**Problema**

En la actualidad, el mundo de Internet es enormemente enorme teniendo en cuenta las páginas web con gran cantidad de sustancias explicativas que se pueden obtener con diseños disímiles como texto, gráfico, audio vídeo, etc. que se centrarán en la contradicción en la reposesión de los hechos debido a la insignificancia con respecto al hecho de que el usuario está viendo. Los datos que muestran los sitios web sólo pueden verse mediante un navegador web. No ofrecen la funcionalidad de guardar una copia de estos datos para uso personal. La única opción entonces es copiar y pegar manualmente los datos expuestos en el sitio web en el navegador en el disco duro de nuestro ordenador, lo cual es un trabajo bastante pesado. Aquí es donde entra en juego el web scraping.

Aproximadamente el 70% de los datos estimulados en Internet se obtienen de documentos PDF, que no están estructurados y son difíciles de manejar. Por otra parte, una página web es un formato organizado (formado por código HTML), aunque con un enfoque no reutilizable. Esta gran cantidad de información distribuida pero confinada de este tipo de diseño se suele denominar "la tiranía del PDF". Asegurando la posibilidad clave de este documento (HTML documentos), su naturaleza organizada reproduce las potencialidades expuestas con los procedimientos de scraping. Los procedimientos y las herramientas de herramientas confían en la organización, así como en los activos del lenguaje HTML del lenguaje HTML. Esto permitirá el trabajo de raspado para herramientas y robots concluyendo así a los humanos aparte de las tediosas, monótonas y defectuosas tareas de recuperación de datos físicos. Para Para concluir, estas herramientas comprometen los datos en formatos accesibles formatos accesibles para una dispensación e integración avanzadas: JSON, XML, CSV, XLS o RSS.

Dentro de la web existe muchas páginas web que contienen información acercar de las criptomonedas y sus variaciones de precio que es muy difícil de entender para personas que quieren iniciar en este tipo de actividades que facilita el internet además esta información en su mayoría está en inglés y esto es otro obstáculo para las personas, por lo que se necesita transformar esta información compleja en algo más entendible para las personas que se quieren dedicar a la compra y venta de criptomonedas.

**Que es el web scraping**

El raspado web (también conocido como Screen Scraping, Web Data Extraction y Web Harvesting, etc.) es un procedimiento de extracción automática de datos web en lugar de copiarlos manualmente. Es una técnica en la que se extraen datos significativos del HTML de los sitios web y se almacenan en una base de datos local central o en una hoja de cálculo. Para ello se utiliza la URL del sitio web. Lo llevan a cabo los "web scrapers" con la ayuda de programas especialmente codificados. Puede estar montado tradicionalmente para algún sitio web preciso o puede ser uno que se organice fácilmente para trabajar con cualquier sitio web. El objetivo de un Web scraper se concentra en la conversión de los datos no estructurados mientras se conservan en bases de datos organizadas. Algunos procedimientos de raspado web procedimientos de raspado web son la programación HTTP, el análisis sintáctico DOM y los parsers. Los datos generados se utilizan posteriormente para su recuperación o análisis. Es una gran ventaja ya que nos proporciona datos sin errores, ahorra nuestro tiempo para dar resultados rápidos como un rayo y almacena todos los datos en un solo lugar. También podemos elegir el formato en el que que debe estar a nuestra disposición. Esto permite una facilidad de acceso y facilita el análisis de los datos.

En la actualidad, el raspado web se aplica a varios aspectos, como la comparación de precios en línea, el seguimiento de los datos meteorológicos, la detección de cambios en los sitios web, el mashup web, la investigación web y la integración de datos web. Además, cabe señalar que el raspado de la web puede ser paralelo a los periodos de uso de algunos sitios web.

**Fuente de conocimiento**

Se intentará hablar con un experto en Criptomonedas (fuente de conocimiento primaria), pero en caso de no poder conseguir el experto para este proyecto se procederá a la búsqueda de proyectos y artículos similares.

**Representación del conocimiento**

Web Scraping, Análisis exploratorio de datos

Es una forma de analizar datos definido por John W. Tukey (E.D.A.: Exploratory data analysis) es el tratamiento estadístico al que se someten las muestras recogidas durante un proceso de investigación en cualquier campo científico.

El objetivo de este análisis es explorar, describir, resumir y visualizar la naturaleza de los datos recogidos en las variables aleatorias del proyecto o investigación de interés, mediante la aplicación de técnicas simples de resumen de datos y métodos gráficos sin asumir asunciones para su interpretación.

**Resumen**

El sistema experto tomará la información de la página web elegida (ya sea de su HTML o de sus data sets) acerca de las criptomonedas y se hará uso de la técnica de Análisis exploratorio de datos y web scraping que no es más que bots programados para examinar bases de datos y extraer información de páginas web.