

Inteligencia Artificial

Machine Learning

Martin Marchetta

martin.marchetta@ingenieria.uncuyo.edu.ar



UNCUYO
UNIVERSIDAD
NACIONAL DE CUYO



- **Qué es Machine Learning?**
 - Se dice que un sistema aprende de la experiencia E , respecto a una clase de tareas T y medida de desempeño P si su performance al realizar las tareas en T medidas por P mejora con la experiencia E (Tom Mitchell)
- **Por qué necesitamos aprendizaje?**
 - Parámetros que no podemos/sabemos calcular
 - Cálculo exacto/tradicional intratable
 - Entornos con aspectos impredecibles o desconocidos
 - Dificultad o imposibilidad de especificar completamente relaciones causa-efecto
 - Mejorar el rendimiento de otras técnicas

Introducción



UNCUYO
UNIVERSIDAD
NACIONAL DE CUYO

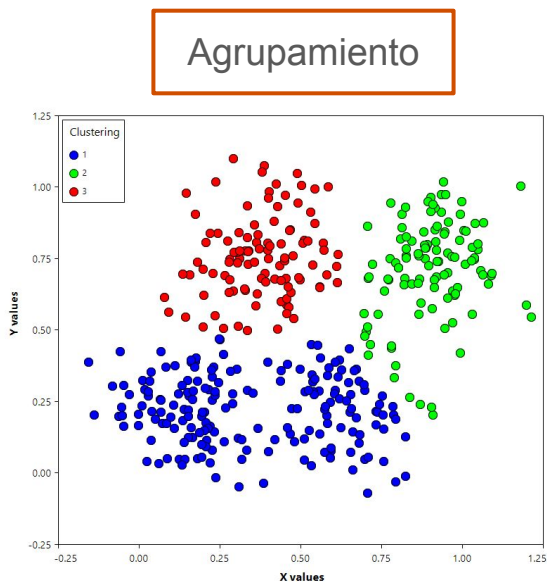


LABSIN

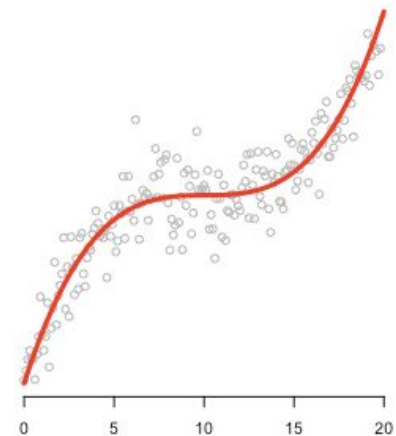
- Problemas de aprendizaje: algunas intuiciones



Clasificación



Agrupamiento



Regresión

Introducción

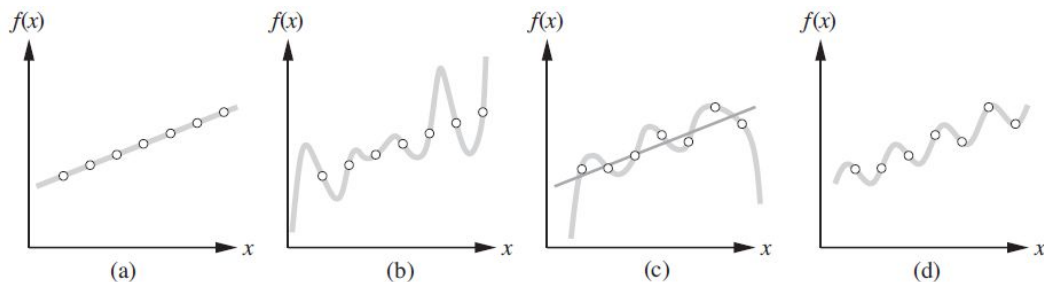


UNCUYO
UNIVERSIDAD
NACIONAL DE CUYO



LABSIN

- Los modelos a aprender se llaman hipótesis o conceptos
- Las hipótesis se aprenden a partir de los datos
- Qué “forma” tienen las hipótesis?
 - Depende del sesgo (bias): pueden ser reglas, árboles de decisión, funciones continuas, instancias, etc. → depende del problema a resolver



Introducción

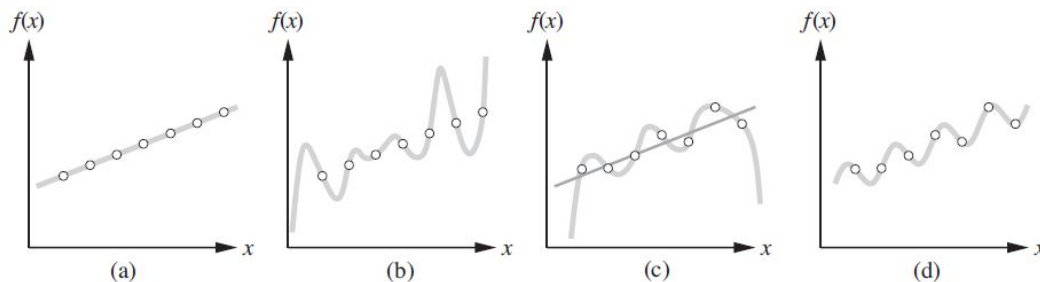


UNCUYO
UNIVERSIDAD
NACIONAL DE CUYO



LABSIN

- Aprendizaje inductivo
 - Dado un conjunto de ejemplos e , buscar una hipótesis h que aproxime los ejemplos en e .
- Las hipótesis deben ser generales
 - Deben comportarse de manera similar con los ejemplos de entrenamiento y con las instancias no vistas
 - A veces se prefiere un ajuste peor a cambio de una mejor generalización
 - Navaja de Ockham: generar la hipótesis más simple que es consistente con los datos



Introducción



- Entradas: cómo son “los datos”?

Ejemplo o
Instancia

Outlook	Temperature	Humidity	Windy	Play
Sunny	85	85	False	No
Sunny	80	90	True	No
Overcast	83	86	False	Yes
Rainy	70	96	False	Yes
Rainy	68	80	False	Yes
Rainy	65	70	True	No
Overcast	64	65	True	Yes
Sunny	72	95	False	No
Sunny	69	70	False	Yes
Rainy	75	80	False	Yes
Sunny	75	70	True	Yes
Overcast	72	90	True	Yes
Overcast	81	75	False	Yes
Rainy	71	91	True	No

Atributo

- Tipos de atributos
 - Numéricos
 - Categóricos
 - Nominales
 - Soleado, Nublado, Lluvioso
 - Ordinales
 - Frio < Caliente < Muy Caliente
 - Intervalos
 - Además de tener un orden, se los mide en unidades iguales y fijas (ej: temperaturas cada un grado, fechas, etc)
 - Ratio
 - Similar a los intervalos pero definen un “cero inherente”
 - La mayoría de los algoritmos permiten datos categóricos nominales y ordinales

- Tipos de aprendizaje: se define en buena medida por los datos con que se cuenta
 - Supervisado
 - Se cuenta con datos de entrada asociados a la salida deseada
 - No supervisado
 - Sólo se cuenta con datos de entrada, sin asociarlos a una salida deseada
 - Por refuerzo
 - Los datos suelen ser secuencias, no instancias individuales
 - En lugar de asociar una salida deseada a cada dato, se provee una recompensa (refuerzo) → buen desempeño: recompensa positiva, mal desempeño: penalización

Introducción



UNCUYO
UNIVERSIDAD
NACIONAL DE CUYO



LABSIN

- Salidas: qué aprender
 - Tablas de decisión
 - Árboles de decisión
 - Reglas de decisión
 - Reglas de asociación
 - Modelos lineales
 - Modelos no lineales
 - Representaciones basadas en instancias
 - Clusters

- Cómo evaluar el aprendizaje?
 - Aprender el modelo con el conjunto de entrenamiento
 - Evaluar el modelo con el conjunto de test
 - Procedimiento
 - Recolectar ejemplos
 - Dividir los ejemplos en conjuntos disjuntos: training y test
 - Aprender el modelo con el conjunto de training
 - Calcular el % de clasificaciones correctas del modelo aprendido sobre los casos del conjunto de test
 - Repetir con distintos tamaños de conjuntos de ejemplos y combinaciones para obtener rendimientos promedio
 - k-fold cross validation: k experimentos dejando $1/k$ ejemplos para el set de test

Introducción



- Algunas medidas
 - Clasificación
 - Error rate: cantidad de errores / cantidad de ejemplos

- Regresión

p: valor predicho

a: valor real

Table 5.8 Performance Measures for Numeric Prediction

Mean-squared error	$\frac{(p_1 - a_1)^2 + \dots + (p_n - a_n)^2}{n}$
Root mean-squared error	$\sqrt{\frac{(p_1 - a_1)^2 + \dots + (p_n - a_n)^2}{n}}$
Mean absolute error	$\frac{ p_1 - a_1 + \dots + p_n - a_n }{n}$
Relative squared error	$\frac{(p_1 - a_1)^2 + \dots + (p_n - a_n)^2}{(a_1 - \bar{a})^2 + \dots + (a_n - \bar{a})^2}$
	(in this formula and the following two, \bar{a} is the mean value over the training data)
Root relative squared error	$\sqrt{\frac{(p_1 - a_1)^2 + \dots + (p_n - a_n)^2}{(a_1 - \bar{a})^2 + \dots + (a_n - \bar{a})^2}}$
Relative absolute error	$\frac{ p_1 - a_1 + \dots + p_n - a_n }{ a_1 - \bar{a} + \dots + a_n - \bar{a} }$
Correlation coefficient	$\frac{S_{PA}}{\sqrt{S_P S_A}}, \text{ where } S_{PA} = \frac{\sum_i (p_i - \bar{p})(a_i - \bar{a})}{n - 1}, S_P = \frac{\sum_i (p_i - \bar{p})^2}{n - 1},$ $S_A = \frac{\sum_i (a_i - \bar{a})^2}{n - 1} \text{ (here, } \bar{a} \text{ is the mean value over the test data)}$

- Cost-sensitive learning:
 - Cuando una clase C es más importante que la otra
 - Precision

Cantidad de ejemplos clasificados como C **correctamente**

Cantidad total de ejemplos clasificados como C

- Recall

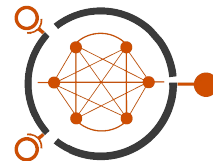
Cantidad de ejemplos clasificados como C **correctamente**

Cantidad total de ejemplos en la clase C

Preguntas? Opiniones?



UNCUYO
UNIVERSIDAD
NACIONAL DE CUYO



LABSIN