entrenamiento_modelos

July 11, 2025

1 ENTRENAMIENTO DE LOS MODELOS

[2]:

```
# INSTALAR LIBRERÍAS NECESARIAS
%pip install scikit-learn
```

```
Collecting scikit-learn
  Using cached scikit_learn-1.7.0-cp312-cp312-win_amd64.whl.metadata (14 kB)
Requirement already satisfied: numpy>=1.22.0 in
c:\users\carolina\documents\proyectos_programacion\predestu\.venv\lib\site-
packages (from scikit-learn) (2.3.1)
Collecting scipy>=1.8.0 (from scikit-learn)
  Using cached scipy-1.16.0-cp312-cp312-win_amd64.whl.metadata (60 kB)
Collecting joblib>=1.2.0 (from scikit-learn)
  Using cached joblib-1.5.1-py3-none-any.whl.metadata (5.6 kB)
Collecting threadpoolctl>=3.1.0 (from scikit-learn)
  Using cached threadpoolctl-3.6.0-py3-none-any.whl.metadata (13 kB)
Using cached scikit_learn-1.7.0-cp312-cp312-win_amd64.whl (10.7 MB)
Using cached joblib-1.5.1-py3-none-any.whl (307 kB)
Using cached scipy-1.16.0-cp312-cp312-win_amd64.whl (38.4 MB)
Using cached threadpoolctl-3.6.0-py3-none-any.whl (18 kB)
Installing collected packages: threadpoolctl, scipy, joblib, scikit-learn

   ---------- ----------------------------- 1/4 [scipy]
   ---------- ----------------------------- 1/4 [scipy]
   ---------- ----------------------------- 1/4 [scipy]
   ---------- ----------------------------- 1/4 [scipy]
   ---------- ----------------------------- 1/4 [scipy]
   ---------- ----------------------------- 1/4 [scipy]
   ---------- ----------------------------- 1/4 [scipy]
   ---------- ----------------------------- 1/4 [scipy]
   ---------- ----------------------------- 1/4 [scipy]
   ---------- ----------------------------- 1/4 [scipy]
   ---------- ----------------------------- 1/4 [scipy]
   ---------- ----------------------------- 1/4 [scipy]
   ---------- ----------------------------- 1/4 [scipy]
   ---------- ----------------------------- 1/4 [scipy]
   ---------- ----------------------------- 1/4 [scipy]
   ---------- ----------------------------- 1/4 [scipy]
```

```
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
```

```
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
---------- ----------------------------- 1/4 [scipy]
-------------------- ------------------- 2/4 [joblib]
```

```
-------------------- ------------------- 2/4 [joblib]
-------------------- ------------------- 2/4 [joblib]
-------------------- ------------------- 2/4 [joblib]
-------------------- ------------------- 2/4 [joblib]
-------------------- ------------------- 2/4 [joblib]
-------------------- ------------------- 2/4 [joblib]
-------------------- ------------------- 2/4 [joblib]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
------------------------------ --------- 3/4 [scikit-learn]
```

```
   ------------------------------ --------- 3/4 [scikit-learn]
   ------------------------------ --------- 3/4 [scikit-learn]
   ------------------------------ --------- 3/4 [scikit-learn]
   ------------------------------ --------- 3/4 [scikit-learn]
   ------------------------------ --------- 3/4 [scikit-learn]
   ------------------------------ --------- 3/4 [scikit-learn]
   ------------------------------ --------- 3/4 [scikit-learn]
   ------------------------------ --------- 3/4 [scikit-learn]
   ------------------------------ --------- 3/4 [scikit-learn]
   ------------------------------ --------- 3/4 [scikit-learn]
   ------------------------------ --------- 3/4 [scikit-learn]
   ------------------------------ --------- 3/4 [scikit-learn]
   ------------------------------ --------- 3/4 [scikit-learn]
   ------------------------------ --------- 3/4 [scikit-learn]
   ------------------------------ --------- 3/4 [scikit-learn]
   ------------------------------ --------- 3/4 [scikit-learn]
   ------------------------------ --------- 3/4 [scikit-learn]
   ------------------------------ --------- 3/4 [scikit-learn]
   ------------------------------ --------- 3/4 [scikit-learn]
   ------------------------------ --------- 3/4 [scikit-learn]
   ------------------------------ --------- 3/4 [scikit-learn]
   ------------------------------ --------- 3/4 [scikit-learn]
   ------------------------------ --------- 3/4 [scikit-learn]
   ------------------------------ --------- 3/4 [scikit-learn]
   ------------------------------ --------- 3/4 [scikit-learn]
   ------------------------------ --------- 3/4 [scikit-learn]
   ------------------------------ --------- 3/4 [scikit-learn]
   ------------------------------ --------- 3/4 [scikit-learn]
   ---------------------------------------- 4/4 [scikit-learn]

Successfully installed joblib-1.5.1 scikit-learn-1.7.0 scipy-1.16.0
threadpoolctl-3.6.0
Note: you may need to restart the kernel to use updated packages.
```

[17]:

```python
# IMPORTACIÓN DE LIBRERÍAS
from sklearn.linear_model import LogisticRegression
from sklearn.model_selection import train_test_split
from sklearn.metrics import classification_report
import pandas as pd
import joblib
from sklearn.metrics import classification_report, confusion_matrix,␣
 ↪accuracy_score
from sklearn.ensemble import RandomForestClassifier
```

[6]:

```python
# CARGAR DATASET
df_entrenamiento = pd.read_csv('../fuentes_datos/datasets_finales/CSV/
 ↪dataset_entrenamiento.csv', delimiter=';')
```

1.1 Regresión Logística Multiclase

[7]:
```
X = df_entrenamiento[['id_escuela', 'id_plan', 'id_ciclo', 'n_reprobaciones',␣
 ↪'id_semestre', 'tutorias', 'uso_acra']]
y = df_entrenamiento['riesgo_academico']

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3,␣
 ↪random_state=42)

modelo_lr = LogisticRegression(multi_class='multinomial', solver='lbfgs',␣
 ↪max_iter=1000)
modelo_lr.fit(X_train, y_train)
y_pred_lr = modelo_lr.predict(X_test)

print("=== Logistic Regression ===")
print(classification_report(y_test, y_pred_lr))
```

```
c:\Users\carolina\Documents\Proyectos_programacion\PredEstu\.venv\Lib\site-
packages\sklearn\linear_model\_logistic.py:1264: FutureWarning: 'multi_class'
was deprecated in version 1.5 and will be removed in 1.7. From then on, it will
always use 'multinomial'. Leave it to its default value to avoid this warning.
  warnings.warn(
```

```
=== Logistic Regression ===
              precision    recall  f1-score   support

           0       0.00      0.00      0.00     40339
           1       0.00      0.00      0.00     42759
           2       0.34      0.11      0.16     82819
           3       0.33      0.90      0.48     86884
           4       0.00      0.00      0.00     13197

    accuracy                           0.33    265998
   macro avg       0.13      0.20      0.13    265998
weighted avg       0.21      0.33      0.21    265998
```

```
c:\Users\carolina\Documents\Proyectos_programacion\PredEstu\.venv\Lib\site-
packages\sklearn\linear_model\_logistic.py:470: ConvergenceWarning: lbfgs failed
to converge after 1000 iteration(s) (status=1):
STOP: TOTAL NO. OF ITERATIONS REACHED LIMIT

Increase the number of iterations to improve the convergence (max_iter=1000).
You might also want to 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(
c:\Users\carolina\Documents\Proyectos_programacion\PredEstu\.venv\Lib\site-
packages\sklearn\metrics\_classification.py:1706: UndefinedMetricWarning:
Precision is ill-defined and being set to 0.0 in labels with no predicted
samples. Use `zero_division` parameter to control this behavior.
  _warn_prf(average, modifier, f"{metric.capitalize()} is", result.shape[0])
c:\Users\carolina\Documents\Proyectos_programacion\PredEstu\.venv\Lib\site-
packages\sklearn\metrics\_classification.py:1706: UndefinedMetricWarning:
Precision is ill-defined and being set to 0.0 in labels with no predicted
samples. Use `zero_division` parameter to control this behavior.
  _warn_prf(average, modifier, f"{metric.capitalize()} is", result.shape[0])
c:\Users\carolina\Documents\Proyectos_programacion\PredEstu\.venv\Lib\site-
packages\sklearn\metrics\_classification.py:1706: UndefinedMetricWarning:
Precision is ill-defined and being set to 0.0 in labels with no predicted
samples. Use `zero_division` parameter to control this behavior.
  _warn_prf(average, modifier, f"{metric.capitalize()} is", result.shape[0])
```

[11]:

```
# Guarda el modelo en un archivo .pkl
joblib.dump(modelo_lr, "../modelos_generados/
 ↪modelo_regresion_logistica_multiclase.pkl")
```

[11]:

```
['../modelos_generados/modelo_regresion_logistica_multiclase.pkl']
```

[15]:

```
# Predicciones
y_pred = modelo_lr.predict(X_test)

# Reporte general
print("=== Evaluación del Modelo ===")
print(classification_report(y_test, y_pred))
```

```
=== Evaluación del Modelo ===
              precision    recall  f1-score   support

           0       0.00      0.00      0.00     40339
           1       0.00      0.00      0.00     42759
           2       0.34      0.11      0.16     82819
           3       0.33      0.90      0.48     86884
           4       0.00      0.00      0.00     13197

    accuracy                           0.33    265998
   macro avg       0.13      0.20      0.13    265998
weighted avg       0.21      0.33      0.21    265998

c:\Users\carolina\Documents\Proyectos_programacion\PredEstu\.venv\Lib\site-
packages\sklearn\metrics\_classification.py:1706: UndefinedMetricWarning:
Precision is ill-defined and being set to 0.0 in labels with no predicted
samples. Use `zero_division` parameter to control this behavior.
  _warn_prf(average, modifier, f"{metric.capitalize()} is", result.shape[0])
c:\Users\carolina\Documents\Proyectos_programacion\PredEstu\.venv\Lib\site-
```

```
packages\sklearn\metrics\_classification.py:1706: UndefinedMetricWarning:
Precision is ill-defined and being set to 0.0 in labels with no predicted
samples. Use `zero_division` parameter to control this behavior.
  _warn_prf(average, modifier, f"{metric.capitalize()} is", result.shape[0])
c:\Users\carolina\Documents\Proyectos_programacion\PredEstu\.venv\Lib\site-
packages\sklearn\metrics\_classification.py:1706: UndefinedMetricWarning:
Precision is ill-defined and being set to 0.0 in labels with no predicted
samples. Use `zero_division` parameter to control this behavior.
  _warn_prf(average, modifier, f"{metric.capitalize()} is", result.shape[0])
```

[16]:

```
# Precisión total
print("Precisión del modelo:", accuracy_score(y_test, y_pred))
```

```
Precisión del modelo: 0.32824306949676313
```

1.2 Random Forest Classifier

[18]:

```
modelo_rf = RandomForestClassifier(n_estimators=100, random_state=42)
modelo_rf.fit(X_train, y_train)
y_pred_rf = modelo_rf.predict(X_test)

print("=== Random Forest ===")
print(classification_report(y_test, y_pred_rf))
```

```
=== Random Forest ===
              precision    recall  f1-score   support

           0       0.50      0.39      0.44     40339
           1       0.52      0.36      0.43     42759
           2       0.54      0.60      0.57     82819
           3       0.54      0.66      0.59     86884
           4       0.57      0.30      0.39     13197

    accuracy                           0.53    265998
   macro avg       0.53      0.46      0.48    265998
weighted avg       0.53      0.53      0.53    265998
```

[19]:

```
# Guarda el modelo en un archivo .pkl
joblib.dump(modelo_rf, "../modelos_generados/modelo_random_forest.pkl")
```

[19]:

```
['../modelos_generados/modelo_random_forest.pkl']
```

[20]:

```
# Predicciones
y_pred = modelo_rf.predict(X_test)

# Reporte general
print("=== Evaluación del Modelo ===")
print(classification_report(y_test, y_pred))
```

```
=== Evaluación del Modelo ===
              precision    recall  f1-score   support

           0       0.50      0.39      0.44     40339
           1       0.52      0.36      0.43     42759
           2       0.54      0.60      0.57     82819
           3       0.54      0.66      0.59     86884
           4       0.57      0.30      0.39     13197

    accuracy                           0.53    265998
   macro avg       0.53      0.46      0.48    265998
weighted avg       0.53      0.53      0.53    265998
```

[21]:

```
# Precisión total
print("Precisión del modelo:", accuracy_score(y_test, y_pred))
```

```
Precisión del modelo: 0.53356792156332
```