Boston_Dataset - Analysis

January 17, 2022

[1]:
```
import warnings
warnings.filterwarnings('ignore')
```

[2]:
```
import scrapbook as sb
import pandas as pd
import numpy as np
import seaborn as sns
import numpy as np
from statistics import mean
import matplotlib.pyplot as plt
```

1 Catboost Baseline

[3]:
```
books = sb.read_notebooks("./BaseLine_Model_Output")
baseLine_data = []
for nb in books.notebooks:
    nbList=[nb.scraps['Catboost MSE'].data]
    baseLine_data.append(nbList)
df = pd.DataFrame(baseLine_data, columns = ["Catboost"])
baseLine_data = np.reshape(baseLine_data,(1,10))[0]
display(df)
mse = mean(baseLine_data)
print("Average MSE (Catboost Model): "+ str(mse))
```

```
   Catboost
0  0.081287
1  0.115721
2  0.120925
3  0.196560
4  0.065258
5  0.116890
6  0.117933
7  0.094344
8  0.086313
9  0.085721
```

```
Average MSE (Catboost Model): 0.10809530632925027
```

2 GAN Analysis

[4]:

```
book = sb.read_notebooks("./GAN_Output")
gan_data = []
gan_mse = []
for nb in book.notebooks:
    metrics = nb.scraps['GAN_1 Metrics'].data
    for i in range(1000):
        gan_mse.append(metrics[0][i])
    nbList = [nb.scraps['GAN Model MSE'].data,
              nb.scraps['GAN Model MAE'].data,
              nb.scraps['GAN Model Euclidean distance'].data,
              nb.scraps['GAN Model Manhattan Distance'].data]
    gan_data.append(nbList)

df = pd.DataFrame(gan_data, columns = ['MSE','MAE','Euclidean␣
 ↪Distance','Manhattan Distance'])
display(df.style)
print("MEAN:")
print(df.mean(axis = 0))
gan_data = np.array(gan_data)
```

```
<pandas.io.formats.style.Styler at 0x7fee68b8bc40>

MEAN:
MSE                    0.241390
MAE                    0.319134
Euclidean Distance     4.838867
Manhattan Distance    32.232509
dtype: float64
```

3 ABC_GAN Analysis

3.1 ABC Pre-generator

1. Prior Model is Catboost Model
2. ABC Pre-generator is Catboost Model with gaussian noise -> N(0,variance) where variance : 1, 0.1, 0.01

[5]:

```
book = sb.read_notebooks("./ABC_GAN_Catboost_Output")
paramVal = [1,0.1,0.01]
abc_mse = [[] for i in range(3)]
abc_mse_skip = [[] for i in range(3)]
abc_mse_mean = [[] for i in range(3)]
abc_mse_skip_mean = [[] for i in range(3)]
abc_weights = [[] for i in range(3)]
prior_model = [[] for i in range(3)]
abc_pre_generator = [[] for i in range(3)]
```

```

for nb in book.notebooks:
    metrics1 = np.array(nb.scraps['ABC_GAN_1 Metrics'].data)
    metrics3 = np.array(nb.scraps['ABC_GAN_3 Metrics'].data)
    paramVar = float(nb.papermill_dataframe.iloc[0]['value'])

    #Divide data according to parameters
    for i in range(3):
        if paramVar == paramVal[i]:
            for j in range(100):
                abc_mse[i].append(metrics1[0,j])
                abc_mse_skip[i].append(metrics3[0,j])
            abc_weights[i].append(nb.scraps['Skip Connection Weight'].data)
            prior_model[i].append(nb.scraps['Prior Model MSE'].data)
            abc_pre_generator[i].append(nb.scraps['ABC Pre-generator MSE'].data)
            abc_mse_mean[i].append(mean(metrics1[0,:]))
            abc_mse_skip_mean[i].append(mean(metrics3[0,:]))
```

```
[6]: for i in range(3):
         data = []
         for j in range(len(abc_weights[i])):
             data.append([paramVal[i],prior_model[i][j],
                          abc_pre_generator[i][j],abc_weights[i][j],
                          abc_mse_mean[i][j],abc_mse_skip_mean[i][j]])

         df = pd.DataFrame(data, columns = ['Variance','Prior Model MSE',
                                            'ABC pre-generator MSE','Skip Node␣
      ↪weight',
                                            'ABC_GAN MSE','ABC_GAN MSE (skip␣
      ↪connection)'])
         display(df.round(5))
         print(df.mean(axis=0))
         print("-------------------------------------------------------------")
```

```
   Variance  Prior Model MSE  ABC pre-generator MSE  Skip Node weight  \
0         1          0.07664                1.41763           1.00000
1         1          0.09810                1.07649           0.99588
2         1          0.08852                1.17904           1.00000
3         1          0.12121                1.07123           0.99195
4         1          0.11223                0.97762           1.00000
5         1          0.10835                0.98674           0.99846
6         1          0.08598                1.03053           0.96760
7         1          0.10016                1.21188           0.99153
8         1          0.09026                0.98423           1.00000
9         1          0.10565                0.87045           0.99148

   ABC_GAN MSE  ABC_GAN MSE (skip connection)
0      1.87390                        0.17363
```

```
1      0.19311                        0.18018
2      0.22607                        0.09878
3      0.28449                        0.32872
4      0.46896                        0.15830
5      0.18262                        0.19808
6      0.20929                        0.30623
7      0.28374                        0.14020
8      0.09356                        0.10143
9      0.23253                        0.12707

Variance                         1.000000
Prior Model MSE                  0.098709
ABC pre-generator MSE            1.080584
Skip Node weight                 0.993690
ABC_GAN MSE                      0.404827
ABC_GAN MSE (skip connection)    0.181262
dtype: float64
-------------------------------------------------------------

   Variance  Prior Model MSE  ABC pre-generator MSE  Skip Node weight  \
0       0.1          0.07862                0.08873           0.12253   
1       0.1          0.12962                0.12612           0.00002   
2       0.1          0.07715                0.09444           0.02943   
3       0.1          0.08187                0.09043           0.00000   
4       0.1          0.09539                0.10172           0.01422   
5       0.1          0.16040                0.17203           0.01009   
6       0.1          0.07169                0.08536           0.00712   
7       0.1          0.07879                0.07907           0.00706   
8       0.1          0.11022                0.11378           0.04363   
9       0.1          0.12492                0.12813           0.02369   

   ABC_GAN MSE  ABC_GAN MSE (skip connection)  
0      0.16360                        0.09379  
1      0.51599                        0.13948  
2      0.93204                        0.08948  
3      0.45847                        0.09221  
4   1794.20627                        0.10681  
5      0.37114                        0.17087  
6      1.03522                        0.08438  
7      0.10154                        7.49635  
8      0.31481                    10947.00884  
9      0.39105                        2.92419  

Variance                            0.100000
Prior Model MSE                     0.100868
ABC pre-generator MSE               0.107982
Skip Node weight                    0.025779
ABC_GAN MSE                       179.849014
ABC_GAN MSE (skip connection)    1095.820639
```

```
dtype: float64
-------------------------------------------------------------

   Variance  Prior Model MSE  ABC pre-generator MSE  Skip Node weight  \
0      0.01          0.12767                0.12715           0.00000
1      0.01          0.11476                0.11569           0.00000
2      0.01          0.12687                0.12736           0.00362
3      0.01          0.22559                0.22545           0.02525
4      0.01          0.12226                0.12263           0.04966
5      0.01          0.09726                0.09884           0.14222
6      0.01          0.33550                0.33424           0.09684
7      0.01          0.11077                0.11096           0.23222
8      0.01          0.13501                0.13419           0.00977
9      0.01          0.09971                0.09871           0.12485

   ABC_GAN MSE  ABC_GAN MSE (skip connection)
0      0.12475                   1.276700e-01
1      0.14163                   1.149100e-01
2      0.16125                   1.261300e-01
3      0.43171                   5.927900e-01
4      0.13184                   1.297300e-01
5      0.12617                   7.319190e+06
6      0.30524                   1.339171e+05
7      0.16980                   6.460739e+08
8      0.69948                   1.432400e-01
9      0.11995                   9.240118e+04

Variance                         1.000000e-02
Prior Model MSE                  1.495399e-01
ABC pre-generator MSE            1.495217e-01
Skip Node weight                 6.844368e-02
ABC_GAN MSE                      2.411803e-01
ABC_GAN MSE (skip connection)    6.536194e+07
dtype: float64
-------------------------------------------------------------
```