

Tue May 24 07:22:31 2022

Processes:						
GPU	GI	CI	PID	Type	Process name	GPU Memory Usage
	ID	ID				
=====						
No running processes found						

downloading python_cabnet-3.1.1-py3-none-any.whl (33 KB)

Downloading optuna-2.10.0-py3-none-any.whl (308 kB)

```
Requirement already satisfied: scikit-learn>0.21 in /usr
```

Requirement already satisfied: numpy<2.0,>=1.17 in /usr/local/lib/python3.7/dist-p

Requirement already satisfied: scipy>1.4 in /usr/local/lib/python3.7/dist-packages

Requirement already satisfied: joblib>=0.11 in /usr/local/lib/python3.7/dist-packages

Requirement already satisfied: typing-extensions in /usr/local/lib/python3.7/dist-

```
Requirement already satisfied: PyYAML in /usr/local/lib/python3.7/dist-packages (1.7.1)
Collecting colorlog
```

```
Downloading colorlog-6.6.0-py2.py3-none-any.whl (11 kB)
```

Downloading cmaes-0.8.2-py3-none-any.whl (15 kB)

Collecting alembic

210 kB 58.2 MB/s

Collecting cliff

81 kB 11.4 MB/s

Requirement already satisfied: greenlet!=0.4.17 in /usr/local/lib/python3.7/dist-p

Collecting Mako

```
Installing Make 1.2.0 py3-none-any.whl (70 KB)
```

78 kB 2.9 MB/s

Collecting autopage>=0.4.0

```

    Downloading autopage-0.5.0-py3-none-any.whl (29 kB)
Collecting cmd2>=1.0.0
    Downloading cmd2-2.4.1-py3-none-any.whl (146 kB)
    |████████████████████████████████████████| 146 kB 52.8 MB/s
Collecting pbr!=2.1.0,>=2.0.0
    Downloading pbr-5.9.0-py2.py3-none-any.whl (112 kB)
    |████████████████████████████████████████| 112 kB 58.0 MB/s
Requirement already satisfied: PrettyTable>=0.7.2 in /usr/local/lib/python3.7/dist-packages
Collecting stevedore>=2.0.1
    Downloading stevedore-3.5.0-py3-none-any.whl (49 kB)
    |████████████████████████████████████████| 49 kB 6.3 MB/s
Collecting pyperclip>=1.6
    Downloading pyperclip-1.8.2.tar.gz (20 kB)
Requirement already satisfied: wcwidth>=0.1.7 in /usr/local/lib/python3.7/dist-packages
Requirement already satisfied: attrs>=16.3.0 in /usr/local/lib/python3.7/dist-packages
Requirement already satisfied: zipp>=0.5 in /usr/local/lib/python3.7/dist-packages
Requirement already satisfied: MarkupSafe>=0.9.2 in /usr/local/lib/python3.7/dist-packages
Building wheels for collected packages: pyperclip
  Building wheel for pyperclip (setup.py) ... done
  Created wheel for pyperclip: filename=pyperclip-1.8.2-py3-none-any.whl size=1113
  Stored in directory: /root/.cache/pip/wheels/9f/18/84/8f69f8b08169c7bae2dde6bd7d
Successfully built pyperclip
Installing collected packages: pyperclip, pbr, stevedore, Mako, cmd2, autopage, co
Successfully installed Mako-1.2.0 alembic-1.7.7 autopage-0.5.0 cliff-3.10.1 cmaes-

```

```

import pandas as pd
import numpy as np
from pytorch_tabnet.tab_model import TabNetClassifier
from sklearn.metrics import accuracy_score, classification_report
import optuna as opt
import torch
import os
import joblib

def make_save_cv_model(i,model_name,model,best_params,optim,output_path="./drive/MyDrive/S

''' This function saves cross validation model in the corresponding directory ( if the

if os.path.exists(os.path.join(output_path,f"{i}_{model_name}_{optim}")):
    joblib.dump(model, os.path.join(output_path,f"{i}_{model_name}_{optim}/{i}_model.z
    with open(os.path.join(output_path,f"{i}_{model_name}_{optim}/model_params.txt"),'
        file.write(str(best_params))
else:
    os.mkdir(os.path.join(output_path,f"{i}_{model_name}_{optim}"))
    joblib.dump(model, os.path.join(output_path,f"{i}_{model_name}_{optim}/{i}_model.z
    with open(os.path.join(output_path,f"{i}_{model_name}_{optim}/model_params.txt"),'
        file.write(str(best_params))

def train(fold_dict,fold,model_name,sc_df,tar_col,optim,optim_trial,k_folds=10,tar_cols=""

''' this function is used to train the model with parameters optimization using optuna

y = sc_df[tar_col]

```

```

x = sc_df.drop([tar_col],axis=1)
model_name = model_name
def objective(trial):
    train_index = fold_dict[fold]["train"]
    test_index = fold_dict[fold]["test"]
    clf = TabNetClassifier(n_d=trial.suggest_int("n_d", 8, 64),
                           n_a=trial.suggest_int("n_a", 8, 64),
                           n_steps = trial.suggest_int("n_steps",3,10),
                           gamma =trial.suggest_float("gamma", 1.0, 2.0),
                           n_independent = trial.suggest_int("n_independent",1,5),
                           n_shared = trial.suggest_int("n_shared",1,5),
                           momentum = trial.suggest_float("momentum", 0.01, 0.4),
                           optimizer_fn = torch.optim.Adam,
                           # scheduler_fn = torch.optim.lr_scheduler,
                           # scheduler_params = {"gamma" :trial.suggest_float("sch-gamm
                           verbose = verbose,
                           device_name = "auto"
                           )

    # print(f" train_index :: {train_index}")
    # print(f" test_index :: {test_index}")
    X_train,X_test = x.iloc[train_index,:], x.iloc[test_index,:]
    # print(X_train.shape, X_test.shape)
    X_train, X_test = X_train.to_numpy(dtype=np.float64), X_test.to_numpy(dtype=np.float
    Y_train, Y_test = y.iloc[train_index], y.iloc[test_index]
    Y_train, Y_test = Y_train.to_numpy(dtype=np.float64), Y_test.to_numpy(dtype=np.float
    print(Y_train.shape, Y_test.shape)
    clf.fit(X_train, Y_train,
            eval_set= [(X_test, Y_test)],
            eval_metric=['accuracy'])
    Y_pred = clf.predict(X_test)
    print(classification_report(Y_test, Y_pred, labels=[x for x in range(6)]))
    acc = accuracy_score(Y_pred, Y_test)
    return acc

print(f"Starting optimization for fold : [{fold}/{k_folds}]")
study = opt.create_study(direction='maximize')
study.optimize(objective, n_trials=optim_trial)
best_params = study.best_params
print(f" Best params for fold: [{fold}/{k_folds}]")
print(best_params)
joblib.dump(best_params,f"./drive/MyDrive/SOLAR_CELL/ML_PROCESSED_DATA/outputs/{model_
with open(f"./drive/MyDrive/SOLAR_CELL/ML_PROCESSED_DATA/outputs/{model_name}/best_par
print(f"Saved best_params at : outputs/{model_name}/best_params/fold_{fold}_best_param
train_index = fold_dict[fold]["train"]
test_index = fold_dict[fold]["test"]
X_train,X_test = x.iloc[train_index,:], x.iloc[test_index,:]
# print(X_train.shape, X_test.shape)
X_train, X_test = X_train.to_numpy(dtype=np.float64), X_test.to_numpy(dtype=np.float64
Y_train, Y_test = y.iloc[train_index], y.iloc[test_index]
Y_train, Y_test = Y_train.to_numpy(dtype=np.float64), Y_test.to_numpy(dtype=np.float64
clf_model = TabNetClassifier(**study.best_params)
clf_model.fit(X_train,Y_train)
Y_pred = clf_model.predict(X_test)
clf_report = classification_report(Y_test, Y_pred, labels=[x for x in range(6)])
with open(f"./drive/MyDrive/SOLAR_CELL/ML_PROCESSED_DATA/outputs/classification_report

```

```

accuracy = accuracy_score(Y_pred, Y_test)
with open(f"./drive/MyDrive/SOLAR_CELL/ML_PROCESSED_DATA/outputs/{model_name}/{model_n
try:
    print("[++] Saving the model and parameters in corresponding directories")
    make_save_cv_model(fold,model_name,clf_model,best_params,optim=optim)
except:
    print("[-] Failed to save the model")

use_df = pd.read_csv("./drive/MyDrive/SOLAR_CELL/ML_PROCESSED_DATA/outputs/data/trainable_
tar_col = "PCE_categorical"
model_name = "pytorch_tabnet"
optimizer = "Adam"
fold_dict = joblib.load("./drive/MyDrive/SOLAR_CELL/ML_PROCESSED_DATA/inputs/fold_vals/fol
fold = 6

train(fold_dict = fold_dict,
      fold = fold,
      model_name=model_name,
      sc_df=use_df,
      tar_col=tar_col,
      optim=optimizer,
      optim_trial = 15)
print(f"[++] Ended the training process for fold {fold}")

epoch 45 | loss: 0.48845 | 0:01:58s
epoch 46 | loss: 0.47144 | 0:02:00s
epoch 47 | loss: 0.46496 | 0:02:03s
epoch 48 | loss: 0.46084 | 0:02:05s
epoch 49 | loss: 0.45356 | 0:02:08s
epoch 50 | loss: 0.44674 | 0:02:10s
epoch 51 | loss: 0.44621 | 0:02:13s
epoch 52 | loss: 0.4578 | 0:02:16s
epoch 53 | loss: 0.45311 | 0:02:18s
epoch 54 | loss: 0.44346 | 0:02:21s
epoch 55 | loss: 0.44408 | 0:02:23s
epoch 56 | loss: 0.44867 | 0:02:26s
epoch 57 | loss: 0.43741 | 0:02:28s
epoch 58 | loss: 0.44922 | 0:02:31s
epoch 59 | loss: 0.44798 | 0:02:33s
epoch 60 | loss: 0.44962 | 0:02:36s
epoch 61 | loss: 0.43741 | 0:02:38s
epoch 62 | loss: 0.4282 | 0:02:41s
epoch 63 | loss: 0.41778 | 0:02:44s
epoch 64 | loss: 0.41568 | 0:02:46s
epoch 65 | loss: 0.41298 | 0:02:49s
epoch 66 | loss: 0.41474 | 0:02:51s
epoch 67 | loss: 0.42709 | 0:02:54s
epoch 68 | loss: 0.42211 | 0:02:56s
epoch 69 | loss: 0.41249 | 0:02:59s
epoch 70 | loss: 0.40765 | 0:03:01s
epoch 71 | loss: 0.40954 | 0:03:04s
epoch 72 | loss: 0.40961 | 0:03:06s
epoch 73 | loss: 0.39756 | 0:03:09s
epoch 74 | loss: 0.39453 | 0:03:11s
epoch 75 | loss: 0.39214 | 0:03:14s
epoch 76 | loss: 0.40054 | 0:03:16s
epoch 77 | loss: 0.39724 | 0:03:19s

```

```
epoch 78 | loss: 0.40694 | 0:03:21s
epoch 79 | loss: 0.40329 | 0:03:24s
epoch 80 | loss: 0.39786 | 0:03:26s
epoch 81 | loss: 0.39452 | 0:03:29s
epoch 82 | loss: 0.39643 | 0:03:31s
epoch 83 | loss: 0.39305 | 0:03:34s
epoch 84 | loss: 0.3882 | 0:03:36s
epoch 85 | loss: 0.388 | 0:03:39s
epoch 86 | loss: 0.3891 | 0:03:41s
epoch 87 | loss: 0.37989 | 0:03:44s
epoch 88 | loss: 0.39215 | 0:03:46s
epoch 89 | loss: 0.38917 | 0:03:49s
epoch 90 | loss: 0.38255 | 0:03:51s
epoch 91 | loss: 0.37401 | 0:03:54s
epoch 92 | loss: 0.37217 | 0:03:56s
epoch 93 | loss: 0.37285 | 0:03:59s
epoch 94 | loss: 0.37109 | 0:04:01s
epoch 95 | loss: 0.37071 | 0:04:04s
epoch 96 | loss: 0.36848 | 0:04:06s
epoch 97 | loss: 0.36445 | 0:04:09s
epoch 98 | loss: 0.36235 | 0:04:11s
epoch 99 | loss: 0.36406 | 0:04:14s
[++] Saving the model and parameters in corresponding directories
[++] Ended the training process for fold 6
```

Fold 0 has started running on 20-05-22

Fold 0 has completed sucessfully on 17:00 20-05-22

Fold 1 has started running at 15:15 21-05-22

Fold 2 has started running at 09:45 22-05-22

Fold 2 has completed sucessfully on 10:58 22-05-22

Fold 3 has started running at 18:40 22-05-22

Fold 3 has completed sucessfully on 22-05-22

Fold 4 completed sucessfully on 21:04 on 22-05-22

Fold 5 started at 18:21 on 23-05-22

Fold 5 completed sucessfully on 19:44 on 23-05-22

Fold 6 started at 12:53 on 24-05-22

Fold 6 has completed at 14:14 on 24-05-22

✓ 1h 19m 58s completed at 2:13 PM ● ✕