kaggle   Search kaggle   Q   **Competitions**   **Datasets**   **Kernels**   **Discussion**   **Jobs**   Sign In

🏆 Featured Prediction Competition

Zillow Prize: Zillow's Home Value Prediction (Zestimate)

Can you improve the algorithm that changed the world of real estate?

$1,200,000

Prize Money

Zillow · 3,499 teams · 4 months to go (2 days to go until merger deadline)

Overview   Data   Kernels   Discussion   Leaderboard   Rules

Overview

Description

Evaluation

Prizes

Competition Overview

Timeline

Zillow's Zestimate home valuation has shaken up the U.S. real estate industry since first released 11 years ago.

A home is often the largest and most expensive purchase a person makes in his or her lifetime. Ensuring homeowners have a trusted way to monitor this asset is incredibly important. The Zestimate was created to give consumers as much information as possible about homes and the housing market, marking the first time consumers had access to this type of home value information at no cost.

"Zestimates" are estimated home values based on 7.5 million statistical and machine learning models that analyze hundreds of data points on each property. And, by continually improving the median margin of error (from 14% at the onset to 5% today), Zillow has since become established as one of the largest, most trusted marketplaces for real estate information in the U.S. and a leading example of impactful machine learning.

Zillow Prize, a competition with a one million dollar grand prize, is challenging the data science community to help push the accuracy of the Zestimate even further. Winning algorithms stand to impact the home values of 110M homes across the U.S.

In this million-dollar competition, participants will develop an algorithm that makes predictions about the future sale prices of homes. The contest is structured into two
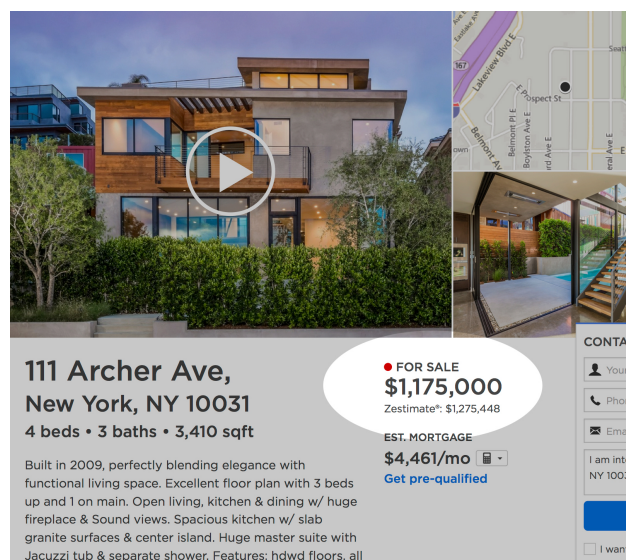
rounds, the qualifying round which opens May 24, 2017 and the private round for the 100 top qualifying teams that opens on Feb 1st, 2018. In the qualifying round, you'll be building a model to improve the Zestimate residual error. In the final round, you'll build a home valuation algorithm from the ground up, using external data sources to help engineer new features that give your model an edge over the competition.

Because real estate transaction data is public information, there will be a three-month sales tracking period after each competition round closes where your predictions will be evaluated against the actual sale prices of the homes. The final leaderboard won't be revealed until the close of the sales tracking period.

Leaderboard

1 Peng S
2 Nima Shahbazi | mchahhou
3 ctlaltdefeat
4 R & D
5 WZS
6 Ivonik
7 Zidmie & Kostadinov & L
8 Hadi Ramezani

Kernels

Keras NN with Q4 validation
10 votes · 10 hours ago

ensemble-stacking（LB~0.6...
11 votes · 11 hours ago

XGBoost using 4th quarter fo...
14 votes · 13 hours ago

Primer for the Zillow - Pred A...
5 votes · a day ago

XGBoost, LightGBM, and OL...
40 votes · a day ago

308 discussion topics

How many rows for submissi...
0 replies · 6 hours ago

...How to construct a convinc...
2 replies · 9 hours ago

Something i'm trying ...
5 replies · 10 hours ago

Is 96273 a valid ZIP code?
2 replies · 11 hours ago

cannot find test data in publi...
1 reply · 15 hours ago

Launch
4 months ago

Close
4 months

2 days
Rules Acceptance
Deadline

3,499 **3,859**
Teams   Competitors

Points This competition awards standard **ranking points**
Tiers   **This competition counts towards tiers**

Tags   housing   real estate   custom metric   small