



Programación Python para Big Data

Lección 3: PyCaret y AutoML

Programación Python para Big Data

ACTIVIDAD LECCIÓN 3

Objetivos

- | En el presente tema hemos hablado de algunos conceptos básicos de PyCaret y el AutoML.
De modo que el objetivo de esta tarea, en este tema 3, será aprender un poco más respecto de ello.
- | Kaggle es una herramienta imprescindible para cualquier profesional en Data Science. Por lo cual trabajaremos con ello en la presente actividad.
- | Otra cosa que trataremos de hacer es fomentar la proactividad de el/la alumno/a. En muchas ocasiones tendremos que formarnos sin ayuda de nadie, de modo que puede servirnos como “entrenamiento” o “práctica”.

Contenido correspondiente a Lección 3:

1. PyCaret y AutoML
 - 1.1. Introducción a PyCaret
 - 1.2. Resolución de Titanic Dataset con PyCaret
 - 1.3. Kaggle

Actividad relacionada con la Lección 3:

Primera parte de la Actividad (Hasta 5 puntos)

El/la alumno/a deberá enviar un archivo .ipynb con el ejercicio resuelto en clase. (Si alguien quiere mejorarlo también se tendrá en cuenta).

A continuación deja una celda en blanco o un par de ellas en el propio .ipynb y continúa con la siguiente parte de la actividad..

Al igual que siempre.. Pregunta y respuesta deberán ir en celdas diferentes.

-Segunda parte de la actividad

-Pregunta 1..

-Respuesta 1..

-Pregunta 2..

-Respuesta 2..

y así sucesivamente..

Segunda parte de la actividad (Hasta 2 puntos)

Símplemente queremos saber tu opinión al respecto de AutoML,

Queremos saber:

- | Pregunta 1 - ¿Te costó esfuerzo aprender el presente tema?
- | Pregunta 2 - ¿Qué te ha llamado la atención de PyCaret?
- | Pregunta 3 - ¿Notas que es más simple haciéndolo así que como se hizo en Fundamentos de Big Data? Si opinas diferente la respuesta no es correcta/incorrecta, siempre y cuando argumentes tu respuesta.
- | Pregunta 4 - Busca más herramientas AutoML que existan y enumera alguna (simplemente queremos que te familiarices con ello, nada más).

NOTA: Esta actividad se puntuará independientemente de lo que se responda.

Tercera parte de la actividad (Hasta 1 punto)

Hemos mencionado un parámetro llamado AUC.

Haz un pequeño resumen con información que encuentres en Internet acerca de AUC y la curva ROC.

Puedes añadir la fuente donde obtuviste la información.

No es relevante todavía pero, estaría bien si lees algo al respecto.

Si te cuesta entenderlo, puedes, sin problema comentarlo.

(Habíamos comentado que se explicará en asignaturas sucesivas, pero si aprendes algo al respecto ahora te vendrá bien para entender un poco mejor el tema).

Cuarta parte de la actividad (Hasta 1 punto)

Hemos usado la Confusion Matrix en esta lección, pero no hemos explicado mucho,

Haz un pequeño resumen con información que encuentres en Internet acerca de esto.

No es relevante todavía pero, estaría bien si lees algo al respecto.

Si te parece complicado, o no lo entiendes puedes comentarlo.

Quinta parte de la actividad (Hasta 1 punto)

Envía tu predicción a Kaggle y adjunta un pantallazo con tu resultado.

Formato: .png o cualquier formato compatible con Moodle.