

Health Insurance Cross Sell Prediction





Tim Data Scientist yang bekerja di perusahaan Asuransi
PT. Paradential Life Assurance

Our Team



Herald



Hansen



Bunga



Putri



Arif



Arsyadana



Erdin

Agenda

Business Understanding

Exploratory Data Analysis

Data Preprocessing

Modeling

Business Recommendation

Business Understanding



Problem Statement

PT Paradential Life Assurance ingin menawarkan produk asuransi kendaraan mereka kepada pelanggan mereka yang sebelumnya sudah mempunyai asuransi kesehatan.



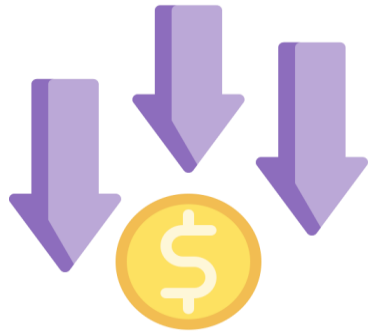
Memerlukan biaya ***marketing campaign*** sebanyak **\$800.000** apabila ingin menawarkan kepada seluruh pelanggan



Response rate atau ketertarikan pelanggan yang rendah terhadap penawaran yang perusahaan lakukan, jumlahnya hanya **12%** dari total seluruh pelanggan



Goal



Menurunkan biaya ***marketing campaign***
sebesar **30 %**



Meningkatkan **Response Rate**
sebanyak **20 %**

Objective



Mendesain **model Machine Learning** agar bisa **memprediksi** apakah pelanggan tertarik pada asuransi kendaraan yang perusahaan tawarkan atau tidak



Memberikan **solusi** dan **rekomendasi bisnis** berdasarkan insight yang kita dapat

Business Metrics



Campaign Cost



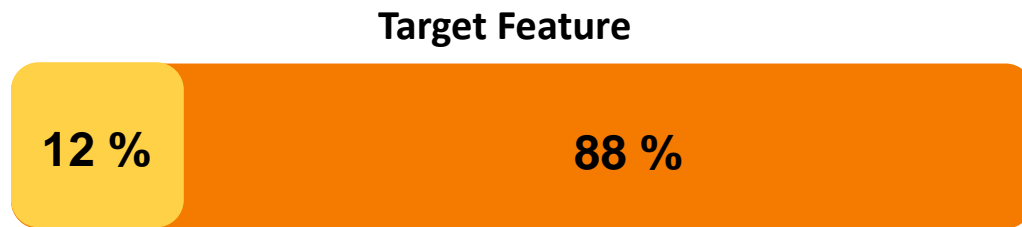
Response Rate



Exploratory Data Analysis



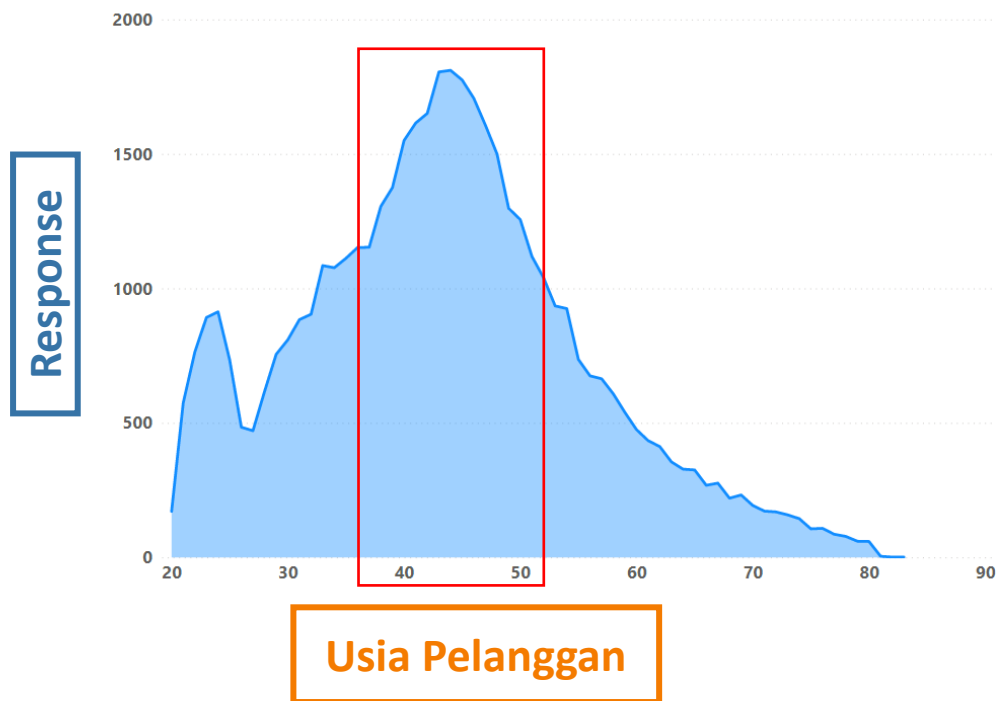
Data Overview

- 1 Dataset terdiri dari **381109 baris** dan **12 kolom**
- 2 **9** kolom numerik dan **3** kolom kategorik
- 3 Tidak ada **missing values**
- 4 Tidak ada **data duplikat**



-  Pelanggan **Tidak tertarik**
-  Pelanggan **Tertarik**

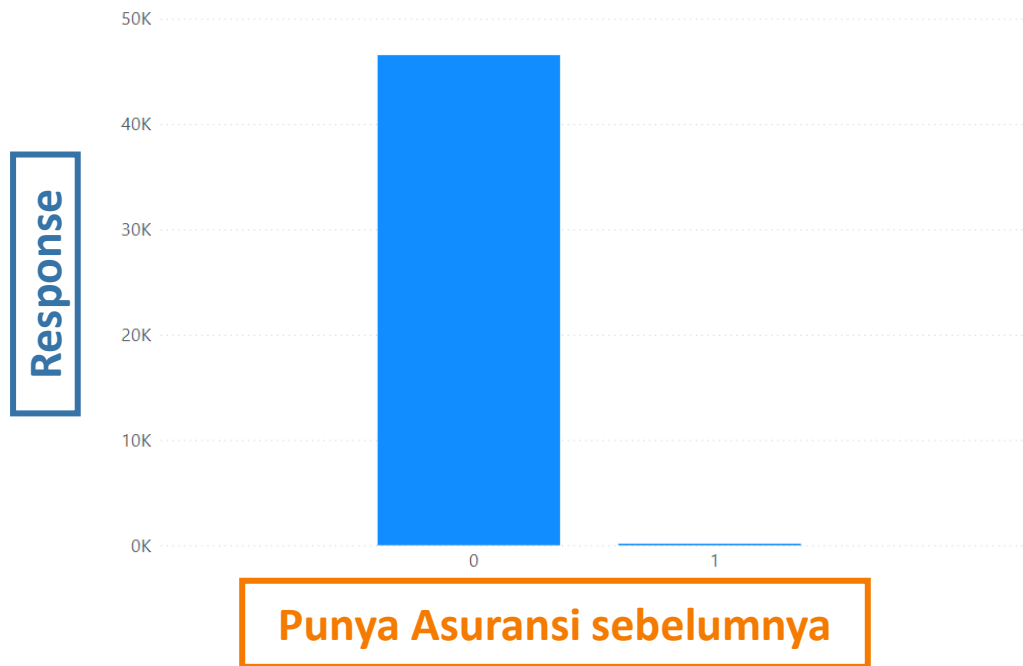
Data Insight



Pengaruh usia pelanggan terhadap Response

Pelanggan yang tertarik di dominasi oleh pelanggan dengan rentang usia antara **40 – 50 tahun**

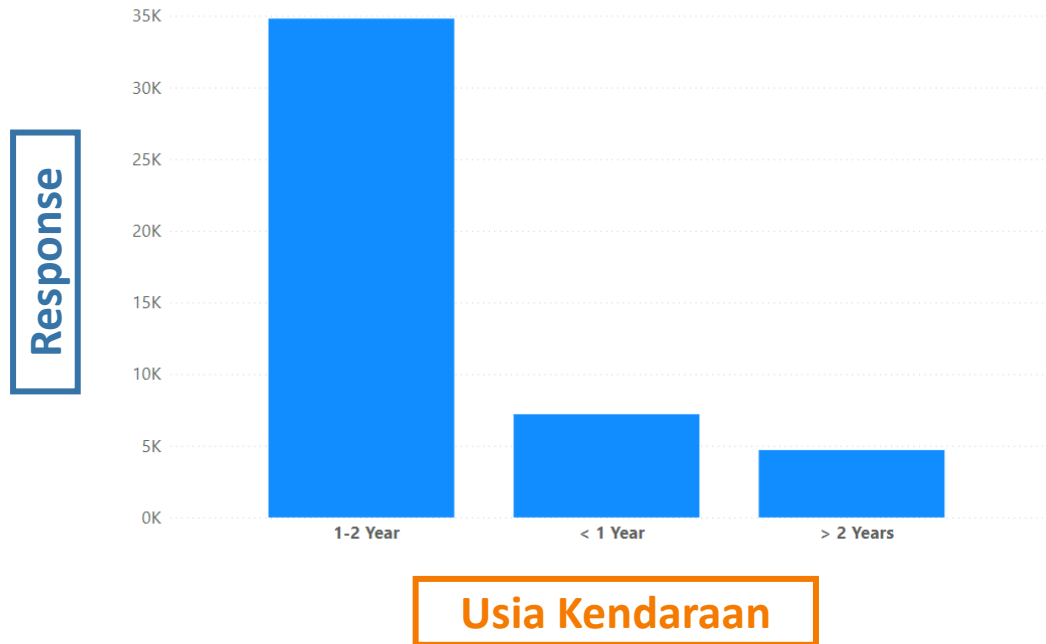
Data Insight



Pengaruh punya asuransi sebelumnya terhadap Response

Pelanggan yang memang sebelumnya **belum pernah** memiliki asuransi kendaraan sangat tertarik terhadap penawaran kita, sebanyak **99,6%** dari total keseluruhan pelanggan yang tertarik

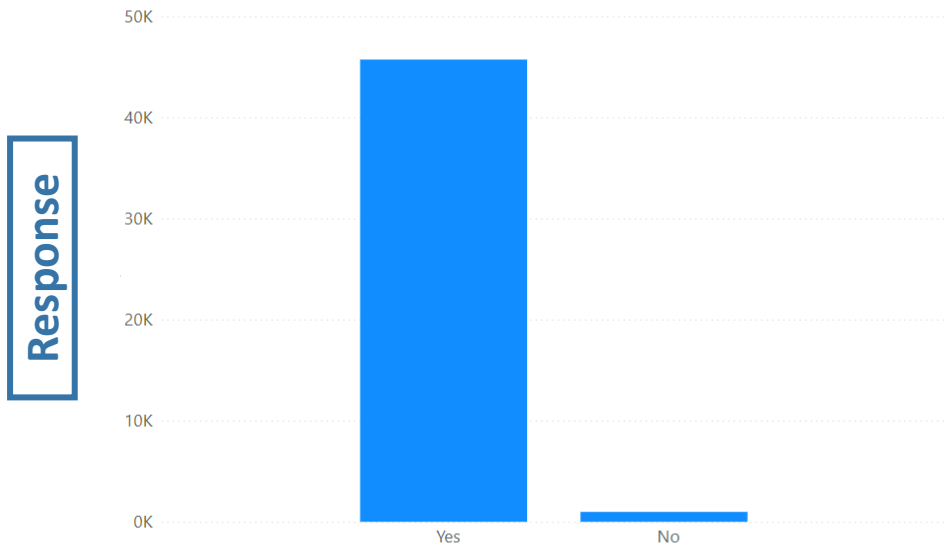
Data Insight



Pengaruh Usia Kendaraan terhadap Response

Pelanggan dengan usia kendaraan **1 – 2 tahun** banyak yang tertarik dengan penawaran yang kita berikan, sebanyak **74,5 %** dari total keseluruhan pelanggan yang tertarik

Data Insight



Kendaraan pernah kecelakaan

Pengaruh Kendaraan pernah kecelakaan terhadap Response

Pelanggan yang kendaraannya pernah mengalami kecelakaan sangat tertarik, dengan kontribusi **97,9 %** dari total keseluruhan pelanggan yang tertarik

Data Preprocessing



Data Preprocessing



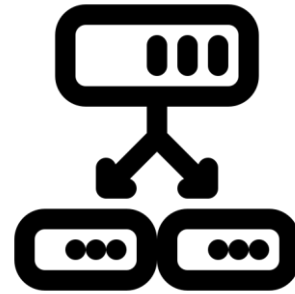
Feature Encoding

- Label Encoding
- One Hot Encoding



Feature Selection

- Drop kolom



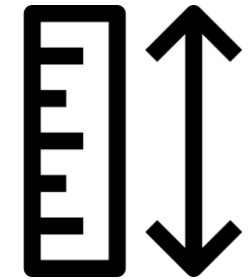
Train Test Split

- Train set 70 %
- Test set 30 %



Handling Imbalance

- Undersampling



Feature Scaling

- Standarization

Modeling



Model Result

Model	Accuracy	F1 Score	ROC-AUC	AUC
RandomForest	0.78	0.80	0.78	0.83
XGBoost	0.78	0.81	0.78	0.84
LightGBM	0.79	0.81	0.79	0.84



Terlihat dari evaluasi model diatas, semua algoritma model hampir memiliki skor yang sama. Kita akan melakukan **Hyperparameter tuning** untuk semua algoritma model tersebut dan melihat mana dari ketiga model diatas yang mendapatkan hasil terbaik

Model Result

Hyperparameters Tuning – Randomized Search CV

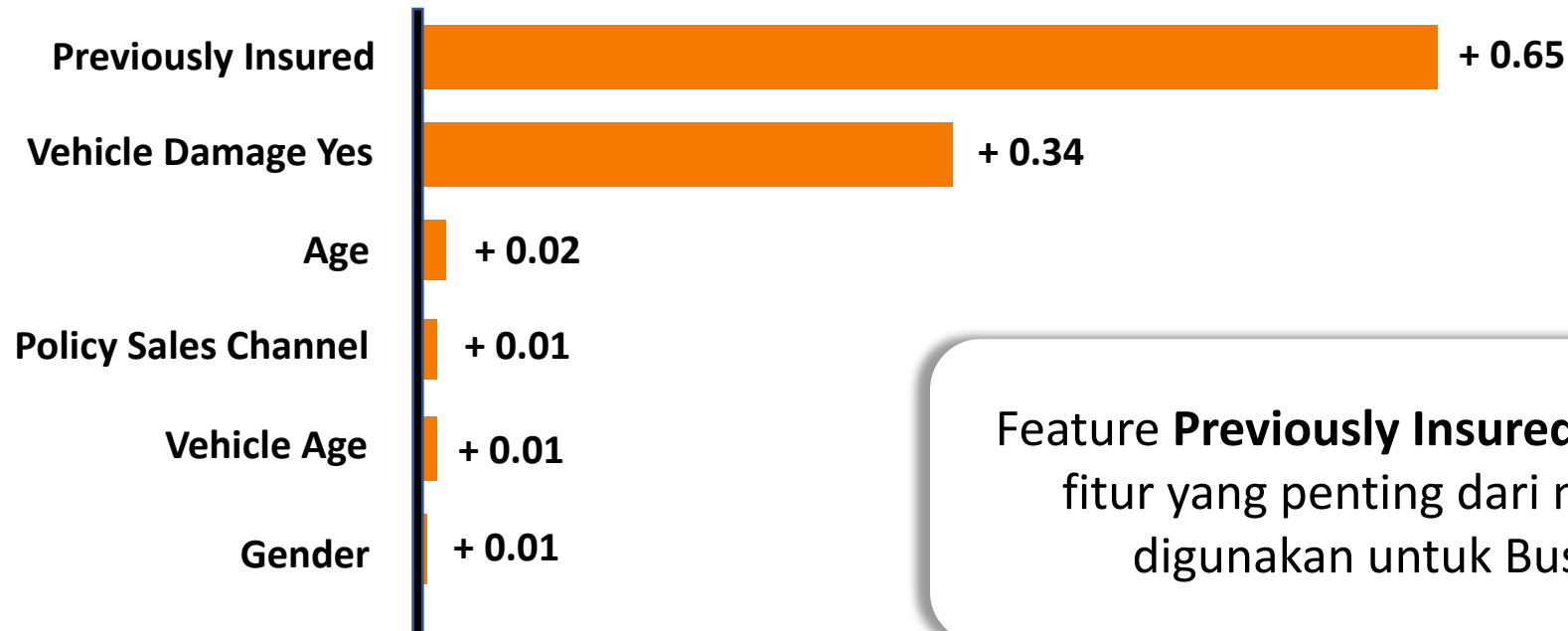
Model	Accuracy	F1 Score	ROC-AUC	AUC
RandomForest	0.79	0.81	0.79	0.84
XGBoost	0.79	0.82	0.79	0.85
LightGBM	0.79	0.81	0.79	0.84

Setelah melakukan **hyperparameter tuning**, model dengan skor yang paling bagus adalah **XGBoost Classifier** dengan nilai **F1 Score 82 %**

Business Recommendation



Feature Importance



Feature **Previously Insured** dan **Vehicle Damage** menjadi fitur yang penting dari model XGBoost ini dan akan digunakan untuk Business Recommendation

Business Problem



Perusahaan ingin
menawarkan produk
asuransi kendaraan mereka



Pelanggan

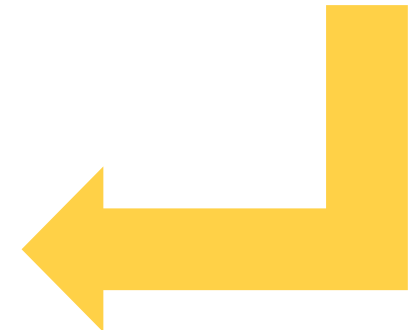


Perusahaan tidak tau
apakah pelanggan
tertarik atau tidak



Marketing
Campaign

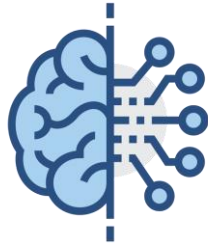
- Biaya Marketing Campaign **tinggi**
- Response Rate hanya **12%**
- ROI yang tidak maksimal



Business Simulation



Perusahaan ingin
menawarkan produk
asuransi kendaraan mereka



Machine
Learning Model

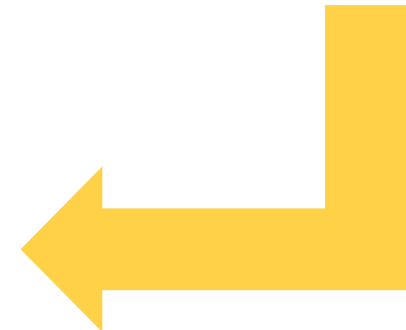


Pelanggan
Tertarik



Marketing
Campaign

- Biaya marketing campaign *lebih efektif*
- Response Rate naik **50 %**
- ROI yang maksimal



Business Simulation – Before Model



- **Campaign Cost :**
\$ 800.000 (381.109 Pelanggan)
- **Response Rate:**
12,26 % (46.723 Pelanggan tertarik)
- **Return of Investment:**
\$ 135.963

Campaign Cost / Customer : \$ 2.09
Return / Customer : \$ 5
ROI : \$ 2.91

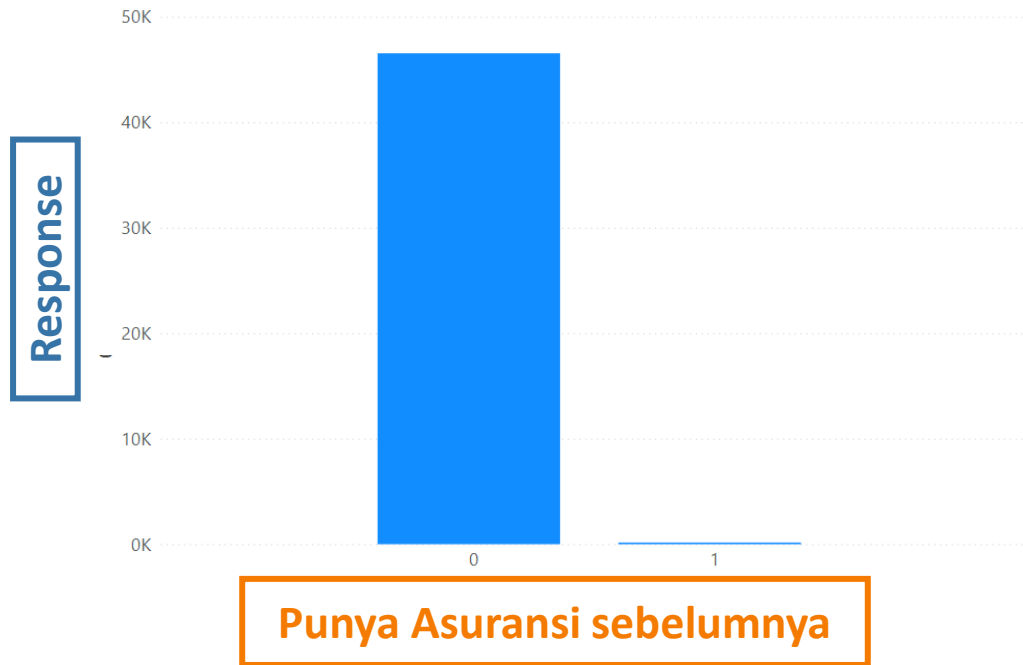
Business Simulation – After Model



- **Campaign Cost :**
\$ 627.000 (300.000 Pelanggan)
- **Response Rate:**
50 % (150.000 Pelanggan tertarik)
- **Return of Investment:**
\$ 436.500

Campaign Cost / Customer : \$ 2.09
Return / Customer : \$ 5
ROI : \$ 2.91

Business Recommendation



Feature Previously Insured

Memfokuskan marketing campaign atau penawaran kepada pelanggan yang sudah memiliki asuransi kendaraan sebelumnya dengan memberikan keunggulan asuransi kendaraan yang kita punya yaitu prosedur **klaim asuransi** yang **mudah** dan **cepat** tidak sampai 1 hari langsung cair.

Business Recommendation



Feature Vehicle Damage

Memberikan **campaign** kepada pelanggan yang kendaraanya belum pernah mengalami kerusakan tentang pentingnya asuransi kendaraan dan besarnya biaya perbaikan kendaraan apabila terjadi kecelakaan. Dan juga memberikan promo menarik yaitu akan cover seluruh biaya perbaikan kendaraan apabila terjadi kerusakan **tanpa syarat apapun.**

Thank You!
