Automating Car Price Estimates

A solution for Discount Motors

Business Goals

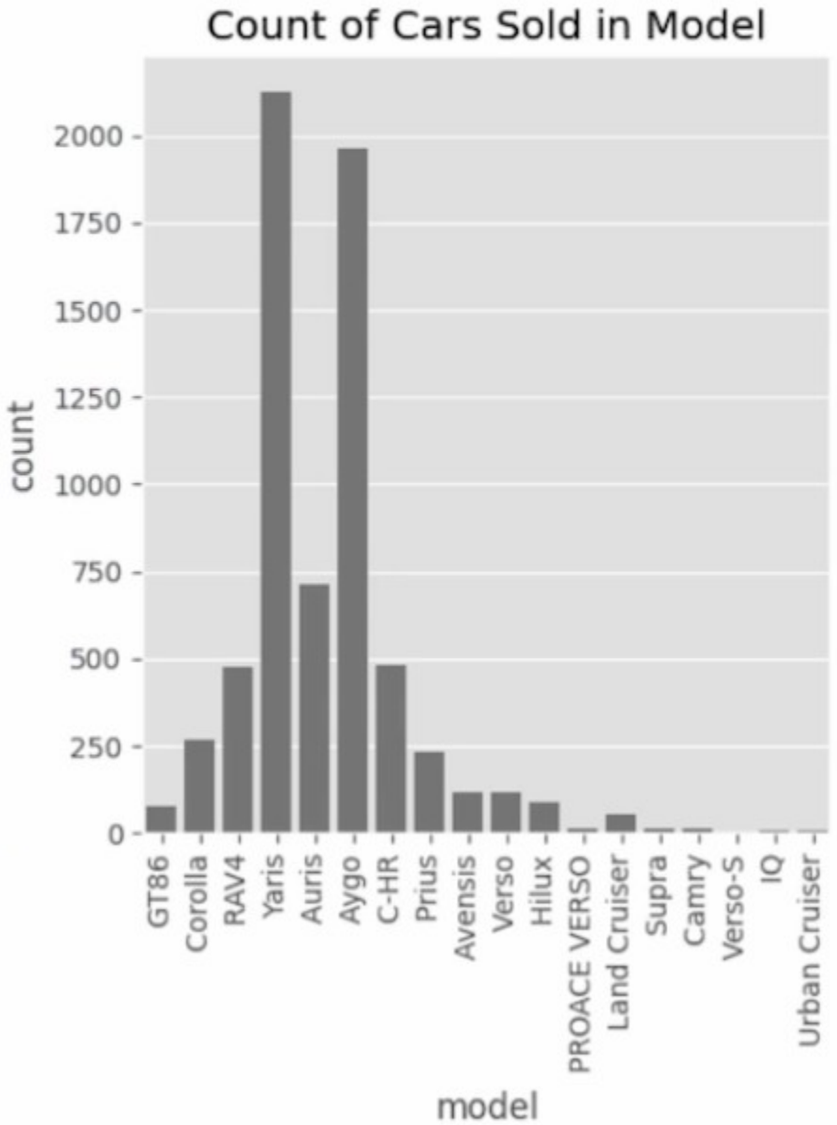
- A solution to replace the most experienced sales team member who is retiring next month
- A solution to automate the entire process to sell cars quicker
- A solution to predict the listed price within 10% the price the used cars would sell (currently about 30% in the team members)

Data

From the web listings of the used cars sold last 6 months.

	model	year	price	transmission	mileage	fuelType	tax	mpg	engineSize
0	GT86	2016	16000	Manual	24089	Petrol	265	36.2	2
1	GT86	2017	15995	Manual	18615	Petrol	145	36.2	2
2	GT86	2015	13998	Manual	27469	Petrol	265	36.2	2
3	GT86	2017	18998	Manual	14736	Petrol	150	36.2	2
4	GT86	2017	17498	Manual	36284	Petrol	145	36.2	2

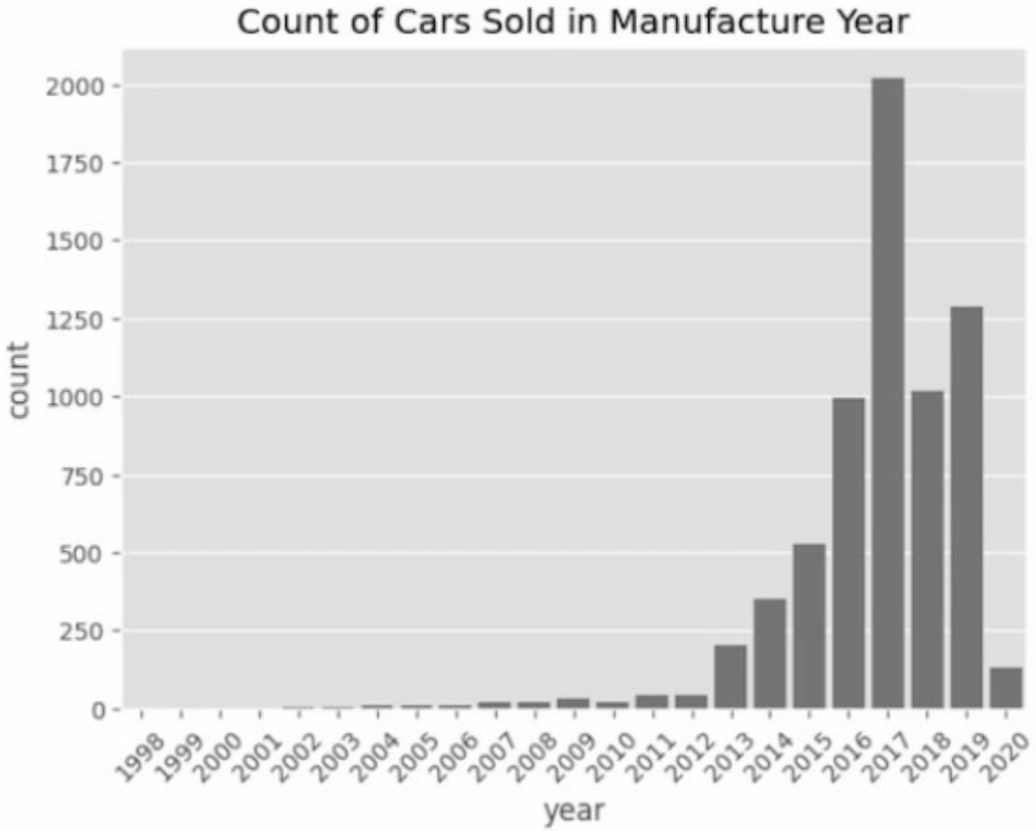
Key Findings



In this dataset, we have 18 car models.

- Yaris
- Arygo
- Auris
- .....

Key Findings

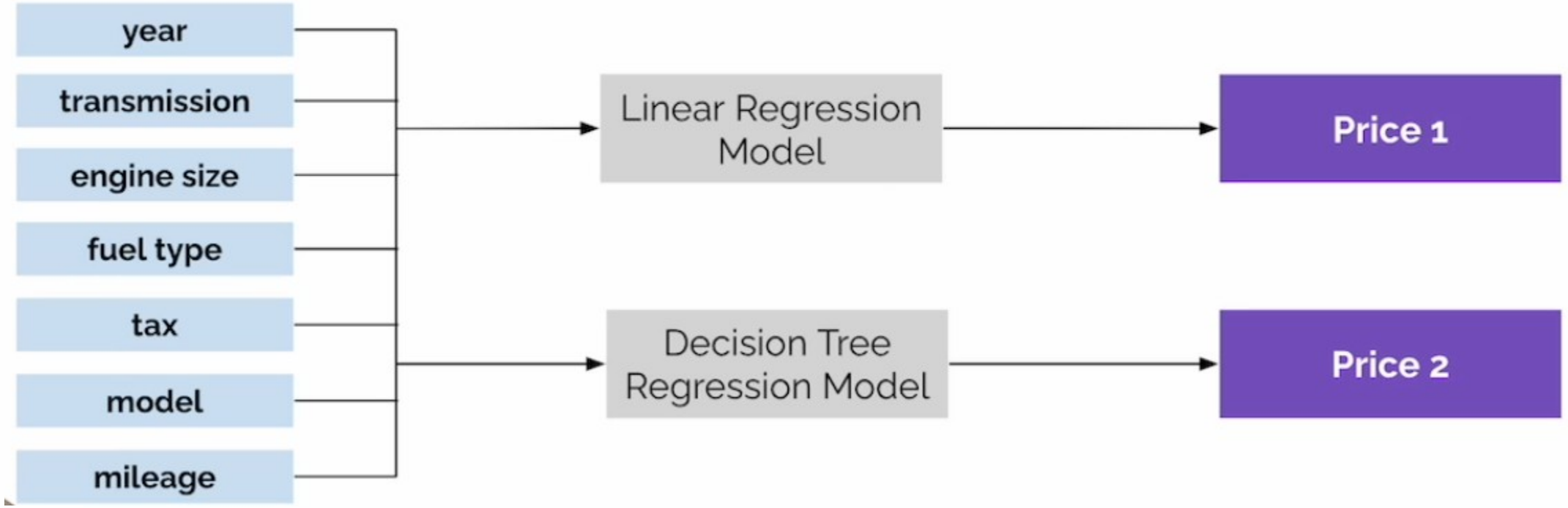


In this dataset, we have cars which are manufactured from 1998 to 2020.

The used cars manufactured in 2016 sold the most in last 6 months.

Outcomes

Two models - Linear Regression and Decision Tree Regression model



Outcomes

Two metrics - R squared and RMSE (Root Mean Squared Error)

Metrics	**R2 Score**	**RMSE (Root Mean Squared Error)**
Purpose	Good fit of data	Accuracy
Range	0 - 1	Depend on the scale of the target variable

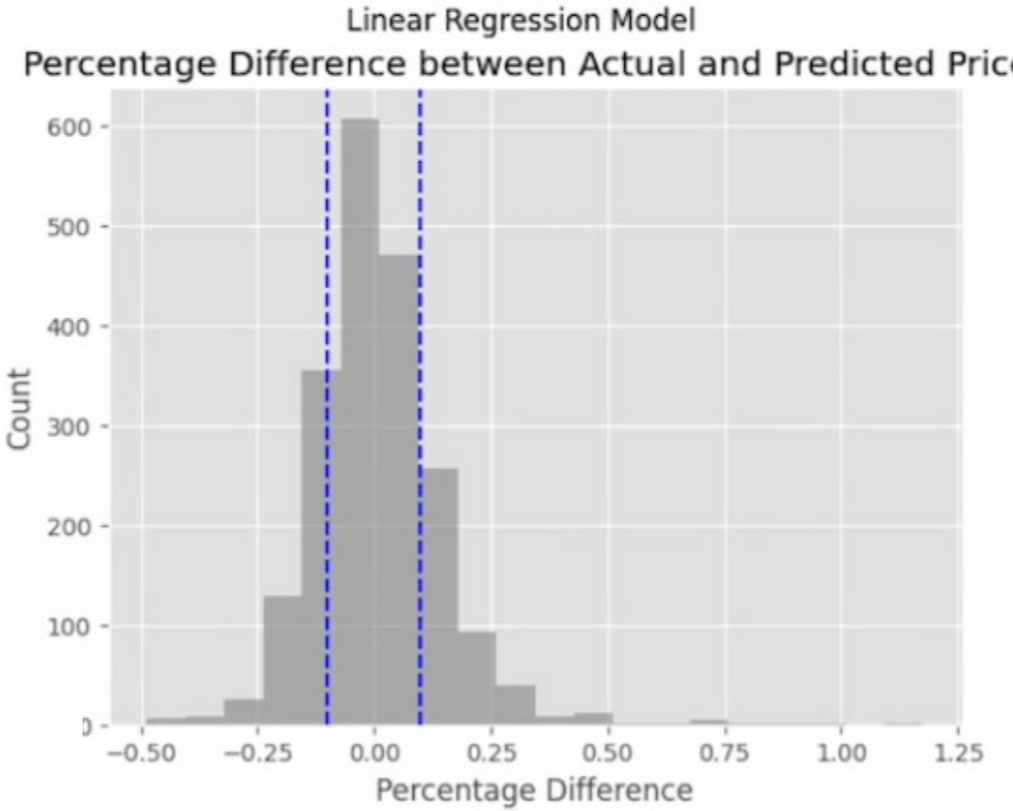
Outcomes

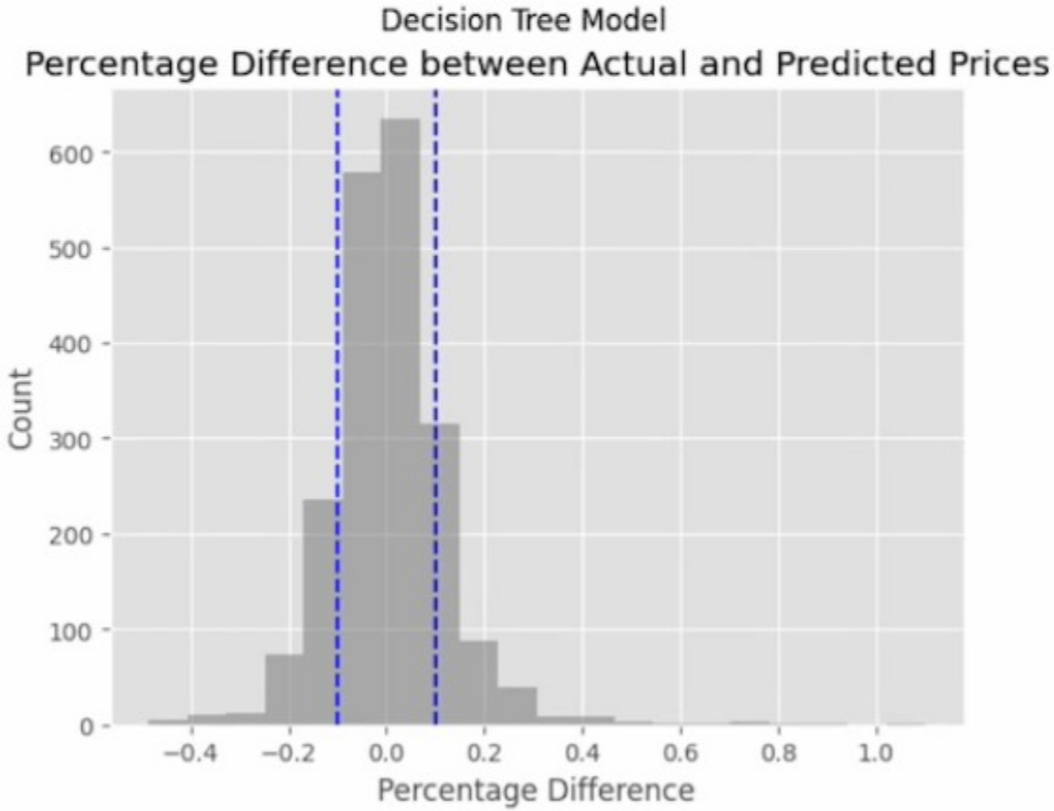
The Decision Tree Regression model is performing better.

Model Name	R2 Score	RMSE (Root Mean Squared Error)
Linear Regression	0.86	2491
Decision Tree Regression	0.94	1611

Outcomes

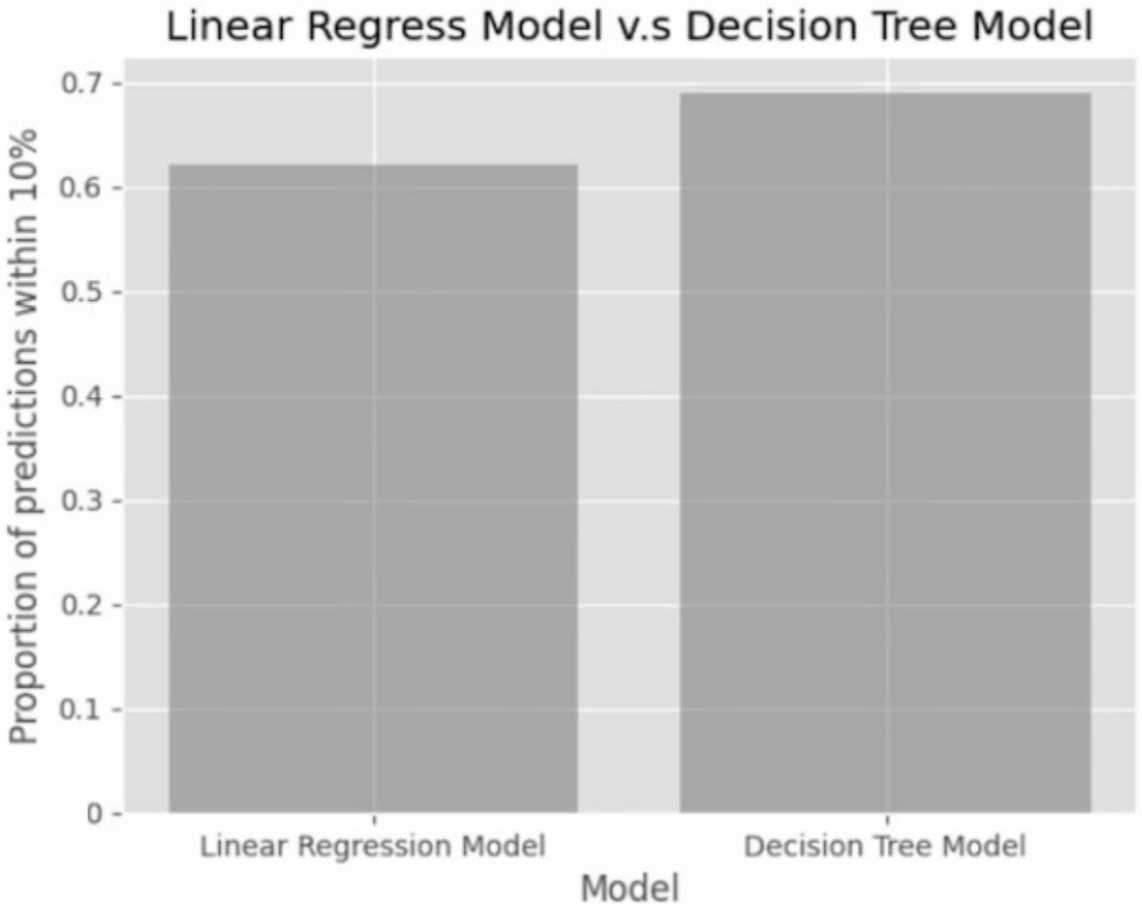
Our KPI - Proportion of the predictions within 10% of the price





Outcomes

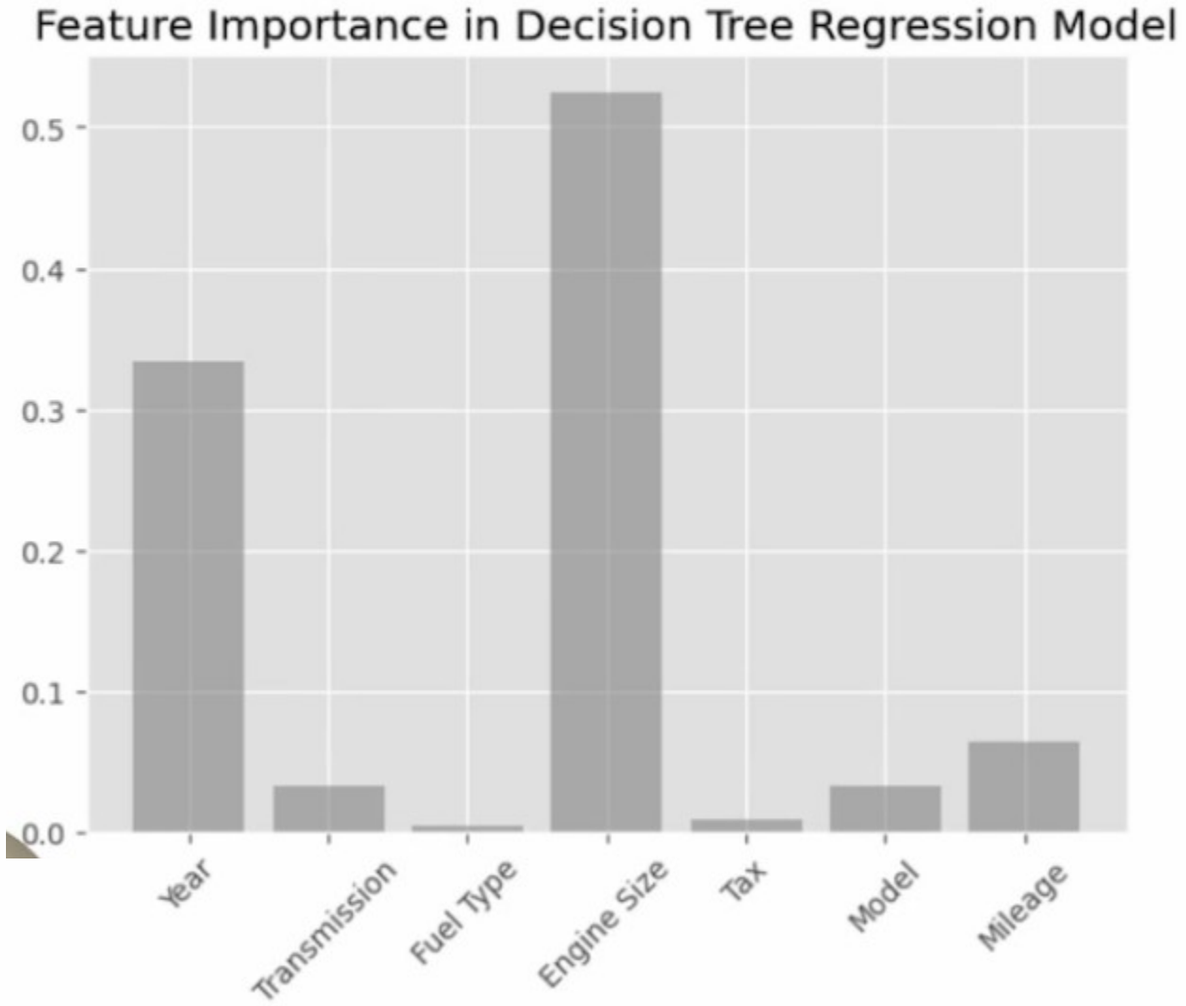
Our KPI - Proportion of the predictions within 10% of the price



Decision Tree Regression Model is performing better.

70% of the prediction within 10% of the price the used cars sell for.

Outcomes



Decision Tree Regression Model

Features

- **Engine Size**
- **Year**
- Model
- Mileage
- Transmission
- Fuel Type
- Tax

Recommendation

- **Test the Decision Tree Regression Model** by comparing its prediction to the estimations of the experienced member on new cars in the recent month.
- **Identify and fix errors** to improve accuracy after testing.
- **Fully deploy the model** using efficient deployment strategy, such as a web or mobile application or API.
- **Continuously improve the model** by collecting more data, feature engineer and fine tuning parameter. This would be the key points to address the limitation from the year and car model variable.