END SEM FINAL PROJECT SUMBISSION

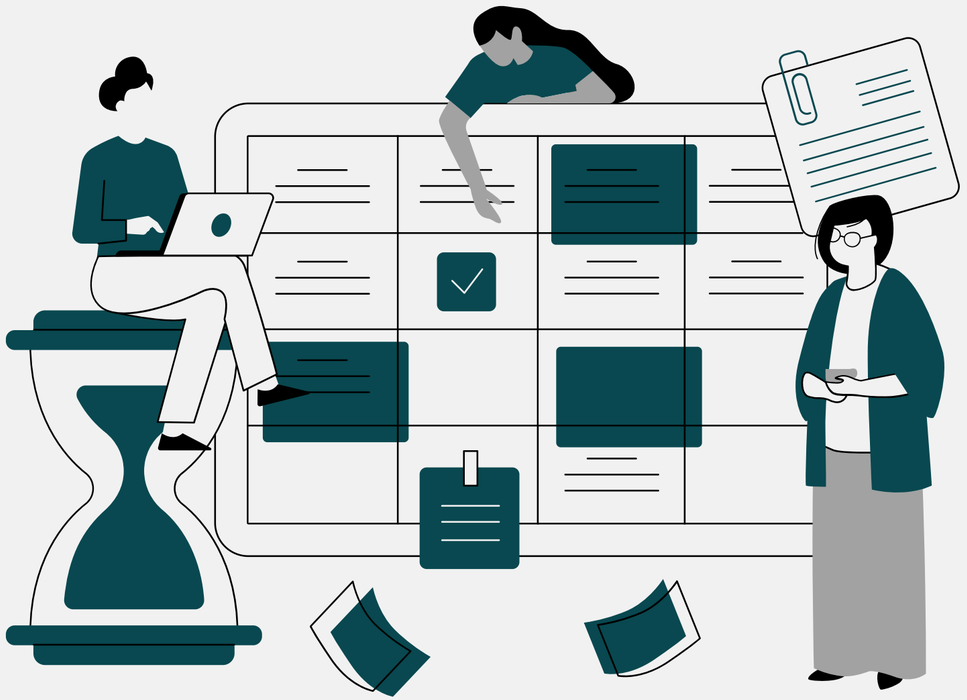
DISRUPTIVE TECHNOLOGY

Research is done by
JOYDEEP SEN

Prepared by
MAYANK SINGH RANA

Matters on the Docket

A brief look at what we will discuss on this report



About our project and problem statement

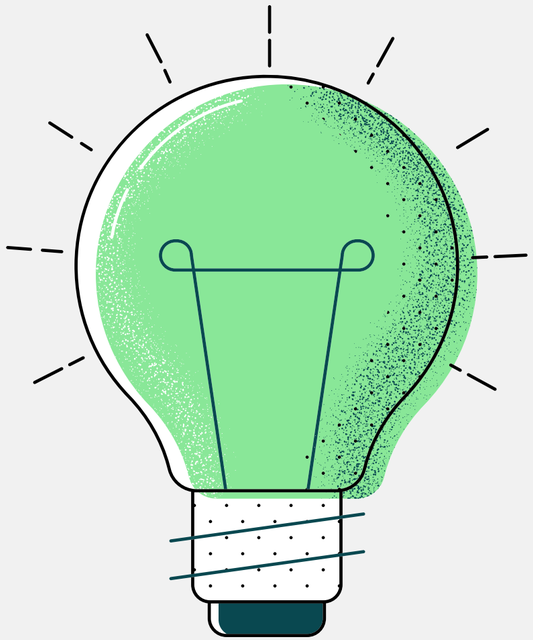
Problem Statement and Overview

We are want to help people that have less closure to the technical field and want ti to choose a device for themselves so they can use our application called LAPTOP INSIDER.

Solution and Brief About Project

So, the problem is that if any user wants to buy a laptop or any other electronic accessories, then our application should be compatible to provide a tentative price of laptop according to user configurations. Although it looks like a simple project or just developing a model, the dataset we have is noisy and needs lots of feature engineering and pre-processing that will drive your interest in our project

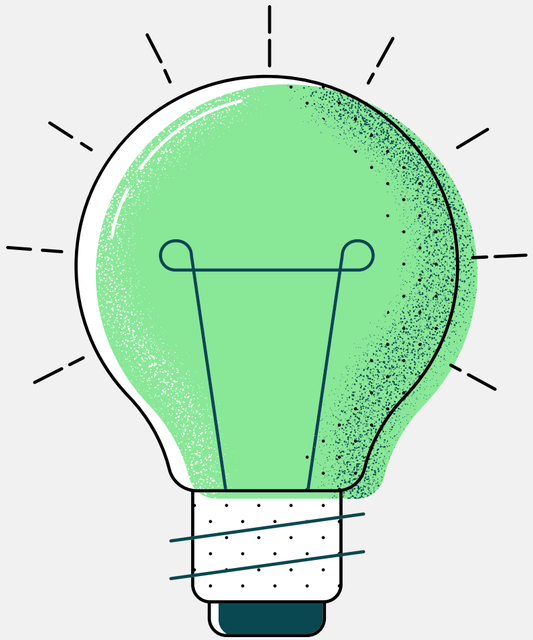
Highlights and Key Updates



Save time and energy

Basically, it will save the time for the research part of the customer that we will provide all the information of price and other things user just have to put his needs in the site and he is good to go.

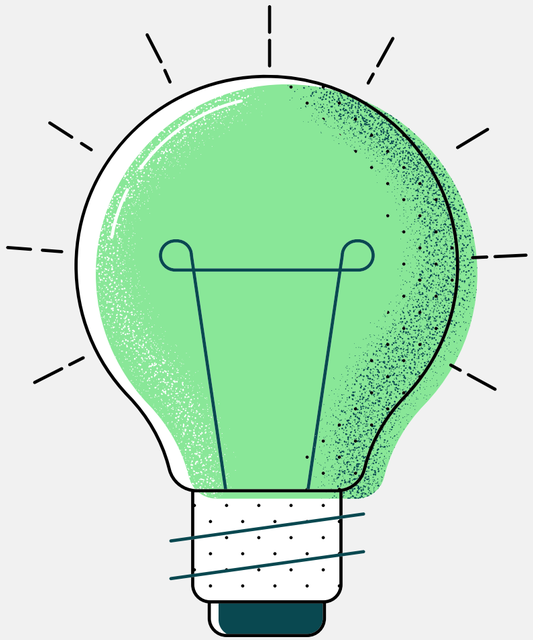
Highlights and Key Updates



Building a custom laptops and pc

As most the customers need their custom laptop and pc according to the work they do such as some do graphic designing video editing rendering high-end games and some do the basic task on their pc so our model can help the customer and also the company to build custom laptops and pc for them.

Highlights and Key Updates



Used as a Guide

Our model can also give suggestions to customers that how will you buy this laptop instead of this laptop by MACHINE LEARNING model by comparing configuration and price of other laptops and suggest the user/customer buy this item.

Methodology

A brief look at how we develop a model using different method.

DESCRIPTION OF THE DATASET

The dataset used for this research purpose was the dataset which is taken from Kaggle and further filtered by JOYDEEP SEN and it consists of 12 databases (Company, TypeName, Inches, screen resolution, CPU, Ram, memory, GPU, Weight, op sys, weight, price).

Methodology

A brief look at how we develop a model using different method.

PRE-PROCESSING OF THE DATASET

No null values exist in the dataset. However, there were numerous outliers to deal with, and the dataset was not distributed evenly. We employed two different ways. One that is free of outliers and uses a feature selection method instead of applying the results directly. inputs to machine learning algorithms, as well as the outcomes, weren't looking good. However, following the use of the dataset's normal distribution.

Methodology

A brief look at how we develop a model using different method.

CHECKING THE DISTRIBUTION OF THE DATA

When it comes to predicting or classifying a problem, the distribution of the data is crucial. As a result, we must balance the dataset, or it will get overfit. This will aid the model in identifying a pattern in the dataset that can be used to predict prices.

Methodology

A brief look at how we develop a model using different method.

Dataset Cleaning

As in every machine learning or any other project, first thing is to do is clean data according to the need of the projects. So in our projects, we have to clean data according to technical terms such as IPS display panel, RAM, ROM, SSD, etc.

Methodology

A brief look at how we develop a model using different method.

FEATURE SELECTION

When doing feature selection, the Lasso algorithm, which is a part of embedded techniques, is used to pick the features and just the significant ones. It outperforms the filter approaches in terms of predicted accuracy. It generates the best feature subsets for the algorithm in use. Select the selected features from the model in the scikit-learn library, which is a part of feature selection.

Methodology

A brief look at how we develop a model using different method.

CHECKING DUPLICATE VALUES ON THE DATA

Duplicate objects must be handled, else the generalisation of the constructed model will be harmed. There's a potential that if duplicates aren't handled properly, they'll show up in the test dataset as well as the training dataset.

Timeline

How we complete our project and divide the tasks

01

Data cleaning and pre-processing

This task is done by JOYDEEP he found the dataset from Kaggle and then clean the dataset ACCto our needs

02

Work on main code and development of model

This task is done by MAYANK he used the knowledge that is taught by Supreet sir in class and develop the model

03

Preparing a UI to host and deploy app

This task is done by Joydeep and develops a UI but due later on we get to know streamlit and Heroku and we used that technology.

04

Hosting of our ML project

This task is done by MAYANK he post our whole code on GITHUB and also on HEROKU.

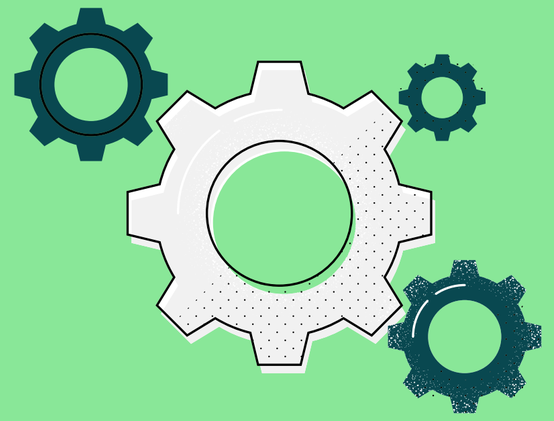
05

Submission of project

Final submission to the respected teacher.

Code and Output

CHECK HERE: https://rb.gy/lnvthg



Libraries used

- Numpy
- Pandas
- Matplotlib
- Sklearn
- Seaborn
- Sklearn
- Pickle

Dataset used

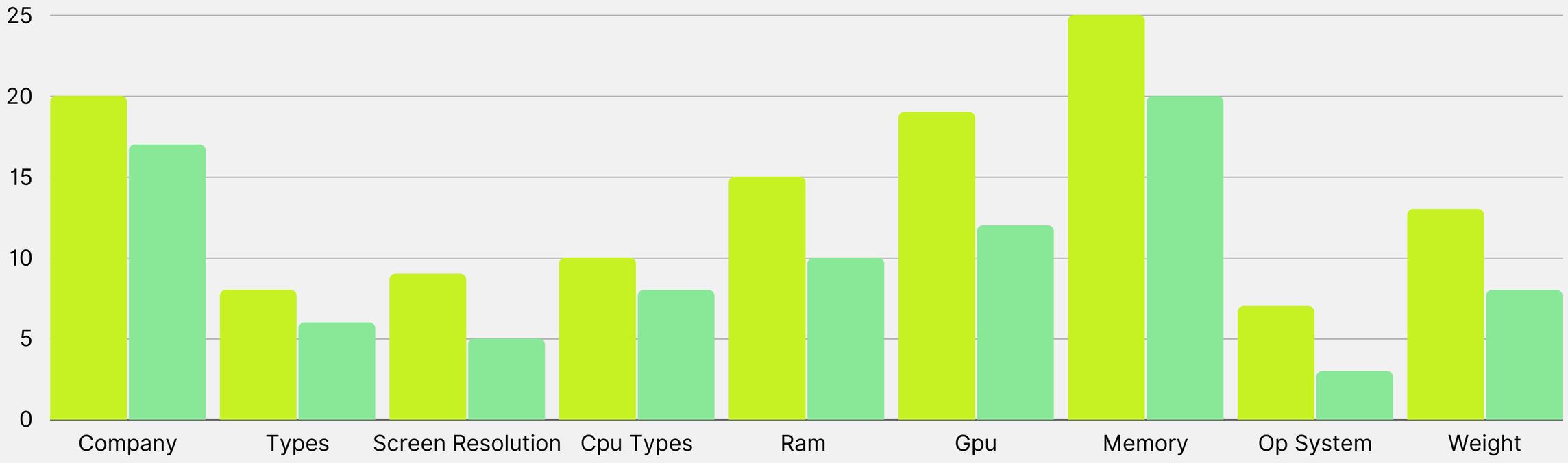
We have used a dataset from Kaggle. The dataset used for this research purpose was the dataset which is taken from Kaggle and further filtered by JOYDEEP SEN and it consists of 12 databases (Company, TypeName, Inches, screen resolution, CPU, Ram, memory, GPU, Weight,op sys, weight, price).

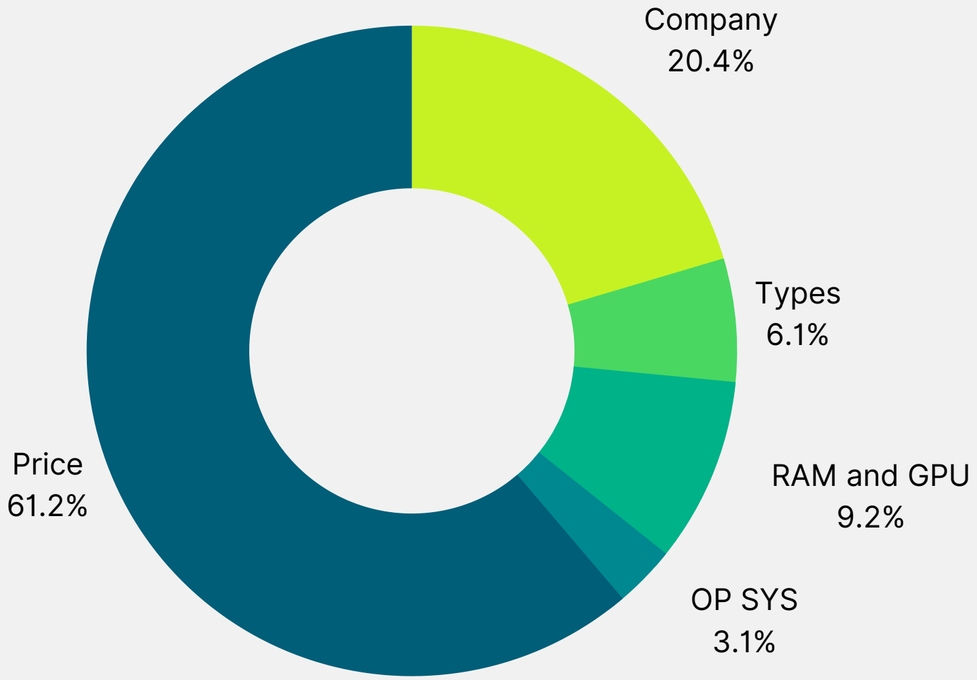
Outputs

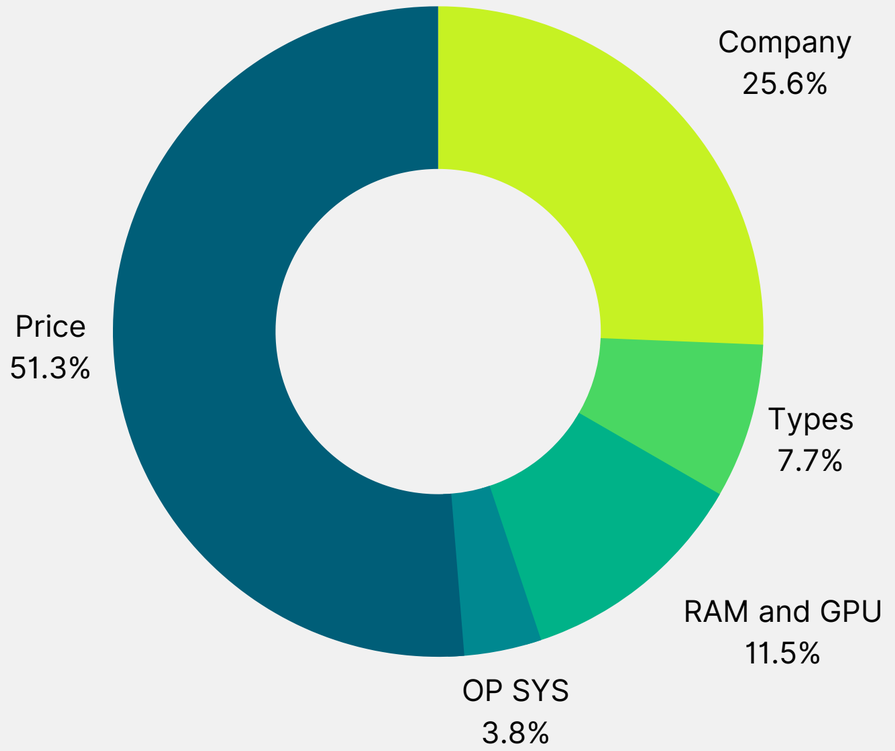
The output of code we can see here: **https://lpp-kaigon.herokuapp.com/**

Dataset Cleaning

Data after and before feature selection





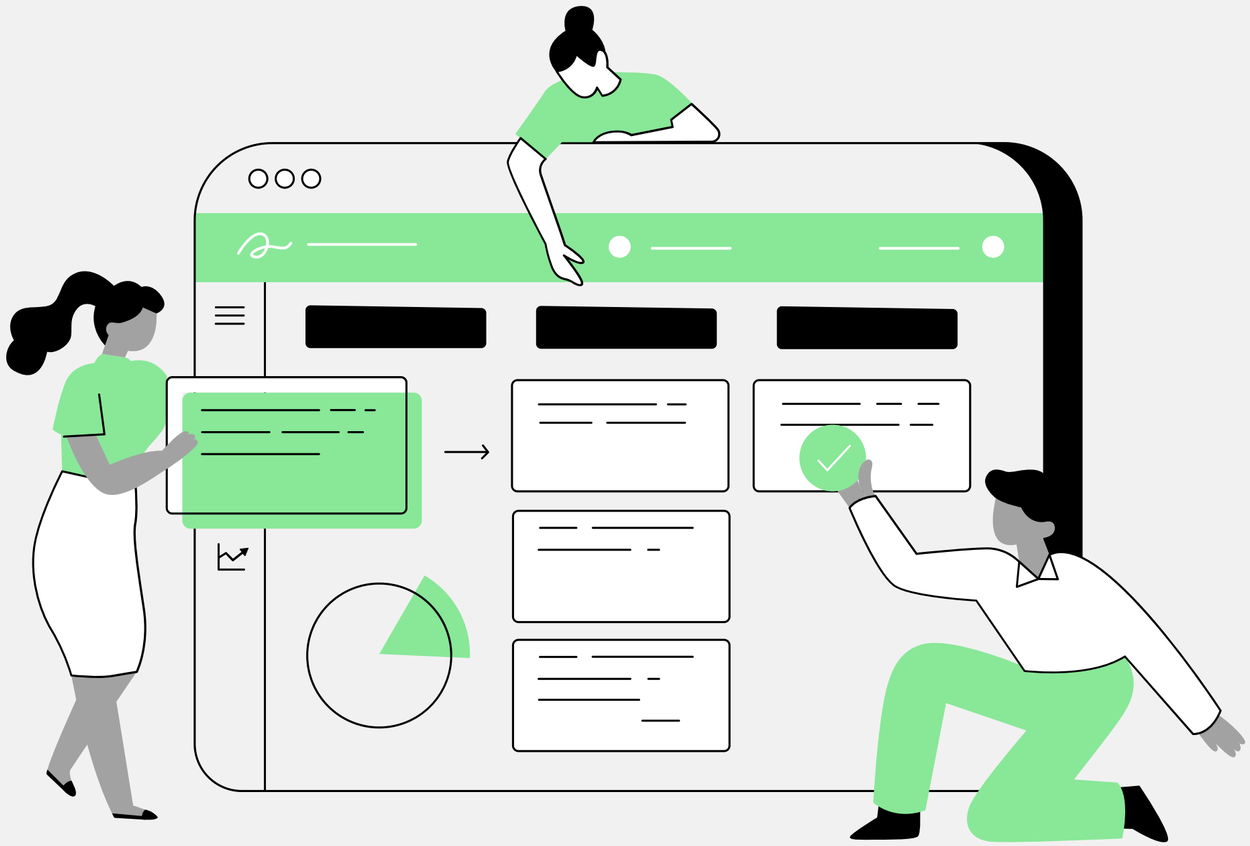


Predicted price

Actual price

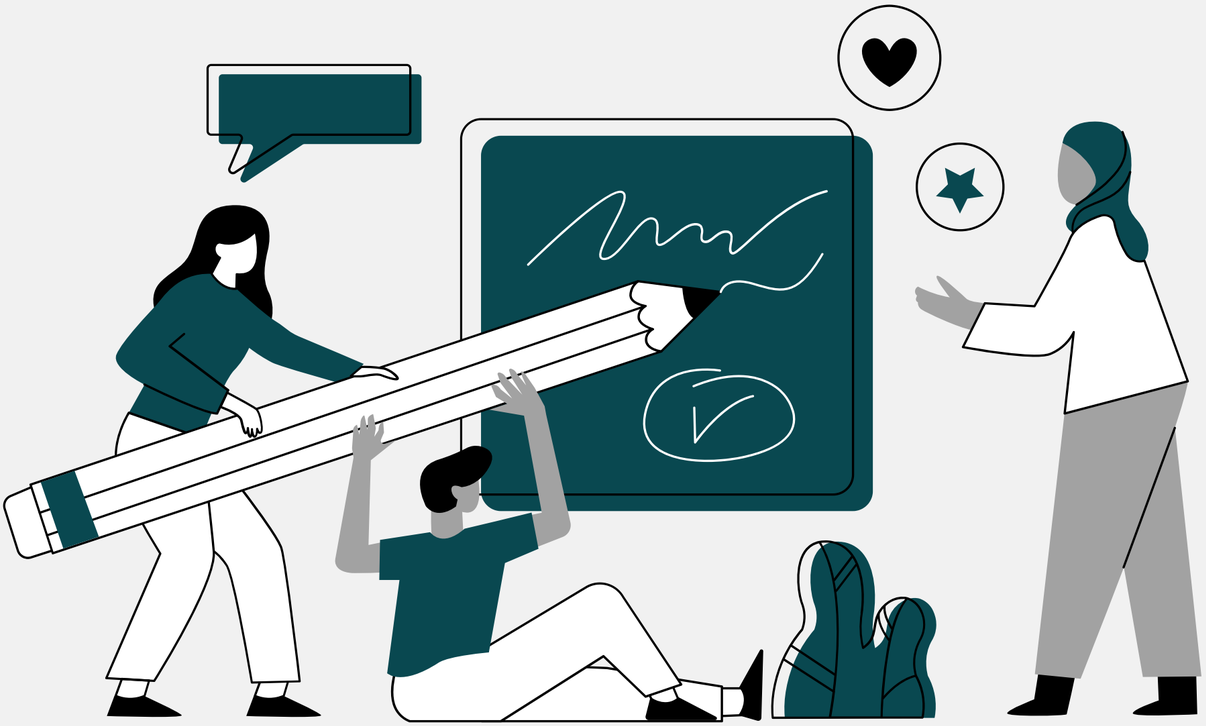
Software Used

Technology and apps that we used to develop an model



REFRENCES

Contact us anytime!



Real-time automatic price prediction for eBay online trading

https://axon.cs.byu.edu/papers/raykhel.iaai09.pdf

Web scrapping laptop prices to estimate hedonic models and extensions to other predictive methods.

hhttps://eventos.fgv.br/sites/eventos.fgv.br/files/arquivos/u161/ottawa_insee_v1.pdf

YouTube

1. https://www.youtube.com/c/CodeWithHarry
2. https://www.youtube.com/channel/UCBwmMxybNva6P_5VmxjzwqA
3. https://www.youtube.com/c/CampusX-official

Articles

https://www.analyticsvidhya.com/blog/2021/11/laptop-price-prediction-practical-understanding-of-machine-learning-project-lifecycle/

THANK YOU SO MUCH !