

3DVisualization

May 13, 2024

1 Loading dependencies

```
[1]: %pip install rdkit tensorrt py3Dmol
```

```
Requirement already satisfied: rdkit in /opt/conda/lib/python3.11/site-packages
(2023.9.6)
Requirement already satisfied: tensorrt in /opt/conda/lib/python3.11/site-
packages (10.0.1)
Requirement already satisfied: py3Dmol in /opt/conda/lib/python3.11/site-
packages (2.1.0)
Requirement already satisfied: numpy in /opt/conda/lib/python3.11/site-packages
(from rdkit) (1.26.4)
Requirement already satisfied: Pillow in /opt/conda/lib/python3.11/site-packages
(from rdkit) (10.2.0)
Requirement already satisfied: tensorrt-cu12 in /opt/conda/lib/python3.11/site-
packages (from tensorrt) (10.0.1)
Note: you may need to restart the kernel to use updated packages.
```

1.1 Get library

```
[2]: # Normal Lib
import os
import numpy as np
from typing import cast
```

```
[3]: # keras
from keras import Model
from keras.layers import Conv3D, Input, MaxPooling3D, BatchNormalization,
    Dense, Dropout, Flatten, Activation, GlobalAveragePooling3D
import tensorflow as tf
import keras
```

```
2024-05-12 06:17:16.355246: I tensorflow/core/platform/cpu_feature_guard.cc:210]
This TensorFlow binary is optimized to use available CPU instructions in
performance-critical operations.
To enable the following instructions: AVX2 FMA, in other operations, rebuild
TensorFlow with the appropriate compiler flags.
2024-05-12 06:17:17.168338: W
```

```
tensorflow/compiler/tf2tensorrt/utils/py_utils.cc:38] TF-TRT Warning: Could not find TensorRT
```

```
[4]: print(tf.config.list_physical_devices('GPU'))
```

```
[PhysicalDevice(name='/physical_device:GPU:0', device_type='GPU')]
```

```
[5]: import py3Dmol
```

```
[6]: import csv
```

1.2 Get data and labels zip

```
[7]: protein_folder = './protein'  
protein_list = os.listdir(protein_folder)
```

```
ligand_folder = './ligand'  
ligand_list = os.listdir(ligand_folder)
```

```
label_folder = './label'  
label_list = os.listdir(label_folder)
```

```
[8]: hit_Protein = './protein/3qzr.pdb'
```

```
[9]: hitLead_Protein = './protein/5c1u.pdb'
```

```
[10]: no_Protein = './protein/3qzr.pdb'
```

```
[11]: # ligand_14748_path = os.path.join(ligand_folder, ligand_14748)  
hitLead_Ligand = './ligand/5c1u-CR_model122.pdb'
```

```
[12]: # ligand_14748_path = os.path.join(ligand_folder, ligand_14748)  
hit_Ligand = './ligand/5c1u-GC376_model197.pdb'
```

```
[13]: no_Ligand = './ligand/5c1u-FOPMC_model171.pdb'
```

```
[14]: csv_path = './Final/CSV/resultSFCNN.csv'
```

```
[15]: totalSize = len(ligand_list)  
totalSize
```

```
[15]: 16786
```

```
[16]: # Generate random array of integers from 0 to totalSize - 1  
permu = np.random.RandomState(seed=69).permutation(totalSize)
```

```
[17]: train_num, validate_num, test_num = 0,0,0
```

```
[18]: iDataset_num = totalSize
ratio = (60,20,20)
```

```
[19]: train_num = int(iDataset_num * (ratio[0]/ (ratio[0]+ratio[1]+ratio[2])))
# val_num = int(iDataset_num * (ratio[1]/ (ratio[0]+ratio[1]+ratio[2])))
# test_num = int(iDataset_num * (ratio[2]/ (ratio[0]+ratio[1]+ratio[2])))

val_num = 100
test_num = 500
last_num = 2000
```

```
[20]: train_list_IDs = permu[:train_num]
val_list_IDs = permu[train_num:(train_num+val_num)]
test_list_IDs = permu[(train_num+val_num):(train_num+val_num+test_num)]
last_list_IDs = permu[(train_num+val_num+test_num):
    ↪(train_num+val_num+test_num+last_num)]
```

```
[21]: train_list_IDs
```

```
[21]: array([ 1592, 14669,  5474, ...,  3035,  6871,  6499])
```

```
[22]: dataset_path = './Final/CSV/dataset.csv'
```

1.3 Get features

```
[23]: from utils.preprocessing import get_data_batch
```

1.4 Get metric

```
[24]: from utils.evaluation import regression_metrics, classification_metric
```

1.5 Get Validate

```
[25]: def model_val_dataset(val_dataset_idx, protein_folder, label_folder,
    ↪label_list, batch_size, save_path):
    dataset_len = len(val_dataset_idx)
    runs = dataset_len // batch_size
    last_batch = dataset_len - batch_size*runs
    model = keras.models.load_model(save_path+'.keras')
    # Casting model to keras model
    model = cast(keras.models.Model, model)

    ba_Actual, stat_Actual, ba_Pred, stat_Pred = [], [], [], []
    print("----- Start ValDataset -----")
    for i in range(int(runs+1)):
```

```

print("Get dataset on batch "+str(i+1))
# If it is not the last batch
if i != runs+1:
    gridList, baList, statList = get_data_batch(val_dataset_idx,
    ↪protein_folder, ligand_folder, ligand_list, label_folder, label_list,
    ↪batch_size, i)
    else:
        gridList, baList, statList = get_data_batch(val_dataset_idx,
    ↪protein_folder, ligand_folder, ligand_list, label_folder, label_list,
    ↪last_batch, i)

print("----- Predict ValDataset -----")
result = model.predict(gridList, verbose=2)
gridList = []
pred_reg, pred_class = result

baList = [value for value in baList.tolist()]
statList = [value for value in statList.tolist()]

pred_reg = [value[0] for value in pred_reg.tolist()]
pred_class = [value[0] for value in pred_class.tolist()]

if ba_Actual == []:
    ba_Actual = baList
    stat_Actual = statList
    ba_Pred = pred_reg
    stat_Pred = pred_class
else:
    ba_Actual.extend(baList)
    stat_Actual.extend(statList)
    ba_Pred.extend(pred_reg)
    stat_Pred.extend(pred_class)

baList, statList = [], []
print(ba_Actual, ba_Pred)
PearsonR, MSE = regression_metrics(ba_Actual, ba_Pred)
print("-----")
precision, recall, specificity, NPV, MCC = classification_metric(stat_Actual,
    ↪stat_Pred)
return PearsonR, MSE, precision, recall, specificity, NPV, MCC, ba_Actual,
    ↪ba_Pred, stat_Actual, stat_Pred

```

1.6 Get Train

```
[26]: def model_train_dataset(model: keras.Model, train_dataset_idx, protein_folder,
    ↪label_folder, label_list, batch_size, epochs, save_path: str, best_path):
    dataset_len = len(train_dataset_idx)
    runs = dataset_len // batch_size
    last_batch = dataset_len - batch_size*runs
    cur = 1
    # PearsonR, MSE, RMSE, precision, recall, auc, f1_score, MCC = 0,0,0,0,0,0,0,0
    # checkPearsonR, checkMCC, checkRMSE = 0,0,999
    log_txt = "log.txt"
    log_path = os.path.join(save_path, log_txt)
    readline = ''
    if os.path.exists(log_path):
        log_file = open(log_path, "r+")
        readline = log_file.readline()
        log_file.close()
    else:
        with open(log_path, 'w+') as f:
            f.write('0/'+str(runs))
            f.close()

    if readline == '' or int(readline.split('/')[0]) > runs or int(readline.
    ↪split('/')[0]) == 0:
        cur = 1
        model.save(save_path + '.keras')
    else:
        cur = int(readline.split('/')[0])

    check = 0

    print("----- Start TrainDataset -----")
    for i in range(int(cur-1),int(runs+1)):
        print("=====Batch "+
    ↪str(i+1)+"=====")
        model = keras.models.load_model(save_path + '.keras') # type: ignore
        print("Get dataset")
        # get_data_batch(dataset_idx, protein_folder, ligand_folder, ligand_list,
    ↪label_folder, label_list, batch_size, index)
        # get_data_batch(train_list_IDs, protein_folder, ligand_folder,
    ↪ligand_list, label_folder, label_list, 32, 0)
        if i != runs+1:
            gridList, baList, statList = get_data_batch(train_dataset_idx,
    ↪protein_folder, ligand_folder, ligand_list, label_folder, label_list,
    ↪batch_size, i)
        else:
```

```

        gridList, baList, statList = get_data_batch(train_dataset_idx,
↳protein_folder, ligand_folder, ligand_list, label_folder, label_list,
↳last_batch, i)

        if len(gridList) == 0 or len(baList) == 0 or len(statList) == 0:
            raise ValueError(f'Empty List: gridList: {len(gridList)}, baList:
↳{len(baList)}, statList: {len(statList)}')
        print("----- Train TrainDataset -----")
        model.fit(gridList, [baList, statList], epochs= epochs, verbose=0)
        gridList, baList, statList = [], [], []
        # PearsonR, MSE, RMSE, precision, recall, auc, f1_score, MCC =
↳model_val_dataset(val_dataset_idx, protein_folder, label_folder, label_list,
↳batch_size, epochs, save_path, best_path)
        print("Save")
        model.save(save_path+'.keras')
        log_file = open(log_path, "r+")
        readline = log_file.write(str(i)+'/'+str(runs))
        log_file.close()
        # if PearsonR > checkPearsonR and MCC > checkMCC and RMSE < checkRMSE:
        #     model.save(best_path)
        #     checkPearsonR = PearsonR
        #     checkMCC = MCC
        #     checkRMSE = RMSE
        if check == 0 and batch_size*i >= 2000:
            check +=1
            model.save(best_path + '.keras')
    return model

```

1.7 Get Model

```

[27]: def combine_embedding(drop_rate, input_shape= (52,52,52,11)):
    inp = Input(shape= input_shape, name='Input_Complexes')

    ## Check there are atoms
    x1 = Conv3D(filters= 8, kernel_size=(1,1,1), padding='same',
↳bias_initializer='zeros', kernel_initializer='glorot_uniform')(inp)
    x1 = BatchNormalization()(x1)
    x1 = Activation('relu')(x1)

    x1 = Conv3D(8, kernel_size=(3,3,3))(x1)
    x1 = BatchNormalization()(x1)
    x1 = Activation('relu')(x1)

    x1 = Conv3D(32, kernel_size=(3,3,3), padding='same')(x1)
    x1 = BatchNormalization()(x1)
    x1 = Activation('relu')(x1)

```

```

x1 = MaxPooling3D(pool_size=2)(x1)

x1 = Conv3D(64,kernel_size=(1,1,1),padding='same')(x1)
x1 = BatchNormalization()(x1)
x1 = Activation('relu')(x1)

x1 = Conv3D(64,kernel_size=(3,3,3),padding='same')(x1)
x1 = BatchNormalization()(x1)
x1 = Activation('relu')(x1)

x1 = MaxPooling3D(pool_size=2)(x1)

x1 = Conv3D(128,kernel_size=(1,1,1),padding='same')(x1)
x1 = BatchNormalization()(x1)
x1 = Activation('relu')(x1)

x1 = Conv3D(128,kernel_size=(3,3,3),padding='same')(x1)
x1 = BatchNormalization()(x1)
x1 = Activation('relu')(x1)
x1 = MaxPooling3D(pool_size=2)(x1)

x1 = Conv3D(256,kernel_size=(1,1,1),padding='same')(x1)
x1 = BatchNormalization()(x1)
x1 = Activation('relu')(x1)

x1 = Conv3D(256,kernel_size=(3,3,3),padding='same')(x1)
x1 = BatchNormalization()(x1)
x1 = Activation('relu')(x1)
x1 = MaxPooling3D(pool_size=2)(x1)

# Global Pooling
x2 = GlobalAveragePooling3D()(x1)

x2 = Dense(256)(x2)
x2 = BatchNormalization()(x2)
x2 = Activation('relu')(x2)
x2 = Dropout(0.5)(x2)

# Flattening
x1 = Flatten()(x1)

x1 = Dense(256)(x1)
x1 = BatchNormalization()(x1)
x1 = Activation('relu')(x1)
x1 = Dropout(0.5)(x1)

# Regression Output

```

```

d1 = Dense(1, kernel_regularizer=keras.regularizers.l2(0.01))(x1)
# Classification Output
d2 = Dense(1, activation='sigmoid')(x2)

return Model(inputs=[inp], outputs=[d1,d2], name='Embedding')

```

1.8 Create model

```

[28]: from utils.preprocessing import set_grid
fake_grid, x,y,z = set_grid('./protein/3qzq.pdb')
shape = np.shape(fake_grid)
shape

```

```

[28]: (52, 52, 52, 14)

```

```

[29]: model = combine_embedding(0.5, shape)
model.summary()

```

```

2024-05-12 06:17:19.055664: I
tensorflow/core/common_runtime/gpu/gpu_device.cc:1928] Created device
/job:localhost/replica:0/task:0/device:GPU:0 with 10367 MB memory: -> device:
0, name: NVIDIA GeForce RTX 3060, pci bus id: 0000:81:00.0, compute capability:
8.6

```

Model: "Embedding"

Layer (type)	Output Shape	Param #	Connected to
Input_Complexes (InputLayer)	(None, 52, 52, 52, 14)	0	-
conv3d (Conv3D)	(None, 52, 52, 52, 8)	120	Input_Complexes[...]
batch_normalization (BatchNormalizatio...	(None, 52, 52, 52, 8)	32	conv3d[0][0]
activation (Activation)	(None, 52, 52, 52, 8)	0	batch_normalizat...
conv3d_1 (Conv3D)	(None, 50, 50, 50, 8)	1,736	activation[0][0]
batch_normalizatio... (BatchNormalizatio...	(None, 50, 50, 50, 8)	32	conv3d_1[0][0]

activation_1 (Activation)	(None, 50, 50, 50, 8)	0	batch_normalizat...
conv3d_2 (Conv3D)	(None, 50, 50, 50, 32)	6,944	activation_1[0][...
batch_normalizatio...	(None, 50, 50, (BatchNormalizatio... 50, 32)	128	conv3d_2[0][0]
activation_2 (Activation)	(None, 50, 50, 50, 32)	0	batch_normalizat...
max_pooling3d (MaxPooling3D)	(None, 25, 25, 25, 32)	0	activation_2[0][...
conv3d_3 (Conv3D)	(None, 25, 25, 25, 64)	2,112	max_pooling3d[0]...
batch_normalizatio...	(None, 25, 25, (BatchNormalizatio... 25, 64)	256	conv3d_3[0][0]
activation_3 (Activation)	(None, 25, 25, 25, 64)	0	batch_normalizat...
conv3d_4 (Conv3D)	(None, 25, 25, 25, 64)	110,656	activation_3[0][...
batch_normalizatio...	(None, 25, 25, (BatchNormalizatio... 25, 64)	256	conv3d_4[0][0]
activation_4 (Activation)	(None, 25, 25, 25, 64)	0	batch_normalizat...
max_pooling3d_1 (MaxPooling3D)	(None, 12, 12, 12, 64)	0	activation_4[0][...
conv3d_5 (Conv3D)	(None, 12, 12, 12, 128)	8,320	max_pooling3d_1[...
batch_normalizatio...	(None, 12, 12, (BatchNormalizatio... 12, 128)	512	conv3d_5[0][0]
activation_5 (Activation)	(None, 12, 12, 12, 128)	0	batch_normalizat...
conv3d_6 (Conv3D)	(None, 12, 12, 12, 128)	442,496	activation_5[0][...

batch_normalizatio... (BatchNormalizatio...	(None, 12, 12, 12, 128)	512	conv3d_6[0][0]
activation_6 (Activation)	(None, 12, 12, 12, 128)	0	batch_normalizat...
max_pooling3d_2 (MaxPooling3D)	(None, 6, 6, 6, 128)	0	activation_6[0][...
conv3d_7 (Conv3D)	(None, 6, 6, 6, 256)	33,024	max_pooling3d_2[...
batch_normalizatio... (BatchNormalizatio...	(None, 6, 6, 6, 256)	1,024	conv3d_7[0][0]
activation_7 (Activation)	(None, 6, 6, 6, 256)	0	batch_normalizat...
conv3d_8 (Conv3D)	(None, 6, 6, 6, 256)	1,769,728	activation_7[0][...
batch_normalizatio... (BatchNormalizatio...	(None, 6, 6, 6, 256)	1,024	conv3d_8[0][0]
activation_8 (Activation)	(None, 6, 6, 6, 256)	0	batch_normalizat...
max_pooling3d_3 (MaxPooling3D)	(None, 3, 3, 3, 256)	0	activation_8[0][...
flatten (Flatten)	(None, 6912)	0	max_pooling3d_3[...
global_average_poo... (GlobalAveragePool...	(None, 256)	0	max_pooling3d_3[...
dense_1 (Dense)	(None, 256)	1,769,728	flatten[0][0]
dense (Dense)	(None, 256)	65,792	global_average_p...
batch_normalizatio... (BatchNormalizatio...	(None, 256)	1,024	dense_1[0][0]
batch_normalizatio... (BatchNormalizatio...	(None, 256)	1,024	dense[0][0]
activation_10 (Activation)	(None, 256)	0	batch_normalizat...

activation_9 (Activation)	(None, 256)	0	batch_normalizat...
dropout_1 (Dropout)	(None, 256)	0	activation_10[0]...
dropout (Dropout)	(None, 256)	0	activation_9[0][...
dense_2 (Dense)	(None, 1)	257	dropout_1[0][0]
dense_3 (Dense)	(None, 1)	257	dropout[0][0]

Total params: 4,216,994 (16.09 MB)

Trainable params: 4,214,082 (16.08 MB)

Non-trainable params: 2,912 (11.38 KB)

```
[30]: optimizer=keras.optimizers.Adam(learning_rate=1e-4)

loss=['mean_squared_error', "binary_crossentropy"]
model.compile(optimizer= optimizer,
              loss= loss, run_eagerly=True)
```

1.9 Save paths

```
[31]: save_path = './SFCNN/Model1'
      best_path = './SFCNN/Best1'
```

1.10 Set hyperparameters

```
[32]: batch_size = 32
      epochs = 150
```

1.11 Train

```
[33]: trained_model = model_train_dataset(model, train_list_IDs, protein_folder,
      ↪label_folder, label_list, batch_size, epochs, save_path, best_path)

----- Start TrainDataset -----
=====Batch 1=====

/opt/conda/lib/python3.11/site-packages/keras/src/saving/saving_lib.py:396:
UserWarning: Skipping variable loading for optimizer 'adam', because it has 98
```

```

variables whereas the saved optimizer has 2 variables.
    trackable.load_own_variables(weights_store.get(inner_path))

Get dataset
----- Train TrainDataset -----

2024-05-12 06:17:27.244799: I
external/local_xla/xla/stream_executor/cuda/cuda_dnn.cc:465] Loaded cuDNN
version 8907

Save
=====Batch 2=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 3=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 4=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 5=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 6=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 7=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 8=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 9=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 10=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 11=====
Get dataset

```

```

----- Train TrainDataset -----
Save
=====Batch 12=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 13=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 14=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 15=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 16=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 17=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 18=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 19=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 20=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 21=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 22=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 23=====
Get dataset

```

```

----- Train TrainDataset -----
Save
=====Batch 24=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 25=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 26=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 27=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 28=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 29=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 30=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 31=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 32=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 33=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 34=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 35=====
Get dataset

```

```

----- Train TrainDataset -----
Save
=====Batch 36=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 37=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 38=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 39=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 40=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 41=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 42=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 43=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 44=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 45=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 46=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 47=====
Get dataset

```

```

----- Train TrainDataset -----
Save
=====Batch 48=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 49=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 50=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 51=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 52=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 53=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 54=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 55=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 56=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 57=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 58=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 59=====
Get dataset

```



```

----- Train TrainDataset -----
Save
=====Batch 60=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 61=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 62=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 63=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 64=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 65=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 66=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 67=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 68=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 69=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 70=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 71=====
Get dataset

```

```

----- Train TrainDataset -----
Save
=====Batch 72=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 73=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 74=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 75=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 76=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 77=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 78=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 79=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 80=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 81=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 82=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 83=====
Get dataset

```

```

----- Train TrainDataset -----
Save
=====Batch 84=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 85=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 86=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 87=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 88=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 89=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 90=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 91=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 92=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 93=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 94=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 95=====
Get dataset

```

```

----- Train TrainDataset -----
Save
=====Batch 96=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 97=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 98=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 99=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 100=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 101=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 102=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 103=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 104=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 105=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 106=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 107=====
Get dataset

```

```

----- Train TrainDataset -----
Save
=====Batch 108=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 109=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 110=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 111=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 112=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 113=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 114=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 115=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 116=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 117=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 118=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 119=====
Get dataset

```

```

----- Train TrainDataset -----
Save
=====Batch 120=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 121=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 122=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 123=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 124=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 125=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 126=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 127=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 128=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 129=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 130=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 131=====
Get dataset

```

```

----- Train TrainDataset -----
Save
=====Batch 132=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 133=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 134=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 135=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 136=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 137=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 138=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 139=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 140=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 141=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 142=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 143=====
Get dataset

```

```

----- Train TrainDataset -----
Save
=====Batch 144=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 145=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 146=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 147=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 148=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 149=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 150=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 151=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 152=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 153=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 154=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 155=====
Get dataset

```



```

----- Train TrainDataset -----
Save
=====Batch 156=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 157=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 158=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 159=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 160=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 161=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 162=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 163=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 164=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 165=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 166=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 167=====
Get dataset

```

```

----- Train TrainDataset -----
Save
=====Batch 168=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 169=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 170=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 171=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 172=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 173=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 174=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 175=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 176=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 177=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 178=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 179=====
Get dataset

```

```

----- Train TrainDataset -----
Save
=====Batch 180=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 181=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 182=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 183=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 184=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 185=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 186=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 187=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 188=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 189=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 190=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 191=====
Get dataset

```

```

----- Train TrainDataset -----
Save
=====Batch 192=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 193=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 194=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 195=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 196=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 197=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 198=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 199=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 200=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 201=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 202=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 203=====
Get dataset

```

```

----- Train TrainDataset -----
Save
=====Batch 204=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 205=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 206=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 207=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 208=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 209=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 210=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 211=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 212=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 213=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 214=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 215=====
Get dataset

```

```

----- Train TrainDataset -----
Save
=====Batch 216=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 217=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 218=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 219=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 220=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 221=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 222=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 223=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 224=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 225=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 226=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 227=====
Get dataset

```

```

----- Train TrainDataset -----
Save
=====Batch 228=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 229=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 230=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 231=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 232=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 233=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 234=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 235=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 236=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 237=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 238=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 239=====
Get dataset

```

```

----- Train TrainDataset -----
Save
=====Batch 240=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 241=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 242=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 243=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 244=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 245=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 246=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 247=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 248=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 249=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 250=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 251=====
Get dataset

```



```

----- Train TrainDataset -----
Save
=====Batch 252=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 253=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 254=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 255=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 256=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 257=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 258=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 259=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 260=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 261=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 262=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 263=====
Get dataset

```

```

----- Train TrainDataset -----
Save
=====Batch 264=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 265=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 266=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 267=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 268=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 269=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 270=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 271=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 272=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 273=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 274=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 275=====
Get dataset

```

```

----- Train TrainDataset -----
Save
=====Batch 276=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 277=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 278=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 279=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 280=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 281=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 282=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 283=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 284=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 285=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 286=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 287=====
Get dataset

```

```

----- Train TrainDataset -----
Save
=====Batch 288=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 289=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 290=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 291=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 292=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 293=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 294=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 295=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 296=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 297=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 298=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 299=====
Get dataset

```

```

----- Train TrainDataset -----
Save
=====Batch 300=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 301=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 302=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 303=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 304=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 305=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 306=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 307=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 308=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 309=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 310=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 311=====
Get dataset

```

```

----- Train TrainDataset -----
Save
=====Batch 312=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 313=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 314=====
Get dataset
----- Train TrainDataset -----
Save
=====Batch 315=====
Get dataset
----- Train TrainDataset -----
Save

```

2 New test

```
[34]: model = cast(keras.models.Model, keras.models.load_model(save_path+'.keras'))
```

```
[35]: model.summary()
```

Model: "Embedding"

Layer (type)	Output Shape	Param #	Connected to
Input_Complexes (InputLayer)	(None, 52, 52, 52, 14)	0	-
conv3d (Conv3D)	(None, 52, 52, 52, 8)	120	Input_Complexes[...]
batch_normalization (BatchNormalizatio...	(None, 52, 52, 52, 8)	32	conv3d[0][0]
activation (Activation)	(None, 52, 52, 52, 8)	0	batch_normalizati...
conv3d_1 (Conv3D)	(None, 50, 50, 50, 8)	1,736	activation[0][0]
batch_normalizatio... (BatchNormalizatio...	(None, 50, 50, 50, 8)	32	conv3d_1[0][0]

activation_1 (Activation)	(None, 50, 50, 50, 8)	0	batch_normalizat...
conv3d_2 (Conv3D)	(None, 50, 50, 50, 32)	6,944	activation_1[0][...
batch_normalizatio... (BatchNormalizatio...	(None, 50, 50, 50, 32)	128	conv3d_2[0][0]
activation_2 (Activation)	(None, 50, 50, 50, 32)	0	batch_normalizat...
max_pooling3d (MaxPooling3D)	(None, 25, 25, 25, 32)	0	activation_2[0][...
conv3d_3 (Conv3D)	(None, 25, 25, 25, 64)	2,112	max_pooling3d[0]...
batch_normalizatio... (BatchNormalizatio...	(None, 25, 25, 25, 64)	256	conv3d_3[0][0]
activation_3 (Activation)	(None, 25, 25, 25, 64)	0	batch_normalizat...
conv3d_4 (Conv3D)	(None, 25, 25, 25, 64)	110,656	activation_3[0][...
batch_normalizatio... (BatchNormalizatio...	(None, 25, 25, 25, 64)	256	conv3d_4[0][0]
activation_4 (Activation)	(None, 25, 25, 25, 64)	0	batch_normalizat...
max_pooling3d_1 (MaxPooling3D)	(None, 12, 12, 12, 64)	0	activation_4[0][...
conv3d_5 (Conv3D)	(None, 12, 12, 12, 128)	8,320	max_pooling3d_1[...
batch_normalizatio... (BatchNormalizatio...	(None, 12, 12, 12, 128)	512	conv3d_5[0][0]
activation_5 (Activation)	(None, 12, 12, 12, 128)	0	batch_normalizat...
conv3d_6 (Conv3D)	(None, 12, 12, 12, 128)	442,496	activation_5[0][...

batch_normalizatio... (BatchNormalizatio...	(None, 12, 12, 12, 128)	512	conv3d_6[0][0]
activation_6 (Activation)	(None, 12, 12, 12, 128)	0	batch_normalizat...
max_pooling3d_2 (MaxPooling3D)	(None, 6, 6, 6, 128)	0	activation_6[0][...
conv3d_7 (Conv3D)	(None, 6, 6, 6, 256)	33,024	max_pooling3d_2[...
batch_normalizatio... (BatchNormalizatio...	(None, 6, 6, 6, 256)	1,024	conv3d_7[0][0]
activation_7 (Activation)	(None, 6, 6, 6, 256)	0	batch_normalizat...
conv3d_8 (Conv3D)	(None, 6, 6, 6, 256)	1,769,728	activation_7[0][...
batch_normalizatio... (BatchNormalizatio...	(None, 6, 6, 6, 256)	1,024	conv3d_8[0][0]
activation_8 (Activation)	(None, 6, 6, 6, 256)	0	batch_normalizat...
max_pooling3d_3 (MaxPooling3D)	(None, 3, 3, 3, 256)	0	activation_8[0][...
flatten (Flatten)	(None, 6912)	0	max_pooling3d_3[...
global_average_poo... (GlobalAveragePool...	(None, 256)	0	max_pooling3d_3[...
dense_1 (Dense)	(None, 256)	1,769,728	flatten[0][0]
dense (Dense)	(None, 256)	65,792	global_average_p...
batch_normalizatio... (BatchNormalizatio...	(None, 256)	1,024	dense_1[0][0]
batch_normalizatio... (BatchNormalizatio...	(None, 256)	1,024	dense[0][0]
activation_10 (Activation)	(None, 256)	0	batch_normalizat...

activation_9 (Activation)	(None, 256)	0	batch_normalizat...
dropout_1 (Dropout)	(None, 256)	0	activation_10[0]...
dropout (Dropout)	(None, 256)	0	activation_9[0][...
dense_2 (Dense)	(None, 1)	257	dropout_1[0][0]
dense_3 (Dense)	(None, 1)	257	dropout[0][0]

Total params: 12,645,160 (48.24 MB)

Trainable params: 4,214,082 (16.08 MB)

Non-trainable params: 2,912 (11.38 KB)

Optimizer params: 8,428,166 (32.15 MB)

```
[36]: test1_list_IDs = permu[12000:12100]
```

```
[37]: PearsonR, MSE, precision, recall, specificity, NPV, MCC, ba_Actual, ba_Pred,
      ↪ stat_Actual, stat_Pred = model_val_dataset(val_list_IDs, protein_folder,
      ↪ label_folder, label_list, batch_size, save_path)
```

```
----- Start ValDataset -----
Get dataset on batch 1
----- Predict ValDataset -----
1/1 - 0s - 279ms/step
Get dataset on batch 2
----- Predict ValDataset -----
1/1 - 0s - 276ms/step
Get dataset on batch 3
----- Predict ValDataset -----
1/1 - 0s - 275ms/step
Get dataset on batch 4
----- Predict ValDataset -----
1/1 - 1s - 781ms/step
[-6.2, -6.5, -5.9, -6.7, -7.2, -6.4, -5.9, -6.3, -6.7, -6.0, -6.4, -7.1, -6.2,
-7.5, -6.9, -7.3, -6.9, -7.3, -6.3, -6.8, -7.1, -6.2, -7.0, -6.8, -6.7, -7.6,
-7.5, -6.3, -6.7, -6.3, -7.9, -7.4, -6.4, -5.8, -7.1, -6.5, -6.9, -6.7, -7.1,
-6.3, -6.7, -6.6, -5.9, -7.1, -6.3, -6.7, -6.5, -7.2, -5.8, -8.1, -7.7, -6.4,
-7.4, -7.0, -6.5, -6.1, -7.0, -6.4, -7.1, -6.4, -7.0, -7.0, -6.2, -6.6, -6.3,
```

```

-6.1, -6.6, -6.7, -6.1, -7.0, -7.2, -6.6, -7.6, -6.6, -6.2, -7.7, -6.4, -7.2,
-6.8, -6.9, -7.5, -7.2, -6.0, -6.6, -6.2, -6.2, -6.9, -6.8, -7.3, -7.0, -7.7,
-7.2, -7.5, -7.2, -7.6, -6.4, -7.2, -6.3, -6.2, -7.0] [-7.5687665939331055,
-6.028539657592773, -7.235199928283691, -7.5687665939331055, -6.280409812927246,
-6.854605674743652, -5.382704734802246, -5.382704734802246, -6.901060104370117,
-7.235199928283691, -7.027408599853516, -6.901060104370117, -7.235199928283691,
-6.220447540283203, -6.318724632263184, -6.854605674743652, -7.5687665939331055,
-7.027408599853516, -7.011621475219727, -6.293680191040039, -6.854605674743652,
-6.177022933959961, -7.235199928283691, -7.5687665939331055, -7.027408599853516,
-6.2170000076293945, -7.011621475219727, -7.235199928283691, -7.011621475219727,
-5.382704734802246, -7.5687665939331055, -7.235199928283691,
-7.5687665939331055, -5.382704734802246, -7.235199928283691, -7.027408599853516,
-6.854605674743652, -6.048478126525879, -7.027408599853516, -7.011621475219727,
-7.011621475219727, -7.011621475219727, -7.5687665939331055, -7.011621475219727,
-7.011621475219727, -7.5687665939331055, -5.382704734802246, -6.854605674743652,
-7.5687665939331055, -6.901060104370117, -7.011621475219727, -7.011621475219727,
-6.854605674743652, -6.854605674743652, -7.027408599853516, -7.235199928283691,
-7.5687665939331055, -7.027408599853516, -6.854605674743652, -6.854605674743652,
-6.854605674743652, -7.5687665939331055, -7.5687665939331055,
-7.235199928283691, -6.156284332275391, -7.235199928283691, -6.854605674743652,
-7.235199928283691, -7.027408599853516, -7.011621475219727, -7.235199928283691,
-7.5687665939331055, -7.011621475219727, -6.854605674743652, -6.901060104370117,
-6.901060104370117, -7.235199928283691, -7.027408599853516, -7.235199928283691,
-7.235199928283691, -7.235199928283691, -7.235199928283691, -7.235199928283691,
-7.5687665939331055, -7.235199928283691, -6.854605674743652, -6.18455696105957,
-5.996206283569336, -7.235199928283691, -6.213775634765625, -6.854605674743652,
-5.382704734802246, -6.901060104370117, -7.5687665939331055,
-7.5687665939331055, -6.854605674743652, -7.568770408630371, -7.235197067260742,
-6.901067733764648, -6.8546037673950195]

```

+++++Regression+++++

Pearson Correlation Coefficient: 0.0705369693363294

Mean Absolute Error: 0.62199855

+++++Classification+++++

TP Result: 29.0

FP Result: 7.0

Precision: 0.8055556

Recall: 0.50877196

Specificity: 0.8372093

NPV: 0.5625

Phi coefficient:0.3568477471006598

```

[38]: PearsonR, MSE, precision, recall, specificity, NPV, MCC, ba_Actual, ba_Pred,
      stat_Actual, stat_Pred = model_val_dataset(test_list_IDs, protein_folder,
      label_folder, label_list, batch_size, save_path)

```

----- Start ValDataset -----

Get dataset on batch 1

```

----- Predict ValDataset -----
1/1 - 0s - 279ms/step
Get dataset on batch 2
----- Predict ValDataset -----
1/1 - 0s - 276ms/step
Get dataset on batch 3
----- Predict ValDataset -----
1/1 - 0s - 275ms/step
Get dataset on batch 4
----- Predict ValDataset -----
1/1 - 0s - 275ms/step
Get dataset on batch 5
----- Predict ValDataset -----
1/1 - 0s - 281ms/step
Get dataset on batch 6
----- Predict ValDataset -----
1/1 - 0s - 277ms/step
Get dataset on batch 7
----- Predict ValDataset -----
1/1 - 0s - 275ms/step
Get dataset on batch 8
----- Predict ValDataset -----
1/1 - 0s - 278ms/step
Get dataset on batch 9
----- Predict ValDataset -----
1/1 - 0s - 274ms/step
Get dataset on batch 10
----- Predict ValDataset -----
1/1 - 0s - 280ms/step
Get dataset on batch 11
----- Predict ValDataset -----
1/1 - 0s - 279ms/step
Get dataset on batch 12
----- Predict ValDataset -----
1/1 - 0s - 280ms/step
Get dataset on batch 13
----- Predict ValDataset -----
1/1 - 0s - 278ms/step
Get dataset on batch 14
----- Predict ValDataset -----
1/1 - 0s - 276ms/step
Get dataset on batch 15
----- Predict ValDataset -----
1/1 - 0s - 277ms/step
Get dataset on batch 16
----- Predict ValDataset -----
1/1 - 3s - 3s/step
[-6.5, -7.2, -7.3, -5.6, -7.5, -5.6, -7.2, -6.3, -8.1, -6.2, -6.5, -7.2, -6.5,

```

-6.2, -6.7, -7.0, -7.2, -6.9, -7.8, -6.7, -7.0, -7.2, -7.6, -7.0, -8.3, -6.4,
 -6.6, -6.8, -6.8, -6.7, -7.4, -6.1, -7.9, -7.6, -7.8, -7.2, -6.4, -6.4, -6.2,
 -5.7, -6.8, -6.7, -7.1, -6.5, -7.1, -6.3, -6.3, -6.8, -8.1, -6.1, -7.2, -7.2,
 -6.5, -6.7, -6.4, -7.9, -7.1, -6.2, -6.5, -7.6, -7.3, -6.4, -6.6, -6.5, -6.1,
 -6.2, -6.4, -6.9, -6.7, -7.0, -6.2, -7.7, -7.4, -6.6, -6.9, -7.1, -7.2, -6.9,
 -6.7, -7.5, -7.1, -6.4, -6.9, -7.6, -7.0, -6.7, -6.3, -6.0, -7.4, -6.8, -7.3,
 -6.5, -7.0, -7.4, -7.8, -6.4, -6.5, -6.7, -6.2, -7.2, -6.1, -7.6, -6.1, -6.4,
 -7.2, -6.8, -7.9, -9.0, -6.4, -6.3, -6.4, -6.4, -6.2, -6.9, -6.3, -7.0, -6.6,
 -6.0, -6.5, -7.6, -6.7, -7.1, -7.1, -6.7, -7.1, -7.4, -6.1, -6.9, -9.2, -6.8,
 -7.5, -7.3, -7.7, -6.3, -6.1, -6.4, -6.5, -6.3, -6.3, -7.7, -6.4, -7.3, -6.6,
 -7.4, -6.3, -6.6, -6.8, -6.8, -6.1, -6.6, -7.1, -6.4, -7.6, -6.2, -7.3, -7.8,
 -6.2, -7.2, -6.6, -7.6, -8.5, -7.2, -6.6, -6.5, -6.7, -7.1, -7.2, -6.4, -6.5,
 -6.2, -7.0, -6.7, -6.4, -7.2, -7.3, -6.5, -9.3, -7.1, -6.5, -6.8, -7.5, -8.3,
 -7.2, -7.0, -6.4, -6.9, -6.9, -6.3, -6.2, -7.1, -7.3, -7.1, -6.8, -6.2, -7.2,
 -7.1, -7.1, -7.0, -6.3, -6.3, -6.7, -6.4, -8.0, -7.1, -6.5, -6.3, -6.8, -6.4,
 -5.8, -6.1, -6.2, -6.2, -7.6, -6.7, -6.8, -7.4, -7.1, -6.7, -8.0, -7.8, -6.6,
 -7.1, -6.3, -6.1, -6.7, -6.7, -6.2, -5.8, -7.1, -6.7, -7.2, -6.5, -6.4, -7.0,
 -6.5, -6.3, -6.9, -6.5, -7.2, -8.1, -6.1, -7.6, -6.3, -6.8, -7.8, -6.9, -6.6,
 -7.3, -6.7, -6.7, -5.9, -6.5, -7.2, -7.5, -6.1, -7.8, -6.8, -6.6, -6.6, -6.0,
 -6.1, -6.0, -7.1, -7.0, -7.0, -6.3, -6.6, -6.7, -5.9, -6.6, -6.3, -7.2, -7.2,
 -7.6, -6.4, -6.4, -7.1, -6.8, -6.8, -6.7, -6.6, -6.3, -7.7, -6.8, -6.7, -7.1,
 -8.4, -6.3, -6.6, -6.8, -7.5, -6.7, -7.6, -6.5, -6.4, -7.0, -7.4, -6.7, -7.2,
 -7.2, -6.7, -7.1, -6.5, -7.9, -7.6, -6.7, -6.7, -7.8, -6.7, -6.8, -7.1, -7.2,
 -6.5, -7.2, -7.5, -6.9, -8.5, -6.3, -6.9, -7.1, -6.0, -7.2, -7.0, -6.3, -6.8,
 -5.7, -6.7, -6.4, -7.2, -6.6, -7.0, -6.9, -7.1, -8.0, -7.6, -7.3, -6.5, -6.8,
 -7.0, -7.4, -6.1, -6.4, -6.7, -6.4, -6.4, -6.9, -7.0, -6.9, -6.6, -5.7, -6.9,
 -6.3, -6.5, -6.6, -7.2, -6.0, -6.4, -6.5, -6.5, -6.0, -7.7, -6.4, -6.8, -6.4,
 -6.7, -7.1, -6.2, -6.1, -6.2, -6.4, -5.6, -7.1, -6.2, -7.2, -6.6, -6.2, -6.7,
 -7.2, -7.1, -7.3, -7.1, -6.2, -8.0, -6.8, -6.9, -5.9, -7.4, -6.2, -6.3, -6.4,
 -6.0, -7.2, -6.8, -6.3, -6.8, -6.3, -6.2, -6.5, -6.3, -6.7, -6.4, -6.9, -6.5,
 -7.2, -7.2, -6.4, -7.9, -6.6, -6.4, -6.6, -6.9, -7.1, -6.2, -6.1, -6.9, -7.2,
 -7.4, -6.5, -7.5, -6.3, -7.2, -6.4, -6.7, -5.9, -7.3, -7.7, -8.3, -7.1, -6.8,
 -6.3, -6.4, -5.7, -7.1, -6.4, -7.1, -7.4, -7.1, -7.4, -6.8, -6.5, -5.9, -5.5,
 -6.7, -7.4, -7.1, -6.9, -6.2, -6.5, -6.2, -7.1, -6.7, -7.3, -7.3, -6.5, -7.8,
 -5.9, -6.8, -6.3, -5.9, -7.4, -6.6, -7.4, -7.7, -6.5, -6.8, -6.7, -5.6, -8.1,
 -6.4, -7.2, -6.1, -6.6, -7.3, -7.7, -6.2, -6.2, -6.2, -6.6, -6.8, -6.8, -6.8,
 -7.2, -6.4, -6.7, -6.0, -7.0, -6.1, -6.4, -6.0, -7.0, -6.5, -6.3, -6.5, -8.2,
 -6.1, -7.1, -6.1, -6.4, -6.6, -7.1] [-6.901060104370117, -7.011621475219727,
 -7.235199928283691, -5.382704734802246, -7.235199928283691, -5.382704734802246,
 -7.027408599853516, -6.854605674743652, -7.011621475219727, -5.382704734802246,
 -7.5687665939331055, -6.288928031921387, -6.854605674743652, -6.901060104370117,
 -7.235199928283691, -7.027408599853516, -6.901060104370117, -7.5687665939331055,
 -6.225028991699219, -7.235199928283691, -7.5687665939331055,
 -7.5687665939331055, -6.854605674743652, -7.027408599853516,
 -7.5687665939331055, -6.901060104370117, -6.126924514770508, -7.027408599853516,
 -6.167058944702148, -6.854605674743652, -6.901060104370117, -7.011621475219727,
 -6.901060104370117, -6.901060104370117, -7.235199928283691, -6.291990280151367,
 -7.235199928283691, -7.5687665939331055, -5.382704734802246, -5.382704734802246,

-6.122589111328125, -7.011621475219727, -6.095429420471191, -7.011621475219727,
-7.235199928283691, -7.027408599853516, -7.235199928283691, -6.178979873657227,
-7.5687665939331055, -5.382704734802246, -7.027408599853516, -7.235199928283691,
-5.382704734802246, -5.382704734802246, -7.5687665939331055, -7.027408599853516,
-5.382704734802246, -6.854605674743652, -6.854605674743652, -6.854605674743652,
-6.854605674743652, -5.382704734802246, -7.235199928283691, -7.011621475219727,
-6.854605674743652, -7.011621475219727, -6.901060104370117, -6.901060104370117,
-7.027408599853516, -6.854605674743652, -6.854605674743652, -5.882505416870117,
-6.592851638793945, -5.382704734802246, -7.027408599853516, -6.854605674743652,
-6.854605674743652, -7.011621475219727, -6.854605674743652, -7.011621475219727,
-7.027408599853516, -7.011621475219727, -6.854605674743652, -7.011621475219727,
-7.5687665939331055, -5.382704734802246, -7.011621475219727, -6.854605674743652,
-7.235199928283691, -6.854605674743652, -6.901060104370117, -7.5687665939331055,
-7.027408599853516, -6.854605674743652, -7.027408599853516, -6.854605674743652,
-7.027408599853516, -7.011621475219727, -7.011621475219727, -7.235199928283691,
-5.382704734802246, -7.027408599853516, -7.011621475219727, -7.5687665939331055,
-7.5687665939331055, -7.5687665939331055, -6.555332183837891,
-6.901060104370117, -7.235199928283691, -7.027408599853516, -6.144098281860352,
-5.382704734802246, -7.011621475219727, -6.142760276794434, -6.901060104370117,
-7.011621475219727, -6.854605674743652, -5.382704734802246, -6.854605674743652,
-7.011621475219727, -5.382704734802246, -7.5687665939331055, -7.235199928283691,
-6.854605674743652, -7.5687665939331055, -6.901060104370117, -5.382704734802246,
-6.854605674743652, -6.517940521240234, -6.854605674743652, -6.186088562011719,
-7.011621475219727, -7.027408599853516, -7.027408599853516, -7.011621475219727,
-7.235199928283691, -7.235199928283691, -5.382704734802246, -6.854605674743652,
-6.901060104370117, -7.027408599853516, -6.901060104370117, -6.16949462890625,
-7.027408599853516, -7.027408599853516, -7.235199928283691, -6.901060104370117,
-7.011621475219727, -6.901060104370117, -5.382704734802246, -7.235199928283691,
-7.235199928283691, -6.901060104370117, -6.1381025314331055, -7.011621475219727,
-7.011621475219727, -6.854605674743652, -7.235199928283691, -7.027408599853516,
-7.235199928283691, -6.901060104370117, -6.288946151733398, -6.901060104370117,
-7.027408599853516, -7.027408599853516, -6.854605674743652, -6.901060104370117,
-5.382704734802246, -5.382704734802246, -6.854605674743652, -6.1739654541015625,
-7.5687665939331055, -7.027408599853516, -7.5687665939331055,
-6.186761856079102, -7.5687665939331055, -6.463435173034668, -6.344721794128418,
-7.011621475219727, -7.011621475219727, -6.901060104370117, -7.5687665939331055,
-6.567815780639648, -6.172545433044434, -7.235199928283691, -5.382704734802246,
-7.011621475219727, -5.382704734802246, -5.382704734802246, -6.901060104370117,
-7.011621475219727, -7.5687665939331055, -7.235199928283691, -5.382704734802246,
-6.2775421142578125, -7.027408599853516, -6.901060104370117, -6.901060104370117,
-6.854605674743652, -7.027408599853516, -6.854605674743652, -6.901060104370117,
-7.5687665939331055, -6.901060104370117, -7.027408599853516, -6.854605674743652,
-7.235199928283691, -7.027408599853516, -5.382704734802246, -7.011621475219727,
-5.382704734802246, -5.382704734802246, -6.901060104370117, -7.235199928283691,
-6.0696001052856445, -7.235199928283691, -6.854605674743652, -6.901060104370117,
-7.235199928283691, -6.901060104370117, -6.901060104370117, -7.5687665939331055,
-6.226378440856934, -5.382704734802246, -7.235199928283691, -7.235199928283691,
-5.382704734802246, -5.382704734802246, -7.027408599853516, -5.382704734802246,

-7.235199928283691, -5.382704734802246, -6.901060104370117, -6.2296905517578125,
-6.854605674743652, -7.011621475219727, -6.854605674743652, -7.235199928283691,
-6.854605674743652, -6.901060104370117, -7.011621475219727, -6.901060104370117,
-5.382704734802246, -6.854605674743652, -7.027408599853516, -7.027408599853516,
-6.901060104370117, -6.901060104370117, -7.235199928283691, -7.027408599853516,
-6.246232986450195, -7.5687665939331055, -6.854605674743652, -7.027408599853516,
-6.854605674743652, -6.901060104370117, -6.901060104370117, -7.011621475219727,
-6.901060104370117, -6.854605674743652, -5.382704734802246, -7.027408599853516,
-5.382704734802246, -6.901060104370117, -7.011621475219727, -5.382704734802246,
-6.214449882507324, -6.901060104370117, -7.5687665939331055, -6.901060104370117,
-7.5687665939331055, -5.917938232421875, -7.027408599853516, -7.027408599853516,
-7.011621475219727, -7.011621475219727, -7.011621475219727, -7.027408599853516,
-7.011621475219727, -6.854605674743652, -6.854605674743652, -7.011621475219727,
-6.901060104370117, -6.854605674743652, -7.011621475219727, -5.382704734802246,
-6.901060104370117, -7.011621475219727, -6.854605674743652, -7.235199928283691,
-6.2167768478393555, -6.901060104370117, -7.027408599853516, -6.901060104370117,
-7.027408599853516, -6.854605674743652, -6.901060104370117, -7.235199928283691,
-6.182711601257324, -6.901060104370117, -7.235199928283691, -7.011621475219727,
-6.151956558227539, -6.901060104370117, -7.235199928283691, -7.5687665939331055,
-6.191339492797852, -6.665726661682129, -7.235199928283691, -6.901060104370117,
-7.235199928283691, -6.854605674743652, -6.854605674743652, -7.027408599853516,
-6.854605674743652, -6.0322370529174805, -7.5687665939331055,
-7.027408599853516, -7.5687665939331055, -7.5687665939331055,
-7.5687665939331055, -6.901060104370117, -7.011621475219727, -5.382704734802246,
-6.18198299407959, -7.5687665939331055, -5.382704734802246, -5.382704734802246,
-7.027408599853516, -7.5687665939331055, -6.901060104370117, -6.400323867797852,
-6.901060104370117, -6.854605674743652, -7.027408599853516, -6.901060104370117,
-7.011621475219727, -6.901060104370117, -7.027408599853516, -6.901060104370117,
-7.235199928283691, -6.12346076965332, -6.854605674743652, -7.011621475219727,
-5.382704734802246, -7.235199928283691, -7.011621475219727, -6.2027435302734375,
-7.235199928283691, -5.382704734802246, -7.5687665939331055,
-7.5687665939331055, -7.027408599853516, -6.854605674743652, -5.382704734802246,
-5.382704734802246, -7.027408599853516, -6.854605674743652, -6.129155158996582,
-7.235199928283691, -6.854605674743652, -7.5687665939331055, -7.011621475219727,
-5.382704734802246, -7.027408599853516, -6.854605674743652, -5.382704734802246,
-7.011621475219727, -7.011621475219727, -7.235199928283691, -5.382704734802246,
-7.5687665939331055, -6.854605674743652, -6.51932430267334, -6.2963457107543945,
-7.011621475219727, -7.5687665939331055, -6.901060104370117, -6.901060104370117,
-7.235199928283691, -7.011621475219727, -7.235199928283691, -7.011621475219727,
-5.382704734802246, -7.5687665939331055, -7.5687665939331055,
-7.235199928283691, -7.027408599853516, -6.854605674743652, -7.011621475219727,
-7.235199928283691, -7.5687665939331055, -6.901060104370117, -5.770956039428711,
-6.854605674743652, -7.011621475219727, -6.854605674743652, -5.382704734802246,
-6.901060104370117, -7.027408599853516, -6.901060104370117, -7.5687665939331055,
-7.235199928283691, -7.027408599853516, -7.011621475219727, -6.854605674743652,
-6.901060104370117, -6.901060104370117, -6.854605674743652, -6.141613006591797,
-5.948367118835449, -7.235199928283691, -5.382704734802246, -7.235199928283691,
-7.027408599853516, -6.146937370300293, -6.901060104370117, -5.382704734802246,

```

-6.901060104370117, -7.235199928283691, -7.235199928283691, -5.382704734802246,
-7.235199928283691, -5.382704734802246, -6.854605674743652, -6.901060104370117,
-7.5687665939331055, -5.932804107666016, -7.5687665939331055,
-7.011621475219727, -7.011621475219727, -7.5687665939331055, -6.901060104370117,
-5.382704734802246, -6.854605674743652, -7.011621475219727, -7.011621475219727,
-6.854605674743652, -6.193777084350586, -6.901060104370117, -5.382704734802246,
-5.382704734802246, -6.901060104370117, -6.901060104370117, -7.027408599853516,
-7.011621475219727, -6.854605674743652, -7.011621475219727, -7.027408599853516,
-7.235199928283691, -7.027408599853516, -7.011621475219727, -6.157717704772949,
-7.011621475219727, -6.813736915588379, -5.382704734802246, -6.901060104370117,
-6.854605674743652, -6.854605674743652, -7.011621475219727, -7.5687665939331055,
-7.235199928283691, -7.027408599853516, -6.901060104370117, -7.235199928283691,
-7.5687665939331055, -5.382704734802246, -7.5687665939331055,
-6.854605674743652, -6.193192481994629, -7.235199928283691, -5.382704734802246,
-6.901060104370117, -7.027408599853516, -5.382704734802246, -7.235199928283691,
-6.901060104370117, -7.5687665939331055, -5.382704734802246, -7.027408599853516,
-7.235220909118652, -6.901043891906738, -7.235220909118652, -7.235220909118652,
-5.382718086242676, -6.8545942306518555, -7.5687761306762695,
-7.5687761306762695, -7.5687761306762695, -7.235220909118652,
-6.8545942306518555, -5.382718086242676, -6.8545942306518555,
-6.177238464355469, -6.8545942306518555, -7.027400970458984,
-7.5687761306762695, -7.235220909118652, -7.027400970458984, -7.011674880981445]

```

+++++Regression+++++

Pearson Correlation Coefficient: 0.24535654491605263

Mean Absolute Error: 0.6146272

+++++Classification+++++

TP Result: 121.0

FP Result: 26.0

Precision: 0.82312924

Recall: 0.44485295

Specificity: 0.88596493

NPV: 0.57223797

Phi coefficient:0.3616552637712131

Save the result into csv files

```

[39]: def to_csv(rows, fields, path):
    with open(path, 'w') as csvfile:
        # creating a csv writer object
        csvwriter = csv.writer(csvfile)

        # writing the fields
        csvwriter.writerow(fields)

        # writing the data rows
        csvwriter.writerows(rows)

```

```
[40]: def prediction_preprocessing(_val_list_IDs):
    rows = []
    for idx, value in enumerate(_val_list_IDs):
        row = []
        row.append(ligand_list[value])
        row.append(stat_Actual[idx])
        row.append(stat_Pred[idx])
        row.append(ba_Actual[idx])
        row.append(ba_Pred[idx])
        # row = (ligand_list[value], stat_Actual[idx], stat_Pred[idx],
        ↪ ba_Actual[idx], ba_Pred[idx])
        # print(type(row))
        rows.append(row)
    print("Rows:", len(rows))
    return rows

def protein_ligand_preprocessing(input_list):
    rows = []
    for idx, value in enumerate(input_list):
        row = []
        row.append(idx+1)
        ligand = ligand_list[value]
        pro = ligand.split('-')[0]
        protein = [value for value in protein_list if pro in value][0]
        protein_name = protein.split('.')[0]
        ligand_name = ligand.split('.')[0]
        row.append(protein_name)
        row.append(ligand_name)

        rows.append(row)
    print("Rows:", len(rows))
    return rows
```

```
[41]: # Write Model', 'Hit_Label', 'Hit_Prediction', 'BindingAffinity_Label',
        ↪ 'BindingAffinity_Prediction -> resultSFCNN.csv
to_csv(
    prediction_preprocessing(val_list_IDs),
    fields = ['Model', 'Hit_Label', 'Hit_Prediction', 'BindingAffinity_Label',
        ↪ 'BindingAffinity_Prediction'],
    path = csv_path
)
```

Rows: 100

```
[42]: to_csv(
    protein_ligand_preprocessing(test1_list_IDs),
    fields = ['Model', 'Protein', 'Ligand'],
```



```

    path = dataset_path
)

```

Rows: 100

3 Draw

```

[43]: view = py3Dmol.view()
view.removeAllModels()
view.setViewStyle({'style':'outline','color':'black','width':0.1})

view.addModel(open(hitLead_Protein,'r').read(),format='pdb')
Prot=view.getModel()
Prot.setStyle({'cartoon':{'arrows':True, 'tubes':True, 'style':'oval', 'color':
    ↪'white'}})
view.addSurface(py3Dmol.VDW,{'opacity':0.8,'color':'white'})

view.addModel(open(hitLead_Ligand,'r').read(),format='mol2')
ref_m1 = view.getModel()
ref_m1.setStyle({},{'stick':{'colorscheme':'redCarbon','radius':0.2}})

view.addModel(open(hit_Ligand,'r').read(),format='mol2')
ref_m2 = view.getModel()
ref_m2.setStyle({},{'stick':{'colorscheme':'blueCarbon','radius':0.2}})

view.addModel(open(no_Ligand,'r').read(),format='mol2')
ref_m3 = view.getModel()
ref_m3.setStyle({},{'stick':{'colorscheme':'yellowCarbon','radius':0.2}})

# ref_m.setStyle({},{'cartoon':{'arrows':True, 'tubes':True, 'style':'oval',
    ↪'color':'red'}})
# ref_m.setStyle({'cartoon':{'arrows':True, 'tubes':True, 'style':'oval',
    ↪'color':'red'}})

# results=Chem.SDMolSupplier('1AZ8_lig_vina_out.sdf')

# p=Chem.MolToMolBlock(results[0],False)

# print('Reference: Magenta | Vina Pose: Cyan')
# print ('Pose: {} | Score: {}'.format(results[0].GetProp('Pose'),results[0].
    ↪GetProp('Score'))))

# view.addModel(p,'mol')
# x = view.getModel()

```

```
# x.setStyle({},{'stick':{'colorscheme':'cyanCarbon','radius':0.2}})

view.zoomTo()
view.show()
```

```
[44]: view = py3Dmol.view()
view.removeAllModels()
view.setStyle({'style':'outline','color':'black','width':0.1})

# view.addModel(open(hitLead_Protein,'r').read(),format='pdb')
# Prot=view.getModel()
# Prot.setStyle({'cartoon':{'arrows':True, 'tubes':True, 'style':'oval',
# ↪ 'color':'white'}})
# view.addSurface(py3Dmol.VDW,{'opacity':0.8,'color':'white'})

view.addModel(open(hitLead_Ligand,'r').read(),format='mol2')
ref_m = view.getModel()
ref_m.setStyle({},{'stick':{'colorscheme':'magentaCarbon','radius':0.2}})
# ref_m.setStyle({},{'cartoon':{'arrows':True, 'tubes':True, 'style':'oval',
# ↪ 'color':'red'}})
# ref_m.setStyle({'cartoon':{'arrows':True, 'tubes':True, 'style':'oval',
# ↪ 'color':'red'}})

# results=Chem.SDMolSupplier('1AZ8_lig_vina_out.sdf')

# p=Chem.MolToMolBlock(results[0],False)

# print('Reference: Magenta / Vina Pose: Cyan')
# print ('Pose: {} | Score: {}'.format(results[0].GetProp('Pose'),results[0].
# ↪ GetProp('Score'))

# view.addModel(p,'mol')
# x = view.getModel()
# x.setStyle({},{'stick':{'colorscheme':'cyanCarbon','radius':0.2}})

view.zoomTo()
view.show()
```

```
[45]: view = py3Dmol.view()
view.removeAllModels()
view.setStyle({'style':'outline','color':'black','width':0.1})

view.addModel(open(hitLead_Protein,'r').read(),format='pdb')
Prot=view.getModel()
```

```

Prot.setStyle({'cartoon':{'arrows':True, 'tubes':True, 'style':'oval', 'color':
↳'white'}})
view.addSurface(py3Dmol.VDW,{'opacity':0.8,'color':'white'})

view.addModel(open(hitLead_Ligand,'r').read(),format='mol2')
ref_m = view.getModel()
ref_m.setStyle({},{'stick':{'colorscheme':'magentaCarbon','radius':0.2}})
# ref_m.setStyle({},{'cartoon':{'arrows':True, 'tubes':True, 'style':'oval',
↳'color':'red'}})
# ref_m.setStyle({'cartoon':{'arrows':True, 'tubes':True, 'style':'oval',
↳'color':'red'}})

# results=Chem.SDMolSupplier('1AZ8_lig_vina_out.sdf')

# p=Chem.MolToMolBlock(results[0],False)

# print('Reference: Magenta | Vina Pose: Cyan')
# print ('Pose: {} | Score: {}'.format(results[0].GetProp('Pose'),results[0].
↳GetProp('Score'))))

# view.addModel(p, 'mol')
# x = view.getModel()
# x.setStyle({},{'stick':{'colorscheme':'cyanCarbon','radius':0.2}})

view.zoomTo()
view.show()

```

```

[46]: view = py3Dmol.view()
view.removeAllModels()
view.setViewStyle({'style':'outline','color':'black','width':0.1})

view.addModel(open(hit_Protein,'r').read(),format='pdb')
Prot=view.getModel()
Prot.setStyle({'cartoon':{'arrows':True, 'tubes':True, 'style':'oval', 'color':
↳'white'}})
view.addSurface(py3Dmol.VDW,{'opacity':0.8,'color':'white'})

view.addModel(open(hit_Ligand,'r').read(),format='mol2')
ref_m = view.getModel()
ref_m.setStyle({},{'stick':{'colorscheme':'magentaCarbon','radius':0.2}})
# ref_m.setStyle({},{'cartoon':{'arrows':True, 'tubes':True, 'style':'oval',
↳'color':'red'}})

```

```

# ref_m.setStyle({'cartoon':{'arrows':True, 'tubes':True, 'style':'oval',
↳ 'color':'red'}})

# results=Chem.SDMolSupplier('1AZ8_lig_vina_out.sdf')

# p=Chem.MolToMolBlock(results[0],False)

# print('Reference: Magenta | Vina Pose: Cyan')
# print ('Pose: {} | Score: {}'.format(results[0].GetProp('Pose'),results[0].
↳ GetProp('Score'))))

# view.addModel(p, 'mol')
# x = view.getModel()
# x.setStyle({},{'stick':{'colorscheme':'cyanCarbon', 'radius':0.2}})

view.zoomTo()
view.show()

```

```

[47]: view = py3Dmol.view()
view.removeAllModels()
view.setViewStyle({'style':'outline', 'color':'black', 'width':0.1})

view.addModel(open(no_Protein, 'r').read(), format='pdb')
Prot=view.getModel()
Prot.setStyle({'cartoon':{'arrows':True, 'tubes':True, 'style':'oval', 'color':
↳ 'white'}})
view.addSurface(py3Dmol.VDW, {'opacity':0.8, 'color':'white'})

view.addModel(open(no_Ligand, 'r').read(), format='mol2')
ref_m = view.getModel()
ref_m.setStyle({},{'stick':{'colorscheme':'magentaCarbon', 'radius':0.2}})
# ref_m.setStyle({},{'cartoon':{'arrows':True, 'tubes':True, 'style':'oval',
↳ 'color':'red'}})
# ref_m.setStyle({'cartoon':{'arrows':True, 'tubes':True, 'style':'oval',
↳ 'color':'red'}})

# results=Chem.SDMolSupplier('1AZ8_lig_vina_out.sdf')

# p=Chem.MolToMolBlock(results[0],False)

# print('Reference: Magenta | Vina Pose: Cyan')
# print ('Pose: {} | Score: {}'.format(results[0].GetProp('Pose'),results[0].
↳ GetProp('Score'))))

```

```
# view.addModel(p, 'mol')  
# x = view.getModel()  
# x.setStyle({}, {'stick': {'colorscheme': 'cyanCarbon', 'radius': 0.2}})  
  
view.zoomTo()  
view.show()
```