

## SPRINT 3

### Project Deliverables (Model Building Code & Evaluation)

|              |  |
|--------------|--|
| Team ID      | PNT2022TMID37191   |
| Project Name | Efficient Water Quality Analysis & Prediction using Machine Learning |

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

The screenshot shows a Google Colab notebook titled 'Untitled3.ipynb'. The notebook contains the following code and output:

```
[33] macro avg      0.59      0.62      0.56      800
      weighted avg      0.59      0.62      0.56      800
```

**Support vector Machine**

```
# 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
```

```
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     |

```
[36] cm = confusion_matrix(y_test, y_pred_scv)
      plt.title('Heatmap of Confusion Matrix', fontsize = 12)
```

completed at 12:05

The bottom of the image shows a Windows taskbar with the date 18-11-2022 and time 12:20.

IBM

IBM

IBM-EPBL/IBM-Projec

IBM-Project-12969-1

Online HTML Editor

Water Quality

Untitled3.ipynb - Cola

+

colab.research.google.com/drive/16-oR9CY8sgFGnz6nJ9Qtg\_BefUuMXOJR#scrollTo=IZxX1SP-bW6X

Untitled3.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment Share

RAM Disk

Editing

+ Code + Text

[35]

10.500.500.50303

accuracy0.62800

macro avg0.60800

weighted avg0.62800

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()

Heatmap of Confusion Matrix

01

345152

150153

01

150325

Hyperparameter Tuning with Support vector Machine

0s completed at 12:05

model training an....pdf

Untitled3.ipynb

Show all

29°C Haze

Search

ENG IN

12:22 18-11-2022

IBM

IBM

IBM-EPBL/IBM-Projec

IBM-Project-12969-16

Online HTML Editor

Water Quality

Untitled3.ipynb - Colab

colab.research.google.com/drive/16-oR9CY8sgFGnz6nJ9Qtg\_BefUuMXOJR#scrollTo=IZxX1SP-bW6X

Untitled3.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

RAM Disk

Editing

Hyperparameter Tuning with Support vector Machine

[37] # 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] from sklearn.model\_selection import GridSearchCV

[39] grid = GridSearchCV(SVC(), param\_grid, refit = True, verbose = 3)  
  
# fitting the model for grid search  
grid.fit(X\_train\_final, y\_train)  
  
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= 0.2s  
[CV 4/5] END .....C=0.1, gamma=1, kernel=rbf; score=0.630 total time= 0.2s  
[CV 5/5] END .....C=0.1, gamma=1, kernel=rbf; score=0.627 total time= 0.2s  
[CV 1/5] END .....C=0.1, gamma=0.1, kernel=rbf; score=0.628 total time= 0.1s  
[CV 2/5] END .....C=0.1, gamma=0.1, kernel=rbf; score=0.630 total time= 0.1s  
[CV 3/5] END .....C=0.1, gamma=0.1, kernel=rbf; score=0.630 total time= 0.2s  
[CV 4/5] END .....C=0.1, gamma=0.1, kernel=rbf; score=0.630 total time= 0.1s  
[CV 5/5] END .....C=0.1, gamma=0.1, kernel=rbf; score=0.627 total time= 0.1s  
[CV 1/5] END .....C=0.1, gamma=0.01, kernel=rbf; score=0.628 total time= 0.1s  
[CV 2/5] END .....C=0.1, gamma=0.01, kernel=rbf; score=0.630 total time= 0.1s

0s completed at 12:05

model training an....pdf

Untitled3.ipynb

Show all

29°C Haze

Search

ENG IN

12:24 18-11-2022

IBM x IBM x IBM-EPBL/IBM-Projec x IBM-Project-12969-1 x Online HTML Editor x Water Quality x Untitled3.ipynb - Colab x +

colab.research.google.com/drive/16-oR9CY8sgFGnz6nJ9Qtg\_BefUuMXOJR#scrollTo=IZxX1SP-bW6X

Untitled3.ipynb ☆

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

RAM Disk Editing

```
[CV 5/5] END ...C=600, gamma=0.0001, kernel=rbf;; score=0.627 total time= 0.3s
[CV 1/5] END .....C=800, gamma=1, kernel=rbf;; score=0.610 total time= 0.2s
[CV 2/5] END .....C=800, gamma=1, kernel=rbf;; score=0.609 total time= 0.2s
[CV 3/5] END .....C=800, gamma=1, kernel=rbf;; score=0.609 total time= 0.2s
[CV 4/5] END .....C=800, gamma=1, kernel=rbf;; score=0.609 total time= 0.2s
[CV 5/5] END .....C=800, gamma=1, kernel=rbf;; score=0.627 total time= 0.2s
[CV 1/5] END .....C=800, gamma=0.1, kernel=rbf;; score=0.543 total time= 1.6s
[CV 2/5] END .....C=800, gamma=0.1, kernel=rbf;; score=0.539 total time= 1.9s
[CV 3/5] END .....C=800, gamma=0.1, kernel=rbf;; score=0.550 total time= 1.2s
[CV 4/5] END .....C=800, gamma=0.1, kernel=rbf;; score=0.558 total time= 1.2s
[CV 5/5] END .....C=800, gamma=0.1, kernel=rbf;; score=0.571 total time= 1.2s
[CV 1/5] END .....C=800, gamma=0.01, kernel=rbf;; score=0.628 total time= 0.6s
[CV 2/5] END .....C=800, gamma=0.01, kernel=rbf;; score=0.638 total time= 0.6s
[CV 3/5] END .....C=800, gamma=0.01, kernel=rbf;; score=0.670 total time= 0.8s
[CV 4/5] END .....C=800, gamma=0.01, kernel=rbf;; score=0.649 total time= 0.7s
[CV 5/5] END .....C=800, gamma=0.01, kernel=rbf;; score=0.657 total time= 0.6s
[CV 1/5] END .....C=800, gamma=0.001, kernel=rbf;; score=0.647 total time= 0.2s
[CV 2/5] END .....C=800, gamma=0.001, kernel=rbf;; score=0.638 total time= 0.2s
[CV 3/5] END .....C=800, gamma=0.001, kernel=rbf;; score=0.654 total time= 0.3s
[CV 4/5] END .....C=800, gamma=0.001, kernel=rbf;; score=0.649 total time= 0.2s
[CV 5/5] END .....C=800, gamma=0.001, kernel=rbf;; score=0.654 total time= 0.2s
[CV 1/5] END .....C=800, gamma=0.0001, kernel=rbf;; score=0.628 total time= 0.3s
[CV 2/5] END ...C=800, gamma=0.0001, kernel=rbf;; score=0.630 total time= 0.3s
[CV 3/5] END ...C=800, gamma=0.0001, kernel=rbf;; score=0.630 total time= 0.3s
[CV 4/5] END ...C=800, gamma=0.0001, kernel=rbf;; score=0.630 total time= 0.3s
[CV 5/5] END ...C=800, gamma=0.0001, kernel=rbf;; score=0.627 total time= 0.3s
GridSearchCV(estimator=SVC(),
               param_grid={'C': [0.1, 1, 10, 100, 200, 400, 600, 800],
                           'gamma': [1, 0.1, 0.01, 0.001, 0.0001],
                           'kernel': ['rbf']},
               verbose=3)
```

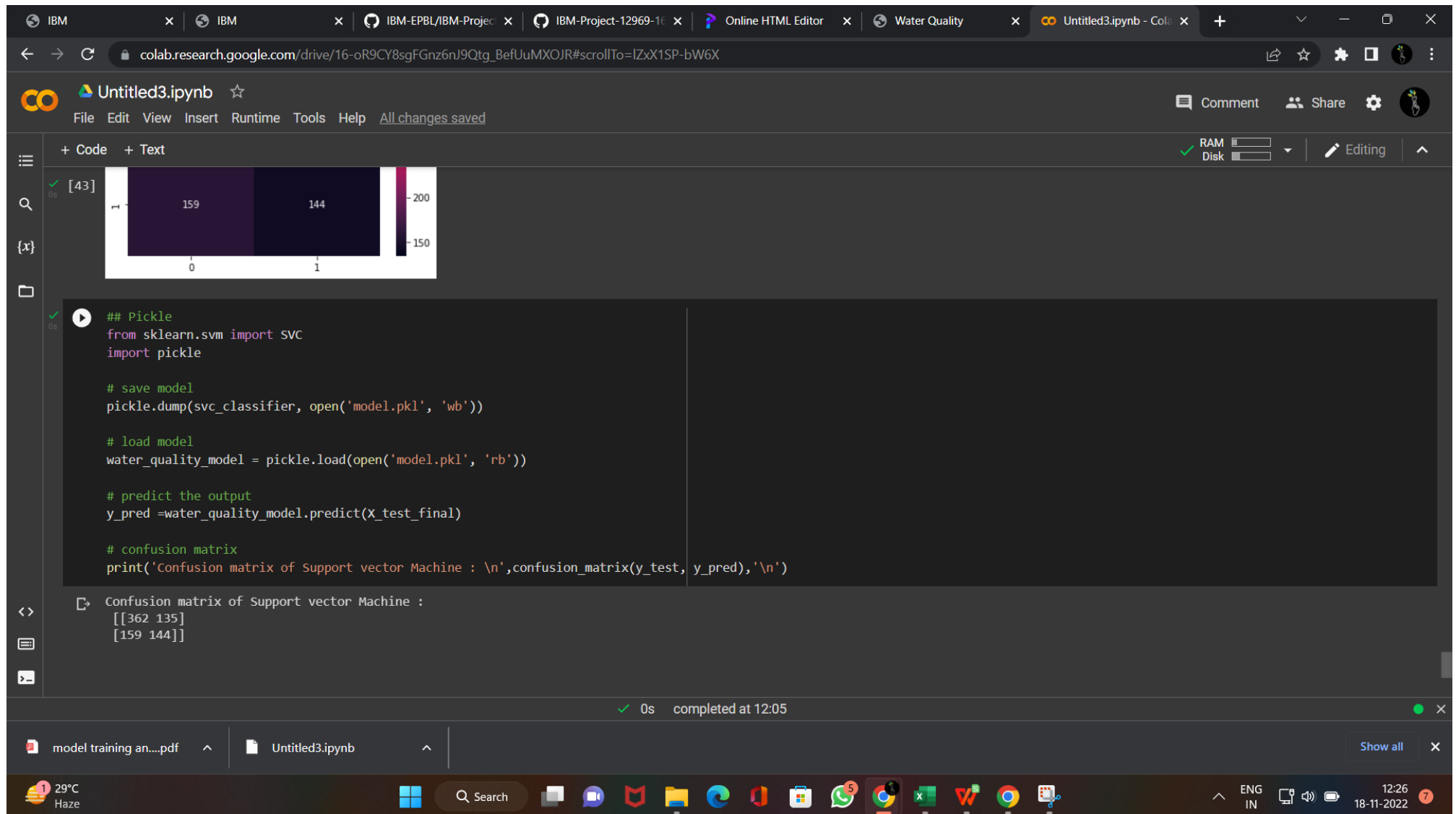
0s completed at 12:05

model training an....pdf ^ Untitled3.ipynb ^ Show all x

29°C Haze Search 12:25 18-11-2022

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 )**



The screenshot displays a Google Colab notebook titled "Untitled3.ipynb". The interface includes a top navigation bar with tabs for "IBM", "Online HTML Editor", "Water Quality", and the current notebook. The notebook's menu bar shows "File", "Edit", "View", "Insert", "Runtime", "Tools", and "Help". On the left, a sidebar contains icons for file management and a search bar. The main workspace is divided into two sections: a code editor and an output area.

The code editor contains the following Python code:

```
## Pickle
from sklearn.svm import SVC
import pickle

# save model
pickle.dump(svc_classifier, open('model.pkl', 'wb'))

# load model
water_quality_model = pickle.load(open('model.pkl', 'rb'))

# predict the output
y_pred = water_quality_model.predict(X_test_final)

# confusion matrix
print('Confusion matrix of Support vector Machine : \n', confusion_matrix(y_test, y_pred), '\n')
```

The output area shows the result of the confusion matrix calculation:

```
Confusion matrix of Support vector Machine :
[[362 135]
 [159 144]]
```

Below the code editor, a status bar indicates "0s completed at 12:05". At the bottom, a taskbar shows the Windows Start button, a search bar, and various application icons, including a weather widget showing "29°C Haze". The system clock in the bottom right corner displays "12:26 18-11-2022".