* 18-6-2020

GBM with out-of-fold treatment of high-cardinality feature performs very well.

O. Zhang. Winning data science competitions. <https://www.slideshare.net/ShangxuanZhang/winning-data-science-competitions-presented-by-owen-zhang>, 2015.

* 22-06-2020

The main variation between many boosting algorithms is their method of [weighting](https://en.wikipedia.org/wiki/Weighting) [training data](https://en.wikipedia.org/wiki/Training_data) points and [hypotheses](https://en.wikipedia.org/wiki/Hypothesis). [AdaBoost](https://en.wikipedia.org/wiki/AdaBoost) is very popular and the most significant historically as it was the first algorithm that could adapt to the weak learners. It is often the basis of introductory coverage of boosting in university machine learning courses.[[10]](https://en.wikipedia.org/wiki/Boosting_(machine_learning)#cite_note-11) There are many more recent algorithms such as [LPBoost](https://en.wikipedia.org/wiki/LPBoost" \o "LPBoost), TotalBoost, [BrownBoost](https://en.wikipedia.org/wiki/BrownBoost" \o "BrownBoost), [xgboost](https://en.wikipedia.org/wiki/Xgboost" \o "Xgboost), MadaBoost, [LogitBoost](https://en.wikipedia.org/wiki/LogitBoost" \o "LogitBoost), and others.

<https://en.wikipedia.org/wiki/Boosting_(machine_learning)>

* 28-06-2020

Realised TargetEncoding requires y to be numberic or binary only, not multiclass. Binary needs to be in the form of 1/0. Probably, no function for multiclass is there. Gained more clarity on multiclass from the paper: A preprocessing scheme..: For multiclass, make as many features as are classes from just one high cardinal feature, each feature representing probability of that class for that category.

* 29-06-2020

<https://stats.stackexchange.com/questions/276067/whats-considered-a-good-log-loss>

logloss

[dumb-logloss depending of number of classes - balanced case](https://i.stack.imgur.com/q5Dck.png)