SenSIP
Sensor, Signal and Information Processing Center

NSF Research Experience for TEACHERS (RET)

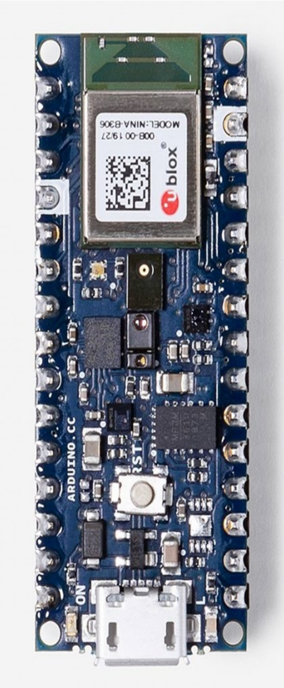
RET Project: The Effect of Bias in Training Data using the APDS996 Color Sensor on the Arduino Nano 33 BLE Sense Board

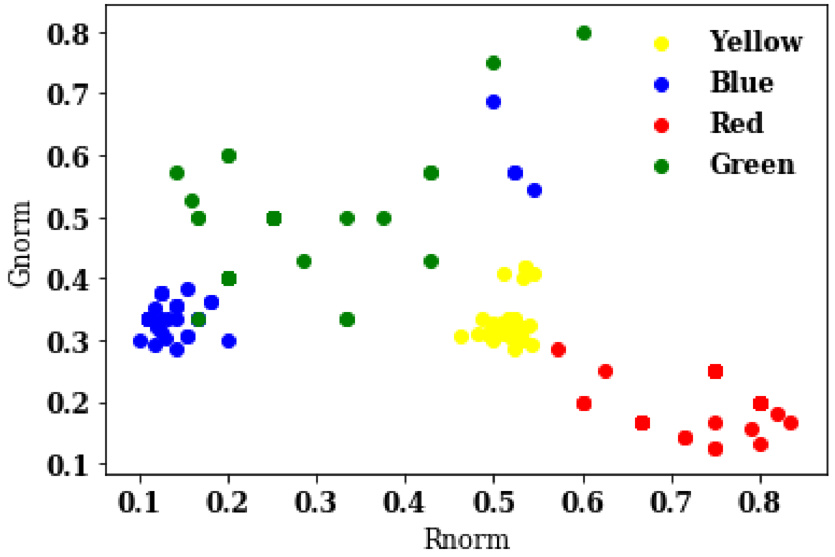
Brian Hawkins

High School Engineering
Corona Del Sol High School, Tempe, AZ
Mentors: Michael Stanley, Dr. Andreas Spanias

NSF Award 1953745

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



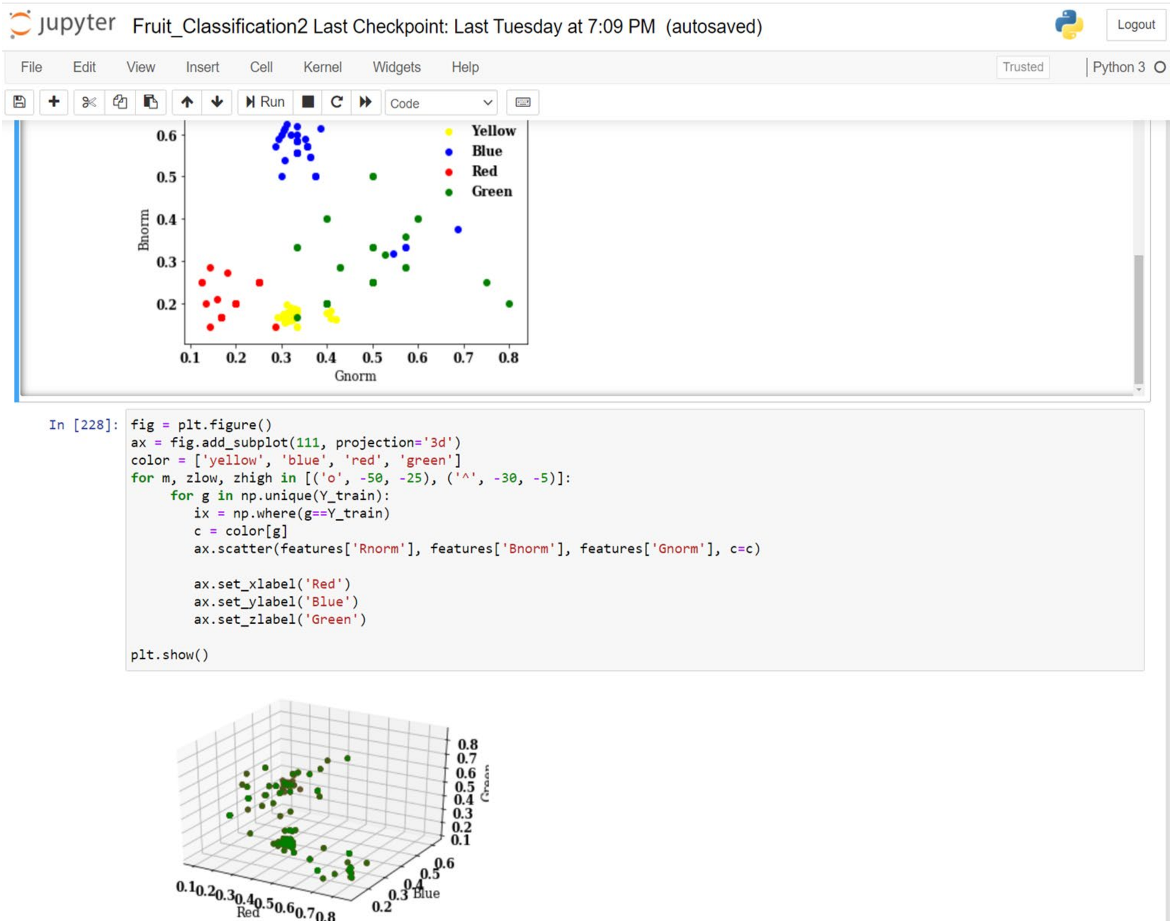


RET Research and Training

RET Schedule and Training

Hands On Technical Training

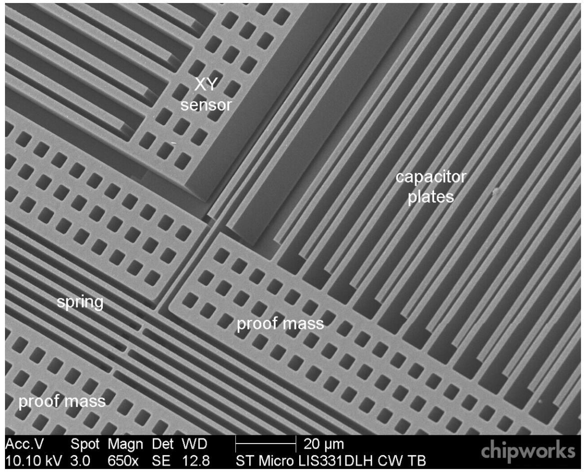
- Arduino Integrated Development Environment (IDE)
  - Color sensor
  - Proximity sensor
- Jupyter was used for machine learning and graphing
  - Code management
  - Python
  - Support Vector Machine
    - Support vectors, C, gamma

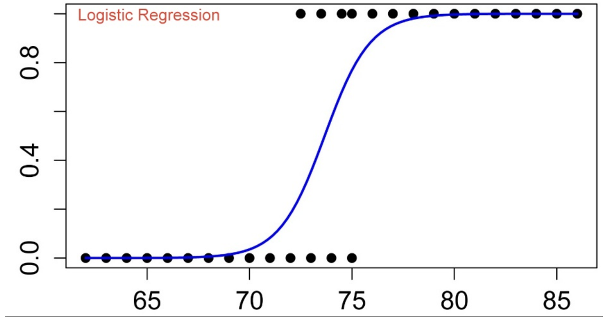


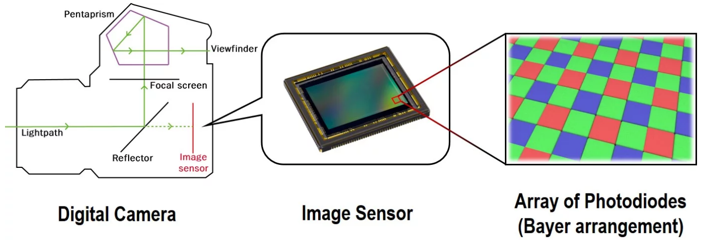
RET Schedule and Training

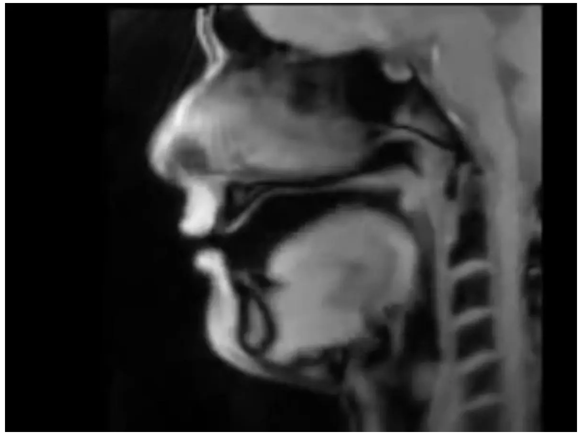
Technical Exposition

- How sensors worked with derivation of formulas
- Kernal and sigmoid with probabilities
- Applications to speech and the complexity of speech
- Indoor air quality in cars
- How photodiodes scan and process data





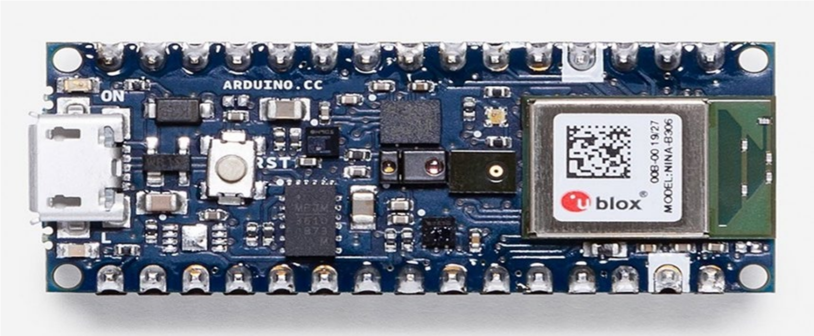




RET Schedule and Training

Research Materials

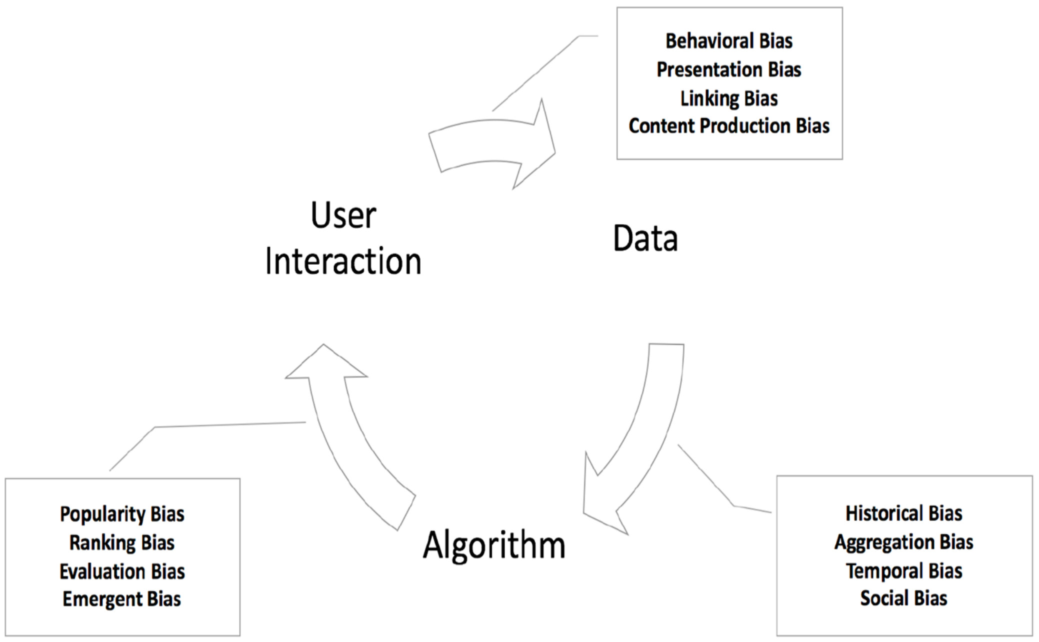
- Mentor (Michael Stanley) - used extensively sometimes several times a day
- SciKit Learn - terminology and syntax of libraries
- Capabilities of Nano Sense board
- ASU online library - research potential topics of interest



RET Lab Experience Research Summary

Research Objectives

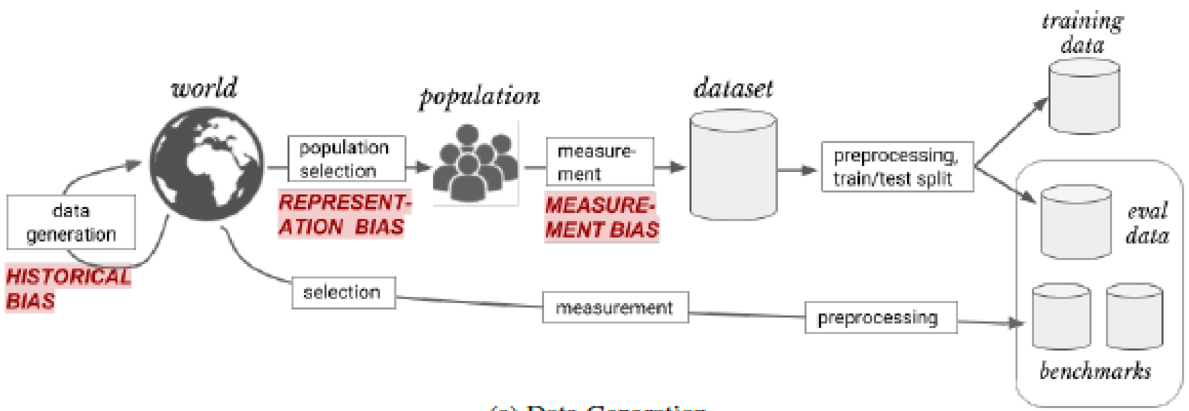
- What are the different ways that machine learning can be biased?
- What metrics are used to determine bias?
  - How does the proximity sensor work?
  - Types of data to collect with RGB sensor.



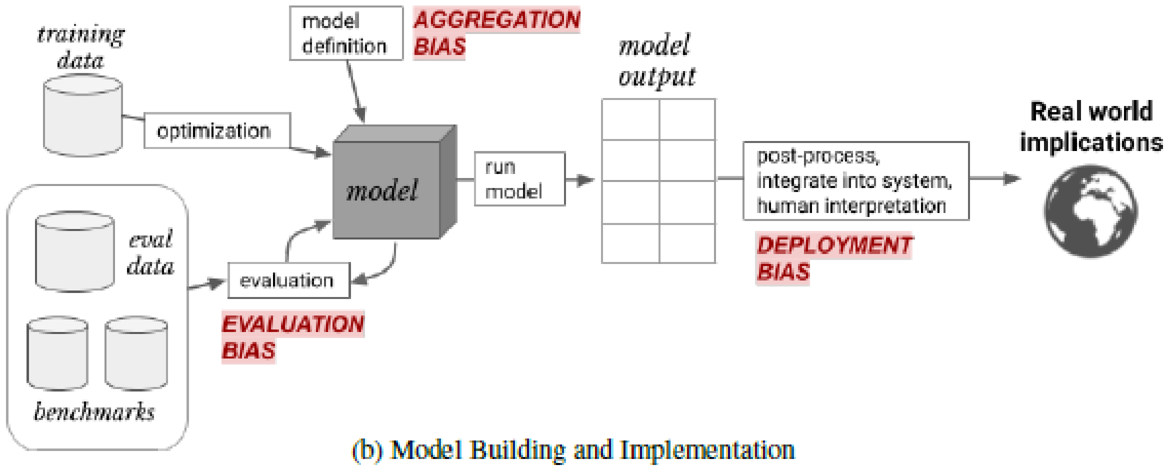
RET Lab Experience Research Summary

Research Background

- Normalization
- Under and Overfitting
- Equal data sets
- Sampling bias
- Data snooping
- Population bias



(a) Data Generation



(b) Model Building and Implementation

RET Lab Experience Research Summary

Research Proposal

- Difficult to pick a topic initially
- Focused too much on data collection instead of machine learning
- Abstract was straightforward

The Effect of Bias in Training Data using the APDS996 Color Sensor on the Arduino BLE 33 Sense Board

Brian Hawkins, RET

Abstract – **Machine learning is only as good as the data that creates it and given the mistrust of AI and big data by the public recently, it is useful to explore the scenarios where machine learning might give misleading results.**

Index Terms---**machine learning, supervised learning, classification**

1. PROJECT DESCRIPTION

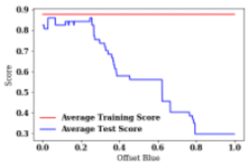
In supervised classification, training samples from known data distributions are tested with the goal to identify a classification boundary that separates these classes. New data samples with known labels can then be identified as belonging to one of these classes. Support vector machines (SVMs), decision tree, random forest, and others are used to solve this problem.

The training data in this scenario is provided by standard data sets, common objects meeting the color criteria, or phone applications capable of creating specific color profiles. Each of these situations will be explored for unbiased and biased data. Biased data will be created by not normalizing the data, sampling unequal data sets, and intentionally creating training sets that are not representative of the population to be modeled.

Many problems are susceptible to various types of bias. Measurement bias is well known in the recidivism risk prediction tool COMPAS which was a factor leading to higher false positive rates for black versus white defendants. [1]. Population basis has been documented in ImageNet where 45% of the images are from the United States and a majority of the remaining portion are from North America or Western Europe while 3.2% are from China and India combined. [1]

In this project, I first performed a literature review of the existing work in this area, examples of its use, and possible future application. I then explored the limitations of the color sensor on the Arduino BLE 33 Sense Board in an attempt to design a scenario that would produce repeatable results during training and testing.

In testing support vector machines and decision trees on a 4 color classification problem, it was determined that a default state needed to be determined to prevent incorrect identification of colors during testing. It was also determined that while the decision tree consistently gave accurate results, the support vector machine did not until the radius basis function was used as the kernel and the values of gamma and C were adjusted. So while it is expected that unequal and non-normalized data will yield a non-ideal confusion matrix, it may not always occur depending on choices made for parameters during machine learning.



[1] Suresh, H., & Guttag, J. (2019). A Framework for Understanding Unintended Consequences of Machine Learning. *ArXiv, abs/1901.10002.*

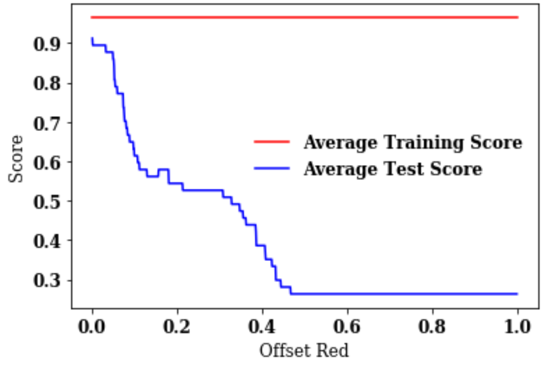
[2] Mehrabi, Ninareh & Morstatter, Fred & Saxena, Nripsuta & Lerman, Kristina & Galstyan, Aram. (2019). A Survey on Bias and Fairness in Machine Learning.

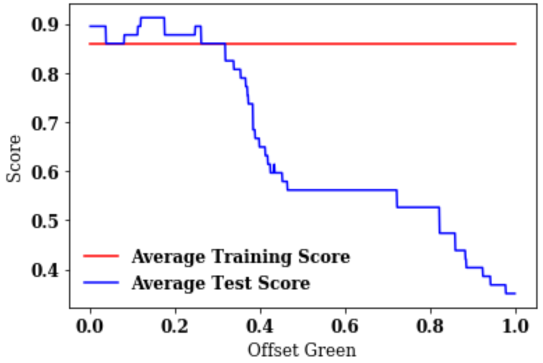
[3] Olteanu, A., Castillo, C., Diaz, F., & Kıcıman, E. (2019). Social Data: Biases, Methodological Pitfalls, and Ethical Boundaries. *Frontiers in big data*, *2*, 13.

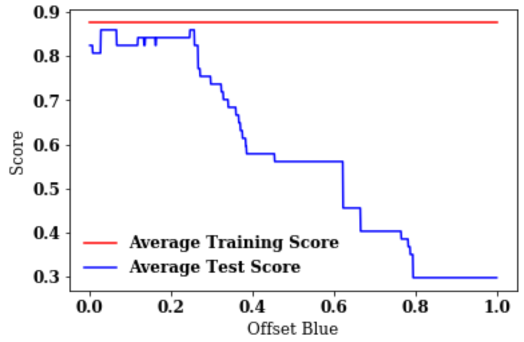
RET Lab Experience Research Summary

Research Conclusions

- Defining bias and fairness is difficult
- Color sensor data is uniform and has clearly defined classes so noise or an offset was necessary
- For students, it is meant to expose them to data that may not be clearly separated that so they can see possible indicators in the graphs
- Metrics are necessary to tell the user if their machine learning is doing what they expect it to do
- Generalization error

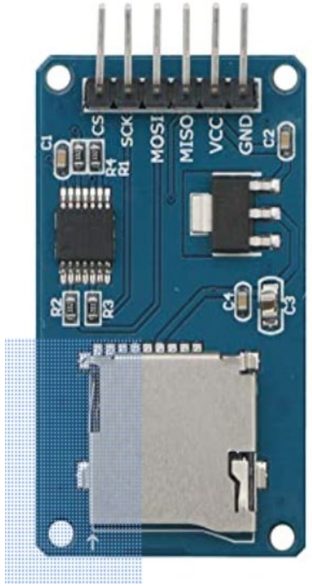


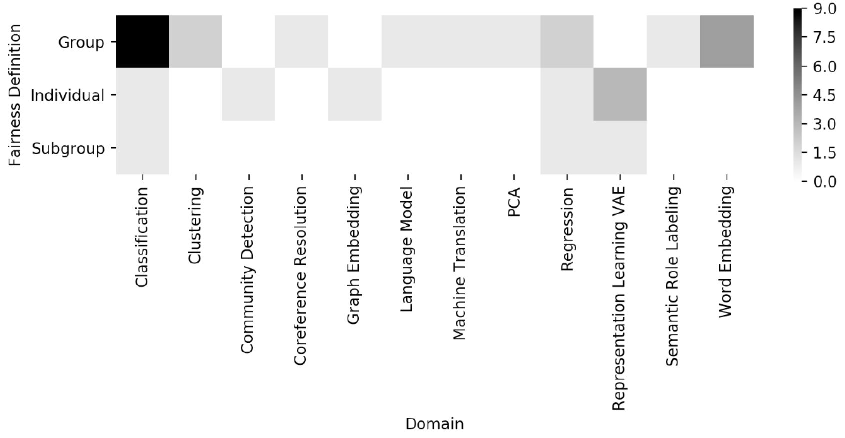




Next STEPS in Research

- Examine more specific metrics that could be used to quantify the margin
- I mostly need to work on programming and creating additional graphs
- Possible work on margin distribution
- Original plan - Nano board, battery, micro SD card reader.
  - 3D printed enclosure

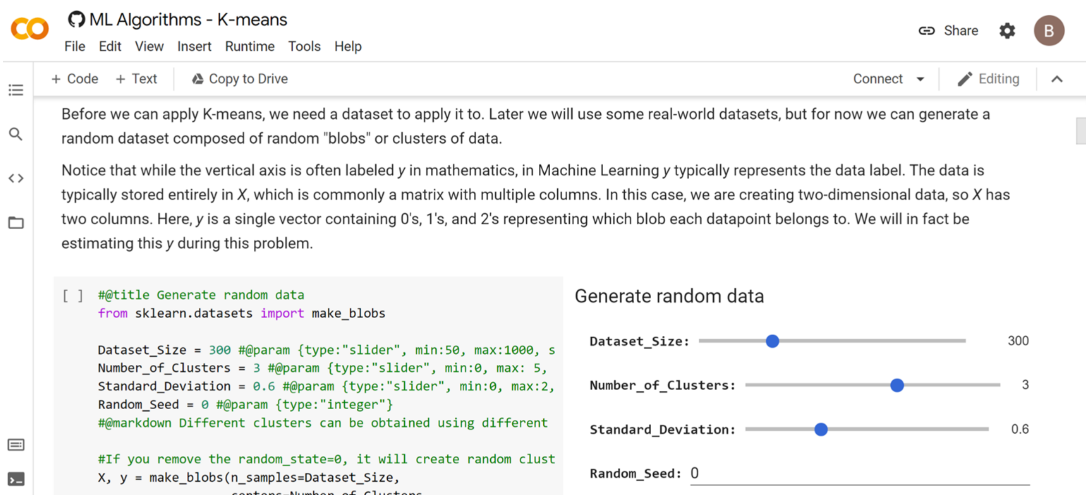


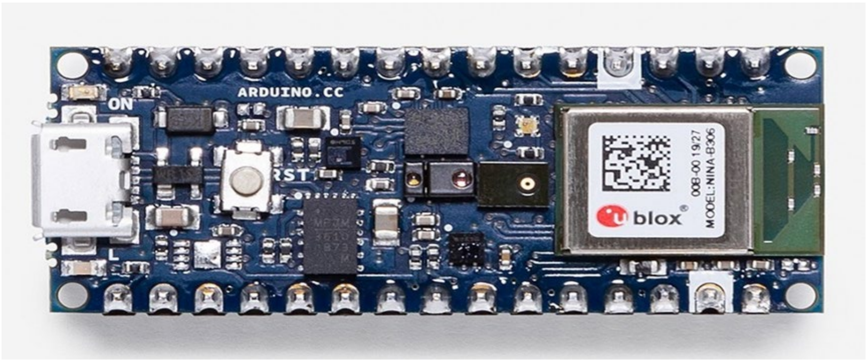


RET Instructional Lesson Implementation

Lesson Objectives

- Run the code in Google Colab to help students understand machine learning by examining confusion matrices and corresponding scatter plots.
- Test various data sets using the Support Vector Machine algorithm on Google Colab to create a confusion matrix and scatter plots.
- Extension - test their machine learning algorithms using items of student's choice



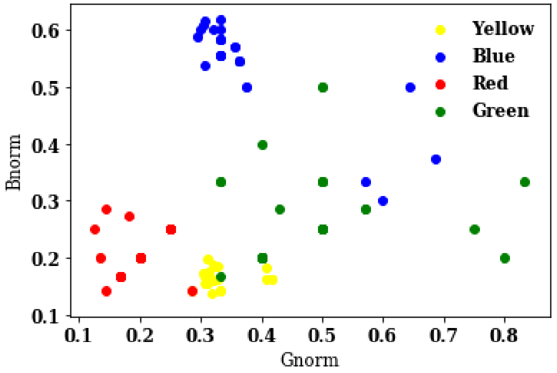


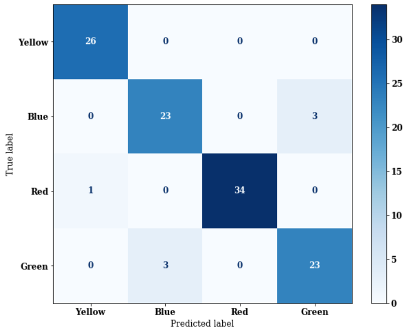
RET Instructional Lesson Implementation

Lesson Description

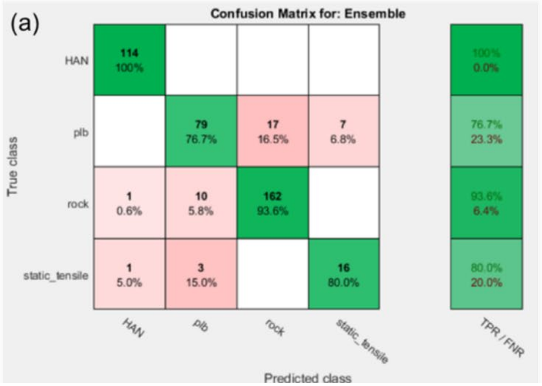
Bias in Machine Learning

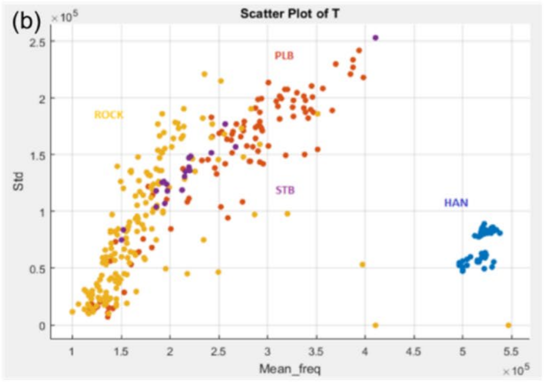
- Students will be working in groups to use the support vector algorithm to create confusion matrices and scatter plots on a set of given data. Groups will compare their results to the results of other groups.
- My goal is to have students understand a confusion matrix and corresponding scatter plots and how data bias can affect the outcome of a machine learning algorithm
- Most of the assessment will be formative while students are working, but students can have a quiz on confusion matrices and a short evaluation of the data that I provided their group.





Example of what I want students to see





Questions & Feedback

- What sorts of scaffolding opportunities do you think I should provide students, or any other suggestions?

Self Assessment

- The open ended nature of the program worked well and also didn’t.
  - Difficulty picking a problem, misunderstanding of machine learning, too much focus on the sensor and data collection, Python background
  - Early access to code samples
- Machine learning algorithms and terminology
- Code management in Jupyter
- ASU online libraries and Mendeley
- Flexible nature of the program
- One page summary and elevator pitches for students

References

Research Objectives

[1] Suresh, H., & Guttag, J. (2019). A Framework for Understanding Unintended Consequences of Machine Learning. ArXiv, abs/1901.10002.

[2] Mehrabi, Ninareh & Morstatter, Fred & Saxena, Nripsuta & Lerman, Kristina & Galstyan, Aram. (2019). A Survey on Bias and Fairness in Machine Learning.

[3] Olteanu, A., Castillo, C., Diaz, F., & Kıcıman, E. (2019). Social Data: Biases, Methodological Pitfalls, and Ethical Boundaries. *Frontiers in big data*, *2*, 13. https://doi.org/10.3389/fdata.2019.00013

[4]Clark, C., Yatskar, M., & Zettlemoyer, L. (2019). Don't take the easy way out: Ensemble based methods for avoiding known dataset biases. *arXiv preprint arXiv:1909.03683*.

Thank You

- Andreas Spanias
- Michael Stanley
- Kristen Jaskie
- Jean Larson
- Ruby Sayed

Best excerpt from my research

Disclaimer: No graduate students were harmed in the production of this paper. Authors are listed in order of increasing procrastination ability.

If you are interested in looking at bias at all:
https://github.com/dssg/aequitas
https://dssg.github.io/fairness_tutorial/
https://textbook.coleridgeinitiative.org/chap-bias.html
https://fairmlbook.org/
https://www.amazon.com/Ethics-Data-Science-Mike-Loukides-ebook/dp/B07GTC8ZN7
https://www.solveforgood.org/