**Práctico Aprendizaje Supervisado**

**Olariaga Sandra, Pesce Cristian**

**Cuadro comparativo**

|  |  |  |  |
| --- | --- | --- | --- |
| **Modelo** | **Mejor Modelo y Mejores Parametros / Estimador** | **roc\_auc\_score** | **f1\_score** |
| **Random Forest con GridSearchCV**  GridSearchCV(cv=3, estimator=RandomForestClassifier(), n\_jobs=-1,  param\_grid={'class\_weight': ['balanced', 'balanced\_subsample'],  'criterion': ['gini'], 'min\_samples\_leaf': [1],  'min\_samples\_split': [2, 3],  'n\_estimators': [100, 150], 'random\_state': [0]},  scoring='accuracy', verbose=4)  Best Random Forest accuracy: 0.7299394158603852 | **Mejor Modelo**  RandomForestClassifier(class\_weight='balanced', n\_estimators=150,  random\_state=0)  **Mejores Parametros**  {'class\_weight': 'balanced', 'criterion': 'gini', 'min\_samples\_leaf': 1, 'min\_samples\_split': 2, 'n\_estimators': 150, 'random\_state': 0} | 0.5278587845944813 | 0.1410733937187844 |
| **BayesSearchCV**  BayesSearchCV(cv=3, estimator=RandomForestClassifier(), n\_iter=3, n\_jobs=-1, scoring='accuracy',  search\_spaces={'bootstrap': Categorical(categories=(True, False), prior=None),  'max\_depth': Integer(low=6, high=20, prior='uniform', transform='identity'),  'max\_features': Categorical(categories=('auto', 'sqrt', 'log2'), prior=None),  'min\_samples\_leaf': Integer(low=2, high=10, prior='uniform', transform='identity'),  'min\_samples\_split': Integer(low=2, high=10, prior='uniform', transform='identity'),  'n\_estimators': Integer(low=100, high=500, prior='uniform', transform='identity')},  verbose=4)  **Best Score** 0.7852759441781957 | **Mejores parámetros**  OrderedDict([('bootstrap', False),  ('max\_depth', 16),  ('max\_features', 'sqrt'),  ('min\_samples\_leaf', 5),  ('min\_samples\_split', 10),  ('n\_estimators', 339)])  **Mejor Estimador**  RandomForestClassifier(bootstrap=False, max\_depth=16, max\_features='sqrt',  min\_samples\_leaf=5, min\_samples\_split=10,  n\_estimators=339) | 0.6474280879625258 | 0.04461893188758023 |
| XGBClassifier  XGBClassifier(alpha=5, base\_score=0.5, booster='gbtree', colsample\_bylevel=1,  colsample\_bynode=1, colsample\_bytree=1, cv=3, gamma=0, gpu\_id=-1, importance\_type='gain', interaction\_constraints='', learning\_rate=0.1, max\_delta\_step=0, max\_depth=10, min\_child\_weight=1, missing=nan, monotone\_constraints='()', n\_estimators=50, n\_jobs=-1, num\_parallel\_tree=1, random\_state=0, reg\_alpha=5, reg\_lambda=1, scale\_pos\_weight=1, scoring='accuracy',  subsample=1, tree\_method='exact', validate\_parameters=1, verbose=3, verbosity=None)  **Boosting accuracy:** 0.7876330814122771 |  | 0.6599377515393277 | 0.06306751497930693 |
| Lineal SVC  Pipeline(steps=[('standardscaler', StandardScaler()),  ('linearsvc',  LinearSVC(class\_weight='balanced', loss='hinge',  multi\_class='crammer\_singer', random\_state=42))])  Accuracy: 15 |  | 0.62 | 0.3886221368615735 |

De la tabla anterior podemos deducir que el modelo XGB Classifier tiene el mejor desempeño según el accuracy, el roc\_auc\_score es bastante similiar para todos los modelos.