

Analisis Kinerja Kredit Nasabah (Final Project)

Project-Based Internship Virtual: Data Scientist
Home Credit Indonesia x Rakamin Academy

Ahmad Miftahul Farohi



Table of Content

- Problem Research
- Data Pre-Processing
- Data Vizualization and Business Insight
- Machine Learning Implementation and Evaluation
- Business Recommendation

Problem Reserach

- **Masalah:**

Menganalisis kinerja kredit nasabah berdasarkan data pada tabel "**bureau**" dan "**bureau_balance**" untuk mengidentifikasi pola-pola yang dapat membantu dalam pengambilan keputusan terkait penilaian kredit.

- **Tujuan:**

Meningkatkan akurasi penilaian kredit, mengurangi risiko kredit macet, dan meningkatkan profitabilitas perusahaan.

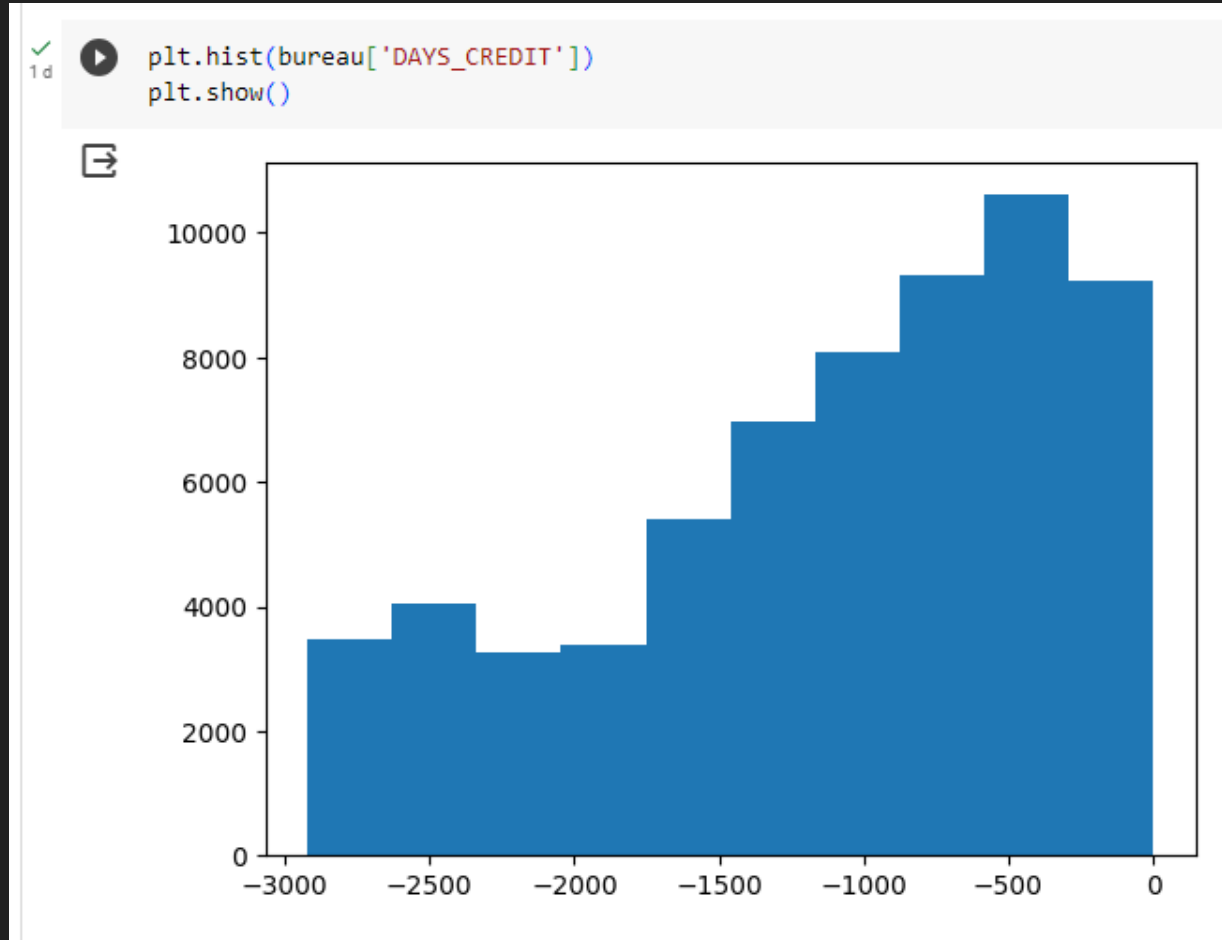
Data Pre-Processing

- Tools yang digunakan adalah Google Colaboratory
- Import Library. Library yang digunakan adalah pandas, matplotlib.pyplot, dan scikitlearn.
- Load Dataset. Tabel yang digunakan adalah **bureau** dan **bureau_balance**.
- Cek dan handle missing value.
- Cek dan handle data duplikat.
- Gabungkan semua data berdasarkan kolom yang sama, yakni kolom **SK_ID_BUREAU**

Data Visualization and Business Insight

Insight:

Permohonan kredit terbanyak ada pada kisaran 500 hari sebelum data direkam.



Machine Learning Implementation and Evaluation

- Menggunakan regresi logistik untuk memprediksi CREDIT_ACTIVE

```
✓ [17] X = bureau[['DAYS_CREDIT', 'AMT_CREDIT_SUM']]  
      y = bureau['CREDIT_ACTIVE']  
  
      # Split data  
      X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=0)
```

Model

```
✓ [18] model = LogisticRegression()  
      model.fit(X_train, y_train)  
      y_pred = model.predict(X_test)
```

Machine Learning Implementation and Evaluation

○ Evaluas performa model

```
✓ [19] print('Accuracy:', accuracy_score(y_test, y_pred))  
1d      print('Confusion Matrix:', confusion_matrix(y_test, y_pred))  
      print('Classification Report:', classification_report(y_test, y_pred))
```

```
Accuracy: 0.7027450980392157  
Confusion Matrix: [[1426    0 3398    0]  
 [   0     0   1     0]  
 [ 342     0 7534     0]  
 [  11     0   38     0]]
```

Classification Report:		precision	recall	f1-score	support
Active	0.80	0.30	0.43		4824
Bad debt	0.00	0.00	0.00		1
Closed	0.69	0.96	0.80		7876
Sold	0.00	0.00	0.00		49
accuracy			0.70		12750
macro avg	0.37	0.31	0.31		12750
weighted avg	0.73	0.70	0.66		12750

Insight:

Dari hasil model yang diberikan, terdapat beberapa insight yang dapat diambil, antara lain: Akurasi model sebesar 70.27%, yang menunjukkan bahwa model tersebut cukup baik dalam memprediksi kategori-kategori yang ada.

Namun, terdapat beberapa kategori yang sulit diprediksi, seperti "Bad debt" dan "Sold", yang memiliki nilai precision dan recall sebesar 0. Hal ini menunjukkan bahwa model perlu diperbaiki untuk dapat memprediksi kategori-kategori ini dengan lebih baik.

Business Recommendations

Berdasarkan insight-insight yang ditemukan dari analisis data, beberapa rekomendasi bisnis yang dapat diberikan adalah:

1. Perlu dilakukan evaluasi ulang terhadap kebijakan penilaian kredit yang digunakan. Hal ini dapat meliputi peninjauan kembali terhadap faktor-faktor yang digunakan dalam penilaian kredit, serta penyesuaian terhadap model penilaian kredit yang digunakan.
2. Perlu dilakukan penelitian lebih lanjut untuk mengetahui penyebab dari kategori-kategori yang sulit diprediksi. Hal ini dapat meliputi peninjauan kembali terhadap data yang digunakan, serta penyesuaian terhadap fitur-fitur yang digunakan dalam model penilaian kredit.
3. Perlu dilakukan monitoring secara berkala terhadap model penilaian kredit yang digunakan. Hal ini dapat meliputi pemantauan terhadap performa model penilaian kredit, serta peninjauan kembali terhadap model penilaian kredit yang digunakan.

Tautan Lengkap di Github

- Silakan kunjungi github untuk lebih lengkapnya

<https://github.com/mieffarohi/pbi-hci>



Terima Kasih