



Imagine living in a house where every single watt of electricity you use is meticulously recorded, each of which contributes to a vast pool of data. By analyzing this detailed household power consumption data recorded over nearly 4 years, an energy company can help customers achieve sustainable energy usage while balancing their energy generation. With predictive models, the company can optimize energy usage, forecast future consumption, and provide tailored recommendations. Your task is to use this dataset to build a model that predicts power consumption, benefiting both the energy provider and its customers.


## The Data

Available in `df_train.csv` and `df_test.csv` :

Column	Type	Description
date	chr	Date of the measurement
power_consumption	dbl	Daily power consumption (in kilowatts)
year	int	Year of the measurement
semester	int	Semester of the measurement (1 for Jan-Jun, 2 for Jul-Dec)
quarter	int	Quarter of the measurement (1 for Q1, 2 for Q2, 3 for Q3, 4 for Q4)
day_in_week	chr	Day of the week of the measurement (e.g., Monday, Tuesday)
week_in_year	int	Week number in the year of the measurement
day_in_year	int	Day number in the year of the measurement
month	int	Month of the year of the measurement

This dataset was donated to the UCI Machine Learning Repository. For detailed information about the dataset and the preprocessing steps, please refer to the License and Data Preprocessing Details ([Invalid URL](#)) notebook.

...	↑↓	da...	...	↑↓	power_consum...	...	↑↓	...	↑↓	:	...	↑↓	...	↑↓	day...	...
1		12/16/2006			1209.176			2006			2		4		Sat	
2		12/17/2006			3390.46			2006			2		4		Sun	
3		12/18/2006			2203.826			2006			2		4		Mon	
4		12/19/2006			1666.194			2006			2		4		Tue	
5		12/20/2006			2225.748			2006			2		4		Wed	
6		12/21/2006			1716.624			2006			2		4		Thu	
Rows: 6															↗ Expand	

index	...	↑↓	Model
1			Linear Regression
2			Random Forest
3			XGBoost
Rows: 3			 Expand
selected_rmse: 392.4813 kW			

