

t1

January 3, 2023

```
[ ]: import os
import pandas as pd
import numpy as np

[ ]: all_data = os.listdir("./data/ASM_Features")

[ ]: df = pd.read_csv(r'./data/trainLabels.csv')

[ ]: i = 0

[ ]: all_code = ()
for ind in df.index:
    i = i + 1
    if i % 500 == 0 :
        print(f"number {i}")
    fields = ['OPcodes', 'count']
    name = df['Id'][ind]
    object_data = pd.read_csv(f"./data/ASM_Features/{name}.csv",
↪usecols=fields)
    for object_index in object_data.index :
        code = object_data['OPcodes'][object_index]
        all_code = all_code + (code,)

[ ]: all_code_list = list(all_code)
df[all_code_list] = np.nan

[ ]: i = 0
for ind in df.index:
    i = i + 1
    if i % 500 == 0 :
        print(f"number {i}")
    fields = ['OPcodes', 'count']
    name = df['Id'][ind]
    object_data = pd.read_csv(f"./data/ASM_Features/{name}.csv",
↪usecols=fields)
    for object_index in object_data.index :
```

```
code = object_data['OPcodes'][object_index]
# df[f'{code}'] = np.nan
df.at[np.where(df['Id']==name)[0][0], code] =
↳object_data['count'][object_index]
```

```
# except:
#     print(f"error code ={code}")
# try :
#     print(df["OPcodes"][object_index])
# except :
#     pass
```

```
[ ]: df.to_csv("final data.csv")
```