

Background:**E-Mail Marketing**

You are working for an analytics consultancy company. An online retailer, “Retail Max”, requested you to pitch to win a major contract with them to develop and deploy an e-mail marketing system.

Retail Max aims to use direct e-mail marketing to reach out to their customers and increase sales. Marketing decisions are critical to their success. Currently, their decisions about who to mail the marketing promotions are mostly done through rules of thumb or randomly selecting their customers. However, targeting uninterested customers costs to the company. In the past, some of their e-mail marketing campaigns were not successful. Retail Max would like to implement a direct marketing system to identify their target group for the marketing campaign. With this system, they want to predict which customers will visit the shop as a result of a direct e-mail campaign. This way, they can target the right customers. Several other consultancy companies have been approached as well, and the final decision on who will get the contract will largely depend on the outcome of a demonstration of the approach to this problem based on a dataset that Retail Max provided.

The dataset¹ contains 64,000 customer records collected after an email marketing campaign.

Instructions:

Conduct data mining on the provided data to identify a suitable approach for predicting customers who will visit the shop as a result of an e-mail marketing campaign. Outline your approach considering the steps in CRISP-DM methodology and justify the selected methods for data preparation, modelling and evaluation steps used in this project.

The focus of this assignment is to give you a go at thinking about the business problem, data preparation for modelling, implementing and interpreting several models using real data, describing your approach to the data mining project clearly and providing clear justification of the steps followed and the methods used.