1E3S (1 Embedding & 3 Stacking layers)
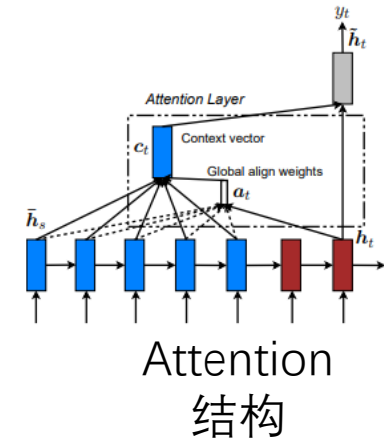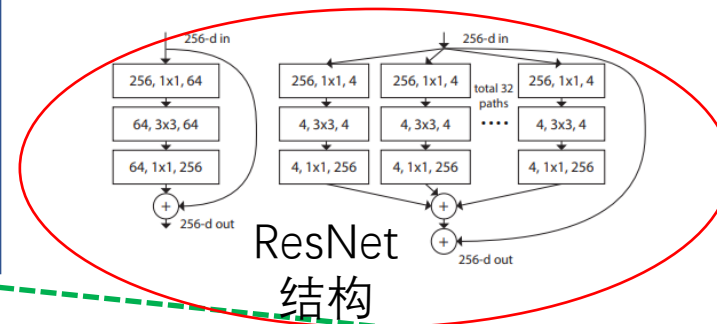
2016 PNN

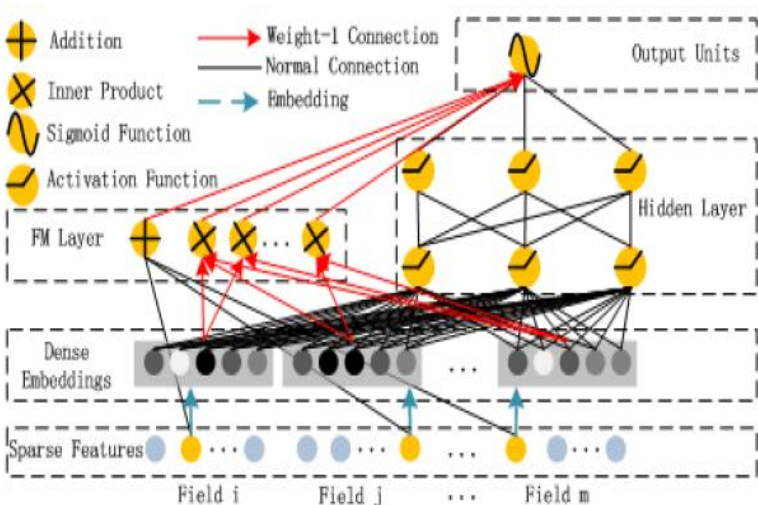**2016 Wide & Deep**

ResNet 结构

Attention 结构

2017 DeepFM

2017 DCN

2017 NCF - NeuMF

LR → Embedding based FM Layer

- ✚ Addition
- ✖ Inner Product
- Sigmoid Function
- Activation Function
- → Weight-1 Connection
- — Normal Connection
- - - → Embedding

Output Units

FM Layer

Hidden Layer

Dense Embeddings

Sparse Features

Field i   Field j   · · ·   Field m

Factorization-Machine based
neural network, DeepFM

LR → Embedding based Cross Netwo

Combination output layer p

$p = \text{sigmoid}(W_{\text{logit}} x_{\text{stack}} + b_{\text{logit}})$

x_{stack}

x_{L_1} h_{L_2}

Cross network   Deep netwrok

x_2 h_2

x_1 h_1

$x_1 = x_0 x_0^T w_{c,0} + b_{c,0} + x_0$ $h_1 = \text{ReLu}(W_{h,0} x_0 + b_{h,0})$

x_0

Embedding and stacking layer

- Dense feature
- Sparse feature
- Embedding vec
- Cross layer
- Deep layer
- Output

$x_{l+1} = x_0 x_l^T w_l + b_l + x_l$

Output   Feature Crossing   Bias   Input
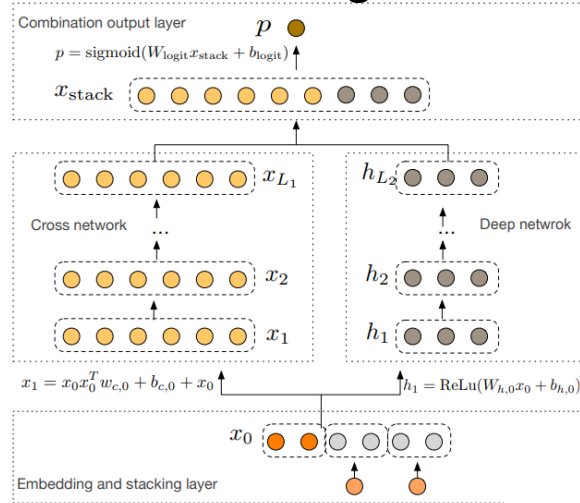
$y = x_0 * x' * w + b + x$

Deep & Cross Network (DCN)

Score $\hat{y}_{ui}$ Training Log loss y_{ui} Target

σ

NeuMF Layer

Concatenation

GMF Layer   MLP Layer X

Element-wise Product   ...... ReLU

MLP Layer 2

ReLU

MLP Layer 1

Concatenation

MF User Vector   MLP User Vector   MF Item Vector   MLP Item Vector

| 0 | 0 | 0 | 1 | 0 | 0 | ...... |

User (u)

| 0 | 0 | 0 | 0 | 1 | 0 | ...... |

Item (i)
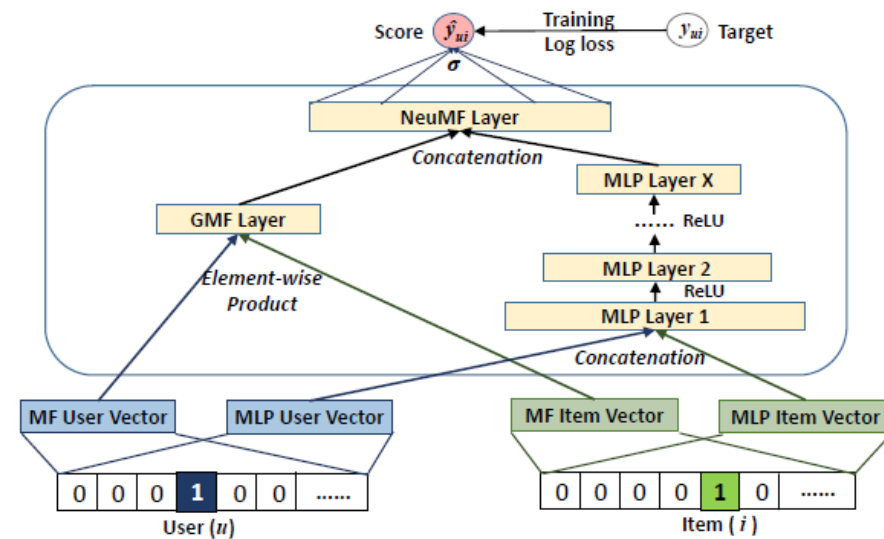
$\phi^{GMF} = \mathbf{p}_u^G \odot \mathbf{q}_i^G,$

$\phi^{MLP} = a_L(\mathbf{W}_L^T(a_{L-1}(...a_2(\mathbf{W}_2^T \begin{bmatrix} \mathbf{p}_u^M \\ \mathbf{q}_i^M \end{bmatrix} + \mathbf{b}_2)...)) + \mathbf{b}_L)$

$\hat{y}_{ui} = \sigma(\mathbf{h}^T \begin{bmatrix} \phi^{GMF} \\ \phi^{MLP} \end{bmatrix}),$

Neural Collaborative Filtering framework, NCF
Neural matrix factorization model, NeuMF

256-d in   256-d in

256, 1x1, 64   256, 1x1, 4   256, 1x1, 4   256, 1x1, 4
64, 3x3, 64   4, 3x3, 4   4, 3x3, 4   total 32 paths   4, 3x3, 4
64, 1x1, 256   4, 1x1, 256   4, 1x1, 256   4, 1x1, 256

256-d out   256-d out

Attention Layer

c_t Context vector

Global align weights

$\bar{h}_s$ a_t h_t

y_t $\bar{h}_t$