



# PREDICT CUSTOMER WITH PAYMENT DIFFICULTIES

Faiz Naida Salimah

# 1. Latar Belakang

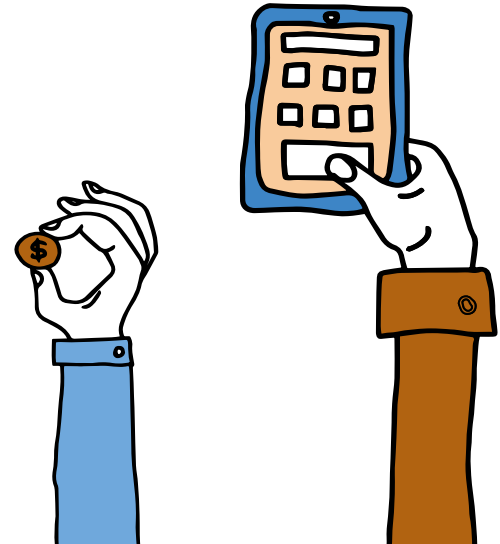
Home Credit Indonesia merupakan perusahaan yang memberikan layanan kredit yang mudah, cepat, dan terjangkau kepada masyarakat. Salah satu permasalahan yang terjadi pada perusahaan ini adalah ***adanya nasabah yang mengalami kesulitan dalam pembayaran kredit***. Apabila banyak pelanggan yang mengalami masalah ini, maka akan berdampak secara signifikan bagi perusahaan.

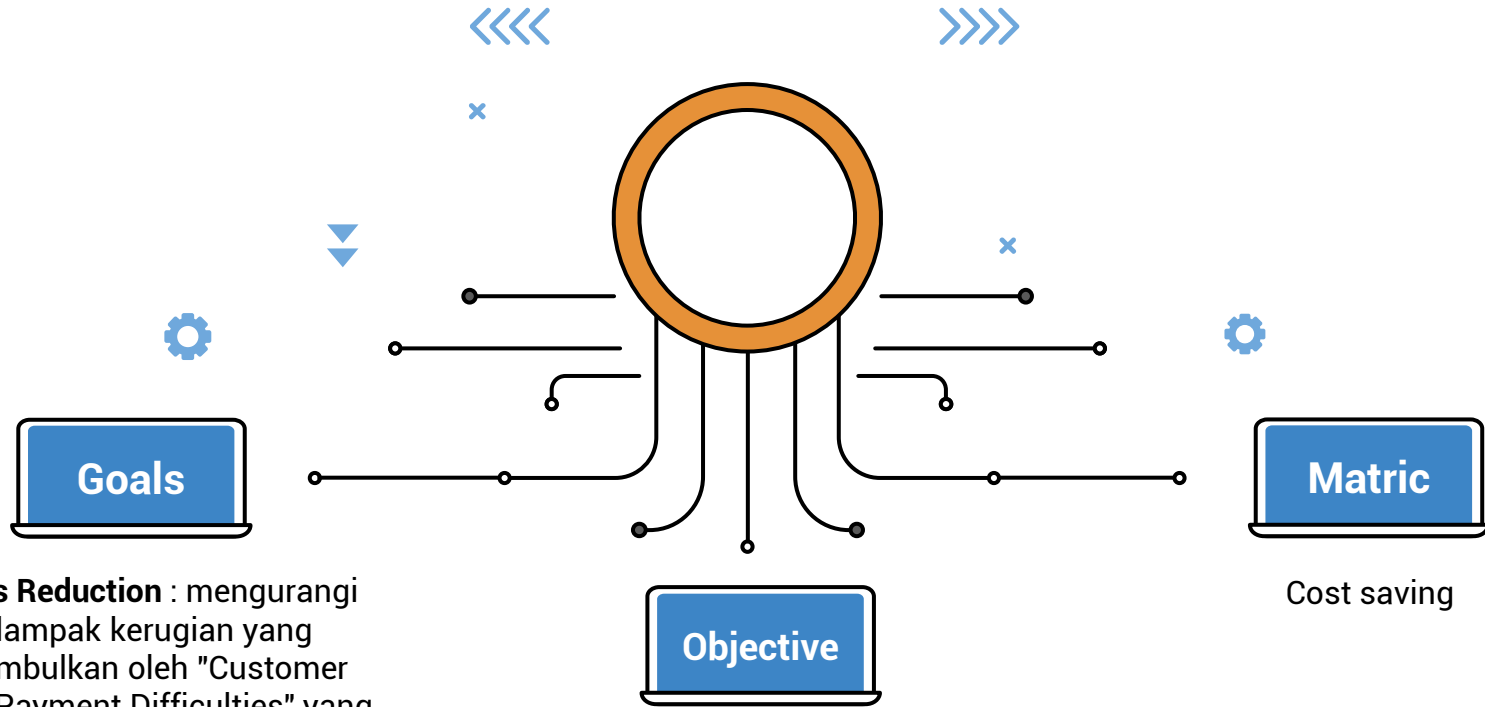
Menurut artikel yang diterbitkan di Harvard Business Review, dinyatakan bahwa ***"Non-payment by consumers can set off a chain reaction of bad debts, lower profits, layoffs, and even bankruptcies, ultimately affecting entire industries and even economies."***

Oleh karena itu, mengidentifikasi nasabah yang memiliki kemungkinan besar mengalami kesulitan dalam membayar kredit penting untuk dilakukan. Hal ini dapat menjadi tindakan preventif bagi perusahaan dan memastikan pelanggan yang mampu melakukan pelunasan tidak ditolak ketika melakukan pengajuan pinjaman.

# 8,07%

nasabah dalam data adalah  
**"Customer with  
Payment Difficulties"**



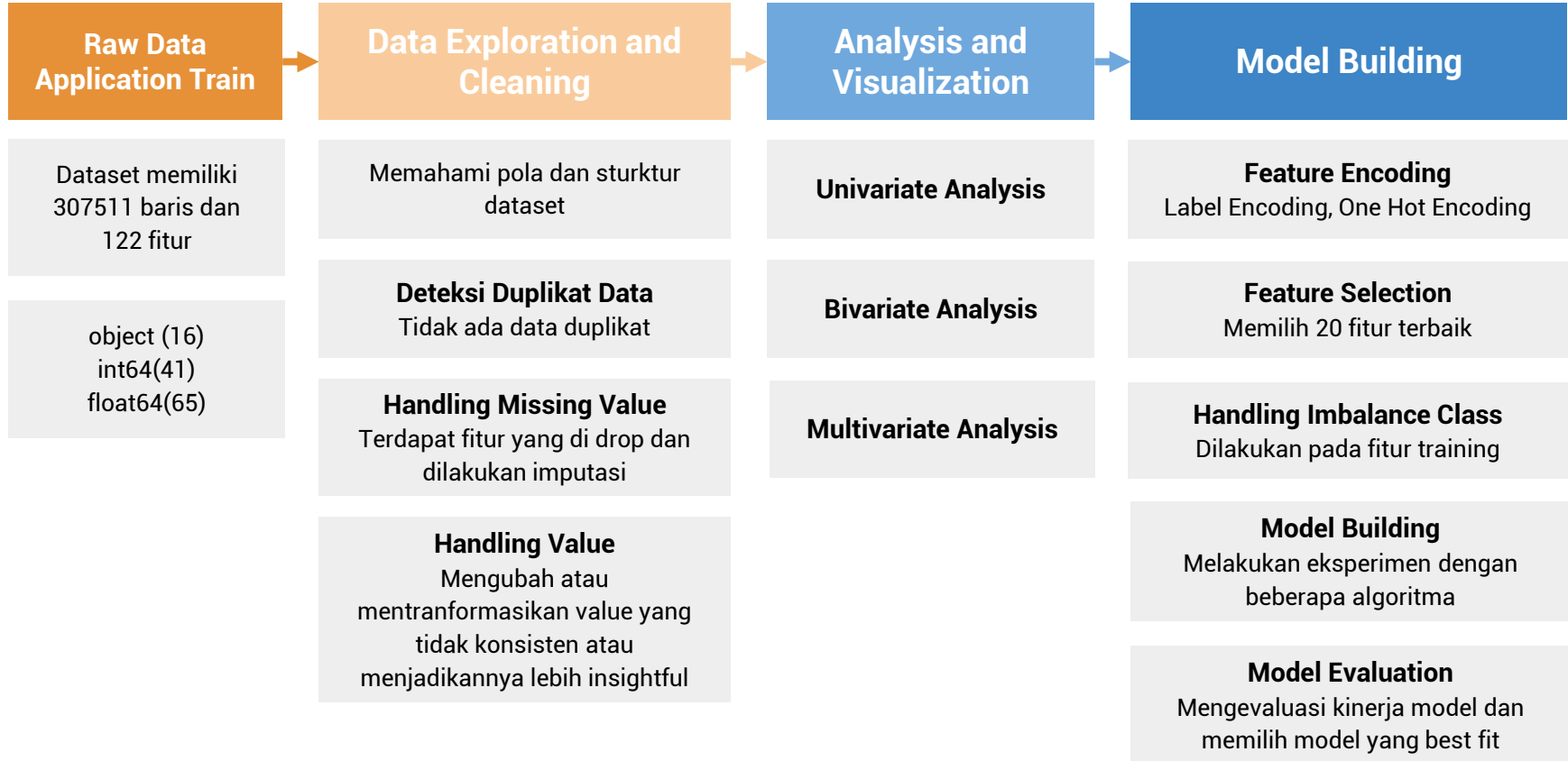


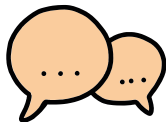
**Loss Reduction** : mengurangi dampak kerugian yang ditimbulkan oleh "Customer with Payment Difficulties" yang memiliki potensi gagal bayar.

Cost saving

- Membuat prediktif model untuk memprediksi dan mengklasifikasikan nasabah yang mampu membayar atau tidak
- Mengidentifikasi karakteristik nasabah yang berpotensi kesulitan dalam pembayaran

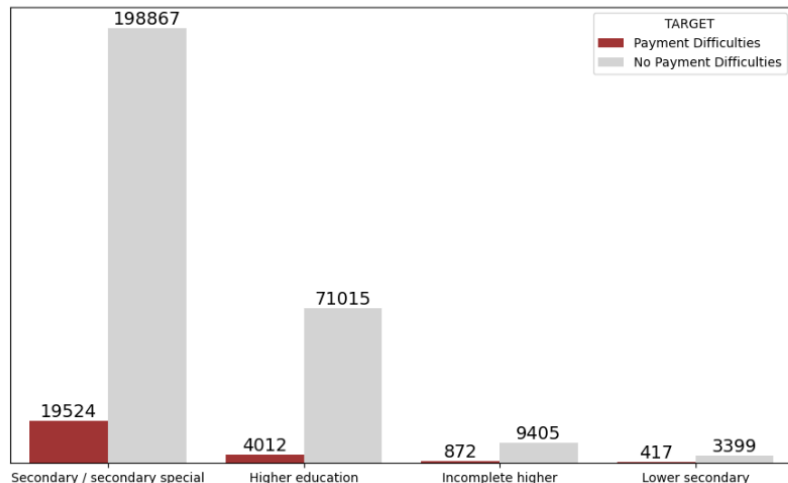
# Data Preprocessing



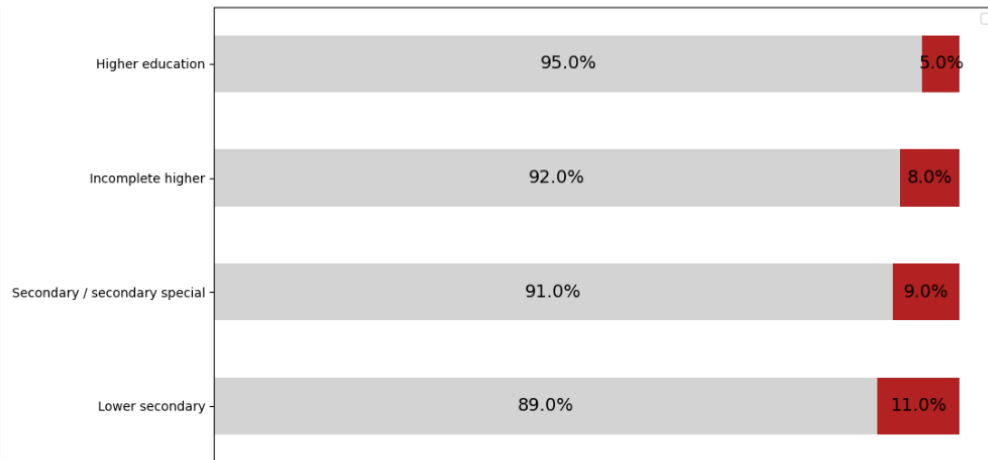


# Data Visualization & Business Insight

Number of Payment Abilities by Education



Percentage of Payment Abilities by Education

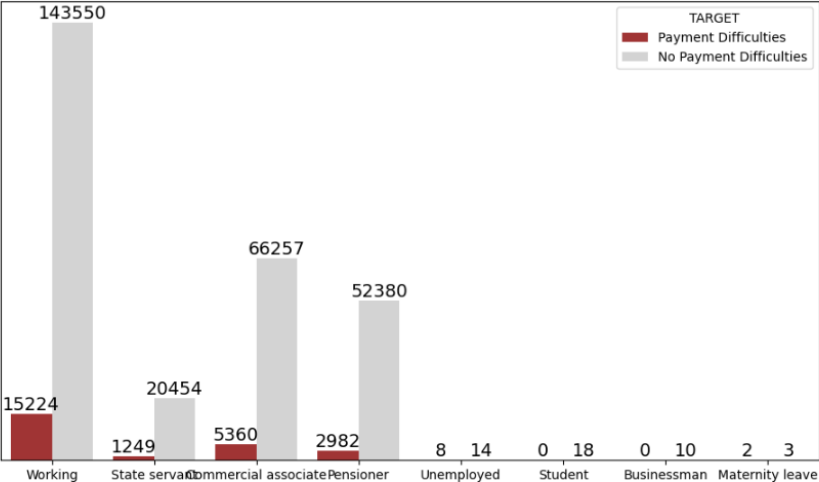


Berdasarkan latar belakang demografi, mayoritas nasabah yang mengajukan pinjaman berpendidikan pada tingkat *secondary education* atau tingkat menengah atas. Namun persentase kemampuannya dalam pembayaran cicilan cukup rendah daripada *higher education* dan *incomplete higher*. Perusahaan sebaiknya **mentargetkan strategi marketingnya kepada nasabah sarjana** (*higher education*) mengingat kemampuan pembayaran cicilannya paling tinggi.

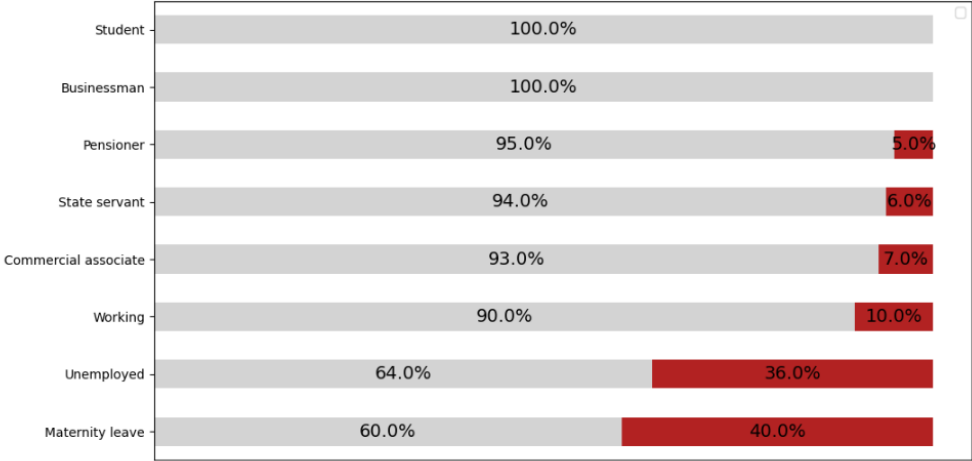


# Data Visualization & Business Insight

Number of Payment Abilities by Income Type



Percentage of Payment Abilities by Income Type



Pegawai Negeri (*state servant*) dan pensioner memiliki kemungkinan tinggi diterima pengajuannya. Namun secara keseluruhan total angka nasabah ini masih rendah daripada pekerja (*working*) yang menjadi mayoritas tipe sumber pemasukan nasabah. **Perlu dibuat campaign yang ditargetkan pada pegawai negeri dan pensioner untuk agar lebih tertarik untuk mengajukan pinjaman.**



# Machine Learning Implementation & Evaluation

```
from sklearn.model_selection import RandomizedSearchCV
from sklearn.linear_model import LogisticRegression
from sklearn import model_selection

# -----
weights = np.linspace(0.0, 0.99, 200)
param_distributions = {'class_weight': [{0:x, 1:1.0-x} for x in weights]}

# -----
logreg = LogisticRegression(random_state=123)
kfold = model_selection.StratifiedKFold(n_splits = 5, shuffle = True, random_state = 123)
logreg_tuned = model_selection.RandomizedSearchCV(logreg, param_distributions, n_jobs = -1,
                                                  n_iter=100, cv=kfold, scoring='roc_auc', random_state=123)

# fit model
logreg_tuned.fit(X_train_ros, y_train_ros)

# -----
print("Best parameters: ", logreg_tuned.best_params_)
```

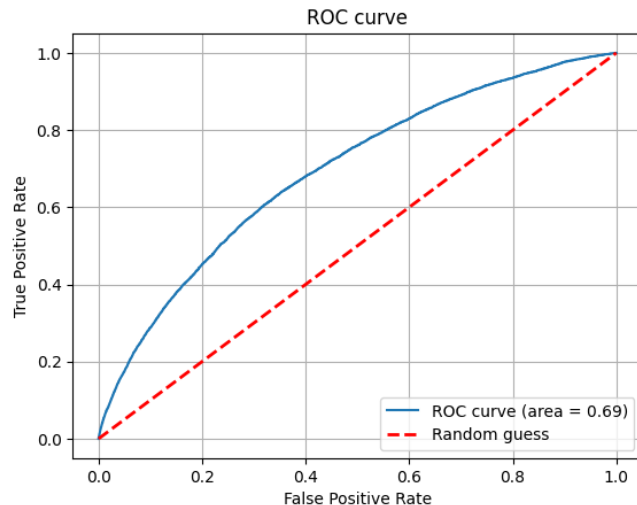
Best parameters: {'class\_weight': {0: 0.6467336683417085, 1: 0.35326633165829147}}

predict(logreg\_tuned)

MODEL EVALUATION :				
	precision	recall	f1-score	support
0	0.93	0.88	0.90	68999
1	0.21	0.33	0.25	6642
accuracy			0.83	75641
macro avg	0.57	0.60	0.58	75641
weighted avg	0.87	0.83	0.85	75641

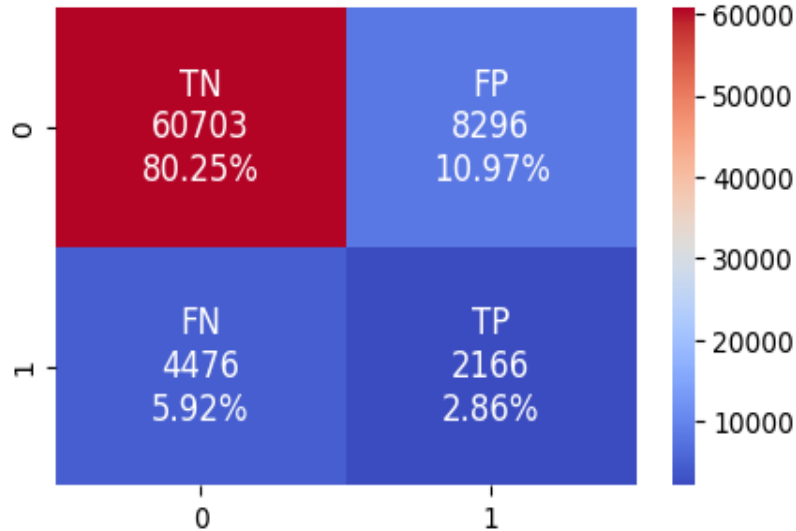
AUC Test score: 0.6935

AUC Train scores: 0.6964



- Berdasarkan nilai precision, recall, dan f1 dapat disimpulkan bahwa model cenderung memprediksi **Customer With No Payment Difficulties** dibandingkan dengan **Customer With Payment Difficulties**. Berdasarkan hasil akurasi model mampu memprediksi **83%** data dengan benar.
- Semakin tinggi nilai AUC, semakin baik kemampuan model dalam membedakan kelas. Dalam hal ini, nilai AUC test score adalah 0.6935, sedangkan AUC train scores adalah 0.6964, yang menunjukkan model memiliki kemampuan yang cukup baik dalam membedakan kedua kelas.

# Machine Learning Implementation & Evaluation



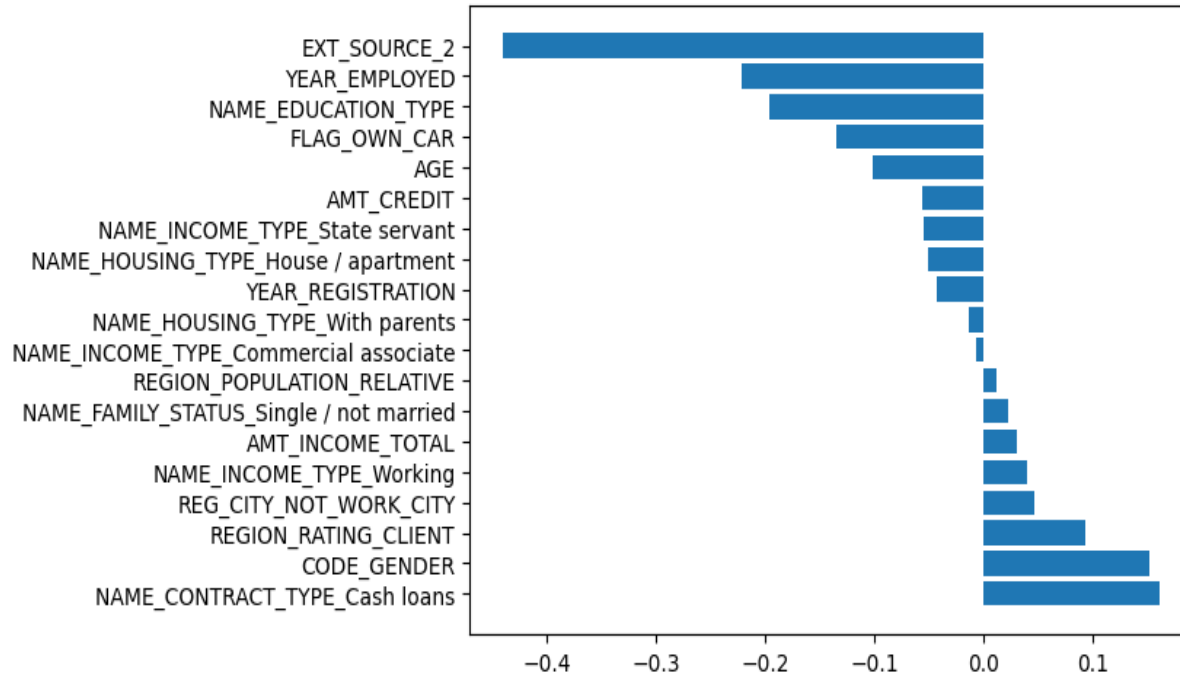
Dari hasil confusion matrix tersebut, dapat **dilihat bahwa model cenderung lebih akurat dalam memprediksi customer tanpa payment difficulties (TN)** dibandingkan dengan customer dengan *payment difficulties* (TP), karena nilai False Positives (FP) yang tinggi menunjukkan adanya kesalahan dalam memprediksi customer yang sebenarnya tidak mengalami *payment difficulties* sebagai customer dengan *payment difficulties*.

Jika terdapat banyak true positives (TP), perusahaan dapat mengambil langkah-langkah yang tepat untuk menyelesaikan tagihan yang tidak terbayar dan mengurangi risiko kehilangan pendapatan. Jika terdapat banyak true negatives (TN), perusahaan tidak perlu mengeluarkan sumber daya untuk menyelesaikan tagihan yang tidak perlu dan tidak akan kehilangan pendapatan dari customer yang tidak membayar.

Kesimpulannya, model tersebut dapat memprediksi customer tanpa *payment difficulties* dengan **cukup baik**, tetapi masih perlu ditingkatkan dalam memprediksi customer dengan *payment difficulties* untuk mengurangi jumlah False Positives (FP) dan False Negatives (FN).



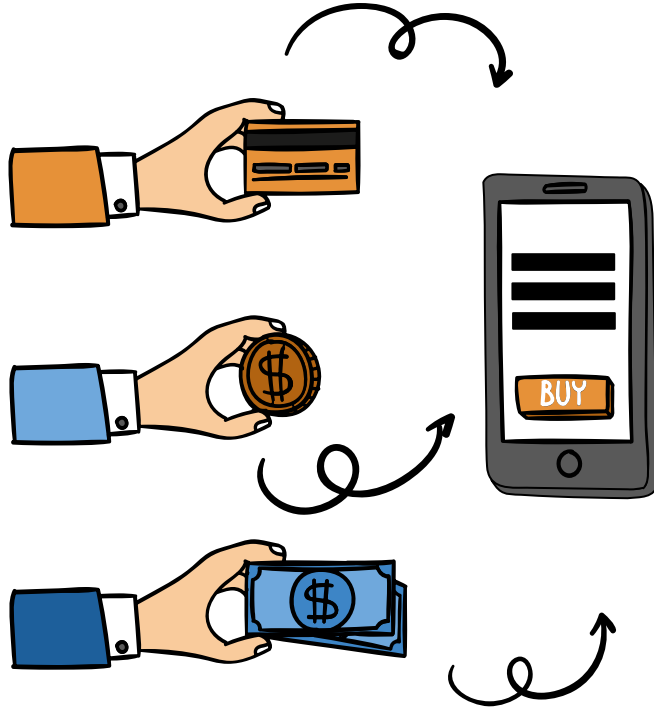
Feature importances obtained from coefficients



Feature importance dalam Logistic Regression digunakan untuk menentukan seberapa penting suatu fitur dalam mempengaruhi kelas target atau hasil prediksi model.

**Sumber eksternal 2, lama masa menjadi pekerja, tingkat pendidikan** memiliki pengaruh negatif terhadap tingkat *payment difficulties* pelanggan. Artinya semakin rendah nilai fitur tersebut maka customer cenderung memiliki *payment difficulties* yang besar.

# Business Recommendations



- Memprioritaskan dan mentargetkan customer yang memiliki tingkat pendidikan sarjana dan telah menikah.
- Meningkatkan seleksi calon peminjam dengan memperketat persyaratan dan mengumpulkan informasi yang lebih lengkap tentang calon peminjam
- Menawarkan opsi pembayaran yang fleksibel, seperti jangka waktu pembayaran yang lebih lama atau pembayaran yang teratur. Hal ini dapat membantu peminjam mengatur keuangan mereka dan menghindari keterlambatan atau gagal membayar cicilan.
- Mengadakan program campaign pinjaman untuk pegawai negeri dan pensioner.

# Thank You!

<https://github.com/faizns/HCI-vix-project>