Data Science Task

A Requirement for a
Tessella Data Science Position

Tessella
ALTRAN GROUP

Introduction

- As part of recruitment process by Tessella, it is required to solve a data analysis task. The purpose of this task is to demonstrate your skills for the role of a data scientist.
- The task is composed of a single problem, in which you need to analyze some data and build a model.
- The full description of the task is included in this document, the data to be used for the task is attached along with this document in a separate file named *Recruiting_Task_InputData.csv*.

Description of the data

- The data represents the reaction taken by population in a large residential area when they receive direct mailing data advertisements, whether they respond to the sent advertisement or not.
- The data is composed of nine features collected for the population in question, the last column in the data contains the target value reaction (“label”).
- The nine feature columns are the ones labeled: "name", "age", "lifestyle", "zip code", "family status", "car", "sports", "earnings" and “living area."

Data preview

name	age	lifestyle	zip code	family status	car	sports	earnings	living area	label
VnSEFOuL	62	cozily	50168	married	practical	athletics	102526	urban	no response
8Tv0hcce	34	active	66479	married	expensive	soccer	33006	urban	no response
Zny9ysbk	69	healthy	16592	single	expensive	badminton	118760	urban	response
HV3xCamM	57	cozily	50068	married	practical	soccer	131429	urban	response
sflRsQ6v	66	cozily	35988	single	practical	badminton	96003	urban	response
w9voDHj0	21	healthy	60039	single	expensive	badminton	83376	urban	no response
VUkT58V5	22	healthy	69662	married	practical	badminton	49826	urban	no response

Description of the task

1. Build a data driven model that can take the input (defined by the nine features) of a resident and predicts his/her reaction.
2. Is it possible to get an ideal resident (described by the nine input features) who would most likely respond to the advertisement?

Requirements and expectations

Please describe

- approach and methodology used for reaching the goals,
- analysis steps you have done and why.

Your line of reasoning is important to us !

Deliverables

- Documented and runnable source code
  - Please use *Python* or *R* for this task. Suggested platforms for coding are *Jupyter Notebook* and *R-Studio*, or simply plain scripts.
  - The resulting code should be fully runnable and documented.
- Presentation
  - The presentation should be prepared for 30 minutes appointment with the customer who requested this project.
  - The presentation should allow an informed third person (customer) to understand your line of reasoning, your approach, and your results.
  - You are not going to present the presentation to us, so please make it self-explanatory.

We are looking forward to receiving your solution!

If you have any question, please contact your recruiter.

Tessella
ALTRAN GROUP