Manual d'usuari del generador de regles d'associació



Carles Capilla Canovas Albert Coma Coma Pere Grau Molina Muhammad Meraj Arshad

Índex de continguts

Funcionalitats de l'aplicació	2
Vista principal	2
Menú de navegació	2
Inici	3
Arxiu	3
Exportar Registres	3
Exportar Regles	4
Preprocés de dades	5
Carregar arxiu de dades	5
Selecció fitxer	5
Generar preprocés de dades	7
Generar preprocés d'un atribut numèric	7
Generar preprocés d'un atribut binari	10
Generar preprocés d'un atribut categòric	12
Finalitzar preprocés	14
Consultar registres	15
Gestió de Regles	17
Generar Regles	17
Validar Regles	20
Consultar Regles	22
Consultar per suport, confiança i nombre de condicions	23
Consultar per condició	25
Consultar per consequent	27
Sortir	28

Funcionalitats de l'aplicació

Aquesta aplicació ens permet obtenir regles d'associació a partir d'un conjunt de dades específic. Abans de generar regles, les dades passen per un preprocés, el qual se li pregunta a l'usuari interactivament per posteriorment aplicar-lo a totes les dades introduïdes.

El sistema et permet consultar tant els registres de dades processats com les regles d'associació generades, amb opció d'aplicar diferents filtres a la consulta. Un cop tenim unes regles generades, el programa ens permet validar un nou conjunt de dades amb el mateix preprocés i comprovar els nous paràmetres obtinguts amb aquestes dades.

A més d'això, s'inclouen funcionalitats per importar i exportar tant els conjunt de registres com els conjunts de regles, per una millor portabilitat i utilitat del programa.

A continuació s'explica cada una de les funcionalitats del programa en detall.

Vista principal

Dins la pantalla inicial del programa podem observar les funcionalitats principals del nostre programa a partir de les quals podem iniciar una generació de regles, validar-ne un conjunt o bé sortir definitivament d'aquest:

- Importar conjunt de regles
- Importar conjunt de registres
- Carregar fitxer de dades
- Sortir

A més disposem d'un menú de navegació a partir del qual accedir a totes les funcionalitats del nostre programa.

Menú de navegació

Inici Arxiu Preprocés de dades Gestió de regles

En aquest menú es poden trobar totes les funcionalitats que té el sistema. Aquestes funcionalitats estan repartides en diferents blocs del menú. Aquests blocs són: *Inici, Arxiu, Preprocés de dades* i *Gestió de regles*. A continuació, s'explicarà més detalladament el contingut de cada un d'aquests components.

Inici

El menú d'Inici ens permet accedir des de qualsevol de les altres vistes on s'inicïi qualsevol de les tasques principals del programa a la vista principal d'aquest. Per fer-ho, primer hem d'anar a Inici, després donar clic a *Pàgina d'Inici*.



Arxiu

Amb aquesta funcionalitat del menú podem accedir a la importació i exportació de regles de les quals se'n detallarà el procés a continuació.



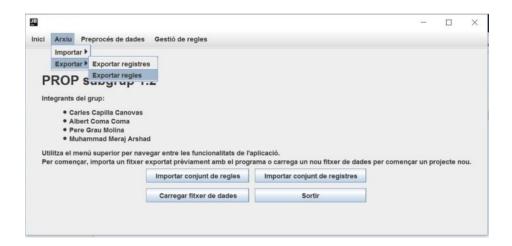
Exportar Registres

Per a poder exportar els registres, primer ha d'haver un conjunt de dades carregat al sistema i s'ha d'haver definit un preprocés, per tal que aquestes dades hagin pogut estar processades.

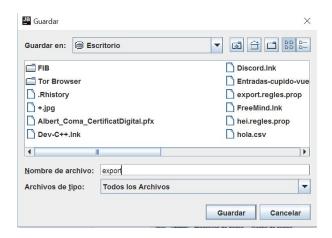
Un cop complert aquest requisit, utilitzant el menú d'exportar registres, es podrà escollir un directori on guardar l'exportació del sistema.

Exportar Regles

Per tal d'importar un arxiu de registres prèviament exportat del sistema, es pot utilitzar aquest item del menú, que demanarà un fitxer de dades i, en cas que aquest sigui vàlid, els registres s'importaran al sistema.



Un cop dins, cal clicar al botó "Selecciona..." i triar la carpeta a on volem exportar les regles tot introduint un nom al fitxer.



Si cliquem el botó "Exporta" s'exporten les regles amb un arxiu .prop i es torna al menú principal.

Si es volen exportar regles però no es selecciona cap carpeta saltarà aquest error:



Preprocés de dades



Carregar arxiu de dades

En aquest bloc tenim la possibilitat de carregar un fitxer de dades en format .csv a partir del qual generarem un preprocés i l'aplicarem per tal de poder discretitzar les seves dades. Això ens permet no haver d'introduir un a un cadascun dels registres que ens interessin per posteriorment obtenir les nostres regles d'associació, és a dir, cada conjunt d'atributs amb el seu valor corresponent.

El format d'un arxiu *csv* és obert i ens permet representar els valors en format d'una taula en la qual les columnes es separen per comes, o punts i coma si consideréssim utilitzar la coma per indicar els decimals. I on cada fila se separa de la següent per un salt de línea.

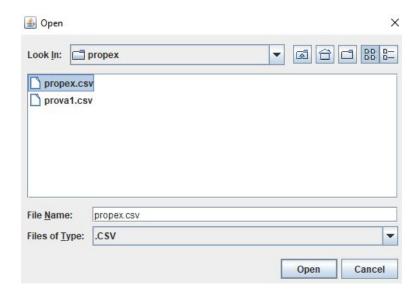
Selecció fitxer

Per tal de poder visualitzar el conjunt de directoris dins la nostra màquina, clicarem en el botó de "Seleccionar...".



Un cop tinguem disponible la vista que ens permeti navegar pels nostres directoris, escollirem el directori on tinguem emmagatzemat el nostre fitxer de dades. Un cop seleccionat el directori, podrem veure quins dels arxius del directori compleixen les condicions per poder ser oberts des del nostre sistema. Aquests arxius seran els que disposin del format .csv, el qual ha estat definit en l'enunciat del projecte.

Un cop decidit quin fitxer utilitzarem, el seleccionem i l'obrim amb la funció "Open" o bé i fem doble click.



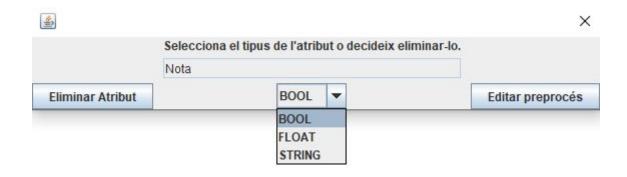
Quan ja disposem del fitxer seleccionat, hi apareixerà el seu path, i si considerem que és el correcte polsarem "Confirmar Selecció del fitxer" per tal de començar a introduir els factors que definiran el seu preprocés atribut a atribut.



Generar preprocés de dades

Un cop obert el fitxer al que vulguem aplicar un preprocés, se'ns presentaran un a un tots els atributs que el componen per tal de poder definir el preprocés a generar. Per cada atribut tindrem dues opcions:

- Eliminar l'atribut ja que per les regles que intentem obtenir posteriorment no ens pot arribar a interessar
- Escollir el tipus de l'atribut per posteriorment editar el seu preprocés en funció d'aquest.



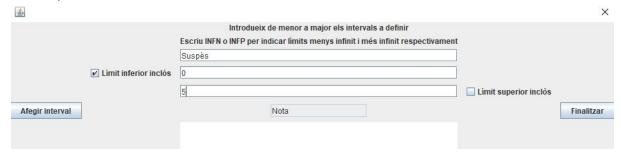
Generar preprocés d'un atribut numèric



En el cas d'escollir que el tipus de l'atribut amb el nom presentat sigui numèric, se'ns presenta la següent vista en la que podrem anar afegint un a un tots els intervals per tal de discretitzar posteriorment els valors de l'arxiu en aquests intervals.

Els paràmetres que podem indicar son:

- El nom de l'interval a generar
- Els límits inferiors i superiors
- La inclusió en l'interval dels seus límits, és a dir, definir si l'interval és obert o tancat pels seus dos extrems.

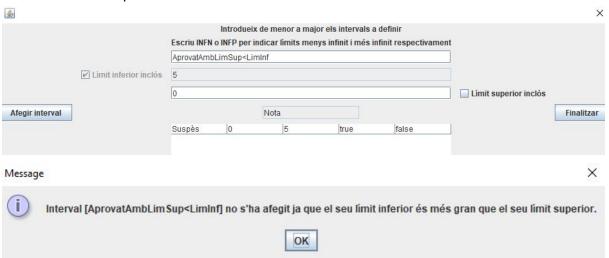


En el cas de definir tots els paràmetres correctament sense que aquest interval intervingui en la creació dels altres, ens apareixerà el següent missatge i apareixeria en la taula inferior com a interval creat amb els seus límits inferior i superior i l'obertura d'aquests.

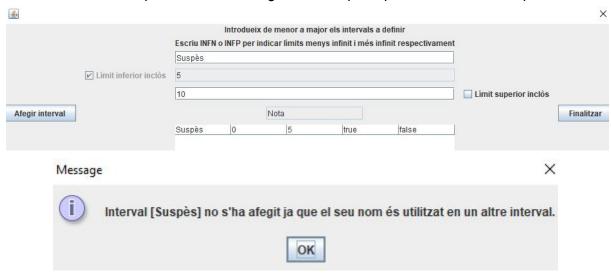


En canvi, en cas que introduïm algun paràmetre incompatible amb la pròpia creació de l'interval ens apareixen diferents missatges segons aquest error a més de no crear-se aquest interval.

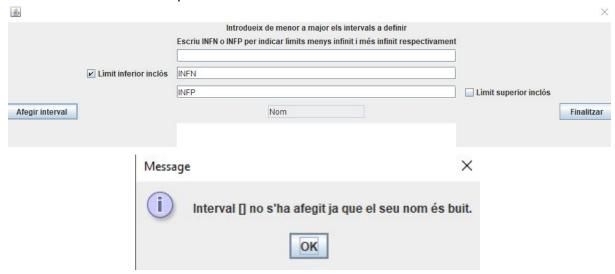
- En el cas que introduïm un límit inferior > límit superior, com veiem a la captura següent, obtindrem l'error que es presenta a continuació que ens indica l'error en aquesta definició de límits.



- En cas que introduïm per nom, el nom d'un interval ja creat per a aquell atribut en concret, ens apareixerà el missatge indicant que aquest nom no està disponible.



- I finalment en cas que afegim un interval sense cap nom o el camp destinat a introduir el nom buit, ens saltarà un missatge d'error indicant que el camp que dóna nom a l'interval no pot ser buit.



Generar preprocés d'un atribut binari



En el cas de voler especificar el preprocés d'un atribut de tipus binari, és a dir, amb dos possibles valors en tot el fitxer, se'ns presentarà la següent vista on definirem dos valors diferents a poder prendre.

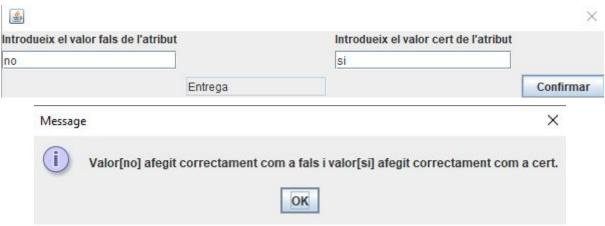
- En el cas de definir dos valors possibles iguals, ens apareixerà el missatge d'error que podem observar a continuació indicant que aquests valors no poden ser iguals.



 En el cas de no definir un dels dos possibles valors a prendre dins el fitxer o bé cap dels dos valors, com podem veure a la següent imatge, ens apareixerà el següent missatge d'error per indicar que hem d'introduir els dos valors per tal de poder binaritzar els valors de l'atribut.



- En canvi, si declarem dos valors diferents ens apareixerà el missatge de confirmació que podem veure a continuació, després de la definició, indicant que s'ha afegit el preprocés d'aquell atribut amb els valors indicats.

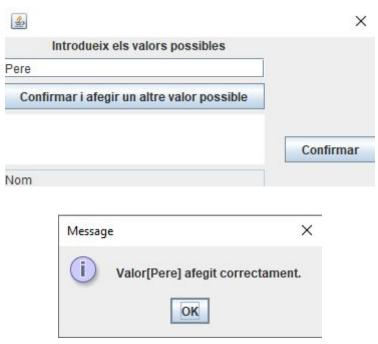


Generar preprocés d'un atribut categòric

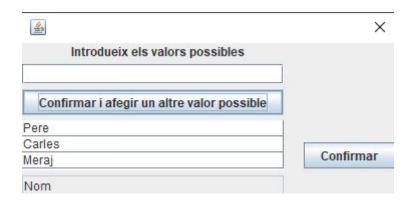


En el cas que vulguem definir un atribut com a categòric se'ns presenta la següent vista en la qual podrem anar afegint un a un els valors possibles o categories que pugui tenir l'atribut amb el nom indicat a sota dins el nostre fitxer.

Un cop haguem escrit el valor que pot prendre farem clic a "Confirmar i afegir un altre valor possible" i si s'afegeix correctament ens apareixerà el missatge de confirmació que tenim sota la vista.



Un cop afegits tots els valors possibles, els quals ens aniran apareixent conforme els anem afegint a la taula de la nostra vista com a l'exemple que tenim just a sota, farem clic en "Confirmar" fet que afagirà tots els valors i ens portarà a definir el preprocés d'un altre atribut si en quedés cap.



Si per error no omplíssim el camp de valor possible o bé intentéssim introduir un valor possible buit ens saltaria el missatge d'error que tenim a sota indicant que no podem afegir un valor buit com a possible dins el nostre atribut.



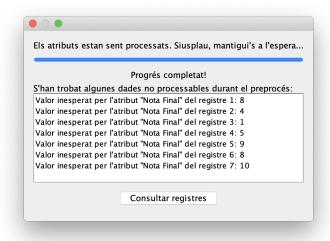
Finalitzar preprocés

Un cop hem acabat d'introduir tots els valors, es demana a l'usuari quina acció vol seguir en cas de trobar un valor d'atribut que no correspongui a cap valor del preprocés. Hi ha dues opcions disponibles:



- Eliminar el registre: en el cas que el valor no es trobi dins de cap valor introduït al preprocés, el registre complet serà eliminat i no es tindrà en compte per la resta del preprocés.
- Avortar el preprocés: el preprocés serà cancel·lat en cas de trobar un valor inesperat.

A la següent imatge podem veure un exemple on hem optat per la primera opció i ens mostra les col·lisions de les dades trobades:



Un cop generat tot el preprocés de les dades aquest s'aplicarà i se'ns dirigirà automàticament a la vista de consulta de registres.

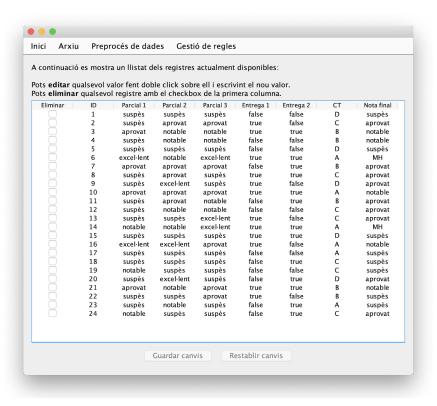
Consultar registres

Per tal de consultar els registres, ens podem adreçar a l'apartat de **Preprocés de dades > Consulta de registres** del menú superior.

Un cop s'ha accedit correctament, se'ns presenta una vista per defecte on ens indica que no hi ha cap registre al sistema. Per a poder carregar dades, es pot carregar un nou arxiu de dades i definir el seu preprocés o importar un conjunt de registres que prèviament hagi estat exportat.

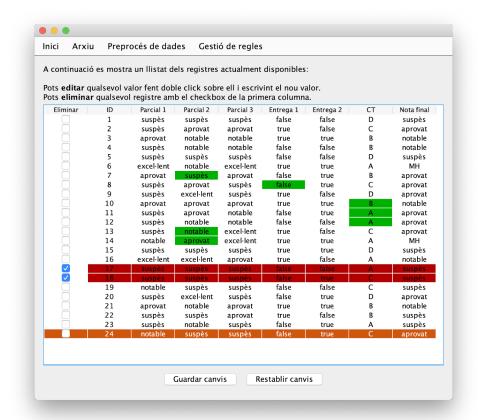


Un cop realitzem una d'aquestes dues opcions i tinguem un conjunt de registres carregat al sistema, el programa ens presentarà una taula amb tot el llistat d'atributs i els seus valors per a cada un dels registres:

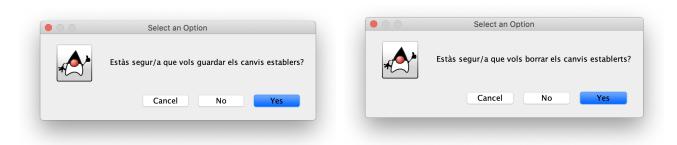


Com podem veure, aquesta taula és perfectament editable. Fent doble click sobre qualsevol valor, podem editar-lo al nostre gust. Amb els *checkboxes* de la columna "eliminar", podem marcar un registre com a eliminat.

Un cop la taula es modifica, els botons inferiors s'activen i ens permeten guardar els canvis al sistema o restablir i borrar totes les modificacions de la taula:

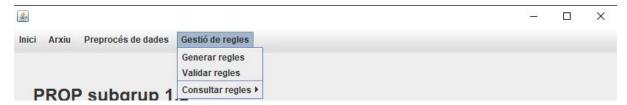


Si decidim guardar els canvis o borrar-los, ens sortirà un missatge confirmant la nostra elecció:



NOTA: si els canvis no són guardats, aquests no seran aplicats al sistema i no es tindran en compte a l'hora de generar regles.

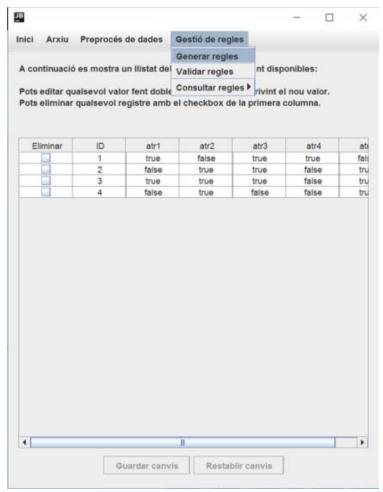
Gestió de Regles



En Gestió de Regles podem trobar totes les funcionalitats relacionades amb les regles d'associació. Concretament, podem generar les regles d'associacions, validar regles i inclus podem consultar regles de tres diferentes formes.

Generar Regles

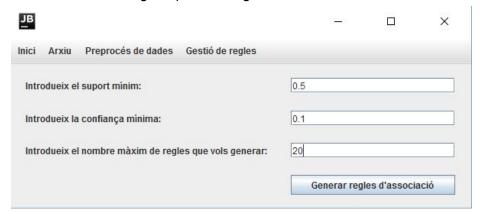
Un cop tenim creats els registres ja podem generar les regles d'associació. Per generar-les s'ha d'accedir al Menú de navegació i dins la pestanya "Gestió de regles" seleccionar l'opció de "Generar regles".



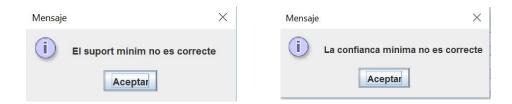
Un cop dins, haurem d'omplir tres camps:

Format decimal amb punt (exemple: 0.2).

- El mínim suport que han de complir les regles (en tant per 1, 1>x>0).
- La mínima confiança que han de complir les regles (en tant per 1, 1>x>0).
- El nombre màxim de regles que volem generar.



Si s'introdueix un valor incorrecta a al suport i confinaça salten aquests errors:



Si s'introdueix un nombre negatiu al nombre màxim de regles a generar salta aquest error:

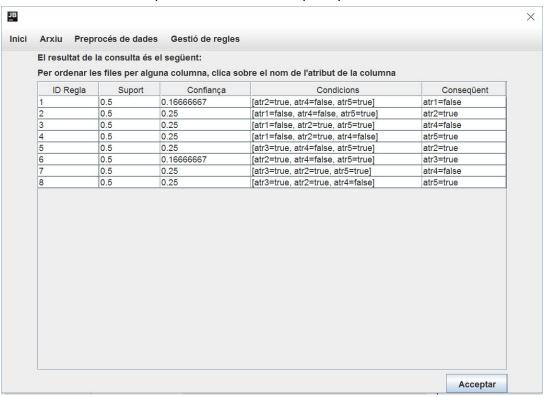


En el cas que no introduíssim algun dels camps anteriors saltaria un error.



Un cop cliquem el botó "Generar regles d'associació" es generen les regles. Aquestes regles apareixen a una taula amb el seu identificador, suport, confiança, atributs condicions i l'atribut conseqüent.

Si es clica el botó "Acceptar" es torna al menú principal.



Validar Regles

Un cop creades les regles d'associació, es poden validar aquestes regles amb un altre fitxer de dades. Per fer-ho, s'ha de tornar a accedir al Menú de navegació a la pestanya "Gestió de regles", allà hem de clicar al botó de "Validar regles".



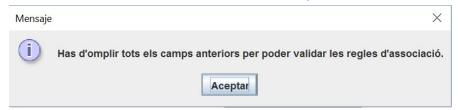
Un cop dins haurem d'omplir dos camps:

- El mínim suport que vols fer servir per validar les regles.
- La mínima confiança que vols fer servir per validar les regles.

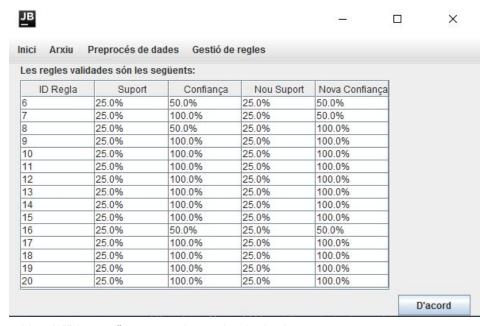
També s'haurà d'importar el nou fitxer .csv amb el que es vol validar les regles d'associació. Aquest fitxer s'importa igual que quan carreguem noves dades clicant al botó "Seleccionar...".



En el cas que no omplíssim el nou mínim suport o confiança saltaria un error.



Si es clica al botó de "Validar regles" es validarien les regles d'associació prèviament generades. Les regles validades apareixen a una taula amb el seu identificador suport, confiança, nou suport obtingut amb les noves dades i la nova confiança obtinguda amb les noves dades.



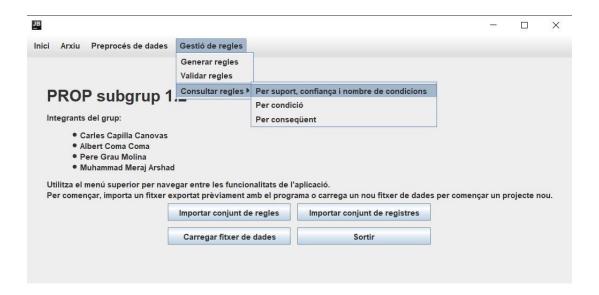
Si es clica el botó "D'acord" tornem al menú principal.

Consultar Regles

En aquest apartat es pot trobar tres diferents maneres de consultar les regles d'associacions. Aquestes regles només es podran consultar si abans s'han generat o importat regles. L'usuari pot escollir l'opció en la que l'interessa consultar les regles.

Aquestes opciones són les següents:

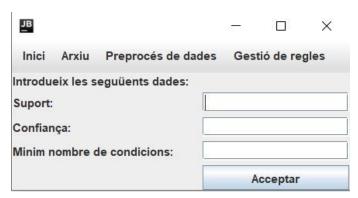
- Consultar per mínim suport, mínima confiança i mínim nombre de condicions.
- Consultar per condició
- Consultar per consequent



A la imatge es pot veure el menú de navegació on es selecciona consultar per suport, confiança i nombre de condicions.

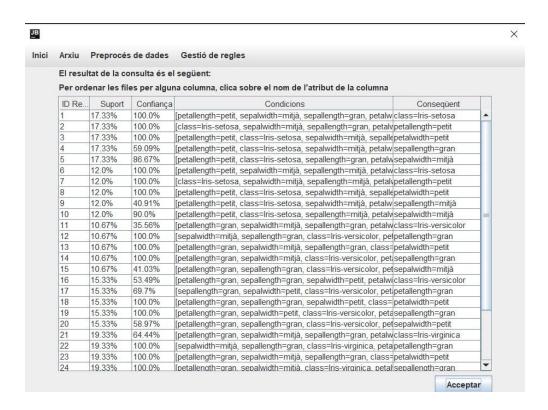
Consultar per suport, confiança i nombre de condicions

Aquí es pot consultar regles d'associació amb un suport i confiança mínima i amb mínim número de condicions que haurien de tenir les regles. Per fer aquesta consulta s'ha de navegar en el menú, buscar aquesta opció i fer clic (com es pot veure a la imatge anterior). A continuació, sortirà una nova vista, com la següent:



En aquesta vista, s'ha d'introduir el suport mínim, un valor entre 0 i 1, després la confiança mínima, un valor entre 0 i 1 i finalment el mínim nombre de condicions, un enter més gran que 0. Un cop introduïdes les dades correctes, s'ha clicar el botó "Acceptar".

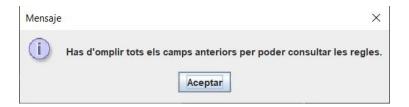
A continuació, si les dades són correctes i el sistema té regles d'associació que compleixen les dades aleshores es podran observar les regles. Un exemple seria el següent:



En aquesta imatge es pot veure les regles d'associació amb tots els seus atributs. Un cop consultades les regles es pot tornar a consultar o fer clic a Acceptar per utilitzar una altra funcionalitat.

Si durant l'entrada del suport, de la confiança o del nombre de les condicions hi hagut un error o s'han introduït incorrectament les dades, llavors el sistema mostrarà un missatge a la pantalla. Així l'usuari podrà corregir i tornar a consultar les regles.

Per exemple, si s'ha deixat algún camp buit aleshores el programa mostrarà el següent missatge:



Després si el valor del suport i de la confiança no està dins de l'interval [0,1] aleshores el sistema retorna un missatge com el següent:



Finalment, si el valor del nombre de condicions no és un enter més gran que zero, llavors el sistema retorna el següent missatge:



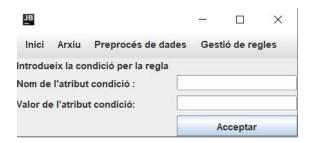
Si no ens interessa tenir en compte el nombre de condicions per la consulta, es pot posar un 1, ja que tota regla té com a mínim una condició. D'aquesta manera, podem seleccionar totes les regles amb aquest número.

Consultar per condició

Per consultar regles per condició s'ha de navegar pel menú i fer clic a "Consultar regles > Per condició", tal com indica la següent imatge.



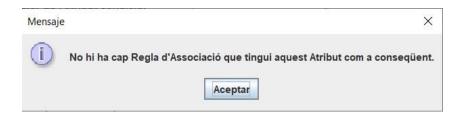
Amb això s'obre una nova vista on el sistema demana el nom i el valor de l'atribut condició.



En aquest cas, s'ha d'introduir un nom i valor d'un atribut que a la generació de regles hagi quedat com a condició d'una o més regles. Per tant, després d'introduir apareixen totes aquelles regles d'associació que tenen aquest nom i valor d'atribut com a condició. En el resultat es podrà veure tots els atributs de la regla. En cas contrari, el sistema informarà amb un missatge que no existeixen regles d'associació. Per exemple, volem consultar per condició com esport=true. Introduïm les dades:



A continuació, li donem clic a acceptar. Si existeix alguna regla amb aquesta condició, es veurà una taula amb tots els atributs. A continuació podem veure el missatge que se'ns mostraria si el sistema no tingués cap regla amb aquesta condició.



Consultar per consequent

Per consultar regles per consequent s'ha de navegar pel menú i fer clic a "Consultar regles>per consequent".



Amb això s'obre una nova vista on el sistema demana el nom i el valor de l'atribut conseqüent.

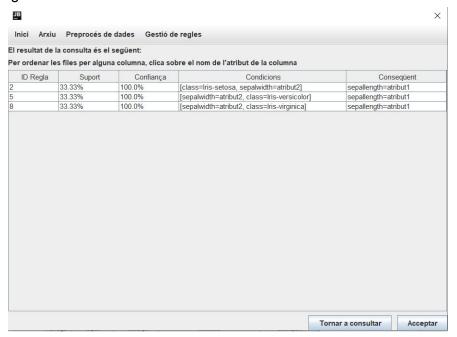


Per consultar s'ha d'introduir el nom i el valor de l'atribut conseqüent en la regla d'associació. Un cop introduïdes aquestes dades el sistema comprova que siguin correctes i mostra les regles d'associació que té aquest conseqüent.

Per exemple, volem consultar per consequent com a nom de l'atribut 'sepallength' i valor 'atribut1' tal com es pot veure a la següent imatge.



A continuació, després de donar-li al botó acceptar, el sistema mostra el llistat de regles d'associacions que tenen aquest atribut com a conseqüent. El resultat es pot observar en la següent imatge.



Sortir

Per tal de finalitzar l'execució del programa des de qualsevol de les pantalles principals de cada funcionalitat disposarem del botó Inici a partir del qual podrem accedir a la pantalla principal d'aquest i dins la pantalla principal tindrem un botó anomenat Sortir que ens permetrà acabar l'execució finalment.