

Aquatic Insights:Cognos- Powered Water portability Analysis

Understand the dataset-Water portability

Introduction:

- Water is a chemical compound composed of two hydrogen atoms and one oxygen atom, known by the molecular formula H_2O . It exists in a liquid state and is commonly referred to as water. It can also exist as a solid called ice or as a gas known as steam. When water is heated and compressed beyond its critical temperature and pressure, it becomes a supercritical fluid, which has applications as a green and environmentally friendly solvent that can be adjusted to specific needs.

Dataset source:

- Your dataset includes information about water sources and their qualities, such as turbidity, hardness, pH, and other parameters. This data was obtained from a crowd-sourced platform called Kaggle.

Data Wrangling:

- The dataset has been loaded into a data frame in the notebook using the `pd.read_csv()` function for further analysis and modeling. To verify the successful reading of the data file and understand its structure, the `head()` function is used to display a few lines of the dataset. This helps in gaining an overview of the data, including its shape, data types, and the presence of any null values.

Properties of Water:

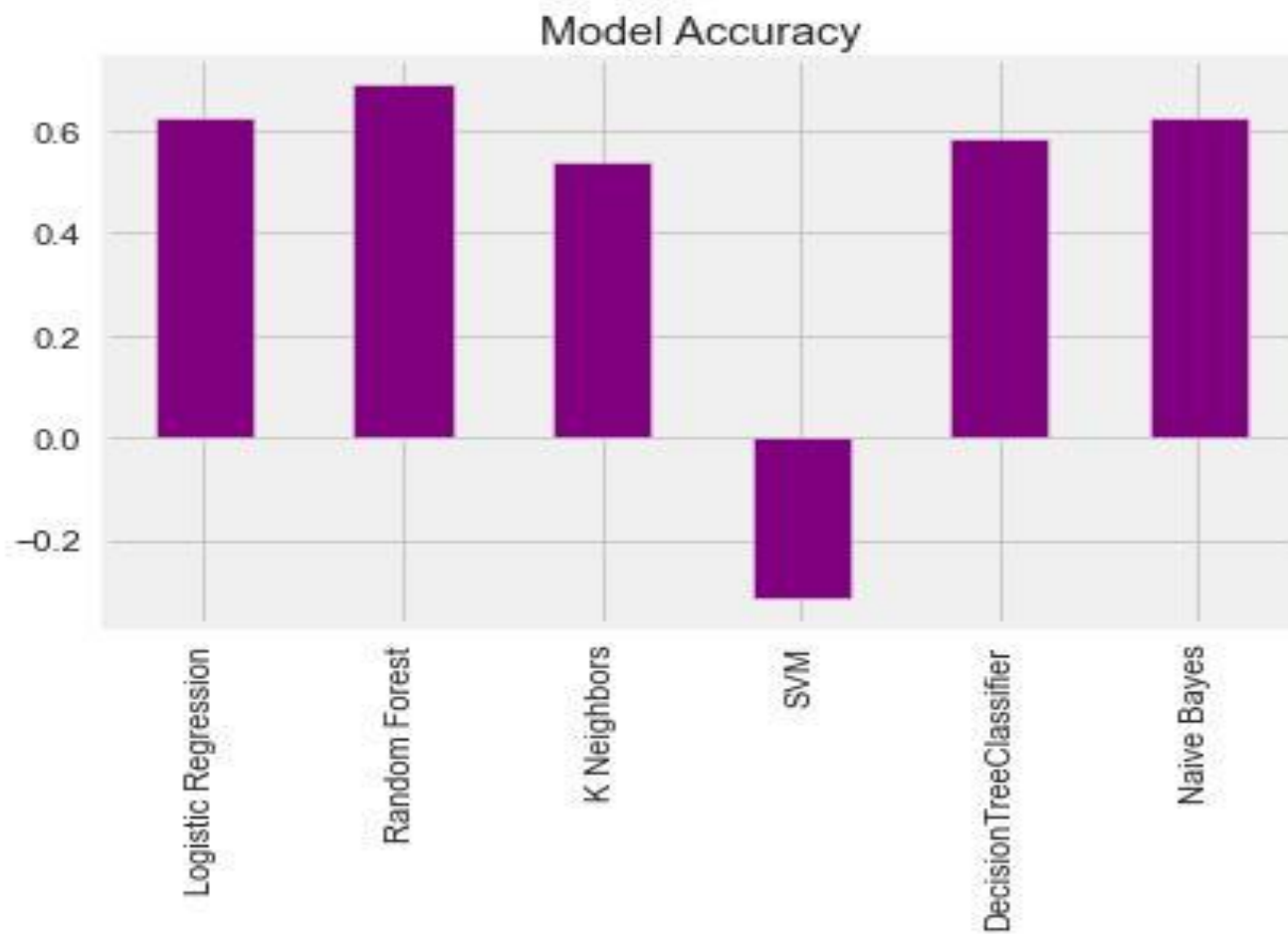
ph	491
Hardness	0
Solids	0
Chloramines	0
Sulfate	781
Conductivity	0
Organic_carbon	0
Trihalomethanes	162
Turbidity	0
Potability	0
dtype:	int64

Program:

```
filename = "water_potability.csv"
sep = ","
dft = AV.AutoViz(
    filename,
    sep=",",
    depVar="",
    dfte=None,
    header=0,
    verbose=1,
    lowess=False,
    chart_format="svg",
    max_rows_analyzed=150000,
    max_cols_analyzed=30,
)
```

Shape of your Data Set loaded: (3076, 10)

Model accuracy chart:



Conclusion:

- To enhance the accuracy of the classifier, it was concluded that building a neural network would be the best choice, potentially achieving up to 90% accuracy. However, it was deemed that the current classifiers provided satisfactory analysis of water
- Furthermore, it is recommended to deploy the model on the web using a Python framework with good compatibility, preferably Flask. This deployment would enable the model to assist in real-life scenarios by determining the potability of water.