

Ad Click Prediction

Ruichen Wang

December 19, 2018

Abstract

There are a lot of algorithms designed for ad ctr prediction. Those algorithms take contextual and historical features as inputs, often combines massive feature engineering work, then predict whether or not the user will click some certain ads. Here I'd like to introduce the traditional solution and recent deep learning technic.

Contents

1	Online Advertising and Linear Model Sparsity	1
1.1	Online Gradient Descent (OGD)	2
1.2	Truncated Gradient (TG)	2
1.3	Forward-Backward Splitting (FOBOS)	2
1.4	Regularized Dual Averaging Algorithm (RDA)	2
1.5	Follow the Regularized Leader (FTRL)	2
2	No Linearity and Feature Engineering	2
2.1	FM	2
2.2	FFM	2
3	Incorporate with Deep Learning	2
3.1	FNN	2
3.2	PNN	2
3.3	Deep and Wide	2
3.4	DeepFM	2
3.5	DNN	2

1 Online Advertising and Linear Model Sparsity

OGD , ,

1.1 Online Gradient Descent (OGD)

1.2 Truncated Gradient (TG)

truncated gradient[1]

1.3 Forward-Backward Splitting (FOBOS)

FOBOS[2]

1.4 Regularized Dual Averaging Algorithm (RDA)

RDA[4]

1.5 Follow the Regularized Leader (FTRL)

FTRL [3]

2 No Linearity and Feature Engineering

2.1 FM

2.2 FFM

3 Incorporate with Deep Learning

3.1 FNN

3.2 PNN

3.3 Deep and Wide

3.4 DeepFM

3.5 DNN

References

- [1] John C. Duchi, Elad Hazan, and Yoram Singer. Adaptive subgradient methods for online learning and stochastic optimization. *Journal of Machine Learning Research*, 12:2121–2159, 2011.
- [2] John C. Duchi and Yoram Singer. Efficient online and batch learning using forward backward splitting. *Journal of Machine Learning Research*, 10:2899–2934, 2009.
- [3] H. Brendan McMahan. Follow-the-regularized-leader and mirror descent: Equivalence theorems and implicit updates. *CoRR*, abs/1009.3240, 2010.

- [4] Lin Xiao. Dual averaging methods for regularized stochastic learning and online optimization. *Journal of Machine Learning Research*, 11:2543–2596, 2010.