

Aprendizaje Automático

Depto. de Computación. FCEyN. UBA
2º cuatrimestre de 2016

Segundo Parcial

El parcial se aprueba con 60 puntos o más.
Poner nombre y apellido en todas las hojas.

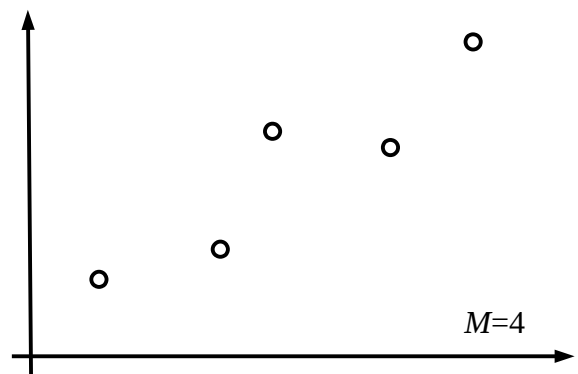
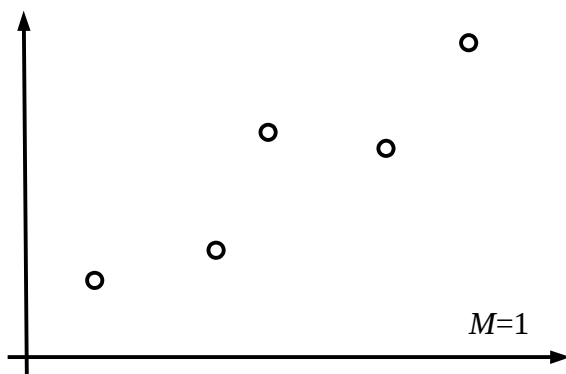
Nombre:				
Problema 1	Problema 2	Problema 3	Problema 4	Problema 5
Total:			Nota:	

Problema 1. [40 puntos] Definir brevemente exactamente 6 de los siguientes términos:

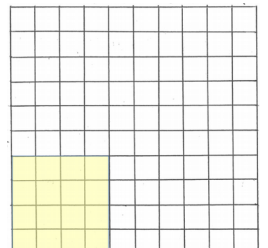
- i. R^2
- ii. Regularización en regresión lineal
- iii. Función logística
- iv. Dendrograma
- v. Algoritmo *model-free* vs. *model-based*.
- vi. *K-armed bandits*
- vii. Coeficiente Silhouette
- viii. Backpropagation

Problema 2. [15 puntos] Dado el siguiente dataset con 5 puntos bidimensionales:

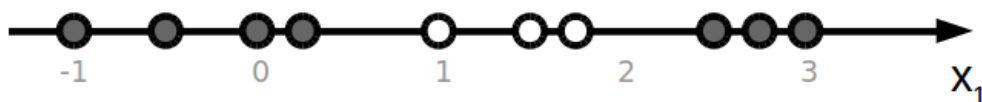
- a) Graficar cómo sería (aproximadamente) un modelo polinomial de grado M ajustado a estos datos, con $M=1$ y $M=4$. Justificar.
- b) ¿Cómo podría esperarse que sea el RSS de cada modelo sobre datos frescos (de validación)?



Problema 3. [15 puntos] Supongamos que tenemos un *grid world* como el de la figura. Corremos Q-learning durante muchos miles de iteraciones, hasta que los valores convergen (no se modifican más) y el algoritmo termina. En ese momento, observamos que todo un sector del *grid world* (en gris en la figura), resultó muy poco explorado y conserva casi intactos los valores iniciales. ¿A qué puede deberse esto, y cómo podría solucionarse?



Problema 4. [15 puntos] Se tienen el siguiente dataset de puntos unidimensionales:



¿Podrá un perceptrón simple discriminar correctamente los puntos blancos de los grises? Justificar.

Problema 5. [15 puntos] ¿Qué resultará de ejecutar *K-Means* (con $K=2$) y *BDSCAN* para los siguientes datasets? Justificar.

