# Entregable 2

Alex Pérez

30 de octubre de 2024

# Título del Proyecto

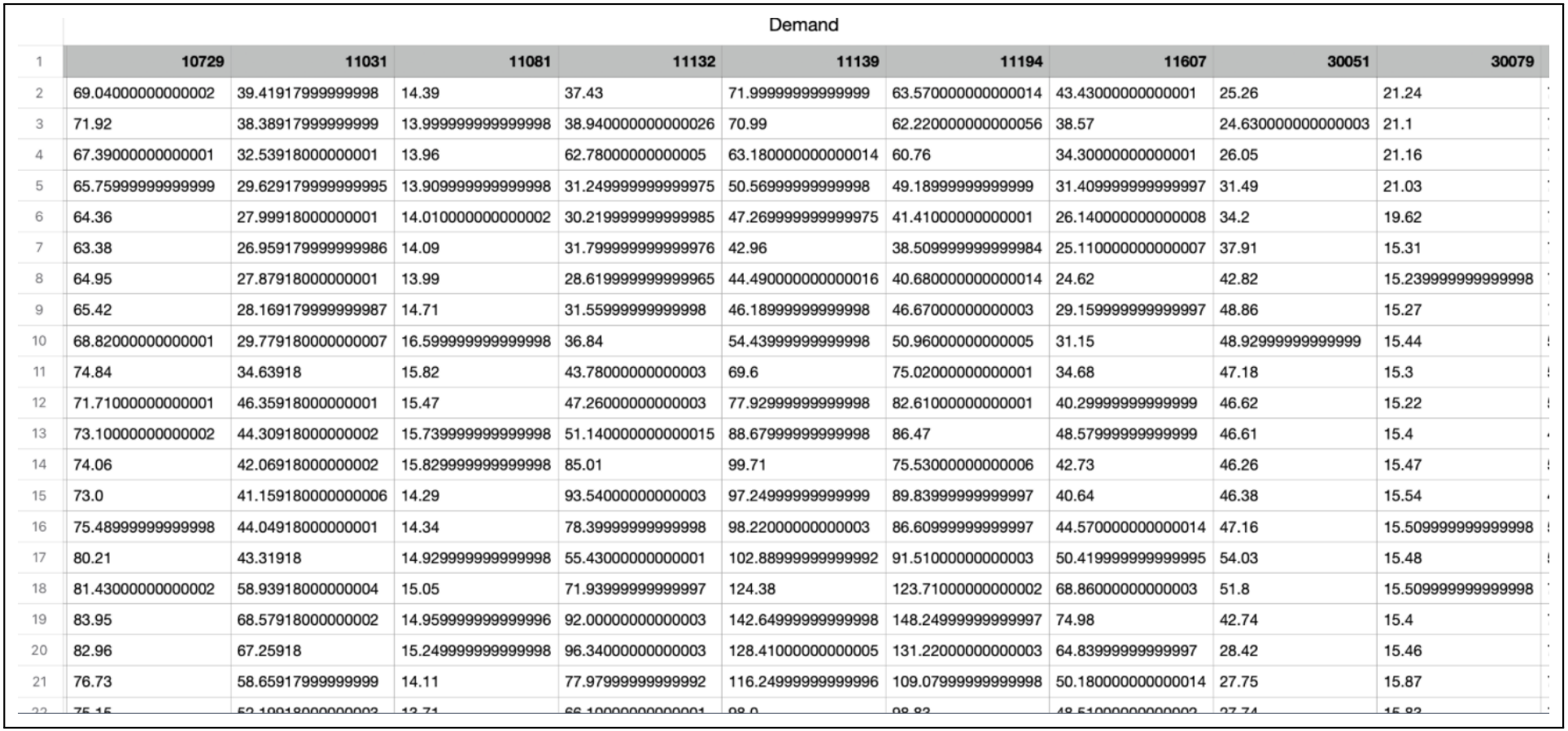
**Machine Learning Applied for Cybersecurity of Energy Management Systems**

*(Aprendizaje Automático Aplicado a la Ciberseguridad del Manejo de Sistemas Energéticos)*

**Tutor:** Felipe Grijalva

**Autor:** Alex Pérez

# Visualización del Dataset Original



| | Demand | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 10729 | 11031 | 11081 | 11132 | 11139 | 11194 | 11607 | 30051 | 30079 |
| 2 | 69.04000000000002 | 39.41917999999998 | 14.39 | 37.43 | 71.99999999999999 | 63.570000000000014 | 43.43000000000001 | 25.26 | 21.24 |
| 3 | 71.92 | 38.38917999999999 | 13.999999999999998 | 38.940000000000026 | 70.99 | 62.220000000000056 | 38.57 | 24.630000000000003 | 21.1 |
| 4 | 67.39000000000001 | 32.53918000000001 | 13.96 | 62.78000000000005 | 63.180000000000014 | 60.76 | 34.30000000000001 | 26.05 | 21.16 |
| 5 | 65.75999999999999 | 29.629179999999995 | 13.909999999999998 | 31.249999999999975 | 50.56999999999998 | 49.18999999999999 | 31.409999999999997 | 31.49 | 21.03 |
| 6 | 64.36 | 27.99918000000001 | 14.010000000000002 | 30.219999999999985 | 47.269999999999975 | 41.41000000000001 | 26.140000000000008 | 34.2 | 19.62 |
| 7 | 63.38 | 26.959179999999986 | 14.09 | 31.799999999999976 | 42.96 | 38.509999999999984 | 25.110000000000007 | 37.91 | 15.31 |
| 8 | 64.95 | 27.87918000000001 | 13.99 | 28.619999999999965 | 44.490000000000016 | 40.680000000000014 | 24.62 | 42.82 | 15.239999999999998 |
| 9 | 65.42 | 28.169179999999987 | 14.71 | 31.55999999999998 | 46.18999999999998 | 46.67000000000003 | 29.159999999999997 | 48.86 | 15.27 |
| 10 | 68.82000000000001 | 29.779180000000007 | 16.599999999999998 | 36.84 | 54.43999999999998 | 50.96000000000005 | 31.15 | 48.92999999999999 | 15.44 |
| 11 | 74.84 | 34.63918 | 15.82 | 43.78000000000003 | 69.6 | 75.02000000000001 | 34.68 | 47.18 | 15.3 |
| 12 | 71.71000000000001 | 46.35918000000001 | 15.47 | 47.26000000000003 | 77.92999999999998 | 82.61000000000001 | 40.29999999999999 | 46.62 | 15.22 |
| 13 | 73.10000000000002 | 44.30918000000002 | 15.739999999999998 | 51.140000000000015 | 88.67999999999998 | 86.47 | 48.57999999999999 | 46.61 | 15.4 |
| 14 | 74.06 | 42.06918000000002 | 15.829999999999998 | 85.01 | 99.71 | 75.53000000000006 | 42.73 | 46.26 | 15.47 |
| 15 | 73.0 | 41.159180000000006 | 14.29 | 93.54000000000003 | 97.24999999999999 | 89.83999999999997 | 40.64 | 46.38 | 15.54 |
| 16 | 75.48999999999998 | 44.04918000000001 | 14.34 | 78.39999999999998 | 98.22000000000003 | 86.60999999999997 | 44.570000000000014 | 47.16 | 15.509999999999998 |
| 17 | 80.21 | 43.31918 | 14.929999999999998 | 55.43000000000001 | 102.88999999999992 | 91.51000000000003 | 50.419999999999995 | 54.03 | 15.48 |
| 18 | 81.43000000000002 | 58.93918000000004 | 15.05 | 71.93999999999997 | 124.38 | 123.71000000000002 | 68.86000000000003 | 51.8 | 15.509999999999998 |
| 19 | 83.95 | 68.57918000000002 | 14.959999999999996 | 92.00000000000003 | 142.64999999999998 | 148.24999999999997 | 74.98 | 42.74 | 15.4 |
| 20 | 82.96 | 67.25918 | 15.249999999999998 | 96.34000000000003 | 128.41000000000005 | 131.22000000000003 | 64.83999999999997 | 28.42 | 15.46 |
| 21 | 76.73 | 58.65917999999999 | 14.11 | 77.97999999999992 | 116.24999999999996 | 109.07999999999998 | 50.180000000000014 | 27.75 | 15.87 |

Figura: Dataset de Demanda

[9].

# Dataset Utilizado

**Datos de Entrenamiento:**

- Cada transformador tiene 25,000 mediciones temporales.
- Dataset total: 17 transformadores.

**Distribución de Datos:**

- **Entrenamiento:** Datos 0 a 17,500.
- **Validación:** Datos 17,501 a 20,000.
- **Testing (prototipo FDIA):** Datos 20,001 a 22,500.
- **Evaluación final FDIA:** Últimos 2,500 datos.

**Frecuencia de las Mediciones:**

- Cada medición representa 1 hora de datos, lo que implica un total de 25,000 horas.

# Frecuencia Fundamental y Ciclo Diario (Análisis Espectral

**Frecuencia Fundamental:**

- Frecuencia identificada: $f \approx 0.0418\,\text{Hz}$.
- Esta frecuencia corresponde a un ciclo completo cada 24 horas.

**Relación Matemática:**

$$T = \frac{1}{f} = \frac{1}{0.0418} \approx 24\,\text{horas} \tag{1}$$

**Interpretación:**

- Relación con el tamaño de ventana: un ciclo completo captura la variabilidad cíclica diaria de la demanda energética.

# Ventana de Aprendizaje (sin Análisis Espectral)

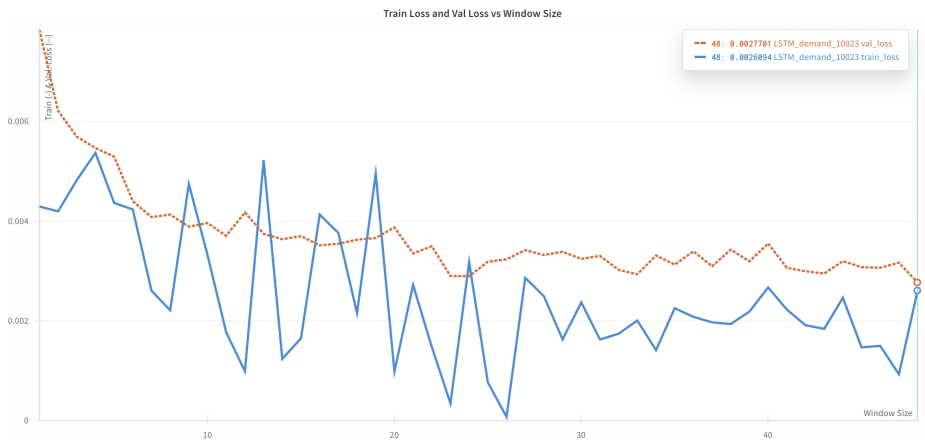



Figura: 10023 T&V vs WindowSize

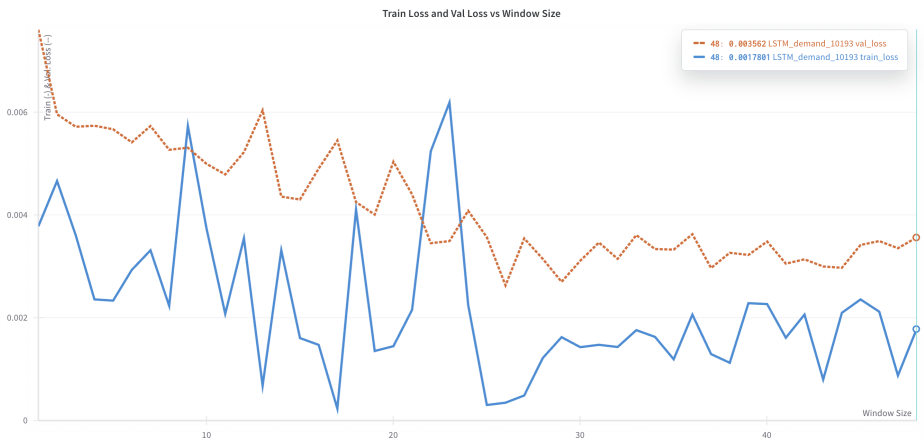



Figura: 10193 T&V vs WindowSize

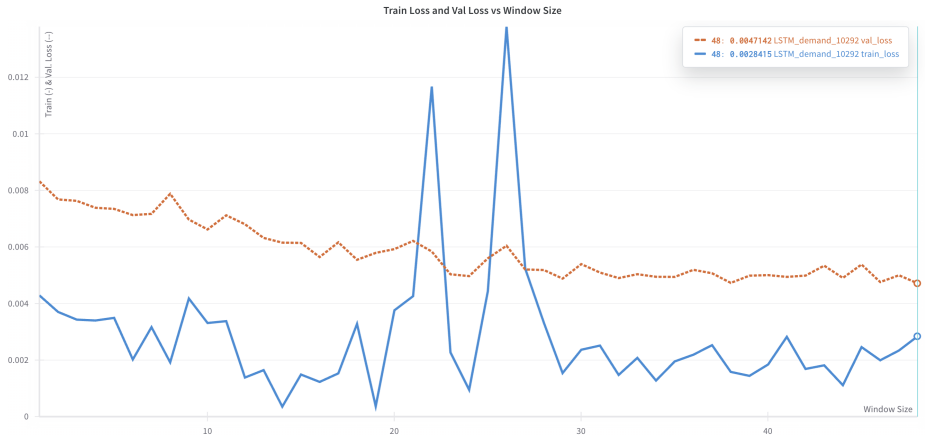



Figura: 10292 T&V vs WindowSize

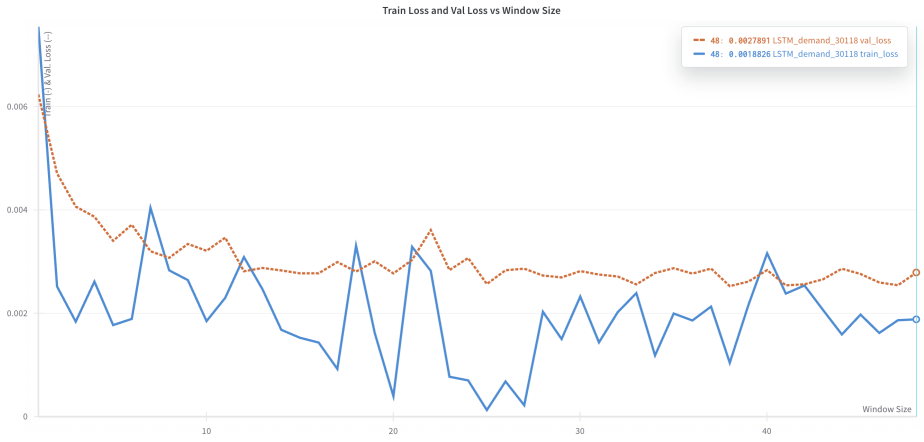



Figura: 30118 T&V vs WindowSize

# Resultados de Evaluación de Modelos por Transformador (Conjunto de Test / Ventana de 24 horas)

| M./T. | 10023 | 10193 | 10292 | 10370 | 10729 | 11031 | 11081 | 11132 | 11139 | 11194 | 11607 | 30051 | 30079 | 30086 | 30095 | 30110 | 30118 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **LSTM (improved: 128 hidden dim., 3 layers, dropout 30 %)** | | | | | | | | | | | | | | | | |
| **MSE** | 0.0048 | 0.0059 | 0.0091 | 0.0045 | 0.0035 | 0.0058 | 0.0046 | 0.0096 | 0.0047 | 0.0035 | 0.0091 | 0.0025 | 0.0013 | 0.0070 | 0.0016 | 0.0055 | 0.0030 |
| **MAE** | 0.0507 | 0.0556 | 0.0720 | 0.0528 | 0.0430 | 0.0589 | 0.0330 | 0.0753 | 0.0526 | 0.0438 | 0.0725 | 0.0354 | 0.0207 | 0.0539 | 0.0279 | 0.0532 | 0.0360 |
| **RMSE** | 0.0691 | 0.0771 | 0.0956 | 0.0670 | 0.0589 | 0.0762 | 0.0675 | 0.0979 | 0.0685 | 0.0590 | 0.0956 | 0.0500 | 0.0356 | 0.0835 | 0.0397 | 0.0743 | 0.0543 |
| **$R^2$** | 0.8807 | 0.7732 | 0.6468 | 0.9199 | 0.8315 | 0.8070 | 0.7660 | 0.7712 | 0.8216 | 0.8641 | 0.6539 | 0.9617 | 0.9804 | 0.8941 | 0.9819 | 0.8175 | 0.9402 |
| | **TRANSFORMERS** | | | | | | | | | | | | | | | | |
| **MSE** | 0.0093 | 0.0080 | 0.0100 | 0.0089 | 0.0048 | 0.0087 | 0.0068 | 0.0173 | 0.0067 | 0.0064 | 0.0100 | 0.0070 | 0.0014 | 0.0121 | 0.0084 | 0.0086 | 0.0094 |
| **MAE** | 0.0745 | 0.0679 | 0.0780 | 0.0755 | 0.0541 | 0.0723 | 0.0542 | 0.0996 | 0.0601 | 0.0607 | 0.0763 | 0.0552 | 0.0195 | 0.0712 | 0.0563 | 0.0713 | 0.0792 |
| **RMSE** | 0.0964 | 0.0893 | 0.1002 | 0.0945 | 0.0690 | 0.0934 | 0.0824 | 0.1317 | 0.0820 | 0.0797 | 0.1002 | 0.0838 | 0.0375 | 0.1102 | 0.0919 | 0.0926 | 0.0971 |
| **$R^2$** | 0.7676 | 0.6958 | 0.6123 | 0.8404 | 0.7690 | 0.7104 | 0.6518 | 0.5859 | 0.7447 | 0.7521 | 0.6195 | 0.8925 | 0.9783 | 0.8157 | 0.9032 | 0.7161 | 0.8091 |
| | **TCN** | | | | | | | | | | | | | | | | |
| **MSE** | 0.0091 | 0.0081 | 0.0100 | 0.0096 | 0.0043 | 0.0080 | 0.0069 | 0.0158 | 0.0071 | 0.0068 | 0.0098 | 0.0079 | 0.0021 | 0.0109 | 0.0081 | 0.0079 | 0.0059 |
| **MAE** | 0.0755 | 0.0700 | 0.0762 | 0.0776 | 0.0498 | 0.0691 | 0.0580 | 0.0963 | 0.0606 | 0.0596 | 0.0734 | 0.0668 | 0.0363 | 0.0663 | 0.0596 | 0.0660 | 0.0483 |
| **RMSE** | 0.0954 | 0.0902 | 0.1001 | 0.0980 | 0.0658 | 0.0897 | 0.0832 | 0.1257 | 0.0841 | 0.0822 | 0.0989 | 0.0886 | 0.0457 | 0.1046 | 0.0901 | 0.0889 | 0.0770 |
| **$R^2$** | 0.7723 | 0.6896 | 0.6133 | 0.8285 | 0.7898 | 0.7331 | 0.6450 | 0.6231 | 0.7316 | 0.7364 | 0.6294 | 0.8799 | 0.9677 | 0.8338 | 0.9070 | 0.7384 | 0.8799 |

# Comparación de Modelos Predictivos (LSTM, TCN y Transformers

- Se analizaron los modelos en términos de las métricas RMSE y $R^2$.
- Los valores promedio de las métricas fueron calculados para cada modelo en todos los transformadores.

```
RMSE Promedio - LSTM: 0.0688
RMSE Promedio - Transformers: 0.0901
RMSE Promedio - TCN: 0.0887
-----------------------------------
R² Promedio - LSTM: 0.8419
R² Promedio - Transformers: 0.7567
R² Promedio - TCN: 0.7646
```

- **LSTM** resultó ser el mejor modelo con un RMSE promedio más bajo y un $R^2$ promedio más alto.
- **TCN** mostró un RMSE promedio y $R^2$ promedio ligeramente mejores que **Transformers** $\rightarrow$ se propone un **Modelo Híbrido** que combine **LSTM y TCN**.

# Resultados de Evaluación de Modelos por Transformador (Conjunto de Test / Ventana de 24 horas)

| M./T. | 10023 | 10193 | 10292 | 10370 | 10729 | 11031 | 11081 | 11132 | 11139 | 11194 | 11607 | 30051 | 30079 | 30086 | 30095 | 30110 | 30118 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **HYBRID (TCN + LSTM (standard: 64 hidden dim., 1 layer, dropout 20 %))** | | | | | | | | | | | | | | | | |
| **MSE** | 0.0047 | 0.0064 | 0.0089 | 0.0046 | 0.0040 | 0.0060 | 0.0072 | 0.0091 | 0.0050 | 0.0032 | 0.0103 | 0.0023 | 0.0031 | 0.0064 | 0.0015 | 0.0050 | 0.0030 |
| **MAE** | 0.0498 | 0.0610 | 0.0711 | 0.0523 | 0.0461 | 0.0581 | 0.0584 | 0.0734 | 0.0536 | 0.0417 | 0.0759 | 0.0353 | 0.0405 | 0.0496 | 0.0275 | 0.0522 | 0.0366 |
| **RMSE** | 0.0687 | 0.0801 | 0.0946 | 0.0677 | 0.0632 | 0.0775 | 0.0849 | 0.0955 | 0.0705 | 0.0568 | 0.1017 | 0.0484 | 0.0554 | 0.0797 | 0.0392 | 0.0707 | 0.0546 |
| **$R^2$** | 0.8820 | 0.7552 | 0.6547 | 0.9181 | 0.8064 | 0.8008 | 0.6299 | 0.7825 | 0.8111 | 0.8743 | 0.6082 | 0.9642 | 0.9525 | 0.9035 | 0.9823 | 0.8348 | 0.9397 |
| | **HYBRID (TCN + LSTM (improved: 128 hidden dim., 3 layers, dropout 30 %))** | | | | | | | | | | | | | | | | |
| **MSE** | 0.0051 | 0.0060 | 0.0094 | 0.0059 | 0.0038 | 0.0060 | 0.0091 | 0.0085 | 0.0054 | 0.0040 | 0.0099 | 0.0023 | 0.0025 | 0.0066 | 0.0027 | 0.0046 | 0.0030 |
| **MAE** | 0.0521 | 0.0589 | 0.0732 | 0.0596 | 0.0475 | 0.0593 | 0.0736 | 0.0707 | 0.0550 | 0.0458 | 0.0762 | 0.0354 | 0.0312 | 0.0531 | 0.0366 | 0.0506 | 0.0355 |
| **RMSE** | 0.0714 | 0.0776 | 0.0968 | 0.0769 | 0.0619 | 0.0773 | 0.0954 | 0.0924 | 0.0734 | 0.0631 | 0.0995 | 0.0484 | 0.0501 | 0.0811 | 0.0516 | 0.0681 | 0.0548 |
| **$R^2$** | 0.8723 | 0.7698 | 0.6382 | 0.8943 | 0.8144 | 0.8015 | 0.5331 | 0.7960 | 0.7952 | 0.8447 | 0.6250 | 0.9641 | 0.9611 | 0.9001 | 0.9694 | 0.8465 | 0.9391 |

# Comparación de Modelos Híbridos con Modelo LSTM

- Se evaluaron las métricas RMSE y $R^2$.

```
RMSE Promedio - LSTM: 0.0688
RMSE Promedio - Hybrid ST: 0.0711
RMSE Promedio - Hybrid Improved: 0.0729
----------------------------------------
R² Promedio - LSTM: 0.8419
R² Promedio - Hybrid ST: 0.8294
R² Promedio - Hybrid Improved: 0.8215
```

- **LSTM** mantuvo el mejor rendimiento global.
- Se propone utilizar solamente esta arquitectura para modelo MIMO.

# One-Class SVM para la Detección de Anomalías

- No requiere datos etiquetados.
- Menos dependiente de la distribución.
- Enfoque en una sola clase.
- Fácil de ajustar.

Lo que se va a hacer:

- Obtener las predicciones, usando el modelo LSTM, de nuevos datos *normales* proporcionados.
- Entrenar el modelo One-Class SVM con dichas predicciones.
- Usar One-Class SVM para detectar posibles anomalías en nuevos datos.

# Detección de anomalías - Nuevos Datos *normales*

- Para el transformador 30095, con $R^2 = 0.9839$ y RMSE $= 0.0375$.

```
Cantidad total de datos analizados: 2476
Cantidad total de anomalías detectadas: 25
Porcentaje de datos marcados como anomalías: 1.01%
```

# Detección de anomalías - Nuevos Datos *infectados*

- Para el transformador 30095, con $R^2 = 0.9839$ y RMSE $= 0.0375$.
- Con un factor de ruido de 0.3. El factor de ruido de 0.3 indica que se han añadido desviaciones aleatorias a los valores originales, con una amplitud del 30 % de la magnitud del dato original.

```
Cantidad total de datos analizados: 2476
Cantidad total de anomalías detectadas: 916
Porcentaje de datos marcados como anomalías: 37.00%
```

# Referencias


[1] Zhang, Y., Wang, X., & Liu, J. (2020). Detecting False Data Injection Attacks in Smart Grids Using CNNs and LSTMs. *IEEE Transactions on Smart Grid*, 11(4), 3043-3051. https://arxiv.org/pdf/2006.11477.

[2] Stanford University. (2024). Recurrent Neural Networks cheatsheet. Recuperado de `https://stanford.edu/~shervine/teaching/cs-230/cheatsheet-recurrent-neural-networks`.

[3] Cayci, S., & Eryilmaz, A. (2024). Convergence of Gradient Descent for Recurrent Neural Networks: A Nonasymptotic Analysis. *arXiv preprint arXiv:2402.12241*. Recuperado de `https://arxiv.org/abs/2402.12241`.

[4] Staudemeyer, R. C., & Morris, E. R. (2019). Understanding LSTM – a tutorial into Long Short-Term Memory Recurrent Neural Networks. *arXiv preprint arXiv:1909.09586*. Recuperado de `https://arxiv.org/abs/1909.09586`.