

Final Project Report

Fundamentals of Machine Learning

May 6th, 2023

Srilaya Valmeekam

Executive Summary:

The goal of this project is to assess several parameters associated with electricity generation in the United States and identify a fuel type that can be phased out of the power generation process, in accordance with the US government's goal of conserving fossil fuels. To accomplish this purpose, an examination of expenditures on various fuel kinds was performed in order to determine which one receives the least support. This sort of fuel is predicted to be used the least, and the government intends to preserve it for future use.

By utilizing a machine learning technique, the data was segregated into three categories, each representing a distinct fuel type. Further inquiry was carried out by analyzing the mean fuel price, the quantity of fuel received at power plants, and the chemical composition of each fuel variant. The research revealed that the highest expenses incurred by the government were related to gas, whereas the least expenditures were made on oil. Despite having a lower MMBtu cost per unit, coal is not commonly used, which may be due to the harmful substances such as ash, mercury, and sulfur present in it.

Introduction:

The data contains information about monthly fuel contracts, purchases, and costs. It has about 20 variables, ranging from the mine_ids from which fuels will be provided to various power plants by various suppliers to the mode of transportation used to supply these fuels. Data was cleansed for effective analysis, and certain variables with missing or redundant data were eliminated. A set was used to sample 2% of the data from 608,565 rows. Seed (2467), which aids in better data interpretation. In addition, the fuel type code pudl variable has been turned into a numerical variable by creating three dummy variables for three different types of fuel.

Problem Statement:

The U.S. Power Generation Unit hired a data analyst to examine past data of monthly_fuel_contract, purchases, and cost information to determine the type of fossil fuel on which they are spending the least money. The fossil fuel type on which they spend the least money will be reduced in cost, and it will no longer be used to generate electricity. This new strategy's declared goal is to conserve fossil fuels for use in the future because they are currently not being used extensively.

Questions to focus on:

- Which fuel do they spend the most and least money on?
- Are there any additional factors to consider when excluding a specific type of fuel from power generation?
- Do you have any other alternative recommendations as a data analyst?

Analysis and Discussion:

After thoroughly analyzing the data using a machine learning algorithm, below are my findings:

The average amount spent on Gas, Coal, and Oil are as follows:

Gas – 963,957.95 dollars

Oil – 53,227 dollars

Coal- 82,689.86 dollars

*To get the total cost, multiply the average number of fuel units received by the average cost per unit of fuel.

- The U.S.A. power generation spends the most on gas since it is the fuel that is obtained in the greatest number of units at the lowest cost per unit.
- Oil is the fuel type on which they spend the least money. According to the US Energy Information Administration, oil is used not just in electricity generation but also to drive vehicles and in the petrochemical industry to create products such as plastics, Solvents, and hundreds of other intermediate and end-user products.

Source: [Link](#)

Therefore, it is not recommended to exclude Oil from power generation as this oil can be used for multiple reasons.

- In comparison to Oil, the amount spent on **Coal** is higher and as per the chemical composition of coal, it contains percentages of ash, mercury, and Sulfur impurities. In power plants, when power is generated by this type of fuel with impurities, additional expenses are to be covered as the greater the degree of these impurities, the higher would be the associated cost. The government should also consider these additional costs.
- When talking about preserving fossil fuels for future generations, as per the article posted by International Energy Agency, all unabated coal generation ends by 2040. To get on track with the Net Zero by 2050 Scenario, an annual average reduction of emissions from coal-fired power plants of around 8% is needed through 2030.

Source: [Link](#)

Hence, there is a need to preserve Coal for future generations.

I would like to recommend the Government deploy more efficient technologies like Carbon Capture Utilization and Storage to generate power to preserve fossil fuels.

Conclusion:

The type of fuel to be excluded from power generation is Coal as there is an immediate need to cut down the usage of Coal and preserve it for future generations. Moreover, the amount spent on Coal is way too higher than on Gas and Oil.

Based on the findings, I propose that the government does not focus on removing the fuel that receives the least funding. Alternatively consider the most expensive petrol, especially if it is not the most efficient variety. As a result, rather than wasting lots of money on coal, the government can redirect the money to oil, which is available in its purest form and can be used for a variety of purposes.