A estas alturas, después de ver la forma en la que scikit-learn nos permite semi-automatizar la optimización de nuestros modelos con GridSearchCV y RandomizedSearchCV es posible que te estés preguntando ¿Cuál es el límite de esta automatización?

Pues te sorprenderás,

Automated Machine Learning (AutoML), es un concepto relativamente nuevo que en general pretende la completa automatización de todo el proceso de Machine Learning, desde la extracción de los datos hasta su publicación final de cara a los usuarios.

Sin embargo, este ideal aún está en desarrollo en la mayoría de las etapas del proceso de Machine Learning y aún se depende bastante de la intervención humana. Aún con esto, es importante que seamos conscientes de que ya existen varias herramientas que nos acercan un poco a esta meta casi tomada de la ciencia ficción.

Puedes encontrar más información leyendo el siguiente enlace:

<https://itmastersmag.com/noticias-analisis/que-es-automated-machine-learning-la-proxima-generacion-de-inteligencia-artificial/>

La herramienta que te quiero presentar en esta clase se llama auto-sklearn, y nos ayudará a llevar aún un paso más lejos nuestro proceso de selección y optimización de modelos de machine learning. Dado que automáticamente prueba diferentes modelos predefinidos y configuraciones de parámetros comunes hasta encontrar la que más se ajuste según los datos que le pasemos como entrada. Con esta herramienta podrás entrenar modelos tanto de clasificación como de regresión por igual.

Para una lista de los clasificadores disponibles consulta:

<https://github.com/automl/auto-sklearn/tree/master/autosklearn/pipeline/components/classification>

Y para una lista de los regresores disponibles consulta:

<https://github.com/automl/auto-sklearn/tree/master/autosklearn/pipeline/components/regression>

Ten en cuenta que podrás añadir modelos personalizados al proceso siguiendo los pasos descritos en la documentación.

**auto-sklearn:**

Esta herramienta es una librería basada en los algoritmos de scikit-learn, aunque hay que tener presente que es una librería externa y se debe instalar siempre por aparte. En todo caso al ser una librería de Python se puede combinar sin ningún problema con el resto de nuestro código desarrollado para scikit-learn, incluso permitiendo la exportación de modelos ya entrenados para su posterior uso.

Enlace a la documentación: <https://automl.github.io/auto-sklearn/master/index.html>

Como lo puedes ver en su página web, los requerimientos para probar autosklearn son:

* Se requiere un sistema operativo basado en Linux.
* Python (>=3.5) .
* Compilador para C++ (con soporte para C++11), por ejemplo GCC.
* SWIG (versión 3.0 o superior).

La forma de hacer funcionar nuestro algoritmo no podría ser más fácil. Nos resultará bastante familiar a estas alturas después de haber trabajado tanto con sklearn.

import autosklearn.classification

cls = autosklearn.classification.AutoSklearnClassifier()

cls.fit(X\_train, y\_train)

predictions = cls.predict(X\_test)

¡Te invito a conocer a fondo esta herramienta a través de su documentación y decidir si es la estrategia que estás buscando para tu problema específico!