# network architecture

# active function

## sigmoid

tf

y = 1 / (1 + exp(-x))

## tanh

## ReLU

https://www.tensorflow.org/versions/r0.9/api\_docs/python/nn.html

max(features, 0)

## softplus

tf

log(exp(features) + 1)

## maxout

# hypothesis function

# loss function

mean square error

absoluter

huber

cross entropy

contrastive loss

# optimization

SGD

# pre Train

# init

distribution

# normaliztion

batch\_normalization

tf

# regularization

dropout

# 模型选择

## 分validation set和test set,

两边的err都低的才是好的.

v err高表明bias高, v err低同时t err高表明bias低的variance高

## ROC,PRC,F1

roc的形状受正负样本分布影响较少

prc考虑到了precision

F1兼顾了precision并且是数字,可用于最优化