

# Detección de Fallos Utilizando Deep Learning

## Análisis de Datos de Vibraciones para un Diagnóstico Preciso de Fallos

Ávalos, Silvio  
Orzusa, Emanuel  
Parodi, Giulio

6 de junio de 2024

# Introducción

- ▶ Importancia de la detección de fallos en entornos industriales.
- ▶ Breve introducción al conjunto de datos: 39 escenarios con 25 pruebas cada uno, total 975 archivos CSV por cada RPM.
- ▶ Conjunto de datos de escenarios de fallas simples y dobles utilizando señales de vibración de una máquina rotativa

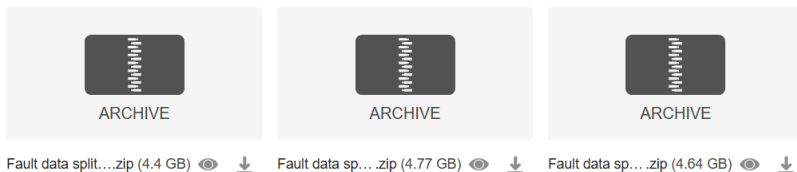


Figura: Dataset original sin descomprimir.

# Introducción

- Origen de los datos: Un simulador de averías está compuesto por un motor que hace girar un eje apoyado entre dos rodamientos. Además, una polea y una correa están unidas al eje y se utilizan para accionar una caja de engranajes. Un tacómetro medía las rpm del eje. Los acelerómetros piezoeléctricos se fijaron en 3 direcciones perpendiculares a las carcasas de los cojinetes y en dos direcciones en la carcasa de la caja de cambios.

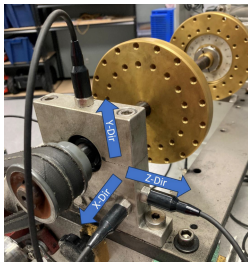


Figura: Simulador de fallas

# Introducción

- Descripción de la recopilación de datos: Para recopilar datos para un escenario de fallo, se reemplazó un componente nominal en el simulador de fallas de la máquina con un componente defectuoso especificado. En el caso de escenarios de dos fallos, se reemplazaron dos componentes nominales por componentes defectuosos específicos conocidos. En el caso del escenario nominal, no se sustituyó ningún componente. El motor se encendió y se ajustó a uno de los tres valores de rpm para cada escenario. Después de alcanzar una velocidad de funcionamiento constante, la recopilación de datos se realizó desde sensores durante 10 segundos. Después de la recopilación de datos, el simulador de la máquina se apagó y los componentes defectuosos se reemplazaron por los componentes nominales o por el siguiente componente defectuoso que se probaría.

# Introducción

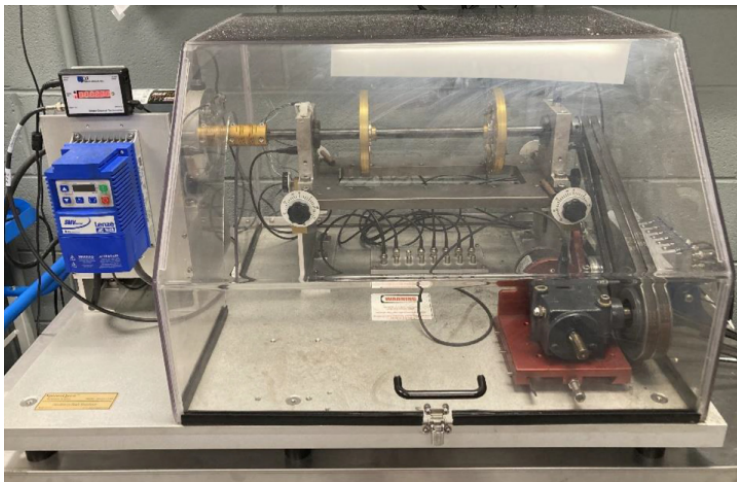


Figura: Simulador de fallas

# Planteamiento del Problema

- ▶ Descripción del problema: Detección temprana de fallos en maquinaria industrial.
- ▶ Pregunta clave: ¿Cómo podemos aprovechar el deep learning para detectar fallos con precisión en maquinaria basada en datos de vibración?



**Figura:** Detección temprana de fallas en máquinas rotativas mediante deep learning en un entorno industrial.

## Descripción de los Datos

- Visualización de muestra de datos: cada uno de los archivos contaba con 64,000 datos tomados por un acelerómetro en un periodo de 10 segundos.

```

# Muestra las primeras filas
data.head()

python-input-3-b6139d99584:4: DtypeWarning: Columns (0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17) have mixed types. Specify dtype option on import or set low_memory=False.
data = pd.read_csv('content/drive/myDrive/Fault data split 25% Fault Trial 1.csv')

Timestamp      7/18/22      7/18/22      7/18/22      7/18/22      7/18/22      7/18/22      7/18/22      7/18/22      7/18/22      7/18/22      7/18/22      7/18/22
              14:09           14:09           14:09           14:09           14:09           14:09           14:09           14:09           14:09           14:09           14:09           14:09
0 Interval  0.00015625 Interval  0.00015625 Interval  0.00015625 Interval  0.00015625 Interval  0.00015625 Interval  0.00015625 Interval  0.00015625 Interval  0.00015625 Interval  0.00015625 Interval  0.00015625 Interval  0.00015625
1 Channel name Tachometer Channel name Motor Channel name Bearing 1 Z Channel name Bearing 1 Y Channel name Bearing 1 X Channel name Bearing 2 Z Channel name Bearing 2 Y Channel name Bearing 2 X Channel name Gearbox
2 Unit V Unit g Unit g Unit g Unit g Unit g Unit g Unit g Unit g Unit g Unit g
3 0 0.862248258 0 0.006744791 0 0.001761446 0 0.004679333 0 0.013791326 0 0.003338325 0 0.022431522 0 0.011058366 0 0.008016799
4 0.00015625 0.862027298 0.00015625 0.004383926 0.00015625 0.003282426 0.00015625 0.004445694 0.00015625 0.011916785 0.00015625 0.00992303 0.00015625 0.020547591 0.00015625 0.013693532 0.00015625 0.009252525

```

**Figura:** Visualización de los primeros datos de uno de los archivos originales.

# Preprocesamiento de Datos

- ▶ Pasos realizados para el preprocesamiento de datos:
  - ▶ Eliminación de las tres primeras filas.
  - ▶ Eliminación de columnas innecesarias (3, 5, 7, 9, 11, 13, 15, 17).
  - ▶ Renombrar columnas a: Tiempo, Tacómetro, Motor, Rodamiento 1 Z, Rodamiento 1 Y, Rodamiento 1 X, Rodamiento 2 Z, Rodamiento 2 Y, Rodamiento 2 X, Caja de engranajes.
- ▶ Herramientas utilizadas para el preprocesamiento.

```
df.head()
```

	Time	Tachometer	Motor	Bearing 1 Z	Bearing 1 Y	Bearing 1 X	Bearing 2 Z	Bearing 2 Y	Bearing 2 X	Gearbox
0	0	0.0013176483849999992	0.0082665870495049399	0.021089014198019798	-0.0023447754257425752	-0.015215976540000016	0.0047812919603960335	-0.024906489900990106	0.018778289811881191	0.0117598404899999979
1	0.00015625	0.0010742086099999997	0.0077682636633663177	0.014745417039603959	-0.0021224366336633667	-0.017427934330000013	0.0057829561386138553	-0.029819486198019808	0.012988346722772277	0.0097760287599999857
2	0.0003125	0.0010779520849999994	0.0039178613663366224	0.012593787594059404	-0.0067174383366336646	-0.0068180351000000153	-0.0026755413663366381	-0.02590386508910892	0.018332909574257428	0.0113106755699999984
3	0.00046875	0.0011145366654999996	0.0070648247821782084	0.015969619999999997	-0.0040864292970297037	-0.012179220930000012	0.0019246941188118746	-0.028526592435643573	0.018592714712871286	0.0125833095099999988
4	0.000625	0.0011828192049999986	0.0077682636633663177	0.01318734054454452	-0.0070138900594059412	-0.013228963610000012	0.0058200548118811854	-0.020141252891089118	0.006567448297029703	0.0061827093999999986

**Figura:** Visualización de primeros datos de uno de los archivos optimizados.



# Visualización de datos

## ► Series temporales

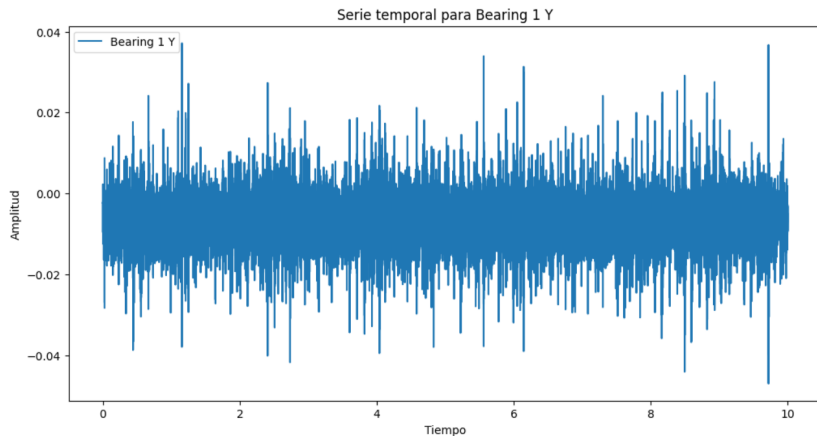


Figura: Serie temporal para cojinete en una de las direcciones

# Visualización de datos

- Dominio de frecuencias del cojinete en uno de los escenarios sin falla.

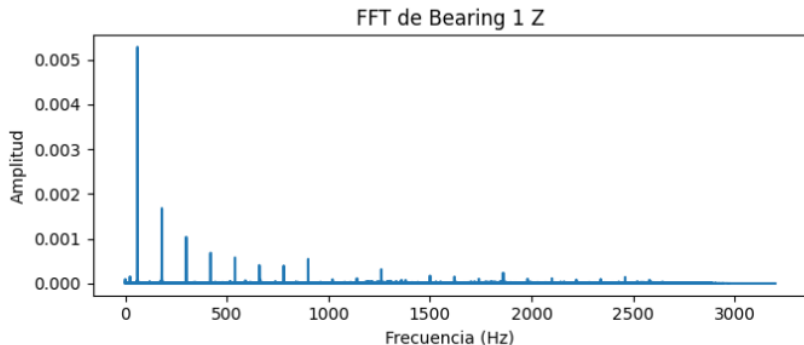


Figura: Dominio de frecuencia mediante FFT

# Visualización de datos

- Dominio de frecuencias del cojinete en uno de los escenarios con falla.

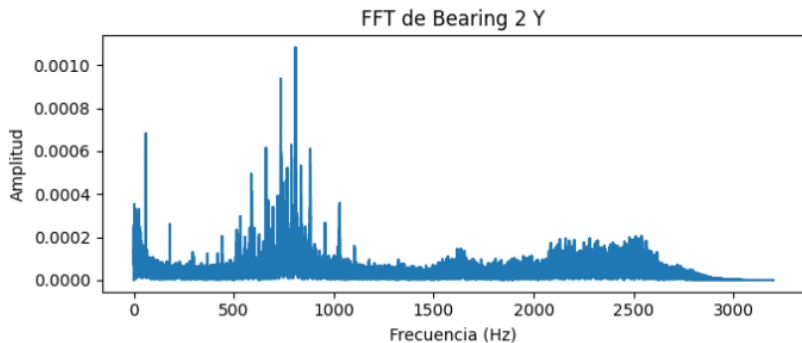


Figura: Dominio de frecuencia mediante FFT

# Extracción de Características

- ▶ Explicación del proceso de extracción de características.
- ▶ Características clave consideradas para el análisis:  
Características en el dominio del tiempo, características en el dominio de la frecuencia.
- ▶ Importancia de cada característica en la detección de fallos.

`df_features.head()`

	mean	std	max	min	ptp	kurtosis	skewness	rms	crest_factor	fft_mean	...	harmonic_3_amplitude	harmonic_4_amplitude	harmonic_5_amplitude	envelope_mean	envelope_std	wavelet_mean	wavelet_std	column	File	label
0	0.000018	0.000432	0.002020	-0.000791	0.003371	0.491288	0.380136	0.000674	3.737075	3.382385e-07	...	0.000518	0.000518	0.000518	0.000724	0.000342	8.358660e-09	0.000790	Tachometer	(Fault) Bearing (1) Ball & Bearing (2) Combina...	1
1	-0.001712	0.004595	0.012692	-0.019407	0.032099	-0.013154	0.008711	0.004904	3.957529	4.663922e-06	...	0.000918	0.000006	0.000343	0.000217	0.002352	8.508267e-08	0.001379	Motor	(Fault) Bearing (1) Ball & Bearing (2) Combina...	1
2	0.006132	0.005580	0.060412	-0.041679	0.102091	1.040749	-0.003990	0.008290	7.287053	1.039403e-05	...	0.006132	0.006132	0.006132	0.009382	0.004380	9.672076e-08	0.003565	Bearing 1 Z	(Fault) Bearing (1) Ball & Bearing (2) Combina...	1
3	-0.006815	0.004025	0.037357	-0.047035	0.084192	4.766746	-0.026289	0.008181	5.749631	1.212229e-05	...	0.006815	0.006815	0.006815	0.008383	0.004139	-9.416609e-08	0.004598	Bearing 1 Y	(Fault) Bearing (1) Ball & Bearing (2) Combina...	1
4	-0.015174	0.006313	0.060778	-0.076701	0.137479	6.155039	-0.001890	0.016434	4.667125	1.535959e-05	...	0.015174	0.015174	0.015174	0.016581	0.005916	8.284537e-10	0.007329	Bearing 1 X	(Fault) Bearing (1) Ball & Bearing (2) Combina...	1

**Figura:** Extracción de features de los 975 archivos para un régimen de 25 RPM.

# Modelo Preliminar CNN

- Resumen de las capas y parámetros del modelo.

Layer (type)	Output Shape	Param #
conv1d_4 (Conv1D)	(None, 98, 64)	256
max_pooling1d_4 (MaxPooling1D)	(None, 49, 64)	0
conv1d_5 (Conv1D)	(None, 47, 128)	24704
max_pooling1d_5 (MaxPooling1D)	(None, 23, 128)	0
flatten_2 (Flatten)	(None, 2944)	0
dense_4 (Dense)	(None, 100)	294500
dense_5 (Dense)	(None, 1)	101
Total params: 319561 (1.22 MB)		
Trainable params: 319561 (1.22 MB)		
Non-trainable params: 0 (0.00 Byte)		

Figura: Estructura del modelo en formato de texto

# Modelo Preliminar CNN

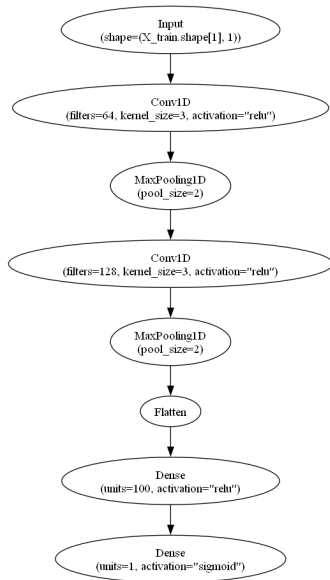
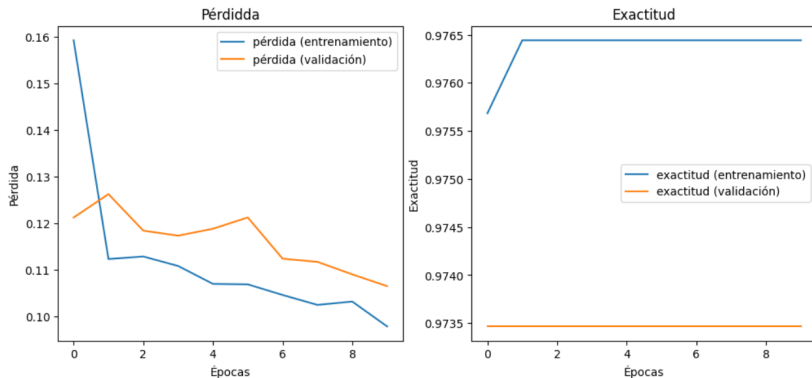


Figura: Estructura del modelo

# Entrenamiento y Evaluación del Modelo

- ▶ Primeramente el modelo fue entrenado con todos los escenarios (38 de fallas y 1 solo de no falla).
- ▶ Resultados preliminares y métricas de rendimiento.



**Figura:** Pérdida y exactitud vs. cantidad de épocas, para los conjuntos de validación y entrenamiento.

# Entrenamiento y Evaluación del Modelo

- Resultados preliminares y métricas de rendimiento.

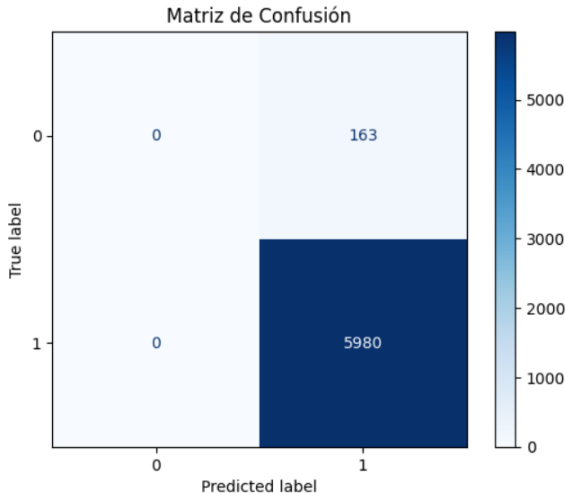
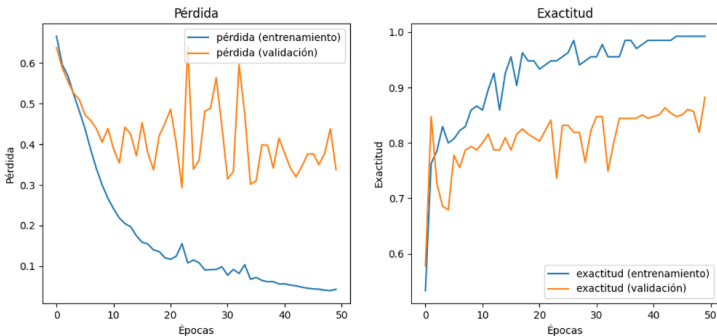


Figura: Matriz de confusión con todos los escenarios.



# Entrenamiento y Evaluación del Modelo

- El modelo estaba claramente sobreajustado, por lo que balanceamos la cantidad de datos comparando 1 falla vs. 1 no falla.



**Figura:** Pérdida y exactitud vs. cantidad de épocas, para los conjuntos de validación y entrenamiento 25 RPM.

# Entrenamiento y Evaluación del Modelo

- Resultados preliminares y métricas de rendimiento.

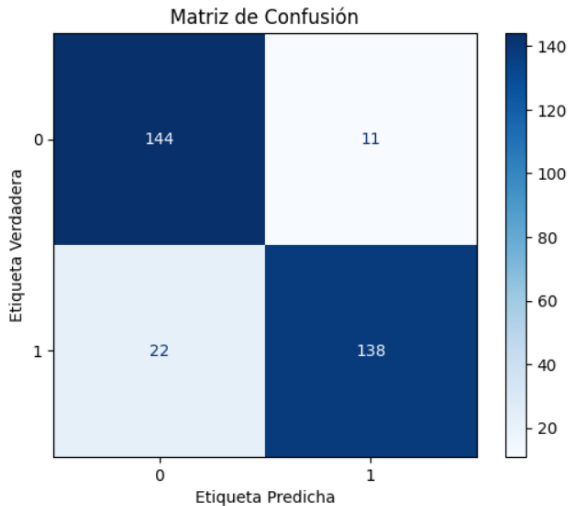


Figura: Matriz de confusión 25 RPM.

# Entrenamiento y Evaluación del Modelo

- Resultados preliminares y métricas de rendimiento.

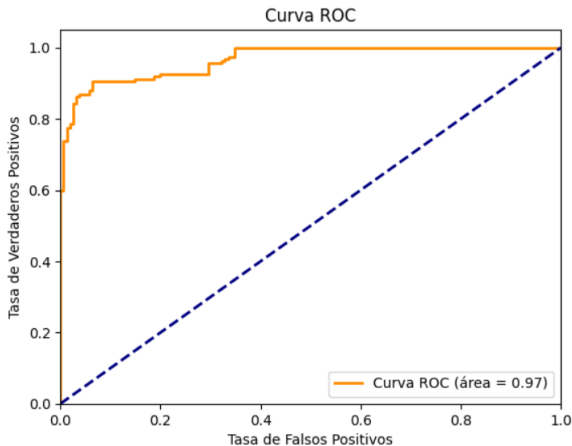


Figura: Curva ROC 25 RPM.

# Entrenamiento y Evaluación del Modelo

- Resultados preliminares y métricas de rendimiento.

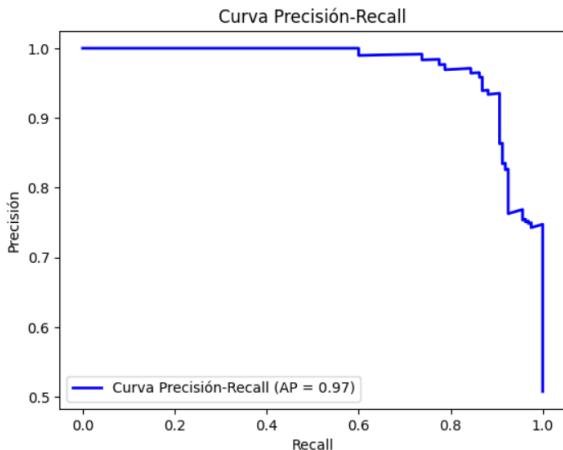
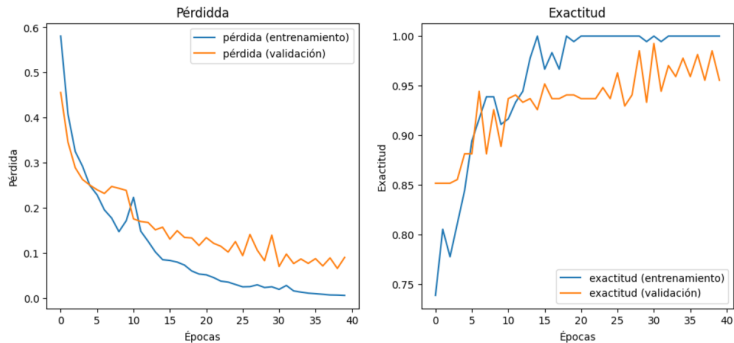


Figura: Curva de Precisión - Recall 25 RPM.

# Entrenamiento y Evaluación del Modelo

- El modelo estaba claramente sobreajustado, por lo que balanceamos la cantidad de datos comparando 1 falla vs. 1 no falla.



**Figura:** Pérdida y exactitud vs. cantidad de épocas, para los conjuntos de validación y entrenamiento 50 RPM.

# Entrenamiento y Evaluación del Modelo

- Resultados preliminares y métricas de rendimiento.

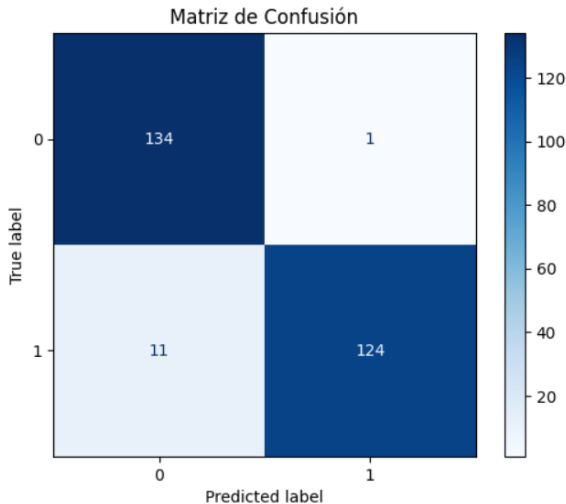


Figura: Matriz de confusión 50 RPM.

# Entrenamiento y Evaluación del Modelo

- Resultados preliminares y métricas de rendimiento.

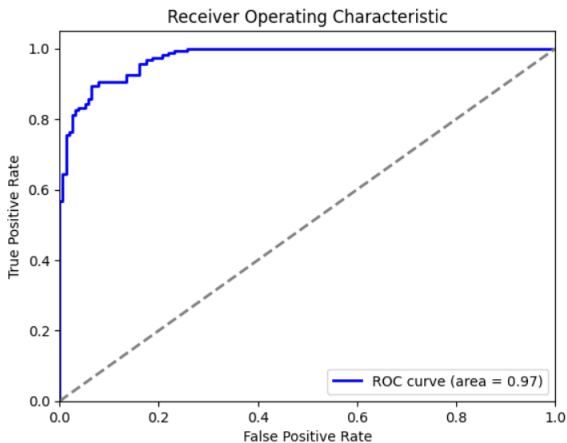


Figura: Curva ROC 50 RPM.

# Entrenamiento y Evaluación del Modelo

- Resultados preliminares y métricas de rendimiento.

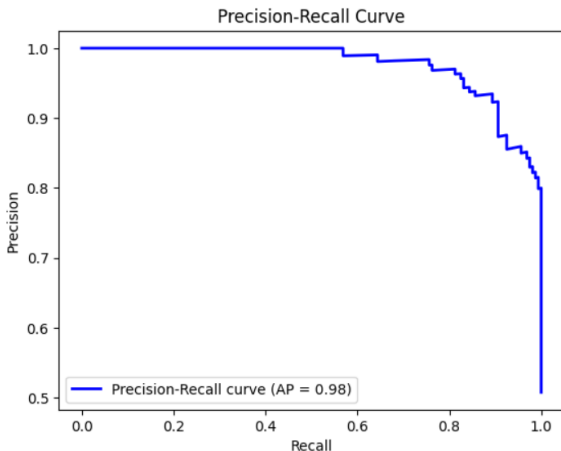
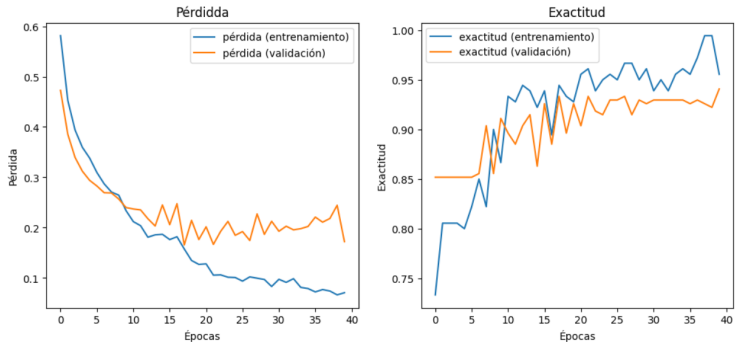


Figura: Curva de Precisión - Recall 50 RPM.



# Entrenamiento y Evaluación del Modelo

- El modelo estaba claramente sobreajustado, por lo que balanceamos la cantidad de datos comparando 1 falla vs. 1 no falla.



**Figura:** Pérdida y exactitud vs. cantidad de épocas, para los conjuntos de validación y entrenamiento 75 RPM.

# Entrenamiento y Evaluación del Modelo

- Resultados preliminares y métricas de rendimiento.

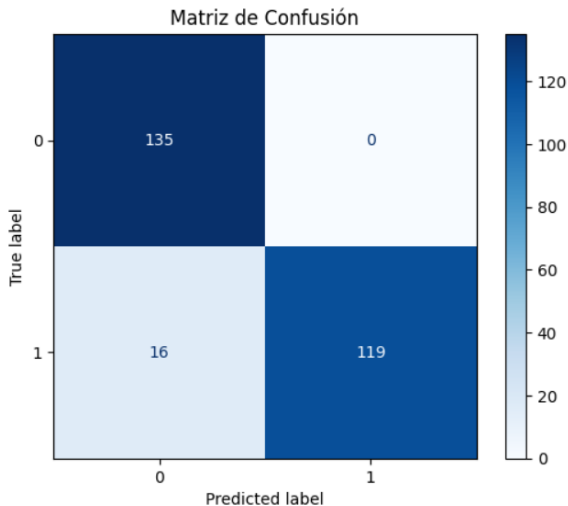


Figura: Matriz de confusión 75 RPM.

# Conclusiones

- ▶ Resumen de los hallazgos.
- ▶ Eficacia del modelo preliminar.
- ▶ Trabajo futuro: Mejoras, pruebas adicionales, posible implementación.

# Preguntas y Respuestas

- ▶ Espacio para preguntas y discusión adicional.