

推荐系统入门系列(五)-NeuralFFM算法理论与实战

何无涯 何无涯的技术小屋 7月11日

点击蓝字，带你发现更大的世界

绳锯木断，水滴石穿！
—— 宋代罗大经

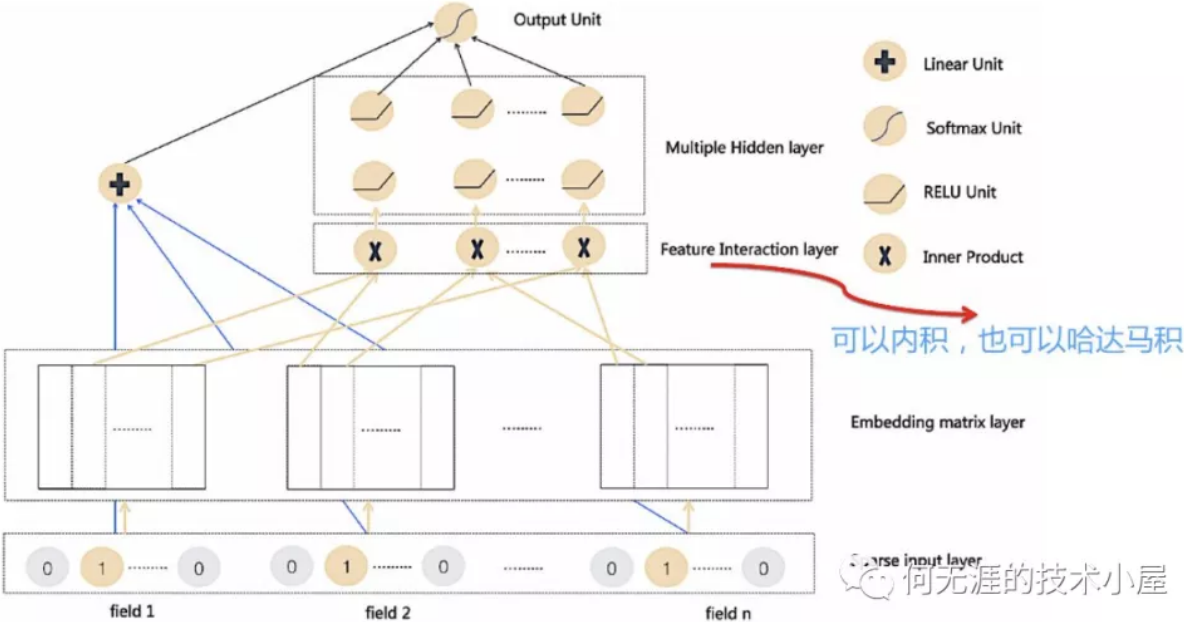
一、NeuralFFM算法背景

DeepFM融合了FM和DNN，分别负责学习低阶特征交互和高阶特征交互。那么很简单的思路，有没有深度神经网络版本的FFM呢，也就是将FFM和DNN结合，当然是有的，这就是NeuralFFM。

二、NeuralFFM算法思想

NeuralFFM，这是2017年南京大学杨毅等人提的，他们在参加腾讯的竞赛中提出了这个模型，这个模型效果比较好，是单模型的第三名。我们知道DeepFM中，为每一个特征学习一个向量，然后在上面做两两特征组合；那么Neural FFM中，为每一个特征学习F个向量，在做特征组合的时候需要做特征交叉，如下NeuralFFM模型结构图：

FFM的神经网络版本：NeuralFFM



其中特征交互层可以做内积也可以做哈达马乘积。

三、NeuralFFM算法实战

我们使用PyTorch实现NeuralFFM算法如下：

```
1 class NeuralFieldAwareFactorizationMachineModel(nn.Module):
2     """
```

```
3 A pytorch implementation of NeuralFFM.
4 """
5 def __init__(self, field_dims, embed_dim, mlp_dims, dropout):
6     super().__init__()
7     self.linear = FeaturesLinear(field_dims)
8     # self.fm = FactorizationMachine(reduce_sum=True)
9     # self.embedding = FeaturesEmbedding(field_dims, embed_dim)
10    self ffm = FieldAwareFactorizationMachine(field_dims, embed_dim)
11    self ffm_output_dim = embed_dim
12    self.mlp = MultilayerPerception(self ffm_output_dim, mlp_dims, dropout)
13
14    def forward(self, x):
15        """
16        :param x: Long tensor of size ``(batch_size, num_fields)``
17        """
18        embed_xs = self ffm(x) # (batch_size, num_combines, embed_dim)
19        # ffm term
20        ffm_term = torch.sum(torch.sum(embed_xs, dim=1), dim=1, keepdim=True)
21        # feature interaction
22        feature_interaction = self.mlp(embed_xs.view(-1, self ffm_output_dim))
23        x = self.linear(x) + ffm_term + torch.sum(feature_interaction.view(-1, self ffm_output_dim), dim=1)
24        return torch.sigmoid(x.squeeze(1))
```

详细的代码请参考我的github：
<https://github.com/yyHaker/RecommendationSystem>。



何无涯的技术小屋

微信号码：leyanyuanyu

机器学习 | 深度学习 | 推荐算法 | NLP | 投资