

Penerapan Data Science dalam Industri Perbankan : Studi Kasus Sistem Credit Scoring

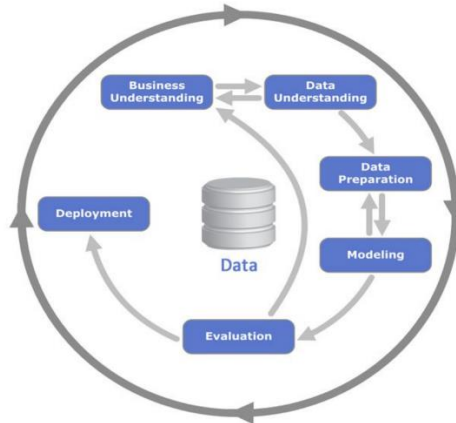
Perkembangan informasi sangat berdampak pada kehidupan manusia, salah satunya adalah penggunaan data. Setiap orang saat ini menggunakan dan menghasilkan data baru setiap harinya. Jika kondisi ini pada setiap orang digabungkan akan menghasilkan data yang besar. Pertumbuhan data yang sangat cepat dan bervariasi dalam skala besar membutuhkan pengelolaan agar menghasilkan suatu informasi baru yang lebih berguna. Ilmu modern yang secara spesifik mampu menjawab dan memberikan informasi dari banyaknya data adalah data science. Ada banyak sekali industri saat ini yang menggunakan data science untuk menghasilkan sesuatu yang lebih insightful. Salah satunya adalah industri perbankan. Penerapan data science dalam industri perbankan dapat dilihat pada sistem credit scoring. Lembaga keuangan seperti Bank biasanya menggunakan sistem credit scoring dalam pemberian kredit kepada nasabah. Sistem ini membantu lembaga keuangan dalam menentukan siapa nasabah yang layak mengusulkan dan menerima kredit dan berapa banyak jumlah yang harus diberikan demi meminimalisasi resiko kerugian yang bisa timbul apabila terjadi kesalahan pemberian pinjaman kredit. Dalam mengelompokkan layak atau tidaknya nasabah dengan keragaman karakteristik calon peminjam, biasanya didasarkan pada indeks-indeks tertentu seperti ketaatan, rekam jejak calon peminjam, dan sebagainya.

Untuk lebih detail mengenai credit scoring, saya akan mengambil salah satu literatur jurnal yang berjudul Sistem Pendukung Keputusan Kredit Usaha Rakyat PT. Bank Rakyat Indonesia Unit Kaliangkrik Magelang. Penelitian tersebut mencoba memanfaatkan data riwayat pembayaran kredit debitur lama yang ada di obyek penelitian serta menggunakan Decision Support System (DSS) dengan metode Mining Classification Rule menggunakan algoritma k-nearest neighbor untuk membentuk model prediksi pembayaran kredit calon nasabah. Output yang dihasilkan berupa nilai kedekatan nasabah baru dengan nasabah lama berdasarkan dari hasil hitung kemiripan variabel kasus baru dengan kasus nasabah lama. Nilai kedekatan kasus terbesar akan dijadikan acuan dalam mengambil keputusan.

Proses CRISP-DM

Cross-Industry Standard Process for Data Mining (CRISP-DM) merupakan standar yang digunakan untuk memproses data. Suatu informasi dari sejumlah data tidak pernah hadir dari proses sesaat, namun dari proses matang yang dipandu oleh standard yang terverifikasi. CRISP-DM adalah salah satu dari banyaknya standar yang digunakan sebagai acuan dalam mengolah data.

CRISP-DM Process Diagram



source: Kenneth Jensen

1. Business Understanding

Tahap pertama dari CRISP DM adalah Business Understanding yang merupakan tahap dimana masalah bisnis didefinisikan dengan tepat dengan memahami tujuan dan kebutuhan dari sudut pandang bisnis. Dalam studi kasus yang dibahas pada jurnal *Sistem Pendukung Keputusan Kredit Usaha Rakyat PT. Bank Rakyat Indonesia Unit Kaliangkrik Magelang*, masalah yang ingin dijawab adalah terkait seringnya terjadi kemacetan pada nasabah saat membayar kreditnya. Dengan adanya sistem credit scoring, nantinya akan memudahkan Bank dalam menyeleksi calon nasabah yang layak diberikan kredit untuk mengurangi risiko kemacetan seperti yang sering terjadi.

2. Data Understanding

Data Understanding merupakan proses memahami data apa yang diperlukan. Dalam tahap ini, peneliti melakukan review terhadap beberapa penelitian terdahulu terkait pembentukan sistem credit scoring untuk membantu memahami data yang diperlukan. Sebagai bagian dalam memahami data, peneliti juga melakukan observasi secara langsung ke lokasi penelitian untuk mengamati bagaimana syarat dan proses pengajuan kredit usaha rakyat yang dilakukan oleh pihak BRI Unit Kaliangkrik Cabang Magelang.

3. Data Preparation

Tahap ketiga adalah tahap data preparation. Tahap ini adalah proses mempersiapkan data untuk nantinya dilakukan modelling. Pada tahap ini juga mencakup pemilihan tabel, record, dan atribut-atribut data, termasuk proses pembersihan dan transformasi data. Atribut dalam penelitian ini dipilih berdasarkan faktor-faktor yang diturunkan dari prinsip-prinsip yang digunakan dalam pertimbangan pengajuan kredit menggunakan prinsip penilaian “5C” yaitu

character, capital, capacity, conditions of economy, dan collateral. Pada tahap transformasi data dilakukan pembobotan pada tiap-tiap atribut pada data training.

Tabel 1. Tabel Data Atribut dan Nilainya

No	Atribut	Nilai Atribut
1	Penghasilan Perbulan	<ul style="list-style-type: none"> • >3x angsuran • 3x angsuran • 2x angsuran • 1x angsuran
2	Nilai Jaminan	<ul style="list-style-type: none"> • > Jumlah Pinjaman • = Jumlah Pinjaman • < Jumlah Pinjaman
3	Jumlah tanggungan	<ul style="list-style-type: none"> • Tidak ada • 1 orang • 2-3 orang • > 3 orang
4	Pendidikan terakhir	<ul style="list-style-type: none"> • Universitas • SLTA • SLTP • SD
5	Lama Usaha	<ul style="list-style-type: none"> • >5 tahun • 3 - 5 tahun • 1 - 2 tahun • < 1 tahun
6	Status Perusahaan	<ul style="list-style-type: none"> • Swasta besar • Swasta menengah • Swasta kecil • Perorangan
7	Usia	<ul style="list-style-type: none"> • < 20 tahun / > 60 tahun • 20 - 50 tahun • 51 - 60 tahun
8	Status	<ul style="list-style-type: none"> • Menikah • Belum menikah • Janda/Duda
9	Kelas	<ul style="list-style-type: none"> • Bad • Normal • Good

Proses akhir dari tahap ini menghasilkan data *training* seperti pada tabel 2.

Tabel 2. Data Training Hasil Proses *Cleaning*

N o	K D T	Nama	Ala mat	JK	A	B	C	D	E	F	G	H	T	S
1	00 1	xxxx	xxxx	L	2 x	=	2	SMA	1	Kecil	20-50	Meni kah	4	Baik
2	00 2	xxxx	xxxx	L	2 x	=	3	UNI	1-2	Menengah	20-50	Meni kah	3	Baik
3	00 3	xxxx	xxxx	P	1 x	=	2	SMA	1	Kecil	51-60	Meni kah	2	Baik
4	00 4	xxxx	xxxx	L	1 x	>	1	SMA	1	Kecil	51-60	Meni kah	1	Baik
5	00 5	xxxx	xxxx	L	2 x	>	1	SMA	3-5	Menengah	20-50	Meni kah	0	Baik
6	00 6	xxxx	xxxx	L	2 x	<	1	UNI	1-2	Menengah	20-50	Meni kah	5	Nor mal
7	00 7	xxxx	xxxx	L	3 x	<	2	UNI	>5	Menengah	<20 >60	Meni kah	6	Nor mal
8	00 8	xxxx	xxxx	P	2 x	>	1	SMA	1-2	Menengah	20-50	Meni kah	7	Nor mal

Keterangan

No : Nomor Urut

KDT : Kode Data Training

JK : Jenis Kelamin

A : Pendapatan Perbulan

B : Nilai Jaminan

C : Jumlah Tanggungan

S : Status Debitur

D : Pendidikan Terakhir

E : Lama Usaha

F : Kondisi Usaha

G : Usia

H : Status Pernikahan

T : Jumlah Tunggalan

Setelah diperoleh data *training*, kemudian dilakukan pendefinisian bobot pada setiap variable yang telah ditentukan. Pendefinisian bobot didasarkan pada atribut yang paling berpengaruh untuk pemberian kredit di BRI Kaliangkrik Bagian KUR.

Tabel 3. Tabel Bobot Atribut

Nama Atribut	Tingkat Kepentingan	Bobot
Penghasilan Perbulan	1	0,340
Nilai Jaminan	2	0,215
Jumlah Tanggungan Keluarga	3	0,152
Pendidikan Terakhir	4	0,111
Lama Usaha	5	0,079
Kondisi Usaha	6	0,054
Usia	7	0,034
Status Pernikahan	8	0,016

4. Modelling

Dalam tahap ini akan dilakukan pemilihan dan penerapan berbagai teknik pemodelan. Ada banyak teknik yang dapat digunakan, peneliti menggunakan metode Mining Classification Rule menggunakan algoritma K-Nearest Neighbor (KNN) untuk membentuk model prediksi pembayaran kredit calon nasabah. Algoritma KNN merupakan algoritma klasifikasi mengklasifikasikan data berdasarkan similarity atau kemiripan atau kedekatannya terhadap data lainnya.

5. Evaluasi

Evaluasi merupakan tahap validasi model yang dibentuk. Pada tahap ini akan dilakukan evaluasi terhadap keefektifan dan kualitas model sebelum digunakan dan menentukan apakah model dapat mencapai tujuan yang ditetapkan pada fase awal (Business Understanding). Data testing digunakan untuk mengevaluasi model yang terbentuk. Evaluasi algoritma KNN biasanya dilakukan menggunakan n-fold Cross Validation. n-fold cross validation atau k-fold cross validation adalah salah satu metode yang digunakan untuk mengetahui rata-rata keberhasilan dari suatu sistem dengan cara melakukan perulangan dengan mengacak atribut masukan sehingga sistem tersebut teruji untuk beberapa atribut input yang acak. Secara keseluruhan, tingkat akurasi yang diberikan model dalam penelitian ini adalah 93.023%. Tingkat akurasi berada dalam level sangat baik, sehingga dapat dilanjutkan untuk membentuk/deployment sistem.

6. Deployment

Pada tahap ini, pengetahuan atau informasi yang telah didapatkan akan diatur dan dinyatakan dalam bentuk khusus sehingga dapat digunakan. Berikut ini tampilan sistem yang telah berhasil dibentuk oleh peneliti :

Analisi Kasus

Nama

Tampil

DTR	Nama	Omset	Jaminan	Tanggungan	Pendidikan	Lama	Kondisi	Umur	Status	Remark
TR001	xxxx	2x angsuran	= jumlah pinjaman	2-3 orang	SLTA	1-2 tahun	kecil	20-50	menikah	baik
TR002	xxxx	2x angsuran	= jumlah pinjaman	2-3 orang	Universitas	1-2 tahun	menengah	20-50	menikah	baik
TR003	xxxx	1x angsuran	= jumlah pinjaman	1-2 orang	SLTA	1-2 tahun	kecil	51-60	menikah	baik
TR004	xxxx	1x angsuran	> jumlah pinjaman	1 orang	SLTA	1-2 tahun	kecil	51-60	menikah	baik
TR005	xxxx	2x angsuran	> jumlah pinjaman	1 orang	SLTA	3-5 tahun	menengah	20-50	menikah	baik
TR006	xxxx	2x angsuran	< jumlah pinjaman	1 orang	Universitas	1-2 tahun	menengah	20-50	menikah	normal
TR007	xxxx	3x angsuran	< jumlah pinjaman	2-3 orang	Universitas	>5 tahun	menengah	<20 / >60	menikah	normal
TR008	xxxx	2x angsuran	> jumlah pinjamn	1 orang	SLTA	1-2 tahun	menengah	20-50	menikah	normal
TR009	xxxx	1x angsuran	< jumlah pinjaman	tidak ada	SMP	<1 tahun	menengah	51-60	belum menikah	buruk
TR010	xxxx	1x angsuran	= jumlah pinjaman	2-3 orang	SD	<1 tahun	perorangan	51-60	menikah	buruk

Showing 1 to 10 of 43 entries

[← Previous](#) [1](#) [2](#) [3](#) [4](#) [5](#) [Next →](#)

Berdasarkan contoh kasus diatas, bisa dilihat bagaimana penggunaan data dapat memberikan solusi atas permasalahan yang ada. Dengan menggunakan data rekam jejak nasabah yang ada sebelumnya, dibentuk sistem credit scoring menggunakan pendekatan data science. Sistem credit scoring yang terbentuk dapat digunakan pihak bank untuk memutuskan apakah seorang nasabah layak untuk diberikan kredit. Tentunya akan sulit untuk menentukan layak atau tidaknya nasabah diberikan kredit jika dilakukan secara manual, karena beragamnya karakteristik data setiap nasabah. Selain menawarkan kemudahan dan keefektifan, sistem credit scoring ini juga memberikan akurasi yang sangat baik. Terlihat dari nilai akurasi yang didapatkan dalam penelitian ini 93.023%.

Penerapan data science di dunia perbankan bukan hanya dalam bentuk sistem credit scoring. Sistem fraud detection, sistem CLV (Customer Lifetime Value), dan lain sebagainya menerapkan data science untuk keberlangsungan bisnis perbankan. Data is the new oil, data sangat berharga dibandingkan dengan minyak dan termasuk kekayaan baru. Di era modern seperti sekarang ini, data dapat menjadi kunci dalam pembangunan berbagai segmen kehidupan.

RUJUKAN :

Nugroho, A., Kusrini, K., & Arief, M. R. (2015). Sistem Pendukung Keputusan Kredit Usaha Rakyat PT. Bank Rakyat Indonesia Unit Kaliangkrik Magelang. *Creative Information Technology Journal*, 2(1), 1. <https://doi.org/10.24076/citec.2014v2i1.33>