Introduction and Problem Statement:

* Renewable energy beneficial but causing problems
* Limited data and knowledge of the grid making it difficult to assess impact renewable have
* Methods show it is possible to build models using imperfect data (my paper, yangs topology paper).
* Utilities are limited by their control over behind the meter devices
* Hosting capacity problem
* Home energy storage is going to become a necessary in order to increase hosting capacity and has the potential to reduce costs for utilities and customers.
* Limited or costly communication networks
* How can utilities use the data they have available to increase the hosting capacity and maintain safety and reliability? Specifically with behind the meter devices that they have little or no control over.
* Machine learning has the potential to be able to do this.
* Customers won’t want to give control of their battery to the utility.

Hypothesis:

Using machine learning, AMI data can be analyzed to identify customers who have behind the meter generation or storage devices. Knowing the topology of the system, utilities will be able to better plan for renewable generation and involve customers to improve the reliability and operation of the grid.

Methodology and Testing:

* Machine learning can be used to analyze data that utilities have available and recreate topology only using AMI data.
* Machine learning also has the potential to identify customer who have renewable generation or storage by analyzing AMI data
* Machine learning program to identify customers with high probability of behind the meter devices
  + Using ami data
  + Historical solar data
* Test on AMI data from DLC
* Test in lab at ASU using different control schemes for batteries.

1. Create machine learning program to group customers into different groups
2. Train the algorithm to be able to identify normal customers vs. PV and storage
3. Not a lot of behind the meter storage data,
   1. Create a battery control algorithm. Model a system with a battery. Test add battery to normal AMI data to create AMI data that would look like a customer with storage. Compare the results if the customer had storage/PV of different sizes and didn’t.
   2. Test control on battery in lab.
4. Refine the algorithm to be able to determine the modes of operation of the battery.
5. Generate PV generation curve at the customers location
6. Once identification works, begin applying it to distribution circuit models
7. Test hosting capacity on a circuit using conventional methods
8. Import AMI data into distribution circuit.
9. A

Expected significance and Broader Impact:

* Increase the hosting capacity of solar on the grid, and will save customers and utilities money
* Cheaper energy. Cleaner energy. Greener world.
* Ability for utilities to provide recommendations to customers who could benefit from energy devices at their location.
* Energy eventually becomes a commodity or a free resource.
* Broaden Machine learning techniques