

Car Insurance Claim Prediction

Project 2

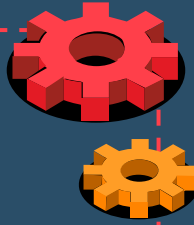
José Cunha – up201905451

Mariana Teixeira – up201905705

Raquel Carneiro - up202005330



Specification of the Problem



The Dataset contains information on policyholders having the attributes like policy tenure, age of the car, age of the car owner, the population density of the city, make and model of the car, power, engine type, etc, and the target variable indicating whether the policyholder files a claim in the next 6 months or not.

Goal:

train an AI to **predict** whether the policyholder will file a claim in the next 6 months or not, based on information on policyholders.

Related Work:

- <https://www.kaggle.com/datasets/ifteshanajni/carinsuranceclaimprediction-classification?select=train.csv>





Tools and algorithms



- Linear Regression
Commonly used to predict insurance based on features.
- Random Forests
Combines multiple decision trees to improve accuracy.
- Gradient Boosting
Combines multiple weak models to create a strong model.
- Libraries:
 - Jupiter Notebook
 - Pandas library
 - Mathplotlib library



Work Carried Out

- Pre-process the data.
- Check for missing values or outliers.

