

# Workshop II 01: Model Training With Early Stop

## Assignment Workshop 01

1. ให้นักศึกษาทำโจทย์ของไฟล์ Workshop II 01.ipynb ให้เรียบร้อยพร้อมทั้งมาตอบคำถามต่อไปนี้

1.1 อธิบายวิธีการเตรียมข้อมูล (Data Preparation) ของนักศึกษาที่ใช้ในการจัดการข้อมูลก่อนที่จะนำтренโนเมเดลครั้งนี้ สามารถแคปภาพหน้าจอแล้วมาอธิบายได้

1.2 หลักการทำงานของ Early Stopping ในการเทรนโนเมเดล และมีไว้ใช้ทำอะไรได้บ้าง

1.3 แสดงประสิทธิภาพของโนเมเดลที่ได้มีอิปวัดผลกับข้อมูลทดสอบที่ให้มา

ธนากร วงศ์อร 663380301-9

1.1 อธิบายวิธีการเตรียมข้อมูล (Data Preparation) ของนักศึกษาที่ใช้ในการจัดการข้อมูล ก่อนที่จะนำเท treff โนเมเดลครั้งนี้ สามารถแคปภาพหน้าจอแล้วมาอธิบายได้

```
import pandas as pd
import os

def read_image_dir(path):
    filename = []
    label = []
    #read all file from subfolders
    for dirname, _, filenames in os.walk(path):
        for file in filenames:
            f = os.path.join(dirname, file)
            if '.png' in f:
                filename.append(f)
                label.append(dirname.split('/')[4])
    df = pd.DataFrame({'filename':filename, 'label':label})
    return df
```

```
train_img_df = read_image_dir('/content/test/Small_Data_CoV2_train')
test_img_df = read_image_dir('/content/test/Small_Data_CoV2_test')
val_img_df = read_image_dir('/content/test/Small_Data_CoV2_val')
```

```
for df in [train_img_df, test_img_df, val_img_df]:
    le = LabelEncoder()
    df['label'] = le.fit_transform(df['label'])
```

} ກ່າວມຕະຫຼາມ path ໃນນັ້ນຍິນ  
ເປັນບັນຫຼາມເກີນ Qn ຢູ່ນັ້ນ  
Data frame

```
class CustomImageDataset(Dataset):
    def __init__(self, im_filename, y, transform=None):
        self.im_filename = im_filename
        self.y = y
        self.transform = transform

    def __len__(self):
        return len(self.y)

    def __getitem__(self, idx):
        img_path = self.im_filename.iloc[idx]
        image = Image.open(img_path).convert('RGB')
        label = self.y.iloc[idx]

        if self.transform:
            image = self.transform(image)

        image_np = np.array(image)

        return image_np, label
```

} Encode Label

ໃນ Dataframe ມີ  
ໃນ Dataset ທຸລະນະກ່ຽວຂ້ອງ  
ດໍວຍ torch

```
train_img_dataset = CustomImageDataset(train_img_df['filename'], train_img_df['label'], transform = train_transform)
test_img_dataset = CustomImageDataset(test_img_df['filename'], test_img_df['label'], transform = test_transform)
val_img_dataset = CustomImageDataset(val_img_df['filename'], val_img_df['label'], transform = test_transform)
```

```
batch_size=16
train_img_dataloader = DataLoader(train_img_dataset, batch_size = batch_size, shuffle = True, num_workers=num_workers)
test_img_dataloader = DataLoader(test_img_dataset, batch_size = batch_size, shuffle = False, num_workers=num_workers)
val_img_dataloader = DataLoader(val_img_dataset, batch_size = batch_size, shuffle = True, num_workers=num_workers)
```

} in DataLoader ເພື່ອ feed ບັນຫຼາມ Qn ອີກ train model ມາຍ

## 1.2 หลักการทำงานของ Early Stopping ในการเทรนโมเดล และมีไว้ใช้ทำอะไรได้บ้าง

### แนวคิดการทำงาน

1. Quando train model จะเก็บ dn validation loss กี่ต่อกี่ครั้ง  
พิจารณา ค่า parameters ทุก epoch นั้นๆ แล้วกำหนดค่า  
Patience เพื่อกำหนด จำนวนครั้งที่ val loss ยังคงเพิ่มต่อไป
2. หากด่า Val loss หันเพิ่มขึ้น จนครบ ค่า patience  
จะหยุด train และ ย้อนกลับ state ของ model  
ถ้าด่า Val loss ลดลงก็สุด

### ปรับแต่งการทำงาน

- กำหนด train Neural Network ให้ uz ณ train  
เมื่อ model ด่า val loss หันเพิ่มขึ้น

1.3 แสดงประสิทธิภาพของโมเดลที่ได้เมื่อไปวัดผลกับข้อมูลทดสอบที่ใหม่มา

```
Epoch 6 Train loss: 1.128429  Val loss: 1.098781
Epoch 7
-----
Epoch 7 Train loss: 1.127664  Val loss: 1.098327
Epoch 8
-----
Epoch 8 Train loss: 1.126960  Val loss: 1.097971
Epoch 9
-----
Epoch 9 Train loss: 1.126414  Val loss: 1.097588
Epoch 10
-----
Epoch 10 Train loss: 1.125708  Val loss: 1.097212
Done!
```

```
test_loop(test_img_dataloader, model, loss_fn)
```

Test Error:

Accuracy: 41.5%, Avg loss: 1.098190