



FINAL TASK DATA SCIENTIST ID/X PARTNER

YUNITA SALSABILA



Business Problem Understanding



Mengurangi risiko kredit macet

Hal ini penting untuk menjaga stabilitas keuangan perusahaan dan menghindari kerugian finansial. Contohnya, jika perusahaan memberikan pinjaman kepada pelanggan berisiko tinggi, ada kemungkinan besar mereka tidak akan dapat membayar pinjaman tersebut. Hal ini dapat menyebabkan perusahaan kehilangan uang dan mengalami kesulitan keuangan.

Meningkatkan pengambilan keputusan kredit

Dengan memprediksi risiko kredit secara akurat, perusahaan dapat membuat keputusan yang lebih baik tentang siapa yang akan diberi pinjaman dan berapa banyak yang akan dipinjamkan. Hal ini dapat membantu perusahaan menghindari memberikan pinjaman kepada pelanggan yang berisiko tinggi dan meningkatkan peluang untuk mendapatkan keuntungan dari pinjaman.

Meminimalkan kerugian perusahaan

Kerugian dapat terjadi ketika perusahaan memberikan pinjaman kepada pelanggan yang tidak dapat membayarnya. Dengan memprediksi risiko kredit secara akurat, perusahaan dapat meminimalkan risiko kerugian dan meningkatkan profitabilitas.

Business Problem Understanding



MASALAH

Memprediksi risiko kredit dari pelanggan yang mengajukan pinjaman

Hal ini dapat menjadi tugas yang sulit karena ada banyak faktor yang dapat memengaruhi kemampuan pelanggan untuk membayar pinjaman. Faktor-faktor ini termasuk riwayat kredit pelanggan, pendapatan, pekerjaan, dan situasi keuangan.



PENDEKATAN

Berkomunikasi dengan tim manajemen perusahaan, tim risiko, dan tim keuangan untuk memahami kebutuhan dan perspektif mereka

Hal ini penting untuk memastikan bahwa model prediksi risiko kredit memenuhi kebutuhan semua pemangku kepentingan.

Mengembangkan model prediksi risiko kredit

Model ini dapat menggunakan berbagai teknik statistik dan data untuk memprediksi probabilitas pelanggan gagal membayar pinjaman.

Business Problem Understanding



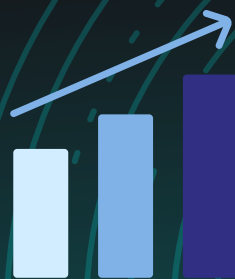
KRITERIA KEBERHASILAN

Meningkatkan akurasi prediksi risiko kredit

Hal ini dapat dicapai dengan menggunakan model prediksi yang akurat dan data yang up-to-date

Mengurangi risiko kredit macet

Hal ini dapat dicapai dengan menggunakan model prediksi risiko kredit untuk membuat keputusan kredit yang lebih baik



ANALYTIC APPROACH

- Analisis data akan dilakukan untuk menemukan pola pada fitur-fitur yang terdapat dalam dataset, yang membedakan karakteristik setiap pengguna.
- Mengingat permasalahan yang ada, kami akan membangun model machine learning karena kami perlu membuat prediksi yang lebih akurat dibandingkan hanya menggunakan pendekatan analisis inferensial dan/atau deskriptif.

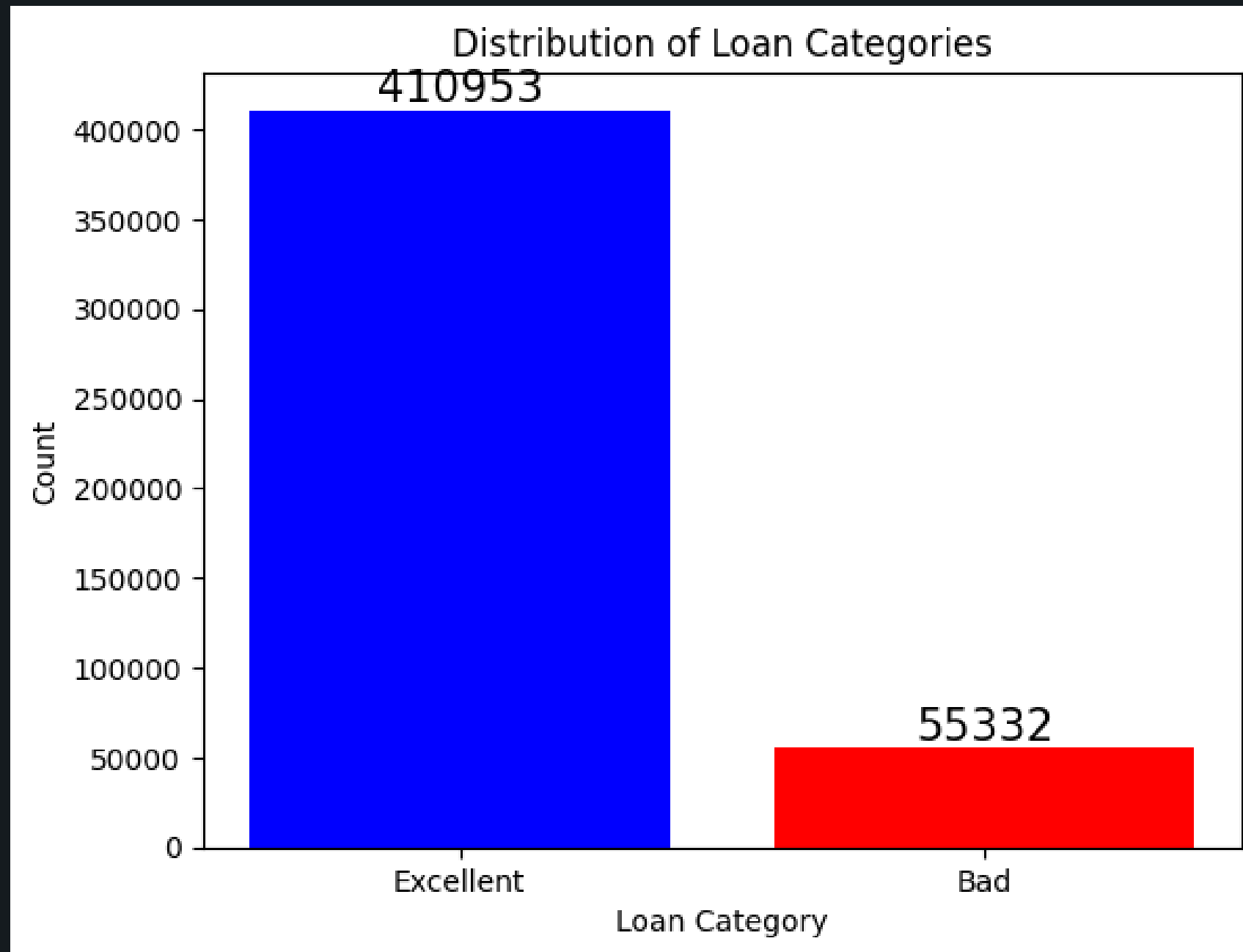
Dataset General Info

	Unnamed: 0	id	member_id	loan_amnt	funded_amnt	funded_amnt_inv	term	int_rate	installment	grade	...	total_bal_il	il_util
0	0	1077501	1296599	5000	5000	4975.0	36 months	10.65	162.87	B	...	NaN	NaN
1	1	1077430	1314167	2500	2500	2500.0	60 months	15.27	59.83	C	...	NaN	NaN
2	2	1077175	1313524	2400	2400	2400.0	36 months	15.96	84.33	C	...	NaN	NaN
3	3	1076863	1277178	10000	10000	10000.0	36 months	13.49	339.31	C	...	NaN	NaN
4	4	1075358	1311748	3000	3000	3000.0	60 months	12.69	67.79	B	...	NaN	NaN

5 rows × 75 columns

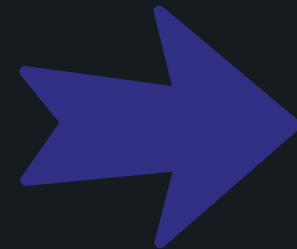
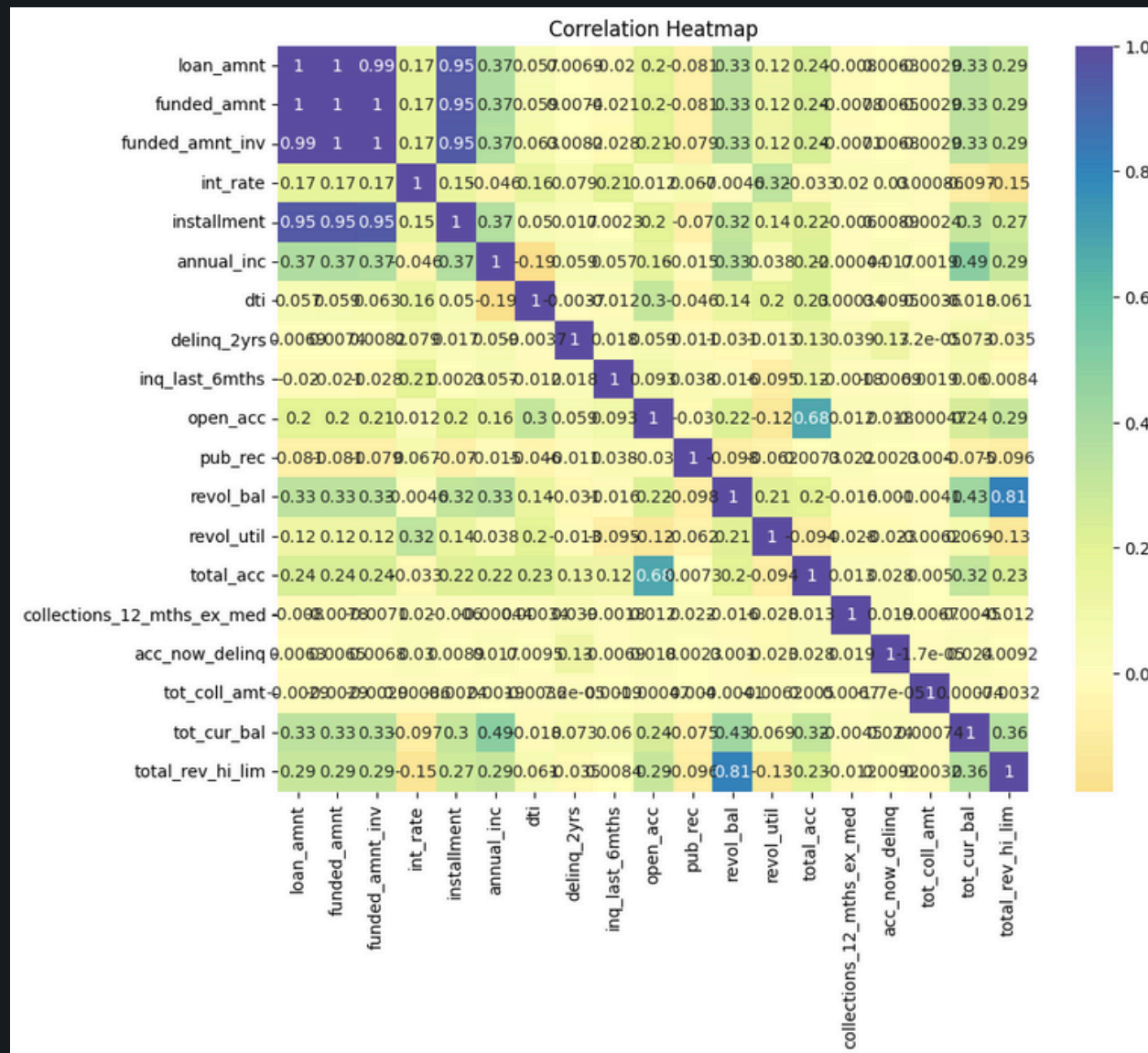
	Unnamed: 0	id	member_id	loan_amnt	funded_amnt	funded_amnt_inv	int_rate	installment	annual_inc
count	466285.000000	4.662850e+05	4.662850e+05	466285.000000	466285.000000	466285.000000	466285.000000	466285.000000	4.662810e+05
mean	233142.000000	1.307973e+07	1.459766e+07	14317.277577	14291.801044	14222.329888	13.829236	432.061201	7.327738e+04
std	134605.029472	1.089371e+07	1.168237e+07	8286.509164	8274.371300	8297.637788	4.357587	243.485550	5.496357e+04
min	0.000000	5.473400e+04	7.047300e+04	500.000000	500.000000	0.000000	5.420000	15.670000	1.896000e+03
25%	116571.000000	3.639987e+06	4.379705e+06	8000.000000	8000.000000	8000.000000	10.990000	256.690000	4.500000e+04
50%	233142.000000	1.010790e+07	1.194108e+07	12000.000000	12000.000000	12000.000000	13.660000	379.890000	6.300000e+04
75%	349713.000000	2.073121e+07	2.300154e+07	20000.000000	20000.000000	19950.000000	16.490000	566.580000	8.896000e+04
max	466284.000000	3.809811e+07	4.086083e+07	35000.000000	35000.000000	35000.000000	26.060000	1409.990000	7.500000e+06

Data Preparation



- Excellent Loan: Untuk pinjaman yang lunas tepat waktu.
- Bad Loan: Untuk pinjaman yang gagal bayar atau default.

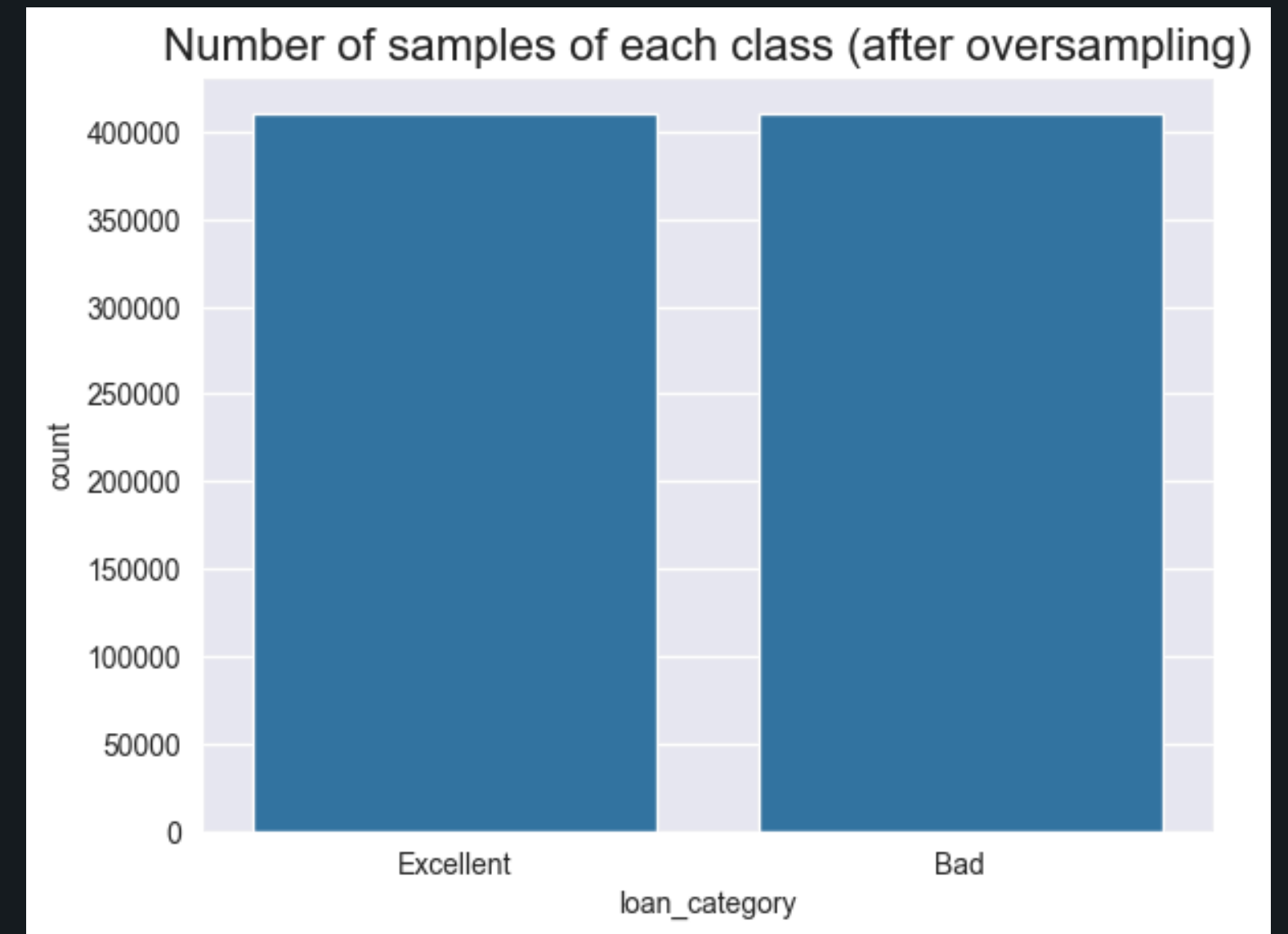
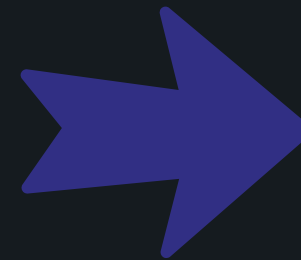
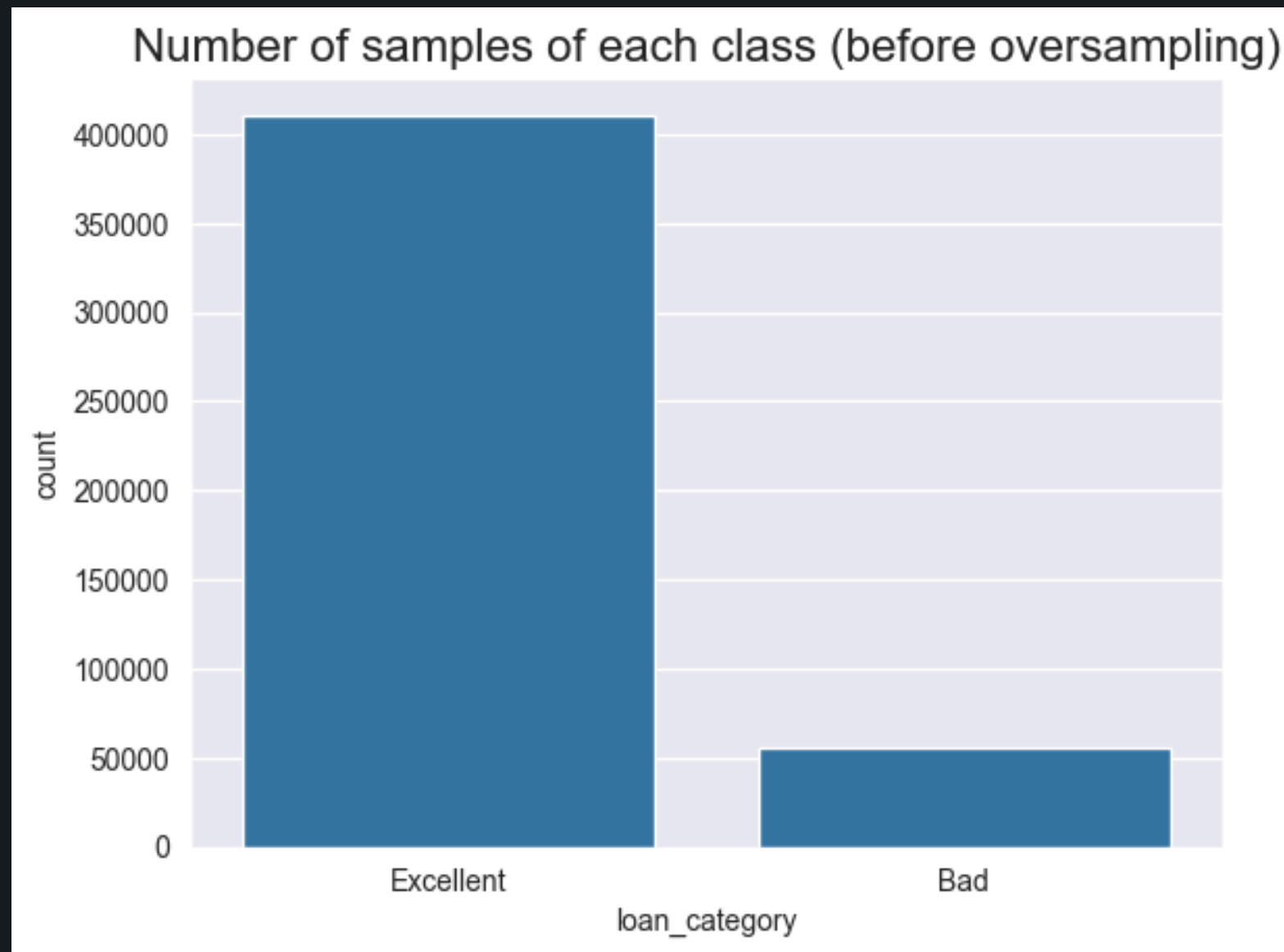
Data Preparation



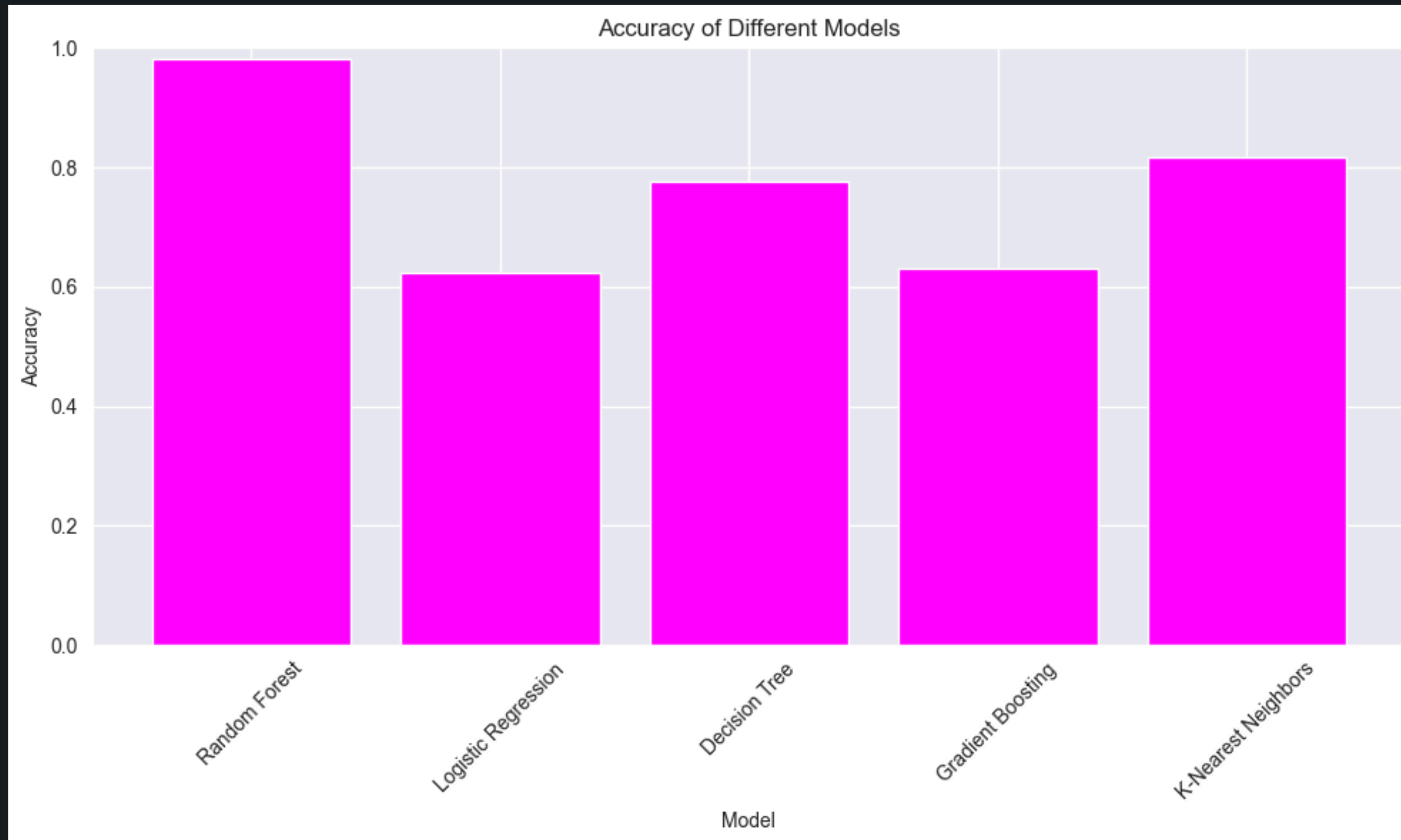
Skala Korelasi:

- Korelasi memiliki nilai antara -1 dan 1.
- Nilai 1 menunjukkan korelasi positif sempurna: ketika satu variabel naik, variabel lainnya juga naik.
- Nilai -1 menunjukkan korelasi negatif sempurna: ketika satu variabel naik, variabel lainnya turun.
- Nilai 0 menunjukkan tidak ada korelasi: perubahan pada satu variabel tidak berpengaruh pada variabel lainnya.

Data Cleaning



Modelling



RANDOM FOREST

0.982

LOGISTIC REGRESSION

0.623

DECISION TREE

0.77

GRADIENT BOOSTING

0.630

K-NEAREST NEIGHBORS

0.816

Thank you!

 +6285869784202

 yunitasalsabila29699@gmail.com

 <https://www.linkedin.com/in/yunitasalsabila/>

 <https://github.com/yunitasalsabila290699>