Vehicle Resale Value Prediction

Team Member : Rashika N
Navami B G
Anusha A R
Ashikhatter S
Arpitha M C

INTRODUCTION AND OVERVIEW

- Machine learning is an application of artificial intelligence(AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed.

- Determining whether the listed price of a used car is a challenging task, due to the many factors that drive a used vehicle's price on the market. The focus of this project is developing machine learning models that can accurately predict the price of a used car based on its features, in order to make informed purchases.

- We implement and evaluate various learning methods on a dataset consisting of the sale prices of different makes and models. Our results show that Random Forest model regression yield the best results.

- Multiple regression algorithms like Simple Linear, Multi Linear, Polynomial, Decision Tree, Random Forest were compared and evaluated based on their accuracy to select the best vehicle.
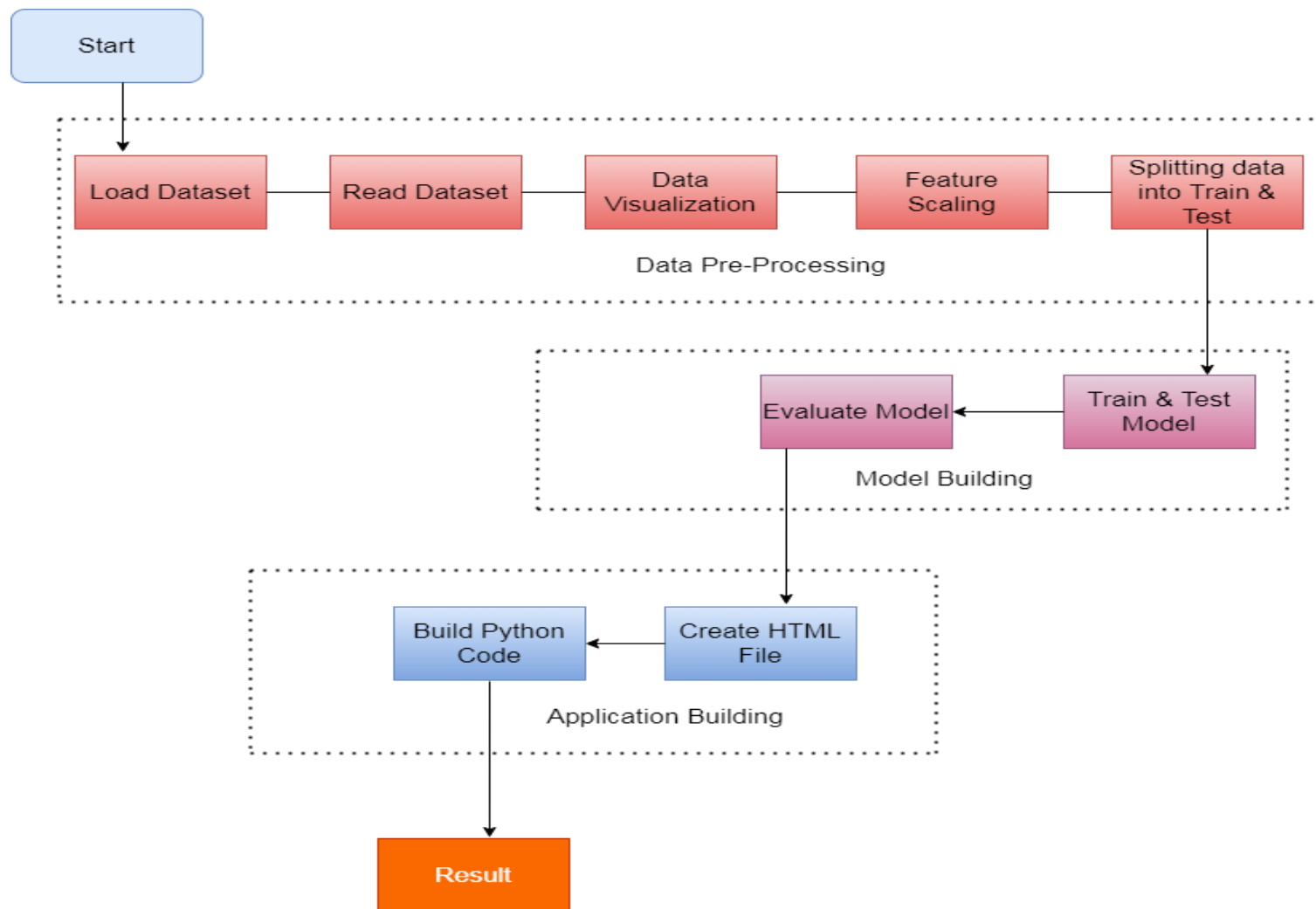
PROJECT DESCRIPTION:

We investigate the application of supervised machine learning techniques to predict the price of used cars. The predictions are based on the data. Different techniques like simple linear regression, multi linear regression, polynomial regression, random tree regression and decision trees have been used to make the predictions. The predictions are then evaluated and compared in order to find those which provide the best performances.

PROPOSED SOLUTION:

In this Project, we discuss the results obtained from applying our selected regressor on the pre-processsed dataset.. All supervised machine learning algorithms applied in this study have been developed using the open-source, object-oriented programming language Python 3.0 and its many packages. The cross-validated accuracy score and cross-validated AUC scores for all the regression algorithms we used on the vehicle resale data.

Flow Chart

Start

Data Pre-Processing

Load Dataset → Read Dataset → Data Visualization → Feature Scaling → Splitting data into Train & Test

Model Building

Evaluate Model ← Train & Test Model

Application Building

Build Python Code ← Create HTML File

Result

Software Design:

- Jupyter Notebook Environment
- Spyder Ide
- Machine learning algorithms.
- Python (pandas, numpy, matplotlib, seaborn, sklearn)
- HTML
- Flask

EXPERIMENTAL INVESTIGATION:

dateCrawl	name	seller	offerType	price	abtest	vehicleTyp	yearOfReg	gearbox	powerPS	model	kilometer	monthOfR	fuelType	brand	notRepaire	dateCreat	nrOfPictur	postalCod	lastSeen
#######	Golf_3_1.(	privat	Angebot	480	test		1993	manuell	0	golf	150000	0	benzin	volkswagen		#######	0	70435	#######
#######	A5_Sportb	privat	Angebot	18300	test	coupe	2011	manuell	190		125000	5	diesel	audi	ja	#######	0	66954	#######
#######	Jeep_Gran	privat	Angebot	9800	test	suv	2004	automatik	163	grand	125000	8	diesel	jeep		#######	0	90480	#######
#######	GOLF_4_1	privat	Angebot	1500	test	kleinwager	2001	manuell	75	golf	150000	6	benzin	volkswage	nein	#######	0	91074	#######
#######	Skoda_Fal	privat	Angebot	3600	test	kleinwager	2008	manuell	69	fabia	90000	7	diesel	skoda	nein	#######	0	60437	#######
#######	BMW_316	privat	Angebot	650	test	limousine	1995	manuell	102	3er	150000	10	benzin	bmw	ja	#######	0	33775	#######
#######	Peugeot_2	privat	Angebot	2200	test	cabrio	2004	manuell	109	2_reihe	150000	8	benzin	peugeot	nein	#######	0	67112	#######
#######	VW_Derby	privat	Angebot	0	test	limousine	1980	manuell	50	andere	40000	7	benzin	volkswage	nein	#######	0	19348	#######
#######	Ford_C__	privat	Angebot	14500	control	bus	2014	manuell	125	c_max	30000	8	benzin	ford		#######	0	94505	#######
#######	VW_Golf_	privat	Angebot	999	test	kleinwager	1998	manuell	101	golf	150000	0		volkswagen		#######	0	27472	#######
#######	Mazda_3_	privat	Angebot	2000	control	limousine	2004	manuell	105	3_reihe	150000	12	benzin	mazda	nein	#######	0	96224	#######
#######	Volkswage	privat	Angebot	2799	control	kombi	2005	manuell	140	passat	150000	12	diesel	volkswage	ja	#######	0	57290	#######
#######	VW_Passa	privat	Angebot	999	control	kombi	1995	manuell	115	passat	150000	11	benzin	volkswagen		#######	0	37269	#######
#######	VW_PASSA	privat	Angebot	2500	control	kombi	2004	manuell	131	passat	150000	2		volkswage	nein	#######	0	90762	#######
#######	Nissan_Na	privat	Angebot	17999	control	suv	2011	manuell	190	navara	70000	3	diesel	nissan	nein	#######	0	4177	#######
#######	KA_Luftha	privat	Angebot	450	test	kleinwager	1910		0	ka	5000	0	benzin	ford		#######	0	24148	#######
#######	Polo_6n_1	privat	Angebot	300	test		2016		60	polo	150000	0	benzin	volkswagen		#######	0	38871	#######
#######	Renault_T	privat	Angebot	1750	control	kleinwager	2004	automatik	75	twingo	150000	2	benzin	renault	nein	#######	0	65599	#######
#######	Ford_C_M	privat	Angebot	7550	test	bus	2007	manuell	136	c_max	150000	6	diesel	ford	nein	#######	0	88361	#######
#######	Mercedes_	privat	Angebot	1850	test	bus	2004	manuell	102	a_klasse	150000	1	benzin	mercedes_	nein	#######	0	49565	#######
#######	Volkswage	privat	Angebot	10400	control	coupe	2009	manuell	160	scirocco	100000	4	benzin	volkswage	nein	#######	0	75365	#######
#######	BMW_530	privat	Angebot	3699	test	limousine	2002	automatik	231	5er	150000	7	benzin	bmw	nein	#######	0	68309	#######
#######	Opel_Meri	privat	Angebot	2900	test		2018	manuell	90	meriva	150000	5	benzin	opel	nein	#######	0	49716	#######
#######	Stadtflitze	privat	Angebot	450	test	kleinwager	1997	manuell	50	arosa	150000	5	benzin	seat	nein	#######	0	9526	#######
#######	MERCEDE!	privat	Angebot	500	test	limousine	1990	manuell	118	andere	150000	10	benzin	mercedes_	ja	#######	0	35390	#######
#######	BMW_530	privat	Angebot	2500	control	kombi	2002	automatik	193	5er	150000	9	diesel	bmw	ja	#######	0	73765	#######

HEAT MAP:
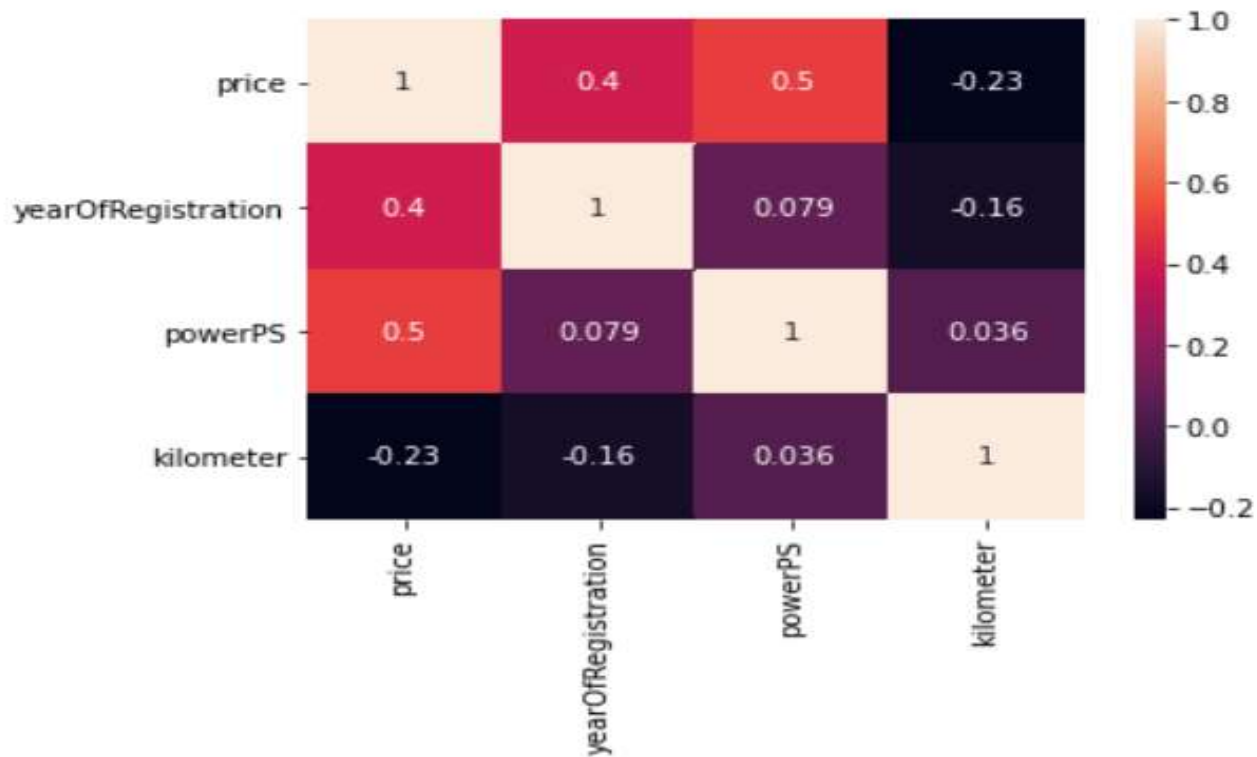
Model Building and Evaluation:

Here random forest regression is used to predict. Accuracy is evaluated.

```
In [51]:  #random forest regression
          from sklearn.ensemble import RandomForestRegressor
          rf=RandomForestRegressor(n_estimators=100,criterion='mse',random_state=0)
```

```
In [52]:  rf.fit(x_train,y_train)
```

```
Out[52]:  RandomForestRegressor(random_state=0)
```

```
In [53]:  joblib.dump(rf,"random.save")
```

```
Out[53]:  ['random.save']
```

```
In [54]:  y_pred=rf.predict(x_test)
          y_pred
```

```
Out[54]:  array([ 928.15125   , 2177.1600974 , 6564.15457131, ...,  566.04522143,
                 7661.12527882,  831.72857143])
```

```
In [55]:  y_test
```

```
Out[55]:  array([ 500, 3850, 6500, ...,  500, 7850, 1890], dtype=int64)
```

```
In [56]:  from sklearn.metrics import r2_score
          r2_score(y_test,y_pred)
```

```
Out[56]:  0.8128590913835658
```

CONCLUSION:

- The main objective of this research was to develop a prototype of the system that can be used by the buyer.

- Multiple machine learning algorithms were developed and used for this research. Random forest regression proved to best-fit for development of the vehicle resale system when compared by giving highest accuracy.

THANK YOU