Prediction-1:

Target Variable: Primary Fuel(the source of energy generation of the power plant)

To predict the primary fuel of a power plant based on the provided attributes, we can use machine learning algorithms such as classification models. As we can observe from the database that the attribute is a categorical variable, we can use algorithms like Logistic Regression or Decision Trees.

Logistic regression:

Let x be the probability of a power plan having the primary fuel. X1,x2,x3,….-attributes

X1-capacity_mw

X2-commissioning_year

X3-generation_gwh_2013 x4- generation_gwh_2014.,,2018,etc….

Log(p/1-p) =b0 + b1x1 + b2x2 +……

Prediction-2:

Target Variable: capacity_mw(electrical generating capacity)

We can observe from the data description that the target variable is numeric. In Machine learning, we use regression type of algorithms to predict. We can use linear regression, Random Forest etc.

Linear Regression:

Let Y be the target variable “capacity_mw”

Y = b0 + b1x1 + b2x2 ……

b0-primary_fuel

b1-commissioning_year

b2generation_gwh_2013-2018…etc

The b0,b1,b2,…. Are estimated using Least squares