

Practical Assessment Objectives

Problem Statement

Airbnb allows its hosts to price their properties. There is a lack of indicators that allow hosts to compare similar listings in their neighbourhood. This has resulted in a suboptimal pricing. As a co-founder of a promising start-up, you intend to solve this problem using state of the art machine learning and natural language processing techniques (optional) to advise Airbnb on how to market their property for the best returns.

Questions of Interest

Some Questions of interest (NOT Exhaustive list). You could add your own questions of interest.

The hosts on Airbnb experiment and charge an optimal price. So, can we recommend an optimal the host should charge for the new listing?

It is uncertain when a listing is hosted or when it becomes unavailable. Are we able to predict whether certain listing is going to be more popular than the others?

Key Deliverables

About 15 pages but not exceeding 30 pages report (excluding annexes) consisting of:

1. Analysis of the datasets (you can use other data sources if appropriate)

- a. Statistical and visualisation

Note: Include your visualisation in the report for key findings and analysis. Do not abuse the annexes.

- b. Comment on the features including feature selection model

2. Key findings and insights

- a. Please use at least 3 different machine learning models (example is regression) covered in class. In addition, you may use other models that you feel suitable.
- b. Important: Please explain your choice of model and the pros and cons of using such a model

3. 2 Recommendations

- a. You can provide recommendations to Airbnb Organization
- b. Evaluate the feasibility of your recommendations
- c. Provide evidence for your recommendations based on the data

Please include your codes in your submission.

Tools

Feel free to use any tools including Python, R and non-coding tools such as Weka, Power BI, Tableau.

Note: Citation is important.

Note: Accuracy is not critical in this assignment

Hint : AirBnB price is not very efficient