

ML report

20170329

edited by GC

In the past two weeks, I briefly learnt the basics and principle of classification.

定堯 and I did some practice few days ago.



Getting Started Prediction Competition

Titanic: Machine Learning from Disaster

Start here! Predict survival on the Titanic and get familiar with ML basics



Kaggle · 6,095 teams · 3 years to go

Overview

Data

Kernels

Discussion

Leaderboard

More

My Submissions

Submit Predictions

Overview

Description

Evaluation

Frequently Asked Questions

Tutorials

Start here if...

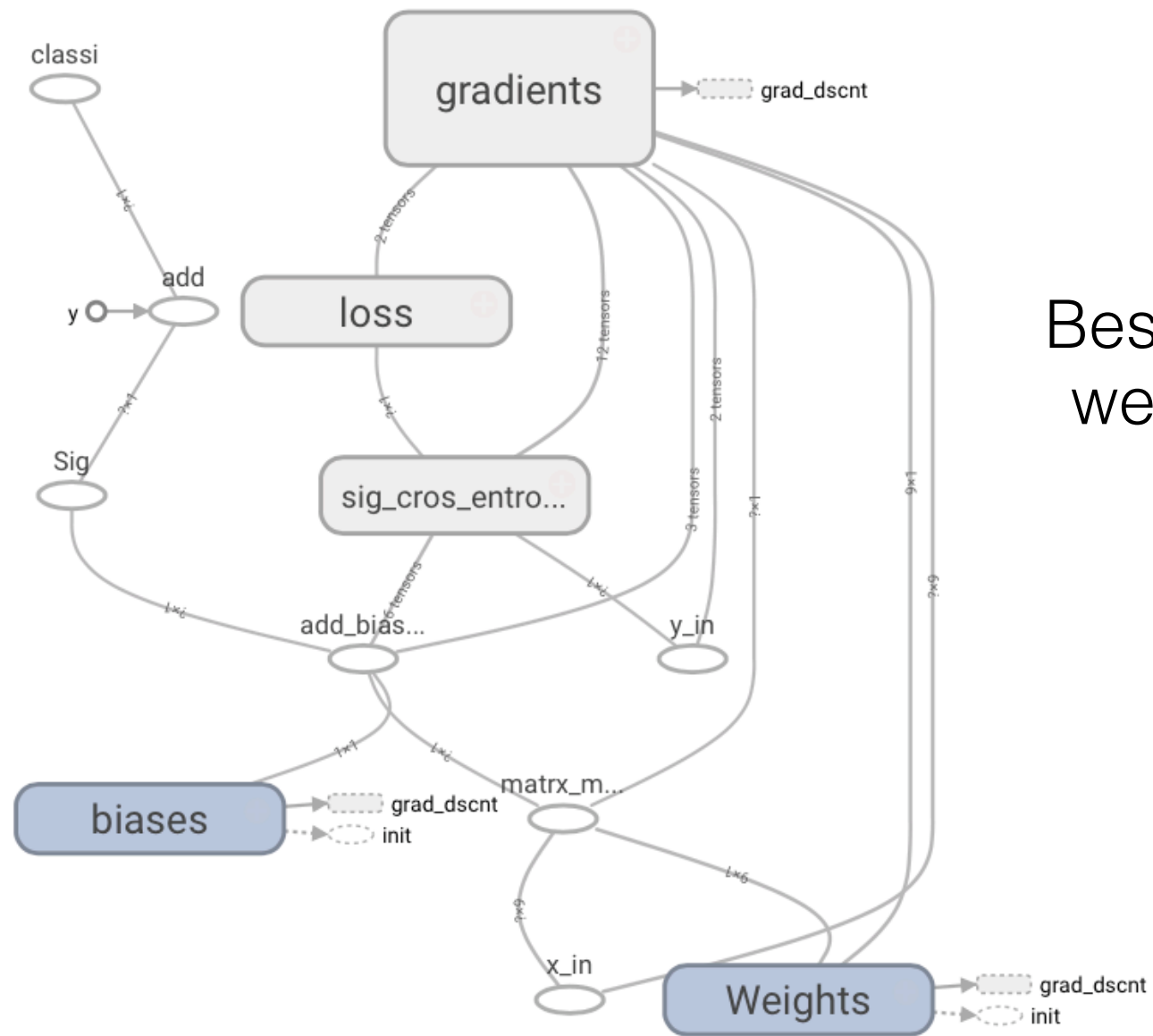
You're new to data science and machine learning, or looking for a simple intro to the Kaggle prediction competitions.

We practiced a common competition for beginner on Kaggle.

<https://www.kaggle.com/c/titanic>

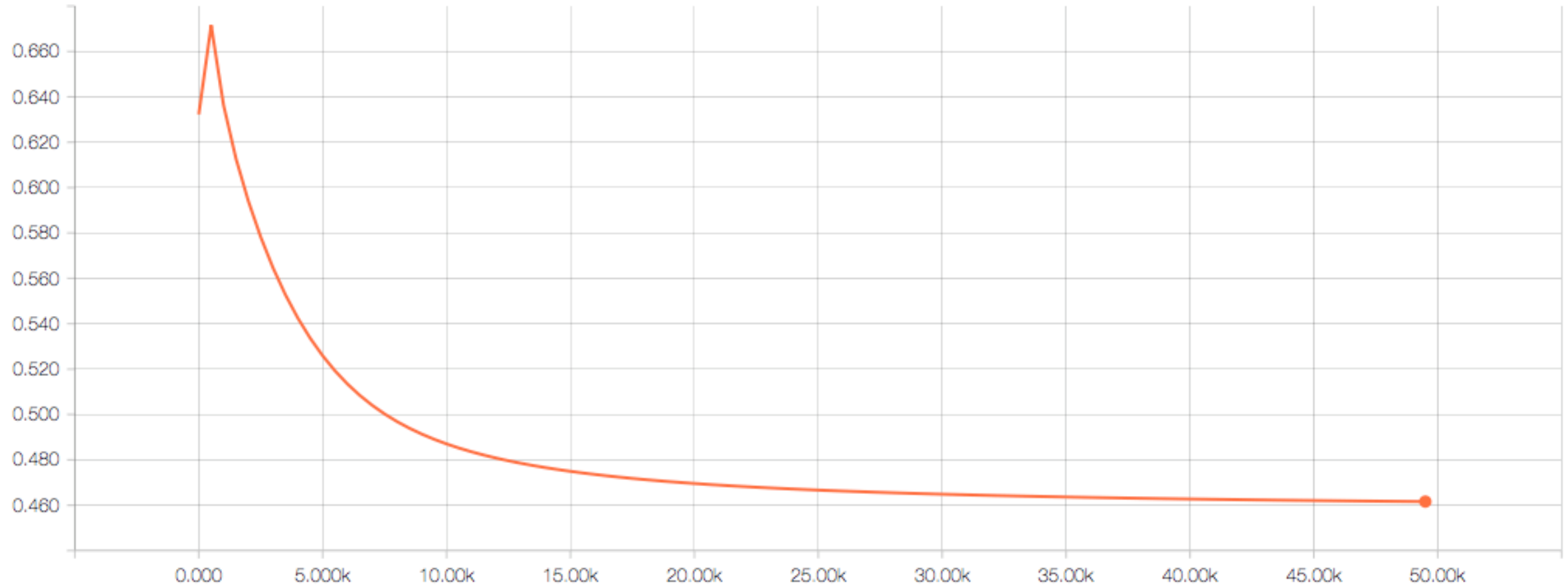
The data processes

- 'name', 'ticket', 'cabin' are affectless so we just drop them.
- the value in 'sex' are changed into 1/0 for the convenience to analyze
- Some values in 'age' are lost, so we fill it with the mean of other ages
- some values in 'embarked' are also lost and we fill it with 'S' due to it's highest appear frequency



Besides the gradient descent technique, we add sigmoid and cross entropy into the training network.

loss/loss



The graph of loss during training.

2475 new GC WhiteShadow



0.78469

6

7m

The score in Kaggle at the end.
Umm..... not bad!

And the source code had already uploaded onto github:

<https://goo.gl/xe4iZG>