HouseSigma

Senior Engineer, Machine Learning

Skills Assessment

How To Complete This Assessment

This is a test to gauge specific skill requirements related to the Senior Engineer, Machine Learning role at HouseSigma. You can copy the questions to a Google doc and share it (open to the internet) or use any document editing software of your choice and send us the PDF.

Problem 1:

HouseSigma is a prop-tech product with a back-end brokerage. We have an amazing group of real estate agents who will service any buyer or seller with real estate needs.

Each agent has a unique skill set and knowledge. What makes an agent stand out is their specialized knowledge of certain regions and communities or house types.

Goal:
We would like to intelligently match any potential home buyer or seller with the real estate agent that has the highest chance of completing a transaction.
(Assume we have no current solutions)

Data:
Property details: We have all the real estate listings, for example:

- Listing id 1000:
    - Location of the listing
    - Home type (detached, town-home, etc..)
    - Listing characteristics (# of bedroom, # of bathroom, etc…)
- We have the listing details data sample downloadable by clicking here (Assume we have millions of listing data)

Agent and Buyer/Seller interaction: We have the agent's historical interactions, at the transaction level. We know whether:

1. Buyer/seller reached out to our agent
2. Went to a property appointment (house showing)

3. Transacted (bought or sold) with that agent.

- We have the agent interaction data sample downloadable by clicking here
- Assume we have:
    - Thousands of transaction data.
    - Hundreds of thousands of appointment data.
    - Tens of millions of “buyer/seller reach out” data

Deliverable:
Note the purpose of this test is to come up with a plan on how you would tackle this problem in a pragmatic and effective way. The goal is to create a solution that can be practically implemented and effectively measured. (Assume we have no system or very simple systems currently).

Feel free to use ML or non-ML methods.

Please send us:

- How you would create the solution, and the experiment to validate the solution offline.
- How you would design an A/B test to validate the solution’s results online.
- How you would monitor the solution’s performance after it’s been implemented.
- *Bonus if you can list out a plan of attack in different launch phases. Going for the easily implementable and low hanging fruits first before moving on to complex solutions.

HouseSigma

Problem 2

Problem Description:

HouseSigma's Sigma estimate home valuation has been one of the more popular features on the HouseSigma app or website.

A home is often the largest and most expensive purchase a person makes in their lifetime. Ensuring homeowners have a trusted way to monitor this asset is incredibly important. The Sigma estimate 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.

Goal:

Use Machine Learning to create your own home price prediction model with the data set below

Data:

We have 1 table downloadable by clicking:

https://drive.google.com/file/d/1L4jMiip74j7DwXyvoR8p-AkiliP2nHKA/view?usp=drive_link

Features & Columns:

Target: price_sold

Id: id_listing

For other features (we have a lot more other features), please try to make your own assumptions.

Deliverable:

Note that the purpose of this test is not about how many hours you are going to spend on it but rather a test on your existing modeling pipeline and thinking process.

HouseSigma

Please send us

- How you would model this problem
- How you interpret the results.
- How you would monitor this in production