**Ejercicio: Predicción de abandono escolar**

**🔹 1. Tipo de aprendizaje**

* **Aprendizaje supervisado**
  + Porque se cuenta con datos históricos donde se conoce si un estudiante abandonó o no.

**🔹 2. Tipo de tarea**

* **Clasificación**
  + El objetivo es predecir si un estudiante **abandonará (sí/no)**, lo que implica una categoría.

**🔹 3. Datos necesarios y posibles fuentes**

| **Tipo de dato** | **Fuente posible** |
| --- | --- |
| Historial académico | Sistemas internos de gestión escolar |
| Asistencia | Registros de clases |
| Datos socioeconómicos | Encuestas, bases gubernamentales |
| Participación en actividades | Plataformas educativas |
| Edad, género, ubicación | Datos demográficos institucionales |

**🔹 4. Etapas a seguir**

1. **EDA (Análisis exploratorio de datos)**
   * Identificar patrones, valores faltantes, correlaciones.
2. **División de datos**
   * Separar en conjunto de entrenamiento y prueba (por ejemplo, 80/20).
3. **Entrenamiento del modelo**
   * Usar algoritmos de clasificación para aprender de los datos.
4. **Evaluación del modelo**
   * Medir precisión, recall, F1-score, matriz de confusión.

**🔹 5. Algoritmos candidatos**

| **Algoritmo** | **Justificación** |
| --- | --- |
| **Random Forest** | Robusto ante datos ruidosos y desequilibrados. |
| **Logistic Regression** | Simple, interpretable, buena base inicial. |
| **Gradient Boosting (XGBoost)** | Alta precisión, útil en competencias. |
| **SVM (Support Vector Machine)** | Eficiente en espacios de alta dimensión. |