Analyse Discriminante Quadratique - OVA

Analyse Discriminante Quadratique (QDA) - One-Versus-All (OVA)

Théorie

L'**Analyse Discriminante Quadratique (QDA)** est une technique de classification qui, contrairement à LDA, permet aux classes d'avoir des **matrices de covariance différentes**. Cela le rend plus flexible mais peut aussi augmenter le risque de sur-apprentissage.

L'approche **One-Versus-All (OVA)** consiste à entraîner un modèle pour chaque classe, en distinguant chaque classe des autres combinées.

Hyperparamètres

Nous allons tester les hyperparamètres suivants : - **Régularisation (`reg_param`)** : contrôle la variance de la covariance estimée (valeurs entre `0` et `1`). - **Standardisation des données** : normalisation des features avant l'entraînement.

Exemple en Python

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.discriminant_analysis import QuadraticDiscriminantAnalysis
from sklearn.multiclass import OneVsRestClassifier
from sklearn.metrics import classification_report, confusion_matrix, accuracy_score
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler

# Chargement des ensembles de données
```

```python
train_data = pd.read_csv('covertype_train.csv')
val_data = pd.read_csv('covertype_val.csv')
test_data = pd.read_csv('covertype_test.csv')

# Préparation des données
X_train = train_data.drop('Cover_Type', axis=1)
y_train = train_data['Cover_Type']

X_val = val_data.drop('Cover_Type', axis=1)
y_val = val_data['Cover_Type']

X_test = test_data.drop('Cover_Type', axis=1)
y_test = test_data['Cover_Type']

# Standardisation des données
scaler = StandardScaler()
X_train = scaler.fit_transform(X_train)
X_val = scaler.transform(X_val)
X_test = scaler.transform(X_test)

# Recherche des meilleurs hyperparamètres
reg_params = np.linspace(0, 1, 10)
train_accuracies = []
val_accuracies = []

for reg_param in reg_params:
    qda_ova = OneVsRestClassifier(QuadraticDiscriminantAnalysis(reg_param=reg_param))
    qda_ova.fit(X_train, y_train)

    y_train_pred = qda_ova.predict(X_train)
    y_val_pred = qda_ova.predict(X_val)

    train_accuracies.append(accuracy_score(y_train, y_train_pred))
    val_accuracies.append(accuracy_score(y_val, y_val_pred))

# Sélection du meilleur reg_param
best_reg_param = reg_params[val_accuracies.index(max(val_accuracies))]
print(f"Meilleur reg_param QDA (OVA): {best_reg_param}")

# Affichage du graphique
plt.figure(figsize=(8, 6))
plt.plot(reg_params, train_accuracies, marker='o', linestyle='dashed', label='Train Accuracy
```

```python
plt.plot(reg_params, val_accuracies, marker='s', linestyle='dashed', label='Validation Accura
plt.xlabel("Régularisation (reg_param)")
plt.ylabel("Précision")
plt.title("Impact de la régularisation sur la performance de QDA (OVA)")
plt.legend()
plt.show()

# Modèle final avec le meilleur reg_param
qda_ova = OneVsRestClassifier(QuadraticDiscriminantAnalysis(reg_param=best_reg_param))
qda_ova.fit(X_train, y_train)
y_test_pred = qda_ova.predict(X_test)

# Affichage de la matrice de confusion
conf_matrix = confusion_matrix(y_test, y_test_pred)
print("\nMatrice de confusion (OVA) :")
print(conf_matrix)

print("\nÉvaluation sur l'ensemble de test")
print(classification_report(y_test, y_test_pred))
```

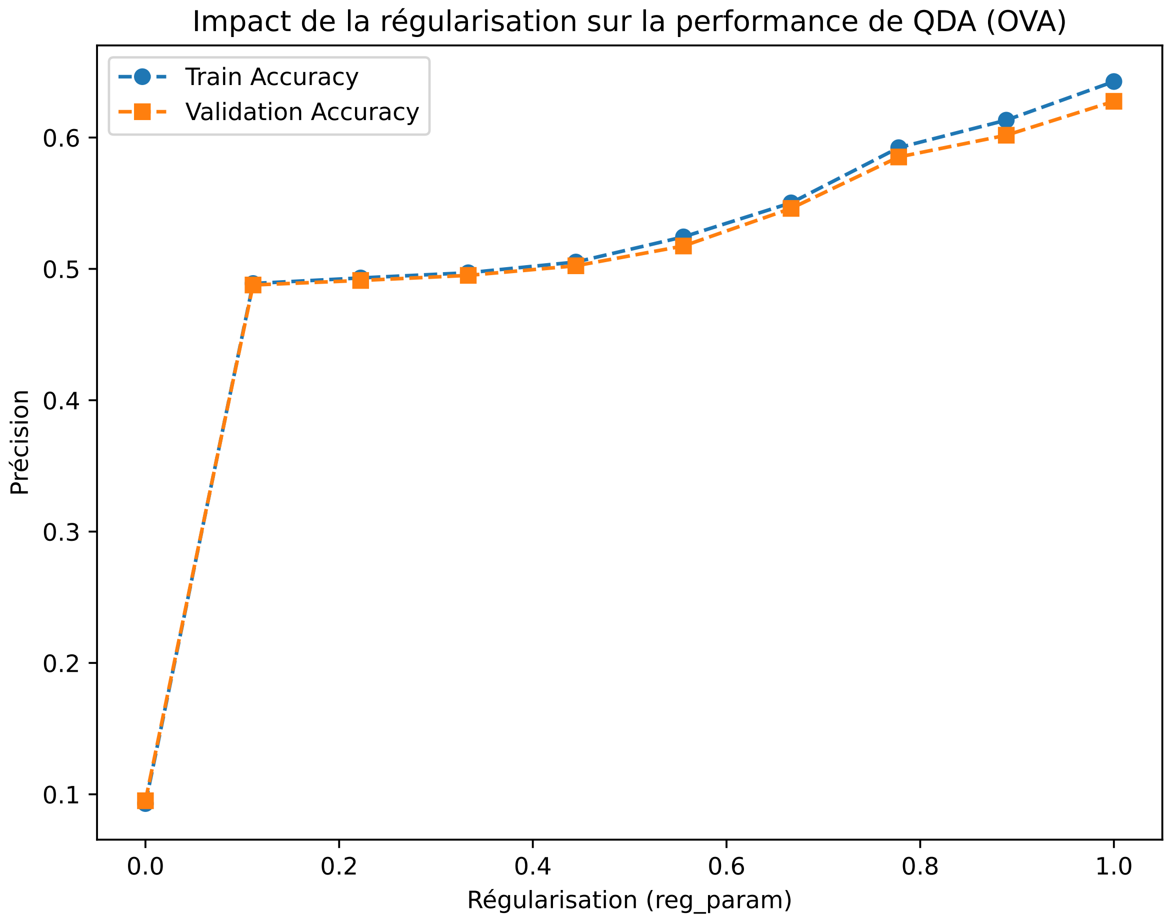
```
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
```

```
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
```

```
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
```

```
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")

Meilleur reg_param QDA (OVA): 1.0
```




```
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
```

```
Matrice de confusion (OVA) :
[[ 851  666    0    0   12    4  175]
 [ 370 1726   35   12   35   73   10]
```

```
 [   0   18  140   16   14   93    0]
 [   0    0    5   12    0    4    0]
 [   5   51   10    0    7    1    0]
 [   0   24   51    2    6   61    0]
 [  19    7    0    0    0    0  133]]

Évaluation sur l'ensemble de test
              precision    recall  f1-score   support

           1       0.68      0.50      0.58      1708
           2       0.69      0.76      0.73      2261
           3       0.58      0.50      0.54       281
           4       0.29      0.57      0.38        21
           5       0.09      0.09      0.09        74
           6       0.26      0.42      0.32       144
           7       0.42      0.84      0.56       159

    accuracy                           0.63      4648
   macro avg       0.43      0.53      0.46      4648
weighted avg       0.65      0.63      0.63      4648

/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
/home/ensai/.local/lib/python3.10/site-packages/sklearn/discriminant_analysis.py:947: UserWa
  warnings.warn("Variables are collinear")
```