

INTELIGENCIA ARTIFICIAL 2025-2026

Práctica 1 (OBLIGATORIA, ENTREGABLE):

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, por ejemplo, obtenidos de la siguiente dirección:

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

empleando las cotizaciones desde el primer día posterior al 1 de enero de 2023 (por ejemplo, 4 de enero) al 1 de septiembre de 2025 o último día disponible antes de esa fecha, emplee los datos de cierre (CLOSE).

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

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

3. Tareas a realizar

- A. Realice un preprocesado de datos consistente en eliminar o corregir las observaciones inexistentes y en eliminar observaciones influyentes alejadas en mas/menos dos desviaciones típicas. Presente una tabla de resultados de la media,

varianza, sesgo y kurtosis de la serie tratada. Dibuje también un histograma de esta. (2 puntos)

Explique los resultados

- B. 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, Evidentemente solo es posible construir dicha matriz a partir de las SEXTA observación de rendimientos, debiendo descartar las filas no completas. A partir de dicha matriz el alumno debe construir un modelo de regresión lineal múltiple. 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*. Todos estos cálculos se deben realizar dentro de la muestra, es decir, se refieren a la calidad de calibrado del modelo. (4 puntos)

Suponga que comparamos las predicciones frente a un paseo aleatorio $r_t=0$ evalúe la calidad del modelo considerando las métricas anteriores y comente los resultados obtenidos

- C. Para la siguiente pregunta el alumno debe repetir el cálculo de los parámetros del modelo lineal, pero de manera iterada, es decir, considerando una ventana DESLIZANTE de UN DÍA 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.

Evalúe la calidad predictiva del modelo frente a un paseo aleatorio. (2 puntos)

- D. En el caso de la ventana deslizante, suponga ahora que quiere emplear el modelo para invertir (considere que no hay costes de transacción ni ventas en corto, es decir, solo se puede invertir en caso de una predicción de precio superior al actual)

Calcule la rentabilidad de la estrategia de inversión y evalúe los resultados (2 puntos)

4. Formato

A partir de los resultados anteriores, cada grupo deberá producir un conjunto de transparencias (obligatoriamente en formato *power point*) que sirvan de soporte para una presentación de estilo *pechakucha*.

5. Fecha

La práctica se realizará a través de un buzón de entrega del código (cuaderno jupyter) y de la presentación en formato .ppt (NO pdf u otros, ni archivos comprimidos) disponible en *Blackboard* hasta las **22.00 horas del martes 14 de octubre**. En la clase

de laboratorios del día **16 de octubre** se presentarán en clase los trabajos de cada uno de los grupos. A las presentaciones solo puede acudir el grupo correspondiente. La secuencia de presentaciones se anunciará oportunamente en Blackboard.

Todas las presentaciones entregadas serán convenientemente convertidas por el profesor a vídeos de 20 segundos de duración por transparencia, para una duración total de 6 minutos, 40 segundos. Estos videos serán usados por los grupos para realizar la defensa de su trabajo y estarán a su disposición el día de la presentación.

6. Criterios de Valoración

- 1) Seguimiento de las normas
- 2) Originalidad
- 3) Corrección de los resultados
- 4) Presentación
- 5) Profundidad de los análisis realizados
- 6) Comprensión general de los objetivos y resultados