

Práctica 2 – Ingeniería del conocimiento

Algoritmo ID3

Marcelino Pérez Durán

Árbol de decisión (ID3) ⓘ

TiempoExterior	Temperatura	Humedad	Viento	Jugar
soleado	caluroso	alta	falso	no
soleado	caluroso	alta	verdad	no
nublado	caluroso	alta	falso	si
lluvioso	templado	alta	falso	si
lluvioso	frio	normal	falso	si
lluvioso	frio	normal	verdad	no
nublado	frio	normal	verdad	si
soleado	templado	alta	falso	no
soleado	frio	normal	falso	si
lluvioso	templado	normal	falso	si
soleado	templado	normal	verdad	si
nublado	templado	alta	verdad	si
nublado	caluroso	normal	falso	si
lluvioso	templado	alta	verdad	no

Marcelino Pérez Durán

Opciones

Uso por archivo

Archivo .txt con los atributos

Examinar...

Archivo .txt con los valores

Examinar...

Iniciar

Memoria

- Descripción de la práctica
- Implementación
- Guía de uso

Descripción de la práctica

El algoritmo ID3 se basa en la búsqueda de hipótesis o reglas dado un conjunto de ejemplos (atributos), uno de los cuales será el objetivo y será de tipo binario. De esta forma, el algoritmo trata de obtener las hipótesis que clasifiquen si dicho ejemplo va a ser positivo o negativo.

El algoritmo ID3 realiza lo anterior mediante la construcción de un árbol de decisión, el cual utiliza:

- Nodos: Los cuales contendrán atributos.
- Arcos: Los cuales contienen valores posibles del nodo padre
- Hojas: Nodos que clasifican el ejemplo como positivo o negativo.

TiempoExterior	Temperatura	Humedad	Viento	Jugar
soleado	caluroso	alta	falso	no
soleado	caluroso	alta	verdad	no
nublado	caluroso	alta	falso	si
lluvioso	templado	alta	falso	si
lluvioso	frio	normal	falso	si
lluvioso	frio	normal	verdad	no
nublado	frio	normal	verdad	si
soleado	templado	alta	falso	no
soleado	frio	normal	falso	si
lluvioso	templado	normal	falso	si
soleado	templado	normal	verdad	si
nublado	templado	alta	verdad	si
nublado	caluroso	normal	falso	si
lluvioso	templado	alta	verdad	no

Figura 1: Tabla, con atributos títulos

Implementación

El algoritmo se ha implementado íntegramente en el lenguaje JavaScript, representando los datos de los archivos como una tabla mediante los lenguajes HTML y CSS.

La implementación, ampliaciones y mejoras realizadas son las siguientes:

1. Algoritmo ID3 con los datos proporcionados mediante archivos de texto.
2. Todos los niveles de recursividad

Para entender el correcto funcionamiento de todas las funcionalidades se recomienda leer la “Guía de uso” detallado a continuación.

Guía de uso

El software utilizado ha sido probado en Mozilla Firefox versión 59.0, Safari versión 11.03 y Google Chrome versiones 64.0 y 65 .0, funcionando sin problemas aparentes en las plataformas y versiones probadas.

Para usar la aplicación debemos abrir el archivo index.html, adjuntado en la misma entrega que este documento, haciendo doble click sobre él o arrastrándolo sobre el navegador web. A continuación, se detalla el uso de la aplicación.

Primer contacto

Una vez abierta la aplicación, nos encontraremos con un primer plano en blanco, acompañado de la parte “OPCIONES” en la derecha de la ventana, que veremos en profundidad más adelante. (Véase la figura 2).

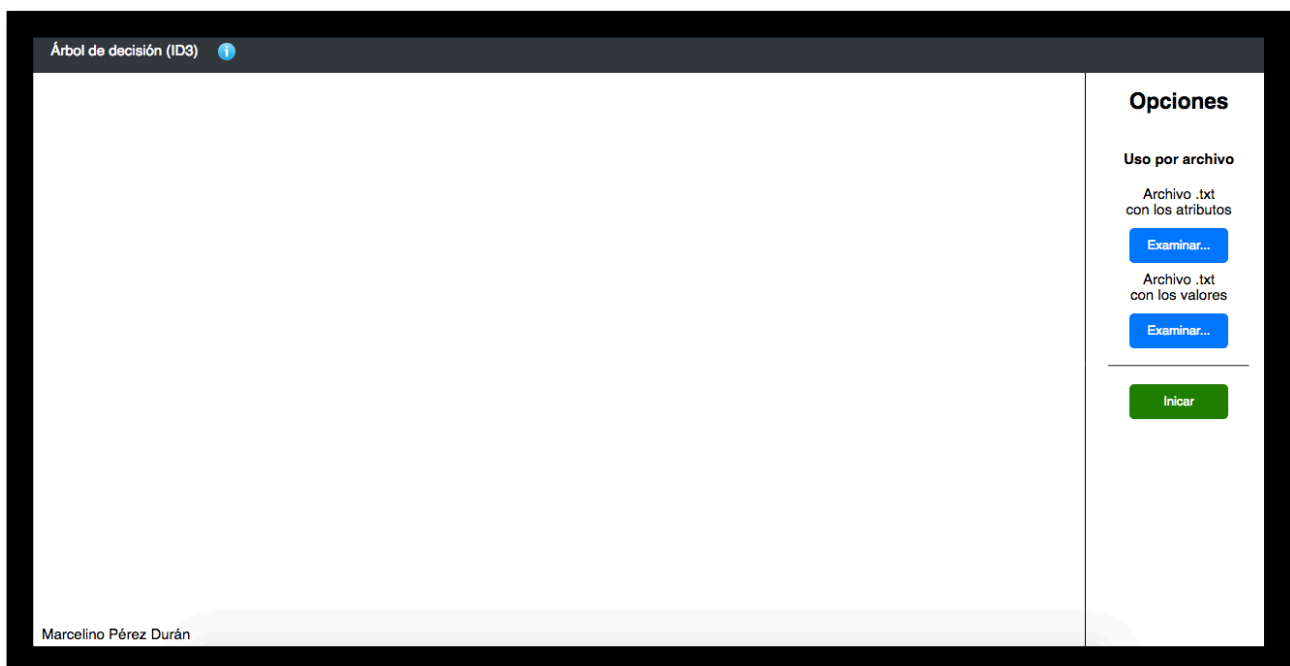


Figura 2: Pantalla inicial

Menú opciones

A la derecha de la pantalla se encuentra el menú “OPCIONES” (Véase figura 3). De arriba a abajo:

- **Archivo .txt con los atributos:** Al pinchar en “Examinar...” tendremos que seleccionar el archivo .txt con los atributos que queramos añadir para realizar el algoritmo. Este archivo debe consistir en una sola línea de texto con el nombre de los encabezados de la tabla separado por comas. Una vez seleccionado un archivo válido se pintarán los atributos en pantalla y el botón “Examinar...” aparecerá con una tonalidad grisácea.
- **Archivo .txt con los valores:** Al pinchar en “Examinar...” tendremos que seleccionar el archivo .txt con los valores que queramos añadir para realizar el algoritmo. Este archivo debe consistir en una o varias líneas de texto, cada línea debe estar separada por comas, con el mismo número de valores que atributos, cada salto de línea representa una nueva fila en la tabla. Una vez seleccionado un archivo válido se pintarán los valores en pantalla y el botón “Examinar...” aparecerá con una tonalidad grisácea.

Opciones

Uso por archivo

Archivo .txt
con los atributos

Examinar...

Archivo .txt
con los valores

Examinar...

Iniciar

Figura 3: Opciones

IMPORTANTE: Antes de añadir los valores debemos añadir los atributos

- **Botón iniciar:** Pulsando sobre el botón iniciar se procederá al cálculo del algoritmo, excepto si no se han añadido los archivos válidos.

Pantalla de resultado

Una vez calculado el algoritmo, se ocultará la tabla con los valores y atributos y aparecerá un cuadrado con el resultado. En primer lugar, aparecerá los méritos de cada atributo, en segundo lugar, el árbol de decisión completo desde la raíz. (Véase la figura 4). Presionando la tecla “esc” volverá a aparecer la tabla.

Resultado:

Pulse "esc" para volver a la tabla

Mérito de atributos

Mérito de TiempoExterior = 0.6935361388961918

Mérito de Temperatura = 0.9110633930116763

Mérito de Humedad = 0.7884504573082896

Mérito de Viento = 0.8921589282623617

Árbol de decisión

-> TiempoExterior -> soleado -> Humedad -> alta -> no

-> TiempoExterior -> soleado -> Humedad -> normal -> si

-> TiempoExterior -> nublado -> si

-> TiempoExterior -> lluvioso -> Viento -> falso -> si

-> TiempoExterior -> lluvioso -> Viento -> verdad -> no

Figura 4: Pantalla de resultado