



Progress Report FinPro Stage 3

Dino Kuning

- Fany Okpiani
- Nadhilah Farhana
- Raditya Satria Gantara
- Rafindra Prihaztama



Dino Kuning's Group Member



Fany Okpiani

Business / Data Analyst



Nadhilah Farhana

Data Scientist



Rafindra Prihaztama

Data Engineer



Raditya Satria G.

Project Manager



Outline - Stage 3

Final Model Testing



**Discuss Model
Evaluation**



Error Analysis

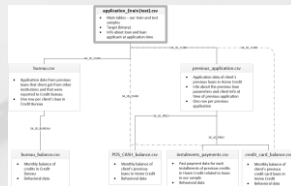


**Final Report
Drafting**



Adjustment

Skema 1



Merge semua dataset
Ukuran data: 307511×150

Model	Akurasi	ROC-AUC	Recall (0)	Precision (0)
XGBoost	92%	52%	99%	92%
Random Forest (RF)	92%	50%	100%	92%

Skema 2

Data application train
Ukuran data: 307511×20

Model	Akurasi	ROC-AUC	Recall (0)	Precision (0)
XGBoost	93%	50%	100%	93%
Random Forest (RF)	93%	50%	100%	93%

- **Recall 100%:** semua nasabah yang benar-benar mampu bayar (kelas 0) berhasil terdeteksi oleh model
- **Precision 92–93% :** mayoritas dari prediksi “mampu bayar” memang benar-benar mampu bayar, sehingga sangat efisien untuk skrining awal
- **Akurasi juga tinggi (92–93%),** menunjukkan bahwa model dapat diandalkan untuk seleksi otomatis nasabah potensial
- **Nilai ROC-AUC masih rendah,** artinya kemampuan model untuk membedakan antara nasabah bayar dan tidak bayar secara umum masih kurang tajam
- **Skema 2 lebih efisien** karena hanya menggunakan fitur dari `application_train`, tetapi tetap memberi hasil recall dan precision maksimal

- **Ketidakseimbangan Kelas (Class Imbalance):**
 - **Masalah:** Model lebih cenderung memprediksi kelas mayoritas (kelas 0) dan mengabaikan kelas 1 (nasabah tidak potensial).
 - **Solusi:** Bisa menggunakan teknik oversampling atau undersampling, atau sesuaikan bobot kelas.
- **Overfitting:**
 - **Masalah:** Model terlalu fit pada data pelatihan dan tidak generalisasi dengan baik pada data uji.
 - **Solusi:** Bisa menggunakan **cross-validation**, **regularization**, dan **early stopping**.
- **Model Tidak Mampu Menangkap Pola Kompleks:**
 - **Masalah:** Model sederhana seperti Logistic Regression tidak menangkap interaksi fitur yang kompleks.
 - **Solusi:** Bisa menggunakan model non-linear seperti XGBoost atau LightGBM, dan lakukan **feature engineering**.

Tuning Hyperparameter XGBoost

```
param_grid = {
    'n_estimators': [50, 100, 200],
    'learning_rate': [0.01, 0.1, 0.2],
    'max_depth': [3, 5, 7],
    'subsample': [0.8, 1.0],
    'colsample_bytree': [0.8, 1.0]
}
```

Fitting 3 folds for each of 108 candidates, totalling 324 fits

Best parameters: {'colsample_bytree': 1.0, 'learning_rate': 0.2, 'max_depth': 7, 'n_estimators': 200, 'subsample': 0.8}

Best ROC-AUC score: 0.9683626205423782

Accuracy (XGBoost): 0.9133863388777783

ROC-AUC (XGBoost): 0.5252367764380541

Recall (XGBoost): 0.06263891695291979

Precision (XGBoost): 0.3106212424849699

	precision	recall	f1-score	support
0	0.92	0.99	0.95	56554
1	0.31	0.06	0.10	4949
accuracy			0.91	61503
macro avg	0.62	0.53	0.53	61503
weighted avg	0.87	0.91	0.89	61503

Confusion Matrix (XGBoost):

```
[[55866  688]
 [ 4639  310]]
```

Hasil tuning hyperparameter tidak menunjukkan perubahan hasil evaluasi yang signifikan, namun model sempat mendapatkan nilai best ROC-AUC yang cukup tinggi

Apa itu Home Credit?

HOME CREDIT

Home Credit adalah perusahaan pembiayaan berbasis teknologi yang menyediakan layanan kredit konsumen, terutama untuk pembelian barang elektronik, furnitur, hingga pinjaman tunai. Home Credit beroperasi di berbagai negara, termasuk Indonesia, dengan fokus pada **nasabah unbanked dan underbanked** – yaitu mereka yang belum terjangkau layanan perbankan formal.

1

Apa Tantangan yang Dihadapi Home Credit Saat Ini?

» Tantangan Proses Seleksi Awal Nasabah Secara Manual

Proses manual dalam menyeleksi calon nasabah **memerlukan waktu dan sumber daya yang besar**. Setiap aplikasi kredit harus diperiksa secara individual, yang **meningkatkan risiko kesalahan manusia dan inkonsistensi dalam penilaian**. Selain itu, metode manual sulit untuk mengakomodasi volume aplikasi yang besar secara efisien.



Home Credit Indonesia
2020: 4,7 Juta Pelanggan

Sumber: homecredit.co.id

Dengan meningkatnya jumlah aplikasi kredit, Home Credit dihadapkan pada kebutuhan untuk menyaring aplikasi secara cepat tanpa mengorbankan akurasi. **Penyaringan yang lambat dapat menurunkan kepuasan pelanggan dan meningkatkan risiko kehilangan peluang bisnis**. Sebaliknya, penyaringan yang tidak akurat dapat menyebabkan pemberian kredit kepada individu yang tidak memenuhi syarat, meningkatkan risiko kredit macet.

2



Thank you!