

```
In [1]: import numpy as np
import pandas as pd
import seaborn as sns
import itertools
import matplotlib.pyplot as plt
import string
import re
import collections
from sklearn import preprocessing

%matplotlib inline
```

```
In [2]: # READ DATA
train_df = pd.read_json('train.json.zip')
test_df = pd.read_json('test.json.zip')
```

TRAIN DATA FEATURE ENGINEERING

```
In [3]: # convert TARGET to the numeric
train_df['interest_level'] = train_df['interest_level'].apply(lambda x: 0 if x=='low'
                                                             else 1 if x=='medium'
                                                             else 2)

# REMOVE UNNECESSARY WORDS FROM DESCRIPTION
train_df['description'] = train_df['description'].apply(lambda x: x.replace("<br />"
train_df['description'] = train_df['description'].apply(lambda x: x.replace("br", ""
train_df['description'] = train_df['description'].apply(lambda x: x.replace("<p><a",

#basic features
train_df['rooms'] = train_df['bedrooms'] + train_df['bathrooms']

# count of photos #
train_df["num_photos"] = train_df["photos"].apply(len)

# count of "features" #
train_df["num_features"] = train_df["features"].apply(len)

# count of words present in description column #
train_df["num_description_words"] = train_df["description"].apply(lambda x: len(x.sp

# description contains email
regex = r'[\w\.-]+\@[\w\.-]+'
train_df['has_email'] = train_df['description'].apply(lambda x: 1 if re.findall(regex

# description contains phone
train_df['has_phone'] = train_df['description'].apply(lambda x: re.sub('[\+string.pun
               .apply(lambda x: [s for s in x if s.isdigit()])\
               .apply(lambda x: len([s for s in x if len(str(s))==10]))\
               .apply(lambda x: 1 if x>0 else 0)

# CONVERT LOWER ALL OF WORDS
train_df[["features"]] = train_df[["features"]].apply(
    lambda _: [list(map(str.strip, map(str.lower, x))) for x in _])
```

TEST DATA FEATURE ENGINEERING

```
In [4]: # REMOVE UNNECESSARY WORDS FROM DESCRIPTION
test_df['description'] = test_df['description'].apply(lambda x: x.replace("<br />",
```

```

test_df['description'] = test_df['description'].apply(lambda x: x.replace("br", ""))
test_df['description'] = test_df['description'].apply(lambda x: x.replace("<p><a", ""))

#basic features
test_df['rooms'] = test_df['bedrooms'] + test_df['bathrooms']

# count of photos #
test_df["num_photos"] = test_df["photos"].apply(len)

# count of "features" #
test_df["num_features"] = test_df["features"].apply(len)

# count of words present in description column #
test_df["num_description_words"] = test_df["description"].apply(lambda x: len(x.split()))

# description contains email
regex = r'[\w\.-]+@[ \w\.-]+'
test_df['has_email'] = test_df['description'].apply(lambda x: 1 if re.findall(regex, x) else 0)

# description contains phone
test_df['has_phone'] = test_df['description'].apply(lambda x: re.sub('[\s\+string.punctuation]', '', x).isdigit())

# CONVERT LOWER ALL OF WORDS
test_df[["features"]] = test_df[["features"]].apply(
    lambda _: [list(map(str.strip, map(str.lower, x))) for x in _])

```

MOST FREQUENT FEATURES EXTRACTION

In [5]:

```

feature_value_train = train_df['features'].tolist()
feature_value_test = test_df['features'].tolist()

feature_value_train
feature_value_test

feature_lst_train = []
feature_lst_test = []

for i in range(len(feature_value_train)):
    feature_lst_train += feature_value_train[i]

for i in range(len(feature_value_test)):
    feature_lst_test += feature_value_test[i]

uniq_feature_train = list(set(feature_lst_train))
uniq_feature_test = list(set(feature_lst_test))

# see the frequency of each feature
def most_common(lst):
    features = collections.Counter(lst)
    feature_value = features.keys()
    frequency = features.values()
    data = [('feature_value', feature_value),
            ('frequency', frequency),]
    df = pd.DataFrame.from_dict(dict(data))
    return df.sort_values(by = 'frequency', ascending = False)

df_features_train = most_common(feature_lst_train)
df_features_test = most_common(feature_lst_test)

```

```
def newColumn(name, df, series):
    feature = pd.Series(0, df.index, name = name) # data : 0
    for row, word in enumerate(series):
        if name in word:
            feature.iloc[row] = 1
    df[name] = feature # feature : series ; value in series : 1 or 0
    return df

# select features based on frequency
facilities = ['elevator', 'cats allowed', 'hardwood floors', 'dogs allowed', 'doorma']
for name in facilities:
    train_df = newColumn(name, train_df, train_df['features'])
    test_df = newColumn(name, test_df, test_df['features'])
```

LABEL ENCODING

```
In [6]: categorical = ["display_address", "manager_id", "building_id", "street_address"]
for f in categorical:
    if train_df[f].dtype == 'object':
        #print(f)
        lbl = preprocessing.LabelEncoder()
        lbl.fit(list(train_df[f].values) + list(test_df[f].values))
        train_df[f] = lbl.transform(list(train_df[f].values))
        test_df[f] = lbl.transform(list(test_df[f].values))
```

```
In [7]: train_df['price'] = np.log10(train_df['price'])
test_df['price'] = np.log10(test_df['price'])
```

```
In [8]: train_df['price']
```

```
Out[8]: 4          3.380211
        6          3.579784
        9          3.543447
       10          3.477121
       15          3.446382
        ...
    124000          3.447158
    124002          3.379306
    124004          3.267172
    124008          3.622732
    124009          3.631444
Name: price, Length: 49352, dtype: float64
```

DROP UNNECESSARY COLUMNS

```
In [9]: # TRAINING DATASET
train_df.drop('created', axis=1, inplace=True)
train_df.drop('description', axis=1, inplace=True)
train_df.drop('features', axis=1, inplace=True)
train_df.drop('photos', axis=1, inplace=True)

# TEST DATASET
test_df.drop('created', axis=1, inplace=True)
test_df.drop('description', axis=1, inplace=True)
```

```
test_df.drop('features', axis=1, inplace=True)
test_df.drop('photos', axis=1, inplace=True)
```

REGRESSION FOR PRICE

In [10]:

```
from sklearn.model_selection import train_test_split
from sklearn.metrics import make_scorer, mean_absolute_error, mean_squared_error
from sklearn.preprocessing import StandardScaler
from sklearn.pipeline import Pipeline
import xgboost as xgb
from sklearn.linear_model import LinearRegression
from sklearn.ensemble import RandomForestRegressor
from sklearn.svm import SVR
import optuna
import math
from keras import callbacks

X = train_df.drop(['price'], axis = 1)
y = train_df.price
X_train, X_test, y_train, y_test = train_test_split(X, y,
                                                    test_size = .3,
                                                    random_state = 5)
```

In [11]:

```
class Optimizer:
    def __init__(self, metric, trials=100):
        self.metric = metric
        self.trials = trials

    def objective(self, trial):
        model = create_model(trial)
        model.fit(X, y)
        preds = model.predict(X_test)
        return mean_absolute_error(y_test, preds)

    def optimize(self):
        study = optuna.create_study(direction="minimize")
        study.optimize(self.objective, n_trials=self.trials)
        return study
```

In [12]:

```
from keras.layers import Dense
from keras.models import Sequential

def create_model(trial):
    model = Sequential()
    model.add(Dense(
        (
            units = trial.suggest_int("units", 32, 256),
            kernel_initializer=trial.suggest_categorical("kernel_initializer", ["glorot_uniform", "he_normal", "lecun_normal", "xavier_uniform", "xavier_normal", "zeros", "ones", "variance_scaling"]),
            input_dim = X.shape[1],
            activation= trial.suggest_categorical("activation", ["relu", "sigmoid", "tanh", "leaky_relu", "selu", "swish"])
        )
    )
    model.add(Dense(1))
    model.compile(
        loss="mean_absolute_error"
    )
    return model

optimizer = Optimizer('mae')
```

```

keras_study = optimizer.optimize()

print("Number of finished trials: ", len(keras_study.trials))
print("Best trial:")
keras_trial = keras_study.best_trial

print("  Value: {}".format(keras_trial.value))
print("  Params: ")
for key, value in keras_trial.params.items():
    print("    {}: {}".format(key, value))

```

```

[I 2021-09-28 17:53:36,049] A new study created in memory with name: no-name-a331b82
c-4720-4590-9c92-fed25c0a8156
1543/1543 [=====] - 2s 889us/step - loss: 27321.9727
[I 2021-09-28 17:53:39,674] Trial 0 finished with value: 16515.011623708084 and para
meters: {'units': 177, 'kernel_initializer': 'normal', 'activation': 'silu'}. Best i
s trial 0 with value: 16515.011623708084.
1543/1543 [=====] - 2s 781us/step - loss: 54704.8359
[I 2021-09-28 17:53:41,675] Trial 1 finished with value: 90180.31850626413 and param
eters: {'units': 203, 'kernel_initializer': 'glorot_uniform', 'activation': 'relu'}.
Best is trial 0 with value: 16515.011623708084.
1543/1543 [=====] - 2s 939us/step - loss: 57048.8086
[I 2021-09-28 17:53:43,949] Trial 2 finished with value: 80783.38150766604 and param
eters: {'units': 243, 'kernel_initializer': 'glorot_uniform', 'activation': 'silu'}.
Best is trial 0 with value: 16515.011623708084.
1543/1543 [=====] - 2s 784us/step - loss: 51089.2109
[I 2021-09-28 17:53:45,938] Trial 3 finished with value: 9333.99197716002 and parame
ters: {'units': 143, 'kernel_initializer': 'glorot_uniform', 'activation': 'leaky_re
lu'}. Best is trial 3 with value: 9333.99197716002.
1543/1543 [=====] - 2s 789us/step - loss: 66806.7266
[I 2021-09-28 17:53:47,941] Trial 4 finished with value: 75658.22424878449 and param
eters: {'units': 216, 'kernel_initializer': 'glorot_uniform', 'activation': 'leaky_r
elu'}. Best is trial 3 with value: 9333.99197716002.
1543/1543 [=====] - 2s 800us/step - loss: 67179.4141
[I 2021-09-28 17:53:49,956] Trial 5 finished with value: 7424.058743536404 and param
eters: {'units': 214, 'kernel_initializer': 'glorot_uniform', 'activation': 'leaky_r
elu'}. Best is trial 5 with value: 7424.058743536404.
1543/1543 [=====] - 2s 843us/step - loss: 25007.1777
[I 2021-09-28 17:53:52,075] Trial 6 finished with value: 9722.378897348204 and param
eters: {'units': 145, 'kernel_initializer': 'normal', 'activation': 'silu'}. Best is
trial 5 with value: 7424.058743536404.
1543/1543 [=====] - 2s 740us/step - loss: 31083.6055
[I 2021-09-28 17:53:53,991] Trial 7 finished with value: 10187.92243860196 and param
eters: {'units': 83, 'kernel_initializer': 'glorot_uniform', 'activation': 'relu'}.
Best is trial 5 with value: 7424.058743536404.
1543/1543 [=====] - 2s 742us/step - loss: 41807.4648
[I 2021-09-28 17:53:55,911] Trial 8 finished with value: 40956.81029221663 and param
eters: {'units': 68, 'kernel_initializer': 'glorot_uniform', 'activation': 'leaky_re
lu'}. Best is trial 5 with value: 7424.058743536404.
1543/1543 [=====] - 2s 761us/step - loss: 26427.2793
[I 2021-09-28 17:53:57,861] Trial 9 finished with value: 5143.597371840069 and param
eters: {'units': 94, 'kernel_initializer': 'normal', 'activation': 'leaky_relu'}. Be
st is trial 9 with value: 5143.597371840069.
1543/1543 [=====] - 2s 749us/step - loss: 18262.5195
[I 2021-09-28 17:53:59,795] Trial 10 finished with value: 7741.619826115889 and para
meters: {'units': 37, 'kernel_initializer': 'normal', 'activation': 'leaky_relu'}. B
est is trial 9 with value: 5143.597371840069.
1543/1543 [=====] - 2s 756us/step - loss: 30943.8340
[I 2021-09-28 17:54:01,743] Trial 11 finished with value: 5721.640498166707 and para
meters: {'units': 109, 'kernel_initializer': 'normal', 'activation': 'leaky_relu'}.
Best is trial 9 with value: 5143.597371840069.
1543/1543 [=====] - 2s 764us/step - loss: 36068.6914
[I 2021-09-28 17:54:04,035] Trial 12 finished with value: 54769.80959808292 and para
meters: {'units': 109, 'kernel_initializer': 'normal', 'activation': 'leaky_relu'}.

```

Best is trial 9 with value: 5143.597371840069.
1543/1543 [=====] - 2s 757us/step - loss: 27747.5215
[I 2021-09-28 17:54:05,980] Trial 13 finished with value: 17666.23294827709 and parameters: {'units': 114, 'kernel_initializer': 'normal', 'activation': 'leaky_relu'}.
Best is trial 9 with value: 5143.597371840069.
1543/1543 [=====] - 2s 769us/step - loss: 16912.9648
[I 2021-09-28 17:54:07,933] Trial 14 finished with value: 18301.660984604212 and parameters: {'units': 70, 'kernel_initializer': 'normal', 'activation': 'relu'}. Best is trial 9 with value: 5143.597371840069.
1543/1543 [=====] - 2s 756us/step - loss: 15365.1514
[I 2021-09-28 17:54:09,873] Trial 15 finished with value: 26945.909688945576 and parameters: {'units': 33, 'kernel_initializer': 'normal', 'activation': 'leaky_relu'}.
Best is trial 9 with value: 5143.597371840069.
1543/1543 [=====] - 2s 763us/step - loss: 32037.3457
[I 2021-09-28 17:54:11,832] Trial 16 finished with value: 40552.65903491385 and parameters: {'units': 114, 'kernel_initializer': 'normal', 'activation': 'leaky_relu'}.
Best is trial 9 with value: 5143.597371840069.
1543/1543 [=====] - 2s 795us/step - loss: 35930.0859
[I 2021-09-28 17:54:13,855] Trial 17 finished with value: 10569.08211923935 and parameters: {'units': 149, 'kernel_initializer': 'normal', 'activation': 'leaky_relu'}.
Best is trial 9 with value: 5143.597371840069.
1543/1543 [=====] - 2s 764us/step - loss: 24883.5723
[I 2021-09-28 17:54:15,812] Trial 18 finished with value: 11579.790785106967 and parameters: {'units': 92, 'kernel_initializer': 'normal', 'activation': 'relu'}. Best is trial 9 with value: 5143.597371840069.
1543/1543 [=====] - 2s 790us/step - loss: 13132.8770
[I 2021-09-28 17:54:17,833] Trial 19 finished with value: 19686.915943272368 and parameters: {'units': 53, 'kernel_initializer': 'normal', 'activation': 'silu'}. Best is trial 9 with value: 5143.597371840069.
1543/1543 [=====] - 2s 777us/step - loss: 36510.5312
[I 2021-09-28 17:54:19,835] Trial 20 finished with value: 47414.51819475656 and parameters: {'units': 168, 'kernel_initializer': 'normal', 'activation': 'leaky_relu'}.
Best is trial 9 with value: 5143.597371840069.
1543/1543 [=====] - 2s 811us/step - loss: 63960.9727
[I 2021-09-28 17:54:21,871] Trial 21 finished with value: 93590.47520912152 and parameters: {'units': 252, 'kernel_initializer': 'glorot_uniform', 'activation': 'leaky_relu'}. Best is trial 9 with value: 5143.597371840069.
1543/1543 [=====] - 2s 769us/step - loss: 57920.1406
[I 2021-09-28 17:54:23,842] Trial 22 finished with value: 105114.55650692731 and parameters: {'units': 127, 'kernel_initializer': 'glorot_uniform', 'activation': 'leaky_relu'}. Best is trial 9 with value: 5143.597371840069.
1543/1543 [=====] - 2s 795us/step - loss: 54381.1641
[I 2021-09-28 17:54:25,853] Trial 23 finished with value: 9696.88126831215 and parameters: {'units': 186, 'kernel_initializer': 'glorot_uniform', 'activation': 'leaky_relu'}. Best is trial 9 with value: 5143.597371840069.
1543/1543 [=====] - 2s 760us/step - loss: 31248.1992
[I 2021-09-28 17:54:27,803] Trial 24 finished with value: 19785.04346982452 and parameters: {'units': 92, 'kernel_initializer': 'normal', 'activation': 'leaky_relu'}. Best is trial 9 with value: 5143.597371840069.
1543/1543 [=====] - 2s 799us/step - loss: 45502.0977
[I 2021-09-28 17:54:29,824] Trial 25 finished with value: 41307.75690472844 and parameters: {'units': 224, 'kernel_initializer': 'normal', 'activation': 'leaky_relu'}.
Best is trial 9 with value: 5143.597371840069.
1543/1543 [=====] - 2s 772us/step - loss: 53962.8711
[I 2021-09-28 17:54:31,799] Trial 26 finished with value: 6426.2445702727155 and parameters: {'units': 131, 'kernel_initializer': 'glorot_uniform', 'activation': 'leaky_relu'}. Best is trial 9 with value: 5143.597371840069.
1543/1543 [=====] - 2s 784us/step - loss: 32292.0996
[I 2021-09-28 17:54:33,784] Trial 27 finished with value: 43416.01004904071 and parameters: {'units': 127, 'kernel_initializer': 'normal', 'activation': 'leaky_relu'}.
Best is trial 9 with value: 5143.597371840069.
1543/1543 [=====] - 2s 754us/step - loss: 46856.1289
[I 2021-09-28 17:54:35,717] Trial 28 finished with value: 61678.71735229826 and parameters: {'units': 101, 'kernel_initializer': 'glorot_uniform', 'activation': 'rel

u'}). Best is trial 9 with value: 5143.597371840069.
1543/1543 [=====] - 2s 858us/step - loss: 33009.7695
[I 2021-09-28 17:54:37,867] Trial 29 finished with value: 14453.937165137655 and parameters: {'units': 162, 'kernel_initializer': 'normal', 'activation': 'silu'}. Best is trial 9 with value: 5143.597371840069.
1543/1543 [=====] - 2s 866us/step - loss: 25004.6621
[I 2021-09-28 17:54:40,029] Trial 30 finished with value: 4632.522870892131 and parameters: {'units': 135, 'kernel_initializer': 'normal', 'activation': 'silu'}. Best is trial 30 with value: 4632.522870892131.
1543/1543 [=====] - 2s 843us/step - loss: 25181.0957
[I 2021-09-28 17:54:42,542] Trial 31 finished with value: 10980.277434570655 and parameters: {'units': 130, 'kernel_initializer': 'normal', 'activation': 'silu'}. Best is trial 30 with value: 4632.522870892131.
1543/1543 [=====] - 2s 793us/step - loss: 22022.1230
[I 2021-09-28 17:54:44,574] Trial 32 finished with value: 2928.144715932625 and parameters: {'units': 75, 'kernel_initializer': 'normal', 'activation': 'silu'}. Best is trial 32 with value: 2928.144715932625.
1543/1543 [=====] - 2s 796us/step - loss: 25327.3301
[I 2021-09-28 17:54:46,608] Trial 33 finished with value: 6419.799125528503 and parameters: {'units': 79, 'kernel_initializer': 'normal', 'activation': 'silu'}. Best is trial 32 with value: 2928.144715932625.
1543/1543 [=====] - 2s 775us/step - loss: 25934.7734
[I 2021-09-28 17:54:48,616] Trial 34 finished with value: 41298.14299345514 and parameters: {'units': 59, 'kernel_initializer': 'normal', 'activation': 'silu'}. Best is trial 32 with value: 2928.144715932625.
1543/1543 [=====] - 2s 815us/step - loss: 18252.1543
[I 2021-09-28 17:54:50,682] Trial 35 finished with value: 26940.275950269523 and parameters: {'units': 100, 'kernel_initializer': 'normal', 'activation': 'silu'}. Best is trial 32 with value: 2928.144715932625.
1543/1543 [=====] - 2s 912us/step - loss: 31918.0547
[I 2021-09-28 17:54:52,916] Trial 36 finished with value: 32329.118003190426 and parameters: {'units': 189, 'kernel_initializer': 'normal', 'activation': 'silu'}. Best is trial 32 with value: 2928.144715932625.
1543/1543 [=====] - 2s 766us/step - loss: 12480.5811
[I 2021-09-28 17:54:54,906] Trial 37 finished with value: 2051.6157451483236 and parameters: {'units': 52, 'kernel_initializer': 'normal', 'activation': 'silu'}. Best is trial 37 with value: 2051.6157451483236.
1543/1543 [=====] - 2s 757us/step - loss: 13681.3984
[I 2021-09-28 17:54:56,873] Trial 38 finished with value: 5329.171856996109 and parameters: {'units': 46, 'kernel_initializer': 'normal', 'activation': 'silu'}. Best is trial 37 with value: 2051.6157451483236.
1543/1543 [=====] - 2s 790us/step - loss: 20067.7148
[I 2021-09-28 17:54:58,905] Trial 39 finished with value: 14832.836937347824 and parameters: {'units': 83, 'kernel_initializer': 'normal', 'activation': 'silu'}. Best is trial 37 with value: 2051.6157451483236.
1543/1543 [=====] - 2s 794us/step - loss: 12439.1562
[I 2021-09-28 17:55:00,956] Trial 40 finished with value: 11058.084726264162 and parameters: {'units': 66, 'kernel_initializer': 'normal', 'activation': 'silu'}. Best is trial 37 with value: 2051.6157451483236.
1543/1543 [=====] - 2s 809us/step - loss: 17988.2188
[I 2021-09-28 17:55:03,018] Trial 41 finished with value: 5299.957124428625 and parameters: {'units': 45, 'kernel_initializer': 'normal', 'activation': 'silu'}. Best is trial 37 with value: 2051.6157451483236.
1543/1543 [=====] - 2s 770us/step - loss: 13752.2598
[I 2021-09-28 17:55:05,066] Trial 42 finished with value: 22548.955066453003 and parameters: {'units': 45, 'kernel_initializer': 'normal', 'activation': 'silu'}. Best is trial 37 with value: 2051.6157451483236.
1543/1543 [=====] - 2s 796us/step - loss: 13857.8740
[I 2021-09-28 17:55:07,110] Trial 43 finished with value: 16829.103555201757 and parameters: {'units': 59, 'kernel_initializer': 'normal', 'activation': 'silu'}. Best is trial 37 with value: 2051.6157451483236.
1543/1543 [=====] - 2s 787us/step - loss: 18598.0566
[I 2021-09-28 17:55:09,142] Trial 44 finished with value: 26058.72848482055 and parameters: {'units': 77, 'kernel_initializer': 'normal', 'activation': 'silu'}. Best is

trial 37 with value: 2051.6157451483236.
1543/1543 [=====] - 2s 754us/step - loss: 21076.9492
[I 2021-09-28 17:55:11,109] Trial 45 finished with value: 4744.706757013969 and parameters: {'units': 45, 'kernel_initializer': 'normal', 'activation': 'silu'}. Best is trial 37 with value: 2051.6157451483236.
1543/1543 [=====] - 2s 783us/step - loss: 22372.1855
[I 2021-09-28 17:55:13,124] Trial 46 finished with value: 3023.8700002230544 and parameters: {'units': 70, 'kernel_initializer': 'normal', 'activation': 'silu'}. Best is trial 37 with value: 2051.6157451483236.
1543/1543 [=====] - 2s 766us/step - loss: 9951.4502
[I 2021-09-28 17:55:15,105] Trial 47 finished with value: 1082.722668145338 and parameters: {'units': 37, 'kernel_initializer': 'normal', 'activation': 'silu'}. Best is trial 47 with value: 1082.722668145338.
1543/1543 [=====] - 2s 741us/step - loss: 10979.1729
[I 2021-09-28 17:55:17,054] Trial 48 finished with value: 2771.9410263718573 and parameters: {'units': 36, 'kernel_initializer': 'normal', 'activation': 'silu'}. Best is trial 47 with value: 1082.722668145338.
1543/1543 [=====] - 2s 752us/step - loss: 12094.3516
[I 2021-09-28 17:55:19,426] Trial 49 finished with value: 19559.083900413898 and parameters: {'units': 38, 'kernel_initializer': 'normal', 'activation': 'silu'}. Best is trial 47 with value: 1082.722668145338.
1543/1543 [=====] - 2s 779us/step - loss: 16654.5391
[I 2021-09-28 17:55:21,439] Trial 50 finished with value: 3539.899575330315 and parameters: {'units': 57, 'kernel_initializer': 'normal', 'activation': 'silu'}. Best is trial 47 with value: 1082.722668145338.
1543/1543 [=====] - 2s 744us/step - loss: 11831.1426
[I 2021-09-28 17:55:23,383] Trial 51 finished with value: 16250.40572353195 and parameters: {'units': 32, 'kernel_initializer': 'normal', 'activation': 'silu'}. Best is trial 47 with value: 1082.722668145338.
1543/1543 [=====] - 2s 764us/step - loss: 15174.0020
[I 2021-09-28 17:55:25,354] Trial 52 finished with value: 20497.44996769659 and parameters: {'units': 57, 'kernel_initializer': 'normal', 'activation': 'silu'}. Best is trial 47 with value: 1082.722668145338.
1543/1543 [=====] - 2s 790us/step - loss: 20755.6426
[I 2021-09-28 17:55:27,375] Trial 53 finished with value: 15371.60851471764 and parameters: {'units': 72, 'kernel_initializer': 'normal', 'activation': 'silu'}. Best is trial 47 with value: 1082.722668145338.
1543/1543 [=====] - 2s 774us/step - loss: 17296.2754
[I 2021-09-28 17:55:29,379] Trial 54 finished with value: 5020.3144121020405 and parameters: {'units': 52, 'kernel_initializer': 'normal', 'activation': 'silu'}. Best is trial 47 with value: 1082.722668145338.
1543/1543 [=====] - 2s 747us/step - loss: 15011.9199
[I 2021-09-28 17:55:31,300] Trial 55 finished with value: 22809.7591546609 and parameters: {'units': 65, 'kernel_initializer': 'normal', 'activation': 'relu'}. Best is trial 47 with value: 1082.722668145338.
1543/1543 [=====] - 2s 752us/step - loss: 9571.4609
[I 2021-09-28 17:55:33,262] Trial 56 finished with value: 13319.44896950012 and parameters: {'units': 40, 'kernel_initializer': 'normal', 'activation': 'silu'}. Best is trial 47 with value: 1082.722668145338.
1543/1543 [=====] - 2s 771us/step - loss: 15888.2334
[I 2021-09-28 17:55:35,253] Trial 57 finished with value: 7522.018213229435 and parameters: {'units': 51, 'kernel_initializer': 'normal', 'activation': 'silu'}. Best is trial 47 with value: 1082.722668145338.
1543/1543 [=====] - 2s 791us/step - loss: 21934.4785
[I 2021-09-28 17:55:37,288] Trial 58 finished with value: 5665.0814336141175 and parameters: {'units': 63, 'kernel_initializer': 'normal', 'activation': 'silu'}. Best is trial 47 with value: 1082.722668145338.
1543/1543 [=====] - 2s 790us/step - loss: 42130.3867
[I 2021-09-28 17:55:39,323] Trial 59 finished with value: 61021.58478819799 and parameters: {'units': 83, 'kernel_initializer': 'glorot_uniform', 'activation': 'silu'}. Best is trial 47 with value: 1082.722668145338.
1543/1543 [=====] - 2s 739us/step - loss: 10692.3438
[I 2021-09-28 17:55:41,238] Trial 60 finished with value: 4092.831026987899 and parameters: {'units': 32, 'kernel_initializer': 'normal', 'activation': 'relu'}. Best is

trial 47 with value: 1082.722668145338.
1543/1543 [=====] - 2s 737us/step - loss: 11669.2080
[I 2021-09-28 17:55:43,138] Trial 61 finished with value: 14209.514560250278 and parameters: {'units': 32, 'kernel_initializer': 'normal', 'activation': 'relu'}. Best is trial 47 with value: 1082.722668145338.
1543/1543 [=====] - 2s 737us/step - loss: 21738.0938
[I 2021-09-28 17:55:45,041] Trial 62 finished with value: 5957.685105244416 and parameters: {'units': 73, 'kernel_initializer': 'normal', 'activation': 'relu'}. Best is trial 47 with value: 1082.722668145338.
1543/1543 [=====] - 2s 746us/step - loss: 15823.5723
[I 2021-09-28 17:55:46,964] Trial 63 finished with value: 1044.2942941830224 and parameters: {'units': 53, 'kernel_initializer': 'normal', 'activation': 'relu'}. Best is trial 63 with value: 1044.2942941830224.
1543/1543 [=====] - 2s 750us/step - loss: 13048.6240
[I 2021-09-28 17:55:48,886] Trial 64 finished with value: 1568.733906626859 and parameters: {'units': 56, 'kernel_initializer': 'normal', 'activation': 'relu'}. Best is trial 63 with value: 1044.2942941830224.
1543/1543 [=====] - 2s 743us/step - loss: 20172.3828
[I 2021-09-28 17:55:50,803] Trial 65 finished with value: 32413.610445812927 and parameters: {'units': 51, 'kernel_initializer': 'normal', 'activation': 'relu'}. Best is trial 63 with value: 1044.2942941830224.
1543/1543 [=====] - 2s 763us/step - loss: 17346.3145
[I 2021-09-28 17:55:52,752] Trial 66 finished with value: 25610.937852538875 and parameters: {'units': 90, 'kernel_initializer': 'normal', 'activation': 'relu'}. Best is trial 63 with value: 1044.2942941830224.
1543/1543 [=====] - 1s 727us/step - loss: 18835.5723
[I 2021-09-28 17:55:54,639] Trial 67 finished with value: 3410.3496567841744 and parameters: {'units': 41, 'kernel_initializer': 'glorot_uniform', 'activation': 'relu'}. Best is trial 63 with value: 1044.2942941830224.
1543/1543 [=====] - 2s 741us/step - loss: 20187.2324
[I 2021-09-28 17:55:56,567] Trial 68 finished with value: 4635.826816703268 and parameters: {'units': 67, 'kernel_initializer': 'normal', 'activation': 'relu'}. Best is trial 63 with value: 1044.2942941830224.
1543/1543 [=====] - 2s 752us/step - loss: 18388.2930
[I 2021-09-28 17:55:58,499] Trial 69 finished with value: 15543.562669573568 and parameters: {'units': 52, 'kernel_initializer': 'normal', 'activation': 'relu'}. Best is trial 63 with value: 1044.2942941830224.
1543/1543 [=====] - 2s 756us/step - loss: 14280.2061
[I 2021-09-28 17:56:00,436] Trial 70 finished with value: 1888.8348079191933 and parameters: {'units': 75, 'kernel_initializer': 'normal', 'activation': 'relu'}. Best is trial 63 with value: 1044.2942941830224.
1543/1543 [=====] - 2s 743us/step - loss: 16550.2969
[I 2021-09-28 17:56:02,820] Trial 71 finished with value: 3448.349923564614 and parameters: {'units': 63, 'kernel_initializer': 'normal', 'activation': 'relu'}. Best is trial 63 with value: 1044.2942941830224.
1543/1543 [=====] - 2s 732us/step - loss: 22077.0840
[I 2021-09-28 17:56:04,717] Trial 72 finished with value: 33658.23314232321 and parameters: {'units': 78, 'kernel_initializer': 'normal', 'activation': 'relu'}. Best is trial 63 with value: 1044.2942941830224.
1543/1543 [=====] - 2s 749us/step - loss: 19817.3340
[I 2021-09-28 17:56:06,636] Trial 73 finished with value: 5696.682805361001 and parameters: {'units': 87, 'kernel_initializer': 'normal', 'activation': 'relu'}. Best is trial 63 with value: 1044.2942941830224.
1543/1543 [=====] - 2s 766us/step - loss: 24858.6445
[I 2021-09-28 17:56:08,591] Trial 74 finished with value: 39163.27128353772 and parameters: {'units': 101, 'kernel_initializer': 'normal', 'activation': 'relu'}. Best is trial 63 with value: 1044.2942941830224.
1543/1543 [=====] - 2s 748us/step - loss: 20006.3496
[I 2021-09-28 17:56:10,512] Trial 75 finished with value: 24233.888692781005 and parameters: {'units': 71, 'kernel_initializer': 'normal', 'activation': 'relu'}. Best is trial 63 with value: 1044.2942941830224.
1543/1543 [=====] - 2s 744us/step - loss: 12005.3545
[I 2021-09-28 17:56:12,427] Trial 76 finished with value: 844.9858726442183 and parameters: {'units': 42, 'kernel_initializer': 'normal', 'activation': 'relu'}. Best is

trial 76 with value: 844.9858726442183.
1543/1543 [=====] - 2s 745us/step - loss: 14307.5479
[I 2021-09-28 17:56:14,340] Trial 77 finished with value: 1541.7290582079538 and parameters: {'units': 41, 'kernel_initializer': 'normal', 'activation': 'relu'}. Best is trial 76 with value: 844.9858726442183.
1543/1543 [=====] - 2s 744us/step - loss: 31815.1172
[I 2021-09-28 17:56:16,277] Trial 78 finished with value: 8483.10794473189 and parameters: {'units': 44, 'kernel_initializer': 'glorot_uniform', 'activation': 'relu'}. Best is trial 76 with value: 844.9858726442183.
1543/1543 [=====] - 2s 774us/step - loss: 20267.0996
[I 2021-09-28 17:56:18,271] Trial 79 finished with value: 1540.1521487327038 and parameters: {'units': 36, 'kernel_initializer': 'normal', 'activation': 'relu'}. Best is trial 76 with value: 844.9858726442183.
1543/1543 [=====] - 2s 753us/step - loss: 11604.5742
[I 2021-09-28 17:56:20,212] Trial 80 finished with value: 16019.953358552064 and parameters: {'units': 49, 'kernel_initializer': 'normal', 'activation': 'relu'}. Best is trial 76 with value: 844.9858726442183.
1543/1543 [=====] - 2s 743us/step - loss: 18130.2988
[I 2021-09-28 17:56:22,122] Trial 81 finished with value: 9527.928814822539 and parameters: {'units': 39, 'kernel_initializer': 'normal', 'activation': 'relu'}. Best is trial 76 with value: 844.9858726442183.
1543/1543 [=====] - 2s 738us/step - loss: 22253.5820
[I 2021-09-28 17:56:24,035] Trial 82 finished with value: 35620.44610788046 and parameters: {'units': 58, 'kernel_initializer': 'normal', 'activation': 'relu'}. Best is trial 76 with value: 844.9858726442183.
1543/1543 [=====] - 2s 735us/step - loss: 13405.9648
[I 2021-09-28 17:56:25,934] Trial 83 finished with value: 21660.368629915865 and parameters: {'units': 37, 'kernel_initializer': 'normal', 'activation': 'relu'}. Best is trial 76 with value: 844.9858726442183.
1543/1543 [=====] - 2s 738us/step - loss: 14583.1973
[I 2021-09-28 17:56:27,850] Trial 84 finished with value: 8098.180757526298 and parameters: {'units': 37, 'kernel_initializer': 'normal', 'activation': 'relu'}. Best is trial 76 with value: 844.9858726442183.
1543/1543 [=====] - 2s 737us/step - loss: 14404.0742
[I 2021-09-28 17:56:29,757] Trial 85 finished with value: 5172.574298725549 and parameters: {'units': 47, 'kernel_initializer': 'normal', 'activation': 'relu'}. Best is trial 76 with value: 844.9858726442183.
1543/1543 [=====] - 2s 753us/step - loss: 20487.5156
[I 2021-09-28 17:56:31,691] Trial 86 finished with value: 23359.957063373593 and parameters: {'units': 57, 'kernel_initializer': 'normal', 'activation': 'relu'}. Best is trial 76 with value: 844.9858726442183.
1543/1543 [=====] - 2s 736us/step - loss: 12311.6748
[I 2021-09-28 17:56:33,604] Trial 87 finished with value: 16033.852183286865 and parameters: {'units': 39, 'kernel_initializer': 'normal', 'activation': 'relu'}. Best is trial 76 with value: 844.9858726442183.
1543/1543 [=====] - 1s 728us/step - loss: 13892.3652
[I 2021-09-28 17:56:35,502] Trial 88 finished with value: 12312.979130937416 and parameters: {'units': 44, 'kernel_initializer': 'normal', 'activation': 'relu'}. Best is trial 76 with value: 844.9858726442183.
1543/1543 [=====] - 2s 751us/step - loss: 21866.1875
[I 2021-09-28 17:56:37,442] Trial 89 finished with value: 4568.773469425719 and parameters: {'units': 54, 'kernel_initializer': 'normal', 'activation': 'relu'}. Best is trial 76 with value: 844.9858726442183.
1543/1543 [=====] - 2s 745us/step - loss: 15327.4785
[I 2021-09-28 17:56:39,367] Trial 90 finished with value: 23936.22120631501 and parameters: {'units': 61, 'kernel_initializer': 'normal', 'activation': 'relu'}. Best is trial 76 with value: 844.9858726442183.
1543/1543 [=====] - 1s 727us/step - loss: 10934.7031
[I 2021-09-28 17:56:41,263] Trial 91 finished with value: 7921.239030455557 and parameters: {'units': 36, 'kernel_initializer': 'normal', 'activation': 'relu'}. Best is trial 76 with value: 844.9858726442183.
1543/1543 [=====] - 2s 770us/step - loss: 12418.1328
[I 2021-09-28 17:56:43,687] Trial 92 finished with value: 8220.341535484136 and parameters: {'units': 47, 'kernel_initializer': 'normal', 'activation': 'silu'}. Best is

```

trial 76 with value: 844.9858726442183.
1543/1543 [=====] - 2s 734us/step - loss: 10737.2812
[I 2021-09-28 17:56:45,584] Trial 93 finished with value: 5019.033403170801 and parameters: {'units': 53, 'kernel_initializer': 'normal', 'activation': 'relu'}. Best is trial 76 with value: 844.9858726442183.
1543/1543 [=====] - 2s 764us/step - loss: 15400.4824
[I 2021-09-28 17:56:47,556] Trial 94 finished with value: 2125.3077831993783 and parameters: {'units': 43, 'kernel_initializer': 'normal', 'activation': 'silu'}. Best is trial 76 with value: 844.9858726442183.
1543/1543 [=====] - 2s 732us/step - loss: 15003.5859
[I 2021-09-28 17:56:49,454] Trial 95 finished with value: 17891.298006097823 and parameters: {'units': 43, 'kernel_initializer': 'normal', 'activation': 'relu'}. Best is trial 76 with value: 844.9858726442183.
1543/1543 [=====] - 2s 750us/step - loss: 21143.8711
[I 2021-09-28 17:56:51,410] Trial 96 finished with value: 9571.66137961033 and parameters: {'units': 35, 'kernel_initializer': 'glorot_uniform', 'activation': 'silu'}. Best is trial 76 with value: 844.9858726442183.
1543/1543 [=====] - 2s 980us/step - loss: 31723.8496
[I 2021-09-28 17:56:53,769] Trial 97 finished with value: 30763.805999193533 and parameters: {'units': 225, 'kernel_initializer': 'normal', 'activation': 'silu'}. Best is trial 76 with value: 844.9858726442183.
1543/1543 [=====] - 2s 744us/step - loss: 17194.4453
[I 2021-09-28 17:56:55,723] Trial 98 finished with value: 14620.5028092472 and parameters: {'units': 41, 'kernel_initializer': 'normal', 'activation': 'relu'}. Best is trial 76 with value: 844.9858726442183.
1543/1543 [=====] - 2s 799us/step - loss: 20250.0156
[I 2021-09-28 17:56:57,790] Trial 99 finished with value: 20931.296469822406 and parameters: {'units': 49, 'kernel_initializer': 'normal', 'activation': 'silu'}. Best is trial 76 with value: 844.9858726442183.
Number of finished trials: 100
Best trial:
  Value: 844.9858726442183
  Params:
    units: 42
    kernel_initializer: normal
    activation: relu

```

In [13]:

```

keras_params = keras_study.best_params

model2 = Sequential()
model2.add(Dense(**keras_params))
model2.add(Dense(**keras_params))
model2.add(Dense(**keras_params))
model2.add(Dense(1, activation=keras_params['activation']))

model2.compile(loss="mean_absolute_error")
earlystopping = callbacks.EarlyStopping(monitor="loss",
                                         mode="min", patience=10,
                                         restore_best_weights=True)
model2.fit(X_train, y_train, validation_data=(X_test, y_test), epochs=100, batch_s

```

```

Epoch 1/100
3455/3455 [=====] - 4s 1ms/step - loss: 9.4946 - val_loss: 3.5171
Epoch 2/100
3455/3455 [=====] - 4s 1ms/step - loss: 3.5175 - val_loss: 3.5171
Epoch 3/100
3455/3455 [=====] - 4s 1ms/step - loss: 3.5175 - val_loss: 3.5171
Epoch 4/100
3455/3455 [=====] - 4s 1ms/step - loss: 3.5175 - val_loss: 3.5171
Epoch 5/100
3455/3455 [=====] - 4s 1ms/step - loss: 3.5175 - val_loss:

```

```

3.5171
Epoch 6/100
3455/3455 [=====] - 3s 1ms/step - loss: 3.5175 - val_loss:
3.5171
Epoch 7/100
3455/3455 [=====] - 4s 1ms/step - loss: 3.5175 - val_loss:
3.5171
Epoch 8/100
3455/3455 [=====] - 4s 1ms/step - loss: 3.5175 - val_loss:
3.5171
Epoch 9/100
3455/3455 [=====] - 4s 1ms/step - loss: 3.5175 - val_loss:
3.5171
Epoch 10/100
3455/3455 [=====] - 4s 1ms/step - loss: 3.5175 - val_loss:
3.5171
Epoch 11/100
3455/3455 [=====] - 4s 1ms/step - loss: 3.5175 - val_loss:
3.5171
Epoch 12/100
3455/3455 [=====] - 4s 1ms/step - loss: 3.5175 - val_loss:
3.5171
Epoch 13/100
3455/3455 [=====] - 4s 1ms/step - loss: 3.5175 - val_loss:
3.5171
Epoch 14/100
3455/3455 [=====] - 4s 1ms/step - loss: 3.5175 - val_loss:
3.5171
Epoch 15/100
3455/3455 [=====] - 4s 1ms/step - loss: 3.5175 - val_loss:
3.5171
Epoch 16/100
3455/3455 [=====] - 3s 1ms/step - loss: 3.5175 - val_loss:
3.5171

```

Out[13]: <keras.callbacks.History at 0x270b0629fd0>

In [14]: `print(model2.get_config())`

```

{'name': 'sequential_100', 'layers': [{'class_name': 'InputLayer', 'config': {'batch_input_shape': (None, 25), 'dtype': 'float64', 'sparse': False, 'ragged': False, 'name': 'dense_200_input'}}, {'class_name': 'Dense', 'config': {'name': 'dense_200', 'trainable': True, 'dtype': 'float32', 'units': 42, 'activation': 'relu', 'use_bias': True, 'kernel_initializer': {'class_name': 'RandomNormal', 'config': {'mean': 0.0, 'stddev': 0.05, 'seed': None}}, 'bias_initializer': {'class_name': 'Zeros', 'config': {}}, 'kernel_regularizer': None, 'bias_regularizer': None, 'activity_regularizer': None, 'kernel_constraint': None, 'bias_constraint': None}}, {'class_name': 'Dense', 'config': {'name': 'dense_201', 'trainable': True, 'dtype': 'float32', 'units': 42, 'activation': 'relu', 'use_bias': True, 'kernel_initializer': {'class_name': 'RandomNormal', 'config': {'mean': 0.0, 'stddev': 0.05, 'seed': None}}, 'bias_initializer': {'class_name': 'Zeros', 'config': {}}, 'kernel_regularizer': None, 'bias_regularizer': None, 'activity_regularizer': None, 'kernel_constraint': None, 'bias_constraint': None}}, {'class_name': 'Dense', 'config': {'name': 'dense_202', 'trainable': True, 'dtype': 'float32', 'units': 42, 'activation': 'relu', 'use_bias': True, 'kernel_initializer': {'class_name': 'RandomNormal', 'config': {'mean': 0.0, 'stddev': 0.05, 'seed': None}}, 'bias_initializer': {'class_name': 'Zeros', 'config': {}}, 'kernel_regularizer': None, 'bias_regularizer': None, 'activity_regularizer': None, 'kernel_constraint': None, 'bias_constraint': None}}, {'class_name': 'Dense', 'config': {'name': 'dense_203', 'trainable': True, 'dtype': 'float32', 'units': 1, 'activation': 'relu', 'use_bias': True, 'kernel_initializer': {'class_name': 'GlorotUniform', 'config': {'seed': None}}, 'bias_initializer': {'class_name': 'Zeros', 'config': {}}, 'kernel_regularizer': None, 'bias_regularizer': None, 'activity_regularizer': None, 'kernel_constraint': None, 'bias_constraint': None}}]}

```

- BEFORE PREDICTION INVERSE LOG10

In [15]: `train_df['price'] = 10 ** train_df['price']`

```
test_df['price'] = 10 ** test_df['price']
```

```
In [16]: train_df['price']
```

```
Out[16]: 4          2400.0
        6          3800.0
        9          3495.0
       10          3000.0
       15          2795.0
        ...
     124000         2800.0
     124002         2395.0
     124004         1850.0
     124008         4195.0
     124009         4280.0
Name: price, Length: 49352, dtype: float64
```

```
In [17]: y_test = 10 ** y_test
```

```
In [18]: score = model2.evaluate(X_test, y_test, verbose=0)
         print(score)
```

```
3963.275390625
```