Intro

Competitions' concepts

Data

Model

Submission

Evaluation

Leaderboard

Competitions' concepts

Data

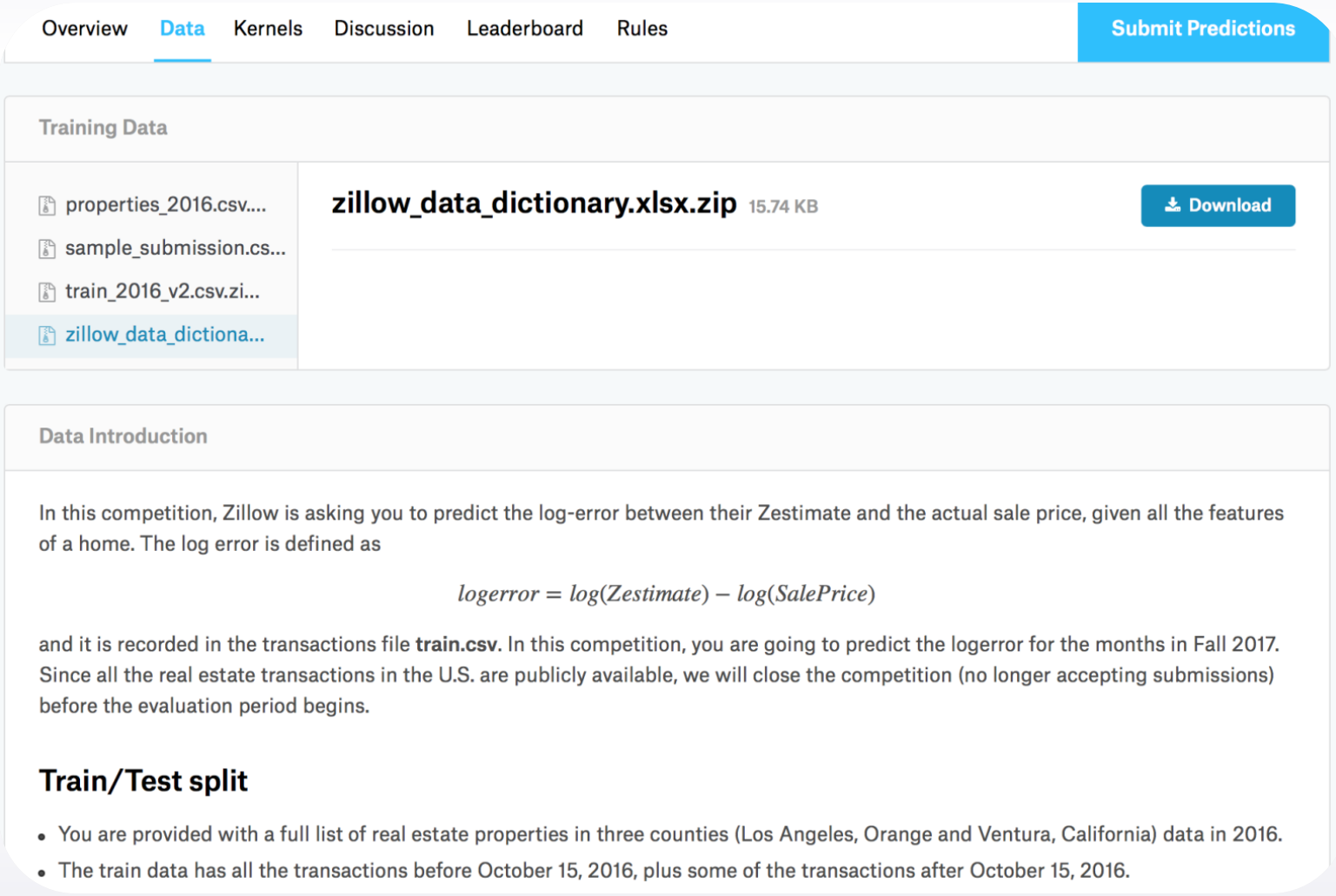
Model

Submission

Evaluation

Leaderboard

Example: data



Competitions' concepts

Data

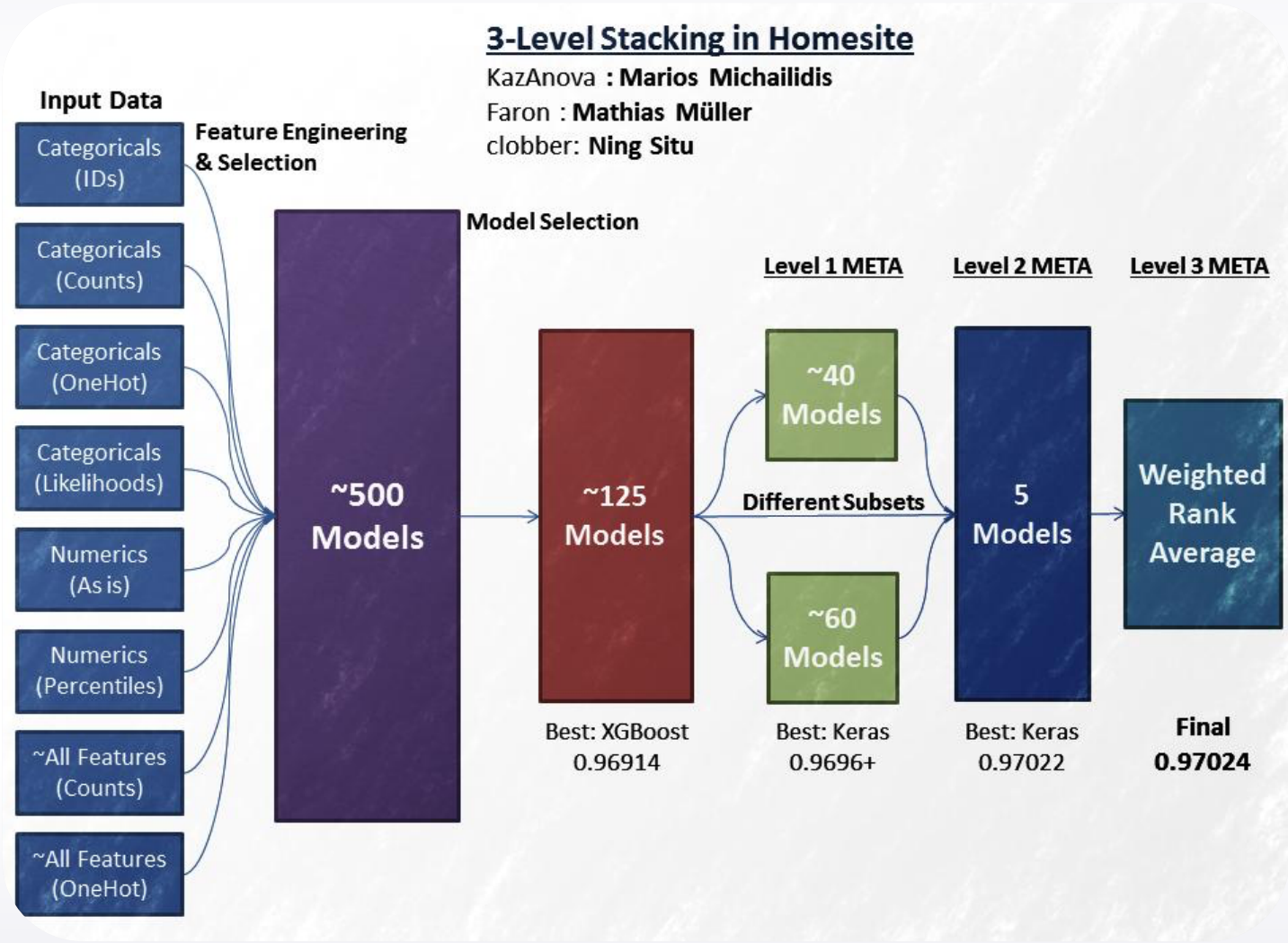
Model

Submission

Evaluation

Leaderboard

Model Example



http://blog.kaggle.com/2016/04/08/homesite-quote-conversion-winners-write-up-1st-place-kazanova-faron-clobber

Competitions' concepts

Data

Model

Submission

Evaluation

Leaderboard

Submission

Usually you are asked to submit only predictions.

Submission

Usually you are asked to submit only predictions.

Sample submission usually looks like:

```
ParcelId,201610,201611,201612,201710,201711,201712
10754147,0.1234,1.2234,-1.3012,1.4012,0.8642-3.1412
10759547,0,0,0,0,0,0
etc.
```

Screenshot kaggle.com

Competitions' concepts

Data

Model

Submission

Evaluation

Leaderboard

Evaluation function

You need to know how good is your model.

The quality of the model is defined by evaluation function:

(predictions, right answers) -> score

Evaluation function

Examples:

- Accuracy
- Logistic loss
- AUC
- RMSE
- MAE

Competitions' concepts

Data

Model

Submission

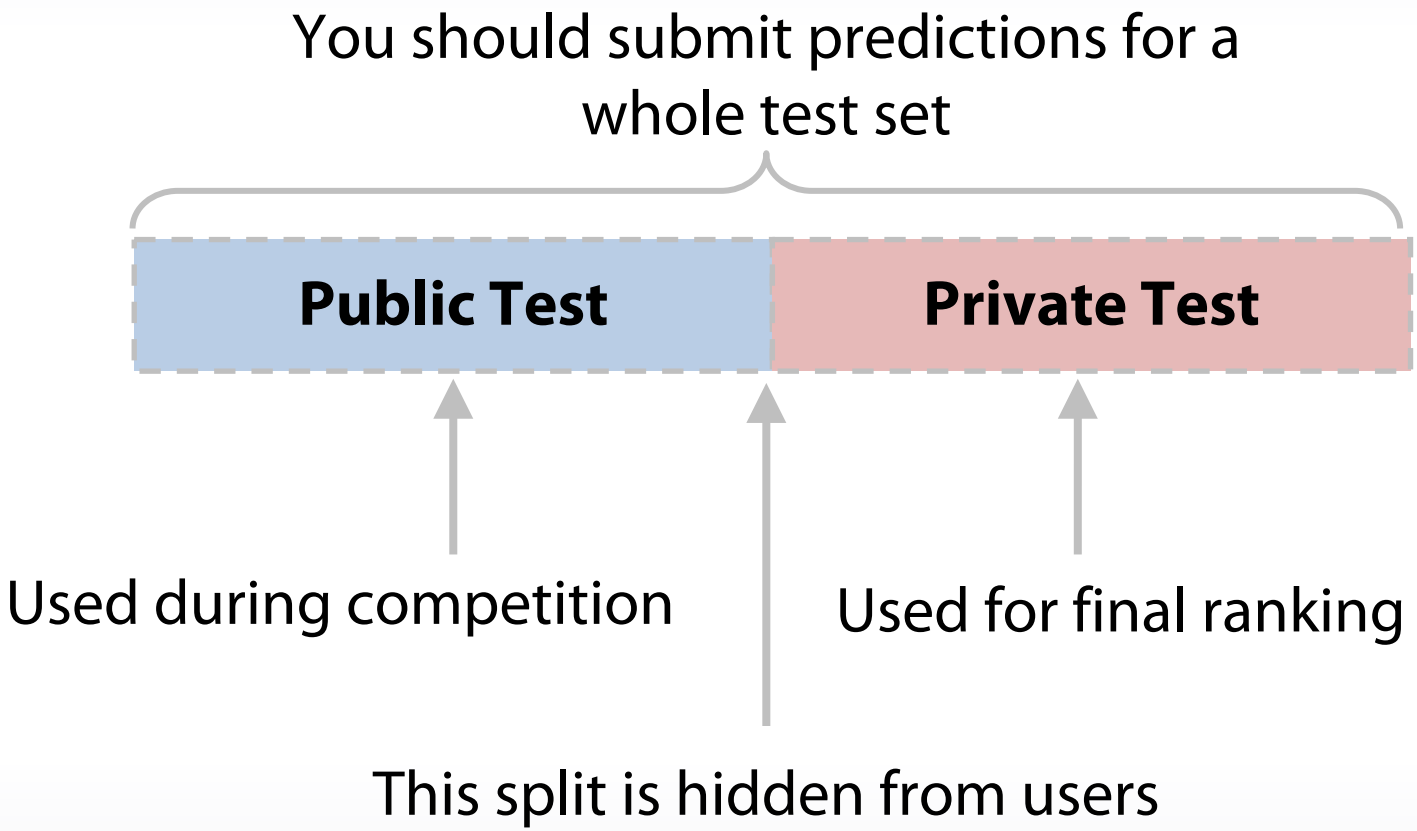
Evaluation

Leaderboard

Leaderboard

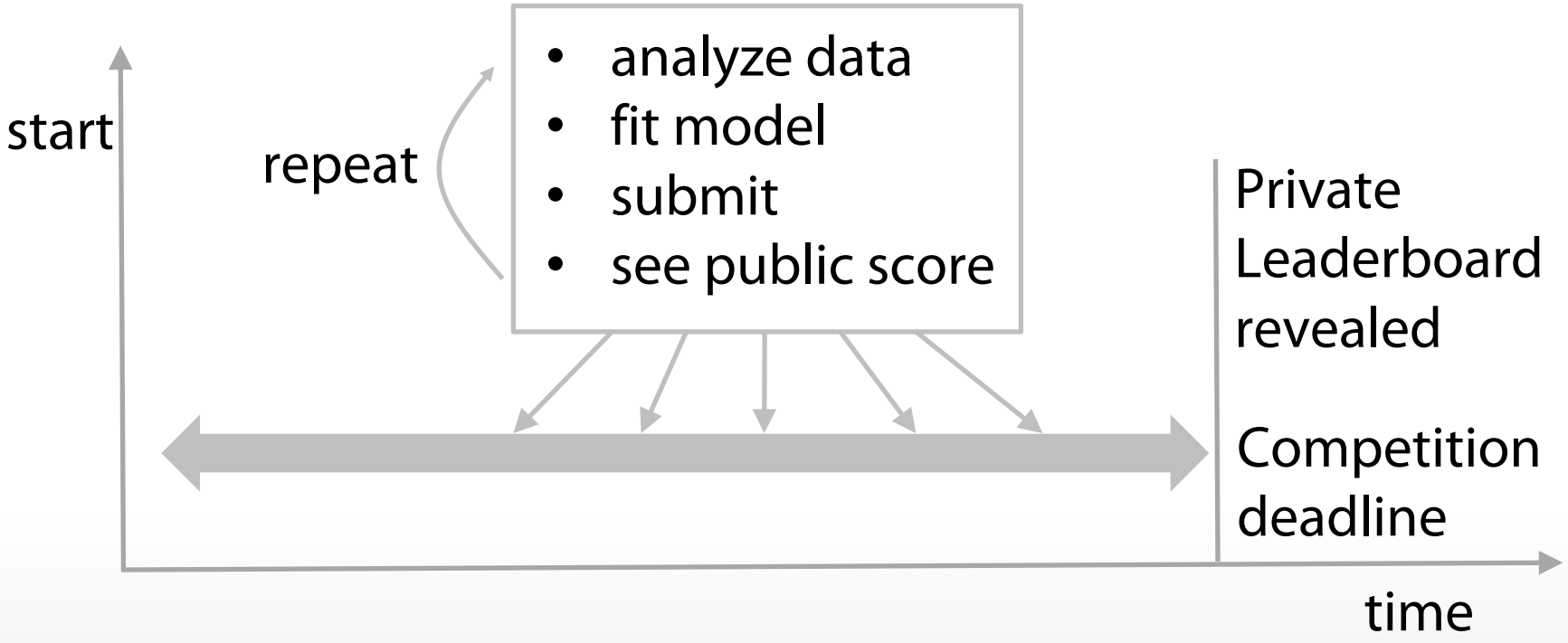
#	△pub	Team Name	Score	Entries	Last
1	▲ 2	Chenglong Chen	0.72189	160	2y
2	▲ 4	Mikhail & Stanislav & Dmitry	0.71871	83	2y
3	▼ 2	Quartet	0.71861	279	2y
4	▲ 1	Shize & Shail & Phil	0.71802	252	2y
5	▲ 8	I love Phở Bò	0.71700	48	2y
6	▼ 2	Gzs_iceberg	0.71681	122	2y
7	▲ 1	YDM	0.71374	283	2y
8	▲ 10	A & A & G	0.71297	229	2y
9	▲ 7	ë	0.71265	96	2y
10	▲ 4	Alexander D'yakonov (PZAD, ...	0.71262	93	2y
11	▼ 9	SearchSearchSearch	0.71022	58	2y
12	▲ 8	woshialex	0.70889	52	2y
13	▲ 43	Alexander Ryzhkov (PZAD, Ru...	0.70777	64	2y
14	▼ 7	Jianmin Sun	0.70711	145	2y
15	▼ 6	I survived Glastonbury (just)	0.70705	119	2y

Public/Private Tests



Example of competition mechanics

Only public leaderboard is available at this stage



Competitions' concepts

Data

Model

Submission

Evaluation

Leaderboard

Platforms:

- Kaggle
- DrivenData
- CrowdAnalityx
- CodaLab
- DataScienceChallenge.net
- Datascience.net
- Single-competition sites (like KDD, VizDooM)

Why to participate?

- Great opportunity for learning and networking

Why to participate?

- Great opportunity for learning and networking
- Interesting non-trivial tasks and state-of-the-art approaches

Why to participate?

- Great opportunity for learning and networking
- Interesting non-trivial tasks and state-of-the-art approaches
- A way to get famous inside data science community

Why to participate?

- Great opportunity for learning and networking
- Interesting non-trivial tasks and state-of-the-art approaches
- A way to get famous inside data science community
- A way to earn some money

Conclusion

- Main concepts:
  - Data
  - Model
  - Submission
  - Evaluation
  - Leaderboard
- Competition platforms
- Reasons for participating