A screenshot of a web page

Description automatically generated with low confidence

GOAL: 7 days hourly forecast (for the following week)(7x24 rows)

<https://pvlib-python.readthedocs.io/en/stable/>

A screenshot of a computer

Description automatically generated with medium confidence

A screenshot of a computer

Description automatically generated

A picture containing text, screenshot, document, font

Description automatically generated

Forecasting GHI

A screenshot of a computer

Description automatically generated with medium confidence

A picture containing text, screenshot, document, font

Description automatically generated

we usually want a set of inputs that are adding value to the forecast, not noise. So so think about it and think signal versus noise for every variable you add. Is there signal here? Is there noise, the other one is are the input variables

You choose bringing in new information to the model or they correlate it O. If you've got 2 variables and 11 variable is a mathematical transformation of the other, it's not adding any new value. If you have two variables and, they're exceedingly highly correlated, they might be bringing in a lot of redundant information.

think through the **causation** and **correlation** component, right? Like we provided and meshed was given some general indications on like GHI and that is a very good indicator. But you know a big part of this project is the **exploration page**, right. Let's go and figure out what else could have an impact.

**Process to evaluate** each approach before you pick the one that makes the most sense, right? Because there are significant advantages to each of those approaches from a standpoint of speed of deploying the model and testing. If you know test out different models, or if you go through a different modeling philosophy, you're trying to account for a much more smaller or relatively smaller amount of variable in predictive variables. If you're looking at smaller areas, right? So if you take those smaller pieces and create an aggregate ensemble model.

**Think about the variables that you're putting in** and think about the different ways of generating this forecast before picking one. And you know, going all the way through and testing different philosophies around itself.

A screenshot of a computer

Description automatically generated with medium confidence

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated with medium confidence

Dataset

A screenshot of a computer

Description automatically generated with medium confidence

Actual meseasurement

A screenshot of a computer

Description automatically generated with low confidence

A screenshot of a computer

Description automatically generated with medium confidence

A screenshot of a website

Description automatically generated with medium confidence

A screenshot of a computer

Description automatically generated with medium confidence

Time series forecast error metrics you should know from towardsdatascience

A screenshot of a computer

Description automatically generated

A picture containing text, font, screenshot

Description automatically generated