



GRADO EN INGENIERÍA EN TECNOLOGÍAS INDUSTRIALES

TRABAJO FIN DE GRADO MODELO DE PREDICCIÓN EN LOS MERCADOS DE AJUSTE

Autor: Lucas Otero-Novas Gómez

Director: Andrés Fernández Ramos

Madrid

Declaro, bajo mi responsabilidad, que el Proyecto presentado con el título

Modelo de predicción en los mercados de ajuste

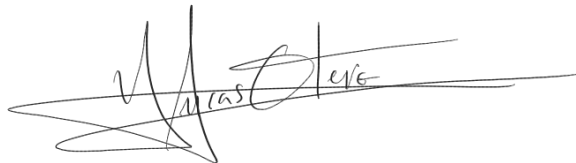
en la ETS de Ingeniería - ICAI de la Universidad Pontificia Comillas en el

curso académico 2021/2022 es de mi autoría, original e inédito y

no ha sido presentado con anterioridad a otros efectos.

El Proyecto no es plagio de otro, ni total ni parcialmente y la información que ha sido

tomada de otros documentos está debidamente referenciada.



Fdo.: Lucas Otero-Novas Gómez

Fecha: 31 / 08 / 2022

Autorizada la entrega del proyecto

EL DIRECTOR DEL PROYECTO

Fdo.: Andrés Fernández Ramos

Fecha: 31 / 08 / 2022





COMILLAS

UNIVERSIDAD PONTIFICIA

ICAI

GRADO EN INGENIERÍA EN TECNOLOGÍAS
INDUSTRIALES

TRABAJO FIN DE GRADO
MODELO DE PREDICCIÓN EN LOS MERCADOS
DE AJUSTE

Autor: Lucas Otero-Novas Gómez

Director: Andrés Fernández Ramos

Madrid

MODELO DE PREDICCIÓN EN LOS MERCADOS DE AJUSTE

Autor: Otero-Novas Gómez, Lucas.

Director: Fernández Ramos, Andrés.

Entidad Colaboradora: Invesyde S.L

RESUMEN DEL PROYECTO

Análisis y modelos de predicción de los mercados intradiarios en el mercado eléctrico español, estudiando el comportamiento de los agentes de mercado y su influencia sobre éste. Se desarrollan tanto modelos fundamentales como modelos basados en redes neuronales.

Palabras clave: intradiario, mercado, modelo, predicción

1. Introducción

El mercado eléctrico español está compuesto de un mecanismo que permite la libre competencia de los agentes a la vez que se asegura el cumplimiento de un suministro con ciertos requisitos de calidad y seguridad. Cada día vendedores y compradores establecen sus ofertas para cada hora del día siguiente a través de la OMIE. Una vez presentadas, se ordenan por el algoritmo EUPHEMIA para obtener las curvas de oferta y demanda y poder concretar el precio del mercado al próximo día. Con este precio y la suma de otros costes como la potencia contratada, los costes regulados y los impuestos establecidos por el gobierno, las comercializadoras establecen los precios de venta al consumidor final.

Este mercado esta formado principalmente por tres elementos principales: el mercado diario, el intradiario y el de los servicios complementarios.

- El mercado diario: corresponde al grupo de transacciones compraventa del día siguiente. En este mercado los productores ofertan la cantidad de energía que están dispuestos a generar al día siguiente a un precio. Por otro lado, comercializadoras y consumidores directos establecen la cantidad de energía que requieren y el precio al que estarían dispuestos a pagar. Estas ofertas se realizan para cada una de las 24 horas del día siguiente. Una vez obtenidas todas las ofertas, se realiza la casación, partiendo de la oferta de venta más barata hasta igualarla a la demanda.
- El mercado intradiario: una vez terminado el programa diario debido a los ajustes necesarios que deben realizar los agentes en sus ofertas o demandas se realizan los mercados intradiarios, cuyo funcionamiento es similar al diario. Se organiza en seis sesiones, las cuales cambiarán el resultado del diario final al día siguiente. Solo podrán participar en el intradiario aquellos agentes que ya hayan participado en el diario previamente. Existen varios mercados intradiarios que se conciertan a lo largo del día anterior. Los horarios seguidos son:

	SESIÓN 1ª	SESIÓN 2ª	SESIÓN 3ª	SESIÓN 4ª	SESIÓN 5ª	SESIÓN 6ª
Apertura de Sesión	14:00	17:00	21:00	1:00	4:00	9:00
Cierre de Sesión	15:00	17:50	21:50	1:50	4:50	9:50
Casación	15:00	17:50	21:50	1:50	4:50	9:50
Publicación del programa acumulado (PIBCA)	15:07	17:57	21:57	1:57	4:57	9:57
Publicación PHF de los OSs	16:20	18:20	22:20	2:20	5:20	10:20
Horizonte de Programación (Períodos horarios)	24 horas (1-24 D+1)	28 horas (21-24 y 1-24 D+1)	24 horas (1-24 D+1)	20 horas (5-24)	17 horas (8-24)	12 horas (13-24)

Tabla 1- Horario de los mercados intradiarios

- El mercado de servicios de balance: la finalidad de este mercado es resolver los desequilibrios que puedan surgir entre generación y demanda, además de otros servicios complementarios.

2. Definición del proyecto

La posibilidad de predicción del comportamiento del mercado eléctrico, como de cualquier otro mercado, supone tener una capacidad de anticipación gracias a la cual se pueden obtener beneficios regularmente. Actualmente existen numerosos modelos con la capacidad de predecir el precio spot del mercado diario. Sin embargo, no es el caso con los mercados intradiarios, ya que éstos dependen más del comportamiento de los agentes, que de condiciones ajenas a ellos como puede ser la previsión del tiempo, la predicción de la demanda, etc. Por ello, en este proyecto se intentará predecir el desenlace de los intradiarios basándose en el estudio de las estrategias e impacto que tienen los agentes de mercado.

3. Modelos fundamentales

Estos modelos están basados en el análisis de las estrategias diarias de los agentes utilizando los ficheros I90DIA y el histórico del precio spot diario como referencia. Todos tienen como finalidad predecir el comportamiento del mercado del día siguiente conociendo los movimientos de los agentes del día actual, bajo la suposición de que los movimientos que realizarán los agentes el día posterior serán similares a los del día actual.

Se realizan tres modelos fundamentales:

- Primer modelo: se intenta predecir el precio mediante la creación de un ranking según el porcentaje de veces que las unidades de programación venden barato y compran caro.
- Segundo modelo: se basa en los movimientos de los agentes más influyentes presentes en el día, considerando agente influyente aquel que tenga un alto valor de sensibilidad corregida, denominada sensibilidad relativa. La sensibilidad es un parámetro que se calcula previamente a los modelos, se trata de un indicador

de la influencia que ha tenido cada agente en el precio con los movimientos históricos que ha realizado.

- Tercer modelo: se basa en el porcentaje de agentes categorizados como peligrosos y no peligrosos presentes ese día, una UP se considera más peligrosa cuanto mayor sea su valor de sensibilidad cuando el mercado encarece. Por el contrario, será menos peligrosa cuanto mayor sea su valor de sensibilidad cuando el mercado abarata.

4. Modelo basado en redes neuronales

Tras comprobar que los valores calculados en los modelos fundamentales para la predicción no tienen una relación lineal con la misma, o pueden realizar la clasificación de la información en reducidas ocasiones, se decide emplear una red neuronal para intentar buscar si estos valores mantienen alguna relación no lineal respecto a si el mercado encarece o no. Por ello, como entradas de esta red neuronal, se utilizarán los valores resultantes calculados en cada uno de los modelos fundamentales, además de otras entradas que puedan tener relación con el comportamiento del mercado.

5. Resultados

Resultados de la predicción del mercado intradiario 2 por horas el 9 de mayo de 2021. El modelo genera las probabilidades estimadas de si el mercado encarece (1) o no (0).

Intradiario 2, 9 Mayo 2021								
Hour	Prob de 0	Prob de 1	Pred T = 0.5	Result	True Positive	True Negative	False Positive	False Negative
00:01	0.99978824	0.00021176	0	0	0	1	0	0
01:02	6.7082E-07	0.99999933	1	1	1	0	0	0
02:03	0.99997125	2.8751E-05	0	0	0	1	0	0
03:04	0.99999648	3.5209E-06	0	0	0	1	0	0
04:05	0.99999755	2.4492E-06	0	1	0	0	0	1
05:06	1.2667E-06	0.99999873	1	1	1	0	0	0
06:07	0.27228057	0.72771943	1	1	1	0	0	0
07:08	0.99997095	2.9049E-05	0	1	0	0	0	1
08:09	0.00077242	0.99922758	1	1	1	0	0	0
09:10	0.02381544	0.97618456	1	1	1	0	0	0
10:11	0.99877886	0.00122114	0	1	0	0	0	1
11:12	0.86392693	0.13607307	0	0	0	1	0	0
12:13	0.00202068	0.99797932	1	1	1	0	0	0
13:14	3.2898E-06	0.99999671	1	1	1	0	0	0
14:15	0.999998	1.9968E-06	0	0	0	1	0	0
15:16	0.9997224	0.0002776	0	0	0	1	0	0
16:17	0.99654248	0.00345752	0	0	0	1	0	0
17:18	0.14698937	0.85301063	1	0	0	0	1	0
18:19	2.7419E-05	0.99997258	1	1	1	0	0	0
19:20	2.0234E-05	0.99997977	1	1	1	0	0	0
20:21	9.9945E-06	0.99999001	1	1	1	0	0	0

21:22	0.99999964	3.6027E-07	0	0	0	1	0	0
22:23	0.99999859	1.4118E-06	0	0	0	1	0	0
23:24	0.99125039	0.00874961	0	0	0	1	0	0

10	10	1	3
----	----	---	---

Tabla 2- Ejemplo de simulación del mercado intradiario 2 el 9 de mayo de 2021

Se obtienen los siguientes porcentajes de acierto:

Porcentaje de acierto a que el mercado encarece = 90.91 %

Porcentaje de acierto a que el mercado abarata = 76.92 %

Porcentaje de acierto medio ponderado = 83.33 %

6. Conclusiones

- La predicción de los mercados intradiarios es una tarea más compleja que la del mercado diario, al tener más peso el comportamiento de los agentes.
- Aunque ningún modelo fundamental pueda predecir el comportamiento de los mercados intradiarios de forma frecuente y acertada, se han encontrado relaciones no lineales entre los valores calculados y el desenlace de los mercados, por ello se decide realizar una red neuronal que utilice estos valores para hacer las predicciones.
- Se puede considerar que la red neuronal es capaz de predecir el desenlace de los mercados intradiarios, aunque con precisión variante en función del día y mercado, fluctuando entre el 65 % y el 84 %.

7. Referencias

- [1] [1] Bendezu, K.. "DIAGRAMA UML Y ARQUITECTURA DEL SISTEMA". Sistemas Distribuidos 2013. Febrero, 2013. <http://comparape.blogspot.com.es/2013/02/diagrama-uml-y-arquitectura-del-sistema.html>.
- [2] ¿Cómo funciona el mercado eléctrico español? (2021, 9 junio). CYE. <https://www.cye-energia.com/como-funciona-el-mercado-electrico-espanol/#:%7E:text=El%20Mercado%20Ib%C3%A9rico%20Espa%C3%B1ol%20es,d%C3%ADa%20a%20trav%C3%A9s%20del%20MIE>
- [3] *Euphemia, el algoritmo que establece el precio de la luz (electricidad)*. (2021, 22 agosto). Naukas. <https://naukas.com/2021/08/23/euphemia-el-algoritmo-que-establece-el-precio-de-la-luz-electricidad/>

PREDICTION MODEL FOR THE SPANISH BALANCING MARKET

Author: Otero-Novas Gómez, Lucas.

Supervisor: Fernández Ramos, Andrés.

Collaborating Entity: Invesyde S.L

ABSTRACT

Analysis and prediction models of intraday markets in the Spanish electricity market, studying the behavior of market agents and their influence on it. Both fundamental and based on neural networks models are developed.

Keywords: model, prediction, intraday, market

1. Introduction

The Spanish electricity market is made up of a mechanism that allows agents to compete freely while ensuring that a supply meets certain quality and safety requirements. Everyday sellers and buyers establish their offers for each hour of the following day through the OMIE. Once presented, they are ordered by the EUPHEMIA algorithm to obtain the supply and demand curves and to be able to specify the market price for the next day. With this price and the sum of other costs such as the contracted power, the regulated costs and the taxes established by the government, the marketers establish the sale prices to the final consumer.

This market is mainly made up of three main elements: the daily market, the intraday market and the complementary services market.

- The daily market: corresponds to the group of trading transactions of the following day. In this market, producers offer the amount of energy they are willing to generate the next day at a price. On the other hand, marketers and direct consumers establish the amount of energy they require and the price at which they would be willing to pay. These offers are made for each of the 24 hours of the following day. Once all the offers have been obtained, the matching is carried out, starting from the cheapest sale offer until it is equal to the demand.
- The intraday market: once the daily program has finished due to the necessary adjustments that the agents must make in their bids or requests, the intraday markets are carried out, whose operation is similar to the daily one. It is organized in six sessions, which will change the result of the final diary the next day. Only those agents who have

	SESIÓN 1ª	SESIÓN 2ª	SESIÓN 3ª	SESIÓN 4ª	SESIÓN 5ª	SESIÓN 6ª
Apertura de Sesión	14:00	17:00	21:00	1:00	4:00	9:00
Cierre de Sesión	15:00	17:50	21:50	1:50	4:50	9:50
Casación	15:00	17:50	21:50	1:50	4:50	9:50
Publicación del programa acumulado (PIBCA)	15:07	17:57	21:57	1:57	4:57	9:57
Publicación PHF de los OSs	16:20	18:20	22:20	2:20	5:20	10:20
Horizonte de Programación (Periodos horarios)	24 horas (1-24 D+1)	28 horas (21-24 y 1-24 D+1)	24 horas (1-24 D+1)	20 horas (5-24)	17 horas (8-24)	12 horas (13-24)

Tabla 3 - Intraday Spanish Electric Market Schedule

previously participated in the daily may participate in the intraday. There are several intraday markets that take place throughout the previous day. The hours followed are:

2. Project Definition

The possibility of predicting the behavior of the electricity market, like any other market, implies having a capacity for anticipation thanks to which profits can be obtained regularly. Currently there are numerous models with the ability to predict the spot price of the daily market. However, this is not the case with intraday markets, since they depend more on the behavior of the agents than on conditions unrelated to them, such as weather forecasts, demand forecasts, etc. For this reason, in this project we will try to predict the outcome of intraday trading based on the study of the strategies and impact of market agents.

3. Fundamental models

These models are based on the analysis of the agents' daily strategies using the I90DIA files and the daily spot price history as a reference. The purpose of all of them is to predict the behavior of the market on the following day knowing the movements of the agents of the current day, under the assumption that the movements that the agents will make the following day will be similar to those of the current day.

The three fundamental models are:

- First model: an attempt is made to predict the price by creating a ranking according to the percentage of times that the agents sell low and buy high.
- Second model: it is based on the movements of the most influential agents present on the day, considering an influential agent to be one with a high value of corrected sensitivity, called relative sensitivity. Sensitivity is a parameter that is calculated prior to the models, it is an indicator of the influence that each agent has had on the price with the historical movements that it has made.
- Third model: based on the percentage of agents categorized as dangerous and non-dangerous present that day, an agent is considered more dangerous the higher its sensitivity value when the market becomes more expensive. On the contrary, it will be less dangerous the higher its sensitivity value when the market cheapens.

4. Model based on Neural Networks

After verifying that the values calculated in the fundamental models for prediction do not have a linear relationship with it, or can perform the classification of the information on a few occasions, it is decided to use a neural network to try to find if these values maintain some relationship not linear with respect to whether the market becomes more expensive or not. Therefore, as inputs to this neural network, the resulting values calculated in each of the fundamental models will be used, in addition to other inputs that may be related to market behavior.

5. Results

Results of the prediction of the intraday market 2 per hour on May 9, 2021. The model generates the estimated probabilities of whether the market becomes more expensive (1) or not (0).

Intradiario 2, 9 Mayo 2021								
Hour	Prob de 0	Prob de 1	Pred T = 0.5	Result	True Positive	True Negative	False Positive	False Negative
00:01	0.99978824	0.00021176	0	0	0	1	0	0
01:02	6.7082E-07	0.99999933	1	1	1	0	0	0
02:03	0.99997125	2.8751E-05	0	0	0	1	0	0
03:04	0.99999648	3.5209E-06	0	0	0	1	0	0
04:05	0.99999755	2.4492E-06	0	1	0	0	0	1
05:06	1.2667E-06	0.99999873	1	1	1	0	0	0
06:07	0.27228057	0.72771943	1	1	1	0	0	0
07:08	0.99997095	2.9049E-05	0	1	0	0	0	1
08:09	0.00077242	0.99922758	1	1	1	0	0	0
09:10	0.02381544	0.97618456	1	1	1	0	0	0
10:11	0.99877886	0.00122114	0	1	0	0	0	1
11:12	0.86392693	0.13607307	0	0	0	1	0	0
12:13	0.00202068	0.99797932	1	1	1	0	0	0
13:14	3.2898E-06	0.99999671	1	1	1	0	0	0
14:15	0.999998	1.9968E-06	0	0	0	1	0	0
15:16	0.9997224	0.0002776	0	0	0	1	0	0
16:17	0.99654248	0.00345752	0	0	0	1	0	0
17:18	0.14698937	0.85301063	1	0	0	0	1	0
18:19	2.7419E-05	0.99997258	1	1	1	0	0	0
19:20	2.0234E-05	0.99997977	1	1	1	0	0	0
20:21	9.9945E-06	0.99999001	1	1	1	0	0	0
21:22	0.99999964	3.6027E-07	0	0	0	1	0	0
22:23	0.99999859	1.4118E-06	0	0	0	1	0	0
23:24	0.99125039	0.00874961	0	0	0	1	0	0

10	10	1	3
----	----	---	---

Tabla 4- Intradary Market 2 Simulation Example

The following success percentages are obtained

Successful predictions when predicted that the market becomes more expensive = 90.91 %

Successful predictions when predicted that the market becomes less expensive = 76.92 %

Average successful predictions = 83.33 %

8. Conclusions

The prediction of the intraday markets is a more complex task than that of the daily market, since the behavior of the agents has more weight.

- Although no fundamental model can predict the behavior of the intraday markets frequently and accurately, non-linear relationships have been found between the calculated values and the outcome of the markets, which is why it is decided to create a neural network that uses these values to make the predictions.
- It can be considered that the neural network is capable of predicting the outcome of the intraday markets, although with variable accuracy depending on the day and market, fluctuating between 65% and 84%.

9. References

- [4] Bendezu, K.. "DIAGRAMA UML Y ARQUITECTURA DEL SISTEMA". Sistemas Distribuidos 2013. Febrero, 2013. <http://comparape.blogspot.com.es/2013/02/diagrama-uml-y-arquitectura-del-sistema.html>.
- [5] *¿Cómo funciona el mercado eléctrico español?* (2021, 9 junio). CYE. <https://www.cye-energia.com/como-funciona-el-mercado-electrico-espanol/#:%7E:text=El%20Mercado%20Ib%C3%A9rico%20Espa%C3%B1ol%20es,d%C3%ADa%20a%20trav%C3%A9s%20del%20OMIE>
- [6] *Euphemia, el algoritmo que establece el precio de la luz (electricidad)*. (2021, 22 agosto). Naukas. <https://naukas.com/2021/08/23/euphemia-el-algoritmo-que-establece-el-precio-de-la-luz-electricidad/>

Índice de la memoria

Capítulo 1. Introducción	8
1.1 El mercado eléctrico	8
1.2 Modelos basados en redes neuronales	10
1.2.1 Función de activación	11
1.3 Entrenamiento de una red neuronal	13
1.3.1 Backpropagation	13
Capítulo 2. Descripción de las Tecnologías.....	16
2.1 Descarga de datos desde E•SIOS	16
2.1.1 I90DIA	16
2.1.2 CSV Histórico Precio Spot	19
2.1.3 CSV Información Adicional UP	20
2.2 Gestión de los datos.....	21
2.2.1 Python.....	21
2.2.2 Tablas Dinámicas Excel.....	22
Capítulo 3. Estado de la Cuestión	23
3.1 Modelos existentes	23
3.1.1 Modelo de previsión del precio del mercado diario de la electricidad en el corto plazo	23
3.1.2 Desarrollo de un modelo de predicción del precio de la energía eléctrica para el mercado a plazo mediante redes neuronales.	24
3.1.3 Econometric modelling and forecasting of intraday electricity prices.....	25
3.1.4 Short-Term Price Forecasting Models Based on Artificial Neural Networks for Intraday Sessions in the Iberian Electricity Market.....	25
3.2 Conclusiones del Estado de la Cuestión.....	26
Capítulo 4. Definición del Trabajo	27
4.1 Justificación.....	27
4.2 Objetivos	27
4.3 Metodología.....	28
Capítulo 5. Infraestructura de los datos	30
5.1 Adaptación de formatos de los archivos.....	30

5.1.1	<i>I90DIA</i>	30
5.1.2	<i>CSV Información Adicional UP</i>	31
5.2	Esquema Infraestructura de datos	31
5.2.1	<i>Archivos CSV procedentes de I90DIA</i>	32
5.2.2	<i>Archivo CSV Información Adicional UP</i>	38
5.2.3	<i>Archivo CSV Precio Spot Histórico</i>	38
Capítulo 6. Análisis previos		40
6.1	Volumen de movimientos por UP	40
6.2	Histórico Unidades de Programación.....	41
Capítulo 7. Modelos fundamentales		43
7.1	Introducción a los modelos.....	43
7.1.1	<i>Estrategia utilizada</i>	43
7.2	Parámetros calculados	44
7.2.1	<i>Relación Movimiento – Precio</i>	44
7.2.2	<i>Sensibilidad del precio respecto a los movimientos</i>	45
7.2.3	<i>Función para calcular los Parámetros</i>	46
7.2.4	<i>Análisis de los parámetros resultantes</i>	47
7.3	Primer Modelo – Rankings por tendencias	52
7.3.1	<i>Cálculos realizados sobre los movimientos del día actual</i>	52
7.3.2	<i>Predicción en base al valor resultante</i>	54
7.3.3	<i>Resultados</i>	54
7.4	Segundo Modelo – Agentes Influyentes.....	56
7.4.1	<i>Cálculos realizados sobre los movimientos del día actual</i>	57
7.4.2	<i>Predicción en base al valor resultante</i>	57
7.4.3	<i>Resultados</i>	58
7.5	Tercer modelo – Agentes peligrosos vs no peligrosos	61
7.5.1	<i>Cálculos realizados sobre los movimientos del día actual</i>	61
7.5.2	<i>Predicción en base a los porcentajes obtenidos</i>	62
7.5.3	<i>Resultados</i>	62
Capítulo 8. Modelo basado en Redes Neuronales		65
8.1	Introducción al Modelo	65
8.2	Variables de entrada	65

8.2.1 Valores resultantes de los modelos fundamentales	65
8.2.2 Otras entradas.....	65
8.2.3 Preparación de los datos de entrada.....	66
8.3 Diseño de la red neuronal	66
8.3.1 Tuning de Hiperparámetros	67
8.3.2 Resultados tuning de hiperparámetros.....	70
8.3.3 Resultados del modelo.....	71
8.3.4 Simulación situación real.....	75
Capítulo 9. Conclusiones y Trabajos Futuros.....	79
9.1 Conclusiones	79
9.2 Trabajos futuros.....	79
Capítulo 10. Bibliografía.....	81

Índice de figuras

Ilustración 1- Esquema transmisión de información en red neuronal. Fuente: towardsdatascience	11
Ilustración 2- Representación función tanh. Fuente: Wikipedia	12
Ilustración 3 - Representación arctan. Fuente: Wikipedia.....	12
Ilustración 4- Representación de backpropagation. Fuente: Medium	13
Ilustración 5 - Representación visual del funcionamiento de SGD.....	14
Ilustración 6 - Índice hoja de cálculo I90DIA	17
Ilustración 7 - Punto de descarga de archivos precio spot en la página web oficial de E•SIOS	20
Ilustración 8 - Esquema arquitectura carpeta principal I90DIA.....	30
Ilustración 9 - Esquema Infraestructura de datos	32
Ilustración 10 - Tabla explicativa tipos de fallos.....	35
Ilustración 11- Explicación cálculo archivo CSV precios.....	39
Ilustración 12 - Esquema Estrategia Modelos Fundamentales	43
Ilustración 13 - Esquema descriptivo relación precio – movimientos	44
Ilustración 14 - Esquema explicativo criterio cálculo sensibilidad relativa	46
Ilustración 15 - Estudio de los porcentajes compra-vende barato o caro de las unidades de programación que tienen un mayor volumen	47
Ilustración 16 - Unidades de programación que tienen el mayor valor absoluto de sensibilidad relativa	48
Ilustración 17 - Unidades de programación que tienen el mayor valor absoluto de sensibilidad.....	49
Ilustración 18 - Tipo de producción de las unidades con mayor sensibilidad.....	49
Ilustración 19- Diagrama de las unidades con mayor sensibilidad cuando el mercado encarece	50
Ilustración 20 - Tipo de producción de las unidades con mayor sensibilidad cuando el mercado encarece	51

Ilustración 21- Diagrama de las unidades con mayor sensibilidad cuando el mercado abarata	51
Ilustración 22 - Tipo de producción de las unidades con mayor sensibilidad cuando el mercado abarata.....	52
Ilustración 23- Parcelas de violín de los valores resultantes del primer modelo.....	55
Ilustración 24 - Diagrama de barras del valor medio resultante diario del primer modelo.	56
Ilustración 25 - Parcelas de violín de los valores resultantes del segundo modelo	58
Ilustración 26 - Zoom Parcelas de violín de los valores resultantes del primer modelo	59
Ilustración 27 - - Diagrama de barras del valor medio resultante diario del segundo modelo	60
Ilustración 28 - Parcelas de violín de los porcentajes resultantes del tercer modelo cuando el mercado encarece.....	62
Ilustración 29 - Parcelas de violín de los porcentajes resultantes del tercer modelo cuando el mercado abarata.....	63
Ilustración 30 - Parcelas de violín de los porcentajes resultantes de las unidades coincidentes del tercer modelo	64
Ilustración 31 - Estrategia Modelo basado en redes neuronales.....	66
Ilustración 32 - Estrategia de tuning de hiperparámetros	68
Ilustración 33 - Ejemplo de “Grid Search”	68
Ilustración 34 - Relación Precision recall en el set de Train	70
Ilustración 35 - Matriz de confusión Test.....	71
Ilustración 36 - Curva Precision recall set Test.....	72
Ilustración 37 - Roc Curve Test.....	72
Ilustración 38 - Matriz de confusión Validación	74
Ilustración 39 - Resultados performance validación	74

Índice de tablas

Tabla 1- Horario de los mercados intradiarios	8
Tabla 2- Ejemplo de simulación del mercado intradiario 2 el 9 de Mayo de 2021	10
Tabla 3 - Intradiary Spanish Electric Market Schedule.....	12
Tabla 4- Intradiary Market 2 Simulation Example.....	13
Tabla 5- Horario de los mercados intradiarios	9
Tabla 6 - Pestañas I90DIA Relevantes	18
Tabla 7 - Extracto de tabla I90DIA01 del archivo I90DIA_20201101	19
Tabla 8 - Extracto de I90DIA26, de la hoja de cálculo I90DIA_20201101	19
Tabla 9 - – Listado de campos disponibles en CSV información adicional UP.....	21
Tabla 10 - Ejemplo de visualización de dataframe en Pycharm Community Edition.....	22
Tabla 11 – Ejemplo de CSV desagregado	33
Tabla 12 - Ejemplo de CSV SYB Completo	37
Tabla 13 - Ejemplo de CSV Desagregado.....	37
Tabla 14 - Ejemplo de CSV SYB.....	38
Tabla 15 - Extracto del archivo sum_syb_intra1_intra2_intra3_from_20201101_to_20211118.csv, resultado de un estudio de volumen	41
Tabla 16 - Tabla del archivo “ABA1_rh_restricciones_intra1_intra2_intra3_intra4_intra5_intra6_intra7_syb_from_20201101_to_20201118.csv”, resultado de la función histórico unidades de programación	42
Tabla 17 - Criterio asignación de valores por rankings.....	53
Tabla 18 - Mejores Combinaciones Tuning Hiperparámetros	71
Tabla 19 - Resultados Performance Test.....	73
Tabla 20 - Tabla de resultados simulación Intradiario 1	76
Tabla 21 - Tabla de resultados simulación Intradiario 2	77
Tabla 22 - Tabla de resultados simulación Intradiario 3	78

Capítulo 1. INTRODUCCIÓN

1.1 EL MERCADO ELÉCTRICO

El mercado eléctrico español empieza el 1 de enero de 1998, cuando el Gobierno de España aprueba la Ley 54/1997, empezando así el proceso de liberalización del sector eléctrico. Dicho mercado está compuesto de un mecanismo que permite la libre competencia de los agentes a la vez que se asegura el cumplimiento de un suministro con ciertos requisitos de calidad y seguridad.

Los principales agentes implicados en el mercado eléctrico español son los siguientes:

- Productores de energía [mercado libre]: aquellas empresas que se encargan de la generación de energía eléctrica.
- Transporte de la energía [mercado regulado]: regulado por Red Eléctrica de España (REE). Se encarga del transporte de la energía generado por los productores a alta tensión y largas distancias hasta las subestaciones.
- Distribuidoras de Energía [mercado regulado]: se encargan de la distribución de la energía desde las subestaciones hasta los puntos de consumo. Además, se encargan también del mantenimiento de estas redes de distribución para garantizar la calidad de estas mismas y de tomar las medidas en los puntos de suministro. El 90 % de este sector está compuesto por las 5 principales distribuidoras.
- Comercializadoras del mercado [mercado libre]: son los intermediarios entre el cliente final y los agentes del mercado. Compran la energía al Operador del Mercado Ibérico Español y la venden a los consumidores directamente, estableciendo los precios con libertad.

El Operador del Mercado Ibérico Español se encarga de supervisar la compraventa de la energía siendo el punto de encuentro entre las comercializadores y los productores. Estos precios son abiertos y transparentes a cualquier usuario.

Cada día vendedores y compradores establecen sus ofertas para cada hora del día siguiente a través de la OMIE. Una vez presentadas, se ordenan por el algoritmo EUPHEMIA para obtener las curvas de oferta y demanda y poder concretar el precio del mercado al próximo día. Con este precio y la suma de otros costes como la potencia contratada, los costes regulados y los impuestos establecidos por el gobierno, las comercializadoras establecen los precios de venta al consumidor final.

Este mercado está formado principalmente por tres elementos principales: el mercado diario, el intradiario y el de los servicios complementarios.

- El mercado diario: corresponde al grupo de transacciones compraventa del día siguiente. En este mercado los productores ofertan la cantidad de energía que están dispuestos a generar al día siguiente a un precio. Por otro lado, comercializadoras y consumidores directos establecen la cantidad de energía que requieren y el precio al que estarían dispuestos a pagar. Estas ofertas se realizan para cada una de las 24 horas del día siguiente. Una vez obtenidas todas las ofertas, se realiza la casación, partiendo de la oferta de venta más barata hasta igualarla a la demanda.
- El mercado intradiario: una vez terminado el programa diario debido a los ajustes necesarios que deben realizar los agentes en sus ofertas o demandas se realizan los mercados intradiarios, cuyo funcionamiento es similar al diario. Se organiza en seis sesiones, las cuales cambiarán el resultado del diario final al día siguiente. Solo podrán participar en el intradiario aquellos agentes que ya hayan participado en el diario previamente. Existen varios mercados intradiarios que se concertan a lo largo del día anterior. Los horarios seguidos son:

	SESIÓN 1ª	SESIÓN 2ª	SESIÓN 3ª	SESIÓN 4ª	SESIÓN 5ª	SESIÓN 6ª
Apertura de Sesión	14:00	17:00	21:00	1:00	4:00	9:00
Cierre de Sesión	15:00	17:50	21:50	1:50	4:50	9:50
Casación	15:00	17:50	21:50	1:50	4:50	9:50
Publicación del programa acumulado (PIBCA)	15:07	17:57	21:57	1:57	4:57	9:57
Publicación PHF de los OSs	16:20	18:20	22:20	2:20	5:20	10:20
Horizonte de Programación (Periodos horarios)	24 horas (1-24 D+1)	28 horas (21-24 y 1-24 D+1)	24 horas (1-24 D+1)	20 horas (5-24)	17 horas (8-24)	12 horas (13-24)

Tabla 5- Horario de los mercados intradiarios

- El mercado de servicios de balance: la finalidad de este mercado es resolver los desequilibrios que puedan surgir entre generación y demanda, además de otros servicios complementarios.

Llegados a este punto es interesante mencionar el comportamiento de varios agentes específicos. Las centrales nucleares suelen realizar sus ofertas a 0, asegurando de esta forma la venta de toda la energía que ofertan, ya que es una energía cuya cantidad es difícil de regular a corto plazo. De esta forma, dejan que el comportamiento de otros agentes influya en el precio de casación final, precio al que acabarán vendiendo su energía. Además de la nuclear, otros agentes como los correspondientes a solar o fotovoltaica, también suelen ofertar a 0, pues el coste del viento o el sol es 0 y es una energía que no se puede almacenar, por ello les conviene vender toda la energía que son capaces de producir.

Otro aspecto a tener en cuenta es el “Market Splitting”, esto se da cuando se sobrepasa la capacidad de interconexión entre los mercados español y portugués, forzando así la separación de ambos mercados y posterior resolución independiente, llegando cada uno a su casación, o cuando no se cumplen las condiciones complejas (las ofertas pueden estar condicionadas por una serie de causas, como por ejemplo una condición de indivisibilidad). Las consecuencias de estos problemas sería el desplazamiento de la curva de demanda o de la curva de oferta respectivamente.

Además del Mercado Ibérico de Energía, los agentes podrán llevar a cabo contratos bilaterales, es decir, acuerdos en los cuales productor y consumidor acuerden de forma independiente una potencia a un precio.

Los programas se obtienen después de cada mercado, plasmando la cantidad de energía que se comprometen a comprar o vender los agentes tras el mercado. Por ejemplo, tras el mercado diario el PBDF, tras el intradiario 1, el PHF1.

1.2 MODELOS BASADOS EN REDES NEURONALES

El funcionamiento de las redes neuronales, como su nombre indica, replica el funcionamiento del cerebro de organismos vivos, es decir, se tratan de redes de neuronas

interconectadas entre sí transmitiendo información. De forma similar a un cerebro, estas redes aprenden con experiencia por sí solas.

Las neuronas de estas redes se estructuran por capas: una capa de entrada, una de salida, y las capas intermedias. La información se transmite a través de pesos y sesgos entre ellas.

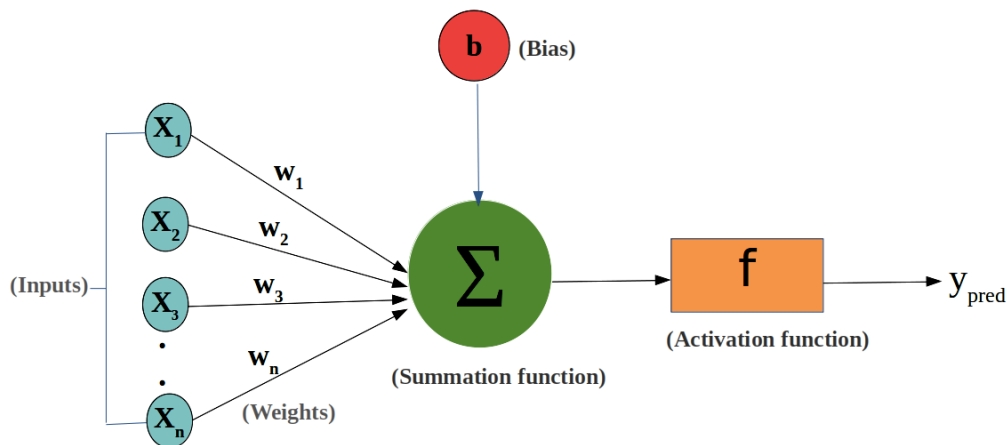


Ilustración 1- Esquema transmisión de información en red neuronal. Fuente: towardsdatascience

Se puede observar en la Figura 1 cómo se transmite la información de una capa, a una neurona de la siguiente capa. Las funciones de activación que se llevan a cabo tras el sumatorio de los valores de la capa anterior multiplicados por los pesos y sumados al sesgo otorgan a estas redes el carácter la posibilidad aprender relaciones no lineales.

1.2.1 FUNCIÓN DE ACTIVACIÓN

Como ya se ha mencionado, su característica principal es la de otorgar la no linealidad al modelo. Pero también tiene otras dos características fundamentales:

- Rango de la función: son funciones de rango finito, lo que facilita el trabajo de los algoritmos de descenso para el aprendizaje.
- Son diferenciables, permitiendo los algoritmos basados en descenso de gradiente.

Existen muchos tipos de funciones de activación en función de la finalidad que tenga la red, en este proyecto, se planteará el uso de los siguientes:

1.2.1.1 Tangente Hiperbólica

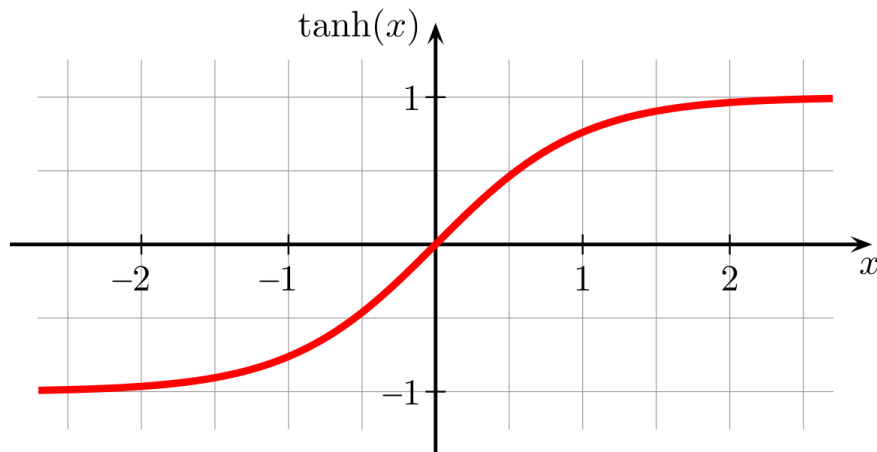


Ilustración 2- Representación función tanh. Fuente: Wikipedia

1.2.1.2 Arcotangente

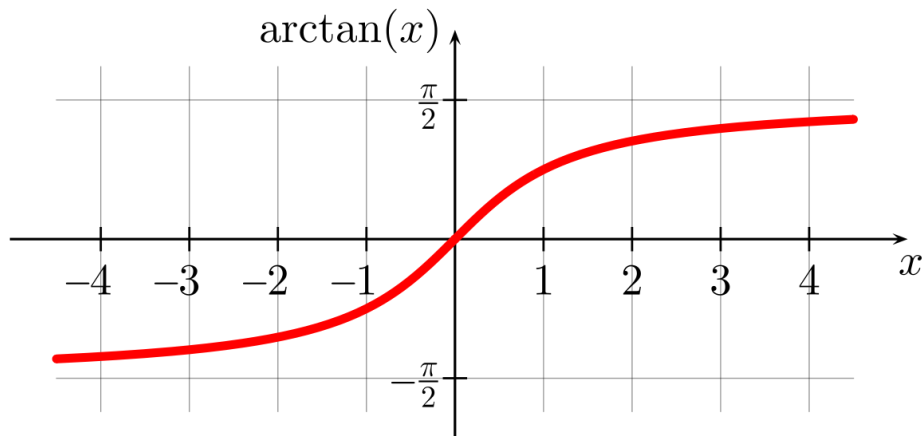


Ilustración 3 - Representación arctan. Fuente: Wikipedia

La característica principal que une a ambas funciones como candidatas frente al resto es la capacidad de trasladar valores negativos a las siguientes neuronas, aspecto importante en los modelos a desarrollar más adelante.

1.3 ENTRENAMIENTO DE UNA RED NEURONAL

El entrenamiento de una red consiste en ajustar el valor de los pesos y sesgos para minimizar el error cometido en la predicción (determinado por la función de coste). Para minimizar este error existen diferentes algoritmos, todos ellos parten del proceso denominado “backpropagation”.

1.3.1 BACKPROPAGATION

El funcionamiento consiste en el cálculo de la derivada del error respecto a cada peso, identificando cuánto afecta ese peso al valor obtenido. Estas derivadas se van calculando realizando un proceso llamado “backpropagation”, teniendo en cuenta también la derivación de la función de activación.

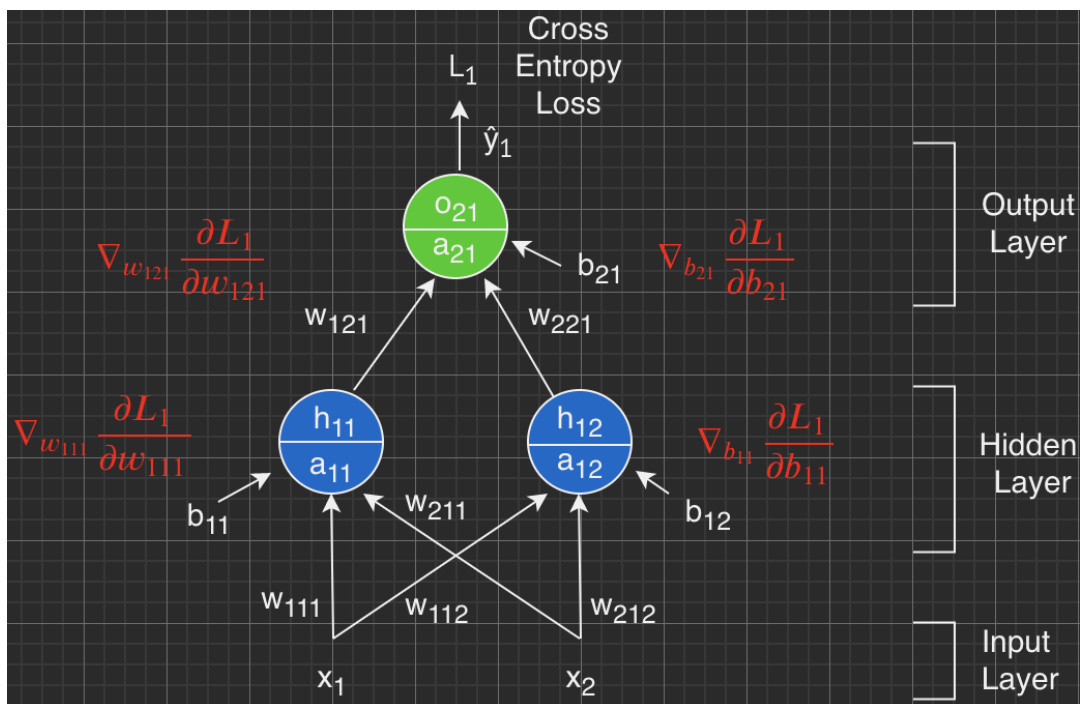


Ilustración 4- Representación de backpropagation. Fuente: Medium

El criterio con el que se actualizan los pesos y sesgos dependerá del algoritmo optimizador escogido. A continuación, se describirán los algoritmos optimizadores utilizados en este proyecto.

1.3.1.1 SGD – Descenso del gradiente

La idea es computar el error cometido por cada parámetro (peso o sesgo) y actualizar su valor en base a la siguiente fórmula:

$$w^+ = w - \eta \frac{\partial E}{\partial w}$$

Considerando w^+ como el peso actualizado, w como el peso anterior, η como el “learning rate” y $\frac{\partial E}{\partial w}$ como la derivada del error respecto al peso.

Estas actualizaciones de los pesos se realizarán de forma iterativa hasta que se alcancen uno de los siguientes escenarios:

- Se alcanza el error mínimo
- Se alcance el número máximo de iteraciones

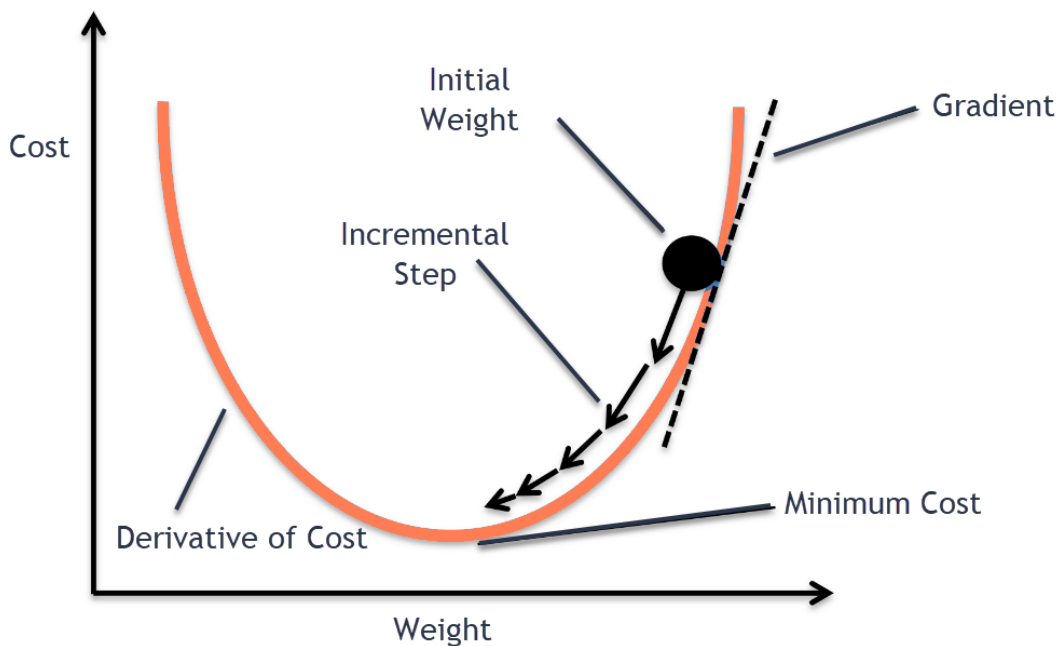


Ilustración 5 - Representación visual del funcionamiento de SGD

1.3.1.2 Adam

Es parecido al descenso de gradiente, como principal diferencia, el “learning rate” cambia a lo largo del proceso iterativo. Para más información sobre el método utilizado como criterio para variar el “learning rate”, véase el siguiente [enlace](#). Por otro lado, es más ligero que el descenso de gradiente a nivel computacional.

Capítulo 2. DESCRIPCIÓN DE LAS TECNOLOGÍAS

2.1 *DESCARGA DE DATOS DESDE E•SIOS*

E•SIOS se trata del sistema de información que proporciona la Red Eléctrica de España, S. A. U. (REE). A lo largo del proyecto, se accederá a la siguiente plataforma para obtener los datos necesarios para realizar los cálculos a partir de ellos. De esta plataforma se descargará información correspondiente a la energía transaccionada (tanto del mercado mayorista como de los contratos bilaterales), información sobre las distintas unidades de programación y sobre los precios de spot de cada mercado. Toda esta información se descargará para cada hora de cada mercado cada día. Los ficheros principales utilizados son los siguientes:

2.1.1 I90DIA

En este archivo se puede obtener información sobre la energía transaccionada cada día siendo el último día disponible es 90 días antes de la fecha actual. El archivo se descarga como un comprimido dentro del cual se pueden encontrar carpetas distintas para cada día, dentro de cada carpeta hay un fichero Excel tipo xls. En la siguiente figura queda reflejado el índice de la información presente en cada hoja del fichero.



P.O. 9 - Información Intercambiada por REE. Información a los tres meses (90 días)

Fecha del Informe
01-Nov-2020

Fecha de publicación
22-Mar-2021

Contenido del Libro

90DIA00	Fecha del Informe y Contenido del Libro
90DIA01	Resultado de la Programación Horaria del PDVP
90DIA02	Resultado de la Programación Horaria del P48
90DIA03	Resultado de la Resolución de Restricciones en el Mercado Diario
90DIA04	Reservada
90DIA05	Resultado de la Programación Horaria del Mercado de Secundaria. Valores horarios de la asignación de Banda de Regulación Secundaria
90DIA06	Resultado de la Programación Horaria del servicio de Energías de balance de tipo RR. Valores horarios de la energía asignada en el servicio de Energías de balance de tipo RR
90DIA07	Resultado de la Programación Horaria del Mercado de Terciaria. Valores horarios de la energía asignada en el Mercado de Terciaria
90DIA08	Resultado de la Resolución de Restricciones en Tiempo Real
90DIA09	Precios de la Resolución de Restricciones en el Mercado Diario
90DIA10	Precios de la Programación Horaria del servicio de Energías de Balance de tipo RR, Mercado de Terciaria y de la Resolución de Restricciones en Tiempo Real
90DIA11	Precios cuarto horarios de la programación del mercado de Energías de Balance de tipo RR
90DIA12	Energía Indisponible
90DIA13	Ofertas de Banda de Regulación Secundaria. Valores horarios de potencia ofertada y precio de la oferta.
90DIA14	Ofertas de energía en el servicio de Energías de Balance de tipo RR. Valores horarios de la energía ofertada y precio de la oferta.
90DIA15	Ofertas de energía en el Mercado de Terciaria. Valores horarios de la energía ofertada y precio de la oferta.
90DIA16	Indisponibilidades en Unidades Físicas
90DIA17	Ofertas Simples para solución de Restricciones técnicas del sistema. Valores horarios de la energía ofertada y precio de la oferta
90DIA18	Ofertas Complejas para solución de Restricciones técnicas del sistema.
90DIA19	Resultado de la Programación Horaria del PHF-1
90DIA20	Resultado de la Programación Horaria del PHF-2
90DIA21	Resultado de la Programación Horaria del PHF-3
90DIA22	Resultado de la Programación Horaria del PHF-4
90DIA23	Resultado de la Programación Horaria del PHF-5
90DIA24	Resultado de la Programación Horaria del PHF-6
90DIA25	Resultado de la Programación Horaria del PHF-7
90DIA26	Resultado de la Programación Horaria del PDBF
90DIA27	Energía gestionada mediante Contratos Bilaterales
90DIA28	Resultado de la Programación Horaria del Mercado de Secundaria. Valores horarios de la reasignación de Banda de Regulación Secundaria
90DIA29	Reservada
90DIA30	Reservada
90DIA31	Reservada
90DIA32	Ofertas Simples actualizadas para solución de Restricciones técnicas del sistema. Valores horarios de la energía ofertada y precio de la oferta
90DIA33	Ofertas Complejas actualizadas para solución de Restricciones técnicas del sistema.
90DIA34	Limitaciones Establecidas en Resolución de Restricciones en el Mercado Diario
90DIA35	Limitaciones Establecidas en Resolución de Restricciones en Tiempo Real
90DIA36	Resultado de la Programación Horaria del PHFC

Ilustración 6 - Índice hoja de cálculo I90DIA

De toda la información proporcionada se seleccionarán las hojas I90DIA01, I90DIA26, I90DIA19, I90DIA20, I90DIA21, I90DIA22, I90DIA23, I90DIA24, I90DIA25. La información correspondiente a cada hoja queda representada en la siguiente tabla:

I90DIA01	Resultado de la Programación Horaria del PDVP (Programa Diario Viable Provisional).
I90DIA26	Resultado de la Programación Horaria del PDBF (Programa Diario Base de Funcionamiento).
I90DIA19	Resultado de la Programación Horaria del PHF-1
I90DIA20	Resultado de la Programación Horaria del PHF-2
I90DIA21	Resultado de la Programación Horaria del PHF-3
I90DIA22	Resultado de la Programación Horaria del PHF-4
I90DIA23	Resultado de la Programación Horaria del PHF-5
I90DIA24	Resultado de la Programación Horaria del PHF-6
I90DIA25	Resultado de la Programación Horaria del PHF-7

Tabla 6 - Pestañas I90DIA Relevantes

Tal y como queda descrito en la página oficial de red eléctrica, los diferentes programas quedan definidos:

- Programa diario viable provisional (PDVP): Es el programa diario, con desglose horario, que incorpora las modificaciones introducidas en el PDBC para resolver las restricciones técnicas.
- Programa Diario Base de Funcionamiento (PDBF): Programa de intercambio internacional establecido el día previo a la entrega de energía mediante contratación en los Mercados Diarios y/o contratación bilateral con entrega física.
- Programa horario final (PHF): Es la programación establecida por el OS con posterioridad a cada una de las sucesivas subastas del mercado intradiario de unidades de programación correspondientes a ventas y adquisiciones de energía en el sistema eléctrico peninsular español, como resultado de la agregación de todas las transacciones firmes formalizadas para cada periodo de programación como consecuencia del programa diario viable y de la casación de ofertas en el mercado intradiario.

Estas hojas contienen la información repartida de la siguiente manera:

		Indicadores	Programa (MWh)		
		Día	01-Nov-2020		
		Programa	PDVP		
Unidad de Programación	Tipo Oferta	Hora	Total	00-01	01-02
A2GAME	10		9.9	0.4	0.4
ABA1	1		9.1	0.3	0.3
ABA2	1		9.1	0.3	0.3
ABO1	1		1,973.0	187.0	186.0

Tabla 7 - Extracto de tabla I90DIA01 del archivo I90DIA_20201101

Como se puede observar, obtenemos el total diario y el desagregado por horas de la cantidad de energía en MWh con la que cada Unidad de Programación entró en el mercado. En este caso por ejemplo se podría decir que A2GAME casó 0.4 MWh entre las 00:00 y la 01:00 el día 1 de noviembre de 2020 en el PDVP. Todas las hojas de I90DIA escogidas son de idéntico formato, excepto la hoja I90DIA26, en la que además también aparecen las transacciones bilaterales. Véase en la siguiente tabla:

APRIC01	8	Mercado	-90.4	-3.7
AQUEC01	7	Mercado	-79.0	-3.5
AQUIC01	8	Bilateral	-22.7	-0.9
AQUOC01	7	Mercado	-45.7	-1.9

Tabla 8 - Extracto de I90DIA26, de la hoja de cálculo I90DIA_20201101

2.1.2 CSV HISTÓRICO PRECIO SPOT

Para obtener los valores del precio spot de forma desagregada en horas para cada mercado se accede a una herramienta de E•SIOS en la que se posibilita la adición de estos parámetros:

PRECIO MERCADO SPOT DIARIO ESPAÑA

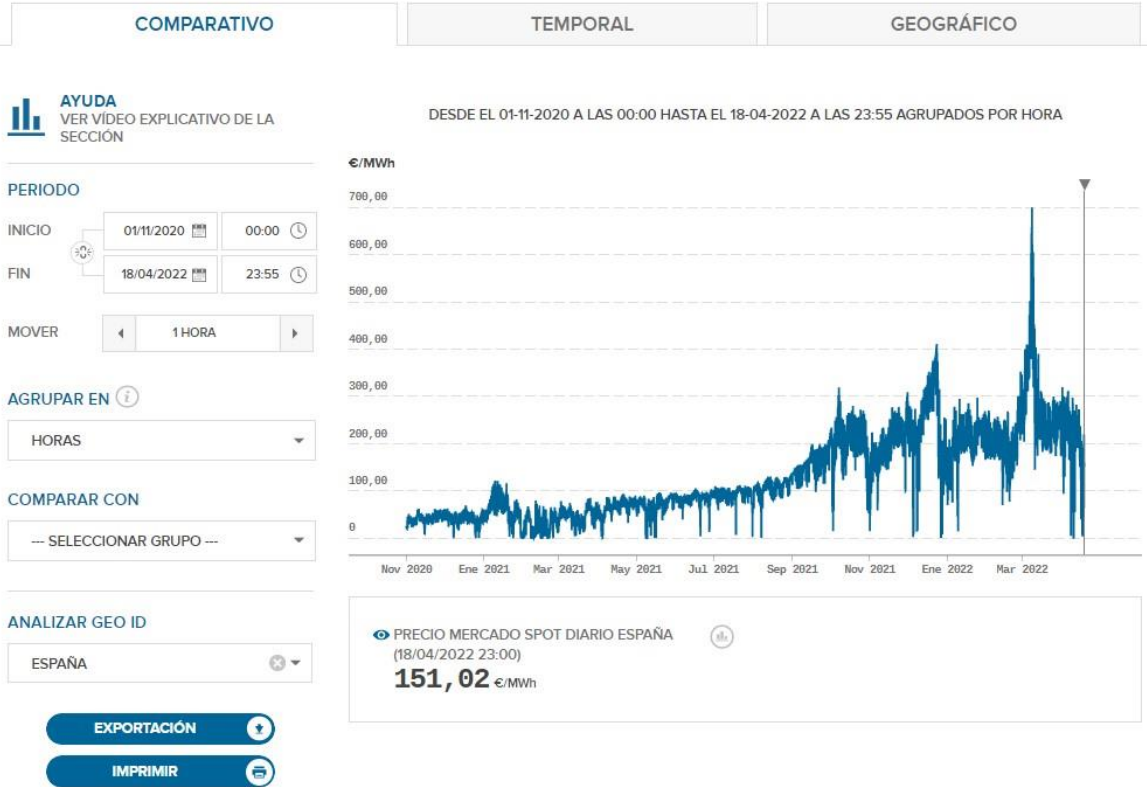


Ilustración 7 - Punto de descarga de archivos precio spot en la página web oficial de E•SIOS

Estos archivos se pueden descargar en formato CSV facilitando así su posterior lectura. Debe descargarse un archivo distinto para cada mercado en el horizonte temporal deseado.

2.1.3 CSV INFORMACIÓN ADICIONAL UP

De cara a poder estudiar las Unidades de Programación más en profundidad y entender mejor su comportamiento, E•SIOS facilita también la descarga de un archivo CSV el cual contiene las siguientes columnas de información para cada UP:

Código de UP
Descripción corta
Descripción larga
Potencia máxima MW
Código EIC
Tipo de producción

Negocio
Zona de Regulación
Sujeto del Mercado
Tipo de UP

Tabla 9 - – Listado de campos disponibles en CSV información adicional UP

2.2 GESTIÓN DE LOS DATOS

El lenguaje de programación a utilizar para el manejo de los datos es Python, tanto de los archivos de Excel, como de los CSV. Una vez se han realizado las lecturas y los cálculos oportunos, generalmente se exportan los resultados a otro archivo CSV. Para analizar los resultados obtenidos, se hará uso de Excel, obteniendo como input el “CSV resultado” y utilizando las herramientas de las tablas dinámicas. Este método proporciona una mayor velocidad de cálculo en los datos, por un lado, y un manejo fácil para el análisis de los resultados por otro.

2.2.1 PYTHON

Las librerías principales empleadas han sido:

- Datetime: para facilitar el manejo de la navegación por las distintas fechas y horas.
- Pandas: se ha utilizado la clase Pandas Dataframe (equivalente de tablas de información) para el manejo de los datos, lectura y exportación de csv.
- Pyexcel: para la lectura y transformación de formato de los ficheros xls procedentes de la carpeta de los I90DIA, explicado en el apartado 2.1.1.
- Sklearn: para facilitar la implementación de los modelos basados en redes neuronales, así como las métricas usadas comúnmente en estos modelos.
- Matplotlib: construcción de gráficas de los resultados.

El entorno de desarrollo integrado (IDE, Integrated Development Environment) utilizada ha sido Pycharm Community Edition, principalmente por la facilidad de visualización de los dataframes (tablas de información) debuggeando:



	Dato	00-01	01-02	02-03	03-04	04-05	05-06	06-07	07-08	08-09	09-10	10-11	11-12	12-13	
0	per_pe_intra1	0.00861	0.00994	0.00928	0.00861	0.00994	0.00928	0.00994	0.00994	0.00795	0.01060	0.00861	0.00928	0.00928	0.
1	per_no_pe_intra1	0.00861	0.00795	0.00795	0.00729	0.00861	0.00928	0.00994	0.00994	0.00994	0.00928	0.00861	0.00861	0.00795	0.
2	per_coinc_intra1	0.02386	0.02187	0.02187	0.02319	0.02386	0.02054	0.02319	0.02121	0.02187	0.02452	0.02253	0.02187	0.02518	0.
3	encarece_intra1	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	0.00000	1.00000	1.00000	1.00000	1.
4	per_pe_intra2	0.00722	0.00919	0.01051	0.01116	0.01051	0.00919	0.01116	0.00919	0.01182	0.00854	0.00919	0.01051	0.00985	0.
5	per_no_pe_intra2	0.00657	0.00788	0.00788	0.00722	0.00722	0.00722	0.00788	0.00657	0.00854	0.00657	0.00788	0.00525	0.00722	0.
6	per_coinc_intra2	0.02495	0.02955	0.02626	0.02364	0.02364	0.02364	0.02298	0.01970	0.02364	0.02298	0.02495	0.02101	0.02035	0.
7	encarece_intra2	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	0.00000	1.00000	0.00000	0.00000	0.00000	0.00000	0.00000	0.
8	per_pe_intra3	0.01242	0.01307	0.01111	0.01111	0.00980	0.01111	0.01176	0.01046	0.01307	0.01242	0.01176	0.01176	0.01176	0.
9	per_no_pe_intra3	0.00980	0.00915	0.00980	0.00915	0.00784	0.00719	0.00784	0.00850	0.00980	0.00915	0.00915	0.00915	0.00850	0.
10	per_coinc_intra3	0.03007	0.03137	0.02941	0.02941	0.02614	0.02745	0.02549	0.02680	0.02810	0.02418	0.02157	0.02157	0.02288	0.
11	encarece_intra3	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.

Tabla 10 - Ejemplo de visualización de dataframe en Pycharm Community Edition

2.2.2 TABLAS DINÁMICAS EXCEL

Principalmente se utilizan los filtros en las columnas, de esta forma se pueden analizar y ordenar los resultados de una forma cómoda, visual y rápida. Además, se pueden hacer cálculos rápidos a partir de los datos en la tabla y obtener un resultado inmediato.

Capítulo 3. ESTADO DE LA CUESTIÓN

Existen numerosos modelos para intentar predecir tanto el precio spot del mercado diario, como los posibles cambios en el mismo, a continuación, se muestran algunos de ejemplo. Sin embargo, no existe ningún modelo capaz de predecir con exactitud los cambios que se producen en los mercados intradiarios. Se ha realizado un extenso estudio de estos modelos destacando los siguientes.

3.1 *MODELOS EXISTENTES*

3.1.1 **MODELO DE PREVISIÓN DEL PRECIO DEL MERCADO DIARIO DE LA ELECTRICIDAD EN EL CORTO PLAZO**

Esta persona desarrolla dos posibles modelos:

- Modelo 1: se trata de hacer una previsión de las ofertas y la demanda en distintos periodos para intentar llegar al precio de casación. En las conclusiones se menciona la baja efectividad de este modelo debido a la falta de datos:

“Se ha podido apreciar en la comparativa de ambos modelos que el margen de fiabilidad del segundo es mucho mejor que el primero. Esto puede deberse tanto a la falta de datos en su elaboración como a su poco rango de obtención de datos.”

- Modelo 2: en este modelo se tienen en cuenta varios inputs interesantes: precio del Brent, índices euro/dólar, índices de gas natural (Henry Hub), y Emisiones de CO₂. El autor realiza unos de cómo afectan estos inputs al precio del mercado eléctrico, pero concluye confirmando que, pese a que estos inputs pueden explicar el comportamiento del precio, los datos extraídos no han sido suficientes para realizar un modelo:

“Una vez finalizado el estudio se realizó el análisis de los resultados y se concluyó que los datos extraídos no eran suficientes para poder realizar un modelo cuyos resultados fueran fiables. Aunque se hayan obtenido los polinomios que aproximan el comportamiento tanto del precio como de la energía en función de las distintas variables, se concluye que no es suficiente y que, por tanto, habrá que encontrar otra forma de prever la curva de la oferta.”

3.1.2 DESARROLLO DE UN MODELO DE PREDICCIÓN DEL PRECIO DE LA ENERGÍA ELÉCTRICA PARA EL MERCADO A PLAZO MEDIANTE REDES NEURONALES.

El autor desarrolla tres redes neuronales para predecir el precio de la energía eléctrica en el medio y largo plazo. Los factores que se tienen en cuenta para estas redes son los siguientes:

- El día de la semana.
- Días festivos.
- Reservas hidroeléctricas un año antes en hm³
- Demanda eléctrica real un año antes en MW.
- Generación eléctrica de los siguientes tipos en MW de un año antes: carbón, ciclo combinado, nuclear, solar, eólica e hidráulica. Se han tomado en consideración por sus características, bien el peso en la mezcla de generación o por favorecer su uso costes mayores o menores en el mercado diario de la energía.
- Precios de los futuros del petróleo Brent y el gas natural en los mercados internacionales.

Analizando los resultados parece que el modelo desarrollado funciona:

“En el futuro a un mes, el error medio ha sido superior en todos los casos al 10%. En el resto de caso siempre existe un modelo que obtenga un error inferior al 5% que se toma como referencia, simplemente

considerando las previsiones para un número de días desde el inicio determinado.”

Los modelos planteados aún así son bastante complejos y el autor concluye indicando su intención por utilizar el output de su modelo para otro.

3.1.3 ECONOMETRIC MODELLING AND FORECASTING OF INTRADAY ELECTRICITY PRICES

Se trata de un modelo de predicción para los mercados intradiarios en Alemania. El autor prueba modelos basados en series temporales, como por ejemplo predicciones por Naive. En las conclusiones finaliza remarcando que el rendimiento del modelo no es lo suficientemente bueno., además, menciona la siguiente frase:

“It is certain that due to the growth of the intraday market and the scarcity of the literature on this subject, still a lot of research needs to be conducted.”

Donde menciona que existe poca investigación del comportamiento de los mercados intradiarios a día de hoy.

3.1.4 SHORT-TERM PRICE FORECASTING MODELS BASED ON ARTIFICIAL NEURAL NETWORKS FOR INTRADAY SESSIONS IN THE IBERIAN ELECTRICITY MARKET

Se trata de un modelo basado en redes neuronales a partir de:

- Serie temporal histórica por horas
- Potencia generada y demandada el día anterior
- Predicciones de demanda
- Meteorología y generación eólica del día siguiente
- Otras variables históricas

Los modelos desarrollados tienen un valor del error medio del porcentaje absoluto (MAPE) del 7.5 % al 13 %, lo que indica que tienen un rendimiento bastante elevado.

3.2 CONCLUSIONES DEL ESTADO DE LA CUESTIÓN

Se han observado cantidad de modelos con la capacidad de realizar predicciones con precisión del precio spot del mercado diario utilizando predicciones de demanda, datos meteorológicos y/o series temporales. Sin embargo, no se han encontrado variedad de modelos para predecir la evolución del precio de los mercados intradiarios en España.

Capítulo 4. DEFINICIÓN DEL TRABAJO

4.1 JUSTIFICACIÓN

La posibilidad de predicción del comportamiento del mercado eléctrico, como de cualquier otro mercado, supone tener una capacidad de anticipación gracias a la cual se pueden obtener beneficios regularmente. En el apartado anterior [Capítulo 3.] queda reflejado como existen actualmente numerosos modelos para predecir el precio spot del mercado diario con buena precisión. Sin embargo, no es el caso con los mercados intradiarios, ya que éstos dependen más del comportamiento de los agentes, que de condiciones ajenas a ellos como puede ser la previsión del tiempo, la predicción de la demanda, etc. Por ello, en este proyecto se intentará predecir el desenlace de los intradiarios basándose en el estudio de las estrategias e impacto que tienen los agentes de mercado.

4.2 OBJETIVOS

- Obtener histórico de los movimientos que han realizado las unidades de programación a partir de los programas que se pueden descargar de la página oficial de red eléctrica de cada mercado. Estos movimientos se registrarán para cada unidad de programación, cada hora de cada día en cada mercado.
- Obtener histórico de los cambios en los precios que han ocurrido a cada hora cada día en cada mercado.
- Hacer un estudio del volumen de cambios que han realizado de forma histórica cada unidad de programación para centrar el estudio en este número reducido y simplificar el modelo.
- Con los datos obtenidos, llevar a cabo un análisis de las unidades de programación entendiendo cómo afecta al precio los movimientos que realizan.
- Alineación con los ODS, garantizar el acceso a una energía asequible, segura, sostenible y moderna. El entendimiento de la casación del precio del mercado eléctrico diario ayudará al desarrollo de la energía asequible y no contaminante, pues

al poder realizar mejores predicciones sobre la futura demanda y generación eléctrica, se podrá evitar recurrir a agentes que ofrezcan energía eléctrica de fuentes más convenientes, pero menos limpias.

- Objetivo principal: predicción del desenlace de los mercados intradiarios a corto plazo basándose en las acciones de los distintos agentes. Esto significa analizar el comportamiento de las unidades de y sus consecuencias en el mercado cuando actúan. De esta forma conociendo las acciones de éstas, poder realizar una previsión del comportamiento del mercado.

4.3 METODOLOGÍA

La metodología de trabajo seguirá los puntos concretados en la parte de objetivos. Sin embargo, es complicado organizar un cronograma, pues a partir de cierto punto el procedimiento no es claro ya que se irán tomando decisiones sobre por donde continuar en función de los resultados que se vayan obteniendo. Por ello, se toma la decisión de seguir una estructura para los pasos iniciales, y luego continuar el estudio sin seguir un cronograma definido.

En orden cronológico, los pasos iniciales serán:

1. Obtención y registro de los datos sobre los movimientos de cada unidad de programación, para cada mercado, cada hora de cada día. Guardado de estos datos en formato cómodo y rápido de utilizar para posterior utilización.
2. Obtención y registro sobre las variaciones en los precios de cada día, cada mercado para cada hora. Guardado de estos datos en formato cómodo y rápido de utilizar para posterior utilización.
3. Estudio del volumen de movimientos realizados por cada unidad de programación en los mercados deseados en un horizonte temporal concretado para poder escoger un grupo más reducido que tenga mayor actividad.
4. Estudiar la consecuencia de estos movimientos en el mercado para poder identificar las unidades de programación más “peligrosas” o influyentes.

A partir de aquí, la dirección del proyecto es abierta por lo que no se puede concretar un horizonte fijo.

Capítulo 5. INFRAESTRUCTURA DE LOS DATOS

Se comprende la necesidad de crear una infraestructura de datos sobre la que los modelos puedan obtener la información en menor tiempo y con la mayor facilidad posible. Como se ha podido observar en el Capítulo 2. los archivos descargados directamente no tienen ni el formato deseado ni el dato final tras el cálculo apropiado.

5.1 ADAPTACIÓN DE FORMATOS DE LOS ARCHIVOS

5.1.1 I90DIA

En el Capítulo 2. se menciona que los archivos I90DIA son de formato xls, lo que dificulta mucho la navegación dentro del archivo en Python. Por ello se ha construido un programa con la capacidad de abrir el archivo xls, cambiar su formato, y guardar una copia en la misma ubicación de tipo.xlsx.

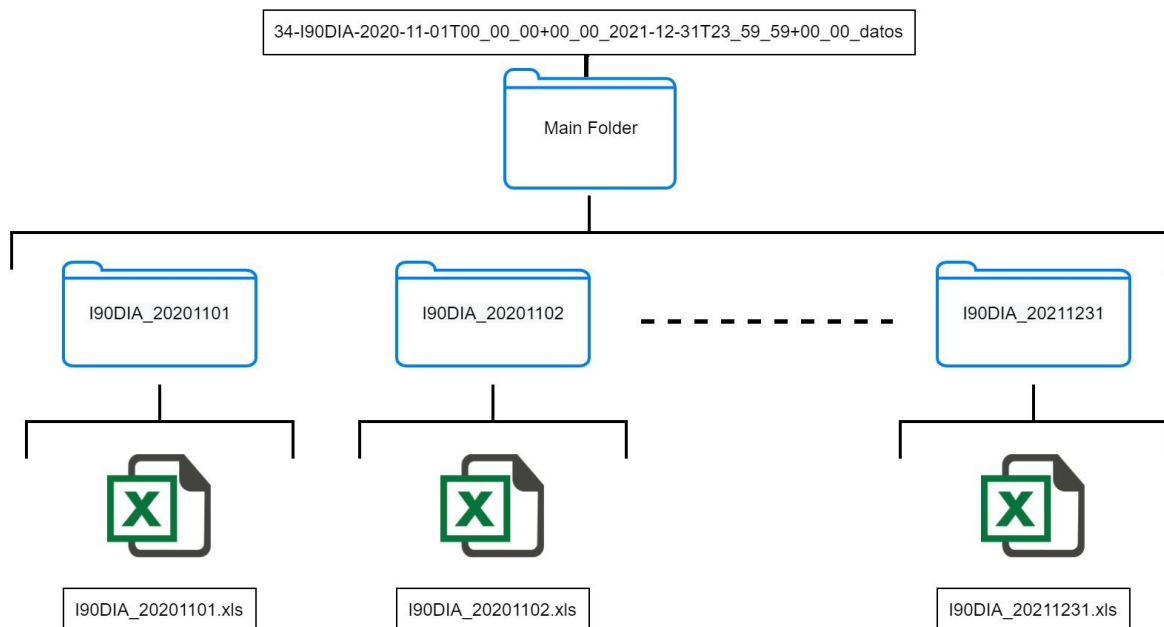


Ilustración 8 - Esquema arquitectura carpeta principal I90DIA

Para poder entender la dinámica del programa que cambia todos los archivos en un horizonte temporal en la misma ejecución, hay que entender antes la arquitectura con la que estos archivos son descargados de E•SIOS.

Como se puede observar en la anterior ilustración, la estructura es:

- Carpeta principal: contiene en el nombre fecha inicial y final.
- Subcarpetas diarias
- Archivo xls dentro de cada subcarpeta

La idea del programa es recorrer un bucle de fechas en el que, en cada vuelta, se cambia el nombre del directorio para acertar con la subcarpeta y archivo xls, abrir el archivo, extraer los datos, y guardarlos en el mismo directorio en un archivo.xlsx. La librería utilizada para realizar este proceso es “pyexcel”. Esta función tarda aproximadamente unos 19 segundos en transformar cada archivo.

5.1.2 CSV INFORMACIÓN ADICIONAL UP

Como se ha explicado en el Capítulo 2. se realiza la descarga de un archivo que contiene información adicional sobre las unidades de programación. El CSV descargado tiene los formatos de número con puntos en los separadores de millares, y comas en los decimales, lo que da error en su lectura con Python. Por ello se crea una función destinada a leer el CSV, reemplazar estos caracteres leyendo estos datos como una cadena de caracteres, transformarlo a formato float (número con decimales), y volver a guardar el CSV.

5.2 ESQUEMA INFRAESTRUCTURA DE DATOS

Lo más cómodo y rápido para realizar lecturas de datos son los archivos CSV, ya que la librería utilizada para gestionar el manejo de datos, pandas, tiene implementado un “method” que facilita esta lectura. Por ello, se deciden almacenar todos los datos que se necesitarán para los modelos en formato CSV. El esquema que seguirá la infraestructura queda reflejado en la siguiente ilustración:

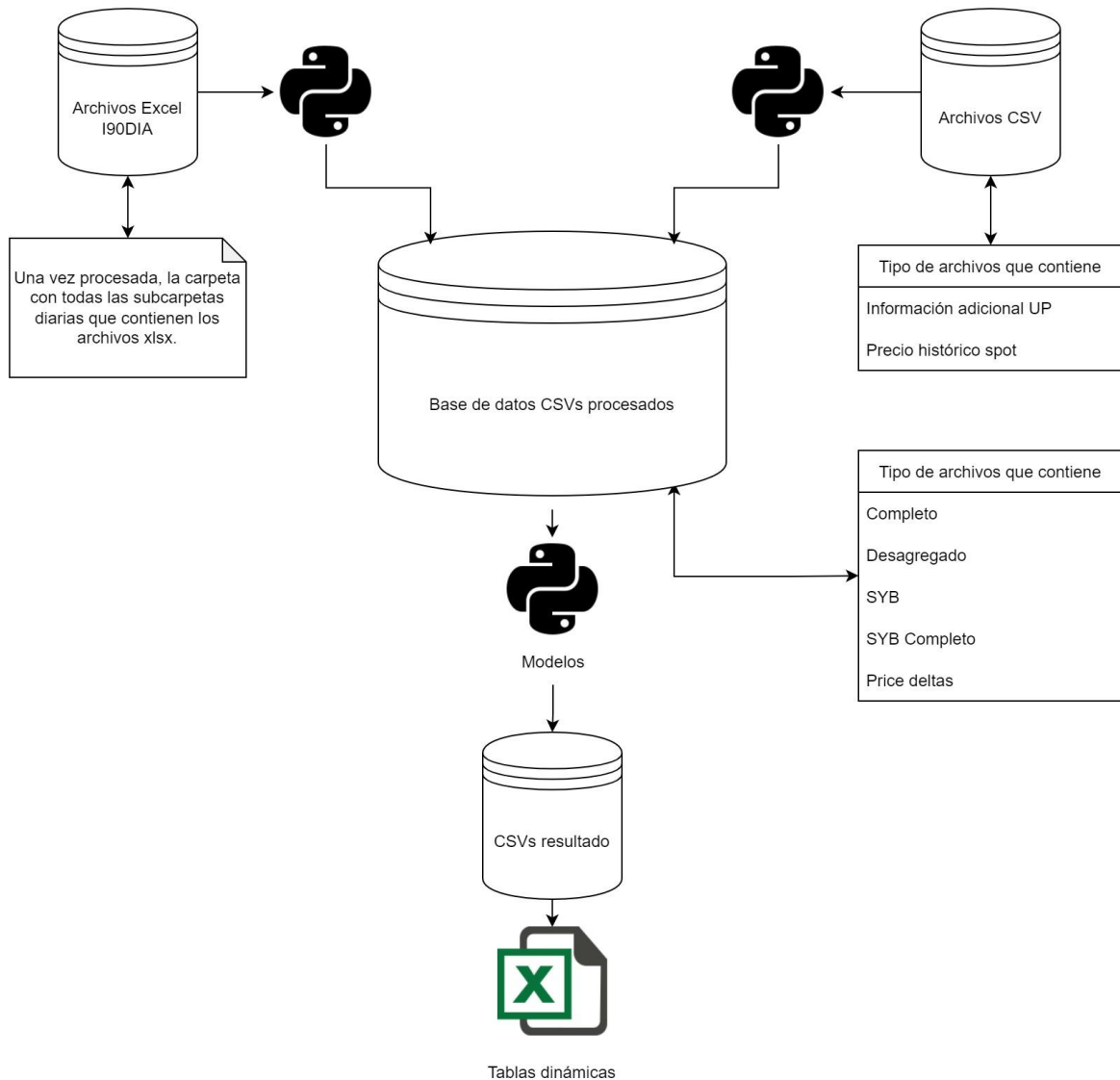


Ilustración 9 - Esquema Infraestructura de datos

Los tipos de archivo que se guardan en la base de datos quedan explicados a continuación.

5.2.1 ARCHIVOS CSV PROCEDENTES DE I90DIA

Cabe destacar que estos archivos tienen columnas repetidas. La finalidad de esto es, en el caso de necesitar información de dos tipos de CSV distintos en el entrenamiento de un modelo, no tener que abrir dos, ya que aumentaría sustancialmente el tiempo de procesamiento para el modelo.

5.2.1.1 Desagregado

El propósito de los modelos a crear es predecir los diferenciales de precios que se van a producir en restricciones y los mercados intradiarios. Por ello, lo que interesa no es el programa resultante de cada mercado (lo que contiene el I90DIA), sino la diferencia entre dos programas ya que esto indica el volumen de energía movido por cada agente cada hora y en cada mercado. A esta energía movida en cada mercado se la ha denominado delta.

Por ejemplo, si la unidad de programación ACE3 el día 01/11/2020 de 00:00 a 01:00 tiene un valor de 121.5 MW en el PHF1, y 160 MW en el PHF2, supondría que en el archivo CSV desagregado correspondiente al mercado intradiario 2 en la fecha 01/11/2020 de 00:00 a 01:00 la unidad de programación ACE3 tendrá registrado un delta de +38.5 MW. Esto significa que la UP ACE3 en el mercado MI2 vendió 38,5 MW.

Con esto se consigue averiguar el volumen movido por cada agente en cada hora y en cada mercado.

Este cálculo se realiza para todas las horas, todos los mercados, todos los días.

Unidad de Programación	00-01	01-02	02-03
A2GAME	0.0	0.0	0.0
ABA1	0.0	0.0	0.0
ABA2	0.0	0.0	0.0
ABO1	0.0	0.0	10.0
ABOUC01	0.0	0.0	0.0
ACCEC01	0.0	0.0	0.0
ACE3	38.5	38.5	38.5
ACENC01	0.0	0.0	0.0
ACSOC01	0.0	0.0	0.0

Tabla 11 – Ejemplo de CSV desagregado

5.2.1.1.1 Corrección Bilaterales

Además del Mercado Ibérico de Energía, los agentes podrán llevar a cabo contratos bilaterales, es decir, acuerdos en los cuales productor y consumidor pactan de forma independiente una potencia a un precio. Estos acuerdos quedan reflejados en el mercado de

restricciones (PDBF, dentro del I90DIA, la pestaña I90DIA26). Debido a que estos acuerdos distorsionan el cálculo de los deltas, la forma de corregir el problema para cada unidad de programación es:

- En el caso de que solo tenga un tipo de movimiento en el mercado: no hace falta hacer nada al respecto.
- En el caso de que tenga programa en el mercado diario de tipo mercado y bilateral también, sumar ambos movimientos y computar el neto.
- En el caso de que solo haya transacción bilateral, no hacemos nada al respecto.

Este parche se realizó ya que se estaban contabilizando los movimientos de mercado y bilaterales de las unidades de programación por separado, por lo que el volumen de los deltas que realizaban las unidades se veía claramente aumentado en algunas.

Anteriormente a llegar a esta solución, lo que se hacía era simplemente omitir los acuerdos bilaterales, sin embargo, esto no solucionaba realmente el problema, ya que la cantidad pujada por aquellas unidades con transacciones de mercado y bilateral a la vez era falsa, por lo que los deltas respecto al intradiario 1 eran falsos.

5.2.1.1.2 Fallos “tipo 1” y “tipo 2”

El cálculo del delta de los programas entre las fotos finales de dos mercados consecutivos se realiza con una función completamente ajena a los programas finales de mercados anteriores, por lo que cabe la posibilidad de que se generen los fallos que se muestran en la siguiente ilustración simplificada.

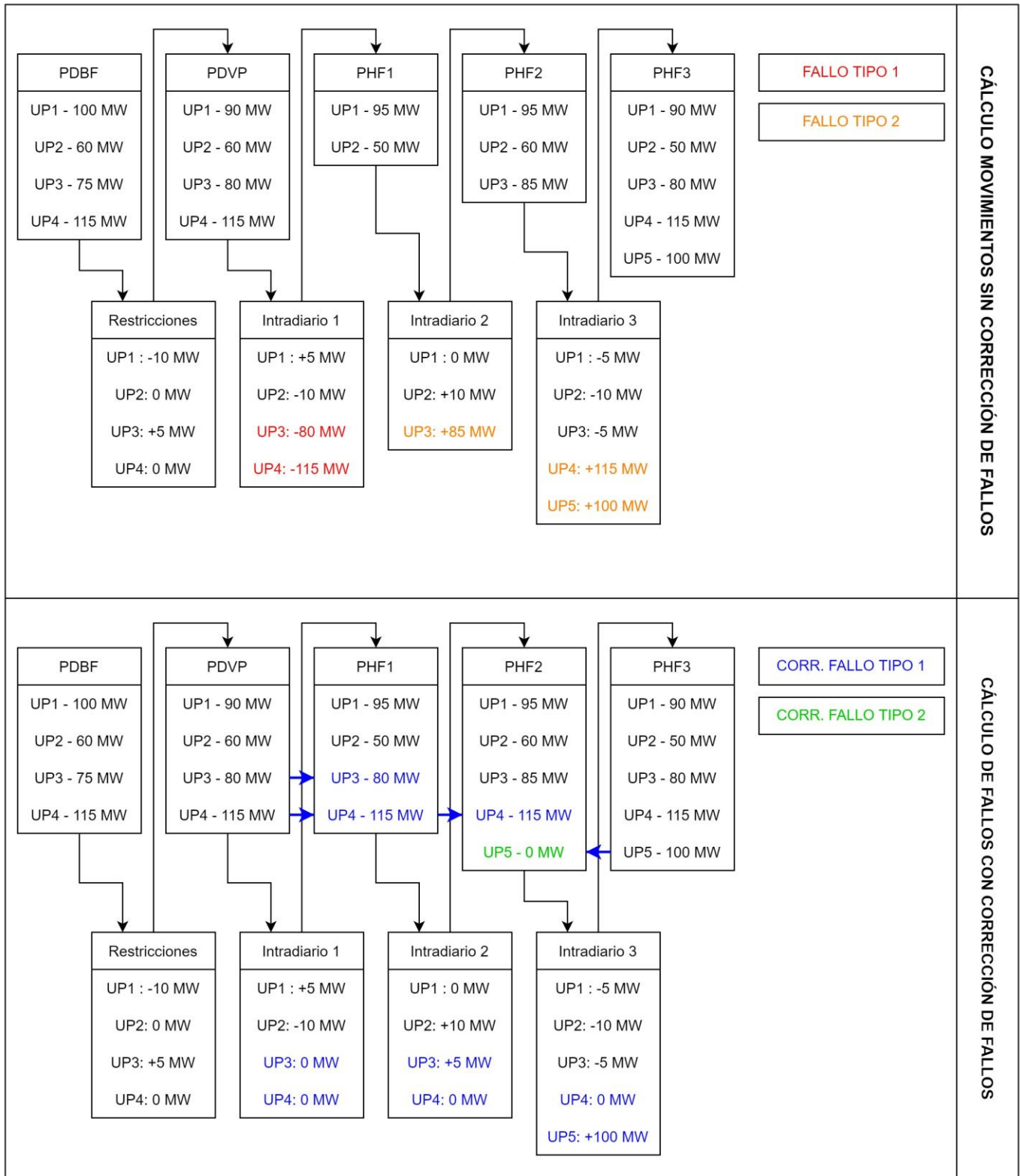


Ilustración 10 - Tabla explicativa tipos de fallos

- Fallo tipo 1: como se puede observar en la ilustración, ocurre cuando una unidad de programación aparece en el programa PDBF, pero no en la final, si no se corrigiese este fallo, computaría como que su puja final es 0. Realmente lo que ocurre es que esta unidad no ha entrado en el siguiente mercado, por lo que mantiene su puja anterior a esa hora, y se debe computar como un movimiento de 0 MW. En la anterior ilustración puede verse este error en UP3 y UP4 en el mercado intradiario 1.
- Fallo tipo 2: ocurre cuando la unidad aparece en el programa final pero no en el inicial. En este caso, aunque el valor del movimiento no se ve afectado, es importante identificarlos ya que, en la mayor parte de los casos, cuando se corrige el fallo tipo 1, prácticamente se deberían eliminar los fallos tipo 2. Nótese en la anterior ilustración que, de tener tres fallos tipo 2, al corregir fallos tipo 1, pasa a haber solo un fallo tipo 2.

Desde el punto de vista del programa, el método implementado se basa en añadir al retorno de la un dataframe que contiene la foto final con las unidades de fallo tipo 1 añadidas. De esta forma, en la próxima llamada a la función, la foto inicial será este dataframe recibido en la anterior iteración. En definitiva, se está otorgando “memoria” a la función para solucionar estos fallos.

La corrección de estos fallos es crucial para el cálculo correcto de los deltas en las pujas de las unidades de programación.

5.2.1.2 SYB Completo

En el archivo CSV desagregado no quedan diferenciadas las subidas de las bajadas, es decir, para poder calcular por separado la cantidad “subida” y “bajada”, de deben separar en columnas distintas. Para crear estos archivos, simplemente se recorren todos los desagregados por horas, si el delta es positivo, se considera subida, se guarda en la columna total subidas en la hora correspondiente, si no, en total bajadas en signo positivo.

Por ejemplo, si la unidad de programación ACE3 el día 01/11/2020 de 00:00 a 01:00 en el archivo CSV correspondiente al mercado intradiario 2 tiene registrado un delta de +38.5 MW. Entonces en la columna Total Subidas 00-01 tendrá +38.5 MW, mientras que en Total Bajadas 00-01 tendrá 0 MW.

Unidad de Programación	Total subidas 00-01	Total bajadas 00-01
A2GAME	0.0	-0.0
ABA1	0.0	-0.0
ABA2	0.0	-0.0
ABO1	0.0	-0.0
ABOUC01	0.0	-0.0
ACCEC01	0.0	-0.0
ACE3	38.5	0.0

Tabla 12 - Ejemplo de CSV SYB Completo

5.2.1.3 SYB

Se calcula la suma de todas las subidas en cada hora y bajadas por separado, de forma que no se tienen en cuenta las horas, si no la suma de lo ha ocurrido en las 24 horas juntas de cada día, para esa unidad de programación, en cada mercado.

Por ejemplo, si los movimientos de la unidad de programación ACE3 el día 01/11/2020 en el mercado intradiario 2 tiene los siguientes deltas:

Unidad de Programación	00-01	01-02	02-03	03-04	04-05	05-06	06-07	07-08	08-09	09-10	10-11	11-12	12-13	13-14	14-15	15-16	16-17	17-18	18-19	19-20	20-21	21-22	22-23	23-24
A2GAME	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ABA1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ABA2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ABO1	0.0	0.0	10.0	80.0	80.0	80.0	80.0	80.0	80.0	80.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	80.0	80.0	80.0	80.0	0.0	0.0	20.0
ABOUC01	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ACCEC01	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ACE3	38.5	38.5	38.5	38.5	38.5	38.5	38.5	38.5	38.5	0.0	0.0	14.3	0.0	0.0	0.0	0.0	0.0	38.5	0.0	0.0	0.0	0.0	0.0	0.0

Tabla 13 - Ejemplo de CSV Desagregado

Entonces se calculará de la siguiente manera:

$$Total\ Subidas = 38.5 + 38.5 + 38.5 + 38.5 + 38.5 + 38.5 + 38.5 + 38.5 + 38.5 + 38.5 + 14.3 + 38.5 = 399.3$$

$$Total\ Bajadas = 0$$

Se puede observar en la siguiente ilustración el archivo CSV resultante:

Unidad de Programación	Total subidas	Total bajadas
A2GAME	0.0	0.0
ABA1	0.0	0.0
ABA2	0.0	0.0
ABO1	990.0	0.0
ABOUC01	0.0	0.0
ACCEC01	0.0	0.0
ACE3	399.3	0.0

Tabla 14 - Ejemplo de CSV SYB

De paso que se realiza este cálculo y que se realiza realmente a partir del archivo SYB completo, se aprovecha y se añaden estas dos columnas al final del último CSV.

5.2.1.4 Completo

Se trata básicamente del CSV desagregado con las dos columnas de SYB añadidas.

5.2.2 ARCHIVO CSV INFORMACIÓN ADICIONAL UP

Este archivo ya se descarga en formato CSV de E•SIOS, como se explica en el 2.1. El motivo de tener que hacer un procesamiento previo a la subida a la base de datos de archivos CSV procesados se describe en el 5.1.

5.2.3 ARCHIVO CSV PRECIO SPOT HISTÓRICO

Como se indica en el 2.1.2, desde la página de E•SIOS, se descargan distintos archivos CSV para cada mercado el histórico, por horas (viene dado en la primera columna en formato datetime). De forma similar al cálculo de los archivos CSV procedentes del I90DIA, descrito en el apartado 5.2.1, lo que interesa es el cambio en el precio por efecto de un mercado determinado, cada día, en cada hora, respecto al Por ello, se debe realizar la resta para obtener el delta del cambio en el precio.

Además, se tiene la posibilidad de estudiar el cambio relativo respecto al precio establecido en el mercado diario, de esta forma, un mismo delta destaca más cuando el precio establecido en el diario es menor, es decir, es un cambio más notable, y queda reflejado en el archivo.

Datetime	Precio Spot diario	Precio Spot tras intradiario 1	Precio Spot tras intradiario 2
2020-11-01 00:00:00	26	28	29.5
2020-11-01 01:00:00	23.9	24.4	25.9
2020-11-01 02:00:00	20.03	20.24	20.79
2020-11-01 03:00:00	19.7	19.79	20.34

Datetime	Delta precio intradiario 2
2020-11-01 00:00:00	$28 - 29.5 = -1.5$
2020-11-01 01:00:00	$24.4 - 25.9 = -1.5$
2020-11-01 02:00:00	$20.24 - 20.79 = -0.55$
2020-11-01 03:00:00	$19.79 - 20.34 = -0.55$

Todos los datos en € / MWh

Datetime	Cambio relativo precio intradiario 2
2020-11-01 00:00:00	$-1.5 / 26 = -0.058$
2020-11-01 01:00:00	$-1.5 / 23.9 = -0.063$
2020-11-01 02:00:00	$-0.55 / 20.03 = -0.028$
2020-11-01 03:00:00	$-0.55 / 19.7 = -0.028$

Ilustración 11- Explicación cálculo archivo CSV precios

Por último, se combinan todos los mercados en el mismo archivo CSV, separado por columnas.

Capítulo 6. ANÁLISIS PREVIOS

6.1 VOLUMEN DE MOVIMIENTOS POR UP

La finalidad es poder estudiar la cantidad de MW que cada unidad de programación ha “cambiado”, es decir, tiene un valor distinto de cero en los archivos CSV calculados del I90DIA, explicado en el 5.2.1. La suma se realiza de forma desagregada para subidas y bajadas, y también la adición de ambas. La función genera este informe en un horizonte temporal determinado y en los mercados que se indique. Además, también se puede escoger si se desea hacer la suma por horas, o por días, al tener ambos archivos CSV disponibles en la base de datos. Por último, se calcula la participación de cada UP en función de el número de días que aparezca en algún punto de los mercados escogidos entre el número total de días.

Antes de exportar el dataframe, se añaden varias columnas con datos procedentes del archivo CSV Información Adicional UP, para tener todo combinado en un mismo archivo CSV resultado, el cual será exportado a la carpeta de los archivos CSV resultado.

Este programa queda como una función por si en futuros modelos se quiere realizar un estudio para seleccionar aquellas unidades que tengan un volumen relevante en un horizonte temporal seleccionado, y descartar el resto, reduciendo el tiempo de procesamiento en el modelo.

Los parámetros escogidos para la ejecución del programa quedan recogidos en el nombre del archivo. Por ejemplo:

“sum_syb_intra1_intra2_intra3_from_20201101_to_20211118.csv”.

UP	Total subidas	Total bajadas	Participación	Días participados	Potencia máxima MW
A12PAME	0	0	0.01305483	5	0.5
A2GAME	405.9	133.6	0.770234987	295	3.8
ABA1	77.9	1383.4	1	383	0.3

	ICAI	ICADE	CIHS		
ABA2	77.9	1385.7	1	383	0.4
ABO1	17201.4	9113.2	0.070496084	27	341.7
ABO2	390970	14248.2	0.960835509	368	561.8
ABOUC01	0	539.8	1	383	0.1

Tipo de producción				
	Negocio	Zona de Regulación	Sujeto del Mercado	Tipo de UP
Derivados del petróleo ó carbón	Venta	SZR	AME	Generación
Energía residual	Venta	SZR	AME	Generación
Hidráulica UGH	Venta	SZR	BARER	Generación
Hidráulica UGH	Venta	SZR	BARER	Generación
Hulla antracita	Venta	HC	HC G	Generación
Hulla antracita	Venta	HC	HC G	Generación
Comercializadores mercado libre	Compra	SZR	ABOUT	Comercializadores

Tabla 15 - Extracto del archivo *sum_syb_intra1_intra2_intra3_from_20201101_to_20211118.csv*, resultado de un estudio de volumen

6.2 HISTÓRICO UNIDADES DE PROGRAMACIÓN

La finalidad es generar un dataframe que contenga, una vez escogido:

- Unidades de Programación
- Horizonte temporal
- Mercados
- Tipo de resultado
 - Desagregado
 - Subidas y Bajadas (por horas, o total diario)

El dataframe una primera columna de fechas, seguida de la información escogida en los parámetros, de esta forma obtenemos un histórico de las acciones de una UP determinada. Si se escoge más de una unidad de programación, se generarán distintos archivos. De forma

similar al archivo de volumen de movimientos por UP, explicado en el 6.1, los parámetros escogidos quedan recogidos en el nombre del archivo. Por ejemplo:

“ABA1_rh_restricciones_intra1_intra2_intra3_intra4_intra5_intra6_intra7_syb_from_20201101_to_20201118.csv”

Fecha	Total subidas	Total bajadas
2020-11-01	0	7.2
2020-11-02	0	7.2
2020-11-03	0	7.2
2020-11-04	0	7.2
2020-11-05	0	7.2
2020-11-06	0	7.2
2020-11-07	3.2	9.3
2020-11-08	4.1	10.8
2020-11-09	2.4	9.6
2020-11-10	0	7.2
2020-11-11	0	7.2
2020-11-12	0	7.2
2020-11-13	0	7.2
2020-11-14	0	7.2
2020-11-15	0	7.2
2020-11-16	0	7.2
2020-11-17	0	7.2
2020-11-18	0	7.2

Tabla 16 - Tabla del archivo

“ABA1_rh_restricciones_intra1_intra2_intra3_intra4_intra5_intra6_intra7_syb_from_20201101_to_20201118.csv”, resultado de la función histórico unidades de programación

Capítulo 7. MODELOS FUNDAMENTALES

7.1 INTRODUCCIÓN A LOS MODELOS

Estos modelos están basados en el análisis de las estrategias diarias de los agentes utilizando los ficheros I90DIA y el histórico del precio spot diario como referencia. Todos tienen como finalidad predecir el comportamiento del mercado del día siguiente conociendo los movimientos de los agentes del día actual, bajo la suposición de que los movimientos que realizarán los agentes el día posterior serán similares a los del día actual.

7.1.1 ESTRATEGIA UTILIZADA

Todos los modelos comparten la misma estrategia:

- Se calculan una serie de parámetros a partir de datos obtenidos del I90DIA y el histórico del precio spot, es decir, datos históricos como máximo hasta tres meses previos a la fecha actual.
- Se utilizan los parámetros calculados previamente para hacer las predicciones conociendo los movimientos de hoy para mañana.

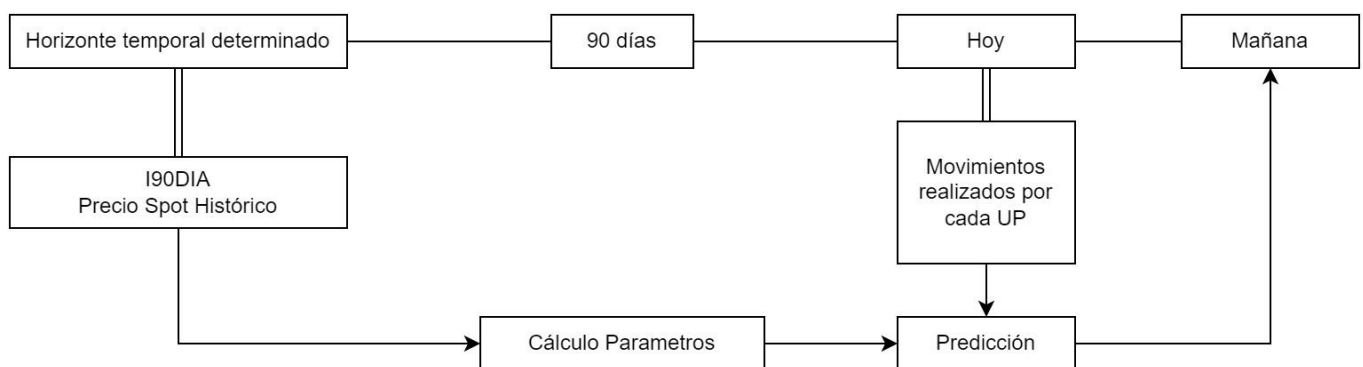


Ilustración 12 - Esquema Estrategia Modelos Fundamentales

7.2 PARÁMETROS CALCULADOS

7.2.1 RELACIÓN MOVIMIENTO – PRECIO

En función de los movimientos en la cantidad de energía por agente y del cambio en el precio a esa hora en ese mercado, se calcula el porcentaje de veces que cada unidad de programación “compra caro”, “compra barato”, “vende caro”, “vende barato”.

- Se considera “comprar” tanto aumentar el valor de compra, como reducir el valor de venta.
- Se considera “vender” tanto aumentar el valor de venta, como reducir el valor de compra.
- Se considera “encarece” cuando el precio tras el mercado es superior a la foto anterior.
- Se considera “abarata” cuando el precio tras el mercado es inferior al precio del mercado anterior.

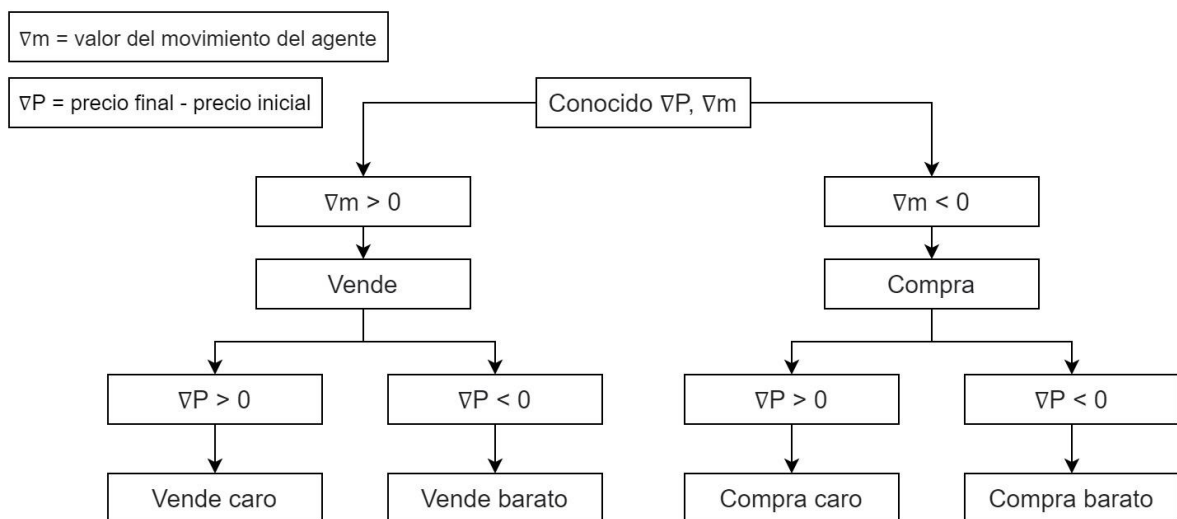


Ilustración 13 - Esquema descriptivo relación precio – movimientos

Se obtiene el porcentaje de veces que cada UP se clasifica en uno de estos cuatro casos.

7.2.2 SENSIBILIDAD DEL PRECIO RESPECTO A LOS MOVIMIENTOS

Se estudia la relación entre la magnitud del movimiento realizado por un agente, y el cambio que se desencadena en el precio tras este movimiento, estudiando así la sensibilidad del precio spot ante los movimientos de cada UP.

$$\text{Sensibilidad} = \frac{\Delta P}{\Delta m}$$

$$\Delta P = \text{Precio Final} - \text{Precio Inicial};$$

$$\Delta m \rightarrow \text{Movimiento del agente}$$

Nótese que el valor resultado puede ser positivo, negativo o cero. Cuando es positivo significa que el agente “afecta” al mercado como debería, por ejemplo:

$$\Delta P > 0; \Delta m > 0 \rightarrow S > 0$$

En el ejemplo mostrado anteriormente, el agente decide comprar más, y el mercado encarece. Tiene sentido que el signo de la sensibilidad sea positivo, ya que un incremento de demanda debería desencadenar una subida de precio. Esta operación se realiza para cada hora de cada día en cada mercado.

7.2.2.1 Sensibilidad calculada de forma desagregada

Esta métrica se calcula también de forma separada para situaciones en las que el mercado encarece o abarata. En estos casos, las unidades sólo suman cuando su movimiento afecta a la evolución del mercado de forma coherente, es decir, si cuando compran más o venden menos el mercado encarece. Por ello, el sumatorio de las sensibilidades será siempre mayor o igual a cero. En el siguiente esquema queda plasmado el criterio utilizado para cada unidad de programación, cada hora de cada día en cada mercado:

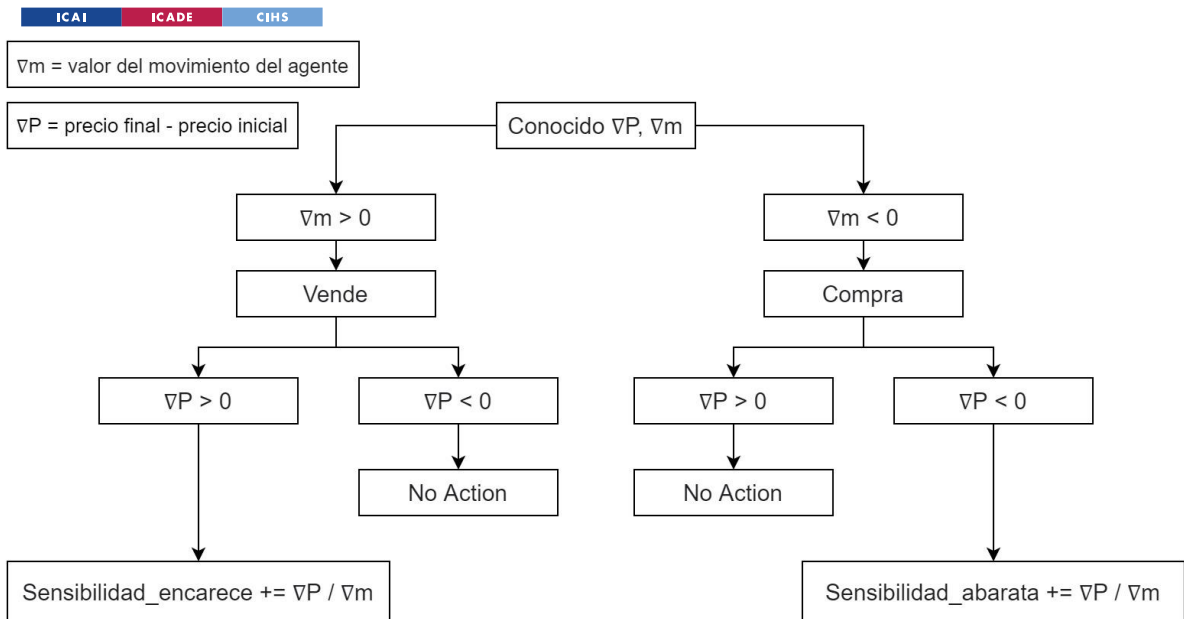


Ilustración 14 - Esquema explicativo criterio cálculo sensibilidad relativa

Categorizaremos a aquellos agentes que obtengan un valor elevado de sensibilidad a encarecer como “unidades peligrosas”, y a aquellos que tengan un valor elevado de sensibilidad a abaratar como “unidades no peligrosas”.

Además, se calculan también los valores corregidos de cada sensibilidad, es decir, divididos entre la participación del agente. De esta forma, una unidad que tiene un gran impacto sobre el precio, pero aparece pocas veces en el mercado queda resaltada por igual en los valores relativos.

7.2.3 FUNCIÓN PARA CALCULAR LOS PARÁMETROS

Tras especificar el horizonte temporal, mercados, y porcentaje de UP a escoger por volumen (descrito en el 6.1) deseados. El código desarrolla las siguientes acciones:

- Creación de un listado del top UP por volumen en el horizonte temporal escogido.
- Bucle que recorre cada hora, de cada día, de cada mercado, abriendo el archivo CSV desagregado apropiado, y haciendo los cálculos especificados en el **¡Error! No se encuentra el origen de la referencia..**
- Adición de información adicional procedente del Archivo CSV Información Adicional UP.

- Exportación de la tabla resultante a un archivo CSV, para su posterior utilización en modelos.

7.2.4 ANÁLISIS DE LOS PARÁMETROS RESULTANTES

Se realiza un estudio de los parámetros resultantes al ser calculados en un horizonte temporal proporcionando visualización de los mismos siguiendo los siguientes filtros:

- Movimientos de los mercados intradiario uno, dos y tres.
- Horizonte temporal desde el 11 de noviembre de 2020 al 1 de marzo de 2021.
- Porcentaje de selección del 25 % de unidades con mayor volumen transaccionado.

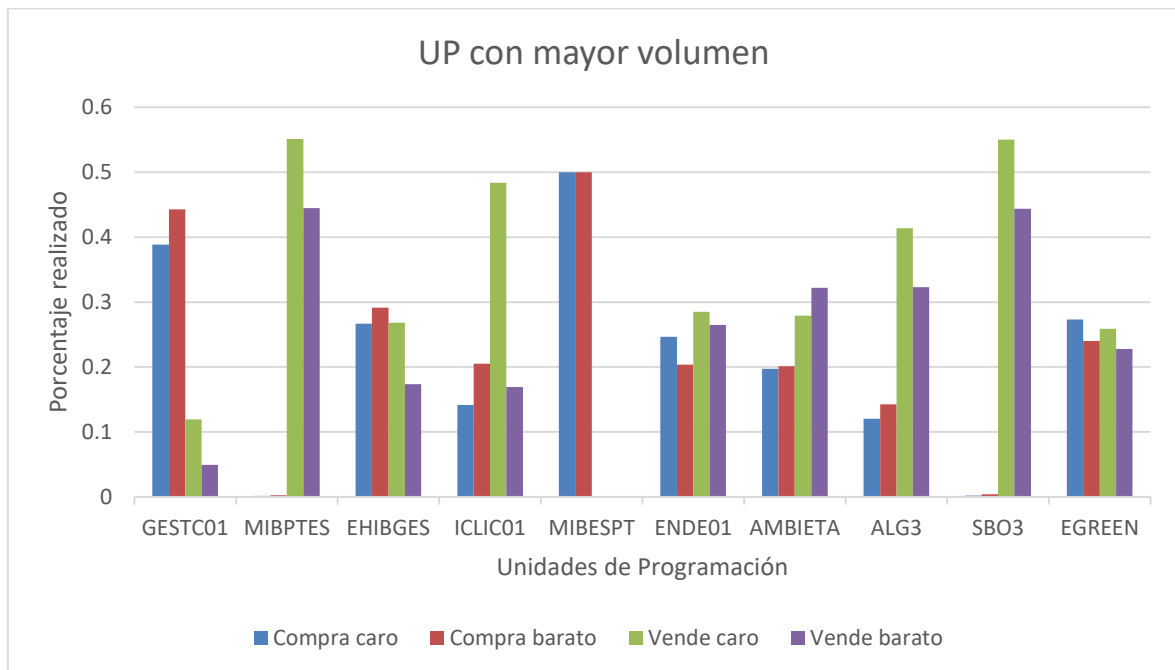


Ilustración 15 - Estudio de los porcentajes compra-vende barato o caro de las unidades de programación que tienen un mayor volumen

En la anterior gráfica se realiza un estudio de los porcentajes compra-vende barato o caro de las unidades de programación que tienen un mayor volumen de movimientos. Tiene sentido que tanto MIBPTES como MIBESPT solo tengan o venta o compra respectivamente al ser saldo de acoplamiento del mercado.

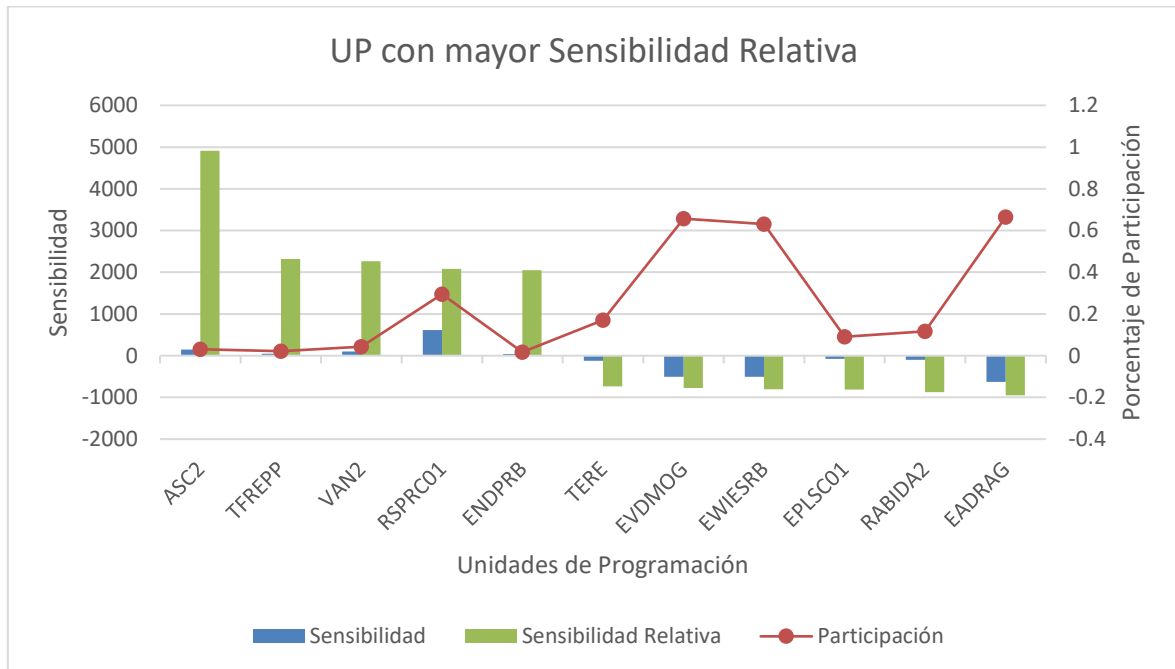


Ilustración 16 - Unidades de programación que tienen el mayor valor absoluto de sensibilidad relativa

En la anterior gráfica quedan representadas las unidades de programación que tienen el mayor valor absoluto del sumatorio de Sensibilidad del precio respecto a los movimientos dividido entre su participación. Cabe destacar como muchas de estas UP tienen una participación muy reducida, lo que significa que cuando entran en el mercado causan un gran impacto.

Sensibilidad del precio respecto a los movimientos

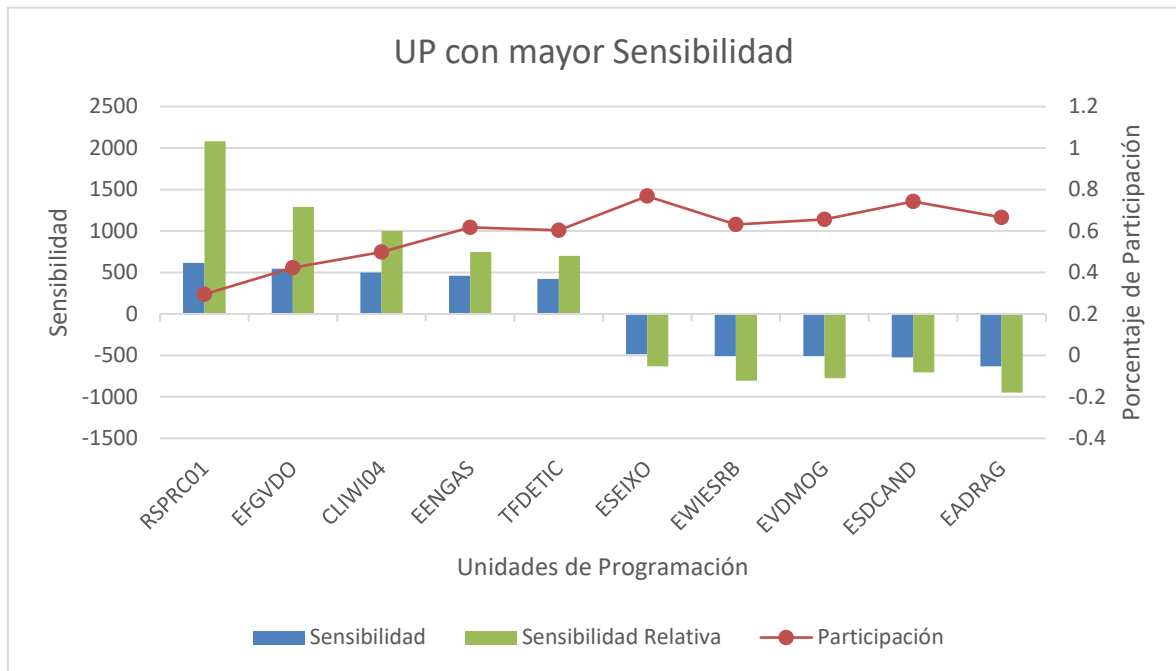


Ilustración 17 - Unidades de programación que tienen el mayor valor absoluto de sensibilidad

En esta gráfica se realiza el mismo proceso, pero para unidades con mayor sensibilidad absoluta. Cabe valorar como EADRAG se encuentra en ambas gráficas, lo que indica que es una Unidad de Programación que afecta mucho y de forma frecuente al mercado.

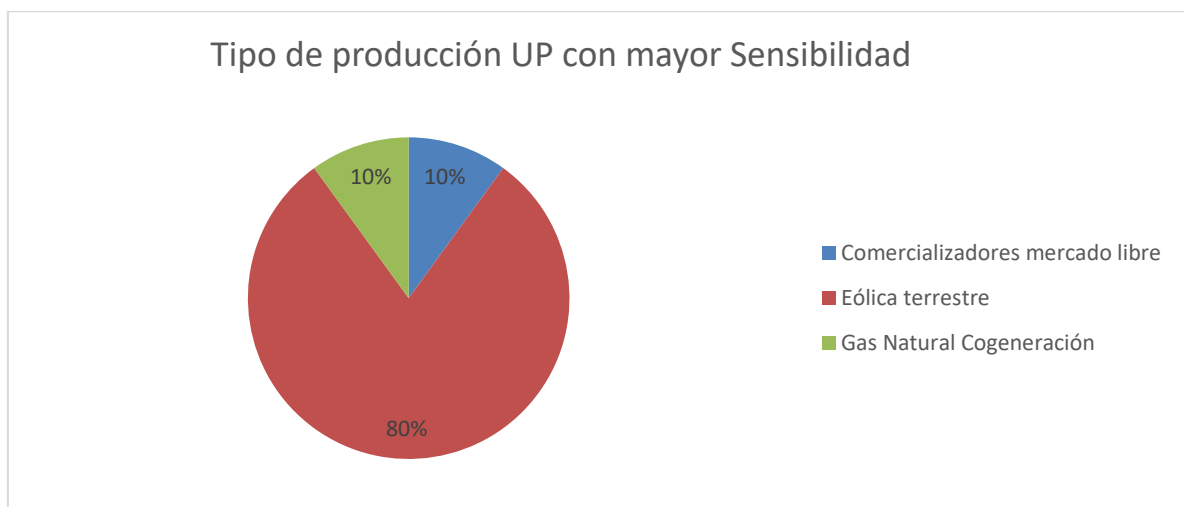


Ilustración 18 - Tipo de producción de las unidades con mayor sensibilidad

En esta última gráfica se analiza el tipo de producción de las unidades de programación que tienen una mayor sensibilidad, se puede observar como en su mayoría son Eólica terrestre, observación a tener en cuenta a la hora de predecir el comportamiento del mercado.

Las gráficas mostradas a continuación, además de los filtros impuestos previamente, se añade:

- Porcentaje de selección para unidades peligrosas y no peligrosas del 20 % superior sobre el 25 % de unidades con mayor volumen transaccionado. Lo que implica que se escogerá un 5 % de las unidades totales en ambos listados.

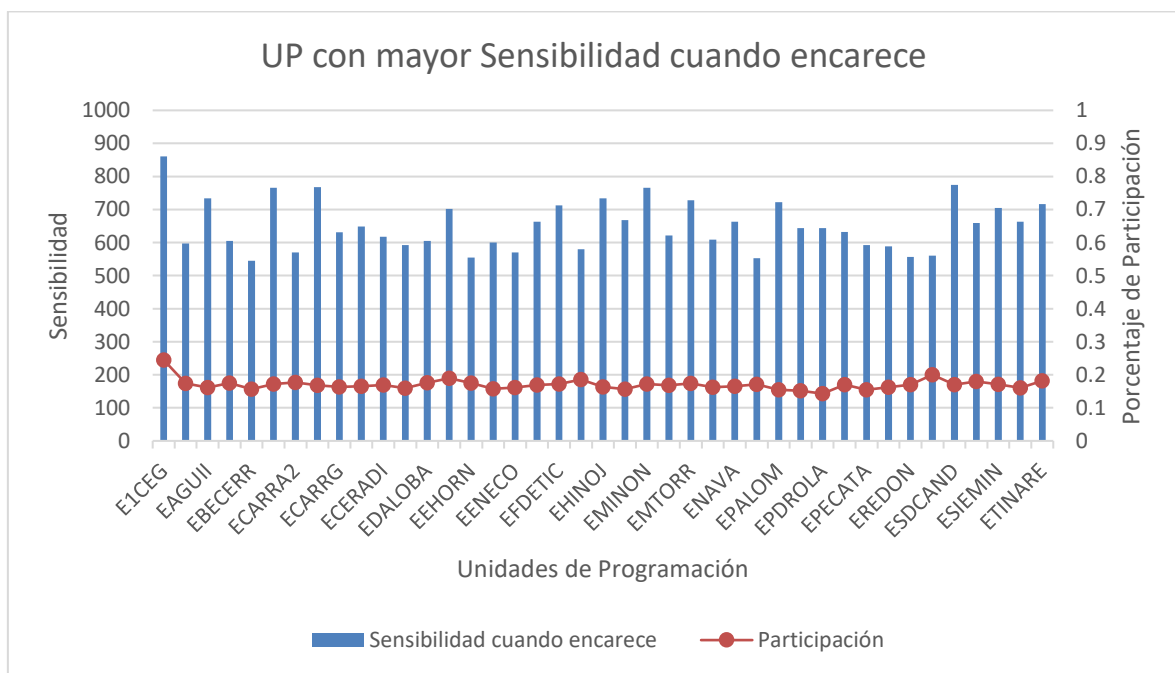


Ilustración 19- Diagrama de las unidades con mayor sensibilidad cuando el mercado encarece

Se puede observar cómo, pese a tener una participación reducida, estas unidades suelen lograr que el mercado encarezca cuando sus transacciones son de tipo comprar más o vender menos, es decir, aumento de demanda.

Tipo de producción UP con mayor Sensibilidad cuando encarece

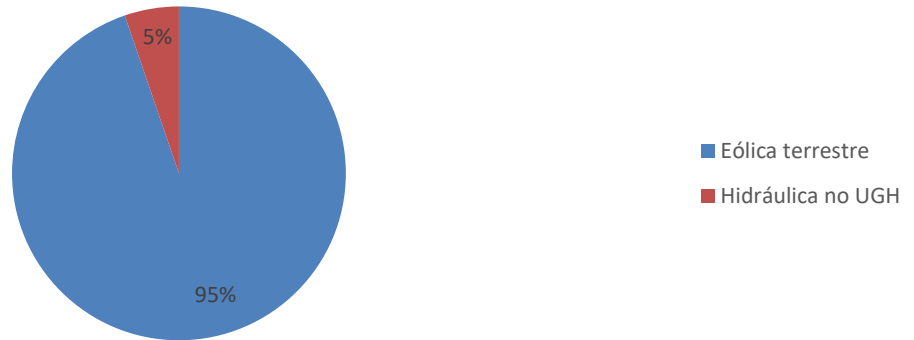


Ilustración 20 - Tipo de producción de las unidades con mayor sensibilidad cuando el mercado encarece

En la gráfica anterior se puede apreciar cómo el 95 % de las unidades que más afectan a que el mercado encarezca son Eólica terrestre.

UP con mayor Sensibilidad cuando abarata

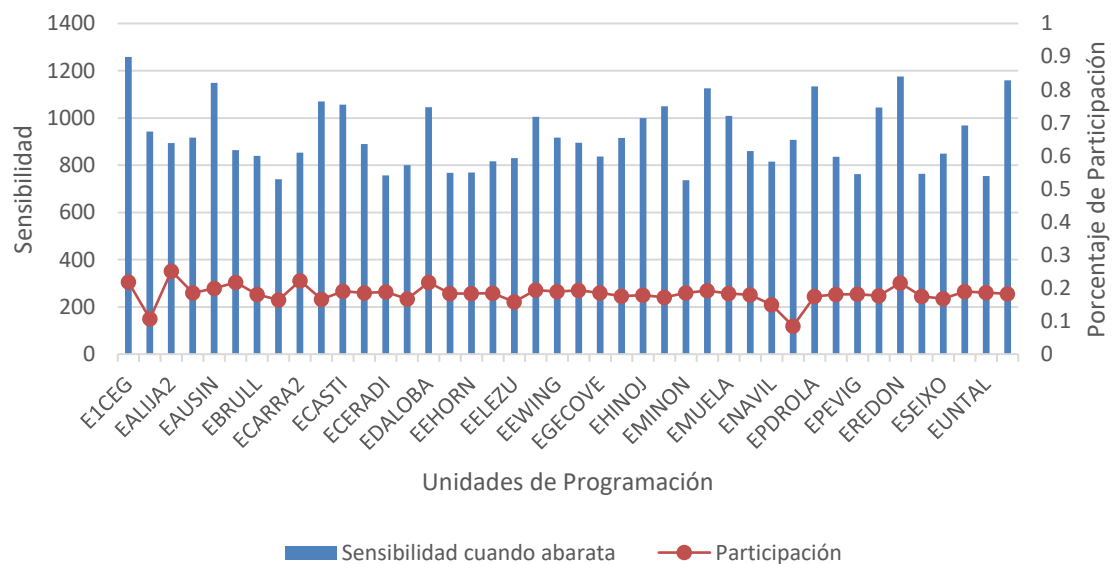


Ilustración 21- Diagrama de las unidades con mayor sensibilidad cuando el mercado abarata

Gráfica similar a la Ilustración 19 salvo con las principales unidades cuando el mercado abarata.

Tipo de producción UP con mayor Sensibilidad cuando abarata

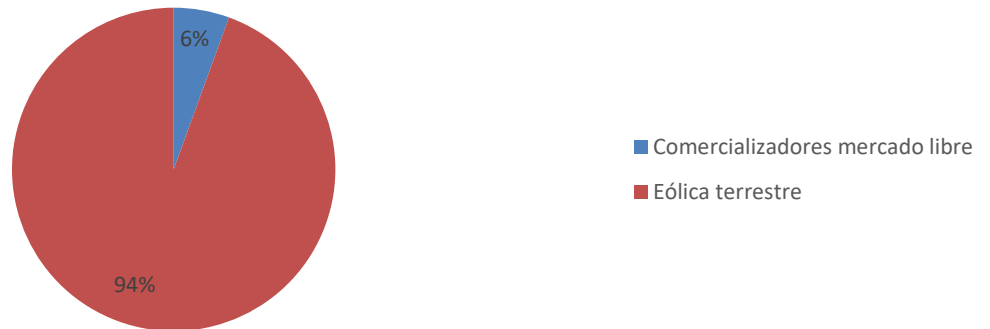


Ilustración 22 - Tipo de producción de las unidades con mayor sensibilidad cuando el mercado abarata

El caso no difiere casi nada respecto al tipo de producción cuando encarece, cuando abarata se obtienen un 94 % de Eólica terrestre.

7.2.4.1 Conclusiones de los parámetros resultantes

- Gracias al cálculo de estos parámetros se pueden identificar aquellas unidades de programación que tienen un gran impacto en el comportamiento del mercado, en este ejemplo particular, EADRAG resalta como la unidad que más impacto se considera que tiene.
- Eólica terrestre es el tipo de producción que mayor impacto tiene sobre las fluctuaciones de precio en los mercados intradiarios.

7.3 PRIMER MODELO – RANKINGS POR TENDENCIAS

Este primer modelo predice el precio mediante la creación de un ranking según el porcentaje de veces que las unidades de programación venden barato y compran caro.

7.3.1 CÁLCULOS REALIZADOS SOBRE LOS MOVIMIENTOS DEL DÍA ACTUAL

Utilizando los parámetros calculados a raíz del I90DIA y el precio spot histórico se calculan los siguientes valores **para las unidades de programación que han estado presentes hoy.**

$$\text{Valor A} = (\text{Total Volumen Movimientos}) * \frac{\% \text{Vende barato}}{\% \text{Vende}}$$

$$\text{Valor B} = (\text{Total Volumen Movimientos}) * \frac{\% \text{Compra Caro}}{\% \text{Compra}}$$

Una vez obtenidos ambos valores para todas las unidades de programación participantes en el mercado del presente día, se crean dos rankings independientes: en uno se encuentran las UP ordenadas de mayor a menor por su valor A, el otro ranking sigue el mismo criterio salvo que con el Valor B.

A cada UP en estos rankings se le asigna un valor correspondiente al inverso de su posición. Es decir, la UP que ocupe la primera posición en una lista de 356 UP totales, obtendrá el valor 356.

UP	Valor A	Valor B	Valor Asignado por A	Valor Asignado por B
UP1	2658	1598	3	1
UP2	1265	5245	1	4
UP3	1585	3572	2	3
UP4	9872	2587	4	2

Tabla 17 - Criterio asignación de valores por rankings

De esta forma, se tienen registradas aquellas UP que:

- A. Cuando venden el mercado tiende a abaratar
- B. Cuando compran el mercado tiende a encarecer

Además de tener un valor registrado que indica la “cantidad de tendencia” que generan.

$$\text{Valor resultante} = \sum \text{Valores Ranking B} - \sum \text{Valores Ranking A}$$

7.3.2 PREDICCIÓN EN BASE AL VALOR RESULTANTE

Este valor resultante viene a indicar la cantidad e importancia de las UP que hay en el sistema, siendo este valor mayor cuanto mayor presencia tienen las UP que cuando compran el precio sale caro. Esto puede indicar en un día en concreto que estén activas en el mercado más unidades con predisposición a comprar y encarecer que unidades con predisposición a vender y abaratar. Por ello, un valor resultante mayor debería suponer un encarecimiento del mercado.

7.3.3 RESULTADOS

Se realiza un estudio del modelo del 1 al 10 de mayo de 2021 que cumple los siguientes criterios:

- Se estudian mercados intradiario uno, dos y tres.
- El horizonte temporal escogido para el cálculo de los parámetros es: 11 de noviembre de 2020 al 1 de marzo de 2021. Nótese cómo el último día del horizonte temporal no sobrepasa los 90 días previos a la fecha que se quiere predecir.

7.3.3.1 Análisis del valor resultante cuando encarece y cuando no encarece el mercado

Se hace un estudio del valor resultante en función del comportamiento del mercado para estudiar si existe algún tipo de relación entre ambos. El horizonte temporal en el que se calculan estos valores resultantes será del 1 al 10 de mayo de 2021.

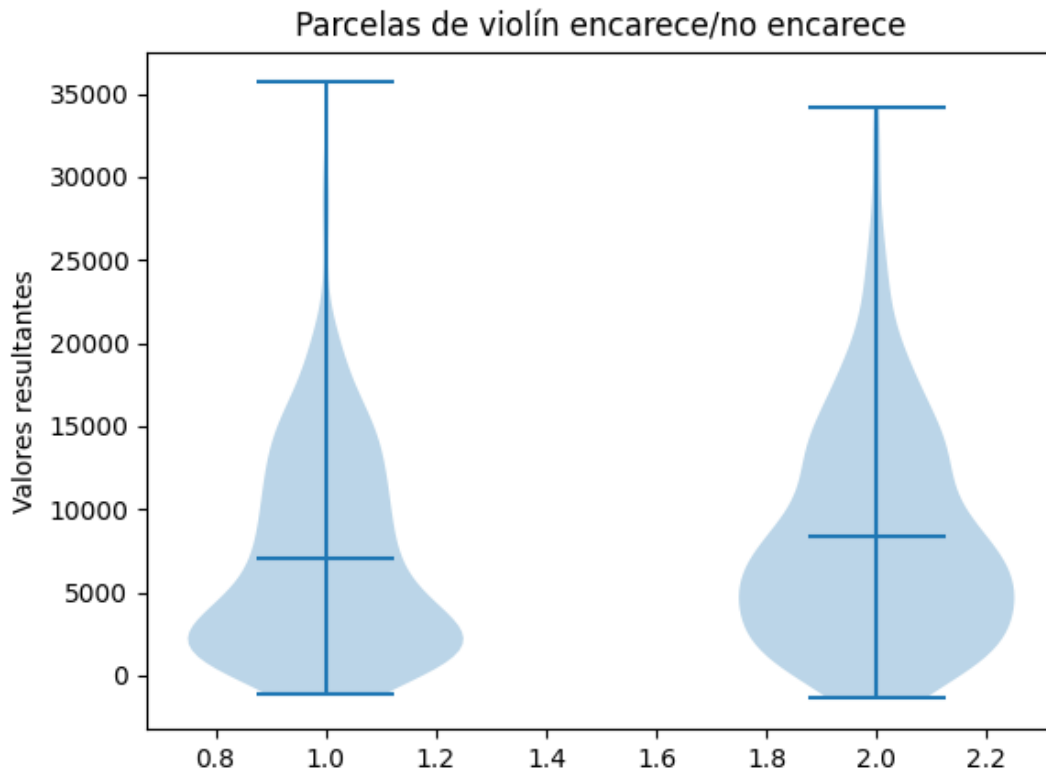


Ilustración 23- Parcelas de violín de los valores resultantes del primer modelo

Como se puede observar en la ilustración 23, en contra de lo esperado, el valor resultante suele ser mayor cuando el mercado no se encarece. Esta salida no sigue el criterio esperado ya que un mayor valor resultante significa que hay más unidades con predisposición a vender encareciendo el mercado, sin embargo, en las parcelas de violín se observa el resultado opuesto. Además, no existe una clara diferencia de los valores obtenidos en función de si el mercado encarece o abarata ya que ambas distribuciones se encuentran por la misma zona de valores.

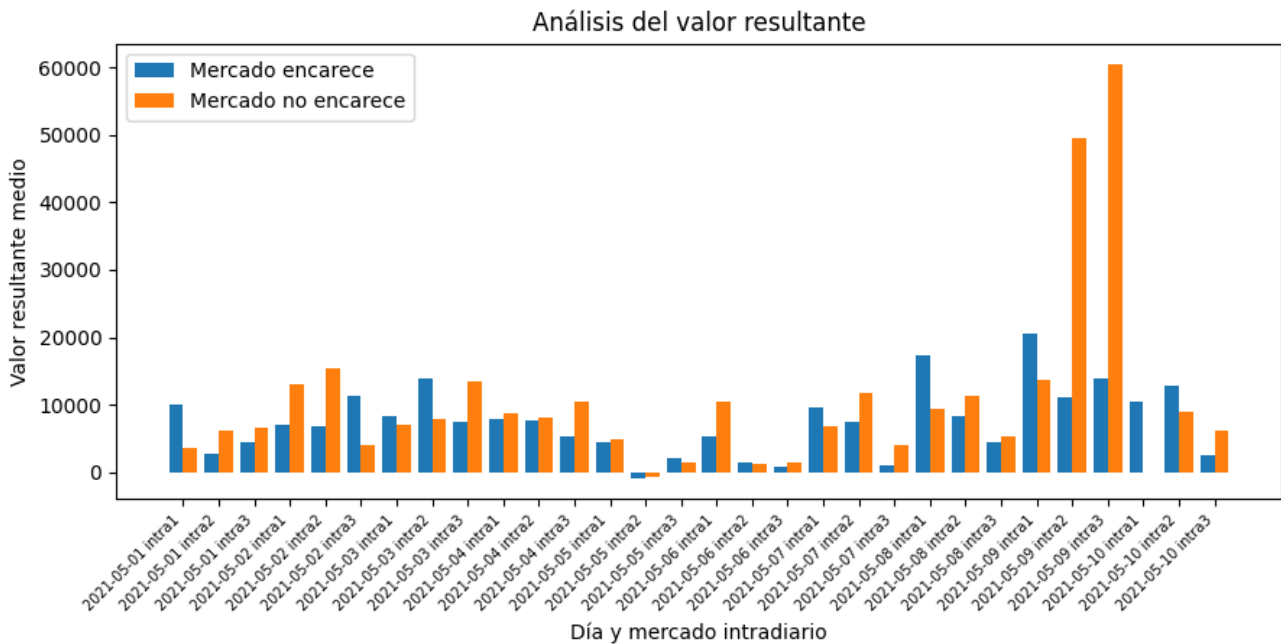


Ilustración 24 - Diagrama de barras del valor medio resultante diario del primer modelo

Se puede confirmar en la ilustración 24 cómo este valor no permite marcar una diferenciación sobre si el mercado encarece o no. Queda plasmado como la relación entre el valor resultante y el comportamiento del mercado es completamente trivial.

7.3.3.2 Conclusiones del modelo

Tal y como se ha observado en el análisis del valor resultante en función del comportamiento del mercado, no existen relaciones entre ambos, por lo que se decide no continuar desarrollando el modelo a partir de este punto.

7.4 SEGUNDO MODELO – AGENTES INFLUYENTES

Predice el precio basándose en los movimientos de los agentes más influyentes presentes en el día, considerando agente influyente aquel que tenga un alto valor de sensibilidad corregida, denominada sensibilidad relativa.

7.4.1 CÁLCULOS REALIZADOS SOBRE LOS MOVIMIENTOS DEL DÍA ACTUAL

Utilizando los parámetros calculados a raíz del I90DIA y el precio spot histórico se realiza una selección de los agentes a considerar como más influyentes. Para ello, se hace una selección del porcentaje de unidades de programación con un mayor valor (en valor absoluto) de sensibilidad relativa.

Con aquellas unidades presentes en el día actual que estén catalogadas como influyentes, se realiza el siguiente sumatorio para obtener el valor resultante:

$$\text{Valor resultante} = \sum_{UP \text{ influyentes}} (\text{Valor Sensibilidad Relativa}) * (\text{Valor del movimiento del agente})$$

7.4.2 PREDICCIÓN EN BASE AL VALOR RESULTANTE

Cuando un agente tiene un valor de sensibilidad relativa positivo, es indicativo de que el mercado tiende a encarecer cuando compra, y a abaratar cuando vende.

De forma contraria, si un agente tiene un valor de sensibilidad relativa negativo, es indicativo de que el mercado tiende a abaratar cuando compra, y a encarecer cuando vende.

En el producto del sumatorio del valor resultante existen cuatro escenarios posibles:

1. Agente compra y tiene valor de sensibilidad positivo: el mercado es influenciado a encarecer.
2. Agente compra y tiene valor de sensibilidad negativo: el mercado es influenciado a abaratar.
3. Agente vende y tiene valor de sensibilidad positivo: el mercado es influenciado a abaratar.
4. Agente vende y tiene valor de sensibilidad negativo: el mercado es influenciado a encarecer.

Por ello, siguiendo este criterio, se hipotetiza lo siguiente:

- A un mayor valor resultante, hay más peso en influencia a encarecer, la predicción debe ser que el mercado encarece.

- A un menor valor resultante, hay más peso en influencia a abaratar, la predicción debe ser que el mercado abarate.

7.4.3 RESULTADOS

Se establecen los mismos criterios y horizontes temporales utilizados para el análisis de resultados del primer modelo, véase 7.3.3.

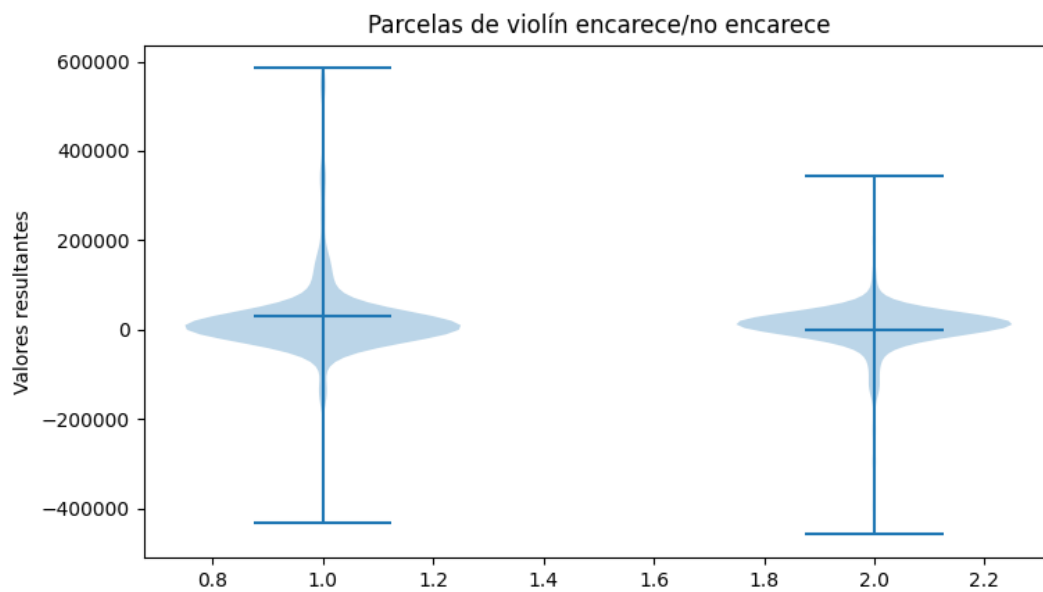


Ilustración 25 - Parcelas de violín de los valores resultantes del segundo modelo

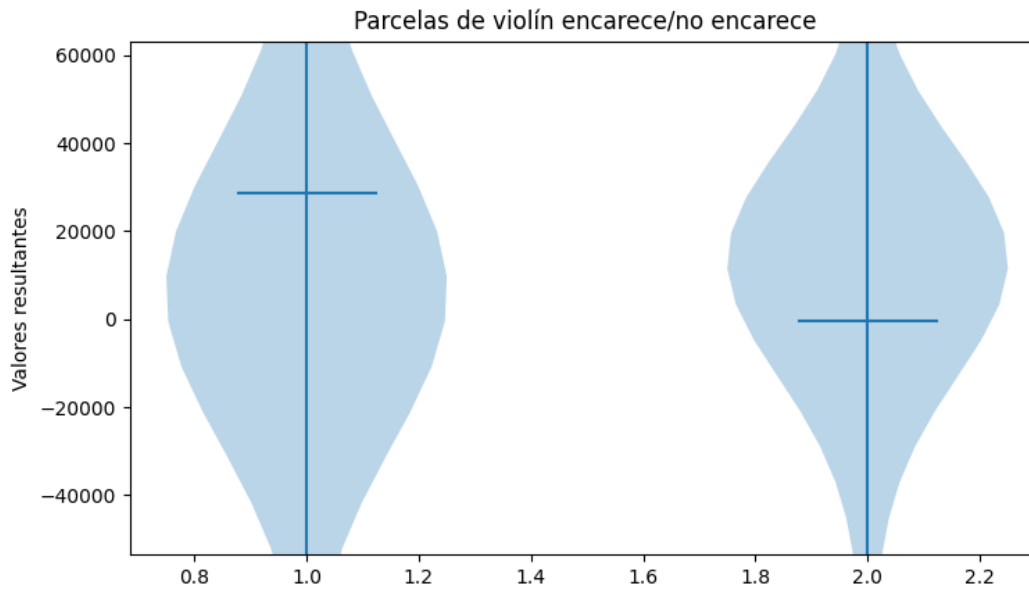


Ilustración 26 - Zoom Parcelas de violín de los valores resultantes del primer modelo

En las anteriores parcelas de violín se observa el resultado esperado. En la figura X1 quedan reflejadas una cola para valores más elevados en el caso de que el mercado encarezca, y en valores reducidos en el caso de que no encarezca, tal y como se había hipotetizado en la introducción del 7.4.2. Estas colas en las distribuciones hacen que exista una clara diferencia en las medias, tal y como queda plasmado en la figura X2. Sin embargo, las distribuciones

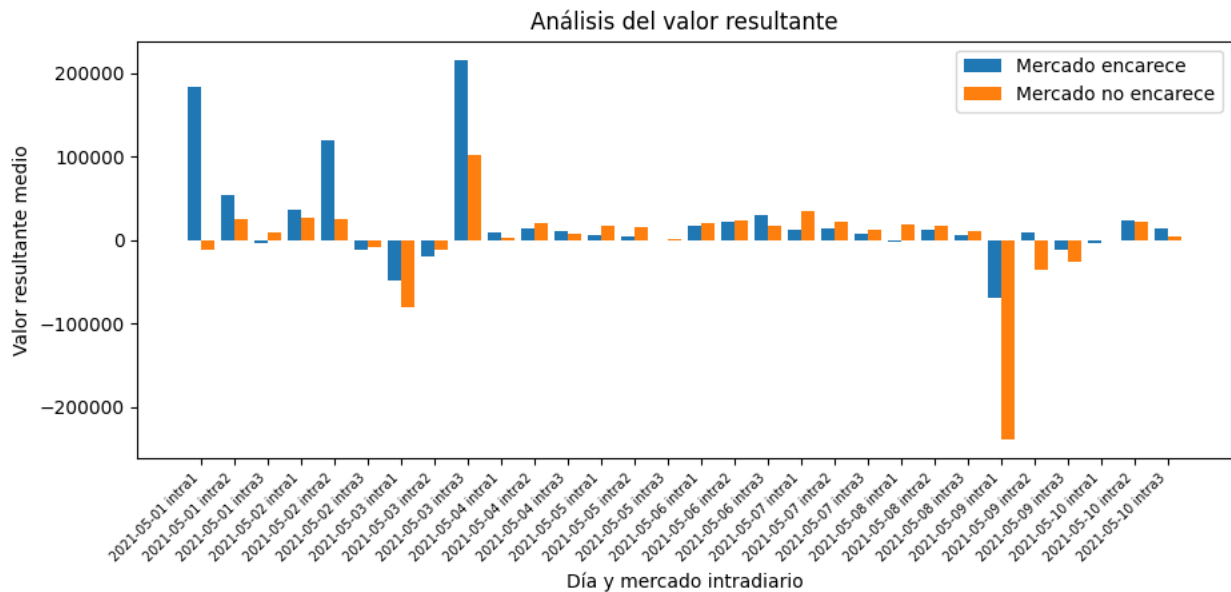


Ilustración 27 - - Diagrama de barras del valor medio resultante diario del segundo modelo

se encuentran menos distanciadas de lo esperado, lo que desembocará en un ratio de error elevado.

En la anterior gráfica queda plasmado como los valores resultantes suelen ser superiores (o menos negativos) en el caso de que el mercado encarezca, con ciertas excepciones, aunque es cierto que, en los casos de fallo, la diferencia entre los valores es mínima, factor que indica que eran casos en los que el modelo tendría “dudas” a la hora de realizar una predicción.

7.4.3.1 Predicción en base al valor resultante

Tal y como se ha podido analizar en la Figura 26, las distribuciones no se encuentran significativamente distanciadas, por lo que se decide crear un umbral de incertidumbre en el que este modelo fundamental no haga ninguna predicción. Fuera del umbral, valores superiores se interpretarán como que el mercado debe encarecer, y valores inferiores, el caso contrario.

Dado que este estudio se ha realizado del 1 al 10 de mayo de 2021, las predicciones se conducirán los próximos tres días, por lo que habrá 24 predicciones por cada uno de los 3 mercados 3 días, es decir, 216 predicciones que se pueden realizar. El umbral de

incertidumbre se escoge para valores de entre -50000 y 50000. Para valores dentro de este rango, el modelo no hará ninguna predicción.

Los resultados obtenidos son:

$$\text{Porcentaje predecido} = \frac{30}{216} = 13.89 \%$$

Este valor indica el porcentaje de veces que el modelo es capaz de realizar una predicción. De este porcentaje, el rendimiento ha sido el siguiente:

$$\text{Porcentaje de acierto} = \frac{28}{30} = 93.33 \%$$

7.4.3.2 Conclusiones del modelo

Pese a la gran precisión del modelo cuando hace una predicción, el hecho de que sólo se pueda garantizar dicha precisión cuando se hacen predicciones un 10-15 % de las veces, hace que este modelo por sí solo no se pueda utilizar.

7.5 TERCER MODELO – AGENTES PELIGROSOS VS NO PELIGROSOS

Predice el precio basándose en el porcentaje de agentes categorizados como peligrosos y no peligrosos presentes ese día, una UP se considera más peligrosa cuanto mayor sea su valor de sensibilidad cuando el mercado encarece. Por el contrario, será menos peligrosa cuanto mayor sea su valor de sensibilidad cuando el mercado abarata.

Se recogen en dos listados por separado las unidades más peligrosas y no peligrosas para poder analizar su presencia en el día actual. Debido a que pueden existir duplicados entre ambos listados, se crea un tercer listado que contenga las coincidentes, eliminando estas unidades de los dos listados en los que estaban previamente.

7.5.1 CÁLCULOS REALIZADOS SOBRE LOS MOVIMIENTOS DEL DÍA ACTUAL

Partiendo de las unidades que han entrado al mercado ese día, se calcula el porcentaje de unidades peligrosas, no peligrosas y coincidentes que se encuentran en el mercado.

7.5.2 PREDICCIÓN EN BASE A LOS PORCENTAJES OBTENIDOS

Según los criterios con los que se han calculado los parámetros:

- A. Las unidades peligrosas suelen influir en un encarecimiento del mercado cuando están presentes.
- B. Las unidades no peligrosas suelen influir en un abaratamiento del mercado cuando están presentes.
- C. No se conoce a priori que influencia pueden tener las coincidentes en el mercado.

Por lo tanto, una mayor presencia de peligrosas vs no peligrosas debería desencadenar un encarecimiento del mercado, y viceversa.

7.5.3 RESULTADOS

Se establecen los mismos criterios y horizontes temporales utilizados para el análisis de resultados del primer modelo, véase 7.3.3.

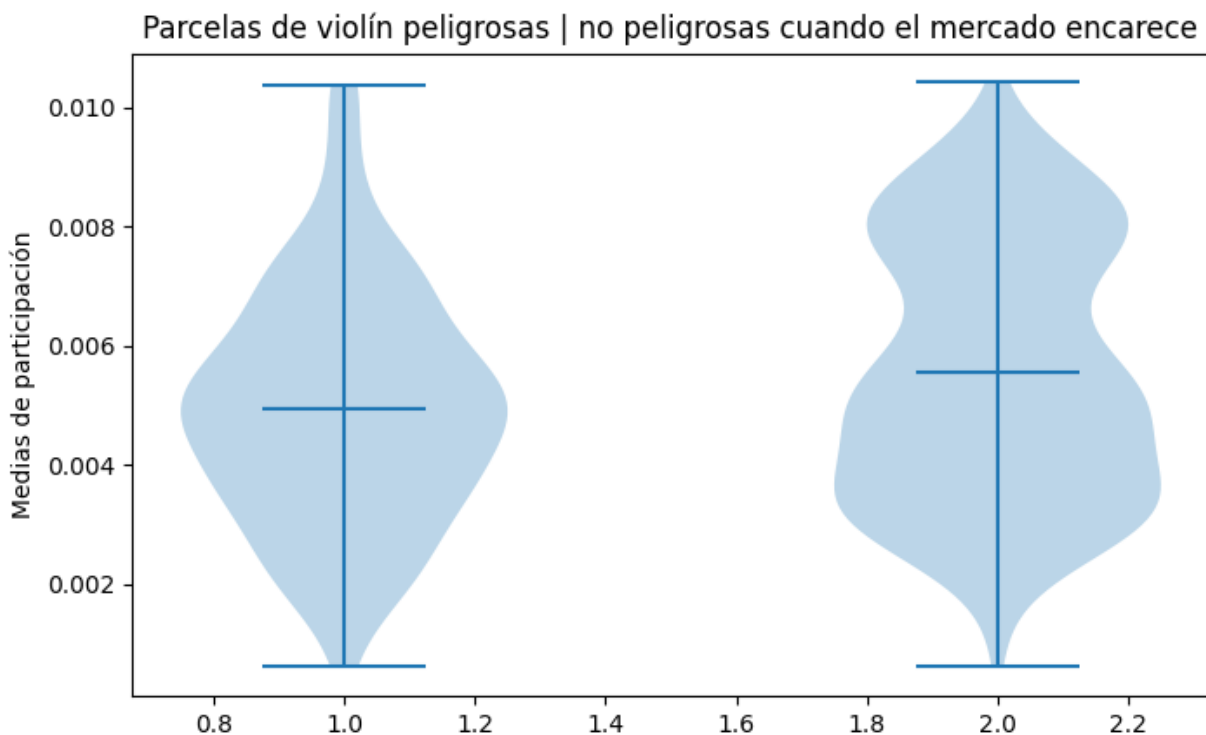


Ilustración 28 - Parcelas de violín de los porcentajes resultantes del tercer modelo cuando el mercado encarece

Los resultados obtenidos en las parcelas son muy interesantes, si bien no existe una relación lineal clara entre los valores y el comportamiento del mercado, se puede observar como las distribuciones difieren en gran medida. La distribución cuando el mercado encarece de unidades peligrosas tiende a concentrarse alrededor de la media, mientras que la de no peligrosas se polariza en los extremos.

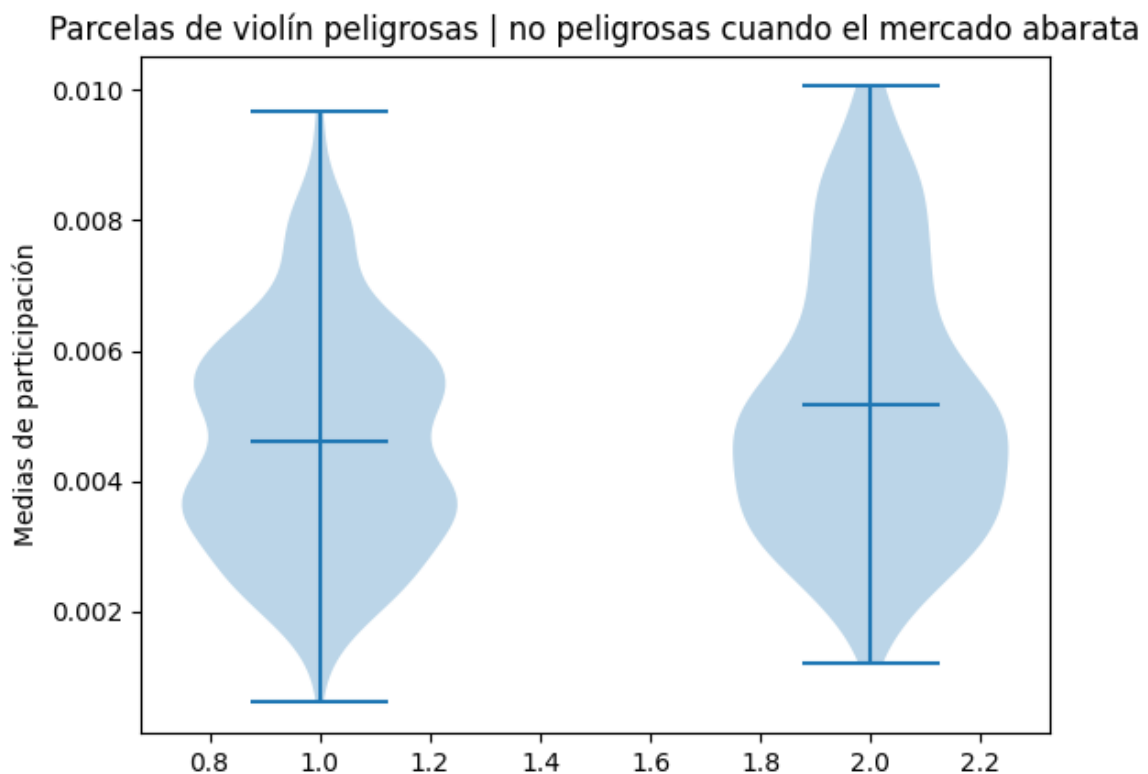


Ilustración 29 - Parcelas de violín de los porcentajes resultantes del tercer modelo cuando el mercado abarata

En la anterior parcela se observa una situación similar respecto a cuando el mercado encarece, pero al revés. Cuando el mercado abarata, los porcentajes de unidades peligrosas se polarizan, mientras que los de no peligrosas tienden a concentrarse alrededor de la media.

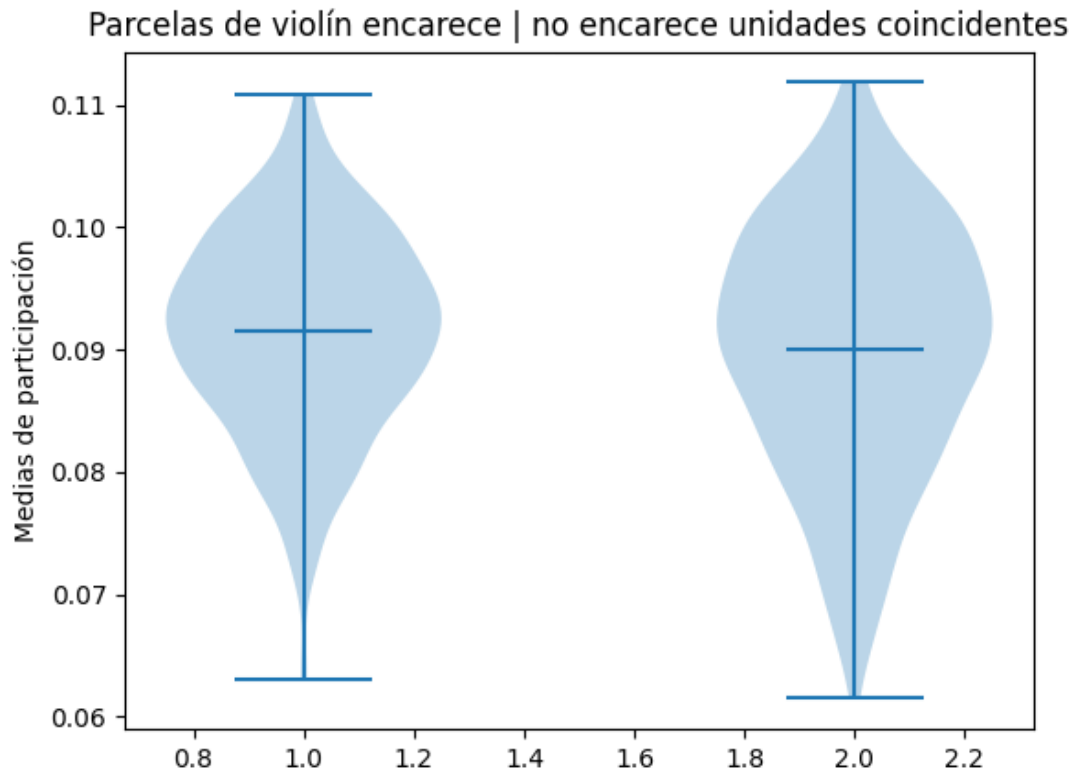


Ilustración 30 - Parcelas de violín de los porcentajes resultantes de las unidades coincidentes del tercer modelo

Respecto a la ilustración 30, se puede extraer la información de que valores muy reducidos de porcentaje de unidades coincidentes sí o sí desemboca en un abaratamiento del mercado.

7.5.3.1 Conclusiones del modelo

Con el estudio realizado de los porcentajes de unidades peligrosas, no peligrosas, y coincidentes en función del comportamiento del mercado, se puede concluir afirmando que no existe una relación lineal clara con la que crear un modelo fundamental. Sin embargo, se han podido observar comportamientos en las distribuciones que indican que existen relaciones no lineales. Por ello, se plantea la creación de modelos con capacidad de estudiar relaciones no lineales, véase Capítulo 8.

Capítulo 8. MODELO BASADO EN REDES NEURONALES

8.1 INTRODUCCIÓN AL MODELO

Se ha podido comprobar que los valores calculados en los modelos fundamentales para la predicción no tienen una relación lineal con la misma, o pueden realizar clasificar la información en reducidas ocasiones, por ello, se decide emplear una red neuronal para intentar buscar si estos valores mantienen alguna relación no lineal respecto a si el mercado encarece o no. Por ello, como entradas de esta red neuronal, se utilizarán los valores resultantes calculados en cada uno de los modelos fundamentales, véase Capítulo 7. , además de otras entradas que puedan tener relación con el comportamiento del mercado.

8.2 VARIABLES DE ENTRADA

Como se ha mencionado en la introducción, se utilizarán los valores resultantes utilizados para hacer las predicciones en los modelos fundamentales, son los siguientes:

8.2.1 VALORES RESULTANTES DE LOS MODELOS FUNDAMENTALES

- Valor resultante primer modelo – Rankings por tendencias, véase 7.3.1.
- Valor resultante segundo modelo – Agentes influyentes, véase 7.4.1.
- Valores resultantes tercer modelo – Agentes peligrosos vs no peligrosos, véase 7.5.1. Serán tres, correspondientes a los porcentajes de peligrosos, no peligrosos, y coincidentes.

8.2.2 OTRAS ENTRADAS

- Hora del día, identificado como número entero del 0 al 23.
- Número del mercado intradiario, identificado como número entero del 1 al 3.

8.2.3 PREPARACIÓN DE LOS DATOS DE ENTRADA

Al estar utilizando como datos de entrada los valores resultantes de los modelos fundamentales, para poder tener un rango de tiempo que utilizar como entrenamiento de la red neuronal, se tendrán que realizar las predicciones a más de 90 días.

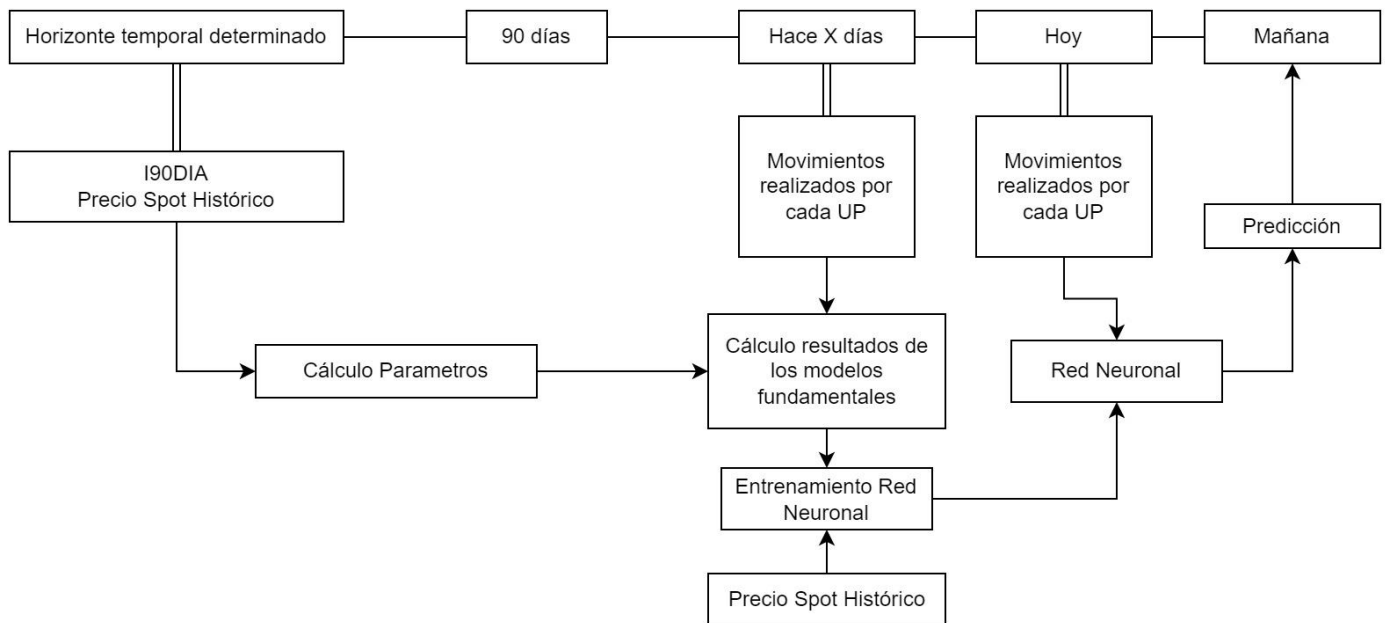


Ilustración 31 - Estrategia Modelo basado en redes neuronales

Además, se escalan todos los datos de entrada de tal forma que se encuentren todos entre -1 y 1, de esta forma todos tienen el mismo peso en la red. Se utiliza la librería StandardScaler para facilitar el proceso.

8.3 DISEÑO DE LA RED NEURONAL

El diseño de la primera y última capa de neuronas viene dado por las entradas y salidas del modelo. En este caso, tendremos siete neuronas en la primera capa, y una neurona en la salida. Sin embargo, la estructura de las capas internas, la función de activación, el ‘solver’ y otras opciones se escogerán para aumentar la precisión del modelo como criterio principal, y reducir el tiempo de procesado como objetivo secundario.

Al tener que realizar una predicción binaria (1, el mercado encarece, 0, el mercado no encarece), se utilizará un clasificador “Multi-layer Perceptron”, un tipo de red neuronal que se entrena utilizando “back propagation”. Se utilizará también “regularization”, variando el valor alfa con la finalidad de evitar overfitting o underfitting.

Los hiperparámetros a probar son los siguientes:

- Función de activación: función utilizada para activar neuronas de la capa siguiente a partir de los valores de las neuronas de la capa anterior.
- Solver: algoritmo de resolución para ajustar los parámetros de la red (pesos y sesgos).
- Learning rate: se puede escoger si debe ser constante, con “inverse scaling” o que se adapte de forma automática.
- Shuffle: mezclar los datos de entrada o no.
- Alpha: valores a probar, se debe ajustar para evitar overfitting o underfitting.
- Número de capas internas
- Tamaño de cada capa interna

8.3.1 TUNNING DE HIPERPARÁMETROS

El proceso de selección de los hiperparámetros está compuesto por dos fases principales: “grid search” y valoración manual. La función “grid search” lleva a cabo el entrenamiento de todas las combinaciones de hiperparámetros deseados escogiendo la combinación que consiga una mayor puntuación. En la valoración manual se analizan los resultados del “grid search” para decidir qué hiperparámetros probar en la siguiente vuelta. La valoración manual es necesaria ya que el proceso de “grid search” se alarga exponencialmente cuantos más hiperparámetros se quieran probar, por lo que reduce el tiempo de tuning considerablemente.

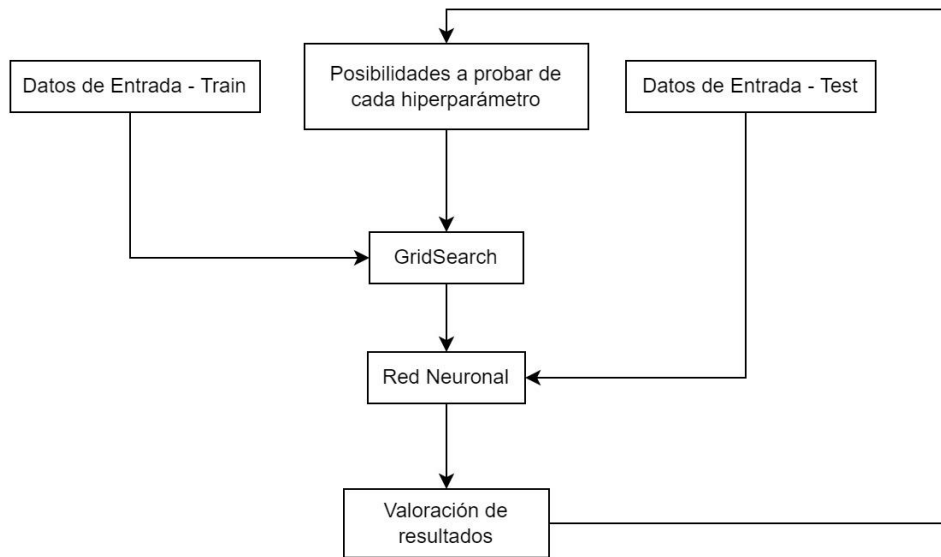


Ilustración 32 - Estrategia de tuning de hiperparámetros

```

0.704 (+/-0.114) for {'activation': 'tanh', 'alpha': 0.05, 'hidden_layer_sizes': (20, 10, 5), 'learning_rate': 'adaptive', 'shuffle': True, 'solver': 'adam'}
0.707 (+/-0.070) for {'activation': 'tanh', 'alpha': 0.05, 'hidden_layer_sizes': (20, 10, 5), 'learning_rate': 'adaptive', 'shuffle': False, 'solver': 'adam'}
0.688 (+/-0.079) for {'activation': 'tanh', 'alpha': 0.05, 'hidden_layer_sizes': (20, 10, 5), 'learning_rate': 'invscaling', 'shuffle': True, 'solver': 'adam'}
0.719 (+/-0.124) for {'activation': 'tanh', 'alpha': 0.05, 'hidden_layer_sizes': (20, 10, 5), 'learning_rate': 'invscaling', 'shuffle': False, 'solver': 'adam'}
0.694 (+/-0.100) for {'activation': 'tanh', 'alpha': 0.075, 'hidden_layer_sizes': (20, 10, 5), 'learning_rate': 'adaptive', 'shuffle': True, 'solver': 'adam'}
0.729 (+/-0.097) for {'activation': 'tanh', 'alpha': 0.075, 'hidden_layer_sizes': (20, 10, 5), 'learning_rate': 'adaptive', 'shuffle': False, 'solver': 'adam'}
0.673 (+/-0.176) for {'activation': 'tanh', 'alpha': 0.075, 'hidden_layer_sizes': (20, 10, 5), 'learning_rate': 'invscaling', 'shuffle': True, 'solver': 'adam'}
0.688 (+/-0.052) for {'activation': 'tanh', 'alpha': 0.075, 'hidden_layer_sizes': (20, 10, 5), 'learning_rate': 'invscaling', 'shuffle': False, 'solver': 'adam'}

Best parameters found:
{'activation': 'tanh', 'alpha': 0.075, 'hidden_layer_sizes': (20, 10, 5), 'learning_rate': 'adaptive', 'shuffle': False, 'solver': 'adam'}
  
```

Ilustración 33 - Ejemplo de “Grid Search”

8.3.1.1 Evitar Overfitting

En el proceso de tuning de hiperparámetros se debe estar pendiente de la precisión que se obtiene en el set de entrenamiento respecto a cómo se comporta con el set de test. En la búsqueda de optimizar el modelo es fácil caer en overfitting con el set de entrenamiento al perseguir una mejor precisión en el mismo. Los métodos utilizados para obtener un fitting óptimo son:

- Aumento de los datos de entrada: se escoge un rango temporal mayor de tal forma que los outliers o casos específicos pierdan peso respecto a la visión general,

obligando a la red a ajustar su predicción a los casos generales reduciendo así overfitting.

- Regularization: Valores elevados en los pesos y sesgos de la red suelen indicar que hay overfitting. El hiperparámetro alfa ajustado en el tuneado es una penalización a los pesos y sesgo de la red. Cuanto mayor es alfa, mayor penalización tendrán pesos y sesgos con valores elevados.
- Simplificación de la red: un mayor número de capas y neuronas constituye una red más compleja, lo que concluye en una mayor probabilidad de realizar overfitting. Como último recurso, se reduciría el número de capas y/o neuronas.

8.3.1.2 Threshold

El valor del threshold determinará si el modelo va a minimizar falsos positivos o negativos. En el caso de este modelo, se tendrá que escoger cuál de los siguientes fallos minimizar:

- Predecir que el mercado encarece y luego no encarece.
- Predecir que el mercado no encarece y luego encarece.

Se prefiere minimizar el falso positivo, es decir, reducir el fallo de la predicción a encarecer y que luego no encarezca. Este criterio se basa en que es preferible dejar pasar una ganancia que coger una pérdida.

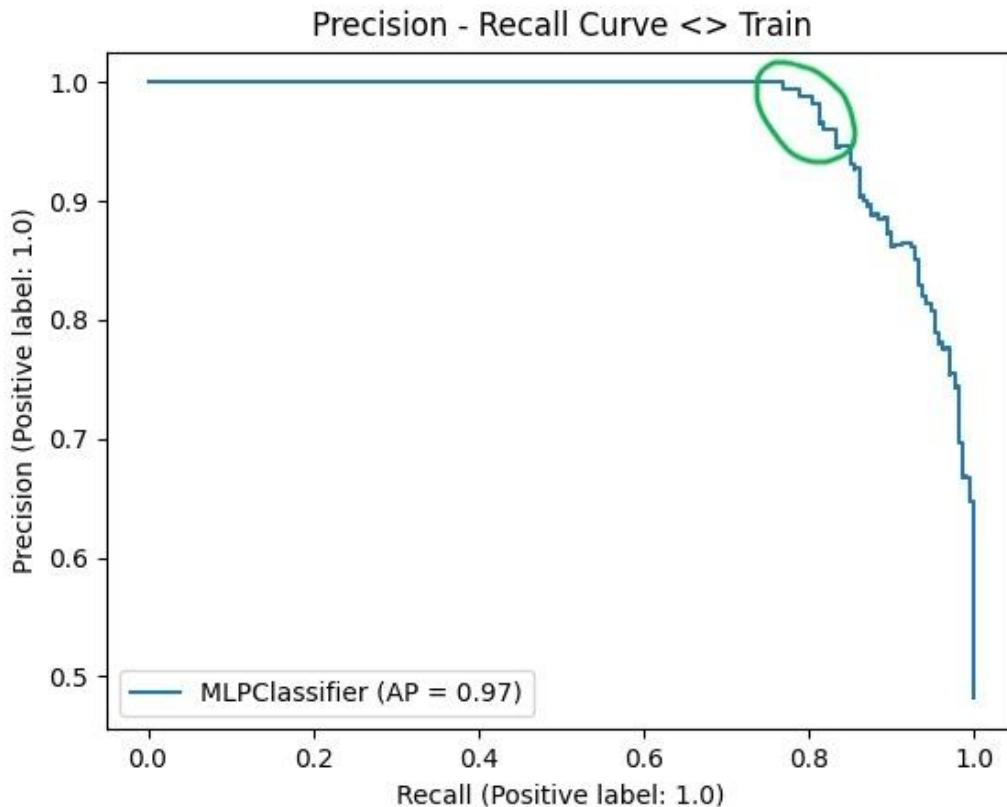


Ilustración 34 - Relación Precision recall en el set de Train

En la ilustración 34 se puede observar la relación entre precisión y “recall”, en el caso de este modelo, se escogerán valores de threshold que beneficien la precisión frente a “recall”, lo que significa que se prefiere acertar mejor cuando el mercado encarece, que cuando no encarece.

8.3.2 RESULTADOS TUNNING DE HIPERPARÁMETROS

Mejores Combinaciones Tunning Hiperparámetros

Activation	Alpha	Hidden layers	HL1 size	HL2 size	HL3 size	HL4 size	Learning rate	Shuffle	Solver	Weighted accuracy	(±) Weighted acc	Test Acc
tanh	0.1	3	20	10	5		adaptive	TRUE	adam	0.755	0.061	0.76
tanh	0.1	3	22	8	4		adaptive	TRUE	adam	0.711	0.061	0.72
tanh	0.075	4	10	10	10	5	adaptive	TRUE	adam	0.693	0.028	0.69
tanh	0.075	3	20	10	5		adaptive	TRUE	adam	0.694	0.059	0.69
tanh	0.05	3	20	10	5		adaptive	TRUE	adam	0.694	0.059	0.68

tanh 0.05 4 10 10 10 5 adaptive TRUE adam 0.674 0.074 0.68

Tabla 18 - Mejores Combinaciones Tuning Hiperparámetros

8.3.3 RESULTADOS DEL MODELO

Se realiza un estudio del modelo del 1 al 8 de mayo de 2021 que cumple los siguientes criterios:

- Se estudian mercados intradiario uno, dos y tres.
- El horizonte temporal escogido para el cálculo de los parámetros para los modelos fundamentales es: 11 de noviembre de 2020 al 1 de marzo de 2021. Nótese cómo el último día del horizonte temporal no sobrepasa los 90 días previos a la fecha que se quiere predecir.
- De los datos disponibles del 1 al 8 de mayo, se utiliza un 64 % de los datos para entrenar la red neuronal, un 20 % para testarla, y el 16 % restante para hacer una validación final.

8.3.3.1 Set de Test

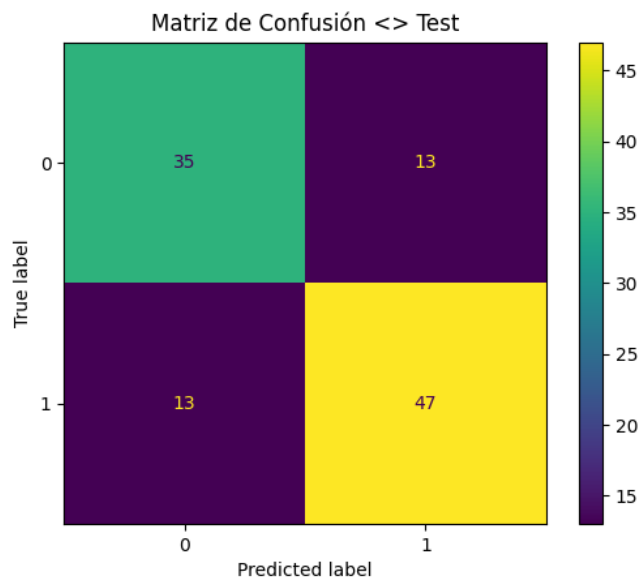


Ilustración 35 - Matriz de confusión Test

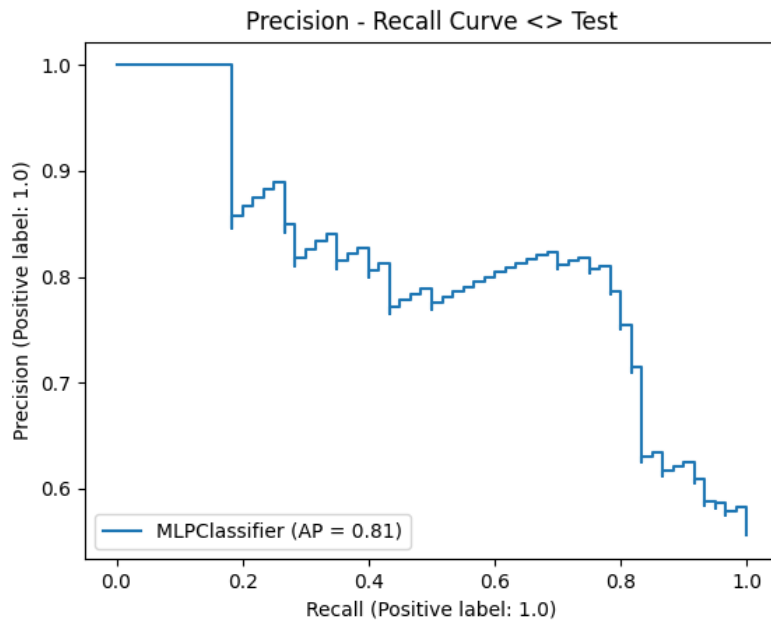


Ilustración 36 - Curva Precision recall set Test

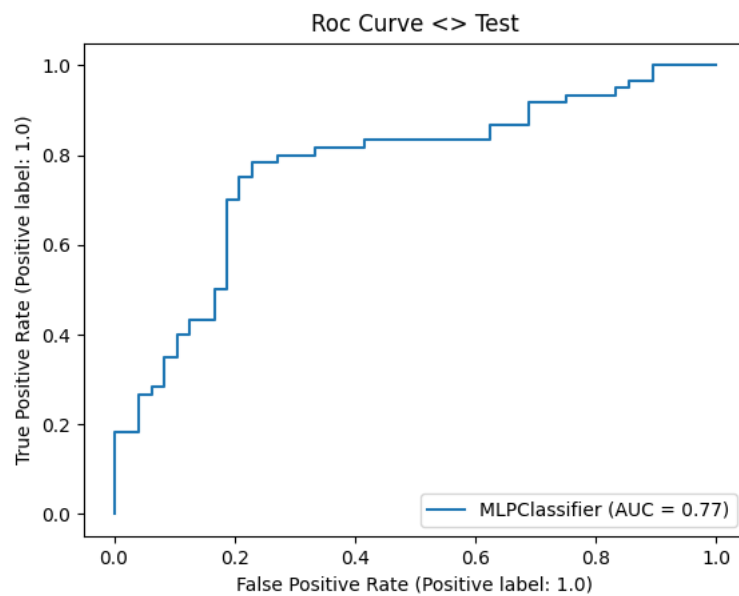


Ilustración 37 - Roc Curve Test

Results on Test set

	Precision	Recall	f1-score
0	0.73	0.73	0.73
1	0.78	0.78	0.78
Accuracy			0.76
Macro AVG	0.76	0.76	0.76
Weighted AVG	0.76	0.76	0.76

Tabla 19 - Resultados Performance Test

La precisión en test es aceptable, obteniendo alrededor de un 76 %. Sin embargo, el porcentaje de verdaderos positivos, es decir, porcentaje de aciertos de que el mercado encarece se encuentra en torno al 78 %.

Por otro lado, como se ha mencionado anteriormente, se prefiere mejorar el porcentaje de aciertos cuando el mercado encarece, en la figura 36, se puede observar como sacrificando aciertos cuando el mercado abarata, podríamos llegar a tener casi un 90 % de aciertos a que el mercado encarece (con un 25 % a que abarata).

8.3.3.2 Set de validación

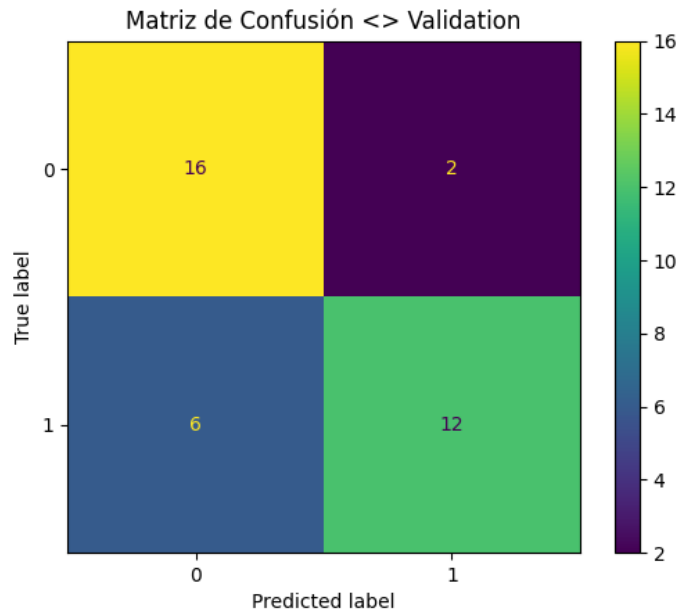


Ilustración 38 - Matriz de confusión Validación

Results on Val set			
	Precision	Recall	f1-score
0	0.73	0.89	0.8
1	0.86	0.67	0.75
Accuracy			0.78
Macro AVG	0.79	0.78	0.77
Weighted AVG	0.79	0.78	0.77

Ilustración 39 - Resultados performance validación

En validación los resultados son más satisfactorios, acertando un 86 % de las veces que el modelo predice que el mercado encarece. Se obtiene una precisión media de en torno al 78 %, más que aceptable considerando la complejidad de estas predicciones.

8.3.4 SIMULACIÓN SITUACIÓN REAL

En el apartado anterior (8.3.3) queda demostrada la capacidad de predicción del modelo. Sin embargo, para poder confirmar si es efectivo, se debe poder hacer una predicción del día siguiente utilizando los movimientos del día actual. Además, en vez de hacer la predicción directa, se considera más útil informar de las probabilidades de encarecimiento o no.

Con la finalidad de que la situación sea lo más real posible, se realizará el entrenamiento del 1 al 8 de mayo de 2021. Sin embargo, las predicciones se harán sobre el 9 de mayo, fecha que todavía el modelo no ha tocado, con la finalidad de conseguir unos resultados realistas.

Los resultados obtenidos son los siguientes:

Intradiario 1, 9 Mayo 2021								
Hour	Prob de 0	Prob de 1	Pred T = 0.5	Result	True Positive	True Negative	False Positive	False Negative
00:01	0.020254758	0.979745242	1	1	1	0	0	0
01:02	0.033551559	0.966448441	1	1	1	0	0	0
02:03	0.00032112	0.99967888	1	1	1	0	0	0
03:04	0.01546734	0.98453266	1	0	0	0	1	0
04:05	0.000100655	0.999899345	1	0	0	0	1	0
05:06	0.999999703	2.97454E-07	0	0	0	1	0	0
06:07	0.876456895	0.123543105	0	0	0	1	0	0
07:08	5.95459E-06	0.999994045	1	1	1	0	0	0
08:09	6.89051E-05	0.999931095	1	0	0	0	1	0
09:10	0.950390079	0.049609921	0	0	0	1	0	0
10:11	0.999915414	8.45864E-05	0	0	0	1	0	0
11:12	0.999995764	4.23553E-06	0	0	0	1	0	0
12:13	0.00807669	0.99192331	1	1	1	0	0	0
13:14	0.999997636	2.36355E-06	0	1	0	0	0	1
14:15	4.10388E-06	0.999995896	1	1	1	0	0	0
15:16	4.48832E-06	0.999995512	1	1	1	0	0	0
16:17	0.000137579	0.999862421	1	1	1	0	0	0
17:18	0.979836129	0.020163871	0	0	0	1	0	0
18:19	0.998789733	0.001210267	0	0	0	1	0	0
19:20	1.71092E-05	0.999982891	1	1	1	0	0	0
20:21	0.983340301	0.016659699	0	0	0	1	0	0
21:22	0.052701119	0.947298881	1	0	0	0	1	0

	ICAI	ICADE	CIHS						
22:23	0.999998508	1.49227E-06	0	1	0	0	0	0	1
23:24	0.983366955	0.016633045	0	0	0	1	0	0	0

9	9	4	2
---	---	---	---

Tabla 20 - Tabla de resultados simulación Intradía 1

En el intradía 1 se han obtenido los siguientes porcentajes de acierto para un threshold de 0.5:

Porcentaje de acierto a que el mercado encarece = 69.23 %

Porcentaje de acierto a que el mercado abarata = 81.82 %

Porcentaje de acierto medio ponderado = 75 %

Intradía 2, 9 Mayo 2021								
Hour	Prob de 0	Prob de 1	Pred T = 0.5	Result	True Positive	True Negative	False Positive	False Negative
00:01	0.99978824	0.00021176	0	0	0	1	0	0
01:02	6.7082E-07	0.99999933	1	1	1	0	0	0
02:03	0.99997125	2.8751E-05	0	0	0	1	0	0
03:04	0.99999648	3.5209E-06	0	0	0	1	0	0
04:05	0.99999755	2.4492E-06	0	1	0	0	0	1
05:06	1.2667E-06	0.99999873	1	1	1	0	0	0
06:07	0.27228057	0.72771943	1	1	1	0	0	0
07:08	0.99997095	2.9049E-05	0	1	0	0	0	1
08:09	0.00077242	0.99922758	1	1	1	0	0	0
09:10	0.02381544	0.97618456	1	1	1	0	0	0
10:11	0.99877886	0.00122114	0	1	0	0	0	1
11:12	0.86392693	0.13607307	0	0	0	1	0	0
12:13	0.00202068	0.99797932	1	1	1	0	0	0
13:14	3.2898E-06	0.99999671	1	1	1	0	0	0
14:15	0.999998	1.9968E-06	0	0	0	1	0	0
15:16	0.9997224	0.0002776	0	0	0	1	0	0
16:17	0.99654248	0.00345752	0	0	0	1	0	0

	ICAI	ICADE	CIHS						
17:18	0.14698937	0.85301063	1	0	0	0	1	0	
18:19	2.7419E-05	0.99997258	1	1	1	0	0	0	
19:20	2.0234E-05	0.99997977	1	1	1	0	0	0	
20:21	9.9945E-06	0.99999001	1	1	1	0	0	0	
21:22	0.99999964	3.6027E-07	0	0	0	1	0	0	
22:23	0.99999859	1.4118E-06	0	0	0	1	0	0	
23:24	0.99125039	0.00874961	0	0	0	1	0	0	

10	10	1	3
----	----	---	---

Tabla 21 - Tabla de resultados simulación Intradiviario 2

En el intradiario 2 obtienen los siguientes porcentajes de acierto:

Porcentaje de acierto a que el mercado encarece = 90.91 %

Porcentaje de acierto a que el mercado abarata = 76.92 %

Porcentaje de acierto medio ponderado = 83.33 %

Intradiviario 3, 9 Mayo 2021								
Hour	Prob de 0	Prob de 1	Pred T = 0.5	Result	True Positive	True Negative	False Positive	False Negative
00:01	3.5928E-06	0.99999641	1	1	1	0	0	0
01:02	4.0915E-07	0.99999959	1	1	1	0	0	0
02:03	4.5105E-08	0.99999995	1	0	0	0	1	0
03:04	1.5678E-05	0.99998432	1	1	1	0	0	0
04:05	0.96396255	0.03603745	0	1	0	0	0	1
05:06	0.00214704	0.99785296	1	0	0	0	1	0
06:07	0.9989386	0.0010614	0	0	0	1	0	0
07:08	0.0172666	0.9827334	1	1	1	0	0	0
08:09	0.99999923	7.7049E-07	0	0	0	1	0	0
09:10	4.0573E-06	0.99999594	1	0	0	0	1	0
10:11	0.00017108	0.99982892	1	1	1	0	0	0
11:12	0.02833883	0.97166117	1	0	0	0	1	0
12:13	4.2577E-06	0.99999574	1	1	1	0	0	0
13:14	0.99999442	5.575E-06	0	0	0	1	0	0
14:15	0.1369175	0.8630825	1	0	0	0	1	0
15:16	5.3097E-05	0.9999469	1	0	0	0	1	0

	ICAI	ICADE	CIH5						
16:17	7.5421E-06	0.99999246	1	1	1	0	0	0	0
17:18	0.62943576	0.37056424	0	0	0	1	0	0	0
18:19	1.3975E-06	0.9999986	1	1	1	0	0	0	0
19:20	6.556E-06	0.99999344	1	1	1	0	0	0	0
20:21	0.07675838	0.92324162	1	0	0	0	1	0	0
21:22	0.99971274	0.00028726	0	0	0	1	0	0	0
22:23	3.5033E-05	0.99996497	1	1	1	0	0	0	0
23:24	0.0005805	0.9994195	1	1	1	0	0	0	0

11	5	7	1
----	---	---	---

Tabla 22 - Tabla de resultados simulación Intradía 3

En el intradía 3 obtienen los siguientes porcentajes de acierto:

Porcentaje de acierto a que el mercado encarece = 61.11 %

Porcentaje de acierto a que el mercado abarata = 83.33 %

Porcentaje de acierto medio ponderado = 66.67 %

Capítulo 9. CONCLUSIONES Y TRABAJOS FUTUROS

9.1 CONCLUSIONES

- Se ha podido analizar la participación y volumen de compraventa de los agentes en los mercados intradiarios. Así como la influencia que tienen en las variaciones de precio y sus tendencias.
- Gracias al cálculo de los parámetros para los modelos fundamentales, se ha podido observar que las unidades de programación de eólica terrestre son las que más influencia tienen sobre las fluctuaciones de precio en los mercados intradiarios.
- Se ha podido verificar que la predicción de los mercados intradiarios es una tarea más compleja que la del mercado diario, al tener más peso el comportamiento de los agentes.
- Aunque ningún modelo fundamental pueda predecir el comportamiento de los mercados intradiarios de forma frecuente y acertada, se han encontrado relaciones no lineales entre los valores calculados y el desenlace de los mercados, por ello se decide realizar una red neuronal que utilice estos valores para hacer las predicciones.
- Se puede considerar que la red neuronal es capaz de predecir el desenlace de los mercados intradiarios, aunque con precisión variante en función del día y mercado, fluctuando entre el 65 % y el 84 %.

9.2 TRABAJOS FUTUROS

- Análisis del modelo basado en redes neuronales en un espacio temporal de forma dinámica. Esto significa que en un espacio temporal determinando, ir desplazando las fechas de entrenamiento para hacer predicciones a corto plazo día a día, obteniendo el rendimiento del modelo durante un tiempo, estudiando además la capacidad de adaptación de la red.

- Implementar el histórico de gradientes en el precio del día anterior como entrada de la red neuronal.
- Automatización del modelo basado en redes neuronales, utilizando una plataforma como Google Cloud para ejecutar los programas continuamente, descargando nuevos datos disponibles con la API de ESIOS y enviando notificaciones rutinarias a los agentes de mercado escogidos con las predicciones.

Capítulo 10. BIBLIOGRAFÍA

- [7] Bendezu, K.. “DIAGRAMA UML Y ARQUITECTURA DEL SISTEMA“. Sistemas Distribuidos 2013. Febrero, 2013. <http://comparape.blogspot.com.es/2013/02/diagrama-uml-y-arquitectura-del-sistema.html>.
- [8] Bosch Alonso, Vicente, V. B. A. (2021). *DESARROLLO DE UN MODELO BASADO EN REDES NEURONALES ADAPTATIVAS PARA OPTIMIZAR EL FUNCIONAMIENTO DE LA NUEVA GENERACIÓN DE MCIA*.
<https://riunet.upv.es/bitstream/handle/10251/170890/Bosch%20-%20Desarrollo%20de%20un%20modelo%20basado%20en%20redes%20neuronaes%20adap%20tativas%20para%20optimizar%20el%20funciona...pdf?sequence=1>
- [9] ¿Cómo funciona el mercado eléctrico español? (2021, 9 junio). CYE. <https://www.cye-energia.com/como-funciona-el-mercado-electrico-espanol/#:~:text=El%20Mercado%20Ib%C3%A9rico%20Espa%C3%B1ol%20es,d%C3%ADa%20a%20trav%C3%A9s%20del%20OMIE>
- [10] Escuela Técnica Superior Ingenieros Industriales Valencia, & Querol Herrá, J. Q. H. (2019). *DESARROLLO DE UN MODELO DE PREDICCIÓN DEL PRECIO DE LA ENERGÍA ELÉCTRICA PARA EL MERCADO A PLAZO MEDIANTE REDES NEURONALES*.
https://riunet.upv.es/bitstream/handle/10251/121891/29198105L_TFM_15524741021117331948730569227088.pdf?sequence=3
- [11] *Euphemia, el algoritmo que establece el precio de la luz (electricidad)*. (2021, 22 agosto). Naukas. <https://naukas.com/2021/08/23/euphemia-el-algoritmo-que-establece-el-precio-de-la-luz-electricidad/>
- [12] Monteiro, C., Ramirez-Rosado, I., Fernandez-Jimenez, L., & Conde, P. (2016). Short-Term Price Forecasting Models Based on Artificial Neural Networks for Intraday Sessions in the Iberian Electricity Market. *Energies*, 9(9), 721. <https://doi.org/10.3390/en9090721>
- [13] Narajewski, M., & Ziel, F. (2020). Econometric modelling and forecasting of intraday electricity prices. *Journal of Commodity Markets*, 19, 100107.
<https://doi.org/10.1016/j.jcomm.2019.100107>
- [14] Repositorio Comillas. (2014, abril). *ANÁLISIS DEL MERCADO ELÉCTRICO ESPAÑOL*. <https://repositorio.comillas.edu/rest/bitstreams/799/retrieve>

- [15] Repositorio Comillas, & García Mate, P. G. M. (2015, junio). *MODELO DE PREVISIÓN DEL PRECIO DEL MERCADO DIARIO DE LA ELECTRICIDAD EN EL CORTO PLAZO*. <https://repositorio.comillas.edu/rest/bitstreams/19054/retrieve>