

**Computer machine learning**

**COMP4388**

Machine learning

Dr: [Mustafa Jarrar](https://ritaj.birzeit.edu/bzu-msgs/type?mttid=104&classid=183185)

Second Project

* **Prepared by:** 
  + Amir Eleyan-1191076.
* **Date:** Tuesday, May 24, 2022

**Introduction**

The goal of this project is to build a model using machine learning to predict “What is the predominant material used in the construction of the external walls of the dwelling” based on data collected under a survey called “Palestinian Expenditure and Consumption Survey” in 2011. About 30 columns were used as features (inputs to be used to predict the target value).

The machine learning algorithm used is the decision tree, especially C5.0, and this was done using the language R, which is a an open-source programming language that is widely used as a statistical software and data analysis tool.

Features

File name: Dwelling.dta

Input features:

|  |  |
| --- | --- |
| Feature name | Description |
| H1 | Type of housing unit |
| H2 | Tenure of the housing unit |
| H5 | How many rooms are there in dwelling |
| H6 | How many sleeping rooms are used in dwelling |
| H8a | What is estimated rent value each month |
| H8b | Specify type of currency |
| H9a | Connection to Water |
| H9b | Connection to Electricity |
| H9c | Connection to Sewage system |
| H10 | Availability of a kitchen |
| H11 | Availability of a bathroom |
| H12 | Availability of a toilet (WC): |
| H13\_1 | Main source of Cooking |
| H13\_2 | Main source of Heating |
| H13\_3 | Main source of Conditioner |
| H13\_4 | Main source of Oven |
| H13\_5 | Main source of Water heater |
| H14\_1 | Do several or all of house rooms and corridors, and kitchen suffer from the Dampness |
| H14\_2 | Do several or all of house rooms and corridors, and kitchen suffer from the  Cold and difficult heating in winter |
| H14\_3 | Do several or all of house rooms and corridors, and kitchen suffer from the  Poor ventilation |
| H14\_4 | Do several or all of house rooms and corridors, and kitchen suffer from the High heat in summer |
| H14\_5 | Do several or all of house rooms and corridors, and kitchen suffer from the Difficulty heating in winter |
| H18\_1 | Are family members in the dwelling or its surroundings exposed to Smoke, exhaust from cars |
| H18\_2 | Are family members in the dwelling or its surroundings exposed to Smoke, exhaust from industry |
| H18\_3 | Are family members in the dwelling or its surroundings exposed to Odors resulting from animals |
| H18\_4 | Are family members in the dwelling or its surroundings exposed to Odors resulting from sewage system water |
| H18\_5 | Are family members in the dwelling or its surroundings exposed to Odors resulting from |
| H18\_6 | Are family members in the dwelling or its surroundings exposed to General dust |
| H18\_7 | Are family members in the dwelling or its surroundings exposed to Dust or smells resulting from other sources |
| H18\_8 | Are family members in the dwelling or its surroundings exposed to Noise |
| H21\_1 | Availability of a car for the family |
| H21\_5 | Availability of a Cooking stove for the family |
| H21\_6 | Availability of a Dish washer for the family |
| H21\_9 | Availability of a Dehumidifier for the family |
| H21\_16 | Availability of a Computer for the family |
| H21\_20 | Availability of a Filter for the family |

Output feature:

|  |  |
| --- | --- |
| What is the main material used in building outside walls of housing unit | |
| 1 | Cleaned stone |
| 2 | Stone & cement |
| 3 | Old stone |
| 4 | Cement cob |
| 5 | Concrete |
| 6 | Mud |
| 7 | Other (specify) |

Results:

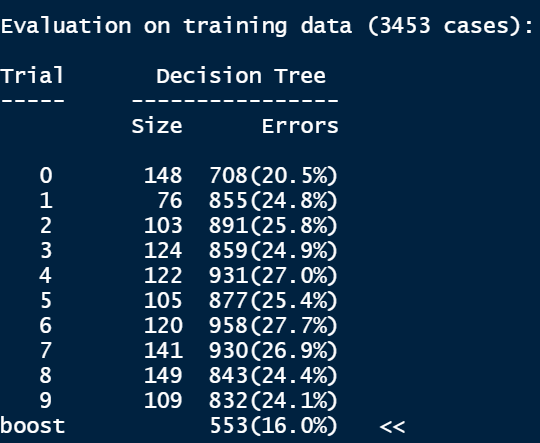


Figure 1: Model after 10 of boosting iterations

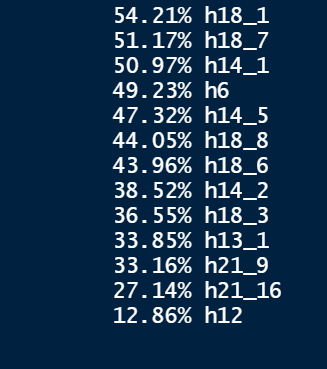
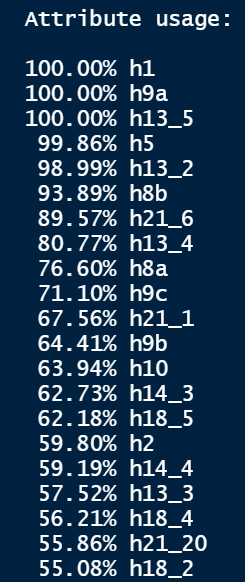


Figure 2: Percentage of each attribute was used, part-1

Figure 3: Percentage of each attribute was used, part-2

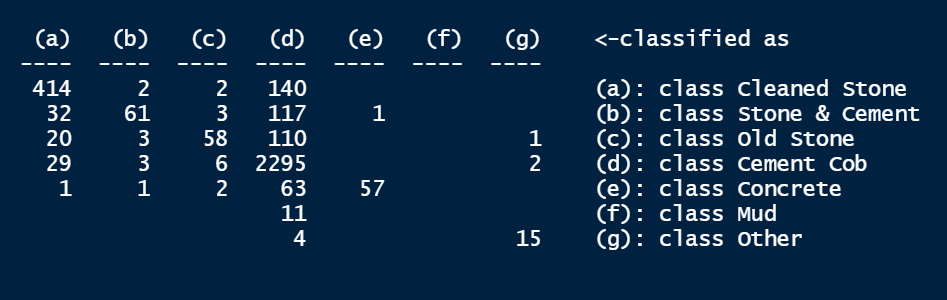


Figure 4: Confusion matrix

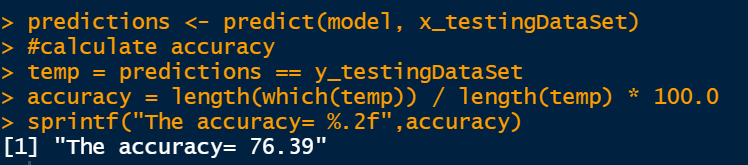


Figure 5: Accuracy of the model in the testing data