



Universidad Autónoma de Baja California

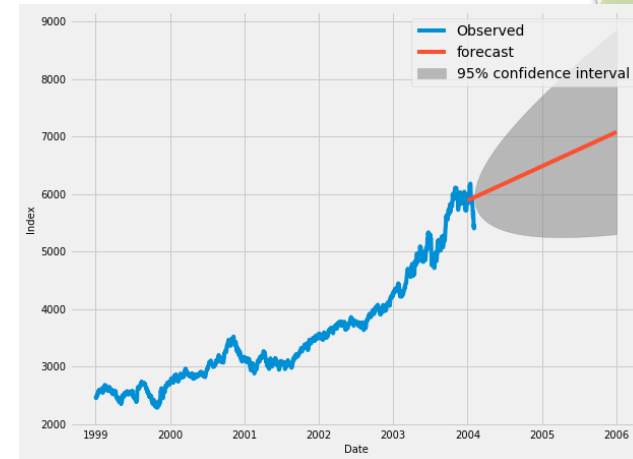
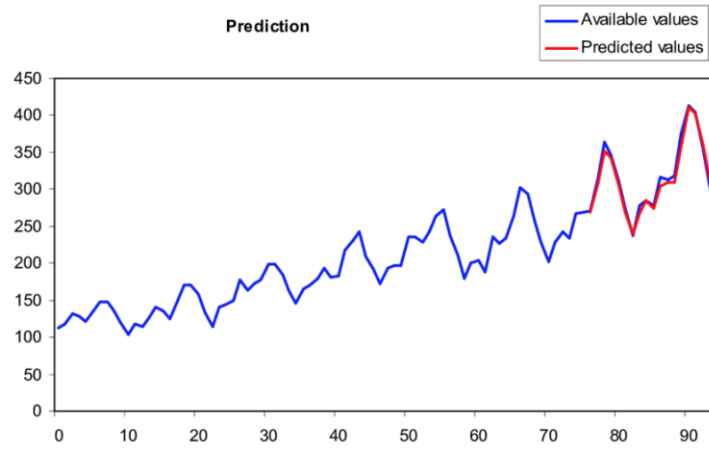
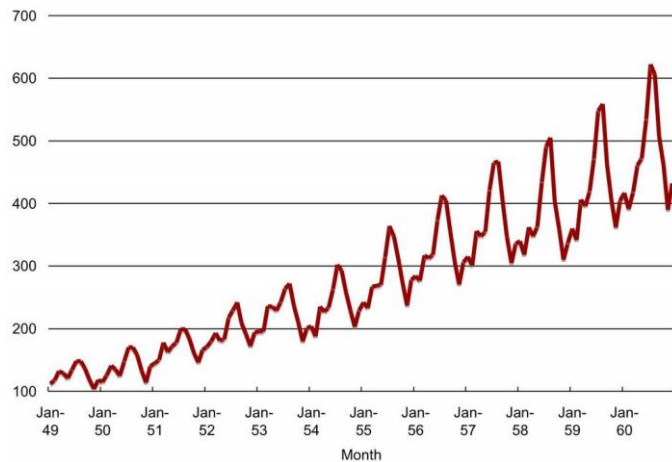
Maestría y Doctorado en Ciencias e Ingeniería
Ingeniería en Computación

6. Pronóstico

Minería de Datos

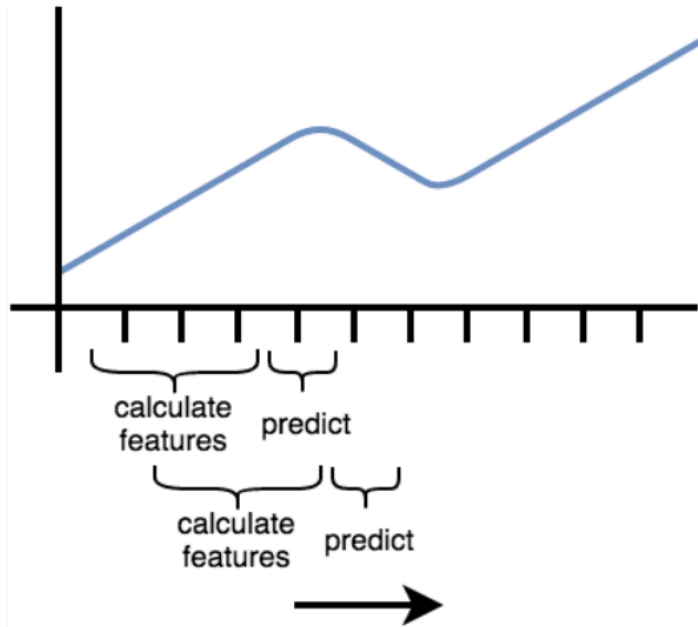
Pronóstico de series de tiempo

- El proceso de pronóstico en series de tiempo consiste en predecir qué va a suceder en un futuro más allá de los datos con el cuál se entrenó el modelo. Es decir, estos modelos extrapolan.
- De manera simplista, se suele trabajar con solamente una variable que varía a través del tiempo



Preparación de datos

- Para convertir una secuencia lineal a través del tiempo hacia un dataset que pueda entrenar un modelo de ML, se requiere utilizar el concepto de ventanas deslizables.



2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Training 1				T1								
		Training 2			T2							
			Training 3			T3						
				Training 4			T4					
					Training 5			T5				
						Training 6			T6			
							Training 7			T7		
								Training 8			T8	
									Training 9			T9

Note: "Tn" stands for "Testing n" (n=1...9).

Ventanas

<i>Tiempo</i>	T1	T2	T3	T4	T5	T6	T7	T8	...	Tn
Ventana 1	t-2	t-1	t+1							
Ventana 2	t-3	t-2	t-1	t+1	t+2					
Ventana 3	t-5		t-3		t-1	t+1				
Ventana 4	t-5		t-3		t-1		t+2			

Actividad

- ▶ Utilizar 2 datasets
 - ▶ Uno de ellos será el de Mackey-Glass (está en Bb)
 - ▶ El otro será diferente por cada alumno (ponerse de acuerdo en el foro de Bb)
- ▶ Proponer 4 ventanas diferentes por cada dataset
 - ▶ Proponer con 2 y 4 tiempos como entrada
 - ▶ Proponer con 1 y 2 tiempos de pronóstico
 - ▶ El alumno propondrá las distancias entre pasos dentro de la ventana
- ▶ Utilizar Arima, o alguna de sus variantes, para pronóstico, así como una técnica avanzada de machine learning
 - ▶ Ponerse de acuerdo con la variante de Arima y la técnica de ML a utilizar en el foro de Bb
 - ▶ Exponer cómo funciona la variante de Arima que eligieron
- ▶ Identificar el mejor resultado y las características para lograr dicho resultado