

MEMORIA

Practica 2: Implementación del algoritmo ID3

Autor: Luis Jaramillo Pulido

INDICE

1. Introducción.....	3
2. Diseño.....	4
3. Representación del Árbol.....	6

PRESENTACIÓN DE LA PRÁCTICA

1.1 INTRODUCCIÓN

El objetivo en esta práctica es la implementación del algoritmo ID3, el lenguaje utilizado es java.

La aplicación se ha implementado usando el IDE eclipse, es una aplicación que interactúa con el usuario por consola, se podía haber hecho una implementación usando una interfaz gráfica pero debido a factores de tiempo y también a que fui incapaz de poder dibujar el árbol gráficamente, al final decidí hacer uso de la consola.

La aplicación empezará leyendo 2 ficheros, "AtributosJuego.txt" y "Juego.txt", y el número de filas de la tabla, todos esos datos lo indicará el usuario de la aplicación.

Una vez hecho esto, se mostrará la tabla, y se ejecutará el algoritmo ID3, he supuesto que el atributo de interés/clase (Jugar), sea la última columna de la tabla.

El comportamiento de la aplicación será la siguiente:

Empezará calculando la entropía de cada ejemplo/variables del atributo, para posteriormente calcular el mérito del atributo.

Así se hará con el resto de los atributos, hasta seleccionar el atributo con el menor mérito.

Después se crearán las ramas que corresponde a las variables del atributo, estas ramas apuntarán a una subtabla, si todos los valores del atributo de interés/clase(Jugar) de la subtabla tienen el mismo valor será una hoja con valor (Si/No) en caso contrario se irán haciendo llamadas recursivas repitiendo el mismo procedimiento del cálculo de la entropía y merito, con la salvedad que la subtabla construida solo contendrá las filas que contengan las variables de esa rama.

El algoritmo acaba hasta que todos los subárboles alcancen una hoja que indica el valor de la clase.

DISEÑO

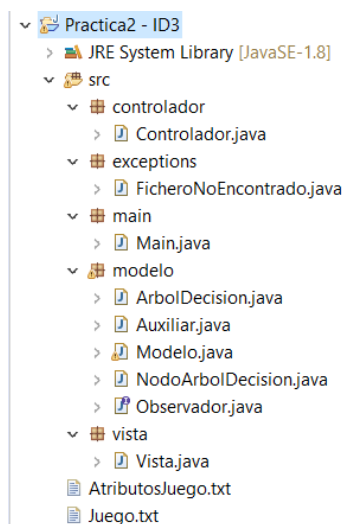
El patrón de diseño aplicado es el modelo vista y controlador al modo de simplificar el diseño y hacerlo lo más simple posible.

En el paquete Modelo contiene la implementación del algoritmo, la clase ArbolDecision es la clase principal del algoritmo que hacen las llamadas a los métodos para el cálculo de la entropía y mérito.

El método correrAlgoritmo es donde se ejecuta todo el algoritmo, se obtienen la entropía general, la entropía para cada atributo, se hace la selección del merito (mínimo) del atributo, se obtienen las variables del atributo seleccionado, se generan sus subtablas y se hacen las llamadas recursivamente hasta alcanzar una hoja que indica el valor de la clase.

Patrón de diseño utilizado Modelo Vista y Controlador







La aplicación que implementa el algoritmo ID3 consta de 5 paquetes



Modelo



Está compuesto por 5 clases: ArbolDecision, Auxiliar, Modelo, NodoArbolDecision, y la clase interfaz Observador.

La clase Modelo es donde se notificará a la vista el resultado de todo el proceso del algoritmo haciendo uso del método “notificaResultado”, en el método ejecuta es donde se crea el objeto árbolDecision para posteriormente notificar el resultado a los observadores (en esta aplicación solo tengo un observador que es la vista)

- ▼  modelo
 - >  ArbolDecision.java
 - >  Auxiliar.java
 - >  Modelo.java
 - >  NodoArbolDecision.java
 - >  Observador.java



Controlador

Compuesto por una sola clase, que es la encargada de interactuar con la vista y el modelo.

- ▼  controlador
 - >  Controlador.java

Vista

Compuesta por una sola clase, aquí es donde recoge los datos del usuario.

- ▼  vista
 - >  Vista.java

Representación del Árbol

