Método de la ingeniera

Santiago Gutiérrez

Esteban Mendoza

Martin Pérez

**PROFESOR:**

Uram Aníbal Sosa Aguirre

**ALGORITMOS Y ESTRUCTURAS DE DATOS**

**UNIVERSIDAD ICESI 2021-2**

***Contexto Problemático:*** Las tiendas de videojuego en la ciudad de cali han mantenido un mismo sistema de negocio durante mucho tiempo, debido a esto, un millonario planea abrir una tienda con una innovadora forma de facilitarle y automatizarle la búsqueda de los videojuegos a los compradores.

El baloncesto es un deporte bastante popular a nivel mundial, de forma que la gran cantidad de datos que este deporte genera se ha incrementado de gran manera. Entre estos datos se incluye la información estadística de los jugadores profesional la cual se requiere poder ser analizada correctamente.

***Desarrollo de la solución:*** Para encontrar la solución adecuada a este problema se ha optado por usar el método de la ingeniería para la resolución de problemas, la cual consta de los siguientes 7 pasos:

· ***1)La identificación del problema:*** Una de las partes mas importantes del proceso, una identificación incorrecta del problema causa una perdida de tiempo y puede llevar una solución inadecuada, por lo que se debe tener claro que es lo que se pide para empezar el desarrollo las soluciones.

· ***2) recopilación de la información necesaria:***  Búsqueda de la información necesaria para la resolución correcta del problema identificado anteriormente. En este caso conocer si existen otras herramientas con un sistema similar de análisis de información y cual ha sido la eficiencia de estos en caso de haber.

· ***3) Búsqueda de soluciones Creativas:*** se buscan ideas por medio de múltiples métodos, entre los cuales se puede destacar la lluvia de ideas, lista de revisión y lista de atributos.

· ***4) Transición de la formulación de ideas a los diseños preliminares:*** Se descartan las ideas menos factibles del punto anterior y adicionalmente se hace un mayor enfoque en las mas viables.

· ***5) Evaluación y selección de la idea preferente:*** Se decide por una idea de las anteriores después de hacer un análisis de profundo de las ventajas y desventajas de aquellas que pasaron el filtro del paso 4.

· ***6) Preparación de informes y especificaciones:*** Se documentan los aspectos mas importantes del proyecto.

· ***7) Implementación del diseño:***  Se efectúa la etapa experimental con la idea obtenidos en los pasos anteriores.

A continuación, se elaborará de forma mas profunda cada paso del método de la ingeniería presentado anteriormente:

***1)La identificación del problema:***

***Debido a la gran cantidad de jugadores profesionales y la cantidad de información estadística importante que posee cada uno de estos, es necesario una forma de poder analizar esta gran cantidad de información además de tener la capacidad de añadir o modificar algún dato.***

***2) Marco teórico:***

***En la actualidad existen múltiples herramientas que permiten un análisis de grandes cantidades información como la que se desea hacer para los jugadores profesionales de baloncesto, entre estas herramientas se pueden destacar las basadas en el análisis del big Data, donde se destacan el software de análisis basado en la nube, como lo es BSA, la cual posee entre sus funciones destacables la captura datos de bienestar, estado físico y entrenamiento de jugadores de registro, además de crear etiquetas personalizadas para cualquier evento, juego, jugador y estadísticas del equipo. Las herramientas similares a la anterior han tenido un gran éxito pero debido a que están adaptadas para múltiples deportes no poseen el énfasis en el baloncesto que se requiere para el proyecto y adicionalmente no brinda la capacidad para la modificación de la información.***

***Ya habiendo realizado una recopilación de las herramientas que cumplen una función similar es necesario recopilar información acerca de las herramientas que podrían ser usadas en el proyecto:***

Listas enlazadas: son una forma de estructura de datos simple, donde cada elemento tiene una referencia al siguiente. De tal forma que con tener un elemento se poseen todos los siguientes.

Generics: Son un tipo de dato parametrizado, que permite la generalización de los métodos de una clase para que puedan utilizarse en múltiples ocasiones , adicionalmente permite la identificación de errores decompilación de una mejor manera.

* + - Queues: Son un tipo de estructura de datos lineal, con el cual solamente se puede acceder al último elemento. es de tipo last in first out. es una estructura que tiene similitudes con las colas comunes del dia a dia donde sale el primero y los nuevos van al final.
    - Tipo Abstracto de Dato: Es un modelo matemático que consta de un conjunto de datos a los cuales se les asignan una lista de operaciones permitidas.
    - Github: github es un servicio de repositorio online que permite gestionar y controlar las versiones del codigo ademas que brinda una gran cantidad de herramientas para el trabajo en equipo en un proyecto.
    - HashTable: Es una matriz asociada de llaves con un valor. su operacion principal que soporta de una manera eficaz es la busqueda. por medio de las claves se puede acceder a la informacion del valor asociado. para esto convierte una la clave en un valor hash para ubicarlo en la tabla.
    - Arbol binario de Busqueda: Un árbol binario de búsqueda es un tipo particular de árbol binario donde se tiene que cada nodo es su propio su árbol binario de búsqueda y adicionalmente tiene dos hijos, uno izquierdo y derecho, donde se cumple siempre que el izquierdo es menor que el nodo padre y el derecho es mayor que el nodo padre.
    - Arbol Autobalanceado: Es un árbol binario que mantiene una diferencia entre sus alturas de 1, 0 o -1.

Fase 3: Búsqueda de soluciones creativas.

Para seguir encontrando forma de dar con soluciones a la problemática planteada, es posible definir una lluvia de ideas en función del uso de las TIC’s para probar distintas formas de solucionar dicha situación. Es por esto, que por parte de cada uno de los integrantes del grupo se produjeron ciertas ideas que acoplan de manera directa la problemática a trabajar. Dichas propuestas son:

* Usar las tablas hash como estructura de datos para almacenar a los jugadores.
* Usar múltiples colas ordenadas para organizar a los jugadores en las múltiples categorías
* Usar múltiples arboles binarios, uno para cada categoría de estadística de los jugadores.
* Usar un árbol binario que tenga como parámetro numérico la categoría estadística que se desea.
* Usar un árbol n-ario para guardar la información de los jugadores.
* Usar listas enlazadas como estructura de datos para almacenar la información.
* Usar arboles binario balanceados para guardar los jugadores.

**Fase 4: *Transición de la formulación de ideas a los diseños preliminares:***

Es necesario encontrar la mejor solución para resolver el manejo de la gran cantidad de datos que se tendrá, por esta razón es necesario es necesario revisar todas las propuestas que se hicieron.

Se empezó revisando cada una de las soluciones propuestas, donde se descartaron aquellas diferentes al uso de arboles binarios y arboles n-arios. Esto debido a que para este proyecto se tendrá una gran cantidad de datos, de forma que tanto las tablas hash, colas y listas enlazadas no son eficientes a comparación de los ya mencionados arboles.

***5) Evaluación y selección de la idea preferente***

Ya habiéndose descartado las ideas menos útiles se procede a elegir una entre las que pasaron a la siguiente fase, es decir, la creación de un único árbol binario según el momento, el uso de un árbol n-ario y el uso de cinco arboles auto balanceados.

Para maximizar la eficiencia del programa al buscar valores que cumplan cierta condición de búsqueda lo ideal seria el uso de arboles binarios y auto balanceados, esto con tal de evitar la búsqueda incensaría en nodos para la obtención de datos o para encontrar el sitio adecuado donde agregar un nuevo nodo.

Ya descartado el árbol n-ario se considere que debido a la gran cantidad de datos que poseerá el programa este seguramente tendrá una gran diferencia entre sus hijos izquierdos y derechos, esto debido a que su raíz será el primer dato ingresado, que en el peor de los casos podría ser el menor o el mayor, haciendo que todo el peso del árbol recaiga en un solo lado. Debido a lo anterior y con tal de evitar lo mayor posible este caso, es necesario que el árbol este siendo balanceado constantemente, esto debido a que el costo de balancear el árbol una vez que estén todos los datos será mucho mayor a que se vaya balanceado después de agregar cada nodo. Debido a lo anterior es que se utilizaran los arboles auto balanceados. Para finalizar con tal de decidir si debería usarse un único árbol o múltiples se considere que el programa va estar cargando múltiples veces los mismos datos en caso de hacerse múltiples búsquedas filtradas, por lo que es preferible un único mayor esfuerzo inicial que a largo plazo aumenta la eficiencia a comparación de la carga constante de datos.

***6) Preparación de informes y especificaciones***

Es necesario analizar los documentos en la carpeta docs. de este mismo proyecto.

***7) Implementación del diseño***

La implementación del diseño de nuestro proyecto está montada en GitHub con un enlace al repositorio.