

Método de la ingeniería

Fase 1: Identificación del problema

Estados Unidos es el tercer país más poblado del mundo con un total de 330 millones de habitantes. En él, aproximadamente se encuentran 106 millones de viviendas donde las personas de los diferentes estados habitan. Además, se tiene en cuenta que en un futuro este mercado avance de forma directa con el aumento de habitantes en cada una de las regiones nativas. Es por esta razón que en algunas ocasiones las compañías de seguro, impuestos de registro y la administración gubernamental, se ven paralizadas por el número de peticiones o encargos que cada día manejan.

Tomando en cuenta esta situación, se necesita realizar un software de aprendizaje autónomo con la idea de agilizar los procesos de compra y venta de cada una de las viviendas solicitadas, dominando los siguientes requerimientos funcionales:

1. Leer y cargar el dataset “House Sales in King County, USA” proveniente de Kaggle.
2. Visualizar los datos de manera correcta y ordenada desde el momento en que el usuario cargue el dataset, hasta cuando deje de utilizar el software.
3. Codificar e implementar un árbol de decisión como la estructura primaria del programa.
4. Clasificar un nuevo datapoint designando un rango de precio con base a las opciones establecidas con anterioridad.
5. Diseñar una interfaz gráfica de usuario donde se logre realizar cada una de las funcionalidades deseadas para el proyecto.
6. Diagramar la base del software por medio de clases (UML), y objetos (con base al primero respectivamente)
7. Generar gráficos estadísticos analizando los datos brindados por el dataset seleccionado con anterioridad.
8. Implementar un algoritmo que permita generar un árbol de decisiones con una buena efectividad para predecir el precio del inmueble a partir de los demás factores presentados en el dataset.

Fase 2: Recolección de la información necesaria

Para solucionar este problema y realizar un correcto software, es indispensable conocer algunas herramientas importantes que permitan cumplir con cada uno de los requerimientos funcionales anteriormente descritos. Algunas de estas son:

Machine Learning: Es una habilidad indispensable para hacer sistemas capaces de identificar patrones entre los datos para hacer predicciones. Esta tecnología está presente en un sinnúmero de aplicaciones como las recomendaciones de Netflix y Spotify o las respuestas inteligentes de Gmail y el habla de Siri y Alexa. Además, es capaz de convertir una muestra de datos en un programa informático capaz de extraer inferencias de nuevos conjuntos de datos para los que no ha sido entrenado previamente

Árbol de decisiones: Esta herramienta nos ayuda a la hora de tomar decisiones correctas cuando se tienen grandes cantidades de datos, y se presenta como posibles ramas en las que se plantean diferentes soluciones a un escenario continuo de sucesos. Asimismo, tomando en cuenta cada complejidad situada en diferentes planteamientos, la tecnología hablada presenta diferentes uso como:

- Plantear problemas desde diferentes puntos de vista.
- Analizar todas las posibles soluciones a una misma decisión.
- Estimar factores cuantitativos cuando el problema lo permita.
- Tomar decisiones en base a información real y confiable.
- Analizar diferentes alternativas y probabilidades que cada rama puede generar.

Gráficos: Una parte indispensable del programa es generar gráficos estadísticos que permitan entender el problema de manera significativa. Por ende, es de gran importancia reconocer cuales pueden ser algunos tipos de gráficos que podemos utilizar, y cuales son de gran necesidad para el problema planteado.

Lector de archivos: Para utilizar el sistema de forma eficaz, es necesario leer rápidamente un dataset con toda la información existente hasta el momento. De esta manera no solo se evita perder datos significativos, sino que también permite filtrar casillas por cualquier tipo de limitación que el mismo usuario interponga.

Además de esas funcionalidades, se complementa la información con los requerimientos no funcionales encontrados en el sistema:

1. Depositar el código fuente, y cualquier otro cambio en el sistema en la plataforma GitHub.
2. Grabar el uso Beta del sistema para mostrar su funcionamiento a las personas interesadas en utilizar o adquirir el software.
3. Programar e implementar el programa en el lenguaje C# (Visual Studio).

Fase 3: Búsqueda de soluciones creativas