

**POSTWORK
SESIÓN 01****Objetivo:**

Definir los proyectos que te acompañarán durante todo el módulo de Machine Learning. El (los) proyecto(s) que elijas deberán cumplir con todos los requisitos que enlistaremos abajo. Te recomiendo utilizarlos como un checklist.

Requisitos a tomar en cuenta antes de elegir un proyecto:

- Todos los proyectos deben ser revisados y evaluados previamente por el experto.
- Todos los proyectos deben tocar al menos un tema de los vistos en el módulo de ML.
- Los proyectos deben estar dentro de un rango realizable: crear un proyecto que conduzca tu auto por sí solo podría escaparse de las manos y necesitar herramientas que no veremos en el curso.

Requisitos del postwork de hoy:

- Tu proyecto debe ser programado en Python 3. Muchos otros proyectos pueden ser contruidos con otros lenguajes de programación; sin embargo, al construirlos en Python 3 nos permitirás ayudarte a mejorarlo. Si lo construyes en Go o Java, quizás no pueda ayudarte a mejorarlo, ya sea por desconocimiento del lenguaje, o porque no existen herramientas que utilizaremos durante todo el curso.
- Puedes elegir si desarrollar uno o varios proyectos tuyos, o bien, puedes desarrollar los mini proyectos que enlistaremos en cada uno de los módulos. Sin embargo, debes estar consciente que los mini proyectos podrían no ser del todo de tu agrado. Utilizaremos datos y problemas genéricos en los mini proyectos, para abordar el algoritmo de machine learning.
- Puedes elegir un solo proyecto o varios, pero para cada proyecto necesitarás evaluar lo siguiente:
 - o Si tu proyecto puedes desarrollarlo con un programa estándar sin utilizar Machine Learning y te da una buena solución, quizás no valga la pena aplicarle Machine Learning (por ejemplo: contar cuantas letras hay en un texto).
 - o Si tu proyecto es un problema abierto, quizás no vamos a poder ayudarte con ello. Léase: un problema abierto es aquellos que actualmente investigan los

Postwork - Introducción a Machine Learning

académicos y centros de investigación y son sumamente difíciles como “reconoce a todas las personas que se encuentran en el metro”.

- o Debes contar con una buena cantidad de datos de entrada para clasificar o predecir. Si, por ejemplo, en tu proyecto quieres saber cuántas personas usan las ecobici de la Ciudad de México, pero tienes solamente datos de 2 horas, no hay mucho que podamos hacer.
- o Debes estar consciente que el crear un proyecto de Machine Learning puede producirte cierta frustración: Machine Learning es un proceso diferente a programar, en el cual no siempre obtendrás resultados satisfactorios. Esto está bien y es normal. Lo importante es que puedes intentarlo varias veces, con diferentes experimentos. Te recomiendo manejar una bitácora de experimentos.

Bitácora de experimentos:

Una bitácora de experimentos registra aquellas cosas que has intentado y el resultado que has tenido. Todos los científicos de datos deberían utilizar una bitácora para evitar frustrarse (créeme, es muy común que pase y muy pocos Data Scientists las usan), por lo que te recomiendo que la sigas.

Una entrada en tu bitácora de experimentos se ve algo así:

*Experimento 1**Fecha y hora: 12 / enero / 2021 -- 12:24pm**Algoritmo: Red neuronal artificial**Dataset: dataset1.csv**Configuración:*

- *Algoritmo RMSProp*
- *Learning Rate: 0.01*
- *Arquitectura [10,5,2]*

Resultado:

- *Accuracy 95.2%*
- *Sensibilidad 88.9%*
- *Especificidad 98.4%*

De esta manera, ya no tienes que utilizar la misma configuración, porque ya sabes qué resultado vas a obtener. Eso te permite tomar todos los experimentos y solamente conservar el que mejor ha funcionado. Tomate un momento para crear una adecuada bitácora y mantenla contigo durante todos tus Postworks.