```
from google.colab import drive
drive.mount('/content/drive')
```

```
Mounted at /content/drive
```

```
dir_path = '/content/drive/MyDrive/dataset/lab 6/'
```

```
import pandas as pd
import numpy as np
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy_score, classification_report
import cv2
import xgboost as xgb
```

Load dataset

```
X_train = pd.read_csv(dir_path + "N_X_train.csv",header=None)
Y_train = pd.read_csv(dir_path + "N_Y_train.csv",header=None)
X_test = pd.read_csv(dir_path + "N_X_test.csv",header=None)
# Y_test = pd.read_csv(dir_path + "N_Y_test.csv",header=None)
```

```
print(X_train.shape)
print(Y_train.shape)
print(X_test.shape)
# print(Y_test.shape)
```

```
(250202, 115)
(250202, 1)
(23040, 115)
```

XGBoost

```
# Định nghĩa và huấn luyện mô hình XGBoost
model = xgb.XGBClassifier(n_estimators=100, random_state=42, use_label_encoder=False, eval_metric='logloss')
model.fit(X_train, Y_train.values.ravel())
```

```
XGBClassifier(base_score=None, booster=None, callbacks=None,
              colsample_bylevel=None, colsample_bynode=None,
              colsample_bytree=None, device=None, early_stopping_rounds=None,
              enable_categorical=False, eval_metric='logloss',
              feature_types=None, gamma=None, grow_policy=None,
              importance_type=None, interaction_constraints=None,
              learning_rate=None, max_bin=None, max_cat_threshold=None,
              max_cat_to_onehot=None, max_delta_step=None, max_depth=None,
              max_leaves=None, min_child_weight=None, missing=nan,
              monotone_constraints=None, multi_strategy=None, n_estimators=100,
              n_jobs=None, num_parallel_tree=None, random_state=42, ...)
```

```
# Dự đoán nhãn cho tập kiểm thử
y_pred = model.predict(X_test)

# Chuyển đổi dự đoán thành DataFrame
Y_test_pred = pd.DataFrame(y_pred)
```

```
f = np.zeros((180, 128))
```

```
f
```

```
array([[0., 0., 0., ..., 0., 0., 0.],
       [0., 0., 0., ..., 0., 0., 0.],
       [0., 0., 0., ..., 0., 0., 0.],
       ...,
       [0., 0., 0., ..., 0., 0., 0.],
       [0., 0., 0., ..., 0., 0., 0.],
       [0., 0., 0., ..., 0., 0., 0.]])
```

```
from google.colab.patches import cv2_imshow
a = np.array(y_pred).reshape(23040, 1)
a = a.reshape(180,128)
a[a == 1] = 255
cv2_imshow(a)
```

Random Forest

```
# Định nghĩa và huấn luyện mô hình
model = RandomForestClassifier(n_estimators=100, random_state=42)
model.fit(X_train, Y_train.values.ravel())
```

```
RandomForestClassifier
RandomForestClassifier(random_state=42)
```

```
# Dự đoán nhãn cho tập kiểm thử
y_pred = model.predict(X_test)

# Chuyển đổi dự đoán thành DataFrame
Y_test_pred = pd.DataFrame(y_pred)
```

```
f = np.zeros((180, 128))
```

```
f
```

```
array([[0., 0., 0., ..., 0., 0., 0.],
       [0., 0., 0., ..., 0., 0., 0.],
       [0., 0., 0., ..., 0., 0., 0.],
       ...,
       [0., 0., 0., ..., 0., 0., 0.],
       [0., 0., 0., ..., 0., 0., 0.],
       [0., 0., 0., ..., 0., 0., 0.]])
```

```
from google.colab.patches import cv2_imshow
a = np.array(y_pred).reshape(23040, 1)
a = a.reshape(180,128)
a[a == 1] = 255
cv2_imshow(a)
```