SPRINT 3
Project Deliverables (Model Building Code & Evaluation)

Team ID	PNT2022TMID50914
Project Name	Efficient Water Quality Analysis & Prediction using Machine Learning

Using the best accuracy algorith (SVC) we are going to train our model for deployment:

Support vector Machine

```python
# Support vector classifier
from sklearn.svm import SVC
svc_classifier = SVC(class_weight = "balanced" )
svc_classifier.fit(X_train_final, y_train)
y_pred_scv = svc_classifier.predict(X_test_final)
accuracy_score(y_test, y_pred_scv)
```

0.6225

```python
print(classification_report(y_test, y_pred_scv))
```

```
              precision    recall  f1-score   support

           0       0.70      0.69      0.70       497
           1       0.50      0.50      0.50       303

    accuracy                           0.62       800
   macro avg       0.60      0.60      0.60       800
weighted avg       0.62      0.62      0.62       800
```

TERMINAL

```
WARNING: This is a development server. Do not use it in a production deployment. Use a production WSGI server instead.
 * Running on http://127.0.0.1:5000
Press CTRL+C to quit
 * Restarting with stat
C:\Users\HOME\AppData\Local\Programs\Python\Python39\lib\site-packages\sklearn\base.py:329: UserWarning: Trying to unpickle estimator StandardScaler from version 0.24.0 when using version 1.1.3. This might l
ead to breaking code or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainability-limitations
  warnings.warn(
 * Debugger is active!
 * Debugger PIN: 873-600-839
127.0.0.1 - - [11/Nov/2022 13:57:12] "GET / HTTP/1.1" 200 -
C:\Users\HOME\AppData\Local\Programs\Python\Python39\lib\site-packages\sklearn\base.py:443: UserWarning: X has feature names, but StandardScaler was fitted without feature names
  warnings.warn(
127.0.0.1 - - [11/Nov/2022 13:57:43] "POST /predict HTTP/1.1" 200 -
```

File   Edit   Selection   View   Go   Run   Terminal   Help

Water_quality.ipynb - Water Quality - Visual Studio Code

app.py   Water_quality.ipynb ×   home.html 2   water_potability.csv

EXPLORER

WATER QUALITY

> static
> templates
  <> home.html   2
  app.py
  model.pkl
  my_scaler.save
  Procfile
  requirements.txt
  water_potability.csv
  Water_quality.ipynb

Water_quality.ipynb > Problem Statement > Task 1

+ Code   + Markdown | ▷ Run All   ≡ Clear Outputs of All Cells   ↻ Restart   ⊟ Variables   ≡ Outline   ···

Python 3.9.1 64-bit

```python
cm = confusion_matrix(y_test, y_pred_scv)
plt.title('Heatmap of Confusion Matrix', fontsize = 12)
sns.heatmap(cm, annot = True, fmt = "d")
plt.show()
```

[37]   ✓ 1.2s

Python

Heatmap of Confusion Matrix

	0	1
0	345	152
1	150	153

PROBLEMS ❷   OUTPUT   DEBUG CONSOLE   TERMINAL   JUPYTER

cmd
Code

WARNING: This is a development server. Do not use it in a production deployment. Use a production WSGI server instead.
 * Running on http://127.0.0.1:5000
Press CTRL+C to quit
 * Restarting with stat
C:\Users\HOME\AppData\Local\Programs\Python\Python39\lib\site-packages\sklearn\base.py:329: UserWarning: Trying to unpickle estimator StandardScaler from version 0.24.0 when using version 1.1.3. This might l
ead to breaking code or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainability-limitations
  warnings.warn(
 * Debugger is active!
 * Debugger PIN: 873-600-839
127.0.0.1 - - [11/Nov/2022 13:57:12] "GET / HTTP/1.1" 200 -
C:\Users\HOME\AppData\Local\Programs\Python\Python39\lib\site-packages\sklearn\base.py:443: UserWarning: X has feature names, but StandardScaler was fitted without feature names
  warnings.warn(
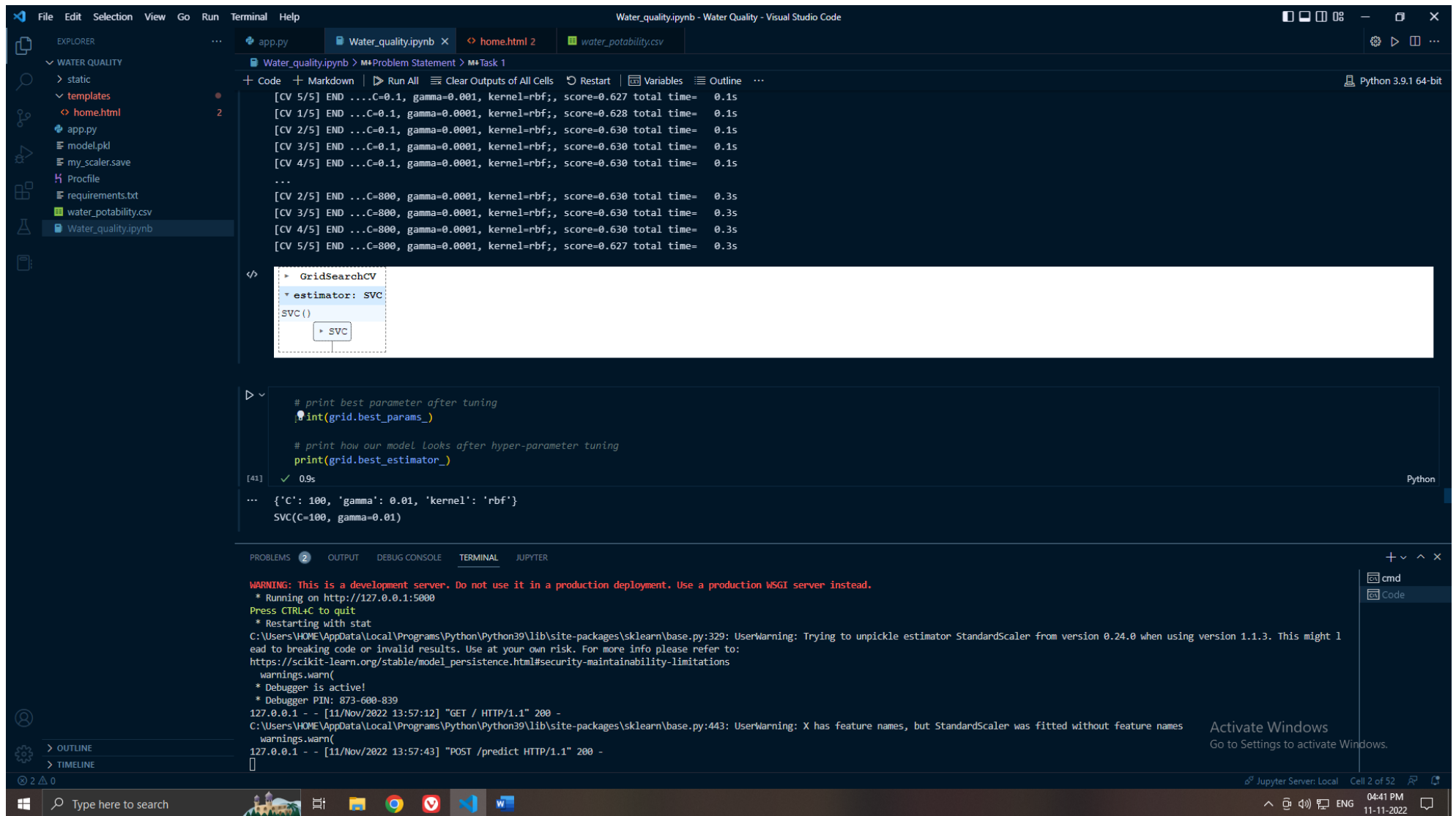127.0.0.1 - - [11/Nov/2022 13:57:43] "POST /predict HTTP/1.1" 200 -

> OUTLINE
> TIMELINE

⊗ 2 ⚠ 0   Jupyter Server: Local   Cell 2 of 52

Type here to search

04:41 PM
11-11-2022
ENG

File   Edit   Selection   View   Go   Run   Terminal   Help

app.py   Water_quality.ipynb ✕   home.html 2   water_potability.csv

Water_quality.ipynb > M↓ Problem Statement > M↓ Task 1

+ Code   + Markdown | ▷ Run All   ⊟ Clear Outputs of All Cells   ↻ Restart | ⊟ Variables   ⬚ Outline   ⋯   Python 3.9.1 64-bit

Hyperparameter Tuning with Support vector Machine

```python
# defining parameter range
param_grid = {'C': [0.1, 1, 10, 100, 200 , 400 , 600 , 800],
              'gamma': [1, 0.1, 0.01, 0.001, 0.0001],
              'kernel': ['rbf']}
```
[38]   ✓   0.9s   Python

```python
from sklearn.model_selection import GridSearchCV
```
[39]   ✓   0.6s   Python

```python
grid = GridSearchCV(SVC(), param_grid, refit = True, verbose = 3)

# fitting the model for grid search
grid.fit(X_train_final, y_train)
```
[40]   ✓   1m 14.9s   Python

⋯   Output exceeds the size limit. Open the full output data in a text editor

Fitting 5 folds for each of 40 candidates, totalling 200 fits
[CV 1/5] END ....C=0.1, gamma=1, kernel=rbf;, score=0.628 total time=   0.2s
[CV 2/5] END ....C=0.1, gamma=1, kernel=rbf;, score=0.630 total time=   0.2s
[CV 3/5] END ....C=0.1, gamma=1, kernel=rbf;, score=0.630 total time=   0.1s
[CV 4/5] END ....C=0.1, gamma=1, kernel=rbf;, score=0.630 total time=   0.1s
[CV 5/5] END ....C=0.1, gamma=1, kernel=rbf;, score=0.627 total time=   0.1s
[CV 1/5] END ....C=0.1, gamma=1, kernel=rbf;, score=0.628 total time=   0.1s

PROBLEMS ②   OUTPUT   DEBUG CONSOLE   TERMINAL   JUPYTER

WARNING: This is a development server. Do not use it in a production deployment. Use a production WSGI server instead.
 * Running on http://127.0.0.1:5000
Press CTRL+C to quit
 * Restarting with stat
C:\Users\HOME\AppData\Local\Programs\Python\Python39\lib\site-packages\sklearn\base.py:329: UserWarning: Trying to unpickle estimator StandardScaler from version 0.24.0 when using version 1.1.3. This might l
ead to breaking code or invalid results. Use at your own risk. For more info please refer to:
https://scikit-learn.org/stable/model_persistence.html#security-maintainability-limitations
  warnings.warn(
 * Debugger is active!
 * Debugger PIN: 873-600-839
127.0.0.1 - - [11/Nov/2022 13:57:12] "GET / HTTP/1.1" 200 -
C:\Users\HOME\AppData\Local\Programs\Python\Python39\lib\site-packages\sklearn\base.py:443: UserWarning: X has feature names, but StandardScaler was fitted without feature names
  warnings.warn(
127.0.0.1 - - [11/Nov/2022 13:57:43] "POST /predict HTTP/1.1" 200 -

cmd
Code

Activate Windows
Go to Settings to activate Windows.

> OUTLINE
> TIMELINE

Now we have successfully trained our model , then we will save our trained model using _PICKLE library_

Model.pkl file: (this is the file where we will be saving our trained model)

Now it contains all the necessary files ..