Creado por:

Isabel Maniega

AutoML-Pycaret

https://pycaret.gitbook.io/docs/get-started/installation

In [1]:
```
# pip install pycaret
```

In [2]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

In [3]:
```
# importamos todo de Pycaret classification

from pycaret.classification import *
```

Importamos el Titanic Dataset

In [4]:
```
# https://www.kaggle.com/competitions/titanic

# Obtenemos los distintos dataset correspondientes a:

# train.csv,
# test.csv y
# gender_submission.csv
```

In [5]:
```
df = pd.read_csv("./data/train.csv")
df.head()
```

Out[5]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	NaN	S
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C85	C
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	NaN	S
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C123	S
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	NaN	S

setup()

In [6]:
```
# En este caso concreto, haríamos el drop()
# las columnas que no queremos las ignoramos con ignore_features
# la data es el propio dataframe obtenido con train.csv
# la columna (y) es "Survived"
# Si estamos de acuerdo con los tipos de datos, click en Enter.
# (Cuando  ejecutamos la celda, debemos esperar y después hacer click en Enter...)
# Pclass podría ser "ordinal", pero lo ponemos como Categorical
# SibSp y Parch con categóricas con un limitado numero de posibles valores
# pudimos considerar (quizá) como numéricas, no obstante.

clf = setup(data=df, target="Survived", ignore_features=["Name", "Ticket", "PassengerId"])
```

	Description	Value
0	session_id	4597
1	Target	Survived
2	Target Type	Binary
3	Label Encoded	None
4	Original Data	(891, 12)
5	Missing Values	True
6	Numeric Features	2
7	Categorical Features	6
8	Ordinal Features	False
9	High Cardinality Features	False
10	High Cardinality Method	None
11	Transformed Train Set	(623, 137)
12	Transformed Test Set	(268, 137)
13	Shuffle Train-Test	True
14	Stratify Train-Test	False
15	Fold Generator	StratifiedKFold
16	Fold Number	10
17	CPU Jobs	-1
18	Use GPU	False
19	Log Experiment	False
20	Experiment Name	clf-default-name
21	USI	3152
22	Imputation Type	simple
23	Iterative Imputation Iteration	None
24	Numeric Imputer	mean
25	Iterative Imputation Numeric Model	None
26	Categorical Imputer	constant
27	Iterative Imputation Categorical Model	None
28	Unknown Categoricals Handling	least_frequent
29	Normalize	False
30	Normalize Method	None
31	Transformation	False
32	Transformation Method	None
33	PCA	False
34	PCA Method	None
35	PCA Components	None
36	Ignore Low Variance	False
37	Combine Rare Levels	False
38	Rare Level Threshold	None
39	Numeric Binning	False
40	Remove Outliers	False
41	Outliers Threshold	None
42	Remove Multicollinearity	False
43	Multicollinearity Threshold	None
44	Remove Perfect Collinearity	True
45	Clustering	False
46	Clustering Iteration	None
47	Polynomial Features	False
48	Polynomial Degree	None
49	Trignometry Features	False
50	Polynomial Threshold	None
51	Group Features	False
52	Feature Selection	False
53	Feature Selection Method	classic
54	Features Selection Threshold	None
55	Feature Interaction	False
56	Feature Ratio	False
57	Interaction Threshold	None
58	Fix Imbalance	False
59	Fix Imbalance Method	SMOTE

compare_models()

In [7]:
```
# K-fold cross validation: revisar que es?¿
compare_models()
```

	Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC	TT (Sec)
gbc	Gradient Boosting Classifier	0.8220	0.8648	0.6826	0.8201	0.7421	0.6085	0.6165	0.0160
lr	Logistic Regression	0.8140	0.8556	0.7246	0.7748	0.7469	0.6002	0.6029	0.2360
et	Extra Trees Classifier	0.8092	0.8477	0.7159	0.7681	0.7378	0.5886	0.5924	0.0340
lightgbm	Light Gradient Boosting Machine	0.8077	0.8641	0.7118	0.7608	0.7334	0.5838	0.5862	0.0410
lda	Linear Discriminant Analysis	0.8042	0.8459	0.6777	0.7811	0.7239	0.5734	0.5784	0.0050
ridge	Ridge Classifier	0.7995	0.0000	0.6821	0.7699	0.7203	0.5650	0.5701	0.0020
ada	Ada Boost Classifier	0.7979	0.8264	0.7072	0.7480	0.7247	0.5655	0.5683	0.0130
rf	Random Forest Classifier	0.7948	0.8608	0.7029	0.7443	0.7195	0.5586	0.5619	0.0360
dt	Decision Tree Classifier	0.7706	0.7564	0.7156	0.6939	0.7012	0.5157	0.5189	0.0040
knn	K Neighbors Classifier	0.7179	0.7574	0.5386	0.6614	0.5915	0.3795	0.3858	0.1570
nb	Naive Bayes	0.6565	0.7883	0.1319	0.8452	0.2218	0.1295	0.2262	0.0020
svm	SVM - Linear Kernel	0.6212	0.0000	0.5415	0.5431	0.4636	0.2082	0.2406	0.0020
dummy	Dummy Classifier	0.6212	0.5000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0020
qda	Quadratic Discriminant Analysis	0.4272	0.5146	0.8728	0.4048	0.5413	0.0339	0.0110	0.0040

Out[7]:
```
GradientBoostingClassifier(ccp_alpha=0.0, criterion='friedman_mse', init=None,
                           learning_rate=0.1, loss='deviance', max_depth=3,
                           max_features=None, max_leaf_nodes=None,
                           min_impurity_decrease=0.0, min_impurity_split=None,
                           min_samples_leaf=1, min_samples_split=2,
                           min_weight_fraction_leaf=0.0, n_estimators=100,
                           n_iter_no_change=None, presort='deprecated',
                           random_state=4597, subsample=1.0, tol=0.0001,
                           validation_fraction=0.1, verbose=0,
                           warm_start=False)
```

create_model()

Para los 2 o 3 mejores:

- En mi caso el mejor modelo es GradientBoostingClassifier
- Lo parámetros que muestra son ACCURACY y AUC, dando los mejores porcentajes.

In [13]:
```
gbc = create_model("gbc")
```

Fold	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC
0	0.8413	0.8568	0.7500	0.8182	0.7826	0.6580	0.6595
1	0.8254	0.8007	0.6250	0.8824	0.7317	0.6078	0.6277
2	0.6984	0.7393	0.4583	0.6471	0.5366	0.3226	0.3331
3	0.8387	0.9167	0.7500	0.8182	0.7826	0.6548	0.6564
4	0.8387	0.9375	0.7083	0.8500	0.7727	0.6493	0.6558
5	0.7903	0.8788	0.7083	0.7391	0.7234	0.5547	0.5550
6	0.8710	0.8785	0.7826	0.8571	0.8182	0.7185	0.7203
7	0.8710	0.8690	0.6957	0.9412	0.8000	0.7079	0.7255
8	0.8226	0.9119	0.6087	0.8750	0.7179	0.5945	0.6154
9	0.8226	0.8584	0.7391	0.7727	0.7556	0.6164	0.6168
Mean	0.8220	0.8648	0.6826	0.8201	0.7421	0.6085	0.6165
Std	0.0469	0.0553	0.0910	0.0791	0.0754	0.1064	0.1058

In [11]:
```
lr = create_model("lr")
```

Fold	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC
0	0.7937	0.8579	0.7500	0.7200	0.7347	0.5660	0.5663
1	0.7937	0.7746	0.6667	0.7619	0.7111	0.5517	0.5547
2	0.7302	0.7495	0.5833	0.6667	0.6222	0.4138	0.4160
3	0.8548	0.9254	0.7917	0.8261	0.8085	0.6917	0.6921
4	0.8710	0.8849	0.8333	0.8333	0.8333	0.7281	0.7281
5	0.7903	0.8871	0.7083	0.7391	0.7234	0.5547	0.5550
6	0.8548	0.8796	0.7826	0.8182	0.8000	0.6862	0.6866
7	0.8710	0.8874	0.6957	0.9412	0.8000	0.7079	0.7255
8	0.8065	0.8841	0.6957	0.7619	0.7273	0.5778	0.5792
9	0.7742	0.8255	0.7391	0.6800	0.7083	0.5246	0.5259
Mean	0.8140	0.8556	0.7246	0.7748	0.7469	0.6002	0.6029
Std	0.0446	0.0528	0.0677	0.0780	0.0603	0.0950	0.0965

In [12]:
```
et = create_model("et")
```

Fold	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC
0	0.8730	0.8787	0.8333	0.8333	0.8333	0.7308	0.7308
1	0.7143	0.7473	0.6667	0.6154	0.6400	0.4038	0.4047
2	0.7302	0.7185	0.5000	0.7059	0.5854	0.3939	0.4068
3	0.9032	0.9282	0.8750	0.8750	0.8750	0.7961	0.7961
4	0.8065	0.8986	0.7500	0.7500	0.7500	0.5921	0.5921
5	0.8548	0.8739	0.7083	0.8947	0.7907	0.6819	0.6928
6	0.8387	0.8618	0.6957	0.8421	0.7619	0.6416	0.6483
7	0.8226	0.8807	0.7391	0.7727	0.7556	0.6164	0.6168
8	0.7581	0.8735	0.6087	0.7000	0.6512	0.4674	0.4700
9	0.7903	0.8161	0.7826	0.6923	0.7347	0.5624	0.5653
Mean	0.8092	0.8477	0.7159	0.7681	0.7378	0.5886	0.5924
Std	0.0586	0.0636	0.1031	0.0867	0.0849	0.1277	0.1264

tune_model() para los 2 o 3 mejores

In [14]:
```
tune_gbc_auc = tune_model(gbc, optimize="AUC")
```

Fold	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC
0	0.7778	0.8429	0.7500	0.6923	0.7200	0.5363	0.5375
1	0.7460	0.7548	0.5833	0.7000	0.6364	0.4437	0.4481
2	0.7778	0.7559	0.6250	0.7500	0.6818	0.5132	0.5183
3	0.8710	0.9178	0.7500	0.9000	0.8182	0.7195	0.7266
4	0.8548	0.9024	0.7500	0.8571	0.8000	0.6869	0.6906
5	0.8710	0.8931	0.7500	0.9000	0.8182	0.7195	0.7266
6	0.8710	0.8623	0.7826	0.8571	0.8182	0.7185	0.7203
7	0.9032	0.9013	0.7391	1.0000	0.8500	0.7809	0.8004
8	0.8065	0.9086	0.5652	0.8667	0.6842	0.5534	0.5797
9	0.7742	0.8434	0.6957	0.6957	0.6957	0.5162	0.5162
Mean	0.8253	0.8583	0.6991	0.8219	0.7523	0.6188	0.6264
Std	0.0519	0.0572	0.0747	0.1006	0.0722	0.1116	0.1136

In [17]:
```
tune_gbc_acc = tune_model(gbc, optimize="Accuracy")
```

Fold	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC
0	0.7778	0.8429	0.7500	0.6923	0.7200	0.5363	0.5375
1	0.7460	0.7548	0.5833	0.7000	0.6364	0.4437	0.4481
2	0.7778	0.7559	0.6250	0.7500	0.6818	0.5132	0.5183
3	0.8710	0.9178	0.7500	0.9000	0.8182	0.7195	0.7266
4	0.8548	0.9024	0.7500	0.8571	0.8000	0.6869	0.6906
5	0.8710	0.8931	0.7500	0.9000	0.8182	0.7195	0.7266
6	0.8710	0.8623	0.7826	0.8571	0.8182	0.7185	0.7203
7	0.9032	0.9013	0.7391	1.0000	0.8500	0.7809	0.8004
8	0.8065	0.9086	0.5652	0.8667	0.6842	0.5534	0.5797
9	0.7742	0.8434	0.6957	0.6957	0.6957	0.5162	0.5162
Mean	0.8253	0.8583	0.6991	0.8219	0.7523	0.6188	0.6264
Std	0.0519	0.0572	0.0747	0.1006	0.0722	0.1116	0.1136

In [15]:
```
tune_lr_auc = tune_model(lr, optimize="AUC")
```

Fold	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC
0	0.7778	0.8536	0.7917	0.6786	0.7308	0.5435	0.5482
1	0.7619	0.7767	0.7083	0.6800	0.6939	0.4992	0.4995
2	0.7143	0.7527	0.6667	0.6154	0.6400	0.4038	0.4047
3	0.8548	0.9221	0.8750	0.7778	0.8235	0.7010	0.7045
4	0.8548	0.8816	0.8750	0.7778	0.8235	0.7010	0.7045
5	0.7903	0.8882	0.7500	0.7200	0.7347	0.5615	0.5618
6	0.8387	0.8807	0.7826	0.7826	0.7826	0.6544	0.6544
7	0.8871	0.8907	0.8261	0.8636	0.8444	0.7559	0.7564
8	0.8387	0.8941	0.7826	0.7826	0.7826	0.6544	0.6544
9	0.7258	0.8333	0.7391	0.6071	0.6667	0.4376	0.4437
Mean	0.8044	0.8574	0.7797	0.7286	0.7523	0.5912	0.5932
Std	0.0559	0.0518	0.0639	0.0784	0.0668	0.1138	0.1134

In [18]:
```
tune_lr_acc = tune_model(lr, optimize="Accuracy")
```

Fold	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC
0	0.7937	0.8600	0.7500	0.7200	0.7347	0.5660	0.5663
1	0.8095	0.7628	0.7083	0.7727	0.7391	0.5896	0.5910
2	0.7460	0.7249	0.6250	0.6818	0.6522	0.4528	0.4538
3	0.8710	0.9232	0.8333	0.8333	0.8333	0.7281	0.7281
4	0.8548	0.8772	0.8333	0.8000	0.8163	0.6964	0.6968
5	0.7903	0.8925	0.7083	0.7391	0.7234	0.5547	0.5550
6	0.8548	0.8807	0.7826	0.8182	0.8000	0.6862	0.6866
7	0.8548	0.8584	0.6522	0.9375	0.7692	0.6683	0.6917
8	0.8226	0.8718	0.7391	0.7727	0.7556	0.6164	0.6168
9	0.7903	0.8166	0.7391	0.7083	0.7234	0.5547	0.5550
Mean	0.8188	0.8468	0.7371	0.7784	0.7547	0.6113	0.6141
Std	0.0378	0.0581	0.0650	0.0705	0.0504	0.0799	0.0816

In [16]:
```
tune_et_auc = tune_model(et, optimize="AUC")
```

Fold	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC
0	0.7460	0.7906	0.7917	0.6333	0.7037	0.4862	0.4955
1	0.7460	0.7735	0.7083	0.6538	0.6800	0.4700	0.4711
2	0.6667	0.7623	0.7500	0.5455	0.6316	0.3408	0.3553
3	0.8065	0.9128	0.7917	0.7308	0.7600	0.5983	0.5996
4	0.8548	0.8925	0.8750	0.7778	0.8235	0.7010	0.7045
5	0.7903	0.8827	0.8750	0.6774	0.7636	0.5806	0.5960
6	0.7742	0.8551	0.7826	0.6667	0.7200	0.5328	0.5376
7	0.8710	0.8963	0.8261	0.8261	0.8261	0.7235	0.7235
8	0.8065	0.9119	0.7826	0.7200	0.7500	0.5926	0.5939
9	0.7419	0.8456	0.7826	0.6207	0.6923	0.4751	0.4846
Mean	0.7804	0.8523	0.7966	0.6852	0.7351	0.5501	0.5562
Std	0.0566	0.0547	0.0487	0.0771	0.0588	0.1087	0.1057

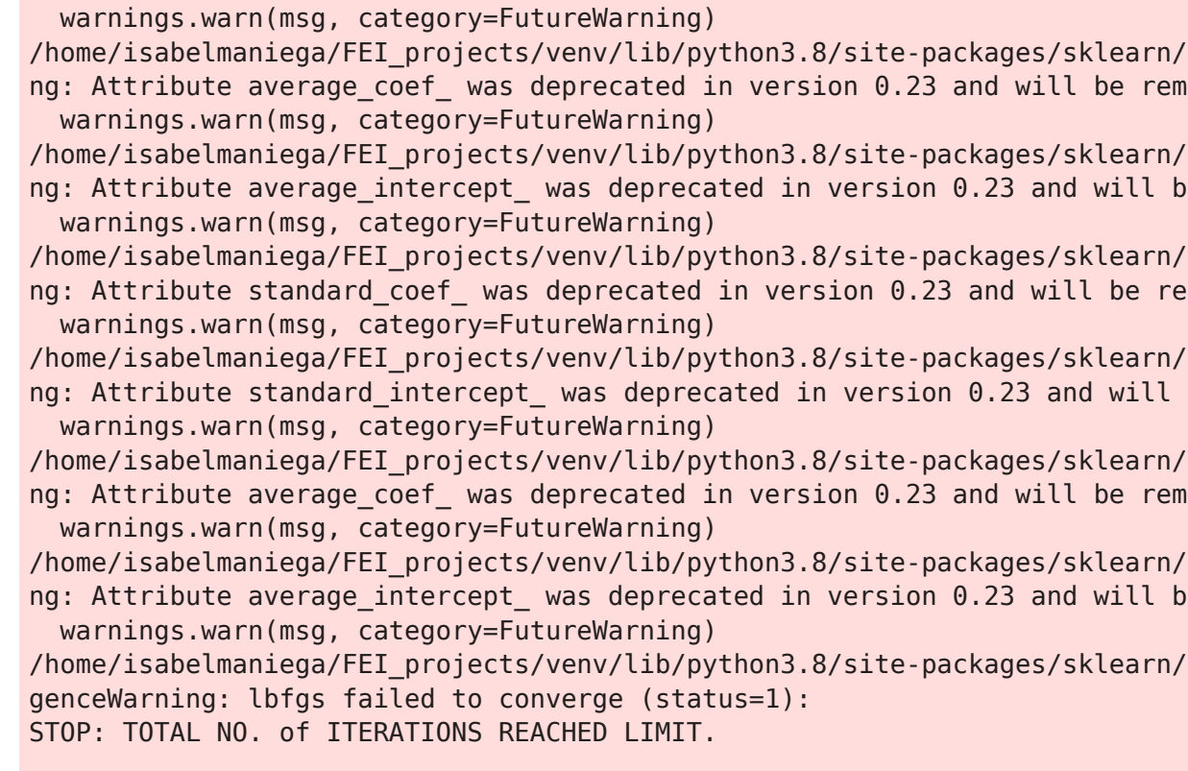
In [19]:
```
tune_et_acc = tune_model(et, optimize="Accuracy")
```

	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC
Fold							
0	0.7460	0.7906	0.7917	0.6333	0.7037	0.4862	0.4955
1	0.6667	0.7532	0.7500	0.5455	0.6316	0.3408	0.3553
2	0.6984	0.7644	0.6667	0.5926	0.6275	0.3756	0.3774
3	0.8226	0.9073	0.8750	0.7241	0.7925	0.6399	0.6487
4	0.8710	0.9030	0.8750	0.8077	0.8400	0.7322	0.7338
5	0.8226	0.8876	0.8750	0.7241	0.7925	0.6399	0.6487
6	0.8065	0.8361	0.8696	0.6897	0.7692	0.6063	0.6184
7	0.8387	0.8718	0.7826	0.7826	0.7826	0.6544	0.6544
8	0.8065	0.8969	0.7826	0.7200	0.7500	0.5926	0.5939
9	0.7258	0.8456	0.7391	0.6071	0.6667	0.4376	0.4437
Mean	0.7805	0.8457	0.8007	0.6827	0.7356	0.5506	0.5570
Std	0.0635	0.0552	0.0681	0.0810	0.0701	0.1250	0.1234

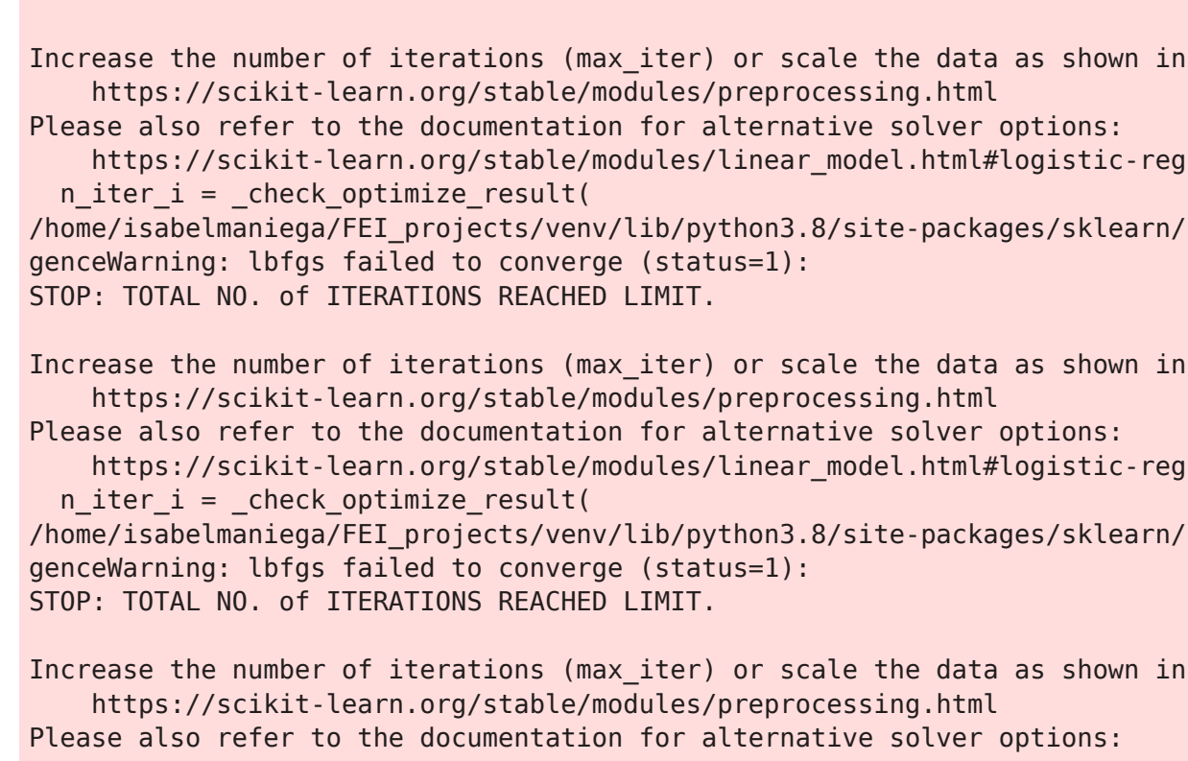
```
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/linear_model/_logistic.py:762: Conver
genceWarning: lbfgs failed to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
  n_iter_i = _check_optimize_result(
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/linear_model/_logistic.py:762: Conver
genceWarning: lbfgs failed to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
  n_iter_i = _check_optimize_result(
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/linear_model/_logistic.py:762: Conver
genceWarning: lbfgs failed to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
  n_iter_i = _check_optimize_result(
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/linear_model/_logistic.py:762: Conver
genceWarning: lbfgs failed to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
  n_iter_i = _check_optimize_result(
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute standard_coef_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute standard_intercept_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute average_coef_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute average_intercept_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/linear_model/_logistic.py:762: Conver
genceWarning: lbfgs failed to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
  n_iter_i = _check_optimize_result(
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/linear_model/_logistic.py:762: Conver
genceWarning: lbfgs failed to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
  n_iter_i = _check_optimize_result(
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute standard_coef_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute standard_intercept_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute average_coef_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute average_intercept_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/linear_model/_logistic.py:762: Conver
genceWarning: lbfgs failed to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
  n_iter_i = _check_optimize_result(
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute standard_coef_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute standard_intercept_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute average_coef_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute average_intercept_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/linear_model/_logistic.py:762: Conver
genceWarning: lbfgs failed to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
  n_iter_i = _check_optimize_result(
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/linear_model/_logistic.py:762: Conver
genceWarning: lbfgs failed to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
  n_iter_i = _check_optimize_result(
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/linear_model/_logistic.py:762: Conver
genceWarning: lbfgs failed to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
  n_iter_i = _check_optimize_result(
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute standard_coef_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute standard_intercept_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute average_coef_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute average_intercept_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/linear_model/_logistic.py:762: Conver
genceWarning: lbfgs failed to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
  n_iter_i = _check_optimize_result(
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/linear_model/_logistic.py:762: Conver
genceWarning: lbfgs failed to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
  n_iter_i = _check_optimize_result(
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute standard_coef_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute standard_intercept_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute average_coef_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute average_intercept_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/linear_model/_logistic.py:762: Conver
genceWarning: lbfgs failed to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
  n_iter_i = _check_optimize_result(
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute standard_coef_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute standard_intercept_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute average_coef_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute average_intercept_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/linear_model/_logistic.py:762: Conver
genceWarning: lbfgs failed to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
  n_iter_i = _check_optimize_result(
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/linear_model/_logistic.py:762: Conver
genceWarning: lbfgs failed to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
  n_iter_i = _check_optimize_result(
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/linear_model/_logistic.py:762: Conver
genceWarning: lbfgs failed to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
  n_iter_i = _check_optimize_result(
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/linear_model/_logistic.py:762: Conver
genceWarning: lbfgs failed to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
  n_iter_i = _check_optimize_result(
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/linear_model/_logistic.py:762: Conver
genceWarning: lbfgs failed to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
  n_iter_i = _check_optimize_result(
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute standard_coef_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute standard_intercept_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute average_coef_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute average_intercept_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/metrics/_classification.py:1221: Unde
finedMetricWarning: Precision is ill-defined and being set to 0.0 due to no predicted samples. Use `zero_divisi
on` parameter to control this behavior.
  _warn_prf(average, modifier, msg_start, len(result))
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/metrics/_classification.py:846: Runti
meWarning: invalid value encountered in double_scalars
  mcc = cov_ytyp / np.sqrt(cov_ytyt * cov_ypyp)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/linear_model/_logistic.py:762: Conver
genceWarning: lbfgs failed to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
  n_iter_i = _check_optimize_result(
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/linear_model/_logistic.py:762: Conver
genceWarning: lbfgs failed to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
  n_iter_i = _check_optimize_result(
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute standard_coef_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute standard_intercept_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute average_coef_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute average_intercept_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/linear_model/_logistic.py:762: Conver
genceWarning: lbfgs failed to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
  n_iter_i = _check_optimize_result(
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute standard_coef_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute standard_intercept_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute average_coef_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute average_intercept_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute standard_coef_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute standard_intercept_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute average_coef_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute average_intercept_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/linear_model/_logistic.py:762: Conver
genceWarning: lbfgs failed to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
  n_iter_i = _check_optimize_result(
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/linear_model/_logistic.py:762: Conver
genceWarning: lbfgs failed to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
  n_iter_i = _check_optimize_result(
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/linear_model/_logistic.py:762: Conver
genceWarning: lbfgs failed to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
  n_iter_i = _check_optimize_result(
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/linear_model/_logistic.py:762: Conver
genceWarning: lbfgs failed to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
  n_iter_i = _check_optimize_result(
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/linear_model/_logistic.py:762: Conver
genceWarning: lbfgs failed to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
  n_iter_i = _check_optimize_result(
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute standard_coef_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute standard_intercept_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute average_coef_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/utils/deprecation.py:101: FutureWarni
ng: Attribute average_intercept_ was deprecated in version 0.23 and will be removed in 0.25.
  warnings.warn(msg, category=FutureWarning)
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/linear_model/_logistic.py:762: Conver
genceWarning: lbfgs failed to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
  n_iter_i = _check_optimize_result(
/home/isabelmaniega/FEI_projects/venv/lib/python3.8/site-packages/sklearn/linear_model/_logistic.py:762: Conver
genceWarning: lbfgs failed to converge (status=1):
STOP: TOTAL NO. of ITERATIONS REACHED LIMIT.

Increase the number of iterations (max_iter) or scale the data as shown in:
    https://scikit-learn.org/stable/modules/preprocessing.html
Please also refer to the documentation for alternative solver options:
    https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression
  n_iter_i = _check_optimize_result(
```

plot_model() para el mejor modelo

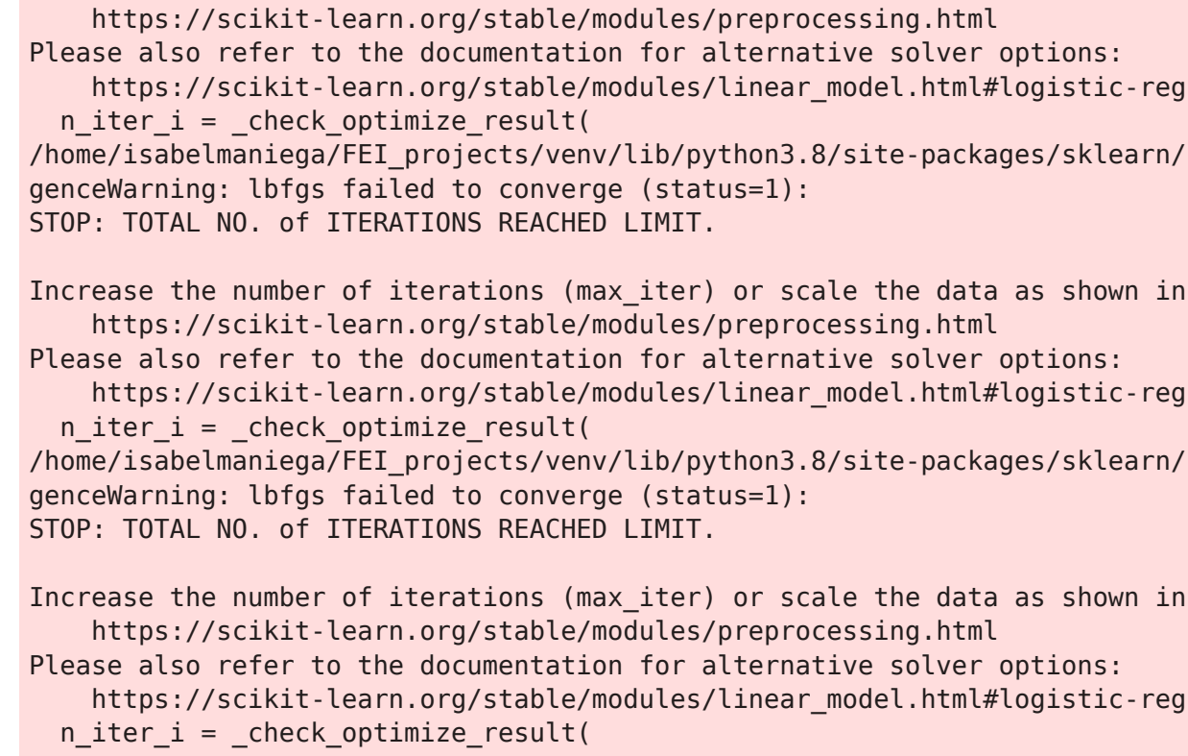
In [20]:
```
# tune_gbc_acc consigue:
# 0.8583 (Accuracy) - 0.8583 (AUC)
plot_model(tune_gbc_acc, plot="error")
```



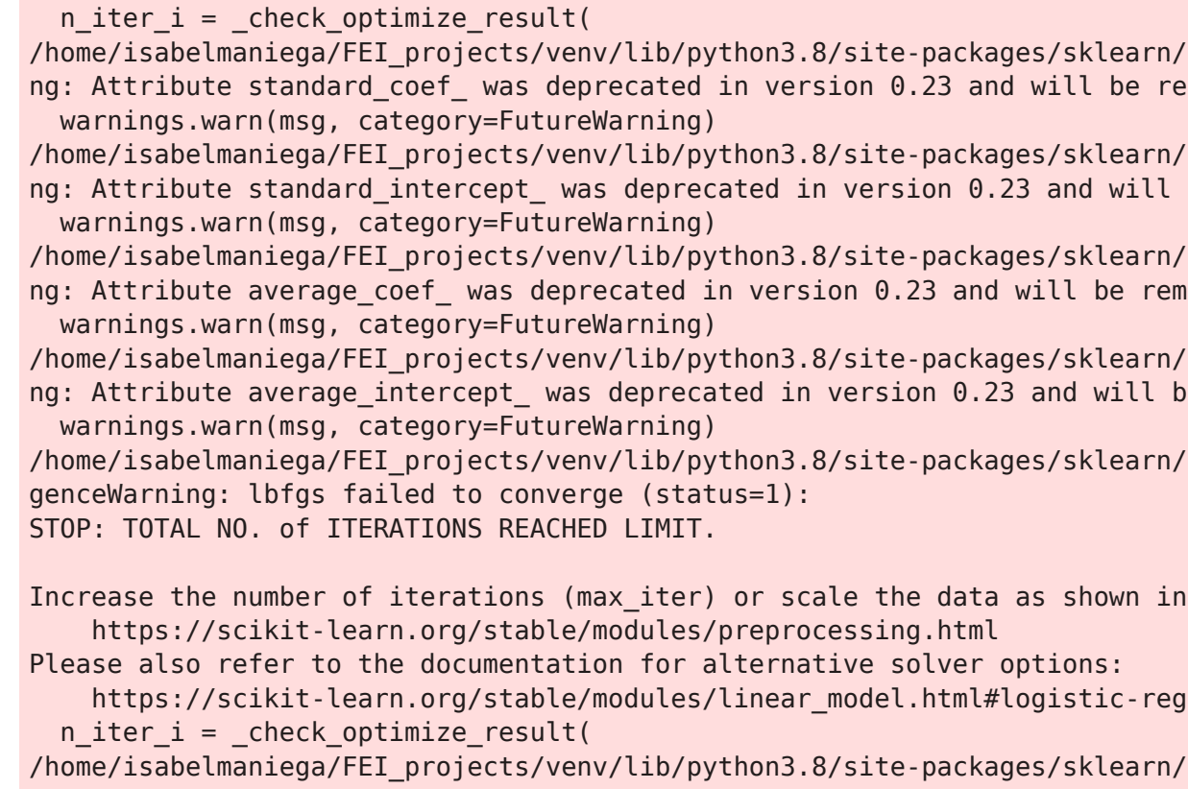
In [21]:
```
plot_model(tune_gbc_acc, plot="pr")
```



In [22]:
```
plot_model(tune_gbc_acc, plot="feature")
```



In [23]:
```
plot_model(tune_gbc_acc, plot="confusion_matrix")
```



evaluate_model() para el mejor

In [24]:
```
evaluate_model(tune_gbc_acc)
```

	Parameters
ccp_alpha	0.0
criterion	friedman_mse
init	None
learning_rate	0.05
loss	deviance
max_depth	2
max_features	log2
max_leaf_nodes	None
min_impurity_decrease	0.0002
min_impurity_split	None
min_samples_leaf	3
min_samples_split	9
min_weight_fraction_leaf	0.0
n_estimators	280
n_iter_no_change	None
presort	deprecated
random_state	4597
subsample	0.25
tol	0.0001
validation_fraction	0.1
verbose	0
warm_start	False

```
interactive(children=(ToggleButtons(description='Plot Type:', icons=('',), options=(('Hyperparameters', 'param…
```

nos vamos a test.csv

In [25]:
```
test = pd.read_csv("./data/test.csv")
test
```

Out[25]:

	PassengerId	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	892	3	Kelly, Mr. James	male	34.5	0	0	330911	7.8292	NaN	Q
1	893	3	Wilkes, Mrs. James (Ellen Needs)	female	47.0	1	0	363272	7.0000	NaN	S
2	894	2	Myles, Mr. Thomas Francis	male	62.0	0	0	240276	9.6875	NaN	Q
3	895	3	Wirz, Mr. Albert	male	27.0	0	0	315154	8.6625	NaN	S
4	896	3	Hirvonen, Mrs. Alexander (Helga E Lindqvist)	female	22.0	1	1	3101298	12.2875	NaN	S
...	...	...	...	...	...	...	...	...	...	...	...
413	1305	3	Spector, Mr. Woolf	male	NaN	0	0	A.5. 3236	8.0500	NaN	S
414	1306	1	Oliva y Ocana, Dona. Fermina	female	39.0	0	0	PC 17758	108.9000	C105	C
415	1307	3	Saether, Mr. Simon Sivertsen	male	38.5	0	0	SOTON/O.Q. 3101262	7.2500	NaN	S
416	1308	3	Ware, Mr. Frederick	male	NaN	0	0	359309	8.0500	NaN	S
417	1309	3	Peter, Master. Michael J	male	NaN	1	1	2668	22.3583	NaN	C

418 rows × 11 columns

Realizamos la predicción

In [26]:
```
predicciones = predict_model(tune_gbc_acc, data=test)
predicciones
```

Out[26]:

	PassengerId	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked	Label	Score
0	892	3	Kelly, Mr. James	male	34.5	0	0	330911	7.8292	NaN	Q	0	0.7977
1	893	3	Wilkes, Mrs. James (Ellen Needs)	female	47.0	1	0	363272	7.0000	NaN	S	0	0.6635
2	894	2	Myles, Mr. Thomas Francis	male	62.0	0	0	240276	9.6875	NaN	Q	0	0.7774
3	895	3	Wirz, Mr. Albert	male	27.0	0	0	315154	8.6625	NaN	S	0	0.8759
4	896	3	Hirvonen, Mrs. Alexander (Helga E Lindqvist)	female	22.0	1	1	3101298	12.2875	NaN	S	1	0.5369
...	...	...	...	...	...	...	...	...	...	...	...	...	...
413	1305	3	Spector, Mr. Woolf	male	NaN	0	0	A.5. 3236	8.0500	NaN	S	0	0.8759
414	1306	1	Oliva y Ocana, Dona. Fermina	female	39.0	0	0	PC 17758	108.9000	C105	C	1	0.8663
415	1307	3	Saether, Mr. Simon Sivertsen	male	38.5	0	0	SOTON/O.Q. 3101262	7.2500	NaN	S	0	0.9040
416	1308	3	Ware, Mr. Frederick	male	NaN	0	0	359309	8.0500	NaN	S	0	0.8759
417	1309	3	Peter, Master. Michael J	male	NaN	1	1	2668	22.3583	NaN	C	0	0.6552

418 rows × 13 columns

In [27]:
```
predicciones_submission = predicciones[["PassengerId", "Label"]]
predicciones_submission
```

Out[27]:

	PassengerId	Label
0	892	0
1	893	0
2	894	0
3	895	0
4	896	1
...	...	...
413	1305	0
414	1306	1
415	1307	0
416	1308	0
417	1309	0

418 rows × 2 columns

Submissions

In [29]:
```
submission = pd.read_csv("./data/gender_submission.csv")
submission
```

Out[29]:

	PassengerId	Survived
0	892	0
1	893	1
2	894	0
3	895	0
4	896	1
...	...	...
413	1305	0
414	1306	1
415	1307	0
416	1308	0
417	1309	0

418 rows × 2 columns

In [30]:
```
submission.PassengerId = predicciones_submission["PassengerId"]
submission.Survived = predicciones_submission["Label"]
submission.head()
```

Out[30]:

	PassengerId	Survived
0	892	0
1	893	0
2	894	0
3	895	0
4	896	1

Creo el archivo csv para enviar al Kaggle

In [31]:
```
# guardamos el archivo en la ruta donde podamos tenerlo localizado.
submission.to_csv("./data/gender_submission_pycaret.csv", index=False)
```

In [32]:
```
# Sólo queda mejorar las predicciones....
```

Creado por:

Isabel Maniega