**Comparación de algoritmos de análisis de asociación**

Como se observó en la etapa 5 del método de la ingeniería, entre los métodos seleccionados para implementar se encuentra el análisis de asociación. Para esto se decidió que se implementaría dos algoritmos que usaran esta técnica: A Priori y fuerza bruta. En esta sección haremos una pequeña comparación basada en el análisis de la media de 20 datos obtenidos como resultado en tiempos de estos algoritmos. Los datos de pruebas se encuentran en los documentos Articulos2.csv y Ventas2.csv, para los objetos artículos y los objetos ventas correspondientemente.

El escenario planteado para estas pruebas consiste en un conjunto de 100 transacciones, y 20 artículos con los cuales se generarán los itemset candidatos. Además, para hacer la prueba mucho más precisa, usaremos el namespace Diagnostico, y Threading. Con estos podremos realizar una secuencia de pasos que nos permitirán ajustar las pruebas y controlar factores externos a nuestra prueba, relacionados con el hardware del equipo.

Usando estos namespace, primero escogeremos el segundo núcleo de nuestro procesador, y además le daremos la mayor prioridad a nuestro programa. Esto con el objetivo de evitar que otras aplicaciones interfieran en nuestra prueba. Posteriormente realizaremos un “warm-up” por parte del procesador, con una duración de 1000 milisegundos (ms), con el objetivo de estabilizar el contador de tics del namespace diagnostico (Stopwatch).

Una vez realizado todo esto, se realiza la ejecución del algoritmo 20 veces. Los resultados serán presentados a continuación.

**Datos**

Las hipótesis planteadas fueron:

**Ejecución de los algoritmos:**

|  |  |
| --- | --- |
| Algoritmo A priori | |
| Intento | Tiempo (ms) |
| 1 | 1002 |
| 2 | 995 |
| 3 | 968 |
| 4 | 974 |
| 5 | 982 |
| 6 | 960 |
| 7 | 961 |
| 8 | 959 |
| 9 | 958 |
| 10 | 962 |
| 11 | 960 |
| 12 | 956 |
| 13 | 954 |
| 14 | 953 |
| 15 | 955 |
| 16 | 954 |
| 17 | 954 |
| 18 | 957 |
| 19 | 962 |
| 20 | 961 |

**Tabla 1.** Resultado de los tiempos del algoritmo a priori.

|  |  |
| --- | --- |
| Algoritmo Fuerza Bruta | |
| Intento | Tiempo (ms) |
| 1 | 12073 |
| 2 | 12038 |
| 3 | 11952 |
| 4 | 11875 |
| 5 | 12112 |
| 6 | 12325 |
| 7 | 12156 |
| 8 | 12132 |
| 9 | 12057 |
| 10 | 12345 |
| 11 | 12284 |
| 12 | 12418 |
| 13 | 12084 |
| 14 | 12038 |
| 15 | 12309 |
| 16 | 12298 |
| 17 | 12362 |
| 18 | 12403 |
| 19 | 12150 |
| 20 | 12209 |

**Tabla 2.** Resultados de los tiempos del algoritmo fuerza bruta.

Como se puede observar a simple vista, la diferencia es de la escala de . Para verificar esto calcularemos las medias.

|  |  |
| --- | --- |
| Algoritmo | Media Tiempo (ms) |
| A priori | 964.35 |
| Fuerza Bruta | 12181 |

**Tabla 3.** Resultados del cálculo de medias de los   
tiempos de ambos algoritmos.

Los resultados que se muestran en la tabla 3 indican claramente que se rechaza la hipótesis nula, y se acepta la hipótesis alternativa. Es decir, la media en tiempos del algoritmo fuerza bruta es estrictamente mayor que la del algoritmo a priori.

**Conclusiones de la prueba**

El algoritmo a priori es mucho más eficiente por lo menos en términos de tiempo. La diferencia entre ambos algoritmos es bastante amplia, y se puede notar incluso a tamaños de datos bastantes pequeños.