

TFM (Trabajo Fin de Máster) del TECI

Curso académico: 2019/2020

Título:

MLOps y cómo industrializar el uso de *Machine Learning* en las empresas: un caso de uso

Tipo (marca una casilla):
Profesional

☐

Académico

☒

Institución:

Persona:

Ponente (si procede):

Observación:

Al rellenar los siguientes puntos hay que considerar que la carga de trabajo no debe superar las 300 horas para el estudiante

Problema a tratar:

Hoy en día, gracias al avance tecnológico, contar con aplicaciones de *Machine Learning* (ML) en las empresas no se considera un reto. De hecho, el número de personas que hacen uso de dichas técnicas y el número de modelos disponibles en instituciones como los bancos, es cada vez más grande. Sin embargo, la puesta en producción del ML es una preocupación y un desafío en sí mismo.

Las buenas prácticas establecidas por los ingenieros de desarrollo de software se están adecuando a la puesta en producción del ML. *MLOps* es la unión de *Machine Learning* con *Operationalization*, propuesta metodológica que se está investigando e implantando para garantizar la utilización de los modelos ML, automatizando tareas y garantizando la calidad de los modelos a lo largo del tiempo.

Al mismo tiempo, otra de las grandes ventajas que trae el avance tecnológico es la aparición de herramientas que facilitan la obtención de previsiones a partir de algoritmos ML - *AutoML*¹ es un ejemplo de ello. Esas herramientas permiten disponer de miles de modelos, sin apenas la intervención de una analista. Aun cuando se trata de una gran noticia, no se puede dejar de lado la importancia que tiene el análisis de los datos por parte de un profesional (científico de datos) que entienda del problema de negocio, garantice la calidad de los datos utilizado, interprete las salidas de los algoritmos/herramientas y sobre todo verifique que los resultados son coherentes, proporcionando el valor esperado y, por lo tanto, pueden ser puestos a disposición de toda la organización (es decir, en producción).

En este trabajo se desarrolla el enfoque *MLOps*, desde la perspectiva de las empresas en España. Además de abordar los aspectos metodológicos, se proponen varias alternativas

¹ <https://www.microsoft.com/en-us/research/project/automl/>

de arquitectura (herramientas y configuración del sistema) para llevar a cabo un *MLOps* de éxito. También se presenta el *MLOps* utilizado en casos reales de empresas donde el objetivo de negocio es detectar la ocurrencia de un evento - como el fraude de transacciones online, el malware informático, etc. En estos casos, se mostrará cómo se desarrolla cada una de las etapas del proceso de modelización, incluyendo varias opciones de creación de variables para el modelo y métricas de diagnóstico y selección de modelos de clasificación.

Objetivos:

- Entender cómo se utiliza la ciencia de datos en un entorno empresarial
- Conocer las diferentes alternativas que existen en el mercado para la adopción de la ciencia de datos en las empresas con tecnología *cloud*.
- Presentar una propuesta de industrialización de modelos de *Machine Learning: MLOps*
- Desarrollar un caso de uso de clasificación predictiva con algoritmos de *Machine Learning*