|  |  |
| --- | --- |
| **Predicción y cobertura del precio de los servicios de ajuste**  Manual de uso |  |

## Introducción

El modelo desarrollado para la predicción y cobertura de los Servicios de Ajuste está desarrollado en Python. Para ello, se ha desarrollado un entorno virtual específico que contiene todos los paquetes necesarios para utilizar el modelo.

## Estructura

El fichero *Final Deliverable* contiene las siguientes carpetas:

* **Coberturas:** Este es el entorno virtual de Python con todos los paquetes instalados para correr el modelo.
* **Data:** Esta carpeta contiene los archivos de datos que sirven como input al modelo.
* **Src:** Esta carpeta contiene funciones específicas para el cálculo de la predicción y cobertura. Todas las funciones están comentadas y su funcionamiento explicado en el código.
  + Archivo *config*: Este archivo contiene los parámetros que se han tomado en el desarrollo del modelo: la ruta de los archivos de input, los conceptos utilizados para definir los Servicios de Ajuste y los índices utilizados como datos spot y forward para cada mes.
* **test\_coberturas.py:** Es un script de Python en el que se muestra el uso correcto y el funcionamiento del modelo.
* **Instructions.**txt: Contiene una guía de los comandos a utilizar para inicial el modelo
* **Packages.**txt: Contiene una lista de los paquetes instalados en el entorno virtual

## Instrucciones para correr el modelo

1. Entra en la carpeta de Final deliverable
2. En la terminal (Powershell), corre este comando para entrar en el entorno virtual:

\coberturas\Scripts\Activate.ps1

(En caso de entrar desde el cmd de Windows, corre el siguiente comando, utilizando la ruta correcta: \coberturas\Scripts\Activate.bat )

1. Para probar el método de evaluación de coberturas, corre este comando en la carpeta de Final Deliverable:

python -m test\_coberturas

1. El script test\_coberturas se debe haber ejecutado correctamente

## Datos input

Los datos de input se ubican en la carpeta de *Data.* Estos archivos deben ser expandidos o sustituidos por archivos con el mismo formato, p

ara mantener el correcto funcionamiento del código.

Son tres archivos Excel:

* *OMEL*: Contiene el histórico de datos diarios sobre el Pool. Está compuesto a su vez por dos hojas:
  + OMEL: Contiene el precio de energía spot de forma horaria, con una última columna con las medias diarias
  + OMIP: Contiene los futuros del Pool con granularidad diaria
* *Liquicomun*: Este archivo contiene los conceptos en los que se disgrega el coste de los Servicios de Ajuste, ubicados en las distintas columnas.
* *Reikon*\_*commodities*: Este archivo contiene el histórico del resto de commodities. Cada commodity está en una hoja diferente, pero todas comparten el mismo formato
* *PBF\_2014\_2023*: Este archivo contiene información procedente de Nexus, es utilizada para extraer los datos del Hueco térmico.
  + El Hueco térmico es una variable que se puede utilizar en los modelos que no realizan coberturas, se ha comprobado que es un buen predictor de los SSCC. Para realizar regresiones out-of-sample y realizar coberturas para los meses futuros, se utiliza como ‘futuro’ la media de los últimos 24 valores conocidos del Hueco Térmico

## Casos de uso

El script test\_coberturas contiene diversos casos de uso para poner en funcionamiento al modelo, se compone de diferentes partes

* **Data ingestion and preprocessing**

En esta parte se leen los archivos de input y se prepara el dataframe para preprocesarlo. Además, se añaden las variables sintéticas 601 y 603 de todas las variables spot.

Tras esta parte, se obtiene un dataframe con todas las variables spot, forward y la variable objetivo

* **User inputs**

Aquí se encuentran todos los parámetros configurables por el usuario:

* *X\_cols*: Variables a incluir en el modelo, estas variables tienen el nombre de los ricks de Reuters para cada índice
* *Y\_var*: Es el nombre de la variable objetivo a utilizar
* *Regression\_type*: El tipo de regresión a utilizar, linear o huber (Robusta)
* *Significativas*: Es un parámetro boolean, en True, el modelo realizará la técnica de Step Wise regression para obtener sólo las variables significativas
* *Prima*: Es el valor que añadir al precio estimado de los Servicios de Ajuste. En el ejemplo se ha utilizado la distribución de Cash Flows históricos con los Forwards a 1 mes con un 80% de fiabilidad
* *Volumen*: Lista con la cantidad de Servicios de Ajuste a cubrir por cada mes. Es una lista con el volumen
* *N\_vars*: Al realizar la técnica de Step Wise, es el máximo número de variables a incluir en el modelo
* *Window*: Es el tamaño de la ventana de entrenamiento del modelo
* *Step\_ahead*: Es el número de meses sobre el cual se va a realizar una cobertura. Un valor de 1 significa que la cobertura se realiza a 1 mes vista y un valor de 12 que se realiza a 12 meses vista
* *Initial\_month:* Es el extremo inicial del intervalo de fechas, es decir, el primer mes sobre el que se van a realizar coberturas
* *final\_month:* Es el extremo final del intervalo, es decir, el primer mes sobre el que se van a realizar coberturas
* **Perform hedging F1Y**

En este caso de uso, se utiliza la función *rw\_hedging\_forwards* para obtener los swaps a realizar para cada mes y para cada variable, los nombres de las variables es el Rick de Reuters para el Spot. Esta información se extrae en un Excel, con el siguiente formato:



* **Evaluate hedging F1M**

En este caso de uso, se desarrolla el modelo con la variables y parámetros indicados por el usuario. El código ejecutará la función *rw\_hedging\_forwards\_eval* para las variables definidas por el usuario y volverá a ejecutar la función para el modelo sin coberturas. El resultado de ejecutar ambas funciones será un grupo de dataframes que serán descritos a continuación:

* *d1:* Este dataframe contiene los resultados de las predicciones para el espectro de tiempo indicado, desde el *final\_month* hasta el número el número de meses indicados en *step\_ahead.*
* *Stat:* Extrae de *d1* las variables más relevantes para el análisis de los resultados:
* *d1\_pool:* Este dataframe contiene los resultados de las predicciones para el espectro de tiempo indicado, desde el *final\_month* hasta el número el número de meses indicados en *step\_ahead* para el modelo Pool*.*
* *Pool\_df:* Extrae de *d1\_pool* las variables más relevantes para el análisis de los resultados. Son las mismas variables que las definidas en *stat,* además genera un par de columnas con distinto nombre que las de *stat:*
  + *cash\_flow\_sin\_coberturas\_EUR\_MWh:* Indica el cash flow del modelo Pool, dividido entre el volumen, medido en EUR/MWh.
* *Stat\_pool:*  Extrae de *d1* las variables más relevantes para el análisis de los resultados junto con las de *pool*\_*df* de *d1* extrae:
  + *Cash\_flow\_inicial\_EUR\_MWh:* Indica el cash flow del modelo inicial (basado en la predicción), dividido entre el volumen, medido en EUR/MWh. No incluye
  + *cash\_flow\_EUR\_MWh:* Indica el cash flow del modelo con coberturas, dividido entre el volumen, medido en EUR/MWh.
  + *cash\_flow\_prima\_EUR\_MWh:* Indica el cash flow del modelo con coberturas más la prima indicada por el usuario en la variable *Prima*, dividido entre el volumen, medido en EUR/MWh
* *df\_results:* Este es el dataframe que incorpora la información consolidada que se obtuvo en *stat* y *pool\_df.* Esta tabla contiene las siguientes columnas:
  + *year:* Indica el año de análisis, según lo seleccionado por el usuario
  + *Variables:* Indica las variables seleccionadas por el usuario para construir el modelo de regresión.
  + *MEAN Cash Flow sin coberturas:* Es el promedio de la variable *cash\_flow\_sin\_coberturas\_EUR\_MWh*.
  + *MEAN Cash Flow EUR/MWh:* Es el promedio de la variable *cash\_flow \_EUR\_MWh*.
  + *MEAN Cash Flow Prima EUR/MWh:* Es el promedio de la variable *cash\_flow\_prima\_EUR\_MWh*.
  + *% Mejora: Esta métrica se define como:*

*1 – CF con coberturas^2/CF sin coberturas^2*

* + **Outputs:**

En este apartado se extrae a Excel el resultado de las coberturas, contenido en el dataframe *d1*.

Además, se compara frente al modelo inicial con la variable Pool (La base sobre la que el modelo se enfrenta) y se extrae el dataframe con los resultados de la comparación y el gráfico de Cash Flow

A graph of different colored squares

Description automatically generated

* **Evaluate hedging F1Y**

Este caso de uso es similar al anterior, en el cual, se cambia el step\_ahead para realizar coberturas a 12 meses vista. Las coberturas en este caso se realizan el 12-2021 para los siguientes 12 meses: 01-2022 hasta el 12-2021

En este caso, se agrupan los resultados según *real\_date*, para mostrar el cash Flow medio de todas las coberturas realizadas en los siguientes 12 meses.

Además, para la gráfica se extrae el cash Flow con la adición de la prima (Columna roja) y el cash Flow inicial del modelo econométrico sin coberturas para comparar

El código sigue una estructura similar al caso de uso anterior.

* + **Outputs:**

En este apartado se extrae a Excel el resultado de las coberturas, contenido en el dataframe *d1*.

Además, se compara frente al modelo inicial con la variable Pool (La base sobre la que el modelo se enfrenta) y se extrae el dataframe con los resultados de la comparación y el gráfico de Cash Flow

A graph of different colored squares

Description automatically generated