Project2_3L_NN(Expedia)

April 3, 2022

```
********************** Exploratory Data Analysis ************************


------------------ Displaying head ------------------


      id     region  latitude  longitude accommodation_type  cost  \
0  13232  Manhattan  40.71854  -74.00439    Entire home/apt   170
1    246   Brooklyn  40.64446  -73.95030    Entire home/apt    65
2  19091     Queens  40.78573  -73.81062       Private room    85
3  34305  Manhattan  40.73863  -73.98002       Private room   210
4    444  Manhattan  40.82426  -73.94630        Shared room    75

   minimum_nights  number_of_reviews  reviews_per_month   owner_id  \
0               5                  7               0.56     929983
1               3                238               2.30     281764
2               1                  0                NaN   19923341
3              30                  0                NaN  200380610
4               3                 38               0.42     745069

   owned_hotels  yearly_availability
0             1                    0
1             1                    0
2             1                    1
3            65                    1
4             3                    1


------------------ Dimensions of dataset ------------------


(2870, 12)


------------------ Contents of dataset  ------------------

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2870 entries, 0 to 2869
Data columns (total 12 columns):
 #   Column              Non-Null Count  Dtype
```

```
---  ------                --------------  -----
 0   id                    2870 non-null   int64
 1   region                2870 non-null   object
 2   latitude              2870 non-null   float64
 3   longitude             2870 non-null   float64
 4   accommodation_type    2870 non-null   object
 5   cost                  2870 non-null   int64
 6   minimum_nights        2870 non-null   int64
 7   number_of_reviews     2870 non-null   int64
 8   reviews_per_month     2194 non-null   float64
 9   owner_id              2870 non-null   int64
 10  owned_hotels          2870 non-null   int64
 11  yearly_availability   2870 non-null   int64
dtypes: float64(3), int64(7), object(2)
memory usage: 269.2+ KB
```

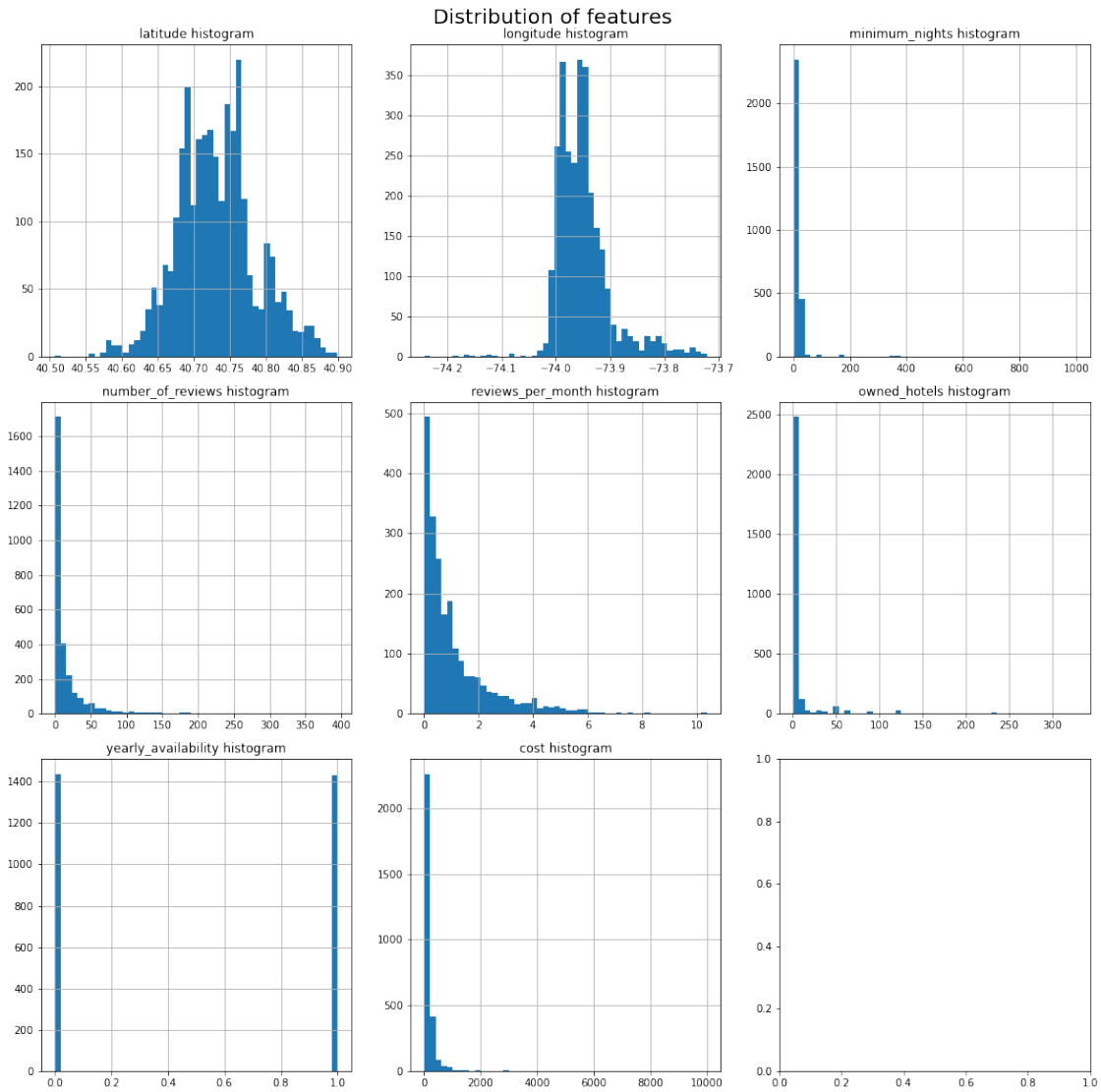
None

---------- Summary of Numerical feature ----------

	Feature_name	datatype	Count	min	quartile1	Mean	Median	quartile3	max	Std dev	Skewness	Kurtosis	Range	IQR	skewness comment
0	id	int64	2870	0.00000	1.593175e+04	2.676066e+04	2.894650e+04	3.847850e+04	4.889300e+04	14140.93	-0.31	-1.08	4.889300e+04	2.254675e+04	Fairly symmetric(negative)
1	latitude	float64	2870	40.50708	4.069246e+01	4.073122e+01	4.072825e+01	4.076266e+01	4.089873e+01	0.05	0.17	0.21	3.916500e-01	7.019500e-02	Fairly symmetric(positive)
2	longitude	float64	2870	-74.24285	-7.398400e+01	-7.395016e+01	-7.395672e+01	-7.393420e+01	-7.372173e+01	0.05					
3	cost	int64	2870	10.00000	7.500000e+01	1.959432e+02	1.200000e+02	2.000000e+02	9.999000e+03	406.18					
4	minimum_nights	int64	2870	1.00000	1.000000e+00	1.153031e+01	3.000000e+00	6.000000e+00	9.990000e+02	37.97					
5	number_of_reviews	int64	2870	0.00000	1.000000e+00	1.631533e+01	4.000000e+00	1.600000e+01	3.950000e+02	32.48					
6	reviews_per_month	float64	2194	0.01000	2.400000e-01	1.157502e+00	6.500000e-01	1.530000e+00	1.037000e+01	1.36					
7	owner_id	int64	2870	2787.00000	7.388002e+06	7.202195e+07	3.352708e+07	1.207625e+08	2.738123e+08	80765157.57					
8	owned_hotels	int64	2870	1.00000	1.000000e+00	8.411498e+00	1.000000e+00	3.000000e+00	3.270000e+02	27.11					
9	yearly_availability	int64	2870	0.00000	0.000000e+00	4.986063e-01	0.000000e+00	1.000000e+00	1.000000e+00	0.50					

```
2      1.36      4.43  5.211200e-01  4.980000e-02        High positive skewed
3     13.01    232.35  9.989000e+03  1.250000e+02        High positive skewed
4     11.87    210.77  9.980000e+02  5.000000e+00        High positive skewed
5      4.27     25.44  3.950000e+02  1.500000e+01        High positive skewed
6      2.16      5.81  1.036000e+01  1.290000e+00        High positive skewed
7      1.05     -0.25  2.738095e+08  1.133744e+08        High positive skewed
8      6.95     62.60  3.260000e+02  2.000000e+00        High positive skewed
9      0.01     -2.00  1.000000e+00  1.000000e+00  Fairly symmetric(positive)

  outlier comment
0     No outliers
1    Has outilers
2    Has outilers
3    Has outilers
4    Has outilers
5    Has outilers
6    Has outilers
7     No outliers
8    Has outilers
9     No outliers
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2870 entries, 0 to 2869
Data columns (total 10 columns):
 #   Column               Non-Null Count  Dtype
---  ------               --------------  -----
 0   region               2870 non-null   object
 1   latitude             2870 non-null   float64
 2   longitude            2870 non-null   float64
 3   accommodation_type   2870 non-null   object
 4   cost                 2870 non-null   int64
 5   minimum_nights       2870 non-null   int64
 6   number_of_reviews    2870 non-null   int64
 7   reviews_per_month    2194 non-null   float64
 8   owned_hotels         2870 non-null   int64
 9   yearly_availability  2870 non-null   int64
dtypes: float64(3), int64(5), object(2)
memory usage: 224.3+ KB

------------------ Summary of Numerical feature  ------------------

          Feature_name datatype  Count       min  quartile1        Mean  \
0             latitude  float64   2870  40.50708  40.692463   40.731224
1            longitude  float64   2870 -74.24285 -73.984003  -73.950158
2                 cost    int64   2870  10.00000  75.000000  195.943206
3       minimum_nights    int64   2870   1.00000   1.000000   11.530314
4    number_of_reviews    int64   2870   0.00000   1.000000   16.315331
```

```
5    reviews_per_month  float64   2194   0.01000   0.240000    1.157502
6         owned_hotels    int64   2870   1.00000   1.000000    8.411498
7  yearly_availability    int64   2870   0.00000   0.000000    0.498606

      Median   quartile3         max  Std dev  Skewness  Kurtosis       Range  \
0   40.72825   40.762658    40.89873     0.05      0.17      0.21     0.39165
1  -73.95672  -73.934202   -73.72173     0.05      1.36      4.43     0.52112
2  120.00000  200.000000  9999.00000   406.18     13.01    232.35  9989.00000
3    3.00000    6.000000   999.00000    37.97     11.87    210.77   998.00000
4    4.00000   16.000000   395.00000    32.48      4.27     25.44   395.00000
5    0.65000    1.530000    10.37000     1.36      2.16      5.81    10.36000
6    1.00000    3.000000   327.00000    27.11      6.95     62.60   326.00000
7    0.00000    1.000000     1.00000     0.50      0.01     -2.00     1.00000

          IQR            skewness comment outlier comment
0    0.070195  Fairly symmetric(positive)    Has outilers
1    0.049800        High positive skewed    Has outilers
2  125.000000        High positive skewed    Has outilers
3    5.000000        High positive skewed    Has outilers
4   15.000000        High positive skewed    Has outilers
5    1.290000        High positive skewed    Has outilers
6    2.000000        High positive skewed    Has outilers
7    1.000000  Fairly symmetric(positive)     No outliers

None
```

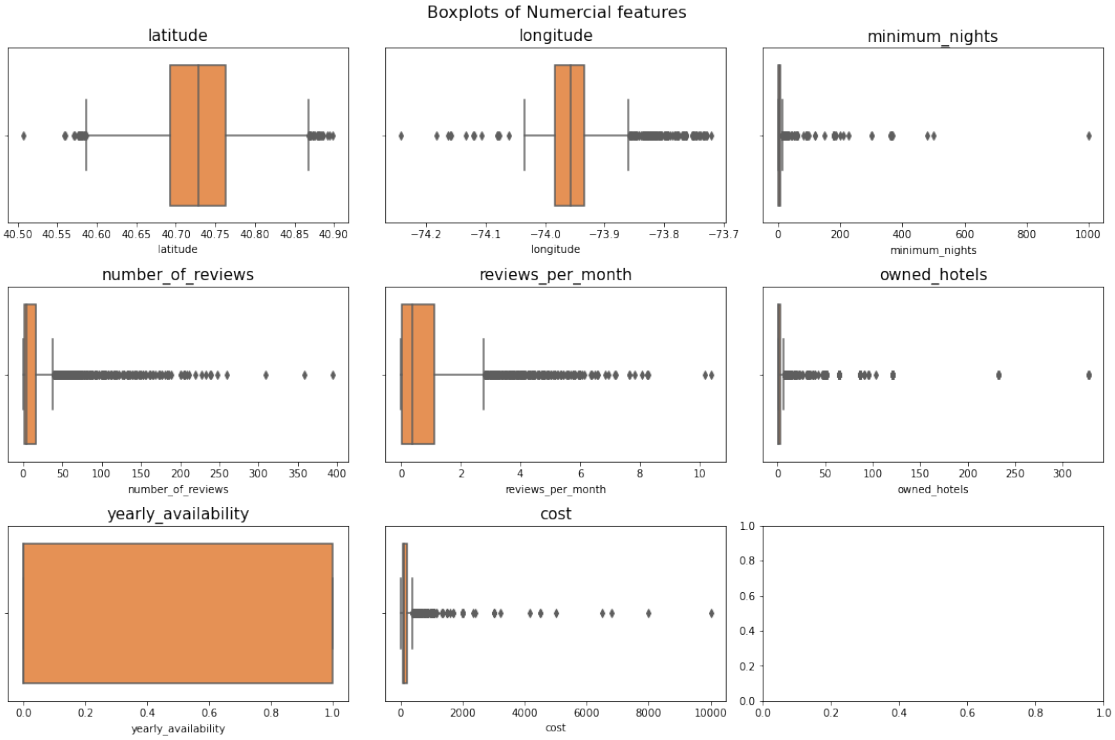


[9]:
```
latitude                0
longitude               0
minimum_nights          0
number_of_reviews       0
reviews_per_month     676
owned_hotels            0
yearly_availability     0
cost                    0
dtype: int64
```

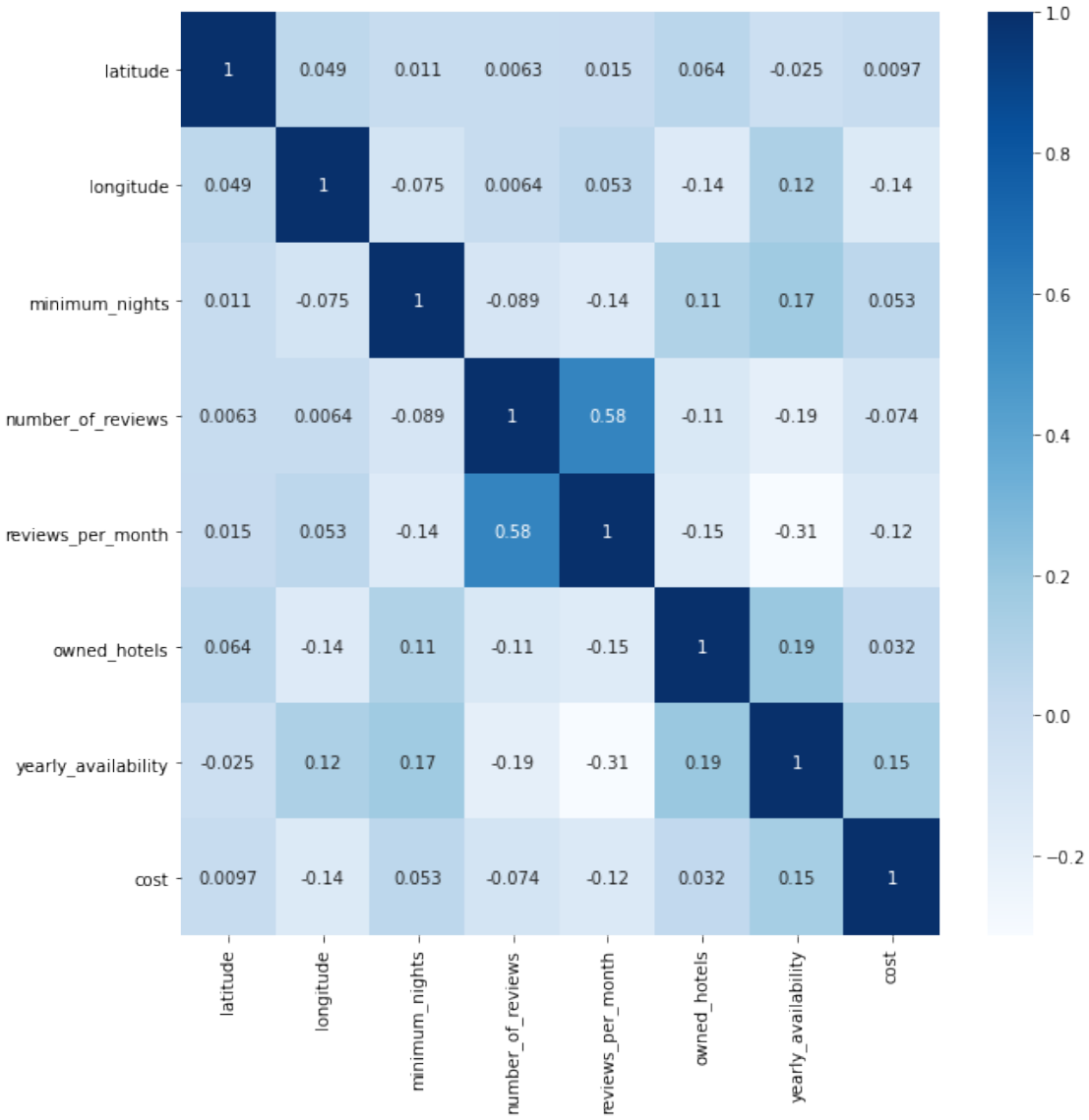
[10]:
```
latitude              0
longitude             0
minimum_nights        0
```

```
number_of_reviews      0
reviews_per_month      0
owned_hotels           0
yearly_availability    0
cost                   0
dtype: int64
```

```
None
```



[12]: <AxesSubplot:>



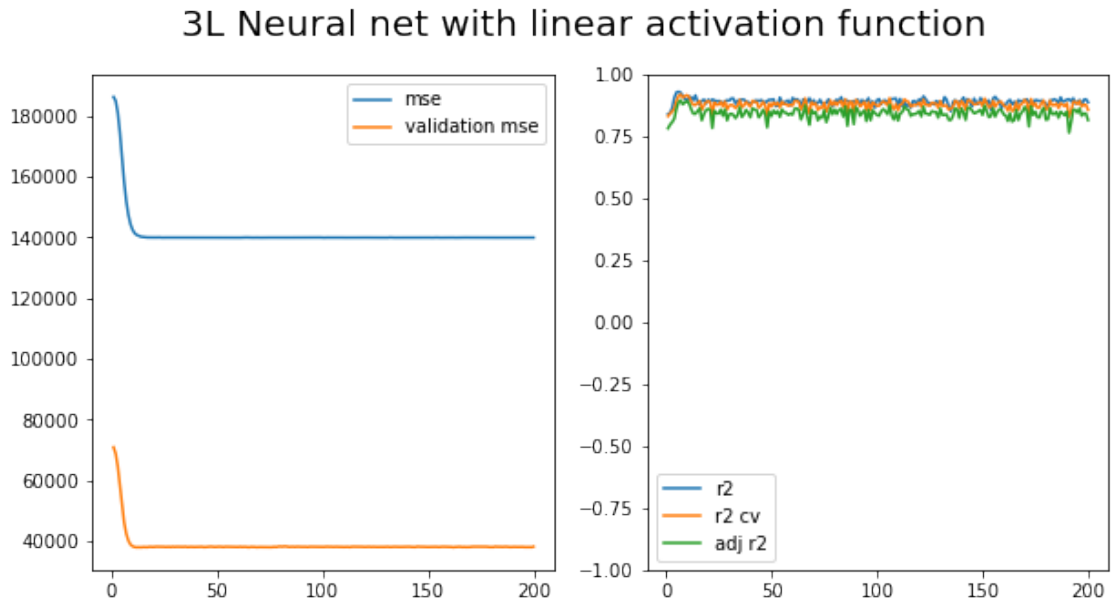
[14]:

	latitude	longitude	minimum_nights	number_of_reviews	\
0	40.71854	-74.00439	5	7	
1	40.64446	-73.95030	3	238	
2	40.78573	-73.81062	1	0	
3	40.73863	-73.98002	30	0	
4	40.82426	-73.94630	3	38	
…	…	…	…	…	
2865	40.74316	-73.98038	2	0	
2866	40.73523	-73.99465	3	2	

```
2867  40.76619  -73.98987               3                  17
2868  40.74637  -73.97207              30                   0
2869  40.79208  -73.96482              30                  24

      reviews_per_month  owned_hotels  yearly_availability  cost
0                  0.56             1                    0   170
1                  2.30             1                    0    65
2                  0.00             1                    1    85
3                  0.00            65                    1   210
4                  0.42             3                    1    75
…                   …               …                    …    …
2865               0.00             1                    1   400
2866               0.07             1                    1   180
2867               0.67             1                    0   179
2868               0.00            49                    1   200
2869               0.33            11                    1  1000

[2870 rows x 8 columns]
```

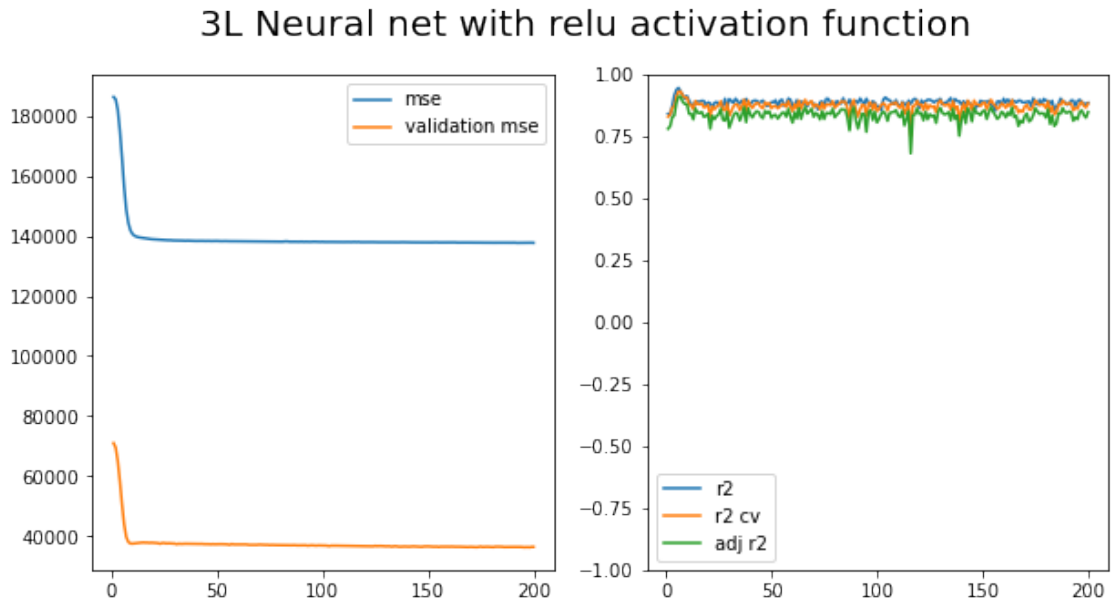
[18]:
```
['latitude',
 'longitude',
 'number_of_reviews',
 'reviews_per_month',
 'yearly_availability']
```

Building Neural nets with linear activation function



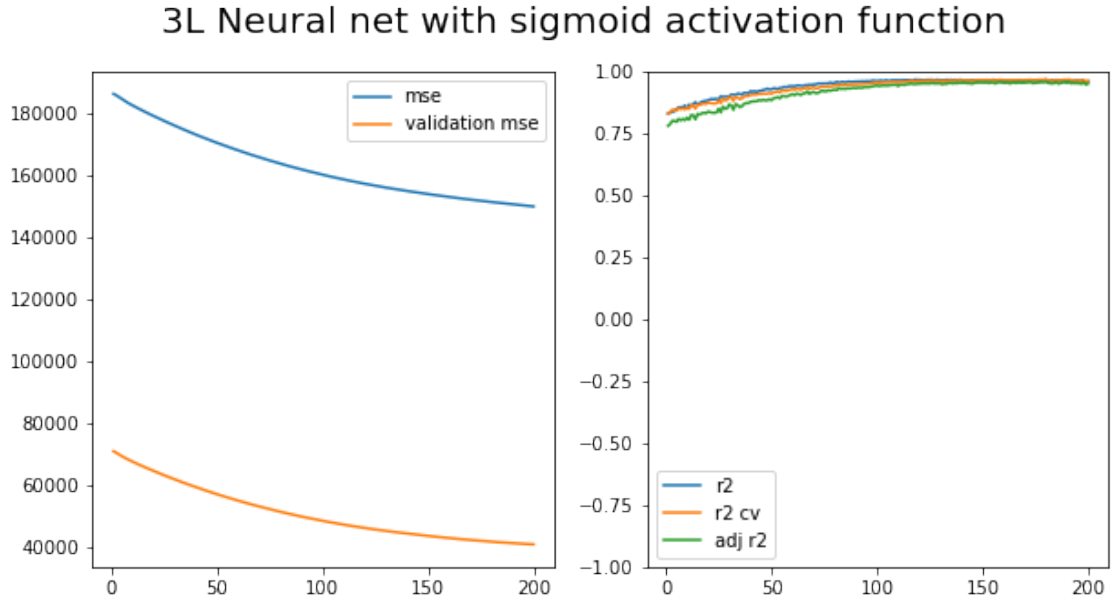
```
**********************************************************
```

```
Building Neural nets with relu activation function
```



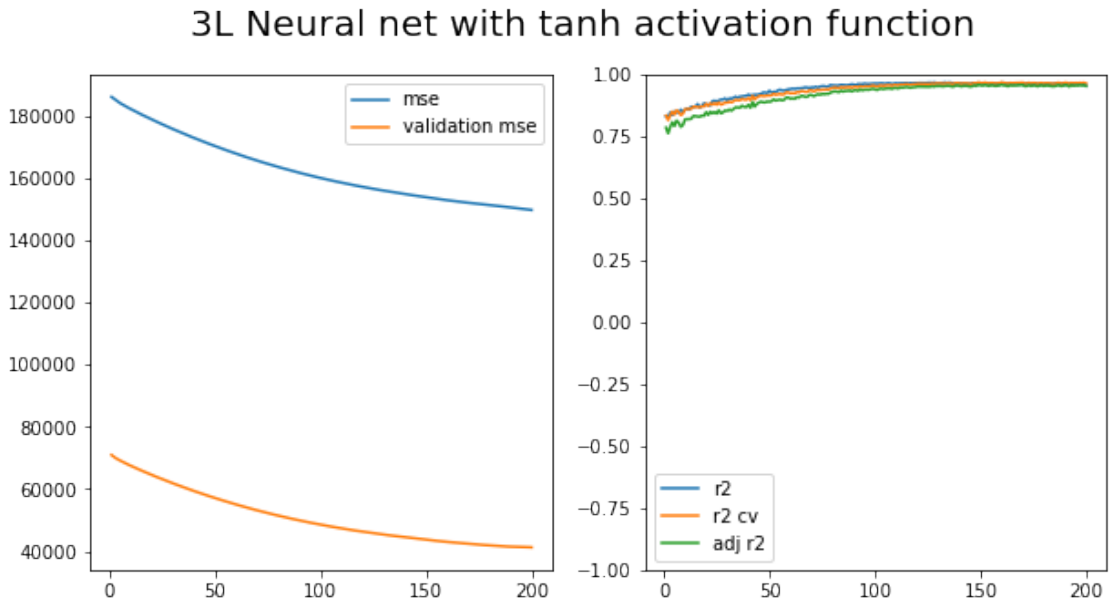

```
**********************************************************
```

```
Building Neural nets with sigmoid activation function
```



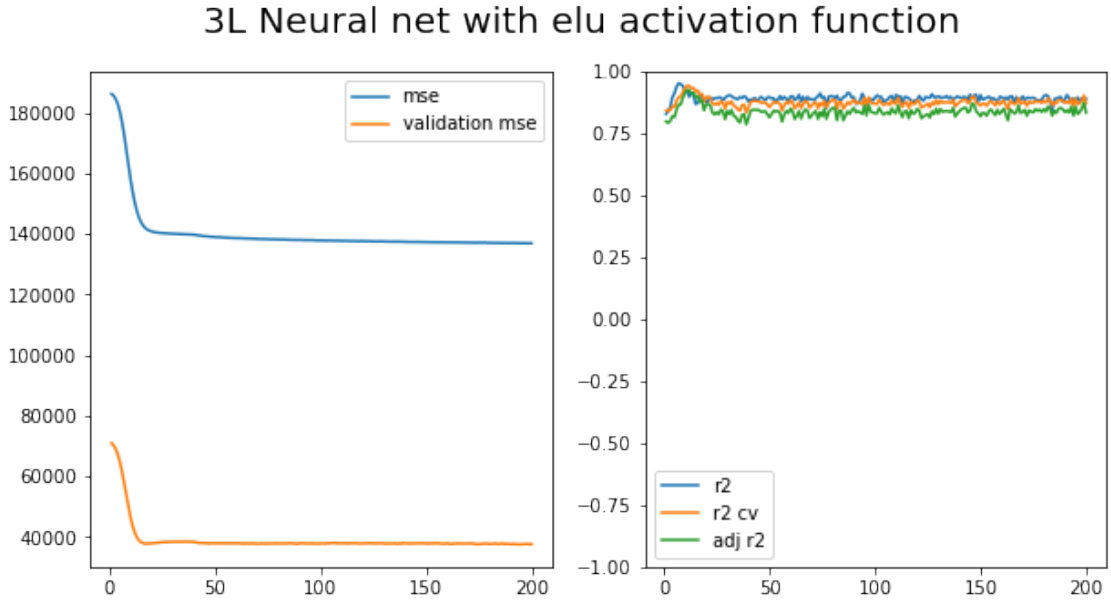
```
**********************************************************
```

```
Building Neural nets with tanh activation function
```



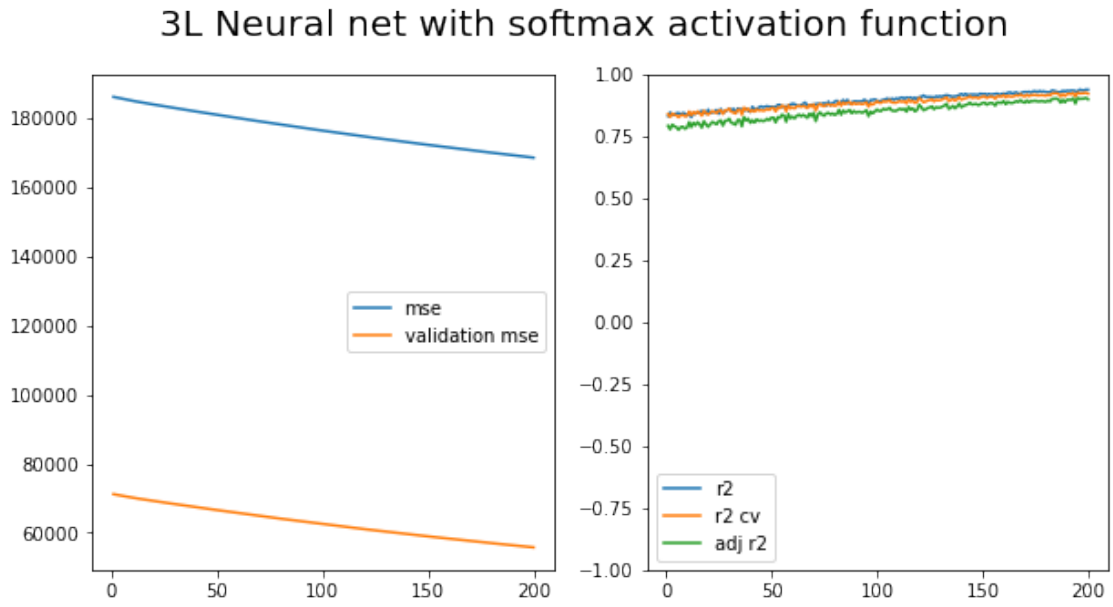

```
**********************************************************
```

```
Building Neural nets with elu activation function
```



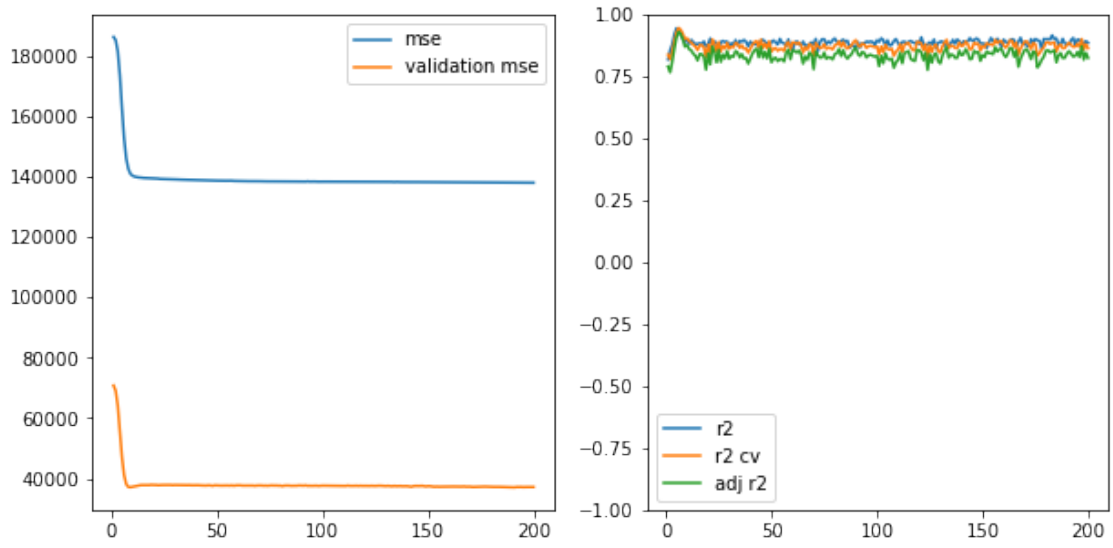
****

Building Neural nets with softmax activation function



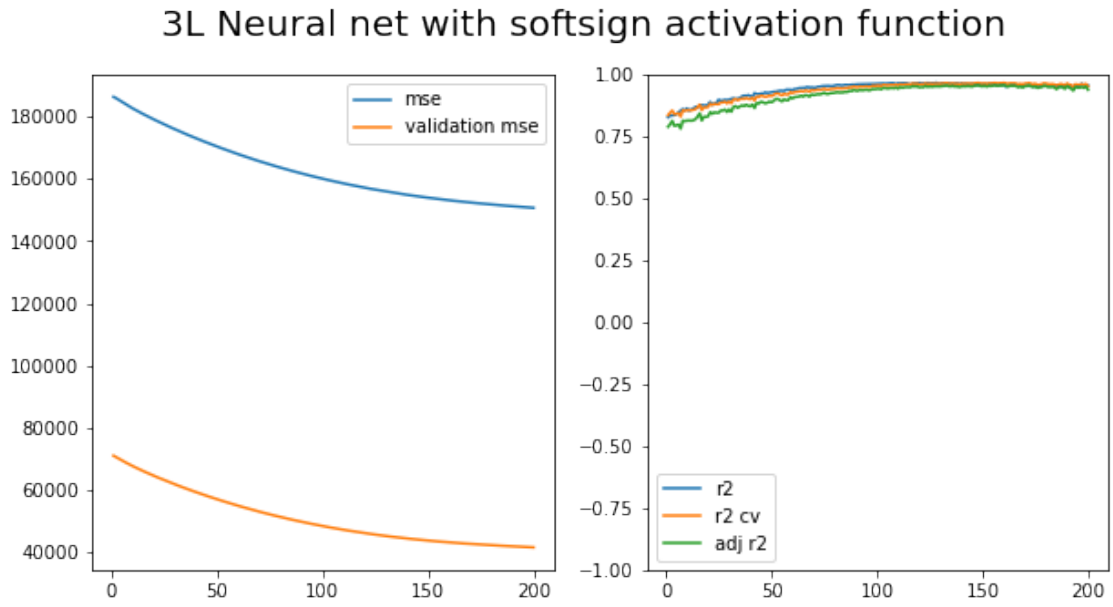
****

Building Neural nets with softplus activation function



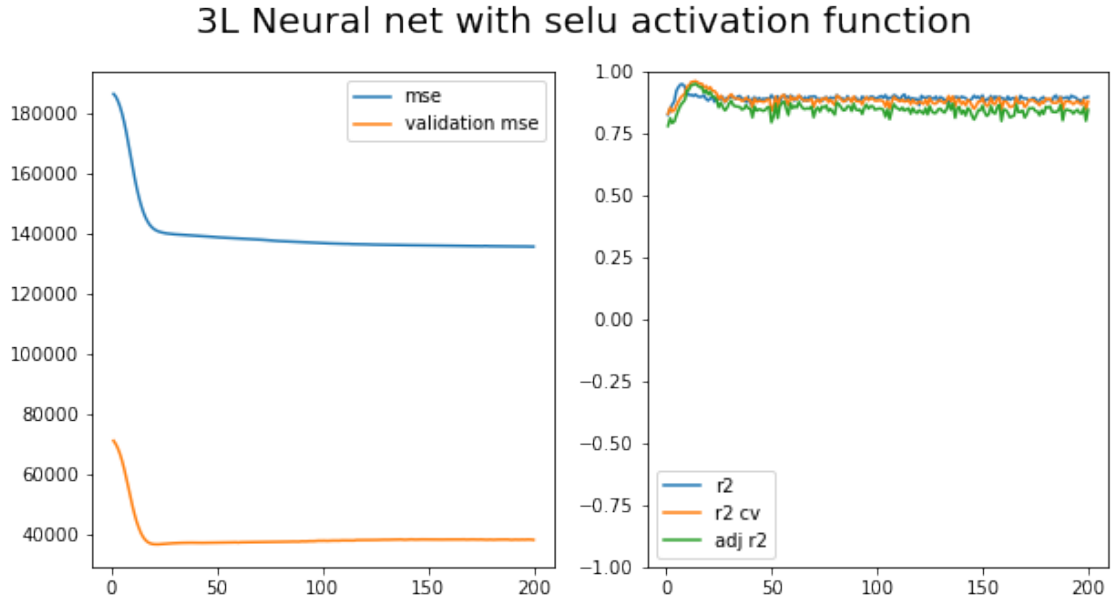
```
**********************************************************
```

```
Building Neural nets with softsign activation function
```




```
**********************************************************
```

```
Building Neural nets with selu activation function
```



```
******************************************************

Building Neural nets with exponential activation function

************************ R2 scores of various activation functions
************************

R2 for NN using linear activation is  4.103879218669126
Adj R2 for NN using linear activation is  3.912374234582594
R2 CV for NN using linear activation is  4.162251719513099


R2 for NN using relu activation is  5.99632274924674
Adj R2 for NN using relu activation is  5.808596983394009
R2 CV for NN using relu activation is  5.922106731234111


R2 for NN using sigmoid activation is  0.08397019808000783
Adj R2 for NN using sigmoid activation is  -0.11556256238314244
R2 CV for NN using sigmoid activation is  -1.2838785157971877


R2 for NN using tanh activation is  -0.021890618552489194
Adj R2 for NN using tanh activation is  -0.22163478354211286
R2 CV for NN using tanh activation is  -1.228467023926183


R2 for NN using elu activation is  6.766079708935624
Adj R2 for NN using elu activation is  6.579891151190109
R2 CV for NN using elu activation is  7.122269331685304


R2 for NN using softmax activation is  -7.963731815602482
Adj R2 for NN using softmax activation is  -8.179335873147364
R2 CV for NN using softmax activation is  -9.469286332801152


R2 for NN using softplus activation is  5.773359667472933
Adj R2 for NN using softplus activation is  5.5851886433440745
R2 CV for NN using softplus activation is  5.873168402499596


R2 for NN using softsign activation is  -0.8099693351650616
Adj R2 for NN using softsign activation is  -1.0112872968927933
R2 CV for NN using softsign activation is  -1.5426560871556028


R2 for NN using selu activation is  6.357356240894796
```

```
Adj R2 for NN using selu activation is  6.170351460547108
R2 CV for NN using selu activation is  1.452417648353499


R2 for NN using exponential activation is  4.8608961979719645
Adj R2 for NN using exponential activation is  4.67090298019458
R2 CV for NN using exponential activation is  6.398381141586129


[NbConvertApp] Converting notebook Project2_3L_NN(Expedia).ipynb to pdf
[NbConvertApp] Writing 399873 bytes to Project2_3L_NN(Expedia).pdf
```