

Deep Factorization Machines

Đỗ Hoàng Thuấn

October 2019

Table of contents

- 1 Table of contents
- 2 Introduction
- 3 Factorization Machine
- 4 Deep Factorization Machine
 - Compare to Wide&Deep
- 5 References

Introduction

- Deep factorization machine is SOTA in Click Through Rate probasdalelem.
- DeepFM makes a good result on Zalo Hit Song Challenge.

- 1 Table of contents
- 2 Introduction
- 3 Factorization Machine**
- 4 Deep Factorization Machine
 - Compare to Wide&Deep
- 5 References

Problem statement

Deep FM is build based on FM.

Suppose we have a dataset.

Clicked	Publisher(P)	Advertiser(A)	Gender(G)
Yes	ESPN	Nike	Male
No	NBC	Adidas	Male

Logistic model and limitation

Apply Logistic algorithm.

$$\sigma = w_0 + \sum_{i=1}^n w_i x_i \quad (1)$$

where:

w_0 is bias

w_i is weight of field i

x_i is feature vector of field i

Logistic algorithm does not get combination features.

Can not get the interaction between features.

Factorization Machine

Extends from logistic regression.

Equation of 2-order FM.

$$\sigma = w_0 + \sum_{i=1}^n w_i x_i + \sum_{i=1}^n \sum_{j=i+1}^n \langle v_i, v_j \rangle x_i x_j$$

Where:

w_0 is bias

w_i is weight of field i .

x_i is feature vector of field i .

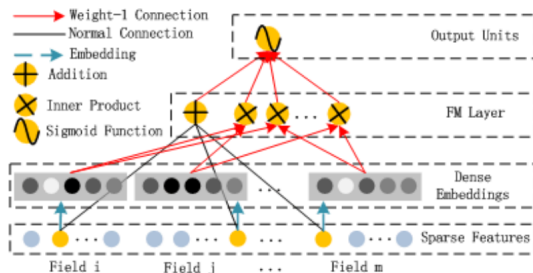
v is latent feature vector.

$\hat{w}_{i,j} = n \langle v_i, v_j \rangle$ is interaction between the i -th and j -th variable.

- 1 Table of contents
- 2 Introduction
- 3 Factorization Machine
- 4 Deep Factorization Machine**
 - Compare to Wide&Deep
- 5 References

Factorization Machine Layer

Integrate the architecture of FM and deep neural networks.

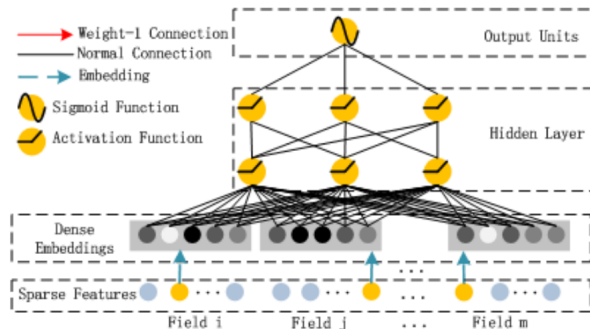


Hình: FM layer in DeepFM model

$$y_{out} = \sigma\left(w_0 + \sum_{i=1}^n w_i x_i + \sum_{i=1}^n \sum_{j=i+1}^n \langle v_i, v_j \rangle x_i x_j\right)$$

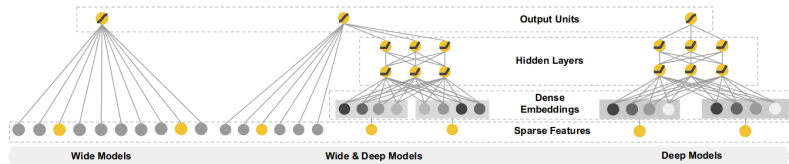
Deep component

Deep component is a feed-forward neural network.
It is used to learn high-order feature interactions.



Hình: DNN layer in DeepFM model

Wide and deep model



Hình: Wide and Deep model

- W&D can model low and high feature interactions simultaneously.
- It needs for manual expertise feature engineering on the input.
- W&D do not generalize to features pairs that have appeared in the training data.

Comparison

	Wide& Deep	DeepFM
High-order feature	✓	✓
Low-order feature	✓	✓
No feature engineering	✓	x

References

- Factorization Machines & their application on huge datasets
- Factorization Machine
- Factorization Machine Paper
- Deep Factorization Machine Paper