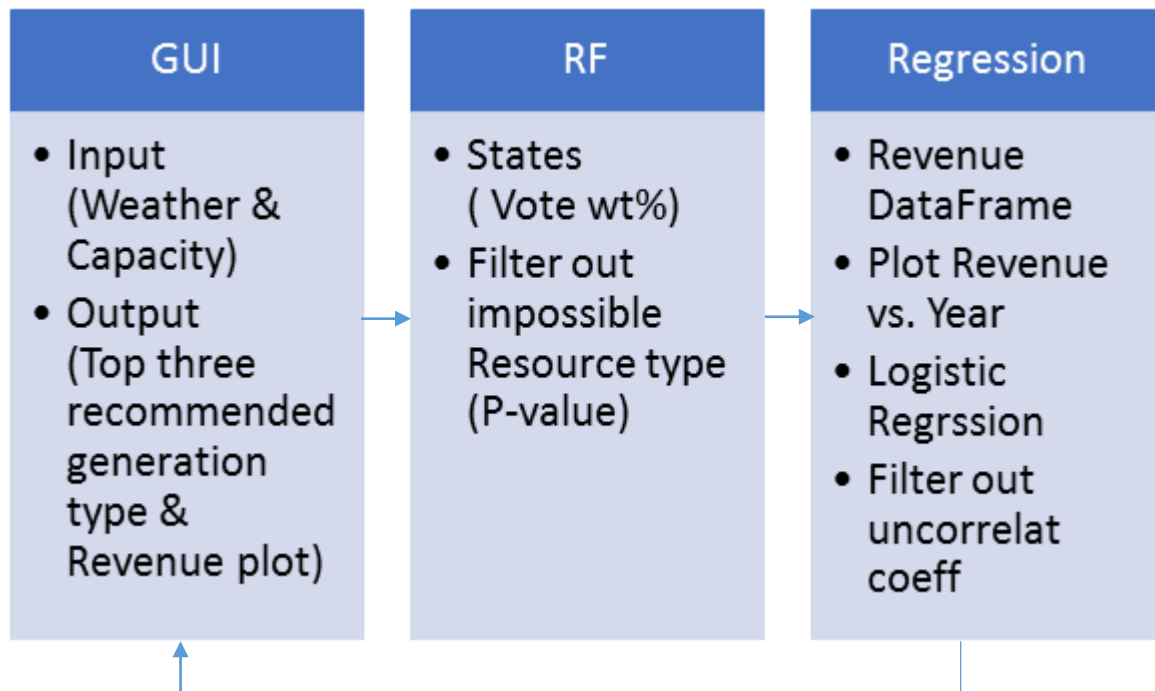


EASE Workflow

Target Users:

Industry users who want to build a self-sustained electricity power plant for their own factories for money saving purpose



GUI:

Input weather data

Fill in Temperature Data:

Local AvgTemp in Summer:	<input type="text"/>	°F
Local AvgTemp in Winter :	<input type="text"/>	°F

Fill in Precipitation Data:

How many days does it rains:
Almost Never Rain: < 1
A few days a year: 1~1.5
A few weeks a year:1.5~2
A few months a year:2~2.5
Half of the time:2.5~3
More than half the time:3~4
Almost every single day: >4

Fill in Wind Speed Data:

How do you feel about wind:
Light Air (Smoke drifts and leaves rustle): <5
Light Breeze (Wind felt on face): 5~6
Gentle Breeze (Flags extended, leaves move): 6~7
Moderate Breeze (Dust and small branches move):7~8
Fresh Breeze (Small trees begin to sway):8~9
Strong Wind (Umbrellas are difficult to control):9~10
Gale (Inconvenience in walking): 10~14

Reference Source: <https://www.unc.edu/~rowlett/units/scales/beaufort.html>

Input electricity generation capacity:

Fill in capacity you need

Capacity you need:
Tiny Capacity: Capacity < 2,500 MW
Small Capacity: 2,500 MW < Capacity < 20,000 MW
Medium Capacity: 20,000 MW < Capacity < 150,000 MW
Large Capacity: 150,000MW < Capacity < 1,000,000 MW
Extreme Capacity: Capacity > 1,000,000 MW

Note: How many homes can 1 megawatt of power supply?

1MW of capacity will produce electricity that equals to about the same amount of electricity consumed by 400 to 900 homes in a year

Machine Learning with RandomForest

Goal: According to the input weather data, filter out the impossible clean energy resources.

Method: Random Forest (Classification)

Procedure:

1. Data mining

- Pre-calculate highest and lowest temperature weather data, termed as summer and winter temp.

Find the highest and lowest temperature and their accordingly months,

Situation one:

If it contains all the data for 12 months, use the temperature data of nearby 3 months to calculate the average temperature as the final Ts and Tw.

Eg.For 2015,CA:

Month	T(°F)
201501	46.5
201502	50.5
201503	54.5
201504	58.5
201505	65.5
201506	71.5
201507	75
201508	74.5
201509	71.5
201510	64
201511	53.5
201512	46
201601	48

highest } $T_{\max} = (71.5 + 75 + 74.5) = 73.3$

lowest } $T_{\min} = (53.5 + 46 + 48) = 49.2$

Situation two:

If it doesn't contain all the data for 12 months, directly use the highest and lowest temperature as their Ts and Tw.

Final DataFrame to be produced prior feeding in for RF training:

Make the final weather dataframe containing Precipitation, Windspeed and Ts & Tw for each state in 1996--2016

	State	Year	TotalMonthlyPrecip	TempSummer	TempWinter	Avgwindspeed
0	AK	1996	2.326250	55.000000	31.320000	13.684615
1	AK	1997	5.604615	53.386970	27.537821	9.979630
2	AK	1998	2.905821	53.293705	17.716720	8.070060
3	AK	1999	3.389773	54.316565	8.305729	8.382022
4	AK	2000	3.018342	52.225000	20.123256	8.205013
5	AK	2001	2.844638	53.287839	15.282611	8.080039

2. Random Forest

Results: get all the possible states and weight percent that fits the user input weather data

Eg.

When Precip = 2; Ts = 67; Tw = 35; Windspeed = 6.5

Result: 'CT': 0.01, 'OR': 0.12, 'WA': 0.87

Equation to be applied will then be:

State1(source1 total electricity output) * RF_state1_weight + State2(source1 total electricity output) * RF_state2_weight + ...

State1(source2 total electricity output) * RF_state1_weight + State2(source2 total electricity output) * RF_state2_weight + ...

Output:

6 by 1 matrix containing all the sources with the sum weighted electricity total output in MW.

3. Use P-value to filter out the impossible resources

Based on the 6 by 1 matrix output from RF (sample mean), compared with the average electricity output of the states with its respective sources (population mean) using P-value at alpha = 0.05.

Output:

Filtered energy sources that are suitable for the location, with 95% confidence interval.

Regression Analysis

Goal: With all the suggested energy source types, prioritize clean energy and fill in the differences with conventional. Each suggested source will be analyzed with cost to give an estimate of the user how much will he/she saved over the course of 10 years. Using regression (logistic likely) to estimate the cost sector with 96-16 cost database constructed for this project.

Method: Regression (Logistic)

Procedure:

1. Construct the cost dataframe

Calculated average cost of certain resource by the voting weight of states

Eg. 4 votes for WA and 2 votes for CA,;The cost of NG in CA is 5.4, in WA is 5.09
so $\text{avgcost_NG} = 5.4 \times 33.3\% + 5.09 \times 66.7\% = 5.19$

All the Possible Resources	Cost(cents/MWh)	CO2 Tax (Cents/MetricTon)	Sales Price (GovernPrice) (cents/kWh) (industrial)
Solar			
Wind			
Coal			
NG			
Petroleum			

*cost sector takes in installation cost, maintenance cost, and operation cost into account and it is an average number.

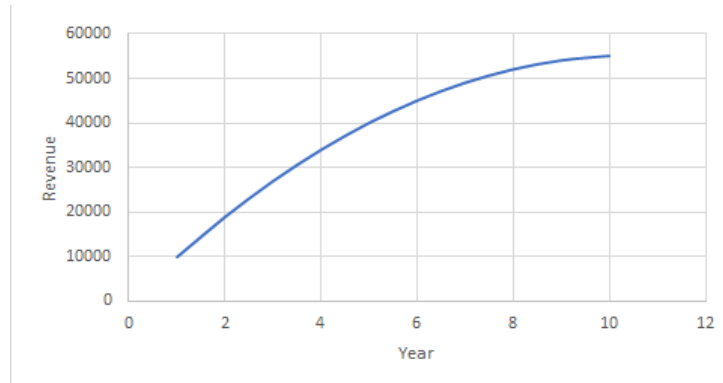
*CO2 tax values are based on a federal proposal in requiring all CO2 tax law in 2015.

2. Calculate the predicted amount saved of different resource

- Fit the cost model (logistic model, using F-statistics to filter out useless coefficients to get better fit) and get predicted cost and sales price for 2018-2028
Cost of buying electricity from Government = Sales Price (GovernPrice)*Capacity
Cost of self-generated electricity = self-generation cost (material, etc.) * Capacity
- Get the predicted Carbon Dioxide Tax for 2018-2028:
Carbon Dioxide Tax Estimate Source:
<https://www.c2es.org/publications/options-considerations-federal-carbon-tax>
- Calculate the amount saved in 2018 – 2028
Amount saved = Cost of Buying electricity from Government - Cost of self-generated electricity – Carbon Dioxide Tax * emission amount

Output:

For Clean energies (Wind, Solar, and Hydro), gives an estimate amount saved over the course of 10 years, the expected output graph should be looking as below with data points supporting.



For conventional energies (Petro, NG, and Coal), gives an estimate amount saved over the course of 10 years with it being used to sub in the differences. The expected output graph should look like below. A secondary graph calculating the amount CO2 emitted will also be plot, this additional amount of CO2 will be added onto the total CO2 emission to give an reference on how much CO2 the user will be contributing to if choose conventional energy source.

