TasaCarros

# Overview

The \*\*TasaCarros\*\* project is a Python-based tool designed to scrape car information from the web, specifically from vehicle listing websites, and store both metadata and images in an HDF5 file. The project can be expanded to scrape data from multiple pages and convert price information between currencies.

# Features

- Scrapes car data (name, year, price) from a vehicle listing website.  
- Downloads and processes car images using `PIL` and stores them as NumPy arrays.  
- Saves metadata and images in HDF5 format.  
- Handles price conversion from USD to DOP.

# Main File

- `API.py`: This is the main entry point of the project. It integrates all components, from data scraping to image processing and saving in the HDF5 format.

# Requirements

Python 3.x  
- `requests`  
- `BeautifulSoup`  
- `h5py`  
- `numpy`  
- `PIL`  
  
You can install the required libraries using:  
```bash  
pip install -r requirements.txt  
```

# Installation

1. Clone this repository:  
```bash  
git clone https://github.com/Banco-FiHogar/TasaCarros.git  
cd TasaCarros  
```  
2. Install dependencies:  
```bash  
pip install -r requirements.txt  
```

# Usage

1. Run the main script:  
```bash  
python API.py  
```  
This script will:  
- Scrape car listings from the specified website.  
- Download images for each car.  
- Store car data (name, year, price) and image arrays into an HDF5 file named `car\_data.h5`.  
  
2. The script supports scraping multiple pages by adjusting the range:  
```python  
for i in range(1, 100): # Adjust the page range as necessary  
```

# Files

- `API.py`: The main script that orchestrates scraping and saving car data.  
- `DBcreator.py`: Handles the logic for scraping data and storing it in an HDF5 file.  
- `CrearModelo.py`: To be used for model creation (assumed).  
- `Main.py`: An additional main entry point for broader project execution.  
- `readme.txt`: Placeholder for the project's readme.

# Contributing

Feel free to fork the repository and submit pull requests for new features, bug fixes, or improvements.

# License

This project is licensed under the MIT License.