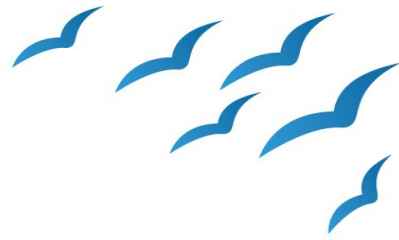# How To Win Any
Data Science Competition

by Kaggle Grandmaster
Pavel Pleskov

Bio

Education

- MSU (2010)
- NES (2012)

Work

- Financial Consultant (1 year)
- Quantitative Researcher (2 years)
- HFT Fund Partner (2 years)

Competitions Grandmaster

Current Rank	Highest Rank
17	15
of 91,335	

7	10	5

Content

- Why participate: pros and cons
- How to choose a competition
- Winning tricks and hacks
- What to do next
- Q&A

T-2: Why participate in competitions - pros

- Rapid knowledge and experience growth - x10 faster than Coursera
- Practical business tasks
- Portfolio improvement
- Money

T-2: Why participate in competitions - cons

- Hard to sell achievements
- Far from real work
- Takes a lot of time
- Harsh competition
- It's addictive
- Cheating

T-1: How to choose a competition
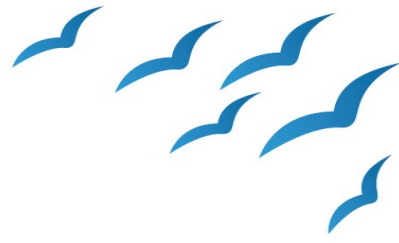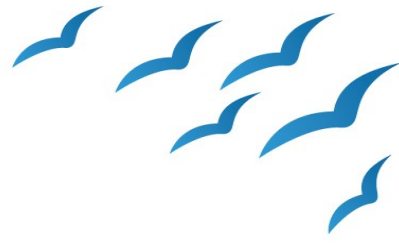
T-1: How to choose a competition

- Check the rules for eligibility (Crimea, Cuba, Iran, Syria, North Korea, Sudan)
- The harder to enter – the better (large data sets, complex registration, etc)
- Fewer participants – less competition (aim for less than 500 people)
- Look at the prize sizes

T-1: How to choose a competition

Red flags:

- Easy registration
- Anonymized features
- Small/no private data set
- Non automated/too many submissions
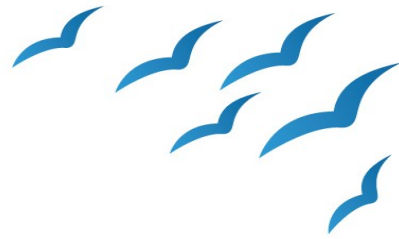- Binary classification with small AUC
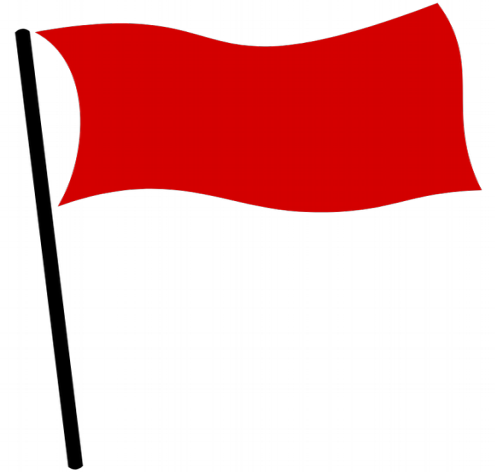
T-0: Winning tricks and hacks

- ods.ai
- TTA
- Folds
- Team work
- Data cleaning
- Looking at errors
- Parameters tuning
- Leaks (meta info for images)
- Stacking (kaz-Anova StackNet, Giba presentation, Stanislav Semenov lectures)

T+1: What to do next

- Write a blog post (fb, twitter, medium, linkedin, habr)
- Make a video (Yandex ML trainings)
- Share your code on github
- Upgrade your computer using prize money
- Choose the next competition

Q&A

THANKS!