

Mini-Project 3: House Price Estimation using Both Visual and Textual Data

CSC 215 Artificial Intelligence Under the guidance of Prof. Haiquan Chen Swapnali Shrikhande Asmita Shrivastav

Set-up with Google colab

```
In [63]: from google.colab import drive  
drive.mount('/content/drive')
```

Drive already mounted at /content/drive; to attempt to forcibly remount, call
drive.mount("/content/drive", force_remount=True).

Imports & utility functions

```
In [0]: # image pre-processing
import glob
import os
import pandas as pd
import numpy as np
import cv2
from matplotlib import pyplot as plt

# data normalization using zscore
def encode_numeric_zscore(df, name, mean=None, sd=None):
    if mean is None:
        mean = df[name].mean()
    if sd is None:
        sd = df[name].std()
    df[name] = (df[name] - mean) / sd

# train-test split
from sklearn.model_selection import train_test_split

# cnn-nn
from keras.layers import Input
from keras.layers.convolutional import Conv2D
from keras.layers.pooling import MaxPooling2D
from keras.layers import Dense
from keras.layers import Flatten
from keras.layers.merge import concatenate
from keras.models import Model
from keras.callbacks import ModelCheckpoint
from keras.callbacks import EarlyStopping

# keras model plot
from keras.utils import plot_model

# rmse
from sklearn import metrics

# Regression chart.
def chart_regression(pred,y,sort=True):
    t = pd.DataFrame({'pred' : pred, 'y' : y.flatten()})
    if sort:
        t.sort_values(by=['y'],inplace=True)
    a = plt.plot(t['y'].tolist(),label='expected')
    b = plt.plot(t['pred'].tolist(),label='prediction')
    plt.ylabel('output')
    plt.legend()
    plt.show()

# Encode text values to dummy variables(i.e. [1,0,0],[0,1,0],[0,0,1] for red,green,blue)
def encode_text_dummy(df, name):
    dummies = pd.get_dummies(df[name])
    for x in dummies.columns:
        dummy_name = "{}-{}".format(name, x)
        df[dummy_name] = dummies[x]
    df.drop(name, axis=1, inplace=True)
```

```
# Encode a numeric column as zscores
def encode_numeric_zscore(df, name, mean=None, sd=None):
    if mean is None:
        mean = df[name].mean()

    if sd is None:
        sd = df[name].std()

    df[name] = (df[name] - mean) / sd

# optimizer tuning
from keras.optimizers import adam
```

Fetch dataset & re-structure it

Create groups of bathroom, bedroom, frontal and kitchen images

```
In [114]: # Load textual data
cols = ["Bedrooms", "Bathrooms", "area", "zipcode", "price"]
df = pd.read_csv("/content/drive/My Drive/Colab Notebooks/Houses Dataset/HouseInfo.txt", sep=" ", header=None, names=cols)

# Bathroom.jpg
bathroom_images = []
for number in range(1, 536):
    for path in glob.glob("/content/drive/My Drive/Colab Notebooks/Houses Dataset/" + str(number) + "_bathroom.jpg"):
        if os.path.isfile(path):
            bathroom_images.append(path)
img = pd.DataFrame(bathroom_images, columns = ['bathroom_img'])
bathroom_img= pd.DataFrame(bathroom_images, columns = ['bathroom_img'])

# Bedroom.jpg
bedroom_images = []
for number in range(1, 536):
    for path in glob.glob("/content/drive/My Drive/Colab Notebooks/Houses Dataset/" + str(number) + "_bedroom.jpg"):
        if os.path.isfile(path):
            bedroom_images.append(path)
img['bedroom_img'] = bedroom_images
bedroom_img= pd.DataFrame(bedroom_images, columns = ['bedroom_images'])

# Frontal.jpg
frontal_images = []
for number in range(1, 536):
    for path in glob.glob("/content/drive/My Drive/Colab Notebooks/Houses Dataset/" + str(number) + "_frontal.jpg"):
        if os.path.isfile(path):
            frontal_images.append(path)
img['frontal_img'] = frontal_images
frontal_img= pd.DataFrame(frontal_images, columns = ['frontal_images'])

# Kitchen.jpg
kitchen_images = []
for number in range(1, 536):
    for path in glob.glob("/content/drive/My Drive/Colab Notebooks/Houses Dataset/" + str(number) + "_kitchen.jpg"):
        if os.path.isfile(path):
            kitchen_images.append(path)
img['kitchen_img'] = kitchen_images
kitchen_img= pd.DataFrame(kitchen_images, columns = ['kitchen_images'])

img.head()
```

Out[114]:

	bathroom_img	bedroom_img	frontal_img	kitchen_img
0	/content/drive/My Drive/Colab Notebooks/Houses...	/content/drive/My Drive/Colab Notebooks/Houses...	/content/drive/My Drive/Colab Notebooks/Houses...	/content/drive/My Drive/Colab Notebooks/Houses...
1	/content/drive/My Drive/Colab Notebooks/Houses...	/content/drive/My Drive/Colab Notebooks/Houses...	/content/drive/My Drive/Colab Notebooks/Houses...	/content/drive/My Drive/Colab Notebooks/Houses...
2	/content/drive/My Drive/Colab Notebooks/Houses...	/content/drive/My Drive/Colab Notebooks/Houses...	/content/drive/My Drive/Colab Notebooks/Houses...	/content/drive/My Drive/Colab Notebooks/Houses...
3	/content/drive/My Drive/Colab Notebooks/Houses...	/content/drive/My Drive/Colab Notebooks/Houses...	/content/drive/My Drive/Colab Notebooks/Houses...	/content/drive/My Drive/Colab Notebooks/Houses...
4	/content/drive/My Drive/Colab Notebooks/Houses...	/content/drive/My Drive/Colab Notebooks/Houses...	/content/drive/My Drive/Colab Notebooks/Houses...	/content/drive/My Drive/Colab Notebooks/Houses...

In [115]: img.shape

Out[115]: (535, 4)

In [98]: bathroom_img.head()

Out[98]:

bathroom_img

0	/content/drive/My Drive/Colab Notebooks/Houses...
1	/content/drive/My Drive/Colab Notebooks/Houses...
2	/content/drive/My Drive/Colab Notebooks/Houses...
3	/content/drive/My Drive/Colab Notebooks/Houses...
4	/content/drive/My Drive/Colab Notebooks/Houses...

In [99]: bedroom_img.head()

Out[99]:

bedroom_images

0	/content/drive/My Drive/Colab Notebooks/Houses...
1	/content/drive/My Drive/Colab Notebooks/Houses...
2	/content/drive/My Drive/Colab Notebooks/Houses...
3	/content/drive/My Drive/Colab Notebooks/Houses...
4	/content/drive/My Drive/Colab Notebooks/Houses...

In [100]: frontal_img.head()

Out[100]:

frontal_images

-
- 0 /content/drive/My Drive/Colab Notebooks/Houses...
 - 1 /content/drive/My Drive/Colab Notebooks/Houses...
 - 2 /content/drive/My Drive/Colab Notebooks/Houses...
 - 3 /content/drive/My Drive/Colab Notebooks/Houses...
 - 4 /content/drive/My Drive/Colab Notebooks/Houses...

In [101]: kitchen_img.head()

Out[101]:

kitchen_images

-
- 0 /content/drive/My Drive/Colab Notebooks/Houses...
 - 1 /content/drive/My Drive/Colab Notebooks/Houses...
 - 2 /content/drive/My Drive/Colab Notebooks/Houses...
 - 3 /content/drive/My Drive/Colab Notebooks/Houses...
 - 4 /content/drive/My Drive/Colab Notebooks/Houses...

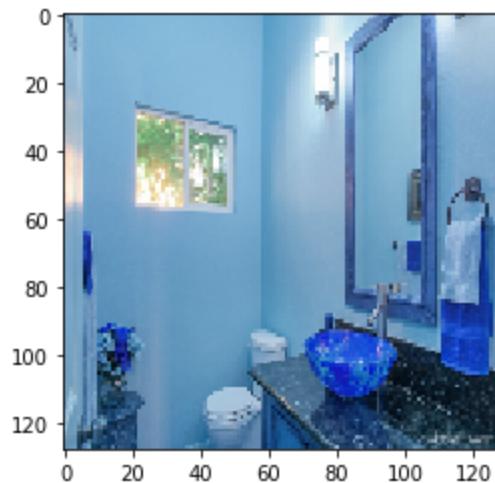
Changing shape of all image array

```
In [116]: print(bathroom_img.shape)
bathroom_output_img = []
for row_index, row in bathroom_img.iterrows():
    outputImage = np.zeros((128, 128, 3), dtype="uint8")
    image_temp = cv2.imread(row.bathroom_img)
    image = cv2.resize(image_temp, (128 , 128))
    outputImage[0:128, 0:128] = image
    bathroom_output_img.append(outputImage)

plt.figure(figsize=(4,4))
plt.imshow(bathroom_output_img[123], interpolation='nearest')
plt.show()

bathroom_arr=np.asarray(bathroom_output_img)
print(bathroom_arr.shape)
```

(535, 1)



(535, 128, 128, 3)

```
In [117]: print(bedroom_img.shape)
bedroom_output_img=[]
for row_index, row in bedroom_img.iterrows():

    outputImage = np.zeros((128, 128, 3), dtype="uint8")
    image_temp = cv2.imread(row.bedroom_images)
    image = cv2.resize(image_temp, (128 , 128))
    outputImage[0:128, 0:128] = image
    bedroom_output_img.append(outputImage)

plt.figure(figsize=(4,4))
plt.imshow(bedroom_output_img[56], interpolation='nearest')
plt.show()

bedroom_arr=np.asarray(bathroom_output_img)
print(bedroom_arr.shape)
```

(535, 1)



(535, 128, 128, 3)

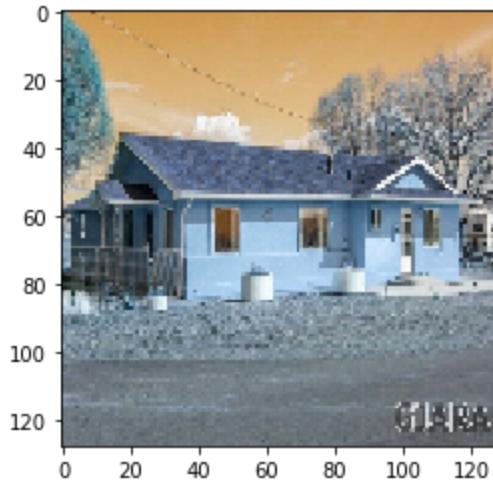
```
In [118]: print(frontal_img.shape)
frontal_output_img=[]
for row_index, row in frontal_img.iterrows():

    outputImage = np.zeros((128, 128, 3), dtype="uint8")
    image_temp = cv2.imread(row.frontal_images)
    image = cv2.resize(image_temp, (128 , 128))
    outputImage[0:128, 0:128] = image
    frontal_output_img.append(outputImage)

plt.figure(figsize=(4,4))
plt.imshow(frontal_output_img[56], interpolation='nearest')
plt.show()

frontal_arr=np.asarray(frontal_output_img)
print(frontal_arr.shape)
```

(535, 1)



(535, 128, 128, 3)

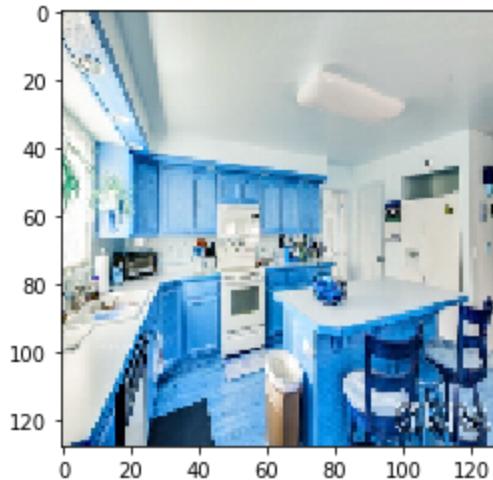
```
In [119]: print(kitchen_img.shape)
kitchen_output_img=[]
for row_index, row in kitchen_img.iterrows():

    outputImage = np.zeros((128, 128, 3), dtype="uint8")
    image_temp = cv2.imread(row.kitchen_images)
    image = cv2.resize(image_temp, (128 , 128))
    outputImage[0:128, 0:128] = image
    kitchen_output_img.append(outputImage)

plt.figure(figsize=(4,4))
plt.imshow(kitchen_output_img[56], interpolation='nearest')
plt.show()

kitchen_arr=np.asarray(kitchen_output_img)
print(kitchen_arr.shape)
```

(535, 1)



(535, 128, 128, 3)

Concatenate house images into one image for each house

```
In [0]: images_output=[]
for row_index, row in img.iterrows():
    inputImages=[]
    outputImage = np.zeros((128, 128, 3), dtype="uint8")
    image_temp1 = cv2.imread(row.bathroom_img)
    image1 = cv2.resize(image_temp1, (64 , 64))

    image_temp2 = cv2.imread(row.bedroom_img)
    image2 = cv2.resize(image_temp2, (64 , 64))

    image_temp3 = cv2.imread(row.frontal_img)
    image3 = cv2.resize(image_temp3, (64 , 64))

    image_temp4 = cv2.imread(row.kitchen_img)
    image4 = cv2.resize(image_temp4, (64 , 64))

    inputImages.append(image1)
    inputImages.append(image2)
    inputImages.append(image3)
    inputImages.append(image4)

    outputImage[0:64, 0:64] = inputImages[0]
    outputImage[0:64, 64:128] = inputImages[1]
    outputImage[64:128, 64:128] = inputImages[2]
    outputImage[64:128, 0:64] = inputImages[3]

    images_output.append(outputImage)
```

```
In [76]: for i in images_output:  
    plt.figure(figsize=(8,8))  
    plt.imshow(i, interpolation='nearest')  
    plt.show()
```





