Model Card

Model Description

Input features:

pH, Hardness, TDS, Chloramines, Sulphate, Conductivity, Organic Carbon, Trihalomethanes, Turbidity

Output predictor:

Potability

Model Architecture:

The model used is sklearn's support vector classification, *svm.SVC()*

Best performance came from the library's default setting, *kernel='rbf', C=1, gamma='scale'*.

Performance

The model is trained using a training dataset that has been split from the testing dataset prior to any pre-processing. Three performance metrics are given, using *sklearn import metrics*. All performance metrics are derived from the test dataset, as follows:

metrics.accuracy_score(y_test, y_pred)

Accuracy: 0.6581892166836215

metrics.recall_score(y_test, y_pred)

Recall: 0.255

metrics.precision_score(y_test, y_pred)

Precision: 0.7285714285714285

Limitations

The model has been trained and tested using a small dataset that is suspected to be synthetic in origin.

Trade-offs

Since we are unsure as to the origin and validity of the dataset, it must be assumed that it is of dubious origin and therefore only to be used for educational purposes.