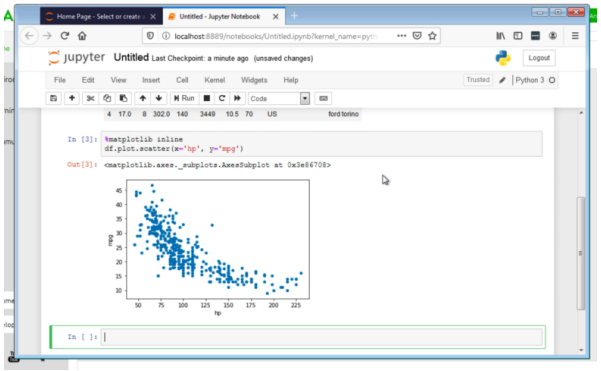




RET Schedule and Training

Research Materials

- This summer was my first experience with canvas, Colab, and scikit.
- The scikit notebooks and the practice exercises using it was an experience that left me wanting to learn more.
- The 5 weeks went so quickly. I will be continuing with my learning journey and share it with my students.



RET Lab Experience Research Summary

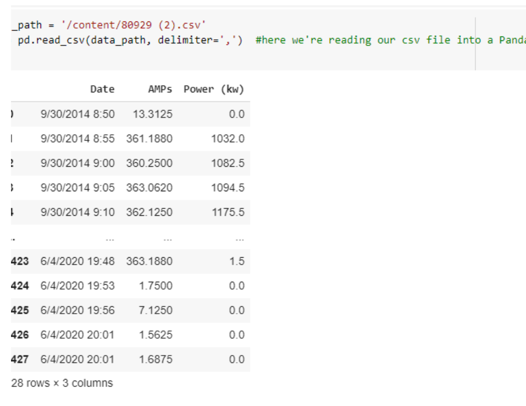
Research Objectives

- The collection of my data is where the bulk of time was spent. I have 30 months of data saved one day at a time.
- My problem is to find how much energy is produced on average for each day during different weather condition.
- Output data is then used to predict future energy production.

RET Lab Experience Research Summary

Research Background

- There is little to no information on residential property solar energy production.
- Most solar energy information found was on commercial properties and energy stations.
- Frequently, articles on load forecasting was found.



```
_path = '/content/80929 (2).csv'
 pd.read_csv(data_path, delimiter=',')  #here we're reading our csv file into a Pandas
```

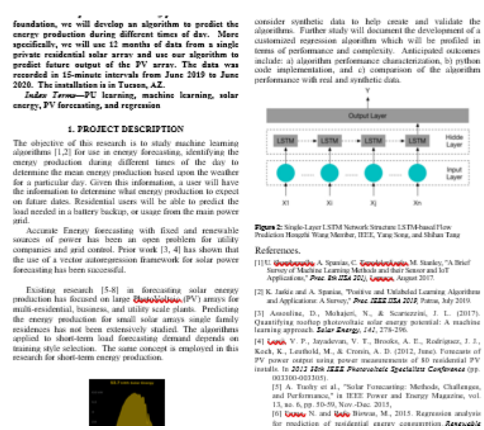
	Date	AMPs	Power (kw)
	9/30/2014 8:50	13.3125	0.0
	9/30/2014 8:55	361.1880	1032.0
	9/30/2014 9:00	360.2500	1082.5
	9/30/2014 9:05	363.0620	1094.5
	9/30/2014 9:10	362.1250	1175.5
...	...	...	...
423	6/4/2020 19:48	363.1880	1.5
424	6/4/2020 19:53	1.7500	0.0
425	6/4/2020 19:56	7.1250	0.0
426	6/4/2020 20:01	1.5625	0.0
427	6/4/2020 20:01	1.6875	0.0

28 rows × 3 columns

RET Lab Experience Research Summary

Research Proposal

- Literature review was performed for two weeks. Finding 8 references was difficult.
- The abstract portion was difficult. This is my first proposal and all my previous experiences in writing abstracts were after completion of work.



foundation, we will develop an algorithm to predict the energy production during different times of day. More specifically, we will use 12 months of data from a single private residential solar array and use our algorithm to predict future output of the PV array. The data was recorded in 15-minute intervals from June 2019 to June 2020. The installation is in Tucson, AZ.

***Index Terms*—PU learning, machine learning, solar energy, PV forecasting, and regression**

1. PROJECT DESCRIPTION

The objective of this research is to study machine learning algorithms [1,2] for use in energy forecasting, identifying the energy production during different times of the day to determine the mean energy production based upon the weather for a particular day. Given this information, a user will have the information to determine what energy production to expect on future dates. Residential users will be able to predict the load needed in a battery backup, or usage from the main power grid.

Accurate Energy forecasting with fixed and renewable sources of power has been an open problem for utility companies and grid control. Prior work [3, 4] has shown that the use of a vector autoregression framework for solar power forecasting has been successful.

Existing research [5-8] in forecasting solar energy production has focused on large PhotoVoltaic (PV) arrays for multi-residential, business, and utility scale plants. Predicting the energy production for small solar arrays single family residences has not been extensively studied. The algorithms applied to short-term load forecasting demand depends on training style selection. The same concept is employed in this research for short-term energy production.

consider synthetic data to help create and validate the algorithms. Further study will document the development of a customized regression algorithm which will be profiled in terms of performance and complexity. Anticipated outcomes include: a) algorithm performance characterization, b) python code implementation, and c) comparison of the algorithm performance with real and synthetic data.

Figure 2: Single-Layer LSTM Network Structure LSTM-based Flow Prediction Hongzhi Wang Member, IEEE, Yang Song, and Shihan Tang

References.

[1] U. Shanthamallu, A. Spanias, C. Tepedelenlioglu, M. Stanley, "A Brief Survey of Machine Learning Methods and their Sensor and IoT Applications," *Proc. 8th IISA 2017*, Larnaca, August 2017.

[2] K. Jaskie and A. Spanias, "Positive and Unlabeled Learning Algorithms and Applications: A Survey," *Proc. IEEE IISA 2019*, Patras, July 2019.

[3] Assouline, D., Mohajeri, N., & Scartezzini, J. L. (2017). Quantifying rooftop photovoltaic solar energy potential: A machine learning approach. *Solar Energy, 141*, 278-296.

[4] Lonij, V. P., Jayadevan, V. T., Brooks, A. E., Rodriguez, J. J., Koch, K., Leuthold, M., & Cronin, A. D. (2012, June). Forecasts of PV power output using power measurements of 80 residential PV installs. In *2012 38th IEEE Photovoltaic Specialists Conference* (pp. 003300-003305).

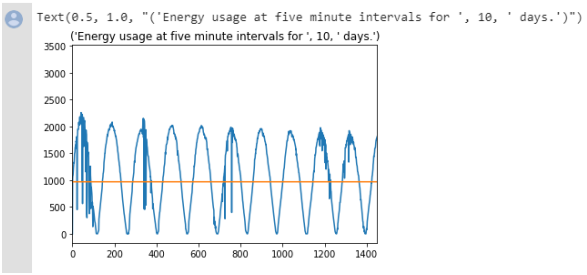
[5] A. Tuohy et al., "Solar Forecasting: Methods, Challenges, and Performance," in IEEE Power and Energy Magazine, vol. 13, no. 6, pp. 50-59, Nov.-Dec. 2015,

[6] Fumo, N. and Rafe Biswas, M., 2015. Regression analysis for prediction of residential energy consumption. *Renewable*

RET Lab Experience Research Summary

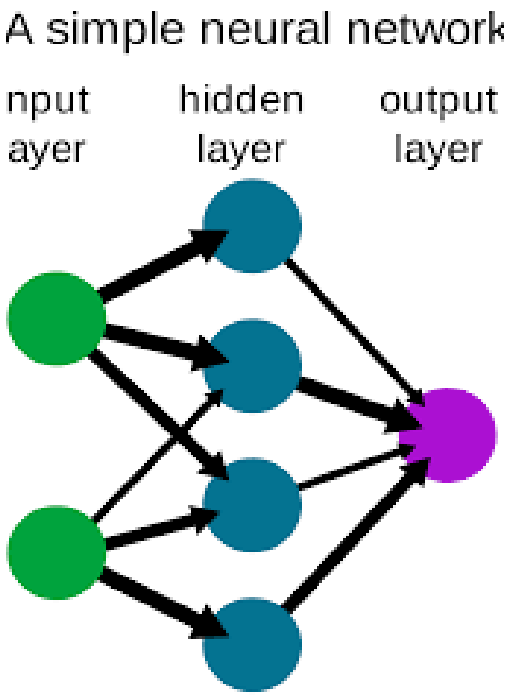
Research Conclusions

- Data from Cyprus was used to begin the process of finding the algorithms to use to learn, validate and predict.
- Linear progression did not appear work well on a large array of data.
- I still have work to do. But, the ability to predict the energy production can aid in buying the proper storage facility.
- Future of solar may not be so grid dependent.



Next STEPS in Research

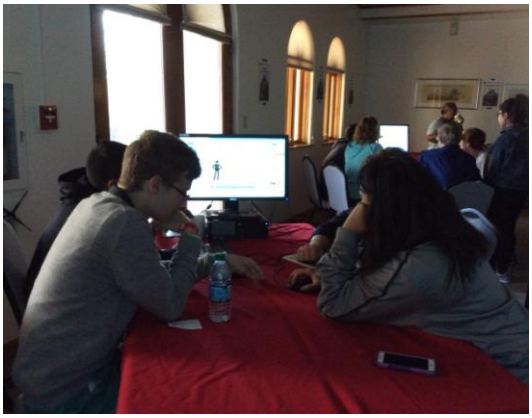
- Retrieve the remainder of the solar data dating back to 15 July 2018
- Use data with existing algorithms and newly developed ML algorithms
- Plan to write final report
- Plan for completing results for publication
- Engagement with SenSIP



RET Instructional Lesson Implementation

Lesson Objectives

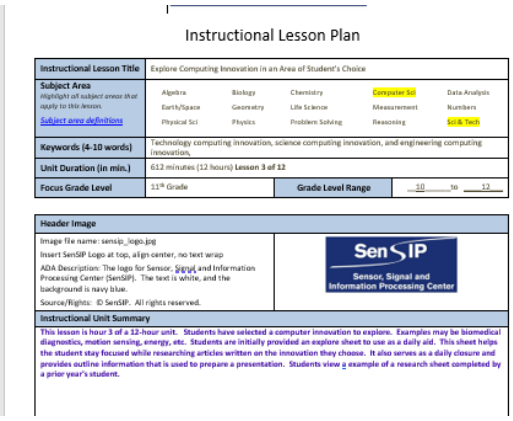
- The information learned will be the basis of a 12-lesson unit on exploring IoT and ML research.
- Students will explore an emerging technology of their choice.
- Students will learn the vocabulary of computer programming.
- Students will learn to write a technical paper of their findings.



RET Instructional Lesson Implementation

Lesson Description

- Lesson 3 of 12 has been created and focuses on researching literature.
- Three current articles (less than 12 months old) are required.
- Students will predict the type of algorithms, functions, and data that relate to their topic.
- A daily checklist is used for lessons 1 – 12 as formative assessment.



Instructional Lesson Plan

Instructional Lesson Title	Explore Computing Innovation in an Area of Student's Choice				
Subject Area *Highlight all subject areas that apply to this lesson.* *Subject area definitions*	Algebra	Biology	Chemistry	Computer Sci	Data Analysis
	Earth/Space	Geometry	Life Science	Measurement	Numbers
	Physical Sci	Physics	Problem Solving	Reasoning	Sci & Tech
Keywords (4-10 words)	Technology computing innovation, science computing innovation, and engineering computing innovation.				
Unit Duration (in min.)	612 minutes (12 hours) **Lesson 3 of 12**				
Focus Grade Level	11th Grade	**Grade Level Range**	10 to 12		

Header Image	
Image file name: sensip_logo.jpg Insert SenSIP Logo at top, align center, no text wrap ADA Description: The logo for Sensor, Signal and Information Processing Center (SenSIP). The text is white, and the background is navy blue. Source/Rights: © SenSIP. All rights reserved.	SenSIP Sensor, Signal and Information Processing Center
Instructional Unit Summary	
This lesson is hour 3 of a 12-hour unit. Students have selected a computer innovation to explore. Examples may be biomedical diagnostics, motion sensing, energy, etc. Students are initially provided an explore sheet to use as a daily aid. This sheet helps the student stay focused while researching articles written on the innovation they choose. It also serves as a daily closure and provides outline information that is used to prepare a presentation. Students view a example of a research sheet completed by a prior year's student.	

Questions & Feedback

- Optional: Is there an open-ended question you can ask the audience to help provide useful feedback?
- For example: What would you recommend I be thinking about as I prepare to implement this lesson with students?
- For example: What's one thing you liked about the lesson and one thing you think I should still be thinking about?

Self Assessment

- What worked well in this program, what did not work so well?
- Skill building gained - theory / software
- Research knowledge gained
- Ability to express research in abstract and presentation
- etc

References

- [1] Assouline, D., Mohajeri, N., & Scartezzini, J. L. (2017). Quantifying rooftop photovoltaic solar energy potential: A machine learning approach. *Solar Energy*, *141*, 278-296.
- [2] Lonij, V. P., Jayadevan, V. T., Brooks, A. E., Rodriguez, J. J., Koch, K., Leuthold, M., & Cronin, A. D. (2012, June). Forecasts of PV power output using power measurements of 80 residential PV installs. In *2012 38th IEEE Photovoltaic Specialists Conference* (pp. 003300-003305).
- [3] A. Tuohy et al., "Solar Forecasting: Methods, Challenges, and Performance," in IEEE Power and Energy Magazine, vol. 13, no. 6, pp. 50-59, Nov.-Dec. 2015,
- [4] Fumo, N. and Rafe Biswas, M., 2015. Regression analysis for prediction of residential energy consumption. *Renewable and Sustainable Energy Reviews*, 47, pp.332-343.
- [5] H. Yang, C. Huang, Y. Huang and Y. Pai, "A Weather-Based Hybrid Method for 1-Day Ahead Hourly Forecasting of PV Power Output," in IEEE Transactions on Sustainable Energy, vol. 5, no. 3, pp. 917-926, July 2014,
- [6] J. Huang, Y. Liang, H. Bian and X. Wang, "Using Cluster Analysis and Least Square Support Vector Machine to Predicting Power Demand for the Next-Day," in IEEE Access, vol. 7, pp. 82681-82692, 2019
- [7] U. Shanthamallu, A. Spanias, C. Tepedelenlioglu, M. Stanley, "A Brief Survey of Machine Learning Methods and their Sensor and IoT Applications," *Proc. 8th IISA 201),* Larnaca, August 2017.
- [8] K. Jaskie and A. Spanias, "Positive and Unlabeled Learning Algorithms and Applications: A Survey," *Proc. IEEE IISA 2019,* Patras, July 2019.
- [9] A. Spanias, Machine Learning Workforce Development Programs on Health and COVID-19*, Proc. IEEE IISA 2020*, Piraeus, July 2020.
- [10] D. Smith, K. Jaskie, J.Cadigan, J. Marvin, and A. Spanias, "Machine Learning for Fast Short-Term Energy Load Forecasting," *IEEE ICPS 2020*, Tampere, June 2020.

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