

AutoML

(Machine Learning para vagos)

Manuel Garrido

#whoami (Manuel Garrido)

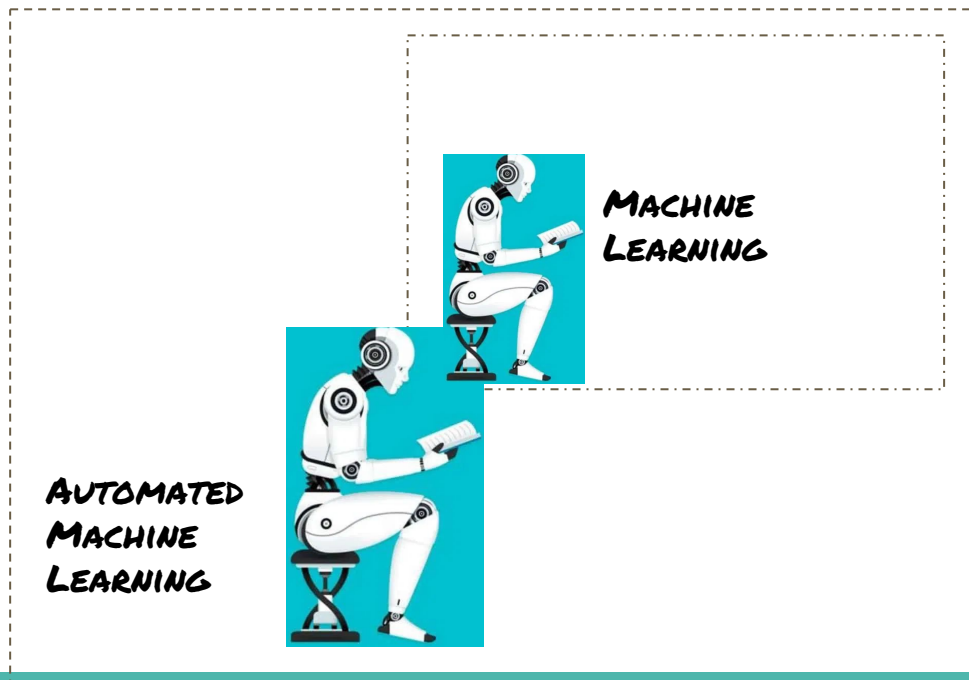
- Lead Data Architect en [Daltix](#)
- Profesor en [ENAE](#) y [Lisbon Data Science Academy](#)
- Autor del curso de Udemty “[Machine Learning y Data Science con Python](#)”
- UPV: Ingeniería Industrial
- IE: Master In Management
- Consultor → Analista → Data Scientist
- Excel - Excel VBA - R - Python

hola@manugarri.com



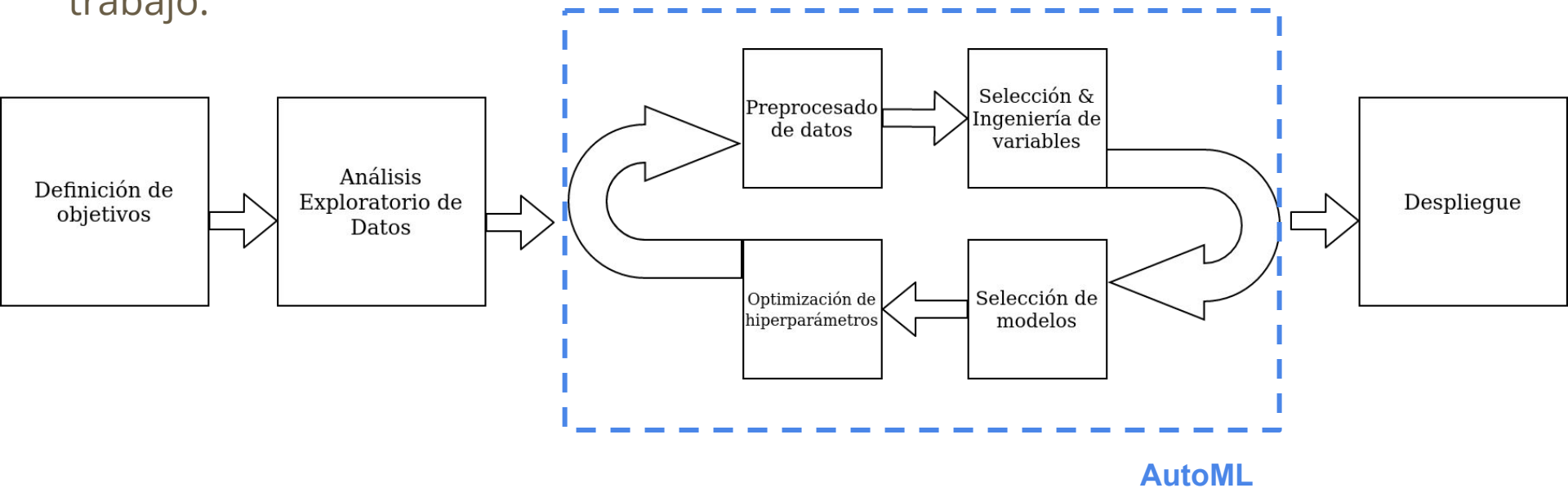
Que es AutoML

Llamamos AutoML al conjunto de técnicas que ayudan al desarrollo automatizado de modelos predictivos.



Que es AutoML

En un proyecto de Machine Learning hay varios pasos que suelen ser repetitivos, AutoML se enfoca en facilitar dichos pasos para que los Data Scientists puedan concentrarse en las partes más complejas y creativas de su trabajo:



Que NO es AutoML

- *Una solución plug and play.* Las soluciones Open Source/comerciales actuales necesitan que nuestro pipeline de datos tenga ciertos requisitos. Además, AutoML es un dominio bastante nuevo y en constante evolución (es bastante inestable).
- *Un reemplazo para un Data Scientist!* Gran parte del valor que aporta un Data Scientist no está en el desarrollo de Modelos predictivos, sino en trasladar necesidades de negocio a un problema de ML.

AutoSklearn

<https://automl.github.io/auto-sklearn/master/>

AutoML enfocado a ser un reemplazo de un estimador de sklearn.

Proporciona:

- Selección de modelos (scikit-learn)
- Optimización de modelos (via [SMAC](#))

```
# from autosklearn.classification import AutoSklearnClassifier
from autosklearn.regression import AutoSklearnRegressor
from autosklearn.metrics import mean_squared_error

automl = AutoSklearnRegressor(
    time_left_for_this_task=180, # busca modelos durante 3 minutos
    metric=mean_squared_error
)

# Ajustamos el modelo de AutoML
automl.fit(X_train, y_train)

# Predecimos como con cualquier estimador de sklearn
y_hat = automl.predict(X_test)

# podemos usar joblib para exportar modelos como en sklearn
joblib.dump(automl, "autosklearn.joblib")
```

TPot

<http://epistasislab.github.io/tpot/>

AutoML enfocado a ser un reemplazo de un estimador de sklearn.

Proporciona:

- Selección de modelos (scikit-learn, xgboost)
- Optimización de modelos (via DEAP)

```
# from tpot import TPOTClassifier
from tpot import TPOTRegressor

tpot = TPOTRegressor(
    max_time_mins=3,
    scoring="neg_mean_squared_error"
)

# ajustamos modelo
tpot.fit(X_train, y_train)

# inferencia
tpot.predict(X_test)

# exportamos el mejor pipeline (un pipeline normal de sklearn)
joblib.dump(tpot.fitted_pipeline_, "tpot.joblib")

# tpot guarda el pipeline como un script!
tpot.export('tpot_boston_pipeline.py')
```

AutoKeras

<https://autokeras.com/>

AutoML que implementa modelos de Deep Learning de la popular libreria Keras Proporciona:

- Selección de modelos (modelos de keras para diversos tipos de inputs, texto, imagenes, estructurados o multimodales) ([NAS](#))
- Optimización de modelos

```
import autokeras as ak
# ak.StructuredDataClassifier
reg = ak.StructuredDataRegressor(
    max_trials=3 # prueba 3 estimadores
)
# Ajustamos a los datos de entrenamiento
reg.fit(
    x=X_train,
    y=y_train,
    epochs=5, # entrena durante 5 epocas
)
# Predecimos en datos de test
predicted_y = reg.predict(X_test)

# exporta como modelo keras
best_model = reg.export_model()
```


PyCaret

<https://pycaret.org/>

Librería que ayuda en prácticamente todos los aspectos de ML. Proporciona:

- Preprocesado de datos
- Selección de variables
- Selección de modelos
(sklearn,xgboost, lightgbm)
- Optimización de modelos
- Despliegue de modelos

```
# from pycaret.classification import setup, compare_models, automl
from pycaret.regression import setup, compare_models, automl
#preparacion de datos
setup(
    data=df,
    target=OBJETIVO,
    train_size=0.9
)

# Realiza la busqueda de modelos
compare_models(
    budget_time=3, #finaliza busqueda despues de 3 minutos
    fold=5,
    sort="MSE" # ordenalos en funcion del MSE
)

# finaliza el mejor modelo encontrado
best_automl = automl(optimize='MSE')

# inferencia
best_automl.predict(X_test)
```

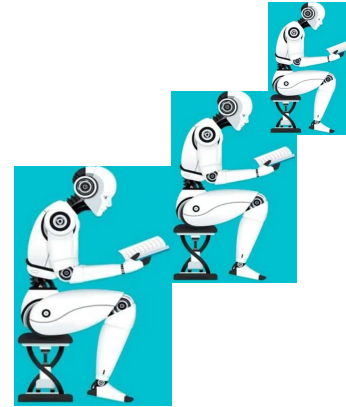
Y muchos más!

- [H2O AutoML](#)
- [Ludwig](#)
- [DEvol](#)
- [ADANet](#)
- [MLBox](#)
- [Dabl](#)
- [Libra](#)
- [OptimalFlow](#)
- [AutoPytorch](#)
- [prophet](#)*
- [Pandas-profiling](#)*

AutoML as a Service

- [Google Cloud AutoML \(AutoML Tables\)](#)
- [AWS SageMaker AutoPilot](#)
- [Azure Automated Machine Learning](#)
- [DataRobot](#)
- [Obviously.ai](#)
- [Fritz AI studio](#)

AutoAutoML



**AUTO
AUTO MACHINE LEARNING**

<https://github.com/manugarri/autoautoml>

A cartoon illustration of a light brown dog with large, expressive eyes and a small black hat, sitting on a wooden stool. The dog is positioned in front of a large, stylized campfire with bright yellow and orange flames. To the right of the dog is a small wooden table with a white mug on it. The background is a simple greenish-yellow color, suggesting an outdoor setting. The text '¿Preguntas?' is overlaid in the center of the image.

¿Preguntas?