

Universidad Complutense de Madrid

Ingeniería del conocimiento

12/04/2019



PRÁCTICA 2

ALGORITMO ID3

Paulo J. Colombo

Y3129411-E

ÍNDICE

INTRODUCCIÓN	3
LENGUAJES Y TECNOLOGIAS UTILIZADAS	4
DETALLES RELATIVOS A LA IMPLEMENTACIÓN	5
AMPLIACIONES REALIZADAS	7
MANUAL DE USUARIO	8

INTRODUCCIÓN

El ID3 es un algoritmo empleado dentro del campo de la inteligencia artificial que busca construir y definir recursivamente modelos de clasificación en forma de estructura de árbol basado en reglas y dado un conjunto de ejemplos; por medio de un procedimiento de arriba y abajo haciendo un recorrido voraz por el espacio de posibles soluciones. El resultado final del algoritmo será un árbol decisión compuesto por nodos decisión y nodos hojas.

El presente trabajo describe una implementación del algoritmo ID3, comentando las tecnologías y lenguajes de programación utilizados, detalles relativos a la propia implementación y ofrece, además, un manual de uso.

LENGUAJES Y TECNOLOGIAS UTILIZADAS

Se ha utilizado HTML5, CSS3 y la librería Bootstrap para diseñar y crear la vista principal de la aplicación, esto debido a que se ha considerado que estos lenguajes de marcado facilitan el desarrollo de la UI y permiten una mejor interacción del usuario con la aplicación.

Para implementar la lógica de la aplicación se ha utilizado JavaScript ES6 en conjunto con jQuery que ofrece una interfaz más sencilla a la hora de manipular el DOM.

Se ha utilizado la herramienta de agrupación de recursos web Webpack para empaquetar todos los ficheros escritos en uno solo y permitir la modularización del programa.

Por último, se ha utilizado la herramienta Mermaid.js para generar los diagramas de los árboles de decisión que se obtienen como resultado de ejecutar el algoritmo.

DETALLES RELATIVOS A LA IMPLEMENTACIÓN

El código fuente de la aplicación se ha organizado en los siguientes ficheros:

1. Modelo:

- a. **ID3.js**: Este fichero JavaScript contiene la lógica principal recursiva del algoritmo. Define una clase ID3 que posee, como uno de sus atributos más importantes, el árbol decisión que se va construyendo a lo largo de cada iteración recursiva. Posee dos métodos importantes: run y execute; el primero llama al segundo método que es el que contiene la lógica recursiva.

2. Estructura de datos:

- a. **Tree.js**: Este fichero posee la clase Tree que posee un atributo que contiene la raíz de este y una colección con todos los hijos posibles de la raíz.
- b. **Node.js**: Este fichero JavaScript define la clase Node que constituye cada uno de los nodos del árbol de decisión del algoritmo. Contiene tanto la lista de atributos como la tabla de datos, así que como el nombre del nodo y el nombre de la rama de la que proviene. Por último, posee un objeto InfoNode información relevante para el cálculo de los nodos hijos en el ID3.
- c. **InfoNode.js**: Este fichero JavaScript define la clase InfoNode que posee el nombre del atributo que representa el nodo, el número total de filas de la tabla de datos en la que aparece el atributo, una colección de los distintos valores que puede tomar ese atributo e información relativa a cada valor y el valor del mérito calculado para dicho atributo.
- d. **InfoValue.js**: Este fichero define la clase InfoValue que proporciona información relativa a un valor determinado de un atributo, como por ejemplo: el número de filas positivas y negativas que contienen el valor, el porcentaje de estas, etc.

3. Vista:

- a. **Index.html**: Este fichero html posee la estructura principal de la aplicación, incluyendo los campos a los que se pueden subir los ficheros con la lista de atributos y la tabla de datos, y los paneles donde se muestra el árbol de decisión y las reglas lógicas.
 - b. **Style.css**: Contiene parte del estilo de la aplicación. El resto del estilo lo aporta Bootstrap.
4. Controlador:
- a. **Index.js**: Contiene el código jQuery encargado de comunicar la vista con el modelo. Además, posee todos los event listeners relacionados con los campos del formulario y el botón “Run”.

AMPLIACIONES REALIZADAS

Se ha hecho el algoritmo ID3 completamente recursivo, pudiendo generar el árbol de decisión completo.

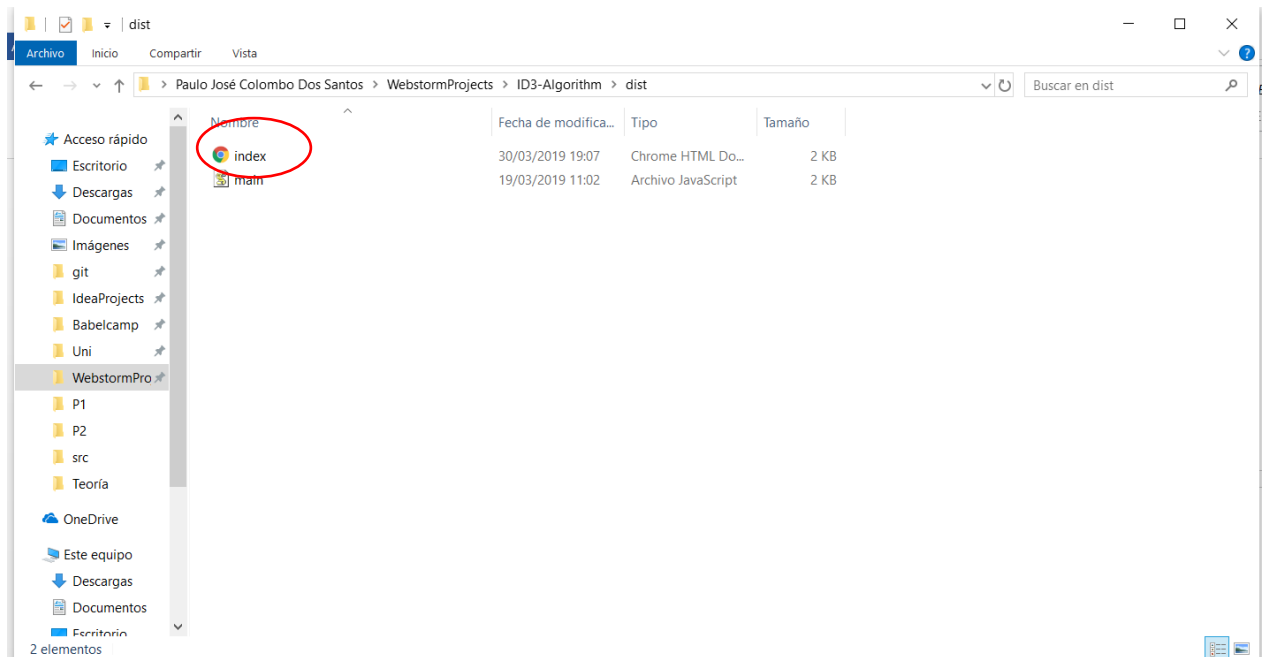
El algoritmo también se ha programado de tal forma que admita cualquier tipo de atributos y tablas de datos, generalizándolo para cualquier posible caso.

También se ha agregado, como ya se ha comentado anteriormente, un diagrama grafico del árbol de decisión con la finalidad de mostrar los resultados al usuario de una forma más descriptiva. Aparte de esto, se realiza la conversión del árbol en reglas lógicas con su respectivo antecedente consecuente.

MANUAL DE USUARIO

Para poder ejecutar esta aplicación, se necesita un navegador web, por ejemplo: Google Chrome, Mozilla Firefox o Safari.

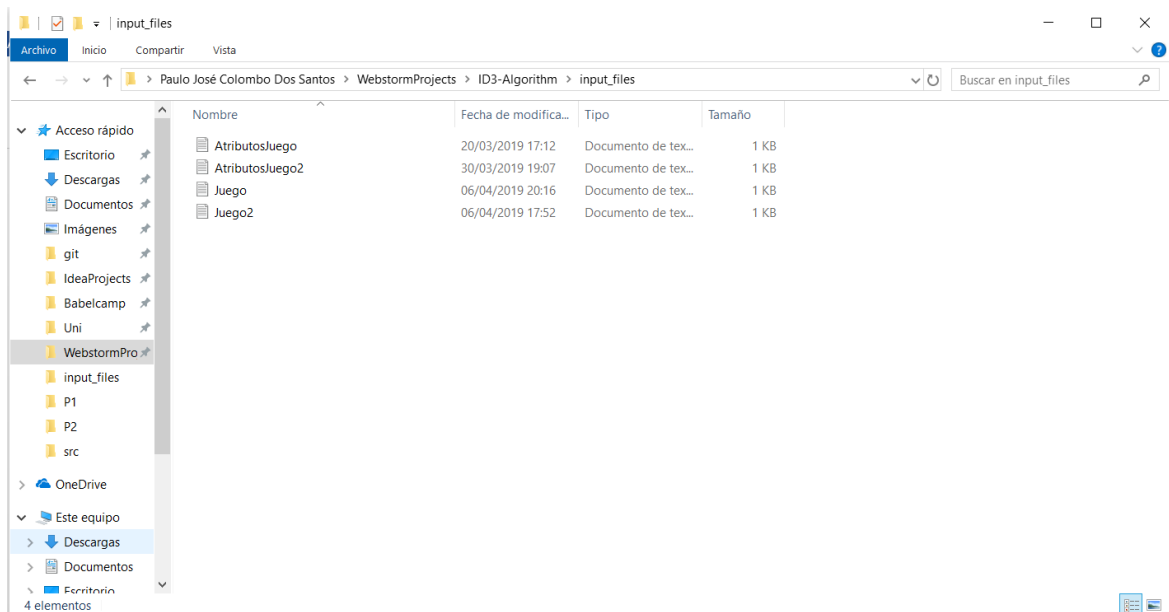
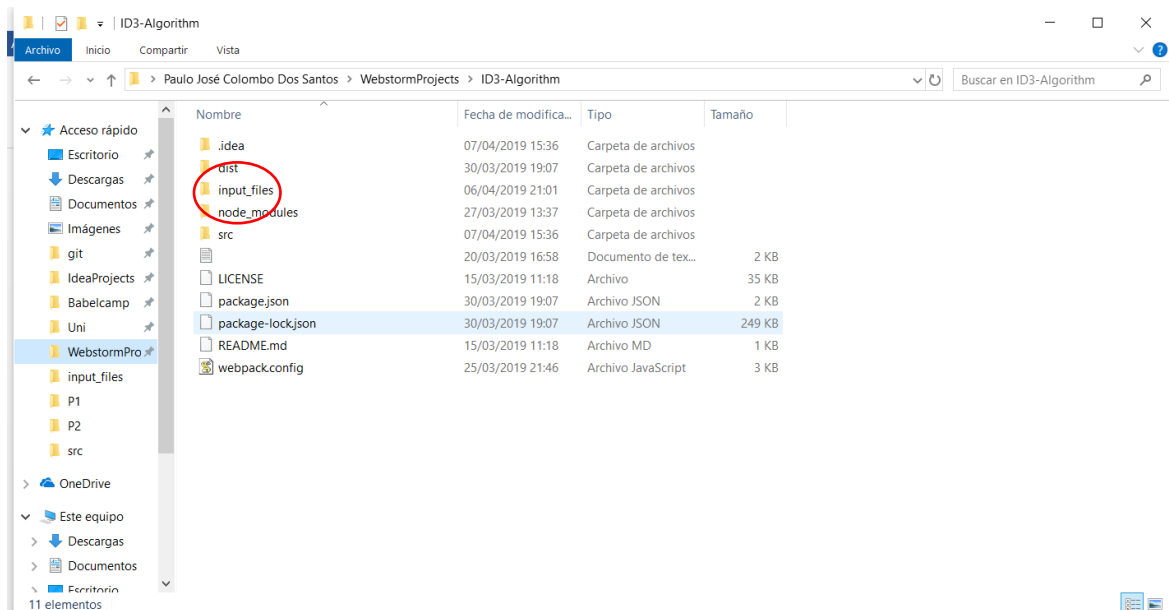
Para poder ejecutar la aplicación hay que hacer doble clic sobre el fichero html llamado index, que se encuentra dentro del directorio dist.



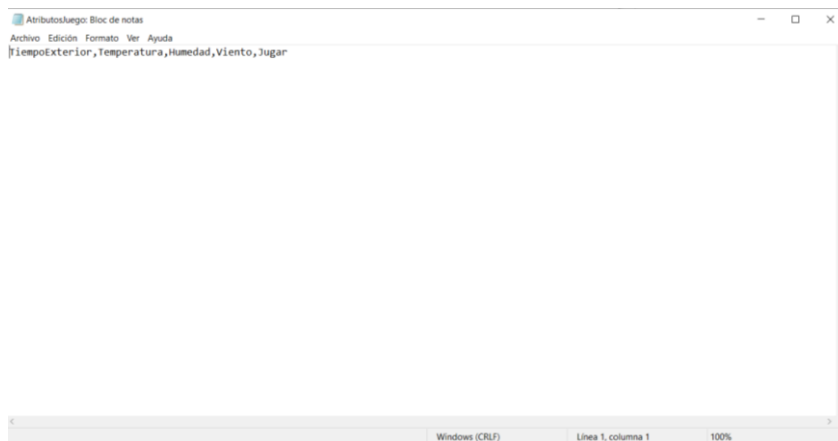
Una vez ejecutada la aplicación, se mostrara un panel con dos campos donde se pueden subir los ficheros con la lista de atributos y la tabla de datos



En el directorio `input_files` dentro de la carpeta de la aplicación se encuentran dos ejemplos que se pueden usar como datos de entrada.

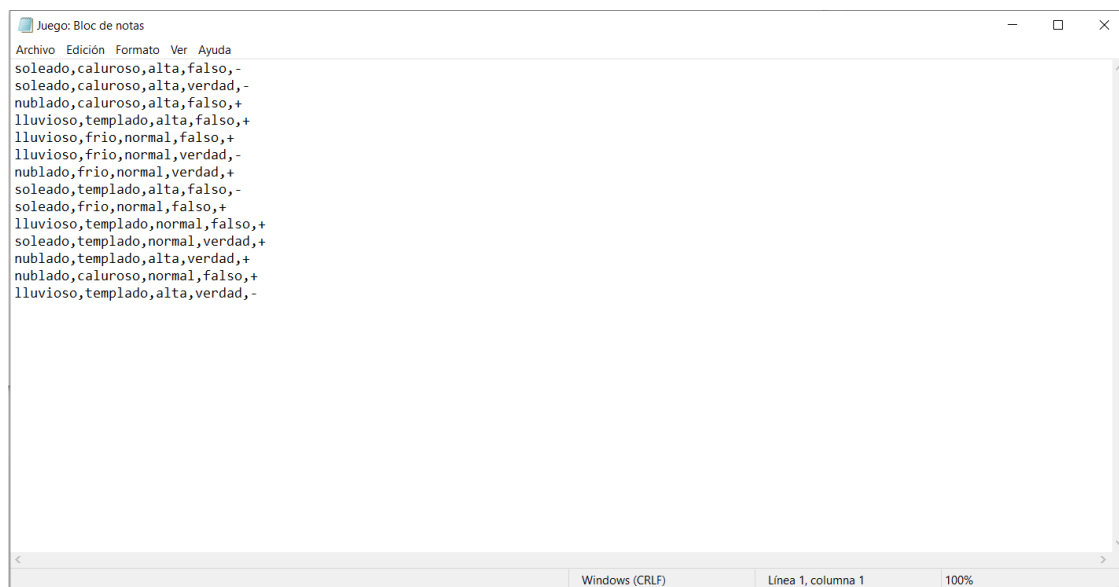


El fichero de atributos debe contener todos los atributos en una única línea separados por coma y sin espacios. Un ejemplo sería el siguiente:



El fichero con la tabla de datos debería contener una línea por fila donde los valores para cada atributo se separan mediante comas y cada fila se representa como positiva o negativa utilizando el carácter “+” o “-”. No obstante, si introducimos datos erróneos o en el formato incorrecto, la aplicación nos lo hará saber por medio de una ventana de texto emergente.

Un ejemplo sería el siguiente:



Véase como todas las filas terminan en + o -.

Una vez hayamos subido ambos ficheros obtendremos algo parecido a esto:

ID3 Algorithm

UPLOAD ATTRIBUTE FILE

Seleccionar archivo

AttributosJuego.txt

UPLOAD DATA FILE

Seleccionar archivo

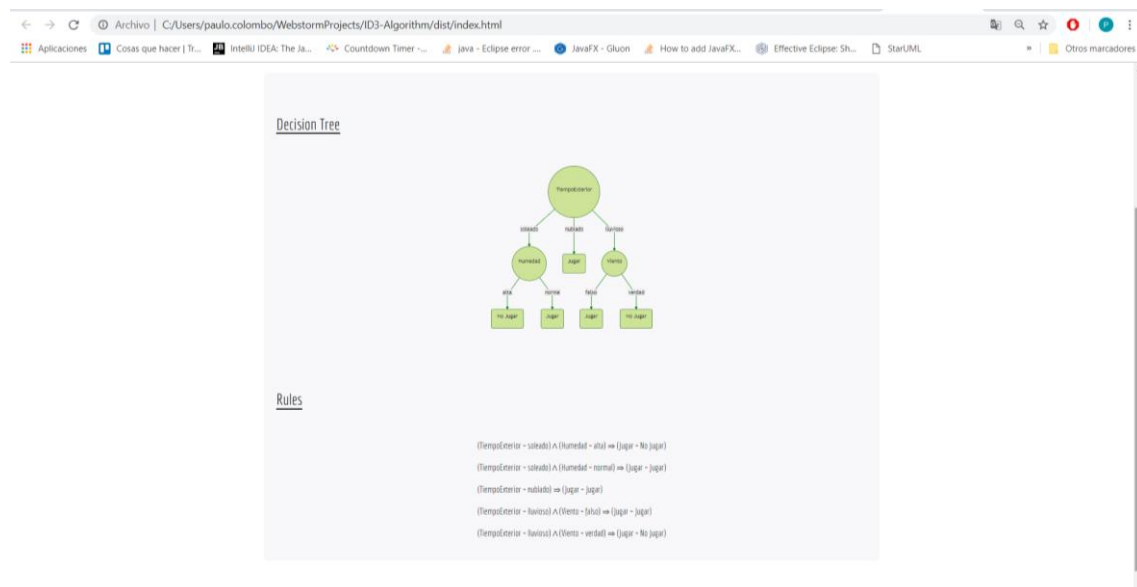
Juego.txt

Run

TempoExterior	Temperatura	Humedad	Viento	Jugar
soleado	cálida	alta	falso	-
soleado	cálida	alta	verdad	-
nublado	cálida	alta	falso	+
lluvioso	templado	alta	falso	+
lluvioso	frío	normal	falso	+
lluvioso	frío	normal	verdad	-
nublado	frío	normal	verdad	+
soleado	templado	alta	falso	-
soleado	frío	normal	falso	+
lluvioso	templado	normal	falso	+
soleado	templado	normal	verdad	+
nublado	templado	alta	verdad	+
nublado	cálida	normal	falso	-
lluvioso	templado	alta	verdad	-

Aparecerá una tabla con los datos que hayamos subido.

Al presionar el botón “Run” obtendremos el árbol de decisión y el conjunto de reglas lógicas:



Si queremos introducir nuevos datos de entrada podemos simplemente añadir nuevos ficheros a los campos del formulario y la aplicación automáticamente se reiniciará.