

CMT307 Coursework 1 Data Description

File name: CW1_data_202223.csv

A health insurance provider is looking to expand into car insurance. In the given dataset, you may find data relating to a campaign involving existing health customers where they have been inquired regarding their interest in purchasing car insurance. Your task will be to create machine learning models to predict **Result** whether a customer will buy car insurance based on these variables. *Result*=1 indicates interested, and *Result*=0 indicates not interested.

The dataset has the following column names, which are largely self-explanation, but no further information is available as the dataset contains all the information that can be shared with third-party data scientists.

id	Gender	Age	HasDrivingLicense	RegionID	Switch	VehicleAge	PastAccident	AnnualPremium	SalesChannelID	DaysSinceCreated	Result
----	--------	-----	-------------------	----------	--------	------------	--------------	---------------	----------------	------------------	--------

Based on the information from the column names and associated data values, the dataset contains categorical and numerical features as follows.

Categorical features

- Gender
- HasDrivingLicense
- RegionID
- Switch
- VehicleAge
- PastAccident
- SalesChannelID

Numerical features

- Age
- AnnualPremium
- DaysSinceCreated

Please refer to the marking criteria in the coursework proforma when completing this coursework.