***Two versions:***

1.English:

# ****PROJECT REPORT: MILLING FAILURE PREDICTION****

## ****1. Introduction****

This project aims to develop a Machine Learning model to predict failures in milling machines. The goal is to identify both the occurrence of failures and the specific type of failure, improving process efficiency and reducing costs associated with downtime and tool damage.

## ****2. Data Description****

The dataset consists of records collected every 3 minutes, including various machine operational parameters such as:

* Rotational speed
* Torque
* Power
* Tool wear
* Environmental conditions
* Tool type
* Failure labels (Machine Failure and its subtypes: TWF, HDF, PWF, OSF)

The dataset includes approximately 400 failure occurrences from a significantly larger total.

## ****3. Modeling Strategy****

Two prediction approaches were developed:

1. **General Failure Prediction (Machine Failure)**
   * Class balancing was applied using **oversampling** and **undersampling**.
   * Different feature selection techniques were evaluated, with **SelectKBest** performing best.
   * Models with and without dimensionality reduction via **PCA** were tested, with better results obtained without PCA.
   * The final model was chosen based on **ROC-AUC** and **Recall**, prioritizing failure detection.
2. **Failure Type Prediction**
   * Data was filtered to include only instances where a failure occurred.
   * A multi-label classification strategy was implemented to predict different failure modes.
   * **Keras** neural networks were used to handle this problem.

## ****4. Model Evaluation****

* Various models were compared using key metrics such as **precision, recall, and ROC-AUC**.
* The selected model for general failure prediction provided a good balance between **recall and precision**, favoring the reduction of undetected failures.
* For failure type classification, a **probability-to-binary label conversion** strategy was applied to enhance result interpretability.

## ****5. Streamlit Implementation****

To facilitate model interaction, a **Streamlit** application was developed, allowing users to upload data and obtain real-time predictions.

## ****6. Conclusion****

The developed models enable failure anticipation in the milling process with an acceptable level of accuracy. The importance of class balancing and feature selection in improving predictive capabilities is highlighted. The developed application facilitates model integration into a production environment, aiding better decision-making for preventive maintenance.

2.Español:

**MEMORIA DEL PROYECTO: PREDICCIÓN DE FALLOS EN PROCESOS DE MILLING**

**1. Introducción**

Este proyecto tiene como objetivo desarrollar un modelo de Machine Learning para predecir fallos en máquinas de fresado (milling). Se busca identificar tanto la ocurrencia de fallos como el tipo de fallo que se produce, con el fin de mejorar la eficiencia del proceso y reducir costos asociados a paradas y daños en herramientas.

**2. Descripción de los Datos**

El dataset utilizado contiene registros cada 3 minutos de distintas variables operativas de la máquina, incluyendo:

* Velocidad de rotación
* Torque
* Potencia
* Desgaste de la herramienta
* Parámetros ambientales
* Tipo de herramienta
* Etiquetas de fallo (Machine Failure y sus subtipos: TWF, HDF, PWF, OSF)

Se cuenta con aproximadamente 400 registros en los que ocurrió una falla, de un total 10.000 registros.

**3. Estrategia de Modelado**

Se desarrollaron dos enfoques de predicción:

1. **Predicción de fallo general (Machine Failure)**
   * Se aplicó balanceo de clases mediante **oversampling** y **undersampling**.
   * Se evaluaron distintas técnicas de selección de características, destacando **SelectKBest**.
   * Se probaron modelos con y sin reducción de dimensionalidad mediante **PCA**, obteniendo mejores resultados sin PCA.
   * El modelo final fue elegido con base en **ROC-AUC** y **Recall**, priorizando la detección de fallos.
2. **Predicción del tipo de fallo**
   * Se filtraron los datos para considerar únicamente registros donde ocurrió una falla.
   * Se aplicó una estrategia de clasificación multi-etiqueta para predecir los diferentes modos de fallo.
   * Se utilizaron redes neuronales en **Keras** para abordar este problema.

**4. Evaluación de Resultados**

* Se compararon distintos modelos mediante métricas clave como **precision, recall y ROC-AUC**.
* El modelo seleccionado para la predicción de fallos generales mostró un buen equilibrio entre **recall y precisión**, favoreciendo la reducción de fallos no detectados.
* Para la clasificación de tipo de fallo, se aplicó una estrategia de **conversión de probabilidades en etiquetas binarias** para mejorar la interpretabilidad de los resultados.

**5. Implementación en Streamlit**

Para facilitar la interacción con el modelo, se desarrolló una aplicación en **Streamlit**, permitiendo cargar datos y obtener predicciones en tiempo real.

**6. Conclusión**

Los modelos desarrollados permiten anticipar fallos en el proceso de milling con un nivel de precisión aceptable. Se destaca la importancia del balanceo de clases y la selección de características para mejorar la capacidad predictiva. La aplicación desarrollada facilita la integración del modelo en un entorno productivo, permitiendo una mejor toma de decisiones en mantenimiento preventivo.