

チーム参加のすすめ ～seed71をとりまく議論から～

2016/3/5
KEI HARADA
(キャッシュ)

自己紹介

- ▶ 某データ分析企業勤務
- ▶ Twitter : @cash_FEG
- ▶ 何でも不器用にこなし、特に数学が得意
 - 数理学博士
 - Project Euler(2013ごろはガチ勢だった)
 - KDDCUP(2011-) 2015は2位になった！
 - ITLeadersに連載してます
 - 2016も頑張ります
 - Kaggle Masterではない
 - @Keiku先輩に先を越された
 - 統計あんまりよくわからないけど統計検定1級とってきた
- ▶ 仕事はSASメインだけど、R/Pythonも使う
 - Rが好き
 - Rcpp, Shiny, Rmarkdownあたりが特に好き、Rstudio最高

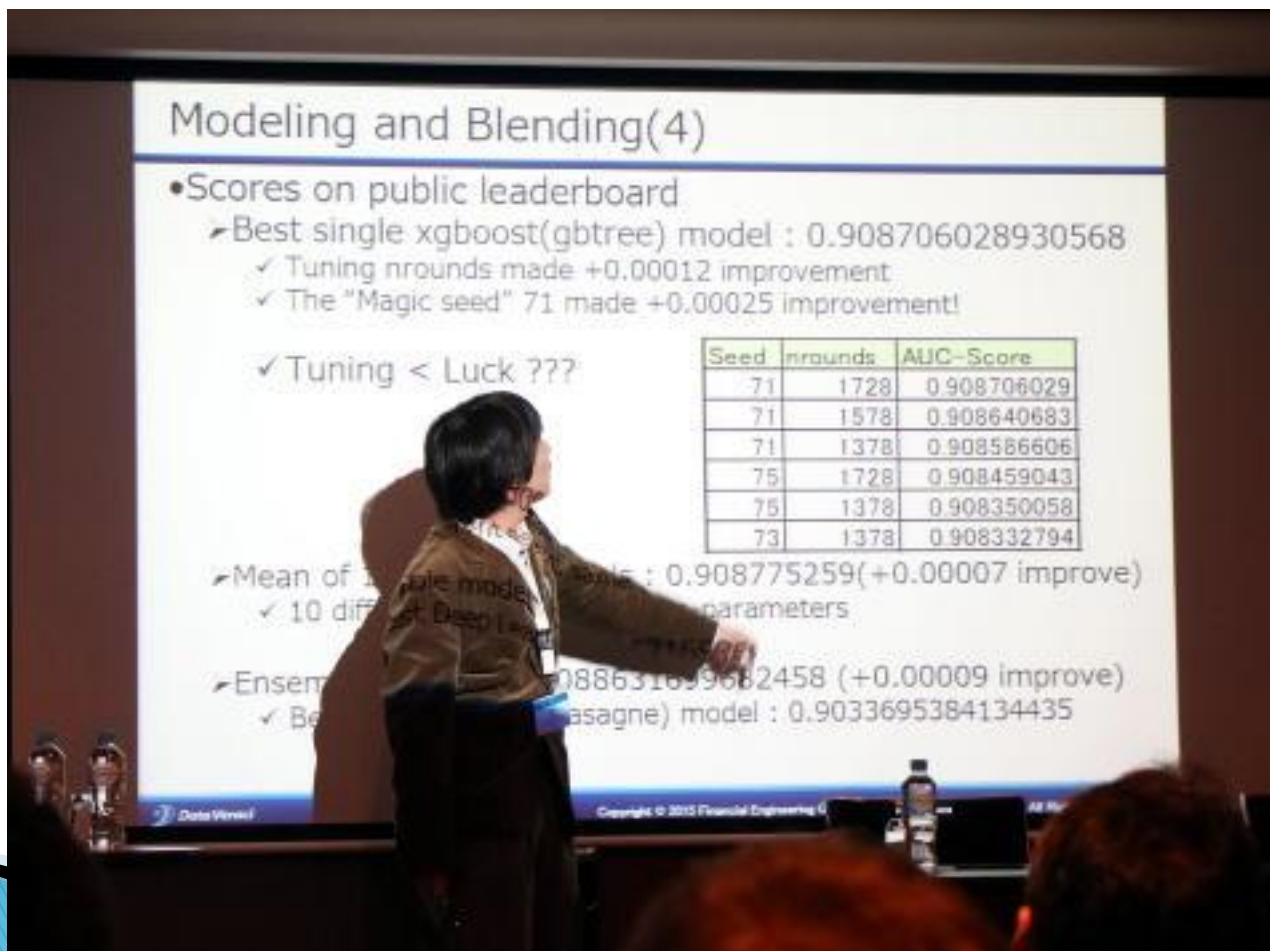


KDDCUP2015で2位になった(1)

- ▶ 開始直後にTOP5入り
 - ばらばらに活動して独立な特徴量を増やす
- ▶ みんな集まったら、そのうち暫定1位になった
 - その後3位まで落ちる
 - 優勝チームのいう”Magic Feature”で1位返り咲き
 - Seedを71にしたらスコア向上、突き放した
- ▶ 終了28時間前に優勝チームに抜かれる

KDDCUP2015で2位になった(2)

- ▶ Seed71の素晴らしさについて世界に発信してきた



Modeling and Blending(4)

- Scores on public leaderboard
 - Best single xgboost(gbtree) model : 0.908706028930568
 - ✓ Tuning nrounds made +0.00012 improvement
 - ✓ The "Magic seed" 71 made +0.00025 improvement!
 - ✓ Tuning < Luck ???
 - Mean of 10 single models : 0.908775259(+0.00007 improve)
 - ✓ 10 different Deep Learning parameters
 - Ensemble : 0.9088631699682458 (+0.00009 improve)
 - ✓ Best (Masagne) model : 0.9033695384134435

Seed	nrounds	AUC-Score
71	1728	0.908706029
71	1578	0.908640683
71	1378	0.908566606
75	1728	0.908459043
75	1378	0.908350058
73	1378	0.908332794

Copyright © 2015 Financial Engineering

今日の結論

- ▶ チーム組んでコンペ参加するとたのしいぞう
- ▶ ソロ活動メインの皆さんもぜひ！

チームを組むと起こるよいこと

- ▶ アイディアが増える
 - 一発目のサブミットが大コケすると心が折れる
 - チームの誰かが成功すればよい
 - 特徴量のアイディアが増える
- ▶ 作業する手が増える
 - アンサンブルモデルではもろに差が出る
 - 最後は力勝負なのでこれも重要
- ▶ 仲間から技術が盗める
 - 何かしら自分よりできるところがあるはず、学べることは多い
 - 例えばRstudio Server使うと自然に盗める

Seed71にみるチームプレー

- ▶ アイディアが増えた
 - 特徴量はすごく増えた
 - XGBoostのseedを変える発想は私にはない
 - seedを変えてうまくいくななんてとんでもない！
 - ちなみに100個seed変えて平均したら71に勝って安心した
- ▶ 作業する手が増えた
 - パラメータチューニングがさくっと終わった
 - 私はチューニングがとても嫌い、だいたい一発目に勝てない
- ▶ 仲間から技術が盗めた
 - Seedを変える根性
 - 効率の良い探索、実験計画

Seed71の他に楽しかったこと

- ▶ 特徴量を全部投入する”genkidama”
 - 結構強い
- ▶ 投稿するときに「よろしく願いしまあああああす」
 - 特に良い効果はなかった
- ▶ 2位に転落してからの28時間は一番団結した
 - 地方の公立校が高校野球の決勝で負けたような感じ

チームを組むと起こるよくないこと

- ▶ 投稿数が減る
 - 自分が投稿したいときにできない
 - 遠慮しあった結果無駄にすることもありがち
- ▶ 経験値が減る
 - 1人でやる方が勉強にはなる
- ▶ モチベーションが下がるときは下がる
 - 何かと自由にできなくなる
 - 協力してもうまくいかなかったときの絶望感
 - やらなくては、というプレッシャー
 - モチベーションが低い人がチームを組んでも急に上がりはしない
 - 逆にうまくやるとモチベーションが上がる？

よいチームを組むために

- ▶ ムードメーカーは大事です
 - コンペ中はうまくいくことよりも、いかないことの方が多い
 - ・ 順位はズルズル下がるもの
- ▶ 作戦も大事です
 - 主にスケジュール
 - ・ ぶれないことはもっと大事
- ▶ チームマネジメント
 - 皆で同じことをやっても意味がない
 - ・ 自分の強みをチームに活かす、活かす方法を見つける
 - 「できる人がやれることをやる」
 - ・ 学生時代にネトゲで学んだこと