Project Summary

The methodologies used to predict Solar farm performance depends on multiple metrics. Our fundamental interest is to evaluate the impact of various locations (geographical impact) and the modeling technique used. The project is designed to run as an experimental matrix where different teams work within different parameters to allow us to contrast and compare different modeling scenarios combining input data sources, and specific modelling techniques.

Organization background & information

Enverus is an energy technology company that provides innovative software, data, and services to help companies optimize their operations and gain greater insight into the energy markets. Founded in 1999, the company has become the leading provider of energy market data, analytics, and technology solutions. It has a comprehensive platform of data, tools, and applications that help energy companies make better-informed decisions and stay ahead of the competition. Enverus also offers a range of services, including expert guidance, data analysis, and market intelligence. With its innovative solutions, Enverus is helping energy companies of all sizes to thrive in a rapidly changing market.

Problem statement & objective, clearly stating desired goals and outcomes:

Objective: Compare and Contrast different design options between location specific source data and alternative modelling techniques. Enverus will provide anonymized data sets with input, and target variables. Enverus will suggest a range of modelling techniques. The design matrix is in the format where there are 1 … L data sources (think of data sources as having spatial proximity to the area under study from being macro , meso, and micro regions to the area under study_ and 1... M model classes (think of Linear Regression, Neural Networks, Tree Models as examples of model classes)

	Model Class				
Data Source					
	DS1, MC1	**DS2, MC1**	**...**	**DSM, MC1**	
	DS1, MC2				
	...				
	DS1 MCM				

Desired/required skills

Required: Solid understanding of Regression, Neural Network (2-5 layers) , and Tree model development , training, and evaluation.

Desired: experience with automated hyperparamater libraries, and Python ML practices. If you program in a different language but have all the skills under required – that is acceptable.

Names of key contacts

- Akash Sharma (Product & DS)
- Mishtu Banerjee (DS)

Nature, size, and availability of data set to be analyzed

Data sets will be less than or equal to 50K rows and 20 input variables per Data Source

Technologies/tools to be used

- Open Source ML and Visualization libraries

Resources and/or reference materials

- Any good reference book on time series forecasting.
- We will provide a short reading list of Solar Generation forecasting using each of the key modelling classes.