

# Introducción al Aprendizaje automático

Dra Ana Georgina Flesia

Optativa Ciencias de la Computación  
FaMAF-UNC  
Oficina 370  
[georgina.flesia@unc.edu.ar](mailto:georgina.flesia@unc.edu.ar)

2020

# Aprendizaje de conceptos

- ▶ Usualmente aprendemos, a partir de situaciones y ejemplos, conceptos o categorías del estilo de

# Aprendizaje de conceptos

- ▶ Usualmente aprendemos, a partir de situaciones y ejemplos, conceptos o categorías del estilo de
  - pájaro,

# Aprendizaje de conceptos

- ▶ Usualmente aprendemos, a partir de situaciones y ejemplos, conceptos o categorías del estilo de
  - pájaro,
  - auto,

# Aprendizaje de conceptos

- ▶ Usualmente aprendemos, a partir de situaciones y ejemplos, conceptos o categorías del estilo de
  - pájaro,
  - auto,
  - temas que tengo que estudiar para aprobar

# Aprendizaje de conceptos

- ▶ Usualmente aprendemos, a partir de situaciones y ejemplos, conceptos o categorías del estilo de
  - pájaro,
  - auto,
  - temas que tengo que estudiar para aprobar
- ▶ Cada concepto puede ser visto como descriptor de un subconjunto de objetos o eventos de un conjunto mas grande (el conjunto de animales que son aves).

# Aprendizaje de conceptos

- ▶ Usualmente aprendemos, a partir de situaciones y ejemplos, conceptos o categorías del estilo de
  - pájaro,
  - auto,
  - temas que tengo que estudiar para aprobar
- ▶ Cada concepto puede ser visto como descriptor de un subconjunto de objetos o eventos de un conjunto mas grande (el conjunto de animales que son aves).
- ▶ Un concepto es una función booleana sobre el conjunto mas grande que es verdadera sobre el subconjunto y falsa en otra parte.

# Aprendizaje de conceptos

Se busca inferir la definición de un concepto a partir de ejemplos rotulados como pertenecientes y no pertenecientes al concepto.



# Aprendizaje de conceptos

Se busca inferir la definición de un concepto a partir de ejemplos rotulados como pertenecientes y no pertenecientes al concepto.

Para ello se aproxima una función booleana que representa al concepto a partir del conjunto de ejemplos de entrenamiento

# Ejemplo de aprendizaje de concepto por inferencia inductiva:

## Concepto

Días en los cuales se puede practicar un deporte acuático

# Ejemplo de aprendizaje de concepto por inferencia inductiva:

## Concepto

Días en los cuales se puede practicar un deporte acuático

## Objetivo

Dado un día nuevo, con ciertas características, queremos dar una respuesta de SI si se puede practicar deporte o NO si no se puede.

# Ejemplo de aprendizaje de concepto por inferencia inductiva:

## Concepto

Días en los cuales se puede practicar un deporte acuático

## Objetivo

Dado un día nuevo, con ciertas características, queremos dar una respuesta de SI si se puede practicar deporte o NO si no se puede.

## Muestra

Tenemos un conjunto de ejemplo de días SI y días NO, y los valores que tomaron las características en esos días.

## Ejemplo de aprendizaje de concepto por inferencia inductiva:

Con esa información tenemos que inferir la función booleana que representa el concepto  $c : X \rightarrow \{0, 1\}$

# Ejemplo de aprendizaje de concepto por inferencia inductiva:

Con esa información tenemos que inferir la función booleana que representa el concepto  $c : X \rightarrow \{0, 1\}$

$X$  es el conjunto de todos los posibles datos (vectores de características observadas)

# Ejemplo de aprendizaje de concepto por inferencia inductiva:

Con esa información tenemos que inferir la función booleana que representa el concepto  $c : X \rightarrow \{0, 1\}$

$X$  es el conjunto de todos los posibles datos (vectores de características observadas)

El valor  $c(x) = 1$  se dá cuando el día  $x$  puede hacerse deporte acuático.

# Ejemplo de aprendizaje de concepto por inferencia inductiva:

- ▶ Las seis variables que se estudian por día son:



# Ejemplo de aprendizaje de concepto por inferencia inductiva:

- ▶ Las seis variables que se estudian por día son:
  - Cielo: sol, nublado, lluvia

# Ejemplo de aprendizaje de concepto por inferencia inductiva:

- ▶ Las seis variables que se estudian por día son:
  - Cielo: sol, nublado, lluvia
  - Temperatura del aire: cálido, frío

# Ejemplo de aprendizaje de concepto por inferencia inductiva:

► Las seis variables que se estudian por día son:

- Cielo: sol, nublado, lluvia
- Temperatura del aire: cálido, frío
- Humedad Ambiente: normal, alta

# Ejemplo de aprendizaje de concepto por inferencia inductiva:

► Las seis variables que se estudian por día son:

- Cielo: sol, nublado, lluvia
- Temperatura del aire: cálido, frío
- Humedad Ambiente: normal, alta
- Viento: débil, fuerte

# Ejemplo de aprendizaje de concepto por inferencia inductiva:

► Las seis variables que se estudian por día son:

- Cielo: sol, nublado, lluvia
- Temperatura del aire: cálido, frío
- Humedad Ambiente: normal, alta
- Viento: débil, fuerte
- Temperatura del agua: caliente, fría

# Ejemplo de aprendizaje de concepto por inferencia inductiva:

► Las seis variables que se estudian por día son:

- Cielo: sol, nublado, lluvia
- Temperatura del aire: cálido, frío
- Humedad Ambiente: normal, alta
- Viento: débil, fuerte
- Temperatura del agua: caliente, fría
- Pronóstico: estable, cambio

# Ejemplo de aprendizaje de concepto por inferencia inductiva:

- ▶ Las seis variables que se estudian por día son:
  - Cielo: sol, nublado, lluvia
  - Temperatura del aire: cálido, frío
  - Humedad Ambiente: normal, alta
  - Viento: débil, fuerte
  - Temperatura del agua: caliente, fría
  - Pronóstico: estable, cambio
- ▶  $3 \times 2 \times 2 \times 2 \times 2 \times 2 = 96$  , por lo cual  $X$  tiene 96 vectores de datos.

# Ejemplo de aprendizaje de concepto por inferencia inductiva:

- ▶ Las seis variables que se estudian por día son:
  - Cielo: sol, nublado, lluvia
  - Temperatura del aire: cálido, frío
  - Humedad Ambiente: normal, alta
  - Viento: débil, fuerte
  - Temperatura del agua: caliente, fría
  - Pronóstico: estable, cambio
- ▶  $3 \times 2 \times 2 \times 2 \times 2 \times 2 = 96$  , por lo cual  $X$  tiene 96 vectores de datos.
- ▶ La muestra de entrenamiento  $D$  es un subconjunto de  $X$  para el cual tengo los valores de la variable  $c$  que determina el concepto.



# Aprendizaje de conceptos

## Hipótesis

Conjunto de restricciones sobre las variables categóricas medidas.

# Aprendizaje de conceptos

## Hipótesis

Conjunto de restricciones sobre las variables categóricas medidas.

## Restricciones posibles

$\emptyset$ , ?, valor

# Aprendizaje de conceptos

## Hipótesis

Conjunto de restricciones sobre las variables categóricas medidas.

## Restricciones posibles

$\emptyset$ , ?, valor

- ▶  $\emptyset$  indica que ningún valor de la variable es aceptable

# Aprendizaje de conceptos

## Hipótesis

Conjunto de restricciones sobre las variables categóricas medidas.

## Restricciones posibles

$\emptyset$ , ?, valor

- ▶  $\emptyset$  indica que ningún valor de la variable es aceptable
- ▶ ? indica que cualquier valor de la variable es aceptable.

# Aprendizaje de conceptos

## Hipótesis

Conjunto de restricciones sobre las variables categóricas medidas.

## Restricciones posibles

$\emptyset$ , ?, valor

- ▶  $\emptyset$  indica que ningún valor de la variable es aceptable
- ▶ ? indica que cualquier valor de la variable es aceptable.
- ▶ `valor` es uno de los valores posibles de la variable.

# Aprendizaje de conceptos

Si un dato  $x$  satisface todas las restricciones de la hipótesis  $h$ , entonces  $h$  clasifica  $x$  como un ejemplo positivo  $h(x) = 1$ .

# Aprendizaje de conceptos

Si un dato  $x$  satisface todas las restricciones de la hipótesis  $h$ , entonces  $h$  clasifica  $x$  como un ejemplo positivo  $h(x) = 1$ .

- ▶ Hipótesis mas específica:  $h = \{\emptyset, \emptyset, \emptyset, \emptyset, \emptyset, \emptyset\}$  dice que ningún dato es un ejemplo positivo.

# Aprendizaje de conceptos

Si un dato  $x$  satisface todas las restricciones de la hipótesis  $h$ , entonces  $h$  clasifica  $x$  como un ejemplo positivo  $h(x) = 1$ .

- ▶ Hipótesis mas específica:  $h = \{\emptyset, \emptyset, \emptyset, \emptyset, \emptyset, \emptyset\}$  dice que ningún dato es un ejemplo positivo.
- ▶ Hipótesis mas general:  $h = \{?, ?, ?, ?, ?, ?\}$  dice que todos los datos son ejemplos positivos.



# Aprendizaje de conceptos

Si un dato  $x$  satisface todas las restricciones de la hipótesis  $h$ , entonces  $h$  clasifica  $x$  como un ejemplo positivo  $h(x) = 1$ .

- ▶ Hipótesis mas específica:  $h = \{\emptyset, \emptyset, \emptyset, \emptyset, \emptyset, \emptyset\}$  dice que ningún dato es un ejemplo positivo.
- ▶ Hipótesis mas general:  $h = \{?, ?, ?, ?, ?, ?\}$  dice que todos los datos son ejemplos positivos.

Se busca una hipótesis  $h$  tal que  $h(x) = c(x)$  para todo  $x$  en la muestra de entrenamiento.

# Aprendizaje de conceptos

► Ejemplos de entrenamiento:

Cielo	Aire	Humedad	Viento	Agua	Pronóstico	Deporte	<i>c</i>
sol	cálido	normal	fuerte	caliente	estable	si	1
sol	cálido	alta	fuerte	caliente	estable	si	1
lluvia	frío	alta	fuerte	caliente	cambio	no	0
sol	cálido	alta	fuerte	fría	cambio	si	1

# Aprendizaje de conceptos

► Ejemplos de entrenamiento:

Cielo	Aire	Humedad	Viento	Agua	Pronóstico	Deporte	$c$
sol	cálido	normal	fuerte	caliente	estable	si	1
sol	cálido	alta	fuerte	caliente	estable	si	1
lluvia	frío	alta	fuerte	caliente	cambio	no	0
sol	cálido	alta	fuerte	fría	cambio	si	1

- Los datos con Deporte = si son llamados datos positivos, i.e.  $c(x) = 1$ .

# Aprendizaje de conceptos

- ▶ Ejemplos de entrenamiento:

Cielo	Aire	Humedad	Viento	Agua	Pronóstico	Deporte	$c$
sol	cálido	normal	fuerte	caliente	estable	si	1
sol	cálido	alta	fuerte	caliente	estable	si	1
lluvia	frío	alta	fuerte	caliente	cambio	no	0
sol	cálido	alta	fuerte	fría	cambio	si	1

- ▶ Los datos con Deporte = si son llamados datos positivos, i.e.  $c(x) = 1$ .
- ▶  $h_1 = \{ \text{sol, ?, normal, ?, ?, cambio} \}$

# Aprendizaje de conceptos

► Ejemplos de entrenamiento:

Cielo	Aire	Humedad	Viento	Agua	Pronóstico	Deporte	$c$
sol	cálido	normal	fuerte	caliente	estable	si	1
sol	cálido	alta	fuerte	caliente	estable	si	1
lluvia	frío	alta	fuerte	caliente	cambio	no	0
sol	cálido	alta	fuerte	fría	cambio	si	1

- Los datos con Deporte = si son llamados datos positivos, i.e.  $c(x) = 1$ .
- $h_1 = \{ \text{sol, ?, normal, ?, ?, cambio} \}$
- Ningun  $x$  satisface las restricciones, la hipótesis es falsa para el concepto.

# Aprendizaje de conceptos

► Ejemplos de entrenamiento:

Cielo	Aire	Humedad	Viento	Agua	Pronóstico	Deporte	$c$
sol	cálido	normal	fuerte	caliente	estable	si	1
sol	cálido	alta	fuerte	caliente	estable	si	1
lluvia	frío	alta	fuerte	caliente	cambio	no	0
sol	cálido	alta	fuerte	fría	cambio	si	1

- Los datos con Deporte = si son llamados datos positivos, i.e.  $c(x) = 1$ .
- $h_1 = \{ \text{sol}, ?, \text{normal}, ?, ?, \text{cambio} \}$
- Ningun  $x$  satisface las restricciones, la hipótesis es falsa para el concepto.
- $h_2 = \{ ?, \text{cálido}, \text{alta}, ?, ?, \text{estable} \}$

# Aprendizaje de conceptos

## ► Ejemplos de entrenamiento:

Cielo	Aire	Humedad	Viento	Agua	Pronóstico	Deporte	$c$
sol	cálido	normal	fuerte	caliente	estable	si	1
sol	cálido	alta	fuerte	caliente	estable	si	1
lluvia	frío	alta	fuerte	caliente	cambio	no	0
sol	cálido	alta	fuerte	fría	cambio	si	1

## ► Los datos con Deporte = si son llamados datos positivos, i.e.

$$c(x) = 1.$$

## ► $h_1 = \{ \text{sol, ?, normal, ?, ?, cambio} \}$

- Ningun  $x$  satisface las restricciones, la hipótesis es falsa para el concepto.

## ► $h_2 = \{ \text{?, cálido, alta, ?, ?, estable} \}$

- $x = \{ \text{sol, cálido, alta, fuerte, caliente, estable} \}$  satisface la hipótesis, es verdadera para el concepto.

# Aprendizaje de conceptos

- ▶ El ejemplo tiene  $3 \times 2^5 = 96$  posibles datos. Y si agrego a los valores los símbolos  $?$  y  $\emptyset$  quedan  $5 \times 4^5 = 5120$  posibles hipótesis en  $H$ .



# Aprendizaje de conceptos

- ▶ El ejemplo tiene  $3 \times 2^5 = 96$  posibles datos. Y si agrego a los valores los símbolos  $?$  y  $\emptyset$  quedan  $5 \times 4^5 = 5120$  posibles hipótesis en  $H$ .
- ▶ Como toda hipótesis con uno o mas valores  $\emptyset$  clasifica a todo el conjunto  $X$  como negativo, solo se tiene  $1 + (4 \times 3^5) = 973$  hipótesis con sentido.

# Aprendizaje de conceptos

- ▶ El ejemplo tiene  $3 \times 2^5 = 96$  posibles datos. Y si agrego a los valores los símbolos  $?$  y  $\emptyset$  quedan  $5 \times 4^5 = 5120$  posibles hipótesis en  $H$ .
- ▶ Como toda hipótesis con uno o mas valores  $\emptyset$  clasifica a todo el conjunto  $X$  como negativo, solo se tiene  $1 + (4 \times 3^5) = 973$  hipótesis con sentido.
- ▶ Como se encuentra la hipótesis que coincide con  $c$  en la muestra? hay mas de una?

# Aprendizaje de conceptos

- ▶ El ejemplo tiene  $3 \times 2^5 = 96$  posibles datos. Y si agrego a los valores los símbolos ? y  $\emptyset$  quedan  $5 \times 4^5 = 5120$  posibles hipótesis en  $H$ .
- ▶ Como toda hipótesis con uno o mas valores  $\emptyset$  clasifica a todo el conjunto  $X$  como negativo, solo se tiene  $1 + (4 \times 3^5) = 973$  hipótesis con sentido.
- ▶ Como se encuentra la hipótesis que coincide con  $c$  en la muestra? hay mas de una?
- ▶ Se busca el espacio  $H$  usando el orden de hipótesis de lo general a lo específico.  $h_1$  es más general que  $h_2$  si todo dato que satisface  $h_2$  también lo satisface  $h_1$ .

# Aprendizaje de conceptos

Métodos de búsqueda de las hipótesis candidatas:

- ▶ **FIND-S**: Usando solo los ejemplos positivos encuentra la hipótesis específica maximal, yendo de la mas específica a la mas general que permitan los datos.

# Aprendizaje de conceptos

Métodos de búsqueda de las hipótesis candidatas:

- ▶ **FIND-S**: Usando solo los ejemplos positivos encuentra la hipótesis específica maximal, yendo de la mas específica a la mas general que permitan los datos.
- ▶ **List and Eliminate (Rote Learning)**: Barre el árbol completo de hipótesis comenzando desde la mas general para encontrar todas las hipótesis verdaderas, usando muestras positivas y negativas, lo cual puede resultar muy complicado.

# Aprendizaje de conceptos

Métodos de búsqueda de las hipótesis candidatas:

- ▶ **FIND-S**: Usando solo los ejemplos positivos encuentra la hipótesis específica maximal, yendo de la mas específica a la mas general que permitan los datos.
- ▶ **List and Eliminate (Rote Learning)**: Barre el árbol completo de hipótesis comenzando desde la mas general para encontrar todas las hipótesis verdaderas, usando muestras positivas y negativas, lo cual puede resultar muy complicado.
- ▶ **Candidate Elimination**: Realiza la búsqueda exhaustiva en una versión acotada del espacio de hipótesis, mejorando ambos algoritmos anteriores.

# Aprendizaje de conceptos

► FIND-S:

# Aprendizaje de conceptos

- ▶ FIND-S:
- ▶ Se empieza con la hipótesis más específica

$$h_0 = \{\emptyset, \emptyset, \emptyset, \emptyset, \emptyset, \emptyset\}$$



# Aprendizaje de conceptos

- ▶ FIND-S:
- ▶ Se empieza con la hipótesis más específica

$$h_0 = \{\emptyset, \emptyset, \emptyset, \emptyset, \emptyset, \emptyset\}$$

- ▶ Se elije el primer dato y si es negativo no se cambia la hipótesis.

# Aprendizaje de conceptos

- ▶ FIND-S:
- ▶ Se empieza con la hipótesis más específica

$$h_0 = \{\emptyset, \emptyset, \emptyset, \emptyset, \emptyset, \emptyset\}$$

- ▶ Se elije el primer dato y si es negativo no se cambia la hipótesis.
- ▶ Si el dato es positivo y la hipótesis inicial es muy específica se actualiza la hipótesis con la condición más general.

# Aprendizaje de conceptos

- ▶ FIND-S:
- ▶ Se empieza con la hipótesis más específica

$$h_0 = \{\emptyset, \emptyset, \emptyset, \emptyset, \emptyset, \emptyset\}$$

- ▶ Se elije el primer dato y si es negativo no se cambia la hipótesis.
- ▶ Si el dato es positivo y la hipótesis inicial es muy específica se actualiza la hipótesis con la condición más general.
  - Para cada restricción de atributo  $a_i$  de  $h$ , si el dato  $x$  satisface  $a_i$  no hacer nada, si no reemplazar  $a_i$  por la restricción mas general que sea satisfecha por  $x$ .

# Aprendizaje de conceptos

- ▶ FIND-S:
- ▶ Se empieza con la hipótesis más específica

$$h_0 = \{\emptyset, \emptyset, \emptyset, \emptyset, \emptyset, \emptyset\}$$

- ▶ Se elije el primer dato y si es negativo no se cambia la hipótesis.
- ▶ Si el dato es positivo y la hipótesis inicial es muy específica se actualiza la hipótesis con la condición más general.
  - Para cada restricción de atributo  $a_i$  de  $h$ , si el dato  $x$  satisface  $a_i$  no hacer nada, si no reemplazar  $a_i$  por la restricción mas general que sea satisfecha por  $x$ .
- ▶ Se repite hasta que se completa la revisión de todas las muestras.

# Aprendizaje de conceptos

- ▶ FIND-S:
- ▶ Se empieza con la hipótesis más específica

$$h_0 = \{\emptyset, \emptyset, \emptyset, \emptyset, \emptyset, \emptyset\}$$

- ▶ Se elije el primer dato y si es negativo no se cambia la hipótesis.
- ▶ Si el dato es positivo y la hipótesis inicial es muy específica se actualiza la hipótesis con la condición más general.
  - Para cada restricción de atributo  $a_i$  de  $h$ , si el dato  $x$  satisface  $a_i$  no hacer nada, si no reemplazar  $a_i$  por la restricción mas general que sea satisfecha por  $x$ .
- ▶ Se repite hasta que se completa la revisión de todas las muestras.
- ▶ Una vez terminado, se tiene una función booleana que puede clasificar nuevos ejemplos de esta forma, si la hipótesis es verdadera para la muestra, se clasifica como positivo, y si no es negativo.

# Find S

- ▶ Empieza por  $h_0 = \{\emptyset, \emptyset, \emptyset, \emptyset, \emptyset, \emptyset\}$

# Find S

- ▶ Empieza por  $h_0 = \{\emptyset, \emptyset, \emptyset, \emptyset, \emptyset, \emptyset\}$
- ▶ Toma el primer  $x$  positivo y reemplaza el anterior por esta instancia

$$h_1 = \{ \text{sol, cálido, normal, fuerte, caliente, estable} \}$$

# Find S

- ▶ Empieza por  $h_0 = \{\emptyset, \emptyset, \emptyset, \emptyset, \emptyset, \emptyset\}$
- ▶ Toma el primer  $x$  positivo y reemplaza el anterior por esta instancia

$$h_1 = \{ \text{sol, cálido, normal, fuerte, caliente, estable} \}$$

- ▶ Toma el segundo positivo  $x = ( \text{sol, cálido, alta, fuerte, caliente, estable} )$  y generaliza  $h_1$  por

$$h_2 = \{ \text{sol, cálido, ?, fuerte, caliente, estable} \}$$



# Find S

- ▶ Empieza por  $h_0 = \{\emptyset, \emptyset, \emptyset, \emptyset, \emptyset, \emptyset\}$
- ▶ Toma el primer  $x$  positivo y reemplaza el anterior por esta instancia

$$h_1 = \{ \text{sol, cálido, normal, fuerte, caliente, estable} \}$$

- ▶ Toma el segundo positivo  $x = ( \text{sol, cálido, alta, fuerte, caliente, estable} )$  y generaliza  $h_1$  por

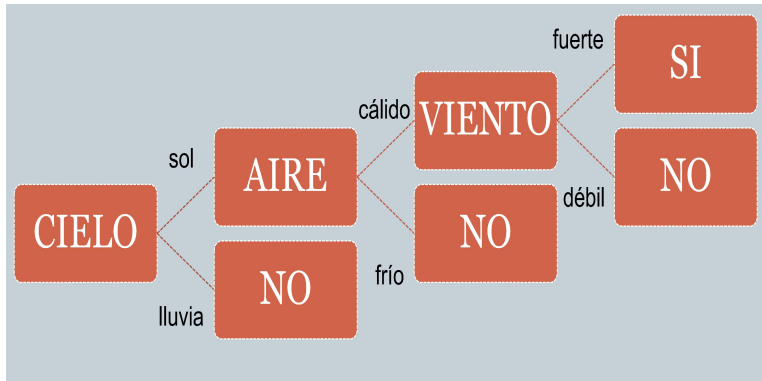
$$h_2 = \{ \text{sol, cálido, ?, fuerte, caliente, estable} \}$$

- ▶ Toma el tercer positivo  $x = ( \text{sol, cálido, alta, fuerte, fría, cambio} )$  y generaliza

$$h_3 = \{ \text{sol, cálido, ?, fuerte, ?, ?} \}$$

el cual es el resultado.

# Aprendizaje de conceptos



# Aprendizaje de conceptos

Problemas....

# Aprendizaje de conceptos

Problemas....

- ▶ **Memorizador (Rote Learning)**: Responde solo si la nueva instancia esta previamente almacenada. No tiene sesgo inductivo pero enumera todas las hipótesis.

# Aprendizaje de conceptos

## Problemas....

- ▶ **Memorizador (Rote Learning)**: Responde solo si la nueva instancia esta previamente almacenada. No tiene sesgo inductivo pero enumera todas las hipótesis.
- ▶ **Candidate elimination**: Responde solo si para la nueva instancia hay acuerdo entre todas las hipótesis. Hay sesgo inductivo pues supone que el espacio considerado contiene el concepto.

# Aprendizaje de conceptos

## Problemas....

- ▶ **Memorizador (Rote Learning)**: Responde solo si la nueva instancia esta previamente almacenada. No tiene sesgo inductivo pero enumera todas las hipótesis.
- ▶ **Candidate elimination**: Responde solo si para la nueva instancia hay acuerdo entre todas las hipótesis. Hay sesgo inductivo pues supone que el espacio considerado contiene el concepto.
- ▶ **FIND-S**: Responde siempre ante una nueva instancia positiva con sesgo inductivo muy grande. Solo encuentra un tipo de hipótesis, pero puede haber otras.

# Aprendizaje de conceptos

Problemas....

- ▶ **Memorizador (Rote Learning)**: Responde solo si la nueva instancia esta previamente almacenada. No tiene sesgo inductivo pero enumera todas las hipótesis.
- ▶ **Candidate elimination**: Responde solo si para la nueva instancia hay acuerdo entre todas las hipótesis. Hay sesgo inductivo pues supone que el espacio considerado contiene el concepto.
- ▶ **FIND-S**: Responde siempre ante una nueva instancia positiva con sesgo inductivo muy grande. Solo encuentra un tipo de hipótesis, pero puede haber otras.

El aprendizaje de conceptos adolece de intolerancia al ruido !

# Bibliográfica

- ▶ Mitchell, T. Machine Learning, McGraw Hill, 1997.



# Bibliográfica

- ▶ Mitchell, T. Machine Learning, McGraw Hill, 1997.
- ▶ Marsland, S. Machine Learning: an algorithmic perspective. CRC Press 2009