

```
[13]: from scipy.io import loadmat
import pandas as pd

# Carga el archivo .mat
data = loadmat('hsefeatures.mat')

# Muestra las claves (variables) almacenadas en el archivo
print("Variables en el archivo:", data.keys())

# Explora todas las variables y sus shapes
for key in data:
    if not key.startswith('__'):
        print(f'{key}: {type(data[key])}, shape: {getattr(data[key], "shape", "N/A")}')

# Accede a la variable 'features' (ajusta el nombre si es diferente)
heart = data['heart']
print("Shape de heart:", heart.shape)

# Si quieres convertirlo a DataFrame (si es 2D)
df = pd.DataFrame(heart)
print(df.head())
```

```
In [14]: print(df[27].unique())
```

```
[0. 1.]
```

```

Meia de cada coluna:
0      1.747012e-17
1      -1.310259e-17
2      -1.921714e-16
3      -2.183766e-16
4      3.857272e-17
5      6.552978e-18
6      1.093802e-17
7      -5.241037e-16
8      5.197362e-16
9      -6.332920e-16
10     1.747012e-17
11     2.27116e-16
12     -5.241037e-17
13     -7.424803e-17
14     -2.375937e-15
15     -6.988095e-17
16     6.988095e-17
17     1.747012e-17
18     9.608569e-17
19     7.863556e-17
20     1.834636e-16
21     5.241037e-17
22     2.896415e-16
23     2.140090e-16
24     1.747012e-16
25     1.747012e-17
26     +3.615987e-16
dtype: float64

```

```

Mínimo y máximo de cada columna:
0    -11.084172
1    -11.555400
2     -1.219734
3     -0.931188
4     -8.899812
5     -7.726383
6     -0.637346
7     -7.015449
8     -0.095783
9     -6.334600
10    -2.970653
11    -1.032308
12    -0.655224
13    -1.588776
14    -3.577081
15    -5.088412
16    -4.556387
17    -3.308155
18    -4.392735
19    -3.976942
20    -5.345528
21    -5.136915
22    -6.612269
23    -5.745571
24    -4.956272
25    -8.273734
26    -5.774492
dtype: float64

```

```
[6]: from sklearn.preprocessing import MinMaxScaler

scaler = MinMaxScaler()
X_normalized = scaler.fit_transform(X)
print("Mínimo:", X_normalized.min(axis=0))
print("Máximo:", X_normalized.max(axis=0))
```

Mínimo: [0. 0.]
Máximo: [1. 1.]

```
def LogisticRegressionTraining(Xtrain, Ytrain, Xtest, Ytest, max_iter=500):
    mdl = LogisticRegression(max_iter=max_iter, solver='lbfgs')
    mdl.fit(Xtrain, Ytrain)
    Ypred = mdl.predict(Xtest)
    return accuracy_score(Ytest, Ypred)

def QDATraining(Xtrain, Ytrain, Xtest, Ytest):
    mdl = QuadraticDiscriminantAnalysis()
    mdl.fit(Xtrain, Ytrain)
    Ypred = mdl.predict(Xtest)
    return accuracy_score(Ytest, Ypred)

def KNNTraining(Xtrain, Ytrain, Xtest, Ytest, n_neighbors=5):
    mdl = KNeighborsClassifier(n_neighbors=n_neighbors)
    mdl.fit(Xtrain, Ytrain)
    Ypred = mdl.predict(Xtest)
    return accuracy_score(Ytest, Ypred)

def DecisionTreeTraining(Xtrain, Ytrain, Xtest, Ytest, max_depth=5):
    mdl = tree.DecisionTreeClassifier(max_depth=max_depth)
    mdl.fit(Xtrain, Ytrain)
    Ypred = mdl.predict(Xtest)
    return accuracy_score(Ytest, Ypred)

def GaussianNBTraining(Xtrain, Ytrain, Xtest, Ytest):
    mdl = GaussianNB()
    mdl.fit(Xtrain, Ytrain)
    Ypred = mdl.predict(Xtest)
    return accuracy_score(Ytest, Ypred)

def MLPTraining(Xtrain, Ytrain, Xtest, Ytest, hidden_layer_sizes=(90,
                                                                    10,
                                                                    max_iter=300)):
    mdl = MLPClassifier(hidden_layer_sizes=hidden_layer_sizes, max_iter=max_iter)
    mdl.fit(Xtrain, Ytrain)
    Ypred = mdl.predict(Xtest)
    return accuracy_score(Ytest, Ypred)

def BaggingKNNTraining(Xtrain, Ytrain, Xtest, Ytest, n_neighbors=5, n_estimators=10):
    mdl = BaggingClassifier(
        estimator=KNeighborsClassifier(n_neighbors=n_neighbors),
        n_estimators=n_estimators,
        random_state=42
```

```
mdl.fit(Xtrain, Ytrain)
Ypred = mdl.predict(Xtest)
return accuracy_score(Ytest, Ypred)
# ...existing code...

In [ ]: # ...existing code...
import numpy as np
from sklearn.linear_model import LogisticRegression
from sklearn.discriminant_analysis import QuadraticDiscriminantAnalysis
from sklearn.neighbors import KNeighborsClassifier
from sklearn.neural_network import MLPClassifier
from sklearn.ensemble import BaggingClassifier
from sklearn.svm import SVC
from sklearn.metrics import accuracy_score, confusion_matrix, ConfusionMatrixDisplay

def get_metrics(y_true, y_pred):
    cm = confusion_matrix(y_true, y_pred, labels=[0, 1])
    tn, fp, fn, tp = cm.ravel()
    acc = (tp + tn) / (tp + tn + fp + fn)
    tnr = tn / (tn + fp) if (tn + fp) > 0 else 0 # Especificidad
    return acc, tnr, cm

n_iter = 10 # Puedes aumentar para mayor robustez

model_results = {
    "LogisticRegression": [],
    "QDA": [],
    "KNN": [],
    "MLP": [],
    "Bagging": [],
    "SVM": []
}

model_params = {
    "LogisticRegression": [],
    "QDA": [],
    "KNN": [],
    "MLP": [],
    "Bagging": [],
    "SVM": []
}

model_cms = {
    "LogisticRegression": [],
    "QDA": [],
    "KNN": [],
    "MLP": [],
    "Bagging": [],
    "SVM": []
}

# Entrenamiento
for i in range(n_iter):
    # Logistic Regression
    max_iter = np.random.randint(100, 2001)
    mdl = LogisticRegression(max_iter=max_iter, solver='lbfgs')
    mdl.fit(Xtrain_flat, Ytrain)
    Ypred = mdl.predict(Xtest_flat)
    acc, tnr, cm = get_metrics(Ytest, Ypred)
    model_results["LogisticRegression"].append((acc, tnr))
    model_params["LogisticRegression"].append({"max_iter": max_iter})
    model_cms["LogisticRegression"].append(cm)

    # QDA
    reg_param = np.random.uniform(0, 0.2)
    mdl = QuadraticDiscriminantAnalysis(reg_param=reg_param)
    mdl.fit(Xtrain_flat, Ytrain)
    Ypred = mdl.predict(Xtest_flat)
    acc, tnr, cm = get_metrics(Ytest, Ypred)
    model_results["QDA"].append((acc, tnr))
    model_params["QDA"].append({"reg_param": reg_param})
    model_cms["QDA"].append(cm)

    # KNN
    n_neighbors = np.random.randint(1, 21)
    mdl = KNeighborsClassifier(n_neighbors=n_neighbors)
    mdl.fit(Xtrain_flat, Ytrain)
    Ypred = mdl.predict(Xtest_flat)
    acc, tnr, cm = get_metrics(Ytest, Ypred)
    model_results["KNN"].append((acc, tnr))
    model_params["KNN"].append({"n_neighbors": n_neighbors})
    model_cms["KNN"].append(cm)

    # MLP
    hidden_layer_sizes = tuple(np.random.randint(50, 200, size=np.random.randint(1, 3)))
    max_iter = np.random.randint(200, 1001)
    learning_rate_init = 10*np.random.uniform(-4, -1)
    mdl = MLPClassifier(hidden_layer_sizes=hidden_layer_sizes, max_iter=max_iter, learning_rate_init=learning_rate_init)
    mdl.fit(Xtrain_flat, Ytrain)
    Ypred = mdl.predict(Xtest_flat)
    acc, tnr, cm = get_metrics(Ytest, Ypred)
    model_results["MLP"].append((acc, tnr))
    model_params["MLP"].append({"hidden_layer_sizes": hidden_layer_sizes, "max_iter": max_iter, "learning_rate_init": learning_rate_init})
    model_cms["MLP"].append(cm)

    # Bagging
    n_neighbors = np.random.randint(1, 11)
    n_estimators = np.random.randint(5, 21)
    mdl = BaggingClassifier(
        estimator=KNeighborsClassifier(n_neighbors=n_neighbors),
        n_estimators=n_estimators,
        max_samples=0.5,
        max_features=0.5,
        random_state=42
    )
    mdl.fit(Xtrain_flat, Ytrain)
    Ypred = mdl.predict(Xtest_flat)
    acc, tnr, cm = get_metrics(Ytest, Ypred)
    model_results["Bagging"].append((acc, tnr))
    model_params["Bagging"].append({"n_neighbors": n_neighbors, "n_estimators": n_estimators})
    model_cms["Bagging"].append(cm)

    # SVM
    C = 10*np.random.uniform(-2, 2)
    gamma = 10*np.random.uniform(-3, 1)
    mdl = SVC(C=C, gamma=gamma, kernel='rbf')
    mdl.fit(Xtrain_flat, Ytrain)
    Ypred = mdl.predict(Xtest_flat)
    acc, tnr, cm = get_metrics(Ytest, Ypred)
    model_results["SVM"].append((acc, tnr))
    model_params["SVM"].append({"C": C, "gamma": gamma, "kernel": "rbf"})
    model_cms["SVM"].append(cm)

# Muestra resultados por modelo
print(f"=== Resultados por modelo ===")
best_acc_global = -1
best_tnr_global = -1
best_acc_model = None
best_tnr_model = None
best_acc_params = None
best_tnr_params = None
best_acc_cm = None
best_tnr_cm = None

for model in model_results:
    arr = np.array(model_results[model])
    acs = arr[:,0]
    tnrs = arr[:,1]
    best_acc_idx = np.argmax(acs)
    best_tnr_idx = np.argmax(tnrs)
    print(f"Modelo: {model}")
    print(f"Accuracy promedio: {acs.mean():.4f}")
    print(f"Mejor accuracy: {acs[best_acc_idx]:.4f} con params {model_params[model][best_acc_idx]}")
    print(f"Especificidad promedio (TNR): {tnrs.mean():.4f}")
    print(f"Mejor especificidad: {tnrs[best_tnr_idx]:.4f} con params {model_params[model][best_tnr_idx]}")

    # Guardar mejor global
    if acs[best_acc_idx] > best_acc_global:
        best_acc_global = acs[best_acc_idx]
        best_acc_model = model
        best_acc_params = model_params[model][best_acc_idx]
        best_acc_cm = model_cms[model][best_acc_idx]
    if tnrs[best_tnr_idx] > best_tnr_global:
        best_tnr_global = tnrs[best_tnr_idx]
        best_tnr_model = model
        best_tnr_params = model_params[model][best_tnr_idx]
        best_tnr_cm = model_cms[model][best_tnr_idx]

# Muestra matriz de confusión del mejor modelo en accuracy
print(f"=== Mejor modelo en accuracy general ===")
print(f"Modelo: {best_acc_model}")
print(f"Hiperparámetros: {best_acc_params}")
print(f"Accuracy: {best_acc_global:.4f}")
disp = ConfusionMatrixDisplay(confusion_matrix=best_acc_cm, display_labels=[0, 1])
disp.plot()

# Muestra matriz de confusión del mejor modelo en especificidad
print(f"=== Mejor modelo en especificidad general ===")
print(f"Modelo: {best_tnr_model}")
print(f"Hiperparámetros: {best_tnr_params}")
print(f"Especificidad: {best_tnr_global:.4f}")
disp = ConfusionMatrixDisplay(confusion_matrix=best_tnr_cm, display_labels=[0, 1])
disp.plot()
```

c:\Users\G_Laptop\AppData\Local\Programs\Python\Python312\Lib\site-packages\sklearn\neural_network_multilayer_perceptron.py:691: ConvergenceWarning: Stochastic Optimizer: Maximum iterations (978) reached and the optimization hasn't converged yet.

```
warnings.warn(
=== Resultados por modelo ===

Modelo: LogisticRegression
Accuracy promedio: 0.8615
Mejor accuracy: 0.8615 con params {'max_iter': 582}
Especificidad promedio (TNR): 0.6723
Mejor especificidad: 0.6723 con params {'max_iter': 582}

Modelo: QDA
Accuracy promedio: 0.8400
Mejor accuracy: 0.8472 con params {'reg_param': 0.1890161796804776}
Especificidad promedio (TNR): 0.9120
Mejor especificidad: 0.9356 con params {'reg_param': 0.015942282030095112}

Modelo: KNN
Accuracy promedio: 0.9122
Mejor accuracy: 0.9265 con params {'n_neighbors': 4}
Especificidad promedio (TNR): 0.8026
Mejor especificidad: 0.8902 con params {'n_neighbors': 4}

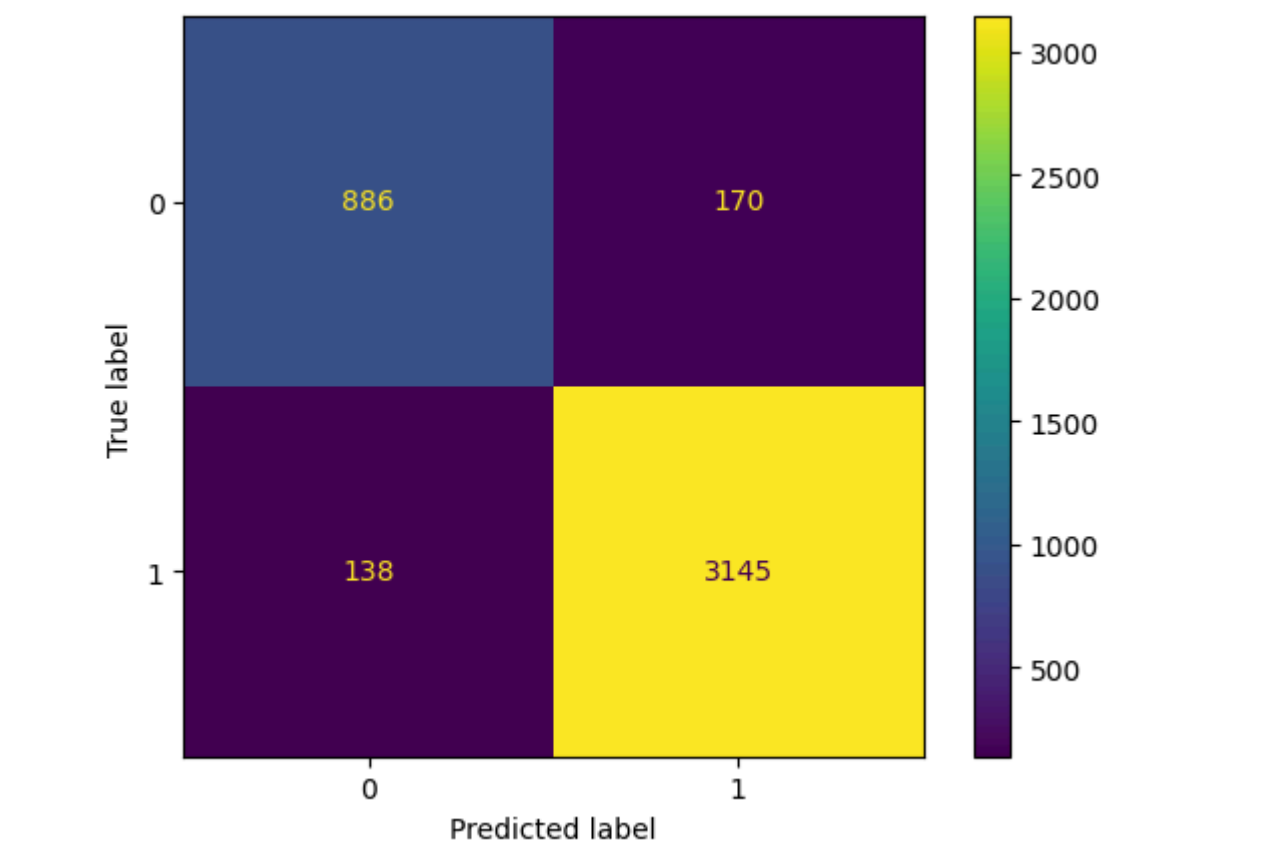
Modelo: MLP
Accuracy promedio: 0.9186
Mejor accuracy: 0.9286 con params {'hidden_layer_sizes': (130, 173), 'max_iter': 486, 'learning_rate_init': 0.0018516696315852466}
Especificidad promedio (TNR): 0.8244
Mejor especificidad: 0.8712 con params {'hidden_layer_sizes': (166, 100), 'max_iter': 850, 'learning_rate_init': 0.00826068770973839}

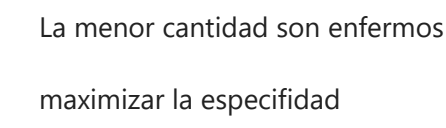
Modelo: Bagging
Accuracy promedio: 0.9130
Mejor accuracy: 0.9290 con params {'n_neighbors': 1, 'n_estimators': 16}
Especificidad promedio (TNR): 0.7734
Mejor especificidad: 0.8390 con params {'n_neighbors': 1, 'n_estimators': 16}

=== Mejor modelo en accuracy general ===
Modelo: Bagging
Hiperparámetros: {'n_neighbors': 1, 'n_estimators': 16}
Accuracy: 0.9290

=== Mejor modelo en especificidad general ===
Modelo: QDA
Hiperparámetros: {'reg_param': 0.015942282030095112}
Especificidad: 0.9356
```

Out[]: <sklearn.metrics._plot.confusion_matrix.ConfusionMatrixDisplay at 0x1296a649a30>





Cross validation

```
Out[11]: <sklearn.metrics._plot.confusion_matrix.ConfusionMatrixDisplay at 0x1296a3af3e0>
```

Modelo: LogisticRegression
Accuracy promedio: 0.8610
Especificidad promedio (TNR): 0.6447



```
=== Resultados por modelo ===
Modelo: LogisticRegression
Accuracy promedio: 0.8594
Mejor accuracy: 0.8594 con params {'max_iter': 1268}
Especificidad promedio (TNR): 0.6285
Mejor especificidad: 0.6285 con params {'max_iter': 1268}

Modelo: QDA
Accuracy promedio: 0.8298
Mejor accuracy: 0.8366 con params {'reg_param': 0.14807051562378964}
Especificidad promedio (TNR): 0.8928
Mejor especificidad: 0.9207 con params {'reg_param': 0.003945502721216321}

Modelo: KNN
Accuracy promedio: 0.9124
Mejor accuracy: 0.9447 con params {'n_neighbors': 1}
Especificidad promedio (TNR): 0.7895
Mejor especificidad: 0.9331 con params {'n_neighbors': 2}

Modelo: MLP
Accuracy promedio: 0.9129
Mejor accuracy: 0.9299 con params {'hidden_layer_sizes': (180, 120), 'max_iter': 318, 'learning_rate_init': 0.001775830862646847}
Especificidad promedio (TNR): 0.7946
Mejor especificidad: 0.8663 con params {'hidden_layer_sizes': (193, 64), 'max_iter': 782, 'learning_rate_init': 0.00202434462616948}

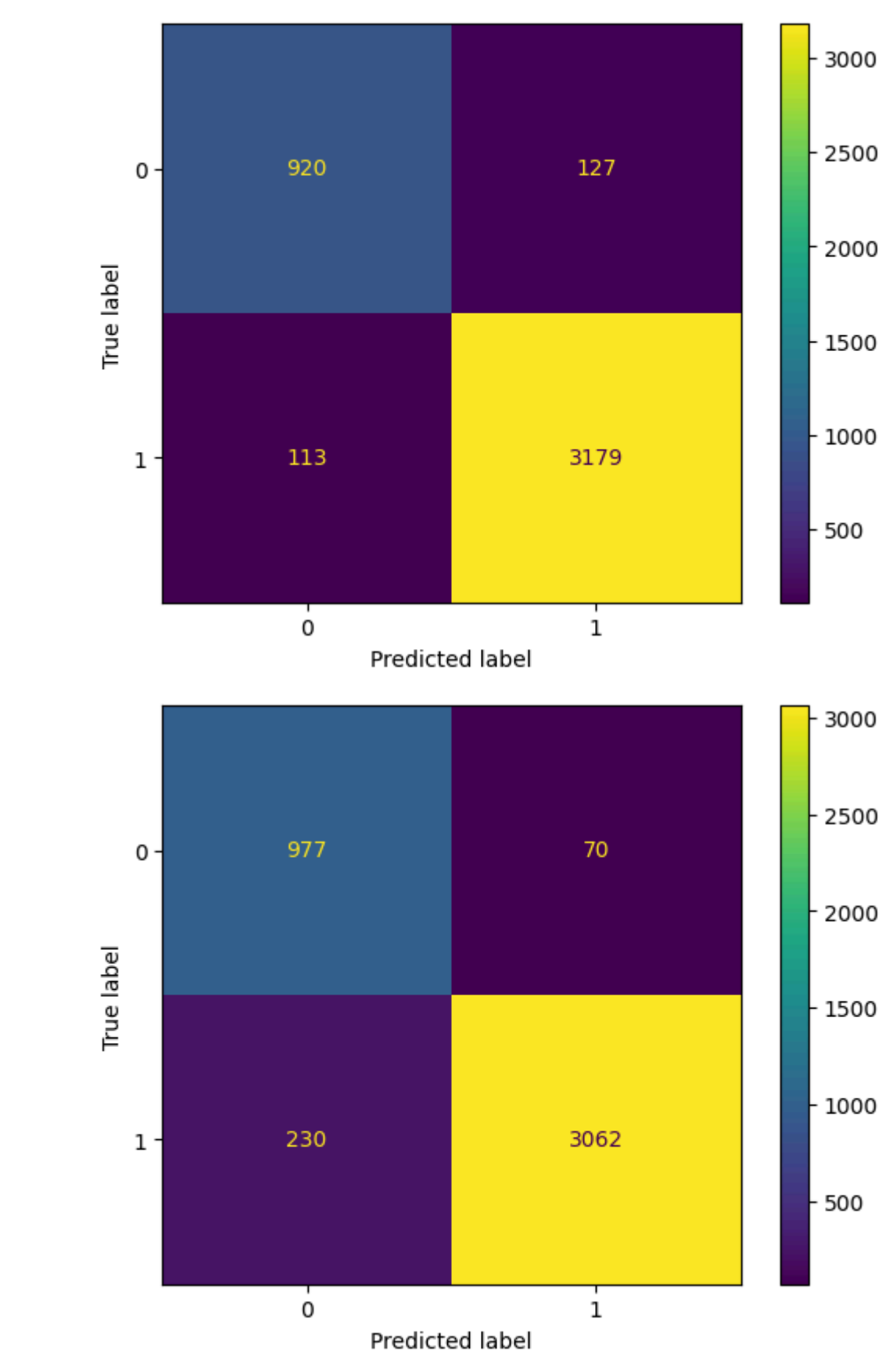
Modelo: Bagging
Accuracy promedio: 0.9091
Mejor accuracy: 0.9265 con params {'n_neighbors': 1, 'n_estimators': 12}
Especificidad promedio (TNR): 0.7423
Mejor especificidad: 0.8367 con params {'n_neighbors': 1, 'n_estimators': 6}

Modelo: SVM
Accuracy promedio: 0.8476
Mejor accuracy: 0.9336 con params {'C': 2.765666448028526, 'gamma': 0.2712132263039592, 'kernel': 'rbf'}
Especificidad promedio (TNR): 0.4544
Mejor especificidad: 0.8558 con params {'C': 45.841954312918844, 'gamma': 0.058704729132111465, 'kernel': 'rbf'}

--- Mejor modelo en accuracy general ---
Modelo: KNN
Hiperparámetros: {'n_neighbors': 1}
Accuracy: 0.9447

--- Mejor modelo en especificidad general ---
Modelo: KNN
Hiperparámetros: {'n_neighbors': 2}
Especificidad: 0.9331
```

Out[37]: <sklearn.metrics._plot.confusion_matrix.ConfusionMatrixDisplay at 0x1296ad4d730>



```
In [ ]: # Exportar solo a HTML
import os
notebook_path = os.path.abspath('Guevara-EvaluarModelosTaller.ipynb')
!jupyter nbconvert --to html --notebook_path"

# Muestra la ruta del archivo HTML generado
html_path = notebook_path.replace('.ipynb', '.html')
print(f"HTML generado en: {html_path}")
```