

## Failure Report

23기 오재현

### 0. 개요

구현에 실패하였습니다. 토치 활용을 잘 해볼 일이 없었어서, 다소 어렵게 느껴졌습니다. 어느 부분에서 막혔는지, 어떤 방식으로 도전했는지 오답노트라도 적어보겠습니다... 미제 출로 체크하셔도(벌금이 부과되더라도) 불만은 없습니다. (제가 못 한 거니까요.)

### 1. 실패 양상

Train.py를 실행하였을때, Input image size가 Patch\_size로 나눌 수 있어야 한다는 메시지가 나왔습니다.그러나 input image의 규격을 조절하는 방법을 확인하지 못했습니다. 애초에 input\_size가 65\*1인데, patch size를 65로 바꾸어봐도 같은 오류메세지,

```
File "/Users/jaehyunoh/miniforge3/envs/dsweek1/lib/python3.9/site-packages/torch/nn/modules/module.py", line 1553, in _wrapped_call_impl
    return self._call_impl(*args, **kwargs)
File "/Users/jaehyunoh/miniforge3/envs/dsweek1/lib/python3.9/site-packages/torch/nn/modules/module.py", line 1562, in _call_impl
    return forward_call(*args, **kwargs)
File "/Users/jaehyunoh/Desktop/학교/학회/YBIGTA/2024-2DS/ViT_for_finance/architectures/vit.py", line 24, in forward
    raise ValueError(f"Input image size ({H}x{W}) must be divisible by the patch size ({self.patch_size}).")
ValueError: Input image size (65x1) must be divisible by the patch size (65).
```

1로 나누면 540800사이즈의 입력값과 호환되지 않는다는 메시지가 나왔습니다.

```
return forward_call(*args, **kwargs)
File "/Users/jaehyunoh/Desktop/학교/학회/YBIGTA/2024-2DS/ViT_for_finance/architectures/vit.py", line 28, in forward
    x = x.permute(0, 2, 3, 1, 4, 5).contiguous().view(B, -1, self.patch_dim)
RuntimeError: shape '[128, -1, 3]' is invalid for input of size 540800
```

### 2. 문제 진단

```
#ViT
class VisionTransformer(nn.Module):
    def __init__(self, img_size=224, patch_size=65, in_chans=3, num_classes=1000, embed_dim=768, depth=12, num_heads=12, mlp_ratio=4.,
                 **kwargs):
        super(VisionTransformer, self).__init__()
        self.num_classes = num_classes
        self.num_features = self.embed_dim = embed_dim
        self.patch_embed = EmbeddingLayer(img_size, patch_size, in_chans, embed_dim)
        self.pos_drop = nn.Dropout(p=drop_rate)

        self.blocks = nn.ModuleList([
            Block(embed_dim, num_heads, mlp_ratio, qkv_bias, drop=drop_rate, attn_drop=attn_drop_rate)
            for _ in range(depth)
        ])
        self.norm = nn.LayerNorm(embed_dim)
        self.head = nn.Linear(embed_dim, num_classes)
```

아무래도 문제는 이 부분이나, train.py와의 어떤 행렬이 안맞아서 인 것 같은데, 데이터 크기 65\*1과, 1\*N(단, 540800과 관련된 문제를 해결 가능한) 지점을 찾지 못하여 생긴 문제 같습니다.

### 3. 하고 싶은 말

많이 부족하지만 열심히 따라가겠습니다...