

Data-Science-Capstone-Project

=====

Repo containing the final Capstone Project for Data Science with Python Career Program by Skill Academy. The Capstone project is based on Car Details dataset, and the aim is to perform EDA on the given Dataset and train and evaluate a model to predict the price of car & give results based on the parameters given by the user. & deploy on streamlit.

Features

1. Explore the car dataset & analyze the data.
2. Visualize the given data.
3. Pre-process & clean the dataset.
4. Make a machine learning model following different modelling techniques and handling all the params to train the model.
5. Selecting the best model to evaluate results.
6. Check with sample data the actual prediction of the model.
7. Deployed a webapp to get the prediction of price using Streamlit.

Model Deployment (Using Streamlit)

To test and use the best model in predicting the cars from a dataset using a webapp by utilising the [Streamlit](https://streamlit.io/) interface.
To try & test the prediction visit -

```
<!-- [Car Prediction Streamlit Page](https://mitsu-ds-capstone-project.streamlit.app/) or
```

```
>
```

```
> [Alternate WebApp]([Streamlit App](https://static.streamlit.io/badges/streamlit_badge_black_white.svg)](https://sumit-ml-capstone-project.streamlit.app/))
-->
```

```
> <a href="url">`Car Prediction Streamlit Page :`</a> <a href="https://mitsu-ds-capstone-project.streamlit.app/"></a>
```

```
>
```

```
> <a href="url">`Alternate Webapp :`</a><a href="https://sumit-ml-capstone-project.streamlit.app/"></a>
```

```
---
```

Demo Video of Streamlit App

```
<p align="center">Stream Lit Demo
<video src="https://github.com/Sumit-SC/Data-Science-Capstone-Project/assets/43643943/f22682df-8879-4132-ad05-94746ea20d98" width="600" height="400" controls>Streamlit Demo</video>
</p>
```

```
---
```

Notebook Exploration

```
#### EDA [![Binder](https://mybinder.org/badge_logo.svg)](https://mybinder.org/
v2/gh/Sumit-SC/Data-Science-Capstone-Project/HEAD?labpath=%2Fnotebooks
%2FUsed_Cars_DA%28Graphical%2526Cleaning%29.ipynb)
```

```
#### ML Modelling Steps
[![Binder](https://mybinder.org/badge_logo.svg)](https://nbviewer.org/github/Sumit-
SC/Data-Science-Capstone-Project/blob/main/notebooks/Used_Cars_ML.ipynb)
```

Project Organization

```
├── README.md          <- The top-level README for developers using this
project.
├── data
│   ├── interim        <- Intermediate data that has been transformed.
│   ├── processed      <- The final, canonical data sets for modeling.
│   └── raw            <- The original, immutable data dump.
├── models
├── model summaries    <- Trained and serialized models, model predictions, or
& Model Evaulation (ML).
├── notebooks          <- Jupyter notebooks - Data analysis(Graphical analysis)
├── reports
│   └── figures        <- Generated analysis as HTML, PDF, LaTeX, etc.
reporting              <- Generated graphics and figures to be used in
├── requirements.txt    <- The requirements file for reproducing the analysis
environment, e.g.      generated with`pip freeze > requirements.txt`
├── setup.py           <- makes project pip installable (pip install -e .) so
src can be imported
├── src (Not in these Project) <- Source code for use in this project (Blank
for these project only used viz to store images.)
│   ├── data           <- Scripts to download or generate data
│   │   └── make_dataset.py
│   ├── features       <- Scripts to turn raw data into features for modeling
│   │   └── build_features.py
│   ├── models         <- Scripts to train models and then use trained models
to make                predictions
│   │   ├── predict_model.py
│   │   └── train_model.py
│   └── visualization  <- Scripts to create exploratory and results oriented
visualizations
│       └── visualize.py
└── app.py             <- Streamlit app to run a web app of the model.
```

<p><small>Project based on the cookiecutter data science project template. #cookiecutterdatascience</small></p>