

PENERAPAN ALGORITMA NAÏVE BAYES UNTUK MEMPREDIKSI TINGKAT PERSETUJUAN KARTU KREDIT

PROPOSAL TUGAS AKHIR



Untuk Memenuhi Persyaratan Penulisan
Pada Mata Kuliah Tugas Akhir

Disusun Oleh :
Fareza Syahwaldi
8020190104

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS DINAMIKA BANGSA
2022**

IDENTITAS PROPOSAL PENELITIAN

Judul Proposal : Penerapan Algoritma Naïve Bayes Untuk Memprediksi
Tingkat Persetujuan Kartu Kredit

Program Studi : Teknik Informatika

Jenjang Pendidikan : Strata 1 (S1)

Peneliti :

- a. Nama Lengkap : Fareza Syahwaldi
- b. NIM : 8020190104
- c. Jenis Kelamin : Laki-Laki
- d. Tempat/Tgl Lahir : Jambi/13 Januari 2002
- e. Alamat : Jl.Kopral Umar, RT.19,
Kec, Kota Baru, Jambi
- f. No. Telepon : 082172257901
- g. Email : farezasyahwaldi37@gmail.com

BAB I

PENDAHULUAN

1.1 LATAR BELAKANG

Masyarakat di negara dunia memiliki kecenderungan untuk berperilaku konsumtif. Hal ini juga dipengaruhi oleh beberapa faktor seperti perkembangan dunia teknologi yang memunculkan berbagai variasi produk keuangan inovatif yang ditawarkan oleh perbankan dengan semakin mudahnya transaksi, perkembangan dunia industri yang menyebabkan masyarakat memiliki dorongan untuk selalu mengedepankan gaya hidup sehingga seseorang bisa mengeluarkan uang banyak dalam sekejap hanya untuk berbelanja [1].

Pada dunia perbankan pemberian kartu kredit kepada nasabah merupakan hal yang sulit dan dengan resiko yang sangat tinggi. Dalam pengajuannya, harus analisis data yang jelas dan tidak pernah mengalami masalah ke bank yang lain. Dan berkas yang meyakinkan pihak bank untuk memberikan kartu kredit seperti data pekerjaan, penghasilan perbulan, tempat tinggal, usia dan masih banyak yang perlu dipertimbang oleh bagian analisis bank. Oleh karna itu banyak nasabah yang gagal dan pihak bank salah untuk menganalisis sehingga mengakibatkan kerugian yang besar seperti macet, lari nasabah. Maka dengan masalah tersebut perlu menggunakan sistem prediksi yang jelas dan yang bisa dipercayai kepada bagian analisis bank. maka dibangun sistem prediksi penerapan *Algoritma Naïve Bayes* dalam memprediksi kelayakan penerimaan kartu kredit [2].

Banyak Penelitian yang memprediksi pembayaran kartu kredit dengan teknik klasifikasi *data mining*. Diantaranya dilakukan oleh [3]. dengan cara membandingkan nilai akurasi dari setiap

algoritma, dimana *Naïve Bayes* nilai akurasi 80,6%, *K-Nearest Neighbor* 79,2%, dan *C4.5* 79%. Hasil dari penelitian tersebut *Naïve Bayes* dan *K-Nearest Neighbor* memiliki nilai akurasi yang lebih tinggi dibanding dengan metode lainnya.

Berdasarkan latar belakang yang telah dikemukakan, maka judul yang diambil dalam skripsi ini adalah “Penerapan *Algoritma Naïve Bayes* Untuk Memprediksi Tingkat Persetujuan Kartu Kredit”.

1.2 RUMUSAN MASALAH

Berdasarkan dari latar belakang, maka didapatkan rumusan masalah sebagai berikut :

1. Bagaimana penerapan *Algoritma Naïve Bayes* dalam memprediksi tingkat persetujuan kartu kredit ?
2. Seberapa besar tingkat akurasi *Algoritma Naïve Bayes* apabila digunakan untuk memprediksi tingkat persetujuan kartu kredit ?

1.3 BATASAN MASALAH

Berdasarkan dari latar belakang, maka adapun batasan masalah dalam penelitian ini, yaitu :

1. Data yang diperlukan adalah dataset Persetujuan Kartu Kredit.
2. Variabel yang digunakan dalam penelitian ini adalah *Gender, Age, Debt, Married, BankCustomer, Industry, Ethnicity, YearsEmployed, PriorDefault, Employed, CreditScore, DriversLicense, Citizen, ZipCode, Income, Approved*.

1.4 TUJUAN DAN MANFAAT

1.4.1 Tujuan Penelitian

Adapun tujuan dari penelitian yang akan dilakukan oleh penulis, yaitu :

1. Menerapkan metode *Algoritma Naïve Bayes* untuk memprediksi tingkat persetujuan Kartu Kredit

2. Mendapatkan akurasi yang tepat untuk memprediksi tingkat persetujuan Kartu Kredit dengan menggunakan *Algoritma Naïve Bayes*.
3. Mengetahui berapa besar tingkat akurasi *Algoritma Naïve Bayes* dalam memprediksi tingkat persetujuan Kartu Kredit.

1.4.2 Manfaat Penelitian

Adapun beberapa manfaat yang akan didapatkan dalam melakukan penelitian ini, yaitu :

1. Dapat memberikan kemudahan dalam memprediksi tingkat persetujuan Kartu Kredit dalam dunia perbankan.
2. Menambah wawasan bagi penulis dalam memprediksi dataset Persetujuan Kartu Kredit

BAB II

LANDASAN TEORI

2.1 KARTU KREDIT

Kartu kredit adalah sebuah alat pembayaran yang dikeluarkan oleh bank tertentu berbahan plastik dan berguna sebagai alat pembayaran secara kredit yang dilakukan oleh pemilik kartu/sesuai dengan nama yang tertera pada kartu kredit tersebut pada saat melakukan pembelian barang atau jasa [4].

2.2 DATA MINING

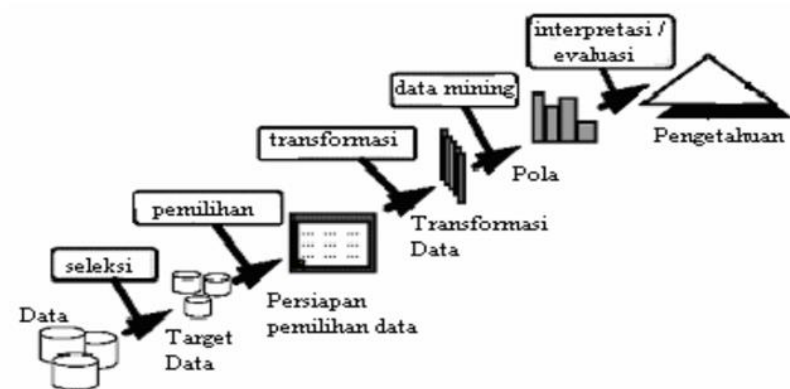
Data mining sering juga disebut *Knowledge Discovery in Database*, adalah analisa terhadap data (biasanya data yang berukuran besar) untuk menemukan hubungan yang jelas serta menyimpulkannya yang belum diketahui sebelumnya dengan cara terkini dipahami dan berguna bagi pemilik data tersebut [5].

Hal penting yang terkait dengan *data mining* adalah :

1. *Data mining* merupakan suatu proses otomatis terhadap data yang sudah ada.
2. Data yang akan di proses berupa data yang sangat besar.
3. Tujuan *data mining* adalah mendapatkan hubungan atau pola yang mungkin memberikan indikasi yang bermanfaat.

2.2.1 Tahapan Proses dalam Data Mining

Ada beberapa tahapan dalam proses *data mining*. Diagram dibawah ini menggambarkan beberapa tahap/proses yang berlangsung dalam *data mining*.



Gambar 2.1 Fase – Fase

Tahapan dalam proses *data mining* dapat dijelaskan sebagai berikut:

1. Seleksi data

Pemilihan (seleksi) data baru sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam *data mining* dimulai. Data hasil seleksi yang akan digunakan untuk proses *data mining*, disimpan dalam suatu berkas, terpisah dari basis data operasional.

2. Pre-processing/Cleaning (pemilihan data)

Proses *cleaning* mencakup antara lain membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data.

3. Transformasi

Coding adalah proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses *data mining*.

4. Data Mining

Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma dalam *data mining* sangat

bervariasi. Pemilihan metode dan algoritma sesuai dengan kebutuhan dan tujuan.

5. Interpretasi/Evaluasi

Pola informasi yang dihasilkan dari informasi *data mining* perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesa yang ada sebelumnya.

2.2.2 Pengelompokan Data Mining

Pengelompokan *data mining* dapat dikategorikan berdasarkan fungsi yang dilakukan atau berdasarkan jenis aplikasi yang menggunakannya [6] :

1. Klasifikasi

Dalam klasifikasi, terdapat target variabel kategori. Sebagai contoh, penggolongan pendapatan dapat dipisahkan dalam tiga kategori, yaitu pendapatan tinggi, pendapatan sedang, dan pendapatan rendah.

2. Pengklusteran (Clustering)

Clustering adalah pengelompokan set data, pengamatan, atau catatan untuk membentuk kelas objek yang memiliki kesamaan. Cluster adalah kumpulan record yang memiliki kesamaan dan perbedaan dengan record di cluster lain.

3. Asosiasi

Tugas asosiasi dalam *data mining* adalah untuk menemukan atribut-atribut yang terjadi bersama-sama. Salah satu implementasi asosiasi adalah analisis keranjang belanja.

4. Deskripsi

Peneliti dan analisis secara sederhana ingin mencoba mencari cara untuk menggambarkan pola dan kecenderungan yang terdapat dalam data. Deskripsi dari pola kecenderungan sering

memberikan kemungkinan penjelasan untuk suatu pola atau kecenderungan.

5. Estimasi

Estimasi hampir sama dengan klasifikasi, kecuali variabel target untuk estimasi adalah numerik daripada kategori. Sebuah model dibangun menggunakan baris data (dataset) lengkap yang memberikan nilai-nilai variabel target sebagai nilai prediksi. Selanjutnya, pada peninjauan berikutnya estimasi nilai dari variabel target dibuat berdasarkan nilai variabel prediksi.

6. Prediksi

Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada di masa mendatang.

2.3 KLASIFIKASI

Klasifikasi didefinisikan sebagai bentuk analisis data untuk mengekstrak model yang akan digunakan untuk memprediksi label kelas [7]. Kelas dalam klasifikasi merupakan atribut paling unik dari dataset dan variabel bebas dalam statistik.

2.4 ALGORITMA NAÏVE BAYES

Algoritma Naïve Bayes merupakan salah satu algoritma yang terdapat pada teknik klasifikasi. *Naïve Bayes* merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris *Thomas Bayes*, yaitu memprediksi peluang di masa depan berdasarkan pengalaman dimasa sebelumnya sehingga dikenal sebagai Teorema Bayes. Teorema tersebut dikombinasikan dengan *Naïve Bayes* dimana diasumsikan kondisi antar atribut saling bebas. Klasifikasi *Naïve Bayes* diasumsikan bahwa ada atau tidak ciri tertentu dari sebuah kelas tidak ada hubungannya dengan ciri dari kelas lainnya[8].

Persamaan dari teorema Bayes adalah :

$$P(H|X) = \frac{P(X|H).P(H)}{P(X)}$$

Keterangan :

X : Data dengan class yang belum diketahui

H : Hipotesis data X merupakan suatu class spesifik

$P(H|X)$: Probabilitas hipotesis H berdasarkan kondisi X
(posteriori probability)

$P(H)$: Probabilitas hipotesis H (prior probability)

$P(X|H)$: Probabilitas X berdasarkan kondisi pada hipotesis
H

$P(X)$: Probabilitas X

2.5 PENELITIAN SEJENIS

Tabel 2.1 Penelitian Sejenis

No.	Judul, Penulis dan Tahun	Metode	Atribut	Akurasi
1	"Penerapan Naïve Bayes Untuk Prediksi Kelayakan Kredit [9]".	<i>Algoritma Naïve Bayes</i>	Jaminan, Total Penghasilan, Pinjaman Lain, Total Pengeluaran, Nominal Kredit, Status Rumah, Tempo Pengembalian, Kelayakan Kredit	Akurasi data yang diperoleh dalam penelitian ini adalah sebesar 92,5% dengan error sebesar 7,5% dari 160 data yang digunakan untuk training dan 40 data untuk testing.
2	"Evaluasi Telemarketing Kartu Kredit Bank Menggunakan Algoritma Genetika	<i>Algoritma Genetika Untuk Seleksi Fitur dan Naïve Bayes</i>	Balance, Rumah, Pinjaman, Media, Bulan, Durasi, Persuasi, Previous, Poutcome, Target	Evaluasi telemarketing kartu kredit bank menggunakan Algoritma Genetika untuk Seleksi Fitur dan Naive Bayes

	Untuk Seleksi Fitur dan Naïve Bayes [8]”			memberikan hasil algoritma NB nilai akurasi 86,71% menjadi 90,27% dengan algoritma GAFS dan NB untuk prediksi nasabah bank yang mengambil kartu kredit.
3	“Klasifikasi Penentuan Pengajuan Kartu Kredit Menggunakan K-Nearest Neighbor [4]”	<i>Algoritma K-Nearest Neighbor</i>	Jenis Kelamin, Status Rumah, Status, Jumlah Tanggungan, Profesi, Penghasilan per Tahun, Kartu Kredit	K-Nearest Neighbor membuktikan bahwa aplikasi yang telah dibuat adalah benar dan sesuai dengan perhitungan manual. Sedangkan berdasarkan pengujian 3 kali pada data training dan data testing, didapatkan hasil rata-rata maksimal dari nilai precision sebesar 92%, nilai recall sebesar 83%, dan nilai accuracy sebesar 93%.
4	“Klasifikasi Peminjaman Nasabah Bank Menggunakan Metode Neural Network [10]”	<i>Neural Network</i>	Id, Age, Experience, Income, Zip Code, Family, CCAvg, Education, Mortgage, Securities Account, Cd Account, Online, Credit Card, Personal Loan	algoritma backpropagation dalam Neural Network menggunakan struktur 12 – 15 – 8 – 1, training cycles sejumlah 500, learning rate 0.01 dan momentum 0.1 menghasilkan nilai

				akurasi = 98.24%, dengan AUC sebesar =0,979 yang menunjukkan bahwa klasifikasi yang dihasilkan sangat baik
5	“Prediksi Kelayakan Pemberian Fasilitas Kartu Kredit Kepada Nasabah Dengan Metode Klasifikasi Data Mining(Studi Kasus : Bank XYZ) [11]”	<i>Discriminant Analysis dan Naïve Bayes</i>	Penghasilan/Tahun, Status Rumah, Jenis Kelamin,Pekerjaan, Status Kawin	Hasil perhitungan dengan model Discriminant Analysis dihasilkan nilai akurasi sebesar 61.80 (poor).sedangkan Pada model Naïve Bayes dihasilkan nilai akurasi sebesar 90,61
6	“Analisa Klasifikasi Kartu Kredit Menggunakan Algoritma Naïve Bayes [12]”	<i>Algoritma Naïve Bayes</i>	<i>Limit_Bal,Sex, Pendidikan,Marriage, Umur,Pay_0,Pay_2, Pay_3,Pay_4,Pay_5, Pay_6,Bill_Amt1, Bill_Amt2, Bill_Amt3, Bill_Amt4, Bill_Amt5, Bill_Amt6,Pay_Amt1, Pay_Amt2, Pay_Amt3 Pay_Amt4, Pay_Amt5, Pay_Amt6,Default Payment</i>	penelitian ini, ketika menggunakan 23 variabel didapatkan akurasi klasifikasi untuk 18001 data yaitu 69.11%. 11 variabel dan mendapatkan tingkat akurasi data sebanyak 79.7%. 6 variabel dan mendapat akurasi sebanyak 80.47%. 4 variabel akurasi data sebanyak 80.52%. Dan terakhir menggunakan 3 variabel memiliki akurasi data sebanyak 80.59%.

7	“Aplikasi Klasifikasi Penentuan Pengajuan Kartu Kredit Menggunakan Metode Naïve Bayes di Bank BNI Syariah Surabaya [13]”	<i>Algoritma Naïve Bayes</i>	Jenis Kelamin, Status Rumah, Status, Jumlah Tanggungan, Profesi, Penghasilan Tahunan, Kartu Kredit	Berdasarkan pengujian 3 kali pada setiap jumlah data training maupun testingnya, didapatkan nilai rata-rata maksimal dari nilai precision sebesar 100%, nilai recall sebesar 95%, dan nilai accuracy sebesar 98,67%
8	“Prediksi Tingkat Kelancaran Pembayaran Kredit Bank Dengan Menggunakan Algoritma Naïve Bayes dan K-Nearest Neighbor [14]”	<i>Naïve Bayes dan K-Nearest Neighbor</i>	3,4,5,6,7,8,9,10, 11,12,13,14, Dari 10 atribut tersebut Ada 3 Data yang Berbentuk Kategori yaitu Atribut 3,12,14	Pengujian yang dilakukan dengan mengkombinasikan algoritma Naïve Bayes dan K-Nearest Neighbor dengan pengujian yang dilakukan sebanyak 3 kali, hasil akurasi yang tertinggi yang dihasilkan 80%
9	“Penerapan Algoritma SVM Berbasis PSO Untuk Tingkat Pelayanan Marketing Terhadap Loyalitas Pelanggan Kartu Kredit [15]”	<i>Algoritma SVM</i>	<i>Komunikasi, Service, Komitmen, Loyalitas, Keamanan, Kemudahan, Jenis Kartu</i>	Dengan menggunakan support vector machine didapat nilai accuracy adalah 92.25 % dan nilai AUC adalah 0.976.
10	“Analisa Data Mining Default Kartu Kredit Menggunakan Algoritma Neural Network [16]”	<i>Algoritma Neural Network</i>	<i>Limit_Bal, Sex, Pendidikan, Marriage, Umur, Pay_0, Pay_2, Pay_3, Pay_4, Pay_5, Pay_6, Bill_Amt1, Bill_Amt2, Bill_Amt3, Bill_Amt4,</i>	visualisasi model Neural Network tersebut memiliki 3 hidden layer. Sehingga dari proses training diatas variable

			<i>Bill_Amt5, Bill_Amt6, Pay_Amt1, Pay_Amt2, Pay_Amt3, Pay_Amt4, Pay_Amt5, Pay_Amt6, Default.Payment</i>	yang paling berpengaruh ada 3 yaitu X14(jumlah pernyataan tagihan pada bulan Juli 2005), X23(jumlah yang dibayarkan pada bulan April 2005) dan X20(jumlah yang dibayarkan pada bulan Juli 2005). Hasil akurasi menggunakan confusion matrix nilai RMSE prediksi 2 sebesar 0.1938333333. Dari prediksi 2 memiliki nilai RMSE yang sangat kecil, sehingga keakurasian prediksi sangat bagus.
--	--	--	--	---

BAB III

METODOLOGI PENELITIAN

3.1 KERANGKA KERJA PENELITIAN

Untuk membantu dalam penyusunan penelitian ini, maka perlu adanya kerangka kerja (*framework*) yang jelas tahapan-tahapannya. Kerangka kerja ini merupakan langkah- langkah yang akan dilakukan dalam penyelesaian masalah yang akan dibahas. Adapun kerangka kerja penelitian yang akan digunakan adalah sebagai berikut :



Gambar 3.1 Tahapan Penelitian

Berdasarkan kerangka kerja penelitian yang telah di gambar di atas, maka dapat di uraikan pembahasan masing masing tahapan dalam penelitian adalah sebagai berikut :

1. Perumusan Masalah

Masalah yang dirumuskan dalam penelitian ini adalah bagaimana penerapan metode *Naïve Bayes* dalam memprediksi tingkat persetujuan Kartu Kredit.

2. Penentuan Tujuan

Tujuan yang akan dicapai dalam penelitian ini adalah ingin mengetahui berapa besar tingkat akurasi *Algoritma Naïve Bayes* dalam memprediksi tingkat persetujuan Kartu Kredit.

3. Mempelajari Literatur

Mempelajari literature-literatur yang dapat mencapai tujuan penelitian, literature-literatur bersumber dari buku-buku perpustakaan dan jaringan internet. Literatur-literatur yang digunakan nanti dilampirkan dalam daftar pustaka.

4. Pengumpulan Data dan Informasi

Dalam pengumpulan data, penulis mendapatkan dataset online yang terdapat pada sebuah website di internet.

5. Proses Data Mining

Data Mining adalah proses pengekstrasian *knowledge* yang tersimpan dalam dataset bervolume besar. Untuk mendapatkan *knowledge* dalam dataset digunakanlah *Algoritma Naïve Bayes*.

6. Pengujian

Pada tahap ini dilakukan pengujian dari hasil yang didapat dari tahap sebelumnya sebagai pedoman untuk mendapatkan hasil prediksi tingkat persetujuan Kartu Kredit.

7. Pembuatan Laporan

Pada tahap ini dilakukan pembuatan laporan, membuat hasil akhir dari suatu kegiatan penelitian berdasarkan data dan fakta yang telah diamati pada saat meneliti.

3.2 ANALISA TEKNIK ALGORITMA NAÏVE BAYES

Pada bagian ini data dan informasi yang diperoleh dan diproses dengan menggunakan *Algoritma Naïve Bayes* untuk mendapatkan hasil yang sesuai.

3.3 METODE PENELITIAN

Metode pengumpulan data yang digunakan dalam penelitian ini adalah dengan mencari dan memperoleh dataset online dari sebuah website bernama “*www.kaggle.com*” yang digunakan sebagai pengujian dalam penelitian ini. Pengumpulan data dilakukan untuk memperoleh hasil dan informasi yang diperlukan untuk mencapai tujuan penelitian.

3.4 ALAT PENELITIAN

Penulis menggunakan beberapa alat yang digunakan untuk melakukan pengolahan data penelitian, yaitu :

1. Hardware, dengan spesifikasi sebagai berikut :
 - a. Komputer, dengan processor Ryzen 7 2700X (3.7 Ghz)
 - b. RAM : 16.00 GB
 - c. SSD : 1.00 TB
2. Software, dengan keterangan sebagai berikut :
 - a. OS Windows 10 (64 bit)
 - b. Microsoft Excel & Word 2016
 - c. RapidMiner Studio Professional 7.1

3.5 JADWAL PENELITIAN

No	Jenis Kegiatan	Bulan																											
		Sept				Okt				Nov				Des				Jan				Feb							
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4				
1	Perumusan Masalah																												
2	Penentuan Tujuan																												
3	Mempelajari Literatur																												
4	Pengumpulan Data & Informasi																												
5	Proses Data Mining																												
6	Pengujian																												
7	Pembuatan Laporan																												

DAFTAR PUSTAKA

- [1] S. Prayoginingsih and R. P. Kusumawardani, "Inspirasi Profesional Sistem Informasi," *J. Sisfo*, vol. 06, no. 03, pp. 347–382 Sistem, 2017.
- [2] U. Budi *et al.*, "PENERAPAN ALGORITMA NEAREST NEIGHBOR DALAM MEMPREDIKSI Pengumpulan Data Studi Pustaka Analisa Masalah Perancangan Sistem Pengujian," no. 338.
- [3] D. Ph, "INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH APPLICATION OF MACHINE LEARNING ALGORITHMS IN CREDIT CARD Admel Dino Keco Zerina Masetic," no. 10, pp. 1–2, 2018.
- [4] Y. I. Kurniawan and T. I. Barokah, "Klasifikasi Penentuan Pengajuan Kartu Kredit Menggunakan K-Nearest Neighbor," *J. Ilm. Matrik*, vol. 22, no. 1, pp. 73–82, 2020, doi: 10.33557/jurnalmatrik.v22i1.843.
- [5] Y. Mardi, "Data Mining : Klasifikasi Menggunakan Algoritma C4.5," *Edik Inform.*, vol. 2, no. 2, pp. 213–219, 2017, doi: 10.22202/ei.2016.v2i2.1465.
- [6] G. Gunadi and D. I. Sensuse, "Penerapan Metode Data Mining Market Basket Analysis Terhadap Data Penjualan Produk Buku Dengan Menggunakan Algoritma Apriori Dan Frequent Pattern Growth (Fp-Growth) :," *Telematika*, vol. 4, no. 1, pp. 118–132, 2012.
- [7] D. Sartika and D. I. Sensuse, "Perbandingan Algoritma Klasifikasi Naive Bayes, Nearest Neighbour, dan Decision Tree pada Studi Kasus Pengambilan Keputusan Pemilihan Pola Pakaian," *JATISI (Jurnal Tek. Inform. dan Sist. Informasi)*, vol. 3, no. 2, pp. 151–161, 2017, [Online]. Available: <https://jurnal.mdp.ac.id/index.php/jatisi/article/view/78>
- [8] C. Industry *et al.*, "Evaluasi Telemarketing Kartu Kredit Bank Menggunakan Algoritma Genetika untuk Seleksi Fitur dan Naive Bayes (Evaluation of Bank Credit Card Telemarketing Using Genetic Algorithms for Feature Selection and Naive Bayes)," *J. Apl. Pelayaran dan Kepelabuhanan*, vol. 10, no. 1, pp. 12–22, 2019.
- [9] D. A. Kurniawan and D. Kriestanto, "Penerapan Naï • Ve Bayes Untuk

- Prediksi Kelayakan Kredit,” *JIKO (Jurnal Inform. dan Komputer)*, vol. 1, no. 1, pp. 19–23, 2016, doi: 10.26798/jiko.2016.v1i1.10.
- [10] N. Hadiananto, H. B. Novitasari, and A. Rahmawati, “Klasifikasi Peminjaman Nasabah Bank Menggunakan Metode Neural Network,” *J. Pilar Nusa Mandiri*, vol. 15, no. 2, pp. 163–170, 2019, doi: 10.33480/pilar.v15i2.658.
- [11] F. Stevanus Alvian, S. Carbini, S. Likmi Bandung, and S. Mardira Indonesia, “PREDIKSI KELAYAKAN PEMBERIAN FASILITAS KARTU KREDIT KEPADA NASABAH DENGAN METODE KLASIFIKASI DATA MINING (Studi Kasus : Bank XYZ),” *J. Comput. Bisnis*, vol. 14, no. 2, pp. 123–128, 2020, [Online]. Available: www.forbes.com
- [12] R. Wijayanti and S. Sulastri, “Analisa Klasifikasi Kartu Kredit Menggunakan Algoritma Naive Bayes,” *Sintak 2018*, vol. 10, no. September 2005, pp. 443–447, 2018.
- [13] T. I. Barokah and Y. I. Kurniawan, “Aplikasi Klasifikasi Penentuan Pengajuan Kartu Kredit Menggunakan Metode K-Nearest Neighbor Di Bank Bni Syariah Surabaya,” *Komputa J. Ilm. Komput. dan Inform.*, vol. 8, no. 1, pp. 36–43, 2019, doi: 10.34010/komputa.v8i1.3047.
- [14] F. Gultom and T. Simanjuntak, “Prediksi Tingkat Kelancaran Pembayaran Kredit Bank Dengan Menggunakan Algoritma Naïve Bayes Dan K-Nearest Neighbor,” *METHOMIKA J. Manaj. Inform. dan Komputerisasi Akunt.*, vol. 4, no. 2, pp. 98–102, 2021, doi: 10.46880/jmika.vol4no2.pp98-102.
- [15] W. Erawati, “Vol. XII No. 2, September 2015 Jurnal Techno Nusa Mandiri,” *Techno Nusa Mandiri*, vol. XII, no. 2, pp. 21–26, 2015.
- [16] S. Aisyah and Sulastri, “Analisa Data Mining Default Kartu Kredit Menggunakan Algoritma Neural Network,” *Din. Inform. Vol.10*, vol. 10, no. 1, pp. 47–52, 2018.

LAMPIRAN DATASET

Gender	Age	Debt	Married	BankCustomer	Industry	Ethnicity	YearsEmployed	PriorDefault	Employed	CreditScore	DriversLicense	Citizen	ZipCode	Income	Approved
Male	36.83	0.0	Yes	has a bank account	Industrials	White	1.25	Yes	Yes	1	No	ByBirth	202	0	Yes
Female	58.67	4.46	Yes	has a bank account	Materials	Black	3.04	Yes	Yes	6	No	ByBirth	43	580	Yes
Female	24.5	0.5	Yes	has a bank account	Materials	Black	1.5	Yes	No	0	No	ByBirth	280	824	Yes
Male	27.83	1.54	Yes	has a bank account	Industrials	White	3.75	Yes	Yes	5	Yes	ByBirth	100	3	Yes
Male	26.17	5.625	Yes	has a bank account	Industrials	White	1.71	Yes	No	0	No	ByCitizenMeans	120	0	Yes
Male	32.08	4.0	Yes	has a bank account	CommunicationServices	White	2.5	Yes	No	0	Yes	ByBirth	360	0	Yes
Male	33.17	1.04	Yes	has a bank account	Transport	Black	6.5	Yes	No	0	Yes	ByBirth	164	31285	Yes
Female	22.92	11.585	Yes	has a bank account	InformationTechnology	White	0.04	Yes	No	0	No	ByBirth	80	1349	Yes
Male	54.42	0.5	No	does not have a bank account	Financials	Black	3.96	Yes	No	0	No	ByBirth	180	114	Yes
Male	42.5	4.915	No	does not have a bank account	Industrials	White	3.165	Yes	No	0	Yes	ByBirth	52	1442	Yes
Male	22.08	0.83	Yes	has a bank account	Energy	Black	2.165	No	No	0	Yes	ByBirth	128	0	Yes
Male	29.92	1.835	Yes	has a bank account	Energy	Black	4.335	Yes	No	0	No	ByBirth	260	200	Yes
Female	38.25	6.0	Yes	has a bank account	Financials	White	1.0	Yes	No	0	Yes	ByBirth	0	0	Yes
Male	48.08	6.04	Yes	has a bank account	Financials	White	0.04	No	No	0	No	ByBirth	0	2690	Yes
Female	45.83	10.5	Yes	has a bank account	Materials	White	5.0	Yes	Yes	7	Yes	ByBirth	0	0	Yes
Male	36.67	4.415	No	does not have a bank account	Financials	White	0.25	Yes	Yes	10	Yes	ByBirth	320	0	Yes
Male	28.25	0.875	Yes	has a bank account	CommunicationServices	White	0.96	Yes	Yes	3	Yes	ByBirth	396	0	Yes
Female	23.25	5.875	Yes	has a bank account	Materials	White	3.17	Yes	Yes	10	No	ByBirth	120	245	Yes
Male	21.83	0.25	Yes	has a bank account	Real Estate	Black	0.665	Yes	No	0	Yes	ByBirth	0	0	Yes
Female	19.17	8.585	Yes	has a bank account	InformationTechnology	Black	0.75	Yes	Yes	7	No	ByBirth	96	0	Yes
Male	25.0	11.25	Yes	has a bank account	Energy	White	2.5	Yes	Yes	17	No	ByBirth	200	1208	Yes
Male	23.25	1.0	Yes	has a bank account	Energy	White	0.835	Yes	No	0	No	ByCitizenMeans	300	0	Yes
Female	47.75	8.0	Yes	has a bank account	Energy	White	7.875	Yes	Yes	6	Yes	ByBirth	0	1260	Yes
Female	27.42	14.5	Yes	has a bank account	Utilities	Black	3.085	Yes	Yes	1	No	ByBirth	120	11	Yes
Female	41.17	6.5	Yes	has a bank account	Materials	White	0.5	Yes	Yes	3	Yes	ByBirth	145	0	Yes
Female	15.83	0.585	Yes	has a bank account	Energy	Black	1.5	Yes	Yes	2	No	ByBirth	100	0	Yes
Female	47.0	13.0	Yes	has a bank account	ConsumerDiscretionary	Asian	5.165	Yes	Yes	9	Yes	ByBirth	0	0	Yes
Male	56.58	18.5	Yes	has a bank account	Real Estate	Asian	15.0	Yes	Yes	17	Yes	ByBirth	0	0	Yes
Male	57.42	8.5	Yes	has a bank account	Education	Black	7.0	Yes	Yes	3	No	ByBirth	0	0	Yes
Male	42.08	1.04	Yes	has a bank account	Industrials	White	5.0	Yes	Yes	6	Yes	ByBirth	500	10000	Yes
Male	29.25	14.79	Yes	has a bank account	ConsumerStaples	White	5.04	Yes	Yes	5	Yes	ByBirth	168	0	Yes
Male	42.0	9.79	Yes	has a bank account	Utilities	Black	7.96	Yes	Yes	8	No	ByBirth	0	0	Yes
Male	49.5	7.585	Yes	has a bank account	ConsumerDiscretionary	Asian	7.585	Yes	Yes	15	Yes	ByBirth	0	5000	Yes
Female	36.75	5.125	Yes	has a bank account	Education	White	5.0	Yes	No	0	Yes	ByBirth	0	4000	Yes
Female	22.58	10.75	Yes	has a bank account	Materials	White	0.415	Yes	Yes	5	Yes	ByBirth	0	560	Yes
Male	27.83	1.5	Yes	has a bank account	Industrials	White	2.0	Yes	Yes	11	Yes	ByBirth	434	35	Yes
Male	27.25	1.585	Yes	has a bank account	InformationTechnology	Black	1.835	Yes	Yes	12	Yes	ByBirth	583	713	Yes
Female	23.0	11.75	Yes	has a bank account	Utilities	Black	0.5	Yes	Yes	2	Yes	ByBirth	300	551	Yes
Male	27.75	0.585	No	does not have a bank account	InformationTechnology	White	0.25	Yes	Yes	2	No	ByBirth	260	500	Yes
Male	54.58	9.435	Yes	has a bank account	Healthcare	Latino	14.415	Yes	Yes	11	Yes	ByBirth	30	300	Yes
Male	34.17	9.17	Yes	has a bank account	Energy	White	4.5	Yes	Yes	12	Yes	ByBirth	0	221	Yes
Male	28.92	15.0	Yes	has a bank account	Energy	Black	5.335	Yes	Yes	11	No	ByBirth	0	2283	Yes
Male	29.67	1.415	Yes	has a bank account	Industrials	Black	0.75	Yes	Yes	1	No	ByBirth	240	100	Yes
Male	39.58	13.915	Yes	has a bank account	Industrials	White	8.625	Yes	Yes	6	Yes	ByBirth	70	0	Yes
Male	56.42	28.0	No	does not have a bank account	Energy	White	28.5	Yes	Yes	40	No	ByBirth	0	15	Yes
Male	54.33	6.75	Yes	has a bank account	Energy	Black	2.625	Yes	Yes	11	Yes	ByBirth	0	284	Yes
Female	41.0	2.04	No	does not have a bank account	Materials	Black	0.125	Yes	Yes	23	Yes	ByBirth	455	1236	Yes
Male	31.92	4.46	Yes	has a bank account	InformationTechnology	Black	6.04	Yes	Yes	3	No	ByBirth	311	300	Yes
.....															
.....															
.....															
Female	25.75	0.5	Yes	has a bank account	Energy	Black	0.875	Yes	No	0	Yes	ByBirth	491	0	Yes
Male	26.0	1.0	Yes	has a bank account	Materials	White	1.75	Yes	No	0	Yes	ByBirth	280	0	Yes
Male	37.42	2.04	Yes	has a bank account	Industrials	White	0.04	Yes	No	0	Yes	ByBirth	400	5800	Yes
Male	34.92	2.5	Yes	has a bank account	Industrials	White	0.0	Yes	No	0	Yes	ByBirth	239	200	Yes
Male	34.25	3.0	Yes	has a bank account	InformationTechnology	Black	7.415	Yes	No	0	Yes	ByBirth	0	0	Yes

