Instituto Politécnico Nacional

Escuela Superior de Cómputo

Unidad de Aprendizaje: Minería de datos

Grupo: 5CDM1 Ciclo escolar: 2022–2

**Proyecto No. 2. Reglas de Asociación**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Grupo: | 5CDM1 |  | Equipo: |  |

Nombre de los integrantes del equipo:

|  |  |
| --- | --- |
| 1) |  |
| 2) |  |

*Inserte una portada indicando el tema que aborda*

**PARTE 1. Ejercicio descriptivo de la generación de reglas de asociación**

Realice el proceso (paso a paso) para generar las reglas de asociación del ejercicio descrito y los parámetros: Fi = 2. min.sup = 2. min.conf = 75% (modificable)

Genere un conjunto de datos con 7 productos y 10 transacciones (2 con dos productos, 1 de tres productos, 4 de cuatro productos, 2 de cinco productos y 1 de seis productos).

Agregue los siguientes apartados:

1. Introducción
2. Desarrollo
   1. Listado de transacciones
   2. Representación tabular
3. Primera Etapa
4. Segunda Etapa
5. Evaluación de las reglas
6. Conclusiones

*Inserte una portada indicando el tema que aborda*

**PARTE 2. Aprendizaje supervisado**

Considere el siguiente conjunto de datos:

|  |  |
| --- | --- |
| Nombre: | Conjunto de datos de registros de votación del Congreso  (*Congressional Voting Records Data Set*) |
| Objetivo: | Este conjunto de datos incluye los votos de cada uno de los congresistas de la Cámara de Representantes de los Estados Unidos sobre los 16 votos clave identificados por la CQA.  La CQA enumera nueve tipos diferentes de votos: votado a favor, emparejado a favor y anunciado a favor (estos tres simplificados a sí), votado en contra, emparejado en contra y anunciado en contra (estos tres simplificados a no), votado presente, votado presente para evitar conflicto de intereses, y no votó ni dio a conocer una posición (estos tres simplificados a una disposición desconocida). |
| Créditos: | Congressional Quarterly Almanac, 98th Congress, 2nd session 1984, Volume XL: Congressional Quarterly Inc. Washington, D.C., 1985.  Donante:  Jeff Schlimmer (Jeffrey.Schlimmer '@' a.gp.cs.cmu.edu) |
| Enlace de acceso: | <https://archive.ics.uci.edu/ml/datasets/congressional+voting+records> |
| Clases: | Demócrata, Republicano |

1. Crear el diccionario de datos con los siguientes campos:

|  |  |  |  |
| --- | --- | --- | --- |
| **Nombre** | **Significado** | **Tipo** | **Dominio** |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

1. Realice el tratamiento de datos necesario para aplicar la técnica de minería de datos de reglas de asociación. (Ustedes establezcan la configuración de los parámetros para la generación de las reglas).
2. Elija las reglas de asociación que tengan como consecuente las clases indicadas (Demócrata, Republicano) para identificar las características de cada partido político.
3. Describa las características generadas por las reglas de asociación que distinguen cada partido.

|  |  |
| --- | --- |
| **Partido** | **Características (listado)** |
| Demócrata |  |
| Republicano |  |

Agregue las reglas de asociación que generó tales características

*Inserte una portada indicando el tema que aborda*

**PARTE 3. Aprendizaje No supervisado**

Crear un reporte (tipo catálogo) de ofertas mediante la herramienta BIRT de KNIME que sugiera productos *de un conjunto de datos que usted elija*. Debe generar reglas de asociación.

Diseñe un catálogo recomendando los productos de *elegidos* más vendidos o menos vendidos (ustedes determinarán dicho criterio). Los datos mostrados en el catálogo los generará considerando las reglas generadas:

Antecedente 🡪 Consecuente

En el reporte mostrará los nombres y los precios, marca (puede incluir algún otro dato) de los productos tanto de antecedente como de consecuente y con base en el significado de soporte, confianza, lift o algún otro parámetro que calcule, realice una oferta de esa asociación.

Para el conjunto de datos elegido, agregue lo siguiente:

|  |  |
| --- | --- |
| Nombre: |  |
| Objetivo: |  |
| Créditos: |  |
| Enlace de acceso: |  |

1. Crear el diccionario de datos con los siguientes campos:

|  |  |  |  |
| --- | --- | --- | --- |
| **Nombre** | **Significado** | **Tipo** | **Dominio** |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

1. Realice el tratamiento de datos necesario para aplicar la técnica de minería de datos de reglas de asociación. (Ustedes establezcan la configuración de los parámetros para la generación de las reglas).

Elija una muestra de 500 registros.

1. Generar un archivo de transacciones y elegir los criterios para generar las reglas de asociación, por ejemplo: recomendar productos de la misma marca,

**PARTE 4. Flujo de trabajo en Knime**

Sólo un integrante de cada equipo enviará el reporte en formato pdf, el archivo generado por el Knime, a la asignación correspondiente.

La entrega de este reporte representa la guía de la evaluación