**Overview**

**Weekend hackathons are fun, aren't they!** *In our last weekend hackathon, we introduced a new and unique problem statement using*[*UCI*](https://archive.ics.uci.edu/ml/index.php)*open dataset. But, we were big-time disappointed as some of the participants ended up probing the leaderboard*. However, we decided to host an open [UCI](https://archive.ics.uci.edu/ml/index.php) dataset competition again this weekend. So In this weekend hackathon, we have trained a machine learning model to perturb the target column instead of manually adding the noise. Yes, you heard it right, In this hackathon, we are challenging all the MachineHackers to capture our leaderboard's and prove their mettle by competing against MachineHack's AI.

The dataset was collected from a [**Combined Cycle Power Plant**](https://en.wikipedia.org/wiki/Combined_cycle_power_plant) over 6 years (2006-2011) when the power plant was set to work with a full load. Features consist of hourly average ambient variables Temperature (**T**), Ambient Pressure (**AP**), Relative Humidity (**RH**), and Exhaust Vacuum (**V**) to predict the net hourly electrical energy output (**PE**) of the plant.  
A combined-cycle power plant ([**CCPP**](https://www.youtube.com/watch?v=KVjtFXWe9Eo&feature=youtu.be&list=PLpSQBjC0ANgxIJ4YaFxITCSfLQQPsKD5s)) is composed of gas turbines (**GT**), steam turbines (**ST**), and heat recovery steam generators.

In a [**CCPP**](https://www.youtube.com/watch?v=KVjtFXWe9Eo&feature=youtu.be&list=PLpSQBjC0ANgxIJ4YaFxITCSfLQQPsKD5s), the electricity is generated by gas and steam turbines, which are combined in one cycle, and is transferred from one turbine to another. While the Vacuum is collected from and has an effect on the Steam Turbine, the other three of the ambient variables affect the GT performance.

**Dataset Description:**

* **Train.csv -**9568 *rows* x 5 *columns*
* **Test.csv -**38272 *row*s x 4 *columns*
* **Sample\_Submission.csv - Accepted submission format**

**Attribute Information:**

Features consist of hourly average ambient variables

1. **Temperature** (T) in the range 1.81°C and 37.11°C
2. **Ambient Pressure** (AP) in the range 992.89-1033.30 millibar
3. **Relative Humidity** (RH) in the range of 25.56% to 100.16%
4. **Exhaust Vacuum** (V) in the range 25.36-81.56 cm Hg
5. Net hourly electrical energy output (PE) 420.26-495.76 MW

*The averages are taken from various sensors located around the plant that record the ambient variables every second. The variables are given without normalization.*

**ACKNOWLEDGMENTS:**

* *Pınar Tüfekci,*[*Prediction of full load electrical power output of a*baseload*operated combined cycle power plant using machine learning methods*](http://dx.doi.org/10.1016/j.ijepes.2014.02.027)*, International Journal of Electrical Power & Energy Systems, Volume 60, September 2014, Pages 126-140, ISSN 0142-0615*
* *Heysham Kaya, Pınar Tüfekci, Sadık Fikret Gürgen: Local and Global Learning Methods for Predicting Power of a Combined Gas & Steam Turbine, Proceedings of the International Conference on Emerging Trends in Computer and Electronics Engineering ICETCEE 2012, pp. 13-18 (Mar. 2012, Dubai*