# 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.

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
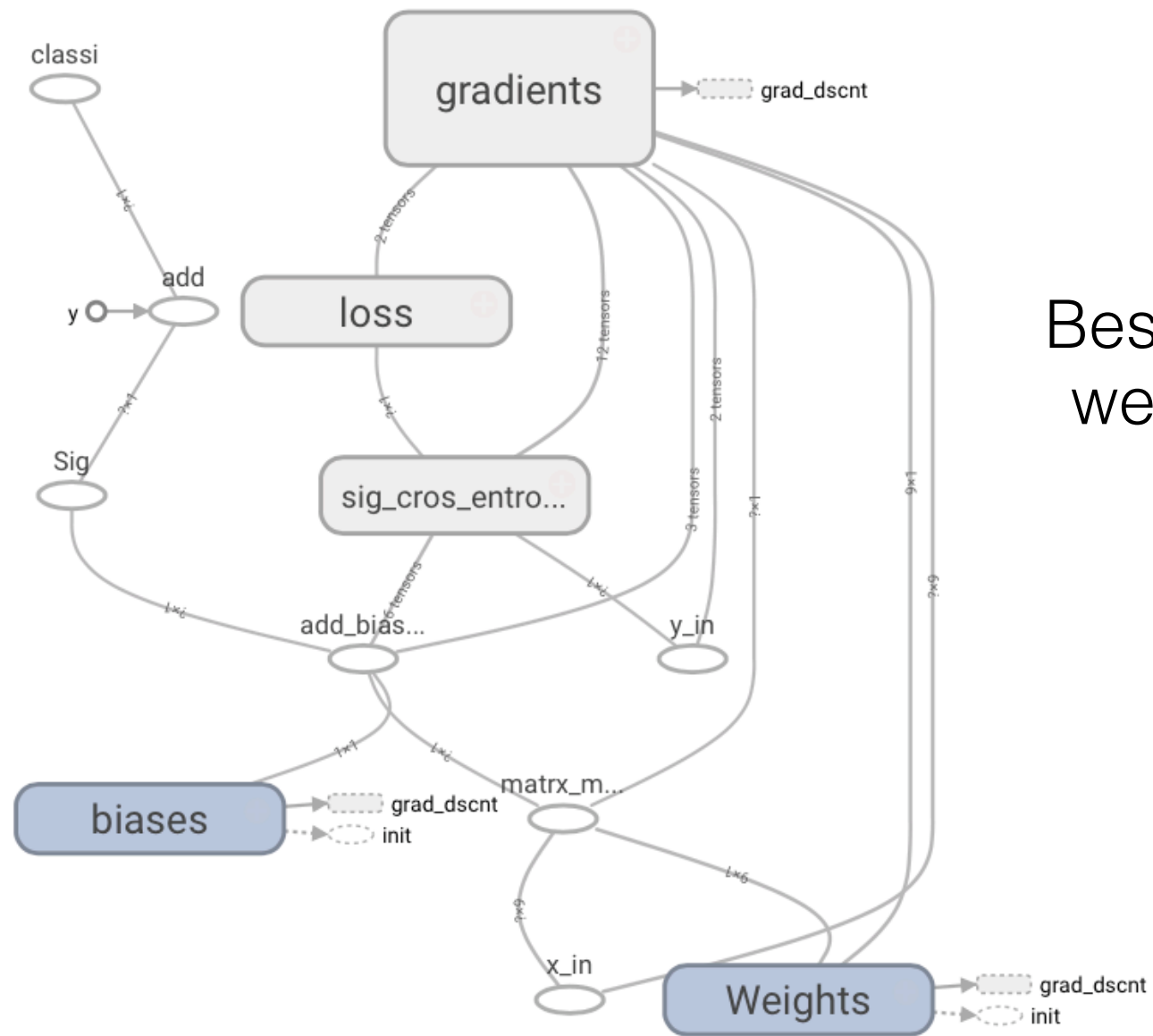
Besides the gradient descent technique, we add sigmoid and cross entropy into the training network.
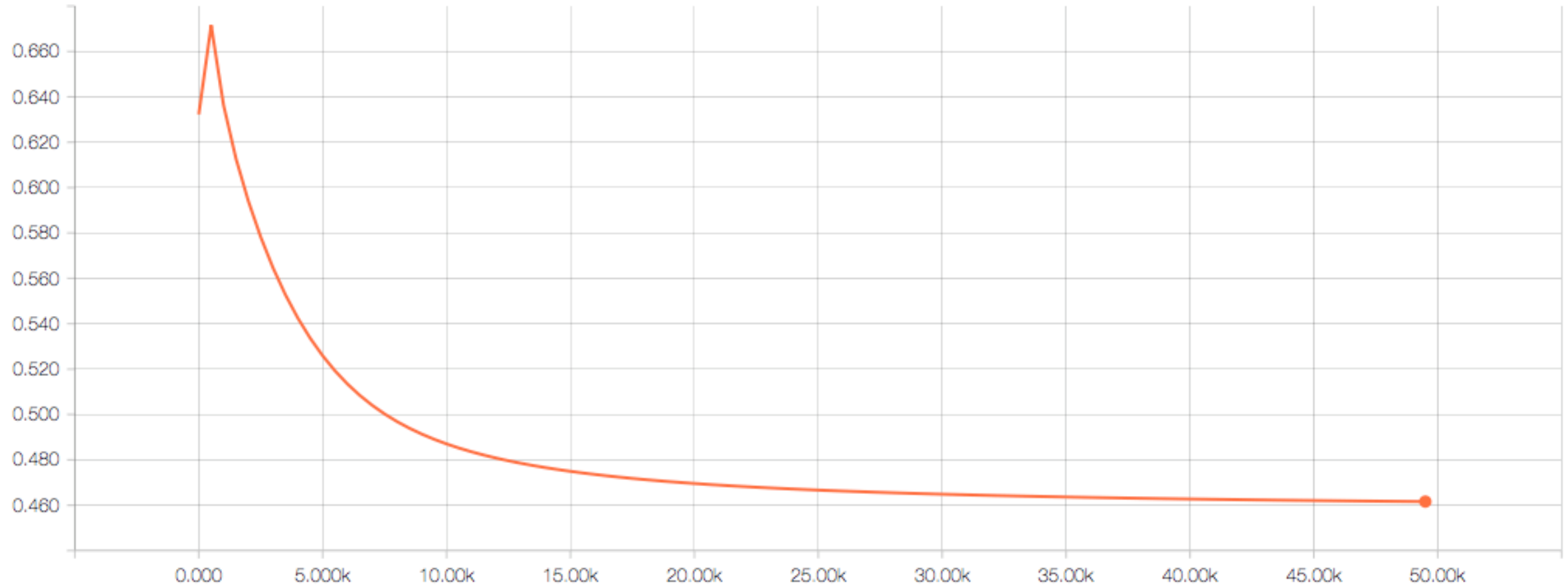
The graph of loss during training.

| 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: