**Diseño de experimentos**

**Planeación y realización**

Para el proyecto que se está realizando, se hace presente un concepto muy amplio e importante llamado “Big Data”, los cuales son datos que contienen una mayor variedad y que se presentan en volúmenes crecientes y a una velocidad superior (Gartner, 2001). Se pretende hacer uso de estas grandes cantidades de datos para llegar a la solución requerida por la empresa, cabe anotar que para lograr hacer un buen análisis sobre estos, en la informática se requiere de diversas estrategias o algoritmos que sean capaces de realizar una lectura y diferentes interacciones entre datos con un tiempo de respuesta razonable, convirtiéndose en un problema en el caso de que el proceso tome mucho tiempo para brindar una respuesta, además es de gran importancia enfocarse en la optimización de estos procesos computacionales debido a que es una solución que va a brindar ganancias a la empresa. Por esta razón el experimento se desarrollará entorno a los diferentes tiempos que brinden las diversas alternativas de estrategias propuestas.

1. **Unidad experimental**

* Cantidad de combinaciones por conjunto de artículos (2,3,4 y 5)

1. **Variables de respuesta**

* Tiempo de respuesta del algoritmo

Con esta variable se refleja una excelente comparación entre las diferentes estrategias implementadas, ya que se quiere llegar a la mejor solución en cuanto a la optimización del software, además brinda unos niveles de medición con una alta exactitud debido a que los tiempos serán tomados con la herramienta de Visual Studio llamada stopWatch.

1. **Factores controlables**

* Procesador del equipo donde se procede la ejecución
* Memoria Ram del equipo donde se procede la ejecución
* Cantidad de núcleos con los que cuenta el equipo
* Cantidad de datos en la base de datos
* Valor del soporte (%)
* Valor de la confianza (%)

1. **Factores no controlables**

* Procesos que ejecuta el sistema operativo mientras se lleva a cabo el experimento
* Fragmentación de la memoria del disco duro del equipo
* Fragmentación de la memoria RAM

1. **Factores estudiados**

* Tipo de algoritmo
* Tamaño de la base de datos

1. **Niveles**

En este experimento se decidió elegir como factores de estudio los siguientes niveles:

La siguiente tabla (Tabla.1) representa los algoritmos disponibles a evaluar:

|  |
| --- |
| **Tipos de algoritmos** |
| Fuerza bruta |
| A-priori |
| F(K-1)xF1 |

La siguiente tabla (Tabla.2) representa los tamaños de la cantidad de artículos:

|  |
| --- |
| **Cantidad de combinaciones** |
| 2 |
| 3 |
| 4 |
| 5 |

1. **Tratamientos**

En el siguiente experimento se medirá el nivel de efectividad en términos de velocidad de los métodos mencionados. Los tratamientos se dividirán entre los niveles ya previamente mencionados siendo los siguientes:

|  |  |  |
| --- | --- | --- |
| **Tipos de algoritmos** | **Cantidad de combinaciones** | **Tratamientos** |
| Fuerza Bruta | 2 | 1 |
| 3 | 2 |
| 4 | 3 |
| 5 | 4 |
| A-priori | 2 | 5 |
| 3 | 6 |
| 4 | 7 |
| 5 | 8 |
| F(K-1)xF1 | 2 | 9 |
| 3 | 10 |
| 4 | 11 |
| 5 | 12 |

Con base en el diseño de experimento planeado se realizará las diferentes comparaciones entre los algoritmos de Fuerza Bruta, A-priori y F(K-1)xF1. Este proceso será registrado por medio de la implementación de tablas ANOVA las cuales brindarán una información detallada con la que se podrá concluir de manera exacta la mejor alternativa que arroja el experimento.