



Universitat
de les Illes Balears

GRADO EN INGENIERÍA INFORMÁTICA

Aplicación práctica ACSI curso 2022-2023

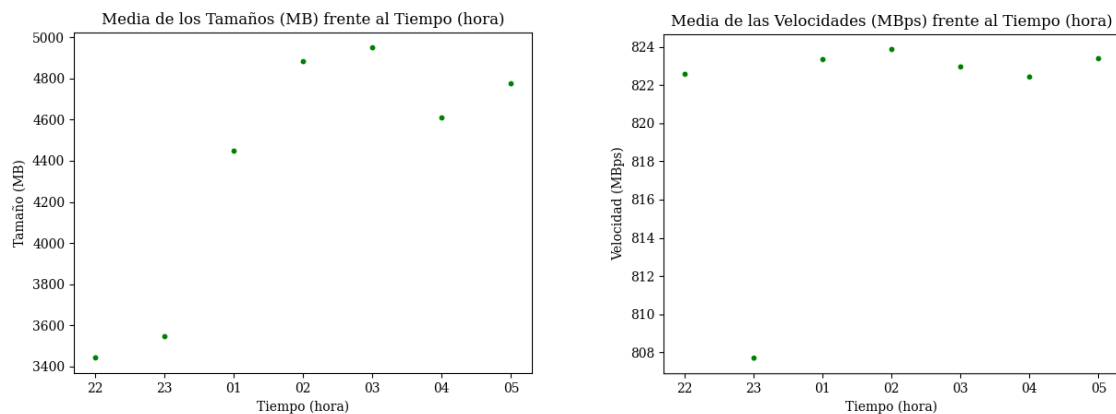
Carlos Lozano Alemañy

Práctica 7

Una empresa de almacenamiento en la nube monitoriza la actividad de sus usuarios, es decir, se guarda la hora de acceso del cliente, el tamaño del fichero al que se ha accedido y la cantidad de información transmitida por unidad de tiempo (hacer uso de los datos del fichero “data.txt”, proporcionado en la práctica 6).

El director del departamento de informática de la empresa solicita calcular la cantidad de información transmitida por la red y el tamaño del fichero accedido para las 6 a.m. (recordar que la última hora monitorizada son las 5 a.m.).

- ¿Qué patrón siguen los datos monitorizados? Proporciona una representación gráfica.



Analizando ambas gráficas nos damos cuenta que por una parte, entre las 22:00 y las 23:00, los archivos son relativamente pequeños, mientras que por otra, de 1:00 a 5:00 de la mañana, los archivos son considerablemente más grandes. Esto es común en sistemas que realizan tareas pesadas durante la noche para no perjudicar a los usuarios ya que suelen frecuentar su uso en horas diurnas.

En cuanto a la velocidad de transferencia, es alta en general, pero se observa una disminución a las 23:00, que posiblemente se deba al cambio a archivos de mayor tamaño en ese momento, como ya se había comentado antes.

- Calcula los valores solicitados para las 6 a.m. haciendo uso de la regresión lineal, medias móviles (usar los 4 últimos valores) y suavizado exponencial (peso fijo del 60%).

REGRESIÓN LINEAL

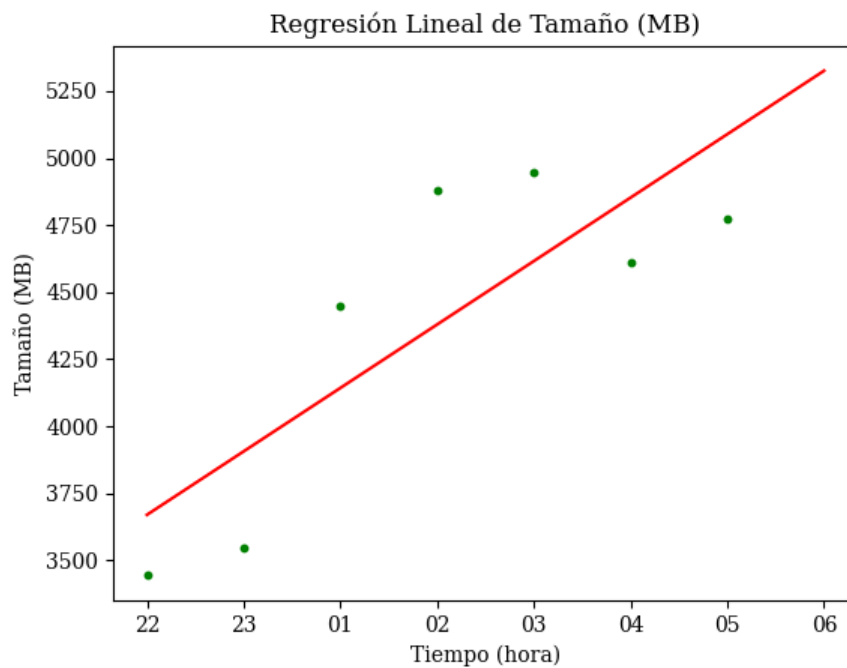
La construcción de la línea que representa la regresión lineal sigue la ecuación de la forma lineal $y = a + bx$, donde...

$$b = \frac{\sum_{i=1}^n x_i \times y_i - n \times \bar{x} \times \bar{y}}{\sum_{i=1}^n x_i^2 - n \times \bar{x}^2}$$

$$a = \bar{y} - b \times \bar{x}$$

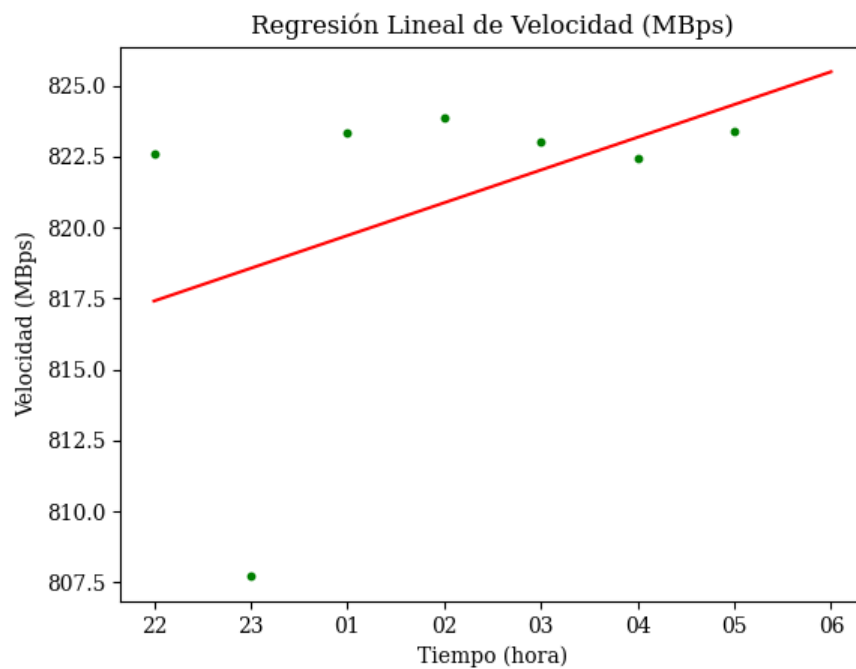
Cálculo para el tamaño:

- $y = 3669,2544 + 236,6961x$
- $\bar{x} = 4$; $\bar{y} = 4379,3427$
- La predicción para la hora 6 es:
- $y = 3669,2544 + 236,6961 \times (8) = \mathbf{5562,8232 \text{ MB}}$



Cálculo para la velocidad:

- $y = 817,4075 + 1,1535x$
- $\bar{x} = 4$; $\bar{y} = 820,8680$
- La predicción para la hora 6 es:
- $y = 817,4075 + 1,1535 \times (8) = \mathbf{826,6355 \text{ MBps}}$



MEDIAS MÓVILES

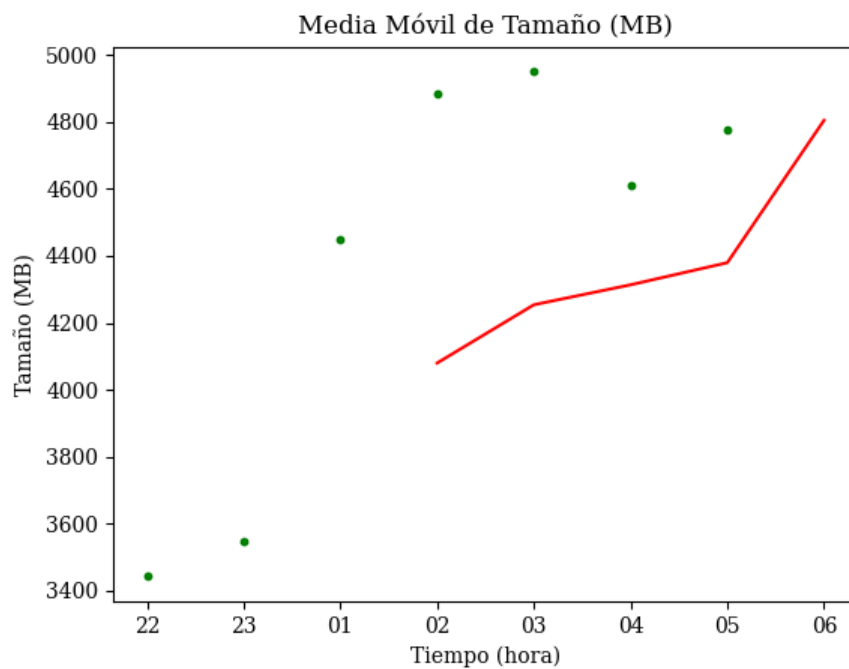
El cálculo de las medias móviles se realiza utilizando la siguiente ecuación:

$$f_{t+1} = \frac{y_t + y_{t-1} + \dots + y_{t-n+1}}{n}$$

Para obtener el resultado de la predicción (f_{t+1}) para la sexta hora se sumarán los valores de y_t de las últimas cuatro horas y se dividirá la suma obtenida en el paso anterior entre 4, ya que estamos promediando los valores.

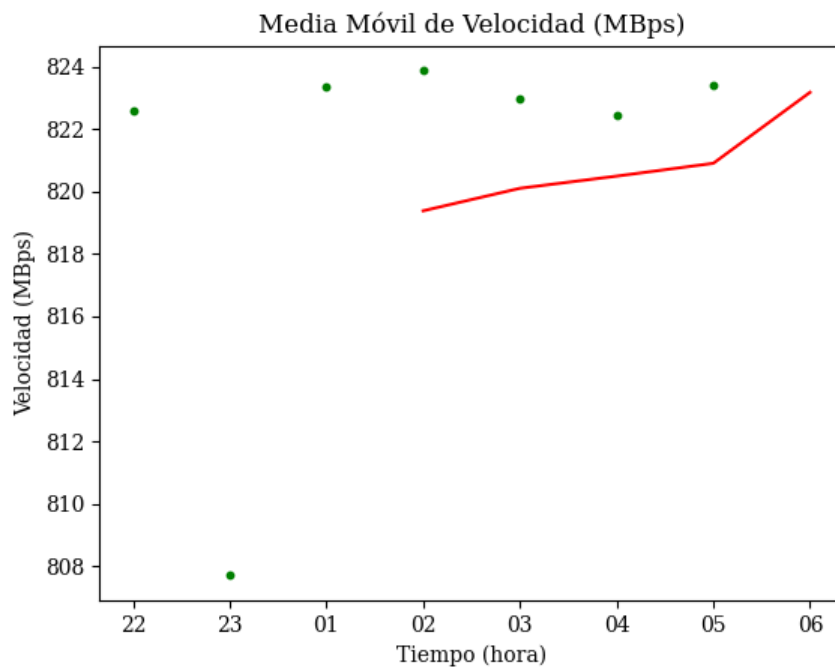
Cálculo para el tamaño:

$$f_6 = \frac{4881,8906 + 4949,225 + 4611,4016 + 4775,2293}{4} = 4804,4366MB$$



Cálculo para la velocidad:

$$f_6 = \frac{823,8627 + 822,9925 + 822,4515 + 823,3838}{4} = 823,1726 \text{ MBps}$$



SUAVIZADO EXPONENCIAL

El cálculo del suavizado exponencial se realiza utilizando la siguiente ecuación:

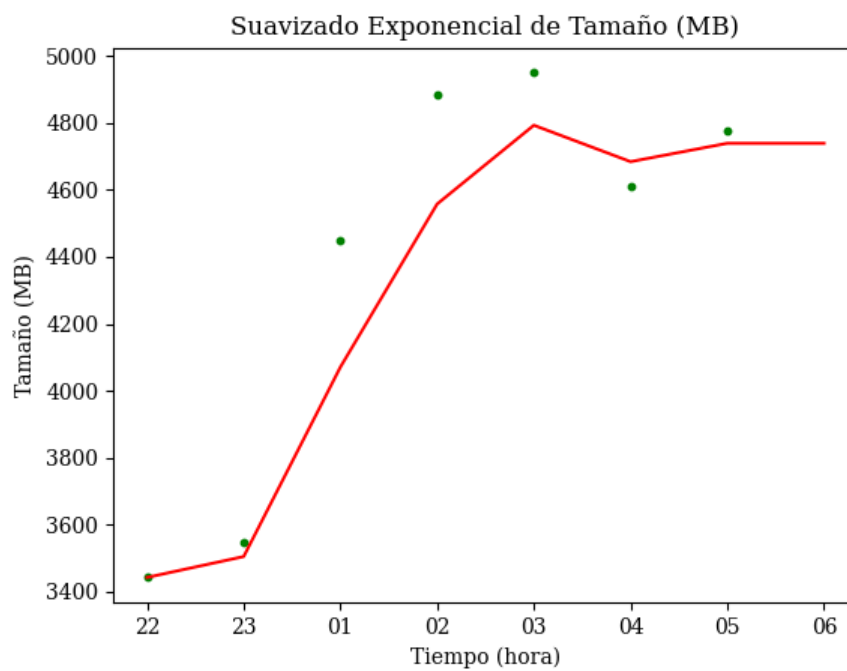
$$f_{t+1} = f_t + \alpha(y_{t+1} - f_t)$$

Para cada valor de y , se aplicará el cálculo utilizando un $\alpha = 0,6$.

Cálculo para el tamaño:

Hora	Media del Tamaño (MB)	Predicción
1	3443,8934	3443,8934
2	3546,7858	3505,2288
3	4447,9735	4070,8756
4	4881,8906	4557,4846
5	4949,225	4792,5288
6	4611,4016	4683,8524
7	4775,2293	4738,67

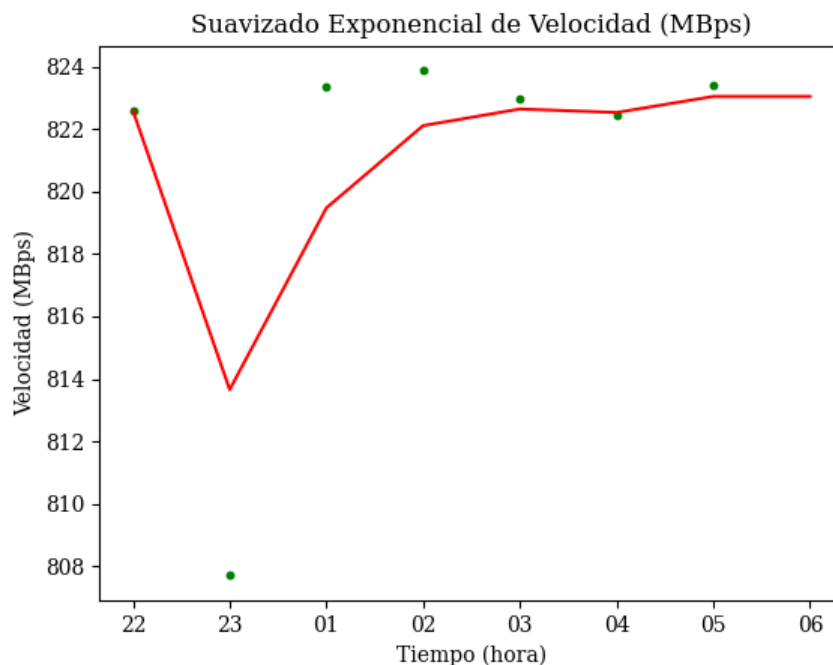
Como resultado, el valor correspondiente a la hora 6 es de 4738.67 MB.



Cálculo para la velocidad:

Hora	Media del Tamaño (MB)	Predicción
1	822,5798	822,5798
2	807,7088	813,6572
3	823,3496	819,4726
4	823,8627	822,1066
5	822,9925	822,6381
6	822,4515	822,5261
7	823,3838	823,0407

Como resultado, el valor correspondiente a la hora 6 es de 823.0407 MBps.



- ¿Qué técnica de predicción funciona mejor? ¿Por qué? ¿Cuál es la más adecuada para los datos con los que contamos?

La técnica de producción que funciona mejor es el suavizado exponencial, se ve claramente en las gráficas puesto que la línea que traza se encuentra más próxima a los puntos que las otras dos, aunque la técnica de medias móviles se le asemeja, esto ya lo vimos en teoría, pues la técnica del suavizado exponencial es similar a la técnica de medias móviles con respecto a la forma en que ambas técnicas calculan el valor de predicción, con la diferencia de que el suavizado exponencial coloca más peso en las observaciones recientes.

Por otra parte, la más adecuada vuelve a ser el suavizado exponencial, puesto que al tratarse de datos no estacionales que no muestran una tendencia sistemática esta es la técnica más adecuada.