

Business objective

Background

Airbnb provides filters and key metrics to see how Airbnb is being used to compete with the residential housing market.

Inventory of Resources

Required personnel:

- Business Unit;
- Project Manager;
- Data Scientist;
- ML Engineer;
- Business Analyst.

Computing resources (In case of GPUs):

- Kaggle Notebooks;
- Google Colab.

Risks and contingencies

During the Data Mining process, the following problems may occur:

- Improper data analysis (outliers are not dealt with, etc.) - Punish Data Scientist and make a proper analysis again;
- Data Leakage during training procedure - accurately check the data preprocessing steps & ensure that there is no contribution of "test" data in the training procedure;
- Lack of computational power - Consider increasing the storage (if it is not enough), optimize the "pipeline" or increase the amount of GPUs used.

Data Mining Goals

`Business Goal` - Increase number of satisfied clients for Airbnb landlords.

`Data Mining Goal` - Predict the Review Scores Rating based on a subset of features that describe the house and living conditions.

Success criteria - The model accurately predicts the Review Scores Rating based on subset of the selected features.
There is linear/non-linear dependency between the features and the target.

Project Plan

The overall project plan contains:

- Data Mining & Preprocessing tools selection - 1 week;
- Data Preprocessing - 2 weeks for different dataset versions;
- Model architecture selection & Review - 1 week;
- Modeling - 2 weeks;
- Evaluation & Model fine-tuning & Review - 2 weeks;
- Final evaluation and results presentation - 1 week;
- Deployment - 1 week.