

# Actividades 4.6 y 6.11

## Predicción demanda y precio

Basque Center for Applied Mathematics-BCAM



INTELIGENCIA ARTIFICIAL  
PARA LA TRANSICIÓN  
ENERGÉTICA SOSTENIBLE



Plan de Recuperación,  
Transformación  
y Resiliencia

(bcam)



EXCELENCIA  
SEVERO  
OCHOA

[www.bcamath.org](http://www.bcamath.org)  
basque center for applied mathematics

# Actividades 4.6 y 6.11

## Predicción demanda y precio



Santiago Mazuelas



Verónica Alvarez



Paula Martin



Onintze Zaballa

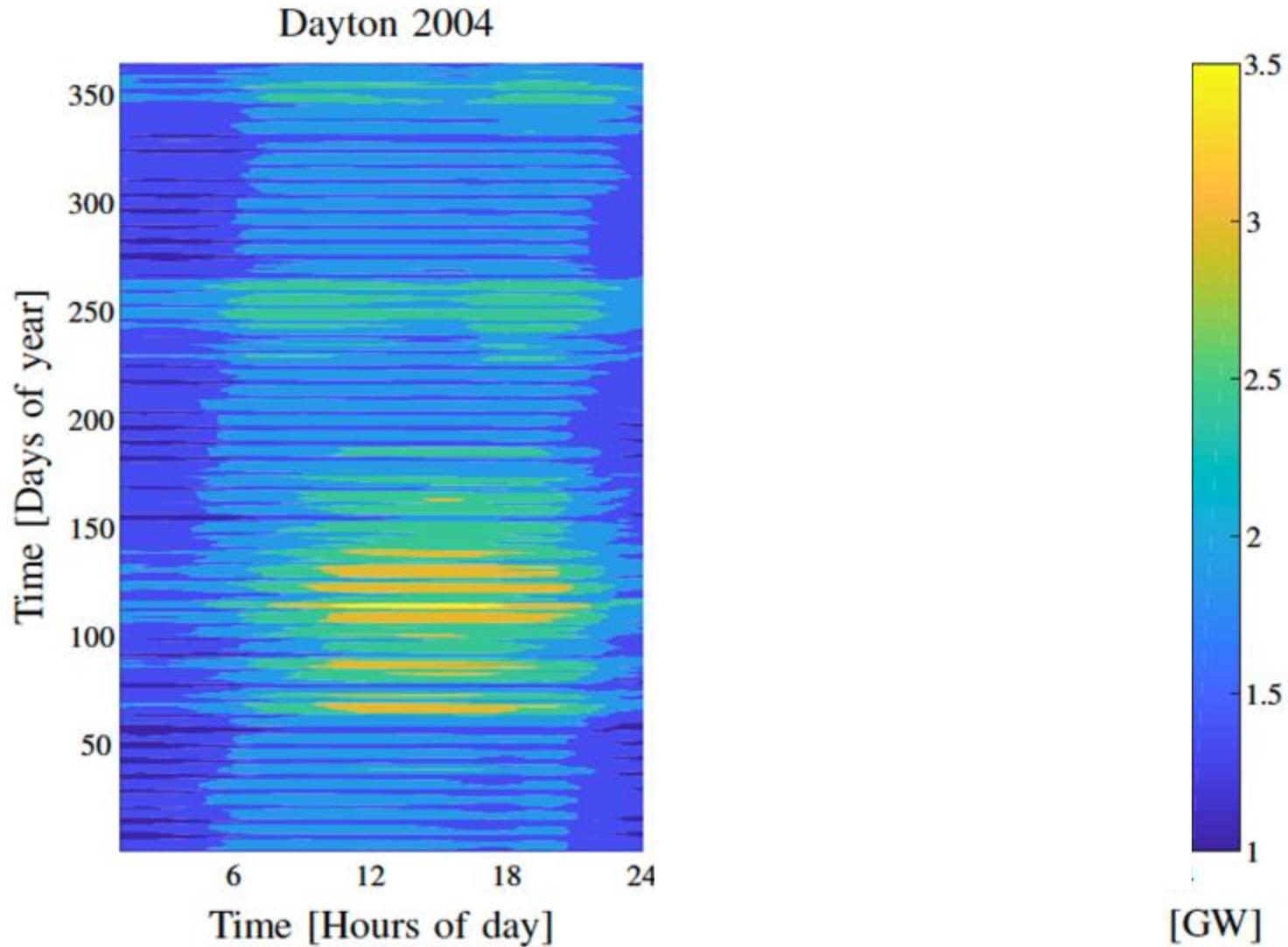


Xabier de Juan

# Predicción demanda



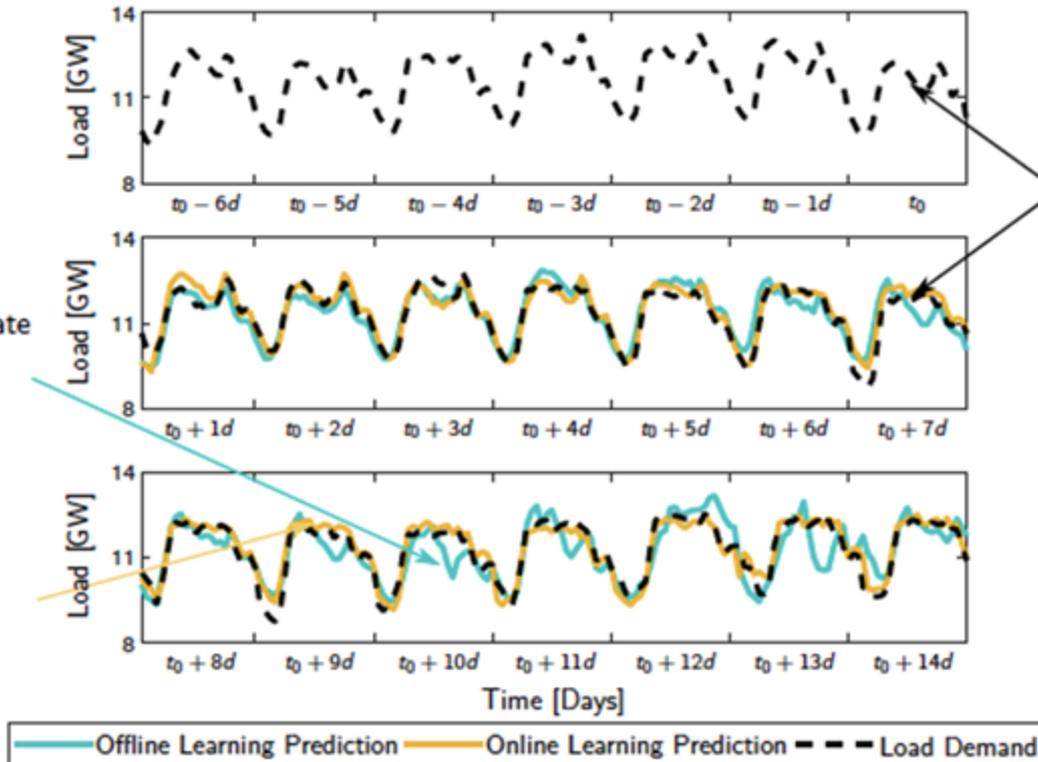
# Los datos cambian con el tiempo



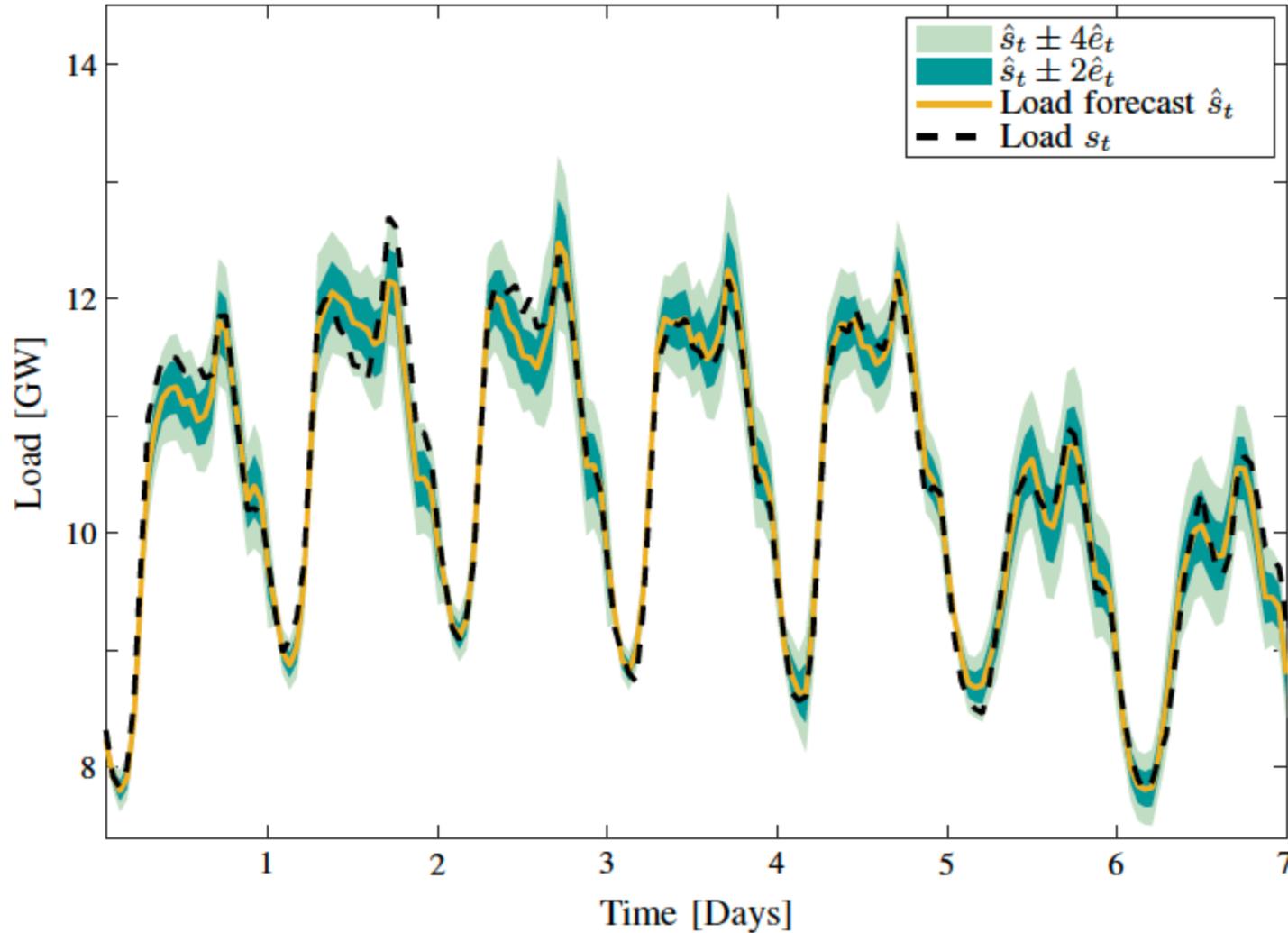
# Online learning

Methods based on offline learning cannot accommodate to such changes

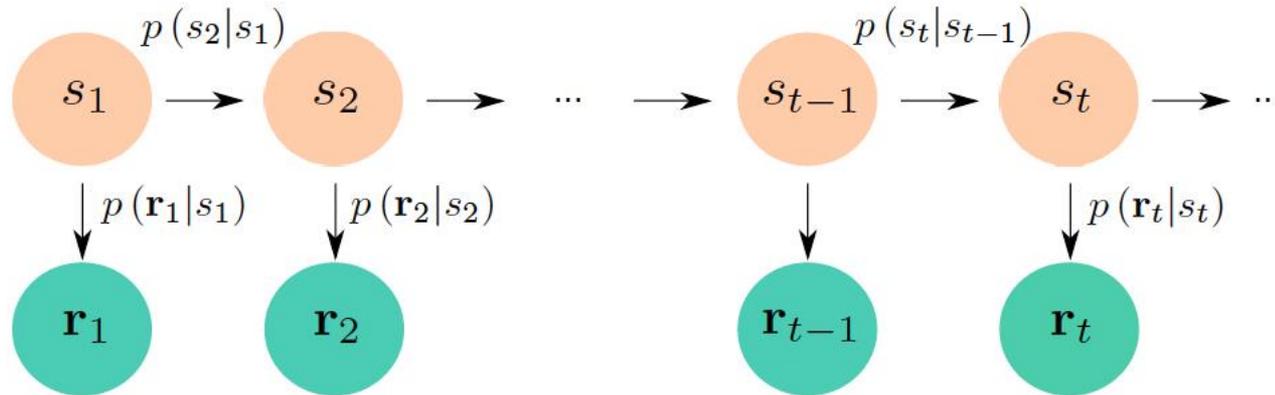
Methods based on online learning correctly adapt models after few days



# Predicciones probabilísticas



# Probabilistic Load Forecasting Based on Adaptive Online Learning (APLF)



Model relationship  
consecutive loads

$$p(s_t | s_{t-1}) = N(s_t; \mathbf{u}_s^T \boldsymbol{\eta}_{s,c}, \sigma_{s,c})$$

Model relationship  
loads and observations

$$p(\mathbf{r}_t | s_t) \propto p(s_t | \mathbf{r}_t) = N(s_t; \mathbf{u}_r^T \boldsymbol{\eta}_{r,c}, \sigma_{r,c})$$

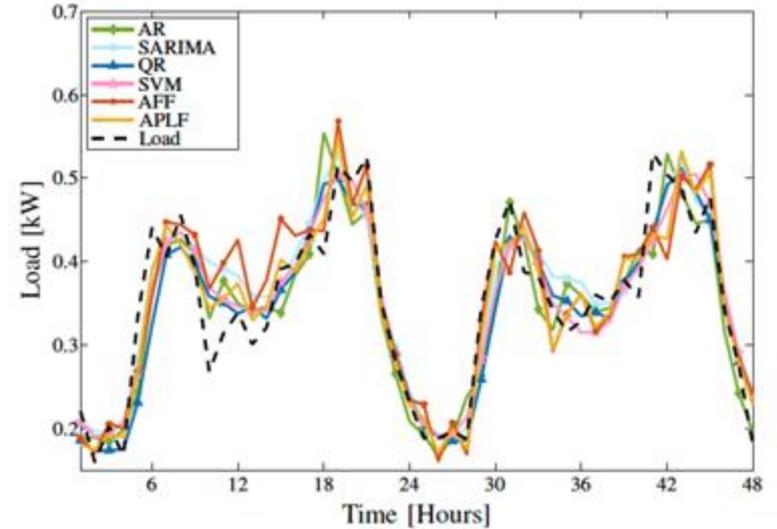
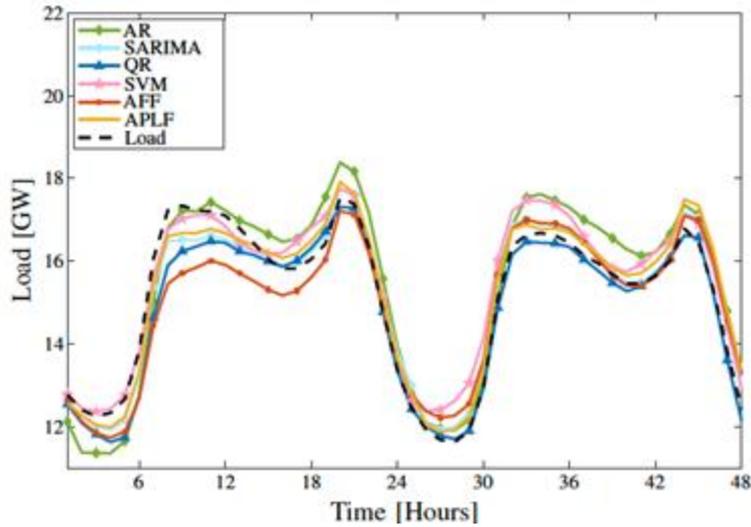
Online learning

$$\boldsymbol{\eta}_i = \boldsymbol{\eta}_{i-1} + \frac{\mathbf{P}_{i-1} \mathbf{u}_{t_i}}{\lambda + \mathbf{u}_{t_i}^T \mathbf{P}_{i-1} \mathbf{u}_{t_i}} (s_{t_i} - \mathbf{u}_{t_i}^T \boldsymbol{\eta}_{i-1})$$

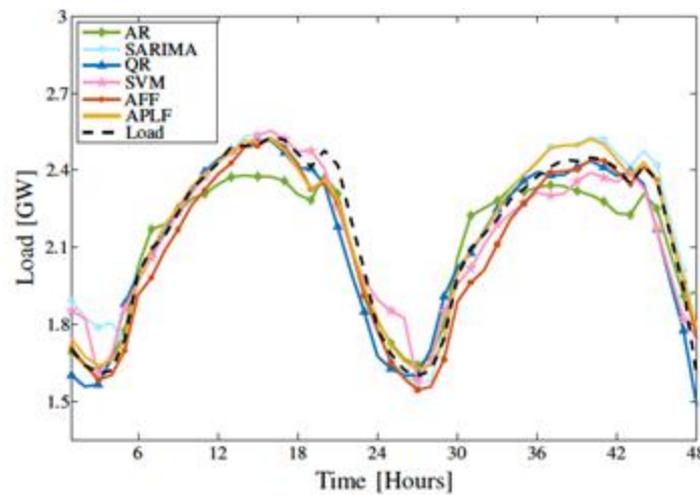
Probabilistic forecasts

$$p(s_{t+i} | s_t, \mathbf{r}_{t+1}, \dots, \mathbf{r}_{t+i}) = N(s_{t+i}; \hat{s}_{t+i}, \hat{e}_{t+i})$$

# Resultados



New  
England



100  
buildings

Dayton

# Resultados

TABLE I  
RMSE AND MAPE OF PREDICTION ERRORS FOR APLF AND 11 STATE-OF-THE-ART TECHNIQUES ON 7 DATASETS.

Method	Large-size region				Medium-size region						Small-size region			
	Belgium		New Engl.		GEFCom12		GEFCom2014		Dayton		400 builds.		100 builds.	
	[GW]	[%]	[GW]	[%]	[MW]	[%]	[MW]	[%]	[GW]	[%]	[kW]	[%]	[kW]	[%]
LR	1.47	11.8	1.73	8.0	5.48	20.6	0.30	15.0	0.46	15.5	0.06	9.1	0.07	14.3
SARIMA	0.81	5.5	1.22	5.4	3.42	11.6	0.25	12.0	0.22	7.7	0.05	9.9	0.07	16.5
QR	1.05	9.2	1.17	5.6	5.46	25.5	0.29	14.7	0.20	7.1	0.08	17.6	0.12	26.8
GP	<b>0.52</b>	<b>4.0</b>	0.89	4.2	2.50	<b>8.5</b>	<b>0.24</b>	<b>10.6</b>	0.19	6.3	<b>0.04</b>	<b>6.8</b>	<b>0.05</b>	<b>11.2</b>
SVM	0.69	4.7	1.11	5.5	3.28	12.2	<b>0.24</b>	12.3	<b>0.17</b>	<b>5.7</b>	<b>0.04</b>	8.0	<b>0.05</b>	11.4
DRN	1.74	13.0	<b>0.52</b>	<b>2.3</b>	<b>2.17</b>	7.6	0.27	19.5	0.31	11.3	<b>0.04</b>	7.0	0.07	14.7
AR	0.66	5.1	1.28	5.6	3.94	16.2	0.30	18.5	0.38	13.6	<b>0.04</b>	8.9	0.07	16.9
ARNFS	1.08	9.4	1.95	10.9	4.41	17.4	0.33	18.4	0.31	14.1	<b>0.04</b>	8.1	<b>0.05</b>	11.4
ARRFFS	1.18	10.3	2.05	11.2	4.54	17.2	0.34	17.7	0.28	10.3	0.08	17.5	0.10	23.4
SFDA	1.14	8.9	1.41	10.2	5.04	16.8	0.35	14.5	0.39	21.8	0.06	13.0	0.08	18.6
AFF	0.95	6.7	1.23	5.8	2.91	10.7	0.25	15.2	0.26	9.6	0.05	10.2	0.07	14.6
APLF	<b>0.33</b>	<b>2.3</b>	<b>0.86</b>	<b>3.9</b>	<b>2.15</b>	<b>8.1</b>	<b>0.20</b>	<b>9.6</b>	<b>0.16</b>	<b>5.5</b>	<b>0.03</b>	<b>6.3</b>	<b>0.05</b>	<b>11.0</b>

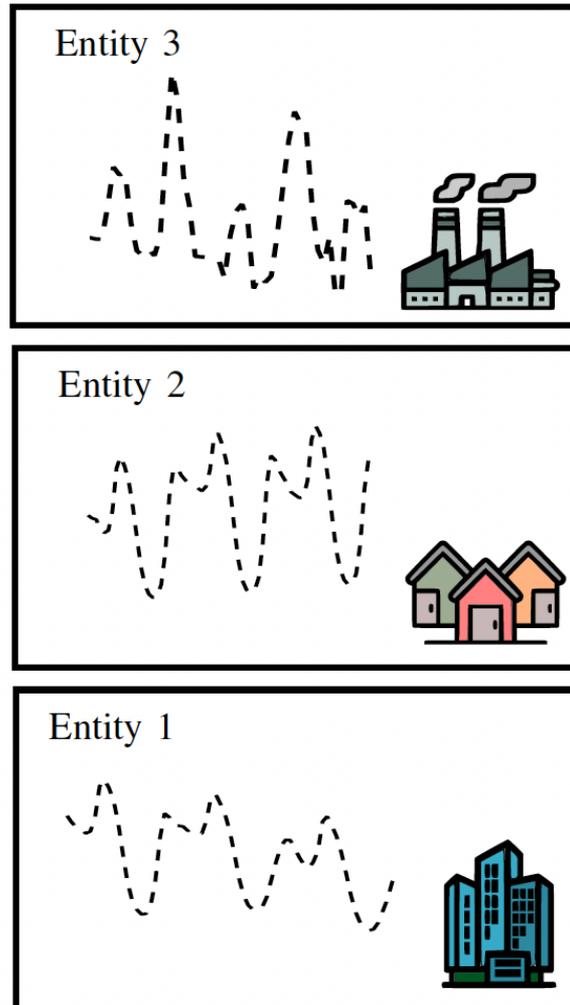
# Premio Sociedad de Estadística e Investigación Operativa



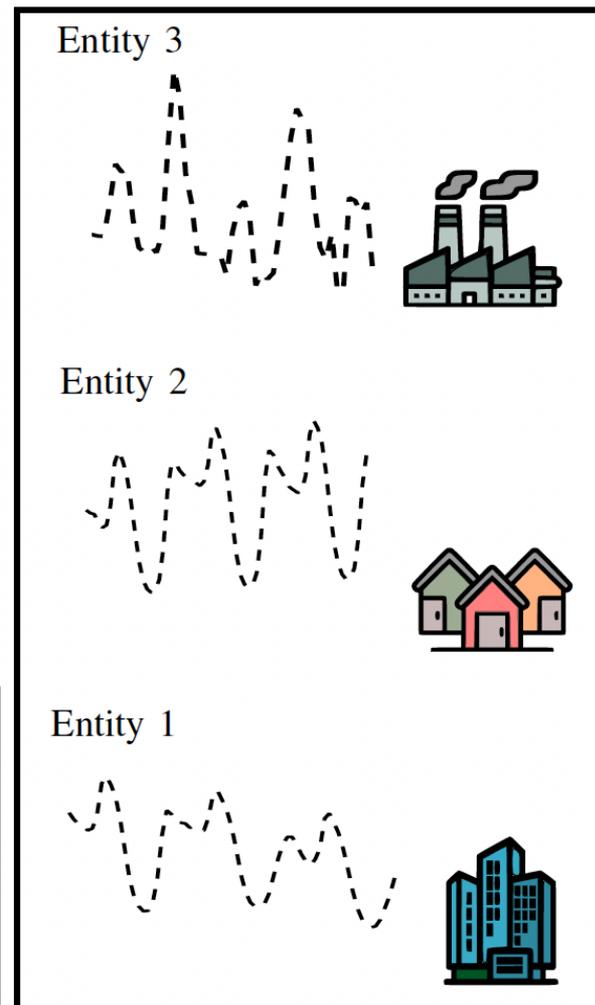
Mejor Contribución Aplicada con un Impacto en el Ámbito Social, la Innovación o la Transferencia del Conocimiento en el Campo de la Estadística, 2022

# Utilizar datos de otras regiones

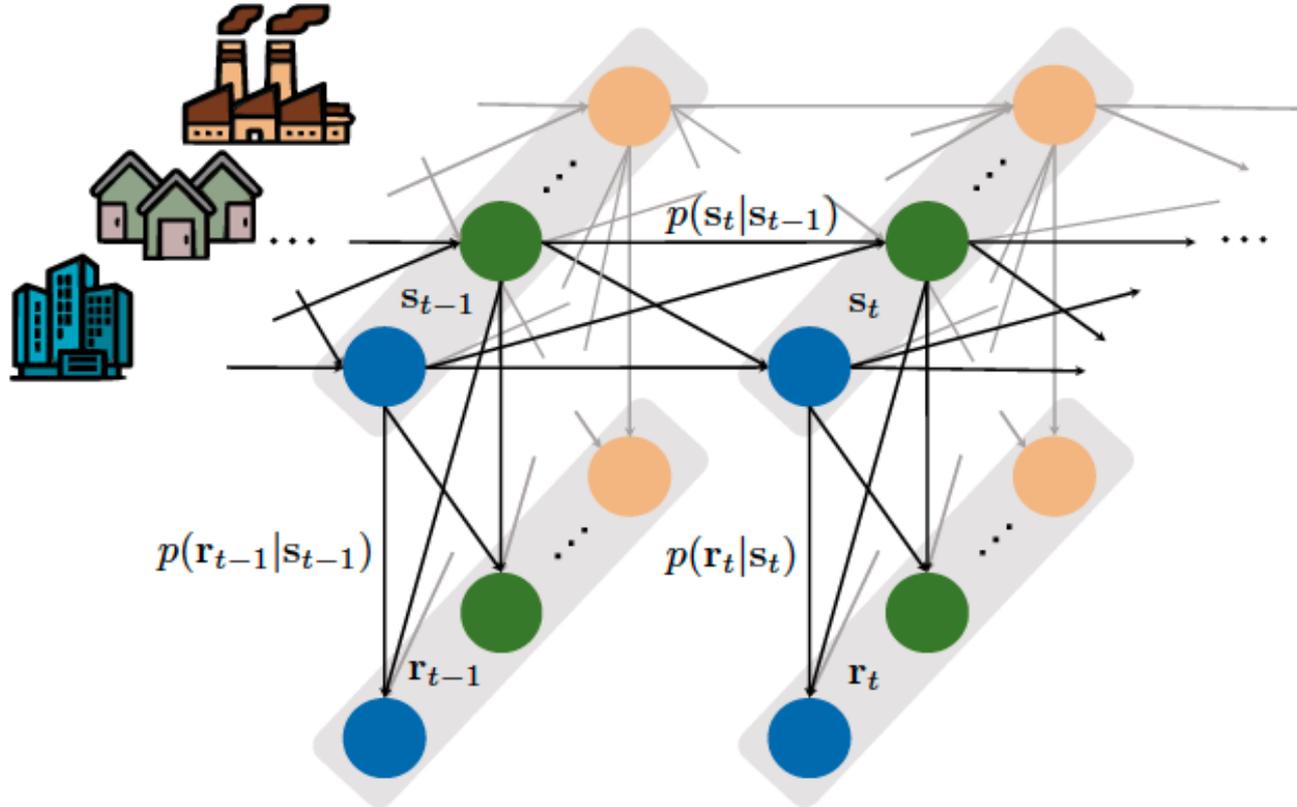
Single-task learning



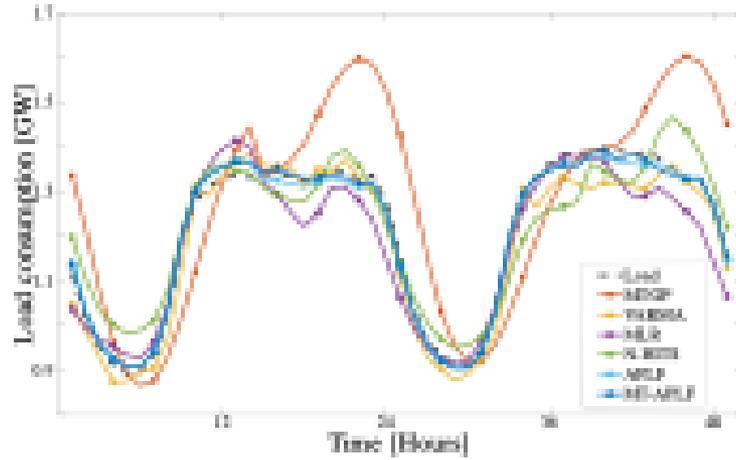
Multi-task learning



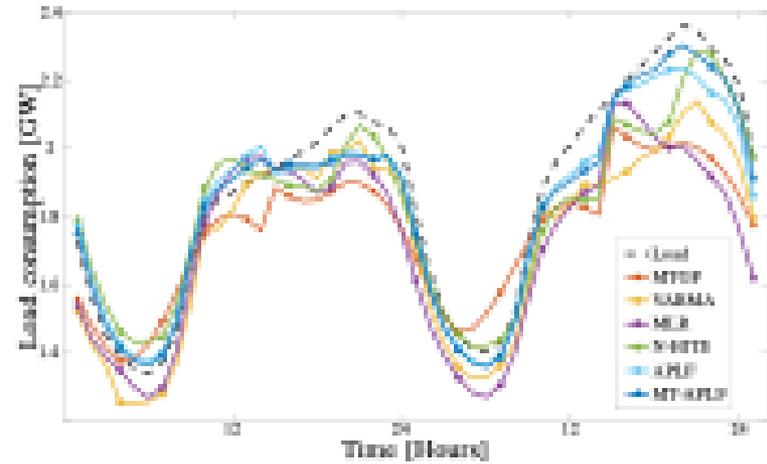
# Multitask APLF



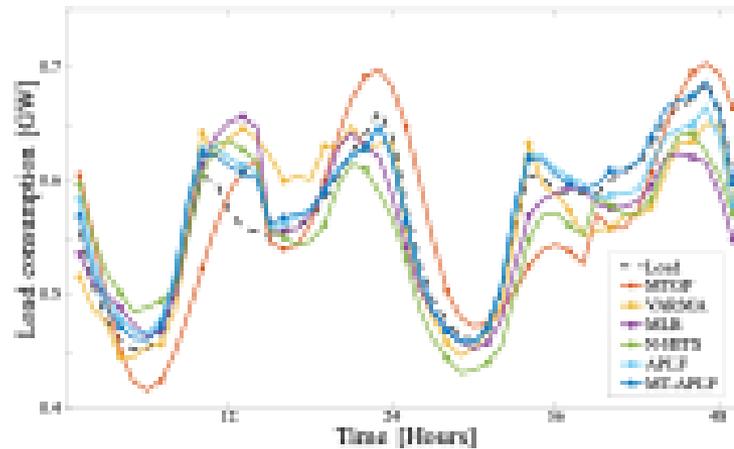
# Resultados



Maine



Vermont



Western Massachusetts

# Resultados

	APLF		MTGP		VAR		MLR		N-HITS		MT-APLF		
	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE	
<b>GEFCom2017</b>	Entity 1	5.04	0.25	6.33	0.28	7.07	0.34	11.01	0.62	7.52	0.35	<b>4.57</b>	<b>0.22</b>
	Entity 2	3.18	<b>0.05</b>	4.55	0.07	5.28	0.09	6.87	0.13	4.77	0.07	<b>3.04</b>	<b>0.05</b>
	Entity 3	4.53	0.19	6.15	0.24	6.98	0.29	8.92	0.40	6.72	0.25	<b>3.90</b>	<b>0.16</b>
	Entity 4	4.29	0.08	5.95	0.01	6.97	0.13	9.83	0.21	6.52	0.11	<b>3.73</b>	<b>0.07</b>
	Entity 5	4.41	0.06	5.96	0.07	6.57	0.09	10.99	0.16	6.83	0.08	<b>4.04</b>	<b>0.05</b>
	Entity 6	5.15	0.12	6.34	0.14	7.03	0.17	12.75	0.35	7.16	0.16	<b>4.43</b>	<b>0.10</b>
	Entity 7	4.64	<b>0.03</b>	5.64	0.04	6.48	0.05	8.66	0.07	6.05	0.04	<b>4.37</b>	<b>0.03</b>
	Entity 8	4.86	0.13	6.43	0.16	7.18	0.19	9.63	0.29	6.29	0.16	<b>4.06</b>	<b>0.10</b>
	TOTAL	4.51	0.11	5.92	0.14	6.70	0.17	9.83	0.28	6.48	0.15	<b>4.02</b>	<b>0.10</b>
<b>NewEngland</b>	Entity 1	9.80	0.13	11.34	0.14	12.91	0.16	16.15	0.23	10.47	0.14	<b>9.29</b>	<b>0.12</b>
	Entity 2	5.58	0.09	6.81	0.10	7.44	0.11	11.50	0.20	7.32	0.11	<b>5.11</b>	<b>0.08</b>
	Entity 3	11.67	0.07	14.01	0.08	19.68	0.10	31.51	0.15	16.30	0.09	<b>11.12</b>	<b>0.06</b>
	Entity 4	5.52	0.23	6.95	0.29	7.63	0.32	12.72	0.61	7.33	0.31	5.34	<b>0.22</b>
	Entity 5	7.52	0.08	9.07	0.09	9.62	0.11	15.79	0.20	9.54	0.10	<b>7.00</b>	<b>0.07</b>
	Entity 6	6.16	0.12	7.63	0.15	8.85	0.17	14.81	0.34	8.36	0.16	<b>6.00</b>	<b>0.12</b>
	Entity 7	8.11	0.22	9.21	0.23	10.32	0.26	18.31	0.50	8.95	0.23	<b>7.36</b>	<b>0.19</b>
	Entity 8	4.71	0.16	6.03	0.21	6.93	0.24	11.41	0.42	6.42	0.22	4.47	<b>0.15</b>
	TOTAL	7.38	0.14	8.88	0.16	10.42	0.18	16.53	0.33	9.34	0.17	<b>6.96</b>	<b>0.13</b>
<b>PJM</b>	Entity 1	4.08	0.82	6.39	1.13	7.94	1.42	6.16	1.11	5.52	1.04	<b>3.78</b>	<b>0.74</b>
	Entity 2	5.25	0.95	8.47	1.24	9.20	1.15	7.21	0.97	6.41	0.99	<b>4.48</b>	<b>0.69</b>
	Entity 3	5.20	0.14	10.05	0.25	9.46	0.22	7.16	0.17	6.94	0.18	<b>4.65</b>	<b>0.11</b>
	Entity 4	5.10	0.20	7.44	0.27	9.20	0.34	7.32	0.27	6.79	0.27	<b>4.72</b>	<b>0.18</b>
	Entity 5	5.16	1.04	8.16	1.41	6.46	1.15	6.36	1.15	6.19	1.18	<b>4.51</b>	<b>0.83</b>
	Entity 6	4.26	0.09	5.43	0.11	6.44	0.12	6.41	0.13	5.48	0.11	<b>4.12</b>	<b>0.08</b>
	Entity 7	8.78	0.18	12.79	0.23	24.64	0.48	22.51	0.44	10.45	0.20	<b>8.34</b>	<b>0.17</b>
	Entity 8	4.01	0.44	10.37	1.00	6.62	0.63	5.80	0.59	5.34	0.54	<b>3.66</b>	<b>0.37</b>
	TOTAL	5.23	0.48	8.64	0.70	9.99	0.69	8.62	0.60	6.64	0.55	<b>4.78</b>	<b>0.40</b>

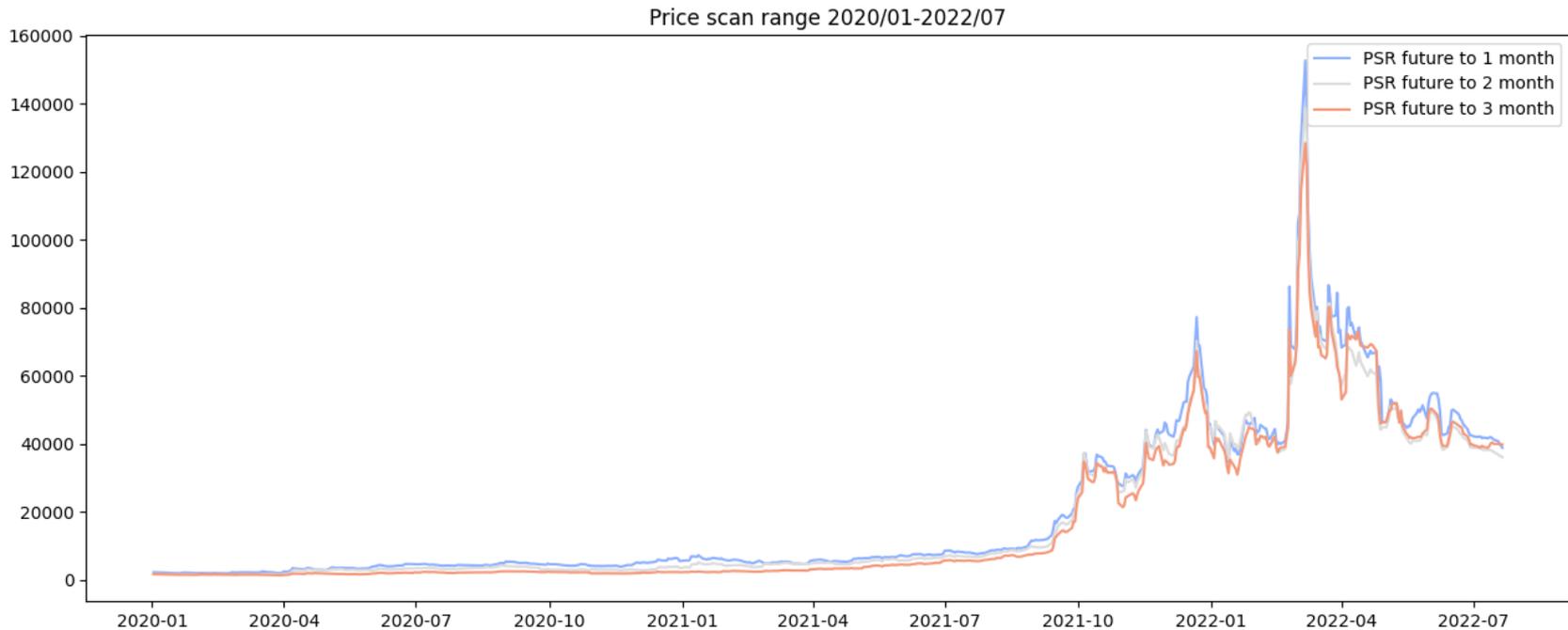
# Best Paper Award IEEE Sustainable Power and Energy Conference 2024



IEEE ISPEC 2024 “Advancing Sustainable Power and Energy Solutions for a Greener Future”

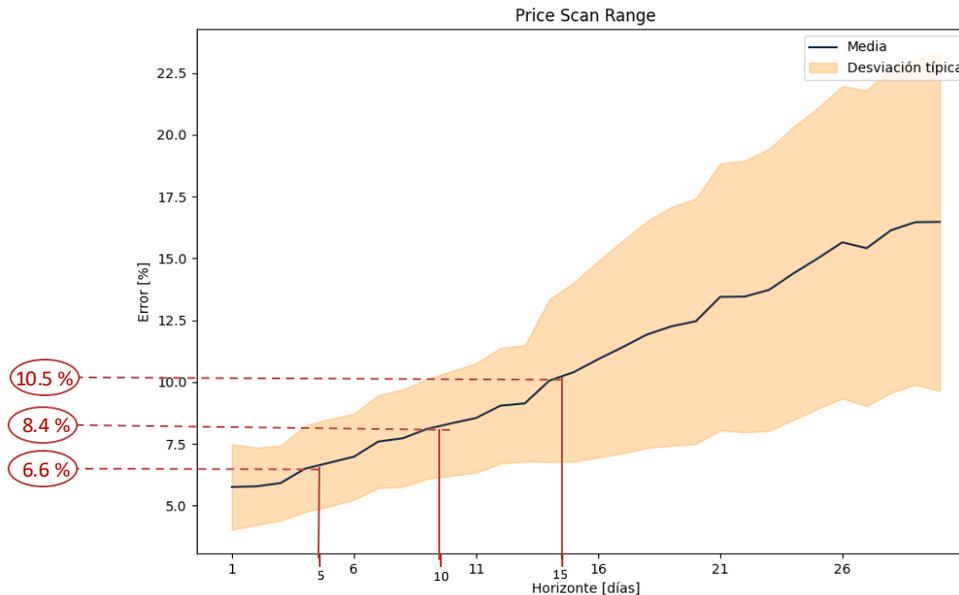
# Predicción precios

- Actividad realizada en colaboración con el departamento de Finanzas de Iberdrola: Vanessa Verde (Iberdrola Finanzas), Andrea Oteo (Iberdrola Finanzas), Julián Calvo (Iberdrola)



# Predicción precios

Herramientas de inteligencia artificial capaces de predecir con errores del 8% a un horizonte de 10 días



# Github con el código y documentación

Navigation bar: IberdrolaBCAM-IA4TES / IberdrolaFinanzasIA4TES

Navigation links: [Code](#) [Issues](#) [Pull requests](#) [Actions](#) [Projects](#) [Security](#) [Insights](#) [Settings](#)

Repository: **IberdrolaFinanzasIA4TES** Private Watch 0

Branches: [main](#) [1 branch](#) [0 tags](#) [Go to file](#) [Add file](#) [Code](#)

<b>rodriguez96</b> Update Example.py	bfa8b9e 1 hour ago	28 commits
Data	Add files via upload	18 hours ago
FitPredict	Update fit_predict_functions.py	18 hours ago
IntercommoditySpreads	Update getIntercomData.py	18 hours ago
PriceScanRange	Update getMarketData.py	18 hours ago
docs	Update set_up.md	1 hour ago
Example.ipynb	Add files via upload	11 hours ago
Example.md	Add files via upload	11 hours ago
Example.py	Update Example.py	1 hour ago
README.md	Update README.md	11 hours ago
environment.yml	Add files via upload	yesterday

# Github con el código y documentación

☰ README.md ✎

## Iberdrola Finanzas IA4TES

---

### Introducción

En este repositorio se presenta el código desarrollado por el grupo de Machine Learning del Basque Centre of Applied Mathematics (BCAM) para la predicción las variables consideradas como más relevantes a la hora de determinar el *Initial Margin*.

Antes de entrar a describir las funciones, vamos a definir los conceptos más importantes:

**Initial Margin:** es la cantidad de dinero que una empresa tiene depositada para negociar posiciones futuras en los mercados energéticos, con efecto negativo sobre la deuda neta. Este impacto sobre la deuda neta y la volatilidad creciente de los precios genera la necesidad de desarrollar metodologías que predigan esa cantidad, así como las subidas de precio y/o de la demanda de energía. Para ello se han desarrollado predictores para las variables más relevantes:

- **Price Scan Ranges.**
- **Spread Credits.**
- **Intercommodity Ratios.**

Las técnicas empleadas han permitido obtener predicciones para 60 productos que abarcan electricidad y gas en diversos mercados, considerando futuros mensuales, trimestrales y anuales. Para ello, se han utilizado redes neuronales recurrentes de memoria a corto y largo plazo (LSTM), que aprovechan la información pasada normalizada mediante regresiones lineales dinámicas. Estas redes tienen una capacidad de retroalimentación que les permite mantener una memoria oculta con información contextual relevante de las entradas actuales, lo que les otorga una memoria a largo plazo al seleccionar la información más pertinente. Estas redes neuronales se reentrenan diariamente con los datos más recientes y generan predicciones para múltiples horizontes temporales. Las predicciones obtenidas son adecuadas para pronosticar los Initial Margin en los mercados de energía, especialmente considerando la evaluación realizada con datos

[Publish your first package](#)

---

### Languages



• Python 81.1%

• Jupyter Notebook 18.9%

---

### Suggested Workflows

Based on your tech stack

 **Actions Importer** Set up

Automatically convert CI/CD files to YAML for GitHub Actions.

 **Python package** Configure

Create and test a Python package on multiple Python versions.

 **Publish Python Package** Configure

Publish a Python Package to PyPI on release.

[More workflows](#) Dismiss suggestions

# Github con el código y documentación

## Definición de funciones

---

En esta sección se presentan las funciones desarrolladas para el preprocesamiento de los datos, para el entrenamiento y para la ulterior predicción de las variables anteriormente descritas.

### Fit

---

Esta función es la encargada de entrenar la red neural y nos devuelve el modelo entrenado con el que posteriormente se realizan las predicciones de precios y *price scan ranges*. Dentro de esta función, antes de llevar a cabo el ajuste de la red neuronal, se lleva a cabo un preprocesamiento de los datos. Dicho procesamiento se compone de los siguientes pasos:

- **Selección de los datos:** se seleccionan los datos de aquellos productos que cuenten con datos durante al menos los 26 días anteriores a la fecha de predicción. Adicionalmente, con el objetivo de facilitar la computación de la red neuronal, se utilizan para el entrenamiento los datos de los últimos dos meses.
- **Estandarización:** antes del procesamiento los datos son normalizados aplicando la siguiente formula, donde  $i$  denota el momento temporal y  $j$  el producto.:

$$\frac{x_{i,j} - \bar{x}_j}{\sigma_j}$$

- **Funciones de preprocesamiento:**
  - **GetMarketData:** Es la función encargada del preprocesado, toma como inputs las variables *Price* y *PriceScanrange* y, realizando regresiones lineales, devuelve los residuos de cada valor respecto a la regresión lineal entre *PriceScanrange* y *Price*. Además, para controlar la volatilidad, se incluyen en el preprocesado la máxima variación de precios de cada producto en los últimos 25 días, sin contar los 5 más inmediatos.
  - **GetIntercomData:** Es la función encargada de preprocesar los datos del fichero de Intercommodities y que se encarga de

# Github con el código y documentación

## Definición de base de datos

---

En el siguiente documento se describen los ficheros de datos utilizados, se definen sus variables y formato.

## Intercommodity

---

Este dataframe que contiene los datos de los precios, ratios y crédito por pares de producto y fecha. El fichero cuenta con un multiindex y con 5 variables.

Los índices que estructuran el DataFrame son los siguientes:

- Fecha ['ReportDate']
- Producto A ['CombinedCommodityAproduct']
- Producto B ['CombinedCommodityBproduct']

Las variables contenidas en el DataFrame son las siguientes:

- Ratio Producto A ['RatioA']
- Ratio Producto B ['RatioB']
- Credito ['Credit']
- Precio Producto A ['PriceA']
- Precio Producto B ['PriceB']

Los datos siguen una estructura similar a de la siguiente tabla:

Indexes			Variables				
ReportDate	CombinedCommodityAproduct	CombinedCommodityBproduct	RatioA	RatioB	Credit	PriceA	Pri

# Github con el código y documentación

## Instrucciones instalación

---

### Instalación environment

---

Los paquetes y librerías necesarios para la ejecución del proyecto *Initial Marginal* se encuentran en el entorno virtual que puede encontrarse en el fichero `environment.yml` de este repositorio. para instalar el entorno virtual hay que ejecutar el siguiente código en la consola de Anaconda.

```
conda env create --file environment.yml
```



(Recomendamos ejecutar el código `.py` en VisualStudio)

### [🔗](#) Librerías necesarias

---

Para la ejecución del código es necesario instalar y cargar las siguientes librerías de Python, las versiones necesarias quedan detalladas en la información del entorno descrita anteriormente.

```
import pandas as pd
import numpy as np
import matplotlib as plt
import pylab
import os
import sklearn
import sys
import datetime
import datetime
import tensorflow
import paths
```



# Github con el código y documentación

## Ejemplo de ejecución

### Set Up

#### Cargamos el directorio donde tenemos los ficheros del repositorio

Aquí habría que cambiar la ruta por la ruta en la que habeis clonado el repositorio del Github

```
In [15]: import os
os.chdir('/Users/rgonzalez/Downloads/IberdroLaFinanzasIA4TES-main')
```

En caso de querer hacerlo desde una carpeta Drive habría que fijar el directorio de la siguiente manera:

```
In [ ]: ...
from google.colab import drive
drive.mount('/content/drive')
import os
os.chdir('/content/drive/MyDrive/Example')
...
```

#### Cargamos las librerías necesarias

```
In [17]: import pandas as pd
import numpy as np
import matplotlib as plt
import pylab
import os
import sklearn
import sys
import datetime
import datetime
import tensorflow
#import paths
import scipy
import tqdm
```

# Conclusiones

- Métodos de **aprendizaje online probabilístico** que se adaptan a los cambios temporales
- Las técnicas desarrolladas tienen **garantías teóricas** de rendimiento
- Pasos futuros: **búsqueda financiación** para continuar con la línea iniciada sobre predicciones energéticas
- Implementación de los métodos desarrollados accesible al público en los repositorios:
  - <https://github.com/MachineLearningBCAM/Load-forecasting-IEEE-TPWRS-2020>
  - <https://github.com/MachineLearningBCAM/MRCpy>