

Insurance companies invest a lot of **time and money** [↗](#) into optimizing their pricing and accurately estimating the likelihood that customers will make a claim. In many countries insurance it is a legal requirement to have car insurance in order to drive a vehicle on public roads, so the market is very large!

Knowing all of this, On the Road car insurance have requested your services in building a model to predict whether a customer will make a claim on their insurance during the policy period. As they have very little expertise and infrastructure for deploying and monitoring machine learning models, they've asked you to identify the single feature that results in the best performing model, as measured by accuracy, so they can start with a simple model in production.

They have supplied you with their customer data as a csv file called `car_insurance.csv`, along with a table detailing the column names and descriptions below.

## The dataset

Column	Description
<code>id</code>	Unique client identifier
<code>age</code>	Client's age: <ul style="list-style-type: none"><li><code>0</code> : 16-25</li><li><code>1</code> : 26-39</li><li><code>2</code> : 40-64</li><li><code>3</code> : 65+</li></ul>
<code>gender</code>	Client's gender: <ul style="list-style-type: none"><li><code>0</code> : Female</li><li><code>1</code> : Male</li></ul>
<code>driving_experience</code>	Years the client has been driving: <ul style="list-style-type: none"><li><code>0</code> : 0-9</li><li><code>1</code> : 10-19</li><li><code>2</code> : 20-29</li><li><code>3</code> : 30+</li></ul>
<code>education</code>	Client's level of education: <ul style="list-style-type: none"><li><code>0</code> : No education</li><li><code>1</code> : High school</li><li><code>2</code> : University</li></ul>
<code>income</code>	Client's income level: <ul style="list-style-type: none"><li><code>0</code> : Poverty</li><li><code>1</code> : Working class</li><li><code>2</code> : Middle class</li><li><code>3</code> : Upper class</li></ul>
<code>credit_score</code>	Client's credit score (between zero and one)

Column	Description
vehicle_ownership	Client's vehicle ownership status: <ul style="list-style-type: none"> <li>0 : Does not own their vehilce (paying off finance)</li> <li>1 : Owns their vehicle</li> </ul>
vehcile_year	Year of vehicle registration: <ul style="list-style-type: none"> <li>0 : Before 2015</li> <li>1 : 2015 or later</li> </ul>
married	Client's marital status: <ul style="list-style-type: none"> <li>0 : Not married</li> <li>1 : Married</li> </ul>
children	Client's number of children
postal_code	Client's postal code
annual_mileage	Number of miles driven by the client each year
vehicle_type	Type of car: <ul style="list-style-type: none"> <li>0 : Sedan</li> <li>1 : Sports car</li> </ul>
speeding_violations	Total number of speeding violations received by the client
duis	Number of times the client has been caught driving under the influence of alcohol
past_accidents	Total number of previous accidents the client has been involved in
outcome	Whether the client made a claim on their car insurance (response variable): <ul style="list-style-type: none"> <li>0 : No claim</li> <li>1 : Made a claim</li> </ul>

index	...	↑↓	best_feature	...	↑↓	best_accuracy	...
		0	driving_experience				0.77