Problem Identification

The energy consumption growth needs more energy production, which has environmentally harmful byproduct: CO and NOx emission which may have irreparable effects on the environment, gradually.

The data set that I am going to analyze here has been collected over five years and is the result of monitoring and reading the sensors output in a gas turbine power plant. The dataset includes features such as:

1. AT -> Ambient temperature (C)
2. AP -> Ambient pressure (mbar)
3. AH -> Ambient humidity (%)
4. AFDP -> Air filter difference pressure (mbar)
5. GTEP -> Gas turbine exhaust pressure (mbar)
6. TIT -> Turbine inlet temperature (C)
7. TAT -> Turbine after temperature (C)
8. CDP -> Compressor discharge pressure (mbar)
9. TEY -> Turbine energy yield (MWH)

And the goal of this project is to find a model that shows the relation between these factors to predict the gas emission. Therefor the target variables in this dataset are:

1. CO -> Carbon monoxide (mg/m3)
2. NOx -> Nitrogen oxides (mg/m3)

This dataset includes total 11 attributes and 36733 record which has been downloaded from the UCI website, here is the link:

https://archive.ics.uci.edu/ml/datasets/Gas+Turbine+CO+and+NOx+Emission+Data+Set

There is also a paper published in Turkish Journal of Electrical Engineering & Computer Sciences related to this problem, here is the link:

https://journals.tubitak.gov.tr/elektrik/issues/elk-19-27-6/elk-27-6-54-1807-87.pdf

By building the prediction model we will be able to run this gas turbine power plant with less gas emission.