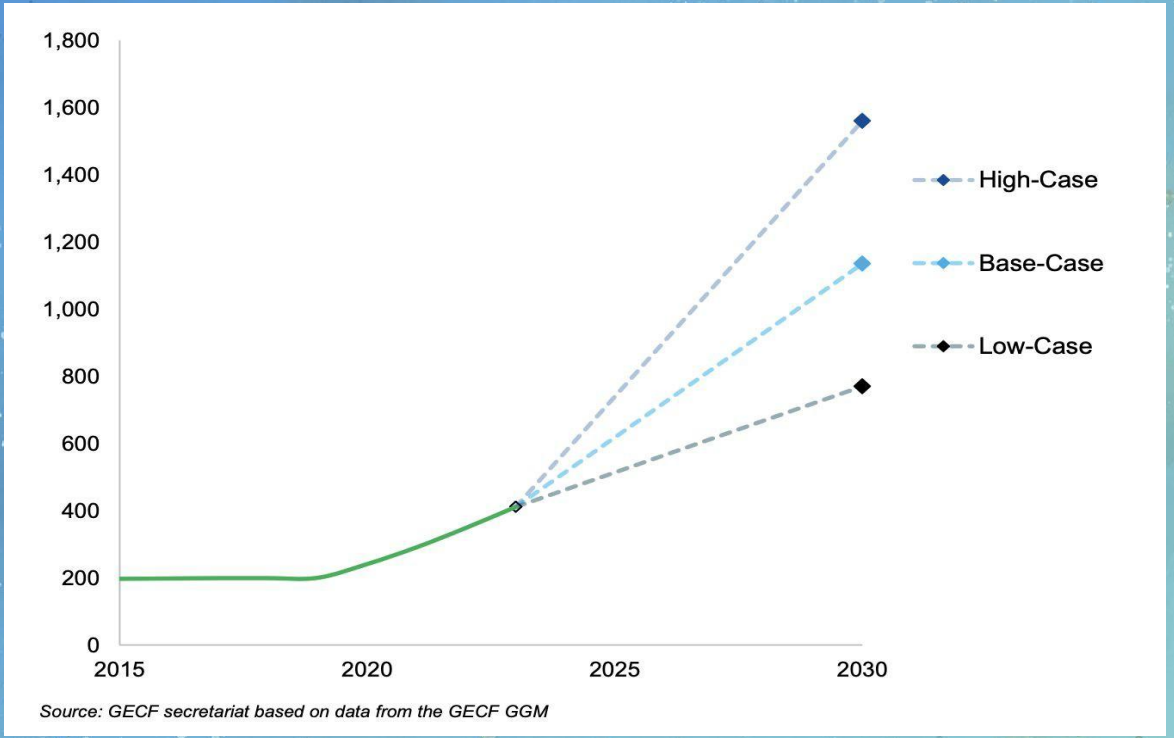
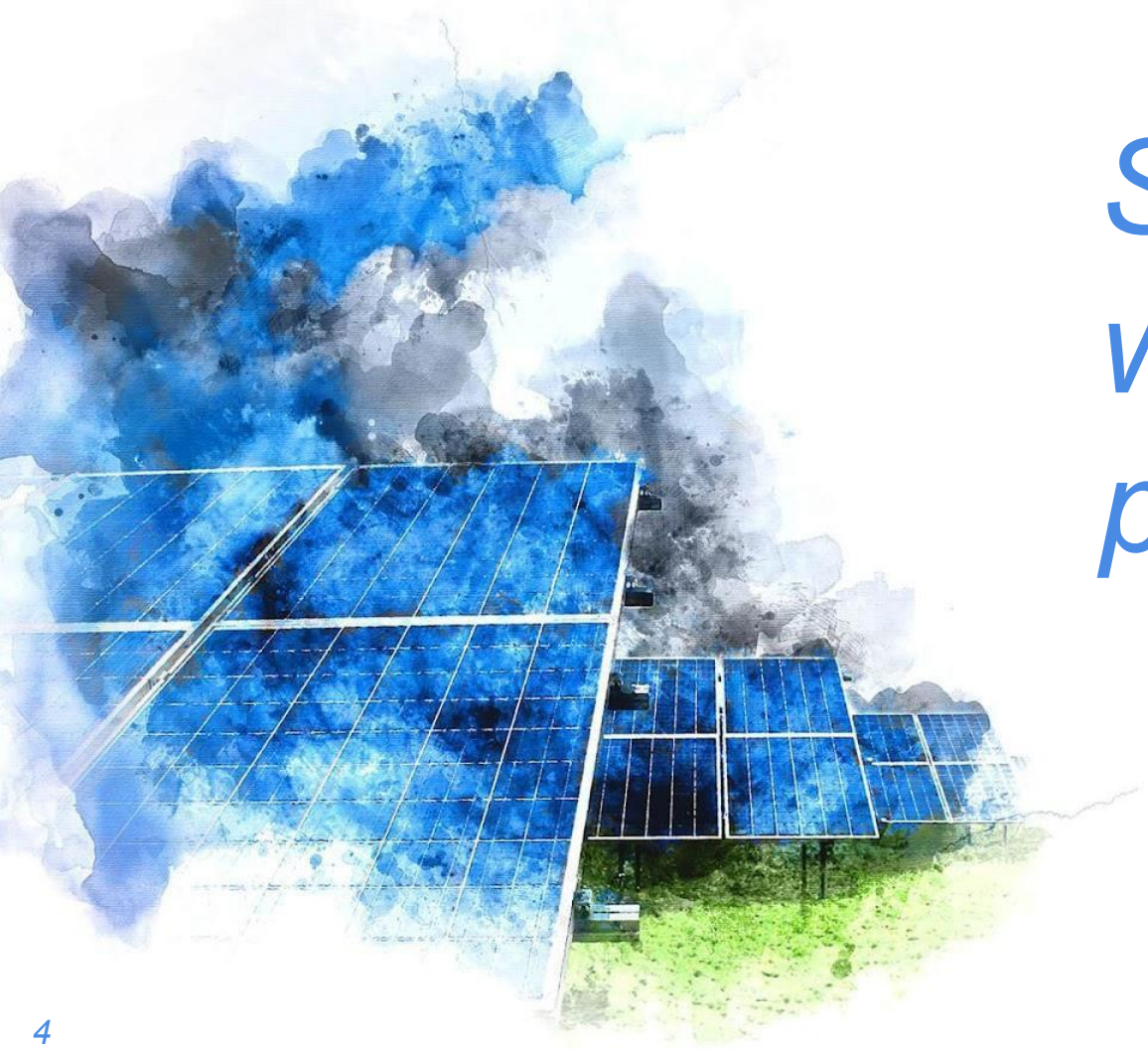
Reducing the Environmental Impact of Data Centers Using Machine Learning: How We Can Innovate Without Harming the Planet

Saanvi Uppalapati


NETFLIX
RECOMMENDS?



Source: GECF secretariat based on data from the GECF GGM



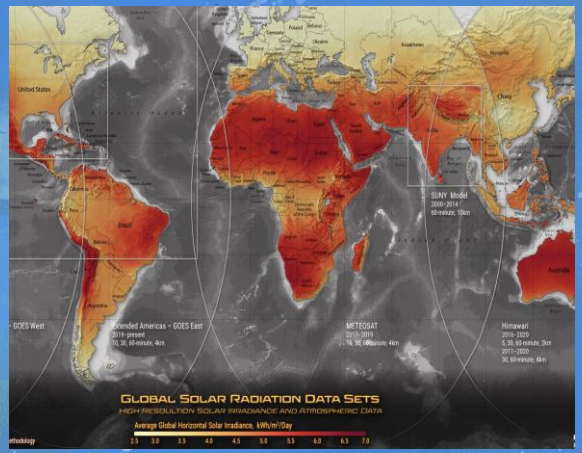
Solutions with problems…

VS.

Main goals

- ***Predict the next 24 hours of solar energy production in and around a given data center's location***
- ***The prediction would be analyzed to determine at what times throughout the day there is peak solar radiation in the atmosphere***
- ***Data centers can train their models during those times, relying more on renewable sources of energy and reducing their carbon footprint by burning less fossil fuels***

```
historical_patterns = self._get_historical_patterns(
    latitude, longitude, start_time, start_year, end_year
```



NSRDB Global Viewer

```
"hour_sin",
"hour_cos",
"day_of_year_sin",
"day_of_year_cos",
"temperature",
"pressure",
"relative_humidity",
"cloud_type",
"solar_zenith_angle",
"clearsky_ghi",
```

Chosen features from dataset utilized in model

Model Evolution Timeline and Improvements

Initial Version	Enhanced Version
$P = \frac{1}{N}\sum_{i=1}^{N} T_i(x)$	$F(x) = F_0(x) + \eta \sum_{t=1}^{T} h_t(x)$
• P = Final prediction of the Random Forest model • N = Number of decision trees in the forest • $T_i(x)$ = Prediction of the i^{th} decision tree for input x	• $F_0(x)$ = The initial prediction (often the mean or mode) • η = Learning rate • $h_t(x)$ = Weak learner (tree) at iteration t • T = Total number of boosting iterations (trees)
• Random Forest model had basic predictors and limited feature analysis. • Surge prediction accuracy: 30-40%. • Majority vote prediction method reduced accuracy.	• Gradient Boosting offered advanced prediction and feature analysis. • Accuracy: 50-55%. • Trees trained sequentially to correct previous errors, boosting accuracy.

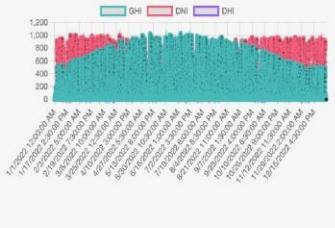
Dashboard

Please note: Due to server resource constraints, data fetching is currently limited to one year at a time. Select a single year between 1998-2022.

Latitude	Longitude	Year	
39.011	-77.471	2022	Fetch Forecast

Total Records	Average GHI (W/m²)	Max GHI (W/m²)
17520	173.73	1034.00

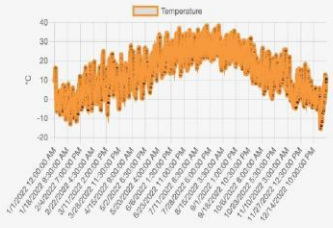
Solar Irradiance Components



Understanding Solar Irradiance Components:

- GHI (Global Horizontal Irradiance): Total solar radiation received on a horizontal surface
- DNI (Direct Normal Irradiance): Solar radiation received directly from the sun
- DHI (Diffuse Horizontal Irradiance): Solar radiation scattered by the atmosphere

Temperature Variation



About the Data:

The chart shows the temperature variation over the selected year. Use this data to understand how temperature affects solar energy generation.

Solar Irradiation Forecast (Next 24 Hours)

Peak Solar Hours — Best times for solar generation

Peak Period 1 — 1/1/2023

01:30 PM
Expected Irradiance: 26.29 W/m²
Estimated Power: 3.94 W

Peak Period 2 — 1/1/2023

02:30 PM
Expected Irradiance: 19.05 W/m²
Estimated Power: 2.86 W

Peak Period 3 — 1/1/2023

12:30 PM
Expected Irradiance: 16.99 W/m²
Estimated Power: 2.55 W

Peak Period 4 — 1/1/2023

11:30 AM
Expected Irradiance: 13.99 W/m²
Estimated Power: 2.10 W

Peak Period 5 — 1/1/2023

03:30 PM
Expected Irradiance: 13.12 W/m²
Estimated Power: 1.97 W

Understanding Peak Hours:

Future Improvements

- *Adding wind energy data*
- *automate data integration from data center locations*
- *test my model using real-world data centers.*

“What can WE do?

Thank you!