

Actividad 3. Implementación de un modelo de aprendizaje automático en un contexto real para la predicción

End-to-End Machine Learning Workflow

1 Descripción

En general, el objetivo de un proyecto de aprendizaje automático es construir un modelo utilizando los datos recogidos y aplicando a ellos algoritmos de aprendizaje automático. Por lo tanto, todo software basado en ML (*machine learning*) incluye tres tareas principales: datos, modelo de ML y código. En correspondencia con estas tareas, el flujo de trabajo típico del aprendizaje automático consta de tres fases principales:

- Ingeniería de datos: adquisición y/o preparación de datos (extracción de características).
- Ingeniería del modelo de ML: entrenamiento y servicio del modelo de ML.
- Ingeniería de código: integración del modelo de ML en el producto final (evaluación, validación y selección).

La siguiente figura muestra los pasos principales de un flujo de trabajo típico de ML.

End-to-End Supervised Machine Learning

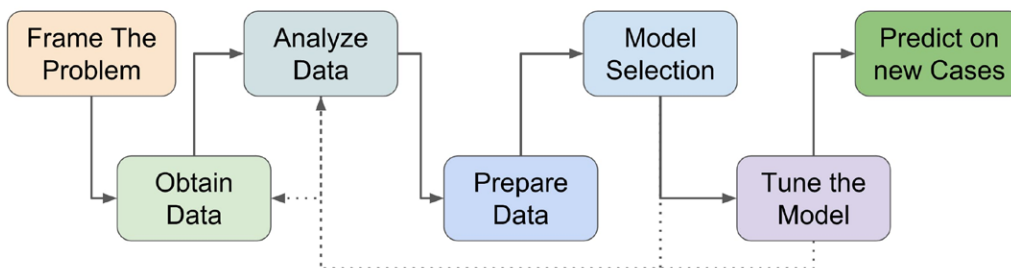


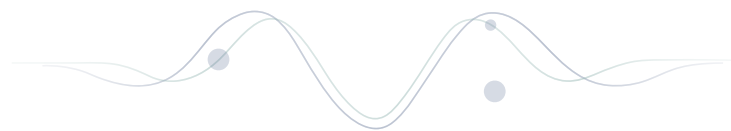
Figura 1. Esquema general de un modelo de aprendizaje automático.

el

2 Actividad de aprendizaje

Para esta actividad final de la asignatura, el objetivo es realizar la implementación de un proceso de aprendizaje automático completo. Identifica un tema de interés en el cual consideres que los sistemas de aprendizaje automático pueden ser utilizados. Para ello realiza:

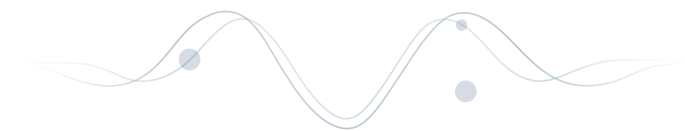
1. Identificar un problema de aplicación de interés para aplicar las técnicas de tratamiento de señales.
2. Identificar el tipo de señal a procesar y la estrategia de aprendizaje (regresión, clasificación y/o agrupamiento).
¿Cuento con las etiquetas para realizar el proceso?
3. Definir el tipo de esquema a abordar (extracción de características de la señal).



3 Entregables

- Notebook en el cual se explique el desarrollo del problema y se realice una evaluación cuantitativa y cualitativa de los resultados, utilizando las métricas de desempeño para cada caso. Recordar realizar una discusión acorde a cada experimento y resultado obtenido.
- Construir un video *pitch* de máximo 8 minutos en el cual evidencien el desarrollo del sistema. Este pitch deberá contener
 - Presentación: portada
 - Motivación: ¿por qué es importante el problema a abordar?
 - Planteamiento del problema: ¿cuál es el problema a resolver?
 - Objetivo general del sistema a desarrollar
 - Materiales y métodos: explicar bases de datos y esquema del modelo de tratamiento de señal a utilizar
 - Resultados y discusiones
 - Conclusiones
 - Trabajos futuros: ideas que consideren relevantes para futuras investigaciones
- La actividad se deberá desarrollar en los equipos de trabajo y se debe construir un repositorio en **GitHub**, en el cual carguen los archivos y demás elementos en el proyecto. Se requiere que existan al menos 3 commits para realizar el seguimiento de las versiones del proyecto. En la siguiente página pueden encontrar los pasos básicos para la gestión del repositorio **Git-guide**.
- A continuación se presenta la rúbrica de evaluación de la actividad:

Criterios a evaluar	Excelente (1.0)	Sobresaliente (0.8)	Aceptable (0.5)	Regular (0.3)	Sin evidencia
Funcionamiento y estructura del Notebook	El funcionamiento del notebook y su estructura son contruidos y funcionan correctamente	El funcionamiento del notebook y su estructura son contruidos y funcionan debidamente	El funcionamiento del notebook y su estructura son contruidos y funcionan parcialmente	El funcionamiento del notebook y su estructura son contruidos y funcionan erróneamente	No se presenta evidencia
Desempeño del algoritmo	El desempeño del algoritmo desarrollado es exacto	El desempeño del algoritmo desarrollado es sobresaliente	El desempeño del algoritmo desarrollado es parcial	El desempeño del algoritmo desarrollado es bajo	No se presenta evidencia
Presentación y discusión apropiada de los resultados	Se presenta y discute con mucha claridad, todos y cada uno de los conceptos e ideas claves del tema	Aparecen recogidos, aunque no con mucha claridad, todos y cada uno de los conceptos e ideas claves del tema	Apenas se esbozan todos y cada uno de los conceptos e ideas claves del tema	Aparecen algunos conceptos e ideas claves del tema	No se presenta evidencia
Conclusiones suficientes	Las conclusiones son suficientes y dan cuenta de los aspectos relevantes del tema	Las conclusiones son abordadas correctamente y discuten los temas	Las conclusiones son abordadas parcialmente y se discuten algunos temas	Las conclusiones no se relacionan con los resultados evidenciados y no dan cuenta de los aspectos del tema	No se presenta evidencia



Nota: en la plataforma de la asignatura se cargará el link del repositorio que contiene la actividad del mini-proyecto.
Fecha máxima de entrega - **Semana 16**.

¡Éxitos!