



Inteligencia Artificial para profesionales TIC

Ciclo de vida de ML

Índice



01. El ciclo de vida del ML

02. MLOps

03. Enfoque centrado en los datos

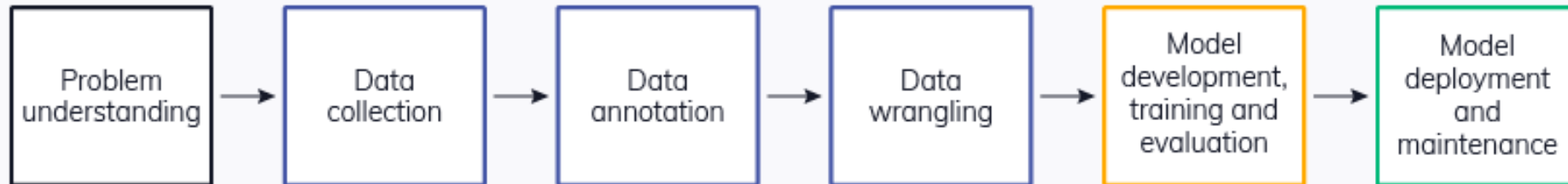
04. Desafíos en el ciclo de vida del ML



1. El ciclo de vida del ML

Por qué?

El ciclo de vida del aprendizaje automático es un proceso estructurado que ayuda a las empresas a crear productos de IA sostenibles, rentables y de calidad.





1. El ciclo de vida del ML

Etapas principales

Planificación: Esta fase implica evaluar el alcance, la métrica de éxito y la viabilidad de la aplicación de ML. Las consideraciones clave incluyen comprender el negocio y cómo el aprendizaje automático puede mejorar el proceso actual, evaluar la disponibilidad y aplicabilidad de los datos, considerar las restricciones legales y garantizar la solidez, la escalabilidad y la explicabilidad.

Recogida de datos: Este paso implica la recopilación de los datos necesarios para el entrenamiento del modelo. La cantidad y la calidad de los datos son cruciales para crear un modelo de aprendizaje automático exitoso.

Preparación de datos: En esta etapa, los datos se limpian, se preprocesan y se transforman en un formato adecuado para el entrenamiento del modelo.

Anotación de datos: Este paso implica etiquetar los datos, lo cual es esencial para los algoritmos de aprendizaje supervisado. La calidad de la anotación de datos afecta directamente al rendimiento del modelo, por lo que es importante invertir tiempo en la creación de directrices de anotación claras y detalladas.

Construcción de modelos y entrenamiento: Esta es la parte más intensiva en código del ciclo de vida del aprendizaje automático, donde los programadores diseñan y ensamblan el algoritmo en sí. El modelo se entrena con los datos preparados, y los desarrolladores supervisan el proceso y ajustan los hiperparámetros según sea necesario para optimizar el enfoque del modelo en aspectos específicos de los datos.

Evaluación del modelo: En esta fase, el rendimiento del modelo se evalúa mediante métricas de éxito predefinidas, como la precisión, la puntuación F1 o el AUC. Esto ayuda a determinar si el modelo cumple con los objetivos establecidos y está listo para su implementación.

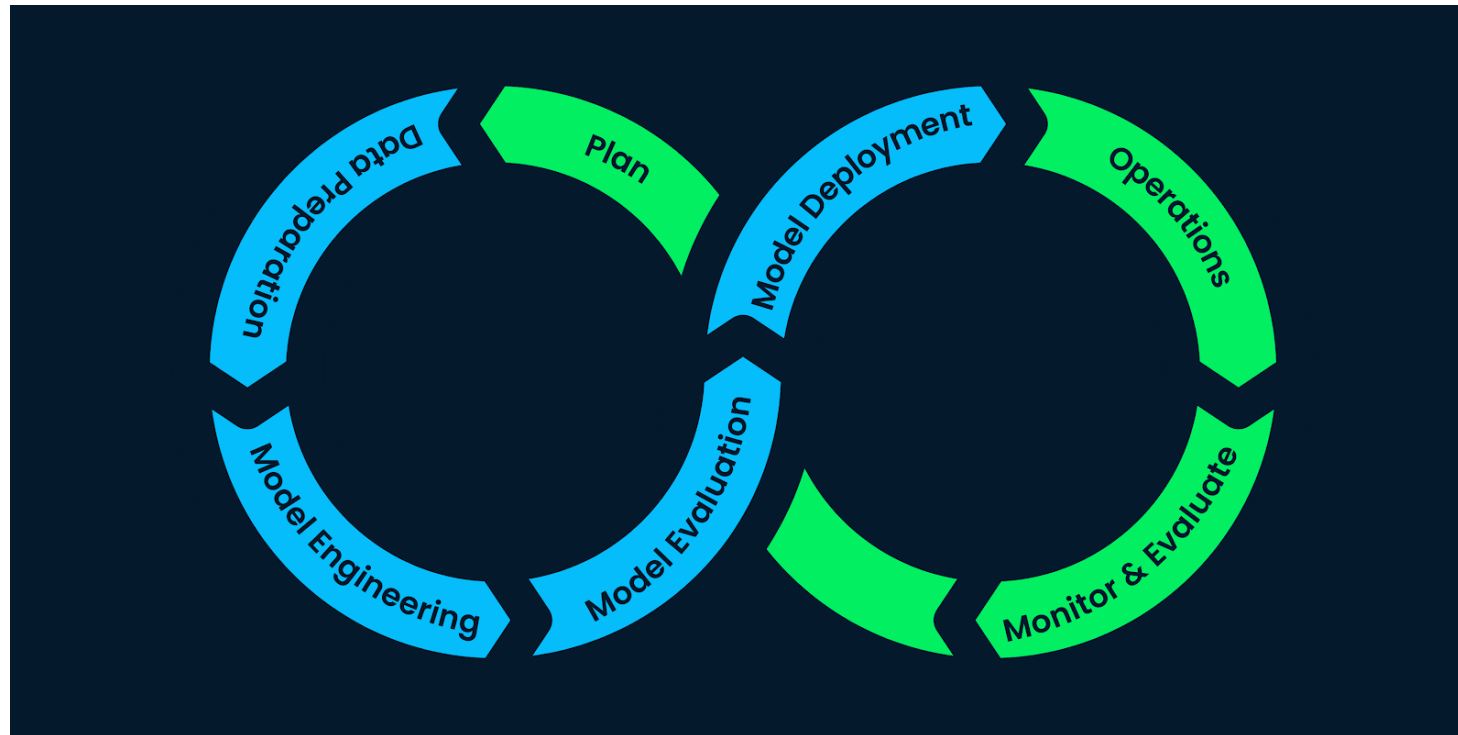
Implementación y mantenimiento de modelos: Una vez que el modelo se implementa en un entorno de producción, debe supervisarse y mantenerse constantemente para garantizar un rendimiento óptimo. Los desarrolladores pueden aprender del rendimiento del modelo e iterar en el proceso para mejorarlo aún más.



1. El ciclo de vida del ML

Etapas principales

Más que una secuencia de etapas, es un ciclo continuo:





2. MLOps

Un nuevo rol?

MLOps, u operaciones de aprendizaje automático, se refiere al conjunto de prácticas y principios que agilizan y automatizan la gestión de los modelos de aprendizaje automático a lo largo de su ciclo de vida.

En el contexto del ciclo de vida de ML, MLOps tiene como objetivo cerrar la brecha entre el desarrollo, la implementación y el mantenimiento mediante la integración de modelos de aprendizaje automático con los procesos de desarrollo, TI y negocios existentes.

MLOps se centra en la colaboración, la integración y entrega continuas, la supervisión y la gobernanza para garantizar la implementación eficiente, confiable y escalable de los modelos de ML, lo que conduce a una toma de decisiones más rápida y precisa.

MLOps no es necesariamente un rol nuevo, sino más bien una disciplina en crecimiento dentro de los campos de la ciencia de datos, el aprendizaje automático y la ingeniería de software.

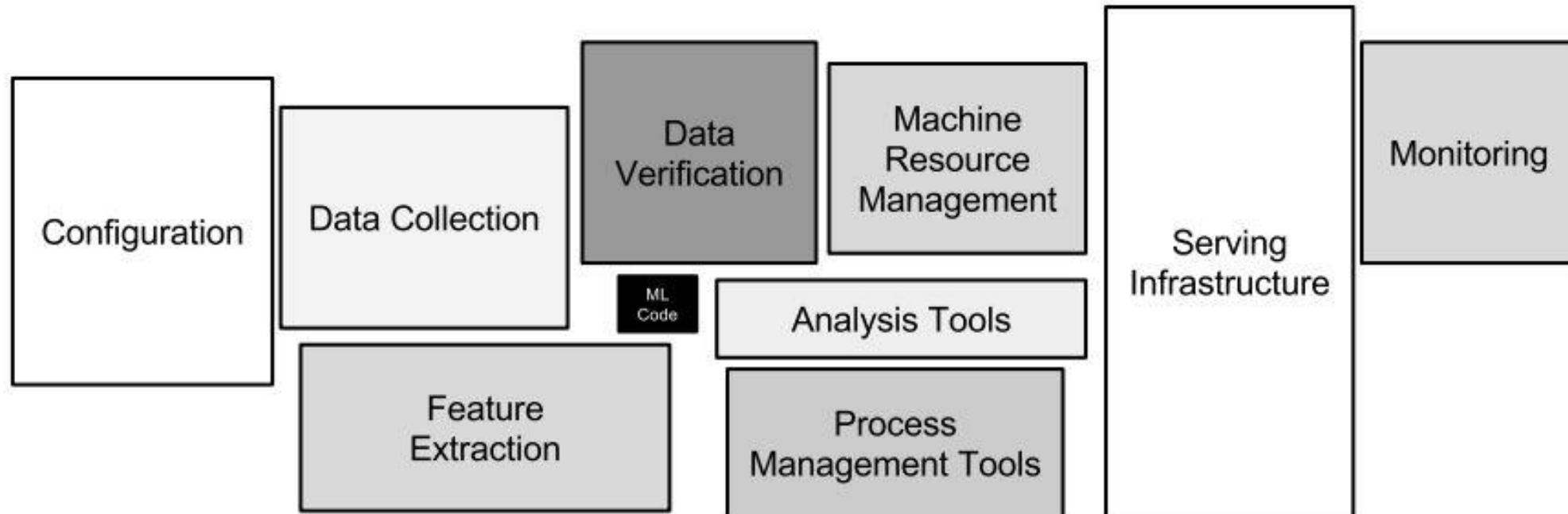
Estos roles suelen requerir una combinación de habilidades de ciencia de datos, ingeniería de software y DevOps, así como experiencia específica del dominio.



3. Enfoque centrado en los datos

Los datos son la clave

El enfoque centrado en los datos es fundamental:





4. Desafíos en el ciclo de vida del ML

Bueno, casi un milagro para que funcione

Datos de calidad: La falta de datos de calidad puede dar lugar a predicciones erróneas, lo que hace que la limpieza de los datos sea crucial para el éxito del aprendizaje automático.

Integración de tecnología: Las tecnologías incompatibles y los sistemas heredados pueden obstaculizar los proyectos de aprendizaje automático y la escalabilidad.

Exhaustividad de los datos: Garantizar datos sin procesar fiables y exhaustivos para los conjuntos de datos de entrenamiento es vital para obtener análisis y conclusiones precisos.

Colaboración humana: Los proyectos se benefician de la colaboración de los científicos de datos y los usuarios finales en los procesos de planificación y limpieza de datos.

Interpretabilidad del modelo: Hacer que los modelos complejos de aprendizaje automático sean comprensibles para las partes interesadas es esencial para el éxito del proyecto.

Generalización de modelos: Evitar el sobreajuste es crucial para garantizar que los modelos funcionen bien con datos nuevos e invisibles.

Implementación y mantenimiento: Se necesita un monitoreo y una actualización constantes para un rendimiento óptimo, lo que requiere un equipo diverso de expertos.

Consideraciones éticas: Garantizar la equidad y la imparcialidad de los modelos es importante para abordar los posibles sesgos y preocupaciones éticas.



4. Desafíos en el ciclo de vida del ML

Bueno, casi un milagro para que funcione

"[...] essentially, all models are wrong, but some are useful [...]"

George E. P. Box. 1987



