

Técnicas de inteligencia artificial

Adriana Cervantes Castillo

Repaso examen

Contenido

- Repaso Examen. Temas 1-10

Tema 1: Introducción

- Definición del aprendizaje:
- Un programa de ordenador aprende de la experiencia E con respecto a una clase de tareas T y una **medida de rendimiento P** , si su rendimiento en las tareas T , medido en base a la medida P , mejora con la experiencia E .

Ejemplo 1: Aprender a detectar robos de tarjetas de crédito.

- T : detectar robos de tarjetas de crédito.
- P : porcentaje de robos detectados.
- E : base de datos de hábitos de compra con la tarjeta de crédito.

- Los ordenadores pueden ser muy eficaces en el **aprendizaje de conceptos** que pueden ser representados mediante distintos tipos de estructuras como árboles, reglas o funciones matemáticas.
- En la búsqueda de las mejores hipótesis en base a los datos de entrenamiento, muchas veces se encuentran soluciones demasiado específicas, que cubren de manera muy exacta los datos de entrenamiento. Esto puede generar un problema de **sobreajuste (*overfitting*)**, siendo la solución adoptada demasiado específica respecto a los datos de entrenamiento y no se generaliza bien.

- Clasificación supervisada
 - Datos etiquetados
 - Conjunto de datos
 - Conjunto de entrenamiento
 - Conjunto de prueba.

¿Qué etapas comprende el aprendizaje de un concepto?

La Figura 7 muestra las etapas de aprendizaje de un concepto que se describen a continuación.

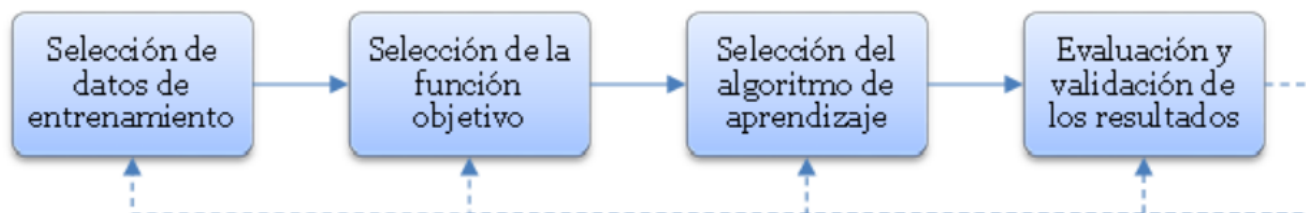


Figura 7. Etapas del aprendizaje de un concepto

- Etapas en el descubrimiento de conocimiento
 - Pre-procesamiento de los datos: los datos ruidosos y datos inconsistentes son tratados. Por ejemplo, una base de datos puede contener datos duplicados, valores incorrectos o desconocidos. Se pueden utilizar métodos automáticos estadísticos como la aplicación de *outliers* (por ejemplo, eliminar el 5 % de los valores más lejanos a la predicción). Este paso es denominado limpieza de datos por algunos autores y requiere conocer muy bien los datos, habiendo un riesgo de descarte de ejemplos de clases poco frecuentes.

Tema 2. Python

- [Scikit-learn](#) es una de las librerías de *machine learning* más populares para los algoritmos de *machine learning* clásico. Está construida sobre dos bibliotecas básicas de Python, NumPy y SciPy. Scikit-learn soporta la mayoría de los algoritmos de aprendizaje supervisado y no supervisado. Scikit-learn también puede ser usado para la minería de datos y el análisis de datos, lo que lo convierte en una gran herramienta para los que empiezan con el *machine learning*.
- [Matplotlib](#) es una biblioteca de Python muy popular para la visualización de datos.

Tema 3 Árboles de decisión

- Cuando el conjunto de datos disponible es pequeño se suele emplear la técnica conocida como validación cruzada de k iteraciones (*K-fold cross-validation*). Mediante esta técnica los datos se dividen en k subconjuntos de igual tamaño. Un subconjunto es utilizado como datos de prueba (o validación), mientras que el resto de los subconjuntos, (los $k-1$ subconjuntos restantes), son utilizados como datos de entrenamiento. La validación cruzada se repite k veces, y en cada una de las repeticiones se utiliza uno de los subconjuntos como datos de prueba.

- En los árboles de decisión, son métodos para la selección de los atributos
 - Tasa de ganancia
 - Índice Gini,
 - Ganancia de información.

Medidas usadas para evaluar un clasificador

- **Sensibilidad** (*sensitivity*): también conocido como ratio de verdaderos positivos (TPR – *True Positive Rate*). Mide la capacidad del clasificador para detectar la clase en una población con instancias de esa clase, la proporción de casos positivos bien detectados. Se calcula como $TP / (TP + FN)$.
- **Especificidad** (*specificity*): mide la capacidad del clasificador para descartar instancias que no pertenecen a la clase en una población con instancias de esa clase, la proporción de casos negativos correctamente detectados. Se calcula como $TN / (TN + FP)$.
- **Ratio de falsos positivos** (FPR – *False Positive Rate*): Proporción de casos negativos que el clasificador detecta como positivos. De forma más concreta $FP / (FP + TN)$.

Aprendizaje integrado

Existen tres métodos principales dentro del aprendizaje integrado en función de cómo se combinan los diferentes algoritmos que componen cada método: el *stacking* (apilamiento), el *bagging* (embolsamiento) y el *boosting* (refuerzo).

Boosting

- En el caso de los métodos *boosting*, los diferentes algoritmos elegidos se entrenan uno por uno de forma secuencial en serie (en lugar de trabajar en paralelo, como sucede en los métodos *stacking* y *bagging*)

Tema 4: Reglas

Específicamente dos medidas populares
que se suelen utilizar con el fin de evaluar una regla son:

Confianza

Soporte

- Existen diversos algoritmos para generar reglas de asociación, uno de los algoritmos más populares se denomina algoritmo *a priori*.
- El algoritmo *apriori* pretende generar *ítem-sets* que cumplan una cobertura mínima de manera eficiente. Un ítem es un par atributo-valor mientras que un *ítem-set* es un conjunto de *pares atributo-valor*. Un *k-ítem-set* es un conjunto de *k pares atributo-valor*. La cobertura de un *ítem-sets* se refiere al número de instancias que cumplen los valores en el *ítem-set* y va a determinar la cobertura de las reglas generadas a partir de dicho *ítem-set*.

Tema 5. Redes neuronales

- La red neuronal artificial más sencilla es el perceptrón, formado por una sola capa con las entradas y salidas que se necesiten para resolver el problema planteado. La máxima simplificación de este perceptrón es lo que se conoce como perceptrón simple, el cual está formado por una sola neurona.

El funcionamiento de la red neuronal vendrá determinado por los siguientes parámetros:

- Arquitectura de la red, esto es, el número de capas, número de neuronas por capa y las conexiones entre neuronas entre diferentes capas.
- Función de activación: función signo, función escalón, etc.
- Algoritmo de aprendizaje determinado principalmente por la regla de aprendizaje para ajustar pesos

Tema 6 – Deep learning

- Test de Turing (Turing, 1950).
 - un sistema se denomina inteligente si es capaz de responder a una serie de preguntas escritas y un humano no sería capaz de diferenciar si las respuestas han sido dadas por un humano o por un ordenador.
- *Aprendizaje profundo o deep learning*, cuando tenemos datos complejos con características poco claras o cuando queremos alcanzar niveles más altos de precisión que el aprendizaje automático clásico.
 - Se puede definir el aprendizaje profundo o *deep learning* como una clase de algoritmos de aprendizaje automático que utiliza múltiples capas para extraer progresivamente características de nivel superior de la entrada bruta.

- **Redes prealimentadas (*Feedforward Networks*)**. Todas las neuronas de una capa están conectadas con todas las neuronas de la capa siguiente, no existe conexiones entre neuronas de la misma capa y no existe retroalimentación.
- **Redes neuronales recurrentes**. Tienen neuronas específicas que incluyen una memoria en la que almacenan las entradas correspondientes a ciclos anteriores. De esta manera, son capaces de producir una salida dependiente de la entrada actual y de las entradas anteriores.

- Los autocodificadores o *autoencoders* (AE) son redes neuronales simétricas en forma de reloj de arena en las que las capas ocultas son más pequeñas que las capas de entrada y de salida (que son células de entrada y de salida que coinciden).

Tema 7: Clustering: agrupamiento o clasificación no supervisada

- **Clustering**. método de aprendizaje no supervisado que permite agrupar objetos en clústers o agrupamientos, cuyos miembros son similares entre si en cierto modo.
- **Cluster**: colección de objetos similares entre si y diferentes de los objetos que pertenecen a otros clústeres
- **k-means** . Algoritmo de partición que utiliza una **medida de distancia** para generar los clústeres. Método iterativo basado en distancia que define k centroides de los k clústeres que genera en cada iteración. Un objeto pertenece a un clúster o a otro en base a su distancia a los centroides de los distintos clústeres. El centroide de un clúster lo calcula como el valor medio de los puntos del clúster.

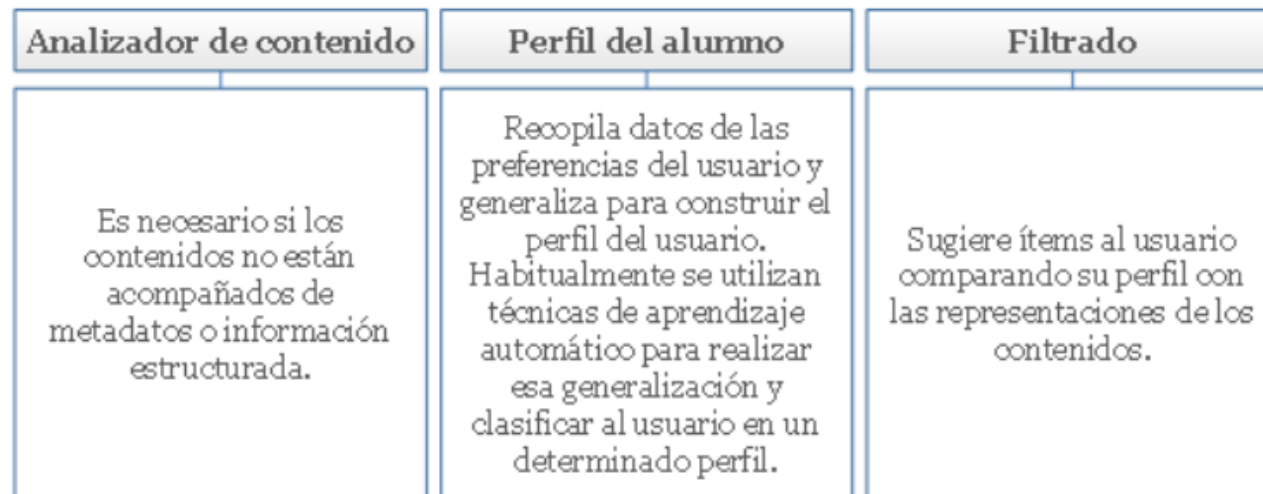
- Los **algoritmos de clustering basados en probabilidad**, en vez de situar las instancias en clústeres de una manera exclusiva, indican la probabilidad de que las instancias pertenezcan a cada uno de los clústeres.
- La base de estos algoritmos es el modelo estadístico denominado mezclas finitas (*finite mixtures*).
- **Algoritmo Fuzzy C-means** es un método muy utilizado para obtener clústeres solapados. Permite encontrar una serie de conjuntos difusos representativos de cada clúster y los grados de pertenencia de cada instancia a los mismos.

Tema 8. Sistemas de recomendación

- Según Montaner *et al.* (2003). El primer paso al crear un sistema recomendador es diseñar cómo se va a **representar el perfil** del usuario.
- **Filtrado colaborativo (recomendación colaborativa)**: realiza recomendaciones usualmente en base a la retroalimentación o el *feedback* que los conjuntos de usuarios dan explícitamente sobre un ítem. También se pueden basar en información capturada implícitamente como, por ejemplo, el historial de compras de los usuarios.
 - Los sistemas de recomendación colaborativa recomiendan los ítems a un usuario basándose en las preferencias o historial de otros usuarios.

- Uno de los algoritmos de predicción más empleado y muy simple en los sistemas de Filtrado Colaborativo basado en usuario es *Slope One*.

- Se pueden distinguir tres componentes principales en los recomendadores de contenidos (Lops *et al.*, 2011):



- Los **sistemas de recomendación híbridos** combinan dos o más estrategias de recomendación para, así, beneficiarse de sus ventajas complementarias
 - **Ponderados (*Weighted*):** Método híbrido en el que las puntuaciones obtenidas por varias técnicas se combinan para producir una sola recomendación.

Tema 9. Resolución de problemas mediante búsqueda

- **Problemas de búsqueda.** Tratan de encontrar el camino que lleve del estado inicial al estado objetivo a través del espacio de estados disponible
- En un problema de búsqueda, esta búsqueda puede hacerse en dos direcciones:
 - Hacia adelante (del inicio al fin).
 - Hacia atrás (del fin al inicio).

- **Búsqueda exhaustiva o a ciegas:** no se tiene en cuenta la posible localización del objetivo. Estos algoritmos no utilizan ninguna información del problema e ignoran hacia dónde se dirigen hasta que encuentran el objetivo.
 - Búsqueda en amplitud
 - Búsqueda en profundidad
 - Búsqueda en profundidad acotada

- **Búsqueda heurística:** se tiene una estimación de lo que falta para llegar al estado objetivo. Los algoritmos utilizan información del problema para localizar el objetivo.
 - Escalada simple o *hill climbing*
 - Escalada por máxima pendiente
 - Búsqueda «mejor el primero» o *best-first*
 - Algoritmo A*

Tema 10. Gestión de la incertidumbre e imprecisión en sistemas expertos

- Una de las aplicaciones más conocidas del razonamiento bayesiano son los clasificadores probabilísticos como naive Bayes.
- Factores de certeza es una alternativa al razonamiento bayesiano
 - El factor de certeza puede tener un valor entre -1 y +1.
-1 indica falsedad total mientras que el valor +1 indica certeza total. Un valor positivo indica cierto grado de creencia en la expresión mientras que un valor negativo indica cierto grado de incredulidad respecto a la veracidad de la expresión.

- **Método de decodificación o *defuzzification*** . último paso, en cualquier sistema de control difuso.
- **Bisector**: establece como salida el punto donde el conjunto difuso es dividido en dos subregiones de igual área. En muchos casos, coincide con el resultado obtenido a través del centroide.



www.unir.net