

TEMA 1. Introducción

Francisco José Ribadas Pena, Santiago Fernández Lanza

Modelos de Razonamiento y Aprendizaje

5º Informática

ribadas@uvigo.es, sflanza@uvigo.es

28 de enero de 2013

- 1.1 Aprendizaje automático
- 1.2 Conceptos básicos
- 1.3 Tareas/Aplicaciones típicas
- 1.4 Tipos de Aprendiziz. Automático

1.1 Aprendizaje automático

¿Qué es aprender?

- Memorizar (Ej. una poesía, la lista de los reyes godos, ...)
- Adquirir una destreza o una capacidad
 - Psicomotriz (Ej. conducir un coche, caminar, ...)
 - de cálculo (Ej. hacer operaciones aritméticas)
 - Lingüística (Leer, escribir, ...)
- Adquirir un concepto (Ej. animal, perro, ...)
- ...

Aprendizaje Natural vs Aprendizaje Automático

- Polémica análoga a Inteligencia Natural vs Inteligencia Artificial

¿Se puede dar una definición de aprendizaje?

- Difícil respuesta
- Para el aprendizaje automático necesitamos comprometernos con una definición

(a) DEFINICIÓN

APRENDIZAJE AUTOMÁTICO: "Proceso mediante el cual un sistema mejora su rendimiento en una tarea determinada en base a una experiencia."

En todo sistema de aprendizaje automático será necesario definir y modelizar:

- Cúal es la tarea a mejorar
- De qué experiencia se dispone
- Cómo se mide el rendimiento

Ejemplos:

Tarea:	Experiencia:	Medida Rendimiento
jugar a las damas reconoc. texto manuscrito diagnóstico médico	jugar contra uno mismo muestra textos clasificados conj. sintomas+diagnostico	% partidas ganadas n. letras erroneas % diagnósticos correctos

CONCLUSIÓN: En el diseño de un sistema que aprenda se debe elegir:

- Lo que se va aprender (hipótesis/concepto objetivo)
- Lenguaje de representación de la hipótesis objetivo
- Experiencia a usar en el aprendizaje (y como y donde obtenerla)
- Algoritmo para aprender la hipótesis objetivo a partir de esa experiencia

(b) UTILIDAD

1. Desarrollo de sistemas demasiado difíciles o imposibles de construir de forma manual
 - Entornos que requieran conocimiento muy específico y detallado (no disponible)
 - Ej.: problemas de percepción (reconocer caracteres manuscritos, reconcim. de voz, aparcar un coche, ...)
2. Desarrollo de sistemas que puedan adaptarse y modificarse por si mismos
 - Sistemas capaces de actuar en entornos cambiantes o nuevos
 - Ej.: sistemas que se adapten a las necesidades de un usuario (asistentes personales, filtros web, conducir un coche, etc...)
3. Descubrimiento de conocimiento en grandes bancos de datos
 - Técnicas de *Data Mining*
 - Ej.: identificación de patrones de compra
4. Modelos computacionales del aprendizaje biológico y del conocimiento humano

(c) ÁREAS RELACIONADAS

1. Inteligencia Artificial
 - Representación conocimiento
 - Aprendizaje como problema de búsqueda,
2. Matemáticas
 - Estadística
 - teorema de Bayes (vincula la probabilidad de A dado B con la probabilidad de B dado A)
 - técnicas de estimación (técnicas que permiten dar un valor aproximado de un parámetro de una población a partir de los datos proporcionados por una muestra)
 - contraste de hipótesis (procedimiento para juzgar si una propiedad que se supone cumple una población estadística es compatible con lo observado en una muestra de dicha población)
 - ...
 - Lógica (formalismos representación y razonamiento)
3. Psicología: modelos del razonamiento humano
4. Biología: Neurobiología (modelos del cerebro y neuronas)
5. Teoría de la información (entropía y contenido de información)
6. Filosofía
7. ...

(d) EVOLUCION HISTÓRICA

- **50's:** Juego de Damas de Samuel (1er sistema de aprendizaje)
 - Aprendía(mejoraba) los parámetros de la función de evaluación de posiciones del juego, jugando contra si mismo o con humanos
- **60's:** Desarrollo y 1er. auge de las redes de neuronas artificiales (perceptrones)
- **70's:** Inducción de conceptos simbólicos.
Algoritmo ID3 (mediante el uso de árboles decisión, se buscan hipótesis o reglas, dado un conjunto de ejemplos)
- **80's:** Resurgir de RNA (nuevos algoritmos de aprendizaje y modelos de redes)
 - Progreso en aprendizaje de árboles de decisión y de reglas.
 - Aprendizaje deductivo (basado en la explicación-demostración)
- **90's:** Aplicaciones reales a mayor escala: data mining, agentes adaptativos.
 - Programación lógica inductiva, aprendizaje redes bayesianas

1.2 Conceptos básicos

(a) CONCEPTOS o HIPÓTESIS

- Tipos de “cosas” que pueden ser aprendidas
- Salida de algoritmo de aprendizaje: descripción/representación del concepto/hipótesis aprendido
- En la mayoría de los casos, se puede modelar como una función $c(e)$ que ofrece una salida para un ejemplo e .
- El algoritmo de aprendizaje debe encontrar/generar una hipótesis h , que haga $h(e) = c(e)$ para los ejemplos de entrenamiento.
- **OBJETIVO:** El algoritmo de aprendizaje debe ser capaz de generalizar y crear una hipótesis h que sea “suficientemente” correcta para ejemplos distintos a los del conjunto de entrenamiento. ($h(e) \approx c(e)$)

(b) REPRESENTACIÓN HIPÓTESIS

- Múltiples posibilidades para representar las hipótesis que maneja y aprende el sistema de aprendizaje
- Ejemplos: $\left\{ \begin{array}{l} \text{árboles de decisión} \\ \text{reglas de producción} \\ \text{lógica de predicados} \\ k \text{ vecinos más próximos y función de proximidad} \\ \text{pesos de conexiones en redes neuronales} \\ \text{secuencias de “genes”} \end{array} \right.$
- El tipo de representación condiciona: $\left\{ \begin{array}{l} \text{qué se puede aprender} \\ \text{cómo se realiza ese aprendizaje} \end{array} \right.$
- En general, podrá modelizarse el aprendizaje como una búsqueda de las hipótesis que mejor se adapten a los datos de entrenamiento.
 - Análogo a búsquedas en espacios de estados
estados \approx representaciones de hipótesis
 - Distintos algoritmos de aprendizaje \leftrightarrow $\left\{ \begin{array}{l} \text{distintas representaciones y} \\ \text{distintas estrategias de búsqueda} \end{array} \right.$

(c) EJEMPLOS DE ENTRENAMIENTO

- Elementos de entrada al sistema de aprendizaje
 - Casos individuales e independientes del concepto objetivo.
 - Elementos sobre los que se aprende a clasificar, asociar, agrupar, predecir, etc..
 - Caracterizados por un conjunto predeterminado de atributos.
 - En tareas de clasificación: atributo especial clase : objetivo del aprendizaje
- En aprendizaje supervisado : incluyen indicación de la salida correcta
 - Si el resultado es booleano (o discreto)
 - Ejemplos positivos: son *verdaderos* (pertenecen a la clase)
 - Ejemplos negativos: son *falsos* (no pertenecen a la clase)
- Entrada algoritmo de aprendizaje: conjunto de ejemplos (*dataset*)
 - Normalmente: fichero plano con todos los ejemplos y sus atributos
 - Estructura análoga a bases de datos relacionales

(d) ATRIBUTOS

- Características de interés que describen un ejemplo.
- Problemas: $\left\{ \begin{array}{l} \text{Dependencias entre atributos} \\ \text{Atributos irrelevantes} \\ \text{Valores desconocidos o incorrectos} \end{array} \right.$
- Tipos de atributos:
 - **Nominales**: los valores del atributos son etiquetas,
 - No es posible definir ordenación ni distancia entre atributos
 - Sólo operaciones de igualdad
 - Ej.: culpable/inocente, gripe/ébola/bronquitis/...
 - **Ordinales**: valores nominales o numéricos
 - Existe un relación de orden entre los valores
 - No posible definir distancia de forma inmediata y evidente
 - Ej.: enero/febrero/..., suspenso/aprobado/notable/..., 0/5/7.5/9/10
 - **Numéricos**: valores numéricos
 - Posible definir orden y distancia

1.3 Tareas/aplicaciones típicas

(a) APRENDIZAJE DE CLASIFICACIONES

- Asignación de un ejemplo a una clase o categoría.
 - Salida: clase
 - Frecuentemente salida booleana
- Aprendizaje supervisado: cada ejemplo figura acompañado de la clase a la que pertenece
- Evaluación: usando nuevos ejemplos cuya clase sea conocida
- *Ejemplos:* predicción meteorológica, decisión créditos bancarios, etc..

(b) APRENDIZAJE DE ASOCIACIONES

- Detectar asociaciones entre distintas características de los datos de entrenamiento.
- Generalización del problema de la clasificación:
 - Se predice sobre cualquier valor de atributo
 - Puede predecirse más de un valor
- Necesidad de definir restricciones (muchas reglas posibles)
 - Maximizar/minimizar cobertura y/o exactitud
- *Ejemplos:* Asociaciones entre productos en las compras de clientes prototipo

(c) CLUSTERING (Agrupamiento)

- Identificación de grupos de elementos similares.
- Se busca maximizar la similitud dentro del grupo y minimizarla con respecto a elementos de otros grupos.
- Aprendizaje no supervisado: la clase de cada ejemplo es desconocida
- *Ejemplos:* { identificación de tipo de perfiles generales de clientes
agrupamiento de documentos por temática similar

(d) PREDICCIÓN NUMÉRICA

- Predecir una cantidad numérica en función de los atributos del ejemplo
- Similar a clasificación (con salida numérica y continua)
- Aprendizaje supervisado
- *Ejemplos:* predicción de ingresos, valores bursátiles, etc...

1.4 Tipos de Aprendizaje Automático

(a) En función de cómo sea la experiencia que se utiliza durante el entrenamiento:

1. APRENDIZAJE SUPERVISADO

- Algoritmo de aprendizaje tiene información del resultado asociado a cada ejemplo.
 - Conoce el ejemplo e y el resultado correcto de $h(e)$
- Conoce lo que es correcto y lo que no lo es.
 - existe un “maestro” que indica lo que es correcto
 - conocido *a priori*: disponible antes de iniciar entrenamiento
 - típico en problemas de clasificación
 - Ejemplo: lista de síntomas y diagnóstico correcto
 - conocido *durante aprendizaje*: existe un mecanismo (de tipo premio/castigo) que informa de la bondad del resultado obtenido
 - mezcla aprendizaje y actuación
 - usado en aprendizaje por refuerzo
 - se aprende como maximizar las recompensas
 - Ejemplo: robot corrige trayectoria (a golpes)

2. APRENDIZAJE NO SUPERVISADO

- Algoritmo no tiene acceso a información sobre el resultado
- Sólo se conocen los ejemplos, no sus resultados
- Ejemplo: problemas de clustering (clasificación sin clases pre-definidas)

(b) En función del método utilizado para aprender (representación hipótesis + algoritmo aprendizaje):

1. APRENDIZAJE INDUCTIVO

- Sigue el esquema del razonamiento inductivo:
 - Identifica conocimiento general (hipótesis) a partir de conocimiento particular (ejemplos)
 - Inicialmente, no hay hipótesis de partida disponibles, sólo ejemplos.

Raz. Deductivo

Raz. Inductivo

Premisas

Premisas

Conclusión (se sigue
NECESARIAMENTE de las
premisas)

Conclusión (se sigue
PROBABLEMENTE de las
premisas)

(NOTA: No confundir el razonamiento inductivo con la inducción matemática.)

- Incluye:
 - Aprendizaje de árboles de decisión
 - Programación lógica inductiva
 - Redes neuronas artificiales
 - ...

2. APRENDIZAJE DEDUCTIVO/ANALÍTICO

- Sigue el modelo de razonamiento deductivo
 - Se parte de un conocimiento inicial correcto (teoría del dominio) y se aprende a mejorarlo

Conocimiento General (c)		
Premisas (e)	→	Nuevo Conocimiento General (n)
Conclusiones (c(e))		(h = c + n)

- El sistema no aprende nada nuevo
- Encuentra formas más eficientes de usar lo que sabe (mejora la forma de “razonar”)
- Incluye:
 - Aprendizaje basado en la explicación
 - utiliza las demostraciones lógicas usadas en las deducciones para crear nuevas expresiones lógicas que aceleren la deducción
 - Aprendizaje de MacroOperadores
 - usado en sistemas de búsqueda en espacio de estados
 - genera nuevos operadores combinando los utilizados en búsquedas anteriores

3. APRENDIZAJE BASADO EN EJEMPLOS (Instancias)

- Aprendizaje por analogía
- Puede verse como una aproximación mixta (inductiva/deductiva)
- No se generan representaciones explícitas de las hipótesis.
- Simplemente se almacenan los ejemplos de entrenamiento.
- Se utiliza una función de distancia.
- Ante la entrada de un ejemplo nuevo se aplica esa función y se recupera el ejemplo o ejemplos más cercanos y a partir de ellos se genera el resultado.

4. APRENDIZAJE POR REFUERZO

- Aprendizaje propio de agentes autónomos que aprenden a elegir las acciones óptimas para sus objetivos en función del resultado de sus acciones.
 - Ante cada acción del sistema, un entrenador externo premia o castiga su acción
 - Sistema aprende a generar secuencias de acciones que provean mayores recompensas.

5. OTROS TIPOS

- APRENDIZAJE PROBABILISTICO/BAYESIANO
 - Incorporación de técnicas estadísticas y de razonamiento probabilístico.
 - Ejemplos: Clasificadores Bayesianos
- APRENDIZAJE CONEXIONISTA
 - Tipo especial de aprendizaje inductivo
 - Uso de modelos computacionales del cerebro humanos
 - Sistemas complejos formados por elementos simples (neuronas artificiales)
 - Sistema RNA para aprendizaje supervisado y no supervisado
- APRENDIZAJE EVOLUTIVO
 - Emulación de la evolución biológica.
 - Hipótesis codificadas en forma de cadenas de genes sobre los que realizan cruces y mutaciones.

EN GENERAL:

- Distintos métodos tienen propiedades específicas que los hacen más apropiados para determinadas tareas.
- Muy difícil definir algoritmos de aprendizaje generales.