



Competitions > Reto ensena oracle espana > **Modelo predictivo reto 3**

Modelo predictivo valor EUR/USD- Reto 3

Cuándo

Desde: Miércoles, 26 de abril de 2023, 15:00:58
Hasta: Lunes, 29 de mayo de 2023, 11:59:58

Categoría

Data Science

Nivel

Junior

Siguientes pasos



Crea o únete a un equipo

2

Resuelve el reto

3

Mi puntuación

ATRÁS

SIGUIENTE

2. Resuelve el reto

Suggested stack



Acabó el

26-04-2023

F1-score drag and drop



Check your F1-score by dropping your JSON file here

Contexto:

NUWE EVA es una cadena de restaurantes innovadora y disruptiva que emplea robots para realizar todas las tareas humanas. Además de esta característica única, también buscan diferenciarse de los demás restaurantes en la forma en que asignan los precios a su carta. Han decidido vincular los precios del menú a las fluctuaciones del mercado de divisas, específicamente en el valor EUR/USD.

Si bien esta estrategia de pricing tiene el potencial de atraer a muchos comensales curiosos, también presenta un riesgo significativo para la estabilidad del negocio debido a la imprevisibilidad de los ingresos.



Dataset:

Para este reto tendréis dos archivos descargables: un **'training_set.csv'** y **'testing_set.csv'**. El primero se debe emplear para el entrenamiento del modelo, y el segundo para introducirlo como input a tu modelo y entregar las predicciones. Abajo los parámetros que hay tener en cuenta:

- **Time**: Indica la fecha y hora en la que se registró el precio del par de divisas EUR/USD. Esto es fundamental para analizar las fluctuaciones y tendencias del mercado en diferentes períodos de tiempo.
- **Open**: Representa el precio de apertura del par EUR/USD en el período de tiempo especificado. Es el primer precio registrado al comienzo de dicho intervalo, como una hora, un día o una semana.
- **High**: Muestra el precio más alto alcanzado por el par EUR/USD durante el período de tiempo analizado. Este

valor es útil para identificar los niveles de resistencia en el análisis técnico.

- **Low**: Refleja el precio más bajo alcanzado por el par EUR/USD en el período de tiempo estudiado. Este valor es relevante para determinar los niveles de soporte en el análisis técnico.
- **Close**: Indica el precio de cierre del par EUR/USD en el período de tiempo especificado. Es el último precio registrado antes de que finalice dicho intervalo, proporcionando una idea de cómo terminó la sesión de negociación.
- **Volume**: Muestra la cantidad de unidades de EUR/USD negociadas durante el período. El volumen puede ayudar a identificar la fuerza de una tendencia o confirmar la validez de un movimiento de precios.
- **Label**: Es la clase que les permitirá predecir ciertas dinámicas del mercado, permitiéndoles así estimar los márgenes de beneficio que tendrán en sus menús con antelación. Esta será la variable que se tendrá que predecir en el dataset de test.

Hay dos archivos descargables:

- **training_set.csv**: [Descargar training_set.csv](#).
- **testing_set.csv**: Se trata de una tabla que relaciona el EUR/USD del conjunto de datos de prueba con los atributos descritos anteriormente, excepto la etiqueta.
[Descargar testing_set.csv](#)



Objetivo:

Crea un modelo predictivo de clasificación para poder entender las dinámicas de movimiento del EUR/USD, únicamente en base a valores del pasado y presente y por último integra el uso de APEX en el mismo.

1. Realiza un análisis exploratorio de los datos empleando APEX.

2. Crea un modelo predictivo que maximice la puntuación F1-score (macro). Entrenando tu modelo con el dataset de training (trainig_test.csv), y efectuando las predicciones sobre el dataset de testing (testing_set.csv).
3. Crea una presentación (máx 9 slides) con la siguiente estructura:
 - Portada.
 - Análisis exploratorio de los datos:
 - Resumen y visualizaciones clave.
 - Preprocesamiento de los datos:
 - Pasos principales y criterios utilizados.
 - Modelos y técnicas de predicción:
 - Modelos seleccionados y justificación.
 - Entrenamiento, validación y resultados:
 - Proceso y métricas principales.
 - Conclusiones.

Se pueden usar fuentes de datos externas para la resolución del reto, como por ejemplo datos históricos de tweets, noticias, etc.

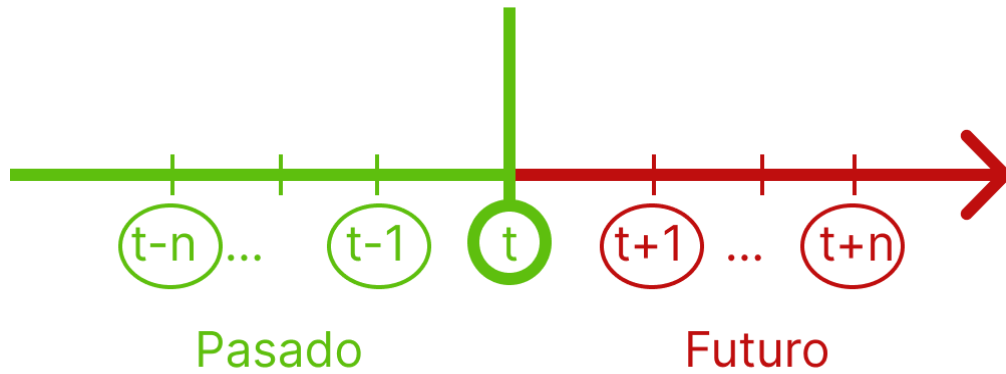
IMPORTANTE:

Hay que tener cuidado de no incurrir en el look-ahead bias. Esto será penalizado. El modelo solo puede utilizar valores actuales (t) o pasados (t-1,...t-n) como elementos predictivos. Si el modelo emplea datos futuros, es decir (t+n), el modelo estaría accediendo a información futura, por lo que la predicción no será válida.

EJEMPLO:

Para predecir el label del día 18 de Mayo de 2018, el modelo tan solo podrá usar las variables de predicción referentes al 18 de Mayo de 2018 y de los días previos (17,16,15... de Mayo de 2018).

Día en el que se va a predecir el Label



(Donde la variable t es tiempo.)

Os dejamos un vídeo con trucos para el desarrollo del tercer reto, de la mano de **Joan Itarte, Technical Manager - Emerging Technology en Oracle España**: [VER VÍDEO \(ESP\)](#).

APEX:

Pasos para solicitar un workspace en APEX: Los equipos participantes del reto deberán cursar los siguientes pasos para acceder a la prueba gratuita de APEX y solicitar un workspace:

- **Acceder URL:** <https://apex.oracle.com/es/>
- **Seleccionar opción** “Comenzar prueba gratuita hoy mismo”.
- **Solicitar un espacio de trabajo gratuito** en apex.oracle.com.
- **Identificar** nombre, apellido, correo electrónico, nombre del workspace
- **Responder cuestionario.**
- **Submit Request**
- **Acceder correo electrónico identificado** para realizar el “cambio de contraseña” y acceder al workspace generado.

Aquí tenéis un **video explicativo** de los pasos a seguir para solicitar workspace en APEX: [VER VÍDEO \(ENG\)](#). SQL para crear la tabla:

```
CREATE TABLE RELACION_EUR_USD
( "ID" NUMBER NOT NULL ENABLE,
  "TIME" TIMESTAMP (6) NOT NULL ENABLE,
  "OPEN" NUMBER,
  "HIGH" NUMBER,
  "LOW" NUMBER,
```

```
"CLOSE" NUMBER,  
"VOLUME" NUMBER,  
"LABEL" NUMBER,  
CONSTRAINT "RELACION_EUR_USD_PK" PRIMARY  
KEY ("ID")  
USING INDEX ENABLE  
);
```



Formato de entrega:

Debe de haber al menos los siguientes 3 archivos:

1. Archivo predictions.json

Las predicciones del dataset 'testing_set.csv' deben estar en un archivo JSON llamado predictions.json, un ejemplo se puede encontrar en el siguiente [link](#).

En este fichero de predicciones, en formato json, cada fila corresponderá al valor predicho del test_idx , es decir, si el primer valor es un 2 significa que este 2 corresponde al primer fichero del conjunto de datos de prueba.

Es **IMPORTANTE** llamar a la columna target tal y como se especifica en el formato. Recuerda que puedes usar la función to_json de pandas para convertir tu dataframe a json, la longitud de las predicciones tiene que ser la misma que en testing_set.csv.

2. Archivo 2: presentation.pdf: Este archivo debe de contener la presentación.

3. Archivo 3: main.py or main.ipynb: Este archivo debe contener el main de tu código fuente.



Evaluación:

En la evaluación se tendrá en cuenta lo siguiente:

- **1000/1500 (presentación)**: La presentación será evaluado por un jurado compuesto por el equipo técnico de Oracle. Los campos en los que se valoraran en esta presentación son:

1. 250 pts: Claridad en la exposición del EDA (Análisis exploratorio de los datos) y empleo de APEX.
 2. 250 pts: Coherencia y organización del proceso de resolución.
 3. 250 pts: Calidad y justificación de las técnicas y métodos empleados.
 4. 250 pts: Interpretación de resultados y aplicabilidad de las soluciones.
- **500/1500 (F1 score)**: A partir de F1-score (macro) del modelo predictivo. Comparando las predicciones de tu modelo del dataset de testing, con las del ground truth.



Está permitido el uso de transfer learning y de datos externos como tweets, noticias, etc... **IMPORTANTE:** Hay que tener cuidado de no incurrir en el look-ahead bias. Esto será penalizado.

REFERENCIAS DE QUANTITATIVE FINANCE:

- [Beginner's Guide to Quantitative Trading I: Useful skills and where to find them](#)
- [Beginner's Guide to Quantitative Trading](#)
- [Python for Finance, Part 2: Intro to Quantitative Trading Strategies](#)
- [Quantitative Finance Books](#)

REFERENCIAS APEX:

- [Curso de "learning.path" en mylearn de oracle.](#)
- [Laboratorios prácticos de Oracle sobre el uso de APEX](#)
- [Guia introducción al uso de aplicaciones de Oracle APEX](#)
- [Libros artículos referenciados por Oracle de terceros](#)

- [Creating charts \(App builder User Guide\)](#)
- [Sample charts \(App builder User Guide\)](#)