Insurance Prediction Web Application

By: Yee Hang

Dataset Info,

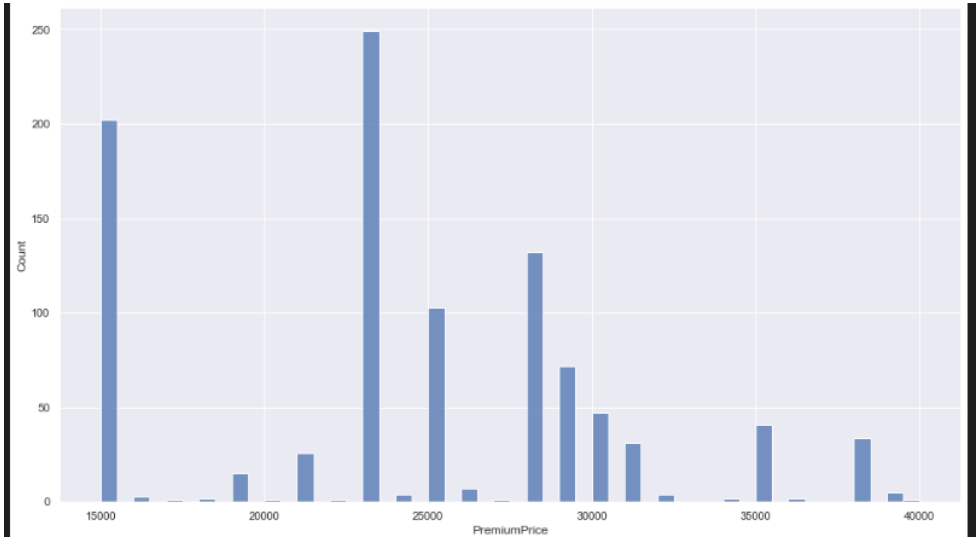
	Age	Diabetes	BloodPressureProblems	AnyTransplants	AnyChronicDiseases	Height	Weight	KnownAllergies	HistoryOfCancerInFamily	NumberOfMajorSurgeries	PremiumPrice
0	45	0	0	0	0	155	57	0	0	0	25000
1	60	1	0	0	0	180	73	0	0	0	29000
2	36	1	1	0	0	158	59	0	0	1	23000
3	52	1	1	0	1	183	93	0	0	2	28000
4	38	0	0	0	1	166	88	0	0	1	23000
...	...	...	...	...	...	...	...	...	...	...	...
981	18	0	0	0	0	169	67	0	0	0	15000
982	64	1	1	0	0	153	70	0	0	3	28000
983	56	0	1	0	0	155	71	0	0	1	29000
984	47	1	1	0	0	158	73	1	0	1	39000
985	21	0	0	0	0	158	75	1	0	1	15000

986 rows × 11 columns

Insurance Data:

- Contains health information and the insurance price they have to pay yearly

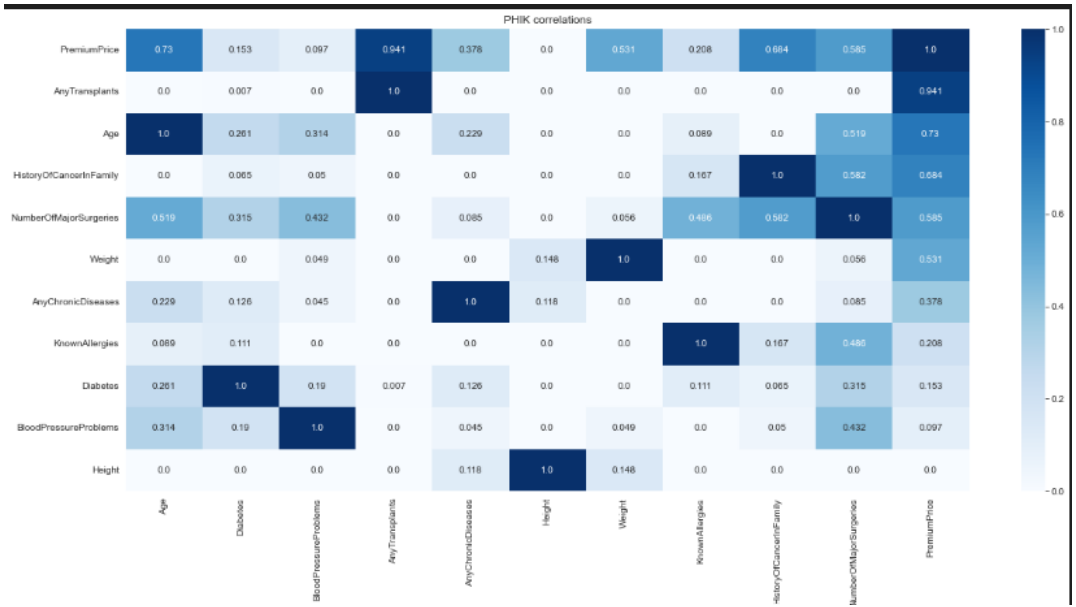
- Nowadays, a lot of insurance companies and agents offering plans
- The web application is useful for customers who want to buy health insurance but want to know the approximate premium they have to pay



Price distribution

EDA phik correlation

- Any transplants have highest correlation with the target
- - Followed by age and the history of transplant in family
- Height has small amount of correlation hence, dropped from data



Feature Engineering

Data Engineering - BMI

> weight (kg) / [height (m)]**2

According to the NHS, BMI indicates whether someone is obese. Insurance companies use this information , to determine the cost of the premium as obese have higher chance of dying.

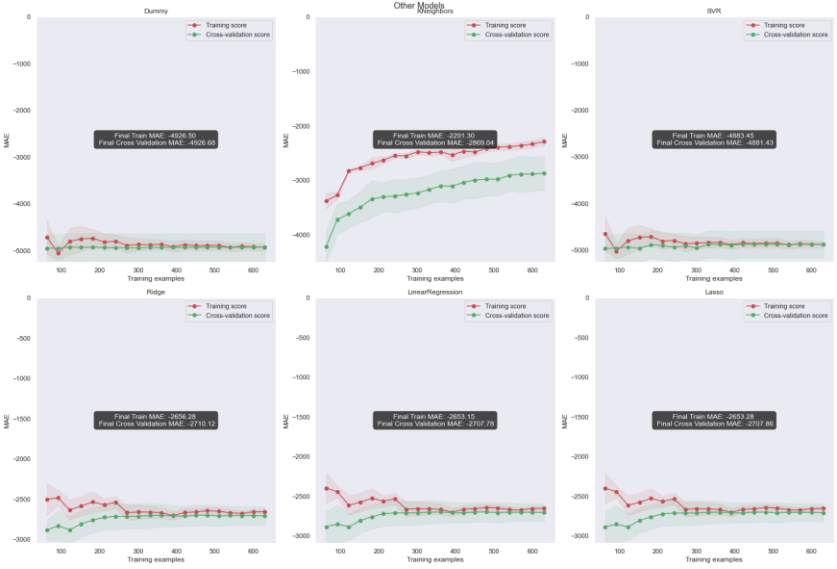
BMI range	Indication
BMI < 18.5	Underweight
18.5 < BMI < 24.9	Healthy
25 < BMI < 29.9	Overweight
30 < BMI < 39.9	Obese

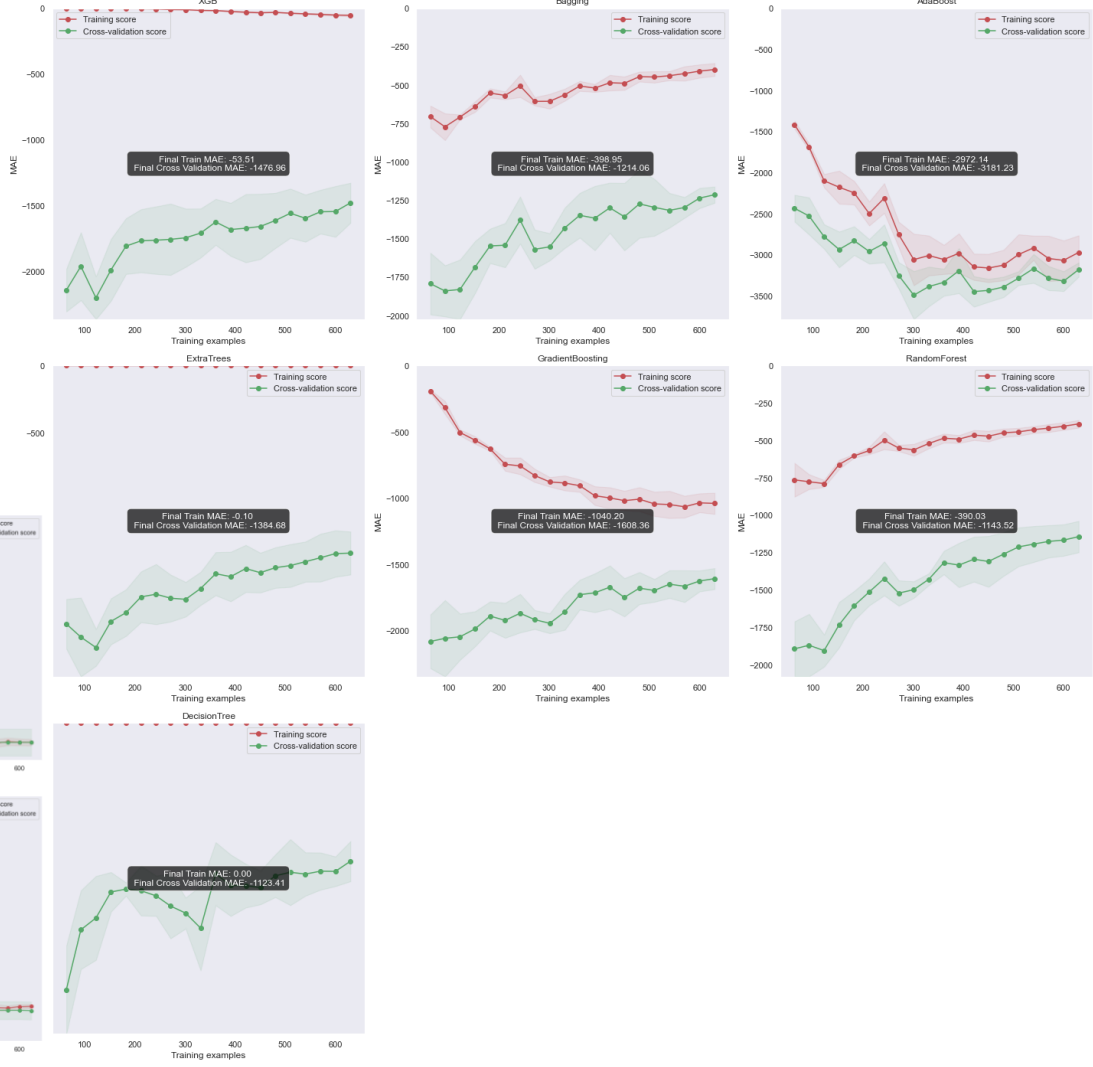
```python
data['bmi'] =data.Weight / ((data.Height/100)**2 )
data.drop(columns = [ 'Height'], inplace = True)
```

Python

Models

- Testing and evaluation
- Decision Tree, best validation score





Tuning and results

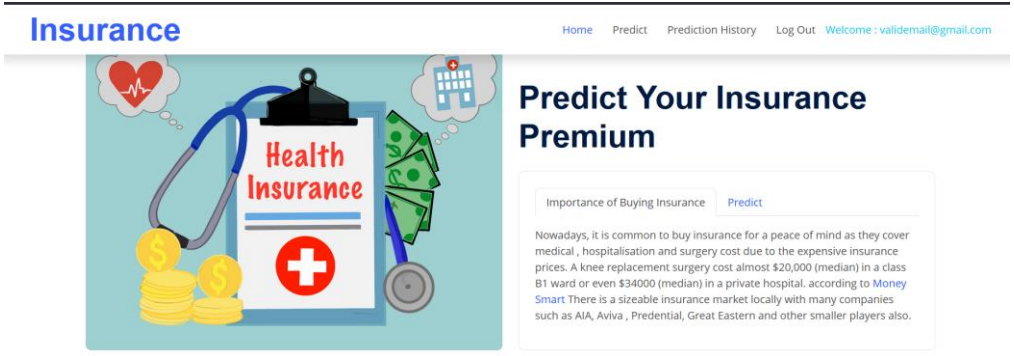
- After tuning the decision tree, Achieved better validation score
- Final MAE is 1271
- although the MAE, looks very high, the percentage error on the test set is 5% which is decent
- - Thus, such predictions will also likely to be useful
- Mean average percentage error is 5%

```
mae 1271.7171717171716

mape 0.05637117797151796
```

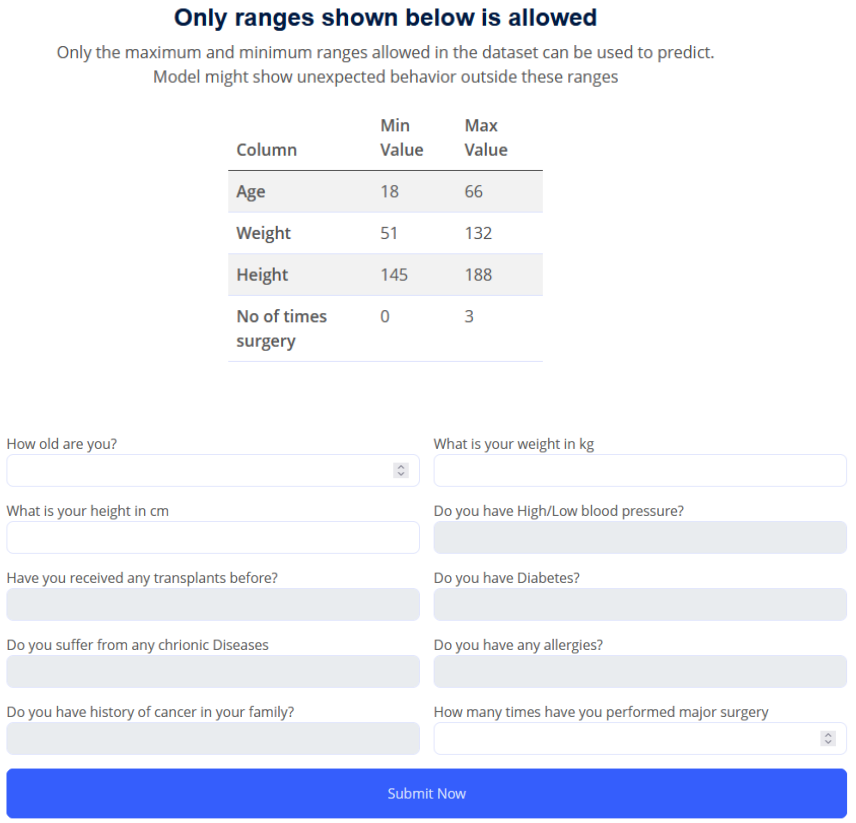
Application – Front end - Homepage

- Convince user why important to buy insurance
  - Peace of mind
- Convince user why get a prediction first
  - Avoid buying plans that don’t worth/too expensive



Application – Prediction Page

- Show User allowed values for numerical features
  - Easier for user to make a prediction
- Use Dropdown instead of radio button to save space and more aesthetically pleasing (personal view)



Only ranges shown below is allowed

Only the maximum and minimum ranges allowed in the dataset can be used to predict.
Model might show unexpected behavior outside these ranges

Column	Min Value	Max Value
Age	18	66
Weight	51	132
Height	145	188
No of times surgery	0	3

How old are you?

What is your weight in kg

What is your height in cm

Do you have High/Low blood pressure?

Have you received any transplants before?

Do you have Diabetes?

Do you suffer from any chrionic Diseases

Do you have any allergies?

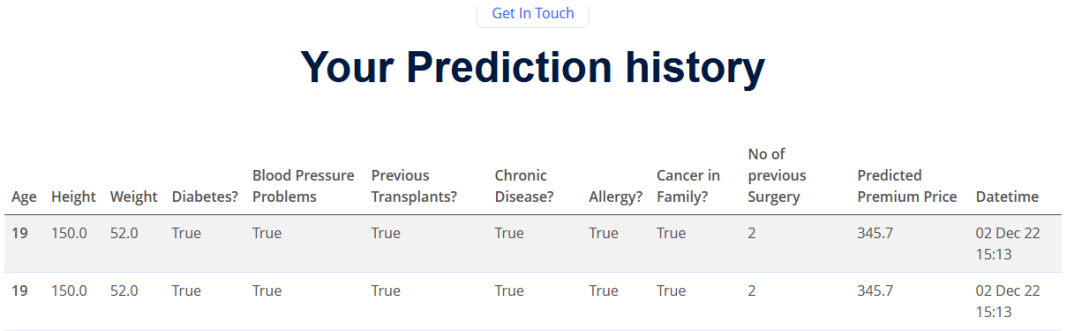
Do you have history of cancer in your family?

How many times have you performed major surgery

Submit Now

Application – Prediction History Page

- Show the history of the prediction with a date and time



Application – Testing

- Expected fail test, test inserting invalid input
- Consistency testing, whether model return same result every time same set of input given
- Range testing, different ranges of inputs are given