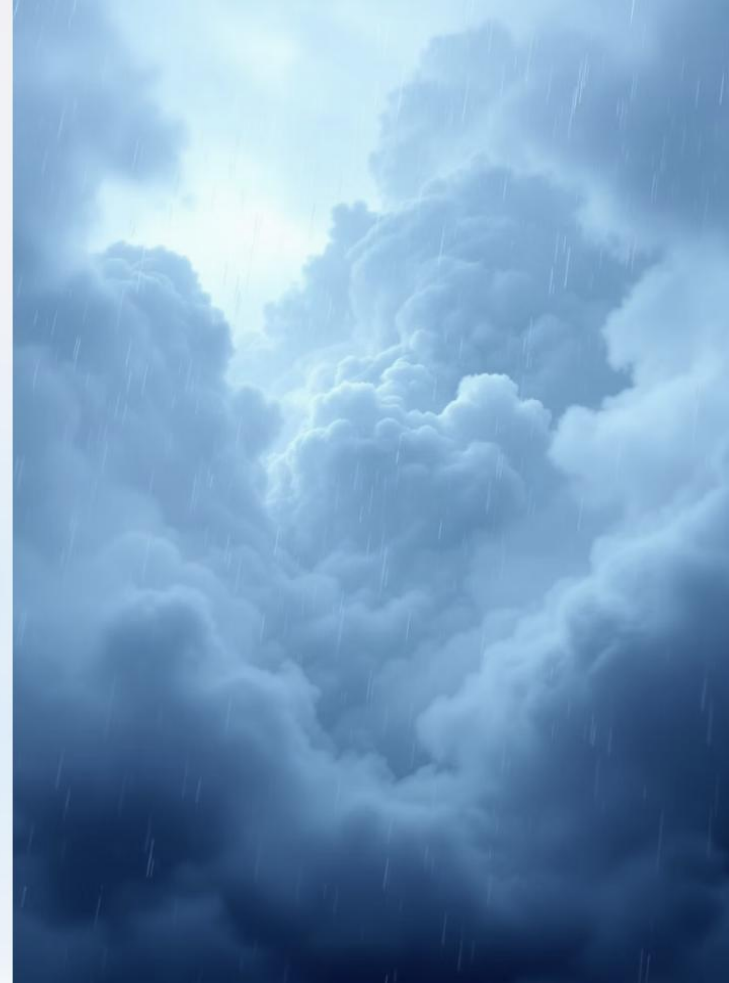


Predicción de Lluvia en Bilbao con Machine Learning

Proyecto académico para predecir lluvia en Bilbao usando variables meteorológicas históricas.

Alex Moura



Aprendizaje del Proyecto



Estructura End-to-End

Proyecto ML completo de inicio a fin



Análisis Exploratorio

Visualizaciones efectivas de datos



Preprocesamiento

Manejo de nulos y variables categóricas



Modelos y Métricas

Clasificación binaria e interpretación

Machine Learning Workflow

Putse are too with machine learning in the place.

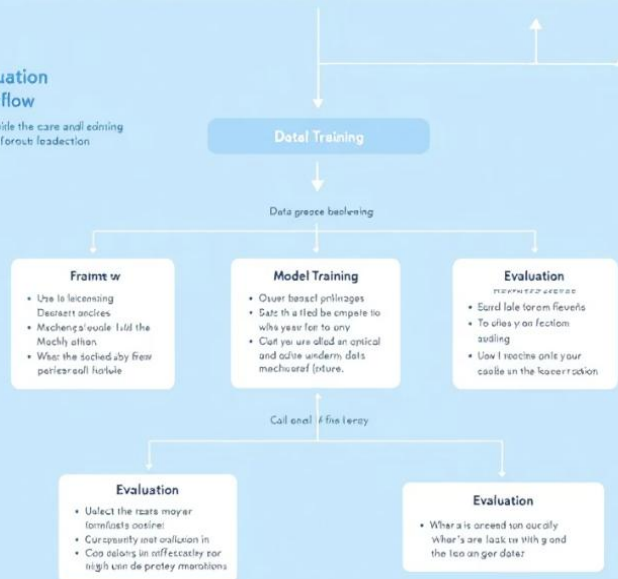
Data 1

Fetempix an day Data
Planning Workflow



Evaluation Morkflow

Hugh exite the care and eding for this forst leadecton



Distribución de Variables Meteorológicas

Temperatura

Distribución normal entre
15-20°C

Humedad

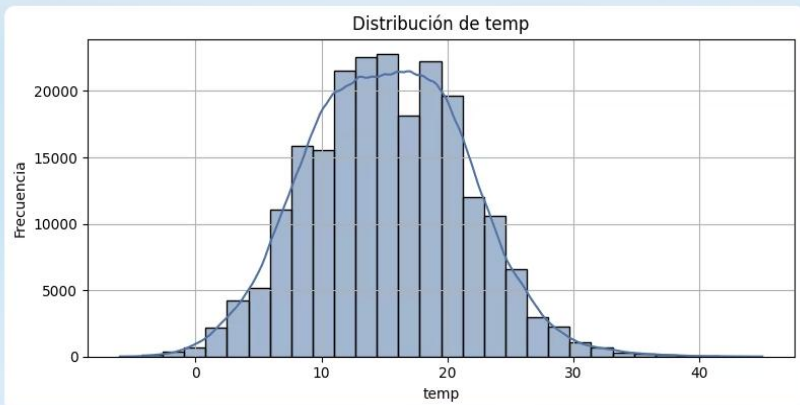
Alta concentración entre
70-100%

Velocidad Viento

Mayoría entre 0-10 km/h

Presión

Normal entre 1000-1030
hPa



Comparación por Clase de Lluvia

Temperatura

Más baja cuando llueve

Humedad

Significativamente más alta con lluvia

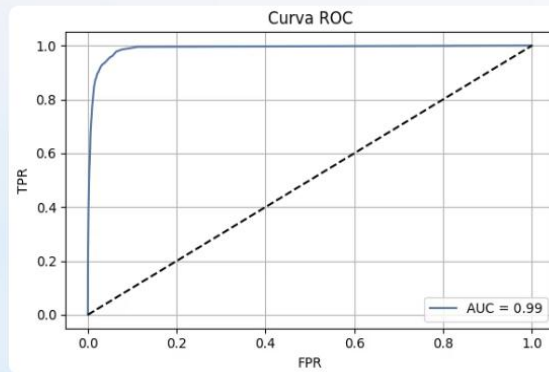
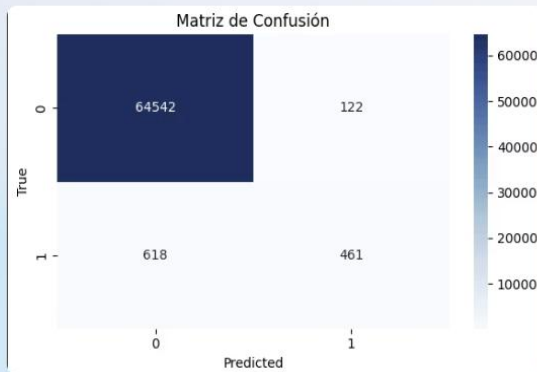
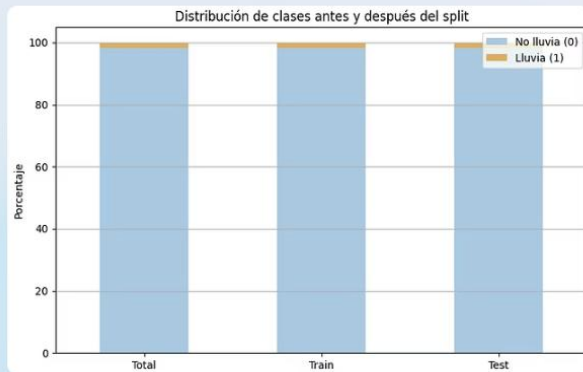
Viento

Ligeramente mayor en horas lluviosas

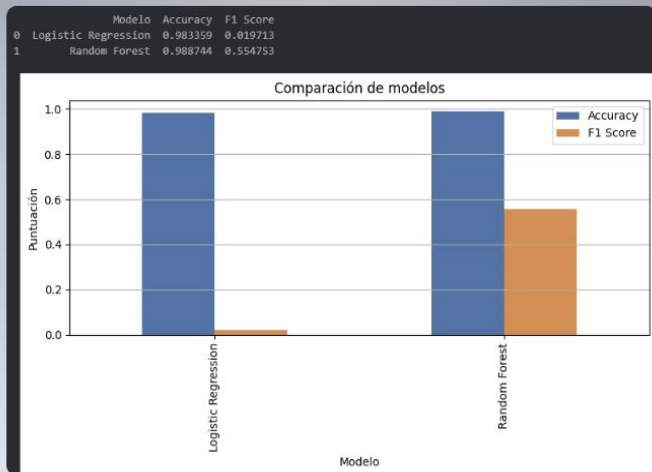
Presión

Tiende a ser más baja con lluvia

Balance y Comparación de Modelos



Dataset equilibrado entre clases. Random Forest logra mejor F1-score para lluvia. Matriz muestra necesidad de ajustar sensibilidad.



Evaluación del Modelo

Curva ROC

Excelente discriminación ($AUC \approx 0.99$)

Análisis de Threshold

Mejor equilibrio cerca de 0.3

Evaluación Final

F1-score de 0.66, recall ≈ 0.70

Predicción con Nuevos Datos

```
# Nuevo ejemplo: 14 de febrero de 2025, 15:00, 85% humedad, 7.2 km/h viento, 1012.5 hPa, 11.3 °C
nuevo_registro = {
    'day': 14,
    'month': 2,
    'year': 2025,
    'hora': 15,
    'rhum': 85.0,
    'wspd': 7.2,
    'pres': 1012.5,
    'temp': 11.3
}
```

```
# Predicción
pred, proba = predict_new(model.models['Random Forest'], features, nuevo_registro, threshold=0.3)

print(f"¿Lloverá? {'Sí 🌧️' if pred == 1 else 'No ☀️'}")
print(f"Probabilidad de lluvia: {proba:.2%}")
```

✓ 0.0s

¿Lloverá? Sí 🌧️
Probabilidad de lluvia: 38.00%



Entrada de Datos

Variables meteorológicas actuales



Procesamiento

Preprocesamiento y aplicación del modelo



Resultado

Predicción y probabilidad de lluvia

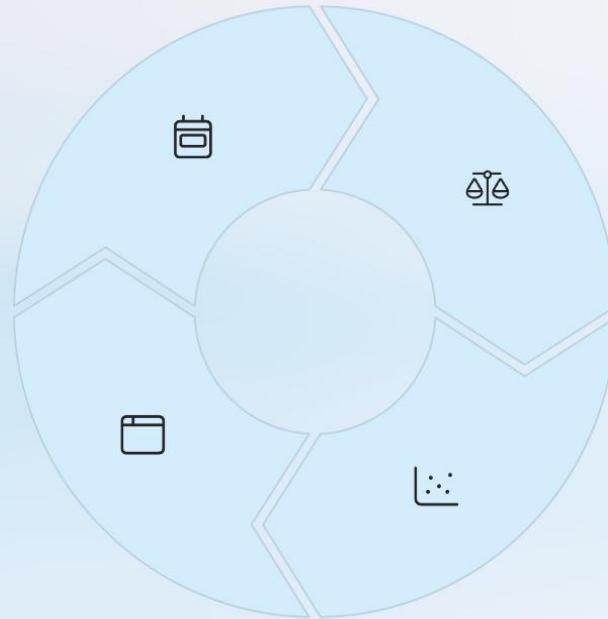
Conclusión

Datos Reales

Dataset complejo con limpieza y transformación

Aplicación Práctica

Sistema de predicción implementado



Balance Óptimo

Equilibrio entre precisión y sensibilidad

Alto Rendimiento

$AUC \approx 0.99$, $F1\text{-score} \approx 0.66$