



**AKN WATER**  
FOR WATER QUALITY

Prepared by:

Abdultawwar Safarji

Najla Bin-Melha

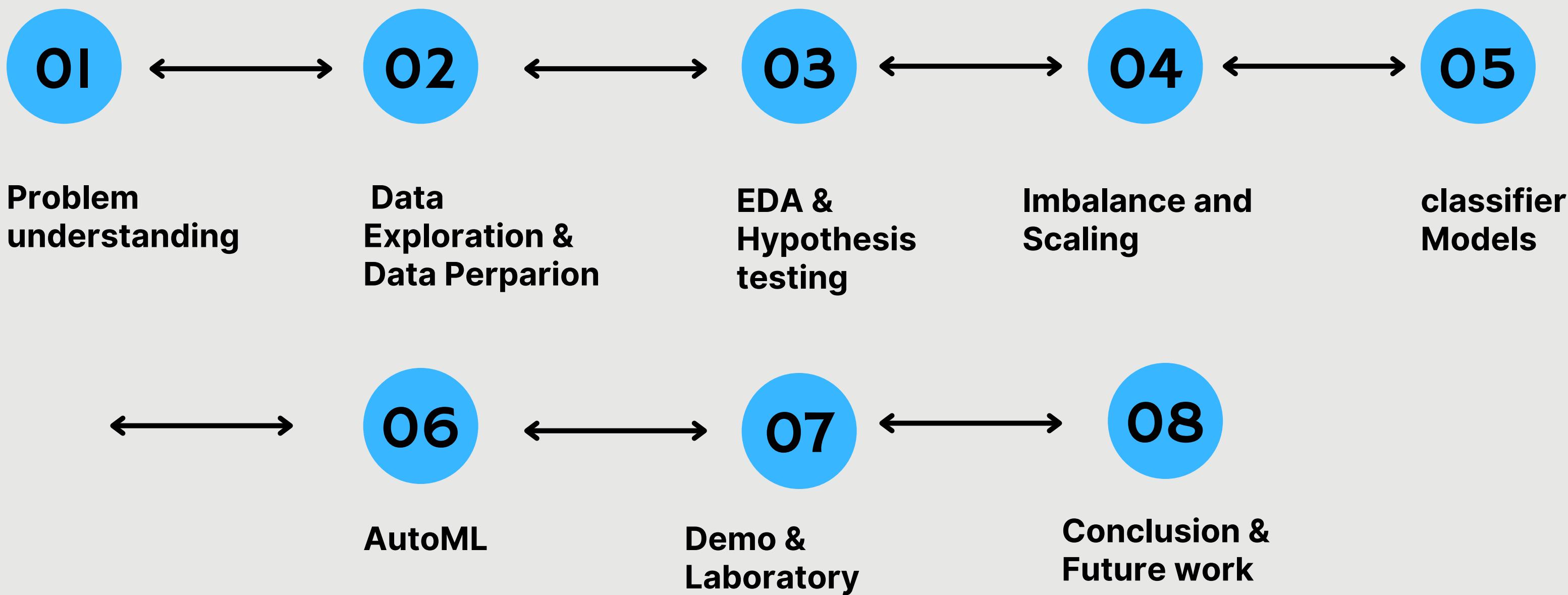
Khalid Alrashed



A clear glass filled with water is positioned on the left side of the frame. The glass is set against a solid black background, which creates a strong contrast with the transparent glass and the light-colored water. The lighting highlights the edges of the glass and the surface of the water, showing small ripples and reflections. The overall composition is minimalist and focuses on the purity and clarity of the water.

Have you ever asked yourself if  
the water you drink is really  
**safe or not?**

# Methodology



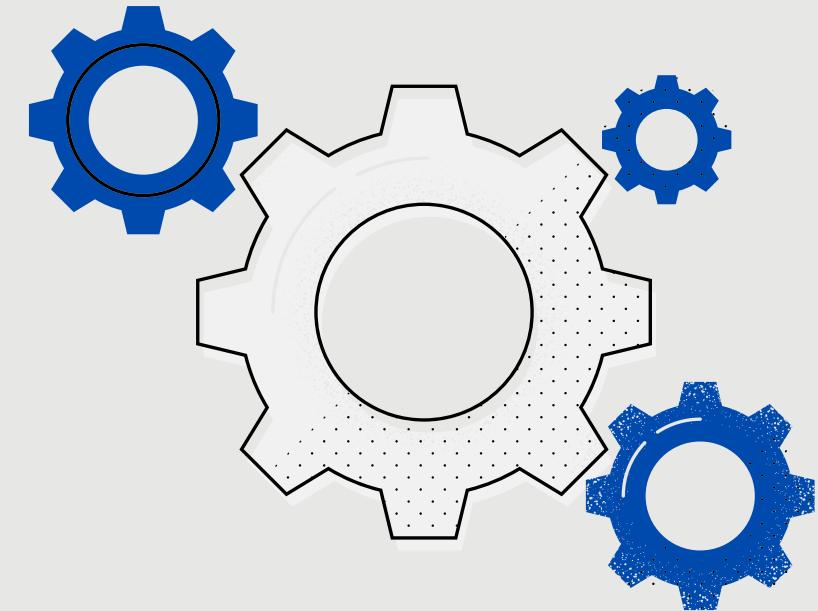
# Problem Understanding

## **Classifying water is critical task to do :**

The weight of wrongly classifying sample to potable is equal to wrongly classifying to not potable, if not how to find a good measure for the relation.



# Data Preparation & Explorarion



**Using water potability the following features:**

Chemical properties (pH, Hardness, Total solids, chloramines , sulfate , conductivity Trihalomethane's , Organic carbon)

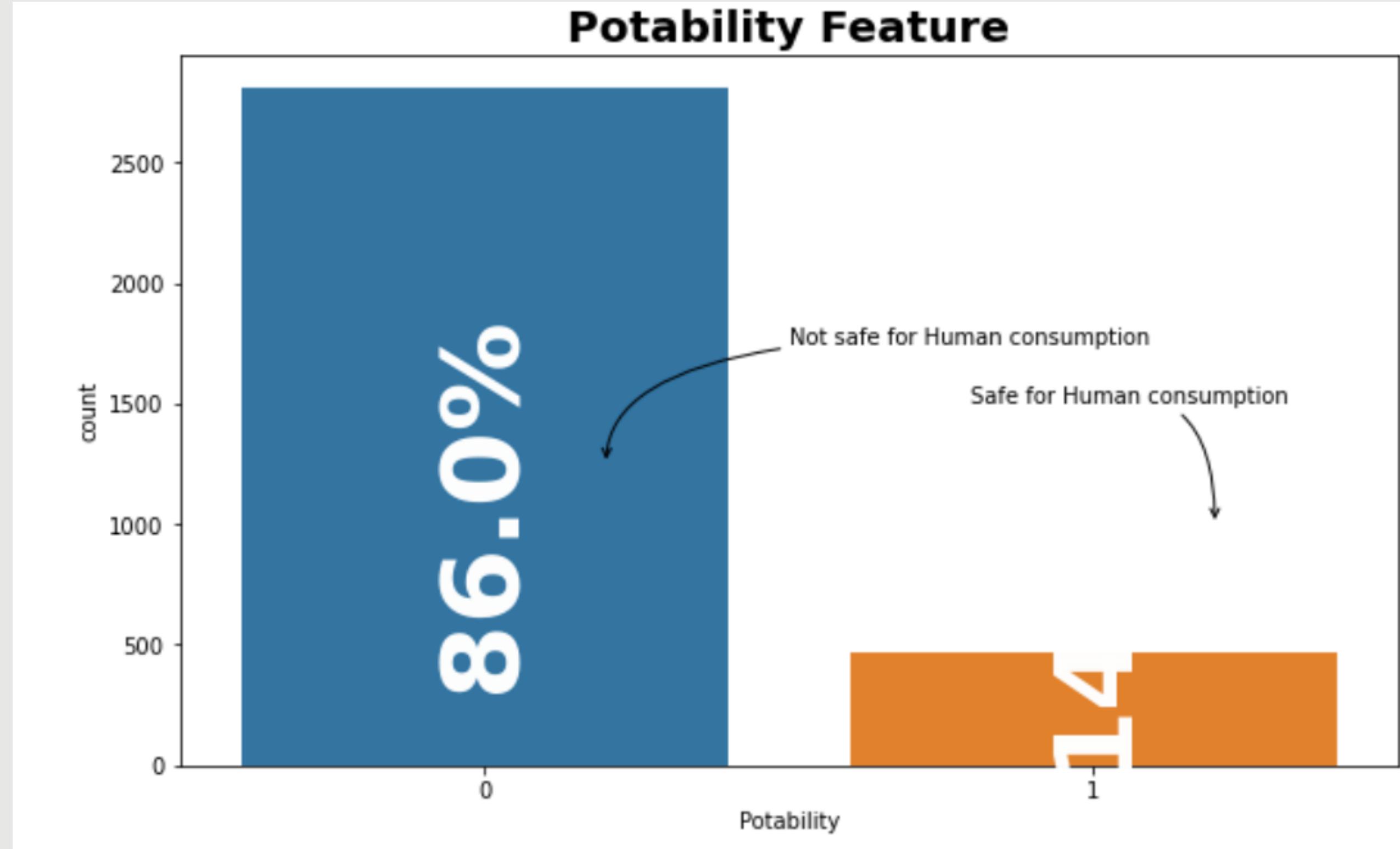
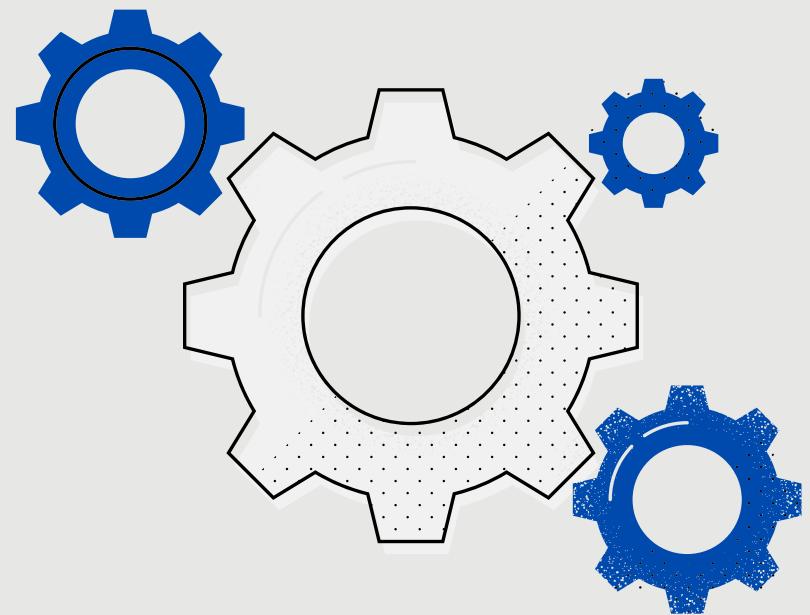
Physical: Turbidity.

**Target: safe or not safe**

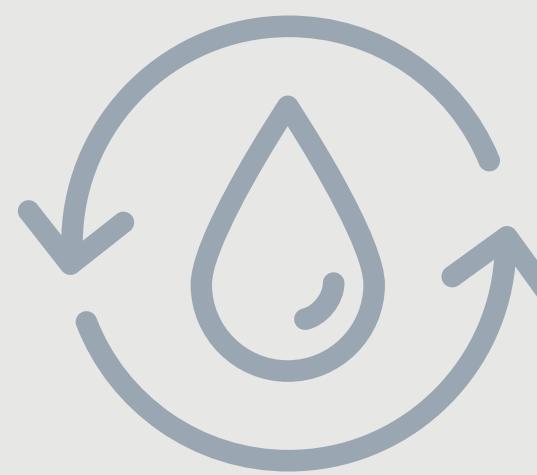
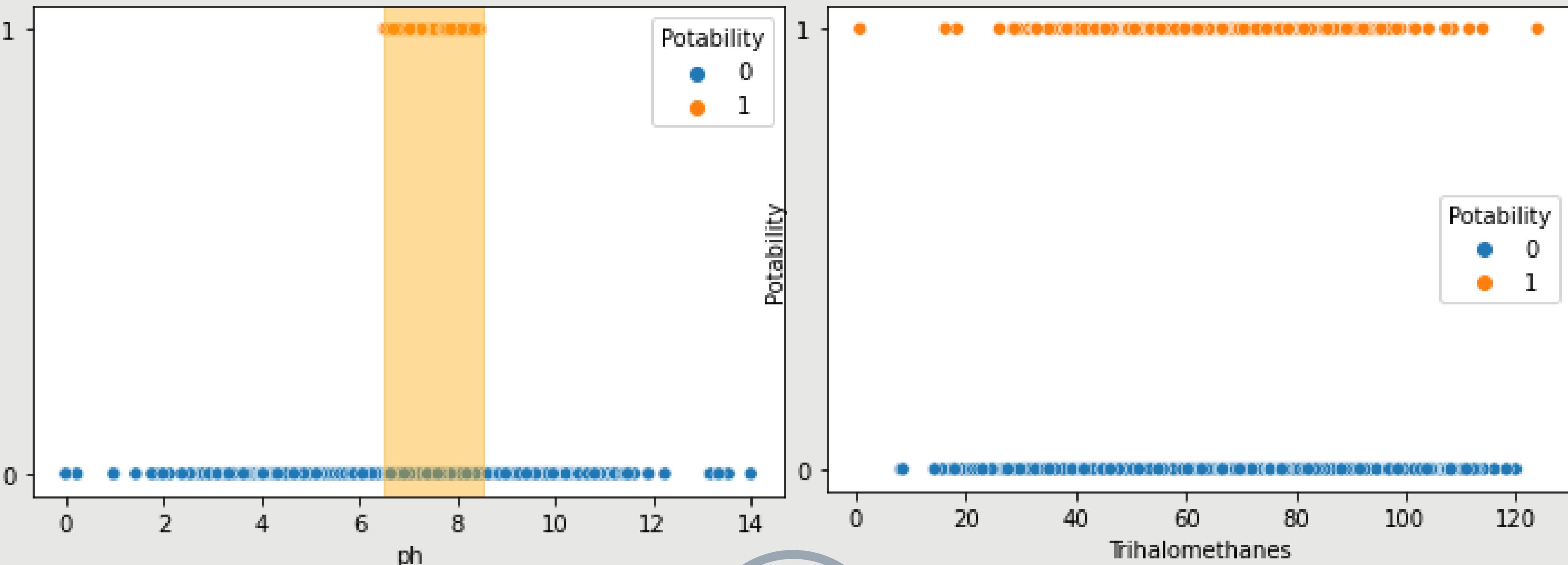
---

# Data Exploration & Validation

- Data Imputation
- Imbalance data
- Missing values
- Duplicate Data

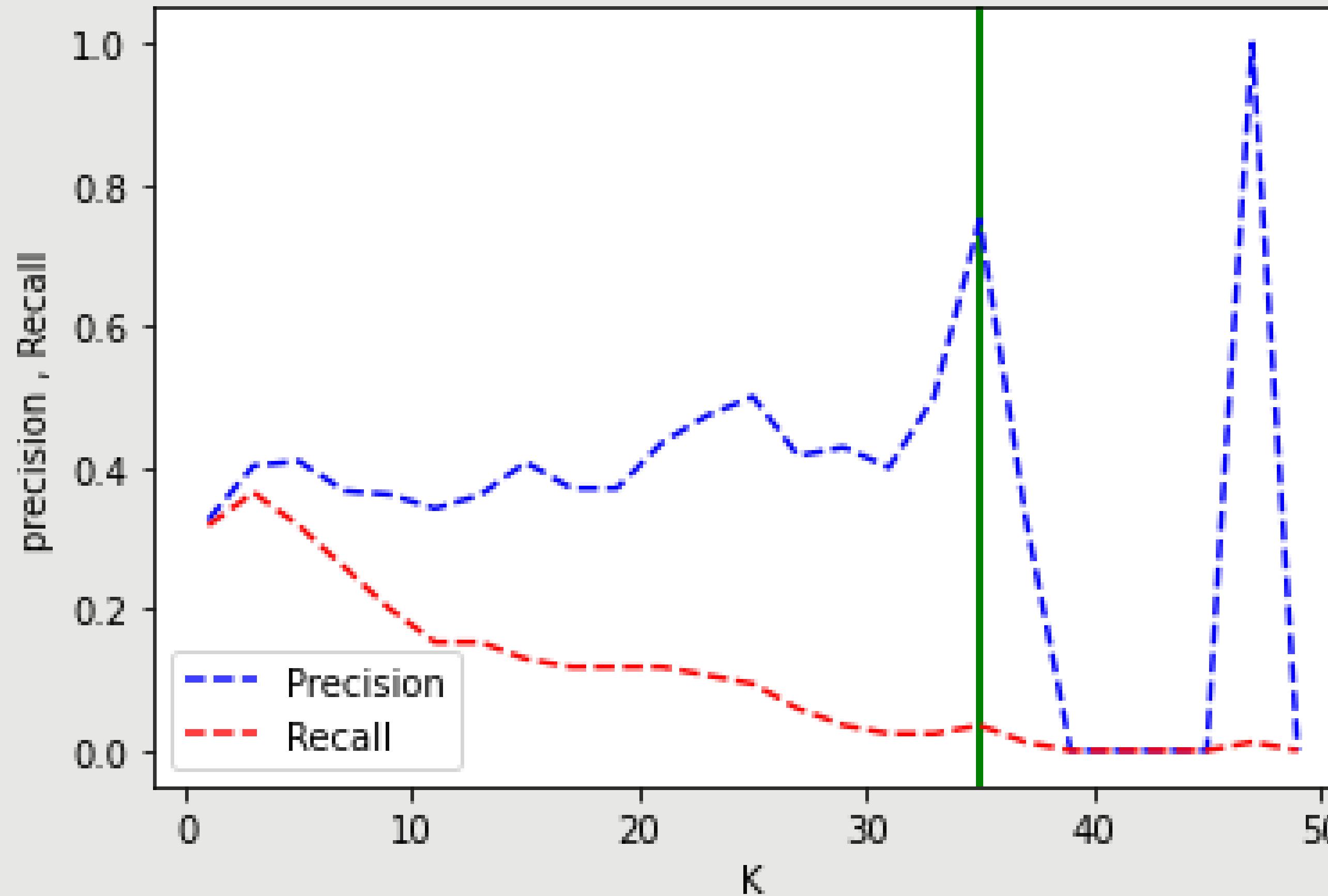


# Exploratory Data Analysis

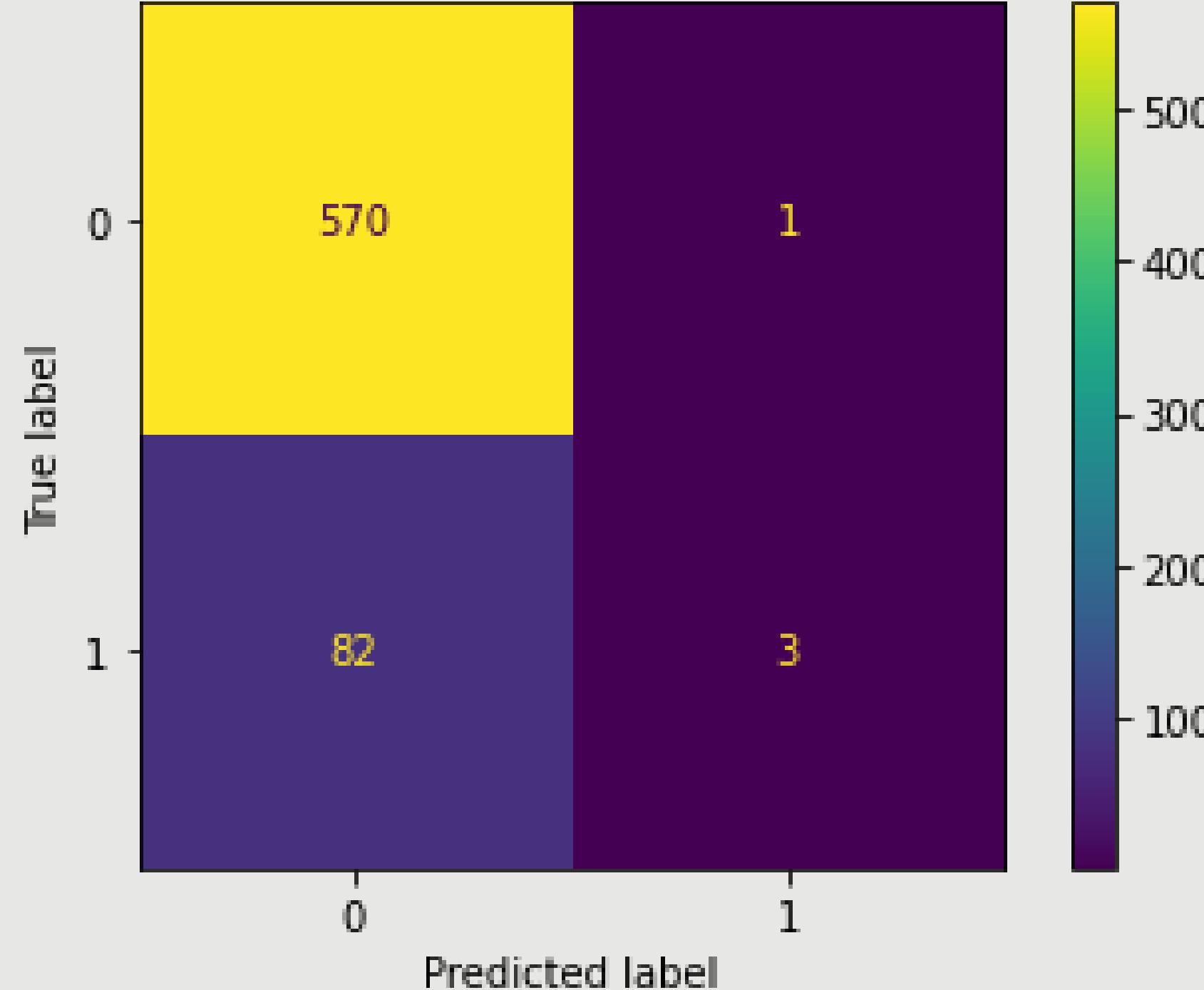


# KNN Model

Precision-Recall vs number of neighbors

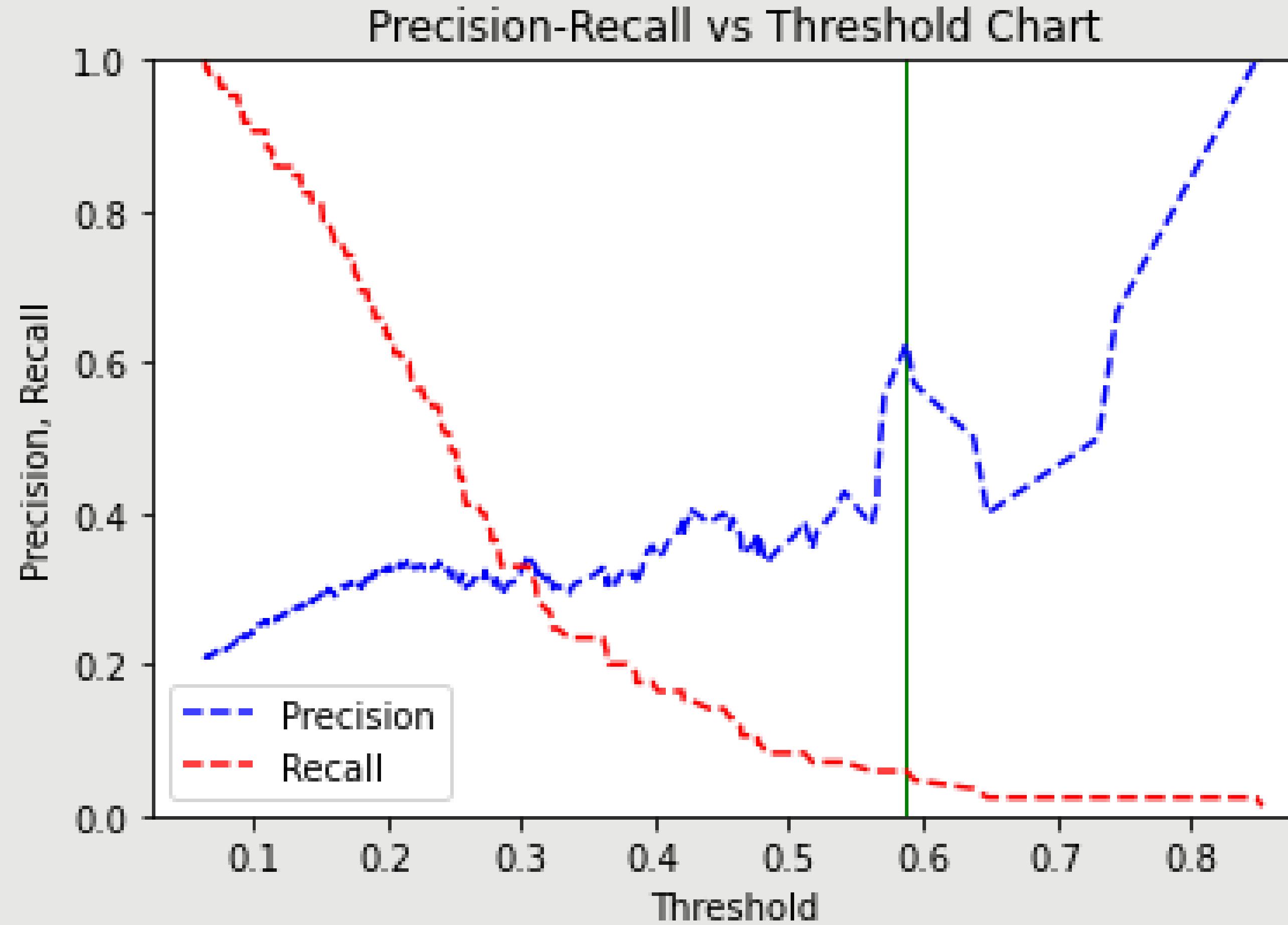


# KNN Model

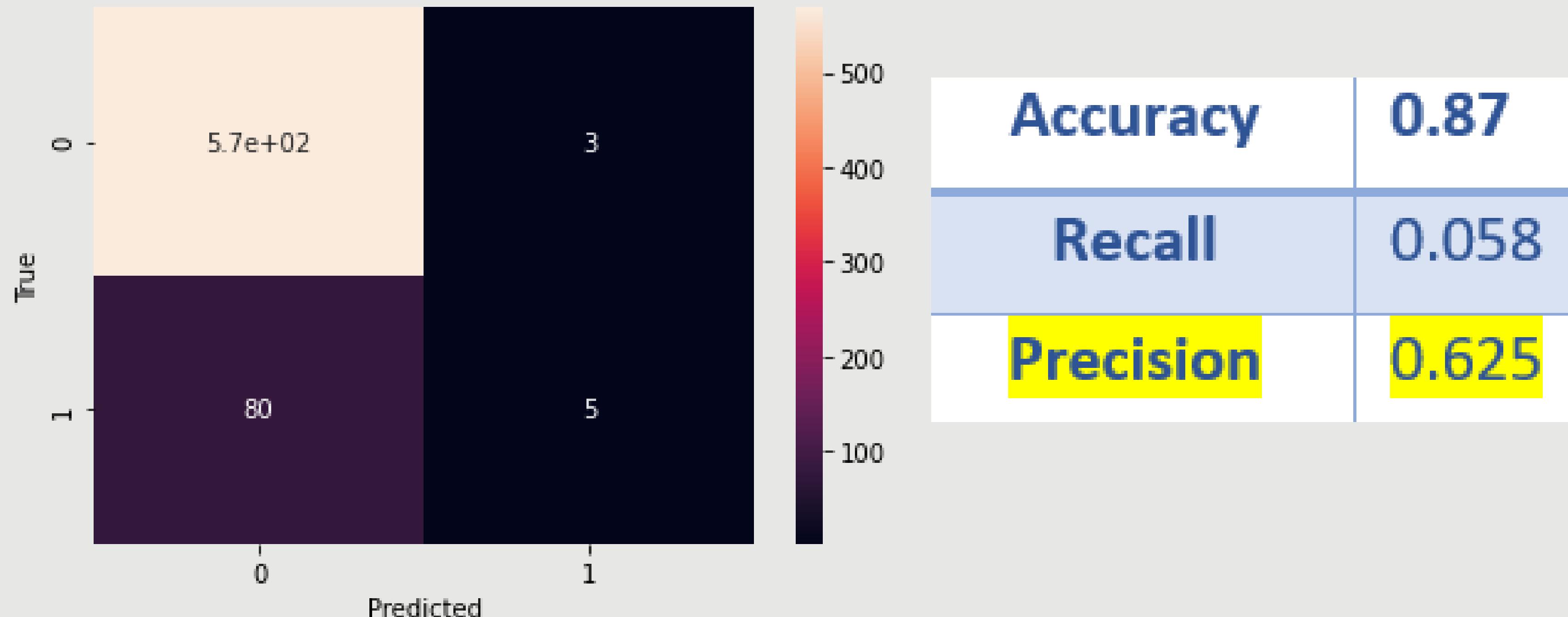


<b>Accuracy</b>	<b>0.87</b>
<b>Recall</b>	<b>0.035</b>
<b>Precision</b>	<b>0.75</b>

# Logistic Regression Model



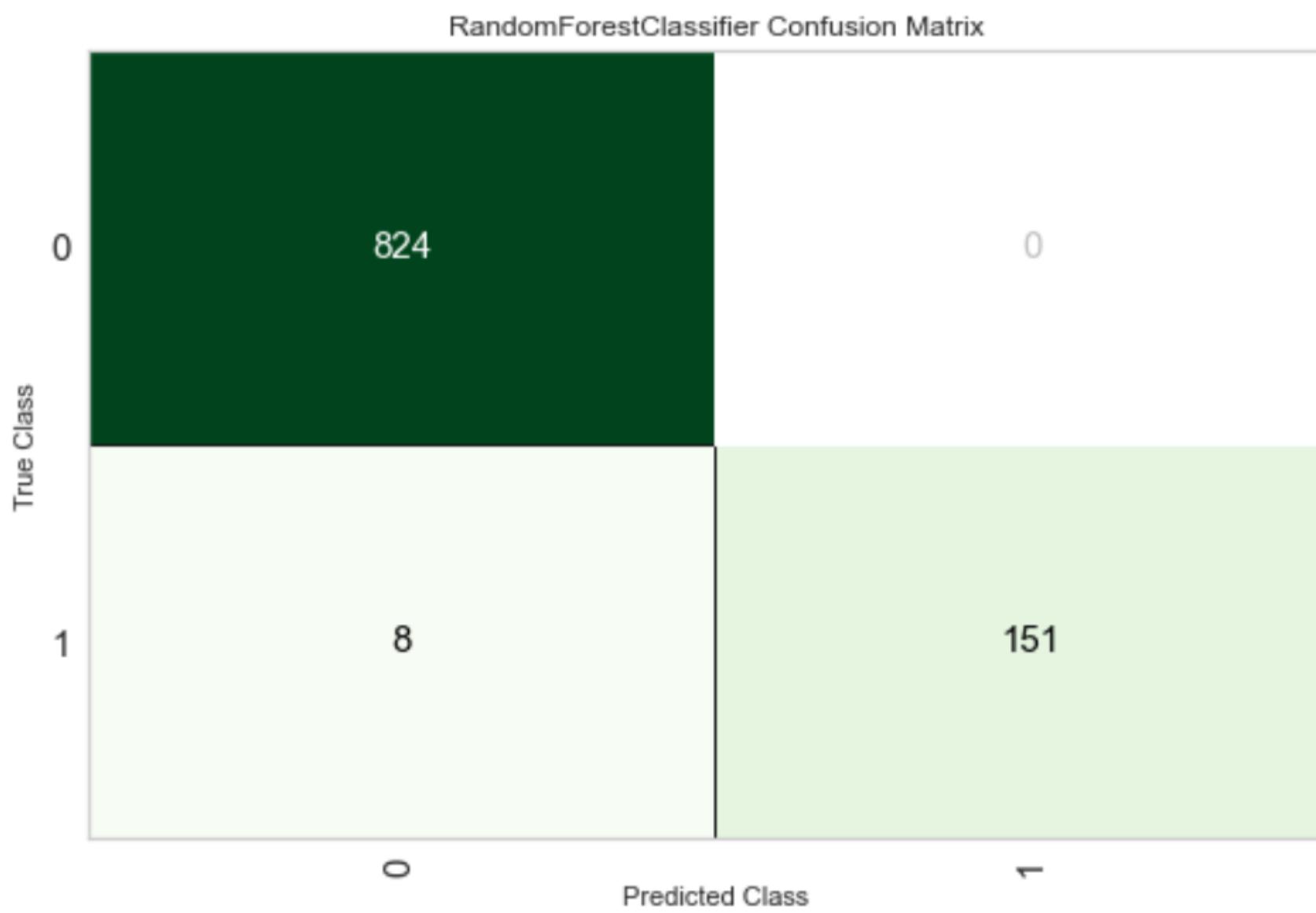
# Logistic Regression Model



# Explore Model Classifiers

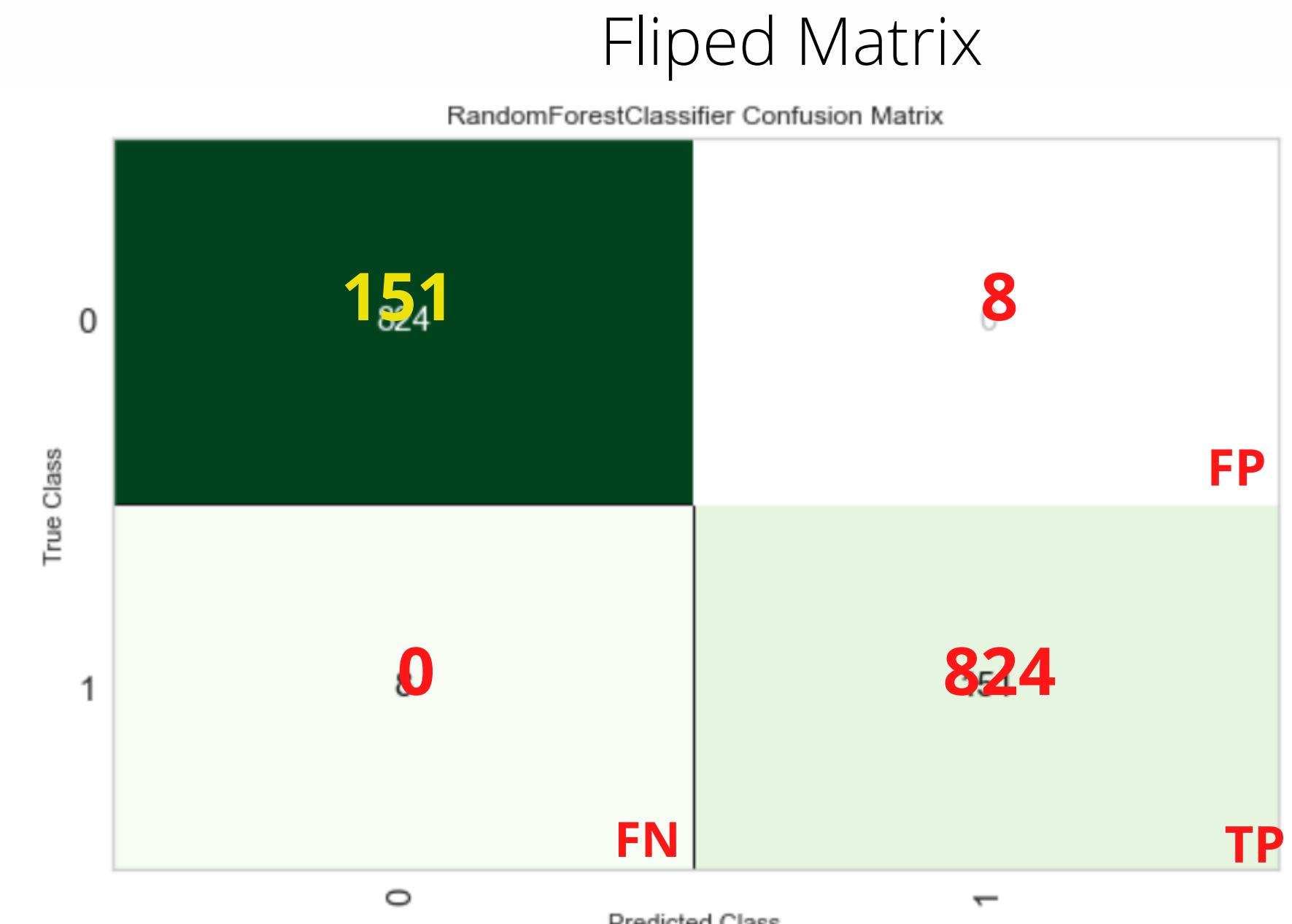
Model		Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC	TT (Sec)
<b>gbc</b>	Gradient Boosting Classifier	0.9983	0.9994	0.9968	0.9905	0.9935	0.9925	0.9926	0.1090
<b>ada</b>	Ada Boost Classifier	0.9974	0.9975	0.9839	0.9968	0.9902	0.9887	0.9888	0.0380
<b>rf</b>	Random Forest Classifier	0.9970	0.9996	0.9774	1.0000	0.9885	0.9867	0.9869	0.0730
<b>xgboost</b>	Extreme Gradient Boosting	0.9965	0.9971	0.9774	0.9969	0.9869	0.9849	0.9850	0.1540
<b>catboost</b>	CatBoost Classifier	0.9961	1.0000	0.9839	0.9876	0.9855	0.9832	0.9834	0.4430
<b>lightgbm</b>	Light Gradient Boosting Machine	0.9956	0.9966	0.9839	0.9842	0.9838	0.9813	0.9814	0.4370
<b>dt</b>	Decision Tree Classifier	0.9948	0.9874	0.9774	0.9842	0.9806	0.9776	0.9777	0.0080
<b>et</b>	Extra Trees Classifier	0.9765	0.9962	0.8609	0.9625	0.9074	0.8940	0.8967	0.0570
<b>nb</b>	Naive Bayes	0.8849	0.9548	0.9028	0.5510	0.6823	0.6177	0.6470	0.1150
<b>qda</b>	Quadratic Discriminant Analysis	0.8849	0.9498	0.8738	0.5515	0.6745	0.6095	0.6343	0.0060
<b>knn</b>	K Neighbors Classifier	0.7658	0.8912	0.9060	0.3596	0.5133	0.3965	0.4702	0.1550
<b>lr</b>	Logistic Regression	0.7514	0.8440	0.8027	0.3292	0.4664	0.3401	0.3975	0.3800

# Classification Metrics



**F1=0.988,**  
**MCC=0.98**

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$$



**F1=0.995**  
**MCC=0.98**

# AutoML



**LuciferML:**

Preprocessing data

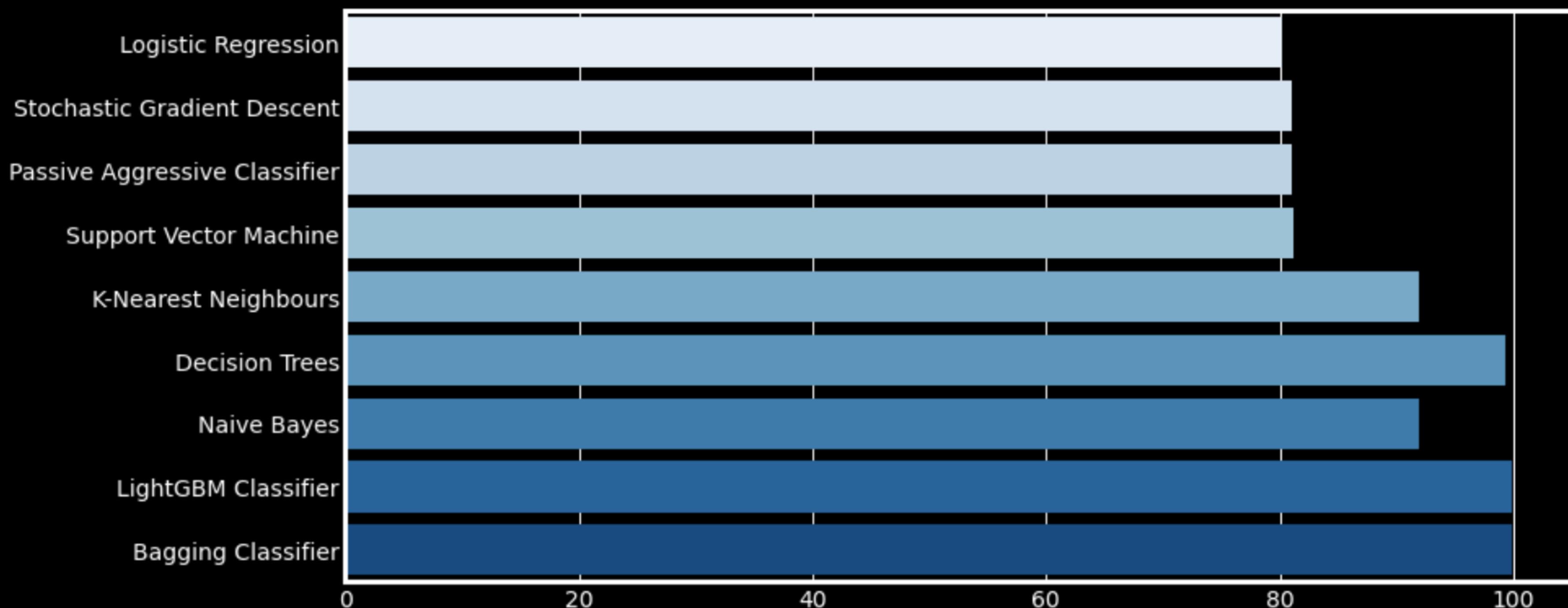
Fix imbalance(ADASYN)

**TPOTClassifier:**

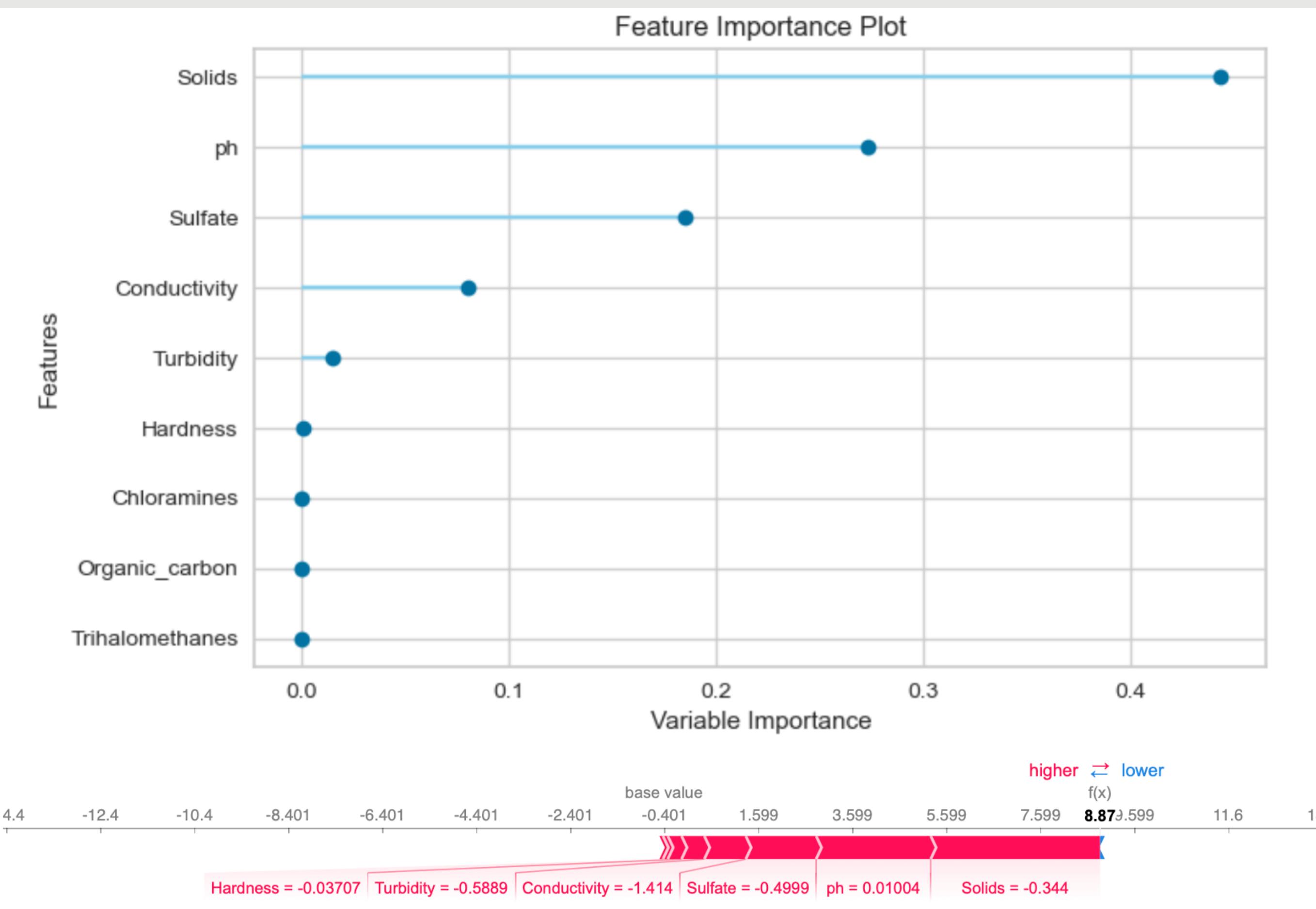
pipelines to find the best

**Hyperopt:**

parameter optimization



# Feature Importance



# Demo



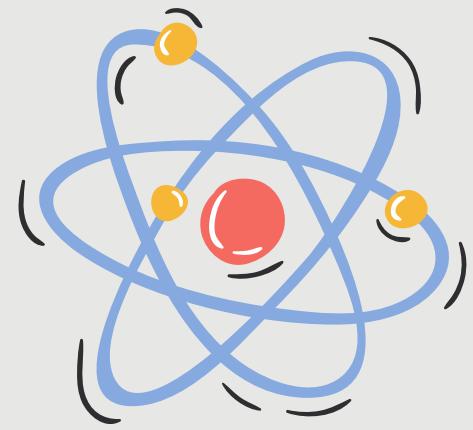
# Water Quality Check



AKN is a material testing laboratory that focuses on water quality tests and classification. Our next step, to expand the business research and development that to start a project to automate the water classification process. Access to safe drinking-water is essential to health, a basic human right and a component of effective policy for health protection.

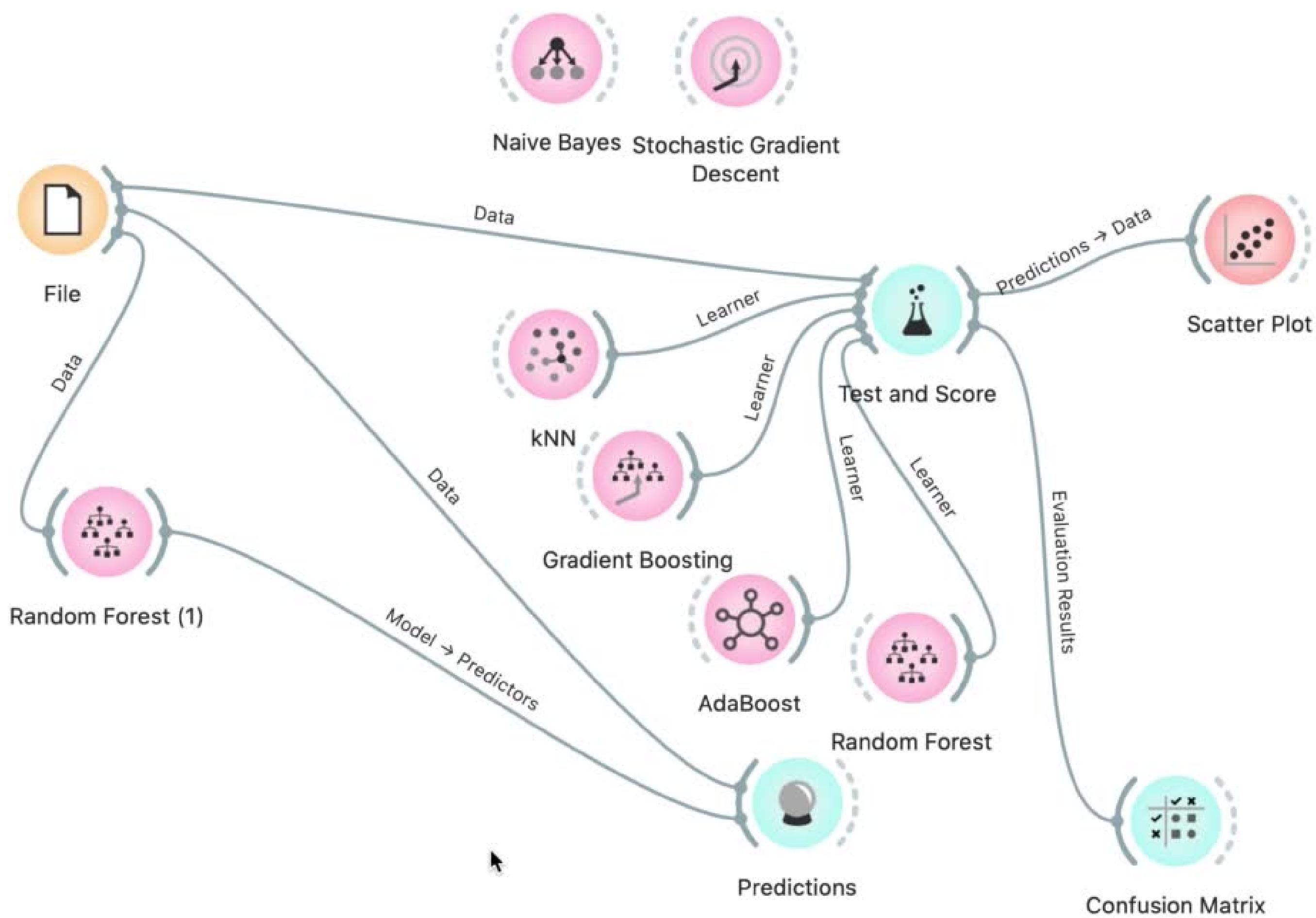
Sample values for the input

< Manage app



# The Laboratory





# Conclusion

- Companies need to focus on water quality and standers.
- Substances are sensitive to change results of safe water.
- Increase the accuracy of the model by adding more standers



# Future Work



- Accessing more data for water quality.
- Integrate the lab with our app.
- Extract quality by real sampling from AKN water lab.



Thank You For  
Listening.

Do you have  
any questions?