Two Sigma Connect: Rental Listing Inquiries

How much interest will a new rental listing on RentHop receive?

2,488 teams · 6 months ago

Competition Data

Edit

- Kaggle-renthop.torre...
- images_sample.zip
- sample_submission.cs...
- test.json.zip
- train.json.zip

train.json.zip 19.99 MB

Download

Data Description

In this competition, you will predict how popular an apartment rental listing is based on the listing content like text description, photos, number of bedrooms, price, etc. The data comes from renthop.com, an apartment listing website. These apartments are located in New York City.

The target variable, **interest_level**, is defined by the number of inquiries a listing has in the duration that the listing was live on the site.

File descriptions

- **train.json** - the training set
- **test.json** - the test set
- **sample_submission.csv** - a sample submission file in the correct format
- **images_sample.zip** - listing images organized by listing_id (a sample of 100 listings)
- **Kaggle-renthop.7z** - (optional) listing images organized by listing_id. Total size: 78.5GB compressed. Distributed by BitTorrent (Kaggle-renthop.torrent).

Data fields

- bathrooms: number of bathrooms
- bedrooms: number of bathrooms
- building_id
- created
- description
- display_address
- features: a list of features about this apartment
- latitude
- listing_id
- longitude
- manager_id
- photos: a list of photo links. You are welcome to download the pictures yourselves from renthop's site, but they are the same as imgs.zip.
- price: in USD
- street_address
- interest_level: this is the target variable. It has 3 categories: 'high', 'medium', 'low'