Milan rent prices forecasting

1 Problem Description

The dataset consists of information from 7334 **rent announcements** in Milan posted in `Immobiliare.it`. For 4500 of these announcements you know the corresponding rent price [denoted as `y`], and you have additional input variables describing several features of the house/apartment. For the other 2834 rent announcements, you have only information on the inputs and not on the rent price.

Your goal is to predict `y` for the held–out 2834 rent announcements.

There are **11 input variables**, which are described below.

1. `square_meters`: dimension of the house/apartment in square meters
2. `contract_type`: type of rental contract
3. `availability`: if the house/apartment is already available, or, if not, when it will be available
4. `description`: description of the rooms in the house/apartment
5. `other_features`: list of additional features of the house/apartment
6. `conditions`: current conditions of the house/apartment
7. `floor`: in which floor of the building the house/apartment is located
8. `elevator`: if an elevator is present or not in the building where the house/apartment is located
9. `energy_efficiency_class`: energy efficiency class of the house/apartment
10. `condominium_fees`: total amount of condominium fees
11. `zone`: area of Milan where the house/apartment is located

Note: the variable `w` refers to weights and hence can be discarded.