

Project Proposal Template

Question/need:

- The purpose of this model is to predict whether or not someone will buy car insurance from a company that also sells health insurance to them
- The insurance provider may benefit from this model because it will allow them to estimate how many of their current clients may be interested in getting car insurance from them as well.

Data Description:

- I will be downloading a publicly available stroke prediction dataset off of kaggle.

Link:

<https://www.kaggle.com/anmolkumar/health-insurance-cross-sell-prediction?select=train.csv>

- The target for this model is whether or not the person (each row) someone is interested in buying car insurance(response), while features include: gender, age, whether they have a driver's license, region code, whether or not they've been previously insured, vehicle age, vehicle damage, annual premium, policy sales channel, whether a car is vintage or not.
- There are a total of 381109 rows and 10 columns.

Tools:

- In order to complete this project I plan on using the following tools: pandas, numpy, seaborn, matplotlib, scikit learn, and tableau

MVP Goal:

- An MVP for this project would be a baseline classification model such as logistic regression with a few features and evaluation metrics such as confusion matrix and f1 score.