

## INTELIGENCIA ARTIFICIAL GRADO EN INGENIERÍA DE COMPUTADORES

2021-2022

### **Práctica 1 (OBLIGATORIA):**

#### **1. Objetivo**

Esta práctica consiste en aplicar los conceptos vistos en el tema relativos a la construcción de modelos lineales en un problema de predicción dinámica en un contexto financiero. La idea es que el alumno sea capaz de implantar dicho modelo de Aprendizaje Automático en un problema financiero, siendo para ello necesario que el alumno implemente en Python dichos modelos así como las métricas asociadas y otros análisis.

#### **2. Planteamiento**

La práctica plantea el uso y análisis de un modelo de Aprendizaje Automático y permite ilustrar cómo algunos modelos de Aprendizaje Automático pueden ser empleados en problemas de modelización dinámica, contextualizando conceptos vistos. Se espera también que el alumno desarrolle una visión crítica sobre dicho modelo.

Se trata de un problema de predicción dinámica. Para el que alumno deberá emplear datos del S&P500 de Yahoo Finance

(<http://finance.yahoo.com/q/hp?s=%5EGSPC+Historical+Prices>)

empleando las cotizaciones desde el primer día posterior al 1 de enero de 2021 (por ejemplo, 4 de enero) al 31 de septiembre de 2021 o último día disponible antes de esa fecha.

Una vez obtenidas las cotizaciones, el alumno debe transformarlas en rentabilidades (*rendimientos*) tomando la diferencia (a partir del segundo día disponible) de precios y dividiéndola por el precio anterior:

$$r_t = \frac{P_t - P_{t-1}}{P_{t-1}}, t = 2, 3, \dots, n$$

A partir de dichos rendimientos deberá construir una matriz de variables dependientes ( $r_t$ ) e independientes teniendo en cuenta las rentabilidades de los cinco días anteriores además de un vector de unos ( $1, r_{t-1}, \dots, r_{t-5}$ ). Evidentemente solo es posible construir dicha matriz a partir de las SEXTA observación de rendimientos (o SEPTIMA de precios), debiendo descartar las filas no completas. Las matrices de las variables dependientes e independientes se muestran en el siguiente ejemplo con los colores rojo y naranja, respectivamente.

dia	P(t)	r(t)		r(t-1)	r(t-2)	r(t-3)	r(t-4)	r(t-5)
1	12							
2	13	0,083	1,000					
3	11	-0,154	1,000	0,083				
4	12	0,091	1,000	-0,154	0,083			
5	14	0,167	1,000	0,091	-0,154	0,083		
6	13	-0,071	1,000	0,167	0,091	-0,154	0,083	
7	11	-0,154	1,000	-0,071	0,167	0,091	-0,154	0,083
8	12	0,091	1,000	-0,154	-0,071	0,167	0,091	-0,154
9	13	0,083	1,000	0,091	-0,154	-0,071	0,167	0,091
10	15	0,154	1,000	0,083	0,091	-0,154	-0,071	0,167
....	11	-0,267	1,000	0,154	0,083	0,091	-0,154	-0,071

### 3. Tareas a realizar y preguntas

A partir de dicha matriz el alumno debe construir un modelo de regresión lineal múltiple en Python del tipo  $Y=a+b_1X_1+b_2X_2+\dots+b_5X_5$  y contestar a las preguntas mas abajo indicadas.

1. De cuantos parámetros consta el modelo? (1 punto)
  - a. Cinco
  - b. Seis

Explicar en cuatro líneas los resultados

2. Los parámetros correspondientes al primero y sexto regresor son iguales? (1 punto)
  - a. Si
  - b. No

Explicar en cuatro líneas los resultados

Una vez implementado dicho modelo el alumno deberá producir dos columnas, una con el rendimiento real del activo y otra con el rendimiento predicho y sobre ambas calcular tres métricas: el *mse*, el *mae* y el *mape*.

3. Las tres métricas, proporcionan los mismos resultados (el mismo valor)? (2 puntos)
- a. Si
  - b. No

Explicar en ocho líneas los resultados

4. Suponga que comparamos las predicciones frente a un paseo aleatorio  $r_t=0$  empleando el mape, las predicciones del modelo lineal son mejores? (2 puntos)
- a. Si
  - b. No

Explicar en ocho líneas los resultados

Para la siguiente pregunta el alumno debe repetir la secciones el calculo de los parámetros del modelo lineal pero de manera iterada, es decir, considerando una ventana DESLIZANTE de UN DIA y considerando las treinta observaciones anteriores, es decir, se emplean los últimos 30 días para predecir el siguiente, mover un día la ventana (manteniendo los 30 últimos días) y volver a predecir, etc..

5. El valor del parámetro correspondiente al tercer regresor varia? (2 puntos)
- a. Si
  - b. No

Explicar en ocho líneas los resultados

En el caso de la ventana deslizante, suponga ahora que quiere emplear el modelo para invertir, suponga que la rentabilidad total es igual a la suma de las diferencias de rentabilidades entre las predichas y reales a lo largo de los días de predicción,

6. es dicha rentabilidad positiva? (2 puntos)
- a. Si
  - b. No

Explicar en ocho líneas los resultados

## 5. Fecha

La práctica se realizar a través de un cuestionario y buzón de entrega del código en *Blackboard* que estará abierto hasta las **22.00 horas del lunes 18 de octubre.**