

LR, GRU+LR, FM

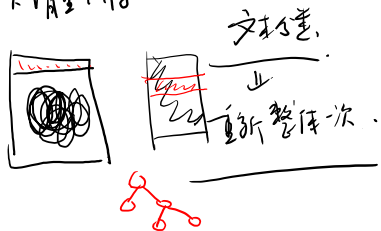
特征: wide&deep, deepfm.
DIN, pnn, deepcross.
FFM, Fi-SiNet.

LR > GRU+LR / FM > deepfm,
 > DIN / FFM

FM:

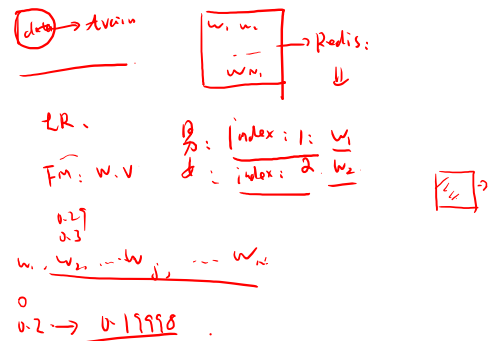
$$\sigma \left(\underbrace{\sum_i \underline{w_i x_i}}_{\substack{w_i \quad v_i \\ \cdot bin \\ \cdot txt}} + \underbrace{\sum_{i=1} \sum_{j=1}^n \langle \underline{v_i v_j} \rangle x_i x_j}_{\substack{t: bin \\ \downarrow \\ t: txt}} \right) \quad \underbrace{\text{field}}_{\substack{(\text{特征id}) \\ \text{男, 女}}}$$

增量训练



LR, FM, ~~GRU~~, wide&deep, deepfm

不同级别的增量训练:



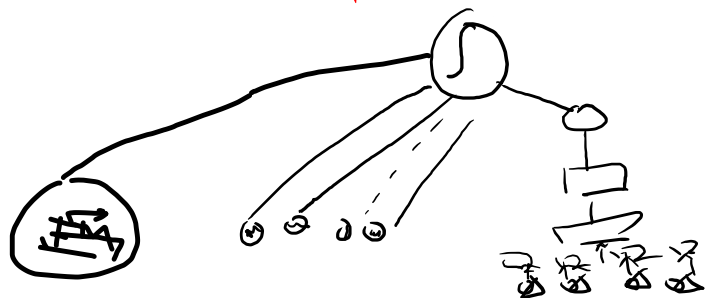
w_1, w_2, w_3

0.3 0.2 -0.5

hp: 0.2117 0.2001 -0.496. $\rightarrow h_2$

sp: 0 0.19 -0.497 $\rightarrow h_1$

wide ~~half~~ deep } w&deep



o o o o

deepctr ; deepmatch , Graph

deepfm :

匹配特征: 词语文章的集合后.

文章的一级分类, 与用户长期兴趣
一级分类的匹配.

user-id: $\square, \square, \square, \dots, \square$

对: 兴趣, weight
用户喜欢的程度, 文章来源

词频, 词性 (类似, 20)

(权重, 类似等), 10个维度的词

20, $\frac{100}{100} > 10$, 50, 5

词频: 分词, 词频的归一化
 $[1, 2, \dots, 10] \times 100000$, 100000

10: \bigcirc
 \bigcirc
 \bigcirc

article-id: simHash
article-id: 相似的文章

文章来源: W, V, X

特征: $\left(\sum W_i x_i + \sum_{j=1}^n \langle v_i, y_j \rangle x_j \right)$

特征: V, F_m
 W, X

特征: 一级, 二级, 三级, 四级

50-200 300 100 Fm

$\left(\frac{100-150}{100} \right) \left[100 \times \text{Number of features} \right]$

1-1 model: 20维 FFM

(1) 特征: V, W

特征: U, W, V, X

特征: U, W, V, X

特征: U, W, V, X

特征: U, W, V, X

特征: U, W, V, X

特征: U, W, V, X

1. Match (2e)

宏观: 增加多样性
微观: embedding

ctr. 时长. ---

多样性 新线文章

思路: 1. 基础行 $\left\{ \begin{array}{l} \text{deepwalk} \\ \text{line} \\ \text{node2vec} \end{array} \right.$

2. side-info: EgEs | Ripple

3. 图种 图种

Gen. Kacm

