

# DeepFM [2017]

来自WDL的启发：需要去组合低阶(low-order, Wide 端)和高阶(high-order, Deep端)的特征，方能够既有记忆性又有泛化性。这与搜索中用exact-match和深度模型同时做召回异曲同工。

我们知道FM只能去显式地捕捉二阶交叉信息，而对于高阶的特征组合却无能为力。DeepFM就是在FM模型的基础上，增加DNN部分，进而提高模型对于**高阶组合特征**的信息提取。DeepFM包含了FM和NN两部分，这两部分**共享**了Embedding层（这也是ONN的优化动力，当然了，共享embedding也是有好处的，比如两个塔都可以对共享的embedding进行优化，达到正则化的目的）：

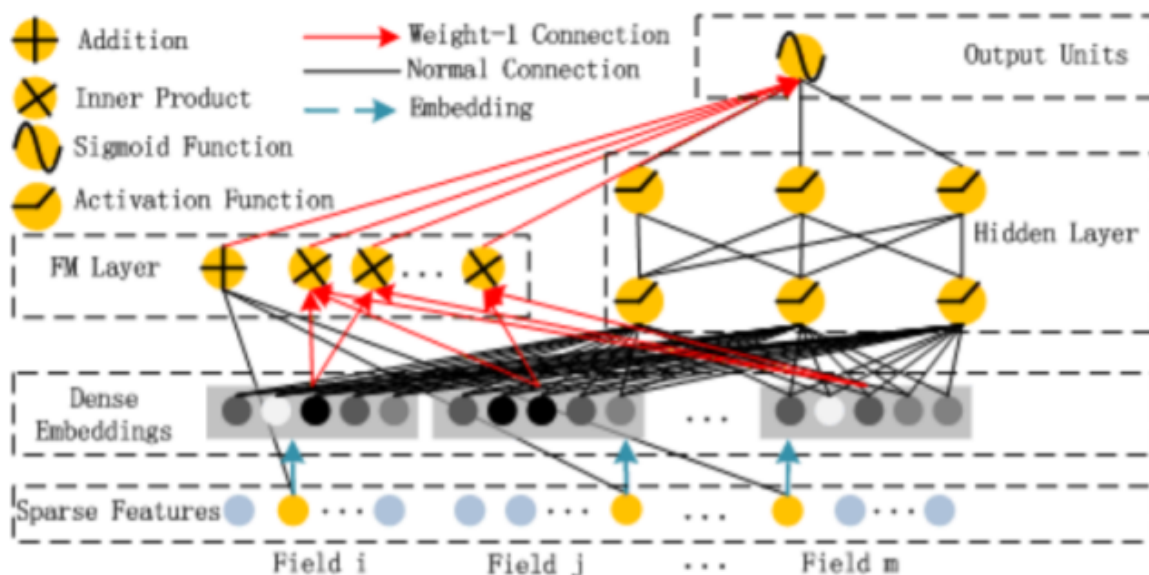


Figure 1: Wide & deep architecture of DeepFM. The wide and deep component share the same input raw feature vector, which enables DeepFM to learn low- and high-order feature interactions simultaneously from the input raw features.

左侧就是一个FM: +号那里，表示的是FM中的一阶项，就是把各个

$w$

加起来。乘号就是内积，最后把所有结果加起来，就是FM这部分的输出。

$$y_{FM} = \langle w, x \rangle + \sum_{i=1}^d \sum_{j=i+1}^d \langle v_i, v_j \rangle x_i x_j$$

右侧是Embedding+MLP，也得到logit输出。最后把两边的输出加起来即可。

$$\hat{y} = \sigma(\text{logits}_{FM} + \text{logits}_{NN})$$

优点：

- 模型具备同时学习低阶与高阶特征的能力
- 共享embedding层，共享了特征的信息表达

**不足：**

- DNN部分对于高阶特征的学习仍然是隐式的，为了解决此问题，可以使用显式有限阶交叉网络DCN、xDeepFM.