## Evaluation: Submissions are evaluated using the multi-class logarithmic loss.

You must submit a csv file with the listing_id, and a probability for each class.

The order of the rows does not matter. The file must have a header and should look like the following:

```
listing_id,high,medium,low
7065104,0.07743170693194379,0.2300252644876046,0.6925430285804516
7089035,0.0, 1.0, 0.0
...
```

# Data Introduction

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'