INSURANCE CLAIMS

Problem Statement

Insurance is a critical component of financial risk management, offering protection against unexpected losses from accidents, damage, or liability.

However, accurately predicting which policyholders are likely to file a claim remains a major challenge for insurers.

This case involves an insurance company aiming to improve its **risk assessment** by predicting the likelihood of a customer making a claim based on vehicle, customer, and regional information. By identifying high-risk customers in advance, the company can optimize its underwriting strategies, reduce financial losses, and design personalized policy offerings.

Currently, only a small percentage of customers file claims, making the prediction problem highly imbalanced. Therefore, advanced **Python-based Machine Learning techniques** will be applied to build a reliable prediction model that supports data-driven decision-making and improves operational efficiency.

Objective

By analyzing key features such as **customer age, vehicle age, subscription length, vehicle specifications, safety features, fuel and engine type, and regional density**,
the goal is to **build various machine learning models capable of predicting the likelihood of insurance claims**.

This predictive model aims to **enhance accuracy and efficiency** in identifying high-risk policyholders, enabling insurance companies to **minimize financial losses, tailor premium pricing**, and optimize risk management strategies ultimately benefiting both the **insurer and the customer**.

Column Name	Description
customer_age	Age of the policyholder
vehicle_age	Age of the insured vehicle
subscription_length	Duration of the insurance subscription
region_density	Population density of the policyholder's region
airbags	Number of airbags in the vehicle
displacement	Engine displacement (cc) of the vehicle
cylinder	Number of cylinders in the vehicle's engine
turning_radius	Turning radius of the vehicle (meters)
length	Length of the vehicle (mm)
width	Width of the vehicle (mm)
gross_weight	Gross weight of the vehicle (kg)
ncap_rating	NCAP safety rating (0–5) of the vehicle
fuel_type	Type of fuel used (encoded)
engine_type	Type of engine (encoded)
rear_brakes_type	Type of rear braking system (encoded)
transmission_type	Type of vehicle transmission (encoded)
steering_type	Steering type (encoded)
is_esc	Electronic Stability Control (1 = Yes, 0 = No)
is_adjustable_steering	Adjustable steering available (1 = Yes, 0 = No)
is_tpms	Tyre Pressure Monitoring System present (1 = Yes, 0 = No)
is_parking_sensors	Parking sensors available (1 = Yes, 0 = No)
is_parking_camera	Parking camera available (1 = Yes, 0 = No)
is_power_steering	Power steering available (1 = Yes, 0 = No)
is_front_fog_lights	Front fog lights present (1 = Yes, 0 = No)
is_rear_window_wiper	Rear window wiper present (1 = Yes, 0 = No)

Column Name	Description
is_rear_window_washer	Rear window washer present (1 = Yes, 0 = No)
is_rear_window_defogger	Rear window defogger present (1 = Yes, 0 = No)
is_brake_assist	Brake assist system present (1 = Yes, 0 = No)
is_power_door_locks	Power door locks present (1 = Yes, 0 = No)
is_central_locking	Central locking available (1 = Yes, 0 = No)
is_child_safety_lock	Child safety lock present (1 = Yes, 0 = No)
is_day_night_rear_view_mirror	Day-night rearview mirror present (1 = Yes, 0 = No)
is_ecw	Emergency Call Warning system present (1 = Yes, 0 = No)
is_speed_alert	Speed alert system present (1 = Yes, 0 = No)
segment	Segment or category of the vehicle (encoded)
claim_status	Target variable: 1 = Claim made, 0 = No claim