Marcos Grobas, José Romero

P0 - Sistemas de Recomendación

1 Introducción

2 ¿Cómo hicimos?

$$B = AA^{'}$$

3 Evaluacion

```
1
2  # marcos pongo esto por si nos hace falta usar codigo python, si no
       lo quitamos
3
4  class Attention(nn.Module):
5      def __init__(self, dim, heads=8, dim_head=64, dropout=0., order
       ='first'):
6          super().__init__()
7          inner_dim = dim_head * heads
8          project_out = not (heads == 1 and dim_head == dim)
9
10         self.heads = heads
11         self.scale = dim_head ** -0.5
12         self.order = order  # 'first' or 'second'
13
14         self.attend = nn.Softmax(dim=-1)
15         self.dropout = nn.Dropout(dropout)
16
17         self.qkv = nn.Linear(dim, inner_dim, bias=False)
18
19         self.to_out = nn.Sequential(
20             nn.Linear(inner_dim, dim),
21             nn.Dropout(dropout)
22         ) if project_out else nn.Identity()
23
24     def forward(self, x):
25         w = rearrange(self.qkv(x), 'b n (h d) -> b h n d', h=self.
       heads)
26
27         # Compute (U^T Z)^T (U^T Z)
28         dots = torch.matmul(w, w.transpose(-1, -2)) * self.scale
29
```

```
        if self.order == 'first':
            # First-order Neumann approximation
            # out = (U^T Z) * softmax((U^T Z)^T (U^T Z))
            attn = self.attend(dots)
            attn = self.dropout(attn)
            out = torch.matmul(attn, w)

        elif self.order == 'second':
            # Second-order Neumann approximation
            # out = out_1st - out_2nd

            # First order term: (U^T Z) * softmax((U^T Z)^T (U^T Z))
            attn_1st = self.attend(dots)
            attn_1st = self.dropout(attn_1st)
            out_1st = torch.matmul(attn_1st, w)

            # Second order term: (U^T Z) * softmax(((U^T Z)^T (U^T Z))^2)
            # Compute ((U^T Z)^T (U^T Z))^2
            dots_2nd = torch.matmul(dots, dots)
            attn_2nd = self.attend(dots_2nd)
            attn_2nd = self.dropout(attn_2nd)
            out_2nd = torch.matmul(attn_2nd, w)

            # Combine: subtract second order correction
            out = out_1st - out_2nd

        else:
            raise ValueError(f"order must be 'first' or 'second', got {self.order}")

        out = rearrange(out, 'b h n d -> b n (h d)')
        return self.to_out(out)
```

Listing 1: la atencion