**Dataset split:**

* **¿Para qué sirve?**
  + Sirve para dividir un conjunto de datos en tres grupos: entrenamiento, pruebas y validación. Cada uno de estos, se utiliza para entrenar, evaluar y ajustar un modelo apropiadamente.
    - **Entrenamiento:** Entrena el modelo.
    - **Prueba:** Evaluar el rendimiento del modelo entrenado, con datos no vistos durante el proceso anterior.
    - **Validación:** Ajustar hiperparámetros probando diferentes combinaciones de valores y, posteriormente, midiendo el rendimiento del modelo. Todo con el fin de encontrar la combinación más optima de estos.
* **¿Qué parámetros tiene?**
  + No se trata de un proceso con parámetros en sí, sino de una estrategia en la que se determina la proporción de datos asignados a cada conjunto.
* **Cuando selecciona un modelo, ¿qué se tiene que hacer?**
  + Se debe dividir el dataset en tres partes: entrenamiento, validación y prueba.
* **Cuando se evalúa un modelo, ¿qué se tiene que hacer?**
  + Para evaluar un modelo después de haberlo entrenado, se debe utilizar el conjunto de datos de prueba que no ha sido visto previamente por el modelo. Calcular métricas de rendimiento, como precisión, exhaustividad, F1-score o la métrica apropiada para tu problema, para determinar qué tan bien se desempeña el modelo en datos no vistos.

**Modelo (creación/selección):**

* **¿Para qué sirve?**
  + La creación o selección del modelo se refiere a la elección de la arquitectura y tipo de modelo que se utilizará para resolver un problema de aprendizaje automático. El objetivo es seleccionar un modelo que se ajuste adecuadamente al problema en cuestión.
* **¿Qué parámetros tiene?**
  + Los parámetros dependen del tipo de modelo seleccionado. Por ejemplo, en redes neuronales, se incluyen parámetros como el número de capas, el número de unidades en cada capa, la función de activación, etc.
* **Cuando selecciona un modelo, ¿qué se tiene que hacer?**
  + Al seleccionar un modelo, se debe tener en cuenta la naturaleza del problema (clasificación, regresión, etc.) y la disponibilidad de datos. Se debe elegir una arquitectura de modelo que sea apropiada para el problema, como redes neuronales, árboles de decisión, máquinas de soporte vectorial, etc.
* **Cuando se evalúa un modelo, ¿qué se tiene que hacer?**
  + Después de seleccionar un modelo y entrenarlo, se debe evaluar su rendimiento utilizando métricas adecuadas. Esto te ayudará a determinar si el modelo es lo suficientemente bueno para resolver el problema o si se necesita ajustar parámetros o probar diferentes modelos.

**Model Compile:**

* **¿Para qué sirve?**
  + La compilación del modelo es el proceso en el que se especifican detalles adicionales sobre cómo se entrenará el modelo, como el optimizador, la función de pérdida y las métricas de evaluación. Sirve para configurar la forma en que el modelo aprenderá de los datos.
* **¿Qué parámetros tiene?**
  + Algunos de los parámetros comunes incluyen el optimizador (como el descenso de gradiente estocástico), la función de pérdida (como el error cuadrático medio) y las métricas (como precisión o precisión/recuperación en clasificación).
* **Cuando selecciona un modelo, ¿qué se tiene que hacer?**
  + Al compilar un modelo, se debe definir la función de pérdida apropiada para tu problema, el optimizador que controlará cómo se actualizan los pesos del modelo durante el entrenamiento y las métricas que se utilizarán para evaluar el rendimiento durante el entrenamiento.
* **Cuando se evalúa un modelo, ¿qué se tiene que hacer?**
  + Durante la evaluación del modelo, no se realiza ninguna acción relacionada con la compilación. La compilación es una fase previa al entrenamiento.

**Model Training:**

* **¿Para qué sirve?**
  + El entrenamiento del modelo implica alimentarlo con datos de entrenamiento y ajustar sus parámetros para minimizar la función de pérdida. Sirve para que el modelo aprenda a hacer predicciones precisas.
* **¿Qué parámetros tiene?**
  + Los parámetros de entrenamiento incluyen el número de épocas (iteraciones), el tamaño del lote (batch size), y otros hiperparámetros relacionados con la regularización y la tasa de aprendizaje.
* **Cuando selecciona un modelo, ¿qué se tiene que hacer?**
  + Al entrenar un modelo, se debe proporcionar el conjunto de datos de entrenamiento y los parámetros de entrenamiento, como el número de épocas y el tamaño del lote. Además, debes supervisar el proceso de entrenamiento para evitar el sobreajuste o subajuste.
* **Cuando se evalúa un modelo, ¿qué se tiene que hacer?**
  + Durante la evaluación del modelo, se deben calcular las métricas de rendimiento utilizando el conjunto de datos de prueba. Esto implica hacer predicciones con el modelo y compararlas con las etiquetas reales para medir su rendimiento.

**Model Persistence:**

* **¿Para qué sirve?**
  + La persistencia del modelo se refiere a guardar el modelo entrenado en disco para poder utilizarlo en el futuro sin tener que volver a entrenarlo desde cero.
* **¿Qué parámetros tiene?**
  + No se trata de parámetros, sino de acciones para guardar y cargar modelos, que pueden variar según la biblioteca o el entorno de desarrollo que estés utilizando.
* **Cuando selecciona un modelo, ¿qué se tiene que hacer?**
  + La persistencia del modelo implica guardar el modelo entrenado en disco una vez que esté satisfecho con su rendimiento. Esto se hace utilizando métodos o funciones proporcionados por la biblioteca de aprendizaje automático que estés utilizando.
* **Cuando se evalúa un modelo, ¿qué se tiene que hacer?**
  + La evaluación del modelo y la persistencia del modelo son acciones separadas. La persistencia ocurre después de que hayas evaluado y entrenado un modelo y decidas que es lo suficientemente bueno para su uso futuro.

**Model Evaluation:**

* **¿Para qué sirve?**
  + La evaluación del modelo se realiza para medir su rendimiento en un conjunto de datos de prueba independiente. Sirve para determinar cuán bien el modelo generaliza a datos no vistos.
* **¿Qué parámetros tiene?**
  + Los parámetros de evaluación incluyen métricas de rendimiento como la precisión, la exhaustividad, el F1-score, la matriz de confusión, etc.
* **Cuando selecciona un modelo, ¿qué se tiene que hacer?**
  + Antes de seleccionar un modelo, se debe tener en cuenta cómo evaluarás su rendimiento en función del problema. Esto ayudará a elegir las métricas adecuadas que se utilizarán durante la evaluación.
* **Cuando se evalúa un modelo, ¿qué se tiene que hacer?**
  + Durante la evaluación del modelo, se debe calcular las métricas de rendimiento previamente seleccionadas utilizando el conjunto de datos de prueba. Luego, se debe analizar estas métricas para comprender cuán bien se desempeña el modelo y si cumple con los objetivos del proyecto.

**Model Predict:**

* **¿Para qué sirve?**
  + La predicción del modelo se utiliza para hacer inferencias en nuevos datos una vez que el modelo ha sido entrenado y evaluado. Sirve para obtener predicciones basadas en el aprendizaje del modelo.
* **¿Qué parámetros tiene?**
  + Los parámetros dependen de la función de predicción específica de cada modelo, como las entradas de datos de prueba.
* **Cuando selecciona un modelo, ¿qué se tiene que hacer?**
  + La predicción en sí no está relacionada con la selección del modelo. Sin embargo, se debe tener en cuenta la capacidad de hacer predicciones una vez que hayas seleccionado y entrenado el modelo.
* **Cuando se evalúa un modelo, ¿qué se tiene que hacer?**
  + La evaluación del modelo se realiza utilizando datos de prueba para medir su rendimiento general. La predicción se utiliza después de haber evaluado y ajustado el modelo para hacer inferencias en nuevos datos no vistos.